New Physics Models in the Diphoton Final State at CMS

Thesis by

Ann Miao Wang

In Partial Fulfillment of the Requirements

for the Degree of

Bachelor of Science

California Institute of Technology

Pasadena, California

2015

(Submitted May 15, 2015)

Acknowledgments

First, I want to express my gratitude to Professor Maria Spiropulu for being the most inspirational leader, mentor, and scientist I could have ever hoped to work for. My deep interest in experimental high energy physics is a result of her passion and tireless enthusiasm for the subject.

I would also like to express my deep appreciation for Professor Harvey Newman for being a wonderful teacher and listener. His diligence and love for physics have painted him as a true role model in my life.

I am extremely grateful to Javier Duarte for always answering my questions and meeting with me for a countless number of hours over the past three years. His kindness and patience have made an indescribable impact on my research experience.

I would also like to Cristian Peña, Dustin Anderson, Alex Mott, and Maurizio Pierini, as well as the entire Caltech CMS group, for their invaluable guidance throughout this thesis.

These past four years at Caltech have truly been life-changing in completely surprising ways. One was my unexpected passion for studying physics, which I could not have anticipated. I owe that to the professors and students at this remarkable institution, especially Professor Steven Frautschi for instilling his enthusiasm in me through Ph 1 recitations. Without his extraordinary teaching, I would never have switched my major to physics.

Finally, I would like to express my gratitude to my parents and my sister for their love and contagious curiosity.

Contents

List of Figures

Chapter 1

Introduction

With the discovery of the Higgs boson as the final particle necessary to complete the standard model (SM) picture [1, 2], the field of particle physics is looking towards the search for new physics outside of the SM. One problem with the SM is that it does not include a viable particle candidate for dark matter [3]. At the Large Hadron Collider (LHC) in Geneva, Switzerland, scientists are looking for beyond-the-standard-model (BSM) events resulting from proton-proton collisions. One possible BSM physics scenario is weak-scale supersymmetry (SUSY) which provides a particle candidate for dark matter and solves the heirarchy problem [4].

A major difficulty in any search for new physics is to characterize what the signatures of the BSM physics would look like at the detector level. In the search for these events, there are many components of a physical experiment that must be studied and tuned. This thesis addresses two of the most important components: (1) trigger studies and (2) model kinematics. The overarching theme of this thesis is the use of the newly discovered Higgs particle as a probe for searching for BSM physics. The first half of the thesis focuses on studies related to the Compact Muon Solenoid (CMS) detector, while the second half of the thesis is more broadly applicable to other experiments.

Chapter 2 gives a short overview of the razor variables, which have been used to distinguish SUSY from SM background, producing strong limits on SUSY models [5]. The thesis will then employ these razor variables in the design of the custom hardware and software algorithms for deciding whether to record a collision event, and will investigate the distribution of these variables for certain SUSY simplified models.

Chapter 3 focuses on trigger design for BSM searches. It first tackles the study of the software-level trigger (known as the High Level Trigger) paths using the razor variables. Then, it includes a proposal of a new trigger path targeting BSM events that include a Higgs boson decaying to a $b\bar{b}$ pair in the final state. Chapter 4 is motivated by a previous search for supersymmetry with a Higgs decaying to two photons [6]. It discusses the kinematics of Higgs-aware BSM models, with particular focus on bottom squark production. This thesis attempts to fully characterize the range of signatures that bottom squark production would produce.

Chapter 2

The Razor Variables

The razor variables, M_R and R^2, are variables that are useful in searches for supersymmetry (SUSY). They have been successfully employed in previous searches with the CMS Collaboration to produce strong limits on SUSY parameter space [7]. These variables are employed to help distinguish SUSY signal events from the Standard Model (SM) background.

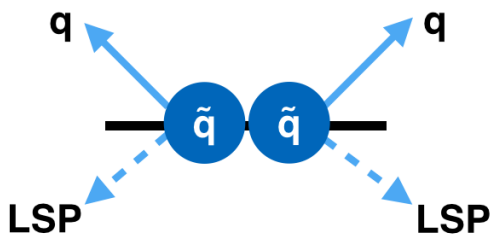


Figure 2.1: A cartoon of a disquark production scenario.

If we consider the scenario where two squarks (supersymmetric partners of quarks) each decay to a quark and a stable, weakly interacting lightest supersymmetric particle (LSP), the missing energy carried away by the LSP makes it difficult to reconstruct the kinematic attributes of all the particles in the decay. Thus, it is a challenge to identify this as a SUSY decay. For an all-hadronic analysis, we form the two razor variables M_R, R^2 as follows: first, we cluster the hadronic jets into two hemispheres using an algorithm in which we minimize the sum in quadrature of the invariant masses of the two hemispheres. These jets often have some baseline thresholds for certain kinematic events, such as transverse momentum (p_T) or pseudorapidity (η). Then, we define M_R as

$$M_R = \sqrt{(|p^{j1}| + |p^{j2}|)^2 - (p_z^{j1} + p_z^{j2})^2},$$

where p_{ji} is the momentum of the ith hemisphere. R is defined as

$$R = M_T^R / M_R,$$

where

$$M_T^R = \sqrt{\frac{\not{E}_T(p_T^{j1} + p_T^{j2}) - \vec{\not{E}}_T \cdot (\vec{p}_T^{j1} + \vec{p}_T^{j2})}{2}},$$

and $\not{E}_T$ is the missing transverse energy of the event.

M_R is contains information about the characteristic SUSY mass scales. It is an estimator, in the case of the squark, for M_Δ, which we define to be

$$M_\Delta = \frac{m_{\tilde{q}}^2 - m_{\tilde{\chi}}^2}{m_{\tilde{q}}}.$$

$m_{\tilde{q}}$ is the mass of the squark, and $m_{\tilde{\chi}}$ is the mass of the LSP. Experimentally, M_R peaks higher for SUSY events. R^2 contains information about the missing energy in the event and helps reduce QCD backgrounds [7]. Because of these advantages, the razor variables have been implemented successfully for new physics searches.

In the following chapters, we will use these razor variables extensively.

Chapter 3

CMS Trigger Studies

3.1 The CMS Detector and the Trigger System

The CMS detector is a general purpose detector at the LHC with various sub-detector components. It includes a lead-tungstate-crystal electromagnetic calorimeter (ECAL), a hadronic calorimeter (HCAL), a silicon tracker, and a muon drift tube system, cathode strip chambers, and resistive plate chambers [8]. The next run of the LHC, scheduled to begin in 2015, requires both software and hardware improvements due to higher energies (13 and 14 TeV) of the proton-proton collisions as well as increased underlying event activity (known as pileup) [9]. The events detected by CMS are too large in number to record all of the data, so a trigger system is required to decide which collision events to record.

This trigger system is composed of two levels; the system results in data reduction from the LHC collision rate of 15 MHz down to a rate of 400 Hz for Run 1 (and 1 kHz for Run 2). The two levels are the hardware-based Level-1 (L1) system and the software-based High Level Trigger (HLT). In order to minimize CPU time, particle and event kinematics are only partially reconstructed since triggering is conducted in real-time [10, 11].

The L1 system has three main components: global, muon, and calorimeter. The L1 calorimeter system is based on a set of trigger towers, which are made up of 4×4 regions. These regions are each 20° in ϕ (azimuthal angle), and are also sectioned by η (pseudorapidity). Jets, which are the result of quark and gluon production in the detector, are calorimeter objects at the L1 level.

One or more L1 filters can feed into an HLT filter. For the HLT system, muons are reconstructed using isolation algorithms and track fitting, which involves matching inner tracker tracks with outer muon spectrometer tracks. There are two possible collections of jets and $\not{E}_T$ at the HLT: calorimeter collections and Particle Flow (PF) collections. Calorimeter objets are crude objects reconstructed from energy deposits in the calorimeters, while PF objects are more sophisticated and rely on the PF algorithm [12]. The PF algorithm takes information from the multiple sub-detectors to produce PF candidates, resulting in tracking and better ECAL measurements to be incorporated into the

particle information. These particles are then clustered using the anti-kT jet clustering algorithm to reconstruct the jets [13, 14].

The selected kinematic thresholds for both the L1 and HLT must be thoroughly tested to (1) ensure sufficient data reduction and to (2) maintain the efficiency of keeping events which are interesting for our physics analyses. We will focus on the HLT in the following section.

3.2 Razor High Level Triggers

A set of trigger options involving razor variables were implemented for the LHC Run I. Due to higher energies, increased pileup, and shorter bunch spacing, these triggers need to be modified for Run II. In addition, PF objects are now available at the HLT, so the option of using PF objects instead of calorimeter objects must be discussed.

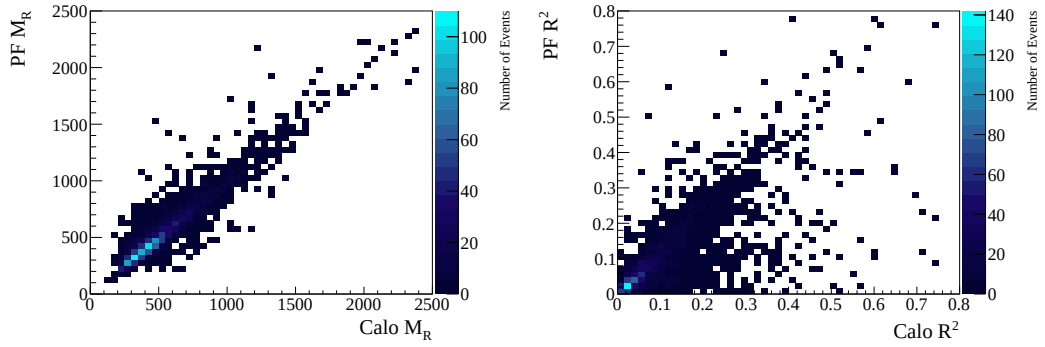
Although PF objects have been shown to have a much improved accuracy [15], they are more costly in CPU time at the HLT. As a result, one tactic of using PF objects but still minimizing CPU time is to reduce the number of events that the PF algorithm acts on by imposing loose thresholds on calorimeter-level objects.

One important detail before implementing this tactic is to study the correlation of PF objects to calorimeter objects.

3.2.1 Calorimeter Objects versus PF Objects Investigation

Because calorimeter objects and PF objects are constructed differently, we conducted a study of PF versus calorimeter variables. We conducted this study with a sample of $t\bar{t}$+jets, common background to many SUSY signals. This particular sample forces $t\bar{t}$ to decay leptonically, and details are in Table A.3. The main variables we were concerned with were the razor variables since the HLT triggers include cutoffs based on these variables. Thus, we studied the razor variables constructed with PF objects versus calorimeter objects. Examining Figure 3.1, the R^2 correspondence is much worse than the M_R correspondence. In addition to using the jet object information, the R^2 variables are constructed using the $\not{E}_T$ information. We then investigated the correspondence between PF and calo $\not{E}_T$, shown in Figure 3.2.

When the HLT $\not{E}_T$ is calculated using energy deposit information from the hadronic and electromagnetic calorimeters [10], the information from the muon chambers is excluded. As a result, we suspected that the muon energies were not correctly incorporated into the calorimeter-level $\not{E}_T$. To investigate this, we replotted the distributions with the additional requirement that muons with $p_T >$ 30 GeV be vetoed. A remarkable improvement of the correspondence between PF and calorimeter $\not{E}_T$ and R^2 distributions was exhibited in Figure 3.3.



(a) (b)

Figure 3.1: Comparisons of the razor variables formed with PF objects versus calorimeter objects.

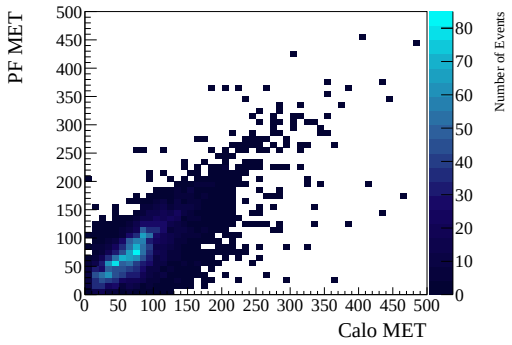
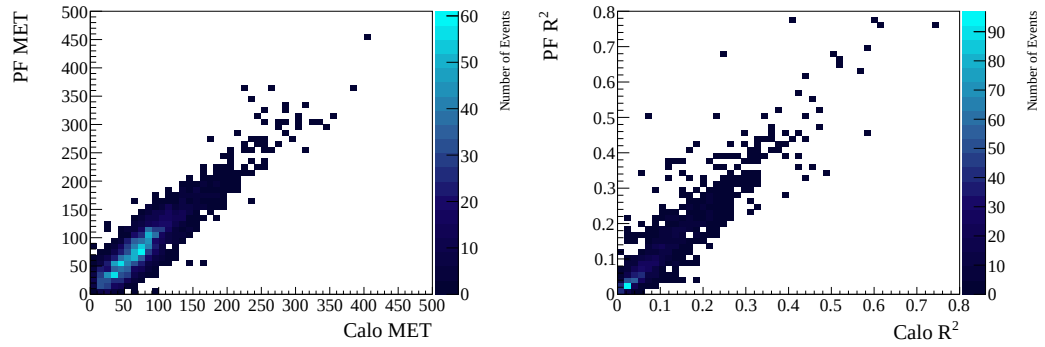


Figure 3.2: Comparisons of the $\not{E}_T$ distributions formed with PF objects versus calorimeter objects.



(a) (b)

Figure 3.3: Comparisons of the $\not{E}_T$ and R^2 distributions formed with PF objects versus calorimeter objects, vetoing all events that have muons with $p_T > 30$ GeV.

Following these studies, recommendations were made for muon information to be incorporated into the calorimeter level razor variables. Muon information was then added to the hemisphere decisions by other members of the Caltech CMS group (Javier Duarte).

3.2.2 Trigger Flow Path

For a trigger involving the razor variables, the trigger path is as follows:

1. Input from the Level 1 CMS Hardware Trigger. For these triggers, a combination of all hadronic triggers available were used.
2. Algorithms to create calorimeter objects, which use information from the calorimeter [13].
3. Thresholds implemented on the calorimeter jet kinematic properties.
4. Thresholds implemented on the razor variables constructed from calorimeter objects.
5. Algorithms to create Particle Flow (PF) objects, which use an increased amount of available information from CMS [13].
6. Thresholds implemented on the PF jet kinematic properties.
7. Thresholds implemented on the razor variables.

3.2.3 Rates of Trigger Menu

Various triggers were investigated to optimize the thresholds at the PF level and the calorimeter level. To calculate these rates, the triggers were evaluated within the CMS HLT software framework [16] over a weighted average of a set of QCD Monte Carlo samples (see Table A.1) with various jet p_T bins from 30 to 100 GeV, with 40 pile-up (pp collisions per bunch) and bunch-crossings happening every 25 ns. These samples essentially consisted of jets and pileup events.

The triggers are in part named according to CMS convention. A short code follows:

1. Rsq0pX indicates $R^2 > 0.X$.
2. MRX indicates $M_R >$ X.
3. RsqMRX indicates $(R^2 - 0.25) * (M_R - 300) >$ X. These offsets of 0.25, 300 approximate the $R^2 * M_R$ iso-probability contours observed in the CMS 8 TeV dataset [17].
4. NoRazorCaloCut indicates all calorimeter-level filters on the razor variables are removed.

Trigger Name	**Rate (Hz)**
HLT_RsqMR300_Rsq0p09_MR200	12.471 ± 1.244
HLT_RsqMR300_Rsq0p09_MR200_Muon	12.472 ± 1.244

Table 3.1: Table of the razor triggers discussed, where the rate is calculated for an instantaneous luminosity of 1.4e34 $\text{cm}^{-2}\text{s}^{-1}$. These all have a calorimeter-level filter of hltRsqMR200Rsq0p01MR100Calo. The second trigger has the muon sequence enabled.

Trigger Name	Rate (Hz)
HLT_RsqMR300_Rsq0p09_MR200_Muon_NoRazorCaloCut	12.967 ± 1.249
HLT_RsqMR300_Rsq0p09_MR200_Muon_CaloRsq0p0196	11.188 ± 1.145
HLT_RsqMR300_Rsq0p09_MR200_Muon_CaloRsq0p0289	10.761 ± 1.132

Table 3.2: Table of the razor triggers discussed, where the rate is calculated for an instantaneous luminosity of 1.4e34 $\text{cm}^{-2}\text{s}^{-1}$. These have a calorimeter-level filter of hltRsqMR240RsqXMR100Calo, where X = 0.0196 for the second trigger and X = 0.0289 for the third trigger. All triggers have the muon sequence enabled.

