

**Search for supersymmetry using boosted Higgs bosons and
missing transverse momentum in proton-proton collisions
at 13 TeV**

by

Frank Jensen

M.S., University of Colorado Boulder, 2014

B.A., University of California Berkeley, 2008

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Physics

2018

This thesis entitled:
Search for supersymmetry using boosted Higgs bosons and missing transverse momentum in
proton-proton collisions at 13 TeV
written by Frank Jensen
has been approved for the Department of Physics

Professor Kevin Stenson

Professor William Ford

Professor Jamie Nagle

Professor John Cumalat

Professor Senarath de Alwis

Date _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Jensen, Frank (Ph.D.,experimental high-energy physics)

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

Thesis directed by Professor Kevin Stenson

A search for physics beyond the Standard Model in events with one or more high-momentum Higgs bosons, H , decaying to pairs of b quarks in association with missing transverse momentum is presented. The data, corresponding to an integrated luminosity of 35.9 fb^{-1} , were collected with the CMS detector at the LHC in proton-proton collisions at the center-of-mass energy $\sqrt{s} = 13 \text{ TeV}$. The analysis utilizes a new b quark tagging technique based on jet substructure to identify jets from $H \rightarrow b\bar{b}$. Events are categorized by the multiplicity of H -tagged jets, jet mass, and the missing transverse momentum. No significant deviation from standard model expectations is observed. In the context of supersymmetry (SUSY), limits on the cross sections of pair-produced gluinos are set, assuming that gluinos decay to quark pairs, H (or Z), and the lightest SUSY particle, LSP, through an intermediate next-to-lightest SUSY particle, NLSP. With large mass splitting between the NLSP and LSP, and 100% NLSP branching fraction to H , the lower limit on the gluino mass is found to be 2010 GeV.

Dedications

The taxpayers.

Contents

Chapter

1	Introduction	1
2	The Standard Model of Particle Physics	2
2.1	Electroweak	3
2.2	Spontaneous Symmetry Breaking	3
2.3	Quantum Chromodynamics	3
2.4	and beyond... Supersymmetry	3
3	The Large Hadron Collider	5
4	The CMS Detector	6
4.1	Silicon Tracker	6
4.1.1	Pixel Detector	6
4.1.2	Microstrips	9
4.2	Electromagnetic Calorimeter	9
4.2.1	Preshower	9
4.3	Hadronic Calorimeter	9
4.4	Solenoidal Magnet	9
4.5	Muon System	9
4.5.1	Drift Tubes	9

4.5.2	Cathode Strip Chambers	9
4.5.3	Resistive Plate Chamber	9
4.6	Trigger System	9
5	Object Definition and Event Reconstruction	10
6	Search for physics beyond the Standard Model using boosted H bosons and missing transverse momentum in proton-proton collisions at 13 TeV	11
6.1	Motivation & Strategy	11
6.2	Object Definition & Event Selection	13
6.3	Dataset & Trigger	14
6.4	Event Simulation	14
6.4.1	Standard Model Processes	16
6.4.2	Signal Models	17
6.5	Event Binning & Background Estimation	17
6.5.1	Control Regions within Data	20
6.5.2	κ as a Correction to the Estimation	21
6.5.3	Sideband Yields & Predictions	30
6.6	Signal Systematics	31
6.7	Yields in the Signal Regions and Exclusion Curves	35
6.8	Reinterpretation	37
6.9	Conclusions	42
7	Conclusions	43
Bibliography		44

Tables

Table

6.1	SM MC samples used in the analysis.	16
6.2	Summary of the control region scale-factors integrated over p_T^{miss}	24
6.3	Summary of the Single Lepton control region scale-factors the p_T^{miss} in the anti-tag sideband region.	24
6.4	The κ factor computed by throwing Gaussian toys for the scale factors.	28
6.5	Corrected MC Yields in the signal regions.	29
6.6	Corrected MC Yields in each of the signal and sideband regions.	29
6.7	Observed yields in the sideband regions, κ , and background predictions for the 6 signal bins.	31
6.8	Fit results for W-mass resolution in data and MC	33
6.9	Summary of signal shape and normalization uncertainties.	34
6.10	Signal yields and SM background predictions	35
6.11	Comparison of the true reco-level event yield with those obtained via the prediction following the prescription above. The columns with the RECO or GEN labels are the prediction using RECO or GEN event variables only, respectively. The prediction was made using the T5HH MC with a gluino mass of 2200 GeV.	39
6.12	Signal efficiencies for an event to land in a given analysis bin. The efficiencies were derived using the T5HH MC with a gluino mass of 2200 GeV. Choosing a gluino mass of 1800 GeV decreases the efficiencies by a relative 5%.	39

Figures

Figure

2.1	The periodic table of elements.	2
2.2	The particles of the Standard Model.	4
3.1	The LHC complex; CMS is seen on the left side of the image.	5
4.1	A depiction of the CMS detector.	7
4.2	A schematic view of the CMS detector and the particles it interacts with.	8
6.1	Diagrams of the benchmark models used for motivation of the targeted signal.	12
6.2	Distributions of bb-tagging discriminator (top row) and the jet mass (bottom row) for the leading (left column) and subleading (right column) AK8 jets. These are used in H/Z tagging of AK8 jets.	15
6.3	Generator level distributions for the leading (top row) and subleading (bottom row) H boson in the T5HH model. The plot on the left shows the p_T of the H boson. The plot on right shows ΔR between the b-quark daughters - for large H p_T the daughters become collimated.	18
6.4	A diagram of the partitioned phase space. An additional binning in p_T^{miss} brings the number of analysis bins to 6x3=18.	19
6.5	p_T^{miss} distributions and predictions in the signal regions using simulation only.	20
6.6	Closure observed within the data control regions.	22

6.7 p_T^{miss} distribution in the signal and sideband bins within the single photon control region. The hashed red band denotes the prediction from simulation; the solid black points denote the observed yields in data. The Data/MC ratio in the lower panel of each plot represents the scale factor for that bin.	25
6.8 p_T^{miss} distribution in the signal and sideband bins within the single lepton control region. The hashed red band denotes the prediction from simulation; the solid black points denote the observed yields in data. The Data/MC ratio in the lower panel of each plot represents the scale factor for that bin. In order to improve statistics, the low- $\Delta\phi$ requirement has been removed.	26
6.9 p_T^{miss} distribution in the signal and sideband bins within the low- $\Delta\phi$ control region. The hashed red band denotes the prediction from simulation; the solid black points denote the observed yields in data. The Data/MC ratio in the lower panel of each plot represents the scale factor for that bin.	27
6.10 p_T^{miss} distributions and the predictions in the signal regions using simulation with scale factors from data.	28
6.11 p_T^{miss} distribution in each of the control regions comparing data and the scale-factor corrected simulation. The hashed red distribution denote the prediction from simulation; the solid points denote the observed yields in data.	30
6.12 Pruned jet mass in semi-leptonic $t\bar{t}$ events. The mass peak for the W-jets is used to derive the mass resolution uncertainty.	34
6.13 The single event in the A_2 region.	36
6.14 Observed and expected limits on the gluino cross section.	37

6.15 Efficiencies for an AK8 jet originating from Higgs boson decay, relative to baseline selection. The "signal mass" curve represents the probability the jet will fall within the mass region [85, 135 GeV]. The "sideband mass" curve represents the probability the jet will fall within the mass region [50, 85 GeV] or [135, 250 GeV]. The "H-tag" curve represents the probability the jet have a double-b discriminator value greater than 0.3, for jets within the mass region [50, 250 GeV]. The efficiencies were derived using the T5ZH MC with a gluino mass of 2200 GeV.	40
6.16 Efficiencies for an AK8 jet originating from Z boson decay, relative to baseline selection. The "signal mass" curve represents the probability the jet will fall within the mass region [85, 135 GeV]. The "sideband mass" curve represents the probability the jet will fall within the mass region [50, 85 GeV] or [135, 250 GeV]. The "H-tag" curve represents the probability the jet have a double-b discriminator value greater than 0.3, for jets within the mass region [50, 250 GeV]. The efficiencies were derived using the T5ZH MC with a gluino mass of 2200 GeV.	41

Chapter 1

Introduction

A description of the Standard Model is presented in Chapter 2. A description of the Large Hadron Collider (LHC) is presented in Chapter 3. A description of the CMS detector is presented in Chapter 4. A description of the object definition and event reconstruction is presented in Chapter 5. A description of the physics analysis is presented in Chapter 6. The conclusions are discussed in Chapter 7.

Chapter 2

The Standard Model of Particle Physics

The objects responsible for chemistry are found in the periodic table of the elements seen in Figure 2.1. Quantum mechanics was developed in the early 20th century and gave an explanation for the experimentally determined structure to the period table of elements.

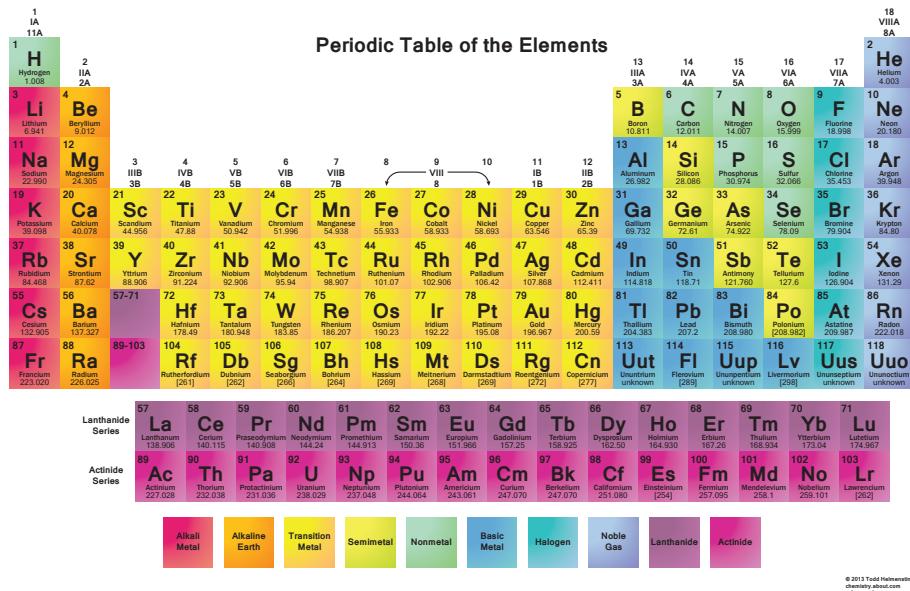


Figure 2.1: The periodic table of elements.

But the elements themselves are made of more fundamental “atoms” described by the Standard Model of particle physics, the current knowledge is summarized in another table, seen in Figure 2.2.

The discovery of the Higgs (H) boson in 2012 was a monumental achievement solidifying the

SM as it found the final fundamental particle within the theory.[1]

2.1 Electroweak

2.2 Spontaneous Symmetry Breaking

2.3 Quantum Chromodynamics

2.4 and beyond... Supersymmetry

Standard Model of Elementary Particles

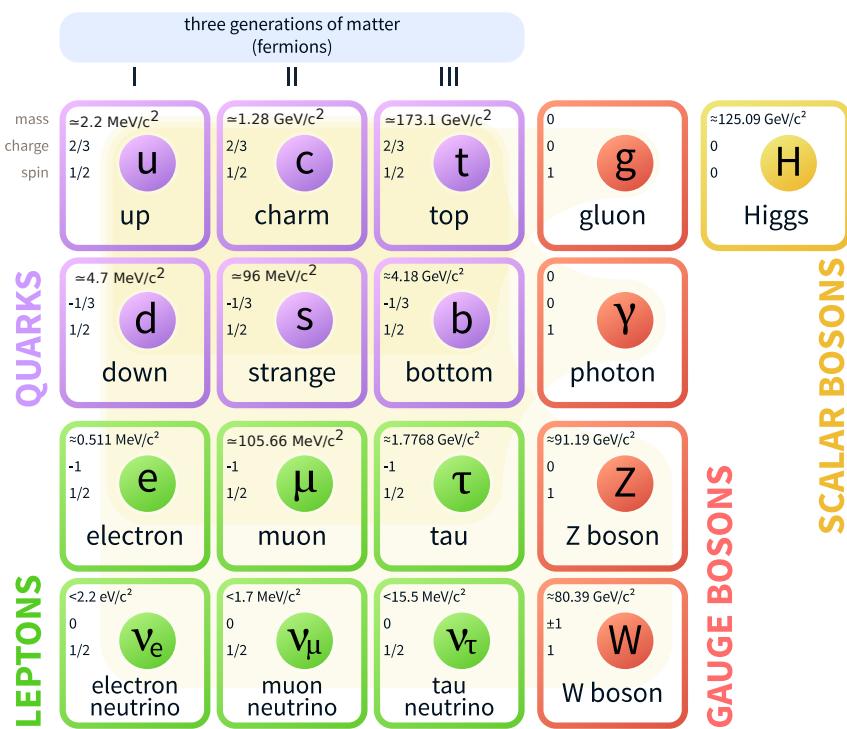


Figure 2.2: The particles of the Standard Model.

Chapter 3

The Large Hadron Collider

An image of the LHC complex is seen in Figure 3.1.



Figure 3.1: The LHC complex; CMS is seen on the left side of the image.

Chapter 4

The CMS Detector

The detector is cylindrically symmetric, with a modular design consisting of different subsystems specialized for detection or identification of specific particles. Many of the Standard Model particles are unstable. In fact only about 8 particles are directly detected in our experiment, the rest of those being reconstructed via their decay into these more stable elements. Of all the Standard Model particles,

Figure 4.1 is a depiction of the CMS detector. Figure 4.2 is a schematic view of the CMS detector and the particles it interacts with.

4.1 Silicon Tracker

At the center of the detector lies the silicon tracker which is responsible for momentum reconstruction of charged particles and reconstruction of the primary vertices created from a proton-proton interaction. Neighboring elements are clustered together creating what is known as a “hit”. These hits are then used to reconstruct the particle’s trajectory through the tracker. To make a measurement of the momentum and charge of the particle, the tracker is immersed in a 3.8 T magnetic field which bends the trajectory of the particle.

