CMS Draft Analysis Note

2019/08/12
Archive Hash: 61fbd1e-D
Archive Date: 2019/08/12

Search for supersymmetry using boosted Z Bosons and missing transverse momentum in proton-proton collisions at 13 TeV

R. Patel[1] and U. Sarkar[2]

[1] University of Colorado Boulder (US)
[2] Tata Institute of Fundamental Research, Mumbai

Abstract

A search for supersymmetry is presented with bosons and large missing transverse momentum in proton-proton collisions collected by the CMS experiment at the LHC at $\sqrt{s} = 13$ TeV. We target Z bosons produced with large momentum, where the hadronization byproducts of a pair of quarks can be reconstructed in a single jet. The analysis uses the jet mass to identify jets reconstructed from overlapping $q\bar{q}$. The background from SM processes is significanctly reduced with requirements on the jet mass and large missing transverse energy. The data sample corresponds to an integrated luminosity of 137 fb^{-1} collected by CMS in Run 2.

Contents

1 Introduction

One of the primary motivations to the CERN LHC was to determine the source of electroweak symmetry breaking and search for physics beyond the standard model (SM). A major milestone was achieved with the discovery of the Higgs boson by the ATLAS and CMS collaborations with a mass at the electroweak scale. Supersymmetry (SUSY) could be a potential solution to explain the low mass Higgs boson without fine tuning of the SM. Supersymmetry is a widely sutdied extension of the SM that posits for each SM particle a new particle, called a superpartner, with a spin that differs from that of its SM counterpart by a half unit. The superpartners of quarks and gluons are squarks ($\tilde{q}$) and gluinos ($\tilde{g}$) respectively. The superpartners of electroweak gauge bosons are neutralinos ($\tilde{\chi}^0$) and charginos ($\tilde{\chi}^\pm$). In this note, we will focus on a simplified model scenario (SMS) where gluinos decay to quarks and the next-to-lightest SUSY particle (NLSP) $\tilde{\chi}^0_2$. If the mass difference is small between the gluino and $\tilde{\chi}^0_2$, the decay products of the SUSY particles can have large lorentz boost.

This note presents a search for Supersymmetry (SUSY) in events with boosted electroweak (EW) bosons that decay to quarks, in particular targetting hadronic decays of the Z-boson. Substructure jet mass is used to identify wide cone jets (R=0.8) that contain the decay products of the Z-boson. The SUSY scenarios assume R-parity conservation, so the event topology also has large missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) from the lightest SUSY particle (LSP). The final analysis categories all require at least two tagged jets coming from the Z-boson decays and increasing values of $E_{\mathrm{T}}^{\mathrm{miss}}$. The narrow mass peak of the Z-decay allows it to be resolved over the non-resonant background from SM processes. Also the proxmitiy of the Z-mass and the W-mass allow this search to be generalized to SUSY models with vector bosons in the final state.

2 Event samples

2.1 Standard model MC samples

Several CMSSW releases were used to process the SM Monte Carlo (MC) samples. The 2016 samples were reconstructed mainly in 9_4_X (RunIISummer16MiniAODv3). The 2017 MC samples were reconstructed in a 9_4_X (RunIIFall17MiniAODv2) release while the 2018 were reconstructed in a 10_2_X release. The SM samples are listed in Tables 1-**??**. The cross sections listed correspond to next-to-next-to-leading-order (NNLO) calculations unless otherwise noted.

2.2 Signal models

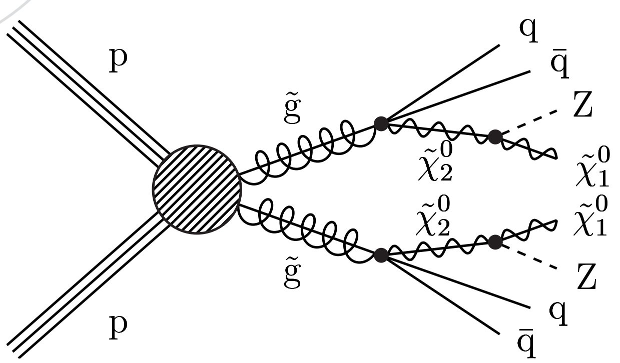


Figure 1: Signal diagrams for the boosted Z-boson search via gluino strong production. We consider 100% branching fraction to the Z boson(left).

Table 1: SM $t\bar{t}$ MC samples used in the analysis. The cross sections are calculated to NNLO.

Year	Dataset	σ (pb)	$\int \mathcal{L}\,dt$ (fb^{-1})
2016	TTJets_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	831.76	12.19
	TTJets_SingleLeptFromT_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	182.72	337.24
	TTJets_SingleLeptFromTbar_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	182.72	330.25
	TTJets_DiLept_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	88.34	349.06
	TTJets_SingleLeptFromT_genMET-150_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	9.684	1792.22
	TTJets_SingleLeptFromTbar_genMET-150_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	9.658	1760.63
	TTJets_DiLept_genMET-150_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	5.919	1647.82
	TTJets_HT-600to800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	2.685	5343.28
	TTJets_HT-800to1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.096	9607.90
	TTJets_HT-1200to2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.194	15097.94
	TTJets_HT-2500toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.002	646450.58
2017/ 2018	TTJets_TuneCP5_13TeV-madgraphMLM-pythia8	831.76	12.19
	TTJets_SingleLeptFromT_TuneCP5_13TeV-madgraphMLM-pythia8	182.72	337.26
	TTJets_SingleLeptFromTbar_TuneCP5_13TeV-madgraphMLM-pythia8	182.72	309.66
	TTJets_DiLept_TuneCP5_13TeV-madgraphMLM-pythia8	88.34	344.63
	TTJets_SingleLeptFromT_genMET-150_TuneCP5_13TeV-madgraphMLM-pythia8	9.684	2904.10
	TTJets_SingleLeptFromTbar_genMET-150_TuneCP5_13TeV-madgraphMLM-pythia8	9.658	2908.44
	TTJets_DiLept_genMET-150_TuneCP5_13TeV-madgraphMLM-pythia8	5.919	2697.85
	TTJets_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8	2.685	5221.29
	TTJets_HT-800to1200_TuneCP5_13TeV-madgraphMLM-pythia8	1.096	9282.72
	TTJets_HT-1200to2500_TuneCP5_13TeV-madgraphMLM-pythia8	0.194	14819.34
	TTJets_HT-2500toInf_TuneCP5_13TeV-madgraphMLM-pythia8	0.002	641714.58

Table 2: SM QCD MC samples used in the analysis. All cross sections are calculated to LO.

Year	Dataset	σ (pb)	$\int \mathcal{L}\, dt$ (fb^{-1})
2016	QCD_HT200to300_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1717000	0.03
	QCD_HT300to500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	351300	0.15
	QCD_HT500to700_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	31630	1.98
	QCD_HT700to1000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6802	2.30
	QCD_HT1000to1500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1206	12.61
	QCD_HT1500to2000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	120.4	98.33
	QCD_HT2000toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	25.24	238.49
2017/ 2018	QCD_HT200to300_TuneCP5_13TeV-madgraph-pythia8	1717000	0.03
	QCD_HT300to500_TuneCP5_13TeV-madgraph-pythia8	351300	0.17
	QCD_HT500to700_TuneCP5_13TeV-madgraph-pythia8	31630	1.77
	QCD_HT700to1000_TuneCP5_13TeV-madgraph-pythia8	6802	6.96
	QCD_HT1000to1500_TuneCP5_13TeV-madgraph-pythia8	1206	13.58
	QCD_HT1500to2000_TuneCP5_13TeV-madgraph-pythia8	120.4	94.55
	QCD_HT2000toInf_TuneCP5_13TeV-madgraph-pythia8	25.24	226.31

Table 3: SM $Z \to \nu\nu$+jets MC samples used in the analysis. The cross sections are calculated to NNLO.

Year	Dataset	σ (pb)	$\int \mathcal{L}\, dt$ (fb^{-1})
2016	ZJetsToNuNu_HT-100To200_13TeV-madgraph	344.83	70.39
	ZJetsToNuNu_HT-200To400_13TeV-madgraph	95.53	259.19
	ZJetsToNuNu_HT-400To600_13TeV-madgraph	13.20	747.31
	ZJetsToNuNu_HT-600To800_13TeV-madgraph	3.148	1831.10
	ZJetsToNuNu_HT-800To1200_13TeV-madgraph	1.451	1495.71
	ZJetsToNuNu_HT-1200To2500_13TeV-madgraph	0.355	1447.84
	ZJetsToNuNu_HT-2500ToInf_13TeV-madgraph	0.009	47414.35
2017/ 2018	ZJetsToNuNu_HT-100To200_13TeV-madgraph	344.83	65.74
	ZJetsToNuNu_HT-200To400_13TeV-madgraph	95.53	225.69
	ZJetsToNuNu_HT-400To600_13TeV-madgraph	13.20	686.16
	ZJetsToNuNu_HT-600To800_13TeV-madgraph	3.148	1789.28
	ZJetsToNuNu_HT-800To1200_13TeV-madgraph	1.451	1396.10
	ZJetsToNuNu_HT-1200To2500_13TeV-madgraph	0.355	929.88
	ZJetsToNuNu_HT-2500ToInf_13TeV-madgraph	0.009	722.32

Table 4: SM $W \to \ell\nu$+jets MC samples used in the analysis. The cross sections are calculated to NNLO.

Year	Dataset	σ (pb)	$\int \mathcal{L}\, dt$ (fb^{-1})
2016	WJetsToLNu_HT-100To200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1627.45	6.11
	WJetsToLNu_HT-200To400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	435.24	89.57
	WJetsToLNu_HT-400To600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	59.18	131.12
	WJetsToLNu_HT-600To800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	14.58	1281.72
	WJetsToLNu_HT-800To1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6.660	1175.76
	WJetsToLNu_HT-1200To2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.608	4273.91
	WJetsToLNu_HT-2500ToInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.039	67792.88
2017/ 2018	WJetsToLNu_HT-100To200_TuneCP5_13TeV-madgraphMLM-pythia8	1627.45	21.96
	WJetsToLNu_HT-200To400_TuneCP5_13TeV-madgraphMLM-pythia8	435.24	48.56
	WJetsToLNu_HT-400To600_TuneCP5_13TeV-madgraphMLM-pythia8	59.18	239.73
	WJetsToLNu_HT-600To800_TuneCP5_13TeV-madgraphMLM-pythia8	14.58	1471.62
	WJetsToLNu_HT-800To1200_TuneCP5_13TeV-madgraphMLM-pythia8	6.660	3020.20
	WJetsToLNu_HT-1200To2500_TuneCP5_13TeV-madgraphMLM-pythia8	1.608	12269.07
	WJetsToLNu_HT-2500ToInf_TuneCP5_13TeV-madgraphMLM-pythia8	0.039	508831.27

Table 5: SM single-top MC samples used in the analysis. The cross sections are calculated to NLO.

Year	Dataset	σ (pb)	$\int \mathcal{L}\,dt$ (fb^{-1})
2016	ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	3.340	116.20
	ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	136.02	493.35
	ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	80.95	479.44
	ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	19.47	167.27
	ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	19.47	167.29
2017/ 2018	ST_s-channel_4f_leptonDecays_TuneCP5_PSweights_13TeV-amcatnlo-pythia8	3.340	1154.17
	ST_t-channel_top_4f_inclusiveDecays_TuneCP5_13TeV-powhegV2-madspin-pythia8	136.02	43.13
	ST_t-channel_antitop_4f_inclusiveDecays_TuneCP5_13TeV-powhegV2-madspin-pythia8	80.95	48.67
	ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_PSweights_13TeV-powheg-pythia8	19.47	272.59
	ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_PSweights_13TeV-powheg-pythia8	19.47	237.87

Table 6: SM diboson and other rare process MC samples used in the analysis. The cross sections are calculated to NNLO.

