Search for Production of Supersymmetric Top Quarks in Hadronic and Multi-leptonic Final States, Using a Deep Neural Network Based Soft B-tagger For Compressed Mass Scenarios

Erich Schmitz

Submitted to the graduate degree program in Department of Physics and Astronomy and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Committee members

Alice Bean, Chairperson

Ian Lewis

Russell Ostermann, External Reviewer

Christopher Rogan

Stephen Sanders

Graham Wilson

Date defended: February 8, 2021

The Dissertation Committee for Erich Schmitz certifies
that this is the approved version of the following dissertation :

Search for Production of Supersymmetric Top Quarks in Hadronic and Multi-leptonic Final States, Using a Deep Neural Network Based Soft B-tagger For Compressed Mass Scenarios

Alice Bean, Chairperson

Date approved: February 11, 2021

Abstract

A search is performed for pair produced supersymmetric top (stop) quarks in hadronic and multi-leptonic final states. The search uses a sample of proton-proton collision data at $\sqrt{s} = 13$ TeV, corresponding to 137 fb^{-1}, recorded by the Compact Muon Solenoid (CMS) experiment at the Large Hadron Collider (LHC). The searches are focused on events with a high transverse momentum system from initial-state-radiation jets recoiling against a potential supersymmetric particle (sparticle) system with significant missing transverse momentum. Stop signals which have small mass splittings between the stop and the lightest supersymmetric particle (LSP) on the order of 10s of GeV are studied for stop masses ranging from 400 to 1500 GeV. This dissertation probes the compressed mass phase space through the use of Recursive Jigsaw Reconstruction (RJR) by assigning reconstructed objects to the initial state radiation or sparticle system following a generic decay tree, and using this assignment to take advantage of mass sensitive variables in different rest frames. A new Deep Neural Network based b quark tagger has been developed to find low p_T b quarks using secondary vertices. The signal regions are defined by the multiplicity of reconstructed objects in each of the two systems, including leptons, jets, soft b-tagged secondary vertices, and b-tagged jets. Limits are placed on the pair production of stops quarks and are interpreted within the framework of simplified models. Exclusions at 95% Confidence Level (CL) are expected for stop masses up to 675 GeV for neutralino masses up to 665 GeV, where the neutralino is assumed to be the lightest supersymmetric particle.

The last part of the dissertation details a project, independent of the stop search, which looks at calculating the location of the CMS beam spot using tracking independent methods. A method was developed, making use of a maximum likelihood fit, which only uses the cluster occupancy and x, y, and z positions of the read out chips located in the first layer of the barrel pixel detector, and is accurate to within 1 mm of the true beam spot when tested on simulated Monte Carlo (MC).

Acknowledgements

It has been a long journey leading up to the writing of this dissertation, and there are many people without whom I would not have been able to finish.

The first, and most important person I would like to thank is my research advisor, Dr. Alice Bean, for being my guide throughout my graduate career. From the beginning she allowed me to immediately get involved in research before I even officially became a graduate student, and at the end was willing to get involved with a SUSY search, a topic for which she has been a vocal opponent. She has been a constant support while I have been at KU and has been immensely helpful with all of the gentle reminders and prodding to get my work done, despite my forgetfulness, and by doing simple things such as checking up on me and asking how my day is going. I would also like to thank Drs. Ian Lewis, Christopher Rogan, Stephen Sanders, Graham Wilson, and Russell Ostermann for agreeing to serve on my committee.

Secondly, I would like to thank the KU HEP analysis group for the work they have done to help put the analysis together, and especially to Dr. Christopher Rogan, who not only agreed to be on my committee, but also without whom we would not have this SUSY analysis. His drive and dedication has been an integral factor in getting this analysis to where it is today.

The whole PhD would not have been possible without funding, so I would first like to thank the National Science Foundation, who has supported much of my work through multiple grants. This work includes my first project working on the Phase I pixel detector upgrade, the first published analysis that I was a main contributor to, my many trips to CERN, and finally my work on this dissertation. Secondly I would like to thank the University of Kansas for providing me a dissertation fellowship in the Fall 2020 semester. I would also like to thank the Lowry family and the KU Physics Department for awarding me the Lowry Graduate Fellowship and Carl and Catherine Chaffee scholarship in my first year, and again for the C.C. Chaffee Scholarship in the Fall of 2019.

Lastly I would like to thank my family. For my parents, Anthony and Theresa Schmitz, who have always been there for me, and will continue to far into the future. For my wife, Katherine, who has been with me for a majority of my graduate schooling, and has provided me with all the love and companionship (and food) I could ever need. Finally, for my son, Benedict, who started off as a thesis baby, and has since turned into a thesis toddler, and who has been a light that has kept me from working too hard, especially during naptime, where we could both just lay down and rest.

Thank you all, without your help and support, I would not be where I am today.

Contents

List of Figures

List of Tables

Chapter 1

The Standard Model and Supersymmetry

1.1 Introduction

In particle physics there are many theories that attempt to explain how matter in its smallest form interacts with other matter through the fundamental forces. The Standard Model (SM) of particle physics stands at the forefront of these theories in describing the interactions of three of the four fundamental forces and the elementary particles. As a well tested theory, it can explain many of the particle phenomena we see today, but it is still incomplete. Some theories that attempt to account for the deficiencies in the SM look for physics "Beyond the Standard Model", a popular one being Supersymmetry (SUSY), which introduces a new class of elementary particles.

For this dissertation a search is performed for one of these new particles, called the stop, which is the supersymmetric partner to the third generation top quark. This search looks for stops which have a small mass splitting between it and the lightest supersymmetric particle (LSP). This scenario, commonly referred to as a compressed mass scenario, creates soft visible decays, where b quarks can be expected. One of the mainstays of this search is the production of a new discriminator, using a Deep Neural Network (DNN) that trains on low-momentum objects, that is used to identify these soft b quarks.

This dissertation starts with a summary of the Standard Model, including a brief history, and then describes the limitations of the SM, leading into a summary of Supersymmetry and how it supplements the SM. Following that, the current state of experimental searches in SUSY is given. In Chapter 2 the Large Hadron Collider (LHC) and Compact Muon Solenoid (CMS) are described in some detail, with some focus given to the Pixel Detector of the CMS experiment

and its Phase I upgrade. Chapter 3 gives an overview of the analysis, and includes descriptions of the event reconstruction, the use of Recursive Jigsaw Reconstruction (RJR), and the pertinent stop signal modes. Chapter 4 goes through the production of a DNN based soft b-tagger using low p_T secondary vertices (SVs), and its use within the analysis. Chapter 5 details the event selection, the object reconstruction, and the building of the analysis regions using object counting and binning in mass sensitive variables produced using RJR. The results of the analysis are given in Chapter 6, and includes the signal interpretations. A summary of the dissertation is then given in Chapter 7. Lastly, Appendix A goes through a project independent of the analysis, that involves the measurement of the CMS beam spot, without making use of track reconstruction.

1.2 A Brief History of the Standard Model

The current formulation of the Standard Model, to be described in more detail in Section 1.3, began to take form with the development of a non-Abelian gauge theory by Chen Ning Yang and Robert Mills [9], which was later used to describe the interactions of the strong force to form the theory that is now known as Quantum Chromodynamics (QCD) [10]. In the 1960s, Glashow put forth a partially symmetric theory of the weak and electromagnetic interactions [11], which would be the basis for the standard model of electroweak interactions. With the addition of the Higgs mechanism to describe the mass generation of the SM fields [12, 13, 14, 15, 16], the basis for the current formulation of the SM was completed.

To date, most experiments that have tested the interactions of the strong, weak, and electromagnetic forces have had results which are shown to be in agreement with the predictions of the SM. This is seen most clearly with the discovery of the elementary particles that are predicted by the SM, such as the discovery of the quarks (including the top quark) [17, 18, 19, 20, 21, 22, 23], the different flavors of neutrinos [24, 25], and most recently the Higgs boson [26, 27]. While experiments have supported what the SM predicts, there are many phenomena that the SM is unable to explain. Some problems of note that the SM cannot account for are the gravitational force, the matter-antimatter asymmetry, the existence of dark matter and dark energy, and the gauge hierar-

chy problem, all of which will be detailed in Section 1.4.

These problems naturally lead to the idea of theories which explore physics "Beyond the Standard Model" and that can be used to explain the basis for the unexplained phenomena. One such theory is Supersymmetry (SUSY), an extension to the Standard Model which adds a new class of particles and provides an elegant solution to many of the problems found within the SM.

1.3 The Standard Model

The Standard Model of particle physics is a gauge theory based on the gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$, and is used to describe the interactions of the strong, weak, and electromagnetic forces. Each force corresponds to a gauge symmetry within the SM gauge group; $SU(3)_C$ for the strong interaction and $SU(2)_L \times U(1)_Y$ for the electroweak interaction, with the sub group $U(1)_{EM}$ corresponding to the electromagnetic interaction. There are four sets of quantum fields associated with the SM, these are the gluon fields G_μ^C, which transform under $SU(3)_C$ [28]; the electroweak fields W_μ^1, W_μ^2, W_μ^3, and B_μ, which transform under $SU(2)_L \times U(1)_Y$; the fermion fields ψ, which represent matter particles; and lastly the Higgs field ϕ, a complex scalar field which is used to generate mass under a spontaneous symmetry breaking of the $SU(2)_L \times U(1)_Y$ gauge group [29]. The interactions of these fields can be described using a field theory Lagrangian, given in a very simplified short form in Eq. (1.1), a form used to summarize the equation as far as possible without losing too much information.

$$\begin{aligned} \mathscr{L} = & -1/4 F_{\mu\nu} F^{\mu\nu} \\ & + i\bar{\psi} \not{D} \psi \\ & + \bar{\psi}_i y_{ij} \psi_j \phi + h.c. \\ & + |D_\mu \phi|^2 - V(\phi) \end{aligned} \tag{1.1}$$

The first line of Eq. (1.1) is for the interactions of the gauge fields, corresponding to the G_μ^C, W_μ^i,

and B_μ fields, with $F_{\mu\nu}$ being the field strength tensor. The second line describes the interactions of the fermion fields with the gauge fields, with $\not{D}$ used as a shorthand for $\gamma^\omega D_\omega$, the gauge covariant derivative. The third line gives the mass terms for the fermion fields, with y_{ij} being the Yukawa mass coupling, which describes the interaction between the fermion fields and the Higgs field. The last line gives the mass terms for the gauge fields, with the first term showing the interaction between the gauge fields and the Higgs field, and the second term showing the Higgs self coupling.

Each of the SM fields correspond to physical particles. Of the force carriers there are eight gluons which directly correspond to the gluon fields, and mediate the strong force. The $W^\pm$, and Z bosons mediate the weak force, and the photon mediates the electromagnetic force. The $W^\pm$, Z, and photon arise out of a combination of the W^i_μ and B_μ fields through the Higgs mechanism. The Higgs boson is the resulting physical particle that is left over from the scalar field after the spontaneous symmetry breaking. The remaining particles are the fermions, consisting of the quarks and leptons. There are a total of 6 quarks (u, d, s, c, b, t) and 6 leptons $(e, \mu, \tau, \nu_e, \nu_\mu, \nu_\tau)$, separated into 3 generations each. Each quark and lepton has its own anti-particle. The quarks mainly interact through the strong force, but also have charges which allow them to interact through the weak and electromagnetic forces. The leptons only interact through the electromagnetic and weak forces. As fermions, both quarks and leptons have a weak hypercharge, and for the case of the left-handed particles, a $SU(2)_L$ charge, which allows them to interact through the electroweak forces.

1.4 Limitations of the Standard Model

Experiments have shown that up to the electroweak energy scale set at approximately 246 GeV, the value of the vacuum expectation value, the Standard Model is an accurate theory and even then, it still has some deficiencies. There are two main deficiencies related to the SM; the first corresponding to the phenomena that cannot be explained by the current theory, and the second being inconsistent scales within the model.

1.4.1 Unexplained Phenomena

On the subject of unexplained phenomena, the Standard Model is unable to deal with gravity, dark matter, dark energy, or the matter-antimatter asymmetry. The SM as it is does not deal with gravity at all. At the current observable energy scales the gravitational force has an insignificant effect on the elementary particles, with the force being 10^{24} times weaker than the weak force, the next weakest of the forces in terms of effective field strength. Gravity also mainly works on a macroscopic scale, requiring large mass and distance scales for its effects to be seen in any significant way.

For dark matter and dark energy, the SM does not have a viable candidate that could take the place of either of these phenomena. Dark matter requires a non-interacting, or at least a weakly-interacting particle. The closest candidate to this would potentially be the neutrino, but as it is close to massless, it would not be able to account for the amount of matter necessary according to recent observations. The amount of matter that the SM can account for in the universe is approximately 5% of the mass-energy content, the remaining amount is dark matter (20%) and dark energy (75%) [30].

Lastly, there is the matter-antimatter asymmetry. Astronomical observations show that the universe is mostly made out of matter. The main observational evidence for this is that there are almost no byproducts of matter-antimatter annihilation events seen in the visible universe. This leads to the conclusion that in the early universe there must have been a small imbalance between matter and antimatter such that when most of the particles annihilated with each other, only regular matter was left over. The problem with this as it pertains to the SM is that there are no sources that can adequately explain the size of the imbalance necessary to create the matter in the universe. While direct CP violation, a type of symmetry breaking that could be a source of the matter-antimatter asymmetry, is allowed through a complex phase in the Cabibbo-Kobayashi-Maskawa (quark mixing) matrix, which gives information on the flavor-changing weak interaction of the quarks [31], the size of such a complex phase would not be enough to explain the asymmetry. Experimentally, there have been observations of CP violation in neutral kaons [32, 33, 34], B-mesons [35, 36],

strange B mesons [37], D mesons [38], and most recently there have been indications of CP violation in neutrinos [39]. The size of the CP violation in each of these cases, though, has so far been unable to explain the source of the asymmetry.

1.4.2 The Problem of Scale

In addition to the unexplained phenomena that the SM does not explain, there are also some potential problems within the model itself, granted these problems may be more aesthetic than physical depending on the interpretation and what energy scale is being considered. These problems stem from the idea of naturalness, a property where the dimensionless ratios between the free parameters or physical constants of a theory should have values within an order of magnitude. In addition, any free parameters should not have any fine-tuning applied.

Within the Standard Model the main problem of note is the Hierarchy problem, which refers to how the Higgs mass is multiple orders of magnitude smaller than the Planck scale. How this theoretically affects the SM can be seen in the scalar potential of the Higgs boson and its subsequent mass term, where experimentally this is currently not an observable effect. The Higgs potential can be written in terms of the physical Higgs boson as given in Eq. (1.2),

$$V(H) = \lambda v^2 H^2 + \lambda v H^3 + \frac{\lambda}{4} H^4 \tag{1.2}$$

where the first term gives the physical mass, the second contains the trilinear self coupling and the third the quadrilinear self coupling [40]. These terms can also be represented by the three tree level diagrams of the Higgs self coupling as seen in Figure 1.1.

The quadratic term would represent a Higgs traveling through space from one point to another. The cubic term would correspond to the interaction of a Higgs scattering off of a second Higgs. The quartic term would then be two Higgs interacting to produce another two Higgs. For the purpose of the hierarchy problem, the quartic coupling can be converted to a one-loop level correction to the physical mass term as seen in Figure 1.2, where a Higgs absorbs and then emits a Higgs as it travels

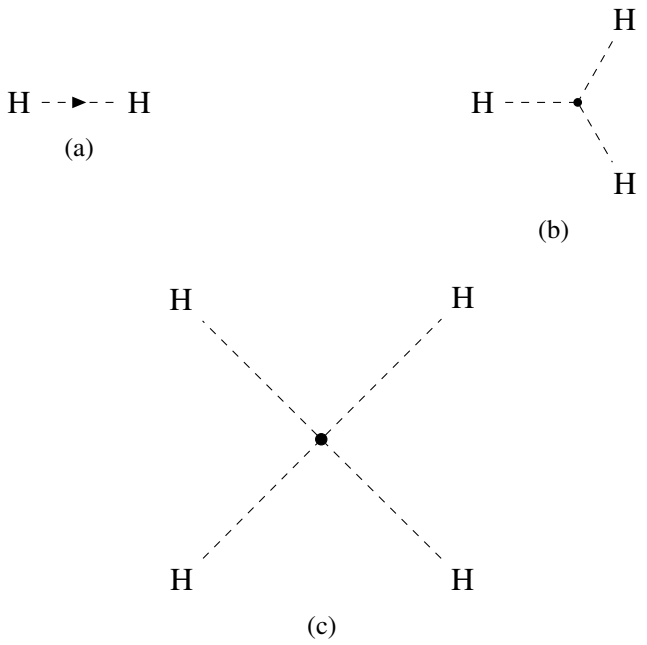


Figure 1.1: Feynman diagrams depicting the quadratic (propagator) (a), cubic (b), and quartic (c) couplings of the Higgs boson. These were drawn using [1].

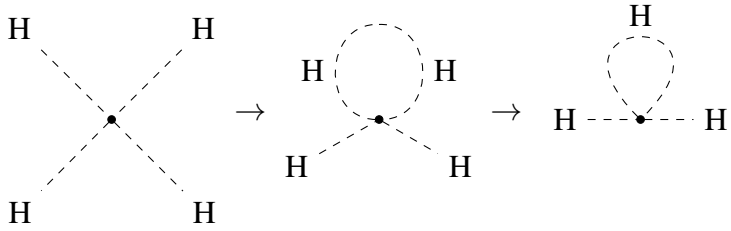


Figure 1.2: Feynman diagrams depicting the conversion of the quartic Higgs coupling to a one-loop propagator where the lines of the two upper Higgs combine to form a loop. These were drawn using [1].

from one point to another. Adding this correction, called the self-energy radiative correction, we get Eq. (1.3) for the Higgs mass.

$$M^2_{physical} = M^2_H + \frac{3\lambda}{8\pi}\Lambda^2 - \frac{3\lambda}{8\pi}M^2_H \log\left(\frac{\Lambda^2 + m^2_H}{M^2_H}\right) \tag{1.3}$$

The Λ parameter in Eq. (1.3) is the Ultra Violet cutoff, the point at which the Standard Model can be considered valid. This cutoff is presumed to be on the order of the Planck scale. Such a scale produces a quadratic divergence in the correction, by which the Higgs boson mass would be made large to a degree where $M^2_{physical} \gg M^2_H$. With the observed mass of the Higgs being of the order of 100 GeV there would need to be some fine-tuning applied to the relation between the radiative corrections and the bare Higgs mass, such that the correction is canceled out. Fermions can also contribute to the Higgs mass through their one loop contributions, similar to the Higgs self-energy correction, which further introduces quadratic divergences.

In the SM the Higgs sector lacks a symmetry that can protect against such a large radiative correction to its mass. A solution then would need to be found beyond the Standard Model. There have been a couple attempts to deal with the quadratic divergence which includes having the elementary particles be composite objects, the idea of Technicolor [41, 42], where fundamental scalars are instead composites of new fermions, and lastly there is the addition of a higher symmetry which can be used to eliminate the quadratic divergence present within the Higgs mass [7]. Supersymmetry falls under this third option and will be the focus of this dissertation.

1.5 Supersymmetry

Supersymmetry is an extension of the Standard Model that introduces a new symmetry that relates bosons and fermions. It provides a framework for incorporating gravity, an explanation for the large gauge hierarchy, a source for dark matter, and in some cases can account for the matter-antimatter asymmetry through leptogenesis.

In SUSY each SM particle will have an associated "superpartner" (sparticle) which, in the sim-

plest theories, will have the same quantum numbers with the exception of its spin, which will differ by a half-integer. SM fermions will have a boson as a partner, and the SM bosons will similarly have a fermion as a partner. As a symmetry, the masses of the sparticles should be equivalent to their SM counterparts, but only if it is an exact symmetry. From experimental observations it is not possible for SUSY to be an exact symmetry, as there have been no discoveries of fundamental scalar particles that have the same mass as the known fundamental fermions. This means that it is a broken symmetry, where the masses of the sparticles will be heavier than the equivalent SM particles. As a broken symmetry there is a limit to how big the sparticle masses can get before the SM problems that were initially solved by SUSY, come back in a new form. The main problem that would reoccur is the hierarchy problem, as new large radiative divergences would be introduced by the sparticles. This would require the sparticles to have masses on the order of the TeV scale, and a spontaneous symmetry breaking that will preserve the gauge invariance and renormalizability of the theory [43].

There are multiple SUSY theories available, and the main difference between each of them is how they deal with the symmetry breaking. This affects how the sparticles interact, what masses they have after the symmetry breaking, and the content of the superfields that partner to the SM fields. The most economic theory available is the Minimal Supersymmetric Standard Model (MSSM), which takes the minimal number of gauge groups, particle groups, Yukawa interactions, and soft symmetry breaking terms. The MSSM will be discussed in detail in section 1.5.3, after going through a general list of the sparticles found in SUSY (sec. 1.5.1), along with an explanation of how the Hierarchy problem is taken care of (sec. 1.5.2).

1.5.1 The S-particle Spectrum

Supersymmetry (SUSY) relates the fermionic and bosonic degrees of freedom [7]. This relation predicts the sparticles as partners to SM particles. A requirement of this symmetry is that the number of degrees of freedom in the bosons and fermions must be equal between SUSY and the SM. This means that at a minimum there must be one supersymmetric particle for every SM

particle. For example, for every quark and anti-quark in the SM, there will be a pair of complex scalar quarks and anti-quarks in SUSY. These two scalars are for the left and right handed chiralities of the SM quarks. As a way to differentiate particles between SUSY and the SM, the equivalent SUSY particles will have an *s*-prefix for the partners to fermions (where the s stands for scalar), and an *-ino* suffix for the partners to bosons.

At a minimum any Supersymmetric theory will have three generations of scalar quarks (s-up, s-down, s-charm, s-strange, s-top, s-bottom), and leptons (s-electron, s-muon, s-tau, s-neutrinos). For the SM bosons there will be the fermionic winos, bino, zino, photino, and gluino; collectively called the gauginos, and then the higgsinos. The winos consist of three states, two charged and one neutral, related to the W bosons and the $SU(2)$ gauge fields. The bino is then the superpartner of the $U(1)_Y$ gauge field. In some cases there will also be the gravitino, should gravity be explicitly incorporated. In addition to the gauginos there are also the neutralino and chargino, which are the generalized mass eigenstates of the gauginos. If the mass eigenstates have specific couplings then they will have a more specific name such as wino, higgsino, or wiggsino (a mixture of higgsino and wino) for the charginos; corresponding to a W-like, Higgs-like or intermediate couplings respectively, and the photino, zino, higgsino, ziggsino (a mixture of zino and higgsino), wino, and bino equivalently for the neutralinos. Table 1.1 gives a list of the sparticles and their corresponding symbols.

There is an up-type and down-type pair of higgsinos in SUSY, corresponding to a pair of scalar Higgs doublets that are added on the SM side. This two doublet requirement is motivated by the fact that certain Higgs-fermion interaction terms in the SM are not allowed in SUSY. Specifically, the complex conjugates of the Higgs doublets cannot be used in the mass generation term in SUSY. As it is, having one Higgs doublet would only allow for either the up- or down-type quarks to acquire mass, since in the SM one quark type will get its mass from the complex conjugate of the Higgs doublet, whereas by adding another Higgs doublet on top of the one already present, both types will attain masses [7].

The sparticles will generally be produced and decay in the same manner as their SM partners,

SM Particles		SUSY Particles					
		Weak eigenstates		generic mass eigenstates		specific mass eigenstates	
name	symbol	name	symbol	name	symbol	name	symbol
quarks	q	scalar quarks	$\tilde{q}_L$, $\tilde{q}_R$	-	$\tilde{q}_1$, $\tilde{q}_2$	-	-
leptons	ℓ	scalar leptons	$\tilde{\ell}_L$, $\tilde{\ell}_R$	-	$\tilde{\ell}_1$, $\tilde{\ell}_2$	-	-
neutrinos	ν	scalar neutrinos	$\tilde{\nu}$	-	-	-	-
gluons	g	gluino	$\tilde{g}$	-	-	-	-
W boson	$W^{\pm}$,	wino	$\tilde{W}^{\pm}$	charginos	$\tilde{\chi}^{\pm}_{1,2}$	wino	$\tilde{w}^{\pm}$
Higgs boson (charged)	H_u^+, H_d^-	higgsino	$\tilde{H}_u^+$, $\tilde{H}_d^-$			higgsino	$\tilde{h}^{\pm}$
						wiggsino	$\tilde{\omega}_{1,2}$
photon	γ	photino	$\tilde{\gamma}$	neutralinos	$\tilde{\chi}^0_{1,2,3,4}$	photino	$\tilde{\gamma}$
Z boson	Z	zino	$\tilde{Z}$			zino	$\tilde{z}$
Higgs boson (neutral)	$H^0_{u,d}$	higgsino	$\tilde{H}^0_{u,d}$			higgsino	$\tilde{h}_{1,2}$
						ziggsino	$\tilde{\zeta}_{1,2}$
W gauge field	W^0	wino	$\tilde{W}^0$			wino	$\tilde{w}^0$
B gauge field	B	bino	$\tilde{B}$			bino	$\tilde{b}$

Table 1.1: The table gives a list of the Supersymmetric particles. Also shown are the symbols used to denote each particle, along with detailing the associated SM particle. This table is made following a similar table given in [7]. A dash denotes that the entry remains unchanged with respect to the previous entry.

with some minor differences and special cases. These decay modes will be detailed in section 1.5.3. The last sparticle to be mentioned is the generically named Lightest Supersymmetric Particle (LSP), which is one that is stable, has a neutral charge and lightly interacts with SM particles. The most likely candidate for the LSP would then be a neutralino, which would correspond to a mixture of the neutral higgsino, wino, and/or bino. Due to the features of the LSP it makes a good dark matter candidate, falling under the category of weakly interacting massive particle (WIMP).

1.5.2 Solving the Hierarchy Problem

As mentioned in section 1.4.2, one of the problems in the Standard Model is the existence of quadratic divergences in radiative corrections to the Higgs mass. This essentially disappears in a supersymmetric extension to the SM through the inclusion of a new symmetry which creates super-partners. These new particles will also introduce quadratic divergences due to radiative corrections, but of opposite sign to the radiative corrections from the SM particles.

Following the description of the Hierarchy problem given in ref. [43], fermions will provide a correction to the Higgs mass, in the same way that the Higgs introduces its own radiative correction, as given in Eq. (1.3). For a number of heavy fermions, on the order of the top mass, a one-loop correction, following the work done by [44], is given by Eq. (1.4).

$$\Delta M_H^2 = N_f \frac{\lambda_f^2}{8\pi^2}\left[-\Lambda^2 + 6m_f^2\frac{\Lambda}{m_f} - 2m_f^2\right] + \mathscr{O}\left(1/\Lambda^2\right) \tag{1.4}$$

As in Eq. (1.3), the correction from the heavy fermions introduce a quadratic divergence in the cutoff scale parameter, Λ. Now one can consider the set of scalar fermions (sfermions) with a trilinear coupling $v\lambda_S$, quadrilinear coupling λ_S, and mass m_S. These scalars will give two contributions to the Higgs mass, corresponding to the two one-loop diagrams of a scalar particle contributing to the Higgs self-energy (one loop for the trilinear coupling, and a second loop for the quadrilinear coupling). The correction to the Higgs mass then becomes Eq. (1.5).

$$\Delta M_H^2 = \frac{\Lambda_S N_S}{16\pi^2}\left[-\Lambda^2 + 2m_S^2\log\left(\frac{\Lambda}{m_S}\right)\right] - \frac{\Lambda_S^2 N_S}{16\pi^2}v^2\left[-1 + 2\log\left(\frac{\Lambda}{m_S}\right)\right] + \mathscr{O}\left(\frac{1}{\Lambda^2}\right) \tag{1.5}$$

A relation between the coupling parameters of the fermions and sfermions can be found such that the quadratic divergences cancel out, in this case $\lambda_f^2 = 2m_f^2/v^2 = -\lambda_S$ and $N_S = 2N_f$. Adding the contributions of Eqs. (1.4) and (1.5) together after making the relevant substitutions will give Eq. (1.6).

$$\Delta M_H^2 = \frac{\lambda_f^2 N_f}{4\pi^2}\left[\left(m_f^2 - m_S^2\right)\log\left(\frac{\Lambda}{m_S}\right) + 3m_f^2\log\left(\frac{m_S}{m_f}\right)\right] + \mathscr{O}\left(\frac{1}{\Lambda^2}\right) \tag{1.6}$$

From this equation it is seen that the quadratic divergence disappears from the Higgs mass correction, while a logarithmic divergence of a much smaller scale remains. It was mentioned at the beginning of sec. 1.5 that the sparticles would have to be on the order of 1 TeV, else another radiative divergence would be introduced. The logarithmic divergence seen in Eq. (1.6) is such a divergence. As the sparticles get larger in mass, so does this divergence, while it remains small

for masses on the same order as the SM fermions and even disappears completely with a perfect mass symmetry between the sparticles and SM particles. With the divergences to the Higgs mass removed and/or reduced, the symmetry found in SUSY protects the Higgs mass, and explains the difference in scale between the mass and the UV-cutoff scale. Trying to maintain this protection also sets the scale for what masses can be expected from the sparticles after the symmetry breaking.

1.5.3 Minimal Supersymmetric Standard Model

The Minimal Supersymmetric Standard Model is an extension to the Standard Model that adds a minimal number of parameters on top of the existing SM parameters. There are four basic assumptions pertaining to this [40, 7]; a minimal gauge group, minimal particle content, R-parity conservation, and minimal soft symmetry breaking terms.

The gauge group of MSSM is the same as for the Standard Model, $SU(3)_C \times SU(2)_L \times U(1)_Y$. Each symmetry in this gauge group has a corresponding superfield consisting of a vector supermultiplet. There is the $\hat{G}$ superfield of $SU(3)$ consisting of the 8 gluons (g) and 8 gluinos ($\tilde{g}$), the $\hat{W}$ superfield of $SU(2)$ consisting of the W^0 and $W^\pm$ fields and the three winos ($\tilde{W}^0$, $\tilde{W}^\pm$), and then the $\hat{B}$ superfield of $U(1)$ consisting of the B field and the bino ($\tilde{B}$). The Z, zino, photon, and photino are not considered amongst these, as they are just linear combinations of the neutral components of the $\hat{W}$ and $\hat{B}$ superfields. The mixtures and pure states of higgsinos, winos, and binos are collectively referred to as electroweakinos.

The particle content will almost directly mirror what is in the Standard Model. There are three generations of quarks/squarks and leptons/sleptons, and no right handed neutrinos/sneutrinos. The left and right handed quarks and leptons will be in chiral supermultiplets with their scalar superpartners. There are also the chiral superfields of the Higgs, one for each of the two complex doublets and their superpartners. All of the superfields and their particle content are given in table 1.2.

In order to ensure lepton and baryon number conservation with the added superpartners, a new symmetry called R-parity is introduced [45], given by Eq. (1.7),

Super-multiplet	Super-field	Boson field	Fermion partner
gluon, gluino	$\hat{G}$	g	$\tilde{g}$
W gauge field, wino	$\hat{W}$	$W^{\pm}, W^0$	$\tilde{W}^{\pm}, \tilde{W}^0$
B gauge field, bino	$\hat{B}$	B	$\tilde{B}$
slepton, lepton	$\hat{L}$	$\tilde{\nu}, \tilde{e}_L$	ν, ℓ_L
$\times 3$	$\hat{\ell}$	$\tilde{\ell}^*_R$	$\ell^{\dagger}_R$
squark, quark	$\hat{Q}$	$\tilde{u}_L, \tilde{d}_L$	u_L, d_L
$\times 3$	$\hat{u}$	$\tilde{u}^*_R$	$u^{\dagger}_R$
	$\hat{d}$	$\tilde{d}^*_R$	$d^{\dagger}_R$
Higgs, higgsino	$\hat{H}_u$	H^+_u, H^0_u	$\tilde{H}^+_u, \tilde{H}^0_u$
	$\hat{H}_d$	H^0_d, H^-_d	$\tilde{H}^0_d, \tilde{H}^-_d$

Table 1.2: A list of the superfields in the MSSM, written in a format similar to a table given in [8], with the exception of a change in nomenclature for some of the superfields, to make their associations more explicit

$$R_p = (-1)^{2S+3B+L} \tag{1.7}$$

where S is the spin and L and B are the lepton and baryon numbers respectively. This equation gives a quantum number to each of the particles, with a value of $R_p = +1$ for SM particles and $R_p = -1$ for the SUSY partners. A consequence of this symmetry is that sparticles will always be produced in pairs, and must decay to an absolutely stable LSP [43].

The interactions of the superfields will be described by a renormalizable supersymmetric Lagrangian, made up of a Kähler potential, a superpotential, and a gauge kinetic function [8]. Each of these are functions of the chiral superfields, where the gauge kinetic function and Kähler potential will have the simple forms of $f_{ab} = \delta_{ab}(1/g_a^2 - i\Omega_a/8\pi^2)$ and $K = \phi_i \tilde{\phi}^{i*}$ at tree level respectively [46]. The superpotential is more complicated in that it includes the Yukawa couplings and supersymmetric Higgs mass term. The superpotential for the MSSM is given in Eq. (1.8).

$$W = \sum_{i,j=gen} -Y^u_{ij}\hat{u}_{Ri}\hat{H}_u\hat{Q}_j + Y^d_{ij}\hat{d}_{Ri}\hat{H}_d\hat{Q}_j + Y^{\ell}_{ij}\hat{\ell}_{Ri}\hat{H}_d\hat{L}_j + \mu\hat{H}_u\hat{H}_d \tag{1.8}$$

with Y_{ij} being the Yukawa coupling, and μ the SUSY conserving higgsino mass parameter. This superpotential has inherent R-parity conservation, and is compatible with gauge invariance.

Aside from the SUSY-conserving terms of the Lagrangian, there are also the soft SUSY-breaking terms which are added by hand to explicitly break the symmetry. These terms, Eqs. (1.10-1.12), correspond to the mass terms for the gauginos, sfermions, Higgs bosons and the trilinear couplings between the sfermions and Higgs bosons [43].

$$-\mathscr{L}_{gaugino} = \frac{1}{2}\left[M_1\tilde{B}\tilde{B} + M_2\sum_{a=1}^{3}\tilde{W}^a\tilde{W}_a + M_3\sum_{a=1}^{8}\tilde{G}^a\tilde{G}_a + h.c.\right] \tag{1.9}$$

$$-\mathscr{L}_{sfermions} = \sum_{i=gen} m^2_{\tilde{Q}_i}\tilde{Q}_i^\dagger\tilde{Q}_i + m^2_{\tilde{L}_i}\tilde{L}_i^\dagger\tilde{L}_i + m^2_{\tilde{u}_i}\left|\tilde{u}_{Ri}\right|^2 + m^2_{\tilde{d}_i}\left|\tilde{d}_{Ri}\right|^2 + m^2_{\tilde{\ell}_i}\left|\tilde{\ell}_{Ri}\right|^2 \tag{1.10}$$

$$-\mathscr{L}_{Higgs} = m^2_{H_u}H_u^\dagger H_u + m^2_{H_d}H_d^\dagger H_d + B\mu\left(H_u\cdot H_d + h.c.\right) \tag{1.11}$$

$$-\mathscr{L}_{trilinear} = \sum_{i,j=gen}\left[A^u_{ij}Y^u_{ij}\tilde{u}^*_{Ri}H_u\cdot\tilde{Q}_j + A^d_{ij}Y^d_{ij}\tilde{d}^*_{Ri}H_d\dot{Q}_j + A^\ell_i jY^\ell_{ij}\tilde{\ell}^*_{Ri}H_d\cdot\tilde{L}_j + h.c.\right] \tag{1.12}$$

Summing together Eqs. (1.10-1.12) gives the soft SUSY-breaking scalar potential, as seen in Eq. (1.13).

$$V_{soft} = -\mathscr{L}_{sfermions} - \mathscr{L}_{Higgs} - \mathscr{L}_{trilinear} \tag{1.13}$$

Combining the SUSY-breaking scalar potential with the supersymmetric Lagrangian and gaugino mass terms gives the unconstrained MSSM (MSSM-124), which adds 105 unknown parameters in addition to those already in the SM. There are inherent phenomenological problems with the unconstrained MSSM, in that a generic parameter set will show no individual lepton number conservation, will allow flavor-changing neutral currents, and will introduce new sources of CP violation beyond that allowed by experimental bounds [8]. Constraints can be added to MSSM-124

to make a phenomenologically viable parameter space, with one example being the phenomenological MSSM (pMSSM), which makes three assumptions thereby reducing the parameter space to 22 input parameters [43]:

- All soft SUSY-breaking parameters are real, which means no new sources of CP-violation.
- mass matrices and trilinear couplings for the sfermions are all diagonal, which implies that there are no flavor-changing neutral currents at tree level.
- soft SUSY breaking masses and trilinear couplings of first and second generation sfermions are the same at low energy.

Different constrained MSSM models, in addition to the pMSSM, will be based on assumptions like these in order to reduce the parameter space and make a more easily testable and phenomenologically viable model. The number of assumptions will determine the number of parameters, where more constrained models with a larger number of assumptions will have a smaller number of parameters that can be studied. Other examples of constrained MSSM models include anomaly mediated SUSY-breaking (AMSB) and gauge mediated SUSY-breaking (GMSB). Both of these types of models describe how SUSY-breaking occurs in hidden sectors and are then transmitted to the MSSM fields through either the super-Weyl anomaly in AMSB models or SM gauge interactions in GMSB models.

An important component of these constrained models will be the composition of the electroweakinos, which will determine what mass differences can be expected between the SM and SUSY particles. Additionally these will determine whether a given decay mode will be suppressed, affecting the production cross sections. For experimental searches simplistic minimal models are used such that the decay mode being looked at for a given sparticle will have a 100% branching ratio and in addition, the minimal number of sparticles are considered in order to further simplify the model.

1.5.4 Mass Couplings and Decays

In talking about the mass couplings and decays of the sparticles, a focus is given to the stop, and electroweakino states, as the main production modes studied correspond to pair production of the stops which subsequently decay to charginos or neutralinos.

Each sparticle will have a mass matrix that describes the mass eigenstates and mixing of the relevant sparticles. These also give the unknown parameters that these masses depend on. The electroweakinos will have two separate matrices, one for the chargino and one for the neutralino. The charginos will be described at tree level by a 2×2 complex matrix given by Eq. (1.14).

$$\begin{bmatrix} M_2 & \sqrt{2}M_W s_\beta \\ \sqrt{2}M_W c_\beta & \mu \end{bmatrix} \tag{1.14}$$

This describes the possible chargino states corresponding to the mixtures of the charged winos $\left(\tilde{W}^{\pm}\right)$ and charged higgsinos $\left(\tilde{H}^{-}, \tilde{H}^{+}\right)$, where M_2 is the mass parameter of the wino, from the gaugino term of the soft susy breaking Lagrangian, Eq. (1.9), and μ is the higgsino mass parameter found in the superpotential, Eq. (1.8). The remaining terms $s_\beta \equiv \sin\beta$ and $c_\beta \equiv \cos\beta$ are related to the ratio of the vacuum expectation values (vevs), $\tan\beta$ from the Higgs doublet, while M_W is the W boson mass term. The matrix can be diagonalized using singular value decomposition, $U\mathscr{M}V^{-1}$, with U being the rotation matrix for the negatively charged states and V the positively charged states [43]. The diagonalized matrix gives two mass eigenstates $m_{\tilde{C}_1}$ and $m_{\tilde{C}_2}$. These physical states are generally denoted by $\tilde{\chi}_1^{\pm}$ and $\tilde{\chi}_2^{\pm}$, which are a linear combination of wino and higgsino states, with the ultimate composition depending on the matrix elements of U and V [8]. The masses are ordered such that $M_{\tilde{\chi}_1^{\pm}} < M_{\tilde{\chi}_2^{\pm}}$.

The neutralino will have a 4×4 mass matrix at tree level, due to the four possible neutral states, given by Eq. (1.15).

$$\begin{bmatrix} M_1 & 0 & -M_Z s_W c_\beta & M_Z s_W s_\beta \\ 0 & M_2 & M_Z c_W c_\beta & -M_Z c_W s_\beta \\ -M_Z s_W c_\beta & M_Z c_W c_\beta & 0 & -\mu \\ M_Z s_W s_\beta & -M_Z c_W s_\beta & -\mu & 0 \end{bmatrix} \tag{1.15}$$

This matrix will describe the mixings of the neutral higgsinos, $\left(\tilde{H}_u^0, \tilde{H}_d^0\right)$, and the neutral gauginos, the bino and wino $\left(\tilde{B}, \tilde{W}^0\right)$. The terms not present in the chargino mass matrix correspond to the bino mass term, M_1 from Eq. (1.9), the Z boson mass, M_Z, and the weak mixing angle, $s_W \equiv \sin\theta_W$ and $c_W \equiv \cos\theta_W$. The matrix can also be diagonalized, but in this case by a single unitary matrix, N, such that $NM_{\tilde{N}}N^{-1} = diag(M_{\tilde{N}_1}, M_{\tilde{N}_2}, M_{\tilde{N}_3}, M_{\tilde{N}_4})$, where the diagonal terms are the masses of the four neutralino states. Like the charginos, the neutralinos are usually denoted by $\tilde{\chi}_1^0$, $\tilde{\chi}_2^0$, $\tilde{\chi}_3^0$, and $\tilde{\chi}_4^0$, with the masses ordered in the same way: $M_1 < M_2 < M_3 < M_4$. The four neutralinos are linear combinations of the neutral higgsino, wino, and bino states, with the composition determined by the unitary matrix N [8].

Considering only the third generation of squarks, the tree level mass matrix for the stop will be a $2x2$ squared mass matrix, seen in Eq. (1.16), which describes the mixing between the left and right handed stops that come from the SUSY Lagrangian in Eqs. (1.8, 1.10).

$$\begin{bmatrix} m_t^2 + m_{\tilde{Q}}^2 + L_t & m_t X_t^* \\ m_t X_q & M_{\tilde{u}}^2 + m_t^2 + R_t \end{bmatrix} \tag{1.16}$$

,

$$X_q \equiv A_q - \mu^* \cot\beta \tag{1.17}$$

The diagonal mass terms correspond to the soft SUSY breaking mass terms, $M_{\tilde{Q}}^2$ and $M_{\tilde{u}}^2$, from Eq. (1.10), the top quark mass, m_t, and then the left and right handed electroweak correction terms L_t and R_t [8]. The off diagonal terms depend on the ratio of the Higgs doublet vevs, the higgsino mass parameter, μ, and the A parameters from the trilinear coupling terms of the soft

SUSY breaking Lagrangian in Eq. (1.12). Using the rotation matrices of the angle, θ_t, the mass matrix can be diagonalized such that it turns the left and right handed eigenstates into the mass eigenstates, $\tilde{t}_1$ and $\tilde{t}_2$, corresponding to the physical stops. As with the gauginos, the masses of the two stops is ordered as $1 < 2$.

The gauginos and stops will have a similar decay pattern to their standard model counterparts, with the exception that all of the decay chains will end with the LSP, which is generally assumed to be the lightest neutralino, $\tilde{\chi}_1^0$. Following the mass matrix of the $\tilde{\chi}_i^{\pm}$'s and $\tilde{\chi}_i^0$'s, each given state will be some admixture of the bino, wino, or higgsino. With these mixtures they will have weak interaction couplings to sfermion/fermion pairs provided that they are light enough and in addition can decay to any lighter gaugino plus electroweak gauge boson or Higgs [46]. The two body decays of the charginos and neutralinos will then be given by Eqs. (1.18, 1.19).

$$\tilde{\chi}_i^0 \to Z\tilde{\chi}_j^0, W\tilde{\chi}_i^{\pm}, h^0\tilde{\chi}_j^0, \ell\tilde{\ell}, \nu\tilde{\nu} \tag{1.18}$$

$$\tilde{\chi}_i^{\pm} \to W\tilde{\chi}_j^0, Z\tilde{\chi}_1^{\pm}, h^0\tilde{\chi}_1^{\pm}, \ell\tilde{\nu}, \nu\tilde{\ell} \tag{1.19}$$

In the case that the two body decay modes are not kinematically feasible there are also the three-body decays to two fermions and a neutralino or chargino given in Eq. (1.20).

$$\tilde{\chi}_i^0 \to ff\tilde{\chi}_j^0, \tilde{\chi}_i^0 \to ff'\tilde{\chi}_j^{\pm}, \tilde{\chi}_i^{\pm} \to ff'\tilde{\chi}_j^0, \tilde{\chi}_2^{\pm} \to ff\tilde{\chi}_1^{\pm} \tag{1.20}$$

The stop will usually decay to a top plus gluino, $\tilde{t} \to t\tilde{g}$, with the other decays being a top plus a neutralino, $\tilde{t} \to t\tilde{\chi}_i^0$ and a b plus a chargino, $\tilde{t} \to b\tilde{\chi}_i^+$. Depending on the composition of the charginos and neutralinos, the decay to the LSP will be favored, especially for a bino like $\tilde{\chi}_1^0$ in the case of right handed squarks, while the left handed quarks may prefer decaying into the heavier charginos and neutralinos due to larger wino couplings. Due to the large Yukawa couplings, the stops will also favor decaying into higgsino like charginos and neutralinos, which will not be the case for other squarks. Depending on the allowed kinematic boundary, it may not be possible for the lighter top squark to decay to a top plus gluino or LSP, narrowing the possible decays to

charginos, $\tilde{t}_1 \to b\tilde{\chi}_1^+$; three body decays, $\tilde{t}_1 \to bW\tilde{\chi}_1^0$; flavor suppressed decay to charm quarks, $\tilde{t}_1 \to c\tilde{\chi}_1^0$; or four body decays, $\tilde{t}_1 \to bff'\tilde{\chi}_1^0$ [46]. For the cases where the stop decays to a gluino or heavier chargino or neutralino then these will invariably decay to the LSP through cascade decays.

1.5.5 Experimental Searches

There are two ways in which experiments can place constraints on SUSY models. There are the indirect constraints, which correspond to searching for or measuring physics that is not directly related to SUSY, but still places constraints on the SUSY parameter space. Data taken for the purpose of measuring SM electroweak observables, or studying dark matter are such examples. There are also the direct constraints, which simply correspond to searches for sparticle production. For this overview of experimental searches, a focus will be placed on the direct searches performed at the various colliders especially those done at the LHC by the CMS and ATLAS experiments.

For direct SUSY searches the most important component is the model that is used to interpret results. It is not feasible to perform searches on the full MSSM parameter space, as the number of free parameters would be too large. Instead searches are performed using constrained or simplified models that reduce the number of free parameters being explored. Constrained models, as mentioned in section 1.5.3, assume a SUSY breaking mechanism and have additional constraints to reduce the parameter space. Their draw is a large selection of experimental signatures that can be observed, with the downside that they do not cover all possible kinematic signatures or mass relations [47]. For such models though, measurements are generally made using gluinos, and first and second generation squark production, which have already been significantly constrained by past SUSY searches at the LHC. Alternatives to the constrained models are the pMSSM and other simplified models. The pMSSM takes the indirect constraints made using SM experiments in order to reduce the parameter set of the full MSSM down to on order 19 free parameters, which will allow for a large selection of experimental signatures, while not being as heavily constrained as some of the constrained SUSY models. The simplified models will only look at specific sparticle productions and decay modes, which allows for a free variation of sparticle masses for a more

focused study of the production in question. For the SUSY searches at the LHC, simplified models have been the main set of models used for the interpretation of results, as they are more convenient when looking at specific sparticle topologies, which analyses are generally built around. An important distinction for experimental searches of SUSY particles is that two separate masses must be jointly considered: the mass of the SUSY particle created in the decay, and the mass of the LSP. When exclusions are determined, they are determined in this 2-dimensional plane of masses, which means that the mass splitting between these two particles is as equally important in a search as the mass of the particles themselves. Cases where the mass splitting of the two particles is small are called compressed mass scenarios, and are important corridors of study due to the possible low momentum signatures of the decays.

Searches have been performed that have placed exclusion limits on various sparticles, with such limits depending on the model used to interpret results. All of the results mentioned here assume that R-parity is conserved. First, there are the limits placed on gluino production, which will depend on what is assumed for the LSP and intermediate decay sparticle masses. Current limits exclude gluinos with masses up to 2 TeV, assuming the LSP is a neutralino with a mass on the order of 1.0 TeV [48, 49, 50, 51]. The searches assume pair production of gluinos which decay to a quark pair and the lightest neutralino, or, for the more recent one, pair production of gluinos which decay to a quark pair and boosted Z bosons, mediated by a massive neutralino. For the earlier searches there are also the possibility of an intermediate decay with the quark pair decaying to an off shell lightest chargino, second lightest neutralino, or stop, producing gauge bosons and the LSP. Example Feynman diagrams are given in Figure 1.3

For chargino decays to neutralino LSPs, the limits will depend on the mass difference between the chargino and neutralino. Starting with large to intermediate mass splittings, masses of up to 0.6 TeV are excluded, assuming a massless LSP where pair produced charginos decay to the LSP with gauge bosons or sleptons as intermediaries [52]. There are also limits on the lightest chargino, and second lightest neutralino of up to 345 GeV with mass splittings corresponding to small values of the LSP [53, 54]. This assumed production of a chargino, neutralino pair decaying to the LSP with

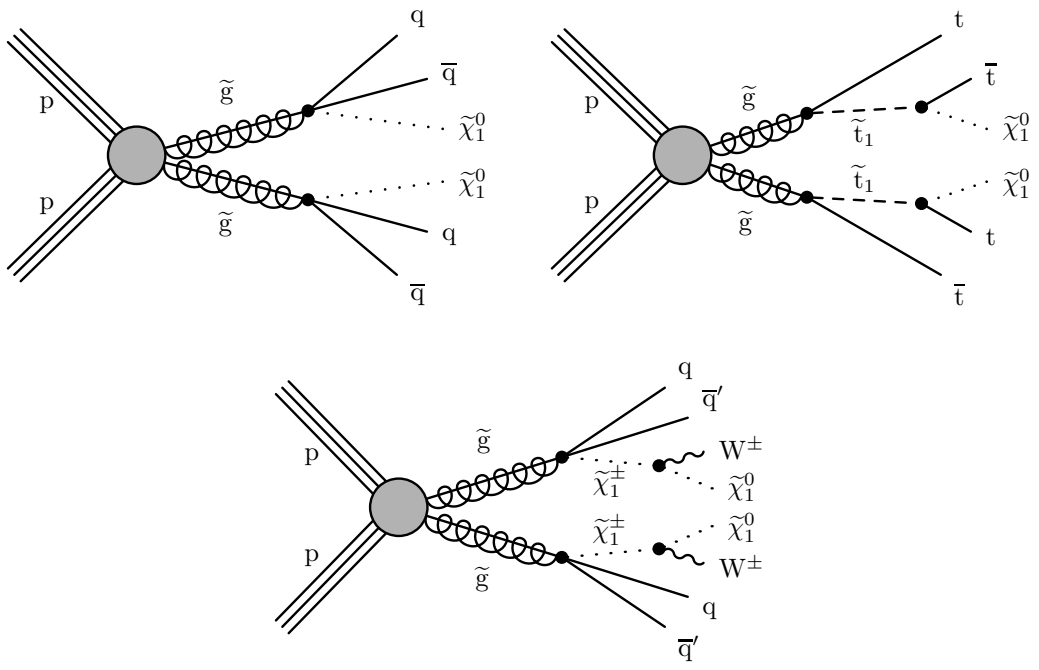


Figure 1.3: Example Feynman diagrams for gluino pair production

gauge bosons as intermediaries. For the more compressed mass scenarios the CMS experiment has placed mass limits of up to 112, 215, and 290 GeV on the lightest chargino, with mass splittings between the LSP of 1, 30, and 50 GeV, respectively [55, 56]. The ATLAS experiment has placed mass limits on the chargino of up to 240 GeV for a mass splitting of 7 GeV, down to a mass splitting of 1.5 GeV at the mass limit bounded by the LEP experiment of 92.4 GeV [57]. Feynman diagrams for chargino or neutralino production are given in Figure 1.4.

Lastly, there are the limits placed on squark production, namely the stop. For large mass splittings the stop mass has been excluded up to 1.2 TeV for close to massless LSPs, and up to 1 TeV for LSP masses up to 0.6 TeV for a pair of stops decaying to a pair of third generation quarks, with the possibility of off shell chargino intermediaries decaying to a gauge boson and the LSP [58, 48, 49]. In another search scenario, a pair of stops decay to pairs of b quarks and taus mediated by off shell charginos or staus with the LSP in the final state. This search excludes stop masses up to 1.1 TeV for LSP masses ranging from nearly massless up to about 0.4 TeV [59]. A more recent search that has been accepted for publication sets limits on stop pair production in a

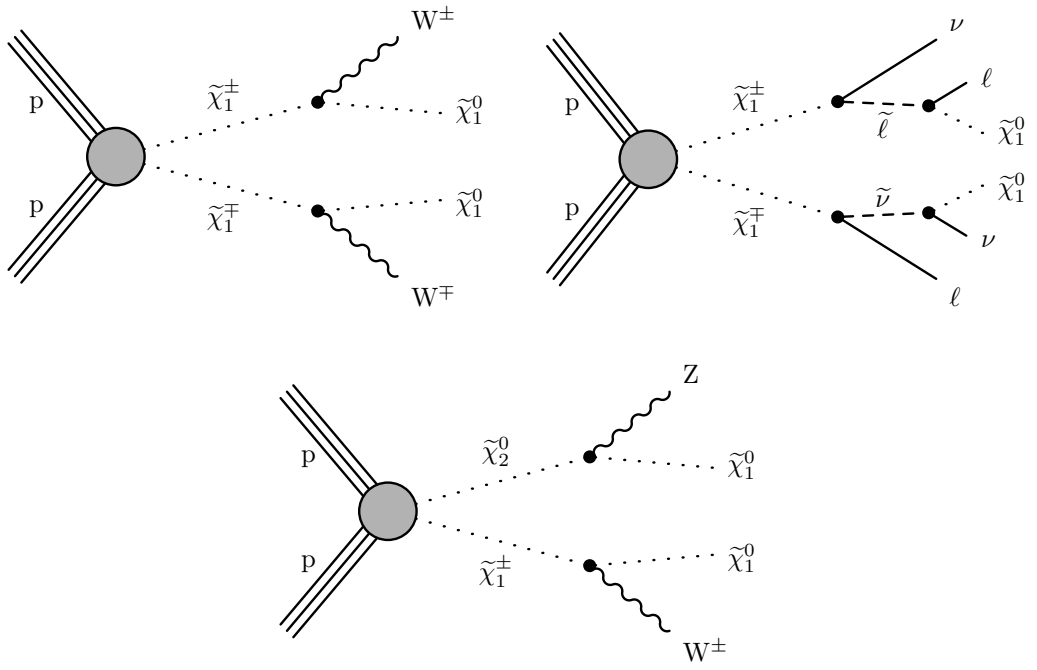


Figure 1.4: Example feynman diagrams for chargino/neutralino pair production

di-lepton final state with b jets and missing transverse momentum. Exclusions are made on the masses of the stop quark for different simplified model scenarios. For stops decaying to top quarks and the LSP, stop masses up to 925 GeV with neutralino masses up to 450 GeV are excluded. For models where the stop decays to a b quark, W boson and neutralino, and mediated by charginos, stop masses up to 850 GeV with neutralino masses up to 420 GeV are excluded. Finally, for models where the stop decays to b quarks and leptons, with the possibility of mediation by charginos and sleptons, stop masses up to 1.4 TeV with neutralino masses up to 900 GeV are excluded [60]. For the more compressed mass scenarios the stop mass has been excluded up to 450 GeV for a mass splitting of 40 GeV with the LSP, for pair production of stops decaying to a pair of b quarks, gauge bosons, and the LSP, mediated by off shell charginos [61]. A second search has excluded stop masses between 420 and 560 GeV for a mass splitting ranging from 10 to 80 GeV between the LSP, for pair production of stops decaying to a pair of b quarks, two pairs of fermions and the LSP, with the possibility of being mediated by off shell charginos [62].

This analysis will focus on the pair production of stops in compressed decay scenarios. The

possible compressed mass scenarios can fall under two main categories, with the smallest splittings being the most difficult to study. The first category corresponds to a mass region between the W and top corridors, where $m_{\tilde{t}} - m_{\tilde{\chi}_1^0} = m_W$ (m_t) in the W (top) corridor. These boundaries define the phase space in which certain particles in a decay transition between being on-shell and off-shell. This first category would then have the possibility of on-shell W's, but all of the tops in any given decay chain would be forced off-shell. The second category corresponds to splittings below m_W, where the W's would also be forced off shell. The regions bounded by these corridors, especially below the W corridor, are not as well studied due to energy constraints on detector acceptance of any soft decay products, in addition to the abundance of high energy decay products expected from colliders with large center of mass energies that would overshadow softer signatures. In current analyses this region is getting more attention, where difficulties in studying the soft decay products are mitigated by requiring large amounts of initial state radiation, giving the LSP a boost in momentum, and allowing for events with large values of missing transverse energy, and softer decay products. Figure 1.5 gives a summary of stop searches from CMS which shows exclusions for the compressed regions. For the regions below the W corridor and in between that and the top corridor, current limits do not place exclusions beyond a stop mass of 600 GeV, with mass splittings ranging from below 10 GeV up to 100-200 GeV. For stop masses above 500 GeV, there are then various sections of this parameter space where certain splittings are not yet covered, allowing for stop searches with compressed mass splittings ranging from 5 GeV, up to (and slightly above) the top corridor. Related to this, there are CMS analyses in the publication process with results that are expected to fill in some of the described phase space.

The main stop decay modes that are under consideration for this Dissertation are a pair of stops decaying to a pair of third (or second) generation quarks and either a pair of gauge bosons mediated by charginos, or two pairs of fermions which can also be mediated by charginos. The masses being focused on have splittings of ≤ 80 GeV, which is below the W corridor. The soft b-tagging, to be described in Chapter 4, helps with these compressed splittings due to its usefulness in tagging b-quarks with p_T less than 20 GeV. While the focus of the analysis is on these smaller

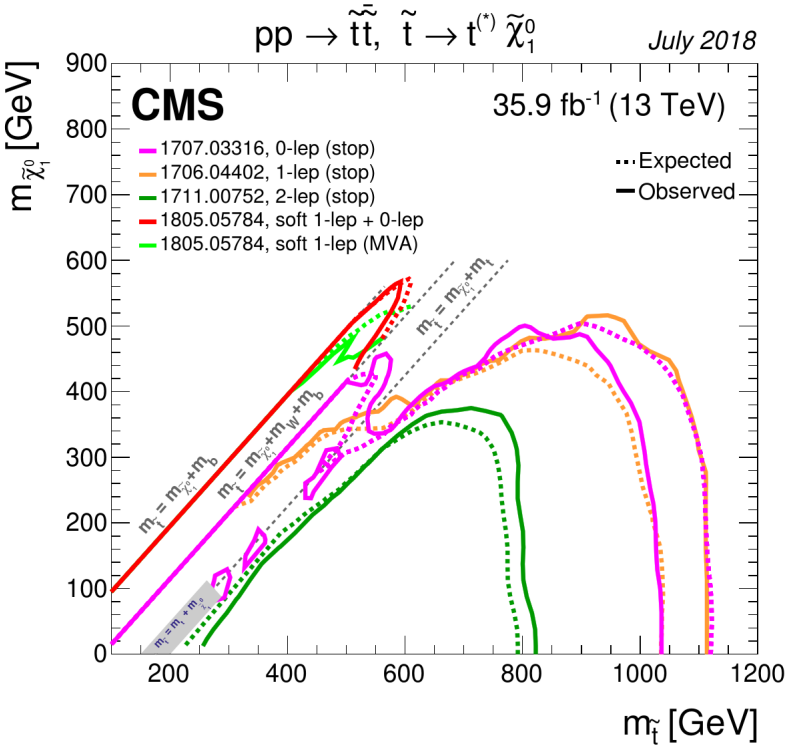


Figure 1.5: Public limits from CMS of SUSY searches for stops with on or off shell W's or top quarks. The dotted lines detail the regions bounded by the top and W corridors [2].

mass splittings, other splittings of up to 200 GeV are also considered. These regions correspond to a section of parameter space where there are not extensive exclusions on stop mass and LSP pairs for stop masses above 500 GeV.

Chapter 2

The CMS Experiment and the Large Hadron Collider

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [3] is a two ring superconducting hadron accelerator and collider with a design center of mass energy of 14 TeV and design luminosity of $10^{34}\ cm^{-2}s^{-1}$. It straddles the border of France and Switzerland and is situated in the tunnel that used to house the Large Electron-Positron Collider [63]. It is 26.7 km in diameter at a depth ranging from 45 m to 170 m below the surface and there are two rings, which hold counter-rotating beams that interact at four points along the tunnel. At these four interactions points are situated detectors which study the collisions of the beams: LHCb, ALICE, ATLAS, and CMS. LHCb and ALICE are detectors dedicated to studying heavy flavor physics and ion collisions, respectively. ATLAS and CMS are general purpose detectors used to study both proton and ion collisions. There have been two operational run periods for proton-proton collisions at the LHC, colloquially referred to as Run I and Run II. The first run period, Run I, was from 2009-2013 and used a center of mass energy of 7 TeV. The second run period, Run II, was from 2015-2018, and operated at a center of mass energy of 13 TeV. In the following section the operations of the LHC will be briefly described. Following that, an in depth view of the CMS experiment will be given, with the most detail being given to the pixel detector, and its recent upgrade.

2.1.1 The LHC Machine

There are two rings in the LHC which are used to contain counter-rotating beams. Along the rings there are eight arcs and straight sections as seen in Figure 2.1. Each of the straight sections serve as

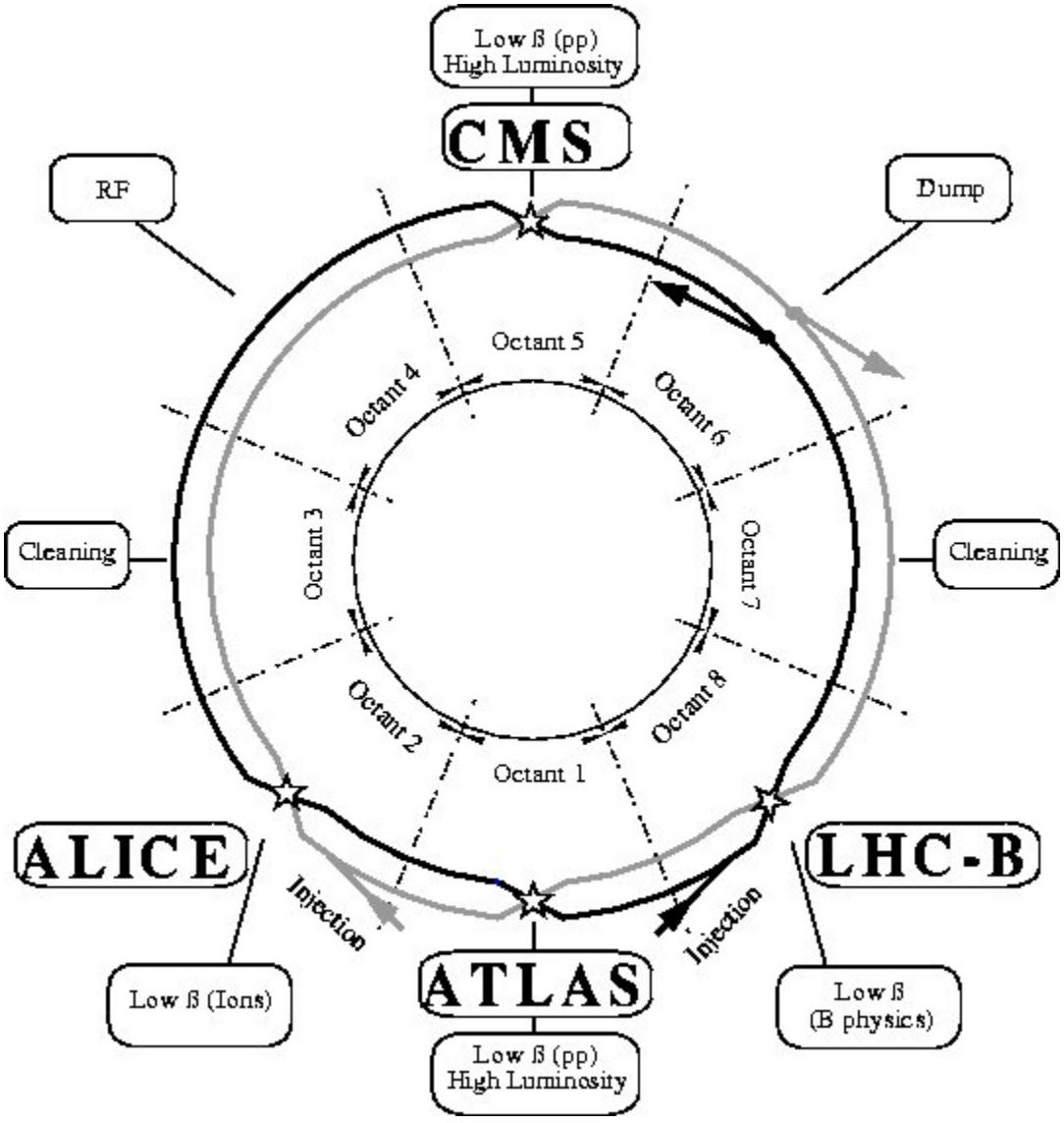


Figure 2.1: A Schematic layout of the LHC rings [3]

either an experimental or utility insertion. The sections are labeled following their octant number, going from Point 1-8. At four of the eight points there are beam crossings where the previously mentioned detectors are located (Points 1, 2, 5 and 8). The remaining points are used for beam upkeep. At Points 3 and 8 are collimation systems for cleaning the beams. At Point 6 is a beam dump insertion and Point 4 has RF cavities for accelerating the beams.

Along the rings are superconducting magnets which are used to guide the beams. There are two main types of magnets situated in the rings, those being the twin-bore dipole and the quadrupole

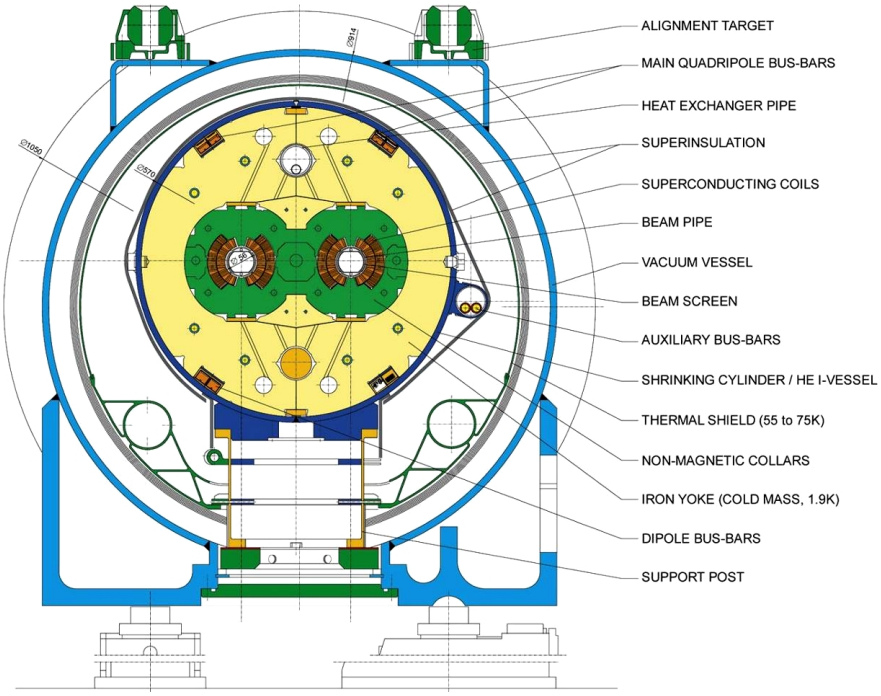


Figure 2.2: A cross section of the LHC dipole magnets [3]

magnets. They have a unique design due to space limitations, where the beams share the same cold mass and cryostat that are used to cool down the magnets to a superconducting state. The cold mass operates at 1.9 K, and is cooled by superfluid helium. The combination of magnet, cold mass and cryostat in the dipole and quadrupole system creates a complicated magnet structure, such that the beam corridors are coupled both magnetically and mechanically [3]. A cross section of the dipole is given in Figure 2.2. There are 1232 main dipoles situated around the arcs of the LHC rings, which are used for guiding the beams around the arcs and there are 392 main quadrupoles which are used for focusing the beams as they move around the ring.

With a design center of mass energy at 14 TeV, and an actual center of mass energy set at 13 TeV for the Run II data taking period, there is a need to accelerate the beams up to this energy through a fairly involved process. Each beam is made up of 'packets' of protons which are accelerated in groups (bunches) and then fed into the LHC ring up to the designated spacing of the bunches, which was 25ns for Run II. Each bunch starts with a hydrogen tank as a source, which feeds 40 MeV protons into the Linac2 (Linear accelerator 2). These protons are injected into the

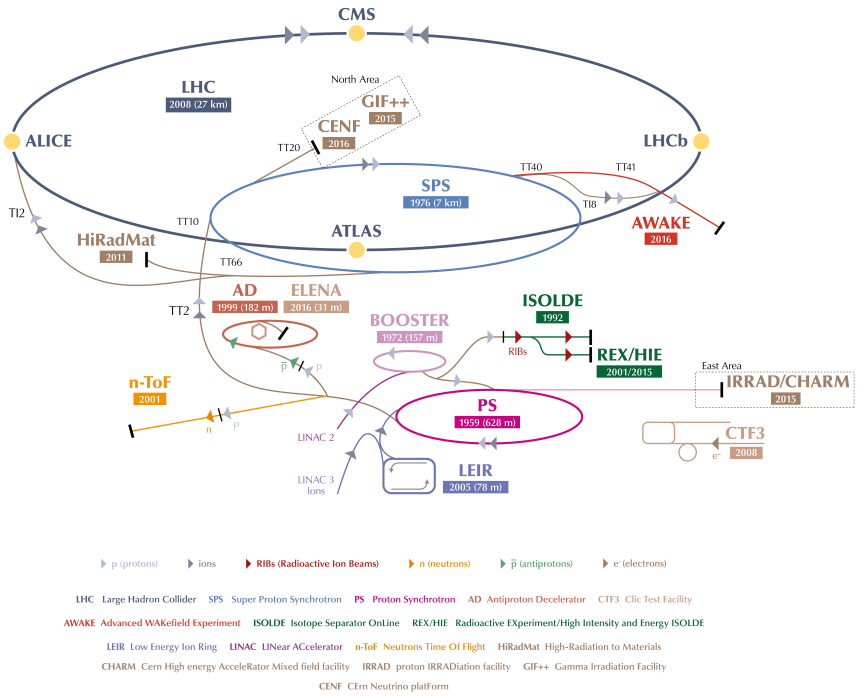


Figure 2.3: A schematic of the full CERN accelerator complex [4]

Proton Synchrotron Booster (PSB) and accelerated up to 1.4 GeV, then injected into the Proton Synchrotron (PS) and accelerated to an energy of 25 GeV. The last step is an injection into the Super Proton Synchrotron (SPS) to accelerate the bunches to 450 GeV, where they are finally injected into the LHC. After the necessary number of bunches are injected into the LHC rings, the counter-rotating beams are accelerated to the final energy of 6.5 TeV per beam, through the use of an electric field produced by a 400 MHz (RF) superconducting cavity system. The bunches in these beams then collide every 25 ns at the 4 interaction points, producing the interactions that are then detected at each of the experiments. The full setup of the accelerator complex can be seen in Figure 2.3.

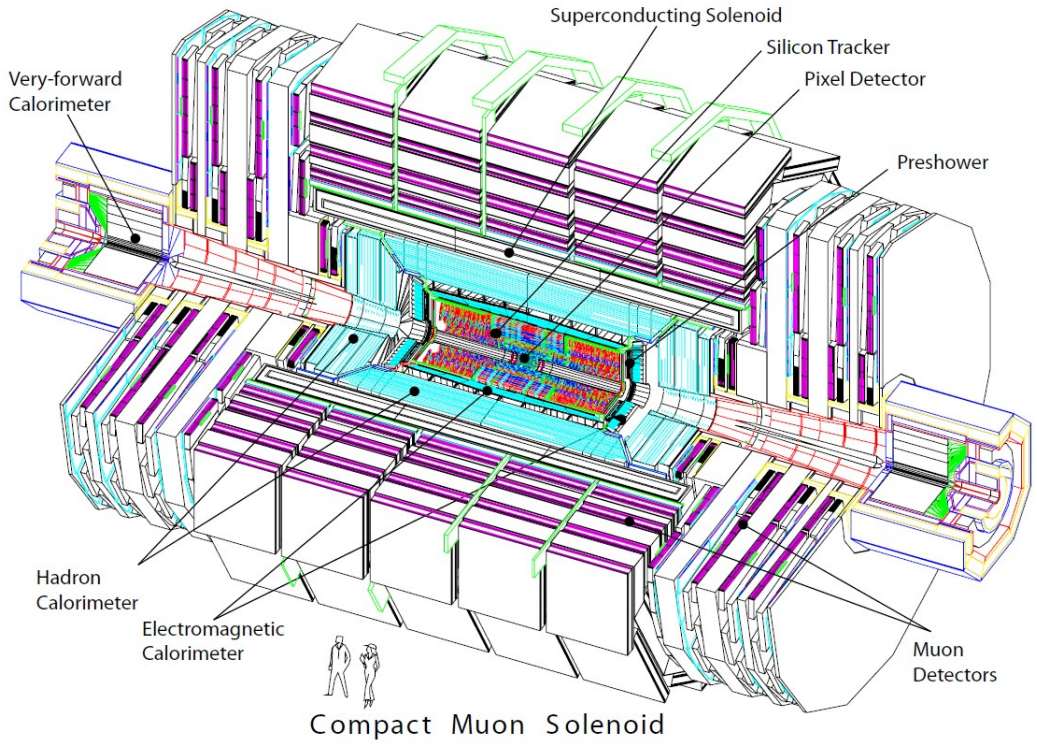


Figure 2.4: A schematic of the CMS detector, showing the location of each of the sub-detectors [5].

2.2 The CMS Experiment

The Compact Muon Solenoid (CMS) is a multi-purpose detector located at the LHC situated at Point 5 [5]. The primary motivation for the detector was to clear up the nature of electroweak symmetry breaking, specifically in the study of the Higgs mechanism, and to study interactions at the TeV scale, where new physics is presumed to occur. To meet these goals, there were four requirements for the CMS detector: good muon identification and momentum resolution over a wide range of energies and angles, good charged-particle momentum resolution and reconstruction efficiency, good electromagnetic energy resolution, and good missing transverse energy ($\not\!\!E_T$) and dijet mass resolution. These requirements are met by the muon system and superconducting solenoid, the silicon tracker system, the electromagnetic calorimeter (ECAL), and the hadronic calorimeter (HCAL), respectively. A schematic view of the CMS detector can be seen in Figure 2.4. The following information on the CMS detector is taken extensively from [5].

2.2.1 Superconducting Magnet

The main feature of the CMS detector, and its namesake, is the superconducting solenoid magnet. With a diameter of 6 m and a length of 12.5 m, it is made up of 4 layers of superconducting NbTi coils, separated into 5 modules. The entire cold mass system that the coils are a part of is situated inside of a 10,000 ton return yoke, consisting of 5 wheels and 2 endcaps. The magnet is designed to create a 4 T field, with an operating field of 3.8 T and a stored energy of 2.6 GJ. The tracker system, ECAL, and HCAL are located in the bore of the magnet coil, while the muon system is interlaced throughout the return yoke outside of the magnet. The main purpose of the magnet is to measure the momentum of high-energy charged particles, namely that of muons, through the bending power provided by the strength of the magnetic field.

2.2.2 Muon System

The detection of muons is of great importance to the CMS experiment, and thus the design of the muon system was a central theme in the design of the overall detector. There are 3 functions assigned to the muon system which are muon identification, momentum measurement, and muon triggering (not to be confused with the event level electronic triggering system). The momentum measurement is enabled by the strength of the magnetic field, with the muon system measuring the curvature of the path of the muons, in addition to the separate curvature that is measured in the inner tracking system. The triggering is enabled by the steel return yoke, which acts as a hadron absorber, helping to prevent particles other than muons from passing through the muon system. The muon system is made of three types of gaseous detectors, which are used to track the paths of muons through the detector. These are designed to reconstruct the momentum over the entire kinematic range of the LHC, and with a pseudorapidity coverage of up to $|\eta| < 2.4$ [5]. There is a cylindrical barrel section, consisting of the drift tube (DT) chambers, and 2 endcaps consisting of cathode strip chambers (CSCs). Lastly there are the resistive plate chambers (RPC), which have a faster response time than the DT and CSCs and are used as a dedicated trigger system in both the barrel and endcaps.

The DTs are made of standard rectangular drift cells and cover a pseudorapidity up to $|\eta| < 1.2$. The choice of drift cells in the barrel region is due to a low neutron-induced background and muon rate, and a uniform magnetic field contained within the return yoke. The DTs are placed among the flux return plates, and are organized into 4 stations of 12 chambers in the first 3 stations, and 8 chambers in the fourth. These measure the muon coordinates in the $r-\phi$ bending plane and in the z direction, where the fourth station does not have chambers corresponding to measurements in the z direction.

CSCs are used in the end-cap regions of CMS, due to high muon and background rates along with a non-uniform magnetic field. The CSCs cover the pseudorapidity $0.9 < |\eta| < 2.4$. There are 4 stations on either end caps, consisting of chambers placed perpendicular to the beam line. Each chamber will have segments of cathode strips and anode wires. The strips are placed pointing radially outward from the beam line to provide measurements in the $r-\phi$ bending plane, and the wires are perpendicular to the strips to provide measurements of η and beam-crossing time.

The RPC was introduced to the muon system as a complementary and dedicated trigger system. Due to the uncertainties in background rates and proper measurements of the beam-crossing time, the RPC was given the purpose of supplementing the measurements from the DTs and CSCs with a more precise time resolution and correspondingly coarser position resolution. In addition they are used to correctly make tracks from multiple hits in any given chamber. There are 6 layers embedded in the barrel system, and 3 layers in each of the endcaps with the layers consisting of double-gap chambers operated in avalanche mode.

2.2.3 Hadronic Calorimeter

The hadronic calorimeter has the purpose of measuring hadron jets, and indirectly, neutrinos or exotic particles through missing transverse energy. It is also used to supplement energy measurements from the ECAL. The HCAL system consists of 4 separate calorimeter systems: the barrel (HB), outer (HO), endcap (HE), and forward (HF). These systems are made of alternating layers of brass or steel absorbers and plastic scintillators. A cross section of HCAL can be seen in Figure

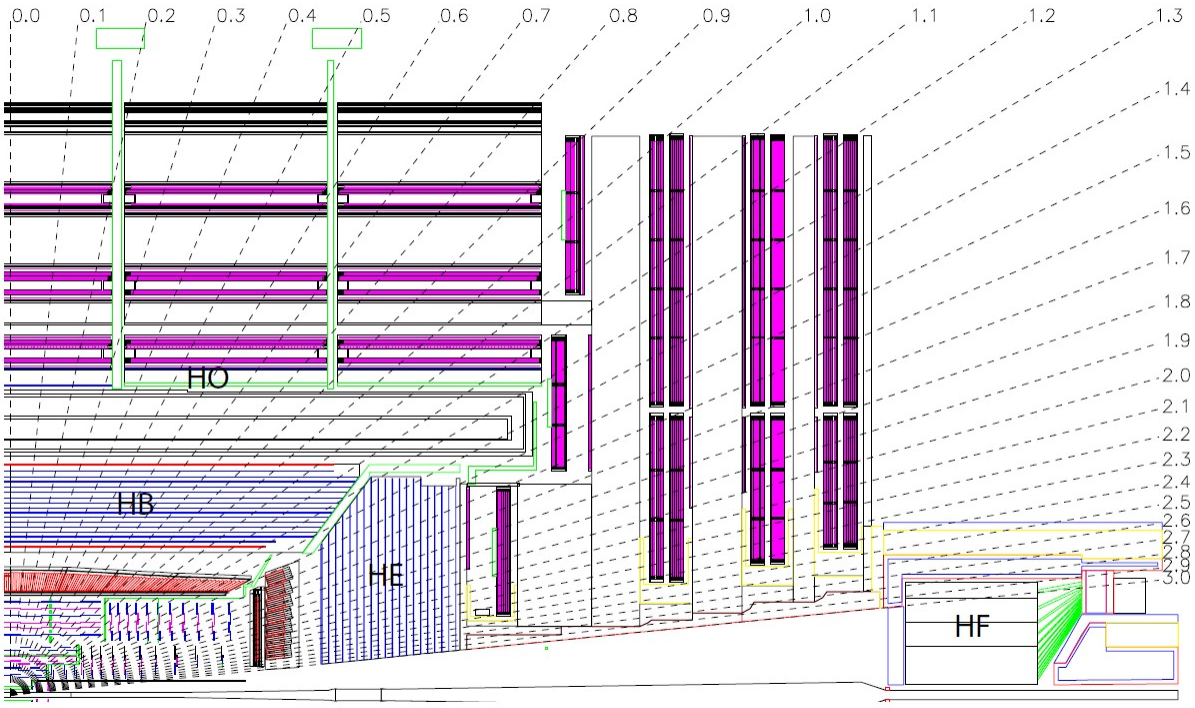


Figure 2.5: A longitudinal view of CMS, showing the location of each sub-component of HCAL [5].

