SEARCH FOR SAME-SIGN DILEPTON SUSY SIGNATURES USING TOP DECAY EVENTS WITH CMS AT THE LHC

by

MARC WEINBERG

A dissertation submitted in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

(PHYSICS)

at the

UNIVERSITY OF WISCONSIN – MADISON

2011

SEARCH FOR SAME-SIGN DILEPTON SUSY SIGNATURES USING TOP DECAY EVENTS WITH CMS AT THE LHC

Marc Weinberg

Under the supervision of Professor Wesley H. Smith

At the University of Wisconsin — Madison

A search for new physics is performed in conjunction with a measurement of the $t\bar{t}$ production cross section in pp collisions at a center-of-mass energy of 7 TeV. The analysis uses a data sample corresponding to an integrated luminosity of $36 \pm 1\,\text{pb}^{-1}$ collected at the LHC with the CMS detector. The $t\bar{t}$ cross section measurement uses events with one isolated muon and multiple hadronic jets. The measured cross section is found to be $\sigma(t\bar{t}) = 159.1 \pm 12.1(\text{stat})^{+33.8}_{-28.2}(\text{syst}) \pm 6.4(\text{lumi})\text{pb}$, consistent with the expectations of the Standard Model. This measurement is used in the new physics search, which requires two same-sign isolated leptons, multiple hadronic jets and missing transverse energy in the final state. The observed event yields agree with Standard Model predictions, and no evidence of new physics is found.

To my friends and family.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In answer to the question of why it happened, I offer the modest proposal that our universe is simply one of those things that happen from time to time.

— Edward Tryon

1.1 The Standard Model of Elementary Particle Physics

Particle physics, at its heart, concerns itself with two very fundamental questions: *What are the ultimate building blocks of matter?* And *how do these fundamental constituents interact?* It began as what might generously be described as a "philosophical tradition" in ancient Greece, but only in the last two centuries has it attained the level of scientific rigor necessary to place it among the greatest intellectual endeavors of the human race. Since 1802, when John Dalton established the atomic theory in its modern form with his observation that each element was composed of a single, unique type of particle, our answers to these two fundamental questions have grown increasingly sophisticated. Startlingly, though the two questions are very different, our answers to them have also grown increasingly similar.

The language of particle physics is quantum field theory [1], in which particles are

represented by excitations in a quantized field (with each kind of particle getting its own field), and their interactions are mediated by other, force-carrying particles. Thus far, field theories have been very successful at describing three of the four known fundamental forces: electromagnetism, the weak interaction, and the strong interaction, with gravity being the only holdout. These theories are collectively referred to as the Standard Model (SM). In fact, all three of these forces are described by a particular *type* of quantum field theory, known as gauge theories due to their symmetry under gauge transformations (that is, transformations on degrees of freedom in the system that should not affect physical results). In each case, the force between particles is caused by the exchange of force carriers called *gauge bosons*. In the case of electromagnetism, this force carrier is the photon (γ). The weak force is carried by two different intermediate vector bosons, the $W^{\pm}$ and the Z^0, and the strong force is carried by particles called gluons (g).

1.1.1 Fermions

While all the force carriers are bosons, matter is composed of spin-$\frac{1}{2}$ fermions (see Section 1.1.4 below). These are further subdivided based on whether they participate in the strong interaction (quarks) or not (leptons). They are generally organized into three "generations" with regularly repeating properties, as shown in Figure 1.1. In each generation, there is a quark with an electric charge of $+2/3$ (in units of the electron charge), a quark with a charge of $-1/3$, a massless (or very light) neutral lepton called a neutrino, and a lepton with -1 electric charge. In addition to each particle shown here, there exists a corresponding antiparticle with identical mass but opposite charge and quantum numbers. The quantum numbers are repeated in each generation, with the only differences between corresponding particles being their mass. The first generation is stable, and all ordinary baryonic matter

is composed from it. The second and third generation particles, by virtue of their greater mass, can decay into particles of the first generation, and often have very short half lives.

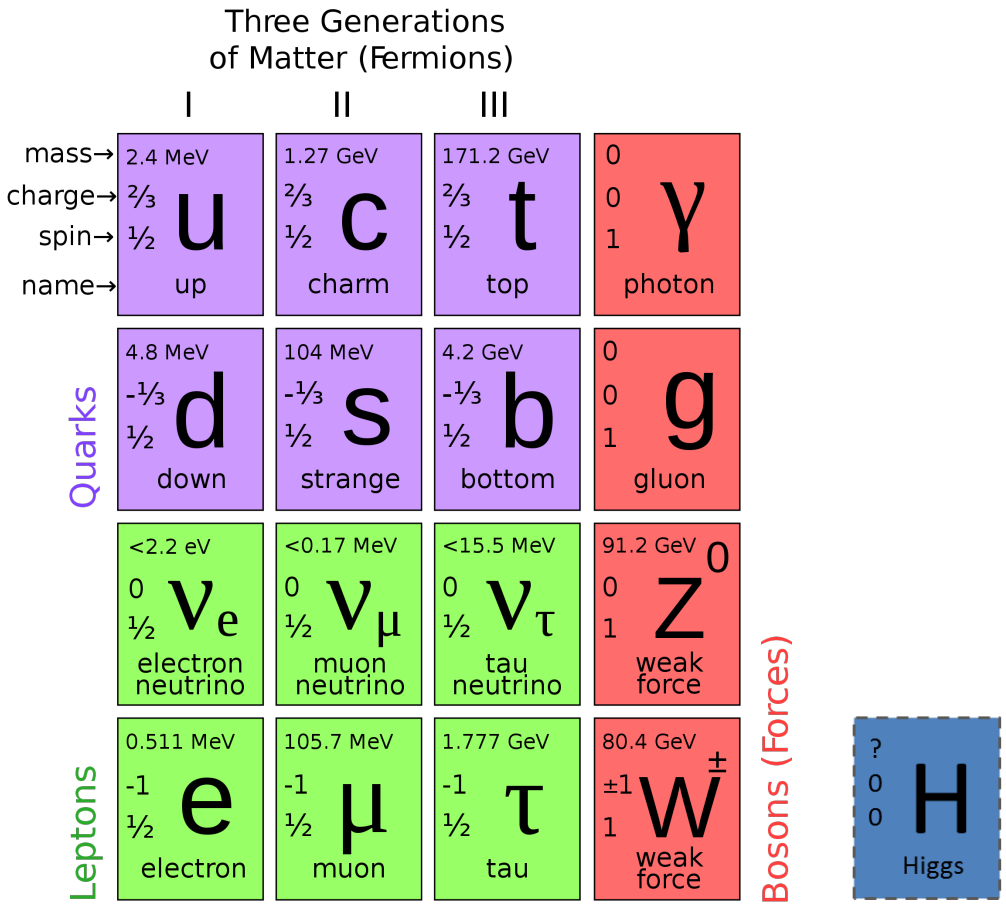


Figure 1.1: The Standard Model of elementary particles, with the gauge bosons in the rightmost column. In addition, for each matter particle shown, there is a corresponding antiparticle with identical mass and spin (see Section 1.1.4) but opposite quantum numbers. The Higgs boson, which has not yet been discovered, is also shown.

Quarks carry the charge associated with strong interactions, called "color", and experience the phenomenon of *color confinement*, which ensures that they are never found free in nature, but are instead confined in colorless composite particles called *hadrons*, of

which protons and neutrons are the most common examples [2]. There are three distinct types of color charge and consequently two different ways to make a color neutral combination: *Baryons* are composed of three quarks (or three antiquarks), one with each color charge; protons (uud) and neutrons (udd) are examples of baryons. *Mesons* are composed of a quark and an antiquark, such that the color charge and its opposite cancel.

1.1.2 Gauge bosons

Forces are transmitted between particles via the exchange of virtual gauge bosons, as illustrated in the diagrams in Figure 1.2. The intermediate state violates energy conservation (via the Heisenberg uncertainty principle), and thus the exchanged boson must be virtual. Though the strengths and characteristics of the three forces vary widely, they are all caused by closely analogous processes, with the differences attributable to the different characteristics of the gauge bosons. The potential energy of the interactions varies as e^{-mr}/r, where m is the mass of the gauge boson and r is the spatial separation the between fermions [3]. The photon is both massless and chargeless, and as a result the range of the electromagnetic force is infinite and follows the familiar inverse-square law. The intermediate vector bosons $W^{\pm}$ and Z^0 are exponentially suppressed due to their mass ($M_W = 81\,\text{GeV}$, $M_Z = 92\,\text{GeV}$) and therefore have a very short effective range (of order 10^{-16}cm).

Gluons, like photons, are massless, and at sufficiently short distances the strong force behaves analogously to electromagnetism. However, after a limiting distance on the order of 10^{-16}cm, the strong force remains at a constant strength, independent of the distance between the quarks. This is the source of color confinement and is the reason individual free quarks are never observed. This dramatically different behavior arises from the fact that, while photons are neutral in the field they mediate (that is, they have no electric charge),

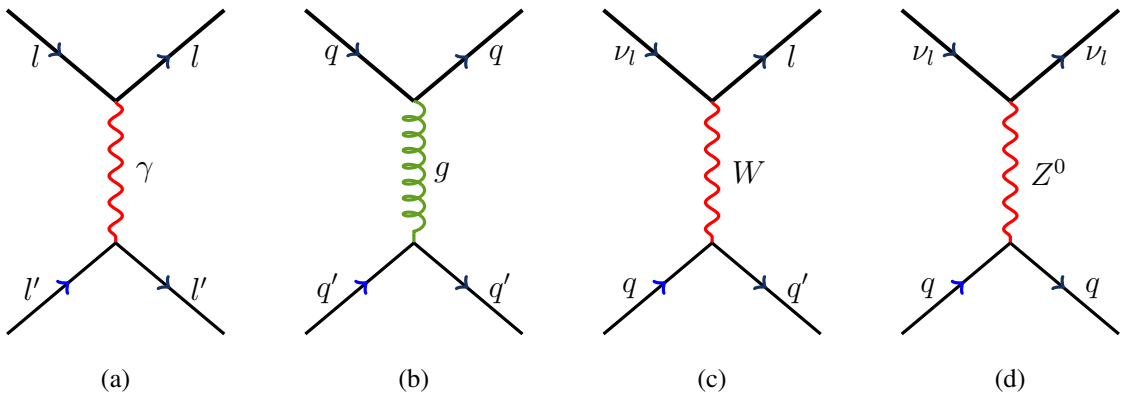


Figure 1.2: (a) Lepton-lepton scattering via the emission and reabsorption of a virtual photon, (b) quark-quark scattering via the emission and reabsorption of a virtual gluon, and neutrino-quark scattering via (c) charged current W-exchange and (d) neutral current Z-exchange.

gluons carry a color charge and an anti-color charge, and hence couple to themselves, as illustrated in Figure 1.3.

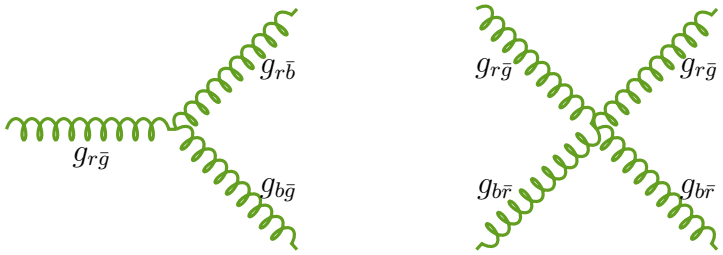


Figure 1.3: Gluon self-coupling vertices with color flow indicated. (a) shows a 3-gluon vertex and (b) shows a 4-gluon vertex.

No such self-coupling exists for photons (which only couple to electrically charged particles), and this additional color interaction accounts for the different behavior of the strong force. Due to photon pair-production, electrons are surrounded by a cloud of virtual

electron/positron pairs, and because oppposite charges attract, this cloud is polarized such that the positrons are preferentially nearer to the core electron, effectively "screening" the electron charge at longer distances. The situation is reversed for quarks due to the additional gluon self coupling: here the gluons can pair-produce other *gluons* [2]. Since the gluons themselves carry color charge, this has the effect of spreading out the charge of the original quark. This "anti-screening" causes the effective quark charge to *increase* with distance, exactly the opposite case of the electron charge. A second quark will experience a reduced force as it nears the first quark, a phenomenon known as *asymptotic freedom* [4], and increasing force as it moves away, explaining why quarks are always found in the colorless bound states of baryons or mesons.

1.1.3 Higgs boson

The final piece of the Standard Model is a hypothesized scalar (spin-0) particle called the Higgs boson. It has not yet been observed, but it is expected that it will be seen in pp collisions at the LHC. The Higgs boson is a consequence of electroweak symmetry breaking, which is believed to be responsible for generating the masses of the $W^{\pm}$ and Z^0 gauge bosons (while leaving the photon massless). The Higgs field possesses a non-zero vacuum expectation value, and this is responsible for generating masses for every particle that couples to the Higgs boson, including the Higgs boson itself. Figure 1.4 neatly summarizes the Standard Model by providing a diagramatic representation of the possible interactions between every particle.

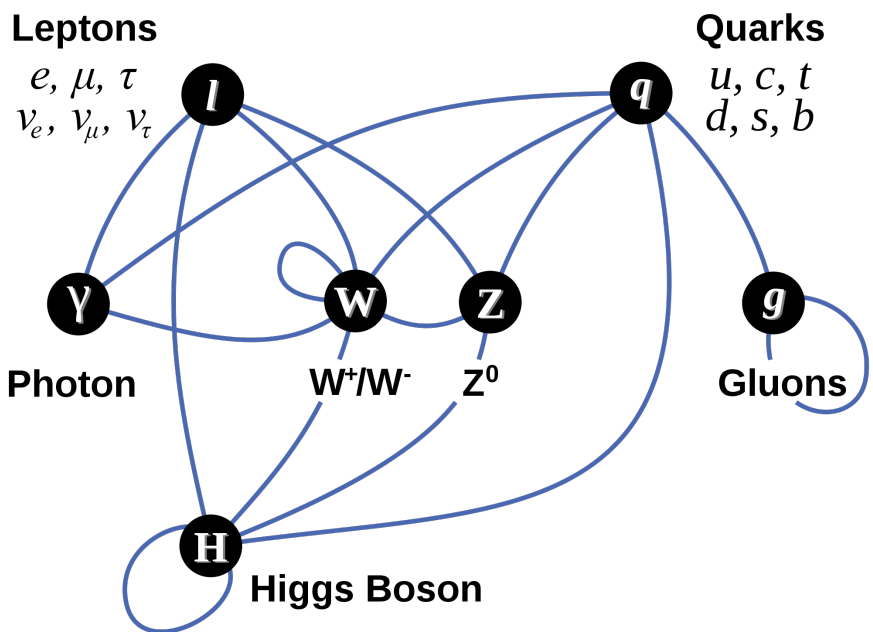


Figure 1.4: Summary of interactions between particles described in the Standard Model.

1.1.4 Spin

Spin plays a very fundamental role in quantum mechanics (and consequently in particle physics) and deserves an explicit discussion. Note that the particle content of the Standard Model is divided by spin: spin-1/2 fermions comprise the particles that make matter, while forces are composed of spin-1 vector bosons, and the scalar Higgs boson is responsible for generating mass. Spin, or intrinsic angular momentum, is a fundamental property of elementary particles, and dictates the statistical behavior of collections of particles due to the *spin-statistics theorem*, which states: *The wavefunction* $\psi(x_1, ..., x_n)$ *of any system of identical bosons (fermions) is symmetric (antisymmetric) under the exchange of any two particles.*

There are a number of important consequences of this theorem. Pauli's exclusion

principle, that no two fermions can occupy the same state, follows directly, since

$$\psi(x_1, x_2) = -\psi(x_2, x_1) \Rightarrow \psi(x_1, x_1) = 0 \tag{1.1}$$

which, fundamentally, explains why ordinary bulk matter is stable and occupies volume. Conversely, bosons are not subject to this exclusion principle. This is why it is possible to formulate superposition principles for macroscopic forces such as electricity and gravity, and accounts for why macroscopic forces exist in the first place.

The spin of a particle is directly related to its *chirality* or "handedness". In the limit of a massless particle, chirality is identical to helicity, which is the sign of the projection of the spin vector onto the momentum vector of the particle (a negative sign corresponds to left-handed helicity, a positive sign corresponds to right-handed helicity). Chirality is important in particle physics because it has been observed in experiment that only left-handed fermions (or right-handed antifermions) interact via the weak force. Thus, the theory of weak interactions is a *chiral* theory, which is to say, a theory in which the couplings of particles differ depending on their chirality.

1.2 Physics of Top Quark Pairs

At the time of this writing, the top quark is the last fundamental particle to have been discovered experimentally. At $172.0 \pm 2.2\,\mathrm{GeV}/c^2$ (roughly the mass of a tungsten atom) it is also the heaviest observed elementary particle. The predicted mean lifetime for the top quark is approximately 5×10^{-25} s, which is about 20 times shorter than the timescale for strong interactions. As a result, the top quark does not form hadrons. Its existence was demonstrated in 1996 at the Tevatron [5, 6], and prior to the start of LHC operation, all direct measurements of top production and decay were performed by either the CDF or D0

experiments.

An understanding of the top quark properties, including its production cross section, is vital for new physics searches for several reasons. Many new physics models, including a variety of SUSY models, predict events with "top-like" signatures, and events with top pairs represent the most tenacious background for these searches. Moreover, the high mass of the top quark means that it couples more strongly to many undiscovered particles, particularly the Higgs boson. Because of this, high-mass resonances in top-pair production represent an excellent channel for the discovery of new physics.

1.2.1 Production of Top Quark Pairs

In particle physics, the likelihood of an interaction between particles is often expressed, rather abstractly, in the form of an area, called a *cross section*. In the case of the protons being collided, the cross section represents the "effective area" of the proton, analogous to the cross sectional area of a classical object. Physics processes, such as the production of $t\bar{t}$ quark pairs, are said to have a production cross section. This is some fraction of the total interaction cross section of the initial colliding protons, and represents the probability of this process occurring. It is a particularly convenient quantity in physics, as it can be multiplied by the luminosity L (*i.e.* the total number of interactions in a given time period) to obtain the event rate:

$$\text{Rate} = \frac{dN}{dt} = \sigma L \tag{1.2}$$

Figure 1.5 shows the cross sections for various physics processes, and how they change as a function of the center-of-mass energy of the colliding particles: $\sqrt{s} = \sqrt{p_1^2 + p_2^2}$. In general, the cross sections are functions of the coupling strength of their corresponding forces. For this reason, QCD decays, which proceed through the strong force, have the

highest cross sections.

Top quark production at hadron colliders is dominated by the production of $t\bar{t}$ pairs [7]. At the Tevatron, top pairs were primarily produced through quark-antiquark annihilation, but the dominant mechanism for producing top pairs at the LHC is through gluon fusion: $gg \to t\bar{t}$ [8].

1.2.2 Semileptonic Decays of Top Quark Pairs

Top quarks decay almost exclusively to a W boson and a b quark, with very small contributions from the (heavily suppressed) decays $t \to Ws$ and $t \to Wd$. At next-to-leading order, the width predicted by the SM for top quark decay is

$$\Gamma_t = \frac{G_F m_t^3}{8\pi\sqrt{2}} \left(1 - \frac{M_W^2}{m_t^2}\right)^2 \left(1 + 2\frac{M_W^2}{m_t^2}\right) \left[1 - \frac{2\alpha_s}{3\pi}\left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right] \tag{1.3}$$

where m_t is the pole mass of the top quark [9]. For a pole mass $m_t = 171\,\text{GeV}/c^2$, the width is $1.29\,\text{GeV}/c^2$, which gives a corresponding lifetime 0.5×10^{-24} s. This short lifetime is the reason $t\bar{t}$ pairs fail to form bound states.

Because top quarks decay almost exclusively into b quarks, the decay modes can be classified into three general types according to the decays of the W boson, as indicated in Figure 1.6:

- The all-hadronic decay mode can yield up to six jets from quark decays, two of them from b quarks:

$$t\bar{t} \to W^+ b W^- \bar{b} \to q\bar{q}' b q'' \bar{q}''' \bar{b} \qquad (46.2\%) \tag{1.4}$$

- In semileptonic decays one W decays leptonically and the other decays hadronically:

$$t\bar{t} \to W^+ b W^- \bar{b} \to q\bar{q}' b l \bar{\nu}_l \bar{b} + \bar{l}\nu_l b q\bar{q}' \bar{b} \qquad (43.5\%) \tag{1.5}$$

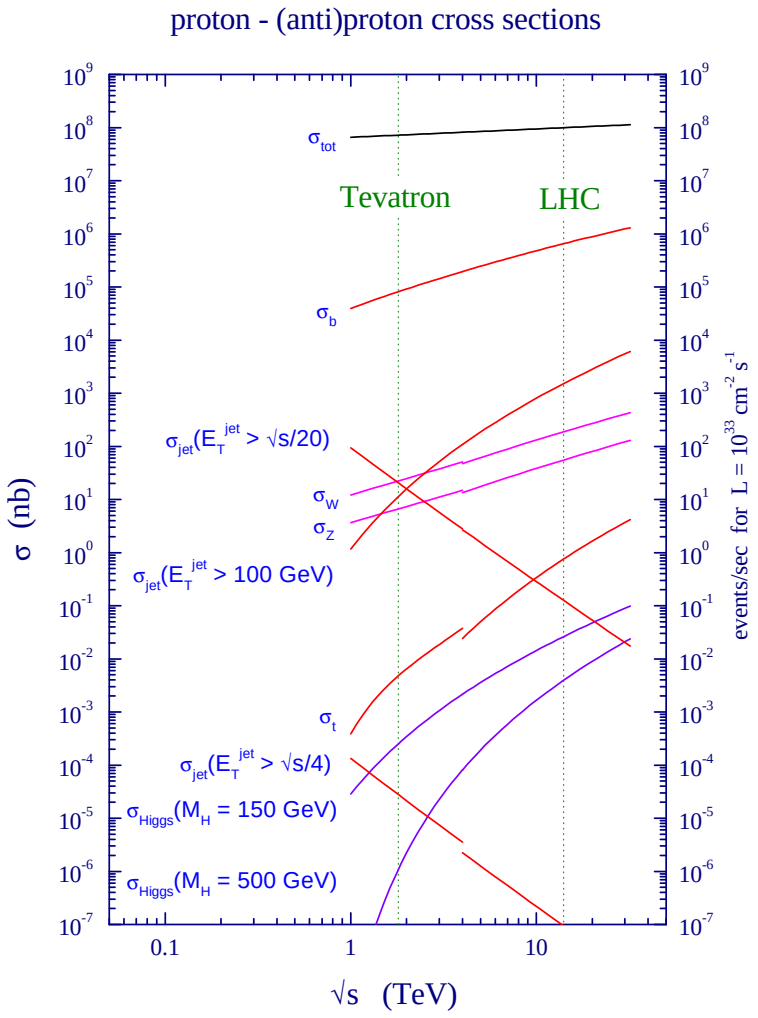


Figure 1.5: Production cross sections as a function of center-of-mass energy for several processes at the LHC (pp collider) and Tevatron ($p\overline{p}$ collider). The cross section for top pair production is approximately 160 pb at $\sqrt{s} = 7\,\text{TeV}$ the current operational energy of the LHC.

- For the all-leptonic decay, both W bosons decay leptonically, giving this decay the lowest branching fraction:

$$t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow \bar{l}\nu_l b l'\bar{\nu}_{l'}\bar{b} \qquad (10.3\%) \qquad (1.6)$$

For the top pair production cross section measurement, this analysis relies on the semileptonic decay of the top quarks. The final state is therefore expected to contain four jets from quarks (2 from b quarks), one lepton (in this case a muon) and missing energy from the neutrino, which escapes the detector without interacting.

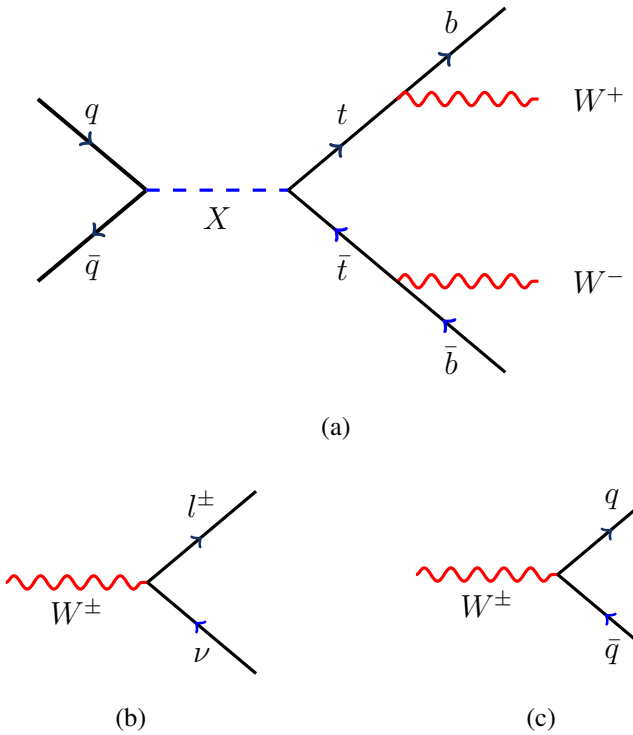


Figure 1.6: Example diagram for (a) top production with subsequent W-boson (b) leptonic decay and (c) hadronic decay.

1.3 Problems with the Standard Model

The Standard Model can be credited with considerable successes, including predicting the existence of the $W^{\pm}$ and Z^0 bosons, the gluons, and the top and charm quarks. Quantitatively, the SM has yielded extremely accurate predictions for the properties of these particles. For example, the masses of the $W^{\pm}$ and Z^0 bosons were predicted to be 80.390 ± 0.018 GeV and 91.1874 ± 0.0021 GeV, respectively. When discovered, their masses were measured to be 80.398 ± 0.025 GeV and 91.1876 ± 0.0021 GeV, in almost exact agreement with prediction.

Despite it's predictive power, the Standard Model faces serious problems. It cannot explain the repeating pattern of generations observed in matter, including why the masses of the particles are what they are, why there are three generations, and why generations exist in the first place. Further, the Standard Model does not include the force of gravity, and no self-consistent quantum theory of gravity has yet been formulated.

Of particular importance to this thesis are the so-called "hierarchy problem" [10] and the inability of the SM to explain the presence of dark matter in the universe. The hierarchy problem concerns astonishingly huge difference between the Planck scale ($M_P = 1/\sqrt{8\pi G_{Newton}} \simeq 2.4 \times 10^{18}$ GeV) and the electroweak scale M_W: $M_P/M_W \sim 10^{16}$. The Higgs boson interacts with every particle with mass, and without an incredible fine-tuning of parameters, loop-order quantum effects from these particles would give enormous corrections to the Higgs mass, driving it from the electroweak scale up to the Planck scale.

To make this more quantitative, consider the Feynman diagram in Figure 1.7, showing a Higgs boson H coupling to a virtual fermion loop f with mass m_f, where the fermion can

represent any lepton or quark in the Standard Model. If the Lagrangian for the interaction of the Higgs with a fermion is given by $-\lambda_f H \bar{f} f$, this process yields a second-order correction to the square of the mass of the Higgs, m_H^2

$$\Delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2}\Lambda_{\mathrm{UV}}^2 + ... \tag{1.7}$$

where Λ_{UV} is the ultraviolet momentum cutoff, representing the scale up to which the Standard Model is valid; beyond this point new physics alters the behavior of the theory. If we naively assume that no new physics enters until we reach the Planck scale, then this second-order correction to m_H^2 will be some *30 orders of magnitude* greater than the final, electroweak-scale value.

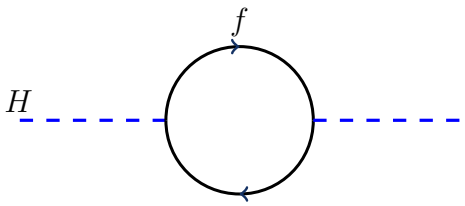


Figure 1.7: One-loop quantum correction to the Higgs squared mass parameter m_H^2 due to a coupling to a hypothetical scalar boson S.

Finally, current cosmological data provides a large amount of evidence for the existence of *dark matter* in the universe [11]. Dark matter does not emit or scatter electromagnetic radiation, and thus has not been directly observed, but from gravitational effects on visible matter it is believed to account for $\sim 80\%$ of the total matter in the universe. Figure 1.8 shows an example of the distribution of dark matter and the distribution of normal, visible matter in an example galaxy, obtained from gravitational lensing, in which the mass of the galaxy bends light from galaxies behind it.

Dark matter is important to particle physics because, despite considerable evidence of

Figure 1.8: Mass distribution of dark matter (blue) and ordinary luminous matter (red) seen in a merger of two galactic subclusters [12].

its existence, there seems to be no particle that exists within the Standard Model that could serve as an appropriate dark matter candidate. Charged leptons and quarks are inadmissible due to their electromagnetic interactions, while neutrinos are too light and would fail to be bound by the gravitational pull of galaxies.

Chapter 2

SUSY and Top Quark Production

> *It is often stated that of all the theories proposed in this century, the silliest is quantum theory. In fact, some say that the only thing that quantum theory has going for it is that it is unquestionably correct.*
>
> — Michio Kaku

When the framework of quantum field theory was originally being formulated and applied to particle physics (in the form of quantum electrodynamics, a quantum theory of the electromagnetic field), it was generally thought of as a complete, fundamental description of nature. It was supposed that field theory could correctly describe the phenomena of particle physics down to *any* scale, without need for further modification. Our modern view of field theory, informed by a close association with condensed matter physics, is very different. Today the quantum field theories that describe the forces of the Standard Model are viewed as *effective* field theories [13], which is to say, the very low energy manifestation of some deeper idea.

It is not necessary to understand everything about this "final" theory in order to make accurate predictions using the effective theory. This is, in a sense, the justification of the technique of *renormalization* [1], the method for consistently removing the infinities and di-

vergences that arise in field theory calculations. For these calculations, we do not integrate up to infinite energy, but rather up to a parameter Λ, which represents the scale at which our approximate theory breaks down. The success of quantum field theory relies on renormalization to guarantee that physical quantities never depend on this parameter Λ, but rather on other physical quantities that can be measured.

Practically, renormalizability is a process of eliminating "unphysical" quantities like, say, the bare mass of the electron (and corrections to this mass) in favor of physical ones, like the measured mass of the electron (which includes all its interactions with the vaccuum). It is possible to cancel the bare mass against the parameter Λ, leaving the actual, measurable electron mass. This final mass contains the bare mass plus all higher-order corrections from loops.