4.1.1 Pixel Detector

The tracker is composed of a number of concentric cylindrical layers each instrumented with silicon sensors. As a charged particle traverses an individual sensor a small amount of ionization

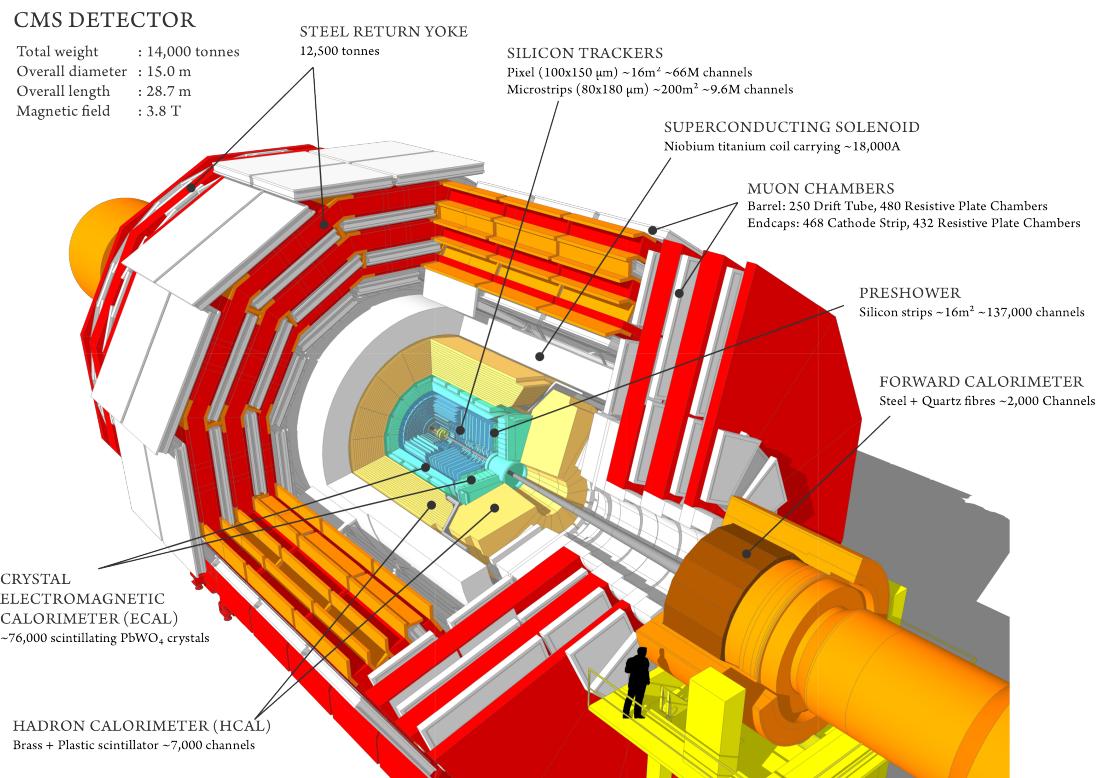


Figure 4.1: A depiction of the CMS detector.

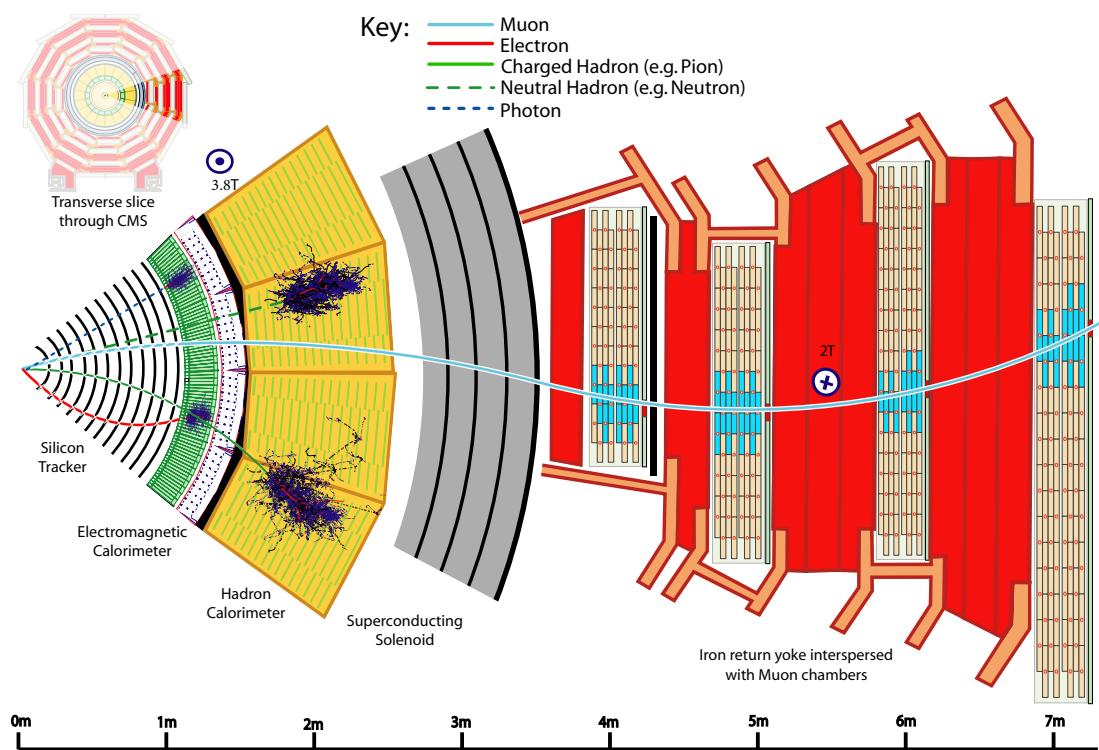


Figure 4.2: A schematic view of the CMS detector and the particles it interacts with.

energy is deposited in that element. A bias voltage is applied across the sensor to direct the charge to either end of the sensor. Bonded to the sensors are front-end electronics responsible for the amplification and shaping of the resultant signal.

4.1.2 Microstrips

The tracker is further divided into an inner and outer parts which make use of slightly different technologies. The inner part of the tracker is known as the pixel detector and consists of 4 layers which the sensitive elements are silicon pixels of size $100 \times 150 \mu\text{m}$.

4.2 Electromagnetic Calorimeter

4.2.1 Preshower

4.3 Hadronic Calorimeter

4.4 Solenoidal Magnet

4.5 Muon System

4.5.1 Drift Tubes

4.5.2 Cathode Strip Chambers

4.5.3 Resistive Plate Chamber

4.6 Trigger System

The trigger system is pretty good.[2]

Chapter 5

Object Definition and Event Reconstruction

Our jets are reconstructed using the “anti-kt” clustering algorithm with cone sizes of $\Delta R = 0.4, 0.8$ [3], denoted as AK4 and AK8 jets respectively. AK4 jets subtend less solid angle and are used to capture the hadronisation of single quarks and gluons. AK8 jets subtend a larger solid angle and are used for reconstruction of boosted objects (e.g. t , H, Z, W).

Particle Flow event reconstruction is pretty good.[4]

Chapter 6

Search for physics beyond the Standard Model using boosted H bosons and missing transverse momentum in proton-proton collisions at 13 TeV

6.1 Motivation & Strategy

If a more unifying theory than the SM exists it certainly has not been forthright in its manifestation. One possibility for the lack of discoveries of phenomena not explained within the SM is that there are indeed particles existing in Nature which have not been observed, but they have such a large mass the energy of the proton-proton collisions provided by the LHC is insufficient to directly create them. The outcome of many searches for new particles is thus the setting of lower limits placed on the mass – if the particle were any lighter than this limit it would have been produced copiously enough for its unambiguous detection (see for example [5], [6]). As these particles become more massive more momentum is imparted upon the particles involved in the final state; any SM particles resulting from the decay of higher mass states will be produced with large momentum (this is called high boost). As a particle becomes more boosted its decay products are emitted at smaller angles, eventually collimating sufficiently to be reconstructed as a single jet. If new physics exists with masses achievable by the LHC, one could suspect that there exists non-zero coupling with the electroweak H, Z, or W bosons. Observation of events containing high- p_T (>300 GeV) electroweak bosons are thus of considerable interest for some hints of something unseen.

The Minimal Supersymmetric SM contains a \mathbb{Z}_2 symmetry in which all SM particles have charge -1 and all supersymmetric particles have charge +1, this is called R-parity[7]. One direct consequence of R-parity is that the decay of a massive supersymmetric particle must include at

least one supersymmetric particle in the final state. Necessarily this is the lightest such particle in the theory, denoted the lightest supersymmetric particle (LSP). If the LSP is electrically neutral it may escape detection, creating an imbalance in the net momentum of the event (as would a neutrino). Events with a large momentum imbalance are also interesting as potential regions for SUSY.

With this as motivation, we designed an analysis searching for hints of new physics beyond the SM in events with boosted H or Z bosons and a large transverse momentum imbalance of the event. We reconstruct the H and Z bosons in the $b\bar{b}$ channel, with 57% and 15% branching fractions respectively. Although our analysis is sensitive to any new physics with this final state, we have adopted two benchmark models (known as SMS models[8], which are phenomenological models of SUSY production), seen in Figure 6.1, to give motivation to the analysis. In this scenario, the proton-proton interaction produces a pair of gluinos \tilde{g} which decay to a neutralino $\tilde{\chi}_2^0$ by the emission of SM quarks. A small mass splitting between the gluino \tilde{g} and neutralino $\tilde{\chi}_2^0$ will result in low- p_T SM quarks and a high- p_T neutralino $\tilde{\chi}_2^0$. This neutralino $\tilde{\chi}_2^0$ further decays into neutralino $\tilde{\chi}_1^0$ with the emission of a H or Z boson. The neutralino $\tilde{\chi}_1^0$ is the LSP and escapes detection.

Past searches targeting similar final states (but different production scenario, i.e. another SMS model) have been performed in which the H bosons are produced with low- p_T , in this case the H bosons are reconstructed as a resolved pair of b-tagged AK4 jets [9].

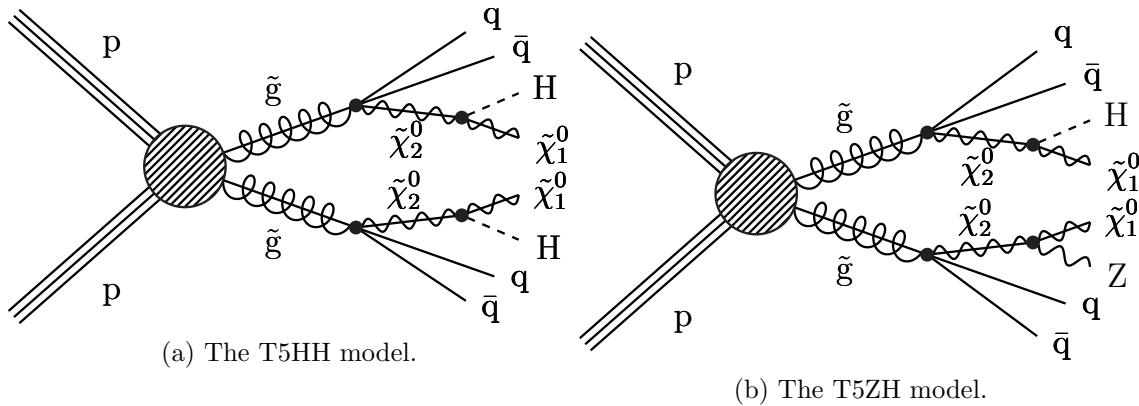


Figure 6.1: Diagrams of the benchmark models used for motivation of the targeted signal.

6.2 Object Definition & Event Selection

We establish a baseline selection choosing events with all-hadronic final states and missing transverse momentum (p_T^{miss}), as motivated by Figure 6.1. The baseline selection is as follows:

- ≥ 2 AK8 jets; $p_T > 300 \text{ GeV}$ and $50 < \text{mass} < 250 \text{ GeV}$
- $p_T^{\text{miss}} > 300 \text{ GeV}$; $p_T^{\text{miss}} \equiv | - \sum_{\text{PF candidates}} p_{T,i}^{\vec{\gamma}} |$
- isolated electron veto; $p_T > 10 \text{ GeV}$

To remove events with top or W production in which the W decays to an electron.

- isolated muon veto; $p_T > 10 \text{ GeV}$

To remove events with top or W production in which the W decays to a muon.

- isolated track veto

To remove events with top or W production in which the W decays to a tau. The tau branching fraction to states containing at least one charged particle is 85%. As an isolated track is defined by looser criteria than that of an electron or muon, this cut also serves to increase the efficiency of the isolated electron and muon vetoes.

- $\Delta\phi_{1,2,3,4} > 0.5, 0.5, 0.3, 0.3$; $\Delta\phi_i \equiv \Delta\phi(p_T^{\text{miss}}, \text{AK4 jet}_i)$

The $\Delta\phi$ cut is designed to mitigate QCD events in which a jet is under-measured leading to an artificial imbalance in the event momentum. This cut requires that the difference in ϕ between the p_T^{miss} vector and each of the four leading AK4 jets is sufficiently large to remove events in which a jet has been under-measured giving rise to fake p_T^{miss} pointing in the same direction. If less than four AK4 jets are available the additional cuts are removed.

A dedicated heavy tagging algorithm is used to identify AK8 jets arising from the decay of two b quarks [10]. The distribution of the output discriminator for the lead and sub-leading jets are seen in the left and right panels of Figure 6.2, respectively; signal-like events peak towards larger values. To $b\bar{b}$ tag the AK8 jets we choose the loose working-point (>0.3) corresponding to a signal

efficiency of approximately 80% per AK8 jet (see Section 6.8). The stacked histogram and solid lines shows the distribution after baseline selection for simulation and two representative signal points, respectively.

