University of California
Santa Barbara

Search for supersymmetry at CMS in proton-proton collisions with center-of-mass energy of 13 TeV in the single-lepton final state using the sum of masses of large-radius jets

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

Doctor of Philosophy
in
Physics

by

Alex Dorsett

Committee in charge:

Professor Jeffrey D. Richman, Chair
Professor Claudio Campagnari
Professor Nathaniel Craig

June 2022

The Dissertation of Alex Dorsett is approved.

Professor Claudio Campagnari

Professor Nathaniel Craig

Professor Jeffrey D. Richman, Committee Chair

May 2022

Search for supersymmetry at CMS in proton-proton collisions with center-of-mass energy of 13 TeV in the single-lepton final state using the sum of masses of large-radius jets

by

Alex Dorsett

To those who educate for education's sake.

Acknowledgements

Well, here I finally am. Despite it coming first, I wrote this section last because properly acknowledging all the people who have helped me along the way was a very daunting task. Here goes nothing...

I would like to start by giving my deepest thanks to my mom and sister. Without their love and support over the years, none of this work would have been possible. Words can not express my gratitude for the endless encouragement they have given me throughout my academic journey.

I would like to thank all my teachers at Butte Community College. In particular, Michael Panunto, Robert White, Jason Trento, Rich Bigler, John Marhenke, and Lisa Nichols. The breadth and quality of education I received at Butte gave me the confidence to change my major to Physics, and prepared me for the rigor of coursework at the university level.

I wouldn't have made it far in grad school without the research experience I had at UC Davis, so I would like to thank all those whose generosity made this experience possible. Rena Zieve, whose effort in organizing the Davis REU gave me my first chance to do university research; Shirley Chiang, my first advisor, who taught me the patience required to do research; Manuel Calderon de la Barca Sanchez and Daniel Cebra, for sparking my love for high-energy physics and giving me invaluable advice on applying to graduate school; and Chad Flores and Christopher Flores, for being great mentors and making me a better programmer.

I would also like to thank the many people at UC Santa Barbara that have made my time here so wonderful. First and foremost, my advisor Jeffrey Richman, whose focus on my enrichment throughout my graduate career has left me a much better scientist and communicator. Our mutual enthusiasm for education led to the most fulfilling project I

had in grad school. To the graduate students who came before me: Jack Bradmiller-Feld, my past and future mentor; Rohan Bhandari and Ryan Heller, who warmly welcomed me into their office space; and Adam Dishaw, for leaving behind an elegant and well-documented code base that has benefited me greatly. To the post-docs I have worked with I have worked with: Ana Ovcharova, my infinite questions were no match for your infinite patience (and answers); Manuel Franco Sevilla, who inspired me to make pretty slides and climb mountains; and JaeBak Kim, for helping me figure out how combine does what it does.

In the end, the real discoveries were the friends I found along the way, so I would like to thank all those I was fortunate enough to find: Dylan Nesbit and Jamison Burke, the first friends I made in grad school who have become my lifelong smash bros; Melissa Quinnan, for keeping me sane during coursework and making sure I experienced some of what Europe had to offer; Michael Oshiro, for humoring me when I was confused about high-energy physics, our chats improved this thesis greatly; and Nicholas Zobrist, for making sure I exercised regularly and enabling my obsessions—from ping pong to disc golf—I can't wait to see what we get into next.

Lastly, I would like to thank my amazing girlfriend Jennifer Smith for her support during the prolonged period over which this thesis was written. The last couple of years have been weird, but I wouldn't have wanted to spend them with anyone else. I'm so excited for the future you, Cosmo, and I will have together.

Curriculum Vitæ

Alex Dorsett

Education

2022	Ph.D. in Physics, University of California, Santa Barbara
2019	M.A. in Physics, University of California, Santa Barbara
2016	B.S. in Physics, *summa cum laude*, University of California, Davis

Selected Publications

CMS Collaboration, "Search for supersymmetry in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with $137\,\text{fb}^{-1}$ in final states with a single lepton using the sum of masses of large-radius jets," *Phys. Rev. D* **101**, no. 5, 052010 (2020) doi:10.1103/PhysRevD.101.052010 arXiv:1911.07558.

Abstract

Search for supersymmetry at CMS in proton-proton collisions with center-of-mass energy of 13 TeV in the single-lepton final state using the sum of masses of large-radius jets

by

Alex Dorsett

Results are reported from a search for supersymmetric particles in proton-proton collisions in the final state with a single, high transverse momentum lepton; multiple jets, including at least one b-tagged jet; and large missing transverse momentum. The search uses a $136\,\mathrm{fb}^{-1}$ sample of proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ accumulated by the CMS experiment in 2016, 2017, and 2018. The search focuses on processes leading to high jet multiplicities, such as gluino pair production with $\tilde{\mathrm{g}} \to \mathrm{t\bar{t}}\,\tilde{\chi}_1^0$. The quantity M_J, defined as the sum of the masses of the large-radius jets in the event, is used in conjunction with other kinematic variables to provide discrimination between signal and background and as a key part of the background estimation method. The observed event yields in the signal regions in data are consistent with those expected for standard model backgrounds, estimated from control regions in data. Exclusion limits are obtained for a simplified model corresponding to gluino pair production with three-body decays into top quarks and neutralinos. Gluinos with a mass below about 2150 GeV are excluded at a 95% confidence level for scenarios with $m(\tilde{\chi}_1^0) < 700\,\mathrm{GeV}$, and the highest excluded neutralino mass is about 1250 GeV.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The work presented in this thesis concerns the details and results of an *analysis*. This term is ubiquitous in and outside of the physical sciences, and has a variety of meanings. In high-energy physics, analysis refers to a particular procedure of searching for a signal model in data. Throughout my time in grad school I have learned a lot about what this procedure entails, but have never found it succinctly described in one place. Therefore, this introduction will serve as my attempt to provide such a high-level description. I will list the fundamental steps of an analysis, and provide a brief summary for each. Though these steps will be presented in the context of high-energy physics, I have tried to keep them general enough to apply to any kind of analysis. For a detailed example of an analysis, feel free to read the rest of the thesis.

1.1 What is an analysis?

Outside of high-energy physics, experiments are usually designed and conducted to measure quantities associated with a single process, or to quantify the behavior of some system. This involves making measurements of a system when the process of interest

occurred, and contrasting them with measurements when the process didn't occur. At the LHC, however, the data collected contain a mixture of all possible processes, and it is impossible to know what process is occurring at a particular time. This makes the task of analyzing the data fundamentally different. It is not a matter of counting the number of times a particular process occurs. Rather, it is determining if events with a particular set of properties occur as frequently as you would expect. This may seem like a subtle difference, but it requires a different approach to analyzing the data.

When designing an analysis, there is an order to how you study the data available to you. You start by only use simulated data, until you have determined the primary features of the signal of interest. Once you have the main selections and signal defined, you can look at regions of data which are free of the expected signal. This is called *blinding*, and is done to avoid introducing bias into the analysis. If you look at data too early you may notice a feature in the data which, in reality, is nothing but a statistical fluctuation. If you interpret it as evidence for a signal, however, you may tune your analysis strategy in order to isolate and enhance this feature. In doing so, you could create an artificial signal in your data out of nothing but a fluctuation. This false result would get disproved once another data sample is studied, but it would likely cause a lot of headaches in the mean time. For this reason, it is forbidden to look at search regions (where we expect signal to be) in data before the complete analysis strategy has already been determined. The procedure of finally looking at search regions in data is called *unblinding*, and is always one of the last steps of an analysis. To further familiarize ourselves with this methodology, we will go over the basic steps one goes through when conducting an analysis at the LHC. This is not meant to be an all-encompassing guide, but rather a brief overview of the important steps and procedures which go into an analysis. These steps are represented as a flowchart in Figure 1.1.

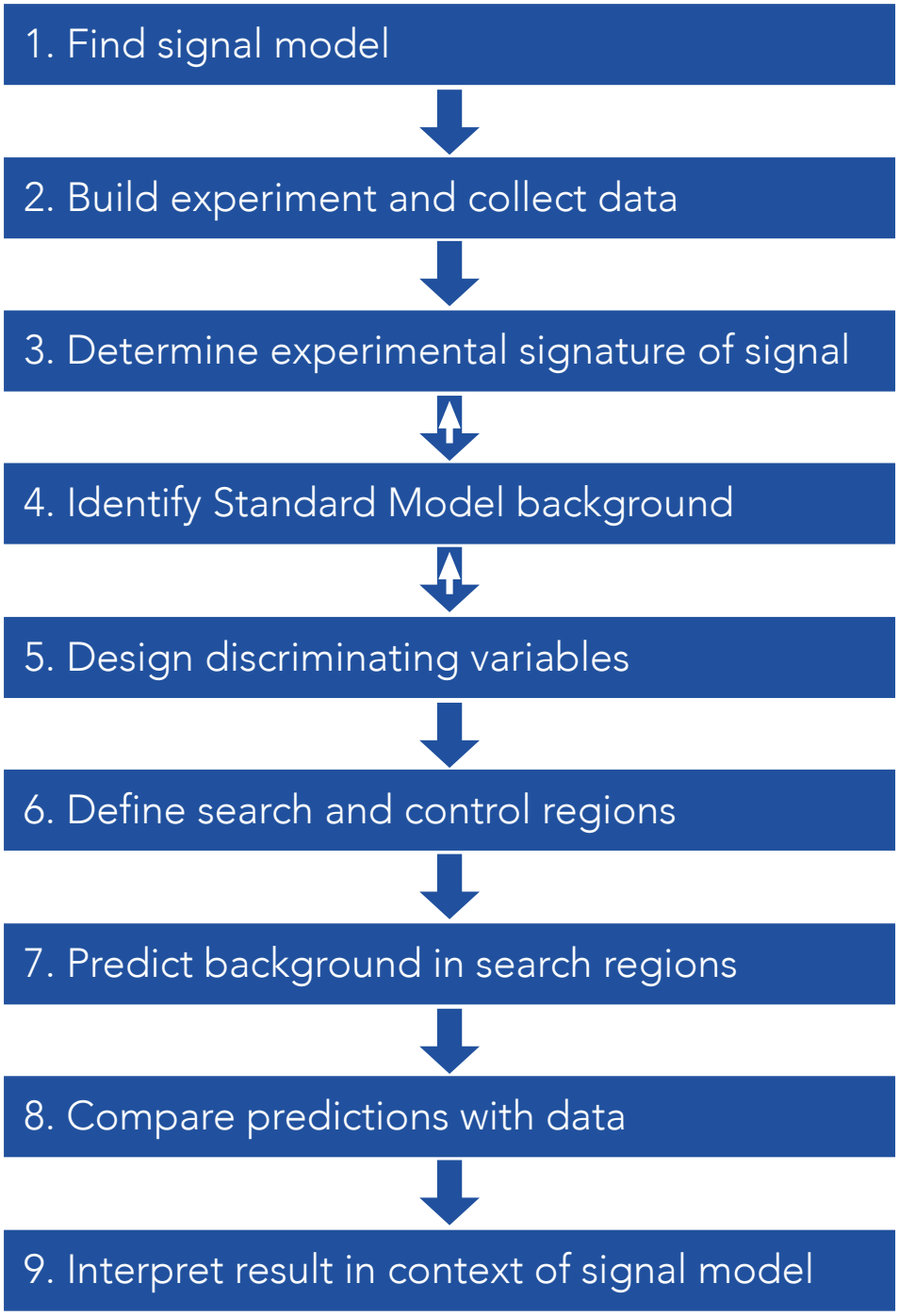


Figure 1.1: Flowchart of an analysis.

Step 1: Find signal model

The first step to take when starting an analysis is to select the model you will be searching for. The range of models you have to choose from is determined by your data. Many experiments are designed with a particular signal type in mind, so the data they collect may only be suited to search for that signal and not much else. In proton-proton collisions, however, the dataset contains a myriad of processes. Furthermore, CMS and ATLAS are general purpose detectors, meaning they collect a wide range of information about each event. Because of this, the data of these experiments can be used to search for many different types of models. Once you've determined the span of models you can search for, it's time to select the model you're interested in. Depending on the models available, this can fall into a couple categories. Namely, a model in the standard model—either an interaction that is allowed but hasn't been observed before, or one that has been observed but not fully characterized—or a model outside the standard model—such as supersymmetry, dark matter, etc. For those outside the standard model, not all models are created equal. When choosing from this category, the decision is usually made based on which models are the most well-motivated. The exact criteria that go into this label are quite complicated, and it is good practice to consult a theorist when deciding on a model of this type. Additionally, you don't need to only select one model to search for. Some analyses are *inclusive*, which means they search for a broad class of models that give similar signatures.

Step 2: Build experiment and collect data

Every kind of analysis requires data of some form or another. In some cases, data is generated by an experiment that induces interactions in which the signal may occur and then be measured by a detector. In other cases, the signal interactions are already

occurring and a detector is designed to record them. In either case, the quantities measured by the detector must be sensitive to the presence of signal. The collections of measurements made by the detector in a given time period is known as an *event*. Events are the fundamental unit of data in high-energy physics.

Step 3: Determine experimental signature

Given a well-motivated model, determining how it would manifest in data still isn't necessarily straightforward. In the case of the Higgs boson, it took decades before its properties could be pinned down enough to predict a signature within the reach of experiments. Even when you have a good understanding of the final states produced by your model, there is still some work to be done to decide which of these channels to search for. This entails determining how well you would be able to measure the products in each final state, how common each final state is, and what standard model processes would give similar products (discussed more in the next section). Finding a final state with a high branching fraction, small background, and good detector resolution is ideal, but rarely possible. Often it is a matter of finding the happy medium between these features, the relative importance of which will be different for each model. For the Higgs, a unique final state with a very small branching fraction gave much more sensitivity than a more common final state with a large branching fraction. We'll see more about determining sensitivity in the next section. Once you've chosen a final state, you can define a *baseline selection*, which is a set of cuts that captures the primary features of the signature. The goal of the baseline selection shouldn't be to absolutely minimize the amount of background, but rather to define the broad category your signal inhabits (i.e. 1 or more leptons, high number of jets, missing energy, etc.).

Step 4: Identify standard model background

Now that you have selected a model and signature, you next need to determine how unique this signature it is. To do this, you need to consider what standard model processes would produce a similar final state. When doing this, you should also take into account the role the detector plays in event reconstruction. So, on top of those processes which directly produce the final state, you should also consider processes that could get *measured* to have the same final state (i.e. if you are searching for a final state with one lepton, you should consider processes where 2 leptons are produced, but only one of them is measured). For each relevant background, acquire simulated samples and study their properties, to get a quantitative measure of how much they will contribute to the overall background. Once you have done this for each background, bring all of your results together, and determine how sensitive this background is to the signal. *Sensitivity* is a measure of the potential for a deviation from the expected background to be noticeable in data. This takes into account the uncertainty in the total background, as well as the expected yield of the signal. Once you've determined the sensitivity for a channel, compare it with the sensitivity to a few other final states for your model. Though this requires having to perform this procedure multiple times, it will help make sure that your search is done in the most sensitive channel.

Step 5: Design discriminating variables

Now that you've determine what your primary backgrounds will be, you can start to think about how to further distinguish them from your signal. Given you determined these backgrounds to be those which most resemble the signal, this isn't always a simple task. It will involve a lot of trial and error: design a variable and then study how it behaves in background and signal. Furthermore, many times it is impossible to find

a single variables which separates the signal from background. Rather, it is a matter of finding the right combination of cuts on multiple variables in order to accomplish separation. It is also important to note that sometimes the primary background has a signature that is identical to the signal process, and there isn't a set of cuts that can be applied to separate the two. This is called an irreducible background. In these cases, it becomes more important to understand this irreducible background well, and accurately predict its behavior in the signal region. This is precisely what control regions are for.

Step 6: Define search and control regions

At this point, we have determined an experimental signature for our signal, identified the corresponding background, and designed variables to discriminate between the two. Now it's time to define where we will be conducting the search, know as the *search region*. This is where we expect the signal events to be found, and is defined using the optimal combination of the discriminating variables found in the previous step. Next, we need to define *control regions*, which have no overlap with our search regions, and can be used to study the leading backgrounds. Usually, there are several control regions, each of which is dominated by a different background process. These control regions will be used to compare simulations with data to verify the background is well-modeled in the simulation. Significant discrepancies between simulations and data may also indicate that there is a contributing background process which was not included. If the simulations agree with the data in the control regions, this gives us confidence that the simulation models the backgrounds well in the search region as well. Finally, it is important to make sure that your control regions will not have any signal events in them. This is called signal contamination, and it will cause you to overestimate the amount of background in your search region because the control region you use for this estimation isn't just background.

Step 7: Predict background in search region

In order to tell if there is any signal present in the real data, we first need to predict how much background we expect there to be. In the last section, we learned about defining control regions which are dominated by the background processes. Now, it's just a matter of extrapolating the background from these control regions to the search regions. Though it's tempting to rely heavily on the simulations, care should be taken when doing so. Though we have simulations of all the physics processes we know about, these aren't always perfect representations of what we measure in data. Small differences in behavior between the simulation and data can lead to incorrect background estimation or a high amount of uncertainty. There are, however, a range of data-driven background estimation methods that can be used. We will describe one such method in Chapter 8

Step 8: Compare prediction with data

Up to this point, we have been relying on simulations, as well as control regions in data, to design and refine our analysis. We have defined signal-rich search regions and have found a prescription for predicted the expected behavior in them from background-dominated control regions. Now is the moment of truth: it's time to unblind. On top of predicting what we expect the background to look like in the search region, we will also need to predict what we would expect the signal would look like. After all, the data needs to disagree with the predicted background in a particular way to indicate the presence of signal. Once we have these two predictions, we look at the data, and determine which option it is more consistent with. How exactly we determine this is covered in the next section.

Step 9: Interpret result in context of signal

Now that everything is said and done, it's time to step back and determine what we have found. We set out to search for a particular model, and so we need to interpret our results in the context of this model. Depending on what we found, this can look a bit different. If evidence for the signal is observed in data, then we will report the how significant this evidence is. If the data shows no evidence of signal, then we will report what effect this has on our understanding of the model. The procedure for each scenario is a bit different, so let's explore both.

In the case in which an excess is observed in the search region in the data, we must quantify how significant this excess is. To do this, we invoke the null hypothesis, which states that the observed excess is simply a fluctuation in the background (after all, predictions always have uncertainty). We then compare this with the signal hypothesis, and assess which one is more consistent with the observed data. From this test, we can calculate a significance level, which quantifies the likelihood that an observation is caused by a signal model compared to a background fluctuation. Significance levels are often quoted as a number of standard deviations of a Gaussian distribution (number of σ). For example, reporting a 1σ result means that the probability that the observed data is caused by background fluctuation is 16%. In order to claim discovery of a new particle, an analysis must have a significance level of 5σ or higher, corresponding to a probability of 3×10^{-10} that the data is caused by a fluctuation. It's worth noting that there isn't anything particularly special about a significance of 5σ, but it is the standard of high-energy physics. That's not to say that you can't report your results if you only have 4σ, but you just can't use the word "discovery" in your title (often "observed excess" or "evidence for" are used).

So what is one supposed to do with a null result? Though all the analyses we have

looked at thus far have been successful in their searches, this is far from the norm. In fact, there are many more searches that did not find anything at all. Fortunately, null results are not meaningless. They can be used to set cross section limits on the model being searched for, and are usually expressed in terms of upper limits on the production cross section (in the case the search is for a decay mode of a known particle, then a limit on the branching fraction is given). Often when we are searching for a new particle, we don't know the exact value of its mass, so we calculate an upper limit for each value of the mass separately. When the limit set on the cross section is below the value predicted by a model for a given value of particle, it is said that this mass value is "excluded". Though it's not as exciting as having a press event and a Nobel prize, limit setting is crucial to future researchers. Limits on the productions cross section of particles in a model are still worth publishing, as they offer results that go beyond the model being searched for. In fact, new models can be rejected based on limits placed by previous searches for other models. When the data results for an analysis are published, they can reinterpreted—re-analyzed from the context of another model. This means that data from a search for one model can be used to set limits for other models as well (as long as the analysis uses general enough methods to allow for this). This is why it is important to set the best limits we can when a null results is found.

Chapter 2

The Standard Model

2.1 Experimental foundations of the standard model

This section briefly reviews the experiments and discoveries that laid the foundations of the standard model, and gives a sense for the searches being conducted today. In this review, many of the tribulations the field has gone through to get where it is today will be omitted. Instead, this review will interpret the results from the privileged position of having a framework everything (mostly) fits into. The scientists who made the discoveries weren't so lucky, and occasionally results gave consternation rather than clarity. Thus the field progressed, with experimental results informing new theories, which in turn motivated the design of further experiments.

A timeline of the prediction and discovery of each particle in the standard model is shown in Figure 2.1. Here we will discuss a few of the discoveries and the theoretical implications that came with them. We will start with the discovery of the positron, which introduced a new type of matter. Next, we'll discuss the prediction and discovery of the neutrino, which demonstrated some particles can't be readily detected. We'll also show how this invisibility was accounted for in the discovery of the W. Finally,

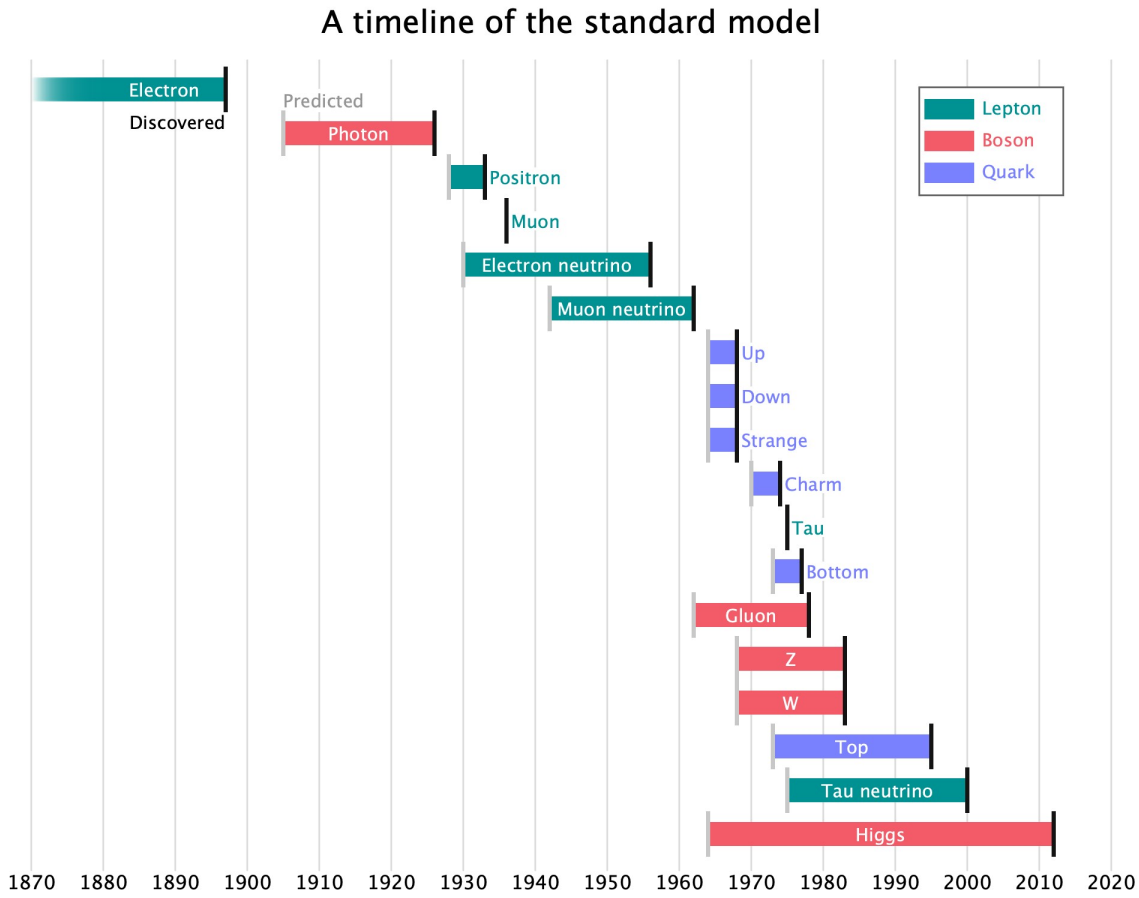


Figure 2.1: Experimental timeline of the standard model. For each particle, the date of theoretical prediction is marked with a gray dash, while the date of discover is marked in black. In the case of the electron, the exact date of prediction is unclear, while for the μ and τ, they were discovered before any predictions of their existence had been made.

we'll recount the saga of the top quark, with its many limits and the new energy scale that came with its eventual discovery. These results were chosen in particular because they highlight many of the key features present in supersymmetric models: new types of matter, invisible particles, and increasing energy scales.

The first paradigm shift of particle physics came with the discovery of the positron. After developing a relativistic wave equation for free electrons ,which we will explore later, Paul Dirac wasn't sure what to make of the negative energy solutions it allowed [13]. This problem was soon solved by results from an experiment conducted in 1932

by Carl Anderson [14], when Anderson was studying the tracks of charged cosmic rays in a cloud chamber in a magnetic field. He observed a positively charged particle that appeared to have a mass close to the electron. After determining the track could not have been caused by a proton, Anderson concluded it was caused by a new positive particle. Thus, the perceived flaw of the Dirac equation turned out to be a profound truth: for every particle it described, there is an antiparticle with the same mass but opposite charge. The discovery of the positron legitimized Dirac's description of spin-1/2 particles, which was eventually developed into the extremely successful theory of quantum electrodynamics (QED). Furthermore, the discovery introduced the concept of *antimatter*: a new form of matter that mirrored normal matter. Thus every fermion was expected to have an antiparticle with the same mass, but opposite charge. This introduction nearly doubled the number of particles that would come to be expected in the standard model.

Though all particles but two were predicted before they were discovered, these predictions were often attempts to explain previous experimental results. This was the case for the neutrino, whose existence was inferred from studies of nuclear beta decay. At the time, it was understood that an unstable nucleus could decay into a lighter nucleus with a different atomic number, emitting an electron in the process. When looking at the energies of the electrons being emitted from a particular beta decay, experiments observed a spectrum rather than a single value. This was contrary to the behavior expected from a two-body decay, in which the products should have a definite energy. The result produced a myriad of responses in the community, with Niels Bohr going so far as conjecturing that energy was only conserved on average. Fortunately, Wolfgang Pauli had a better take, proposing that there was another particle produced that escaped detection. The energies of decay products are not fixed in three-body decays, so the presence of a third particle would explain the electron energy spectrum observed. Pauli initially named the

new particle the *neutron*, but the "ino" suffix was later added due to its predicted mass being extremely small.

Though the prediction of a weakly interacting neutral particle was attractive in that it accounted for missing energy observed in nuclear beta decay, it took two decades before these particles could be directly detected. There had been many experiments in which the neutrino's presence could be inferred, but it had never been directly observed. For particles that pass through the earth by the trillions without interacting, direct detection is a challenge to say the least. As is often the case with such a rare signal process, the solution was to set up an experiment with favorable interaction conditions and then wait. Designed by Clyde Cowan and Frederick Reines, the experiment consisted of two 200 L tanks of water placed next to a nuclear reactor [15]. Antineutrinos produced in the reactor would interact with hydrogen nuclei in the water through inverse beta decay ($\bar{\nu}_e + N \rightarrow N' + e^+$). The resulting positron then annihilated with an electron and produced two photons, which were measured. Between the high number of targets afforded by the water and the high neutrino flux of the reactor, they were able to achieve an estimated three signal events per hour. Over a period of five months in 1956, data were collected for ∼900 hours with the reactor on, and ∼250 hours with the reactor off. After comparing the two data samples, Cowan and Reines measured a cross section for inverse beta decay within 5% of the theoretically predicted value.

The discovery of the charm quark in 1974 completed the two generations of fermions known at that time: two charged leptons with their corresponding neutrinos, and two pairs of up- and down-type quarks. This two-generation picture didn't last long, however, as the τ unexpectedly appeared on the scene a year later, revealing there was at least another generation of leptons. Two years after that in 1977, the down-type bottom quark was discovered, maintaining the symmetry between the number of lepton and quark generations. With a mass of 4.2 GeV, the bottom was the heaviest elementary particle

observed at the time. It was expected the yet-observed top quark would follow the trend of having a higher mass than its down-type counterpart, though it wasn't clear how much heavier it should be. While neutrinos were easy to produce but hard to measure, the top was hard to produce but would decay into measurable products (mostly). The search for the top quark became a matter of developing an experiment able to produce it. In the 18 years following the discovery of the bottom quark, experiments with higher and higher energies searched for the top quark and, not finding it, put higher and higher limits on its mass.

That isn't to say there weren't any successful experiments in this period. In 1983, both the W and Z were discovered in proton-antiproton collisions by the UA1 and UA2 experiments at the Super Proton-Antiproton Synchrotron (S$p\bar{p}$S) at CERN [16, 17, 18, 19]. With masses of 80.4 GeV and 91.2 GeV, respectively, the W and Z were the first particles discovered at a hadron collider. Furthermore, as it was identified through its decay to a lepton-neutrino pair, the W was the first particle discovered using missing energy. This achievement was made possible by the hermeticity of the detector, which covered nearly all 4π steradians around the interaction point with calorimeters. This level of coverage became the norm for general-purpose detectors at collider experiments.

Despite S$p\bar{p}$S reaching a center-of-mass energy of 540 GeV, the top quark remained undiscovered. By the end of its life cycle in 1990, the UA2 experiment had put a lower limit of 69 GeV on the top mass [20]. Around this time, the Tevatron began running at Fermilab, colliding protons and antiprotons at a center-of-mass energy of 1800 GeV. After two years of collecting $4\,\mathrm{pb}^{-1}$ of data, the CDF experiment increased the limit to 91 GeV [21]. Finally, in 1995, with integrated luminosities of $67\,\mathrm{pb}^{-1}$ and $56\,\mathrm{pb}^{-1}$, respectively, the CDF [22] and DØ [23] experiments discovered the top quark at a mass of 173 GeV—40 times heavier than the second-heaviest fundamental fermion. After 18 years of incremental progress and limit-setting, the top had been found. By far the

heaviest elementary particle, this discovery motivated further exploration at even higher energies.

This has been a very brief survey of the experimental developments of the standard model. It's a story of new types of matter, invisible particles, and ever-increasing energy scales. If you enjoy having these qualities in your model of nature, you'll be big fan of supersymmetry. Before we get into that, though, we'll need to do some math. Far from just a list of particles, the standard model is a gauge theory describing particle fields and all the interactions between them. As an introduction to gauge theories, we will examine the most simple pieces of the standard model, Quantum Electrodynamics. We will see how gauge symmetries are introduced into the model, and how they lead to interactions between particles.

2.2 QED from the Dirac Lagrangian

In simple terms, a gauge theory is one in which the Lagrangian doesn't change under a local transformation. This invariance is achieved by including additional fields in the Lagrangian. These new fields, referred to as *gauge fields*, introduce interactions into the model. In this section, we will see how this procedure can be applied to the Dirac Lagrangian to give Quantum Electrodynamics (QED).

Based on the form of the free particle Dirac equation, the Dirac Lagrangian is

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(x)(i\gamma^\mu \partial_\mu - m)\psi(x), \tag{2.1}$$

where $\psi(x)$ and $\bar{\psi}(x) \equiv \psi^\dagger(x)\gamma^0$ are bispinor fields representing spin-1/2 particle-antiparticle degrees of freedom and γ^μ are the Dirac matrices—4×4 matrices which satisfy the anti-commutation relation $[\gamma^\mu, \gamma^\nu] = 2\eta^{\mu\nu}$. We will check this Lagrangian for invariance under U(1) gauge transformations [1] where U(1) is the mathematical group of 1×1 unitary ma-

[1] Though "gauge transformation" implies some change in coordinate scale, it actually refers to a

trices (complex numbers that can be written as $e^{i\chi}$, where $\chi \in \mathbb{R}$). Before checking for local invariance, let's start by checking for global U(1) invariance—in which all points are transformed in the same way. To do this, we introduce the transformation

$$\psi(x) \to \psi'(x) = e^{iq\chi}\psi(x), \tag{2.2}$$

and check if the Lagrangian changes. Since χ is a constant, the $e^{iq\chi}$ from $\psi(x)$ can simply commute through ∂_μ and cancel the $e^{-iq\chi}$ introduced in the transformation of $\bar{\psi}$. Once this is done, the Lagrangian returns to it's original form. Therefore, this Lagrangian is invariant under global U(1) transformations, thus having a *global symmetry*. Applying Noether's theorem, which states there is a physical quantity observed for every symmetry in the Lagrangian, we find this symmetry leads to the conservation of electric charge. However, we haven't modified the Lagrangian in any way—it still only describes free spin-1/2 particles.

To make things more interesting, let's impose a more strict condition: *gauge invariance*. We will perform the same transformation as before, but allow χ to be a function of the spacetime coordinate x. Whereas the previous transformation was global, this transformation will be local—different at every point in spacetime. After performing this transformation, we can quickly see that things aren't going to workout as nicely as before. Since χ is now a function of x, we can't commute it through ∂_μ without picking up an extra term, so we end up with the expression

$$\mathcal{L}'_{\text{Dirac}} = \bar{\psi}(\gamma^\mu(i\partial_\mu + q\partial_\mu\chi) - m)\psi. \tag{2.3}$$

Though this may seem like a quick and underwhelming end to our foray into gauge invariance, it is actually where things get interesting. When a Lagrangian fails to be invariant under a given transformation, gauge fields are included to fix it. The number of gauge fields included is determined by the dimensionality of the transformation group

change in phase.

(for U(1) this is 1). In this case, we will salvage the Dirac Lagrangian by adding the electromagnetic field (we will see soon why we chose this field in particular). This field is added by modifying the derivative operator to include the electromagnetic potential A_μ (the resulting D_μ is called the *gauge covariant derivative*),

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + iqA_\mu(x). \tag{2.4}$$

With this change, our new Lagrangian is

$$\mathcal{L}_{\text{Dirac+EM}} = \bar{\psi}(i\gamma^\mu\partial_\mu - q\gamma^\mu A_\mu(x) - m)\psi. \tag{2.5}$$

Before checking for gauge invariance again, we need to determine how the new field A_μ will transform. Fortunately, since we chose a field with well-established transformation properties, this isn't too hard to figure out. Under gauge transformation, the potential A_μ transforms as follows:

$$A_\mu \rightarrow A'_\mu = A_\mu - \partial_\mu\chi. \tag{2.6}$$

Notice the transformation yields a term very similar to the additional term we had before. This behavior is why this the electromagnetic potential was chosen as the gauge field. With all the transformations established, we can perform the gauge transformation on Equation 2.5. After a bit of algebra, we will find that $\mathcal{L}'_{Dirac+EM} = \mathcal{L}_{Dirac+EM}$. By introducing the electromagnetic potential into the Lagrangian, we have achieved gauge invariance! Not only that, but this addition yields the term $q\bar{\psi}\gamma^\mu A_\mu\psi$. This may not seem particularly exciting, but this term describes interactions between the charged fermions of the Dirac equation and the electromagnetic field. To complete the incorporation of electromagnetism into the Dirac Lagrangian, we just add a term corresponding to the propagation of A_μ. With this, the Lagrangian becomes

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\gamma^\mu\partial_\mu - q\gamma^\mu A_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{2.7}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field tensor. By design, the additional $F^{\mu\nu}F_{\mu\nu}$ term maintains the all-important gauge invariance. Another detail worth noting

is the lack of an explicit mass term for the gauge field, which would have the form $\frac{1}{2}m^2 A^\mu A_\mu$. A quick check, however, will show that this term violates gauge invariance. Therefore, the photon field is required to be massless. This feature is typical of gauge theories, and we will see later how it is reconciled in the case of the W and Z.

To summarize, in our quest to make the Dirac Lagrangian gauge invariant, we added the gauge field A_μ. In doing so, we introduced interactions between the charged fermions of the Dirac equation and the electromagnetic field! The quantum field theory based on Equation 2.7 is called *Quantum Electrodynamics* (QED). Though it was developed to describe the interaction of electrons with the electromagnetic field, it works just as well for all charged fermions: both leptons and quarks. Many experiments have verified the predictions made by QED. The crown jewel of these is the measurement of the magnetic moment of the electron, which agrees with the QED prediction up to 10 significant digits.

Though it would later be superseded by electroweak theory, the success of QED set an important precedent. It established quantum field theory as the best way to describe fundamental particles and their interactions. The same methodology used to construct QED was subsequently applied to the other known interactions, and thus the construction of the standard model began.

2.3 Symmetries of the standard model

Now that we have seen how requiring U(1) gauge invariance introduces interactions between charged fermions and photons, we will see where the other interactions of the standard model come from. All of these interactions will arise in a similar fashion to those in QED. In fact, all interactions in particle physics are derived from gauge symmetries. Let's break down these symmetries piece by piece. The standard model Lagrangian can be represented by a product of its three symmetry groups—and thus separated by

interaction—like so:

$$\mathcal{L}_{SM} = \underbrace{\mathrm{SU}(3)_C}_{\text{QCD}} \times \underbrace{\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y}_{\text{EW}} . \tag{2.8}$$

Here, SU(n) refers to the special unitary group of degree n, which is the Lie group of $n \times n$ matrices with a determinant of 1. Additionally, the label "EW" represent the *electroweak* interaction, which is the unification of the electromagnetic and weak interactions. We will discuss this in further detail later. First, let's start by looking at the term responsible for quantum chromodynamics (QCD).

2.3.1 Quantum chromodynamics

The strong interaction is responsible for the interactions between quarks and gluons. As its name implies, it is the strongest interaction present in the standard model, and it is responsible for the formation of hadrons. Unlike QED, whose charge is represented as a scalar, the charge of the strong interaction is represented as a vector of three numbers. Referred to as *color charge*, quarks can have a charge of red, blue, or green, while antiquarks carry the corresponding anticolor charge.

The formulation of the strong interaction, QCD, is a non-abelian (non-commutative) gauge theory with the symmetry group $\mathrm{SU}(3)_C$ (C for color charge). Whereas the U(1) group for QED only had a single generator, $\mathrm{SU}(3)_C$ has 8 generators, each one being a traceless 3×3 hermitian matrix. Therefore, rather than acting on a single bispinor, the generators will act on a vector of *three* bispinors (one for each color). This gives the gauge transformation the form

$$\Psi(x) = \begin{pmatrix} \psi_R(x) \\ \psi_G(x) \\ \psi_B(x) \end{pmatrix} \rightarrow \Psi'(x) = e^{ig\alpha_a(x)T_a} \begin{pmatrix} \psi_R(x) \\ \psi_G(x) \\ \psi_B(x) \end{pmatrix} , \tag{2.9}$$

where T_a are the generators (typically represented by the Gell-Mann matrices), $\alpha_a(x)$

are the transformation parameters, with the sum over the a index going from 1 to 8. As in QED, to make the Lagrangian invariant under this transformation, gauge fields will need to be introduced. In this case, however, because there are 8 generators for SU(3), 8 fields are required to ensure gauge invariance. As before, this is done by replacing the derivative with the gauge covariant derivative, defined as

$$D_\mu = \partial_\mu I_3 - igG^a_\mu(x)T_a, \tag{2.10}$$

where I_3 is the 3×3 identity matrix and $G^\mu_a(x)$ are the 8 gauge fields. As in QED, these gauge fields correspond to the force-carrying particle of the interaction, meaning there are 8 gluons (each with a different combination of color and anticolor). With the covariant derivative established, we find that the Lagrangian is invariant under the gauge transformation of Eq. (2.9) if the gluon fields transform as

$$G^a_\mu(x) \to G^{a\prime}_\mu(x) = G^a_\mu(x) - \partial_\mu\alpha_a(x) - gf_{abc}\alpha_b(x)G^c_\mu(x), \tag{2.11}$$

where f_{abc} are the structure constants of the gauge group. With gauge invariance preserved, we can write the Lagrangian of QCD as

$$\mathcal{L}_{\text{QCD}} = \bar{\Psi}_q(i\gamma^\mu\partial_\mu I_3 - g\gamma^\mu G^a_\mu T_a)\Psi_q - \frac{1}{4}\mathcal{G}^c_{\mu\nu}\mathcal{G}^{\mu\nu}_c, \tag{2.12}$$

where $\mathcal{G}^c_{\mu\nu} = \partial_\mu G^c_\nu - \partial_\nu G^c_\mu - gf_{abc}G^b_\mu G^c_\nu$ is the gluon field strength tensor, q is summed over quark types, and c is summed over color. We now have a Lagrangian with interactions between quarks and gluons! This expression looks similar to that which we found for QED, with a few notable exceptions. First and foremost, the terms involved are 3×3 matrices rather than complex numbers. Another important difference is the additional term present in the field tensor: $gf_{abc}G^b_\mu G^c_\nu$. This term gives interactions between the gluon fields themselves, thus implying that gluons carry color charge. This self-coupling behavior is not present in QED, in which the photon carries no electric charge.

2.3.2 Electroweak interaction

For much of their history, the electromagnetic and weak interactions were treated as separate entities, each governed by their own gauge theory and symmetries. As we saw earlier, the electromagnetic force was the first of the fundamental forces to be described successfully as a quantum field theory. This was in no small part due to the simplicity of QED's U(1) gauge group and the wealth of previous work done on understanding electromagnetic interactions. The weak interaction, on the other hand, was not nearly as well understood. It has many properties that hadn't been observed in either QED or QCD, such as: the ability to change quark flavor, violation of particle-antiparticle symmetry and parity (as well as their combination CP), and mediating bosons that have mass. Despite all these apparent differences, however, it was shown that they are, in fact, different aspects of the same force. This combination, known as *electroweak unification*, was one of the most significant breakthroughs in the formulation of the standard model, earning those who developed it the Nobel Prize.

The Lagrangian of the electroweak interaction is based on the product of two symmetry groups, SU(2) and U(1), meaning it will be invariant under the transformations of either group. The U(1) symmetry group yields a similar structure to that of QED, with a gauge field and an associated conserved quantity Y called *weak hypercharge* (this symmetry group is often referred to as $\mathrm{U}(1)_Y$). The SU(2) symmetry group gives three gauge fields and a corresponding conserved quantity known as *weak isospin*. Weak isospin has a structure analogous to particle spin, with T and T^3 denoting total isospin and its third component.

A key feature of the electroweak interaction is that it is *chiral*. This means that

fermion fields are split into left- and right-handed components, defined as

$$\psi_L = \frac{1}{2}(1-\gamma^5)\psi \tag{2.13}$$

$$\psi_R = \frac{1}{2}(1+\gamma^5)\psi, \tag{2.14}$$

where $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$. This separation is not superficial; left- and right-handed components have very different interactions. Namely, the weak interaction only couples to left-handed particles and right-handed antiparticles. Because of this, left-handed fermions are joined together in pairs to form weak isospin doublets with $(T, T^3) = (1/2, \pm 1/2)$, containing a neutrino and a lepton or an up-type quark and its corresponding down-type. As they only interact electromagnetically, right-handed fermions remain in weak isospin singlet states with $T = 0$. The distinction between left- and right-handed particles has a profound effect on neutrinos: because they have no electric charge, and the weak interaction only couples to the left-handed component of particles, there is no production mechanism for right-handed neutrinos. Therefore, right-handed neutrinos are not allowed in the standard model.

With this arrangement of fermion fields, the interaction terms in the Lagrangian take the form

$$\mathcal{L}_{\text{int}} = \bar{\Psi}^j_L \gamma^\mu \partial_\mu I_2 \Psi^j_L + \bar{\psi}^k_R \gamma^\mu \partial_\mu \psi^k_R, \tag{2.15}$$

where I_2 is the 2×2 identity matrix; j is summed over the weak isospin doublets (with particles denoting the corresponding bispinor field)

$$\Psi^j_L \in \left\{ \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, \begin{pmatrix} \mathrm{u} \\ \mathrm{d} \end{pmatrix}_L, \begin{pmatrix} \mathrm{c} \\ \mathrm{s} \end{pmatrix}_L, \begin{pmatrix} \mathrm{t} \\ \mathrm{b} \end{pmatrix}_L \right\}; \tag{2.16}$$

and k is summed over the weak isospin singlets

$$\psi^k_R \in \{e_R,\ \mu_R,\ \tau_R,\ \mathrm{u}_R,\ \mathrm{d}_R,\ \mathrm{c}_R,\ \mathrm{s}_R,\ \mathrm{t}_R,\ \mathrm{b}_R\}. \tag{2.17}$$

Now that we've separated the right- and left-handed fermion fields in the Lagrangian,

we can introduce the $\mathrm{U}(1)_Y \times \mathrm{SU}(2)_L$ gauge transformation

$$\Psi_L \to \Psi'_L = e^{i\alpha_j(x)\tau_j + i\chi(x)YI_2}\Psi_L \tag{2.18}$$

$$\psi_R \to \psi'_R = e^{i\chi(x)Y}\psi_R, \tag{2.19}$$

where τ_j are the 2×2 generating matrices of $\mathrm{SU}(2)_L$ (usually represented as the Pauli matrices). As we did with QED and QCD, we will introduce gauge fields in order to keep the Lagrangian invariant under this transformation. Since two group transformations are being performed, we will need to introduce two distinct gauge field terms with different coupling constants. As usual, these fields will be introduced in the gauge covariant derivative. However, because the right- and left-handed fermion fields are transformed differently, a different covariant derivative will be substituted for each. Denoting these as D_μ^L and D_μ^R, they are defined as

$$D_\mu^L = I_2\partial_\mu + ig\tau_a W_\mu^a(x) + iI_2 g' Y B_\mu(x) \tag{2.20}$$

$$D_\mu^R = \partial_\mu + ig' Y B_\mu(x), \tag{2.21}$$

where g and g' are the coupling constants for the W and B fields, respectively. These covariant derivatives introduce four gauge fields: W_μ^1, W_μ^2, W_μ^3, and B_μ. Working out how these fields must transform to achieve invariance in the Lagrangian, we get

$$W_\mu^a(x) \to W_\mu^{a\prime}(x) = W_\mu^a(x) - \partial_\mu\alpha_a(x) - g\epsilon_{abc}\alpha_b(x)W_\mu^c(x) \tag{2.22}$$

$$B_\mu(x) \to B'_\mu(x) = B_\mu(x) - \partial_\mu\beta(x), \tag{2.23}$$

where ϵ_{abc} are the structure constants of $\mathrm{SU}(2)_L$.

Putting this all together and including the field strength tensors, the interaction Lagrangian becomes

$$\mathcal{L}_{\mathrm{EW}} = \bar{\Psi}_L^j\gamma^\mu D_\mu^L\Psi_L^j + \bar{\psi}_R^k\gamma^\mu D_\mu^R\psi_R^k - \frac{1}{4}\mathcal{W}_{\mu\nu}^a\mathcal{W}_a^{\mu\nu} - \frac{1}{4}\mathcal{B}_{\mu\nu}\mathcal{B}^{\mu\nu} \tag{2.24}$$

where the gauge field strength tensors are defined as

$$\mathcal{W}^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - g\epsilon_{abc} W^b_\mu W^c_\nu \tag{2.25}$$

$$\mathcal{B}_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \tag{2.26}$$

Phew! Let's take a step back and review what we've done. In unifying the electromagnetic and weak interactions into one electroweak theory, we required the Lagrangian to be invariant under gauge transformations in the $\mathrm{U}(1)_Y \times \mathrm{SU}(2)_L$ group. To accomplish this, 4 gauge fields were introduced as well as the corresponding conserved quantities of hypercharge and weak isospin. Previously, each gauge field we introduced into the Lagrangian directly corresponded to the force-carrying particles of that interaction. In the electroweak interaction, however, this is not the case. The force-carrying particles of the electroweak interaction correspond to *mixtures* of the underlying gauge particles. The W^+ and W^- fields are linear combinations of the W^1 and W^2 gauge fields, given by

$$W^+_\mu = \frac{1}{\sqrt{2}}(W^1_\mu - iW^2_\mu) \tag{2.27}$$

$$W^-_\mu = \frac{1}{\sqrt{2}}(W^1_\mu + iW^2_\mu). \tag{2.28}$$

The γ and Z fields are related to the B and W^3 fields via the rotation

$$\begin{pmatrix} \gamma \\ \mathrm{Z} \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B \\ W^3 \end{pmatrix} \tag{2.29}$$

where $\theta_W = \arctan(g'/g)$ is known as the Weinberg angle.

Additionally, the conserved quantities of the electroweak interaction combine to give observed properties. Namely, the electric charge Q is a function of the hypercharge and the third component of weak isospin, given by

$$Q = T^3 + \frac{Y}{2}. \tag{2.30}$$

So far, we have introduced gauge fields into the Lagrangian to make it invariant under the transformations of the $\mathrm{SU}(3)_C$ and $\mathrm{U}(1)_Y \times \mathrm{SU}(2)_L$ gauge groups. Table 2.1 lists the charges of every particle under each gauge group, as well as their masses. Though these

fields explain the couplings of particles, they do not explain another fundamental particle feature: mass. In their current form, the fields of Equations 2.12 and 2.24 are all massless. Explicit inclusion of a mass term such as $\frac{1}{2}m^2A^\mu A_\mu$ would break the gauge invariance these fields were meant to fix. In the following section, we will see that an additional field must be incorporated to generate mass.

2.3.3 The Higgs mechanism

The origin of gauge boson mass was the cause of much anxiety in the development of the standard model. The very symmetries the theory relied on to generate interactions also required the gauge bosons mediating these interactions to be massless (as we saw in the case of QED, an explicit mass term would violate gauge invariance). From experimental results, it was known that the W and Z had mass, quite a lot of it, in fact. The standard model was at a crossroads: either gauge invariance was the wrong approach, or the particles were acquiring mass through some unknown mechanism. This unknown mechanism would have to be unusual, as it needed to get non-symmetric results from a symmetric Lagrangian.

Eventually, it was discovered that this behavior was achievable through a process known as *spontaneous symmetry breaking* (SSB). SSB describes a system in which the Lagrangian is invariant under a symmetry group, but the lowest-energy vacuum state is not. We will look at a complex scalar field exhibiting this behavior, and examine the consequences of coupling it to our electroweak fields.

A complex scalar field has a Lagrangian of the form

$$\mathcal{L}_{Scalar} = (\partial_\mu\phi)^\dagger(\partial^\mu\phi) - V(\phi). \tag{2.31}$$

To allow for electroweak interactions, the field ϕ will be a weak isospin doublet of two

Table 2.1: Particles of the standard model and their properties. Antiparticles are not listed in this table, but would have identical mass and opposite charges as their corresponding particle. In cases where the left- and right-handed particles differ, the values for the left-handed version are shown without parentheses, while the values for the right-handed particle are shown with parenthesis. Particle masses m are those reported by Reference [11] (although the standard model assumes neutrinos are massless, they are known to have a small non-zero mass). J is the spin of each particle. The $\mathrm{SU}(3)_C$ representations indicate the number of possible color charges for each particle and their transformation properties, with **8** denoting the adjoint representation, **3** the fundamental representation, and **1** the trivial representation. T_3 is the third component of weak isospin and Y the weak hypercharge, the two of which are summed to find the conserved electric charge $Q = T_3 + \frac{Y}{2}$.

Particle	mc^2	J	$\mathrm{SU}(3)_C$ Rep.	T_3	Y	$Q = T_3 + \frac{Y}{2}$
g	0	1	**8**	0	0	0
γ	0	1	**1**	0	0	0
$\mathrm{W}^{\pm}$	80.385 GeV	1	**1**	± 1	0	± 1
Z	91.187 GeV	1	**1**	0	0	0
H	125.09 GeV	0	**1**	0	0	0
e	511.00 keV	$\frac{1}{2}$	**1**	$-\frac{1}{2}$ (0)	-1 (-2)	-1
μ	105.66 MeV	$\frac{1}{2}$	**1**	$-\frac{1}{2}$ (0)	-1 (-2)	-1
τ	1.7769 GeV	$\frac{1}{2}$	**1**	$-\frac{1}{2}$ (0)	-1 (-2)	-1
$\nu_{\mathrm{e},\mu,\tau}$	< 2 eV	$\frac{1}{2}$	**1**	$\frac{1}{2}$	-1	0
u	2.2 MeV	$\frac{1}{2}$	**3**	$\frac{1}{2}$ (0)	$\frac{1}{3}$ $\left(\frac{4}{3}\right)$	$\frac{2}{3}$
c	1.27 GeV	$\frac{1}{2}$	**3**	$\frac{1}{2}$ (0)	$\frac{1}{3}$ $\left(\frac{4}{3}\right)$	$\frac{2}{3}$
t	173.21 GeV	$\frac{1}{2}$	**3**	$\frac{1}{2}$ (0)	$\frac{1}{3}$ $\left(\frac{4}{3}\right)$	$\frac{2}{3}$
d	4.7 MeV	$\frac{1}{2}$	**3**	$-\frac{1}{2}$ (0)	$\frac{1}{3}$ $\left(-\frac{2}{3}\right)$	$-\frac{1}{3}$
s	96 MeV	$\frac{1}{2}$	**3**	$-\frac{1}{2}$ (0)	$\frac{1}{3}$ $\left(-\frac{2}{3}\right)$	$-\frac{1}{3}$
b	4.18 GeV	$\frac{1}{2}$	**3**	$-\frac{1}{2}$ (0)	$\frac{1}{3}$ $\left(-\frac{2}{3}\right)$	$-\frac{1}{3}$

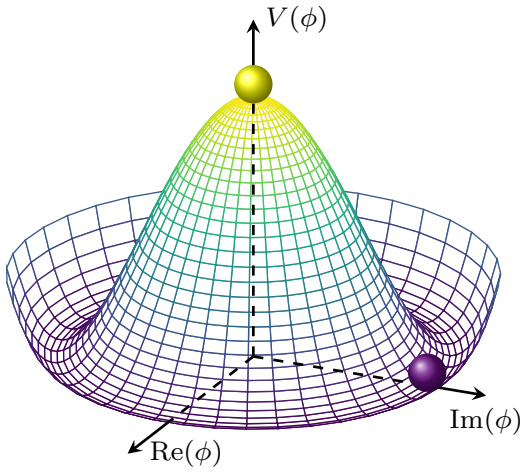


Figure 2.2: The Higgs potential $V(\phi)$, which satisfies gauge invariance and is symmetric about its center (yellow sphere). The symmetry is spontaneously broken when the system enters its ground state (purple sphere). Figure from reference [1].

complex scalar fields

$$\phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1(x) + i\phi_2(x) \\ \phi_3(x) + i\phi_4(x) \end{pmatrix}. \tag{2.32}$$

Before introducing electroweak interactions into the Lagrangian, let's examine the potential function. Illustrated in Figure 2.2, $V(\phi)$ is assumed to have the form

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2. \tag{2.33}$$

Depending on the sign of μ^2 and λ, this potential will behave differently. In the case of $\mu^2 > 0$ and $\lambda > 0$, the minimum of the potential occurs at a nonzero ϕ. This minimum is referred to as the vacuum expectation value (VEV). The nonzero value of the VEV causes the symmetry present in the potential function to be broken when the system spontaneously decays into the lower-energy ground state. Looking at the shape of the potential in Figure 2.2, we see that the rotational symmetry about the origin is no longer present in the ground state.