In this sense, the hierarchy problem is not technically a problem with the Standard Model itself, as the one-loop correction (1.7), regardless of its size, can simply be incorporated into the measured Higgs mass. However, given that the SM is in fact an effective low-energy limit, the Higgs mass is disturbingly sensitive to almost any imaginable new physics. Indeed, if the Higgs mass is calculable in this new theory, an incredible degree of fine-tuning in the coupling of the Higgs to other particles would seem to be required to get all higher-order contributions to the Higgs mass to precisely cancel.

2.1 Supersymmetry and Mass Spectrum of the MSSM

Even assuming that the ultraviolet cutoff Λ_{UV} in equation (1.7) is not too large doesn't necessarily fix the problem. In order to keep Δm_H^2 small, whatever new physics exists at this scale would actually have to completely cut off the loop integral. Consider a heavy complex scalar particle S with a mass m_S. If the Lagrangian term describing the interaction

of S with the Higgs is $-\lambda_S |H|^2 |S|^2$, then this new particle gives a loop-order correction to the Higgs mass of

$$\Delta m_H^2 = \frac{\lambda_S}{16\pi^2}\left[\Lambda_{\mathrm{UV}}^2 - 2m_S^2 \ln\left(\frac{\Lambda_{\mathrm{UV}}}{m_S}\right) + ...\right] \tag{2.1}$$

The Feynman diagram corresponding to this process is shown in Figure 2.1.

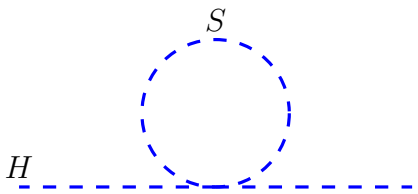


Figure 2.1: One-loop quantum correction to the Higgs squared mass parameter m_H^2 due to a coupling to a hypothetical scalar boson S.

A comparison of (1.7) to (2.1) shows that both leading terms are proportional to Λ_{UV}^2, but with a relative minus sign between the fermion and boson contributions. This suggests an immediate solution to the hierarchy problem: if there exists a pair of complex scalars with $\lambda_S = |\lambda_f|^2$ corresponding to each fermion in the Standard Model, then each contribution to the Higgs mass from its couplings to fermions could be systematically cancelled against the contributions from these scalar bosons [14]. Of course, the existence of a boson state corresponding to each fermion state would imply the presence of a symmetry; specifically, a symmetry relating fermions and bosons, called *supersymmetry* (SUSY).

Formally constructing the rules for supersymmetry is beyond the scope of this paper (see, for example [10]), but we will state certain general results. If a supersymmetric transformation is one that converts a fermion state into a boson state, then up to some overall phase

$$Q\,|\text{Boson}\rangle = |\text{Fermion}\rangle \qquad\qquad Q\,|\text{Fermion}\rangle = |\text{Boson}\rangle \tag{2.2}$$

where Q is the generator of supersymmetric transformations. Since it changes the spin of a field by $1/2$, Q must itself be a fermionic spinor, carrying an intrinsic angular momentum of $1/2$. Such a generator will not commute with Lorentz transformations (unlike other generators of internal symmetries), so in this sense supersymmetry must be a spacetime symmetry.

Without spinor indices, the supersymmetry algebra [15] reads

$$\left\{Q, Q^\dagger\right\} = P^\mu \tag{2.3}$$

$$\{Q, Q\} = \left\{Q^\dagger, Q^\dagger\right\} = 0 \tag{2.4}$$

$$[P^\mu, Q] = \left[P^\mu, Q^\dagger\right] = 0 \tag{2.5}$$

where P^μ is the four-momentum generator of spacetime translations. Single-particle states in a supersymmetric theory are grouped into irreducible representations of this supersymmetry algebra, called *supermultiplets*, with each supermultiplet containing both fermion and boson states, called *superpartners* of each other.

From this algebra we can immediately deduce a couple of important results. Suppose the two states $|b\rangle$ and $|f\rangle$, with respective masses m_b and m_f, are members of the same supermultiplet (that is, $Q|b\rangle = |f\rangle$). Since $P^\mu P_\mu |b\rangle = m_b^2 |b\rangle$ and $P^\mu P_\mu |f\rangle = m_f^2 |f\rangle$, it follows from (2.5) that

$$P^\mu P_\mu Q |b\rangle = P^\mu P_\mu |f\rangle = m_f^2 |f\rangle \tag{2.6}$$

$$= Q P^\mu P_\mu |b\rangle = m_b^2 Q |b\rangle = m_b^2 |f\rangle \tag{2.7}$$

$$\Rightarrow m_b = m_f \tag{2.8}$$

which is to say, in (unbroken) supersymmetry, superpartners must have equal masses. This is an important result, because experimentally no scalar particle has been found with a mass

equal to, for example, the electron. Thus, if supersymmetry is a symmetry of nature, it must be spontaneously broken. This gives rise to the mass differences between particles and their superpartners in much the same way that electroweak symmetry breaking gives rise to the masses of the gauge bosons.

2.1.1 The MSSM

One fundamental requirement of any supersymmetric theory is that it must have the same number of bosonic and fermionic degrees of freedom. In order to be phenomenologically viable, the theory must also be consistent with the known features of the Standard Model, including the presence of chiral fermions (fermions whose right-handed and left-handed states couple differently to different gauge fields) and parity violation, which further restricts its form.

There are two basic supermultiplets in viable supersymmetric theories: *Chiral* supermultiplets contain one fermion (with two spin helicity states) and one complex scalar boson (also with two states, one for each component). *Gauge* supermultiplets contain one vector boson (with two spin states; in the unbroken theory this boson is massless, and hence has no transverse polarization) and one spin-1/2 fermion (again with two spin helicity states). There are other possible combinations of particles with equal degrees of fermionic and bosonic degrees of freedom, but these can always be reduced to some combination of chiral and gauge supermultiplets.

If the Standard Model can be extended by supersymmetry then all the particles in the SM must be a member of either a chiral or a gauge supermultiplet and have a superpartner with a spin difference of 1/2 [16]. All of the Standard Model fermions have left-handed pieces that transform differently from their right-handed pieces, so they must be members

Names		spin 0	spin $1/2$
squarks, quarks	Q	$\left(\widetilde{u}_L\, \widetilde{d}_L\right)$	$(u_L\, d_L)$
($\times$ 3 families)	$\overline{u}$	$\widetilde{u}^*_R$	$u^\dagger_R$
	$\overline{d}$	$\widetilde{d}^*_R$	$d^\dagger_R$
sleptons, leptons	L	$(\widetilde{\nu}\, \widetilde{e}_L)$	$(\nu\, e_L)$
($\times$ 3 families)	$\overline{e}$	$\widetilde{e}^*_R$	$e^\dagger_R$
Higgs, higgsinos	H_u	$(H^+_u\, H^0_u)$	$\left(\widetilde{H}^+_u\, \widetilde{H}^0_u\right)$
	H_d	$\left(H^0_d\, H^-_d\right)$	$\left(\widetilde{H}^0_d\, \widetilde{H}^-_d\right)$

Table 2.1: Chiral (matter) supermultiplets in the Minimal Supersymmetric Standard Model.

of chiral supermultiplets, and hence have scalar superpartners. These are generically called *squarks* and *sleptons* (collectively, *sfermions*), and are denoted with a tilde (e.g. $\widetilde{e}$, $\widetilde{\mu}$, etc). They are generally written with an L or R subscript; this does not refer to the helicity of the sfermion (which is a scalar particle), but rather to the helicity of its Standard Model partner. The gauge interactions of the sfermion fields are the same as for their corresponding partners; left-handed sfermions couple to the weak interaction, whereas right-handed sfermions do not.

The Higgs boson is a special case. In order to ensure that the electroweak interaction is anomaly-free in a supersymmetric extension of the Standard Model, there must be two Higgs supermultiplets [17]. Further, because of the additional degrees of freedom in supersymmetric theories, it is necessary to postulate two Higgs bosons with different quantum numbers. One of these gives mass to quarks of charge $+2/3$ and is therefore labeled H_u while the other gives mass to quarks of charge $-1/3$ and the charged leptons and is labeled H_d. Both doublets contain a charged and neutral component, denoted (H^+_u, H^0_u) and $\left(H^0_d, H^-_d\right)$, respectively.

The SM vector bosons are members of gauge supermultiplets, and their fermionic

Names	spin $1/2$	spin 1
gluino, gluon	$\widetilde{g}$	g
winos, W bosons	$\widetilde{W}^{\pm}\,\widetilde{W}^{0}$	$W^{\pm}\,W^{0}$
bino, B boson	$\widetilde{B}^{0}$	B^{0}

Table 2.2: Gauge supermultiplets in the Minimal Supersymmetric Standard Model$S$.

superpartners are called "gauginos". The gauge eigenstates of the strong force carriers are gluons (g), and their superpartners are *gluinos* ($\widetilde{g}$). Similarly, the gauge eigenstates of the electroweak force carriers (distinct from the *mass* eigenstates shown in Figure 1.1 due to symmetry breaking) are $W^{\pm}$, W^{0}, and B^{0}. The corresponding supersymmetric spin-$1/2$ states are $\widetilde{W}^{\pm}$, $\widetilde{W}^{0}$, and $\widetilde{B}^{0}$, named *winos* and *binos*. The physical states Z^{0} and γ are mixtures of W^{0} and B^{0}, with Z^{0} acquiring mass via the Higgs mechanism, and their superpartners (mixtures of $\widetilde{W}^{0}$ and $\widetilde{B}^{0}$) are the *zino* ($\widetilde{Z}^{0}$) and *photino* ($\widetilde{\gamma}$).

Collectively, these chiral and gauge supermultiplets comprise the content of the Minimal Supersymmetric Standard Model (MSSM), and are summarized in Tables 2.1 and 2.2. In the case of Table 2.1, the particles and their superpartners are grouped into *superfields* (labeled Q_i, L_i, etc, where $i = 1, 2, 3$ is a family index).

The presence of these new particles in supersymmetry poses an experimental problem: there now exist renormalizable gauge couplings in the theory that can violate baryon and lepton number conservation, in contrast to the Standard Model couplings. In particular, for coupling constants λ, there are additional allowed terms in the superpotential

$$W_{\Delta B=1} = \lambda \overline{u}\overline{d}\overline{d} \tag{2.9}$$

$$W_{\Delta L=1} = \lambda' L \cdot L\overline{e} + \lambda'' L \cdot Q\overline{d} \tag{2.10}$$

The first term violates baryon number B by 1, and the next two terms violate lepton number

L by 1. All of the couplings corresponding to these potential terms involve either one or three sfermions. The general form of the single-sfermion interactions is shown in Figure 2.2. Given the precise experimental limits on violation of B and L [18, 19], it is necessary that these terms have very small amplitudes.

One solution to this problem is to postulate that these B- and L-violating couplings are forbidden entirely, by a new multiplicative symmetry called R-parity (or sometimes "matter" parity), defined for each particle as

$$P_R = (-1)^{3(B-L)+2s} \tag{2.11}$$

where s is the spin of the particle. Quarks (and squarks) have $B = 1/3$, $L = 0$, while leptons (and sleptons) have $B = 0$, $L = 1$. Since normal matter particles have $s = 1/2$ and their scalar superpartners have $s = 0$, their R-parities will differ by a factor of -1: quarks and leptons have $P_R = +1$ while squarks and sleptons have $P_R = -1$. If R-parity is indeed an exact symmetry of nature (that is, the product of P_R for all final-state particles is equal to the product of P_R for all initial-state particles), then vertex terms such as those in Figure 2.2 are prohibited, as they intrinsically have overall $P_R = -1$ and must therefore flip the R-parities of the initial and final states.

There are several consequences of R-parity conservation. First, since supersymmetric particles (or "sparticles") have opposite R-parity from normal matter particles, there can be no mixing between them. Also, any allowed vertex in the theory must have an even number of sparticles, which preserves the necessary terms of the SUSY Lagrangian while prohibiting the B- and L-violating terms of Eqs. 2.9 and 2.10.

This fact has several very important phenomenological consequences. First, the lightest supersymmetric particle (called the LSP) must be *completely* stable. Sparticles with

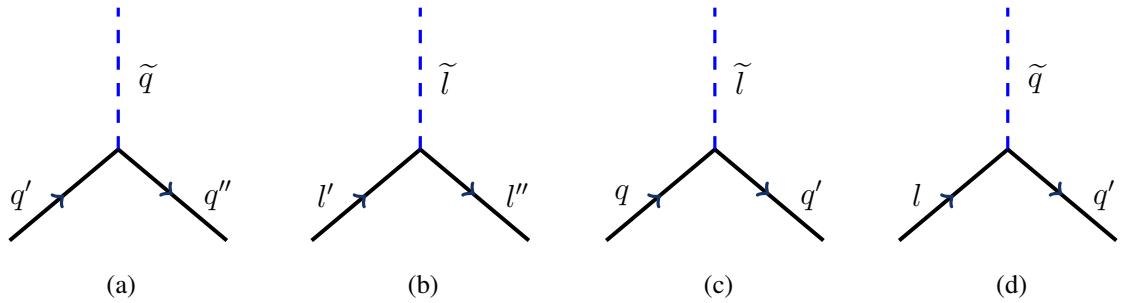


Figure 2.2: Renormalizable R-parity violating couplings of the MSSM involving a single sfermion. (a) Vertex corresponding to superpotential term $\lambda\overline{u}\overline{d}\overline{d}$ violating B by 1, (b) vertex corresponding to $\lambda' L \cdot L\overline{e}$ violating L by 1, (c) and (d) correspond to $\lambda'' L \cdot Q\overline{d}$, also violating L by 1. In addition, there are analogous couplings involving 3 fermions that violate B and L conservation.

higher mass can decay into the LSP, but the lightest sparticle would have to decay into normal particles, which is prohibited by R-parity. This is especially important in the context of the dark matter problem if the LSP is electrically neutral: massive, weakly-interacting, and absolutely stable, the LSP provides a very attractive dark matter candidate [20].

Conservation of R-parity provides distinct experimental signatures. Each higher-mass sparticle must eventually decay into a state that contains an odd number of LSPs (typically one). Further, if SUSY is accessible to pp collisions at the LHC, then the sparticles from the collision must be produced in pairs. Taken together, these facts imply that any experimental signature of SUSY will include a pair of massive LSPs in the final state.

The gauge eigenstates of the MSSM are summarized in Tables 2.1 and 2.2, but these will not, in general, be equal to the mass eigenstates of the theory, which will depend greatly on the details of how SUSY is broken. (This is analogous to the situation in the Standard Model: due to electroweak symmetry breaking, the gauge eigenstates of the electroweak

theory, W^0 and B^0, mix to form the mass eigenstates γ and Z^0.) In general, mixing can occur between gauginos and higgsinos with the same charge. There is also mixing between various sets of sfermions with the same charge, analogous to the quark mixing of the CKM matrix of the Standard Model [21], which relates the mass eigenstates of the quarks to their gauge eigenstates. In SUSY, only the gluino is exempt from this, as its color charge prohibits it from mixing with any other particles.

The neutral higgsinos ($\widetilde{H}_u^0$ and $\widetilde{H}_d^0$) mix with the neutral gauginos ($\widetilde{B}^0$ and $\widetilde{W}^0$) to form four mass eigenstates called *neutralinos*, denoted $\widetilde{\chi}_i^0$ (with index i = 1, 2, 3, 4). The charged higgsinos ($\widetilde{H}_u^+$ and $\widetilde{H}_d^-$) combine with the charged winos ($\widetilde{W}^\pm$) to form two mass eigenstates called *charginos*, denoted $\widetilde{\chi}_i^\pm$ (with index i = 1, 2). The labels go in ascending order of mass, such that the LSP (often assumed to be a neutralino) would be $\widetilde{\chi}_1^0$.

Experimental limits on sparticle contributions to flavor violating processes (*e.g.* [22, 23]) place stringent limits on possible mixing among squarks and sleptons, particularly for the first two generations, so the mass eigenstates of most sfermions are equivalent to those of the normal fermions (that is, $\widetilde{e}_R$, $\widetilde{e}_L$, $\widetilde{\mu}_R$, $\widetilde{\mu}_L$, etc). The exception to this is the third generation of sfermions, which can have substantial mixing between the left- and right-handed states. Hence, these are distinguished by an index 1 or 2 rather than L or R: $\widetilde{t}_1$, $\widetilde{t}_2$, $\widetilde{b}_1$, $\widetilde{b}_2$, $\widetilde{\tau}_1$, $\widetilde{\tau}_2$. The relationship between the gauge eigenstates and mass eigenstates in the MSSM is shown in Table 2.3.

2.1.2 Experimental Leptonic Signature and Backgrounds

This analysis includes a search for new physics based on final states including two same-sign leptons (here defined as either electrons or muons) [24, 25, 26]. This can occur if the gluino decays with a significant branching fraction to hadrons and a chargino, which

Names	Spin	P_R	Gauge eigenstates	Mass eigenstates
Higgs bosons	0	+1	$H_u^0\ H_d^0\ H_u^+\ H_d^-$	$h^0\ H^0\ A^0\ H^\pm$
squarks	0	–1	$\widetilde{u}_L\ \widetilde{u}_R\ \widetilde{d}_L\ \widetilde{d}_R$ $\widetilde{c}_L\ \widetilde{c}_R\ \widetilde{s}_L\ \widetilde{s}_R$ $\widetilde{t}_L\ \widetilde{t}_R\ \widetilde{b}_L\ \widetilde{b}_R$	(same) (same) $\widetilde{t}_1\ \widetilde{t}_2\ \widetilde{b}_1\ \widetilde{b}_2$
squarks	0	–1	$\widetilde{e}_L\ \widetilde{e}_R\ \widetilde{\nu}_e$ $\widetilde{\mu}_L\ \widetilde{\mu}_R\ \widetilde{\nu}_\mu$ $\widetilde{\tau}_L\ \widetilde{\tau}_R\ \widetilde{\nu}_\tau$	(same) (same) $\widetilde{\tau}_1\ \widetilde{\tau}_2\ \widetilde{\nu}_\tau$
neutralinos	1/2	–1	$\widetilde{B}^0\ \widetilde{W}^0\ \widetilde{H}_u^0\ \widetilde{H}_d^0$	$\widetilde{\chi}_1^0\ \widetilde{\chi}_2^0\ \widetilde{\chi}_3^0\ \widetilde{\chi}_4^0$
charginos	1/2	–1	$\widetilde{W}^\pm\ \widetilde{H}_u^+\ \widetilde{H}_d^-$	$\widetilde{\chi}_1^\pm\ \widetilde{\chi}_2^\pm$
gluino	1/2	–1	$\widetilde{g}$	(same)

Table 2.3: Relationship of gauge eigenstates to mass eigenstates of particles in the MSSM.

subsequently decays into a lepton, a neutrino and a neutralino LSP. The gluino is a Majorana particle (*i.e.* a particle that is its own antiparticle) and hence has no "preferred" charge in its decays. It is thus equally likely to produce a lepton of either charge in this example. This final state can also come about through squark production, for example via $\widetilde{q} \to q\widetilde{g}$. In either case, this decay can lead to final states with same-sign leptons (though possibly different flavors) along with jets (from the decay of the initial colored SUSY particles) and missing energy (carried away by the massive LSPs and the neutrinos). The same-sign dilepton final state is particularly attractive because it has so few Standard Model backgrounds: the primary sources of lepton pairs in the SM come from Drell-Yan and $t\bar{t}$ production, both of which produce opposite-sign leptons.

The specific distribution of the masses of the SUSY particles, and hence the exact decay channels through which an interaction can proceed, depend heavily on the choice of SUSY-breaking mechanism, as well as the particular values one chooses for the as-yet-unmeasured free parameters of the system. To illustrate the possible SUSY processes ca-

pable of producing our signature same-sign dilepton final state, we assume a typical mass spectrum with three characteristic scales:

- The most massive particles are the squarks and gluinos, with a mass scale $m_{\widetilde{q}}$.
- An intermediate mass scale for the charginos, denoted $m_{\widetilde{\chi}^{\pm}}$.
- The LSP mass, assumed here to be a neutralino, $m_{\widetilde{\chi}^{0}}$.

We also take $m_{\widetilde{g}} > m_{\widetilde{q}}$, such that the decay $\widetilde{g} \to \widetilde{q}q$ is kinematically allowed. Given this mass spectrum, Figure 2.3 illustrates several possible decay modes that lead naturally to final states with same-sign dileptons. Hadronic activity in the event comes from the quarks produced by the gluino and squark decays, while the massive neutralinos give rise to the missing energy. Figure 2.3 (d) shows a typical semileptonic squark cascade decay, with the lepton produced from the decay of the chargino. Because there are two such squark decays in each event, we expect that a significant fraction of the time they should produce same-sign dileptons.

2.2 Short Overview of Previous Results

The CDF same-sign analysis [27] is a search for new physics using $1\ \text{fb}^{-1}$ of data from $p\overline{p}$ collisions at a center-of-mass energy of $1.96\,\text{TeV}$. It considers two separate selections: the *inclusive* selection imposes no further restrictions beyond requiring two high momentum same-sign leptons, while the *tight* selection adds a Z boson veto (no two same-flavor, opposite-sign leptons in the event with an invariant mass $66\,\text{GeV}/c^2 < m_{ll} < 116\,\text{GeV}/c^2$ and a requirement on the missing transverse energy, $\not\!\!E_T > 15\,\text{GeV}$ (defined in Chapter 5). These requirements are particularly useful for removing background events resulting from the process $Z/\gamma^* \to l^+l^-$, where the charge of one lepton is mismeasured.

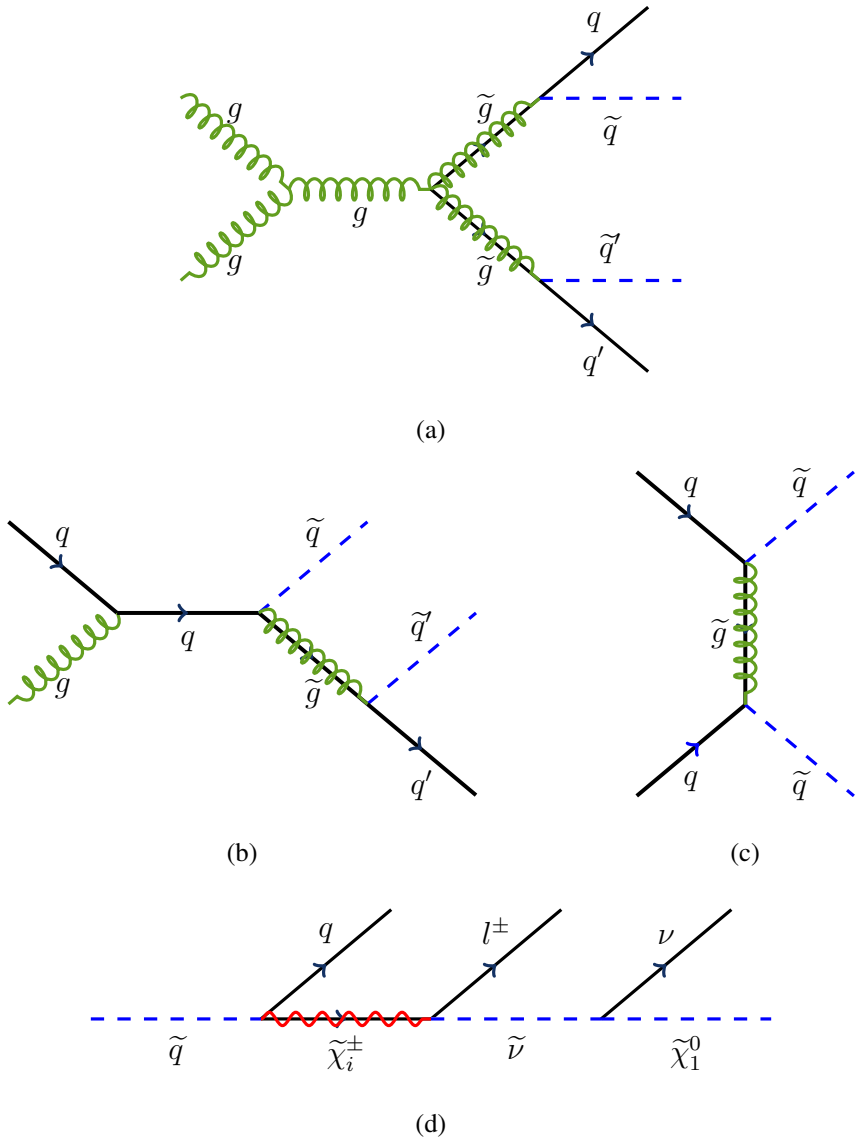
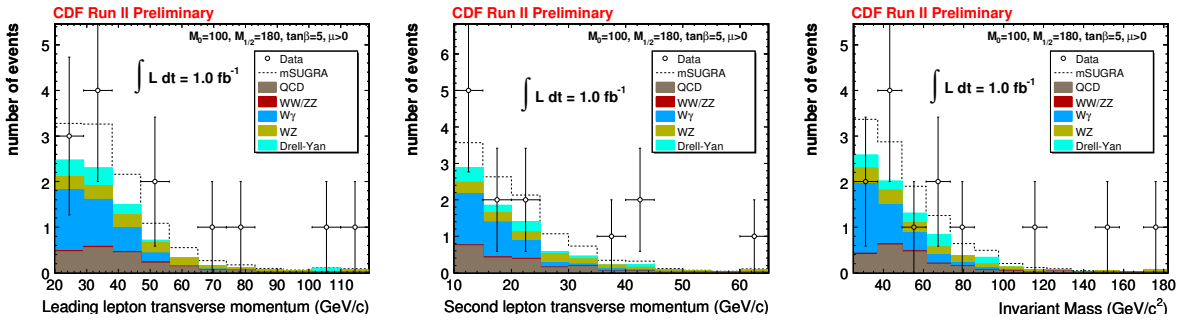


Figure 2.3: Example diagrams for (a) gluino-gluino production, (b) gluino-squark production and (c) squark-squark production, as well as (d) an example squark cascade decay.

The result of applying the tight selection is shown in Figure 2.4. The first two plots show the transverse momenta of the of the two leptons, while the third shows the resulting invariant mass from combining the two (with the Z veto and $\not\!\!E_T$ requirements removed). In total, 7.9 ± 1.0 events are expected to pass the tight selection solely from SM sources, and 13 events are observed in data. For the inclusive selection, 33.7 ± 3.5 SM events are expected to pass the selection requirements, and 44 are observed. In both cases, the excess of events is observed in the high momentum region of the leading lepton distribution.



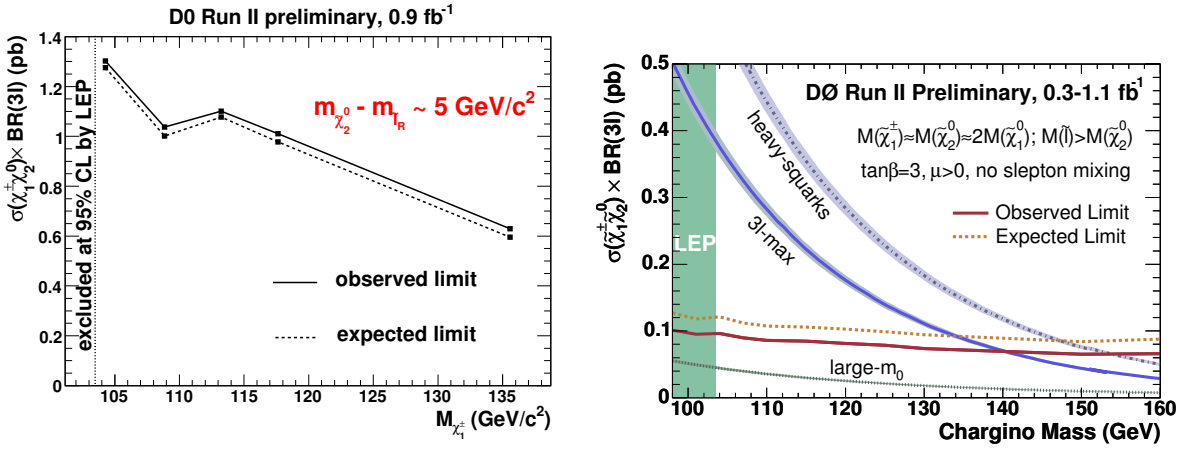
(a) (b) (c)

Figure 2.4: CDF results for (a) Transverse momentum of the leading lepton, (b) transverse momentum of the sub-leading lepton and (c) invariant mass of the lepton pair.

The corresponding result from D0 [28] restricts its focus to events with two same sign muons. The analysis uses approximately $1\,\mathrm{fb}^{-1}$ and considers multiple scenarios with the MSSM model broken by gravity effects (called *minimal supergravity*, or *mSUGRA* [29]) in order to place a limit on chargino and neutralino production. Events are selected with two well-isolated muons with matching charge and transverse momenta of at least $13\,\mathrm{GeV}/c$ for the leading muon and at least $8\,\mathrm{GeV}/c$ for the sub-leading muon. Multiple additional requirements are applied to reduce the contribution from SM background events such as

dibosons, $W \to \mu\nu$ and $Z/\gamma^* \to l^+l^-$.

This selection is an excellent channel to search for the process $\sigma(p\bar{p} \to \widetilde{\chi}_1^{\pm}\widetilde{\chi}_2^0) \times BR(3l)$ for scenarios where $m_{\widetilde{l}_R} \lesssim m_{\widetilde{\chi}_2^0}$, because the third lepton typically has a very soft momentum spectrum. Limits on this quantity are shown in Figure 2.5 as a function of the mass of the $\widetilde{\chi}_1^{\pm}$, with the region above the curve excluded at the 95% confidence level [28].



