UNIVERSITY OF CALIFORNIA
RIVERSIDE

A Search for A Simplified Model of Supersymmetry With Light Sbottoms and Higgsinos Using Higgs to Diphoton Decays

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Physics

by

Kira Burt

June 2017

Dissertation Committee:

Professor Gail Hanson, Chairperson
Professor John Ellison
Professor Jose Wudka

The Dissertation of Kira Burt is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

I would like to thank my advisor, Gail Hanson, for opportunities provided through the years. I am also indebted to Anthony Barker and Yuri Gershtein for laying the groundwork for this research in their previous analysis. Finally, a massive thank you to everyone I worked with on the FPIX system. It was truly the experience of a lifetime and I am fortunate to have met and worked with so many talented scientists and fine people.

For family, especially Mom. I couldn't have done this without you.

ABSTRACT OF THE DISSERTATION

A Search for A Simplified Model of Supersymmetry With Light Sbottoms and Higgsinos Using Higgs to Diphoton Decays

by

Kira Burt

Doctor of Philosophy, Graduate Program in Physics
University of California, Riverside, June 2017
Professor Gail Hanson, Chairperson

We present a search for a simplified supersymmetric model with pair-produced light supersymmetric bottom quarks decaying to neutralinos. Higgs-type neutralinos (Higgsinos) decay to the Higgs boson and LSP, with at least one Higgs boson decaying to a diphoton system. Events with at least two b-jets and a photon pair in the Higgs-tagged invariant mass window are considered. In 36.2 fb^{-1} of proton-proton collision data collected at the CMS experiment at $\sqrt{s} = 13$ TeV, we find no evidence of signal and set lower limits on the production of the bottom squark at a 95% confidence level at masses of below 350 GeV, with a Higgsino mass of 150 GeV or less.

Contents

List of Figures

List of Tables

Chapter 1

The Standard Model

The Standard Model of particle physics represents our current understanding of elementary particles and the forces that govern their interactions. This description of the most basic constituents of nature is robust and well-tested but as yet incomplete.

The history of the Standard Model began with the discovery of the electron by Thomson in 1897, and continued through the 20th and 21st centuries as scientists probed deeper into the structure of the atom, the nucleus, and finally the nucleons themselves. Atomic substructure was first discovered by Rutherford in 1913, when the "plum-pudding" model of the atom was replaced by the nucleus-and-electrons theory. Subsequent discoveries included the proton, also due to Rutherford; the neutron, proposed by Rutherford and discovered in 1931 by Chadwick; and the positron, antiparticle to the electron and the first evidence of the existence of antimatter, in 1932 by Anderson. The 20th century eventually saw the discovery of all three generations of the quark families, the massive leptons and neutrinos, and the vector bosons, as pictured in Figure 1.1.

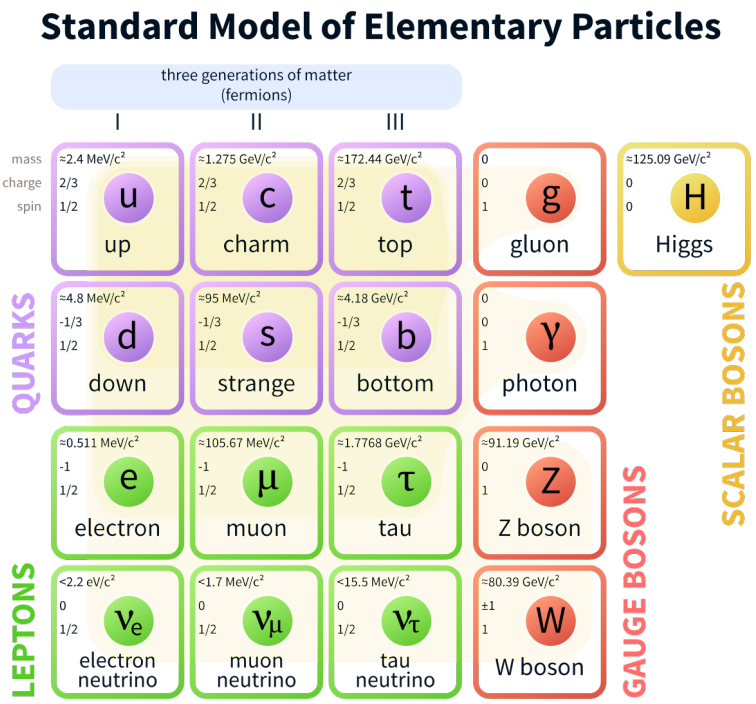


Figure 1.1: The Standard Model of particle physics [22].

1.1 Elementary Particles

Matter, at its most fundamental level, is composed of quarks and leptons. Though their interactions are governed by different forces, quarks and leptons share some similarities. Both particle types carry spin in half-integer units (making them fermions, which are defined as having half-integer spin, versus bosons which have integer spin) and are sorted into three generations based on mass. The lightest of these generations is comprised of the up and down quarks and the electron and its neutrino. The next generation contains the charm and strange quarks and the muon and its neutrino, and the final and heaviest generation is made up by the top and bottom quarks in addition to the tau lepton and its neutrino.

The division of each generation into two quark types is based on the distribution of

isospin and fractional charge among the quarks. Both isospin and hypercharge have strong and weak versions, with isospin **I** a property of quarks only, and weak isospin **T** a property of quarks and leptons. In both cases, isospin and hypercharge combine to yield electric charge, which is a conserved quantity in both strong and weak interactions. Hypercharge refers to the flavor quantum numbers of quarks and is detailed in Eq. 1.1. In the case of both **I** and **T**, only one component of the isospin vector (I_3 or T_3) specifies the flavor state of a multiplet.

$$Y = B + S + C + T + B' \tag{1.1}$$

Term-by-term, the hypercharge equation breaks down to: B = baryon number, S = strangeness, C = charmness, B' = bottomness and T = topness. An individual quark has baryon number 1/3 regardless of its other quantum numbers.

Hypercharge and isospin (up and down quarks form an isospin doublet with $I_3 = \pm\frac{1}{2}$, with other quarks having $I_3 = 0$) combine to yield the total electric charge of the particle:

$$Q = I_3 + \frac{Y}{2} \tag{1.2}$$

Up- and down-type quarks form weak isospin doublets according to their generations, with up-type (u, c, t) quarks having $T_3 = +\frac{1}{2}$ and down-type quarks (d, s, b) having $T_3 = -\frac{1}{2}$. Leptons form a similar set of three doublets, with the neutrinos possessing $T_3 = +\frac{1}{2}$ while the electron, muon and tau possess $T_3 = -\frac{1}{2}$. Weak isospin and another quantum number, weak hypercharge (Y_W), add to form the familiar electric charge Q as

Quark	**Mass (GeV)**	**Electric Charge (Q/e)**	**Flavor**
u	2.3×10^{-3}	$+\frac{2}{3}$	no flavor, $I_3 = +\frac{1}{2}$
d	4.8×10^{-3}	$-\frac{1}{3}$	no flavor, $I_3 = -\frac{1}{2}$
s	9.5×10^{-2}	$-\frac{1}{3}$	$S = -1$
c	1.275	$+\frac{2}{3}$	$C = 1$
b	4.18	$-\frac{1}{3}$	$B' = -1$
t	173.5	$+\frac{2}{3}$	$T = 1$
Lepton	**Mass (GeV)**	**Electric Charge (Q/e)**	**Flavor**
e	0.511×10^{-3}	-1	$L_e = 1$
$\bar{\nu_e}$	0	0	$L_e = -1$
μ	186×10^{-3}	-1	$L_\mu = 1$
$\bar{\nu_\mu}$	0	0	$L_\mu = -1$
τ	1.77	-1	$L_\tau = 1$
$\bar{\nu_\tau}$	0	0	$L_\tau = -1$

Table 1.1: Quark and lepton properties.

follows:

$$Q = T_3 + \frac{Y_W}{2} \tag{1.3}$$

Other important properties of the quarks and leptons, including their mass, electric charge, and flavor quantum numbers, are detailed in Table 1.1.

Quarks may combine to form composite particles. The most common constituents of everyday matter, excluding the electron, are the proton and neutron, which are baryons - fermions composed of quarks. Quarks may also combine to form mesons, which have integer spin and are shorter-lived than the proton and neutron - indeed the proton decay time is currently computed to be longer than the lifetime of the universe at 10^{33} years. In general a particle will decay to another set of lighter particles unless it is forbidden by some conservation law. Examples of these laws include energy and angular momentum, which hold at macro scales as well.

On the scale of elementary particles, interactions among quarks conserve baryon number, which is ± 1 for baryons and 0 for mesons. Total lepton number $L = n_l - n_{\bar{l}}$ is also conserved in interactions, with each family also possessing its own (electronic, muonic, tauonic) lepton number. Leptons have $n_l = 1$ and antileptons $n_l = -1$. Prior to the discovery of neutrino flavor oscillation, the lepton family number was considered to be a conserved quantity. The solar neutrino deficit, observed in the late 1960s, was eventually explained to be oscillations in neutrino flavor, demonstrating that neutrinos in fact have a very small nonzero mass. Neutrino mass and flavor eigenstates mix, and the oscillation of flavor shows a breaking of lepton family symmetry.

1.2 Interactions

Interactions between particles are governed by gauge fields. The quanta of these fields, also called "force carriers", are the vector bosons. All have spin 1 (distinct from the scalar Higgs, which has spin 0). Properties of the force-carrying particles are detailed in Table 1.2.

The quarks and leptons differ in the forces to which they can couple. Both particle types interact via the electromagnetic and the weak force: decays from one generation to other, less massive generations within a quark or lepton family are mediated by the weak force via its vector bosons the $W^{\pm}$. An example of a charged-current interaction - the beta decay of a neutron - is shown in Figure 1.2. The weak force also allows for non-flavor-changing or neutral current decays via the chargeless Z^0, which mediates elastic scattering interactions. The weak force couples only to fermions of negative chirality - left-handed

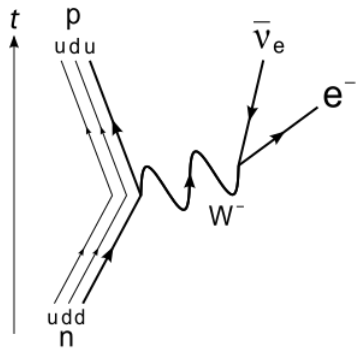


Figure 1.2: A neutron decays into a proton, electron, and electron anti-neutrino.

fermions and right-handed antifermions.

Both the weak vector bosons are massive and therefore relatively short-lived. Quarks and leptons also interact via the electromagnetic force, which acts on charged particles and is mediated by the massless, electrically neutral photon (γ).

Quarks are unique in that they interact with each other via Quantum Chromodynamics (QCD), colloquially known as the "strong" or "nuclear" force. The strong sector is an $SU(3)$ gauge group, and its carrier is the massless gluon, which is electrically neutral but has intrinsic "color" charge.

There are three color charges - red, green, and blue - and each has a corresponding anticolor. Naturally, these quantities do not represent real-world colors in terms of light wavelength, but are a convenient way to represent allowed mixing states of quarks and gluons.

Stable, longer-lived particles such as protons and neutrons exist as color singlets, or are "colorless", with their color state represented as in Eq. 1.4.

$$(r\bar{r} + b\bar{b} + g\bar{g})/\sqrt{3} \tag{1.4}$$

Boson	Mass (GeV)	Electric Charge (Q/e)	Spin
γ	0	0	1
gluon	0	0	1
$W^{\pm}$	80.4	± 1	1
Z^0	90.2	0	1
H^0	125.0	0	0

Table 1.2: Gauge bosons.

Mesons, generally made up of quark-antiquark pairs ($q_i\bar{q}_j$), are color-neutral and combine as $r\bar{r}$, $b\bar{b}$, or $g\bar{g}$. The photon carries no electric charge and so cannot self-interact, but the gluon carries its own color charge, allowing for self-interaction and effectively shortening the range of the strong force to sub-nuclear levels. This restriction on the range of interaction is known as "color confinement".

1.3 Electroweak Symmetry Breaking

At higher energies the electromagnetic and weak forces are unified into the electroweak sector, an $SU(2) \times U(1)$ group, with three gauge bosons of weak isospin (the $W^{\pm}$ and W^0 bosons) and one of weak hypercharge (the B boson). When the Higgs field acquires a vacuum expectation value, three degrees of freedom of the field combine with the gauge bosons, which acquire mass and become the weak gauge bosons ($W^{\pm}$, Z^0), with the Z^0 acquiring components of the neutral weak isospin boson W^0 and the weak hypercharge boson B as follows:

$$Z^0 = W^0 - \frac{B}{2} \tag{1.5}$$

The fourth degree of freedom is then the Standard Model Higgs boson, a scalar

(zero-spin) boson. The photon remains as the massless carrier of the electromagnetic field - an unbroken (since it does not interact with the Higgs) $U(1)$ group represented as a linear combination of the neutral W^0 and B bosons:

$$\gamma = W^0 + \frac{B}{2} \tag{1.6}$$

1.4 Beyond the Standard Model

While the Standard Model offers a robust description of natural processes at the fundamental level, it is not complete. Thus far no fundamental particles have been discovered at higher than the electroweak scale, with the heaviest known particle being the top quark at 172.04 GeV. At what is known as the Planck scale, $M_P = 2.48 \times 10^{18}$ GeV, quantum gravitational effects are expected to contribute, but between the electroweak and Planck scales there exists a "particle desert" some 10^{16} GeV across.

Additionally, the low mass of the scalar Higgs particle is difficult to reconcile with the possible existence of particles in this "desert". The Standard Model as it is today only allows for the observed low H^0 mass if a large number of loop corrections are carried out to its mass calculation, and then subsequently individually cancelled. This is contrary to assumptions of naturalness in physical theory, wherein the fine-tuning of parameters is limited as much as possible. In addition to the problematic H^0 mass, the difficulty of identifying the particle constituents of dark matter has puzzled physicists for decades. A possible solution to all these dilemmas is posited in the introduction of a new symmetry to the physical laws, known as supersymmetry, which is broken at higher energy scales than

are currently accessible by experiment.

Chapter 2

Supersymmetry

2.1 The Minimal Supersymmetric Standard Model - MSSM

Problematic loop corrections to the Higgs boson mass term m_H^2 arise due to the fact that every particle that couples to the Higgs field contributes such a correction. Should we make the reasonable assumption that some particles exist in the unexplored energies between the electroweak and Planck scales, the corrections due to these high-mass particles would additionally contribute directly to the final measured value of m_H^2. Why, then, is this measured mass so low? Either these particles do not exist, or there is some cancellation for every loop term that adds to m_H^2. Such one-to-one cancellations would imply the existence of a new physical symmetry. Supersymmetry applies such cancellations by relating bosons and fermions on a fundamental level, due to the difference in sign on fermion versus boson contributions to the m_H^2 term seen in Eqs. 2.1 [21] and 2.2 [21]. The Λ_{UV} in both equations is an ultraviolet momentum cutoff term that regulates the loop integral, and represents the energy scale at which new physics is likely to manifest. Eq. 2.1 refers to the loop mass

correction of fermions with mass m_f, and Eq. 2.2 shows the correction term for a heavy scalar particle.

$$\Delta m_H^2 = -\frac{|\lambda_F^2|}{8\pi^2}\Lambda_{UV}^2 + ... \tag{2.1}$$

$$\Delta m_H^2 = \frac{\lambda_S}{16\pi^2}\left[\Lambda_{UV}^2 - 2m_S^2 ln(\frac{\Lambda_{UV}}{m_S}) + ...\right] \tag{2.2}$$

The assumption of a symmetry between fermions and bosons neatly cancels all such loop corrections, and gives a sensible output for m_H^2 compatible with the measured value. An example particle list for a supersymmetric model is shown in Figure 2.1.

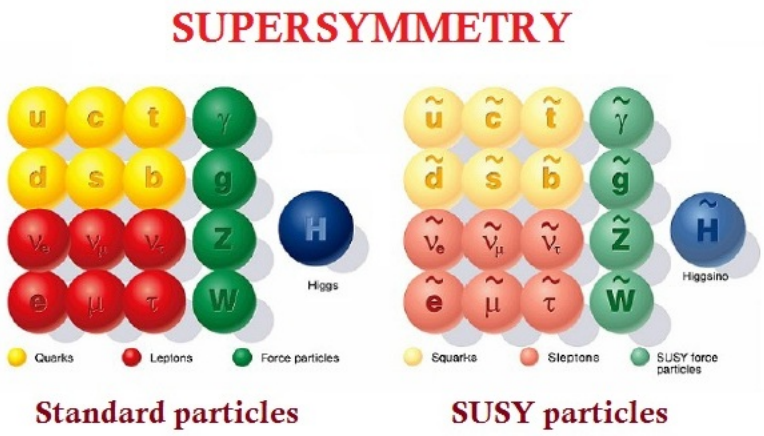


Figure 2.1: The Standard Model particles with their supersymmetric partners [39].