2.5, showing the location of each of the calorimeters.

The HB is made of 36 wedges arranged azimuthally around the detector covering a range of $|\eta| < 1.3$. It is located inside the bore of the solenoid, bordering the inner wall. The plates in each wedge are placed parallel to the beam line, with alternating plates of brass and scintillating tiles placed in a staggered geometry to minimize dead material, where the brass absorbers have a total thickness of about 6 interaction lengths. The front and back plates are made of steel for structural support. The HO is an extension of the HB, located just outside of the solenoid, and used as a tail catcher to increase the sampling depth of the barrel calorimeter. The shape of the HO matches the muon system due to its location, and is separated into 5 rings, with 12 sections in ϕ for each ring. The central ring holds two layers of scintillating tiles separated by a layer of tail catcher iron, while the 2 outer rings on either side hold a single layer of scintillating tiles. The solenoid coil is treated as an initial absorber for the HO. The addition of the HO to the barrel calorimeter system increases the total absorber width to about 12 interaction lengths at a minimum.

The HE is mounted on the iron yoke endcap of the muon system and covers a range of $1.3 < |\eta| < 3$. The absorber is made of cartridge brass due to the need for a non-magnetic material with a maximal number of interaction lengths in order to contain hadron showers in the covered pseudorapidity. The shape of the HE is designed to minimize cracks in pseudorapidity between the HB and HE. Absorber plates are bolted together in a staggered geometry to minimize dead material. The ECAL endcap and preshower detector are attached to the front of the HE, with the total length of the calorimeter system amounting to 10 interaction lengths. Trapezoidal shaped scintillators are arranged in trays which are then placed between the gaps in the absorbers, for up to 18 layers of alternating scintillator trays and absorbers.

The HF is located 11.2 m from the interaction point of CMS, extending the pseudorapidity range of HCAL down to $|\eta| = 5.2$. With its close proximity to the beam line, an inner radius of 12.5 cm from the beam line, radiation hard quartz fibers were used for the active material instead of scintillating tiles. The absorber is made of a steel structure containing grooved plates, with the quartz fibers inserted into these grooves. The fibers have lengths ranging the full depth of the absorber to 22 cm from the front of the detector in order to differentiate between electromagnetic and hadron showers. Surrounding the calorimeter is a hermetic radiation shielding made of steel, concrete and polyethylene, with a plug structure in the back of the calorimeter for additional shielding. The photomultiplier tubes and front end electronics are protected by a steel-lead-polyethylene shielding matrix.

In 2017 the HCAL was partially upgraded, where the photomultiplier tubes were replaced with Silicon Photomultipliers in the HB and HE, and multi-anode photomultiplier tubes in the HF. These new components with dual readout, in addition to upgraded readout electronics, enables better discrimination from anomalous signals due to hits on the photomultiplier tube windows. This upgrade allows for better discrimination from anomalous signals through the measurement of the hit timing and charge asymmetry between the two readouts now produced [64].

In 2018, a couple months into the run period, there were two modules who had their power supply fail, causing these modules to be unable to take data. This affected jet energy and missing

transverse energy measurements in the regions $-3.2 < \eta < -1.3$ and $-1.57 < \phi < -0.87$. Any impact on jets and missing transverse energy in this region is studied in order to determine any action that needs to be taken.

2.2.4 Electromagnetic Calorimeter

The electromagnetic calorimeter is a hermetic homogeneous detector made of lead tungstate crystals, and consists of a barrel (EB), endcap (EE) and preshower detector (ES). The choice of lead tungstate crystals provides a radiation hard detector that is fast with a fine granularity all in a compact size. The EB is placed in the bore of the solenoid magnet in between the tracker system and HB, while EE and ES are attached to the front of the HE system. The main purpose of ECAL is the measurement of photon and electron energies, with the driving motivation in the design being the detection of a di-photon event, which is one of the decay modes of the Higgs boson. A layout of ECAL is given in Figure 2.6.

The EB is made up of 61200 crystals with a granularity 360-fold in ϕ and (2×85)-fold in η and covers a pseudorapidity range up to $|\eta| < 1.479$. Signals from the crystals are read out by avalanche photodiodes. The crystals are arranged in aluminum alveolar structures called submodules, which are further arranged into modules that contain 400 to 500 crystals. Four modules then further make up a supermodule of which there are 36, each containing 1700 crystals.

The EE has 7324 crystals in each endcap and covers a pseudorapidity of $1.479 < |\eta| < 3.0$. As opposed to the EB crystals, the signals from the EE crystals are read out using vacuum phototriodes. The crystals are arranged in groups of 5x5 crystals called a supercrystal, and are held together with a carbon-fiber alveola structure. The endcaps are separated into halves, called Dees, each containing 138 supercrystals and 18 partial supercrystals arranged in a rectangular x-y grid.

The ES is a sampling detector covering the pseudorapidity range $1.653 < |\eta| < 2.6$. Its purpose is the identification of neutral pions and electrons from minimum ionizing particles along with improving position resolution for leptons. There are two layers, consisting of lead radiators which will induce electromagnetic showers from photons and electrons, and followed by silicon strip

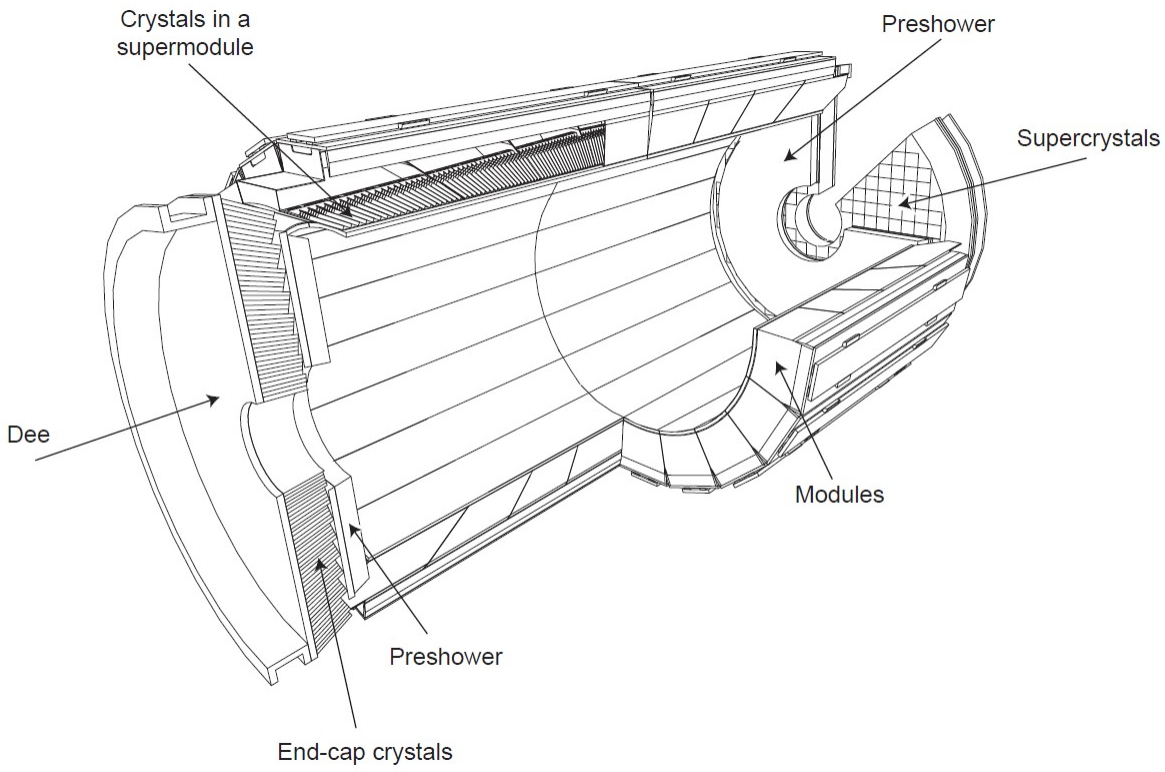


Figure 2.6: The layout of ECAL, showing how the crystal modules are arranged [5].

sensors to measure deposited energy and shower profiles. The silicon sensors are combined with front end electronics and ceramic supports to form micromodules. A set of 7-10 micromodules are connected to an electronics system motherboard to form a ladder, where about 500 of these ladders are attached to the lead radiators in an x-y configuration.

2.2.5 Tracking System

The tracking system of CMS has the purpose of measuring the trajectories of charged particles as well as the reconstruction of secondary vertices in a precise and efficient manner. The system directly surrounds the interaction point with a length and diameter of 5.8 m and 2.5 m, respectively. As it is located inside the bore of the solenoid magnet, it is immersed in a homogeneous magnetic field, which is ideal for consistent track curvature of charged particles. The design luminosity of $10^{34} cm^{-2} s^{-1}$ necessitates a robust system with both high granularity and good time resolution so that particles from any given interaction will have their trajectories and bunch crossing correctly identified; coupled with this is a need for any sensors to be radiation hard due to proximity to the interaction point. These requirements led to the choice of a silicon based detector.

The tracking system consists of 2 separate detectors: the pixel detector and the silicon strip tracker. The silicon strip tracker is made of 10 barrel layers, extending out to a radius of 1.1 m, and 12 disks in the endcap on either side of the barrel. The pixel detector initially consisted of 3 barrel layers with radii between 4.4 and 10.2 cm from the beam line. Each endcap had 2 disks. In 2017 a new pixel detector was installed as part of the Phase 1 upgrade to CMS, which increased the number of barrel layers to 4, with radii ranging between 2.9 and 16.0 cm, and increased the number of endcap disks from 2 to 3, with each disk made of an inner and outer ring. In the description of the pixel detector, the one that was initially installed will be called the phase-0 detector, while the one that was installed as part of the upgrade will be called the phase-1 detector. The phase-0 detector was used in the collection of data during the first half of Run II, during the years 2015 and 2016, while the phase-1 detector has been used for data taking in the later half of Run 2, in the years 2017 and 2018.

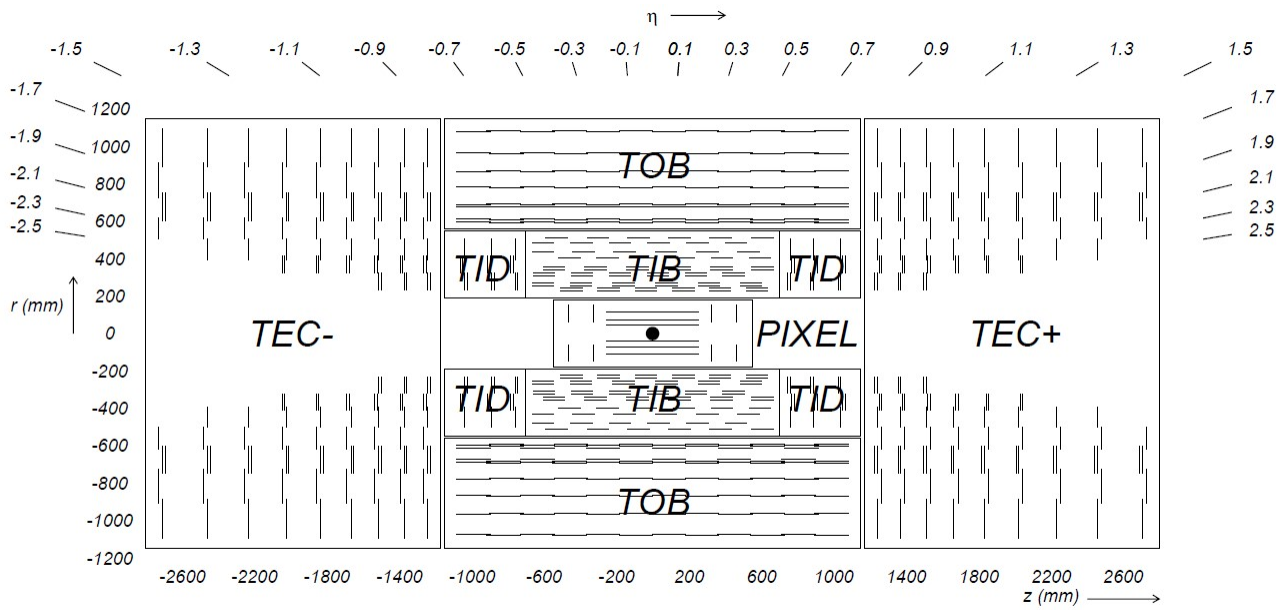


Figure 2.7: A Schematic cross section of the CMS tracker, with each line corresponding to a detector module [5].

2.2.5.1 Silicon Strip Tracker

The silicon strip tracker is made up of three subsystems, the inner, outer and endcap trackers. The inner tracker consists of two sections: the Tracker Inner Barrel (TIB) and Tracker Inner Disks (TID), with an inner and outer radius of 20 cm and 55 cm, respectively. The outer system is The Tracker Outer Barrel (TOB), directly surrounding the TIB, and extending out to 116 cm in radius, and $\pm$118 cm in z. The last subsystem is the Tracker End Caps (TEC$\pm$), which covers a radius from 22.5 cm to 113.5 cm and a z range of 124 cm to 282 cm. Each of these detectors are made up of micro-strip sensors with thicknesses of either 320 or 500 μm amounting to 9.3 million strips with 198 m^2 of active silicon area. The micro-strip sensors have strip pitches with thicknesses varying between 80 and 183 μm. A schematic of the tracking system can be found in Figure 2.7.

The TIB/TID subsystem has 3 layers in the barrel region and 3 disks on either end. The placement of the layers and disks with respect to each other can provide up to 4 measurements of a particle's trajectory in r and ϕ. The layers are made up of 320 μm thick silicon micro-strip sensors, with strip pitches of 80 μm for the first two layers, 120 μm for the last two layers of the TIB and

between 100 μm and 141 μm for the TID. The modules are placed parallel to the beam axis for the TIB and radially around the beam axis for the TID. A second set of modules are placed on the back side of the first two layers and rings to provide a measurement of z and r, respectively.

The TOB consists of 6 layers of micro-strip sensors surrounding the TIB and TID. All 6 layers have sensors that are 500 μm thick, while the first 4 have pitches of 183 μm and the remaining 2 have pitches of 122 μm. The first 2 layers, like the TIB and TID, have a second micro-strip module on the backside of the layers. These layers provide an additional 6 measurements in r and ϕ.

The TEC is placed on either end of the barrel, with each endcap made up of 9 disks of overlapping rings, with the inner disks containing up to 7 rings. The 4 inner rings contain 320 μm thick micro-strip detectors, with the remaining outer rings having 500 μm thick detectors. These rings have strips with pitches ranging from 97 μm to 184 μm. The first, second, and fifth rings of the disks, following the double lines seen in Figure 2.7, have a second micro-strip module on the back-side of the ring. The disks altogether provide up to 9 measurements in ϕ. With the full layout of the silicon strip tracker, there are at least 9 hits in any given trajectory up to a pseudorapidity of $|\eta| < 2.5$, with at least 4 of those hits being a 2-dimensional measurement in r and ϕ

2.2.5.2 Pixel Detector

The pixel detector is the innermost detector of CMS, and thus is the closest to the interaction region. Its proximity required a setup that was radiation hard, while providing the resolution necessary to cope with the high interaction rate. These considerations led to the use of n+ pixels on n-substrate, which allowed for operation in a state of partial depletion at high particle rates.

The phase-0 detector consisted of 3 barrel layers (BPix), and two end-cap disks (FPix) on either side of the barrel. It covered the pseudorapidity range of $|\eta| < 2.5$, matching the acceptance of the silicon tracking system. The detector had pixels that had dimensions of $100 \times 150 \mu m^2$ in order to facilitate good track resolution in the three directions of r, ϕ, and z. The BPix had 48 million pixels, while the FPix had 18 million, with an arrangement of the pixel sensors which allowed tracks to be made up of 3 points over most of the full η range.

The modules which make up the detector are hybrid pixel detector modules. They are a hybrid in the sense that the pixel sensor and readout electronics are produced separately and then bump bonded together to get the completed product. In FPix, a disk was populated with 7 sensor tiles, using either 1 column and 2 rows (1×2) read out chip (ROC) modules or 2×5 ROC modules, with each ROC corresponding to a pixel sensor. In BPix there were two different sensor geometries, with 708 2×8 ROC modules and 96 1×8 ROC modules, also known as half modules.

The pixel sensors are controlled through a system consisting of three parts; there is the data link which is used for readout of the sensors, a fast control link for sending triggers and resets, and a slow control link using an I^2C protocol to configure the modules and blades. Readout of signals from the pixel sensors starts with the read out chips (ROCs), which are custom application specific integrated circuits (ASICs) that cover the readout of 52×80 pixels. The ROCs are made to amplify and buffer the signal from the sensors, perform zero suppression in the pixels, level 1 trigger verification, sending hit and configuration information, and adjusting voltage levels, offsets, and currents.

Readout of the ROCs is controlled by the token bit manager (TBM), a chip which is located on the pixel modules. For phase-0 BPix, the TBMs controlled 8 or 16 ROCs depending on the layer, and in the case of FPix a single TBM controlled 21 or 24 ROCs, depending on the blade. The TBMs have four main functions: controlling the readout of the ROCs, writing headers and trailers to identify where hit information comes from in the modules, distributing Level 1 triggers and the clock, and encoding and sending an analog signal to the front end drivers (FEDs). The analog signals from the TBM are sent through an Analog Optical Hybrid (AOH) which produces an optical signal that is sent to the FED.

The FEDs receive data from the TBMs, digitizes them, and does some light processing, including transmitting errors, before sending them on to data acquisition (DAQ). Connected to the front end are pixel front end controllers (pFECs), which provide the clock and trigger information used by the FEDs and TBMs in the data readout.

In the phase-1 detector, this readout structure remains, with the main differences being in some

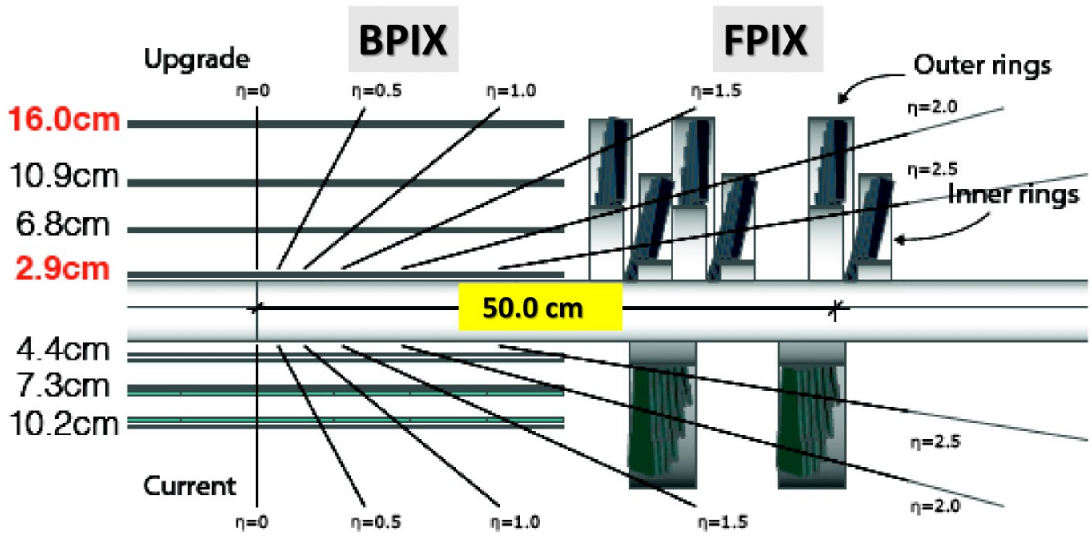


Figure 2.8: A comparison of the layout between the phase-0 (bottom) and phase-1 (top) pixel detectors. [6].

of the components used. Apart from the basic readout chain essentially remaining the same, the major changes correspond to an increase in the number of layers, a digital read out chip with increased buffers and data transmission bandwidth, and a change in the layout of the modules used for FPix. This will be described in more detail in Section 2.2.5.3. A comparison of the geometrical layout between the phase-0 and phase-1 detectors can be seen in Figure 2.8.

2.2.5.3 Phase 1 Upgrade of the Pixel Detector

The phase-1 pixel detector [65] was installed in the beginning of 2017 as part of an effort to improve tracking efficiency in anticipation of the increased luminosity that would occur in the following data taking period. The original detector was designed to operate at a maximum instantaneous luminosity of $1 \times 10^{34} cm^{-2} s^{-1}$. In the second half of Run 2, when the phase-1 detector was put into use, the instantaneous luminosity reached above $2 \times 10^{34} cm^{-2} s^{-1}$, which the phase-0 detector would not have been able to handle efficiently. At such a luminosity, it was foreseen that the phase-0 detector would experience high tracking inefficiencies due to buffer overflow in the ROCs. In order to prevent the degradation of performance of the CMS detector, the phase 1 upgrade was

installed.

In making new ROCs for the detector, new modules also had to be made, as the readout electronics are bonded to the pixel sensors. Instead of using different types of modules between BPix and FPix, the full 2×8 modules are used throughout the phase-1 detector. These modules introduce new ROCs which have an increased buffer size from 32 to 80 for the buffer depths of the hits, and from 12 to 24 for the time-stamp buffers. To increase the bandwidth of data readout, the ROCs moved from an analog to a digital signal, changing from 40 MHz to 160 Mbit/s. Each module is separated into two banks of ROCs, corresponding to the two columns. The data from these two banks are merged in the TBM using a 4-to-5 bit encoder, increasing the transmission rate to 400 Mbit/s.

The layer 1 modules contain a different version of ROC than the other layers in BPix and FPix, due to the higher hit rates. Layer 1 of BPix contains ROCs which are able to read out clusters of 2×2 pixels instead of the normal single pixel readout. These ROCs are also adjusted so that a reset signal is no longer required to continue after a readout, along with allowing seven columns of signal readout instead of the three in a normal ROC. The ROCs in the other layers and disks are just digital upgrades of the ROCs that were in the phase-0 detector, with 26 double columns of 80 pixels, each column having their own data and time stamp buffer [66]. Layers 1 and 2 also contain a special version of TBM which allows two data streams to be transmitted, with layer 1 modules using two of these special TBMs.

Apart from the new ROCs, other new changes included the replacement of the AOHs with pixel optical hybrids (POH), which are used to send the digital signals received from the ROCs and TBMs optically to the FEDs. The DAQ also received an upgrade, where the original VME electronics crates were replaced with μTCA crates, which required a redesign of these back-end boards, including the FEDs and FECs.

The last major change was the increase in the number of layers and forward disks by one for both parts. This increase has resulted in almost doubling the number of channels in the detector, and provides up to 4 hits for track reconstruction in the acceptance range of the pixel detector of

$|\eta| < 2.5$. With up to 4 available hits, a larger number of tracks can be reconstructed, and better momentum resolution of the tracks can be obtained. This allows for a better vertex resolution, as more tracks can be reconstructed at higher η and lower momentum. In addition, with a layer closer to the beam line, the impact parameter resolution of the tracks is also improved, as the path of a track does not need to be extrapolated as far of a distance to the vertex, due to it having a closer point of reference.

As part of the phase-1 upgrade, the University of Kansas was involved in the testing of pixel modules that were to be used in FPix. High rate studies using an x-ray box were performed on completed modules as part of their grading process, where each module would be given a grade of A, B, or C. In the high rate studies the x-rays were used to simulate hits in a detector, where single-hit efficiencies were measured, among other things. In addition, fluorescent materials were used in the x-ray box to measure the energy-response of the ROCs [6]. Modules with grades of A and B were used in the detector, with preference given to A modules, especially in the layers closer to the interaction points. Modules with a grade of C were discarded, with the production yield from working sensor to fully functioning module being about 80% [66].

The phase-1 detector was fully installed during the extended year-end technical stop at the end of 2016 and the beginning of 2017. It took data for the 2017 and 2018 data taking periods of Run 2, where for the second half of the running, it was made to run degraded due to the failure of components related to power distribution. In this case running degraded would mean that some of the modules were taken offline due to power-cycling failures. In the Run 2 data this introduces a decrease in the tracking resolution for the areas where these 'dead' modules exist, since the number of hits for any given track would be decreased from 4, to 3 or less.

Chapter 3

Compressed SUSY Search

In this dissertation a search is performed for pair produced stop quarks decaying to third or second generation quarks. The search uses $\sqrt{s} = 13$ TeV pp collision data collected by the CMS experiment during 2016-2018 for a total integrated luminosity of 137 fb^{-1}. It is geared towards compressed decay scenarios, where the mass difference between the stop and the LSP is small, on the order of the W mass or less. The stop search, is a part of a more generic analysis, the makeup of which targets generic SUSY signatures in compressed decay scenarios, by taking advantage of the high transverse momentum produced by initial state radiation (ISR) recoiling against a sparticle system. These generic signatures are massive LSPs, with R-parity conservation, which means the sparticles will always be pair produced, and decay to the LSP. Some of the design decisions for the stop search are based on what to expect from the other signal models in the generic search, but these regions that were designed more specifically for the stop search are those that concern secondary vertex tagging, an important part of this dissertation. What makes the generic search possible is the use of Recursive Jigsaw Reconstruction [67] to iteratively rebuild the kinematics in the event in order to approximate the different reference frames of a generic decay. In the following sections a description is given about the kinematic reconstruction of the events, including how Recursive Jigsaw Reconstruction is used. After that, is an overview of the analysis and a description of the signal models used. The following chapters then describe the SV tagging used (Chapter 4), a description of the event selection and regions (Chapter 5), and then the analysis results (Chapter 6).

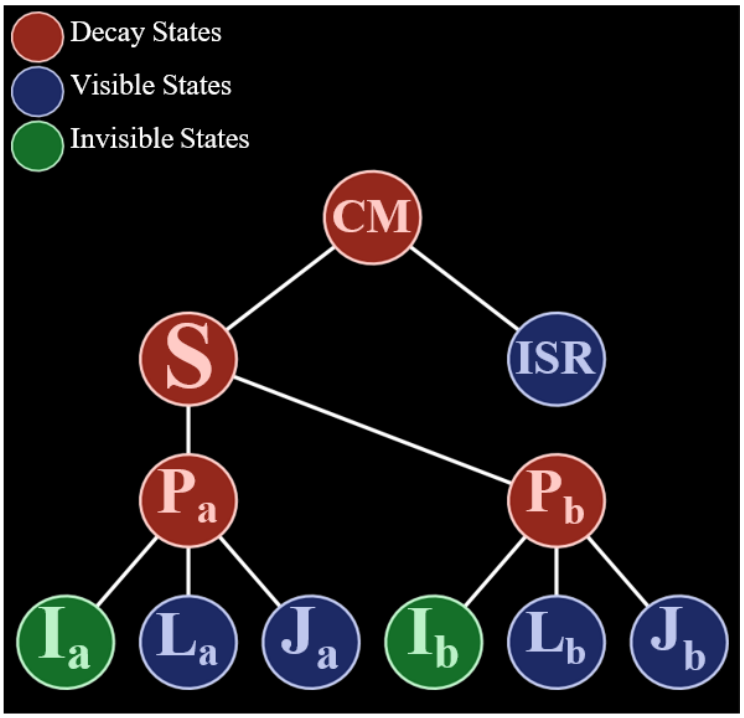


Figure 3.1: A graphic showing the decay tree used by the analysis. Starting from the top, each circle corresponds to the center of mass (CM), the sparticle system (S), the ISR system (ISR), the sparticle pairs (P_a, P_b), and then the Invisible (I_a, I_b) and Visible (V_a, V_b) systems. The Visible systems are further broken down into the lepton (L_a, L_b) and jet/SV (J_a, J_b) systems.

3.1 Kinematic Reconstruction

An important part of this analysis is how to interpret the kinematic information for each event. In the events under consideration there are two distinct systems, what is visible and what is invisible. The difficulty in interpreting the event is figuring out how the visible portion should be arranged based on possible configurations of the invisible portion, which includes the missing energy determination. This interpretation first requires the assumption of a decay tree, given in Figure 3.1, in which the visible objects (leptons, jets, SVs), are based, and from which kinematic observables can be calculated. The next step is then determining a process to assign the visible objects in the event, while accounting for the kinematic and combinatoric unknowns in the invisible portion. This is done in the analysis through the use of Recursive Jigsaw Reconstruction [68, 67], which can approximate the rest frames of the decay tree to measure mass sensitive variables.

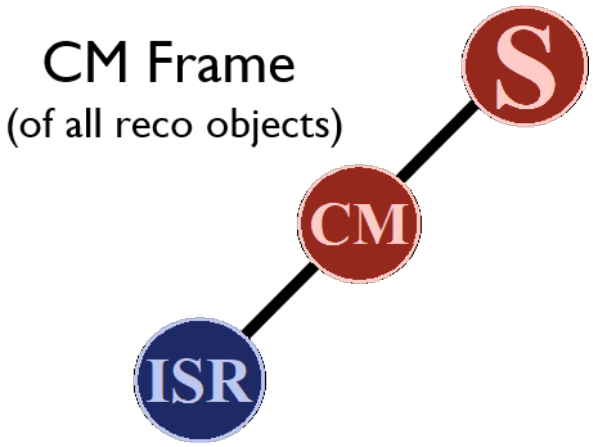


Figure 3.2: A graphic showing the center of mass frame of the decay tree in Figure 3.1. The center of mass is the rest frame of the proton-proton collision.

3.1.1 Recursive Jigsaw Reconstruction

The purpose of Recursive Jigsaw Reconstruction is to analyze an event according to a decay tree, and produce reference frames recursively starting from the final state. The end result of this is a kinematic basis for the event, from which one can transform the observable momenta in the final state between any of the reference frames of the decay tree. The available observables then consist of angles, energies, and masses in any of these reference frames, from which mass sensitive variables can be produced. The reference frames of particular note are the center of mass frame, see Figure 3.2, the lab frame, see Figure 3.3, and the sparticle frame, see Figure 3.4.

Starting from the decay tree in Figure 3.1, there is a collision from the center of mass which produces a sparticle (S) and ISR system. The ISR system gives a kick to the S system, depicted in Figure 3.5, and the S system produces a pair of sparticles, P_a, and P_b. These sparticles each decay to an invisible (I) and visible (V) system, to which the missing transverse momentum, ($\not{p}_T$), is assigned to the invisible systems, and the reconstructed objects are assigned to the visible system. The visble system is further divided into the jet (J) system, which includes both jets and SVs, and the leptons (L). The ISR system has no other divisions, so it is a single system, and does not contain a further divided J system (So J only refers to the jets or SVs in the S system). All objects in the event are assigned following a set of Jigsaw rules, which are a collection of kinematic min-

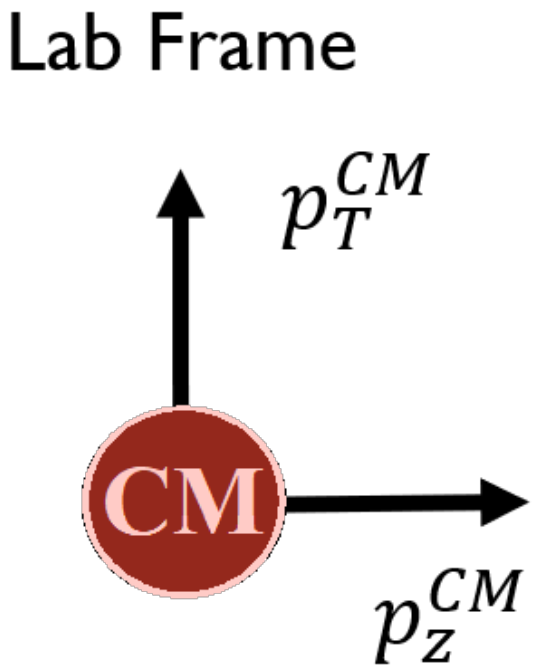


Figure 3.3: A graphic showing the laboratory frame of the decay tree in Figure 3.1.

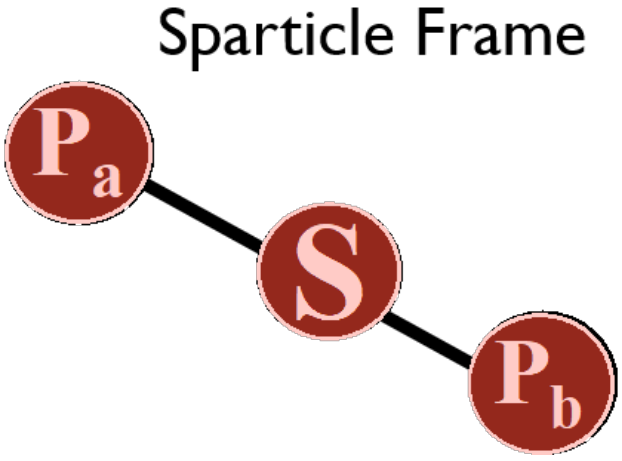


Figure 3.4: A graphic showing the sparticle frame of the decay tree in Figure 3.1.

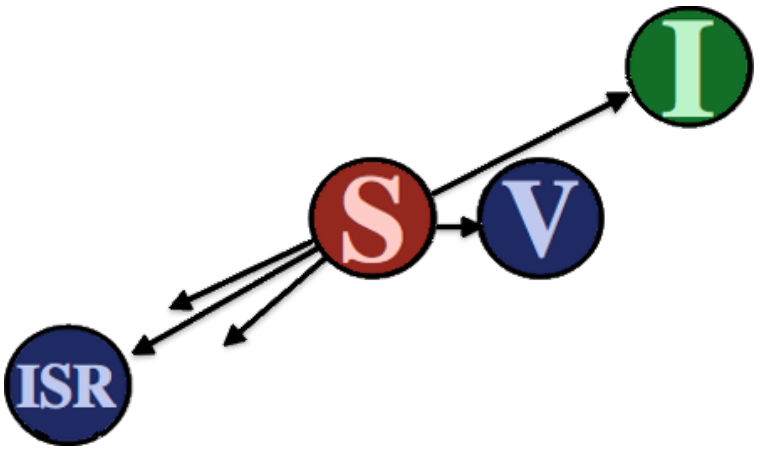


Figure 3.5: A graphic showing a sparticle system after being boosted by ISR recoil. S represents the sparticle system, V the visible system, and I the Invisible system.

imizations which determine how each system is treated with respect to the objects that contribute to it. The jigsaw rules used by the analysis are as follows:

- The Leptons are always assigned to the S system
- The Jets and SVs are assigned to either the S or ISR systems, such that the masses of the S and ISR systems, m_S and m_{ISR}, respectively, are minimized, which also corresponds to maximizing the momentum of the sparticle system, p_S, in the CM frame.
- The kinematics of I_a and I_b are approximated using the knowledge of $\not{p}_T$ coupled with the minimization of the sum of the masses squared of the P_a and P_b systems, $m_{P_a}^2 + m_{P_b}^2$
- In the S system, the Jets, SVs, (J) and Leptons (L) of the V system are subdivided into the P_a and P_b systems, such that the sum of the masses squared of the P_a and P_b systems, $m_{P_a}^2 + m_{P_b}^2$, is minimized.

As these four rules are dependent upon each other, they are all simultaneously minimized following Equation 3.1. The equation describes which quantities need to be simultaneously minimized, or maximized, and which parameters feed into those minimizations/maximizations. For example, the min term requires the minimum of $(m_{P_a}^2 + m_{P_b}^2)$, and uses the parameters located below the term, so the J, L, and I systems for the a and b hemispheres. The argmax term then, in

conjunction with the categorization produced by the min term, maximizes p_S^{CM} over the splitting of jets and SVs between the ISR and J systems. Running these jigsaw rules gave the object categorization on an event by event basis, which then allowed for the production of many mass sensitive variables used by the analysis.

$$\begin{pmatrix} ISR, J \\ J_a, J_b \\ L_a, L_b \\ I_a, I_b \end{pmatrix} = arg\max_{ISR,J} p_S^{CM}\Big(\min_{\begin{pmatrix} J_a, J_b \\ L_a, L_b \\ I_a, I_b \end{pmatrix}} \left[m_{P_a}^2 + m_{P_b}^2 \right] \Big) \tag{3.1}$$

3.1.2 Sensitive Variables

Analyzing the events through Recursive Jigsaw Reconstruction gives a set of kinematic observables for each of the reference frames in the decay tree; these are angles, energies and masses. Variables can then be calculated that are sensitive to the mass splitting of the pair produced sparticles and LSP.

The first, and most important of these variables is R_{ISR} with its definition given in Equation 3.2. This variable takes a simplified view of the event, where there are the visible (V), invisible (I), and ISR systems. In the CM frame, the ratio of the missing momentum, projected along the ISR kick direction and the magnitude of the ISR momentum approximates the ratio of the LSP and sparticle masses, without specific knowledge of what those masses are. This means that it is sensitive to the mass splitting, where the more compressed cases will fall at higher values of R_{ISR}. Take, for example, Figure 3.6, which shows a 2D plot with R_{ISR} on the x-axis for a stop signal with stop and LSP masses of 500 and 490 GeV, respectively for a 1 lepton $\geq$ 1 SV final state. It can be seen that a majority of this signal falls around an R_{ISR} value of 0.98, which is roughly close to the ratio of the two masses. Due to its sensitivity to this ratio, it is one of the two variables used in the 2D binning for each final state.

$$R_{ISR} = \frac{\left|\vec{p}_I^{CM} \cdot \hat{p}_{ISR}^{CM}\right|}{\vec{p}_{ISR}^{CM}} \sim \frac{\not{p}_T}{p_T^{ISR}} \tag{3.2}$$

The second most important variable is $M_\perp$, defined in Equation 3.3. This variable is derived in the sparticle rest frame, which ignores the momentum components parallel to the boost between the CM and S frames, becoming an ISR boost invariant mass. This variable is then sensitive to the magnitude of the mass splittings between the sparticles and LSP, as seen in Figure 3.6. Similar to R_{ISR}, it is sensitive to the LSP masses, but this time in the magnitude of the mass splitting, rather than the ratio of the masses. One caveat to this sensitivity is that it is not absolute, in that the final state chosen plays a large role in how close the distribution matches with the magnitude of the splitting. For example, in Figure 3.6, the $M_\perp$ distribution is localized around a value of 5 GeV, where the magnitude of the mass splitting for a stop mass of 500 GeV and an LSP of 490 GeV is approximately $(M_{\tilde{t}}^2 - M_{\tilde{\chi}_1^0})/M_{\tilde{t}} \sim 20$ GeV. This is due to how, for this 1 lepton $\geq$ 1 SV final state, there is an imbalance on the two sides of the event, with one side only having soft b-tagged SVs, forcing $M_\perp$ to a lower value. These imbalances will generally shift $M_\perp$ values slightly below the magnitude of the mass splitting, while still being sensitive. Due to its sensitivity, $M_\perp$ is then the second mass sensitive variable that is used in the 2D binning for each final state. Together the $M_\perp$ and R_{ISR} variables are sensitive to both the ratio and the difference of the sparticle and LSP masses, while also being largely uncorrelated, making it ideal for 2D binning, as depending on where the divisions are, each final state can be sensitive to multiple different sparticle masses and splittings.

$$M_\perp = \sqrt{\frac{M_{P_a}^2 + M_{P_b}^2}{2}} \tag{3.3}$$

The remaining sensitive variables that are used in the analysis are listed here:

- p_T^{ISR}: The transverse momentum of the ISR system, necessary for producing a large enough ISR kick such that the missing transverse energy becomes large enough to support a massive LSP.

- p_T^{CM}: The difference between $\not{p}_T$ and $\not{H}_T$, the missing transverse momentum of jets. When

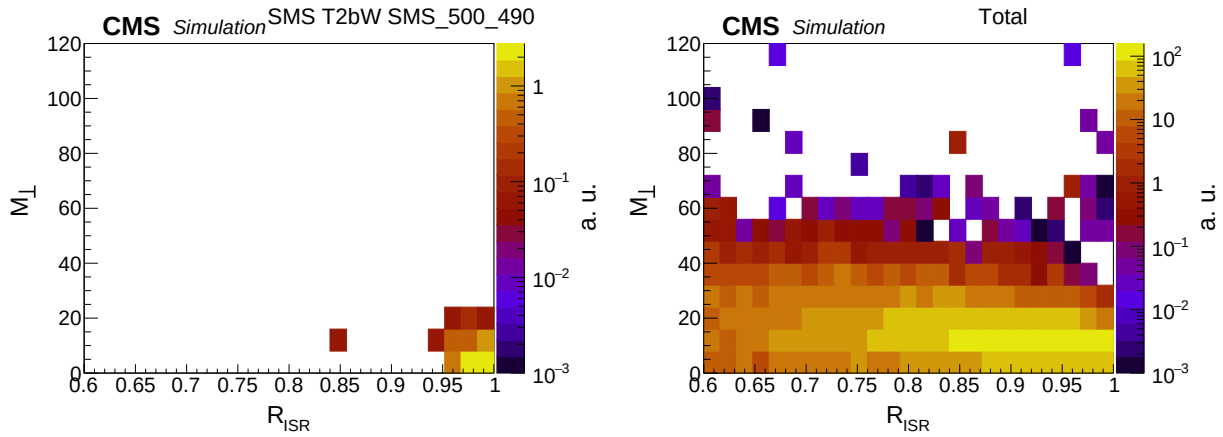


Figure 3.6: Some 2D distributions of $M_\perp$ versus R_{ISR} for a pair produced stop signal (left) with 500 and 490 GeV masses for the stop and LSP, respectively and the total background (right), not including QCD multijets. This is for a $1L_G - 0J_{0b}^{g1SV} - X_{0b}^{P_{300}}$ region. These plots show how the signal is very localized when binned in these two variables, while this is not the case for the background.

used in conjunction with $\Delta\phi_{CM,I}$ it can isolate poorly reconstructed events in data through an induced correlation.

- $\Delta\phi_{CM,I}$: In the CM frame, the difference in ϕ between the I system and the direction of the boost between the CM and lab frames. This produces a correlation between $\not{p}_T$ and the p_T^{CM} boost.

- $\Delta\phi_{\not{E}_T,V}$: The difference in ϕ between $\not{p}_T$ and the V system.

- γ_T: Equal to $2M_\perp/M_S$, and is sensitive to an event imbalance, where a kinematically balanced event is expected for pair produced sparticle decay. It is used for categorization in higher multiplicity final states.

3.2 Analysis overview

The analysis searches for generic SUSY signatures making use of ISR recoil to boost the sparticle system. These generic signatures are massive LSPs, with R-parity conservation, which means the sparticles will always be pair produced, and decay to the LSP. The analysis is run on the

NanoAOD data format, which stores all of the necessary reconstructed objects in a flat data tree for ease of use. It looks at multi-lepton final states corresponding to 0, 1, 2, and 3 leptons, where events are kinematically reconstructed with RJR following the decay tree in Figure 3.1, using lepton, jet, SV, and $\not{p}_T$ objects. The full set of regions in the analysis was determined by their reconstructed object multiplicity, categorizing by lepton flavor and charge for the leptonic side, and by jet, b-jet, and secondary vertex multiplicities for the hadronic side. Each event could only be categorized into one region, and then the number of events was counted for each region. The object categorization for the hadronic side involved enumerating the jets, b-jets and SVs over the S system, and enumerating the b-jets over the ISR system. The categorizations involving SVs were further subdivided into central and forward η. Apart from the categorizations based on objects, there were also categorizations based on two mass sensitive kinematic observables, p_T^{ISR}, and γ_T, which were described in Section 3.1.2. A full region would then be a specific combination of the previously mentioned categorizations, with an example being $1L_G^{\mu} - 0J_{0b}^{g1SVc} - X_{0b}^{P_{300}}$, which translates to a selection of 1 golden muon, of any charge, 0 jets, 0 b-jets, and ≥ 1 central SV in the S system, and then 0 b-jets in the ISR system, with $p_T^{ISR} > 300$ GeV, and no further γ_T categorization. The nomenclature for the region names is given in Table 5.8 of Section 5.4. Lastly, there was then a binning of these final states in $M_{\perp}$ and R_{ISR}, which are used consistently between the different final states. The 2D bins in $M_{\perp}$ and R_{ISR} for all the regions were flattened into a 1D histogram where each bin corresponded to an $R_{ISR} - M_{\perp}$ bin, and each histogram corresponds to a region with a set of categorizations applied.

Once the region histograms were defined, before the actual fitting, the systematics needed to be determined. The main systematics to be considered are those related to the background normalization and object reconstruction. For the background, normalization scale factors were implemented. For the leptons, data to MC efficiencies and scale factors were measured, and in addition a data driven lepton fake estimate was produced. For the jets there were uncertainties related to jet energy scale and resolution measurements, which were also propagated to $\not{p}_T$. For the b-jets, b-efficiencies in MC samples were measured and used to apply centrally produced data

to MC scale factors. For SVs, scale factors were measured and constrained by data within the analysis fit. There were also systematics related to MC simulation and data collection, such as MC scale factors and uncertainties applied due to the trigger turn on, all to be described in Section 6.2.

When both the regions and systematics were determined, the region histograms were used as input into the Higgs Combine analysis package, which is a command line interface for doing analyses using the statistical frameworks, RooFit and RooStats [69], inside of the ROOT analysis framework [70], in order to perform the limit setting to produce the expected and observed limits for the various stop mass grids. In the fit, all of the regions are connected by the systematic scale factors. The major backgrounds have floating normalizations and are constrained by the data, while the rare processes have scale factors constrained by simulation. As already mentioned, the b-tag scale factors have floating shape variations and are constrained by data, while the SVs scale factors are derived from data in the fit. The lepton fakes have scale factors and shapes that are determined from the data. Lastly there are the kinematic normalizations constrained by data, and the ISR scale factors that are derived from the data. The fits were then performed and limits were derived using a modified frequentist approach, CL_S criterion, and an asymptotic method [71], for the mass grids of five simplified models of stop pair production.

3.3 Signal Models

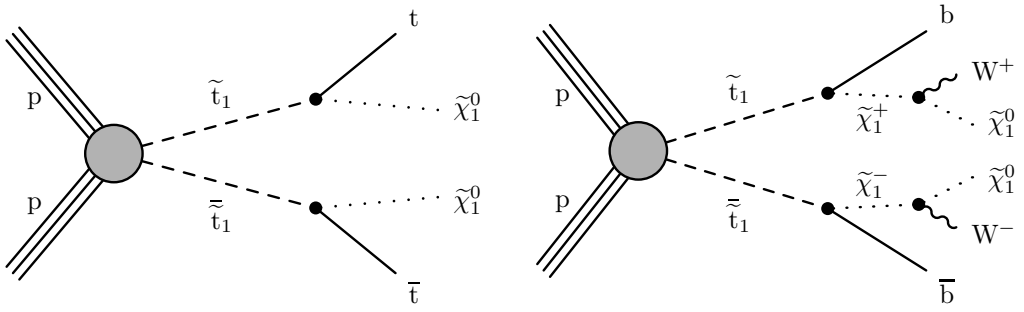
There are 5 signal models under consideration within this dissertation, which is only a part of the overall analysis. These are all stop pair produced signals with masses ranging from 500 to 1000 GeV and mass splittings ranging from 10 GeV up to the top corridor.

The signals use the simplified model spectra (SMS) [72, 73, 74, 75] which assume a 100% branching ratio to a particular final state, always includes the LSP as $\tilde{\chi}_1^0$, and may contain $\tilde{\chi}_1^\pm$ as an intermediary in the decay. For the models where there is an intermediary, its mass is assumed to be $m_{\tilde{\chi}_1^\pm} = 0.5(m_{\tilde{t}} + m_{\tilde{\chi}_1^0})$. The stop signal model decays are then:

- **T2tt**: $\tilde{t}\bar{\tilde{t}} \rightarrow t\tilde{\chi}_1^0 \bar{t}\tilde{\chi}_1^0$

- **T2bW**: $\tilde{t}\bar{\tilde{t}} \to b\tilde{\chi}_1^+ \left(W^+\tilde{\chi}_1^0\right) \bar{b}\tilde{\chi}_1^- \left(W^-\tilde{\chi}_1^0\right)$
- **T2-4bd**: $\tilde{t}\bar{\tilde{t}} \to bf\bar{f}\tilde{\chi}_1^0\bar{b}f\bar{f}\tilde{\chi}_1^0$
- **T2cc**: $\tilde{t}\bar{\tilde{t}} \to c\tilde{\chi}_1^0\bar{c}\tilde{\chi}_1^0$

The Feynman diagrams for these signals are given in Figure 3.7. With the exception of the T2cc signal, all of these stop signals will have the same final state $b\bar{b} + W^+W^- + \tilde{\chi}_1^0\tilde{\chi}_1^0$, as the top quark will decay to bW. The major differences between each of these are the mass points that should be targeted. For T2tt, a search would be ideal for mass splittings at or beyond the top corridor, as that is the only way to get on-shell tops in the decays. Next, is then T2bW which is geared towards signals with mass splittings between the top and W corridors, and finally, T2-4bd which is geared towards signals below the W corridor, due to the direct production of b quarks and fermions (f), rather than a W boson. T2cc, as it has a final state that is on the lower end of possible energies due to the lower mass of the c quark, compared to the b quark, is geared towards mass splittings even smaller than that of T2-4bd.



(a) T2tt

(b) T2bW

(c) T2-4bd

(d) T2cc

Figure 3.7: The Feynman diagrams of the studied SMS stop samples.

Chapter 4

Soft b tagging

In a compressed SUSY search, it is important to be able to reconstruct low p_T objects as decay candidates due to the low amount of energy available. When looking at a pair of stops decaying, an important set of decay chains ends with b-quarks in the final state, for both hadronic and leptonic versions of the decay. In a compressed scenario, it is expected that these b quarks will be soft, for which current b tagging methods are unable to explore for the smallest mass splittings, with current jet b tagging at CMS only in use for jets with a $p_T > 20$ GeV [76].

The ideal p_T range for soft decay products is less than 20 GeV, which precludes the use of current jet b tagging methods. This can be seen in Figure 4.1, which gives the p_T spectrum of generator b's, which shows that for SUSY signals with a mass splitting less than 80 GeV, corresponding to T2-4bd and T2bW production, a majority of the b quarks are found with $p_T < 20$ GeV. This can be compared to $t\bar{t} + jets$ production, which is highly likely to have b quarks in the final state, and where a majority of the b quarks are found above this p_T range, as they are not limited by any mass splitting, and have much more energy to work with.

With a p_T range of less than 20 GeV, it is necessary to determine the reconstructed object that can be best used in this regime. For hadronic activity, the possible choices are tracks, secondary vertices, and jets (along with their associated activity in the HCAL). For b quarks, single tracks by themselves would not give enough information on how far away a particle decayed from the primary vertex, which is necessary for a b quark, due to its relatively long and definite lifetime. Secondary Vertices (SVs) on the other hand, have at least two tracks with which a decay vertex, displaced from the primary interaction, can be more reasonably determined and their transverse momentum can have a range down to at least 0.5 GeV. Jets will generally have at least one sec-

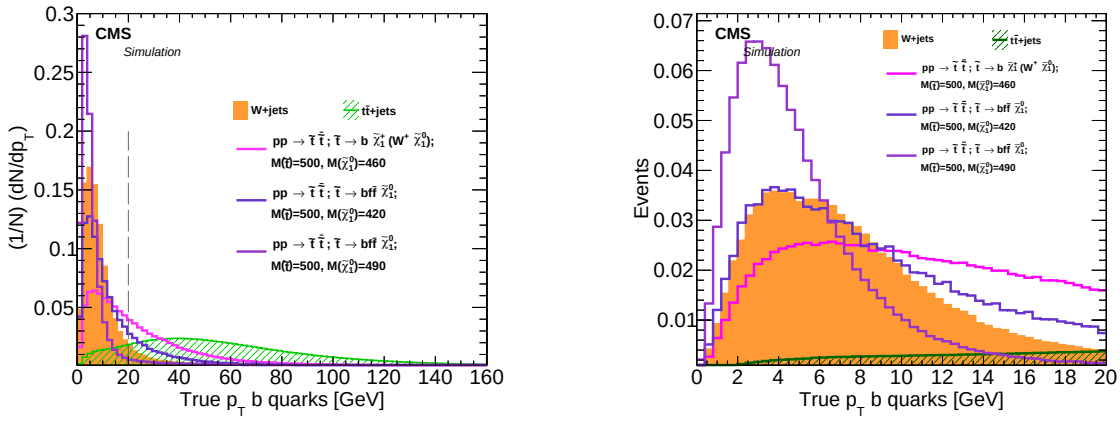


Figure 4.1: An overlay plot of the p_T spectrum, normalized to unity, of generator b quarks for different simulated samples: $t\bar{t}+jets$, $W+jets$, SMS-T2bW $m_{\tilde{t}} = 500; m_{\tilde{\chi}_1^0} = 460$, and SMS-T2-4bd $m_{\tilde{t}} = 500; m_{\tilde{\chi}_1^0} = 420, 490$. The plot on the left is the same distribution zoomed in on the 0-20 GeV p_T range. These plots show how the compressed SUSY signals have a majority of their generator b quarks at a p_T less than 20 GeV.

ondary vertex, so it is possible to determine a decay length and they are commonly used in b tagging. The jet momentum, though, only goes down to 10 GeV, a point at which it is still too high for a lot of the desired b activity, as seen in Figure 4.1. The obvious choice is then to use secondary vertices (SVs) as a base for tagging low p_T objects as b quarks. Section 4.1, then, describes the variables that were considered in the production of such a tagger.

4.1 Secondary Vertex (SV) variables

For the creation of the soft b-tagger there were 10 different SV variables that were considered. These variables were sourced from the `slimmedSecondaryVertices` collection in the MiniAOD data format, which is also the same collection used by the NanoAOD data format to produce its own SV collection, but drops a lot of information for the sake of reducing event size. The observables that are stored in the base vertex collection are position, error, χ^2, number of degrees of freedom, and reconstructed tracks. The rest of the variables such as four-momentum, are then calculated from the base collection. The variables used, which are all present in, or calculable from the MiniAOD dataset format are:

- p_T: The transverse momentum of the sum of the tracks associated with the SV.
- η: The pseudo-rapidity of the sum of the tracks associated with the SV.
- M: The mass of the sum of the tracks associated with the SV.
- N_{trk}: The number of tracks associated with the SV that have a minimum track weight of 0.5
- N_{dof}: The sum of the track weights for the tracks associated to the SV
- d_{xy}: The transverse (2D) distance between the SV and the primary vertex (PV), also referred to as the 2D impact parameter, 2D decay length or 2D flight distance.
- SIP_{xy}: The 2D impact parameter divided by its uncertainty.
- d_{3D}: The 3D distance between the SV and PV, also referred to as the 3D impact parameter, decay length or flight distance.
- SIP_{3D}: The 3D impact parameter divided by its uncertainty.
- $\cos\theta_{SV,PV}$: The cosine of the angle between the SV momentum and the vector pointing from the primary vertex to the secondary vertex, where the angle is referred to as the flight direction.

The SVs were studied using the two different data formats previously mentioned. The format used to produce the training sample made use of the MiniAOD data format, while the one used by the analysis is the NanoAOD format. These two formats have slightly different SV content, where the NanoAOD does not have N_{trk} or any other track information stored, but it has all of the other listed variables stored, and has an additional selection requirement of $SIP_{3D} > 3$. The SIP_{3D} cut is important, as it greatly changes the SV content between MiniAOD and NanoAOD, causing a substantial difference in the performance of the soft b-tagger. For the N_{trk} variable, it is replaced by N_{dof} where relevant, as N_{dof} is closely related to N_{trk}, due to it being the sum of the track weights.

4.2 Previous soft b-tagging with square cuts

To get a baseline of where to start, a study was performed on what has been done for soft b tagging in other SUSY analyses [77]. This previous b tagging went with simplicity as its design, and applied flat cuts on a set of SV observables in order to tag SVs. This tagging, which will be dubbed as the 'square cut' method, had a selection on the SVs of $N_{trk} > 2$, $d_{xy} < 2$, $SIP_{3D} > 4$, and $\cos\theta_{SV,PV} > 0.98$. When comparing results with the new soft b-tagger, the number of tracks cut is changed to $N_{dof} > 1.8$ in order to match the variables present in NanoAOD. The square cuts method also required a jet disambiguation in which all SVs matched to jets with $p_T > 15$ GeV failed the selection. For this dissertation, the definition was changed so that only SVs matched to jets which are considered by b tagging, i.e. jets with $p_T > 20$ GeV, fail the selection. The baseline performance of the square cuts method is given in Figures 4.2 and 4.3. The plots show the performance for different selections of SVs, one set with and one without a $SIP3D > 3$ cut. Improving this performance is the main goal in producing the new soft b-tagger.

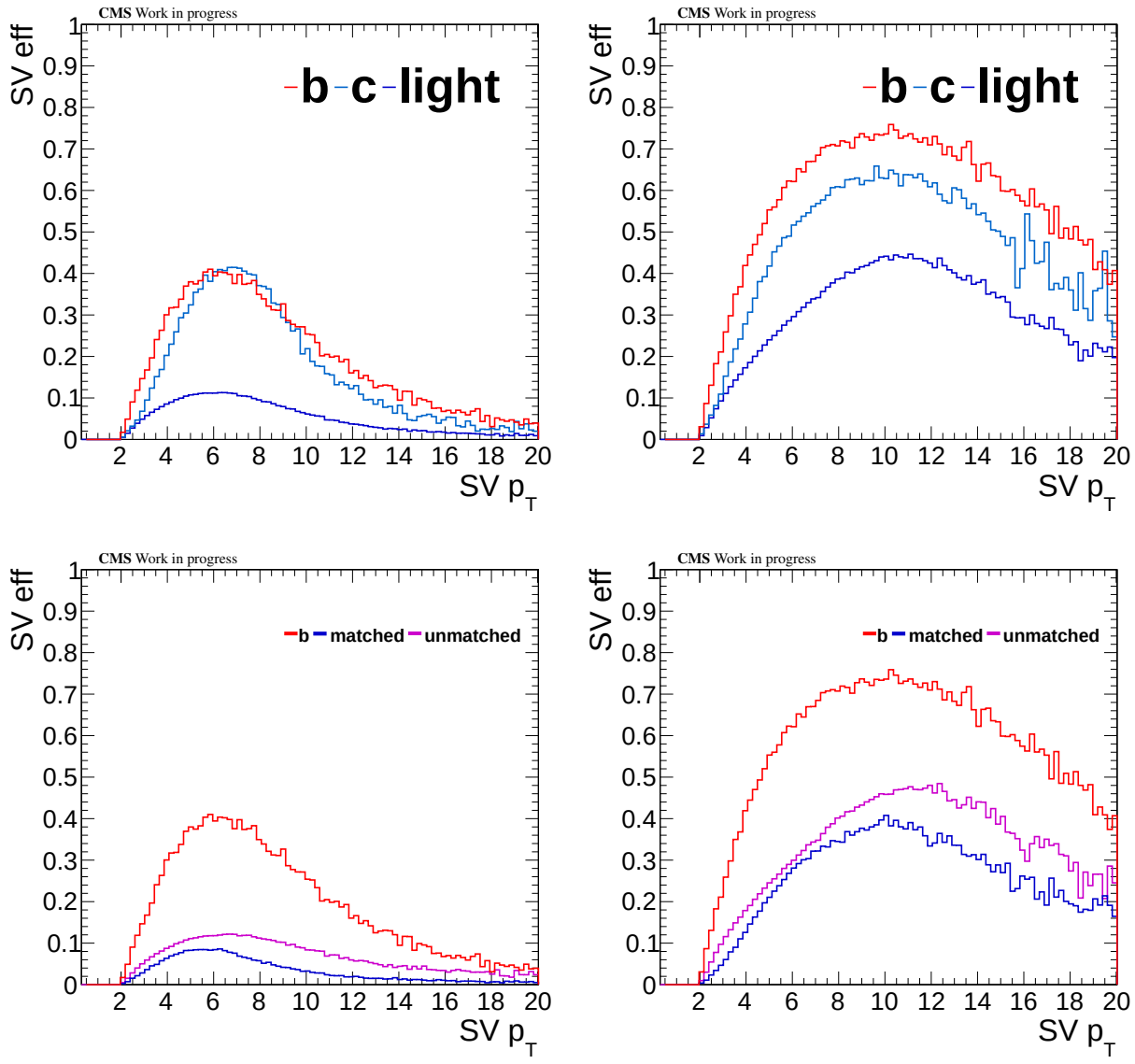


Figure 4.2: Performance over p_T of b, *c* and *light* flavor SVs (TOP) and b, matched, and unmatched flavor SVs (BOTTOM) using the square cut selection. The left plots show the flavor efficiency of selected SVs, with the SVs in the denominator only passing a p_T cut, and corresponding to how SVs are stored in MiniAOD. The right plots show the same performance, except the SVs in the denominator have an additional cut on the decay length significance (SIP3D) > 3, which matches the SVs as they are stored in the NanoAOD format.

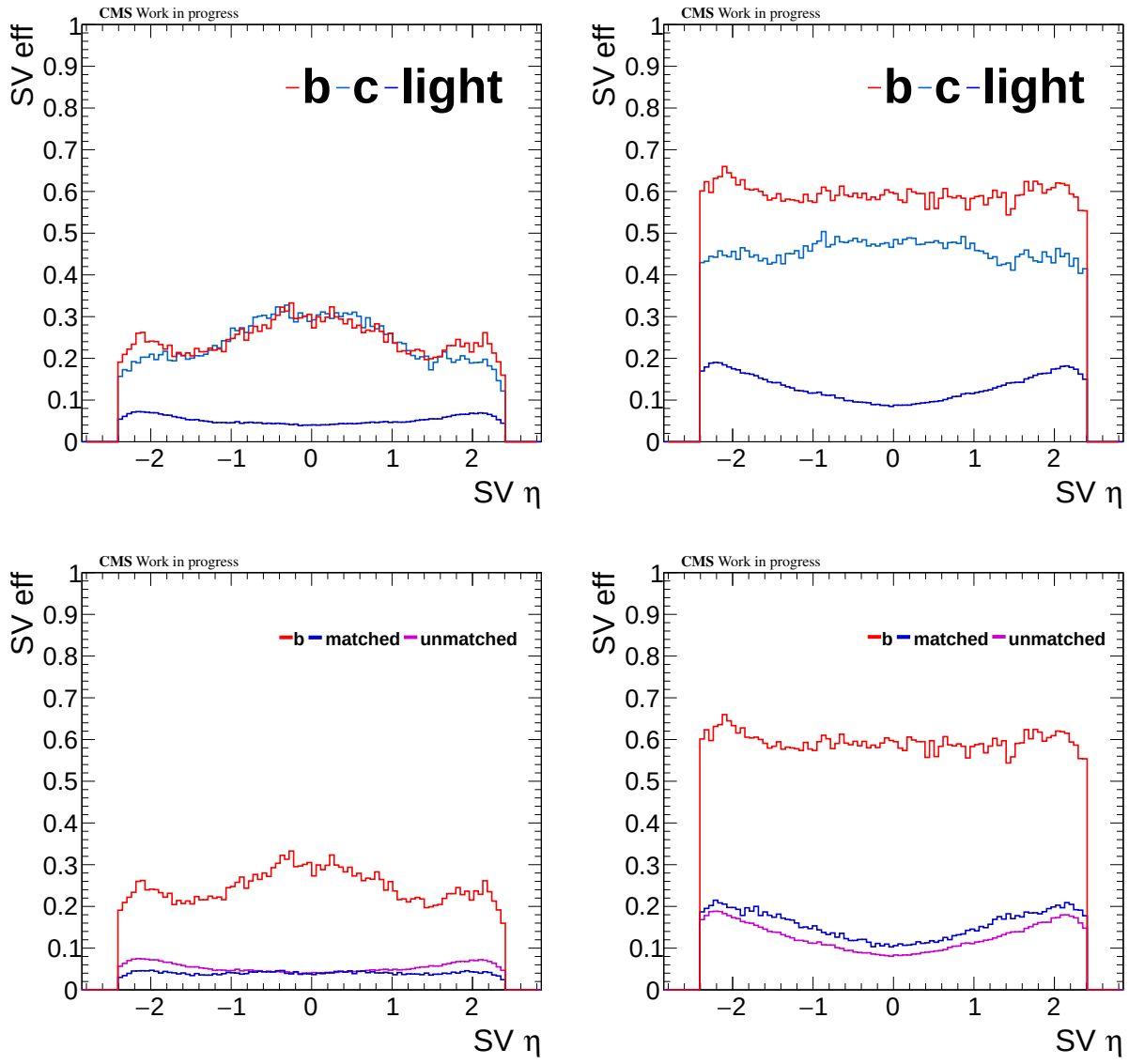


Figure 4.3: Performance over η of b, *c* and *light* flavor SVs (TOP) and b, matched, and unmatched flavor SVs (BOTTOM) using the square cut selection. The left plots show the flavor efficiency of selected SVs, with the SVs in the denominator only passing a p_T cut, and corresponding to how SVs are stored in MiniAOD. The right plots show the same performance, except the SVs in the denominator have an additional cut on the decay length significance (SIP3D) > 3, which matches the SVs as they are stored in the NanoAOD format.

4.3 Making a soft b tagger

Following what is done for jet *b* tagging, it was decided to use the secondary vertices as inputs into a deep neural network (DNN) to produce a discriminator that best makes use of the available SV observables.

In producing a new tagger, this dissertation starts with what was used by the square cuts, only considering the 4 SV variables and only using fully isolated SVs. This base was built upon and adjusted, changing the type of SV that is used, and adding or changing variables used as input into the training algorithm until one arrives at the final product that is used in the analysis. Before going through all of these changes, one important distinction between the initial SVs used in training, and the final SVs used in training, is the difference in how the SVs are treated between two of the data formats used by CMS. As previously described, the NanoAOD format, stores slightly different SVs than the MiniAOD format. The difference in these two formats can best be seen in Figures 4.2 and 4.3, which gives the square cut performance for SVs following the difference in the SIP_{3D} requirement between the MiniAOD and NanoAOD data formats. From these figures, the difference between the two sets of SVs is substantial. This difference is addressed in the last set of trainings to be described in Section 4.5.

4.4 SV preparation

Before doing any training, it was necessary to prepare the SVs that were to be input into the DNN. In order to ensure access to all of the necessary SV variables, and to streamline the calculation of some variables, it was decided to use the MiniAOD data format with the `slimmedSecondaryVertices` object. This was ideal due to the readily available plugins for calculating variables found in CMSSW, which are only usable with EDM objects.

MC Samples
TTJets_DiLept_TuneCP5_13TeV-madgraphMLM-pythia8
TTJets_SingleLeptFromT_TuneCP5_13TeV-madgraphMLM-pythia8
TTJets_SingleLeptFromTbar_TuneCP5_13TeV-madgraphMLM-pythia8
TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8
TTToHadronic_TuneCP5_13TeV-powheg-pythia8
WJetsToLNu_HT-70To100_TuneCP5_13TeV-madgraphMLM-pythia8
WJetsToLNu_HT-100To200_TuneCP5_13TeV-madgraphMLM-pythia8
WJetsToLNu_HT-200To400_TuneCP5_13TeV-madgraphMLM-pythia8
WJetsToLNu_HT-400To600_TuneCP5_13TeV-madgraphMLM-pythia8
WJetsToLNu_HT-600To800_TuneCP5_13TeV-madgraphMLM-pythia8
WJetsToLNu_HT-800To1200_TuneCP5_13TeV-madgraphMLM-pythia8
WJetsToLNu_HT-1200To2500_TuneCP5_13TeV-madgraphMLM-pythia8
WJetsToLNu_HT-2500ToInf_TuneCP5_13TeV-madgraphMLM-pythia8
WJetsToQQ_HT400to600_qc19_3j_TuneCP5_13TeV-madgraphMLM-pythia8
WJetsToQQ_HT600to800_qc19_3j_TuneCP5_13TeV-madgraphMLM-pythia8
WJetsToQQ_HT-800ToInf_qc19_3j_TuneCP5_13TeV-madgraphMLM-pythia8

Table 4.1: A list of the MC samples used in the training and testing of the SVs. These are produced simulating 2017 run conditions using the global tag `94X_mc2017_realistic_v14`

4.4.1 Samples and selection

The simulated samples that were used in the training are $t\bar{t} + jets$ with leptonic, semi-leptonic and hadronic decay modes, and $\hat{H}_T$ binned $W + jets$ in leptonic and hadronic decay modes. These samples, given in Table 4.1, were centrally produced for use by CMS analyses, using the global tag `94X_mc2017_realistic_v14`, and stored in the MiniAOD data format. As denoted in the global tag, these samples are simulated using a realistic CMS detector with Run II 2017 conditions and 2017 pile-up. It was decided to use 2017 MC as it is the intermediate set of samples between 2016 and 2018 data, and should be best representative of overall conditions between the three data taking periods. It also includes the phase I pixel detector, which is especially important since SVs are track based objects. The $t\bar{t} + jets$ samples were chosen so as to guarantee events that have b quark content, making it more likely for low p_T SVs to be matched to a b. The $W + jets$ samples were added in to make the events that the SVs are sourced from more diverse, while also not introducing too much *light* flavor activity, as would be the case with QCD multijets samples.

For the event selection, the only requirement was the presence of at least one good primary

vertex that was classified as not fake, with the number of degrees of freedom > 4, $|z| < 24$, and $\rho < 2$. The SVs selection initially just had a requirement on their transverse momentum of $p_T < 20$ GeV, But this was later changed to include the cut, $SIP_{3D} > 3$, so as to match SVs used in the NanoAOD format.

4.4.2 Flavor assignment

The most important part of the SV selection and preparation, was the assigning of particle flavor, as SVs do not have a centrally assigned flavor like leptons and jets generally do. For the case of the SVs, the flavor needs to be assigned using a generator particle collection, in this case the collection used was `prunedGenParticles`. Following how flavor is assigned to jets, SVs were assigned flavor in a similar manner. The overall method involves ΔR matching the SVs to generator particles, and incrementing a counter depending on whether the gen particle is a bottom, charm, gluon, strange, up/down, other or no match. After looping through all of the generator particles, a set of if-else statements are used sequentially to determine the SV flavor, which forces each SV to have a single assigned flavor, making sure that the SVs will be assigned to exclusive categories in the training. Starting with the heavy flavors, if the SV has at least one b quark then it is assigned as having b flavor, if it fails this then there is a check on the number of c quarks, then the number of gluons, then the number of s quarks, then the number of u or d quarks, then the number of matched particles not having any of these flavors, and lastly whether the SV is unmatched to any generator particle.

As part of the flavor assignment, it was necessary to determine the magnitude of the ΔR matching, in order to say that a generator particle comes from an SV. There were three different magnitudes looked at, $\Delta R = 0.01$, 0.05, and 0.1. In looking at these values, a balance needed to be found between having enough SVs categorized as heavy flavor, while also increasing the likelihood that a given SV truly comes from the matched generator particle. In general, while looking at both $t\bar{t} + jets$ and $W + jets$ samples, it was found that a $\Delta R = 0.01$ is too tight of a matching criteria, where a majority of the SVs were getting categorized as unmatched, as can be seen in Figure 4.4.

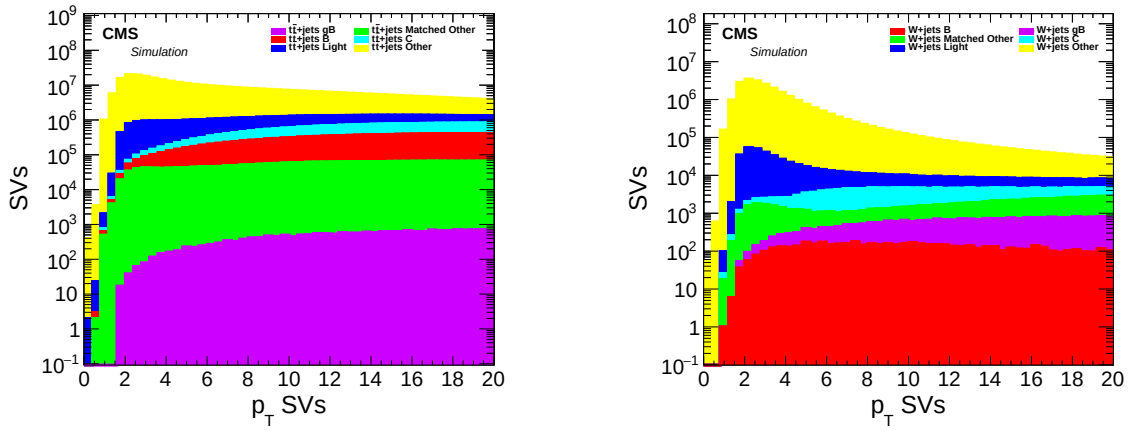


Figure 4.4: Stacked plots in log scale of the p_T of SVs matched to generator particles within a $\Delta R = 0.01$. The flavors shown are b (red), b from a gluon (magenta), c (light blue), *light* (dark blue), matched but other flavor (green), and unmatched to a generator particle (other, yellow). $t\bar{t} + jets$ simulation is on the left, and $W + jets$ simulation is on the right

On the other hand, $\Delta R = 0.1$ was too loose, where, while the heavy flavor content was increased, there is much less of a guarantee that the SVs came from a given matched generator particle, as the most desirable case is where the generator particle's direction is in line with the direction of the SV as close as possible. In the end, the rough midpoint between the two of $\Delta R = 0.05$ was chosen as a good balance. This ends up being tight enough such that the directional axis of the generator particles and SVs are more in line with each other, but not so tight that SVs that would otherwise be matched to a heavy flavor, instead go unmatched. Figures 4.5 and 4.6 give the flavors of SVs over p_T for the other two cases of $\Delta R = 0.05$ and 0.1, respectively.

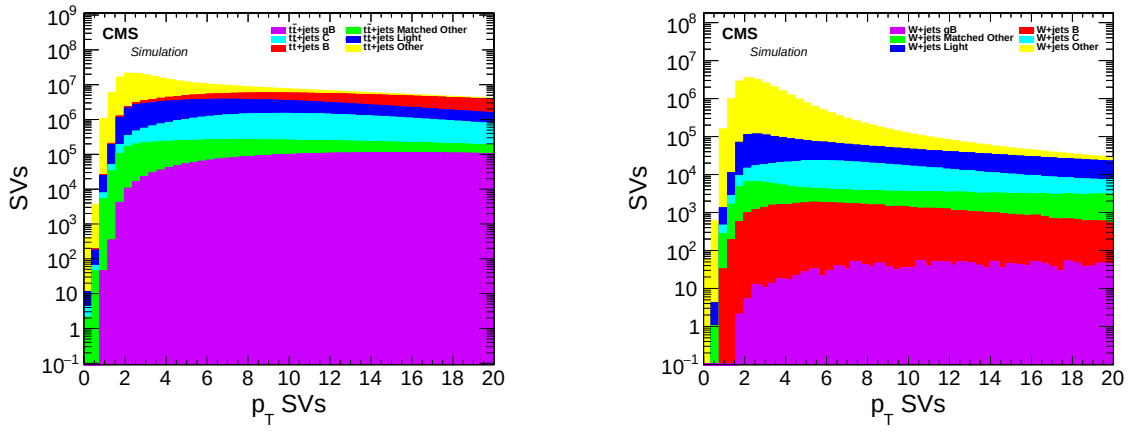


Figure 4.5: Stacked plots in log scale of the p_T of SVs matched to generator particles within a $\Delta R = 0.05$. The flavors shown are b (red), b from a gluon (magenta), c (light blue), $light$ (dark blue), matched but other flavor (green), and unmatched to a generator particle (other, yellow). $t\bar{t} + jets$ simulation is on the left, and $W + jets$ simulation is on the right

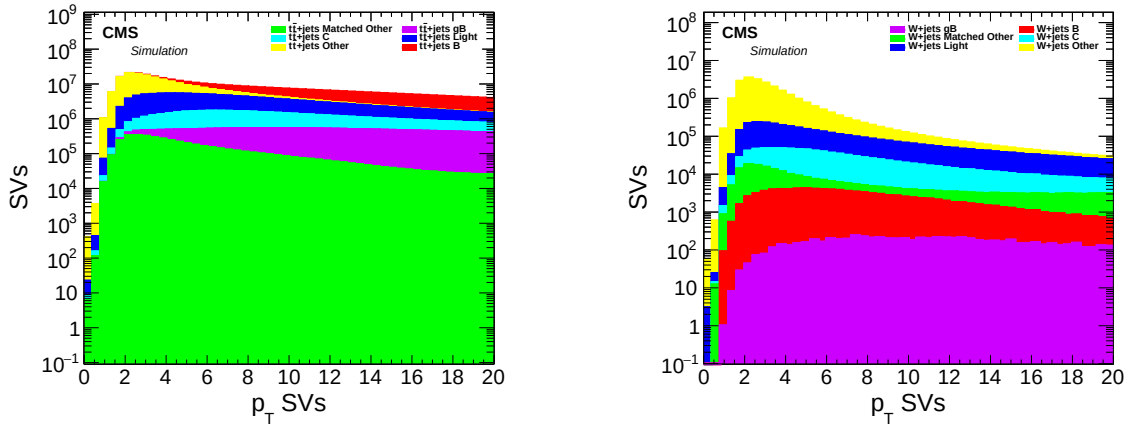


Figure 4.6: Stacked plots in log scale of the p_T of SVs matched to generator particles within a $\Delta R = 0.1$. The flavors shown are b (red), b from a gluon (magenta), c (light blue), $light$ (dark blue), matched but other flavor (green), and unmatched to a generator particle (other, yellow). $t\bar{t} + jets$ simulation is on the left, and $W + jets$ simulation is on the right

4.5 Training

The training of the SVs was done using the DeepJet framework [78, 79], which is a software package that is used for the training of the jet taggers used by the CMS experiment. This framework was adjusted slightly so that it could run solely on SVs, instead of jets. Multiple trainings were performed using the SVs as input, with each one adjusting inputs to compare how different changes affect the overall performance. These changes include sample size, sample content, number of classifiers, and number and type of variables. Most of the trainings to be shown used a common set of data, on the order of $70,000,000$ secondary vertices (SV), with a mixture of SVs from leptonic, semi-leptonic and hadronic $t\bar{t} + jets$ and $W + jets$ Monte Carlo simulation. These samples were chosen due to their b content, with more of the $light$ content coming from the $W + jets$ simulation. QCD multijets was not chosen as a sample due to the abundant amount of $light$ content over b content already present in the two chosen samples. The testing set contained a mixture of SVs, similar to the training sample, but on the order of $10,000,000$ SVs. The training algorithm is duplicated from what was used for the DeepCSV tagger [76, 80], a jet b tagger used by the CMS experiment, due to its relative simplicity.

The algorithm consisted of 7 layers, with 5 dense hidden layers of 100 nodes each. The dense layers used a rectifier activation function with a dropout rate of 0.1, and the output layer used a soft max activation function. The data was trained over 50 epochs, with a batch size of 5000. It was decided to use similar machine learning parameters as those used in the DeepCSV training, so as to focus on just the SV inputs that were going into the training, keeping everything else constant. As the DeepCSV tagger has gone through its own iterations, the hyperparameters of the algorithm may have some differences now compared to when the soft b-tagger was produced, but they were the same at the time of the training of the soft b-tagging discriminator.

Before running on the SVs specifically, some tests were done with adjusting some training parameters using the original DeepCSV training algorithm including low p_T jets. This included changing the number of nodes per layer, the weights applied to specific flavors of jets, and adjusting the batch size. Beyond increasing the run time with respect to changing the batch size, the results

were inferior or comparable to when keeping all the parameters the same, and large differences were not seen when adjusting these parameters.