(a) (b)

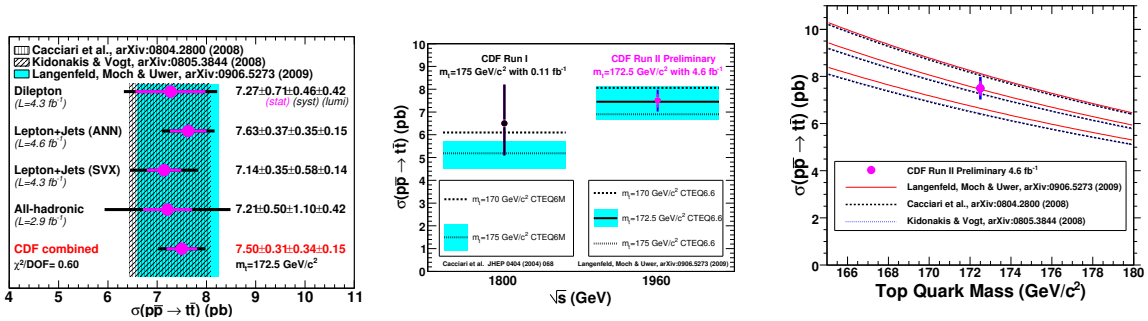
Figure 2.5: Limits on $\sigma(p\bar{p} \to \widetilde{\chi}_1^{\pm}\widetilde{\chi}_2^0) \times BR(3l)$ for $m_{\widetilde{l}_R} \lesssim m_{\widetilde{\chi}_2^0}$ mSUGRA points. The $\widetilde{\chi}_1^{\pm}$ mass region below $103.5\,\mathrm{GeV}/c^2$ is excluded by LEP. The region above the curve is excluded at 95% CL. (a) indicates the limit set in the same-sign dimuon channel, while (b) shows the limit obtained by combining this result with the results of trilepton searches (eel, $e\mu l$ and $\mu\mu l$).

CDF measures the production cross section of $t\bar{t}$ events in the lepton + jets channel through the use of a "neural network" which discriminates between top pair production and background processes [30]. The ratio $\sigma(t\bar{t})/\sigma(Z)$ is computed, and this is multiplied by the theoretical Z cross section, in effect replacing the uncertainty in luminosity with the theoretical uncertainty on the Z cross section.

The measured top pair production cross section at CDF (assuming a top mass of $172.5\,\mathrm{GeV}/c^2$) is [30]

$$\sigma_{t\bar{t}} = 7.63 \pm 0.37(\mathrm{stat}) \pm 0.35(\mathrm{syst}) \pm 0.15(\mathrm{Z\ theory})\mathrm{pb} \tag{2.12}$$

The result of the calculation is shown in the left plot of Figure 2.6, along with the results from multiple other channels. The middle plot shows the combined result as a function of the center-of-mass collision energy, and the right plot shows the theoretical prediction for the $t\bar{t}$ cross section as a function of the top mass, along with the measured value.



(a) (b) (c)

Figure 2.6: (a) Combination of top quark production cross section results from various CDF analyses, (b) combination of cross section results compared to theoretical predictions as a function of $\sqrt{s}$ and (c) combination of cross section results compared to theoretical predictions as a function of top quark mass. A top quark mass value of 172.5 GeV is assumed for the experimental result.

The top pair production cross section measurement for D0 combines the use of kinematic information and an algorithm for tagging jets from b quarks in the event [31]. For a

top quark mass of $172.5\,\mathrm{GeV}/c^2$, the measured cross section is

$$\sigma_{t\bar{t}} = 7.78^{+0.77}_{-0.64}(\text{stat + syst + lumi})\text{pb} \tag{2.13}$$

This result is shown in Figure 2.7 along with the expected and theoretical cross sections as a function of top quark mass.

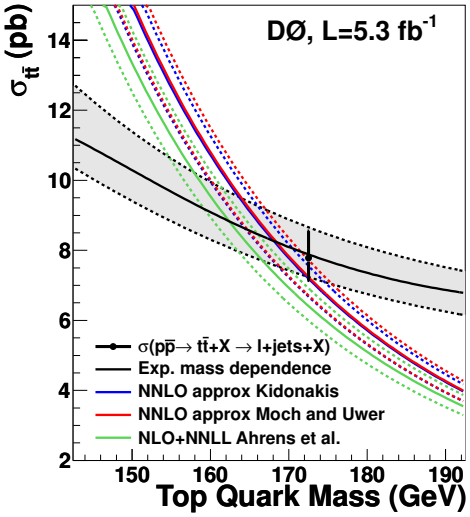


Figure 2.7: D0 combined $t\bar{t}$ production cross section in the lepton + jets channel as a function of top quark mass compared to the theoretical calculations.

Chapter 3

Experimental Setup

Stonehenge really was the most incredible accomplishment. It took five hundred men just to pull each sarsen, plus a hundred more to dash around positioning the rollers. Can you imagine trying to talk six hundred people into helping you drag a fifty-ton stone eighteen miles across the countryside and muscle it into an upright position, and then saying, "Right, lads! Another twenty like that, plus some lintels and maybe a couple of dozen nice bluestones from Wales, and we can party!"

— Bill Bryson

3.1 Large Hadron Collider

The Large Hadron Collider (LHC) is the largest and highest-energy particle accelerator in the world. It is installed in the tunnel originally constructed for the Large Electron-Positron Collider (LEP) between 1984 and 1989. The tunnel is roughly circular, consisting of eight straight sections and eight arcs, with a circumference of 27 km, at a depth ranging from 45 m to 170 m underground, as shown in Figure 3.1. It lies beneath the Franco-Swiss border near Geneva, Switzerland [8].

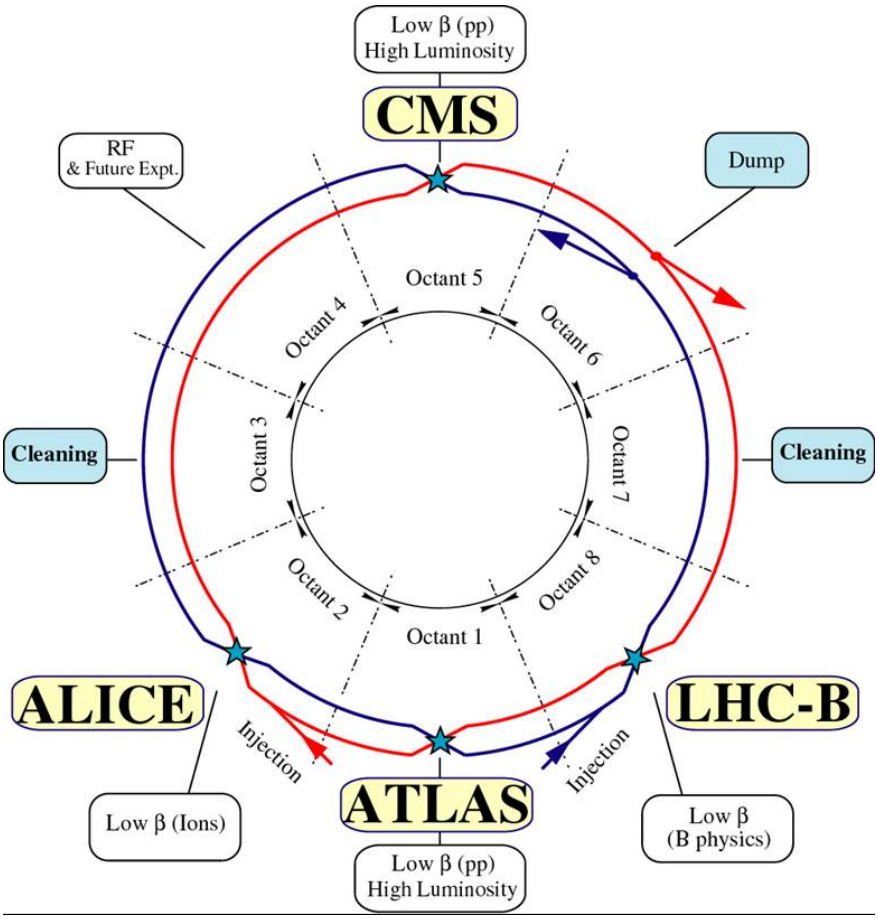


Figure 3.1: Schematic layout of the LHC (Beam 1 – clockwise, Beam 2 – counterclockwise).

The LHC is designed to collide proton beams with a center-of-mass energy of 14 TeV and an instantaneous luminosity of $\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The high expected beam intensities preclude the use of an antiproton beam, which could circulate in the same pipe as the proton beam. Instead, the LHC has two separate parallel rings, each containing a proton beam. These beams rotate in opposite directions and require opposite magnetic dipole fields, and thus two sets of dipole magnets and two separate vacuum systems. However, the LEP tunnel is only 3.8 m wide, which is insufficient for two separate rings of magnets, so the LHC uses twin bore magnets consisting of two sets of coils and two separate beam channels that can be placed within the same structure and can share a common cryostat. The beams then intersect at four different points along the circumference, where the experimental detectors are located.

The benefit of higher luminosity lies in the increased rate of potentially rare events, given by

$$N_{\text{event}} = \mathcal{L}\sigma_{\text{event}} \tag{3.1}$$

where $\mathcal{L}$ is the luminosity and σ_{event} is the cross section for the event. In particular, the luminosity is

$$\mathcal{L} = \frac{kN^2 f}{4\pi\sigma_x^*\sigma_y^*} \tag{3.2}$$

Here, k is the number of bunches (nominally 2,808) in the beam, N is the bunch population (nominally 1.15×10^{11} protons), f is the revolution frequency (11.246 kHz, given protons moving around the 27 km ring at speed $\simeq c$), and σ^* is the size of the beam at the collision point ($\sigma_x^* = \sigma_y^* = 16\ \mu$m).

The beam optics define the size of the beam, as shown in Figure 3.2. The size of the beam oscillates regularly in the arc segments, while in the long straight sections it varies

such that it is most strongly collimated at the collision point. At CMS and ATLAS, the collision point size is 16 μm.

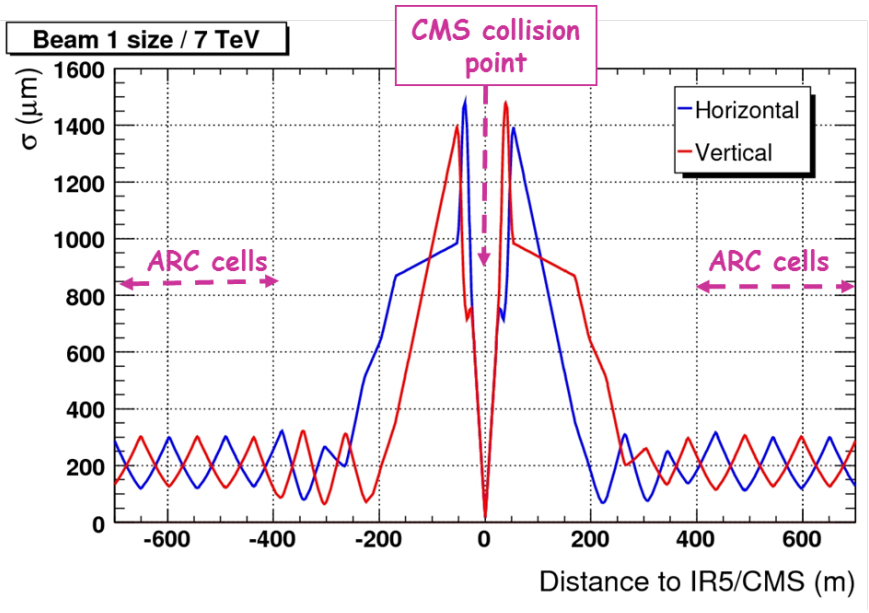


Figure 3.2: Beam size as a function of the distance to the collision point.

The beams are accelerated in stages, with the linear particle accelerator LINAC 2 first generating 50 MeV protons. These protons are then accelerated to 1.4 GeV in the Proton Synchrotron Booster (PSB) and injected into the Proton Synchrotron (PS). There they are accelerated to 26 GeV before being injected into the Super Proton Synchrotron (SPS). Here they are further accelerated to an energy of 450 GeV and then injected into the main ring of the LHC, where they are finally accelerated to their peak beam energy.

The proton beams are not continuous, but rather consist of "bunches" of protons spaced at intervals in a pattern created by combining or splitting bunches at each stage in the injector chain. At design luminosity, 6 "booster" bunches are injected into the PS, where each one is split into 12 smaller bunches, for a total of 72. These bunches are then injected into the SPS; between two and four injections occur, resulting in $(2-4) \times 72 = 144 - 288$

bunches. Finally, a sequence of 12 such collections of bunches from the SPS are injected into the LHC, resulting in a final (nominal) set of 2808 in the LHC. These are spaced into 39 groups of 72 bunches, each spaced 25 ns apart.

The beam pipe is a continuous vacuum chamber that passes through a large number of magnets. Magnetic fields in excess of 8 T are required due to the high momentum of the beam, which in turn require high electric current. To reduce resistance losses the magnets used in the LHC are superconducting and must be cooled to temperatures below 2 K with liquid helium. The ring contains 1,232 dipole magnets, 1,104 in the arcs and 128 in the straight sections, which serve to steer the beam along the beam pipes. Additionally, there are 392 quadrupole magnets whose function is to collimate the beams, to maximize the probability of interactions at the intersection points.

Operation of the LHC begins with the injection of bunches of protons into the LHC, which typically takes 20 minutes. From there, the beam can be ramped up to its operational energy in another 20 minutes. Once the beams reach their full operational energy, the optics are changed to squeeze the beams at the interaction points, and the magnets separating the beams are switched off, resulting in collisions. If the beams remain stable, this collision phase lasts on the order of 10 hours, during which time the luminosity is constantly decreasing. Eventually, the luminosity decreases to the point that it becomes advantageous to dump the beam, at which point the process can start over with fresh beams. Thus, faster turn-around time (limited by magnet hysteresis, injection, ramping, and squeezing) makes it possible to have shorter runs with higher-luminosity beams.

The LHC was in the very initial stages of its operation, with first beams circulating in the machine on 27th February 2010, and first collisions at 3.5 TeV/ beam (7 TeV center-of-

mass energy) taking place on 30th March 2010. Throughout 2010 a multi-stage commissioning procedure was employed in which the number of bunches and machine luminosity were steadily increased (see Figure 3.3). On 25th October 2010, the LHC reached its peak 2010 instantaneous luminosity of $2.07 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, colliding 348 bunches in each beam. The total integrated luminosity recorded at the LHC for 2010 is $43\,\mathrm{pb}^{-1}$ [32].

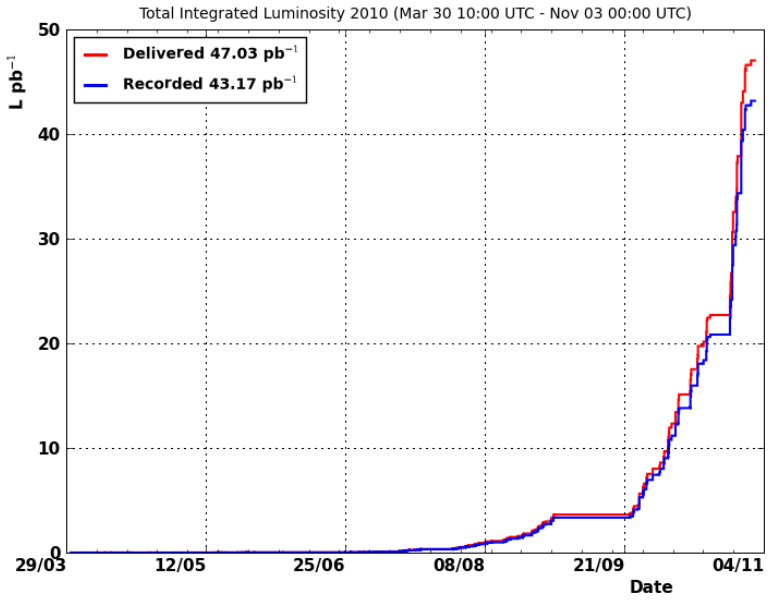


Figure 3.3: Beam size as a function of the distance to the collision point.

3.2 Compact Muon Solenoid

The Compact Muon Solenoid (CMS) experiment is a large general-purpose particle physics detector that lies at a point along the northern part of the circumference of the LHC. It is located in the town of Cessy, on the French side of the border.

One of the primary motivations of the experiment is to evaluate the consistency of the Standard Model (SM) at the new energy frontier opened up by the LHC. There are a wide

variety of proposed new physics models, many of which predict the SM will break down in precisely this region, invoking as-yet-unseen extensions such as new symmetries of nature or new forces and particles. The investigation of possible new physics at this TeV scale is a compelling part of the reason the LHC was built, and requires a detector that will be sensitive to a broad range of potential signals.

The overall design of CMS is shown in Figure 3.4. The detector is constructed as a series of layers going outward from the interaction point of the colliding protons [33]. By convention, we use a coordinate system with the origin located in the center of the detector at the collision point. The x- and y-axes point horizontally south and vertically up, respectively, so that the z-axis points west. We may then define a polar angle θ with respect to the z-axis, with $\theta = 0$ lying along the positive z-axis and $\theta = \pi$ along the negative z-axis, and an azimuthal angle ϕ in the xy-plane, where $\phi = 0$ is along the positive x-axis and $\phi = \pi/2$ is along the positive y-axis. Positions within the detector are generally defined in terms of the *pseudorapidity* η, rather than the polar angle, where the two are related by

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \tag{3.3}$$

The pseudorapidity is particularly useful in particle physics because particle production is roughly constant as a function of η.

3.2.1 Inner tracking system

The innermost part of CMS, closest to the interaction point, is the tracking system (see Figure 3.5). This is designed to measure precisely the trajectories of charged particles coming from collisions, as well as reconstruct secondary vertices so they can be distinguished from the primary interaction. This second feature is particularly important for the identifi-

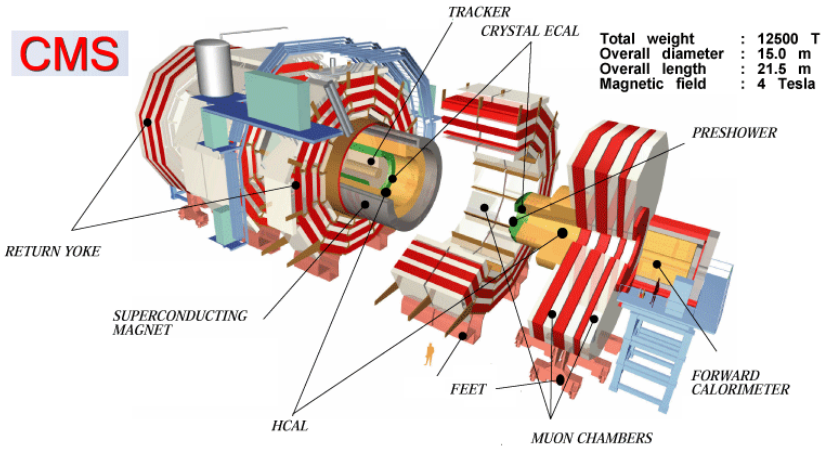


Figure 3.4: A perspective view of the CMS detector.

cation of B mesons, which can travel a detectable distance before decaying. The first three layers of the tracker are the silicon pixel detectors, which provide the highest precision and granularity of tracker channels. This is most important closest to the beam spot, where the density of particles from the interaction is highest. The pixel detectors extend out to a radius of 11 cm, and comprise 66 million pixels in total, each with an area of 100×150 μm. The resolution is $\sim$ 10 μm in the r-ϕ plane and $\sim$ 20 μm in the r-z plane. In total, the pixel detector covers the range $|\eta| < 2.5$

The next 10 layers of the detector, going out to a radius of 1.1 m, are composed of silicon microstrip detectors. The particle density decreases with increasing radius, so it becomes possible to achieve the required granularity with some 11.4 million long strips, 10 or 25 cm in length, but only 180 μm in height. The tracker barrel is divided into a tracker inner barrel (TIB) for $r < 65$ cm, with a resolution of 230 μm, and a tracker outer barrel (TOB) for $65 < r < 110$ cm, with resolution 530 μm. The tracker endcaps (TEC) cover the region out to $|\eta| < 2.5$ and contain strips oriented radially to the beam.

In total the tracker comprises 76 million channels with a total area of 205 m^2 (roughly

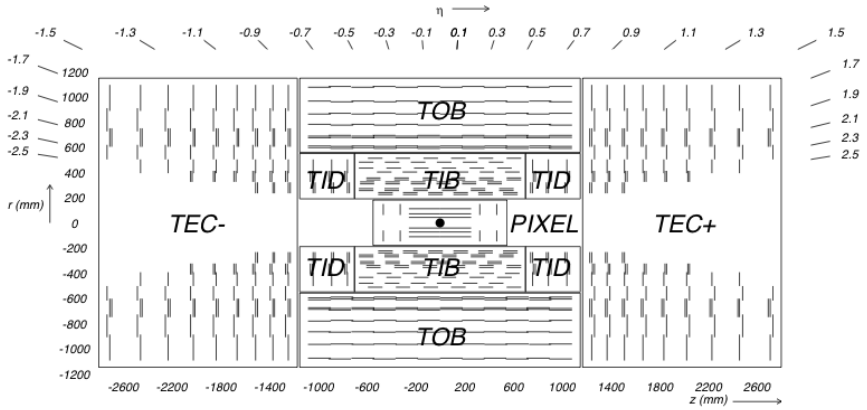


Figure 3.5: Cross-sectional diagram of the CMS tracker. Each line represents a detector module.

the size of a tennis court), making it the largest silicon detector in the world. The signal-to-noise ratio is better than 25:1, and the rate of dead or noisy strips is less than 3/1000. As the innermost layer of CMS, the tracker is subject to the strong magnetic field generated by the surrounding solenoid. It is therefore possible to measure the transverse momentum of any charged particle from the curvature of its path in the tracker, via

$$p_{\mathrm{T}} = qrB \tag{3.4}$$

where p_{T} is the transverse momentum, q is the charge of the particle, r is the radius of the track made by the particle, and B is the strength of the magnetic field (3.8 T). The large magnetic field and the fine granularity of the tracker elements yield excellent momentum resolution:

$$\frac{\sigma p_{\mathrm{T}}}{p_{\mathrm{T}}} = (15 p_{\mathrm{T}} \oplus 0.5)\,\%\,(\,\mathrm{TeV}) \qquad |\eta| < 1.6 \tag{3.5}$$

$$\frac{\sigma p_{\mathrm{T}}}{p_{\mathrm{T}}} = (60 p_{\mathrm{T}} \oplus 0.5)\,\%\,(\,\mathrm{TeV}) \qquad |\eta| = 2.5 \tag{3.6}$$

More energetic particles curve less in the magnetic field, which makes their momentum

more difficult to determine. Thus, the uncertainty in increases with increasing p_{T}. The tracker is an important component in the identification of both electrons and muons, which are used in this analysis.

3.2.2 Electromagnetic calorimeter

The next detector component, moving radially outward from the interaction point, is the electromagnetic calorimeter (ECAL), designed to acurately measure the energies of electrons and photons (Figure 3.6). The ECAL is a homogeneous calorimeter composed of crystals of lead tungstate ($PbWO_4$), a dense but optically transparent material which is ideal for stopping high energy particles. In total there are 61 200 crystals in the central barrel (EB), $|\eta| < 1.4442$, and 7 324 in each of the two endcaps (EE), $1.566 < |\eta| < 3.0$. The crystals in the barrel are organized into supermodules consisting of 1 700 crystals, 85 in η and 20 in ϕ. There are 36 supermodules in total, 18 in each half barrel.

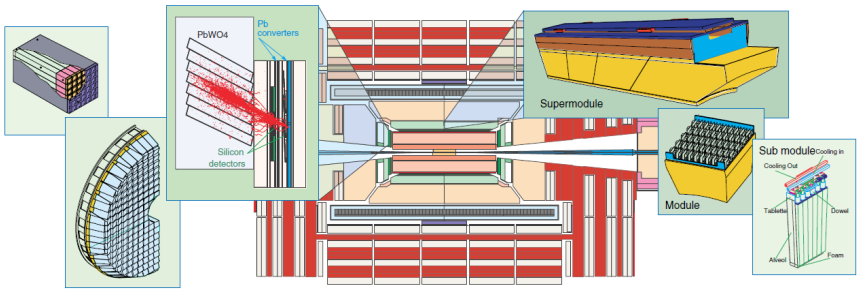


Figure 3.6: Diagram of the components of the ECAL showing the organization of the calorimeter crystals and an example electromagnetic shower.

To understand the performance of the ECAL, it is necessary to briefly discuss the interaction of electromagnetic particles with matter. High-energy electrons predominantly lose energy via bremsstrahlung, or “breaking radiation”, in which electrons in the field of

an atomic nucleus undergo large acceleration, causing the emission of a photon. As the electron passes through the material, it loses energy via bremsstrahlung as

$$\frac{dE}{dx} = -\frac{E}{X_0} \tag{3.7}$$

where X_0 is the characteristic radiation length of the material that the electron is moving through. In the case of the lead tungstate used in the ECAL, $X_0 = 0.89$ cm; the length of each crystal is 23 cm, or $25.8X_0$, ensuring that the incident electrons will undergo bremsstrahlung many times.

Photons tend to interact with matter primarily via pair-production, wherein a photon converts to a l^+l^- lepton pair as it traverses the medium. The combination of this process with bremsstrahlung leads to the creation of an electromagnetic shower. An incident electron will radiate a photon, which then pair-produces two additional electrons. These electrons will subsequently radiate photons, and so forth, until all of the energy of the original electron is contained in the resulting cascade of photons, which can then be measured by the scintillation counter at the base of the crystal. The same process occurs with an incident photon, but the chain begins with a pair-production.

Electromagnetic showers are known empirically to be well collimated in the transverse direction. Showers are generally characterized in the lateral direction by the Moliere radius R_M, defined as the radius of a cylinder containing 90% of the shower energy. In the case of $PbWO_4$, $R_M = 2.2$ cm, the same as the front size of the crystals, ensuring that the ECAL has fine granularity.

The energy resolution of the ECAL, given by [34]

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{41.5\,\text{MeV}}{E}\right)^2 + (0.3\%)^2 \tag{3.8}$$

is parametrized into multiple independent terms. The first term is stochastic, including statistical fluctuations and intrinsic shower-size fluctuations. The second term accounts for electronics noise and pile-up energy, and the third term is a constant covering detector non-uniformity and calibration uncertainty. The overall energy resolution is very good, and is optimal for energies between $1\,\text{GeV}$ and $1\,\text{TeV}$. Additional corrections are required over time, because exposure to radiation causes the crystals to lose transparency, so they must be recalibrated using a laser system.

3.2.3 Hadron calorimeter

Beyond the ECAL is the hadron calorimeter (HCAL), whose purpose is both to measure the energy of hadron jets and to provide a reliable measure of the missing energy in an event. It is divided up into three components by η, with the HCAL barrel (HB) covering the region $|\eta| < 1.305$, the HCAL endcap (HE) extending from $1.305 < |\eta| < 3.0$, and the HCAL forward (HF) region going from $3.0 < |\eta| < 5.0$. The total amount of material available in the barrel is restricted radially by the outer edge of the ECAL and the inner edge of the magnet. This is filled with the interleaved layers of of dense material, in this case brass, and tiles of plastic scintllator. These layers extend outwards 0.96 m, and are divided into 36 azimuthal wedges which compose the two half-barrels of the HB. In the HF steel is used in place of brass, and the scintillator layer is composed of quartz fiber (Figure 3.7).

The function of the metal layers is to provide a dense material with heavy nuclei to cause the hadron to interact, while the scintillator layers measure the visible energy from those interactions. Hadronic interactions with matter can proceed through a variety of processes, including particle production (dominated by pion production), nuclear breakup, and π^0 production, with the subsequent decay $\pi^0 \rightarrow \gamma\gamma$. Because of the photon production,

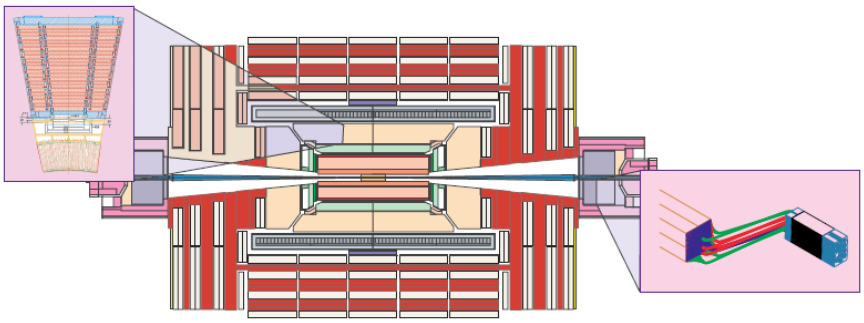


Figure 3.7: Diagram of the components of the ECAL showing the brass and scintillator of the HB/HE and the steel and quartz of the HF.

hadronic showers will also have an electromagnetic component. In addition, decays to neutral particles will fail to produce energy observed by the scintillator layers. As a result of fluctuations in the number of netural particles and π^0s produced, there can be large variations in the size and visible energy of hadronic jets, though they are generally much larger than electromagnetic jets, both longitudinally and transversely.

The HCAL extends 6–8 interaction lengths, with the energy of a hadron decreasing by a factor of $1/e$ in a single interaction length. Brass is optimal in the barrel, as it is non-magnetic and possesses a short interaction length, while steel and quartz are used in the endcaps because they are better able to withstand the high rate of radiation from the beams. The resolution for both the ECAL and HCAL is given by [35]

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{85\%}{\sqrt{E}}\right)^2 + (7.4\%)^2 \qquad |\eta| < 3.0 \tag{3.9}$$

and the resolution of the HCAL in the forward region is

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{198\%}{\sqrt{E}}\right)^2 + (9.0\%)^2 \qquad 3.0 < |\eta| < 5.0 \tag{3.10}$$

The first term is stochastic, including statistical fluctuations and intrinsic shower-size fluctuations. The second term is a constant accounting for detector non-uniformity and calibration

uncertainty. Comparison of (3.9) to (3.8) shows that HCAL resolution is worse than ECAL resolution, both because the HCAL is a sampling calorimeter and because hadronic showers tend to have larger statistical fluctuations.

3.2.4 Superconducting magnet

The main distinguishing feature of CMS is a very large superconducting solenoid magnet, 13 m long and 6 m in diameter, which is capable of providing a magnetic field of 4 T in the interior of the solenoid 3.8. For the collection of data in 2010, the actual operating field of the magnet was 3.8 T. This field will bend the paths of charged particles in the inner layers, allowing their momenta to be determined, and so a high magnetic field is essential for both good momentum resolution and unambiguous determination of particle charge.