2.1.1 Supermultiplets

The supersymmetry operator must therefore change a fermionic state to a bosonic state, and vice versa. Since these states differ in their spin angular momentum, supersymmetry must be a symmetry of spacetime. Each set of fermionic and bosonic superpartners is

known as a supermultiplet. Members of a supermultiplet have the same weak isospin values, electric charge, and color charge due to the commutation of the supersymmetry operators with the gauge generators. Therefore each chiral supermultiplet of the MSSM will have a complete gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$. Squarks and sleptons, being spin-0 particles, do not themselves possess helicity, but are sorted by the chiral group to which their Standard Model partner belongs, and may participate in the electroweak interaction according to their handedness.

Higgsinos present a unique challenge in that they are fermions and must possess weak hypercharge $Y = \pm 1/2$. If only a single chiral supermultiplet for the Higgsino existed, the electroweak symmetry would undergo a gauge anomaly with behavior not currently exhibited by the Standard Model. Therefore a pair of chiral supermultiplets must exist - one for each sign of weak hypercharge. This pair corresponds to the up-type and down-type higgsino multiplets, denoted $\tilde{H}_u$ and $\tilde{H}_d$ [21]. The electrically neutral spin-zero scalar observed from the Standard Model is a linear combination of H_u and H_d. The left-handed squark and slepton and the higgsino chiral supermultiplets are shown in Table 2.1. Each entry for quarks and leptons in Table 2.1 presupposes the three known families - therefore each quark possesses a squark, e.g. a top has a "stop" superpartner and a tau has its "stau". Gauge supermultiplets are shown in Table 2.2.

Particle Type	**Spin 1/2**	**Spin 0**
quarks and squarks	u_L, d_L	$\tilde{u_L}, \tilde{d_L}$
leptons and sleptons	ν, e_L	$\tilde{\nu}, \tilde{e_L}$
higgsinos and Higgs	$\tilde{H_u^+}, \tilde{H_u^0}$	H_u^+, H_u^0
	$\tilde{H_d^-}, \tilde{H_d^0}$	H_d^-, H_d^0

Table 2.1: Left-handed MSSM chiral supermultiplets

Particle Type	Spin 1	Spin 1/2
gluons and gluinos	g	$\tilde{g}$
W bosons and winos	$W^{\pm}, W^0$	$\tilde{W^{\pm}}, \tilde{W^0}$
B boson and bino	B^0	$\tilde{B^0}$

Table 2.2: MSSM gauge supermultiplets

The chiral and gauge supermultiplets may mix - higgsinos may mix with gauginos, for example - when electroweak symmetry breaking effects are taken into account.

It is evident that supersymmetry is a broken symmetry, due to the fact that superpartners have not been observed at current energy scales - for example, no smuons or selectrons have been observed at electroweak energies.

2.1.2 R-parity

One important facet of the MSSM is a quantity known as R-parity. The conservation of this quantity determines the number of supersymmetric particles in an event as well as its final decay products. R-parity is defined in terms of baryon and lepton number, as well as particle spin:

$$P_R = (-1)^{3(B-L)+2s} \tag{2.3}$$

R-parity (alternatively dubbed "matter parity") allows for conservation of baryon and lepton number in ordinary interactions. For example, proton decay would violate both B and L and has as yet not been observed. Sparticles possess $P_R = -1$, while Standard Model particles have $P_R = +1$. R-parity imposes several important phenomenological constraints. It demands that any supersymmetric particle that is not the LSP (Lightest

Supersymmetric Particle) must decay into an odd number of LSPs; non-LSP sparticles must be produced in even numbers; and the LSP itself must be stable. The particle content of dark matter has yet to be determined. A massive, chargeless LSP would be absolutely stable and interact only via gravitational forces, making it an excellent candidate for dark matter.

2.2 Naturalness, Light Squarks, and Light Neutralinos

Many SUSY models rely on assumptions of "naturalness" [23] to choose their parameters. Simply put, a natural theory does not have an overabundance of fine-tuning. A constraint on fine-tuning is exemplified in the MSSM as follows:

$$-\frac{m_z^2}{2} = |\mu|^2 + m_{H_U}^2 \tag{2.4}$$

The higgsino mass is directly controlled by the μ term, while the stop and gluino correct the $m_{H_U}^2$ term at one and two-loop levels respectively. Therefore the masses of the higgsino, stop and gluino must not be too heavy in order to avoid counter-tuning in these terms in order to keep with the known electroweak mass scale given by $\frac{m_{z^2}}{2}$. Left-handed sbottoms must also be light when naturalness is assumed, due to the weak isospin coupling between the top and bottom quarks.

2.3 Simplified Model Scan and Event Signature

Many searches for supersymmetry at the LHC use what is known as a "Simplified Model" [7] wherein a set of hypothetical particles (in our case the sbottom, chargino, and

neutralino) are specified along with their masses and decays. These signal events are produced, and their detector acceptance and selection efficiency are used to generate a signal cross section and branching ratio for the simplified process, which is then employed to set a confidence limit on production of the process in question. This limit can then be interpreted in terms of different theories that share the same basic phenomenology, or used to verify theoretical predictions against data.

2.3.1 Event Signature

Our search focuses on light sbottoms and higgsinos. The simplified model for this scan is known as T2bH - two sbottoms directly produced from the proton-proton collision decay into a bottom quark and neutral higgsino (or a top quark and neutral higgsino if the mass splitting allows, with the top decaying further into a down-type quark and W^+ boson) with the neutral higgsino decaying into a Higgs boson and a final, massive LSP as shown in Fig. 2.2. The varying parameters in the simplified model are the masses of the sbottom and of the LSP, with $m_{\chi_2^0} - m_{\chi_1^0} = 130$ GeV, near the SM Higgs mass. All other SUSY particles are decoupled from the interaction, with masses above 1 TeV.

Our model assumes a branching fraction of 100% higgsino to higgs, stemming from the phenomological assignment of a low value of tanβ as seen in Figure 2.3, with

$$\tan\beta = v_u/v_d \tag{2.5}$$

the ratio of the vacuum expectation values of the Higgs doublets.

The final state in our search exploits the narrow mass resonance of the H $\rightarrow \gamma\gamma$

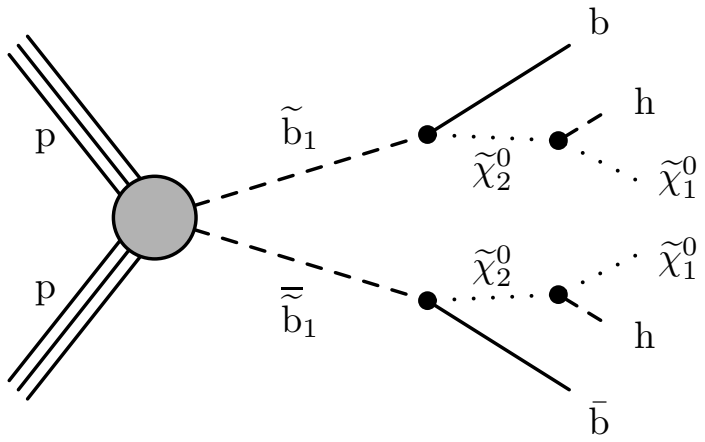


Figure 2.2: The T2bH simplified model. Our search requires one of the Higgs bosons to decay specifically to a pair of photons, while the second Higgs decays purely according to branching ratio.

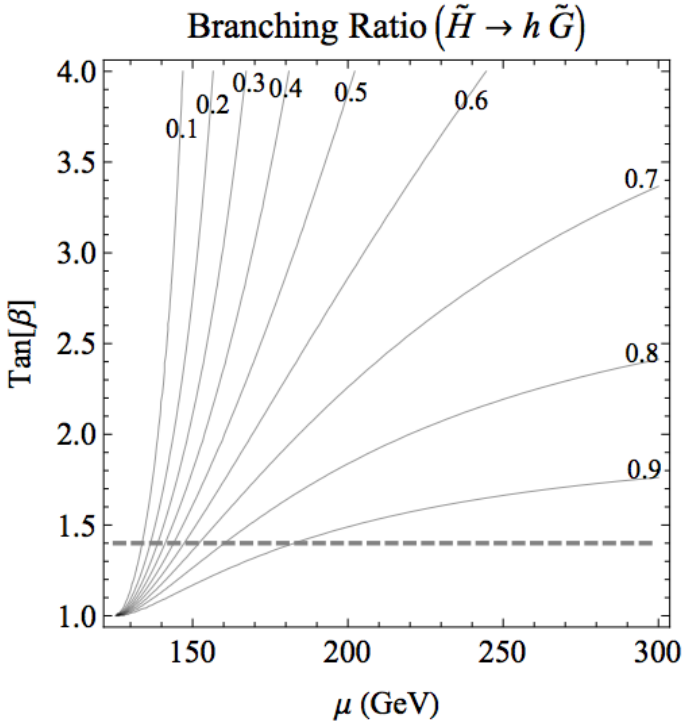


Figure 2.3: Branching ratio of a higgsino NLSP to a Higgs and a gravitino LSP as a function of higgsino mass and tanβ [15].

decay by requiring one of the Higgs to decay to a pair of photons. The H $\rightarrow \gamma\gamma$ decay has a low branching ratio at only 0.23% of total Higgs decays, but the low diphoton background compensates adequately.

The second Higgs may decay freely. Up to four b-quarks are expected per event - two from the sbottom, and potentially two from a H $\rightarrow b\bar{b}$ decay.

2.3.2 Previous and Complementary Searches

Much of the strategy in this analysis is adapted from a Run 1 search based around light stops rather than sbottoms - however, the identical final event signature and diphoton mass channel allows for the sideband background estimation method to be used [9]. This search was performed using the full Run 1 dataset at $\sqrt{s} = 8$TeV.

A complementary search using the T2bH SMS and H $\rightarrow \gamma\gamma$ decay has been performed using a divergent search strategy - no reliance on b-jet multiplicity and instead an emphasis on the resolution of the diphoton invariant mass [10]. This search was performed on the Run 2 "Moriond" dataset.

2.4 Event Simulation

Signal events are simulated at the collision level by MADGRAPH5 [1]. MADGRAPH5 accepts parameters such as particle masses and mixing information, calculates cross sections and event content at leading or next-to-leading order, and finally outputs the hard products of the proton-proton collision. These partons are fragmented and their showering is modeled by PYTHIAv8.2 [31]. Once simulation is complete, the output is reconstructed via the usual CMS detector reconstruction algorithms.

Simulated Monte Carlo samples are produced both for signal events and for SM Higgs boson backgrounds. SM Higgs samples from $t\bar{t}$H, $b\bar{b}$H, VBF (vector boson fusion), VH (associated production with a W or Z boson), and gluon-gluon fusion are produced using the MADGRAPH5_amc@NLOv2.2 generator at next-to-leading order, and showered by PYTHIAv8.2. Cross sections for SM Higgs production are obtained from the Yellow Report 4 of the LHC Higgs Cross Section Group [19]. The signal production cross sections are calculated to next-to-leading order plus next-to-leading-logarithm accuracy. Limits on the possible masses of the sbottom and LSP are derived using these cross sections and their uncertainties in Chapter 10.

Chapter 3

LHC

The Large Hadron Collider, or LHC, accelerates, circulates and collides protons at speeds close to light in a double beam ring nearly 27 km in circumference. It was built, with ground broken in 2001, in the tunnel originally constructed for CERN's LEP (Large Electron-Positron) Collider. This tunnel is ~ 100 meters underground, and passes beneath the border between France and Switzerland multiple times. Prior to the construction of the Channel Tunnel, the LEP tunnel was the largest civil engineering project in Europe.

In the main ring of the LHC, protons, circulating in opposing beams, are fed into four main interaction points (IPs) at an energy of 6.5 TeV per beam, with a final center-of-mass collision energy of 13 TeV. The LHC is currently operating at its design luminosity of 1×10^{34} cm^{-2}s^{-1} after substantial upgrade work performed in the Long Shutdown period of 2013-2014. Beam parameters at time of collision may be seen in Table 3.1.

Proton energy	GeV	6500
RMS bunch length	cm	7.5
Stored energy per beam	MJ	362
Number of particles per bunch	1.5×10^{11}	
Number of bunches	2808	
Beam size at P1 and P5	μm	16.7

Table 3.1: Beam parameters at collision, adapted from [5]

3.1 Proton Production and Initial Boosting

The instantaneous luminosity delivered by the collider is related the number of events produced by the LHC per second as in Eq. 3.1, where σ_{event} represents the production cross-section of the event type under study.

$$N = L\sigma_{event} \tag{3.1}$$

Protons are circulated around the LHC in bunches, with a nominal proton density per bunch of $N_{bunch} = 1.15 \times 10^{11}$. To achieve maximum beam intensity, each beam contains 2808 proton bunches with a bunch spacing (time between collisions) of 25 ns. In order to achieve the final energy of 6.5 TeV per beam, substantial acceleration of the protons is required before they reach the LHC ring, in what as known as the "injection chain". Details on the accelerator complex and injection chain, including the proposed LINAC4 upgrade, can be seen in Figure 3.1.

3.1.1 Proton Production

To produce the protons used in the collisions, hydrogen gas is injected into the Duoplasmatron Proton Ion Source, where electrons are stripped from the atoms in an

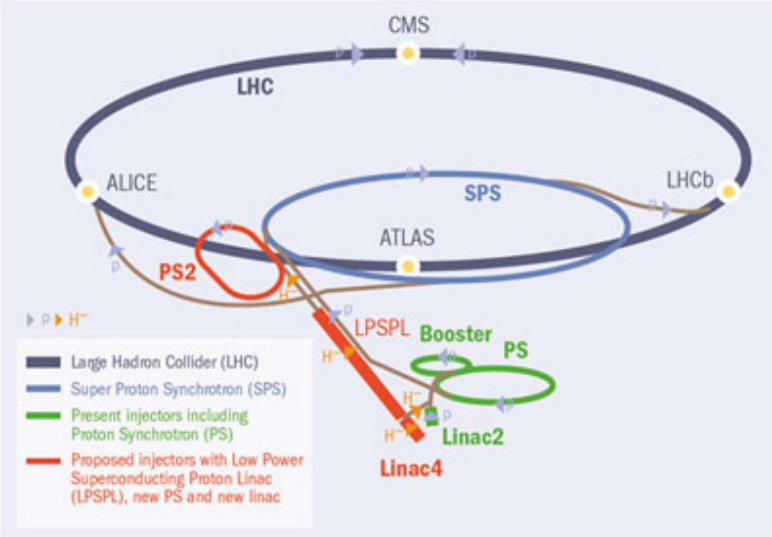


Figure 3.1: LHC injector complex [12].

applied electric field. The positively charged protons are lightly accelerated through a 90 kV potential and then fed into the LINAC2 with a speed of approximately .0014c, where c is the speed of light.

3.1.2 Injection Chain

LINAC2 is a linear accelerator using radiofrequency (RF) cavities to accelerate the protons to 50 MeV. From the LINAC2, the protons enter the Proton Synchrotron Booster. The PSB accelerates the protons from 50 MeV to 1.4 GeV using four synchrotron rings, at which point they are injected into the Proton Synchrotron. The Proton Synchrotron, which entered service at CERN in 1959 as its first synchrotron, takes the protons to 25 GeV and then feeds them further to the Super Proton Synchrotron, the last step before injection into the LHC itself [6].

The Super Proton Synchrotron, at 7 km in circumference, is the second-largest member of the accelerator complex. The protons are accelerated through room-temperature

electromagnets to an energy of 450 GeV.

The protons are finally fed into the main ring by the SPS where they are accelerated to their target energy by the LHC ring itself.

3.2 LHC Acceleration and Beam Storage

High beam energy and intensity are required to provide the luminosities needed to identify rare physics processes. Beam energy is dependent on the field strength of the storage ring - a powerful 8.3 T in the main dipoles.

Superconducting magnets and radiofrequency cavities are employed to accelerate and focus the beams, and to prevent heat loss in the beam itself. Presently 1232 dipole magnets provide the bending field, while 858 quadrupole magnets focus the beam [5]. The dipole magnets act as synchrotron accelerators, and alternating quadrupole fields focus the beam in either the horizontal or vertical direction. A system of three quadrupole magnets called an "inner triplet" is used to squeeze the beam from 0.2 mm to 16 μm at each interaction point [5].

Radiofrequency (RF) cavities use an oscillating electric field both to accelerate the protons in the beam and to coordinate their timing. Tightly packed proton bunches increase the likelihood of a successful collision, and the 16 RF cavities - 8 along each beam line - oscillate a 2 MV potential at 400 MHz. A perfectly timed or "synchronous" proton will not feel any acceleration at all and will remain in its bunch, while protons with incorrect timing will be accelerated.

3.3 Experiments

Several detectors are employed at the interaction points along the collider's circumference. There are four major experiments currently in use on the LHC. The full luminosity of the collider can be delivered to its two major general-purpose experiments, CMS (Compact Muon Solenoid) and ATLAS (A Large Toroidal LHC ApparatuS). LHCB, a smaller experiment focusing on b-quark physics, receives a luminosity of $L = 10^{32}\text{cm}^{-2}\text{s}^{-1}$. The LHC also collides lead ions, with Pb-Pb collisions delivering a luminosity of $L = 10^{27}$ $\text{cm}^{-2}\text{s}^{-1}$ to the ALICE (A Large Ion Collider Experiment) detector.