Year	Dataset	σ (pb)	$\int \mathcal{L}\, dt$ (fb^{-1})
	TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.253	5023.00
	TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.530	310.65
	TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.204	4033.43
	TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.403	551.94
	TTGJets_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	3.697	418.06
	WWTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	50.00	40.69
	WWTo2L2Nu_13TeV-powheg	12.18	164.15
2016	WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	10.71	764.54
	WZTo1L3Nu_13TeV_amcatnloFXFX_madspin_pythia8	3.058	170.33
	ZZTo2Q2Nu_13TeV_amcatnloFXFX_madspin_pythia8	4.040	2845.43
	ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.220	1899.70
	TTTT_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.009	46824.95
	WWZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.165	1188.32
	WZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.056	3408.53
	ZZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.014	13044.95
	TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8	0.253	6665.62
	TTZToQQ_TuneCP5_13TeV-amcatnlo-pythia8	0.530	319.53
	TTWJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	0.204	7130.58
	TTWJetsToQQ_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	0.403	596.93
	TTGJets_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	3.697	571.05
	WWTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	50.00	39.76
2017/ 2018	WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	10.71	629.88
	WZTo1L3Nu_13TeV_amcatnloFXFX_madspin_pythia8	3.058	483.67
	ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.220	3521.70
	TTTT_TuneCP5_13TeV-amcatnlo-pythia8	0.009	34901.16
	WZZ_TuneCP5_13TeV-amcatnlo-pythia8	0.056	3468.14
	ZZZ_TuneCP5_13TeV-amcatnlo-pythia8	0.014	13043.84

Figure 1 shows the event diagrams for the signal considered in this analysis. The mass splitting between $\tilde{g}$ and $\tilde{\chi}_2^0$ is fixed at 50 GeV, thus each of the $\tilde{g}$ produces a low p_{T} quark. The mass of the $\tilde{\chi}_1^0$ is fixed to 1 GeV so that the Z boson p_{T} is proportional to $m_{\tilde{\chi}_2^0}/2 \sim m_{\tilde{g}}/2$. The signal regions for this analysis are the events strictly with 2 Z bosons with 100% branching fraction in the final state, where $Z \rightarrow q\bar{q}$. For most gluino masses, the quarks, $Z \rightarrow q\bar{q}$ are expected to be contained in a large-radius jet, $\Delta R = 0.8$ instead of showing up as two resolved jets due to the boosted topology of the model. Events are generated with the Full Simulation using the reconstruction in CMSSW version `9_4_X`.

2.3 Data samples

We analyze the 13 TeV dataset collected during 2016, 2017, and 2018 with the CMS detector. For 2016 and 2017 we used the `17Jul2018 re-reco` and `31Mar2018 re-reco` versions, respectively, while 2018 we used we used a combination of `17Sep2018 re-reco` and `prompt` (period D only) datasets. Table 7 lists the integrated luminosities for the primary datasets used, split up by data-taking period, for each of the years. The data set is measured to correspond to 137.2 fb^{-1} using the BRIL Work Suite [1].

3 Triggers

In Section 4, the primary offline kinematic selection for the search region is $H_T > 500$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ GeV,along with vetoed the events with leptons. This section will describe the trigger efficiency for the signal and control regions, and check if these offline regions are well above the trigger turn-on [2]. The details on Trigger efficiency measurements and the list of reference triggers are in [2].

3.1 Signal region

Events in the SR, as well as events collected for the single-electron and single-muon validation regions were collected using a set of $p_{\mathrm{T}}^{\mathrm{miss}}$-$H_{\mathrm{T}}^{\mathrm{miss}}$ cross-triggers, denoted by the HLT paths

- `HLT_PFMETX_PFMHTX_IDTight_v* (X=90,100,110,120,130,140)` and
- `HLT_PFMETNoMuX_PFMHTNoMuX_IDTight_v* (X=90,100,110,120,130,140).`

Here, `X` indicates the threshold applied to the online $p_{\mathrm{T}}^{\mathrm{miss}}$ and $H_{\mathrm{T}}^{\mathrm{miss}}$, as calculated by the particle flow (PF) algorithm; the asterisks indicate that more than one version of the same trigger may have been used. During periods of higher instantaneous luminosity, trigger paths with lower thresholds became prescaled to reduce the event rate; in such cases, the search relies on the higher-threshold triggers, which remained un-prescaled throughout all data-taking periods. To compensate for losses in efficiency associated with the higher trigger thresholds, a set of back-up triggers was used when the low-threshold $p_{\mathrm{T}}^{\mathrm{miss}}$-$H_{\mathrm{T}}^{\mathrm{miss}}$ triggers became prescaled:

- `HLT_PFMETX_PFMHTX_IDTight_PFHT60_v* (X=100,110,120,130,140),`
- `HLT_PFMETNoMuX_PFMHTNoMuX_IDTight_PFHT60_v* (X=100,110,120,130,140),`
- `HLT_PFMET120_PFMHT120_IDTight_HFCleaned_v*,`
- `HLT_PFMET120_PFMHT120_IDTight_PFHT60_HFCleaned_v*`, and
- `HLT_PFMETNoMu120_PFMHTNoMu120_IDTight_HFCleaned_v*.`

The logical OR of all of the above trigger paths was taken as the online criterion for selecting events throughout the three years of data-taking.

Table 7: Datasets collected from three years of data-taking. All $\int \mathcal{L}\,\mathrm{d}t$ are listed in fb^{-1} and are calculated using the BRIL Work Suite [1].

Year	Primary Dataset	A	B	C	D	E	F	G	H	Total
2016	HTMHT	–	5746.365	2572.903	4242.289	4024.754	3104.509	7574.961	8650.622	35916.403
	JetHT	–	5750.126	2572.903	4242.292	4024.754	3104.509	7575.824	8650.628	35921.036
	MET	–	5746.370	2572.903	4242.287	3924.254	3104.508	7575.824	8649.019	35815.165
	SingleElectron	–	5746.183	2572.813	4242.201	4025.019	3104.288	7575.483	8650.155	35916.142
	SingleMuon	–	5746.010	2572.903	4242.292	4025.228	3104.509	7575.579	8650.628	35917.149
	SinglePhoton	–	5746.364	2572.903	4242.286	4025.226	3104.509	7575.824	8650.626	35917.738
2017	HTMHT	–	4793.970	9631.262	4247.704	9313.989	13534.500	–	–	41521.425
	JetHT	–	4793.980	9631.323	4247.706	9313.989	13534.525	–	–	41521.523
	MET	–	4793.367	9632.850	4247.706	9313.990	13498.415	–	–	41486.328
	SingleElectron	–	4793.922	9631.008	4247.695	9313.682	13539.222	–	–	41525.529
	SingleMuon	–	4793.980	9631.323	4247.706	9313.682	13538.559	–	–	41525.250
	SinglePhoton	–	4793.980	9631.319	4247.705	9313.682	13539.211	–	–	41525.897
2018	EGamma	13926.173	7091.450	6932.632	31249.311	–	–	–	–	59199.566

3.2 Single-photon and Di-lepton regions

Events in the single-photon CR were collected using a single-photon trigger,

- `HLT_Photon175_v*`

in 2016 and

- `HLT_Photon200_v*`

in 2017 and 2018.

The efficiency of these two triggers has been measured in bins of the offline photon p_{T}, in two independent samples, one collected by the set of H_{T} triggers defined above, and again in a sample collected by a single-muon trigger. In addition to passing the trigger requirements, events are required to have an offline H_{T} greater than 300 GeV and to have at least one loose WP photon. This selection is consistent though slightly looser with the background for the background in the search region, and provides a suitable dataset for a validation region. In addition dilepton events on-Z are also used to validate the background in the search region.

Events in the di-electron CR were collected using a set of single-electron triggers,

- `HLT_EleX_CaloIdVT_GsfTrkIdT_v*(X=105, 115, 135, 145);`
- `HLT_Ele25_eta2p1_WPTight_Gsf_v*;`
- `HLT_EleX_eta2p1_WPLoose_Gsf_v*(X=20, 27);`
- `HLT_Ele15_IsoVVVL_PFHTX_v*(X=350, 400, 450, 600);`
- `HLT_EleX_WPTight_Gsf_v*(X=27, 35);` and
- `HLT_EleX_WPLoose_Gsf_v*(X=20,45).`

Events in the di-muon CR were collected using a set of single-muon triggers that closely resembles the single-muon set. The muon triggers are:

- `HLT_IsoMuX_v* (X=20,22,24,27);`
- `HLT_IsoMuX_eta2p1_v* (X=22,24) ;`
- `HLT_IsoTkMuX_v* (X=22,24) ;`
- `HLT_Mu15_IsoVVVL_PFHTX_v* (X=350,400,450,600) ;`
- `HLT_Mu50_IsoVVVL_PFHTX_v* (400,450) ;` and
- `HLT_MuX_v* (X=50,55) .`

4 Event selection

The final state of the signal model is boosted and fully hadronic, so the search regions for this analysis have a set of baseline selection for targetting a broad range of SUSY hadronic signals. On top of this selection we apply a selection to target events with boosted jets. The search baseline for this analysis require large $E_{\mathrm{T}}^{\mathrm{miss}}$, large H_{T}, and no leptons. This selection makes use of the same kinematic variables as more inclusive SUSY analyses [2].

Jets used in this analysis are reconstructed from charged-hadron subtracted particle-flow (PF) candidates using the anti-k_{T} algorithm [3] with size parameters 0.8 (AK8) and 0.4 (AK4).The PF algorithm is used to individually identify and reconstruct all particles produced in the collision (PF candidates); namely charged hadrons, photons, neutral hadrons, muons, and electrons [4]. This selection is summarized in Section 4.1, and includes a description of filters and corrections designed to improve the modeling of the MC and reduce the number of events with mismeasurement due to noise or known detector performance issues. The boosted selection will be described in Section 4.2, and further tightens the selection to reject the bulk of the SM background.

4.1 Hadronic Baseline

The following requirements define the baseline selection:

- $H_{\mathrm{T}} > 400\,\mathrm{GeV}$, where $H_{\mathrm{T}} = \sum_{\mathrm{AK4jets}} |\vec{p}_{\mathrm{T}}|$
  - AK4 jets are required to pass the loose jet ID requirements:
    For jets with $|\eta| < 2.4$:
    - "loose" working point for 2016:
      - neutral hadron fraction < 0.99,
      - neutral EM fraction < 0.99,
      - number of constituents > 1,
      - charged hadron fraction > 0,
      - charged multiplicity > 0,
      - charged EM fraction < 0.99
    - "tight" (only supported) working point for 2017 (and 2018):
      - neutral hadron fraction < 0.90,
      - neutral EM fraction < 0.90,
      - number of constituents > 1,
      - charged hadron fraction > 0,
      - charged multiplicity > 0,
  - $E_{\mathrm{T}}^{\mathrm{miss}} > 300\,\mathrm{GeV}$ where $E_{\mathrm{T}}^{\mathrm{miss}} = |\sum_{\mathrm{PFcandidates}} \vec{p}_T|$.
  - $\mathrm{H}_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$ where $\mathrm{H}_{\mathrm{T}}^{\mathrm{miss}} = \left|\sum_{\mathrm{AK4jets}} \vec{\mathrm{p}}_{\mathrm{T}}\right|$.
  - Angular cut: The majority of QCD multijet events in our high-$E_{\mathrm{T}}^{\mathrm{miss}}$ search region have jets with under-measured momenta and give rise to momentum imbalance. A signature of such an event is a jet closely aligned in direction with the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector. To suppress this background we require the two leading AK4 jets to be seperated by more than 0.5 radians from the $\mathrm{H}_{\mathrm{T}}^{\mathrm{miss}}$vector in the azimuthal coordinate. If present, the third and fourth highest-p_{T} AK4 jets must be seperated by at least 0.3 radians.
  - Muon veto:
    Muon candidates are selected using the POG-recommended "Medium Muon" selection [5] with the additional requirements:

    $$\begin{aligned} d_{xy}(\mu, \mathrm{PV}) &< 0.2\ \mathrm{cm} \\ d_{z}(\mu, \mathrm{PV}) &< 0.5\ \mathrm{cm} \end{aligned} \tag{1}$$

    Muon candidates are required to have $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and $|\eta| < 2.4$. To distinguish between prompt muons and muons from b-hadron decays, muons are required to satisfy an isolation requirement, $I_{\mathrm{mini}} < 0.2$, where I_{mini} is the mini-isolation variable described in Ref. [6]. Any event with a muon satisfying all of the above criteria is vetoed.
  - Electron veto:
    Electron candidates are selected using the POG-recommended "Cut Based VETO" selection [7]. Electron candidates are required to have $p_{\mathrm{T}} >$

10 GeV and $|\eta| < 2.5$. Electron candidates are required to satisfy an isolation requirement of $I_{\text{mini}} < 0.1$. Any event with an electron satisfying all of the above criteria is vetoed.

- Isolated track vetoes:

  Following the event selection described above, including the muon and electron event vetoes, there is still some background in the search regions from $\mathrm{t\bar{t}}$, single-top, and W+jets events with one $\mathrm{W} \to \ell\nu$ decay. In about half these background events, the W boson decays to a τ lepton and the τ lepton decays hadronically, while in the other half, an electron or muon is not identified or does not satisfy the criteria for an isolated electron or muon candidate given above. To suppress these backgrounds, we reject events with one or more isolated charged track.

  The requirements for the definition of an isolated track differ slightly depending on whether the track is identified as leptonic or hadronic by the PF algorithm. For leptonic tracks, we require:

  - $p_{\mathrm{T}} > 5\,\text{GeV}$,
  - $I_{\text{tk}} < 0.2$,

  where I_{tk} is the scalar p_{T} sum of other charged tracks within $\Delta R \equiv \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} <$ 0.3 of the primary track, divided by the p_{T} value of the primary track. For hadronic tracks, we apply slightly tighter requirements:

  - $p_{\mathrm{T}} > 10\,\text{GeV}$,
  - $I_{\text{tk}} < 0.1$.

