A Search for Supersymmetric Top Quarks in CMS 8TeV Data in the b/τ + Jets + MET + Muon Final State

By

Andrew Malone Melo

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY
in
Physics

May, 2016

Nashville, Tennessee

Approved:

Paul Sheldon, Ph.D.

Will Johns, Ph.D.

Charles Maguire, Ph.D.

Thomas Kephart, Ph.D.

Alan Tackett, Ph.D.

To those before me and those who got me here.

ACKNOWLEDGMENTS

This work wouldn't have been possible without the support of my family and friends, who so patiently supported me through all this work.

I congratulate my colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, I gratefully acknowledge the computing centres and personnel of the Worldwide LHC Computing Grid for delivering so effectively the computing infrastructure essential to my analyses. Finally, I acknowledge the enduring support for the construction and operation of the LHC and the CMS detector provided by the following funding agencies: the Austrian Federal Ministry of Science, Research and Economy and the Austrian Science Fund; the Belgian Fonds de la Recherche Scientifique, and Fonds voor Wetenschappelijk Onderzoek; the Brazilian Funding Agencies (CNPq, CAPES, FAPERJ, and FAPESP); the Bulgarian Ministry of Education and Science; CERN; the Chinese Academy of Sciences, Ministry of Science and Technology, and National Natural Science Foundation of China; the Colombian Funding Agency (COLCIENCIAS); the Croatian Ministry of Science, Education and Sport, and the Croatian Science Foundation; the Research Promotion Foundation, Cyprus; the Ministry of Education and Research, Estonian Research Council via IUT23-4 and IUT23-6 and European Regional Development Fund, Estonia; the Academy of Finland, Finnish Ministry of Education and Culture, and Helsinki Institute of Physics; the Institut National de Physique Nucléaire et de Physique des Particules / CNRS, and Commissariat à l'Énergie Atomique et aux Énergies Alternatives / CEA, France; the Bundesministerium für Bildung und Forschung, Deutsche Forschungsgemeinschaft, and Helmholtz-Gemeinschaft Deutscher Forschungszentren, Germany; the General Secretariat for Research and Technology, Greece; the National Scientific Research Foundation, and National Innovation Office, Hungary; the Department of Atomic Energy and the Department of Science and Technology, India; the Institute for Studies in Theoretical Physics and Mathematics, Iran; the Science Foundation, Ireland; the Istituto Nazionale di Fisica Nucleare, Italy; the Ministry of Science, ICT and Future Planning, and National Research Foundation (NRF), Republic of Korea; the Lithuanian Academy of Sciences; the Ministry of Education, and University of Malaya (Malaysia); the Mexican Funding Agencies (CINVESTAV, CONACYT, SEP, and UASLP-FAI); the Ministry of Business, Innovation and Employment, New Zealand; the Pakistan Atomic Energy Commission; the Ministry of Science and Higher Education and the National Science Centre, Poland; the Fundação para a Ciência e a Tecnologia, Portugal; JINR, Dubna; the Ministry of Education and Science of the Russian Federation, the Federal Agency of Atomic Energy of the Russian Federation, Russian Academy of Sciences, and the Russian Foundation for Basic Research; the Ministry of Education, Science and Technological Development of Serbia; the Secretaría de Estado de Investigación, Desarrollo e Innovación and Programa Consolider-Ingenio 2010, Spain; the Swiss Funding Agencies (ETH Board, ETH Zurich, PSI, SNF, UniZH, Canton Zurich, and SER); the Ministry of Science and Technology, Taipei; the Thailand Center of Excellence in Physics, the Institute for the Promotion of Teaching Science and Technology of Thailand, Special Task Force for Activating Research and the National Science and Technology Development Agency of Thailand; the Scientific and Technical Research Council of Turkey, and Turkish Atomic Energy Authority; the National Academy of Sciences of Ukraine, and State Fund for Fundamental Researches, Ukraine; the Science and Technology Facilities Council, UK; the US Department of Energy, and the US National Science Foundation. Individuals have received support from the Marie-Curie programme and the European Research Council and EPLANET (European Union); the Leventis Foundation; the A. P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the Council of Science and Industrial Research, India; the HOMING PLUS programme of the Foundation for Polish Science, cofinanced from European Union, Regional Development Fund; the OPUS programme of the National Science Center (Poland); the Compagnia di San Paolo (Torino); the Consorzio per la Fisica (Trieste); MIUR project 20108T4XTM (Italy); the Thalis and Aristeia programmes cofinanced by EU-ESF and the Greek NSRF; the National Priorities Research Program by Qatar National Research Fund; the Rachadapisek Sompot Fund for Postdoctoral Fellowship, Chulalongkorn University (Thailand); and the Welch Foundation, contract C-1845.

TABLE OF CONTENTS

Page

Chapter

Appendix

LIST OF TABLES

LIST OF FIGURES

Figure Page

Chapter 1

The Standard Model and the Fine Tuning Problem

The Standard Model (SM)[1, 2, 3] as currently formulated describes (in some cases) with incredible accuracy the properties and behavior of the building blocks of matter. Despite the precision with which the SM describes parts of the known universe, there is experimental evidence suggesting the SM is incomplete. For example, the SM as currently constructed does not provide a viable dark matter candidate which is compatible with measurements from astronomy/cosmology.

The SM states that matter is made up of 12 fermions. The distinguishing characteristic of the SM fermions is their $\frac{1}{2}$ spin. Due to their intrinsic half-integer spin, fermions obey Fermi-Dirac Spin statistics. As a consequence, fermions obey the Pauli Exclusion Principle, restricting their ability to share coordinates in space-time.

The twelve fermions can be divided first into two groups of six quarks and six leptons. The distinguishing characteristic of the quarks is that they have color charge, which means they participate in the strong force. The six leptons are colorless. As a consequence, they don't interact with the strong force. The quarks and three of the six leptons have electric charge and interact with both the electromagnetic force and the weak force. The remaining three leptons, known as neutrinos, have neither color charge nor electric charge. They only interact with the weak force [1], making them incredibly difficult to detect.

These twelve fermions interact by exchanging information mediated by four types of spin-1 bosons. Bosons all have integer spin and therefore obey Einstein-Bose statistics. At the quantum-mechanical level, the properties of the exchange of information via bosons explains the fundamental forces. The spin-1 bosons responsible for force mediation are the photon, the Z boson, the $W^{\pm}$ bosons and the gluon. The gluons mediate the strong force, while the remaining three spin-1 bosons are responsible for the electroweak force. In addition to the four spin-1 bosons, there is an additional spin-0 boson, the Higgs boson[4]. The Higgs boson is predicted by the addition of the Higgs mechanism to the SM. The Higgs mechanism gives rise to the masses of the W and Z bosons and is described in more detail later.

The twelve fermions and five bosons described above form the basis of the SM. The SM describes the universe in terms of the gauge groups $\mathrm{SU}(3)_\mathrm{C} \times \mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_\mathrm{Y}$, which dictate the intrinsic symmetries of the model such that any physical state must be invariant under a space-time dependent phase transformation. Within this formulation, the gauge group $\mathrm{SU}(3)_\mathrm{C}$ represents the interaction of colored particles via the strong force. The remainder of the SM gauge groups, $\mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_\mathrm{Y}$, represents the electroweak force.

As mentioned previously, there are unresolved inconsistencies between SM predictions and experimental observations at the microscopic scale (where the gravitational force has

[1] An astute reader will note neutrinos also interact gravitationally. The SM does not explain gravity and the effects of gravity at the microscopic scale are small enough to be negligible. For these reasons, this paper neglects gravity.

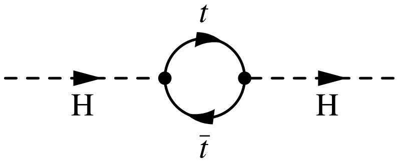


Figure 1.1: Example Higgs self-energy correction due to t quarks

negligible effect). These inconsistencies have led to a number of proposed extensions to the SM. Supersymmetry (SUSY) is one of these extensions and the subject of this analysis

A particularly illustrative issue with the SM that SUSY attempts to address is known as the "Fine-Tuning Problem." This is closely related to another problem named the "Hierarchy Problem." In the sections that follow, the fine-tuning problem is outlined in the context of the Higgs boson. To motivate a possible solution, this paper first describes a similar but analogous example. After examining this example and its solution, we will then introduce SUSY and motivate why it could fix the fine-tuning problem in the context of the Higgs boson.

1.1 The Mass of the Higgs Boson, Attempt 1

The observed mass of a particle is a combination of the true or 'bare' mass and additional contributions from self-energy effects. Each particle creates a field around itself based on its charges; this field then interacts back with the particle that created it. This self-interaction between the particle and the field it generates has the effect of changing the effective mass of the particle, which is what we observe. Since the observed mass is a combination of the bare mass and the self-energy from its fields, the bare mass can't be directly measured. Consider first the bare mass of the newly-discovered Higgs boson. The observed mass of the Higgs (m_H) is known via experimental observation to be $m_H \approx 125\,\text{GeV}$. This observed mass is a combination of the bare Higgs mass (m_0) and contributions from self-energy effects. These self-energy (E_{self}) effects can be described via Quantum Field Theory (QFT). Without loss of generality, we consider only the dominant term, which is the contribution from loops of virtual t quarks (δm_t). An example of this contribution is shown in Figure 1.1. The sum of these contributions results in the following expressions for the value of m_H

$$m_H^2 = m_0^2 + E_{self}^2 \tag{1.1}$$

$$m_H^2 \approx m_0^2 + \delta m_t^2 \tag{1.2}$$

The value of m_H is known experimentally, δm_t is known via QFT and m_0 is an unknown parameter. To find m_0, we compute δm_t via QFT and solve Equation 1.2. The contribution from loops of t quarks depends on the cutoff scale Λ and the quark mass m_t. Λ is a theoretical tool to prevent our model from blowing up when modeling very large momenta (or, equivalently, very small lengths). Since it is assumed this theory breaks down at the Planck scale, we set $\Lambda \approx 10^{19}\,\text{GeV}$. Finally, $g_t \propto h_t$ and $h_t \approx 1$ is the t quark Yukawa coupling constant.

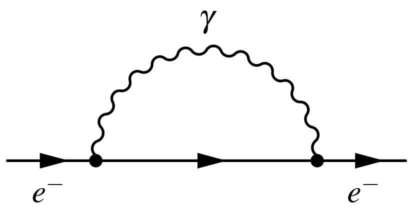


Figure 1.2: Example electron self-energy correction due to the Coloumbic field

$$\delta m_t^2 \approx +g_t^2 \int_{\Lambda_0}^{\Lambda} \frac{d^4k}{k^2} \approx +g_t^2(\Lambda^2 + m_t^2) \tag{1.3}$$

With these values, the contribution from δm_t^2 and ends up being around $10^{36}\,\mathrm{GeV}^2$. Substituting known values, Equation 1.2 becomes:

$$10^4\,\mathrm{GeV}^2 = m_0^2 + 10^{36}\,\mathrm{GeV}^2 \tag{1.4}$$

To produce the observed 125 GeV Higgs mass m_H, m_0^2 is approximately $-10^{36}\,\mathrm{GeV}^2$. Neglecting the fact that m_0^2 is negative (requiring m_0 to be imaginary), this heretofore unknown constant has to be tuned to an incredible precision. Though there is nothing physically restricting such a precise value of m_0^2, this seems unnatural. This unnaturalness is known as the 'Fine-Tuning Problem'.

The electron mass also suffered from fine-tuning, the solution to that problem instructs our efforts with the Higgs mass. The next section describes this case further.

1.2 The Mass of the Electron

The electron mass suffers from a similar fine-tuning problem. Like the Higgs, the observed mass (m_e) of the electron consists of a contribution from both the bare mass (m_0) and its self-energy (E_{self}). Similar to the Higgs case, we consider only the dominant term of the self-energy: the energy resulting from the Coloumbic field produced by the electron's electrical charge. Following the same prescription as Equation 1.2, the relationship which determines the electron mass is:

$$m_e^2 \approx m_0^2 + \delta m_{Coloumb}^2 \tag{1.5}$$

This Coloumbic field consists entirely of the emission and reabsorption of virtual photons, such as Figure 1.2. This Coloumbic field is analogous to a Higgs field generated only from the emission and reabsorption of virtual t quarks.

Like before, to find the bare mass of the electron requires the observed mass and the contribution from self-energy corrections. The observed electron mass is $m_e^2 = 10^{-1}\,\mathrm{GeV}^2$ while the self-energy correction is $\delta m_{Coloumb}^2 = 10^{10}\,\mathrm{GeV}^2$. This yields a bare mass of $m_0^2 = -10^{10}\,\mathrm{GeV}^2$. This m_0^2 is also unnatural, requiring a value 11 orders of magnitude larger than m_e^2, tuned to a precision of 10^{11}.

In both cases, there appears to be a missing contribution to the self-energy correction which drives its value to an unnaturally large range. In a more "natural" scenario, the missing contribution(s) would be roughly the same magnitude as the existing correction(s) but

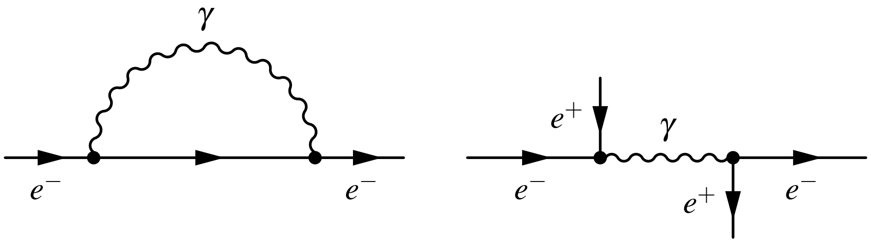


Figure 1.3: Selected electron self-energy correction terms due to the augmented Coloumbic field. The e^+ in the second diagram are produced elsewhere via quantum vacuum fluctuations.

with the opposite sign. The dominant term of our corrections results from loops of a virtual boson. One solution supplements the dominant term by including loops of a new, distinct, possibly fermionic, particle. If an appropriate particle could be found, its contribution(s) might nearly cancel the existing correction. With the dominant term suppressed, the bare mass could then settle to a more natural value.

For electrons, there is an elegant solution. First theorized as a consequence of the Dirac equation, the positron provides the necessary cancellations to $\delta m_{Coloumb}$. The positron is the anti-partner of the electron; it shares all of its properties except the sign of its electric charge is flipped. The first Coloumbic field consisted only of electrons emitting and reabsorbing virtual photons. Adding positrons and the quantum mechanical effect of vacuum fluctuations gives more possibilities. A nearby vacuum fluctuation could produce a virtual electron-positron pair and the positron could interact with our real electron to annihilate and produce a photon. A selection of diagrams for this modified Coloumbic field can be found in Figure 1.3.

Adding the effects of virtual positrons to the Coloumbic field has the desired effect. These positrons interact with electrons and photons which contribute additional terms to $\delta m_{Coloumb}$. These new terms have the effect of canceling the contribution from our old terms. These additional terms cause the fine-tuning problem to vanish, leaving a bare mass consistent with the observed mass.

In effect, the simple existence of a heretofore unknown particle influences the behavior of previously known particles.

1.3 The Mass of the Higgs Boson - A Motivation for SUSY

The treatment of the electron case is instructive to the Higgs case. The bare mass of the Higgs boson is dominated by the seemingly unnatural δm_t. If another particle existed whose contributions cancelled δm_t, the self-energy and bare mass terms of the Higgs mass equation would settle towards more natural values.

At this point, the SM falls short. For the electron mass, the chiral degree of freedom (e.g. left-chiral, right-chiral) produced the positron. Due to the symmetry between the electron and positron, their combined contributions to the electron self-energy nearly exactly cancel out. Within the SM, there are no degrees of freedom which produce particles with the necessary properties to correct the Higgs self-energy. This motivates extensions to the SM like SUSY.

In general terms, SUSY proposes an additional degree of freedom in the form of a new

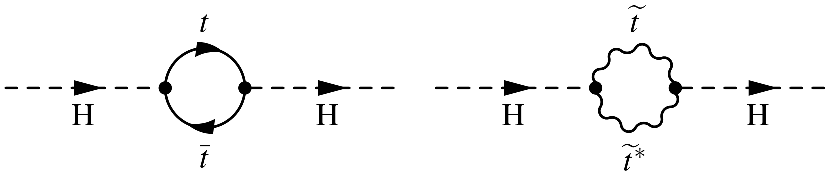


Figure 1.4: Augmented Higgs self-energy loops due to t quarks and hypothesized $\widetilde{t}$ squarks. The $\widetilde{t}$ squarks are bosonic superpartners of t quarks.

symmetry. SUSY isn't a single theory. It could be more accurately described as a collection of theories that all share this common new symmetry, but differ on their mathematical formulations and predicted effects. Being the latest proposed symmetry, this new symmetry is known as a **super**symmetry and is the source of the name SUSY[2]. These supersymmetries produce new **super**partners of the already-known SM particles. High-energy physicists search for SUSY by either detecting superpartners directly or by indirectly searching for the effects of superpartners which are not included in the SM.

A difficulty with searching for SUSY is the sheer number of possible forms it could take. Within the SUSY umbrella are several different mathematical formulations, each with different predictions and signatures. Worse, even if it were known that a specific SUSY theory is correct, SUSY theories typically contain many free parameters, making it difficult to obtain enough independent experimental measurements to discriminate amongst various scenarios. In practice, searching for new SUSY hypothetical particles can be simplified with the knowledge of other similar new particles. However, to date none of these supposed superpartners have been detected[3].

SUSY describes a symmetry between fermions and bosons. If such a symmetry exists, each SM spin-1 boson would gain a spin-$\frac{1}{2}$ fermionic superpartner and each SM spin-$\frac{1}{2}$ fermion would gain a spin-1 bosonic superpartner. A key difference between fermions and bosons is that the loop self-energy corrections to the Higgs mass from fermions and bosons contribute with opposite signs.

Revisiting the Higgs mass fine-tuning problem, adding superpartners to the model adds a new set of self-energy correction terms. These new self-energy correction terms would resemble Figure 1.4. Since superpartners behave like their SM counterparts with their boson/fermion identity flipped, their contributions to the self-energy term of the Higgs mass will resemble the positron's contribution to the electron self energy.

Returning to Equation 1.2, we originally calculated the Higgs mass using the bare mass and contribution to the self-energy from t quark loops. Adding the correction from the proposed t superpartner (known as the top squark, 'stop', or $\widetilde{t}$) leads to:

$$m_H^2 = m_0^2 + \delta m_t^2 + \delta m_{\widetilde{t}}^2 \tag{1.6}$$

[2] See also: wordplay.

[3] Unless the Higgs boson detected in 2012 is itself supersymmetric. At this time, it is unclear whether or not that is the case. For simplicity, this paper presumes the Higgs boson discovered in 2012 is the SM Higgs and not a SUSY Higgs.

where δm_t^2 and $\delta m_{\tilde{t}}^2$ are

$$\delta m_t^2 \approx \boxed{+} g_t^2 \int_{\Lambda_0}^{\Lambda} \frac{d^4k}{k^2} \approx \boxed{+} g_t^2 (\Lambda^2 + m_t^2) \tag{1.7}$$

$$\delta m_{\tilde{t}}^2 \approx \boxed{-} g_{\tilde{t}}^2 \int_{\Lambda_0}^{\Lambda} \frac{d^4k}{k^2} \approx \boxed{-} g_{\tilde{t}}^2 (\Lambda^2 + m_{\tilde{t}}^2) \tag{1.8}$$

and $m_{\tilde{t}}$ is the $\tilde{t}$ mass, $g_{\tilde{t}} \propto h_{\tilde{t}}$, and $h_{\tilde{t}} \approx 1$ is the $\tilde{t}$ squark Yukawa coupling constant.

The difference in sign between $\delta m_{\tilde{t}}^2$ and δm_t^2 is important. The sign change results from the top being a fermion and top squark being a boson. The sum of these self-energy contributions yields their combined cancellation

$$\delta m_t^2 + \delta m_{\tilde{t}}^2 \approx \Lambda^2 (g_t^2 - g_{\tilde{t}}^2) + g_t^2 m_t^2 - g_{\tilde{t}}^2 m_{\tilde{t}}^2 \tag{1.9}$$

If the t and $\tilde{t}$ have similar properties ($g_t^2 \sim g_{\tilde{t}}^2$ and $m_t^2 \sim m_{\tilde{t}}^2$), their combined contribution $\delta m_t^2 + \delta m_{\tilde{t}}^2$ will nearly exactly cancel out. Many SUSY theories predict $g_t^2 \sim g_{\tilde{t}}^2$, while m_t^2 is very loosely bounded. If, for instance, this cancellation is nearly complete,

$$m_H^2 \sim m_0^2 \tag{1.10}$$

demonstrating the naturalness of our new bare Higgs mass. This naturalness argument is based on the assumption $m_t^2 \sim m_{\tilde{t}}^2$, which motivates a 'light' $\tilde{t}$ scenario. This scenario is attractive since lighter superpartners are predicted to have larger cross-sections than superpartners with greater mass. Conversely, if $m_t^2 \ll m_{\tilde{t}}^2$, the cancellation of $\delta m_t^2 + \delta m_{\tilde{t}}^2$ remains incomplete and m_0^2 remains somewhat unnnatural. This residual unnaturalness is known as the little fine-tuning problem, and is the subject of intense theoretical research. Due to its relative simplicity, this analysis focuses on the light $\tilde{t}$ scenario.

For simplicity's sake, this example only considered the corrections due to t quarks. Extending this example to the complete self-energy correction shows similar cancellations between SUSY superpartners and their SM counterparts. These additional terms to the Higgs field cause the total self-energy correction to nearly cancel itself out, leaving the bare mass at a more natural value. This elegant solution to the Higgs mass fine-tuning problem is one of the motivations behind light top-like partners in Beyond the Standard Model (BSM) scenarios.

Chapter 2

Searching for SUSY

Before delving into the details of this analysis, two points need to be made with regards to SUSY:

1. SUSY predicts the existence of many superpartners.
2. No superpartners have been observed.[1].

The fact that no superpartners have been observed after several decades of searches implies that if SUSY were true, these superpartners must be difficult to find. They could be rarely produced at previously studied energies, they could be difficult to detect within the background of SM processes or, more likely, a combination of the two. Producing more superpartners is infeasible, the LHC is built. Increasing the production rate through an increase of the energy or luminosity of the accelerator will have to wait until either the LHC is upgraded or a new accelerator is built. What remains is to increase the signal detection efficiency and to understand and reject the backgrounds so statistically significant signals can be observed.

A common problem with all hadron colliders is the amount and complexity of backgrounds they produce. At a lepton-lepton collider, the energy of the colliding particles can be known reasonably well and the initial hard scattering processes are colorless, leading to less complex events. Compared with lepton colliders, hadron colliders like the LHC have three main problems: the initial state of the collisions aren't well-defined, many gluons are produced via initial and final state radiation, and the collisions themselves are a QCD process.

Though the energy of the protons are known, the collisions occur between the quarks and gluons (collectively referred to as partons) within the protons. The energy of the colliding partons is known only probabilistically – an 8 TeV proton could contribute a 1 TeV quark to the collision, or it could contribute a 1 MeV quark. In practice, the true rate of a process involves the probability of the process itself convoluted with the probability distribution function (PDF) of the parton momenta. Compared with lepton colliders, whose collision energy is very close to the nominal energy of the accelerated particles, hadron colliders have a wide distribution of collision energies for a given energy of accelerated particles.

Along with the ambiguity of the energy of the initial state, hadron collisions produce a lot of hadronic jets, which are difficult to reconstruct and filter out. Both before and after the collision, relativistic particles emit radiation. This radiation is known as Initial State Radiation (ISR) and Final State Radiation (FSR). In a lepton-lepton collider, the predominant ISR/FSR is in the form of photons emitted by the relativistically colliding leptons. The relatively simple signature of photons makes their reconstruction straightforward. During

[1] Excluding the possibility the Higgs boson in 2012 is the Higgs superpartner and not the SM Higgs itself

hadron collisions, the incoming and outgoing partons radiate gluons. These gluons then hadronize into jets which travel to the detector. Compared with photon reconstruction, jet reconstruction has large uncertainties. Worse, many interesting processes have jets in their final states. Accurately reconstructing these processes necessarily involves identifying and suppressing contributions from ISR and FSR. This is a profoundly difficult process.

Finally, after accounting for the difficult conditions leading up to the hard-scattering process, the actual collision of hadrons is a QCD process involving interacting partons. The partons interact overwhelmingly via QCD and only very rarely via the electroweak force. This creates an enormous QCD spectrum which must be understood and suppressed to study processes involving electroweak interactions. For example, at 8 TeV, the production rate of purely QCD hard interactions is three orders of magnitude larger than the W + jets production rate or six orders of magnitude larger than the Higgs production rate.

Despite these difficulties, hadron colliders are still built. The data produced with a hadron collider is more difficult to analyze, but there is a tradeoff. It is much more cost-effective to build a hadron collider at very high energies compared with a similarly equipped lepton collider. Along with the higher energies accessible to hadron colliders, they tend to be more luminous, causing more collisions per unit time than their lepton counterparts.

The Large Electron Positron (LEP) collider at CERN was the most energetic lepton collider before its decommissioning in 2001. Through a series of upgrades from 1989-2001, LEP had a maximum energy of 209 GeV. Even at its peak energy, LEP was unable to produce pairs of Higgs bosons (125 GeV) or top quarks (173 GeV), much less any heavier particles that may exist beyond the SM. By comparison, the LHC's design 14 TeV center-of-mass energy could theoretically produce pairs of 7 TeV particles, extending the reach for new physics far beyond what was previously possible. At the same time, the collisions will be far more complex than a lepton collider. To take advantage of the additional reach of the LHC, a method to accurately classify events in spite of their complexity must be used. This analysis leverages data from the 2012 run of the Compact Muon Solenoid (CMS) detector at CERN and an innovative technique known as Simultaneous Heavy Flavor and Top (SHyFT). SHyFT improves upon previous searches by using as much of the information available. Along with the precision provided by performing a simultaneous fit of all events from all processes, SHyFT's performance has produced the lowest-uncertainty measurements of the top quark cross section at both the CDF experiment at the Tevatron[5] and later with the 2011 7 TeV data from CMS[6]. The use of SHyFT in this analysis is summarized in Section 2.1. SHyFT itself is further described in detail in Chapter 7.

The signal probed in this analysis is a hard-scattering event producing a pair of stop squarks, $(\widetilde{t\,t^*})$ which decay into a pair of b-quarks, neutrinos (ν), τ-leptons, and neutralinos ($\widetilde{\chi}^0$). SUSYs free parameters allow the masses of the superpartners to also be freely chosen. The following decay chain is then simulated

$$\widetilde{t}\widetilde{t}^* \to b\bar{b} + \widetilde{\chi}^+\widetilde{\chi}^- \to b\bar{b} + \nu\bar{\nu} + \widetilde{\tau}\widetilde{\tau}^* \to b\bar{b} + \nu\bar{\nu} + \tau\bar{\tau} + \widetilde{\chi}^0\widetilde{\chi}^0 \qquad (2.1)$$

where * denotes antiparticles of SUSY particles (ex $\widetilde{\tau}^*$ is the antiparticle of $\widetilde{\tau}$). This decay is represented in Feynman diagram form in Figure 2.1

The $b\bar{b}$ in the final state hadronizes into jets whose decay products are detected reconstructed. The $\nu\bar{\nu}$ and $\widetilde{\chi}^0\widetilde{\chi}^0$ pass through the detector undetected, resulting in missing

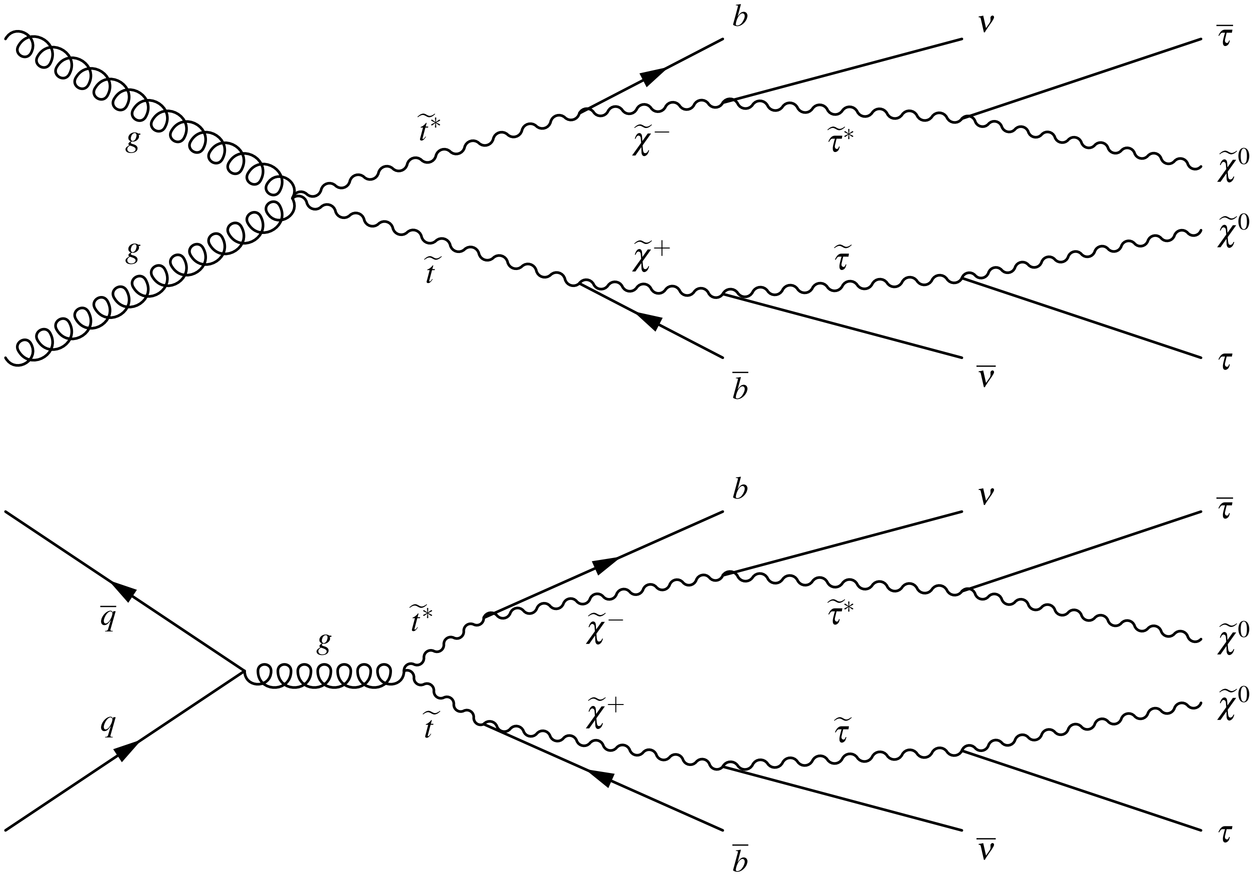


Figure 2.1: Our chosen $\widetilde{t}\,\widetilde{t}^*$ decays. Events with exactly one muon and no electrons are selected by this analysis. The τ-leptons are metastable during the hard interaction and can produce these additional leptons when they decay later.

transverse energy ($\not\!\!E_{\mathrm{T}}$). The τ leptons themselves are semi-stable. They travel a short distance then decay either hadronically (approximately 64% of the time) or leptonically. Taus that decay leptonically are unable to be reconstructed as taus (since the final state is an electron or muon and two neutrinos), however we accept events with exactly one muon and no electrons to allow for the case where one of the taus decays leptonically. A process known as 'b-tagging' identifies the jets as being produced from the hadronization of b-quarks, distinguishing them from jets produced from the hadronization of other quarks or gluons.

$\not\!\!E_{\mathrm{T}}$ is the transverse component of the energy misbalance in the detector. If all particles passing through the detector had their energy perfectly measured, the total $\not\!\!E_{\mathrm{T}}$ of the event would be zero. Any particles which pass through the detector undetected or any energy mis-measurement will show up as a change in $\not\!\!E_{\mathrm{T}}$. The signal final state has two $\overline{\nu}$ and two $\widetilde{\chi}^0$ which causes events to have a very large $\not\!\!E_{\mathrm{T}}$. Several of the predominant SM backgrounds have no 'true' $\not\!\!E_{\mathrm{T}}$, meaning any measured $\not\!\!E_{\mathrm{T}}$ is from detector effects and not from particles escaping the detector undetected. This analysis removes a large fraction of background events with minimal loss of signal events by selecting only events with a lot of $\not\!\!E_{\mathrm{T}}$. Importantly, this $\not\!\!E_{\mathrm{T}}$ cut rejects nearly all QCD events, which would otherwise pose significant difficulties.

With this analysis' event selection, the predominant backgrounds of this analysis are $t\bar{t}$, W + jets, Z + jets, single top, diboson (WW, WZ, ZZ) and multijet QCD production. In bulk, the kinematic distributions for many of these backgrounds are both similar to each other and the signal. Attempting to assign a portion of the overall contribution to two different processes whose kinematic effects are similar leads to degeneracies which are difficult to resolve. SHyFT exploits the distributions of the number of jets, the number of jets determined to be from the decay of a b-quark, and the number of τ leptons (N_{jet},N_b,N_τ) to remove these degeneracies, leading to a more precise estimate of the numbers of events of each process.

2.1 Overview of this Analysis

This paper describes a SHyFT-based search for $\widetilde{t}$ in 8 TeV proton-proton collisions produced in 2012 by the LHC and recorded by the CMS detector. The approach used in this analysis, called Simultaneous Heavy Flavor and Top (SHyFT) is an evolution of a technique pioneered at the Tevatron and later refined at CMS for the measurement of the $t\bar{t}$ cross section[6]. SHyFT performs a simultaneous global fit of all relevant simulations to data. The result of this fit is an estimate of the number of events contributed by each simulated process to the data.