To examine the consequences of the spontaneous symmetry breaking, let's expand ϕ about the ground state, choosing a basis where the VEV is real and in the ϕ^0 component:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{2.34}$$

where $v = \sqrt{\frac{\mu^2}{\lambda}}$ and the expectation value of $h(x)$ is zero. Rewriting the potential as a function of h and defining the mass of the Higgs boson $m_H = \sqrt{2}\mu$, we get

$$V(h) = \frac{m_H^2 v^2}{8} + \frac{m_H^2}{2}h^2 + \frac{m_H^2}{8v^2}h^3 + \frac{m_H^2}{8v^2}h^4. \tag{2.35}$$

This parameterization of the potential gives an explicit mass term and two self-interaction terms (the constant term can be ignored).

Now that we've looked at the potential, let's see how this scalar field interacts with the electroweak gauge fields. To couple this field to our electroweak theory, we substitute in the gauge covariant derivative

$$\partial_\mu \rightarrow D_\mu = I_2\partial_\mu + ig\tau_a W_\mu^a(x) + iI_2 g' Y B_\mu(x) \tag{2.36}$$

into Equation 2.31. Using the new parameterization of $\phi(x)$ given in Equation 2.34, and rewriting the resulting expression in terms of the physical fields W^+, W^-, and Z , we get

$$\mathcal{L}_{\text{Higgs}} = \frac{v^2 g^2}{8}(\mathrm{W}^+)^2 + \frac{v^2 g^2}{8}(\mathrm{W}^-)^2 + \frac{v^2(g^2 + (g')^2)}{8}\mathrm{Z}^2 + \dots , \tag{2.37}$$

where we are only showing the diagonal field terms. From these expressions, we can see that the W^+, W^-, and Z fields have acquired explicit mass terms! These masses are

$$m_{\mathrm{Z}} = \frac{v}{2}\sqrt{g^2 + (g')^2} \qquad m_{\mathrm{W}} = \frac{vg}{2} = m_{\mathrm{Z}} \cos\theta_W. \tag{2.38}$$

Not only do the W and Z have mass, but the masses are related by the Weinberg angle.

When we introduced ϕ as a weak isospin doublet in Equation 2.32, it came with 4 degrees of freedom: $\phi_1(x)$, $\phi_2(x)$, $\phi_3(x)$, $\phi_4(x)$. After the spontaneous breaking of the symmetry and our choice of expanding ϕ about a real and neutral VEV, it appears we only have one degree of freedom: $h(x)$. Thanks to the mass terms for the W^+, W^-, and

Z fields, however, this is not the case. Once these fields acquired mass, they each gained a longitudinal polarization component that was previously absent. Thus, three of the four scalar fields we introduced were transmuted into the longitudinal component of the W and Z fields, maintaining the four degrees of freedom added.

In summary, we introduced a complex scalar field into our electroweak formulation, which led to the generation of mass for the three weak gauge bosons and the remaining scalar field. In doing so, we introduced two additional parameters, μ and λ. Including g and g', there are now four free parameters in our electroweak model. Though this mechanism provides mass for the bosons, it does not do so for the fermions.

Fermion masses arise from terms in the Lagrangian that couple the left- and right-handed component of a fermion field together. Since the left-handed components are in weak isospin doublets and the right-handed components are singlets, however, the complex scalar field doublet is required to ensure the mass term transforms as a singlet under $\mathrm{SU}(2)_L$ transformations. This gives mass terms with the form:

$$\mathcal{L}_{\text{Yukawa}} = y_{ij}(\bar{\Psi}^i_L \phi \psi^j_R + \phi^* \Psi^i_L \bar{\psi}^j_R). \tag{2.39}$$

We can then substitute in our form of ϕ from Equation 2.34, and expand the expression in terms of the fermion fields. Focusing on the charged lepton terms, we get

$$\mathcal{L}_{\text{Yukawa}} = \frac{\lambda_i v}{\sqrt{2}}(\bar{e}^i_L e^i_R + \bar{e}^i_R e^i_L)(1 + \frac{h}{v}), \tag{2.40}$$

where λ_i is the Yukawa coupling term and i is summed over the three generations. We now have a term of the form $m\bar{\psi}\psi$, and can identify the mass $m_i = \frac{1}{\sqrt{2}} v\lambda_i$. The additional term involving h gives couplings between the charged leptons and the Higgs field. Using the scalar field to couple the left- and right-handed components of fermion fields gave us charged lepton mass terms which will transform properly under $\mathrm{SU}(2)_L$.

Quarks follow the same paradigm as charged leptons, but there is a complication. To understand this, we will group the quark fields into up-type and down-type vectors, and

write the Lagrangian as

$$\mathcal{L}_{\text{Yukawa}} = \frac{v}{\sqrt{2}}(\bar{\mathbf{u}}_L \mathcal{M} \mathbf{u}_R + \bar{\mathbf{d}}_L \mathcal{N} \mathbf{d}_R + \bar{\mathbf{u}}_R \mathcal{M}^\dagger \mathbf{u}_L + \bar{\mathbf{d}}_R \mathcal{N}^\dagger \mathbf{d}_L)(1 + \frac{h}{v}) \tag{2.41}$$

where $\mathbf{u}_{L/R} = (u, c, t)_{L/R}$ and $\mathbf{d}_{L/R} = (d, s, b)_{L/R}$. If we had written the Lagrangian in this form for the charged leptons, the matrix joining the left- and right-handed components would have been diagonal with the lepton masses as its eigenvalues. In the case of the quarks, however, these matrices are *not diagonal*. This means that the mass eigenstates of quarks are not the same as their eigenstates in the weak interaction. As we refer to quarks by their mass eigenstates, we will transform the up- and down-type quarks into their mass bases:

$$\mathbf{u}_L \to U_L^u \mathbf{u}_L \qquad \mathbf{d}'_L \to U_L^d \mathbf{d}_L \tag{2.42}$$

$$\mathbf{u}'_R \to U_R^u \mathbf{u}_R \qquad \mathbf{d}'_R \to U_R^d \mathbf{d}_R, \tag{2.43}$$

where the unitary matrices $U_{L/R}^u$ and $U_{L/R}^d$ are *not the same* (we'll see the effect of this in a moment). These bases represent the "physical" (mass-definite) quark states: up, down, charm, strange, top, and bottom. With these transformations, the terms in Eq. (2.41) reduce to the form seen in the charged lepton case:

$$\mathcal{L}_{\text{Yukawa}} = \frac{v}{\sqrt{2}} \left(\lambda_i^u (\bar{u}_L^i u_R^i + \bar{u}_R^i u_L^i) + \lambda_i^d (\bar{d}_L^i d_R^i + \bar{d}_R^i d_L^i) \right) (1 + \frac{h}{v}), \tag{2.44}$$

where i is summed over the three generations and λ_i^u and λ_i^d are the up- and down-type Yukawa couplings. As they only couple quarks of the same type, the interactions with the gluon, photon, and Z boson are unaffected by the transformation into the mass eigenstates ($U^\dagger U = \mathbb{I}$ for any unitary matrix U). The W boson, however, couples quarks of different types, resulting in interaction terms like $\bar{u}_L (U_L^u)^\dagger U_L^d \bar{d}_L$ in the Lagrangian. Because U_L^u and U_L^d are different, $(U_L^u)^\dagger U_L^d$ is non-diagonal, resulting in non-zero couplings between quarks of different generations. These couplings are represented by the Cabibbo-

Kobayashi-Maskawa (CKM) matrix:

$$V_{\text{CKM}} = (U_L^u)^\dagger U_L^d = \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.9740 & 0.2265 & 0.0036 \\ 0.2264 & 0.9732 & 0.0405 \\ 0.0085 & 0.0398 & 0.9992 \end{pmatrix}. \quad (2.45)$$

The coefficients of this matrix are included when calculating amplitudes for qq'W vertices, and so non-diagonal (cross-generational) couplings are suppressed relative to diagonal (inter-generational) couplings. In summary, charged-current ($W^\pm$) interactions are able to change quark flavor!

This section has summarized the mathematical structure of the standard model in terms of its symmetry groups: $SU(3)_C$, $SU(2)_L$, and $U(1)_Y$. The invariance of the Lagrangian under gauge transformations from these groups introduces the gauge bosons, which mediate interactions between particles. The addition of the Higgs field allows for mass terms to appear in the Lagrangian for all particles except neutrinos. Figure 2.3 summarizes the particles of the standard model and their interactions, grouping them by particle type. Including all the terms we have covered in this section, the full standard model Lagrangian is

$$\begin{aligned} \mathcal{L} = \; & \bar{\Psi}_q(i\gamma^\mu \partial_\mu I_3 - g\gamma^\mu G_\mu^a T_a)\Psi_q - \frac{1}{4}\mathcal{G}_{\mu\nu}^c \mathcal{G}_c^{\mu\nu} \\ & + \bar{\Psi}_L^j \gamma^\mu D_\mu^L \Psi_L^j + \bar{\psi}_R^k \gamma^\mu D_\mu^R \psi_R^k - \frac{1}{4}\mathcal{W}_{\mu\nu}^a \mathcal{W}_a^{\mu\nu} - \frac{1}{4}\mathcal{B}_{\mu\nu}\mathcal{B}^{\mu\nu} \\ & + \frac{\lambda_i}{\sqrt{2}}(\bar{e}_L^i e_R^i + \bar{e}_R^i e_L^i)(v+h) + |D_\mu^L \phi|^2 - V(h) \\ & + \frac{v}{\sqrt{2}}\left(\lambda_i^u(\bar{u}_L^i u_R^i + \bar{u}_R^i u_L^i) + \lambda_i^d(\bar{d}_L^i d_R^i + \bar{d}_R^i d_L^i)\right)(1 + \frac{h}{v}). \end{aligned}$$

2.4 Limitations of the standard model

The standard model has been wildly successful at predicting experimental observations at particle colliders, giving good agreement across a wide range of energies. Despite

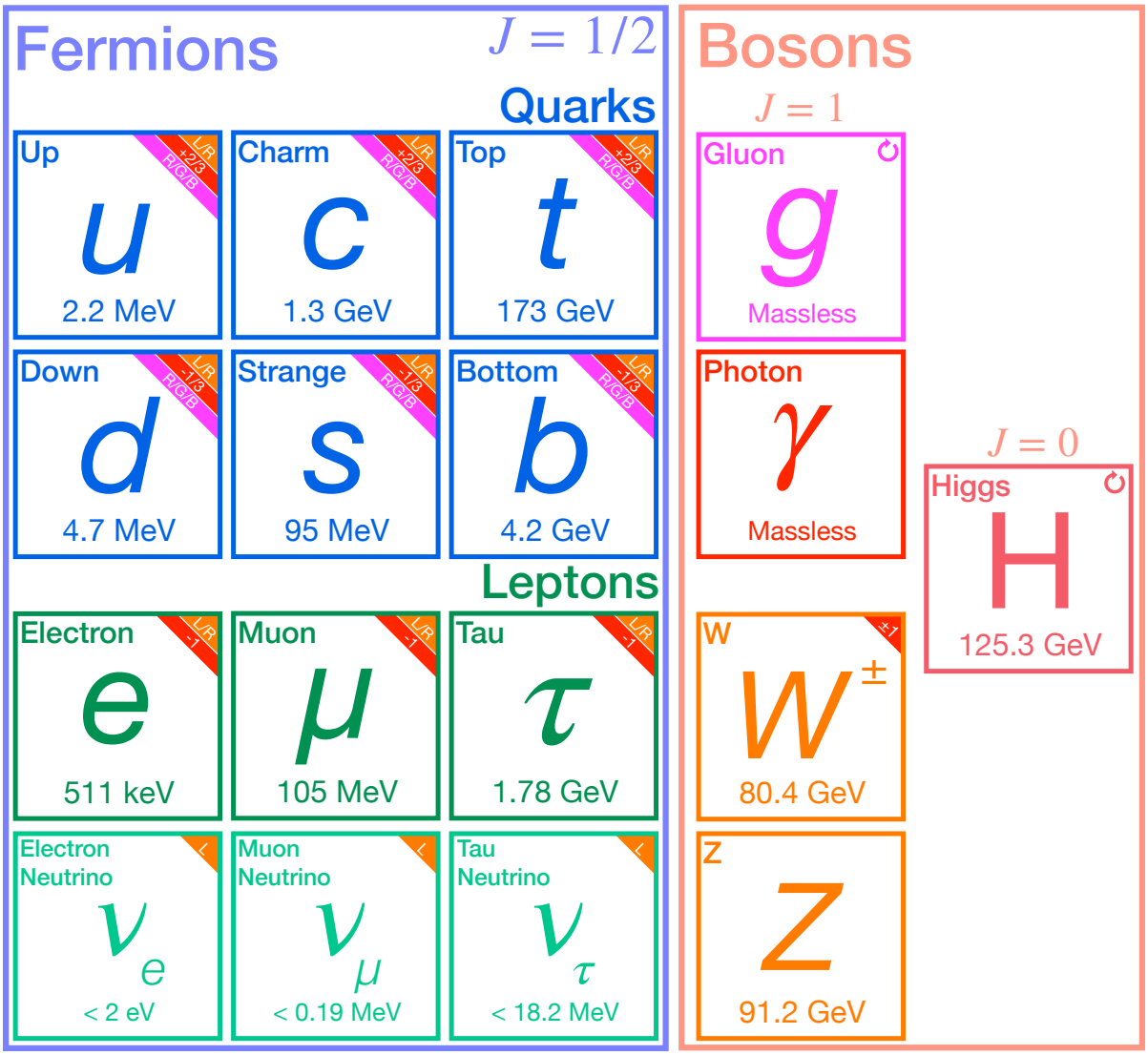


Figure 2.3: Particles of the standard model, separated into fermions and bosons, with the mass given for each. The charges a particle may have under the different interactions are included in the upper right corner, with orange referring to the weak interaction, red the electromagnetic, and magenta the strong. In the case of bosons, self-coupling is represented by a circular arrow.

this agreement, however, there are many results the standard model offers no mechanism for. For example, observations of solar neutrinos by the Super–Kamiokande experiment [24] showed a deficit in the number of muon neutrinos compared to the number of electron neutrinos. Unable to be explained by the standard model, this result was found to be well explained if neutrinos oscillate between flavor states as they propagate through space. Known as *neutrino oscillations* [25], this phenomenon can only occur if neutrinos have mass. Since, as we saw in Section 2.3, neutrinos don't have masses in the standard model, there must be some other phenomenon beyond the standard model through which they acquire mass.

Outside of high-energy experiments, there are also many astrophysical observations that the standard model has no explanation for. Studies of galactic rotation curves [26] and gravitational lensing have shown that galaxies contain much more mass than expected based on direct observation. This additional mass if referred to as *dark matter* due to its lack of interaction with light. Additionally, observations of type Ia supernovae [27] showed the expansion of the universe is accelerating, a result which requires a new form of energy—known as *dark energy*—to permeate through the universe. Together, it is estimated that dark energy and dark matter account for the majority of the mass-energy content of the universe [28]. There is no candidate particle for either of these phenomena in the standard model.

These data demand a new paradigm to be interpreted through. Next chapter, we will introduce a theory that offers solutions for both.

Chapter 3

Supersymmetry

In addition to the experimental observations we saw in Section 2.4 that are not explained by the standard model, there are also theoretical concerns about why the parameters of the standard model have the values they do. One of these problems, known as the hierarchy problem, concerns the mass of the Higgs boson. Just as other particles achieve mass terms in the Lagrangian through their interactions with the Higgs, these interactions also influence the mass of the Higgs itself. Thus, we can express the physical Higgs mass m_{H} as

$$m_{\mathrm{H}}^2 = m_{\mathrm{H},0}^2 + \Delta m_{\mathrm{H}}^2, \tag{3.1}$$

where $m_{\mathrm{H},0}$ is a bare mass parameter and Δm_{H} is a correction term containing contributions from every particle the Higgs couples to. Looking at the contribution from the one-loop corrections of fermions, shown in the left diagram in Figure 3.1, there is a correction to m_{H} of

$$\Delta m_{\mathrm{H},f} = -\frac{|\lambda_f|^2}{8\pi^2}\Lambda^2 + \mathcal{O}(\log\Lambda), \tag{3.2}$$

where Λ is a momentum cutoff parameter used to regulate the loop integral, representing the energy scale at which physics beyond the standard model, such as gravity, become

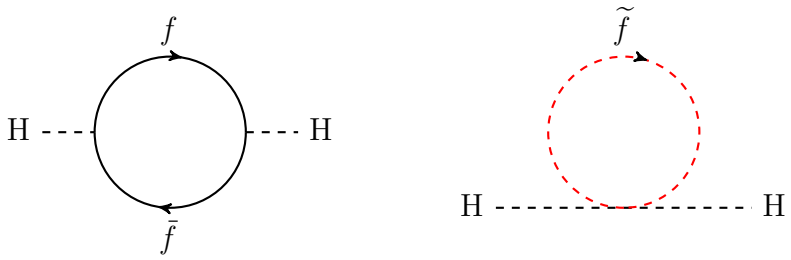


Figure 3.1: Left: correction to the Higgs propagator from a fermion loop. Right: hypothetical correction to the Higgs propagator from the fermion's scalar superpartner. Theoretical particles introduced by supersymmetry are drawn with red lines.

relevant. Considering the energy scale at which the strength of gravity becomes comparable to the forces of the standard model—known as the Planck scale—the corrections to m_{H} would be approximately 1×10^{19}GeV, 17 orders of magnitude larger than it the observed value of m_{H}. These contributions can be canceled out by other correction terms, but this requires the parameters involved to be very fine-tuned, with some parameters being orders of magnitude larger than others. Theories that require such a high level of parameter tuning to agree with observations are referred to as *unnatural*. In the spirit of Occam's razor, naturalness favors models with fewer assumptions or strict constraints on parameter values.

So, how can we restore the naturalness to our theory of nature? Rather than relying on the arbitrary alignment of parameter to miraculously cancel out the $\Delta m_{\mathrm{H},f}$ terms, we can introduce new particles and see what terms they add to m_{H}. If we consider the contribution to the m_{H} caused by one-loop corrections of a new scalar $\tilde{f}$, shown in the right diagram of Figure 3.1, we get

$$\Delta m_{\mathrm{H},S} = \frac{\lambda_{\tilde{f}}}{16\pi^2}\Lambda^2 + \mathcal{O}(\log \Lambda). \tag{3.3}$$

If $\lambda_{\tilde{f}} = |\lambda_f|^2$, this expression only differs in magnitude to Eq. (3.2) by a factor of 2, but has the opposite sign. Therefore, if two new scalars were included with the same coupling factor, then the $\Delta m_{\mathrm{H},f}$ terms in m_{H} would cancel out completely! Wait a minute though,

wouldn't it be considered supremely unnatural to require two bosons with just the right Higgs couplings to cancel out the contributions from a fermion? Absolutely, unless *these fermion-boson pairs are related by a symmetry.*

While the gauge symmetries we studied in Chapter 2 only involved transforming fermions to fermions and bosons to bosons, there also exist transformations between these particle classes. The symmetry associated with such a transformation is known as *supersymmetry* (SUSY). In SUSY, bosonic superpartners are assigned to all the fermions in the standard model and fermionic superpartners are assigned to all the bosons. For each fermionic/bosonic degree of freedom in the standard model particles, there is a corresponding bosonic/fermionic degree of freedom in the superpartners. The new set of supersymmetryic particles introduced with SUSY are sometimes referred to as *sparticles.*

This mapping between degrees of freedom is summarized in Table 3.1, showing the gauge eigenstates of the standard model particles and their corresponding superpartners.

The simplest implementation of SUSY into the standard model, in which the number of SUSY generators has the minimal value ($N = 1$), is known as the Minimal Supersymmetric Standard Model (MSSM) [29]. In the MSSM, all the particles of the standard model we've seen so far are present, with a few additions (outside the new superpartners that is). The structure of supersymmetry requires the inclusion of two complex Higgs doublets to give Yukawa coupling terms that behave well under the supersymmetric transformation. Therefore, rather than the four degrees of freedom that accompanied the Higgs mechanism discussed in Section 2.3.3—three of which were used to give mass to the $W^{\pm}$ and Z bosons—we will instead have 8 degrees of freedom from the Higgs sector. These will be expressed as two electroweak doublet complex scalar fields:

$$H_u = (H_u^+ \; H_u^0) \tag{3.4}$$

$$H_d = (H_d^0 \; H_d^-), \tag{3.5}$$

where the physical Higgs is a linear combination of H_u^0 and H_d^0. H_u and H_d will be mapped to doublets of spin-1/2 fields, $\widetilde{H}_u$ and $\widetilde{H}_d$. Boson superpartners are named by appending *-ino* to the end of the standard model particle name, so $\widetilde{H}_u$ and $\widetilde{H}_d$ are called *higgsinos*.

The mapping for fermions behaves similarly to those for the Higgs doublets, but in reverse. The left-handed doublets transform into doublets of complex scalar fields, and the right-handed singlets transform into complex scalar singlets. The superpartners resulting from this mapping are named by appending *s* (for scalar) to the name of the standard model particles; thus, leptons and quarks are mapped to *sleptons* and *squarks*.

The vector bosons in the standard model are mapped to spin 1/2 superpartners. As we have represented the Higgs doublets in their pre-SSB form, so too shall we represent the electroweak bosons in their massless $SU(2)_L \times U(1)_Y$ eigenstates. The $W^{\pm}$, W^0, and B^0 bosons are mapped to spin-1/2 superpartners $\widetilde{W}^{\pm}$, $\widetilde{W}^0$, and $\widetilde{B}^0$, referred to as *winos* and *bino*. Similarly, the gluon is mapped to the *gluino* $\widetilde{g}$. Similar to the standard model gauge bosons, the supersymmetric *gauginos* are not necessarily mass eigenstates; there can be mixing between the electroweak gauginos and higgsinos.

The pairs of particles and their superpartners are known as *supermultiplets*. The collection of all supermultiplets in the MSSM are shown in Table 3.1. Given all the superpartners seen in the table, an obvious question is: why haven't we discovered any of them? If supersymmetry was an unbroken symmetry, all the superpartners would have the same mass as the corresponding standard model particles. If this were the case, the production rate of superpartners would have been more than sufficient for discovery, whether we can directly measure them or not. Since no such discoveries have been made, SUSY must be a broken symmetry if it is to describe nature.

Though the breaking of supersymmetry allows superpartners to have higher masses than their corresponding particles, there are constraints on how much higher they can be while still giving natural corrections to m_H. For broken SUSY, corrections to m_H from

Table 3.1: Supermultiplets of the Minimal Supersymmetric Standard Model, organized by the standard model particle type, expressed with the pre-SSB gauge fields for simplicity. The degrees of freedom (DOF) associated with each particle are summed for the fermions and bosons to demonstrate they are conserved by the supersymmetric transformation. Complex scalars are represented with $J = 0^*$.

Particle	SM	J	DOF	SUSY	J	DOF
Fermions	$(\mathrm{u_L}\ \mathrm{d_L})$	1/2	4	$(\widetilde{\mathrm{u}}_\mathrm{L}\ \widetilde{\mathrm{d}}_\mathrm{L})$	0^*	4
	$\mathrm{u_R}$	1/2	2	$\widetilde{\mathrm{u}}_\mathrm{R}$	0^*	2
	$\mathrm{d_R}$	1/2	2	$\widetilde{\mathrm{d}}_\mathrm{R}$	0^*	2
	$(\nu_L\ e_L)$	1/2	4	$(\widetilde{\nu}_\mathrm{L}\ \widetilde{\mathrm{e}}_\mathrm{L})$	0^*	4
	e_R	1/2	2	$\widetilde{\mathrm{e}}_\mathrm{R}$	0^*	2
	Total		**14**	Total		**14**
Bosons	$\mathrm{W}^\pm$	1	4	$\widetilde{\mathrm{W}}^\pm$	1/2	4
	W^0	1	2	$\widetilde{\mathrm{W}}^0$	1/2	2
	B^0	1	2	$\widetilde{\mathrm{B}}^0$	1/2	2
	g	1	2	$\widetilde{\mathrm{g}}$	1/2	2
	$(\mathrm{H_u^+}\ \mathrm{H_u^0})$	0^*	4	$(\widetilde{\mathrm{H}}_\mathrm{u}^+\ \widetilde{\mathrm{H}}_\mathrm{u}^0)$	1/2	4
	$(\mathrm{H_d^0}\ \mathrm{H_d^-})$	0^*	4	$(\widetilde{\mathrm{H}}_\mathrm{d}^0\ \widetilde{\mathrm{H}}_\mathrm{d}^-)$	1/2	4
	Total		**18**	Total		**18**

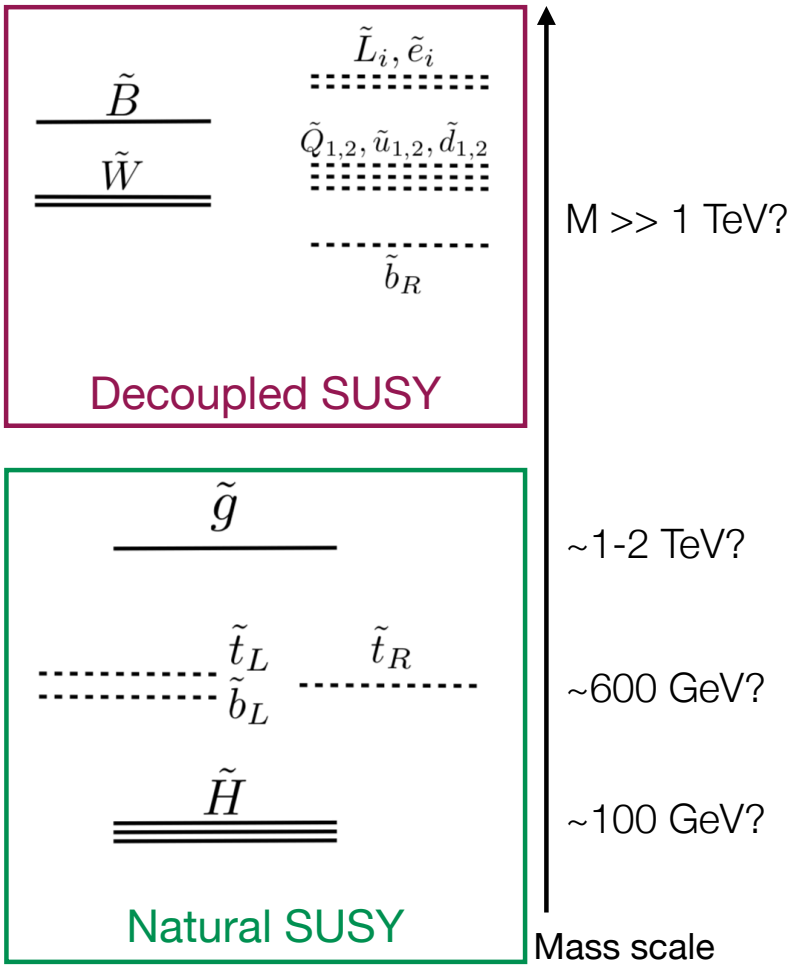


Figure 3.2: Qualitative example of a natural SUSY spectrum from Reference [2].

interactions with a fermion and its superpartner have the leading term

$$\Delta m^2_{\mathrm{H},f\tilde{f}} = \frac{|\lambda_f|^2}{4\pi^2}\left(m^2_{\tilde{f}} - m^2_f\right)\log\left(\frac{\Lambda}{m_{\tilde{f}}}\right). \tag{3.6}$$

Though it now scales as $\log\Lambda$ instead of Λ^2, this correction may be large if $\tilde{f}$ is significantly heavier than f. In addition to the Λ dependence, $\Delta m_{\mathrm{H},f\tilde{f}}$ also scales with the $|\lambda_f|^2$, which is proportional to m^2_f. This dependence makes m_{H} most sensitive to the mass difference between the heaviest fermions and their superpartners, which constrains these superpartners to be the lightest. Thus, natural SUSY models [2] constrain the top and bottom squarks to be the lightest, while the other squark masses may be much larger. An example of a SUSY mass spectrum is shown in Figure 3.2, with the particles constrained to be light referred to as *Natural SUSY* and the particles allowed to be heavy referred to as *Decoupled SUSY*.

With the mass spectrum of a natural SUSY model roughly established, he next question is: which superpartner is the LHC most sensitive to? To answer this, we must consider how frequently each particle will be produced as well as how distinct events with these particles will be compared to standard model background events. Starting with the production rate, Figure 3.3 shows the cross section for pair production of various superpartners as a function of their mass. Like standard model particle production, for a given mass value, superpartners that are strongly produced such as $\widetilde{\mathrm{g}}$ and $\widetilde{\mathrm{q}}$ have the highest cross section. As the cross section decreases exponentially with mass, however, lighter superpartners may provide the most sensitive channel, whether they are strongly produced or not. The LHC is sensitive to a wide range of SUSY scenarios, with most searches targeting the superpartners that are expected, to be the lightest in natural SUSY. The results of SUSY searches are typically presented as exclusion limits on the mass parameters of simplified models. In these models, a production and decay sequence is selected and all parameters but the masses of the searched for particles are fixed. An

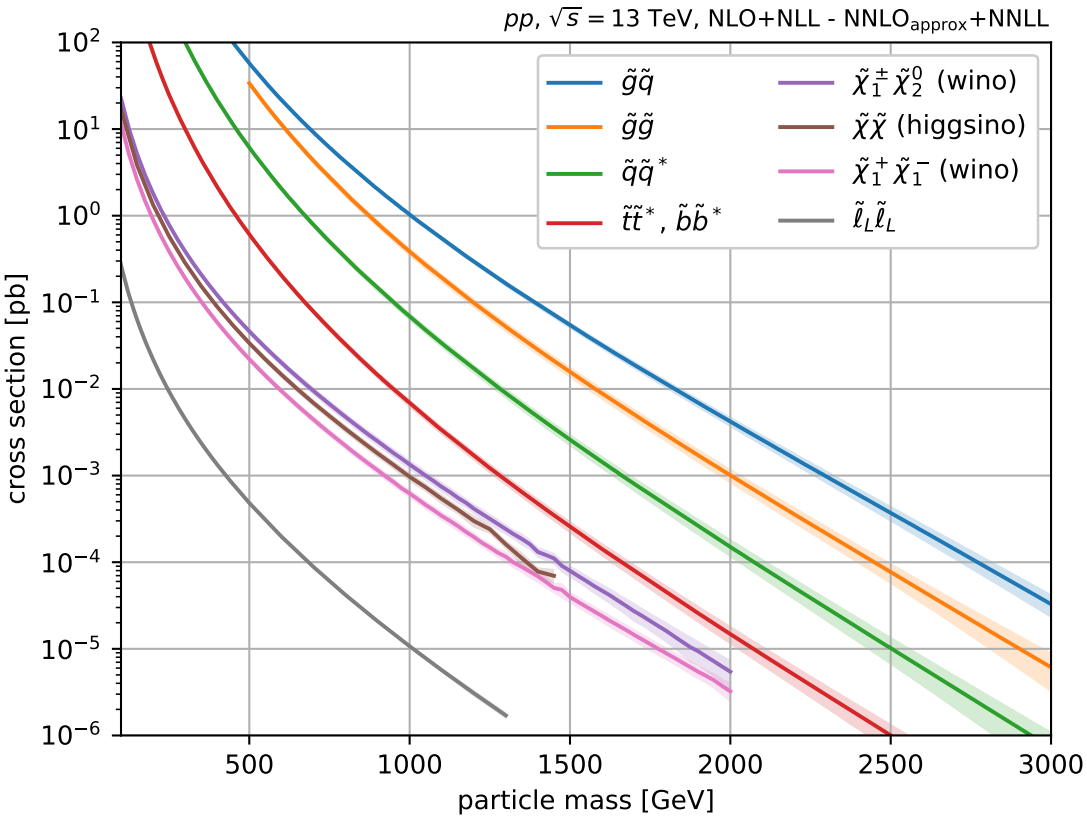
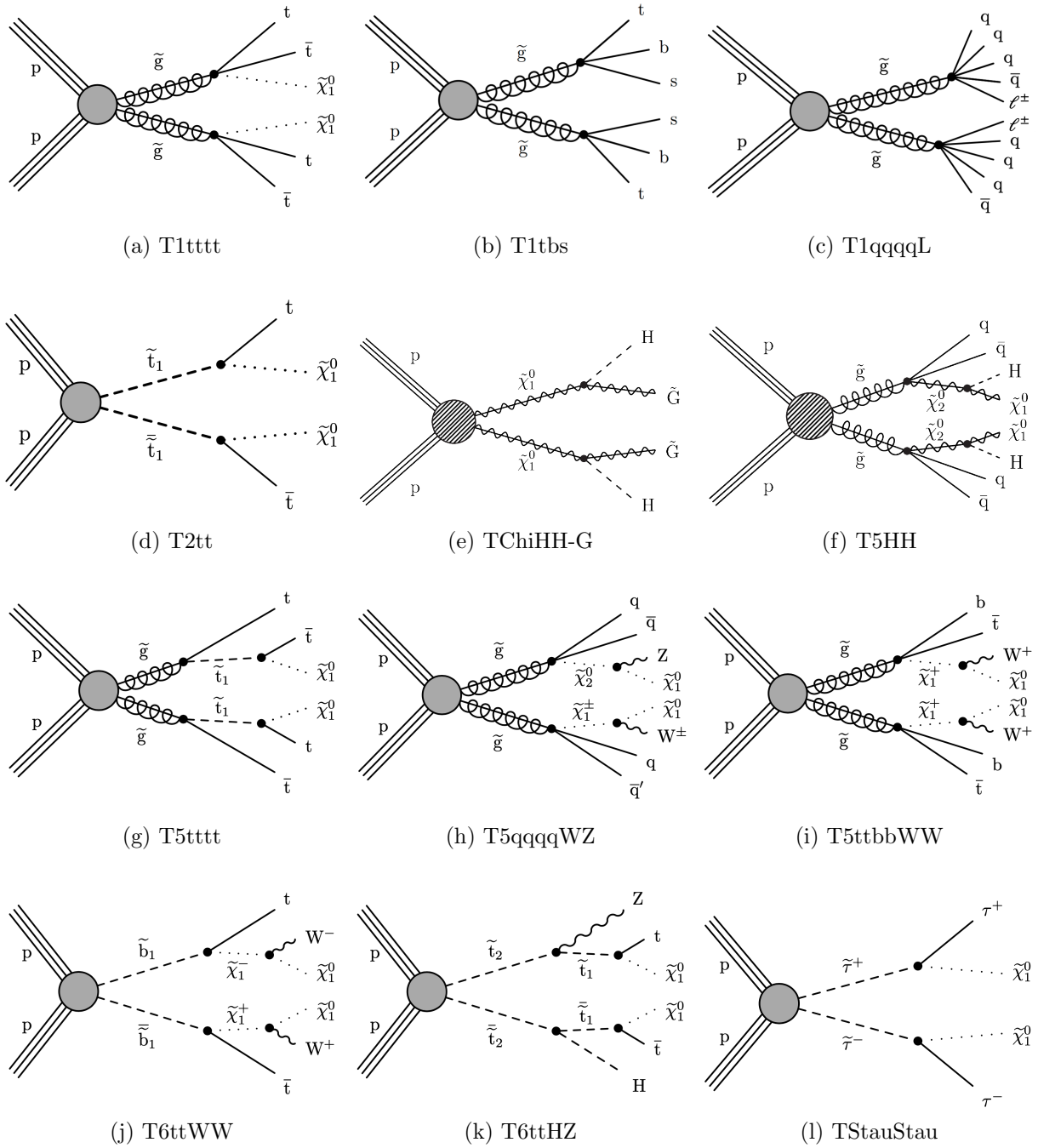


Figure 3.3: Comparison of superpartner production cross sections at a center-of-momentum collision energy of 13 TeV. Squark cross sections assume a ten-fold degeneracy among the squark flavors, not including the top squarks. Figure made with code from Reference [3].

excessive number of examples of simplified models are shown in Figure 3.4, labeled with the model name. The analysis described here targets models with gluino pair production. As they are strongly produced and constrained to be relatively light in natural SUSY models, they make an ideal search candidate. We will explore these models in much more detail in Chapter 7.



(a) T1tttt (b) T1tbs (c) T1qqqqL

(d) T2tt (e) TChiHH-G (f) T5HH

(g) T5tttt (h) T5qqqqWZ (i) T5ttbbWW

(j) T6ttWW (k) T6ttHZ (l) TStauStau

Figure 3.4: Examples of simplified SUSY models.

Chapter 4

Experimental Apparatus

In this chapter, we will introduce the technologies of the Large Hadron Collider (LHC) and the Compact Muon Solenoid (CMS) experiment. We will start by going through the steps required to accelerate beams of protons to an energy of 6.5 TeV. Next, we will see how these beams are controlled and made to collide with one another. Finally, we will review the detector technologies present at CMS that allow us to measure the products of these collisions.

4.1 Large Hadron Collider

The largest particle accelerator in the world, the Large Hadron Collider (LHC) is a 27 km accelerator ring located 100 m underground on the border between France and Switzerland, built and operated by the European Organization for Nuclear Research (CERN). Though the LHC itself was constructed over the course of a decade, it relies on many other accelerators and facilities to support its operation. Fortunately, rather than having to build this pipeline from the underground up for the LHC, CERN was able to utilize infrastructure from past experiments. Ranging from the initial proton extraction

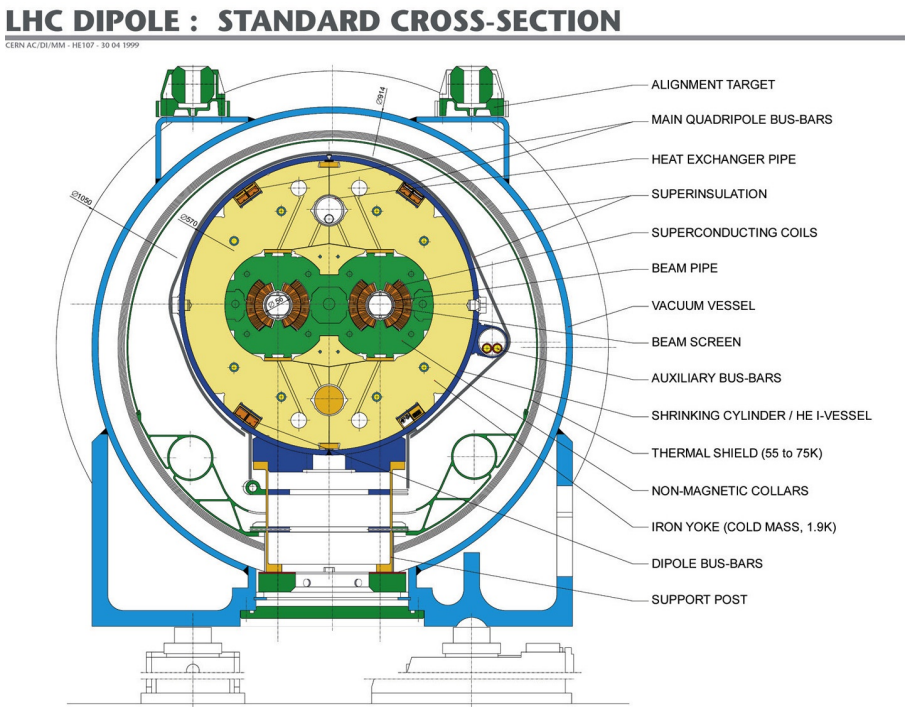


Figure 4.1: Cross-sectional view of a dipole magnet at the LHC. Each dipole contains two sets of superconducting magnet coils surrounding beam pipes, allowing for counter-rotating proton beams. With each dipole measuring 15 m in length, there are 1232 dipoles in total at the LHC. Reference [4]

setup to the 27 km tunnel that houses the LHC itself, these pre-existing facilities greatly reduced the cost and construction time of the LHC.

Accelerating protons to energies of 6.5 TeV requires a myriad of state-of-the-art technologies, most notably thousands of superconducting magnets held in vacuum and cooled to temperatures of $-271\,^{\circ}\mathrm{C}$ with 130 tons of liquid Helium. A cross-section of one of these magnets with its components labeled is shown in Figure 4.1. With an average current density of $400\,\mathrm{A\,mm^{-2}}$, superconducting coils of Niobium-Titanium (NbTi) wire provide a magnetic field of 8 T. While dipole magnets are needed to bend the proton beam into a circular orbit, additional magnets are also required to manage smaller scale beam properties. Quadrupole magnets, for example, are used to maintain the focus of the proton bunches.

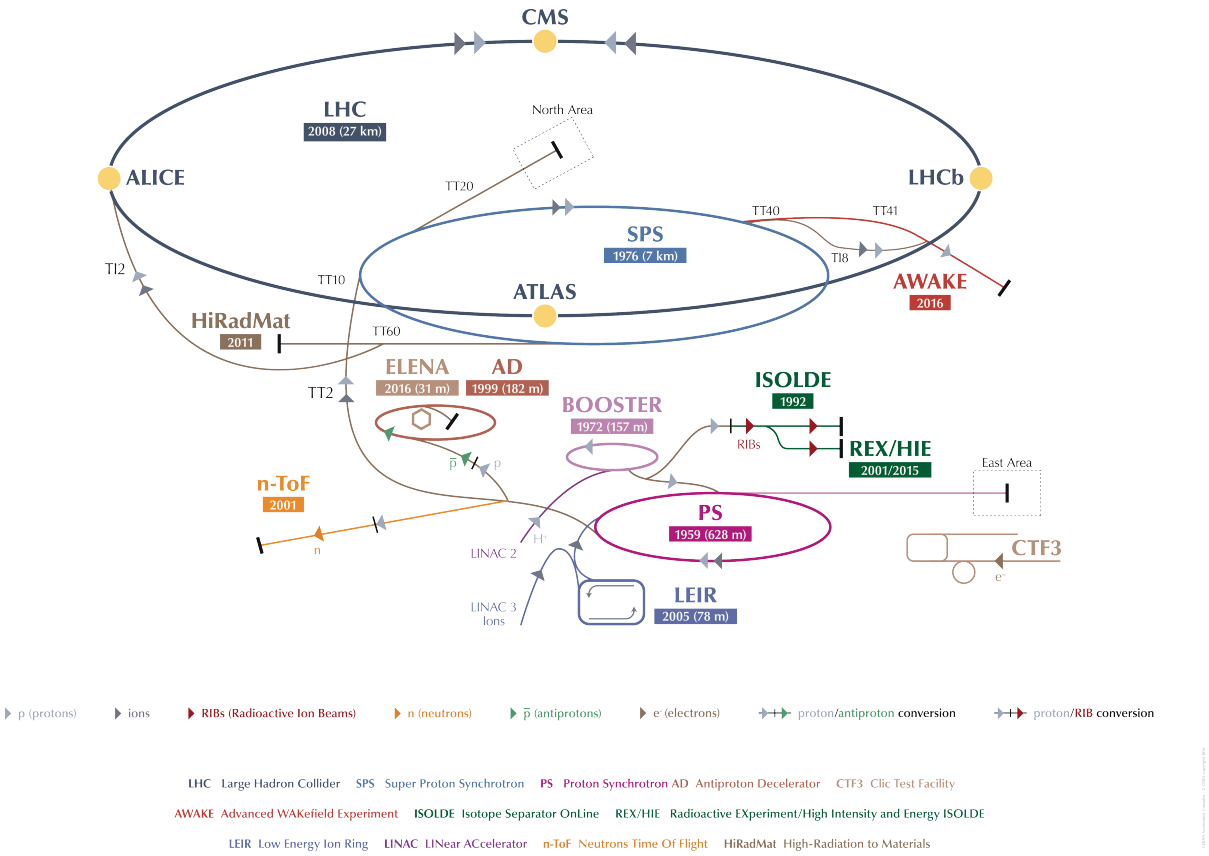


Figure 4.2: Diagram of the various accelerators connected to the LHC. Each accelerator system is labeled, with its first year of operation below. Where applicable, the circumference is also given. Figure from Reference [4]

4.1.1 Life of a proton at the LHC

While the LHC is designed to accelerate protons to extremely high energies, it requires the protons to already be at fairly high energies (hundreds of GeV) to do so. This is accomplished through a series of accelerator facilities, including several past colliders, leading up to the LHC. Figure 4.2 shows the layout of all the accelerators and experiments at CERN, including the types of beams used in each. In this section, we will go through the typical day of a proton at the LHC, from its initial extraction to its final collision energy of 6.5 TeV. A summary of this journey is shown in Table 4.1, including the various stages the protons go through and the energy achieved at each.

Table 4.1: Stages of proton acceleration leading up to the LHC. The length/circumference is given as well as the final proton energy reached at each stage.

Facility	Length [m]	E_{final} [MeV]
Duoplasmatron	-	0.05
LINAC	35	50
Booster Synchrotron	4×157	1400
Proton Synchrotron	628	26000
Super Proton Synchrotron	6900	450000
Large Hadron Collider	26659	6500000

It all starts with Hydrogen gas, which is ionized by the duoplasmatron. The protons are then extracted from the resulting plasma with an electric field. The resulting 50 keV protons are then fed into a 35 m linear accelerator (LINAC) that increases their energy to 50 MeV, followed by a booster synchrotron (PSB) that accelerates them to 1.4 GeV through four 157 m rings. The protons are then injected into the Proton Synchrotron (PS). In addition to accelerating the beam up to 26 GeV, the PS also groups the beam into bunches of 100 billion protons each. With a bunch length of 1.2 m and a separation of 7 m, this bunch structure serves two purposes. First, increasing proton density in the bunches allows for a much higher interaction rate when two bunches cross than would be possible with two diffuse beams. Second, the time in between bunch crossings allows the detectors that measure the collisions to recover, maximizing the amount of data that can be recorded. With the 26 GeV proton beam separated into bunches, it is injected into the Super Proton Synchrotron (SPS). With a radius of over a kilometer, the SPS accelerates the proton beam up to 450 GeV and then injects it into the LHC; as the LHC has two concentric beam lines, this process is performed twice.

While the superconducting magnets mentioned earlier are used to steer and focus the beam, the actual acceleration is accomplished in 16 radio frequency (RF) cavities along the beam line. In each cavity, an external voltage generator induces an electric field inside the cavity, which oscillates at a frequency of 400 MHz. This oscillation is timed so that,

when a proton bunch enters the cavity, the electric field is approaching its maximum strength in the direction the protons are traveling. The effect of this oscillation is two fold. First, as the electric field is pointing in the same direction as the momenta of the protons, the protons will accelerate, gaining about 2 MeV of energy each time they pass through a cavity. Second, the oscillation is tuned such that protons in the front of a bunch experience slightly less acceleration than those in the back, causing the proton bunch to compress over time. So not only do the RF cavities accelerate the protons, but they also maintain bunch density.

Overall, it takes about 15 minutes in the LHC—during which the beams have passed through the RF cavities about 1 million times—for protons to be accelerated from their injection energy of 450 GeV to the final collision energy of 6.5 TeV. The fully accelerated beams are then crossed at four detector points, with bunches cross every 25 ns. Though millions of collisions occur per second, the high number of protons in each bunch allows for this intensity to be maintained for long periods of time, with each beam fill providing collisions for about 10 hours.

With a stored energy of 350 MJ, the proton beams at the LHC contain the equivalent of the kinetic energy of a train traveling at 100 km/h, or more entertainingly, the equivalent of the chemical energy contained in 35 kg of Bircher muesli. Due to the massive amount of energy concentrated in tightly packed bunches, the beam's intensity must be reduced before it can be dumped. To accomplish this, kicker magnets and a 700-meter-long vacuum line are used to sweep the beam in a circular pattern, distributing the deposited energy over a larger area. The beam is then absorbed by a 7 m long cylinder of graphite. Even with the dilution mechanisms in place, a beam with nominal intensity will heat up the graphite absorber by up to 700 °C!

Table 4.2: Overview of typical LHC beam parameters for each year of Run 2 compared to the design values, including the peak luminosity achieved as well as the total integrated luminosity. Values from Reference [12].

Parameter	Design	2016	2017	2018
Energy [TeV]	7	6.5	6.5	6.5
γ	8083	6928	6928	6928
Bunch Spacing [ns]	25	25	25	25
β^* [cm]	55	40	30	25
Half crossing Angle [µrad]	142.5	140	150	130
Orbital frequency f [kHz]	11.25	11.25	11.25	11.25
Emittance ϵ_{n} [µm]	3.75	2.1	1.8	2.0
Max. Protons per Bunch $[1 \times 10^{11}]$	1.15	1.2	1.25	1.1
Max. Bunches per Beam	2808	2076	2556	2556
Max. Stored Energy [MJ]	360	345	320	320
Peak Luminosity $[10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}]$	1.0	1.4	2.06	2.1
Total integrated Luminosity $[\mathrm{fb}^{-1}]$	-	39.7	50.6	66

4.1.2 Luminosity

At an instantaneous luminosity L, a process with cross-section σ_i will occur at a rate of $N_i = L\sigma_i$ events per second. At the LHC, this instantaneous luminosity takes the form

$$L = \frac{N_{\mathrm{p}}^2 N_{\mathrm{bunches}} f \gamma F}{4\pi \epsilon_{\mathrm{n}} \beta^*} \tag{4.1}$$

where N_{p} is the number of protons per bunch, N_{bunches} is the number of bunches per beam, f is the orbit frequency, γ is the relativistic Lorentz factor, ϵ_{n} is the normalized beam emittance, β^* is the beta function evaluated at the collision point, and F is a geometric reduction factor that arises due to nonzero crossing angle of the beams (≤ 1). Changes to these parameters due to throughout Run 2–most often due to technical improvements–resulted in a variable luminosity that increased each year. Typical values for each year, including the peak luminosity achieved, are shown in Table 4.2. To get the total amount of luminosity—referred to as *integrated luminosity*—we simply integrate the instantaneous luminosity over time. The integrated luminosity delivered by the LHC, and recorded by CMS, as a function of time is shown in Figure 4.3, starting from the beginning of Run 1.

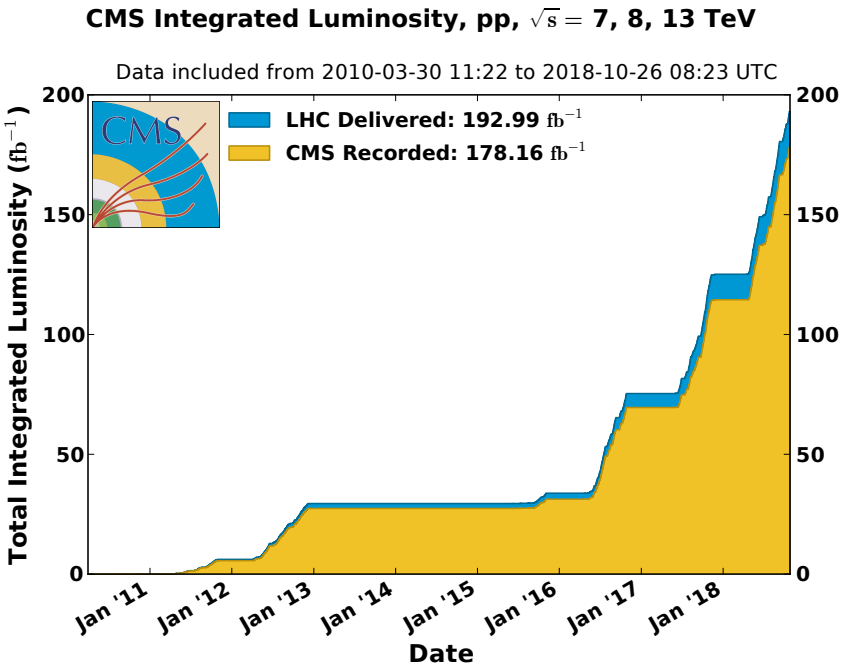


Figure 4.3: Integrated luminosity for proton-proton collisions as a function of time from the beginning of CMS data collection in 2011 to the end of Run 2 in 2018. Figure from Reference [5]

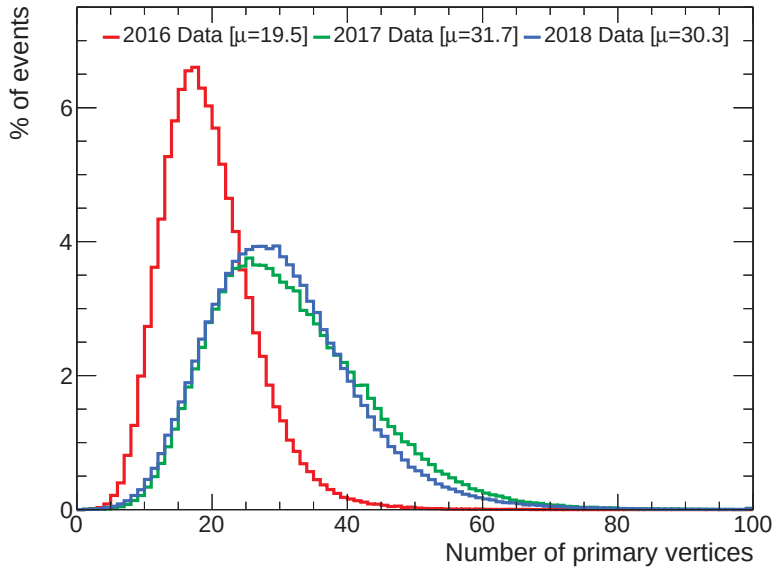


Figure 4.4: Number of vertices per event for events passing the baseline selection of this search, separated by year. There were an average of 20 vertices per event in 2016, and 30 vertices per event in 2017 and 2018.

The extremely high luminosity achieved by the LHC doesn't come without a cost. During each bunch crossing, there are dozens of soft proton-proton interactions, known as *pileup*, that occur. The number of these interactions varies for each bunch crossing, but the average scales with instantaneous luminosity. Represented by the number of primary vertices reconstructed in an event, the distribution of pileup for each year of Run 2 is shown in Figure 4.4. In 2016, there were an average of 20 interactions per bunch crossing. As the luminosity was increased in 2017 and 2018, this average increased to 30, with some events having over 60 primary vertices!

Each pileup interaction produces its own set of final state particles. Therefore, the final state of a given bunch crossing is the superposition of final states produced by all pileup interactions as well as the hard scattering. Resolving the underlying hard scattering from the hundreds of final state particles originating from pileup is a daunting task. In the next section, we will review the detector technologies employed to accomplish this task. Specifically, we will look at one of the general-purpose detectors at the LHC, the Compact Muon Solenoid (CMS).