  Isolated tracks are considered only if they satisfy

  $$m_T(\text{tk}, E_{\mathrm{T}}^{\text{miss}}) = \sqrt{2 p_T^{\text{tk}} E_{\mathrm{T}}^{\text{miss}} (1 - \cos\Delta\phi)} < 100\ \text{GeV}, \quad (2)$$

  where p_T^{tk} is the transverse momentum of the track and $\Delta\phi$ is the azimuthal separation between the track and $\vec{p}_{\mathrm{T}}^{\text{miss}}$.

  To reduce the influence of tracks from extraneous pp interactions (pileup), isolated tracks are considered only if their nearest distance of approach along the beam axis to a reconstructed vertex is smaller for the primary event vertex than for any other vertex.

- Event cleaning:

  We reject events with a jet that satisfies $p_{\mathrm{T}} > 30\,\text{GeV}$ and $|\eta| < 5$ if the jet fails the loose jet ID criteria given above. We apply event filters designed by various POGs to reject events with spurious $E_{\mathrm{T}}^{\text{miss}}$ signals. The current list includes:

  - globalTightHalo2016Filter
  - HBHENoiseFilter
  - HBHEIsoNoiseFilter
  - eeBadScFilter
  - EcalDeadCellTriggerPrimitiveFilter
  - BadChargedCandidateFilter
  - BadPFMuonFilter
  - ecalBadCalibReduced∗ (being finalized, will only apply to 2017 and 2018 data)
  - Good vertex filter (requiring at least one reconstructed vertex satisfying !isFake && $N_{\text{dof}} > 4$ && $|z| < 24$ && $\rho < 2$)

  - To protect against particle flow failures, events are rejected if PFMET/CaloMET > 5.

  We also apply the following "muon jet filter" to reject events with misreconstructed muons as described in [8]:

  - Veto events if any jet in the event has $p_{\mathrm{T}} > 200\,\mathrm{GeV}$, muon energy fraction > 0.5, and $\Delta\phi(\mathrm{jet}, E_{\mathrm{T}}^{\mathrm{miss}}) > \pi - 0.4$.

  Another case of anomalous jets affecting the QCD control regions is handled with the following filter:

  - Veto events if the leading jet has neutral EM energy fraction < 0.03 and $\Delta\phi(j_1,\ \mathrm{H}_{\mathrm{T}}^{\mathrm{miss}}) >$ ß $- 0.4$.

  To protect against particle flow failures, events are rejected if PFMET/CaloMET > 5.

  Events with anomalously energetic jets in the HF are observed in the data between $|\eta|$ values of 3.0 and 3.1, which lead to excess event counts in the QCD-enriched control regions. To reject such events, a cut is placed in the plane of $\Delta\phi(j_1,\ \mathrm{H}_{\mathrm{T}}^{\mathrm{miss}})$ and the quantity $H_{\mathrm{T}}^{5}/H_{\mathrm{T}}$, the ratio of the H_{T} computed using all jets within $|\eta| < 5$ to the standard H_{T} computed with jets within $|\eta| < 2.4$, such that passing events must satisfy

$$H_{\mathrm{T}}^{5}/H_{\mathrm{T}} < 1.2 \text{ or } \Delta\phi(j_1,\ \mathrm{H}_{\mathrm{T}}^{\mathrm{miss}}) \geq 5.3\,\mathrm{H}_{\mathrm{T}}^{5}/\mathrm{H}_{\mathrm{T}} - 4.78 \tag{3}$$

  The impact of this filter on the signal efficiency was determined to be negligible for the considered models, and the filter reduces the fake-$E_{\mathrm{T}}^{\mathrm{miss}}$ background by up to 70% in some search bins. Accounting for the correlation between $H_{\mathrm{T}}^{5}/H_{\mathrm{T}}$ and $\Delta\phi(j_1,\ \mathrm{H}_{\mathrm{T}}^{\mathrm{miss}})$ further rejects miscalibrated jets.

  In lieu of the updated ecalBadCalibReduced filter referenced above, we apply a filter designed to reject events with $\mathrm{H}_{\mathrm{T}}^{\mathrm{miss}}$artificially induced by noise in the ECAL:

  - Veto events if either of the two leading jets with $|\eta| > 2.4$ and $|\eta| < 5.0$ have $p_{\mathrm{T}} > 250\,\mathrm{GeV}$ and $\Delta\phi(\mathrm{jet}, \mathrm{H}_{\mathrm{T}}^{\mathrm{miss}}) > 2.6$ or < 0.1.

  This filter is only applied to 2017 data.

- Corrections: The "ECAL L1 pre-firing issue" causes events with high p_{T} forward objects to suffer from a reduced trigger efficiency. We correct the MC based on the pre-firing inefficiency in [9] to account for this effect in the data. Additional information about these effects is found in Appendix **??**.

  In the 2018 data, starting with Run 319077 (just before the 2018C data-taking period), one sector of the minus side of the HE was disabled unexpectedly. This is often described as the HEM problem. The unmodified particle flow algorithm may generate additional jets and/or electrons in the disabled sector, because there may be energy measured in ECAL, but cannot be any corresponding energy measured in HCAL. To improve the agreement between data and MC, we veto both MC and data events that have any activity in that region. A wider veto region is used for jets (wider by half the jet radius), with an additional cut to minimize the reduction in signal efficiency. This veto is defined as:

  - Veto events with any electron with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$, $-3.0 < \eta < -1.4$, and $-1.57 < \phi < -0.87$.
  - Veto events with any jet with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$, $\Delta\phi(\mathrm{jet},\ \mathrm{H}_{\mathrm{T}}^{\mathrm{miss}}) < 0.5$, $-3.2 < \eta < -1.2$, and $-1.77 < \phi < -0.67$.

The significant background for these high $E_{\mathrm{T}}^{\mathrm{miss}}$ all hadronic events are from QCD multijet with mis-measured $E_{\mathrm{T}}^{\mathrm{miss}}$ and decays of weak vector bosons which produce neutrinos. Most dominant backgrounds that can contribute to the search region phase space would come from W jets decaying semileptonically where the lepton is missed, or from Z jets where Z decays into invisible ν. Also, $t\bar{t}$ events can give a similar boosted topology in this phase space. All other rare background like di-boson or single-top will have different topologies away from the Z boson mass. Figure 2 shows the expected distributions of H_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$ for SM backgrounds after baseline selection.

4.2 Boosted Object Baseline

This section describes the kinematic selection applied to ensure a boosted topology with AK8 jets. These AK8 jets are reclustered from their original jet constituents, and the clustering sequence is modified to remove soft and wide-angle particles or groups of particles. The reclustering method that has been used is softdrop. This "softdrop jet" is used to compute the mass after removing the soft radiation to provide a narrower Z mass window [10]. AK4 jets are used to compute the H_{T}, $\mathrm{H}_{\mathrm{T}}^{\mathrm{miss}}$, and $\Delta\phi$ variables as described in the previous section, on top of this selection we apply the following:

- To ensure events with a boosted topology, events are selected based on high p_{T} AK8 jets with the following criteria:
  - Require the event to have at least two AK8 jets with leading jet $p_{\mathrm{T}} > 200\,\mathrm{GeV}$ and subleading jet $p_{\mathrm{T}} > 200\,\mathrm{GeV}$. Both the boosted object p_{T} s are chosen for a fat jet radius of 0.8 for a Z boson mass near to 90 GeV. This also makes the selection more inclusive to other models.
  - The softdrop mass [10] of the two highest p_{T} AK8 jets are required to be between 40 and 200 GeV
- ΔR cut: This cut is applied mainly to reject backgrounds coming from $t\bar{t}$. We find a b-tagged (DeepCSV, medium working point) AK4 jet near to the sub-lead AK8 jet and veto the events if ΔR between AK4 and AK8 jet is less than 0.8. From, signal topology, we expect this two jets to be widely separated as they come from different vertex, but for $t\bar{t}$, they come from the same vertex in case of a boosted scenario of the background event.

The baseline selections have the effect of selecting events with one or more boosted objects. The p_{T} requirement ensures that bosons with mass $\lesssim 90\,\mathrm{GeV}$ around Z boson mass will have both decay products captured within a single AK8 jet and the slection of at least two AK8 jets ensures that most of the final state events are fully hadronic. The softdrop mass cut ensures that high p_{T} AK8 jets resulting from a single parton are largely rejected. The pruning algorithm of softdrop improves the background rejection power of the mass cut by reducing the effect of pileup and underlying event, and by removing the soft, wide angle radiation that provide the primary mechanism for generating jet masses in QCD jets [10].

Figure 3 shows the MC distributions of jet p_{T}, after baseline selection except p_{T} requirements and the softdrop-jet mass distributions with and without the ΔR cut between AK4 b-tagged jet and the sublead AK8 jet. All MC samples are scaled to the data luminosity of 137 fb^{-1}. These cuts make this analysis particularly unique with respect to other all-hadronic analyses. By targeting boosted jet topologies, this analysis can significantly reduce SM background rates in a way that compliments other more inclusive all-hadronic searches. In particular, since most SM processes do not produce hadronically decaying bosons, the jet masses will typically be below our baseline selection of 50 GeV, even if they have large p_T. A cut flow for each background process and two representative signals is listed in Table 8.

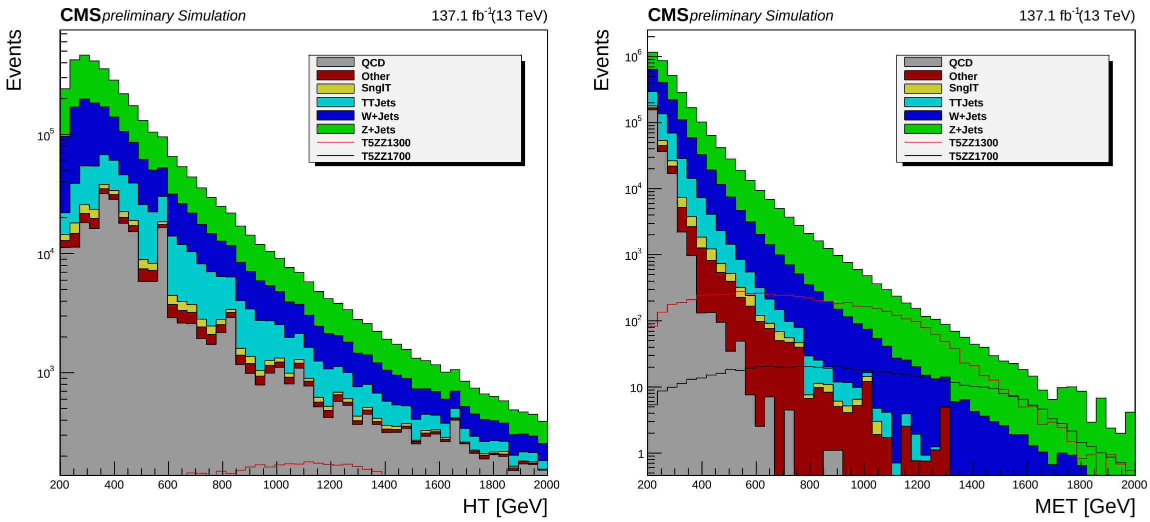


Figure 2: H_T (left) and E_T^{miss} (right) distributions after hadronic baseline selection.