For the classification scheme, SVs are matched to generator particles within a ΔR cone size of 0.05. Based on this matching there are four classifiers used:

- **isB**: SVs matched to a *b* quark.
- **isC**: SVs matched to a *c* quark.
- **isMatched**: SVs matched to anything else
- **isUnMatched**: SVs that are unmatched to generator particles

The last two classifiers, isMatched and isUnmatched, combined, correspond to the *light* flavor category that will be shown in later figures. Some plots look at the b, matched and unmatched flavors, corresponding to the early trainings, while the final trainings combine the matched and unmatched flavors together into one, and add the *c* flavor efficiency in addition to *b* and *light*. For all of these classifiers, there is also a jet disambiguation where SVs are removed if they are matched to a jet within a ΔR cone size of 0.4, whereas any SVs that are included in jets are taken care of by the jet *b* tagging that is performed. Depending on the training, the jets that were removed changed to more closely match the jet disambiguation used in the analysis, where the jets that are not included in jet *b* tagging ($p_T < 20$ GeV) are kept to improve signal acceptance at higher SV p_T. Figure 4.7 gives an example of how including all SVs matched to jets affects the performance of the training, using an arbitrary discriminator value of 0.2 and 0.15. By including all of the SVs, the efficiency of the SVs of interest, those that are isolated from the jets used in *b* tagging, take a severe hit due to the greater number of SVs in jets over isolated SVs.

Presented here is a list of the different trainings that were performed, and how they were setup. The label 'NANO' in these trainings corresponds to the use of SV variables that can be found in the NanoAOD data format, with the number designating how many variables were used as input. For all of these trainings the p_T and η were used to calculate weights in bins of 1 GeV in p_T and

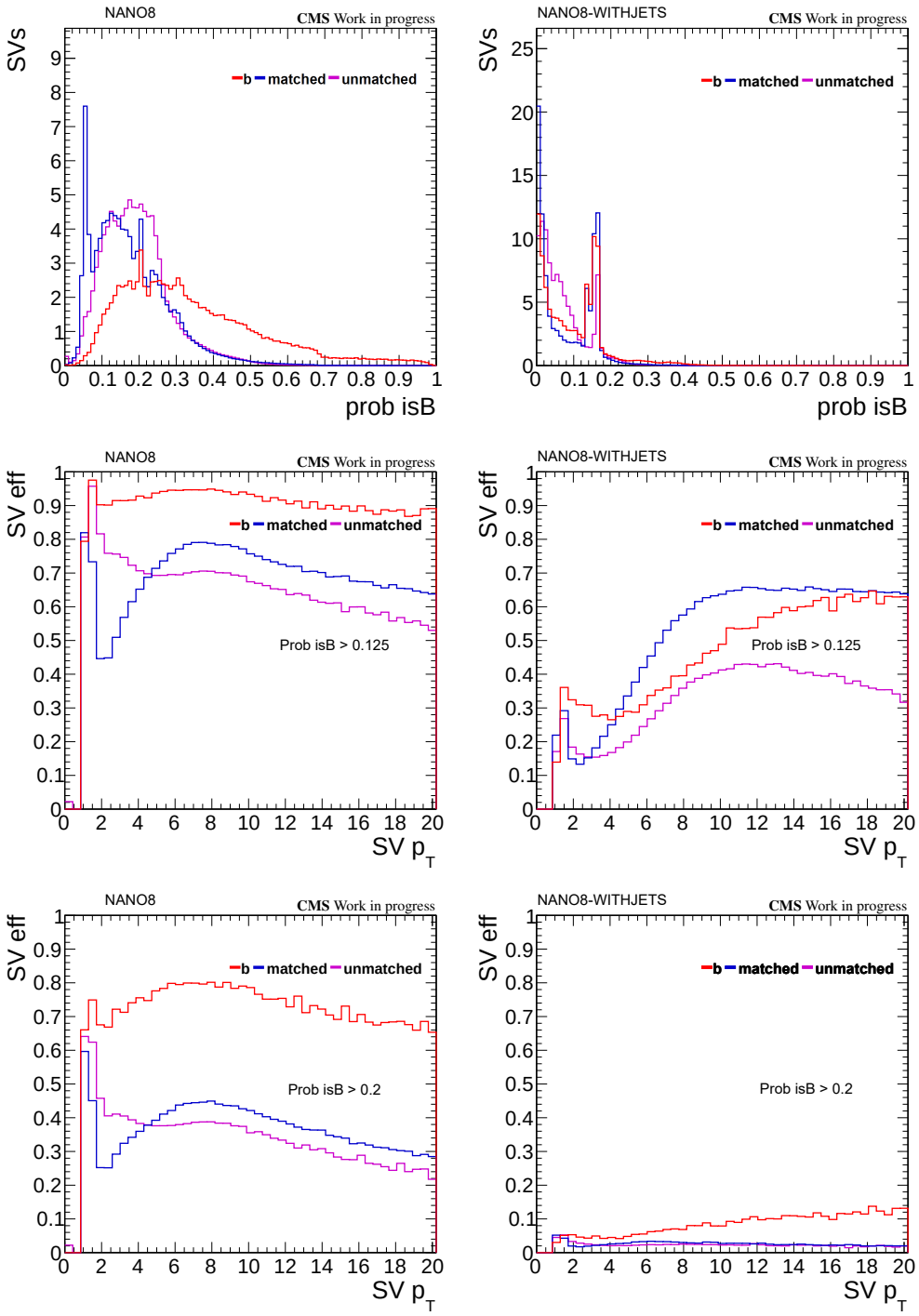


Figure 4.7: TOP: Discriminator distributions from the and NANO-WITHJETS training. MIDDLE: Efficiencies over p_T where the denominator is SVs with the flavor b (red), matched (blue), and unmatched (purple), and the numerator is SVs that pass a probability of being b of 0.125. BOTTOM: Efficiencies over p_T where the denominator is SVs with the flavor b (red), matched (blue), and unmatched (purple), and the numerator is SVs that pass a probability of being b of 0.2. On the left are distributions where SVs matched to jets are not included in any training (NANO8), and the right are distributions where SVs matched to jets are included in the training (NANO8-WITHJETS). The distributions themselves only show the SVs that are not matched to jets, as these are the SVs of interest.

0.5 in η such that the classifiers used in the training are kinematically flattened. In addition, all of the SV variables had a mean normalization and zero padding applied where necessary.

- **4-VAR**: The first training performed, using the 4 variables d_xy, SIP_{3D}, $\cos(\theta_{SV,PV})$, and N_trk that were used in the square cuts.

- **NANO4**: An adjustment of 4-VAR, where the number of tracks is switched with the number of degrees of freedom, as track multiplicity is not stored in the NanoAOD format.

- **NANO6**: An extension of NANO4, where the SV p_T and η are added as inputs in addition to being used for weight calculations, for a total of 6 input variables.

- **NANO7**: An extension of NANO6, where the SV mass is added as an additional input, for a total of 7 input variables.

- **NANO8**: An extension of NANO7, with d_{3D} added as an additional input, for a total of 8 input variables.

- **NANO8-WITHJETS**: An adjustment of NANO8, where SVs matched to jets are included in the training.

- **9-VAR**: An adjustment of NANO8, with N_{dof} changed back to N_{trk}, and with SIP_{xy} added as an additional input, for a total of 9 input variables.

- **NANO8plus**: An extension of NANO8, adding in the cut $SIP_{3D} > 3$, to match the SVs as they are stored in NanoAOD.

- **NANO8plus-JetIso20**: A further extension to NANO8plus, where the jet disambiguation is adjusted such that only SVs matched to jets with $p_T > 20$ GeV are removed. This is the closest match to the SVs used in the analysis, and allows for jets that are not included in b-jet taggers.

4.5.1 Training results

The results of each of the trainings will be given in the form of a plot of the discriminator split into relevant flavors, and then efficiency plots in p_T and η corresponding to a flat discriminator cut of 0.3, which will be a constant cut for each training to show how they change with successive trainings. The efficiencies for multiple different discriminator cuts were looked at, but the value of 0.3 was chosen to compare performances due to it being the value of the cut on the final discriminator that is being used in the analysis. The performance between trainings is compared with side-by-side plots of the efficiencies, along with a ROC curve. The flavor efficiencies that are shown are exclusive of each other, with the denominator being SVs matched to that specific flavor, and the numerator corresponding to those same SVs, which also pass the given discriminator cut.

4.5.1.1 4-VAR Training

The first training was the 4-VAR training. The intent with starting here was to build upon what has been in use by other SUSY analyses, namely the square cuts. The 4-VAR training was setup to match the variables used in the square cuts, d_{xy}, SIP_{3D}, $\cos(\theta_{SV,PV})$, and N_{trk}. The selection of the SVs as inputs was also made to match the square cuts selection up to the point right before the square cuts are actually applied, where only the SVs with a $p_T < 20$ GeV that were not ΔR matched to any jets within a cone size of 0.4 were kept. Figure 4.8 gives the 4-VAR discriminator, separated out into b, matched, and unmatched flavors. Figure 4.9 gives the efficiency plots for those same flavors. Comparing the performance of 4-VAR to the square cuts in Figure 4.10, it can be seen that just by putting the same set of variables in to a simple DNN, the b efficiency is improved, while the $light$ misidentification is maintained at a similar level to the square cuts. This should not be surprising, as the DNN is better able to take advantage of correlations between the variables, that a set of flat cuts would otherwise be unable to account for.

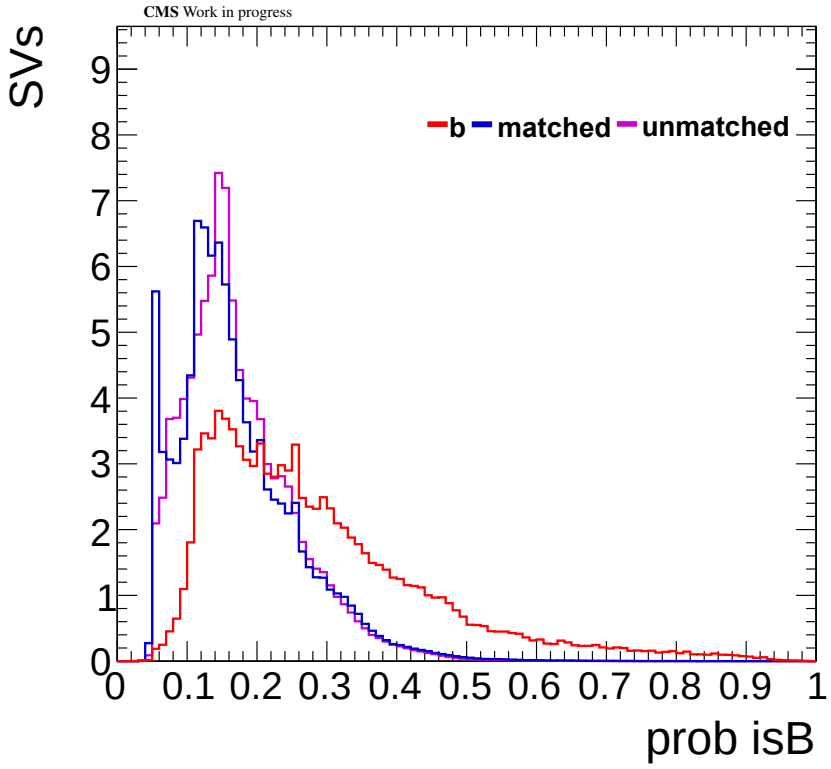


Figure 4.8: Discriminator created using the 4-VAR training, normalized to 100. With this training, one can see the *b* SVs are partially separated from the matched and unmatched SVs, giving some discrimination between the flavors.

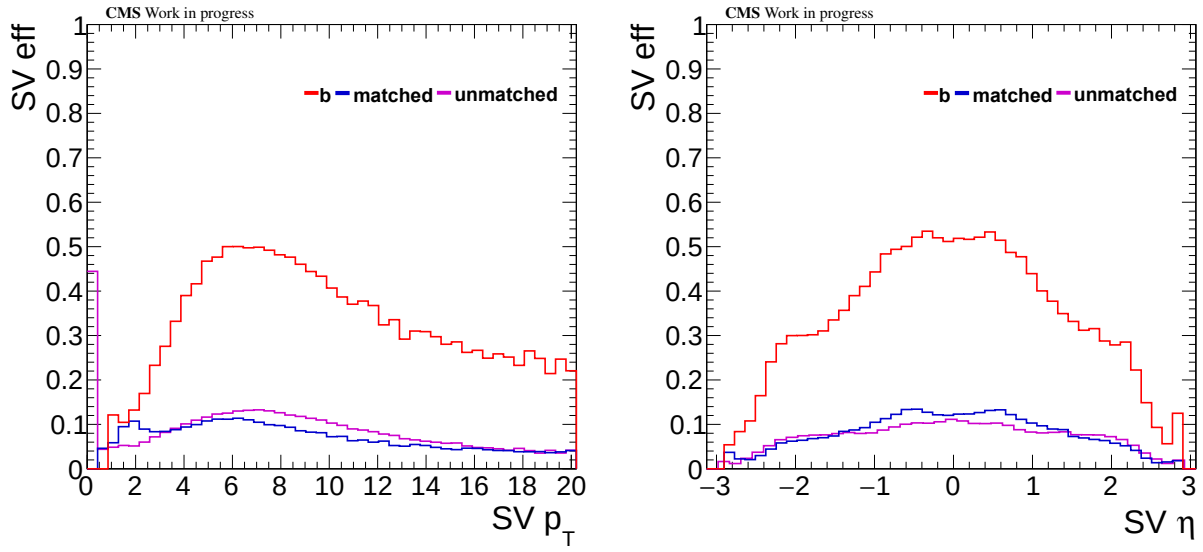


Figure 4.9: Performance efficiencies of the 4-VAR training in p_T (LEFT) and η (RIGHT). The flavors shown are *b* (red), matched (dark blue), and unmatched (magenta).

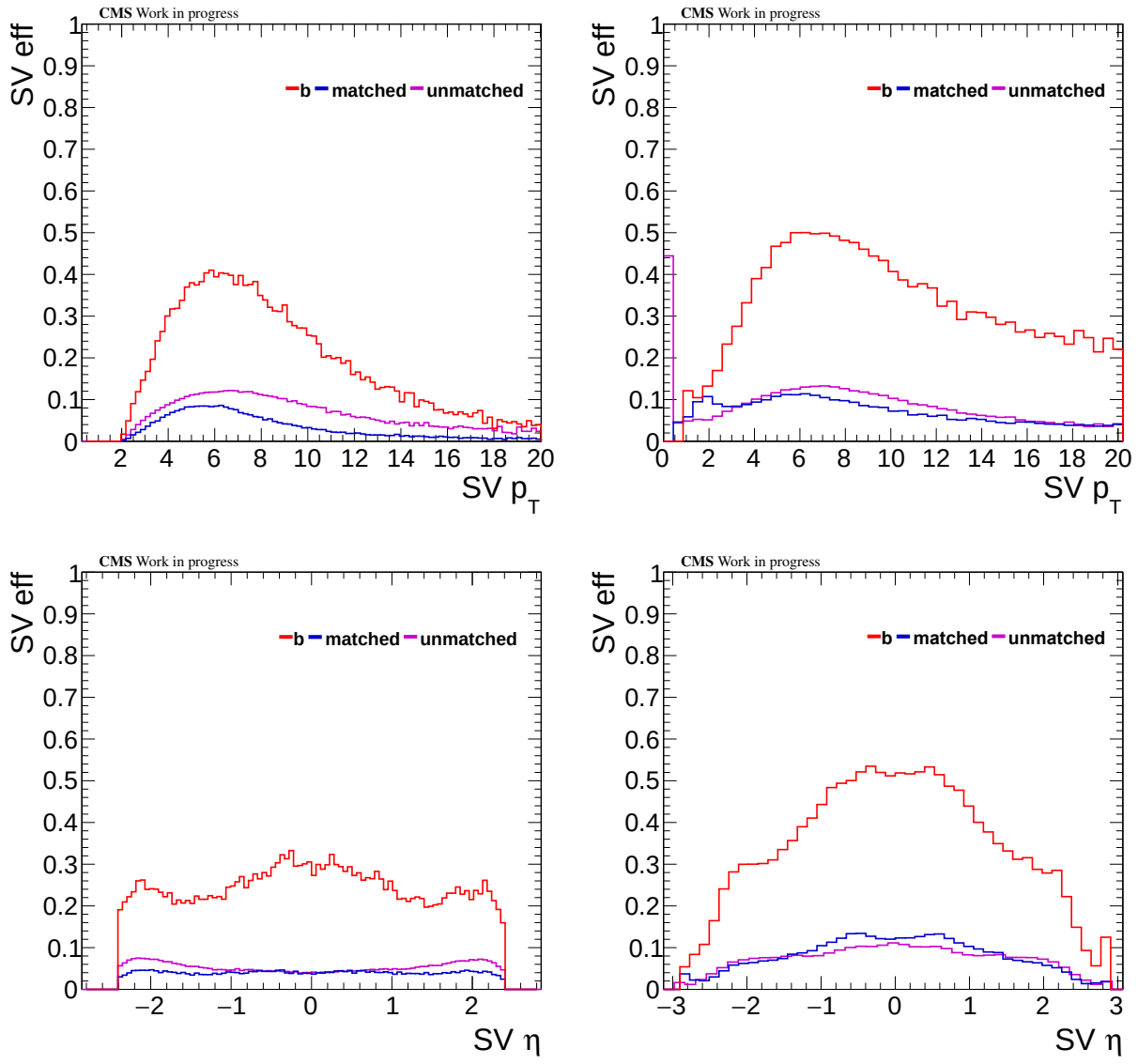


Figure 4.10: Comparison of performance efficiencies over p_T (TOP) and η (BOTTOM) of b, matched and unmatched SVs using the square cuts (LEFT) and 4-VAR training (RIGHT). For the p_T plots, one can see an overall improvement in the b efficiency, especially at higher p_T, ranging from a 5-20% increase in efficiency. The *light* efficiency is roughly unchanged, with the matched efficiency up to 5% higher for the 4-VAR training. The η plots also show an overall improvment in the b efficiency, especially in the central region, with up to a 20% increase in central η and 5-10% at a more forward η. The matched and unmatched efficiencies also saw an increase, but only to the level of 5-10% in the central region, and comparable efficiencies in the forward regions.

4.5.1.2 NANO4 Training

The next training, NANO4, changed the inputs so that it only used variables that were stored in the NanoAOD format. This meant changing the number of tracks to the number of degrees of freedom, due to their close correlation. The SV selection, otherwise, was the same as the 4-VAR training, with the full jet disambiguation applied. Comparing NANO4 to 4-VAR, there is only a small difference in performance, best seen in Figure 4.14, which shows that the NANO4 training does slightly better at low b efficiency values between 0.2 and 0.4. With a discriminator cut of 0.3 applied, shown by the efficiencies seen in Figure 4.13, the cut takes small advantage from the increase in performance, mainly seen in the wings of the η distribution, increasing the b efficiency by upwards of 2%

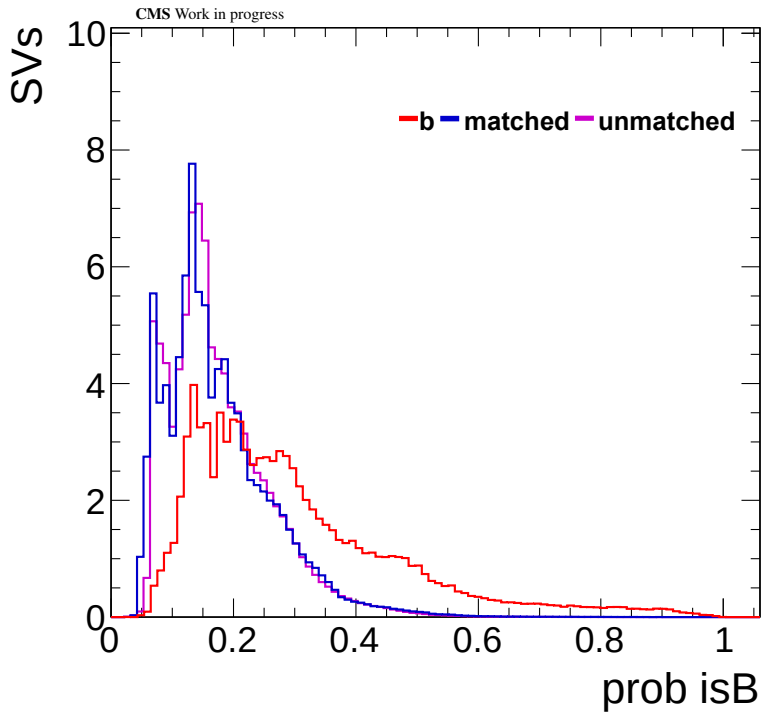


Figure 4.11: Discriminator created using the NANO4 training, normalized to 100.

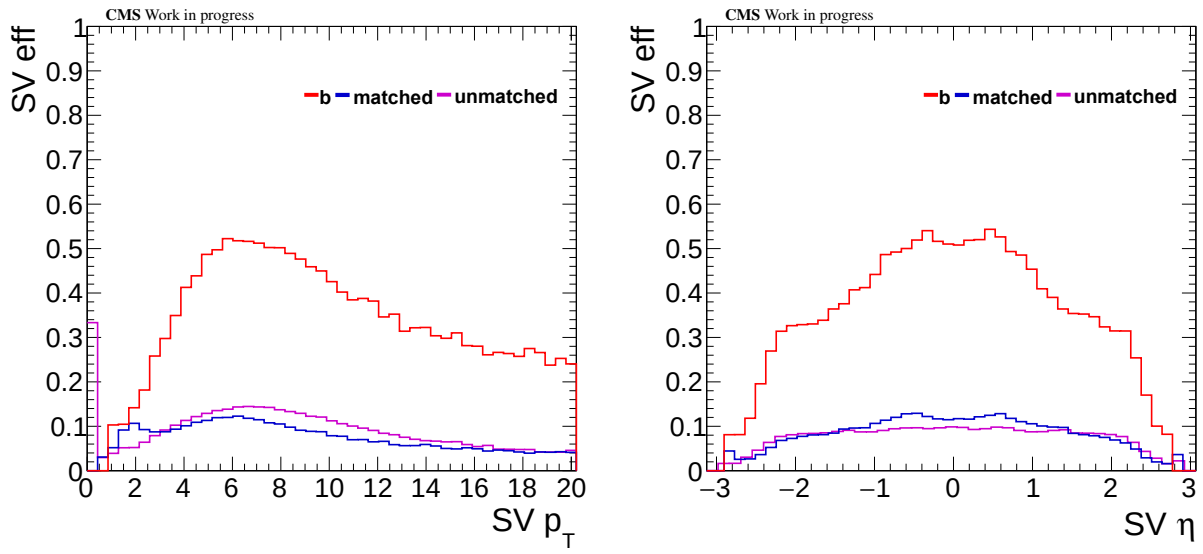


Figure 4.12: Performance efficiencies of the NANO4 training in p_T (LEFT) and η (RIGHT). The flavors shown are *b* (red), matched (dark blue), and unmatched (magenta).

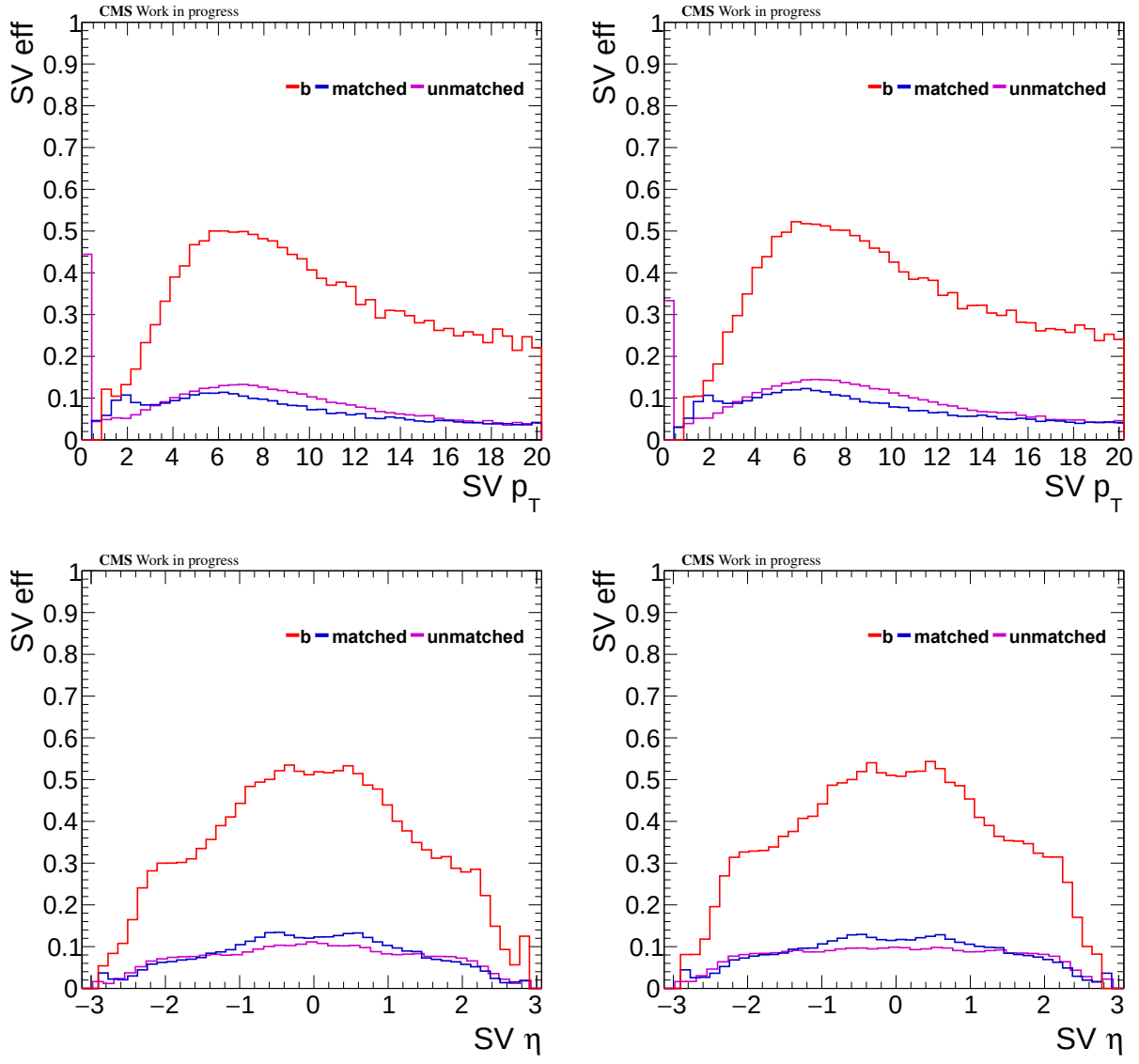


Figure 4.13: Comparison of performance efficiencies over p_T (TOP) and η (BOTTOM) of b, matched and unmatched SVs using the 4-VAR (LEFT) and NANO4 trainings (RIGHT). These trainings show small difference in performance between the two. NANO4 trainings show a slightly better performance related to the b-efficiency, while maintaining a similar *light* efficiency.

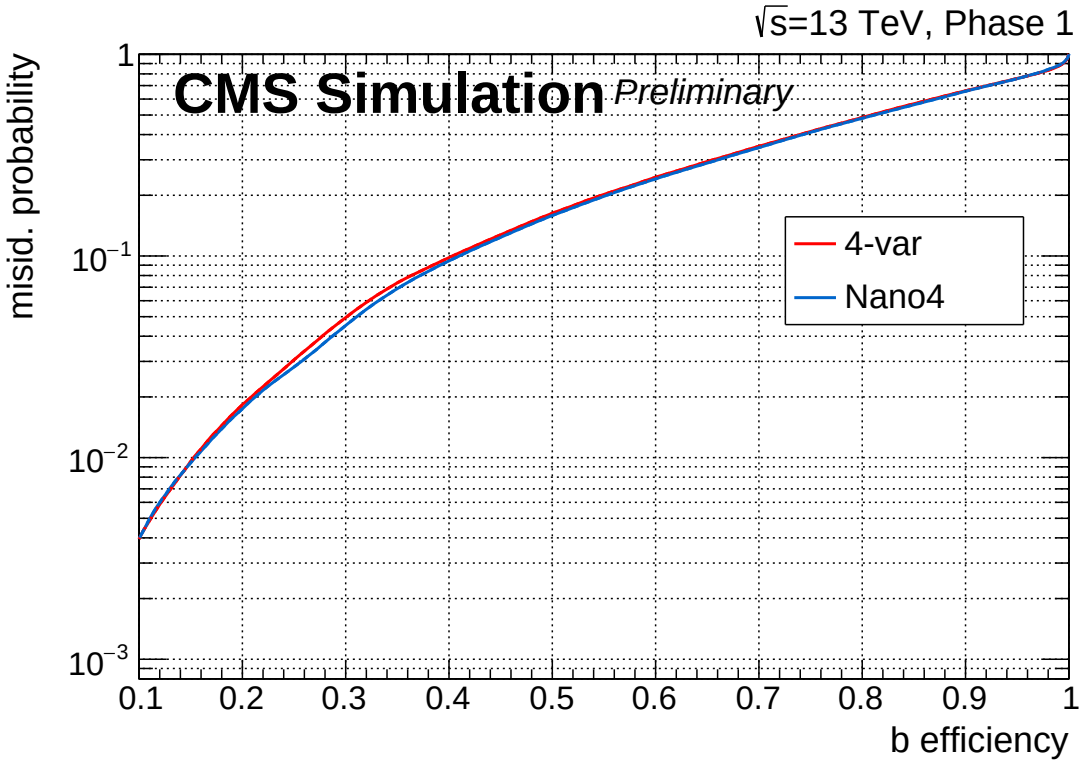


Figure 4.14: ROC curve comparing the *b* efficiency to *light* misidentification probability for the 4-VAR and NANO4 trainings. The curve shows a mild improvement in NANO4 over 4-VAR

4.5.1.3 NANO6 Training

The NANO6 training adds the p_T and η of the SV as inputs into the training, on top of their use to calculate weights. This particular training shows one of the more significant improvements in performance, as seen in Figures 4.17 and 4.18. Looking at the efficiencies in Figure 4.17, improvements can can be seen in the b efficiency at low and high p_T, and at the more forward η. For $p_T > 10$ GeV and $p_T < 5$ GeV, the b efficiency has increased by upwards of 10%, with a similar increase for $|\eta| > 1$.

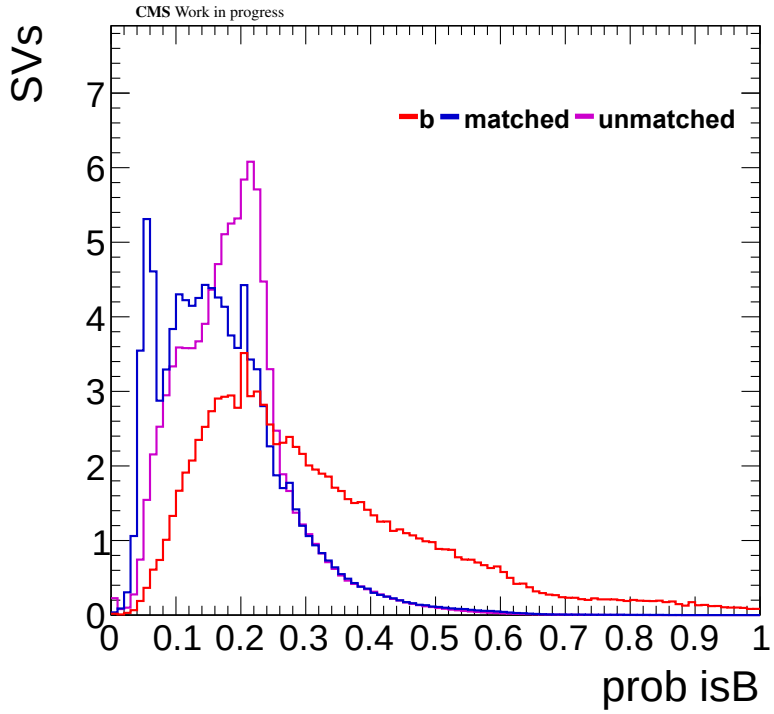


Figure 4.15: Discriminator created using the NANO6 training, normalized to 100.

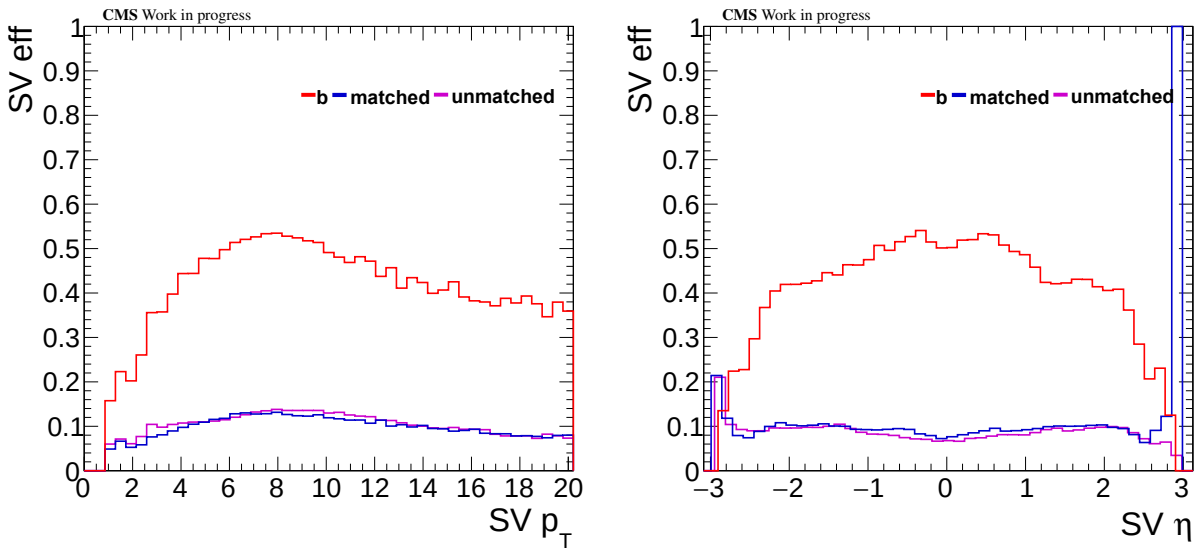


Figure 4.16: Performance efficiencies of the NANO6 training in p_T (LEFT) and η (RIGHT). The flavors shown are *b* (red), matched (dark blue), and unmatched (magenta).

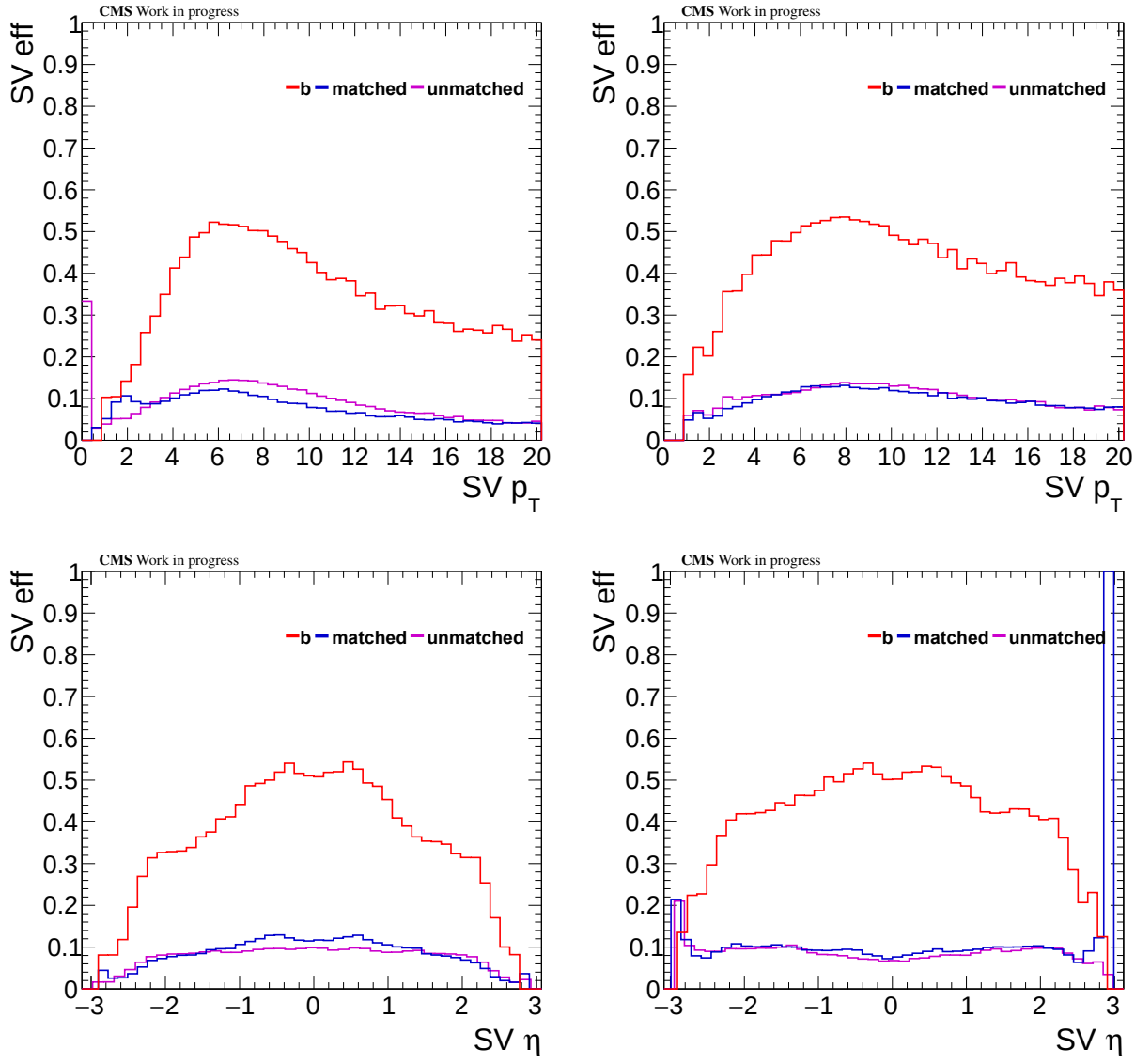


Figure 4.17: Comparison of performance efficiencies over p_T (TOP) and η (BOTTOM) of b, matched and unmatched SVs using the NANO4 (LEFT) and NANO6 trainings (RIGHT).

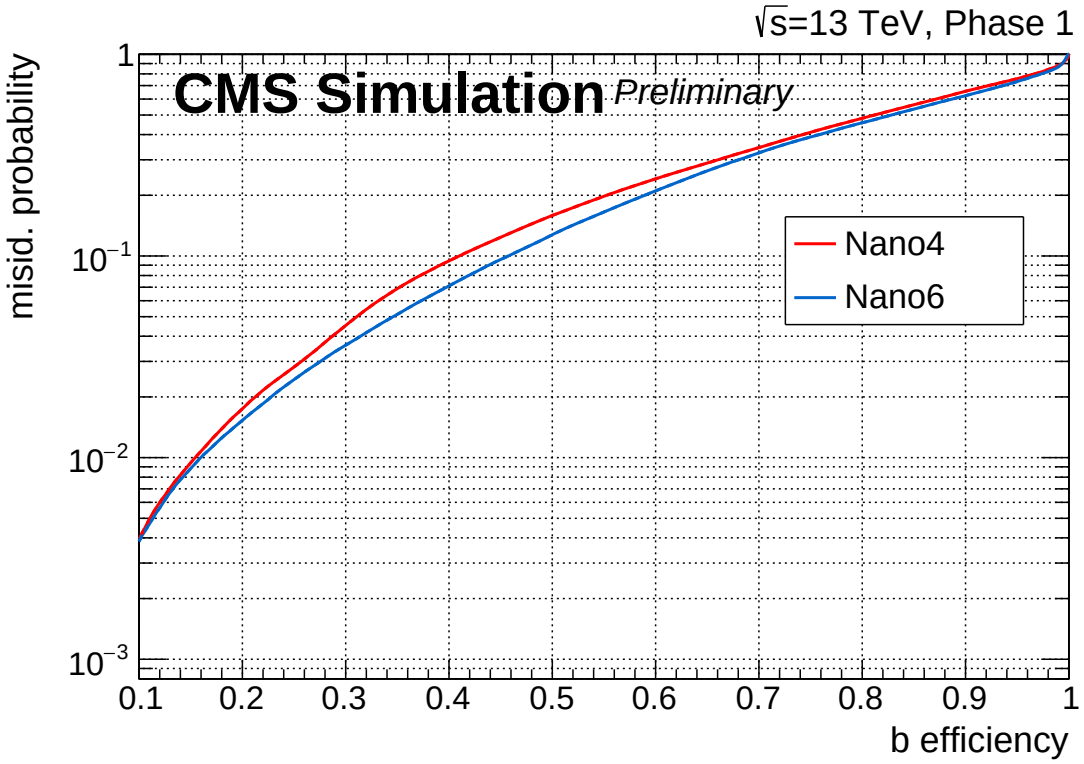


Figure 4.18: ROC curve comparing the *b* efficiency to *light* misidentification probability for the NANO4 and NANO6 trainings.

4.5.1.4 NANO7 Training

The NANO7 training added the SV mass as an additional input. Compared to the NANO6 training, there is overall, a small increase in performance, as can be seen in Figure 4.21. While the shapes of the efficiencies do not change significantly, a more pronounced feature can be seen in the discriminator plot in Figure 4.19, at values of $ProbB < 0.1$. This feature, which also appears at a smaller scale in the previous discriminator versions, is found to be due to the SVs matched to strange quarks, where the mass of the SVs has a small mass peak, corresponding to the mass of a Kaon. There is an overall improvement in the performance over the NANO6 training, as seen in the roc curve in Figure 4.22. This performance increase, while small, can be seen over the full distribution of *b* efficiency.

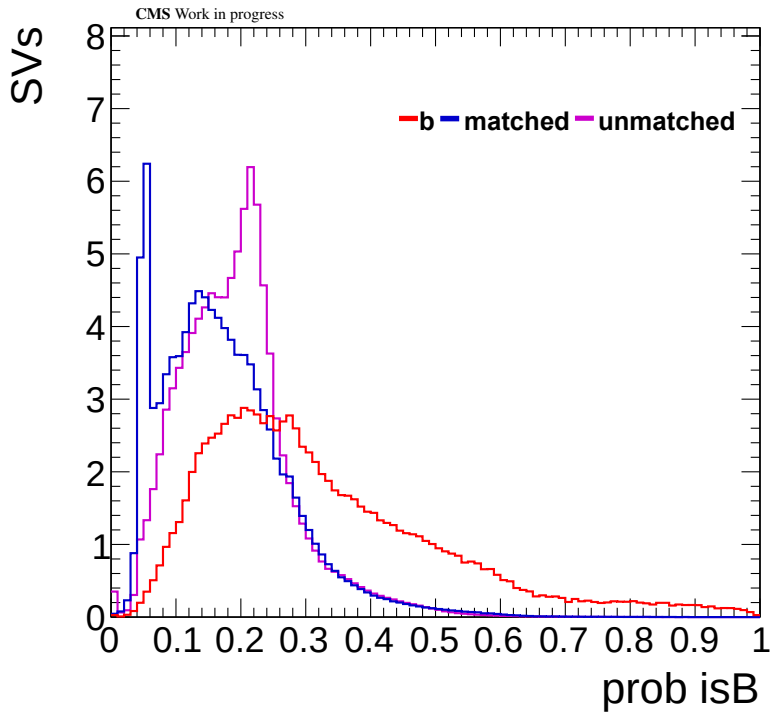


Figure 4.19: Discriminator created using the NANO7 training, normalized to 100.

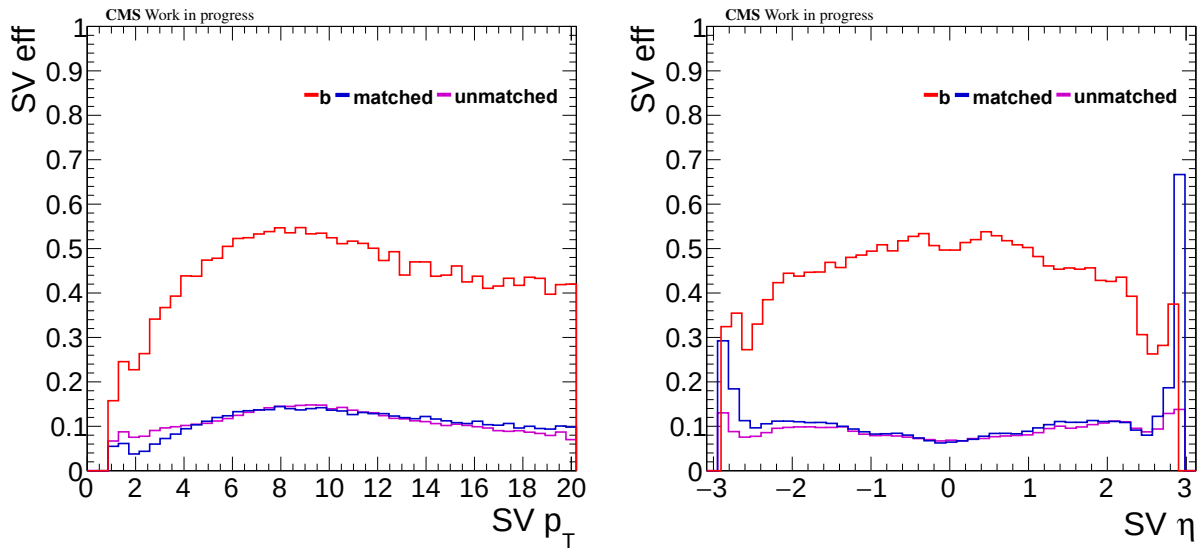


Figure 4.20: Performance efficiencies of the NANO6 training in p_T (LEFT) and η (RIGHT). The flavors shown are *b* (red), matched (dark blue), and unmatched (magenta).

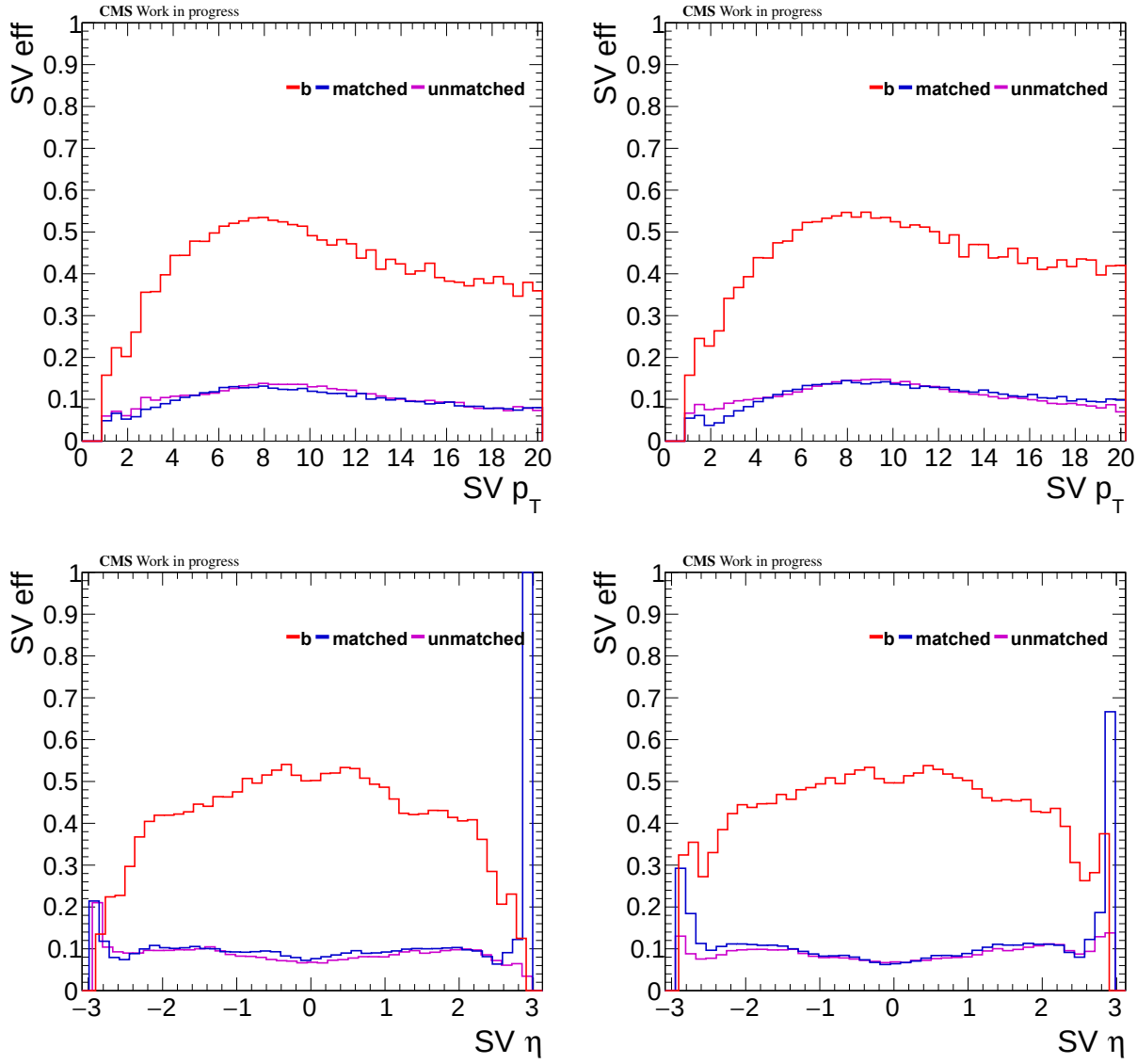


Figure 4.21: Comparison of performance efficiencies over p_T (TOP) and η (BOTTOM) of b, matched and unmatched SVs using the NANO6 (LEFT) and NANO7 trainings (RIGHT).

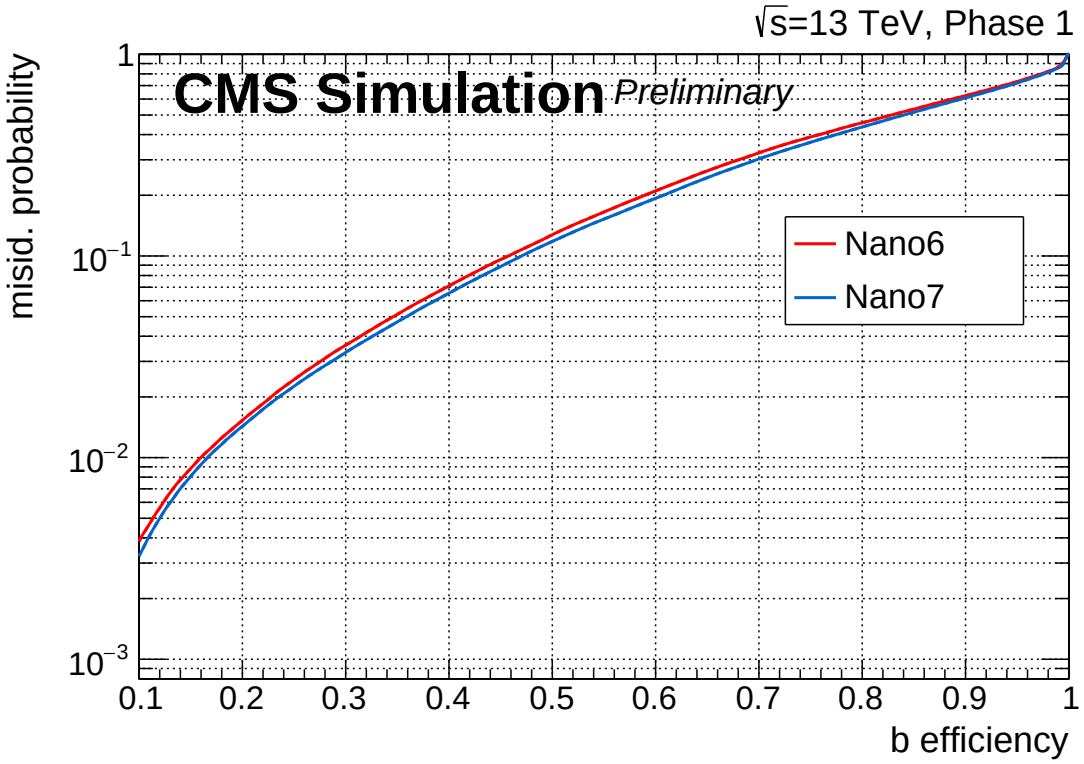


Figure 4.22: ROC curve comparing the *b* efficiency to *light* misidentification probability for the NANO6 and NANO7 trainings

4.5.1.5 NANO8 Training

The NANO8 training adds the 3D flight distance in addition to the 3D flight distance significance, which has been included in the trainings since the beginning. This version improves the overall performance, but by a slightly lesser amount over moving from NANO6 to NANO7. In Figure 4.25, one can see an increase in the b efficiencies over the whole p_T range, and in addition, the b efficiency over η has flattened.

Looking at the overall increase in performance between the NANO4 and NANO8 training, Figure 4.27 shows significant improvement in the ROC curve, with up to a 30% reduction in misidentification probability at a b efficiency of 0.4.

Originally this training was to be the final version of the discriminator, but comparisons of efficiencies showed vastly different shapes, depending on whether the efficiency was made using the data samples that were used as inputs into the training, or using the samples that were used directly by the SUSY analysis. After determining the cause of this difference in efficiency, two extra trainings were done, adjusting the SV content that was used as input into the machine learning algorithm. The first of these trainings, NANO8plus, was to account for the difference in the samples, and the second, NANO8plus-JetIso20, was to match the SV selection of the machine learning inputs to the selection of SVs used in the analysis.

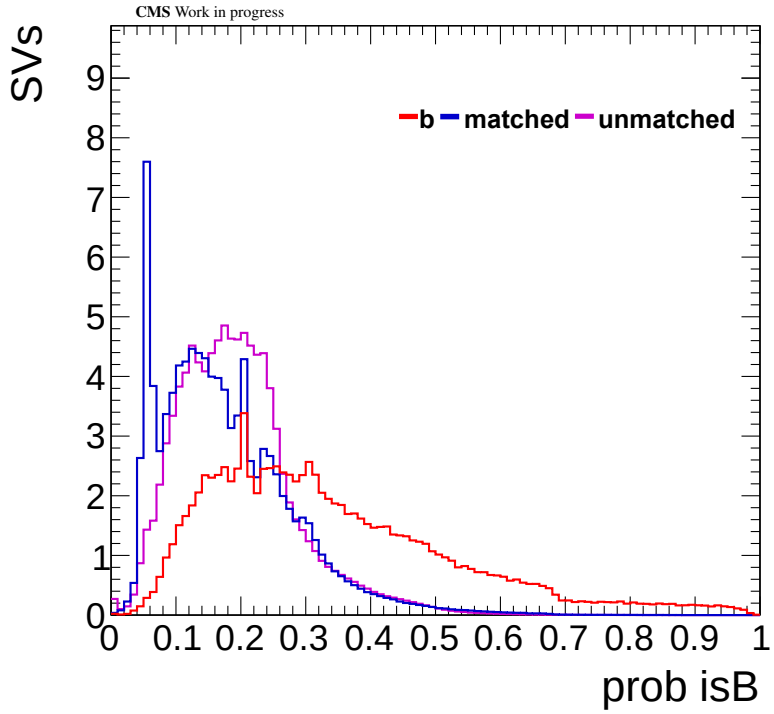


Figure 4.23: Discriminator created using the NANO8 training, normalized to 100.

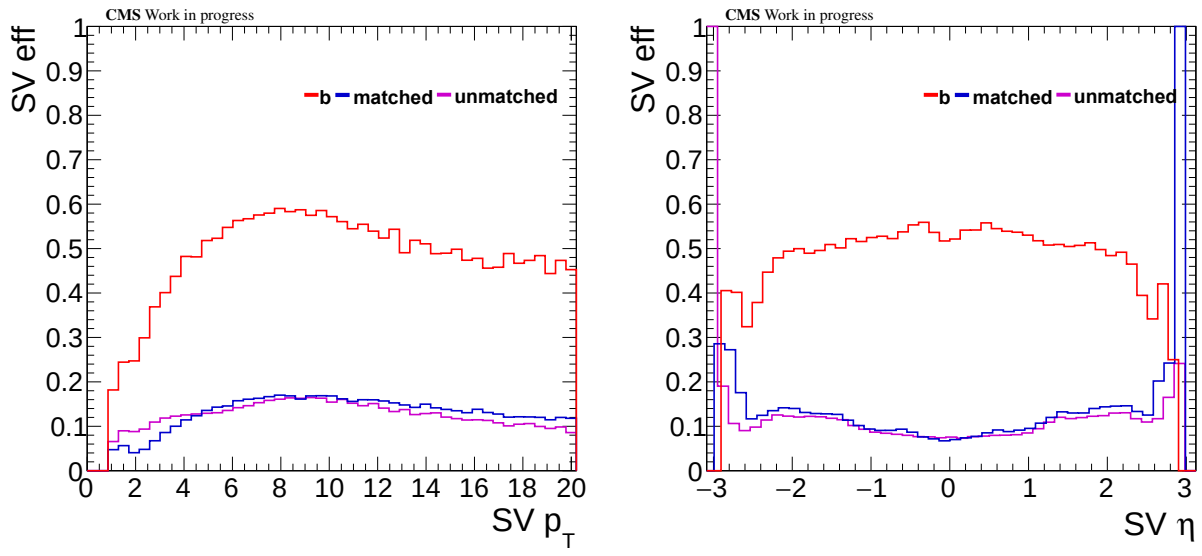


Figure 4.24: Performance efficiencies of the NANO8 training in p_T (LEFT) and η (RIGHT). The flavors shown are b (red), matched (dark blue), and unmatched (magenta).

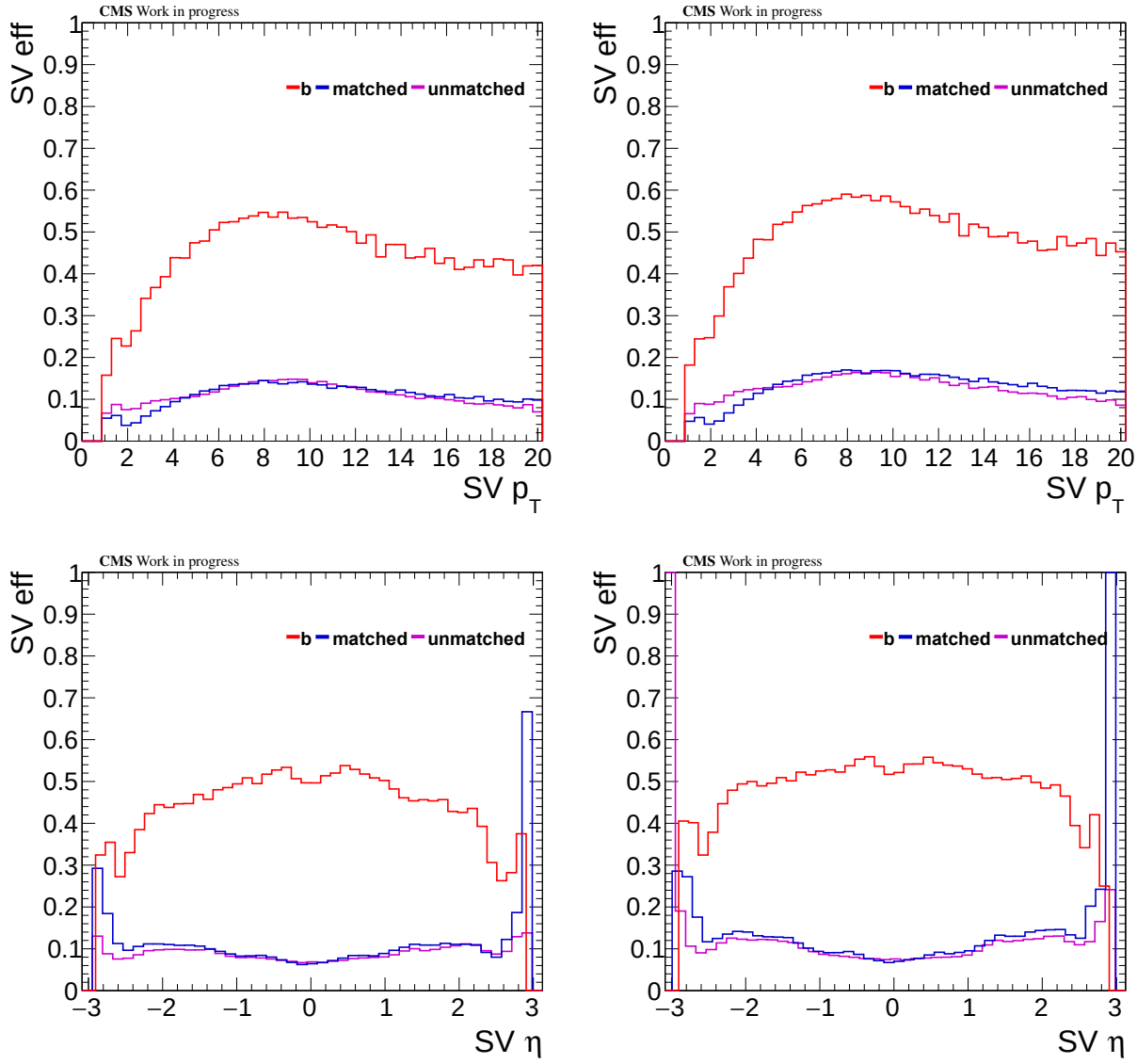


Figure 4.25: Comparison of performance efficiencies over p_T (TOP) and η (BOTTOM) of b, matched and unmatched SVs using the NANO7 (LEFT) and NANO8 trainings (RIGHT).

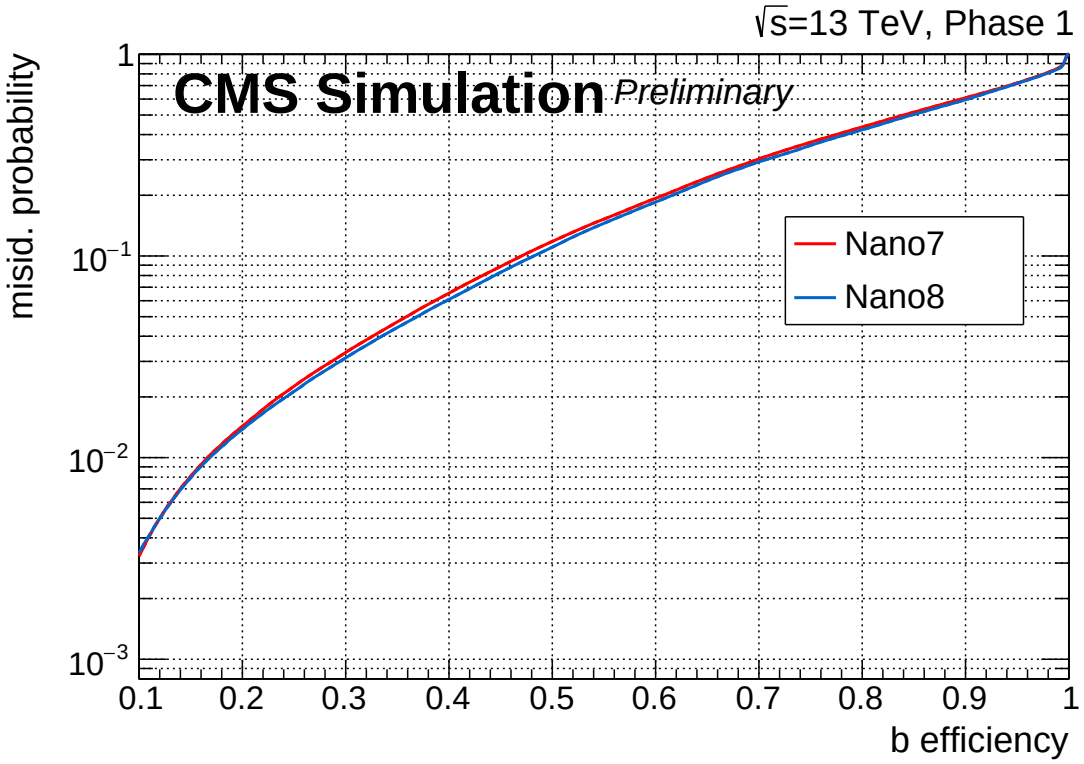


Figure 4.26: ROC curve comparing the *b* efficiency to *light* misidentification probability for the NANO7 and NANO8 trainings

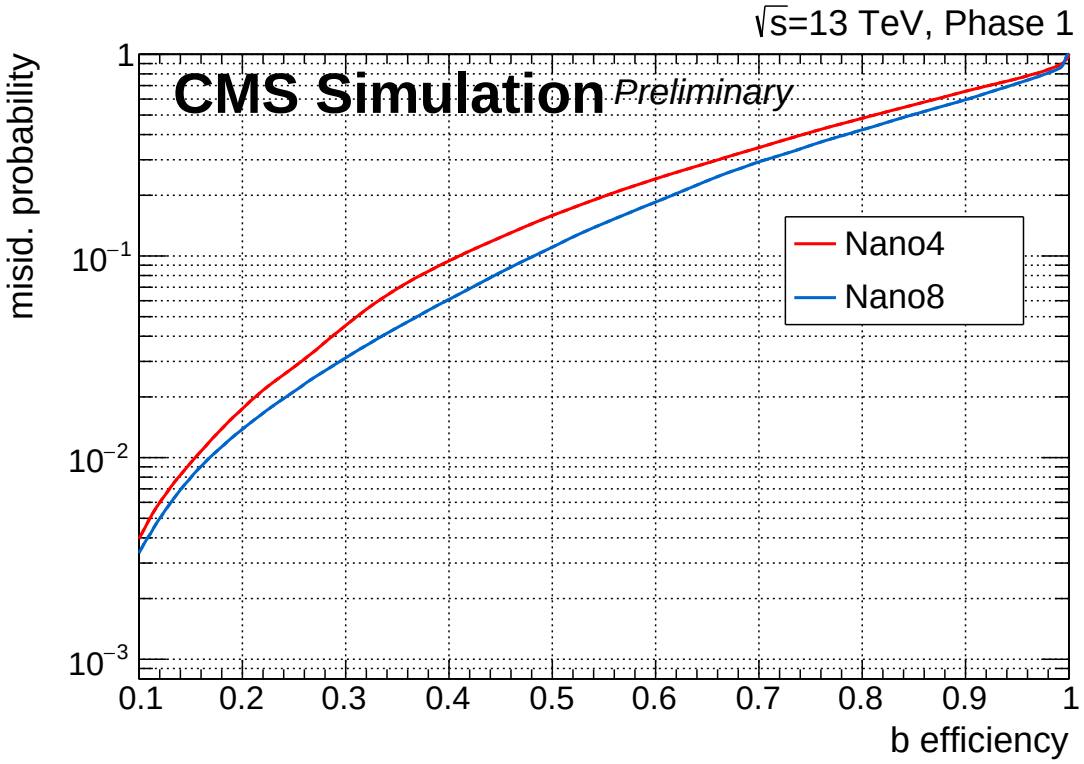


Figure 4.27: ROC curve comparing the *b* efficiency to *light* misidentification probability for the NANO4 and NANO8 trainings

4.5.1.6 NANO8-WITHJETS Training

The purpose of the NANO8-WITHJETS training, was to look at the effect of adding in SVs that are matched to any jet as inputs into the training. The performance plots in Figures 4.28, and 4.29 show how this training performs, using the same testing dataset as the NANO8 version. This means that while the SVs used as input in the training included all SVs, even those matched to jets, the SVs used for testing had the jet disambiguation present. The aforementioned figures show a very poor efficiency at properly tagging isolated SVs. This effect is due to the much larger number of SVs present in jets, than isolated. The features found in the larger population of data, such as SVs matched to jets, are able to dominate in the training, causing isolated SVs to perform much worse, as their features do not necessarily match the features of SVs in higher energy jets. This training highlights the importance of adjusting inputs to match the population of what the tagger is going to be used for. The results of this particular training was a contributing factor in the decision of which SVs to include in the NANO8plus, and NANO8plus-JetIso20 trainings.

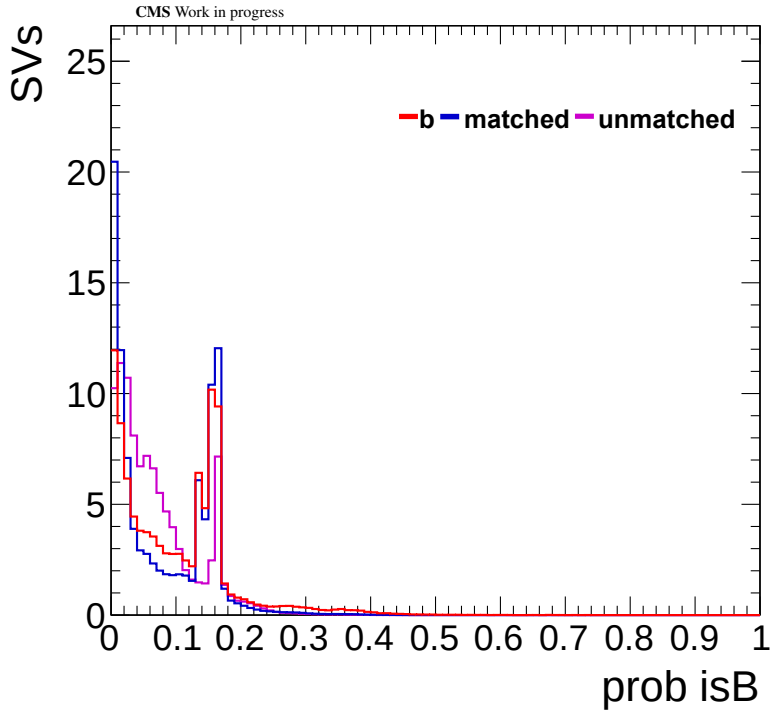


Figure 4.28: Discriminator created using the NANO8-WITHJETS training, normalized to 100.

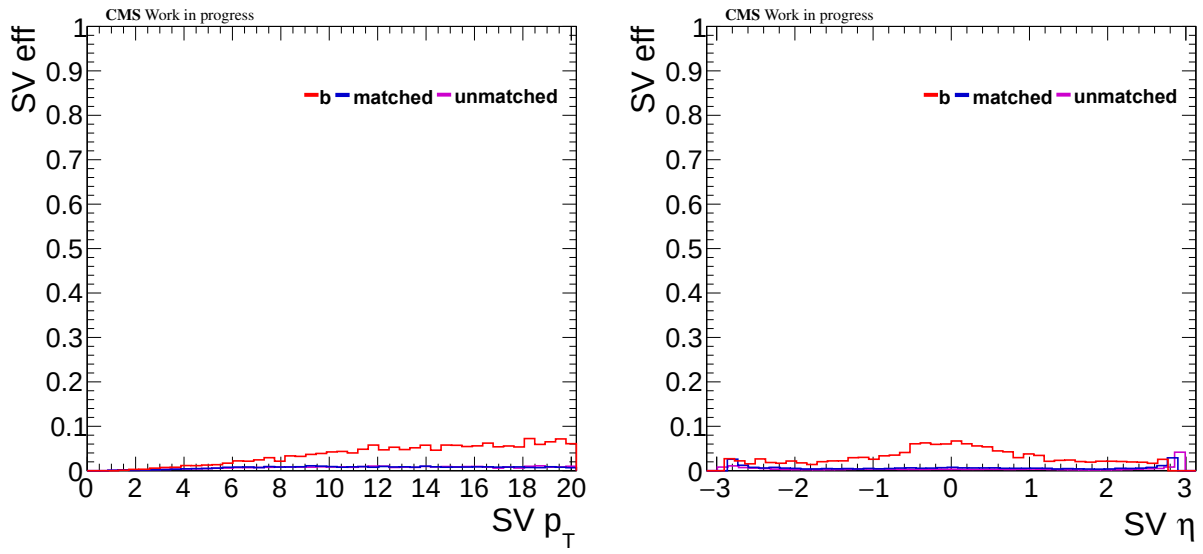


Figure 4.29: Performance efficiencies of the NANO8-WITHJETS training in p_T (LEFT) and η (RIGHT). The flavors shown are *b* (red), matched (dark blue), and unmatched (magenta).

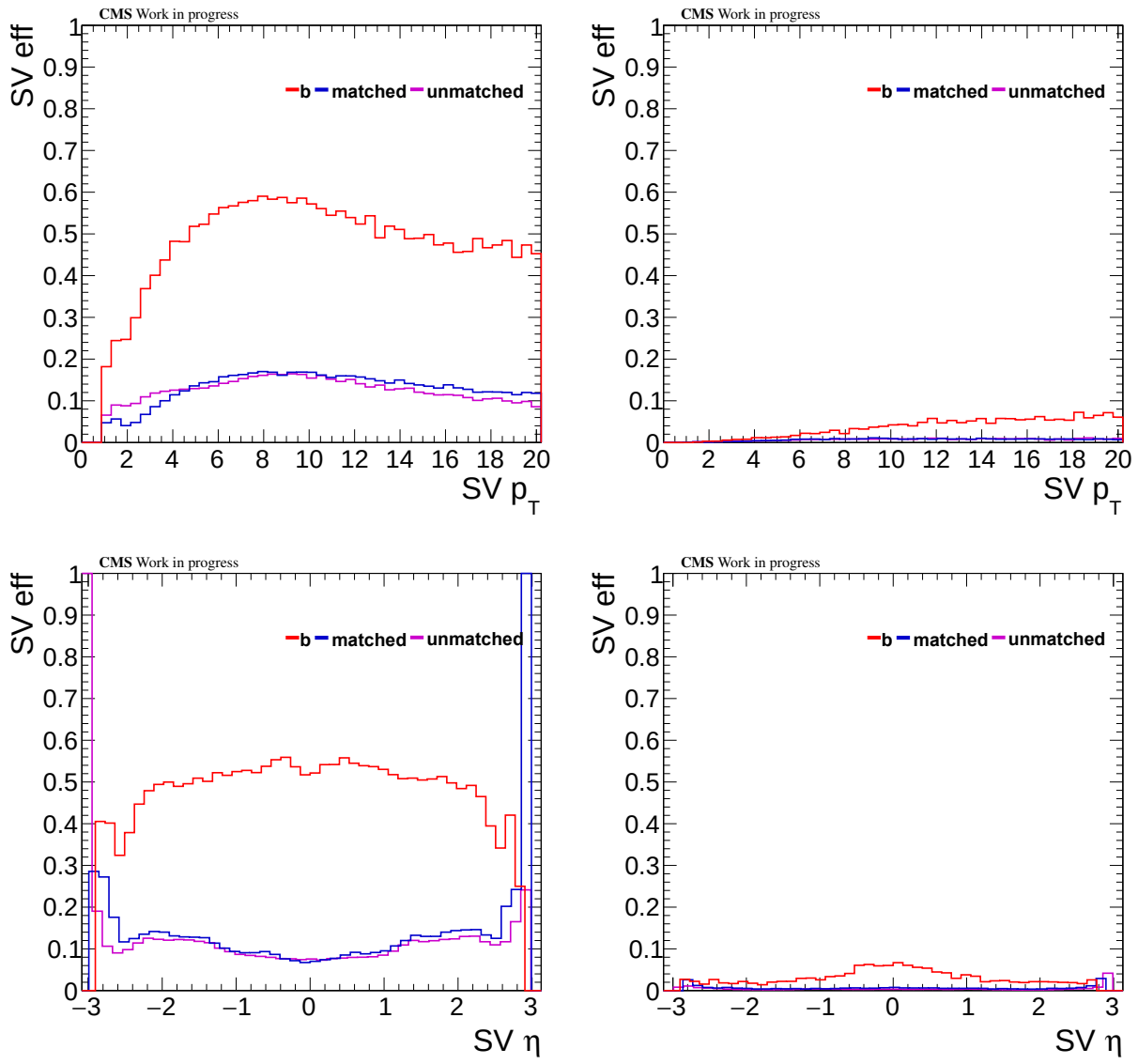


Figure 4.30: Comparison of performance efficiencies over p_T (TOP) and η (BOTTOM) of b, matched and unmatched SVs using the NANO8 (LEFT) and NANO8-WITHJETS trainings (RIGHT).

4.5.1.7 9-VAR Training

The 9-VAR training was made to use all of the variables present in the NANO8 training, but changing the ndof back to number of tracks and adding, in addition, the 2D decay length significance. The main purpose was to see if the results seen in the NANO4 versus the 4-VAR trainings is maintained when compared to the NANO8 version. The 2D decay length significance was added to see if it would, in addition, affect the performance.

Looking at the ROC curve in Figure 4.34, one can see that there is no improvement going from NANO8 to 9-VAR. In fact, the NANO8 training performs slightly better than the 9-VAR training, even though it has an additional variable compared to the NANO8 training. The difference in performance between the NANO8 and 9-VAR trainings is comparable to the difference in the NANO4 and 4-VAR trainings seen in Figure 4.14. From this it can be inferred that the 2D decay length significance will not add any significant increase in performance, and thus is not included in the NANO8 training, and there is not any loss in performance from remaining with N_{dof} over number of tracks.

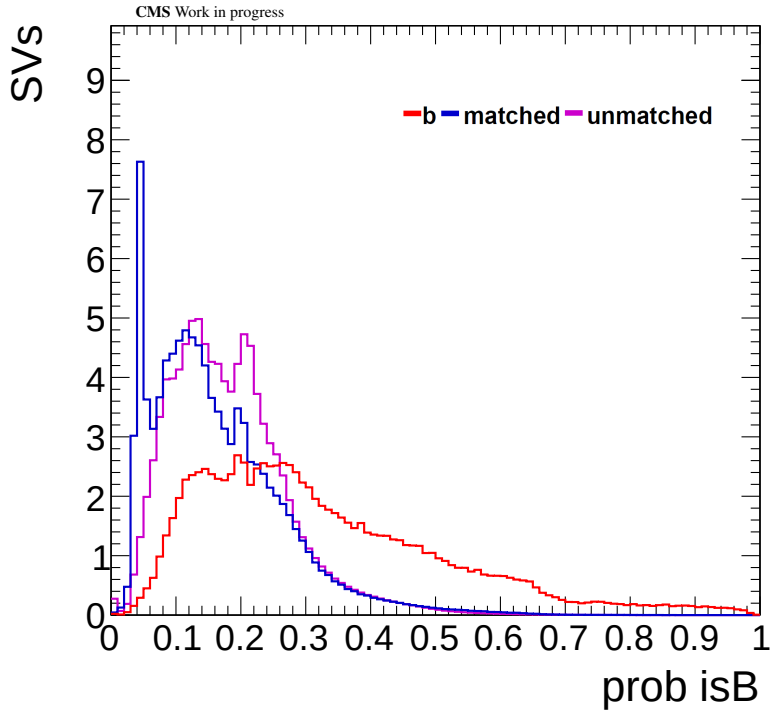


Figure 4.31: Discriminator created using the 9-VAR training, normalized to 100.

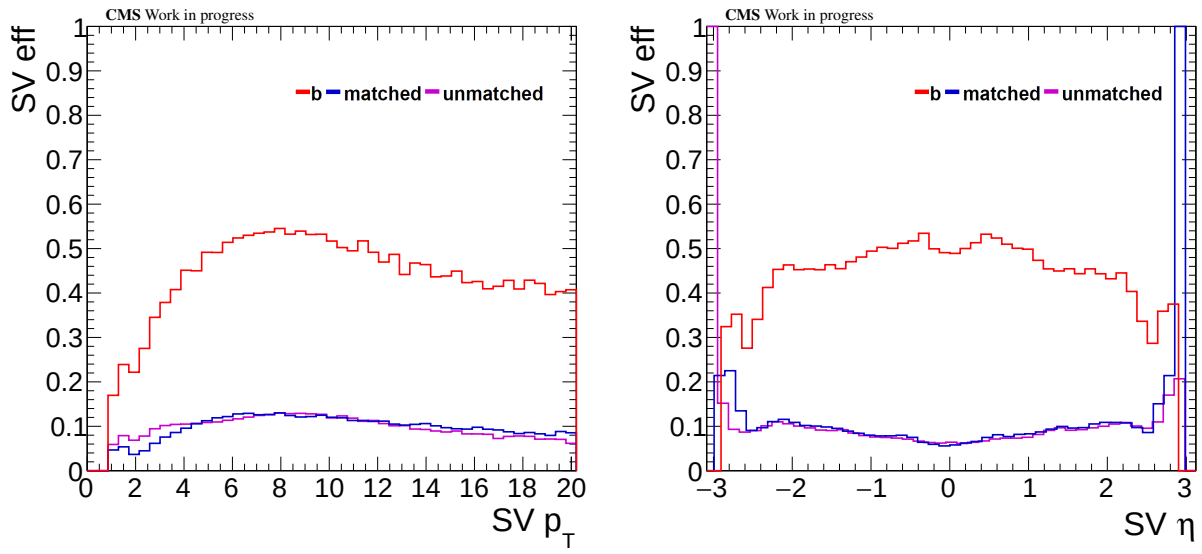


Figure 4.32: Performance efficiencies of the 9-VAR training in p_T (LEFT) and η (RIGHT). The flavors shown are *b* (red), matched (dark blue), and unmatched (magenta).