Trigger Name	Rate (Hz)
HLT_RsqMR300_Rsq0p09_MR200_Muon_CaloRsq0p04	9.513 ± 1.025
HLT_RsqMR300_Rsq0p09_MR200_Muon_CaloRsq0p0576	8.911 ± 1.012

Table 3.3: Table of the razor triggers discussed, where the rate is calculated for an instantaneous luminosity of 1.4e34 $\text{cm}^{-2}\text{s}^{-1}$. These have a calorimeter-level filter of hltRsqMR260RsqXMR100Calo, where X = 0.04 for the first trigger and X = 0.0576 for the second trigger. All triggers have the muon sequence enabled.

We first investigated the addition of the muon sequence mentioned in the previous section. We see that the rate increases marginally in Table 3.1. Therefore, adding the muon sequence to the calorimeter razor variables does not noticeably increase the rate, but would have a beneficial effect on accuracy for backgrounds and signals with a lot of muons, i.e. $t\bar{t}$+jets.

We then investigate the effect of the calorimeter-level razor variable thresholds on rate. Ideally, an increase in calorimeter-level razor variable thresholds would have a minimal effect on the rate when the thresholds are lower than the following PF thresholds. The PF thresholds implemented without any calorimeter-level thresholds on the razor variables are given in line 1 of Table 3.2.

As we increase the calorimeter-level razor thresholds, we see the rate slowly decrease to $\approx$ 10 Hz, which is around the rate that we desire. Note that for Table 3.3, we also raise the $R^2 \times M_R$ threshold in order to match the increase in R^2. This illustrates the fact that there is still a small mismatch between calorimeter objects and PF objects.

3.3 H$\rightarrow b\bar{b}$ HLT Trigger

3.4 Motivation

Investigation of an H$\rightarrow b\bar{b}$ trigger has been conducted for possible addition to the Run II trigger menu. This is motivated by the use of Higgs detection as a tool for finding new physics, which will be further elaborated on in Chapter 3. Since H$\rightarrow b\bar{b}$ is the dominant decay mode (branching ratio 57.8%) for a 125 GeV Higgs [18], this trigger is ideal for capturing the maximum amount of events. However, the high SM cross section for bottom production at the LHC poses a problem for the rate of events passing this trigger [19]. As a result, many kinematic thresholds have to be combined in order to minimize the rate to a reasonable level.

3.5 b-tagging

Bottom quarks, or b quarks, are found in many supersymmetric processes [20]. CMS has several algorithms for b quark identification, or b-tagging. For our study, we will use the Combined Secondary Vertex (CSV) algorithm. This algorithm assigns a value to each jet which indicates the probability of misidentification. There are three working points for this algorithm: loose, medium, and tight, which correspond to ≈10%, ≈1%, and ≈0.1% probabilities for an average jet p_T of 60 GeV [21]. These algorithms rely on tracks as well as secondary vertices, which can indicate the lifetime of quark before decay. This exploits the fact that the b quark has a relatively long lifetime, ≈1 picosecond [22]. CSV loose, medium, and tight points also result in varying b-tag efficiencies. The loose point corresponds to ≈ 80-90% efficiency, medium to ≈ 60-70% efficiency, and tight to ≈ 50% efficiency [21].

3.6 Trigger Flow Path

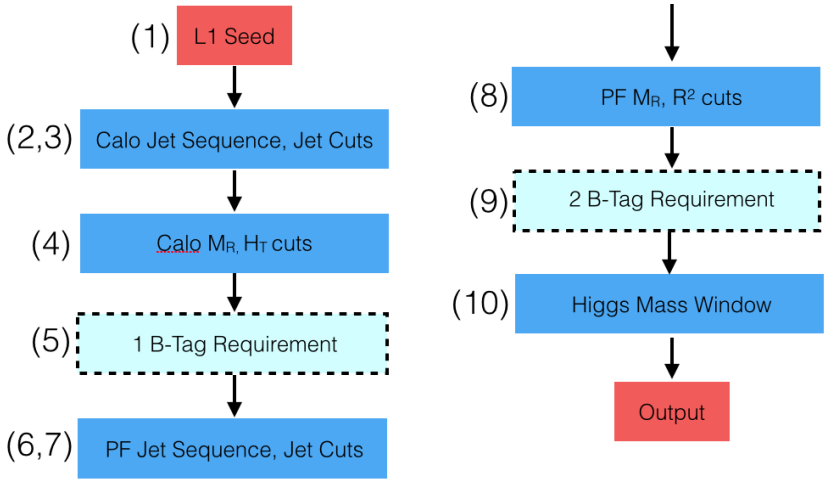


Figure 3.4: Schematic of the H→ $b\bar{b}$ trigger flow path. The numbers in parentheses correspond to the more detailed list described in the text.

The trigger flow path is as follows (Figure 3.4):

1. Input from the Level 1 CMS Hardware Trigger. For these triggers, a combination of all hadronic triggers available were used.
2. Algorithms to create calorimeter objects, which use information from the calorimeter [13].
3. Thresholds implemented on the calorimeter jet kinematic properties.

4. Thresholds implemented on encompassing variables, such as the razor variables or H_T, which is a sum of transverse hadronic energies.

5. At least one b-tagged jet with a specified CSV threshold (see Section 3.5).

6. Algorithms to create Particle Flow (PF) objects, which use an increased amount of available information from CMS [13].

7. Thresholds implemented on the PF jet kinematic properties.

8. Thresholds implemented on the razor variables.

9. At least two b-tagged jets with specified CSV threshold (see Section 3.5).

10. Higgs mass window filter applied to any two b-tagged jets that pass the specified CSV requirements in the previous step.

3.7 Rates of H$\rightarrow b\bar{b}$ Trigger

The naming convention carries from Section 3.2.3. A few more abbreviations are as follows:

1. BTagXCSV0YCSV0Z indicates that there are X b-tags required, with a Combined Secondary Vertex (CSV) cutoff of 0.Y and 0.Z for two PF jets. For calorimeter jets, there is a CSV cutoff of 0.Y - 0.1 required for one jet. See Section 3.5 for further information on b-tagging.

2. DiJetX indicates that we require two PF jets with $p_T >$ X GeV, and for calorimeter jets we require the leading jet to have $p_T >$ X - 10 GeV and the subleading jet to have $p_T >$ X - 20 GeV.

3. TriJetX indicates that we require three PF jets with $p_T >$ X GeV.

4. MqqMinXMaxY indicates that we require that there be two PF b-jets (with the b-tagging CSV requirements from (1)) that have a invariant mass between X and Y GeV.

5. HTX indicates $H_T >$ X GeV.

Trigger Name	**Rate (Hz)**
HLT_BTag2CSV05CSV02_DiJet80_Rsq0p01_MR200_MqqMin70Max190	121.770 ± 7.015
HLT_BTag2CSV05CSV02_DiJet80_HT200_Rsq0p01_MR300_MqqMin70Max190	93.946 ± 5.791
HLT_BTag2CSV05CSV02_DiJet80_CaloMR200_Rsq0p01_MR300_MqqMin70Max190	93.384 ± 5.774
HLT_BTag2CSV05CSV02_DiJet80_TriJet40_CaloMR200_Rsq0p0196_MR300_MqqMin70Max190	61.056 ± 5.333
HLT_BTag2CSV07CSV04_DiJet80_CaloMR200_Rsq0p01_MR300_MqqMin70Max190	37.483 ± 3.807
HLT_BTag2CSV07CSV04_DiJet80_TriJet40_CaloMR200_Rsq0p0196_MR300_MqqMin70Max190	25.827 ± 3.685

Table 3.4: Table of the $H \rightarrow b\bar{b}$ triggers discussed, where the rate is calculated for an instantaneous luminosity of 1.4e34 $\text{cm}^{-2}\text{s}^{-1}$.

3.7.1 Addition of new HCAL local reconstruction method and ECAL Multifit

An update of the calorimeter local reconstruction methods to handle out-of-time pile-up, which is pile-up from adjacent bunch-crossing, was released [23]. This improved energy reconstruction in the hadronic calorimeter endcap and barrel, relies on fitting measured pulses with parameterizations derived from test pulses, then subtracting the out-of-time pileup contribution from the energy measurement. This technique results in better resolution and measurement [24]. A similar update was released for the electromagnetic calorimeter to reduce out-of-time pileup by simulating and fitting the pulse shape from an average of pulses from samples shifted from each other in bunch-crossings [25]. We applied this technique to two of our proposed triggers and measured their rates.

Trigger Name	**Rate (Hz)**
HLT_BTag2CSV05CSV02_DiJet80_TriJet40_CaloMR200_Rsq0p0196_MR300_MqqMin70Max190	59.174 ± 5.156
HLT_BTag2CSV07CSV04_DiJet80_TriJet40_CaloMR200_Rsq0p0196_MR300_MqqMin70Max190	24.112 ± 3.550

Table 3.5: Table of the $H \to b\bar{b}$ triggers discussed, where the rate is calculated for an instantaneous luminosity of 1.4e34 $\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

3.7.2 Discussion

The rates are shown in Table 3.4. We see that the tightening of the CSV thresholds sharply reduces the rate. The addition of the razor variable or H_T cut is also effective at reducing the rate and is low enough to keep high signal events. We refrain from a high R^2 threshold in an effort to select for different events from the regular razor triggers. We also keep a relatively large Higgs mass window since the main goal of this trigger is to keep events that include a Higgs.

The new ECAL and HCAL methods do not reduce the rates significantly (Table 3.5), but should still be implemented due to the improvements in measurement.

Chapter 4

Investigation of Higgs-Aware Decay Models

4.1 The Higgs decay as a tool for new physics searches

Armed with the discovery of the Higgs Boson, we are able to employ the Higgs as a tool for future physics studies. We can study its mass and other properties, which have implications for the possible existence of beyond-the-standard-model (BSM) phenomena including SUSY and dark matter. We can also use the Higgs boson as a probe for searching for new physics. For example, the Higgs invisible decay channel has been ideal for BSM studies since the branching fraction of the purely SM decay, $h \to ZZ^* \to 4\nu$ is small [26]. There have already been studies exploiting this channel [27].

Similarly, we can employ the Higgs to look for SUSY by targeting decays that specifically produce a Higgs, such as those that include a next-to-lightest neutralino decaying to a Higgs and a lightest neutralino, $\chi_2^0 \to h\chi_1^0$ [28]. The decay of the Higgs to two photons has been observed as a clean, sharp signal [29] and is an ideal decay channel to target. A previous search for supersymmetry with this decay channel motivates this thesis work [6], but we will examine different SUSY models from the search.

4.2 Motivation

Many standard SUSY searches rely on missing transverse energy ($\not{E}_T$) in order to separate signal from Standard Model background. $\not{E}_T$ is calculated from energy deposits in the subdetectors [30]. As a result, strong limits have been placed for SUSY models with large $\not{E}_T$ [31]. These standard searches look for models with signal events at R^2 values of greater than 0.2 - 0.3 [28]. Instead, we would like to investigate models with lower $\not{E}_T$ and lower R^2, which may have eluded LHC bounds but are good candidates for Run II.

Furthermore, the use of the Higgs decay channels gives us an alternative method for reducing

background from our analysis. We can then study models with less $\not{E}_T$ and look in areas of the SUSY phase space that have not been extensively studied. We will focus on bottom squark production decaying to Higgses, rather than top squark production. In general, since top quarks decay to a W boson and a bottom quark, top squark models have more complicated topology with higher $\not{E}_T$ due to the W bosons decaying to neutrinos.

Our goal is to study a viable model for a LHC Run II analysis. In order to achieve this, this thesis will explore the different kinematic parameters in bottom squark production and decay. In particular, we desire models which have low $\not{E}_T$, low R^2, and M_R values clustered from 400 to 500 GeV. Higher M_R values have a higher probability of already being excluded, and at too low of M_R we lose events from the trigger thresholds.

We can consider a model where we have a bino-like next-to-lightest supersymmetric particle (NLSP) which would decay predominantly to a Higgs boson and a weakly coupled LSP [32]. The next-to-lightest neutralino can be considered the NLSP, and the lightest neutralino is the LSP. For the dominant mode of bottom squark production, the final state particles will be two bottom quarks and two Higgs bosons.

CMS SUSY searches are interpreted in the context of simplified models [20]. A simplified model consists of only relevant particles, specifying the particle production and subsequent decay. It is defined by the model parameters that would influence how an event would look in the detector. For example, the T2qq model described in [20] is a simplified model that focuses on pair-produced squarks decaying to two quarks and two LSPs, where the free parameters are masses of the squark and the LSP. We will use this interpretation for our studies.

4.3 Bottom squark production model

We investigated a bottom squark production model, where the bottom squark has two possible decays channels: $\tilde{b} \to b\tilde{\chi}_2^0 \to bH\tilde{\chi}_1^0$, and $\tilde{b} \to b\tilde{\chi}_1^0$.

The bottom squark particle can be produced through several mechanisms, see Section A.2. The dominant production modes result in pair production.

We wanted to obtain M_R peak values from 400-500 GeV for these events, and we needed to comply with current exclusion limits [33]. Note that these exclusion limits assume a decay of $\tilde{b} \to b\tilde{\chi}_1^0$. If we examine Fig. 4.1, we see that the contour lines for $M_\Delta = 450 \pm 50$ GeV become approximately linear at above $\tilde{\chi}_1^0 \approx$ 200-250 GeV. Referring to the limits in [33], we see that $m_{\tilde{b}} = 600$ GeV and $m_{\tilde{\chi}_1^0} = 300$ GeV is a good starting point. For all of the follow studies, the $m_{\tilde{\chi}_2^0}$ is set at $m_{\tilde{\chi}_1^0} +$ 130 GeV to minimize the $\not{E}_T$ carried away by the lightest neutralino. After choosing appropriate mass points, several working points of the branching ratios for the two decay channels of the $\tilde{b}$ were chosen for examination: 10%, 50%, 90%, and 100% decay to $bH\tilde{\chi}_2^0$.

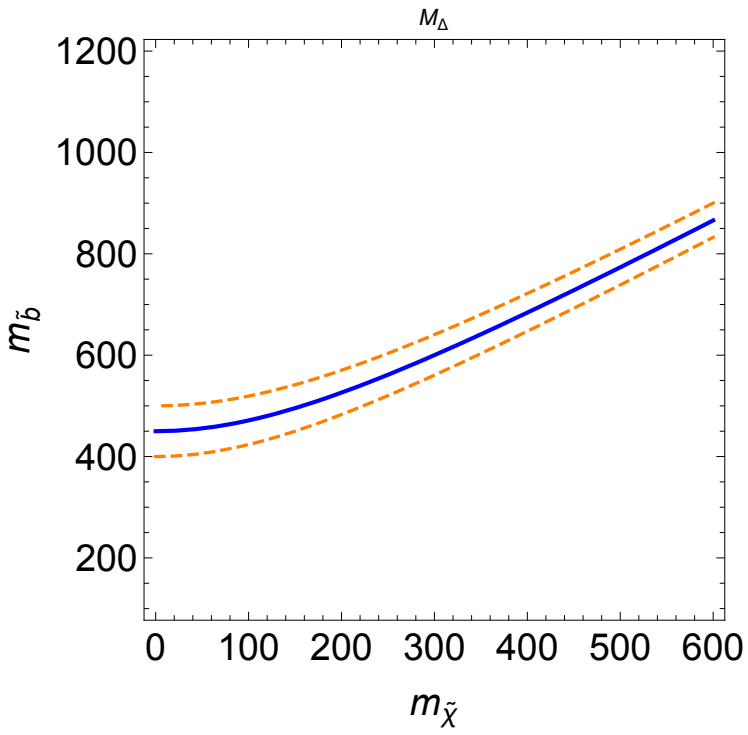


Figure 4.1: The contour lines for $M_\Delta = 450$ GeV, where $M_\Delta = \frac{m_{\tilde{b}}^2 - m_{\tilde{\chi}}^2}{m_{\tilde{b}}}$, with $\pm$ 50 GeV as indicated by the dashed lines.

Events were produced at 8 TeV center-of-mass energy using PYTHIA 8.1 [34, 35] interfaced with SUSY Les Houches Accord (SLHA) files [36] and the BSMatLHC package [37]. They were filtered using a selection process described in Section 4.4. The Higgs boson was forced to decay to two photons.