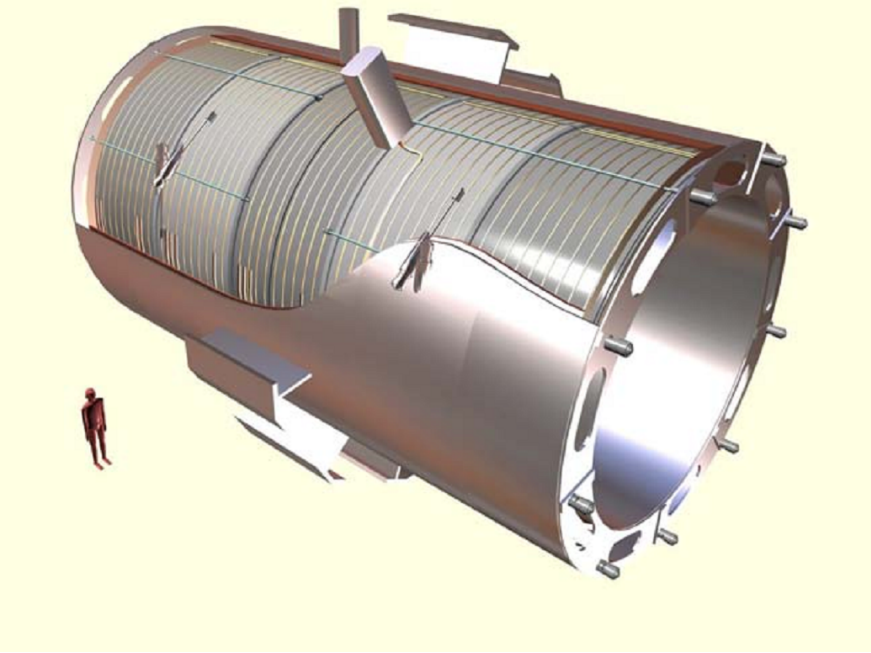


Figure 3.8: A perspective view of the solenoid.

Because the magnet itself is so large, it is possible to accomodate not only a tracking

system, but also all of the calorimetry, entirely in its interior. This allows better electromagnetic energy resolution as well as missing-transverse-energy and jet energy resolution, as these can be measured before the particles pass through the high-density magnet. The magnetic flux through the interior of the solenoid is returned through a supporting iron yoke 1.5 m thick, weighing 10,000 t and comprising 5 "wheels" and 2 "endcaps" with 3 disks each.

3.2.5 Muon system

As suggested by the name of the experiment, muon detection is a major theme of CMS, and the precise and accurate measurement of muons has been given high importance from the earliest design stages. Due to their particular mass, muons have minimal interactions with matter and are consequently very penetrating. Because of this, the muon system can be placed beyond the solenoid in the iron return yoke. There are three separate detector systems for muons (see Figure 3.9): drift tubes (DT) are employed in the central region (out to $|\eta| < 1.2$) for precise trajectory measurements and cathode strip chambers (CSC) are used in the endcaps ($0.9 < |\eta| < 2.4$); resistive plate chambers, which provide precise timing of muons passing through the detector, cover the barrel and much of the endcaps, out to an $|\eta|$ of 1.6.

The DTs are a type of wire chamber, in which an electric field is set up around a thin wire that is immersed in a gas (in this case, a mixture of argon and carbon dioxide). The passage of the charged muon ionizes this gas, and the electric field causes the freed atomic electrons to drift toward the sensing wire. These electrons, accelerated by the field, produce more ions, and the resulting avalanche produces a current in the wire, such that the signal can be recorded. The DTs are organized into 4 stations moving radially outward, and the

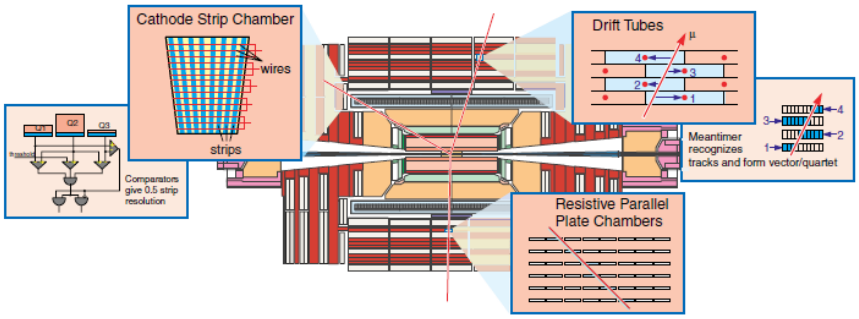


Figure 3.9: Configuration of the muon system showing the three separate muon detectors and their positions.

drift cells of each chamber are offset by half of the cell width with respect to their neighbors. This onfiguration makes it possible to link the signals in each chamber into a single muon track via the mean timing of the hits on each wire. It also eliminates inefficient spots in the DT and improves rejection of background hits.

The muon production rates and neutron-induced backgrounds are higher closer to the beam line, so CSCs, which have a faster response time, fine segmentation, and good resistance to radiation, are preferred for the endcaps. They are composed of cathode strips that run radially outward and anode wires that run perpendicular to the strips. The anode wires detect the ionized electrons from the passage of the muon, and the positive ions induce a pulse in the cathode strips. Thus, the strips provide a precision measurement of the path of the muon in the r-ϕ plane, while the wires measure the η of the muon.

The RPCs comprise a third, complementary muon detection system, designed to make fast and independent muon measurements over a wide η range. The parallel resistive plates create a constant electric field in the gas filled interior, such that when an incident muon ionizes the gas, the electrons are pulled to the plates. The electrodes themselves are transparent to the signal, which is instead picked up by a series of conductive detecting strips on

the other side. RPCs have a very fast response (1 ns), but coarser spatial resolution than the DTs or CSCs, making them optimal for triggering on muons.

3.3 Trigger

For pp collisions, the LHC has a nominal beam crossing interval of 25 ns at peak luminosity, which corresponds to a crossing frequency of 40 MHz (though the maximum frequency achieved in 2010 running was 5 MHz). The actual number of collisions for each proton bunch crossing depends on luminosity, but is approximately 20 at the machine's nominal design luminosity of $\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. It is not possible to store or process the total data from all events, and the production cross section for many interesting physics processes, including supersymmetry and top quark production, is several orders of magnitude smaller than the inelastic pp cross section. It is therefore necessary to perform an initial event selection, both to reduce the overall event rate and to increase the proportion of relevant physics events in the stored data. The overall goal for the rate reduction is a factor of at least 10^6, and this is handled by the trigger system, which reduces the event rate in two steps: the Level-1 (L1) trigger and the high-level trigger (HLT).

3.3.1 Level-1 trigger

The L1 trigger consists primarily of custom-designed programmable electronics and has a maximum output rate of 100 kHz. The hardware for the L1 trigger uses coarsely segmented data from the calorimeters and muon systems to make a fast decision (less than $3.2\ \mu$s) on whether to keep a given event. It is divided into local, regional, and global components, as shown in Figure 3.10.

Local trigger information is produced by so-called Trigger Primitive Generators (TPG),

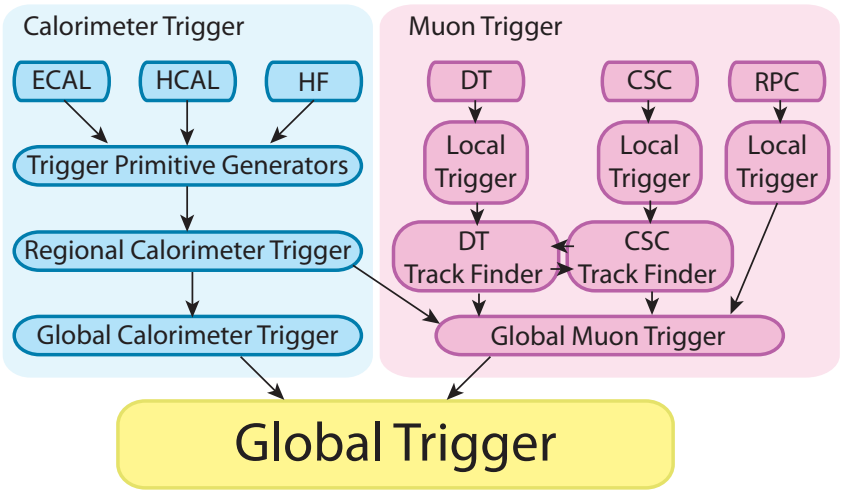


Figure 3.10: Schematic showing the architecture of the L1 trigger.

which come from both the calorimeter and muon triggers; in the former case they are based on energy deposits in the calorimeter towers, and in the latter case they use track segments or hit patterns in the muon chambers. The information from the TPGs is passed to the regional triggers, where it is used to determine a sorted list of trigger objects like electron or muon candidates in specific regions, and to rank them by energy. These candidates are then passed to the global calorimeter and global muon triggers, which determine the overall highest-rank calorimeter and muon objects to send to the global trigger. It is the global trigger which makes the final L1 decision to reject an event or to accept it for further evaluation by the HLT.

3.3.2 High-level trigger

In contrast to the L1 trigger, the HLT is a software system, implemented by approximately a thousand commercial processors. The job of the HLT is to reduce the event rate to a final output rate of approximately 300 Hz, rougly the maximum rate that can be written

to a storage disk. While the L1 trigger uses coarse-binned data, the HLT has access to the complete read-out data of the event, and is able to perform complex calculations similar to those performed in an offline analysis. Consequently, it is possible to construct more sophisticated trigger requirements for event selection.

Chapter 4

Event Simulation

What happens if a big asteroid hits Earth? Judging from realistic simulations involving a sledgehammer and a common laboratory frog, we can assume it will be pretty bad.

— Dave Barry

Data analysis in high energy physics depends on our ability to accurately compare theoretical expectations with experimental output from the detector. For this we rely extensively on computer simulation of data, based on current theoretical models. The simulation of these events typically involves the evaluation of difficult integrals. Rather than attempting to calculate these integrals directly, the simulation uses "Monte Carlo" techniques [36] (named for the city due to its widely known affiliation with gambling) to perform a large number of simulated experiments using random number generators.

4.1 Monte Carlo Simulation

Event generation using Monte Carlo is modular and proceeds in stages. In many cases it is feasible to use different programs for each step [37]. First, relevant information about the collisions to be simulated is given to the program, including the particles to be collided

(in this case protons) and the center-of-mass energy of the collision (in this case $\sqrt{s} = 7\,\text{TeV}$). The structure of the colliding protons is modeled with the use of *parton distribution functions* (PDFs), which define the probability density for finding a parton (a quark or gluon) with a given longitudinal momentum fraction x at a particular value of momentum transfer Q^2 of the collision. The collision occurs between individual partons within each proton, and is referred to as the *hard process*, in which the particles of interest for the event are produced (*e.g.* heavy quarks, vector bosons, hypothetical new particles, etc.). The production of these particles, and their subsequent decays, are described by *matrix elements* corresponding to each interaction. These are calculated from Feynman diagrams in quantum field theory, where the square of the matrix element gives the probability density for the process [13]. In addition, because of QCD confinement (described in Chapter 1), colored remnants from the hard interaction will produce *parton showers*, eventually hadronizing into collections of colorless hadrons seen as jets.

4.1.1 Parton Distribution Functions and the Hard Process

The simulation of events and calculation of production cross sections at the LHC rely on understanding the distribution of the momentum fraction x of the partons in the colliding protons in the relevant kinematic range. In practice, the PDFs containing this information cannot be calculated perturbatively, and are instead determined by fits to data from processes such as deep inelastic scattering (DIS) and Drell-Yan (DY). Groups such as CTEQ [38] and MRST [39] provide fairly regular updates to these experimentally determined PDFs. Distribution functions for various partons are plotted in Figure 4.1 for a typical momentum transfer of $Q = 10\,\text{GeV}$.

The interaction producing the hard scatter occurs between the individual partons in

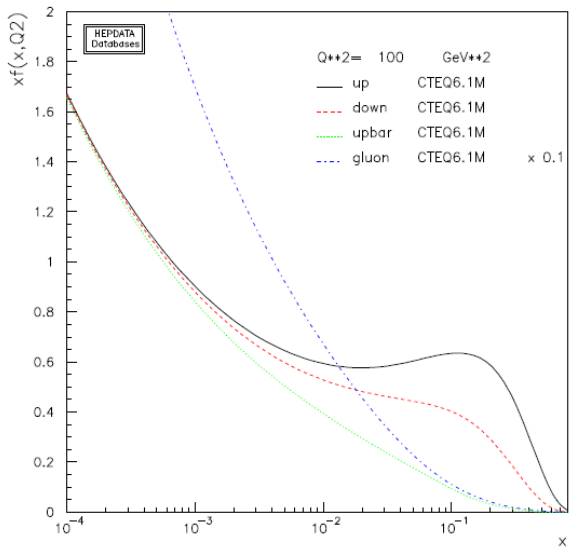


Figure 4.1: The CTEQ 6.1 parton distribution functions evaluated at $Q = 10\,\text{GeV}$.

each proton, with the PDFs determining the energy available for the process. Feynman diagrams, such as the one shown in Figure 1.6 are used to calculate the matrix element for each process, and the production cross section can be derived from this matrix element. In practice, the programs used to generate the simulated events are "tree level" (*i.e.* leading order) generators, so separate programs are used to calculate the cross sections in order to include higher-order effects.

4.1.2 Parton Showers and Hadronization

In addition to the hard scattering process, the collision includes QCD radiation from both the incoming and outgoing partons and referred to as *initial* and *final state radiation* (ISR and FSR), respectively. This radiation depends primarily on the momentum transfer scale Q^2, rather than the details of the particular process being simulated.

Due to color confinement, partons produced in the hard scatter cannot exist on their own. Their kinetic energy is transferred to the color field, where it produces additional

partons from the vacuum. The partons in this shower then hadronize, forming color-neutral combinations. The resulting collinear spray of particles is called a jet.

The process of hadronization is not well described theoretically, and is instead modeled phenomenologically with the *Lund string model* [40, 41]. In this model, the gluons binding two quarks are treated as field lines. Gluons are self-interacting, and hence the color field they produce is compressed into a narrow tube (or string). By contrast, the electromagnetic field tends to spread much more, because the photon has no self-interaction terms, as discussed in Chapter 1. When sufficient energy is stored in the color field to produce new quark-antiquark pairs, the string "snaps", and the newly produced particles form bound-state mesons with the original quarks. In hadronization this process happens repeatedly, and the produced bound states model the kinematics of the original parton. This model of string fragmentation is notably used by the PYTHIA event generator [42] and explains many features of hadronization well, particularly the production of jets.

Two partons, one from each jet, create the hard scatter process, but the remaining partons from the protons cannot be ignored, and represent the *underlying event*. The proton remnants are left behind after the hard scatter, but due to their strong interactions with the hard process, they must be included when considering hadronization. These interactions are typically soft and the description of the underlying event relies on non-perturbative or semi-perturbative phenomenological models.

Additionally, the effect of *pileup* must also be considered, where multiple protons in a single bunch crossing interact. Pileup is simulated in generated events by superimposing minimum bias events on the events in the nominal samples, according to the distribution of number of pileup collisions observed in the data for 2010. For the majority of the data taken

in 2010, there were around 2.2 collisions for each triggered bunch crossing.

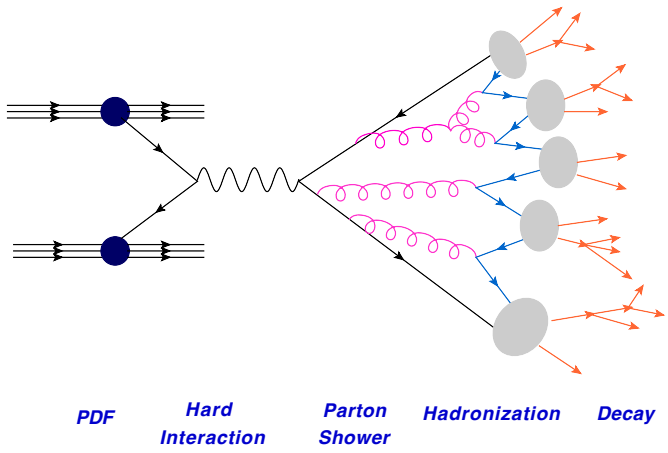


Figure 4.2: A cartoon showing each stage of the MC simulation.

4.1.3 Monte Carlo Generators

Different Monte Carlo programs are used depending on the type of event being simulated. Some generators, such as PYTHIA [42] and HERWIG [43], simulate every step in the event, including the intial interaction, parton shower, underlying event and subsequent hadronization and decay to final state particles. Cross sections are calculated only to leading order (LO), and these generators are most suitable for $2 \rightarrow 2$ processes (those with two initial and two final state particles) such as partonic scattering in QCD.

For many processes, particularly those with complex multi-particle final states such as $t\bar{t}$ and SUSY, LO cross section calculations are insufficient. For these cases we employ more specialized MC generators such as MADGRAPH [44] and ALPGEN [45]. These generators specialize in modeling the hard interaction and are capable of calculating next-to-leading order (NLO) corrections to matrix elements, as well as simulating $2 \rightarrow n$ processes with multiple final state partons. These generators are appropriate for reproducing the correct

kinematics and cross sections of events containing multiple jets. However, because they do not include hadronization, they must be interfaced with other generators to produce the full event.

Table 4.1 shows the simulated events used in this analysis, along with the event generator used to produce them. SM processes (except for QCD) were generated with MADGRAPH, which calculates the matrix element for each object ($t\bar{t}$, W, Z, etc.) and up to four jets. These samples were then interfaced to PYTHIA in order to incorporate the parton shower and hadronization steps. The particular parameter values used for these depend on the "tune" which is used. The "Z2" tune [46] is used consistently for all samples listed in the table, while the "D6T" tune [47] is used for additional systematic studies.

4.2 Detector Simulation

It is necessary not only to model the physics of various processes, but also the detector response to the particles that are produced. The simulation of the CMS detector is done with a program called GEANT4 [48], a toolkit used to model the interactions of particles in the detector. The algorithm incorporates information on the materials, magnetic fields and specific geometry of the CMS detector, and uses these to determine the detector response to the simulated particles.

Monte Carlo samples				
Process	Generator	Kinematic cuts	σ (pb)	Generated events
Common samples				
$t\bar{t}$	MADGRAPH	—	157.5	1165716
$W \to l\nu$	MADGRAPH	—	31314.0	15154787
$Z/\gamma^* \to l^+l^-$	MADGRAPH	$m_{ll} > 50\,\text{GeV}$	3048.0	5257046
t (s-channel)	MADGRAPH	—	1.4	494967
t (t-channel)	MADGRAPH	—	20.93	484060
tW	MADGRAPH	—	10.6	494961
Samples for $t\bar{t}$ production cross section measurement				
QCD (μ-enriched)	PYTHIA	$\hat{p_\text{T}} > 20\,\text{GeV}$, $p_\text{T}^\mu > 15\,\text{GeV}$	84679.3	29504868
Samples for SUSY search				
LM0 SUSY	PYTHIA	—	38.93	219595
QCD (EM-enriched)	PYTHIA	$20\,\text{GeV} < \hat{p_\text{T}} < 30\,\text{GeV}$	2454400	36920244
QCD (EM-enriched)	PYTHIA	$30\,\text{GeV} < \hat{p_\text{T}} < 80\,\text{GeV}$	3866200	71834016
QCD (EM-enriched)	PYTHIA	$80\,\text{GeV} < \hat{p_\text{T}} < 170\,\text{GeV}$	139500	8073559
QCD ($b/c \to e$)	PYTHIA	$20\,\text{GeV} < \hat{p_\text{T}} < 30\,\text{GeV}$	132160	2243439
QCD ($b/c \to e$)	PYTHIA	$30\,\text{GeV} < \hat{p_\text{T}} < 80\,\text{GeV}$	136804	1995502
QCD ($b/c \to e$)	PYTHIA	$80\,\text{GeV} < \hat{p_\text{T}} < 170\,\text{GeV}$	9360	1043390
QCD (μ-enriched)	PYTHIA	$15\,\text{GeV} < \hat{p_\text{T}} < 20\,\text{GeV}$, $p_\text{T}^\mu > 5\,\text{GeV}$	1471168	2884915
QCD (μ-enriched)	PYTHIA	$20\,\text{GeV} < \hat{p_\text{T}} < 30\,\text{GeV}$, $p_\text{T}^\mu > 5\,\text{GeV}$	1224034	11417239
QCD (μ-enriched)	PYTHIA	$30\,\text{GeV} < \hat{p_\text{T}} < 50\,\text{GeV}$, $p_\text{T}^\mu > 5\,\text{GeV}$	578463	11431864
QCD (μ-enriched)	PYTHIA	$50\,\text{GeV} < \hat{p_\text{T}} < 80\,\text{GeV}$, $p_\text{T}^\mu > 5\,\text{GeV}$	144422	10748755
QCD (μ-enriched)	PYTHIA	$80\,\text{GeV} < \hat{p_\text{T}} < 120\,\text{GeV}$, $p_\text{T}^\mu > 5\,\text{GeV}$	29049	3191979
QCD (μ-enriched)	PYTHIA	$120\,\text{GeV} < \hat{p_\text{T}} < 150\,\text{GeV}$, $p_\text{T}^\mu > 5\,\text{GeV}$	4440	998503
QCD (μ-enriched)	PYTHIA	$\hat{p_\text{T}} > 150\,\text{GeV}$, $p_\text{T}^\mu > 5\,\text{GeV}$	2838	1022541

Table 4.1: MC signal and background samples used for each part of the analysis.

Chapter 5

Event Reconstruction

A common mistake that people make when trying to design something completely foolproof is to underestimate the ingenuity of complete fools.

— Douglas Adams

The data from an event is collected by the CMS detector in a very low-level form (*e.g.* hits in the tracker, energy deposits in the calorimeters). This raw data must then be synthesized into the actual objects that are used in the analysis. The top cross section measurement relies on both muons and jets, while the new physics search employs leptons (electrons and muons) as well as jets and $\not\!\!E_T$.

5.1 The Particle Flow Algorithm

Particle flow (PF) [49] is an algorithm designed to reconstruct all possible final state particles in an event (*i.e.* electrons, muons, charged and neutral hadrons, etc.) in a unified way. Its goal is to optimally determine the properties of the reconstructed object, such as type, energy and direction, by using information from all parts of the detector. This information is gathered from each subdetector and linked together to provide a coherent description of each reconstructed particle. All activity in the detector above some noise

threshold is taken to be part of an object and is assigned to one of the reconstructed particles. From these fundamental final state particles, more complex objects like jets and $\not\!\!E_T$ are then constructed.

5.1.1 Track and Vertex

One of the most fundamental components of many physics objects is the reconstruction of tracks of charged particles. Tracks are used in the reconstruction of every object used in this analysis, including electrons, muons, jets and $\not\!\!E_T$. As discussed in Section 3.2.1, the large magnetic field and high resolution of the tracker make it possible to measure the momentum of charged particles extremely accurately. Moreover, it is possible to determine the production vertex of the event with high precision.

Tracks in the CMS detector are reconstructed via the combinatorial track finder (CTF) [50]. Initial estimates (or "seeds") of tracks are constructed either from triplets of hits or from pairs of hits with additional constraints from the beamspot or a vertex. This seed is then propagated outwards, and as new hits are found in the tracker, they are added to the seed trajectory and the entire track is updated with the new information. The process continues until no further compatible hits are found, and this collection of hits is then fit to obtain the best estimate of the track parameters, including the p_{T}, η and ϕ of the reconstructed track.

Multiple iterations of this process are performed, with hits removed from the collection as they become unambiguously assigned to tracks. After each iteration, filtering is applied to remove potential fake tracks. The initial seeding criteria are very tight, to ensure the initial fake rate is negligibly low. The criteria are progressively loosened for subsequent iterations in order to increase the efficiency to match charged hadrons and to find tracks far from the interaction vertex, such as electrons from photon conversion and tracks from

b-jets. The successive removal of matched hits ensures that rates of fake tracks are kept low, and tracks with high efficiency and small fake rate can be constructed down to transverse momenta on the order of $100\,\text{MeV}$.

The primary interaction vertex is reconstructed from the collection of tracks. "Prompt" tracks, coming directly from the primary interaction, are defined based on multiple variables, including the transverse impact parameter d_0, the number of hits in the track, and its normalized χ^2. These tracks are then clustered along the z axis, and this cluster is fit using an "adaptive vertex fit" [51]. The location of the primary interaction vertex is then determined from this fit.

5.1.2 Calorimeter Energy

As described in Sections 3.2.2 and 3.2.3, the calorimeters are important for determining the energy of a number of different particles, including electrons and photons (which deposit all their energy in the ECAL) and hadrons (which deposit energy in both the ECAL and the HCAL). The granularity of the HCAL is about 25 times coarser than the ECAL, but combining information from the two to obtain a single energy estimate gives an energy resolution of order 10% for a typically energetic hadron of $100\,\text{GeV}$. This makes it possible to separate neutral hadrons from charged hadrons by identifying an excess of energy beyond that deposited by the charged hadrons.

The PF algorithm employs a clustering method in the calorimeters to measure the energy from neutral particles, separate them from charged particles, and reconstruct the energy of electrons, which can undergo significant energy loss on their way to the calorimeter due to the process of Bremsstrahlung. First, the cluster is seeded by a single calorimeter cell with a local energy maximum above some threshold, defined as a two standard deviation

excess above the electronics noise in the calorimeter ($80\,\text{MeV}$ in the barrel and $300\,\text{MeV}$ in the endcaps). Next, the algorithm grows the cluster by aggregating cells that share a side with a cell already in the cluster and have an energy greater than the threshold. Each separate seed eventually becomes a "particle flow cluster".

5.1.3 Link Algorithm

In order to construct final state objects, it is necessary to combine information from multiple different subdetectors. For this PF employs a "link algorithm" whose purpose is to combine information from each detector element in a way that ensures that it is not double-counted. A link is established between two elements, with the distance between them determining the link quality, and "blocks" are formed from these linked elements.

A link between a charged particle track and a calorimeter cluster is established by extrapolating the track from the last measured hit in the tracker to the ECAL and HCAL. If this extrapolated track lies within a cluster boundary defined in the calorimeter, that cluster is then linked to the track. The link distance here is defined as the difference in η-ϕ space between the position of the extrapolated track and the position of the cluster. Other clusters are linked by taking positions tangent to the track, in order to include the energy lost due to Bremsstrahlung.

Links between deposits in the ECAL and HCAL are performed similarly. Here, the cluster position is defined in the ECAL, which has greater granularity, and extrapolated into the HCAL. If the cluster position lies within the envelope defined by the HCAL cluster, then the two are linked. The link distance is analogously defined as the difference in η-ϕ space between the two clusters.

For muons, it is also necessary to link tracks in the tracker to tracks in the muon

chambers. This link depends on matching the hits in each detector element to a global fit, and requiring that this fit return an acceptable χ^2 [52]. In the case where a given track in the muon chambers is compatible with multiple tracks in the tracker, the fit that returns the smallest χ^2 is used to define the muon. In this case the link distance is defined as the χ^2 of the global fit.

5.2 Electron Reconstruction

Electrons are reconstructed from a track in the tracker that has been matched to an ECAL energy deposit. It is done via one of two complementary methods, optimized for different situations [53]. "Tracker driven" seeding is appropriate for low-p_{T} electrons, as well as those with poor isolation (such as electrons appearing inside jets). The electrons used in this analysis are required to be well-isolated, and are instead seeded from the ECAL. The algorithm starts from the presence in the calorimeter of a "supercluster" [54], defined as a collection of one or more energy clusters, constructed using a narrow width in η and a spread in ϕ to account for the effect of photons radiated from the electron as it bends in the magnetic field. These superclusters are matched to the track seeds described above to make electron tracks. The reconstructed trajectory of the electron is calculated based on a fit performed with a Gaussian sum filter (GSF) [55]. This reconstruction process is illustrated in Figure 5.1. On the left is a cartoon indicating how the electron is reconstructed from a track in the tracker and a narrow strip of clusters in the calorimeter, while the event display on the right shows an actual event from data which includes an electron created from a track and ECAL deposit.

In addition, the reconstructed electron is subject to several preselection requirements in order to ensure high quality. The ratio of energy deposited in the HCAL to energy de-

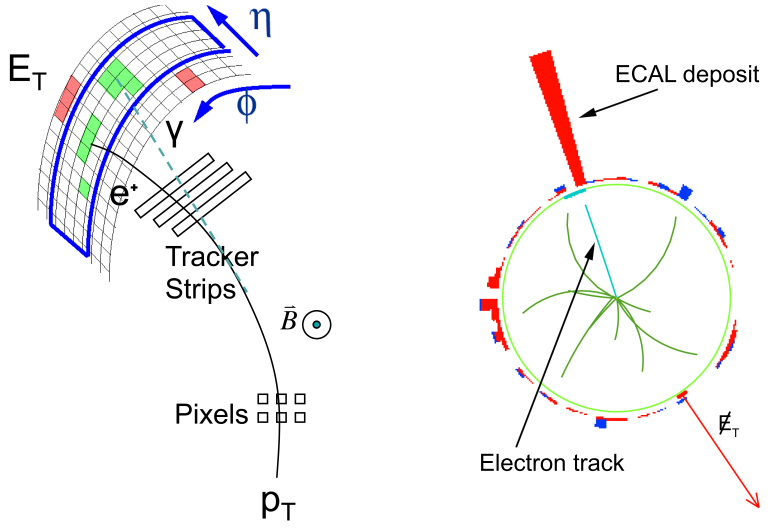
posited in the ECAL (H/E) is required to be less than 0.15 in order to reduce the fake rate. In addition, loose requirements are imposed on the distance between the track and the supercluster in the η and ϕ directions ($\Delta\eta < 0.02$, $\Delta\phi < 0.15$). These are tightened at the event selection step, as described in Chapter 6.

The momentum of the reconstructed electon is computed either using both the tracker and ECAL information, or solely from the supercluster in the ECAL, depending on the precision with which the tracker momentum is known: if $\sigma_{E/p}$ is the uncertainty in the ECAL and track measurements, then for electrons with $|E/p - 1| < 2.5\sigma_{E/p}$ the electron momentum is computed as the weighted mean of the supercluster energy and track momentum. Otherwise it is determined exclusively from the supercluster energy. Typically track momentum is favored at low-p_{T}, as these electrons will bend more in a magnetic field, making more precise measurements possible. At higher energies it is preferable to rely on the energy in the ECAL [56].

5.3 Muon Reconstruction

Muon reconstruction relies on both tracker and muon chamber information. Muon trajectories are calculated from both tracks in the tracker and hits in one or more of the muon systems at at the outermost parts of the detector (drift tubes, cathode strip chambers and resistive plate chambers, depending on the η region, as described in Chapter 3). In addition, muons can leave energy deposits in both the ECAL and the HCAL which are used for muon identification.