Key to the power of SHyFT is the method used to match simulation to the data. The final SHyFT fit is four dimensional. Three of the dimensions encode information about the final states: the number of jets, the number of b-tagged jets, and the number of τ leptons which decay hadronically. These three dimensions form a grid of what are called jet/tag/τ buckets. For example, the (4-jet, 1-tag, 1-tau) bucket contains all of the events with four jets, one b-tagged jet, and one hadronically-decaying τ lepton. Note that '4-jet' is an inclusive count, it includes the number of b-jets as well as jets resulting from the hadronic decays of τ leptons. Finally, each bucket is a one-dimensional histogram of a kinematic quantity of the events in the data and simulated samples.

The events produced by the LHC are both complex and numerous. Identifying signal is difficult, extracting a statistically-significant signal from the numerous backgrounds is daunting. For example, the cross-section of W + jets at $8\,\mathrm{TeV}$ is six orders of magnitude greater than the predicted cross-section of $\widetilde{t}\widetilde{t}^*$-pairs at the same energy. One common technique is to select signal and control regions designed to extract a very pure sample of different processes, then fit each region in sequence to eventually extract the rates of the different processes.

There are a few difficulties with this method. Unless event selection alone can produce a single region with all of the signal, the presumably scant number of signal events will be diluted into the control regions. Those signal events in the control regions are effectively missed sensitivity.

Additionally, propagating rates between the various regions can be problematic. If an event yield is estimated in region A, there's an uncertainty involved with propagating that event yield to region B. There are problems with estimating correlated rates as well. If two similar backgrounds have significant contributions in multiple regions, examining each region in sequence can lead to different results depending on the order in which the regions are examined.

The SHyFT method instead uses as much information as posible about the events' final states to discriminate between the signal and various backgrounds. It does this by separating the events into many non-overlapping buckets based on their final states, and not excluding impure buckets. Separating events this way, then fitting all regions simultaneously allows what would normally be considered contamination to contribute usefully to the measurement of all of the rates. This global simultaneous fit also abrogates the difficulties seen with producing distinct regions and fitting them in sequence.

The final state of the $\widetilde{t}\,\widetilde{t}^*$-pairs proposed in the model used by this analysis consists of two b-jets, two τ leptons and $\not\!\!E_{\mathrm{T}}$ resulting from the two ν_τ and two $\widetilde{\chi}^0$ passing through the detector. This somewhat unique final state is distinct from the predominant backgrounds and directs this analysis. Some backgrounds (W + jets, $t\bar{t}$, etc) produce the same basic signature, but there are several characteristics which distinguish $\widetilde{t}\;\widetilde{t}^*$ events from these backgrounds. W + jets events can produce the required muon, jets, and $\not\!\!E_{\mathrm{T}}$, but produce few energetic jets. Every $t\bar{t}$ event has two b-jets, but they produce comparatively fewer τ leptons and have different kinematic characteristics.

The signal and backgrounds therefore have different distributions in jet/tag/τ-space. Fitting all of the data simultaneously not only informs the fit of the kinematic distributions of each event, but of their relative contributions to different final states. It is this additional information that allows SHyFT to produce measurements with such low statistical uncertainties. In addition, some systematic uncertainties themselves can be modeled by a shift in event content in jet/tag/τ-space. This process helps SHyFT produce very competitive measurements compared with its peers.

SHyFT is described in more detail in Chapter 7.

Chapter 3

b-Tagging

The presence of *b*-quarks in the final state of the hard-scattering process enables better discrimination between the signals and backgrounds. A vast number of backgrounds produce jets that arise from the hadronization of light-flavored (non-*b*) quarks, but comparatively fewer have *b*-jets. Distinguishing between jets produced from the hadronization of *b*-quarks and jets produced from the hadronization of light quarks is critical to reducing the otherwise overwhelming amount of background.

The decays of *b*-quarks are characterized primarily by the large masses, long lifetimes, and daughter particles with hard momentum spectra. These properties, combined with the performance of the CMS detector led to the creation of a number of algorithms which exploit these differences to distinguish between *b*-jets and light-jets . Due to the prevalence of *b*-jets in CMS analyses, these algorithms have been refined extensively to provide not only high efficiency (meaning there are comparatively few *b*-jets that are un-tagged) but a low misidentification probability (meaning comparatively few light-flavor jets are tagged as *b*-jets). The Combined Secondary Vertex (CSV) algorithm uses secondary vertex information combined with additional information from tracking and calorimetry to determine the flavor of jets.

This analysis uses the CSV *b*-tagging algorithm to tag *b*-jets. This algorithm was chosen for its high *b*-tagging efficiency and low mistagging performance over the ranged of jet momenta. The algorithm is described in full below.

3.1 Combined Secondary Vertex *b*-tagging Algorithm

3.1.1 Introduction

The CMS collaboration developed a number of b tagging algorithms each designed to exploit a particular trait of b quarks to aid in their identification. For data recorded in the 2011 run, a menu of different taggers based on the presence of secondary vertices, soft leptons, or energetic tracks were each maintained. The state of the art has continued to evolve and newer taggers that combine multiple of these discriminating traits were developed for the 2012 run and beyond.

The combined secondary vertex (CSV)[7] algorithm is one of these next-generation taggers, exploiting both secondary vertex and track kinematic variables combined with a multivariate analysis (MVA) technique. An offline process run by the btag Physics Object Group (POG) trains the MVA using representative samples to generate calibrations which are stored in the CMS Conditions Database (CondDB). Once this is complete, the b tagger calibrations can be loaded in future jobs by loading the calibrations along with the rest of the detector alignment and calibration conditions from CondDB. The online portion of the tagger uses this calibration to generate a discriminator for each jet, corresponding to the confidence that a particular jet is from the decay of a b quark or not.

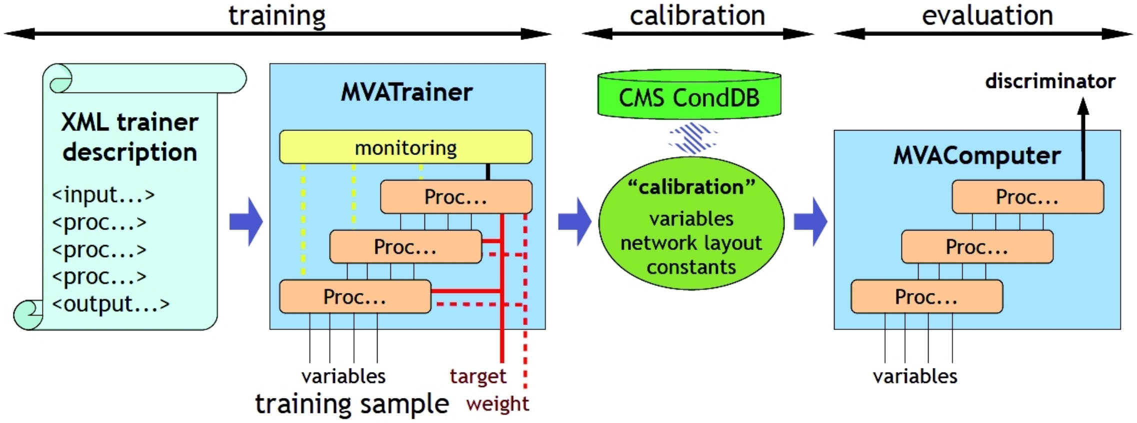


Figure 3.1: Schematic overview of the b tagging framework.

3.1.1.1 Training object selection and vertex type

The training step uses Particle Flow (PF)[8] jets as reconstructed by the AK5 jet algorithm, meaning the jets pass the anti-k_T[9] clustering algorithm with a cone size of $R = 0.5$. Corrections for pile up, electronic noise, η and p_T are applied to the jets. Finally, quality cuts are applied to limit the number of jets that don't originate from the hadronization of quarks. We require:

- neutralHadronEnergyFraction < 0.99
- neutralEmEnergyFraction < 0.99
- nConstituents > 1
- chargedHadronEnergyFraction > 0
- chargedMultiplicity > 0
- chargedEmEnergyFraction < 0.99

The neutral (charged) energy fractions are the proportion of energies deposited in the hadronic (electromagnetic) calorimetry. Requiring a mixture of neutral and charged energy deposits decreases contributions from mesons, who otherwise resemble b-quarks.

One key feature that is used to categorize jets is the vertex category. If there are three of more tracks and a secondary vertex is reconstructed, the jet belongs to the 'RecoVertex' categroy. if there are three or more tracks reconstructed, but there no secondary vertex can be reconstructed, an attempt is made to reconstruct a 'PseudoVertex' consisting of tracks with a signed 2D impact parameter significance of at least 2. If the jet fits into neither of the two previous classifications, it will end up in the 'NoVertex' category.

A key feature of the CMS detector is its very finely segmented silicon tracking subdetector. Particles which interact electromagnetically with the active detector material deposit hits. These hits are reconstructed as tracks and possibly secondary vertices. The MVA has input variables pertaining to the secondary vertex (if any) and impact parameter information. The impact parameter information is stored for tracks fulfilling the requirements:

- $p_T > 1$ GeV
- ≥ 8 valid hits in the tracker
- ≥ 2 valid hits in the pixel detector
- norm. $\chi^2 < 5$
- distance between the track and the primary vertex in the transverse plane is required to be less than 0.2 cm
- distance between the primary vertex and the z-position of the track should be less than 17 cm

Requiring several hits in the tracker and pixel detector ensures the tracks are of good quality. Additional requirements are imposed on the tracks used for secondary vertex reconstruction. Only high purity tracks are used fulfilling the following requirements in addition to the previous requirements:

- $\Delta R(\vec{p}_{jet}, \vec{p}_{track}) < 0.3$
- distance between track and jet axis < 0.2

Secondary Vertices (SV) are reconstructed using the Adaptive Vertex Reconstruction algorithm. During the hard interaction, b-quarks can be produced. These quarks produce intermediate mesons which then travel several microns before the b-quarks decay and hadronize. The distance between the primary and secondary vertex is characteristic of the decay of b-quarks. The CSV algorithm uses the SV information if it exists to enhance the detection of b-quarks.

The reconstructed secondary vertices are filtered according to the following requirements:

- $|m_{\text{vertex(tracki,trackj)}} - m_{K^0}| > 0.05$ GeV
- # tracks ≥ 2
- mass of the weighted vector sum of all tracks < 6.5 GeV
- $\Delta R(\text{vertex,jetaxis}) < 0.5$
- fraction of tracks shared with Primary Vertex (PV) < 0.65
- 2D vertex flight distance > 0.01 cm
- 2D vertex flight distance < 2.5 cm

- 2D vertex flight significance > 3.0

The training trees that contain the track and vertex variables are only filled for jets with at least 3 tracks for which the impact parameters' information was stored. Afterwards, additional track selection cuts are applied, which might result in jets having less than 3 tracks and even jets without tracks, later referred to as 'trackless'. The additional criteria are:

- $\Delta R(\vec{p}_{\mathrm{jet}}, \vec{p}_{\mathrm{track}}) < 0.3$
- distance between track and jet < 0.07 cm
- distance between Principle Component Analysis (PCA) of track and PV < 5 cm. The PCA is roughly the transverse distance between the track and the PV at its closest approach.
- $|m_{\mathrm{tracki,trackj}} - m_{K^0}| > 0.03$ GeV

These additional track selection cuts only affect the variables constructed from tracks. Hence, the variables that are retrieved from the secondary vertex information are not affected.

From this menu of available variables for different vertex classes, the CSV training uses only a subset of these variables which are chosen. Care is taken to keep from using variables that are known a priori to be correlated to each other, since these correlations can bias the fit by providing the same information twice. Table 3.1 provides a list of the available and used variables for each vertex class while Figure 3.2 shows the correlations between selected variables for two different vertex classes. Note the large correlation between the vertex mass and the signed impact parameter significance of the track that raises the total mass of the vertex above the charm quark mass and the pseudo vertex mass. This could indicate that the reconstructed pseudovertex is the vertex of a D-meson decay. Future studies will reevaluate the choice of variables and see if further optimization can be achieved.

3.1.1.2 Jet (p_T,η) reweighting

By design, we don't want the training to explicitly learn anything from the jet p_{T} and η. On the other hand, we have to calculate a weight for each jet depending on its p_{T} and η. This prevents, for instance, high p_{T} jets being tagged more often as b jets simply because jets that originate from b quarks have on average a higher p_{T} than other jets (b quarks are more massive than lighter quarks by about 4 GeV/c^2). The other track and secondary vertex variables might themselves be correlated to the jet kinematics, but this is taken into account by performing the training in a number of (p_{T}, η) bins.

To produce these weights, data is broken into a number of histograms for each vertex class ('RecoVertex','PseudoVertex' and 'NoVertex') and jet flavor (b, c and light). These 9 histograms are then divided into 50 bins of η and 40 bins of p_{T}. The weight of each jet is defined as the inverse of the bin content of the bin described by (vertex class, jet flavor, p_{T}, η).

Table 3.1: The input variables available for the CSV training for the different vertex categories. A variable that is labeled as available could be used in the training, but is not used at the moment. Some variables do not exist for the NoVertex or PseudoVertex categories. These variables are labeled with n/a.

Variable name	RecoVertex	PseudoVertex	NoVertex
jetPt	used	used	used
jetEta	used	used	used
trackSip2dSig	available	available	available
trackSip3dSig	used	used	used
trackSip2dVal	available	available	available
trackSip3dVal	available	available	available
trackSip2dSigAboveCharm	used	used	available
trackSip3dSigAboveCharm	available	available	available
trackMomentum	available	available	available
trackEta	available	available	available
trackPtRel	available	available	available
trackPPar	available	available	available
trackEtaRel	used	used	available
trackDeltaR	available	available	available
trackPtRatio	available	available	available
trackPParRatio	available	available	available
trackJetDistVal	available	available	available
trackDecayLenVal	available	available	available
trackSumJetEtRatio	available	available	available
trackSumJetDeltaR	available	available	available
vertexMass	used	used	n/a
vertexNTracks	used	used	n/a
vertexEnergyRatio	used	used	n/a
vertexJetDeltaR	available	available	n/a
flightDistance2dSig	used	n/a	n/a
flightDistance3dSig	available	n/a	n/a
flightDistance2dVal	available	n/a	n/a
flightDistance3dVal	available	n/a	n/a
jetNSecondaryVertices	available	n/a	n/a

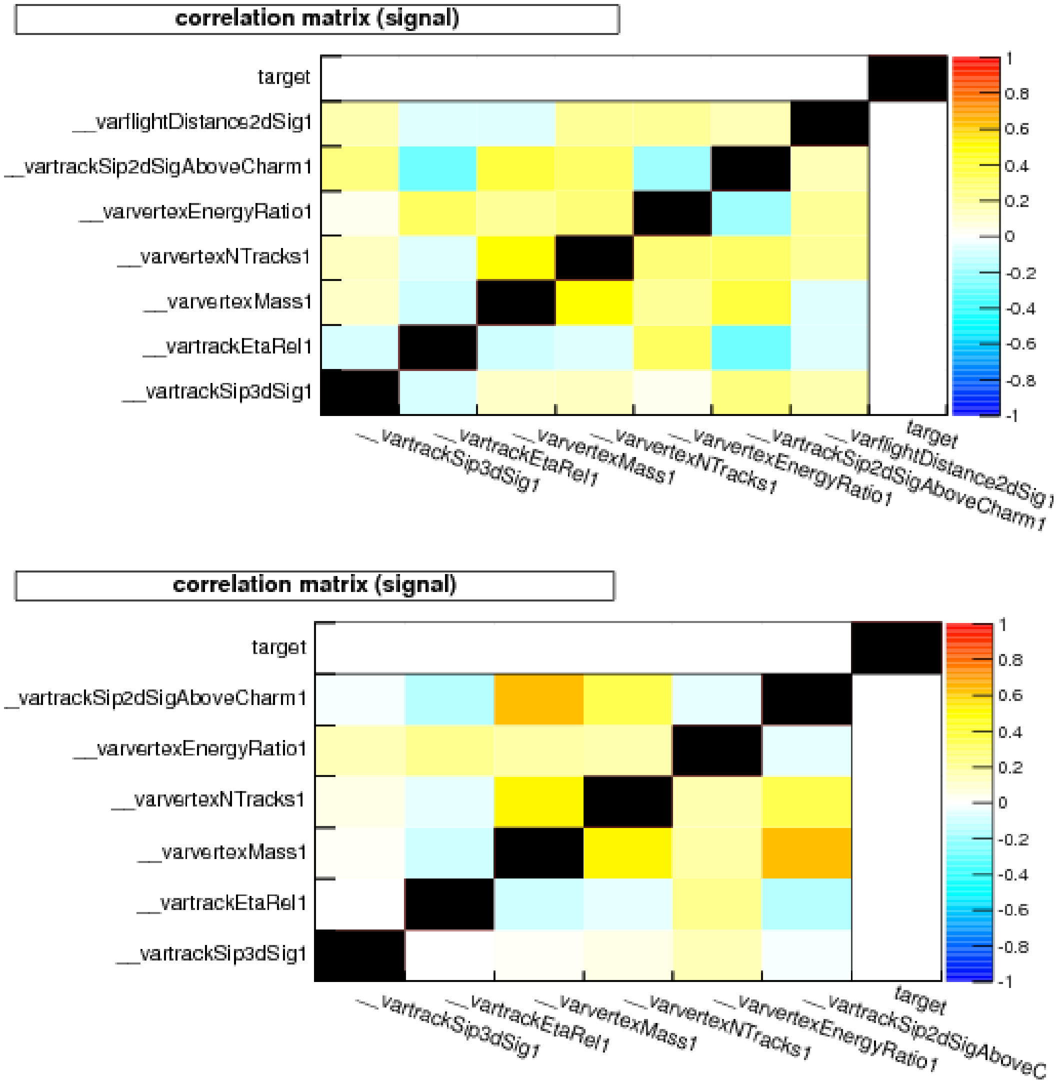


Figure 3.2: Correlation between the variables that are used for the RecoVertex (upper) and PseudoVertex (lower) categories.[7]

Table 3.2: The p_T and η bins used in the training.

Bin number	p_T range (GeV)	$\|\eta\|$ range
0	15 - 40	0 - 1.2
1	15 - 40	1.2 - 2.1
2	15 - 40	2.1 - 2.4
3	40 - 60	0 - 1.2
4	40 - 60	1.2 - 2.1
5	40 - 60	2.1 - 2.4
6	60 - 90	0 - 1.2
7	60 - 90	1.2 - 2.1
8	60 - 90	2.1 - 2.4
9	90 - 150	0 - 1.2
10	90 - 150	1.2 - 2.1
11	90 - 150	2.1 - 2.4
12	150 - 400	0 - 1.2
13	150 - 400	1.2 - 2.1
14	150 - 400	2.1 - 2.4
15	400 - 600	0 - 1.2
16	400 - 600	1.2 - 2.4
17	600 - ∞	0 - 1.2
18	600 - ∞	1.2 - 2.4

3.1.1.3 The different training steps

The CSV tagger trains twice, once for b vs c jets and once for b vs d, u, s, and g jets (b vs dusg). The c jets can form D mesons which have long lifetimes, but their lifetimes are still shorter than the B mesons from b jets. On the other hand, dusg jets don't produce long-lived intermediate mesons. The two trainings are thus optimized for these two cases individually and then are combined to produce a single discriminator.

3.1.1.4 Training in bins of p_T and η

As mentioned, the training is performed in different bins of p_T and η. The bins in η are defined by the detector geometry. $|\eta| < 1.2$ for jets in the barrel, $1.2 < |\eta| < 2.1$ for an intermediate region, and finally $2.1 < |\eta| < 2.4$ for jets in the region with diminished tracker efficiency. The p_T ranges are then defined in a way that each bin has sufficient statistics. For $p_\mathrm{T} > 400$, the forward two η regions are combined since few high p_T jets are produced in the forward region.

3.1.1.5 Likelihood ratio discriminator to combine the variables

Once the variables are translated into float values between 0.0 and 1.0, they are combined with a likelihood ratio:

$$LR = S/(S+B),$$

with

$$S = \prod_{i=1}^{n} pdf_{sig,i}(x_i^j),$$

and

$$B = \prod_{i=1}^{n} pdf_{bkg,i}(x_i^j),$$

where

$$pdf_{sig,i}(x_i^j),$$

and

$$pdf_{bkg,i}(x_i^j)$$

are respectively the probability density functions for the signal and background distributions of variable i for jet j. Figure 3.3 shows an example of the likelihood ratio of b and dusg jets in the RecoVertex category. This likelihood ratio has sharp peaks at 0 and 1, which decrease sensitivity. A normalization step transforms this likelihood ratio into a more widely distributed ratio. The distribution after the transformation is shown in Figure 3.4.

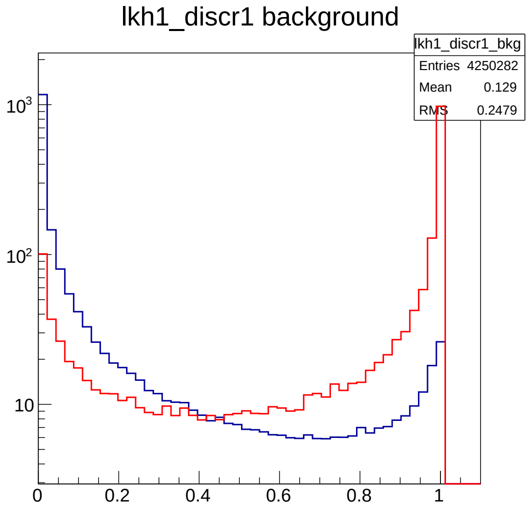


Figure 3.3: The likelihood ratio of b (red) and dusg (blue) jets in the RecoVertex category.

3.1.1.6 Combining the b versus c and b versus dusg trainings

Finally, the two neural networks which each discriminate between b vs c and b vs dusg jets need to be combined. The separate discriminators are simply added with a fraction of 0.25 for b vs c and a fraction of 0.75 for b vs dusg jets. This relative contribution is based on the fraction of c and dusg jets in the decay of W bosons into quarks. This combination (and specific fractions) have been found to be applicable for a wide range of analyses[10].

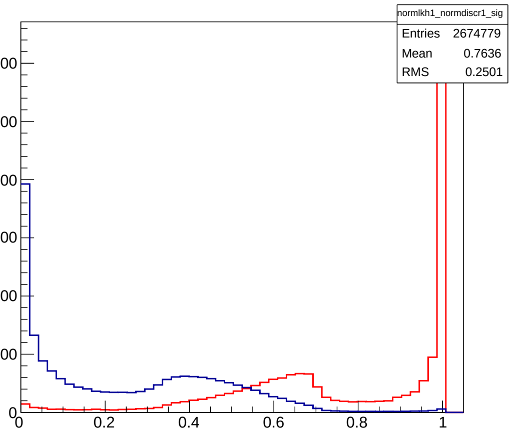


Figure 3.4: The normalized likelihood ratio of b (red) and dusg (blue) jets in the RecoVertex category.

Chapter 4

CMS Detector

This analysis uses data recorded in 2012 by CMS, which is one of two general-purpose detectors for the LHC[1]. The central feature of the CMS apparatus is a superconducting solenoid [11] of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the superconducting solenoid volume are a silicon pixel and strip tracker[12, 13, 14], a lead tungstate crystal electromagnetic calorimeter (ECAL)[15], and a brass and scintillator hadron calorimeter (HCAL)[16], each composed of a barrel and two endcap sections. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid[17]. Extensive forward calorimetry complements the coverage provided by the barrel and endcap detectors.

In the barrel section of the ECAL, an energy resolution of about 1% is achieved for unconverted or late-converting photons in the tens of GeV energy range. The remaining barrel photons have a resolution of about 1.3% up to a pseudorapidity of $|\eta| = 1$, rising to about 2.5% at $|\eta| = 1.4$. In the endcaps, the resolution of unconverted or late-converting photons is about 2.5%, while the remaining endcap photons have a resolution between 3 and 4% [18]. The HCAL, when combined with the ECAL, measures jets with a resolution $\Delta E/E \approx 100\%/\sqrt{E\,[\mathrm{GeV}]} \oplus 5\%$.

In the region $|\eta| < 1.74$, the HCAL cells have widths of 0.087 in pseudorapidity and 0.087 in azimuth (ϕ). In the η-ϕ plane, and for $|\eta| < 1.48$, the HCAL cells map on to 5×5 ECAL crystals arrays to form calorimeter towers projecting radially outwards from close to the nominal interaction point. At larger values of $|\eta|$, the size of the towers increases and the matching ECAL arrays contain fewer crystals. Within each tower, the energy deposits in ECAL and HCAL cells are summed to define the calorimeter tower energies, subsequently used to provide the energies and directions of hadronic jets.

Jets are reconstructed offline from the energy deposits in the calorimeter towers, clustered by the anti-k_t algorithm [19, 20] with a size parameter of 0.5. In this process, the contribution from each calorimeter tower is assigned a momentum, the absolute value and the direction of which are given by the energy measured in the tower, and the coordinates of the tower. The raw jet energy is obtained from the sum of the tower energies, and the raw jet momentum by the vectorial sum of the tower momenta, which results in a nonzero jet mass. The raw jet energies are then corrected to establish a relative uniform response of the calorimeter in η and a calibrated absolute response in transverse momentum p_T.

The particle-flow event algorithm reconstructs and identifies each individual particle with an optimized combination of information from the various elements of the CMS detector. The energy of photons is directly obtained from the ECAL measurement, corrected for zero-suppression effects. The energy of electrons is determined from a combination of the electron momentum at the primary interaction vertex as determined by the tracker, the

[1] The author was not involved with the design, construction, or maintenance of the detector. This brief description is the suggested description produced by the experiment and is here for completeness.

energy of the corresponding ECAL cluster, and the energy sum of all bremsstrahlung photons spatially compatible with originating from the electron track. The energy of muons is obtained from the curvature of the corresponding track. The energy of charged hadrons is determined from a combination of their momentum measured in the tracker and the matching ECAL and HCAL energy deposits, corrected for zero-suppression effects and for the response function of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is obtained from the corresponding corrected ECAL and HCAL energy.

Jet momentum is determined as the vectorial sum of all particle momenta in the jet, and is found from simulation to be within 5% to 10% of the true momentum over the whole p_{T} spectrum and detector acceptance. An offset correction is applied to jet energies to take into account the contribution from additional proton-proton interactions within the same bunch crossing. Jet energy corrections are derived from simulation, and are confirmed with in situ measurements of the energy balance in dijet and photon+jet events. Additional selection criteria are applied to each event to remove spurious jet-like features originating from isolated noise patterns in certain HCAL regions.

The global event reconstruction (also called particle-flow event reconstruction [21, 22]) consists in reconstructing and identifying each single particle with an optimized combination of all subdetector information. In this process, the identification of the particle type (photon, electron, muon, charged hadron, neutral hadron) plays an important rôle in the determination of the particle direction and energy. Photons (e.g. coming from π^0 decays or from electron bremsstrahlung) are identified as ECAL energy clusters not linked to the extrapolation of any charged particle trajectory to the ECAL. Electrons (e.g. coming from photon conversions in the tracker material or from b-hadron semileptonic decays) are identified as a primary charged particle track and potentially many ECAL energy clusters corresponding to this track extrapolation to the ECAL and to possible bremsstrahlung photons emitted along the way through the tracker material. Muons (e.g. from b-hadron semileptonic decays) are identified as a track in the central tracker consistent with either a track or several hits in the muon system, associated with an energy deficit in the calorimeters. Charged hadrons are identified as charged particle tracks neither identified as electrons, nor as muons. Finally, neutral hadrons are identified as HCAL energy clusters not linked to any charged hadron trajectory, or as ECAL and HCAL energy excesses with respect to the expected charged hadron energy deposit.

For each event, hadronic jets are clustered from these reconstructed particles with the infrared and collinear safe anti-k_{t}[9] algorithm, operated with a size parameter R of 0.5. The jet momentum is determined as the vectorial sum of all particle momenta in this jet, and is found in the simulation to be within 5% to 10% of the true momentum over the whole p_{T} spectrum and detector acceptance. Jet energy corrections are derived from the simulation, and are confirmed with in situ measurements with the energy balance of dijet and photon+jet events [23]. The jet energy resolution amounts typically to 15% at 10 GeV, 8% at 100 GeV, and 4% at 1 TeV, to be compared to about 40%, 12%, and 5% obtained when the calorimeters alone are used for jet clustering [21].

Muons are measured in the pseudorapidity range $|\eta| < 2.4$, with detection planes made using three technologies: drift tubes, cathode strip chambers, and resistive plate chambers. Matching muons to tracks measured in the silicon tracker results in a relative transverse momentum resolution for muons with $20 < p_{\mathrm{T}} < 100\,\mathrm{GeV}$ of 1.3–2.0% in the barrel and better than 6% in the endcaps, The p_{T} resolution in the barrel is better than 10% for muons with p_{T} up to 1 TeV [24].

A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [25].

Chapter 5

CMS Computing Model

The luminosity, center of mass energy, and resolution of CMS give access to new unexplored realms in our understanding of the world around us. The actual collisions last for a few tens of nanoseconds; to make this data usable in perpetuity requires a robust and scalable computing infrastructure.

Computing is an integral part of CMS. This infrastructure ushers event candidates from the detector frontend electronics eventually to the end user's analysis. The computing effort is broken into two portions: online and offline. The online portion directly interacts with the detector, recording its data and handling data quality and calibration tasks. These online tasks execute in near-realtime and store the raw data from the detector on permanent storage. Once the data is produced and stored by online, the offline system asynchronously processes and tracks the data for production and analysis tasks.

5.1 Organization of CMS resources

CMS operates a considerable amount of storage and computing resources at CERN, but many other sites including universities and national laboratories provide additional resources. These sites are organized into a number of tiers, which roughly delineate the sites according to their size and responsibilities.

The Tier-0 at CERN is directly attached to the detector, with some resources physically co-located at the interaction point known as Point (P5). The Tier-0 includes the detectors frontend electronics, L1 triggering system[26], High Level Trigger (HLT) farm, tape libraries and sufficient computing resources to perform various prompt workflows like ALignment and CAlibration (ALCA), Data Quality Monitoring (DQM), and prompt Reconstruction (prompt RECO). Additionally, the Tier-0 stores a custodial copy of the raw data to tape as a backup.

Data from the Tier-0 is then divided into a number of event streams based on event content. For example, one stream contains all events where the single muon triggers fired (SingleMu). Several of these streams or Primary Datasets (PDs) are transmitted to each of the Tier-1s, which are large processing facilities usually affiliated with a national laboratory. For safekeeping, the Tier-1s make another copy of the data, ensuring that all raw data from the detector is stored on tape in two locations.

The Tier-1s are dedicated to central CMS use, meaning they only run production workflows and see very little use by end-analyzers. These workflows convert data from intermediate formats into a format that's directly usable by a majority of analyses. This format, the Analysis Object Description (AOD), takes a holistic view of the entire detector and attempts to reconstruct a set of objects which describe the physical objects (muons, electrons, etc..) seen by the detector.

The resulting AOD is divided amongst at least one of the Tier-2s. The Tier-2s are run by the universities and provide the resources most often used by physicists. It would be unwieldy to require each analyzer to manually figure out where the data was stored and how to run their analysis code at the site, so CMS provides the CMS Remote Analysis Builder (CRAB). CRAB gives users a simple command-line interface to access all of the available CMS computing resources. The user merely needs to provide an executable and a dataset to analyze, and CRAB handles data discovery, job submission, job monitoring and returns the results to the location of the user's choosing. 50% of the Tier-2s capacity is dedicated to running user analysis jobs via tools like CRAB, while the other 50% is used centrally to produce Monte Carlo (MC) simulations.

Finally, sites which don't fit in the other three tiers are in the final tier, the Tier-3. Tier-3s are dedicated entirely to user analysis, but provide resources on a best-effort basis. CMS has little manpower to dedicate to supporting the Tier-3s, so the quality of service is determined almost entirely by the local admins. Despite the lack of support, the Tier-3s provide an impressive amount of resources to the experiment and are the home to a number of important physics groups. When the architecture for CMS computing was designed in the early 2000s, it was assumed that network capacity would be the most limiting factor in how data could be distributed amongst the different sites. CMS implemented the Monarch model, which dictated that all of the sites were organized into a tree where each level was a different Tier, and the parent of each site was somewhere geographically close. This meant that if the SingleMu dataset was transferred to the Tier-1 in the US (Fermilab), its products would be transferred down to one of the US Tier-2s and then possibly down to one of the US Tier-3s.

Since then, the availability of high-speed networks has caused a shift in the previous paradigm. Instead of a rigid top-down transfer of data through this tree, sites are increasingly connecting across branches and continents to form a global data transfer mesh, allowing a site in Pakistan to receive data from the University of Nebraska or a site in China to send data to a site in Germany. This allows more choices in data-placement policy, enabling datasets to be replicated to a wider number of sites and in turn being more readily accessible by more people.

Another recent advancement has come out of the Any data, Anytime, Anywhere (AAA) project. Previously to run a job over a dataset meant the data needed to be physically co-located at the same site as where the job was running. AAA deployed an additional interface to each site allowing remote users to performantly access the data stored at the site over the WAN. Each site is then connected to a number of global redirectors, which lets these redirectors know which sites have what files. With AAA, nearly every file in CMS is globally accessible. CMS exploits this to provide fallback functionality for jobs. If a job requests to open a file and it doesn't exist at the site, it can ask the redirector for the location of another replica elsewhere in CMS and transparently open it with only a minor performance hit. Users can even access files directly from their laptop, which is useful for quick studies.