4.4 Selection

After events were produced, several requirements were implemented to filter out the events. This selection was loosely modeled after the aforementioned Higgs to diphoton supersymmetry search [6] to reduce potential background, with modifications specific to a generator-level study.

1. The event must have a pair of photons that has a diphoton mass of > 100 GeV. One of these photons should have $p_T > 40$ GeV, and the other should have $p_T > 25$ GeV.

2. Then, if two pairs of photons that satisfy (1) are found, then one of them is randomly set aside to later be counted as an alternate channel for Higgs decay besides two photons. This is done to retain the maximum amount of events generated by PYTHIA.

3. The pair of photons that is not set aside must have $p_T > 20$ GeV and $|\eta| < 1.44$.

4. The event must have at least one jet with $p_T > 30$ GeV and $|\eta| < 3.0$, and have a $\Delta R > 0.5$ with respect to the either of the pair of photons selected. These jets are produced using the PF algorithm, clustered with an anti-kT algorithm using ΔR = 0.5 [15, 14, 38].

5. We only keep jets with $p_T > 30$ GeV and $|\eta| < 3.0$.

4.5 Box definitions

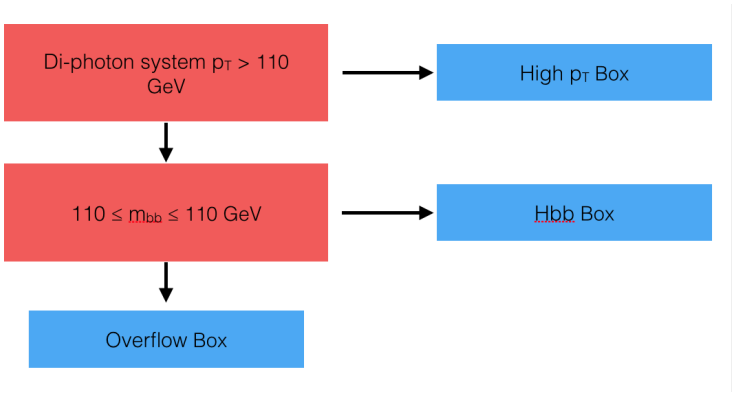


Figure 4.2: A schematic for how events are placed into the three boxes. The red rectangles are the two requirements, and the blue rectangles represent the boxes.

After selection, the events were further categorized into boxes for further discrimination of potential signal from background. These boxes are all restricted to $M_R \in [150, 3000]$ GeV and $R^2 \in [0, 1]$. The boxes are again motivated by the Higgs to diphoton supersymmetry search and are as follows:

1. **High p_T Box** requires p_T of the di-photon system to be > 110 GeV.

2. **Hbb Box** requires $110 \leq m_{bb} \leq 140$ GeV, with m_{bb} being the invariant mass of any two b-tagged jets.

3. **Overflow Box**

The events are grouped into the boxes in a hierarchal system. So, if an event passes the requirements for box 1, then it is grouped into box 1. If it fails and passes the requirements for box 2, then it goes in box 2. Else, it goes into the Overflow Box (see Figure 4.2). These boxes are constructed to characterize a potential SUSY signal from background events. The High p_T box selects for boosted Higgs bosons (with high momentum), while the Hbb box allows us to select for diHiggs production. For the Hbb box, we assume that the second Higgs produced would decay to two b quarks because of the high branching ratio [18].

Because the branching ratio of $H \to \gamma\gamma$ is low at a fraction of 0.00228 [39], events were produced with an artificial branching ratio of 100%. As previously stated, in the analysis one of the $H \to \gamma\gamma$

photon pairs is randomly set aside to be counted as an alternate channel. We apply a 57.8% chance that this pair should be counted as a $H \to b\bar{b}$ decay channel to simulate the branching ratio [18]. Then, with every b-jet in the event, we emulate the b-tagging efficiency by randomly tagging 60% of generator level b-jets. This approximately corresponds to the efficiency of a CSV medium b-tag requirement (see Section 3.5). The reason this artificial handling of bottom quarks is necessary so we do not overestimate the number of events that would fall in the Hbb box. These box definitions are important for future inclusive analyses and will be used in the later statistical study in Section 4.11.

4.6 Branching Ratio to $\tilde{\chi}_2^0$ Study

$m_{\tilde{b}}$	$m_{\tilde{\chi}_2^0}$	$m_{\tilde{\chi}_1^0}$	BR (to $\tilde{\chi}_2^0$)
600 GeV	430 GeV	300 GeV	10%
600 GeV	430 GeV	300 GeV	50%
600 GeV	430 GeV	300 GeV	90%
600 GeV	430 GeV	300 GeV	100%

Table 4.1: Table of the models investigated in the branching ratio study, with their associated values of M_Δ and M_{split}.

In the following section we will examine the effect of the branching ratios listed in Table 4.1 on the kinematics of the $m_{\tilde{b}} = 600$ GeV, $m_{\tilde{\chi}_1^0} = 300$ GeV model. If we examine Fig. 4.3, we see that we obtain the M_R shape in the desired range as predicted by M_Δ, as well as R^2 peak values in the desired range. As the branching ratio to $\tilde{\chi}_2^0$ decreases, we see that the peak region is less defined. Fig. 4.3d shows that the R^2 versus M_R distribution with the loosest clustering in the events corresponds to the smallest branching ratio. This can be understood through the chance that the event we capture is asymmetric.

With an 100% branching ratio to $\tilde{\chi}_2^0$, we see a marked decrease in events at high R^2. Because of how R^2 is defined with $\not{E}_T$ dependence, larger $\not{E}_T$ in general should result in a larger $\not{E}_T$ if the M_R is fixed. The 2D distribution shrinks to lower R^2, although the M_R distribution, as expected, shifts very little. However, regardless of the branching ratio, the plots still maintain a significant portion of events above $R^2 \approx 0.1$. With Fig. 4.4, we see a similar phenomenon as the branching ratio to $\tilde{\chi}_2^0$ increases. The plot shows the strong dependence of R^2 on the $\not{E}_T$ distribution. Although the distribution becomes more defined, the rough slope of the R^2 vs $\not{E}_T$ 2D distribution remains the same.

The M_R versus $\not{E}_T$ 2D plots reveal the robustness of the M_R variable with varying branching ratios. Comparing Fig. 4.4a and b shows that increasing the branching ratio clusters the events closer. Examining Fig. 4.4a and d as well, we see that the $\not{E}_T$ becomes pushed higher. However, we

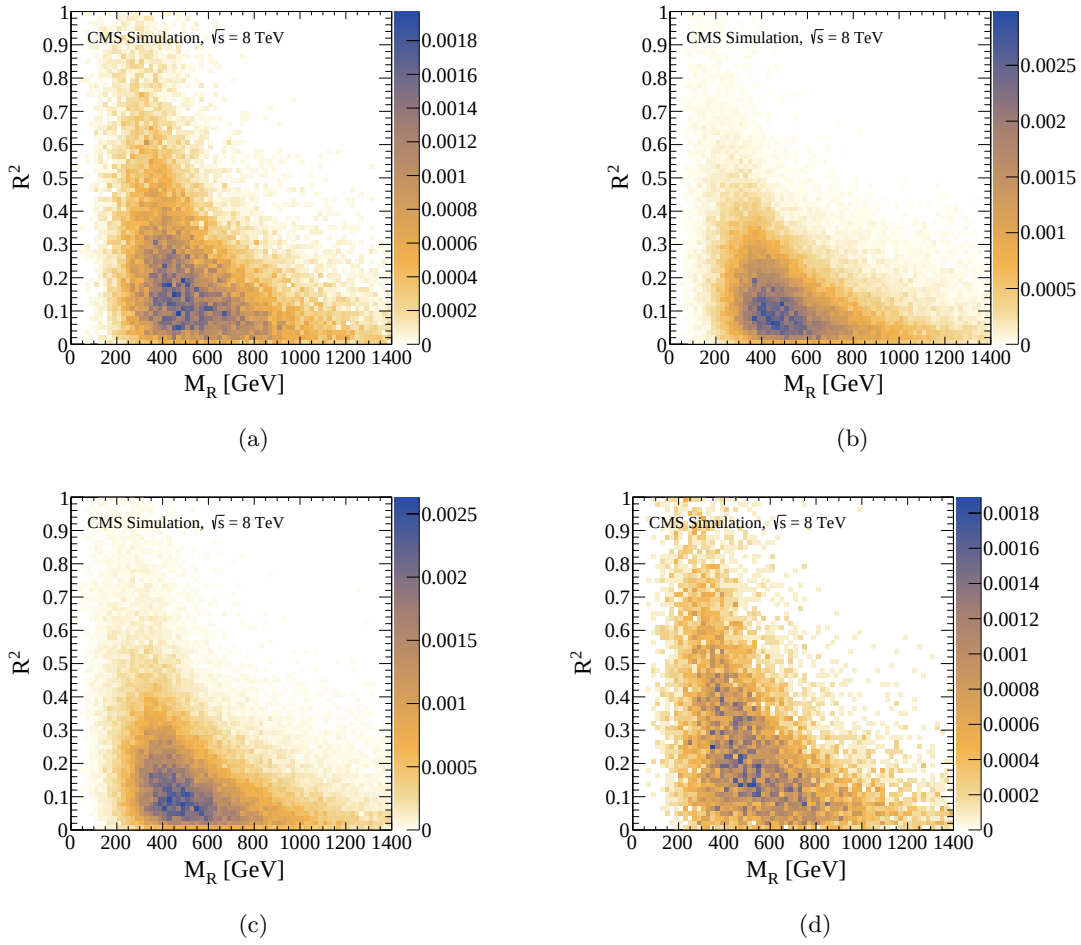


Figure 4.3: Z-axis scale units are such that the total number of events is normalized to 1. $m_{\tilde{b}} = 600$ GeV, $m_{\tilde{\chi}_1^0} = 300$ GeV model. (a) R^2 versus M_R for the 50% (to $\tilde{\chi}_2^0$) working model. (b) R^2 versus M_R for the 100% (to $\tilde{\chi}_2^0$) working model. (c) R^2 versus M_R for the 90% (to $\tilde{\chi}_2^0$) working model. (d) R^2 versus M_R for the 10% (to $\tilde{\chi}_2^0$) working model.

also lose the nicely clustered peak since the decay of the $\tilde{b}$ pair is now mostly asymmetric. However, the M_R spread seems to change minimally.

The conclusion from this study of the branching ratios with this one mass working point is that an increase in branching ratio to $\tilde{\chi}_2^0$ is beneficial for an R^2 distribution that evades the current LHC bounds. Varying the branching ratio has little effect on the M_R. Furthermore, standard bottom squark searches assume that the bottom squark decays directly to $\tilde{\chi}_1^0$ [33], so increasing the branching ratio to $\tilde{\chi}_2^0$ would loosen limits.

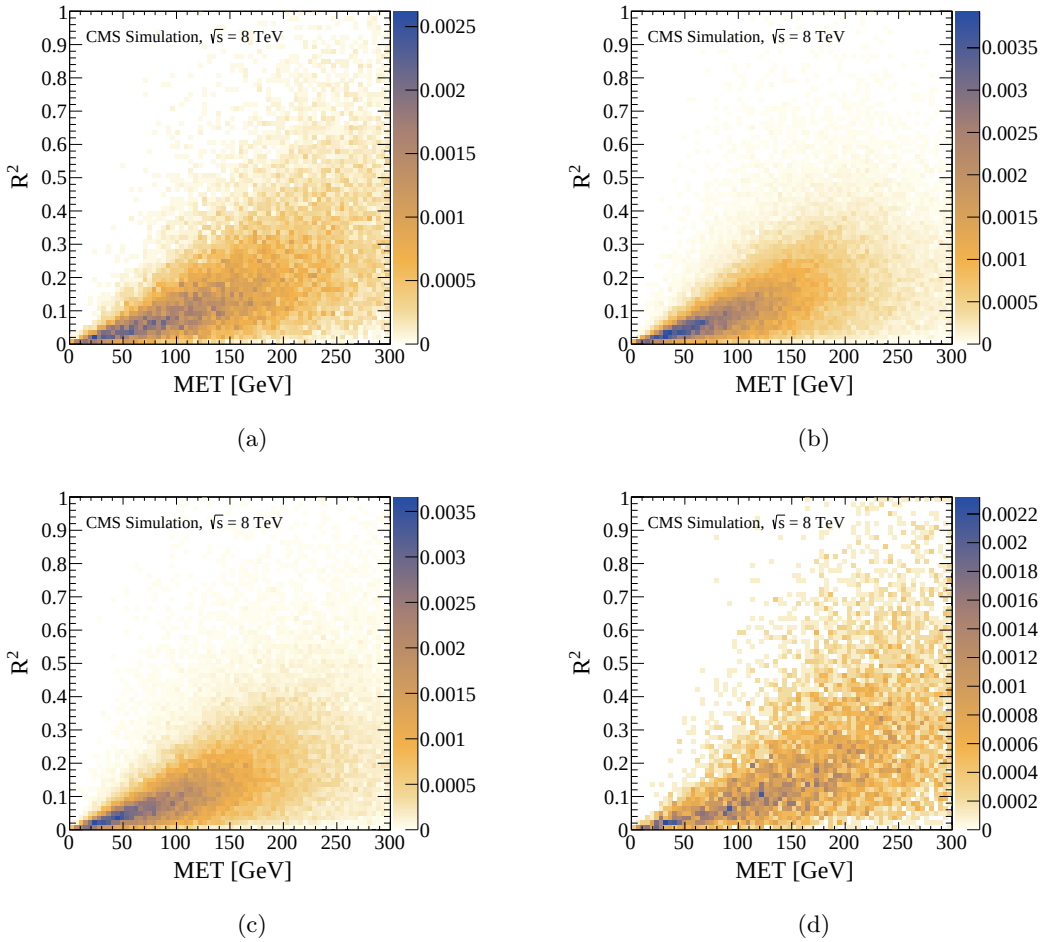


Figure 4.4: Z-axis scale units are such that the total number of events is normalized to 1. $m_{\tilde{b}} = 600$ GeV, $m_{\tilde{\chi}_1^0} = 300$ GeV model. (a) R^2 versus $\not{E}_T$, for the 50% (to $\tilde{\chi}_2^0$) working model. (b) R^2 versus $\not{E}_T$, for the 100% (to $\tilde{\chi}_2^0$) working model. (c) R^2 versus $\not{E}_T$, for the 90% (to $\tilde{\chi}_2^0$) working model. (d) R^2 versus $\not{E}_T$, for the 10% (to $\tilde{\chi}_2^0$) working model.

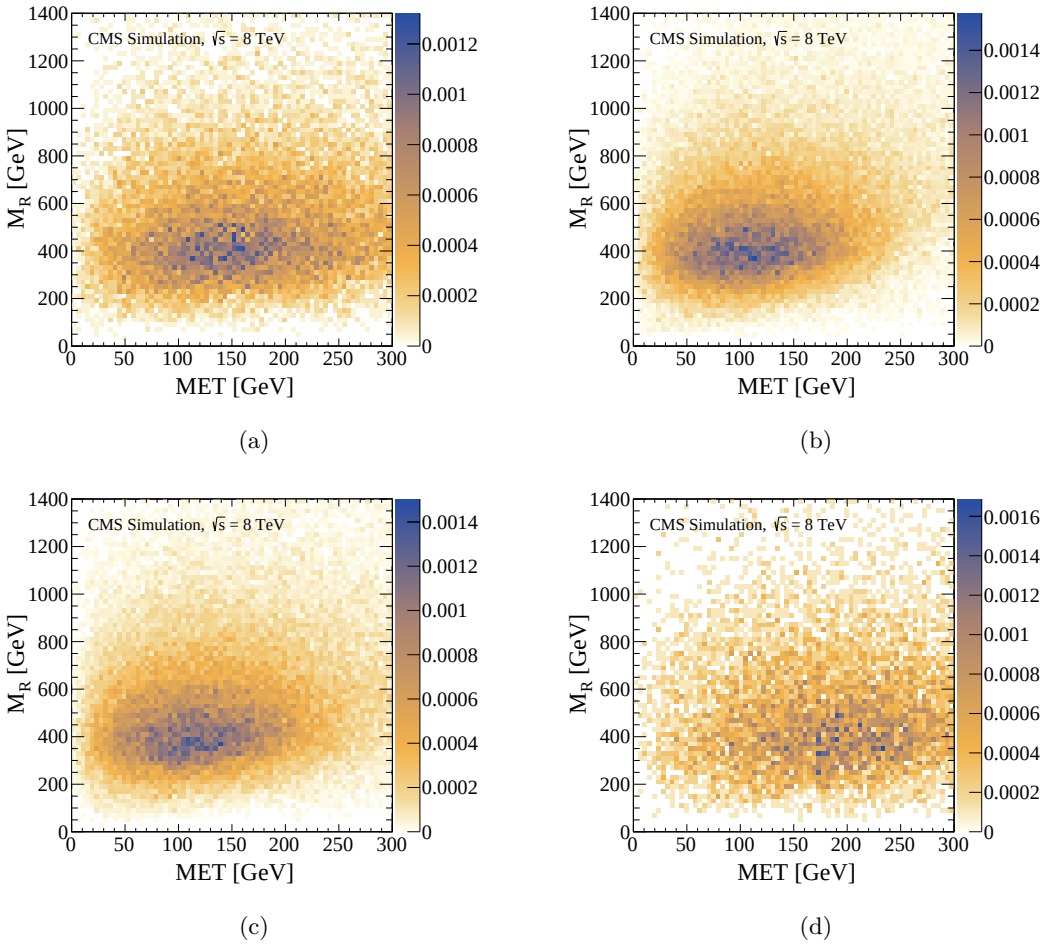


Figure 4.5: Z-axis scale units are such that the total number of events is normalized to 1. $m_{\tilde{b}} = 600$ GeV, $m_{\tilde{\chi}_1^0} = 300$ GeV model. (a) M_R versus $\not{E}_T$ for the 50% (to $\tilde{\chi}_2^0$) working model. (b) M_R versus $\not{E}_T$, for the 100% (to $\tilde{\chi}_2^0$) working model. (c) M_R versus $\not{E}_T$, for the 90% (to $\tilde{\chi}_2^0$) working model. (d) M_R versus $\not{E}_T$, for the 10% (to $\tilde{\chi}_2^0$) working model.

