Available on the CERN CDS information server **CMS PAS SUS-11-006**

CMS Physics Analysis Summary

Contact: cms-pag-conveners-susy@cern.ch 2011/10/15

Search for New Physics in Events with b-quark Jets and Missing Transverse Energy in Proton-Proton Collisions at 7 TeV

The CMS Collaboration

Abstract

We present a search for physics beyond the standard model (SM) based on events with large missing transverse energy, at least three jets, and b-quark jets. The study is performed using 1.1 fb^{-1} of proton-proton collisions collected at $\sqrt{s} = 7$ TeV with the CMS detector at the LHC. All SM backgrounds are evaluated using data-based techniques. The number of events observed is consistent with the SM expectation. We set 95% confidence level upper limits on the cross sections of simplified models in which new particles decay to two b-quark jets plus an undetected particle. We also present exclusion limits in the context of the constrained minimal supersymmetric extension of the SM.

1 Introduction

Final states with large missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) and multiple, high transverse momentum (p_{T}) jets characterize many extensions of the standard model (SM). Within this class of events, those with identified b-quark jets (b-jets) present a distinct topological signature, a different background composition, and different sensitivity to new physics (NP) model parameters. This signature can therefore be used to search for NP in a manner that complements other studies. Examples of NP models with $E_{\mathrm{T}}^{\mathrm{miss}}$ and a b-jet-rich environment include little Higgs models [1] with a new heavy top quark, and models of Supersymmetry (SUSY) [2–4] in which the scalar partners of bottom and top quarks are relatively light.

This note describes a search for NP in events with large $E_{\mathrm{T}}^{\mathrm{miss}}$, no identified leptons, three or more high p_{T} jets, and at least one identified b-jet. The analysis is based on a sample of proton-proton collisions collected at $\sqrt{s} = 7$ TeV with the CMS detector at the Large Hadron Collider (LHC) at CERN. The integrated luminosity is 1.1 fb^{-1}. Recent studies with similar goals are presented in Refs. [5, 6]. Our study is characterized by a strong reliance on data-based techniques to evaluate the SM backgrounds, reducing systematic uncertainties. In particular, we introduce a variable $\Delta\phi_{\mathrm{N}}^{\mathrm{min}}$, described below, that simplifies the data-driven evaluation of the background from QCD interactions.

The principal sources of SM background are (1) events with top (t) quarks, specifically $\mathrm{t\bar{t}}$ pair and single-top events, (2) electroweak (EW) events with a Z or W boson plus jets, and (3) multijet QCD events. Diboson (WW, ZZ, and WZ) events form a smaller class of EW background. For EW events and events with a top quark, significant $E_{\mathrm{T}}^{\mathrm{miss}}$ can arise when a W (Z) decays into a lepton and a neutrino (two neutrinos). The neutrino provides a true source of $E_{\mathrm{T}}^{\mathrm{miss}}$. For QCD events, significant $E_{\mathrm{T}}^{\mathrm{miss}}$ arises primarily from jet p_{T} mismeasurements.

In Sects. 2-3 we present the data sample and event selection. The b-tagging algorithm is described in Sect. 4. Sect. 5 introduces the $\Delta\phi_{\mathrm{N}}^{\mathrm{min}}$ variable. Our data-based techniques to evaluate background are described in Sect. 6. In Sect. 7 we present a likelihood analysis that simultaneously determines the SM backgrounds and tests the consistency of NP models with the data. The likelihood framework allows us to account for NP contamination of the signal and control regions in a comprehensive and consistent manner. Our results are presented in Sect. 8. A summary is given in Sect. 9.

A roadmap of our analysis is shown in Fig. 1. SIG refers to signal regions, SB to an $E_{\mathrm{T}}^{\mathrm{miss}}$ sideband control region and LSB and LDP to additional control regions. The QCD background is evaluated using the $\Delta\phi_{\mathrm{N}}^{\mathrm{min}}$ variable mentioned above. Top and W+jets backgrounds are estimated using a single-lepton control sample (in this note, "top" refers to both $\mathrm{t\bar{t}}$ and single-top events). The Z+jets background with $\mathrm{Z} \to \nu\bar{\nu}$ is evaluated using $\mathrm{Z} \to \ell^+\ell^-$ events with $\ell = \mu$ or e. Beside the standard method to evaluate the top and W+jets background, indicated in Fig. 1, we have a cross-check method based on different procedures.

2 Detector and Trigger

A detailed description of the CMS detector is given elsewhere [7]. The CMS coordinate system is defined with the origin at the center of the detector and the z-axis along the direction of the counter-clockwise beam. The transverse plane is perpendicular to the beam axis, with ϕ the azimuthal angle, θ the polar angle, and $\eta = -\ln[\tan(\theta/2)]$ the pseudorapidity. A superconducting solenoid provides an axial magnetic field of 3.8 T. Within the field volume are a silicon pixel and strip tracker, a crystal electromagnetic calorimeter, and a brass-scintillator hadron

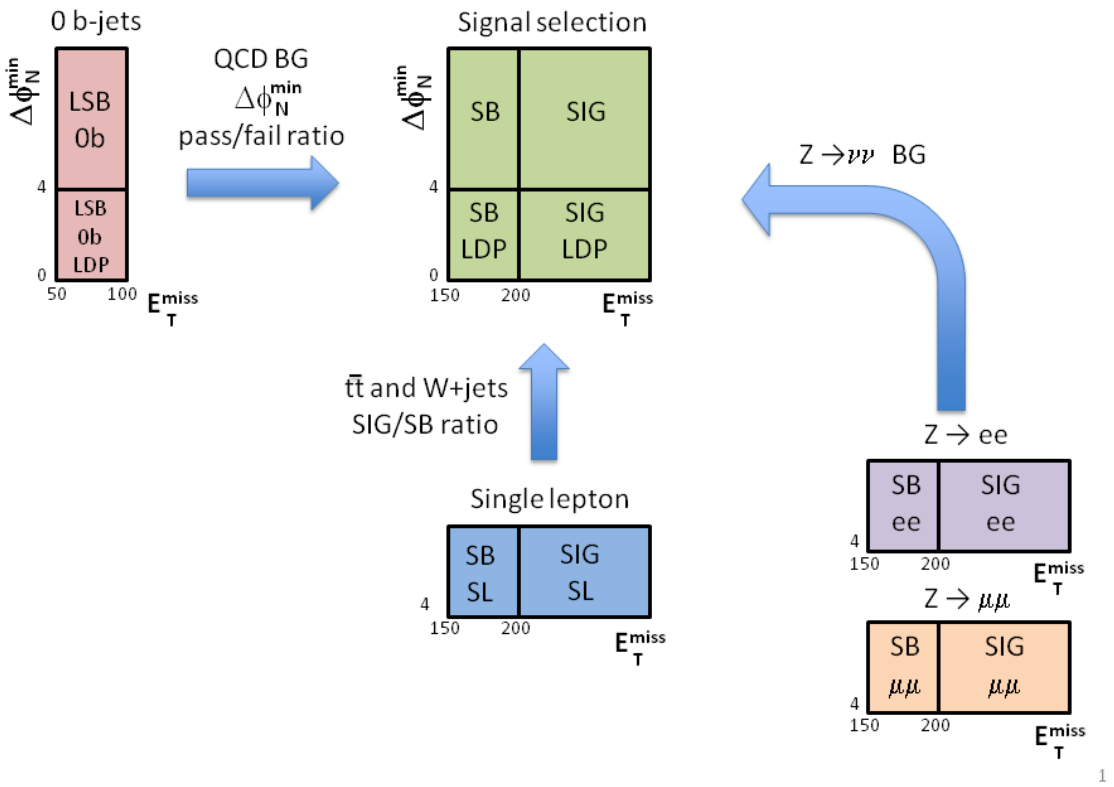


Figure 1: Schematic diagram indicating the various event samples used for background evaluation. SIG refers to signal regions, SB to the sideband region $150 < E_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV, LSB to the low-sideband region $50 < E_{\mathrm{T}}^{\mathrm{miss}} < 100$ GeV, and LDP to the low $\Delta\phi_{\mathrm{N}}^{\mathrm{min}}$ region $\Delta\phi_{\mathrm{N}}^{\mathrm{min}} < 4.0$. The diagram illustrates the loose selection, which requires $H_{\mathrm{T}} > 350$ GeV for all SB and SIG regions and $E_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV for the SIG regions. The tight selection is the same except with $H_{\mathrm{T}} > 500$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ GeV, respectively. Beside the standard method to evaluate the top and W+jets background, indicated in the diagram, we have a cross-check method based on different procedures.

calorimeter. Muons are detected with gas-ionization chambers embedded in the steel flux-return yoke outside the solenoid. The tracker covers the region $|\eta| < 2.5$ and the calorimeters $|\eta| < 3.0$. The region $3 < |\eta| < 5$ is instrumented with a forward calorimeter. The near-hermeticity of the detector permits accurate measurement of energy balance in the transverse plane.

The principal trigger used in the analysis is based on cross-object requirements for H_{T} and MH_T, where H_{T} is the scalar sum of the transverse energy of jets and MH_T is the modulus of the corresponding vector sum. The trigger is found to be 100% efficient for the offline requirements $H_{\mathrm{T}} > 400$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV. Our loosest analysis requirement is $H_{\mathrm{T}} > 350$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV, for which the trigger is about 99% efficient. A correction is applied to account for this small inefficiency. As part of the evaluation of the QCD and Z+jets background, we also employ a pre-scaled pure H_{T} trigger and specialized lepton triggers, respectively. A cross-object muon and H_{T} trigger is used as part of our cross-check analysis of the top and W+jets background.

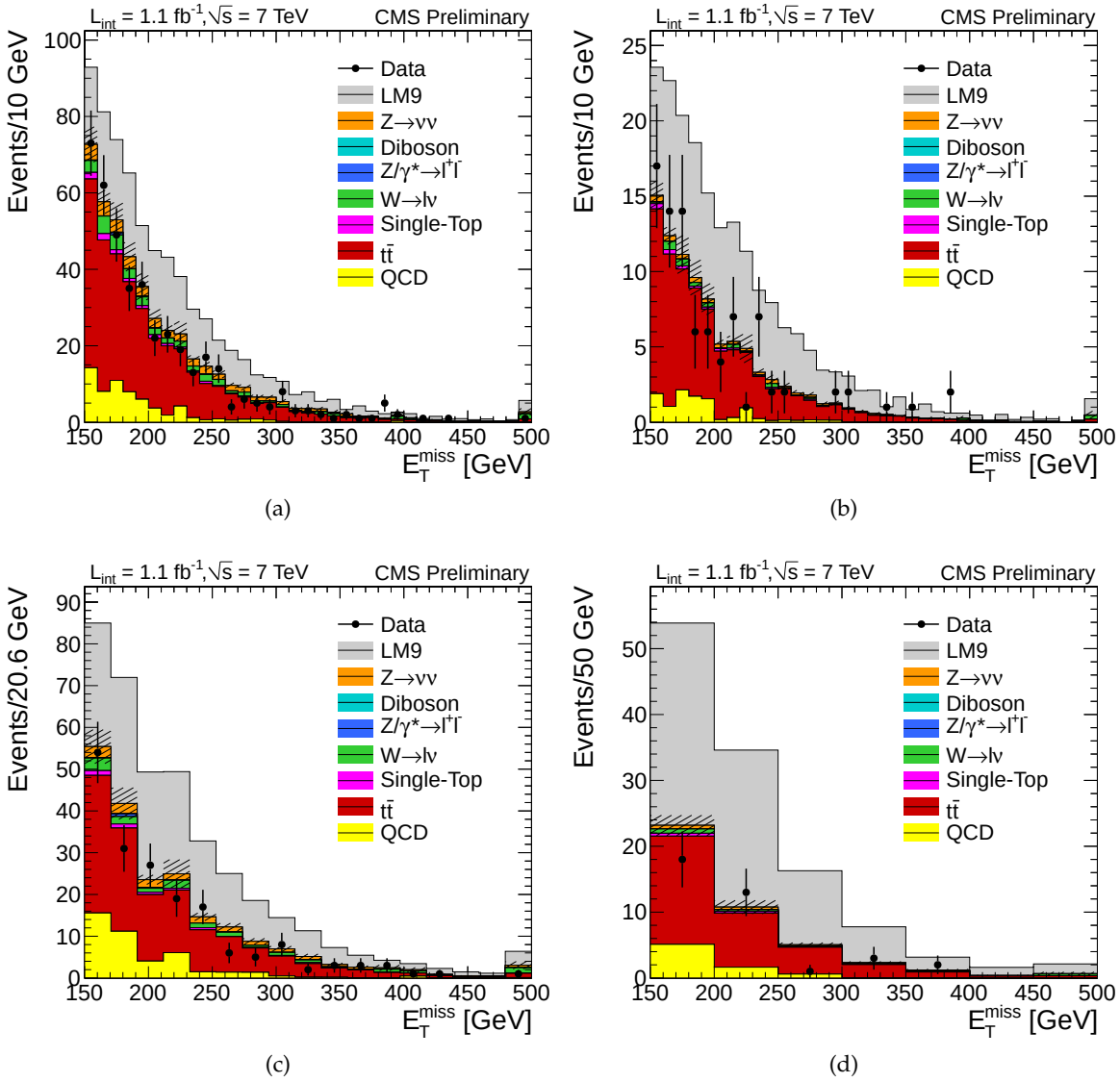


Figure 2: Distribution of $E_{\mathrm{T}}^{\mathrm{miss}}$ for the (a,b) loose and (c,d) tight signal selections in comparison to MC for (a,c) $\geq$ 1 b-jets and (b,d) $\geq$ 2 b-jets. The hatched bands show the statistical uncertainty on the total standard model MC prediction.

3 Event Selection

Physics objects are defined using the particle flow (PF) method [8], which reconstructs and identifies charged and neutral hadrons, muons, electrons (with associated bremsstrahlung photons), and photons, using an optimized combination of information from CMS subdetectors. The PF objects serve as input for jet reconstruction, based on the anti-k_T algorithm [9] with resolution scale 0.5. p_{T}- and η-dependent jet corrections [10] account for residual effects of non-uniform detector response. The missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined as the magnitude of the vector sum of the transverse momenta of all PF objects. The $E_{\mathrm{T}}^{\mathrm{miss}}$ vector is the negative of that same vector sum. Henceforth, the jet, lepton, and $E_{\mathrm{T}}^{\mathrm{miss}}$ results in this note refer to the corresponding PF quantities.

The basic event selection criteria are:

Table 1: Number of data events and corresponding MC predictions for the loose ($H_T > 350$ GeV, $E_T^{miss} > 200$ GeV) and tight ($H_T > 500$ GeV, $E_T^{miss} > 300$ GeV) signal selections. MC results for the CMSSM test point LM9 are also shown. The MC uncertainties are statistical. The normalization is to 1.1 fb^{-1}.

	$(H_T, E_T^{miss}) > (350, 200)$ GeV		$(H_T, E_T^{miss}) > (500, 300)$ GeV	
	≥ 1 b-jets	≥ 2 b-jets	≥ 1 b-jets	≥ 2 b-jets
Data	155	30	20	5
Total SM	183 ± 5	35.7 ± 1.3	25.1 ± 1.6	4.54 ± 0.37
$t\bar{t}$	122 ± 2	28.9 ± 0.7	14.7 ± 0.8	3.49 ± 0.24
Single top	4.54 ± 0.38	0.77 ± 0.09	0.59 ± 0.15	0.12 ± 0.04
W+Jets	17.0 ± 2.1	1.21 ± 0.45	4.20 ± 1.28	0.42 ± 0.28
$Z \to \nu\bar{\nu}$	22.5 ± 0.5	2.23 ± 0.10	4.25 ± 0.20	0.43 ± 0.04
$Z/\gamma^* \to \ell^+\ell^-$	0.17 ± 0.17	0.01 ± 0.01	0	0
Diboson	0.69 ± 0.07	0.10 ± 0.02	0.10 ± 0.02	0.006 ± 0.002
QCD	16.4 ± 3.9	2.5 ± 0.9	1.28 ± 0.40	0.08 ± 0.01
SUSY LM9	147 ± 5	60.0 ± 2.5	27.7 ± 2.2	10.1 ± 1.0

- at least one well-defined primary event vertex;
- at least three jets with $p_T > 50$ GeV and $|\eta| < 2.4$;
- at least one tagged b-quark jet ("b-jet"); the b-tagging algorithm is described in Sect. 4; b-jets used in this analysis are required to have $p_T > 30$ GeV;
- no identified, isolated electron or muon candidate with $p_T > 10$ GeV; electron candidates are restricted to $|\eta| < 2.5$ and muon candidates to $|\eta| < 2.4$;
- $\Delta\phi_N^{min} > 4.0$, where the $\Delta\phi_N^{min}$ variable is defined in Sect. 5.

We select two signal (SIG) event samples, corresponding to a "loose selection" and a "tight selection." Besides the basic criteria, the loose (tight) selection requires:

- $H_T > 350$ GeV (500 GeV), where H_T is calculated using jets with $p_T > 50$ GeV and $|\eta| < 2.4$;
- $E_T^{miss} > 200$ GeV (300 GeV).

The E_T^{miss} distribution of events in the loose selection (except for the E_T^{miss} requirement) is shown in the top half of Fig. 2, for events with (a) ≥ 1 or (b) ≥ 2 b-jets. The corresponding results for the tight selection are shown in the bottom half of Fig. 2. The results are presented in comparison to Monte Carlo (MC) simulations of SM processes, which are processed through GEANT [11] to account for the CMS detector response. The simulated $t\bar{t}$, single-top, W+jets, Z+jets, and WW events are created at the parton level with the MADGRAPH [12] event generator. The cross sections are valid to the next-to-next-to-leading order for the W and Z events, and to next-to-leading order (NLO) for the single-top and WW events. The $t\bar{t}$ events are normalized to the measured cross section [13]. WZ, ZZ, and QCD events are generated with the PYTHIA program [14] and normalized to the leading-order cross sections. For all MC samples, PYTHIA is used to describe subsequent parton showering and hadronization. The jet energy resolution in the MC is corrected to account for a small discrepancy with respect to the data [10]. The pileup distributions in the MC are reweighted to match the measured distribution.