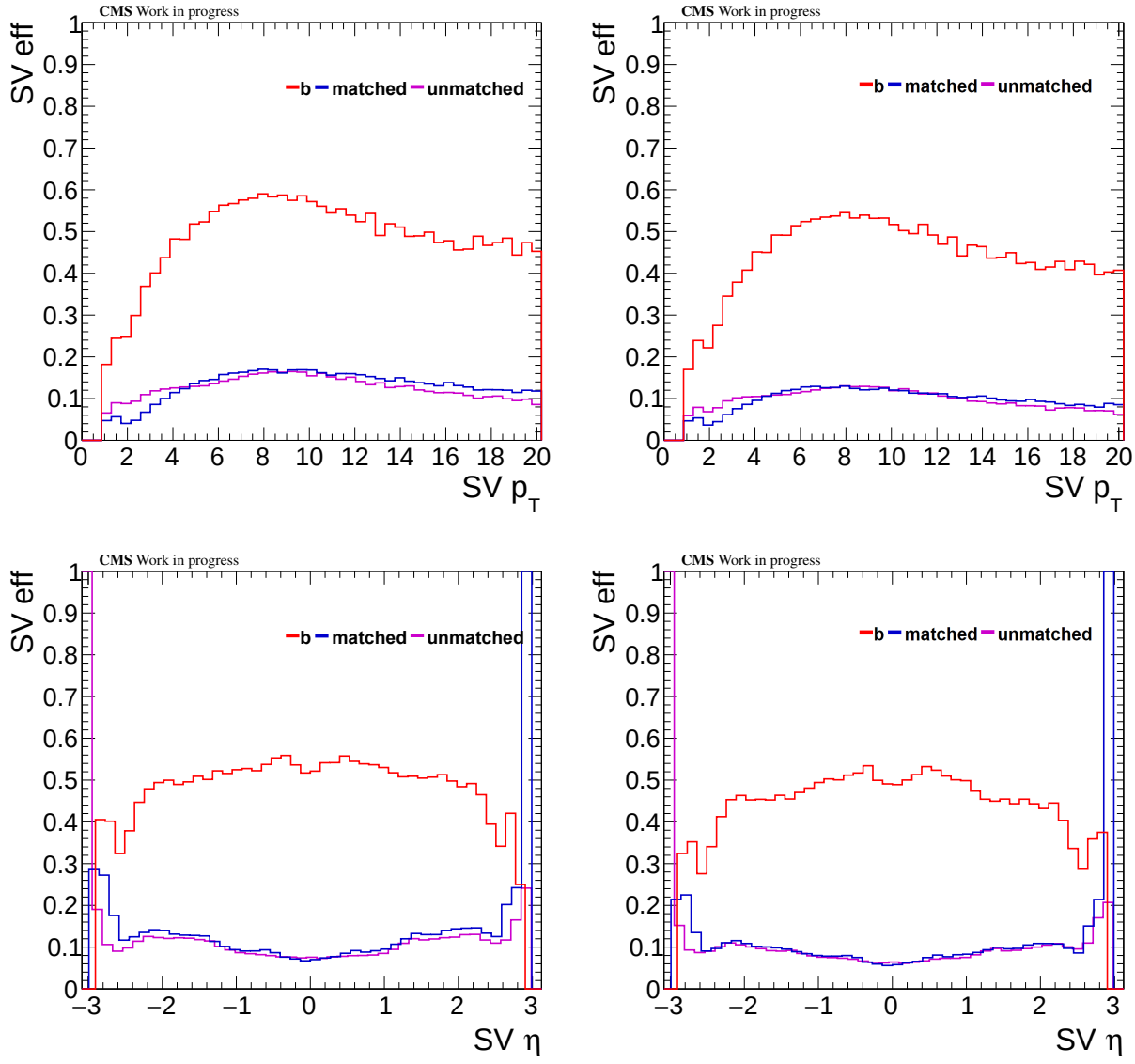


Figure 4.33: Comparison of performance efficiencies over p_T (TOP) and η (BOTTOM) of b, matched and unmatched SVs using the NANO8 (LEFT) and 9-VAR trainings (RIGHT).

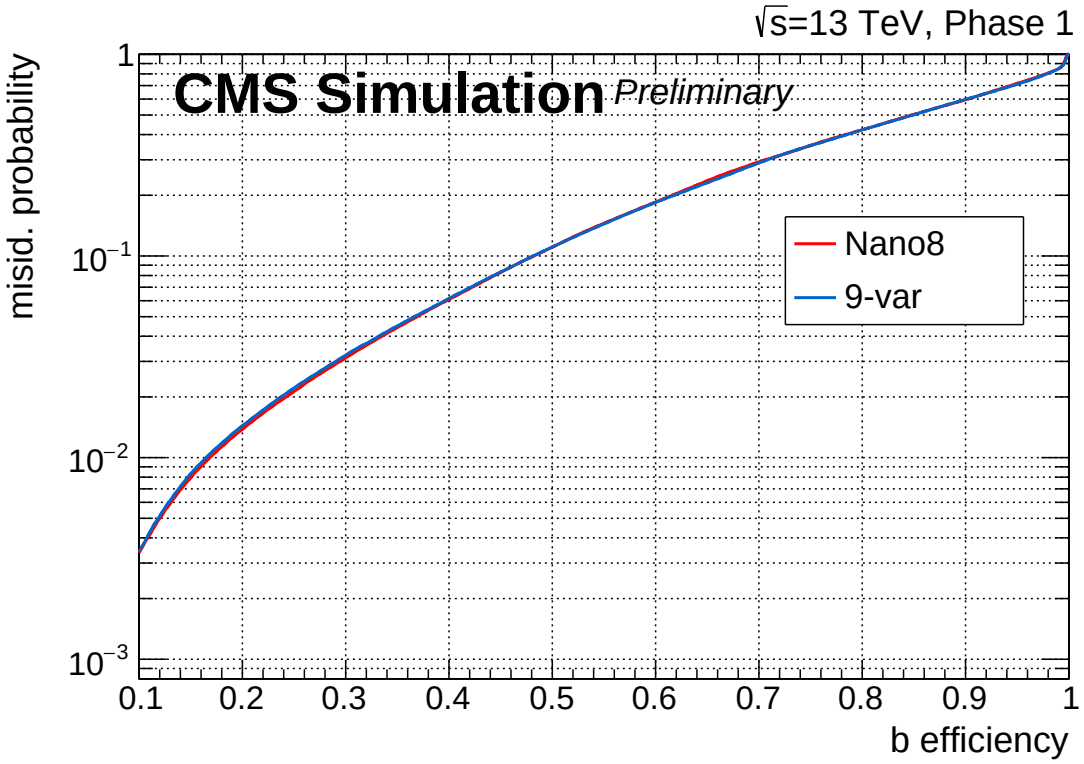


Figure 4.34: ROC curve comparing the *b* efficiency to *light* misidentification probability for the NANO8 and 9-VAR trainings

4.5.1.8 NANO8plus Training

The NANO8plus training involved adjusting the SVs used as input into the training and testing, over adding additional variables. Since the analysis is using the nanoAOD format to process the data and MC, it is desirable to have the inputs going into the training to match the selection that is present in nanoAOD. It was found that the SVs stored in nanoAOD have an additional cut than the SVs used in the previous trainings up to NANO8 did not have, $SIP3D > 3$. The NANO8plus training adds this additional cut to the SVs used as input in order to determine whether the performance is significantly affected. As seen in Figure 4.37, this difference in SV population greatly changes the overall efficiency over both p_T and η.

An interesting factor in this is that the performance of the NANO8 and NANO8plus trainings are nearly identical when both trainings are tested on the same dataset, as seen in the ROC curve in Figure 4.38. This shows that, while the existence of the removed SVs do not affect results when they are included as inputs into the trainings, the opposite is true when including the SVs in the testing. One can then assume that the SVs with $SIP3D <= 3$ do not have very much discriminating power, and since there are so many of them, they bias the efficiency results, showing a good ratio of *b* efficiency to *light* flavor rejection, compared to the true population of SVs that are of interest to the analysis.

With these results, it became more important to make sure that the SV selection in the training and testing, matches the SV selection that is in the analysis, in order to make sure the performance being looked at is the performance that should be expected in the analysis. Going towards this, the final discriminator, from the NANO8plus-JetIso20 training, was made to use SVs that match the jet isolation put on SVs in the analysis.

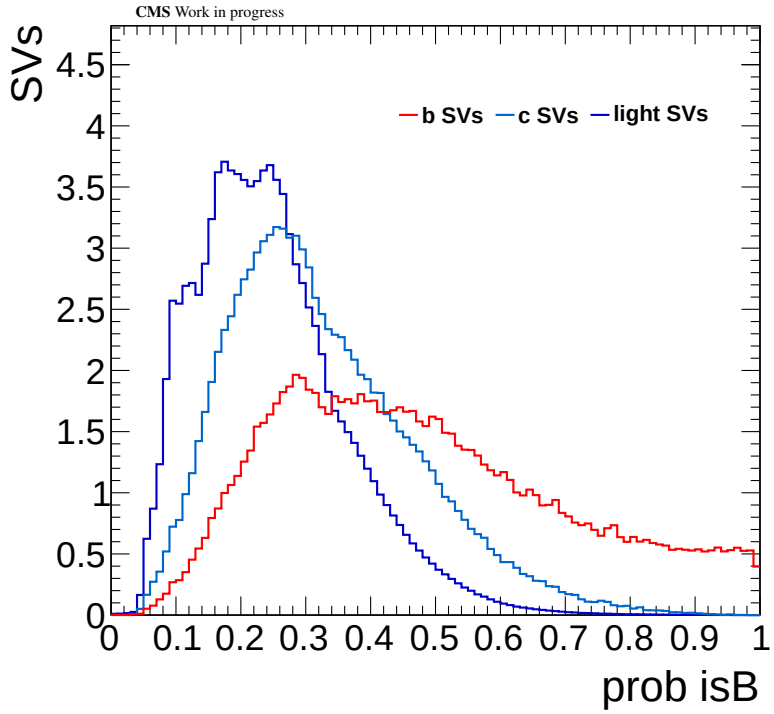


Figure 4.35: Discriminator created using the NANO8plus training, normalized to 100.

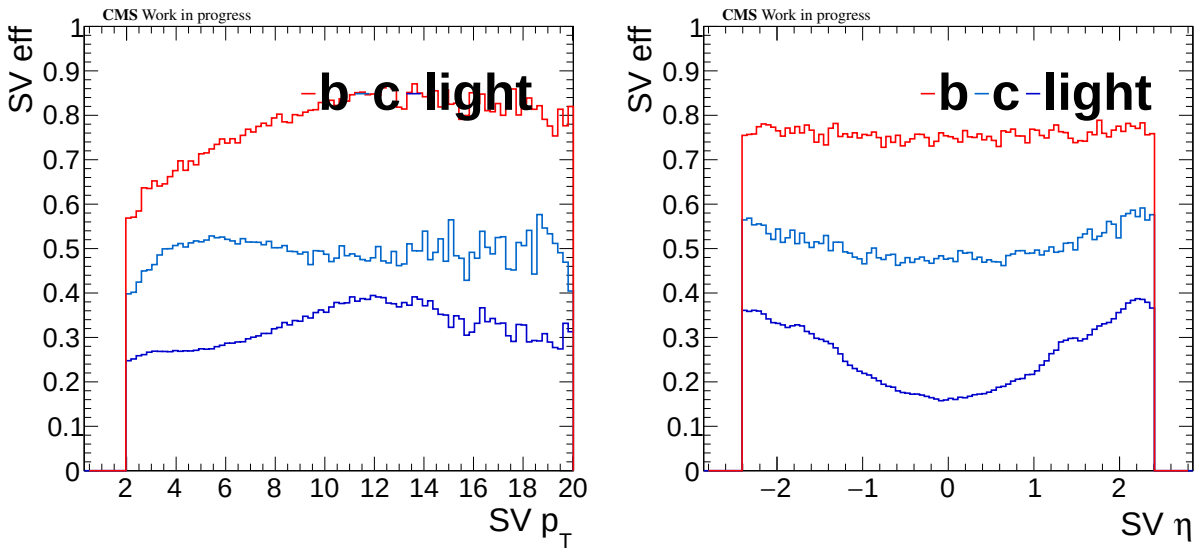


Figure 4.36: Performance efficiencies of the NANO8plus training in p_T (LEFT) and η (RIGHT). The flavors shown are b (red), matched (dark blue), and unmatched (magenta).

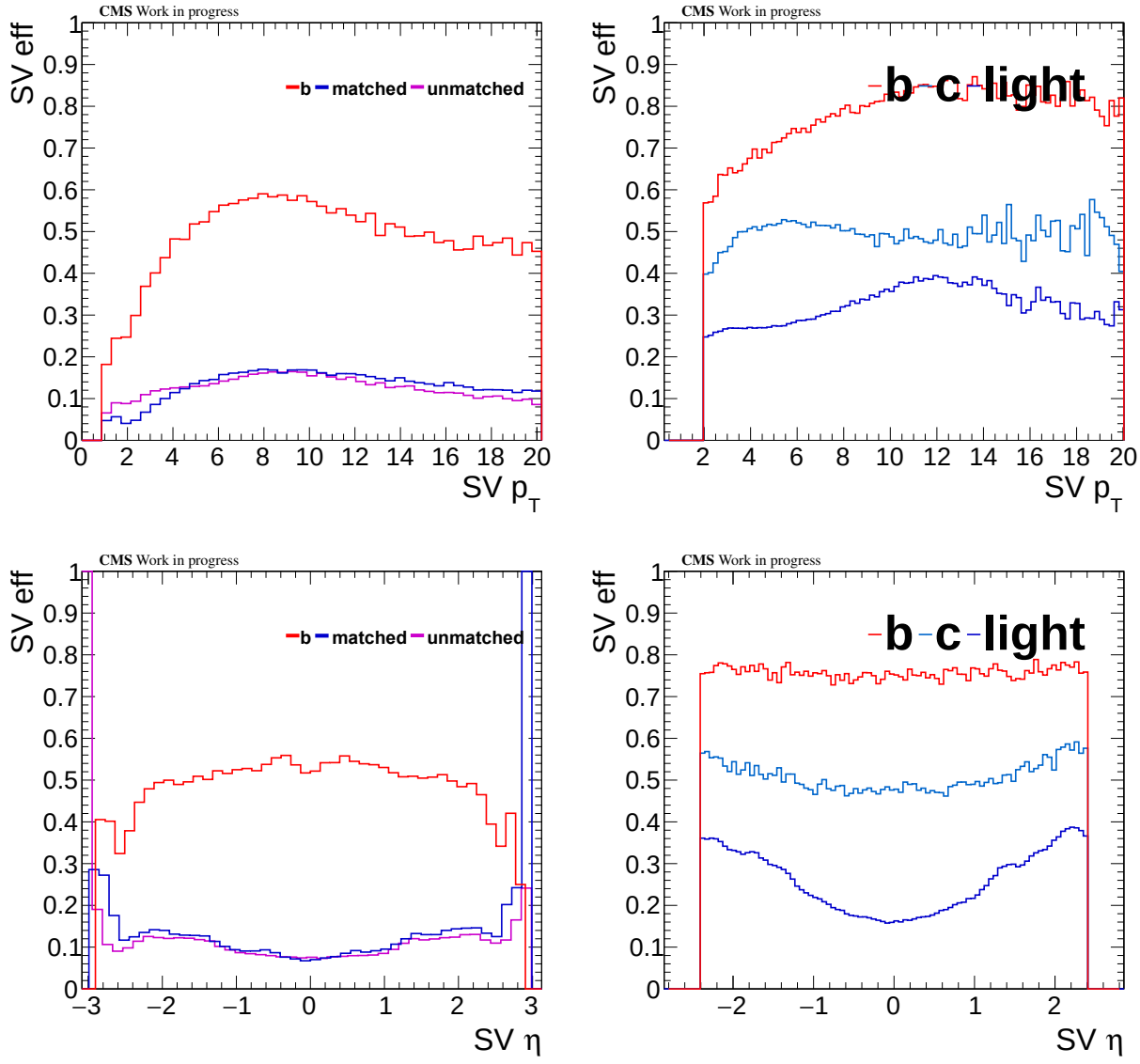


Figure 4.37: Comparison of performance efficiencies over p_T (TOP) and η (BOTTOM) of b, matched and unmatched SVs using the NANO8 (LEFT) and NANO8plus trainings (RIGHT).

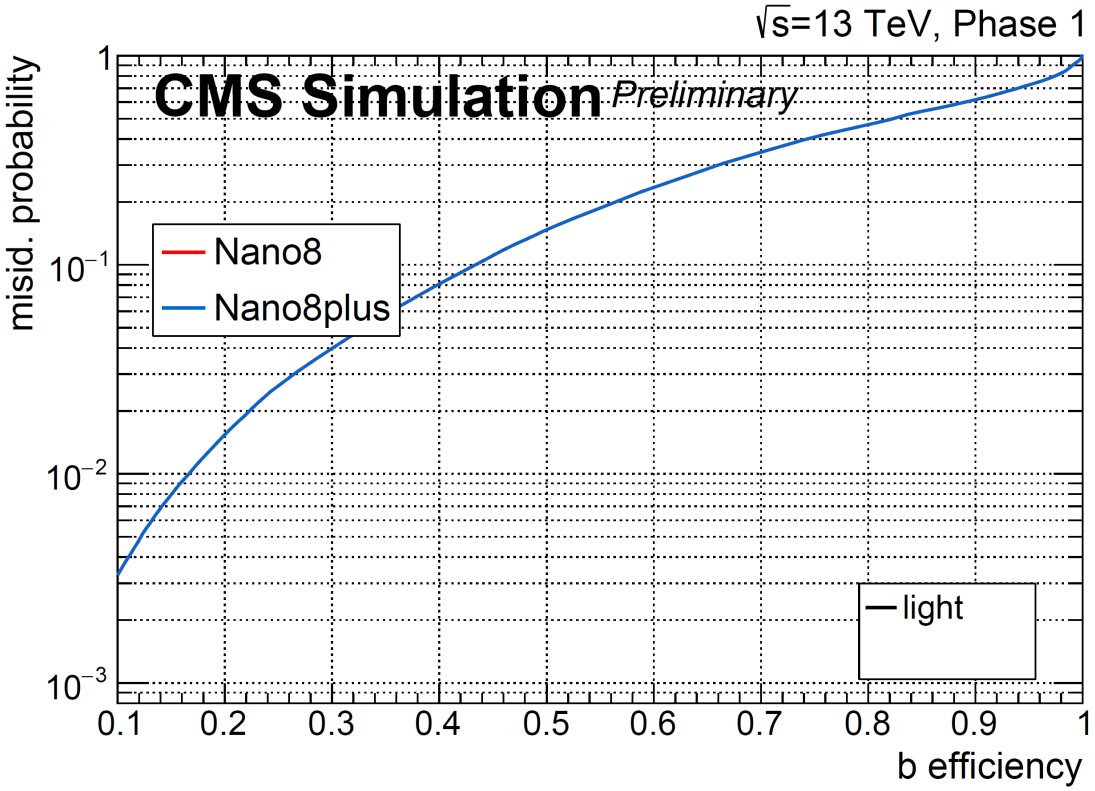


Figure 4.38: ROC curve comparing the *b* efficiency to *light* misidentification probability for the NANO8 and NANO8plus trainings. These two trainings are tested on the same dataset, which corresponds to SVs that have a $SIP3D > 3$.

4.5.1.9 NANO8plus-JetIso20 Training

The NANO8plus-JetIso20 training was used to produce the final sv discriminator used in the analysis. It is a further adjustment of the NANO8 training where the SVs used as inputs into the training were made to match the SVs that were used in the analysis. This adjustment corresponds to loosening the jet isolation that was applied to the SVs used in the training. Initially the SVs matched to any jet were removed from the collection, but for the analysis there is a looser isolation criteria, where only SVs matched to jets that have a $p_T > 20$ GeV are removed from the collection. This isolation corresponds to the divide where jets can be b-tagged, as *b* tagging criteria for jets requires the jet to have a $p_T > 20$ GeV. By including jets with a lower p_T, potential *b* jets that would otherwise be removed will remain through the SVs that are inside of these jets.

Looking at the performance of this training, there is an improvement in *b* efficiency at both low and high p_T, seen in Figure 4.41. For the increase in *b* efficiency at low p_T there is an equivalent increase in *light* flavor efficiency, which does not occur at high p_T. In the η distribution, there is a larger fraction of *light* flavor SVs in the high η regions, with a fairly flat *b* flavor efficiency, improved over the previous training. There is not a lot of concern with the increased *light* flavor efficiency in the high η region, as in the analysis, the SVs are further separated into a low and high η selection, where for the signal regions, only SVs with an $|\eta| < 1.5$ are considered, with the SVs that fail this selection being kept for a control region. With larger fractions of *light* flavor in the high η region, this region can be used to better constrain other regions with higher *light* flavor components, with the low η regions being used to better isolate the *b* flavor SVs. For the overall performance, which can be seen in Figure 4.42, there is a small difference seen at low and high *b* efficiency values. For the lower values, the NANO8plus-JetIso20 training has a slightly lower misidentification probability for an equivalent *b* efficiency, while at the highest *b* efficiencies the opposite is true. Overall this is a fairly small effect, mostly seen at high η, and low/high p_T, as was described previously.

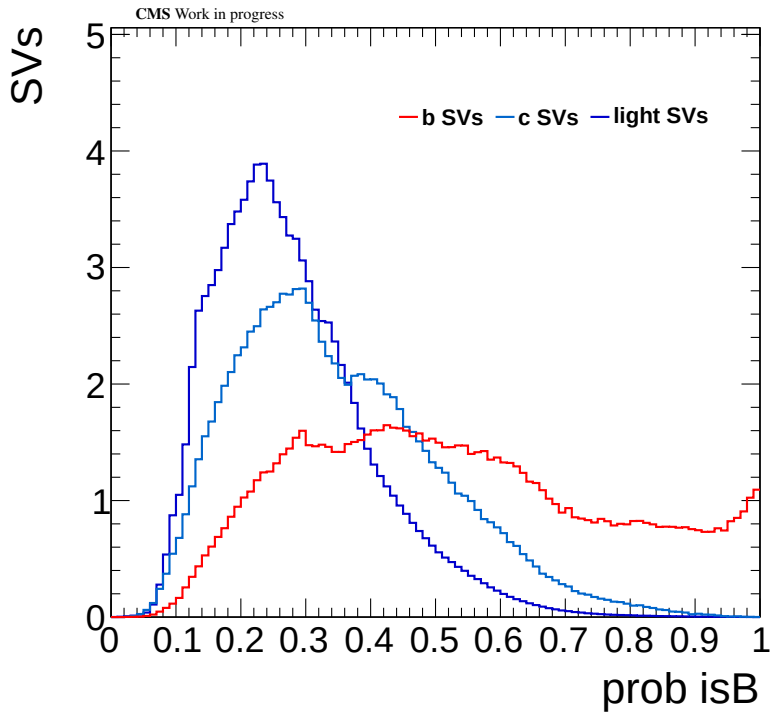


Figure 4.39: Discriminator created using the NANO8plus-JetIso20 training, normalized to 100.

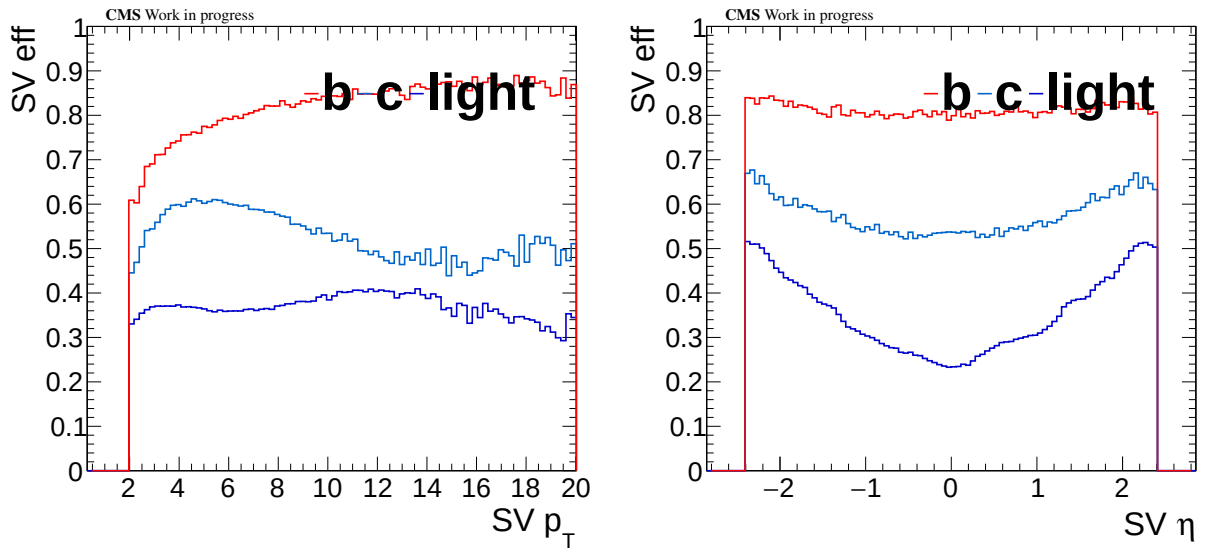


Figure 4.40: Performance efficiencies of the NANO8plus-JetIso20 training in p_T (LEFT) and η (RIGHT). The flavors shown are *b* (red), matched (dark blue), and unmatched (magenta).

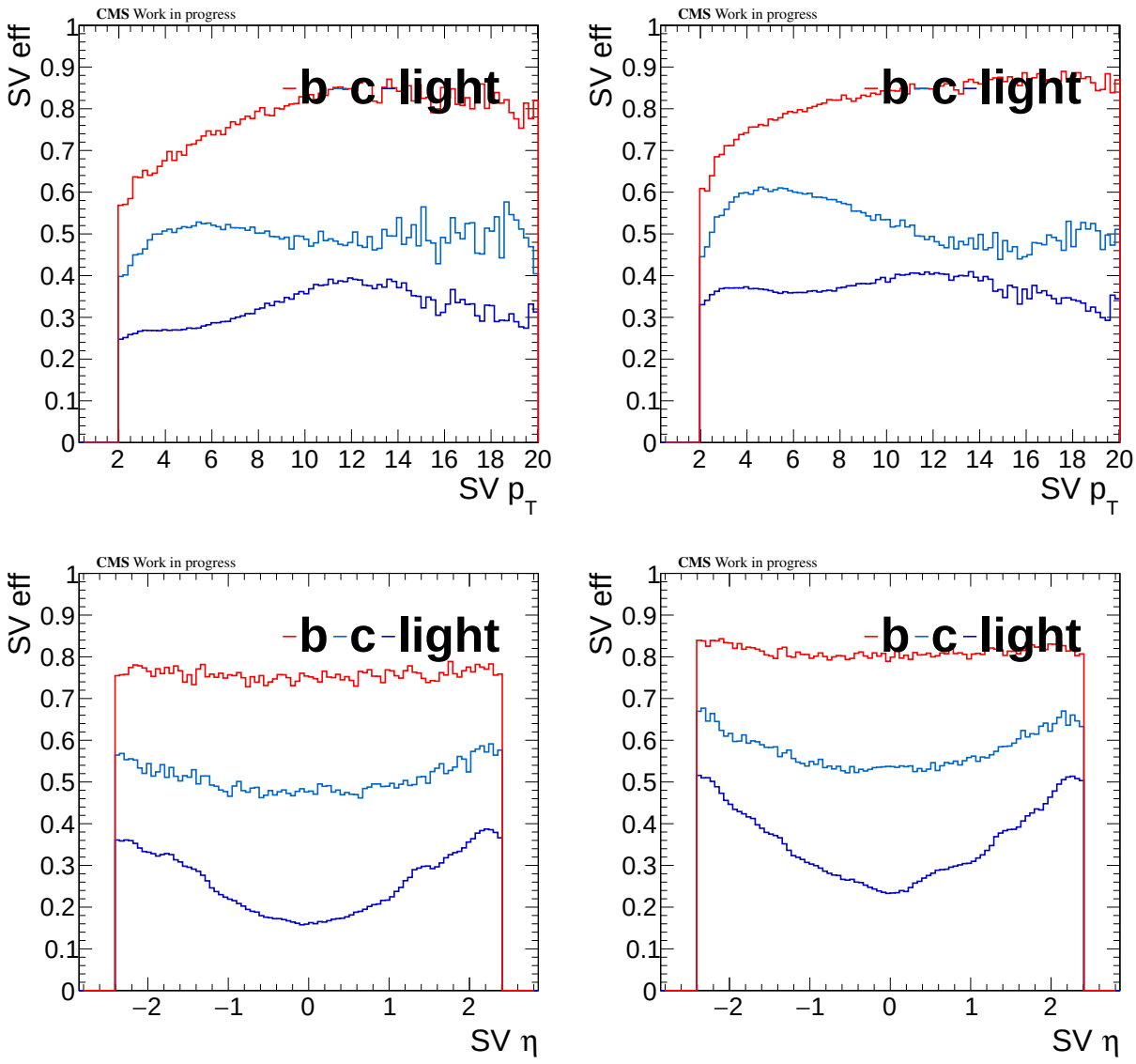


Figure 4.41: Comparison of performance efficiencies over p_T (TOP) and η (BOTTOM) of b, matched and unmatched SVs using the NANO8plus (LEFT) and NANO8plus-JetIso20 trainings (RIGHT).

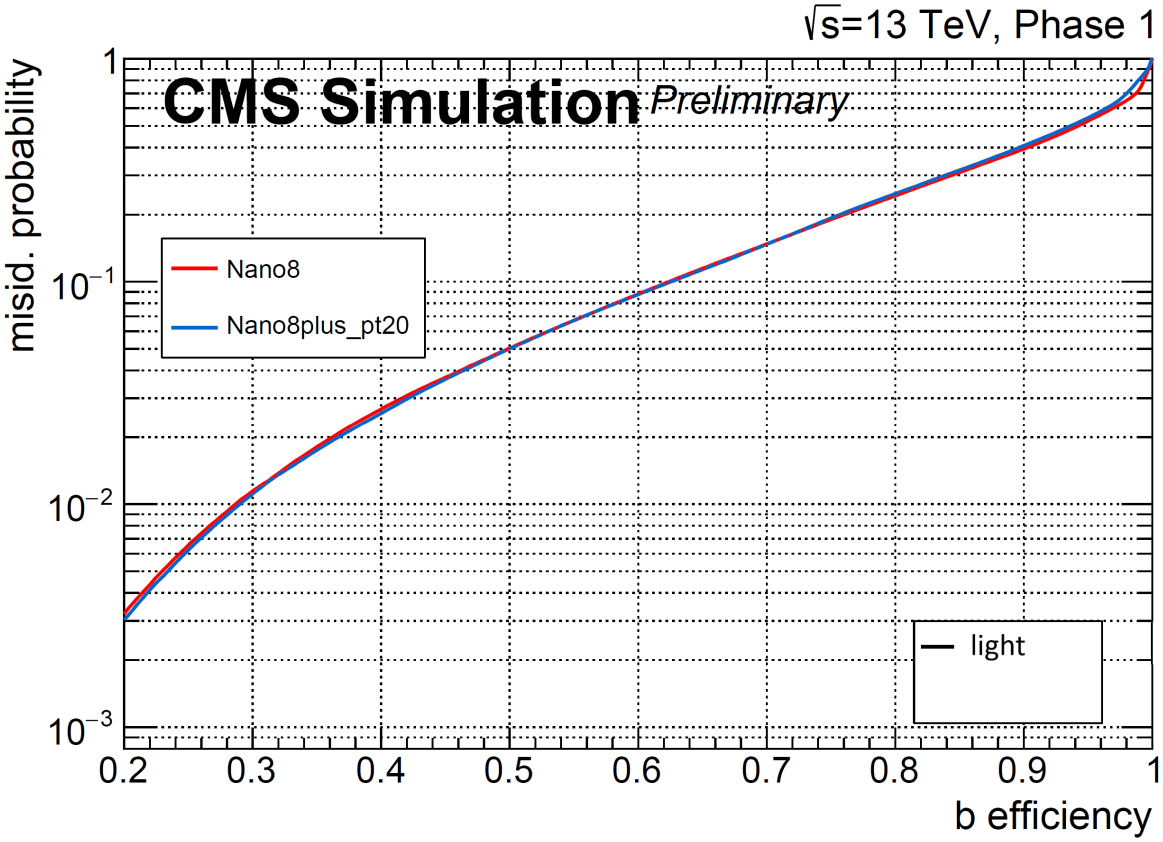


Figure 4.42: ROC curve comparing the *b* efficiency to *light* misidentification probability for the NANO8 and NANO8plus-JetIso20 trainings, using the same testing samples which has the $SIP3D > 3$ cut on the SVs

4.5.2 Training Summary

There were a total of 10 different trainings performed before deciding on a final discriminator. Each discriminator showed improvement over the previous one, with the exception of the NANO8-WITHJETS training, which was only used to confirm that including SVs in jets was a bad idea. The trainings from 4-VAR up to 9-VAR are included in a ROC curve in Figure 4.43, where it can be seen that the largest increase in performance occurs when the p_T and η are added as inputs to the training. Subsequent trainings still improve the results, albeit by smaller amounts. The best discriminators came from the NANO8 collection of discriminators, even better than the 9-VAR which technically provided more information. Of the NANO8 discriminators, there is not a large difference in their performance, but the final discriminator was chosen to be the NANO8plus-JetIso20, as it best matched the SV population that is used in the analysis, as was intended. The SV variables used as inputs in this training were p_T, η, M, N_{dof}, d_{xy}, d_{3D}, SIP_{3D}, and $\cos(\theta_{SV,PV})$. Additionally the p_T and η were used to produce weights that would kinematically flatten the output classifiers, in order to remove normalization bias between the different flavors. The final ROC plot of this discriminator is found in Figure 4.45, and its performance over p_T can be seen in Figure 4.44 with a discriminator cut of > 0.3. The cut of > 0.3 was chosen, as this was the best balance found between rejecting *light* flavor SVs, and retaining events corresponding to the compressed SUSY signals. Applying a tighter cut would remove too many events from the SUSY signals, while applying a looser cut lets in too many events with predominantly *light* flavor objects, mainly seen in W+jets.

The final selection for SVs then requires that the candidate have a p_T between 2 and 20 GeV, and a discriminator cut of > 0.3. Additionally, the SV could not be matched to any jet with a p_T above 20 GeV. There are also η bins that are put into use for the SVs, where it has been found that the b flavor SVs will generally congregate near central η, while the *light* flavor SVs are more prevalent in the forward η regions.

Following the comparison of performance between the final discriminator and the square cuts in Figure 4.46, the final discriminator was able to improve the b-efficiency by up to 50% at low p_T

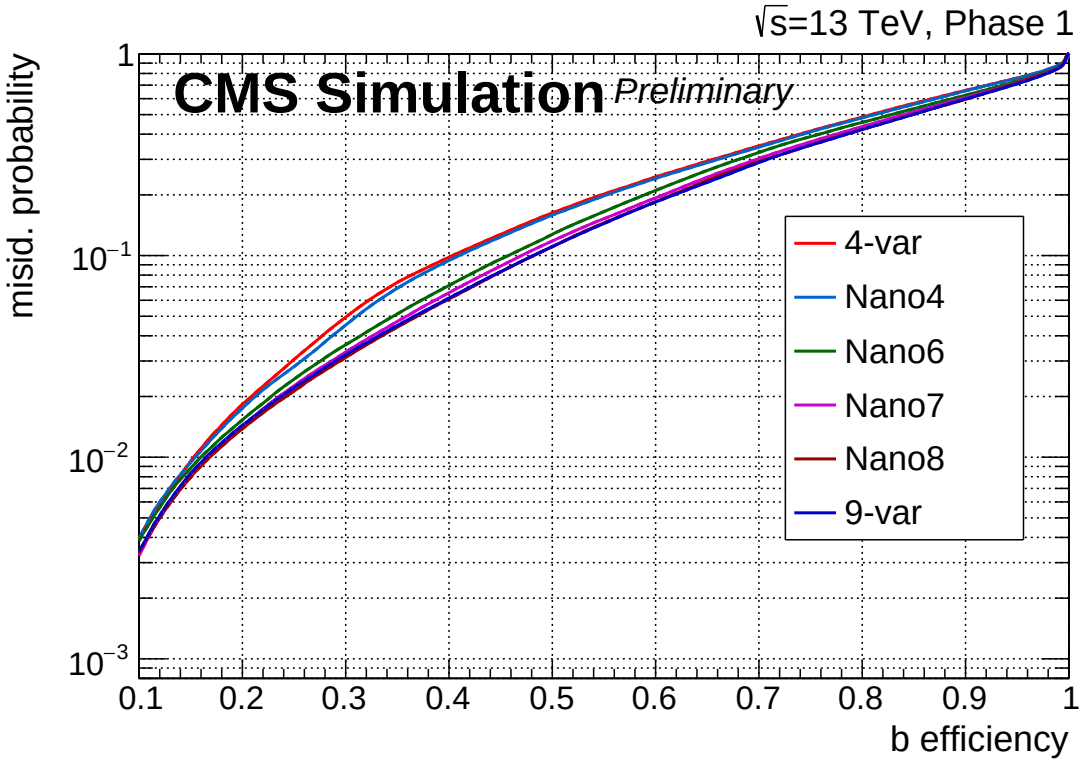


Figure 4.43: ROC curve comparing the b efficiency to *light* misidentification probability for the trainings up to NANO8, and including 9-VAR. This shows that the largest gain in performance comes from the NANO6 training, which adds the p_T and η as inputs. The NANO8 and 9-VAR are almost indistinguishable, except at low efficiency values where NANO8 does slightly better.

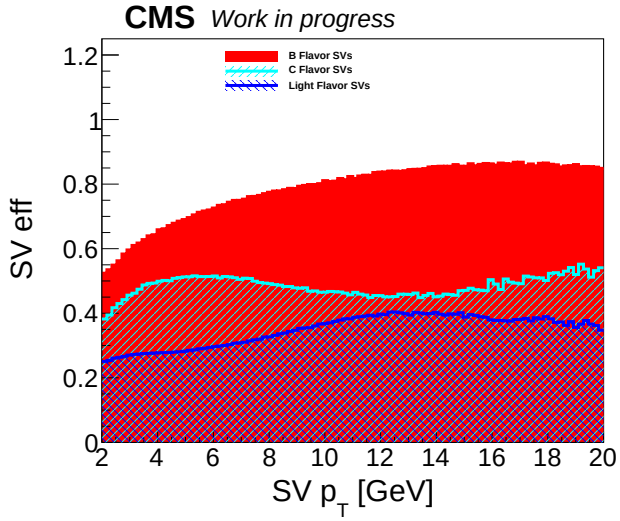


Figure 4.44: Distributions of the b, c, and *light* flavor efficiencies over p_T for the final discriminator. The samples used to create these efficiencies are the full background for the analysis, not including QCD multijets.

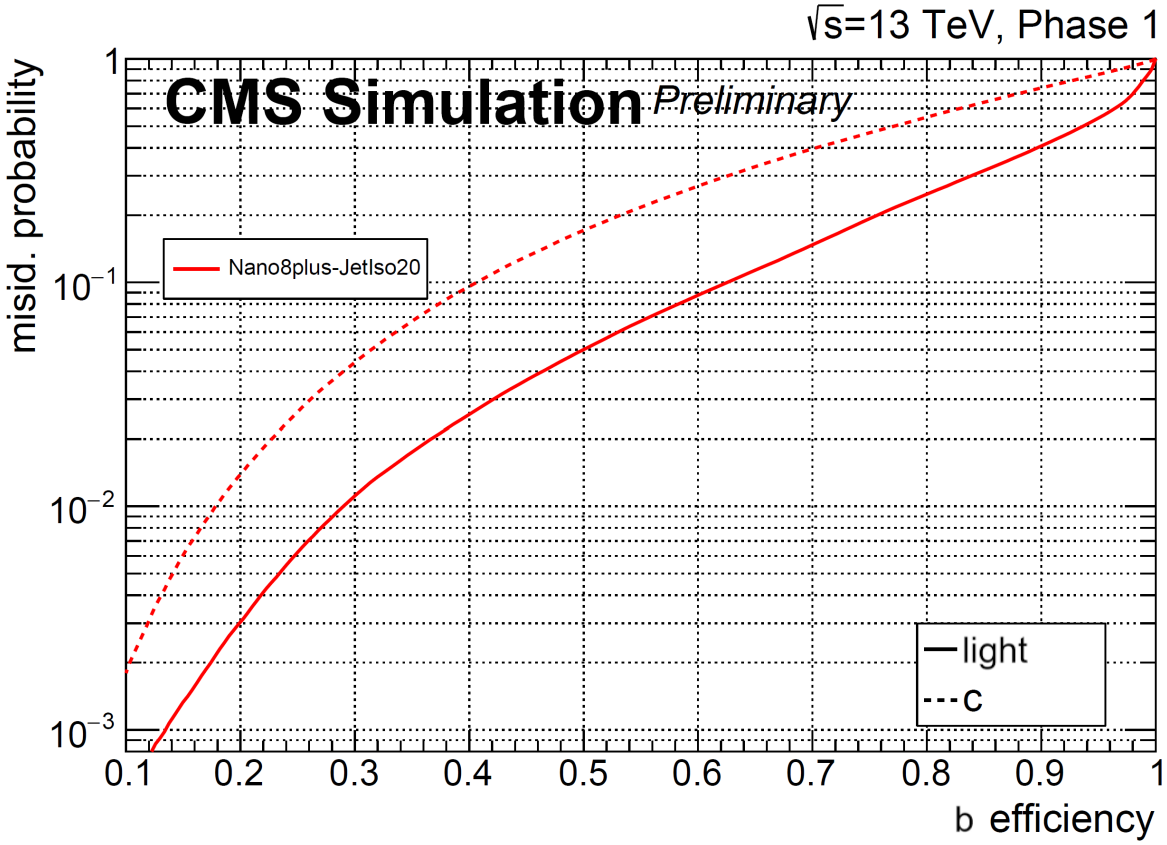


Figure 4.45: A ROC curve comparing the *b* efficiency to light/c misidentification probability for the final discriminator, labeled Nano8plus-JetIso20. The chosen working point for the discriminator is 0.3. The samples used to create the curve correspond to those listed in Table 4.1.

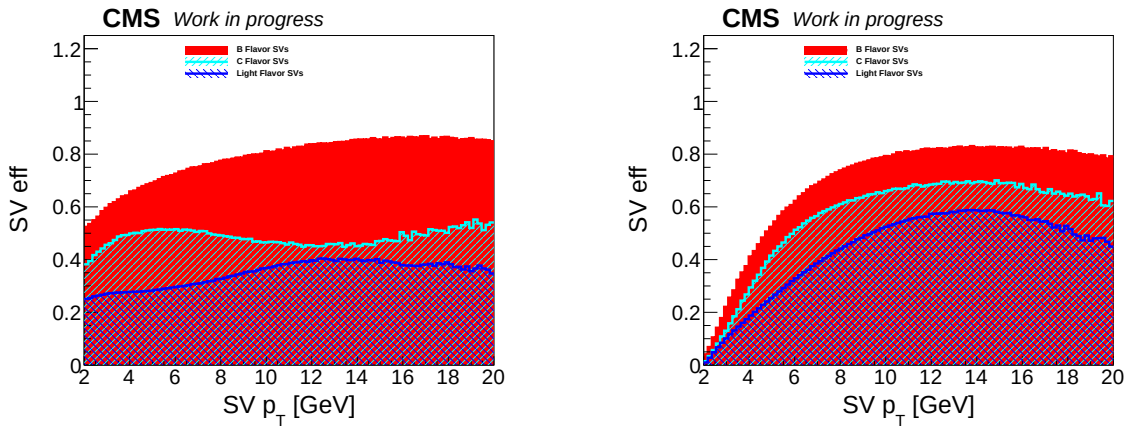


Figure 4.46: Distributions of the b, c, and $light$ flavor efficiencies over p_T for the discriminator on the left and the square cuts on the right. The samples used to create these efficiencies are the full background for the analysis, not including QCD multijets.

over that of the square cuts. Over the p_T range of the secondary vertices, from 2 to 20 GeV, the light flavor misidentification rate was maintained below 40%. This was compared to the square cuts, which had a light misidentification rate above 40% for the p_T range from 8 to 20 GeV, while having a comparable b-efficiency to the final discriminator. At a p_T of about 5 GeV, the misidentification rate between the two taggers intersect, where the square cuts then has a lower misidentification rate, but the trade off is that its b-efficiency takes a proportional dive, falling by about 50% moving from 5 to 2 GeV, while the final discriminator's b-efficiency is better maintained, only falling by about 15% as it moves from 5 to 2 GeV.

As has been mentioned, all flavors drop off sharply at low p_T for the square cuts, and this is due to the removal of 2 track SVs. Looking at Figure 4.48, one can see that a large population of SVs exists in the region corresponding to 2 track SVs ($N_{dof} < 1.8$), in both the b and $light$ flavor SVs. An added bonus is that the 2 track SVs between the two flavors are kinematically separated in both η and p_T, where the b flavor SVs are mostly central, which is opposite for the $light$ flavor ones. In addition, the b flavor SVs are mostly found between 3 and 5 GeV, while the $light$ flavor SVs are just above 2 GeV. When the ≥ 2 tracks cut is added in addition to the discriminator cut, and then compared to the same square cut efficiency plot, we get Figure 4.47, which reflects what is seen in Figure 4.48. Comparing these distributions shows that the 2 track SVs are very important

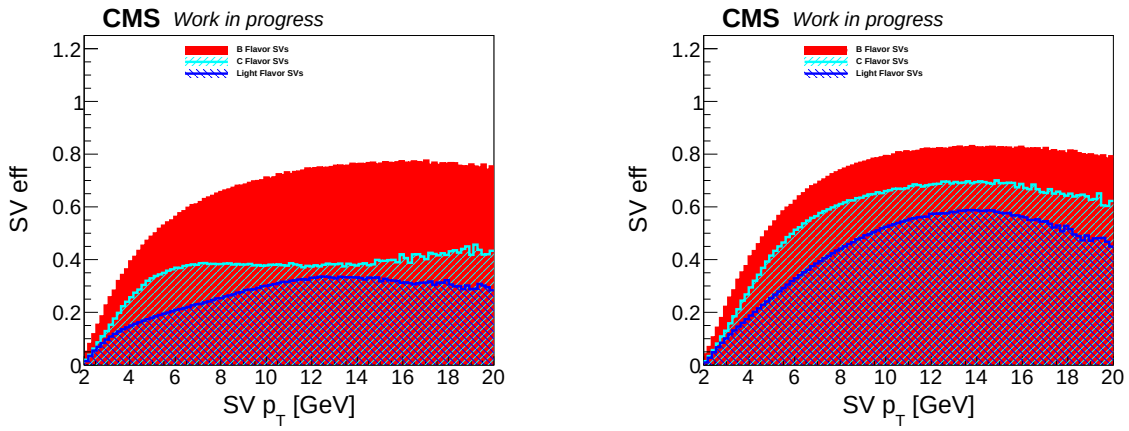


Figure 4.47: Distributions of the b, c, and $light$ flavor efficiencies over p_T for the discriminator on the left and the square cuts on the right. The numerator for the discriminator plot has an additional requirement of $N_{dof} > 1.8$ in order to only look at 3 or greater track vertices. The samples used to create these efficiencies are the full background for the analysis, not including QCD multijets.

for the efficiency gains at low p_T, but it is not the only reason for the increased performance, since the final discriminator still has better performance when considering just the 3 or greater track SVs over the whole population of SVs. Apart from allowing in 2 track SVs, the training is also able to greatly suppress the $light$ misidentification rate over that of the square cuts, while maintaining a similar b efficiency.

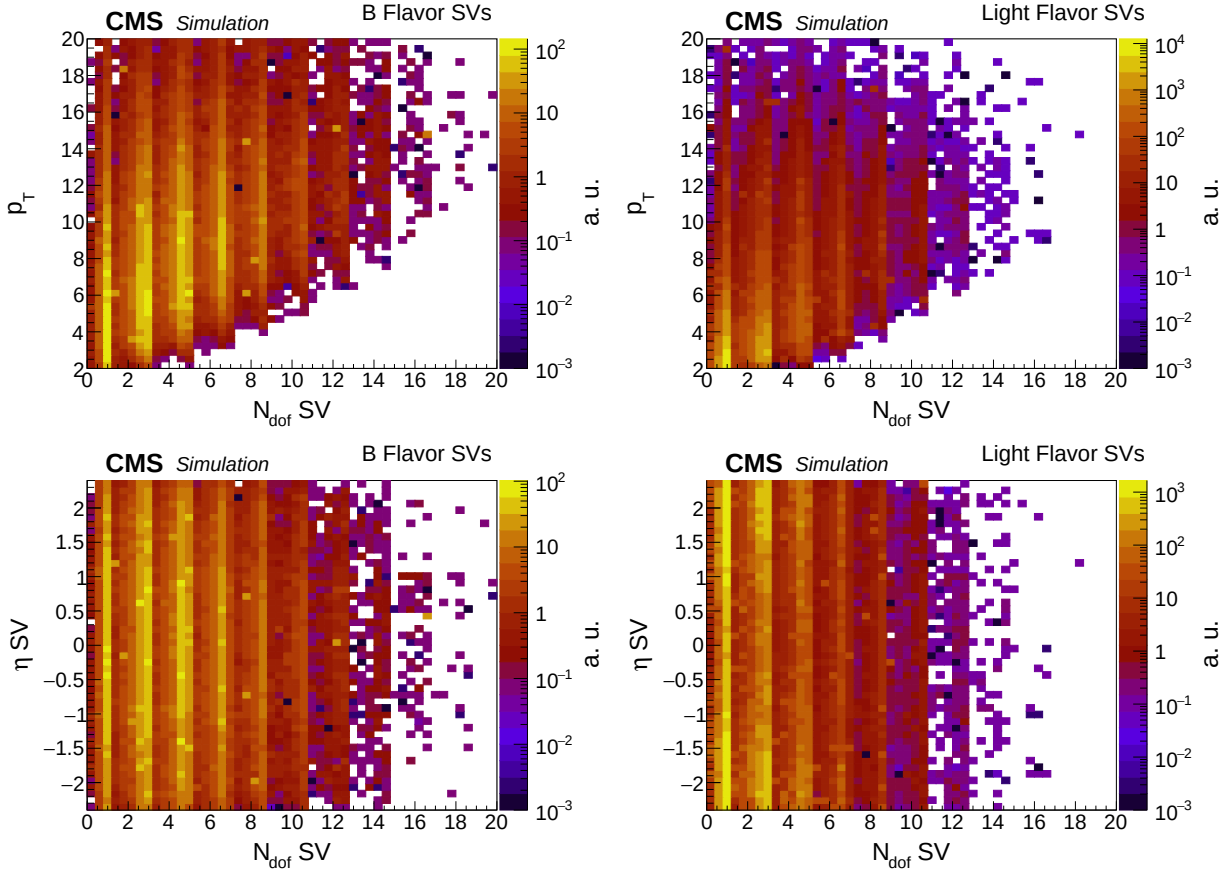


Figure 4.48: 2D histograms of N_{dof} versus p_T (top) and N_{dof} versus η (bottom) for b flavor SVs (left) and *light* flavor SVs (right). These distributions show how there is a large population of SVs below $N_{dof} < 1.8$, in both the b and *light* flavors, while still having discrimination between the two.

4.6 Data/MC Comparisons

After producing a discriminator it is important to look at the data to MC agreement for the discriminator, and the variables used to create it. We look at the agreement in two different regions, due to their differing flavor fractions and backgrounds. The first region is a low-$\not{p}_T$ di-lepton control region that is dominated by $t\bar{t}$ and has a large fraction of b flavor SVs, and the second is an inclusive single lepton region that is dominated by W+jets and has a much higher fraction of $light$ flavor SVs. For the low-$\not{p}_T$ di-lepton control region the selection requires two golden leptons as an opposite sign $e\mu$ pair, $\not{p}_T$ between 50 and 175 GeV, and at least one SV passing the selection as described previously. The single lepton selection requires one golden lepton of any flavor, $\not{E}_T > 175$ GeV, and at least one SV. The MuonEG dataset was used for the low-$\not{p}_T$ di-lepton region, which is why we can go lower in $\not{p}_T$. In addition because of this lower $\not{p}_T$ cut, the selection is complementary to the analysis regions, and contains a higher number of events than the analysis region, allowing for better data to MC comparisons. The single lepton region uses the MET dataset, which necessitates the higher MET cut due to trigger requirements. Figure 4.49 gives the discriminator in the low-$\not{p}_T$ di-lepton region, with one plot having the discriminator cut applied to show the good agreement in data to MC where the cut is applied and Figure 4.50 gives the discriminator in the single lepton region. The data to MC agreement in the low-$\not{p}_T$ di-lepton regions as seen in Figures 4.51 and 4.52, show good agreement. The single lepton region distributions, as seen in Figures 4.53 and 4.54, while not to the same degree as the low-$\not{p}_T$ di-lepton region, also show good agreement between the data and MC.

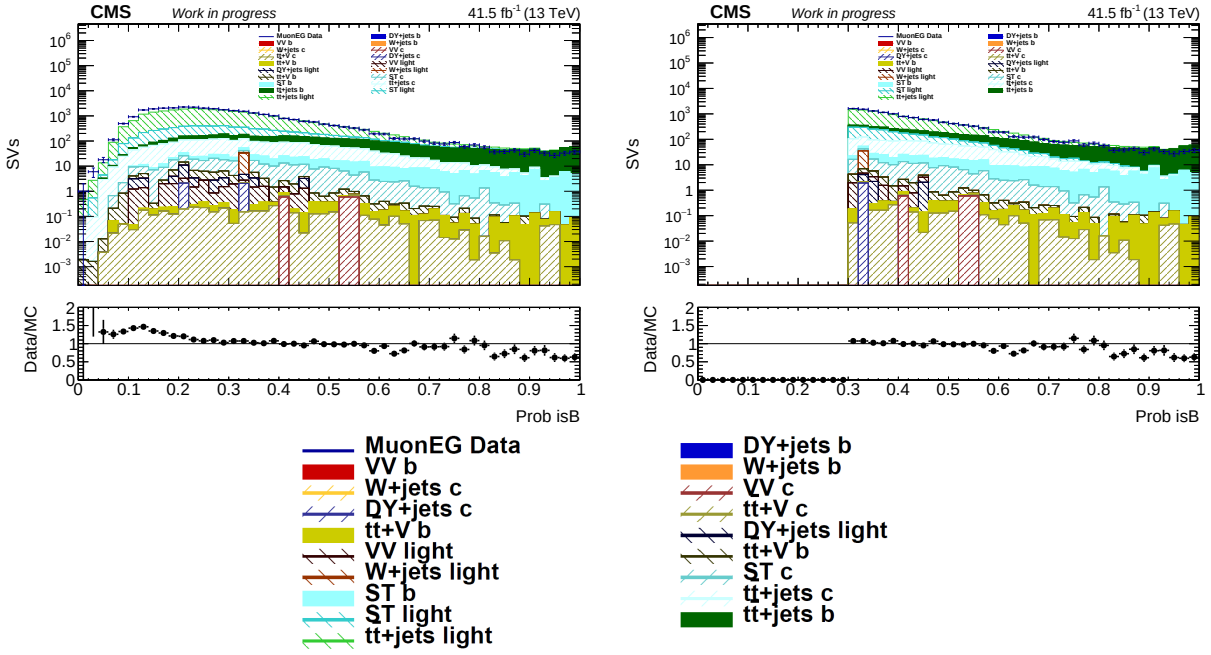


Figure 4.49: Distributions of the discriminator in the low-$\not{p}_T$ di-lepton region, before the discriminator cut on the left, and after the discriminator cut on the right. These distributions show that there is a good data to MC agreement after the discriminator cut. The bottom figure is an enlarged view of the legend.

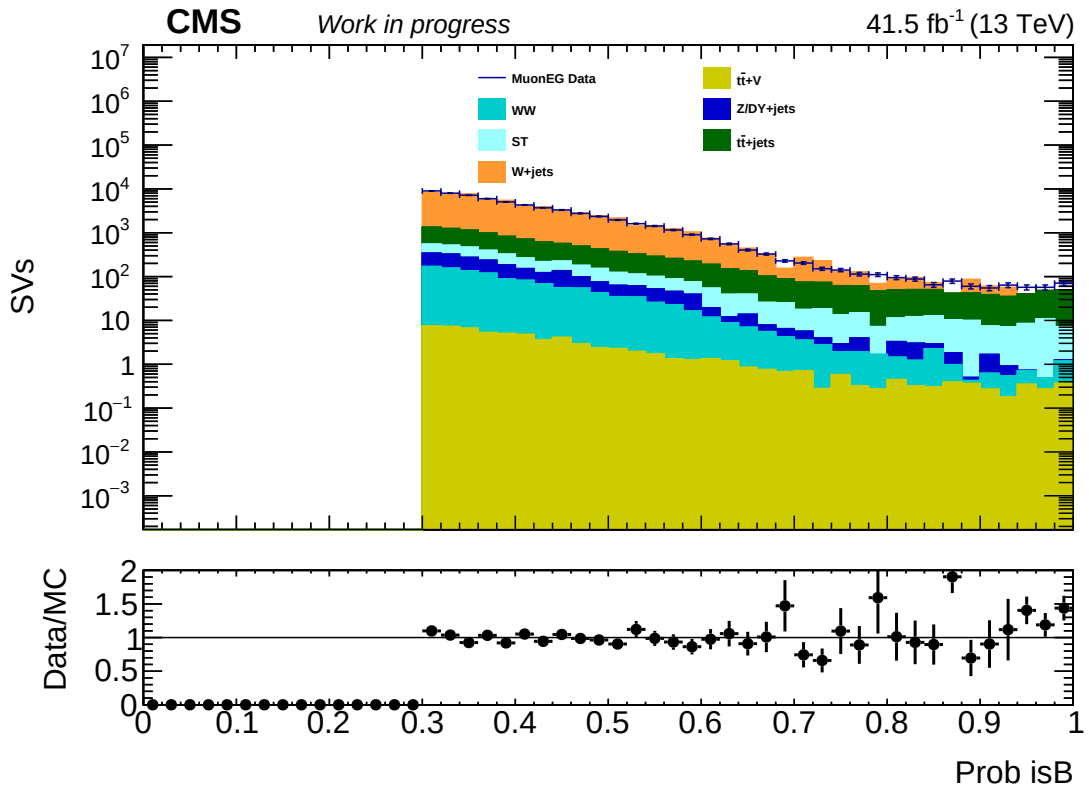


Figure 4.50: A distribution of the discriminator in the single lepton region after applying the discriminator cut. This distribution shows a good agreement after the discriminator cut.

.

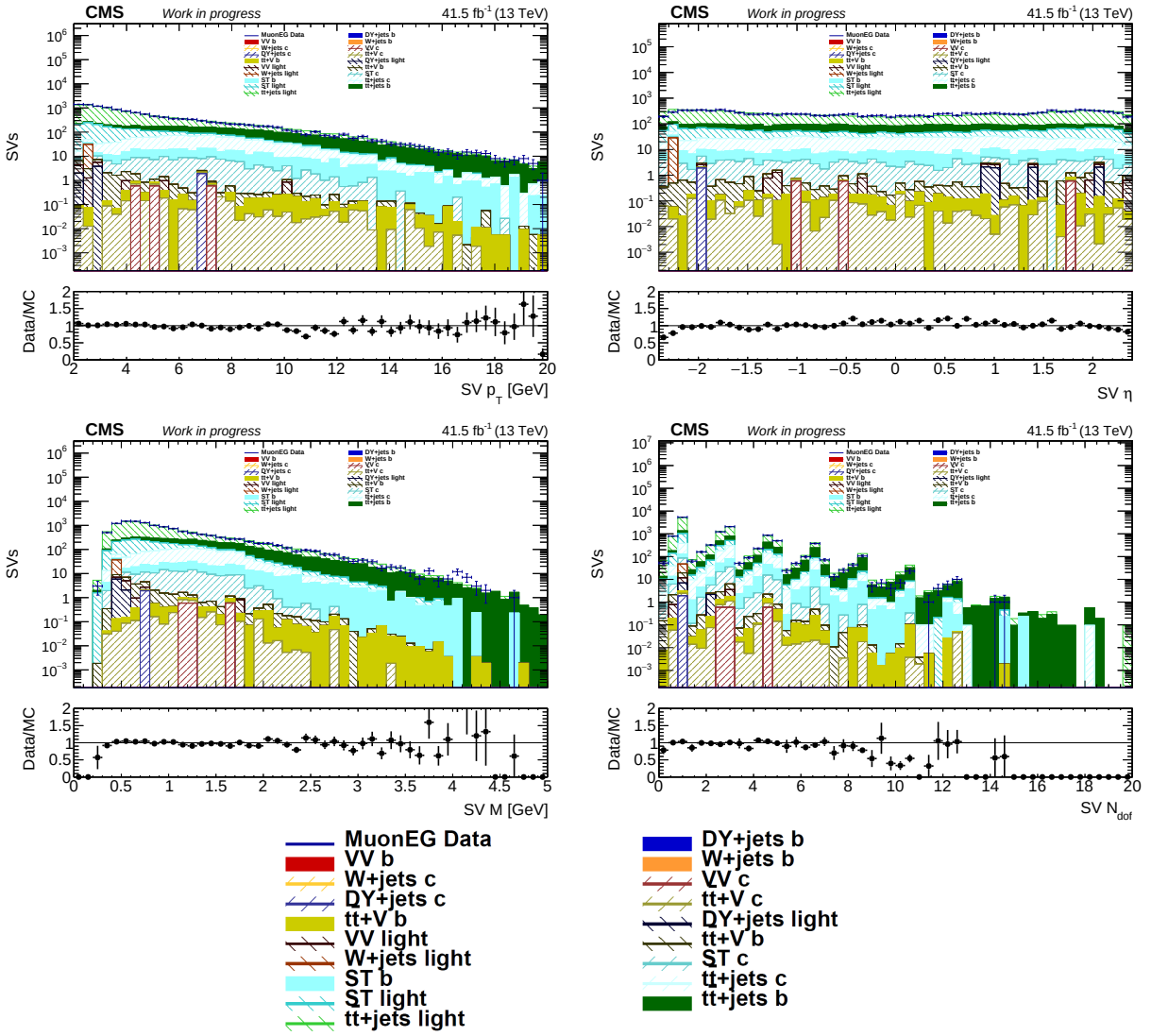


Figure 4.51: Stacked histograms of the p_T (top left), η (top right), mass (bottom left), and N_{dof} (bottom right) of the SVs in the low-$\not{p}_T$ di-lepton region after the discriminator cut. These histograms are separated into the analysis backgrounds, where each background is further separated into their flavor components. The bottom figure is an enlarged view of the legend.

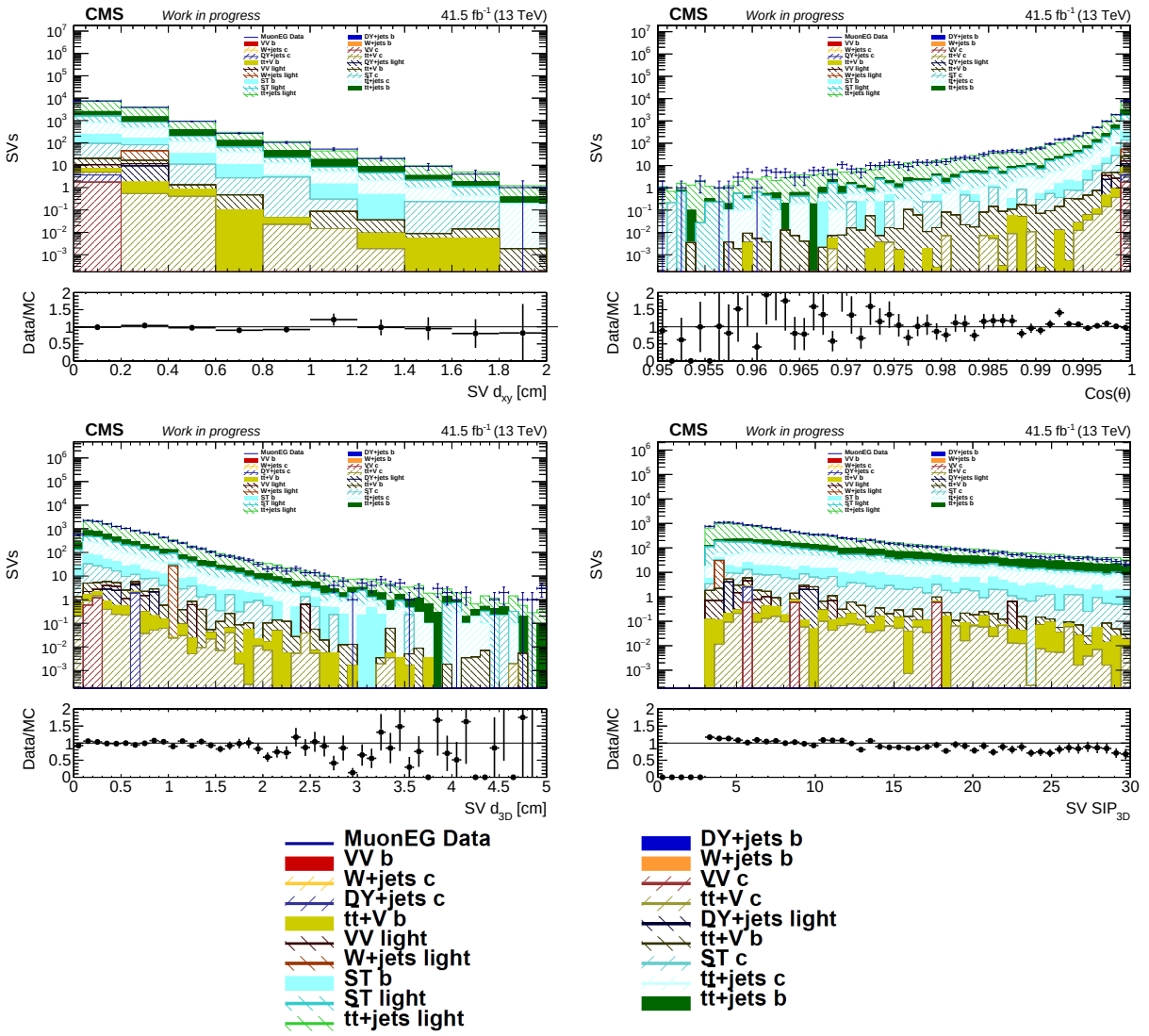


Figure 4.52: Stacked histograms of d_{xy} (top left), $\cos\theta_{SV,PV}$ (top right), d_{3D} (bottom left), and d_{3D}^{sig} (bottom right) of the SVs in the low-$\not{p}_T$ di-lepton region after the discriminator cut. These histograms are separated into the analysis backgrounds, where each background is further separated into their flavor components. The bottom figure is an enlarged view of the legend.

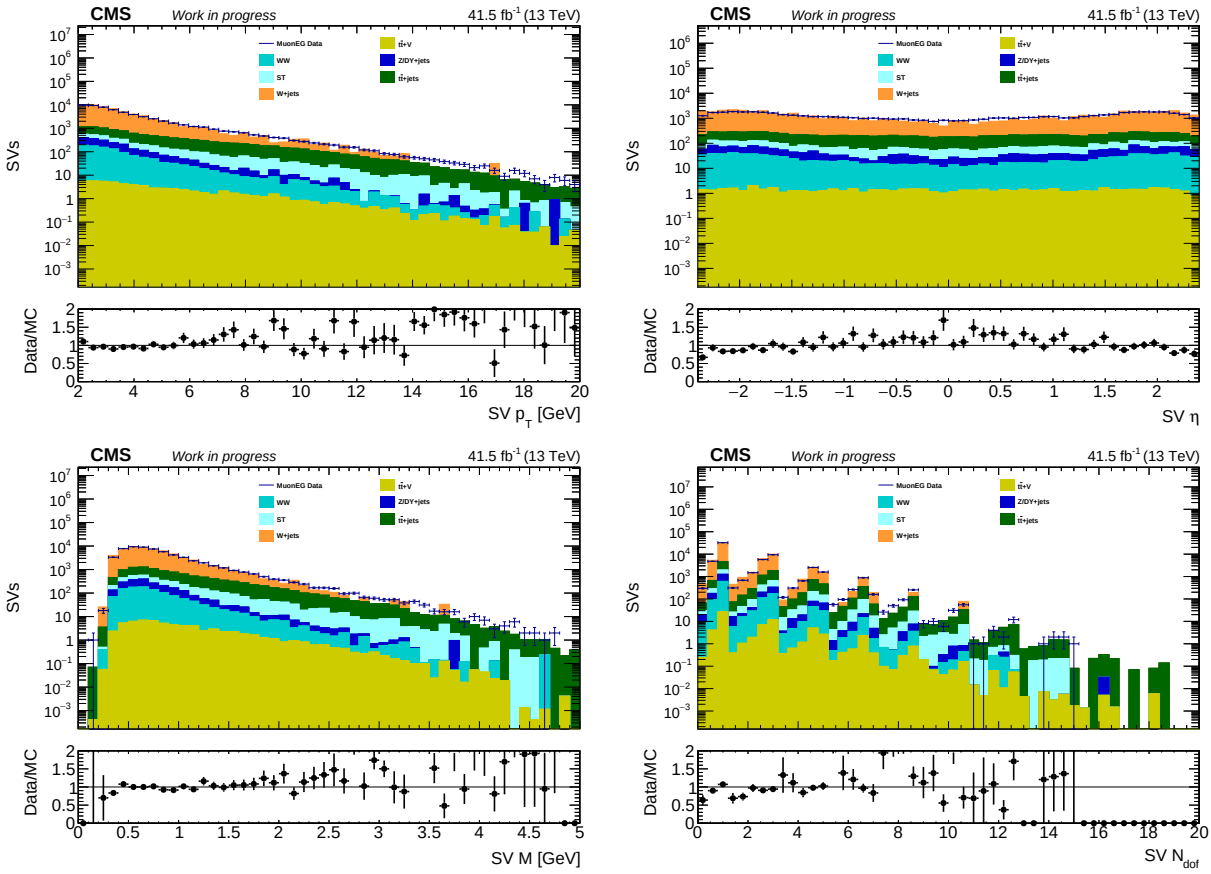


Figure 4.53: Stacked histograms of the p_T (top left), η (top right), mass (bottom left), and N_{dof} (bottom right) of the SVs in the single lepton region after the discriminator cut. These histograms are separated into the analysis backgrounds, where the W+jets has a large fraction of *light* flavor SVs

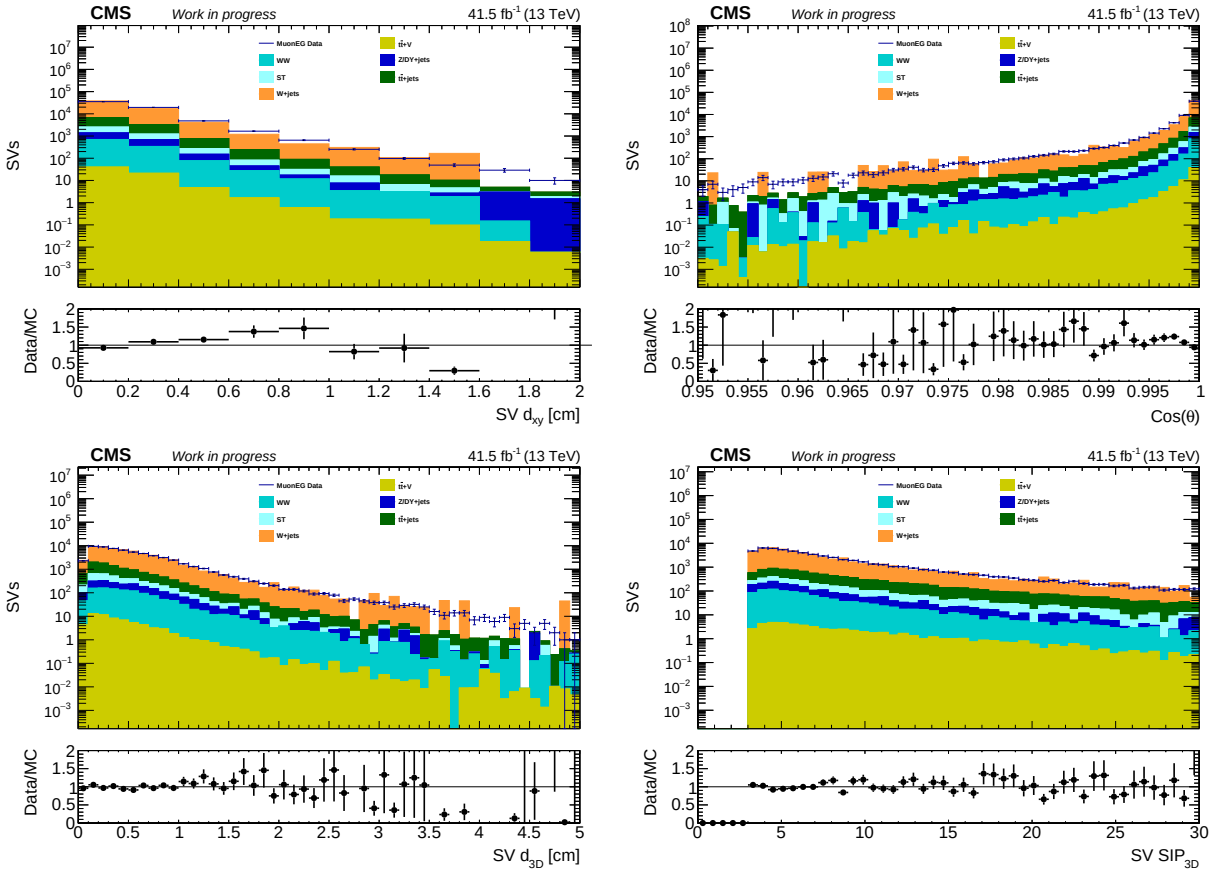


Figure 4.54: Stacked histograms of d_{xy} (top left), $\cos\theta_{SV,PV}$ (top right), d_{3D} (bottom left), and d_{3D}^{sig} (bottom right) of the SVs in the single lepton region after the discriminator cut. These histograms are separated into the analysis backgrounds, where the W+jets has a large fraction of *light* flavor SVs

4.6.1 Data to MC efficiency measurements

In order to use the discriminator to tag SVs in the analysis, it is important to make sure that the MC used to model the SVs is well described in data, and to make proper corrections if that is not the case. Normally this requires comparing the data and MC in a heavy flavor region that also requires a tagged object, in order to determine an efficiency that can be used to scale the MC in this region to match the data in that same region, producing a heavy flavor scale factor. A similar approach would be done for a *light* flavor scale factor, commonly called the mistag scale factor. For this SV tagger, even in a region that is almost guaranteed to have *b* flavor objects, there is still a significant amount of *light* flavor SVs, especially at low p_T. Other regions beside this heavy flavor enriched one, will have different fractions of light and heavy flavor SVs, which cannot be described by a single scale factor.

As a solution to this, the analysis will implement a region specific scale factor measurement that will be determined by the analysis, during the limit setting fit. The method to calculate such a scale factor follows Equation 4.1, where N_{region} is the number of events in the region without explicitly selecting SVs, and $N_{region+SV}$ is that same region, but including an SV requirement. The purpose of the double ratio is to remove data to MC differences that are not explicitly due to an SV selection.

$$SF(SV) = (N^{Data}_{region+SV}/N^{MC}_{region+SV})/(N^{Data}_{region}/N^{MC}_{region}) \tag{4.1}$$

Essentially what this equation does is scale the MC in a region with an SV selection by the data to MC factor from the inclusive version of this region. By doing so, any data to MC disagreement caused by the selection in the inclusive region is removed, and any remaining disagreement in the SV region can be attributed to the selection of the SV object. The scale factor from Equation 4.1 would then be the factor used to correct the MC for the SV selection. Every region would have its own version of this scale factor, as flavor fractions change depending on which background process is dominant.

This method can be used in the analysis due to how regions are treated. Since it is a generic SUSY search, for every cut in the analysis that goes into building a signal region, instead of throwing out the events that do not pass the selection, the events are kept in order to either build a separate signal region, or a control region that can be used to constrain the background. With this, for every region with an SV requirement, the information for the inclusive version of that region is also being used either as a separate signal or control region, and can also be used to calculate a SV scale factor per region.

With the large number of regions in the analysis, the SF measurement is best done during the limit setting, where the scale factor is treated as a nuisance parameter, and is measured using the data driven background. Furthermore, in some of the regions, the SVs are separated into central ($|\eta| < 1.5$) and forward ($|\eta| > 1.5$) SV regions, which better separates both SUSY signal and b flavor SVs, both of which favor the central η regions. For the case of the SV flavor, this can be seen in Figure 4.55, which shows that the *light* flavor SVs are pushed more to the forward regions, creating a larger b flavor fraction in the central regions. For the scale factors, this gives a better idea of the correction needed for higher *light* and b flavor fractions by focusing on the central or forward regions. For the analysis proper, these η regions increase the signal sensitivity in the central regions.

Following Equation 4.1, example SV scale factors are calculated using the two lepton regions described in Section 4.6. These scale factors are given in Table 4.2. The numbers detail the importance of calculating separate SFs for each region. Between the two lepton regions, there are different values for SFs, where we want to underline the fact that these two lepton regions have different background contributions and SV flavor fractions. The same thing is seen for the central and forward η regions, especially in the single lepton region, with the central region giving a SF above 1, and the forward region giving a SF below 1.

Lepton region	η inclusive	central	forward
low-$\not{p}_T$ di-lepton	0.879 ± 0.011	0.892 ± 0.015	0.858 ± 0.016
single lepton	0.929 ± 0.017	1.04 ± 0.03	0.83 ± 0.03

Table 4.2: Example SFs calculated using the low-$\not{p}_T$ di-lepton and single lepton regions given in Section 4.6. Values are calculated following Equation 4.1, in an inclusive, central, and forward η region. Errors are statistical only.

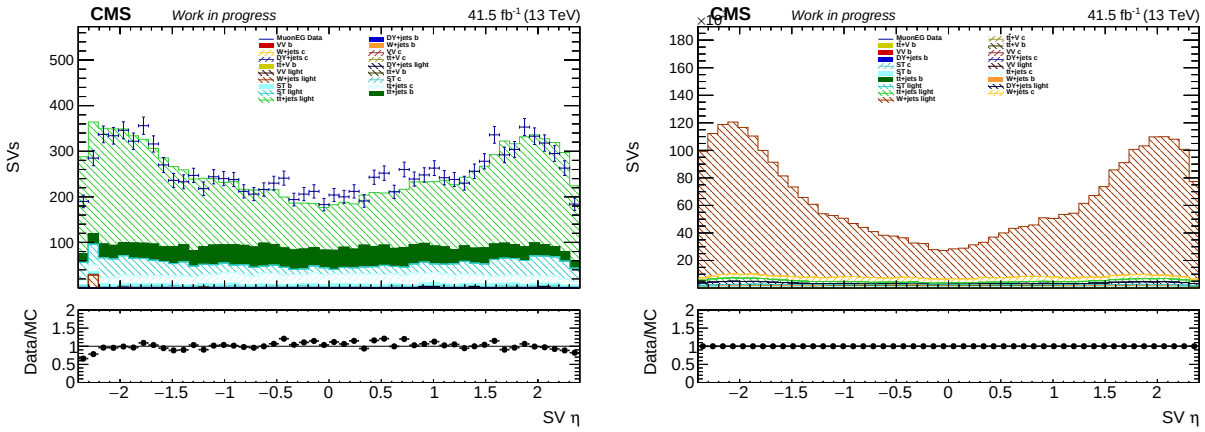


Figure 4.55: Stacked histograms of the η distributions of the low-$\not{p}_T$ di-lepton (left) and single lepton (right) regions, in linear form. The single lepton plot only shows the MC background, without the data, separated into flavor components. These distributions highlight how the *light* flavor SVs concentrate in the forward η regions, creating a larger *b* flavor fraction in the central η regions.