4.7 Mass Splitting Study with Fixed $\tilde{\chi}_1^0$ Mass

The next parameter that would affect the kinematics is the mass splitting of the system, i.e. the difference in mass between the bottom squark and lightest neutralino. We can classify this in two ways: using M_Δ, or simply $M_{split} = m_{\tilde{b}} - m_{\tilde{\chi}_1^0}$. The following subsections investigate the kinematics of models with the characteristics listed in Table 4.2. Notice the caveat that these effects would not necessarily be decoupled from the effects of changing the bottom squark mass.

$m_{\tilde{b}}$	$m_{\tilde{\chi}_2^0}$	$m_{\tilde{\chi}_1^0}$	M_Δ	M_{split}	BR (to $\tilde{\chi}_2^0$)
550 GeV	430 GeV	300 GeV	386.36 GeV	250 GeV	90%
450 GeV	430 GeV	300 GeV	250 GeV	150 GeV	100%
530 GeV	430 GeV	300 GeV	360.19 GeV	230 GeV	100%

Table 4.2: Table of the models investigated in the mass splitting study, with their associated values of M_Δ and M_{split}.

The 1D $\not{E}_T$ distributions have a strong dependence on mass splitting since with a smaller mass splitting there is likely less energy available for the decay particles (Fig. 4.10a. With Figures 4.6, 4.7, 4.8, we can also see a strong dependence of M_R on M_Δ. This is in line with the principle that M_R is an estimator of M_Δ. If we compare Figure 4.6b to Figure 4.4c, the R^2 versus $\not{E}_T$ 2D distributions are roughly the same, indicating that the decrease in mass splitting from the $m_{\tilde{b}} = 600$ GeV model to the $m_{\tilde{b}} = 550$ GeV has a relatively small effect on these variables. However, when we look at Figure 4.7b, we notice a large difference not only in a tighter clustering at small $\not{E}_T$, R^2, but also a change in the slope for the $m_{\tilde{b}} = 450$ GeV model. From Figure 4.7c, we also notice that the M_R clusters at a center of less than 200 GeV, which is lower than we would expect. There is an unexpected shift of the peaks of the M_R distributions if we compare the various models in Figure 4.9a, where the peak for $m_{\tilde{b}} = 450$ GeV is ≈ 100 GeV lower than the corresponding peaks for the other models.

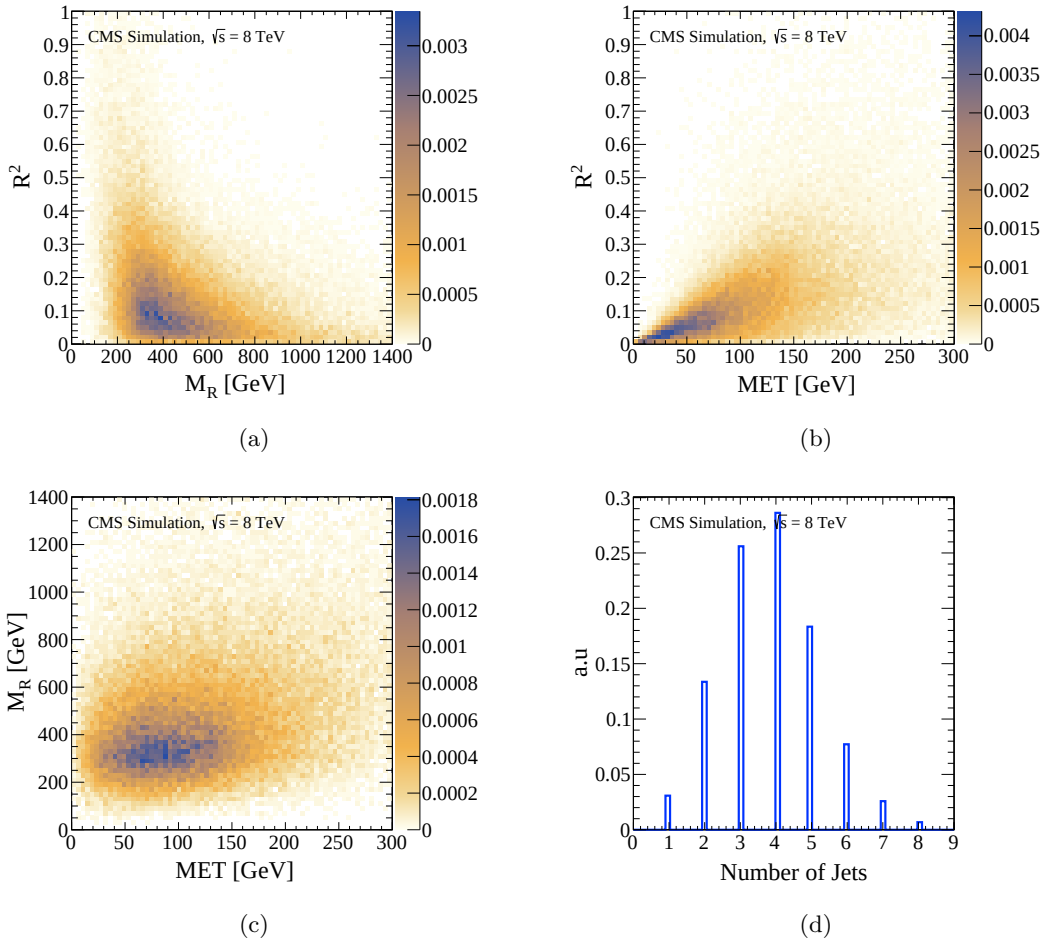


Figure 4.6: Characterization of the $m_{\tilde{b}} = 550$ GeV, $m_{\tilde{\chi}_1^0} = 300$ GeV model with 90% to $\tilde{\chi}_2^0$. Z-scale units are such that the total number of events is normalized to 1. (a) R^2 versus M_R. (b) R^2 versus $\not{E}_T$. (c) M_R versus $\not{E}_T$. (d) Number of PF jets, clustered with the anti-kT algorithm with ΔR = 0.5.

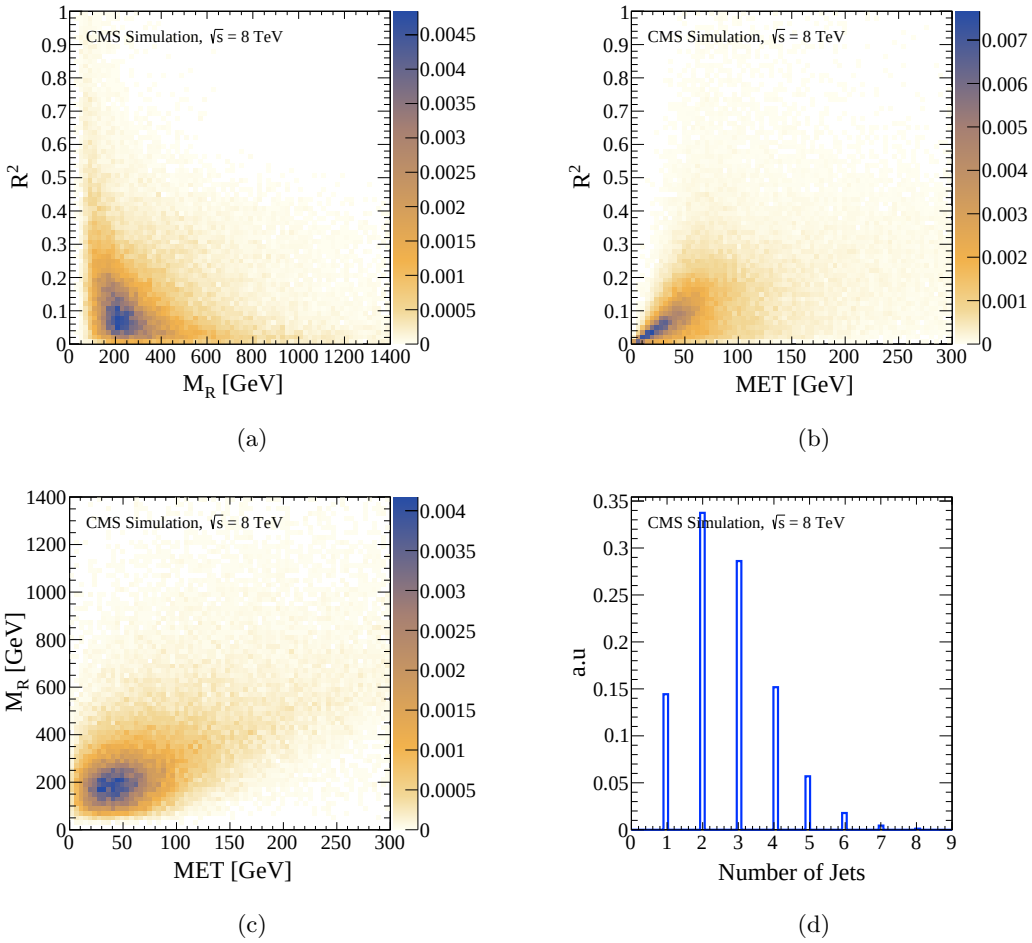


Figure 4.7: Characterization of the $m_{\tilde{b}} = 450$ GeV, $m_{\tilde{\chi}_1^0} = 300$ GeV model with 100% to $\tilde{\chi}_2^0$. Z-scale units are such that the total number of events is normalized to 1. (a) R^2 versus M_R. (b) R^2 versus $\not{E}_T$. (c) M_R versus $\not{E}_T$. (d) Number of PF jets, clustered with the anti-kT algorithm with ΔR = 0.5.

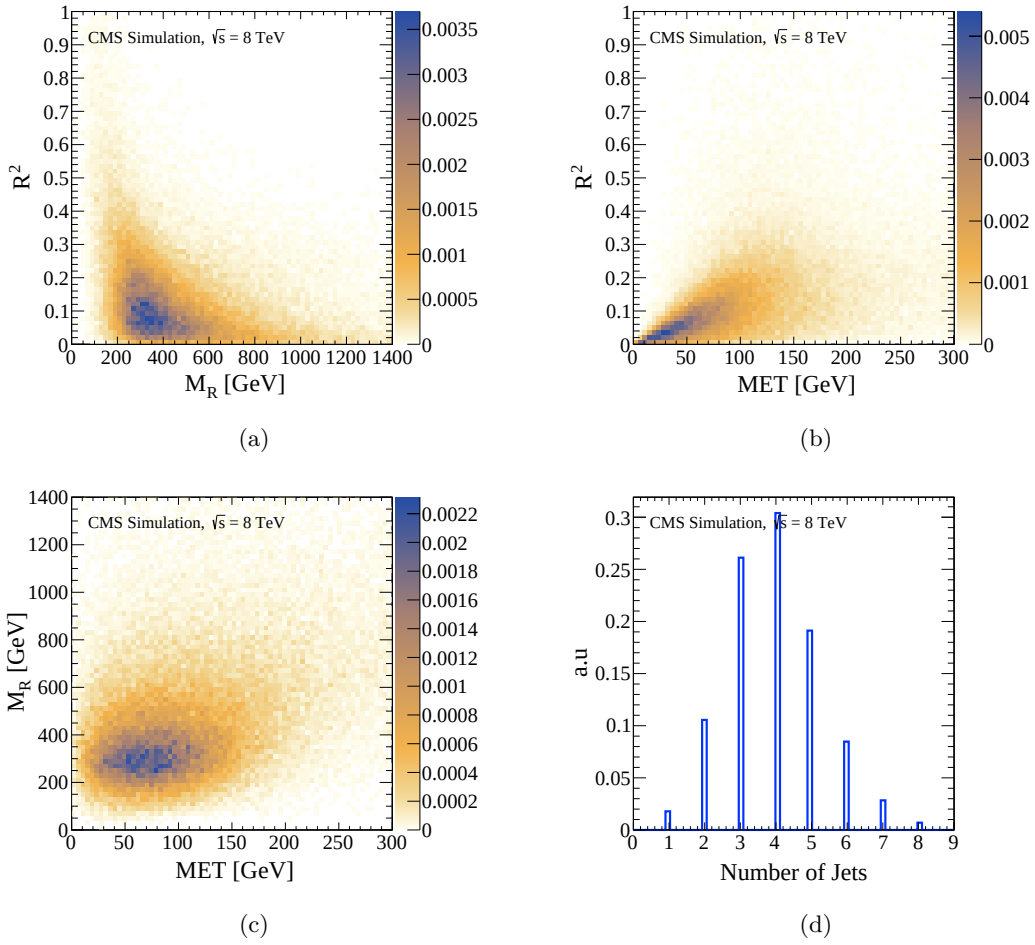
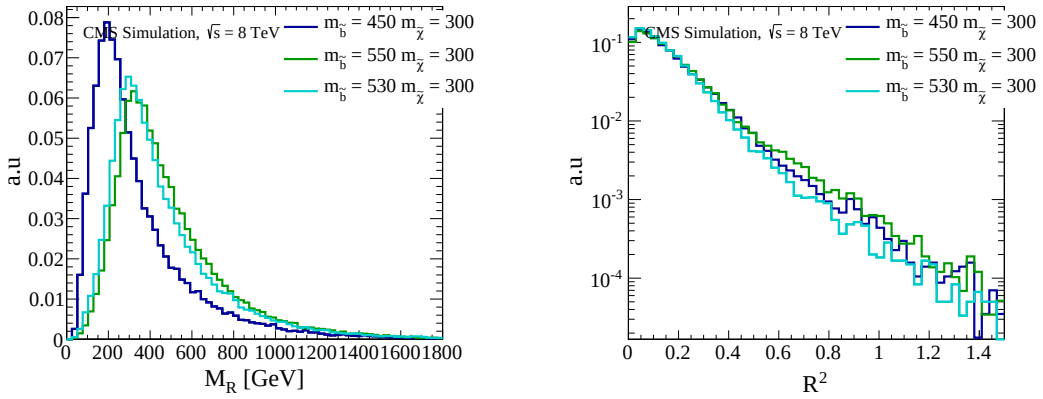


Figure 4.8: Characterization of the $m_{\tilde{b}} = 530$ GeV, $m_{\tilde{\chi}_1^0} = 300$ GeV model with 100% to $\tilde{\chi}_2^0$. Z-scale units are such that the total number of events is normalized to 1. (a) R^2 versus M_R. (b) R^2 versus $\not{E}_T$. (c) M_R versus $\not{E}_T$. (d) Number of PF jets, clustered with the anti-kT algorithm with ΔR = 0.5.