The numbers of events in the loose and tight signal selections are listed in Table 1 for data and MC. The dominant component of the SM background is seen to be $t\bar{t}$. The MC results are informational only. We use data-based techniques to evaluate SM backgrounds as described

Table 2: CMSSM parameter values for LM9.

Process	m_0 (GeV)	$m_{1/2}$ (GeV)	A_0 (GeV)	$\tan\beta$	μ
SUSY LM9	1450	175	0	50	> 0

below.

As an example of results from a NP model, Fig. 2 and Table 1 include predictions from the "low mass" test point LM9 [15], which belongs to a class of SUSY models known as the constrained minimal supersymmetric extension of the standard model (CMSSM) [2–4]. The cross section for CMSSM events at the LHC is dominated by squark-squark, squark-gluino, and gluino-gluino pair production, where squarks (gluinos) are the supersymmetric partners of quarks (gluons). A high p_{T} squark or gluino decays into quarks, gluons, and other SM particles, as well as the lightest supersymmetric particle (LSP), which is presumed to escape detection and lead to significant $E_{\mathrm{T}}^{\mathrm{miss}}$. Amongst CMSSM models, LM9 occupies a favorable position in parameter space for an enhanced production of b-jets. The LM9 parameter values are listed in Table 2. The LM9 MC sample is generated with PYTHIA. The NLO cross section from PROSPINO [16] is used to normalize the results.

4 b-jet Identification

To identify b-jets, we use an algorithm [17] that attempts to reconstruct a secondary vertex with at least three charged tracks within a jet. The discriminating variable is the three-dimensional decay length significance. The nominal misidentification rate for non b-jets with a p_{T} value of 80 GeV is 0.1%. The corresponding b-tagging efficiency is 38%. The MC describes the measured p_{T} dependence of the efficiency well, but a p_{T}-independent efficiency scale factor of $0.90 \pm 0.03\,(\mathrm{stat.}) \pm 0.05\,(\mathrm{syst.})$ [17] is required. The momentum range of the nominal efficiency study, $20 < p_{\mathrm{T}} < 240$ GeV, covers most of the spectrum relevant to our search. With the event selection used here, the fraction of $\mathrm{t\bar{t}}$ (LM9) events with a b-jet satisfying $p_{\mathrm{T}} > 240$ GeV is only about 10% (8%). In addition, a dedicated study of b-tagging efficiency in the range 240-350 GeV is performed by measuring the ratio of double- to single-b tagged events in a single-lepton control sample. The scale factors for the lower p_{T} range are found to be valid, within larger uncertainties determined by the statistical uncertainties of the control sample, for this larger p_{T} region as well. Thus, for p_{T} values above 240 GeV, we use the same scale factor to correct efficiencies but assign an uncertainty that we evaluate to be 32%. Due to the very small number of b-jet candidates at high p_{T}, we conservatively assume the b-tagging efficiency to be zero for $p_{\mathrm{T}} > 350$ GeV when calculating limits (Sect. 8).

5 The $\Delta\phi_{\mathrm{N}}^{\mathrm{min}}$ Variable

The $\Delta\phi_{\mathrm{N}}^{\mathrm{min}}$ variable is defined as follows. We calculate the azimuthal opening angle $\Delta\phi_i$ ($i = 1, 2, 3$) between $E_{\mathrm{T}}^{\mathrm{miss}}$ and each of the three highest p_{T} jets in an event. $\Delta\phi^{\mathrm{min}} \equiv \min(\Delta\phi_i)$ is a common variable used to reject QCD background. $\Delta\phi^{\mathrm{min}}$ is known to be strongly correlated with $E_{\mathrm{T}}^{\mathrm{miss}}$, which hinders its use in a simple data-based evaluation of the QCD background. To reduce this correlation, we divide the $\Delta\phi_i$ by their estimated resolutions $\sigma_{\Delta\phi,i}$, yielding our new variable $\Delta\phi_{\mathrm{N}}^{\mathrm{min}} \equiv \min(\Delta\phi_i/\sigma_{\Delta\phi,i})$.

The resolution $\sigma_{\Delta\phi,i}$ of jet i is evaluated by considering the p_{T} resolution $\sigma_{p_{\mathrm{T}}}$ of the other jets in the event. We consider that jet mismeasurements affect the magnitude of a jet's p_{T} but not

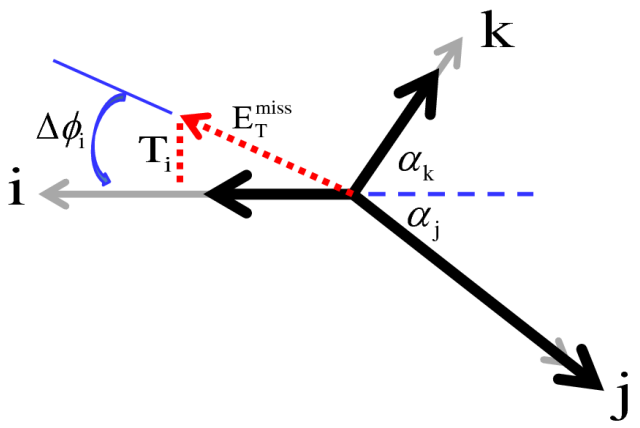
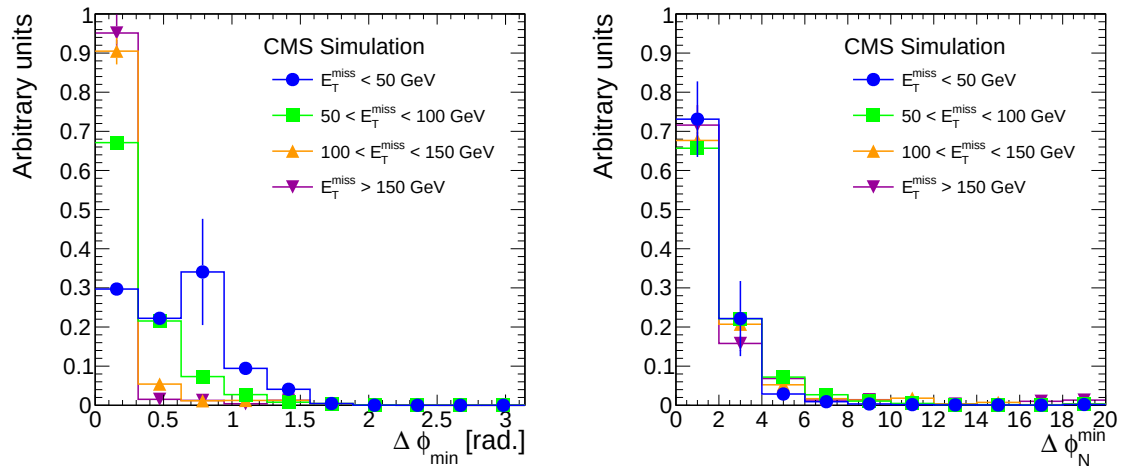


Figure 3: Illustration of variables used to calculate $\Delta\phi_N^{min}$ in the case of an event with exactly three jets. The light-shaded (light gray) solid arrows show the true p_T values of the three jets i, j, and k. The dark-shaded (black) solid arrows show the reconstructed jet p_T values. α_j and α_k are the angles of jets j and k with respect to jet i. The E_T^{miss} for the event is shown by the dotted (red) arrow. The component of E_T^{miss} perpendicular to jet i, denoted T_i, is shown by the dotted (red) line. $\Delta\phi_i$ is the angle between E_T^{miss} and jet i.

its direction, and that most E_T^{miss} in a QCD event arises from the mismeasurement of a single jet. The situation is depicted in Fig. 3 for the case of an event with exactly three jets i, j, and k. The p_T mismeasurements of jets j and k cause E_T^{miss} to point away from jet i's axis and thus $\Delta\phi_i$ to deviate from zero. Let T_i be the component of E_T^{miss} perpendicular to jet i. Then $T_i^2 \approx (\sum_n \sigma_{p_T,n} \sin\alpha_n)^2$, where the sum is over all other jets in the event with $p_T > 30$ GeV. Our estimate of the $\Delta\phi$ resolution is $\sigma_{\Delta\phi,i} = \arctan(T_i/E_T^{miss})$. For the p_T resolution, we use the approximate result $\sigma_{p_T} = 0.10\, p_T$ [10].

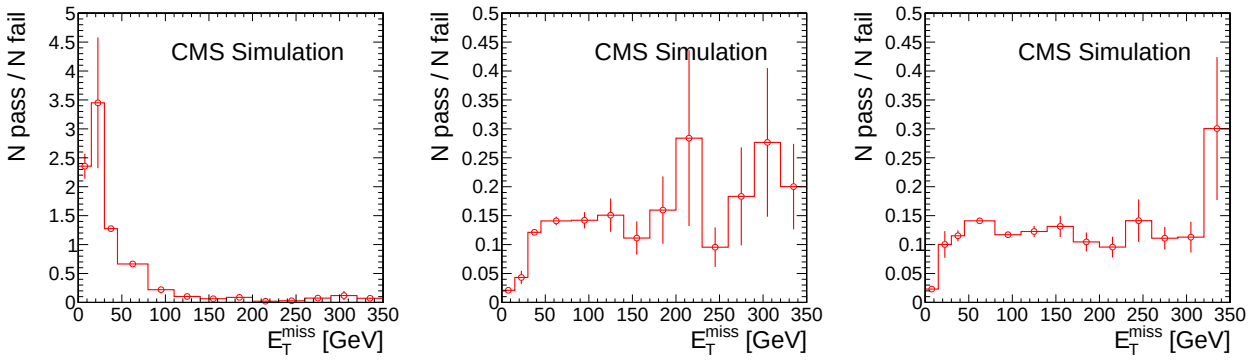
Figure 4 (a) shows the distribution of $\Delta\phi^{min}$ in intervals of E_T^{miss} for a QCD MC sample with ≥ 1 b-jets selected with our loose criteria except for the $\Delta\phi_N^{min}$ requirement. The strong correlation between $\Delta\phi^{min}$ and E_T^{miss} is evident. Figure 5 (a) shows the ratio of the number of QCD MC events with $\Delta\phi^{min} < 0.3$ to the number with $\Delta\phi^{min} > 0.3$, based on the same event selection. (The requirement $\Delta\phi^{min} > 0.3$ or similar is commonly used to reject QCD background, see, e.g., Ref. [6].) The corresponding results for $\Delta\phi_N^{min}$ are shown in Figs. 4 (b) and 5 (b). For the latter figure we choose $\Delta\phi_N^{min} = 4.0$ in place of $\Delta\phi^{min} = 0.3$, which yields a similar selection efficiency for $E_T^{miss} > 100$ GeV. For $E_T^{miss} > 30$ GeV, the distributions based on $\Delta\phi_N^{min}$ are seen to be far less dependent on E_T^{miss} than those based on $\Delta\phi^{min}$.

Figure 5 (c) shows the ratio $N(\Delta\phi_N^{min} \geq 4)/N(\Delta\phi_N^{min} < 4)$ for a QCD MC sample in which there are zero tagged b-jets. By requiring that there not be a b-jet, we reduce the contribution of top events. It is seen that $N(\Delta\phi_N^{min} \geq 4)/N(\Delta\phi_N^{min} < 4)$ is not significantly different between Figs. 5 (b) and (c), i.e., this ratio has an approximately constant value of about 0.13 (for E_T^{miss} values larger than about 30 GeV) irrespective of the number of b-jets. The measured result for $N(\Delta\phi_N^{min} \geq 4)/N(\Delta\phi_N^{min} < 4)$ with zero b-jets is shown in Fig. 6 in comparison to MC predictions. These data are collected with a pre-scaled H_T trigger, which allows us to select events at low E_T^{miss} without introducing a trigger bias. At low E_T^{miss}, the distribution is dominated by QCD.



(a) (b)

Figure 4: QCD MC results: The distribution of (a) $\Delta\phi^{\min}$ and (b) $\Delta\phi_{\mathrm{N}}^{\min}$ in intervals of $E_{\mathrm{T}}^{\mathrm{miss}}$, for events with ≥ 1 b-jets selected with the loose criteria except for the requirement on $\Delta\phi_{\mathrm{N}}^{\min}$.



(a) (b) (c)

Figure 5: QCD MC results: (a) Ratio of the number of events that pass the criterion $\Delta\phi^{\min} > 0.3$ to the number that fail, for events with ≥ 1 b-jets selected with the loose criteria except for the requirement on $\Delta\phi_{\mathrm{N}}^{\min}$. (b) Analogous ratio of events with $\Delta\phi_{\mathrm{N}}^{\min} > 4$ to those with $\Delta\phi_{\mathrm{N}}^{\min} < 4$. (c) Same as part (b) except for events with zero b-jets.

6 Background Evaluation

6.1 QCD Background

The low level of correlation between $\Delta\phi_{\mathrm{N}}^{\min}$ and $E_{\mathrm{T}}^{\mathrm{miss}}$ allows us to employ a simple data-based approach to evaluate the QCD background. As discussed in Sect. 5, the ratio $N(\Delta\phi_{\mathrm{N}}^{\min} \geq 4)/N(\Delta\phi_{\mathrm{N}}^{\min} < 4)$ is approximately independent of $E_{\mathrm{T}}^{\mathrm{miss}}$, and also of the number of b-jets, for QCD events. Furthermore, the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution below around 100 GeV is dominated by QCD. We can therefore measure $N(\Delta\phi_{\mathrm{N}}^{\min} \geq 4)/N(\Delta\phi_{\mathrm{N}}^{\min} < 4)$ in a low $E_{\mathrm{T}}^{\mathrm{miss}}$ region of the zero b-jet sample, selected with the pre-scaled H_{T} trigger, and assume this equals $N(\Delta\phi_{\mathrm{N}}^{\min} \geq 4)/N(\Delta\phi_{\mathrm{N}}^{\min} < 4)$ for QCD events at all $E_{\mathrm{T}}^{\mathrm{miss}}$ values, also for samples with b-jets such as our SIG samples. To make this measurement we use the low-$E_{\mathrm{T}}^{\mathrm{miss}}$ region defined by $50 < E_{\mathrm{T}}^{\mathrm{miss}} < 100$ GeV. We call this region the "low sideband" (LSB).

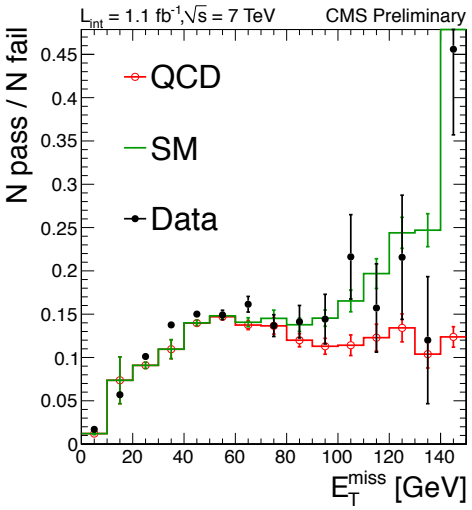
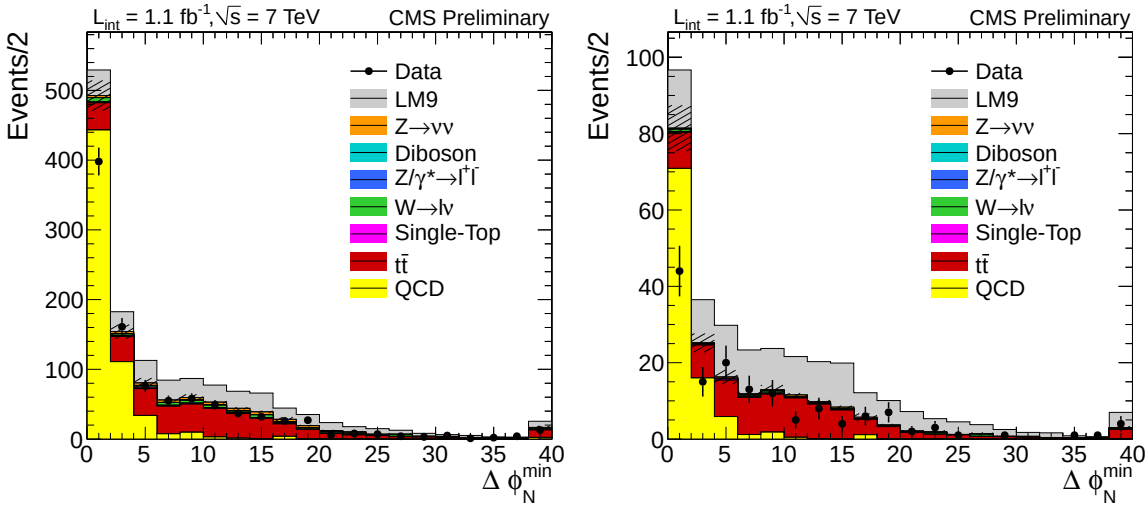


Figure 6: The ratio $N(\Delta\phi_N^{\min} \geq 4)/N(\Delta\phi_N^{\min} < 4)$ as a function of E_T^{miss} for the zero b-jet sample. The histograms show the QCD MC and total SM MC predictions. The solid points show the data, collected with a pre-scaled H_T trigger.



(a) (b)

Figure 7: Distribution of $\Delta\phi_N^{\min}$ in data and MC for the loose selection with $E_T^{\text{miss}} > 150$ GeV for (a) ≥ 1 b-jets and (b) ≥ 2 b-jets. The hatched bands show the statistical uncertainty on the total standard model MC prediction.

We also define "low $\Delta\phi$" (LDP) intervals $\Delta\phi_N^{\min} < 4.0$. We do this not only for the LSB, but also for the sideband (SB) and signal (SIG) regions (see Fig. 1). We denote these regions LSB-LDP, SB-LDP and SIG-LDP. The LSB-LDP is essentially 100% QCD. Similarly, the SB-LDP and SIG-LDP regions are mostly QCD, as shown in Fig. 7. At higher values of E_T^{miss}, top and EW contributions to the SB-LDP and SIG-LDP become more important. This residual contamination is subtracted using MC.

The non-LDP region of the LSB is 97% QCD according to MC. Denoting this latter region "B," the LSB-LDP "A," and the SIG-LDP "C," the QCD background in the SIG region (region "D") is $D = C\,B/A$ in the limit that all regions A, B, and C are pure QCD and that $\Delta\phi_N^{\min}$ is uncorrelated with E_T^{miss}. This is known as the ABCD method. An analogous procedure with the SB-LDP as

Table 3: MC closure test of the QCD background evaluation method. "Loose" and "Tight" refer to the corresponding H_{T} requirement, and also to the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement in the case of the SIG region. h refers to either the SIG or SB region depending on the row. The closure test is based on comparison of the last two columns.