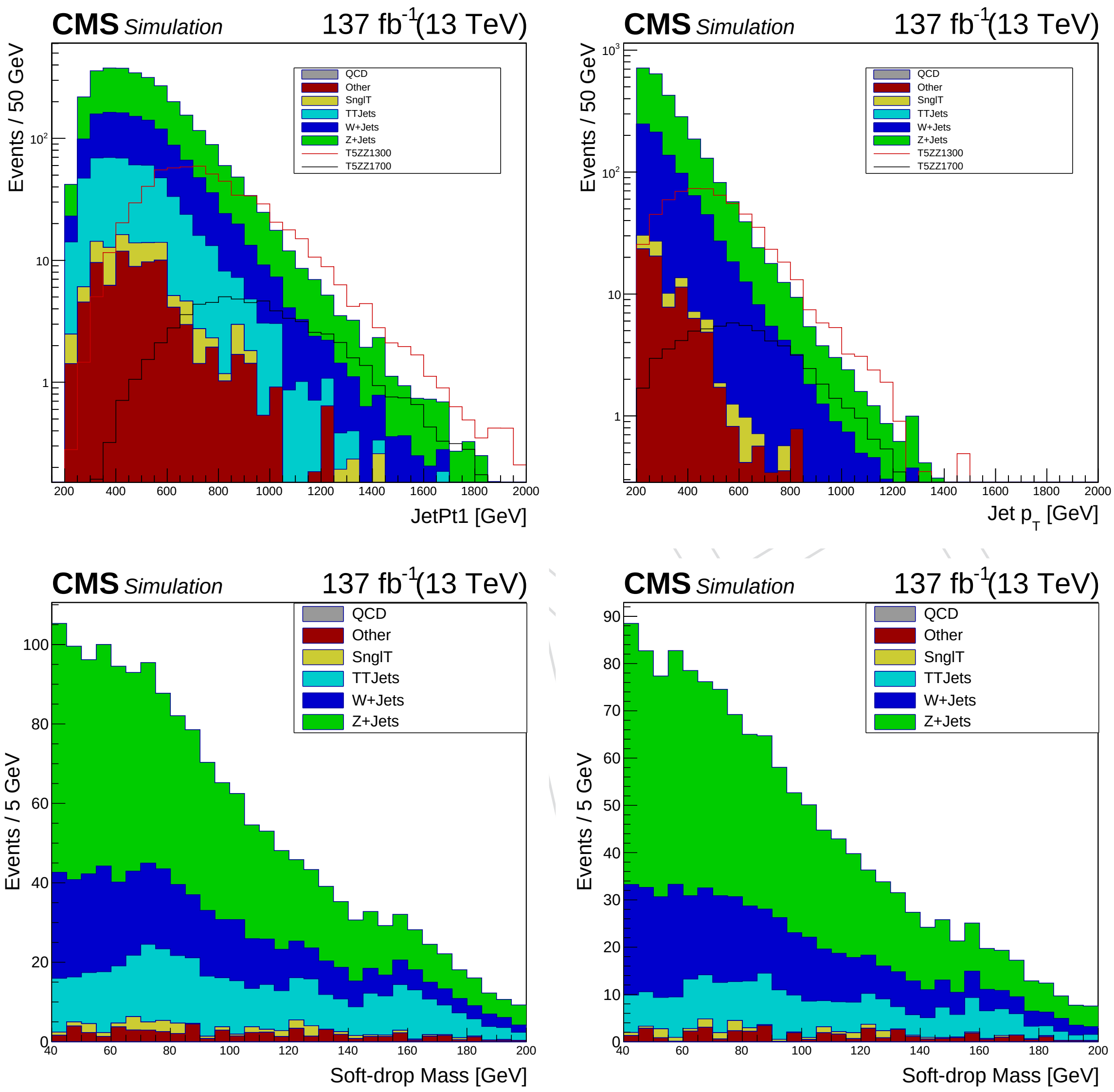


Figure 3: Top: Jet p_T distributions after full boosted selection. Bottom: Softdrop mass shapes before (left) and after (right) veto on the seperation between sublead AK8 jet and deep CSV b-tagged AK4 jet .

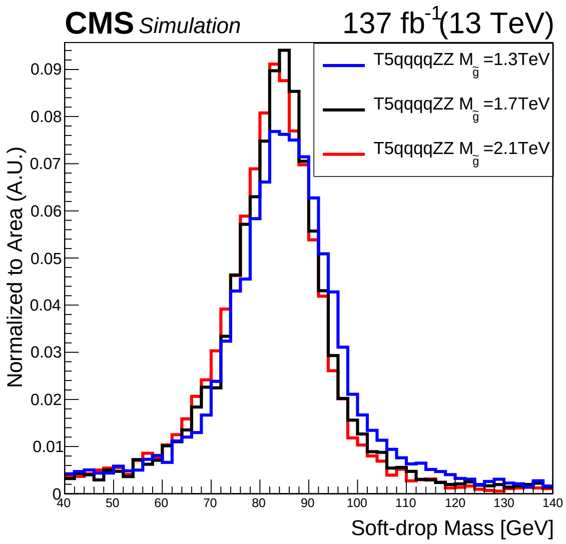


Figure 4: Soft-drop mass shapes for the signal in the SMS T5qqqqZZ model for representative gluino points after applying the boosted event selection. A gaussian fit to the signal fits show that within 1.5σ of the gaussian mass resolution 80% of the signal events within the soft-drop mass window of $[70, 100]$ GeV.

Table 8: Cut flow table for each of the main MC background samples and their sum. The table also includes the signal for two mass points of the 2Z final state. Cuts on the AK8 jets are applied to both the leading and subleading jets.

Cut	Total Bkg.	$Z \to \nu\nu$	$W \to \ell\nu$	$t\bar{t}$	QCD	T5HH1300	T5HH1700
Baseline SUSY Hadronic Skim: $E_{\mathrm{T}}^{\mathrm{miss}} > 300\,\mathrm{GeV}$, $H_T > 300\,\mathrm{GeV}$, $\Delta\phi$ cuts, lepton veto, isolated track veto							
Baseline SUSY Hadronic Skim	792649.0	559899.0	219422.0	4459.5	8869.0	5313.1	551.5
$E_{\mathrm{T}}^{\mathrm{miss}} > 300\,\mathrm{GeV}$	792649.0	559899.0	219422.0	4459.5	8869.0	5313.1	551.5
	318870.9	224238.0	86327.1	3923.8	4382.0	5042.2	536.2
AK8 Lead and sublead Jet $p_{\mathrm{T}} > 300\,\mathrm{GeV}$	59077.1	38863.4	15777.7	3753.1	682.9	4997.9	534.6
AK8 Jet Mass in $[40, 140]$ GeV	4357.6	2542.6	1107.4	663.6	44.0	2362.2	261.3
Δ R > 0.8 for sublead AK8 Jet	3782.8	2317.4	1016.2	415.2	34.0	1221.8	133.7
AK8 Jet Mass in $[70, 100]$ GeV	369.0	216.7	94.0	58.0	0.3	646.2	77.8

4.3 Signal Regions

Applying the hadronic baseline selection in Section 4.1 ensures an event topology with jets and missing energy and then also applying the cuts in Section 4.2 targets an event topology with high p_T boosted jets. The soft-drop mass is used to define signal window around the Z-mass in range [70,100] GeV which is found to preserve 80% signal efficiency as shown in Figure **??**. The two highest p_T AK8 jets are required to be in this mass range for the signal region.

These events are then categorized by E_T^{miss} in the ranges shown in Figure 5. The requirement on the AK8 jet masses and large missing energy gives regions of large signal purity, which contain two jets with resolved Z-boson mass and missing energy from the $\tilde{\chi}_1^0$. The loosest E_T^{miss} bin has a mixed composition of the main SM backgrounds, but the larger E_T^{miss} range is dominated mainly by $Z \to \nu\bar{\nu}$.

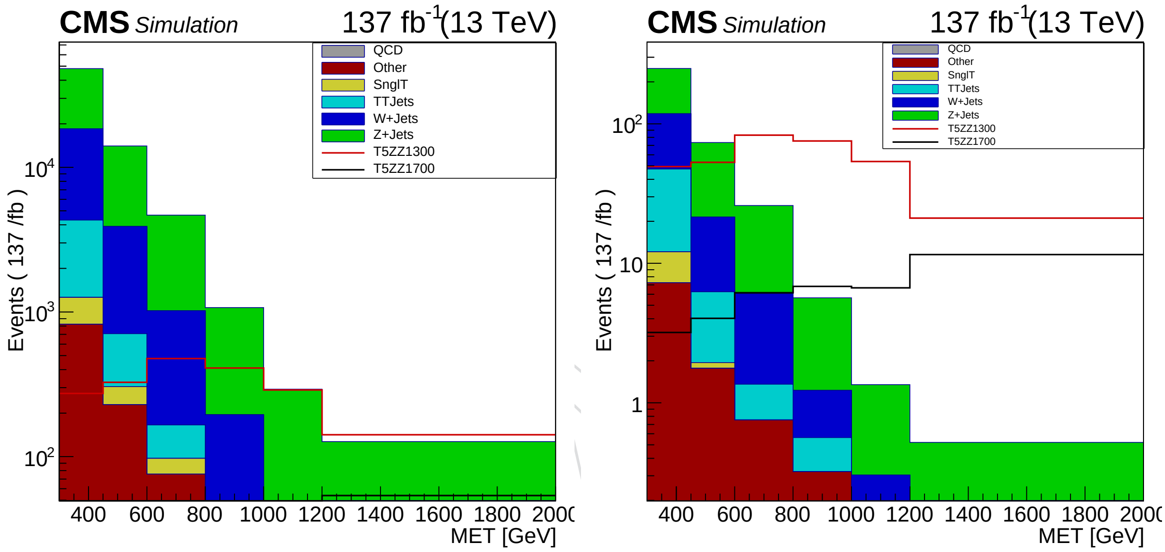


Figure 5: E_T^{miss} Search regions of the boosted Z analysis, on the left with only the hadronic baseline selection on the left and adding the full selection with boosted objects on the right.

4.4 Control Regions

For this search, the control region used to predict the total background in the search regions is based on the mass-sideband of the Z-boson. The search regions require two AK8 jets to be in the signal window, so a background enriched control region is defined by the lead p_T AK8 jet having soft-drop mass outside the Z-boson mass window in [40,70] or [100,140] GeV. The minimum mass range is set to reject the bulk of non-resonant SM processes and the maximum jet mass is set to be away from the soft-drop top mass peak. The mass sideband is used to derive the mass shape of the background PDF, which is used to measure the total background integrating $E_T^{miss} > 300$ GeV. Figure 6 shows the full soft-drop mass range for the lead p_T AK8 jet mass. The full range shows that the bulk of non-resonant SM background which likely have no-substructure has a soft-drop mass near zero. Above 40 GeV the soft-drop mass is modeled well with a smoothly falling function like a polynomial. Centering the window on the Z-mass removes top events which are boosted and allow for a mass shape fit that falls linearly. A key feature of the analysis is that the E_T^{miss} in this sideband region has the same shape as in the search region. This assumption is shown in terms of the change in soft-drop mass shape in Figure 6.

The figure also shows that the precision of the mass fit at high E_T^{miss} is lower in the simulation because of limited statistics. For this reason, we derive the mass fit integrating over all the E_T^{miss} search regions.

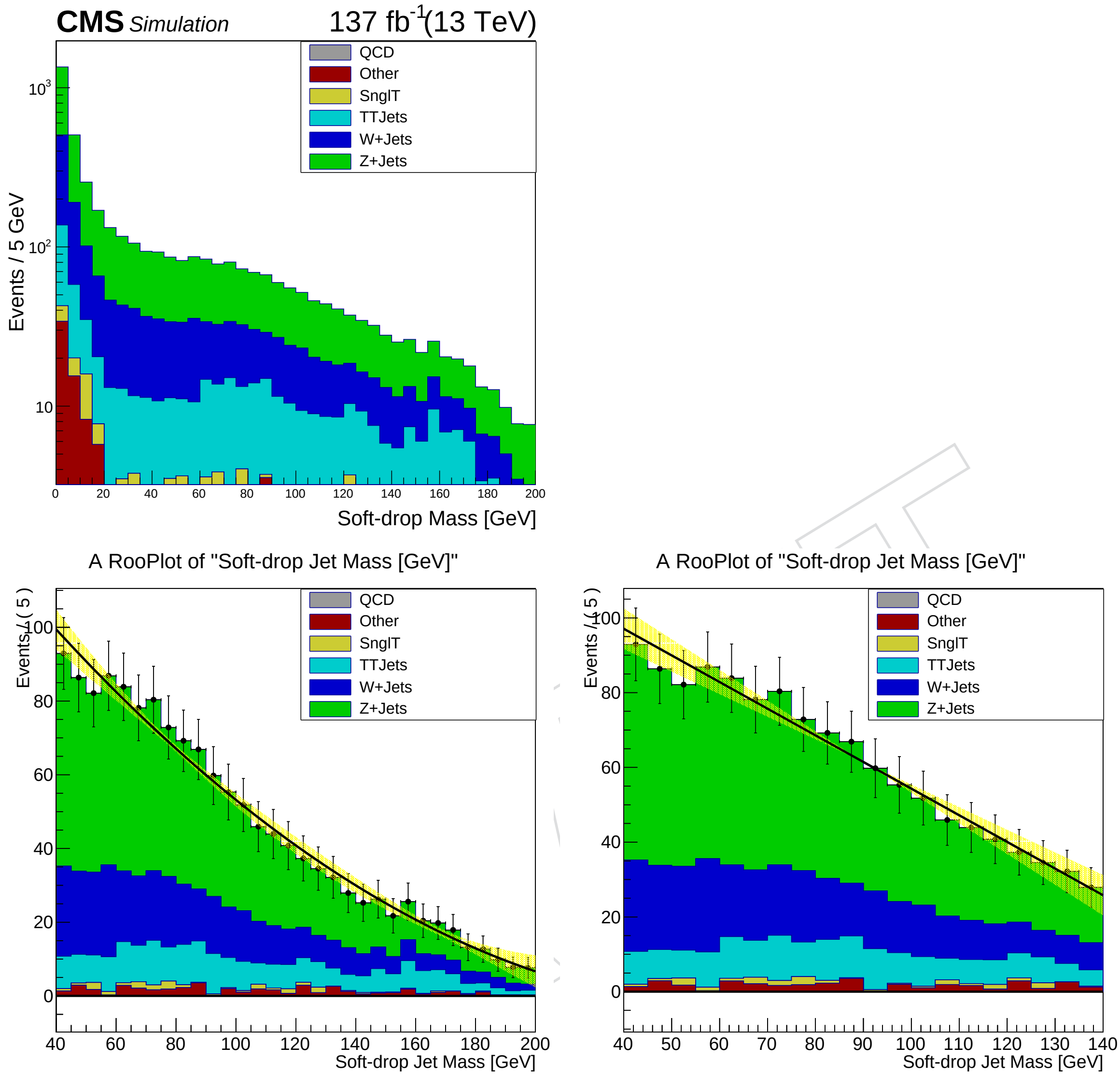


Figure 6: A look at the mass shapes in different ranges of lead AK8 jet soft-drop mass. (top) The full range of soft-drop mass shows a large peak near zero for background events that have little or no sub-structure. Most jets in this range likely will only have a single jet axis. (bottom) Requiring a minimal soft-drop mass of 40 GeV greatly reduces the bulk of the SM background and also makes rare background contributions very small. Above this mass range the background is well-modeled in simulation by a second order (left) or first order (right) polynomial. The narrow side band range (right) removes the top peak from the considered mass range and is centered PDG mass value of the Z-mass of 90 .