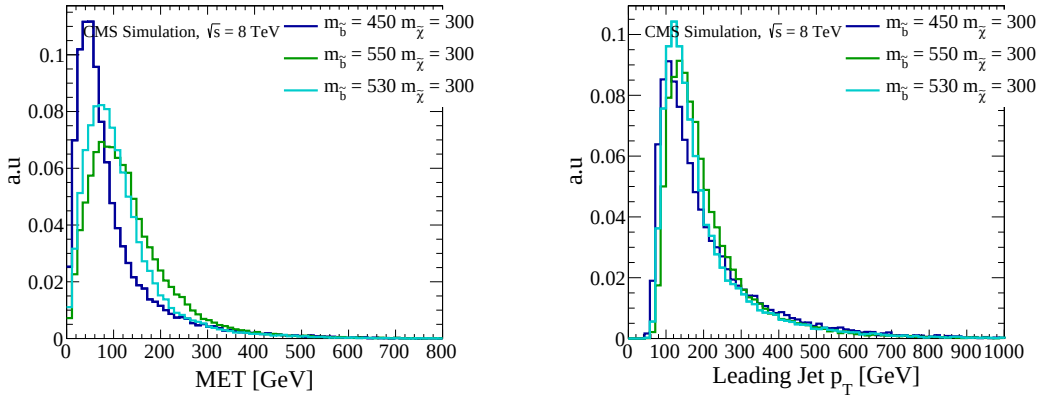
In an effort to decipher these effects, we can think of how M_R is defined. Since M_R is constructed from jets and particles that must have a minimum p_T, soft particles will be excluded. In Figure 4.11, we see that the available energy to impart to the bottom quark is only 20 GeV; thus, these b-jets are likely very soft. As a result, they are not clustered in the hemispheres that go into the calculation of M_R, and M_Δ is not being constructed accurately. Since $R = M_T^R/M_R$, this also affects the distribution of R^2. A smaller M_R than expected thus increases the slope of the R^2 versus $\not{E}_T$ plot.

This explanation is further supported by the plot of the distribution of PF jets that enter the hemispheres (see Figure 4.7d), which has a mode of only two jets. We expect on average four jets for a model with BR = 100% to $\tilde{\chi}_2^0$, with two from the direct decay of the bottom squark particle and two from the Higgs decay. This is validated in Figure 4.6d. Furthermore, the leading jet p_T for



(a) (b)

Figure 4.9: Razor variable distributions for the various mass splitting with fixed $\tilde{\chi}_1^0$ mass. (a) M_R 1D distributions. (b) R^2 1D distributions.



(a) (b)

Figure 4.10: Kinematic distributions for the various mass splitting with fixed $\tilde{\chi}_1^0$ mass. (a) $\not{E}_T$ 1D distributions. (b) Leading jet p_T 1D distributions (excluding H $\rightarrow b\bar{b}$ jets).

the $m_{\tilde{b}} = 450$ GeV model is shown to be very soft in Figure 4.10b.

Following the $m_{\tilde{b}} = 450$ GeV model, the $m_{\tilde{b}} = 530$ GeV model was investigated. Here, we see that although the mass splitting is smaller, the slope of the R^2 versus $\not{E}_T$ plot does not increase dramatically (Figure 4.8b). Thus, we can conclude that there is a threshold of mass splitting in between 530 GeV and 450 GeV, mostly likely around ≈460 GeV, in which the b-jets are too soft to be included in the calculation of M_R. The number of jets in Figure 4.8d is also as expected, and M_R returns to being a good probe for M_Δ (Figure 4.8c).

In conclusion, decreasing the mass splitting changes the M_R as expected, but has little effect on the R^2 up to a certain threshold. After the b-jets are too soft to be included in the calculation of M_R, we see a strong drop in M_R values and an undesirably high R^2.

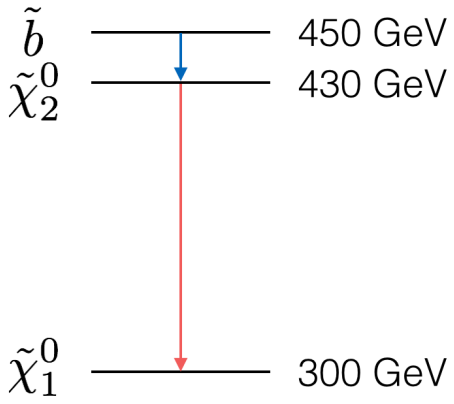


Figure 4.11: The decay spectrum for the $m_{\tilde{b}} = 450$ GeV model.

4.8 Mass Splitting Study with Fixed $\tilde{b}$ Mass

$m_{\tilde{b}}$	$m_{\tilde{\chi}_2^0}$	$m_{\tilde{\chi}_1^0}$	M_Δ	M_{split}	BR (to $\tilde{\chi}_2^0$)
600 GeV	230 GeV	100 GeV	583.33 GeV	500 GeV	100%
600 GeV	330 GeV	200 GeV	533.33 GeV	400 GeV	100%
600 GeV	430 GeV	300 GeV	450 GeV	300 GeV	100%

Table 4.3: Table of the models investigated in the mass splitting study with fixed $\tilde{b}$ mass, and their associated values of M_Δ and M_{split}.

Varying the mass splitting by changing the $\tilde{\chi}_1^0$ mass has a different effect than varying the $\tilde{b}$ mass. We see that larger mass splitting corresponds to larger M_R values, as expected, but causes R^2 to fall faster (Fig. 4.14). The 2D R^2 versus M_R distribution is also clustered more tightly around the peak values (Fig. 4.13a, 4.12a). This is due to an increasing M_R accompanied by a relatively constant $\not{E}_T$ distribution in Figure 4.15a. Although the larger mass splitting correlates with a larger leading jet p_T (Fig. 4.15b), which we expect is the b quark directly decaying from the bottom squark, the $\tilde{\chi}_1^0$ p_T decreases with larger mass splitting. Furthermore, we can see a dependence of the azimuthal angle, $\Delta\Phi$, between the two $\tilde{\chi}_1^0$ particles. The 2D M_R versus $\not{E}_T$ distribution becomes more spread out with larger mass splitting (Fig. 4.13c, 4.12c).

We note that the original squarks are almost always produced back-to-back in $\Delta\Phi$, regardless of the mass splitting (Fig. 4.17b). Also note that as the leading jet p_T increases, the $\tilde{\chi}_1^0$ particles will most likely deviate more from the path of the original bottom squark due to the recoil of the $\tilde{\chi}_1^0$ from the Higgs, which are produced from $\tilde{\chi}_2^0$ particles recoiling against the b quarks. The larger the mass splitting, the larger the leading jet p_T, and the more the $\tilde{\chi}_1^0$ particles deviate from the original path of the bottom squark. This is illustrated in the $\tilde{\chi}_1^0$ $\Delta\Phi$ distributions in Figure 4.16b. The smaller mass splitting correlates with the $\tilde{\chi}_1^0$ particles being produced more often back-to-back in the azimuthal angle.

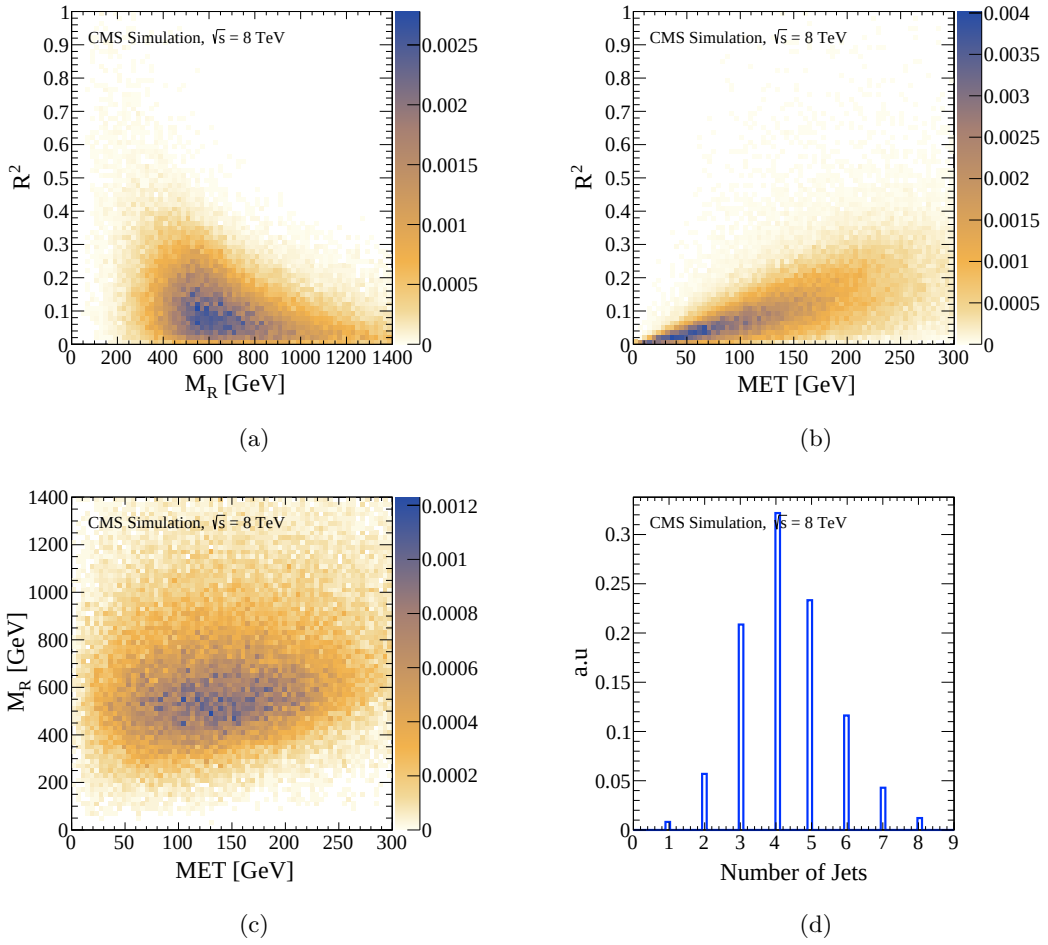


Figure 4.12: Characterization of the $m_{\tilde{b}} = 600$ GeV, $m_{\tilde{\chi}_1^0} = 200$ GeV model with 100% to $\tilde{\chi}_2^0$. Z-scale units are such that the total number of events is normalized to 1. (a) R^2 versus M_R. (b) R^2 versus $\not{E}_T$. (c) M_R versus $\not{E}_T$. (d) Number of PF jets, clustered with the anti-kT algorithm with ΔR = 0.5.

Even though a smaller mass splitting has less energy available to the $\tilde{\chi}_1^0$, its effect is overshadowed by the explanation that the $\tilde{\chi}_1^0$ particles likely have more p_T if they remain on the path of the original bottom squark. Examining Figure 4.16a, we see that a smaller mass splitting results in higher leading $\tilde{\chi}_1^0$ p_T. However, this higher leading $\tilde{\chi}_1^0$ p_T is balanced by the effect that the $\tilde{\chi}_1^0$ particles are produced more often back-to-back in $\Delta\Phi$ with a smaller mass splitting. As a result, the $\not{E}_T$ distribution stays roughly constant due to the two $\tilde{\chi}_1^0$ momentum vectors canceling each other. Overall, this effect can be attributed to the two-step decay chain.

This phenomena was not present in Section 4.7. In the previous section, any effects from variable sampling the $\sqrt{s}$ distribution due to the production of bottom squarks of different masses were not removed. This likely counteracted the effect of the particle decay angles, resulting in a $\not{E}_T$ distribution that increased with mass splitting. More importantly, however, all jets produced were

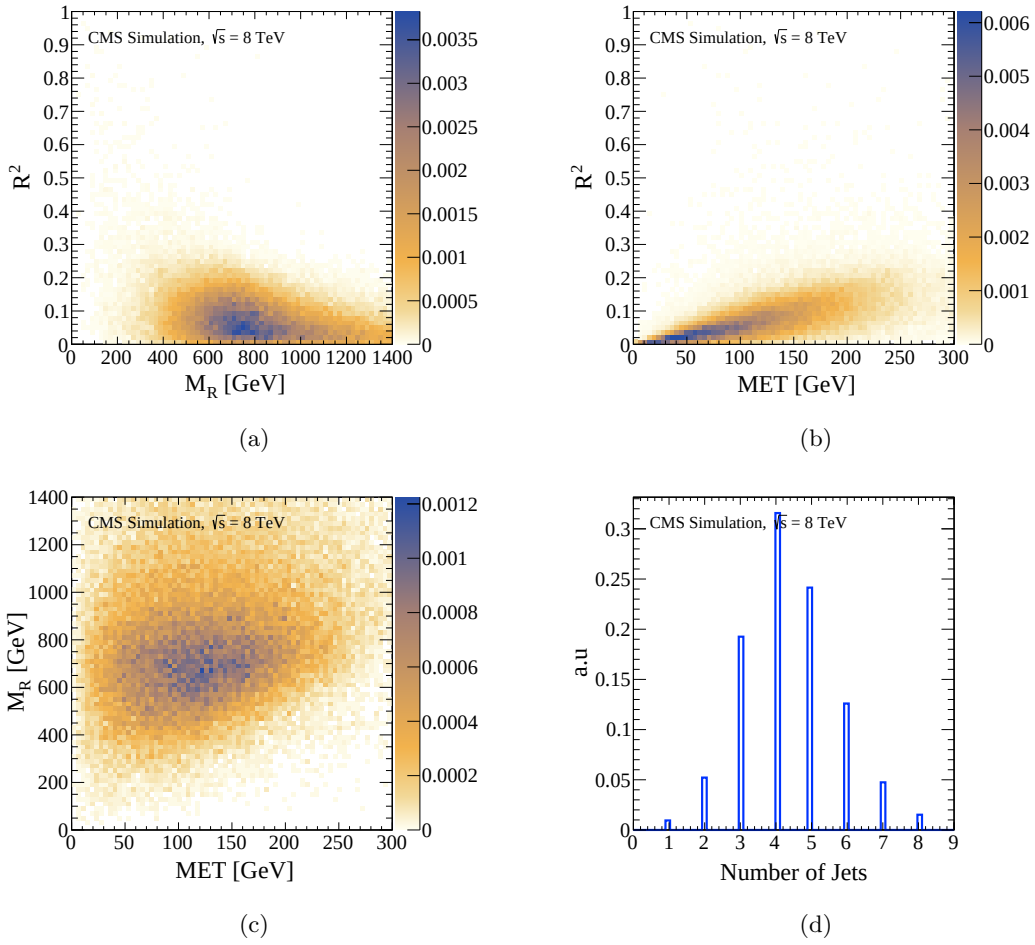
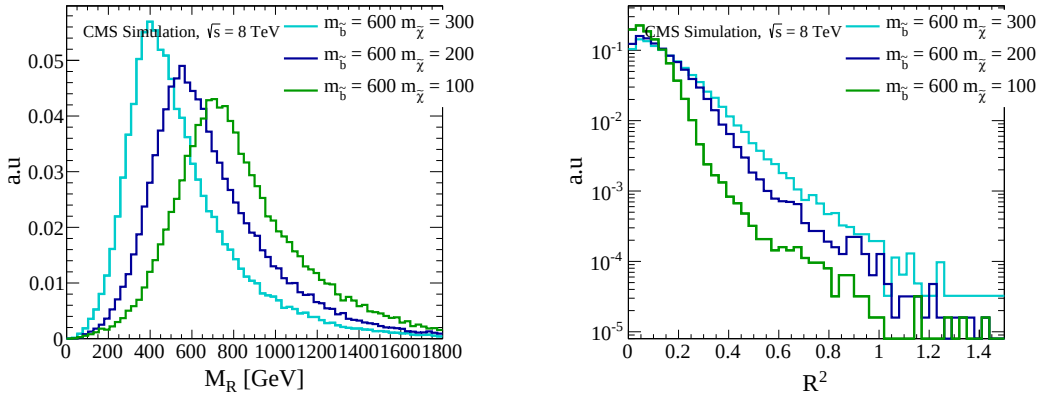


Figure 4.13: Characterization of the $m_{\tilde{b}} = 600$ GeV, $m_{\tilde{\chi}_1^0} = 100$ GeV model with 100% to $\tilde{\chi}_2^0$. Z-scale units are such that the total number of events is normalized to 1. (a) R^2 versus M_R. (b) R^2 versus $\not{E}_T$. (c) M_R versus $\not{E}_T$. (d) Number of PF jets, clustered with the anti-kT algorithm with ΔR = 0.5.