Chapter 5

Event Selection and Reconstruction

The overall analysis makes use of Recursive Jigsaw Reconstruction to approximate the reference frames of a decay tree in order to take advantage of mass sensitive observables that can be produced. In making the search regions, the analysis relies on counting the multiplicities of reconstructed objects, depending on where they were assigned in the decay tree, along with extra regions defined by the mass sensitive variables. There are about 292 regions defined, which are further binned in the mass sensitive variables described in Chapter 3. The overall selection is geared towards increased sensitivities for compressed signals with small mass splittings, which, for the stop quark scenarios are $\Delta m\left(\tilde{t}_1, \tilde{\chi}_1^0\right) < m_t$, where m_t is the mass of the top quark. For categories which make use of the SV tagger, the targeted splittings are more specifically $10 < \Delta m\left(\tilde{t}_1, \tilde{\chi}_1^0\right) < 80$ GeV.

5.1 Samples

The datasets used in the analysis are made up of data and MC simulation samples produced for the 2016, 2017, and 2018 data taking periods. Each year has its own set of samples, using detector conditions specific for that data taking period. In addition, the specific years for the MC samples also simulate the version of the detector used in those years, which is especially important between the years 2016 and 2017, as there were two different pixel detectors in use between 2016 and then 2017 and onward.

The data samples in this analysis made use of the NanoAOD data format [81], which is processed from MiniAOD. The NanoAOD format is a simplified version of what is stored in MiniAOD, only keeping the event and trigger information necessary for most analyses to be able to

use them without any further processing. The format also drops a significant amount of the event content, greatly reducing the size of the dataset, and stores all the information into a flat data tree, making it much simpler to work with than the `edmNtuple` format used by MiniAOD. In addition, many of the various object corrections and IDs, such as energy scale and resolution, and lepton IDs and isolation, are applied and stored as part of the NanoAOD production process, making it easier for analyses to have the most up to date corrections, since they are handled centrally with NanoAOD. The campaign version used was NanoAODv7, which specifies the corrections applied to the objects, and the CMS software version used to produce it.

5.1.1 Data

The analysis makes use of pp collision data at $\sqrt{s} = 13$ TeV, with an integrated luminosity of 137 fb^{-1}. The events are selected from a variety of Primary Datasets (PDs), which are skimmed events that fire a specific set of object triggers during the data taking. These data sets are:

- SingleMuon: single muon channel
- SingleElectron: single electron channel (PD changes to EGamma in 2018 data taking period)
- DoubleMuon: double muon channel
- DoubleMuonLowMass: double muon channel with lower energy threshold
- DoubleEG: double electron channel
- MuonEG: muon, electron channel
- MET: missing transverse momentum channel
- JetHT: jets and hadronic activity channel

The versions of the dataset used by all of the PDs is given in Table 5.1, for all three years. For 2016 data, the samples from Run2016B-H are used. For 2017 data, the samples from Run2017B-F

Year	Dataset
2016	/<PD>/Run2016B-02Apr2020_ver1-v1/NANOAOD
	/<PD>/Run2016B-02Apr2020_ver2-v1/NANOAOD
	/<PD>/Run2016C-02Apr2020-v1/NANOAOD
	/<PD>/Run2016D-02Apr2020-v1/NANOAOD
	/<PD>/Run2016E-02Apr2020-v1/NANOAOD
	/<PD>/Run2016F-02Apr2020-v1/NANOAOD
	/<PD>/Run2016G-02Apr2020-v1/NANOAOD
	/<PD>/Run2016H-02Apr2020-v1/NANOAOD
2017	/<PD>/Run2017B-02Apr2020-v1/NANOAOD
	/<PD>/Run2017C-02Apr2020-v1/NANOAOD
	/<PD>/Run2017D-02Apr2020-v1/NANOAOD
	/<PD>/Run2017E-02Apr2020-v1/NANOAOD
	/<PD>/Run2017F-02Apr2020-v1/NANOAOD
2018	/<PD>/Run2018A-02Apr2020-v1/NANOAOD
	/<PD>/Run2018B-02Apr2020-v1/NANOAOD
	/<PD>/Run2018C-02Apr2020-v1/NANOAOD
	/<PD>/Run2018D-02Apr2020-v1/NANOAOD

Table 5.1: Datasets used when processing the various PDs. All of the PDs use the same processing campaign, so where those names would go is instead replaced with <PD>.

are used. Lastly for 2018 data the samples from Run2018A-B are used. Good events, where the status of all the subdetectors are flagged as 'GOOD' are given in what is called the golden JSON file, and for each of the data taking periods are:

- 2016: Cert_271036-284044_13TeV_23Sep2016ReReco_Collisions16_JSON.txt
- 2017: Cert_294927-306462_13TeV_EOY2017ReReco_Collisions17_JSON.txt
- 2018: Cert_314472-325175_13TeV_17SeptEarlyReReco2018ABC_PromptEraD_Collisions18_JSON.txt

The integrated luminosity for each of the three years, when referring to the golden JSONs, came out to 35.922 fb^{-1} ($\pm$2.5%) for 2016, 41.529 fb^{-1} ($\pm$2.3%) for 2017 and 59.74 fb^{-1} ($\pm$2.5%) for 2018. All of the simulated samples were normalized according to these integrated luminosities, along with the theoretical cross section for each process.

5.1.2 Simulation

5.1.2.1 Background

Due to the number of final states used in this analysis, a large number of background processes needed to be considered. The most important of these are the backgrounds that are related to leptonic final states, but the hadronic backgrounds are necessary as well. The background processes simulated were split into 7 groupings of similar processes, given in Tables 5.2, 5.3, 5.4. There were:

- ZDY: Made of the $Z + jets$ and Drell-Yan backgrounds. Both sets of samples are generated at leading order (LO) using the MG5_aMC@NLO v2.2.2 generator (madgraph) [82].

- $W + jet$: Only the $W + jets$ background, as it is one of the dominant backgrounds for this analysis, alongside $t\bar{t} + jets$. Also generated at LO using Madgraph.

- $t\bar{t} + jets$: The second of the two most dominant backgrounds, generated at LO using Madgraph, the final states are separated into di-lepton, single lepton, and then $\hat{H}_T$-binned for inclusive decays.

- Single t: This covers the three separate single top processes corresponding to s-channel, t-channel, and tW decays. The s-, t-channel and leptonic tW processes are simulated using Madgraph, while the inclusive tW samples are simulated using the `POWHEG v1` generator [83].

- VV: Includes the di-boson processes, and single top quark production in association with bosons. The WW, WZ, VH and some of the ZZ processes are generated using Madgraph, while the remainder are generated using `POWHEG v1/2`.

- VVV+: Includes tri-boson processes, top pair production in association with bosons, and other rare processes. All of these are generated using Madgraph, with the exception of $t\bar{t} + H$, which is generated using `POWHEG v1`.

- QCD: Only includes the QCD multijets background. Due to its poor simulation this background is fully data-driven in the fit. The simulated samples were used in the tuning of the regions. These samples were generated using Madgraph.

All of the 2017 and 2018 samples used the NNPDF 3.1 parton distribution function (PDF) set [84], while the 2016 samples used the NNPDF 3.0 PDF set [85]. The parton showering and hadronization was handled by PYTHIA v8.205 [86], using the underlying event tunes CP5 for the 2017 and 2018 backgrounds, CP2 for the 2017 and 2018 signals, and CUETP8M1 for the 2016 signals and backgrounds. The CMS detector is simulated for all samples using the GEANT4 package [87].

process	Sample 2016	sigma * k-factor * ε
ZDY (LO)	DYJetsToLL_M-5to50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	224.2
	DYJetsToLL_M-5to50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	37.2
	DYJetsToLL_M-5to50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	3.581
	DYJetsToLL_M-5to50_HT-600toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.1241
(NLO)	DYJetsToLL_M-50_HT-70to100_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	301.2
	DYJetsToLL_M-50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	147.40 * 1.23
	DYJetsToLL_M-50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	40.99 * 1.23
	DYJetsToLL_M-50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	5.678 * 1.23
	DYJetsToLL_M-50_HT-600to800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.367 * 1.23
	DYJetsToLL_M-50_HT-800to1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.6304 * 1.23
	DYJetsToLL_M-50_HT-1200to2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.1514 * 1.23 * 0.833
	DYJetsToLL_M-50_HT-2500toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.003565 * 1.23 * 1.015
(NLO)	ZJetsToNuNu_HT-100To200_13TeV-madgraph	280.35 * 1.23 * 0.994
	ZJetsToNuNu_HT-200To400_13TeV-madgraph	77.67 * 1.23 * 0.981
	ZJetsToNuNu_HT-400To600_13TeV-madgraph	10.73 * 1.23 * 0.977
	ZJetsToNuNu_HT-600To800_13TeV-madgraph	2.559 * 1.23 * 0.975
	ZJetsToNuNu_HT-800To1200_13TeV-madgraph	1.1796 * 1.23 * 0.916
	ZJetsToNuNu_HT-1200To2500_13TeV-madgraph	0.28833 * 1.23 * 0.880
	ZJetsToNuNu_HT-2500ToInf_13TeV-madgraph	0.006945 * 1.23 * 1.276
$W + jets$ (NLO)	WJetsToLNu_HT-70To100_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1353.0 * 1.21
	WJetsToLNu_HT-100To200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1345 * 1.21 * 0.993
	WJetsToLNu_HT-200To400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	359.7 * 1.21 * 1.002
	WJetsToLNu_HT-400To600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	48.91 * 1.21 * 1.009
	WJetsToLNu_HT-600To800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	12.05 * 1.21 * 1.120
	WJetsToLNu_HT-800To1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	5.501 * 1.21 * 1.202
	WJetsToLNu_HT-1200To2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.329 * 1.21 * 1.332
	WJetsToLNu_HT-2500ToInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.03216 * 1.21 * 4.200
$t\bar{t} + jets$ (NLO)	TTJets_DiLept_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	56.86
	TTJets_SingleLeptFromT_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	114.0
	TTJets_SingleLeptFromTbar_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	114.0
	TTJets_HT-600to800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.65
	TTJets_HT-800to1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.6736
	TTJets_HT-1200to2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.1194
	TTJets_HT-2500toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.001445
single t (NLO)	ST_s-channel_4f_InclusiveDecays_13TeV-amcatnlo-pythia8	10.12
	ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	71.74
	ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	119.7
	ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.85
	ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.85
	ST_tWll_5f_LO_13TeV-MadGraph-pythia8	0.01104
	ST_tWnunu_5f_LO_13TeV-MadGraph-pythia8	0.02122
VV (NLO)	GluGluHToZZTo4L_M125_13TeV_powheg2_JHUgenV6_pythia8	29.99
	VHToNonbb_M125_13TeV_amcatnloFXFX_madspin_pythia8	0.952
	WWTo2L2Nu_13TeV-powheg	12.178
	WWTo4Q_13TeV-powheg	51.723
	WWToLNuQQ_13TeV-powheg	49.997
	WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	10.71
	WZTo1L3Nu_13TeV_amcatnloFXFX_madspin_pythia8	3.033
	WZTo3LNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	4.42965
	WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8	4.42965
	ZZTo2L2Nu_13TeV_powheg_pythia8_ext1	0.564
	ZZTo2L2Q_13TeV_powheg_pythia8	3.22
	ZZTo2Q2Nu_13TeV_amcatnloFXFX_madspin_pythia8	4.04
	ZZTo4Q_13TeV_amcatnloFXFX_madspin_pythia8	6.912
	tZq_ll_4f_13TeV-amcatnlo-pythia8	0.0758
VVV+ (NLO)	TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.2043
	TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.4062
	TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.2529
	TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.5297
	TTGJets_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	3.697
	WWG_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.2147
	WWZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.1651
	WZG_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.04123
QCD multijets (LO)	QCD_HT100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	28060000.0
	QCD_HT200to300_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1710000.0
	QCD_HT300to500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	347500.0
	QCD_HT500to700_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	32060.0
	QCD_HT700to1000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6829.0
	QCD_HT1000to1500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1207.0
	QCD_HT1500to2000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	120.0
	QCD_HT2000toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	25.25

Table 5.2: 2016 Background MC samples simulated using the `102X_mcRun2_asymptotic_v8` global tag, which is used to describe the 2016 detector conditions. Cross sections are given for each sample, where the process column notes the accuracy to which the cross sections are calculated, which is a mix of LO and NLO. Some of the cross sections have multiplicative factors, and for those cases, the first number is the LO cross section, the second number is the NLO k-factor, and the third number, ε, is a scale factor that is used to combine the $\hat{H}_T$ binned samples.

process	Sample 2017	sigma * k-factor * ϵ
ZDY (LO)	DYJetsToLL_M-4to50_HT-100to200_TuneCP5_13TeV-madgraphMLM-pythia8	204
	DYJetsToLL_M-4to50_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8	54.39
	DYJetsToLL_M-4to50_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8	5.697
	DYJetsToLL_M-4to50_HT-600toInf_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8	1.85
(NLO)	DYJetsToLL_M-50_HT-70to100_TuneCP5_13TeV-madgraphMLM-pythia8	169.9 * 1.23
	DYJetsToLL_M-50_HT-100to200_TuneCP5_13TeV-madgraphMLM-pythia8	161.1 * 1.23
	DYJetsToLL_M-50_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8	48.66 * 1.23 * 0.999
	DYJetsToLL_M-50_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8	6.968 * 1.23 * 0.990
	DYJetsToLL_M-50_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8	1.743 * 1.23 * 0.975
	DYJetsToLL_M-50_HT-800to1200_TuneCP5_13TeV-madgraphMLM-pythia8	0.6304 * 1.23
	DYJetsToLL_M-50_HT-1200to2500_TuneCP5_13TeV-madgraphMLM-pythia8	0.1933 * 1.23 * 0.833
	DYJetsToLL_M-50_HT-2500toInf_TuneCP5_13TeV-madgraphMLM-pythia8	0.003468 * 1.23 * 1.015
(NLO)	ZJetsToNuNu_HT-100To200_13TeV-madgraph	280.35 * 1.23 * 0.994
	ZJetsToNuNu_HT-200To400_13TeV-madgraph	77.67 * 1.23 * 0.981
	ZJetsToNuNu_HT-400To600_13TeV-madgraph	10.73 * 1.23 * 0.977
	ZJetsToNuNu_HT-600To800_13TeV-madgraph	2.559 * 1.23 * 0.975
	ZJetsToNuNu_HT-800To1200_13TeV-madgraph	1.1796 * 1.23 * 0.916
	ZJetsToNuNu_HT-1200To2500_13TeV-madgraph	0.28833 * 1.23 * 0.880
	ZJetsToNuNu_HT-2500ToInf_13TeV-madgraph	0.006945 * 1.23 * 1.276
$W + jets$ (NLO)	WJetsToLNu_HT-70To100_TuneCP5_13TeV-madgraphMLM-pythia8	1353.0 * 1.21
	WJetsToLNu_HT-100To200_TuneCP5_13TeV-madgraphMLM-pythia8	1395.0 * 1.21 * 0.993
	WJetsToLNu_HT-200To400_TuneCP5_13TeV-madgraphMLM-pythia8	407.9 * 1.21 * 1.002
	WJetsToLNu_HT-400To600_TuneCP5_13TeV-madgraphMLM-pythia8	57.48 * 1.21 * 1.009
	WJetsToLNu_HT-600To800_TuneCP5_13TeV-madgraphMLM-pythia8	12.87 * 1.21 * 1.120
	WJetsToLNu_HT-800To1200_TuneCP5_13TeV-madgraphMLM-pythia8	5.366 * 1.21 * 1.202
	WJetsToLNu_HT-1200To2500_TuneCP5_13TeV-madgraphMLM-pythia8	1.074 * 1.21 * 1.332
	WJetsToLNu_HT-2500ToInf_TuneCP5_13TeV-madgraphMLM-pythia8	0.008001 * 1.21 * 4.200
$t\bar{t} + jets$ (NLO)	TTJets_DiLept_TuneCP5_13TeV-madgraphMLM-pythia8	54.23
	TTJets_SingleLeptFromT_TuneCP5_13TeV-madgraphMLM-pythia8	109.6
	TTJets_SingleLeptFromTbar_TuneCP5_13TeV-madgraphMLM-pythia8	114.0
	TTJets_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8	1.65
	TTJets_HT-800to1200_TuneCP5_13TeV-madgraphMLM-pythia8	0.6736
	TTJets_HT-1200to2500_TuneCP5_13TeV-madgraphMLM-pythia8	0.1194
	TTJets_HT-2500toInf_TuneCP5_13TeV-madgraphMLM-pythia8	0.001445
single t (NLO)	ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-madgraph-pythia8	3.36
	ST_s-channel_4f_hadronicDecays_TuneCP5_13TeV-amcatnlo-pythia8	11.24
	ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powhegV2-mapin-pythia8	119.7
	ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powhegVmadspin-pythia8	71.74
	ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythi	35.85
	ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-p	35.85
	ST_tWll_5f_LO_TuneCP5_PSweights_13TeV-madgraph-pythia8	0.01104
VV (NLO)	WWTo2L2Nu_NNPDF31_TuneCP5_13TeV-powheg-pythia8	12.178
	WWToLNuQQ_NNPDF31_TuneCP5_13TeV-powheg-pythia8	49.997
	ZZTo2L2Nu_TuneCP5_13TeV_powheg_pythia8	0.564
	ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.22
	ZZTo2Q2Nu_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8	4.04
	ZZTo4L_TuneCP5_13TeV_powheg_pythia8	1.212
	WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	10.71
	WZTo1L3Nu_13TeV_amcatnloFXFX_madspin_pythia8	3.033
	WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.606
	WZTo3LNu_TuneCP5_13TeV-amcatnloFXFX-pythia8	4.42965
	tZq_ll_4f_ckm_NLO_TuneCP5_13TeV-madgraph-pythia8	0.07358
	VHToNonbb_M125_13TeV_amcatnloFXFX_madspin_pythia8	0.952
	VHToGG_M125_13TeV_amcatnloFXFX_madspin_pythia8	2.162
	GluGluHToZZTo4L_M125_13TeV_powheg2_JHUGenV7011_pythia8	29.99
VVV+ (NLO)	ttHTobb_M125_TuneCP5_13TeV-powheg-pythia8	0.5269
	ttHToNonbb_M125_TuneCP5_13TeV-powheg-pythia8	0.5638
	TTWJetsToLNu_TuneCP5_PSweights_13TFXFX-madspin-pythia8	0.2043
	TTWJetsToQQ_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	0.4062
	TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8	0.2529
	TTZToQQ_TuneCP5_13TeV-amcatnlo-pythia8	0.5297
	TTGJets_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	3.697
	WWW_4F_TuneCP5_13TeV-amcatnlo-pythia8	0.2086
	WWG_TuneCP5_13TeV-amcatnlo-pythia8	0.2147
	WZG_TuneCP5_13TeV-amcatnlo-pythia8	0.04123
	WWZ_TuneCP5_13TeV-amcatnlo-pythia8	0.1651
	WZZ_TuneCP5_13TeV-amcatnlo-pythia8	0.05565
	ZZZ_TuneCP5_13TeV-amcatnlo-pythia8	0.01398
	TTTT_TuneCP5_PSweights_13TeV-amcatnlo-pythia8	0.009103
QCD multijets (LO)	QCD_HT100to200_TuneCP5_13TeV-madgraphMLM-pythia8	28060000.0
	QCD_HT200to300_TuneCP5_13TeV-madgraphMLM-pythia8	1710000.0
	QCD_HT300to500_TuneCP5_13TeV-madgraphMLM-pythia8	347500.0
	QCD_HT500to700_TuneCP5_13TeV-madgraphMLM-pythia8	32060.0
	QCD_HT700to1000_TuneCP5_13TeV-madgraphMLM-pythia8	6829.0
	QCD_HT1000to1500_TuneCP5_13TeV-madgraphMLM-pythia8	1088.0
	QCD_HT1500to2000_TuneCP5_13TeV-madgraphMLM-pythia8	120.0
	QCD_HT2000toInf_TuneCP5_13TeV-madgraphMLM-pythia8	25.25

Table 5.3: 2017 Background MC samples simulated using the `102X_mc2017_realistic_v8` global tag, which is used to describe the 2017 detector conditions. Cross sections are given for each sample, where the process column notes the accuracy to which the cross sections are calculated, which is a mix of LO and NLO. Some of the cross sections have multiplicative factors, and for those cases, the first number is the LO cross section, the second number is the NLO k-factor, and the third number, ϵ, is a scale factor that is used to combine the $\hat{H}_T$ binned samples.

process	Sample 2018	sigma * k-factor * ε
ZDY (LO)	DYJetsToLL_M-4to50_HT-70to100_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	145.5
	DYJetsToLL_M-4to50_HT-100to200_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	204
	DYJetsToLL_M-4to50_HT-200to400_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	54.39
	DYJetsToLL_M-4to50_HT-400to600_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	5.697
(NLO)	DYJetsToLL_M-50_HT-70to100_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	169.9 * 1.23
	DYJetsToLL_M-50_HT-100to200_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	161.1 * 1.23
	DYJetsToLL_M-50_HT-200to400_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	48.66 * 1.23 * 0.999
	DYJetsToLL_M-50_HT-400to600_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	6.968 * 1.23 * 0.990
	DYJetsToLL_M-50_HT-600to800_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	1.743 * 1.23 * 0.975
	DYJetsToLL_M-50_HT-800to1200_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	0.6304 * 1.23
	DYJetsToLL_M-50_HT-1200to2500_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	0.1933 * 1.23 * 0.833
	DYJetsToLL_M-50_HT-2500toInf_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	0.003468 * 1.23 * 1.015
(NLO)	ZJetsToNuNu_HT-100To200_13TeV-madgraph	280.35 * 1.23 * 0.994
	ZJetsToNuNu_HT-200To400_13TeV-madgraph	77.67 * 1.23 * 0.981
	ZJetsToNuNu_HT-400To600_13TeV-madgraph	10.73 * 1.23 * 0.977
	ZJetsToNuNu_HT-600To800_13TeV-madgraph	2.559 * 1.23 * 0.975
	ZJetsToNuNu_HT-800To1200_13TeV-madgraph	1.1796 * 1.23 * 0.916
	ZJetsToNuNu_HT-1200To2500_13TeV-madgraph	0.28833 * 1.23 * 0.880
	ZJetsToNuNu_HT-2500ToInf_13TeV-madgraph	0.006945 * 1.23 * 1.276
$W + jets$ (NLO)	WJetsToLNu_HT-70To100_TuneCP5_13TeV-madgraphMLM-pythia8	1353.0 * 1.21
	WJetsToLNu_HT-100To200_TuneCP5_13TeV-madgraphMLM-pythia8	1395.0 * 1.21 * 0.993
	WJetsToLNu_HT-200To400_TuneCP5_13TeV-madgraphMLM-pythia8	407.9 * 1.21 * 1.002
	WJetsToLNu_HT-400To600_TuneCP5_13TeV-madgraphMLM-pythia8	57.48 * 1.21 * 1.009
	WJetsToLNu_HT-600To800_TuneCP5_13TeV-madgraphMLM-pythia8	12.87 * 1.21 * 1.120
	WJetsToLNu_HT-800To1200_TuneCP5_13TeV-madgraphMLM-pythia8	5.366 * 1.21 * 1.202
	WJetsToLNu_HT-1200To2500_TuneCP5_13TeV-madgraphMLM-pythia8	1.074 * 1.21 * 1.332
	WJetsToLNu_HT-2500ToInf_TuneCP5_13TeV-madgraphMLM-pythia8	0.008001 * 1.21 * 4.200
$t\bar{t} + jets$ (NLO)	TTJets_DiLept_TuneCP5_13TeV-madgraphMLM-pythia8	54.23
	TTJets_SingleLeptFromT_TuneCP5_13TeV-madgraphMLM-pythia8	109.6
	TTJets_SingleLeptFromTbar_TuneCP5_13TeV-madgraphMLM-pythia8	114.0
	TTJets_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8	1.65
	TTJets_HT-800to1200_TuneCP5_13TeV-madgraphMLM-pythia8	0.6736
	TTJets_HT-1200to2500_TuneCP5_13TeV-madgraphMLM-pythia8	0.1194
	TTJets_HT-2500toInf_TuneCP5_13TeV-madgraphMLM-pythia8	0.001445
single t (NLO)	ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-madgraph-pythia8	3.36
	ST_s-channel_4f_hadronicDecays_TuneCP5_13TeV-madgraph-pythia8	11.24
	ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	119.7
	ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	71.74
	ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	35.85
	ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	35.85
	ST_tWll_5f_LO_TuneCP5_PSweights_13TeV-madgraph-pythia8	0.01104
	ST_tWnunu_5f_LO_TuneCP5_PSweights_13TeV-madgraph-pythia8	0.02122
VV (NLO)	WWTo2L2Nu_NNPDF31_TuneCP5_13TeV-powheg-pythia8	12.178
	WWToLNuQQ_NNPDF31_TuneCP5_13TeV-powheg-pythia8	49.997
	ZZTo2L2Nu_TuneCP5_13TeV_powheg_pythia8	0.564
	ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.22
	ZZTo2Q2Nu_TuneCP5_13TeV_amcatnloFXFX_madspin_pythia8	4.040
	ZZTo4L_TuneCP5_13TeV_powheg_pythia8	1.212
	WZTo1L3Nu_13TeV_amcatnloFXFX_madspin_pythia8	3.033
	WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.606
	WZTo3LNu_TuneCP5_13TeV-amcatnloFXFX-pythia8	4.42965
	tZq_ll_4f_ckm_NLO_TuneCP5_13TeV-madgraph-pythia8	0.07358
	VHToNonbb_M125_13TeV_amcatnloFXFX_madspin_pythia8	0.952
	VHToGG_M125_13TeV_amcatnloFXFX_madspin_pythia8	2.162
	GluGluHToZZTo4L_M125_13TeV_powheg2_JHUGenV7011_pythia8	29.99
VVV+ (NLO)	ttHTobb_M125_TuneCP5_13TeV-powheg-pythia8	0.5269
	ttHToNonbb_M125_TuneCP5_13TeV-powheg-pythia8	0.5638
	TTWJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	0.2043
	TTWJetsToQQ_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	0.4062
	TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8	0.2529
	TTZToQQ_TuneCP5_13TeV-amcatnlo-pythia8	0.5297
	TTGJets_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	3.697
	WWW_4F_TuneCP5_13TeV-amcatnlo-pythia8	0.2086
	WWG_TuneCP5_13TeV-amcatnlo-pythia8	0.2147
	WZG_TuneCP5_13TeV-amcatnlo-pythia8	0.04123
	WWZ_TuneCP5_13TeV-amcatnlo-pythia8	0.1651
	WZZ_TuneCP5_13TeV-amcatnlo-pythia8	0.05565
	ZZZ_TuneCP5_13TeV-amcatnlo-pythia8	0.01398
	TTTT_TuneCP5_13TeV-amcatnlo-pythia8	0.009103
QCD multijets (LO)	QCD_HT100to200_TuneCP5_13TeV-madgraphMLM-pythia8	28060000.0
	QCD_HT200to300_TuneCP5_13TeV-madgraphMLM-pythia8	1710000.0
	QCD_HT300to500_TuneCP5_13TeV-madgraphMLM-pythia8	347500.0
	QCD_HT500to700_TuneCP5_13TeV-madgraphMLM-pythia8	32060.0
	QCD_HT700to1000_TuneCP5_13TeV-madgraphMLM-pythia8	6829.0
	QCD_HT1000to1500_TuneCP5_13TeV-madgraphMLM-pythia8	1088.0
	QCD_HT1500to2000_TuneCP5_13TeV-madgraphMLM-pythia8	120.0
	QCD_HT2000toInf_TuneCP5_13TeV-madgraphMLM-pythia8	25.25

Table 5.4: 2018 Background MC samples simulated using the `102X_upgrade2018_realistic_v21` global tag, which is used to describe the 2018 detector conditions. Cross sections are given for each sample, where the process column notes the accuracy to which the cross sections are calculated, which is a mix of LO and NLO. Some of the cross sections have multiplicative factors, and for those cases, the first number is the LO cross section, the second number is the NLO k-factor, and the third number, ε, is a scale factor that is used to combine the $\hat{H}_T$ binned samples.

5.1.2.2 Signal

The signals used are as described in Section 3.3, with the samples used given in Table 5.5. There are two types of SUSY samples produced by CMS: Fast Simulation (FastSim), which is a special type of simulation used to quickly process a large number of masses, and Full Simulation (FullSim), which is the full simulation process over a select few signal masses. The FullSim process takes a significantly longer amount of time to process events, so it quickly becomes intractable when trying to run over a grid of masses. Due to this, most of the samples used are FastSim, but corrections are applied in order to account for detector modeling issues in the simulation. FullSim samples were only used for the T2-4bd signal, corresponding to two different mass points, with a stop mass of 500 GeV and LSPs of 420 and 490 GeV, and only for 2017 data. The signals are all generated using Madgraph with `PYTHIA` for parton showering and hadronization.

The signal cross sections are calculated assuming that the stop will decay to the lightest neutralino with a 100% branching ratio, ignoring the composition of the neutralino. The cross sections are only dependent on the stop mass, and are kept constant between models with pair production of the stop or sbottom. Cross sections are given in Figure 5.1, as a function of stop masses ranging from 100 to 2000 GeV. A distribution of the mass points that were present in the signals is given in Figure 5.2, for the T2tt, T2bW, and T2cc signal models. For the models that had more than one sample, the samples are labeled in different colors.

Signal Model	Dataset
2016	
T2tt	~~SMS-T2tt_dM-10to80_genHT-160_genMET-80_TuneCUETP8M1_13TeV-madgraphMLM-pythia8~~ ~~SMS-T2tt_dM-10to80_genHT-160_genMET-80_mWMin-0p1_TuneCUETP8M1_13TeV-madgraphMLM-pythia8~~ ~~SMS-T2tt_mStop-400to1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8~~
T2-4bd	~~SMS-T2-4bd_genMET-80_mStop-500_mLSP-490_TuneCUETP8M1_13TeV-madgraphMLM-pythia8~~ ~~SMS-T2-4bd_genMET-80_mStop-500_mLSP-420_TuneCUETP8M1_13TeV-madgraphMLM-pythia8~~
T2cc	SMS-T2cc_genHT-160_genMET-80_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
2017	
T2tt	SMS-T2tt_dM-10to80_genHT-160_genMET-80_TuneCP2_13TeV-madgraphMLM-pythia8 SMS-T2tt_dM-10to80_genHT-160_genMET-80_mWMin-0p1_TuneCP2_13TeV-madgraphMLM-pythia8 SMS-T2tt_mStop-400to1200_TuneCP2_13TeV-madgraphMLM-pythia8
T2bW	SMS-T2bW_TuneCP2_13TeV-madgraphMLM-pythia8 SMS-T2bW_X05_dM-10to80_genHT-160_genMET-80_mWMin-0p1_TuneCP2_13TeV-madgraphMLM-pythia8
T2-4bd	SMS-T2-4bd_genMET-80_mStop-500_mLSP-490_TuneCP2_13TeV-madgraphMLM-pythia8 SMS-T2-4bd_genMET-80_mStop-500_mLSP-420_TuneCP2_13TeV-madgraphMLM-pythia8
T2cc	SMS-T2cc_genHT-160_genMET-80_TuneCP2_13TeV-madgraphMLM-pythia8
2018	
T2tt	SMS-T2tt_dM-10to80_genHT-160_genMET-80_TuneCP2_13TeV-madgraphMLM-pythia8 SMS-T2tt_mStop-400to1200_TuneCP2_13TeV-madgraphMLM-pythia8
T2bW	SMS-T2bW_TuneCP2_13TeV-madgraphMLM-pythia8 SMS-T2bW_X05_dM-10to80_genHT-160_genMET-80_mWMin-0p1_TuneCP2_13TeV-madgraphMLM-pythia8
T2-4bd	~~SMS-T2-4bd_genMET-80_mStop-500_mLSP-490_TuneCP2_13TeV-madgraphMLM-pythia8~~ ~~SMS-T2-4bd_genMET-80_mStop-500_mLSP-420_TuneCP2_13TeV-madgraphMLM-pythia8~~
T2cc	SMS-T2cc_genHT-160_genMET-80_TuneCP2_13TeV-madgraphMLM-pythia8

Table 5.5: 2016, 2017, and 2018 Signal samples for stop masses between 400 and 1500 GeV for T2tt, T2bW, T2-4bd and T2cc SMS models. Some samples for certain years do not have available NanoAOD samples, so for those cases (marked by a strike through) the 2017 MC samples are used as a benchmark, and scaled to the sum of the integrated luminosity expected for the relevant years added together, in order to determine potential expected limits.

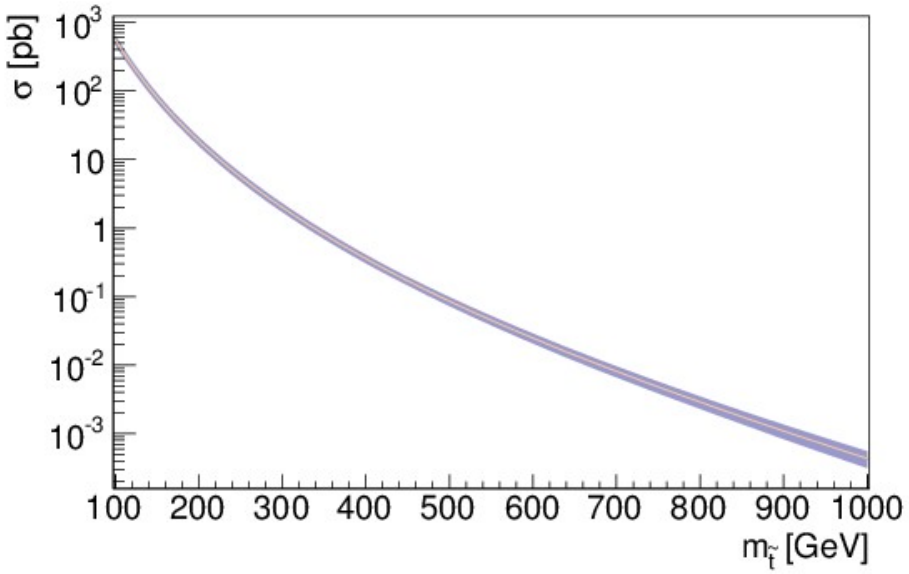
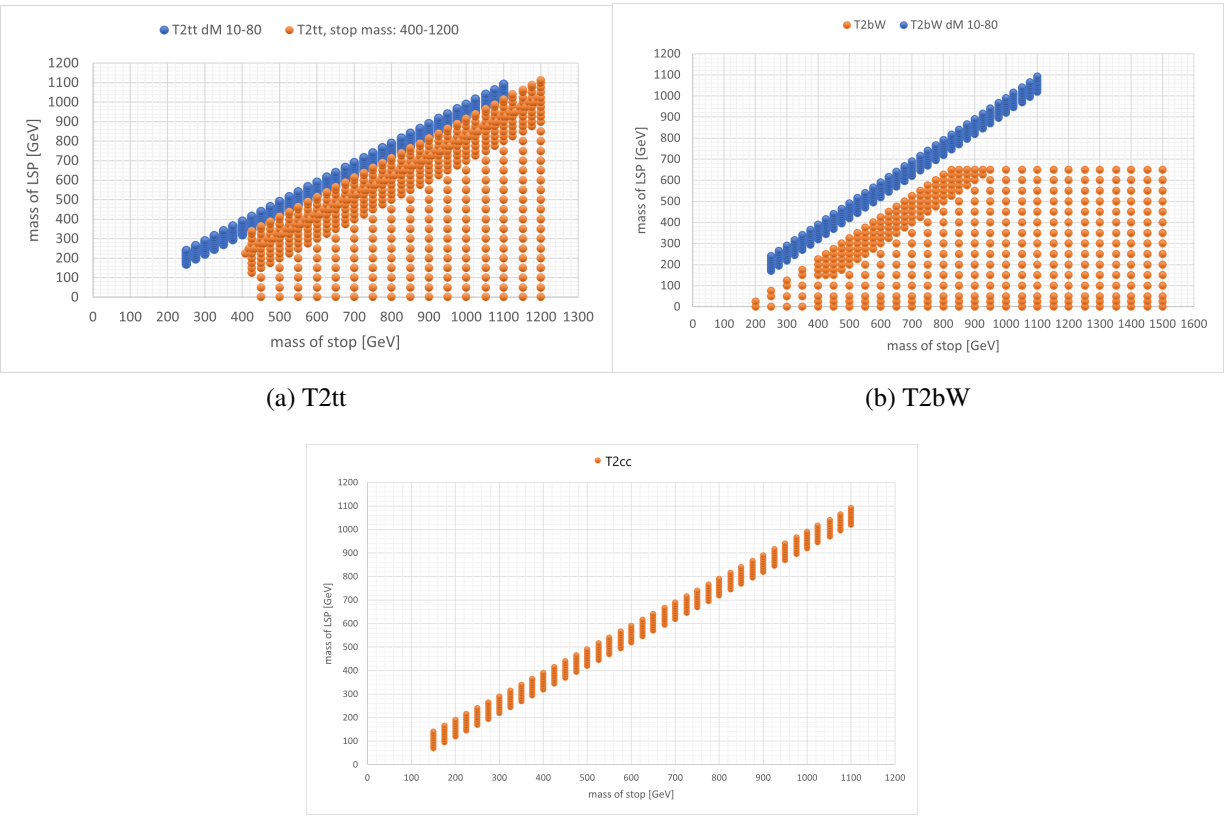


Figure 5.1: Stop cross sections as a function of stop mass, calculated to NLO accuracy. These cross sections are kept constant between signal models which assume the pair production of stops or sbottoms.



(a) T2tt

(b) T2bW

(c) T2cc

Figure 5.2: The mass points present in the T2tt, T2bW, and T2cc stop signals. They are given with the stop mass on the x-axis, and the LSP mass on the y-axis.

5.2 Event Reconstruction

Events are reconstructed using the standard reconstruction algorithms of CMS, which are based on a particle-flow (PF) algorithm [88]. They work by going through and combining the information from all of the subdetectors in CMS in order to identify PF candidates, corresponding to charged and neutral hadrons, photons, and leptons. These PF candidates are then used to reconstruct the objects that are used to define the many final states in the analysis, including the missing transverse momentum.

5.2.1 Triggers

During the data taking, events are categorized based on the firing of triggers, which are flags that specify the contents seen in events and are used as real-time event selection when recording data. There are two trigger systems used at CMS, the Level-1 (L1) trigger, and the High Level Trigger (HLT). The L1 triggers are hardware based triggers used during data taking (online) and implemented in custom-designed Field Programmable Gate Arrays (FPGAs). The HLTs are offline, software based triggers which use streamlined reconstruction software to process events [89].

Events are initially selected in the analysis using a suite of triggers which require large missing transverse momentum ($\not{p}_T$), where $\not{p}_T$ is the negative vector sum of the p_T of all PF candidates. The triggers in general require events with $\not{p}_T > 120$ GeV, and are listed in Table 5.6.

The efficiencies of the triggers were measured as a function of $\not{p}_T$, in data, following Equation 5.1 and compared to the efficiency of the same triggers simulated in MC, shown in Figure 5.3. This figure also shows how the trigger is only 100% efficient at $\not{p}_T$ around 300 GeV. The analysis corrects for the differences in turn on behavior between data and MC by deriving a data to MC scale factor from fits of the turn ons using a Gaussian CDF. This allows the analysis to recover efficiency down to $\not{p}_T > 150$ GeV.

$$\text{Efficiency} = \frac{\text{Events Passing Trigger \& Selection}}{\text{Events Passing Selection}} \tag{5.1}$$

Year	Trigger
2016	HLT_PFMET120_PFMHT120_IDTight HLT_PFMETNoMu120_PFMHTNoMu120_IDTight
2017	HLT_PFMET120_PFMHT120_IDTight HLT_PFMETNoMu120_PFMHTNoMu120_IDTight HLT_PFMET120_PFMHT120_IDTight_PFHT60 HLT_PFMETNoMu120_PFMHTNoMu120IDTight_PFHT60
2018	HLT_PFMET120_PFMHT120_IDTight HLT_PFMETNoMu120_PFMHTNoMu120_IDTight HLT_PFMET120_PFMHT120_IDTight_PFHT60 HLT_PFMETNoMu120_PFMHTNoMu120IDTight_PFHT60

Table 5.6: List of MET triggers used in the analysis. Some triggers are used for all three years, while others are specific to 2017 and 2018.

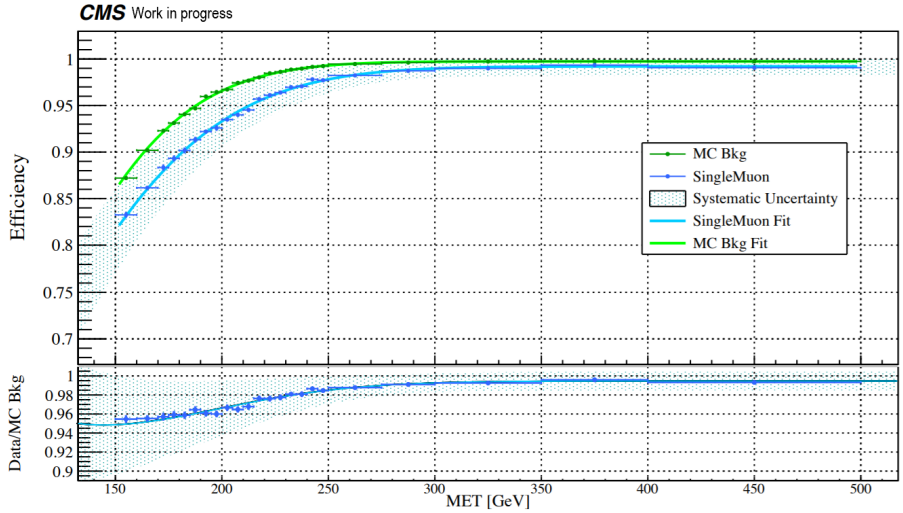


Figure 5.3: Plot of the trigger turn on, for the Single Muon primary dataset, and total MC background. Overlayed are fits which try to model the turn on, in order to later recover efficiency below the plateau.

5.2.2 Leptons

PF candidates which are leptons are initially selected based on some loose criteria, detailed in Table 5.7, with the selection separated between electrons and muons. Selected leptons are then divided into mutually exclusive tiers called gold, silver, and bronze. The highest tier, gold, refers to signal leptons (prompt), silver refers to those that are likely to be from heavy flavor decays, and bronze refers to everything else that passed the initial selection. There are 3 requirements that leptons must pass (or fail) in order to be assigned to the different tiers. The first is the ID requirement, which for muons is to pass the medium working point centrally produced by the Muon object group at CMS [58], with a $p_T > 3.5$ GeV and $|\eta| < 2.4$. For electrons the ID requirement is to pass the tight (loose) MVA ID [76] for $p_T > 10$ ($5 < p_T < 10$) GeV and $|\eta| < 2.4$. The next requirement is that the lepton must be prompt, which means the 3D impact parameter significance is < 2. The last requirement is that the lepton is isolated, which requires that the absolute mini- and PF- isolation, are both < 4 GeV, (MiniIso$_{\text{abs}}$ and PFIso$_{\text{abs}} < 4$ GeV). The two isolation requirements are based on two different algorithms, where for mini-isolation [90, 91], isolation is computed using the scalar sum of the p_T of the PF candidates within a varying cone size based on the lepton p_T, and for PF-isolation [92], isolation is computed using the scalar sum of the p_T of the charged hadrons from the PV, and the scalar sum of the E_T of the neutral hadrons and photons, mitigated by the charged hadron contribution from pileup (referred to as a $\Delta\beta$ correction), all within a cone size (ΔR) of 0.4. The gold tier required that the leptons pass the ID, were prompt, and were isolated. The silver tier required that the leptons pass the ID, were not prompt, but were still isolated. The bronze tier was then the remaining leptons that were not classified as gold or silver.

Cut	Electron	Muon
p_T	> 5 GeV	> 3 GeV
$\|\eta\|$	< 2.4	< 2.4
IVF $\text{IP}_{3D}/\sigma_{\text{IP}_{3D}}$	< 8	< 8
$\|d_{xy}\|$	< 0.05 cm	< 0.05 cm
$\|d_z\|$	< 0.1 cm	< 0.1 cm
$\text{PFIso}_{\text{abs}}$	$< 20 + (300/p_T)$ GeV	$< 20 + (300/p_T)$ GeV
MVA Id	Very Loose	None

Table 5.7: The base lepton selection, split between muons and electrons. $\text{PFIso}_{\text{abs}}$ is a basic isolation criteria, and MVA Id is a lepton Id centrally produced by the CMS electron physics object group, for selecting electrons.

5.2.3 Jets and secondary vertices

Jets were reconstructed from PF candidates using the infrared and collinear safe anti-k_T algorithm [93], and a cone size of $\Delta R = 0.4$. Charged hadron contributions not from the leading primary vertex were removed to mitigate pileup effects. Current jet energy corrections were applied, and the jets in MC were smeared following the recommended jet energy resolution scale factor. The tight jet ID [94] was applied, which rejects fake jets arising from detector and electronics noise. Jets were required to have a $p_T > 20$ GeV and $|\eta| < 5$. Jets that passed the medium working point of the DeepJet tagger [95] were labeled as b-jets. An additional selection required that jets not be matched to a lepton within a cone size of $\Delta R = 0.2$.

Secondary vertices were reconstructed using the Inclusive secondary Vertex Finder (IVF) [96]. These SVs were required to have a p_T between 2 and 20 GeV, and a soft b-tagging discriminator cut of > 0.3, as discussed in the previous chapter. Additionally, the SV could not be matched to any jet with a p_T above 20 GeV within a cone size of $\Delta R = 0.4$, or any lepton within a cone size of $\Delta R = 0.2$.

5.2.4 MET

The missing transverse energy ($\not{p}_T$) is the negative vector sum of the transverse momentum of all PF candidates in the event. Type-1 corrections [97] are applied to $\not{p}_T$, in which the jet energy

corrections applied to the jets are propagated to $\not{p}_T$. The $\not{p}_T$ is also corrected in 2017 data to account for EE noise present during the data taking period. Additionally MET filters [98]. were applied as recommended in order to remove false $\not{p}_T$, including the primary vertex filter, the 2016 CSC beam halo filter, the HBHE noise filter, the HBHEiso noise filter, the ECAL dead cell trigger primitive filter, and the bad PF muon filter.

5.3 Event selection

After passing the trigger selection, events were required to have a primary vertex (PV), which passes the selection:

- $\chi^2 > 0$, is real
- $N_{dof} > 5$, vertex number of degrees of freedom
- $|z| <= 24$ cm, beam spot position along the z direction
- $\rho < 4$ cm, the displacement from the beam spot position along the transverse direction

Of the PVs that pass this selection, the one with the largest value of the summed p_T^2 of the reconstructed objects was taken as the PV of the event. Every event was also required to have either the sum of the number of leptons, jets and SVs to be ≥ 2, or to have at least 1 jet in the event. In the following section, further selections were applied in order to clean events before moving to defining the regions

5.3.1 Preselection

First, event cuts are applied, including:

- $\not{p}_T > 150$ GeV, based on the trigger turn on
- $p_T^{ISR} > 250$ GeV, to provide a moderate ISR kick to resolve compressed signal masses

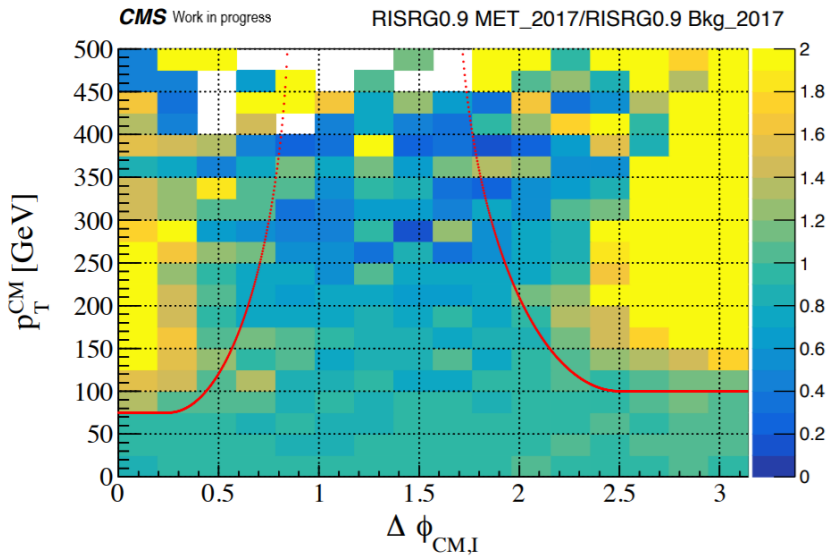


Figure 5.4: A distribution of $\Delta\phi_{CM,I}$ versus p_T^{CM} for the 2017 MET PD for $R_{ISR} > 0.9$ divided by the total background from 2017 samples for $R_{ISR} > 0.9$. Overlayed in red is the 2D cut applied to these two variables, where the central area is kept. The yellow areas are where the poorly modeled events occur.

- $R_{ISR} > 0.5$, to focus on signals with a significant mass ratio
- $|\Delta\phi_{\vec{E}_T,V}| < \pi/2$, to ensure the Visible and Invisible systems are pointing in the same direction
- 2D cut in $\Delta\phi_{CM,I}$ and p_T^{CM}, following Figure 5.4, to remove poorly modeled events

5.4 Event Categorization

The analysis regions were then defined according to a combination of the multiplicity of reconstructed objects in the S and ISR systems, and the kinematics of the remaining mass sensitive variables described in Section 3.1.2. The full list of categories is summarized in Table 5.8, with their naming convention for the full region names. The $M_\perp$ and R_{ISR} binning is summarized in Table 5.9 for the different S system categories.

The first set of categorizations was the multiplicity counting of reconstructed objects in the S system. The categorizations were divided into number of leptons (NL), jets (NJ), b-jets (NJ_{Nb}), and SVs (NJ^{NSV}). Depending on the region, the leptons were always split into flavor (NL^e, NL^μ)

and sometimes split into charge ($NL^{e\pm}$, $NL^{\mu\pm}$), and the SVs were split into central ($|\eta| < 1.5$) and forward ($|\eta| \geq 1.5$) η regions (NJ^{NSVc}, NJ^{NSVf}). These categorizations were used to build the different final states, allowing a range of signals and final states to be studied.

The next set of categorizations was related to the ISR system, where depending on the object multiplicities, events were separated into different p_T^{ISR} categories, instead of just a lower p_T^{ISR} cut, as the major backgrounds sit at lower values of p_T^{ISR}. For higher multiplicities of objects, the category boundaries are adjusted upwards, resulting in 200, 250, 300, 350, 400, 450, and 500 GeV lower bound cuts for the p_T^{ISR} categories, depending on the number of jets and leptons. An additional set of ISR categories was based on the existence of a b jet in the ISR system. This categorization was used to produce $t\bar{t} + jets$ enriched regions. It was found that for events with a b jet in the ISR system, for both 1 and 2 lepton categories, a much larger fraction of b jets were found in the ISR system than not, which is shown in Figure 5.5. It was also seen that the main backgrounds became more suppressed compared to signal when requiring SVs and b jets in the S system.

There was then a categorization involving whether the regions should be split into two γ_T categories. For the regions where the two categories are used, the boundaries correspond to $\gamma_T < 0.5$ and $\gamma_T > 0.5$. Figure 5.6 shows γ_T for a 500, 490 GeV T2bW stop signal and the total background, not including QCD multijets. In this variable, the events above 0.5 have more balanced events, which is more indicative of a sparticle decay. In the figure, it can be seen that the signal peaks just above $\gamma_T = 0.5$, while the total background is peaking around 0.4, showing that the majority of the background events are more unbalanced than the signal.

Last, is the 2D binning in $M_\perp$ and R_{ISR}, which for the final region is unrolled into a 1D histogram of all of the bins. The purpose of this binning is to isolate the compressed signals, where the more compressed ones will generally congregate at high R_{ISR} and low $M_\perp$, with the backgrounds generally being more dispersed over the 2D plane. There are multiple sets of binning, generally corresponding to the number of objects in any given region. For example, regions with 1 lepton and 1 jet would have the same binning as regions with 0 leptons and 2 jets. This was due to the

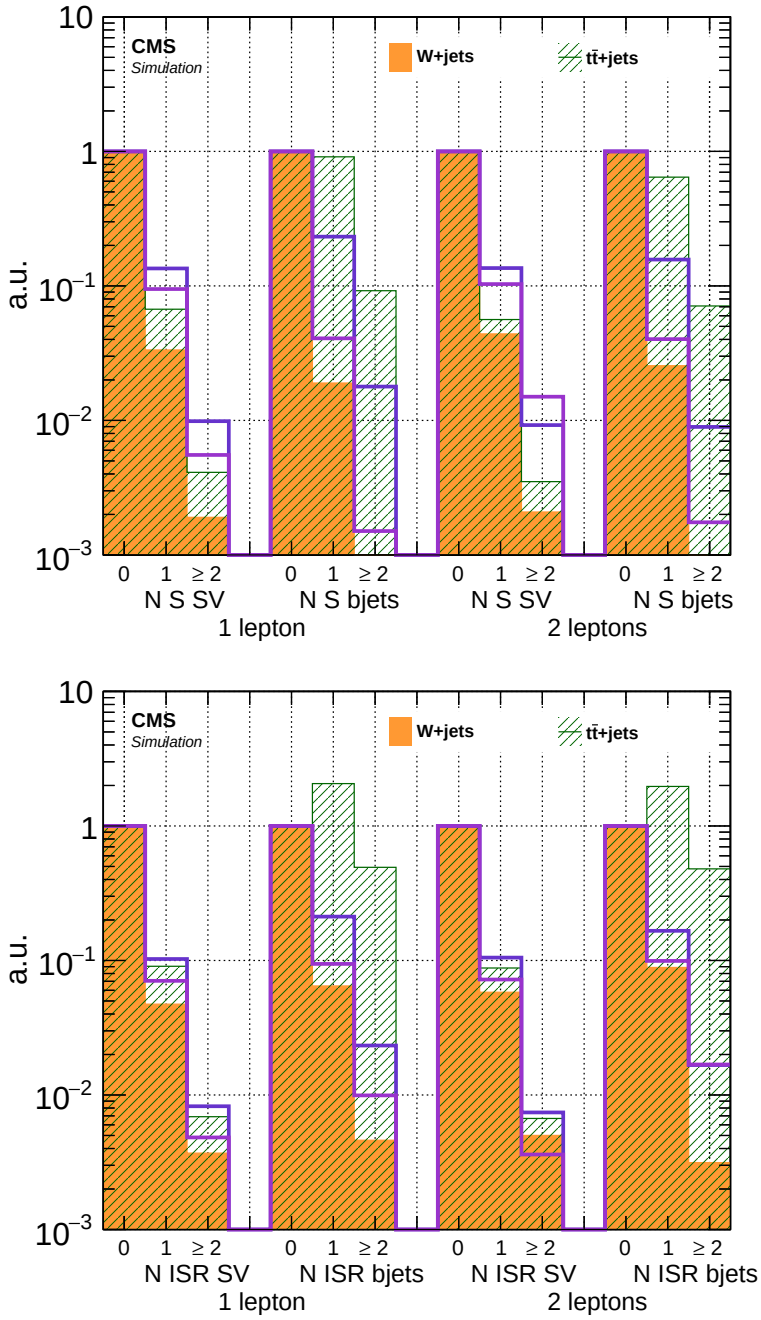


Figure 5.5: Fractional distributions of the number of S (top) and ISR (bottom) SVs and b jets for 1 and 2 golden lepton selections. These plots show how the SVs and b jets are distributed between the S and ISR systems for the two main backgrounds ($t\bar{t} + jets$ and $W + jets$), and the two T2-4bd samples with a stop mass of 500 GeV and LSP masses of 490 (dark purple) and 420 (light purple). The distributions are scaled by the count of the 0 bins for each set of object counting. This means that the 1 and ≥ 2 count bins are fractional values with respect to the 0 count bins.

regions with the same number of objects having similar kinematic scales. The 2D binning for $M_\perp$ and R_{ISR} is given in Table 5.9.

5.4.1 Region Definitions

The regions in the analysis were put together to cover a wide range of final states with varying mass splittings in the signals. This large number of final states allowed for a generic search for signals which follow similar decay trees, with the generic one used by this analysis given in Figure 3.1.

Every region in the analysis is a combination of the categorizations listed in Table 5.8, which culminates in a 1D histogram that contains the $M_\perp$ and R_{ISR} bins. The region definitions were chosen based on signal to background studies for a set of mass splittings, while also trying to define distinctive final states that match different processes, whether signal or background. For example, a $1L-0J$ categorization would be dominantly $W+jets$, and would have larger event counts for the more compressed signals. A region with this categorization would then provide some signal bins for more compressed signals and then $W+jets$ control bins for the less compressed.

The base set of regions are those that have 0, 1, 2, or 3 leptons and up to 5 or more jets, up to 2 b-jets, and up to 2 or more SVs. The full list of regions is given in Tables 5.10, 5.11, 5.12, and 5.13 for the 0, 1, 2, and 3 gold lepton final states, respectively. The regions with silver and bronze leptons are similar to the gold versions, with the exception that all of the categorizations were condensed, meaning that the leptons were only flavor separated, the SVs did not have an η separation, there was no γ_T categorization, and the ISR, and p_T^{ISR} categorizations were set to be inclusive.

An important part of the region definitions was defining which $M_\perp$-R_{ISR} bins were control and which were signal. The problem, as it pertains to this analysis, was that what will be a control bin for one signal, could very well be a signal bin for another. In order to look at the SV tagging in real data, with regards to the fit, a set of control bins were defined which pertain only to the overall SV regions and their inclusive counterpart. With the small portion of phase space that the SV regions

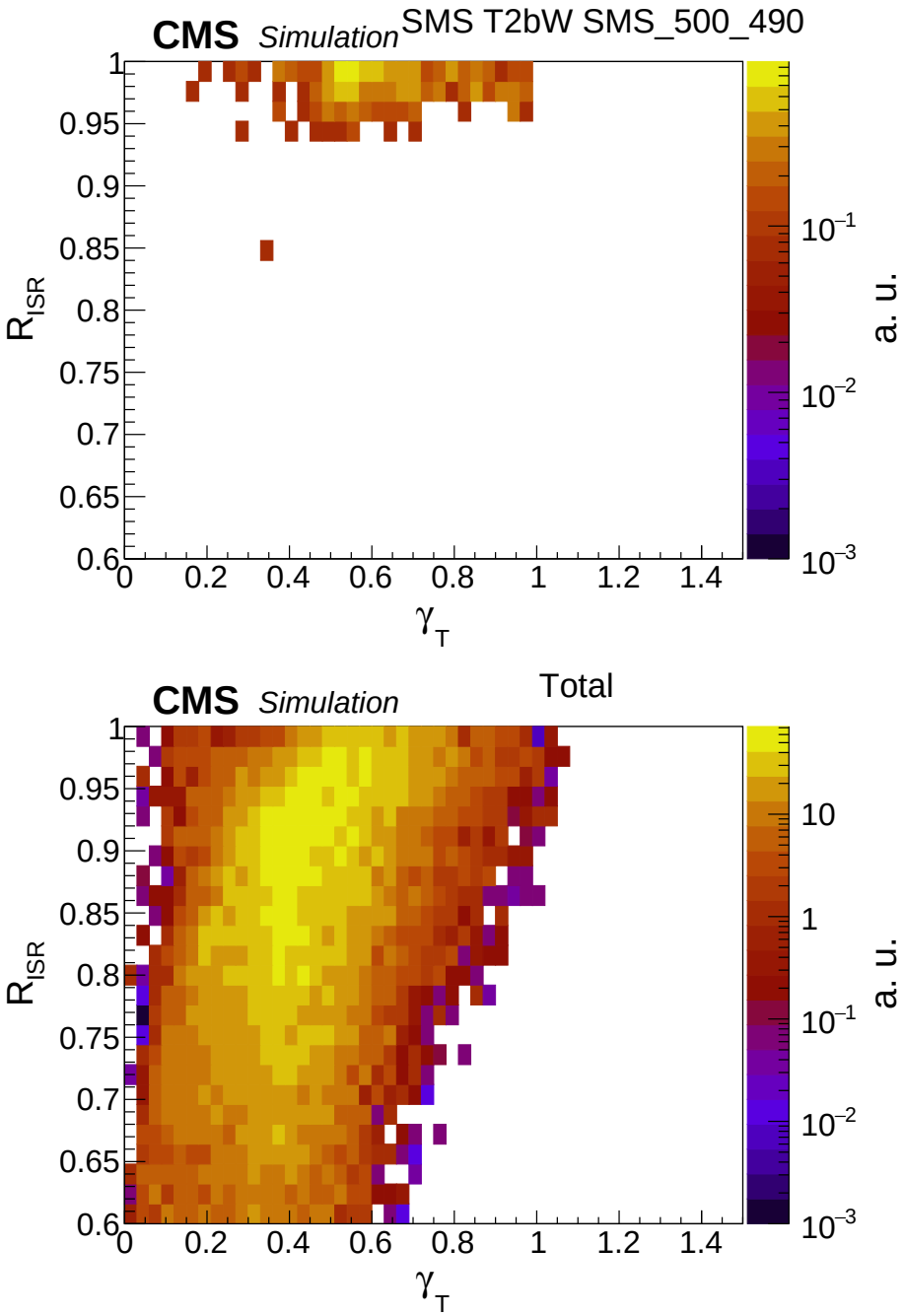


Figure 5.6: Distribution of γ_T for a T2bW stop signal with masses 500, 490 (top) and the total background, not including QCD (bottom). This is for a $1L_G - 0J_{0b}^{g1SVc} - X_{0b}^{P_{300}}$ region. Larger values of γ_T represent a more balanced event in the S system, where smaller values represent an imbalance.

Region Objects	Category Designations	Category
leptons	$0L$	no leptons
	NL_G; NL_S; NL_B	N gold; N silver; N bronze
	L^{ℓ}; $L^{\ell\pm}$	no charge sep.; with charge $\pm$
	$2L^{\ell\ell OS}$; $2L^{\ell\ell SS}$	2 ℓ, OS; SS
	$2L^{\ell\ell\|}$	2 ℓ in P_a
	$2L^{\ell\|\ell}$	2 ℓ split between P_a and P_b
	$3L^{\ell\ell\|\ell OSSF}$; $3L^{\ell\ell\|\ell nOSSF}$	3 ℓ, OSSF in P_a; No OSSF in P_a
	$3L^{\ell\ell\ell OSSF}$; $3L^{\ell\ell\ell nOSSF}$	3 ℓ, OSSF inclusive; No OSSF
S jets	NJ	$N \in \{0,1,2,3,4\}$ jets
	gNJ	$\geq N \in \{1,3,4,5\}$ jets
S b-jets	J	≥ 0 b-jets (inclusive)
	J_{Nb}	$N \in \{0,1,2\}$ b-jets
	J_{gNb}	$\geq N \in 1,2$ b-jets
S SVs	J	≥ 0 SVs (inclusive)
	J^{NSVc}	$N \in \{0,1\}$ central η SVs
	J^{NSVf}	$N \in \{0,1\}$ forward η SVs
	J^{gNSVc}	$\geq N \in \{1,2\}$ central η SVs
	J^{gNSVf}	$\geq N \in \{1,2\}$ forward η SVs
p_T^{ISR}	$X^{P_{250}}$	> 250 GeV
	$X^{P_{350}}$	> 350 GeV
	$X^{P_{400}}$	> 400 GeV
	$X^{P_{500}}$	> 500 GeV
	$X^{P_{550}}$	> 550 GeV
	$X^{P_{250}^{350}}$	$250 < p_T^{ISR} \leq 350$
	$X^{P_{350}^{500}}$	$350 < p_T^{ISR} \leq 500$
	$X^{P_{400}^{550}}$	$400 < p_T^{ISR} \leq 540$
ISR b-jet counting	X	≥ 0 b-jets (inclusive)
	X_{0b}	$= 0$ b-jets
	X_{1b}	≥ 1 b-jets
γ_T	X	inclusive
	$X^{\gamma\downarrow}$	$\gamma_T < 0.5$
	$X^{\gamma\uparrow}$	$\gamma_T > 0.5$

Table 5.8: Table detailing the different categorizations present in the analysis regions. The category designations are the labels that will be used in the region naming scheme, for the leptons ℓ can be replaced with either e or μ. Additionally if charge is added to the designation, then that means the region is charge separated in the fit. The variables P_a and P_b refer to the two hemispheres of the sparticle system, as seen in the decay tree in Figure 3.1. The X in the table refers to extra event kinematic and ISR selections as described in Section 5.4. Each line in the table corresponds to a separate category, and a combination of one category from each object row constitutes a region.

S System Category	Bin Boundaries
$0L-0J$	$\{R_{ISR}$ bin$\} \ni \{M_\perp$ bin edges$\}$ $\{0.95, 0.985\} \ni \{0, 40\}$ $\{0.985, 1.00\} \ni \{0, 5, 10, 40\}$
$1L-0J$	$\{0.9, 0.96\} \ni \{0, 40\}$ $\{0.96, 0.98\} \ni \{0, 10, 40\}$ $\{0.98, 1.00\} \ni \{0, 5, 10, 40\}$
$0L-2J$ $1L-1J$	$\{0.65, 0.75\} \ni \{0, 50, 120\}$ $\{0.75, 0.85\} \ni \{0, 40, 120\}$ $\{0.85, 0.9\} \ni \{0, 30, 120\}$ $\{0.9, 0.95\} \ni \{0, 20, 120\}$ $\{0.95, 1.00\} \ni \{0, 120\}$
$0L-3J$ $1L-2J$	$\{0.55, 0.65\} \ni \{0, 110, 200\}$ $\{0.65, 0.75\} \ni \{0, 90, 200\}$ $\{0.75, 0.85\} \ni \{0, 70, 200\}$ $\{0.85, 0.9\} \ni \{0, 50, 200\}$ $\{0.9, 1.00\} \ni \{0, 200\}$
$0L-4J$ $1L-3J$	$\{0.55, 0.65\} \ni \{0, 150, 300\}$ $\{0.65, 0.75\} \ni \{0, 100, 300\}$ $\{0.75, 0.85\} \ni \{0, 80, 300\}$ $\{0.85, 1.0\} \ni \{0, 300\}$
$0L-g5J$ $1L-g4J$	$\{0.5, 0.6\} \ni \{0, 210, 400\}$ $\{0.6, 0.7\} \ni \{0, 180, 400\}$ $\{0.7, 0.8\} \ni \{0, 150, 400\}$ $\{0.8, 1.00\} \ni \{0, 400\}$
$2L-0J$	$\{0.7, 0.8\} \ni \{0, 40, 120\}$ $\{0.8, 0.9\} \ni \{0, 30, 120\}$ $\{0.9, 0.95\} \ni \{0, 20, 120\}$ $\{0.95, 1.00\} \ni \{0, 15, 120\}$
$2L-1J$	$\{0.55, 0.7\} \ni \{0, 80, 200\}$ $\{0.7, 0.8\} \ni \{0, 60, 200\}$ $\{0.8, 0.9\} \ni \{0, 40, 200\}$ $\{0.9, 1.00\} \ni \{0, 30200\}$
$2L-2J$	$\{0.5, 0.65\} \ni \{0, 100, 300\}$ $\{0.6, 0.7\} \ni \{0, 120, 300\}$ $\{0.7, 0.8\} \ni \{0, 100, 300\}$ $\{0.8, 0.9\} \ni \{0, 80, 300\}$ $\{0.9, 1.00\} \ni \{0, 300\}$
$2L-g3J$	$\{0.5, 0.65\} \ni \{0, 130, 400\}$ $\{0.65, 0.8\} \ni \{0, 100, 400\}$ $\{0.8, 1.0\} \ni \{0, 400\}$
$3L-0J$	$\{0.6, 0.7\} \ni \{0, 250\}$ $\{0.7, 0.8\} \ni \{0, 250\}$ $\{0.8, 0.9\} \ni \{0, 250\}$ $\{0.9, 1.00\} \ni \{0, 250\}$
$3L-g1J$	$\{0.55, 0.7\} \ni \{0, 250\}$ $\{0.7, 0.85\} \ni \{0, 250\}$ $\{0.85, 1.0\} \ni \{0, 250\}$

Table 5.9: Table detailing the different R_{ISR} versus $M_\perp$ binning present in the analysis regions. The bins are determined by the S system category, corresponding to number of leptons and jets ($NL-NJ$). The boundaries describe how they were binned, giving the R_{ISR} bin, and then the bin edges in $M_\perp$ which belong to that R_{ISR} bin.

0L Regions
$0L-0J^{1SVc}-X^{P_{350}}$
$0L-0J^{1SVf}-X^{P_{350}}$
$0L-1J_{0b}^{0SV}-X_b^{P_{400}^{550}}$
$0L-1J_{1b}^{0SV}-X_b^{P_{400}^{550}}$
$0L-1J^{g1SVc}-X^{P_{400}}$
$0L-1J^{g1SVf}-X^{P_{400}}$
$0L-2J_{0b}-X_b^{\gamma P_{350}^{500}}$
$0L-2J_{1b}-X_b^{\gamma P_{350}^{500}}$
$0L-2J_{2b}-X^{\gamma P_{350}^{500}}$
$0L-3J_{0b}-X_b^{\gamma P_{350}^{500}}$
$0L-3J_{1b}-X_b^{\gamma P_{350}^{500}}$
$0L-3J_{2b}-X^{\gamma P_{350}^{500}}$
$0L-4J_{0b}-X_b^{\gamma P_{300}^{450}}$
$0L-4J_{1b}-X_b^{\gamma P_{300}^{450}}$
$0L-4J_{2b}-X^{\gamma P_{350}^{500}}$
$0L-g5J_{0b}-X_b^{\gamma P_{350}^{500}}$
$0L-g5J_{1b}-X_b^{\gamma P_{350}^{500}}$
$0L-g5J_{g2b}-X^{\gamma P_{350}^{500}}$

Table 5.10: List of the 0 lepton regions ($0L$), counting the number of hadronic objects in the S system (i.e. $1J$), with the event system specific categorization ($X...$). The presence of a γ denotes that this categorization is applied and it counts as 2 regions, one for $X^{\gamma\uparrow}$ and one for $X^{\gamma\downarrow}$. When p_T^{ISR} categories are applied, the presence of two numbers, an upper and lower bound, assumes two regions: $X^{P_{low}^{high}}$ and $X^{P_{high}}$. Lastly, X_b means the ISR b-jet categorization is applied. The R_{ISR}-$M_\perp$ binning used for any given region can be found in Table 5.9, following the S system, $NL-NJ$, categorization.

1L Regions	
e	μ
$1L_G^{e\pm} - 0J_{0b}^{0SV} - X_b^{P_{350}^{500}}$	$1L_G^{\mu\pm} - 0J_{0b}^{0SV} - X_b^{P_{350}^{500}}$
$1L_G^{e\pm} - 0J^{g1SVc/f} - X^{P_{350}}$	$1L_G^{\mu\pm} - 0J^{g1SVc/f} - X^{P_{350}}$
$1L_G^{e\pm} - 0J^{g1SVc/f} - X^{P_{350}}$	$1L_G^{\mu\pm} - 0J^{g1SVc/f} - X^{P_{350}}$
$1L_G^{e\pm} - 1J_{0b}^{0SV} - X_b^{\gamma P_{350}}$	$1L_G^{\mu\pm} - 1J_{0b}^{0SV} - X_b^{\gamma P_{350}}$
$1L_G^{e\pm} - 1J_{1b}^{0SV} - X_b^{\gamma P_{350}}$	$1L_G^{\mu\pm} - 1J_{1b}^{0SV} - X_b^{\gamma P_{350}}$
$1L_G^{e\pm} - 1J^{g1SVc/f} - X^{P_{350}}$	$1L_G^{\mu\pm} - 1J^{g1SVc/f} - X^{P_{350}}$
$1L_G^{e} - 2J_{0b} - X_b^{\gamma P_{350}^{500}}$	$1L_G^{\mu} - 2J_{0b} - X_b^{\gamma P_{350}^{500}}$
$1L_G^{e} - 2J_{1b} - X_b^{\gamma P_{350}^{500}}$	$1L_G^{\mu} - 2J_{1b} - X_b^{\gamma P_{350}^{500}}$
$1L_G^{e} - 2J_{2b} - X^{\gamma P_{350}^{500}}$	$1L_G^{\mu} - 2J_{2b} - X^{\gamma P_{350}^{500}}$
$1L_G^{e} - 3J_{0b} - X_b^{\gamma P_{350}^{500}}$	$1L_G^{\mu} - 3J_{0b} - X_b^{\gamma P_{350}^{500}}$
$1L_G^{e} - 3J_{1b} - X_b^{\gamma P_{350}^{500}}$	$1L_G^{\mu} - 3J_{1b} - X_b^{\gamma P_{350}^{500}}$
$1L_G^{e} - 3J_{g2b} - X^{\gamma P_{350}^{500}}$	$1L_G^{\mu} - 3J_{g2b} - X^{\gamma P_{350}^{500}}$
$1L_G^{e} - g4J_{0b} - X_b^{\gamma P_{350}^{500}}$	$1L_G^{\mu} - g4J_{0b} - X_b^{\gamma P_{350}^{500}}$
$1L_G^{e} - g4J_{g1b} - X^{\gamma P_{350}^{500}}$	$1L_G^{\mu} - g4J_{g1b} - X^{\gamma P_{350}^{500}}$

Table 5.11: List of the Gold 1 lepton regions ($1L$), counting the number of hadronic objects in the S system (i.e. $1J$), with the event system specific categorization (X...). The Silver and Bronze regions are similar, with the exception that the lepton and X categorization, and any p_T^{ISR}, γ_T or SV η categories are made to be inclusive. Additionally, the b-jet regions are also made to be inclusive. For the leptons, the presence of $\pm$ means charge separation is applied, for a total of 2 regions. For the SV regions, c/f means that the η categorization is applied, so it counts as 2 regions. The presence of a γ denotes that this categorization is applied and it counts as 2 regions, one for $X^{\gamma\uparrow}$ and one for $X^{\gamma\downarrow}$. When p_T^{ISR} categories are applied, the presence of two numbers, an upper and lower bound, assumes two regions: $X^{P_{low}^{high}}$ and $X^{P_{high}}$. Lastly, X_b means the ISR b-jet categorization is applied. The R_{ISR}-$M_\perp$ binning used for any given region can be found in Table 5.9, following the S system, $NL-NJ$, categorization.

2L Regions			
$eeOS$	$e\mu OS$	$\mu\mu OS$	SS
$\frac{2L_G^{eeOS} - 0J_{0b}^{0SV} - X_b^{\gamma P_{250}^{350}}}{2L_G^{ee} - 0J^{g1SVc/f} - X^{P_{250}}}$	$\frac{2L_G^{e\mu OS} - 0J_{0b}^{0SV} - X_b^{\gamma P_{250}^{350}}}{2L_G^{e\mu} - 0J^{g1SVc/f} - X^{P_{250}}}$	$\frac{2L_G^{\mu\mu OS} - 0J_{0b}^{0SV} - X_b^{\gamma P_{250}^{350}}}{2L_G^{\mu\mu} - 0J^{g1SVc/f} - X^{P_{250}}}$	$2L_G^{eeSS} - 0J_{0b}^{0SV} - X^{P_{250}}$ $2L_G^{e\mu SS} - 0J_{0b}^{0SV} - X^{P_{250}}$ $2L_G^{\mu\mu SS} - 0J_{0b}^{0SV} - X^{P_{250}}$
$2L_G^{e\|eOS} - 1J_{0b} - X_b^{\gamma P_{250}^{350}}$ $2L_G^{ee\|OS} - 1J_{0b} - X_b^{\gamma P_{250}^{350}}$ $2L_G^{e\|eOS} - 1J_{1b} - X_b^{\gamma P_{250}^{350}}$ $2L_G^{ee\|OS} - 1J_{1b} - X_b^{\gamma P_{250}^{350}}$	$2L_G^{e\mu OS} - 1J_{0b} - X_b^{\gamma P_{250}^{350}}$ $2L_G^{e\mu OS} - 1J_{1b} - X_b^{\gamma P_{250}^{350}}$	$2L_G^{\mu\|\mu OS} - 1J_{0b} - X_b^{\gamma P_{250}^{350}}$ $2L_G^{\mu\mu\|OS} - 1J_{0b} - X_b^{\gamma P_{250}^{350}}$ $2L_G^{\mu\|\mu OS} - 1J_{1b} - X_b^{\gamma P_{250}^{350}}$ $2L_G^{\mu\mu\|OS} - 1J_{1b} - X_b^{\gamma P_{250}^{350}}$	$2L_G^{eeSS} - 1J - X^{P_{250}}$ $2L_G^{e\mu SS} - 1J - X^{P_{250}}$ $2L_G^{\mu\mu SS} - 1J - X^{P_{250}}$
$2L_G^{e\|eOS} - 2J_{0b} - X_b^{\gamma P_{250}^{350}}$ $2L_G^{ee\|OS} - 2J_{0b} - X_b^{\gamma P_{250}^{350}}$ $2L_G^{e\|eOS} - 2J_{g1b} - X_b^{\gamma P_{250}^{350}}$ $2L_G^{ee\|OS} - 2J_{g1b} - X_b^{\gamma P_{250}^{350}}$	$2L_G^{e\mu OS} - 2J_{0b} - X_b^{\gamma P_{250}^{350}}$ $2L_G^{e\mu OS} - 2J_{g1b} - X_b^{\gamma P_{250}^{350}}$	$2L_G^{\mu\|\mu OS} - 2J_{0b} - X_b^{\gamma P_{250}^{350}}$ $2L_G^{\mu\mu\|OS} - 2J_{0b} - X_b^{\gamma P_{250}^{350}}$ $2L_G^{\mu\|\mu OS} - 2J_{g1b} - X_b^{\gamma P_{250}^{350}}$ $2L_G^{\mu\mu\|OS} - 2J_{g1b} - X_b^{\gamma P_{250}^{350}}$	$2L_G^{eeSS} - 2J - X^{\gamma P_{250}}$ $2L_G^{e\mu SS} - 2J - X^{\gamma P_{250}}$ $2L_G^{\mu\mu SS} - 2J - X^{\gamma P_{250}}$
$2L_G^{e\|eOS} - g3J_{0b} - X_b^{\gamma P_{250}^{350}}$ $2L_G^{ee\|OS} - g3J_{0b} - X_b^{\gamma P_{250}^{350}}$ $2L_G^{e\|eOS} - g3J_{g1b} - X_b^{\gamma P_{250}^{350}}$ $2L_G^{ee\|OS} - g3J_{g1b} - X_b^{\gamma P_{250}^{350}}$	$2L_G^{e\mu OS} - g3J_{0b} - X^{\gamma P_{250}^{350}}$ $2L_G^{e\mu OS} - g3J_{g1b} - X^{\gamma P_{250}^{350}}$	$2L_G^{\mu\|\mu OS} - g3J_{0b} - X^{\gamma P_{250}^{350}}$ $2L_G^{\mu\mu\|OS} - g3J_{0b} - X^{\gamma P_{250}^{350}}$ $2L_G^{\mu\|\mu OS} - g3J_{g1b} - X^{\gamma P_{250}^{350}}$ $2L_G^{\mu\mu\|OS} - g3J_{g1b} - X^{\gamma P_{250}^{350}}$	$2L_G^{eeSS} - g3J - X^{\gamma P_{250}}$ $2L_G^{e\mu SS} - g3J - X^{\gamma P_{250}}$ $2L_G^{\mu\mu SS} - g3J - X^{\gamma P_{250}}$

Table 5.12: List of the Gold 2 lepton regions ($2L$), counting the number of hadronic objects in the S system (i.e. $1J$), with the event system specific categorization (X...). The Silver and Bronze regions are similar, with the exception that the lepton and X categorizations are condensed so that leptons are only flavor separated, and any p_T^{ISR}, γ_T or SV η categories are made to be inclusive. Additionally, the b-jet regions are also made to be inclusive. For leptons, $\ell\ell SS$ means same sign charge separation is applied, counting for 2 region, and $\ell\ell OS$ means opposite sign charge separation is applied, counting for 2 regions. For the SV regions, c/f means that the η categorization is applied, so it counts as 2 regions. The presence of a γ denotes that this categorization is applied and it counts as 2 regions, one for $X^{\gamma^{\uparrow}}$ and one for $X^{\gamma\downarrow}$. When p_T^{ISR} categories are applied, the presence of two numbers, an upper and lower bound, assumes two regions: $X^{P_{low}^{high}}$ and $X^{P_{high}}$. Lastly, X_b means the ISR b-jet categorization is applied. The R_{ISR}-$M_{\perp}$ binning used for any given region can be found in Table 5.9, following the S system, $NL-NJ$, categorization.