5.2 Subsystems within CMS Computing

Overall, CMS computing has three mildly-separate classes of services and resources it operates:

1. Central bookkeeping, management and workflow services.

2. Resources which provide CPUs.

3. Resources which provide storage.

The central resources are managed centrally by the experiment and coordinate the CPU and storage resources to execute workflows over the data. These services include

1. Dataset Bookkeeping Service (DBS) - Records information about each file and dataset known to CMS, as well as important metadata.

2. Physics Experiment Data Export - Transfers data between sites based on subscriptions to datasets.

3. WMAgent/Workqueue/Request Manager (ReqMgr) - The production operators inject workflows into ReqMgr, which are picked up by workqueues, which are divided and acquired by a number of WMAgents.

4. GlideinWMS - CMS' global job scheduler and resource acquisition system, based on HTCondor.

5. Xrootd redirectors - Implements AAAs goal of making any data accessible anywhere at any time.

6. CRAB servers - The server portion of the CRAB client-server architecture. Receives analysis tasks from users and manages their lifecycle.

7. Asynchronous StageOut (ASO) - Used by CRAB to move user outputs back to the user's home site.

8. SiteDB - A registry of sites and users known to CMS.

9. Virtual Organization Membership Service (VOMS) - Provides CMS authentication services.

10. Web Frontends - Proxies, authenticates and load balances public-facing web services.

Where possible, the online and offline systems use the same infrastructure. Certain critical services are duplicated for online to ensure availability during data-taking. This way, failures in global resources will not affect CMS' ability to record data when the LHC provides collisions.

5.3 CMS Online

When CMS is recording data, the online system interfaces with the detector and handles the products for the first several hundred milliseconds of their existence. The LHC nominally produces collisions every fifty nanoseconds, which corresponds to an event rate of 20 MHz (in 2015, this will be reduced to every 25 nanoseconds or 40 MHz). The CMS detector has approximately 300 million channels. Even in the most optimistic case of one bit per channel, saving all of the data from each event would require accepting and storing $10^{16}\frac{bits}{second}$ of data.

Fortunately, both the event rate and event size can be trimmed by several orders of magnitude. In a typical proton-proton collision, there is a low occupancy of the detector channels, meaning most channels don't have any particles traversing them. With no energy deposition from a particle, the values returned from a segment of the detector are effectively noise above the zero level. A method known as zero suppression drops the channels representing noise from the event output, which reduces the per-event size to a mere tens of kilobytes.

The second optimization necessary to make the data rate from the detector more manageable is to only save interesting events to disk. This process is known as triggering. Conceptually, one would like to output all output channels to some sort of buffer, decide if the event is interesting, then transmit that information onwards to a permanent storage location. The naive approach is unworkable, unfortunately. If the entire contents of the detector need to be stored for each bunch crossing, the problem reverts to what was originally described, 300 million channels triggering at 20 MHz need to be transmitted over a set of links to a buffer, then that buffer needs to be analyzed quickly enough to make a decision before the next bunch crossing arrives (in theory, multiple events could be buffered, but this makes the required buffer size increase as well). There's simply too many bytes being shipped around in too short a time to have a single layer trigger.

A technique adapted by CMS is to have multiple layers of triggers, each both winnowing down the event rate while increasing the amount of data stored at each step. In previous experiments, these layers were implemented using a combination of custom ASICs and FPGAs since existing general purpose CPUs weren't fast enough to provide the necessary operations per second needed to trigger on time. CMS has two levels of triggering, the Level 1 (L1) trigger and the High Level Trigger (HLT). The combined trigger reduces the data rate from 20 MHz to approximately 1,000-200 Hz which is stored, processed, and analyzed downstream.

The L1 trigger is comprised of custom electronics deeply integrated with the muon and calorimetry readout systems. Its goal is to reduce the event rate to under 100 kHz by reading a reduced precision version of these detectors to search for interesting characteristics in the events (e.g. an energetic electron). These characteristics are known as 'L1 seeds' and are the first line of triggering, designed to be broad enough to encompass all possible interesting events while still reducing the event rate by three orders of magnitude. All of this functionality has a deadline of 3.2 μs to make a decision on whether or not to pass the event to the HLT. In the meantime, the full resolution event data is stored in the memories of the frontend readout electronics of each detector subsystem. An example of some L1 seeds from the 2012 LHC run can be seen in Table 5.1.

Trigger	Threshold (GeV)	Rate (kHz)
Single μ ($\eta < 2.1$)	14	7
Double μ ($\eta < 2.4$)	10,0	6
Single e/γ	20	13
Double e/γ	13,7	8
e/γ + μ	12,3.5	3
μ + e/γ	12,7	1.5
Single jet	128	1.5
Quad jet	36	5
H_T	150	5
$\not{E}_{\mathrm{T}}$	40	8

Table 5.1: Rates of selected trigger algorithms at L = 6.66 x $10^{33}cm^2s^{-1}$

If the decision is made at the L1 to keep the event, all of the information from the various detector subsystems has to be read out from the frontends, converted into a common representation, and combined into single events. These streams of events, each representing different L1 seeds are then sent to a farm of several thousand general purpose CPUs known as the HLT farm. Each CPU runs a slimmed down and optimized version of the CMS event reconstruction software which more accurately attempts to reconstruct different physical quantities in the events. The configuration of the HLT, which is known as the ‘HLT menu’, includes possibly hundreds of different trigger paths, each of which can fire to accept the event. These triggers each map to one or several CPUs in the HLT farm and are the final gatekeeper of what data is stored permanently and what is dropped. In order to keep the HLT trigger rate within the predefined budget, some of the paths can be prescaled, which means one out of every N events is stored and the others are dropped. Without prescaling, the triggering thresholds would have to be increased, losing key physics content such as low energy muons. Finally, the luminosity of the machine varies during each fill of the LHC, so some paths can have their prescales vary during the fill to keep a constant rate of events throughout the fill.

Once all the triggering is complete, the HLT writes out approximately 1 kHz of events, which corresponds to $\sim$ 3 Gigabytes/second of data. These data are transferred from the HLT farm at Point 5 to the Tier-0 at the CERN Computing Centre and the offline system is notified, moving these data into offline’s purview.

5.4 CMS Offline

CMS’ compute resources are spread amongst different sites organized in a roughly hierarchical structure based on their logical distance to the detector. There are roughly O(100PB) of disk and O(100k) CPU cores spread globally across different member institutions. The hierarchical structure roughly organizes fewer and larger sites at the lower tiers with the usage patterns becoming more chaotic from the Tier-0 all the way up to the Tier-3s. CMS offline is responsible for enabling access to these resources, generating centrally-produced datasets, providing tools for subsequent analysis, and handling the ancillary computing needs of the experiment.

The Tier-0 is responsible for producing and archiving the raw incoming collision data and performing an initial reconstruction of the data. This reconstruction pass takes the RAW data from the detector, calibrates the values and reconstructs basic traits of the event like the tracks of charged particles or calorimeter depositions. These reconstructed datasets make up the RECO data tier. These RAW and RECO datasets are then each transferred to Tier-1 sites who are responsible for performing a second tape archive and further processing the reconstructed datasets to produce new datasets known as Analysis Object Data (AOD) datasets. The AOD contains calibrated physics objects like muons, photons and jets and is compatible with a large toolkit of user analysis tools known as the Physics Analysis Toolkit. The AOD is then transferred to the $\sim$ 25 Tier-2 sites where users can process and analyze the data for their particular topic.

In parallel to the flow of data from the Tier-0 outwards to higher tiers, CMS also produces and manages several billion simulated events. The CMS event reconstruction software has interfaces to a number of event generators such as Madgraph[27], Pythia[28] or Herwig[29] which allow a wide range of physics processes to be simulated. These interfaces are leveraged to produce GEN datasets with events that describe the 'stable' particles resulting from a particular physical process. The CMS detector is then simulated using the GEANT4 toolkit, which models the magnetic field of the detector as well as the interaction between the generated particles with the detector material. The subsequent readout electronics are also emulated to provide a reasonable facsimile of the actual detector. The simulated detector is then interfaced with the reconstruction software to produce RECOSIM and then AODSIM outputs for eventual use by users.

The AOD and AODSIM datasets, which represent the analysis-level objects of data and simulation are then distributed amongst the Tier-2 sites or, if a user requests it, to a Tier-3 site. To facilitate the access of these datasets by users who might not be members of the institutions where their data is stored, CMS maintains the Cms Remote Analysis Builder (CRAB) tool. CRAB allows a user to take analysis code from their local workstation, transmit it to sites hosting their data, execute the code, and then return the outputs back to their local institution.

Chapter 6

Event Selection

This analysis uses 8 TeV data from the 2012 LHC run and a substantial amount of simulated Monte Carlo events. Within these data, we search for a $\widetilde{t}\widetilde{t}^*$-pair which decays via the following decay chain

$$\widetilde{t}\widetilde{t}^* \to b\overline{b} + \widetilde{\chi}^+\widetilde{\chi}^- \to b\overline{b} + \nu\overline{\nu} + \widetilde{\tau}\widetilde{\tau}^* \to b\overline{b} + \nu\overline{\nu} + \tau\overline{\tau} + \widetilde{\chi}^0\widetilde{\chi}^0 \tag{6.1}$$

where the $\tau\overline{\tau}$ decays into exactly one muon and no electrons.

The input datasets and this analysis' subsequent event and object selections are described in detail beginning in Section 6.1.

6.1 Data Samples

The data are from the January 2013 reprocessing of the 2012 8 TeV data. To limit background contamination, each event is required to have exactly one muon in the final state to be considered. We chose the `SingleMu` Primary Dataset (PD), whose lowest unprescaled muon p_T trigger in the is `HLT_IsoMu24_eta2p1`. Choosing the lowest unprescaled muon p_T trigger maximizes the acceptance of data events. The luminosities in Table 6.1 are the luminosities used in this analysis. The total luminosity analyzed (19.684 fb^{-1}) differs slightly from the total amount of data recorded (19.712 ± 0.513 fb^{-1}) due to computing issues. Some files in an intermediate dataset were lost, leading to missing luminosity.

Dataset	Luminosity (fb^{-1})
/SingleMu/Run2012A-22Jan2013-v1/AOD	0.876
/SingleMu/Run2012B-22Jan2013-v1/AOD	4.400
/SingleMu/Run2012C-22Jan2013-v1/AOD	7.046
/SingleMu/Run2012D-22Jan2013-v1/AOD	7.362
Total	19.684

Table 6.1: Data samples and luminosities

6.2 Signal and Background Monte Carlo Simulation Samples

The backgrounds for this analysis consist of the W + jets, Z + jets, $t\bar{t}$, SingleTop, DiBoson (WW, WZ, ZZ), and QCD multijet production. These backgrounds all contribute to the analysis due to their hadronic activity and real/fake muons/$\not{E}_\mathrm{T}$. Other minor backgrounds are considered to have negligible effects, due to their low cross-sections selection efficiencies. Each of these processes are simulated to Leading Order (LO) with either the

MADGRAPH[27] or PYTHIA[28] event generators. Once these generators complete the simulation of the hard interaction, 'stable' τ leptons can remain. Though τ leptons are themselves unstable, their decay lengths are long enough to be effectively stable at the scales simulated by MADGRAPH and PYTHIA. To simulate the decay of these 'stable' τ leptons at larger scales, the event is modified using the TAUOLA[30] package. The standalone TAUOLA package is designed to ensure the τs decay as realistically as possible.

Next, the interactions between the resulting particles and the CMS detector are simulated using GEANT4[31]. These detector-particle interactions include the interactions of the particles with both active sensor material and inert support material. These effects are important to correctly model certain types of particle mis-identification.

Finally, these LO simulations were then scaled to the best-available Next to Leading Order (NLO) or Next to Next to Leading Order (NNLO) cross-sections. These initial scalings become the initial values of the fit, which are later extracted using a data-driven method described below.

The Monte Carlo samples used are listed in Table 6.2. They are from CMS' Summer12 MC production campaign. They were then reconstructed using CMSSW_5_3_2_patch4, which was the official production software version during the Summer12 MC campaign.

Dataset	process
/DY1JetsToLL_M-50_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	Z+1Jet (m >50)
/DY2JetsToLL_M-50_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	Z+2Jet (m >50)
/DY3JetsToLL_M-50_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	Z+3Jet (m >50)
/DY4JetsToLL_M-50_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	Z+4Jet (m >50)
/DYJetsToLL_M-50_TuneZ2Star_8TeV-madgraph-tarball/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	Z+0Jet (m >50)
/DYJetsToLL_M-10To50_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	Z+Jets (10 <m <50)
/QCD_Pt_20_MuEnrichedPt_15_TuneZ2star_8TeV_pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	MuEnriched QCD
/QCD_Pt_20_MuEnrichedPt_15_TuneZ2star_8TeV_pythia6/Summer12_DR53X-PU_S10_START53_V7A-v3/AODSIM	MuEnriched QCD
/TTJets_MassiveBinDECAY_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7C-v1/AODSIM	TTBarJets
/T_s-channel_TuneZ2star_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	SingleTop_t_sChannel
/T_t-channel_TuneZ2star_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	SingleTop_t_tChannel
/T_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	SingleTop_t_tWChannel
/Tbar_s-channel_TuneZ2star_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	SingleTop_tbar_sChannel
/Tbar_t-channel_TuneZ2star_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	SingleTop_tbar_tChannel
/Tbar_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	SingleTop_tbar_tWChannel
/W1JetsToLNu_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	W+1Jet
/W2JetsToLNu_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	W+2Jet
/W3JetsToLNu_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	W+3Jet
/W4JetsToLNu_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	W+4Jet
/WJetsToLNu_TuneZ2Star_8TeV-madgraph-tarball/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	W+0Jet
/WWJetsTo2L2Nu_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	WWJetsTo2L2Nu
/WW_DoubleScattering_8TeV-pythia8/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	WW_DoubleScattering
/WW_TuneZ2star_8TeV_pythia6_tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	WW
/WZJetsTo2L2Q_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	WZJetsto2L2Q
/WZJetsTo2Q2Nu_TuneZ2star_8TeV-madgraph-tauloa/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	WZJetsTo2Q2Nu
/WZJetsTo3LNu_TuneZ2_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	WZJetsTo3LNu
/ZZJetsTo2L2Nu_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v3/AODSIM	ZZJetsTo2L2Nu
/ZZJetsTo2L2Q_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	ZZJetsTo2L2Q
/ZZJetsTo2Q2Nu_TuneZ2star_8TeV-madgraph-tauloa/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	ZZJetsTo2Q2Nu
/ZZJetsTo4L_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	ZZJetsTo4L

Table 6.2: Monte Carlo samples and processes

$\sigma_{\widetilde{t}\widetilde{t}^*}$(pb)	$\sigma_{\widetilde{t}\widetilde{t}^*} \cdot$ L(events)	$\widetilde{t}$ (GeV)	$\widetilde{\chi}^{\pm}$ (GeV)	$\widetilde{\tau}$(GeV)	$\widetilde{\chi}^{0}$ (GeV)
12.97	$2.55 \cdot 10^5$	250	175	165	100
4.885	$9.61 \cdot 10^4$	300	200	190	100
2.050	$4.03 \cdot 10^4$	350	225	215	100
0.936	$1.84 \cdot 10^4$	400	250	240	100
0.4571	$8.99 \cdot 10^4$	450	275	265	100
0.2291	$4.50 \cdot 10^4$	500	300	290	100
0.06515	$1.28 \cdot 10^4$	600	350	340	100

Table 6.3: The SUSY mass points and associated cross-sections investigated in this analysis. The integrated luminosity recorded by CMS and used by this analysis is $\mathrm{L} = 19.684\,\mathrm{fb}^{-1}$

In addition to centrally-produced background samples, this analysis uses seven privately-produced signal samples which model the targeted SUSY process. First, each sample used `MADGRAPH` to generate one million events. These generated events were then simulated using CMS' standard `FastSim`[32] framework to produce `AODSIM` files. For the processes and kinematic regions examined by this analysis, `FastSim` produces nearly identical results as the `FullSim` framework, but with significantly reduced resource requirements (e.g. $\sim$ 1 second versus $\sim$ 2 minutes per event).

These signal samples simulate the cascade decay chain of $\widetilde{t}\,\widetilde{t}^*$-pairs via Equation 6.1. We've investigated seven different scenarios which differ in the masses of the superpartners. By virtue of the different masses, each scenario has a different predicted cross-section and event kinematics. If the SHYFT fit finds no significant excess compared with the SM, measuring the observed and expected yields for each scenario allows one to extract limits on the existence of this SUSY scenario with respect to the mass of the superpartners. This process is described in detail in Section 9.5.

Each collision between two protons at the LHC has a small chance of being interesting physically. To produce a statistically significant number of interesting collisions, the LHC produces and aligns its bunches in a way that involves many proton-proton collisions per bunch crossing. The number of collisions during a particular bunch crossing is known as the 'pile-up' of the event. During the 2012 run, the typical pile-up was $\sim$ 15. This means that an event selected for an interesting collision will, on average, have another $\sim$ 14 low-energy collisions at the same time. Along with the collisions that occur during the same bunch crossing, collisions from previous and subsequent bunch crossings can also noticeably affect the detector. This additional effect is known as 'out-of-time pile-up'.

When simulating events, Simulating the effect of pile-up is performed by first recording a sample of MinimumBias events. These events are recorded by CMS with event selections/triggering removed, effectively producing a sample of background collisions. The MinimumBias sample is then superimposed on top of the simulated events. This effectively models the effect of multiple collisions per bunch crossing. The MC is produced with a simulated pileup distribution that is different than the observed distribution in data. This discrepancy is corrected by weighing each MC event by the ratio of the simulated and ob

6.3 Object Selection

This analysis applies object-level selection criteria to determine what objects reconstructed by CMS reconstruction algorithms are used. These criteria are designed to both maximize the acceptance of this analysis while minimizing the backgrounds and reconstruction errors. The selection criteria for each object (jet, muons, etc) are described below.

6.3.1 Jets

Partons emitted from the collisions travel and hadronize via the strong force, depositing a distinctive jet of decay products in the detector. The Particle Flow Jets (PFJets)[21] algorithm is used to reconstruct these partons from their daughter particles. The particle flow algorithm uses information from all subdetectors to produce a set of track candidates which are then used as inputs to the anti-kt[9] algorithm with a reconstruction cone size of $R = 0.5$. The anti-kt algorithm takes a list of track candidates and attempts to cluster them into cone-shaped groups to produce a list of PFJets in the event. Next, these PFJets' properties are corrected using the `L1 FastJet`, `L2 Relative`, and `L3 Absolute` corrections. `L1 FastJet` corrects jet energies by removing contributions from pileup events. `L2 Relative` and `L3 Absolute` modify the jet energy response as a function of p_{T} and η. After these corrections, jets are required to satisfy:

1. $p_{\mathrm{T}} > 30\,\mathrm{GeV}$
2. $|\eta| < 2.4$

6.3.2 b-tagged Jets

This analysis uses the Combined Secondary Vertex (CSV)[7] b-tagging algorithm (described in Section 3.1.1) to discriminate between b- and light-flavored jets (light flavored jets are jets that result from the hadronization of anything but a b-quark). Each algorithm has a number of 'working points', which make a tradeoff between efficiency and purity. This analysis uses the medium working point, which is a compromise to accept more *b*-tagged jets (efficiency) at the expense of allowing more mistagged *b*-tag jets (purity). In this analysis, no jets are rejected based on their flavor. The number of reconstructed *b*-tagged jets in the final state determines if the event will be placed in a $= 0$, $= 1$, or ≥ 2 *b*-tag bucket.

6.3.3 Muons (μ)

Muon reconstruction takes advantage of both the silicon tracker and muon stations, which are the innermost and outermost detectors. The design of the CMS detector makes it exceptionally efficient at the detection and reconstruction of muons. First, hits in the muon stations are combined using a Kalman fitting technique to produce a probable muon trajectory. This trajectory is then extrapolated backwards to find potentially matching tracks in the silicon detectors. Finally, a global fit is performed to find candidate muon trajectories compatible with both the muon and silicon detectors. The results of this fit are known as 'global' muons.

The main source of fake background muons are charged hadrons which leave a signal in the tracker, penetrate the hadronic calorimeter and produce tracks in the muon subdetectors. These fake muons leave significant deposits in the calorimetry. A fake-reduction algorithm exploits these deposits to reduce the rate of these muon fakes. In addition, muons are required to be isolated from jets, further reducing the rate of reconstructed jets resulting from either the decay of long-lived hadrons or the hadronization of quarks resulting from the hard scattering process.

This analysis uses the `loose` muon identification requirement with additional kinematic cuts applied on top. These additional cuts restrict muon candidates to regions of high detector and trigger efficiency. In total, we require the following for a muon candidate to be considered:

1. $p_{\mathrm{T}} > 35\,\mathrm{GeV}$
2. $|\eta| < 2.1$
3. Relative Isolation < 0.125
4. Is PF Muon
5. Is either a tracker or global muon. Muons are required to be detected by the silicon tracker. Muons which only are detected by the muon subdetector are rejected.
6. The `HLT_IsoMu24_2p1` trigger requirement itself adds the requirement that the muon is both isolated and satisfies $|\eta| \leq 2.1$.

6.3.4 Taus (τ)

Along with jets produced from the hadronization of quarks, this analysis attempts to identify jets resulting from the hadronic decay of τ leptons. Jets which pass the jet selection criteria and are not tagged as b-jets are examined to see if they satisfy the τ identification requirement. We use the byLooseCombinedIsolationDeltaBetaCorr3Hits τ identification working point and additionally require that the reconstructed τ p_{T} is greater than $20\,\mathrm{GeV}$ and $|\eta|$ is less than 2.4. The number of τs identified in each event determines if the event is assigned to a $= 0$ τ or ≥ 1 τ bucket.

6.3.5 Missing Transverse Energy ($\not{E}_{\mathrm{T}}$)

The targeted signal is expected to have very large $\not{E}_{\mathrm{T}}$ due to the presence of multiple νs and $\widetilde{\chi}^0$s in the final state. Many predicted background events have little-to-no true $\not{E}_{\mathrm{T}}$, so we use $\not{E}_{\mathrm{T}}$ in both the selection criteria and kinematic distributions to increase the signal purity and separate signal from backgrounds. The $\not{E}_{\mathrm{T}}$ content in each event is determined using the `PFMET` algorithm with additional type-0 and type-1 corrections, as suggested by the MET POG .

6.3.6 Electrons

This analysis targets the semileptonic muon channel and rejects events if they contain any electrons. If any potential electron candidates pass the `eidTight` electron selection requirements and have a $p_{\mathrm{T}} > 20\,\mathrm{GeV}$, the event is vetoed. The total electron selection criteria, including the eidTight requirements are

1. $p_{\mathrm{T}} > 20\,\mathrm{GeV}$
2. $|\eta_{supercluster}| < 1.4442$ and $|\eta_{supercluster}| > 1.5660$. This veto eliminates electron candidates which traverse through the transition region in the ECAL subdetector.
3. $|\eta_{electron}| < 2.5$
4. Passes MVA identification
5. Relative Isolation < 0.1. The isolation requirement rejects electrons who are spatially close to hadronic jets. This significantly decreases the amount of fake electrons which are mis-reconstructed due to nearby hadronic activity.
6. $E_t > 30\,\mathrm{GeV}$
7. Passes conversion veto. This veto is designed to reject electron candidates which result from decay products interacting with the detector and producing additional electrons.
8. $dB < 0.02$
9. Number of inner tracker hits ≤ 0

6.4 Event Selection

After selecting objects from an event, the entire event must pass a number of selection criteria to be accepted.

1. Pass the `HLT_IsoMu24_eta2p1` trigger. This trigger requires one muon with $p_{\mathrm{T}} \geq 24\,\mathrm{GeV}$ and $|\eta| \leq 2.1$. In addition, this muon is required to be isolated spatially from additional hadronic activity.
2. Have exactly one muon with $p_{\mathrm{T}} \geq 30\,\mathrm{GeV}$.
3. Have zero electrons with $p_{\mathrm{T}} \geq 20\,\mathrm{GeV}$.
4. Have at least one jet with $p_{\mathrm{T}} \geq 35\,\mathrm{GeV}$.
5. Have at least $20\,\mathrm{GeV}$ of $\not{E}_{\mathrm{T}}$.
6. Have a $m_T(l, \not{E}_{\mathrm{T}})$ of at least $50\,\mathrm{GeV}$ (where $m_T(l, \not{E}_{\mathrm{T}})$ is defined as the norm of the vector sum of the transverse components of the lepton energy 4-vector and MET)

In addition to the analysis-level selections described above, additional cleanup selections are done to the sample as recommended by the Physics Validation Team (PVT) prescription:

- Require events with at least 10 tracks to have at least 25% of them to be high purity to remove backgrounds (the “PKAM” filter).
- Require at least one good primary vertex which
    - Passes the fake veto
    - Has at least 4 degrees of freedom (approximately equal to the number of tracks, plus one)
    - Impact parameter with respect to the *xy* plane of the beamspot < 2cm
    - Impact parameter with respect to the *z* coordinate of the beamspot < 24cm
- Reject events with significant noise in the hadronic calorimeter barrel or endcap (HBHE).

Finally, data samples are selected to only contain events corresponding to ‘good’ periods of data taking, determined by the PVT-produced ‘Golden JSON’ luminosity mask. These golden luminosity periods correspond to periods of stable beam where all detector subsystems were operational.

6.5 Kinematic Distributions or ‘Templates’

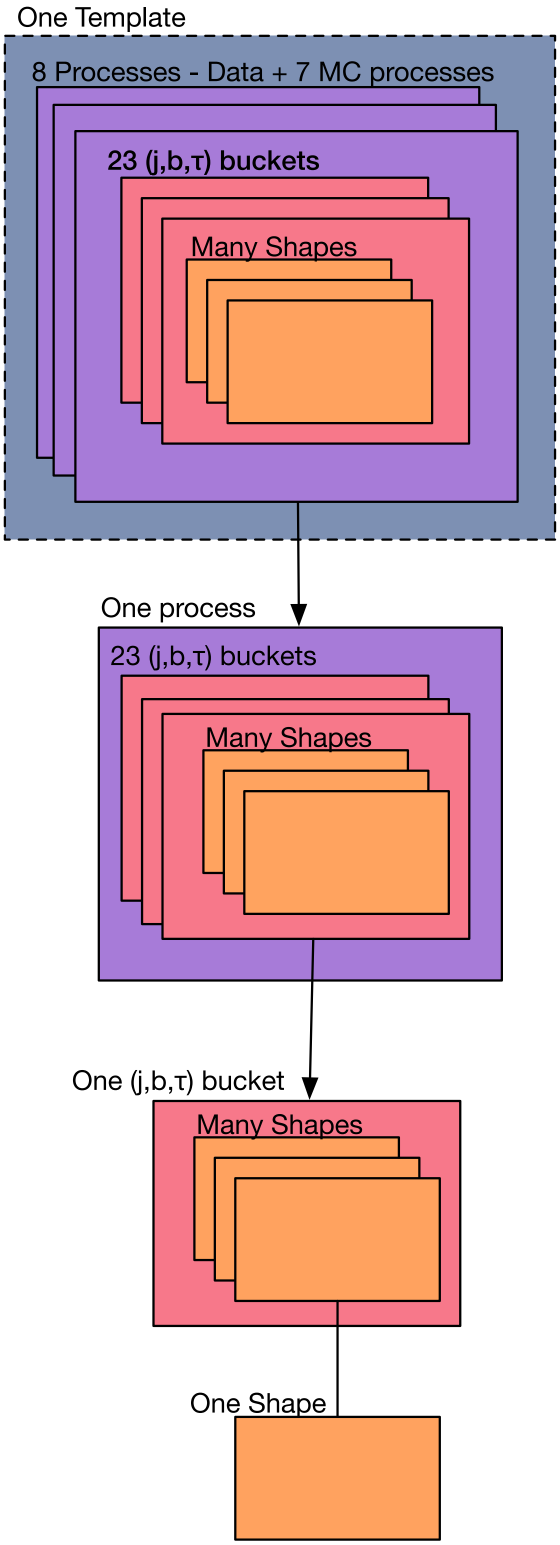
The SHYFT fit combines several hundred shapes into templates. Shapes are kinematic distributions of groups of events. There is a group for each combination of jet/tag bin, kinematic distribution, and process. For example, there is a shape for the $\not\!\!E_T$ distribution of W + jets events with one jet, one b-jet, and zero τ jets in the final state. Templates are a hierarchical grouping of all events considered by this analysis. The hierarchy is as follows: kinematic variable, jet/tag/τ bin, then process (data, signal, or one of the six background processes). CMS simulates different sub-processes separately. To simplify the fit, similar sub-processes are grouped together. For example, the individual ZZ, WZ and WW shapes are all combined into a single DiBoson shape. Figure 6.1 provides a visual representation of this hierarchy.

SHYFT models systematic effects by producing additional templates. Each of these templates are produced by shifting[1] a theoretical parameter away from the nominal value, then reprocessing the MC with this effect. For example, to produce the $+5\%$ Jet Energy Scale (JES) template, the JES parameter in the MC is shifted up by five percent and reprocessed. The template produced with this new MC is then a reasonable facsimile of the effect of a ten percent larger JES. Fit-based systematics are described in detail in Chapter 7.

The contributions from simulation are scaled via the following Scale Factor (SF) to match their predicted yields:

$$SF_{MC} = \frac{L \times \sigma_{MC} \times \varepsilon}{N_{passing}} \tag{6.2}$$

[1] Wordplay.



Each template completely describes the event content for either the nominal selection or a systematic effect (e.g. JES increased by 10%)

Data and the seven MC proceses make up the template

Each process is broken into 23 different buckets based on the (j,b,τ) content of their event final states

Each bucket contains many shapes. Each shape represents a kinematic distribution like MET or sumET

Figure 6.1: Relationship between shapes and templates.

where $L = \int \mathscr{L} dt$ is the integrated luminosity of the processed data (19.684 fb^{-1}), $\mathscr{L}$ is the instantaneous luminosity, σ_{MC} is the theoretical cross-section of the particular sample (either NLO or NNLO), ε is the event selection efficiency, and $N_{passing}$ is the number of MC events which pass all event selections.

After scaling, the sum of the contributions from each all simulated shapes should match the shape seen in data. (ex $\sum_{MC} Shape_{MC} = Shape_{Data}$). These scaled MC shapes are the 'starting point' of the fit, though the fit is free (within constraints) to further scale these contributions. The fit represents this scaling as a multiple of the theoretical expectation. A fitted value of 1.5 represents a contribution of 1.5 times as many events as theoretically predicted.

6.6 Event Counts

The results of this scaling are shown in Table 6.4 along with the observed data rates.

6.7 QCD Prescription

QCD is handled differently than other processes. Current theoretical simulations of multijet QCD suffer with increased jet multiplicity and heavy flavor fraction. Not only is QCD poorly modeled in our signal region, few QCD events pass our event selection. This analysis' $\not{E}_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $m_T(l, \not{E}_{\mathrm{T}}) > 50\,\mathrm{GeV}$ cuts remove nearly all the QCD contribution from the final event selection. Out of 28 million generated events, less than one thousand pass all event selections.

After event selection, each raw MC event represents many expected events in the data. Re-using Equation 6.2,

$$SF_{QCD} = \frac{L \times \sigma_{QCD} \times \varepsilon}{N_{passing}} = 186 \tag{6.3}$$

where σ_{QCD} is $364\mu b$, L is $19635\mathrm{pb}^{-1}$, and $N_{passing}$ is 756. Two separate efficiencies contribute to ε. The first is the filtering efficiency at the theoretical level, $\varepsilon_{filtering} = 3.70 \cdot 10^{-4}$. Most of the possible QCD diagrams are considered unimportant and are excluded as early as possible to minimize the computation required. The second efficiency we consider is the event selection efficiency – the percentage of events which pass our analysis' event selection. The event selection very efficiently rejects QCD events; the efficiency is $\varepsilon_{selection} = 7.02 \cdot 10^{-8}$. We can then directly calculate ε:

$$\varepsilon = \varepsilon_{filtering} \times \varepsilon_{selection} = 3.70 \cdot 10^{-4} \times 7.02 \cdot 10^{-8} = 2.60 \cdot 10^{-11} \tag{6.4}$$

The combination of a large σ_{QCD} and a small $N_{passing}$ produces a large value of $SF_{QCD} = 186$. This means every simulated QCD event must represent 186 events in our data sample. This scale factor is several orders of magnitude larger than the SFs from Table 6.4.