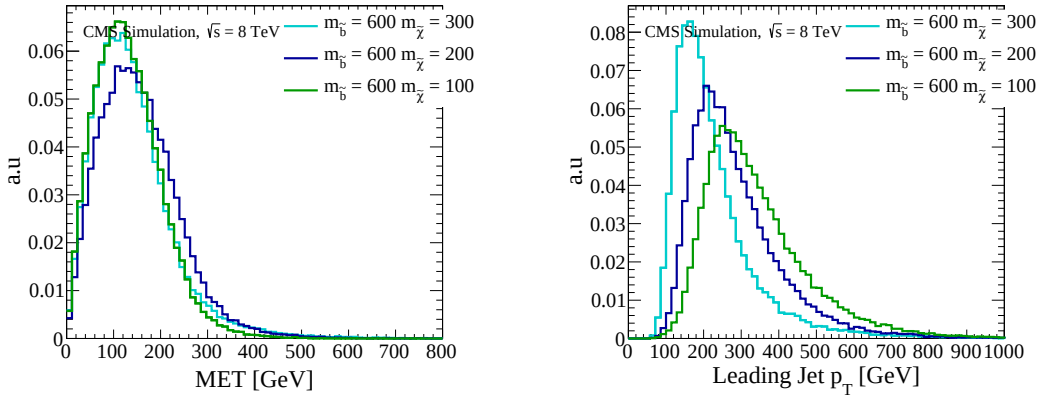
soft due to the relatively smaller mass splittings, which also minimizes the effect of the particle decay angles.

In summary: for the larger mass splitting models, the R^2 distribution is desirable but the M_R distribution is too high. We see a large effect from the two-step decay chain.



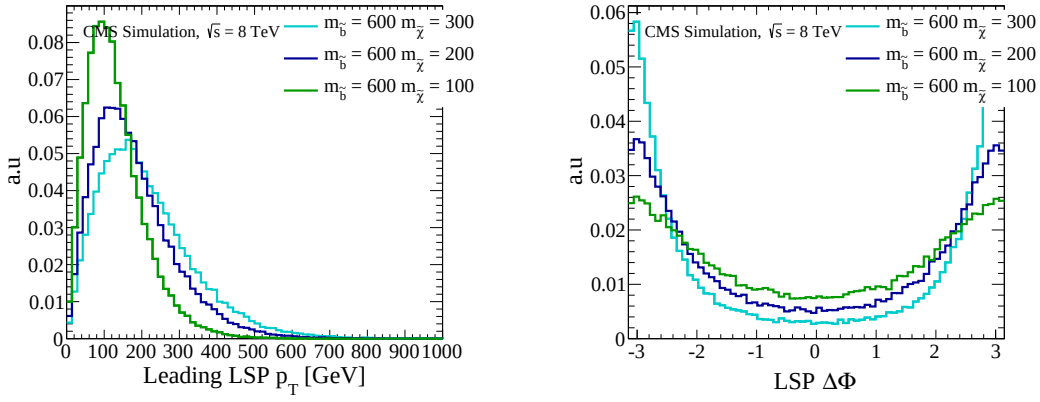
(a) (b)

Figure 4.14: Razor variable distributions for the various mass splitting with fixed $\tilde{b}$ mass. (a) M_R 1D distributions. (b) R^2 1D distributions.



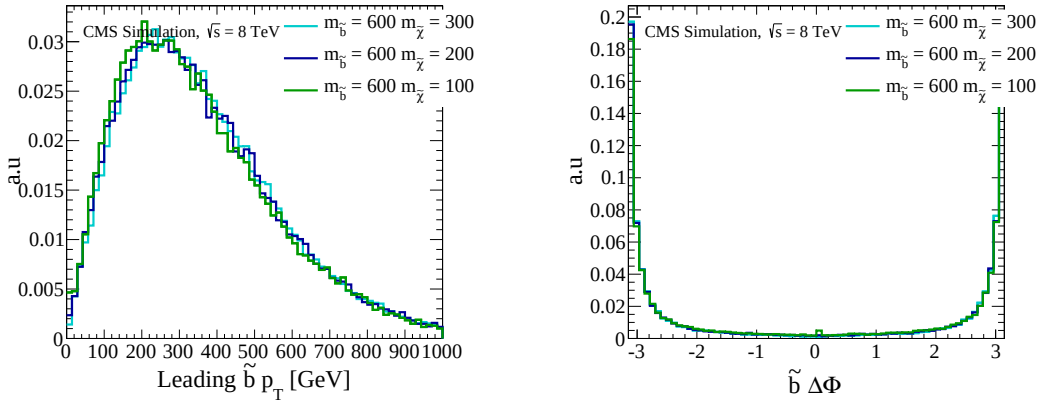
(a) (b)

Figure 4.15: Kinematic distributions for the various mass splitting with fixed $\tilde{b}$ mass. (a) $\not{E}_T$ 1D distributions. (b) Leading jet p_T 1D distributions (excluding H $\rightarrow$ $b\bar{b}$ jets).



(a) (b)

Figure 4.16: $\tilde{\chi}_1^0$ kinematic distributions for the various mass splitting with fixed $\tilde{b}$ mass. (a) Leading $\tilde{\chi}_1^0$ p_T 1D distributions. (b) $\Delta\Phi$ between the two $\tilde{\chi}_1^0$ particles.



(a) (b)

Figure 4.17: Bottom squark kinematic distributions for the various mass splitting with fixed $\tilde{b}$ mass. (a) Leading $\tilde{b}$ p_T 1D distributions. (b) $\Delta\Phi$ between the two bottom squarks particles.

4.9 $\tilde{b}$ Mass Study with Fixed M_Δ

In the following models, we instead vary the bottom squark mass while keeping M_Δ roughly the same. The model characteristics are given in Table 4.4.

Name	$m_{\tilde{b}}$	$m_{\tilde{\chi}_2^0}$	$m_{\tilde{\chi}_1^0}$	M_Δ	M_{split}	BR (to $\tilde{\chi}_2^0$)
CLT2015-AW1	470 GeV	230 GeV	100 GeV	448.72 GeV	370 GeV	100%
CLT2015-AW2	500 GeV	290 GeV	160 GeV	448.8 GeV	340 GeV	100%
CLT2015-AW3	600 GeV	430 GeV	300 GeV	450 GeV	300 GeV	100%

Table 4.4: Table of the models investigated in the bottom squark mass study, with their associated values of M_Δ and M_{split}.

We see an deficit of events in the R^2 tail in the smaller $\tilde{b}$ mass samples (Figures 4.18a, 4.19a, 4.21), as compared to the $m_{\tilde{b}} = 600$ GeV model. The R^2 versus M_R distribution for the $\tilde{b} = 470$ model exhibits very low, clustered R^2 events, as well as a tightly clustered R^2 versus $\not{E}_T$ distribution (see Figure 4.18b). As a result, this model is promising for our desired constraints. From this point, we will refer to it as the **CLT2015-AW1** model for convenience. The other two models, $m_{\tilde{b}} = 500$ GeV and $m_{\tilde{b}} = 600$ GeV, will be called **CLT2015-AW2** and **CLT2015-AW3**, respectively.

This small R^2 tail is likely a mixture of the phenomena from the previous two studies, since we have neither fixed $\tilde{b}$ nor fixed $\tilde{\chi}_1^0$ masses. However, the deficit of R^2 events seems to be more similar to the fixed $\tilde{b}$ mass study. More energy is available for the b quark (Figure 4.20), which is exhibited in the higher leading jet p_T if we compare Figures 4.23a and 4.10b, which increases the role of the angles of decay in the kinematics. We again have a $\tilde{\chi}_1^0$ $\Delta\Phi$ distribution that varies greatly between the different models (Figure 4.23d). The model with the smallest mass splitting, CLT2015-AW3, again has the most back-to-back $\tilde{\chi}_1^0$ particles. This confirms the dependence we saw in Section 4.8.

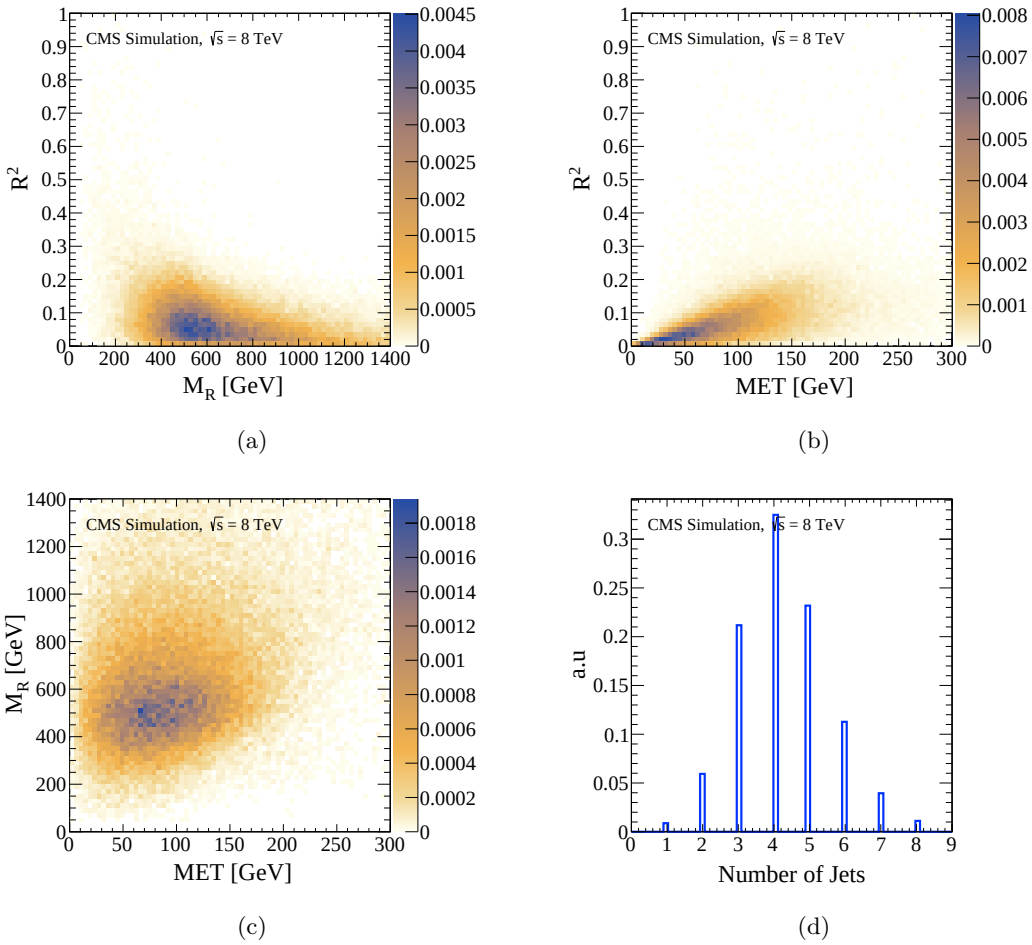


Figure 4.18: Characterization of the $m_{\tilde{b}} = 470$ GeV, $m_{\tilde{\chi}_1^0} = 100$ GeV model with 100% to $\tilde{\chi}_2^0$. Z-scale units are such that the total number of events is normalized to 1. (a) R^2 versus M_R. (b) R^2 versus $\not{E}_T$. (c) M_R versus $\not{E}_T$. (d) Number of PF jets, clustered with the anti-kT algorithm with ΔR = 0.5.

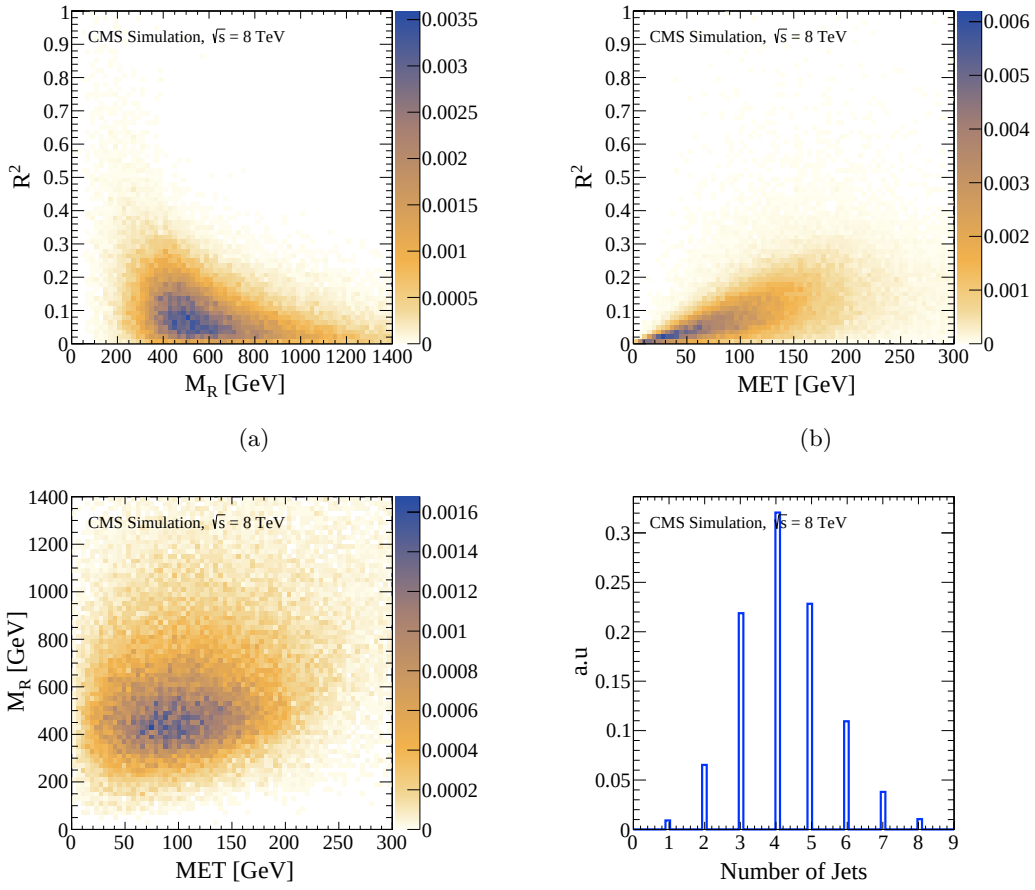


Figure 4.19: Characterization of the $m_{\tilde{b}} = 500$ GeV, $m_{\tilde{\chi}_1^0} = 160$ GeV model with 100% to $\tilde{\chi}_2^0$. Z-scale units are such that the total number of events is normalized to 1. (a) R^2 versus M_R. (b) R^2 versus $\not{E}_T$. (c) M_R versus $\not{E}_T$. (d) Number of PF jets, clustered with the anti-kT algorithm with ΔR = 0.5.