Further H/Z tagging of an AK8 jet is accomplished by restricting the jet mass window to [85, 135 GeV] to be consistent with that of the H boson. The distributions of the jet mass are seen in Figure 6.2. The signal shown in the solid line is the T5HH model (i.e. Figure 6.1a). The same identification criteria are applied to tag an AK8 jet as either an H or Z boson, there is no distinction made.

6.3 Dataset & Trigger

We use a total of 35.9 fb^{-1} of data collected by the CMS experiment in 2016. Events are selected in data using a trigger which requires greater than 100 GeV of p_T^{miss} calculated at high-level trigger (HLT); additionally the logical OR of two other triggers with thresholds of 110 and 120 GeV are applied. The trigger efficiency is derived in data using a single-electron reference trigger requiring a tight-ID electron of $p_T > 27 \text{ GeV}$. We further select events with at least three AK4 jets and exactly one reconstructed electron of $p_T > 25 \text{ GeV}$. The signal region trigger is found to be greater than 98% for events with $p_T^{\text{miss}} > 250 \text{ GeV}$ and $\text{HT} > 300 \text{ GeV}$ [5].

6.4 Event Simulation

Event simulation of the proton-proton collisions proceeds in a step-wise manner. To simulate the hard physics process MadGraph@NLO2.2.2 [11] is used to calculate matrix-element amplitudes. The parton distribution functions (PDFs) used in these calculations are from NNPDF 3.0 [12]. Parton showering and other event dynamics are generated with Pythia. The simulation of the interaction of the final state particles with the detector is performed with GEANT [13]. This “raw” simulation data is then at the same tier as data collected from the physical experiment and is merged into a single pipeline for event reconstruction.

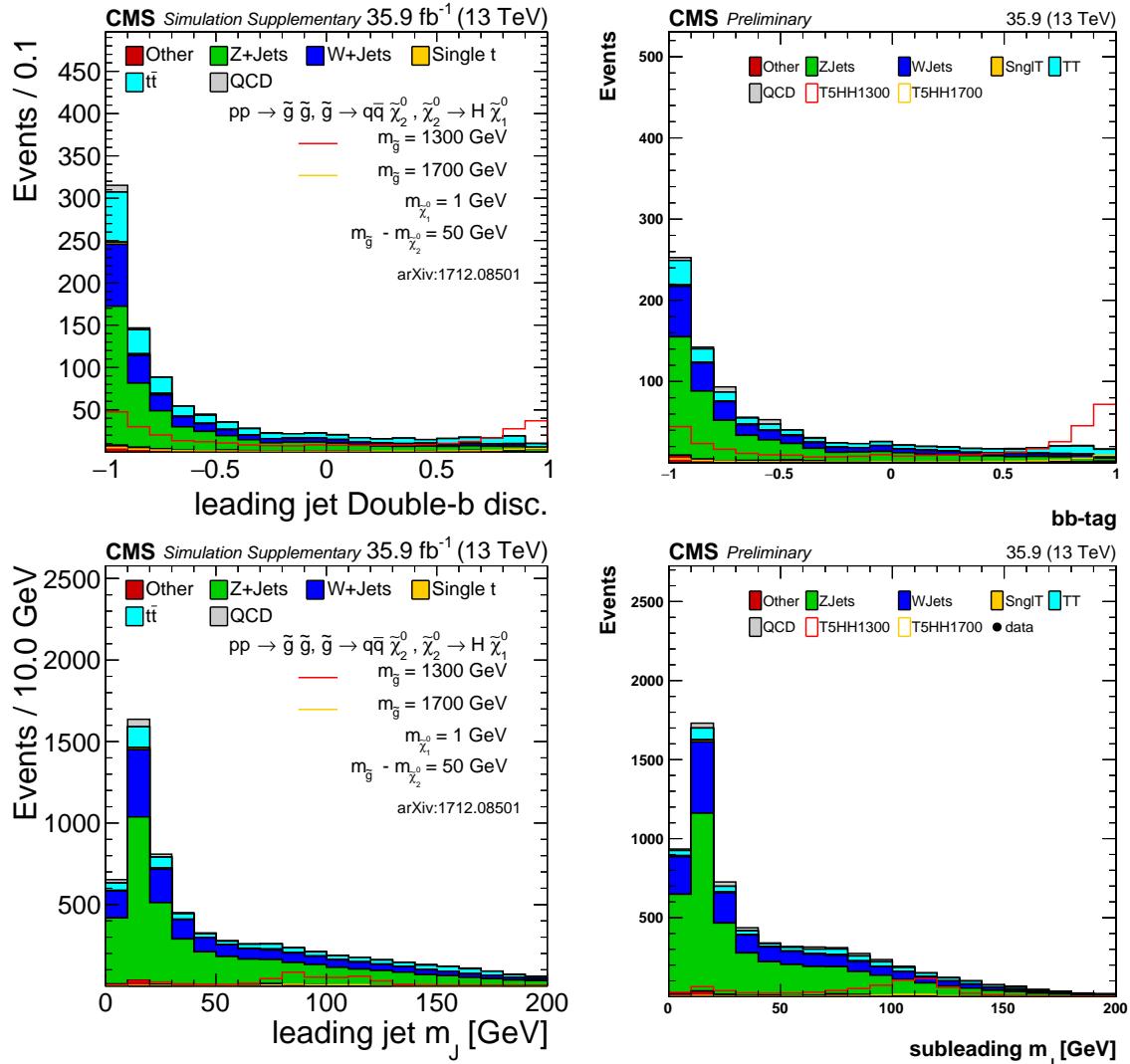


Figure 6.2: Distributions of bb-tagging discriminator (top row) and the jet mass (bottom row) for the leading (left column) and subleading (right column) AK8 jets. These are used in H/Z tagging of AK8 jets.

6.4.1 Standard Model Processes

The SM samples which enter as the primary backgrounds are listed in Table 6.1 (see Section 6.5.1 for a discussion of the SM background). All samples are generated with a pileup distribution with an average of 25 interactions per bunch crossing and a 25ns interval between bunches. For acceptable statistics over a wide range of parameter space, the MC samples are often binned in $\text{HT} \equiv \sum_{\text{AK4 jets}} p_T$. As our event selection requires at least two AK8 jets with $p_T > 300$ GeV we roughly operate in regime of $\text{HT} > 600$ GeV.

Table 6.1: SM MC samples used in the analysis.

process	final state	HT (GeV)	σ (pb)	$\int \mathcal{L}$ (fb^{-1})
t̄t	$t \rightarrow \ell\nu, \bar{t} \rightarrow 2q$	inclusive	182.72	283.90
t̄t	$\bar{t} \rightarrow \ell\nu, t \rightarrow 2q$	inclusive	182.72	326.48
t̄t	2ℓ	inclusive	88.34	346.25
t̄t	inclusive	[600, 800]	2.734	5231.81
t̄t	inclusive	[800, 1200]	1.121	9416.61
t̄t	inclusive	[1200, 2500]	0.198	14819.34
t̄t	inclusive	[2500, ∞]	0.002	221088.29
QCD	inclusive	[200, 300]	1735000	0.03
QCD	inclusive	[300, 500]	366800	0.16
QCD	inclusive	[500, 700]	29370	1.95
QCD	inclusive	[700, 1000]	6524	6.68
QCD	inclusive	[1000, 1500]	1064	12.62
QCD	inclusive	[1500, 2000]	121.5	32.63
QCD	inclusive	[2000, ∞]	25.42	239.30
Z+jets	$\nu\bar{\nu}$	[100, 200]	344.8	54.13
Z+jets	$\nu\bar{\nu}$	[200, 400]	95.53	208.46
Z+jets	$\nu\bar{\nu}$	[400, 600]	13.20	77.30
Z+jets	$\nu\bar{\nu}$	[600, 800]	3.148	1795.26
Z+jets	$\nu\bar{\nu}$	[800, 1200]	1.451	1486.09
Z+jets	$\nu\bar{\nu}$	[1200, 2500]	0.355	1029.81
Z+jets	$\nu\bar{\nu}$	[2500, ∞]	0.0085	47498.87
W+jets	$\ell\nu$	[100, 200]	1627.45	18.16
W+jets	$\ell\nu$	[200, 400]	435.24	45.88
W+jets	$\ell\nu$	[400, 600]	59.18	123.64
W+jets	$\ell\nu$	[600, 800]	14.58	221.32
W+jets	$\ell\nu$	[800, 1200]	6.66	1123.13
W+jets	$\ell\nu$	[1200, 2500]	1.608	153.44
W+jets	$\ell\nu$	[2500, ∞]	0.039	6497.28

6.4.2 Signal Models

For commissioning of the analysis technique (as well as the limit-setting procedure, see Section 6.7) Monte Carlo samples with our final-state signal topology were generated, as in Figure 6.1. The signal sample follows the same processing chain as the SM samples. The mass splitting between the gluino \tilde{g} and neutralino $\tilde{\chi}_2^0$ is fixed to 50 GeV, resulting in low p_T SM quarks produced in the gluino \tilde{g} decays. The mass of the neutralino $\tilde{\chi}_1^0$ (LSP) is fixed to 1 GeV. We have samples with a range of gluino \tilde{g} masses from 750 to 2200 GeV. The p_T distribution for the generated H bosons in these samples is seen in Figure 6.3 for a number of gluino \tilde{g} masses. Additionally the angular separation $\Delta R \equiv \sqrt{\Delta\phi^2 + \Delta\eta^2}$ between the $b\bar{b}$ pair is shown. As the p_T of a parent boson increases the $b\bar{b}$ pair from its decay tend to align, allowing complete reconstruction with a single AK8 jet.

6.5 Event Binning & Background Estimation

The background estimation procedure makes use of what is known as an “ABCD” prediction in which the analysis phase space is divided into signal and sideband regions; scaling relations are applied to sideband yields to make predictions for the SM background (inclusive in all processes) in the signal regions. The events are categorized according to whether the two leading AK8 jets are a) in the signal or sideband mass region and b) have or have not been $b\bar{b}$ tagged. A diagram of this partitioning is seen in Figure 6.4. An additional dimension is added by binning in p_T^{miss} : [300, 500 GeV], [500, 700 GeV], [700, ∞ GeV]. This gives a total of $2 \times 3 = 6$ signal and $4 \times 3 = 12$ sideband bins. The two signal regions $A_{1,2}$ contain events with one (and only one) or two jets being consistent with H/Z boson decay, respectively.

Assuming that there is no correlation between the jet mass and the $b\bar{b}$ tagging one would expect that

$$\frac{A_{1,2}}{B_{1,2}} = \frac{C}{D} \quad (6.1)$$

Rearranging this gives a prediction for the events in the signal regions

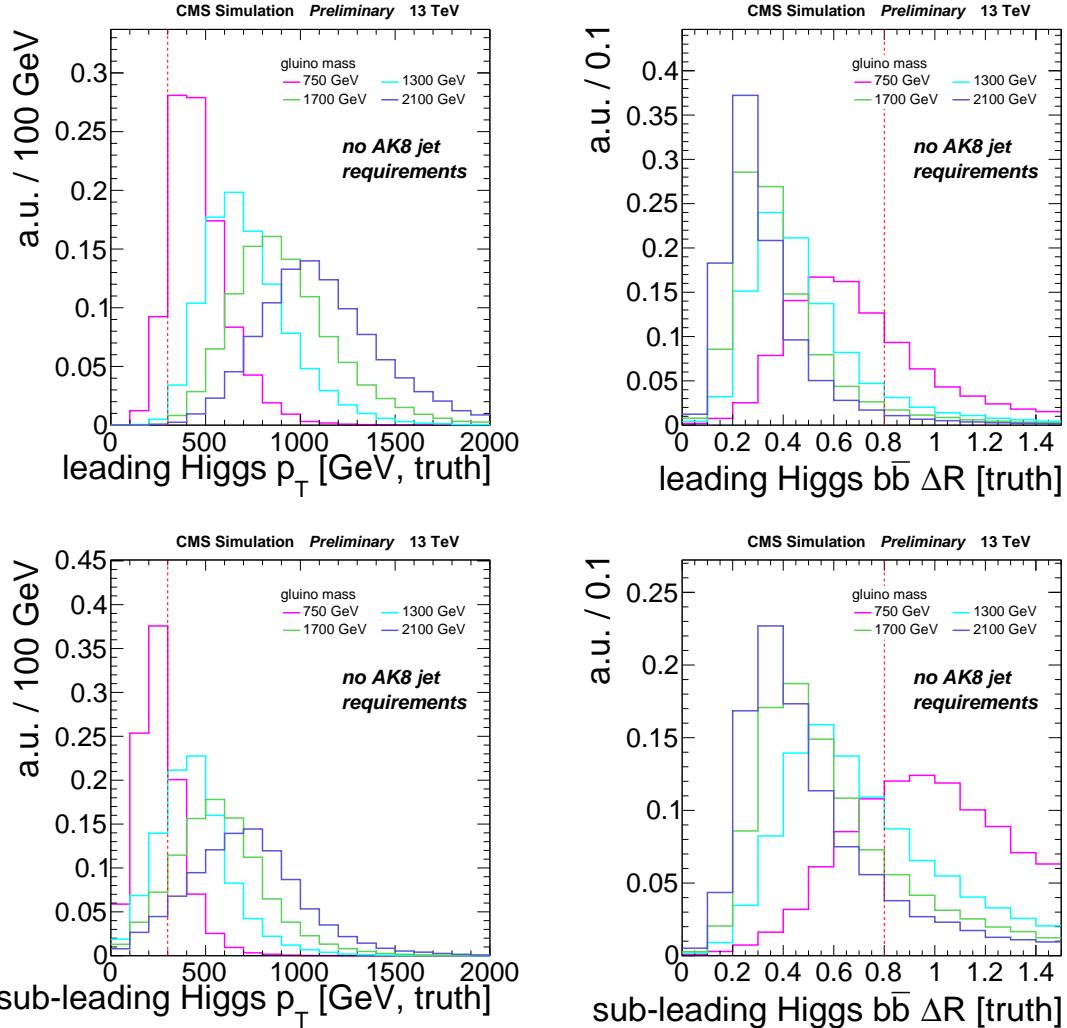


Figure 6.3: Generator level distributions for the leading (top row) and subleading (bottom row) H boson in the T5HH model. The plot on the left shows the p_T of the H boson. The plot on right shows ΔR between the b-quark daughters - for large H p_T the daughters become collimated.