The E_T^{miss} shape in the mass sideband is used to measure the fraction of the total background in each E_T^{miss} search bin. This sideband is taken from events wthere both lead and sublead AK8jets have soft-drop mass outside the Z-boson mass window in [40,70] or [100,140] GeV. To validate that the E_T^{miss} shape does not vary between events in the mass sideband and the Z-boson mass window, we use two main validation regions enriched in each of the main SM background processes. Events

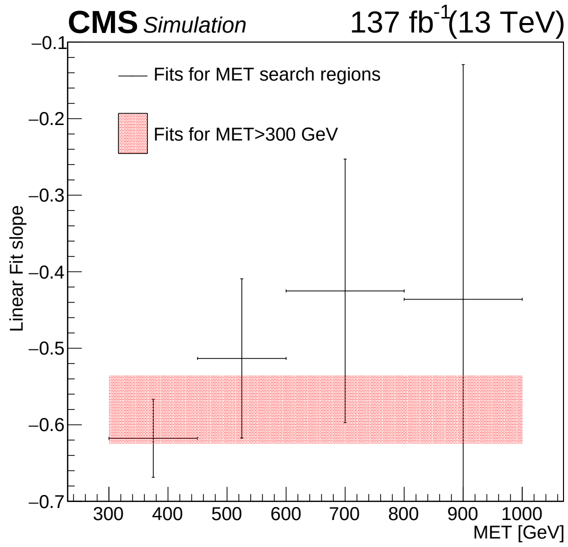


Figure 7: The mass shape fit to the sideband region of [40,70] or [100,140] GeV in sideband mass is done for each of the MET search bins up to $E_{\mathrm{T}}^{\mathrm{miss}} > 800$ GeV. The slope is shown to be consistent with the fit done for $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ GeV, but the assumption that the soft-drop mass and $E_{\mathrm{T}}^{\mathrm{miss}}$ is uncorrelated is further tested in the simulation and the validation region. The bands show the uncertainty on the slope for the fit derived integrated over $E_{\mathrm{T}}^{\mathrm{miss}}$ search regions, while the error on the points are the uncertainty on the fitted slope in four $E_{\mathrm{T}}^{\mathrm{miss}}$ search bins.

triggered with the search region trigger, but including a single electron or a single muon is used validate the $E_{\mathrm{T}}^{\mathrm{miss}}$ shape for W+Jets and TTBar background processes. Events triggered by the single photon trigger is used to validate the $E_{\mathrm{T}}^{\mathrm{miss}}$ shape from $\mathrm{Z} \to \nu\overline{\nu}$, which is the dominant background in this search. Since $\mathrm{Z} \to \nu\overline{\nu}$ is the main component of the background, we also validate this region in events with Z decays to leptons.

5 Background Estimation Method

This section focuses on the estmation of the SM background in the $E_{\mathrm{T}}^{\mathrm{miss}}$ search regions. Our method first uses a fit to the background mass shape in the sideband to derive the normalization scale across the search regions, which will be described in detail in Section 5.1. This sideband is based only on the lead AK8 jet being outside of the soft-drop Z-window [40, 70] and [100, 140] GeV while the sublead jet is in the Z-window [70, 100]. The $E_{\mathrm{T}}^{\mathrm{miss}}$ shape is taken from a non-overlapping sideband where both lead and sublead AK8 jets are in the sideband region of the soft-drop Z-window [40, 70] and [100, 140] GeV. In this way we make use of two exclusive control regions for the background normalization and the $E_{\mathrm{T}}^{\mathrm{miss}}$ shape. The analysis relies on the $E_{\mathrm{T}}^{\mathrm{miss}}$ shape in the sideband being compatible with the signal region within the statistical uncertainty. This will be the focus of Section 5.2.

The correlation between $E_{\mathrm{T}}^{\mathrm{miss}}$ and soft-drop mass is tested in simulation for each of the SM background components in the search region. Figure 8 shows the $E_{\mathrm{T}}^{\mathrm{miss}}$ shapes in the two regions as well as the ratio of events in each region. The ratio shows compatible shapes for each background component. The largest deviation is seen for $t\bar{t}$ which is the smallest background component, but the deviation is still within the assigned statistical uncertainty.

The simulation is only used to test whether there is strong correlation between the soft-drop mass and the E_T^{miss} and closure of the full method. The final estimation of the background comes from the data sidebands. To validate the same assumption in data, we make use of validation regions that are enriched in a SM component. In particular, dilepton and single photon data are used to vlaidate that the E_T^{miss} shapes are similar in the search region and sideband for $Z \to \nu\bar{\nu}$, while a single lepton region is used to validate this assumption for W+jets and $t\bar{t}$. The data-driven background estimate will be described in Section 5.3.

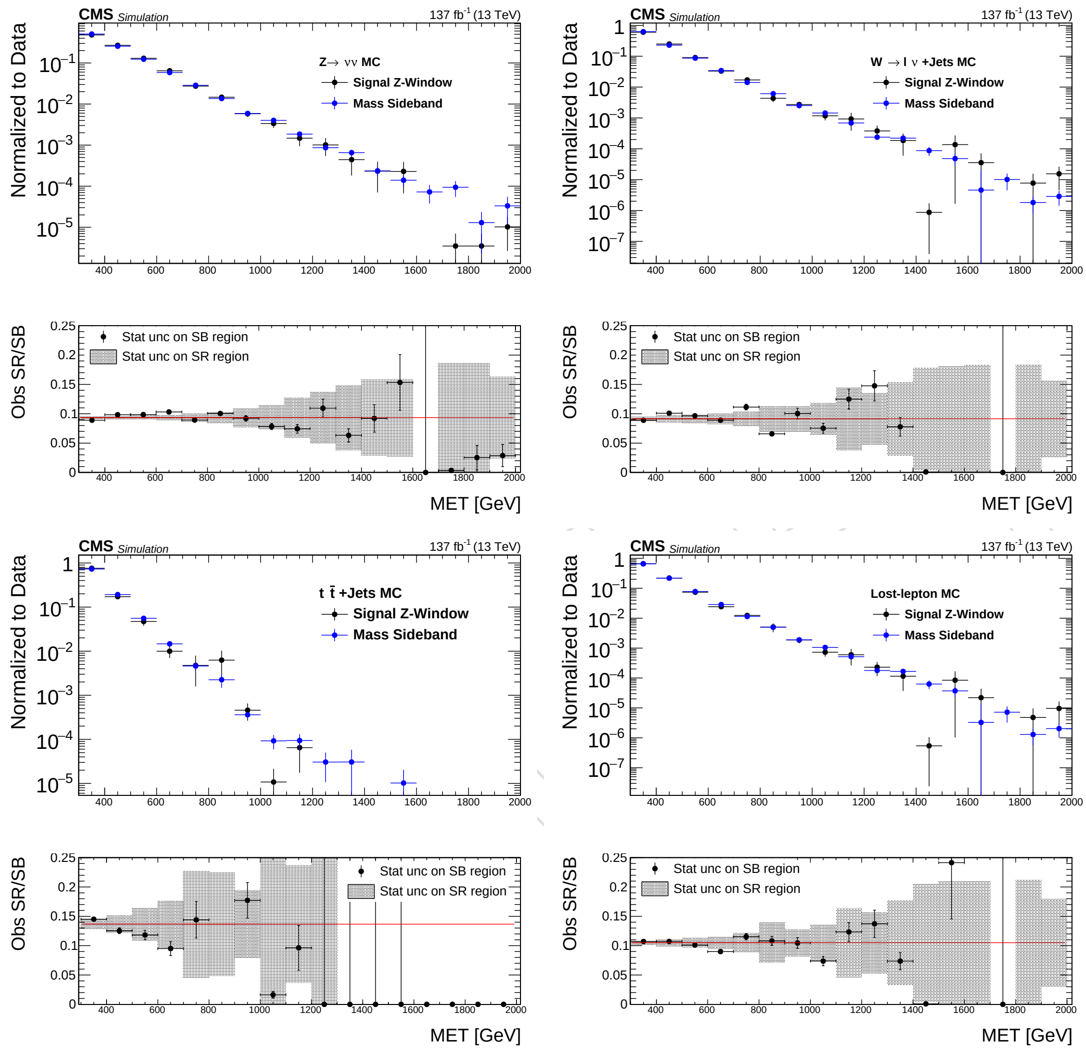
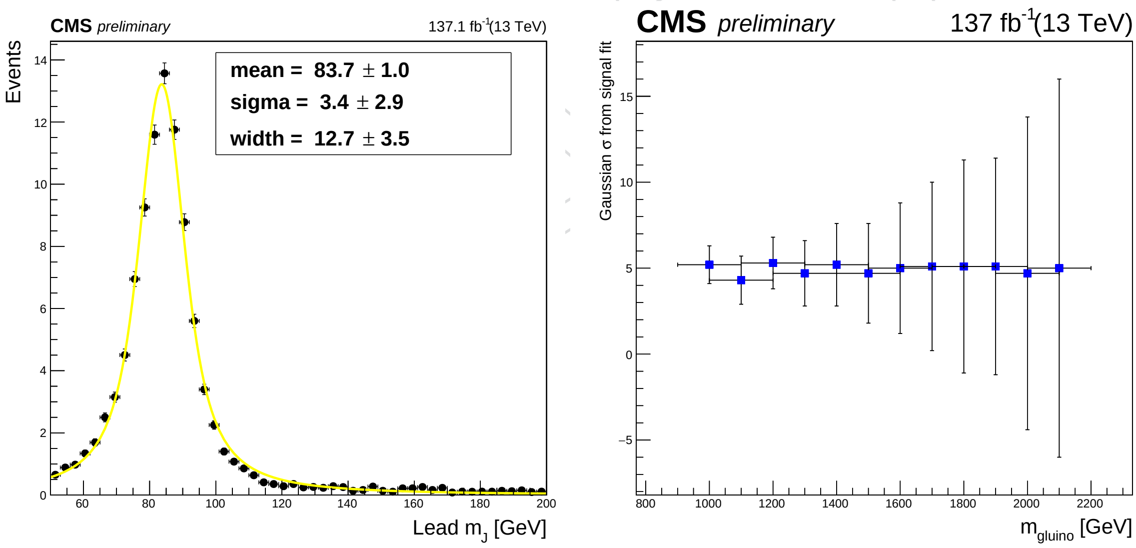


Figure 8: Comparison plots of the E_T^{miss} shape in the sideband control region and the search region. The top panels show the normalized E_T^{miss} shapes in the two region while the bottom panels show the ratio of events in the search window to the sideband. This comparison is done for each of the main background components: (top left) $Z \to \nu\bar{\nu}$, (top right) W+jets, (bottom left) $t\bar{t}$ +jets. Since $t\bar{t}$ is validated along with W+jets in data, we sum the two backgrounds (bottom right) as a single SM background lost-letpon (W decays where the lepton is not measured). A fit to a constant is included in the bottom panel to show the average ratio.

5.1 Mass shape fitting

This section describes the background estimation from the method of soft-drop mass shape fitting. We fit the mass shape from both MC and data to crosscheck our understanding of the background mass shape. For the final uncertainty and results, we will completely rely on data. We fit the MC soft-drop mass shape in the sideband with different functions such as linear, exponential and 2nd order Chevychev ploynomial. We observed that the linear function fits our mass shape more accurately compared to other models. The fitted shape is shown in Figure 9. To determine the bias of our fit, we perform a bias study from toys generated from different background models. In total 1000 toys are generated using each alternate background model. Generated toys are fitted to a test model. The pull distribution of fitted toys gives the uncertainty in mass shape fit.