4.2 Compact Muon Solenoid

The Compact Muon Solenoid (CMS) is a general-purpose detector operating at the LHC. Constructed over the course of a decade starting in 1998, CMS started collecting data in 2010. With a length of 21.6 m and a diameter of 14.6 m, the CMS detector weighs approximately 14,000 tons. To give a sense of the compactness using two other French landmarks, the CMS detector could fit inside the Arc de Triomphe, but weighs twice as much as the Eiffel tower.

In this section, we will discuss how the fundamental quantities used in this analysis are measured. Let's start by introducing what these quantities are. Typically, particle

four-momenta are expressed using energy and the cartesian components of momentum (E, p_x, p_y, p_z). Given the geometry of the CMS detector, however, it is easier to describe the momentum using the three coordinates p_T, η, and ϕ. p_T, also known as transverse momentum, is the component of the momentum perpendicular to the beam axis (we will see later why this is used over the total momentum). The second components, η, is called pseudorapidity. Defining the beam line as the z axis and θ as the angle off this axis, η is defined as:

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]. \tag{4.2}$$

Pseudorapidity is preferred over θ itself because, in the high-energy limit, differences in η are invariant under boosts along the z axis. Lastly, the third component used to describe particle momentum, ϕ, is the typical azimuthal angle used in polar coordinates, and takes values between -π and π. Together with energy, these quantities give the experimental form of the 4-momentum: (E, p_T, η, ϕ). These quantities are reconstructed for each object in an event using the signals measured in the detector. This process will be covered in more detail in the next chapter.

Along with reviewing the key features of each detector system, we will briefly describe the physical phenomena which allow each system to provide the information it does. In doing so, I hope to provide the reader with a basic understanding of modern techniques in high-energy experiments.

Detection of particles involves many methods and technologies, but they all rely on the interaction between particles and matter. When passing through a material, most high-energy particles will lose energy interacting with the medium. The amount of energy lost depends on both the material and the type of particle. By measuring the change in the material caused by the absorption of this energy, we can gain information about the particle that interacted with the detector. This information typically comes in two

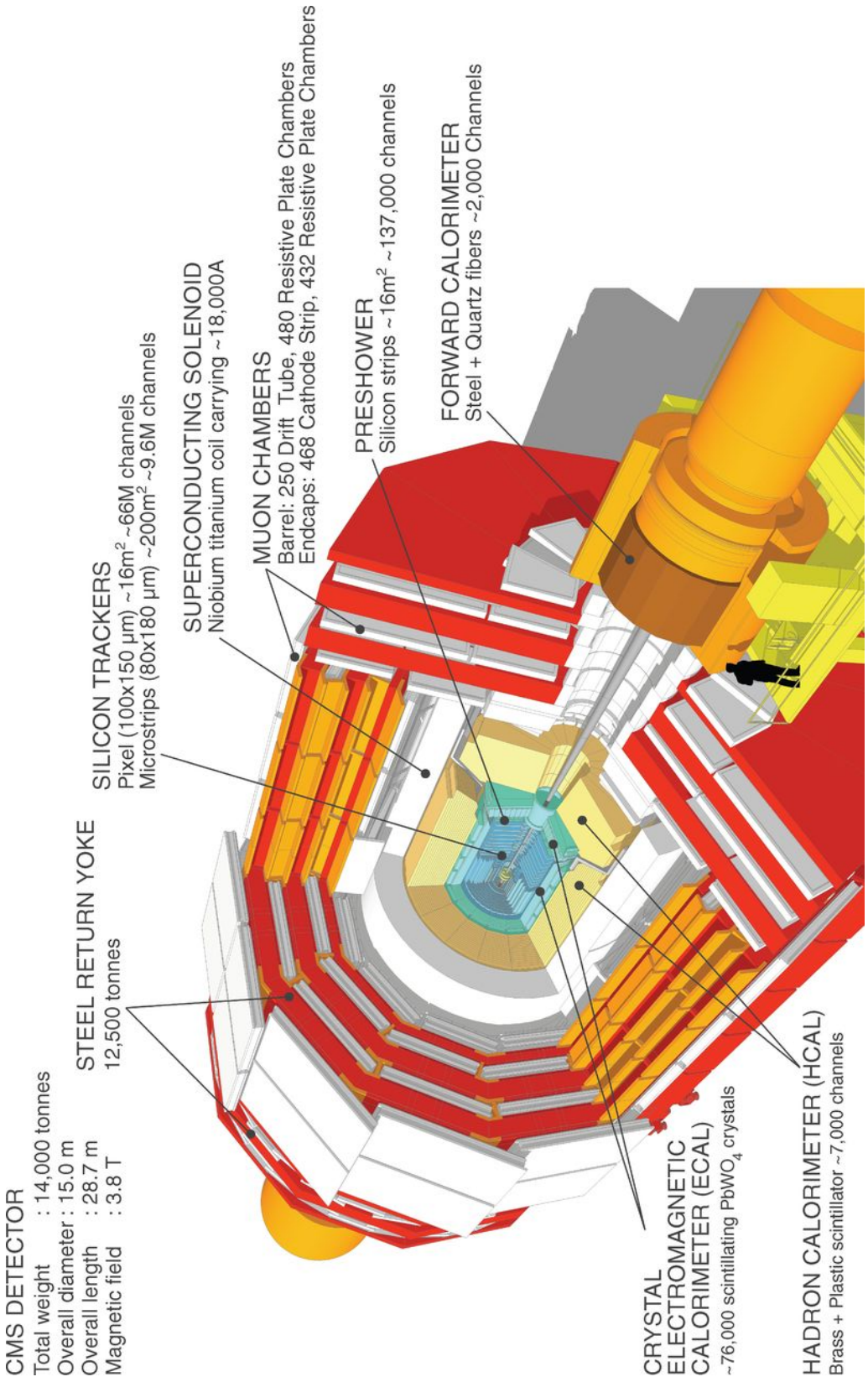


Figure 4.5: Schematic view of the CMS detector showing its main components. Figure from Reference (?).

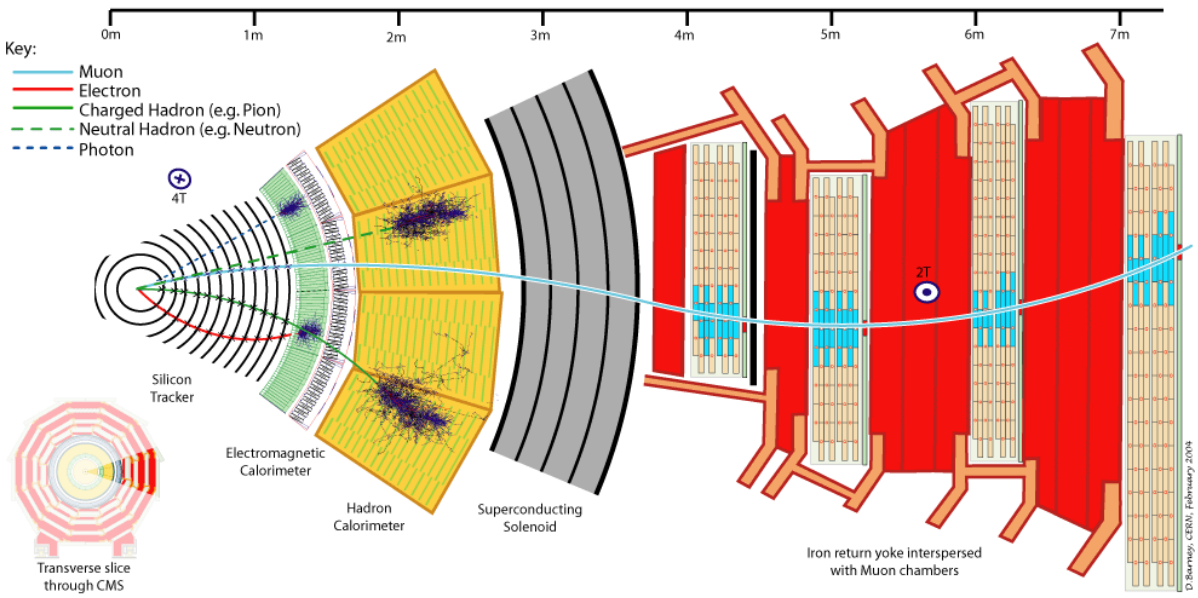


Figure 4.6: Cross sectional wedge of the CMS detector in the R-ϕ plane, showing the detector layers and the interaction of several types of particles with the various layers. Figure from Reference [6].

types: position and energy.

Detectors that gather position information of charge particles are called trackers. Generally, trackers use materials in which particles lose a small but measurable amount of energy in via ionization. The CMS inner tracker employs many thin layers of silicon, allowing for multiple measurements along a charged particle's trajectory. Other types of trackers use liquid or gas as their active media, such as those in the muon system. Though they only provide the position of a charged particle along its path, tracker information can be even more useful in the presence of a magnetic field. The force exerted on a charged particle by a magnetic field is a function of its velocity, and thus its momentum. Therefore, by measuring the curve in the path a charged particle takes, we can calculate its momentum. There is one caveat to this method: it can only be used to calculate the component of the momentum *perpendicular to the magnetic field*. We therefore refer to this as *transverse momentum*. Despite being a projection of the particle's total momentum, it is an extremely useful quantity.

Detectors that measure particle energy are called calorimeters. In contrast to trackers, which particles minimally interact with, calorimeters rely on particles depositing all of their energy within the detector. To achieve this stopping power for a broad range of particles, calorimeters use large volumes of dense material. Not all particles will be stopped by the same material, however, so calorimeters come in two types: electromagnetic and hadronic. Electromagnetic calorimeters (ECALs) measure the energy of electrons and photons, while hadronic calorimeters measure the energy of strongly-interacting particles—such as pions and neutrons.

To record the most information possible of the variety of particles produced in collisions at the LHC, CMS employs all of the previously mentioned detector technologies. We will start with the magnet that lends the detector its name. We will then work outward from the beam line, discussing the tracker and calorimeter systems. Finally, we

will review the other other namesake detector, the muon system.

4.2.1 Solenoid Magnet

We saw earlier that the presence of a magnetic field is required to determine the momenta of charged particles. For this to work, however, their are several requirements. The magnet needs to completely enclose the tracking system, and provide a steady and uniform magnetic field. Furthermore, to give the large bending power needed to precisely measure the momentum of high-energy charged particles, the field strength must be very high. Two magnet configurations that satisfy these criteria are solenoids and toroids. Both of these configurations give there name to the general-purpose detectors at the LHC: Compact Muon Solenoid (CMS) and A Toroidal LHC Apparatus (ATLAS). The choice of magnet field configuration informed the design of the rest of the detector systems in each machine, resulting in two unique experiments.

The solenoidal configuration chosen for CMS provides a large volume in with a uniform magnetic field parallel to the beam line. With an inner diameter of 6.3 m and a length of 13 m, it can accommodate both the tracker and calorimeter systems within it, and is the largest magnet of its kind ever built. In order to provide the current needed to generate a 4 Tesla field in such a large volume, superconducting wires (Niobium-Titanium) are wound in four layers around the central volume. When at the operating temperature of $-267\,^{\circ}\mathrm{C}$, the wires have a current of $\sim$20 kA and a stored energy of 2.6 GJ. To support the 220 ton weight of this system, the superconducting solenoid is housed within a 10,000 ton iron yoke, which also holds the barrel components of the muon system.

4.2.2 Tracker

The innermost component of the CMS detector is its tracking system. The tracker is responsible for measuring the paths of charged particles as they pass through its volume. Not only does this path information include the location of a charged particle within the detector, but because of the presence of a magnetic field, it also provides the momentum of the particle as well. By measuring the curvature of a particle's path through the tracker, the momentum transverse to the magnetic field can be calculated as $p_{\mathrm{T}} = BR$, where B is the magnetic field strength and R is the radius of curvature. The full three-momentum can also be calculated from this, but the transverse momentum is typically more useful.

On top of providing position and momentum information about each particle individually, the tracker data also allows for the reconstruction of the primary interaction point. We saw earlier that the LHC produces bunches of billions of protons, and these bunches are "crossed" to produce the collisions we measure at CMS. When these beams cross each other, a variety of interactions will occur. This means that for each event of interest (I'll discuss what makes an event interesting in Section 4.2.6), there are more than 20 other proton-proton interaction superposed on top of it. This means there are an average of 1000 particles passing through the tracker at every bunch crossing! In order to sort out which particles belong to which interactions, the particle paths are traced back to the beam line. From this, the particles can start to be grouped together based on shared points of origin. A point along the beam axis that multiple particles can be traced back to is referred to as a primary vertex. The number of primary vertices reconstructed for each beam crossing has increased over time as the instantaneous luminosity of the LHC has gone up, reaching as high as 80 during Run 2. Therefore, it is crucial that the tracker is able to provide precise and efficient measurements of particle trajectories.

To be able to handle the intense particle flux near the beam line while still maintaining the resolution and granularity requirements, the tracker system is built with two detector technologies: pixels and silicon strips. The pixel system is the part of the tracker closest to the interaction point, starting at 4 cm away and going out to 20 cm. The silicon strips provide coverage in the radial region between 20 and 116 cm. Both these systems use silicon as their sensor material, provide coverage for $|\eta| < 2.5$, and rely on the same interaction to produce a signal. As we saw earlier, trackers

When a charged particle passes through the silicon sensor, it transfers a small amount of energy into the material. This absorbed energy excites electrons in the material, producing electron-hole pairs (an electron excited to a higher energy state leaves a hole in its old state, which behaves like a positively charged particle). Under an applied electric field, these charge carriers drift towards their respective collection electrodes, where they induce a signal current, which is amplified and readout.

Immediately surrounding the interaction point in the barrel region are three layers of hybrid pixel detector modules. These are supported by disks on either side which provide high η coverage and contain two layers each. These pieces are arranged to provide three tracking points over almost the full η range. With a combined area of $\approx 1\,\mathrm{m}^2$, the pixel detector consists of 66 million pixels with a cell size of $100\,\mu\mathrm{m} \times 150\,\mu\mathrm{m}$. Despite this cell size, the pixel detector can provide a spatial resolution of about 10 µm. This resolution is possible thanks to the presence of the magnetic field, which causes the signal charges to spread out over multiple pixels. By analyzing how the signal charge is distributed, the particle position can be determined to a much higher degree of accuracy than if the charge was localized to a single pixel.

Outside of the pixel detectors there is the silicon strip tracker. Diagramed in Figure **??**, this system is composed of four subsystems: Tracker Inner Barrel and Disks (TIB/TID), Tracker Outer Barrel (TOB), and Tracker EndCaps (TEC±). Each of these

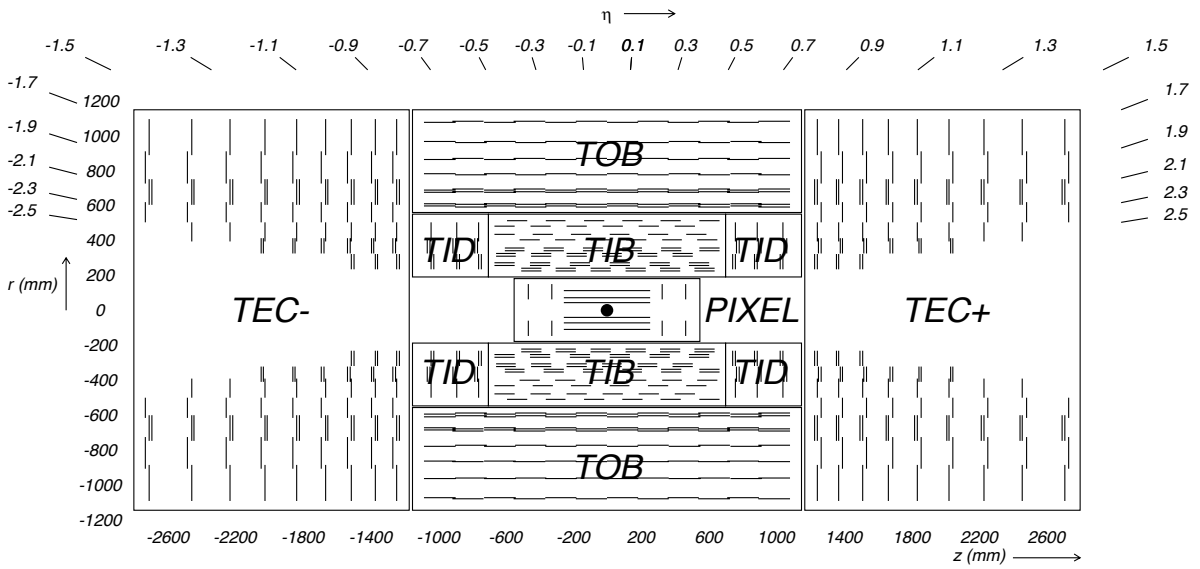


Figure 4.7: Arrangement of the pixel and silicon strip trackers. Figure from Reference [7].

subsystems is made of layers of arrayed silicon micro-strip sensors. In the barrel these strips are oriented parallel to the beam axis, while in the endcap they are arranged radially. In order to provide full coverage, different sensor geometries are used in each subsystem. Typical sensor dimensions are 6 cm × 12 cm in the inner barrel, and 10 cm × 9 cm in the outer barrel. The resolution also varies between subsystems, ranging from 23 µm to 35 µm in the TIB and 35 µm to 53 µm in the TOB. All together, there are 24,244 silicon sensors covering an active area of 198 m^2. With each sensor containing as many as 768 strips, there are a total of 9.3 million silicon strips.

4.2.3 Electromagnetic Calorimeter

Since the tracker is only able to provide positional information, another type of detector is required in order to provide information on the energy of particles. This is accomplished with two types of calorimeters: the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL).

The ECAL is able to measure low-mass particles which interact electromagnetically—electrons and photons—while the HCAL is able to measure hadrons. With the exception of muons and neutrinos, all collision products will deposit their energy in one of these calorimeters.

Providing coverage up to $|\eta| < 3.0$, the electromagnetic calorimeter (ECAL) contains 75,848 lead tungstate crystal scintillators.

Scintillators are materials which produce light when struck by an ionizing particle, such as an electron of photon. The amount of light produced in this interaction is proportional to the energy of the incident particle. By measuring the produced photons with photodetectors placed behind the scintillators, the energy of the incoming particle can be determined. Additionally, by analyzing the cells with the most measured light, the

4.2.4 Hadronic Calorimeter

Located immediately outside the electromagnetic calorimeter, the hadronic calorimeter (HCAL) is responsible for measuring the energies of strongly interacting particles. This includes many particles which would otherwise go undetected, such as neutral hadrons. Accurately measuring the energy of these particles is critical in the determination of the missing transverse energy, a quantity heavily used in SUSY searches. Due to the importance of the energy measurements it provides, the HCAL was designed to be as hermetic as possible.

Unlike the ECAL, which only uses one material, the HCAL uses alternating layers of two materials: a metal absorber and a plastic scintillator. When passing through the detector, hadrons will interact with the absorbing layers and produce a shower of particles. Charged particles in this shower will induce detectable light in the scintillators, which

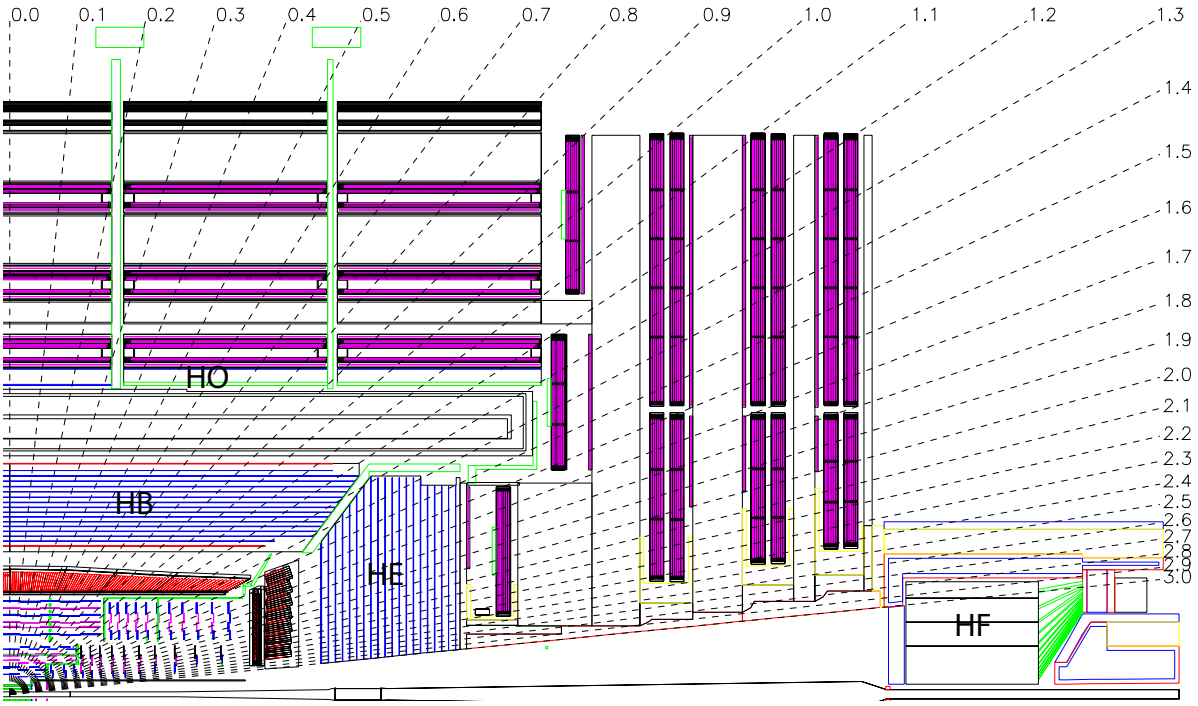


Figure 4.8: R-z view of a quarter section of the CMS detector showing the position of the four HCAL subsystems. The HCAL barrel and endcap (HB and HE) are located within the solenoid, while the outer and forward calorimeters (HO and HF) are outside. Figure from Reference [7].

is guided to the readout system by wavelength-shifting fibers. The energy of the initial shower is then extrapolated from these signals.

The HCAL consists of four sub-systems: the HCAL barrel (HB), HCAL endcap (HE), outer calorimeter (HO), and forward calorimeter (HF). The arrangement of these four subsystems is shown in Figure 4.8.

With the constraint of having to fit between the ECAL and the solenoid, the HCAL barrel (HB) system extends from $R = 1.77\,\text{m}$ to $R = 2.95\,\text{m}$. Composed of 36 identical wedges split into two barrels, the HB has 16 layers of metal absorber interspersed with 17 layers of plastic scintillator. The absorbers vary in thickness from 40 mm to 75 mm. The innermost and outermost absorber layers are made of steel to provide structural support, while the middle 14 layers are made of brass, which is non-magnetic and has

an interaction length of 16.42 cm. All together, these HB absorbers provide 5.82 nuclear interaction lengths. The 17 scintillator layers in between the absorbers vary in thickness from 3.7 mm to 9 mm, and are tiled. These tiles are divided into 2304 towers, with each tower covering 0.087×0.087 region in η and ϕ, corresponding to the same area covered by the 5×5 superclusters in the ECAL.

Despite the 5.82 interaction lengths provides by the HCAL barrel system, hadronic showers will still leak though. For this reason, an outer calorimeter (HO) is located outside the solenoid in the region $|\eta| < 1.26$. The HO uses both the solenoid coil and iron yolk as absorbers, and its scintillators segmentation mirrors the barrel calorimeter. The outer calorimeter increases the effective thickness of the HCAL to around 11 interaction lengths in the barrel.

To provide coverage that is hermetic as possible, forward calorimeters (HF) are placed beyond the muon endcap system on either side of the interaction point, and cover the region $3.0 < |\eta| < 5.2$. Unlike the rest of the calorimeters, which come in ECAL-HCAL pairs, the forward region only has one calorimeter on each side of the interaction point. We will see shortly that the novel design of the forward calorimeter allows for electromagnetic and hadronic energy contributions to approximated.

Due to their proximity to the beam line, these detectors experience extraordinary particle fluxes, with 760 GeV deposited per interaction, compared to 100 GeV for the rest of the detector. Driven by the need for radiation hardness, the detector was designed with quartz fibers as the active medium with steel absorbers. In the quartz fibers, signal is generated when charged shower particles above the Cherenkov threshold generate light. This mechanism makes the HF calorimeters mostly sensitive to the electromagnetic components of showers, and this sensitivity is exploited to estimate the electromagnetic and hadronic components of the shower. Additional quartz fibers are included in the first 22 cm of absorber and read out separately. Showers generated by electrons and photons

will deposit most of their energy in the first 22 cm of material, while hadronic showers deposit energy uniformly throughout the detector volume. With the energy measured in the long fibers as L and the energy measured in the short fibers as S, electromagnetic energy can be approximated by $2S$, while the hadronic component is $L - S$.

4.2.5 Muon System

The muon system consists of several layers of three detector technologies: drift tubes (DTs), cathode strip chambers (CSCs), and resistive plate chambers (RPCs). These technologies provide coverage in different regions of the detector. DTs cover the barrel region ($|\eta| < 1.2$), CSCs are used in the endcaps ($0.9 < |\eta| < 2.4$), and RPCs provide complementary measurements in both ($|\eta| < 1.9$). The arrangement of these detectors is shown in Figure 4.9. In the following sections, I will review each of these detector types.

Drift Tubes

The fundamental unit of drift tubes is the drift cell (shown in Figure 4.10). These cells are rectangular chambers (2.4 m long with a cross-section of 42 mm × 13 mm) filled with Argon gas. An anode wire held at 3600 V runs along the center of the cell, while cathode strips held at −1200 V run along the cell wall. A muon passing through the drift cell will ionize the Argon gas, causing free electrons produced by this ionization then drift to the anode wire where they produce a measurable change in voltage. Four parallel drift cells grouped together are called a superlayer.

In the barrel region, drift tubes are organized into four stations arranged in concentric cylinders with radii of of 4.0, 4.9, 5.9, and 7.0 m from the beam axis. The inner two stations contain three superlayers each: two oriented to measure r and ϕ, and a perpendicular layer them between to measure r and η. The two outer stations have two

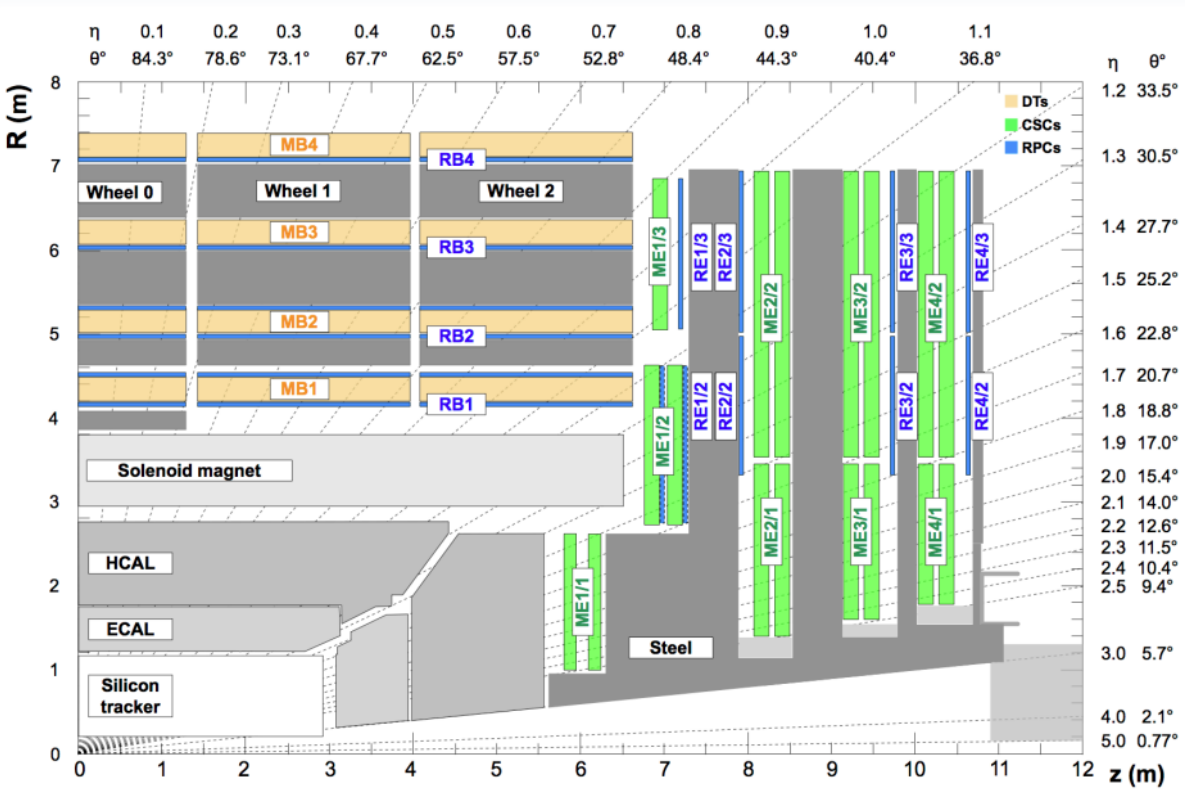


Figure 4.9: *R-z* view of a quadrant of the CMS detector highlighting the muon system layout. Figure from Reference [8].

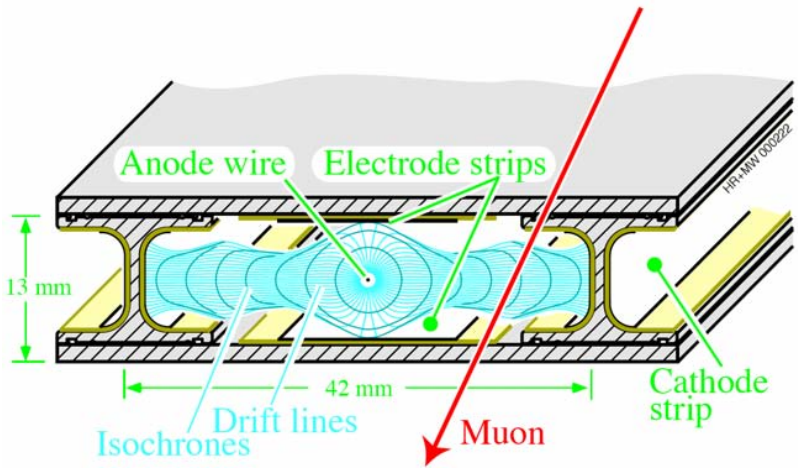


Figure 4.10: Cross-sectional view of a drift cell. Figure from Reference [7]

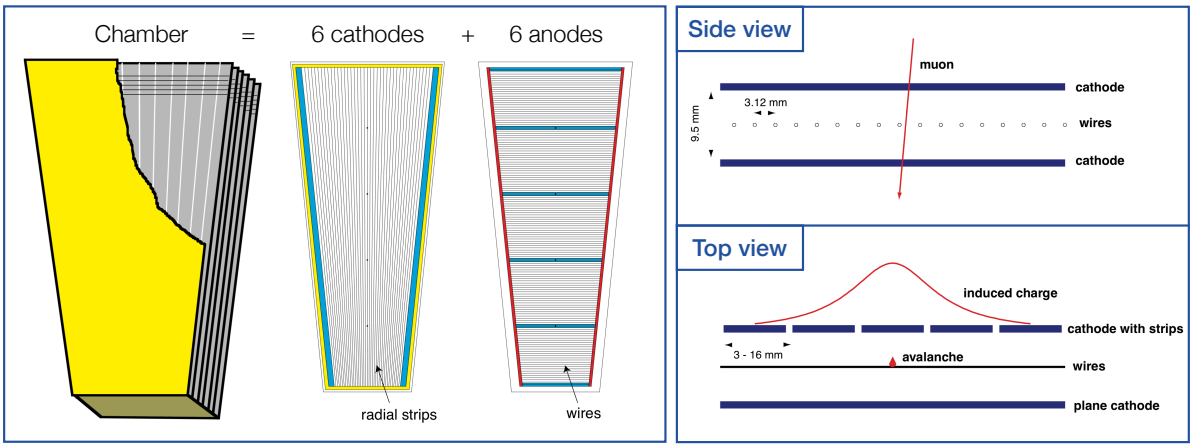


Figure 4.11: The orientation of the anodes and cathodes of a cathode strip chamber, with the dimensions between each shown (right). (ODMB). Image from Reference [7].

superlayers each, excluding the middle η-sensitive superlayer.

Cathode Strip Chambers

The flux of muons increases as you approach the beam line. In order to handle this increased muon rate, cathode strip chambers (CSCs) are used. With low drift times, CSCs have much fast response times than the DTs in the barrel. Diagrammed in Figure 4.11, each CSC consists of six layers of multi-wire proportional chambers—filled with a 50% CO_2, 40% Ar, and 10% CF_4 gas mixture—paired with finely segmented cathode readout strips. When a muon passes through a chamber will ionize the gas and cause a charge buildup on the anode wires, which can be measured to determine $r - \eta$ position. This charge buildup also induces a measurable charge on the cathode strips, providing $r - \phi$ position as well. Analysis of the charge distribution between adjacent strips allows CSCs to achieve a resolution lower than the strip width, with an offline spatial resolution of 75 µm in the inner two rings and 150 µm in the other CSCs.

A diagram of the electronics used to readout data from a CSC chamber are shown in Figure 4.12. Signals produced by the anode wires are amplified by the anode front-end

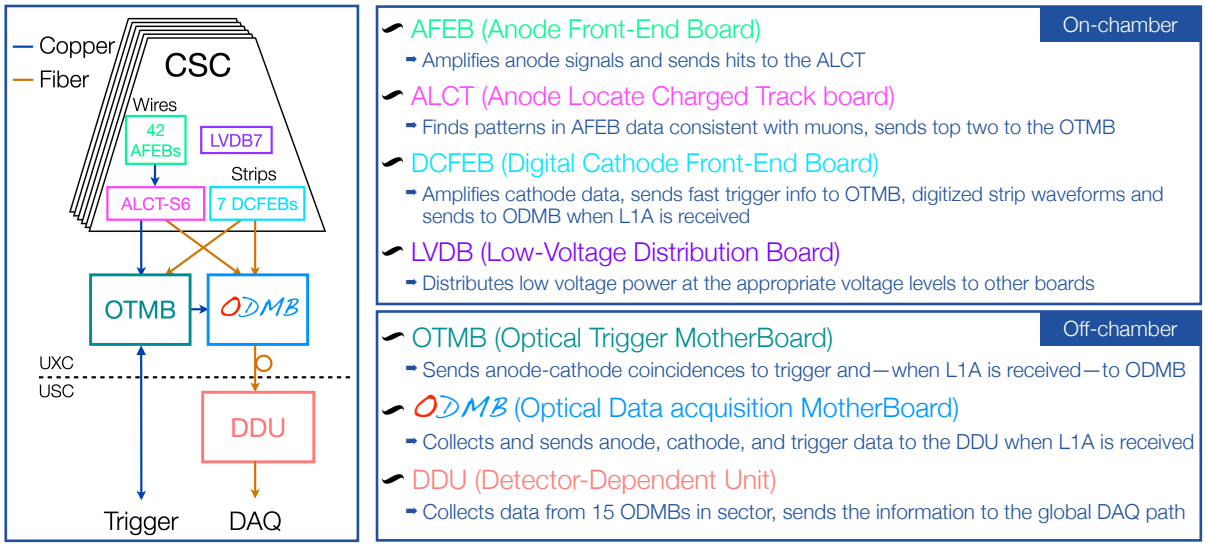


Figure 4.12: Schematic showing connections between on-chamber electronics and the upgraded Optical Data Acquisition Motherboard (ODMB7/5), as well as a description of each component of the readout system.

board (AFEB) and then passed to the local charged track board (ALCT) where hits across anode layers are matched to form muon stubs. The ALCT then sends two candidates with the most layer hits off-chamber to the optical trigger motherboard (OTMB), located in a crate in the CMS cavern. Cathode strip readings are collected by digital cathode front-end boards (DCFEBs), which amplify the signals and send them to the OTMB where they are analyzed. Coincidences between hit patterns in the anode and cathode data found by the OTMB are then are then sent to the main CMS trigger path.

Along with sending out data, the OTMB is also responsible for receiving the Level 1 Accept (L1A) signal from the central trigger system, which indicates the event should be recorded. The OTMB then passes this signal along to the other readout electronics. When an L1A is received, the optical data acquisition motherboard (ODMB) is responsible for collecting all chamber information and sending it to the central CMS DAQ path via the detector-depending unit (DDU). The ODMB gets cathode information from the DCFEBs and combines it with the anode and trigger information from the OTMB,

sending the results to the DDU.

The increase in muon flux that will accompany the start of high-luminosity LHC (HL-LHC) will require the CSC readout electronics to operate with a data rate beyond their current capacity. To handle the increased data rate, the CSC readout electronics are being upgraded. In particular, the ODMBs responsible for the chambers closest to the beam line—of which there are 180—will be completely redesigned. The new ODMBs will include additional optical links to provide the requisite data transfer rate for the on-chamber electronics as well as the DDU.

Resistive Plate Chambers

Resistance plate chambers (RPCs) compliment the DTs and CSCs in the region $|\eta| < 1.9$. Yet another gas detector, RPCs consist of two parallel resistive plates—an anode and a cathode—separated by a gas volume filled with a $C_2H_2F_4$ mixture. When a muon passes through the chamber, the resulting ionization creates an avalanche of electrons that will pass through the resistive plates and be measured by a copper readout strips. Though their spatial resolution is not as good as DTs or CSCs, RPCs provide a time resolution of 1 ns, making them extremely useful for identifying which bunch crossing a muon was produced in.

4.2.6 Trigger and Data Acquisition

As we learned in Section 4.1.2, each bunch crossing produces many final state particles, most of which originate from soft proton collisions. With bunch crossings occurring every 25 ns, there can be up to 3 billion collisions per second, with most of these collisions being of little interest to us. While the detector response is synchronized to the bunch crossing frequency, the speed of reading out the data is much slower, allowing us

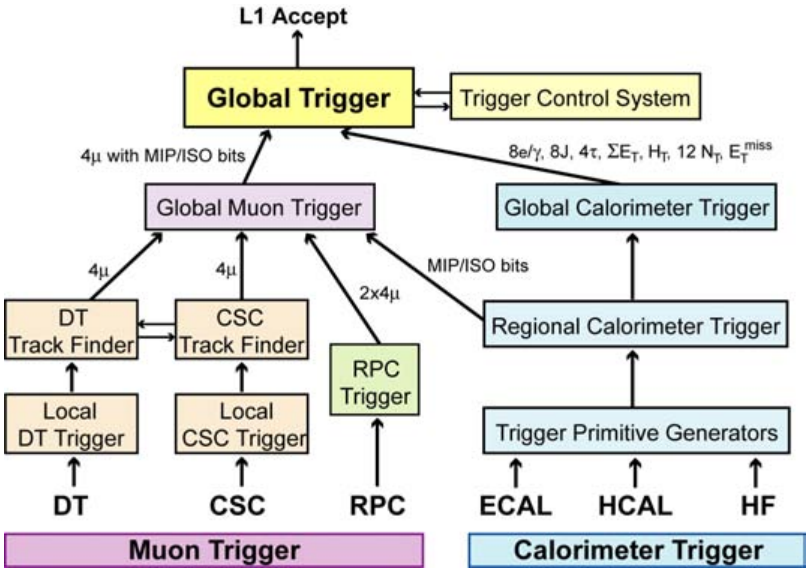


Figure 4.13: Schematic diagram showing how an L1A is generated from trigger primitives from the calorimeter and muon systems [7].

to only record about 1000 events per second. Most bunch crossings will only produce soft collisions that are not worth recording, so determining when events of interest occur is of paramount importance. At CMS, this selection is performed using a trigger system composed of two stages. The first stage, L1, uses fast-response data from the calorimeters and muon system to identify events that contain particular candidate objects such as electrons, muons, photons, and jets. Figure 4.13 shows the pieces from the muon system and calorimeters, known as trigger primitives, that contribute to an L1 accept. This identification occurs at a rate of 100 kHz.

The second level of the CMS trigger system, known as the high-level trigger (HLT), uses data from all subsystems to perform a more complete reconstruction of candidate objects. To increase processing efficiency, this reconstruction is separated into different workflows, known as *HLT paths*, with each path using different algorithms and selection criteria to determine if an event has certain characteristics. Paths may correspond to one or more different objects. For example the `HLT_IsoMu24` path is designed to trigger

events with an isolated muons with $p_{\mathrm{T}} > 24\,\mathrm{GeV}$, while the `HLT_Mu15_IsoVVVL_PFHT450` path is designed to trigger on events with an isolated loose muon with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$ and $H_{\mathrm{T}} > 450\,\mathrm{GeV}$. To determine whether to process an event, each HLT path starts from a set of L1 triggers called *seeds*; if an event was accepted by any of the L1 seeds, it will be processed by the path. In the end, if any HLT paths accept an event, the full event information is stored. The high-level trigger reduces the total event rate to around 1 kHz.

Beyond determining what events to save, the HLT is also used to organize the data. Once an event is accepted by the HLT, it transferred to the CMS computing center for offline processing and storage. Once the full offline event reconstruction is performed, the event is added to datasets based on which HLT paths accepted it. This grouping is non-exclusive, so an event with multiple features of interest—such as a muon and an electron—may be put into several datasets. This organization allows analyzers to restrict their analysis to the subset of the data with the features they are interested in. We will see in Chapter 6 what triggers this analysis uses in data.

Chapter 5

Object and Event Reconstruction

In Chapter 4, we saw how different particles interact with each detector system in CMS, as well as what information we can infer from these interactions, such as particle position and energy. Determining these quantities from detector data and classifying them by particle type is known as *reconstruction*. During reconstruction, raw detector outputs are translated into abstract objects that represent electrons, photons, and muons. In this chapter, we will discuss the techniques used to perform this reconstruction and the criteria used to define the objects used in this analysis. When discussing reconstructed objects, we should always keep in mind that these are actually object *candidates*: a hypothesis of what particle produced the signal in the detector. Though we do our best to reconstruct objects accurately, the objects we are working are not the particles themselves.

We will start by discussing the particle flow (PF) algorithm, the backbone of reconstruction at CMS.

5.1 Particle Flow

The particle flow (PF) algorithm [30] describes events through the comprehensive list of final-state particles it reconstructs and classifies. This task is accomplished by correlating basic detector elements, tracks and clusters, from each detector layer and then combining these measurements to reconstruct particles and their properties.

PF reconstruction is divided into three steps. First, subdetector information is used to produce *particle flow candidates*. PF candidates are either tracks constructed from measurements by the inner tracker and muon system, or energy clusters constructed from calorimeter data. The track objects are reconstructed iteratively using specialized software called the combinatorial track finder (CTF) [31]. This software builds particle trajectories from from groups of hits starting in the innermost tracker layers and moving outward. A final refit is then performed on the resulting trajectories, with any likely spurious tracker hits removed.

PF clusters are formed in a similar manner. Starting from calorimeter cells with a local energy maximum, topological clusters are formed by iterating over surrounding cells and adding those with energy above some threshold. Topological clusters are then fit with a Gaussian-mixture model, which assumes the energy deposits may originate from multiple seeds, to reconstruct individual clusters within a topological cluster.

In the next step of PF reconstruction, a link algorithm iterates through pairs of PF candidates from different subdetectors and links together candidates that are found to be mutually consistent with a single object. The criteria used to determine if two PF candidates should be linked depend on the candidate types being considered, and are specified in [30]. Rather than considering all pairs of PF candidates, the link procedure only considers pairs of candidates that are nearest neighbors in the η-ϕ plane. Linked

objects are then further combined to form *PF blocks*[1]. Due to the high granularity of the trackers at CMS, PF blocks typically contain elements originating from one or two particles. Within each block, higher level identification and reconstruction sequences are executed for particle candidates in the following order: muons, electrons, isolated photon, and jets. The order of reconstruction is based on signature, with muons having the most distinctive detector signature and jets acting as a catch-all for tracks and clusters not reconstructed into the other objects. Once an element is associated with a particle candidate, it is removed from consideration for future candidates. Once all blocks have been processed, the linking step is complete.

With all the reconstructed objects, a final post-processing step is run that calculates global variables such as $p_{\mathrm{T}}^{\mathrm{miss}}$ and H_{T}. Additional checks are made to verify that objects were reconstructed correctly, with special attention paid to $p_{\mathrm{T}}^{\mathrm{miss}}$. The resulting event description—object lists and global variables—are then made available to the analyzers for more strict selections. The following sections will summarize the additional criteria applied for each object type.

5.2 Leptons

Muons are the most distinguishable type of particle measured in CMS. This is due to the minimal interaction they have with the calorimeters, which allows them to be measured by the inner tracker as well as the outer muon system. As such, muons can be reconstructed from measurements in both systems. Those reconstructed from inner tracker hits are known as *tracker muons*, while those reconstructed from muon system hits are known as *standalone muons*. To distinguish them from other charged particles that produce tracks, tracker muons are also required to have at least one matching muon

[1]Why a group of links isn't called a chain is beyond me.

Table 5.1: Kinematic and quality requirements applied to muons. d_{xy} and d_z are the transverse and longitudinal impact parameters of the tracks associated to the muon. The "Is global or tracker muon" excludes standalone muons.

Muon Property	Requirement
p_{T} [GeV]	> 20
$\lvert\eta\rvert$	< 2.4
$\lvert d_{xy}\rvert$ [mm]	< 2
$\lvert d_z\rvert$ [mm]	< 5
$I^{\mathrm{rel}}_{\mathrm{mini}}$	< 0.2
PF muon ID	True
Is global or tracker muon	True
Fraction of valid tracker hits	> 0.8
Segment compatibility	> 0.303
Requirements below apply only if segment compatibility	≤ 0.451
Normalized global track χ^2	< 3
χ^2 of tracker-standalone match	< 12
Track kink χ^2	< 20

segment in the muon system. These reconstructions are performed separately, with the results getting matched later. Tracks found to be consistent between both the inner tracker and muon system are referred to as *global muons*. The additional requirements imposed on muon candidates are shown in Table 5.1, with the resulting efficiency plotted as a function of p_{T} and η in Figure 5.1.

The second object to be reconstructed, electrons are identified from charged particle tracks coupled with ECAL clusters. Unlike muons, electrons lose a significant fraction of their energy in the inner tracker due to bremsstrahlung, the power of which is proportional to $(E/m)^4$. This radiation slightly complicates the reconstruction since the radiated photons, which may not be close to the electron by the time they reach the ECAL, need to be associated with the electron candidate in order to fully reconstruct its four momentum. Due to the magnetic field, however, the radiated photons will only be separated from the electron along the ϕ direction. Therefore, cluster elements in the ECAL are combined along this direction to form *superclusters*, which are then spatially

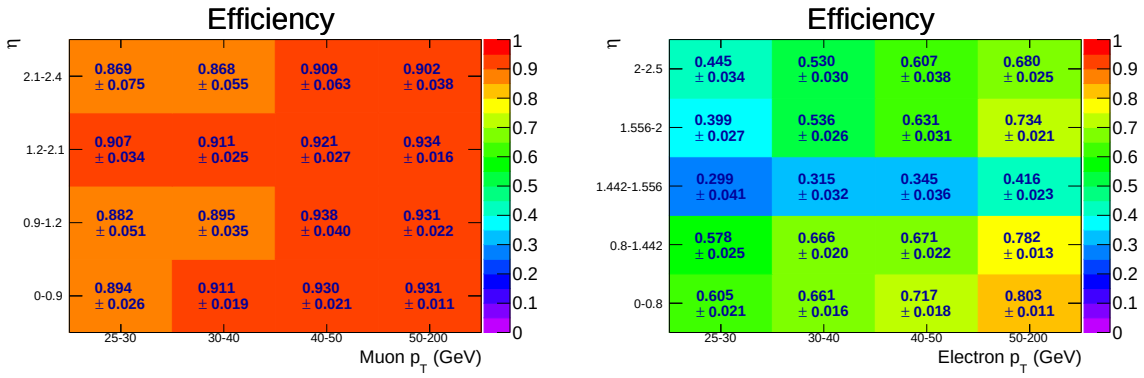


Figure 5.1: Overall efficiency of the reconstruction, identification, isolation, and vertexing requirements for both muons (left) and electrons (right) as a function of p_{T} and η. Plots from supplementary material for Reference [9] and published online in Reference [10].

associated with consistent hits from the inner tracker. Furthermore, to account for the sudden changes in flight path caused by bremsstrahlung, electron tracks are reconstructed with a Gaussian-sum filter (GSF) [32]. The requirements placed on electron tracks and superclusters are listed in Table 5.2, with the resulting efficiency plotted as a function of p_{T} and η in Figure 5.1. Since the detector signature of the electron is not as distinct as that of the muon, electrons suffer from a significantly lower identification efficiency.

5.2.1 Mini Isolation

As we expect signal leptons to originate from the decay of a W boson, we require the leptons to be isolated from other PF candidates. This helps distinguish *prompt leptons*–those produced by the hard-scatter process–from non-prompt leptons, which are produced by secondary decays or hadronization. Lepton isolation is quantified by summing the transverse momenta of all particles within some ΔR of the lepton candidate and then dividing by the p_{T} of the lepton being considered. Specifically, the numerator has the

Table 5.2: Kinematic and quality requirements applied to electrons. Different requirements are applied to electrons in the barrel ($|\eta_{\text{supercluster}}| \leq 1.479$) and endcap ($|\eta_{\text{supercluster}}| > 1.479$). d_{xy} and d_z are the transverse and longitudinal impact parameters of the tracks associated to the electron. $\sigma_{i\eta i\eta}$ is a shower shape variable which measures the width of the ECAL energy deposits in the η direction.

Electron Property	Barrel Requirement	Endcap Requirement
p_{T} [GeV]	> 20	> 20
$\lvert\eta_{\text{supercluster}}\rvert$	< 1.479	< 2.5
$\lvert d_{xy}\rvert$ [mm]	< 0.118	< 0.739
$\lvert d_z\rvert$ [mm]	< 3.73	< 6.02
$I^{\text{rel}}_{\text{mini}}$	< 0.1	< 0.1
$\sigma_{i\eta i\eta}$	< 0.0101	< 0.0283
$\Delta\eta$(supercluster, track)	< 0.0103	< 0.00733
$\Delta\phi$(supercluster, track)	< 0.0336	< 0.114
$E_{\text{hadronic}}/E_{\text{electromagnetic}}$	< 0.0876	< 0.0678
$\frac{1}{E} - \frac{1}{p}$ [GeV^{-1}]	< 0.0174	< 0.0898
Missing hits	≤ 2	≤ 1
Pass photon conversion	True	True

form:

$$
\begin{aligned}
I^{\text{rel}} = \Big(& \sum_{\text{cone}} p_{\text{T}}(\text{charged hadrons from PV}) \\
& + \max\Big[\sum_{\text{cone}} p_{\text{T}}(\text{photons}) + \sum_{\text{cone}} p_{\text{T}}(\text{neutral hadrons}) \\
& \quad - \frac{1}{2} \sum_{\text{cone}} p_{\text{T}}(\text{charged hadrons not from PV}),\, 0 \Big] \Big) / p^{\ell}_{\text{T}},
\end{aligned}
\tag{5.1}
$$

where the last term is included to correct for energy from pileup near the leptons. When computed with a fixed cone radius, the resulting ratio is known as *relative isolation*. In this analysis, isolation is calculated using a cone radius that varies with the p_{T} of the lepton, according to

$$
R_{\text{mini-iso}} = \begin{cases} 0.2, & p^{\ell}_{\text{T}} \leq 50\,\text{GeV}, \\ \frac{10\,\text{GeV}}{p^{\ell}_{\text{T}}}, & p^{\ell}_{\text{T}} \in (50\,\text{GeV},\, 200\,\text{GeV}), \\ 0.05, & p^{\ell}_{\text{T}} \geq 200\,\text{GeV}. \end{cases}
\tag{5.2}
$$

The resulting isolation, I_{mini}, is referred to as *mini isolation* [33]. The choice of using a p_{T}-dependent cone size is motivated by the angular separation expected between two particles produced by the decay of a massive particle. In the limit of a high-momentum parent particle, the separation between the decay products in η-ϕ is $\Delta R \approx 2M_{\text{parent}}/p_{\text{T}}^{\text{parent}}$. Though this is a simplified case, it indicates the separation should scale as $1/p_{\text{T}}$, and the constant in the numerator should be related to the mass of the decaying particle. As we are trying to discriminate from non-prompt leptons produced in b-quark decay, a value of 10 GeV (approximately twice the mass of the b-quark) is chosen for the numerator. An upper limit on the cone size of 0.2 is chosen to avoid overlap with jets, and a lower limit of 0.05 avoids issues related to detector resolution effects.

5.2.2 Veto Tracks

The dominant background of this analysis, as will be discussed in more detail in Chapter 7, is $\text{t}\bar{\text{t}}$ events in which both top decays produce a lepton. These events make it into our search region when one of the leptons is lost, due to being outside detector acceptance or failing identification. To account for these events, we define a collection of *veto tracks* that contain candidates that are likely to be a lepton or a charged hadron from τ lepton decay, but fail lepton identification. PF candidates for leptons and charged hadrons are considered for this collection, with the requirements for each summarized in Table 5.3.

In addition to the two leptons, 2ℓ $\text{t}\bar{\text{t}}$ events also have two neutrinos that contribute to missing transverse energy ($p_{\text{T}}^{\text{miss}}$) in the event. In 1ℓ $\text{t}\bar{\text{t}}$ events, $\vec{p}_{\text{T}}^{\,\text{miss}}$ will originate from a single neutrino. This difference allows for a transverse mass to be calculated between $\vec{p}_{\text{T}}^{\,\text{miss}}$ and the reconstructed lepton ($\vec{\ell}$) that has an upper bound of m_{W} in 1ℓ $\text{t}\bar{\text{t}}$, where

Table 5.3: Selection requirements for veto tracks. In all cases, it is assumed the track has not already been identified as an electron or muon. m_{T2} requires an identified electron or muon to compute. If an event does not contain any leptons, the m_{T2} and charge requirements are omitted. For charged hadron tracks, the definition of mini-isolation is modified to exclude photons and neutral hadron candidates, retaining only the first term in Equation (5.1).