Though there are very few expected QCD events in our high jet multiplicity signal region, accurately modeling its contribution is still important. There is a sizable contribution in the lower jet multiplicity regions, and these contributions can influence other backgrounds that, in turn, can influence the overall contribution in the high jet multiplicity regions. Compounding the problem, the shape from the QCD contribution is similar

Process	$N_{passing}$	σ_{MC}(pb)	ε	SF_{MC}	N_{scaled}
StopStopbar (450 GeV)	52503	0.18	1.11e-06	3.98e-03	208.7
SingleTop_t_sChannel	5926	3.79	3.85e-06	2.87e-01	1699.2
SingleTop_t_tChannel	89438	56.10	2.66e-07	2.94e-01	26251.6
SingleTop_t_tWChannel	27753	11.10	2.01e-06	4.39e-01	12178.9
SingleTop_tbar_sChannel	3422	1.76	7.15e-06	2.47e-01	846.0
SingleTop_tbar_tChannel	49013	30.70	5.17e-07	3.12e-01	15286.6
SingleTop_tbar_tWChannel	27731	11.10	2.03e-06	4.43e-01	12271.3
TTbarJets	384658	234.00	1.44e-07	6.64e-01	255250.0
W+0Jets	19007	37509.00	5.66e-08	4.17e+01	791869.3
W+1Jets	450043	6662.78	4.79e-08	6.27e+00	2820982.6
W+2Jets	986427	2159.24	4.14e-08	1.76e+00	1732294.2
W+3Jets	730985	640.37	7.34e-08	9.23e-01	674402.0
W+4Jets	241456	264.04	2.60e-07	1.35e+00	325623.8
WW	194062	54.84	1.62e-07	1.74e-01	33821.9
WWJetsTo2L2Nu	104828	5.76	5.22e-07	5.90e-02	6184.5
WW_DoubleScattering	4647	0.58	2.63e-06	2.99e-02	139.1
WZJetsTo2Q2Nu	78956	2.49	3.17e-07	1.55e-02	1224.4
WZJetsTo3LNu	78208	1.08	6.07e-07	1.29e-02	1005.9
WZJetsto2L2Q	1	4.49	1.06e-06	9.36e-02	0.1
Z+0Jets (m >50)	39834	3503.71	6.31e-08	4.34e+00	172814.8
Z+1Jets (m >50)	262374	666.30	2.44e-08	3.20e-01	83913.2
Z+2Jets (m >50)	52409	214.97	4.46e-07	1.88e+00	98765.4
Z+3Jets (m >50)	279448	60.69	1.14e-07	1.36e-01	38053.9
Z+4Jets (m >50)	207783	27.36	1.96e-07	1.05e-01	21872.7
Z+Jets (10 <m <50)	4193	13124.07	3.02e-08	7.79e+00	32649.2
ZZJetsTo2L2Nu	17974	0.71	1.96e-06	2.74e-02	492.1
ZZJetsTo2L2Q	10996	2.49	2.78e-06	1.36e-01	1497.2
ZZJetsTo2Q2Nu	1	4.93	1.25e-06	1.21e-01	0.1
ZZJetsTo4L	51236	0.18	4.44e-07	1.57e-03	80.5
QCD	756	3.64e+08	2.60e-11	1.86e+02	140602.0
Total Simulation	4456084	-	-	-	7302281.1
Data	-	-	-	-	7014854

Table 6.4: Input Event SF and Yields

to other backgrounds. Without a separate, data-driven estimate of the QCD yield, these correlated shapes strongly drive the other backgrounds into non-physical regions.

To extract a data-driven estimate of the QCD contribution, we perform a separate fit in a QCD-enriched region using a suitable proxy for the QCD shape. As noted, there are two difficulties to work around – insufficient statistics leading to poor shapes and imprecise theoretical modeling of the underlying QCD hard scattering process. We make three simplifying assumptions. Lower jet multiplicities are more accurately modeled by simulation, the shapes of Z + jets and QCD are similar, and QCD kinematic distributions reasonably similar both with and without b/τ tagging. These assumptions stem from the fact that less jets in the final state involve less vertices and therefore are less sensitive to theoretical uncertainties, while Z + jets and QCD contributions result from similar mismodeling effects. Finally, comparing shapes showed that b/τ tagging requirements did not bias the QCD shape.

6.7.1 QCD Shape

Since few QCD events pass event selection, it is difficult to extract a representative shape from simulation. The QCD shapes are instead obtained using a combination of pre-tagged[2] QCD shapes and Z + jets shapes obtained from simulation. The $N_{jet} = 1$ pre-tagged QCD shape has enough statistics to make smooth shapes, so all $N_{jet} = 1$ QCD buckets use this shape. The higher jet multiplicity buckets suffer from poor statistics, even without b/τ tagging. For these buckets, we use proxy shape derived from Z + jets, which reasonably resembles the QCD shape. In all cases, the augmented QCD shapes are renormalized to the simulated (default) QCD yields. The net effect is that the shapes are replaced, but the overall event rate remains the same.

6.7.2 QCD Normalization

After modifying the QCD shapes, we then estimate the QCD event yield for each jet/tag bucket. Since there are few QCD events in the signal region, we perform a fit to the data in a QCD-enriched sideband and later propagate these values to the signal region. We can't simply use SHYFT for this background because the shapes are substituted with modified shapes whose distribution in each bucket is quite different than the distribution we would expect from using the simulated QCD shape directly.

This sideband uses the normal event selection, but with no $\not{E}_{\mathrm{T}}$ or $m_T(l, \not{E}_{\mathrm{T}})$ cut. With this modified event selection, a fit of the full $\not{E}_{\mathrm{T}}$, $m_T(l, \not{E}_{\mathrm{T}})$ distribution for each jet/tag bucket is used to extract the QCD event yields. To simplify the fit, the Z + jets, DiBoson and SingleTop samples are combined into one Electroweak (EWK) category. Then, QCD, $t\bar{t}$, W + jets and EWK are fit to data, yielding the QCD contribution for each jet/tag bucket. To take into account measured cross-sections, a gaussian penalty term is applied to the $t\bar{t}/W$ + jets/EWK contributions. These gaussians have widths of 15%/10%/10% and are conservatively chosen to be compatible with CMS' measured values. The QCD contribution is allowed to freely vary.

[2] Pre-tagged means before b/τ tagging requirements

A subset of the output distributions are shown in Figure 6.2. The complete normalization results and distributions are located in Appendix B.

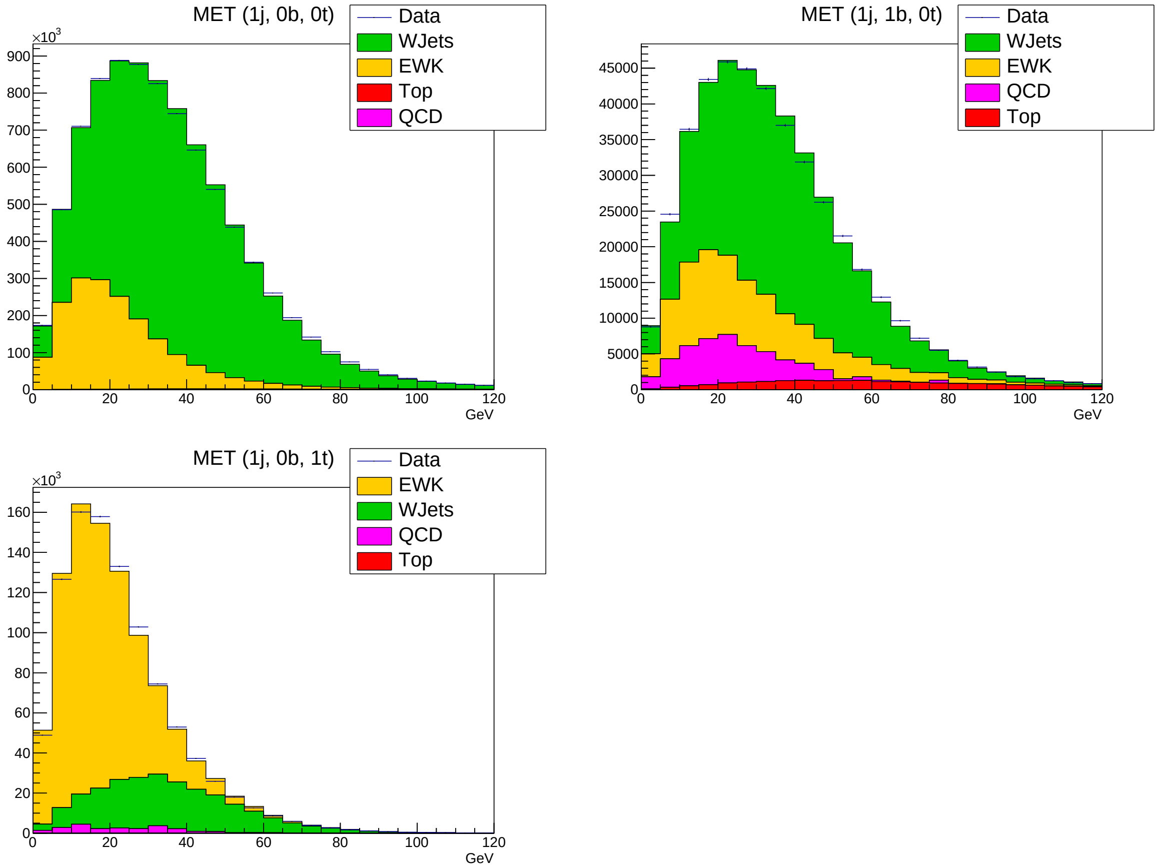


Figure 6.2: Fit results for QCD normalization, $= 1$ jet buckets, no $\not\!\!E_T$ cut.

6.7.3 QCD Prescription in Nominal Fit

The scale factors from the sideband region are then used to estimate the QCD rate in the nominal region. For each jet/tag bucket j, we fix the QCD rate in the signal region using the fit-extracted scale factors:

$$N_{SR'}(j) = SF_{CR}(j) \times N_{SR}(j) \tag{6.5}$$

Where $SF_{CR}(j)$ is the fitted scale factor from bucket j and $N_{SR'}/N_{SR}$ are the augmented/original QCD rates in the signal region, respectively.

Due to the different event selection criteria, the predicted SFs from the sideband region might not necessarily translate into the nominal region. Instead of fixing the signal region QCD yield to the sideband-extracted yield, the yields of each QC jet multiplicity rates are then allowed to vary in a band around the extracted value. Uncertainty penalties of 200% are applied to the QCD contribution to account for shape and rate mismodeling effects.

Chapter 7

SHYFT in Detail

This chapter describes in further detail the SHYFT method, which can extract faint signals from large backgrounds and significantly reduce some systematic uncertainties. Instead of defining separate signal and control region(s), SHYFT divides the entire data sample into buckets based on N_{jets}, N_b, and N_τ. Not only does this allow a better fit, many of the largest uncertainties result in a relative scaling between the different buckets. This scaling is used to constrain systematic uncertainties. Since this analysis uses N_{jets}, N_b, and N_τ to separate the events into different regions, systematic uncertainties which influence these distributions have a large effect. For instance, the N_{jets} classification of an event depends on whether or not jets pass the p_{T} cut. The Jet Energy Scale (JES) affects the measured jet p_{T}. Thus, higher or lower JES can influence the N_{jets} classification of an event. Because of this, correctly estimating the preferred value of JES is important to this analysis.

The SHYFT fitter is able to use data-driven information about the N_{jets} distribution to constrain and make a data-driven estimate of the JES scale. Similarly, the N_b distribution depends on the btagging efficiency (ε_b). Simulating the effect of different ε_b on the N_b distribution lets the fitter extract a preferred ε_b. These uncertainties are simultaneously applied to the simulation to converge in a region of phase space containing the proper amount of background normalization and uncertainty scale factors. Once the sample is divided into different '(j,b,τ)' buckets and the effects of those systematic effects are modeled, a global likelihood fit is performed. This fit simultaneously extracts out rates of all the samples as well as the modeled systematic effects, indirectly extracting shapes as well.

Accurately modeling backgrounds and uncertainties makes SHYFT an improvement over simple counting and shape-based methods. The method is described in more detail below.

7.1 Sample Discrimination with SHYFT

Without decreasing the acceptance too far, it is difficult to construct a sample containing a high-purity of our $\tilde{t}\tilde{t}^*$ signal using event selection alone. Using tighter and tighter event selection criteria reduces contamination from background events but simultaneously reduces the acceptance of signal events. If $\tilde{t}$ exists, we predict very few $\tilde{t}\tilde{t}^*$-pairs will be produced at the LHC. Throwing away already-small signal statistics via event selection is counter-productive. At the same time, trying to estimate the contribution of a rarely-occurring process against much larger backgrounds leads to significant uncertainties. Balancing these two contrary needs via event selection alone is a challenging task.

It *is*, however, more feasible to produce several regions of the lepton+jets data sample, where these regions are heavily enriched in different background processes. This seemingly sidesteps the tradeoff inherent in choosing how narrowly to select events. Instead of tossing statistics to gain extra purity at the expense of acceptance, keep all of the statistics and produce regions enriched in the signal or one of the backgrounds.

The most significant backgrounds in the lepton+jets sample guide the choices made when dividing the sample. As stated in Chapter 6, $W+$jets, $Z+$jets, $t\bar{t}$, SingleTop, DiBoson (WW, ZZ, WZ), and multijet QCD are the backgrounds with the largest contributions. For instance, the hard interaction from $W+$jets events produces less hadronic jets than $t\bar{t}$ events. Because of this, low jet multiplicity buckets have a large contribution of W + jets events and relatively few $t\bar{t}$ events. Higher jet multiplicity buckets have an increasing fraction of $t\bar{t}$ events compared with W + jets events. Similarly, Z + jets and W + jets have different jet flavor and τ-lepton composition. These backgrounds are then separated into different buckets. These (j,b,τ) buckets let SHYFT pull apart contributions that would otherwise be intertwined with each other. An example cartoon can be seen in Figure 7.1.

At this point, it is important to note that SHYFT isn't entirely novel. It is common to define a signal and one (or many) control region(s), extrapolate the background rates from the control regions, and then propagate the background rates to the signal regions. With SHYFT, the difference is two-fold – the different regions are fit simultaneously, and the full statistics of both our 'signal' and 'control' regions are exploited.

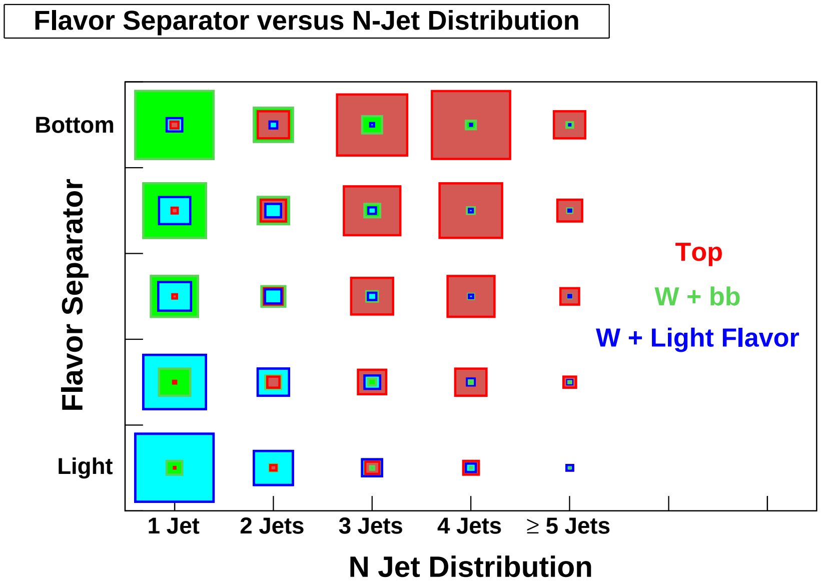


Figure 7.1: Cartoon of jet flavor versus N_{jet} for different samples. The size of each box indicates the contribution of the sample to the bucket. Each variable on its own is not enough to separate the three samples, but together it is easy to distinguish them.

To determine the heavy flavor content of an event, we attempt to *b-tag* each jet using the Combined Secondary Vertex (CSV) algorithm, described in Section 3.1.1. Each N_{jet} category is subdivided into three N_{btag} buckets, each containing events with $(0,1,\geq 2)$ b-tagged jets. Finally, each (N_{jet},N_{btag}) bucket is further subdivided into buckets that contain $(0,\geq 1)$ τ_h candidates in the final state. Separating events events by their τ_h multiplicity improves discrimination of $\widetilde{t}\,\widetilde{t}^*$ due to the two τ_h leptons produced by the decay of intermediate $\widetilde{\tau}$ sleptons and the lack of τ_h in the predominant backgrounds. The backgrounds which have a significant τ_h contribution ($t\bar{t}$, $Z+\text{jets}$) still benefit since the majority of their contribution is in the $\tau_h = 0$ buckets, which is constrained by the $\tau_h \geq 1$ buckets.

If each contribution to a region is estimated sequentially, any cross-contamination between regions are difficult to properly estimate. For instance, assume process X in control region A is estimated, then process Y in control region B is estimated. If regions A and B are very pure, the contributions extracted for X and Y should be uncorrelated with each other, and the order in which the contributions are estimated won't matter. On the other hand, consider the case where both regions A and B have significant contributions from both processes. As more events from each process bleed into both control regions, the measurement of each contribution becomes more correlated with the other. Once correlated, it becomes tricky to accurately perform two measurements one after another. Assuming a contribution for Y, fitting X to the data, then fitting Y to the data using the measured X contribution can yield a different result than fitting in the other order.

Combatting this leads to the second difference between SHYFT and a simple counting experiment. Purifying the regions by excluding cross-contaminated events decreases the acceptance of the analysis. Every event in the excluded region is an event that can't contribute to the statistics of the measurement.

In contrast, SHYFT separates the events into several adjacent regions and fits all of the contributions simultaneously. During this simultaneous fit, regions of high purity help to tightly constrain each process in regions of low purity. Since it's no longer necessary to exclude regions with mixed contributions, the event selection can be broadened to include as much information as possible. These two traits cause SHYFT to exploit as much information as possible, leading to more accurate measurements.

7.2 SHYFT and Systematic Effects

While both the N_{jets}, N_b, N_τ spectrum and kinematic distributions together distinguish the different signal and background processes, they are not sufficient to accurately model the observed data.

For example, the $t\bar{t}$ event yield in a particular bucket depends on the $t\bar{t}$ production cross section. Allowing the fit to scale the cross section causes the $t\bar{t}$ yields in every bucket to change proportionally in the same direction. Scaling whole processes doesn't change the relative normalization between buckets, which means any mismodeling in the simulated (N_jet, N_b, N_τ) spectrum will remain uncorrected.

To take into account this possible mismodeling, we teach the fitter the effect of various systematic parameters on the relative normalization between buckets. In the case of $t\bar{t}$, the number of events in each bucket depends not only on the $t\bar{t}$ production cross section, but also on the b-tagging efficiency. Given a certain b-tagging SF, an event could be placed into a $N_b = 1$ bucket. If we instead assume an increased value of the b-tagging SF, this same event could end up in the $N_b \geq 2$ bucket. Choosing different values of the b-tagging SF effectively moves events between buckets, modifying the N_b spectrum, which could possibly lead to better agreement between simulation and data. After this addition, the $t\bar{t}$ yield in each bin is determined by both the production-cross section and the b-tagging efficiency. A graphical example can be seen in Figure 7.2.

We model the effect of b-tagging and Jet Energy Scale by first reprocessing our simulation with different SFs and then allowing the fitter to vary these SFs during the fit. This concept is expanded in Section 7.4, but we will first describe a simplified version of SHYFT that doesn't take into account systematic effects.

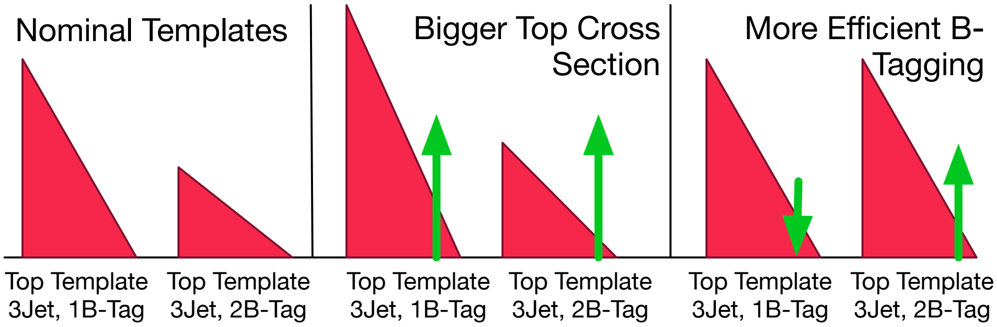

Nominal Templates
Bigger Top Cross Section
More Efficient B-Tagging
Top Template 3Jet, 1B-Tag
Top Template 3Jet, 2B-Tag
Top Template 3Jet, 1B-Tag
Top Template 3Jet, 2B-Tag
Top Template 3Jet, 1B-Tag
Top Template 3Jet, 2B-Tag

Figure 7.2: Cartoon illustrating how increasing the b-tag scale factor or the top cross section affects single and double tags differently. As the top cross section is increased, both 1-tag 2-tag templates get larger. However, if the b-tagging becomes more efficient, some 1-tag events become 2-tag events, leading to fewer events in the single-tag template and more in the double-tag template. Thus, the relative numbers of 2-tag to 1-tag events are directly proportional to ε_b, and because of this our fit can determine it from our own data.

7.3 SHyFT Fitter without Systematic Effects

Consider first a fit that neglects systematic effects. The SHyFT fitter is then quite simple. The sample is divided into buckets based on the number of jets, number of b-tagged jets and number of τ_h candidates (which we call *jet/tag buckets*). We consider buckets for $(1,2,3,4,\geq 5)$ jets, $(0,1,\geq 2)$ b-jets, and $(0,\geq 1)$ τ_h candidates where $N_b + N_\tau \leq N_{jets}$. This splits the sample into 26 buckets of 1-dimensional kinematic distributions.

At this point, each (N_{jets}, N_b, N_τ) bucket contains a 1-dimensional histogram for each process. If we label each histogram's bins as l, the fit becomes

$$\sum_x S_x(N_{jets}, N_b, N_\tau)_l = S_{Data}(N_{jets}, N_b, N_\tau)_l \tag{7.1}$$

where x is each simulated process and $S_x(N_{jets}, N_b, N_\tau)_l$ is the event yield for a specific bucket. Effectively, Equation 7.1 states the fit wants to make every bin in the simulation have the same yield as the data.

Unless our simulation perfectly models the observed data, Equation 7.1 is false. The only inputs are the theoretical simulations and the data. We will complicate things later but, for now, assume the shapes of the kinematic distributions of the simulated events are correct, but the overall normalization of those shapes can vary. Introduce a new parameter K_x which represents the multiplicative scale factor between the theoretical cross-section and the fitted value for process x. The predicted event yield (N_x) for a single sample x then becomes

$$N_x(N_{jets},N_b,N_\tau)_l = K_x \cdot S_x^{MC}(N_{jets},N_b,N_\tau)_l \tag{7.2}$$

Expanding to consider all of the constituent processes (and remembering that by construction $S_{Data} = N_{Data}$)

$$\sum_x K_x \cdot S_x^{MC}(N_{jets},N_b,N_\tau)_l = N_{data}(N_{jets},N_b,N_\tau)_l \tag{7.3}$$

Equation 7.3 is the ideal relationship between our data histograms, simulation histograms and cross-section scale factors. We assume the simulation for each process has the right shape, meaning each histogram is accurate to within a single per-process Scale Factor (SF) given by K_x. Extracting the contribution from each sample x then requires finding the values of K_x that most closely causes Equation 7.3 to be true.

There are many ways to find the set of K_x that causes Equation 7.3 to most accurately be true. This analysis begins by expressing the *likelihood* (L) that specific values of K_x produce simulated shapes that statistically overlap with the data.

$$L = \prod_{N_{jets},N_b,N_\tau,l} P_{oi}(N_{data}(N_{jets},N_b,N_\tau)_l, \sum_x N_x^{MC}(N_{jets},N_b,N_\tau)_l) \tag{7.4}$$

where $P_{oi}(x,y)$ is the poisson probability that x and y are statistically compatible. Finding the best fit for our model means finding the values of K_x that maximize L.

Before performing the actual fit, two more transformations need to be applied to L. Instead of maximizing L, we instead minimize $-\ln L$. Symbolically, maximizing L and minimizing $-\ln L$ are equivalent, but attempting to use L directly suffers from numerical accuracy when executed on a computer. Expressed as L, the maximum likelihood for this fit is on the order of e^{-20000} and even small changes of K_x cause gigantic fluctuations in L. In order to search the parameter space for the maximum, the fitter must repeatedly compute the derivative of L with respect to K_x and these derivatives can exhaust the precision of even double-precision floating point arithmetic.

L in Equation 7.4 has one other undesirable characteristic. Each bin we would like to examine is referenced by (N_{jets},N_b,N_τ,l). In effect, this is a four-dimensional histogram which is cumbersome to manage computationally. Within the fitter, the (N_{jets},N_b,N_τ) indices are unfolded, yielding a one-dimensional histogram whose bins are indexed via l. Instead of having 26 different 1-dimensional histograms, an equivalent "long" 1-dimensional histogram with each of the 26 different input histograms are placed side by side is used internally. To keep the notation clear when fit-based systematics are used, $\ln L$ will continue to show N_x in terms of (N_{jets},N_b,N_τ,l)

Given our prior knowledge of the processes we are simulating, we would like to apply constraints to some of our backgrounds - they are either constrained by other parts of this analysis or by their CMS-measured 8 TeV cross sections. To implement this, we apply a gaussian penalty term which increases as the fit diverges from our chosen central value. These constraints are represented by

$$e^{\frac{-(z-\bar{z})^2}{\sigma^2}} \tag{7.5}$$

where z, $\bar{z}$ and σ are the fitted value, expected value, and width of the constraints, respectively. We are minimizing the least log-likelihood, so the gaussian penalty to the likelihood becomes the following log-gaussian penalty to the least log-likelihood

$$\frac{1}{2}\frac{(z-\bar{z})^2}{\sigma^2} \tag{7.6}$$

With this additional penalty term, the total log-likelihood we wish to minimize is

$$-\ln L = -1 \cdot \left\{ \sum_{N_{jets},N_b,N_\tau,l} \ln P_{oi}(N_{data}(N_{jets},N_b,N_\tau)_l, \sum_x N_x^{MC}(N_{jets},N_b,N_\tau)_l) \right\} + \left\{ \frac{1}{2} \sum_n^{constraints} \frac{(z_n - \overline{z_n})^2}{\sigma_n^2} \right\} \tag{7.7}$$

where the $P_{oi}(q,p)$ is the poisson probability that q and r are statistically compatible and the sum over n is over each constraint applied to the fitter. Taking the natural logarithm of P_{oi} leads to

$$\ln P_{oi}(q,p) = q \ln p - p - \ln\Gamma(q+1) \tag{7.8}$$

where Γ is the gamma function

$$\Gamma(n) = (n-1)! \tag{7.9}$$

7.4 Fit-based Systematics

Letting the fitter modify the overall cross-section of each process isn't enough to guarantee good agreement between simulation and data (see Section 7.2). Since each process is broken into multiple jet/tag buckets, it's possible that mismodeling of various systematics could cause events to be distributed incorrectly among those buckets. The fitter only modifies the overall cross section of each process, it has no way to fix an imbalance between the jet/tag buckets within a process (Figure 7.2).

Modeling these imbalances is a tractable problem. In this analysis, we choose to divide events into buckets based on the $N_{jets}/N_b/N_\tau$ multiplicities, so effects that change the rate these objects are both produced and detected are obvious targets for improvement. An effect which changes the number of jets produced will cause a change in the N_{jets} distribution, so we reprocess our simulation assuming different effects and extract templates whose shapes have been modified from the nominal.

To model these effects for a single jet/tag bucket, the fitter needs a continuous, real-valued function to vary for the fit, but it's computationally infeasible to generate a set of systematic histograms for each value of the systematic parameter. Instead, we generate additional sets of templates for $\pm 1\sigma$ or optionally two additional points at $\pm 2\sigma$ and use a polynomial to smoothly define the changes in normalization for each bucket. One such polynomial might model the change in relative normalization for the contribution of $t\bar{t}$ in the (1-jet, 0-btag, 0-τ) bucket as the systematic effect is shifted from 80% to 120% of the nominal value. Importantly, the systematics considered in this analysis aren't strongly correlated (e.g. b-tagging depends primarily on tracking performance, while JES is influenced primarily by calorimetry), so we can factorize the contribution from each systematic into its own scalar value.

A single polynomial only describes the effect of a systematic on a single jet/tag bucket of a single process. To model the effect for all processes and all buckets, we combine several polynomials together to produce what we call "polynoids". Polynoids are collections of polynomials that describe an effect on all buckets where each individual polynomial represents the effect of that systematic on a single bucket. Figure 7.3 is an example of polynoid that shows the effect of Jet Energy Scale (JES) on the SingleTop process. In this case, 1.0 represents the nominal, or 'unshifted', JES and 1.10 represents JES increased by 10% (to 110%). To produce the 110% sample, we scale the 4-momentum of each jet by a factor of 1.1 and compensate the $\not\!\!E_T$ by the appropriate amount. Performing this scaling pushes some jets from just below the jet p_T threshold to just above the jet p_T threshold and simultaneously pulls some events from just above the $\not\!\!E_T$ threshold to just below the $\not\!\!E_T$ threshold. The net effect causes some events to jump from the 1-jet to 2-jet buckets while other events who previously passed the $\not\!\!E_T$ cut then fail the $\not\!\!E_T$ cut.

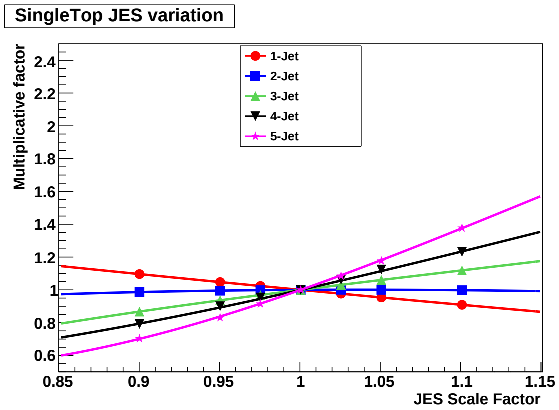


Figure 7.3: Polynoid representing the effect of JES on the SingleTop distribution. A value of 1.0 represents the nominal effect of JES (e.g. 100% of the expected value), 1.05 represents shifting the JES up 5% (105%), etc. As the JES varies, the different jet multiplicities gain and lose events at different relative rates.

Modeling these systematics in the fit provides a direct way to extract the uncertainty on the fit due to the these effects. Before the inclusion of the systematics, the fit error contained only the statistical error. Adding an additional parameter to the likelihood broadens the log-likelihood at the minimum. The uncertainty of the measurement is proportional to the width of this minimum, so any new parameters add additional uncertainty to the measurement. At first glance, it seems counter-intuitive to voluntarily add uncertainties to our measurement. It's important to note that whether or not these systematic effects were integrated within the fit, their effects and associated uncertainties would still need to be quantified. When quantifying these effects out-of-band, their uncertainties are larger than when they are applied to the fit directly. In effect, directly modeling these effects and their uncertainties instead of applying them later results in a net gain of sensitivity.

For comparison with other measurements, it is helpful to divide the combined error on our measurement into statistical and systematic errors. To do this, the fit is run twice. The first fit is done with all parameters floating. Then, the fit is performed again with all the parameters except the one in question fixed to the values obtained in the first fit. Subtracting the uncertainties of the two fits in quadrature yields an estimate of the contribution due to the systematic in question. It is important to note that the statistical and systematic uncertainties are convolved, so the uncertainty extracted from this procedure is just an estimate.

To model the effect of systematic y on sample x, an additional term $R_{x,y}(N_{jets}, N_b, N_\tau)$ is added to the fit.

$$N_x(N_{jets}, N_b, N_\tau)_l = \left(\prod_y R_{x,y}(N_{jets}, N_b, N_\tau)\right) \cdot K_x \cdot S_x^{MC}(N_{jets}, N_b, N_\tau)_l \tag{7.10}$$

where $R_{x,y}(N_{jets}, N_b, N_\tau)$ is the multiplicative scale factor for sample x due to systematic y on the bucket (N_{jets}, N_b, N_τ). This yields our total log-likelihood

$$-\ln L = -1 \cdot \left\{ \sum_{N_{jets}, N_b, N_\tau} \ln P_{oi}(N_{data}(N_{jets}, N_b, N_\tau)_l, \sum_x N_x^{MC}(N_{jets}, N_b, N_\tau)_l) \right\} + \left\{ \frac{1}{2} \sum_n^{constraints} \frac{(z_n - \overline{z_n})^2}{\sigma_n^2} \right\} \tag{7.11}$$

where N_x^{MC} contains an implicit product of each systematic effect y from Equation 7.10. In summary, the fit varies these parameters to attempt to minimize $-\ln L$:

- K_x - The production cross section for each simulated process x
- $R_{x,y}$ - Each systematic parameter y has an associated SF, which is translated via a polynoid to give the relative normalization for each bucket of process x

Along with the fit parameters, the fit is given as inputs

- $S_x(N_{jets}, N_b, N_\tau)_l$ - The per-bucket shape of each simulated process or data x
- $\overline{z_n}$, σ_n - The mean $\overline{z_n}$ and width σ_n of the gaussian constraint on parameter n
- The polynoids describing the normalization on each simulated process and bucket due to each systematic effect

Chapter 8

Method

This analysis first takes data recorded by CMS and simulated events and divides them into buckets based on their jet, b-tag, and τ multiplicities. Then, the events in each bucket are analyzed to produce kinematic distributions. To model systematic effects, the simulated events are processed several times using different assumptions of different theoretical parameters. The distributions in each bucket and their associated systematic effects are then simultaneously fit to extract the event yield of each process and the observed effect of their systematic effects.

8.1 Workflow

This analysis is performed as a sequence of discrete steps, each consuming the output of the previous stage to generate a new set of outputs. The main goal of these stages is to produce kinematic distributions for each sample divided into jet, b-tag, and τ multiplicities. The fitter uses a collection of these distributions (known as templates) along with systematic constraints to estimate the cross sections of the signal and backgrounds.

CMS provides simulation and data in the Analysis Object Definition (AOD) format. Since these data are intended to be used by the majority of analyses, they contain a lot of information about each event. The generality AOD data provides makes them broadly usable but the amount of data stored makes them both slow to read and occupy a lot of disk space. Worse, the AOD-level objects aren't directly usable by analyses. Since each group treats different physics objects (e.g. muons) somewhat uniquely, the responsibility for these differences are pushed down to the individual analyzers, meaning that a significant amount of reprocessing has to be done on AOD level objects to produce the specific objects each analysis needs.

Converting AOD to analysis-appropriate objects was performed by the Physics Analysis Toolkit (PAT). PAT takes AOD, performs standard cleaning and reconstruction algorithms and generates PAT-tuples with values that are usable by a group of analyses. This analysis uses PAT-tuples created by the RA2Tau analysis group. To give a sense of scale involved in generating these PAT-tuples, there are approximately one hundred million events and it takes nearly two seconds to generate a single PAT event. Two hundred million seconds is more than six years of CPU time. Even with massive parallelization, groups studying similar processes share the PAT-tupleization steps to lower the total resources needed to produce and store these PAT-tuples.