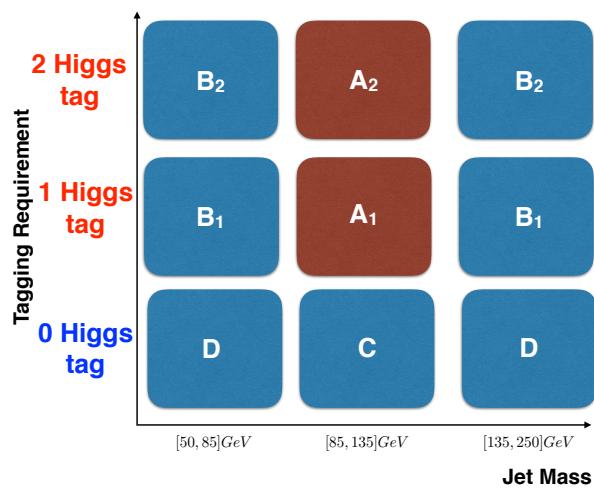


Figure 6.4: A diagram of the partitioned phase space. An additional binning in p_T^{miss} brings the number of analysis bins to $6 \times 3 = 18$.

$$A_{1,2}^{\text{predicted}} = (B_{1,2} \cdot \frac{C}{D})^{\text{observed}} \quad (6.2)$$

The expected p_T^{miss} distribution from simulation is seen in the stacked histograms of Figure 6.5. The prediction using the ABCD method is seen in the red hash. The expected level of closure in simulation can be determined by dividing the true content in the signal region with the prediction. This ratio, denoted κ , is seen in the bottom panel of Figure 6.5. $\kappa = 1$ represents perfect closure. As will be discussed in Section 6.5.2, κ is used as a correction in the background estimation procedure.

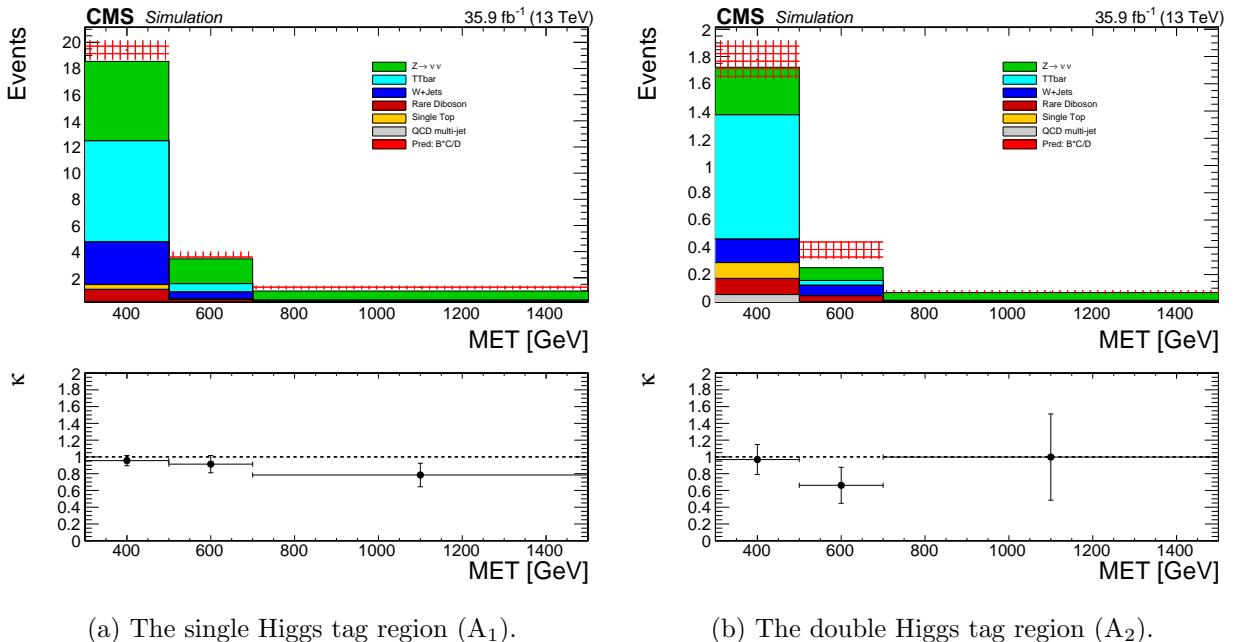


Figure 6.5: p_T^{miss} distributions and predictions in the signal regions using simulation only.

6.5.1 Control Regions within Data

We treat the SM backgrounds in the analysis as consisting of three main components (hinted at the list of MC samples seen in Table 6.1):

- $Z \rightarrow \nu\bar{\nu}$ in which the invisible Z gives true p_T^{miss} ('Z-invisible').
- Semi-leptonic W or t production in which the lepton is not identified, the associated neutrino from the leptonic decay creates true p_T^{miss} ('lost-lepton').

- Jet production via QCD in which the p_T of a jet is substantially under-measured, this creates a fake source of p_T^{miss} .

Three control regions are defined to serve as a proxy for each source. They are defined with the same event selection as the signal and control regions, with the exception of the inversion of a single cut.

- A control region with a single-photon, after artificially removing the photon from event reconstruction, closely mimics the Z-invisible background. For high- p_T , both photons and Z bosons become massless, neutral, gauge bosons whose kinematics are expected to be similar.
- A control region with a single-lepton, mimics the lost-lepton background.
- A control region defined by the logical inversion of the low- $\Delta\phi$ cuts, most closely mimics the QCD background. This enriches our events with those likely in which p_T^{miss} is aligned with an under-measured AK4 jet.

As they are orthogonal to the analysis region, we are able to test the closure of the background estimation technique independently within each of the three control regions. By comparing the prediction of the SM yields (using the ABCD method) with those observed, the validity of the technique can be verified for that particular background category. The comparisons for the single-photon, single-lepton and low- $\Delta\phi$ control regions can be seen in Figure 6.6. κ in the bottom panel is defined as the ratio of the true event yield to the prediction. $\kappa = 1$ represents the case of perfect closure. These comparisons are used for commissioning of the background estimation technique only.

6.5.2 κ as a Correction to the Estimation

To correct for the observed non-closure of the background estimation method a correction κ is applied to the SM background prediction. κ is obtained by dividing the MC yields for the signal

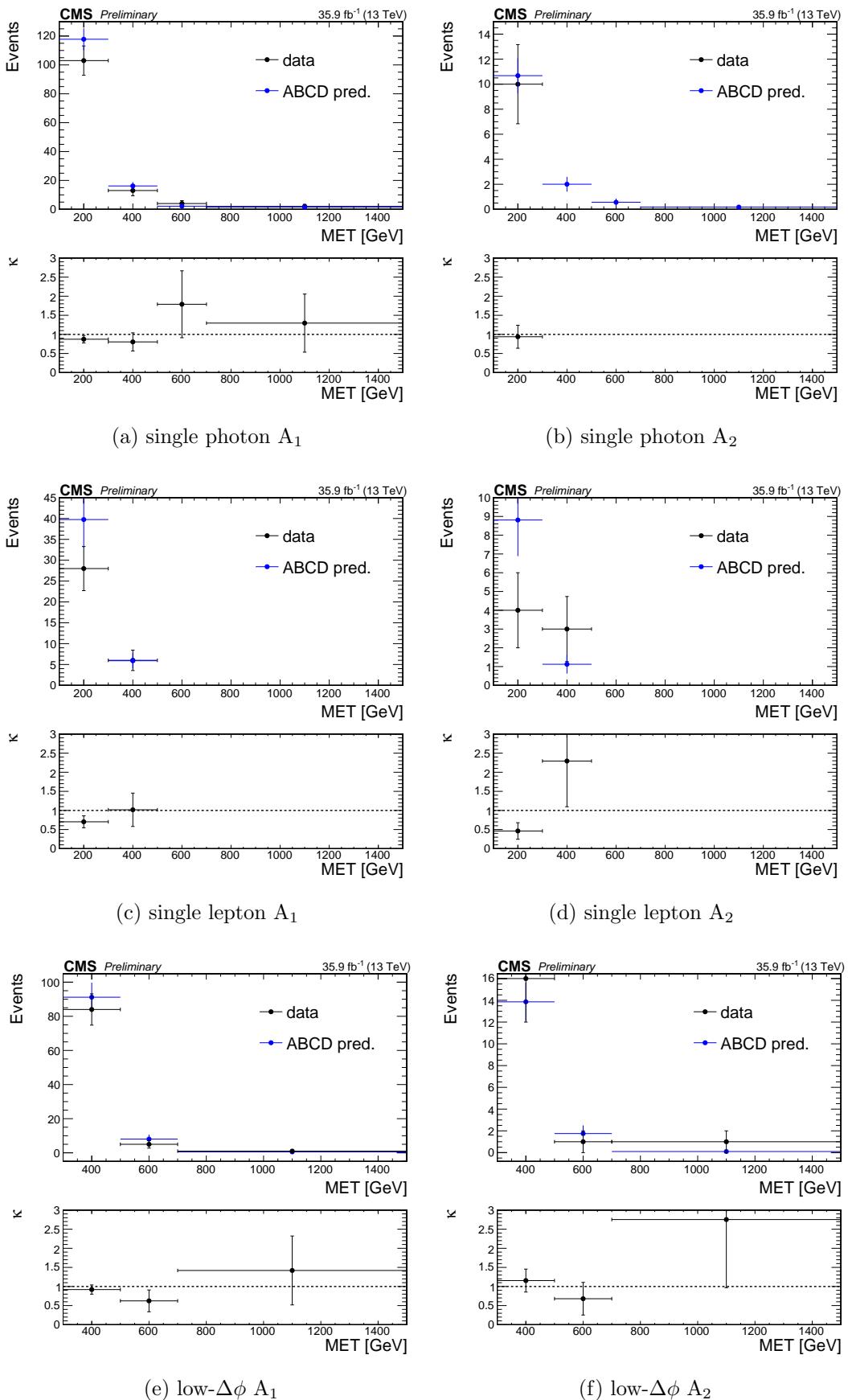


Figure 6.6: Closure observed within the data control regions.

region by that predicted:

$$\kappa \equiv A^{mc} / \left(B \cdot \frac{C}{D} \right)^{mc} \quad (6.3)$$

There are $2 \times 3 = 6$ values of κ , one for each signal bin. $\kappa = 1$ represents the case of perfect closure. The corrections are then applied as follows:

$$A_{1,2}^{\text{predicted}} = \kappa \cdot \left(B_{1,2} \cdot \frac{C}{D} \right)^{\text{observed}} \quad (6.4)$$

These values of κ are those which we have already seen in Figure 6.5.

The value of κ is dependent on the yields of each analysis bin and is therefore sensitive to the accuracy of the modeling of MC in each of the 18 analysis bins. To improve the determination of κ , scale factors are derived using the data control regions to correct the normalization of MC in each of these bins. Different scale factors are assigned separately to the Z-invisible, lost-lepton, and QCD background MC samples.

First consider how the yield N predicted by MC in an arbitrary bin (of 18) is the sum of the yields in the different MC datasets ($t\bar{t}$ and $W \rightarrow \ell\nu$ are grouped as they together represent the lost-lepton background):

$$N^{mc} = N_{Z \rightarrow \nu\bar{\nu}}^{mc} + N_{t\bar{t}, W \rightarrow \ell\nu}^{mc} + N_{QCD}^{mc} \quad (6.5)$$

Scale factors are defined for this bin using the corresponding control regions in data and forming the ratio of events in simulation to that observed. They are then applied as follows:

$$N_{\text{corrected}}^{mc} = \left(\frac{N_{\text{single}-\gamma}^{\text{data}}}{N_{\text{single}-\gamma}^{mc}} \right) \cdot N_{Z \rightarrow \nu\bar{\nu}}^{mc} + \left(\frac{N_{\text{single}-\ell}^{\text{data}}}{N_{\text{single}-\ell}^{mc}} \right) \cdot N_{t\bar{t}, W \rightarrow \ell\nu}^{mc} + \left(\frac{N_{\text{low}-\Delta\phi}^{\text{data}}}{N_{\text{low}-\Delta\phi}^{mc}} \right) \cdot N_{QCD}^{mc} \quad (6.6)$$

The p_T^{miss} distribution within the control regions is shown for both data and MC in Figures 6.7, 6.8, 6.9 for the single photon, single lepton, and low- $\Delta\phi$ control regions, respectively. The ratio in the bottom panel of each plot represents the scale factor for that p_T^{miss} bin. The dotted

horizontal line shows the average scale factor inclusive in p_T^{miss} . The scale factors for the single-photon and low- $\Delta\phi$ control regions show no p_T^{miss} dependence and are determined integrated over $p_T^{\text{miss}} > 300 \text{ GeV}$. The values of the scale factors are summarized in Table 6.2. The Single-lepton region, shown in Figure 6.8, does show p_T^{miss} dependence and are summarized in Tables 6.2 and 6.3. In order to improve statistics for the single-lepton region, the low- $\Delta\phi$ requirement has been removed.

Table 6.2: Summary of the control region scale-factors integrated over p_T^{miss} .

Low $\Delta\phi$					
A_{SF}^{1H}	A_{SF}^{2H}	C_{SF}	B_{SF}^1	B_{SF}^2	D_{SF}
1.1 ± 0.33	0.85 ± 0.12	0.93 ± 0.1	0.88 ± 0.04	1.2 ± 0.16	0.71 ± 0.027
Single Lepton					
A_{SF}^{1H}	A_{SF}^{2H}	C_{SF}	B_{SF}^1	B_{SF}^2	D_{SF}
0.61 ± 0.04	0.59 ± 0.08	p_T^{miss} dependent	0.59 ± 0.016	0.71 ± 0.04	p_T^{miss} dependent
Photon					
A_{SF}^{1H}	A_{SF}^{2H}	C_{SF}	B_{SF}^1	B_{SF}^2	D_{SF}
0.61 ± 0.088	0.75 ± 0.29	0.5 ± 0.07	0.98 ± 0.094	2.58 ± 0.63	0.71 ± 0.035

Table 6.3: Summary of the Single Lepton control region scale-factors the p_T^{miss} in the anti-tag sideband region.