Chapter 4

CMS

CMS - the Compact Muon Solenoid - is an experiment located at one of the four major interaction points inside the LHC. It is an all-purpose detector designed to capture as much physical information as possible about collsions of interest that take place within it, in order to refine our current knowledge of the Standard Model as well as expand our understanding for processes beyond the Standard Model.

To meet its physics goals, the CMS detector was designed and built with the following in mind:

- Good muon identification and momentum resolution; good dimuon mass resolution; the ability to unambiguously determine the charge of muons with $p < 1$ TeV/c.
- Good charged particle momentum resolution and reconstruction efficiency in the inner tracker. Efficient triggering and offline tagging of τs and b-jets, requiring pixel detectors close to the interaction region.
- Good electromagnetic energy resolution, good diphoton and dielectron mass resolu-

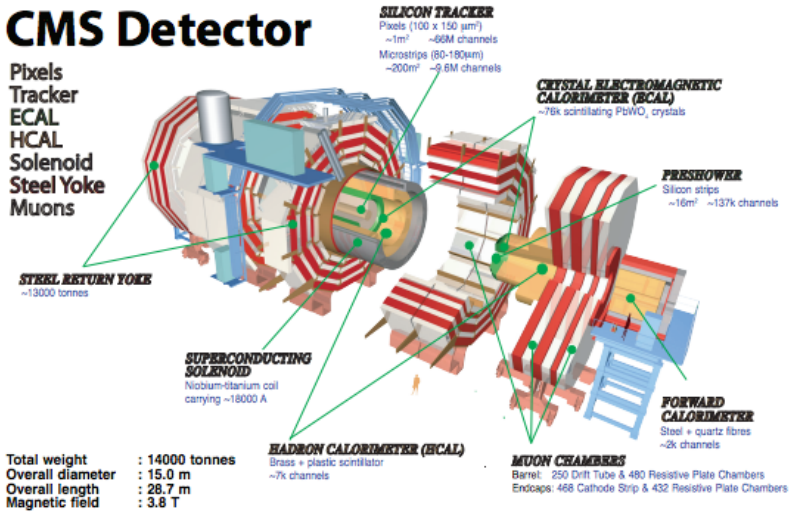


Figure 4.1: The CMS detector, exploded view. All figures in this chapter are from [32] unless otherwise noted.

tion, wide geometric coverage, measurement of the direction of photons and/or correct localization of the primary interaction vertex, π^0 rejection and efficient photon and lepton isolation at high luminosities.

- Good E_T^{miss} and dijet mass resolution, require hadron calorimeters with large hermetic geometric coverage and fine lateral segmentation ($\Delta\eta \times \Delta\phi < 0.1 \times 0.1$) [32].

As shown in Fig. 4.1, CMS contains 4 main sub-detector systems: a silicon tracker, lead tungstate electromagnetic calorimeter, brass and plastic hadronic calorimeter, and finally copper and plastic muon chambers. These sub-detectors are all cylindrically symmetric and are divided by a large solenoidal magnet as well as an iron return yoke. The 3.8 T solenoidal magnetic field bends charged particles, giving precise information on their momentum.

The coordinate system conventions used by CMS have an origin located at the

nominal collision point in the very center of the detector. The z-axis points along the beam direction, while the y-axis points vertically up and the x-axis points inward toward the center of the LHC ring. Generally, event coordinates are given in terms of ϕ, the azimuthal angle in the x-y plane measured from the x-axis, and η, the pseudorapidity. Pseudorapidity is a function of the polar angle θ, measured from the z-axis, as follows:

$$\eta = -\ln[\tan\left(\frac{\theta}{2}\right)] \tag{4.1}$$

Pseudorapidity is preferred to polar angle because differences in rapidity are Lorentz invariant under longitudinal boosts. Thus, particles moving under differing levels of boost in the longitudinal direction will be able to share a coordinate system.

The detector is divided into barrel and endcap sections, with the endcaps sealing the ends of the cylindrical volume while providing additional detector coverage.

4.1 Magnet and Return Yoke

The superconducting solenoid provides the large bending field used to obtain high momentum resolution for charged particles, particularly muons, in the detector. By passing 20 kA of current through 2168 turns of a superconducting NB_3Sn cable, a 3.8 T field is produced which provides $\Delta p/p \approx 10\%$ at $p = 1$ TeV. The steel return yoke is divided into five wheel-shaped sections in the barrel, and three disks in the endcap. The yoke conducts magnetic flux back through the detector, and contains more iron at 12500 tons than the Eiffel Tower.

4.2 Muon Chambers

Muon chambers enclose the detector, since the relatively long lifetime and high mass of the muon allow it to travel through much of the inner detector volume without decaying.

Muon chambers around the barrel are composed of 250 drift tubes organized in four layers, with one or two resistive plate chambers for each tube. Drift tubes are a gas-wire system, where electrons resulting from the ionization of the gas caused by the muon's passing will drift to a nearby wire. Resistive plate chambers are also a gas system composed of a pair of oppositely charged, high-resistivity plastic plates. The extremely fast response time of the RPC system is employed in the online triggering system.

The endcaps are covered with a combination of cathode strip chambers and RPCs, with 468 CSCs . The CSCs are very large ($\sim$ 1 m) chambers where ionized electrons encounter anode wires, which in turn send positive ions to copper cathode strips. This allows for 2-dimensional position resolution of $\sim$ 200 μm for muons passing through the chamber. RPCs supplement the CSCs on the endcap.

4.3 Hadronic Calorimeter

The hadronic calorimeter is designed to absorb almost all the hadronic energy of an interaction (excluding that carried away by muons and by non-interacting particles), meaning that interactions with the calorimeter material must be maximized before the outgoing particles reach the magnet-yoke system [32]. Brass is chosen for its short interaction length, while the active medium (in the case of the HCAL a plastic scintillator, with embedded

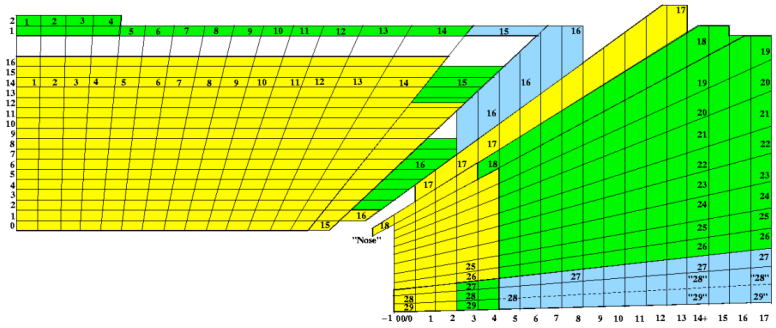


Figure 4.2: The CMS hadronic calorimeter.

wavelength-shifting optical fibers) must take up as little space as possible. The geometry of the HCAL is designed with this in mind, with 15 5 cm brass plates layered with 3.7 mm scintillator tiles. The barrel coverage extends from $-1.4 < \eta < 1.4$, with a segmentation of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$, with up to 11 interaction lengths of thickness.

Endcap coverage is divided into the HE (endcap) and HF (forward) sections. HE covers the region $1.4 < |\eta| < 3.0$, and is composed of brass and plastic layers with a segmentation in ϕ of 5°. In the HF, covering the $3.0 < |\eta| < 5.0$ region, different materials are used. A steel absorber and quartz fiber structure is segmented by 10° in ϕ.

The HCAL is shown in detail in Figure 4.2.

4.4 Electromagnetic Calorimeter

The high granularity and energy resolution of the ECAL, with its 61,200 scintillating crystals, provides for accurate reconstruction of electron and photon showers in an event. The ECAL covers a pseudorapidity range of $|\eta| < 3.0$, with an array of lead tungstate ($PbWO_4$) crystals with attached silicon avalanche photodiodes as light collectors and amplifiers. Lead tungstate was chosen for its short radiation length and fast showering-

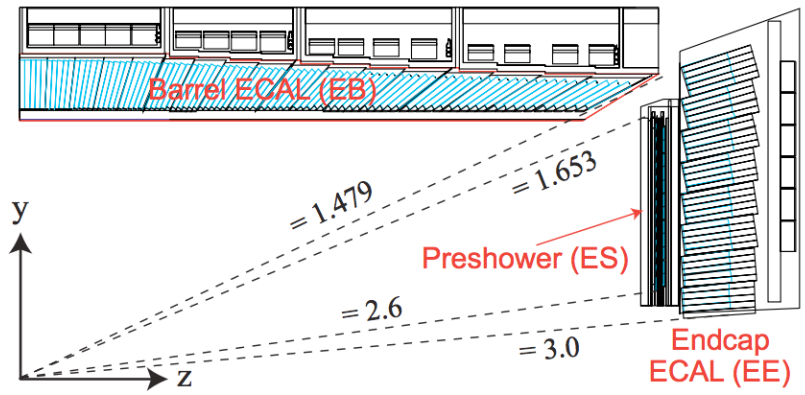


Figure 4.3: The CMS electromagnetic calorimeter. Diagonal dashed lines represent η values.

80% of the light contained in a shower will be emitted within 25 ns [32]. The full coverage of the ECAL encompasses some 25 X_0 (radiation lengths).

In the barrel ($|\eta| < 1.4442$), the ECAL has an inner radius of 129 cm, with 36 "supermodules" composed of 1700 crystals each. This gives a granularity of $\Delta\eta \times \Delta\phi = 0.0174 \times 0.0174$.

The region from $1.508 < |\eta| < 3.0$ is covered by the ECAL preshower and the endcap crystals. The preshower system consists of a lead plate to trigger immediate showering, followed by an array of silicon strips that detect the converted e^+e^- pairs. This system allows the entire EM energy of the event to be recorded, and allows for higher position resolution in the forward region.

The endcap consists of a series of 5×5 crystal structures about 314 cm from the interaction point. Figure 4.3 shows the ECAL in detail.

4.5 Silicon Tracker

High-precision tracking of charged particles is necessary for accurate event reconstruction, especially for the identification of primary and secondary interaction vertices in the event. To this end, a dual silicon tracker system is placed at the center of the detector, as close as possible to the interaction region. The outer layer of the tracker is composed of silicon microstrips, with a multi-layer architecture enabling the detection of multiple tracks in an event with high precision. The inner layer consists of 3 layers of pixel detectors, designed to accurately measure the impact parameter of charged particles as well as to locate secondary or displaced vertices in an event. The 65 million separate channels in the pixel detector provide an outstanding degree of momentum and position resolution.

4.5.1 Strip Tracker

The silicon strip tracker consists of four major sections: The tracker inner barrel (TIB), tracker outer barrel (TOB), tracker inner disk (TID) and tracker endcap (TEC). Each section is composed of the same basic module architecture, with each module consisting of a frame, Kapton tape delivering bias voltage to the sensors, a front-end electronics system, and one or two silicon sensors. The sensor geometry - primarily the pitch of the strip - is dependent on its placement in the detector. A sensor itself is a semiconducting silicon microstrip, with a p-type strip embedded in an n-type bulk.

In the TIB, modules are arranged in four cylindrical layers, while the TOB system has six layers mounted on a carbon fiber support frame, with three modules (either double or single-sided) on either side of the frame. The TID has three disks on each end of the

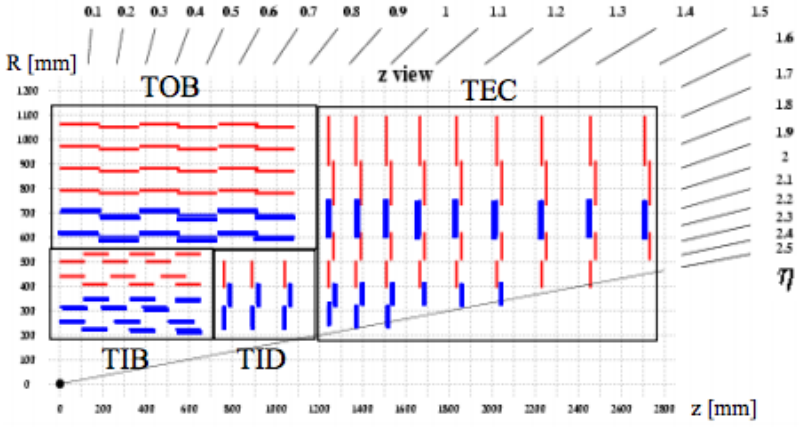


Figure 4.4: The CMS silicon strip tracker, one-quarter cross section. Blue lines represent double-sided modules and red lines represent single-sided modules.

cylinder, with modules sorted into three rings per disk. Finally, the TEC system is composed of nine disks mounted on carbon fiber support in a unique "petal" geometry, with each disk carrying eight front and eight back petals and each petal (a carbon fiber wedge) carrying up to 28 modules. A cross-section of the strip tracker geometry is shown in Figure 4.4.

The tracker system requires a reliable, fast, radiation-hard readout system[17]. The strip tracker utilizes a complex of either four or six readout chips mounted on the front-end hybrid system of each module. Each chip has 128 channels with a sampling frequency of 40 MHz (in line with the optimal 25 ns proton bunch crossing window).

4.5.2 Pixel Tracker

The pixel tracker, closest to the collision point, is divided into barrel and endcap sections much like the strip tracker. The barrel section (BPIX) consists of three layers 53.3 cm in length, with inner radii of 4.4 cm, 7.3 cm, and 10.2 cm respectively. Each layer is organized into a ladder (carbon fiber support system) with eight sensor modules per ladder.

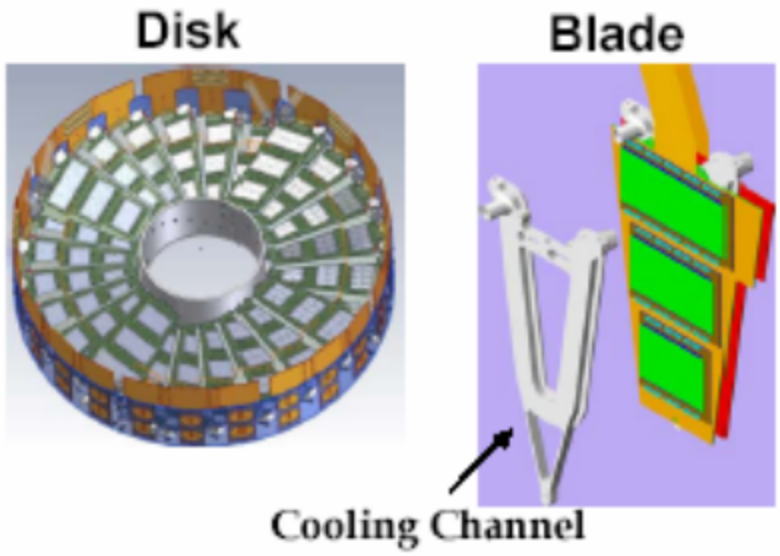


Figure 4.5: Layout of the CMS Forward Pixel Detector disk system as well as a single blade with plaquettes [20].

An individual module contains 16 readout chips (ROCs), with each ROC containing 4160 pixels arranged into 52 columns and 80 rows. Each full module contains 66,560 individual n-on-n silicon pixels. The full barrel region contains 672 full modules and 96 half modules, for a total pixel count of 47,923,200.

The pixel endcap section (FPIX) is composed of a pair of disks located at the ends of the BPIX, at $z = \pm 35.5$ cm and $z = \pm 48.5$ cm. Each half-disk contains 12 blades in a tilted geometry similar to a turbine in order to ensure maximum hit coverage for passing charged particles. A blade is composed of a pair of panels set back-to-back, with each panel covered in an arrangement of ROCs called "plaquettes". Plaquettes may contain 2-10 ROCs, with 672 plaquettes and a total of 17,971,200 pixels in the FPIX. Plaquette layout differs on each side of the blade in order to minimize gaps. Each panel also carries a Token Bit Manager (TBM) that coordinates the readout signals from the ROCs while delivering

trigger decisions and clock information. The disk and blade layout are shown in Figure 4.5.

When hit by a charged particle, a pixel will record its address in the detector, the time of the hit and the analog signal and then store this information in the ROC. The TBM, on reception of a trigger, passes a token ordering the ROCs to read out their signal. These electrical signals are converted to optical pulses in the Analog Opto-Hybrid (AOH) and sent to the Front-End Driver (FED) for cleanup and digitization.

Electronics readout must be fast, accurate, radiation-hard, and free of interference. High luminosity and therefore high radiation exposure closest to the collision point means that the pixel detector in particular must be able to adapt its response to the effects of irradiation. This is done by periodically recalibrating the sensor bias voltages to compensate for charge collection losses due to irradiation. The pixel online software is also designed to deal with occasional bit flips or single event upsets (SEUs) due to high-energy particles by quickly resetting the ROC [8].