After the PAT level, more significant cuts are needed to both reduce the physics content of the events and remove extraneous data. This analysis has an EDM-tuple step which takes the PAT-tuples, performs some selection criteria (defined in Chapter 6), and writes out simple data structures that are merely a list of the important kinematic variables. These EDM-tuples are both compact and fast to read. Packing all the final calibrations and selections in the EDM-tuples makes it so that the subsequent steps can iterate quickly - they're mostly just lists of floating-point numbers at this point.

Next, a python script using CMS' FWLite library performs a significant translation. Until this point, all of the data is organized on a per-event basis, meaning there is a large list of events where each row of the list is the characteristics of a single event. The analysis doesn't need information about each individual event, it observes the aggregate effects of whole processes. The FWLite script handles that impedance mismatch. It iterates over each event and outputs a number of histograms describing the kinematic distributions of each bucket.

Finally, the few thousand FWLite files are combined and scaled according to the processes modeled to generate one template per systematic variation. These histograms become input to the fitter, parameters for the systematic calculator, and are used to extract the QCD contribution.

8.2 Fit-based Systematics

The following subsections describe each fit-based systematic in turn.

8.2.1 *b*-tagging

The *b*-tagging algorithm chosen for this analysis is the Combined Secondary Vertex (CSV)[7] tagger operating at the medium working point. In general, the efficiencies for successfully tagging or mistagging a *b*-jet vary between data and simulation, so a data to simulation SF has to be used to match the simulated performance with the actual performance. Instead of directly computing the *b*-tagging efficiency for data ($\varepsilon_{btag}^{data}$), it is more convenient to compute and use the *b*-tagging efficiency scale factor $SF_{btag} = \frac{\varepsilon_{btag}^{data}}{\varepsilon_{btag}^{MC}}$ since simulated events store the true quark composition of each jet. We can vary this ε_{btag}^{MC} and SF_{btag} by selectively weighing events depending on the flavor determined by the CSV tagger and the Monte Carlo truth.

CMS uses *b*-tagging for a wide variety of analyses, so a dedicated group both implements the algorithms and quantifies their performances. The *b*-tag Physics Object Group (POG) measured $SF_{btag} = 0.953 \pm 0.012$[10]. Since SF_{btag} was produced centrally with different event topologies and selections than this analysis, this calculated scale factor may not be appropriate. In particular, Instead of using the centrally produced SF_{btag}, the SHYFT method extracts SF_{btag} directly from the same events used to fit the central value.

Along with SF_{btag}, the performance of *b*-tag algorithms are characterized the mistag scale factor (SF_{lftag}), which represents how often light flavor jets are mistakenly tagged as a heavy flavor jet.

To model the data to MC SF, this analysis weighs each jet in simulation according to

$$W = \begin{cases} SF_{btag} & \text{Heavy flavor jet tagged during simulation} \\ 0 & \text{Heavy flavor jet not tagged during simulation} \\ SF_{lftag} & \text{Light flavor jet tagged during simulation} \\ 0 & \text{Light flavor jet not tagged during simulation} \end{cases} \tag{8.1}$$

The per-event weight is then the combinatoric probability of all of the jet weights in the event. Two groups of four sets of distributions are made, each scaling either SF_{btag} or SF_{lftag} up or down by ten or twenty percent to model the effect of these scale factors on the overall fit. These additional templates allow the fitter to effectively move jets back and forth between *b*-jet multiplicities and provide a data-driven estimate of these scale factors using the exact same event topologies, reducing the systematic uncertainty associated with the SFs. Effectively, making the MC templates match the data templates provides a data-driven estimate of the SF.

The resulting polynoids for SF_{btag} for the 2-Jet buckets can be found in Figure 8.1. The entire set of SF_{btag} and SF_{lftag} polynoids are in Appendix A.

8.2.2 Jet Energy Corrections (JEC)

This analysis depends heavily on accurate measurements of jet p_{T}. A number of calibrations, collectively known as Jet Energy Corrections are used to compensate for various non-linearities in both data and simulation. These corrections result in the scaling of each jet 4-momentum by a scale factor depending on the η and p_{T} of the jet. This scale factor is known as the Jet Energy Scale (JES).

If the JES is increased, low p_{T} jets which were previously under the selection threshold may gain enough momentum to be selected. Conversely, a lower JES my cause low p_{T} jets which marginally passed the selection threshold to fail. As the JES SF is varied up and down, the N_{jet} distribution of the SHYFT fit varies as well. We generate MC templates by varying JEC by ±5% and producing polynoids to characterize the effect of JES on the N_{jet} normalizations [23]. The resulting polynoids can be seen in Figure A.11

Along with the affecting the normalization, varying the JEC causes the shapes of the MC templates to vary as well. To account for this effect, the SHYFT fitter additionally interpolates the shapes between

8.3 Fit Constraints

The fitter extracts the process normalizations and systematic effect parameters that cause the simulation to most closely match the data. Without any external information, the fit could conceivably converge into a region of parameter space that is mathematically consistent but physically nonsensical. As long as the output from the model matches the data, the fitter is free to choose any values for the input parameters. For instance, nothing prevents the fit from choosing negative event counts or efficiencies greater than 100%.

SHYFT attempts to make processes with similar effects distinct by dividing the total event content into a number of buckets. This process isn't complete, meaning some processes still resemble one another. If two parameters X and Y can both cause simulation to more accurately match the data, the fitter is in a difficult position.

At their cores, these problems stem from the model having too little information. Until this point, the model treated its input parameters strictly as a set of floating-point numbers. An outside observer, however has *a priori* knowledge of their allowed ranges and *a posteriori* knowledge of their predicted values. Providing the model with this knowledge will

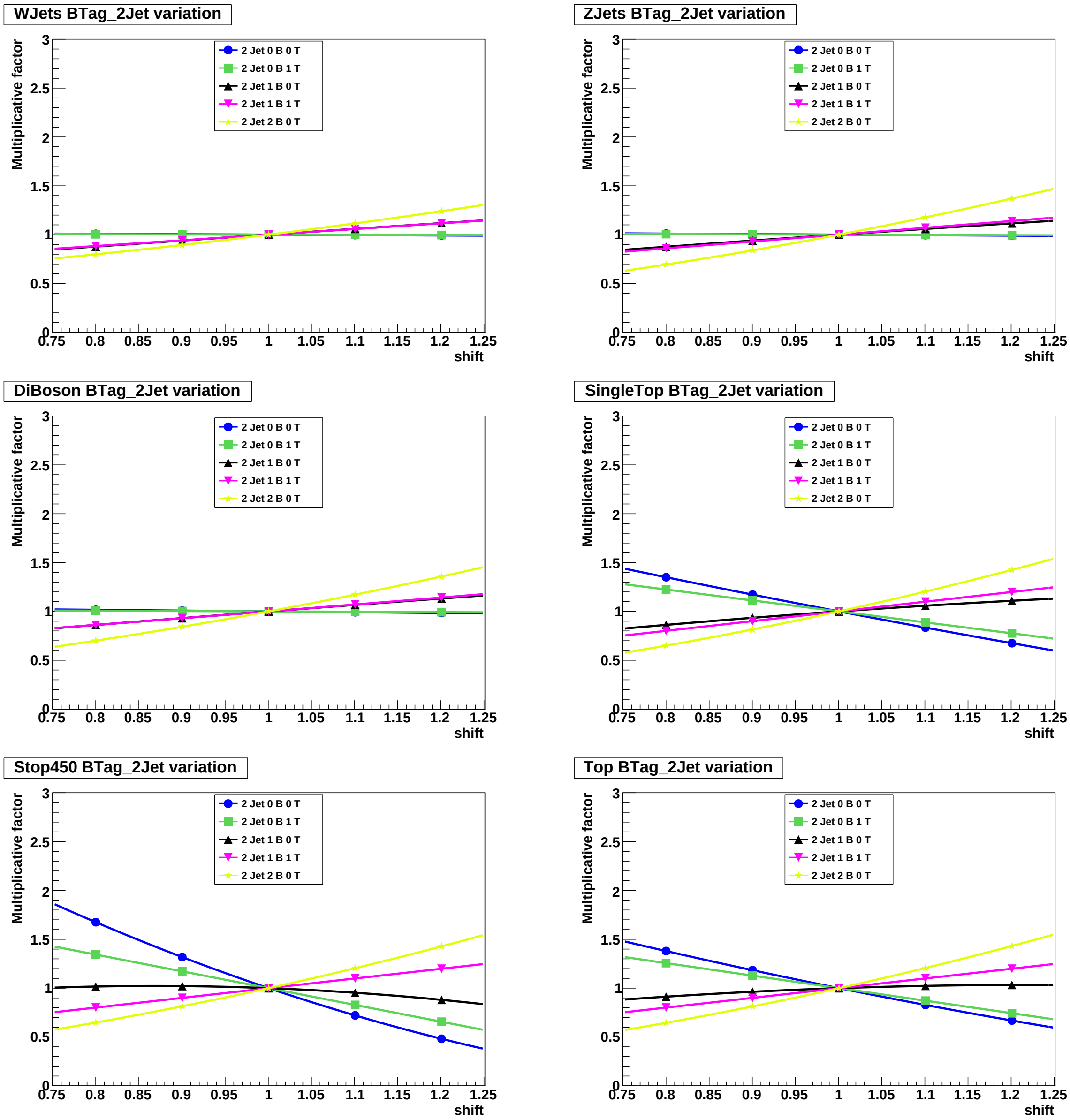


Figure 8.1: Polynoids Characterizing Effect of *b*-tagging on 2-jet multiplicity events. Each shift of 0.1 corresponds to a 10% change in SF_{btag}.

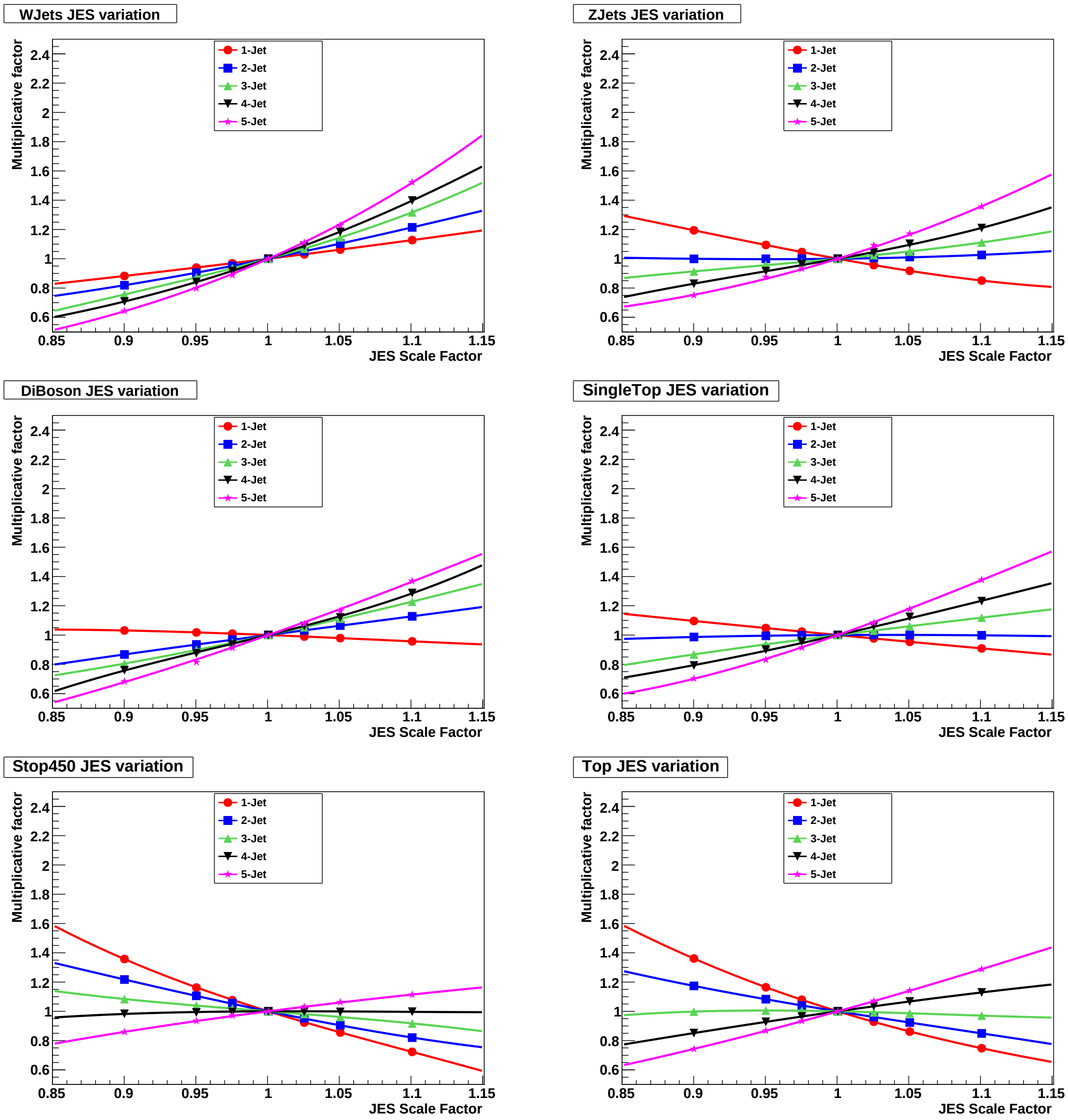


Figure 8.2: Polynoids Characterizing Effect of JES. Each shift of 0.05 corresponds to a 5% change in JES.

guide its results to physical realms of parameter space.

It would be tautological to simply fix all of the parameters to their initial values, then extract those values back out from the fit. Instead, we augment the likelihood function near the initial values. Without this modification, the likelihood depends only on how well the model matched the data, causing both unphysical and physical regions to have similar likelihoods. The augmented likelihood superimposes a peak in parameter space around the initial values. This new likelihood nudges the fit towards the initial values, where the central values and steepness of the peak are determined externally.

Using this new likelihood, fit parameters are allowed to vary around their initial values. Per-process scale factors (K_x) of 1.0 represents the theoretical event yield of sample x, while $R_{x,y}$ of 0.0 represents no effect from systematic y on sample x (See Equation 7.7). For each parameter $K_x/R_{x,y}$ we wish to constrain (z_l), we extract a data-driven estimate of the initial value ($\overline{z_l}$) and that measurement's uncertainty (σ_l). Using $\overline{z_l}$ and σ_l, we add an additional multiplicative term

$$e^{\frac{-(z_l-\overline{z_l})^2}{\sigma_l^2}} \tag{8.2}$$

This gaussian curve peaks at $\overline{z_l}$ with a width of σ_l. This width, which is the uncertainty of our measurement, influences how tightly the around the initial value the likelihood will match. A more precise measurement results in a small uncertainty, which makes a smaller value of σ_l, which causes this peak to be narrow. As the fit travels away from the initial value, this term influences the likelihood to decrease.

This analysis applies constraints for each background and systematic parameter (Table 8.1). These constraints use analysis-provided data-driven measurements where possible. Otherwise, existing published results are used. These constraints are described in detail below.

Parameter	$\overline{z_l}$	σ_l
$K_{t\bar{t}}$	1.0	0.20
K_{W+Jets}	1.0	0.10
K_{Z+Jets}	1.0488	0.046
$K_{SingleTop}$	1.7608	0.20
$K_{DiBoson}$	1.0	0.264
$K_{1JetQCD}$	1.0	2.0
$K_{2JetQCD}$	1.0	2.0
$K_{3JetQCD}$	1.0	2.0
$K_{4JetQCD}$	1.0	2.0
$K_{5JetQCD}$	1.0	2.0
$R_{x,btag}$	100.0%	10.0%
$R_{x,mistag}$	100.0%	10.0%
$R_{x,JES}$	100.0%	5.0%

Table 8.1: Fit Constraints

8.3.1 Z + jets and DiBoson (WW, ZZ, WZ) Constraints

Z + jets and DiBoson events which pass this analysis' one muon event selection criteria result from muon misidentification. Ten percent of the time, Z bosons decay into a pair of muons. Similarly, many DiBoson events with muons in the final state produce more than one muon. CMS reconstructs muons with a high efficiency, meaning most of these muons are properly identified, causing many of these events to be rejected.

Errors in detectors or event reconstruction can cause events which would normally rejected to pass the event criteria. If a Z boson decays into two muons, it has no true $\not{E}_{\mathrm{T}}$. Both the number of muons and the lack of $\not{E}_{\mathrm{T}}$ cause the event to fail event selection. However, if one of the muons isn't reconstructed, the E_{T} of that muon would be added to $\not{E}_{\mathrm{T}}$. The event would then have exactly one muon and possibly enough $\not{E}_{\mathrm{T}}$ to be accepted.

The similarities between Z + jets and DiBoson events cause their kinematic and N_jet distributions to resemble one another. As described before, these similarities cause K_{Z+Jets} and $K_{DiBoson}$ to be correlated, which poses difficulties for the fit. Z + jets events occur approximately an order of magnitude more often than DiBoson events. Without any constraints, even small changes in K_{Z+Jets} cause huge swings in $K_{DiBoson}$. This is because K_x is the scale factor between the theoretical and fitted event yields. Since Z + jets occurs much more often, every unit change in K_{Z+Jets} adds many more events than the same change to $K_{DiBoson}$.

We obtain $\overline{z_l}$ and σ_l using samples with either two or three muons. The two muon sample is comprised of roughly 99/1% Z + jets/DiBoson. The three muon sample is 40/60% Z + jets/DiBoson. Then we extract the di-muon invariant mass kinematic distribution for each sample, not separating into buckets based on final state. If there are more than two muons, we select the pair whose invariant mass is closest to the Z boson mass.

Finally, we perform a simplified 'mini-SHyFT' fit of both distributions simultaneously and extract K_{Z+Jets}, $K_{DiBoson}$ and corresponding uncertainties. These values are then used to constrain the contribution of Z + jets and DiBoson in the main fit.

8.3.2 QCD Constraints

Estimating the contribution from QCD processes poses difficulties. Section 6.7 describes the data-driven method used to estimate these contributions. We use this estimate as the initial value in our constraint. The estimate is performed with wildly different event selections and uses the shape of Z + jets as a proxy for the QCD shape. Instead of a single constraint for the entire QCD contribution, we apply a constraint per jet-multiplicity. This takes into account the effect of using pretagged shapes to model the QCD contribution. To be conservative, we set the constraint width to 200% of the uncertainty of the QCD measurement.

8.3.3 Other Constraints

The remaining background event yields are scaled to the most recent CMS-measured 8 TeV cross sections as their initial values ($\overline{z_l} = 1.0$). The constraint widths are chosen to be two times the total CMS-reported uncertainty on the measurement.

8.4 Summary of Method

To estimate the event rates of our signal and backgrounds, the SHYFT fitter exploits a finely-segmented, detailed description of as many properties of the events as possible. First, a number of templates are produced, where each template breaks down a single systematic situation into several hundred kinematic shapes. There is a shape for each (jet,b-tag,τ) final state, kinematic distribution, and data/simulated process. Next, the QCD contribution is estimated using the low mT sideband. Then, each group of templates describing a systematic effect is combined to produce a polynoid, which is used by the fitter to model these effects. Finally, the contribution from Z + jets and DiBosons are constrained using a data-driven multi-lepton sideband measurement.

With these inputs, the fit attempts to simultaneously find all the simulated event rates and systematic parameters which make the modeled $(N_j, N_b, N_\tau, Kinematic)$ distribution most closely resemble the data. The results of this fit and this analysis are described in the following chapter.

Chapter 9

Results and Analysis

This analysis searches for new physics in the muon+jets+$\not{E}_{\mathrm{T}}$ +τ channel resulting from the decay chain

$$\widetilde{t}\widetilde{t}^* \to b\bar{b} + \widetilde{\chi}^+\widetilde{\chi}^- \to b\bar{b} + \nu\bar{\nu} + \widetilde{\tau}\widetilde{\tau}^* \to b\bar{b} + \nu\bar{\nu} + \tau\bar{\tau} + \widetilde{\chi}^0\widetilde{\chi}^0 \tag{9.1}$$

where the $\tau\bar{\tau}$ decay into exactly one muon and no electrons.

We explore seven different $\widetilde{t}\,\widetilde{t}^*$ scenarios with different masses of the SUSY partners. These different scenarios simulate the above cascade decay of $\widetilde{t}$-pairs with a mass of 250, 300, 350, 400, 450, 500, or 600 GeV. A full description of the different mass points can be found in Table 6.3. To investigate these seven scenarios, the SHYFT fit is performed seven different times. Each mass point is fit independently, sharing only the input templates extracted from data or MC. In the subsequent sections, simplified results are presented for all seven points and detailed information for the 450 GeV mass point are provided.

The data are divided into a number of buckets based on their reconstructed jet, *b*-jet, and τ content. These buckets are simultaneously fit by a joint likelihood, which is described in Chapter 7

The SHYFT fit finds no significant deviation from the SM, so the CMS-developed Higgs combination tool is used to extract 95% Confidence Limits (CLs) on the signal strength. We describe the systematic effects, results from the SHYFT fit, and our expected limits below.

9.1 Systematic effects

This analysis uses three different methods to quantify systematic effects and their associated uncertainties.

- Modeling effects within the fit
- Pseudoexperiments
- Fixed penalties

Effects which can be parameterized and modeled as a relative scaling of the kinematic shapes are referred to as ‘fit-based‘ effects by this analysis. The *b*-tagging scale factor, light flavor tagging scale factor and jet energy scale factor are handled in this manner. These effects are exposed to the fitter as additional parameters it can vary to find a more accurate match between simulation and data.

To do this, we first generate templates with these systematic parameters shifted up and down. Using the these templates, we produce polynomials which relate the scaling of each bucket to the shift of each parameter. Finally, the fitter is taught how to manipulate these parameters to affect the relative normalizations of each bucket. This technique is described in detail in Section 7.4

In a log-likelihood fit, the width of the likelihood curve at the minimum is used to estimate the statistical uncertainty of the fit. Each fit-based systematic modeled in the fitter contributes to a further broadening of the likelihood at the minimum. As a result, the uncertainties estimated by the fitter are no longer simply the statistical uncertainty. These uncertainties become a convolution of the statistical and systematic uncertainties.

This combined statistical and systematic uncertainty makes quantifying the systematic uncertainties associated with each systematic effect difficult. To compare with previous estimates, we perform the fit twice. The first time, we fit normally to get the preferred scales. Then we perform the same fit again with the one systematic fixed to the values found in the first step. To estimate the statistical uncertainty neglecting the effect of fit-based systematics, we perform one last fit where all three fit-based parameters are fixed to the values extracted from the first fit. These two statistical uncertainties subtracted in quadrature provide an estimate of the contribution of each uncertainty. Since these effects are intermingled within the likelihood function, each of their contributions can not rigorously be extracted. However, the estimates are useful to for comparisons with existing results.

Some systematic effects can be quantified using pseudo experiments. We again produce templates with parameters shifted up and down, but instead of introducing an additional fit parameter, we generate toy MC from these alternate templates and fit them to the data. The difference between the measured and generated event rates are used to estimate the effect of the systematic effect. We quote the largest of the up/down variation or their average as the final systematic uncertainty.

Finally, the remaining systematic effects assess fixed penalties to the systematic uncertainty. These uncertainties result from the lepton trigger, identification and isolation data to MC scale factors, as well as the global uncertainty on the delivered LHC luminosity.

Each considered uncertainty is described in detail below.

9.1.1 B-Tag / Light Flavor-Tag Scale Factor

The performance of b-tagging algorithms with both MC and data are reasonably good. The agreement between MC and data b-tagging is known as the b-tag Scale Factor (SF_b). The b-tag Physics Object Group (POG) produces SF_b for several different scenarios, but since this analysis extracts SF_b directly from the fit, we choose not to use it. Instead, we generate templates with SF_b shifted $\pm$ $1/2\sigma$. These templates are produced by examining the Monte Carlo truth and b-tag status of each jet and weighing them by $\pm 10\%$ or $\pm 20\%$. The total event weight is the combinatorial probability to tag

In addition to the b-tag Scale Factor, we consider the rate of light flavor quarks being mis-tagged as b-quarks (SF_{lf}). We consider this effect in a similar manner to SF_b.

9.1.2 Jet Energy Corrections/Jet Energy Scale

The high luminosity and center of mass energy of the LHC will provide incredible opportunities for both furthering our understanding of the SM and probing the existence of possible new physics beyond the standard model. To achieve this, however, the LHC collides hadrons (protons) instead of leptons (e.g. electrons). The combined effects of high luminosity and colliding colored particles results an incredibly large amount of hadronic activity within the detector (Chapter 2 describes in more detail the consequences).

Consequently, much care has been taken in CMS' ability to accurately detect and reconstruct jets which pass through the detector. On the hardware side, the finely-segmented electromagnetic calorimeters (ECAL) and hadronic calorimeters (HCAL) accurately measure the energies deposited by particles in their detector media. The software then takes these energy deposits and known detector performance and attempts to reconstruct the trajectory and energy of the mother parton.

In order to faithfully reconstruct jets, a number of calibrations must be made in software. The raw ECAL and HCAL measurements themselves are influenced by a number of sources of error. For instance, during collisions the detector is subjected to an enormous flux of radiation. This radiation induces defects in the lattice structure of the crystals within the electromagnetic calorimeter. Over time, these defects cause the crystals to become progressively more opaque, causing less light to enter the photomultipliers, leading to less detected energy for the same particles.

In addition to the corrections relating to the physical detector's performance, known differences between simulated and real detector responses must be factored out. The effect of transiting particles on detector and readout hardware are modeled via GEANT4 [31]. Though the model is quite accurate, a small residual correction is applied to simulation to compensate for known inconsistencies.

These corrections are collectively known as Jet Energy Corrections (JEC). These corrections modify the energy of each jet and consequently the $\not\!\!E_{\mathrm{T}}$ of each event, the resulting scaling is known as the Jet Energy Scale (JES). Different amounts of JES will change the jet multiplicity spectrum. For example, a jet previously just under the p_{T} threshold could to meet the p_{T} requirement once its corresponding correction is scaled up. Alternately, the same event reconstructed using different JES could cause the $\not\!\!E_{\mathrm{T}}$ of the event to cross the $\not\!\!E_{\mathrm{T}} > 20\,\mathrm{GeV}$ cut required by this analysis. Since the data is left unchanged while the simulated samples' JES is allowed to vary, the fitter is effectively allowed to find the value of JES which causes the jet multiplicity spectra for simulation and data to most closely agree.

The dependence on JES is modeled within the SHYFT fit by producing shapes with JEC shifted by $\pm 5\%$. Care is taken to offset the $\not\!\!E_{\mathrm{T}}$ after each jet is rescaled with the new JEC. Like the b-tagging SFs, the relative normalizations between buckets are input to the fitter to allow it to simulate the effect of JES on the simulation. However, since the fit uses the $\sum E_{\mathrm{T}}$ and $m_T(l, \not\!\!E_{\mathrm{T}})$ distributions, modifying the JES also changes the shapes. The fitter therefore additionally morphs the shapes as well.

9.1.3 W + jets and $t\bar{t}$ Q^2 Energy Scale

The Q^2 energy scale is the combination of two independent energy scales: the vertex energy scale and the renormalization and factorization energy scale. These energy scales, in turn, depend on either qfac or ktfac, which are scalar constants. The effect of Q^2 is to determine the 'jettiness' of an event. A decrease in Q^2 will lead to an enhanced amount of partons being radiated. To estimate the uncertainty due to Q^2, we generate alternate templates using the `scaleup`/`scaledown` samples, which generates events with qfac and ktfac simultaneously set to 0.5 and 2.0, respectively. We generate toy MC from these alternate templates and fit to the nominal templates to assess this uncertainty.

9.1.4 W + jets and $t\bar{t}$ Matrix Element to Particle Shower Matching

Samples generated with `MADGRAPH` [27] interfaced to `PYTHIA` [28] must perform Matrix Element (ME) to Particle Shower (PS) matching, which incurs an associated uncertainty. The matrix element and particle shower simulation techniques represent and evolve the simulated event using vastly different methods. The `MADGRAPH` generator calculates matrix elements of the hard interaction, while `PYTHIA` handles the showering of outgoing partons. Partons stop being evolved in `MADGRAPH` once the momentum of the parton drops below a threshold known as xqcut. Since `MADGRAPH` and `PYTHIA` model the partons differently, modifying xqcut changes the hadronic composition of the final state. To estimate the uncertainty due to Q^2, we generate alternate templates using the `matchingup`/`matchingdown` samples, which generate events with xqcut scaled up and down. We generate toy MC from these alternate templates and fit to the nominal templates to assess this uncertainty.

9.1.5 Muon Trigger/Identification/Isolation

Simulated events are corrected by a data to MC scale factor, which takes into account the difference in muon triggering, identification and isolation performance in MC and data. These SFs have a combined uncertainty of 2%[33]. These uncertainties are applied globally to the measured yields.

9.1.6 Luminosity

The total integrated luminosity of the analyzed data was produced using the CMS-provided `pixelLumiCalc.py` tool. This luminosity is used to scale MC to their theoretical event yields. The uncertainty associated with this value is 2.6%.

9.1.7 Complete Systematic Uncertainty

The statistical uncertainty quoted by the SHYFT fitter is a combination of the true statistical uncertainty and the uncertainties resulting from the addition of fit-based uncertainties to a fit. Performing multiple fits and subtracting the uncertainties in quadrature allows us to estimate their individual contributions to the overall fit. A summary of the uncertainties in the SHYFT fit are in Table 9.1.

	Relative uncertainty on $\tilde{t}\,\tilde{t}^{*}$ yield						
Source	250 GeV $\tilde{t}$	300 GeV $\tilde{t}$	350 GeV $\tilde{t}$	400 GeV $\tilde{t}$	450 GeV $\tilde{t}$	500 GeV $\tilde{t}$	600 GeV $\tilde{t}$
SF$_b$	6.2%	4.4%	2.6%	2.6%	2.0%	6.3%	8.1%
SF$_{lf}$	1.0%	1.3%	3.4%	0.9%	0.9%	0.5%	0.4%
SF$_{jes}$	6.2%	6.2%	6.0%	3.7%	2.4%	1.5%	1.7%
Statistical	11.7%	19.3%	17.7%	19.3%	21.5%	20.8%	21.2%
Total Unc.	13.3%	20.3%	18.7%	19.6%	21.7%	20.8%	21.3%

Table 9.1: Summary of SHYFT uncertainties. The individual uncertainties are estimated by fixing each uncertainty and subtracting in quadrature the resulting uncertainty from the uncertainty on a fit with all systematic parameters fixed. Correlations between the individual uncertainties cause their sum to differ from the total uncertainty, which is extracted directly from the fit.

9.2 Validation of SHYFT method

In order to validate the SHYFT method, I perform an additional fit in a region with no signal contribution. This region is produced by inverting the $\not\!\!E_{\mathrm{T}}$ requirement to require $\not\!\!E_{\mathrm{T}} < 20\,\mathrm{GeV}$. As expected (Table 9.2), this region both shows an insignificant contribution from signal and good ($\sim 99\%$) agreement between data and simulation.

Sample	Event Yield
WJets	1.87E+06 ±0.25%
Top	3.89E+04 ±1.24%
ZJets	1.37E+06 ±0.31%
SingleTop	3.99E+04 ±2.58%
DiBoson	1.80E+04 $(+1.09/-0.00)\%$
QCD	9.01E+04 ±3.83%
Total Bkg.	3.42E+06 (+0.52/–0.31)%
450 GeV $\tilde{t}\,\tilde{t}^{*}$	2.63E+01 (+3.49/–2.10)%
Total Sim.	3.42E+06 (+0.52/–0.31)%
Data	3.38E+06
SF$_{sim}$	0.988

Table 9.2: Results of SHYFT validation fit in the $\not\!\!E_{\mathrm{T}} < 20\,\mathrm{GeV}$ sideband, $m_{\tilde{t}} = 450\,\mathrm{GeV}$ scenario. The quoted uncertainties include the contributions from statistics, SF$_b$, SF$_{lf}$, and SF$_{jes}$.

9.3 Result of SHyFT fit

Using SHyFT, we perform a fit in our signal region which includes the effects of systematic parameters on our model. The fit factors are listed in Table 9.3. The final event yields, divided into individual jet/tag buckets are listed below in Table 9.4. Finally, the correlation matrix for this fit is located at Table 9.5. The resulting kinematic distributions are located at Figures 9.1, 9.2, and 9.3.