Track Property	Lepton PF Cand.	Charged Hadron PF Cand.
p_{T} [GeV]	> 10	> 15
d_{xy} [mm]	< 0.5	< 0.5
d_z [mm]	< 0.7	< 0.7
I_{mini}	< 0.2	N.A.
$I_{\mathrm{mini}}^{\text{chg. trk. only}}$	N.A.	< 0.1
Required only if event contains a lepton		
$m_{\mathrm{T2}}(\text{track}, l, p_{\mathrm{T}}^{\mathrm{miss}})$ [GeV]	< 80	< 60
Opposite charge as lepton	True	True

the transverse mass between two particles a and b is defined as

$$m_{\mathrm{T}}(a,\, b) \equiv m_a^2 + m_b^2 + 2(E_{\mathrm{T},a}E_{\mathrm{T},b} - \vec{p}_{\mathrm{T},a} \cdot \vec{p}_{\mathrm{T},b}). \tag{5.3}$$

If two neutrinos are contributing to $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$, however, this transverse mass has no upper bound. Therefore, in events with both a lepton and veto track, a new mass variable is constructed. Known as m_{T2} [34, 35], it is defined as

$$m_{\mathrm{T2}}(\text{track},\, l,\, \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}) \equiv \min_{\vec{p}_{\mathrm{T},a}+\vec{p}_{\mathrm{T},b}=\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}} \{\max[m_{\mathrm{T}}(\text{track},\, \vec{p}_{\mathrm{T},a}),\, m_{\mathrm{T}}(l,\, \vec{p}_{\mathrm{T},b})]\}. \tag{5.4}$$

A scan over pairs of transverse momenta that sum to the reconstructed $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is represents the two neutrinos. For each pair, two transverse masses are calculated, pairing neutrino candidate a with the veto track and neutrino candidate b with the lepton. The maximum of these two values is then minimized across all neutrino candidate combinations. Similar to m_{T} in the single lepton case, m_{T2} is expected to have an upper bound of m_{W} in events with two leptonically decaying W bosons. Therefore, track candidates that give a large value of m_{T2} are not counted as veto tracks.

5.3 Jets

By far the most common type of object produced at the LHC, jets are collimated sprays of particles produced by the hadronization of quarks and gluons. Rather then work with each particle separately, it is more convenient to cluster them together and treat them as a single object. Furthermore, this clustering allows us to reconstruct particles produced in the hard scattering that subsequently decayed.

Jets are formed by clustering PF candidates together into a cone of radius R in η-ϕ space using the anti-k_T algorithm [36]. This clustering is performed using the FastJet package [37]. The cone size of jets can vary depending on the mass of the hadronizing object that is being reconstructed, as this impacts the separation the decay products are expected to have. The standard jet cone size, which we use here, is $R = 0.4$ which allows for b quark decay products to be contained within a single jet[2].

To determine the order in which particles should be clustered, distance between particles i and j is defined as

$$d_{ij} = \min\left(p_{\mathrm{T}i}^{-2}, p_{\mathrm{T}j}^{-2}\right) \frac{\Delta R_{ij}^2}{R^2}, \tag{5.5}$$

where $\Delta R_{ij}^2 = \Delta\eta_{ij}^2 + \Delta\phi_{ij}^2$ and $R = 0.4$ is the jet radius parameter. Additionally, a self-distance is defined for each object:

$$d_i = p_{\mathrm{T}i}^{-2}. \tag{5.6}$$

Starting from the list of unassigned PF candidates, the anti-k_T algorithm iteratively clusters particles starting with the minimum distance parameter and proceeds in ascending order. If the minimum distance comes from a pair of particles, the particles' four momenta are added together, with the result replacing the initial particles in the candidate list. If the minimum distance comes from a single object, then it is removed from the

[2]Using the formula introduced in Section 5.2.1, the decay products of a b quark with a p_{T} of 30 GeV would have a $\Delta R \approx 0.33$.

Table 5.4: Kinematic and quality requirements applied to the jets after removing those matched to leptons. The fraction requirements refer to the fraction of the energy of the jet coming from a particular source as identified by the PF algorithm.

Jet Property	Requirement
p_{T} (corrected)	$> 30\,\mathrm{GeV}$
$\lvert\eta\rvert$	< 2.4
Neutral hadron fraction	< 0.99
Charged hadron fraction	> 0
Neutral electromagnetic fraction	< 0.99
Charged electromagnetic fraction	< 0.99
Number of constituents	> 1
Charged multiplicity	> 0

list and considered a jet. This process continues until all particles have been assigned to a jet. The distance metric used by the anti-k_T algorithm leads to jets being clustered around the highest p_{T} particles first and results in a circular shape in the η-ϕ plane. When using the standard radius of 0.4, jets constructed in this way are known as AK4 jets.

To accurately reconstruct the particle that produced the jet, corrections are applied to account for the effects of pileup and other experimental effects. As was the case for leptons, these corrections are applied to each object independently and are p_{T} and η-dependent. Once these corrections are applied, identification requirements are imposed to remove poorly reconstructed jets. Listed in Table 5.4, these requirements check that the reconstructed jet had more than one constituent and was measured by both calorimeters.

As jets are made from the complete PF candidate list, they may contain candidates previously identified as isolated leptons based on the criteria given in Section 5.2. To avoid double counting the leptons or their energy, jets are removed from the jet collection if they contain a lepton candidate or if any isolated leptons are within $\Delta R = 0.4$ of the jet momentum. Non-isolated leptons, or candidates failing another identification criterion, are clustered into jets.

We define N_{jets} as the number of AK4 jets passing the requirements given above. Additionally, we define H_{T} as the scalar sum of the transverse momenta of these jets, and S_{T} as this sum including reconstructed leptons.

5.3.1 b-tagging

As discussed earlier, jets allow us to reconstruct colored objects originating from the initial hard process. In most cases we are unable to uniquely identify the object that produced a jet, be it a quark or a gluon. The similar couplings and low masses of the first two generations of quarks lead to very small kinematic differences between the resulting hadronization processes. With a mass of 4.6 GeV and CKM-suppressed decay modes, however, hadrons containing b quarks will travel several millimeters in the detector before decaying. This leads to a distinctive secondary decay vertex that can be reconstructed and used to identify a jet originating from a b quark. Known as b tagging, this procedure is critical to many SUSY searches, including the one presented here, as the relative lightness of third-generation squarks lead to the corresponding quarks appearing in the final states of many simplified models.

The procedure of b tagging is performed at CMS using the Combined Secondary vertex (CSV) tagger [38], a multivariate analyzer (MVA) which takes the PF candidates from each jet as an input and returns a discriminator score between 0 and 1, with 1 representing a jet very likely to be from a b quark. If a jet has a score above some threshold, it is considered b tagged. The threshold used is referred to as the working point. Working points are typically defined based on the false positive rate—the rate at which jets not originating from a b quark are b tagged—with the loose, medium, and tight working points corresponding to false positive rates of 10, 1, and 0.1% respectively. The efficiencies for these working points as a function of jet p_{T} are shown in Figure 5.2 for

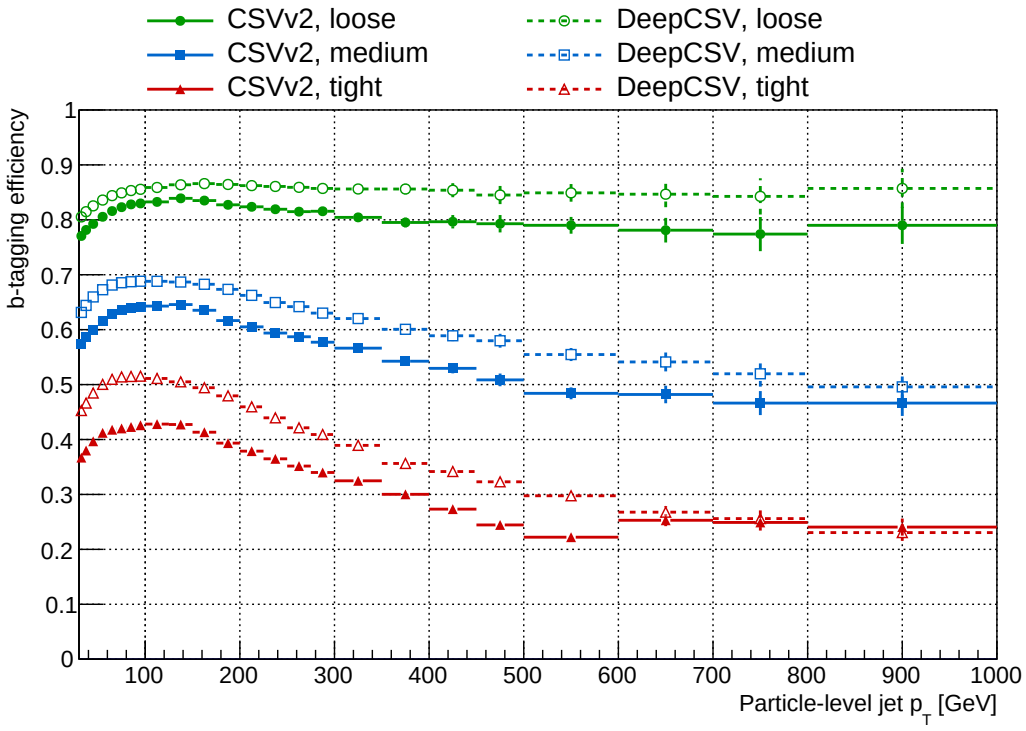


Figure 5.2: Efficiency of the CSVv2 and DeepCSV b-tagging efficiencies, computed in simulated $t\bar{t}$ +jets events. The medium working point of CSVv2 is used in this analysis. Figure from Reference [10].

two implementations of the CSV algorithm: CSVv2 and DeepCSV [39]. In this analysis, we use the medium working point of the DeepCSV algorithm, which gives an efficiency of 68% for b jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ in $t\bar{t}$ events. We define N_{b} as the number of AK4 jets that are b tagged.

5.3.2 Large-Radius Jets

While a jet radius of 0.4 is sufficient for capturing the decay products of a b quark, heavier particles such as top quarks require a larger radius. Due to the extremely high mass of the top quark, its decay products can span across several AK4 jets. In order to recover this kinematic information, we create large-radius jets ($R = 1.4$)—henceforth referred to as large-R jets—by clustering together the AK4 jets described in Section 5.3.

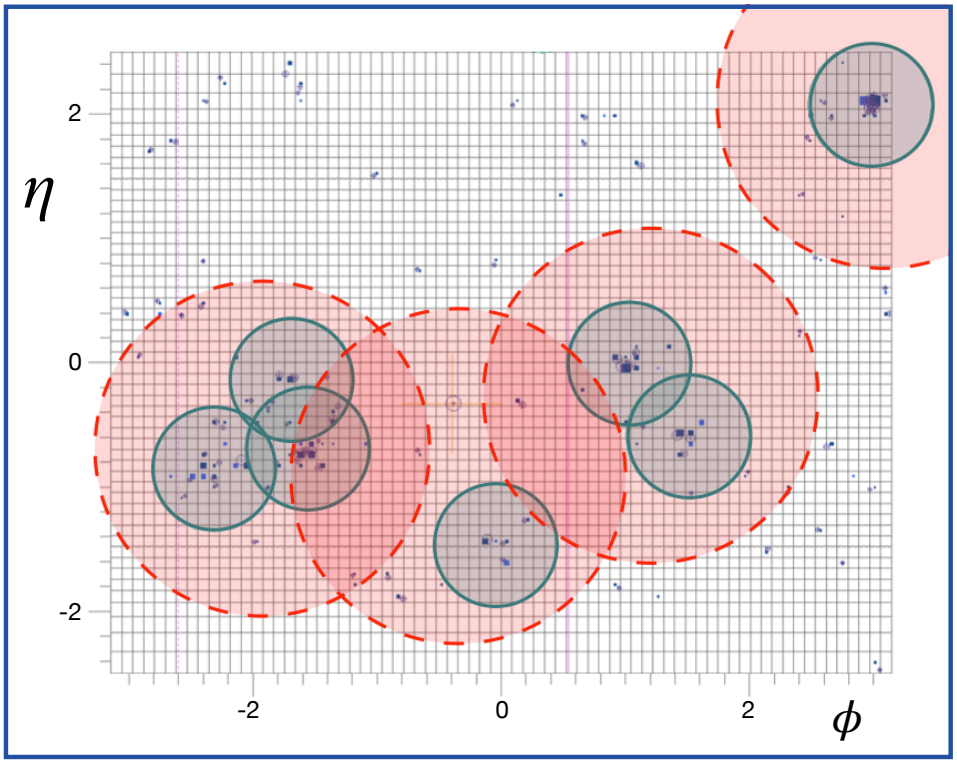


Figure 5.3: η-ϕ plane for a simulated 1ℓ $t\bar{t}$ event showing the measured energy depositions and how they are clustered into small- (green) and large-radius (red) jets.

Figure 5.3 shows this clustering procedure in the η-ϕ plane for a simulated 1ℓ $\mathrm{t\bar{t}}$ event.

When the underlying event contains heavy particles that decay hadronically, such as the four top final state of our signal models, the constructed large-R jets are expected to have correspondingly large masses. To represent the mass scale of the underlying event, we define the variable M_J as the sum of all large-R jet masses:

$$M_J = \sum_{J_i=\text{large-}R\text{ jets}} m(J_i). \tag{5.7}$$

Properties of this variable have been explored in depth in References [40, 41, 42]. We will study how M_J behaves for our signal and backgrounds in Chapter 7.

Instead of clustering AK4 jets, large-R jets could be constructed in the typical fashion using PF candidates and the anti-k_T algorithm with a higher input radius. Furthermore, similar to the b tagging of AK4 jets, top tagging can be performed on large-radius AK

jets. Due to our top-heavy final state, we studied the efficacy of using higher radius AK jets with top tagging—namely, the DeepAK8 jet framework described in [43]. In the end, however, top tagging did not offer a significant enough improvement over the nominal M_J to warrant the added uncertainty in background estimation that its inclusion would have caused.

5.4 $\boldsymbol{p}_{\mathrm{T}}^{\mathbf{miss}}$

A key feature of many SUSY models is the undetectability of the supersymmetric particles in the final state. Though this may seem like an undesirable feature, it provides significant discrimination power at the LHC. Because the momenta of the incoming protons are only along the beam axis, the colliding system has a negligible amount of momentum in the transverse direction. Therefore, assuming all particles are detected, summing the transverse momenta of the final state particles should give zero. In the case particles are not measured, however, the total transverse momenta may differ significantly from zero. To quantify this difference, we define the missing transverse momentum $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ as

$$\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} = -\sum \vec{p}_{\mathrm{T}}, \tag{5.8}$$

where the sum is taken over all jets and PF candidates not clustered into jets. As we will most often refer to the magnitude of this quantity, we define $p_{\mathrm{T}}^{\mathrm{miss}} = |\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}|$. In addition to indicating the presence of unmeasured particles, $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ also has the nice property that, in events where a single particle (such as a neutrino) escapes the detector without being detected, $p_{\mathrm{T}}^{\mathrm{miss}}$ recovers this particles p_{T}. We will see in Chapter 7 that we can use this property to remove a large fraction of the 1ℓ $\mathrm{t\bar{t}}$ background from our search region.

Chapter 6

Data and Simulated Event Samples

In Chapter 4, we learned how triggers are used to select interesting data events while data is being collected. In this chapter, we will review which of the resulting datasets we use in this analysis.

We will start by reviewing the triggers we use to select a subset of the data collected by CMS, as well as how efficient these triggers are with respect to offline kinematic variables. Next, we will discuss how our simulated event samples are generated, and which processes we include to simulate our standard model background. Finally, we will briefly review the processing pipeline that samples go through before we analyze them.

6.1 Data

The CMS Run 2 data sample was collected across 3 years—2016, 2017, and 2018—-and corresponds to $137\,\mathrm{fb}^{-1}$ of total integrated luminosity. Each year's data collection was divided into eras, with detector and beam conditions remaining constant within each era. The luminosity collected in each era for each year is shown in Table 6.1. These eras consist of many runs, which correspond to one beam fill. To ensure that each event in the

Table 6.1: Integrated luminosity of the data eras used in this analysis.

Era	L [fb^{-1}]		
	2016	2017	2018
A	-	-	13.482
B	5.751	4.793	6.785
C	2.573	9.755	6.612
D	4.242	4.320	31.947
E	4.025	9.424	-
F	3.105	13.568	-
G	7.576	-	-
H	8.865	-	-
Total	35.922	41.680	58.826

dataset is well reconstructed, each CMS subsystem is required to monitor data quality and submit a list of verified runs. Thus, the final dataset only consists of runs that were verified by all detector subsystems.

As we discussed in Section 4.2.6, triggers are used to categorize datasets at CMS. This reduces the amount of data each analysis must process, as they only need to download a subset of the primary datasets. These datasets can be further reduced trivially by applying the logical OR of a collection of high-level triggers (HLT). The primary datasets used for this analysis are `MET`, `JetHT`, `SingleElectron` (`EGamma` in 2018), and `SingleMuon`. The HLT paths selected from these datasets are listed in Table 6.2. These triggers fall into three categories: lepton, lepton + H_{T} , and MET. The lepton triggers target isolated leptons with p_{T} requirements that range from 24-32 GeV depending on the lepton type. The lepton + H_{T} triggers allow the lepton p_{T} requirement to be reduced to 15 GeV by also requiring H_{T}—the sum of AK4 jet p_{T}—to be greater than 350–400 GeV. Finally, the MET triggers target events with missing transverse momentum, and require $p_{\mathrm{T}}^{\mathrm{miss}}$ to be greater than 100–120 GeV.

When studying trigger efficiencies, it is important to understand the difference between *online* and *offline* reconstruction. Online reconstruction occurs in two stages: L1

Table 6.2: List of high-level triggers used for 2016 (left) and 2017+2018 (right).

Lepton triggers	
IsoMu24	IsoMu24
IsoTkMu24	IsoMu27
IsoMu27	Mu50
Mu50	Ele27_WPTight_Gsf
Ele27_WPTight_Gsf	Ele35_WPTight_Gsf
Ele115_CaloIdVT_GsfTrkIdT	Ele115_CaloIdVT_GsfTrkIdT
Lepton + H_{T} triggers	
Mu15_IsoVVVL_PFHT350	Mu15_IsoVVVL_PFHT450
Mu15_IsoVVVL_PFHT400	Mu15_IsoVVVL_PFHT600
Mu50_IsoVVVL_PFHT400	Mu50_IsoVVVL_PFHT450
Ele15_IsoVVVL_PFHT350	Ele15_IsoVVVL_PFHT450
Ele15_IsoVVVL_PFHT400	Ele15_IsoVVVL_PFHT600
Ele50_IsoVVVL_PFHT400	Ele50_IsoVVVL_PFHT450
MET triggers	
PFMET110_PFMHT110_IDTight	PFMET120_PFMHT120_IDTight
PFMETNoMu110_PFMHTNoMu110_IDTight	PFMET120_PFMHT120_IDTight_PFHT60
PFMET120_PFMHT120_IDTight	PFMETNoMu120_PFMHTNoMu120_IDTight
PFMETNoMu120_PFMHTNoMu120_IDTight	PFMETNoMu120_PFMHTNoMu120_IDTight_PFHT60

and HLT. L1 reconstruction is performed very quickly (on the time scale of 4 µs) and only uses information from the calorimeters and, sometimes, the muon system. HLT reconstruction uses full detector information, but is limited to a processing time of about 200 ms, requiring trigger decisions to be made with partial event information. Offline reconstruction, on the other hand, is performed using full detector data, no constraint on processing time, and the full suite of reconstruction software and techniques described in Chapter 5. As a result of these differences, online variables may differ significantly from those calculated offline. Therefore, while the efficiency of a trigger will be a step function in its online target variable, it becomes a turn-on curve when using the corresponding offline variable. At a particular value of the offline variable, the trigger efficiency reaches a maximum value and remains constant for higher values of the offline variable. We will refer to the maximum efficiency reached as the *plateau efficiency* and the value of the

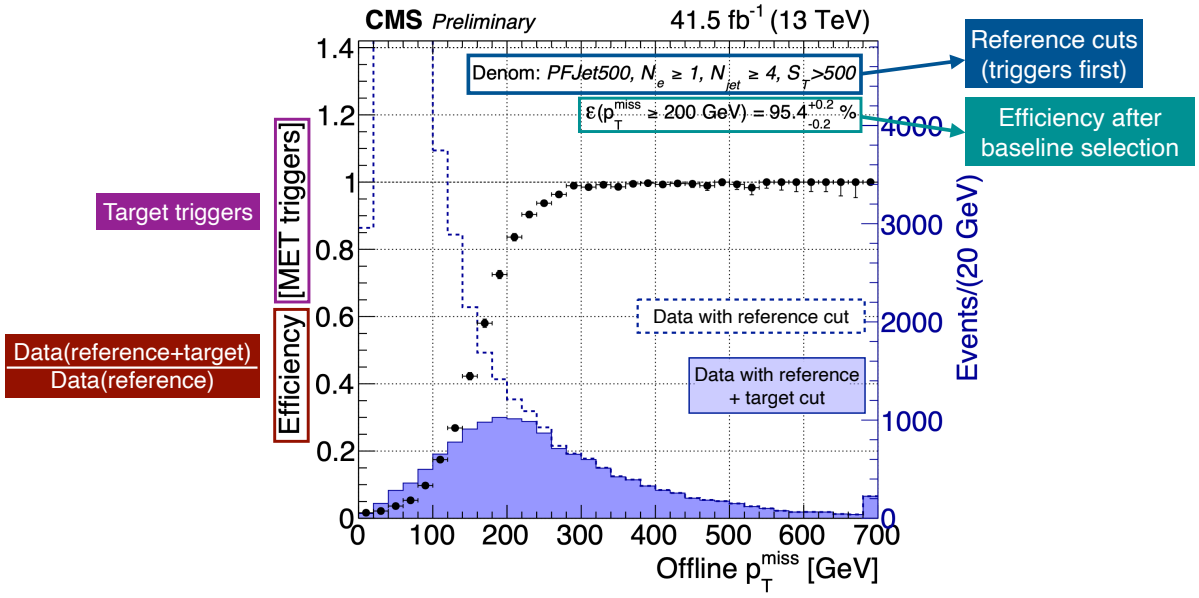


Figure 6.1: A fully documented trigger efficiency plot. The target triggers are specified in brackets on the y-axis, while the reference selection used to define the dataset is shown at the top of the plot in blue. The two distributions used to calculate the efficiency are also plotted, with the filled histogram representing the numerator, and the dashed histogram the denominator. The efficiency for the variable above the value used in the baseline selection is shown in green below the reference cuts.

offline variable at which this efficiency is reached as the *plateau threshold*. We will study both of these parameters for the lepton and $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers as a function of offline $p_{\mathrm{T}}^{\mathrm{miss}}$.

To determine the efficiency for a given trigger, which we will call the *target trigger*, we must first create a dataset to perform the study in. To avoid any bias, this dataset is constructed using an orthogonal trigger, which selects objects of a different type than the target trigger; we will call this the *reference trigger*. We also impose cuts on some offline variables to ensure the baseline trigger is fully efficient. Using this dataset, we compare the distributions of an offline variable before and after applying the target trigger. Taking the ratios of these distributions will give the efficiency of the target trigger as a function of the chosen variable. Figure 6.1 shows a typical plot resulting from this procedure, with the various features labeled.

Figures 6.2 and 6.3 show the $p_{\mathrm{T}}^{\mathrm{miss}}$ and single-lepton triggers as a function of $p_{\mathrm{T}}^{\mathrm{miss}}$

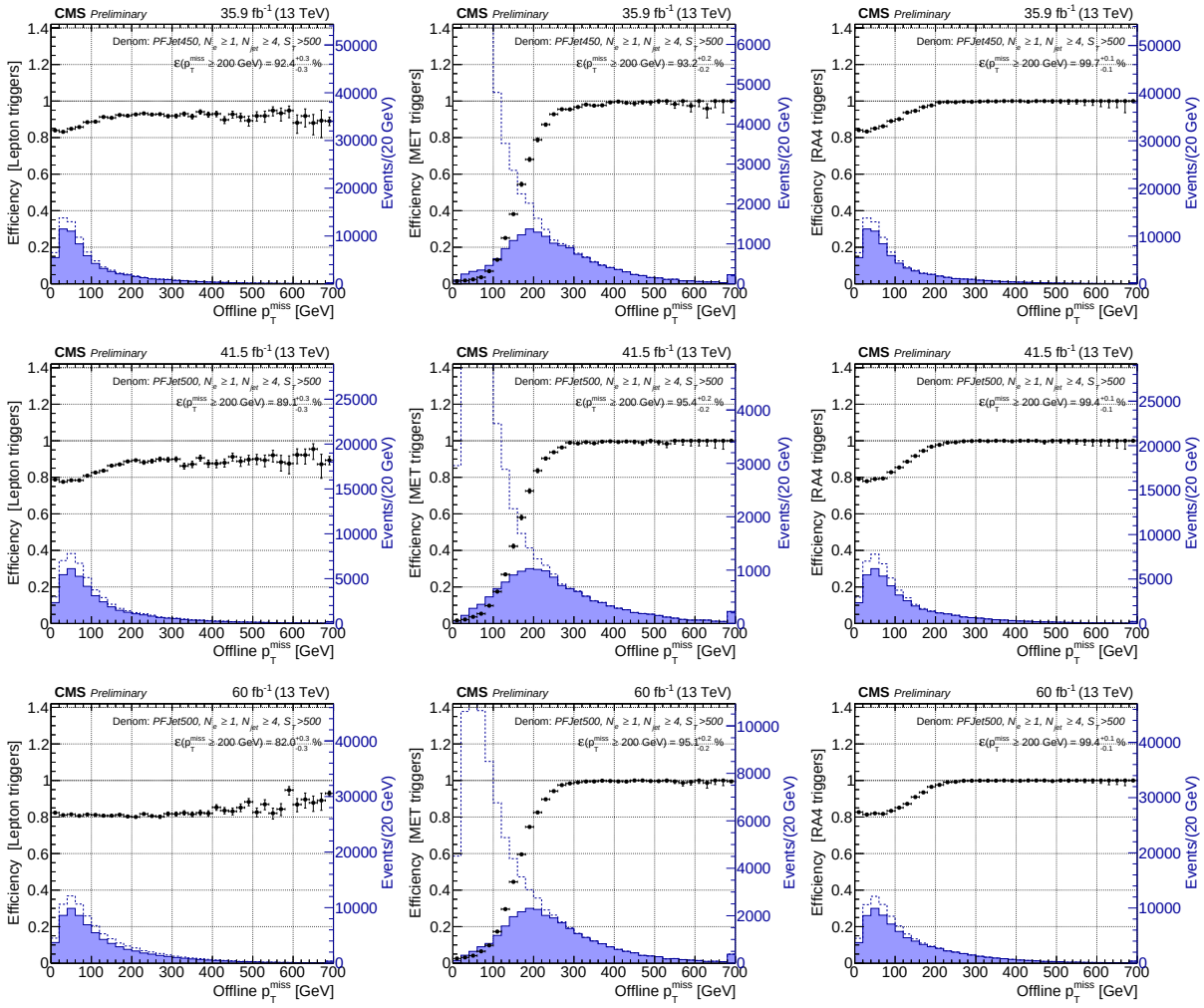


Figure 6.2: Trigger efficiencies as a function of $p_{\mathrm{T}}^{\mathrm{miss}}$ in electron events for the OR of all lepton triggers (left), all $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers (center), and all lepton and $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers together (right). Top, middle, and bottom rows correspond to 2016, 2017, and 2018 data. The efficiencies are measured using a data sample collected using the `HLT_PFJet450(500)` and `HLT_AK8PFJet450(500)` triggers for 2016 (2017 and 2018) data, and offline requirements of one electron, 4 or more jets, and $S_{\mathrm{T}} > 500\,\mathrm{GeV}$.

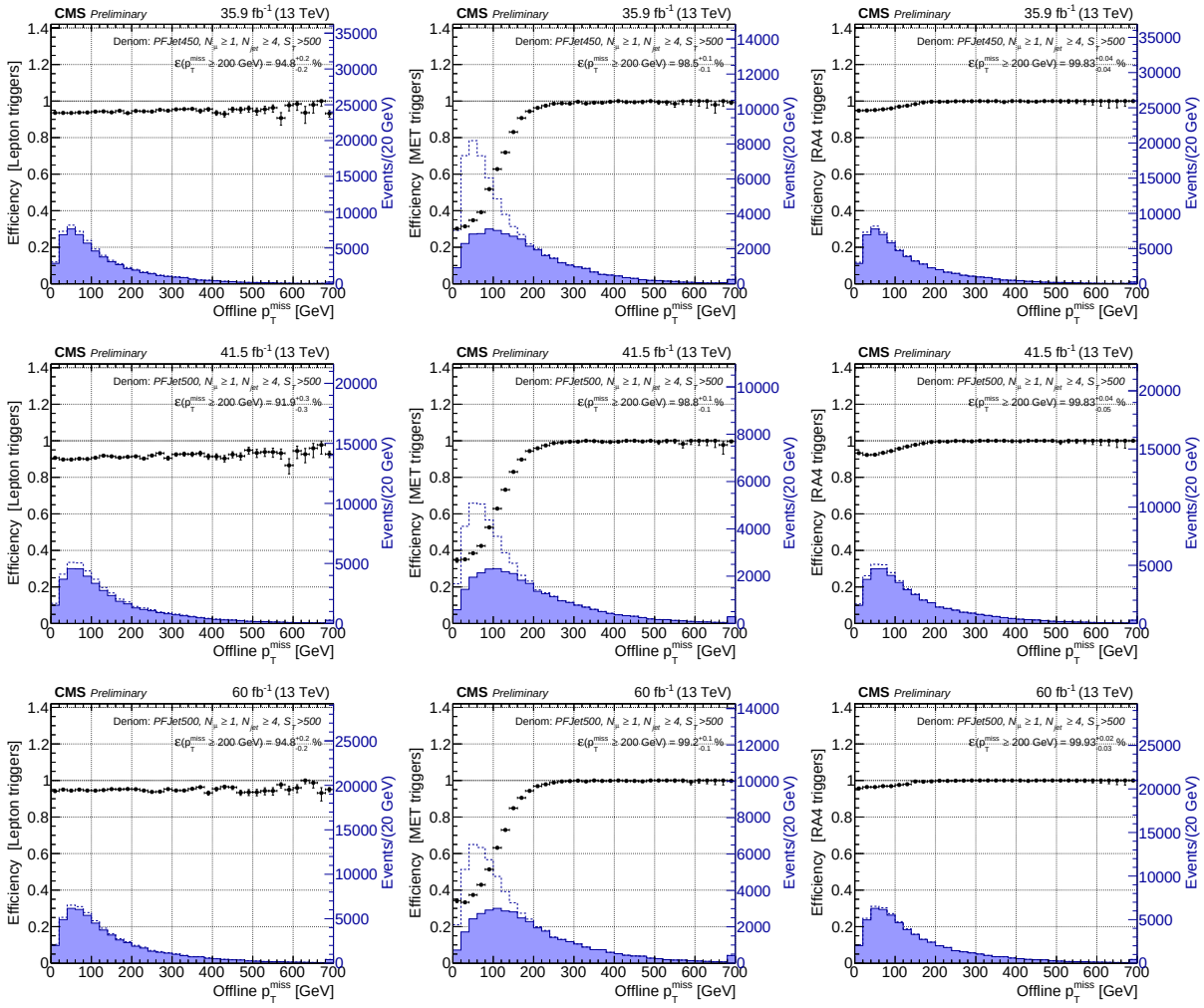


Figure 6.3: Trigger efficiencies as a function of $p_{\mathrm{T}}^{\mathrm{miss}}$ in muon events for the OR of all lepton triggers (left), all $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers (center), and all lepton and $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers together (right). Top, middle, and bottow rows correspond to 2016, 2017, and 2018 data. The efficiencies are measured using a data sample collected using the `HLT_PFJet450(500)` and `HLT_AK8PFJet450(500)` triggers for 2016 (2017 and 2018) data, and offline requirements of one muon, 4 or more jets, and $S_{\mathrm{T}} > 500\,\mathrm{GeV}$.

for each year of Run 2 for the electron and muon channels, respectively. The trigger efficiency is shown as a function of $p_{\mathrm{T}}^{\mathrm{miss}}$ for the single-lepton triggers in the left plots and the $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers in the middle plots. Though the $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers by themselves give a plateau efficiency of around 95-99% depending on the lepton channel, this only occurs above $p_{\mathrm{T}}^{\mathrm{miss}}$ values of 250–300 GeV. To lower this threshold, the $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers are combined with single-lepton triggers. Due to the lack of online $p_{\mathrm{T}}^{\mathrm{miss}}$ requirement, the single-lepton triggers are nearly constant as a function of $p_{\mathrm{T}}^{\mathrm{miss}}$. When the logical OR is taken between $p_{\mathrm{T}}^{\mathrm{miss}}$ and lepton triggers, shown on the right in the figures, the plateau efficiency is nearly 100% in both electron and muon channels, and occurs for $p_{\mathrm{T}}^{\mathrm{miss}}$ values above 200 GeV.

The efficiency of the $p_{\mathrm{T}}^{\mathrm{miss}}$ + lepton trigger scheme was also studied as a function of the other analysis variables, namely N_{jets}, N_{b}, M_J, and m_{T}. Shown for $p_{\mathrm{T}}^{\mathrm{miss}}$ ¿ 200 GeV in Figure 6.4 and Figure 6.5 for the electron and muon channels respectively, the trigger efficiency is near 100% and uniform in each variable.

Based on these results, the overall systematic in the trigger efficiency is estimated to be 0.5%.

6.2 Monte Carlo Simulation

Different aspects of MC events are taken care of by different programs in a multi-step generation process. The parton distribution function (PDF) of the colliding protons, which influences the kinematic properties and cross-sections of the to-be-generated process, is taken from NNPDF 3 [44]. Next, hard scattering process is simulated at leading order (LO) or next-to-leading order (NLO) with the MadGraph5_aMC@NLO [45] or POWHEG [46] MC event generators. The particles produced by the hard scattering or any subsequent decays are passed to the Pythia 8.2 [47] generator, which simulates showing and fragmentation of partons using the CUETP8M1 tune [48] for the underlying

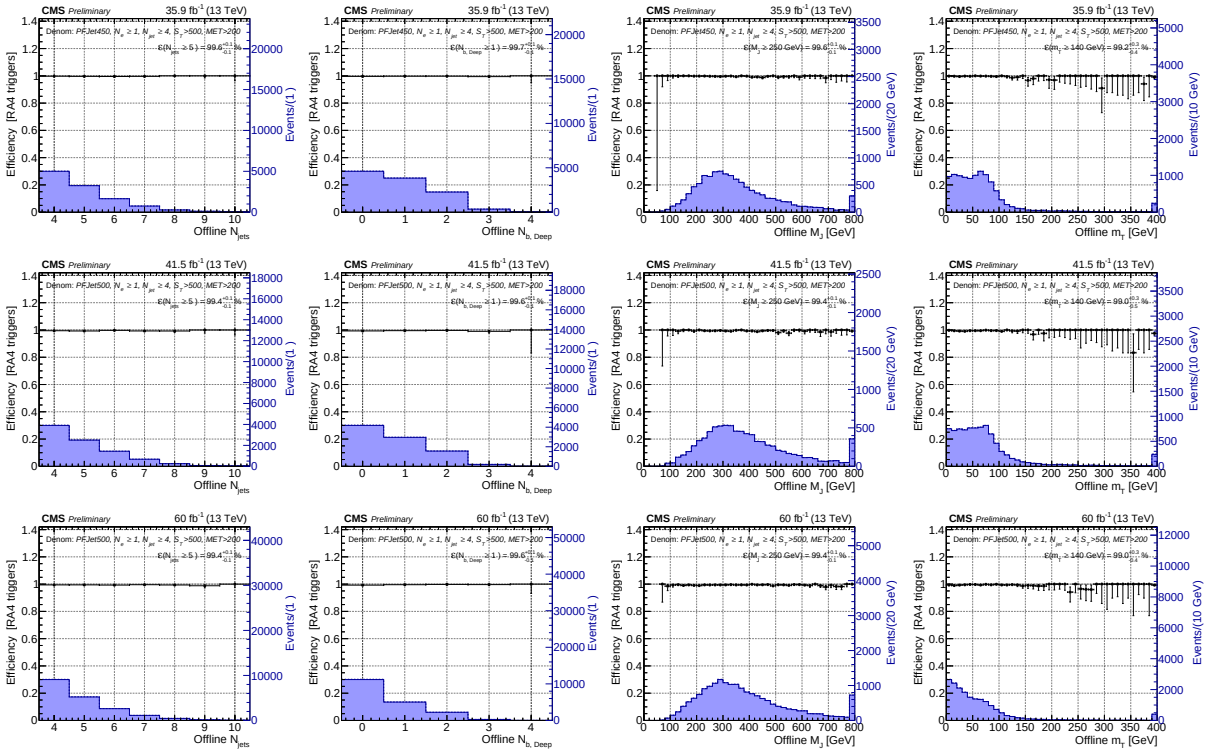


Figure 6.4: Trigger efficiencies in electron events as a function of N_{jets}, N_{b}, M_J, and m_{T} from left to right. Top, middle, and bottow rows correspond to 2016, 2017, and 2018 data. The efficiencies are measured using a data sample collected using the `HLT_PFJet450(500)` and `HLT_AK8PFJet450(500)` triggers for 2016 (2017 and 2018) data, and offline requirements of one electron, 4 or more jets, and $S_{\text{T}} > 500\,\text{GeV}$ and $p_{\text{T}}^{\text{miss}} > 200\,\text{GeV}$.

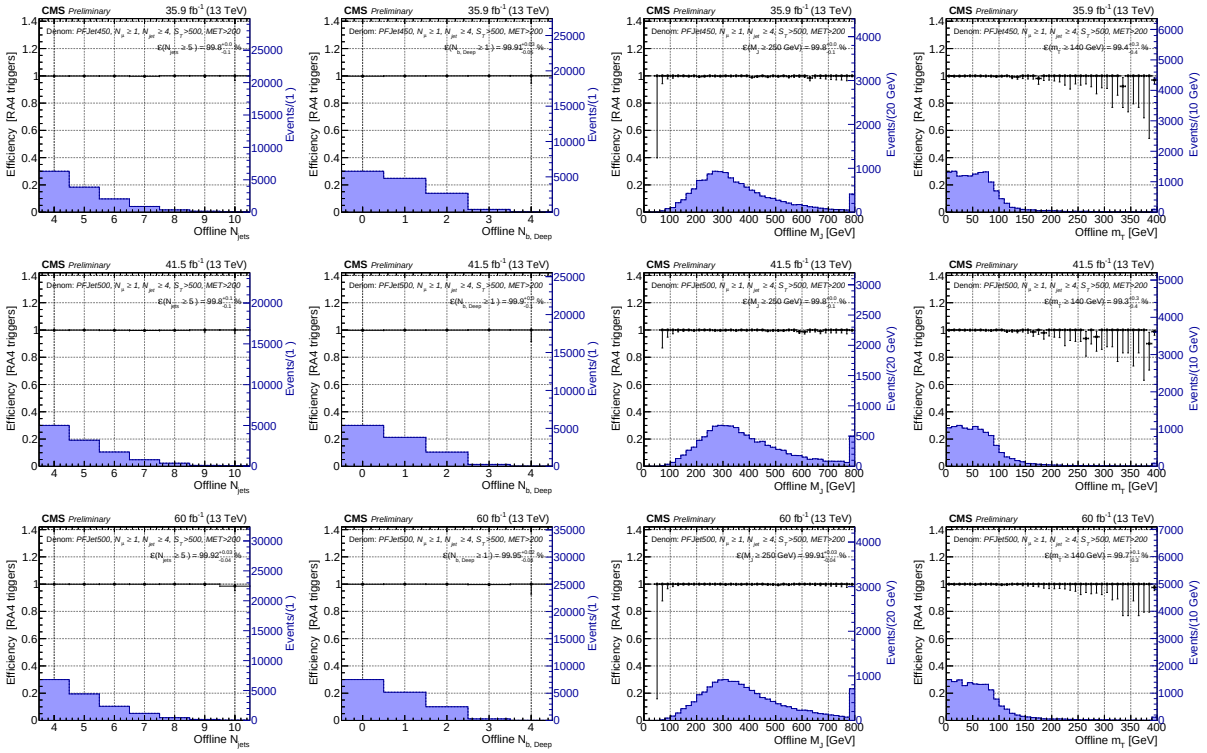


Figure 6.5: Trigger efficiencies in muon events as a function of N_{jets}, N_{b}, M_J, and m_{T} from left to right. Top, middle, and bottow rows correspond to 2016, 2017, and 2018 data. The efficiencies are measured using a data sample collected using the `HLT_PFJet450(500)` and `HLT_AK8PFJet450(500)` triggers for 2016 (2017 and 2018) data, and offline requirements of one electron, 4 or more jets, and $S_{\text{T}} > 500\,\text{GeV}$ and $p_{\text{T}}^{\text{miss}} > 200\,\text{GeV}$.

event model. The resulting final state particles are then processed by the GEANT4 4 [49] software package, which simulates the interactions between the particles and the CMS detector. For signal samples, the CMS fast simulation package [50] is used in lieu of GEANT4. To simulate pileup, additional minimum bias events are generated by Pythia 8.2 and overlaid on top of the underlying event.

The full list of standard model background samples used in this analysis is given in Table 6.3. The production of $t\bar{t}$ +jets, W+jets, Z +jets, and QCD are simulates using MadGraph5_aMC@NLO 2 in LO mode. Single top events are generated at NLO with MadGraph5_aMC@NLO for the s-channel and POWHEG v2 for the t-channel and W-associated production. Additional small backgrounds—such as $t\bar{t}$ production in association with bosons, $t\bar{t}t\bar{t}$, and diboson production—are produce at NLO using MadGraph5_aMC@NLO or POWHEG. For all processes, the highest order cross-section available is used. In the $t\bar{t}$ +jets sample, additional weights are applied based on the number of ISR jets; these weights were derived using a dilepton $t\bar{t}$ sample with two b tagged jets.

The simulated signal samples used are generated by MadGraph5_aMC@NLO 2 in LO mode, corresponding to the SMS-T1tttt_TuneCP2_13TeV-madgraphMLM-pythia8 and SMS-T5tttt_dM175_TuneCP2_13TeV-madgraphMLM-pythia8 samples. Each sample contains approximately 800 mass points, with gluino masses varying from 600 GeV (800 GeV) to 2600 GeV for T1tttt (T5tttt), and $\widetilde{\chi}_1^0$ masses varying from 1 GeV to 1600 GeV. As the detector response is simulated with the fast simulation package for signal events, additional scale factors are applied to account for any differences with respect to the full simulation used for the backgrounds.

Table 6.3: Dataset names for simulated event samples in 2016, with the grouping used in plot legends specified. The corresponding set of samples in 2017 and 2018 are used to model the 2017 and 2018 data, respectively.

Dataset name	Legend
TTJets_SingleLeptFromT_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 TTJets_SingleLeptFromT_genMET-150_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 TTJets_SingleLeptFromTbar_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 TTJets_SingleLeptFromTbar_genMET-150_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	$t\bar{t}$ (1ℓ)
TTJets_DiLept_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 TTJets_DiLept_genMET-150_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	$t\bar{t}$ (2ℓ)
WJetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 WJetsToLNu_HT-[binned]_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	W+jets
ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1 ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1 ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1 ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1 ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	Single t
TTGJets_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8 TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8 TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8 TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8 TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8 TTTT_TuneCUETP8M1_13TeV-amcatnlo-pythia8	$t\bar{t}\,V$
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-50_HT-[binned]_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 QCD_HT[X]_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 X = [200to300, 300to500, 500to700, 700to1000, 1000to1500, 1500to2000, 2000toInf] WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8 WZTo1L3Nu_13TeV_amcatnloFXFX_madspin_pythia8 WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8 WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8 ZJetsToNuNu_HT-[binned]_13TeV-madgraph ZZ_TuneCUETP8M1_13TeV-pythia8 ttHTobb_M125_13TeV_powheg_pythia8	Other
HT bins = [100To200, 200To400, 400To600, 600To800, 800To1200, 1200To2500, 2500ToInf]	

6.2.1 Validation of Monte Carlo with data

With the simulated background samples established, we verified that these samples model the key analysis variables well. This is done to check for any large discrepancies that would indicate either a significant difference in composition between data and simulation, or that the distributions of these variables are not modeled accurately in simulation. To perform this validation, we compare the distributions of six variables—S_{T}, $p_{\mathrm{T}}^{\mathrm{miss}}$, N_{jets}, N_{b}, M_J, and m_{T}—between data and simulation after applying the baseline selection (using a flat N_{jets} requirement of $N_{\mathrm{jets}} \geq 7$ for simplicity). The resulting plots are shown in Figure 6.6, with the full Run 2 data plotted as points and combined simulation represented with a stacked histogram that is normalized to the data yield in each plot.

Overall, there is good agreement between data and simulation. Normalization factors are within 5% for all distributions, ranging from 77-82%. Simulation shows slightly hard spectra for S_{T}, $p_{\mathrm{T}}^{\mathrm{miss}}$, and M_J, as well as a higher average jet multiplicity. A small excess is observed in the m_{T} distribution between 100-200 GeV. Upon further study, this effect was found to be correlated with pileup, becoming more pronounced in the high-pileup runs of 2017 and 2018. Any effect this may have on the background prediction is accounted for using the procedure detailed in Chapter 9.

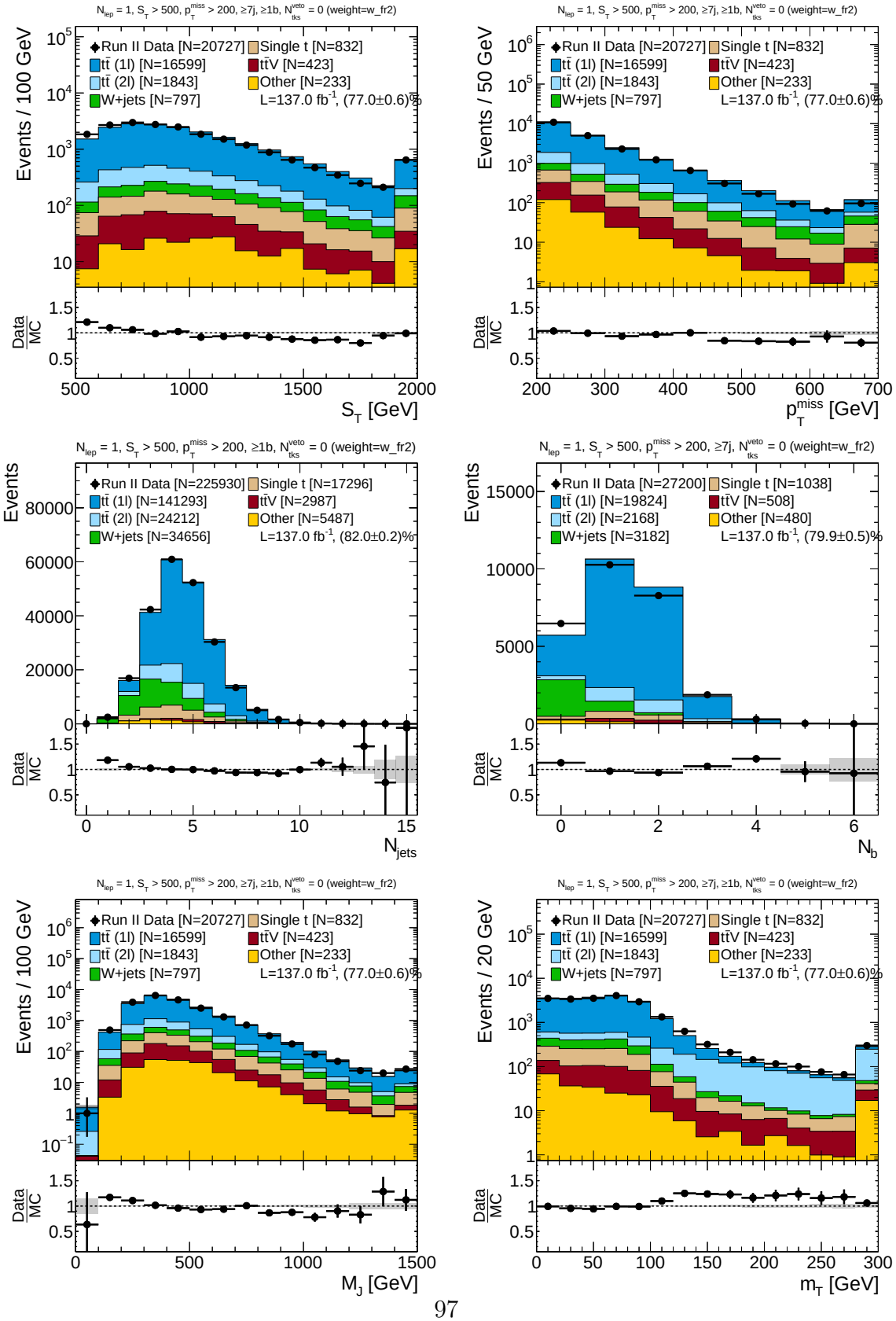


Figure 6.6: Comparison of data vs. MC for key variables. S_{T} and $p_{\mathrm{T}}^{\mathrm{miss}}$ are at the top, N_{jets} and N_{b} in the middle, and M_J and m_{T} in the bottom row.

Chapter 7

Experimental Signatures and Event Selection

This chapter introduces the supersymmetry models targeted in this analysis, focusing on the features that will be used to distinguish them from standard model backgrounds. These features will be used to develop a series of cuts—referred to as the *baseline selection*—that are designed to remove most of the standard model background while retaining as much signal as possible. The remaining backgrounds are classified using the key kinematic variables, m_{T} and M_J.

7.1 Signal Features and the Baseline Selection

As discussed in Chapter 3, searches for SUSY typically target particular simplified models to search for. SUSY models shown in Chapter 3, this analysis will focus on the two shown in Figure 7.1. In each of these processes, denoted as T1tttt and T5tttt, an interaction between the incoming partons produces a pair of gluinos ($\widetilde{\mathrm{g}}$) with mass $m_{\widetilde{\mathrm{g}}}$. These gluinos subsequently decay via an intermediate top squark into a top-antitop

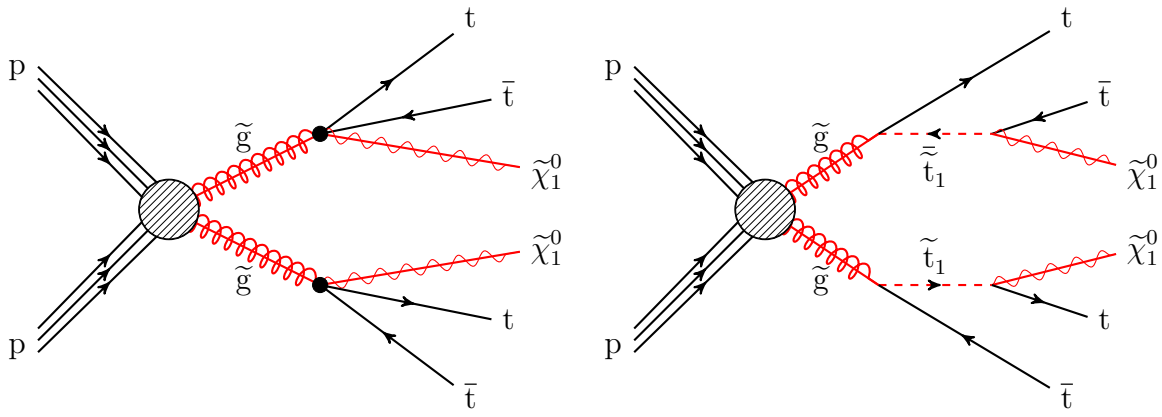


Figure 7.1: Diagrams of the T1tttt (left) and T5tttt (right) simplified models, with the supersymmetric particles shown in red. Both models give two neutralinos and four top quarks, with the T5tttt model including an on-shell top squark in the intermediate state.

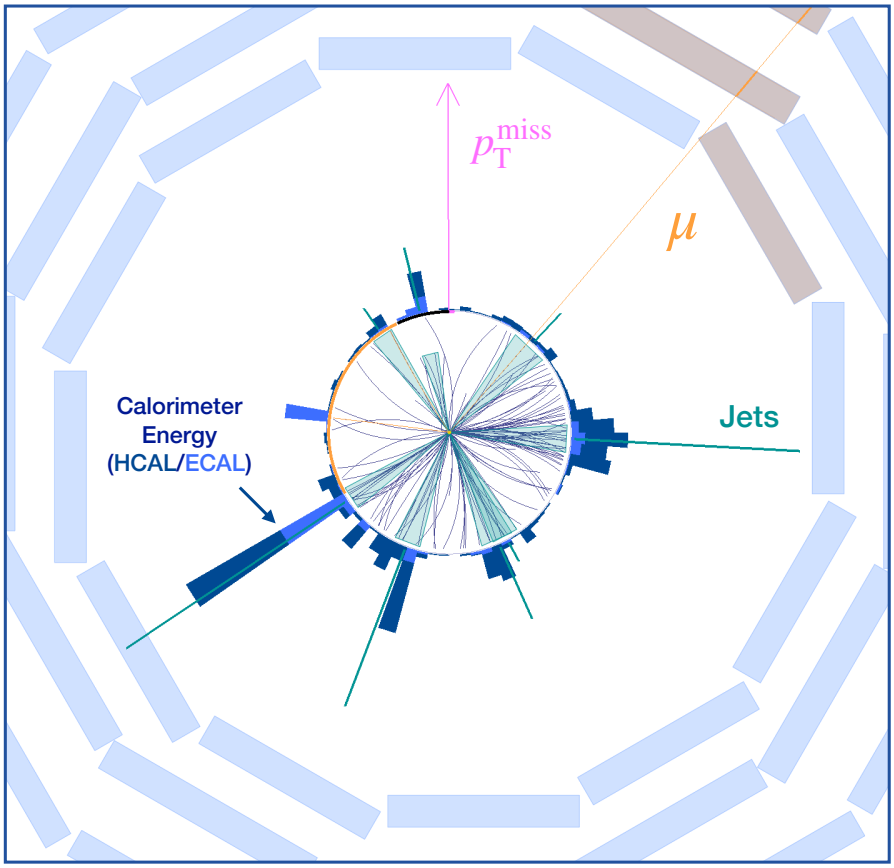


Figure 7.2: A simulated detector response to a T1tttt event, demonstrating the typical features of a signal event. Jet cones are highlighted in teal in the tracker region, while the energy deposits are shown as radial bars outside.

pair and a neutralino ($\widetilde{\chi}_1^0$) with mass $m_{\widetilde{\chi}_1^0}$. The $\widetilde{\chi}_1^0$ is the lightest supersymmetric particle (LSP), making it stable. The difference between the two models is whether the intermediate top squark is on (T5tttt) or off (T1tttt) its mass shell. These models are considered for a range of gluino and LSP masses, where a specific model is denoted as T1tttt($m_{\widetilde{g}}$,$m_{\widetilde{\chi}_1^0}$), with the masses in units of GeV. In T5tttt, the mass of the top squark is set to be greater than the mass of the neutralino by the mass of the top quark (175 GeV). A simulated detector response to a signal event is shown in Figure 7.2. This event exhibits many of the key features of the T1tttt and T5tttt final states, which we will discuss in the following paragraphs.