Single Lepton C_{SF}		
p_T^{miss} [300, 500]	[500, 700]	[700, ∞]
0.47 ± 0.05	0.54 ± 0.15	0.18 ± 0.1
Single Lepton D_{SF}		
0.49 ± 0.02	0.40 ± 0.05	0.35 ± 0.08

The scale factors are then applied to the MC samples to give yields which better reflect data. The p_T^{miss} distributions for the signal regions and expectations from the ABCD background prediction are seen in Figure 6.10 (the data-corrected version of Figure 6.5). The improved value of κ is seen in the lower panel of each plot. The modified values of the MC yields in the signal region (seen in the calculation of κ) are seen in Tables 6.5 and 6.6. Since most of the scale factors are less than one the background decreases in Figure 6.5 relative to Figure 6.10 but still preserves closure so that κ is statistically compatible with unity. A distribution of κ is derived by throwing

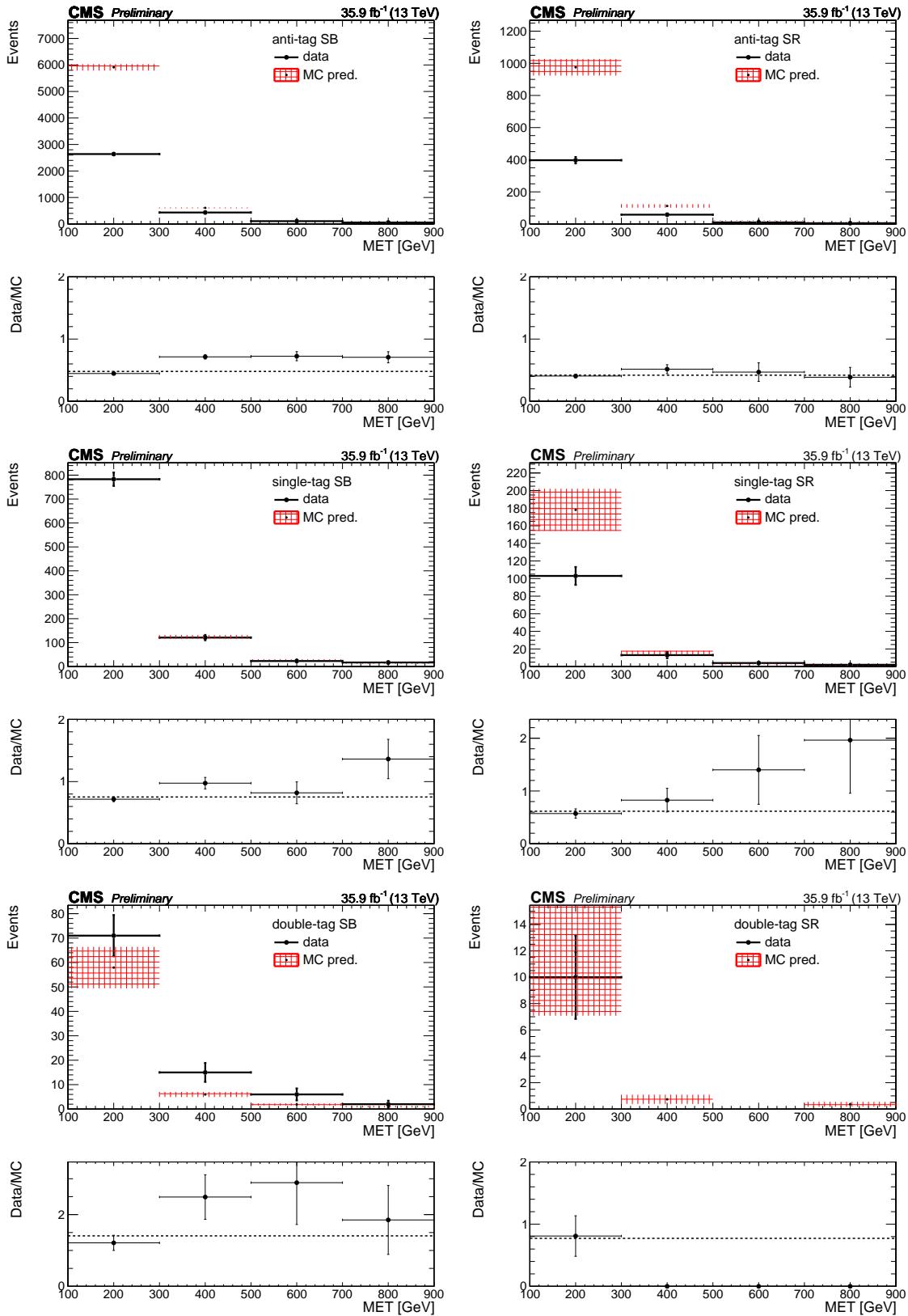


Figure 6.7: p_T^{miss} distribution in the signal and sideband bins within the single photon control region. The hashed red band denotes the prediction from simulation; the solid black points denote the observed yields in data. The Data/MC ratio in the lower panel of each plot represents the scale factor for that bin.

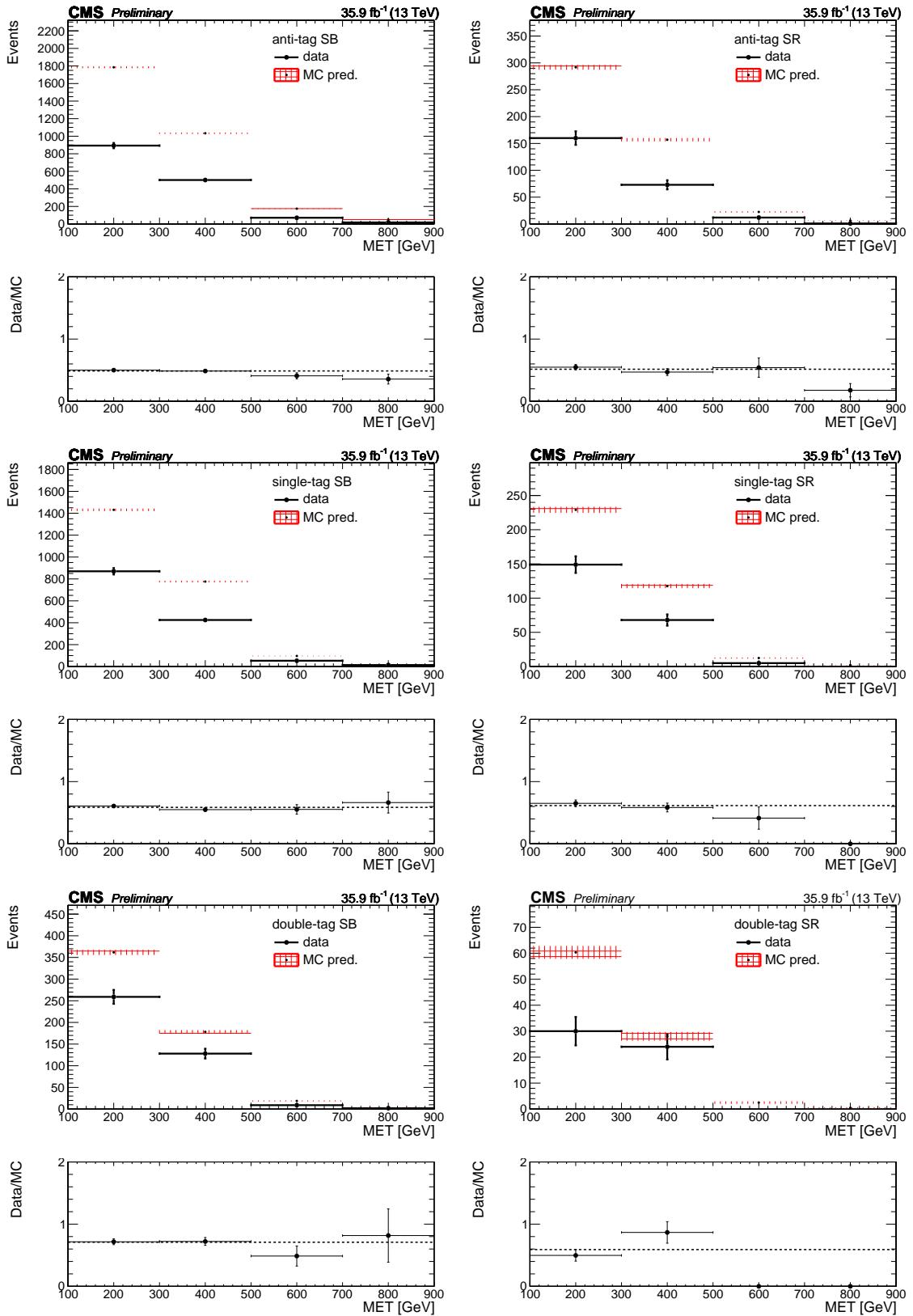


Figure 6.8: p_T^{miss} distribution in the signal and sideband bins within the single lepton control region. The hashed red band denotes the prediction from simulation; the solid black points denote the observed yields in data. The Data/MC ratio in the lower panel of each plot represents the scale factor for that bin. In order to improve statistics, the low- $\Delta\phi$ requirement has been removed.

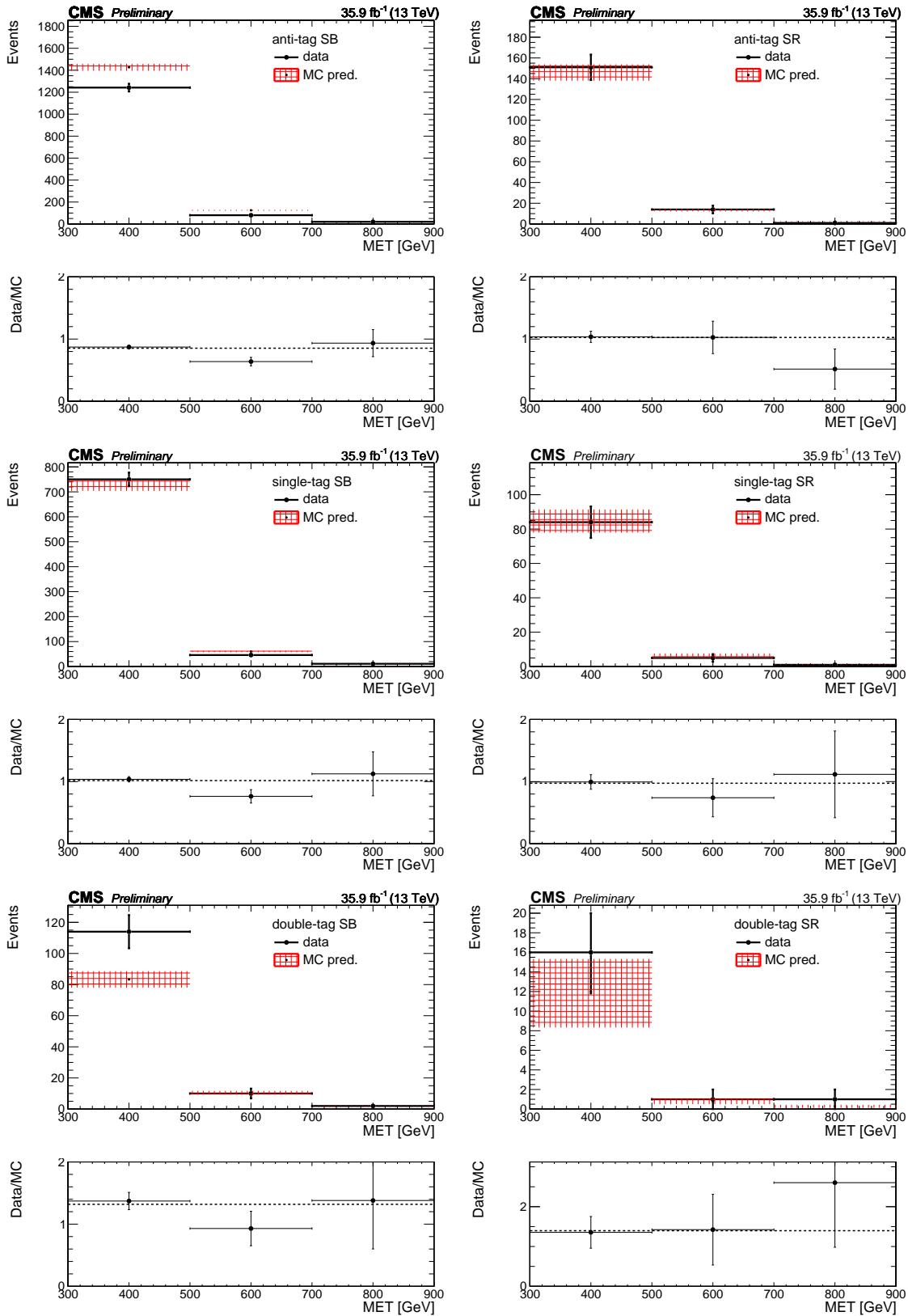
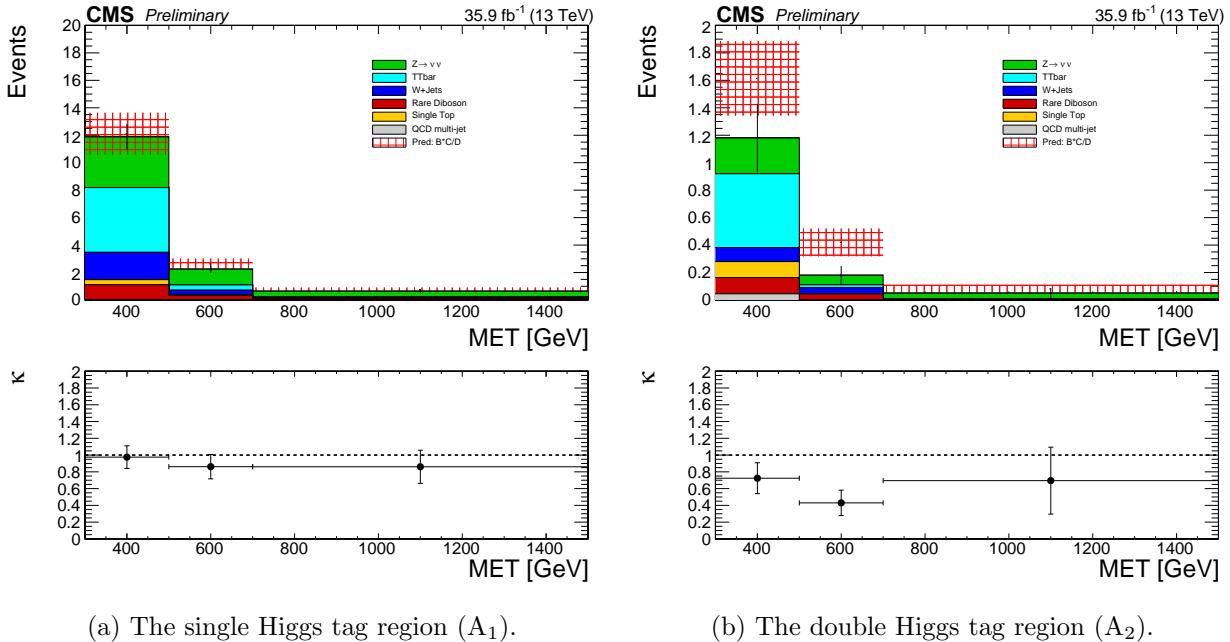


Figure 6.9: p_T^{miss} distribution in the signal and sideband bins within the low- $\Delta\phi$ control region. The hashed red band denotes the prediction from simulation; the solid black points denote the observed yields in data. The Data/MC ratio in the lower panel of each plot represents the scale factor for that bin.

gaussian toys for each of the scale factors, the final results being summarized in Table 6.4.