4.6 Trigger

At peak luminosity, the LHC is capable of delivering over 10^6 collisions per second. Detector readout and archiving systems are capable of recording approximately 10^2 collisions per second of data. Therefore a very fast, robust method for identifying events of physics interest while discarding the rest has been developed in the form of the Level 1 Trigger (L1) system. Further filtering is handled by the High Level Trigger (HLT).

4.6.1 L1 Trigger

The L1 trigger accepts inputs known as "trigger primitives" from subsystems with very fast readout and dedicated trigger electronics, primarily the ECAL and muon system. A decision based on these inputs is computed in the L1 FPGAs and returned to every detector subsystem. The tracker in particular is reliant on the trigger in order to empty its data buffers quickly and reliably, due to its high multiplicity of channels and large data rate.

4.6.2 High Level Trigger

The HLT represents the second step in the event filtering process. Here, events passing L1 criteria are sent to the HLT server farm above the CMS experiment where they are sorted further depending on their particle content. Several HLT paths exist depending on the interest of analysts - as an example, the Higgs to diphoton path sorts events depending on ECAL energy deposits, shower shape, and invariant mass. Once events are filtered through the HLT system they are passed to full digitization and reconstruction as described in Chapter 6.

Chapter 5

Datasets

This analysis makes use of the full Run 2 dataset as delivered by the LHC, recorded by the CMS detector systems, and verified by the CMS Luminosity group. Verified events are listed in the Golden JSON file provided by the CMS DQM (Data Quality Monitoring) group. A list of datasets and their corresponding integrated luminosity is shown in Table 5.1.

Standard Model H $\rightarrow \gamma\gamma$ samples were used to verify background contributions by

Primary Dataset	**Dataset Name**	**Luminosity** ($\mathbf{pb^{-1}}$)
DoubleEG	/DoubleEG/Run2016B-03Feb2017_ver2-v2/MINIAOD	5930
	/DoubleEG/Run2016C-03Feb2017-v1/MINIAOD	2573
	/DoubleEG/Run2016D-03Feb2017-v1/MINIAOD	4248
	/DoubleEG/Run2016E-03Feb2017-v1/MINIAOD	4005
	/DoubleEG/Run2016F-03Feb2017-v1/MINIAOD	3101
	/DoubleEG/Run2016G-03Feb2017-v1/MINIAOD	7721
	/DoubleEG/Run2016H-03Feb2017_ver2-v1/MINIAOD	8391
	/DoubleEG/Run2016H-03Feb2017_ver3-v1/MINIAOD	215
Total		36184

Table 5.1: Datasets used in the analysis. Luminosity values are accurate as of the February 2017 Re-MiniAOD campaign.

Process	Dataset Name
GluGlu	/GluGluHToGG-M125_13TeV_amcatnloFXFX_pythia8/RunIISpring16MiniAODv2-PUSpring16RAWAODSIM-reHLT_80X_mcRun2_asymptotic_v14-v1
VBF	/VBFHToGG_M125_13TeV_amcatnlo_pythia8/RunIISpring16MiniAODv2-PUSpringRAWAODSIM-reHLT_80X_mcRun2_asymptotic_v14-v1
VH	/VHToGG_M125_13TeV_amcatnloFXFX_madspin_pythia8/RunIISpring16MiniAODv2-PUSpring16RAWAODSIM-reHLT_80X_mcRun2_asymptotic_v14-v1
ttH	/ttHJetToGG_M125_13TeV_amcatnloFXFX_madspin_pythia8/RunIISpring16MiniAODv2-PUSpring16RAWAODSIM-reHLT_80X_mcRun2_asymptotic_v14-v1

Table 5.2: SM Higgs datasets

the SM Higgs. Samples used appear in Table 5.2.

Chapter 6

Object Reconstruction

Proper identification of physics objects, including particles and energy deposits, is vital to any analysis effort. Reconstruction of these objects is managed via a multi-tiered offline analytical framework.

Data coming through the detector must first pass the L1 Trigger system, which discards the majority of events in order to speed processing and minimize storage requirements. Events passing the L1 Trigger are then passed through the High Level Trigger (HLT) filtering system, which flags data according to predetermined filters, allowing for later offline sorting of event types. Data is then digitized from the raw electronic sub-detector readouts, and reconstruction occurs post-digitization. The data tier system is shown in Figure 6.1. Physics objects are reconstructed using the CMS Particle Flow [33] algorithm and identified offline using recommendations from the EGamma and Jet/MET POGs. Objects are then sorted into collections via the MiniAOD data structuring system. We are concerned primarily with the reconstruction of photons, hadron jets, and missing transverse energy.

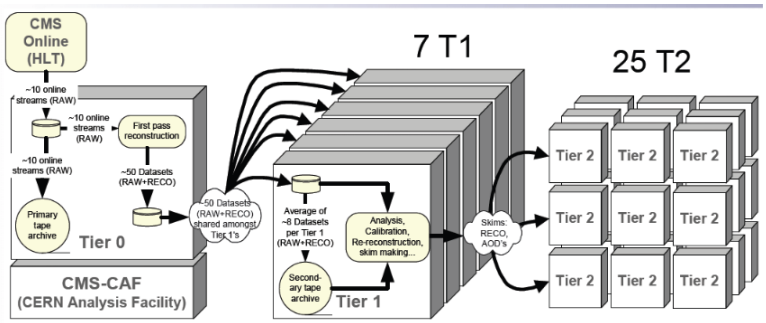


Figure 6.1: CMS detector data tiers[24]

6.1 Photons

Electrons and photons are reconstructed from superclusters in the ECAL. The electron and photon collections have substantial overlap - in fact as of 2015 all ECAL superclusters with electron seeds in the tracker are duplicated in the photon collection - due to the presence of non-prompt objects (decays not from the primary vertex) in the ECAL. In the case of non-prompt objects, photons are primarily from $\pi^0 \rightarrow \gamma + \gamma$ decays and electrons occur as a result of photon conversions ($\gamma \rightarrow e^+ + e^-$). A large fraction - up to 60% - of prompt photons convert in the tracker material [37].

6.1.1 ECAL Clustering

Two algorithms are currently employed to create ECAL superclusters. One, the Island algorithm, creates a cluster from a "seed crystal" - a crystal with energy deposit above a defined threshold - by scanning crystals adjacent to it first in ϕ and then in η. The Island algorithm is used in both the barrel and the endcap of the detector. The second, or Hybrid, algorithm, is used in the barrel exclusively. Hybrid functions by first creating superclusters of crystals over threshold energy and then decomposing them into basic clusters. It searches

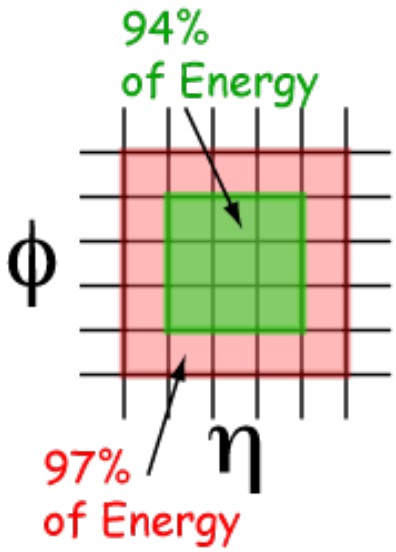


Figure 6.2: Energy cluster in ECAL crystals [27].

using a fixed η parameter of 3 or 5 crystals and then expanding in the ϕ direction. Almost all the energy of a photon or electron shower will be contained in a 3×3 crystal cluster, with 97% contained in a 5×5 cluster as shown in Fig. 6.2.

The measurement of energy deposits in the ECAL is not a perfect reproduction of the actual energy of the particle causing the showers. Several physical limitations of the ECAL must be taken into account when reconstructing the energy of a photon or electron. Electrons, for example, lose a fraction of their total energy in the tracker due to bremsstrahlung radiation. Other factors requiring energy correction include the effects of multiple interaction vertices (pileup) as well as loss of longitudinal shower depth due to intermodule voids. These parameters and others are used to train a regression, which corrects the raw crystal energy in the cluster to a value closer to the true energy of the particle. This method of correction is verified using the well-known $Z \rightarrow e^+ + e^-$ invariant mass and shows an excellent absolute energy resolution precision in the ECAL of approximately 0.4%

in the barrel and 0.8% in the endcap [34].

6.1.2 Photon Identification

Once ECAL cluster reconstruction and energy correction are completed, prompt photons must be distinguished from "fake" photons originating from neutral hadron decays or from conversions in the tracker material. Several properties of a cluster are employed in order to identify prompt photons, including but not limited to the energy ratio of a seed cluster to its supercluster, the shape of the shower, and its isolation from other objects in the ECAL.

R_9 measures the energy of a 3×3 cluster centered on a seed crystal vs the energy of the supercluster created around the seed [37]. This variable is useful in rejecting converted photons, since their energy is more widely spread across a supercluster, leading to a lower R_9 value. Additionally, R_9 may be used to reject showers from π^0 decays, since the outbound decay photons will be separated.

In order to separate prompt photons from neutral hadron showers, the fraction of energy deposited in the ECAL vs the HCAL is measured - a photon will deposit almost all its energy in the ECAL crystals, while a hadron will generally travel through the ECAL material before decaying in the HCAL.

The shape of a shower is also considered in rejecting fake photons. The extent of a supercluster's energy distribution in the η direction is generally more useful, since a wide lateral distribution in ϕ can be due to either a photon shower or the separation of a converted e^+e^- pair due to the magnetic field. This variable, called $\sigma_{i\eta i\eta}$, measures the energy across the 5×5 basic cluster centered around the seed crystal versus the energy of

the supercluster.

Perhaps most important to the differentiation of prompt from fake photons is the isolation of the shower. Isolation is measured in a cone of $\Delta R = 0.3$ ($\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$) around the photon's direction and is separated into charged hadron, neutral hadron, and photon components. Each component makes a distinct p_T contribution inside the cone, reconstructed from the particle-flow algorithm [33]. To prevent the addition of the photon's own energy, a slice of $\eta = 0.015$ is removed from the particle flow sum for photon isolation. For charged hadron isolation, a $\Delta R = 0.02$ region is excluded.

Isolation sums are affected by the contribution of particles from pileup vertices. Charged hadrons may be reliably associated with a primary vertex due to their presence in the tracker, but photon and neutral hadrons from particle flow must be corrected for pileup contributions. This is accomplished by estimating the E_T density per unit area in the detector, which is an event-wide value, and multiplying by the effective area of the measured isolation. This contribution is subtracted from the total isolation sum, leaving only the contributions from the primary interaction of the event.

A powerful method for rejecting electrons is to associate any reconstructed ECAL cluster with a "pixel seed" - a pattern of at least 2 hits in the pixel tracker. Any possible photon object with a pixel seed is rejected. This approach ensures a high purity of the photon sample, particularly in the barrel region. Another method, the conversion-safe electron veto, allows for a higher signal efficiency for photons but allows slightly more electrons into the sample. The two methods are employed dependent on the sensitivity of a selection to electron background and is at the discretion of the analyst. Their performance

	Barrel		Endcap	
	γ	e	γ	e
Conversion-safe electron veto	$99.1 \pm 0.1\%$	5.3%	$97.8 \pm 0.2\%$	19.6%
Pixel track seed veto	$94.4 \pm 0.2\%$	1.4%	$81.0 \pm 0.6\%$	4.3%

Table 6.1: Fractions of photons and electrons, in the ECAL barrel and endcap, passing the two different electron vetos. Statistical uncertainties in the values given for electrons are negligible [37].

is detailed in Table 6.1.

6.2 Jets

Hadron jets are reconstructed using a combination of track information and energy deposits in the HCAL. At high collision energies, the number of interaction vertices in an event can be very high and thus the clustering algorithms employed in the HCAL, as well as the high momentum resolution of the tracker which associates charged particles with a vertex, are used to identify jets as originating from the primary collision vertex or from a pileup vertex.

The approximately 10% energy resolution for hadrons in the combined ECAL-HCAL system allows for the identification of neutral hadrons, which appear as an energy excess in the HCAL once charged hadrons associated with tracks have been reconstructed. Jet E_T resolution is shown as a function of E_T, for barrel jets, in Figure 6.3.

6.2.1 Jet Reconstruction

The outstanding momentum resolution of the silicon tracker system allows for excellent reconstruction of charged particles in the PF algorithm. It also allows for the precise determination of a particle's direction as it exits the primary interaction vertex,

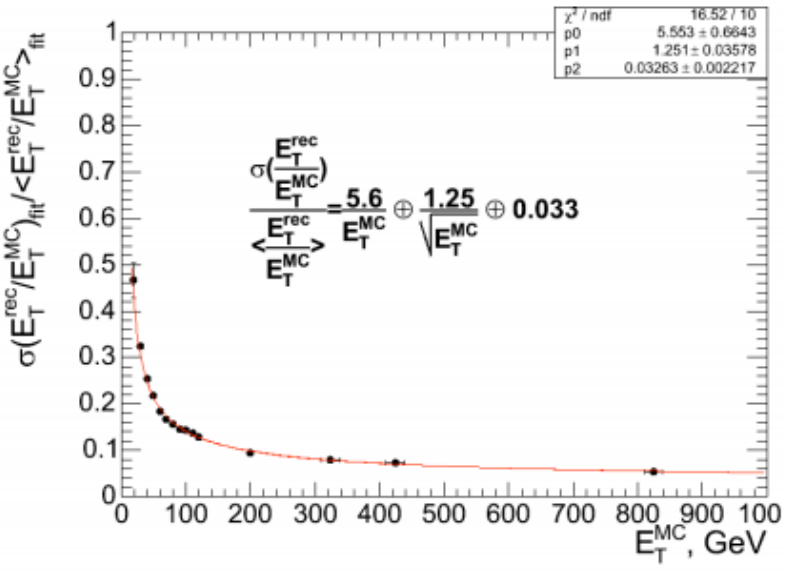


Figure 6.3: Jet transverse energy resolution as a function of generated jet transverse energy for barrel jets $|\eta| < 1.4$ [32].

since it registers hits before the particles are substantially deviated by the 3.8 T magnetic field. An iterative track seeding algorithm allows for a greater than 90% efficiency in charged hadron detection, with p_T as small as 150 MeV/c and a very small fake rate ($\sim 1\%$) [33].

These tracks are subsequently associated with energy deposits in the HCAL. Since hadron jets are composed of several individual stable or semi-stable particles, their reconstruction utilizes a sophisticated sequential clustering algorithm, in our case the AK4 (AK standing for anti-k_T and 4 referring to the $\Delta R = 0.4$ cone used to extract PF particles for further clustering). Anti-k_T refers to a value assigned to a parameter (k_T) in the clustering algorithm. Anti-k_T clustering is preferential to high-p_T objects. It is widely used due to its high resolving power for single jets, but, since it performs poorly with lower-p_T particles, does not resolve jet substructure well.

The effects of pileup are mitigated via the pileup jet ID, a multivariate technique that uses the direction of charged hadrons (if present in the events) as well as jet shape

variables (since pileup jets tend to be lower-p_T and more diffuse). The multivariate analysis gives a final discriminator value which separates jets according to their likelihood of origination in the primary interaction vertex.

Jet energy resolution performs less well in data than it does in simulation, necessitating a correction in the form of jet energy smearing of 10% in simulated jets. Jet energy corrections are applied to data and to simulation to account for the energy deposited in the calorimeters due to pileup jets (L1FastJet) and to counter the effects of differing detector response in p_T and η (L2L3Residuals). Some response corrections are applied only to simulated jets (L2L3) and also correct for response in η and p_T.

6.2.2 *b*-Tagging

b-jets are identified using the Combined Secondary Vertex (CSV) algorithm [35]. Hadrons arising from *b*-quarks have a lifetime on the order of 10^{-12} seconds, allowing them to travel some measurable distance from the primary vertex before decaying. This displaced vertex can be reconstructed using the high momentum resolution of the tracker system by examining the shifted impact parameter of its tracks. A CSV jet is shown in Figure 6.4.