3L Regions	
eee	$\mu\mu\mu$
$3L_G^{ee\|eOSSF} - 0J_{0b} - X^{P_{250}}$	$3L_G^{\mu\mu\|\mu OSSF} - 0J_{0b} - X^{P_{250}}$
$3L_G^{ee\|enOSSF} - 0J_{0b} - X^{P_{250}}$	$3L_G^{\mu\mu\|\mu nOSSF} - 0J_{0b} - X^{P_{250}}$
$3L_G^{eeeOSSF} - g1J - X^{P_{250}}$	$3L_G^{\mu\mu\mu OSSF} - g1J - X^{P_{250}}$
$3L_G^{eeenOSSF} - g1J - X^{P_{250}}$	$3L_G^{\mu\mu\mu nOSSF} - g1J - X^{P_{250}}$
$ee\mu$	$\mu\mu e$
$3L_G^{ee\|\mu OSSF} - 0J_{0b} - X^{P_{250}}$	$3L_G^{\mu\mu\|eOSSF} - 0J_{0b} - X^{P_{250}}$
$3L_G^{ee\mu nOSSF} - 0J_{0b} - X^{P_{250}}$	$3L_G^{\mu\mu\|enOSSF} - 0J_{0b} - X^{P_{250}}$
$3L_G^{ee\mu OSSF} - g1J - X^{P_{250}}$	$3L_G^{\mu\mu eOSSF} - g1J - X^{P_{250}}$
$3L_G^{ee\mu nOSSF} - g1J - X^{P_{250}}$	$3L_G^{\mu\mu enOSSF} - g1J - X^{P_{250}}$

Table 5.13: List of the Gold 3 lepton regions ($3L$), counting the number of hadronic objects in the S system (i.e. $1J$), with the event system specific categorization (X...). The Silver and Bronze regions are similar, with the exception that the lepton and X categorizations are condensed so that leptons are only flavor separated, and any p_T^{ISR}, γ_T or SV η categories are made to be inclusive. Additionally, the b-jet regions are also made to be inclusive. For leptons, when there is a category like $L^{e\pm Xe\mp|\mu}$, it just means that there cannot be an OSSF pair in P_a, the muon is not limited to P_b in this case, as long as the first part is satisfied. The presence of a γ denotes that this categorization is applied and it counts as 2 regions, one for $X^{\gamma\uparrow}$ and one for $X^{\gamma\downarrow}$. The R_{ISR}-$M_\perp$ binning used for any given region can be found in Table 5.9, following the S system, $NL-NJ$, categorization.

Bin Name	Bin type	N bkg	Largest Bkg	N Largest Bkg	largest signal	N signal
$0L-0J_{0b}^{1SVc}-X_{0b}^{P_{300}}, R_{ISR}>0.98$	s	125.8	$W+jets$	95.3	T2bW-500-480	17.5
$0L-0J_{0b}^{1SVc}-X_{1b}^{P_{300}}, R_{ISR}>0.98$	s	19.5	$W+jets$	8.0	T2bW-500-481	2.6
$0L-0J_{0b}^{1SVf}-X_{0b}^{P_{300}}, R_{ISR}>0.98$	s	43.3	$W+jets$	31.7	T2bW-500-480	1.9
$0L-0J_{0b}^{1SVf}-X_{1b}^{P_{300}}, R_{ISR}>0.98$	s	3.7	$W+jets$	1.4	T2tt-550-400	0.3
$0L-0J_{0b}^{1SVc}-X_{1b}^{P_{300}}, R_{ISR}\leq 0.98$	s	663.0	$t\bar{t}+jets$	315.8	T2tt-550-425	13.6
$0L-0J_{0b}^{g2SV}-X^{P_{300}}, R_{ISR}>0.98$	s	10.3	$W+jets$	7.7	T2bW-500-480	1.5
$0L-0J_{0b}^{1SVc}-X_{0b}^{P_{300}}, R_{ISR}\leq 0.98$	c	2880.5	$W+jets$	2557.9	T2tt-550-450	24.4
$0L-0J_{0b}^{1SVf}-X_{0b}^{P_{300}}, R_{ISR}\leq 0.98$	c	1263.3	$W+jets$	1164.9	T2bW-500-480	2.8
$0L-0J_{0b}^{1SVf}-X_{1b}^{P_{300}}, R_{ISR}\leq 0.98$	c	213.8	$W+jets$	104.6	T2tt-550-325	2.8
$0L-0J_{0b}^{g2SV}-X^{P_{300}}, R_{ISR}\leq 0.98$	c	303.5	$W+jets$	237.6	T2bW-500-480	4.2

Table 5.14: Table detailing the number of events in the background and largest signal model. For the defined SV signal (s) and control (c) bins in 0 lepton final states, the numbers of total background events with the largest source of these background events is given. The number of signal events is shown for the signal mass model listed. The statistical errors on the background range from sub-percent for the largest count, and up to 7% for the smallest. The statistical errors on the signals range from 6% for the largest counts, and up to 58% for the smallest.

cover, it was simple to separate out bins that could be considered control bins for every stop signal considered. To be considered a SV control bin, it was required that in a given region, a lower R_{ISR} bin was required, and additionally the fraction of signal to background had to be less than 1.5%. These lower R_{ISR} bins were not separated by $M_{\perp}$. The signal bins were then those that were present when the R_{ISR} cut was reversed. Figure 5.7 gives two distributions of SV bins which give the total background and the stop signal that has the largest contribution for these control and signal bins. The ratio plot in these figures shows that, for the control bins, there is little signal contamination. The information in the figure is also given numerically in Tables 5.14 and 5.15, with the highest background and signal named. The control bins defined here were made to test the SF method for the SVs that will be described in Section 6.2. It will be noticed that some of these bins have a further categorization not present in the SV regions given by Tables 5.10 and 5.11. This extra category corresponds to the ISR b-jet counting categorization. This was added in order to increase the number of bins used for this test, along with adjusting the phase space so that some bins would have a more dominant background other than $W+jets$, for which the b-jet counting increases the fraction of $t\bar{t}+jets$.

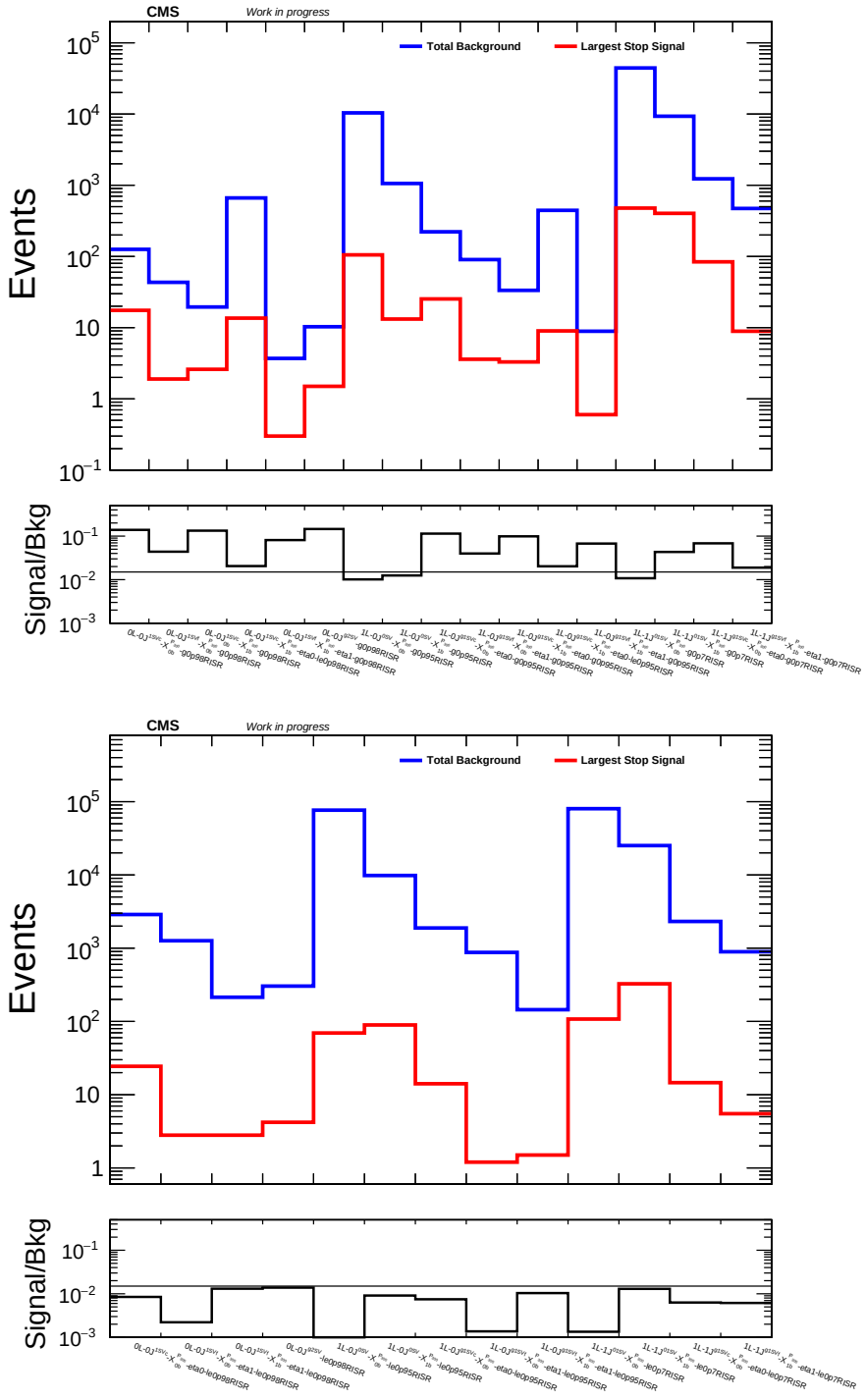


Figure 5.7: Distributions of the number of events in a set of regions that are related to the SV selection in the 0 and 1 lepton categories. These are separated into signal regions (top) and control regions (bottom). The histograms give the total MC background, and the largest of a mix of T2tt and T2bW signals with mass splittings ranging from 10 to 225 GeV for stop masses between 500 and 600 GeV. The ratio plots give the percentage of signal with respect to the background, with a line drawn at the 1.5% mark.

Bin Name	Bin type	N bkg	Largest Bkg	N Largest Bkg	largest signal	N signal
$1L_G - 0J_{0b}^{0SV} - X_{0\bar{b}}^{P_{300}}, R_{ISR} > 0.95$	s	10375.5	$W + jets$	9922.5	T2tt-500-480	105.2
$1L_G - 0J_{0b}^{0SV} - X_{1\bar{b}}^{P_{300}}, R_{ISR} > 0.95$	s	1058.0	$W + jets$	782.0	T2tt-500-480	13.2
$1L_G - 0J_{0b}^{g1SVc} - X_{0\bar{b}}^{P_{300}}, R_{ISR} > 0.95$	s	221.8	$W + jets$	204.1	T2bW-500-480	25.3
$1L_G - 0J_{0b}^{g1SVf} - X_{0\bar{b}}^{P_{300}}, R_{ISR} > 0.95$	s	90.3	$W + jets$	84.5	T2bW-500-480	3.6
$1L_G - 0J_{0b}^{g1SVc} - X_{1\bar{b}}^{P_{300}}, R_{ISR} > 0.95$	s	33.4	$W + jets$	17.9	T2bW-500-480	3.3
$1L_G - 0J_{0b}^{g1SVc} - X_{1\bar{b}}^{P_{300}}, R_{ISR} \leq 0.95$	s	444.0	$t\bar{t} + jets$	213.2	T2tt-518-350	9.0
$1L_G - 0J_{0b}^{g1SVf} - X_{1\bar{b}}^{P_{300}}, R_{ISR} > 0.95$	s	8.9	$W + jets$	5.6	T2tt-550-400	0.6
$1L_G - 1J^{0SV} - X_{0\bar{b}}^{P_{300}}, R_{ISR} > 0.7$	s	44312.4	$W + jets$	40055.1	T2tt-550-450	477.3
$1L_G - 1J^{0SV} - X_{1\bar{b}}^{P_{300}}, R_{ISR} > 0.7$	s	9306.4	$t\bar{t} + jets$	4544.7	T2tt-550-325	402.9
$1L_G - 1J^{g1SVc} - X^{P_{300}}, R_{ISR} > 0.7$	s	1230.6	$W + jets$	856.4	T2tt-550-450	84.0
$1L_G - 1J^{g1SVf} - X^{P_{300}}, R_{ISR} > 0.7$	s	471.7	$t\bar{t} + jets$	363.8	T2tt-550-450	8.9
$1L_G - 0J_{0b}^{0SV} - X_{0\bar{b}}^{P_{300}}, R_{ISR} \leq 0.95$	c	76567.2	$W + jets$	73312.2	T2tt-550-450	69.5
$1L_G - 0J_{0b}^{0SV} - X_{1\bar{b}}^{P_{300}}, R_{ISR} \leq 0.95$	c	9819.0	$W + jets$	6335.6	T2tt-550-325	89.5
$1L_G - 0J_{0b}^{g1SVc} - X_{0\bar{b}}^{P_{300}}, R_{ISR} \leq 0.95$	c	1885.6	$W + jets$	1687.4	T2tt-550-450	14.1
$1L_G - 0J_{0b}^{g1SVf} - X_{0\bar{b}}^{P_{300}}, R_{ISR} \leq 0.95$	c	875.4	$W + jets$	813.2	T2tt-500-420	1.2
$1L_G - 0J_{0b}^{g1SVf} - X_{1\bar{b}}^{P_{300}}, R_{ISR} \leq 0.95$	c	144.6	$W + jets$	69.1	T2tt-558-375	1.5
$1L_G - 1J^{0SV} - X_{0\bar{b}}^{P_{300}}, R_{ISR} \leq 0.7$	c	80194.4	$W + jets$	69347.3	T2tt-550-325	107.7
$1L_G - 1J^{0SV} - X_{1\bar{b}}^{P_{300}}, R_{ISR} \leq 0.7$	c	25146.5	$t\bar{t} + jets$	15215.6	T2tt-550-325	326.3
$1L_G - 1J^{g1SVc} - X^{P_{300}}, R_{ISR} \leq 0.7$	c	2316.7	$W + jets$	1417.1	T2tt-550-325	14.6
$1L_G - 1J^{g1SVf} - X^{P_{300}}, R_{ISR} \leq 0.7$	c	893.5	$W + jets$	598.5	T2tt-550-325	5.5

Table 5.15: Table detailing the number of events in the background and largest signal model. For the defined SV signal (s) and control (c) bins in 1 lepton final states, the numbers of total background events with the largest source of these background events is given. The number of signal events is shown for the signal mass model listed. The statistical errors on the background range from sub-percent for the largest count, and up to 7% for the smallest. The statistical errors on the signals range from 5% for the largest counts, and up to 58% for the smallest.

Chapter 6

Results

After all of the regions were defined, the sensitivity studies were prepared. Expected sensitivities were determined for grids of mass points corresponding to the T2tt, T2bW, and T2cc signals, for stop masses ranging from 400 up to 1500 GeV for mass splittings ranging from 10 to 200 GeV. Before presenting these results, the fitting method is described, along with the systematics that were added as nuisance parameters to the fit. After that, the sensitivity results and interpretations are given.

6.1 Fitting method

The fitting process made use of the Higgs Combine Tool (combine), which provides a command line interface to various statistical techniques that are present inside the RooFit/RooStats packages [69]. For this analysis, a maximum likelihood fit was done following a modified frequentist approach and the asymptotic method [71] was used for calculating the limits.

As input for the fit, datacards are prepared for each signal mass point, for which the datacards, essentially, are the instructions for how the fit should proceed. It details the observables, which correspond to the number of bins in the analysis, each containing the number of weighted events found in the backgrounds, and signals. It then sets the systematic uncertainties and describes how they scale for each observable, and how they should interact between all of the bins.

To prepare the datacard, the 1D histograms of all of the regions, separated into processes and their systematic variations are supplied to a program called BuildFit, which creates the datacard for a set of regions and processes based on a set of command line parameters. This datacard is

then converted into a RooFit workspace, which is just the datacard and histograms converted into RooFit objects and function, and which is what is used as input into Higgs combine.

Simultaneous fits of all of the regions were then performed, floating all of the backgrounds except for single top, $t\bar{t}+V$ and tri-boson processes, which were instead given a 20% systematic, for all of the relevant stop signal mass grids. By floating the main backgrounds in all the bins, the bins that were treated as controls for a particular mass point and region were able to constrain the background in the regions where signal was expected. These control and signal bins would change depending on the signal model and mass point.

6.2 Systematics

There were a number of systematic uncertainties that needed to be accounted for: those related to object selection, corrections, and uncertainty in measurements. This section details these uncertainties, and how they are accounted for in the fit.

6.2.1 Lepton corrections and fake estimation

For the leptons, scale factors were derived based on the data to MC efficiencies measured when making selections on the lepton ID, promptness and isolation, as it is used in the analysis. For both electrons and muons, the Tag-and Probe method [92] (described in more detail below) was used to calculate the data to MC efficiencies. This was done separately for each type of lepton selection. So for the lepton IDs, two ID scale factors were produced, one for bronze, the other for silver and gold. Then one isolation scale factor, and one prompt scale factor were calulated for the remaining selections. The specific combination of scale factors that needed to be applied for a lepton depended on whether it was categorized as gold, silver, or bronze.

The Tag and Probe method is a tool that is used to measure an object efficiency by making use of known resonances that decay into two objects, for which one leg of the resonance is made to pass a tight identification (tag), while the other passes a looser identification (probe), in this

case the IDs, promptness, and isolation used by the analysis would be the probe. Line shapes are then fitted for the cases where the probe passes, and where the probe fails. The efficiency is then calculated by fitting the line shapes of the resonance in the pass and fail groups, and looking at the ratio of the passing probes with the sum of the passing and failing probes. This efficiency is measured in bins of p_T and η and is what was used as the scale factor, where the uncertainty in the scale factor was used as a systematic error.

For the electrons, the di-lepton resonance used is the Z resonance from the Drell-Yan process. The scale factors were binned in 6 p_T and 3 $|\eta|$ categories:

- $5 \leq p_T < 10$ GeV
- $10 \leq p_T < 20$ GeV
- $20 \leq p_T < 30$ GeV
- $30 \leq p_T < 40$ GeV
- $40 \leq p_T < 60$ GeV
- $60 \leq p_T < 100$ GeV
- $0 \leq |\eta| < 0.8$
- $0.8 \leq |\eta| < 1.479$
- $|\eta| \leq 1.479$

An example of the electron 2017 scale factors is given in Figure 6.1.

For the muons, two resonance peaks were used to calculate the scale factors, Z and J/ψ. The Z peak was used for the ID, promptness, and isolation efficiencies for $20 \leq p_T < 100$ GeV, and then to extrapolate efficiencies down to $3 \leq p_T < 20$ GeV for low p_T promptness and isolation scale factors. The J/ψ resonance was then used to calculate ID scale factors for $3 \leq p_T < 20$ GeV. These separate scale factors were then combined, depending on the p_T to get the overall scale factors for

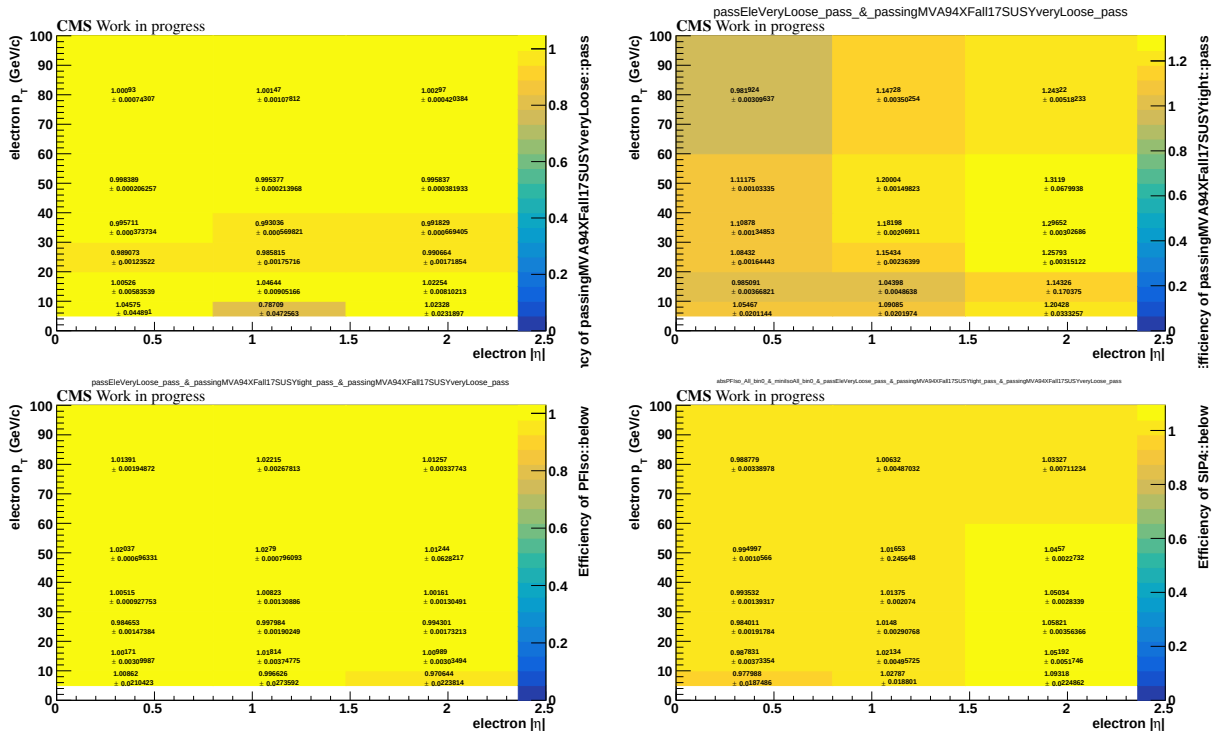


Figure 6.1: Scale factors obtained from the electron tag-and-probe binned in $|\eta|$ and p_T for 2017 data. The four efficiencies shown are: very loose ID (top left), tight ID (top right), Isolation (bottom left) and Promptness (bottom right).

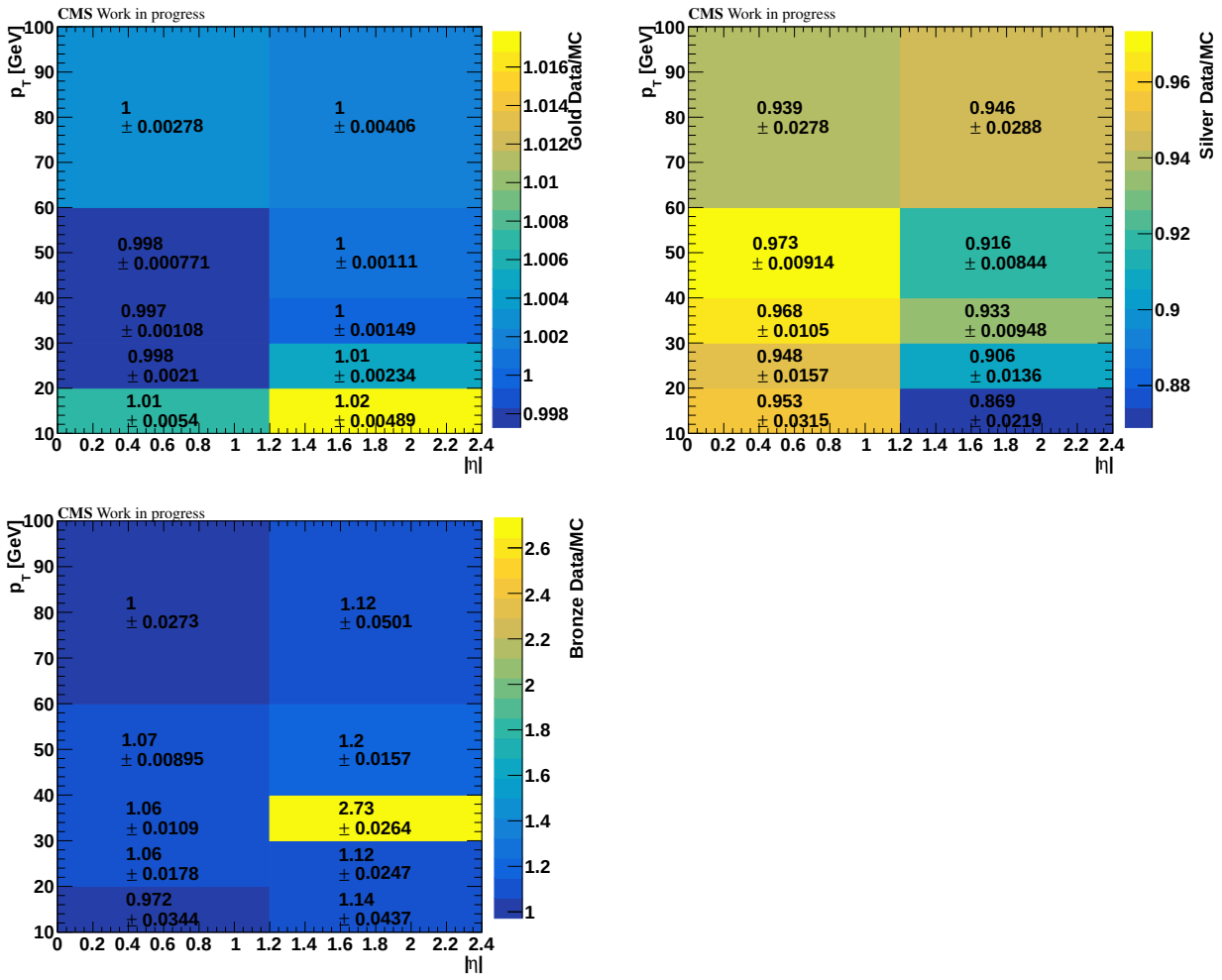


Figure 6.2: The 2017 data/MC scale factors for muons in all three lepton tiers: Gold (top left), Silver (top right), and Bronze (bottom).

the gold, silver, and bronze tiers. Figure 6.2 then gives the gold, silver, and bronze muon 2017 scale factors, binned in p_T and $|\eta|$.

6.2.1.1 Lepton fake estimation

The goal of the fake estimation was to account for the rate and systematic shape differences between data and MC for sources of fake leptons. For this method, the fake sources are separated out into heavy and light flavors, through the use of the silver and bronze categorizations, where silver has more heavy flavor, and bronze has more light. This can be seen in Figure 6.3, where the fractional flavor contributions are given for bronze and silver muons.

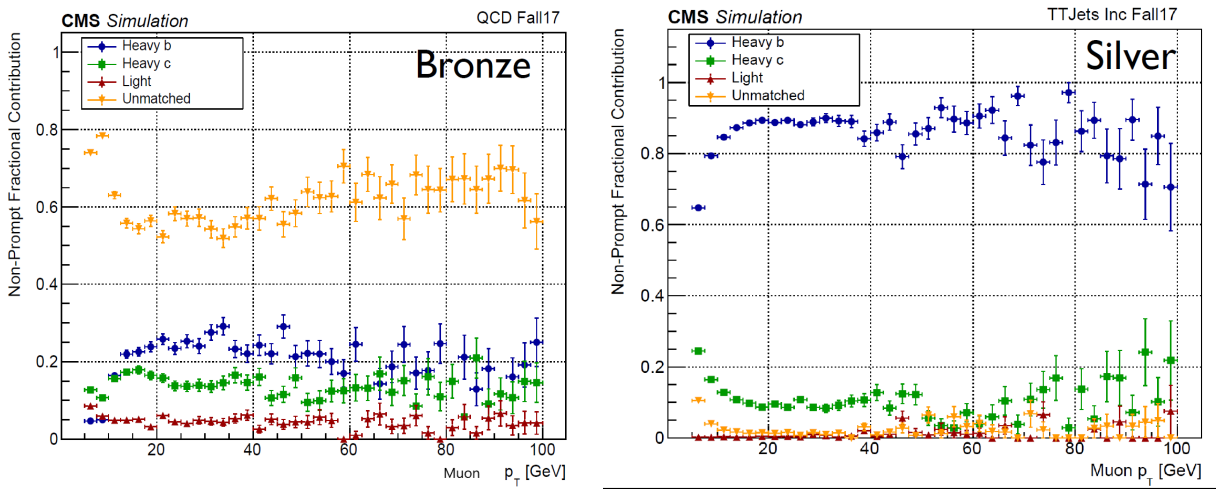


Figure 6.3: The muon fractional flavor contributions for the bronze (left) and silver (right) categorizations.

To account for systematic differences, the lepton fakes were modeled in each bin following Equation 6.1, which then becomes a nuisance parameter in the fit. In this equation, each of the terms for sources (s), categories (c), and process (p) are:

- SF: scale factors for the heavy, light and unmatched flavor sources, connected between processes and categories, and constrained by the silver and bronze categories
- w_{cps}: fractional weights related to the normalization of the processes, categories, and sources
- n_{cps}^{MC}: Central values, connected to the floating normalization of each process, and based on MC predictions
- $\hat{f}(\vec{\alpha}_B, \vec{\alpha}_S, \vec{\alpha}_G)$: Shape morphing based on lepton p_T, promptness and isolation variations.

$$n^{pred} = SF_s \times w_{cps} \times n_{cps}^{MC} \times \hat{f}(\vec{\alpha}_B, \vec{\alpha}_S, \vec{\alpha}_G) \tag{6.1}$$

Figure 6.4 gives an example of the shape correlations between different b-jet categorizations for heavy flavor fake contributions.

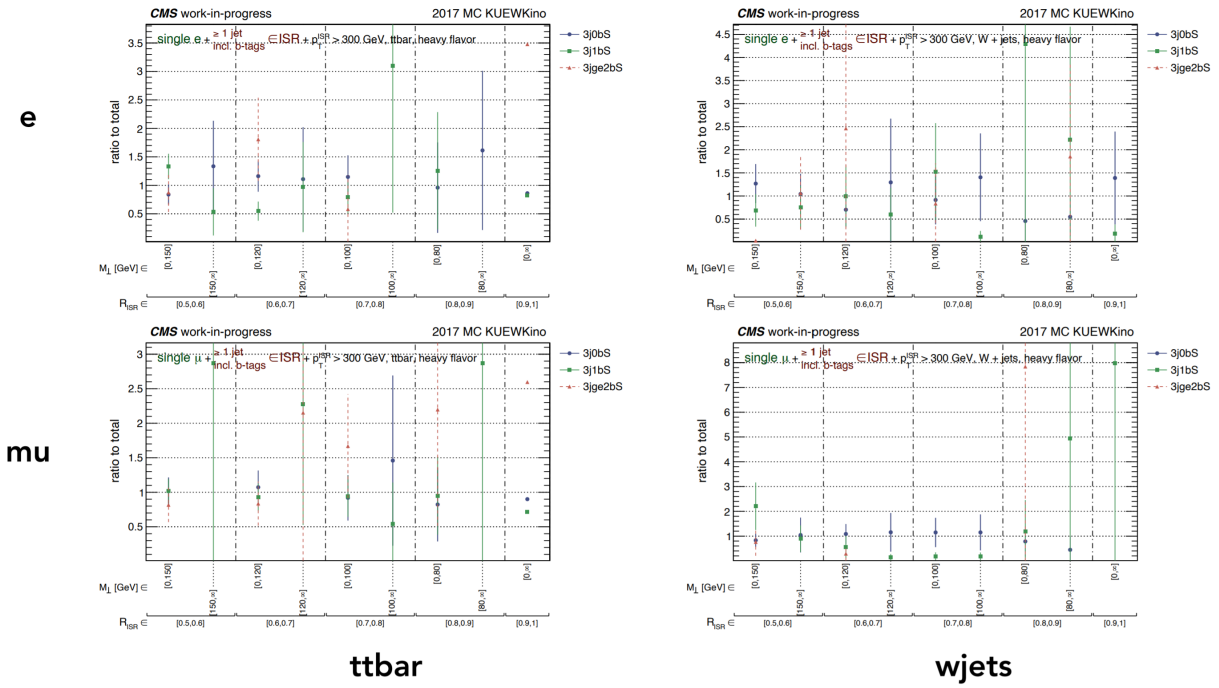


Figure 6.4: Distributions of the shape correlations for $1L-3J$ categories with varying numbers of b-jets. These are looking at the heavy flavor fake contributions from $t\bar{t}+jets$ (left) and $W+jets$ (right) for electrons (top) and muons (bottom)

6.2.2 JetMET corrections

Jet energy scale (JES) and energy resolution (JER) corrections were applied to jets, and propagated to $\not{p}_T$ for the MC samples. For the JES, a scale factor is applied to the jet energy, and propagated to the related objects. For the JER, the jet p_T is smeared with a JER factor that matches the jet resolution in data. The uncertainties on these corrections are treated as systematics in the fit. For the JES uncertainty, the jet energy is scaled up and down by the uncertainty in the JES scale factor. For the JER uncertainty, additional smearing is applied to the jet p_T, corresponding to an up and down variation.

Corrections are also applied based on the $\not{p}_T$ trigger scale factors, with an associated systematic uncertainty. The scale factor was calculated from the ratio of data to MC in the trigger turn on from Figure 5.3. The systematic uncertainty from this scale factor was then determined by Equation 6.2, where a 2nd order polynomial is multiplied by the ratio of a Gaussian CDF fit in data and MC,

where the a, b, and c parameters float in the fit, and are sensitive to the maximum/minimum scale factor, to where the efficiency plateaus, and to the scale factor in the plateau region, respectively.

$$\sigma_{SF} = \left(a\,(x-b)^2 + c\right)\left(\frac{\text{Data Fit}}{\text{MC Fit}}\right) \tag{6.2}$$

6.2.3 B-tagging corrections

For the b-jet tagging, corrections are made to MC based on the true hadron flavor of a b-tagged jet, in conjunction with centrally measured data to MC scale factors, and analysis specific b-tagging efficiencies for the simulated processes.

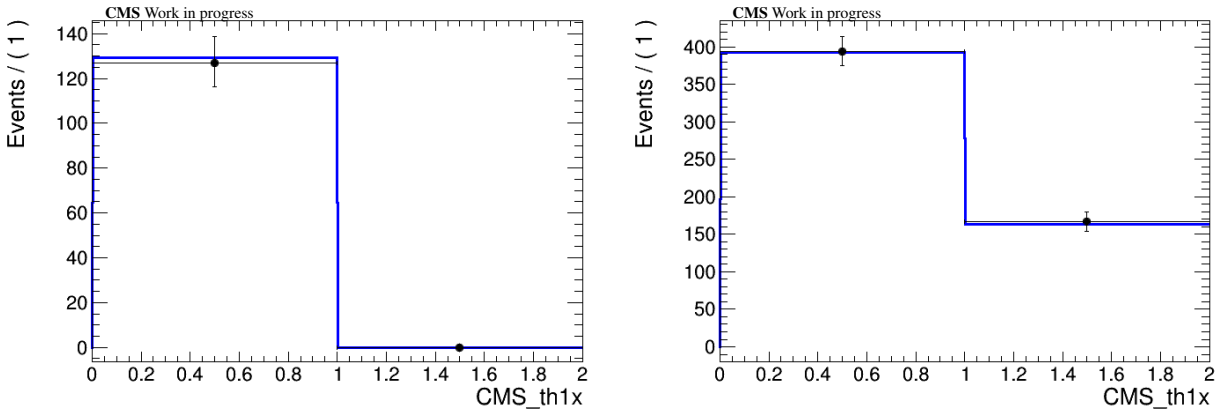
An event weight was applied, and calculated following Equation 6.3, where SF_i is the data to MC scale factor as a function of jet flavor, and ε_i is the MC efficiency as a function of jet flavor.

$$w = \frac{\prod\limits_{i=tagged} \varepsilon_i \prod\limits_{j=nottagged} \left(1-\varepsilon_j\right)}{\prod\limits_{i=tagged} SF_i\varepsilon_i \prod\limits_{j=nottagged} \left(1-SF_j\varepsilon_j\right)} \tag{6.3}$$

The systematic uncertainties in the event weight were separated into heavy and light flavor nuisance parameters, where the up and down variations are determined by centrally varied data to MC scale factors. The size of the corrections can range from $2-8\%$, depending on the p_T and flavor of the jet.

6.2.3.1 SV soft b-tagging

For the tagged SVs, scale factors are to be measured directly in the fit using a data driven method, following the method described in Section 4.6.1. These data driven scale factors are to be checked in the fit through the use of the SV control regions previously shown in Figure 5.7 and Tables 5.14 and 5.15, so that the data could remain blinded in the signal regions. For each of the control bins, there were two associated regions, one with an SV selection and one without. The sum of these two bins then gave the inclusive bin. Then, with an SV category bin and an inclusive category bin, Equation 4.1 would then be used in the fit to get a correction for that bin. After performing the



(a) $1L_G^{\mu} - 0J^{g1SVc} - X_{0b}^{P_{300}}$ (b) $1L_G^{\mu} - 1J^{g1SVc} - X^{P_{300}}$

Figure 6.5: Post-fit distribution of the background only fit of the $1L_G^{\mu} - 0J^{g1SVc} - X_{0b}^{P_{300}}$ and $1L_G^{\mu} - 1J^{g1SVc} - X^{P_{300}}$ SV control regions, compared to real data, for 2017 samples.

fit with this correction factor, the background only fit would be compared to the data such as in Figure 6.5. In the full fit this scale factor would be calculated for every bin with an SV category, with the value being constrained by control bins.

6.2.4 Other systematics

Other systematics and corrections that were necessary to account for in the analysis are:

- Luminosity uncertainty
- Applying pileup re-weighting and systematic uncertainty
- PDF and Q^2, factorization, renormalization corrections and uncertainty
- Corrections for EE Noise in data
- Corrections for ECAL prefire in data
- Corrections for HEM-related effects in data, related to HEM-issue described in the last paragraph of Section 2.2.3
- ISR systematics in signal samples

- Kinematic corrections including $m_{\ell\ell}$ spin corrections and W/Z boson off-shell branching ratio corrections

- FULL to FAST Sim corrections to the stop signals, which account for differences between simulations. These corrections are calculated for each object used in the analysis: leptons, jets, b-jets, and SVs.

6.2.5 Fit closure tests

After the systematics were defined, tests on the fits were performed to ensure closure, and to see which systematics were dominating. The closure tests involved separately varying all of the nuisance parameters in the fit to get the pull and impact of each of them. These tests were done for a couple of toy Asimov datasets [71], one assuming a background only (b-only) fit and the other with signal injected (s+b). Figure 6.6 gives the results of some closure tests for a selection of nuisance parameters using the SV control regions. The ideal expectation for these plots, is that the pull is centered on 0, with $\pm 1\sigma$ uncertainty bands, such as what is seen for the `lumi_2017` nuisance parameter. Smaller or larger uncertainty bands would indicate the parameters being more or less constrained, respectively. The `BTAGSF` nuisance would be an example of a parameter being more constrained by the fit, due to the number of regions and bins that require b-tagged jets. For the impacts, the ideal is that the nuisance parameters do not show lopsided impacts, and if they do it should be well understood. From impact plots shown, there is nothing overly concerning with any of the nuisance parameters shown.

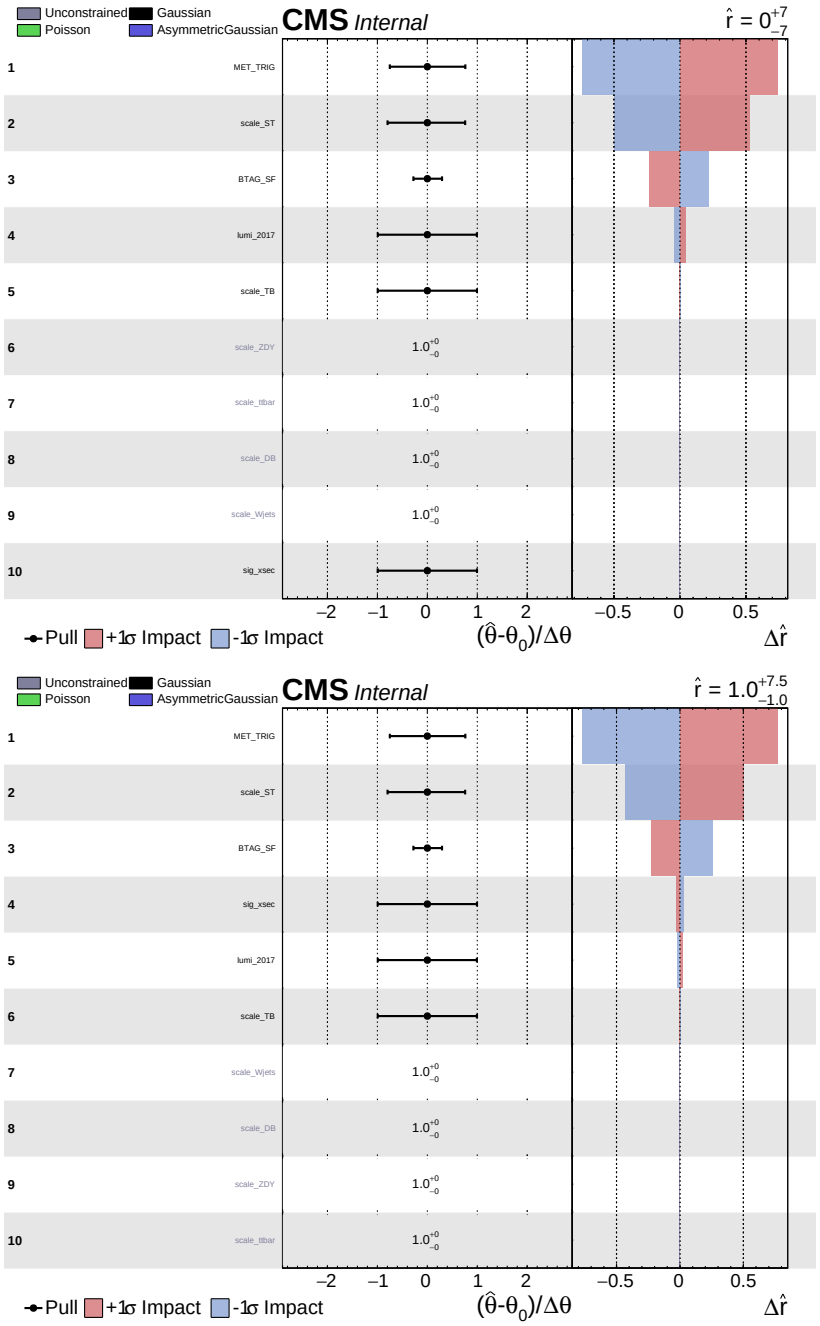


Figure 6.6: Impact and pulls for the nuisance parameters of the SV control region fit, using a b-only (top) and s+b (bottom) fit. The scale parameters that show a 1.0 are the normalizations for the dominant processes that are allowed to float freely. Due to this, the fitting was unable to gauge the pull for those parameters, as the normalizations are not allowed to go into negative values.

6.3 Results and Interpretations

Fits were performed for T2tt, T2bW, T2-4bd, and T2cc signal models for stop masses ranging from 400 to 1500 GeV and mass splittings between 10 and 200 GeV. Expected yields for some of the regions are illustrated in Figure 6.7 for the background, divided into their $M_{\perp}$ and R_{ISR} bins. These expected yields are consistent with a background only hypothesis, and no issues were detected. Additionally, the event yields from the simulated MC for most of the regions are shown in Figures 6.8, 6.9, 6.10, and 6.11 for the background and some select T2tt and T2bW signals. These figures give a compact view of the event yields for all of the regions, where the first set of bins correspond to the events in the inclusive final states, i.e. number of leptons and jets. The following bins then start subdividing these final states into some of their lepton and SV categorizations, and then the last bins subdividing these into their event categorizations.

The expected results are interpreted for stop quark pair production in the previously mentioned models, where the LSP is assumed to be the lightest neutralino. Expected limits are set using a modified frequentist approach using CL_S criterion, and an asymptotic method [71].

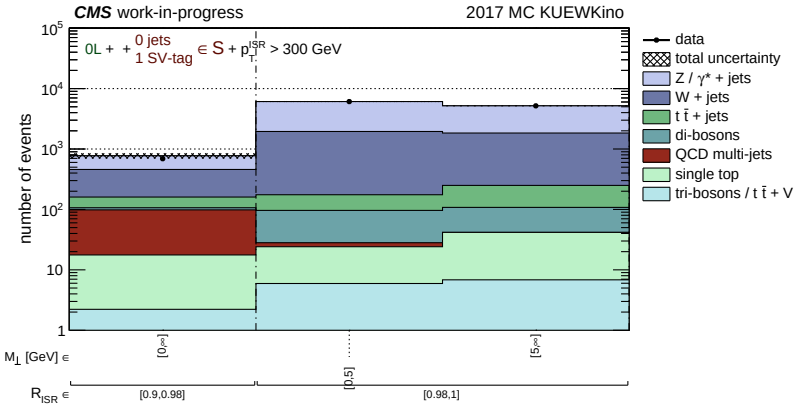
The 95% confidence level (CL) expected upper limits on the cross section for the T2tt process are given in Figure 6.12. Equivalent limits are given for the T2bW, and T2cc processes in Figures 6.13, and 6.14, respectively.

Looking at the smallest mass splittings ($\Delta m_{\tilde{t}\tilde{\chi}_1^0}$) of 10 and 20 GeV, the expected exclusions are as follows:

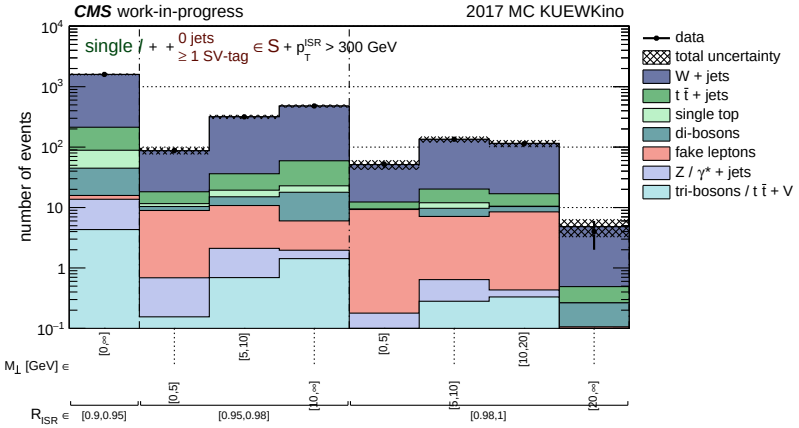
- $\Delta m_{\tilde{t}\tilde{\chi}_1^0} = 10$ GeV: For the T2tt process, exclusions are expected for stop masses up to 625 GeV. For the T2bW process, exclusions are expected for stop masses up to 675 GeV. Lastly, for the T2cc process, exclusions are expected for stop masses up to 675 GeV.

- $\Delta m_{\tilde{t}\tilde{\chi}_1^0} = 20$ GeV: For the T2tt process, exclusions are expected for stop masses up to 700 GeV. For the T2bW process, exclusions are expected for stop masses up to 725 GeV. Lastly, for the T2cc process, exclusions are expected for stop masses up to 725 GeV.

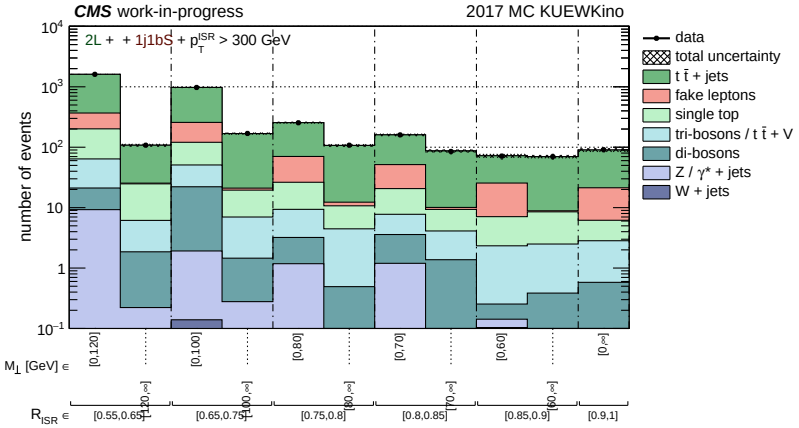
Additionally, for the two mass points available for the T2-4bd model, exclusions are expected



(a) $0L_G - 0J_{0b}^{1SV} - X^{P_{300}}$

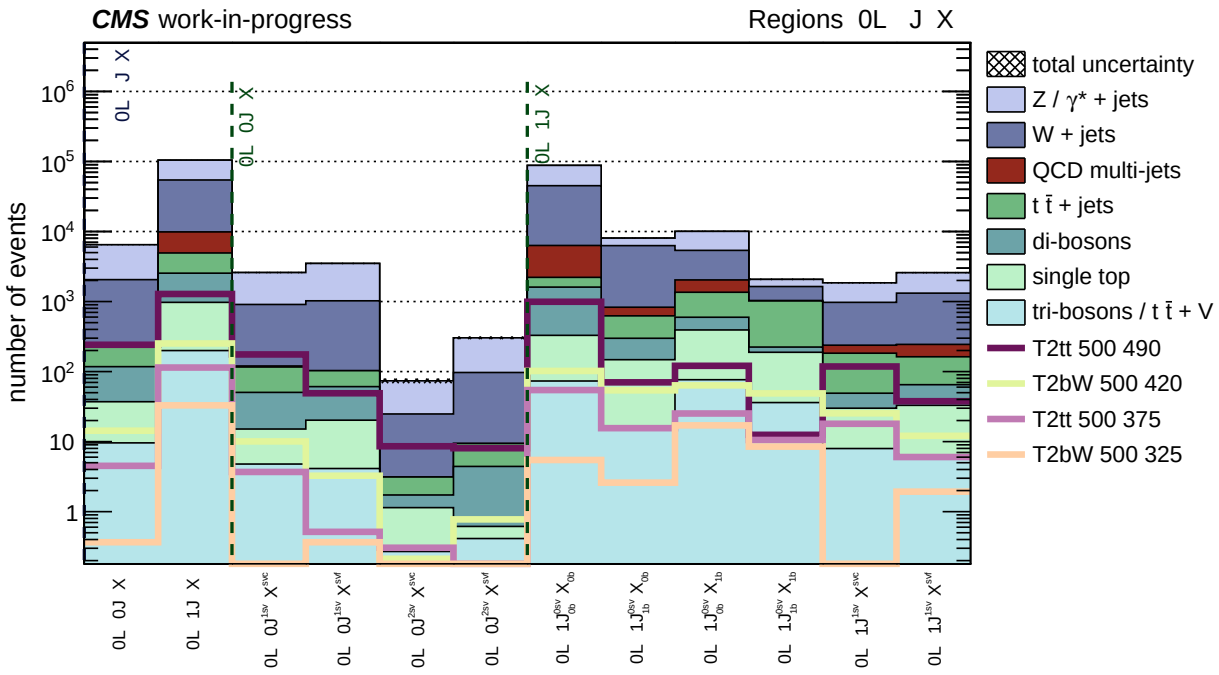


(b) $1L_G - 0J^{g1SV} - X^{P_{300}}$



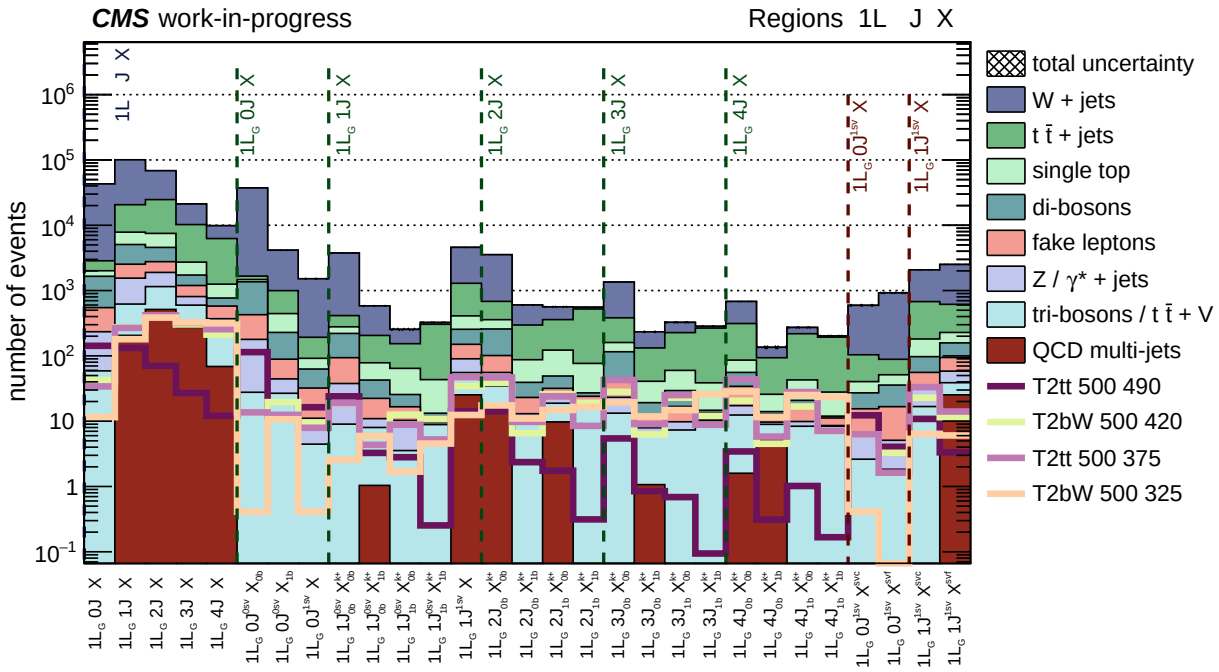
(c) $2L_G - 1J_{1b} - X^{P_{300}}$

Figure 6.7: Post-fit background yields for some 0L, 1L and 2L final states. The 0L and 1L plots correspond to an SV selection, $1L - 0J_{0b}^{1SV}$ (top) and $1L - 0J_{b}^{g1SV}$ (middle), while the 2L corresponds to a 1 jet, 1 b-jet selection, $2L - 1J_{1b}^{SV}$ (bottom). The data correspond to an Asimov dataset, since the real data are currently blinded.



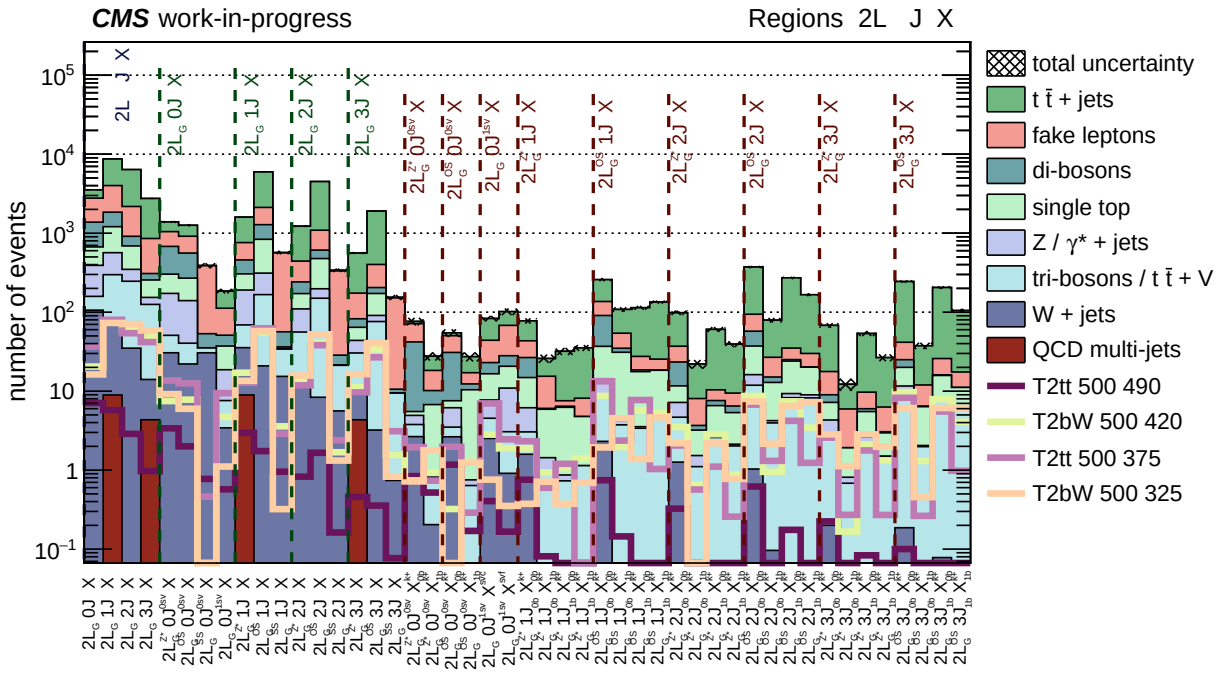
(a) 0*L* regions

Figure 6.8: Simulated event yields for the background and some select T2tt and T2bW signals in the 0 lepton final states. The bins detail how the yields change with further categorization, starting with the overall final state on the left, and adding on the lepton and SV categorizations, and the event categorizations as one moves to the bins on the right.



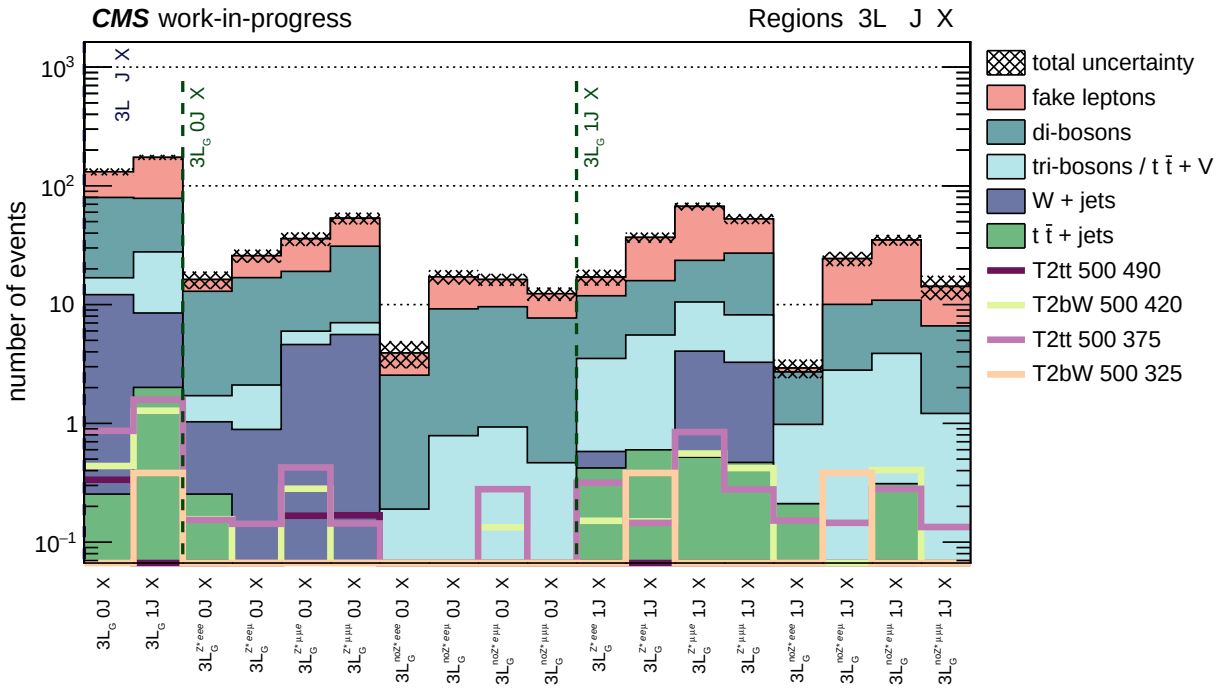
(a) 1L regions

Figure 6.9: Simulated event yields for the background and some select T2tt and T2bW signals in the 1 lepton final states. The bins detail how the yields change with further categorization, starting with the overall final state on the left, and adding on the lepton and SV categorizations, and the event categorizations as one moves to the bins on the right. Some of the regions have a slightly different nomenclature, as the regions are a condensed form of the final regions. For example $k+$ refers to a combination of the highest p_T^{ISR} and γ_T categorizations.



(a) 2L regions

Figure 6.10: Simulated event yields for the background and some select T2tt and T2bW signals in the 2 lepton final states. The bins detail how the yields change with further categorization, starting with the overall final state on the left, and adding on the lepton and SV categorizations, and the event categorizations as one moves to the bins on the right. Some of the regions have a slightly different nomenclature, as the regions are a condensed form of the final regions. For example the presence of a Z^* denotes a Z candidate region, where there are 2 leptons in the same hemisphere, and $k+$ refers to a combination of the highest p_T^{ISR} and γ_T categorizations.



(a) 3L regions

Figure 6.11: Simulated event yields for the background and some select T2tt and T2bW signals in the 3 lepton final states. The bins detail how the yields change with further categorization, starting with the overall final state on the left, and adding on the lepton and SV categorizations, and the event categorizations as one moves to the bins on the right. Some of the regions have a slightly different nomenclature, as the regions are a condensed form of the final regions. For example the presence of a Z^* denotes a Z candidate region, where there are 2 leptons in the same hemisphere, and $k+$ refers to a combination of the highest p_T^{ISR} and γ_T categorizations.

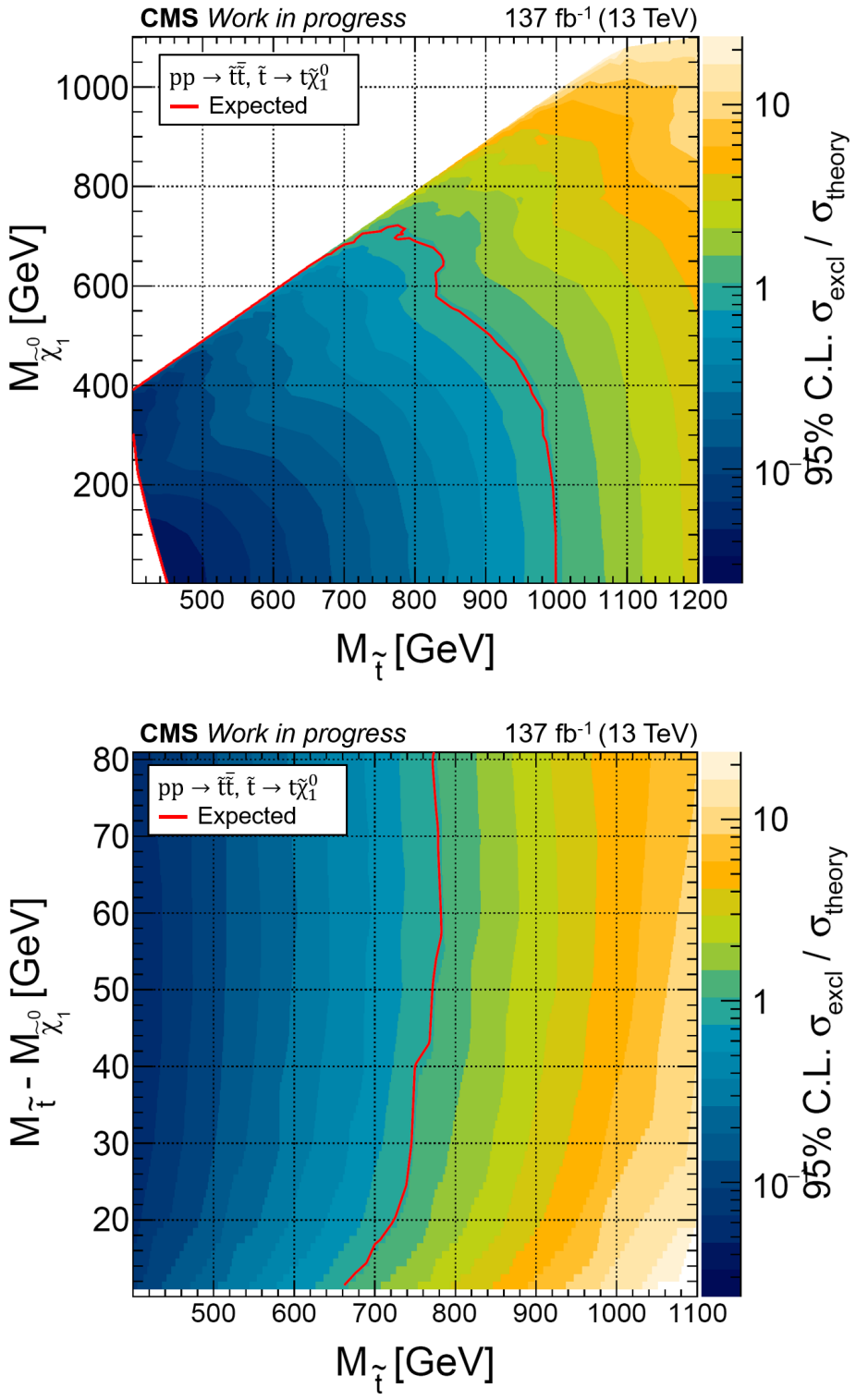


Figure 6.12: T2tt expected limits. The top plot shows expected limits for the stop mass against the LSP. The bottom plot shows expected limits for the stop mass against the mass difference between the stop and LSP, and is focused on the smallest mass splittings, below the W-corridor.

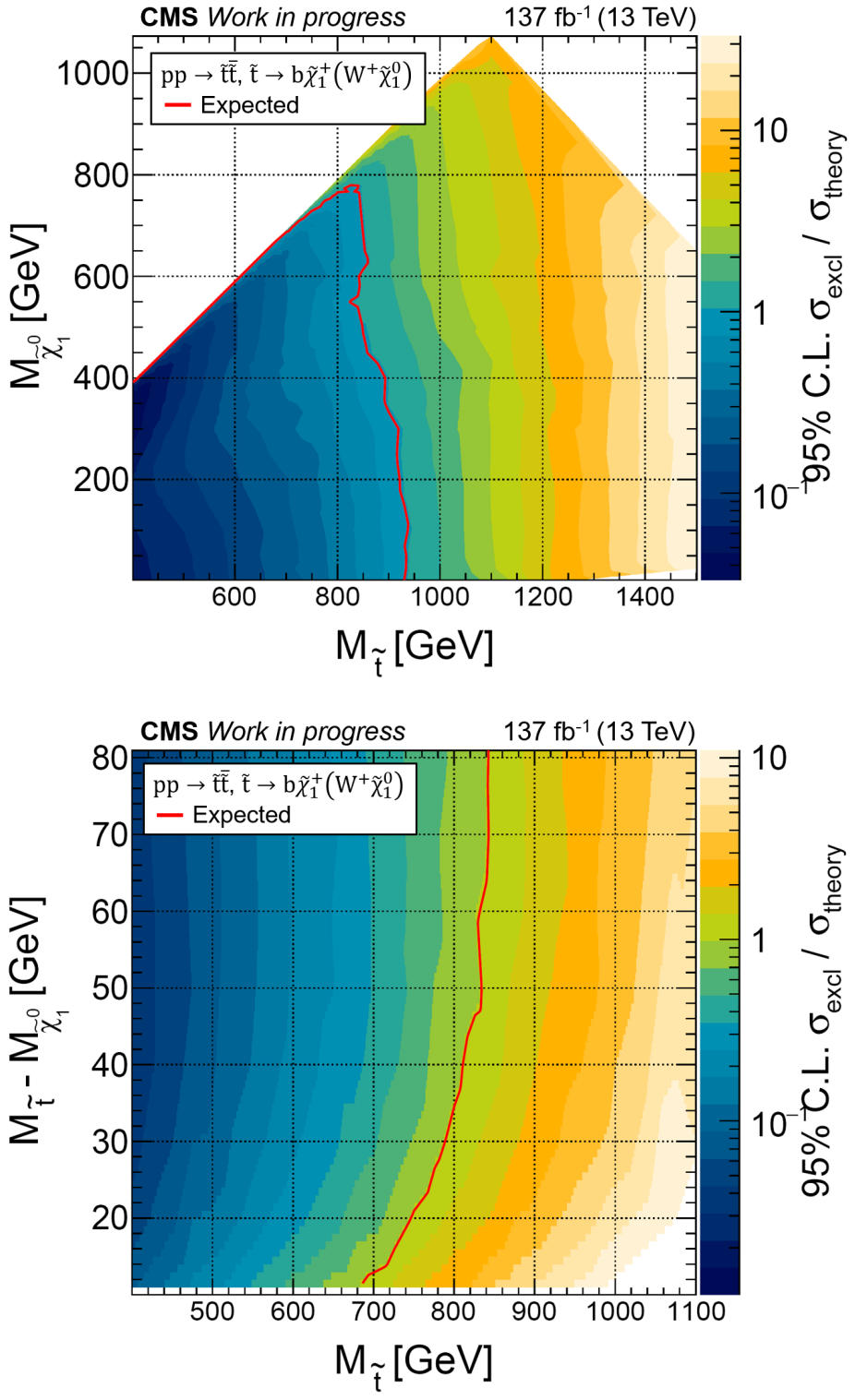


Figure 6.13: T2bW expected limits. The top plot shows expected limits for the stop mass against the LSP. The bottom plot shows expected limits for the stop mass against the mass difference between the stop and LSP, and is focused on the smallest mass splittings, below the W-corridor.

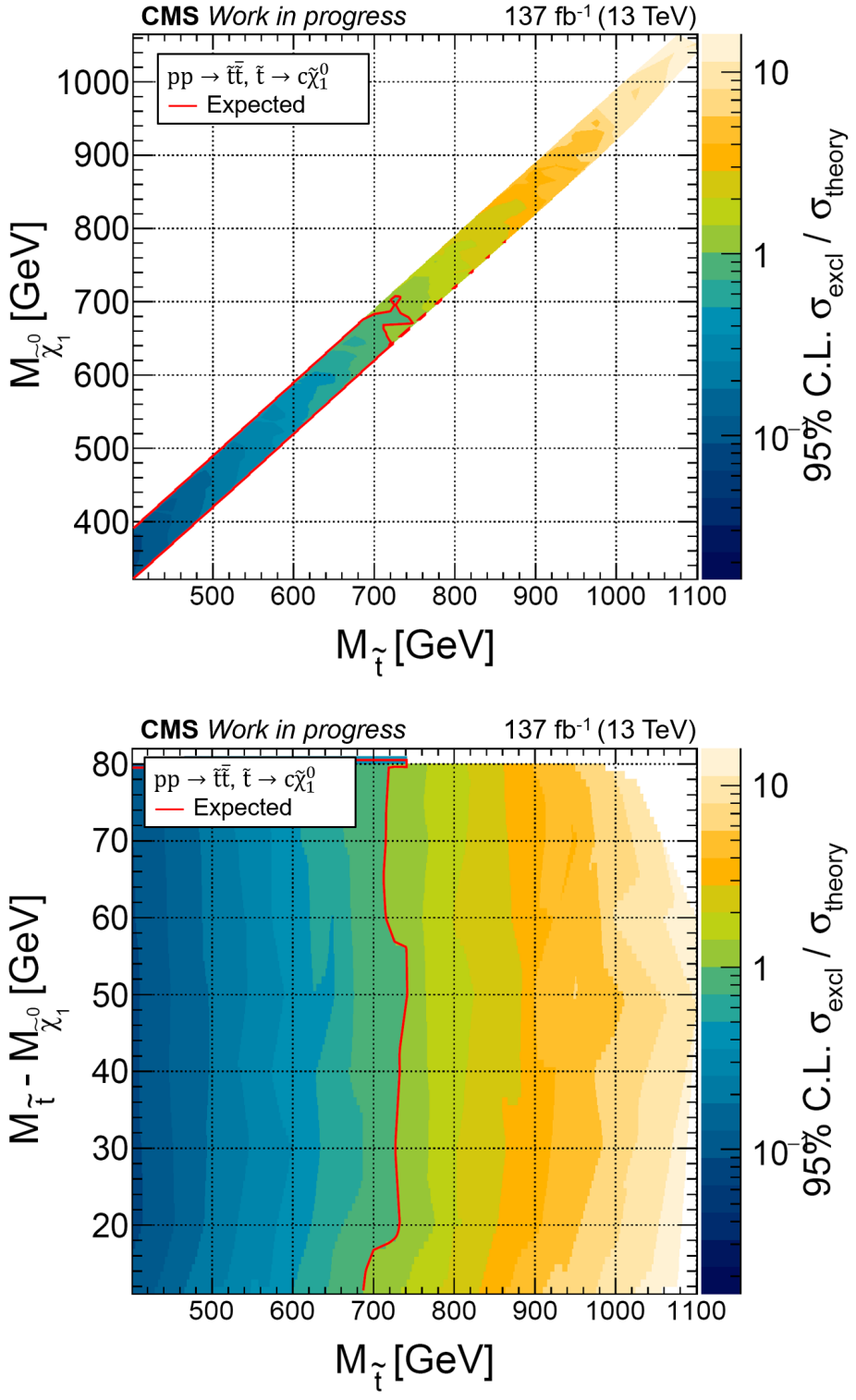


Figure 6.14: T2cc expected limits. The top plot shows expected limits for the stop mass against the LSP. The bottom plot shows expected limits for the stop mass against the mass difference between the stop and LSP, and is focused on the smallest mass splittings, below the W-corridor.

for a stop mass of 500 GeV with LSP masses of 420 and 490, with 95% CL upper limits of $0.18^{+0.07}_{-0.05}$ and $0.40^{+0.16}_{-0.11}$ pb, respectively.

Chapter 7

Summary

In this dissertation, a search was performed for supersymmetric top quarks in multiple different final states, ranging from fully hadronic, to 3 or more leptons, with varying multiplicities of jets, b-tagged jets, and soft b-tagged secondary vertices. The search uses a sample of proton-proton collision data at $\sqrt{s} = 13$ TeV, corresponding to a luminosity of 137 fb^{-1}. The data were collected by the CMS experiment at the LHC from 2016-2018. The focus of the search was compressed mass splittings between the stop and lightest supersymmetric particle, the neutralino. These splittings could range from 10 GeV up to slightly above the top corridor at 200 GeV. The compressed scenarios are accessed in the high energy events produced by 13 TeV collisions through the use of ISR recoil, wherein one or more jets are used to provide a kinematic 'kick' to the missing transverse momentum, to support the high mass of the invisible LSP, leaving the remaining energy in the event to the other visible reconstructed objects, ensuring that they have relatively low energy. The events are analyzed following a generic decay tree through the use of Recursive Jigsaw Reconstruction, which assigns reconstructed objects to different branches of said decay tree, corresponding to an ISR system and sparticle system, further divided into a visible and invisible system for the sparticle side. The object assignment is then used to create mass sensitive observables related to the approximated rest frames of the different decay tree systems.

As this search made use of b-tagging, where low energies are expected, a main component of the dissertation was the creation of a Deep Neural Network based soft b quark tagger, as the available jet b-tagging was not suited to the lower energies expected for the more compressed mass scenarios. This soft b-tagger made use of secondary vertices, and as their name implies, are an important part of any b-tagging due to there being a vertex displaced from the interaction

point, denoting particle lifetime. While a version of soft b-tagging was used in previous SUSY analyses by CMS, it only applied a set of flat cuts to select secondary vertex variables, and it was believed that this could be improved upon by using a DNN to increase the performance at very low transverse momentum on the order of a couple GeV. The most important part of improving performance at low p_T was including 2 track vertices, which are a large population of the SVs below 4 GeV. A driving factor in using a DNN to tag secondary vertices is the existence of a machine learning framework called the DeepJet framework, which is used by the CMS experiment to create the current jet b-tagging discriminators used by many CMS analyses. With the way the framework was built, it was able to be adapted to run on secondary vertices only, instead of jets, while keeping all of the necessary machinery intact such that it could make use of the same DNN algorithm used by one of the jet b-taggers, called the DeepCSV tagger. Multiple different trainings were performed over a set of simulated secondary vertices, focusing on the variables used, rather than the DNN algorithm. Nine different trainings were done, with inputs ranging from the same four variables used by the previous soft b-tagger, to the nine variables generally associated to secondary vertices. The final version of the discriminator made use of 8 secondary vertex variables, and was able to improve the b-tagging efficiency by up to 50% at low p_T over that of the previous soft b-tagger. Over the p_T range of the secondary vertices, from 2 to 20 GeV, the light flavor misidentification rate was maintained below 40%. This was compared to the previous soft b-tagging, which had a light misidentification rate above 40% for the p_T range from 8 to 20 GeV, while having a comparable b-efficiency to the DNN based tagger. At a p_T of about 5 GeV, the misidentification rate between the two taggers intersect, where the older b-tagger then has a lower misidentification rate, but the trade off is that its b-efficiency takes a proportional dive, falling by about 50% moving from a p_T of 5 down to 2 GeV. In this same p_T range, the DNN based tagger's b-efficiency only falls by 15%.

The soft b-tagging was included in the analysis through object counting for the different regions. For certain low object multiplicity regions, in the 0 or 1 lepton, and 0 or 1 jet final states, there was a requirement of either 1, ≥ 1, or ≥ 2 secondary vertices which pass the soft b-tagging

selection. Due to the low p_T of the selected secondary vertices, these regions were geared towards very compressed signals, with mass splittings below the W corridor. The low multiplicity secondary vertex regions were used in conjunction with many other regions, with both high and low multiplicities of reconstructed objects to set limits on the pair production of supersymmetric top quarks. Expected limits were given for three different stop production modes in the simplified framework, labeled T2bW, T2tt, and T2cc. Stop masses ranged from 400 to 1500 with mass splittings ranging from 10 to 200 GeV. Expected exclusions were found for $\Delta m_{\tilde{t}\tilde{\chi}_1^0} = 10$ GeV, for stop masses of up to 675 GeV in the T2bW model, 625 in the T2tt model, and 675 in the T2cc model.

Apart from the analysis, an independent project was also done related to finding a way to measure the CMS beam spot without making use of tracking reconstruction. Multiple fitting methods were tried to find the beam spot. Ultimately a method was made using a maximum likelihood fit which has an inverse power law function convolved with a Gaussian term. This method made use of a simple set of variables, corresponding to the cluster occupancy of the ROCs of the pixel detector, focusing on layer 1, and the global position of those ROCs. The method, when tested on simulated MC samples with a design detector, gave results that were accurate to within 1 mm of the true beam spot. Future work on this project would involve improving the fit so that it can run on real data, and making use of more than just the ROCs of the inner ladders of layer 1 as inputs to the fit.

For continuing work related to the soft b-tagging, there are a couple of proposals. The first is making the DNN based tagger centrally available through the B-tag and Vertexing (BTV) object group in CMS. This would involve making and testing a module that can be used by CMS Software to produce the discriminator in centrally produced data samples. In addition it would be necessary to create a method and framework for the calculation of scale factors and the validation of the tagger in data and MC samples. This would then lead to ongoing work for maintenance of the tagger, with the need for validation and scale factor calculations for every major data processing campaign. The second proposal, which is potentially tied closely with the first, is the improvement of the tagger itself. There are two immediate things that come to mind when considering improving

the tagger, and those are optimizing the machine learning algorithm so that it is better tailored to using secondary vertices as inputs, rather than using a basic algorithm, and then adding in track related variables as inputs into the training, along with any other useful low-level information that may be pertinent to a secondary vertex. There is a lot of promise for continued work on a machine learning based soft b-tagger.