	N_{LSB}	$N_{LSB-LDP}$	N_{LDP-h}	$N_{predicted}$	N_{true}
$\geq$ 1 b, Loose, SB	68077 ± 1266	478813 ± 2915	340 ± 25	48 ± 4	43 ± 10
$\geq$ 1 b, Loose, SIG	68077 ± 1266	478813 ± 2915	85 ± 9	12 ± 1	18 ± 6
$\geq$ 1 b, Tight, SB	17275 ± 322	207018 ± 971	241 ± 15	20 ± 1	31 ± 8
$\geq$ 1 b, Tight, SIG	17275 ± 322	207018 ± 971	6.1 ± 0.5	0.51 ± 0.04	1.1 ± 0.3
$\geq$ 2 b, Loose, SB	68077 ± 1266	478813 ± 2915	49 ± 11	7 ± 2	6 ± 5
$\geq$ 2 b, Loose, SIG	68077 ± 1266	478813 ± 2915	5.0 ± 0.8	0.7 ± 0.1	6 ± 5
$\geq$ 2 b, Tight, SB	17275 ± 322	207018 ± 971	34 ± 8	2.8 ± 0.7	6 ± 5
$\geq$ 2 b, Tight, SIG	17275 ± 322	207018 ± 971	0.29 ± 0.05	0.025 ± 0.004	0.07 ± 0.04

Table 4: Systematic uncertainties for the QCD background estimate, in percentage. In the case marked *, the percent change is ill-defined because the nominal estimate is zero; the estimate is 0.5 ± 0.4 when the MC subtraction is varied by -50%. The systematic uncertainty from closure is then calculated using 0.5 as the central value.

Selection	MC	Closure	LSB range	Total
$\geq$ 1 b, Loose, SB	10	28	2	30
$\geq$ 1 b, Loose, SIG	29	102	2	106
$\geq$ 1 b, Tight, SB	8	71	10	72
$\geq$ 1 b, Tight, SIG	73	213	10	225
$\geq$ 2 b,Loose, SB	21	69	2	72
$\geq$ 2 b,Loose, SIG	*	1156	*	*
$\geq$ 2 b,Tight, SB	19	199	10	200
$\geq$ 2 b,Tight, SIG	34	370	10	371

region "C" provides the estimate of the QCD background in the SB.

Applying corrections for the non-QCD components of the SIG-LDP and SB-LDP, our estimates of the QCD yields in the SIG and SB regions are thus

$$N_{SIG}^{QCD} = \frac{N_{LSB}}{N_{LSB-LDP}} \times (N_{SIG-LDP} - N_{SIG-LDP}^{top,MC} - N_{SIG-LDP}^{EW,MC}), \quad (1)$$

$$N_{SB}^{QCD} = \frac{N_{LSB}}{N_{LSB-LDP}} \times (N_{SB-LDP} - N_{SB-LDP}^{top,MC} - N_{SB-LDP}^{EW,MC}), \quad (2)$$

where the LSB results are derived from the zero b-jet, pre-scaled H_{T} trigger sample as stated above. The N_{SB}^{QCD} result is used in Sect. 6.3.

An MC closure test of the method is presented in Table 3. To verify that the closure results are not sensitive to details of the MC, we repeat this test after reweighting the MC to account for discrepancies in the jet multiplicity distributions between data and MC. A similar level of closure is observed. We also repeat the test for a QCD sample generated with the HERWIG [18] Monte Carlo and find similar closure.

The systematic uncertainties are summarized in Table 4. We consider:

- Use of MC to subtract top and EW contamination from the SIG-LDP and SB-LDP: We evaluate systematic uncertainties for the MC predictions as described in Sect. 8.1

Table 5: Estimates of the QCD background in the sideband (SB) and signal (SIG) regions. N^{MC}_{LDP-h} is the sum of the top and EW contributions from MC. The uncertainties are statistical only.

Selection	N_{LSB}	$N_{LSB-LDP}$	N_{LDP-h}	N^{MC}_{LDP-h}	$N_{predicted}$
≥ 1 b, Loose, SB	1462	9671	393	66 ± 3	49 ± 3
≥ 1 b, Loose, SIG	1462	9671	89	33 ± 2	9 ± 1
≥ 1 b, Tight, SB	341	4204	297	42 ± 2	21 ± 2
≥ 1 b, Tight, SIG	341	4204	6	3.6 ± 0.4	0.2 ± 0.2
≥ 2 b, Loose, SB	1462	9671	47	13.9 ± 0.8	5 ± 1
≥ 2 b, Loose, SIG	1462	9671	7	7.0 ± 0.8	0.0 ± 0.4
≥ 2 b, Tight, SB	341	4204	36	10.0 ± 0.7	2.1 ± 0.5
≥ 2 b, Tight, SIG	341	4204	2	0.8 ± 0.2	0.1 ± 0.1

and assign an uncertainty of $\pm 50\%$, where the largest contributions are associated with the jet energy scale and b-tagging efficiency.

- Assumption that $E_{\mathrm{T}}^{\mathrm{miss}}$ and $\Delta\phi_{\mathrm{N}}^{\mathrm{min}}$ are uncorrelated: This is evaluated using the closure test. We compute $(N_{true} - N_{predicted})/N_{predicted}$ (see Table 3) and assign the result added in quadrature with its statistical uncertainty as a symmetric systematic uncertainty. This test is performed for the standard MC and also for the reweighted MC described above; we take the larger discrepancy as the uncertainty.
- Definition of the LSB: We shift the LSB by $\pm 10\,\mathrm{GeV}$, which alters the number of events in the LSB by more than a factor of two for each shift. We take the larger of the observed changes as the systematic uncertainty.

Our estimates of the QCD background are presented in Table 5.

6.2 Z+jets Background

Events with a Z boson and one or more b jets present an irreducible background when the Z decays to two neutrinos. We evaluate this background by reconstructing $\mathrm{Z} \to \ell^+\ell^-$ events ($\ell = \mu$ or e) and removing the reconstructed leptons.

Fits are performed to determine the yields of $\mathrm{Z} \to \mu^+\mu^-$ and $\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$ events. The yields are corrected for background and efficiency. The overall efficiency ϵ is

$$\epsilon = \mathcal{A} \cdot \epsilon^2_{\ell\,\mathrm{reco}} \cdot \epsilon_{\mathrm{trig}} \cdot \epsilon^2_{\ell\,\mathrm{sel}}. \tag{3}$$

The acceptance $\mathcal{A}$ is determined from MC, while the lepton reconstruction $\epsilon_{\ell\,\mathrm{reco}}$, trigger ϵ_{trig}, and lepton selection $\epsilon_{\ell\,\mathrm{sel}}$ efficiencies are determined from data. The efficiency-corrected $\mathrm{Z} \to \ell^+\ell^-$ yields are used to estimate the $\mathrm{Z} \to \nu\bar{\nu}$ background through scaling by the ratio of branching fractions, $\mathrm{BR}(\mathrm{Z} \to \nu\bar{\nu})/\mathrm{BR}(\mathrm{Z} \to \mu^+\mu^-) = 5.95 \pm 0.02$ [19].

In the case of the loose ≥ 1 b-jet selection, the $\mathrm{Z} \to \ell^+\ell^-$ event yields are obtained directly after applying all nominal event selection criteria. For the other, more restrictive, selection conditions, there are no reconstructed $\mathrm{Z} \to \ell^+\ell^-$ events in data after all criteria are applied. In these cases, we loosen either the H_{T} or the H_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$ restrictions to increase the available statistics. The MC is used to scale the observed yields to those expected for the full selection. This method relies on an accurate MC representation of the H_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions. A comparison of data to MC is presented in Fig. 8: To retain enough events for a meaningful comparison, the following selection is applied for all variables except the one plotted: ≥ 2 jets, ≥ 1 b-jet, $H_{\mathrm{T}} \geq 100\,\mathrm{GeV}$, $E_{\mathrm{T}}^{\mathrm{miss}} \geq 50\,\mathrm{GeV}$, and $\Delta\phi_{\mathrm{N}}^{\mathrm{min}} \geq 4$.

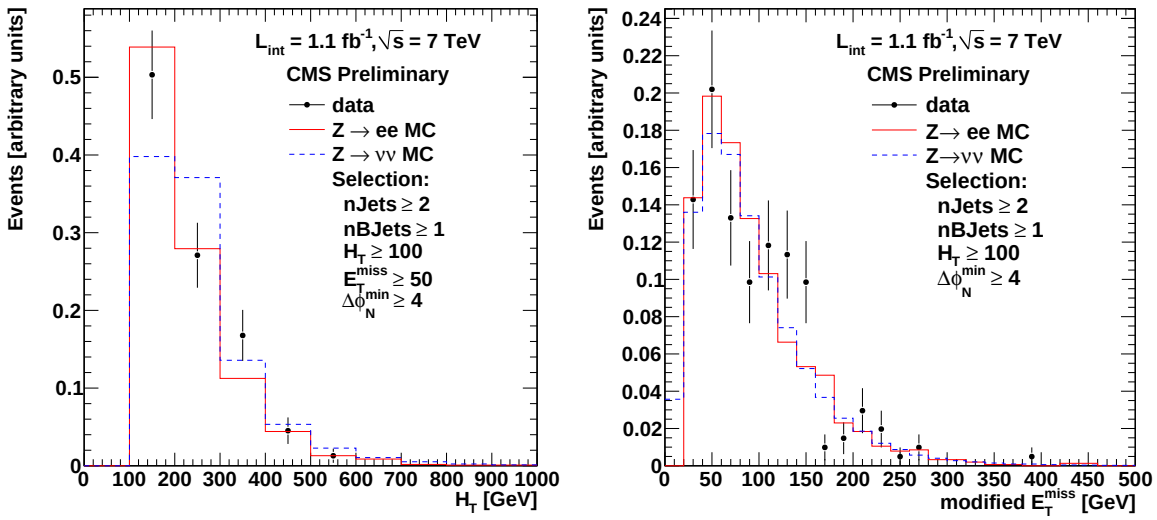


Figure 8: Comparison of data and MC for reconstructed (a,b) $Z \to e^+e^-$ and (c,d) $Z \to \mu^+\mu^-$ events, with loosened selection criteria as described in the text for all but the variable plotted: (a,c) H_T, and (b,d) E_T^{miss}.

A total of 4 $Z \to \mu^+\mu^-$ and 2 $Z \to e^+e^-$ events are found in the loose ≥ 1 b-jet SIG region. For the corresponding SB region we find 2 $Z \to \mu^+\mu^-$ and 3 $Z \to e^+e^-$ events.

For the tight ≥ 1 b-jet selection, we use the sum of the $Z \to \ell^+\ell^-$ results in the loose SIG and SB regions (i.e., $H_T \geq 350$ GeV and $E_T^{miss} \geq 150$ GeV, plus all other nominal selection criteria) and scale the result by the ratio of $Z \to \nu\bar{\nu}$ MC events reconstructed in the tight SIG region to the number in the $H_T \geq 350$ GeV, $E_T^{miss} \geq 150$ GeV region. A similar procedure is used to obtain the tight SB $Z \to \nu\bar{\nu}$ background estimate.

For both the loose and tight ≥ 2 b-jet selections, we perform the $Z \to \ell^+\ell^-$ reconstruction in the region with $H_T \geq 100$ GeV and $E_T^{miss} \geq 150$ GeV, with all the other nominal selection criteria.

Table 6: MC closure test of the $Z \to \nu\bar{\nu}$ background evaluation method. nJets and nB refer to the minimum numbers of jets and b-jets required, respectively. The closure test is based on comparison of the last two columns.

Selection					$Z \to \mu^+\mu^-$ channel			MC truth
nJets	H_T (GeV)	E_T^{miss} (GeV)	$\Delta\phi_N^{min}$	nB	N observed	ϵ	$Z \to \nu\bar{\nu}$ estimate	
≥ 3	≥ 350	≥ 150	≥ 4	≥ 1	30 ± 5	0.569 ± 0.063	314 ± 67	386 ± 8
≥ 3	≥ 350	≥ 200	≥ 4	≥ 1	16 ± 4	0.569 ± 0.063	167 ± 46	218 ± 6
≥ 1	≥ 300	≥ 150	≥ 4	≥ 1	101 ± 10	0.569 ± 0.063	1056 ± 157	1092 ± 13
≥ 3	≥ 300	≥ 50	≥ 4	≥ 1	116 ± 11	0.469 ± 0.055	1392 ± 181	1365 ± 14
≥ 3	≥ 300	≥ 150	≥ 4	≥ 1	42 ± 6	0.569 ± 0.063	439 ± 83	502 ± 9
≥ 1	≥ 300	≥ 150	≥ 4	≥ 2	5 ± 2	0.569 ± 0.063	52 ± 24	85 ± 4
≥ 3	≥ 300	≥ 50	≥ 4	≥ 2	14 ± 4	0.469 ± 0.055	168 ± 47	165 ± 5
Selection					$Z \to e^+e^-$ channel			MC truth
nJets	H_T (GeV)	E_T^{miss} (GeV)	$\Delta\phi_N^{min}$	nB	N observed	ϵ	$Z \to \nu\bar{\nu}$ estimate	
≥ 3	≥ 350	≥ 150	≥ 4	≥ 1	32 ± 6	0.443 ± 0.035	430 ± 83	386 ± 8
≥ 3	≥ 350	≥ 200	≥ 4	≥ 1	16 ± 4	0.443 ± 0.035	215 ± 56	218 ± 6
≥ 1	≥ 300	≥ 150	≥ 4	≥ 1	105 ± 10	0.443 ± 0.035	1410 ± 177	1092 ± 13
≥ 3	≥ 300	≥ 50	≥ 4	≥ 1	98 ± 10	0.416 ± 0.049	1402 ± 218	1365 ± 14
≥ 3	≥ 300	≥ 150	≥ 4	≥ 1	44 ± 7	0.443 ± 0.035	591 ± 101	502 ± 9
≥ 1	≥ 300	≥ 150	≥ 4	≥ 2	12 ± 3	0.443 ± 0.035	161 ± 48	85 ± 4
≥ 3	≥ 300	≥ 50	≥ 4	≥ 2	16 ± 4	0.416 ± 0.049	229 ± 63	165 ± 5

Table 7: Systematic uncertainties for the $Z \to \nu\bar{\nu}$ background estimate, in percentage.

Contribution	$Z \to \mu^+\mu^-$	$Z \to e^+e^-$
MC extrapolation	$0 - 100$	$0 - 100$
Background subtraction	18	20
Acceptance	2	2
Trigger efficiency	3	3
Lepton selection efficiency	5	5
MC closure	19	11
Total without extrapolation	27	24
Total with 50% extrapolation uncertainty	57	55
Total with 100% extrapolation uncertainty	104	103

The MC is used to obtain scale factors between this fit region and the loose and tight ≥ 2 b-jet selection regions.

Results from an MC closure test are shown in Table 6 for a variety of selection conditions. In comparison to the true MC values, the closure results are generally low for $Z \to \mu^+\mu^-$ events and high for $Z \to e^+e^-$ events. Since all of the selections overlap, rows in the $\mu^+\mu^-$ closure table are correlated with each other and can be expected to show the same general trend, and similarly for the e^+e^- table. The closure is generally valid to within 1 standard deviation and therefore does not indicate an observable bias.

The systematic uncertainties are summarized in Table 7. We consider:

- MC extrapolation: We assign an uncertainty of 50% (100%) for MC scale factors greater (less) than 0.10 based on the consistency of the results with alternative, independent scaling procedures. This is the dominant systematic uncertainty.
- Background subtraction: We quote the typical statistical uncertainty that we obtain in fits to the Z mass peak. We estimate an 18% (20%) uncertainty on the fraction of background in the fitted $Z \to \mu^+\mu^-$ ($Z \to e^+e^-$) yields.

Table 8: Estimates for the $Z \to \nu\bar{\nu}$ background compared to MC predictions for various conditions. The uncertainties are statistical only. nB refers to the minimum number of b-jets required.

Selection	Fit region			$Z \to \mu^+\mu^-$ channel				$Z \to \nu\bar{\nu}$
	H_T (GeV)	E_T^{miss} (GeV)	$\geq$ nB	Events	Yield	scale factor	$Z \to \nu\bar{\nu}$ estimate	MC truth
"Loose" SB $\geq$ 1 b	≥ 350	$150 - 200$	1	2	1.4	1.0	16.0 ± 11.4	10.2 ± 0.9
"Loose" Sig $\geq$ 1 b	≥ 350	≥ 200	1	4	2.8	1.0	32.0 ± 16.4	15.5 ± 1.1
"Tight" SB $\geq$ 1 b	≥ 350	≥ 150	1	6	4.2	0.145	7.0 ± 2.9	4.2 ± 0.6
"Tight" Sig $\geq$ 1 b	≥ 350	≥ 150	1	6	4.2	0.103	4.9 ± 2.1	2.7 ± 0.5
"Loose" SB $\geq$ 2 b	≥ 100	≥ 150	2	4	2.8	0.209	6.7 ± 3.5	1.7 ± 0.1
"Loose" Sig $\geq$ 2 b	≥ 100	≥ 150	2	4	2.8	0.258	8.3 ± 4.3	2.1 ± 0.2
"Tight" SB $\geq$ 2 b	≥ 100	≥ 150	2	4	2.8	0.050	1.6 ± 0.8	0.4 ± 0.1
"Tight" Sig $\geq$ 2 b	≥ 100	≥ 150	2	4	2.8	0.034	1.1 ± 0.6	0.3 ± 0.1
Selection	Fit region			$Z \to e^+e^-$ channel				$Z \to \nu\bar{\nu}$
	H_T (GeV)	E_T^{miss} (GeV)	$\geq$ nB	Events	Yield	scale factor	$Z \to \nu\bar{\nu}$ estimate	MC truth
"Loose" SB $\geq$ 1 b	≥ 350	$150 - 200$	1	3	2.0	1.0	30.2 ± 17.8	10.2 ± 0.9
"Loose" Sig $\geq$ 1 b	≥ 350	≥ 200	1	2	1.3	1.0	20.1 ± 15.9	15.5 ± 1.1
"Tight" SB $\geq$ 1 b	≥ 350	≥ 150	1	5	3.3	0.145	7.3 ± 3.4	4.2 ± 0.6
"Tight" Sig $\geq$ 1 b	≥ 350	≥ 150	1	5	3.3	0.103	5.2 ± 2.4	2.7 ± 0.5
"Loose" SB $\geq$ 2 b	≥ 100	≥ 150	2	0	0	0.209	0	1.7 ± 0.1
"Loose" Sig $\geq$ 2 b	≥ 100	≥ 150	2	0	0	0.258	0	2.1 ± 0.2
"Tight" SB $\geq$ 2 b	≥ 100	≥ 150	2	0	0	0.050	0	0.4 ± 0.1
"Tight" Sig $\geq$ 2 b	≥ 100	≥ 150	2	0	0	0.034	0	0.3 ± 0.1

Table 9: Estimate of the $Z \to \nu\bar{\nu}$ background in the signal (SIG) regions. The first uncertainty is statistical and the second systematic. The corresponding MC predictions with statistical uncertainties are also shown.