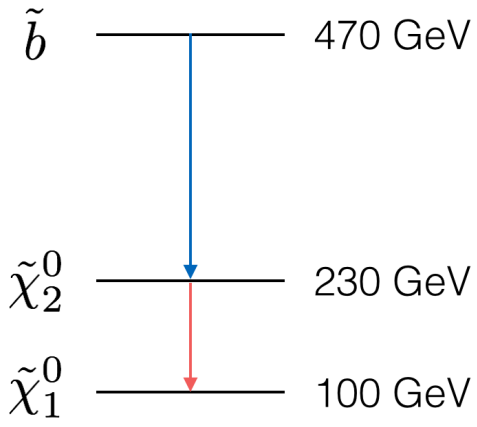
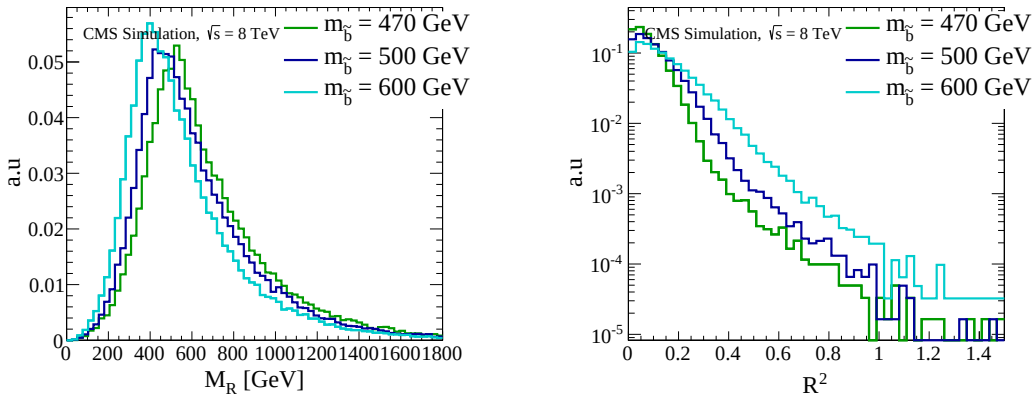
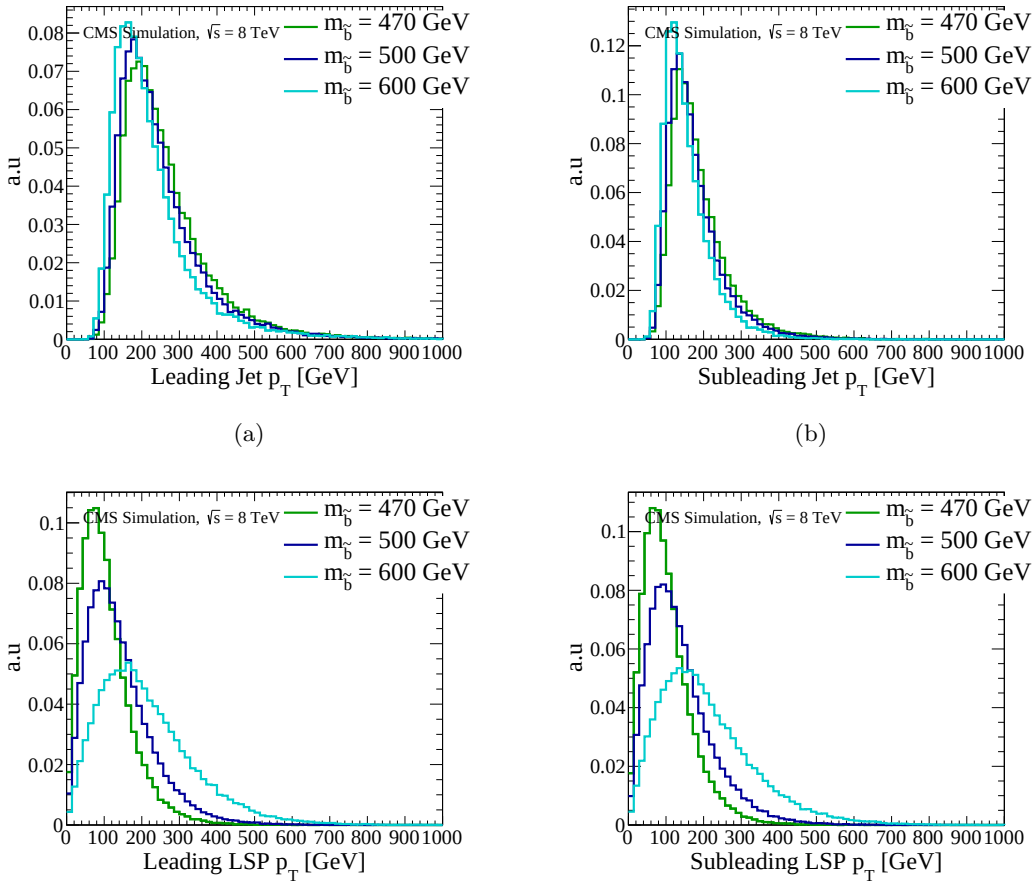


Figure 4.20: The decay spectrum for the $m_{\tilde{b}} = 470$ GeV model.



(a) (b)

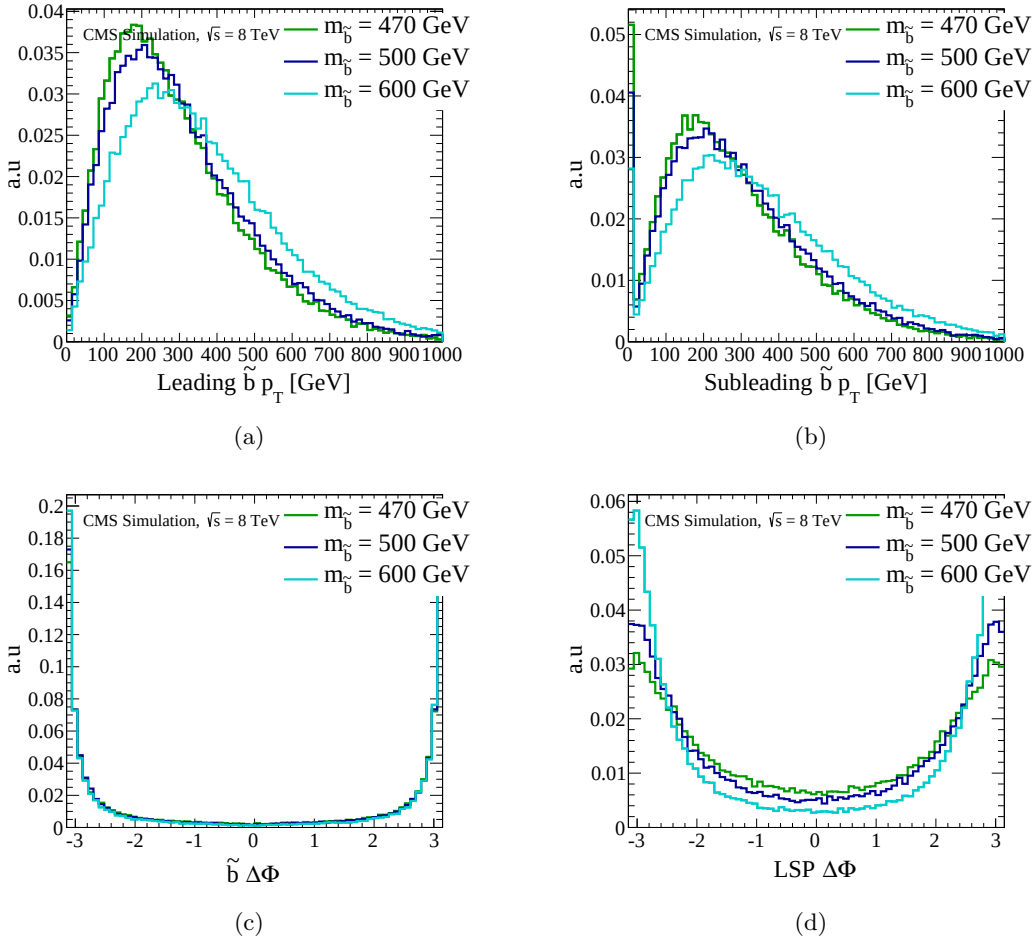
Figure 4.21: Razor variable distributions for the various $\tilde{b}$ models with fixed M_Δ. (a) M_R 1D distributions. (b) R^2 1D distributions.



(a) (b)

(c) (d)

Figure 4.22: Kinematic distributions for the various $\tilde{b}$ models with fixed M_Δ. (a) Leading jet p_T 1D distributions (excluding H $\rightarrow$ $b\bar{b}$ jets). (b) Subeading jet p_T 1D distributions (excluding H $\rightarrow$ $b\bar{b}$ jets). (c) Leading $\tilde{\chi}_1^0$ p_T 1D distributions. (d) Subeading $\tilde{\chi}_1^0$ p_T 1D distributions.



(a) (b) (c) (d)

Figure 4.23: Kinematic distributions for the various $\tilde{b}$ models with fixed M_{Δ}. (a) Leading $\tilde{b}$ p_T 1D distributions. (b) Subleading $\tilde{b}$ p_T 1D distributions. (c) $\Delta\Phi$ between the two bottom squarks particles. (d) $\Delta\Phi$ between the two $\tilde{\chi}_1^0$ particles.

4.10 Investigating the CLT2015-AW1 Model

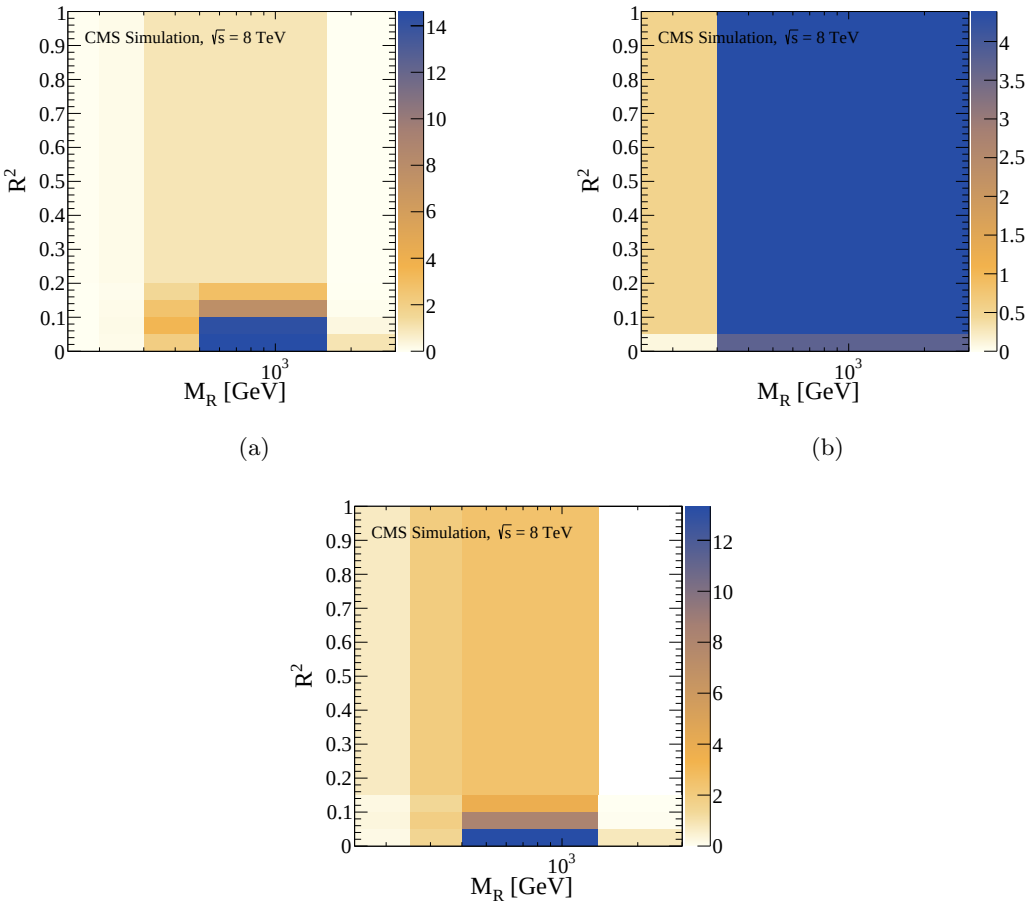


Figure 4.24: Event populations of the CLT2015-AW1 Model at $\sqrt{s} = 8$ TeV in each of the boxes, with the total amount of events passing the selection scaled to 100. (a) High p_T box (b) Hbb box (c) Overflow box.

We examine the event distribution in each of the boxes in Figure 4.24. The binning boundaries are again motivated by the Higgs to diphoton supersymmetry search [6]. They are as follows:

1. High p_T box: M_R: {150, 200, 300, 500, 1600, 3000} GeV. R^2: {0, 0.5, 0.1, 0.15,0 .2,1}.

2. Hbb box: M_R: {150, 300, 3000} GeV. R^2: {0, 0.5, 1}.

3. Overflow box: M_R: {150, 250, 400, 1400, 3000} GeV. R^2: {0, 0.05, 0.1, 0.15, 1}.

We see a high clustering of events in the $M_R \in [500, 1600]$ GeV and $R^2 \in [0, 0.1]$ region for the High p_T box, and in the $M_R \in [400, 1400]$ GeV and $R^2 \in [0, 0.05]$ region for the Overflow box. The decrease in events at high R^2 is apparent and ideal for complying with current bounds.

4.11 Counting Experiment

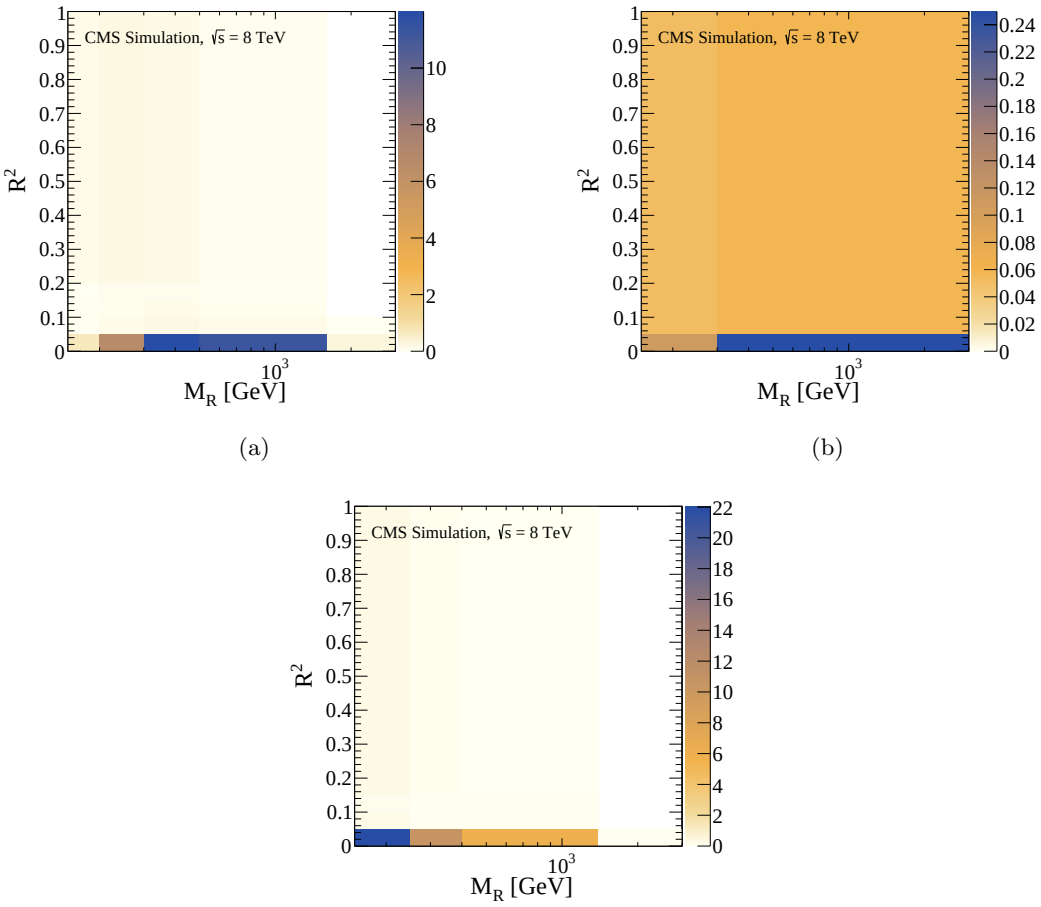


Figure 4.25: Event populations of the SM Higgs production background at $\sqrt{s} = 8$ TeV in each of the boxes, with the total amount of events passing the selection scaled to 100. (a) High p_T box (b) Hbb box (c) Overflow box.

We can do a counting experiment in order to estimate the cross section of the $m_{\tilde{b}} = 470$ GeV model needed in order for discovery at 8 TeV. We make the assumption that the main background is Standard Model Higgs background. Again, we generate SM Higgs events with PYTHIA 8.1 [34] [35] and then filter using the same selection process as before in Section 4.4. For consistency, we force the Higgs boson to decay to two photons.

At 13 TeV, the theoretical total cross section of all SM Higgs production is 22.10 pb [40]. If we also assume an integrated luminosity of 19.3 fb^{-1} for the data collected by CMS in 2012 [41]. We then scale the amount of events produced by the branching ratio of Higgs to two photons (2.28

$\times 10^{-3}$) to get the total number of expected events:

$$\sigma_{\text{SM Higgs}} \times \mathscr{L} \times \text{BR}_{\gamma\gamma}$$

$$= 22.1\, pb \times 19.3\, fb^{-1} \times 0.00228 = 972.5 \text{ events}$$

From the produced PYTHIA events, if we have P events passing out of N produced events, then we can calculate a selection efficiency with an error of $\delta e = e \times \sqrt{(\frac{\delta P}{P})^2 + (\frac{\delta N}{N})^2}$, where e is the selection efficiency. Out of 208505 events produced, 25230 passed, giving $e \approx 0.1210 \pm .0008066$, assuming Poisson errors. We then expect 117.7 ± 0.7844 events to pass the selection.