A secondary vertex must pass the following criteria before its parameters are given to the CSV MVA:

- secondary vertices must share less than 65% of their associated tracks with the primary vertex and the significance of the radial distance between the two vertices has to exceed 3σ;
- secondary vertex candidates with a radial distance of more than 2.5 cm with respect to

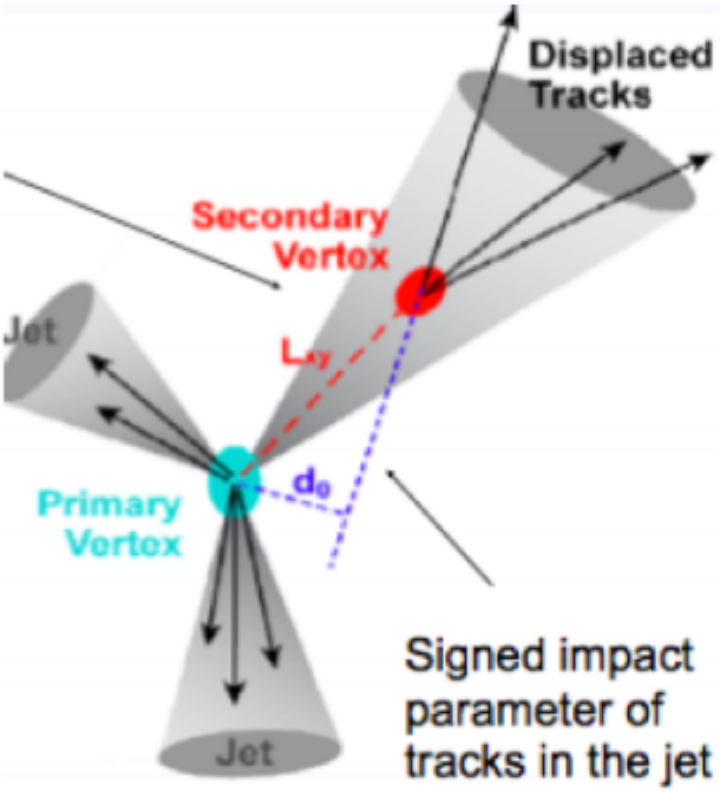


Figure 6.4: A b-jet with shifted impact parameter

the primary vertex, with masses compatible with the mass of K^0 or exceeding 6.5 GeV are rejected, reducing the ocntamination by vertices corresponding to the interactions of particles with the detector material and by decays of long-lived mesons;

- the flight direction of each candidate has to be within a cone of $\Delta R < 0.5$ around the jet direction. [35]

The secondary vertex algorithm uses several variables to train a multivariate analyzer and provide a final discriminant value, which is used to determine the likelihood that a given jet originates from a *b*-quark. Secondary vertices are constructed from particle flow jets passing basic quality cuts, and the vertex and track information (including time-of-flight, impact parameter, secondary vertex mass, and other relevant variables) are fed to the MVA for training. A final discriminant value is produced which gives the likelihood of a given jet originating from a *b*-quark.

The efficiency of the *b*-tag method varies between data and simulation, necessitating the application of a scale factor. The method of this application, and its effect on event weight, is detailed in Appendix A.

6.3 Missing Transverse Energy

Not all the energy in a given interaction will be recorded by the detector. Weakly interacting, chargeless and colorless particles (for example, the stable neutrino or the supersymmetric LSP) will carry energy out of the detector and will not register on any of the subsystems. The inbound protons will collide with a zero transverse momentum sum, since very nearly all their momentum is carried in the longitudinal direction due to their

near light-speed opposing velocities. This zero value should be conserved in the outgoing collision system, and if transverse momentum sums to a non-zero value we may infer that some of the momentum has been lost due to undetectable particles resultant in the collision.

This value, missing transverse energy or E_T^{miss}, is defined as the negative of the vectorial sum of the transverse momentum of an event:

$$\begin{aligned} \vec{E}_T^{miss} &= -\sum \vec{E}_T \\ E_T^{miss} &= \left| \vec{E}_T^{miss} \right| \end{aligned} \tag{6.1}$$

Full reconstruction of every particle and energy deposit in the event is required to make an accurate estimation of this value. The tracker is vital in the reconstruction of missing transverse energy, since any missed charged particle would likely still be detected in the calorimeters, causing problems in the final determination of direction of missing E_T as well as worsened energy resolution.

The E_T^{miss} response of the particle flow algorithm is within approximately $\pm 5\%$ of the true E_T^{miss} value above values of 20 GeV [36].

Chapter 7

Event Selection

As detailed in Section 2.3.1, the expected signal event will contain the following:

- Two photons with invariant mass in the Higgs boson resonance window
- Two or more b-jets, with b-jet multiplicity and invariant mass further dividing the signal categories

The proper selection of photon and jet candidates to best optimize the efficiency of the signal selection, while maximizing rejection of background, is outlined in detail below.

7.1 Photons

Photons, as discussed in Section 6.1.2, are reconstructed in the electromagnetic calorimeter. Careful selections are required in order to distinguish "prompt" - photons from the primary interaction - from "fake" - photons from neutral pion decay or electron conversion. To this end, various quality cuts are applied to candidate photons in the signal sample:

- Higgs to diphoton trigger pass for the event
- Photon p_T of (40, 25) GeV for the (leading, trailing) photon
- Both photons in the barrel of the ECAL
- Both photons have zero seeds in the pixel detector
- Both photons passing the EGamma POG cut-based photon ID at the loose working point
- Photon pair invariant mass in the (120, 131) GeV range

7.1.1 High-Level Trigger

The high-level trigger path used in this analysis is the general Higgs to diphoton trigger, HLT_Diphoton30_18_R9Id_OR_IsoCaloId_AND_HE_R9Id_Mass90, developed by the $H \to \gamma\gamma$ working group. Its selection depends on several variables, including photon p_T, hadronic energy fraction H/E, cluster isolation, and track isolation. A full description of trigger requirements is shown in Table 7.1. Note the similarity between variables used in trigger pre-selection and in the photon ID critera detailed in Section 7.1.5. Every HLT diphoton path is seeded by at least one hardware L1 electromagnetic candidate [14]

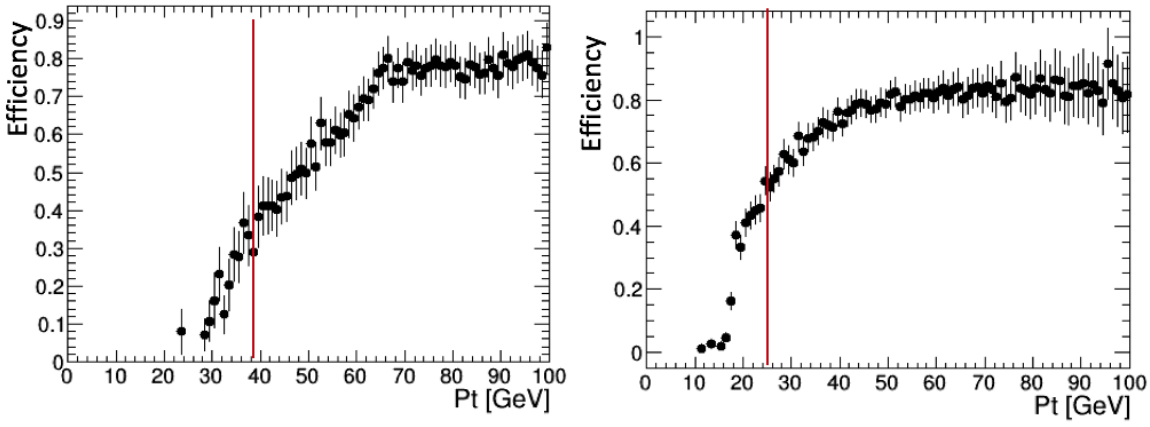
	H/E	$\sigma_{i\eta i\eta}(\mathbf{5\times 5})$	$\mathbf{R_9}(\mathbf{5\times 5})$	**ECAL PF Cluster Iso**	**Track Iso**
EB; $R_9 \geq 0.85$	< 0.12	–	< 0.5	–	–
EB; $R_9 \leq 0.85$	< 0.12	< 0.015	< 0.5	$< (6.0 + 0.012E_T)$	$< (6.0 + 0.002E_T)$
EE; $R_9 \geq 0.90$	< 0.1	–	< 0.8	–	–
EE; $R_9 \leq 0.90$	< 0.1	< 0.035	< 0.8	$< (6.0 + 0.012E_T)$	$< (6.0 + 0.002E_T)$
General Requirements					
$m_{\gamma\gamma} \geq 90$ GeV		HLT Seeded $E_T \geq 30$ GeV		HLT Unseeded$E_T \geq 18$ GeV	

Table 7.1: High-level trigger path selections, from [14]

Trigger requirements are applied in a very specific order. First, The L1-seeded leg of the trigger path is required to have $E_T > 30$ GeV, with the cluster located in the region $|\eta| < 2.5$ and $R_9 < 0.5$ in the ECAL barrel (EB), or $R_9 < 0.8$ if in the ECAL endcap. A hadronic fraction cut of H/E $< 0.12(0.1)$ in the EB(EE) region is applied as well. Once these selections are passed, each cluster must pass a further R_9 check at $> 0.85(0.9)$for EB(EE). If all these criteria are met, the trigger path proceeds to the unseeded leg. If the R_9 selection is not met, the $\sigma_{i\eta i\eta} < 0.015(0.035)$ and ECAL isolation $< (6.0 + 0.012(0.02)E_T)$ GeV are checked before moving to the unseeded leg. In addition, both legs of the trigger must pass the track isolation requirement $< (6.0 + .002E_T)$.

7.1.2 P_T

The p_T threshold is designed to minimize efficiency losses due to trigger turn-on effects. Efficiency plots are shown for leading and sub-leading photons in Figure 7.1.



(a) Trigger efficiency for leading photon

(b) Trigger efficiency for sub-leading photon

Figure 7.1: Trigger efficiency for leading and trailing photon in the event for signal point $m_{\tilde{b}} = 500$ GeV, $m_{\tilde{\chi}_2^0} = 200$ GeV.

7.1.3 η

Photons are required to be inside the ECAL barrel volume, $|\eta| < 1.4442$. Electron rejection is improved by this requirement, primarily due to the additional conversion electrons present in the endcap due to the preshower system. Both electron vetoes available (see Table 6.1) perform poorly in the endcap, with the conversion-safe veto providing a poor purity sample by allowing too many electrons through, and the pixel seed veto suffering a 13% efficiency loss while maintaining relatively high purity.

7.1.4 Pixel Seeds

Though this analysis is not a priori sensitive to electron background, the pixel seed cut ensures a higher purity signal sample at the expense of only a 4% efficiency loss in the barrel as seen in Table 6.1. A sensitivity comparison between the pixel seed and conversion-safe veto may be seen in Figure 7.2.

7.1.5 Cut-Based ID

The criteria for the EGamma POG cut-based (VID) ID for Run 2 is shown in Table 7.2. The working points of the ID (loose, medium, and tight) are predicated on signal efficiency (91%, 81%, and 71%) and misidentification rate (10%, 5%, and 1%). We choose the loose working point, as its high efficiency combined with the rejection power of the pixel seed veto provide an optimized selection [28].

Cut-Based ID WP	**Loose**	**Medium**	**Tight**
H/E	0.0597	0.0396	0.0269
$\sigma_{i\eta i\eta}$	0.01031	0.01022	0.00994
charged hadron iso	1.295	0.441	0.202
neutral hadron iso	$10.910 + 0.0148 * p_T + 0.000017 * p_T^2$	$2.725 + 0.0148 * p_T + 0.000017 * p_T^2$	$0.264 + 0.0148 * p_T + 0.000017 * p_T^2$
photon iso	$3.630 + 0.0047 * p_T$	$2.571 + 0.0047 * p_T$	$2.362 + 0.0047 * p_T$

Table 7.2: Values for the EGamma cut-based photon ID. Isolation values are ρ-corrected particle flow p_T sums in a cone of $\Delta R = 0.3$. [28]

7.1.6 Invariant Mass

The invariant mass "tag" window is based around the known H $\rightarrow \gamma\gamma$ resonance. This is a narrow, clean channel with good mass resolution. The window width (120, 131) GeV is chosen to enclose a majority of the resonance, as seen in Figure 7.3.

7.2 B-Jets

An event must have at least two jets that pass a set of quality cuts similar but not identical to photons, and are then assessed for their b-tag status. Each event must have two jets with the following:

- Both jets with $p_T > 30$ GeV
- Both jets with $|\eta| < 2.4$
- Both jets pass the PF loose ID [29]
- Both jets pass the pileup ID
- Both jets pass the CSV loose ID, with at least one passing the medium ID

Jet quality cuts are taken from the JetMET POG recommendations. Jets do not suffer from purity losses in the endcap as photons do, so acceptance extends to $|\eta| < 2.4$. The particle flow jet ID and pileup ID are both covered in detail in Chapter 6. The pileup ID discriminant is set to 0.8 - the loose working point - at the recommendation of the JetMET POG [30].

7.2.1 b-Tag ID

After jet candidates pass the previous set of criteria, their b-tag status is tested. Candidates must both pass the CSV Loose working point of 0.5426, with at least one of the jets passing the Medium working point at 0.8484 [26]. The choice of Medium/Loose is based on an evaluation of Q score, a measure of significance developed in [3]. Q scores for the signal point $m_{\tilde{b}} = 250$ GeV, $m_{\tilde{\chi}_2^0} = 50$ GeV are shown in Table 7.3. In this case there is no gain in significance associated with a tighter working point combination that would offset the loss of statistics. Q scores are tested separately in the distinct signal regions developed in the next section.

	$Q = (\sqrt{(S+B)} - \sqrt{B})$		
	2 b-jets on H mass	2 b-jets off H mass	3+ b-jets
CSV Medium, Loose	1.781	1.55461	4.80496
CSV Medium, Medium	1.69216	1.46833	4.83165
CSV Tight, Medium	1.73133	1.52369	4.74599
CSV Tight, Tight	1.04701	1.05779	2.40181

Table 7.3: Significance in Q score for various combinations of b-jet working points in the CSV algorithm.

Event weights are scaled according to the method developed in Appendix A. The CSV discriminant represents the likelihood that a given jet is a b-jet, and a scale factor is

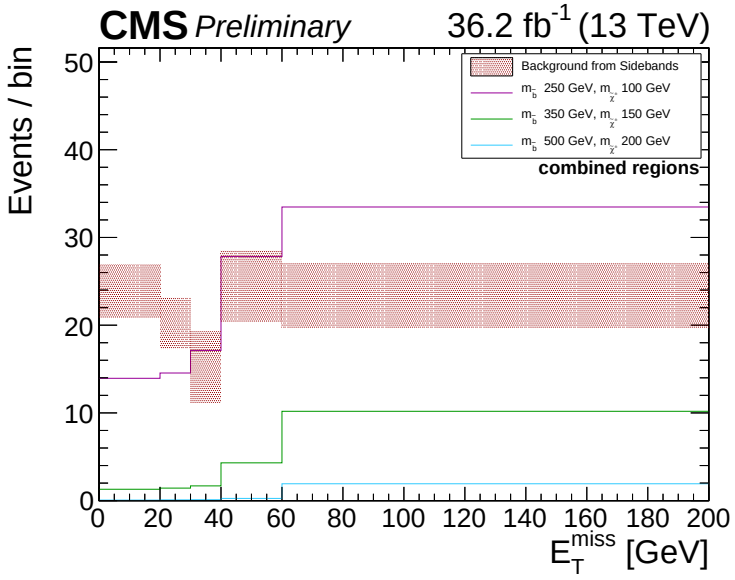
applied to account for possible mistagging as well as the data/simulation mismatch in the tag rate [25].

7.3 Search Categories

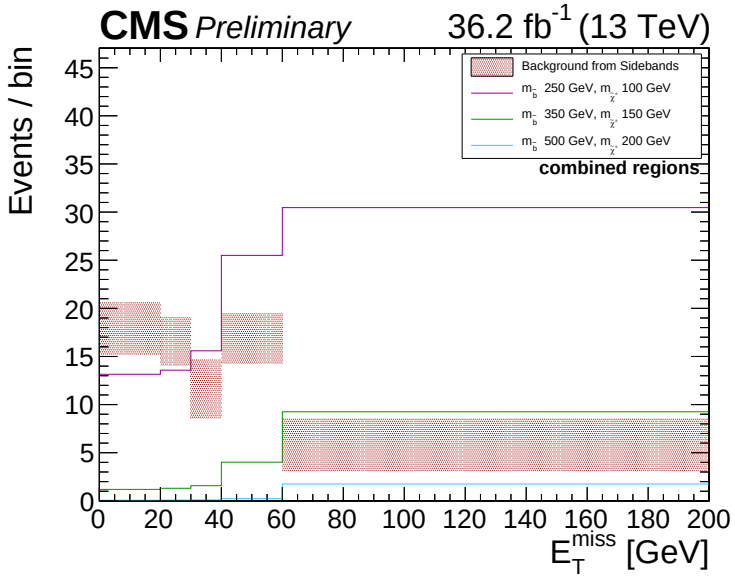
Once a candidate event has been determined to pass all signal criteria, it is sorted into one of three mutually exclusive search categories. These are:

- 2 b-jets on the H $\rightarrow b\bar{b}$ mass window
- 2 b-jets off the H $\rightarrow b\bar{b}$ mass window
- 3 or more b-jets

The first category takes advantage of a possible H $\rightarrow b\bar{b}$ resonance. The invariant mass window here extends from (95, 155) GeV given the broader width of this resonance compared to the H $\rightarrow \gamma\gamma$ decay. The second category accounts for possible detection of b-jets from the sbottom decays, and the 3+ b-jet category searches for events with a H $\rightarrow b\bar{b}$ decay as well as b-jets from sbottoms - possibly up to four per event. This categorization improves the sensitivity of the analysis, particularly in the 3+ b-jet category, as shown in Figure 7.4.