We observe that the highest deviation in pull comes from toys generated using exponential fit as shown in Figure 10. In MC, the fits are performed from the total background shape assuming that signal mass shape is always fixed for all gluino mass points. To verify this, we fitted the signal shape separately with a Gaussian convoluted with Breit-Wigner and observe that the mean and sigma of the fitting do not vary along different gluino mass points, shown in Figure 5.1. Based on the largest observed bias in the pull distributions in Figure 10, the shift in the fitted background gives a small bias in the signal region $\approx 2\%$ of the statistical uncertainty of the sideband. Figure 11, shows similar features that the bias terms are small when considering background models fit to data. Given these results, only the statistical uncertainty of the sideband is assigned for the background normalization uncertainty.



5.2 MET Shape Validation

This section describes the background estimation from the control regions and the measured systematics on the E_T^{miss} shape in the validation region. The closure in simulation will test if the E_T^{miss} shape in the sideband and Z-signal window are compatible within the statistical uncertainties. Only the data is used for the final results and the systematics, but the MC is used to show that the background estimation strategy is robust. This search relies on minimal correlation between E_T^{miss} and the soft-drop mass so that the E_T^{miss} shape can be taken from the background-rich mass-sideband. The validation regions are enriched in the main background processes and verify that the E_T^{miss} shape in the sideband is compatible with Z-boson mass-window.

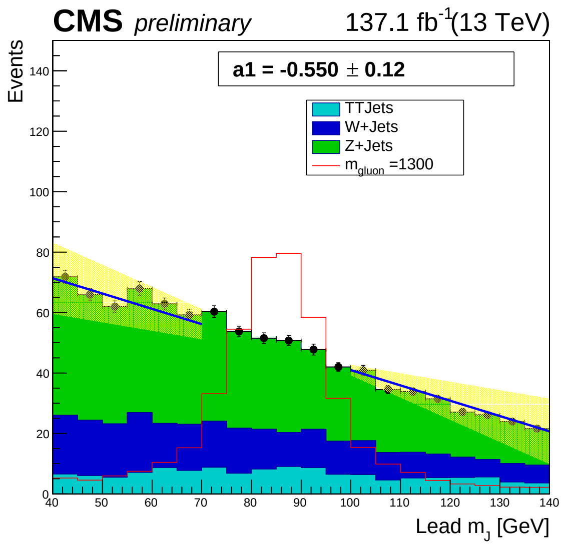


Figure 9: MC soft-drop mass shape fit of total background with linear function in the sidebands, blinding the signal region.

Two validation regions are used for the Z $\rightarrow \nu\bar{\nu}$ $E_{\mathrm{T}}^{\mathrm{miss}}$ shapes a single photon region where the photon is removed from the event to emulate the $E_{\mathrm{T}}^{\mathrm{miss}}$ and a dilepton region enriched in Z $\rightarrow \ell^{+}\ell^{-}$. The Z $\rightarrow \ell^{+}\ell^{-}$ region requires a dilepton Z p_{T} larger than 100 GeV and require the dilepton mass to be within 15 GeV of the Z-boson mass (90 GeV). The same selection is applied as for the signal region but the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut is lowered to 100 GeV to allow to probe any trends in the $E_{\mathrm{T}}^{\mathrm{miss}}$ shape from the low to high range. The two validation regions allow to crosscheck any trends in $E_{\mathrm{T}}^{\mathrm{miss}}$ betweeen the sideband and the signal region for the dominant background Z $\rightarrow \nu\bar{\nu}$. The remaining part of the standard model background are events where a lepton from a W-decay is not measured. These lost-lepton events mainly result in W+jets and $t\bar{t}$ processes. A single-lepton region is used to validate the $E_{\mathrm{T}}^{\mathrm{miss}}$ shapes for these processes combined. For the single lepton validation region the same cuts are applied as for the search region because the same $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is used.

Figure 12 shows the validation for events in simulation. The top figures show the $E_{\mathrm{T}}^{\mathrm{miss}}$ shapes for the photon MC and the Drell-Yan Z $\rightarrow \ell^{+}\ell^{-}$ events. These plots show some upward trends in the ratio panel, but the photon MC shows that these trends are not significant within the MC statistical uncertainty of the true $E_{\mathrm{T}}^{\mathrm{miss}}$ shape. The single lepton region on the bottom shows compatible results between the signal Z-window and the mass sideband.

The background estimations stratgy is shown in Figure 13. The first plot shows the fit performed to the MC lead jet mass sideband. The PDF used for the background model is a first order Chebychev polynomial. The fitted function is shown in Figure 9 and the integral in the signal window gives the total background across all search bins. The uncertainty on this background normalization is given by the statistical uncertainty on the events in the mass sideband. This gives 3.6% uncertainty for this the MC closure test. Figure 13 also shows the $E_{\mathrm{T}}^{\mathrm{miss}}$ shape from the mass sideband along with the uncertainty on these fractions based on the statistical uncertainty in the $E_{\mathrm{T}}^{\mathrm{miss}}$ sideband region. The fraction of events in each $E_{\mathrm{T}}^{\mathrm{miss}}$ search bin is scaled by the integral of the fitted mass shaape in the search region to give the background prediciton. The main uncertainties on the background method come from the uncertainty in the $E_{\mathrm{T}}^{\mathrm{miss}}$ shape. The MC closure shown Figure 13 shows good agreement in this method within the stat uncertainties.

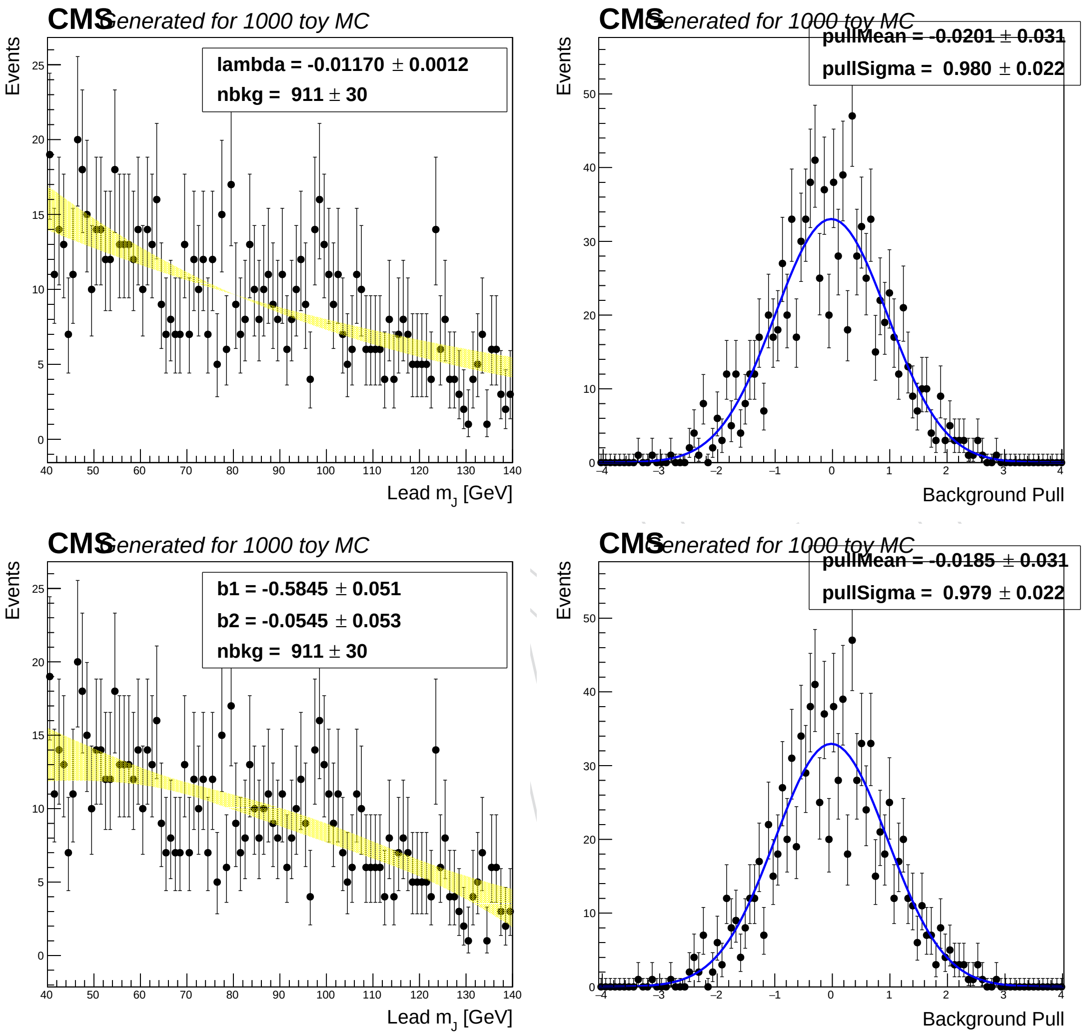


Figure 10: Soft-drop mass fit of one of the generated pseudoexperiments from a fitted exponential (top) or a 2nd order chebychev polynomial (bottom). The pull distributions on the right show the bias estimated from the choice of a linear fit. The background models are derived from fits to the SM Monte-carlo.

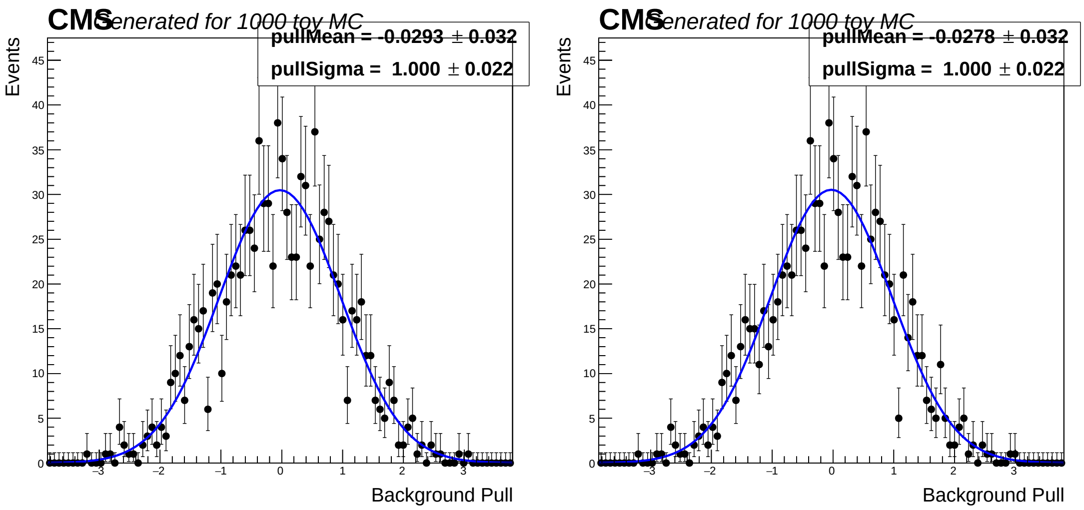


Figure 11: The pull distributions of the bias estimated from the choice of a linear fit. The background models are derived from fits to the data sideband in the MET dataset.

Table 9 shows the full set of measurements from the MC that go into the prediction: the integral from the fitted lead jet mass PDF along with the statistical uncertainty from the mass sideband, the MET shape from the mass sideband region along with its statistical uncertainty. These numbers are multiplied and their uncertainties are combined to give the total systematic error on the prediction. The MC truth is also included in the table for comparison along with its statistical uncertainty. The statistical uncertainties on the prediction are small in MC where the sum of weights still give gaussian uncertainties that are small. For the data-driven approach, the uncertainties on the $E_{\mathrm{T}}^{\mathrm{miss}}$ shape will be driven by poisson uncertainties in the tight $E_{\mathrm{T}}^{\mathrm{miss}}$ regions.

Table 9: Predictions in the signal regions

$E_{\mathrm{T}}^{\mathrm{miss}}$ bin	Bkg in Z-window	MET Fraction	Total Pred.	MC Truth
$E_{\mathrm{T}}^{\mathrm{miss}}[300, 450]$	356.61 ± 12.95	0.71 ± 0.0039	253.0 ± 9.3	261.9 ± 4.5
$E_{\mathrm{T}}^{\mathrm{miss}}[450, 600]$	356.61 ± 12.95	0.20 ± 0.0018	71.7 ± 2.7	74.1 ± 2.7
$E_{\mathrm{T}}^{\mathrm{miss}}[600, 800]$	356.61 ± 12.95	0.069 ± 0.00081	24.6 ± 0.9	25.4 ± 0.9
$E_{\mathrm{T}}^{\mathrm{miss}}[800, 1000]$	356.61 ± 12.95	0.015 ± 0.00036	5.23 ± 0.2	5.5 ± 0.5
$E_{\mathrm{T}}^{\mathrm{miss}}[1000, 1200]$	356.61 ± 12.95	0.0042 ± 0.00021	1.51 ± 0.09	1.2 ± 0.2
$E_{\mathrm{T}}^{\mathrm{miss}} > 1200$	356.61 ± 12.95	0.0015 ± 0.00012	0.52 ± 0.04	0.48 ± 0.12

5.3 Data Driven Background Estimate

Given that the simulation based tests work out to be compatible with the truth, we apply the same procedure to data for the final background predictions. From the fit in data in Figure 14, we derive the normalization for the $E_{\mathrm{T}}^{\mathrm{miss}}$ shape. The data is overlayed on top of the MC for the full mass window in [40, 140] GeV the shapes are in reasonably good agreement. As in the previous section, the mass shape is fit to a linear background PDF.