Reconstructed muons are formed from tracks in the tracker and tracks in the muon chambers using one of two different methods. *Global muons*, or "outside in" muons are seeded from the standalone track in the muon systems. This is then matched to a track in



(a) (b)

Figure 5.1: (a) Cartoon of an electron leaving a track in the tracker and depositing energy in the ECAL crystals, including a photon γ from bremsstrahlung. (b) Display in the ρ-ϕ plane showing an event with a high-p_{T} electron leaving a track and an ECAL deposit.

the tracker and the combined track is fit using hits in both systems. This method of reconstruction is particularly useful for high transverse momentum muons ($p_{\mathrm{T}} \gtrsim 200\,\mathrm{GeV}/c$) where momentum resolution is improved over that obtained solely from the tracker [57].

Tracker muons, or "inside out" muons are seeded from the tracker. All tracks of $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}/c$ and $p > 2.5\,\mathrm{GeV}/c$ are treated as muon candidates and extrapolated into the muon detectors. If at least one muon segment (*i.e.* a short track seed made of muon hits) can be matched to the extrapolated track, the combined fit is taken to be the trajectory of the tracker muon. This approach can be more efficient for low-p_{T} muons, as they require only a single segment in the muon chambers.

To ensure that the muons used in this analysis are of high quality, they are required to be reconstructed by both the global and tracker algorithms. They are taken to be the same muon if they share the same track in the tracker and are merged into a single reconstructed muon.

In addition, several quality requirements are imposed on the reconstructed muons used in this analysis, referred to as the "tight muon" selection. The normalized χ^2 of the global muon fit must be less than 10 and at least one hit in the muon chambers must be included in the final fit. Further, the muon must be matched to at least two muon segments, and the corresponding track in the tracker must include at least 11 hits (including at least one from the pixels).

5.4 Jet Reconstruction

The next step in the PF algorithm is to reconstruct the individual hadrons and photons in the event. These are then amalgamated into collinear collections called "jets". Charged hadrons are identified by linking their tracks in the tracker to both their ECAL and HCAL

energy deposits. Then neutral hadrons and photons can be identified by comparing the track momenta to energy deposits in the calorimeters; quantities in excess of those predicted by the tracks are attributed to neutral particles.

Because the partons emerging from the hard interaction are color-connected to the underlying event, the hadrons they fragment into are expected to be clustered into jets consisting of collinear groups of particles. In the PF algorithm, jets are not constructed from subdetector elements (as in most other jet reconstruction algorithms) but rather from the reconstructed PF particles found in the previous steps. As a result, the PF jets more closely resemble the jets constructed in MC simulations, where the true particles from the interaction are grouped via the same algorithm. In a typical jet, the fractions of charged particles, photons and neutral hadrons are 65%, 25% and 10%, respectively. Only the neutral hadrons are affected by the poor resolution of the HCAL, which ensures that the PF algorithm is capable of reconstructing roughly 90% of the jet energy with high precision.

5.4.1 The Anti-k_t Algorithm

Particles are clustered into jets via an algorithm known as "anti-k_t [58], which is infrared and collinear (IRC) safe. Matrix element calculations at the parton level can include divergences due to soft gluon emmission (called infrared divergences) and from a single parton splitting into two collinear partons. Because the jet reconstruction algorithm is robust against the presence of infinitely soft gluons, and because it correctly recombines collinear partons into a single parton, it is considered IRC "safe" [59]. Because it possesses these properties, it can be compared to theoretical predictions of any order.

The anti-k_t algorithm takes the locally highest-momentum particle and clusters nearby particles into a jet around it based on their proximity in both position and momentum space.

As each new particle is added, the resulting "pseudojet" is adjusted based on the resulting position and momentum. We define a distance metric d_{ij} between any two entities (*e.g.* particles or pseudojets) i and j, as well as d_{iB}, the distance between the entity i and the beam B. If the smallest distance is d_{ij}, we recombine i and j into a single pseudojet, weighting the position by the momenta of the entities, while if the smallest distance is d_{iB}, we define i as a jet and remove it from the list of entries. This procedure is iterated until no entries remain. The distance metric used for this algorithm is

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} \tag{5.1}$$

$$d_{iB} = k_{ti}^{2p} \tag{5.2}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and k_{ti}, y_i and ϕ_i are the transverse momentum, rapidity and azimuthal angle of particle i, respectively. R indicates the radius of the cone used in the clustering algorithm, and p is a parameter governing the relative power of the energy versus the geometrical scales (Δ_{ij}^2). In the inclusive "k_t" algorithm we set $p = 1$, and for $p = 0$ we recover the Cambridge/Aachen algorithm [58]. In the case of the anti-k_t algorithm, we set $p = -1$. Referring to Eq. 5.1, we see that for negative values of p the algorithm starts with higher-momentum particles and adds successively softer particles. The cone size used for this analysis is $R = 0.5$.

The behavior of this algorithm depends on how hard and soft particles are distributed in the event. Softer particles tend to cluster with harder ones preferentially over other soft particles. If a hard particle has no other hard particles nearby, it will cluster all the soft particles around it, resulting in a jet of radius R. If another hard particle exists in the region $R < \Delta_{ij} < 2R$, then the softer particles surrounding them will be split between the two, weighted by the relative momenta of each hard particle. If two hard particles exist

with $\Delta_{ij} < R$, the algorithm combines them into a single jet whose center is weighted by the relative momenta of the two particles. The shape of the final jet is determined by the cone around each hard particle, plus a cone centered around the final jet. As a result, this algorithm tends to produce jets whose shape is resistant to the distribution of softer particles, while still being adaptable to harder radiation. Figure 5.2 shows an example of the anti-k_t clustering algorithm applied to a parton-level event with additional soft radiation.

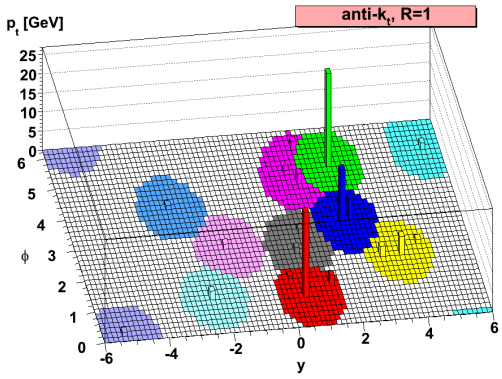


Figure 5.2: A sample parton-level event (generated with Herwig [43]) along with additional soft radiation. The diagram shows the results of applying the anti-k_T algorithm to this event, with the differently colored regions indicating the areas defined by each jet. Higher energy jets take up a larger portion of the surrounding area than nearby soft jets.

5.4.2 Jet Energy Corrections

The energy of jets needs to be corrected for a variety of factors. The largest of these is the non-uniform and non-linear response of the calorimeters, with additional contributions from noise in the electronics and pileup from multiple interactions occuring in the same bunch crossing. The goal of these jet energy corrections is to relate the jet energy measured in the detector to the energy of the original particle jet.

These corrections are factorized into three separate steps. The level 1 (L1) correction is a flat energy offset which compensates for the additional energy in a jet due to noise and pile up events. The level 2 (L2) correction is a relative correction that removes variation in detector response as a function of η. Lastly, the level 3 (L3) correction is an absolute correction independent of η and adjusts the detector response to remove variations due to jet p_{T}. Applying all three steps yields a corrected jet energy

$$E_{\text{Corrected}} = (E_{\text{Uncorrected}} - E_{\text{Offset}}) \times C_{\text{Rel}}(\eta, p''_{\mathrm{T}}) \times C_{\text{Abs}}(p'_{\mathrm{T}}) \tag{5.3}$$

Here, p''_{T} is the transverse momentum of the jet corrected for the L1 offset and $p'_{\mathrm{T}} = p''_{\mathrm{T}} \times C_{\text{Rel}}(\eta, p''_{\mathrm{T}})$ is the transverse momentum of the jet corrected for both the L1 offset and the L2 relative η dependence [60].

These corrections can be determined either from MC simulations or, more recently, from data-driven techniques from known physics processes for in-situ jet calibration [61]. For the PF jets used in this analysis, the jet energy scale corrections were found to be of order 5% in the barrel region, and of order 10% in the endcap region.

5.5 Missing Transverse Energy

As discussed in Chapter 2, the lightest supersymmetric particle (LSP) in most SUSY models interacts only weakly with matter. As a result, any LSPs produced in collisions at the LHC will leave no signature in any of the CMS subdetectors. It is possible to detect their presence, however, in the form of missing transverse energy, $\not{E}_T$. Because the incoming particles in a collision possess no component of momentum in the transverse plane, conservation of momentum requires that the vector sum of the p_{T} of all objects in the event must vanish. Any imbalance in this sum may therefore be attributed to particles which leave

no signature the detector. We define the particle flow $\not\!\!E_T$ as the magnitude of the vector sum of the p_{T} for all PF objects in the event:

$$\not\!\!E_T = \left| -\sum_{i}^{n} \overrightarrow{p_{\mathrm{T}i}} \right| \tag{5.4}$$

where i is the index of each PF object.

The calculation of the $\not\!\!E_T$ involves every particle in the event, and is therefore sensitive to mismeasurements in the p_{T} of any reconstructed object. In the case of QCD multijet events, $\not\!\!E_T$ reconstructed in the event is exclusively a consequence of jet momentum mismeasurements. The resolution of $\not\!\!E_T$ is lower than that of leptons in the event, but is improved by the use of noise reduction techniques and PF objects [62].

Chapter 6

Event Selection

To show this diagram properly, I would really need a four dimensional screen. However, because of government cuts, we could manage to provide only a two dimensional screen.

— Stephen Hawking

The first step of the analysis is to select events based on the reconstructed physics objects defined in Chapter 5. The goal of this analysis is twofold: first, to measure the $t\bar{t}$ production cross section using events with a single muon and jets, and second, to search for possible SUSY-motivated new physics signatures using events with same-sign dileptons. In many cases, the event selection criteria for these two channels are quite similar, and can be discussed in parallel.

The purpose of the selection is to attempt to extract, from all events in data, the small set of events taken to be the signal. For the first part of this analysis, we select events with a muon and multiple jets in order to isolate potental $t\bar{t}$ events for a cross section measurement. In the second part of this analysis, we apply the much more stringent requirement of including a second lepton with the same charge as the first. Because same-sign dileptons are so rare in the Standard Model, this requirement is designed to remove almost all potential

backgrounds to the SUSY search.

6.1 Data Samples and Trigger Selection

The data used in this analysis were collected from April to November of 2010 in pp collisions at $\sqrt{s} = 7\,\text{TeV}$. They are divided into two data-taking periods, separated by a technical stop of the LHC at the beginning of September, and are referred to as "Run2010A" and "Run2010B". It is necessary that we use only the subset of these runs for which the entire CMS detector is fully operational. To ensure this we apply a certified list of "good" runs to the full datasets, where the runs defined by this list correspond to a total integrated luminosity of $36 \pm 1\,\text{pb}^{-1}$ [63].

The data is further sorted into primary datasets (PDs) based broadly on the type of trigger that the events pass, as shown in Table 6.1. Events in the "electron/photon" (EG) PD, for example, must pass some trigger based on the presence of an electron or photon, while events in the MultiJet PD are required to pass triggers based on multiple jets or the H_{T} variable (defined below). The H_{T} triggers that are used for parts of this analysis have been sorted into several different PDs, and in some cases there is significant overlap, with the same event appearing in more than one PD. To avoid possible double counting of events, only the two largest disjoint samples are used, corresponding to a total luminosity of of $31 \pm 1\,\text{pb}^{-1}$.

For each PD there is a corresponding set of triggers used in this analysis, as shown in Table 6.2. For the $t\bar{t}$ production cross section measurement, which relies on the presence of a muon, events are required to pass one of two the single muon triggers in the middle column, depending on when the data were taken. For the SUSY-based search, trigger requirements are defined by search region: For the high H_{T} / low lepton p_{T} region, all events are required

Data samples		
Time period	2010A	2010B
Electrons	/EG/Run2010A-Dec22ReReco_v1	/Electron/Run2010B-Dec22ReReco_v1
Muons	/Mu/Run2010A-Dec22ReReco_v1	/Mu/Run2010B-Dec22ReReco_v1
Jets	/JetMET/Run2010A-Dec22ReReco_v1	/MultiJet/Run2010B-Dec22ReReco_v1

Table 6.1: Data samples used for the analysis. The total luminosity is $36 \pm 1\,\text{pb}^{-1}$ for the electron and muon datasets and $31 \pm 1\,\text{pb}^{-1}$ for the jet dataset.

Trigger selection			
Run range	e triggers	μ triggers	H_T triggers
132440 – 140040	HLT_Ele10_LW_L1R	HLT_Mu9	HLT_HT100U
140041 – 143962	HLT_Ele15_SW_L1R	HLT_Mu9	HLT_HT100U
143963 – 146427	HLT_Ele15_SW_CaloEleId_L1R	HLT_Mu9	HLT_HT100U
146428 – 147116	HLT_Ele17_SW_CaloEleId_L1R	HLT_Mu9	HLT_HT100U
147117 – 148058	HLT_Ele17_SW_TightEleId_L1R	HLT_Mu15_v1	HLT_HT140U
148058 – 149442	HLT_Ele17_SW_TightEleId_L1R	HLT_Mu15_v1	HLT_HT150U_v3

Table 6.2: List of electron, muon and H_T triggers and the run ranges in which they have been used.

to pass one of the H_T triggers shown in the rightmost column. For the high lepton p_T / low H_T region, events with two same-sign electrons must pass an electron trigger, while events with same-sign muons must pass a muon trigger. In order to prevent double-counting of events appearing in two separate PDs, events with an electron and a muon are exclusively taken from the muon triggers.

In each case, the trigger requirement depends both on the presence of a physics object and on a p_T threshold. Thus, the trigger HLT_Mu9 requires the presence of a muon in the high level trigger with a transverse momentum of at least $9\,\text{GeV}$. In order to maximize the selection of interesting events, this threshold is set as low as possible. However, due to the steeply increasing instantaneous luminosity during data taking, this threshold frequently had

to be increased, while lower-threshold trigger paths became prescaled. (A prescale factor of n implies that the trigger selects only one out of ever n events passing this path.) For this analysis, maintaining a high trigger efficiency is crucial, so in all cases the lowest-threshold unprescaled trigger was used.

Most trigger paths, particularly those used in the second half of the 2010 run, were not implemented in the simulation, so the trigger requirements imposed on the Monte Carlo samples are slightly different. For electrons, the samples were required to pass HLT_Ele17_SW_L1R_v2, which is similar to the triggers used in data. In the muon path, HLT_Mu9 was available in the simulation, as was HLT_HT150U_v3 in the H_{T} path, so these triggers were also used in the Monte Carlo samples. In both data and simulation, trigger efficiency weights are applied to each event, as discussed in Section 6.7. These weights are measured directly from data, and in all cases the efficiencies are close to 1.0.

6.2 Definitions and SUSY Search Regions

Selection of events for both the $t\bar{t}$ production cross section measurement and the SUSY search rely on several composite variables, defined here:

- The distance between two objects in η-ϕ space is used in a variety of contexts. It is defined as

$$\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{6.1}$$

  where $\Delta\eta$ ($\Delta\phi$) is the difference between the η (ϕ) coordinates of the two objects.

- H_{T} is a measure of the total hadronic energy in the event. It is defined as the scalar

sum of the p_{T} of all jets in the event that pass selection requirements:

$$H_{\mathrm{T}} \equiv \sum_{i=0}^{n} p_{\mathrm{T}}^{i} \tag{6.2}$$

where i is an index that runs over all n jets.

- Several variables are used to determine if a lepton is isolated from other high-p_{T} particles in the event. The tracker isolation variable I_{trk} is the sum of the transverse momenta of all tracks within a cone of $\Delta R < 0.3$ around the lepton (excluding the lepton track itself) and the calorimeter isolation variable I_{E} (I_{H}) is the sum of the energies of all ECAL (HCAL) towers within $\Delta R < 0.3$ around the lepton (excluding the energy from the lepton itself). The total relative isolation is then defined as

$$I_{\mathrm{rel}} \equiv \frac{I_{\mathrm{trk}} + I_{\mathrm{E}} + I_{\mathrm{H}}}{p_T^l} \tag{6.3}$$

- We obtain the invariant mass of a lepton pair by taking the vector sum of their individual four-momenta and using this resulting four-momentum to calculate the invariant mass. Thus,

$$M_{ll} = (E_{l1} + E_{l2})^2 - ||p_{l1} + p_{l2}||^2 \tag{6.4}$$

Our new physics search strategy relies on two different types of triggers: lepton triggers, which are appropriate for triggering on events in which we expect relatively low H_{T}, and H_{T} triggers, which are optimal for events with lower lepton p_{T}. It is necessary to employ both trigger strategies in order to cover the widest possible phase space, as can be illustrated by the SUSY cascade decay shown in Figure 2.3.

The unknown masses of the SUSY particles strongly determine the characteristics of the signature. The difference $\Delta m_{\widetilde{q}\widetilde{\chi}^{\pm}}$ between the chargino mass scale and the mass scale of

the gluinos and squarks determines the momenta of the produced quarks, and hence defines the hadronic activity in the event. Similarly, the difference $\Delta m_{\tilde{\chi}^{\pm}\tilde{\chi}^{0}}$ between the chargino and the LSP influences the lepton p_{T} spectrum. Thus, we expect this type of decay to fall into one of two general regions: high lepton p_{T} and low H_{T} if $m_{\tilde{\chi}^{\pm}}$ is closer to $m_{\tilde{\chi}^{0}}$, or high H_{T} and low lepton p_{T} if $m_{\tilde{\chi}^{\pm}}$ is closer to $m_{\tilde{q}}$.

6.3 Muon Selection

The Muon PD contains roughly 49 million events. Starting from this sample, muon candidates from the reconstructed muon collection must pass several selection requirements in order to be identified as muons for the purpose of this analysis.

- For muon-triggered events, the leading (highest-p_{T}) muon in the event is required to be "matched" to the HLT triggering object in η–ϕ space via the requirement $\Delta R(\mu_{\mathrm{REC}}, \mu_{\mathrm{HLT}}) < 0.1$. This ensures the leading muon triggered the event.

- The leading muon must have $p_{\mathrm{T}}^{\mu 1} > 20\,\mathrm{GeV}/c$ (for muon-triggered events) or $p_{\mathrm{T}}^{\mu 1} > 10\,\mathrm{GeV}/c$ (for electron- or H_{T}-triggered events). If there is a second muon in the event (as for the $\mu\mu$ channel of the SUSY-based search), it must have $p_{\mathrm{T}}^{\mu 2} > 10\,\mathrm{GeV}/c$. The p_{T} requirement must be sufficiently higher than the trigger threshold to ensure that the muon is triggered on with high efficiency, and that the trigger efficiency will not be sensitive to small changes in the muon p_{T}. The second muon is not subject to this requirement, because this muon does not trigger the event.

- All muons must be in the fiducial volume of the detector, with the leading muon within $|\eta| < 2.1$ (for muon-triggered events) or $|\eta| < 2.4$ (for electron- or H_{T}-triggered events). Any subsequent muons in the event must have $|\eta| < 2.4$. The difference here

comes from the fact that the muon systems (and hence reconstructed muons) extend out to $|\eta| = 2.4$, but the muon trigger only reaches $|\eta| = 2.1$. The η range is thus increased for any muons in the event that are not triggered on.

- A muon candidate from either a top or SUSY decay is expected to be isolated from other high-p_{T} particles in the event. For the $t\bar{t}$ cross section measurement, muons are required to have a relative isolation $I_{\text{rel}} < 0.05$, and for the H_{T}-triggered SUSY-based search this is loosened to $I_{\text{rel}} < 0.15$. For the lepton-triggered SUSY-based search, muons of $p_{\mathrm{T}} > 20\,\text{GeV}/c$ must have a relative isolation $I_{\text{rel}} < 0.1$ and muons of $p_{\mathrm{T}} < 20\,\text{GeV}/c$ must have an isolation sum $I_{\text{trk}} + I_{\mathrm{E}} + I_{\mathrm{H}} < 2\,\text{GeV}/c$.

- To ensure a muon comes from the primary interaction, we require that the impact parameter with respect to the beamspot, $d_0(\text{bsp})$, is less than 0.02 cm.

- Muons are required to have a minimum number of "hits" both in the overall silicon tracking system (11 hits) and in the pixel system in particular (one hit).

- Muons used in the $t\bar{t}$ cross section measurement are required to have hits in at least two muon stations, because the DT requires two segments to trigger.

- In order to exclude muons coming from decays in jets, we require for the $t\bar{t}$ cross section measurement that a muon be separated from any selected jet with $p_{\mathrm{T}} > 30\,\text{GeV}/c$ in the event by $\Delta R(\mu, j_{30}) > 0.3$. This requirement was omitted from the SUSY search in order to make the treatment of muons and electrons as uniform as possible.

The muon selection variables are shown in Figures 6.1 and 6.2. They are shown with all selection criteria prior to the variable plotted already applied ("progressive" plots). All

events shown are required to pass the trigger selection and to have a good primary vertex. To further reduce the contribution from QCD, events are required to have $\not\!\!E_T > 30\,\text{GeV}$, analogous to the $\not\!\!E_T$ requirement for the SUSY-based search, discussed in Section 6.5 below. Approximately 110000 events survive all the muon requirements listed.

6.4 Electron Selection

Electrons are used exclusively in the SUSY-based search of this analysis, and play no role in the $t\bar{t}$ cross section measurement. The electron PD used for this portion of the analysis includes approximately 76 million events. In most cases the electron requirements are the same for the lepton- and H_{T}-triggered searches, but both selection criteria are quoted for the cases where they differ.

- For electron-triggered events (i.e. the ee channel of the lepton-triggered search), the leading (highest-E_{T}) electron in the event is required to be "matched" to the HLT triggering object in η–ϕ space via the requirement $\Delta R(e_{\text{REC}}, e_{\text{HLT}}) < 0.1$. This ensures the leading electron triggered the event.

- The leading electron must have $E_{\text{T}}^{e1} > 20\,\text{GeV}$ (for electron-triggered events) or $E_{\text{T}}^{e1} > 10\,\text{GeV}$ (for muon- or H_{T}-triggered events). If there is a second electron in the event (as for the ee channel of the SUSY-based search), it must have $E_{\text{T}}^{e2} > 10\,\text{GeV}$.

- All electrons must be within $|\eta| < 2.4$, to make the fiducial acceptance of the electrons identical to that of the muons.

- Electrons must also pass several identification criteria which make use of shower shape variables to help distinguish real electrons from jets. Several levels of the re-

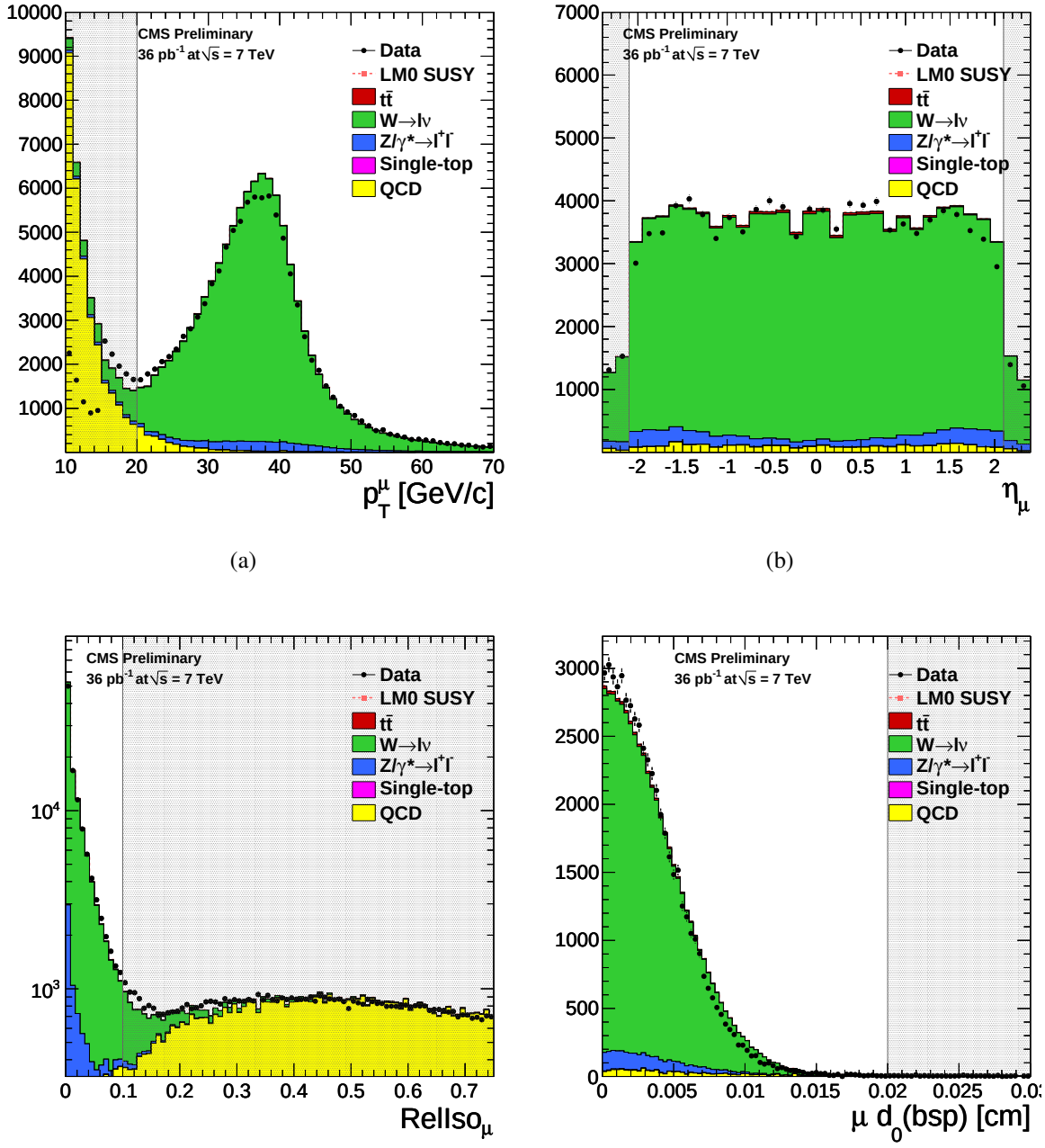
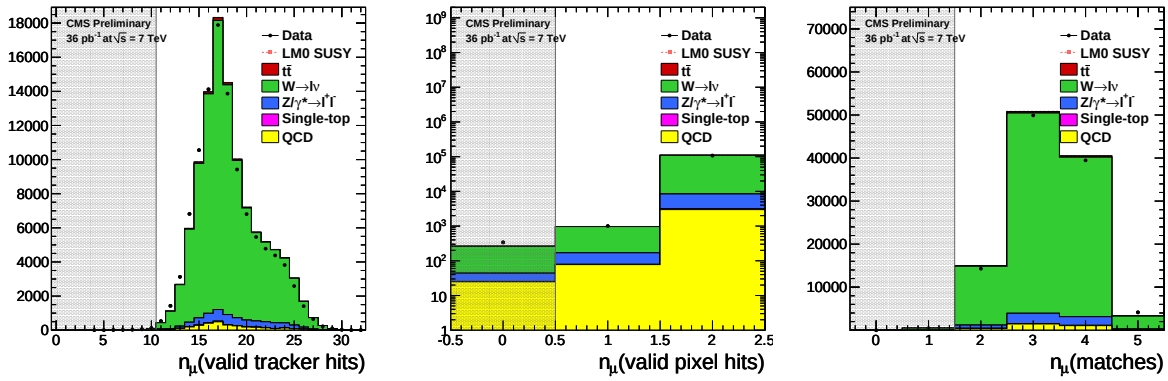


Figure 6.1: Distributions showing the comparison between data from the muon PD and MC simulation using the samples described in Table 4.1 for (a) muon p_{T}, (b) muon η, (c) muon relative isolation and (d) muon $d_0(\mathrm{bsp})$ (the impact parameter with respect to the beam spot). All MC samples have been normalized to the luminosity of the data. The shaded area represents the region rejected by the muon selection requirements.



(a) (b) (c)

Figure 6.2: Distributions showing the comparison between data and MC simulation for (a) the number of valid tracker hits, (b) the number of valid pixel hits and (c) the number of hits in muon stations.

quirements exist, corresponding to different electron ID efficiencies. In this case the "working point 80" criteria were used, designed to be 80% efficient at selecting real electrons. The variables used are shown below.

- The shower shape variable $\sigma_{i\eta i\eta}$ measures the width of the electromagnetic cluster in the ECAL in the η direction. A single electron is expected to leave a deposit with a small spread in the η direction, so $\sigma_{i\eta i\eta}$ is required to be less than 0.01 in the barrel and less than 0.03 in the endcaps. (No corresponding restriction exists for the spread in ϕ due to the effect of Bremsstrahlung.)
- $\Delta\phi$, the difference between the ϕ of the supercluster in the calorimeter and the ϕ of the track, extrapolated from the vertex, must be less than 0.06 in the barrel and less than 0.03 in the endcaps.
- $\Delta\eta$, analogously, is the difference between the η of the supercluster and the η of

the track. This must be less than 0.004 in the barrel and less than 0.007 in the endcaps.

  - The fraction of deposited energy in the HCAL and ECAL, $E_{\text{HCAL}}/E_{\text{ECAL}}$, is required to be less than 0.04 in the barrel and less than 0.025 in the endcaps.

- Electrons are also required to pass an isolation requirement. For the H_{T}-triggered SUSY-based search we require $I_{\text{rel}} < 0.15$. For the lepton-triggered SUSY-based search, electrons of $E_{\text{T}} > 20\,\text{GeV}$ must have a relative isolation $I_{\text{rel}} < 0.1$ and electrons of $E_{\text{T}} < 20\,\text{GeV}$ must have an isolation sum $I_{\text{trk}} + I_{\text{E}} + I_{\text{H}} < 2\,\text{GeV}$.

- To ensure an electron comes from the primary interaction, we require that the impact parameter with respect to the beamspot, $d_0(\text{bsp})$, is less than 0.02 cm.