	250 GeV $\widetilde{t}$	300 GeV $\widetilde{t}$	350 GeV $\widetilde{t}$	400 GeV $\widetilde{t}$	450 GeV $\widetilde{t}$	500 GeV $\widetilde{t}$	600 GeV $\widetilde{t}$
$\widetilde{t}$	$1.28^{+0.17}_{-0.17}$	$0.70^{+0.15}_{-0.14}$	$0.85^{+0.16}_{-0.15}$	$0.90^{+0.18}_{-0.17}$	$1.03^{+0.23}_{-0.21}$	$1.47^{+0.32}_{-0.29}$	$3.09^{+0.68}_{-0.63}$
Top	$0.98^{+0.01}_{-0.01}$	$0.98^{+0.01}_{-0.01}$	$0.98^{+0.01}_{-0.01}$	$0.98^{+0.00}_{-0.00}$	$0.99^{+0.00}_{-0.00}$	$0.99^{+0.00}_{-0.00}$	$1.01^{+0.01}_{-0.01}$
WJets	$1.02^{+0.00}_{-0.00}$	$1.01^{+0.00}_{-0.00}$	$1.01^{+0.00}_{-0.00}$	$1.02^{+0.00}_{-0.00}$	$1.02^{+0.00}_{-0.00}$	$1.01^{+0.00}_{-0.00}$	$1.02^{+0.00}_{-0.00}$
ZJets	$1.07^{+0.00}_{-0.00}$	$1.08^{+0.00}_{-0.00}$	$1.08^{+0.00}_{-0.00}$	$1.08^{+0.00}_{-0.00}$	$1.08^{+0.00}_{-0.00}$	$1.08^{+0.00}_{-0.00}$	$1.08^{+0.00}_{-0.00}$
SingleTop	$2.61^{+0.03}_{-0.03}$	$2.62^{+0.03}_{-0.03}$	$2.62^{+0.03}_{-0.03}$	$2.61^{+0.03}_{-0.03}$	$2.61^{+0.03}_{-0.03}$	$2.64^{+0.03}_{-0.03}$	$2.57^{+0.03}_{-0.03}$
DiBoson	$1.96^{+0.00}_{-0.00}$	$1.96^{+0.00}_{-0.00}$	$1.96^{+0.00}_{-0.00}$	$1.96^{+0.00}_{-0.00}$	$1.96^{+0.00}_{-0.00}$	$1.96^{+0.00}_{-0.00}$	$1.96^{+0.00}_{-0.00}$
btag	$0.98^{+0.03}_{-0.03}$	$1.04^{+0.04}_{-0.05}$	$1.13^{+0.02}_{-0.03}$	$1.06^{+0.03}_{-0.03}$	$1.07^{+0.03}_{-0.03}$	$0.94^{+0.03}_{-0.03}$	$1.00^{+0.02}_{-0.02}$
lftag	$1.60^{+0.02}_{-0.02}$	$1.55^{+0.02}_{-0.03}$	$1.45^{+0.05}_{-0.05}$	$1.54^{+0.03}_{-0.03}$	$1.54^{+0.03}_{-0.03}$	$1.57^{+0.02}_{-0.02}$	$1.66^{+0.01}_{-0.01}$
jes	$0.96^{+0.00}_{-0.00}$	$0.96^{+0.00}_{-0.00}$	$0.96^{+0.00}_{-0.00}$	$0.96^{+0.00}_{-0.00}$	$0.96^{+0.00}_{-0.00}$	$0.96^{+0.00}_{-0.00}$	$0.96^{+0.00}_{-0.00}$
qcd_1j	$1.70^{+0.04}_{-0.04}$	$1.71^{+0.04}_{-0.04}$	$1.70^{+0.04}_{-0.04}$	$1.70^{+0.04}_{-0.04}$	$1.70^{+0.04}_{-0.04}$	$1.71^{+0.04}_{-0.04}$	$1.71^{+0.04}_{-0.04}$
qcd_2j	$0.63^{+0.02}_{-0.02}$	$0.62^{+0.03}_{-0.03}$	$0.65^{+0.02}_{-0.02}$	$0.64^{+0.02}_{-0.02}$	$0.64^{+0.02}_{-0.02}$	$0.62^{+0.02}_{-0.02}$	$0.64^{+0.02}_{-0.02}$
qcd_3j	$1.33^{+0.05}_{-0.05}$	$1.26^{+0.05}_{-0.05}$	$1.29^{+0.05}_{-0.05}$	$1.29^{+0.05}_{-0.05}$	$1.28^{+0.05}_{-0.05}$	$1.25^{+0.05}_{-0.05}$	$1.35^{+0.05}_{-0.05}$
qcd_4j	$1.45^{+0.08}_{-0.08}$	$1.56^{+0.09}_{-0.09}$	$1.63^{+0.08}_{-0.08}$	$1.59^{+0.08}_{-0.08}$	$1.59^{+0.08}_{-0.08}$	$1.54^{+0.08}_{-0.08}$	$1.52^{+0.08}_{-0.08}$
qcd_5j	$1.31^{+0.09}_{-0.09}$	$1.33^{+0.09}_{-0.09}$	$1.35^{+0.09}_{-0.09}$	$1.33^{+0.09}_{-0.09}$	$1.34^{+0.09}_{-0.09}$	$1.31^{+0.09}_{-0.09}$	$1.32^{+0.09}_{-0.09}$

Table 9.3: Fit factors from the nominal SHyFT fit. The quoted uncertainties are a combination of both the statistical and fit-based systematic uncertainties, described in Sec 9.1. The btag, lftag, and jes parameters are defined as percent shift from the nominal, while the other parameters are expressed as multiples of the theoretical yield. The QCD factors are the multiplicative difference relative to the sideband-fitted QCD rates

.

	Data	Total Pred	Stop450	Top	WJets	ZJets	QCD	DiBoson	SingleTop
1 Jet 0 B 0 T	4872103	4862819.2	3.1	12596.8	4437444.2	333104.4	0.0	44309.4	35361.4
1 Jet 0 B 1 T	283007	291574.2	0.3	662.8	145707.5	127738.7	10221.3	6222.1	1021.6
1 Jet 1 B 0 T	237451	225592.1	5.3	12665.3	148959.5	12025.4	10840.4	2510.3	38586.0
2 Jet 0 B 0 T	989927	997544.0	10.4	19495.5	871915.3	64174.7	4509.4	17448.0	19990.9
2 Jet 0 B 1 T	65936	64553.9	2.2	1417.2	38368.0	19493.9	1517.0	2491.7	1263.9
2 Jet 1 B 0 T	130721	134155.3	33.1	35936.7	55714.8	4796.2	1154.6	2002.2	34517.7
2 Jet 1 B 1 T	5404	5712.2	4.8	1969.9	1258.2	801.3	89.1	80.9	1507.9
2 Jet 2 B 0 T	18428	20933.2	17.1	11462.4	2527.2	227.4	114.4	44.7	6540.1
3 Jet 0 B 0 T	187590	186489.6	14.0	15313.8	143457.3	12903.0	4665.3	4136.5	5999.8
3 Jet 0 B 1 T	15810	15059.0	4.3	1306.0	7145.9	4438.7	961.7	687.2	515.2
3 Jet 1 B 0 T	66921	69633.1	81.0	39278.2	13558.6	1338.8	1814.6	681.0	12880.9
3 Jet 1 B 1 T	4857	4943.1	20.0	2956.3	511.2	305.0	277.7	49.2	823.8
3 Jet 2 B 0 T	28068	27969.4	20.1	21768.2	952.4	176.4	163.1	44.2	4845.0
3 Jet 2 B 1 T	1596	1727.5	14.4	1438.4	25.2	26.5	18.4	7.3	197.2
4 Jet 0 B 0 T	36834	35957.3	11.4	6676.0	23986.7	2719.5	376.7	803.0	1383.9
4 Jet 0 B 1 T	3604	3308.2	6.8	550.0	1297.8	912.7	256.4	154.3	130.1
4 Jet 1 B 0 T	27983	27782.8	85.5	20224.7	2971.5	374.6	405.2	182.2	3539.2
4 Jet 1 B 1 T	2390	2298.6	23.3	1669.4	106.9	84.0	146.9	12.3	255.8
4 Jet 2 B 0 T	18878	18147.3	15.8	15591.9	447.3	82.3	0.0	17.3	1992.7
4 Jet 2 B 1 T	1282	1221.3	21.6	1070.7	10.1	8.7	0.0	1.0	109.2
5 Jet 0 B 0 T	8461	7776.3	8.7	2622.3	4031.6	503.1	106.2	159.5	344.9
5 Jet 0 B 1 T	895	800.2	3.4	241.9	217.6	206.3	65.3	36.8	28.9
5 Jet 1 B 0 T	10707	10426.5	76.3	8660.1	676.0	128.2	0.0	40.1	846.0
5 Jet 1 B 1 T	988	906.8	21.4	704.0	26.5	23.8	55.5	4.5	71.1
5 Jet 2 B 0 T	9188	8732.1	11.0	7984.1	80.3	14.9	13.0	9.4	619.5
5 Jet 2 B 1 T	729	695.8	23.4	600.8	5.0	3.5	14.8	1.9	46.5
Total	7029758	7026759.5	538.6	244863.5	5901402.6	586612.1	37786.9	82136.7	173419.0

Table 9.4: Per-bin event yields SHYFT fit in the $m_{\tilde{t}} = 450\,\mathrm{GeV}$ scenario.

	Stop450	Top	WJets	ZJets	DiBoson	SingleTop	btag	jes	lftag	qcd_1j	qcd_2j	qcd_3j	qcd_4j	qcd_5j
Stop450	1.000	-0.181	0.004	0.006	0.000	0.060	0.052	-0.008	-0.098	0.023	0.013	0.030	-0.047	-0.090
Top	-0.181	1.000	0.463	-0.214	0.000	-0.658	-0.147	-0.445	0.274	-0.195	0.240	-0.018	-0.073	-0.003
WJets	0.004	0.463	1.000	-0.389	0.000	-0.381	-0.024	-0.951	0.093	-0.015	0.398	0.346	0.102	0.077
ZJets	0.006	-0.214	-0.389	1.000	0.000	0.242	0.108	0.267	-0.098	0.384	-0.282	-0.164	-0.029	-0.014
DiBoson	0.000	0.000	0.000	0.000	1.000	0.000	-0.000	0.000	0.000	-0.000	0.000	0.000	0.000	-0.000
SingleTop	0.060	-0.658	-0.381	0.242	0.000	1.000	-0.178	0.322	0.002	0.364	-0.452	-0.172	-0.036	-0.038
btag	0.052	-0.147	-0.024	0.108	-0.000	-0.178	1.000	0.017	-0.796	0.017	0.210	0.116	0.183	0.126
jes	-0.008	-0.445	-0.951	0.267	0.000	0.322	0.017	1.000	-0.086	-0.036	-0.436	-0.378	-0.111	-0.084
lftag	-0.098	0.274	0.093	-0.098	0.000	0.002	-0.796	-0.086	1.000	-0.038	-0.117	-0.128	-0.166	-0.088
qcd_1j	0.023	-0.195	-0.015	0.384	-0.000	0.364	0.017	-0.036	-0.038	1.000	-0.172	-0.040	0.008	0.007
qcd_2j	0.013	0.240	0.398	-0.282	0.000	-0.452	0.210	-0.436	-0.117	-0.172	1.000	0.270	0.101	0.066
qcd_3j	0.030	-0.018	0.346	-0.164	0.000	-0.172	0.116	-0.378	-0.128	-0.040	0.270	1.000	0.103	0.056
qcd_4j	-0.047	-0.073	0.102	-0.029	0.000	-0.036	0.183	-0.111	-0.166	0.008	0.101	0.103	1.000	0.045
qcd_5j	-0.090	-0.003	0.077	-0.014	-0.000	-0.038	0.126	-0.084	-0.088	0.007	0.066	0.056	0.045	1.000

Table 9.5: Correlation matrix from SHYFT fit in the $m_{\tilde{t}} = 450\,\text{GeV}$ scenario

Taking into account the fitted event yield and associated uncertainties, Table 9.6 lists $N_{\widetilde{t}\widetilde{t}^*}$ for each of our mass points. Even without considering the full gamut of systematic effects, the measured signal is consistent with the backgrounds. In order to effectively extract and quantify the significance of this measurement, I use a statistical packaged originally developed by CMS during the search for the Higgs boson.

$\widetilde{t}$ Mass (GeV)	Data Yield	Signal Yield	Background Yield
250	7029758	6709.66 ± 869.53	7020209.27 ± 10863.11
300	7029758	2210.53 ± 440.14	7024551.53 ± 10357.22
350	7029758	1460.88 ± 278.14	7025240.23 ± 10301.82
400	7029758	841.08 ± 165.11	7025922.26 ± 10277.80
450	7029758	538.63 ± 116.67	7026220.83 ± 10267.14
500	7029758	419.36 ± 87.28	7026428.27 ± 10286.74
600	7029758	271.03 ± 57.59	7026579.24 ± 10795.59

Table 9.6: Measured yields from SHYFT fit. The quoted uncertainties are a convolution of stat, *b*-tagging, and JES.

9.4 The Higgs Combination Tool

Since its prediction in the 1960s, the Higgs boson eluded detection by generation after generation of detectors and physicists. The mass of the Higgs was initially very weakly constrained, so performing a search for the Higgs was made more difficult by not knowing ‘where’ to look. Hopes were raised when tantalizing glimpses of this mysterious particle were observed, but for decades these glimpses were all determined to be false. Instead of triumphant discoveries, search after search yielded exclusions on the Higgs. Towards the end of the ’00s, however, the search for the Higgs reached a frenzied pitch.

The Tevatron, a proton-antiproton collider operating at $\sqrt{s} = 1.96\,\text{TeV}$ at Fermi National Accelerator Laboratory (FNAL) in Chicago, IL, was reaching the end of it’s lifespan. Originally operational since 1984, the Tevatron was credited with several groundbreaking discoveries including the discovery of the top quark. It’s collision energies were limited to 1.96 TeV, but a series of upgrades to it’s instantaneous luminosity in the early ’00s helped expand the reach of both the Tevatron and it’s two detectors - the Collider Detector at Fermilab (CDF) and D-Zero (DØ). After nearly two decades of operation, the Tevatron experiments had figured out how to wring out every last bit of sensitivity out of their machines. The accelerator itself was operating at one hundred times it’s design luminosity and the detector groups had honed their reconstruction efficiency using their experience of several hundred trillion events. Even more tantalizing, years of successive exclusions had narrowed down the Higgs boson mass to a range that was just on the edge of what was detectable at the Tevatron. Towards the late ’00s, there was growing optimism that the Tevatron would be the one to make this discovery.

On the other side of the world at CERN, the LHC and the scientists affiliated with it’s detectors were biding their time. A series of issues caused the startup date to push back

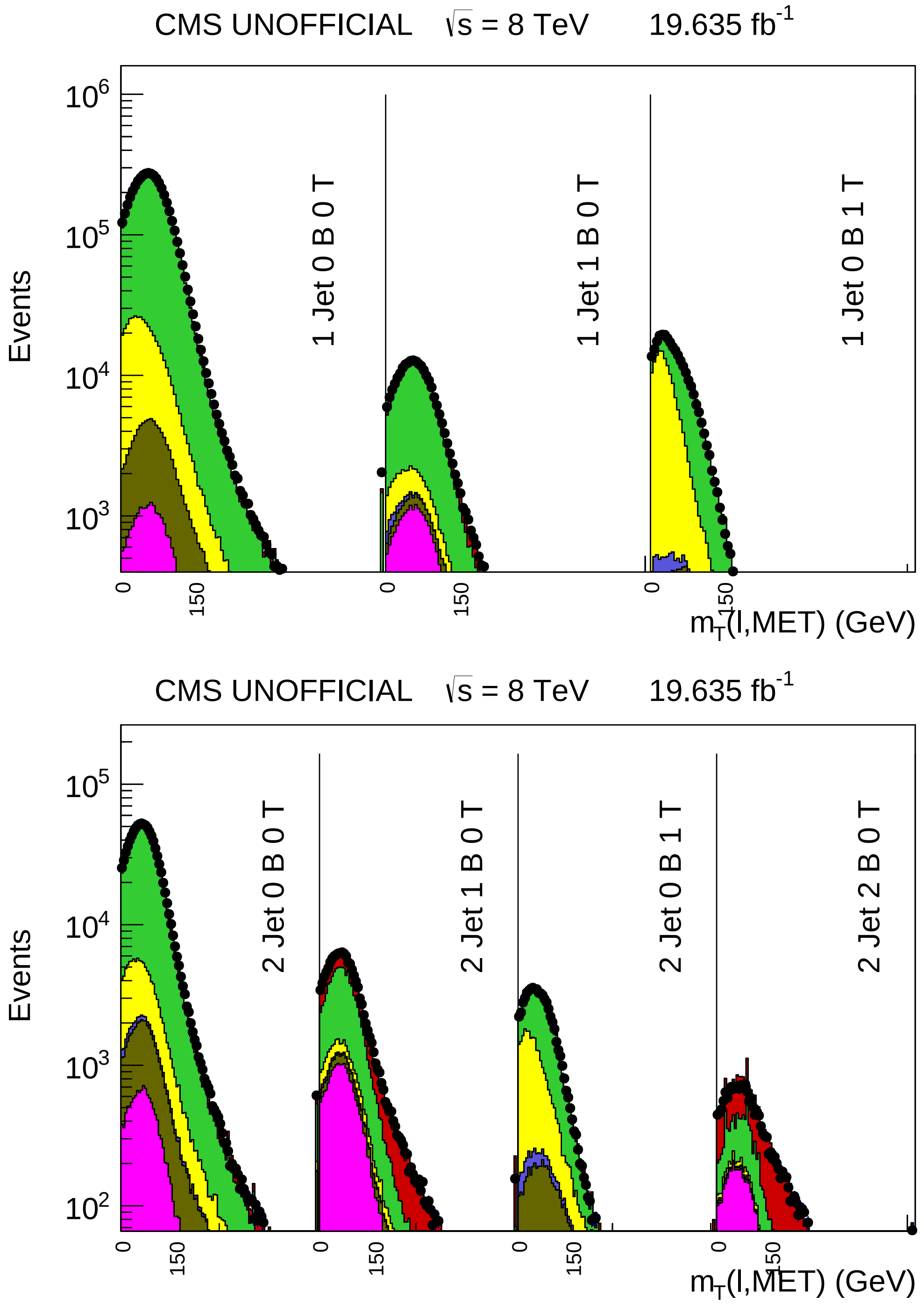


Figure 9.1: Output distributions for 1/2/3 jet bins(log scale)

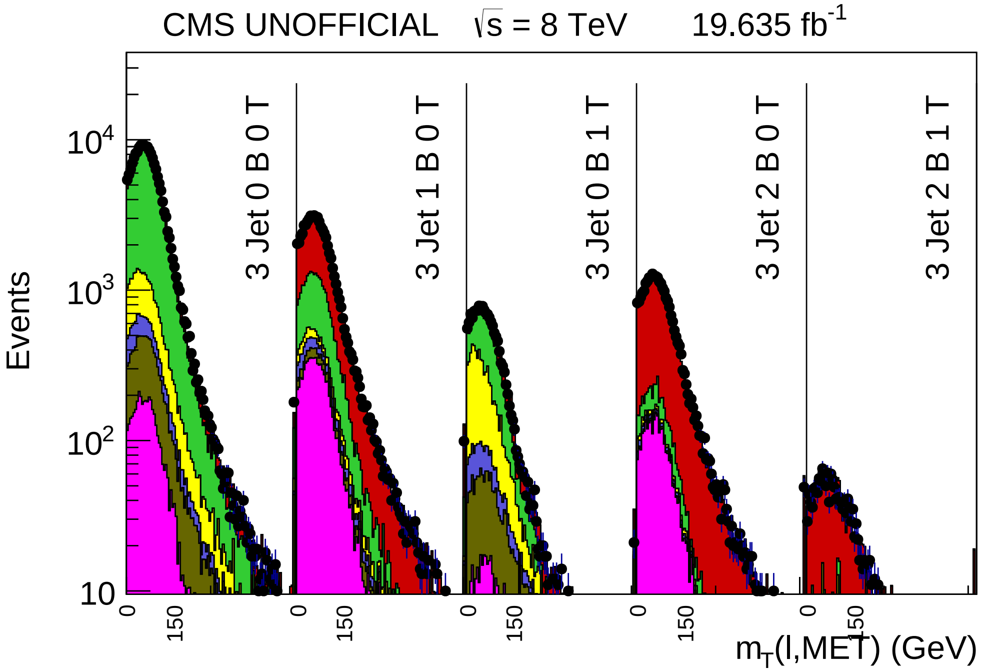


Figure 9.2: Output distributions for 3 jet bins(log scale)

from 2005 to 2009. The relative advantages of the LHC over the Tevatron couldn't be exploited until the accelerator was running and producing usable data.

The LHC began operations at $\sqrt{s} = 7\,\text{TeV}$ in the beginning of 2010 and the hunt began. The higher energies meant each collision had a higher probability of producing massive particles, but the initial periods of data taking were limited while the accelerator and detectors were undergoing tuning during the commissioning runs. Back at the Tevatron, momentum was building for a last-ditch extension to the project. Glimpses of the Higgs were reportedly seen[34], and the hope was that another three years of running would allow the Tevatron to collect enough data to turn these glimpses into a statistically significant discovery.

Even for an experiment as large and complex as CMS, the Higgs search was a massive undertaking[4]. Hundreds of scientists were directly involved with examining what little data the detector had recorded. This workforce was divided up into many groups, each tasked with examining a specific final state (e.g. Higgs decaying into two photons). Individually, these results wouldn't have enough statistics to distinguish the Higgs from the backgrounds. Simply waiting for more data could've meant ceding the potential Higgs discovery to the Tevatron. Instead, combined results which included the results from each group were produced. These combined result, however, could be significant enough to claim a discovery.

In order to combine these disparate results in a statistically rigorous way, the Higgs combination tool was developed. Guided by statistics experts within the collaboration, this tool quickly became a well-tested, fully-featured framework for calculating the significance of the measurement. Results from ATLAS (another LHC detector) would later be combined

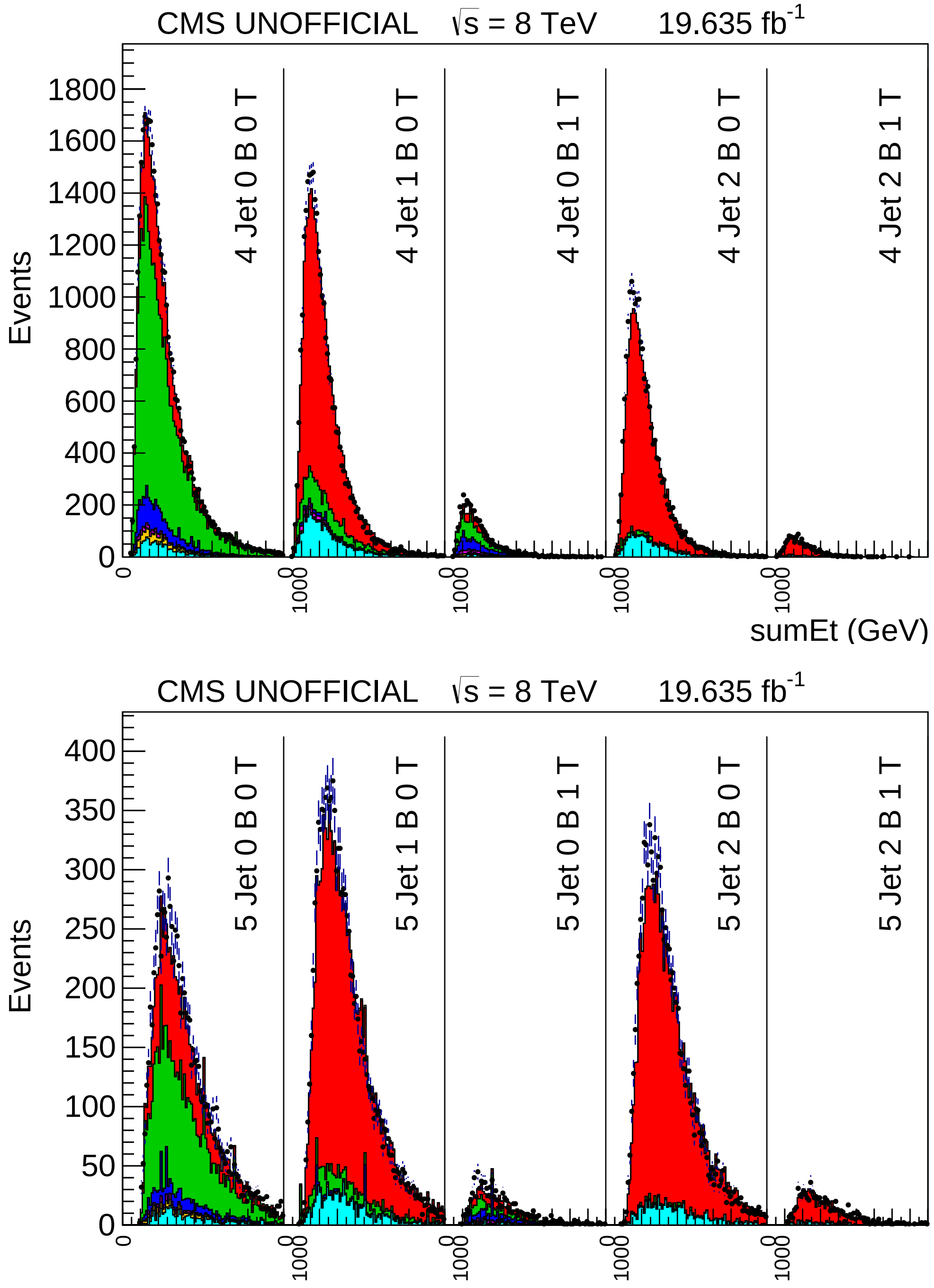


Figure 9.3: Output distributions for 4/5 jet bins

with CMS' measurements using the combination tool as well.

The combination tool is still in use, but instead of being primarily a tool for the Higgs search, many new users are using it to search for new physics like SUSY. Its roots as a tool to compute the significance of a faint signal against a large background makes it apt for these measurements. Having a highly functional and well-developed tool with many experts within the collaboration has been a boon, lowering the barrier for new analyses to be performed on ever-fainter signals.

9.5 Limits

Fitting simulation with the 2012 data using the SHYFT fitter results in a signal yield which is completely within the uncertainty of the background yield (Table 9.6). This analysis seeks to compute the significance of the observed signal yield in order to set 95% C.L. limits for these SUSY scenarios. I use the Higgs Combination Tool[35] to perform this task.

The Higgs Combination Tool calculates the exclusion limit with the CL_s[36, 37] method. The CL_s method constructs a joint likelihood for each bin in the signal, data, and backgrounds. Systematic uncertainties are represented by nuisance parameters which contribute to the joint likelihood. Parameters which affect normalization are modeled as log normals, while parameters which affect shapes are modeled as gaussians. Nuisance parameters are considered either 100% uncorrelated or 100% correlated across processes by default. Implementing partially correlated nuisance parameters requires changing the basis of the uncertainties to a coordinate system where each parameter is linearly independent from one another. The 95% upper limit is then determined by locating the signal yield θ where the integral of the likelihood from $-\infty$ to θ is 95% of the total likelihood. In the case of a gaussian likelihood function, this corresponds to an upward deviation of 1.96σ from the mean.

The combination tool accepts as it's input a datacard, which expresses the analysis in terms of input shapes, uncertainties and the correlations between them. The datacard has pointers to shapes within ROOT files. The combination tool takes the datacard and ROOT files, and extracts the expected and observed significance of the signal. It begins by repeatedly generating pseudodata from simulation, fitting it to the simulation, then recording the signal yield. After several iterations, the distribution of signal yields forms a gaussian peak. The mean value of this gaussian is then the expected limit, and the width is used to compute the $1/2/3\sigma$ bands around the expected limit. Once the expected limit is calculated, the combination tool performs the fit one more time with real data instead of pseudo data.

To get an accurate measurement, the datacard needs to have a complete model of the analysis' shapes and uncertainties. To begin the process, I first start with the simplest datacard possible. The background normalizations are extracted by performing a SHYFT fit with no signal. Once the SHYFT fitter has converged on a solution, it outputs one shape for each process. Instead of a shape per process per bucket, it places each bucket side by side to produce a single long shape per process. Not only are the shapes normalized to the per-process normalizations, but they are also morphed by the fit-based systematics (SF$_b$, SF$_{lf}$, SF$_{jes}$). This is desirable because the SHYFT fit has ostensibly found the values of these

SFs that make the simulation most closely agree with the data. Finally, the signal shapes are morphed by the fit-based systematics, normalized to their theoretical cross-sections and added to the input shapes.

Three different methods were used to implement systematic effects in the combination tool, each corresponding to a category in Section 9.1.

To model systematic effects with pseudo-data, I provide shift up/down shapes for each of these effects. So, in addition to the input shapes from the SHYFT fit, there is a pair of additional shapes for each systematic effect . For example, to add the effect of $t\bar{t}$ Q^2 energy scale, I add two additional shapes: one with the $t\bar{t}$ Q^2 energy scale shifted up one σ and another with the $t\bar{t}$ Q^2 energy scale shifted down one σ. Within the combination tool, these shape nuisance parameters are modeled as gaussian priors.

Effects which scale entire processes up and down (like luminosity) are modeled by log-normal nuisance parameters.

Finally, uncertainties from the SHYFT fit are propagated to the combination tool. This task consists of two independent steps.

To teach the combination tool about how effects change the shapes, it needs the shapes where each systematic effect is shifted up or down by one σ. However, since the combination tool is provided nominal templates whose SFs are scaled away from 1.0, the existing shapes of $1.0 \pm 1\sigma$ don't properly bracket the effect. Instead, the SHYFT fit is run again, fixing every parameter to the fitted values then varying one systematic at a time $\pm 1\sigma$. For example, if $\mathrm{SF}_b = 0.96$ I run the fit where I've fixed everything to the nominal values but change $\mathrm{SF}_b = 0.96 \pm 0.05$.

Next, the remaining statistical uncertainties are added to the fit. To do this, the statistical uncertainty is estimated by rerunning the fit normally except fixing the systematic SFs to their fitted values. Removing the parameters from the fit removes their effects on the likelihood curve, leaving only the uncertainty due to statistics. Each resulting uncertainty on the processes could be added as a nuisance parameter to the combination tool. This isn't strictly accurate, though. Adding a nuisance parameter per-process implies that each of the uncertainties are totally uncorrelated. The correlation matrix for the SHYFT fits show that many processes are strongly correlated with each other. Unfortunately, the combination tool only handles 0% or 100% correlations.

To accurately propagate the uncertainties from the SHYFT fit to the combination tool, the covariance matrix of the SHYFT parameters is produced. The eigenvectors and eigenvalues of the covariance matrix represents the directions and magnitudes of uncertainties in parameter-space. Since covariance matrices are symmetrical, and the eigenvectors of symmetric matrices are orthogonal, each of these eigenvectors are linearly independent of each other. Each eigenvector can then be interpreted as a nuisance parameter in a new basis where each parameter is independent of each other, which allows them to be accepted by the combination tool. As a result, the fitter understands both the SHYFT-provided uncertainties on the normalizations and the correlations between them.

Once the input shapes and their uncertainties modeled within the combination tool, it extracts the limits for each mass point which can be seen in Figure 9.4.

The range of mass points chosen at the beginning of this analysis was not wide enough to bracket the theoretical values, which prevented the ability to set a limit. In order to have enough breadth to set a limit, an additional point was added at 700 GeV. Instead of

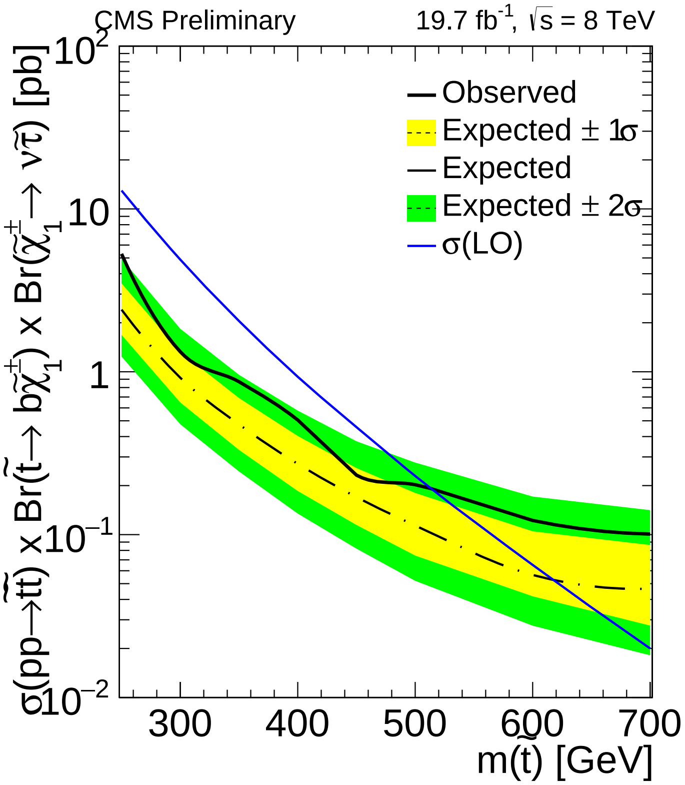


Figure 9.4: Confidence limits for the examined SUSY scenarios.

reprocessing this point from scratch, the final point is an extrapolation from the previous points.

As the stop mass increases, the production cross-section drops, but the shapes remain nearly identical. The 700 GeV limit was produced by taking the 600 GeV shapes, rescaling them to the 700 GeV cross-section, and rerunning the combination tool with this shape estimate.

The observed signal rate in Figure 9.4 deviates from the expected limit by approximately 1σ across the entire mass range. The expected limit and $\pm 1, 2\sigma$ bands represents background-only hypothesis the 1 and 2 σ standard deviations the mean measurement. The observed limit straddles the $+1\sigma$ band, meaning there is roughly a 68% probability the observed limit is due to statistical fluctuations alone. The leading order theoretical calculations of the cross section intersect the observed limit at $m(\widetilde{t}) = 525\,\mathrm{GeV}$, which excludes $m(\widetilde{t})$ less than $525\,\mathrm{GeV}$.

Chapter 10

Conclusion

This analysis describes a search for $\widetilde{t}\,\widetilde{t}^*$ in 8 TeV data recorded using the CMS detector in 2012. To increase sensitivity, the technique known as "Simultaneous Heavy Flavor and Top", or SHYFT, was used. This approach involves dividing up a sample based on event content then simultaneously fitting both the signal and all bckgrounds. Additionally, major sources of systematic uncertainty are constrained by the data as well. By extracting all values *in situ*, the overall measurement error is reduced.