The single photon triggered data and the single lepton events from the search trigger are used to validate that the $E_{\mathrm{T}}^{\mathrm{miss}}$ shapes in the Z-signal window and mass sideband are compatible. The photon p_{T} is used to emulate the $E_{\mathrm{T}}^{\mathrm{miss}}$ from the Z-boson when it decays to neutrinos. Figure **??** shows the $E_{\mathrm{T}}^{\mathrm{miss}}$ shape comparison for the single photon data. Though compared to

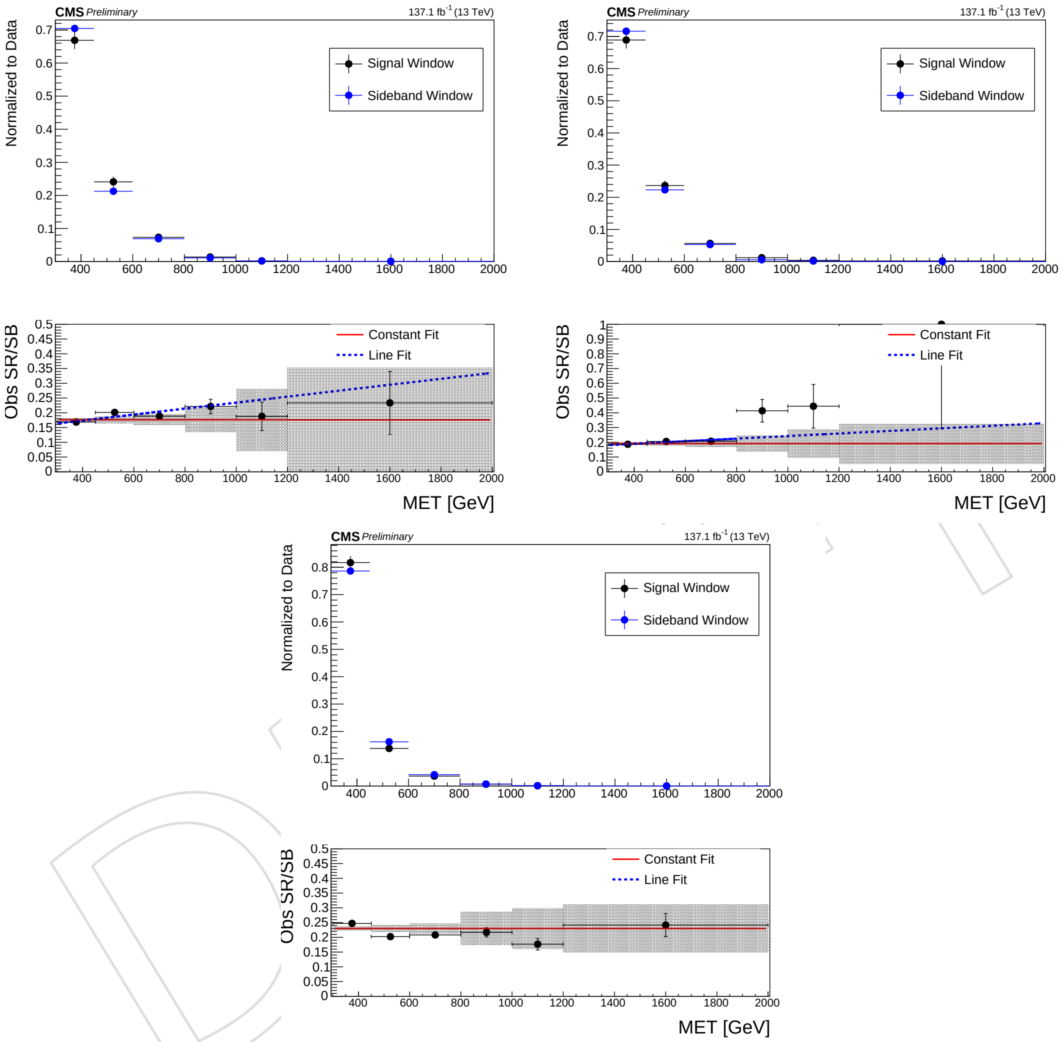


Figure 12: MC comparisons of the MET shape and yields in the Z-boson signal mass window and mass sideband control regions in the validation regions for $Z \to \nu\bar{\nu}$ (top) and the lost-lepton region (bottom). The validation regions for $Z \to \nu\bar{\nu}$ are the single photon region (top right) and the $Z \to \ell^+\ell^-$ region (top left).

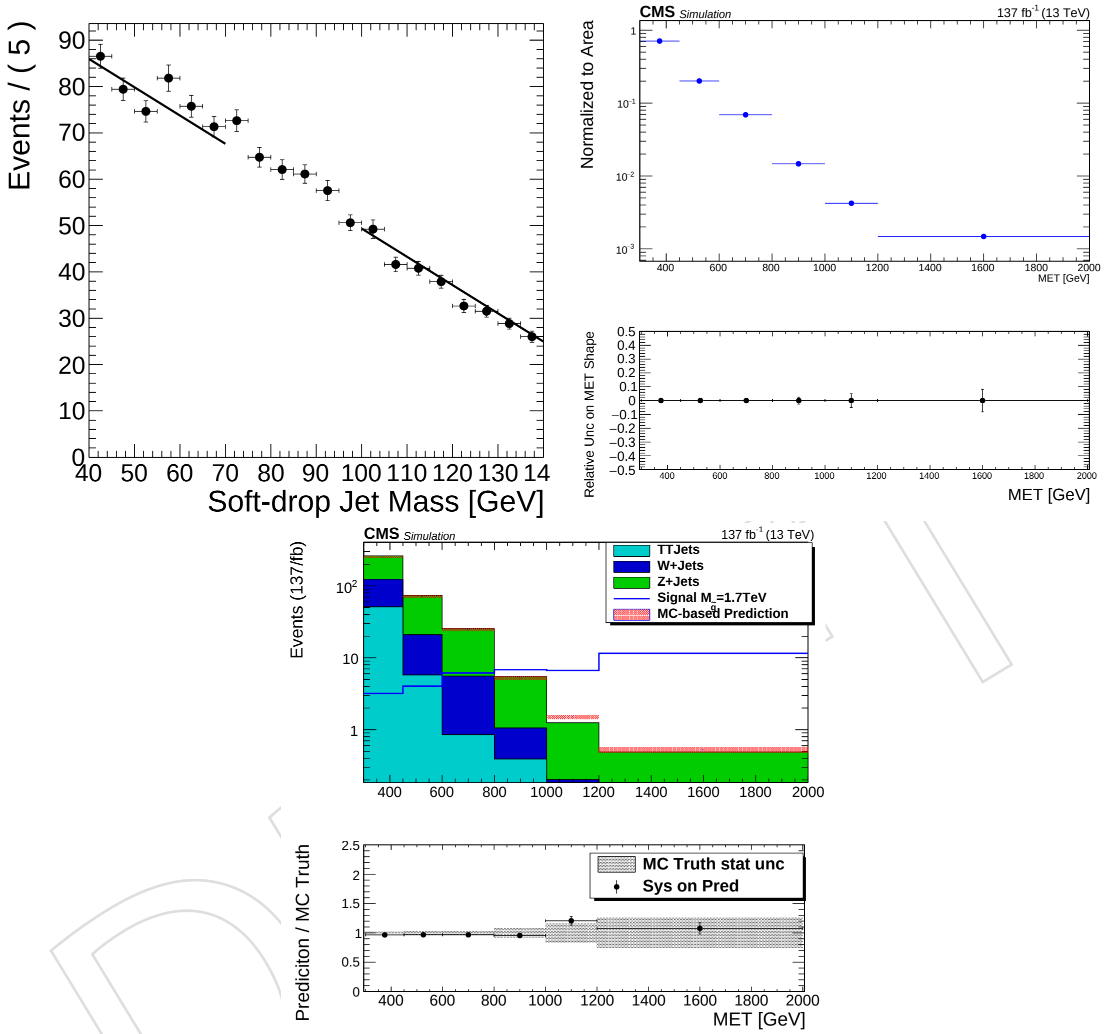


Figure 13: This figure shows the main components that go into the MC-based prediction for the closure test. The mass shape is taken from the mass sideband fit to lead AK8 jet soft-drop mass with a linear function (top left). The $E_{\mathrm{T}}^{\mathrm{miss}}$ shape (top right) is taken from the sideband where the lead and sub-lead mass are in the sideband control region. The bottom panel shows the uncertainties on the $E_{\mathrm{T}}^{\mathrm{miss}}$ shape based on the statistical uncertainty in the sideband. The $E_{\mathrm{T}}^{\mathrm{miss}}$ shape is scaled by the integral from the mass shape PDF to give the full prediciton. (bottom) The prediction is compared to the MC truth in the analysis search bins. The bands show the uncertainty on the MC truth while the error bars on the points show the uncertainty on the ratio from the prediction.

the simulation the photon data is depleted at high E_T^{miss}, we can plot the E_T^{miss} distribution in the lower range starting at 100 GeV. Some deviation is observed by about 33% from the constent fit at E_T^{miss} [600, 800]. As a crosscheck we also perform the same validation in the dilepton data, which shows some deviation in one of the lower E_T^{miss} bins but has good agreement for E_T^{miss} [600, 800]. Figure **??** also shows the single lepton data is compatible between z-search region and sideband up to 800 GeV.

The full data-driven prediction is derived from the fit to the data mass-sideband and E_T^{miss} shape in the sideband. Figure **??** shows the mass-sideband of the lead jet that is used to find the normalization of the background. The integral of the PDF in the search window scales the E_T^{miss} fraction plot shown in Figure **??** to give the background prediction. The bottom panel of the plot shows the relative uncertainty on the E_T^{miss} shape in each analysis search bin. Table 10 shows the data-driven background predictions with each of the systematics tabulated. The E_T^{miss} shape uncertainty is smaller than the uncertainty on the background integral taken as the statistical uncertainty of the data in the sideband.

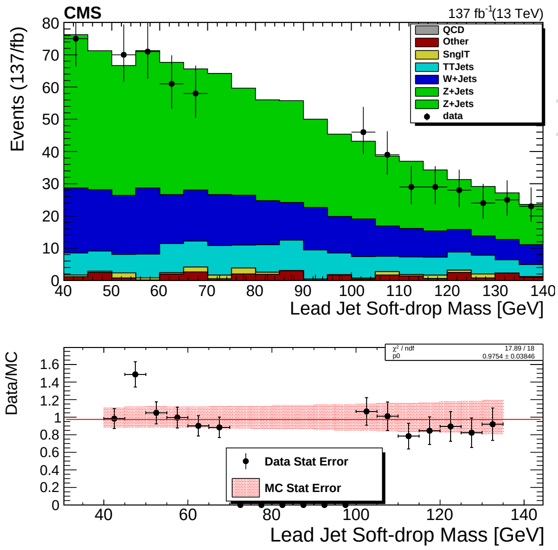


Figure 14: Data soft-drop mass shape fit of total background with chevychev polynomial of order one in the Sidebands blinding the signal region. The bottom panel shows the ratio of data to MC in the same mass sideband. The MC is scaled to the integral of the data to compare mass shapes.