(a) Sensitivity with conversion-safe electron veto



(b) Sensitivity with pixel seed veto

Figure 7.2: Photon purity veto in combined signal regions. Note the substantial reduction in background (red shaded area).

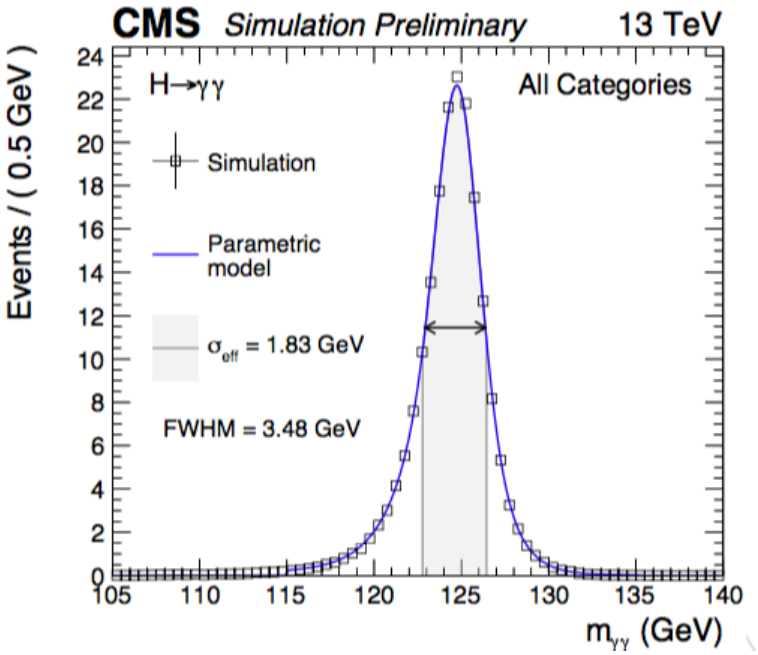
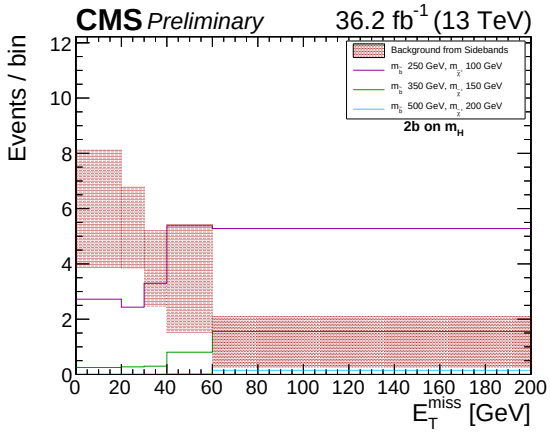
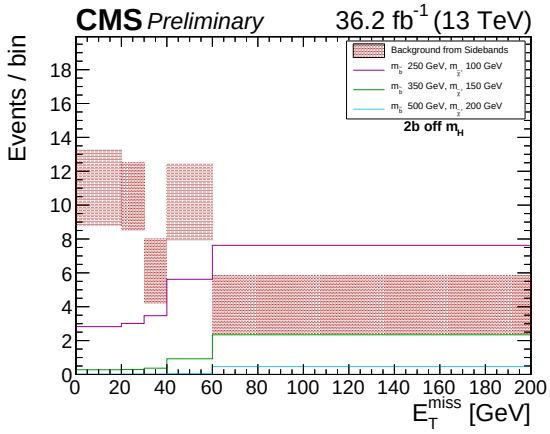


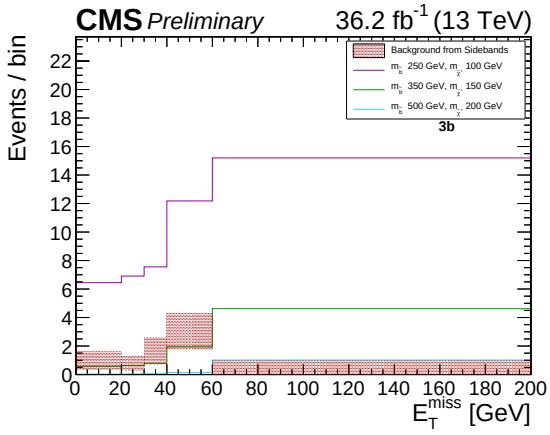
Figure 7.3: The Higgs $\rightarrow \gamma\gamma$ invariant mass. The high resolution allows for the creation of a Higgs "tag" based around the diphoton mass [14].



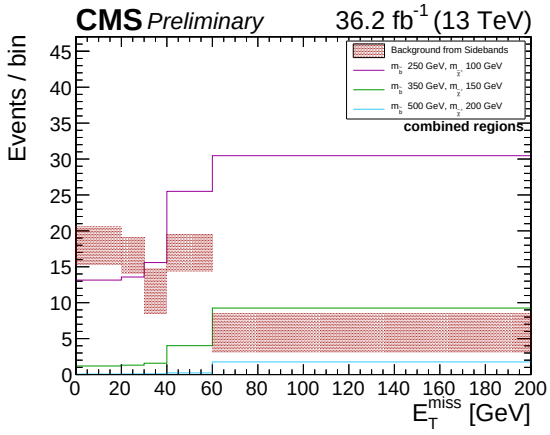
(a) 2 b-jets on Higgs mass



(b) 2 b-jets off Higgs mass



(c) 3+ b-jets



(d) Combined signal regions

Figure 7.4: Sensitivity by signal region. Solid colored lines represent one of three chosen signal points, while the red shaded area represents the background prediction as estimated in Chapter 8.

Chapter 8

Background Estimation

Determination of the background is entirely data-driven and is accomplished by a sideband method. Events are chosen via the criteria established in Chapter 7, but at the invariant mass check they are sorted into mass sidebands, one on either side of the signal or "tag" region of (120, 131) GeV. This method allows for a simple prediction of non-SM Higgs background, as the $H \to \gamma\gamma$ resonance sits as a narrow peak on top of the steadily decreasing SM diphoton invariant mass distribution (Figure 8.1). Event kinematics of background-type events inside this window are very similar to those outside it, allowing for an examination of sensitive variables whether or not they are correlated with the diphoton mass. The sideband method allows for a determination of Standard Model background that is completely agnostic to its composition. This method was originated in the Run 1 stop-higgsino analysis mentioned in Chapter 2.

A fit is performed separately in each sideband, with the results of the fit and the contents of the sidebands used to predict the number of events in the signal region.

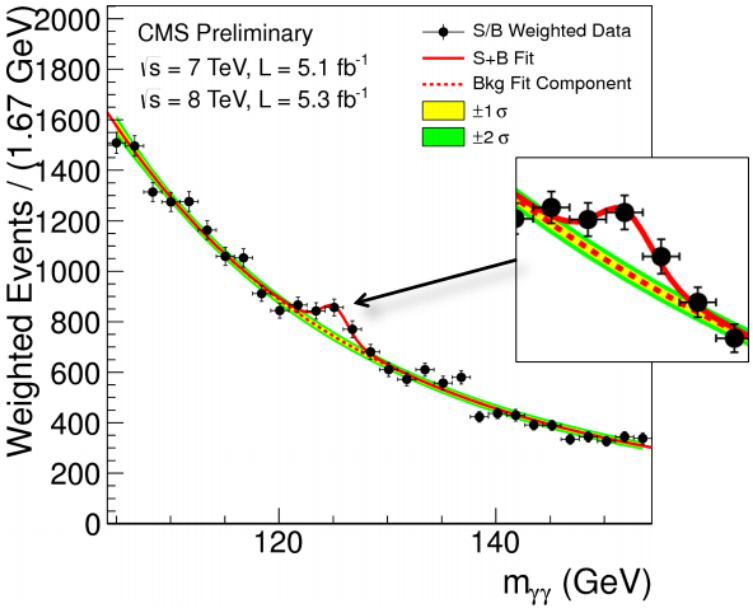


Figure 8.1: Higgs discovery 2012 - diphoton invariant mass [16].

Process	Cross Section (pb)
gluon-gluon fusion	44.01
vector boson fusion	3.782
VH	2.257
ttH	0.5071
bbH	0.4880

Table 8.1: Standard Model Higgs production cross sections at $m_H = 125$ GeV [19].

This method accounts for all possible sources of background except for the Standard Model Higgs boson. Production channels and cross sections for contributing SM Higgs processes are shown in Table 8.1. Since the SM Higgs is singly produced, this contribution is found to be negligible - about 4.7 events in the combined signal regions, mostly in the low E_T^{miss} region where the search is less sensitive. Figure 8.2 shows the number of events from all major SM Higgs processes in the combined signal regions.

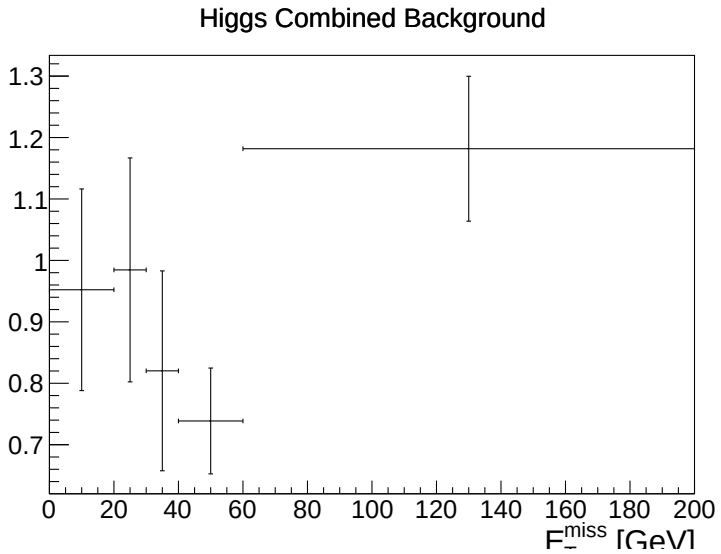


Figure 8.2: Standard Model Higgs background contributions, in combined signal regions.

8.1 Sidebands

The width of the signal region is chosen to contain the width of the $H \to \gamma\gamma$ (see Figure 8.1 as well as Figure 7.3) and to allow for some spread in the energy resolution of the photon measurement in the ECAL. Sideband widths are (103, 118) GeV for the lower sideband and (133, 163) GeV for the upper sideband. A 2 GeV buffer region lies between each sideband and the signal region to reduce contamination of the background measurement by any tails in the signal distribution (see Figure 7.3).

The lower sideband region is chosen such that the $Z \to e^+e^-$ resonance does not contribute to the photon background. The border on the upper sideband is less constrained and therefore the sideband itself is wider - the upper sideband is twice the width of the lower. Depending on the signal category chosen (2 b-jets on the H mass, 2 b-jets off the H mass, 3+ b-jets) the uncertainty of the background estimate is primarily driven by statistical concerns and is elaborated further in Section 8.4.

8.2 Fitting

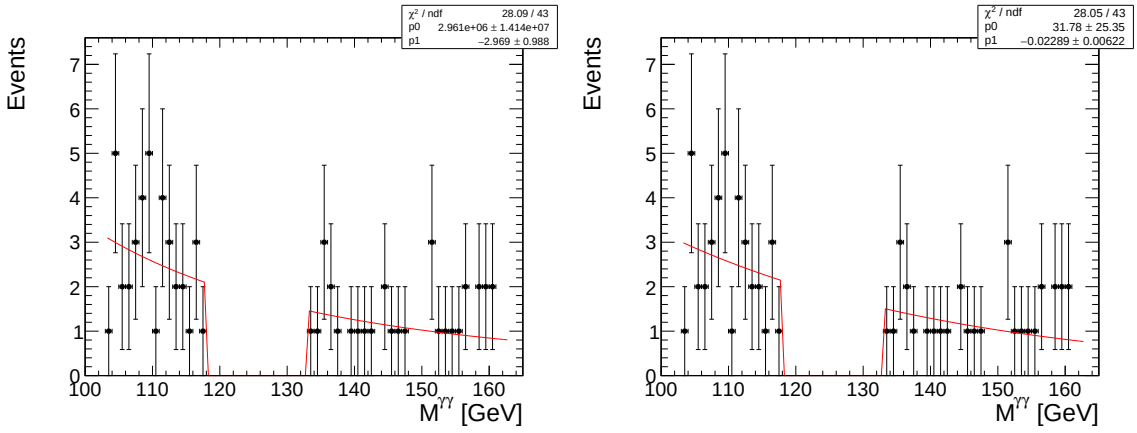
Once sidebands are filled, a fit is performed in each sideband and search category separately (for a total of 6 independent fits). Fits exclude the signal region and the buffers to avoid signal contamination. Four functions - an exponential, a power law, a second-degree polynomial and a third-degree polynomial - were tested to find the best modelling of the background distribution. Plots are shown in Figures 8.3, 8.4, and 8.5 for the 2 b-jet on Higgs mass, 2 b-jet off Higgs mass, and 3+ b-jet search categories respectively.

Both the power law and exponential functions match the distribution well - the low statistics cause unphysical shapes in the 2nd-degree polynomial and 3rd-degree polynomial fit. This is in line with physics expectations, given that QCD $\gamma\gamma + b\bar{b}$ and $b\bar{b} + \gamma + j$ (where the jet is misidentified as a photon) production dominates the background distribution. The power law and exponential functions show similar χ^2 values, and, as shown in Table 8.3 make virtually the same predictions for the signal region within error. The exponential fit shows slightly lower parameter error and so is chosen as the fit function. χ^2 values are given for each fit function in Table 8.2.

χ^2/NDF				
Fit Function	2 b-jets on h mass	2 b-jets off h mass	3+ b-jets	all categories
Exponential	0.652435	1.55266	0.36156	0.936729
Power Law	0.653274	1.5609	0.361221	0.930079

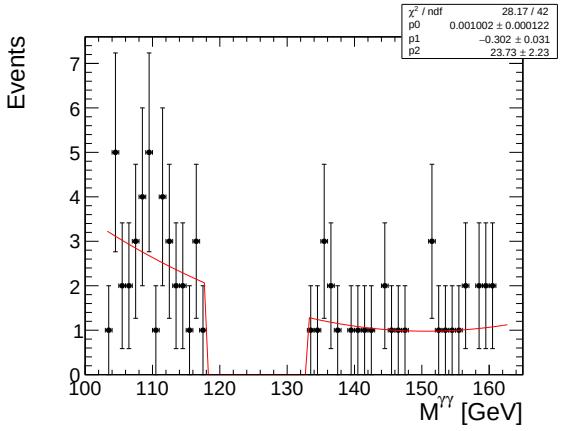
Table 8.2: χ^2 per degree of freedom, by fit and signal region.

The final form of the fit functions are given here for the 2 b-jet on Higgs mass (Eq. 8.1), 2 b-jet off Higgs mass (Eq. 8.2), and 3+ b-jet categories (Eq. 8.3).

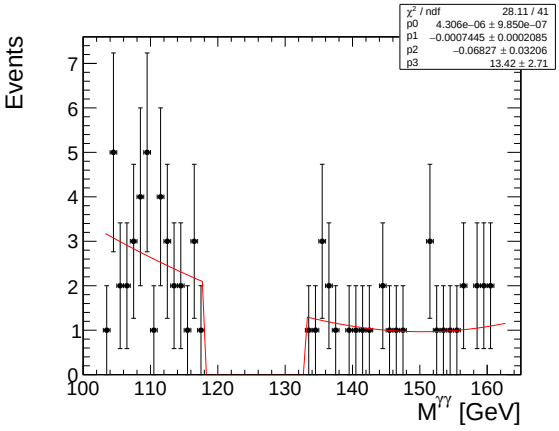


(a) Power Law

(b) Exponential

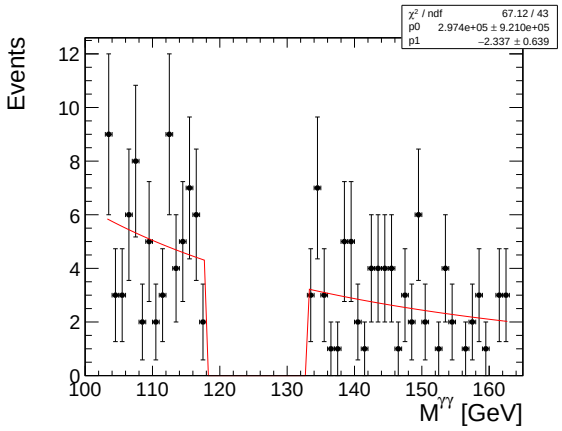


(c) 2nd-degree Polynomial

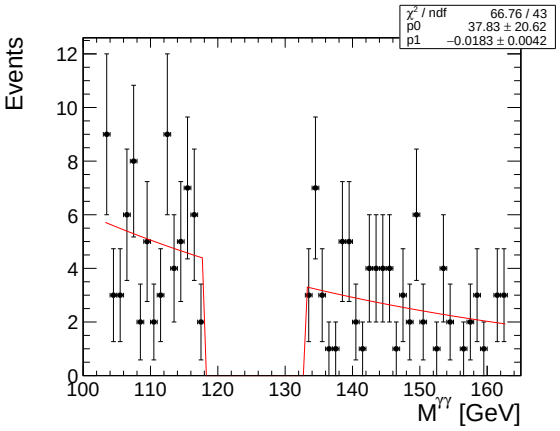


(d) 3rd-degree Polynomial

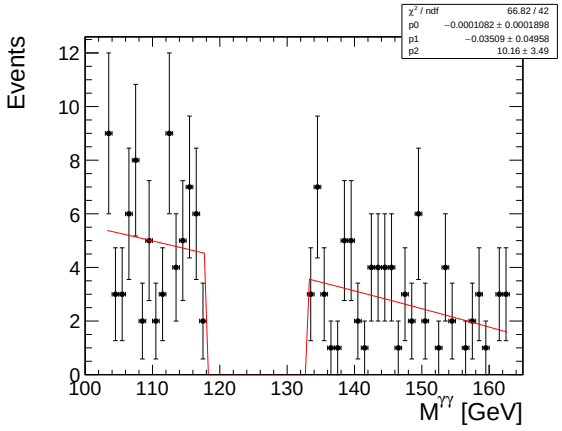
Figure 8.3: Fits to diphoton invariant mass - 2 b-jet on Higgs mass signal region.