The presence of neutralinos in the final state gives these SUSY models the distinct feature of missing transverse momentum ($p_{\mathrm{T}}^{\mathrm{miss}}$). In the T1tttt and T5tttt models, the amount of missing momentum depends on the difference between $m_{\widetilde{g}}$ and $m_{\widetilde{\chi}_1^0}$, leading to two model benchmark models: compressed (C) and non-compressed (NC). The compressed model, T1tttt(1900,1250), represents scenarios where the mass splitting between the gluino and neutralino is small, and will have a softer $p_{\mathrm{T}}^{\mathrm{miss}}$ spectrum. The non-compressed model, T1tttt(2100,100), represents scenarios with a large mass difference, and will have a much harder $p_{\mathrm{T}}^{\mathrm{miss}}$ spectrum. The $p_{\mathrm{T}}^{\mathrm{miss}}$ signature of the signal models is the primary feature used to distinguish them from the standard model background, with $p_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$ required in the baseline selection. The remaining features selected for in the baseline selection are due to the decay of the four top quarks.

The dominant decays of the top quark are shown in Figure 7.3. Due to the small values of the off-diagonal CKM elements for the top quark, it will almost always decay into a b quark and a W boson. The W will then decay into either a lepton-neutrino pair or a pair of quarks (one up-type quark and one down-type antiquark). Given top decays produce a charged lepton 26% of the time, we can calculate the number of leptons (electrons or muons) we would expect to observe in the final state of a four-top system.

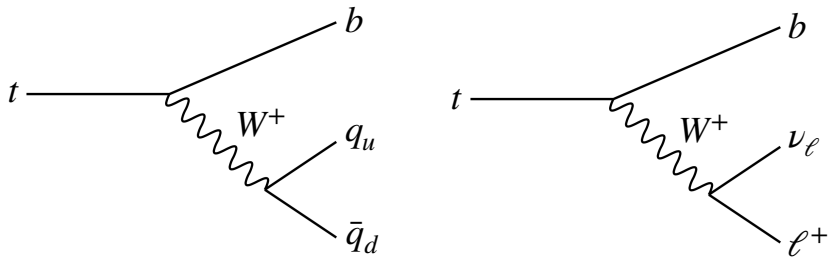


Figure 7.3: Primary decay modes of the top quark. The top will almost always produce a b quark and a W boson. The W will either decay into a pair of quarks (left) or a lepton-neutrino pair (right).

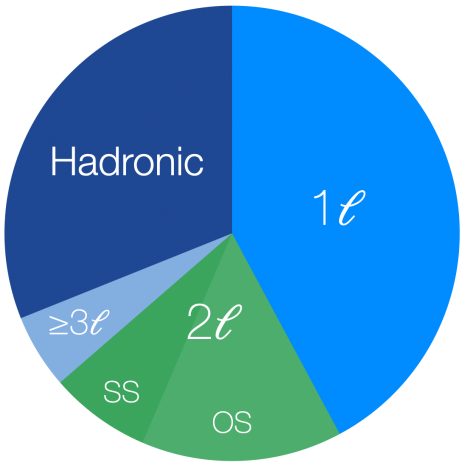


Figure 7.4: Pie chart of the lepton multiplicities expected from the decay of a four top system, assuming the top decays leptonically 26% of the time.

Table 7.1: Cuts applied in the baseline selection.

Object	Baseline selection
Missing momentum	$p_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$
Muons	$N_{\mathrm{leps}} = 1$
Electrons	
Veto tracks	$N_{\mathrm{veto}} = 0$
Jets	$N_{\mathrm{jets}} \geq 7$ (6 for $p_{\mathrm{T}}^{\mathrm{miss}} > 500\,\mathrm{GeV}$)
b jets	$N_{\mathrm{b}} \geq 1$
Total energy	$S_{\mathrm{T}} > 500\,\mathrm{GeV}$

Represented as a pie chart, this distribution is shown in Figure 7.4. With a probability of 42%, the single lepton final state is the most likely, and thus is required in the baseline selection. Furthermore, outside of the produced lepton, we do not expect any other isolated tracks to occur in signal events. Therefore, a track veto is also imposed, which requires there to be no additional isolated tracks in the event. The tracks considered for this cut were defined in Section **??**.

In addition to the potential for lepton production, a four-top final state will also give a large number of jets, four of which originate from bottom quarks. Furthermore, we expect these jets to have high energies. Considering the minimum-energy configuration in which all four top quarks are produced at rest, there is still ∼700 GeV of energy to be distributed among the decay products, most of which will be measured as jets. Therefore, on top of high jet (and b jet) multiplicity, we expect the sum of jet energies to be large as well. The variable S_{T}, defined in Eq. (**??**), captures this kinematic information, and is required to be greater than 500 GeV in the baseline selection.

In summary, the T1tttt and T5tttt models are targeted with the baseline selection, which requires a large missing transverse momentum; a single identified lepton with no other isolated tracks; multiple jets, at least one of which is tagged as a b jet; and a large sum of jet and lepton transverse momenta. These features result in the baseline selection,

shown in Table 7.1. Though the baseline selection rejects most of the standard model background, there are some processes which pass these cuts.

7.2 Standard Model Background

7.2.1 Effect of the baseline selection

Table 7.2 shows the event yields of each simulated event sample as selections are made. After the initial requirement of one lepton, S_{T} >500 GeV, and $p_{\mathrm{T}}^{\mathrm{miss}}$ >200 GeV, there are $\sim$ 840000 standard model background events. The background is then reduced to $\sim$ 28000 events by the remaining baseline requirements, with the N_{jets} selection removing the highest fraction of events (95%). Of these remaining events, the majority (88%) arise from top anti-top pair production ($\mathrm{t\bar{t}}$). Production of a single top quark (Single t) or a W boson in association with jets (W+jets) each make up about 4%. Figure 7.6 shows event displays for $\mathrm{t\bar{t}}$ and W+jets events that pass the baseline selection.

As $\mathrm{t\bar{t}}$ makes up such a large fraction of the background, it is divided into two categories: events where only one of the top quark decays produces a lepton (1ℓ $\mathrm{t\bar{t}}$), and events where both top quark decays produce leptons (2ℓ $\mathrm{t\bar{t}}$). The additional lepton in 2ℓ $\mathrm{t\bar{t}}$ events differentiates them from 1ℓ $\mathrm{t\bar{t}}$ events in two ways. First, this lepton can be reconstructed as an isolated track, causing the event to fail the track veto. Thus, the track veto rejects 38% of 2ℓ $\mathrm{t\bar{t}}$ events and only 8% of 1ℓ $\mathrm{t\bar{t}}$ events. Second, the additional lepton takes the place of a jet-producing quark, resulting in the N_{jets} requirement removing a greater fraction of 2ℓ $\mathrm{t\bar{t}}$ events. Furthermore, the 2ℓ $\mathrm{t\bar{t}}$ events remaining will have a greater fraction of ISR jets than 1ℓ $\mathrm{t\bar{t}}$ events.

The second-largest background contributions come from the production of either a single top quark (Single t) or a W boson in association with jets (W+jets).

Table 7.2: Event yields obtained from simulated event samples, as the event selection criteria are applied. The category *Other* includes Drell-Yan, $t\bar{t}t\bar{t}$, Z+jets, ZZ, $W^{\pm}Z$, $W^{\pm}W^{\pm}$, $t\bar{t}H(\rightarrow b\bar{b})$. The category $t\bar{t}V$ includes $t\bar{t}W^{\pm}$, $t\bar{t}Z$, and $t\bar{t}\gamma$. The event selection requirements listed above the horizontal line in the middle of the table make up the baseline selection. The background estimates before the S_{T} requirement are not specified because some of the simulated event samples do not extend to the low S_{T} region. Given the size of the MC samples described in Section 6.2, rows with zero yield have statistical uncertainties of at most 0.15 events, and below 0.10 events in most cases.

$\mathcal{L} = 137.0$ fb^{-1}	Other	QCD	$t\bar{t}\,V$	Single t	W+jets	1ℓ $t\bar{t}$	2ℓ $t\bar{t}$	SM Bkg.	Eff.	NC	C
$1\ell, S_T > 500$ GeV, $p_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV	29996.8	14127.7	5503.1	32683.5	465578.8	233719.1	57661.4	839270.4	-	28.3	66.8
Track veto	22905.3	13047.6	4599.9	28802.4	445636.2	214954.3	36079.4	766025.1	0.91	25.9	60.3
$N_{\mathrm{jets}} \geq 7$ (6 if $p_{\mathrm{T}}^{\mathrm{miss}}$ >500 GeV)	454.6	202.7	657.7	1425.1	4610.8	25358.7	2811.0	35520.7	0.05	20.8	47.4
$N_{\mathrm{b}} \geq 1$	212.7	107.4	566.4	1178.1	1168.7	21976.5	2473.1	27682.9	0.80	19.8	46.0
$M_J > 250$ GeV	197.8	107.4	528.5	1076.5	1058.2	20393.8	2205.6	25567.9	0.93	19.8	45.7
$m_{\mathrm{T}} > 140$ GeV	22.6	19.3	51.3	68.7	36.2	333.3	860.5	1391.9	0.05	16.8	34.7
$M_J > 400$ GeV	15.9	11.3	30.5	49.0	25.4	230.0	441.3	803.4	0.58	16.6	29.8
$N_{\mathrm{b}} \geq 2$	8.6	1.4	15.4	24.1	5.9	113.9	242.2	411.5	0.52	13.0	25.5
$p_{\mathrm{T}}^{\mathrm{miss}} > 350$ GeV and $M_J > 450$ GeV	1.7	0.5	3.2	2.7	1.2	9.5	48.1	66.9	0.17	11.8	16.1
$p_{\mathrm{T}}^{\mathrm{miss}} > 500$ GeV and $M_J > 500$ GeV	0.5	0.0	0.9	0.7	0.5	2.4	13.4	18.4	0.28	10.0	7.9
$N_{\mathrm{jets}} \geq 8$	0.2	0.0	0.3	0.1	0.1	1.1	4.2	6.0	0.35	5.8	5.6
	Other	QCD	$t\bar{t}\,V$	Single t	W+jets	1ℓ $t\bar{t}$	2ℓ $t\bar{t}$	SM Bkg.	Eff.	NC	C

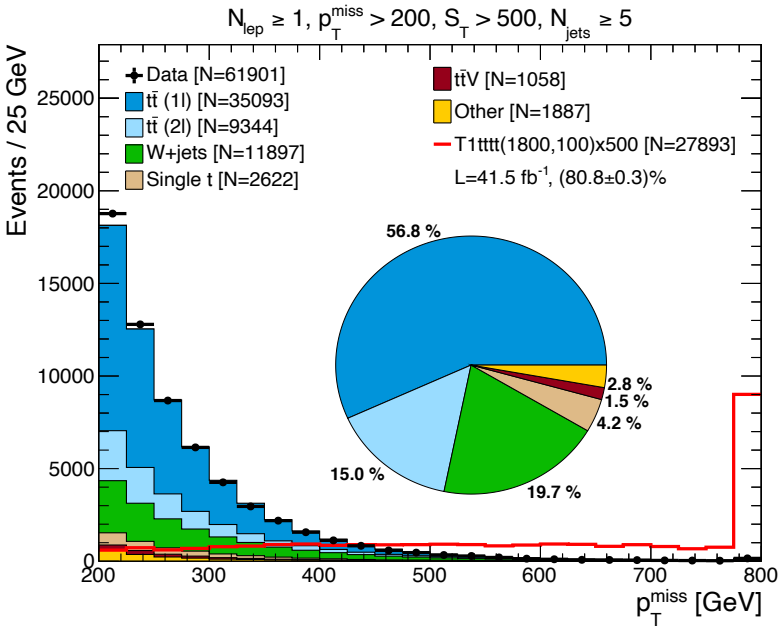


Figure 7.5: $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution for background and signal simulated events with data included. A pie chart is included in the plot to show the relative composition of the background after a loose selection. For visibility, the signal cross section was increased by a factor of 500.

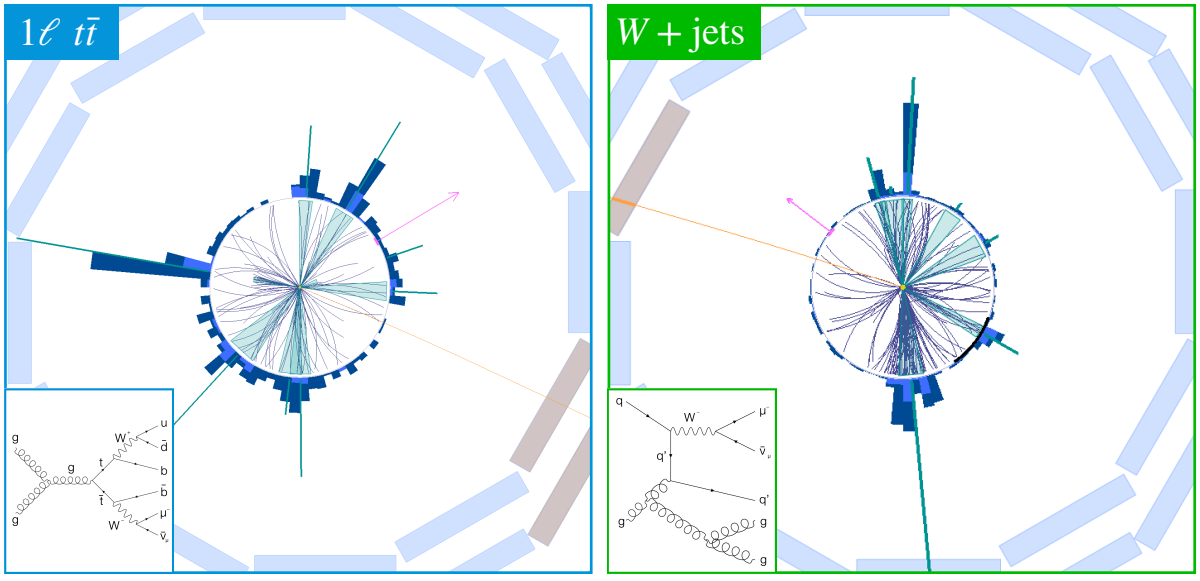


Figure 7.6: Event displays for two standard model backgrounds, 1ℓ t$\bar{\mathrm{t}}$ (left) and W+jets (right). Both events contain a lepton, missing transverse energy, and high jet multiplicity.

7.2.2 Removing single lepton events with m_{T}

In addition to selecting on the number of reconstructed objects and their properties, we also employ variables that represent underlying event kinematics. The first of these variables uses the lepton and $p_{\mathrm{T}}^{\mathrm{miss}}$ vectors to calculate a transverse mass. Referred to as m_{T}, this variable is defined as

$$m_{\mathrm{T}} = \sqrt{2p_{\mathrm{T}}^{\ell} p_{\mathrm{T}}^{\mathrm{miss}} [1 - \cos(\Delta\phi_{\ell, p_{\mathrm{T}}^{\mathrm{miss}}})]}, \tag{7.1}$$

where $\Delta\phi_{\ell, p_{\mathrm{T}}^{\mathrm{miss}}}$ is the difference between the azimuthal angles of the lepton momentum vector and the missing momentum vector. This variable is used to separate backgrounds characterized by the presence of a single W decaying leptonically and no other sources of missing energy. Assuming well-reconstructed events, backgrounds of this kind should have a transverse mass less than m_{W}. In reality, however, effects such as $p_{\mathrm{T}}^{\mathrm{miss}}$-resolution cause this variable to be larger than m_{W}. Therefore, a value of 140 GeV is used to separate events into two m_{T} categories. As seen in Table 7.2, applying this selection on m_{T} removes 98% of 1ℓ $\mathrm{t\bar{t}}$ events. A distribution of m_{T} for simulated event samples is shown in Figure 7.7.

7.2.3 M_J and initial-state radiation

As we saw in Section 5.3, in addition to the standard $R = 0.4$ jets, this analysis also uses large-radius ($R = 1.4$) jets, which are constructed by clustering $R = 0.4$ jets and leptons. The sum of the masses of these large-R jets is denoted by M_J, defined in Eq. (5.7). For both signal and background, the shape of the M_J distribution is sensitive to whether or not the top decays are cleanly clustered into large-R jets. For signal models, this resolution depends on the difference in the masses of the gluino and LSP. Large mass differences allow for the decay products of the top quarks to be more collimated, and thus more likely to be clustered into a single large-R jet. Therefore, M_J will tend to have

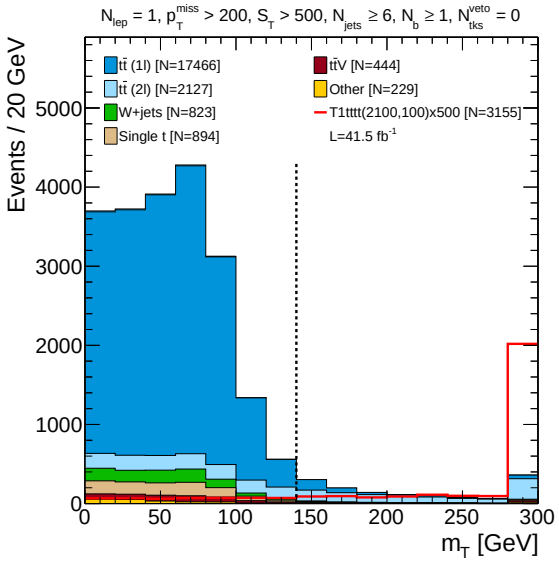


Figure 7.7: Distribution of m_{T} in 2017 simulated event samples after the baseline selection is applied. The vertical dashed line shows the selection made on m_{T} which defines the search region. For visibility, the signal cross section was increased by a factor of 500.

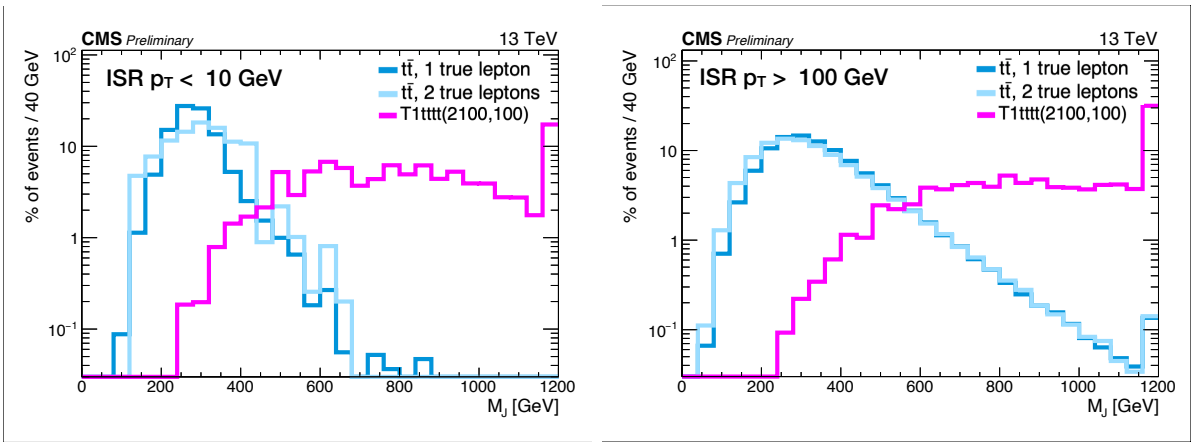


Figure 7.8: M_J distributions for simulated events with a negligible ISR contribution (left) and a significant ISR contribution (right).

higher values in non-compressed signal models.

Since M_J is used for the estimation of background in the search region, it is important to understand the background events that comprise the high-M_J region. In $\mathrm{t\bar{t}}$ +jets events, which make up the majority of the background after the baseline selection is

applied, M_J is very sensitive to the amount of initial-state radiation (ISR) present. Figure 7.8 shows the M_J distribution in $\mathrm{t\bar{t}}$ +jets events when the amount of ISR is negligible and when it is significant. Without ISR, the top quarks will be produced back-to-back in the lab frame, allowing for the products of each top to be grouped into a single large-R jet. In these cases M_J will have a maximum of twice the top mass, with a lower value if not all the decay products fit within a large-R jet. This behavior is reflected in the distribution, which drops off steeply above $2m_{\mathrm{t}}$.

Conversely, events with a significant amount of ISR do not have such an upper limit. Large amounts of ISR can give $M_J > 2m_{\mathrm{t}}$ through one of two mechanisms. First, the ISR jet can be aligned with the top decay products and included in the reconstructed large-R jet. Alternatively, the ISR can cause the products of the two top decays to overlap and be reconstructed into a single large-R jet. The resulting large-R jet mass may exceed m_{t} in both these cases, resulting in $M_J > 2m_{\mathrm{t}}$. While the M_J shape for low-ISR events steeply drops above $2m_{\mathrm{t}}$, the high-ISR shape has a tail. It is also worth noting that, because the mechanisms through which $\mathrm{t\bar{t}}$ events reach high M_J are due to ISR jets, the shape of the tail in high-ISR events is very similar between 1ℓ and 2ℓ $\mathrm{t\bar{t}}$ samples. The independence of M_J and number of leptons is critical to the background estimation of this analysis, which will be described in the next chapter.

Chapter 8

Background Estimation

Last chapter we explored the final state given by our signal models, and used these features to develop the baseline selection to reduce the standard model background. No set of cuts, however, can remove all background. Before we can determine if there is any signal present in data, we must predict the amount of background we expect in our search region. In this chapter, we will introduce the ABCD background estimation method, starting with the foundational principles. We will then discuss the mathematical details of this prediction method as well as improvements that were made to increase signal sensitivity. Finally, we will derive the likelihood function used to determine the background yield.

8.1 Extrapolating from Low m_{T} to High m_{T}

Last chapter, we saw that by cutting on m_{T} we reject most 1ℓ $\mathrm{t\bar{t}}$ events, leaving a search region dominated by 2ℓ $\mathrm{t\bar{t}}$ events in which one of the leptons is lost. The question of background estimation then becomes: can we predict the amount of background in the high-m_{T} region using the background in the low-m_{T} region? Given these regions are

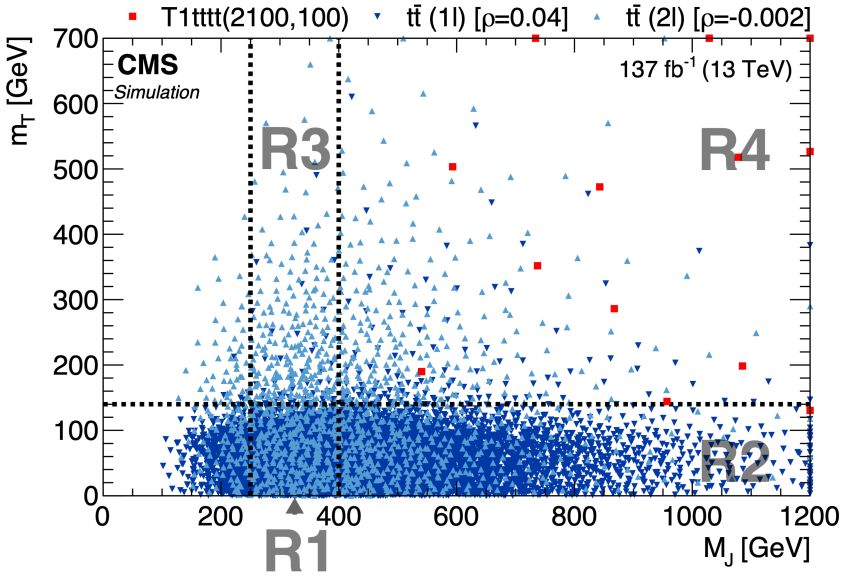


Figure 8.1: Distribution of simulated single-lepton $t\bar{t}$ events (dark-blue triangles), dilepton $t\bar{t}$ events (light-blue inverted triangles), and T1tttt(2100,100) events (red squares) in the M_J-m_T plane after the baseline selection and at least 2 b-jets. Each marker represents one expected event at $137\,fb^{-1}$. Overflow events are placed on the edge of the plot. The values of the correlation coefficients ρ for each background process are given in the legend. Region R4, which is further split into smaller bins as described later, is the nominal signal region, while R1, R2, and R3 serve as control regions. Note that the boundary between R1/R3 and R2/R4 is p_T^{miss}-dependent, the line shown at 400 GeV corresponds to the lowest p_T^{miss} bin.

dominated by different processes, the answer is not obvious.

To make this extrapolation, we need to relate the number of low-m_T 1ℓ $t\bar{t}$ events with the number of high-m_T 2ℓ $t\bar{t}$ events. This can be accomplished by finding a variable uncorrelated with m_T in which the two backgrounds have a similar shape. In Section 7.2.3, we saw that 1ℓ and 2ℓ $t\bar{t}$ have very similar M_J shapes in the high-ISR regime. Additionally, since M_J is primarily constructed from jets, it should be uncorrelated with m_T. Figure 8.1 shows the distribution of 1ℓ and 2ℓ $t\bar{t}$ events in the m_T-M_J plane after the baseline selection, with the correlation coefficients shown in legend. Indeed, the correlation between m_T and M_J is small ($\rho < 4\%$).

As an additional check of similarity between the two samples, Figure 8.2 compares

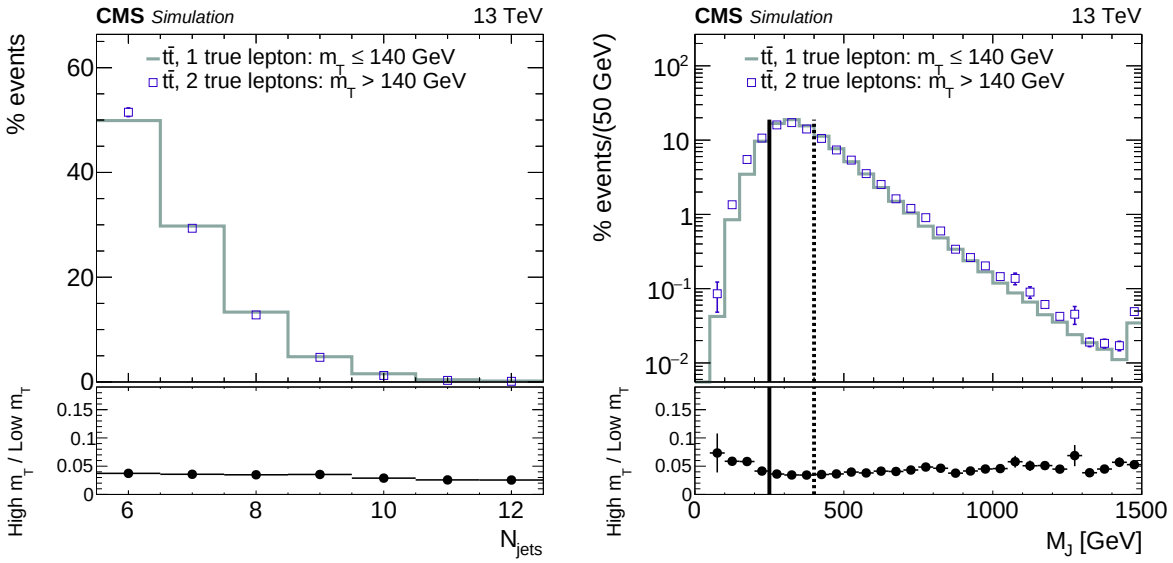


Figure 8.2: Comparison of N_{jets} and M_J distributions, normalized to the same area, in simulated $\text{t}\bar{\text{t}}$ events with two true leptons at high m_{T} and one true lepton at low m_{T}, after the baseline selection is applied. The shapes of these distributions are similar. These two contributions are the dominant backgrounds in their respective m_{T} regions. The dashed vertical line on the right-hand plot indicates the M_J >400 GeV threshold that separates the signal regions from the control samples. The region corresponding to M_J <250 GeV is not used in the background estimation.

the normalized N_{jets} and M_J distributions for high-m_{T} 2ℓ $\text{t}\bar{\text{t}}$ and low-m_{T} 1ℓ $\text{t}\bar{\text{t}}$ after the baseline selection. Despite having one less hadronically decaying W to produce jets, 2ℓ $\text{t}\bar{\text{t}}$ has a similar N_{jets} shape to 1ℓ $\text{t}\bar{\text{t}}$. This agreement arises from the fact that, due to the kinematic similarity between the initial states of 1ℓ and 2ℓ $\text{t}\bar{\text{t}}$ events, additional ISR jets are produced with similar probabilities in each sample. Therefore, even though 2ℓ $\text{t}\bar{\text{t}}$ requires two more ISR jets than 1ℓ $\text{t}\bar{\text{t}}$ to pass the $N_{\text{jets}} \geq 6$ requirement, the N_{jets} distribution will still have the same shape (though a lower overall event yield).

The agreement in N_{jets} means the number of objects contributing to M_J is comparable between the two samples, with the only difference being that 1ℓ $\text{t}\bar{\text{t}}$ events have a three jet system that reconstructs to a top quark. As we saw earlier, however, in the high-ISR regime the mass of large-R jets is dominated by the overlap between ISR jets and top decay products. Indeed, the M_J distribution after the baseline selection very much

resembles the high-ISR case in Figure 7.8. Thus, the M_J shapes of 1ℓ and 2ℓ $t\bar{t}$ agree extremely well after the baseline selection is applied. This agreement, coupled with the absence of a significant correlation between m_{T} and M_J, allows us to extrapolate the M_J distribution from low m_{T} to high m_{T}.

8.2 The ABCD Method

Now that we have shown that the M_J shapes are consistent between the backgrounds which dominate the low- and high-m_{T} regions, we can construct a background estimation method using M_J and m_{T}. This procedure, known as the ABCD method, involves partitioning the M_J-m_{T} plane into four regions: R1, R2, R3, and R4. These regions—diagrammed in Figure 8.3—are defined as

(a) Control Region R1: $250\,\mathrm{GeV} < M_J \leq 400\,\mathrm{GeV}$, $m_{\mathrm{T}} \leq 140\,\mathrm{GeV}$;

(b) Control Region R2: $M_J > 400\,\mathrm{GeV}$, $m_{\mathrm{T}} \leq 140\,\mathrm{GeV}$;

(c) Control Region R3: $250\,\mathrm{GeV} < M_J \leq 400\,\mathrm{GeV}$, $m_{\mathrm{T}} > 140\,\mathrm{GeV}$;

(d) Search Region R4: $M_J > 400\,\mathrm{GeV}$, $m_{\mathrm{T}} > 140\,\mathrm{GeV}$.

Most signal events are expected to reside in R4, and so this is denoted as the search region. The remaining three regions—R1, R2, and R3—are dominated by background, and will be used as control regions to predict the background yield in the search region.

Under the assumption that M_J and m_{T} are uncorrelated, the expected background yield in the search region can be expressed as

$$\mu_{\mathrm{R4}}^{\mathrm{bkg}} = \frac{\mu_{\mathrm{R2}}^{\mathrm{bkg}} \cdot \mu_{\mathrm{R3}}^{\mathrm{bkg}}}{\mu_{\mathrm{R1}}^{\mathrm{bkg}}}, \tag{8.1}$$

where μ^{bkg} denotes the expected background yield in the specified region. Using the maximum likelihood estimator for a single Poisson counting experiment in the control

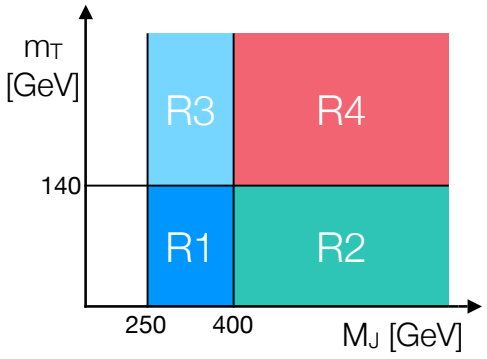


Figure 8.3: M_J-m_{T} plane separated into the four ABCD regions (R1, R2, R3, and R4).

regions ($\hat{\mu} = N$), the estimator for the background in the search region becomes

$$\hat{\mu}_{\mathrm{R4}}^{\mathrm{bkg}} = \frac{N_{\mathrm{R2}} \cdot N_{\mathrm{R3}}}{N_{\mathrm{R1}}}, \tag{8.2}$$

where N is the number of events measured in a region. If the number of events measured in the search region (N_{R4}) is significantly greater than $\hat{\mu}_{\mathrm{R4}}^{\mathrm{bkg}}$, the excess may be interpreted as signal and provide evidence for discovery. On the other hand, if N_{R4} is significantly less than $\hat{\mu}_{\mathrm{R4}}^{\mathrm{bkg}} + \hat{\mu}_{\mathrm{R4}}^{\mathrm{sig}}$, the data is incompatible with the signal model and this signal can be excluded.

The prediction in Eq. (8.2) arises in the simplest implementation of the ABCD method. It assumes that the variables used to separate in regions are completely uncorrelated, and that all signal events will inhabit the search region. In practice, however, neither of these assumptions hold, and a generalization of the method is required. Thus, we will reparameterize the expected yields in each region, including terms that account for signal contamination in the control regions. This gives

$$\mu_{\mathrm{R1}} = r \cdot \mu_{\mathrm{R1}}^{\mathrm{sig}} + \mu_{\mathrm{R1}}^{\mathrm{bkg}} = r \cdot \mu_{\mathrm{R1}}^{\mathrm{sig}} + \theta_1 \tag{8.3a}$$

$$\mu_{\mathrm{R2}} = r \cdot \mu_{\mathrm{R2}}^{\mathrm{sig}} + \mu_{\mathrm{R2}}^{\mathrm{bkg}} = r \cdot \mu_{\mathrm{R2}}^{\mathrm{sig}} + \theta_2 \tag{8.3b}$$

$$\mu_{\mathrm{R3}} = r \cdot \mu_{\mathrm{R3}}^{\mathrm{sig}} + \mu_{\mathrm{R3}}^{\mathrm{bkg}} = r \cdot \mu_{\mathrm{R3}}^{\mathrm{sig}} + \theta_3 \tag{8.3c}$$

$$\mu_{\mathrm{R4}} = r \cdot \mu_{\mathrm{R4}}^{\mathrm{sig}} + \mu_{\mathrm{R4}}^{\mathrm{bkg}} = r \cdot \mu_{\mathrm{R4}}^{\mathrm{sig}} + (\theta_3/\theta_1) \cdot \theta_2 \cdot \kappa. \tag{8.3d}$$

In these expressions, μ^{sig} are the expected number of signal events in each region, determined using MC; r is the signal strength modifier, with a value of 0 corresponding to the absence of signal and a value of 1 corresponding to its presence with the nominal cross section; θ are the expected background yields in each control region; and κ is a correction factor from MC that accounts for any correlation between M_J and m_{T}, given by

$$\kappa = \frac{N_{\text{R4}}^{\text{bkg}}/N_{\text{R2}}^{\text{bkg}}}{N_{\text{R3}}^{\text{bkg}}/N_{\text{R1}}^{\text{bkg}}}. \tag{8.4}$$

A κ value of 1 indicates no correlation between M_J and m_{T}; in this case, the expected background yield in R4 reduces to Eq. (8.2).

The nine parameters of Eq. (8.3) are fit to the observed ABCD yields in three samples—data, signal MC, and background MC. We will see later how exactly these parameters are estimated, but it is approximately done as follows: expected signal yields $\mu_{\text{Ri}}^{\text{sig}}$ are determined from the simulated signal yields $N_{\text{Ri}}^{\text{sig}}$; θ_1, θ_2, and θ_3 are constrained by the control region yields in data $N_{\text{R1}}^{\text{data}}$, $N_{\text{R2}}^{\text{data}}$, and $N_{\text{R3}}^{\text{data}}$; the signal strength r is determined from the search region in data $N_{\text{R4}}^{\text{data}}$; and κ is obtained from simulated background yields $N_{\text{Ri}}^{\text{bkg}}$ as shown in Eq. (8.4).

8.3 Additional Binning

We have established a method for measuring the signal strength using the yields observed in regions of the M_J-m_{T} plane. For a given signal model, the sensitivity of this method to the presence of signal depends on two factors: the number of background events in the search region and the accuracy with which this background is predicted. While the number of background events can be reduced by tightening the requirements in the baseline selection, doing so would also lower the amount of background in the control regions, thereby decreasing the accuracy of the estimation method. One way to alleviate this dilemma is to separate the events passing the baseline selection into bins of

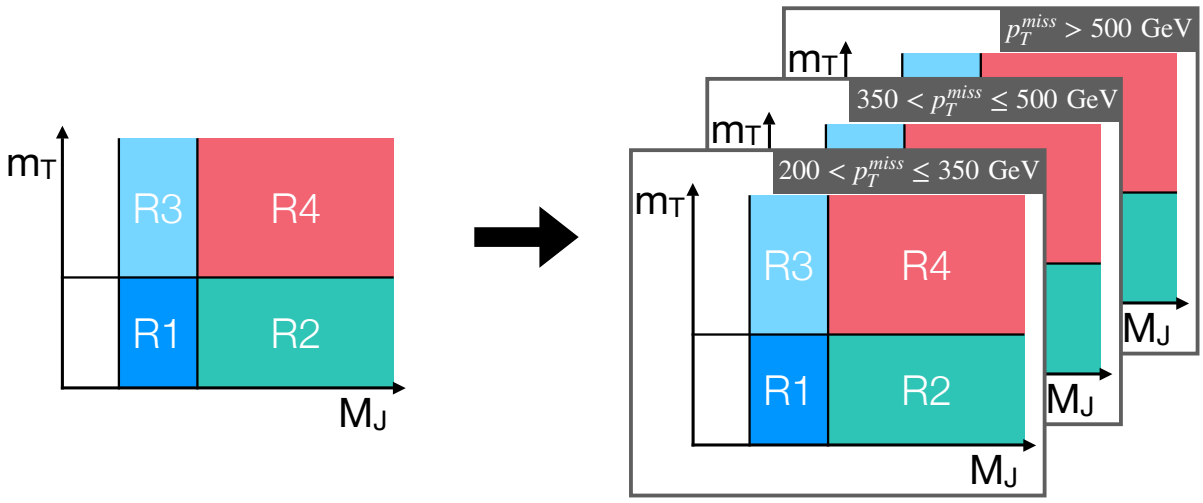


Figure 8.4: M_J-m_{T} plane separated into three $p_{\mathrm{T}}^{\mathrm{miss}}$ bins. In each bin, a separate ABCD background estimation is performed.

$p_{\mathrm{T}}^{\mathrm{miss}}$, N_{jets}, and N_{b}. Bins with tighter requirements will have lower background yields—as shown in Table 7.2—while bins with looser requirements maintain the accuracy of the background estimation.

When introducing bins into the ABCD method, there are two approaches that can be taken. The first is to bin the events after the baseline selection and then perform a separate ABCD background estimation in each bin. This binning method is used to separate events into three $p_{\mathrm{T}}^{\mathrm{miss}}$ bins, as shown in Figure 8.4. Though this approach is straightforward, the control regions become less populated as the number of bins increases. Therefore, this binning treatment is only used for $p_{\mathrm{T}}^{\mathrm{miss}}$.

Another approach to binning is to treat the bins as additional search regions in the ABCD plane. Shown in Figure 8.5, R2 and R4 can be separated into two M_J bins—with A denoting the low-M_J region and B denoting the high-M_J region. When performing the background prediction, each of these bins will use the same transfer factor, calculated with R3 and R1. This binning scheme allows us to target non-compressed signal models with R4B, while maintaining sensitivity to compressed models with R4A. Furthermore,

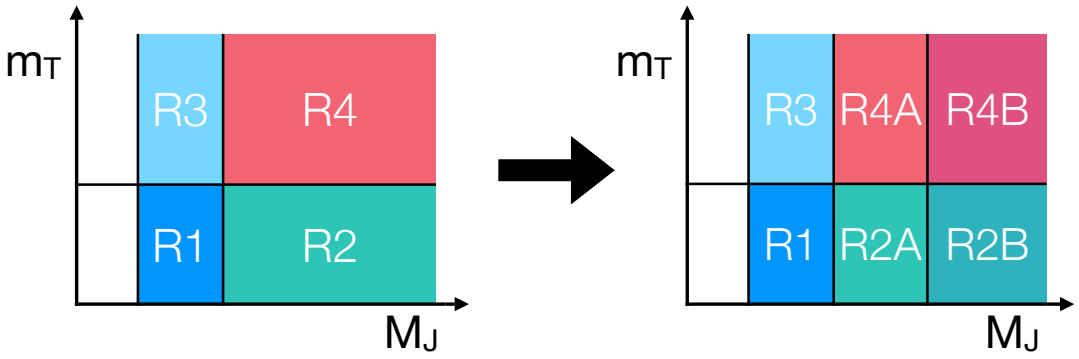
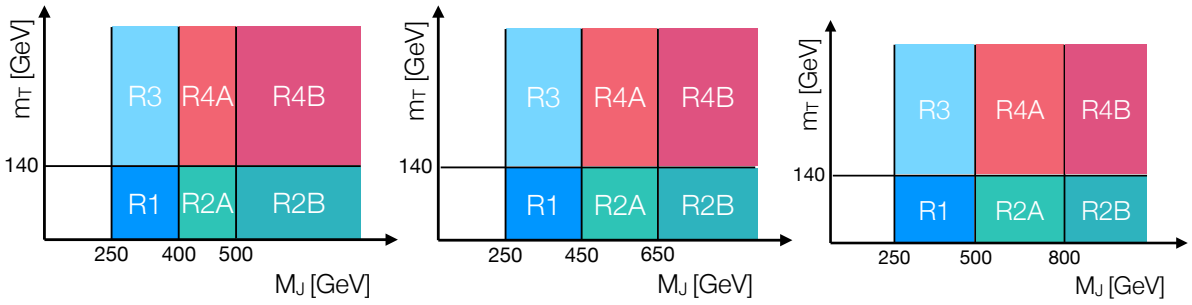


Figure 8.5: The ABCD plane was updated by further binning of the high M_J region. This update gives increased sensitivity to non-compressed signal models, while also maintaining sensitivity to compressed models.



(a) $200 < p_{\mathrm{T}}^{\mathrm{miss}} \leq 350\,\mathrm{GeV}$ (b) $350 < p_{\mathrm{T}}^{\mathrm{miss}} \leq 500\,\mathrm{GeV}$ (c) $p_{\mathrm{T}}^{\mathrm{miss}} > 500\,\mathrm{GeV}$

Figure 8.6: ABCD plane for each $p_{\mathrm{T}}^{\mathrm{miss}}$ bin.

as each $p_{\mathrm{T}}^{\mathrm{miss}}$ bin is treated as a separate ABCD plane, the M_J binning can be done differently in each. The M_J binning for each $p_{\mathrm{T}}^{\mathrm{miss}}$ bin is shown in Figure 8.6.

The ABCD method can be generalized further to allow for binning in more than two variables. Section 8.1 argued that the lack of correlation between M_J and m_{T} allowed us to extrapolate from low m_{T} to high m_{T} across bins of M_J. This argument can be applied to any variable—or combination of variables—that is uncorrelated with m_{T}. Figure 8.2 showed that the shapes of the N_{jets} distribution are almost identical for the dominant backgrounds—1ℓ $\mathrm{t\bar{t}}$ and 2ℓ $\mathrm{t\bar{t}}$—due to the large amount of ISR present. As a result, the m_{T} distribution is approximately independent of N_{jets} and N_{b}. Therefore, R2 and R4 can be binned in these variables. This is shown in Figure 8.7. As before, all bins use the

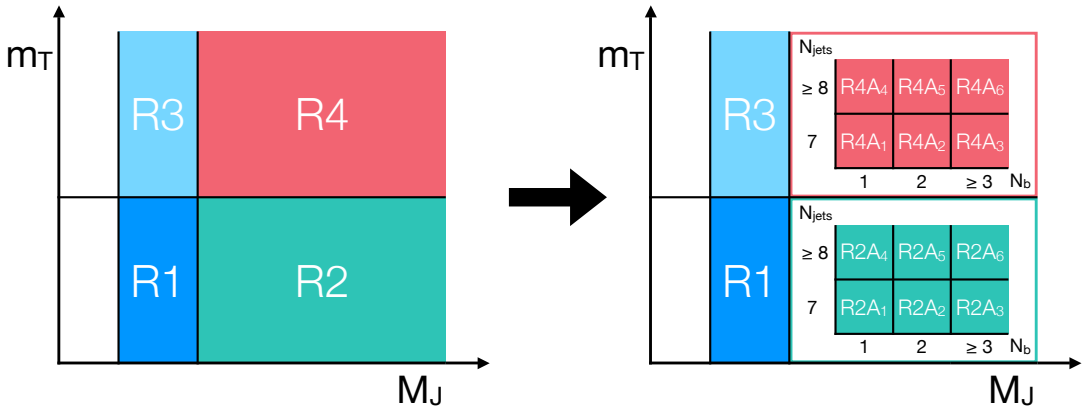


Figure 8.7: ABCD plane updated by introducing bins of N_{jets} and N_{b} into R2 and R4.

transfer factor calculated in the inclusive R3 and R1 regions to predict the background.

Combining all these binning strategies, we end up with the bin configuration shown in Figure 8.8. The search region is separated into three $p_{\text{T}}^{\text{miss}}$ categories (represented as three ABCD planes), two bins of M_J, two bins of N_{jets}, and three bins of N_{b}. This gives a total of $3 \times 2 \times 2 \times 3 = 36$ search region bins.

To incorporate this binning into the expected yields of Eq. (8.3), indices are added to represent the bins. This gives

$$\mu_{\text{R1},i} = r \cdot \mu^{\text{sig}}_{\text{R1},i} + \mu^{\text{bkg}}_{\text{R1},i} \qquad = r \cdot \mu^{\text{sig}}_{\text{R1},i} + \theta_{1,i}, \tag{8.5a}$$

$$\mu_{\text{R2A},i,j} = r \cdot \mu^{\text{sig}}_{\text{R2A},i,j} + \mu^{\text{bkg}}_{\text{R2A},i,j} = r \cdot \mu^{\text{sig}}_{\text{R2A},i,j} + \theta_{2A,i,j}, \tag{8.5b}$$

$$\mu_{\text{R2B},i,j} = r \cdot \mu^{\text{sig}}_{\text{R2B},i,j} + \mu^{\text{bkg}}_{\text{R2B},i,j} = r \cdot \mu^{\text{sig}}_{\text{R2B},i,j} + \theta_{2B,i,j}, \tag{8.5c}$$

$$\mu_{\text{R3},i} = r \cdot \mu^{\text{sig}}_{\text{R3},i} + \mu^{\text{bkg}}_{\text{R3},i} \qquad = r \cdot \mu^{\text{sig}}_{\text{R3},i} + \theta_{3,i}, \tag{8.5d}$$

$$\mu_{\text{R4A},i,j} = r \cdot \mu^{\text{sig}}_{\text{R4A},i,j} + \mu^{\text{bkg}}_{\text{R4A},i,j} = r \cdot \mu^{\text{sig}}_{\text{R4A},i,j} + (\theta_{3,i}/\theta_{1,i}) \cdot \theta_{2A,i,j} \cdot \kappa_{A,i,j}, \tag{8.5e}$$

$$\mu_{\text{R4B},i,j} = r \cdot \mu^{\text{sig}}_{\text{R4B},i,j} + \mu^{\text{bkg}}_{\text{R4B},i,j} = r \cdot \mu^{\text{sig}}_{\text{R4B},i,j} + (\theta_{3,i}/\theta_{1,i}) \cdot \theta_{2B,i,j} \cdot \kappa_{B,i,j}, \tag{8.5f}$$

where the i index denotes the $p_{\text{T}}^{\text{miss}}$ bin and the j index denotes the bin in the N_{jets}-N_{b} plane. For notational clarity, M_J bins are treated separately, though they could be represented using an additional index.

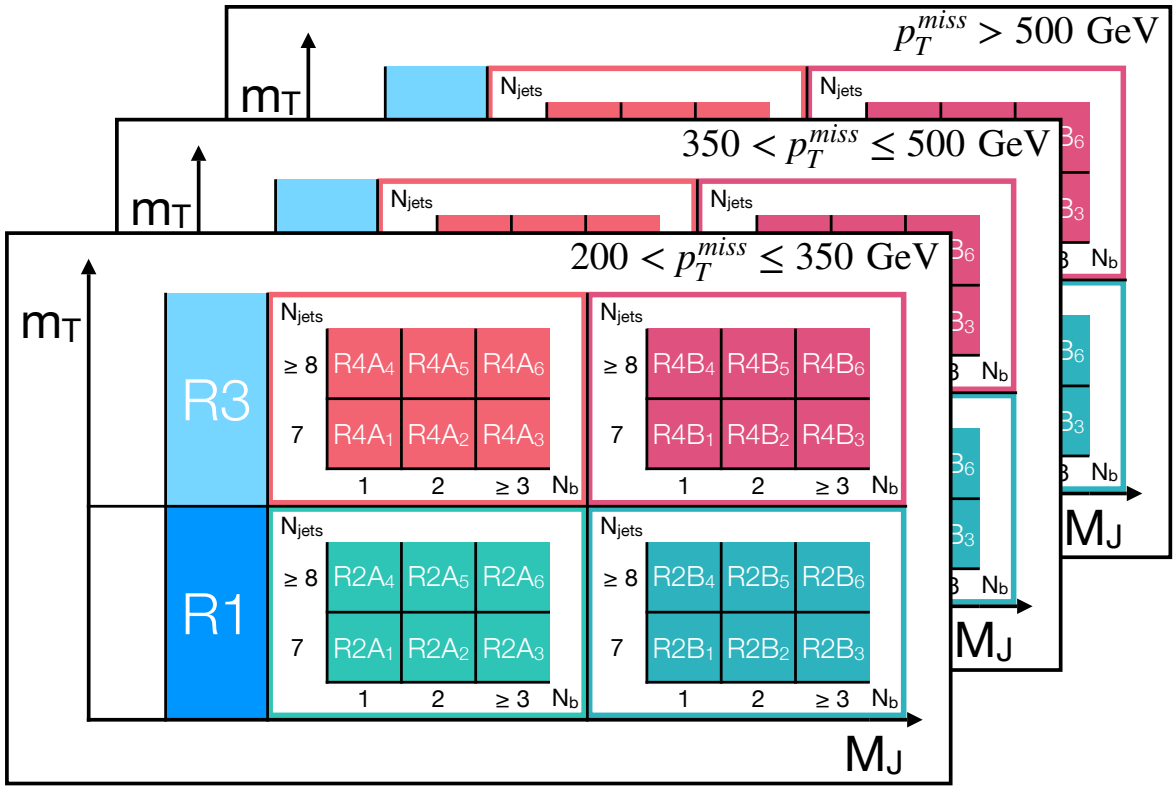


Figure 8.8: The complete ABCD binning used in this analysis. The three $p_{\mathrm{T}}^{\mathrm{miss}}$ bins are represented as separate planes. The R1 and R3 regions are inclusive in each $p_{\mathrm{T}}^{\mathrm{miss}}$ bin, while the R2 and R4 regions are binned in M_J, N_{jets}, and N_{b}. The bounds for the M_J bins are different in each $p_{\mathrm{T}}^{\mathrm{miss}}$ bin. In total, there are 36 search regions.

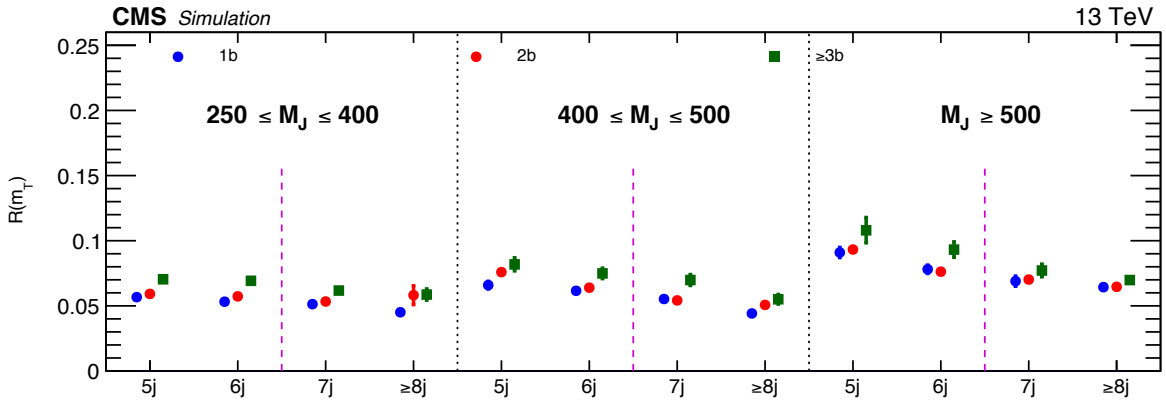


Figure 8.9: The ratio $R(m_{\mathrm{T}})$ of high-m_{T} to low-m_{T} event yields based on $\mathrm{t\bar{t}}$ simulated events in the lowest $p_{\mathrm{T}}^{\mathrm{miss}}$ bin, as a function of N_{jets} and N_{b} for the three M_J regions. The baseline selection requires $N_{\mathrm{jets}} \geq 7$ indicated by the magenta lines. The uncertainties shown are statistical only. Higher $p_{\mathrm{T}}^{\mathrm{miss}}$ bins do not show significant changes in behavior.

As no additional m_{T} bins are added, θ_1 and θ_3 only have one index, which specifies the $p_{\mathrm{T}}^{\mathrm{miss}}$ bin. As discussed earlier, this binning method requires N_{jets} and N_{b} to be approximately independent with m_{T} in each $p_{\mathrm{T}}^{\mathrm{miss}}$ bin. This dependence is studied with $R(m_{\mathrm{T}})$, defined as

$$R(m_{\mathrm{T}}) = \frac{\mu^{\mathrm{bkg}}(m_{\mathrm{T}} > 140\,\mathrm{GeV})}{\mu^{\mathrm{bkg}}(m_{\mathrm{T}} \leq 140\,\mathrm{GeV})}. \tag{8.6}$$

In Figure 8.9, Figure 8.9 shows $R(m_{\mathrm{T}})$ in simulation plotted as a function of N_{jets} and N_{b} across the three M_J regions in the lowest $p_{\mathrm{T}}^{\mathrm{miss}}$ bin. Above the $N_{\mathrm{jets}} \geq 7$ requirement of the baseline selection, $R(m_{\mathrm{T}})$ has no significant dependence on N_{jets}, and a slight dependence on N_{b}. Below this requirement, however, $R(m_{\mathrm{T}})$ increases for lower values of N_{jets}. In low-N_{jets}, high-M_J bins, M_J is less dominated by ISR jets, thus the M_J agreement between high-m_{T} 2ℓ $\mathrm{t\bar{t}}$ and low-m_{T} 1ℓ $\mathrm{t\bar{t}}$ begins to break down.