We then examine the Standard Model Higgs background events in the three boxes from Section 4.5. For the Overflow box, we have a selection efficiency from the total production of 0.04816 ± 0.00049. We can also calculate the selection efficiencies for each box (see Table 4.5). Then we get 46.84 ± 0.48 events in the Overflow box, 0.5550 ± 0.0509 events in the Hbb box, and 38.26 ± 0.43 events in the High p_T box (some events do not pass the M_R, R^2 definitions of the boxes).

Using a likelihood-based statistical test described in [42], we can assume a fixed number of background events and estimate the cross section of total bottom squark production needed for a discovery at 5σ for the fixed background. We will briefly summarize the method that is described in [42].

We assume a Poisson likelihood function of

$$L(\mu) = \frac{(\mu s + b)^n}{n!} e^{-(\mu s + b)},$$

where μ is described as a "signal strength" parameter, s is the number of expected signal events, b is the number of expected background events, and n is the number of observed events. The test statistic is

$$q_0 = \begin{cases} -2\, \ln \frac{L(0)}{L(\hat{\mu})} & \hat{\mu} \geq 0, \\ 0 & \hat{\mu} < 0. \end{cases}$$

Using asymptotic simplifying assumptions described in [42], the authors derive a median significance value of

$$\tilde{Z} = \sqrt{2 * ((s+b)\, \ln(1 + s/b) - s)}.$$

We then solve for the value of the signal that gives us $\tilde{Z} = 5$ given a fixed background.

We will focus on one box, the Overflow box, which would need 38.17 ± 0.18 signal events. The error on the number of signal events needed is most likely overestimated (ignoring systematic error) since we simplified the calculation by inputing b+ δb into our $\tilde{Z}$=5 equation as it cannot be solved explicitly. Our selection efficiency for signal events in the Overflow box is 0.2798 ± 0.0021 (derived

from MC simulation). If we divide the number of events by the Higgs to two photons branching ratio [39], and then by the signal efficiency $\times$ integrated luminosity, we round to two significant figures and arrive at a cross section of 3.1 ± 0.3 pb.

Box Name	**SM Higgs Efficiency**	**Error**	**Signal Efficiency**	**Error**
High p_T	0.03934	0.00044	0.4080	0.0027
Hbb	0.0005707	0.0000523	0.06748	0.00094
Overflow	0.04816	0.00049	0.2798	0.0021

Table 4.5: Selection efficiencies of the SM Higgs production background and the $m_{\tilde{b}} = 470$ GeV signal in the various boxes.

We can improve upon this cross section if we choose a more restricted box with lower background. If we restrict to a sum of bins in the Overflow box where we see a cluster of signal, i.e. $M_R \in [400, 1400]$ GeV and $R^2 \in [0.05, 1]$, we have an efficiency for SM Higgs background of 0.000149 ± 0.000027 with Poisson error. That corresponds to 0.145 ± 0.026 SM Higgs events. Then, using the same likelihood test as before, we require 4.76 ± 0.20 signal events in this defined bin. The signal selection efficiency is estimated as 0.1098 ± 0.0012. Again, dividing by the branching ratio, the integrated luminosity, and the signal efficiency, we get a corresponding cross section of 0.99 ± 0.04 pb.

The error on these cross sections are most likely greatly underestimated. In a future calculation we will consider the systematic errors more carefully, which would include the error on photon identification, jet energies, photon resolution, b-tagging efficiencies, etc. This is a generator-level study and gives a first estimate of the cross-section needed for discovery.

Assuming that the SM Higgs background predominates, the conclusion of this study is that $\sqrt{s}$ = 8 TeV CMS dataset requires a production cross section of 0.99 pb for a 5 sigma discovery. The predicted cross section for bottom squark pair production at 8 TeV is $0.128326 \pm 14.5144\%$ pb for a bottom squark of $m_{\tilde{b}} = 470$ GeV. This is a good indicator for the future at 13 and 14 TeV.

Chapter 5

Conclusion

The goal of this thesis was to examine the potential of using the Higgs boson as a tool in searching for new physics. To accomplish this, we investigated the production of bottom squarks that decay to a Higgs boson, where we varied different working parameters of the model. After tuning these parameters, we pinpointed a desirable model which would not likely have been discovered in Run I, but has a large potential for discovery in Run II. Further work would involve studying the model without an 100% branching ratio to the Higgs and instead consider the situation where a Z boson is sometimes produced. We would also like to reproduce the events with a full simulation of the detector.

Then, an understanding of the High Level Trigger was developed through the design of a trigger to select for events with $H \to b\bar{b}$. This complements the bottom squark model study extremely well since events that have one Higgs boson decaying to two photons will likely have the other Higgs decaying to two b quarks. As a result, the discussed model should also pass the designed trigger; it has > 3 jets, razor variable distributions that cluster at higher values than the trigger thresholds of $M_R > 300$ GeV and $R^2 > 0.0196$, and two Higgs bosons.

This model, CLT2015-AW1, is a non-standard model which does not rely on large $\not{E}_T$ for background reduction. Furthermore, the studies were conducted using a simplified model interpretation, which increases the utility of these studies. CLT2015-AW1 simply describes the production of a particle ($\tilde{p}_1$) largely produced in pairs, that decays $\tilde{p}_1 \to$b $\tilde{p}_2 \to$b H $\tilde{p}_3$, where $\tilde{p}_3$ is an invisible particle and $\tilde{p}_2$ is an intermediate particle.

Appendix A

A.1 Samples

The following tables list the simulated event samples produced by CMS internal members that were used for the trigger studies. These were produced using MadGraph 5 [43], PYTHIA 6 [34], and CMS Fast Simulation [44]. They were made available for members within CMS.

/QCD_Pt-30to50_Tune4C_13TeV_pythia8/Fall13dr-castor_tsg_PU40bx25_POSTLS162_V2-v1/GEN-SIM-RAW
/QCD_Pt-50to80_Tune4C_13TeV_pythia8/Fall13dr-castor_tsg_PU40bx25_POSTLS162_V2-v1/GEN-SIM-RAW
/QCD_Pt-80to120_Tune4C_13TeV_pythia8/Fall13dr-castor_tsg_PU40bx25_POSTLS162_V2-v1/GEN-SIM-RAW
/QCD_Pt-120to170_Tune4C_13TeV_pythia8/Fall13dr-castor_tsg_PU40bx25_POSTLS162_V2-v1/GEN-SIM-RAW
/QCD_Pt-170to300_Tune4C_13TeV_pythia8/Fall13dr-castor_tsg_PU40bx25_POSTLS162_V2-v1/GEN-SIM-RAW
/QCD_Pt-300to470_Tune4C_13TeV_pythia8/Fall13dr-castor_tsg_PU40bx25_POSTLS162_V2-v1/GEN-SIM-RAW
/QCD_Pt-470to600_Tune4C_13TeV_pythia8/Fall13dr-castor_tsg_PU40bx25_POSTLS162_V2-v1/GEN-SIM-RAW
/QCD_Pt-600to800_Tune4C_13TeV_pythia8/Fall13dr-castor_tsg_PU40bx25_POSTLS162_V2-v1/GEN-SIM-RAW
/QCD_Pt-800to1000_Tune4C_13TeV_pythia8/Fall13dr-castor_tsg_PU40bx25_POSTLS162_V2-v1/GEN-SIM-RAW

Table A.1: Data samples used to calculate trigger rates. These can be accessed through the data aggregation system (DAS) system.

/store/user/dalfonso/T2qq_400_150
/store/user/dalfonso/T2qq_450_400
/store/user/dalfonso/T2tt_800_100
/TT_Tune4C_13TeV-pythia8-tauola/Fall13dr-tsg_PU40bx50_POSTLS162_V2-v1/GEN-SIM-RAW

Table A.2: Data samples used to calculate L1 trigger efficiencies. The first three samples are unpublished to the DAS but may be accessed by CMS members with details located at [45]. The others can be accessed through the DAS system.

/RelValTTbarLepton_13/CMSSW_7_3_0-MCRUN2_73_V7-v1/GEN-SIM-DIGI-RAW-HLTDEBUG

Table A.3: Data samples used for the muon study. This can be accessed through the DAS system.

A.2 Bottom squark production mechanism

The following production mechanisms were considered in the PYTHIA event generation.

1. $g\ g \rightarrow \tilde{q}\ \tilde{q}$
2. $q\ \bar{q}' \rightarrow \tilde{q}\ \tilde{q}$
3. $q\ \bar{q}' \rightarrow \tilde{q}\ \tilde{q} + c.\ c.$
4. $q\ g \rightarrow \tilde{\chi}_1^0\ \tilde{q} + c.\ c.$
5. $q\ g \rightarrow \tilde{\chi}_2^0\ \tilde{q} + c.\ c.$
6. $q\ \bar{q}' \rightarrow \tilde{\chi}_1^0\ \tilde{\chi}_1^0$
7. $q\ \bar{q}' \rightarrow \tilde{\chi}_1^0\ \tilde{\chi}_2^0$
8. $q\ \bar{q}' \rightarrow \tilde{\chi}_2^0\ \tilde{\chi}_2^0$

The maximum cross sections for each process, in millibarns, outputed by PYTHIA, were: (1) 4.491×10^{-10}, (2) 2.228×10^{-10}, (3) 6.722×10^{-14}, (4) $2.227\ \times10^{-12}$, (5) 1.714×10^{-11}, (6) $3.285\ \times10^{-14}$, (7) $8.368\ \times10^{-14}$, and (8) $6.914\ \times10^{-14}$; the total is $6.91589\ \times10^{-10}$ mb.

Bibliography

[1] The ATLAS Collaboration. Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC. *Phys. Rev. B*, 716(1), 2012.

[2] The CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Rev. B*, 716(30), 2012.

[3] Bhaskar Dutta. Dark Matter Searches at Accelerator Facilities. *arXiv:1403.6217 [hep-ph]*, 2014.

[4] Stephen P. Martin. A Supersymmetry Primer. *arXiv:hep-ph/9709356v6*, 2011.

[5] The CMS Collaboration. Inclusive Search for Supersymmetry Using Razor Variables in pp collisions at $\sqrt{s} = 7$ tev. *Phys. Rev. Lett.*, 111(081802), 2013.

[6] Alexander Mott. Search for the Higgs boson production beyond the standard model using the razor kinematic variables in pp collisions at 8 TeV and optimization of Higgs boson identification using a quantum annealer, doctoral dissertation (unpublished). 2015.

[7] The CMS Collaboration. Inclusive search for squarks and gluinos in pp collisions at $\sqrt{s} = 7$ TeV. *Phys. Rev. D*, 85(012004), 2012.

[8] Maria Spiropulu and S. Stapnes. LHCs ATLAS and CMS Detectors. *Int. J. Mod. Phys. A*, 23:4081, 2008.

[9] The CMS Collaboration. Study of Pileup Removal Algorithms. *CMS-PAS-JME-14-001*, 2014.

[10] The CMS Collaboration. The CMS High Level Trigger. *Eur. Phys. J. C*, 46(605), 2006.

[11] Tomasz Bawej et al. The New CMS DAQ System for Run-2 of the LHC. *Real Time Conference (RT), 2014 19th IEEE-NPSS*, 2014.

[12] Danielle Trocino (on behalf of the CMS Collaboration). The CMS High Level Trigger. *20th Int. Conf. on Computing in High Energy and Nuclear Physics*, 513(012036), 2014.

[13] The CMS Collaboration. Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS. *Journal of Instrumentation*, 6(11002), 2011.

[14] M. Cacciari, G. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *Journal of High Energy Physics*, 0804(063), 2008.

[15] Florian Beaudette. The CMS Particle Flow. *arXiv:1401.8155 [hep-ex]*, 2014.

[16] The CMS Collaboration. CMS Offline Software. *https://github.com/cms-sw/cmssw*, 2015.

[17] The CMS Collaboration. Search for supersymmetry using razor variables in events with b-tagged jets in pp collisions at a $\sqrt{s} = 8$ TeV. *Phys. Rev. D*, 91(052018), 2015.

[18] LHC Higgs Cross Section Working Group. Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables. *CERN-2011-002*, 2011.

[19] Karl Jakobs. Physics at the LHC – From Standard Model measurements to Searches for New Physics. *arXiv:1206.7024 [hep-ex]*, 2012.

[20] The CMS Collaboration. Interpretation of searches for supersymmetry with simplified models. *Phys. Rev. D*, 88(052017), 2013.

[21] The CMS Collaboration. Identification of b-quark jets with the CMS experiment. *CMS-BTV-13-001*, 2013.

[22] E. Fernandez et al. Lifetime of Particles Containing b Quarks. *Phys. Rev. Lett.*, 51(1022), 1983.

[23] Tai Sakuma. Missing E_T Reconstruction with the CMS Detector. *Journal of Physics: Conference Series*, 404(012010), 2012.

[24] Artur Apresyan. 25ns reconstruction and OOT PU mitigation, presentation (unpublished). *https://indico.cern.ch/event/345149/session/1/contribution/7/material/slides/0.pdf*, 2015.

[25] Josh Bendavid and Emanuele Di Marco. ECAL OOT Pileup Mitigation with Template Fits in Pulse Reconstruction , presentation (unpublished). *https://indico.cern.ch/event/333476/contribution/5/material/slides/0.pdf*, 2014.

[26] Diptimoy Ghosh, Rohini Godbole, Monoranjan Guchait, Kirtimaan Mohan, and Dipan Sengupta. Looking for an invisible Higgs signal at the LHC. *Phys. Lett. B*, 725:344 – 351, 2013.

[27] Abdelhak Djouadi, Adam Falkowski, Yann Mambrini, and Jeremie Quevillon. Direct detection of Higgs-portal dark matter at the LHC. *Eur. Phys. J. C*, 73(2455), 2013.

[28] Maurizio Monaco, Maurizio Pierini, Andrea Romanino, and Martin Spinrath. Phenomenology of Minimal Unified Tree Level Gauge Mediation at the LHC. *arXiv:1302.1305 [hep-ph]*, 2013.

[29] The CMS Collaboration. Observation of the diphoton decay of the Higgs boson and measurement of its properties. *arXiv:1407.0558*, 2014.

[30] Didar Dobur. Jets and Missing Transverse Energy Reconstruction with CMS. *CMS CR-2009/030*, 2009.

[31] The CMS Collaboration. Search for supersymmetry in pp collisions at a center-of-mass energy of 8 TeV in events with two opposite sign leptons, large number of jets, b-tagged jets, and large missing transverse energy. *CMS PAS SUS-13-016*, 2013.

[32] The CMS Collaboration. Excess Higgs Production in Neutralino Decays. *Journal of High Energy Physics*, 1210(065), 2012.

[33] The CMS Collaboration. Search for the direct production of bottom squark pairs. *CMS-PAS-SUS-13-018*, 2014.

[34] T Sjorstrand, S Mrenna, and P Skands. PYTHIA 6.4 Physics and Manual. *Journal of High Energy Physics*, 0605(26), 2006.

[35] T Sjorstrand, S Mrenna, and P Skands. A Brief Introduction to PYTHIA 8.1. *Comp. Phys. Comm.*, 178(852), 2008.

[36] P. Skands et al. SUSY Les Houches Accord: Interfacing SUSY Spectrum Calculator, Decay Packages, and Event Generators. *Journal of High Energy Physics*, 0407(036), 2004.

[37] M. Pierini. Tools for Beyond Standard Model event generation and analysis at LHC and other colliders. *https://github.com/pierinim/BSMatLHC*, 2014.

[38] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet user manual. *arXiv:1111.6097*, 2011.

[39] J. Beringer et al. (Particle Data Group). Status of Higgs Boson Physics. *Phys. Rev. D*, 48(010001), 2014.

[40] The LHC Higgs Cross Section Working Group. Handbook of LHC Cross Sections: 3. Higgs Properties. *CERN-2013-004*, 2013.

[41] The CMS Collaboration. Search for supersymmetry in pp collisions at $\sqrt{s} = 8$ TeV in events with a single lepton, large jet multiplicity, and multiple b jets. *Phys. Lett. B*, 733(328), 2014.

[42] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. Journ.*, 71(1554), 2014.

[43] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, 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. *arXiv:1405.0301 [hep-ph]*, 2014.

[44] Douglas Orbaker. Fast Simulation of the CMS Detector. *Journal of Physics: Conference Series*, 219(032053), 2010.

[45] Maria D'Alfonso. SUSY MC Private Production, unpublished. *https://twiki.cern.ch/twiki/bin/viewauth/CMS/SUSYMCprivateProduction8TeVFall13*, 2014.