Table 6.4: The κ factor computed by throwing Gaussian toys for the scale factors.

	1-Higgs Tag	2-Higgs Tag
	κ	
p_T^{miss} [300, 500 GeV]	0.98 ± 0.11	0.73 ± 0.14
p_T^{miss} [500, 700 GeV]	0.86 ± 0.16	0.43 ± 0.12
p_T^{miss} [700, ∞ GeV]	0.86 ± 0.17	0.62 ± 0.30



(a) The single Higgs tag region (A_1).

(b) The double Higgs tag region (A_2).

Figure 6.10: p_T^{miss} distributions and the predictions in the signal regions using simulation with scale factors from data.

Table 6.5: Corrected MC Yields in the signal regions.

p_T^{miss}	Z+Jets	W+Jets	T#bar{T}bar	QCD	Rare	Total	Data
Region: A^{1H}							
[300, 500] GeV	3.76 ± 0.54	2.05 ± 0.30	4.87 ± 0.72	0 ± 0	1.48 ± 0.40	12.17 ± 1.03	15
[500, 700] GeV	1.18 ± 0.21	0.33 ± 0.08	0.40 ± 0.12	0 ± 0	0.39 ± 0.16	2.30 ± 0.29	2
> 700 GeV	0.43 ± 0.10	0.053 ± 0.025	0.046 ± 0.01	0 ± 0	0.13 ± 0.053	0.66 ± 0.12	1
Region: A^{2H}							
[300, 500] GeV	0.26 ± 0.12	0.11 ± 0.05	0.58 ± 0.21	0.045 ± 0.05	0.23 ± 0.12	1.24 ± 0.28	1
[500, 700] GeV	0.07 ± 0.045	0.049 ± 0.031	0.022 ± 0.0098	0 ± 0	0.045 ± 0.039	0.19 ± 0.068	0
> 700 GeV	0.044 ± 0.032	0.005 ± 0.005	0 ± 0	0 ± 0	0.002 ± 0.016	0.051 ± 0.036	0

Table 6.6: Corrected MC Yields in each of the signal and sideband regions.

p_T^{miss}	Z+Jets	W+Jets	T#bar{T}bar	QCD	Rare	Total	Data
Region: C							
[300, 500] GeV	17.66 ± 2.52	8.23 ± 1.44	3.87 ± 0.73	0.81 ± 0.49	2.78 ± 1.142	33.36 ± 3.24	44
[500, 700] GeV	5.20 ± 0.77	0.57 ± 0.27	0.22 ± 0.11	0 ± 0	0.63 ± 0.16	6.63 ± 0.84	12
> 700 GeV	2.48 ± 0.39	0.12 ± 0.12	0.028 ± 0.031	0 ± 0	0.14 ± 0.06	2.76 ± 0.41	4
Region: B^{1H}							
[300, 500] GeV	42.13 ± 4.17	15.61 ± 10.15	30.99 ± 20.15	1.57 ± 0.54	12.16 ± 1.37	102.47 ± 23.00	112
[500, 700] GeV	12.05 ± 1.28	2.74 ± 1.79	3.04 ± 2.00	0 ± 0	2.55 ± 0.43	20.37 ± 3.00	20
> 700 GeV	5.92 ± 0.69	0.67 ± 0.61	0.49 ± 0.46	0 ± 0	1.93 ± 0.72	9.01 ± 1.25	5
Region: B^{2H}							
[300, 500] GeV	5.51 ± 1.47	0.73 ± 0.44	4.46 ± 2.65	0.33 ± 0.23	2.06 ± 0.32	13.09 ± 3.09	13
[500, 700] GeV	1.80 ± 0.56	0.17 ± 0.11	0.59 ± 0.39	0 ± 0	0.62 ± 0.23	3.17 ± 0.73	1
> 700 GeV	0.67 ± 0.27	0.0084 ± 0.009	0.035 ± 0.031	0 ± 0	0.23 ± 0.073	0.94 ± 0.28	1
Region: D							
[300, 500] GeV	164.82 ± 8.31	61.24 ± 3.70	33.20 ± 2.16	8.50 ± 2.73	20.64 ± 1.78	288.41 ± 9.90	273
[500, 700] GeV	47.37 ± 2.52	6.36 ± 1.39	2.37 ± 0.55	0 ± 0	4.42 ± 1.46	60.51 ± 3.27	60
> 700 GeV	26.79 ± 1.50	0.99 ± 0.53	0.16 ± 0.086	0 ± 0	3.48 ± 1.01	31.42 ± 1.88	28

6.5.3 Sideband Yields & Predictions

Observed yields in data and the MC expectation for the 12 sideband bins (e.g. $B_{1,2}$, C, D in Figure 6.4) are seen in Figure 6.11. Table 6.7 lists these observed yields and calculated background prediction for the 6 signal bins.

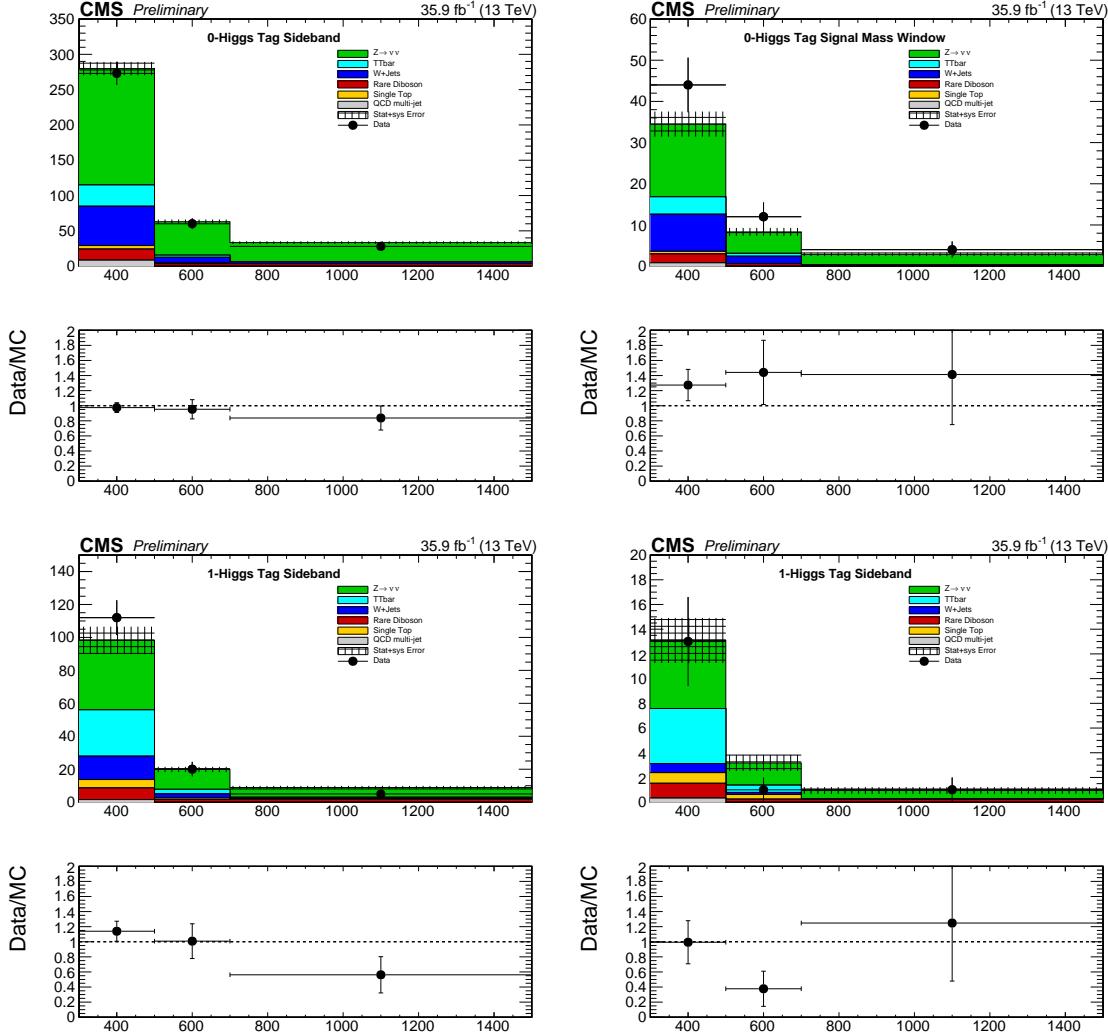


Figure 6.11: p_T^{miss} distribution in each of the control regions comparing data and the scale-factor corrected simulation. The hashed red distribution denote the prediction from simulation; the solid points denote the observed yields in data.

6.6 Signal Systematics

We consider a variety of systematic uncertainties on the signal efficiency and distribution. Some are common to more inclusive SUSY analyses [14] and there are additional systematics related to $b\bar{b}$ tagging efficiency and the effect of the pruned mass scale and resolution on the signal efficiency.

- **Luminosity:** The recommendation for the 2016 dataset is currently a flat uncertainty of 2.5%.
- **Isolated track veto:** A flat uncertainty of 2% is assigned to the signal samples to account for any data/MC differences based on the study from the 2015 analysis [14].
- **MC statistics:** The signal MC sample statistical uncertainty is generally 2-4% .
- **Trigger efficiency:** The effect of the uncertainty on the signal yield is about 2%.
- **Pileup reweighting:** The sensitivity to the pileup distribution was studied for various benchmark signal models by comparing events with $n_{\text{vtx}} < 20$ (low PU) or $n_{\text{vtx}} \geq 20$ (high PU). Accordingly, no pileup reweighting is applied to the signal MC samples and no associated uncertainty is assessed.
- **ISR:** An ISR correction is derived from $t\bar{t}$ events, with a selection requiring two leptons (electrons or muons) and two b-tagged jets, implying that any other jets in the event arise from ISR. The correction factors are 1.000, 0.920, 0.821, 0.715, 0.662, 0.561, 0.511 for NISR

Table 6.7: Observed yields in the sideband regions, κ , and background predictions for the 6 signal bins.

N_H	p_T^{miss} (GeV)	B	C	D	κ	$\kappa \cdot B \cdot C/D$
A_1	[300, 500 GeV]	112	44	273	0.98 ± 0.11	17.7 ± 3.8
A_1	[500, 700 GeV]	20	12	60	0.86 ± 0.16	3.4 ± 1.5
A_1	[700, ∞ GeV]	5	4	28	0.86 ± 0.17	0.61 ± 0.45
A_2	[300, 500 GeV]	13	44	273	0.73 ± 0.14	1.52 ± 0.57
A_2	[500, 700 GeV]	1	12	60	0.43 ± 0.12	0.09 ± 0.08
A_2	[700, ∞ GeV]	1	4	28	0.62 ± 0.30	$0.09^{+0.11}_{-0.09}$

$= 0, 1, 2, 3, 4, 5, 6+$. The corrections are applied to the simulated signal jet samples with an additional normalization factor, typically 1.15 (depending on the signal model), to ensure the overall cross section of the sample remains constant. The systematic uncertainty in these corrections is chosen to be half of the deviation from unity for each correction factor. The effect on the yield ranges from 0.01%, with the largest effect at high MET.

- **Scales:** The uncertainty is calculated using the envelope of the weights from varying the renormalization and factorization scales, μ_R and μ_F , by a factor of 2 [15, 16]. The effect on the yield of is less than 0.1%.
- **Jet Energy Corrections:** The jet energy corrections (JECs) are varied using the p_T - and η -dependent jet energy scale uncertainties from the official database. These variations are propagated into the various jet-dependent variables, including: HT, MET, $\Delta\phi(\text{MET}, j_i)$. The overall effect is less than 1%.
- **Jet Energy Resolution:** The jet momenta in the MC samples are smeared to match the jet energy resolution in data. The smearing factors are varied according to the uncertainties on the jet energy resolution measurements. These variations are propagated into the various jet-dependent variables, including: HT, MET, $\Delta\phi(\text{MET}, j_i)$. The overall effect ranges from 0.01%.
- **PDFs:** The LHC4PDF prescription for the uncertainty on the total cross section is included as ± 1 sigma bands in the results plots. No additional uncertainty is considered for the uncertainty in the acceptance due to PDFs, as per SUSY group recommendation.