Selection	$N_{predicted}$	MC
$\geq$ 1 b, Loose, SIG	$24.4 \pm 11.4 \pm 4.2$	15.5 ± 1.1
$\geq$ 1 b, Tight, SIG	$5.0 \pm 1.6 \pm 2.0$	2.7 ± 0.5
$\geq$ 2 b, Loose, SIG	$2.6 \pm 2.9 \pm 2.0$	2.1 ± 0.2
$\geq$ 2 b, Tight, SIG	$0.2 \pm 0.4 \pm 0.5$	0.3 ± 0.1

- Acceptance: In MC events, we recalculate the acceptance after varying the p_T and η ranges of charged leptons subtracted from E_T^{miss}. The largest difference of 2% from the nominal result is assigned as the systematic uncertainty.
- Trigger efficiency: A 3% uncertainty is assigned to cover the range of results obtained when applying the selection criteria for H_T, E_T^{miss}, $\Delta\phi_N^{min}$, the number of jets, or the number of b jets individually.
- Lepton selection efficiency: The data-based efficiency measurement relies on fitting reconstructed Z yields. A 5% uncertainty is evaluated based on using alternative fitting shapes and varying the selection criteria as described for the trigger efficiency systematic study. A typical variation in fitting shape is to use an exponential instead of a polynomial to describe background.
- MC closure: The MC closure test exhibits discrepancies in the loose SB region on the level of 19% (11%) for $Z \to \mu^+\mu^-$ ($Z \to e^+e^-$), which we assign as a systematic uncertainty.

Results for the separated $Z \to \mu^+\mu^-$ and $Z \to e^+e^-$ channels are summarized in Table 8 for various conditions. We observe good agreement between the predictions from data and MC. Our results for the combined $Z \to \mu^+\mu^-$ and $Z \to e^+e^-$ channels are given in Table 9 along with the corresponding MC results.

Table 10: MC closure test of the top and W+jets background evaluation method. "Loose" and "Tight" refer to the corresponding H_T requirement for the SB and SIG regions, and also to the E_T^{miss} requirement for the SIG region. The closure test is based on comparison of the last two columns.

	$N_{SIG,SL}$	$N_{SB,SL}$	N_{SB}	$N_{predicted}$	N_{true}
≥ 1 b, Loose	131 ± 4	193 ± 4	224 ± 5	153 ± 6	155 ± 4
≥ 1 b, Tight	15 ± 1	68 ± 3	82 ± 3	18 ± 2	18 ± 1
≥ 2 b, Loose	33 ± 2	52 ± 2	56 ± 2	35 ± 2	39 ± 2
≥ 2 b, Tight	4.1 ± 0.8	18 ± 1	21 ± 1	5 ± 1	3.8 ± 0.4

As a cross-check, we evaluate the $Z \to \nu\bar{\nu}$ background using a second approach, in which $Z \to \ell^+\ell^-$ events are selected with all the nominal loose and tight selection criteria except for a very loose b-tagging requirement. We determine a b-tagging scale factor from data, using a sample of events in which there are no Z or lepton candidates (to increase statistical precision), by measuring the ratio of the number of events that pass our nominal b-tagging requirements for ≥ 1 b or ≥ 2 b jets to the number that pass the very loose requirement. We verify that the output of the b-tagging algorithm is independent of H_T, E_T^{miss}, and the presence of a Z. The results from this independent check are consistent with those presented in Table 9.

6.3 Top and W+jets Background

The dominant background at high E_T^{miss} arises from $t\bar{t}$ events, as noted in Sect. 3. Single-top and W+jets events comprise smaller backgrounds with similar signatures. Almost all top and W+jets background arises either when a W decays leptonically to an e or a μ, with the e or μ unidentified or outside the acceptance of the analysis, or else when a W decays to a hadronically-decaying τ. We find empirically that the shape of the E_T^{miss} distribution is very similar for all $t\bar{t}$ backgrounds that enter the SIG or SB samples, e.g., whether the W from top decays to e, μ, or τ, whether a τ decays hadronically or leptonically, etc. Based on this observation, we employ a template method in which the shape of the E_T^{miss} distribution measured in a single-lepton (SL) data control sample is used to describe the shape of the E_T^{miss} spectrum for all $t\bar{t}$ event categories. We find a similar situation for single-top and W+jets events and therefore group $t\bar{t}$, single-top, and W+jets events into a single template. Note that single-top events comprise only a small component of the background. The template is normalized to the number of top plus W+jets events measured in the SB region, where the SB region is defined by $150 < E_T^{miss} < 200$ GeV for both the loose and tight selections.

The SL control sample is formed by inverting the lepton veto, i.e., we require exactly one e or one μ to be present based on the lepton identification criteria of Sect. 3 in a sample whose selection is otherwise the same as the SIG sample. Figure 9 shows the E_T^{miss} distribution of the SL sample, along with MC predictions. The distributions are seen to be dominated by $t\bar{t}$, with some contribution from W+jets.

Using MC, we verify that the E_T^{miss} distribution of the SL sample accurately describes the top and W+jets component of the E_T^{miss} distribution of the SIG sample. As an illustration, Fig. 10 shows a comparison based on the loose selection for the dominant $t\bar{t}$ component. Similar agreement is found for the tight selection.

Contributions to the SB region from QCD and $Z \to \nu\bar{\nu}$ events are given by the data-based estimates presented in Sects. 6.1 and 6.2. Small, residual contributions from miscellaneous backgrounds such as diboson events are subtracted using MC.

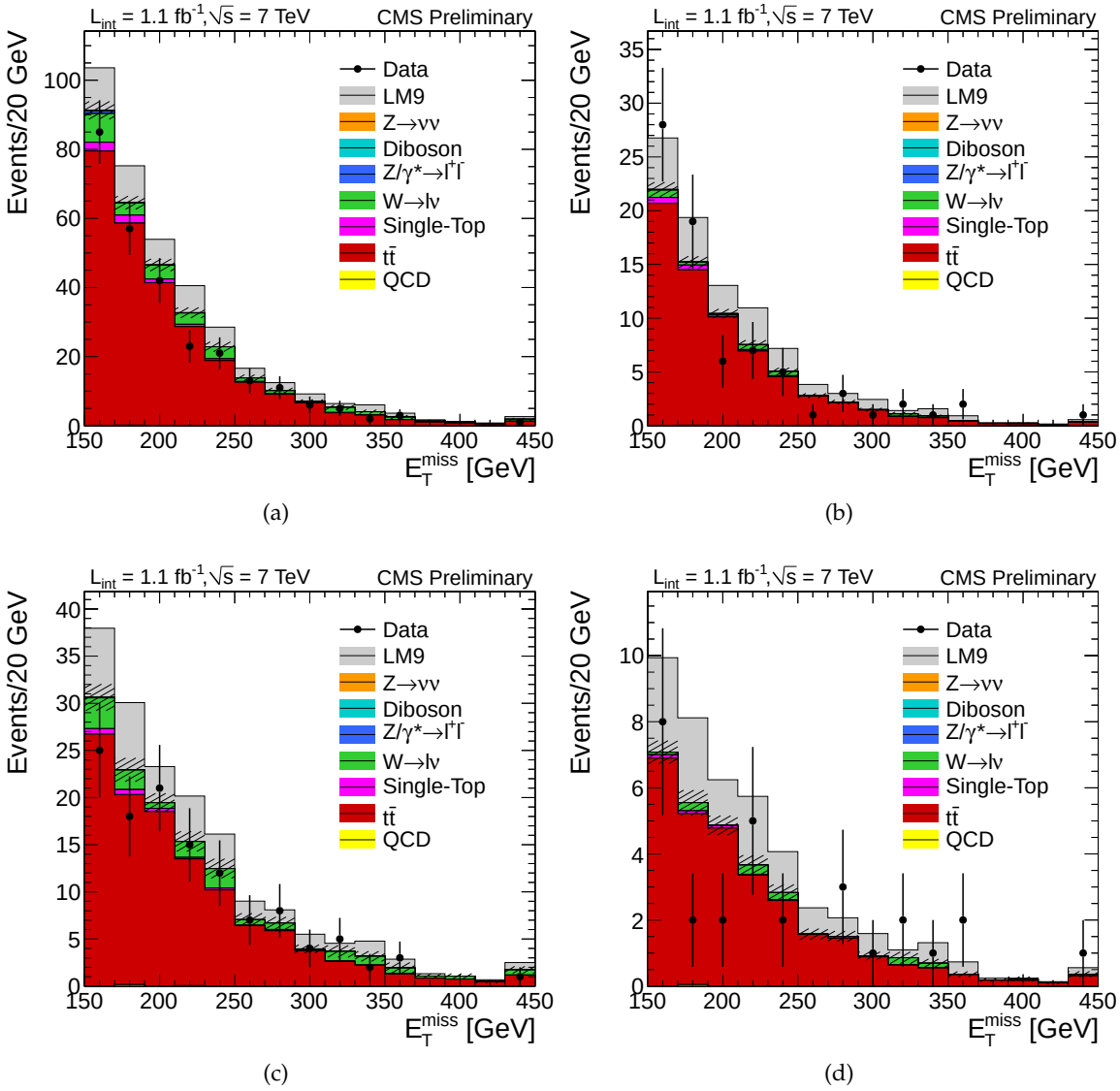


Figure 9: Distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ in data and MC for the single lepton (SL) control sample for the (a,b) loose and (c,d) tight selection with (a,c) ≥ 1 b-jets and (b,d) ≥ 2 b-jets. The hatched bands show the statistical uncertainty on the total standard model MC prediction.

Our estimate of the top and W+jets background in the SIG region is therefore

$$N_{SIG}^{top+W} = \frac{N_{SIG-SL}}{N_{SB-SL}} \times (N_{SB} - N_{SB}^{Z\to\nu\bar{\nu}} - N_{SB}^{QCD} - N_{SB}^{other,MC}). \quad (4)$$

Results from an MC closure test are presented in Table 10.

The systematic uncertainties are summarized in Table 11. We consider:

- MC closure: The systematic uncertainty is evaluated as $(N_{true} - N_{predicted})/N_{predicted}$ added in quadrature with its statistical uncertainty. We evaluate the closure using both the $\mathrm{t\bar{t}}$ sample only and the combined top and W+jets samples. The systematic uncertainty is derived from the result that closes least well.
- Subtraction of the data-driven QCD and $\mathrm{Z} \to \nu\bar{\nu}$ backgrounds: The data-driven

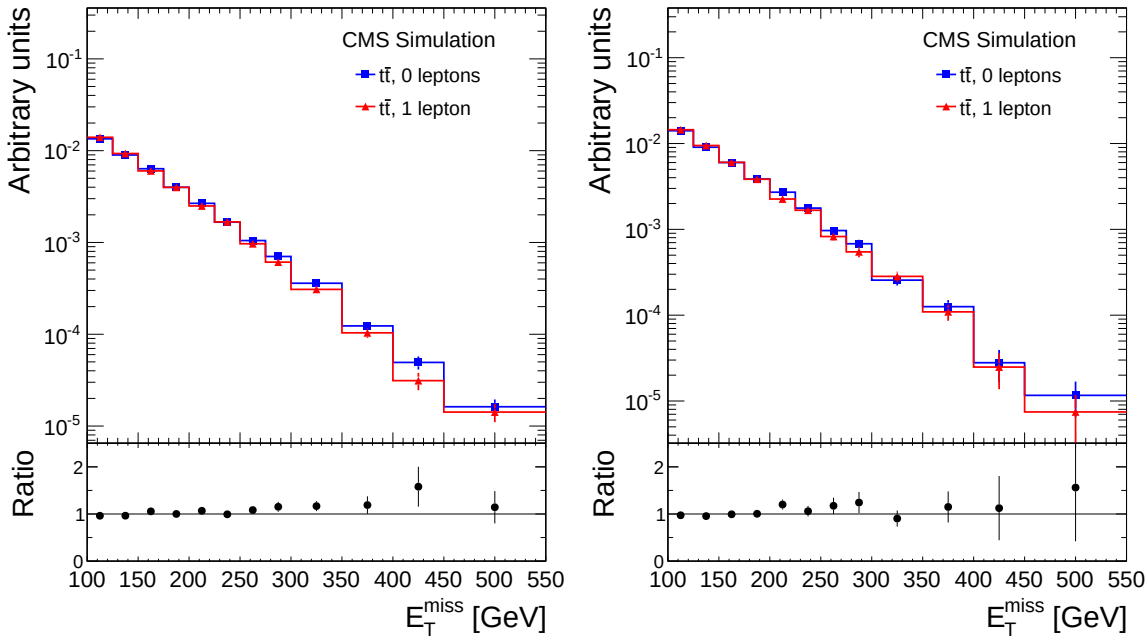


Figure 10: Normalized distributions of E_T^{miss} in $t\bar{t}$ MC for the loose selection with (a) ≥ 1 b-jets and (b) ≥ 2 b-jets. The square (blue) symbols show the results for the signal (SIG) sample. The triangle (red) symbols show those for the single lepton (SL) control sample. The small plots below the main figures show the ratio of the SIG to SL curves.

Table 11: Systematic uncertainties for the top and W+jets background estimate, in percentage.

		Contamination subtraction			
Selection	Closure	QCD	$Z \to \nu\bar{\nu}$	Other	Total
≥ 1 b, Loose	6	9	6	0.4	12
≥ 1 b, Tight	17	22	7	0.2	29
≥ 2 b, Loose	16	8	7	0.1	19
≥ 2 b, Tight	28	30	7	0.1	42

backgrounds in the SB region are varied by their total uncertainties. The fractional change in the top and W+jets prediction is taken as the systematic uncertainty.

- Subtraction of miscellaneous backgrounds: The MC-based backgrounds in the SB region are varied by their uncertainties. We assume $\pm 100\%$ uncertainty for these small terms.

Our estimate of the top and W+jets background is presented in Table 12.

6.4 Cross-Check of the Top and W+jets Background

This section presents results from a largely independent set of methods that use control samples in the data to determine the top and W+jets backgrounds contributing to the signal region. The E_T^{miss} distribution is determined separately for each of three background categories, which are defined according to the number and type of leptons from W-boson decay:

1. $t\bar{t}$, single-top, or W+jets events in which exactly one W decays into an e, μ, or into a τ that

Table 12: Combined top and W+jets background estimates for the signal (SIG) region. The uncertainties are statistical only.

	$N_{SIG,SL}$	$N_{SB,SL}$	N_{SB}	$N_{SB}^{subtracted}$	N_{SIG}
≥ 1 b, Loose	103	165	244	70 ± 10	108 ± 18
≥ 1 b, Tight	11	59	96	28 ± 3	13 ± 5
≥ 2 b, Loose	26	50	55	8 ± 3	24 ± 7
≥ 2 b, Tight	6	12	17	2.9 ± 0.8	7 ± 4

decays into an e or μ;

2. $t\bar{t}$, single-top, or W+jets events in which exactly one W decays into a τ that decays hadronically;

3. $t\bar{t}$ events in which both W bosons decay leptonically. This category is further subdivided, as described below.

When at least one b-tagged jet is required, these three categories represent, respectively, approximately 50%, 45%, and 5% of the expected total SM background from $t\bar{t}$, single top, and W+jets events, according to simulation. (The fractions for the ≥ 2 b-tag case are similar.) Within each category, the largest contribution arises from $t\bar{t}$ events; the W+jets contribution is smaller but is not negligible, especially at high $E_{\mathrm{T}}^{\mathrm{miss}}$. To measure the background in each of these three categories, a method is applied that uses a control sample in the data, in conjunction with certain information from simulated event samples.

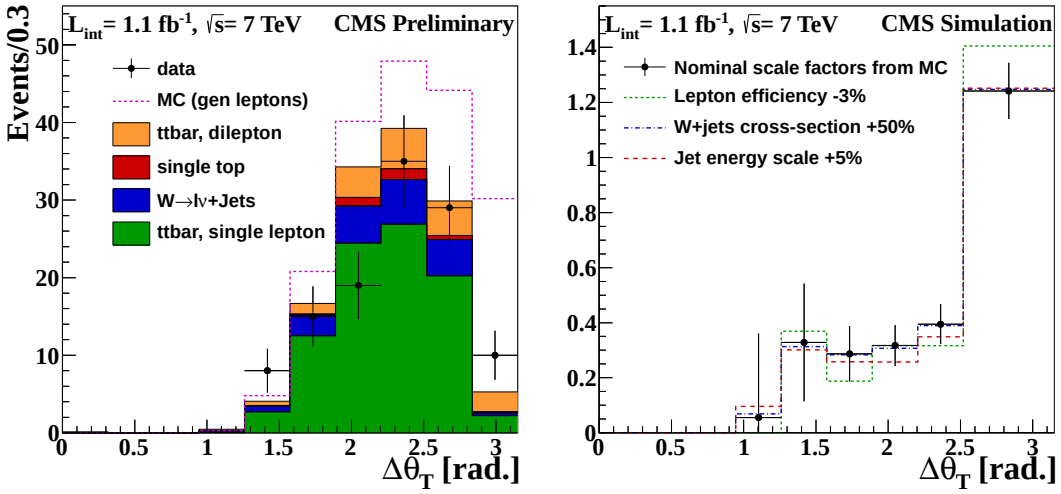
The cross-check analysis is applied only to the tight selection (for the cases of both ≥ 1 b-tagged jets and ≥ 2 b-tagged jets), because the method used to measure the category 2 background requires a control sample with a significantly lower H_{T} trigger threshold than that applied in the analysis, as explained in Sect. 6.4.2.