On the stop search, there are a couple of improvements that can be made, in conjunction with any improvements made to the soft b-tagger. The first, and possibly most obvious improvement is the use of more data, when the Run III data taking period begins. On the improvements to the detectors, those of note are the maintenance on the pixel detector, which involves replacing layer 1, and the DCDC converters, which should produce marked tracking improvements after its deteriorated running by the end of the 2018 data taking period, and the installation of the new readout system for HCAL, which, along with its other upgrades, should improve the energy resolution and data readout for Run III. For the analysis proper, improvements can be made at multiple stages. For the Recursive Jigsaw Reconstruction step, one definite improvement that could be done is taking advantage of counting objects in the different sparticle hemispheres. In this version of the analysis, there was little consideration for whether the leptons, jets, and SVs were assigned to the same or opposite sides of the sparticle system. Also, studies could be done in order to find better mass sensitive variables to bin events in. Additionally, with RJR's ability to target generic decay scenarios, it is possible to look at more complicated SUSY models which have a greater amount of particle content and multiple decay modes in the same model, as RJR is able to account for these by just including the final states corresponding to all the decay modes present for any given sparticle production. The last thing that could be improved upon is the actual limit setting. With the way the analysis was set up, the number of bins was starting to push against the upper limit of what the limit setting software was capable of. The software would either need to be upgraded, or remade such that it can handle more bins efficiently than what was possible in this analysis.

Bibliography

[1] Joshua Ellis, Comput. Phys. Commun. **210**, 103 (2017), arXiv:1601.05437 [hep-ph] .

[2] CMS Supersymmetry Physics Results.

[3] L. Evans and P. Bryant, JINST **3**, S08001.

[4] E. Mobs, The CERN accelerator complex. Complexe des accélérateurs du CERN (2016), general photo.

[5] S. Chatrchyan *et al.* (CMS), JINST **3**, S08004.

[6] A. Saha (CMS), *Proceedings, 14th Topical Seminar on Innovative Particle and Radiation Detectors (IPRD16): Siena, Italy, October 3-6, 2016*, JINST **12** (02), C02033.

[7] H. E. Haber and G. L. Kane, Phys. Rept. **117**, 75 (1985).

[8] M. Tanabashi *et al.* (Particle Data Group), Phys. Rev. **D98**, 790 (2018).

[9] C.-N. Yang and R. L. Mills, Phys. Rev. **96**, 191 (1954), [,150(1954)].

[10] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, Phys. Lett. **47B**, 365 (1973).

[11] S. L. Glashow, Nucl. Phys. **22**, 579 (1961).

[12] P. W. Higgs, Phys. Rev. Lett. **13**, 508 (1964), [,160(1964)].

[13] G. Guralnik, C. Hagen, and T. Kibble, Phys. Rev. Lett. **13**, 585 (1964).

[14] F. Englert and R. Brout, Phys. Rev. Lett. **13**, 321 (1964).

[15] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

[16] A. Salam, *8th Nobel Symposium Lerum, Sweden, May 19-25, 1968*, Conf. Proc. **C680519**, 367 (1968).

[17] E. D. Bloom *et al.*, Phys. Rev. Lett. **23**, 930 (1969).

[18] M. Breidenbach, J. I. Friedman, H. W. Kendall, E. D. Bloom, D. Coward, H. DeStaebler, J. Drees, L. W. Mo, and R. E. Taylor, Phys. Rev. Lett. **23**, 935 (1969).

[19] J. Augustin *et al.* (SLAC-SP-017), Phys. Rev. Lett. **33**, 1406 (1974).

[20] J. Aubert *et al.* (E598), Phys. Rev. Lett. **33**, 1404 (1974).

[21] S. Herb *et al.*, Phys. Rev. Lett. **39**, 252 (1977).

[22] F. Abe *et al.* (CDF), Phys. Rev. Lett. **74**, 2626 (1995), arXiv:hep-ex/9503002 .

[23] S. Abachi *et al.* (D0), Phys. Rev. Lett. **74**, 2632 (1995), arXiv:hep-ex/9503003 .

[24] G. Danby, J. Gaillard, K. A. Goulianos, L. Lederman, N. B. Mistry, M. Schwartz, and J. Steinberger, Phys. Rev. Lett. **9**, 36 (1962).

[25] K. Kodama *et al.* (DONUT), Phys. Lett. B **504**, 218 (2001), arXiv:hep-ex/0012035 .

[26] S. Chatrchyan *et al.* (CMS), Phys. Lett. B **716**, 30 (2012), arXiv:1207.7235 [hep-ex] .

[27] G. Aad *et al.* (ATLAS), Phys. Lett. B **716**, 1 (2012), arXiv:1207.7214 [hep-ex] .

[28] M. Tanabashi *et al.* (Particle Data Group), Phys. Rev. **D98**, 141 (2018).

[29] M. Tanabashi *et al.* (Particle Data Group), Phys. Rev. **D98**, 161 (2018).

[30] D. Griffiths, *Introduction to elementary particles* (Wiley-VCH, 2008) p. 414.

[31] M. D. Schwartz, *Quantum Field Theory and the Standard Model* (Cambridge University Press, 2014) pp. 407–414, cd 595–598.

[32] J. Christenson, J. Cronin, V. Fitch, and R. Turlay, Phys. Rev. Lett. **13**, 138 (1964).

[33] A. Alavi-Harati *et al.* (KTeV), Phys. Rev. Lett. **83**, 22 (1999), arXiv:hep-ex/9905060 .

[34] V. Fanti *et al.* (NA48), Phys. Lett. B **465**, 335 (1999), arXiv:hep-ex/9909022 .

[35] B. Aubert *et al.* (BaBar), Phys. Rev. Lett. **86**, 2515 (2001), arXiv:hep-ex/0102030 .

[36] K. Abe *et al.* (Belle), Phys. Rev. Lett. **87**, 091802 (2001), arXiv:hep-ex/0107061 .

[37] R. Aaij *et al.* (LHCb), Phys. Rev. Lett. **110**, 221601 (2013), arXiv:1304.6173 [hep-ex] .

[38] R. Aaij *et al.* (LHCb), Phys. Rev. Lett. **122**, 211803 (2019), arXiv:1903.08726 [hep-ex] .

[39] K. Abe *et al.* (T2K), Nature **580**, 339 (2020), arXiv:1910.03887 [hep-ex] .

[40] A. Djouadi, Phys. Rept. **457**, 1 (2008), arXiv:hep-ph/0503172 [hep-ph] .

[41] S. Weinberg, Phys. Rev. D **13**, 974 (1976), [Addendum: Phys.Rev.D 19, 1277–1280 (1979)].

[42] L. Susskind, Phys. Rev. D **20**, 2619 (1979).

[43] A. Djouadi, Phys. Rept. **459**, 1 (2008), arXiv:hep-ph/0503173 [hep-ph] .

[44] M. J. G. Veltman, Acta Phys. Polon. **B8**, 475 (1977).

[45] G. R. Farrar and P. Fayet, Phys. Lett. **76B**, 575 (1978).

[46] S. P. Martin, Adv. Ser. Direct. High Energy Phys. **21**, 1 (2010), arXiv:hep-ph/9709356 .

[47] M. Tanabashi *et al.* (Particle Data Group), Phys. Rev. **D98**, 807 (2018).

[48] A. M. Sirunyan *et al.* (CMS), JHEP **10**, 244, arXiv:1908.04722 [hep-ex] .

[49] G. Aad *et al.* (ATLAS), JHEP **06**, 046, arXiv:1909.08457 [hep-ex] .

[50] A. M. Sirunyan *et al.* (CMS), Phys. Rev. D **101**, 052010 (2020), arXiv:1911.07558 [hep-ex] .

[51] A. M. Sirunyan *et al.* (CMS), JHEP **09**, 149, arXiv:2008.04422 [hep-ex] .

[52] G. Aad *et al.* (ATLAS), Eur. Phys. J. C **80**, 123 (2020), arXiv:1908.08215 [hep-ex] .

[53] M. Aaboud *et al.* (ATLAS), Phys. Rev. **D98**, 092012 (2018), arXiv:1806.02293 [hep-ex] .

[54] G. Aad *et al.* (ATLAS), Phys. Rev. D **101**, 072001 (2020), arXiv:1912.08479 [hep-ex] .

[55] A. M. Sirunyan *et al.* (CMS), Phys. Rev. Lett. **124**, 041803 (2020), arXiv:1910.01185 [hep-ex] .

[56] A. M. Sirunyan *et al.* (CMS), JHEP **08**, 150, arXiv:1905.13059 [hep-ex] .

[57] G. Aad *et al.* (ATLAS), Phys. Rev. D **101**, 052005 (2020), arXiv:1911.12606 [hep-ex] .

[58] A. M. Sirunyan *et al.* (CMS), JHEP **05**, 032, arXiv:1912.08887 [hep-ex] .

[59] A. M. Sirunyan *et al.* (CMS), JHEP **02**, 015, arXiv:1910.12932 [hep-ex] .

[60] A. M. Sirunyan *et al.* (CMS), Eur. Phys. J. C **81**, 3 (2021), arXiv:2008.05936 [hep-ex] .

[61] A. M. Sirunyan *et al.* (CMS), Phys. Lett. **B782**, 440 (2018), arXiv:1801.01846 [hep-ex] .

[62] A. M. Sirunyan *et al.* (CMS), JHEP **09**, 065, arXiv:1805.05784 [hep-ex] .

[63] *LEP design report* (CERN, Geneva, 1984) copies shelved as reports in LEP, PS and SPS libraries.

[64] S. I. Cooper (CMS), *Proceedings, 37th International Conference on High Energy Physics (ICHEP 2014): Valencia, Spain, July 2-9, 2014*, Nucl. Part. Phys. Proc. **273-275**, 1002 (2016).

[65] *The CMS Phase-1 Pixel Detector Upgrade*, Tech. Rep. arXiv:2012.14304 (2020).

[66] M. Lipinski (CMS Collaboration), *The Phase-1 Upgrade of the CMS Pixel Detector*, Tech. Rep. CMS-CR-2017-135. 06 (CERN, Geneva, 2017).

[67] P. Jackson and C. Rogan, Phys. Rev. D **96**, 112007 (2017), arXiv:1705.10733 [hep-ph] .

[68] P. Jackson, C. Rogan, and M. Santoni, Phys. Rev. D **95**, 035031 (2017), arXiv:1607.08307 [hep-ph] .

[69] L. Moneta, K. Belasco, K. S. Cranmer, S. Kreiss, A. Lazzaro, D. Piparo, G. Schott, W. Verkerke, and M. Wolf, in *13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT2010)* (SISSA, 2010) PoS(ACAT2010)057, arXiv:1009.1003 [physics.data-an] .

[70] I. Antcheva *et al.*, Comput. Phys. Commun. **180**, 2499 (2009), arXiv:1508.07749 [physics.data-an] .

[71] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Eur. Phys. J. C **71**, 1554 (2011), [Erratum: Eur.Phys.J.C 73, 2501 (2013)], arXiv:1007.1727 [physics.data-an] .

[72] R. Barnett, H. Haber, and G. Kane, Nuclear Physics B **267**, 625 (1986).

[73] M. Chen, C. Dionisi, M. Martinez, and X. Tata, Physics Reports **159**, 201 (1988).

[74] J. Alwall, M.-P. Le, M. Lisanti, and J. G. Wacker, Phys. Rev. D **79**, 015005 (2009), arXiv:0809.3264 [hep-ph] .

[75] J. Alwall, P. Schuster, and N. Toro, Phys. Rev. D **79**, 075020 (2009), arXiv:0810.3921 [hep-ph] .

[76] A. Sirunyan *et al.* (CMS), JINST **13** (05), P05011, arXiv:1712.07158 [physics.ins-det] .

[77] A. M. Sirunyan *et al.* (CMS), JHEP **10**, 005, arXiv:1707.03316 [hep-ex] .

[78] S. S. Mehta, DeepJet: a portable ML environment for HEP. 2nd IML Machine Learning Workshop (2018) Apr, 2018.

[79] J. Kieseler, M. Stoye, M. Verzetti, P. Silva, S. S. MEHTA, A. Stakia, Y. IIYAMA, E. Bols, S. R. QASIM, H. KIRSCHENMANN, H. Qu, M. Rieger, and L. Gouskos, Deepjetcore (2020).

[80] M. Verzetti (CMS), EPJ Web Conf. **214**, 06010 (2019).

[81] K. Ehatäht, EPJ Web Conf. **245**, 06002 (2020).

[82] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, JHEP **07**, 079, arXiv:1405.0301 [hep-ph] .

[83] E. Re, Eur. Phys. J. C **71**, 1547 (2011), arXiv:1009.2450 [hep-ph] .

[84] R. D. Ball *et al.* (NNPDF), Eur. Phys. J. C **77**, 663 (2017), arXiv:1706.00428 [hep-ph] .

[85] R. D. Ball *et al.* (NNPDF), JHEP **04**, 040, arXiv:1410.8849 [hep-ph] .

[86] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, Comput. Phys. Commun. **191**, 159 (2015), arXiv:1410.3012 [hep-ph] .

[87] S. Agostinelli *et al.* (GEANT4), Nucl. Instrum. Meth. A **506**, 250 (2003).

[88] A. Sirunyan *et al.* (CMS), JINST **12** (10), P10003, arXiv:1706.04965 [physics.ins-det] .

[89] M. Tosi (CMS), PoS **EPS-HEP2017**, 523 (2017).

[90] K. Rehermann and B. Tweedie, JHEP **03**, 059, arXiv:1007.2221 [hep-ph] .

[91] A. M. Sirunyan *et al.* (CMS), Phys. Rev. D **101**, 052010 (2020), arXiv:1911.07558 [hep-ex] .

[92] *Muon Identification and Isolation efficiency on full 2016 dataset*, Tech. Rep. CMS-DP-2017-007 (2017).

[93] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **04**, 063, arXiv:0802.1189 [hep-ph] .

[94] *Jet Performance in pp Collisions at 7 TeV*, Tech. Rep. CMS-PAS-JME-10-003 (CERN, Geneva, 2010).

[95] M. Stoye (CMS), J. Phys. Conf. Ser. **1085**, 042029 (2018).

[96] V. Khachatryan *et al.* (CMS), JHEP **03**, 136, arXiv:1102.3194 [hep-ex] .

[97] *Plans for Jet Energy Corrections at CMS*, Tech. Rep. CMS-PAS-JME-07-002 (CERN, Geneva, 2008).

[98] A. M. Sirunyan *et al.* (CMS), JINST **15** (09), P09018, arXiv:2003.00503 [hep-ex] .

[99] T. Miao, N. Leioatts, H. Wenzel, and F. Yumiceva, *Beam Position Determination using Tracks*, Tech. Rep. CMS-NOTE-2007-021 (CERN, Geneva, 2007).

Appendix A

A Tracking Independent Measurement of the CMS Beam Spot

The beam spot of the CMS experiment is the luminous region produced by the collision of the two counter-rotating proton beams. A precise measurement of the beam spot is necessary for the reference point it provides in the calculation of many physics observables. To give more concrete examples, a precise measurement of the beam spot location is needed as an estimate of the primary interaction point for the High Level Trigger, and physics object reconstruction, to improve the momentum resolution of tracks from the primary vertex, determining track impact parameter resolution, checking the global alignment of the tracking detector, and for real time monitoring of the radiation dose to the tracker detector in addition to providing feed-back to the accelerator operators [99].

The official beam spot measurement is done using the d0-ϕ algorithm [99], which makes use of reconstructed tracks in a maximum likelihood fit to determine the location. Due to its need for tracks, there is some computational overhead from the track reconstruction that limits the speed at which the beam spot location can be determined, especially when it needs to be done online during data taking. This led to the question of whether it is possible to measure the location without having to perform track reconstruction, a tracking independent measurement.

Over the course of this study multiple different fitting methods were used to try to calculate the beam spot displacement. Starting with the assumption that hit density should scale as $1/r^2$, the occupancy of layer 1 of the pixel detector was measured and fits using various equations and methods were performed in order to measure beam spot displacement. These fit methods started with a simple sine function and through multiple iterations, arrived at a final method corresponding to a maximum likelihood fit of a power function.

A.1 Production of MC

In order to perform studies on the beam spot location without tracking information, it was important to have enough simulated Monte Carlo with a wide array of beam spots to test over, as one concern with using tracking independent methods was the sensitivity it would have to a minimally displaced beam spot. Following this a wide range of MC samples were produced, with different beam spot locations, and different conditions ranging from the simplest simulations based on a perfect detector (called design MC) with no variation in the z position of the interaction point, to more realistic conditions which contain pileup.

Apart from the centrally produced MC, amounting to a couple of design MC samples with a $(0,0,0)$ displaced beam spot and a slightly shifted one, many more samples were produced allowing for different degrees of beam spot displacement, and spread of interaction points. In the descriptions of the beam spots (x,y,z) refers to the x, y, and z positions in CMS coordinates, and units of centimeters. The simulations were all produced using a 'RelVal TTbar' tagged data card, with the number of events ranging from 10,000 to 125,000. A full list of the produced design MC samples can be found in Table A.1. There were in addition a couple of centrally produced samples that used realistic conditions with a beam spot of $(0.01,0.04)$ cm. These were used in the beginning to see what the modules looked like, but most of the study focuses on design MC, as many of the fitting methods used were exploratory in nature, and the realistic MC introduced too many complications in determining whether a fitting method was working.

Before making the decision to use design MC in the fitting process, the first studies were done using real data from express and rereco datasets, collected during the end of the 2018 run period. The data is from run 324970 in the Express dataset and 325169 in the Rereco dataset. The main reason for using real data instead of MC was due to its ready availability in the needed format. After going through a couple different fitting methods, it was decided that it would be better to simplify the data being used by moving to MC, which removes a lot of the variation that could be seen in data. The main sources of variability in these cases are the bad modules, pileup, and potential shifts of the beam spot during the data taking. By moving to MC, it was possible to make

Beam Spot (cm)	z-smearing	pileup	N Events
$(0,0,0)$	None	None	125,000
$(0,0,10)$	None	None	100,000
$(0,0,-10)$	None	None	100,000
$(0.03,0.06,0)$	None	None	100,000
$(-0.08,0,02,0)$	None	None	100,000
$(0.1,-0.08,0)$	None	None	100,000
$(0.1,-0.08,10)$	None	None	100,000
$(0.1,-0.08,-10)$	None	None	100,000
$(-0.1,0.2,0)$	None	None	100,000
$(0.2,0.19,0)$	None	None	100,000
$(-0.3,-0.32,0)$	None	None	100,000
$(0,0,0)$	4 cm	None	225,000
$(0.01,0.04,0)$	4 cm	None	110,000
$(0.1,-0.08,0)$	4 cm	None	110,000
$(0.1,-0.08,-10)$	4 cm	None	100,000
$(-0.1,0.08,0)$	4 cm	None	100,000
$(0.1,0.08,0)$	4 cm	None	100,000
$(0.2,-0.19,0)$	4 cm	None	100,000
$(0.3,-0.32,0)$	4 cm	None	100,000
$(0.5,-0.48,0)$	4 cm	None	100,000
$(1.0,-0.95,0)$	4 cm	None	100,000
$(0.2,-0.19,0)$	None	2018 PU	100,000
$(0.3,-0.32,0)$	None	2018 PU	100,000
$(0.5,-0.48,0)$	None	2018 PU	100,000
$(1.0,-0.95,0)$	None	2018 PU	100,000

Table A.1: A list of the Design MC samples produced with varying beam spot parameters. The first column details the beam spot position in (x,y,z), the second column whether there is a z-smearing of the interaction point, the third whether pileup was simulated, and lastly the approximate number of events in the sample.

sure that all modules would perform in their ideal case, all interactions would occur at the same z position, and the beam spot would remain constant. This allows for focusing on the fitting method, rather than trying to account for variations before even having a reliable way to calculate the beam spot.

A.2 Nomenclature for the pixel detector

Before continuing on the subject of determining a fit method for measuring the beam spot, it is important to clarify the nomenclature that will be used later on. Some of the terms used are specific to this study, and are not necessarily used by others in the CMS experiment.

The Bpix detector is made up of layers of modules. Each module is made of 16 ROCs, with 8 ROCs on each side. Each layer of Bpix consists of a number of ladders, made up of a line of modules in the z-direction. The ladders are staggered in order to have full coverage over ϕ.

Layer 1, which is the focus of the study, consists of 12 ladders, each containing 8 modules. The 12 ladders are staggered at 3 different radii, with the innermost ladders falling at an average radius of 2.77 cm, called the inner ladders. The set of ladders at the next radius are referred to as the outer ladders, for a total of 4 outer ladders, ignoring the last two which are set at a further out radius. With the arrangement of the ladders for full coverage over ϕ, the modules in each of the ladders of a given z slice align to form module rings, with 8 module rings for the 8 modules in a ladder. Figure A.1 gives a caricature of the module rings following the z direction, along with a numerical designation for each ring. For the latter part of this study, this is further separated into ROC rings, where each set of ROCs in a given z slice form such a ring, for a total of 64 such rings. Also for the purpose of this analysis, each ladder is separated further into a half-ladder, corresponding to a line of ROCs in the z-direction. Figures A.2 and A.3 give examples of the ladders as they appear down

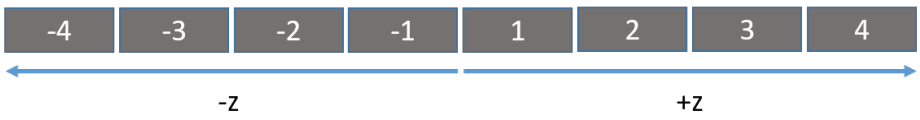


Figure A.1: A figure showing the module rings over z, and their naming convention

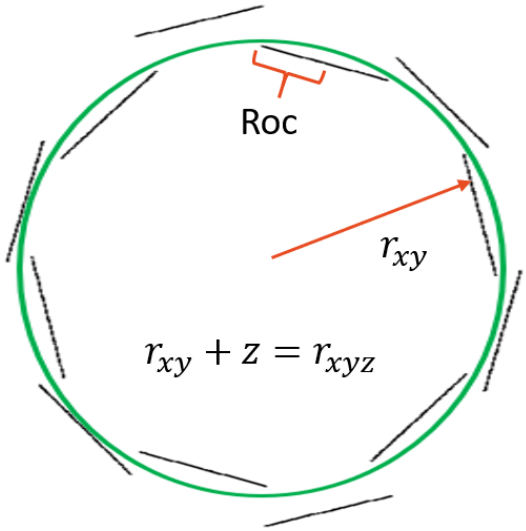


Figure A.2: A plot of the ladders of Bpix layer 1 with cartoons detailing different parts. The black lines are the ladders, showing the length of the modules. The green circle surrounds the inner ladders of layer 1. The red bracket shown outlines a ROC which is half of a module, corresponding to the half-ladders. The red arrow shows r_{xy} which is the 2D radius in xy from the center of the detector to the relevant layer.

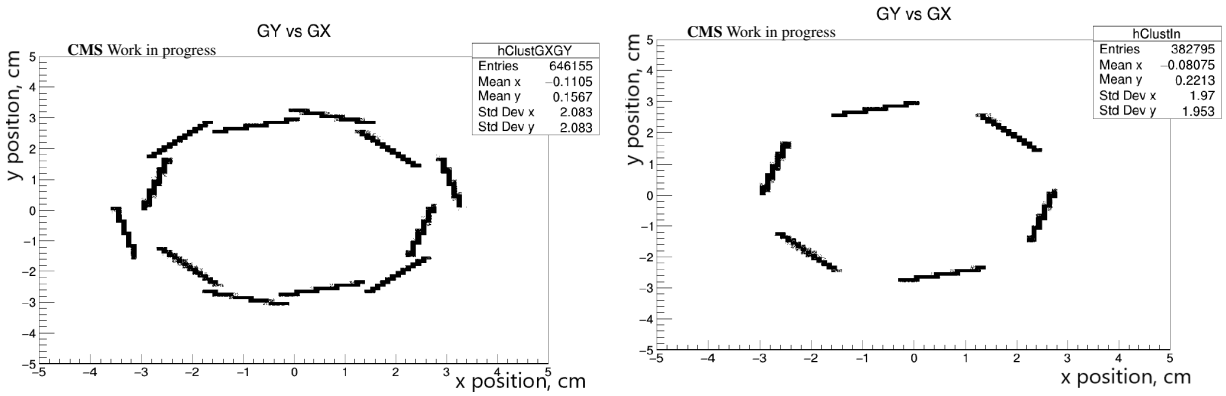


Figure A.3: A plot of the ladders of Bpix layer 1, looking down the z-direction. The plot on the left shows all 12 ladders, while the plot on the right shows just the inner ladders. Cutting each of these bars in half would then give the half-ladders.

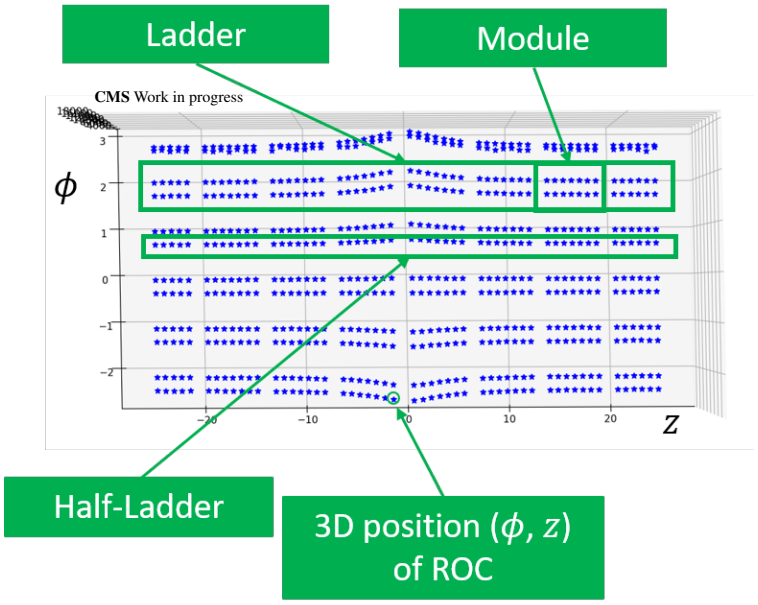


Figure A.4: A figure showing the arrangement of ROCs in the occupancy, ϕ, z space of the inner ladders of BPix layer 1. Highlighted in the boxes are examples of a ladder, half-ladder, module, and ROC.

the z-axis, where the first figure labels some features of layer 1 and the second figure shows layer 1 in the Global x and y plane. These are the ladders of Bpix layer 1, where one can see the three different radii the ladders sit at. The second plot shows just the inner ladders of layer 1, which have been a focus of the study. Another helpful picture is given in Figure A.4, which attempts to label the different features of the pixel detector in a top down view of the 3D space of occupancy, ϕ and z. This figure highlights the arrangement of the ladders, half-ladders, modules, and ROCs as they are spaced in CMS coordinates.

The positions for the occupancies, pixels, and ROCS are determined by the 'Global' classifier, which uses the center of the CMS detector as the origin. This means that the Global x, y, and z positions are with respect to the $(0,0,0)$ position of the CMS detector, otherwise called CMS coordinates. There are two radii used in the following fits, the first is r_{xy} which is the 2D radius in x and y coordinates, shown in Figure A.2, calculated using the measured global positions. The second radius is the 3D radius, r_{xyz}, which is the vector sum of r_{xy} and the z position.

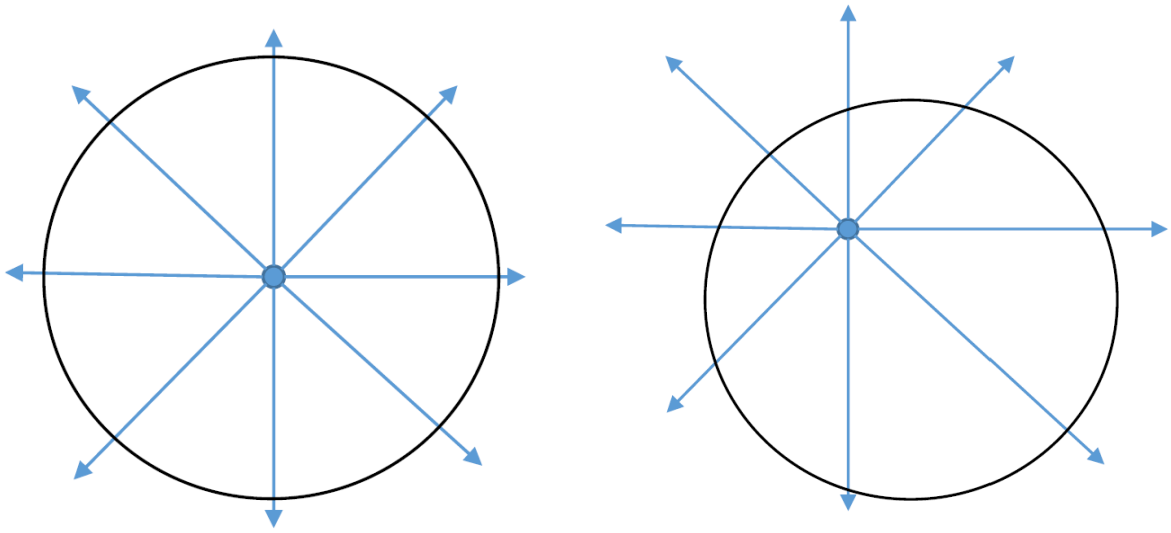


Figure A.5: A simple figure depicting a point source surrounded by a circle. In the left picture, the circle is centered on the point source, with the flux lines equally distributed around the circle. In the right plot, the point source is shifted from the center, showing the flux lines being more concentrated towards the direction of displacement.

A.3 Determining a fit method

By removing tracking reconstruction, one is left with just the hits from the tracker system, with a focus from the pixel detector due to its close proximity to the beam line. In order to measure the location of the beam spot, one must first determine a possible relation between the hits from the pixel detector and their distance from the beam line, the more obvious possible relation being the flux of a point source spherically radiating outward. For such a point source, if a circle is drawn around it, one can note the flux passing through such a circle. If the source is centered on the circle, the flux will be uniformly distributed around it. Moving the point source closer to one edge of the circle will see a larger concentration of flux over a smaller area, in the direction of the source's displacement. This is simply illustrated in Figure A.5. This idea can be directly applied to the pixel detector (the circle) surrounding the particle collisions (the point source), and while this is a very simplified image, since the full detector is a cylindrical shape, it is a good place to start.

Assuming the case of a radiating point source, the hit density from the pixels of the pixel detector should scale by the inverse square of the distance from the interaction point. The next

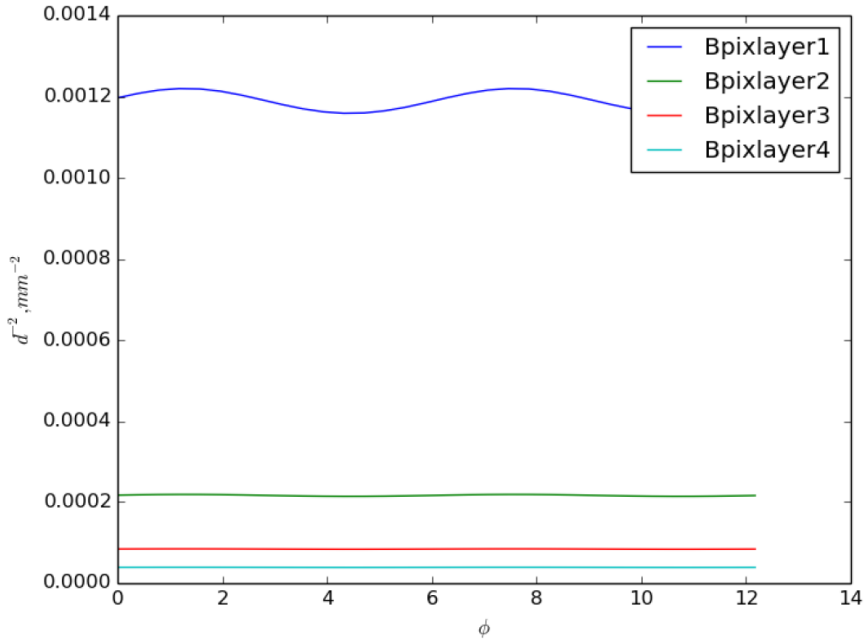


Figure A.6: A plot showing how changes in hit density scale for the radius of each layer of the barrel pixel detector, for a beam spot of $(0.01, 0.04)$ cm. The layers have an average radius of 2.9, 6.8, 10.9, and 16.0 cm, respectively, for layers 1-4.

question to answer, then, is at which distance (radius) can this effect be seen before any small changes in displacement are washed out by a much larger distance to the surrounding detector layers.

A check was made, looking at the distribution of $1/(r_{xy} - d_{bs})^2$ versus ϕ for each pixel layer, where r_{xy} is the average cylindrical radius (only using x and y components, no z) of the layer and d_{bs} is the beam spot displacement in the x-y plane. This was done for a slightly displaced beam spot of $(0.01, 0.04)$ cm. The expectation is that this distribution will create a distinct sinusoidal shape over ϕ, as shown in Figure A.6. The question here, is to what degree a sinusoidal shape is present for each layer. Each line shape in Figure A.6 corresponds to one of the four pixel layers, and as can be seen, the beam displacement would best be seen in layer 1, with each successive layer showing less sensitivity to the chosen displacement. From this it was decided to focus on the first pixel layer, eventually only considering the inner ladders, which have a fairly consistent radius at a smaller value than the other ladders of layer 1. While it would be possible to use the

hit information of the ladders and pixel layers at larger radii, more importance was placed on the inner most ladders in order to reduce the complexity of a measurement and to focus on the hits at the smallest possible radii, for increased sensitivity to any beam spot displacement.

In the following sections, different fit methods used in an attempt to measure the beam spot displacement will be described. This starts with the use of a basic sine function, fitting over occupancy versus ϕ. Following the sine function, an elliptical expression was used, before moving to a ratio of power functions. The final iteration of fits made use of a maximum likelihood scan, where the base equation still assumed a power law, but all of the parameters were allowed to float, removing the assumption that the occupancy scales as $1/r_{xyz}^2$, and allowing the fit to choose the power. The first three methods only made use of r_{xy}, ignoring the z position. The radius was then changed to r_{xyz} in the maximum likelihood fit to take advantage of the z position of the clusters.

A.3.1 Data preparation

Before performing any fitting, the data first had to be prepared and cleaned. Overall the data selection and cleaning was a fairly simple process. First off, only the clusters that were located in layer 1 of the barrel pixel detector were kept, all other clusters were not considered. This was done by only selecting the detector IDs that correspond to the 1st pixel layer. The only other persistent selection was that clusters could not be centered on edge pixels. For the case of this study, edge pixels are defined as any pixel that is within 8 rows or columns of the edge of a ROC. Other cluster requirements that were tried out, but eventually dropped were cuts on the size of the pixel cluster. Initially it was required that the size of the clusters be small, first 1 pixel in size to try to only look at clusters with normal incidence to the detector. This was later changed to ≥ 2 pixels in size, as there were potential problems when looking at 1 pixel clusters. The final selection removes this cut entirely due to it removing a large fraction of necessary clusters at central z position. Charge cuts were not considered when looking at the occupancy, for the same considerations as the cluster size. A charge cut is only considered when taking a look at total charge collected, instead of cluster occupancy, detailed in Section A.4.

In addition to the selection of the clusters, there are some further requirements in order to remove under/over-performing pixels, ROCs, or modules. This particular selection is more particular to the fitting method used, mainly in how the data is provided to the fitting. This selection will be described in the sections corresponding to the fitting method used.

A.3.2 Sine fit

The first fitting method used was a simple sine function, given by Equation A.1, where ϕ is the global ϕ of the detector. For displaced beam spots, a sinusoidal relation would be expected, while a flat line would occur for beam spots centered at $(0,0)$ relative to the pixel detector. The data used as input were real data from the express physics dataset. The information used from this data were the occupancy and ϕ location of the clusters, put into a binned histogram, with some examples given in Figure A.7. The histograms were additionally scaled by the number of events processed, which was later determined to be unnecessary, where the figures shown in this section reflect the scaling that was performed. After removing edge pixels, the histogram had some artifacts on the edges of the filled bins, where these bins correspond to the transition point between edge pixels and more central pixels. As an extra cleaning, these particular bins were removed from consideration in the fit. All of the fits done in this section used the MINUIT minimization program to perform the fitting.

$$a\sin(b\phi - c) + d \tag{A.1}$$

Using Equation A.1, a fit is performed using a cleaned version of the right plot in Figure A.7. The result of the fit is given in Figure A.8, and while it did converge, the χ^2/ndf value did not necessarily indicate a good fit. Visually, the fit equation looks to fit the data, but other factors indicate some problems. The most glaring problem, that is addressed later, is the fact that the results of the fits do not directly give a way to calculate a beam spot value. Too much focus was put on trying to fit a sinusoidal function, and not enough on the end result of a beam spot value.

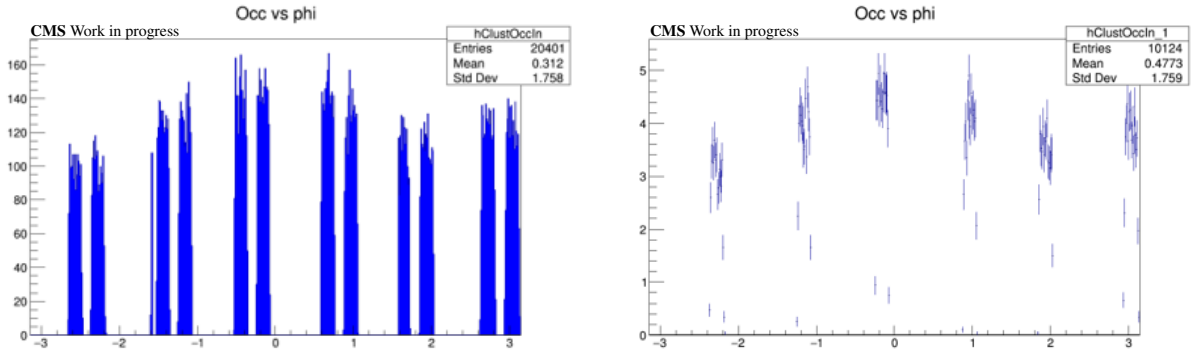


Figure A.7: Two histograms showing the occupancy over phi for the inner ladders of Bpix layer 1. The left plot shows the full ladders, while the right shows a set of half ladders from the left side plotted as points to highlight the low bins. The bins on the edge of the half ladders have lower values than the more central bins due to the edge pixel removal affecting the occupancy in those bins.

The second problem is that the equation that resulted from the fit did not circumscribe a circle as would be expected. Figure A.9 shows the fit equation plotted in polar coordinates, and due to the lack of closure, indicates that the fit is not ideal. An ideal fit would give a closed circle, with the location of the center of the circle being proportional to the beam spot displacement.

From the results of this fitting there were two major questions to answer. The first was how to ensure that the fit will give an enclosed circle, rather than what is given in Figure A.9. The second question is how to calculate the beam spot once a good fit is obtained. For the first question, the answer that was arrived at is to force the circle to close by using the equation for a circle, or rather an ellipse as the general form to try to account for any deformations in the shape of layer one due to small inconsistencies in the radius of the module. This also partially answers the second question in how to obtain the beam spot displacement in reasonable units of measure, where it is fairly simple to write the equation of an ellipse in terms of Cartesian coordinates.

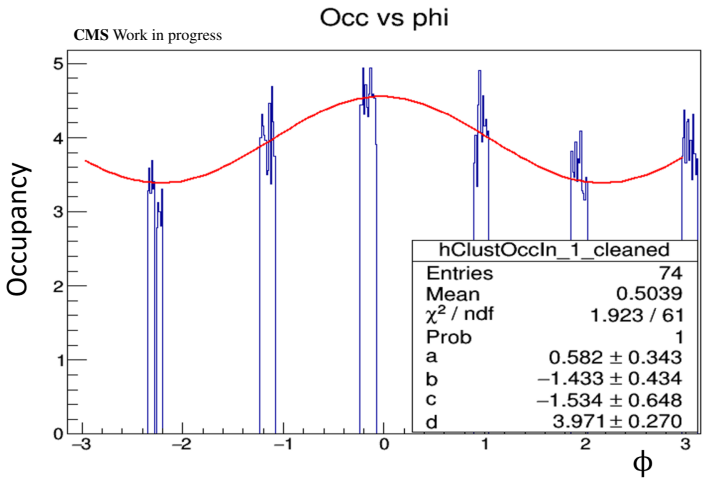


Figure A.8: A histogram of the set of cleaned half ladders from Figure A.7, with a sine fit overlayed. The box gives the fit results for the sine function, Equation A.1.

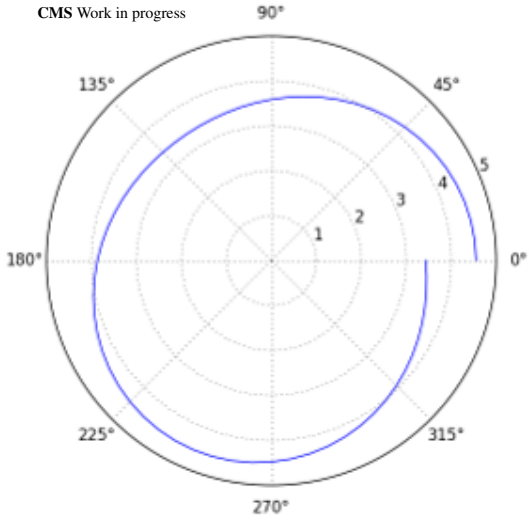


Figure A.9: The fit equation from Figure A.8, plotted in polar coordinates. For an ideal fit, the equation should circumscribe a circle, for which the center would be proportional in some way to the beam spot displacement.

A.3.3 Elliptical Fit

The next equation used in fitting was for an ellipse, which is the next logical step when moving from a sinusoidal function. The main difference from a sine function is that the ellipse forces closure over the angles from 0 to 2π, and in addition it is somewhat easier to work with Cartesian coordinates, which is useful when trying to get the x and y components of the beam spot. For this method, Equation A.2 is what is used for fitting. Instead of using the occupancy over ϕ as in the previous fitting method, this one instead uses the occupancy in x and y coordinates. To do this the occupancy is separated into sections of ϕ, much like the binning seen in Figure A.7. The occupancy in each of these bins is then separated into an x and y component using the central ϕ value of the bin as the direction, and the occupancy value in that bin as the magnitude. As in the previous section, the occupancies were scaled by the number of events processed, where this scaling was not removed until the fitting method described in the following section. All of the fits done in this section used the MINUIT minimization program to perform the fitting.

$$\frac{((x-x_0)\cos a-(y-y_0)\sin a)^2}{r_a^2}+\frac{((x-x_0)\cos a-(y-y_0)\sin a)^2}{r_b^2} \tag{A.2}$$

Before working with the occupancy, the ellipse fitting method first needed to be tested on something that would give definitive results, in an ideal case. For this check, instead of the binned occupancy, and its x and y components, what was used was the x and y position, in CMS coordinates, of each cluster measured in the data. This removed the need to worry about unit conversion, and the results of the x_0 and y_0 parameters would be given in known units, corresponding to where the center of layer 1 is with respect to CMS coordinates. In addition, for this case, there is no need to worry about underperforming ROCs or modules, due to the fit not being dependent on the occupancy. Figure A.10 gives some distributions of the x and y position of the clusters in layer 1, with the ellipse fit overlayed. The two fits are done using module sections 4 and -4. The parameter values, also seen in Figure A.10, give accurate values of the radius expected at layer one, which has an average radius of approximately 2.77 cm. For the center value, a beam spot of $(0.01, 0.04)$ cm

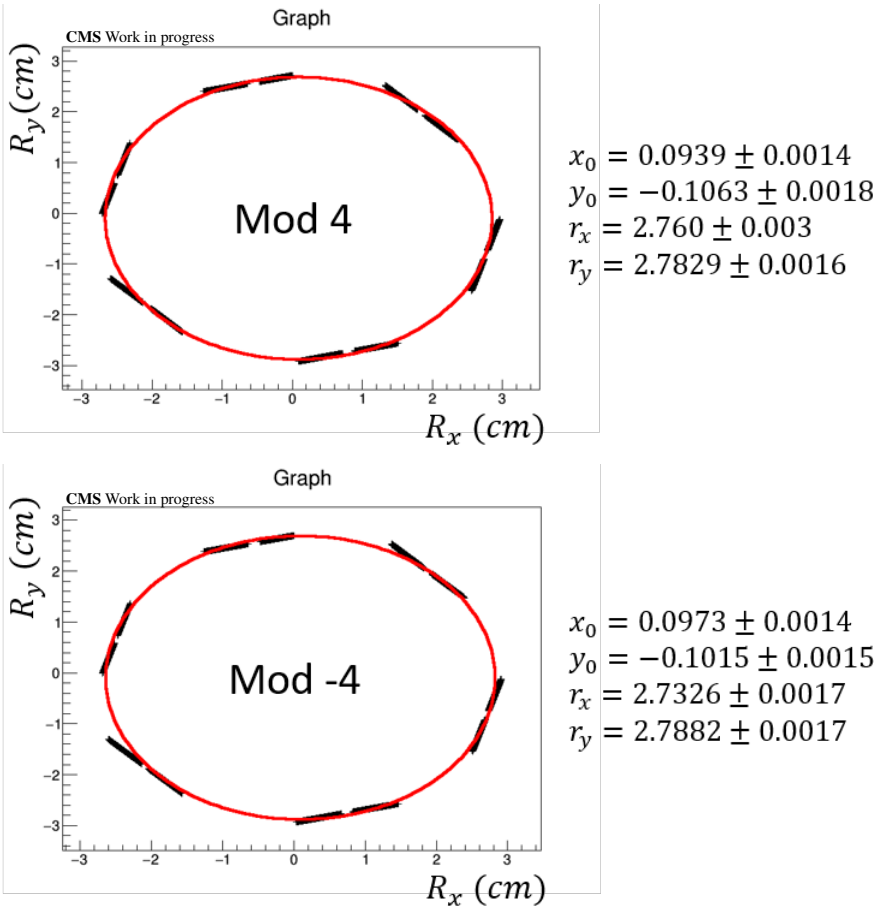


Figure A.10: Positions of the clusters in x-y coordinates, with ellipse fits overlayed. The top plot shows the fit for module section 4, while the bottom plot shows the fit for module section -4. To the right of each plot are the fitted values for x_0, y_0, r_x, and r_y. Each of these fitted values closely match their true values, corresponding to the BPix center in CMS coordinates, and the radii of the inner ladders of BPix layer 1.

with respect to BPix center was expected for the 2018 run period, with a value in CMS coordinates of roughly $(0.096, -0.067)$ cm. This comes out to a BPix center of approximately $(0.086, -0.107)$ cm, give or take a couple 0.01 cm due to beam drift over time, which closely corresponds to what was determined by the fit. So for this ideal case, fitting with Equation A.2 accurately describes the center of an ellipse formed by the inner ladders of BPix layer 1. This leads to the whether this can be applied to occupancy in the same way.

In order to perform the fits on the occupancy, it was necessary to determine how the data should be treated when making the x-y distributions. Since the fitting was initially done on real

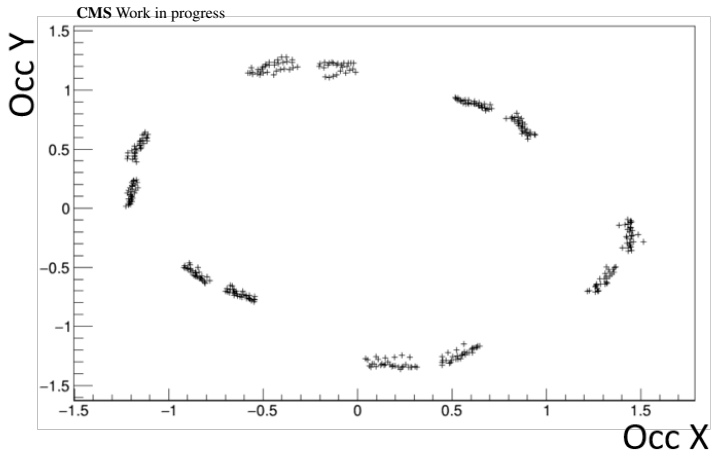


Figure A.11: Occupancy of the inner ladders of layer 1 in x-y coordinates. This occupancy uses data from the Express physics dataset.

data, it was important to separate out data, so that dependence on the poorly performing modules is mitigated as much as possible. Figure A.11 shows the occupancy in x-y coordinates for all of the inner ladders of layer 1 using the Express dataset. One thing that can be noticed from this distribution is that the ellipse formed by the distribution is fairly lopsided, and not what one would expect, considering the distribution of the x and y position of the ladders as seen in Figure A.3. The positions of the ladders in each of these distributions should be roughly proportional to each other, as one should describe the position of BPix layer 1 within CMS coordinates, while the other should be proportional to the displacement of the beam spot from those same CMS coordinates.

To better understand what there is to work with, with the data, the plot in Figure A.11 was separated into module rings to see how each ring is affecting the overall occupancy. Figures A.12 and A.13 show the scaled occupancy for each of the module rings. As can be seen, for the Express dataset, none of the module rings are showing reliable performance that can be used for the ellipse fitting. The data from the Rereco dataset was then checked, and it was found that a couple of the module rings from this dataset showed uniform performance between the ladder sections, rings -3, -2, 2, and 3, with examples seen in Figure A.14. Before performing the ellipse fits, the occupancy distribution was transformed to show $1/\sqrt{occ}$ instead of occupancy, so that the results would be directly proportional to the radius, still assuming a $1/r_{xy}^2$ dependence. The elliptical fits were done

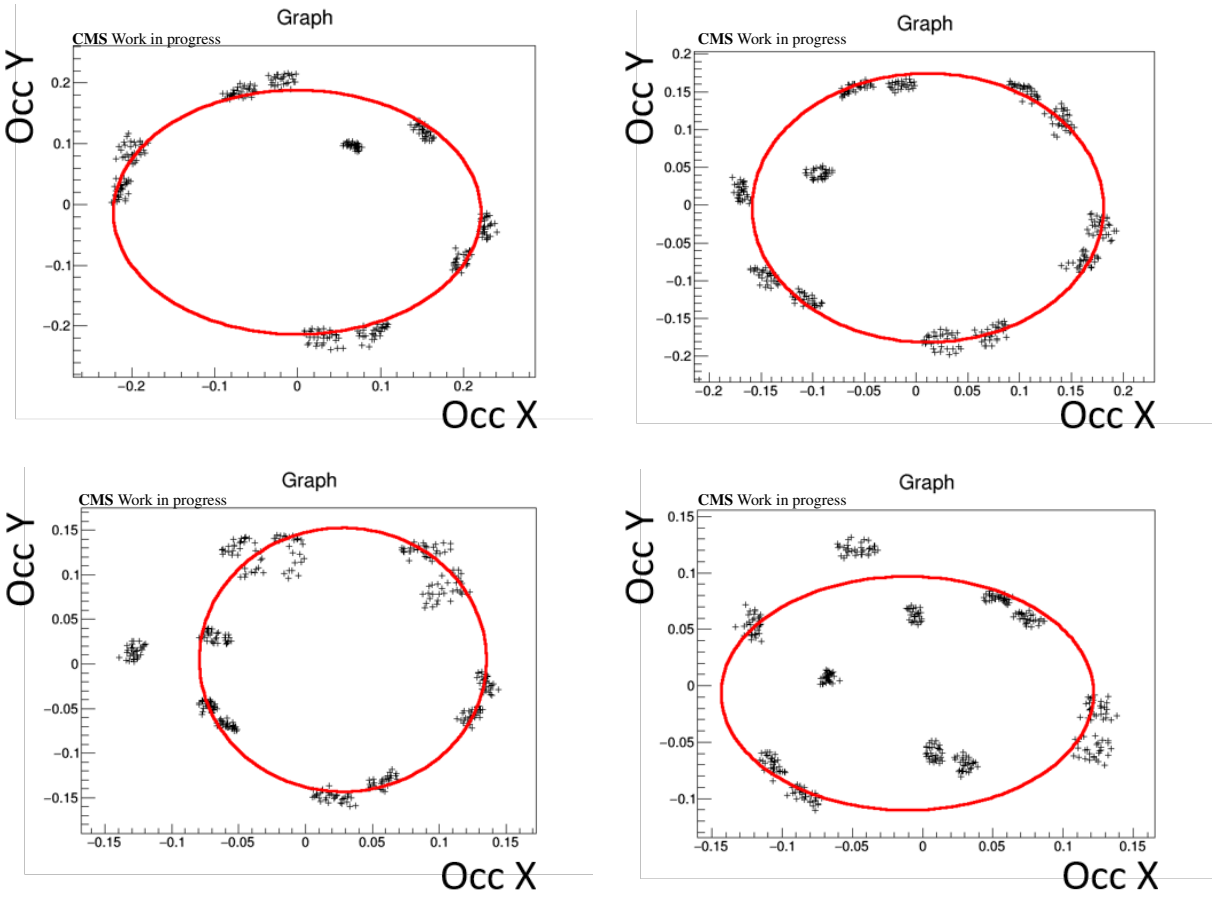


Figure A.12: Scaled occupancy distributions for the 4 module rings on the $+z$ side of layer 1, with an attempted ellipse fit overlayed on top. From top left to bottom right, the plots corresponds to module rings 1, 2, 3, 4

for these four module sections, with the results of x_0 and y_0 given in Table A.2. What may be apparent from this table is that the values found for x_0 and y_0 are inconsistent between the module sections, and there are no units on these numbers. With these considerations, one would need an ideal performance in the modules to be able to get a good fit, where small change in the occupancies for different modules will affect results significantly, additionally there is no trivial way to convert the units from occupancy to centimeters. The next fitting method tries to answer the problem of units by ensuring the fitted results will come out in centimeters through the use of ratios.

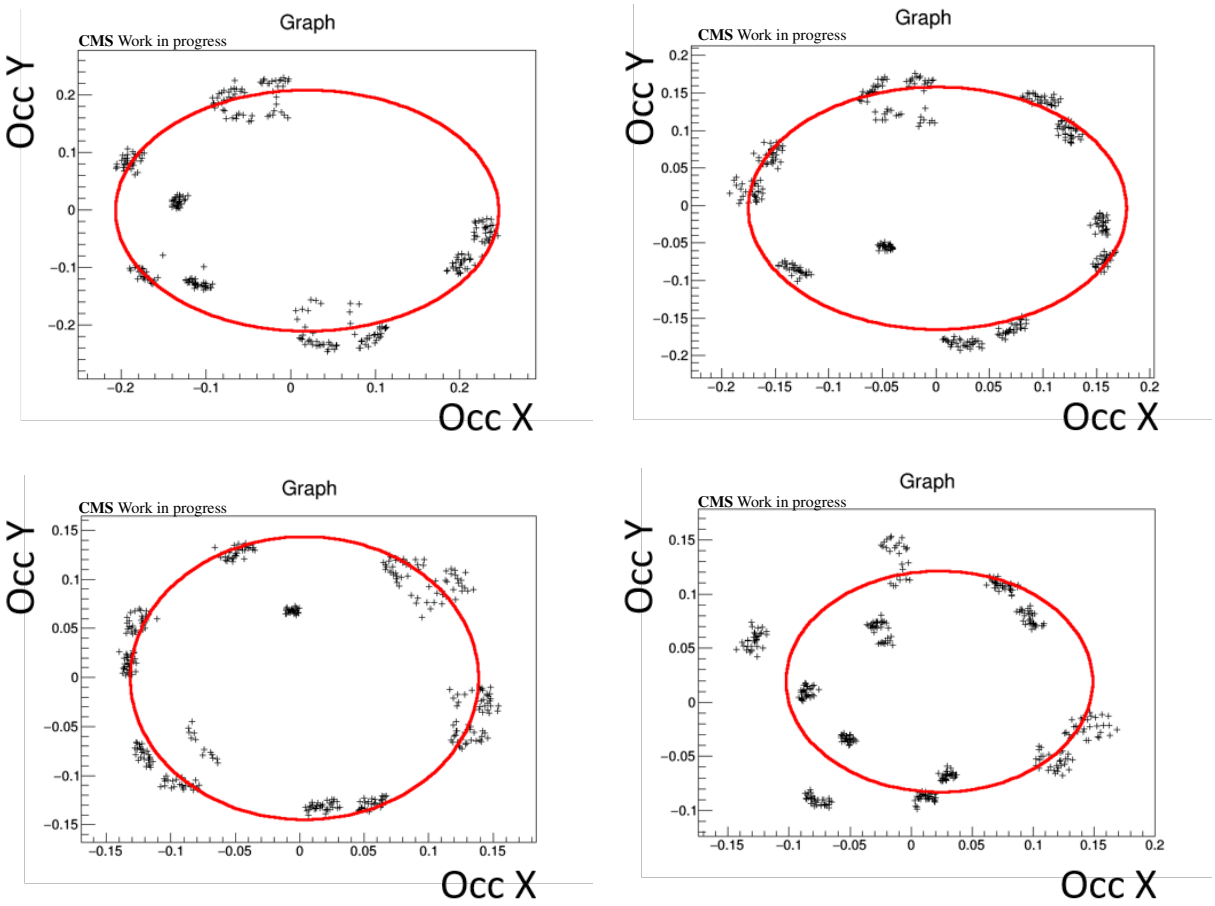


Figure A.13: Scaled occupancy distributions for the 4 module rings on the $-z$ side of layer 1, with an attempted ellipse fit overlayed on top. From top left to bottom right, the plots corresponds to module rings -1, -2, -3, -4.

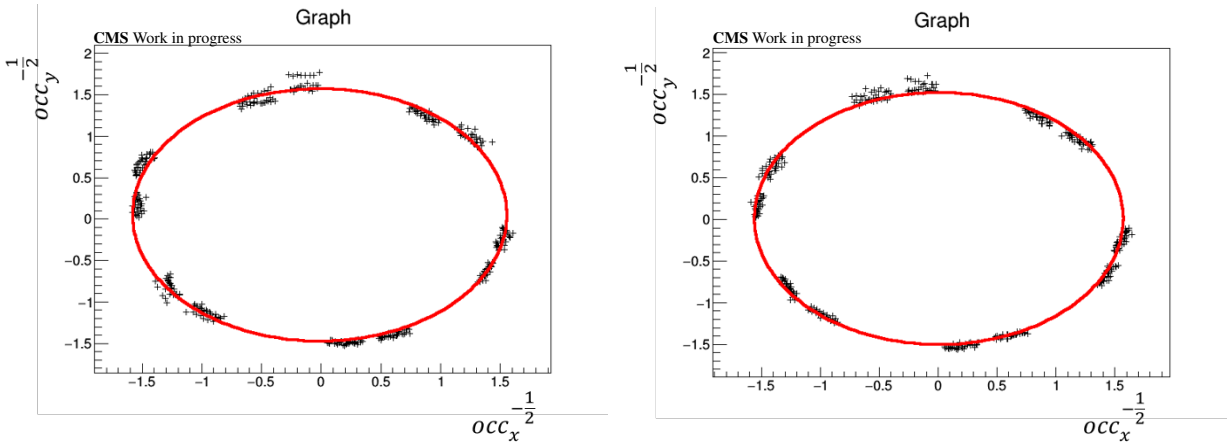


Figure A.14: Distributions of $1/\sqrt{occ}$ for module sections -2 and -3 added together (left), and 2 and 3 added together (right)

Module Section	$x_0\ (1/\sqrt{N})$	$y_0\ (1/\sqrt{N})$
-3	-0.036 ± 0.004	0.022 ± 0.003
-2	0.005 ± 0.025	0.080 ± 0.008
2	-0.0173 ± 0.0017	0.048 ± 0.002
2	0.053 ± 0.006	-0.033 ± 0.004

Table A.2: Values of the x_0 and y_0 parameters from the ellipse fit in the labeled module sections. The units shown are $1/\sqrt{N}$, as the results are only proportional to units of length, where they would need to be multiplied by some unknown factor to get centimeters.

A.3.4 Ratio Method

The ratio method was a simplified fitting method that started with the assumed $1/r_{xy}^2$ dependence of the occupancy. This method, instead of strictly being a fitting method, was a system of equations that gave the approximate location of a point source based on the difference in flux seen on opposite sides of a set of arbitrary axes.

Starting with the $1/r_{xy}^2$ assumption, the occupancies at opposite sides of an arbitrary axis, Occ^+ and Occ^-, are proportional to $1/r_{xy}^2$, following Equation A.3. Taking the ratio of the two relations will give Equation A.4, which relates the change in occupancy between the two sides with a change in the radius from a predetermined point. Equation A.5 is then a quadratic form of Equation A.4, assuming constant r. Solving for Δr then gives the central point between the two sides, based on their occupancy values, where Δr would be in the reference frame of the chosen axis. For the case of the occupancies of layer one, this corresponds to the displacement of the beam spot from the origin.

$$\begin{aligned} Occ^+ &\propto \frac{1}{(r-\Delta r)^2} \\ Occ^- &\propto \frac{1}{(r+\Delta r)^2} \end{aligned} \tag{A.3}$$

$$\frac{Occ^+}{Occ^-} = \frac{(r+\Delta r)^2}{(r-\Delta r)^2} \tag{A.4}$$

$$\left(\frac{Occ^+}{Occ^-} - 1\right)\Delta r^2 - 2r\left(\frac{Occ^+}{Occ^-} + 1\right)\Delta r + r^2\left(\frac{Occ^+}{Occ^-} - 1\right) = 0 \tag{A.5}$$

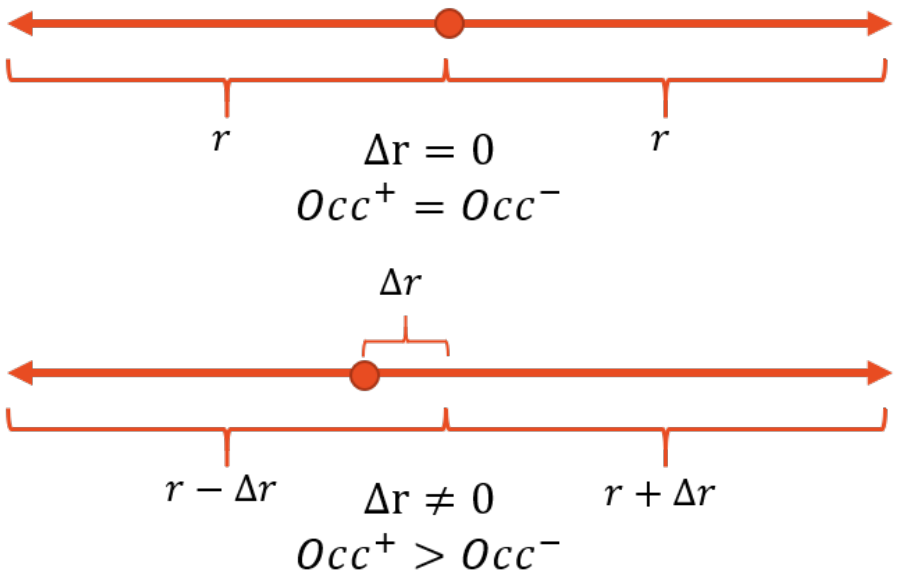


Figure A.15: A simple example of the displacement of a point source based on the occupancy on either side of a predetermined axis. The top picture shows what is expected for a point source between two equivalent occupancies, where the point source is found at the origin of this axis. The bottom picture shows the displacement of the point source from this origin based on two occupancies, with Occ^+ having a larger magnitude than Occ^-.

To better visualize this, Figure A.15 gives a simple example of the location of a point source depending on some occupancy values, and their distance from a predefined central point. When the occupancies on either side of the central point are equal, the location of the point source will be at the central point between the two occupancies, shown in the top part of Figure A.15. When these occupancies are not equal, then the point source will be displaced from the central point by some Δr in the direction of the larger occupancy, shown on the bottom. With the $1/r_{xy}^2$ assumption, Equation A.5 can then be used to find Δr in this example, with Δr being interchangeable with x_0 and y_0 for the case of the beam spot measurement.

For this fitting method, the occupancies in the $\pm x$ and $\pm y$ directions were used with Equation A.5 to determine a beam spot displacement in the x and y directions, respectively. Since the Cartesian coordinate system was used, the data preparation was kept simple by counting the occupancy in each of the four Cartesian quadrants, $+x/+y$, $-x/+y$, $-x/-y$, and $+x/-y$, through the use of different ϕ ranges, with an example given in Figure A.16. The positive and negative directions for the x and y occupancies were then calculated by adding the two relevant quadrants. So for ex-

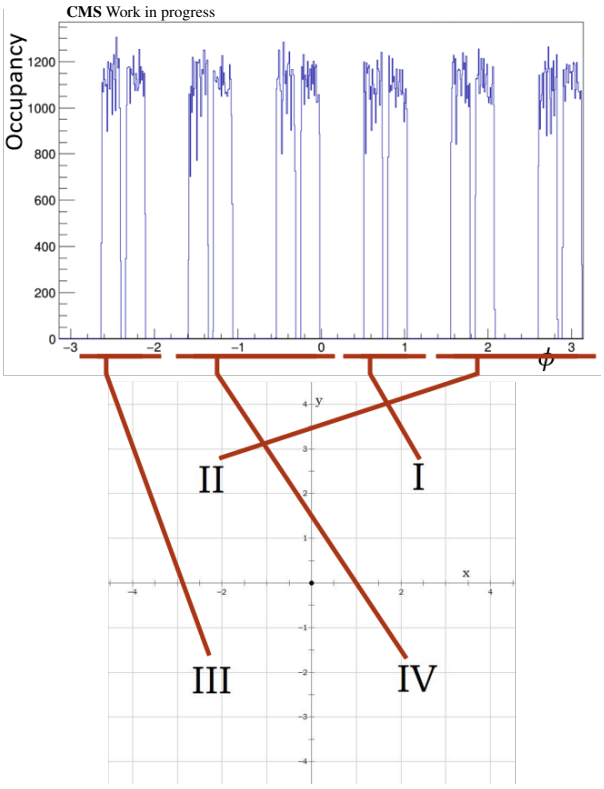


Figure A.16: An visual example of how the occupancy numbers were calculated for each of the Cartesian quadrants, showing which ϕ ranges contributed to each of the quadrants.

ample, the $+x$ component added together the occupancies from the $+x+y$ and $+x/-y$ quadrants. Since all of the measured occupancies were at the same average radius of 2.77 cm, this was seen as a feasible approach that allowed the condensing of occupancies to just four numbers that can be added together to get the necessary $+x$, $-x$, $+y$, and $-y$ occupancies, greatly reducing the complexity of having to bin the occupancy over ϕ. To allow for the measurement of multiple beam spot values for comparisons, and to see how the results change over the z-direction, the measurements of the occupancy were split over the module rings, following Figure A.1, so that each of the eight module rings had a pair of occupancies for the x and y axes.

The first dataset that was tested with the ratio method was a centrally produced design MC with a beam spot at $(0,0,0)$, that had z-smearing of the interaction point applied. The ratio method was tested over 3000 events using pixel clusters of two different sizes: $= 1$ and up to 3 pixel clusters. The sizes refer to how many pixels are included in the cluster. The point of the cluster size cut was to try to get clusters that are from particles with normal incidence to the pixel detectors, which

Cluster Size	x_0 (cm)	σ_x	y_0 (cm)	σ_y
Inner Ladders				
1	0.006 ± 0.039	0.02	0.002 ± 0.039	0.017
≤ 3	-0.002 ± 0.015	0.015	-0.002 ± 0.015	0.008
Outer Ladders				
1	0.015 ± 0.057	0.045	-0.012 ± 0.057	0.029
≤ 3	0.008 ± 0.022	0.014	-0.003 ± 0.022	0.014

Table A.3: The mean values of the calculated beam spot averaged over the different module rings for the two cluster size selections. The errors on the values are statistical only, with the standard deviation of the mean given in a separate column.

was an attempt to remove some of the z-dependence that would be expected when looking at the module rings at larger z-values. It was later decided that 1 pixel clusters would not be used for later measurements, as they were not a large population of the total number of clusters, and that 1 pixel clusters were potentially problematic in that they did not necessarily originate from beam collisions. For this fitting method, both the inner and outer ladders of layer 1 were used, but kept separate, so as to gauge the outer ladders' sensitivity to the beam spot location. The plot in Figure A.17 shows the distribution of measured x_0 and y_0 values for each of the module rings for the 1 pixel clusters, while Table A.3 gives the value averaged over each of the module sections for the two sets of pixel sizes. From the table, the values from the ratio method are consistent with the beam spot value of $(0,0)$ cm, with an expected decrease in the statistical error for the ≤ 3 pixel clusters due to the increase in the number of clusters. One thing to notice is that the outer ladders show both a larger spread in the measured values, and a larger statistical error, denoting problems in both sensitivity, and number of clusters.

After the check with a beam spot centered at $(0,0)$, further measurements were made using datasets with more events and displaced beam spots, while still keeping design detector conditions, and the z-smearing. The number of events is increased, to on the order of $10,000$ for the following three measurements, and $100,000$ after that, so as to reduce poor statistics at the module rings further away from the center of the detector. The next set of measurements were done using three different beam spots, $(0,0)$, using more events than the previous measurement, $(0.01,0.04)$, which is the approximate value of the true beam spot in real data, centered on BPix, and $(0.1,-0.08)$,

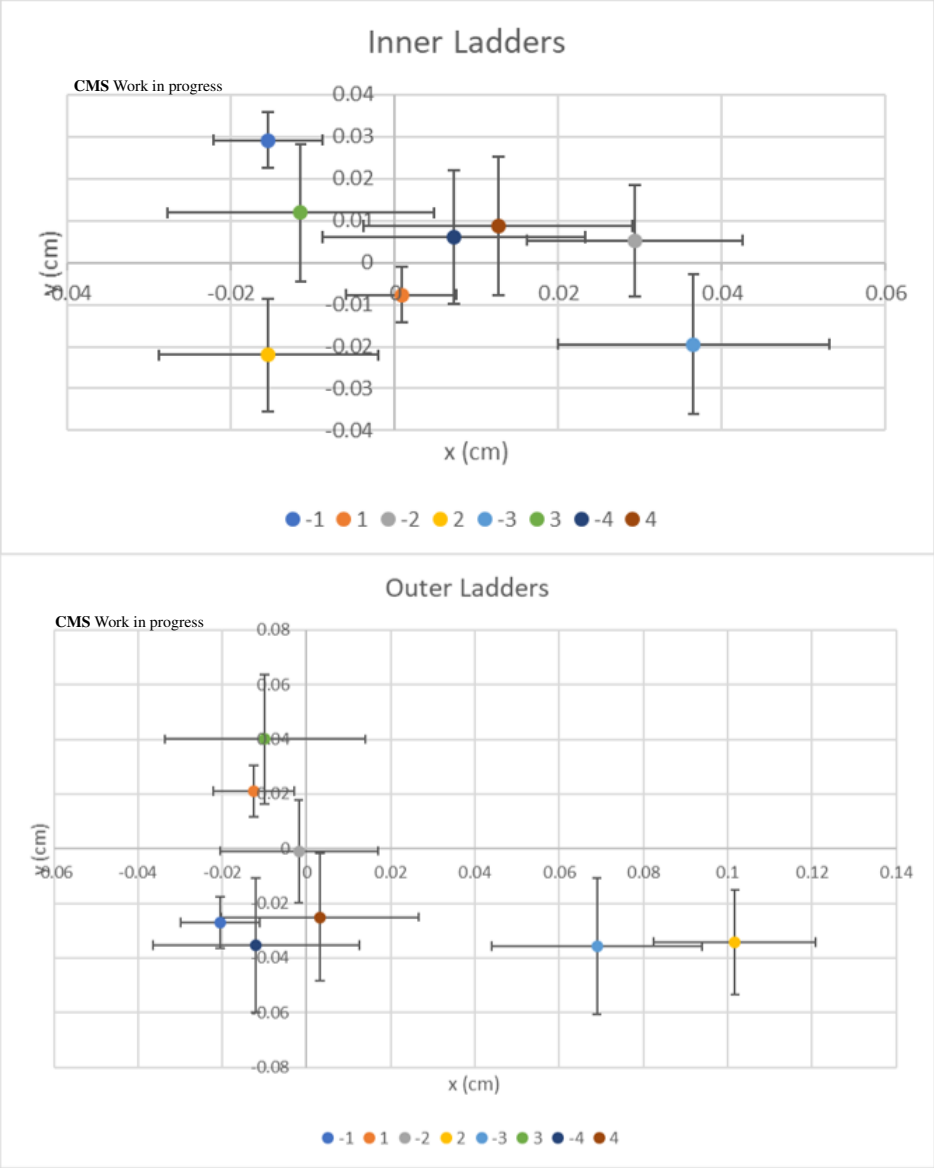


Figure A.17: Distribution of measurements of the beam spot using the ratio method, for a true value of $(0,0)$ using design MC. Each point represents the value calculated using the clusters from a single module ring. Results from the inner ladders are on top, and those from the outer ladders are on the bottom.

which is close to the value of the true beam spot in CMS coordinates, since BPix is not perfectly centered in CMS coordinates. The requirement for 1 pixel clusters was used for the next set of measurements, but for the set of 5 after that, the cut was instead changed to ≥ 2, since it was determined that focusing solely on 1 pixel clusters was problematic, as a majority of the clusters occur with a size at or greater than 2 pixels.

Figure A.18 gives the distribution of the measured beam spot values for each of the module sections, and Table A.4 gives the mean value of the data, including both inner and outer ladders. The mean values are given in three different configurations: with all of the module rings included, with the two inner-most module rings removed, and with the four inner-most module rings removed. These different configurations are used, following what is seen on the distribution corresponding to the $(0.1, -0.08)$ beam spot in Figure A.18. One may notice that there are some outliers sitting on the $+y$ axis, away from the other points. These outliers correspond to module rings -2, -1, 1, and 2. The most that can be said of the results for the three beam spots is that they are sensitive to the location, especially for the $(0.1, -0.08)$ beam spot. One thing to note is that the results seem to improve when removing the innermost ladders from the averaging. One possible explanation is that the clusters that are contributing to the beam spot sensitivity may be predominantly coming from the beam halo. The fact that the inner module rings show such a large discrepancy from the true beam spot is the first indication that a majority of the interactions from the particle collisions do not follow the $1/r_{xy}^2$ dependence as it has been assumed with the ratio method. This leads to the conclusion that the beam halo, which is drowned out in the inner module rings, is what is sensitive to the beam spot when using the ratio method. The ratio method, following this conclusion, is tested again for further displaced beam spots using on the order of $100,000$ events.

The last set of samples used to test the ratio method consisted of a total of 5 different beam spots, with z-smearing applied to the samples. The beam spots had values of $(0.1, -0.08)$, $(0.2, 0.19)$, $(0.3, -0.32)$ $(0.5, -0.48)$, and $(1.0, -0.95)$ cm. For this set of samples, it was apparent that the $1/r_{xy}^2$ assumption being used was sensitive to the beam spot location, but was not entirely correct, so an additional $1/r_{xy}$ assumption was tested, alongside the $1/r_{xy}^2$ assumption, the reason for which

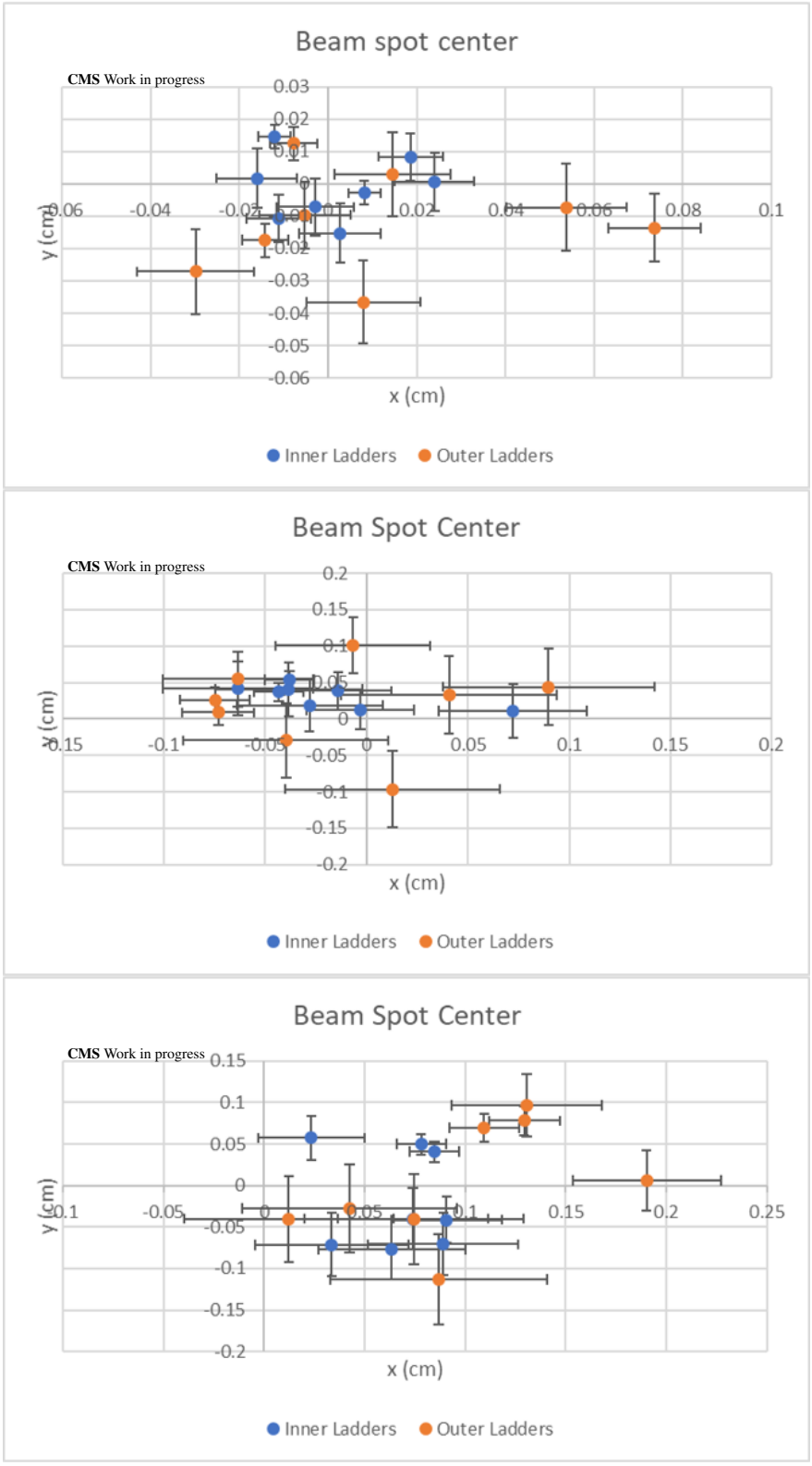


Figure A.18: Distribution of measurements using the ratio method, for a $(0,0)$ (top), $(0.01,0.04)$ (middle), and $(0.1,-0.08)$ (bottom) beam spot using $10,000$ events each. Each plot shows the measurements for the inner ladders in blue, and the outer ladders in orange.

Beam Spot	x_0 (cm)	$\frac{x_{true}-x_0}{\sigma_x}$	y_0 (cm)	$\frac{x_{true}-x_0}{\sigma_x}$
All Module Rings				
$(0,0)$	0.007 ± 0.027	-0.24	-0.007 ± 0.014	0.49
$(0.01, 0.04)$	-0.017 ± 0.05	0.54	0.025 ± 0.043	0.36
$(0.1, -0.08)$	0.082 ± 0.045	0.4	-0.008 ± 0.065	-1.12
Rings ± 1 and ± 2 removed				
$(0,0)$	0.007 ± 0.025	-0.27	-0.011 ± 0.014	0.77
$(0.01, 0.04)$	0.006 ± 0.057	0.07	0.008 ± 0.049	0.66
$(0.1, -0.08)$	0.06 ± 0.039	1.48	-0.06 ± 0.028	-0.71

Table A.4: Mean of the measured values of the beam spot using the Ratio method for a $(0,0)$, $(0.01, 0.04)$, and $(0.1, -0.08)$ beam spot using $10,000$ events each. The measurements from both the inner and outer ladders are averaged together. The mean values are given for different sets of points, corresponding to all of the module rings together, and removing module rings ± 2, and ± 1.

will be explained. The equation used for the $1/r_{xy}$ assumption uses the exact same derivation as Equations A.4 and A.5, only removing the square in Equation A.4. For the results from these samples, the cut on the size of the pixel clusters is required to be ≥ 2.