The above signal systematics are applied as an uncertainty on the signal normalization. These uncertainties are in general small. The main signal systematics come from the AK8 Jet Double-b tagging efficiency data/MC scale factors and the uncertainty on the pruned mass resolution. The AK8 Jet Double-b tagging efficiency has an uncertainty which is propagated to the signal efficiency. This uncertainty is applied as a shape uncertainty across the Higgs tag regions and the anti-tag re-

gion. Also the pruned jet mass scale and resolution uncertainties are propagated to the final signal efficiency using POG recommendations. The pruned mass scale factor is derived using W-jets in semi-leptonic $t\bar{t}$ and extrapolating to the H mass. This uncertainty is assigned a shape uncertainty on the signal mass window and the sideband.

- A data/MC scale-factor is derived from double-muon tag data selected with HLT Trigger `HLT_BTagMu_AK8Jet300_Mu5_v` and muon enriched QCD Monte-Carlo. The scale factors have mainly a statistical error along with a smaller set of systematic errors due to shape systematics, Jet-Energy scale uncertainty, Pile-up corrections, uncertainty on the number of tracks, uncertainty of b-fragmentation and c-fragmentation, and the uncertainty on K_s and Λ fraction.
- The pruned mass scale-factor is derived by comparing the efficiency to select W-jets in data and MC within a mass window of [65, 85] GeV. The fit for the gaussian resolution of the W-mass peak is shown in Figure 6.12 and the fit results are shown in Table 6.8. The mass scale between MC and data is consistent though MC predicts a narrower mass resolution compared to data. The jet mass in each event is smeared to mimic the pruned jet mass resolution in data and an uncertainty is assigned based on the ratio of efficiencies between the smeared and un-smeared cases. [17]

The summary of the signal systematics and their effect on the signal yields is shown in Table 6.9. The dominant effect is from the mass resolution uncertainty.

Data	
Mean	78.2 ± 0.46
Sigma	10.10 ± 0.671
$t\bar{t}$ MC	
Mean	78.4 ± 0.35
Sigma	7.23 ± 0.48

Table 6.8: Fit results for W-mass resolution in data and MC

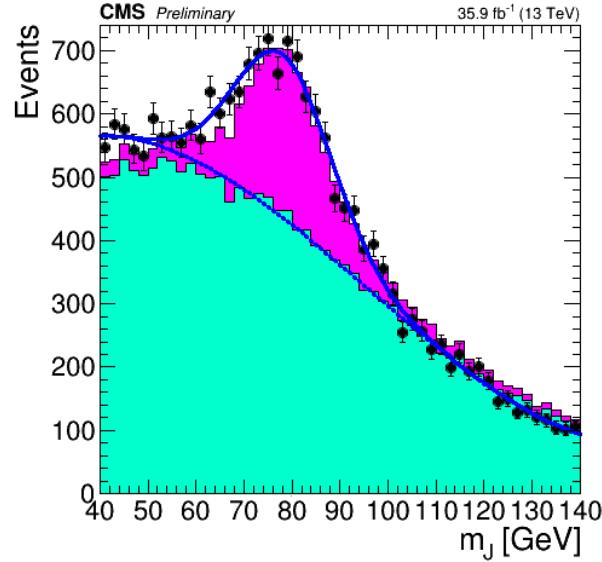


Figure 6.12: Pruned jet mass in semi-leptonic $t\bar{t}$ events. The mass peak for the W-jets is used to derive the mass resolution uncertainty.

Unc. on Normalization	
Systematic	% Effect on yields
Luminosity	2.6%
Trigger Eff.	2.0%
Iso. Track Veto	2%
ISR modeling	0.01%
PDF Scale	0.1%
JEC	1%
JER	0.01%
MC Stat	1-4%
Shape Unc.	
Double-b SF	6%
Mass Resolution	1-15%

Table 6.9: Summary of signal shape and normalization uncertainties.

6.7 Yields in the Signal Regions and Exclusion Curves

After unblinding the 4x3=12 sideband regions and performing the background estimation the 2x3=6 signal regions were unblinded. The observed yields, along with the SM background predictions, are seen in Table 6.10. Our signal region yields are consistent with the SM background expectation. Additionally, Table 6.10 shows the expected signal yields for two model points corresponding to gluino \tilde{g} masses of 2000 or 1800 GeV; the mass of the neutralino $\tilde{\chi}_1^0$ is fixed at 1 GeV; the mass splitting between the gluino \tilde{g} and neutralino $\tilde{\chi}_2^0$ is fixed at 50 GeV.

Table 6.10: Signal yields and SM background predictions

p_T^{miss}	$B \cdot C/D$	κ	$\kappa \cdot B \cdot C/D$	Obs.	T5HH(2000)	T5HZ(1800)
1-Higgs Tag						
[300, 500 GeV]	18.05 ± 3.39	0.98 ± 0.11	17.68 ± 3.85	15	0.24	0.75
[500, 700 GeV]	4 ± 1.54	0.86 ± 0.16	3.44 ± 1.47	2	0.32	0.98
[700, ∞ GeV]	0.71 ± 0.50	0.86 ± 0.17	0.61 ± 0.45	1	2.13	4.34
2-Higgs Tag						
[300, 500 GeV]	2.09 ± 0.67	0.73 ± 0.14	1.52 ± 0.57	1	0.17	0.35
[500, 700 GeV]	0.2 ± 0.20	0.43 ± 0.12	$0.09^{+0.08}_{-0.08}$	0	0.23	0.44
[700, ∞ GeV]	0.14 ± 0.16	0.62 ± 0.30	$0.09^{+0.11}_{-0.09}$	0	1.36	1.98

A visual representation of the one event in the double-H tagged signal bin is seen in Figure 6.13. The purple line represents p_T^{miss} . The three yellow cones represent the AK8 jets (there is one additional low- p_T AK8 jet in the event).

Interpreting our results in the context of the T5HH or T5ZH models, the absence of signal allows us to place lower limits on the mass of the gluino \tilde{g} . For the statistical treatment, we use the Higgs combination tool to encode the ABCD approach in the likelihood. In this approach, the data card for one search bin contains the observed number of events and the expected signal and background in each of the ABCD regions. A likelihood function is built that contains these ABCD regions and explicitly encodes the relation $A = \kappa B \frac{C}{D}$ and a Gaussian nuisance is assigned for the uncertainty on κ . The likelihood for each search bin can be described by:

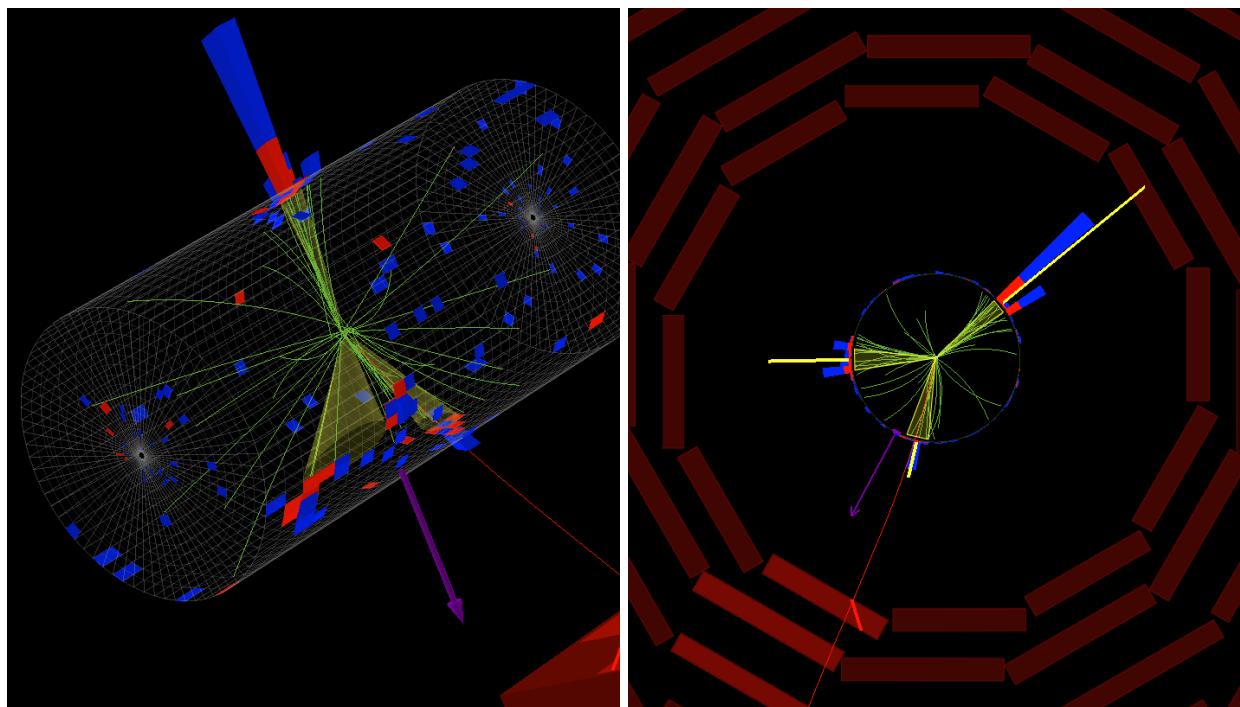


Figure 6.13: The single event in the A_2 region.

$$\mathcal{L} = \prod_i^{ABCD} Poisson(n_i | bkg_i + r \cdot sig_i) \times \prod_j^{nuisances} Constraints(\theta_j, \hat{\theta}_j) \quad (6.7)$$

where the 4 regions are modeled by Poisson distribution and the term Constraints refers to either Gaussian distributions for the κ uncertainties or log normal distributions that model the signal systematics. The expected and observed limits are then calculated based on the asymptotic approximation of the profile likelihood ratio using the CLs criterion to place limits at the 95% confidence level. These exclusion curves are seen in Figure 6.14. We are able to place lower limits at 95% confidence level on the gluino \tilde{g} mass at 2010 and 1825 GeV for the T5HH and T5ZH models, respectively. The weaker limit for the T5ZH model is due to the smaller branching fraction of the Z boson to b-quarks and our choice of signal mass window not being optimal for Z reconstruction.

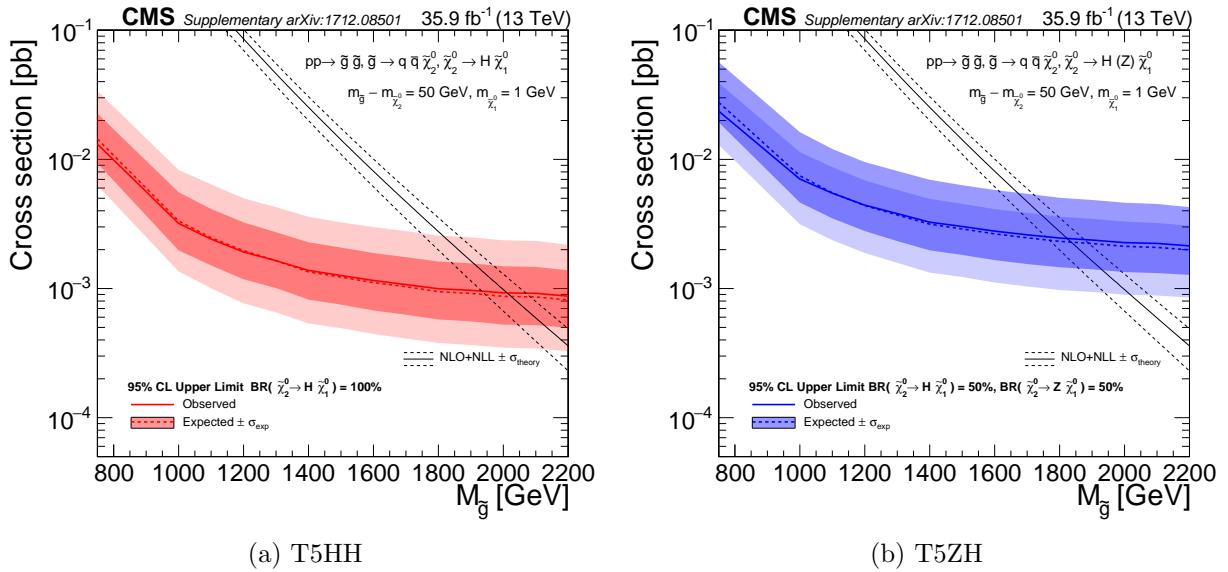


Figure 6.14: Observed and expected limits on the gluino cross section.

6.8 Reinterpretation

In Section 6.7 our results were presented in the context of limit-setting for the T5HH and T5ZH models. Many such SMS models exist within the MSSM which predict the production of high- p_T bosons and it is therefore important to include information necessary to make predictions of

yields for different final states. This is aided by providing the user with efficiencies for $b\bar{b}$ tagging and mass tagging of the AK8 jets. Tagging efficiencies for the five largest decay channels relevant to the analysis for the H boson are seen in 6.15. Tagging efficiencies for the hadronic decay modes of the Z boson are seen in Figure 6.16, the much lower mass tagging efficiency for the Z boson is due to our choice of signal mass window [85, 135 GeV] not being optimal for Z boson reconstruction. These are used to calculate the expected yields in the 6 analysis regions when performing a reinterpretation of the analysis using different final states.

For each event, the yield in each bin can be predicted by first forming the following weights using the tagging efficiencies for the leading two jets, as seen below. j_0 and j_1 represent the leading and subleading AK8 jet, respectively. The weights on the right-hand-side are p_T dependent.