6.4.1 Polarization ($\Delta\theta_T$) Method: Category 1

For category 1 background events to contribute to the signal region, the electron or muon must either be missed entirely by the reconstruction, or it must fail to satisfy one or more of the criteria used to identify and veto events with leptons. The category 1 background contribution is measured using the single-lepton (SL) control sample (see Sect. 6.3).

To relate the event yields in the signal and control regions, we use constraints from knowledge of the polarization of the W boson. The W polarization controls the angular distribution of the leptons in the W rest frame. Because forward-going leptons are Lorentz-boosted to higher momentum, and backward-going leptons are Lorentz-boosted to lower momentum, the W polarization is directly related to the observed lepton momentum spectrum in the laboratory frame. Furthermore, because the neutrino is produced in association with the lepton in the two-body W decay, the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution is related to the polarization in an analogous manner.

The polarization of W bosons produced in $t\bar{t}$ events is predicted to high precision in SM calculations carried out to next-to-next-to leading order (NNLO) [20]. Expressed for the W^+, the polarization fractions are $f_0 = 0.687 \pm 0.005$, $f_{-1} = 0.311 \pm 0.005$, and $f_{+1} = 0.0017 \pm 0.0001$, where the subscripts indicate the W^+ helicity. These calculations have been confirmed by measurements from CDF [21] and D0 [22], although the experimental uncertainties are larger. For W+jets events, which represent only a small part of the background, the polarization fractions have been calculated to NLO [23] and are stable with respect to QCD corrections. The calculations are confirmed by recent measurements from CMS [24]. In contrast to W bosons in $t\bar{t}$



(a) (b)

Figure 11: Cross-check measurements of single-lepton backgrounds for the tight selection with ≥ 1 b-tagged jet. (a) Distribution of $\Delta\theta_T$ for events in the single-lepton control sample. The points with error bars are the data, and the stacked shaded histograms are the predictions from simulated event samples. The dashed histogram shows the prediction from the simulated event samples, but requires only that the lepton be generated, not reconstructed or satisfy the selection criteria. (b) Scale factors for the $\Delta\theta_T$ method. Black points with error bars indicate the nominal scale factor applied in each bin of $\Delta\theta_T$, with statistical errors. The colored histograms show how the scale factors change with the application of systematic variations.

events, the polarization of high p_{T} W bosons produced in W+jets events is nearly independent of their charge, and is predominantly left-handed.

To construct a distribution sensitive to the polarization of the W, we calculate the angle $\Delta\theta_T$ between the direction of the W in the lab frame and the direction of the e or μ in the W rest frame, all defined in the transverse plane. The p_{T} of the W is calculated by combining the lepton p_{T} vector with the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector. For the large values considered here, the $E_{\mathrm{T}}^{\mathrm{miss}}$ provides a good representation of the neutrino transverse momentum.

When $\Delta\theta_T$ is small, the lepton is produced along the transverse momentum direction of the W, typically resulting in a high p_{T} lepton and a low p_{T} neutrino (and therefore low $E_{\mathrm{T}}^{\mathrm{miss}}$) in the laboratory frame. Such high p_{T} leptons are usually observed in the single-lepton control sample, unless the $E_{\mathrm{T}}^{\mathrm{miss}}$ falls below the preselection requirement of $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV (chosen to give a trigger efficiency of nearly 100%), or if other criteria are not satisfied, such as the lepton isolation. Conversely, when $\Delta\theta_T$ is large, the lepton has lower p_{T} and the neutrino higher p_{T} in the laboratory frame. This decay configuration typically leads to larger $E_{\mathrm{T}}^{\mathrm{miss}}$ and to a lepton that is more likely to fall below the lepton p_{T} veto threshold. Thus, the event is more likely to contribute as a background to the signal region.

Figure 11 (a) shows the distribution of $\Delta\theta_T$ in the single-lepton control sample with ≥ 1 b-tagged jet and the tight selection applied, except with a looser $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement, $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV. The data are shown as points with error bars, while the stacked, shaded histograms show the expectations from SM simulated event samples, separated into various processes. These contributions can be compared with the total simulated $\mathrm{t\bar{t}}$, single-top, and W+jets category 1 events (only), regardless of whether the muon was reconstructed. These simulated

events, which are shown as the dashed histogram, are reconstructed in an identical manner to those in the stacked distributions except for the lepton, which need only be present at the generator level in the simulation.

The difference between the generated- and reconstructed-muon events from single-lepton processes represents category 1 events with one lost or ignored lepton. At the low end of the $\Delta\theta_T$ distribution, most of the generated events are in fact observed in the single-lepton control sample. At the upper end, most of the generated events feed down to the signal region.

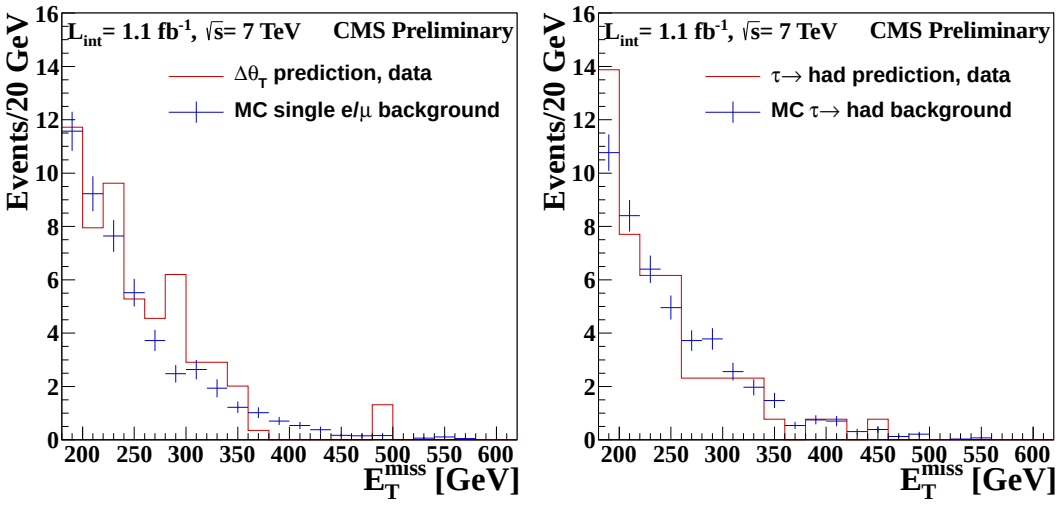
To predict the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of category 1 background events, we obtain the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of the single-lepton control sample in data separately in each bin of $\Delta\theta_T$. These $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions are rescaled according to the ratio of lost (=background) to observed (=SL control) events in the MC. The scale factors are defined in the MC as follows. The numerator is equal to the difference between the total yield from single-lepton processes (regardless of whether the lepton was reconstructed) and the subset of those events that have a reconstructed lepton satisfying the veto requirements. The denominator is equal to the number of MC events observed in the single-lepton channel from all sources. The definition of the denominator corresponds to the same observable in the data and implies that the small dilepton contribution is effectively divided out from this prediction. (It is explicitly included as the category 3 background.)

Figure 11 (b) shows the scale factors that are applied to the $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions from data in each $\Delta\theta_T$ bin. The rescaled $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions from the different $\Delta\theta_T$ bins are then summed to define the total $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for category 1 events. (Due to the small number of events at very high $\Delta\theta_T$, the final two bins are combined, and a common scale factor is applied to the corresponding $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions.) In this way, the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution is taken from the data, and the normalization is effectively determined by the yield of the single-lepton control sample in data, multiplied by ratios taken from MC that give the muon efficiency times acceptance, as a function of $\Delta\theta_T$.

Figure 12 (a) shows the predicted $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution from the method compared with that expected from the simulated event samples.

The systematic uncertainties on the factors used to rescale the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the single lepton control sample are summarized in Table 13. We propagate the effects of uncertainties on the lepton efficiencies, the relative $\mathrm{t\bar{t}}$ and W+jets cross sections, the jet energy scale, the jet energy resolution, and the shape of the W p_{T} spectrum. The lepton reconstruction efficiency systematic is calculated by removing 3% of electrons or muons passing the selection. The p_{T} of the lepton is added vectorially to the $E_{\mathrm{T}}^{\mathrm{miss}}$. The W+jets cross section is varied by $\pm 50\%$ to determine the sensitivity to the relative fractions of $\mathrm{t\bar{t}}$ and W+jets. For the jet energy scale, the p_{T} of all jets is changed by $\pm 5\%$. A corresponding change in $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated and added vectorially to the observed $E_{\mathrm{T}}^{\mathrm{miss}}$. This prescription, which is applied uniformly for all p_{T} and η values, is used for simplicity. The b-tagging efficiency is tested by changing the b-tagging efficiency data-to-MC scale factors by $\pm 1\sigma$. For the W p_{T} spectrum in $\mathrm{t\bar{t}}$ and W+jets, events are reweighted such that the upper 10% of the distribution is changed by $\pm 30\%$. The effects of a few of these systematic variations on the $\Delta\theta_T$ scale factors are shown in Fig. 11 (b). The dominant contributions arise from the lepton efficiency uncertainty and the $E_{\mathrm{T}}^{\mathrm{miss}}$ scale uncertainty.

In addition to the systematic uncertainties considered in Table 13 we also include a systematic uncertainty associated with the ability of the method to predict the correct yields in simulated event samples. The method is applied to an independent sample, and the uncertainty is taken as the sum in quadrature of two terms: the difference between the predicted and the true background yields and the statistical uncertainty on this difference. A similar procedure is also used



(a) (b)

Figure 12: Cross-check predictions for single-lepton top and W+jets backgrounds, starting from a single-lepton control sample. (a) Results from applying the polarization method to data. Red histogram: data-driven prediction. Blue points with error bars: expectation from simulated event samples. (b) Results from applying the τ →hadrons method to data. Red histogram: data-driven prediction. Blue points with error bars: expectation from simulated event samples..

Table 13: Systematic errors on the $\Delta\theta_T$ scale factors (SF), expressed as percent difference in average value.

Source of systematic error	Δ(SF)/SF (%)
Lepton efficiency	7.5
$\sigma(W \to \ell\nu)/\sigma(t\bar{t})$ ratio	0.2
Jet energy scale	$+6.1, -8.6$
b-tagging efficiency	0.5
Jet energy resolution	0.5
W p_T spectrum	$+0.9, -1.4$
Total	$+9.7, -11.5$

for the category 2 and 3 backgrounds. The tests with simulated samples all gave background predictions consistent with the true simulated yields, within the statistical uncertainty on the prediction.

6.4.2 Background from $\tau \to$ hadrons (Category 2)

Category 2 backgrounds, which involve hadronic τ decay, are measured using a procedure that effectively replaces the muon in a single-muon control sample with a jet that simulates a hadronic τ decay. As a consequence, a two-jet event can be transformed into a three-jet event.

Because of the addition of the jet from hadronic τ decay, one has to start with a control sample obtained with looser requirements than those applied in the actual analysis. The sample is obtained using a separate trigger that requires an isolated muon and has an online H_T requirement sufficiently low that all events satisfying the offline H_T requirement are recorded, even after the additional H_T associated with the simulated τ jet is added to the event. To be fully efficient, the offline p_T cut on the muon is increased to 20 GeV, and the relative isola-

tion requirement tightened. The event pre-selection is relaxed to $H_T > 200$ GeV, ≥ 2 jets with $p_T > 50$ GeV, and no E_T^{miss} or $\Delta\phi_N^{min}$ requirements. These quantities are recomputed after the muon-to-jet replacement and the corresponding requirements are applied.

The hadronic properties of events in the muon control sample are similar to those of category 2 background events, except for the visible energy of the τ jet in the detector. To account for this difference, the muons in the control sample are replaced by jets from a τ decay. The τ-jet momentum is obtained by randomly sampling a response template, taken from simulation, that provides the fraction of visible energy from the τ, computed relative to the underlying τ-lepton momentum. The resulting simulated τ-jet is then treated as any other jet in the event and is included in the calculation of H_T, the number of jets, and $\Delta\phi_N^{min}$, if the added jet has sufficient p_T. The nominal signal selection is then applied to the modified sample, and a factor taken from simulation is applied to account for the relative rates of reconstructed muons and hadronic τ decays. The same categories of systematic errors considered for the category 1 backgrounds are also evaluated here. The resulting E_T^{miss} distribution from the method is shown in Fig. 12 (b).

6.4.3 Dilepton Background (Category 3)

To measure the background from $t\bar{t}$ events in which both W bosons decay to leptons (e, μ, or τ), we use dilepton control samples. When both leptons are electrons or both are muons, or when one is an electron and the other a muon (where the electrons and muons can either be from W or from τ decay), we use simulated event samples to describe the shape of the E_T^{miss} distribution but measure the normalization from data. To determine the normalization, we count the number of dilepton events in data that pass loose (preselection) criteria for each of the three classes of events (ee, $\mu\mu$, or eμ) individually. This number is multiplied by a ratio, obtained from the simulated event samples, defined by the number of $t\bar{t}$ dilepton events that satisfy the final selection criteria (including the lepton vetoes) divided by the number of $t\bar{t}$ dilepton events satisfying the same loose requirements that are applied to the data. This case is an exception to our approach in which the shape of the E_T^{miss} distribution is taken from the data, but the number of background events is less than one.

When one or both of the leptons is a hadronically decaying τ, we apply the selection criteria used for the $\tau \rightarrow$ hadronic prediction on $e\mu$ and $\mu\mu$ events. The same procedure as the category 2 prediction is carried out, with one or both μ replaced by a τ jet using the response template. In this way, we obtain separate predictions for $\tau\tau$, $\mu\tau$, and eτ events where the τ decays hadronically.

As noted earlier, the background from dilepton events is small compared with other sources. The estimate has a large statistical uncertainty, which arises from the small size of the dilepton control sample.

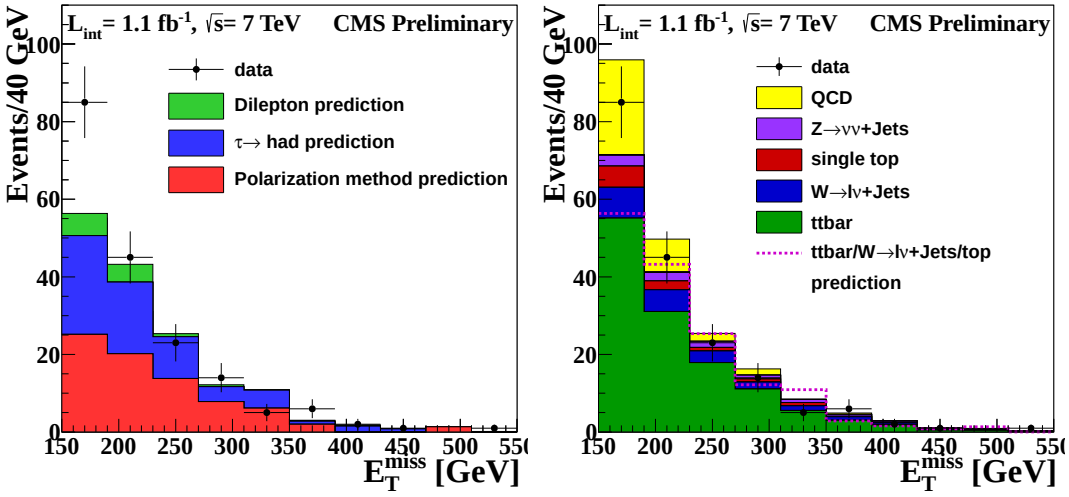
6.4.4 Total Cross-Check Predictions for $t\bar{t}$, Single Top, and W+jets Backgrounds

Table 14 summarizes the predicted backgrounds from the cross-check methods for the tight selection requirements.

Figure 13 (a) shows the predicted E_T^{miss} distributions for category 1, 2, and 3 backgrounds, together with the data, for the tight selection. Apart from the region at low E_T^{miss}, where there is a substantial contribution from QCD, these SM background categories account for most of the observed event yield in the data. Figure 13 (b) shows the expected background contributions from simulation, including the QCD component.

Table 14: Cross-check prediction of the background from $t\bar{t}$, single-top, and W+jets events, using the tight selection.

Category	≥1 b-tag		≥2 b-tags	
	Data Prediction	MC Prediction	Data Prediction	MC Prediction
Single lepton ($\Delta\theta_T$)	$9.5 \pm 3.3 \pm 1.2$	10.2	$4.7 \pm 2.3 \pm 1.2$	2.4
$W \to \tau \to$ had	$7.1 \pm 2.2^{+1.0}_{-1.3}$	9.2	$1.0 \pm 1.0 \pm 0.2$	1.8
Dilepton $t\bar{t}$	$0.4 \pm 0.2 \pm 0.3$	1.0	$0.2 \pm 0.2 \pm 0.2$	0.3
Total $t\bar{t}$ + (W+jets) + single top	$17.0 \pm 5.7 \pm 2.1$	20.4	$5.9 \pm 3.5 \pm 1.3$	4.5



(a) (b)

Figure 13: Cross-check predictions for the SM backgrounds contributing to the signal region with the tight selection (except for the final E_T^{miss} requirement) and ≥ 1 b-tagged jet. (a) Distributions of E_T^{miss} for the data (points with error bars) and for the predicted backgrounds based on control samples in the data (stacked histograms). The data are largely accounted for by this set of backgrounds, except at low values of E_T^{miss}, where the QCD contribution is significant. (The $Z \to \nu\bar{\nu}$ background also contributes across the full E_T^{miss} range.) (b) Distributions of E_T^{miss} for the data (points with error bars), simulated events (stacked histograms), and summed prediction for $t\bar{t}$, W+jets, and single top backgrounds (dashed histogram). The summed prediction matches well with the corresponding background contribution from simulation.

6.5 Summary of Individual Background Estimates

A summary of the background estimates from the individual techniques of Sects. 6.1-6.4 is presented in Table 15. The total number of background events is in agreement with the data. Note that our final SM background estimates are given in Sect. 8 and Table 16.