To account for the small dependence of $R(m_{\mathrm{T}})$ on binning variables, κ factors are included in the background estimation. Calculated in simulation for each search region

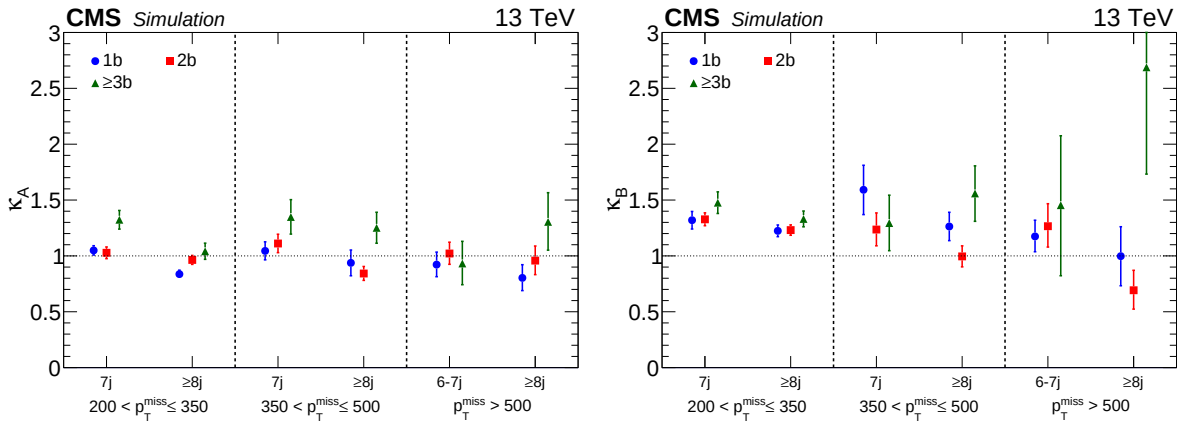


Figure 8.10: Values of κ in each of the 18 signal bins of the low-M_J ABCDs (left), and the 18 signal bins of the high-M_J ABCDs (right), calculated using the simulated SM background. The κ factors are close to unity, indicating the small correlation between M_J and m_{T}. The uncertainties shown are statistical only.

bin, these factors are defined as

$$\kappa^{\mathrm{MC}}_{A,i,j} = \frac{N^{\mathrm{bkg}}_{\mathrm{R4A},i,j}/N^{\mathrm{bkg}}_{\mathrm{R2A},i,j}}{N^{\mathrm{bkg}}_{\mathrm{R3},i}/N^{\mathrm{bkg}}_{\mathrm{R1},i}}, \tag{8.7}$$

$$\kappa^{\mathrm{MC}}_{B,i,j} = \frac{N^{\mathrm{bkg}}_{\mathrm{R4B},i,j}/N^{\mathrm{bkg}}_{\mathrm{R2B},i,j}}{N^{\mathrm{bkg}}_{\mathrm{R3},i}/N^{\mathrm{bkg}}_{\mathrm{R1},i}}. \tag{8.8}$$

Figure 8.10 shows the κ factors calculated in each search region bin. Most κ factors are consistent with unity, indicating the small correlation between M_J and m_{T}. κ factors are the only input from simulation in the background estimation method.

8.4 Maximum Likelihood Parameter Estimation

To determine the parameter values that best describe the observed data, we maximize a function describing the probability of the measured data as a function of these parameters. This function, known as the *likelihood*, is the product of the probability density function (PDF) for each observable evaluated at the measured value. Explicitly, for a given set of independent observations $\mathbf{x} = \{x_1, x_2, ...x_N\}$, the likelihood of a model

Table 8.1: Observables and their corresponding fit value, with the contribution to the likelihood.

Observable	Fit value	Likelihood term
$N^{\text{data}}_{R1,i}$	$r \cdot \mu^{\text{sig}}_{R1,i} + \theta_{1,i}$	$\text{Pois}(N_{\text{Obs}}\|\mu_{\text{Fit}})$
$N^{\text{data}}_{R2A,i,j}$	$r \cdot \mu^{\text{sig}}_{R2A,i,j} + \theta_{2A,i,j}$	
$N^{\text{data}}_{R2B,i,j}$	$r \cdot \mu^{\text{sig}}_{R2B,i,j} + \theta_{2B,i,j}$	
$N^{\text{data}}_{R3,i}$	$r \cdot \mu^{\text{sig}}_{R3,i} + \theta_{3,i}$	
$N^{\text{data}}_{R4A,i,j}$	$r \cdot \mu^{\text{sig}}_{\text{R4A},i,j} + (\theta_{3,i}/\theta_{1,i}) \cdot \theta_{2A,i,j} \cdot \kappa_{A,i,j}$	
$N^{\text{data}}_{R4B,i,j}$	$r \cdot \mu^{\text{sig}}_{\text{R4B},i,j} + (\theta_{3,i}/\theta_{1,i}) \cdot \theta_{2B,i,j} \cdot \kappa_{B,i,j}$	
$N^{\text{sig}}_{R1,i}$	$\mu^{\text{sig}}_{R1,i}$	$\text{Log-Norm}(N_{\text{MC}}\|\mu_{\text{Fit}},\sigma_{\text{MC}})$
$N^{\text{sig}}_{R2A,i,j}$	$\mu^{\text{sig}}_{R2A,i,j}$	
$N^{\text{sig}}_{R2B,i,j}$	$\mu^{\text{sig}}_{R2B,i,j}$	
$N^{\text{sig}}_{R3,i}$	$\mu^{\text{sig}}_{R3,i}$	
$N^{\text{sig}}_{R4A,i,j}$	$\mu^{\text{sig}}_{R4A,i,j}$	
$N^{\text{sig}}_{R4B,i,j}$	$\mu^{\text{sig}}_{R4B,i,j}$	
$\kappa^{\text{MC}}_A = \frac{N^{\text{bkg}}_{\text{R4A},i,j}/N^{\text{bkg}}_{\text{R2A},i,j}}{N^{\text{bkg}}_{\text{R3},i}/N^{\text{bkg}}_{\text{R1},i}}$	κ^{Fit}_A	$\text{Gaus}(\kappa^{\text{MC}}\|\kappa^{\text{Fit}},\sigma^{\text{MC}})$
$\kappa^{\text{MC}}_B = \frac{N^{\text{bkg}}_{\text{R4B},i,j}/N^{\text{bkg}}_{\text{R2B},i,j}}{N^{\text{bkg}}_{\text{R3},i}/N^{\text{bkg}}_{\text{R1},i}}$	κ^{Fit}_B	

with parameters $\boldsymbol{\theta}$ is given by

$$L(\boldsymbol{\theta}|\mathbf{x}) = \prod_i^N f_i(x_i|\boldsymbol{\theta}), \tag{8.9}$$

where $f_i(x|\boldsymbol{\theta})$ is the PDF for the ith observable. Depending on the observable, a different PDF may be used in the likelihood. Given this likelihood function, the model parameters are estimated by finding the values which maximize the likelihood.

Ignoring systematics for now, the likelihood function we maximize can be factored into three terms,

$$L = L^{\text{data}} \cdot L^{\text{MC}}_{\text{sig}} \cdot L^{\text{MC}}_{\kappa}, \tag{8.10}$$

with each term accounting for a different set of observables. Each of these terms will use a different PDF to compare the observed and fit values. We will discuss each of these distributions, and why they are used for the given set of observables.

The probability of observing N events (where N is an integer) in a bin where μ events

are expected is given by the Poisson distribution:

$$\text{Pois}(N|\mu) = \frac{1}{N!}\mu^N e^{-\mu}. \tag{8.11}$$

The shape of the Poisson distribution is completely determined by the expected value μ, with the varian given by μ. Observables corresponding to the number of events in a bin—such as N^{data}—will contribute a Poisson term to the likelihood. Thus, the data term in the likelihood has the form

$$L^{\text{data}} = \prod_{R\in S} \prod_{i=1}^{N_{\text{bins}}(R)} \text{Pois}(N_{R,i}|\mu_{R,i}). \tag{8.12}$$

As we saw in Chapter 6, simulated events are weighted by cross section so that the relative amount of each sample agrees with data for the integrated luminosity of the data set. Thus, a simulated sample with N_{gen} generated events, each having a weight of w, corresponds to $N_{\text{data}} = wN_{\text{gen}}$ data events. These weighted events require a slightly different treatment in the likelihood than the unweighted data yields. Treating N_{gen} as the observed number of events and μ/w as the expectation, the Poisson expression for this would be:

$$\text{Pois}\left(N_{\text{gen}}\middle|\frac{\mu}{w}\right) = \frac{1}{N_{\text{gen}}!}\left(\frac{\mu}{w}\right)^{N_{\text{gen}}} e^{-\mu/w} = \text{Gamma}\left(N_{\text{gen}}|\mu, w\right). \tag{8.13}$$

With the inclusion of the weight parameter, this expression takes the form of another type of PDF, the gamma distribution.

The previous treatment of weighted events assumes that every generated event has the same weight w. This assumption does not hold when considering contributions from multiple processes, as each will have a different cross section. Furthermore, the application of scale factors—which depend on the kinematic quantities of an event—results in a distribution of event weights, even within a single sample. To account for

this, N_{gen} and w are replaced with effective values, defined as

$$N_{\text{eff}} \equiv \frac{\left(\sum_{i=1}^{N_{\text{gen}}} w_i\right)^2}{\sum_{i=1}^{N_{\text{gen}}} w_i^2}, \tag{8.14}$$

$$w_{\text{eff}} \equiv \frac{\sum_{i=1}^{N_{\text{gen}}} w_i^2}{\sum_{i=1}^{N_{\text{gen}}} w_i}. \tag{8.15}$$

These definitions maintain the property $\hat{\mu} = wN_{\text{gen}} = w_{\text{eff}}N_{\text{eff}}$. In other words, these definitions do not bias the mean parameter $\hat{\mu}$.

With this treatment of N and w, the signal yield term in the likelihood has the form:

$$L_{\text{sig}}^{\text{MC}} = \prod_{R\in S} \prod_{i=1}^{N_{\text{bins}}(R)} \text{Gamma}(N_{R,i}^{\text{MC,sig}} | \mu_{R,i}^{\text{sig}}, w_{R,i}^{MC,sig}). \tag{8.16}$$

Though this method gives a close approximation to the Poisson errors in the data, calculating effective yields and weights can be a time-intensive endeavor. Therefore, for the full signal scan—in which the limit is calculating for each of the ∼1000 mass points—the Log-Normal distribution is used for the signal terms in the likelihood. This distribution is given by

$$\text{Log-Norm}(x|\mu,\sigma) = \frac{1}{x\sqrt{2\pi}\ln\sigma} e^{-\frac{1}{2}\left[\frac{\ln(x/\mu)}{\ln\sigma}\right]^2}, \tag{8.17}$$

with μ denoting the median and σ denoting the geometric standard deviation $(1+\Delta x/x)$. Using this method, the signal terms in the likelihood become

$$L_{\text{sig}}^{\text{MC}} = \prod_{R\in S} \prod_{i=1}^{N_{\text{bins}}(R)} \text{Log-Norm}(N_{R,i}^{\text{MC,sig}} | \mu_{R,i}^{\text{sig}}, \sigma_{R,i}^{MC,sig}), \tag{8.18}$$

where the σ value is taken from simulation. This distribution is also used for systematic uncertainties, which we will discuss in the next chapter.

For observables that don't correspond to the number of events in a bin, such as κ, a Gaussian (or normal) distribution is used. This is given by

$$\text{Gaus}(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2}, \tag{8.19}$$

where μ is the mean of the distribution and σ is the standard deviation. Considering

both the A and B region κ values, the term in the likelihood is given by

$$L_{\kappa}^{\mathrm{MC}} = \prod_{i=1}^{3}\prod_{j=1}^{6}\left[\mathrm{Gaus}(\kappa_{A,i,j}^{\mathrm{MC}}|\kappa_{A,i,j}^{\mathrm{Fit}},\sigma_{A,i,j}^{\mathrm{MC}})\cdot\mathrm{Gaus}(\kappa_{B,i,j}^{\mathrm{MC}}|\kappa_{B,i,j}^{\mathrm{Fit}},\sigma_{B,i,j}^{\mathrm{MC}})\right]. \tag{8.20}$$

8.4.1 Systematic Uncertainties in the Likelihood

Before discussing sources of systematic uncertainties, we will cover how they are implemented in the likelihood. Because systematic errors occur alongside statistical error, they will add additional variability to the yields predicted in the ABCD method. This is realized mathematically by multiplying terms in Eq. (8.5) by an exponential factor of the form

$$\exp\left(\sum_{k}\alpha_{i,j,k}Z_k\right), \tag{8.21}$$

where the sum is taken over all sources of systematic uncertainty; $\alpha_{i,j,k}$ are constants specifying the size of the effect of the kth uncertainty on $\kappa_{i,j}$, being related to the relative uncertainty by $\alpha = \ln(1+\sigma)$; and Z_k are fit parameters with standard normal constraints.

The exponential factors include the combined effect of all systematic variations. Generally, there will be an exponential factor for each $\mu^{\mathrm{sig/bkg}}$ term in Eq. (8.5). For the background systematics, however, this factor only needs to be included in either the μ_{R2} or μ_{R4} terms[1]. This is because the background estimation method can only be biased by the mismodeling of κ in simulation, which will be explored further in Section **??**. As we will need to omit the R4 terms for one of our fits in Chapter 10, we will include the background systematics factor in the R2 expression. Including these terms, the expected

[1] As the nuisance parameters Z_k are Gaussian constrained, these two methods are mathematically equivalent.

yields are parameterized as

$$\mu_{\mathrm{R1},i} = r \cdot \mu^{\mathrm{sig}}_{\mathrm{R1},i} \cdot \exp\left(\sum_{k=1}^{S_{\mathrm{syst}}} \alpha^{\mathrm{sig}}_{\mathrm{R1},i,j,k} Z_k\right) + \theta_{1,i}, \tag{8.22a}$$

$$\mu_{\mathrm{R2A},i,j} = r \cdot \mu^{\mathrm{sig}}_{\mathrm{R2A},i,j} \cdot \exp\left(\sum_{k=1}^{S_{\mathrm{syst}}} \alpha^{\mathrm{sig}}_{\mathrm{R2A},i,j,k} Z_k\right) + \theta_{2A,i,j} \cdot \exp\left(\sum_{m=1}^{B_{\mathrm{syst}}} \alpha^{\mathrm{bkg}}_{i,j,m} Z_m\right), \tag{8.22b}$$

$$\mu_{\mathrm{R2B},i,j} = r \cdot \mu^{\mathrm{sig}}_{\mathrm{R2B},i,j} \cdot \exp\left(\sum_{k=1}^{S_{\mathrm{syst}}} \alpha^{\mathrm{sig}}_{\mathrm{R2B},i,j,k} Z_k\right) + \theta_{2B,i,j} \cdot \exp\left(\sum_{m=1}^{B_{\mathrm{syst}}} \alpha^{\mathrm{bkg}}_{i,j,m} Z_m\right), \tag{8.22c}$$

$$\mu_{\mathrm{R3},i} = r \cdot \mu^{\mathrm{sig}}_{\mathrm{R3},i} \cdot \exp\left(\sum_{k=1}^{S_{\mathrm{syst}}} \alpha^{\mathrm{sig}}_{\mathrm{R3},i,j,k} Z_k\right) + \theta_{3,i}, \tag{8.22d}$$

$$\mu_{\mathrm{R4A},i,j} = r \cdot \mu^{\mathrm{sig}}_{\mathrm{R4A},i,j} \cdot \exp\left(\sum_{k=1}^{S_{\mathrm{syst}}} \alpha^{\mathrm{sig}}_{\mathrm{R4A},i,j,k} Z_k\right) + (\theta_{3,i}/\theta_{1,i}) \cdot \theta_{2A,i,j} \cdot \kappa_{A,i,j}, \tag{8.22e}$$

$$\mu_{\mathrm{R4B},i,j} = r \cdot \mu^{\mathrm{sig}}_{\mathrm{R4B},i,j} \cdot \exp\left(\sum_{k=1}^{S_{\mathrm{syst}}} \alpha^{\mathrm{sig}}_{\mathrm{R4B},i,j,k} Z_k\right) + (\theta_{3,i}/\theta_{1,i}) \cdot \theta_{2B,i,j} \cdot \kappa_{B,i,j}, \tag{8.22f}$$

where k is summed over signal systematics and m is summed over background systematics.

With the parameterization of Eq. (8.22) established, we can now introduce the relevant terms into the likelihood. As mentioned previously, the nuisance parameters Z_k are constrained by a standard normal distribution. Therefore, for each of these nuisance parameters, there will be a Gaussian term in the likelihood. Including background and signal systematics, the likelihood term introduced by systematic uncertainties is

$$L^{\mathrm{syst}} = \prod_{m=1}^{B_{\mathrm{syst}}} \mathrm{Gaus}\left(0|Z_m, 1\right) \cdot \prod_{k=1}^{S_{\mathrm{syst}}} \mathrm{Gaus}\left(0|Z_k, 1\right). \tag{8.23}$$

This implementation is equivalent to treating the exponential factors of Eq. (8.21) as fit parameters and adding log-normal terms to the likelihood. In other words, if Z_k is normally distributed, then $e^{\alpha_{i,j,k} Z_k}$ is log-normally distributed. Thus, systematic uncertainties are often said to be treated with symmetric log-normal constraints.

Chapter 9

Systematic Uncertainties for Background Estimation

In Chapter 8, we constructed an ABCD method that predicts the standard model background in our search regions. This method was formulated mathematically using the likelihood function, which accounts for the difference between the value of each fit parameter and the observed value it corresponds to. So far, these differences have only been treated as if they are caused purely by random error. For a complete assessment of the uncertainties, however, we must also consider the effect of *systematic uncertainties* on our observations. Systematic uncertainties are not introduced by chance, but by inaccuracies inherent in the analysis procedure. Typical sources of systematic uncertainties include detector resolution, cross section uncertainty, and MC mismodeling.

The two types of systematic uncertainties considered in this analysis are those arising from the background estimation and the those arising from signal models. These types of systematic uncertainties affect the analysis in different ways, so they are evaluated separately. The background systematics are evaluated using measurements in control samples, while signal systematics are evaluated using dedicated studies.

When we constructed the ABCD method in Section 8.2, we assumed that m_{T} and M_J were approximately independent, accounting for any residual correlation by including MC-derived κ factors in our background estimation. As κ is a double ratio, it is robust against a wide range of mismodeling scenarios. However, there are scenarios when the MC shows significant trends in κ. One example of this can be seen in lower $p_{\mathrm{T}}^{\mathrm{miss}}$ regions, as shown in Figure 9.1. With this extended $p_{\mathrm{T}}^{\mathrm{miss}}$ range, κ increases as we go lower in $p_{\mathrm{T}}^{\mathrm{miss}}$, with the effect most pronounced in the high-M_J ABCDs. In scenarios like this, we need to verify that the MC is accurately modeling κ dependencies. This verification is performed by studying control samples that allow us to compare κ values between data and simulation. Before doing so, however, we introduce a new categorization of events based on the mechanism through which they reach the high-m_{T} region.

9.1 Background Event Categories

Figure 9.1 showed that κ deviates significantly from one in the low $p_{\mathrm{T}}^{\mathrm{miss}}$ bins. We find that this dependence is mainly due to the difference in relative contribution of underlying processes between the bins. This can be seen by examining the composition of the high-m_{T} region after the baseline selection in each of the $p_{\mathrm{T}}^{\mathrm{miss}}$ bins, represented with pie charts in Figure 9.2. For these figures, 2ℓ $\mathrm{t\bar{t}}$ events where one of the leptons is a hadronically decaying τ are plotted separately. Though $\mathrm{t\bar{t}}$ events dominate—accounting for $> 75\%$ in every $p_{\mathrm{T}}^{\mathrm{miss}}$ bin—the largest contributing $\mathrm{t\bar{t}}$ final state changes across the $p_{\mathrm{T}}^{\mathrm{miss}}$ range.

The high $p_{\mathrm{T}}^{\mathrm{miss}}$ bins are primarily composed of 2ℓ $\mathrm{t\bar{t}}$ events, with the second lepton either being undetected or a hadronically decaying τ. In both cases, the multiple neutrinos present in the event allow for m_{T} to reach values well above m_{W}. Because the M_J distribution of these events is very similar to that of low-m_{T} 1ℓ $\mathrm{t\bar{t}}$ events, as shown in Figure 8.2, these contributions to the high-m_{T} region are predicted accurately by the

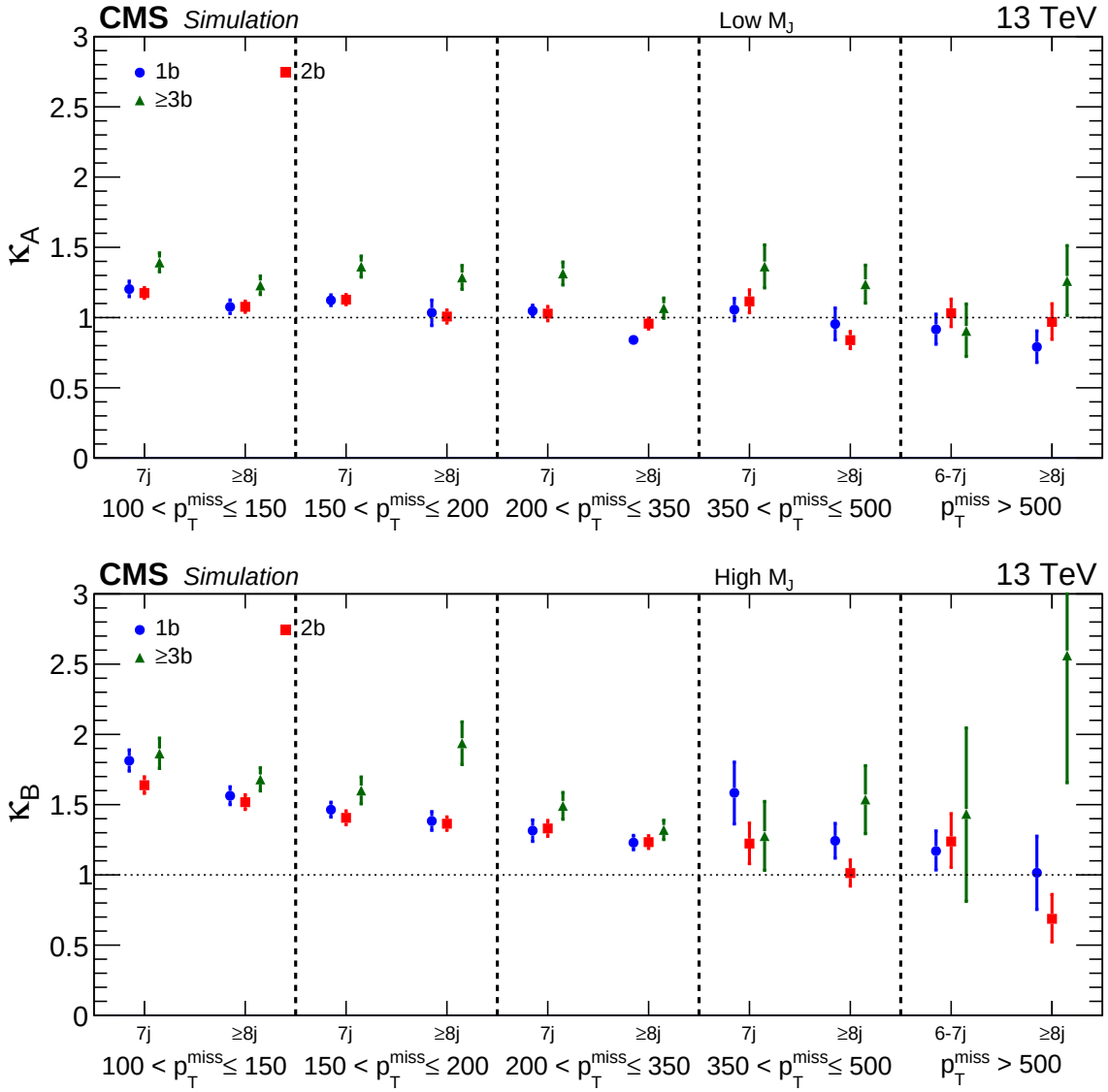


Figure 9.1: Values of κ obtained in simulation as a function of N_{jets} and N_{b} for the full $p_{\text{T}}^{\text{miss}}$ range including the 100-200 GeV validation regions along with the $p_{\text{T}}^{\text{miss}} >$ 200 GeV signal regions for the low-M_J (top) and high-M_J (bottom) ABCD planes.

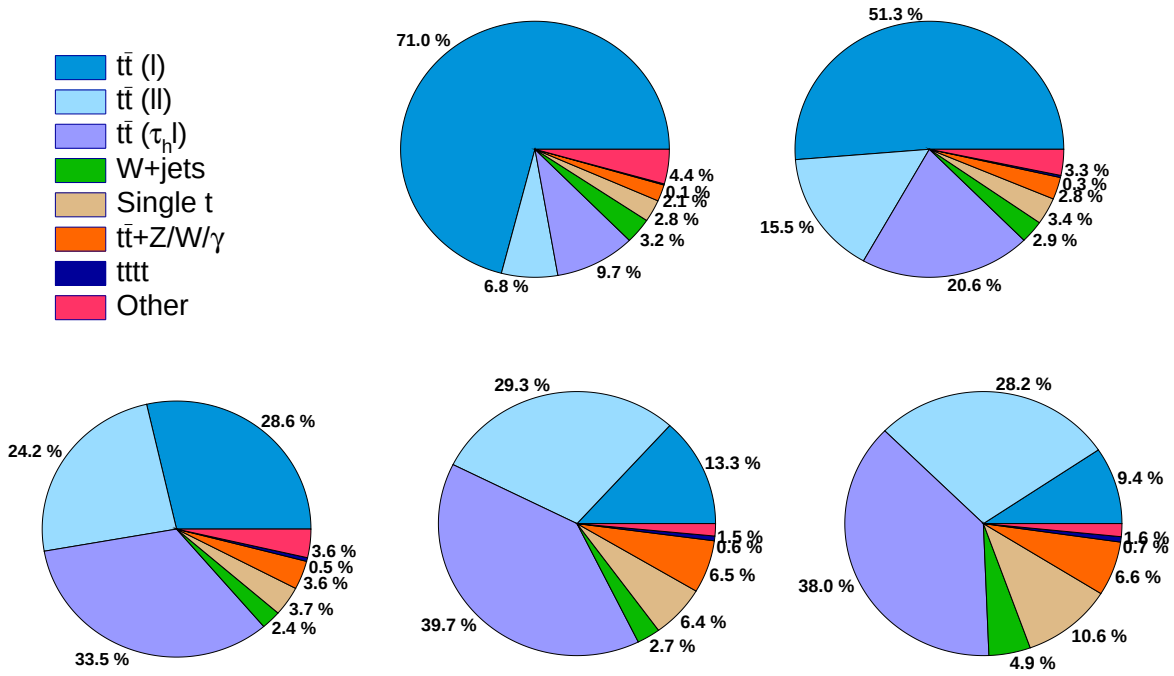


Figure 9.2: Contributions of the different physics processes to the high-m_{T} region in bins of $p_{\mathrm{T}}^{\mathrm{miss}}$ for the baseline selection. Top row (left to right): $100 < p_{\mathrm{T}}^{\mathrm{miss}} < 150\,\mathrm{GeV}$, $150 < p_{\mathrm{T}}^{\mathrm{miss}} < 200\,\mathrm{GeV}$; bottom row: $200 < p_{\mathrm{T}}^{\mathrm{miss}} < 350\,\mathrm{GeV}$, $350 < p_{\mathrm{T}}^{\mathrm{miss}} < 500\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\mathrm{miss}} > 500\,\mathrm{GeV}$.

uncorrected ABCD method, evidenced by κ values near one in these bins.

The low $p_{\mathrm{T}}^{\mathrm{miss}}$ bins, on the other hand, are not as accurately predicted by the uncorrected ABCD method. These bins are dominated by 1ℓ $\mathrm{t\bar{t}}$ events which have a single prompt neutrino produced with the lepton. Though these events typically satisfy $m_{\mathrm{T}} \leq m_{\mathrm{W}}$, there are several mechanisms that allow them to populate the high-m_{T} region. These mechanisms include jet mismeasurement, non-prompt neutrinos, and an off-shell W.

So far we have categorized events based the underlying physical processes in the events. For studying systematic uncertainties, however, events will be categorized by the mechanism through which they populate the high-m_{T} region. This categorization provides a more natural way to study the kinematic dependencies of κ, as it separates events based on kinematic qualities that may introduce correlations between M_J and m_{T}. It also reduces the number of categories that we must consider, simplifying the analysis.

We start by examining the behavior of high-m_{T} $\mathrm{t\bar{t}}$ +jets events. Figure 9.3 compares the M_J distribution of low-m_{T} 1ℓ $\mathrm{t\bar{t}}$ events (black) with four categories of high-m_{T} events. As expected, the categories with two prompt leptons—2ℓ (cyan) and $\ell\tau_{\mathrm{h}}$ (purple)—both have M_J distributions consistent with the low-m_{T} 1ℓ $\mathrm{t\bar{t}}$. The other two categories consist of events with one prompt lepton and a true transverse mass $m_{\mathrm{T}}^{\mathrm{true}}$—calculated using the true lepton momentum and the generated $p_{\mathrm{T}}^{\mathrm{miss}}$—that is either greater than 140 GeV (green) or less than 140 GeV (red). Compared to low-m_{T} 1ℓ $\mathrm{t\bar{t}}$, these events have significantly higher M_J values on average.

The M_J distributions of the high-m_{T} 1ℓ $\mathrm{t\bar{t}}$ categories can be explained by looking at the two distributions shown in Figure 9.4. The plot of $p_{\mathrm{T}}^{\mathrm{miss}}$ resolution (left)—defined as the difference between reconstructed $p_{\mathrm{T}}^{\mathrm{miss}}$ and generated $p_{\mathrm{T}}^{\mathrm{miss}}$—shows events with mismeasured m_{T} (red) typically have a high amount of mismeasured $p_{\mathrm{T}}^{\mathrm{miss}}$. Since $p_{\mathrm{T}}^{\mathrm{miss}}$

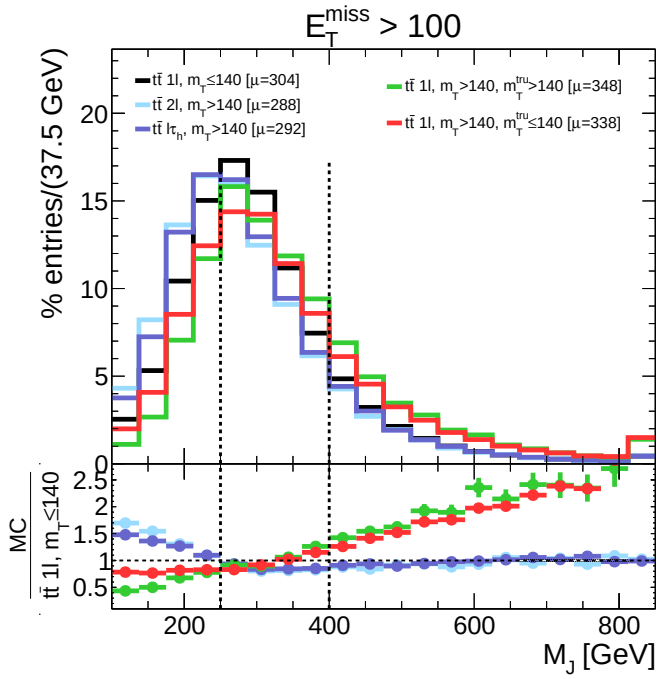


Figure 9.3: Comparison of the M_J distributions for $t\bar{t}$ +jets event categories at low and high m_T after baseline selection with a relaxed p_T^{miss} requirement. The bottom panel shows the ratio between the different high-m_T $t\bar{t}$ categories, and the low-m_T $t\bar{t}$ category plotted in black in the upper panel.

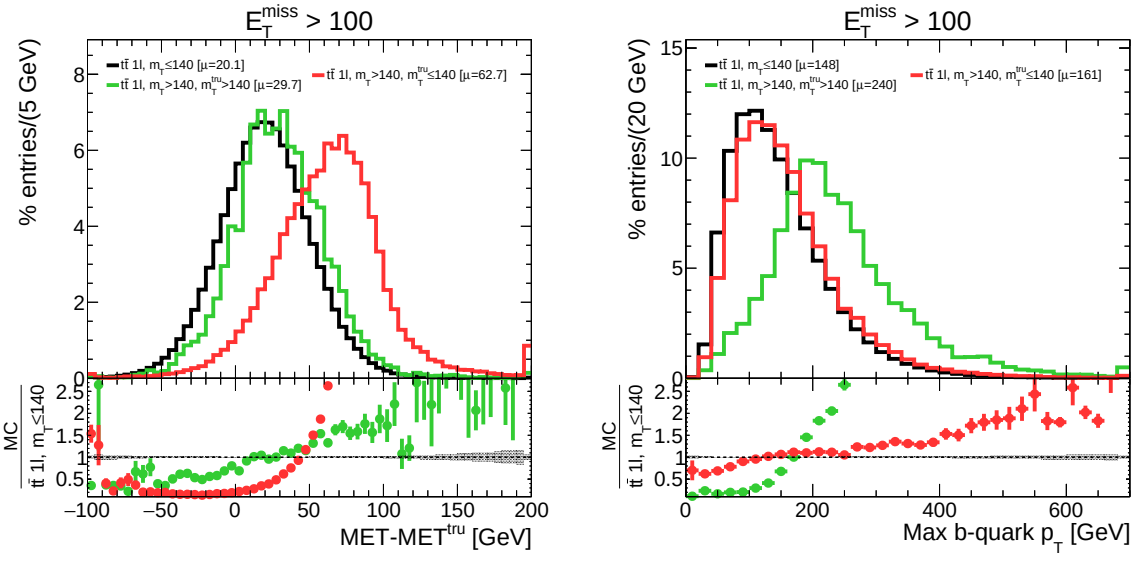


Figure 9.4: Comparison of the p_T^{miss} resolution and maximum b quark p_T distributions for $t\bar{t}$ +jets events at low and high m_T after baseline selection with a relaxed p_T^{miss} requirement.

mismeasurement is positively correlated with jet multiplicity[1] and momentum[2], quantities also correlated with M_J, events with higher average mismeasurement will also have higher average M_J. Thus, M_J and m_{T} become correlated in events with significant mismeasurement.

Well-measured 1ℓ events require a genuine source of additional $p_{\mathrm{T}}^{\mathrm{miss}}$ to reach the high-m_{T} region, which is generated by non-prompt neutrinos, typically originating from b quark decays. In order for these neutrinos to increase m_{T} significantly, however, they need to have momentum comparable to that of the prompt neutrino of the event. As the non-prompt neutrino's momentum is dependent on the parent b quark momentum, events with a high-momentum non-prompt neutrino are also expected to have a high-momentum b quark, which is what we observe in the max b quark p_{T} spectrum (right in Figure 9.4). The higher-p_{T} b jet will give a higher M_J value, and thus the mechanism for reaching the high-m_{T} region also leads to higher average values of M_J.

As m_{T} is intended to reconstruct the mass of the leptonically decaying W, high values of m_{T} can also be achieved if this Wis far off its mass shell. The off-shell Wboson will only significantly effect the kinematics of the lepton and neutrino, however, so the M_J distribution of these events would be consistent with the low-m_{T} region.

So far we have examined the mechanics of high-m_{T} events only in the context of $\mathrm{t\bar{t}}$, but the mechanisms we have explored apply equally well to the subleading backgrounds as well. We therefore define four categories for high-m_{T} events. Shown in Figure 9.5, these categories are:

$\geq$ 2 prompt neutrinos: This category consists primarily of dilepton $\mathrm{t\bar{t}}$ events in which one lepton is not reconstructed or is a hadronically decaying tau. There is a smaller

[1] The more jets there are in an event, the more likely one will be mismeasured.

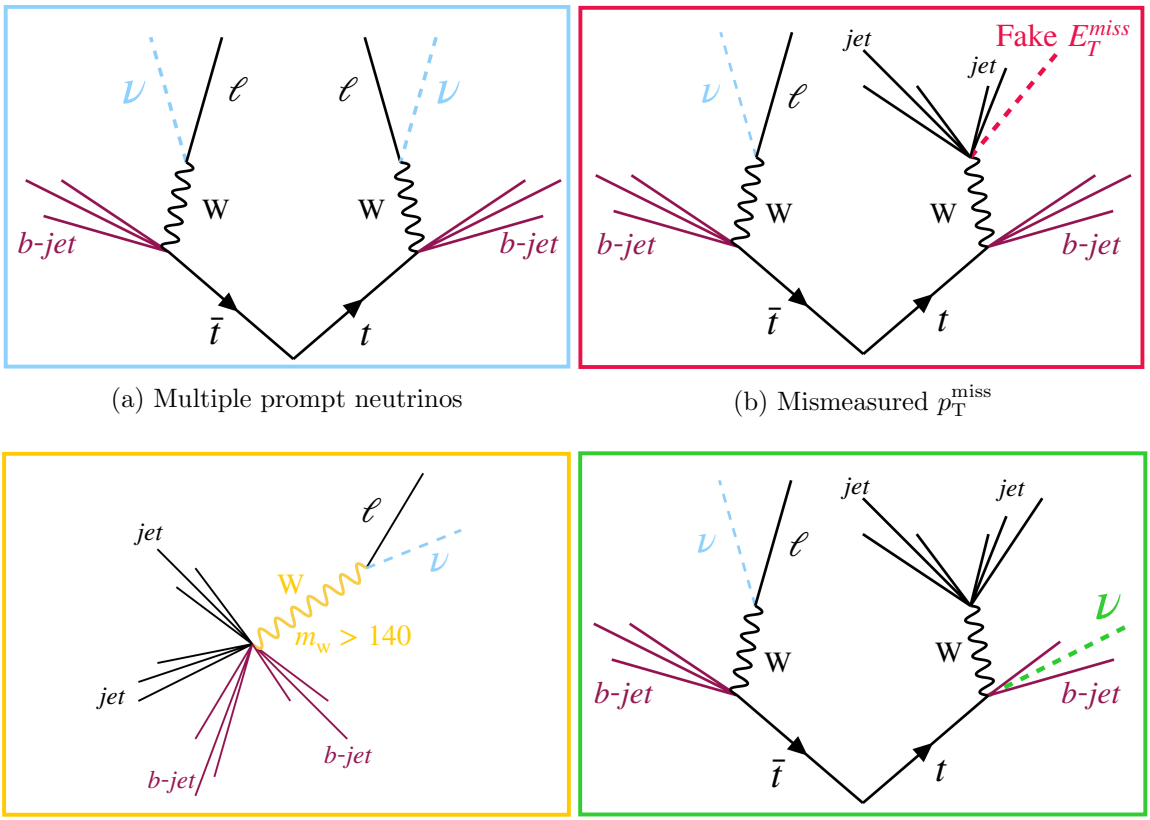
[2] Mismeasuring a high momentum jet will have a larger effect on $p_{\mathrm{T}}^{\mathrm{miss}}$ resolution than mismeasuring a low momentum jet.

contribution from events with 1 charged lepton and extra neutrinos from a Z decay, as in $t\bar{t}Z$. This is the dominant category in the higher-p_T^{miss} search bins.

$\leq 1\ell$, $m_T^{true} \leq 140$ GeV: Events for which the true value of m_T, calculated with the true p_T of the highest p_T prompt lepton and the generated p_T^{miss}, is less than 140 GeV but was mismeasured to be $m_T > 140$ GeV.

$\leq 1\ell$, $m_T^{true} > 140$ GeV, no off-shell $W^{\pm}$: Events with well-measured high m_T originating from non-prompt neutrinos.

$\leq 1\ell$, $m_T^{true} > 140$ GeV, off-shell $W^{\pm}$: Off-shell $W^{\pm}$ allow m_T values greater than the



(a) Multiple prompt neutrinos

(b) Mismeasured p_T^{miss}

(c) Off-shell W (leptonic decay)

(d) Non-prompt neutrinos

Figure 9.5: $t\bar{t}$ background processes categorized by the mechanism through which they reach the high-m_T region.

usual m_{W} cutoff without mismeasurement or correlation with higher boost of the hadronic system. This category contains many W+jets events and can therefore have a different ISR structure than the dominant $\mathrm{t\bar{t}}$ +jets backgrounds.

With these categories defined, we can reexamine the event composition in this new basis. Figure 9.6 shows the relative contribution of each of the four event categories for each of the $p_{\mathrm{T}}^{\mathrm{miss}}$ bins, also separated into four N_{jets} bins. The low $p_{\mathrm{T}}^{\mathrm{miss}}$ bins are dominated by mismeasured events, while the high $p_{\mathrm{T}}^{\mathrm{miss}}$ bins are dominated by events with two prompt neutrinos. The composition across $p_{\mathrm{T}}^{\mathrm{miss}}$ is relatively consistent across the N_{jets} bins.

9.2 Control Samples

In the last section, we separated high-m_{T} background events into four categories, with κ behaving slightly differently in each. In this section, we introduce the two control regions we use to determine the systematic uncertainties associated with the background estimation method. These samples are designed to probe the dependence of κ on key analysis variables, and to determine if the MC is modeling this dependence properly. This is accomplished by comparing κ values found in MC with those measured in data, and checking for any systematic discrepancies. If none are found, the systematic uncertainty associated with the studied variable is taken as the expected uncertainty of the data summed in quadrature with the statistical uncertainty of the MC. Depending on the variable, this may be done on a bin-by-bin basis, or inclusively; the details of the total systematic uncertainty calculation will be given in Section 9.4.

The discussion for each control sample is separated into four sections. We start by motivating the choice of control sample for the behavior we wish to probe. Then, we define the control sample and discuss how it is used. Next, we validate that the control

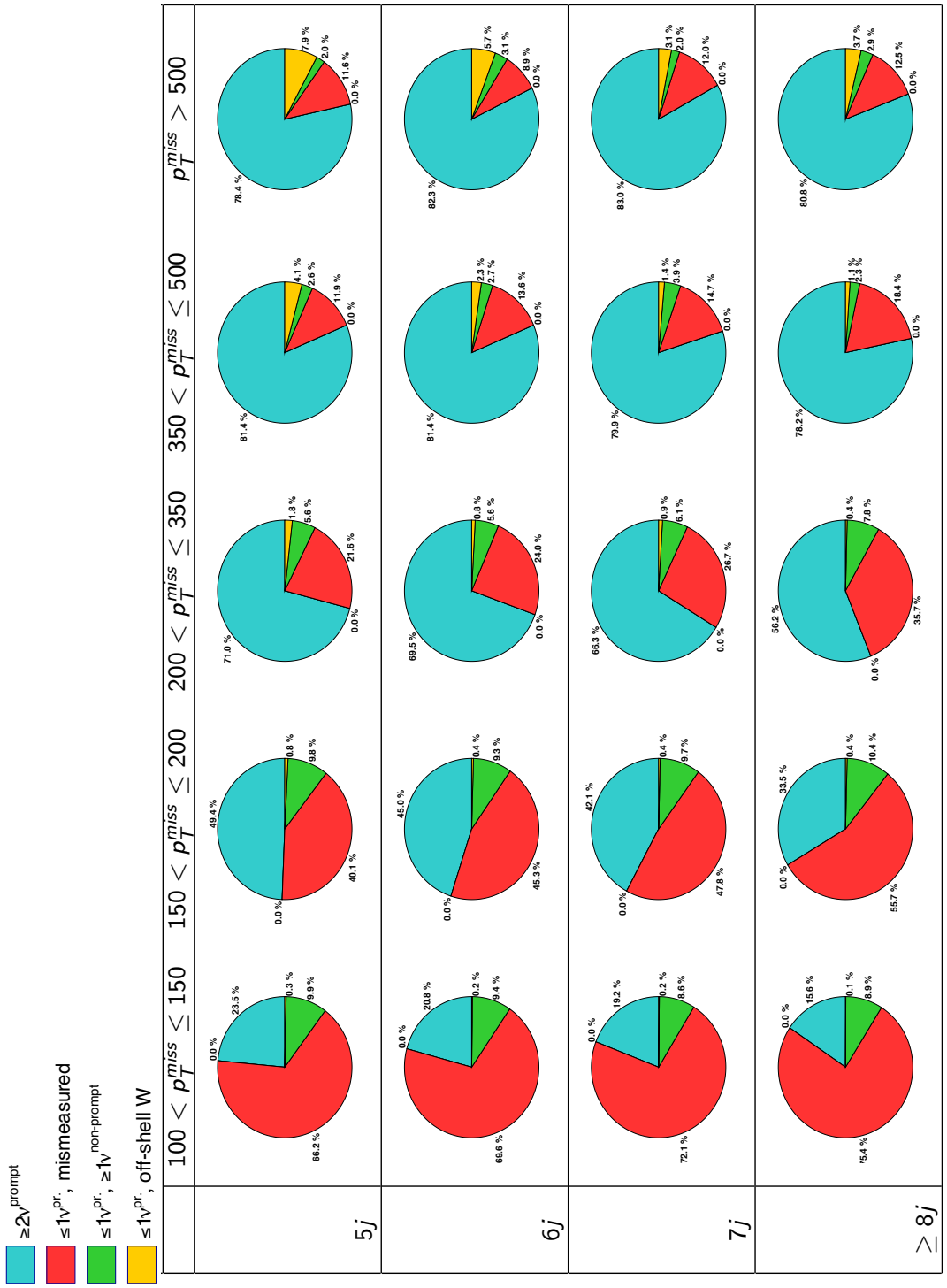


Figure 9.6: Relative background contribution of each of the four event categories in bins of $p_{\mathrm{T}}^{\mathrm{miss}}$ and N_{jets}.

sample is sensitive to the κ dependencies it was designed to account for. Finally, we compare κ values in data and MC to determine the systematic uncertainties that will be assigned for the studied variable.

9.2.1 2ℓ control sample

Motivation

We saw in Figure 8.2 that M_J distributions are similar between high-m_{T} 2ℓ $\mathrm{t\bar{t}}$ and low-m_{T} 1ℓ $\mathrm{t\bar{t}}$. However, since this agreement is fundamental to the background estimation background estimation, it must be verified rigorously. Thus, we define the 2ℓ control sample to confirm this agreement and to quantify the ability of the MC to model any residual shape differences at high M_J. The control sample will be studied in bins of $p_{\mathrm{T}}^{\mathrm{miss}}$, with the lowest $p_{\mathrm{T}}^{\mathrm{miss}}$ bin separated into two bins of N_{jets}. Since the N_{jets} dependence of κ is found to be uncorrelated with $p_{\mathrm{T}}^{\mathrm{miss}}$, the N_{jets} uncertainty found in the lowest $p_{\mathrm{T}}^{\mathrm{miss}}$ bin is used for all bins. The systematic uncertainty associated with $p_{\mathrm{T}}^{\mathrm{miss}}$ is found in the mid- and high-$p_{\mathrm{T}}^{\mathrm{miss}}$ bins for the low- and high-M_J ABCD planes, and applied to the corresponding search region bin.

Definition

A new ABCD plane, shown in Figure 9.7 is constructed that maintains the low-m_{T} regions of the baseline ABCD (R1 and R2A/B) while defining new regions D3 and D4A/B that share the same M_J boundaries, but includes events with a vetoed track (ℓ+v) or an additional lepton (2ℓ). The ℓ+v events are still required to have $m_{\mathrm{T}} > 140\,\mathrm{GeV}$, while the 2ℓ events are not. The baseline selection is applied to ℓ+v events, while for 2ℓ events it is modified in two ways: first, to maintain the number of objects considered for large-R jets, all N_{jets} requirements are reduce by one; second, to increase the sample size, $N_{\mathrm{b}} = 0$

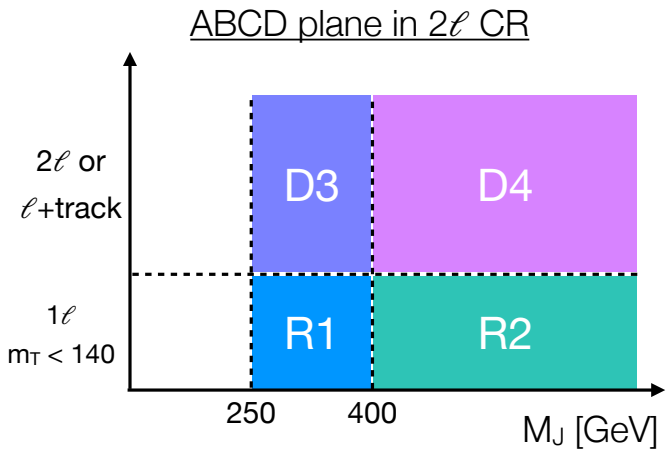


Figure 9.7: Alternate ABCD plane for the 2 lepton control region. The high-m_{T} regions R3/R4 are replaced with D3/D4, which require events having either 2 reconstructed leptons or a reconstructed lepton and an isolated track.

events are included. To avoid any signal contamination, the $N_{\mathrm{b}} \geq 2$ bin is excluded from the control sample. The composition of the 2ℓ and ℓ+v samples are shown in Figure 9.8.

Validation

A study was performed in which the ISR multiplicity corrections applied to MC were lowered by twice their estimated uncertainty for 2ℓ $\mathrm{t\bar{t}}$ events only. The effect this change has on κ values is compared between the 2ℓ control sample and the search region. This was done to demonstrate that mismodeling of jet multiplicity has the same effect on κ values in the 2ℓ control sample as it does for those in the search region, which is shown in Figure 9.9. Specifically, the dependence of κ on jet multiplicity increases. Though mismodeling of ISR multiplicity is unrealistic—ISR corrections are derived and verified in data—this study illustrates how the control sample would be used for other mismodeling scenarios.

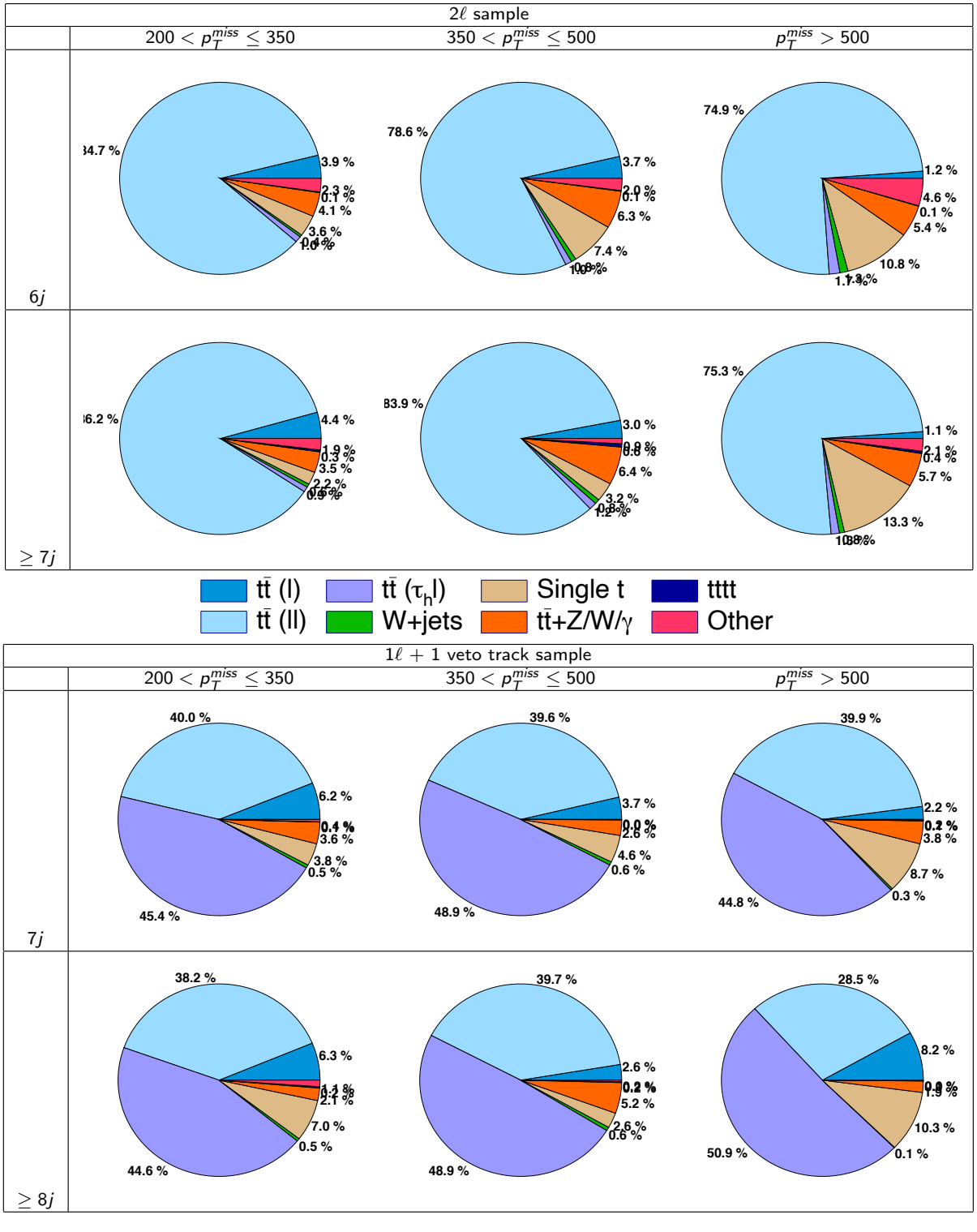


Figure 9.8: Composition of the dilepton (2ℓ) and lepton+track (ℓ+v) control samples.