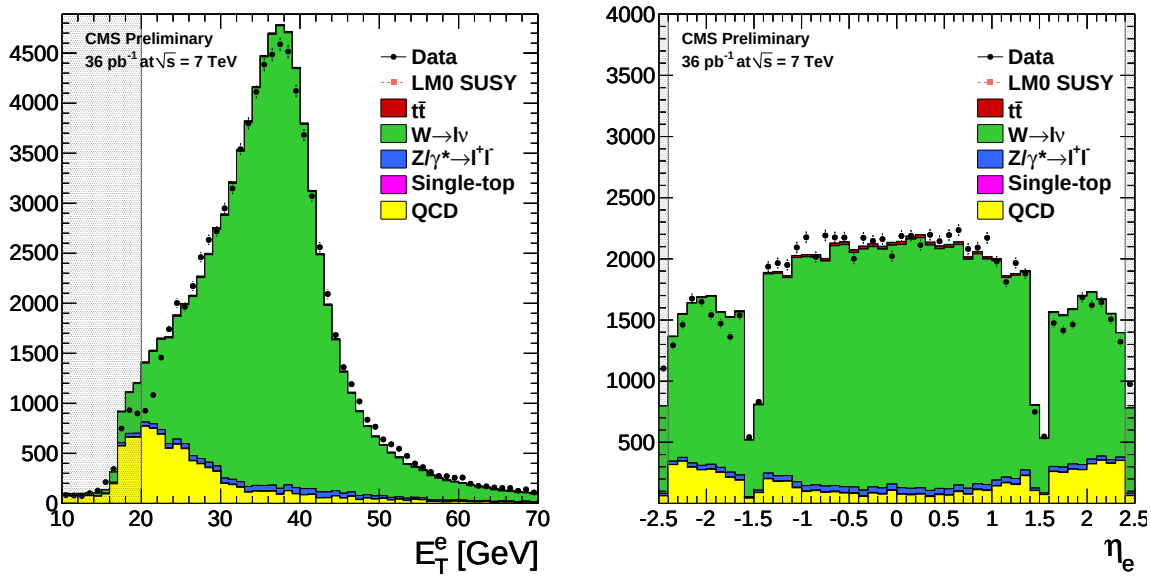
- To remove electrons that come from converted photons, rather than the primary interaction, we impose several conversion rejection criteria:

  - Because photons are neutral and leave no track, electrons coming from converted photons typically miss hits in the tracker. Thus, we require that there be no missed hits in the tracker.

  - If an electron comes from photon conversion, it is expected to have a partner track from a second electron. If a nearby track is found with charge opposite the electron within a distance 0.02 in the R–ϕ plane and within a $\Delta\cot\theta$ of 0.02, then the electron is rejected.

- Electrons within $\Delta R(e,\mu) < 0.1$ of a muon passing all selection requirements above are rejected, to remove electron candidates coming from muon bremsstrahlung.

The electron selection variables are shown in Figures 6.3 through 6.7. They are shown with all selection criteria prior to the variable plotted already applied ("progressive" plots). All events shown are required to pass the trigger selection and to have a good primary vertex. To further reduce the contribution from QCD, events are required to have $\not\!\!E_T > 30\,\text{GeV}$, analogous to the $\not\!\!E_T$ requirement for the SUSY-based search, discussed in Section 6.5 below. In total, about 90,000 events pass all electron selection criteria listed above.



(a) (b)

Figure 6.3: Distributions showing the comparison between data and MC simulation for (a) electron E_{T} and (b) electron η. All MC samples have been normalized to the luminosity of the data. The shaded area represents the region rejected by the electron selection requirements.

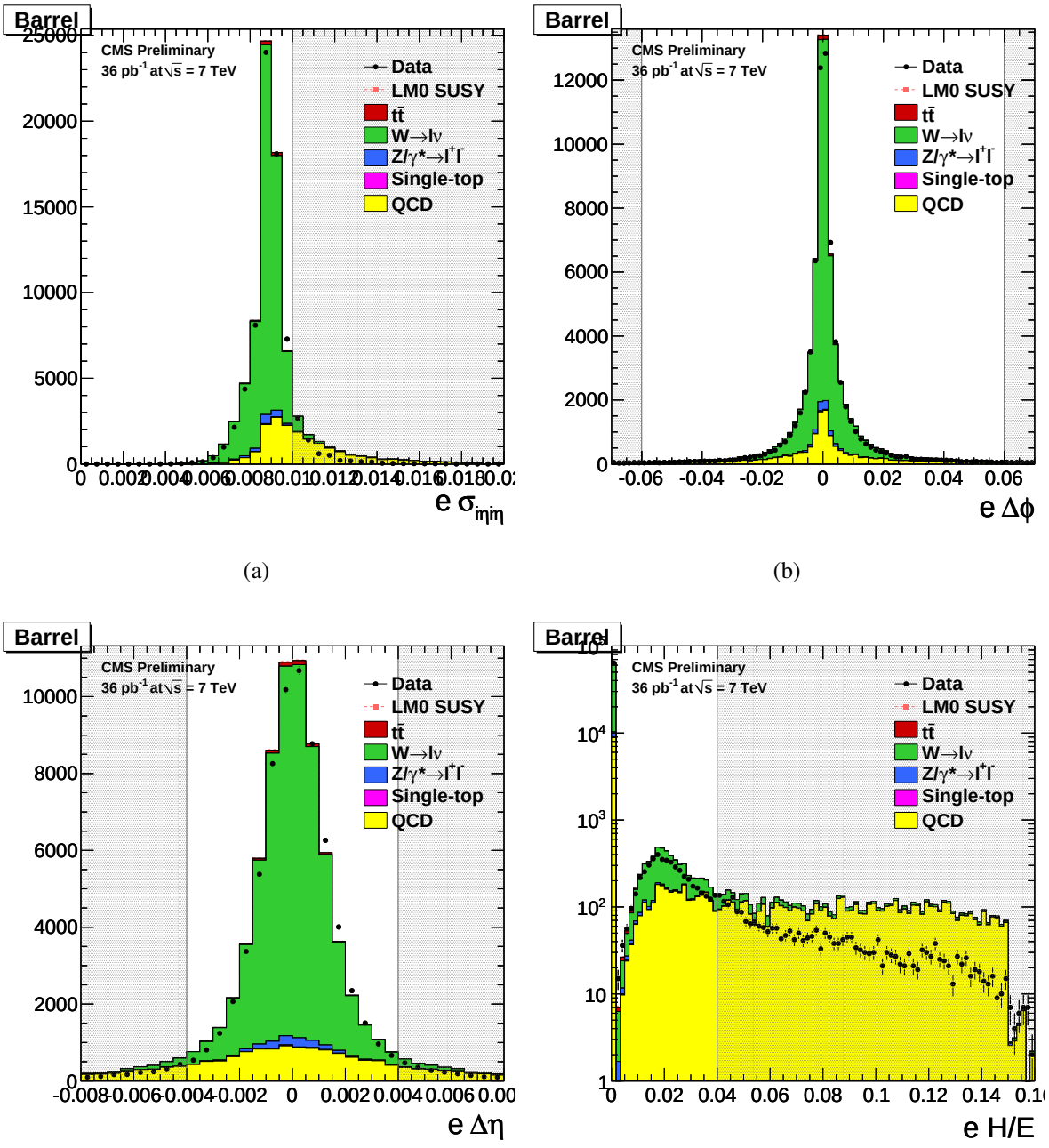


Figure 6.4: Distributions showing the comparison between data and MC simulation for the electron ID variables in the barrel region.

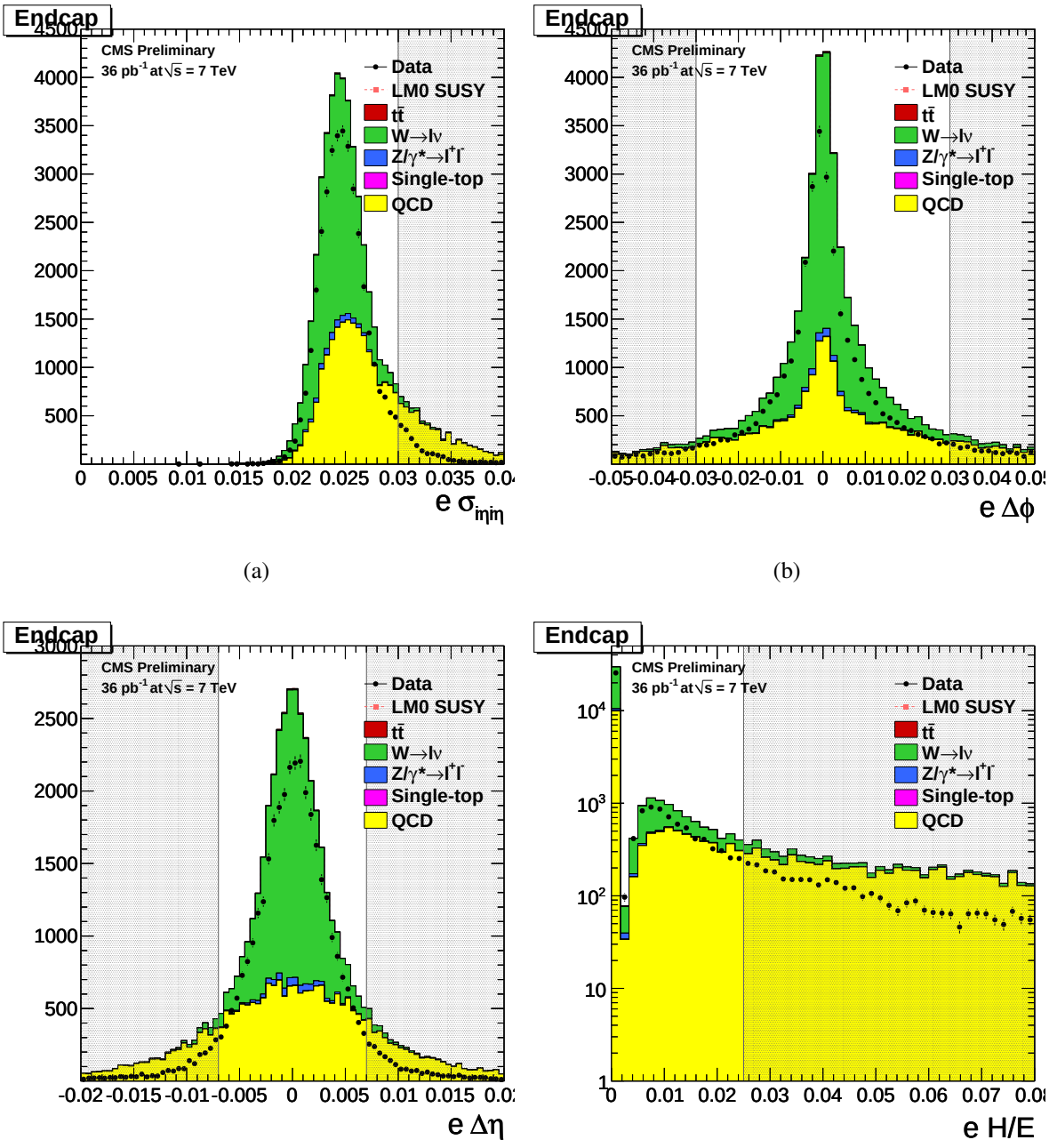
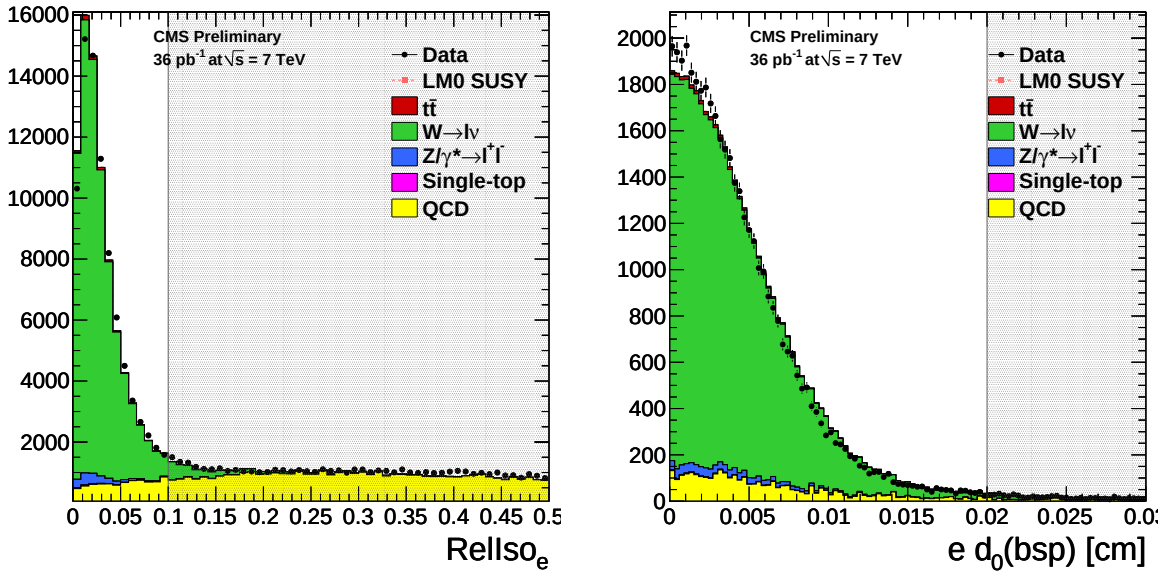
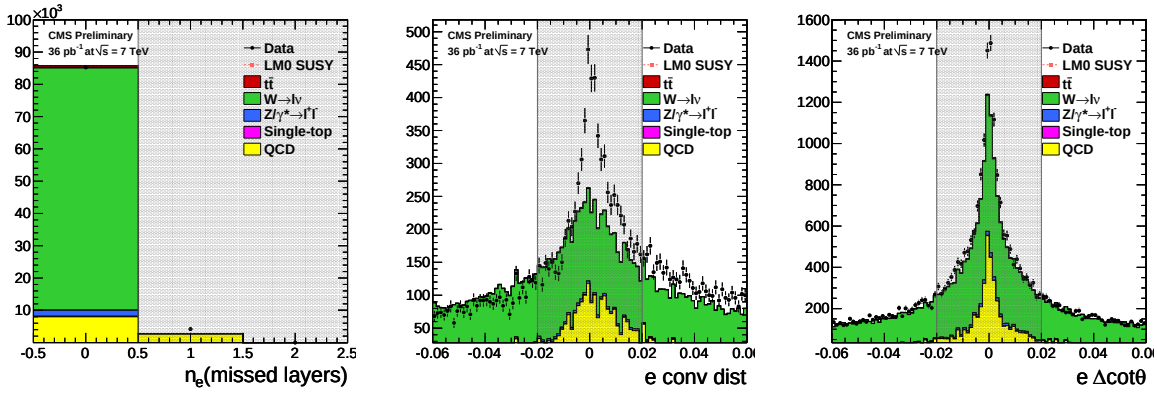


Figure 6.5: Distributions showing the comparison between data and MC simulation for the electron ID variables in the endcap regions.



(a) (b)

Figure 6.6: Distributions showing the comparison between data and MC simulation for (a) electron I_{rel} and (b) electron $d_0(\text{bsp})$.



(a) (b) (c)

Figure 6.7: Distributions showing the comparison between data and MC simulation for the conversion rejection variables.

6.5 Jet, H_{T} and $\not\!\!E_T$ Selection

As discussed in Chapter 5, this analysis uses jets reconstructed by the particle flow algorithm (PF jets) and imposes the following selection requirements

- All jets have the L1 pile-up offset, L2 relative and L3 absolute jet corrections applied, as described in Section 5.4.2. We require that the corrected jet $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$.

- Jets must lie within the fiducial acceptance of the detector. For the $t\bar{t}$ cross section measurement, $|\eta| < 2.4$, while for the same-sign dilepton search the range has been slightly extended to $|\eta| < 2.5$ in order to maintain consistency with the published result [64].

- Several additional ID criteria are imposed on jet candidates: Jets are required to have at least two constituent particles. The fractions of their charged electromagnetic energy (CEF), neutral electromagnetic energy (NEF) and neutral hadronic energy (NHF) are all required to be less than 0.99, while the fraction of their charged hadronic energy (CHF) must be greater than 0.0. Additionally, the jet must contain at least one charged hadron. In all cases, the energy fractions are calculated from uncorrected jets.

- The PF algorithm described in Section 5.1 does not explicitly differentiate jets from leptons, so it is necessary to "clean" the jets to ensure that the same object does not appear in both the jet collection and a lepton collection. For the same-sign dilepton search, we remove any jets which lie within a cone of $\Delta R < 0.4$ of a muon. For the $t\bar{t}$ cross section measurement, this cone is reduced to $\Delta R < 0.1$. Taken together with the muon selection, this requirement indicates that for the cross section measurement,

if $\Delta R(\mu, j_{30}) < 0.1$, the jet is removed, whereas if $0.1 < \Delta R(\mu, j_{30}) < 0.3$, the muon is removed.

The $t\bar{t}$ cross section measurement requires the presence of at least one jet passing all selection criteria, while the SUSY-based search requires the presence of at least two jets.

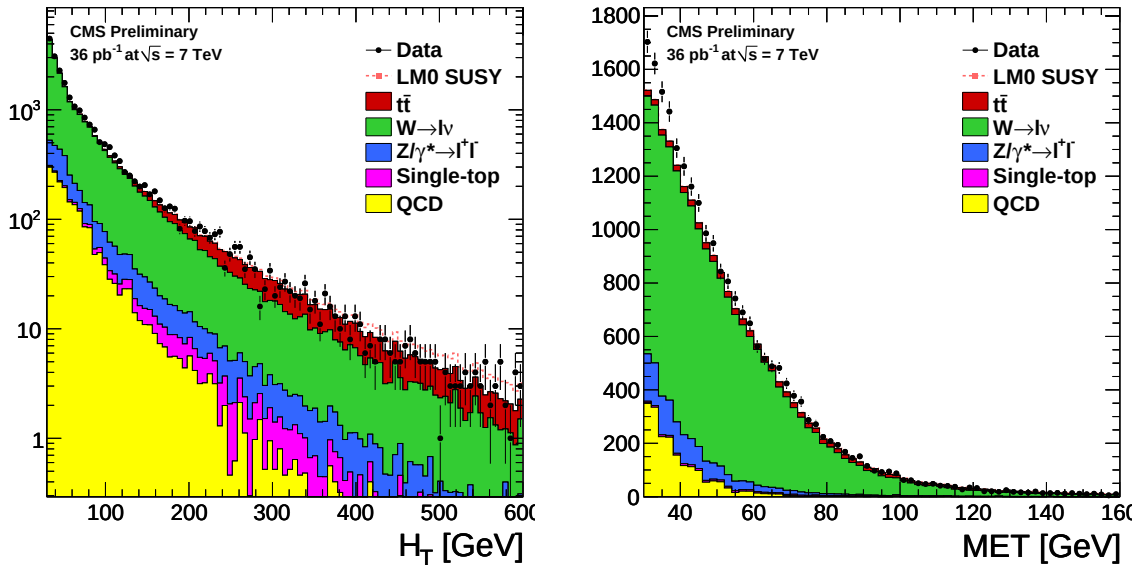
For the cross section measurement, no requirement is imposed on the $H_{\rm T}$ or $\not\!\!E_T$. For the $H_{\rm T}$-triggered SUSY search, we require $H_{\rm T} > 300\,\text{GeV}$ to ensure the trigger selection is fully efficient. For the lepton-triggered search, a baseline selection is imposed of $\not\!\!E_T > 30\,\text{GeV}$ ($20\,\text{GeV}$) for the ee and $\mu\mu$ ($e\mu$) channels. We further subdivide this search into two more search regions by requiring either $H_{\rm T} > 200\,\text{GeV}$ or $\not\!\!E_T > 80\,\text{GeV}$.

The $H_{\rm T}$ and $\not\!\!E_T$ selection variables are shown in Figure 6.8. All events are required to pass the trigger selection and to have a good primary vertex. In addition, events shown must have at least one muon and one jet.

For many of these distributions the shape of the sum of Madgraph MC samples matches that of the data, but there is significant disagreement in the overall number of events. Thus, it is not possible to naively take the normalization of the MC from the NLO calculation, and we must rely instead on a data-driven fitting method (described in Chapter 7) to fix the relative contributions of the different samples.

6.6 Other Selection Requirements

For the $t\bar{t}$ cross section measurement, we require exactly one muon passing all selection requirements. Dilepton decays are excluded by removing events containing a second, more loosely defined high-$p_{\rm T}$ lepton: the event must contain no second muon with $p_{\rm T} > 10\,\text{GeV}/c$, $|\eta| < 2.5$, $I_{\rm rel} < 0.2$, and no electron with $E_{\rm T} > 15\,\text{GeV}$, $I_{\rm rel} < 0.2$.



(a) (b)

Figure 6.8: Distributions showing the comparison between data and MC simulation for (a) H_{T} and (b) $\not\!\!E_T$. All MC samples have been normalized to the luminosity of the data.

For the SUSY-based search, we require at least two leptons ($\mu\mu$, $e\mu$ or ee) and these two leptons are required to have the same charge. In order to reduce the background from $Z \rightarrow l^+l^-$ decays, two separate "Z vetos" are imposed: First, if the two selected leptons have the same flavor, then the event is rejected if the invariant mass of the object reconstructed from them lies within a mass window of $10\,\text{GeV}$ of the Z-boson mass: $81\,\text{GeV}/c^2 < M_{ll} < 101\,\text{GeV}/c^2$. Second, if either of the two selected leptons have the same flavor as a third lepton in the event, then the event is rejected if the invariant mass of those leptons lies within $15\,\text{GeV}$ of the Z-boson mass: $76\,\text{GeV}/c^2 < M_{ll} < 106\,\text{GeV}/c^2$. We further reduce the background from low-mass resonances of heavy-flavor quarks by requiring events with same-flavor leptons to have an invariant mass $M_{ll} > 12\,\text{GeV}$ (for lepton-triggered events) or $M_{ll} > 5\,\text{GeV}$ (for H_{T}-triggered events).

6.7 Selection Efficiencies

The efficiency of our selection requirements is necessary to convert the final event yields into a physical result. The lepton efficiencies are parametrized into trigger, identification and isolation components, and were measured using a lepton sample collected from $Z \rightarrow l^+l^-$ decays. A high purity Z-boson sample can be extracted from data by requiring two leptons with an invariant mass within $15/gev$ of the Z-boson mass: $76\,\text{GeV}/c^2 < M_{ll} < 106\,\text{GeV}/c^2$. From this sample it is possible to determine the selection efficiencies via the "tag and probe" (T&P) method in both data and MC simulation. One lepton is treated as the "tag" and is required to satisfy all selection requirements, while the other lepton is the "probe", selected based on the measured variable. The efficiency is then determined from the number of Z-boson events passing or failing the selection criteria.

For the muon definition used in the $t\bar{t}$ cross section measurement, the efficiency of

finding an isolated muon was measured by this method to be 0.880 ± 0.002 (stat + syst), and the efficiency to trigger on muons was found to be 0.922 ± 0.002 (stat) [65]. The cross section measurement uses a simultaneous fit to the MC samples defined in Table 4.1, so it is particularly important to know the ratio of the efficiencies in data to the efficiencies in MC, as determined by the T&P method. This is applied as a scale factor to the final results obtained from the fit. For muons these scale factors were found to be very close to unity: 0.995 ± 0.003 for the isolation requirement, and 0.970 ± 0.002 for the trigger.

It is also necessary to measure the trigger efficiencies in the same-sign dilepton SUSY search, where the different search regions are defined by their trigger characteristics. For the lepton-triggered region, the trigger efficiency is expected to be very high, because only one of the two leptons in the event needs to trigger. The signal efficiency as measured by the T&P method in Monte Carlo is estimated to be 0.99 ± 0.01 for events passing the baseline selection.

A different method is used for the H_{T}-triggered region, to determine the trigger efficiency directly from data. Because the rate for a muon to fake a jet is low, the set of events passing a muon trigger is statistically independent of the set of events passing an H_{T} trigger. We thus take events from the Mu PD as an orthogonal trigger sample. The H_{T} of these events is then calculated using jets passing all selection criteria. Figure 6.9 shows the efficiency for each H_{T} trigger to fire as a function of the fully reconstructed H_{T} in the event. The three triggers shown correspond to the H_{T} triggers in Table 6.2, with roughly half of the total luminosity being taken with the highest threshold trigger. This measurement indicates that at $H_{\mathrm{T}} = 300\,\mathrm{GeV}$ the trigger efficiency is 0.94 ± 0.05.

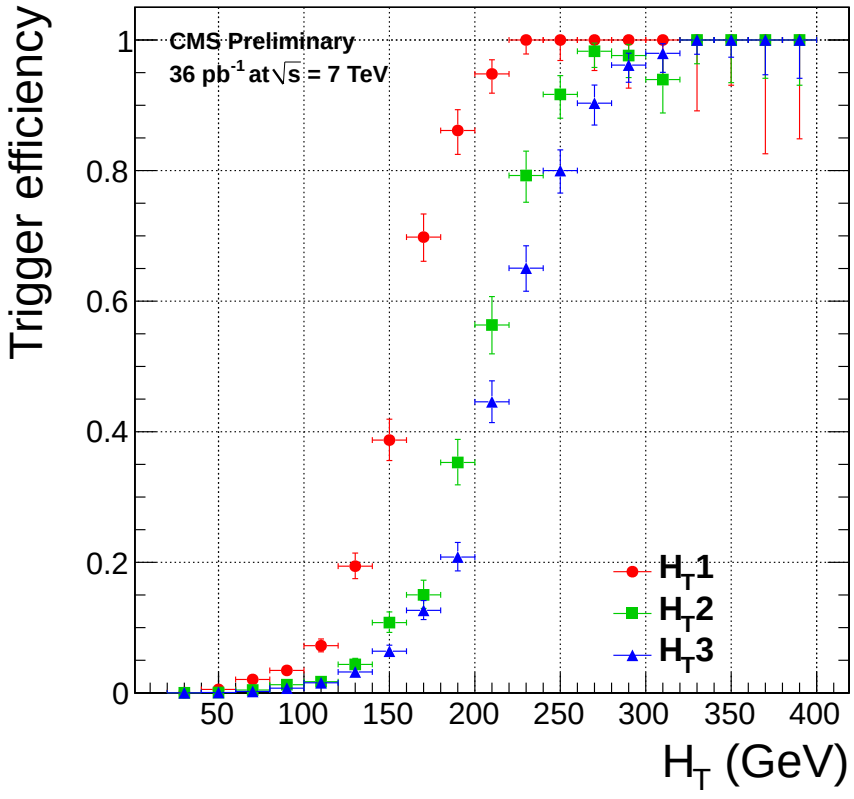


Figure 6.9: H_{T} trigger efficiency as a function of the reconstructed H_{T} for three data-collecting periods: $7\,\mathrm{pb}^{-1}$ with $H_{\mathrm{T}}1$, $10\,\mathrm{pb}^{-1}$ with $H_{\mathrm{T}}2$, and $18\,\mathrm{pb}^{-1}$ with $H_{\mathrm{T}}3$.

Predicted and observed event yields for 36.0 pb^{-1}							
Cut	$t\bar{t}$	$W \to l\nu$	$Z/\gamma^* \to l^+l^-$	Single-top	QCD	Sum MC	Data
1 jet	29	17565	1655	63	6762	26075	25183
2 jets	116	3253	279	69	764	4480	4816
3 jets	199	555	54	31	111	950	1049
$\geq$ 4 jets	224	118	13	12	19	386	406

Table 6.3: Predicted and observed event yields for the $t\bar{t}$ selection in 36 pb^{-1}. Note that the QCD yields assume a cross section that is a factor of two greater than the theoretical cross section. The MC numbers include the muon trigger, ID and isolation scale factors.

6.8 Event Yields for Top Analysis

Based on the event selection criteria described above, we can calculate the predicted event yields from MC simulation as

$$N = \epsilon_{\text{MC}} \cdot \sigma_{\text{theory}} \cdot \mathcal{L}_{\text{int}} \tag{6.5}$$

where N is the number of predicted events, ϵ_{MC} is the Monte Carlo acceptance and efficiency, σ_{theory} is the theoretical cross section from Table 4.1 and $\mathcal{L}_{\text{int}}$ is the integrated luminosity.

Table 6.3 and Figure 6.10 show the results of this calculation for all MC samples as well as data. For QCD we used twice the theoretical cross section to estimate the predicted yield. A comparison of the various MC contributions shows that $t\bar{t}$, $W \to l\nu$ and QCD compose the majority of events in each jet bin.

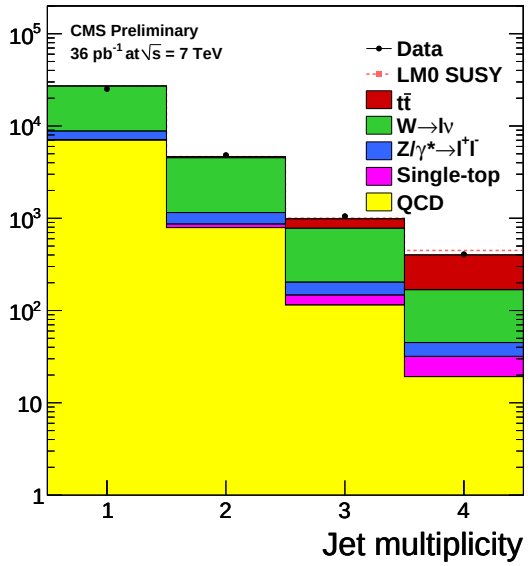


Figure 6.10: Jet multiplicity for the $t\bar{t}$ selection in 36 pb^{-1}.

Chapter 7

Analysis Method and Background Modeling

God does not play dice with the universe; He plays an ineffable game of His own devising, which might be compared, from the perspective of any of the other players [i.e. everybody], to being involved in an obscure and complex variant of poker in a pitch-dark room, with blank cards, for infinite stakes, with a Dealer who won't tell you the rules, and who ***smiles all the time.***

— Terry Pratchett

7.1 Comparison of Data and MC

An initial comparison between data from the muon dataset and Madgraph Monte Carlo samples for all SM processes (including $t\bar{t}$, single top, W/Z and QCD) after the full $t\bar{t}$ event selection discussed in Chapter 6 is shown in Figures 7.1 through 7.5, split between the 2 jets, 3 jets and $\geq$ 4 jets bins. Also shown for comparison is a sample point "LM0" in the constrained Minimal Supersymmetric Standard Model (CMSSM) [66]. LM0 produces squarks and gluinos with relatively low mass, and the production cross section is therefore higher, making this an ideal new physics reference point. Figures 7.1 and 7.2 show the muon variables, including the transverse momentum p_{T}, pseudorapidity η and relative

isolation I_{rel} described in the previous chapter. The Monte Carlo samples used here are "out of the box" in the sense that they have been weighted by the product of the luminosity and their NLO cross sections, but no attempt has been made to match the number of events to that observed in data. There is reasonable agreement between the data points and the total Madgraph MC, but we see that the simulation consistently underestimates the data. Also shown is the transverse mass of the reconstructed W-boson, defined as the transverse mass obtained from the muon and the missing transverse energy:

$$M_T^2(W) = 2p_{\text{T}}^{\mu} \not{E}_T(1 - \cos\phi) \tag{7.1}$$

where ϕ is defined as the angle between the muon p_{T} and the $\not{E}_T$ directions in the transverse plane. This distribution is more sharply peaked around the W-boson mass for leptonic W decays because the $\not{E}_T$ approximates the energy of the neutrino in the decay $W \to l\nu$.