Table 10: Predictions in the signal regions

E_T^{miss} bin	Bkg in Z-window	MET Fraction	Total Pred.	T5ZZ(1700)	
E_T^{miss}[300, 450]	322.9 ± 12.3	0.73 ± 0.014	234.5 ± 10.1		3.17
E_T^{miss}[450, 600]	322.9 ± 12.3	0.20 ± 0.0075	64.1 ± 3.46		4.03
E_T^{miss}[600, 800]	322.9 ± 12.3	0.06 ± 0.0042	19.7 ± 1.55		6.15
E_T^{miss}[800, 1000]	322.9 ± 12.3	0.01 ± 0.0018	3.6 ± 0.59		6.83
E_T^{miss}[1000, 1200]	322.9 ± 12.3	0.002 ± 0.0007	0.56 ± 0.22		6.83
$E_T^{miss} > 1200$	322.9 ± 12.3	0.0009 ± 0.0005	0.27 ± 0.16		18.22

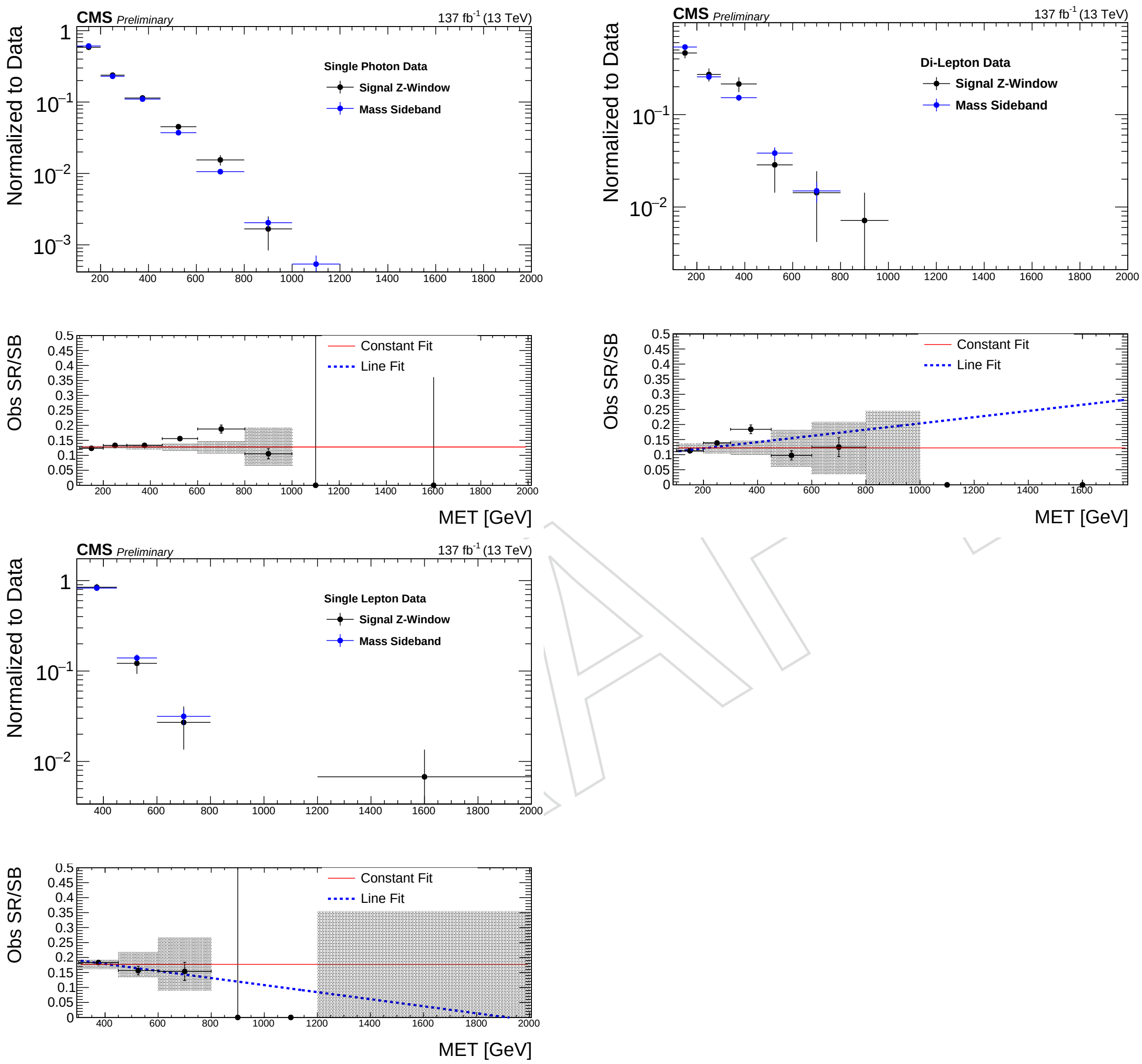


Figure 15: Comparison of $E_{\mathrm{T}}^{\mathrm{miss}}$ shapes in the validation region data for Single Photon (top left) Dilepton (top right) and single lepton (bottom).

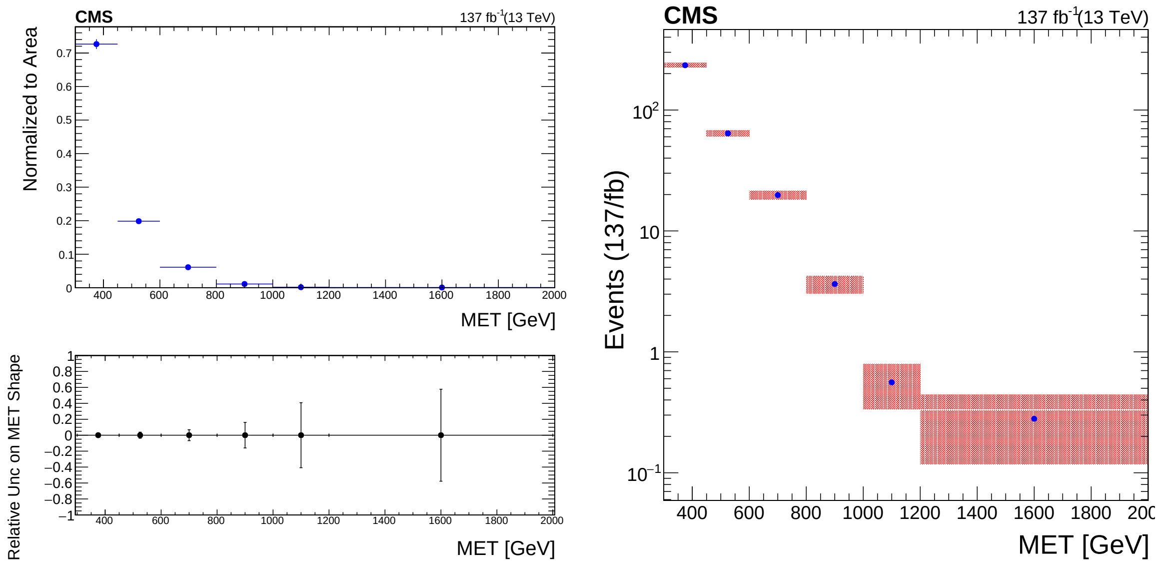


Figure 16: (left) The integral in the Z-window of the backgroun mass shape PDF is used for the normalization for the E_T^{miss} shape (left). The uncertainty on the E_T^{miss} shape is taken from the relative uncertainty in the bottom panel. The data-driven background prediction is shown on the right.

6 Results

Figure 17 shows the exclusion using Asymptotic CLs limits. The likelihood is constructed with the predicted data-driven background in Table 10. The uncertainty on the normalization of the total background is assigned as a log-normal nuisance correlated across all search bins, while the E_T^{miss} sideband statistical uncertainties are assigned as uncorrelated.

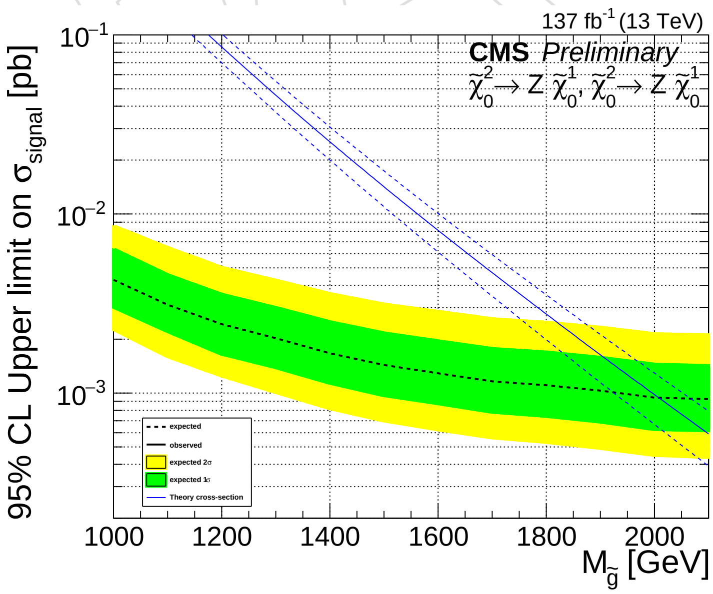


Figure 17: 95% CL limit for the signal model considered in this analysis: 100% branching fraction to the Z boson

7 Summary

8 Effect of Run Dependent Corrections

This section covers the effect of run dependent treatments on either the MC or the data that can affect the $E_{\mathrm{T}}^{\mathrm{miss}}$ shape. Figure 18 shows the effect of applying the pre-fire weights to the 2017 standard model MC. The largest effect is at low $E_{\mathrm{T}}^{\mathrm{miss}}$ where the pre-fire weights can reduce the $E_{\mathrm{T}}^{\mathrm{miss}}$ by 20%. Likewise, the HEM veto treatment in the 2018 dataset shows that the HEM veto is reduces 20% in the lowest $E_{\mathrm{T}}^{\mathrm{miss}}$ region.

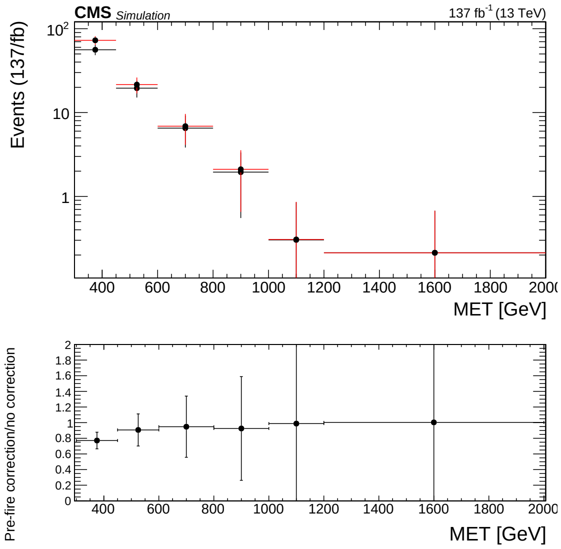


Figure 18: The effect of applying pre-fire weights on 2017 MC.

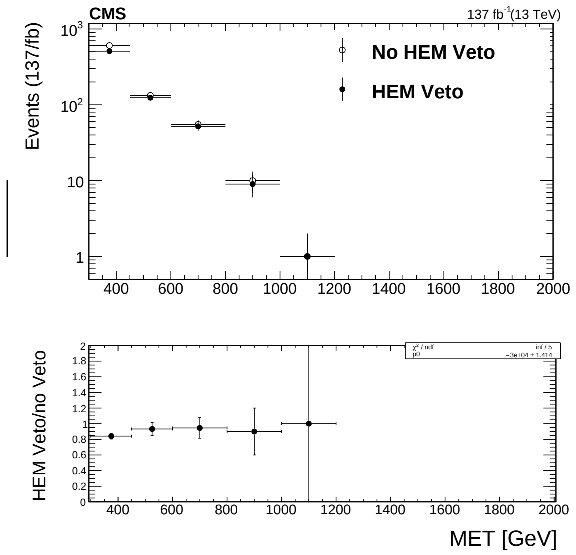


Figure 19: The effect of applying the HEM Veto on 2018 Data.

References

[1] BRIL Group, "Bril work suite", (2017). http://cms-service-lumi.web.cern.ch/cms-service-lumi/brilwsdoc.html.

[2] CMS Collaboration, "Search for supersymmetry in proton-proton collisions at 13 TeV in final states with missing transverse momentum and jets", CMS Physics Analysis Note AN-2018/271, CERN, 2019.

[3] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[4] CMS Collaboration, "Particle–flow event reconstruction in CMS and performance for jets, taus, and E_T^{miss}", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, CERN, 2009.

[5] CMS Muon POG, "Reference muon id, isolation and trigger efficiencies for 2016 legacy re-reco data", (2018). https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2016LegacyRereco.

[6] Rohan Bhandari, Adam Dishaw, Ryan Heller, Ana Ovcharova, Jeffrey Richman, Manuel Franco Sevilla, David Stuart, Chris West, Jae Hyeok Yoo, "Search for supersymmetry in pp collisions at $\sqrt{s} = 13$ tev in the single-lepton final state using the sum of masses of large radius jets", *CMS Analysis Note* **AN-15-139** (2015).

[7] CMS E/gamma POG, "Cut based electron id for run 2", (2014). https://twiki.cern.ch/twiki/bin/viewauth/CMS/CutBasedElectronIdentificationRun2.

[8] CMS JetMET POG, "Met filter recommendations for run 2",. https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFiltersRun2.

[9] C. J. POG, "Reweighting recipe to emulate level 1 ecal prefiring", (2018). https://twiki.cern.ch/twiki/bin/view/CMS/L1ECALPrefiringWeightRecipe.

[10] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft Drop", *JHEP* **05** (2014) 146, doi:10.1007/JHEP05(2014)146, arXiv:1402.2657.