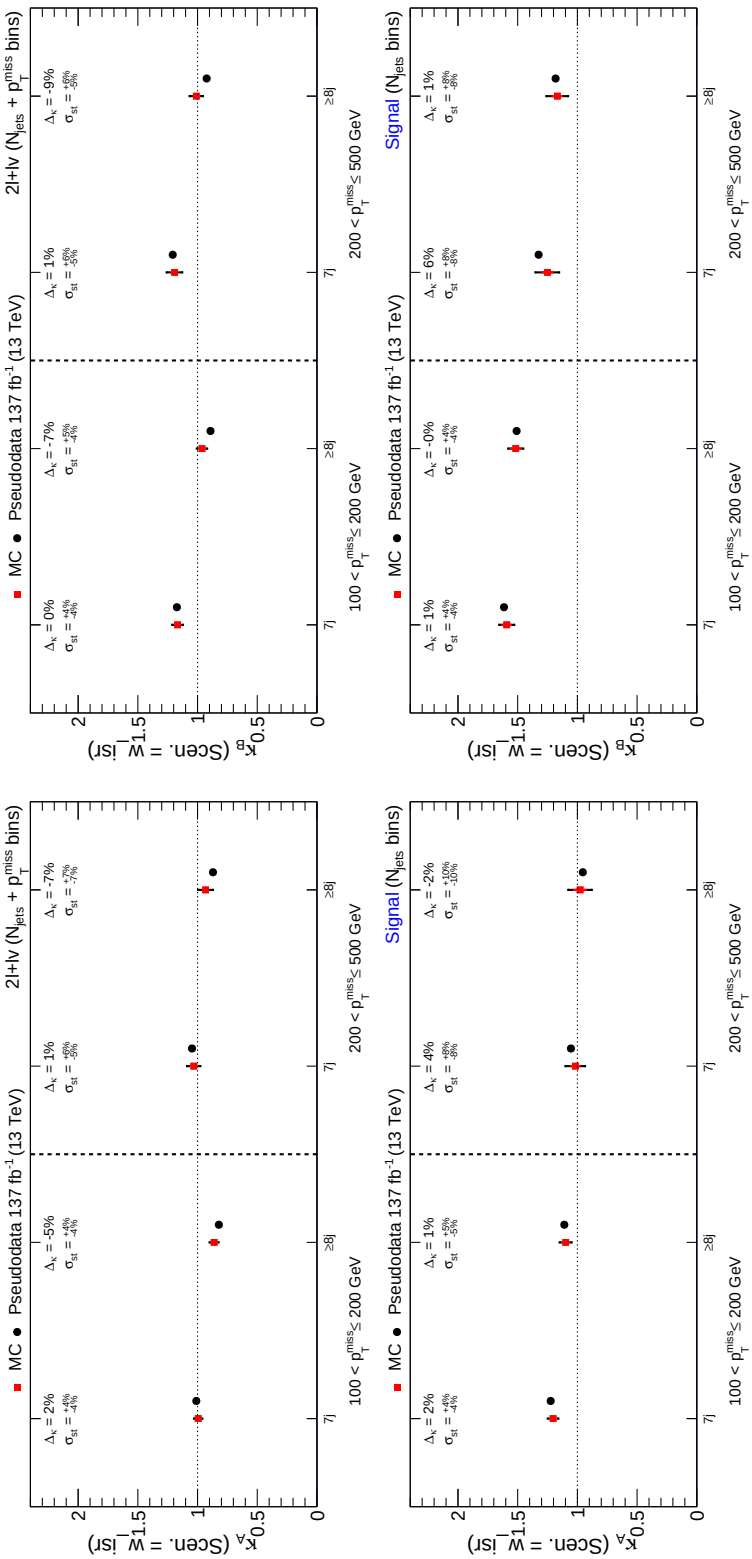


Figure 9.9: Changes in κ caused by applying ISR reweighting to semileptonic $\mathrm{t\bar{t}}$ events. Left and right columns show low-M_J and high-M_J ABCDs, respectively. Comparing the top and bottom row shows that the dilepton sample binned in N_{jets} (top row) correctly models the increased dependence of κ on jet multiplicity observed in the signal region (bottom row).

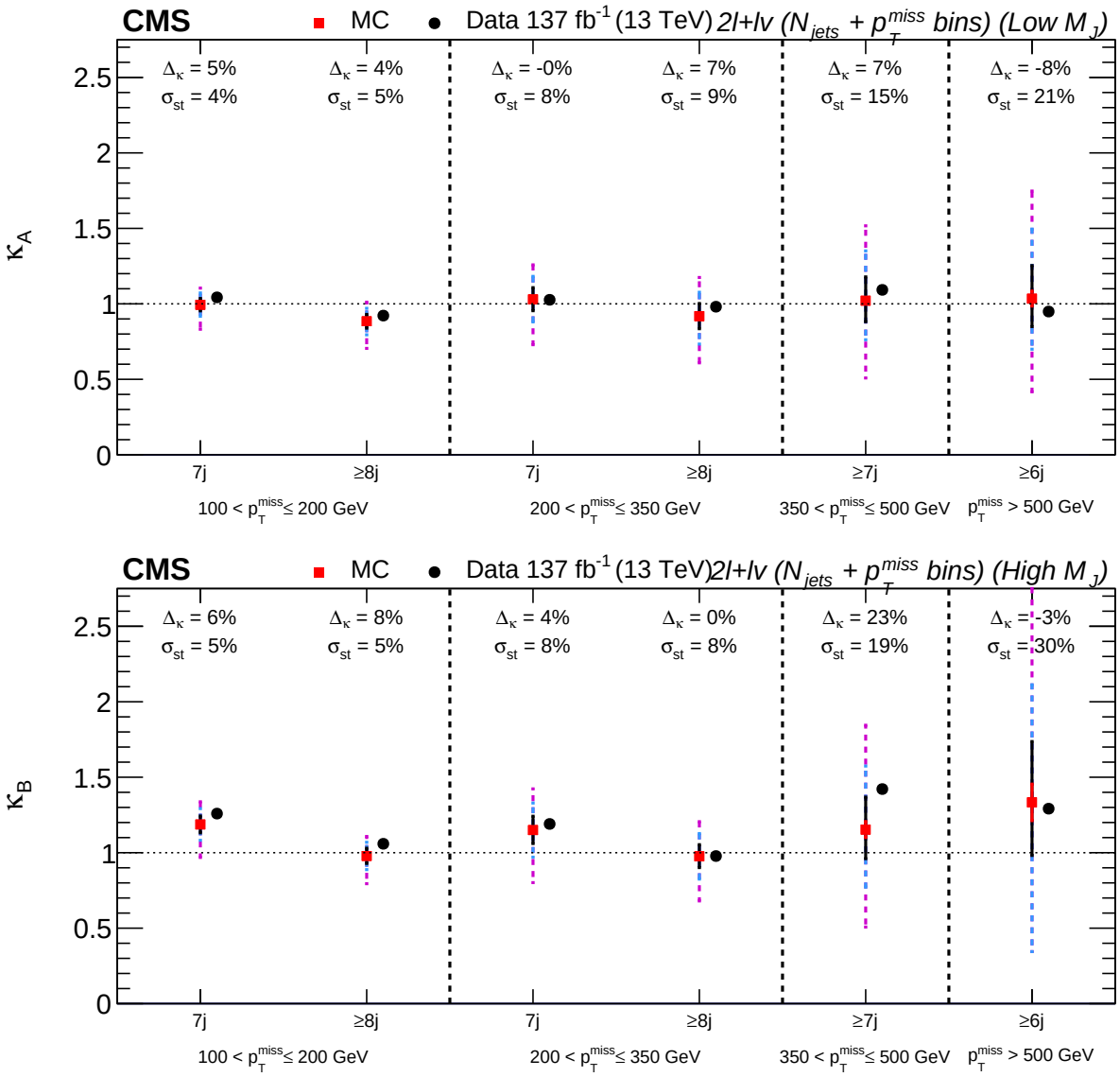


Figure 9.10: 2ℓ control sample κ values found in MC vs. data. The expected uncertainty of the data given the data set integrated luminosity, summed in quadrature with the statistical uncertainty of the simulated samples, is given by the error bar on the red points (σ_{st}). The red portion of the error bar indicates the contribution from the limited statistical power of the simulated samples. The solid black, blue dashed and pink dashed show 1σ, 2σ, and 3σ total statistical uncertainty, respectively. The values of Δ_κ are the relative difference between the κ values found in simulation and in data.

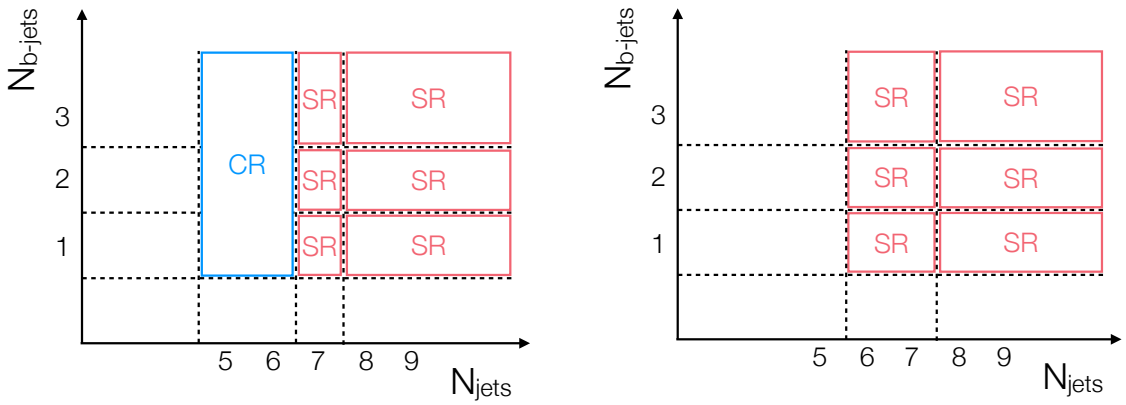
Results

With the 2ℓ control sample defined and validated, we perform the background estimation in data and MC, looking at the low-$p_{\mathrm{T}}^{\mathrm{miss}}$ region in two bins of N_{jets} and the mid- and high-$p_{\mathrm{T}}^{\mathrm{miss}}$ regions integrated over N_{jets}. The resulting κ values found in simulation and data are compared in Figure 9.10, with the ultra-low $p_{\mathrm{T}}^{\mathrm{miss}}$ bin shown as an auxiliary validation. Given the κ values we observe in data are consistent with those we see in MC, we take the statistical uncertainty (σ_{st}) in each N_{jets} bin and $p_{\mathrm{T}}^{\mathrm{miss}}$ bin and include it in the systematic uncertainty for signal bins with the corresponding $p_{\mathrm{T}}^{\mathrm{miss}}$ and jet multiplicity. Taking the low-N_{jets}, high-$p_{\mathrm{T}}^{\mathrm{miss}}$ search regions as an example, the N_{jets} systematic uncertainty is taken from the low-N_{jets} low-$p_{\mathrm{T}}^{\mathrm{miss}}$ bin of the 2ℓ control sample, while the $p_{\mathrm{T}}^{\mathrm{miss}}$ systematic uncertainty is taken from the N_{jets}-inclusive high-$p_{\mathrm{T}}^{\mathrm{miss}}$ bin of the 2ℓ control sample. The various sources of systematic uncertainty are summed in quadrature to get the total systematic uncertainty of a bin. As all N_{jets} bins have similar statistical uncertainties, the maximum (9%) is assigned as the N_{jets} systematic uncertainty in all search region bins. The $p_{\mathrm{T}}^{\mathrm{miss}}$ systematic uncertainties assigned to the mid- and high-$p_{\mathrm{T}}^{\mathrm{miss}}$ bins are taken as 15% and 21% (19% and 30%) for low-M_J (high-M_J). The total systematic uncertainty calculation is shown in Section 9.4.

9.2.2 5-6 jet control sample

Motivation

The 2ℓ control sample allowed us to isolate the dominant background in our search region. Unfortunately, such a pure control sample is not possible for the other high-m_{T} categories. Instead, we find a control sample with the same composition as the search region. As we saw in Figure 9.6, this requirement is satisfied by events with five or six jets. In this sample, both κ values for individual categories and the admixture of



(a) $200 < p_{\mathrm{T}}^{\mathrm{miss}} \leq 500\,\mathrm{GeV}$ (b) $p_{\mathrm{T}}^{\mathrm{miss}} > 500\,\mathrm{GeV}$

Figure 9.11: N_{jets}-N_{b} planes in the low+intermediate bins of $p_{\mathrm{T}}^{\mathrm{miss}}$ (left), and in the high-$p_{\mathrm{T}}^{\mathrm{miss}}$ bin.

background types are similar to the higher N_{jets} search region. This sample, which we will refer to as the 5-6 jet control sample, is used to quantify mismodeling of κ due to mismeasurement or due to misprediction of the background composition as a function of N_{b}. As was done for the N_{jets} uncertainties in the 2ℓ control sample, the N_{b} uncertainties are derived in the low-$p_{\mathrm{T}}^{\mathrm{miss}}$ bin and applied to all $p_{\mathrm{T}}^{\mathrm{miss}}$ regions; This choice was based on studies of simulated events that showed the N_{b} dependence of κ is not correlated with $p_{\mathrm{T}}^{\mathrm{miss}}$. Additionally, as the lowest $p_{\mathrm{T}}^{\mathrm{miss}}$ bin has the highest fraction of mismeasured events, the N_{b} uncertainties will also account for the systematic uncertainty associated with mismeasurement.

Definition

Except for the N_{jets} requirement, the 5-6 jet control sample shares all the selection criteria with the search region. Since the high $p_{\mathrm{T}}^{\mathrm{miss}}$ bin of the search region includes events with $N_{\mathrm{jets}} = 6$, only the low and mid $p_{\mathrm{T}}^{\mathrm{miss}}$ bins are considered in the control sample. Figure 9.11 shows the binning of the search regions and 5-6 jet control sample

in the N_{jets}-N_{b} plane for two $p_{\text{T}}^{\text{miss}}$ regions. As they share the same N_{b} and N_{jets} binning, the low and intermediate $p_{\text{T}}^{\text{miss}}$ bins are combined.

Validation

As we did in the 2ℓ control sample, we perform a study in which we alter the weights of a subset of MC events and verify that this change has the same effect on simulated κ values in the control sample and the search region. In this case, we double the weight of mismeasured 1ℓ events at high-m_{T} and high-M_J, a change that effectively doubles the κ values for the mismeasured event category. The κ values before and after this change are shown in Figure 9.12, with the 2ℓ control sample included to show this sample is unaffected by mismodeling of mismeasured events. As expected, this change has the largest effect on the lowest $p_{\text{T}}^{\text{miss}}$ bin, with the effect becoming smaller as we go higher in $p_{\text{T}}^{\text{miss}}$, reflecting the relative contribution of mismeasured events we saw in Figure 9.6. This behavior is also consistent between the 5-6 jet control sample and the search region, with the Δ_κ introduced by the altered weights being approximately equal between the samples. Though we do not expect mismeasured events to be mismodeled this severely, this study shows the 5-6 jet control sample accurately represents the composition of the search region.

Results

κ values are calculated in three N_{b} bins at low $p_{\text{T}}^{\text{miss}}$ and integrated over N_{b} in the intermediate $p_{\text{T}}^{\text{miss}}$ bin. In the absence of significant disagreement between data and MC, the statistical uncertainties (σ_{st}) of the low $p_{\text{T}}^{\text{miss}}$ bins would be included in the systematic uncertainties of search region bins with the corresponding b jet multiplicity. As discussed in the motivation section, this uncertainty also accounts for the systematic uncertainty associated with mismeasurement. The results are shown in Figure 9.13, with the ultra-low

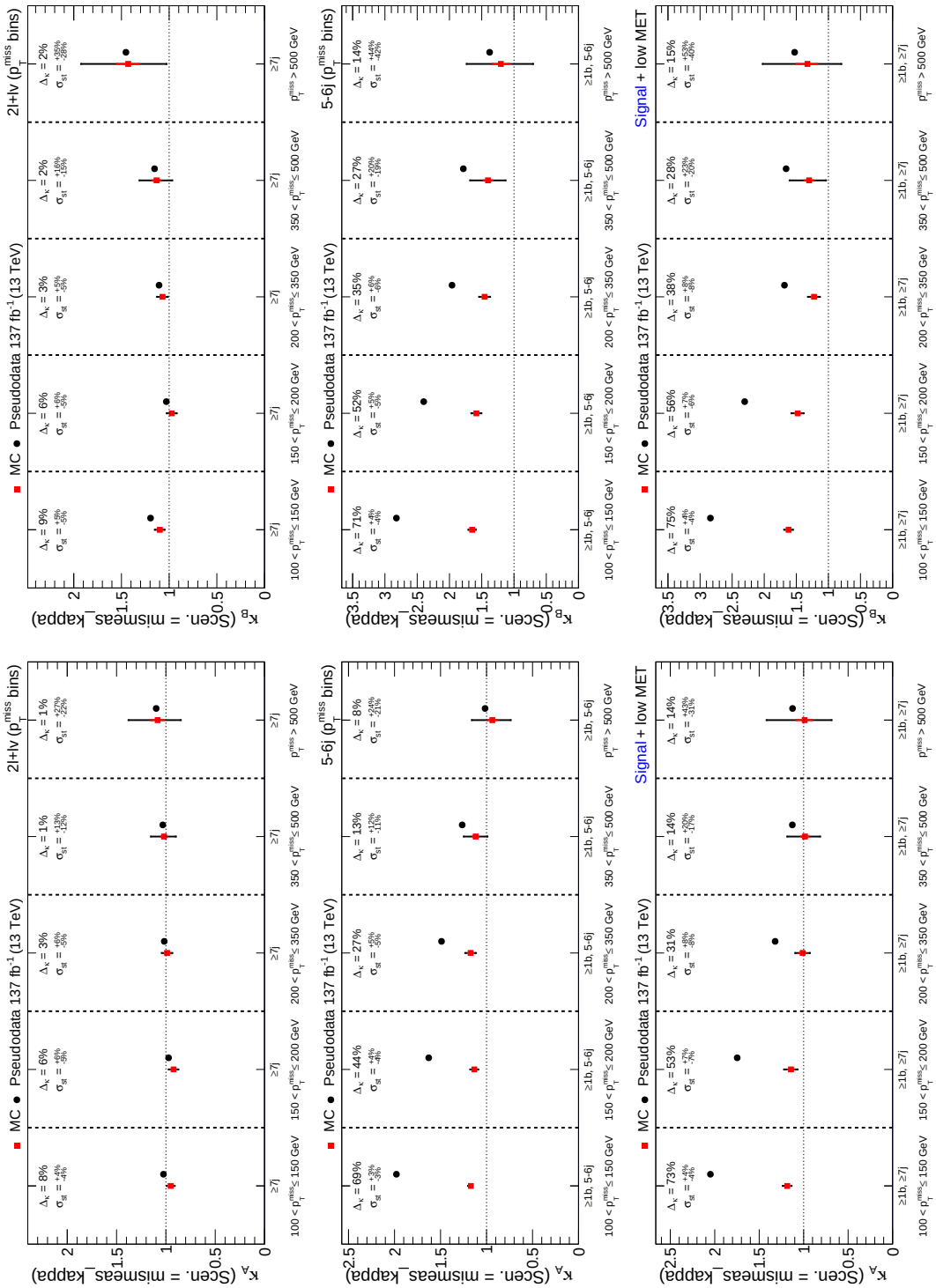


Figure 9.12: Changes in κ caused by doubling the number of mismeasured events in region R4. Left and right columns show low-M_J and high-M_J ABCDs, respectively. The top row shows that the 2ℓ control sample is unaffected. Middle row shows that the 5-6 jet control sample captures the change as seen in the signal sample (bottom row).

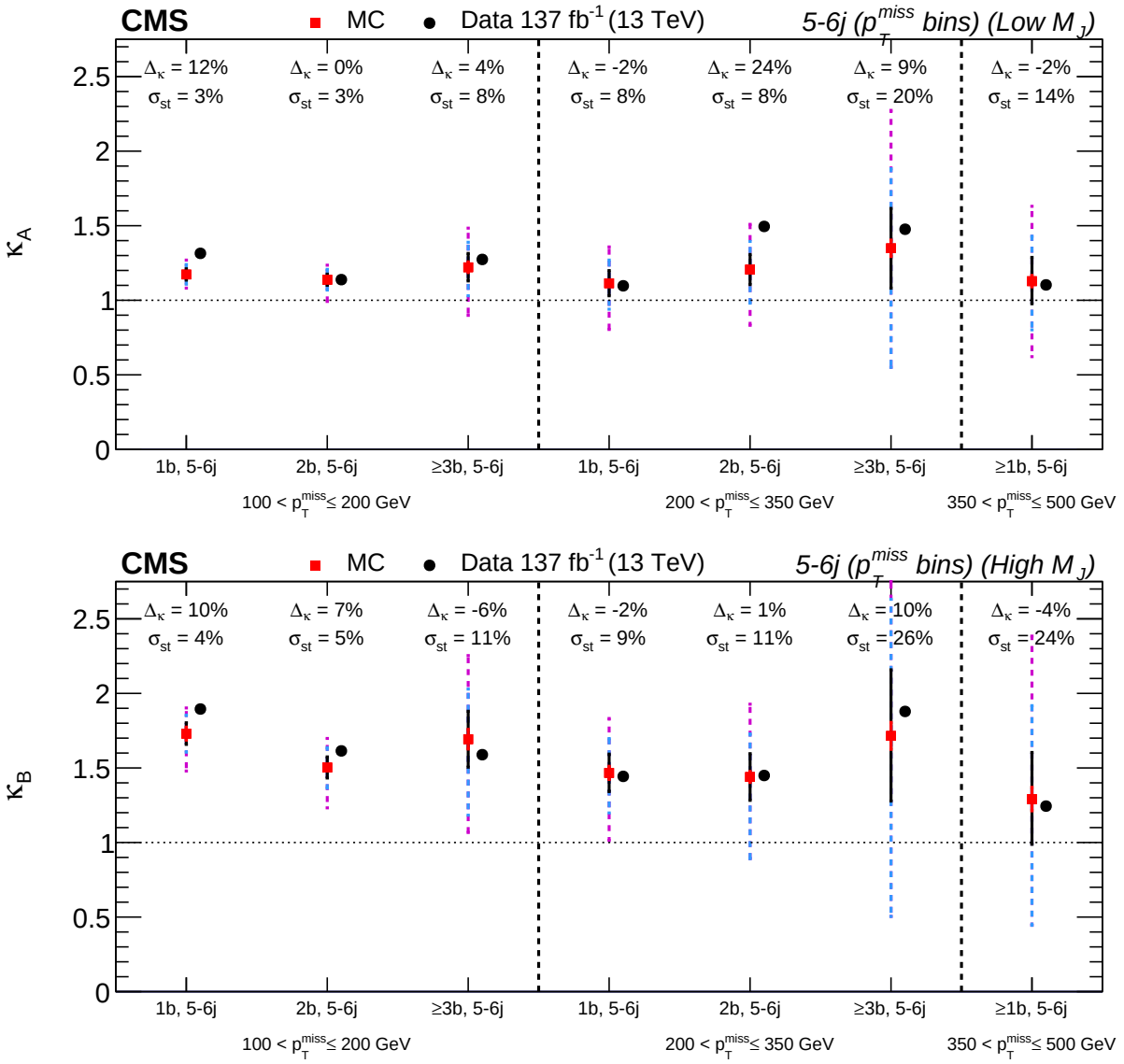


Figure 9.13: Validation of κ values found in MC vs. data. The expected uncertainty of the data given the data set integrated luminosity, summed in quadrature with the statistical uncertainty of the simulated samples, is given by the error bar on the red points (σ_{st}). The red portion of the error bar indicates the contribution from the limited statistical power of the simulated samples. The solid black, blue dashed and pink dashed show 1σ, 2σ and 3σ total statistical uncertainty, respectively. The values of Δ_κ are the relative difference between the κ values found in simulation and in data. The values of Δ_κ are the relative difference between the κ values found in simulation and in data. Validation for each year separately is shown in

$p_{\mathrm{T}}^{\mathrm{miss}}$ region—binned in N_{b}—included as an auxiliary validation. We observe a deviation of κ_{A} in the $N_{\mathrm{b}} = 2$ bin in the low $p_{\mathrm{T}}^{\mathrm{miss}}$ region with $\Delta_{\kappa}/\sigma_{\mathrm{st}} \approx 3$. Further examination of the N_{b} distribution gives no signs of systematic mismodeling, nor is this behavior observed at high-M_J (κ_{B}) or in the other validation regions. Nevertheless, to account for this discrepancy, we assign conservative systematic uncertainties of 10%, 20%, and 30% to events with $N_{\mathrm{b}} = 1$, $N_{\mathrm{b}} = 2$, and $N_{\mathrm{b}} = 3$, respectively.

We also observe a significant difference between data and MC in the $N_{\mathrm{b}} = 1$ bins of the ultra-low $p_{\mathrm{T}}^{\mathrm{miss}}$ validation region; this difference occurs in both M_J regions (for κ_{A} and κ_{B}), being more pronounced at low-M_J with $\Delta_{\kappa}/\sigma_{\mathrm{st}} \approx 4$. Though this is outside our analysis selection, we performed an study of the ultra-low $p_{\mathrm{T}}^{\mathrm{miss}}$ region to investigate it further.

9.3 Ultra-Low $p_{\mathrm{T}}^{\mathrm{miss}}$ Validation

To further study the behavior of κ at ultra-low $p_{\mathrm{T}}^{\mathrm{miss}}$, we use two $p_{\mathrm{T}}^{\mathrm{miss}}$ bins: $100 - 150\,\mathrm{GeV}$ and $150 - 200\,\mathrm{GeV}$. We then calculate κ in bins of N_{jets} to verify that the uncertainties of the $N_{\mathrm{jets}} \geq 7$ region can be extrapolated from the 5-6 jet control sample. If the higher N_{jets} bins show a larger difference between data and MC than the 5-6 bin in the same $p_{\mathrm{T}}^{\mathrm{miss}}$ range, then the uncertainties derived in the 5-6 jet control sample may be underestimating the systematic uncertainty. We also calculate κ in bins of N_{b} to further test the modeling of the N_{b} dependence. The results of these checks are shown in Figure 9.14. As there are two separate binning configurations used, events are being represented twice—once in the N_{jets}-binned κ values in the left two panels, and once in the N_{b}-binned κ values in the right two panels. No significant systematic disagreement between data and MC is observed in the ultra-low $p_{\mathrm{T}}^{\mathrm{miss}}$ sample, so no additional systematic uncertainties are assigned.

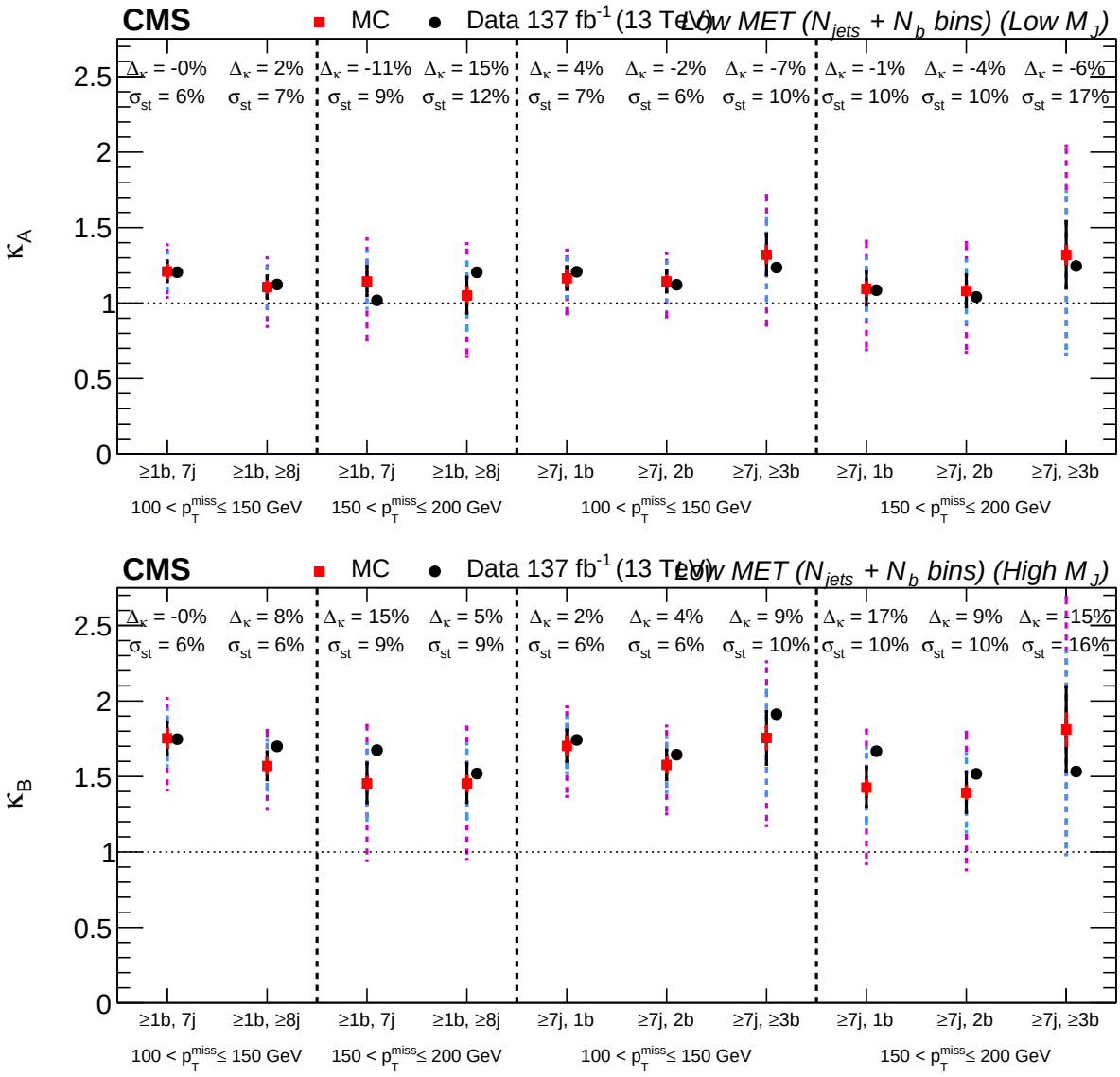


Figure 9.14: Validation of κ values found in MC vs. data. The expected uncertainty of the data given the data set integrated luminosity, summed in quadrature with the statistical uncertainty of the simulated samples, is given by the error bar on the red points (σ_{st}). The red portion of the error bar indicates the contribution from the limited statistical power of the simulated samples. The solid black, blue dashed and pink dashed show 1σ, 2σ and 3σ total statistical uncertainty, respectively. The values of Δ_κ are the relative difference between the κ values found in simulation and in data. Validation for each year separately is shown in

Table 9.1: Systematic uncertainties assigned in each signal bin, as measured in $137\,\text{fb}^{-1}$ of data.

Bin	$200 < p_\text{T}^\text{miss} \leq 350\,\text{GeV}$			$350 < p_\text{T}^\text{miss} \leq 500\,\text{GeV}$			$p_\text{T}^\text{miss} > 500\,\text{GeV}$		
	1 b	2 b	$\geq$ 3b	1 b	2 b	$\geq$ 3b	1 b	2 b	$\geq$ 3b
low-M_J (R4A)	13%	22%	27%	20%	27%	31%	25%	30%	34%
high-M_J (R4B)	13%	22%	27%	22%	28%	32%	32%	36%	39%

9.4 Total Background Systematic Uncertainty

Now that we have established and quantified the sources of systematic uncertainty in the background estimation method, we will combine them to get the total systematic uncertainties. This will be done separately for each bin in the search region by summing all contributing uncertainties in quadrature, assuming there is no correlation between the difference sources of uncertainty. Factored into the uncertainties provided by each control region for each p_T^miss bin, this combination has the form:

$$\sigma^{\text{SR}}_{\text{low } p_\text{T}^\text{miss}\text{, j, b}} = \sigma^{2\ell}_{\text{low } p_\text{T}^\text{miss}\text{, j}} \oplus \sigma^{\text{5-6j CR}}_{\text{low } p_\text{T}^\text{miss}\text{, b}} \tag{9.1}$$

$$\sigma^{\text{SR}}_{\text{mid } p_\text{T}^\text{miss}\text{, j, b}} = \sigma^{2\ell}_{\text{low } p_\text{T}^\text{miss}\text{, j}} \oplus \sigma^{\text{5-6j CR}}_{\text{low } p_\text{T}^\text{miss}\text{, b}} \oplus \sigma^{2\ell}_{\text{mid } p_\text{T}^\text{miss}} \tag{9.2}$$

$$\sigma^{\text{SR}}_{\text{high } p_\text{T}^\text{miss}\text{, j, b}} = \sigma^{2\ell}_{\text{low } p_\text{T}^\text{miss}\text{, j}} \oplus \sigma^{\text{5-6j CR}}_{\text{low } p_\text{T}^\text{miss}\text{, b}} \oplus \sigma^{2\ell}_{\text{high } p_\text{T}^\text{miss}}, \tag{9.3}$$

where the j and b indices represent jet and b jet multiplicity, respectively. This combination is represented graphically in Figure 9.15, which also shows the uncertainties acquired

$$\sigma^2_{tot} = \sigma^2_{2\ell,\, low\, p_T^{miss}}(N_{jets}) + \sigma^2_{5-6j,\, low\, p_T^{miss}}(N_b, p^{miss}_{T,mismeas.}) + \sigma^2_{2\ell}(p_T^{miss})$$

$\sigma_{2\ell,\, low\, p_T^{miss}}(N_{jets})$	
All	9%

$\sigma_{5-6j,\, low\, p_T^{miss}}(N_b, p^{miss}_{T,mimeas.})$	1b	2b	≥ 3b
Low M_J	10%	20%	25%
High M_J	10%	20%	25%

$\sigma_{2\ell}(p_T^{miss})$	Mid p_T^{miss}	High p_T^{miss}
Low M_J	15%	21%
High M_J	18%	29%

Figure 9.15: Combination of background systematic uncertainties, showing the contributions from the control samples for each search region bin. The results of this combination are summarized in Table 9.1.

from each control region. The resulting uncertainties are summarized in Table 9.1, ranging from 13% to 39%, increasing with $p_{\mathrm{T}}^{\mathrm{miss}}$.

In practice, the three terms of Eq. (9.3) are implemented as six log-normal constraints, with a separate nuisance parameter for the low-M_J and high-M_J regions used for each term. This separation is motivated by the observation that the deviation of κ from one has significantly different $p_{\mathrm{T}}^{\mathrm{miss}}$ dependence at low M_J and high M_J.

9.5 Signal Model Systematics

When determining the systematic uncertainties in the background estimation method, we use control regions with the same composition as the search region. This allows us to test the ABCD method in these control regions and assign the resulting uncertainties as systematic uncertainties in the search region. This approach, however, will not work for determining the systematic uncertainties associated with the signal model, as the kinematic properties of the signal change significantly depending on the sparticle masses and model being considered. Instead, dedicated studies are performed for each source of signal uncertainty, and these values are evaluated for each analysis bin separately. Table 9.2 lists the sources of systematic uncertainty associated with the signal model (including detector effects), and gives a range of values for the most sensitive bins for two signal models. We will briefly discuss each of these sources and how the systematic uncertainties are determined.

As we saw in Section 8.4.1, systematic uncertainties are accounted for with a symmetric Log-Normal term in the likelihood. In cases where the derived uncertainties are asymmetric–the sizes of up and down variations are not the same–the larger of the two uncertainties is used in the likelihood. Additionally, some uncertainties have an effect on the signal yield that is anti-correlated between bins. For example, an upward variation

Table 9.2: Characteristic range of values for the systematic uncertainties in the signal efficiency and acceptance across sensitive bins, specifically across high $p_{\mathrm{T}}^{\mathrm{miss}}$ signal bins for T1tttt(2100,100) and high N_{jets} signal bins for T1tttt(1900,1250). Uncertainties due to a particular source are treated as fully correlated among bins, while uncertainties due to different sources are treated as uncorrelated.

Source	Relative uncertainty [%]	
	T1tttt(2100,100)	T1tttt(1900,1250)
MC sample statistics	3–8	7–15
Renormalization and factorization scales	1–2	2–4
Fast sim. $p_{\mathrm{T}}^{\mathrm{miss}}$ resolution	1–2	1–5
Lepton efficiency	3	3
Fastsim lepton efficiency	6–8	3–4
Trigger efficiency	1	1
b tagging efficiency	2–8	2–8
Fast sim. b tagging efficiency	1–2	1–2
Mistag efficiency	1	1
Fast sim. mistag efficiency	<1	1–3
Jet energy corrections	1–5	2–11
Initial-state radiation	1–7	1–10
Jet ID	1	1
Pileup	1–2	1–4
Integrated luminosity	2.5	2.5

in b tagging efficiency may increase the signal yield in high-N_{b} bins and decrease the yield in low-N_{b} bins. These anti-correlated effects are represented by negative values in the bin-by-bin uncertainties given in Table 9.2, where two oppositely signed values in a given row are anti-correlated. This is implemented in the likelihood by translating the fractional uncertainties $\sigma_{i,j,k}$ from Table 9.2 into the fit constants

$$\varsigma_{i,j,k} = \begin{cases} \ln(1 + \left|\sigma_{i,j,k}\right|), & \sigma_{i,j,k} \geq 0, \\ -\ln(1 - \left|\sigma_{i,j,k}\right|), & \sigma_{i,j,k} < 0. \end{cases} \tag{9.4}$$

Uncertainties from different rows are treated as uncorrelated with each other.

Systematic uncertainties arising from experimental effects are typically determined from comparisons between data and simulated event samples. For an example of this, let's consider the systematic uncertainties associated with lepton efficiency. We saw in Section 5.2 how scale factors for lepton identification are calculated and applied in order to make simulated event samples more representative of data. By varying these scale factors within their uncertainties, we can quantify the effect they have on signal yields in the search region. We then take the observed change of 3% as the systematic uncertainties associated with lepton efficiency.

Systematic uncertainties associated with b tagging and mistagging efficiencies are determined in a similar way to those used for leptons. The scale factor for each jet is varied within its uncertainties, changing the event weights and thus the signal yield in each bin. The size of the change induced by this variation in each bin is taken as the systematic uncertainty for that bin. Uncertainties associated with b tagging range from 2-9%, while mistagging uncertainties range from $<$ 1-2%.

As discussed in Chapter 6, a different detector simulation is used for signal samples to reduce the computing resources required to general the full scan of mass points. These FastSim samples model event variables with less accuracy than those used for the back-

ground; therefore, scale factors are applied to make the FastSim match the FullSim with respect to lepton identification, b tagging, and mistagging efficiencies. Since the previously discussed scale factors account for differences between FullSim and data, both sets of scale factors are applied to make FastSim samples match data. Systematic uncertainties for the FastSim scale factors are calculated in the same way as those for the first set of scale factors.

Additional weights are applied to account for differences in the number of primary vertices (pileup) between data and simulated samples. These weights are varied in a similar fashion to scale factors, and the corresponding change in the signal yield in each bin is taken as the systematic uncertainty.

The procedure for calculating systematic uncertainties due to jet-energy corrections (JECs) is slightly more involved than that for scale factors, as changes in jet energy effect key event variables such as M_J, $p_{\mathrm{T}}^{\mathrm{miss}}$, and m_{T}. Jet momenta are varied within the uncertainties of the corrections applied to them. Each event is then reconstructed using the modified jet so the variables sensitive to jet momentum are recalculated. The resulting changes in signal yields for each bin are taken as the systematic uncertainty. Furthermore, as the variables for jet identification are not properly modeled in FastSim, jets in signal samples are assumed to pass the jet identification requirements. This is accounted for with a 1% systematic uncertainty assigned to all bins.

QCD scale systematic uncertainties account for kinematic differences between leading and next-to-leading order contributions to the signal. With the cross section held constant, the renormalization and factorization scales are each varied up and down by a factor of two, in such a way that they stay within a factor of two of each other (if one is doubled, the other can not be halved). These variations result in assigned uncertainties of 1-5%, with the effect being largest in compressed signal models. Similarly, the uncertainties associated with ISR modeling are determined by varying ISR weights within

their uncertainties, with the resulting change in signal yields in each bin taken as the systematic uncertainty in that bin. As in the case of QCD scale, ISR weights have a more significant effect on compressed signal models.

Table 9.3: Bin-by-bin systematic uncertainties for two signal models, showing the relative uncertainties for the low-M_J (black) and high-M_J (blue) regions. Uncertainties from one source are considered fully correlated between bins and opposite signs indicate anti-correlation. Different sources of uncertainties are considered uncorrelated.

Uncertainty [%]	$200 < p_{\mathrm{T}}^{\mathrm{miss}} \leq 350$						$350 < p_{\mathrm{T}}^{\mathrm{miss}} \leq 500$						$p_{\mathrm{T}}^{\mathrm{miss}} > 500$					
	Low N_{jets}			High N_{jets}			Low N_{jets}			High N_{jets}			Low N_{jets}			High N_{jets}		
	$N_{\mathrm{b}}=1$	$N_{\mathrm{b}}=2$	$N_{\mathrm{b}}\geq 3$	$N_{\mathrm{b}}=1$	$N_{\mathrm{b}}=2$	$N_{\mathrm{b}}\geq 3$	$N_{\mathrm{b}}=1$	$N_{\mathrm{b}}=2$	$N_{\mathrm{b}}\geq 3$	$N_{\mathrm{b}}=1$	$N_{\mathrm{b}}=2$	$N_{\mathrm{b}}\geq 3$	$N_{\mathrm{b}}=1$	$N_{\mathrm{b}}=2$	$N_{\mathrm{b}}\geq 3$	$N_{\mathrm{b}}=1$	$N_{\mathrm{b}}=2$	$N_{\mathrm{b}}\geq 3$
	Signal model: T1tttt(2100,100)																	
MC sample statistics	66 16	64 11	49 12	76 13	46 8	49 6	20 13	19 11	19 14	29 10	16 7	15 6	6 5	5 4	5 4	8 4	6 3	5 3
FastSim MET	−15 0	21 −4	4 −5	46 11	19 5	−9 1	4 8	−2 5	11 5	2 −4	6 −4	−1 1	−2 −1	0 −1	−1 −1	−1 −1	−1 0	0 −1
Lepton efficiency	4 3	3 3	4 3	3 3	3 4	3 3	3 3	3 4	3 3	3 3	3 3	3 3	3 3	3 3	3 3	3 3	3 3	3 3
FastSim lepton efficiency	4 9	6 10	3 11	7 8	8 6	3 7	13 10	6 15	7 8	8 7	7 7	9 8	6 8	7 8	6 7	8 8	6 7	6 6
Trigger efficiency	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1
b-tag efficiency	−5 −6	−2 −1	5 7	−8 −7	0 −1	4 7	−7 −6	1 −1	8 7	−6 −9	−1 −2	7 6	−6 −6	0 −1	8 9	−7 −8	−2 −2	6 7
FastSim b-tag efficiency	−1 −2	1 0	2 3	−2 −2	0 0	2 2	−2 −2	1 0	2 3	−1 −3	0 0	2 2	−2 −3	0 0	3 3	−1 −2	0 0	2 2
Mistag efficiency	−2 −1	−1 −1	−1 0	−1 −1	−2 0	−1 1	−1 −1	−1 0	0 2	−2 −1	−1 −1	3 1	−1 −1	0 0	1 1	−1 −1	−1 −1	1 1
FastSim mistag efficiency	−1 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0
Jet energy corrections	−23 7	20 −6	99 −1	−99 −2	17 2	96 1	−8 −8	−7 5	−2 −8	−21 3	−11 4	−5 4	−1 2	−3 2	−4 2	3 4	5 5	1 5
QCD Scales	2 1	6 1	1 2	2 2	3 2	3 1	2 2	2 1	1 2	2 2	2 1	2 2	1 1	1 1	1 1	1 2	1 2	1 2
ISR	−3 −2	−6 −3	−8 −4	9 6	−4 5	0 4	−3 −1	−1 −1	−4 −4	0 5	0 7	0 5	−3 −3	−4 −3	−4 −4	2 7	2 5	0 5
Jet ID	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1
Pile up	−13 6	−17 1	9 0	−8 4	2 1	6 1	1 4	−3 −1	−6 2	−6 0	3 1	0 2	1 −1	0 0	1 1	−1 2	0 1	1 1
Luminosity	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2
	Signal model: T1tttt(1900,1250)																	
MC sample statistics	18 18	12 11	10 9	15 10	10 6	7 4	15 26	12 19	9 20	12 14	7 8	5 6	12 29	8 27	8 24	11 15	7 9	5 7
FastSim MET	6 −2	6 −1	−4 3	1 2	−3 1	3 4	−1 3	−2 2	−1 8	−3 0	2 0	−2 −2	−2 5	0 −1	−4 −4	1 −4	−5 −1	−1 0
Lepton efficiency	3 4	3 3	3 3	3 3	3 3	3 3	3 3	3 3	3 3	3 3	3 3	3 3	3 3	3 3	3 3	3 3	3 3	3 3
FastSim lepton efficiency	3 3	3 4	3 4	3 3	4 4	3 4	4 5	4 5	4 3	3 4	3 4	3 4	3 3	4 6	4 5	4 4	4 4	3 4
Trigger efficiency	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1	−1 −1
b-tag efficiency	−5 −5	−1 −1	4 4	−7 −6	−3 −3	3 3	−5 −5	−1 −2	4 5	−6 −8	−2 −3	3 4	−5 −5	−1 −2	5 6	−6 −8	−3 −3	4 4
FastSim b-tag efficiency	−1 −1	0 0	1 1	−1 −1	0 0	0 1	−1 −1	0 0	1 1	−1 −1	0 −1	1 1	−1 −2	0 −1	1 1	−1 −2	0 −1	1 1
Mistag efficiency	−1 −1	−1 −1	0 0	−1 −1	−1 −1	1 0	0 0	0 −1	0 −1	−1 −1	−1 −1	0 0	−1 −1	1 −1	1 −1	−1 −1	−1 −1	0 1
FastSim mistag efficiency	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 3
Jet energy corrections	11 −9	−10 −8	−11 5	−2 8	2 4	2 4	−2 20	14 −5	2 −5	2 6	7 9	3 7	7 8	4 6	−4 18	9 5	6 5	10 11
QCD Scales	1 2	1 2	2 2	2 2	1 2	2 1	1 2	2 1	2 2	2 3	1 3	1 3	1 4	1 3	2 3	3 4	2 4	2 4
ISR	−1 −2	−2 −3	−3 −3	−2 5	1 3	0 3	−2 −2	−2 −2	−3 −2	1 6	1 4	1 4	−2 −1	−2 0	−3 −2	3 8	2 10	2 10
Jet ID	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1	1 1
Pile up	1 −1	4 −1	2 1	−1 1	2 2	−1 1	1 −1	−2 −3	0 6	3 2	1 2	0 1	−1 3	2 −6	1 −3	1 −4	0 0	1 0
Luminosity	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2

Chapter 10

Results and Interpretations

In Chapter 7 we defined signal-rich search regions based on signal simulations. Until now, we have not examined how these search regions look in data. Instead, we have relied on studies of simulated samples and control regions in data to design and refine the analysis. This analysis strategy is known as *blinding*, and it ensures that the analysis remains unbiased. In Chapter 8 we established a method for predicting the number of standard model background events we expect to observe in our search regions, and checked how these predictions performed in data control regions. We then quantified the systematic uncertainties in the background estimation method, as well as the systematic uncertainties in the yields predicted by the signal model, in Chapter 9.

Now is the moment of truth: it's time to look at the search regions in data. We start by evaluating the background using the ABCD method to obtain an expected number of standard model events for each search region bin. Next, we compare the expected yields in each bin with those actually observed, determining how consistent they are with one another. Finally, we will interpret the level of consistency in the context of our signal models T1tttt and T5tttt, placing upper limits on the cross section as a function of $m_{\tilde{\chi}1}^{0}$ and $m_{\tilde{g}}$.

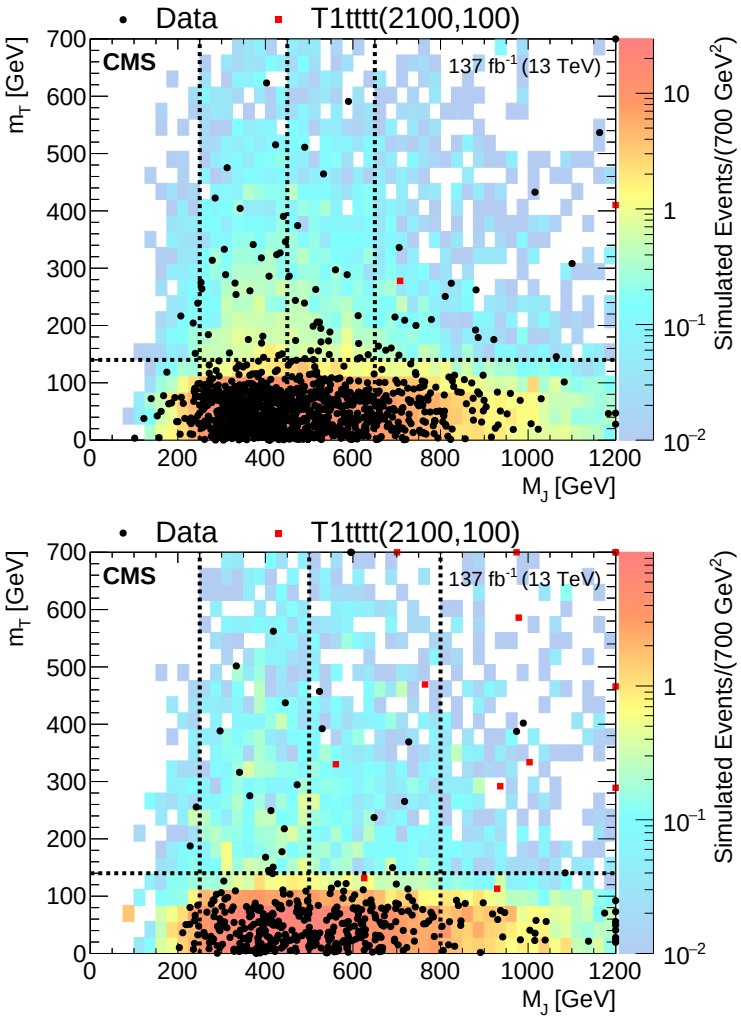


Figure 10.1: Two-dimensional distributions for data and simulated event samples integrated over the N_{jets} and $N_{\text{b}} \geq 2$, shown for the $p_{\text{T}}^{\text{miss}}$ 350-500 GeV bin (top) and the $p_{\text{T}}^{\text{miss}}$ $\geq$500 GeV bin (bottom). The black dots are the data; the colored histogram is the total simulated background, normalized to the data; and the red dots are a particular signal sample drawn from the expected distribution for gluino pair production in the T1tttt model with $m_{\tilde{\text{g}}}$ =2100 GeV and $m_{\tilde{\chi}_1^0}$ =100 GeV for 137 fb^{-1}. Overflow events are shown on the edges of the plot.

10.1 Estimated Backgrounds

Figure 10.1 provides an overview of the observed data, showing the distribution of events in the M_J-m_{T} plane for the two highest $p_{\text{T}}^{\text{miss}}$ bins. The black dots are data, the underlying color map shows the total simulated background normalized to data, and the

red dots show a representative distribution of T1tttt signal events with $m_{\tilde{g}} = 2100\,\text{GeV}$ and $m_{\tilde{\chi}1}^{0} = 100\,\text{GeV}$ for $137\,\text{fb}^{-1}$. The large number of data points in the low $p_{\text{T}}^{\text{miss}}$ bin obscure the other plot features, so this region is not shown.

Before performing the background estimation, we should check if the assumption we made when defining the ABCD method are supported by the data. Namely, are M_J and m_{T} uncorrelated in our search region? Figure 8.1 showed this was true for 1ℓ and 2ℓ $\text{t}\bar{\text{t}}$ simulated samples, but now we will check this in data. The M_J distributions for high- and low-m_{T} events in each $p_{\text{T}}^{\text{miss}}$ bin are shown in Figure 10.2, with the low-m_{T} data (R1, R2A, and R2B) shown in blue and the high-m_{T} data (R3, R4A, and R4B) shown in black. The low-m_{T} events shown are weighted by the corresponding κ value, derived in simulation, and normalized to the high-m_{T} event yield in order to easily compare distribution shapes. The agreement observed between the M_J shapes at high and low m_{T} confirms that the κ factors are able to account for any correlation between M_J and m_{T}.

With the primary assumption of the ABCD method validated, let us estimate the background in our search region. Tables 10.1 and 10.2 show the observed and predicted events yields for all regions in the low-M_J and high-M_J ABCD plane, respectively. Predictions from two different fitting methods are given. The R1–R3 fit is performed using only the observed yields from the R1, R2A, R2B, and R3 regions, while the R1–R4 fit also includes the R4A and R4B regions. In both cases, the signal strength (r) is fixed to zero to perform a background-only fit that is independent of the signal model. Both fits maximize the likelihood function defined in Section 8.4 over all background parameters, with the R4 terms being excluded from the likelihood in the R1–R3 fit. Expected yields from two SUSY models are also included in the tables as an indicator of the most sensitive bins for each model.

One may notice that the R1–R3 fit perfectly predicts the background in the R1, R2,

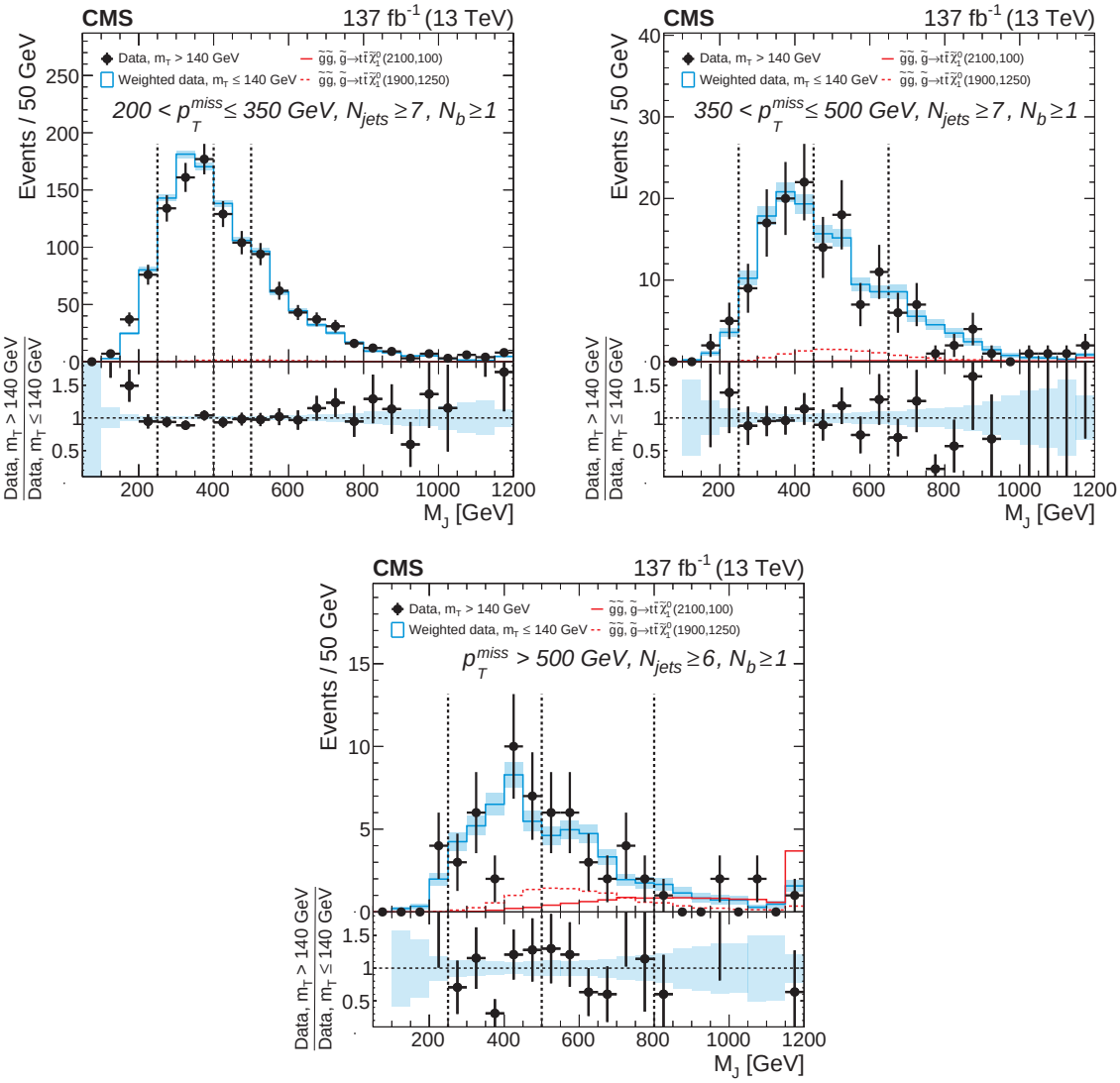


Figure 10.2: M_J distributions observed in data for $200 < p_{\mathrm{T}}^{\mathrm{miss}} \leq 350\,\mathrm{GeV}$ (top left), $350 < p_{\mathrm{T}}^{\mathrm{miss}} \leq 500\,\mathrm{GeV}$ (top right) and $p_{\mathrm{T}}^{\mathrm{miss}} > 500\,\mathrm{GeV}$ (bottom) in the 1ℓ data for low and high m_{T} regions. In each plot, the data at low m_{T} have been weighted by the relevant κ factor and normalized to the yield observed at high m_{T} to facilitate comparison of the shapes of the distributions. The vertical dashed line at $M_J = 250\,\mathrm{GeV}$ shows the lower boundary of regions R1 and R3, while the vertical lines at higher M_J values denote the lower M_J boundaries of the signal regions R4A and R4B. The data are integrated over the N_{b} and N_{jets} signal bins. Two SUSY benchmark models are shown in the solid and dashed red histograms.