7 Likelihood Analysis

We construct a global likelihood function that simultaneously determines the SM background and the yield of a NP model. The background estimation techniques mirror those of Sects. 6.1-6.3. Details of the likelihood function are given in Appendix A. The likelihood analysis permits a unified evaluation of the SM backgrounds. It therefore provides a more consistent framework than the collection of individual results presented in Table 15. More importantly, it allows us to account for NP contamination of the signal and control regions in a comprehensive and

Table 15: SM background estimates from the procedures of Sects. 6.1-6.4 in comparison with the observed number of events in data. Our final SM background estimates are given in Table 16. Top and W+jets cross-check results are not available for the loose selection because of trigger restrictions. The first uncertainties are statistical and the second systematic.

	Loose search region		Tight search region	
	≥ 1 b	≥ 2 b	≥ 1 b	≥ 2 b
QCD	$9 \pm 1 \pm 9$	$0.0 \pm 0.4^{+5.8}_{-0.0}$	$0.2 \pm 0.2^{+0.5}_{-0.2}$	$0.1 \pm 0.1^{+0.4}_{-0.1}$
top and W+jets	$108 \pm 18 \pm 13$	$24 \pm 7 \pm 5$	$13 \pm 5 \pm 4$	$7 \pm 4 \pm 3$
top and W+jets cross-check	—	—	$17.0 \pm 5.7 \pm 2.1$	$5.9 \pm 3.5 \pm 1.3$
$Z \to \nu\overline{\nu}$	$24 \pm 11 \pm 4$	$2.6 \pm 2.9 \pm 2.0$	$5.0 \pm 1.6 \pm 2.0$	$0.2 \pm 0.4 \pm 0.5$
Total SM	$141 \pm 21 \pm 16$	$25.8 \pm 7.4^{+7.8}_{-5.2}$	$18.2 \pm 5.3 \pm 4.5$	$7.3 \pm 4.0 \pm 4.3$
Data	155	30	20	5

Table 16: SM background estimates from the likelihood fit, in comparison with the observed number of events in data. The uncertainties are the combined statistical and systematic terms. 95% CL upper limits on the number of observed CMSSM LM9 signal events evaluated with the CL_s method are also shown.

	Loose search region		Tight search region	
	≥ 1 b	≥ 2 b	≥ 1 b	≥ 2 b
QCD	$9.7\ ^{+10.1}_{-8.4}$	$0.0\ ^{+3.7}_{-0.0}$	$0.2\ ^{+0.8}_{-0.2}$	$0.1\ ^{+0.6}_{-0.1}$
top and W+jets	115 ± 15	24.5 ± 5.5	$13.9\ ^{+4.6}_{-4.3}$	$5.0\ ^{+2.4}_{-1.9}$
$Z \to \nu\overline{\nu}$	$29\ ^{+14}_{-11}$	$5.2\ ^{+4.6}_{-2.9}$	$5.3\ ^{+3.6}_{-2.7}$	$0.6\ ^{+0.9}_{-0.6}$
Total SM (LH)	152.8	29.7	19.5	5.7
Data	155	30	20	5
LM9 95% CL upper limit	91	21	20	7.3
LM9 MC	145	58	27	9.3

consistent manner.

The data are divided into 12 mutually exclusive bins corresponding to the boxes of Fig. 1, which are the 12 observables of the analysis. The likelihood model provides a prediction for the mean expected value of each observable in terms of the parameters of the signal and background components. The likelihood is the product of 12 Poisson probability distribution functions (PDF), one for each observable, and log-normal PDFs that account for systematic uncertainties and uncertainties on external parameters.

NP can contribute significantly to six of the 12 observables, corresponding to the four boxes labeled "Signal selection" and the two labeled "Single lepton" in Fig. 1. In our likelihood analysis, the relative contributions of NP to these six boxes are taken from the NP model under consideration. The NP yield in the SIG box is a free parameter. Thus the NP contribution to the other five boxes is tied to that of the SIG box. We perform MC exercises for both the loose and tight selections by setting the 12 observables to either the SM-only predicted values (based on MC) or to the SM+SUSY (LM9) values. The fitted yields in both cases agree with the true MC values for all the individual SM and NP terms.

95% confidence level (CL) upper limits are evaluated with the CL_s [25] method, taking into account the effects of variation of the external parameters and their correlations. We perform cross-checks (not presented below) using the profile likelihood technique. The upper limits we obtain with the latter method are slightly more conservative than those found with CL_s.

8 Results

The upper portion of Table 16 summarizes the results of the likelihood fit to data for the SM-only hypothesis. The SM backgrounds are consistent with those in Table 15 but have somewhat reduced uncertainties.

Since the number of events observed in the data is consistent with the expectations from the SM, we interpret our results in the context of limits on NP models.

8.1 CMSSM Test Point LM9

To illustrate the application of our likelihood method to a NP model, we consider the CMSSM test point LM9 (Table 2). For ≥ 1 b jets, the loose (tight) signal efficiency is 1.24% (0.24%). For ≥ 2 b jets, the corresponding efficiency is 0.53% (0.09%).

As systematic uncertainties on the signal efficiency, we consider:

- Jet energy scale: The jet energy scale is varied by p_{T} and η dependent uncertainties, which are 3.5% - 2% in the jet p_{T} range from 30 to 500 GeV.
- Jet energy resolution: The correction is varied by its uncertainty, which ranges between 2.4% and 6.6% depending on η.
- Unclustered energy: Transverse energy in an event not clustered into a physics object is varied by $\pm 10\%$ following standard CMS procedures.
- Pileup: The reweighting mentioned in Sect. 3 is varied by its uncertainty.
- b-tagging efficiency: The correction mentioned in Sect. 4 is varied by its uncertainty.
- Parton distribution functions: Uncertainties are evaluated following the PDF4LHC recommendations [26].
- Trigger efficiency: The efficiency is varied by its uncertainty, including a 2.5% uncertainty on the plateau efficiency.
- Lepton veto: The uncertainty is 2%.
- Anomalous $E_{\mathrm{T}}^{\mathrm{miss}}$: Removal of events with anomalous $E_{\mathrm{T}}^{\mathrm{miss}}$ caused by beam background and reconstruction effects yields an uncertainty of 1%.
- Luminosity: The uncertainty is 4.5% [27].

The systematic uncertainties are summarized in Table 17.

95% CL upper limits on the number of observed LM9 signal events are presented in the lower portion of Table 16. We exclude LM9 at 95% CL with all selections.

8.2 CMSSM Scan

We perform a scan in the CMSSM $m_{1/2}$ versus m_0 plane, for $\tan\beta = 40$, $A_0 = -500$ GeV, and $\mu > 0$. For each scan point, the SUSY particle spectrum is determined with the SoftSUSY [28] program. Events are generated at leading order with PYTHIA and normalized to the NLO cross section from PROSPINO. Systematic uncertainties on signal efficiencies are evaluated as described in Sect. 8.1. The jet energy scale, unclustered energy, parton distribution function, and b-tagging efficiency uncertainties are evaluated individually for each point in the scan. Other uncertainties are fixed to the values determined for LM9. At each scan point, we compute the $\mathrm{CL_s}$ value corresponding to the cross-section predicted by the model and exclude points with $\mathrm{CL_s} < 0.05$.

The observed 95% CL exclusion limits for the four selections are shown in Fig. 14. At low

Table 17: Systematic uncertainties, in percent, on the efficiency of the LM9 signal. The "Other" category includes the trigger efficiency, the lepton veto, and the anomalous $E_{\mathrm{T}}^{\mathrm{miss}}$ terms.

	Loose search region		Tight search region	
Source	≥ 1 b	≥ 2 b	≥ 1 b	≥ 2 b
Jet energy scale	7.7	8.6	12.1	13.7
Jet energy resolution	0.1	0.3	3.0	4.2
Unclustered energy	2.0	1.6	5.7	7.5
Pileup	3.4	3.1	4.3	4.2
b-tagging efficiency	6.5	15.8	7.1	17.2
Parton distribution functions	11.1	11.2	11.8	12.1
Other	3.5	3.5	3.5	3.5
Luminosity	4.5	4.5	4.5	4.5
Total uncertainty	16.5	22.2	20.7	27.5

m_0, the tight selection with either ≥ 1 b or ≥ 2 b jets provides the most stringent limits. At higher m_0, the most stringent limits arise from the loose ≥ 2 b selection. As an example, Fig. 15 shows the expected 95% CL exclusion curves based on the ≥ 1 b tight selection, along with the corresponding 1 standard deviation uncertainties and observed limits.

For values of m_0 above around 800 GeV our exclusion curves are similar to those of the inclusive SUSY analyses presented in Refs. [29, 30], which are based on the same data set. As the value of m_0 increases, b-jet production is enhanced. Therefore, in the context of the CMSSM, we expect our analysis to be most sensitive for large values of m_0. (Note that the CMSSM results presented in Refs. [29, 30] are based on $\tan\beta = 10$; however, the inclusive results are not very sensitive to the value of $\tan\beta$.)

8.3 Simplified Model T1bbbb

Simplified models [31–34] (SMS) provide a more general framework than the CMSSM to characterize NP signatures and interpret experimental results. They include only a few new particles and interactions and focus on generic topologies.

We consider the b-rich SMS model denoted T1bbbb. The T1bbbb event diagram is shown in Fig. 16. It is assumed that new strongly-interacting particles are pair-produced and that a weakly-interacting massive particle analogous to the SUSY LSP (labeled $\tilde{\chi}^0$ in Fig. 16) is produced in the subsequent decay chain. For convenience, we express SMS phenomenology using SUSY nomenclature. In T1bbbb, pair-produced gluinos each decay to two b-quarks and the LSP.

We present 95% CL upper limits on T1bbbb cross sections, and observed 95% CL exclusion curves based on NLO SUSY cross sections calculated with PROSPINO. For these latter results, separate limits are shown for the nominal cross section and for variations by a factor of 3 in each direction. Systematic uncertainties at each point are determined as described in Sect. 8.2.

The T1bbbb samples are generated with PYTHIA for a range of particle masses, providing a broader spectrum than the CMSSM where (for example) the ratio of the gluino to the LSP mass is approximately fixed. The cross section upper limits are presented as a function of the gluino and LSP masses.

The results are shown in Fig. 17. For each point in the plot, the selection that yields the best

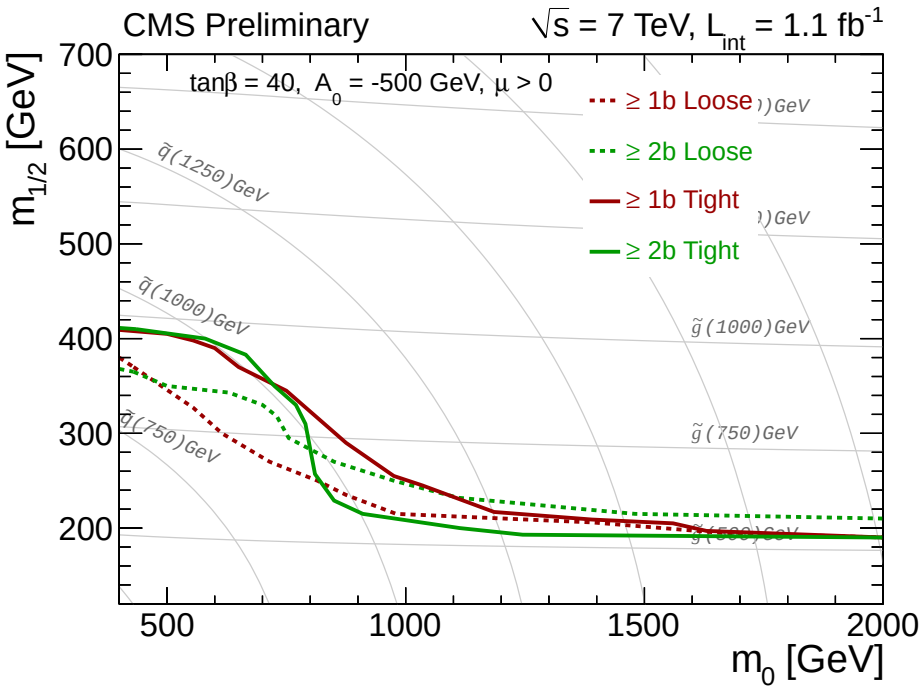


Figure 14: Observed 95% CL upper limits in the CMSSM $m_{1/2}$ versus m_0 mass plane, evaluated with the CL_s method. The fixed CMSSM parameters are $\tan\beta = 40$, $A_0 = -500$ GeV, and $\mu > 0$. Note that for the ≥ 2 b tight case, statistical fluctuations in the observed limits, combined with a conservative approach to drawing the exclusion curve, lead to a steep drop in the excluded region for $m_0 \approx 800$ GeV.

expected limit is chosen. Because the best expected result is dominated by two of the selection options, we simplify the procedure by selecting only from these two options, ≥ 1 b tight and ≥ 2 b loose. Figure 18 shows the signal efficiency for the selection that yields the best expected result, and the best expected selection itself. We do not present results for points near the $m_{\tilde{g}} = m_{\mathrm{LSP}}$ diagonal because we neglect uncertainties from initial-state radiation, which are large in this region. Signal contamination is treated in the same manner as for the rest of the analysis, but because the T1bbbb model contains no leptons, the pattern of signal contamination differs compared to the CMSSM study.

9 Summary

In this note, we present a search for an anomalous rate of events with three or more jets, at least one b-quark jet, no leptons, and large missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$. The principal standard model backgrounds, due to top and W+jets, Z+jets, and QCD events, are evaluated with data-based techniques. We introduce a variable $\Delta\phi_{\mathrm{N}}^{\mathrm{min}}$ that allows us to address the QCD multi-jet background with a simple approach. Our analysis is performed in a comprehensive likelihood framework, which permits us to account for new physics contamination of the signal and control regions in a consistent and unified manner.

We find no evidence for an excess of events above the expectation from the standard model and set limits on new physics in the context of the constrained minimal supersymmetric extension of the standard model, and also in the context of the generic simplified model T1bbbb, in which new particles decay to two b-quark jets plus an undetected particle.

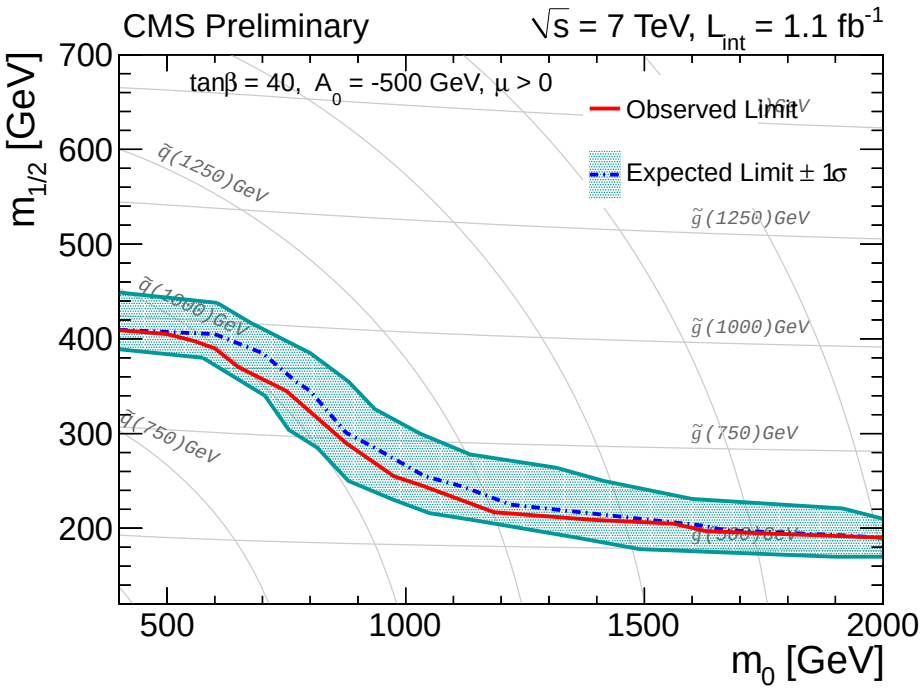


Figure 15: Observed and expected 95% CL upper limits in the CMSSM $m_{1/2}$ versus m_0 mass plane for the ≥ 1 b tight selection, evaluated with the CL_s method. The fixed CMSSM parameters are $\tan\beta = 40$, $A_0 = -500$ GeV, and $\mu > 0$. The 1-standard deviation uncertainties of the expected result are also indicated.

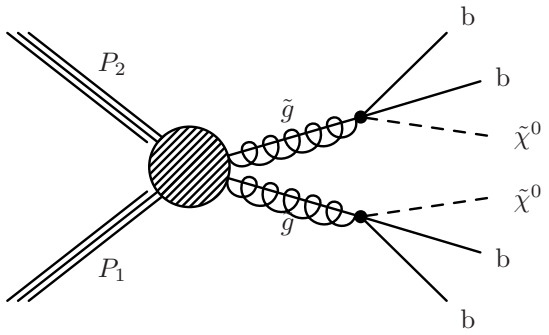


Figure 16: Diagram for the b-enriched simplified model T1bbbb.