- double mass tag weight = $j_{0\text{signalmass}} \cdot j_{1\text{signalmass}}$
- anti mass tag weight = $(j_{0\text{sidebandmass}} \cdot j_{1\text{signalmass}}) + (j_{1\text{signalmass}} \cdot j_{1\text{sidebandmass}}) + (j_{0\text{sidebandmass}} \cdot j_{1\text{sidebandmass}})$
- double bb tag weight = $j_{0\text{bbtag}} \cdot j_{1\text{bbtag}}$
- single bb tag weight = $(j_{0\text{bbtag}} \cdot (1 - j_{1\text{bbtag}})) + ((1 - j_{0\text{bbtag}}) \cdot j_{1\text{bbtag}})$
- anti bb tag weight = $(1 - j_{0\text{bbtag}}) \cdot (1 - j_{1\text{bbtag}})$

These weights are then combined in the following manner to determine the yields across each of the 6 bins for a single event.

- A1 weight = (single bb tag weight) \cdot (double mass tag weight)
- A2 weight = (double bb tag weight) \cdot (double mass tag weight)
- B1 weight = (single bb tag weight) \cdot (anti mass tag weight)
- B2 weight = (double bb tag weight) \cdot (anti mass tag weight)
- C weight = (anti bb tag weight) \cdot (double mass tag weight)

- D weight = (anti bb tag weight) · (anti mass tag weight)

These weights for the 6 analysis bins (inclusive in p_T^{miss}) are then summed over all events to get the expected yields. Following this prescription, the authors performed this prediction using the T5HH model with a gluino mass of 2200 GeV and compared to the true value. The largest deficit was in the D region, with a difference of -36% difference from nominal. The greatest over-prediction is found in the B2 region, with a surplus of +8.2% events relative to nominal. The closure in the other bins fall somewhere in this range. These results are summarized in Table 6.11. As a further cross-check to the yield estimates, Table 6.12 shows the true signal event efficiencies for the T5HH model with a gluino mass of 2200 GeV.

Table 6.11: Comparison of the true reco-level event yield with those obtained via the prediction following the prescription above. The columns with the RECO or GEN labels are the prediction using RECO or GEN event variables only, respectively. The prediction was made using the T5HH MC with a gluino mass of 2200 GeV.

	RECO "truth"	RECO prediction	GEN prediction
Baseline	4.08	3.46 (-15%)	3.53 (-16%)
A1	1.21	1.18 (-2.3%)	1.26 (+3.6%)
A2	0.777	0.748 (-3.7%)	0.815 (+4.7%)
B1	0.802	0.703 (-12%)	0.664 (-21%)
B2	0.322	0.338 (+5.0%)	0.350 (+8.2%)
C	0.498	0.473 (-4.9%)	0.487 (-2.1%)
D	0.478	0.353 (25%)	0.308 (-36%)

Table 6.12: Signal efficiencies for an event to land in a given analysis bin. The efficiencies were derived using the T5HH MC with a gluino mass of 2200 GeV. Choosing a gluino mass of 1800 GeV decreases the efficiencies by a relative 5%.

	Baseline	A1	A2	B1	B2	C	D
$p_T^{\text{miss}} < 300 \text{ GeV}$	32%	9.4%	6.0%	6.2%	2.5%	3.9%	3.7%
$300 < p_T^{\text{miss}} < 500 \text{ GeV}$	2.7%	0.78%	0.52%	0.54%	0.25%	0.31%	0.30%
$500 < p_T^{\text{miss}} < 700 \text{ GeV}$	3.5%	1.0%	0.65%	0.72%	0.28%	0.43%	0.40%
$p_T^{\text{miss}} > 700 \text{ GeV}$	26%	7.6%	4.9%	5.0%	2.0%	3.1%	3.0%

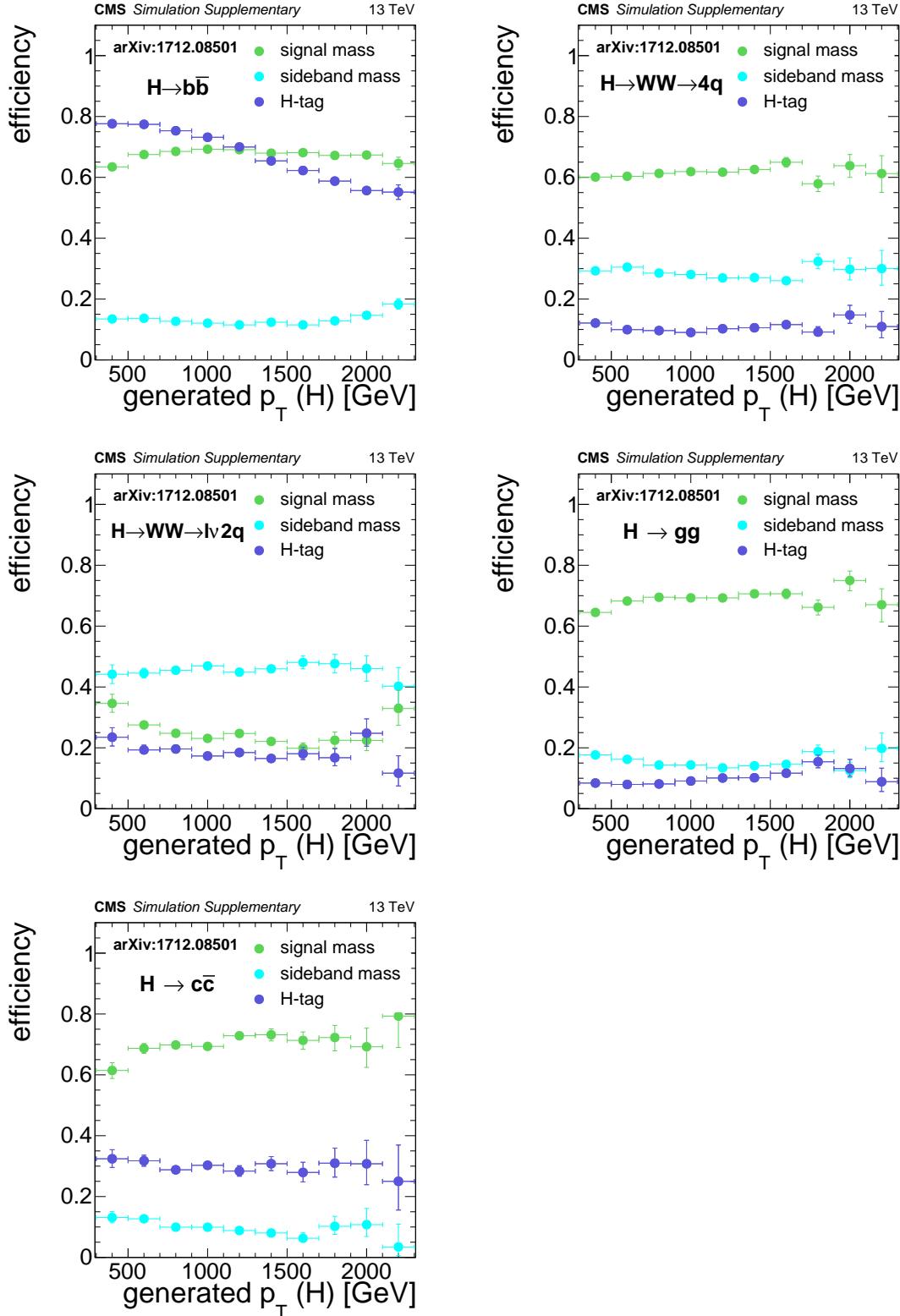


Figure 6.15: Efficiencies for an AK8 jet originating from Higgs boson decay, relative to baseline selection. The "signal mass" curve represents the probability the jet will fall within the mass region [85, 135 GeV]. The "sideband mass" curve represents the probability the jet will fall within the mass region [50, 85 GeV] or [135, 250 GeV]. The "H-tag" curve represents the probability the jet have a double-b discriminator value greater than 0.3, for jets within the mass region [50, 250 GeV]. The efficiencies were derived using the T5ZH MC with a gluino mass of 2200 GeV.

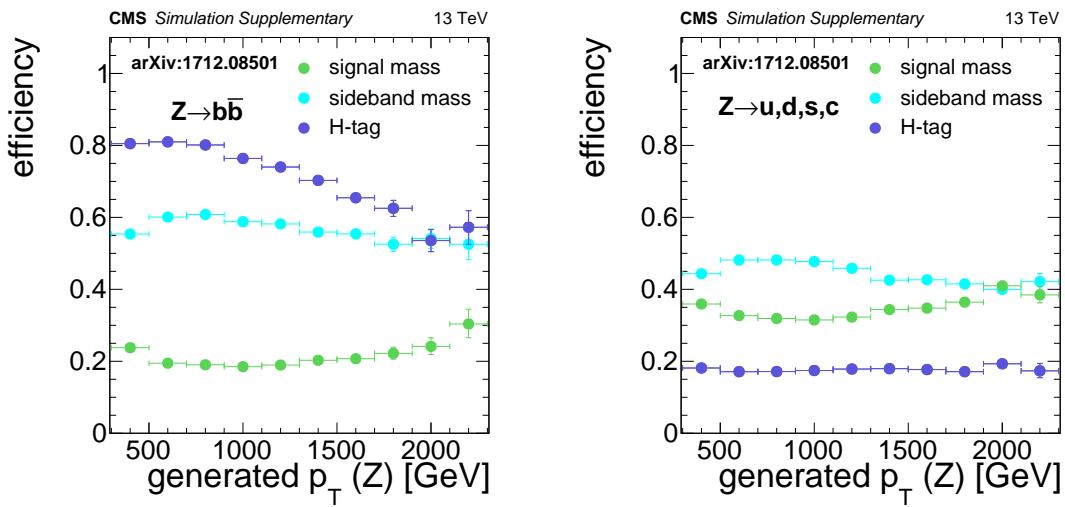


Figure 6.16: Efficiencies for an AK8 jet originating from Z boson decay, relative to baseline selection. The "signal mass" curve represents the probability the jet will fall within the mass region [85, 135 GeV]. The "sideband mass" curve represents the probability the jet will fall within the mass region [50, 85 GeV] or [135, 250 GeV]. The "H-tag" curve represents the probability the jet have a double-b discriminator value greater than 0.3, for jets within the mass region [50, 250 GeV]. The efficiencies were derived using the T5ZH MC with a gluino mass of 2200 GeV.

6.9 Conclusions

A search for physics beyond the SM was presented using events with boosted H bosons and missing transverse energy (p_T^{miss}). The search targeted events with two or more wide-angle jets (AK8) being consistent with the decay of a boosted H or Z boson decaying to $b\bar{b}$. p_T^{miss} could potentially arise in the case a supersymmetric particle escapes detection. An ABCD method uses a sideband region to predict the the SM background in our signal region. Events are categorized according to the $b\bar{b}$ and mass tagging of the leading two AK8 jets in the event. The observed yields in the 6 signal bins are statistically compatible with the SM background expectation and no excess of events is observed. We use these results to set limits on the gluino mass for the SUSY-inspired T5HH or T5ZH models. For the T5HH model we are able to exclude gluino masses below 2010 GeV at 95% confidence level. This is with the assumption the NLSP mass is 50 GeV less than the gluino mass and that the LSP has a mass of 1 GeV. The work presented here has been published in a peer-reviewed journal[18].

Chapter 7

Conclusions

A chapter describing the conclusions.

Bibliography

- [1] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,” *Phys. Lett. B*, vol. 716, no. 1, pp. 30 – 61, 2012.
- [2] CMS Collaboration, “The CMS trigger system,” *JINST*, vol. 12, p. P01020. 122 p, Sep 2016.
- [3] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm,” *Journal of High Energy Physics*, vol. 2008, no. 04, p. 063, 2008.
- [4] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector,” *JINST*, vol. 12, p. P10003, 2017.
- [5] CMS Collaboration, “Search for supersymmetry in multijet events with missing transverse momentum in proton-proton collisions at 13 TeV,” *Phys. Rev. D*, vol. 96, p. 032003, 2017.
- [6] CMS Collaboration, “Search for supersymmetry in the multijet and missing transverse momentum final state in pp collisions at 13 TeV,” 2016.
- [7] S. P. Martin, “A Supersymmetry Primer.” arXiv:hep-ph/9709356v7, Jan 2016.
- [8] CMS Collaboration, “Interpretation of searches for supersymmetry with simplified models,” *Phys. Rev. D*, vol. 88, p. 052017, 2013.
- [9] CMS Collaboration, “Search for higgsino pair production in pp collisions at $\sqrt{s} = 13$ TeV in final states with large missing transverse momentum and two Higgs bosons decaying via $H \rightarrow b\bar{b}$,” *Phys. Rev. D*, vol. 97, p. 032007, 2018.
- [10] CMS Collaboration, “Identification of double-b quark jets in boosted event topologies,” 2016. CMS-PAS-BTV-15-002.
- [11] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations,” *JHEP*, vol. 07, p. 079, 2014.
- [12] R. D. Ball *et al.*, “Parton distributions for the LHC Run II,” *JHEP*, vol. 04, p. 040, 2015.
- [13] S. Agostinelli *et al.*, “GEANT4—a simulation toolkit,” *Nucl. Instrum. Meth. A*, vol. 506, p. 250, 2003.

- [14] CMS Collaboration, “Search for supersymmetry in multijet events with missing transverse momentum in proton-proton collisions at 13 TeV,” CMS Physics Analysis Note AN-2016/350, CERN, 2017.
- [15] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi, “The $t\bar{t}$ cross-section at 1.8 TeV and 1.96 TeV: A Study of the systematics due to parton densities and scale dependence,” *JHEP*, vol. 04, p. 068, 2004.
- [16] S. Catani, D. de Florian, M. Grazzini, and P. Nason, “Soft gluon resummation for Higgs boson production at hadron colliders,” *JHEP*, vol. 07, p. 028, 2003.
- [17] I. Belotelov, I. Golutvin, D. Bourilkov, A. Lanyov, E. Rogalev, M. Savina, and S. Shmatov, “Identification and calibration of boosted hadronic w/z bosons at 13 tev,” CMS Note 2016/215, 2016.
- [18] CMS Collaboration, “Search for physics beyond the standard model in events with high-momentum Higgs bosons and missing transverse momentum in proton-proton collisions at 13 TeV,” *Phys. Rev. Lett.*, vol. 120, p. 241801, 2018.