The results per module ring for these samples are given in Figures A.19 - A.23, with averages over the module rings given in Table A.5 for the $1/r_{xy}^2$ assumption and Table A.6 for the $1/r_{xy}$ assumption. All of these results only look at the inner ladders of layer 1, removing the outer ladders based on the results from Figure A.18, showing that the outer ladders in general have less sensitivity to beam spot displacements. To summarize the results, they all point towards the ratio method being sensitive to, but insufficient in describing how the occupancy changes with displaced beam spots. The method seems to work for the module rings that are further away from the center in z, as shown in Figure A.24, but there are still oddities in the results that do not engender confidence in the method.

For each of the samples, the inner module rings show that they are not as sensitive to the beam spot location as the outer module rings, which becomes more apparent as the beam spots approach 1 cm in displacement. The Figures A.19 - A.23 highlight which points belong to the inner module rings (-2, -1, 1, and 2). There are two other features that also become more apparent at larger beam spot displacements. The first of these is that, with the $1/r_{xy}^2$ assumption, the calculated values

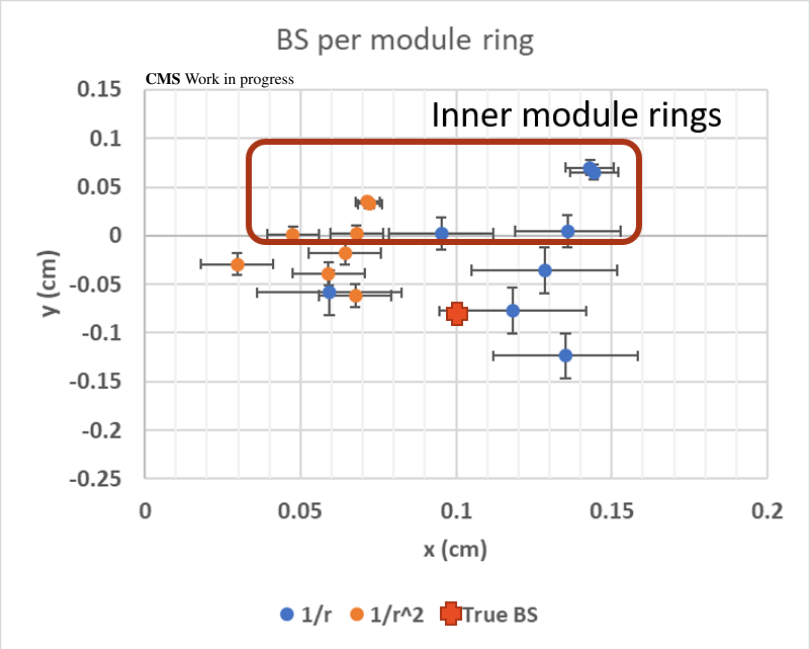


Figure A.19: Distribution of measurements using the ratio method, for a $(0.1, -0.08)$ beam spot using $100,000$ events. The orange markers assume a $1/r_{xy}^2$ dependence, and the blue markers assume a $1/r_{xy}$ dependence. The modules highlighted in the rectangle correspond to module rings ± 2 and ± 1.

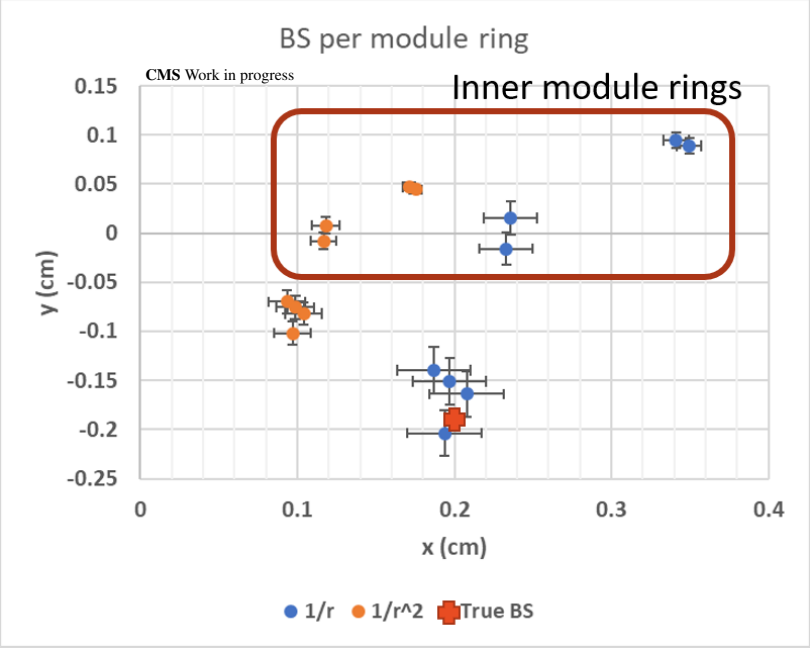


Figure A.20: Distribution of measurements using the ratio method, for a $(0.2, -0.19)$ beam spot using $100,000$ events. The orange markers assume a $1/r_{xy}^2$ dependence, and the blue markers assume a $1/r_{xy}$ dependence. The modules highlighted in the rectangle correspond to module rings ± 2 and ± 1.

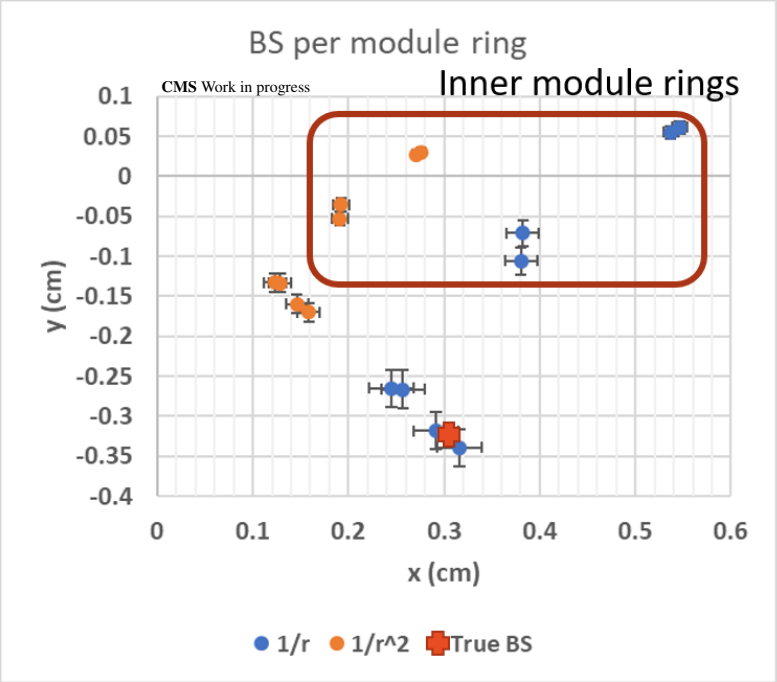


Figure A.21: Distribution of measurements using the ratio method, for a $(0.3, -0.32)$ beam spot using $100,000$ events. The orange markers assume a $1/r_{xy}^2$ dependence, and the blue markers assume a $1/r_{xy}$ dependence. The modules highlighted in the rectangle correspond to module rings ± 2 and ± 1.

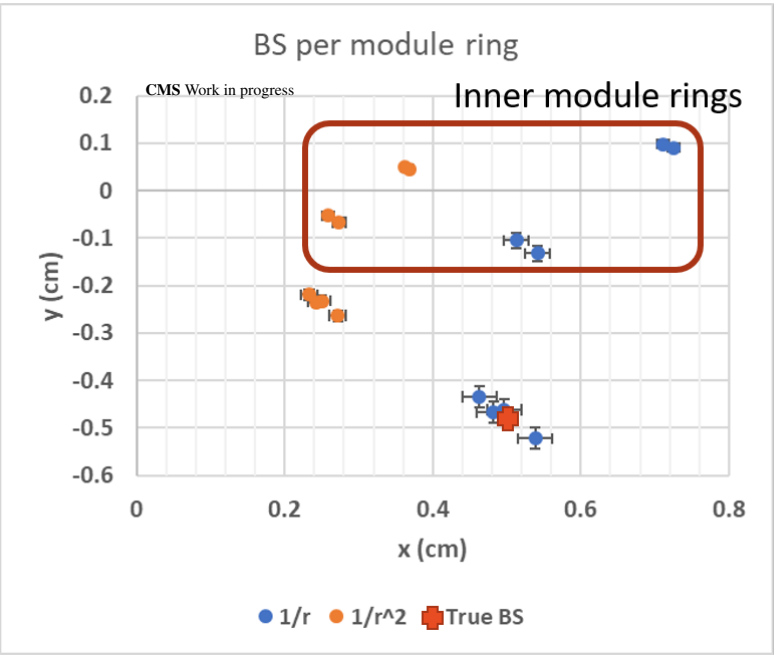


Figure A.22: Distribution of measurements using the ratio method, for a $(0.5, -0.49)$ beam spot using $100,000$ events. The orange markers assume a $1/r_{xy}^2$ dependence, and the blue markers assume a $1/r_{xy}$ dependence. The modules highlighted in the rectangle correspond to module rings ± 2 and ± 1.

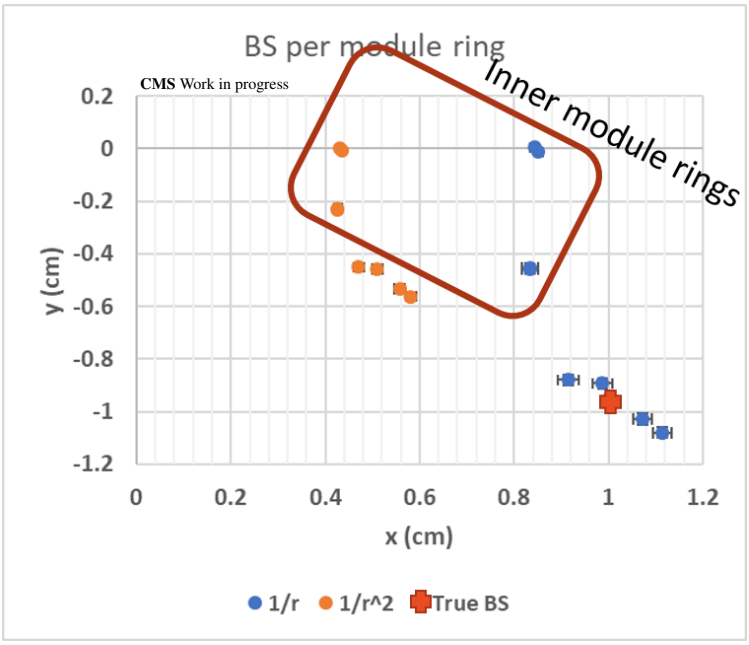


Figure A.23: Distribution of measurements using the ratio method, for a $(1.0, -0.95)$ beam spot using $100,000$ events. The orange markers assume a $1/r_{xy}^2$ dependence, and the blue markers assume a $1/r_{xy}$ dependence. The modules highlighted in the rectangle correspond to module rings ± 2 and ± 1.

Beam Spot	x_0 (cm)	y_0 (cm)
All module rings		
$(0.1, -0.08)$	0.06 ± 0.03	-0.01 ± 0.043
$(0.2, -0.19)$	0.122 ± 0.043	-0.03 ± 0.066
$(0.3, -0.32)$	0.186 ± 0.065	-0.078 ± 0.086
$(0.5, -0.48)$	0.283 ± 0.059	-0.122 ± 0.133
$(1.0, -0.95)$	0.48 ± 0.067	-0.308 ± 0.228
Rings ± 1 and ± 2 removed		
$(0.1, -0.08)$	0.055 ± 0.029	-0.037 ± 0.03
$(0.2, -0.19)$	0.098 ± 0.024	-0.082 ± 0.027
$(0.3, -0.32)$	0.139 ± 0.028	-0.149 ± 0.03
$(0.5, -0.48)$	0.249 ± 0.028	-0.237 ± 0.03
$(1.0, -0.95)$	0.53 ± 0.054	-0.501 ± 0.059

Table A.5: Mean of the measured values of the beam spot using the Ratio method with a $1/r_{xy}^2$ assumption for $(0.1, -0.18)$, $(0.2, -0.19)$, $(0.3, -0.32)$, $(0.5, -0.48)$ and $(1.0, -0.95)$ cm beam spots using $100,000$ events each. The mean values are given for a couple different sets of points, corresponding to all of the module rings together, and removing the inner module rings ± 2, and ± 1. The errors are a combination of the statistical error, and the standard deviation of the values from each of the module rings.

Beam Spot	x_0 (cm)	y_0 (cm)
All module rings		
$(0.1, -0.08)$	0.12 ± 0.061	-0.019 ± 0.086
$(0.2, -0.19)$	0.243 ± 0.085	-0.059 ± 0.131
$(0.3, -0.32)$	0.369 ± 0.129	-0.156 ± 0.171
$(0.5, -0.48)$	0.559 ± 0.115	-0.242 ± 0.265
$(1.0, -0.95)$	0.931 ± 0.124	-0.6 ± 0.439
Rings ± 1 and ± 2 removed		
$(0.1, -0.08)$	0.11 ± 0.058	-0.074 ± 0.06
$(0.2, -0.19)$	0.196 ± 0.048	-0.164 ± 0.055
$(0.3, -0.32)$	0.277 ± 0.057	-0.297 ± 0.06
$(0.5, -0.48)$	0.495 ± 0.056	-0.471 ± 0.058
$(1.0, -0.95)$	1.022 ± 0.098	-0.97 ± 0.108

Table A.6: Mean of the measured values of the beam spot using the Ratio method with a $1/r_{xy}$ assumption for $(0.1, -0.18)$, $(0.2, -0.19)$, $(0.3, -0.32)$, $(0.5, -0.48)$ and $(1.0, -0.95)$ cm beam spots using $100,000$ events each. The mean values are given for a couple different sets of points, corresponding to all of the module rings together, and removing the inner module rings ± 2, and ± 1. The errors are a combination of the statistical error, and the standard deviation of the values from each of the module rings.

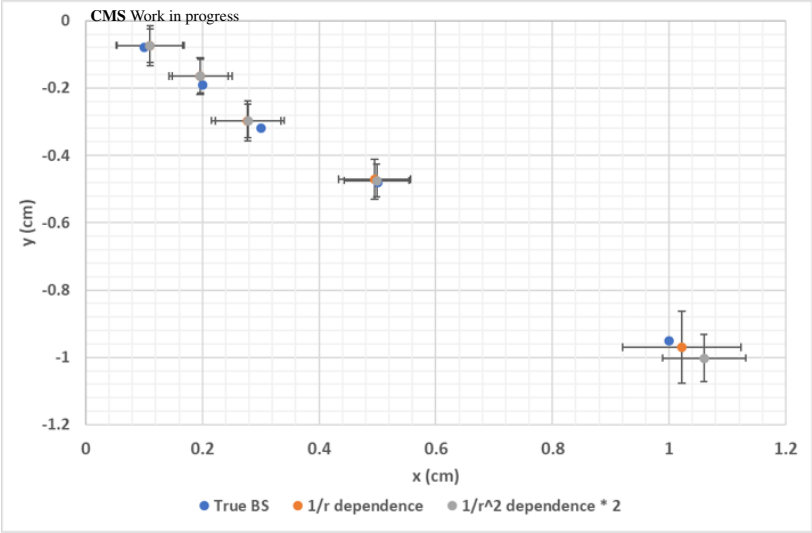


Figure A.24: Averaged beam spot measurements for each of the $(0.1, -0.18)$, $(0.2, -0.19)$, $(0.3, -0.32)$, $(0.5, -0.48)$ and $(1.0, -0.95)$ beam spots, using the $1/r_{xy}$ and $1/r_{xy}^2$ assumptions, and compared to the true beam spot value. The values from the $1/r_{xy}^2$ assumption are scaled up by a factor of 2.

of the beam spot are a factor of 2 smaller than the true value of the beam spot, while with the $1/r_{xy}$ assumption, the calculated values are on par with the true beam spot values. This was also checked with an assumption that the occupancy scales as $1/r_{xy}^{1/2}$, which gave beam spot values that are a factor of 2 larger than the true value. The last odd feature is that at higher displaced beam spots apart from the two pairs of inner module rings, the results from the outer module rings are arranged in a linear pattern, with the true beam spot falling in between module rings ± 3 and ± 4, a feature which is as of yet unexplained.

With the results from the ratio method showing promise, but ultimately being untrustworthy, it was decided to move to a fit that could determine the best occupancy dependence numerically, without using preconceived assumptions on how the occupancy should scale over distance. This newer fit, following from the ratio method that a $1/r_{xy}$ dependence is a step in the right direction, would still use an inverse power relation, but would instead allow the power to float as its basis. In addition, it would make use of the full 3D radius of the clusters in the detector, rather than just the cylindrical radius like in the previous fitting methods. Also, following the other two fit methods, it was important to account for any variance over ϕ, as this would affect many of the floating parameters in a fit, and the importance of working in Cartesian coordinates for the ease in separating out the beam spot values.

A.3.5 Maximum Likelihood Fit

The final method used to study the location of the beam spot made use of a maximum likelihood fit that could be used to find the x, y, and z position of the average location of the interaction point. The driving force behind the use of this method is that it includes the 3D radius of the clusters, so that it now makes full use of the z-position, and from this is closer to modeling the true radius of the clusters from the interaction point, which better accounts for how the occupancy changes with distance.

The first step in creating this likelihood fit was putting together the likelihood function, and testing it. Over the course of using this fitting method there have been a couple of iterations,

which introduced improvements to the overall fitting method, where each change will be covered sequentially.

The data for each iteration of the fit was prepared by counting the number of clusters per ROC in the inner ladders of layer 1. The position of each ROC was variable, in that the ultimate position was determined by a weighted average of the position of the pixels in the ROC. For each cluster there was a position assigned, corresponding to the central pixel of the cluster. The occupancy was incremented for every hit on a specific pixel, and once all the clusters were counted, a weighted average was calculated for the ROC using the position of each pixel and its corresponding occupancy. Essentially what was made was a ROC occupancy map that also stored the 3D position weighted by the occupancy. Initially the selected clusters had a pixel size cut placed on them of ≥ 2, but later iterations removed all of the size cuts, this decision will be detailed in the following section. All of the likelihood fits were done using the MINUIT minimization program.

A.3.5.1 First iteration: a power law

The first iteration of the likelihood fit used a simple inverse power law, given in Equation A.6, as the base of the likelihood function. In this function, the r variable is the 3D radius, r_{xyz}, and a, b, and c are floating parameters in the fit. The data were separated by occupancy per ROC, so that each ROC will have a x, y, and z position, and an occupancy. The full likelihood function is then given in Equation A.7, using a Normal distribution, where N is the number of ROCs, Φ_i is the occupancy at the ith ROC, σ_i is $\sqrt{\Phi_i}$, n is a normalization scaling parameter, and r_i is written in its Cartesian coordinates, as given in Equation A.8. The floating parameters in this function are then n, x_0, y_0, and z_0, of which x_0, y_0, and z_0 would be the measured location of the beam spot in x and y, and the average location of the interaction point in z. For this iteration the parameters a, b, and c from Equation A.6 were not made to float, but instead pre-determined from a separate fit of the occupancy, which is the reason for the n parameter, as it is used to rescale the normalization of the pre-determined a, and c parameters. One caveat to this likelihood fit function is that it will not be able to account for in z-smearing of the interaction point, so this particular function was only

tested on non z-smeared samples, where the interaction point was kept at a constant z value.

$$\Phi(r) = a/r^b + c \tag{A.6}$$

$$\mathscr{L} = \Pi_i^N \exp \frac{(n\Phi(r_i) - \Phi_i)^2}{2\sigma_{\Phi_i}} \tag{A.7}$$

$$r_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2} \tag{A.8}$$

The function was tested using a couple different non z-smeared design MC samples. The beam spots had positions of $(0,0,0)$, $(0,0,10)$, and $(0.1,-0.08,0)$. The only difference between these sets of samples was the beam spot location, all other parameters were kept the same, such as number of events, set at $100,000$ events, and the use of the design detector. Before going through the results of fitting over these samples, the cleaning applied to the data will be described.

For the distribution of ROCs in ϕ and z there was extra cleaning applied, beyond what was already done when processing the clusters to make the occupancy map. This cleaning simply removed ROCs which were seen as outliers when compared to their surrounding neighbors. Figure A.25 highlights the ROCs that were removed from the data for the first iteration. While the figure is integrated over ϕ, the same features are present when unfolded into single ROCs in each ladder. For the outlying ROCs which are labeled as coming from the edge of modules, this is believed to be some artifact related to simulation, as it has not been observed in the real data. For the decreased occupancy at central z values, this was found to be related to the cluster size cut removing those clusters, but for this fit the offending ROCs were simply removed, with the cut instead being removed in the following iterations.

Before running the actual likelihood fit, the inverse power law, Equation A.6, was tested over a distribution of r_{xyz} versus occupancy, with the result given in Figure A.26. The fit result shows that the previous assumptions of the occupancy scaling as $1/r_{xy}$ or $1/r_{xy}^2$ was not correct, and denoting,

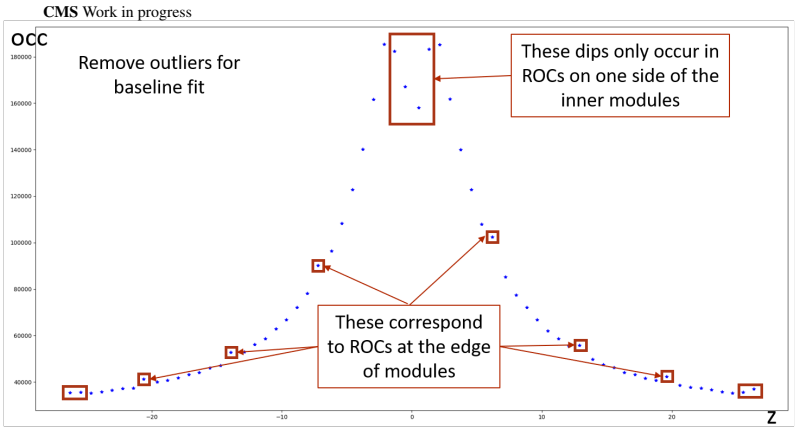


Figure A.25: Figure of the occupancy integrated over the ROC rings, versus z-position for a non z-smeared $(0,0,0)$ beam spot. Highlighted are the outlier ROC sections which are removed for cleaning purposes.

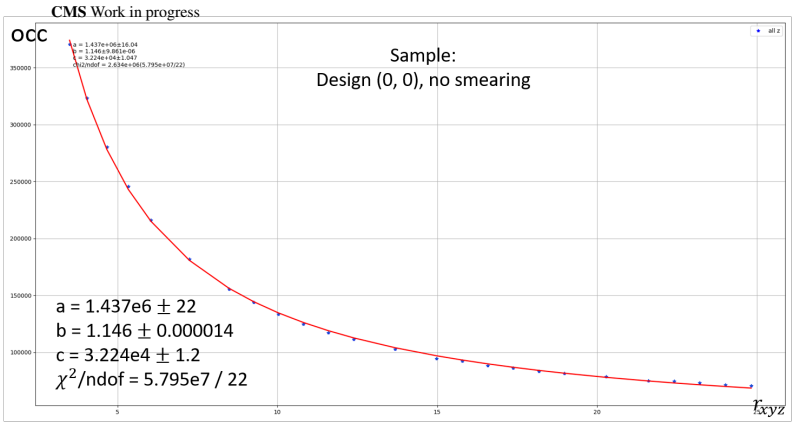


Figure A.26: Figure of the occupancy integrated over the ROC rings, versus r_{xyz} for a non z-smeared $(0,0,0)$ beam spot. Overlayed is a fit of Equation A.6.

instead, that it is closer to a $1/r_{xyz}^{1.15}$ dependence, which would be non-trivial to derive analytically with the ratio method. This fit result also shows that the power law, otherwise, is, at least visually, a good fit for describing how the occupancy is changing with increasing radius, though, only when considering non z-smeared samples. The values of the parameters given in this fit, and shown in Figure A.26, are used as the a, b, and c parameters in for the power function in the likelihood fit.

When running the maximum likelihood fit, the negative log of Equation A.7 was calculated, and then minimized, floating the variables n, x_0, y_0, and z_0. The initial parameters were set to 1 for n, and $(0,0,0)$ for (x_0, y_0, z_0). The likelihood fit was done for the three previously mentioned beam spots, with a distribution of the data given to the fit shown in Figure A.27 for each of the beam

Beam Spot	x_0 (cm)	y_0 (cm)	z_0 (cm)	n
$(0,0,0)$	-0.001 ± 0.018	0.004 ± 0.018	-0.011 ± 0.011	0.0416 ± 0.0001
$(0,0,10)$	-0.0002 ± 0.0018	0.002 ± 0.018	10.063 ± 0.007	0.0416 ± 0.0001
$(0.1,-0.08,0)$	0.308 ± 0.018	-0.244 ± 0.018	0.00 ± 0.01	0.0416 ± 0.0001

Table A.7: Table of the fitted values using the first iteration of the maximum likelihood fit. The parameter errors are calculated by the HESSE method.

spots. The results of the fits are given in Table A.7. Looking to the figures, visually it is clear when there is a displaced beam spot, as the occupancy shows a sinusoidal shape over ϕ, while for the centered beam spot, the occupancy is shown to be flat. The results of the likelihood fits show good results for the first two beam spots at $(0,0,0)$ and $(0,0,10)$, where it finds values fully consistent with the true values, including the z-values. The fit with a displaced beam spot at $(0.1,-0.08,0)$, on the other hand, does not give as accurate results. The signs on the beam spot are correct, but the magnitude is too large by a factor of 3.

The results from this first iteration show that the likelihood fit is on the right track, but needs adjustment. This leads to the second iteration of the likelihood fit, which allows the parameters from Equation A.6 to vary in the fit, and in addition introduces a ϕ modulation of those parameters.

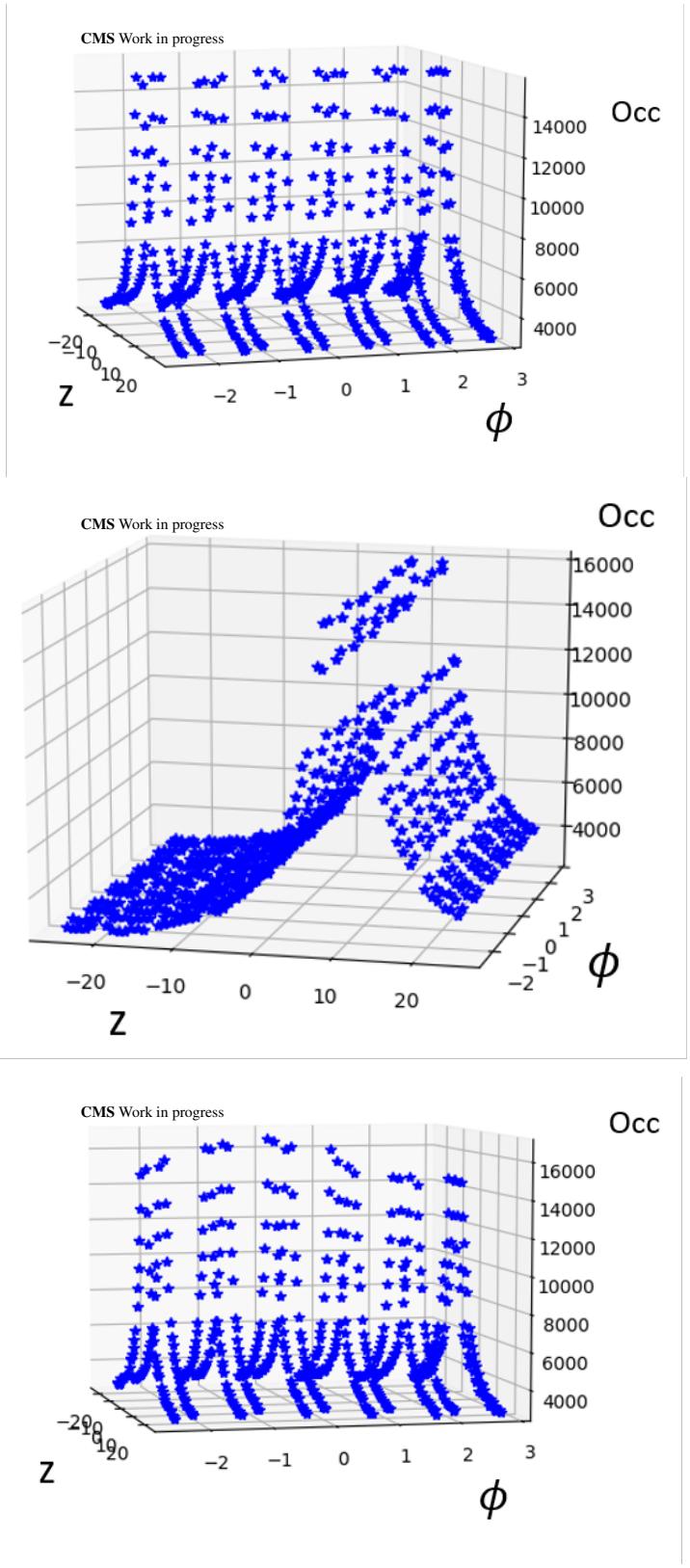


Figure A.27: Distributions of the occupancy per roc in ϕ and z positions for the $(0,0,0)$ (top), $(0,0,10)$, (middle) and $(0.1,-0.08,0)$

A.3.5.2 Second Iteration: ϕ modulation

The changes introduced in this next set of fits were done in order to fix the problems seen previously. The first problem, which is best seen in Figure A.25, was the existence of a dip in occupancy at central z, while the second problem was the large magnitude in the fit result for a displaced beam spot.

The first problem was easy to fix, as the dip in occupancy was actually caused by the cluster size cut applied when processing the clusters to create the occupancy maps. This was discovered with help of some pixel detector experts, who suggested removing the cut. The reason for the cut creating such a dip is due to a large fraction of 1-2 pixel clusters being removed with the application of the cut. This creates a large effect at central z due to those ROCs being almost directly over the interaction point, where at z values further away from the center, the incident particles are hitting the ROCs at angles which cause a greater number of pixels to register hits in a single cluster. Figure A.28 shows the effect of removing the cluster size cut.

For the second problem, there were no obvious changes that could be made, so a focus was placed on improving the fit equation to account for as many variations as possible. The first, more obvious change, was allowing the parameters of Equation A.6 to float, which lets the likelihood fit choose the best values for these parameters. Doing this additionally made the n parameter superfluous, as the fit would now control the normalization of the a, and c parameters naturally, so it was removed.

The next change made was not an apparent one, and almost needed one to stumble upon it. It required that the parameters of Equation A.6 vary over ϕ. This was determined by performing a power law fit using the occupancy distributions over r_{xyz}, and separating the data into half-ladder sections. For the 6 inner ladders of layer 1, this would correspond to 12 different power law fits. Figure A.29 gives an example of these 12 different fits. Looking at the resulting fitted parameters plotted over the ϕ position of the half-ladders, given in Figure A.30, shows a distinct sinusoidal shape for each of the a, b, and c parameters, that is dependent on the beam spot. From these results the parameters were then allowed to vary over ϕ in the fit, following Equation A.9. For each

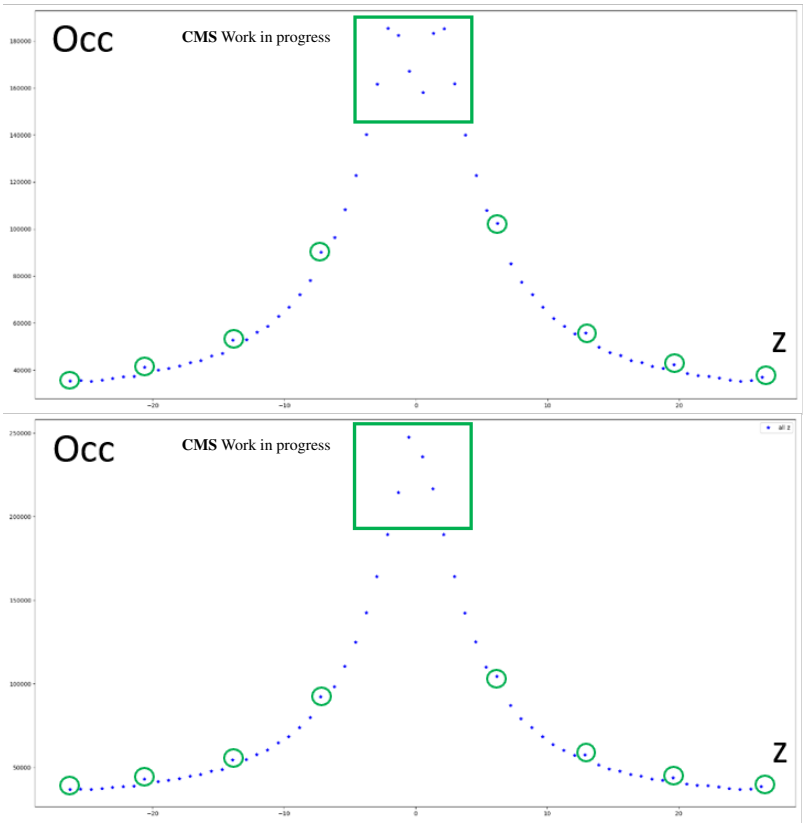


Figure A.28: Distributions of the occupancy per roc over the z position with (top) and without (bottom) a cluster size cut of ≥ 2 pixels. The squares highlight where this cut greatly effects the distributions. The circles show the ROCs that are still removed between the two distributions.

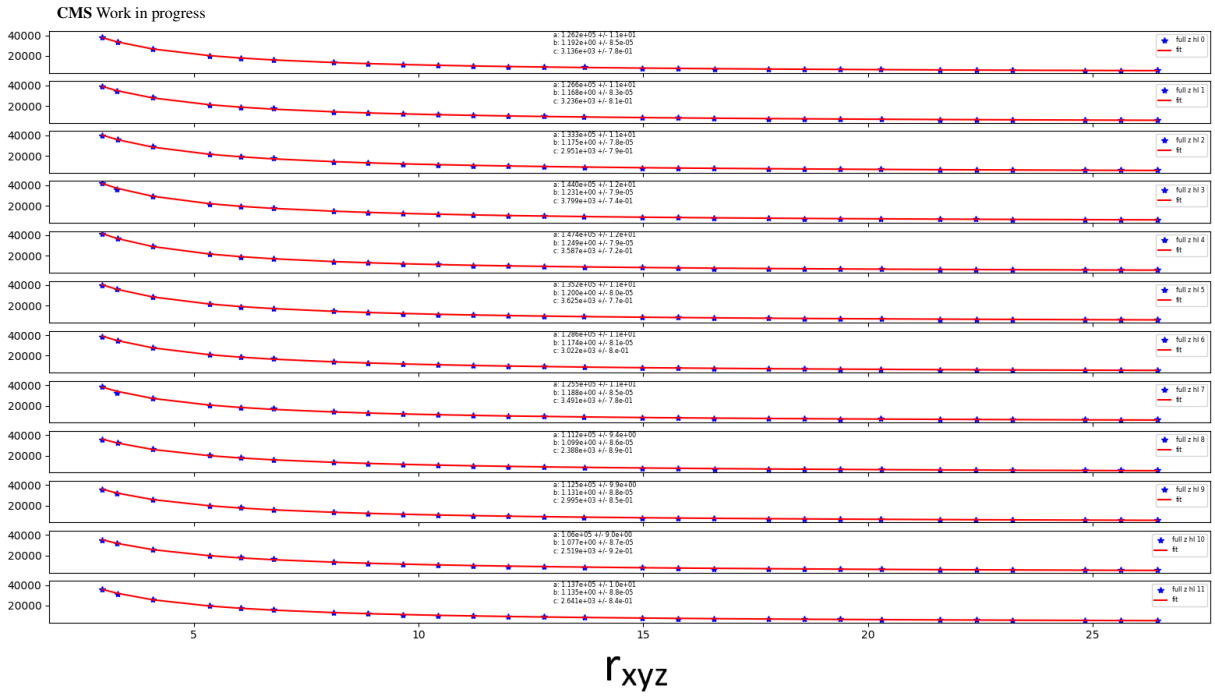


Figure A.29: Distributions of the occupancy per ROC for each of the half ladders in layer 1, using data from the $(0.1, -0.08, 0)$ beam spot. A fit is overlayed corresponding to Equation A.6. This is an example of the fits done to determine that the power law parameters vary over ϕ in a sinusoidal manner.

of the power law parameters, only the amplitude and offset is kept distinct between parameters. The phases on the other hand are kept constant between the a, b, and c parameters, as Figure A.30 shows that the phase is constant between parameters, and only changes when the beam spot location changes.

$$s_1 \sin(\phi - s_2) + s_3 \tag{A.9}$$

With these changes applied, the likelihood function becomes Equation A.10. One thing to notice is that the ϕ values used are not the measured ϕ as it is given in the occupancy maps, but instead is a corrected version. This is due to their being a small change in the angles after displacing the beam spot away from the center. This is to account for that change in angle by recalculating ϕ for where the measured beam spot is located.

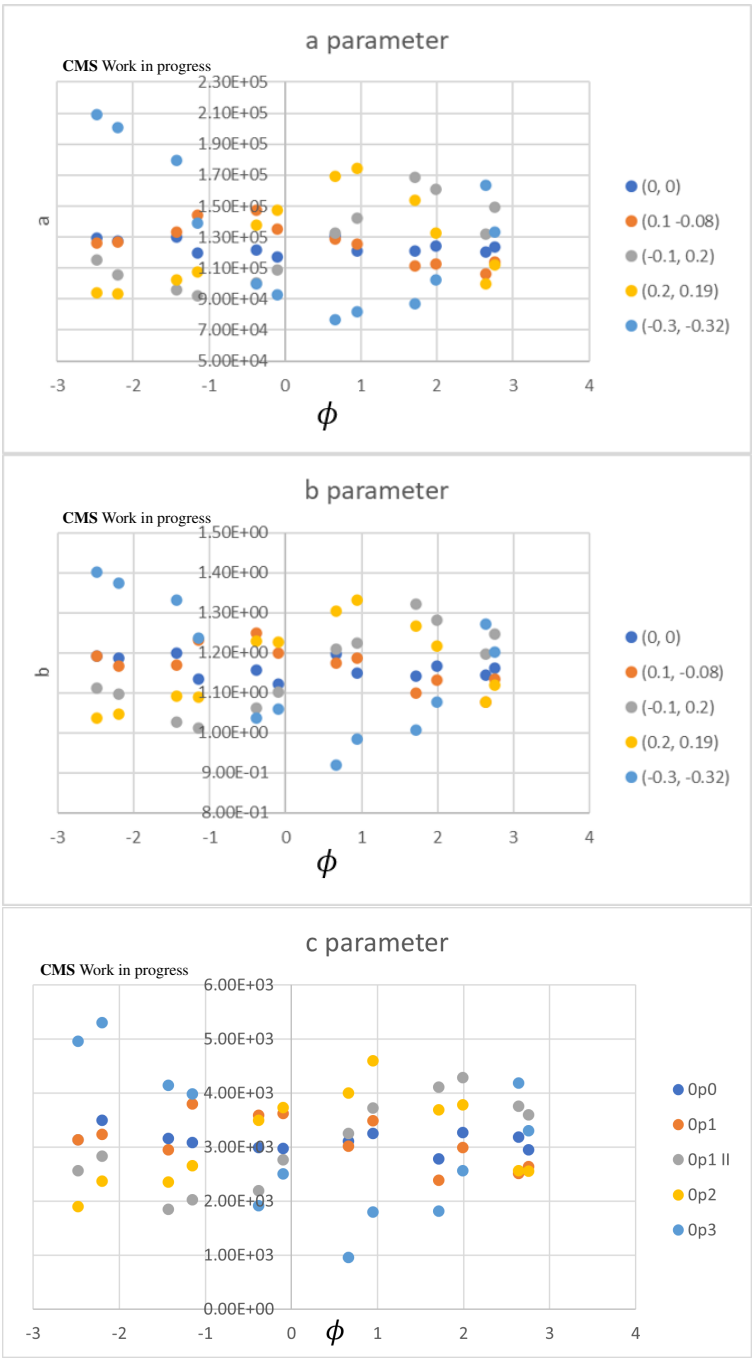


Figure A.30: Distributions of the fitted power law parameters from Equation A.6 over ϕ. Each point corresponds to the fitted parameter from one of the half-ladders in layer 1. The parameter distributions of a (top), b (middle), and c (bottom) are given for 5 different beam spots with no z-smearing applied: $(0,0)$, $(0.1-0.08)$, $(-0.1,0.2)$, $(0.2,0.19)$, and $(-0.3,-0.32)$.

$$\mathscr{L} = \Pi_i^N \exp\left(\frac{\left(\Theta(r_i) - \Theta_i\right)^2}{2\sigma_{\Theta_i}}\right)$$
$$\Theta(r_i) = \left(a_1 \sin\left(\phi_i - s_2\right) + a_3\right) / r^{b_1 \sin(\phi_i - s_2) + b_3} + \left(c_1 \sin\left(\phi_i - s_2\right) + c_3\right) \tag{A.10}$$
$$r_i = \sqrt{\left(x_i - x_0\right)^2 + \left(y_i - y_0\right)^2 + \left(z_i - z_0\right)^2}$$
$$\phi_i = \arctan\left(\left(y_i - y_0\right) / \left(x_i - x_0\right)\right)$$

This newer likelihood function was used to fit data from a wider selection of beam spots than with the first iteration. All are from design MC samples with no z-smearing applied. The beam spots are separated such that they cover each major quadrant, and at varying degrees of displacement. The inital free parameter values follow what was used in the first iteration likelihood. After obtaining full convergence with the $(0,0)$ beam spot, the $(0,0)$ beam spot was fitted again, but fixing x_0, y_0, and z_0 to their expected values. The initial parameters for the displaced beam spots was set to the fitted parameters from the fixed $(0,0)$ beam spot fit as a baseline. The results of the fits are given in Table A.8. What this table shows for the first interation likelihood is that the removal of the cluster size cut has a moderate effect in improving the fitted beam spot value, but is still close to twice as large as expected. When going to the second iteration likelihood fit, with the ϕ modulation included, the fit improves dramatically, such that in all cases, except for the most displaced beam spot at $(-0.3,-0.32)$, the fitted values are accurate to well within 1 mm of the true value of the beam spots. Table A.9 gives the results for the z-smeared samples using the second iteration. The table shows that this particular iteration is not fully capable in determining the beam spot for z-smeared sample. At low displacements it is unable to get the correct signs on the values of the beam spots, though they are accurate otherwise. Figure A.31 gives δ_x and δ_y, the difference between the true and measured values of the beam spots for the results shown in Table A.8. Figure A.32 does the same, but for Table A.9. One can see that with the exception of the $(-0.3,-0.32)$ beam spot, all of the non z-smeared beam spots are well within a difference of 1 mm. The fit does have trouble with the z-smeared samples though, but only in the sign at low beam spot displacement, otherwise the fits are surprisingly accurate.

Beam Spot	x_0 (cm)	y_0 (cm)
Using first iteration Likelihood (w/o cluster size cut)		
$(0,0)$	-0.009 ± 0.009	0.009 ± 0.009
$(0.1,-0.08)$	0.218 ± 0.009	-0.171 ± 0.009
$(-0.1,0.2)$	-0.210 ± 0.009	0.424 ± 0.009
$(0.2,0.19)$	0.429 ± 0.009	0.401 ± 0.009
$(-0.3,-0.32)$	-0.606 ± 0.009	-0.661 ± 0.009
Using second iteration Likelihood		
$(0,0)$	-0.009 ± 0.009	0.026 ± 0.071
$(0.01,0.04)$	0.008 ± 0.016	0.04 ± 0.05
$(0.03,-0.06)$	0.005 ± 0.034	-0.014 ± 0.057
$(-0.08,0.02)$	-0.09 ± 0.06	0.025 ± 0.020
$(0.1,-0.08)$	0.09 ± 0.04	-0.073 ± 0.031
$(-0.1,0.2)$	-0.092 ± 0.021	0.20 ± 0.03
$(0.2,0.19)$	0.247 ± 0.026	0.213 ± 0.026
$(-0.3,-0.32)$	-0.402 ± 0.024	-0.439 ± 0.026

Table A.8: Table of the fitted values of the beam spot using the first iteration of the maximum likelihood fit, but without the cluster size cut applied, and using the second iteration of the maximum likelihood fit, also without the cluster size cut applied. The parameter errors are calculated by the HESSE method.

Beam Spot	x_0 (cm)	y_0 (cm)
z-smeared samples		
$(0,0)$	-0.03 ± 0.05	0.05 ± 0.05
$(0.1,0.08)$	-0.06 ± 0.05	-0.11 ± 0.05
$(-0.1,0.08)$	0.11 ± 0.05	-0.06 ± 0.04
$(0.1,-0.08)$	-0.07 ± 0.05	0.04 ± 0.04
$(0.2,-0.19)$	0.12 ± 0.03	-0.14 ± 0.03
$(0.3,-0.32)$	0.244 ± 0.029	-0.297 ± 0.029

Table A.9: Table of the fitted values for the z-smeared beam spot samples using the second iteration of the maximum likelihood fit. The parameter errors are calculated by the HESSE method.

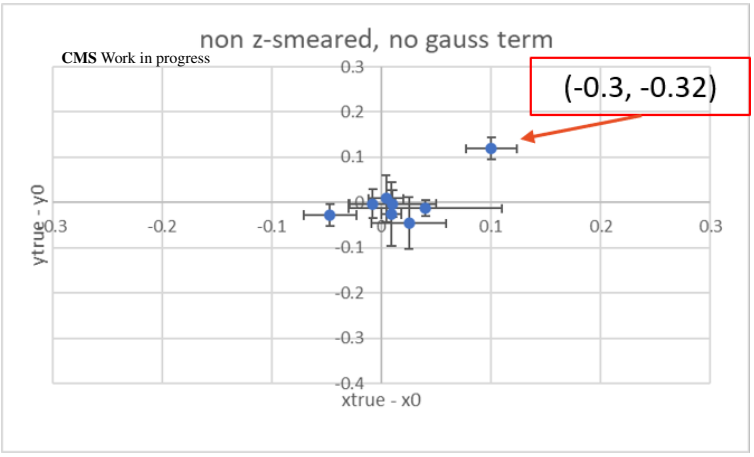


Figure A.31: Distribution of δx and δy in cm for the beam spot values in Table A.8 for the second iteration fits.

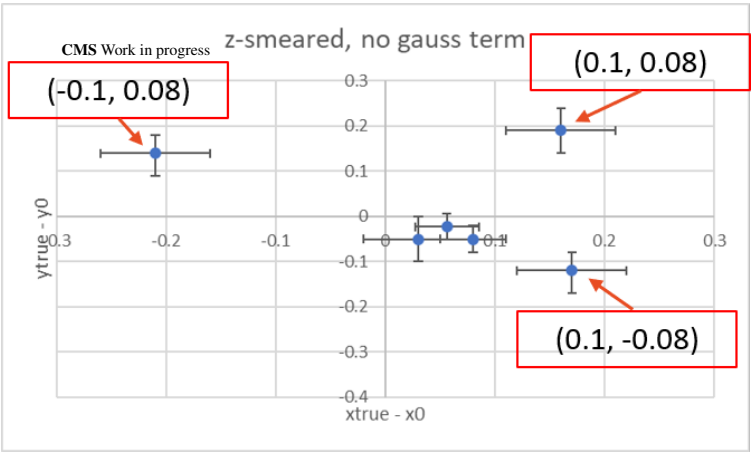


Figure A.32: Distribution of δx and δy in cm for the beam spot values in Table A.9 for the second iteration fits.

With this iteration of the likelihood fit, there was now a method that could find a value of the beam spot that is both accurate, and which one can be fairly confident in. One thing to note is that this method has only so far been tested on the ideal case of no z-smearing in the interaction point. The next step, then, is to further adjust the likelihood function so that it can account for a varying z-interaction point, and eventually be able to account for pileup in real data. This leads to the third and final iteration of the likelihood fit, which adds a Gaussian term to the base power law function, so that the fit can account for the z-variance.

A.3.5.3 Third Iteration: Gaussian term

As stated, the third, and final, iteration of involved the inclusion of a Gaussian term to account for a spread in the location of the interaction point along the z direction. The Gaussian term was chosen because the z-position of the interaction point is simulated using a Gaussian smearing of 4 cm, in order to match what is generally seen in the data. If the simulation used a Gaussian to model the position of the interaction, then it seemed prudent to also have a Gaussian term in the likelihood function to model it as well. Since the spread of the interaction point only depends on z, the Gaussian term was also made so that it only depended on z. The Gaussian term follows Equation A.11, where g_a, g_b, and g_c are the floating parameters, with $g_b = z_0$, as the μ of a Gaussian is the same as what z_0 is trying to find, the expected z position of the interaction point. Same as with the power law, $\Theta(r_i)$ in Equation A.10, the remaining Gaussian parameters are also modulated over ϕ. Also in the same way, the phase of the ϕ modulation is closely related between the power and Gaussian terms. For the case of the Gaussian term, what is found is that the phases of the floating parameters are shifted by π from the phases of the floating parameters in the power term. From this all of the phases for the ϕ modulation are given in terms of s_2, the original phase parameter.

$$G = g_a \exp^{-(z-g_b)/2g_c} \tag{A.11}$$

With the Gaussian term added, the final Likelihood function becomes Equation A.12. The fits are then performed over the same beam spots as in the second iteration, with no z-smearing, and

then samples are also included which have either z-smearing or simulated pileup. For the non z-smeared samples, the results from the third iteration can be compared to the results from the second iteration in Table A.8. Results are also calculated for a set of z-smeared samples and a couple that were simulated with pileup. The shape between the z-smeared and the pileup samples are essentially the same, as seen in Figure A.33. The main difference between the two samples, is that there is more randomness, so to speak, in the pileup sample. The z-interaction point of the event for the z-smeared samples follows a simple Gaussian smearing of its position. The simulated pileup, on the other hand, samples from a dataset of z-smeared zero-bias events and essentially inserts them into an event on top of the base simulated event. This makes it so that the pileup sample has a distribution that is not as smooth as the z-smeared samples, potentially making it more difficult to fit on.

For the fits, the initial parameter set was first determined in the same way is with the second iteration. A good fit was determined with the $(0,0)$ beam spot, in both the non z-smeared and the z-smeared sample. After full convergence the x_0, y_0, and z_0 parameters were fixed to their expected values and the fit was done again to get the initial parameters for the other beam spots. There then ended up being two sets of initial parameters, one set is for the non z-smeared samples, and comes from those samples $(0,0)$ beam spot, then the other set is for the z-smeared samples, which is sourced from the z-smeared $(0,0)$ beam spot.

$$\mathscr{L} = \Pi_i^N \exp\left(\frac{\left(\Theta(r_i) + G(z_i) - \Theta_i\right)^2}{2\sigma_{\Theta_i}}\right) \tag{A.12}$$
$$G(z_i) = \left(g_{a1}\sin\left(\phi_i - (s_2 - \pi)\right) + g_{a3}\right)\exp\left(\left(-\left(z_i - z_0\right)/\left(g_{c1}\sin\left(\phi_i - (s_2 - \pi)\right) + g_{c3}\right)\right)\right)$$

The results of the fits are given in Table A.10 for the non z-smeared samples and Table A.11 for the z-smeared and PU samples. Looking at the results, all of the fits fully converged, but some floating parameters had some large errors, which means that one needs to be careful with these fits, and their initial parameters. On average the results are found to be within 1 mm of the true value.

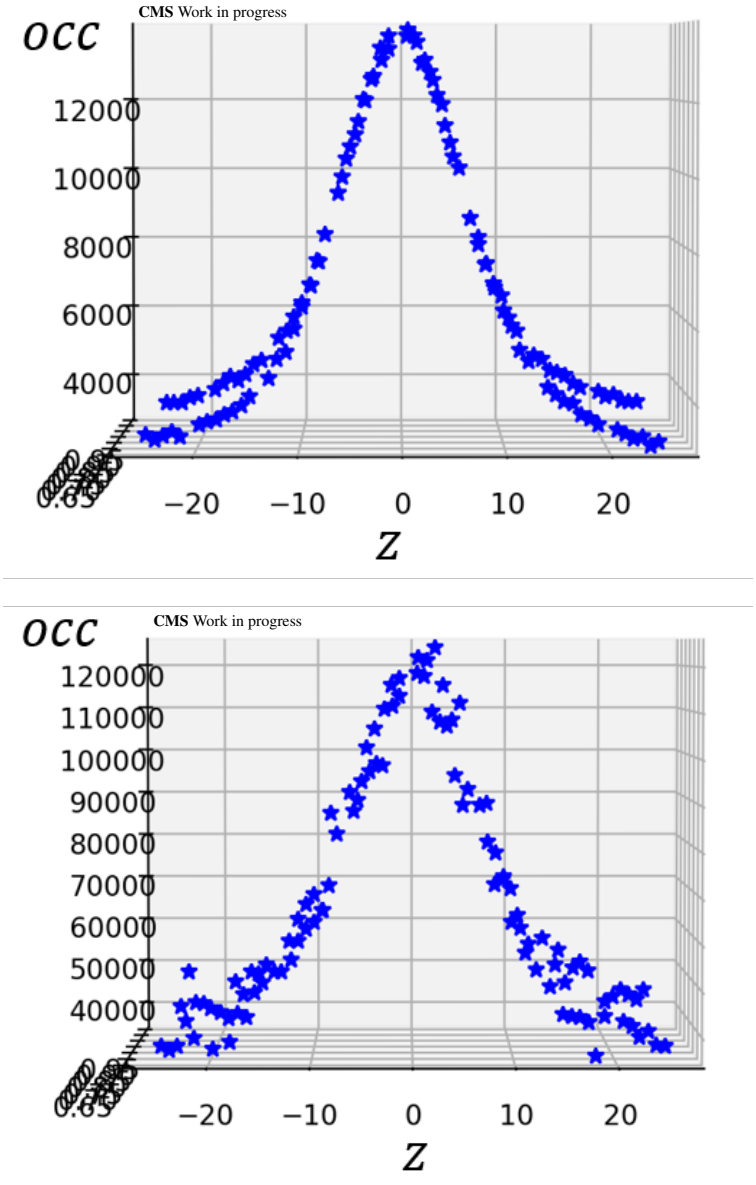


Figure A.33: Distribution of of the 3D ROC occupancy map for a single ladder, looking down the ϕ direction, for a z-smeared (top) and pileup (bottom) sample.

Beam Spot	x_0 (cm)	y_0 (cm)
$(0,0)$	-0.021 ± 0.051	-0.0006 ± 0.04
$(0.01,0.04)$	0.028 ± 0.039	0.028 ± 0.17
$(0.03,-0.06)$	0.005 ± 0.034	-0.014 ± 0.057
$(-0.08,0.02)$	-0.09 ± 0.05	-0.04 ± 0.06
$(0.1,-0.08)$	0.08 ± 0.07	-0.06 ± 0.06
$(-0.1,0.2)$	-0.174 ± 0.01	0.24 ± 0.06
$(0.2,0.19)$	0.27 ± 0.04	0.25 ± 0.06
$(-0.3,-0.32)$	-0.47 ± 0.05	-0.51 ± 0.06

Table A.10: Table of the fitted values for the non z-smeared beam spots using the third iteration of the maximum likelihood fit with a Gaussian term. The parameter errors are calculated by the HESSE method.

When fitting over the non z-smeared samples, the results with the Gaussian term, are not as strong, and for good reason considering that the shape with a non z-smeared sample is fairly different from what a Gaussian would expect. The results for the z-smeared samples with the Gaussian term are all within 1 mm of the true beam spot, and it is also the case for the samples with simulated pileup. One thing that can be seen is that the likelihood fits struggle with the $(0,0)$ beam spots after adding in the ϕ modulation, as the fit does not like forcing the amplitude of the sine functions to zero. In addition, the fitted values are generally underestimated for all of the beam spot samples. Figures A.35 gives the δx and δy for the non z-smeared samples, while A.34 gives the δx and δy for the z-smeared and pileup samples using the third iteration fit. What can be seen from these figures is that, with the exception of the non z-smeared samples, the measured values are all within a difference of 1 mm for the fitted values, though they are not as tightly spaced as for the non z-smeared samples with the second iteration in Figure A.31. One thing to note for the pileup sample results is that the fit did have trouble converging, which matches with what was said earlier about the shape, where full convergence was only achieved after using the same initial parameters as the second iteration, which points to needing to be careful about the initial parameters used in the fits with simulated pileup.

The last check that was done for the third iteration was how the fits are affected for a couple of the z-smeared samples after certain modules and ROCs were removed to simulate some of the dead ones in real data. Figure A.36 gives some top down views of the occupancy map of ϕ against

Beam Spot	x_0 (cm)	y_0 (cm)
z-smeared samples		
$(0,0)$	0.024 ± 0.054	0.025 ± 0.077
$(0.1,0.08)$	0.05 ± 0.05	0.025 ± 0.04
$(-0.1,0.08)$	-0.03 ± 0.034	0.029 ± 0.035
$(0.1,-0.08)$	0.04 ± 0.04	-0.04 ± 0.03
$(0.2,-0.19)$	0.13 ± 0.03	-0.13 ± 0.03
$(0.3,-0.32)$	0.263 ± 0.026	-0.285 ± 0.027
pileup samples		
$(0.2,-0.19)$	0.216 ± 0.012	-0.224 ± 0.012
$(0.3,-0.32)$	0.354 ± 0.008	-0.380 ± 0.008

Table A.11: Table of the fitted values for the z-smeared and pileup beam spot samples using the third iteration of the maximum likelihood fit with a Gaussian term. The parameter errors are calculated by the HESSE method.

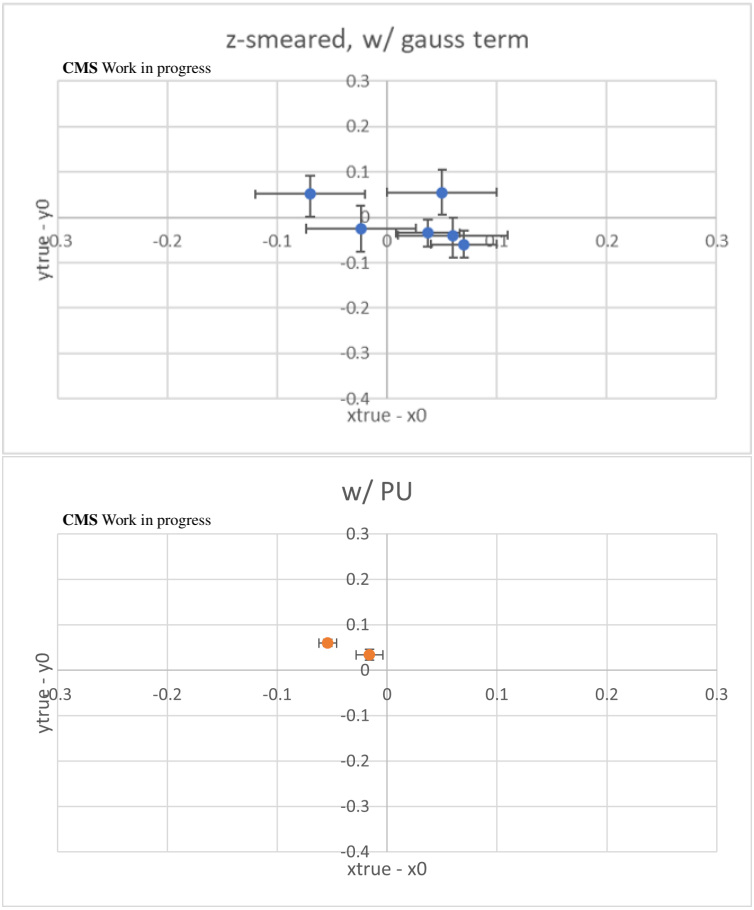


Figure A.34: Distribution of δx and δy in cm for the beam spot values in Table A.9 for the third iteration fits. The δs for the z-smeared samples are on top, and the δs for the pileup samples are on the bottom

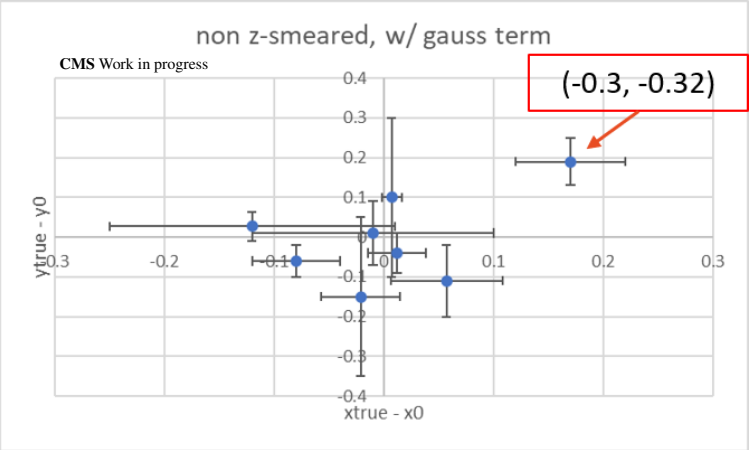


Figure A.35: Distribution of δx and δy in cm for the non z-smeared beam spot values in Table A.8 for the third iteration fits.

z for a z-smeared sample, with certain ROCs and modules removed. The top plot in this figure only removes the dead ROCs and modules that were present in early 2018 data. The bottom plot removes these modules and ROCs, and then on top of that performs a cleaning which removes the outlier ROCs that would correspond to over/under performing ROCs, i.e. ROCs that may have multiple dead or hot pixels. Results for the fits corresponding to different cleanings are given in Table A.12. The results point towards the fitting method having some resistance to the removal of ROCs, but there is a point at which too many ROCs/modules are getting removed, as seen in the row for '2018A cleaning' for the two beam spots. In any case, the error on the fitted values becomes much larger, such that the results, while still within 1 mm of the true beam spot, are also consistent with a beam spot of zero, at least for the $(0.1, -0.08)$ beam spot. The further displaced beam spot shows more resistance to ROC removals, probably due to the greater amplitude in the sinusoidal effect. One problem to be aware of is if too many ROCs and module at central z get removed, the fit will have a tough time finding the beam spot, as most of the sensitivity is coming from this region.

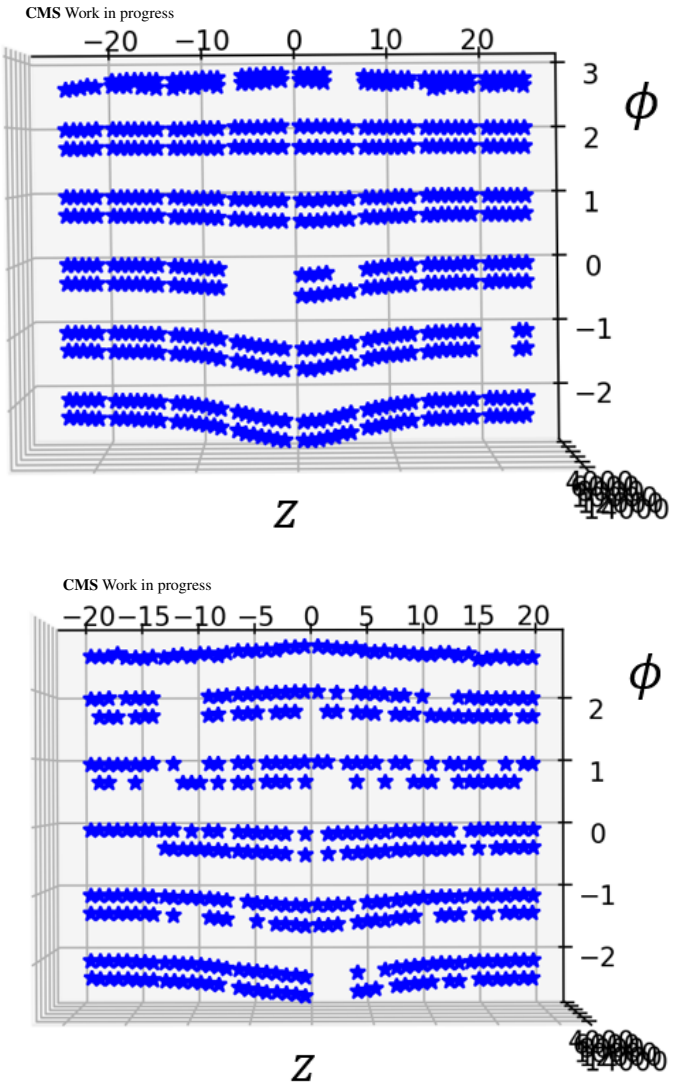


Figure A.36: Distributions of ϕ, z ROC occupancy maps of the $(0.1, -0.08)$ z-smeared beam spot sample, which details which ROCs were removed when performing some fits, where the results are detailed in Table A.12. The label 'Dead modules/ROCs removed' corresponds to the top plot, where only the ROCs and modules that didn't register any clusters in real 2018 data are removed. The label '2018A cleaning' corresponds to the bottom plot, where a cleaning was performed based on which ROCs and modules showed over/under-performance in 2018 data, probably caused by dead/hot pixels in the relevant ROC.

Beam Spot	x_0 (cm)	y_0 (cm)
$(0.1, -0.08)$		
original	0.04 ± 0.04	-0.04 ± 0.03
Dead modules/ROCs removed	0.05 ± 0.05	-0.04 ± 0.03
2018A cleaning	0.018 ± 0.081	-0.013 ± 0.054
$(0.2, -0.19)$		
original	0.13 ± 0.03	-0.13 ± 0.03
Dead modules/ROCs removed	0.13 ± 0.03	-0.133 ± 0.029
2018A cleaning	0.15 ± 0.06	-0.15 ± 0.05

Table A.12: Table of the fitted values for the $(0.1, -0.08)$ and $(0.2, -0.19)$ z-smeared beam spot samples using the third iteration fit after removing different sets of modules and ROCs to simulate real 2018 data conditions. The parameter errors are calculated by the HESSE method.

A.4 Looking at charge collected

Before moving on to the summary, a quick foray is shown into the potential use of the total cluster charge collected per ROC rather than the occupancy. The way the charge is counted is the same as occupancy, except for each cluster, the amount of charge can be variable, while for occupancy it is only counting the clusters. The reason for potentially using the charge collected is that it has a consistent size to the sinusoidal effect over ϕ, unlike the occupancy, where the size of the sinusoidal effect depends heavily on the distance from the interaction point, where the size decreases with increasing distance (as it should). Figure A.37 gives distributions of the occupancy and total charge collected for different sections of z. These distributions highlight the size of the sinusoidal effect, and how the occupancy has a decreasing amplitude, while the charge has a fairly constant amplitude as it moves over z. The potential from this is the larger number of data points that are very sensitive to the beam spot location, as the greater the amplitude, the better the sensitivity, where the amplitude falls off fairly quickly when only looking at the occupancy. A downside to this is that the total charge collected would not follow the power law used for the occupancy, and a new relation would need to be determined.

A second thing that would need to be considered with potentially using the charge collected is the spread in charge collected that is seen in the outer z positions, as shown in Figure A.38. It was found that the cause of this spread in charge collected is due to the size of the charge present

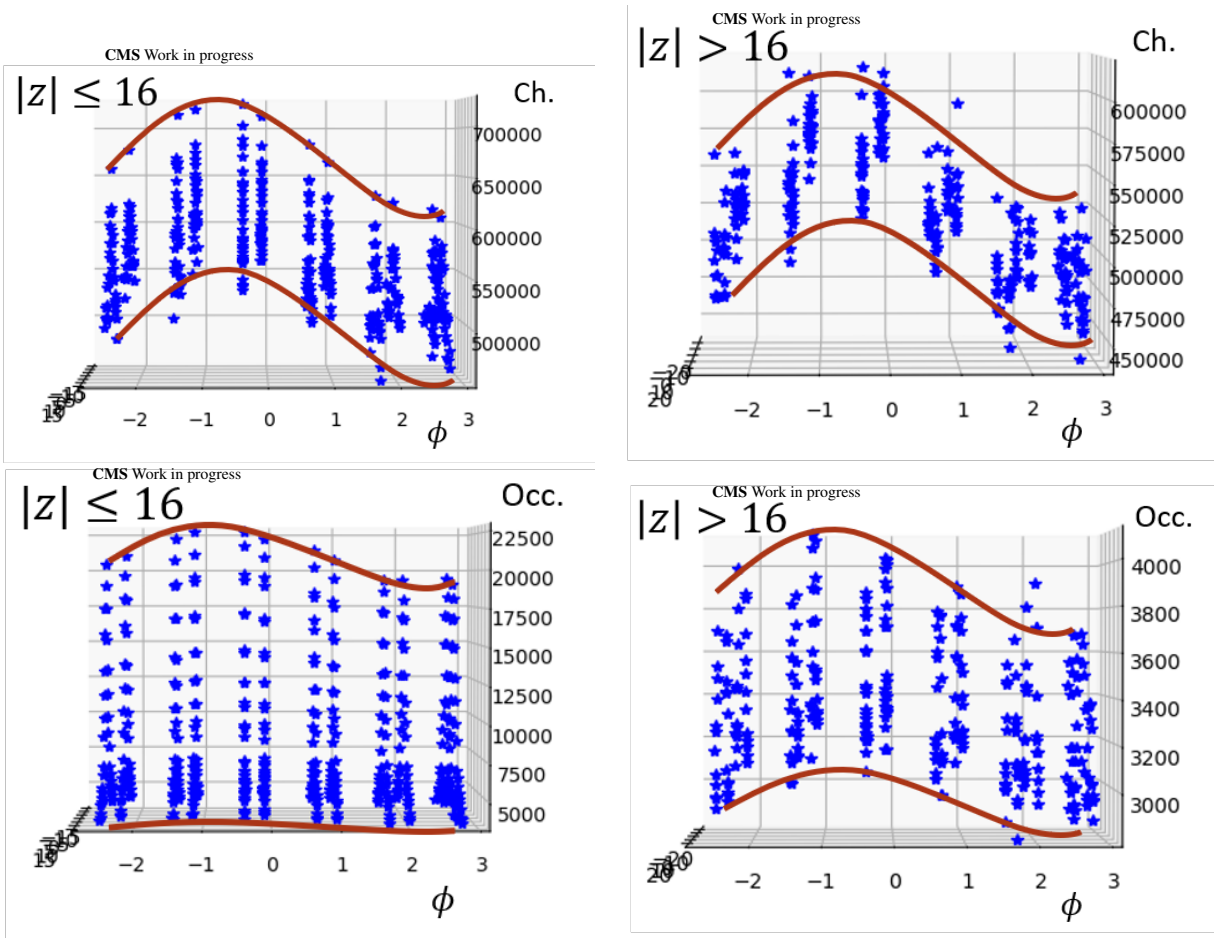


Figure A.37: Distributions of the total charge collected per ROC, looking in the ϕ direction, for a $(0.1, -0.08)$ beam spot without z-smearing applied. The top plots shows the total charge collected for $|z| \leq 16$ cm on the left and $|z| > 16$ cm, while the bottom plots show the same, but for the occupancy instead of charge collected. Sinusoidal shapes are overlaid on the plots to highlight the amplitudes that the sinusoidal shapes have when moving to different z values.

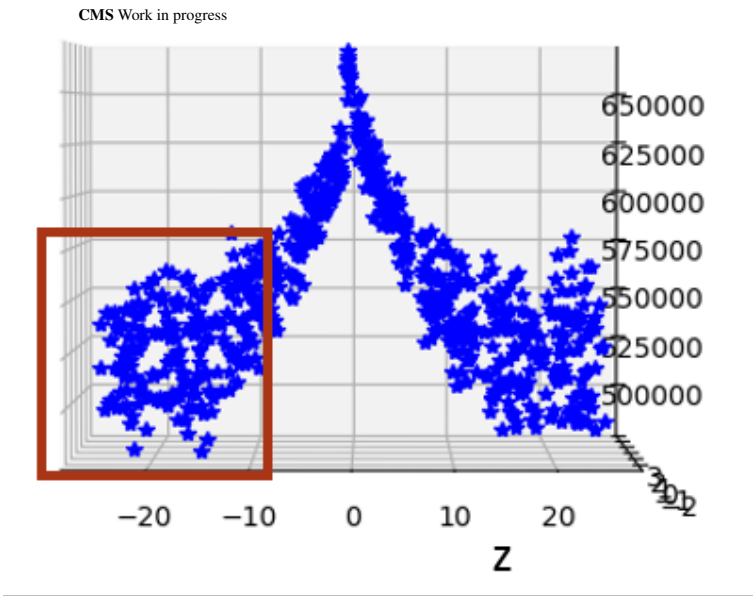


Figure A.38: Distribution of the total charge collected per ROC for a $(0,0)$ beam spot. The red rectangle is to highlight the spread in total charge collected at large values of $|z|$.

in single clusters. Much like how the cluster size is dependent on how far away in z the cluster is from the interaction point, the charge collected depends on the size of the cluster. At higher values of $|z|$ the size of a cluster will be larger due to the particle having a larger angle of incidence, thus making contact with more pixels as it passes through the layer. The larger the number of pixels that are passed through, the greater amount of charge is collected for that cluster.

When the charge collected per cluster is separated out into a high charge and low charge set of clusters, the distribution of total charge collected gains an interesting shape. Figure A.39 shows these shapes, and highlights how the outer z greatly favors clusters with a large charge, while the central z favors clusters with a low charge. In addition, the high charge clusters show a larger sinusoidal amplitude at outer z, while the opposite is true for the low charge clusters.

The potential for using the charge collected is in taking advantage of a large sinusoidal amplitude at many values of z, where for the occupancy, the high amplitude effect is only seen at central z, or small values of r_{xyz}. With the total charge collected, the data can be manipulated such that the high amplitude effect can be seen at multiple values of z or r_{xyz}. The difficulty in using this is how to split the clusters so that a good fit can be obtained between low charge cluster and high charge

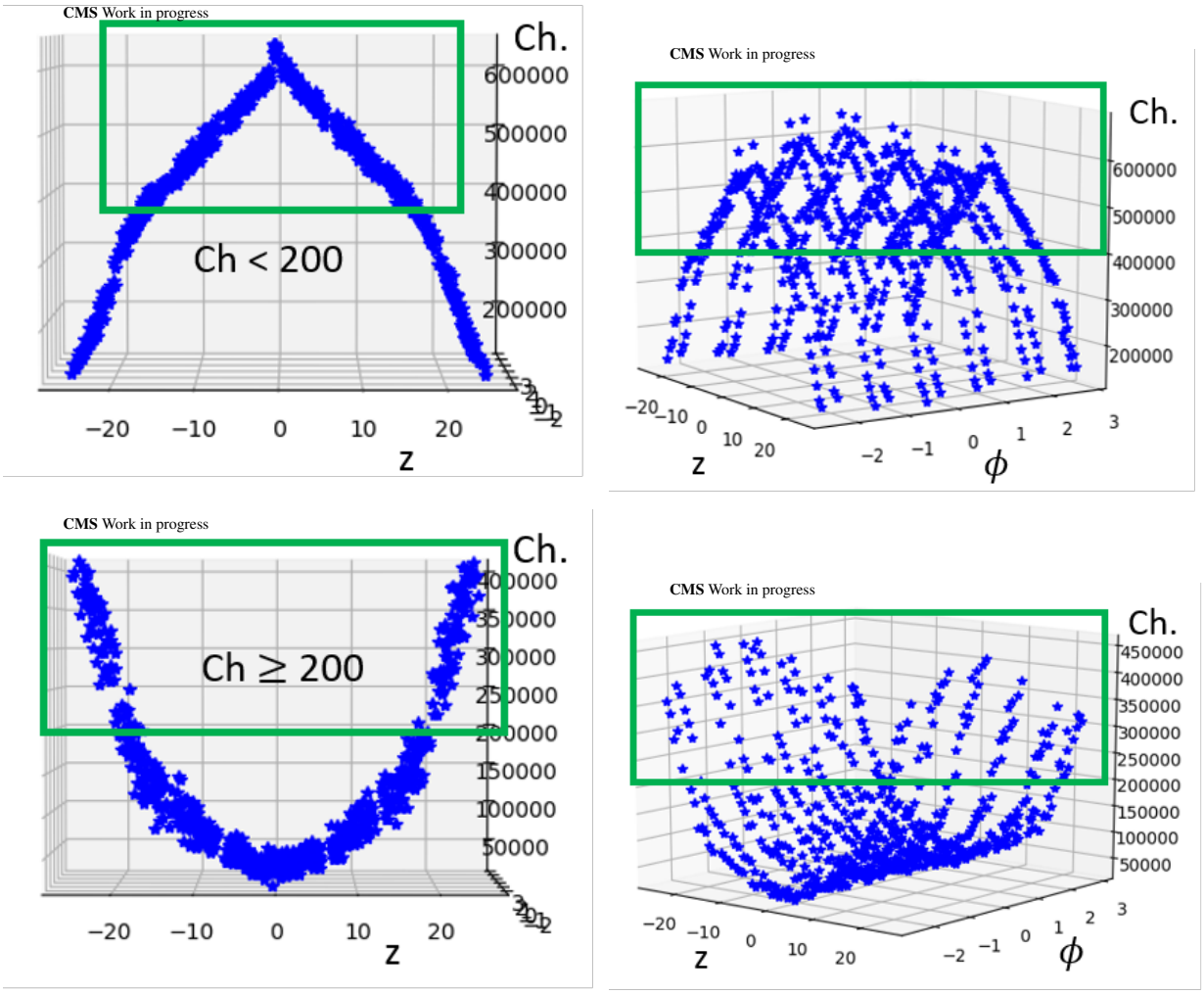


Figure A.39: Distributions of the total charge collected per ROC for a $(0,0)$ and $(0.1,-0.08)$ beam spot. The top plots show the total charge collected for clusters with a charge < 200, while the bottom plots show the total charge collected for clusters with a charge > 200. The left shows the $(0,0)$ beam spot while looking down ϕ. The right shows the $(0.1,-0.08)$ beam spot while looking at a oblique angle between ϕ and z to highlight the sinusoidal shape. The green boxes highlight the area where the sinusoidal effect has the largest amplitude.

clusters. In addition a fit with the total charge would have a greater complexity than the occupancy due to having to simultaneously fit two distinct sets of data. This leaves the use of total charge collected as something that may be useful in the future after more study.

A.5 Beam Spot Summary

A new fitting method was shown for calculating a beam spot without making use of tracking reconstruction. There were multiple fitting methods that were attempted in calculating the beam spot, starting with a simple sine fit, and ending with a maximum likelihood fit. The final method was a maximum likelihood fit that made use of an inverse power law with a Gaussian term to account for changes in the z-position of the interaction point. The inputs to this fit were the occupancy and position of all of the ROCs in the inner ladders of BPix layer 1. Results from fitting showed that the method is able to calculate the beam spot to within 1 mm of the true beam spot, but generally underestimates the true value. It is able to work for samples with either z-smeared interaction points, or simulated pileup.

The next step that would need to be taken for this measurement is finding some real data to test the method on. The choice of dataset is important, as there is a need to make sure that the events aren't too biased such that particles incident upon the detector are no longer relatively equally distributed around the detector. There was an initial check on data with a single muon dataset, but there were problems in which there was no sinusoidal effect seen in the occupancy. After finding a dataset that is usable, there is then the need to figure out the cleaning that would need to be applied. An automated way would be ideal, as removing poor performing ROCs by hand quickly becomes intractable depending on how many need to be removed. Should the method become approved, there will also be a need to port the fitting method into CMS Software, so that it can be calculated directly off of the reconstructed datasets.

Apart from making sure the method still works as the data becomes more complicated, there are also improvements in the process that can be studied. The first one is stabilizing the fit, especially for pileup samples. This would either entail finding a good set of initial parameters that lead to

full convergence, or tweaking the free parameters to improve minimization. Another improvement would be expanding the fit to beyond just the inner ladders of layer 1. Since the fit is working with radii out to 25 cm, it is definitely possible to add in the outer ladders of layer 1, and possibly even the ladders of layer 2. The purpose of this expansion would be to increase the sensitivity to the beam spot location. As it is, the fit probably relies too much on the inner-z modules, since the sensitivity drops fairly quickly as one moves our to larger z values. This is where the potential of the total charge collected comes in, where depending on how the data is manipulated, it is possible to get increased sensitivity to the beam spot at larger values of z. Overall, the fitting method works for ideal cases with MC samples, but there are multiple ways in which the method can be improved.