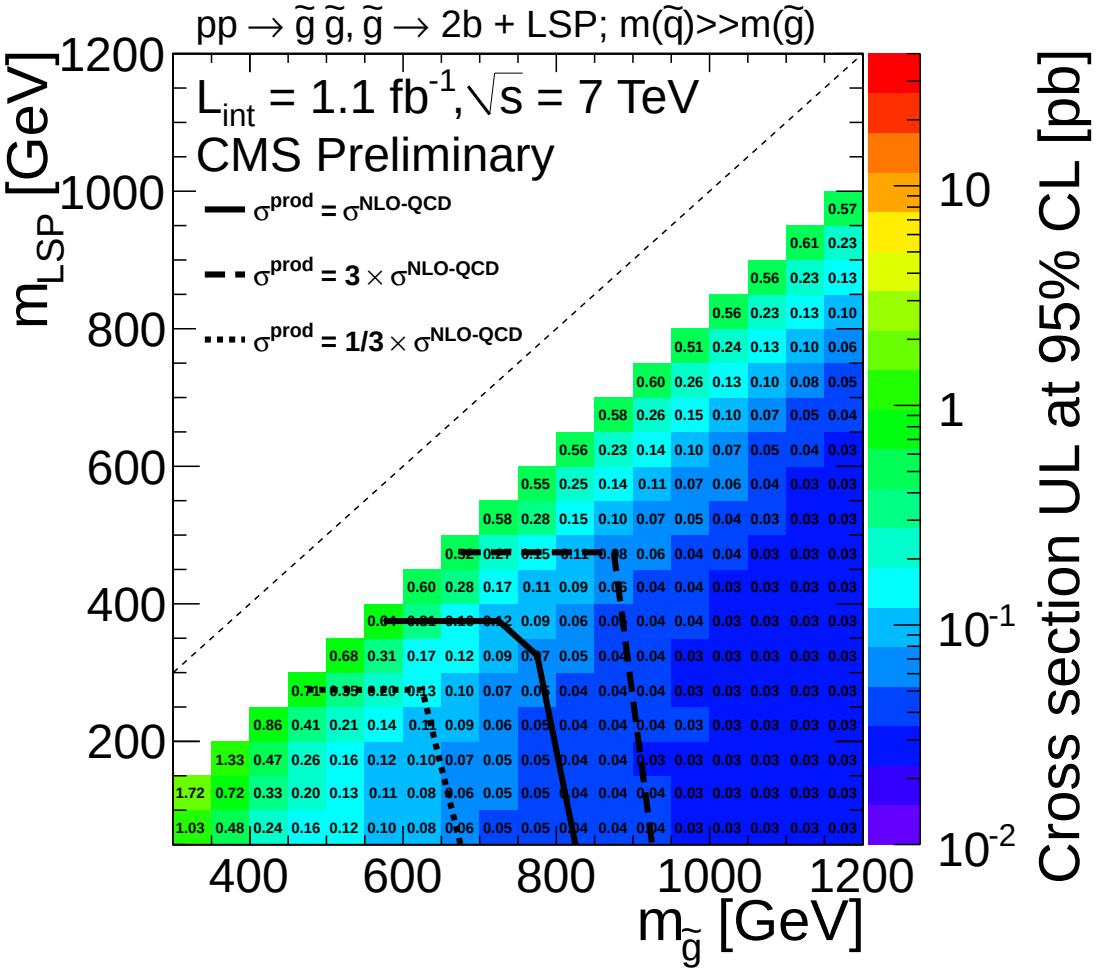
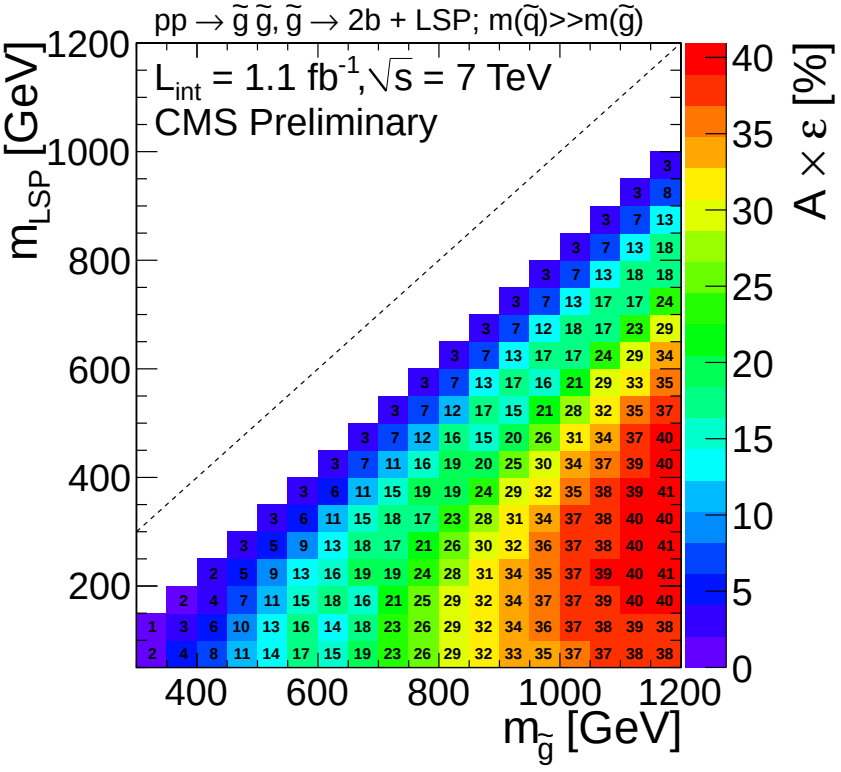
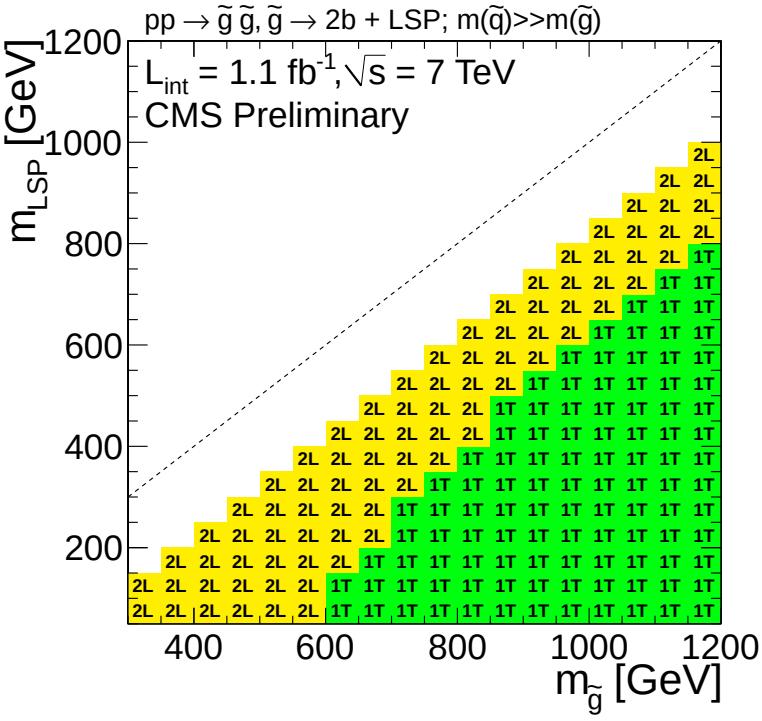


Figure 17: 95% CL cross section upper limits for the T1bbbb simplified model, evaluated with the CL_s method. For each point, we choose the selection that yields the best expected cross section limit, as described in the text. The contours indicate the bounds on the regions where the reference cross sections are excluded at 95% CL.



(a)



(b)

Figure 18: (a) The efficiency of the selection that provides the best expected cross section limit, and (b) the corresponding selection itself, where "1T" and "2L" correspond to the $\geq$ 1 b tight and $\geq$ 2 b loose selections, respectively.

References

[1] N. Arkani-Hamed, A.G. Cohen, and H. Georgi, "Electroweak Symmetry Breaking from Dimensional Deconstruction", *Phys. Lett.* **B513** (2001) 232. doi:10.1016/S0370-2693(01)00741-9.

[2] A.H. Chamseddine, R.L. Arnowitt, and P. Nath, "Locally Supersymmetric Grand Unification", *Phys. Rev. Lett.* **49** (1982) 970. doi:10.1103/PhysRevLett.49.970.

[3] R.L. Arnowitt and P. Nath, "Supersymmetric Mass Spectrum in SU(5) Supergravity Grand Unification", *Phys. Rev. Lett.* **69** (1992) 725. doi:10.1103/PhysRevLett.69.725.

[4] G. L. Kane, C.F. Kolda, L. Roszkowski and J. D. Wells, "Study of Constrained Minimal Supersymmetry", *Phys. Rev.* **D49** (1994) 6173. doi:10.1103/PhysRevD.49.6173.

[5] CMS Collaboration, "Search for Supersymmetry in Events with b Jets and Missing Transverse Energy at the LHC", *JHEP* **1107** (2011) 113. doi:10.1007/JHEP07(2011)113.

[6] ATLAS Collaboration, "Search for Supersymmetry in pp Collisions at $\sqrt{s} = 7$ TeV in Final States with Missing Transverse Momentum and b Jets", *Phys. Lett.* **B701** (2011) 398. doi:10.1016/j.physletb.2011.06.015.

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

[8] CMS Collaboration, "Particle Flow Event Reconstruction in CMS and Performance for Jets, Taus and MET", CMS PAS PFT-09-001.

[9] M. Cacciari, G. Salam, and G. Soyez, "The Anti-kt Jet Clustering Algorithm", *JHEP* **0804** (2008) 063. doi:10.1088/1126-6708/2008/04/063.

[10] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", arXiv:1107.4277. Submitted to JINST.

[11] GEANT 4 Collaboration, "GEANT4: A Simulation Toolkit", *Nucl. Instr. and Meth. A* **506** (2003) 250. doi:10.1016/S0168-9002(03)01368-8.

[12] J. Alwall et al., "MadGraph/MadEvent v4: The New Web Generation", *JHEP* **09** (2007) 28. doi:10.1088/1126-6708/2007/09/028.

[13] CMS Collaboration, "Combination of Top Pair Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV and Comparisons with Theory", CMS PAS TOP-11-001.

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

[15] CMS Collaboration, "CMS Technical Design Report, Volume II: Physics Performance", *J. Phys.* **G34** (2007) 995. doi:10.1088/0954-3899/34/6/S01.

[16] W. Beenakker, R. Hopker, M. Spira et al., "Squark and Gluino Production at Hadron Colliders", *Nucl. Phys.* **B492** (1997) 51. doi:10.1016/S0550-3213(97)80027-2.

[17] CMS Collaboration, "Performance of the b-Jet Identification in CMS", CMS PAS BTV-11-001.

[18] G. Corcella et al., "HERWIG 6: an Event Generator for Hadron Emission Reactions with Interfering Gluons (including Supersymmetric Processes)", *JHEP* **01** (2001) 010. doi:10.1088/1126-6708/2001/01/010.

[19] K. Nakamura et al., "Review of Particle Physics", *J. Phys.* **G37** (2010) 075021. doi:10.1088/0954-3899/37/7A/075021.

[20] A. Czarnecki, J. Körner, and J. Piclum, "Helicity Fractions of W bosons from Top Quark Decays at NNLO in QCD", *Phys. Rev.* **D81** (2010) 111503. doi:10.1103/PhysRevD.81.111503.

[21] CDF Collaboration, "Measurement of W Boson Polarization in Top quark Decay in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV", *Phys. Rev. Lett.* **105** (2010) 042002. doi:10.1103/PhysRevLett.105.042002.

[22] D0 Collaboration, "Measurement of the W Boson Helicity in Top Quark Decays using 5.4 fb^{-1} of $p\bar{p}$ Collision Data", *Phys. Rev. D.* **83** (2011) 032009. doi:10.1103/PhysRevD.83.032009.

[23] Z. Bern et al., "Left-handed W Bosons at the LHC", *Phys. Rev.* **D84** (2011) 034008. doi:10.1103/PhysRevD.84.034008.

[24] CMS Collaboration, "Measurement of the Polarization of W Bosons with Large Transverse Momentum in W+Jets Events at the LHC", *Phys. Rev. Lett.* **107** (2011) 021802. doi:10.1103/PhysRevLett.107.021802.

[25] A. Read, "Presentation of Search Results: the CL_s Technique", *J. Phys.G.* **28** (2002) 2693. doi:10.1088/0954-3899/28/10/313.

[26] J. Huston, R. Thorne, and S. Forte, "PDF4LHC Recommendations", (2011). http://www.hep.ucl.ac.uk/pdf4lhc/PDF4LHCrecom.pdf.

[27] CMS Collaboration, "Absolute Calibration of Luminosity Measurement at CMS: Summer 2011 Update", CMS PAS EWK-11-001.

[28] B. Allanach, "SOFTSUSY: a Program for Calculating Supersymmetric Spectra", *Comput. Phys. Commun.* **143** (2002) 305. doi:10.1016/S0010-4655(01)00460-X.

[29] CMS Collaboration, "Search for Supersymmetry at the LHC in Events with Jets and Missing Transverse Energy", CMS PAPER SUS-11-003, arXiv:1109.2352. Submitted to Phys. Rev. Lett.

[30] CMS Collaboration, "Search for Supersymmetry in All-Hadronic Events with Missing Energy", CMS PAS SUS-11-004.

[31] N. Arkani-Hamed et al., "MARMOSET: The Path from LHC Data to the New Standard Model via On-Shell Effective Theories", arXiv:hep-ph/0703088.

[32] J. Alwall, P. Schuster, and N. Toro, "Simplified Models for a First Characterization of New Physics at the LHC", *Phys. Rev.* **D79** (2009) 075020. doi:10.1103/PhysRevD.79.075020.

[33] J. Alwall, M.-P. Le, M. Lisanti et al., "Model-Independent Jets plus Missing Energy Searches", *Phys.Rev.* **D79** (2009) 015005, arXiv:0809.3264. doi:10.1103/PhysRevD.79.015005.

[34] D. Alves et al., "Simplified Models for LHC New Physics Searches", arXiv:1105.2838.

Table 18: Notation for the 12 observables of the likelihood.

Label	Description
SIG	Signal selection, $E_{\mathrm{T}}^{\mathrm{miss}}$ SIG range
SB	Signal selection, $E_{\mathrm{T}}^{\mathrm{miss}}$ SB range
SIG–LDP	Signal selection, $E_{\mathrm{T}}^{\mathrm{miss}}$ SIG range, low $\Delta\phi_{\mathrm{N}}^{\mathrm{min}}$
SB–LDP	Signal selection, $E_{\mathrm{T}}^{\mathrm{miss}}$ SB range, low $\Delta\phi_{\mathrm{N}}^{\mathrm{min}}$
SIG–SL	Single lepton selection, $E_{\mathrm{T}}^{\mathrm{miss}}$ SIG range
SB–SL	Single lepton selection, $E_{\mathrm{T}}^{\mathrm{miss}}$ SB range
LSB–0b	Zero b-tag sample, $E_{\mathrm{T}}^{\mathrm{miss}}$ LSB range ([50,100] GeV)
LSB–0b–LDP	Zero b-tag sample, $E_{\mathrm{T}}^{\mathrm{miss}}$ LSB range, low $\Delta\phi_{\mathrm{N}}^{\mathrm{min}}$
SIG–ee	$\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ selection, $E_{\mathrm{T}}^{\mathrm{miss}}$ SIG range
SB–ee	$\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ selection, $E_{\mathrm{T}}^{\mathrm{miss}}$ SB range
SIG–$\mu\mu$	$\mathrm{Z} \to \mu^{+}\mu^{-}$ selection, $E_{\mathrm{T}}^{\mathrm{miss}}$ SIG range
SB–$\mu\mu$	$\mathrm{Z} \to \mu^{+}\mu^{-}$ selection, $E_{\mathrm{T}}^{\mathrm{miss}}$ SIG range

A The Likelihood Function

This appendix describes details of the likelihood function. The treatment of the $\mathrm{Z} \to \nu\bar{\nu}$ background differs for the loose selection with ≥ 1 b-jets and the other selections. The default description below is for the former selection. The slightly different treatment we apply for the tighter selections is described in Sect. A.3.

A.1 Definition of the Likelihood Function

The 12 bins that define the observables of the analysis are summarized in Table 18. We use the following abbreviations

- SIG : the $E_{\mathrm{T}}^{\mathrm{miss}}$ signal region.
- SB : the $E_{\mathrm{T}}^{\mathrm{miss}}$ sideband region.
- LSB : the $E_{\mathrm{T}}^{\mathrm{miss}}$ low sideband region.
- LDP : the low $\Delta\phi_{\mathrm{N}}^{\mathrm{min}}$ region.
- SL : the single lepton selection.
- 0b : the zero b-tag selection.
- ee : the $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ selection.
- $\mu\mu$: The $\mathrm{Z} \to \mu^{+}\mu^{-}$ selection.

The notation used is as follows:

- The observed number of events in bin i is N_i.
- The analysis model prediction for the expected number of events in bin i is n_i.
- The contribution of component j to bin i is μ_i^j. For example, the QCD component in the SIG bin is μ_{SIG}^{QCD}.

Table 19 presents the prescription for the expected number of events n_i in each of the 12 bins as a function of the likelihood model parameters. Tables 20 and 21 specify the relationships between the likelihood model parameters for the 12 bins. Tables 22 and 23 list the fixed and floating parameters of the likelihood.

Table 19: Equations for the expected number of events n_i in each of the 12 bins as a function of the likelihood model parameters. The $SF_{\epsilon,i}$ parameters are multiplicative efficiency scale factors that are applied to the signal (SUSY) component and to the contributions from the MC in the SIG–LDP and SB–LDP bins. The SF_{MC} parameter is used to vary the normalization of the MC inputs. The P_{ee} and $P_{\mu\mu}$ parameters give the estimated purity (signal/total) of the $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ samples, respectively.

n_{SIG}	=	$\mu_{SIG}^{ttwj} + \mu_{SIG}^{QCD} + \mu_{SIG}^{Z\nu\nu} + SF_{\epsilon,sig}\,\mu_{SIG}^{SUSY}$
n_{SB}	=	$\mu_{SB}^{ttwj} + \mu_{SB}^{QCD} + \mu_{SB}^{Z\nu\nu} + SF_{\epsilon,sb}\,\mu_{SB}^{SUSY}$
$n_{SIG-LDP}$	=	$\mu_{SIG-LDP}^{QCD} + SF_{\epsilon,sig-ldp}\left\{ SF_{MC}\,(\mu_{SIG-LDP}^{tt-MC} + \mu_{SIG-LDP}^{wj-MC} + \mu_{SIG-LDP}^{Z\nu\nu-MC}) + \mu_{SIG-LDP}^{SUSY}\right\}$
n_{SB-LDP}	=	$\mu_{SB-LDP}^{QCD} + SF_{\epsilon,sb-ldp}\left\{ SF_{MC}\,(\mu_{SB-LDP}^{tt-MC} + \mu_{SB-LDP}^{wj-MC} + \mu_{SB-LDP}^{Z\nu\nu-MC}) + \mu_{SB-LDP}^{SUSY}\right\}$
n_{SIG-SL}	=	$\mu_{SIG-SL}^{ttwj} + SF_{\epsilon,sig-sl}\,\mu_{SIG-SL}^{SUSY}$
n_{SB-SL}	=	$\mu_{SB-SL}^{ttwj} + SF_{\epsilon,sb-sl}\,\mu_{SB-SL}^{SUSY}$
n_{LSB-0b}	=	μ_{LSB-0b}^{QCD}
$n_{LSB-0b-LDP}$	=	$\mu_{LSB-0b-LDP}^{QCD}$
n_{SIG-ee}	=	$\mu_{SIG-ee}^{Zee}/P_{ee}$
n_{SB-ee}	=	μ_{SB-ee}^{Zee}/P_{ee}
$n_{SIG-\mu\mu}$	=	$\mu_{SIG-\mu\mu}^{Z\mu\mu}/P_{\mu\mu}$
$n_{SB-\mu\mu}$	=	$\mu_{SB-\mu\mu}^{Z\mu\mu}/P_{\mu\mu}$

Uncertainties on external input parameters and systematic uncertainties are handled with floating multiplicative factors. In this discussion, we will refer to these parameters as nuisance parameters. All nuisance parameters are constrained by log-normal PDFs in the likelihood. In practice, this is done by deriving the nuisance parameter p from a primary variable g, which is constrained by a Gaussian PDF in the likelihood with zero mean and width one. The relationship between the nuisance parameter p and the primary variable g is given by

$$p = m\,[\exp(s/m)]^g\,, \tag{5}$$

where m and s are the mean and width parameters of the log-normal distribution, which correspond to the mean and width parameters of the Gaussian distribution in in the limit $s \ll m$. When the width parameter is small compared to the mean parameter, the log-normal distribution is quite close to a Gaussian distribution with the same mean and width.