This technique was first developed at the CDF experiment[5] to measure the $t\bar{t}$ cross-section and was later used at CMS to measure the $t\bar{t}$ cross-section at 7 TeV. In both cases, SHYFT was strongly competitive with other results.

For this analysis, additional discrimination is provided by further dividing the data based on τ-lepton content since the targeted signal is expected to be rich in τs.

The SHYFT fit finds no significant excess compared to the standard model for any of the the analyzed mass points. The Higgs combination tool was then used to obtain 95% C.L. upper limits on the cross section for each mass point using a CL_s method. The Higgs combination tool performs a shape based analysis to determine the likelihood of observing signal in the presence of the measured data and background. These significances exclude masses of $\widetilde{t}$ less than 525 GeV.

Appendix A

Input Polynoids

The following sections are the polynoids which describe the affect of each systematic effect on each sample and jet/tag bin. Polynoids are described in Section 7.4.

A.1 B-Tag Efficiency Scale Factor Polynoids

A.2 B-Mistag Efficiency Scale Factor Polynoids

A.3 JES Scale Factor Polynoids

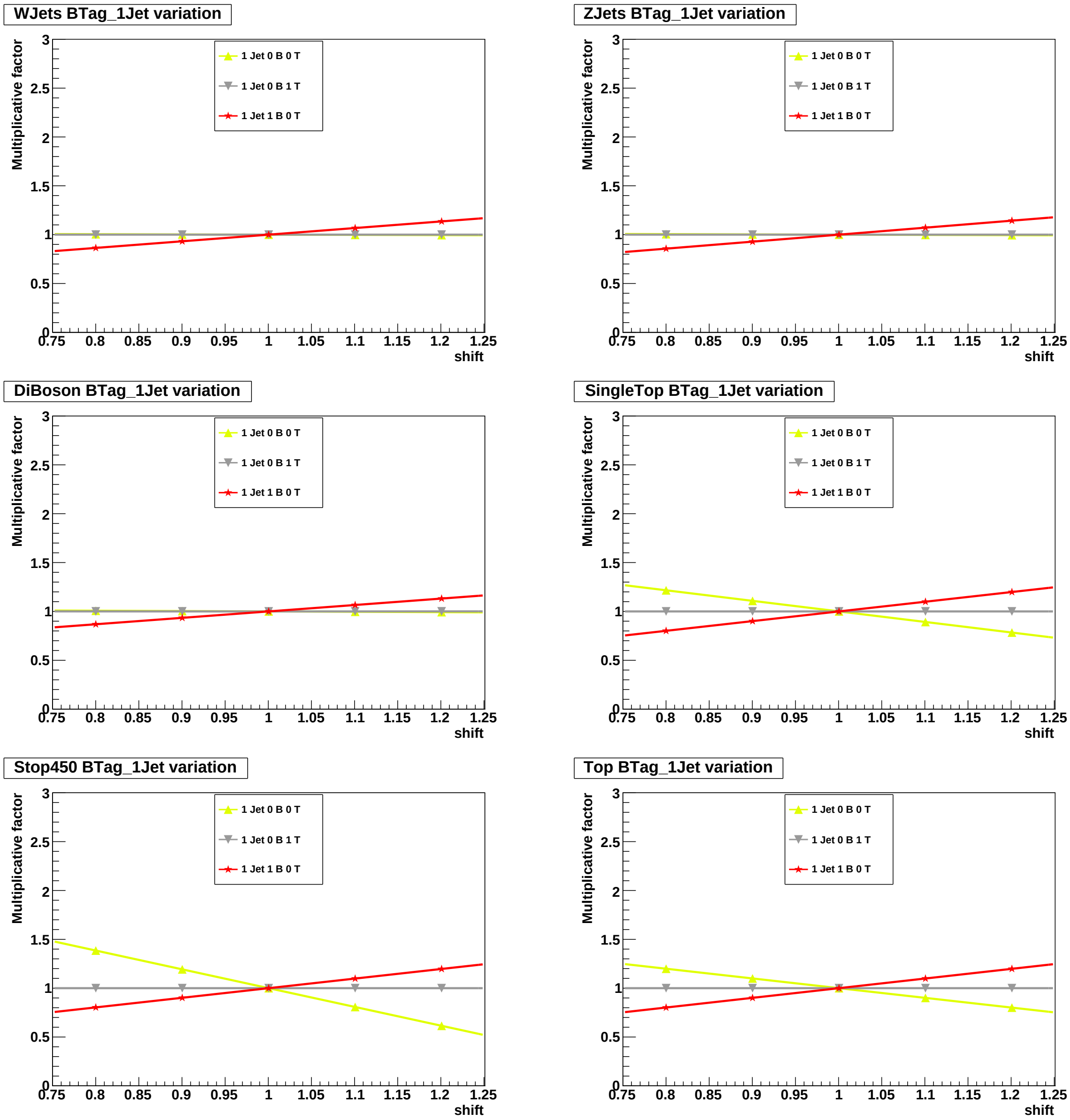


Figure A.1: Polynoids Characterizing Effect of b-tagging on 1-jet multiplicity events. Each shift of 0.1 corresponds to a 10% change in SF_{btag}.

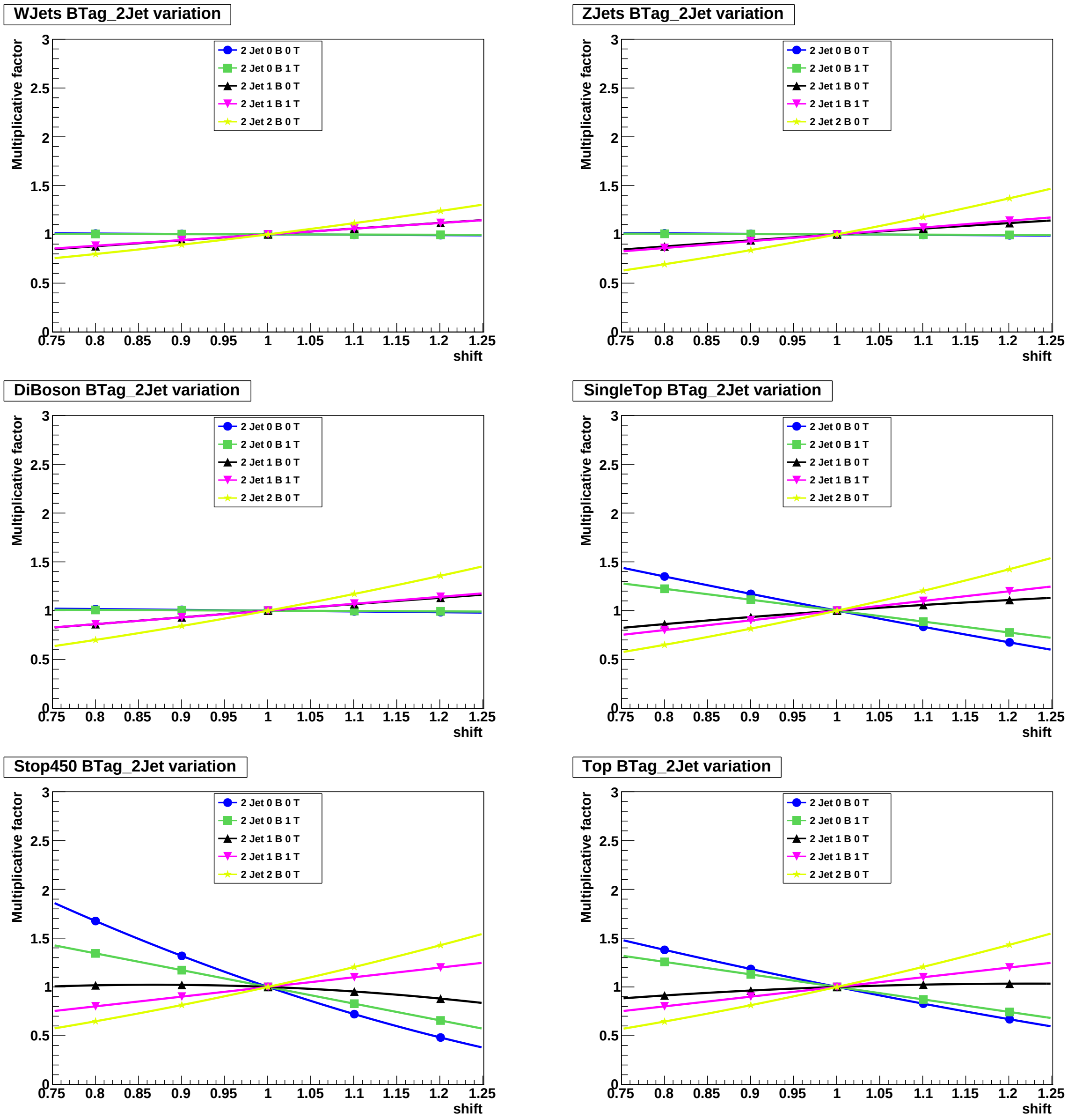


Figure A.2: Polynoids Characterizing Effect of b-tagging on 2-jet multiplicity events. Each shift of 0.1 corresponds to a 10% change in SF_{btag}.

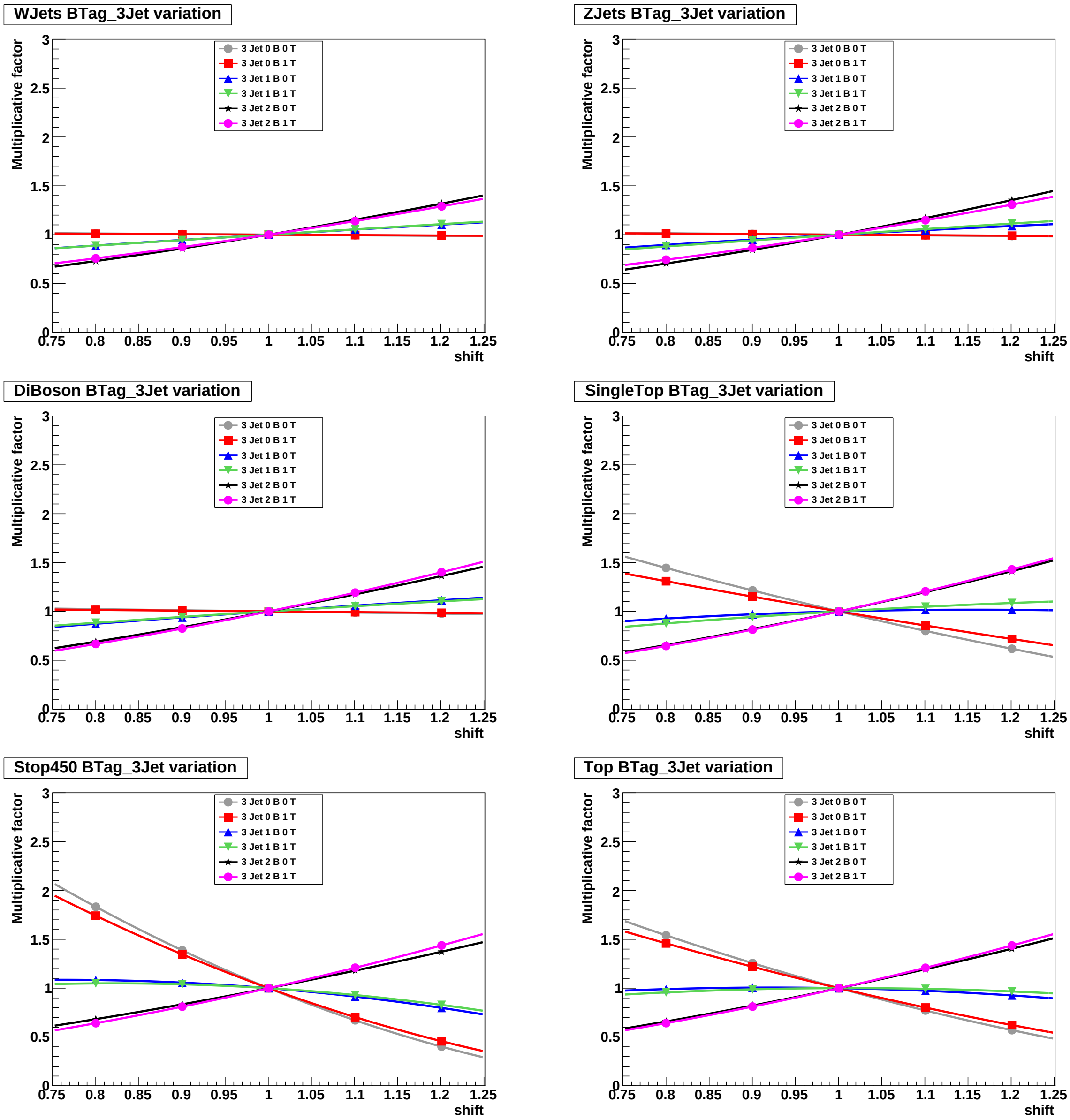


Figure A.3: Polynoids Characterizing Effect of b-tagging on 3-jet multiplicity events. Each shift of 0.1 corresponds to a 10% change in SF_{btag}.

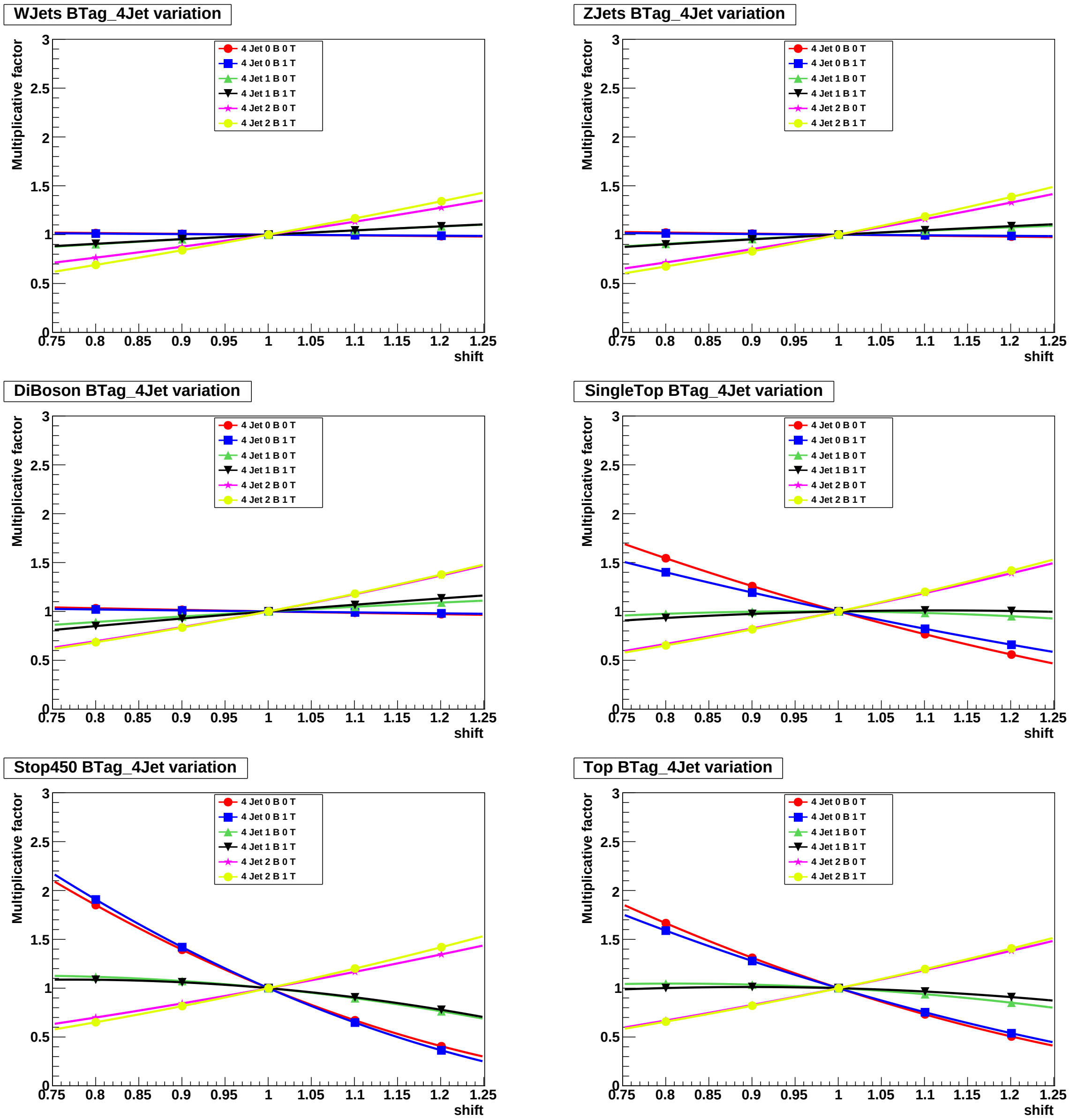


Figure A.4: Polynoids Characterizing Effect of b-tagging on 4-jet multiplicity events. Each shift of 0.1 corresponds to a 10% change in SF_{btag}.

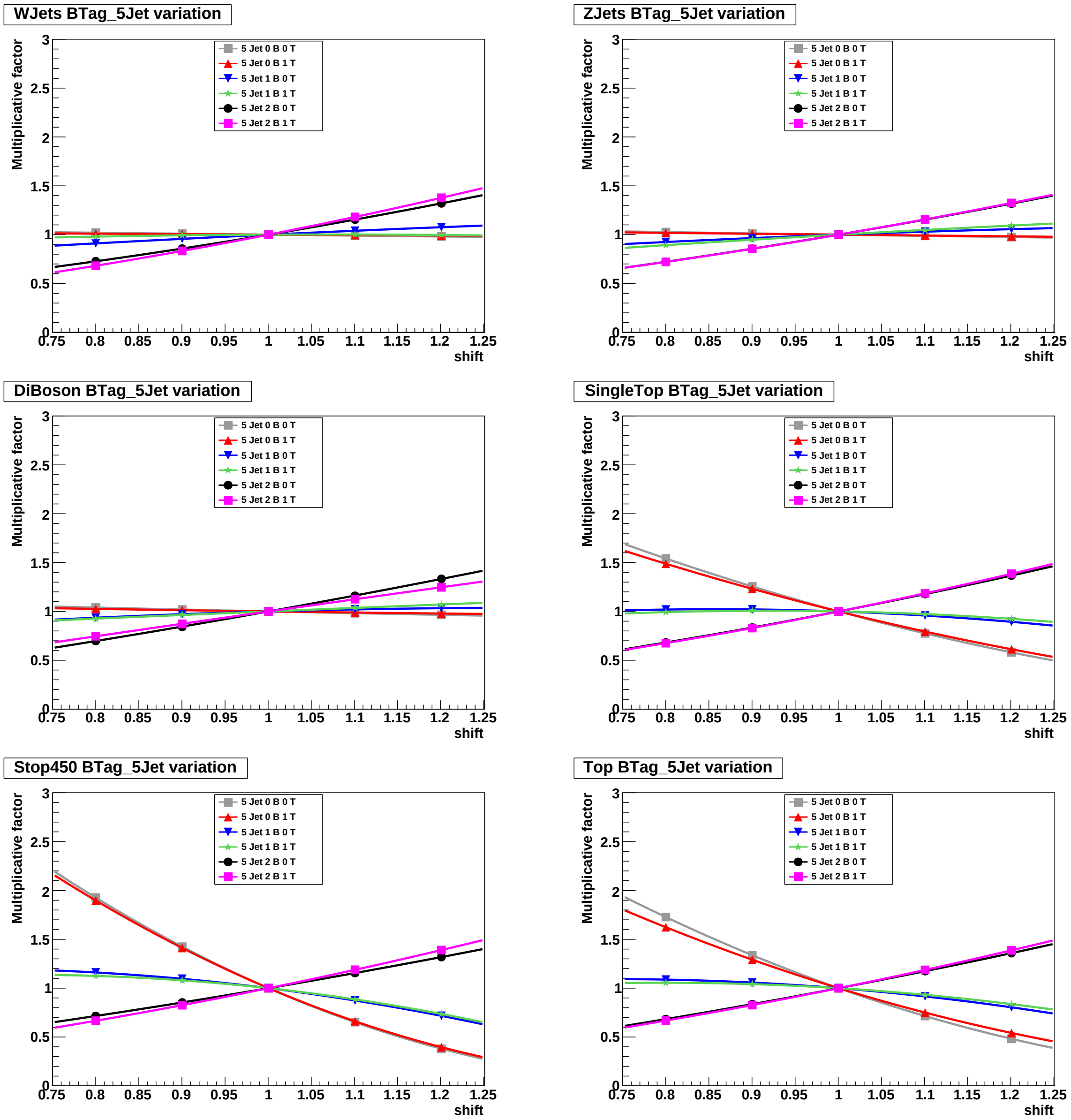


Figure A.5: Polynoids Characterizing Effect of b-tagging on 5-jet multiplicity events. Each shift of 0.05 corresponds to a 5% change in SF_{jes}.

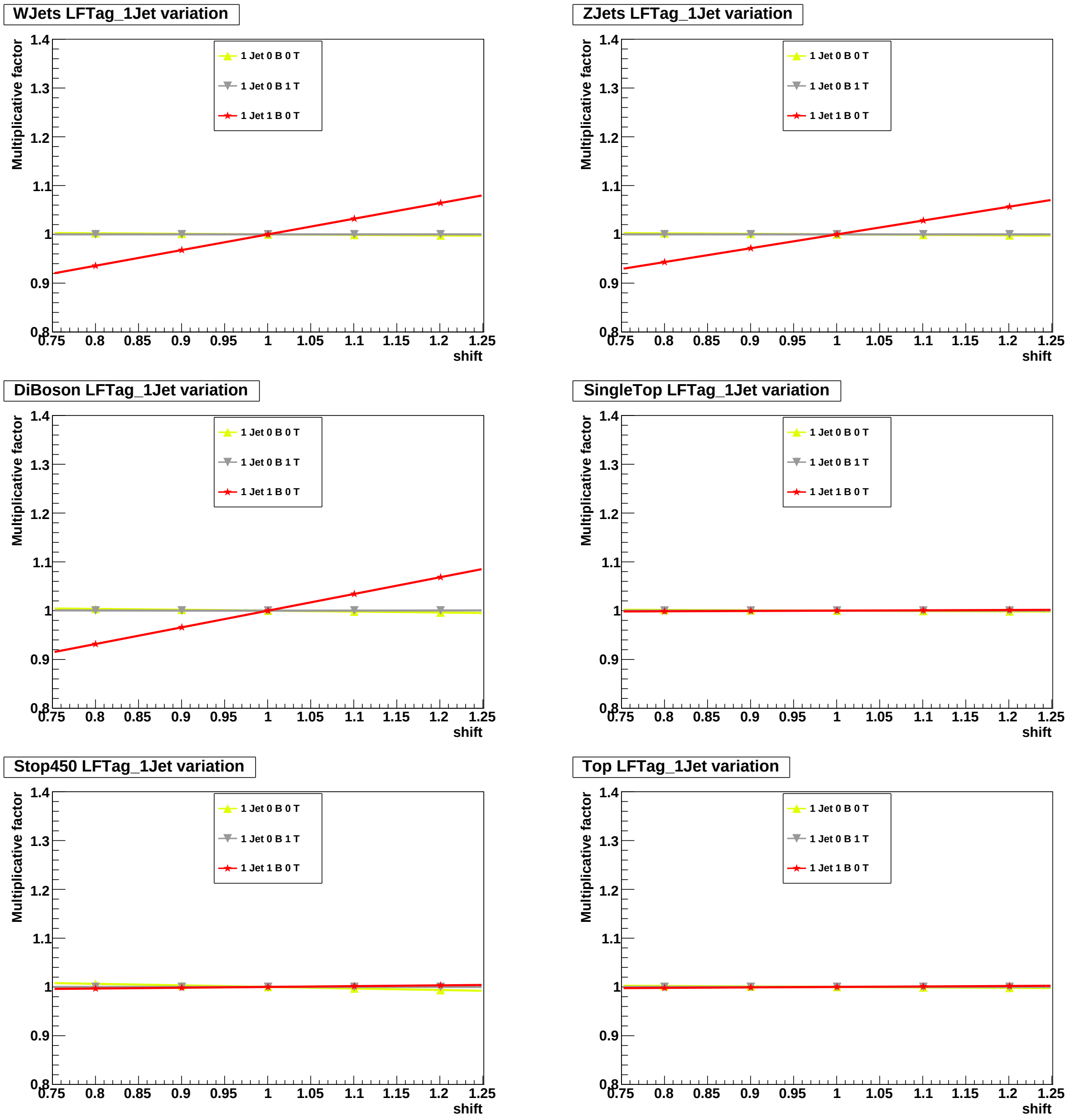


Figure A.6: Polynoids Characterizing Effect of lf-mistagging on 1-jet multiplicity events. Each shift of 0.1 corresponds to a 10% change in SF_{lftag}.

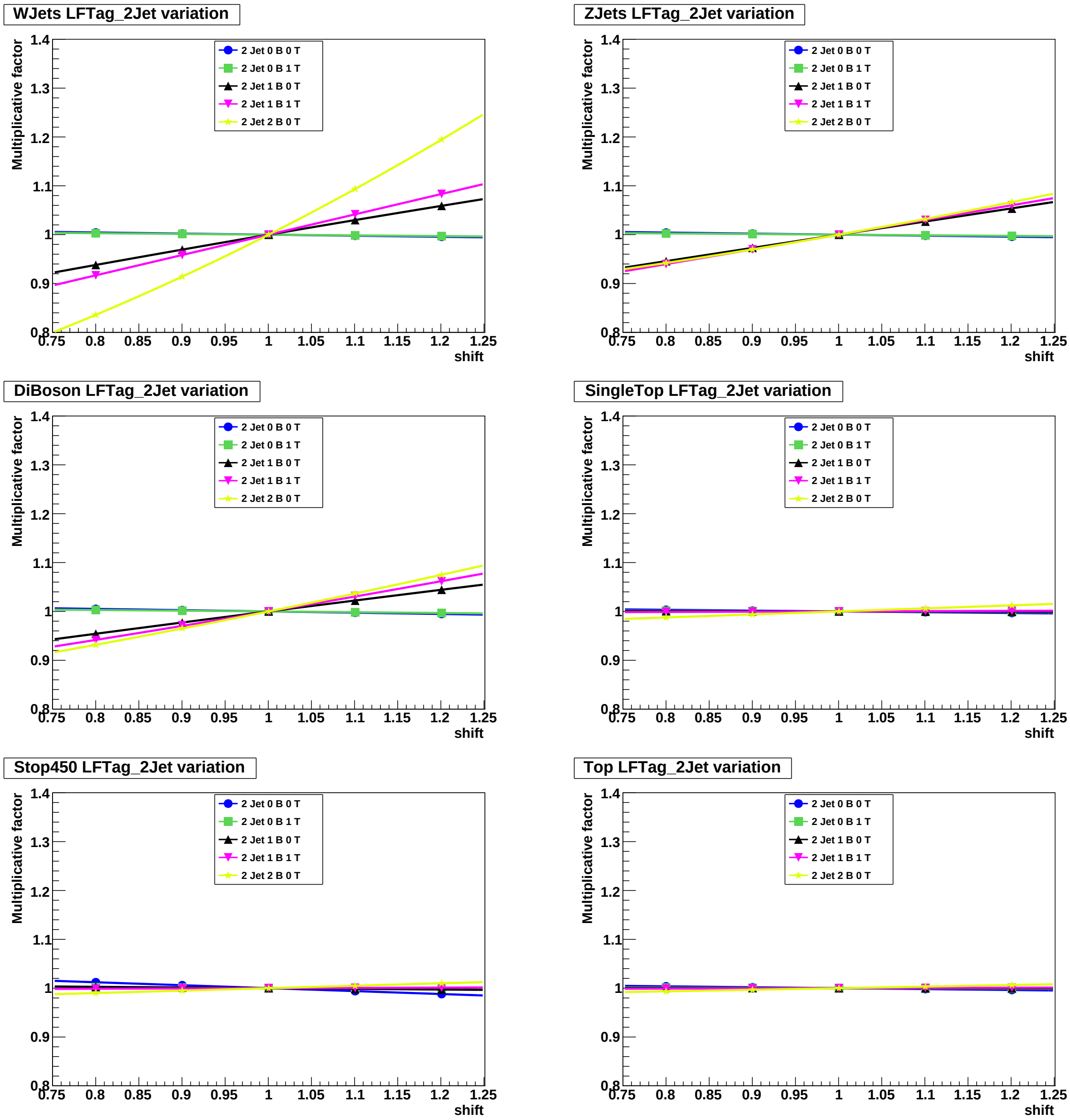


Figure A.7: Polynoids Characterizing Effect of lf-mistagging on 2-jet multiplicity events. Each shift of 0.1 corresponds to a 10% change in SF_{lftag}.

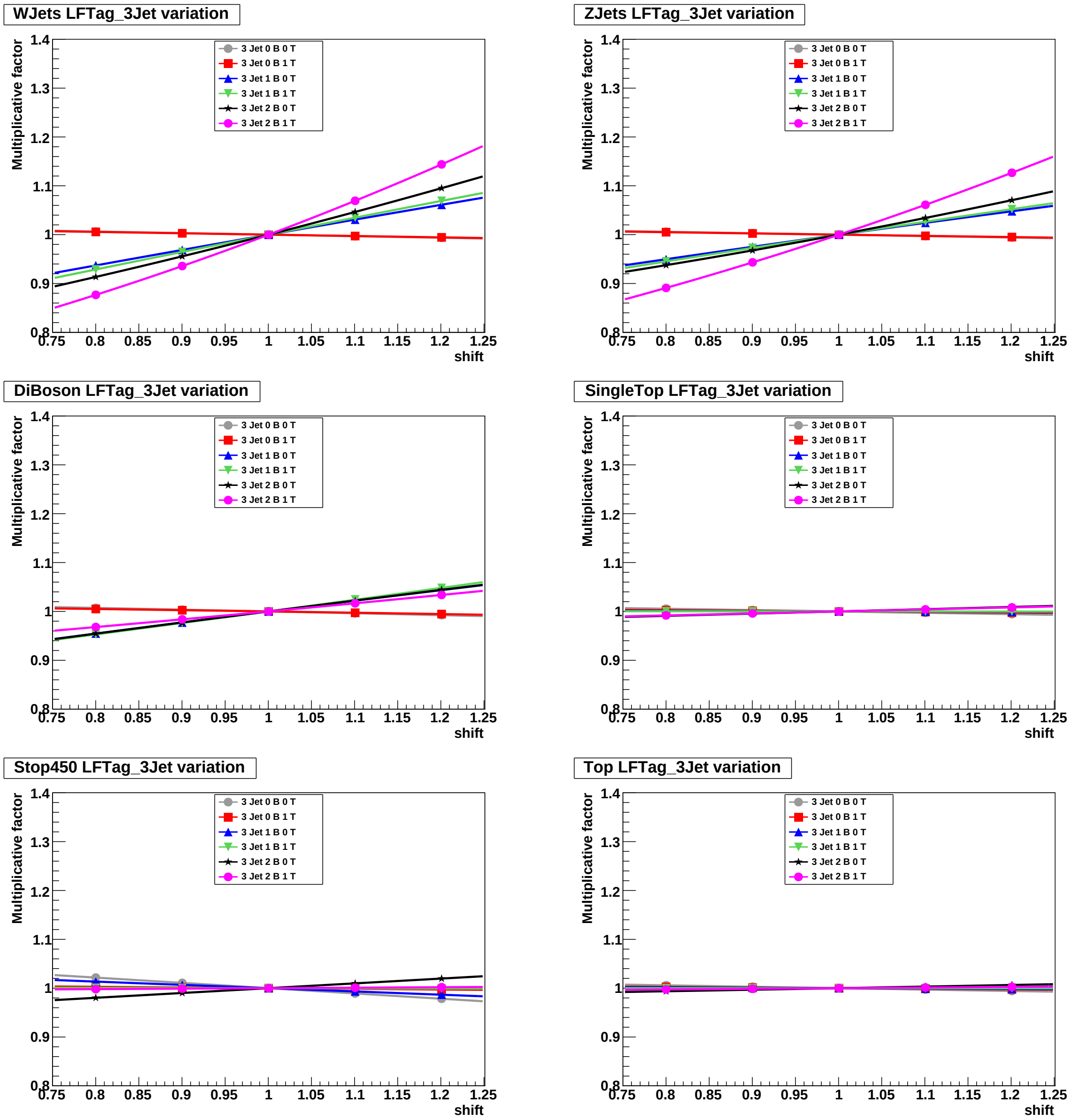


Figure A.8: Polynoids Characterizing Effect of lf-mistagging on 3-jet multiplicity events. Each shift of 0.1 corresponds to a 10% change in SF_{lftag}.