Figures 7.3 through 7.5 show the corresponding jet variables split by jet multiplicity. The p_{T} and η distributions are shown for the lowest-p_{T} jet in the event. Also shown is the variable $M3$, defined as the invariant mass of the combination of the three jets with the largest vectorially summed transverse momentum. It approximates the mass of the hadronically-decaying top quark and thus peaks around the top quark mass in semileptonic $t\bar{t}$ decays.

The yields predicted by the Madgraph MC simulation are calculated by multiplying the selection efficiencies for each process, as determined by simulation, by the appropirate NLO or NNLO cross section times the total integrated luminosity of $36\,\text{pb}^{-1}$. The overall shape of the Madgraph MC is in good agreement with the distributions observed in muon data. However, more events are seen in this data than predicted by the simulation, indicating that the cross sections are underestimated. Indeed, the excess of data over MC, particularly

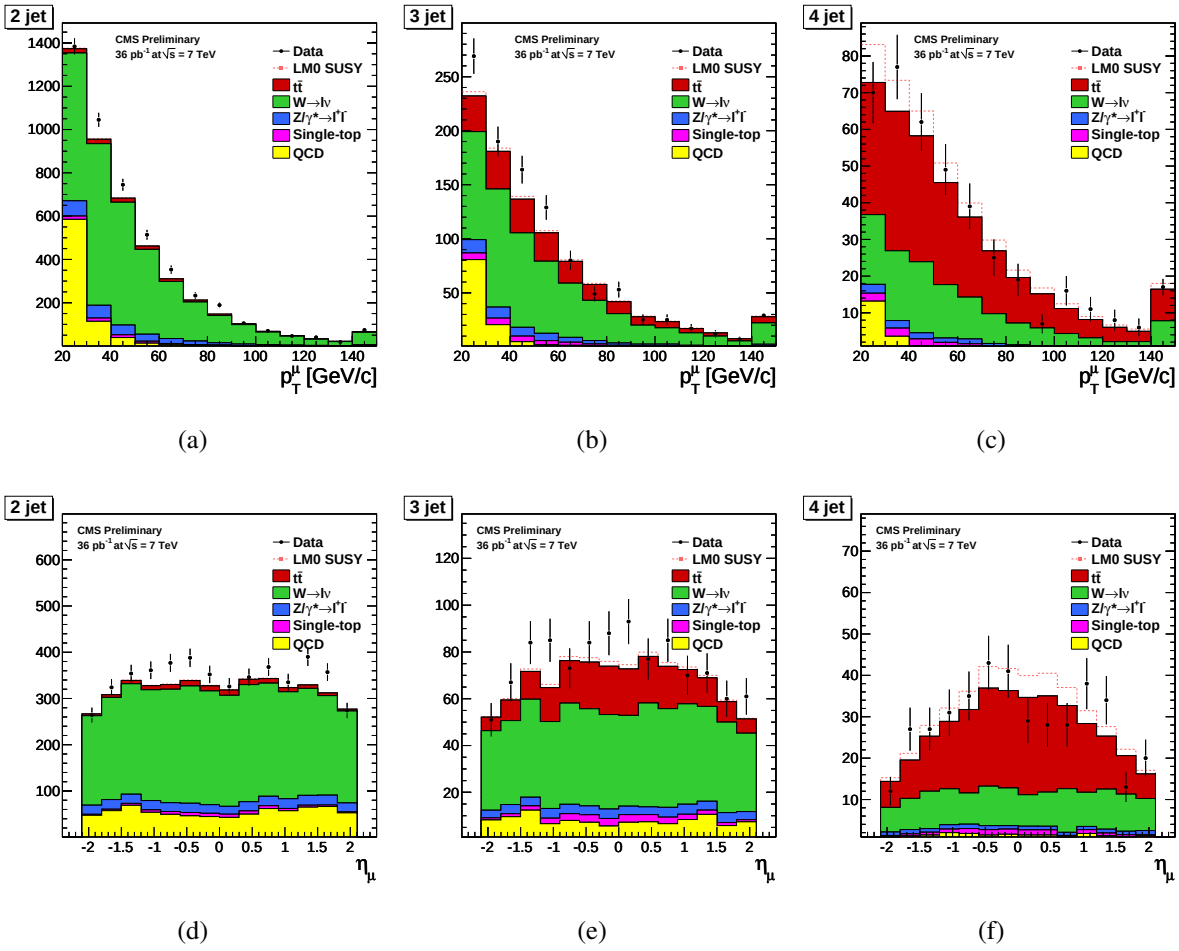


Figure 7.1: Muon p_{T} distribution (top row) and η distribution (bottom row) comparison between muon data and Madgraph Z2 Monte Carlo normalized to the NLO cross section for 2 PF jets (left column), 3 PF jets (middle column) and 4 PF jets (right column) after the full $t\bar{t}$ event selection of Chapter 6.

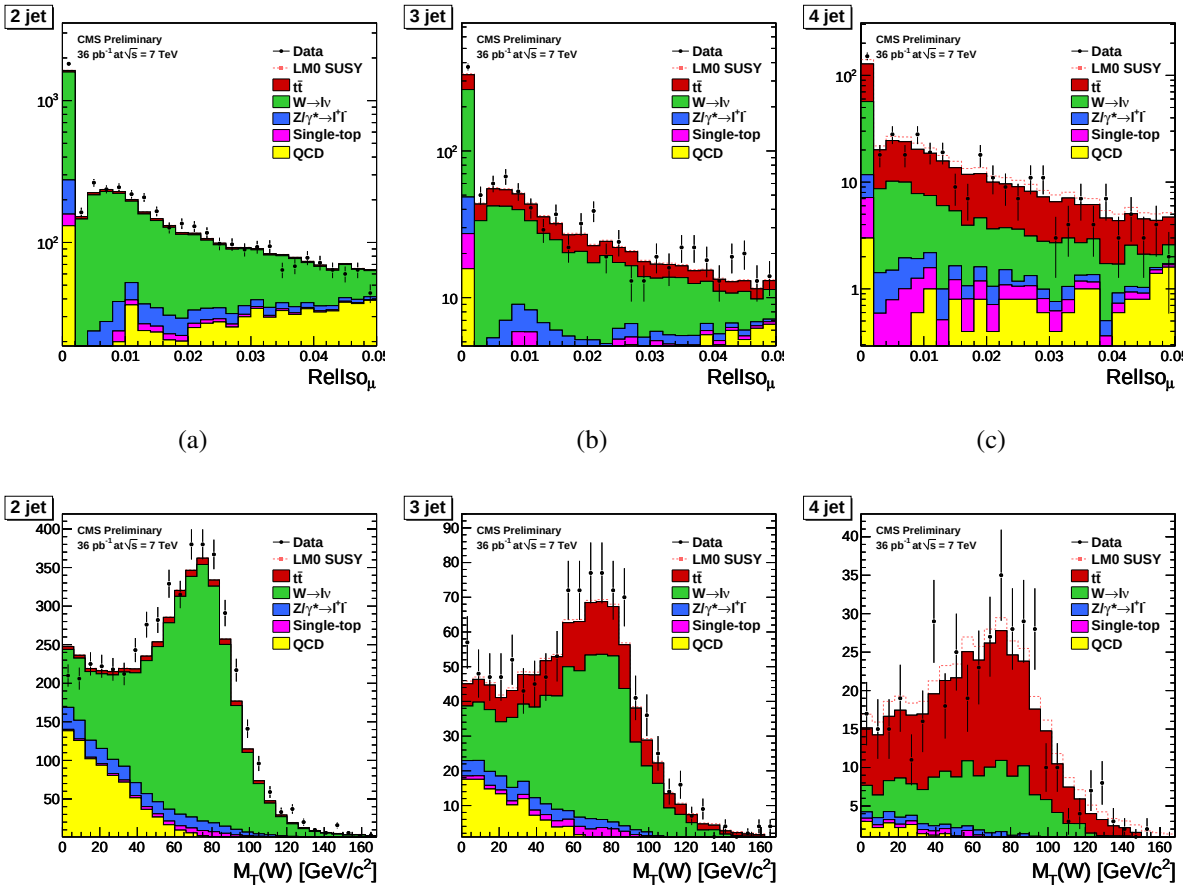


Figure 7.2: Muon relative isolation distribution (top row) and $M_T(W)$ distribution (bottom row) comparison between data and Madgraph Z2 Monte Carlo normalized to the NLO cross section for 2 PF jets (left column), 3 PF jets (middle column) and 4 PF jets (right column) after the full $t\bar{t}$ event selection.

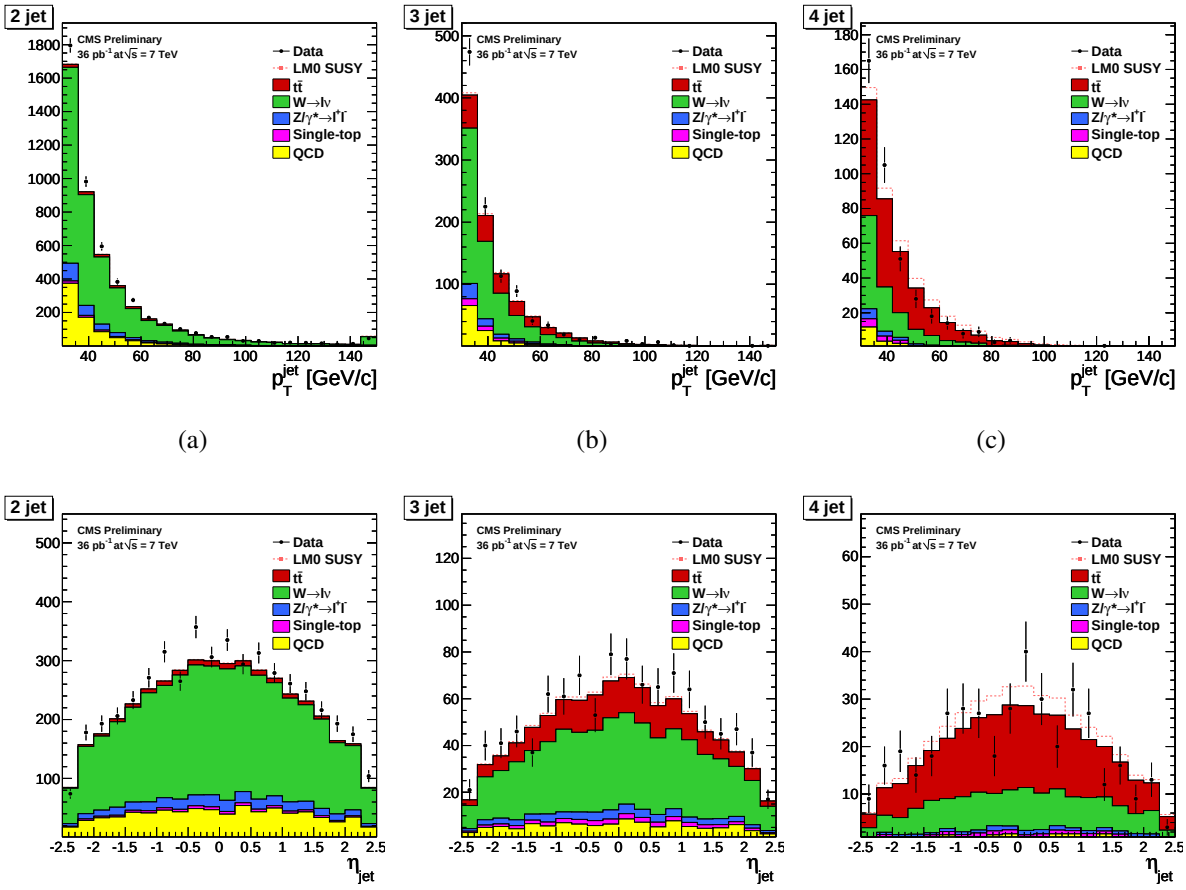
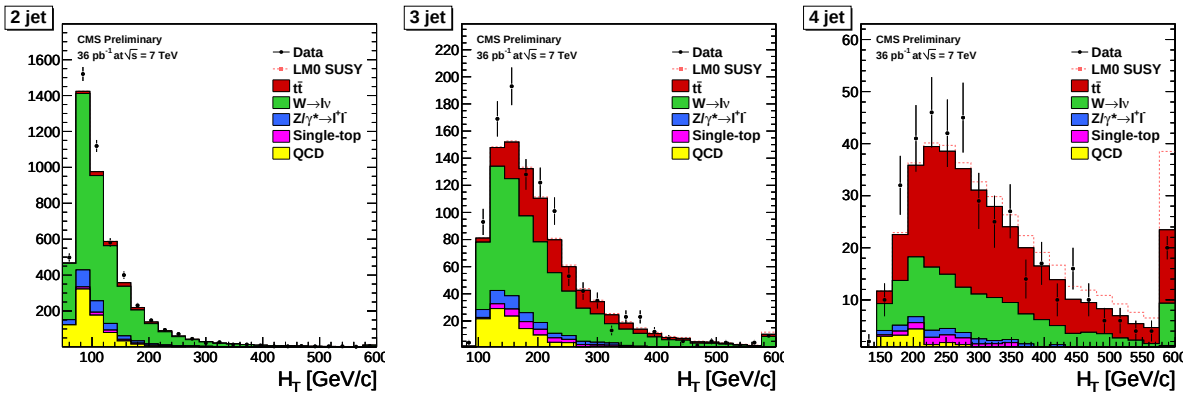
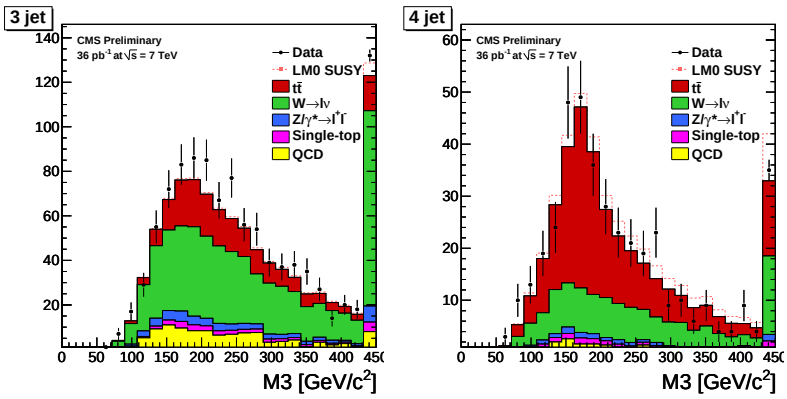


Figure 7.3: Lowest-p_{T} PF jet p_{T} distribution (top row) and η distribution (bottom row) comparison between data and Madgraph Z2 Monte Carlo normalized to the NLO cross section for 2 PF jets (left column), 3 PF jets (middle column) and 4 PF jets (right column) after the full $t\bar{t}$ event selection.



(a) (b) (c)

Figure 7.4: PF jet H_{T} distribution comparison between data and Madgraph Z2 Monte Carlo normalized to the NLO cross section for (a) 2 PF jets, (b) 3 PF jets and (c) 4 PF jets after the full $t\bar{t}$ event selection.



(a) (b)

Figure 7.5: PF jet $M3$ distribution comparison between data and Madgraph Z2 Monte Carlo normalized to the NLO cross section for (a) 3 PF jets and (b) 4 PF jets after the full $t\bar{t}$ event selection.

at low-p_{T}, is quite large compared to contributions from the LM0 signal. Thus, the overall normalization to data must be determined experimentally, or it will not be possible to separate this signal from the effect of MC normalization. Our method for determining this normalization via simultaneous template fits is described in the next section.

7.2 Cross Section Extraction Strategy

The $t\bar{t}$ cross section is measured in a data sample containing events with a single muon and jets in the final state, as described in Chapter 6. We use the kinematic properties of these events to distinguish contributions from each component process by applying a binned maximum likelihood fit to different distributions simultaneously. The MC shapes of the distributions are used as templates for the fit and these templates are then varied by the fit algorithm to determine the overall normalization of the process. From Equation 6.5, this normalization is directly related to the cross section of the process.

As can be seen from Table 6.3, the major contributions to the sample come from the processes $t\bar{t}$, $W \rightarrow l\nu$, and QCD. No single kinematic distribution shown can clearly distinguish between all three of these processes, but several distributions provide distinct shapes that can be used to distinguish one process from the other two. The shape of the jet multiplicity spectrum, shown in Figure 7.6, is very different between $t\bar{t}$, which tends to have high jet multiplicity, and both $W \rightarrow l\nu$ and QCD, which have an exponentially falling spectrum. We expect a high number of jets from semileptonic $t\bar{t}$ decays, due to the two b-quarks and the hadronically-decaying W-boson, making jet multiplicity suitable for separating out $t\bar{t}$ events. Although no new physics sample is used in the fit, the shape of the LM0 jet multiplicity spectrum is shown for comparison. Due to the high jet production rates of SUSY, its spectrum most nearly matches that of the $t\bar{t}$ sample.

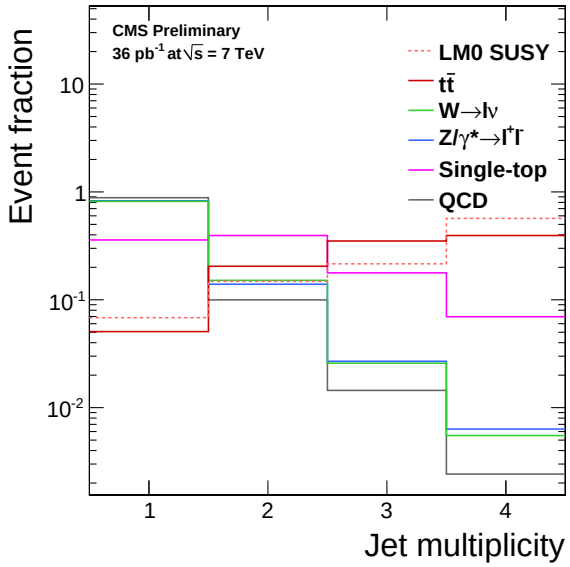


Figure 7.6: Shape comparison between all Monte Carlo samples for jet multiplicity. All samples are normalized to unit area.

Figure 7.7 shows the shapes of the muon p_{T} and η distributions for each MC process. QCD generally produces low-p_{T} leptons inside jets, and hence has a much softer muon p_{T} spectrum than the other processes. Muons produced from W and $t\bar{t}$ come directly from the decay of a W boson, and receive a significant contribution to their momentum from the W rest mass. Further, the muons produced from top decays are more centrally produced, and have lower η values, while those from QCD are more democratically distributed. In each case, the distinguishing power between the three dominant processes is good, and muon variables make a particularly attractive choice of templates because they are unaffected by the major jet systematic uncertainties. The muon p_{T} and η spectra for LM0 are similar to $t\bar{t}$ because they are also produced by heavy particle decays (in this case the chargino, as in Figure 2.3).

Statistical and systematic uncertainties have been studied for simultaneous fits to a

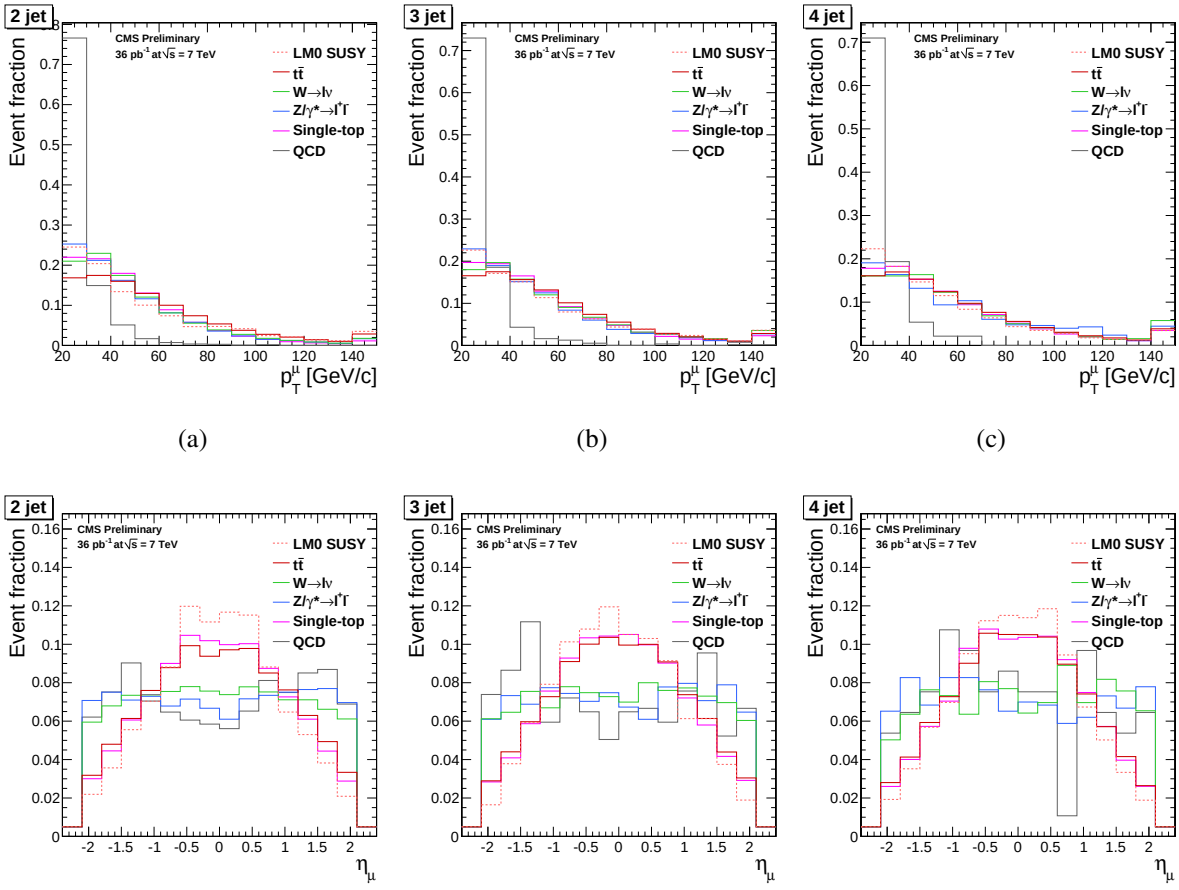


Figure 7.7: Shape comparison between all Monte Carlo samples for the muon p_{T} distribution (top row) and η distribution (bottom row) for 2 PF jets (left column), 3 PF jets (middle column) and 4 PF jets (right column). All samples are normalized to unit area.

wide variety of variable choices and jet multiplicities. Including more jet bins was found to increase the statistical sensitivity of the measurement but also to increase the systematic uncertainty arising from the jet energy scale corrections. Correspondingly, this analysis uses a fit to the muon p_{T} and η distributions for the 3 and ≥ 4 jet bins. Because these variables are fit "simultaneously", each distribution has to be scaled by the same amount, rather than allowing the scale factor to vary independently between the different distributions. Thus, the jet multiplicity spectrum is used implicitly in the fit, since the predicted ratio between the number of events in the 3 and ≥ 4 jet bins is required to remain fixed.

The parameters of the fit are taken to be the number of events, or, equivalently, the cross section of each process (since the luminosity and selection efficiencies are determined experimentally, as discussed in the last chapter). Note that LM0 is provided for comparison and is not included as a fit parameter:

- $\sigma_{t\bar{t}}$, the cross section for $t\bar{t}$
- σ_W, the cross section for $W \to l\nu$
- σ_Z, the cross section for $Z/\gamma^* \to l^+l^-$
- σ_t, the cross section for single-top
- σ_{QCD}, the cross section for QCD

The processes $Z/\gamma^* \to l^+l^-$ and single-top are not well constrained by this fit procedure. However, their contributions, particularly in the high-jet multiplicity bins, are expected to be very small. Therefore, we impose two additional requirements on the fit. The cross sections for $W \to l\nu$ and $Z/\gamma^* \to l^+l^-$ are allowed to float freely when determining the best fit, but the ratio σ_W/σ_Z, which is experimentally well-known, is constrained to lie

within 30% of the theory value via a Gaussian fit constraint. The ratio of the cross sections is weighted by a Gaussian with a mean at the theory value and a standard deviation of 30% of the theory value, such that ratios which deviate more from theory are progressively less likely. The single-top cross section is also required to be within 30% of the theory value via a Guassian fit constraint. These two additional requirements fix the contributions from all five contributing processes.

7.3 Estimation of QCD-Multijet Background

QCD differs from the other processes in that it contains significant contributions from events with muons produced in jets that are misidentified as isolated muons. Additionally, our Monte Carlo simulation of QCD is produced using leading-order PYTHIA, which does not incorporate contributions from higher-parton final states, as MADGRAPH does. Because of these differences, it is necessary to verify that the shape of our QCD simulation accurately reflects the shape of the true contribution from multijet events.

To verify the shape of the QCD, we compare two different templates:

- Data with the standard selection requirements, but with muon relative isolation required to be in the range $0.2 < I_{\text{rel}} < 0.75$. As indicated in Figure 6.1(c), this sample is expected to be dominated by contributions from QCD. This is called *non-isolated data*, referring to the isolation of the muon.

- QCD simulation with the standard selection requirements, but with muon isolation required to be in the range $0.2 < I_{\text{rel}} < 0.75$. This is referred to as *non-isolated QCD*.

Figures 7.8 and 7.9 show the comparison of these two distributions for jet multiplicity and the muon variables used in the fit. The excellent agreement between non-isolated

QCD and non-isolated data (which is expected to be very pure in QCD) demonstrates that the QCD simulation does indeed match the true contribution from multijets in this control region. Based on this, we use the nominal MC simulation to define the QCD template for the fit.

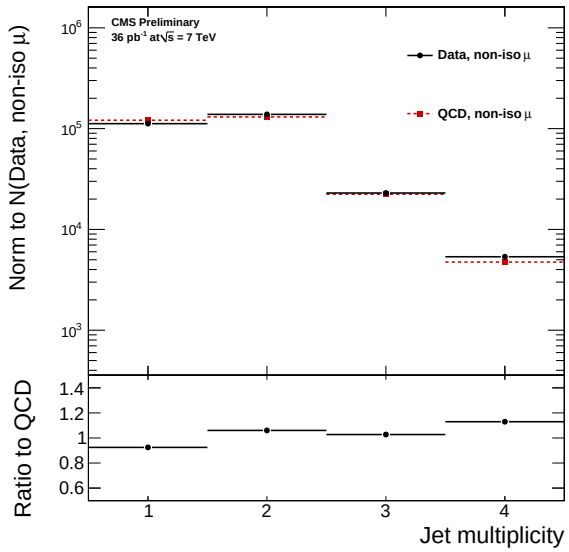


Figure 7.8: A comparison of the jet multiplicity distribution between QCD Monte Carlo with non-isolated muons (*i.e.* muons with $0.2 < I_{\text{rel}} < 0.75$) and the control region of data with non-isolated muons. The bottom of the plot shows the ratio of the data sample with non-isolated muons to the QCD sample. The distributions are normalized to the number of events in the non-isolated muon data sample.

7.4 Estimation of Vector Boson Backgrounds

As this analysis relies on the MC prediction for the shape of the vector boson backgrounds, $W \to l\nu$ and $Z/\gamma^* \to l^+l^-$, it is necessary to verify that the simulation of these processes agrees with their shape in the data, to within the assigned systematic uncertain-

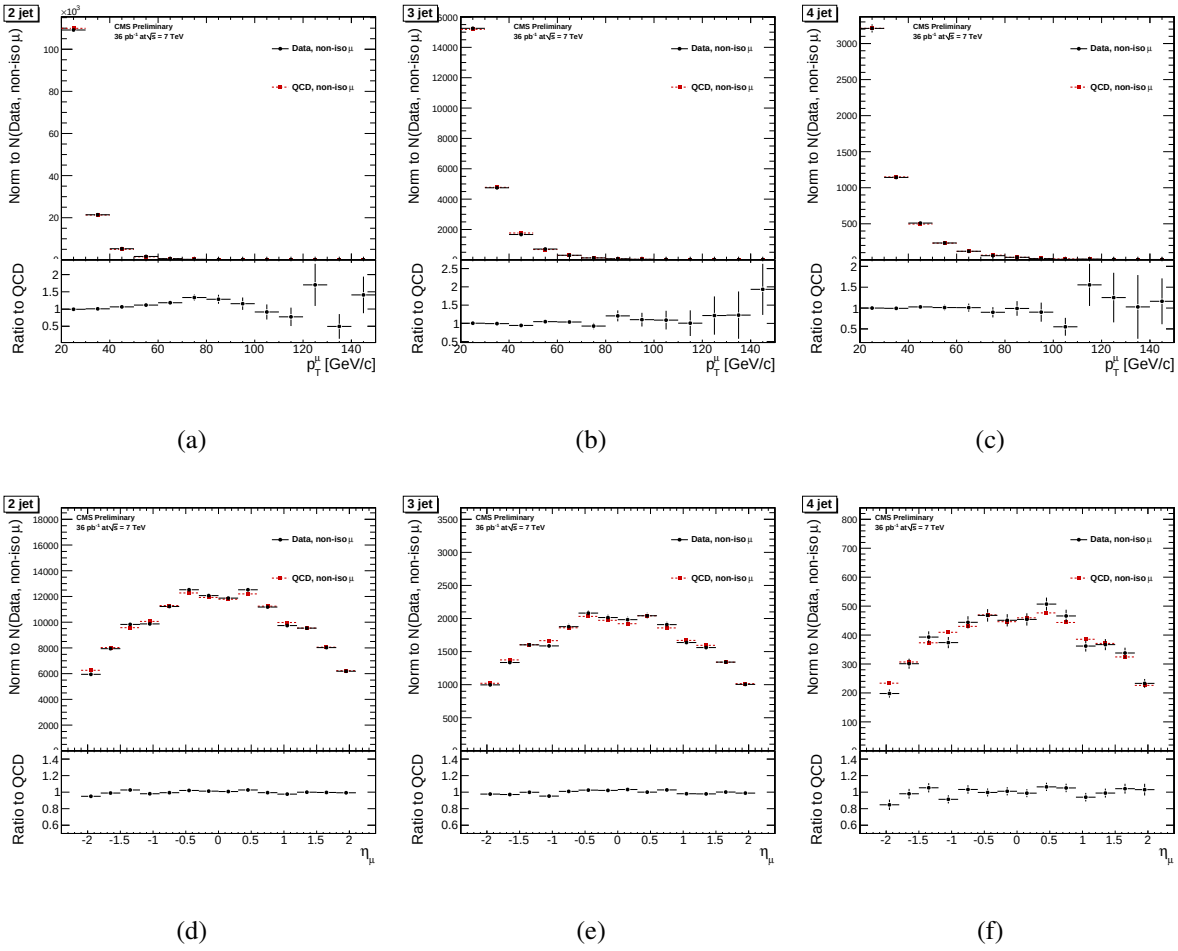


Figure 7.9: A comparison of the muon p_{T} distribution (top row) and the η distribution (bottom row) for 2 PF jets (left column), 3 PF jets (middle column) and 4 PF jets (right column). Comparison is between QCD Monte Carlo with non-isolated muons and data with non-isolated muons. The bottom of each plot shows the ratio of data to the QCD sample with non-isolated muons. The distributions are normalized to the number of events in the non-isolated muon data sample.

ties. In each case, we select a subset of our signal region that is enhanced in W or Z events and depleted in $t\bar{t}$ and QCD. This is compared to the MC prediction to assess how well it is described.