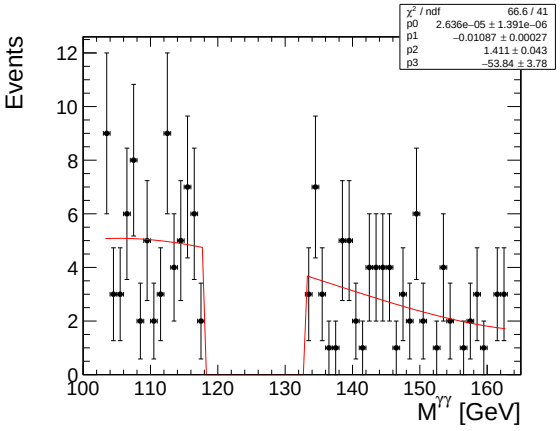
(a) Power Law



(b) Exponential

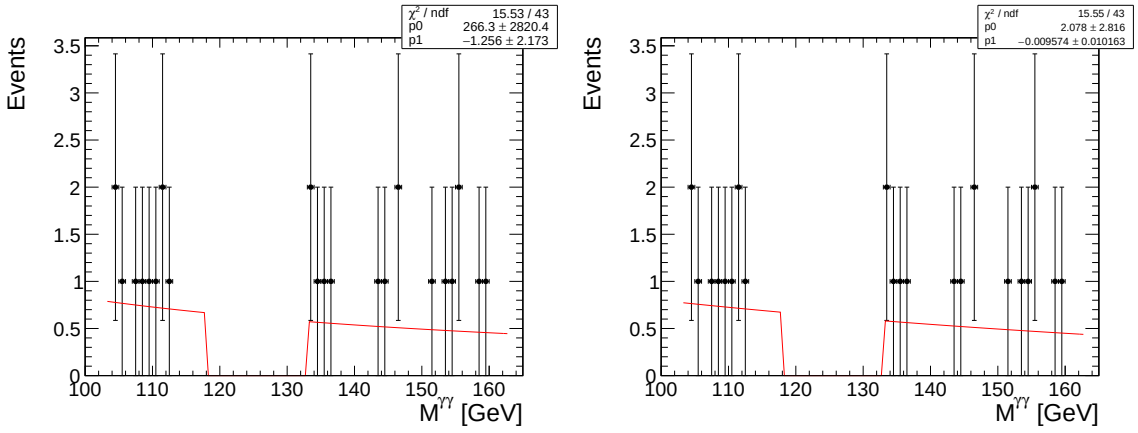


(c) 2nd-degree Polynomial



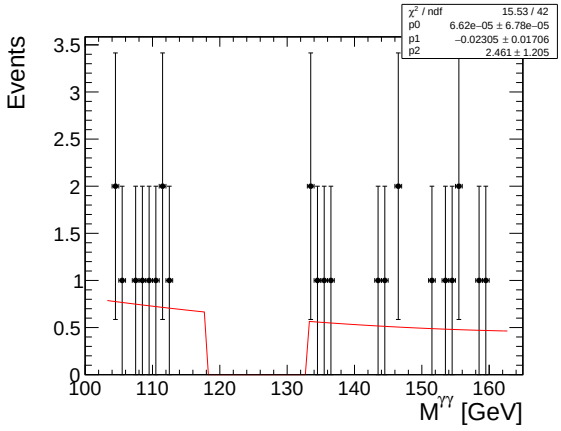
(d) 3rd-degree Polynomial

Figure 8.4: Fits to diphoton invariant mass - 2 b-jet off Higgs mass signal region.

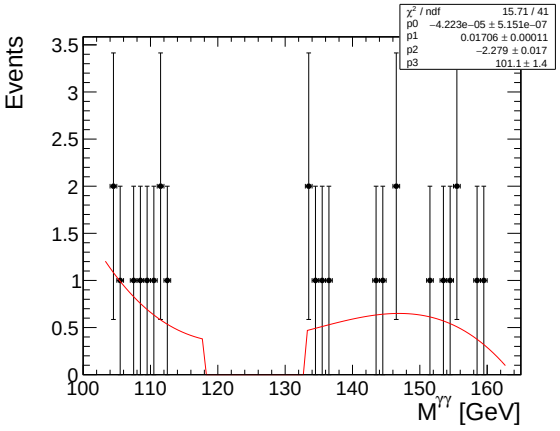


(a) Power Law

(b) Exponential



(c) 2nd-degree Polynomial



(d) 3rd-degree Polynomial

Figure 8.5: Fits to diphoton invariant mass - 3+ b-jet on Higgs mass signal region.

$$N_{events} = (31.78 \pm 25.35) * e^{(-.02289 \pm 0.00622) * M_{\gamma\gamma}} \tag{8.1}$$

$$N_{events} = (37.83 \pm 20.62) * e^{(-.0183 \pm .0042) * M_{\gamma\gamma}} \tag{8.2}$$

$$N_{events} = (2.078 \pm 2.816) * e^{(-.0096 \pm 0.0101) * M_{\gamma\gamma}} \tag{8.3}$$

8.3 Estimation in Tag Region

Once the exponential fit has been completed in each sideband for each search category (excluding the signal region), it may be used within the signal region. Without actually inspecting the event content of the tag region, the fit is integrated across it (again excluding the buffers). The integral of the fit inside the $M_{\gamma\gamma}$ window is employed, along with the contents of a sideband, to make an estimate from each sideband. These estimates are then averaged to give the final prediction. The sideband contribution estimation is computed bin-by-bin for a given kinematic variable as follows for the upper sideband (USB):

$$\mathrm{BG}_{USB} = \mathrm{bin}_{USB} * \frac{\int \mathrm{Fit}_{Tag}}{\sum \mathrm{Events}_{USB}} \tag{8.4}$$

Here, BG_{USB} represents the background estimate from the upper sideband, bin_{USB} is the contents of a single kinematic bin whose diphoton invariant mass lands in the upper sideband, $\int \mathrm{Fit}_{Tag}$ is the integral of the fit in the signal region, and $\sum \mathrm{Events}_{USB}$ is the sum of all events in the upper sideband.

The identical calculation is performed for the lower sideband in a bin-by-bin fashion. Finally, BG_{USB} and BG_{LSB} are averaged to give a final prediction of the number of events per bin in the signal region for the kinematic variable in question.

Summed predictions for each signal region are given in Table 8.3. The background prediction given by the power law fit is compared to the exponential method and found to be in agreement within errors.

Background Prediction (Events)				
Fit Function	2 b-jets on h mass	2 b-jets off h mass	3± b-jets	all categories
Exponential	19.8 ± 2.4	42.0 ± 3.5	6.9 ± 1.4	68.8 ± 4.4
Power Law	19.2 ± 2.3	40.9 ± 3.4	6.8 ± 1.5	67.0 ± 4.3

Table 8.3: Background prediction by fit and signal region.

8.4 Error

The uncertainty involved in the background estimation method is primarily statistical, but must also account for any possible correlation between the diphoton invariant mass and the sensitive variable in question. This correlation uncertainty is given by taking half the difference of the upper and lower sideband predictions, on a bin-by-bin basis, and adding it in quadrature to the statistical uncertainty. The development of statistical uncertainty is taken from [2] and given below.

$$\begin{aligned}\sigma_{\mathrm{BG}_{avg}} &= \mathrm{BG}_{avg}\sqrt{\frac{\sigma_U^2+\sigma_L^2}{(U+L)^2}+\left(\frac{\sigma_{int}}{\int \mathrm{Fit}_{Tag}}\right)^2} \\ &= \frac{1}{2}\sqrt{\left(\sigma_U^2+\sigma_L^2\right)^2\sigma_{int}+(U+L)^2\int \mathrm{Fit}_{Tag}}\end{aligned} \tag{8.5}$$

with

$$U = \frac{\text{bin}_{USB}}{\sum \text{Events}_{USB}}$$
$$L = \frac{\text{bin}_{LSB}}{\sum \text{Events}_{LSB}} \tag{8.6}$$

and

$$\sigma_U = U\sqrt{\frac{1}{\text{bin}_{USB}} - \frac{1}{\sum \text{Events}_{USB}}}$$
$$\sigma_L = L\sqrt{\frac{1}{\text{bin}_{LSB}} - \frac{1}{\sum \text{Events}_{LSB}}} \tag{8.7}$$

Finally, adding half the bin-by-bin difference between lower and upper sidebands, the total uncertainty on the background estimation is

$$\sigma_{\text{BG}_{total}} = \sqrt{\sigma^2_{\text{BG}_{avg}} + \left(\frac{\text{BG}_{USB} - \text{BG}_{LSB}}{2}\right)^2} \tag{8.8}$$

Chapter 9

Systematic Uncertainties

Measurements of both simulated events and data are subject to uncertainty. Statistical uncertainty in the background estimation is treated in Chapter 8. However, systematic uncertainties due to mismeasurement, resolution effects, or misidentification probability must be treated as well. Several of these affect our confidence in the raw signal acceptance of the simulated samples and are elaborated on in the following sections. All systematics listed here affect the observed and expected limits to be set forth in Chapter 10 and are included in limit-setting as a template fit to data [13]. The systematic uncertainties studied here are according to the recommendations of the CMS SUSY Physics Analysis Group.

9.1 Luminosity

A flat 2.5% uncertainty is applied to the luminosity per the recommendation of the CMS luminosity group [38].

9.2 HLT

The uncertainty in trigger efficiency has been measured by the Razor H $\rightarrow \gamma\gamma$ group and found to be 3% [10]. This uncertainty applies across all signal samples.

9.3 Cross Section

Cross section uncertainties are specific to the mass of the generated sbottom quark and fluctuate between 13-14% for the generated masses in the splitting. A higher mass will typically have a slightly lower uncertainty in its cross section [4].

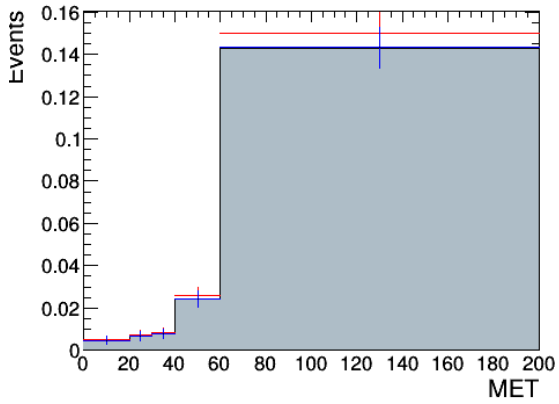
9.4 Photon Identification and Pixel Seed Scale Factor

The cut-based photon ID is found to be slightly less efficient in data than in simulation. A scale factor, dependent on p_T and η, is applied to each event to normalize simulation to data counts. For the Loose ID, scale factors are very close to unity. The uncertainty of the scale factor in the barrel, using the pixel seed veto and an R_9 value of > 0.94, is $\sim 1.7\%$. This uncertainty is calculated based on pileup corrections, choice of generator in simulated samples, and the combined uncertainty of the ID variables [18].

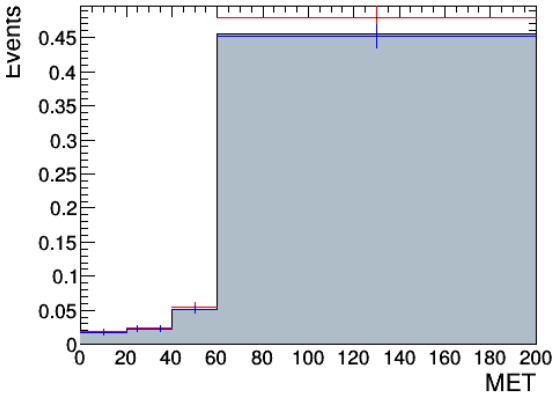
9.5 B-Tag Scale Factor

The b-tag event reweighting method is done on a jet-by-jet basis using efficiencies and mis-tag probabilities, and is explained in detail in Appendix A. The effect of the scale factor uncertainty on signal efficiency is measured by shifting the scale factor up and down

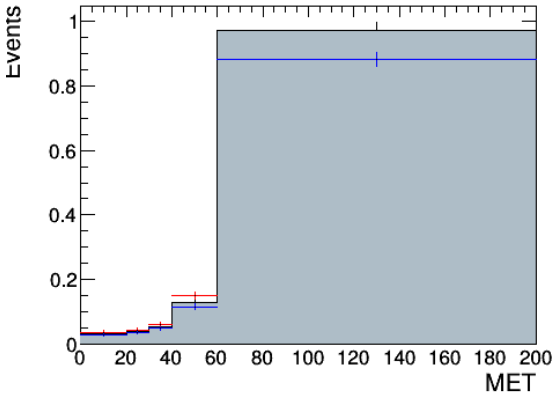
by 1σ and generating a new event weight for each shift. The up-shifted and down-shifted uncertainties are averaged to give a final systematic for the b-tag scale factor. The effect of this shift is shown in Figure 9.1.



(a) 2 b-jet on Higgs mass



(b) 2 b-jet off Higgs mass

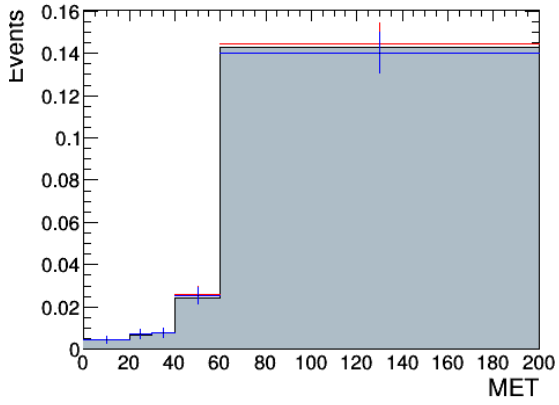


(c) 3+ b-jets

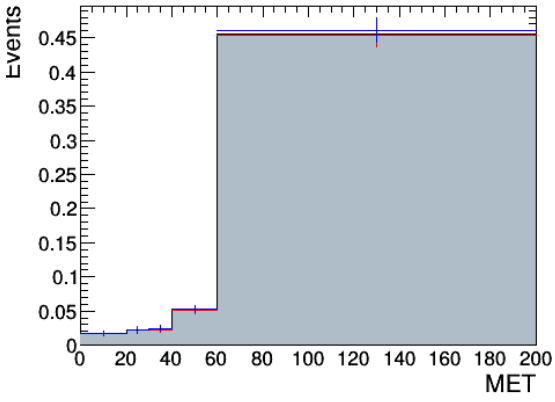
Figure 9.1: The effects of the b-tag scale factor uncertainty on signal acceptance for the signal point $m_{\tilde{b}} = 500\text{GeV}, m_{\tilde{\chi}_2^0} = 200$ GeV. Red lines indicate a 1σ shift up in scale factor, blue lines indicate a 1σ shift down in scale factor, and the grey filled area is the result of the selection based on the central b-tag scale factor. The y-axis is in arbitrary units.

9.6 Jet Energy Correction

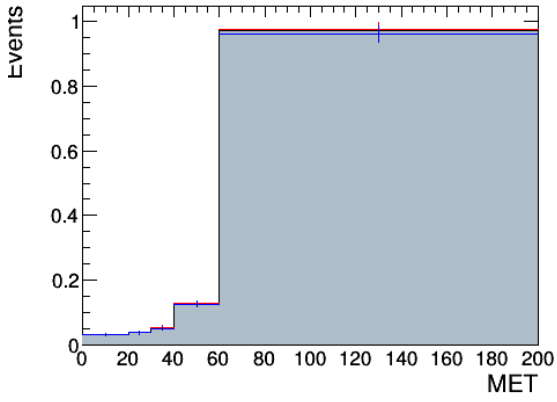
Jet energy correction uncertainty is found in much the same way as the b-tag systematic. Uncertainty shifts of $\pm 1\sigma$ are provided by the JetMET POG. The previously corrected jet p_T is then shifted either upward or downward and its effect on signal efficiency is measured. Plots showing the effect of JEC uncertainty shifting are shown in Figure 9.2.