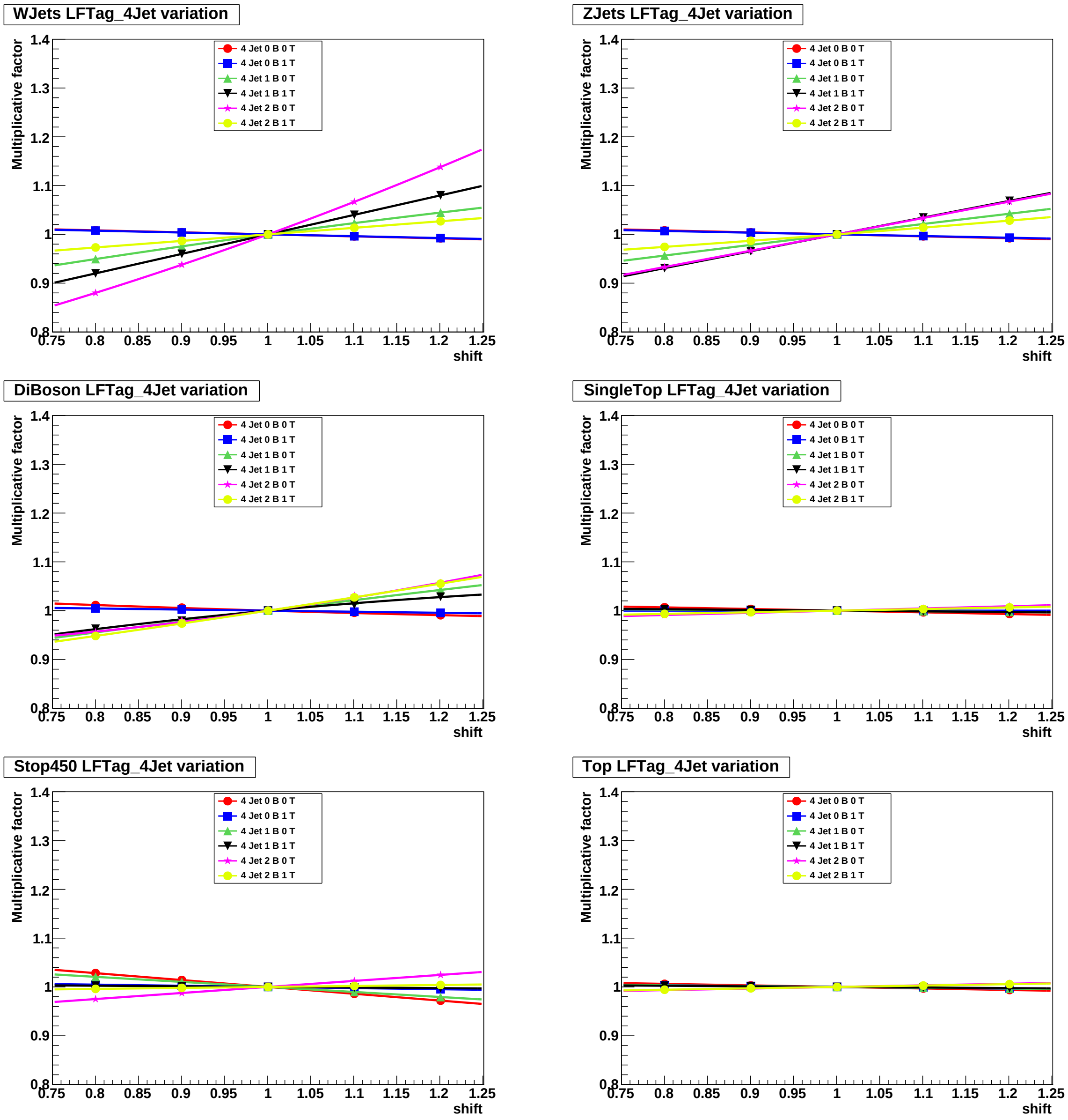


Figure A.9: Polynoids Characterizing Effect of lf-mistagging on 4-jet multiplicity events. Each shift of 0.1 corresponds to a 10% change in SF_{lftag}.

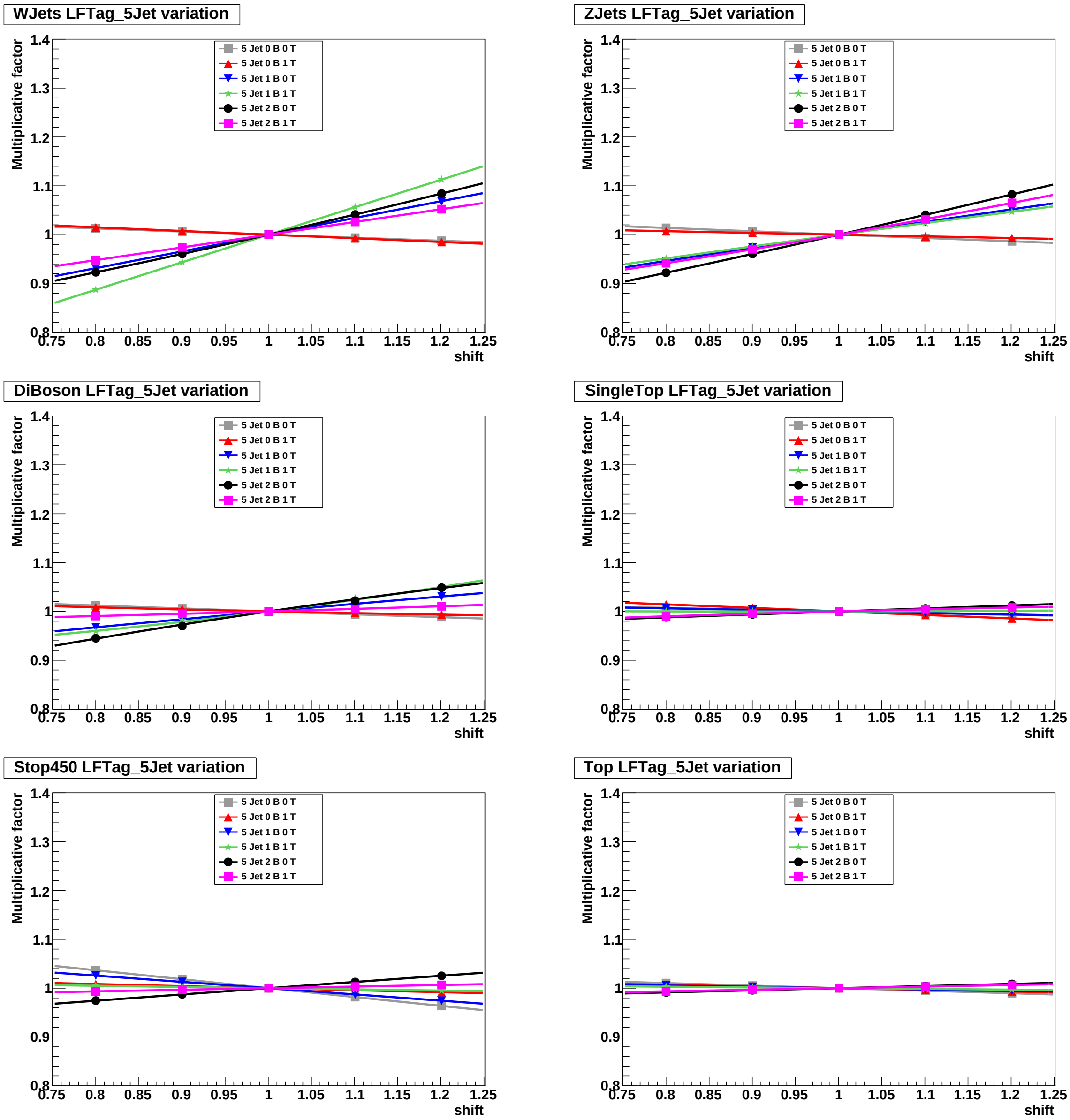


Figure A.10: Polynoids Characterizing Effect of lf-mistagging on 5-jet multiplicity events. Each shift of 0.1 corresponds to a 10% change in SF_{lftag}.

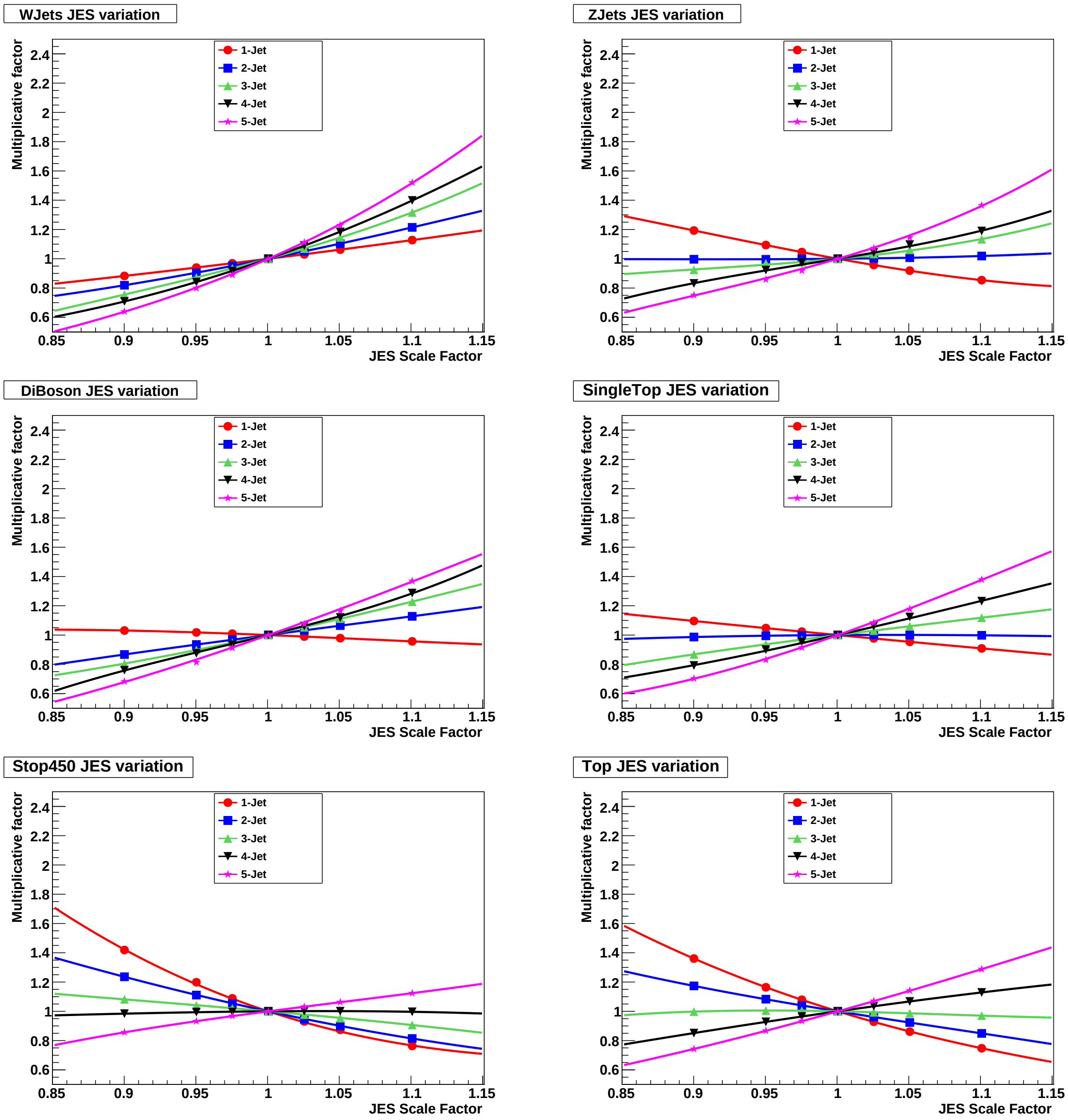


Figure A.11: Polynoids Characterizing Effect of JES on 1-jet multiplicity events. Each shift of 0.1 corresponds to a 10% change in SF_{jes}.

Appendix B

QCD Fit Output Distributions

This analysis performs a data-driven estimate of the QCD contribution The following histograms correspond to the results of the QCD normalization procedure

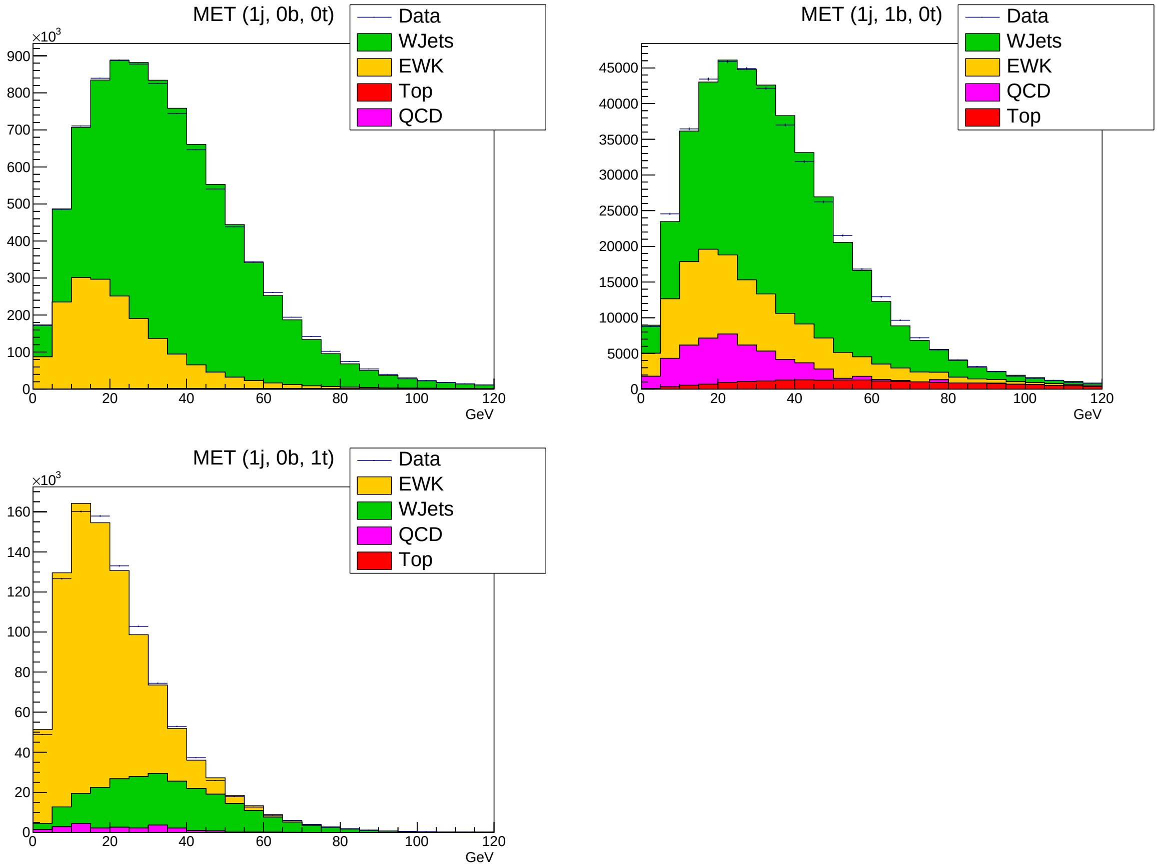


Figure B.1: Fit results for QCD normalization, = 1 jet bins

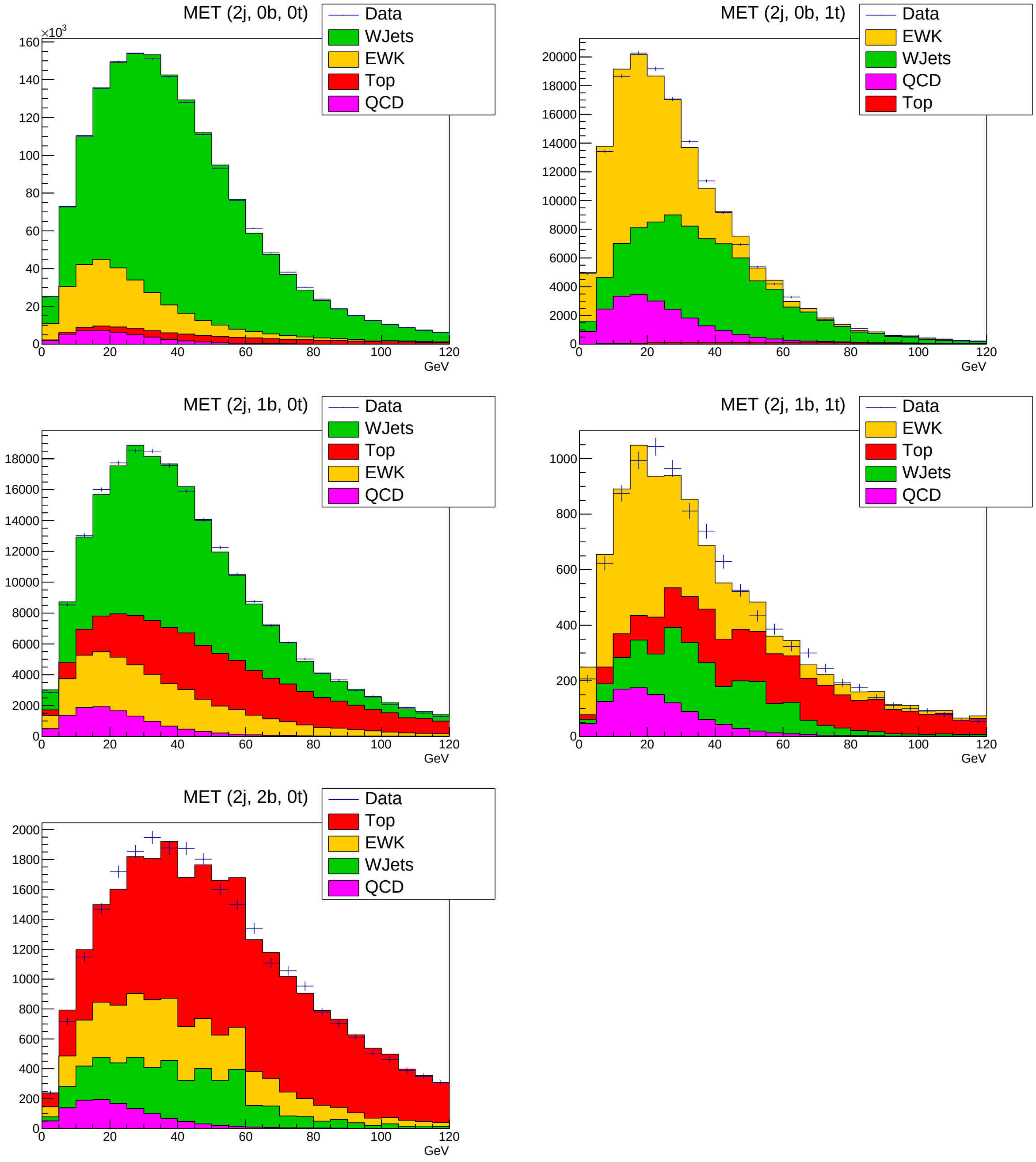


Figure B.2: Fit results for QCD normalization, = 2 jet bins

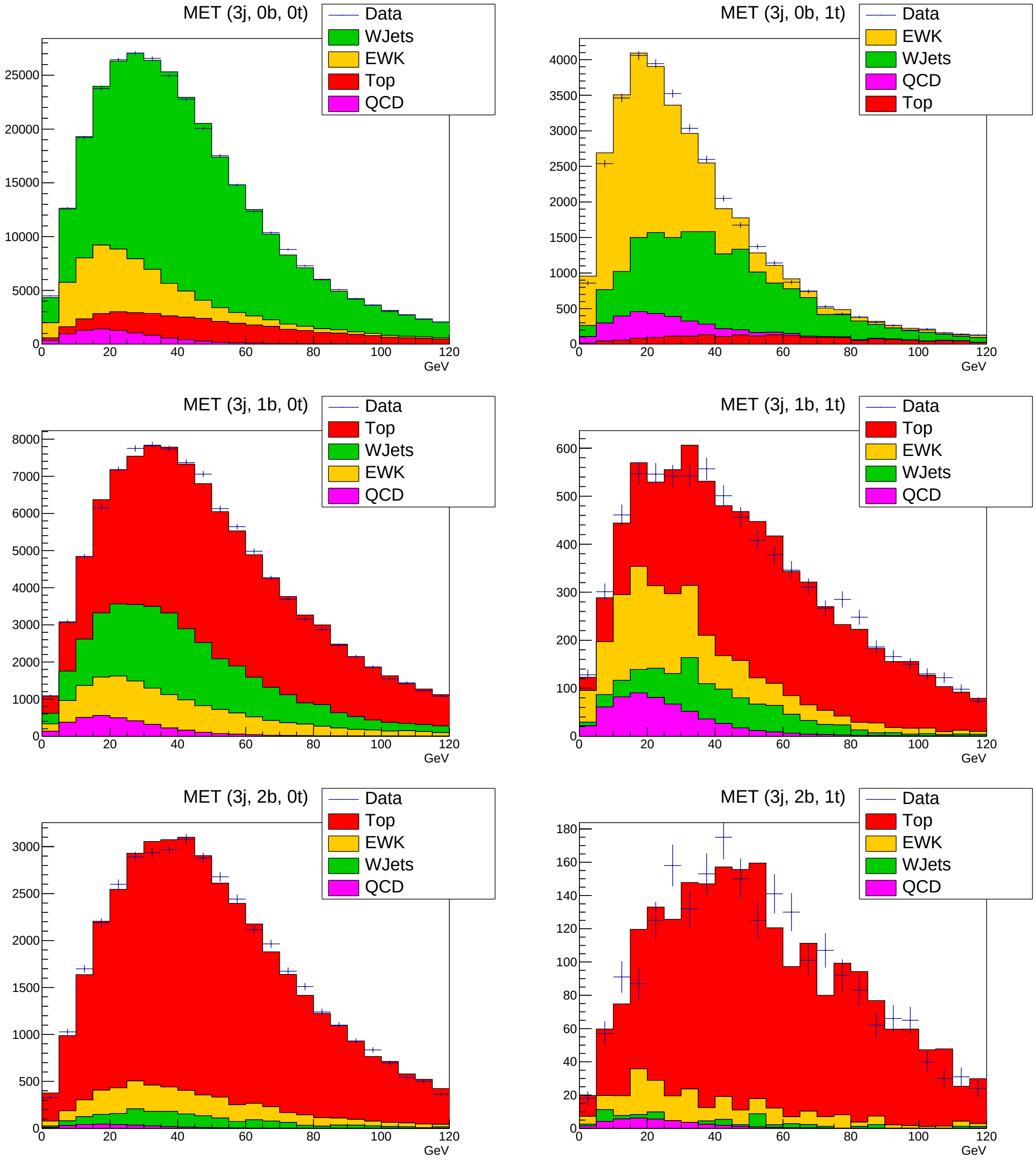


Figure B.3: Fit results for QCD normalization, = 3 jet bins

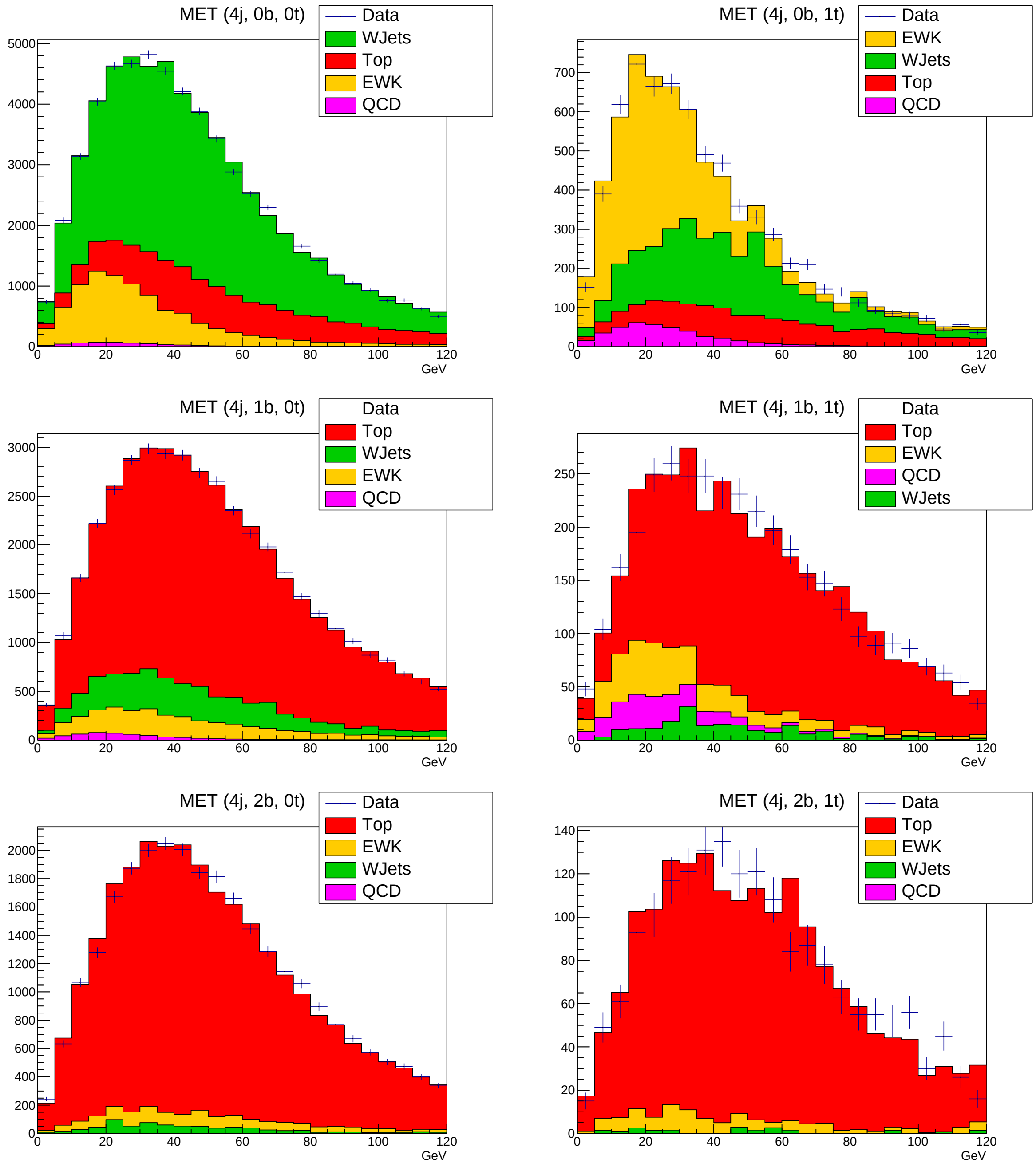


Figure B.4: Fit results for QCD normalization, = 4 jet bins

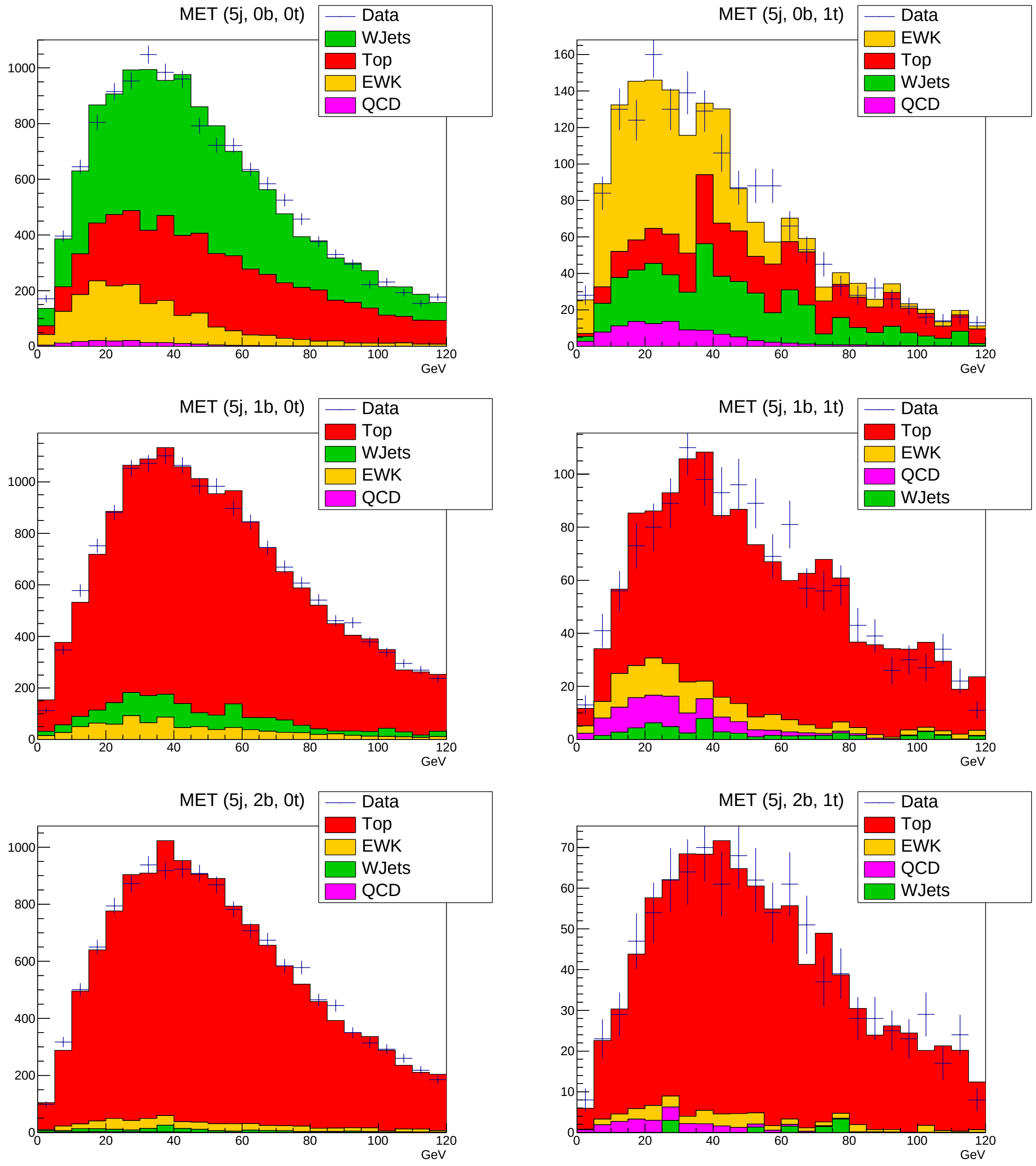


Figure B.5: Fit results for QCD normalization, ≥ 5 jet bins

BIBLIOGRAPHY

[1] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22:579, 1961.

[2] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264, 1967.

[3] Abdus Salam. Weak and electromagnetic interactions. In Nils Svartholm, editor, *Elementary particle physics: relativistic groups and analyticity*, page 367, Stockholm, 1968. Almqvist & Wiksell. Proceedings of the eighth Nobel symposium.

[4] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.

[5] D. Acosta et al. Measurement of the $t\bar{t}$ production cross section in $p\overline{p}$ collisions at $\sqrt{s} = 1.96tev$ using lepton $+$ jets events with secondary vertex b -tagging. *Phys. Rev. D*, 71(5):052003, Mar 2005.

[6] The CMS collaboration. Measurement of the $t\bar{t}$ production cross section in pp collisions at with lepton + jets final states. *Physics Letters B*, 720(13):83 – 104, 2013.

[7] Christian Weiser. A Combined Secondary Vertex Based B-Tagging Algorithm in CMS. Technical Report CMS-NOTE-2006-014, CERN, Geneva, Jan 2006.

[8] The CMS collaboration. Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. Technical Report CMS-PAS-PFT-09-001, CERN, 2009. Geneva, Apr 2009.

[9] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti-k(t) jet clustering algorithm. *JHEP*, 04:063, 2008.

[10] The CMS collaboration. Performance of b tagging at sqrt(s)=8 TeV in multijet, ttbar and boosted topology events. Technical Report CMS-PAS-BTV-13-001, CERN, Geneva, 2013.

[11] The CMS collaboration. *The CMS magnet project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.

[12] V Karimaki, M Mannelli, P Siegrist, H Breuker, A Caner, R Castaldi, K Freudenreich, G Hall, R Horisberger, M Huhtinen, and A Cattai. *The CMS tracker system project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.

[13] D Kotlinski. Status of the CMS Pixel detector. *JINST*, 4(03):P03019, 2009.

[14] A Satpathy. Overview and status of the CMS silicon strip tracker. *JPCS*, 110(9):092026, 2008.

[15] The CMS collaboration. *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.

[16] The CMS collaboration. *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.

[17] The CMS collaboration. *The CMS muon project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.

[18] Vardan Khachatryan et al. Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV. *JINST*, 10:P08010, 2015.

[19] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008.

[20] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet user manual. *Eur. Phys. J. C*, 72:1896, 2012.

[21] The CMS collaboration. Particle-flow event reconstruction in CMS and performance for jets, taus, and MET. CMS Physics Analysis Summary CMS-PAS-PFT-09-001, 2009.

[22] The CMS collaboration. Commissioning of the particle-flow event with the first LHC collisions recorded in the CMS detector. CMS Physics Analysis Summary CMS-PAS-PFT-10-001, 2010.

[23] Serguei Chatrchyan et al. Determination of jet energy calibration and transverse momentum resolution in CMS. *JINST*, 6:P11002, 2011.

[24] Serguei Chatrchyan et al. Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7\,\mathrm{TeV}$. *JINST*, 7:P10002, 2012.

[25] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[26] The CMS collaboration. *CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems*. Technical Design Report CMS. 2000.

[27] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.

[28] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.

[29] Johannes Bellm et al. Herwig 7.0 / Herwig++ 3.0 Release Note. 2015.

[30] Z. Was. TAUOLA the library for tau lepton decay, and KKMC / KORALB / KORALZ /... status report. *Nucl. Phys. Proc. Suppl.*, 98:96–102, 2001. [,96(2000)].

[31] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[32] Andrea Giammanco. The Fast Simulation of the CMS Experiment. *J. Phys. Conf. Ser.*, 513:022012, 2014.

[33] The CMS collaboration. Single Muon efficiencies in 2012 Data. CMS Detector Performance Summary CMS-DP-2013-009, Mar 2013.

[34] T Aaltonen et al. Higgs boson studies at the tevatron. *Phys. Rev. D*, 88:052014, Sep 2013.

[35] Documentation of the roostats-based statistics tools for higgs pag. https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideHiggsAnalysisCombinedLimit, 2014.

[36] Thomas Junk. Confidence level computation for combining searches with small statistics. *Nucl. Instrum. Meth. A*, 434:435, 1999.

[37] A.L. Read. Modified frequentist analysis of search results (the CL_s method). Technical Report CERN-OPEN-2000-005, CERN, 2000.