Table 10.1: Observed and predicted event yields for the signal regions (R4) and background regions (R1–R3) in the low-M_J ABCDs. Expected yields for the two SUSY T1tttt benchmark scenarios (2100, 100) and (1900,1250) are also given. The uncertainties on the prediction account for the available statistics in the data control samples, the precision of κ from MC, but not yet the systematic uncertainties assessed from control samples in data.

$\mathcal{L} = 137$ fb^{-1}	T1tttt(2100,100)	T1tttt(1900,1250)	R1–R3 fit	R1–R4 fit	Obs.
$200 < p_{\mathrm{T}}^{\mathrm{miss}} \leq 350$ GeV					
R1	0.0	1.1		7705.3 ± 86.8	7706
R2: $N_b = 1$, $N_{jets} = 7$	0.0	0.1		1087.8 ± 32.1	1088
R2: $N_b = 1$, $N_{jets} \geq 8$	0.0	0.1		735.7 ± 26.4	732
R2: $N_b = 2$, $N_{jets} = 7$	0.0	0.1		882.3 ± 29.5	879
R2: $N_b = 2$, $N_{jets} \geq 8$	0.0	0.3		641.6 ± 25.0	644
R2: $N_b \geq 3$, $N_{jets} = 7$	0.0	0.2		235.3 ± 15.1	237
R2: $N_b \geq 3$, $N_{jets} \geq 8$	0.0	0.5		199.6 ± 13.9	202
R3	0.0	2.2		472.7 ± 19.8	472
R4: $N_b = 1$, $N_{jets} = 7$	0.0	0.2	70.0 ± 10.0	70.2 ± 4.6	70
R4: $N_b = 1$, $N_{jets} \geq 8$	0.0	0.3	37.7 ± 5.6	38.3 ± 2.8	42
R4: $N_b = 2$, $N_{jets} = 7$	0.0	0.4	55.5 ± 11.8	55.7 ± 4.5	59
R4: $N_b = 2$, $N_{jets} \geq 8$	0.0	0.6	37.9 ± 8.1	37.4 ± 3.1	35
R4: $N_b \geq 3$, $N_{jets} = 7$	0.0	0.4	19.2 ± 4.9	18.7 ± 2.1	17
R4: $N_b \geq 3$, $N_{jets} \geq 8$	0.0	0.9	12.9 ± 3.3	12.4 ± 1.5	10
$350 < p_{\mathrm{T}}^{\mathrm{miss}} \leq 500$ GeV					
R1	0.0	0.9		968.0 ± 30.6	967
R2: $N_b = 1$, $N_{jets} = 7$	0.0	0.1		206.7 ± 14.1	208
R2: $N_b = 1$, $N_{jets} \geq 8$	0.0	0.2		148.3 ± 12.0	150
R2: $N_b = 2$, $N_{jets} = 7$	0.0	0.1		141.7 ± 11.8	139
R2: $N_b = 2$, $N_{jets} \geq 8$	0.0	0.3		112.2 ± 10.5	111
R2: $N_b \geq 3$, $N_{jets} = 7$	0.0	0.2		30.1 ± 5.3	30
R2: $N_b \geq 3$, $N_{jets} \geq 8$	0.0	0.6		37.7 ± 6.0	38
R3	0.1	2.9		67.0 ± 6.5	68
R4: $N_b = 1$, $N_{jets} = 7$	0.1	0.3	15.2 ± 3.7	15.3 ± 2.1	14
R4: $N_b = 1$, $N_{jets} \geq 8$	0.0	0.4	9.9 ± 2.7	9.7 ± 1.6	8
R4: $N_b = 2$, $N_{jets} = 7$	0.1	0.5	10.8 ± 3.1	11.3 ± 1.7	14
R4: $N_b = 2$, $N_{jets} \geq 8$	0.1	1.3	6.6 ± 1.9	6.8 ± 1.1	8
R4: $N_b \geq 3$, $N_{jets} = 7$	0.1	0.7	2.8 ± 1.1	2.9 ± 0.7	3
R4: $N_b \geq 3$, $N_{jets} \geq 8$	0.1	2.1	3.3 ± 1.2	3.3 ± 0.7	3
$p_{\mathrm{T}}^{\mathrm{miss}} > 500$ GeV					
R1	0.1	0.6		434.1 ± 20.6	434
R2: $N_b = 1$, $6 \leq N_{jets} \leq 7$	0.1	0.1		159.8 ± 12.8	158
R2: $N_b = 1$, $N_{jets} \geq 8$	0.0	0.2		41.7 ± 6.4	41
R2: $N_b = 2$, $6 \leq N_{jets} \leq 7$	0.1	0.2		80.5 ± 8.8	80
R2: $N_b = 2$, $N_{jets} \geq 8$	0.1	0.3		32.0 ± 5.5	34
R2: $N_b \geq 3$, $6 \leq N_{jets} \leq 7$	0.1	0.2		19.8 ± 4.5	20
R2: $N_b \geq 3$, $N_{jets} \geq 8$	0.1	0.5		10.1 ± 3.1	10
R3	0.6	3.2		27.9 ± 4.2	28
R4: $N_b = 1$, $6 \leq N_{jets} \leq 7$	0.6	0.5	9.4 ± 3.1	10.2 ± 1.9	12
R4: $N_b = 1$, $N_{jets} \geq 8$	0.3	0.5	2.1 ± 0.8	2.3 ± 0.6	3
R4: $N_b = 2$, $6 \leq N_{jets} \leq 7$	0.9	1.0	5.3 ± 2.0	5.5 ± 1.1	6
R4: $N_b = 2$, $N_{jets} \geq 8$	0.6	1.3	2.1 ± 0.9	2.0 ± 0.5	0
R4: $N_b \geq 3$, $6 \leq N_{jets} \leq 7$	0.8	0.9	1.2 ± 0.6	1.2 ± 0.4	1
R4: $N_b \geq 3$, $N_{jets} \geq 8$	0.8	2.3	0.8 ± 0.4	0.9 ± 0.3	1

Table 10.2: Observed and predicted event yields for the signal regions (R4) and background regions (R1–R3) in the high-M_J ABCDs. Expected yields for the two SUSY T1tttt benchmark scenarios (2100, 100) and (1900,1250) are also given. The uncertainties on the prediction account for the available statistics in the data control samples, the precision of κ from MC, but not yet the systematic uncertainties assessed from control samples in data.

$\mathcal{L} = 137$ fb^{-1}	T1tttt(2100,100)	T1tttt(1900,1250)	R1–R3 fit	R1–R4 fit	Obs.
$200 < p_{\mathrm{T}}^{\mathrm{miss}} \leq 350$ GeV					
R1	0.0	1.1	7706 ± 87.4	7705.3 ± 86.8	7706
R2: $N_b = 1$, $N_{jets} = 7$	0.0	0.1	935 ± 33.8	937.3 ± 29.7	935
R2: $N_b = 1$, $N_{jets} \geq 8$	0.1	0.3	961 ± 33.7	958.7 ± 29.7	961
R2: $N_b = 2$, $N_{jets} = 7$	0.0	0.2	600 ± 31.5	606.4 ± 23.7	600
R2: $N_b = 2$, $N_{jets} \geq 8$	0.1	0.6	832 ± 40.2	821.0 ± 27.8	832
R2: $N_b \geq 3$, $N_{jets} = 7$	0.0	0.2	168 ± 14.8	171.2 ± 12.6	168
R2: $N_b \geq 3$, $N_{jets} \geq 8$	0.1	1.1	306 ± 21.7	307.9 ± 17.0	306
R3	0.0	2.2	472 ± 21.4	472.7 ± 19.8	472
R4: $N_b = 1$, $N_{jets} = 7$	0.1	0.2	75.6 ± 11.3	81.7 ± 5.6	84
R4: $N_b = 1$, $N_{jets} \geq 8$	0.2	0.5	71.8 ± 10.3	76.3 ± 4.9	74
R4: $N_b = 2$, $N_{jets} = 7$	0.2	0.4	48.9 ± 10.3	57.6 ± 4.3	64
R4: $N_b = 2$, $N_{jets} \geq 8$	0.3	1.5	62.7 ± 13.1	70.0 ± 5.1	59
R4: $N_b \geq 3$, $N_{jets} = 7$	0.1	0.6	15.2 ± 3.9	18.8 ± 2.1	22
R4: $N_b \geq 3$, $N_{jets} \geq 8$	0.4	2.6	24.9 ± 6.1	30.1 ± 2.9	32
$350 < p_{\mathrm{T}}^{\mathrm{miss}} \leq 500$ GeV					
R1	0.0	0.9	967 ± 30.9	968.0 ± 30.6	967
R2: $N_b = 1$, $N_{jets} = 7$	0.0	0.0	78 ± 9.6	72.2 ± 8.2	78
R2: $N_b = 1$, $N_{jets} \geq 8$	0.1	0.1	95 ± 10.9	92.4 ± 9.5	95
R2: $N_b = 2$, $N_{jets} = 7$	0.1	0.0	54 ± 8.5	55.8 ± 7.3	54
R2: $N_b = 2$, $N_{jets} \geq 8$	0.1	0.2	65 ± 9.7	66.1 ± 8.1	65
R2: $N_b \geq 3$, $N_{jets} = 7$	0.0	0.1	8 ± 3.0	9.1 ± 2.9	8
R2: $N_b \geq 3$, $N_{jets} \geq 8$	0.1	0.4	16 ± 4.5	18.7 ± 4.2	16
R3	0.1	2.9	68 ± 8.2	67.0 ± 6.5	68
R4: $N_b = 1$, $N_{jets} = 7$	0.2	0.1	8.7 ± 2.6	6.8 ± 1.4	1
R4: $N_b = 1$, $N_{jets} \geq 8$	0.2	0.3	8.4 ± 2.4	7.6 ± 1.4	5
R4: $N_b = 2$, $N_{jets} = 7$	0.2	0.1	4.7 ± 1.6	5.2 ± 1.0	7
R4: $N_b = 2$, $N_{jets} \geq 8$	0.4	0.7	4.6 ± 1.5	4.9 ± 0.9	6
R4: $N_b \geq 3$, $N_{jets} = 7$	0.2	0.1	0.7 ± 0.4	0.9 ± 0.3	2
R4: $N_b \geq 3$, $N_{jets} \geq 8$	0.5	1.3	1.8 ± 0.8	2.3 ± 0.7	5
$p_{\mathrm{T}}^{\mathrm{miss}} > 500$ GeV					
R1	0.1	0.6	434 ± 20.8	434.1 ± 20.6	434
R2: $N_b = 1$, $6 \leq N_{jets} \leq 7$	0.1	0.0	49 ± 8.5	46.9 ± 7.0	49
R2: $N_b = 1$, $N_{jets} \geq 8$	0.2	0.1	13 ± 4.0	13.2 ± 3.7	13
R2: $N_b = 2$, $6 \leq N_{jets} \leq 7$	0.2	0.0	18 ± 5.0	18.5 ± 4.3	18
R2: $N_b = 2$, $N_{jets} \geq 8$	0.3	0.2	7 ± 2.9	7.6 ± 2.8	7
R2: $N_b \geq 3$, $6 \leq N_{jets} \leq 7$	0.2	0.0	4 ± 2.1	4.5 ± 2.1	4
R2: $N_b \geq 3$, $N_{jets} \geq 8$	0.4	0.3	5 ± 2.4	4.3 ± 2.0	5
R3	0.6	3.2	28 ± 5.3	27.9 ± 4.2	28
R4: $N_b = 1$, $6 \leq N_{jets} \leq 7$	1.0	0.1	3.7 ± 1.5	3.1 ± 0.9	1
R4: $N_b = 1$, $N_{jets} \geq 8$	1.1	0.3	0.8 ± 0.4	0.8 ± 0.3	1
R4: $N_b = 2$, $6 \leq N_{jets} \leq 7$	1.4	0.1	1.5 ± 0.7	1.5 ± 0.5	2
R4: $N_b = 2$, $N_{jets} \geq 8$	2.0	0.6	0.3 ± 0.2	0.4 ± 0.2	1
R4: $N_b \geq 3$, $6 \leq N_{jets} \leq 7$	1.1	0.1	0.4 ± 0.3	0.5 ± 0.3	1
R4: $N_b \geq 3$, $N_{jets} \geq 8$	2.4	1.0	0.9 ± 0.6	0.7 ± 0.4	0

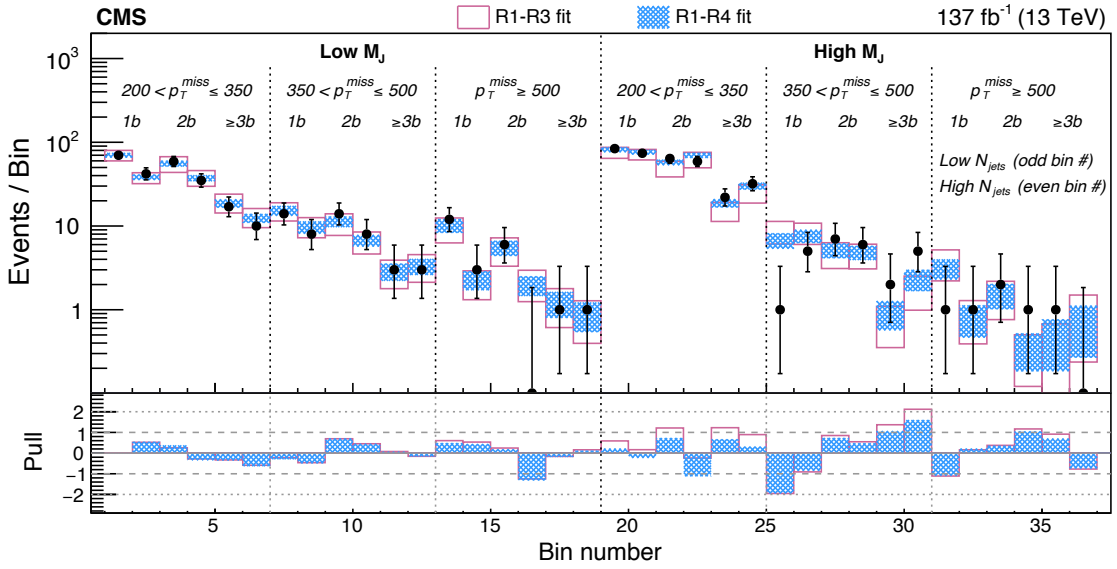


Figure 10.3: Observed and predicted event yields in each signal region. The hollow rectangles represent the prediction and uncertainty obtained from the fit with masked R4 yields, while the hashed rectangles represent the prediction obtained when all regions, R1 through R4, are included in the fit. In both cases, all statistical and systematic uncertainties are included. The bottom panel shows the pull defined as $(N_{obs} - N_{pred})/\sqrt{N_{pred} + (\sigma^{sys}_{pred})^2}$

and R3 regions. This is not a coincidence: after the exclusion of the R4 terms from the likelihood, the number of R1–R3 fit parameters is equal to the number of observables, and the fit will always match the yields in these bins exactly. Thus, the R4 yields predicted by the R1–R3 fit are simply those predicted by the standard ABCD method described in Section 8.2. Additionally, including the r parameter, which was fixed to zero for the background-only fit, would leave the R1–R3 fit under-constrained (more fit parameters than observables). Therefore, the R1–R4 fit will be used for testing signal hypotheses.

A graphical comparison between the observed data and the predictions of the two fitting methods in the search regions is shown in Figure 10.3, including the pull for each signal bin. Overall, the observed yields are consistent with background predictions. This consistency will be interpreted in the context of our signal models in the coming sections.

10.2 Methods of Interpretation

Given the absence of any significant excesses, the observed data will be interpreted as exclusion limits on the production cross section for the T1tttt and T5tttt models. Thus far, this procedure has only been referred to as "interpretation," so before we go through the results we will discuss what this procedure entails. This discussion in intended to make our interpretation a bit more, well, interpretable.

Interpretation, or hypothesis testing, refers to the process of using statistical methods to determine how consistent or inconsistent data are with a particular hypothesis. Depending on the data, results can be reported in several ways. If there is an excess in the data which causes the background-only hypothesis to be rejected, then the significance of this excess is reported, with or without a particular signal hypothesis. If the data are consistent with the background-only hypothesis, then an upper limit on the production cross section for the signal model can be reported. In searches for particles with unknown mass, upper limits are calculated for each mass value being considered. In these cases, the range of mass values excluded by the search may be reported in together with the upper limits on the cross section. As this is the type of search being conducted here, we will explore the machinery used for placing these upper limits, and converting them into mass constraints.

To calculate the upper exclusion limits, we use the test statistic q_r, defined as

$$q_r = -2 \ln \left[\frac{L(r, \boldsymbol{\theta}_r)}{L(\hat{r}, \hat{\boldsymbol{\theta}})} \right], \quad 0 \leq \hat{r} \leq r. \tag{10.1}$$

Here, r is the signal strength parameter, $\hat{r}$ is the best-fit signal strength, $\hat{\boldsymbol{\theta}}$ are the best-fit parameter values, and $\boldsymbol{\theta}_r$ are the parameters that maximize the likelihood for the given value of r. The argument of the natural log in q_r known as a *profile likelihood ratio*, as it is the likelihood maximized (profiled) for a given value of r compared with the likelihood maximized across all values of r. Additionally, the value of $\hat{r}$ is constrained to

be between zero and r. The lower bound on $\hat{r}$ ensures that signal yields are not negative, while the upper bound prevents upwards fluctuations in data from being considered evidence against the signal hypothesis.

In addition to being a function of r, q_r is also a function of the observed data, albeit implicitly. As the measurement of of each bin yield is a fundamentally random process—that is, it is equivalent to sampling a Poisson distribution—the value of q_r is a random variable as well. Depending on whether the data contain signal events or not, q_r will follow a different distribution. These distributions will be referred to as $f(q_r|B)$ and $f(q_r|S+B)$ for the background-only and signal+background scenarios, respectively. Thus, the question becomes: is the value of q_r observed in the data—which we will denote as q_r^{Obs}—more consistent with the $f(q_r|B)$ or $f(q_r|S+B)$?

To answer this question, we need to determine the distribution of q_r in each scenario, which can be done in several ways. One way would be to use the observed data to generate additional pseudodata (also known as toys). These toys would be generated with the following procedure:

1. Using the observed data and a particular value of the signal strength parameter r_{toy}, find the maximum likelihood values for all other parameters.$r_{\text{toy}} = 0$ corresponds to the background-only scenario, while $r_{\text{toy}} = 1$ would include signal with its nominal cross section.

2. From the determined parameter values, calculate Poisson mean parameters, given by Eq. (8.3), for each analysis bins.

3. Sample Poisson distributions with the calculated means to obtain a yield for each bin. This set of generated yields is the toy dataset.

4. Compute the quantity of interest for the toy dataset. In this case, this would be

the test statistic q_r. Note the r here is not necessarily equal to the r_{toy} value used to generate the toy dataset.

5. Repeat this process as many times as needed to determine the distribution of the test statistic for the given value of r_{toy}.

Though generating toys is robust, it is also computationally expensive, as you need to generate a large number of toys for a range of r values. Furthermore, this would need to be done for every mass point we wish to test, increasing the computational cost 1000 fold.

An alternative approach is to approximate the distributions of $f(q_r|B)$ and $f(q_r|S+B)$ analytically. For a sufficiently large sample size, this can be done with the asymptotic approximation [51]:

$$q_r \approx \frac{(r-\hat{r})^2}{\sigma^2} \qquad 0 \leq \hat{r} \leq r, \tag{10.2}$$

where $\hat{r}$ follows a Gaussian distribution with mean r_{data}—the true signal strength present in the data—and standard deviation σ. Using this approximation, and skipping a lot of math, the distribution of q_r is

$$f(q_r|r_{\text{data}}) = \Phi\left(\frac{r_{\text{data}}-r}{\sigma}\right)\delta(q_r) + \begin{cases} \frac{1}{\sqrt{8\pi}}\frac{1}{\sqrt{q_r}}\exp\left[-\frac{1}{2}\left(\sqrt{q_r}-\frac{r-r_{\text{data}}}{\sigma}\right)^2\right] & 0 < q_r \leq r^2/\sigma^2 \\ \frac{1}{\sqrt{8\pi}}\frac{1}{(r/\sigma)}\exp\left[-\frac{1}{2}\frac{(q_r-(r^2-2rr_{\text{data}})/\sigma^2)^2}{(2r/\sigma)^2}\right] & q_r > r^2/\sigma^2 \end{cases} \tag{10.3}$$

where Φ is the cumulative distribution function (CDF) of the standard normal distribution. This expression gives the distribution of q_r assuming it is being measured in a dataset with a true signal strength of r_{data}. Therefore, we get $f(q_r|B)$ by plugging in $r_{\text{data}} = 0$ and $f(q_r|S+B)$ by plugging in $r_{\text{data}} = r$. The asymptotic approximation allows us to quickly calculate the relevant distributions of q_r without having to generate a large number of toy datasets. Furthermore, this approximation performs well even in cases where the sample size is small.

Now that we have the distribution of q_r for the background-only and signal+background scenarios, we determine which values of r can be excluded based on our observation of q_r^{Obs} by calculating the quantity CL_s [52], defined as:

$$CL_s = \frac{CL_{s+b}}{CL_b} = \frac{\int_{q_r^{\text{Obs}}}^{\infty} f(q_r|S+B)\,dq_r}{\int_{q_r^{\text{Obs}}}^{\infty} f(q_r|B)\,dq_r}. \tag{10.4}$$

This quantity compares the probability of measuring a $q_r > q_r^{\text{Obs}}$ given the data had true signal strength of r with the probability of measuring a $q_r > q_r^{\text{Obs}}$ given the data had no signal in it. Thus, lower values of CL_s correspond to cases where q_r^{Obs} is much more consistent with the background-only scenario. High CL_s values, on the other hand, don't necessarily correspond to a q_r^{Obs} that supports the signal hypothesis. For example, if q_r provides no discrimination between signal and background ($f(q_r|S+B) \approx f(q_r|B)$), $CL_s = 1$ for all values of r. In other words, CL_s can only tell us if there is significant evidence against a signal model, not if there is evidence supporting it. So how low is low enough? If $CL_s \leq \alpha$, then the signal strength of r is said to be excluded at the confidence level (CL) of $1-\alpha$. Typically, 95% confidence levels are reported, corresponding to a $CL_s \leq 0.05$. Therefore, the maximum signal strength (r_{max}) for which $CL_s \geq 0.05$ is the lower bound of the excluded signal strength values. As results are reported as upper limits on the cross section, r_{max} is multiplied by the nominal production cross section of the signal model being considered to get the reported upper limit.

In addition to being carried out for the observed data, the above procedure is also carried out for a representative data set that treats the Poisson mean parameter of each bin as the observed yield for that bin. This is known as the Asimov data set, and is used for determining the median sensitivity of a search or measurement.

10.3 T1tttt and T5tttt Limits

Given the absence of any significant excess in the data, upper limits are placed on the production cross section of our two signal models. Using the Higgs combine tool [53], which implements the CL_s method described in the previous section, we place upper limits on the production cross section at a 95% confidence level. Figure 10.4 shows these limits for the T1tttt and T5tttt models in the $m_{\tilde{g}} - m_{\tilde{\chi}^0_1}$ plane for 137 fb^{-1} of data. Limits were calculated assuming the branching fraction for the four-top final state is 100%.

For the T1tttt model, gluinos with masses of up to 2150 GeV are excluded for $\tilde{\chi}^0_1$ masses below 700 GeV. The highest limit on the $\tilde{\chi}^0_1$ mass is 1250 GeV, attained for $m_{\tilde{g}}$ masses between 1700 and 1900 GeV. The observed limits for T1tttt are within 1σ uncertainty of the expected limits (computed from the Asimov data set) across the full scan range.

The T5tttt model allows us to extend the interpretation to scenarios where the top squark is lighter than the gluino. Rather than including the top squark mass as an additional model parameter, we fix the mass to the smallest value that would allow the top squark to decay into a top quark and $\tilde{\chi}^0_1$ ($m_{\tilde{t}} \approx m_t + m_{\tilde{\chi}^0_1}$). As the top quark and $\tilde{\chi}^0_1$ are produced approximately at rest in the top squark frame in these models, they correspond to the lowest signal efficiency for given values of $m_{\tilde{g}}$ and $m_{\tilde{\chi}^0_1}$. As a consequence, the excluded cross section values of $m_{\tilde{g}}$ and $m_{\tilde{\chi}^0_1}$ where $m_{\tilde{g}} > m_{\tilde{t}} \geq m_t + m_{\tilde{\chi}^0_1}$ are minimized. In particular, for low values of $m_{\tilde{\chi}^0_1}$, the neutralino carries little momentum, giving a lower value of m_T, and resulting in a significantly lower sensitivity to T5tttt than T1tttt in this mass region. In fact, the sensitivity to low-$m_{\tilde{\chi}^0_1}$ signal models is dominated by events with at least two leptonically decaying W bosons that produce additional p_T^{miss} as well as a tail in the m_T distribution. Although the dilepton veto is meant to exclude such events from the analysis, a significant number make it into the search region.

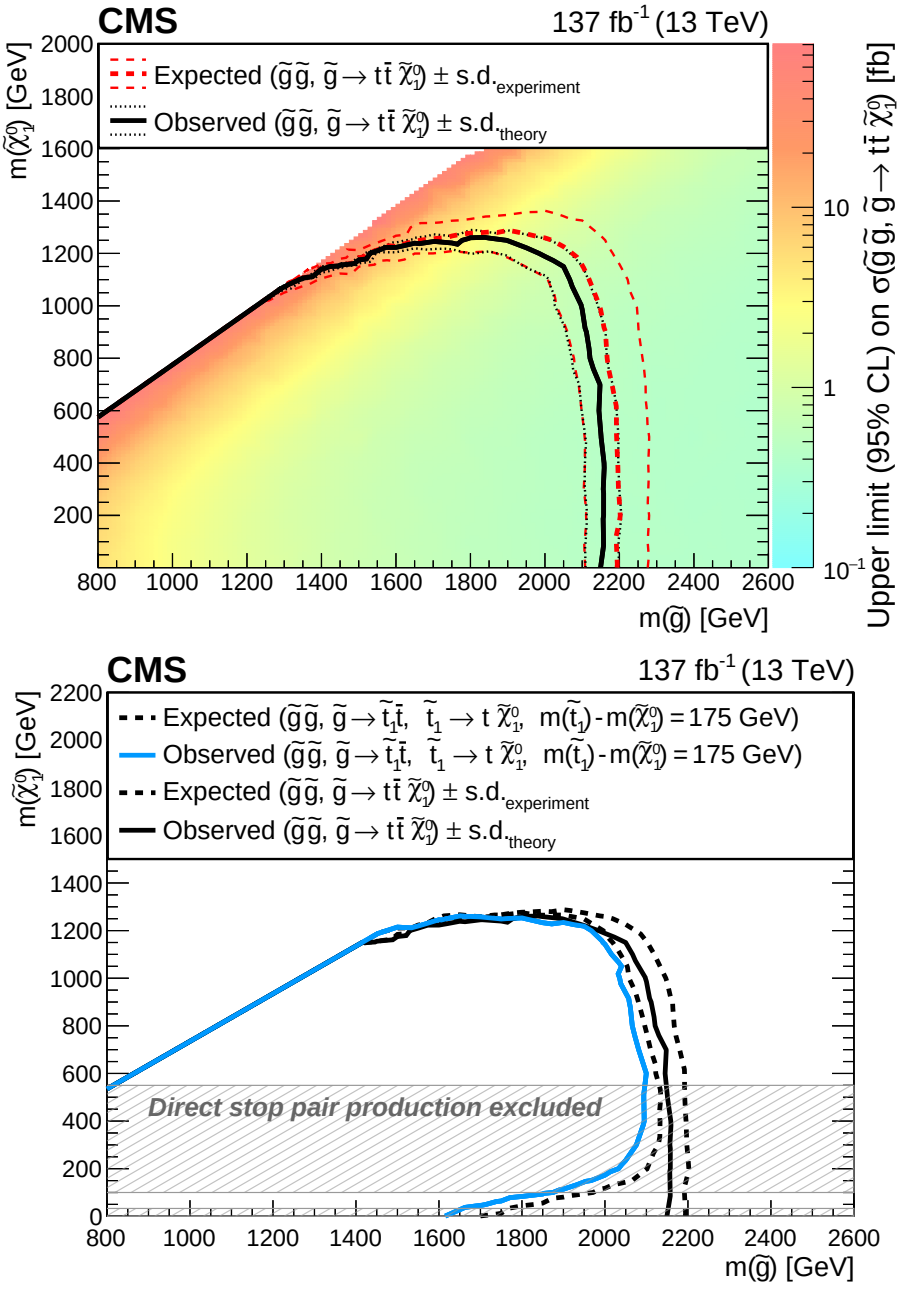


Figure 10.4: Interpretation of results in the T1tttt and T5tttt models. The colored regions show the upper limits (95% CL) on the production cross section for $pp \to \tilde{g}\tilde{g}, \tilde{g} \to t\bar{t}\tilde{\chi}_1^0$ in the $m_{\tilde{g}}$-$m_{\tilde{\chi}_1^0}$ plane. The curves show the expected and observed limits on the corresponding SUSY particle masses obtained by comparing the excluded cross section with theoretical cross sections.

For physical consistency, the signal model used in the T5tttt study should include additional contributions from top squark pair production—$\widetilde{\mathrm{t}}\bar{\widetilde{\mathrm{t}}}$, referred to as T2tt—for values of $m_{\widetilde{\chi}1}^{0}$ that have not been excluded[1] for T2tt by direct searches for $\widetilde{\mathrm{t}}\bar{\widetilde{\mathrm{t}}}$ production [54, 55]. However, we have verified that the addition of the T2tt process (where it is not excluded) does not significantly impact our sensitivity. Therefore, the exclusion curve shown in Figure 10.4 only includes contributions from the T5tttt process.

While the limits presented here are very stringent, there are still ways for natural SUSY models to evade them. In particular, we assume the decay $\widetilde{\mathrm{g}} \to \mathrm{t}\bar{\mathrm{t}}\widetilde{\chi}_1^0$ occurs 100% of the time. Any additional decay modes would reduce the expected signal events, and weaken the reported limit. If interpreted as upper limits on the production cross section times the $\widetilde{\mathrm{g}} \to \mathrm{t}\bar{\mathrm{t}}\widetilde{\chi}_1^0$ branching fraction, however, the limits reported here remain valid in the presence of additional gluino decay channels.

[1] Represented by the hashed out region in the right plot of Figure 10.4, T2tt is excluded for $m_{\widetilde{\chi}1}^{0} < 33\,\mathrm{GeV}$ and $100 < m_{\widetilde{\chi}1}^{0} < 550\,\mathrm{GeV}$ when $m_{\widetilde{\mathrm{t}}} - m_{\widetilde{\chi}1}^{0} = 175\,\mathrm{GeV}$ The $33 < m_{\widetilde{\chi}1}^{0} < 100\,\mathrm{GeV}$ region is not excluded due to the inability to assess the rapidly changing acceptance with limited simulation data.

Chapter 11

Summary and Conclusions

Results have been presented for a search for supersymmetry in a 137 fb^{-1} dataset of $\sqrt{s} =$ 13 TeV proton-proton collisions collected at the Compact Muon Solenoid experiment. This search was conducted in the single-lepton final state, focusing on events with large missing transverse momentum met, high jet multiplicity, and one or more b tagged jets. With this selection, an ABCD method is established using two variables: M_J and m_{T}. M_J is defined as the sum of masses of large-R jets, constructed by clustering anti-k_T jets with $R = 0.4$; and m_{T} is defined as the transverse mass between the lepton momentum and $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$. The independence of these two variables allows for the extrapolation of the M_J shape in the high-m_{T} search region from a low-m_{T} control region. The primary standard model background in these regions is $\mathrm{t\bar{t}}$, with the low-m_{T} region dominated by 1ℓ $\mathrm{t\bar{t}}$ and the high-m_{T} region dominated by 2ℓ $\mathrm{t\bar{t}}$. Despite the slightly different composition of the low- and high-m_{T} regions, the high N_{jets} requirement ensures the presence of a significant amount of initial-state radiation. In this regime, the M_J shapes of these two backgrounds converge. To correct for any residual correlations between M_J and m_{T}, κ correction factors are calculated in simulation and found to be near unity. By comparing these κ values between data and simulation in control regions, we quantify the systematic

uncertainties in the background estimation.

Separate ABCD methods are conducted in three $p_{\mathrm{T}}^{\mathrm{miss}}$ bins, with each bin further divided into 12 bins of N_{jets}, N_{b}, and M_J. These 36 search regions allow the analysis to be sensitive to a wider range of signal models, whose kinematic properties vary with gluino and LSP mass.

Across the 36 search region, the observed yields are consistent with the estimated standard model background yields. The lack of significant excesses is interpreted in the framework of the T1tttt and T5tttt simplified models that describe natural supersymmetry scenarios. For the T1tttt model, in which gluinos pair production is followed by the three-body decay $\widetilde{\mathrm{g}} \to \mathrm{t\bar{t}}\widetilde{\chi}_1^0$, gluinos with masses below 2150 GeV are excluded at 95% confidence level for $m_{\widetilde{\chi}_1^0} < 700\,\mathrm{GeV}$ and the highest excluded neutralino mass is about 1250 GeV. For the T5tttt model, in which gluino pair production is followed by the two-body decays $\widetilde{\mathrm{g}} \to \widetilde{\mathrm{t}}\bar{\mathrm{t}}$ and $\widetilde{\mathrm{t}} \to \mathrm{t}\widetilde{\chi}_1^0$, the limits are similar, except at low neutralino masses, where the excluded neutralino mass is slightly lower. Due to the increase in size of the data samples and the optimization this increase allowed, these results extend previous gluino mass limits [9] by about 250 GeV. These mass limits are among the most stringent constraints on these supersymmetry models to date.

Now that we have seen the mass limits this analysis was able to place on the T1tttt and T5tttt models, let's review the limits placed by other searches for supersymmetry at CMS. Table 11.1 lists the highest limits placed on the gluino mass by published CMS analyses using the $137\,\mathrm{fb}^{-1}$ dataset of Run 2, organized by the channel in which the searches were conducted. For searches that also placed limits on the neutralino mass, the highest limit placed is reported.

Due to differences in final states and kinematic properties, mass limits placed on different simplified models can not be compared one-to-one. Focusing at the limits placed on the T1tttt model, we see a slightly higher sensitivity in the hadronic channel, with

Table 11.1: Limits placed on the gluino mass during Run 2 at CMS, organized by channel and simplified model. Where applicable, limits on the neutralino mass are also shown. The results of this analysis are shown in bold.

Channel	Model	$m_{\tilde{g}}$	$m_{\tilde{\chi}_1^0}$	Ref.
Hadronic	T1tttt	2260	1420	[56]
		2250	1250	[57]
		2170	1370	[58]
	T1bbbb	2310	1420	
		2250	1525	[57]
	T1qqqq	1970	1200	
	T1qqqq-LLChi	2460	2000	
	T5HH	2330	-	[59]
	T5ZZ	1920	-	[60]
Single lepton	**T1tttt**	**2150**	**1250**	[61]
	T5tttt	**2050**	**1250**	
Dilepton	T1qqqqL	2100	-	[62]
	T5qqqqZZ-G	1870	1500	[63]

the highest gluino mass limit at 2260 GeV [57] compared to the limit of 2150 GeV in the single lepton channel (placed by this analysis). Though there are no dilepton searches that targeted T1tttt, limits in this channel are also near 2 TeV. Overall, the limits placed across the various lepton channels provide complimentary sensitivity to a wide range of gluino pair production scenarios.

Though not searched for in this analysis, models with stop quark pair production are popular search targets. Though stop quark pair production has a lower cross section than that of gluinos for the same particle mass, the expected lightness of the stop quark in many natural SUSY models makes it an attractive search candidate. Table 11.2 summarizes the limits placed on the stop mass (and neutralino mass when applicable) by CMS analyses using the 137 fb^{-1} dataset of Run 2, organized in a similar fashion as Table 11.1.

To compare channel sensitivity for stop searches, we can compare the limits placed on the stop mass for the T2tt model, in which the produced stop quarks decay into a neutralino and top quark. As there are only two top quarks produced in this model, 61% events will have zero leptons in the final state, 34% will have one lepton, and 5% will have two leptons. As such, the hadronic and single lepton channels give the highest sensitivity to this model, with limits on the stop mass as high as 1310 GeV and 1200 GeV, respectively. Due to the low fraction of T2tt events that produce two leptons, the dilepton channel is less sensitive to this model, with an upper limit of 850 GeV placed on the stop mass. Looking beyond the T2tt model, however, we see the dilepton channel gives good sensitivity to models that the other two channels do not. Thus, searches conducted across the lepton channels provide complimentary sensitivity to a variety of stop pair production scenarios.

In summary, while SUSY searches conducted at CMS using Run 2 data have not discovered SUSY, they have placed extremely stringent limits on a variety of natural

Table 11.2: Limits placed on the stop mass during Run 2 at CMS, organized by channel and simplified model. Where applicable, limits on the neutralino mass are also shown.

Channel	Model	$m_{\tilde{t}}$	$m_{\tilde{\chi}_1^0}$	Ref.
Hadronic	T2tt	1190	600	[58]
		1310	650	[56]
		1200	580	[57]
	T2tt-LLChi	1660	1210	
Single lepton	T2tt	1200	600	[64]
	T2bbWW	1160	540	
	T2bbW	1100	500	
	RPV UDD	670	-	[65]
	Stealth SYY	870	-	
Dilepton	T2tt	850	450	[66]
	T2bW	850	420	
	T8bbllnunu	1400	900	
	T6ttHZ	900	-	[62]

SUSY models. With limits on the gluino mass surpassing 2 TeV and limits on the stop mass surpassing 1 TeV, natural SUSY is running out of places to hide. Since the cross sections for these models decrease exponentially with particle mass, however, further increases in these limits will require much more data. The increased luminosity offered by the High-Luminosity LHC will provide an estimated $3000\,\text{fb}^{-1}$ of data. Using the limits of this analysis and assuming the observed limit scales inversely with integrated luminosity, a $3000\,\text{fb}^{-1}$ dataset would allow us to exclude gluino masses up to 2.8 TeV. Here's to hoping we won't just be placing limits by then.

Bibliography

[1] M. Rieger, "Search for $t\bar{t}H$ Production in the $H \to b\bar{b}$ Decay Channel: Using Deep Learning Techniques with the CMS Experiment". PhD thesis, RWTH Aachen U., 2019. doi:10.18154/RWTH-2019-06415. [Cited on pages xi and 28]

[2] M. Papucci, J. T. Ruderman, and A. Weiler, "Natural SUSY Endures", *Journal of High Energy Physics* **2012** (2012), no. 9, 35, doi:10.1007/JHEP09(2012)035, arXiv:1110.6926. [Cited on pages xi, 40, and 41]

[3] CMS Collaboration, "LHC SUSY Cross Section Working Group". Accessed: 2021 September 8. [Cited on pages xi and 42]

[4] C. De Melis, "The CERN accelerator complex. Complexe des accélérateurs du CERN", July, 2016. General Photo. [Cited on pages xii, 46, and 47]

[5] CMS Collaboration, "CMS Luminosity – Public Results". Accessed: 2021 May 18. [Cited on pages xii and 51]

[6] CMS Collaboration, S. R. Davis, "Interactive Slice of the CMS detector", August, 2016. [Cited on pages xii and 55]

[7] CMS Collaboration, "The CMS experiment at the CERN LHC", *Journal of Instrumentation* **3** (2008), no. 08, S08004, doi:10.1088/1748-0221/3/08/S08004. [Cited on pages xii, xiii, 60, 62, 65, 66, and 69]

[8] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV", *Journal of Instrumentation* **13** (2018), no. 06, P06015, doi:10.1088/1748-0221/13/06/P06015. [Cited on pages xii and 65]

[9] CMS Collaboration, "Search for Supersymmetry in pp Collisions at $\sqrt{s} = 13$ TeV in the Single-Lepton Final State Using the Sum of Masses of Large-Radius Jets", *Phys. Rev. Lett.* **119** (2017), no. 15, 151802, doi:10.1103/PhysRevLett.119.151802, arXiv:1705.04673. [Cited on pages xiii, 75, and 170]

[10] CMS Collaboration, "CMS SUSY Results: Objects Efficiency", May, 2017. Accessed 2021 May 20. [Cited on pages xiii, 75, and 82]

[11] C. Patrignani and Particle Data Group, "Review of Particle Physics", *Chinese Physics C* **40** (2016), no. 10, 100001, doi:10.1088/1674-1137/40/10/100001. [Cited on pages xix and 27]

[12] R. Steerenberg et al., "Operation and performance of the CERN Large Hadron Collider during proton Run 2", in *10th International Particle Accelerator Conference*, p. MOPMP031. 2019. doi:10.18429/JACoW-IPAC2019-MOPMP031. [Cited on pages xix and 50]

[13] R. N. Cahn and G. Goldhaber, "The Experimental Foundations of Particle Physics". Cambridge University Press, 2 edition, 2009. doi:10.1017/CBO9780511609923. [Cited on page 12]

[14] C. D. Anderson, "The Positive Electron", *Phys. Rev.* **43** (1933) 491–494, doi:10.1103/PhysRev.43.491. [Cited on page 13]

[15] F. Reines et al., "Detection of the free anti-neutrino", *Phys. Rev.* **117** (1960) 159–173, doi:10.1103/PhysRev.117.159. [Cited on page 14]

[16] UA1 Collaboration, "Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV", *Phys. Lett. B* **122** (1983) 103–116, doi:10.1016/0370-2693(83)91177-2. [Cited on page 15]

[17] UA1 Collaboration, "Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider", *Phys. Lett. B* **126** (1983) 398–410, doi:10.1016/0370-2693(83)90188-0. [Cited on page 15]

[18] UA2 Collaboration, "Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider", *Phys. Lett. B* **122** (1983) 476–485, doi:10.1016/0370-2693(83)91605-2. [Cited on page 15]

[19] UA2 Collaboration, "Evidence for $Z^0 \to e^+e^-$ at the CERN $\bar{p}p$ Collider", *Phys. Lett. B* **129** (1983) 130–140, doi:10.1016/0370-2693(83)90744-X. [Cited on page 15]

[20] UA2 Collaboration, "Search for Top Quark Production at the CERN $\bar{p}p$ Collider", *Z. Phys. C* **46** (1990) 179, doi:10.1007/BF01555995. [Cited on page 15]

[21] CDF Collaboration, "A Lower limit on the top quark mass from events with two leptons in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. Lett.* **68** (1992) 447–451, doi:10.1103/PhysRevLett.68.447. [Cited on page 15]

[22] CDF Collaboration, "Observation of top quark production in $\bar{p}p$ collisions", *Phys. Rev. Lett.* **74** (1995) 2626–2631, doi:10.1103/PhysRevLett.74.2626, arXiv:hep-ex/9503002. [Cited on page 15]

[23] D0 Collaboration, "Observation of the top quark", *Phys. Rev. Lett.* **74** (1995) 2632–2637, doi:10.1103/PhysRevLett.74.2632, arXiv:hep-ex/9503003. [Cited on page 15]

[24] Super-Kamiokande Collaboration, "Evidence for oscillation of atmospheric neutrinos", *Phys. Rev. Lett.* **81** (1998) 1562–1567, doi:10.1103/PhysRevLett.81.1562, arXiv:hep-ex/9807003. [Cited on page 34]

[25] L. Wolfenstein, "Neutrino Oscillations in Matter", *Phys. Rev. D* **17** (1978) 2369–2374, doi:10.1103/PhysRevD.17.2369. [Cited on page 34]

[26] E. Corbelli and P. Salucci, "The Extended Rotation Curve and the Dark Matter Halo of M33", *Mon. Not. Roy. Astron. Soc.* **311** (2000) 441–447, doi:10.1046/j.1365-8711.2000.03075.x, arXiv:astro-ph/9909252. [Cited on page 34]

[27] Supernova Search Team Collaboration, "Observational evidence from supernovae for an accelerating universe and a cosmological constant", *Astron. J.* **116** (1998) 1009–1038, doi:10.1086/300499, arXiv:astro-ph/9805201. [Cited on page 34]

[28] Planck Collaboration, "Planck 2018 results. VI. Cosmological parameters", *Astron. Astrophys.* **641** (2020) A6, doi:10.1051/0004-6361/201833910, arXiv:1807.06209. [Erratum: Astron.Astrophys. 652, C4 (2021)]. [Cited on page 34]

[29] S. P. Martin, "A Supersymmetry Primer", volume 21 of *Advanced Series on Directions in High Energy Physics*, ch. 1, pp. 1–153. World Scientific, April, 2010. arXiv:hep-ph/9709356. doi:10.1142/9789814307505_0001. [Cited on page 37]

[30] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017), no. 10, P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965. [Cited on page 72]

[31] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014), no. 10, P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569. [Cited on page 72]

[32] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov, "Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC", *Journal of Physics G: Nuclear and Particle Physics* **31** (January, 2005) N9, doi:10.1088/0954-3899/31/9/N01, arXiv:physics/0306087. [Cited on page 75]

[33] K. Rehermann and B. Tweedie, "Efficient Identification of Boosted Semileptonic Top Quarks at the LHC", *JHEP* **03** (2011) 059, doi:10.1007/JHEP03(2011)059, arXiv:1007.2221. [Cited on page 77]

[34] A. Barr, C. Lester, and P. Stephens, "A variable for measuring masses at hadron colliders when missing energy is expected; m_{T2}: the truth behind the glamour", *Journal of Physics G: Nuclear and Particle Physics* **29** (2003), no. 10, 2343, doi:10.1088/0954-3899/29/10/304, arXiv:hep-ph/0304226. [Cited on page 78]

[35] C. Lester and D. Summers, "Measuring masses of semi-invisibly decaying particle pairs produced at hadron colliders", *Physics Letters B* **463** (1999), no. 1, 99–103, doi:10.1016/S0370-2693(99)00945-4, arXiv:hep-ph/9906349. [Cited on page 78]

[36] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *Journal of High Energy Physics* **2008** (2008), no. 04, 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189. [Cited on page 79]

[37] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *The European Physical Journal C* **72** (2012), no. 3, 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097. [Cited on page 79]

[38] CMS Collaboration, "Identification of b quark jets at the CMS Experiment in the LHC Run 2", Technical Report CMS-PAS-BTV-15-001, CERN, Geneva, 2016. [Cited on page 81]

[39] D. Guest et al., "Jet Flavor Classification in High-Energy Physics with Deep Neural Networks", *Phys. Rev. D* **94** (2016), no. 11, 112002, doi:10.1103/PhysRevD.94.112002, arXiv:1607.08633. [Cited on page 82]

[40] A. Hook, E. Izaguirre, M. Lisanti, and J. G. Wacker, "High multiplicity searches at the LHC using jet masses", *Physical Review D* **85** (Mar, 2012) 055029, doi:10.1103/PhysRevD.85.055029, arXiv:1202.0558. [Cited on page 83]

[41] T. Cohen, E. Izaguirre, M. Lisanti, and H. K. Lou, "Jet substructure by accident", *Journal of High Energy Physics* **2013** (2013), no. 3, 161, doi:10.1007/JHEP03(2013)161, arXiv:1212.1456. [Cited on page 83]

[42] S. El Hedri, A. Hook, M. Jankowiak, and J. G. Wacker, "Learning how to count: a high multiplicity search for the LHC", *Journal of High Energy Physics* **2013** (2013), no. 8, 136, doi:10.1007/JHEP08(2013)136, arXiv:1302.1870. [Cited on page 83]

[43] H. Qu, "A search for Lorentz-boosted Higgs bosons decaying to charm quarks in the CMS experiment using deep neural networks". PhD thesis, UC, Santa Barbara (main), 2019. [Cited on page 84]

[44] R. D. Ball et al., "Parton distributions for the LHC run II", *Journal of High Energy Physics* **2015** (2015), no. 4, 40, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849. [Cited on page 91]

[45] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *Journal of High Energy Physics* **2014** (2014), no. 7, 79, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301. [Cited on page 91]

[46] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method", *The European Physical Journal C* **71** (2011), no. 2, 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450. [Cited on page 91]

[47] T. Sjöstrand et al., "An introduction to PYTHIA 8.2", *Computer Physics Communications* **191** (2015) 159–177, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012. [Cited on page 91]

[48] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *The European Physical Journal C* **76** (2016), no. 3, 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815. [Cited on page 91]

[49] S. Agostinelli et al., "Geant4–a simulation toolkit", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003), no. 3, 250–303, doi:10.1016/S0168-9002(03)01368-8. [Cited on page 94]

[50] CMS Collaboration, "The Fast Simulation of the CMS Detector at LHC", *Journal of Physics: Conference Series* **331** (2011), no. 3, 032049, doi:10.1088/1742-6596/331/3/032049. [Cited on page 94]

[51] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *The European Physical Journal C* **71** (2011), no. 2, 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727. [Cited on page 164]

[52] A. L. Read, "Presentation of search results: the CL_s technique", *Journal of Physics G: Nuclear and Particle Physics* **28** (2002), no. 10, 2693. [Cited on page 165]

[53] The ATLAS Collaboration, The CMS Collaboration, The LHC Higgs Combination Group Collaboration, "Procedure for the LHC Higgs boson search combination in Summer 2011", technical report, CERN, Geneva, Aug, 2011. [Cited on page 166]

[54] CMS Collaboration, "Search for the pair production of light top squarks in the $e^{\pm}\mu^{\mp}$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **03** (2019) 101, doi:10.1007/JHEP03(2019)101, arXiv:1901.01288. [Cited on page 168]

[55] CMS Collaboration, "Search for supersymmetry in proton-proton collisions at 13 TeV in final states with jets and missing transverse momentum", *JHEP* **10** (2019) 244, doi:10.1007/JHEP10(2019)244, arXiv:1908.04722. [Cited on page 168]

[56] CMS Collaboration, "Search for top squark production in fully-hadronic final states in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. D* **104** (2021), no. 5, 052001, doi:10.1103/PhysRevD.104.052001, arXiv:2103.01290. [Cited on pages 171 and 173]

[57] CMS Collaboration, "Searches for physics beyond the standard model with the M_{T2} variable in hadronic final states with and without disappearing tracks in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C* **80** (2020), no. 1, 3, doi:10.1140/epjc/s10052-019-7493-x, arXiv:1909.03460. [Cited on pages 171, 172, and 173]

[58] CMS Collaboration, "Search for supersymmetry in proton-proton collisions at 13 TeV in final states with jets and missing transverse momentum", *JHEP* **10** (2019) 244, doi:10.1007/JHEP10(2019)244, arXiv:1908.04722. [Cited on pages 171 and 173]

[59] CMS Collaboration, "Search for higgsinos decaying to two Higgs bosons and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **05** (2022) 014, doi:10.1007/JHEP05(2022)014, arXiv:2201.04206. [Cited on page 171]

[60] CMS Collaboration, "Search for supersymmetry in proton-proton collisions at $\sqrt{s} = 13$ TeV in events with high-momentum Z bosons and missing transverse momentum", *JHEP* **09** (2020) 149, doi:10.1007/JHEP09(2020)149, arXiv:2008.04422. [Cited on page 171]

[61] CMS Collaboration, "Search for supersymmetry in pp collisions at $\sqrt{s} = 13$ TeV with 137 fb^{-1} in final states with a single lepton using the sum of masses of large-radius jets", *Phys. Rev. D* **101** (2020), no. 5, 052010, doi:10.1103/PhysRevD.101.052010, arXiv:1911.07558. [Cited on page 171]

[62] CMS Collaboration, "Search for physics beyond the standard model in events with jets and two same-sign or at least three charged leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Eur. Phys. J. C* **80** (2020), no. 8, 752, doi:10.1140/epjc/s10052-020-8168-3, arXiv:2001.10086. [Cited on pages 171 and 173]

[63] CMS Collaboration, "Search for supersymmetry in final states with two oppositely charged same-flavor leptons and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **04** (2021) 123, doi:10.1007/JHEP04(2021)123, arXiv:2012.08600. [Cited on page 171]

[64] CMS Collaboration, "Search for direct top squark pair production in events with one lepton, jets, and missing transverse momentum at 13 TeV with the CMS experiment", *JHEP* **05** (2020) 032, doi:10.1007/JHEP05(2020)032, arXiv:1912.08887. [Cited on page 173]

[65] CMS Collaboration, "Search for top squarks in final states with two top quarks and several light-flavor jets in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. D* **104** (2021), no. 3, 032006, doi:10.1103/PhysRevD.104.032006, arXiv:2102.06976. [Cited on page 173]

[66] CMS Collaboration, "Search for top squark pair production using dilepton final states in pp collision data collected at $\sqrt{s} = 13\,\text{TeV}$", *Eur. Phys. J. C* **81** (2021), no. 1, 3, doi:10.1140/epjc/s10052-020-08701-5, arXiv:2008.05936. [Cited on page 173]