(a) 2 b-jet on Higgs mass



(b) 2 b-jet off Higgs mass



(c) 3+ b-jets

Figure 9.2: The effects of JEC uncertainty on signal acceptance. Red lines indicate a 1σ shift up in p_T, blue lines indicate a 1σ shift down in p_T, and the grey filled area is the result of the selection based on the nominal energy-corrected p_T value. The y-axis is in arbitrary units.

Photon ID, b-tag, and JEC scale factor uncertainties are all mass point specific and are applied separately to each as a log-normal uncertainty.

Chapter 10

Results and Limit

10.1 Limit

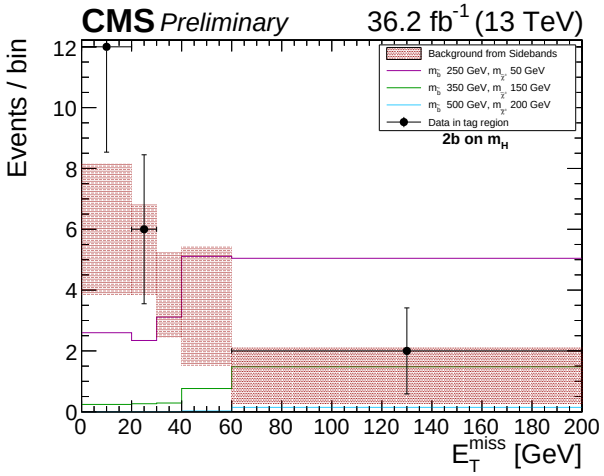
To avoid biasing the choice of event selection criteria, events in the signal region were "blinded" - not counted and observed limits not computed - until selection optimization was fixed. Event selection criteria were not changed after the unblinding of the signal region. Signal criteria were chosen as in Chapter 7 based on their effect on expected sensitivity of the analysis to signal.

Final event counts in the tag region are compared to the background estimation and the signal prediction, for 3 separate signal points. These counts are shown independently for each search category (2 b-jets on Higgs mass, 2 b-jets off Higgs mass, and 3+ b-jets) in Table 10.1. The final yield from data across the E_T^{miss} spectrum is shown in Figure 10.1.

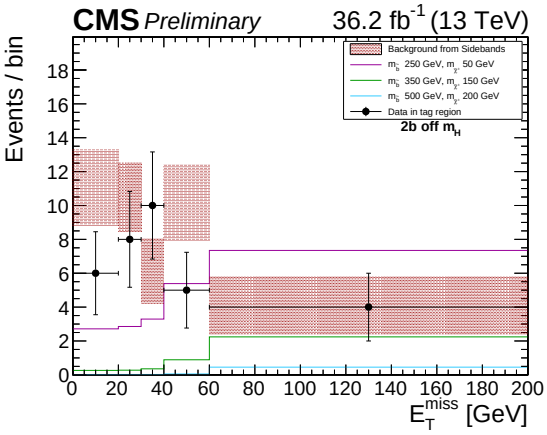
No excess beyond the Standard Model predicted background is observed, and therefore we set a lower limit on the production cross section of sbottom pair production with

$m_{\tilde{b}}, m_{\tilde{\chi}_2^0}$	on Higgs	off Higgs	3+
250 GeV, 50 GeV	18.21	21.60	45.96
350 GeV, 150 GeV	3.03	4.03	8.12
500 GeV, 200 GeV	0.19	0.57	1.22
Predicted Background	19.80 ± 2.36	41.95 ± 3.45	6.88 ± 1.42
Observed	20	33	10

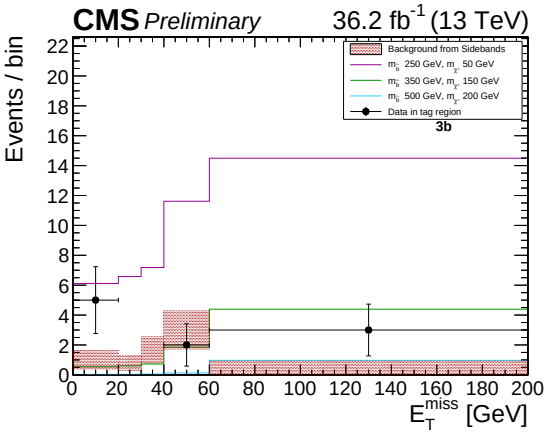
Table 10.1: Predicted signal yield, estimated background, and observed event counts in the $120 < M_{\gamma\gamma} < 131$GeV region.



(a) 2b-jet on Higgs mass channel.



(b) 2b-jet off Higgs mass channel.



(c) 3+ b-jet channel.

Figure 10.1: Final event yields for each search category. Background prediction is the shaded red area, signal prediction for 3 mass points are solid colored lines, and the final data yield in the tag region is represented by black dots.

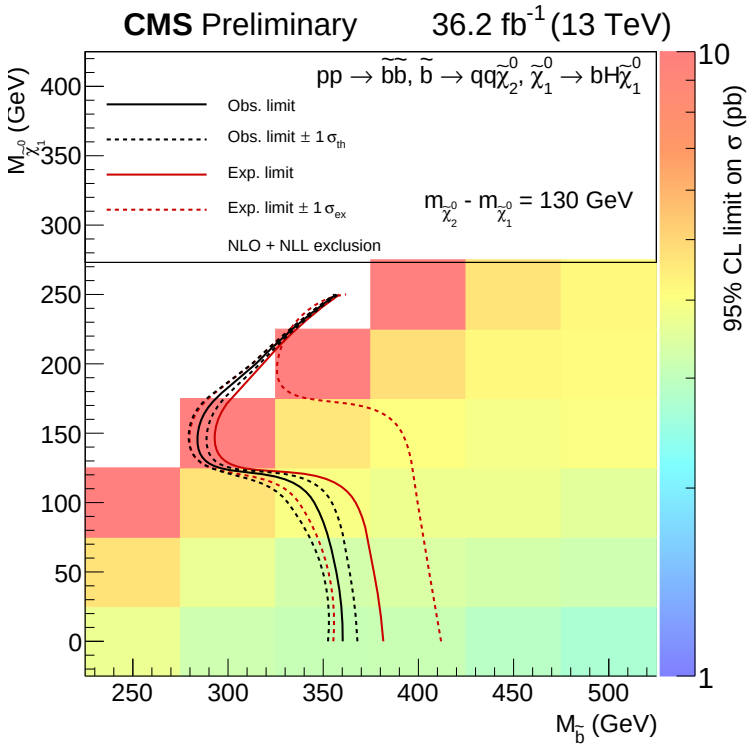


Figure 10.2: The observed 95% confidence level (CL) upper limits on production cross section for sbottom pair production. The solid and dotted red contours represent the expected exclusion region and its $\pm 1\sigma$ bands, while the black contours represent the observed exclusion region and $\pm 1\sigma$ bands. Sbottom production under 350 GeV is excluded for neutralino masses below 150 GeV.

a sbottom decaying to a bottom quark, Higgs boson, and LSP. Limit setting is performed using the RooStats asymptotic CL_s method [11]. This method uses the LHC test-statistic q_μ, where μ represents a signal strength parameter, in this case the production cross section for the process in question. Observed and expected limits are shown in Figure 10.2.

Systematic uncertainties are treated as log-normal nuisance parameters. They are correlated across signal bins, but uncorrelated with each other. All systematics in this analysis are treated in the log-normal method using their multiplicative or scaling effect on signal acceptance and normalization.

(400, 250) GeV	Acceptance (%)	Efficiency (%)
stop-higgsino	80.5	17.0
sbottom-higgsino	77.8	6.6

Table 10.2: Acceptance and efficiency comparison for different signal models with our search strategy. Acceptance is % of simulated photons with $|\eta| < 2.4$ and $p_T > 25$ GeV. Efficiency is of the overall selection, which is the same for both models.

The results are interpreted in terms of the T2bH simplified model outlined in Chapter 2. We exclude bottom squarks with masses below 350 GeV, for neutralinos with masses below 150 GeV. These limits perform somewhat below the limits set in the Run 1 analysis [9] due to the change in signal model. The fine-tuned mass parameter ($\mathrm{m}_{\tilde{\chi}_2^0} - \mathrm{m}_{\tilde{\chi}_1^0} =$ 130 GeV was chosen by the Razor H $\rightarrow \gamma\gamma$ specifically to pursue a slight excess seen in their Run 1 analysis involving a low-p_T Higgs. This low-p_T Higgs in turn negatively impacts acceptance of the Higgs decay products in the form of photons and jets, since they in their turn are softer. Additionally, the original stop-higgsino signal model assumed a massless LSP, whereas the LSP mass in our current model ranges from 1 to 250 GeV.

An unofficial signal sample generated in 2016 using the same methods as the Run 1 stop-higgsino analysis is compared to a sample at the same mass points ($m_{\tilde{t}}/m_{\tilde{b}} = 450$ GeV, $m_{\tilde{\chi}_0^2} = 250$ GeV). Results for simulated photon acceptance and overall selection efficiency are shown in Table 10.2.

Photon reconstruction and selection efficiency suffers due to the low-p_T Higgs and massive LSP, thus lowering selection efficiency and limiting sensitivity. For high-mass LSPs in particular, jet selection efficiency is severely hampered with up to a 60% loss of events in the 2 b-jets off Higgs mass and 3 b-jet categories for a mass difference of 150 GeV between the sbottom and the higgsino.

10.2 Conclusions

A search for a simplified model of supersymmetry with direct production of sbottom quarks with decays to higgsinos and Higgs bosons has been performed using 36.2fb^{-1} of data collected at the CMS experiment at $\sqrt{s} = 13$ TeV. We exploited the narrow invariant mass resonance of the H $\rightarrow \gamma\gamma$ decay, using a sideband method to predict Standard Model background. We find no excess over Standard Model predictions, and set lower limits for the production cross-section of sbottom quarks at a mass of 350 GeV, with higgsino mass 150 GeV or lower.

Bibliography

[1] J Alwall et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.

[2] Anthony Barker. Search for top squark and higgsino production using diphoton Higgs boson decays. CMS Analysis Note, 2013.

[3] S.I. Bityukov and N.V. Krasnikov. Towards the Observation of Signal over Background in Future Experiments. ArXiV report, 1998.

[4] C Borschensky et al. Squark and gluino production cross sections in pp collisions at sqrt(s) = 13, 14, 33 and 100 tev. *EPJ C*, 2014.

[5] O Brüning et al. LHC Design Report Volume 1. Technical report.

[6] O Brüning et al. LHC Design Report Volume 3. Technical report.

[7] S Chatrchyan et al. Interpretation of searches for supersymmetry with simplified models. *Phys. Rev. D*, 88:052017, 2013.

[8] CMS Collaboration. Commissioning and Performance of the CMS Pixel Tracker with Cosmic Ray Muons. Technical report, 2010.

[9] CMS Collaboration. Search for top squark and higgsino production using diphoton Higgs boson decays. *Phys. Rev. Lett.*, 112:161802, 2014.

[10] CMS Collaboration. Search for excess Higgs production in diphoton final states using the Razor variables at sqrt(s) = 13 TeV. CMS Analysis Note, 2016.

[11] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J*, C7:1554, 2011.

[12] Lyn Evans and Lucy Linssen. The Super-LHC is on the starting blocks. *CERN Courier*, 2008. General Photo.

[13] CMS Higgs Working Group. Documentation of the RooStats-based statistics tools for Higgs PAG. Technical report, 2017.

[14] Higgs to gamma gamma Working Group. Further measurement of H to $\gamma\gamma$at sqrt(s) = 13 tev. CMS Analysis Note, 2016.

[15] Kiel Howe and Prashant Saraswat. Excess higgs production in neutralino decays. *JHEP*, 1210:065, 2012.

[16] Joseph Incandela. Status of the CMS SM Higgs Search. Technical report, 2012.

[17] Katja Klein. The CMS Silicon Strip Tracker - Overview and Status. In *International Europhysics Conference on High Energy Physics, Lisbon Portugal*. Proceedings of Science, 2005.

[18] C.M. Kuo, Fabrice Couderc, Pen-Hsuan Wang, RongShyang Lu, and Shilpi Jain. Measurement of 2016 Full dataset SFs of Moriond photon ID. Technical report, 2017.

[19] LHC Higgs Cross Section Working Group. SM Higgs production cross sections at sqrt(s) = 13 TeV (update in CERN Report 4 2016). Technical report, 2016.

[20] Sudhir Malik. The CMS Forward Pixel Detector. In *X Pisa Meeting, Frontier Detectors for Frontier Physics*. Conference Proceedings in Nuclear Instrumentation and Methods, 2006.

[21] Steven A Martin. A Supersymmetry Primer. 1997.

[22] MissMJ. Standard model of particle physics. Jun 2006. General Photo.

[23] Michele Papucci, Joshua Ruderman, and Andreas Weiler. Natural SUSY Endures. *JHEP*, 09:035, 2012.

[24] Kati Lassila Perini. Chapter 2: The Basics of CMS Offline Computing. Technical report, 2010.

[25] CMS BTag POG. Methods to apply b-tagging efficiency scale factors. Technical report, 2015.

[26] CMS BTag POG. Usage of b/c Tag Objects for 13 TeV Data in 2016 and 80X MC. Technical report, 2017.

[27] CMS EGamma POG. ECAL clustering. Technical report, 2009.

[28] CMS EGamma POG. Cut Based Photon ID for Run 2. Technical report, 2016.

[29] CMS JetMET POG. Jet Identification. Technical report, 2016.

[30] CMS JetMET POG. Jet identification in high pile-up environment. Technical report, 2017.

[31] T Sjöstrand et al. An Introduction to PYTHIA 8.2. *Computer Physics Communications*, 2014.

[32] The CMS Collaboration. CMS Physics: Technical Design Report Volume 1: Detector Performance and Software. Technical report, 2006.

[33] The CMS Collaboration. Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus and Missing Transverse Energy. CMS Physics Analysis Summary PFT-09-001, 2009.

[34] The CMS Collaboration. Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at sqrt(s) = 7 TeV. *JINST*, 8:09, 2013.

[35] The CMS Collaboration. Identification of b-quark jets with the CMS experiment. *JINST*, 18:04, 2013.

[36] The CMS Collaboration. Performance of MET reconstruction and pileup mitigation techniques in CMS. *Nuclear Physics B, Proceedings Supplements*, 2014.

[37] The CMS Collaboration. Performance of photon reconstruction and identification with the cms detector in proton-proton collisions at sqrt(s) = 8 tev. *JINST*, 10, 2015.

[38] The CMS Collaboration. CMS Luminosity Measurements for the 2016 Data Taking Period. CMS Physics Analysis Summary, 2017.

[39] Xavier Vidal and Ramon Manzano. Supersymmetry. General Photo.

Appendix A

b-tag Scale Factor Application

Event reweighting according to the b-tag scale factor is accomplished via a jet-by-jet comparison of efficiencies and misidentification rates, with each jet in the event contributing a scale factor to the final product or event weight. Various methods for the computation of this weight are used depending on the preference of the analyst. Our analysis aims to predict the correct event yield in data by only changing the weight of the selected simulation events - meaning events that do not pass selection are not considered in the weighting procedure [25]. The following method was developed by the CMS BTag POG and is designed to keep events with a specific b-tag multiplicity from migrating between categories (number of b-tags). Since the signal categorization is dependent on b-tag multiplicity this specific method is required to provide accurate weighting without affecting the final signal yield.

Events passing selection will have at least two b-tagged jets, with all jets passing the CSV-Loose tag and at least one passing the CSV-Medium tag. The event weight, w, is defined in Eq. A.1:

$$w = \frac{P(\text{DATA})}{P(\text{MC})} \tag{A.1}$$

The probabilities are defined so as to account for multiple working points. It is possible that a jet might be tagged with the CSV-Loose working point but not the CSV-Medium point, which affects the probability according to the efficiency of the tagger at the given working point. For simulation (MC), the probability is given in Eq. A.2 and for data, the probability is given in Eq. A.3.

$$P(\text{MC}) = \prod_{i=\text{tagged M}} \varepsilon_i^M \prod_{j=\text{j = tagged L, not M}} \left(\varepsilon_j^L - \varepsilon_j^M\right) \prod_{k=\text{not tagged}} \left(1 - \varepsilon_k^L\right) \tag{A.2}$$

$$P(\text{DATA}) = \prod_{i=\text{tagged M}} \text{SF}_i^M \varepsilon_i^M \prod_{j=\text{j = tagged L, not M}} \left(\text{SF}_j^L \varepsilon_j^L - \text{SF}_j^M \varepsilon_j^M\right) \prod_{k=\text{not tagged}} \left(1 - \text{SF}_k^L \varepsilon_k^L\right) \tag{A.3}$$

Scale factors ($\text{SF}_{i,j,k}$) are dependent on the p_T, η, and flavor of the jet, largely due to varying detector response based on these values.