Table 20: Relationships between the parameters of the likelihood. Entries in blue are free parameters of the likelihood. All SF_i parameters are floating parameters of the likelihood constrained by log-normal PDFs. Entries in red are inputs from MC.

Observable bin	ttwj	QCD
SIG	$\mu_{SIG}^{ttwj} = \mu_{SB}^{ttwj}\ SF_{ttwj}\ \left(\frac{\mu_{SIG-SL}^{ttwj}}{\mu_{SB-SL}^{ttwj}}\right)$	$\mu_{SIG}^{QCD} = \mu_{SIG-LDP}^{QCD}\ SF_{QCD-SIG}\ \left(\frac{\mu_{LSB-0b}^{QCD}}{\mu_{LSB-0b-LDP}^{QCD}}\right)$
SB	$\boldsymbol{\mu_{SB}^{ttwj}}$	$\mu_{SB}^{QCD} = \mu_{SB-LDP}^{QCD}\ SF_{QCD-SB}\ \left(\frac{\mu_{LSB-0b}^{QCD}}{\mu_{LSB-0b-LDP}^{QCD}}\right)$
$SIG-LDP$	$\mu_{SIG-LDP}^{tt-MC}$, $\mu_{SIG-LDP}^{wj-MC}$	$\boldsymbol{\mu_{SIG-LDP}^{QCD}}$
$SB-LDP$	μ_{SB-LDP}^{tt-MC}, μ_{SB-LDP}^{wj-MC}	$\boldsymbol{\mu_{SB-LDP}^{QCD}}$
$SIG-SL$	$\boldsymbol{\mu_{SIG-SL}^{ttwj}}$	—
$SB-SL$	$\boldsymbol{\mu_{SB-SL}^{ttwj}}$	—
$LSB-0b$	—	$\boldsymbol{\mu_{LSB-0b}^{QCD}}$
$LSB-0b-LDP$	—	$\boldsymbol{\mu_{LSB-0b-LDP}^{QCD}}$
$SIG-ee$	—	—
$SB-ee$	—	—
$SIG-\mu\mu$	—	—
$SB-\mu\mu$	—	—

Table 21: Relationships between the parameters of the likelihood (continued). Entries in blue are free parameters of the likelihood. All SF_i parameters are floating parameters of the likelihood constrained by log-normal PDFs. Entries in red are inputs from MC.

Observable bin	Znn	SUSY
SIG	$\mu_{SIG}^{Z\nu\nu}$	μ_{SIG}^{SUSY}
SB	$\mu_{SB}^{Z\nu\nu}$	$\mu_{SB}^{SUSY} = \mu_{SIG}^{SUSY} \left(\frac{\mu_{SB}^{SUSY-MC}}{\mu_{SIG}^{SUSY-MC}} \right)$
$SIG - LDP$	$\mu_{SIG-LDP}^{Z\nu\nu-MC}$	$\mu_{SIG-LDP}^{SUSY} = \mu_{SIG}^{SUSY} \left(\frac{\mu_{SIG-LDP}^{SUSY-MC}}{\mu_{SIG}^{SUSY-MC}} \right)$
$SB - LDP$	$\mu_{SB-LDP}^{Z\nu\nu-MC}$	$\mu_{SB-LDP}^{SUSY} = \mu_{SIG}^{SUSY} \left(\frac{\mu_{SB-LDP}^{SUSY-MC}}{\mu_{SIG}^{SUSY-MC}} \right)$
$SIG - SL$	—	$\mu_{SIG-SL}^{SUSY} = \mu_{SIG}^{SUSY} \left(\frac{\mu_{SIG-SL}^{SUSY-MC}}{\mu_{SIG}^{SUSY-MC}} \right)$
$SB - SL$	—	$\mu_{SB-SL}^{SUSY} = \mu_{SIG}^{SUSY} \left(\frac{\mu_{SB-SL}^{SUSY-MC}}{\mu_{SIG}^{SUSY-MC}} \right)$
$LSB - 0b$	—	—
$LSB - 0b - LDP$	—	—
$SIG - ee$	$\mu_{SIG-ee}^{Zee} = \mu_{SIG}^{Z\nu\nu}\, SF_{ee} \left(\frac{A_{ee}\epsilon_{ee}}{R_B} \right)$	—
$SB - ee$	$\mu_{SB-ee}^{Zee} = \mu_{SB}^{Z\nu\nu}\, SF_{ee} \left(\frac{A_{ee}\epsilon_{ee}}{R_B} \right)$	—
$SIG - \mu\mu$	$\mu_{SIG-\mu\mu}^{Z\mu\mu} = \mu_{SIG}^{Z\nu\nu}\, SF_{\mu\mu} \left(\frac{A_{\mu\mu}\epsilon_{\mu\mu}}{R_B} \right)$	—
$SB - \mu\mu$	$\mu_{SB-\mu\mu}^{Z\mu\mu} = \mu_{SB}^{Z\nu\nu}\, SF_{\mu\mu} \left(\frac{A_{\mu\mu}\epsilon_{\mu\mu}}{R_B} \right)$	—

Table 22: Fixed parameters of the likelihood.

Parameter	Description
$\mu_{SIG-LDP}^{wj-MC}$	MC prediction for W+jets in SIG–LDP in events, norm to data int. lum.
μ_{SB-LDP}^{wj-MC}	MC prediction for W+jets in SB–LDP in events, norm to data int. lum.
$\mu_{SIG-LDP}^{t\bar{t}-MC}$	MC prediction for $t\bar{t}$ in SIG–LDP in events, norm to data int. lum.
$\mu_{SB-LDP}^{t\bar{t}-MC}$	MC prediction for $t\bar{t}$ in SB–LDP in events, norm to data int. lum.
$\mu_{SIG-LDP}^{Z\nu\nu-MC}$	MC prediction for $t\bar{t}$ in SIG–LDP in events, norm to data int. lum.
$\mu_{SB-LDP}^{Z\nu\nu-MC}$	MC prediction for $t\bar{t}$ in SB–LDP in events, norm to data int. lum.
R_B	Ratio of Z decay branching ratios: $\mathrm{BR}(Z \to \nu\bar{\nu})/\mathrm{BR}(Z \to \mu^+\mu^-) = 5.95$
$\mu_{SIG}^{SUSY-MC}$	MC prediction for SUSY for a given model in SIG in events, norm to data int. lum.
$\mu_{SB}^{SUSY-MC}$	MC prediction for SUSY for a given model in SB in events, norm to data int. lum.
$\mu_{SIG-SL}^{SUSY-MC}$	MC prediction for SUSY for a given model in SIG–SL in events, norm to data int. lum.
$\mu_{SB-SL}^{SUSY-MC}$	MC prediction for SUSY for a given model in SB–SL in events, norm to data int. lum.
$\mu_{SIG-LDP}^{SUSY-MC}$	MC prediction for SUSY for a given model in SIG–LDP in events, norm to data int. lum.
$\mu_{SB-LDP}^{SUSY-MC}$	MC prediction for SUSY for a given model in SB–LDP in events, norm to data int. lum.

Table 23: Floating parameters of the likelihood.

Parameter	Description
μ_{SIG}^{SUSY}	SUSY yield in the SIG bin
$\mu_{SIG}^{Z\nu\nu}$	$Z \to \nu\bar{\nu}$ yield in the SIG bin
μ_{SB}^{ttwj}	Sum of $t\bar{t}$ and W+jets contributions to the SB bin
$\mu_{SB}^{Z\nu\nu}$	$Z \to \nu\bar{\nu}$ yield in the SB bin
$\mu_{SIG-LDP}^{QCD}$	QCD yield in the SIG–LDP bin
μ_{SB-LDP}^{QCD}	QCD yield in the SB–LDP bin
μ_{SIG-SL}^{ttwj}	Sum of $t\bar{t}$ and W+jets contributions to the SIG–SL bin
μ_{SB-SL}^{ttwj}	Sum of $t\bar{t}$ and W+jets contributions to the SB–SL bin
μ_{LSB-0b}^{QCD}	QCD yield in the LSB–0b bin
$\mu_{LSB-0b-LDP}^{QCD}$	QCD yield in the LSB–0b–LDP bin
g_{ϵ}	Primary variable for efficiency scale factors
g_{MC}	Primary variable for Scale factor for SM MC inputs
g_{ttwj}	Primary variable for Scale factor for *ttwj* SIG/SB ratio
$g_{QCD-SIG}$	Primary variable for Scale factor for QCD $\Delta\phi_{\mathrm{N}}^{\mathrm{min}}$ pass/fail ratio, SIG
g_{QCD-SB}	Primary variable for Scale factor for QCD $\Delta\phi_{\mathrm{N}}^{\mathrm{min}}$ pass/fail ratio, SB
g_{ee}	Primary variable for Scale factor, Closure for $Z \to e^+e^-$
$g_{\mu\mu}$	Primary variable for Scale factor, Closure for $Z \to \mu^+\mu^-$
g_{Pee}	Primary variable for Purity (signal/total) of $Z \to e^+e^-$ sample
$g_{P\mu\mu}$	Primary variable for Purity (signal/total) of $Z \to \mu^+\mu^-$ sample
g_{Aee}	Primary variable for Acceptance for $Z \to e^+e^-$
$g_{A\mu\mu}$	Primary variable for Acceptance for $Z \to \mu^+\mu^-$
$g_{\epsilon_{ee}}$	Primary variable for Efficiency for $Z \to e^+e^-$
$g_{\epsilon_{\mu\mu}}$	Primary variable for Efficiency for $Z \to \mu^+\mu^-$

Table 24: Nuisance parameters of the likelihood. The parameter is derived from the primary variable (g) using Eq. (5). The four numbers given in parentheses in the last column for the *ttwj*, QCD-SIG, and QCD-SB systematics are for the ≥ 1 b-loose, ≥ 2 b-tight, ≥ 1 b-loose, and ≥ 2 b-tight selections, respectively.

Parameter	Prim. var	Description	Mean	Width
$SF_{\epsilon,i}$	g_{ϵ}	Efficiency scale factors	1.0	depends on signal model
SF_{MC}	g_{MC}	SM MC inputs	1.0	0.50
SF_{ttwj}	g_{ttwj}	*ttwj* SIG/SB ratio	1.0	(0.06, 0.17, 0.16, 0.28)
$SF_{QCD-SIG}$	$g_{QCD-SIG}$	QCD $\Delta\phi_{\mathrm{N}}^{\mathrm{min}}$ pass/fail ratio, SIG	1.0	(1.02, 2.13, 11.56, 3.70)
SF_{QCD-SB}	g_{QCD-SB}	QCD $\Delta\phi_{\mathrm{N}}^{\mathrm{min}}$ pass/fail ratio, SB	1.0	(0.28, 1.02, 0.72, 2.13)
SF_{ee}	g_{ee}	Closure for $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$	1.0	0.11
$SF_{\mu\mu}$	$g_{\mu\mu}$	Closure for $\mathrm{Z} \to \mu^{+}\mu^{-}$	1.0	0.20
P_{ee}	g_{Pee}	Purity of $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ sample	0.651	0.130
$P_{\mu\mu}$	$g_{P\mu\mu}$	Purity of $\mathrm{Z} \to \mu^{+}\mu^{-}$ sample	0.711	0.130
A_{ee}	g_{Aee}	Acceptance for $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$	0.80	0.05
$A_{\mu\mu}$	$g_{A\mu\mu}$	Acceptance for $\mathrm{Z} \to \mu^{+}\mu^{-}$	0.85	0.04
ϵ_{ee}	$g_{\epsilon_{ee}}$	Efficiency for $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$	0.480	0.052
$\epsilon_{\mu\mu}$	$g_{\epsilon_{\mu\mu}}$	Efficiency for $\mathrm{Z} \to \mu^{+}\mu^{-}$	0.621	0.061

The likelihood $\mathcal{L}$ is the product of 12 Poisson PDFs, one for each observable $\mathcal{P}(N_i|n_i)$, and Gaussian PDFs $\mathcal{G}(g)$ to take into account the uncertainties on each nuisance parameter:

$$
\begin{aligned}
\mathcal{L} \;=\; & \mathcal{P}(N_{SIG}|n_{SIG}) \times \mathcal{P}(N_{SB}|n_{SB}) \\
\times\; & \mathcal{P}(N_{SIG-SL}|n_{SIG-SL}) \times \mathcal{P}(N_{SB-SL}|n_{SB-SL}) \\
\times\; & \mathcal{P}(N_{SIG-LDP}|n_{SIG-LDP}) \times \mathcal{P}(N_{SB-LDP}|n_{SB-LDP}) \\
\times\; & \mathcal{P}(N_{LSB-0b}|n_{LSB-0b}) \times \mathcal{P}(N_{LSB-0b-LDP}|n_{LSB-0b-LDP}) \\
\times\; & \mathcal{P}(N_{SIG-ee}|n_{SIG-ee}) \times \mathcal{P}(N_{SB-ee}|n_{SB-ee}) \\
\times\; & \mathcal{P}(N_{SIG-\mu\mu}|n_{SIG-\mu\mu}) \times \mathcal{P}(N_{SB-\mu\mu}|n_{SB-\mu\mu}) \\
\times\; & \mathcal{G}(g_{\epsilon}) \\
\times\; & \mathcal{G}(g_{MC}) \\
\times\; & \mathcal{G}(g_{ttwj}) \times \mathcal{G}(g_{QCD-SIG}) \times \mathcal{G}(g_{QCD-SB}) \\
\times\; & \mathcal{G}(g_{ee}) \times \mathcal{G}(g_{\mu\mu}) \\
\times\; & \mathcal{G}(g_{Pee}) \times \mathcal{G}(g_{P\mu\mu}) \\
\times\; & \mathcal{G}(g_{Aee}) \times \mathcal{G}(g_{A\mu\mu}) \\
\times\; & \mathcal{G}(g_{\epsilon_{ee}}) \times \mathcal{G}(g_{\epsilon_{\mu\mu}})
\end{aligned}
$$

The nuisance parameters are listed in Table 24.

A.2 Systematic uncertainties

Statistical and systematic uncertainties on external parameters are handled with log-normal PDFs in the likelihood, one for each parameter. The SF_i parameters in Table 24 take into account corrections and systematic uncertainties on the analysis model assumptions. These are

- $\boldsymbol{SF_{\epsilon,i}}$: This scale factor takes into account all uncertainties on the signal reconstruction efficiency. This scale factor is applied to efficiency-corrected signal yield expectations, so the log-normal mean parameter is 1. The six bins with signal contributions have differing uncertainties (log-normal width parameters), but the $SF_{\epsilon,i}$ are assumed to be 100% correlated. The PDF for $SF_{\epsilon,i}$ is a log-normal distribution constructed from the common underlying primary variable g_ϵ using Eq. (5). The efficiency uncertainty is evaluated at every point in the scan of the 2D signal parameter space.
- $\boldsymbol{SF_{ttwj}}$: This covers the assumption that the SIG/SB ratio for the $ttwj$ component is the same in the signal and single lepton selections. We use the ratio $N_{predicted}/N_{true}$ in the $ttwj$ MC closure test from Table 10 to determine the uncertainty. The width parameter is taken as the deviation of $N_{predicted}/N_{true}$ from one added in quadrature with its statistical uncertainty.
- $\boldsymbol{SF_{QCD-SIG}}$ and $\boldsymbol{SF_{QCD-SB}}$: This covers the assumption that the $\Delta\phi_{\mathrm{N}}^{\mathrm{min}}$ pass / fail ratio is the same in the $E_{\mathrm{T}}^{\mathrm{miss}}$ LSB, SB, and SIG regions. We use the QCD MC closure test results in Table 4 for the uncertainty.
- $\boldsymbol{SF_{ee}}$ and $\boldsymbol{SF_{\mu\mu}}$: This covers the systematic uncertainty derived from the closure test in the $\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$ and $\mathrm{Z} \to \mu^+\mu^-$ determinations of the $Z \to \nu\nu$ background. The uncertainty is taken from Table 7.

A.3 Alternative Likelihood Construction for the Tight Selection

The likelihood for the tight selections and for the loose selection with ≥ 2 b-jets is slightly different from the one described above. The observed number of $\mathrm{Z} \to \ell^+\ell^-$ events for the tighter selections is very low or zero. Instead of applying the tighter selections directly to the $\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$ and $\mathrm{Z} \to \mu^+\mu^-$ samples, we apply loosened selections and then use an MC scale factor (K) to predict the number of $\mathrm{Z} \to \nu\bar{\nu}$ events in the tighter selections, as described in Sect. 6.2. In this alternative model, the SIG and SB $\mathrm{Z} \to \ell^+\ell^-$ yields are combined and separate scale factors (K_{SIG} and K_{SB}) are applied to this combined looser selection SIG+SB yield to compute the predicted $\mathrm{Z} \to \nu\bar{\nu}$ yield for the SIG and SB bins.