The $W \rightarrow l\nu$ enhanced region is selected using the same selection criteria as the signal region defined in Chapter 6, with the following additional requirements:

- $M_T(W) < 55\,\text{GeV}$, to reduce QCD contamination, as shown in Figure 7.2
- $H_{\text{T}} < 200\,\text{GeV}$, to reduce $t\bar{t}$ contamination, as shown in Figure 7.4

With these additional criteria applied to data, the sample is expected to be dominated by $W \rightarrow l\nu$, with a small contribution from $Z/\gamma^* \rightarrow l^+l^-$ and negligible contributions from all other processes.

Figures 7.10 and 7.11 show the comparison between the data and the $W \rightarrow l\nu$ MC in this W-enhanced region. Also shown for comparison are the simulations for $W \rightarrow l\nu$ in which the renormalization scale Q^2 (defined in Section 7.7.2 below) has been varied up by a factor of 2 or down by a factor of 1/2. For each jet bin the agreement between the data and the nominal W sample is good, and (to within statistical uncertainty) falls within the envelope defined by the W samples with renormalization scale Q^2 varried up and down.

The $Z/\gamma^* \rightarrow l^+l^-$ sample is selected using the standard selection criteria defined in Chapter 6, except that events are required to have exactly two muons whose combined invariant mass lies in the range $76\,\text{GeV}/c^2 < M_{ll} < 106\,\text{GeV}/c^2$. This sample is completely dominated by $Z/\gamma^* \rightarrow l^+l^-$, and has negligible contamination from all other processes.

Figures 7.12 and 7.13 show the comparison between data and the $Z/\gamma^* \rightarrow l^+l^-$ MC in this Z-enhanced region. Simulations with Q^2 varied up and down are shown for comparison. For each jet bin the agreement between the data and the nominal sample is good, and falls

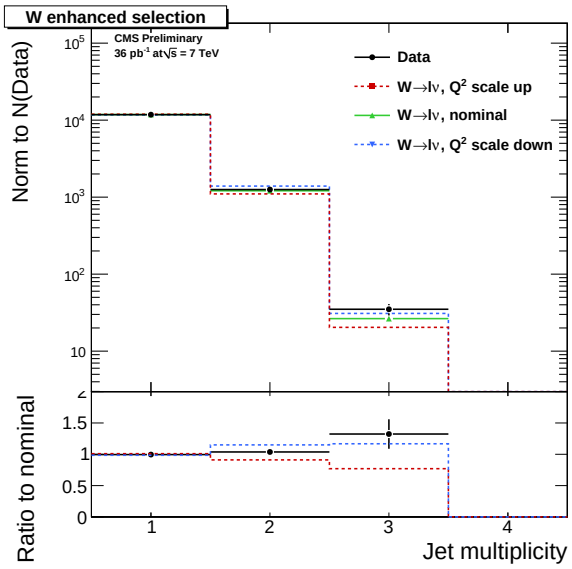
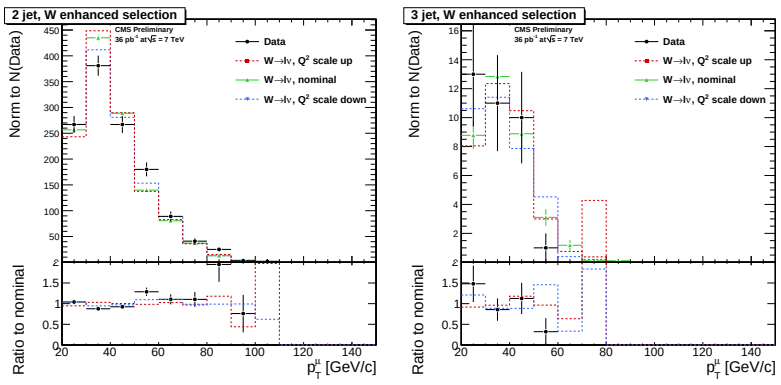


Figure 7.10: A comparison between W-enriched data and W MC for the jet multiplicity distribution. The bottom of the plot shows the ratio of each sample to the nominal W MC.



(a) (b)

Figure 7.11: A comparison between W-enriched data data and W MC for the muon p_{T} for events with (a) 2 and (b) 3 PF jets. The bottom of each plot shows the ratio of each sample to the nominal W MC.

within the envelope defined by the Z sample with Q^2 varied up and down, within statistical uncertainty.

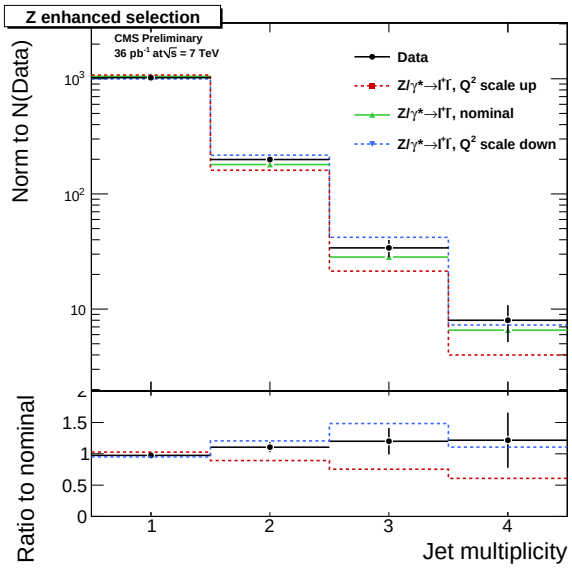
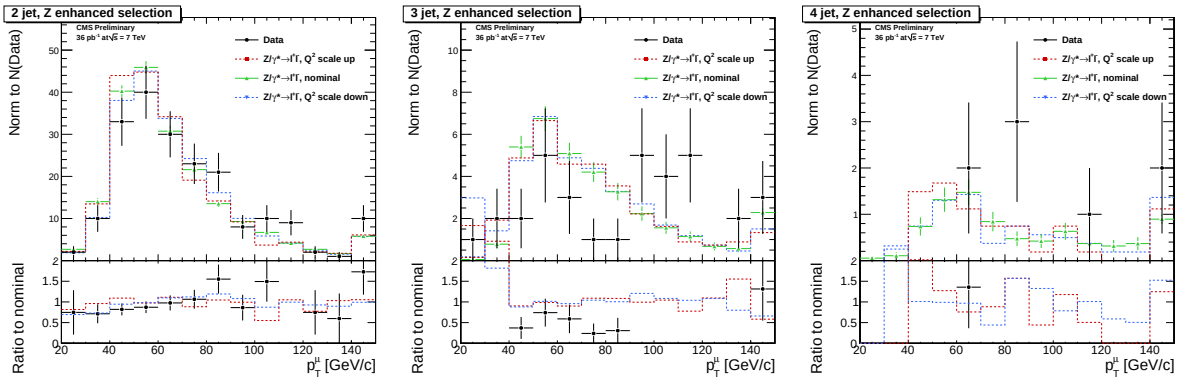


Figure 7.12: A comparison between data and MC for the jet multiplicity distribution with the Z-enhanced requirements. The bottom of the plot shows the ratio of each sample to the nominal MC.

7.5 Effect of Pile-Up

The impact of having multiple interactions in the same bunch crossing, or "pile-up" is emulated in our MC simulations, and is accounted for in both data and MC via the L1 offset corrections to the jet energy described in Section 5.4.2. We study the effect of pile-up on the analysis by comparing events in data with exactly one reconstructed primary vertex to events with two or more reconstructed primary vertices. Figure 7.14 (a) shows the distribution of vertex multiplicities in events passing all selection criteria, while Figure 7.14 (b) shows the corresponding jet multiplicity comparison for events with and without pile-up. Figure 7.15



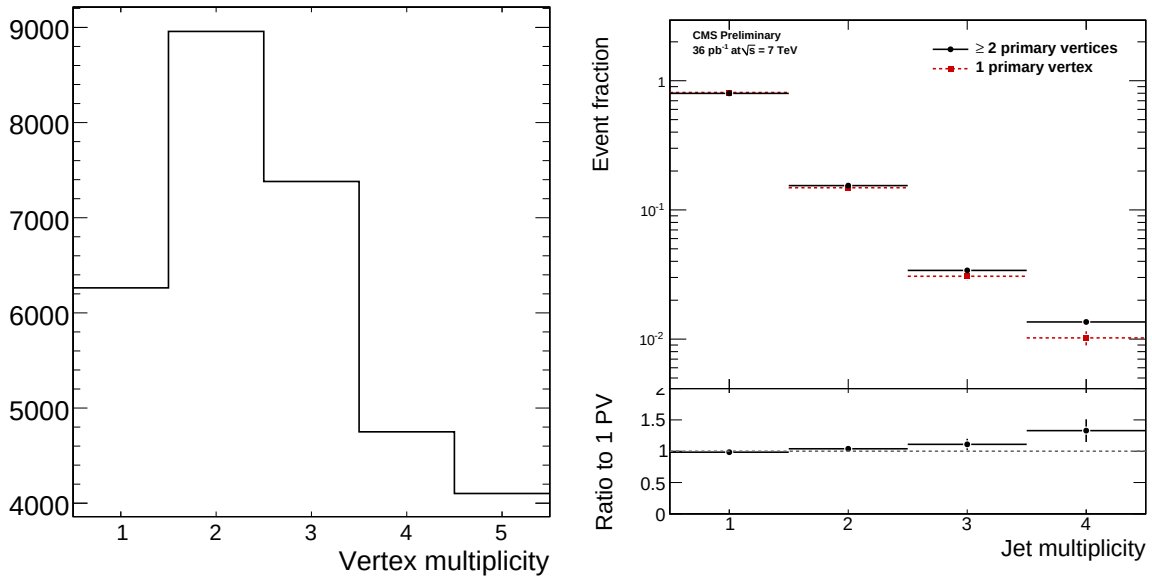
(a) (b) (c)

Figure 7.13: A comparison between data and MC for the muon p_{T} for events in the Z-enhanced region with (a) 2 and (b) 3 PF jets. The bottom of each plot shows the ratio of each sample to the nominal MC.

shows the corresponding effect of pile-up on the muon p_{T} distribution. From the agreement of the shapes, we conclude that pile-up will not have a large impact on this analysis.

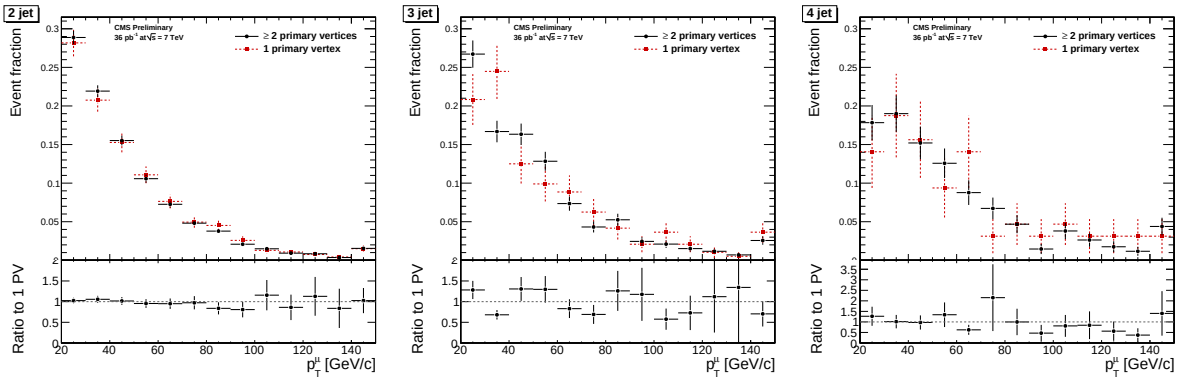
7.6 Estimation of Statistical Sensitivity

We evalulate the statistical uncertainty of the $t\bar{t}$ cross section measurement through the use of *pseudo experiments* to simulate the effect of varying the distribution being fit. First, all MC sample templates are combined, each weighted by their respective NLO cross sections (this cross section was doubledin the case of QCD to account for the large uncertainty). This combined template represents the expected distribution of data. Next, multiple sets of "pseudo data" are then generated from this template by normalizing the shape to the appropriate luminosity and fluctuating the contribution in each bin according to a Poisson distribution. Each set of pseudo data is then fitted with the same MC template shapes used



(a) (b)

Figure 7.14: (a) Distribution of the number of vertices in an events passing the standard selection requirements. (b) Comparison between events with and without pile-up for the jet-multiplicity distribution.



(a) (b) (c)

Figure 7.15: Comparison between events with and without pile-up for the muon p_{T} distribution for (a) 2, (b) 3 and (c) ≥ 4 PF jets.

to fit the actual data, and new values for the fit parameters are found. Each fit to pseudo data represents a pseudo experiment.

The statistical uncertainty was evaluated based on the distribution of fits in 3000 pseudo experiments. Figure 7.16 shows the resulting fit pull distribution, defined as $(\sigma_{\text{fit}} - \sigma_{\text{gen}})/(\text{fit error})$, fit to a Gaussian. On average, the cross section from the fit should equal the generated cross section of the pseudo data, and the standard deviation should match the fit error. Thus, the fact that the pull distribution has a mean of 0 and a standard deviation of 1 indicates that the fit is unbiased and the fit uncertainty is correctly estimated. Also shown in Figure 7.16 is the distribution $(\sigma_{\text{fit}} - \sigma_{\text{gen}})/\sigma_{\text{gen}}$, fit to a Gaussian. The numerator is equal to the pull numerator, so again we expect a mean of 0, but the width now indicates the statistical precision of the fit, as a fraction of the fit parameter. From this, the statistical uncertainty associated with the $t\bar{t}$ cross section measurement is determined to be $7.63 \pm 0.11\%$.

7.7 Estimation of Systematic Uncertainties

Systematic uncertainties are also evaluated by means of pseudo experiments. The procedure is identical to the one described above, but the combined template is instead generated from templates that have been shifted to account for the effect of the systematic. This is then fluctuated according to a Poisson distribution, and the resulting sets of pseudo data are fit to the nominal MC templates to extract the cross section. The systematic uncertainty is the mean difference between the generated and fitted cross sections, and is therefore represented (as a fraction of the cross section) by the mean of the fit to the $(\sigma_{\text{fit}} - \sigma_{\text{gen}})/\sigma_{\text{gen}}$ distribution.

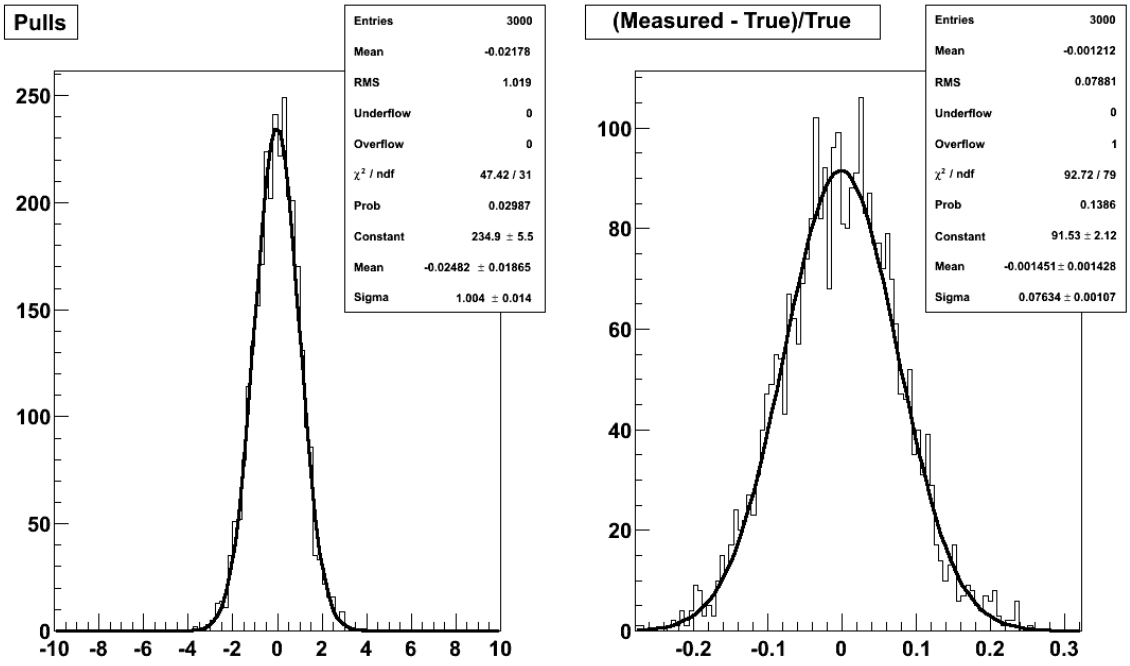


Figure 7.16: On the left is the distribution of $(\sigma_{\text{fit}} - \sigma_{\text{gen}})/(\text{fit error})$ or *pull* for 3000 pseudo experiments, fit to a Guassian. On the right is the distribution of $(\sigma_{\text{fit}} - \sigma_{\text{gen}})/\sigma_{\text{gen}}$, fit to a Gaussian. The width of this Gaussian indicates the statistical precision of the fit.

JES systematic yield change							
Cut	Sys	$t\bar{t}$	Single-top (s)	Single-top (t)	Single-top (tW)	$W \to l\nu$	$Z/\gamma^* \to l^+l^-$
2 jets	JES +	-6.6%	+1.4%	+1.5%	-2.6%	+7.7%	+8.1%
	JES -	+7.4%	-1.5%	-1.7%	+3%	-7.2%	-7%
3 jets	JES +	-1.8%	+5.6%	+5%	+2.9%	+10%	+9.1%
	JES -	+1.1%	-6.1%	-5.4%	-3.7%	-9.6%	-11%
≥ 4 jets	JES +	+6.5%	+13%	+9.7%	+11%	+15%	+15%
	JES -	-6.5%	-11%	-8.5%	-10%	-12%	-12%

Table 7.1: Relative yield variation due to JES shift up / down.

7.7.1 Jet Energy Scale Uncertainty

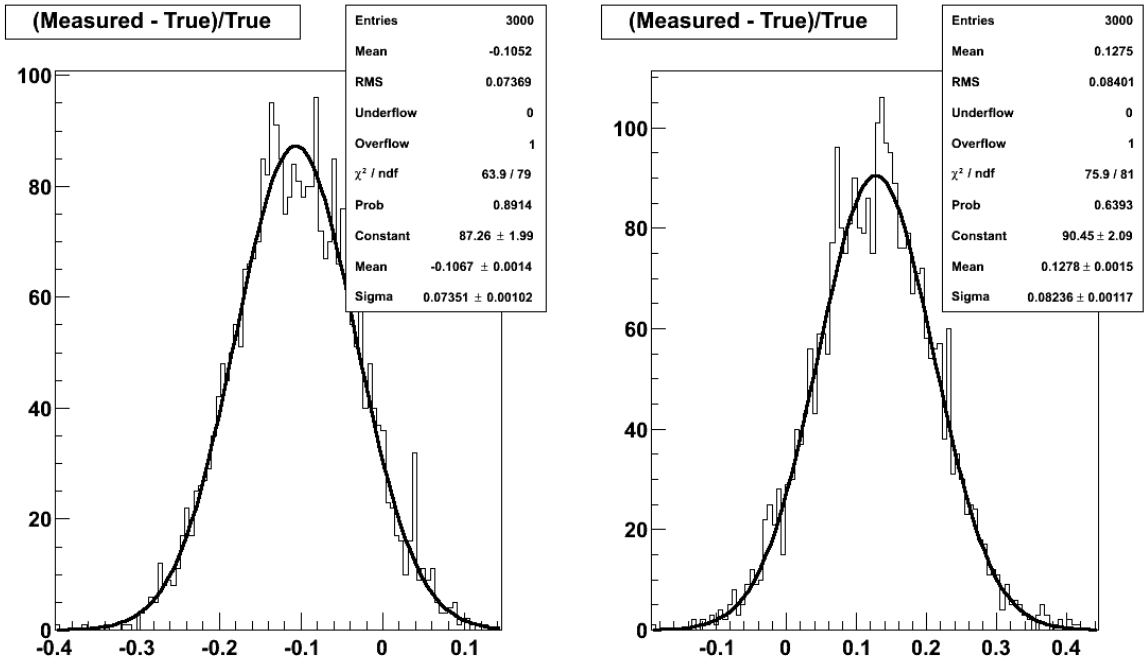
Several independent uncertainties are added in quadrature to obtain the 1σ uncertainty in the jet energy scale (JES). An official list of uncertainties, which depend on the p_{T} and η of the jet, is provided [67]. To these we add in quadrature:

- $C_{SW} = 0.015$ for calibration changes and software release differences
- $C_{PU} = E_{PU} \cdot A_{\text{jet}} \cdot \text{Avg}_{PU}/p_{\mathrm{T}}^{\text{jet}}$ for the effects of pile-up, where $E_{PU} = 0.2\,\text{GeV}$ is the pile-up energy, $A_{\text{jet}} = 0.8$ is the jet area, $\text{Avg}_{PU} = 2.2$ is the average number of pile-up events and $p_{\mathrm{T}}^{\text{jet}}$ is the fully corrected jet p_{T}
- $C_b = 0.01$ for all jets as a conservative upper-bound on flavor-dependent uncertainties for b-jets (since b-jets are not specifically tagged in this analysis)

These values, added in quadrature to the p_{T}– and η-dependent uncertainties (typically of order 2%), give the total 1σ uncertainty. The variations in the event yields corresponding to this shift are shown in Table 7.1.

As described above, the systematic uncertainty quoted for the JES is evaluated by means of pseudo experiments. Figure 7.17 shows the results of these pseudo experiments

prepared with templates for which the jet energy scale has been shifted up or down by 1σ. The mean of the corresponding Gaussian fitted to the $(\sigma_{\text{fit}} - \sigma_{\text{gen}})/\sigma_{\text{gen}}$ distribution gives the systematic uncertainty in the jet energy scale, found to be $+12.8 \pm 0.15\%$ for the upward shift and $-10.67 \pm 0.14\%$ for the downward shift.



(a) (b)

Figure 7.17: Distributions of $(\sigma_{\text{fit}} - \sigma_{\text{gen}})/\sigma_{\text{gen}}$ for 3000 pseudo experiments, fit to Gaussians. The pseudo data are constructed from templates corresponding to shifts of (a) -1σ and (b) $+1\sigma$ in the jet energy scale. This pseudo data is then fit using the nominal MC templates, such that the mean of the Gaussian indicates the systematic uncertainty in the jet energy scale.

Shifted templates are prepared for each distribution used in the fit, for each MC sample. Figures 7.18 and 7.19 show the effect of shifting the JES on the $t\bar{t}$ sample, while

Figures 7.20 and 7.21 show the effect for the $W \to l\nu$ sample.

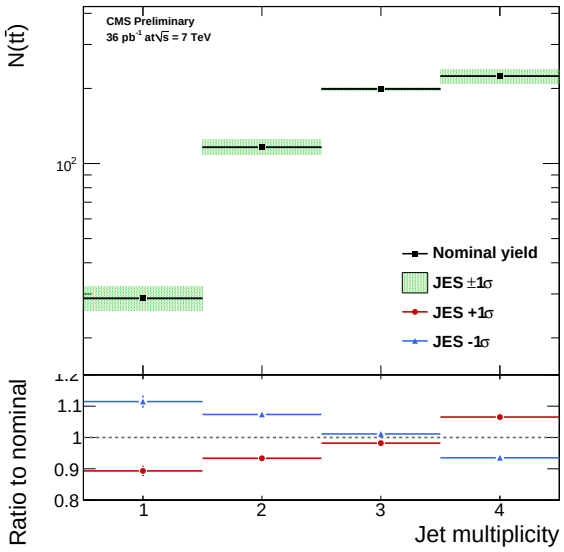
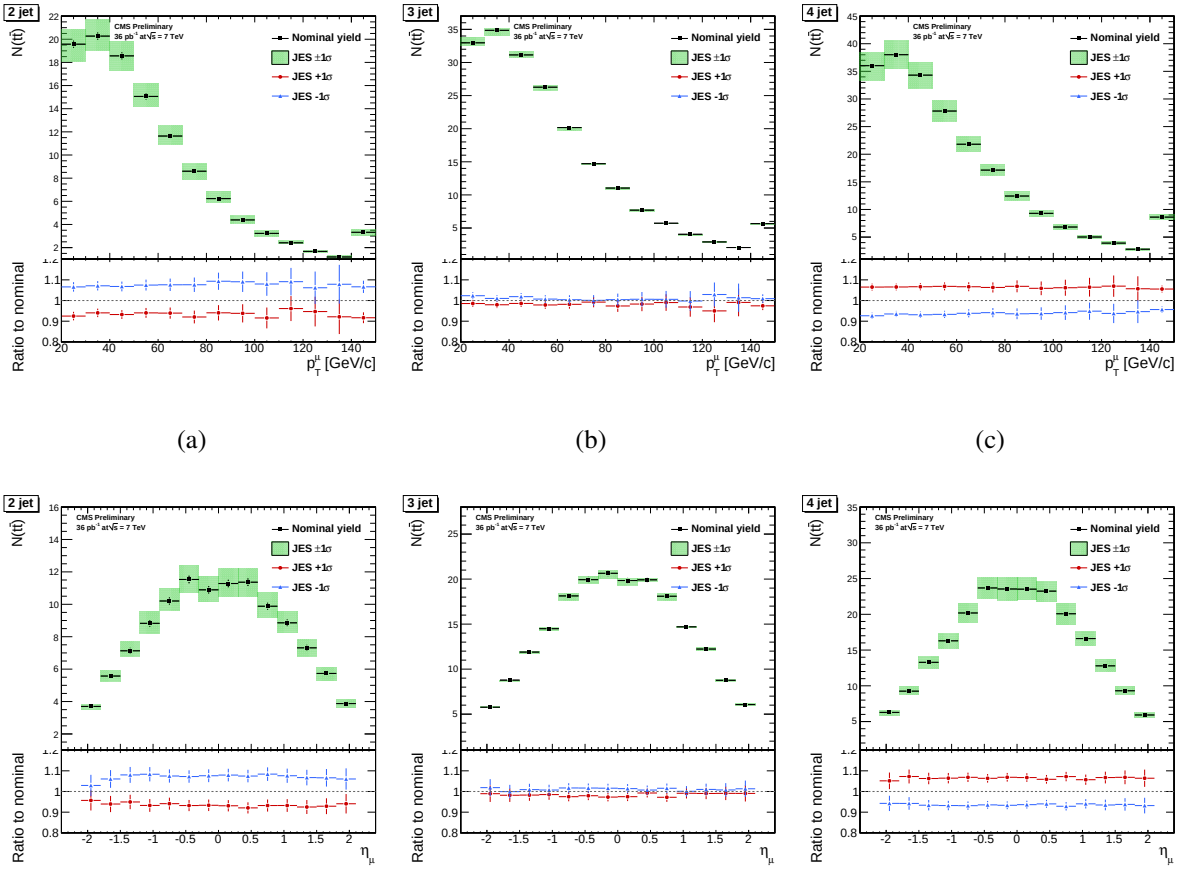


Figure 7.18: Jet multiplicity distribution for $t\bar{t}$ showing the uncertainty due to a 1σ shift in the JES. In the top plot, the green shaded area represents the 1σ spread. The bottom plot gives the ratio of a $+1\sigma$ JES shift (in red) and a -1σ shift (in blue) to the nominal, unshifted $t\bar{t}$ sample.



(a) (b) (c)

(d) (e) (f)

Figure 7.19: Muon p_{T} (top row) and muon η distributions (bottom row) of the $t\bar{t}$ sample for the 2, 3, and ≥ 4 jet bins showing the uncertainty due to a 1σ shift in the JES. In the top plots, the green shaded area represents the 1σ spread. The bottom plots give the ratio of a $+1\sigma$ JES shift (in red) and a -1σ shift (in blue) to the nominal, unshifted $t\bar{t}$ sample.

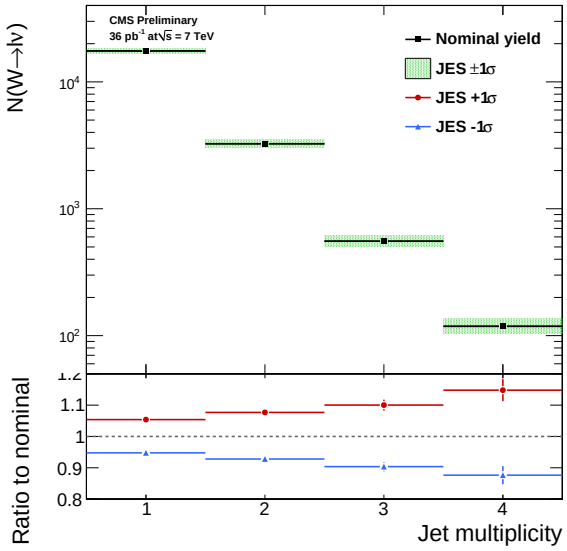
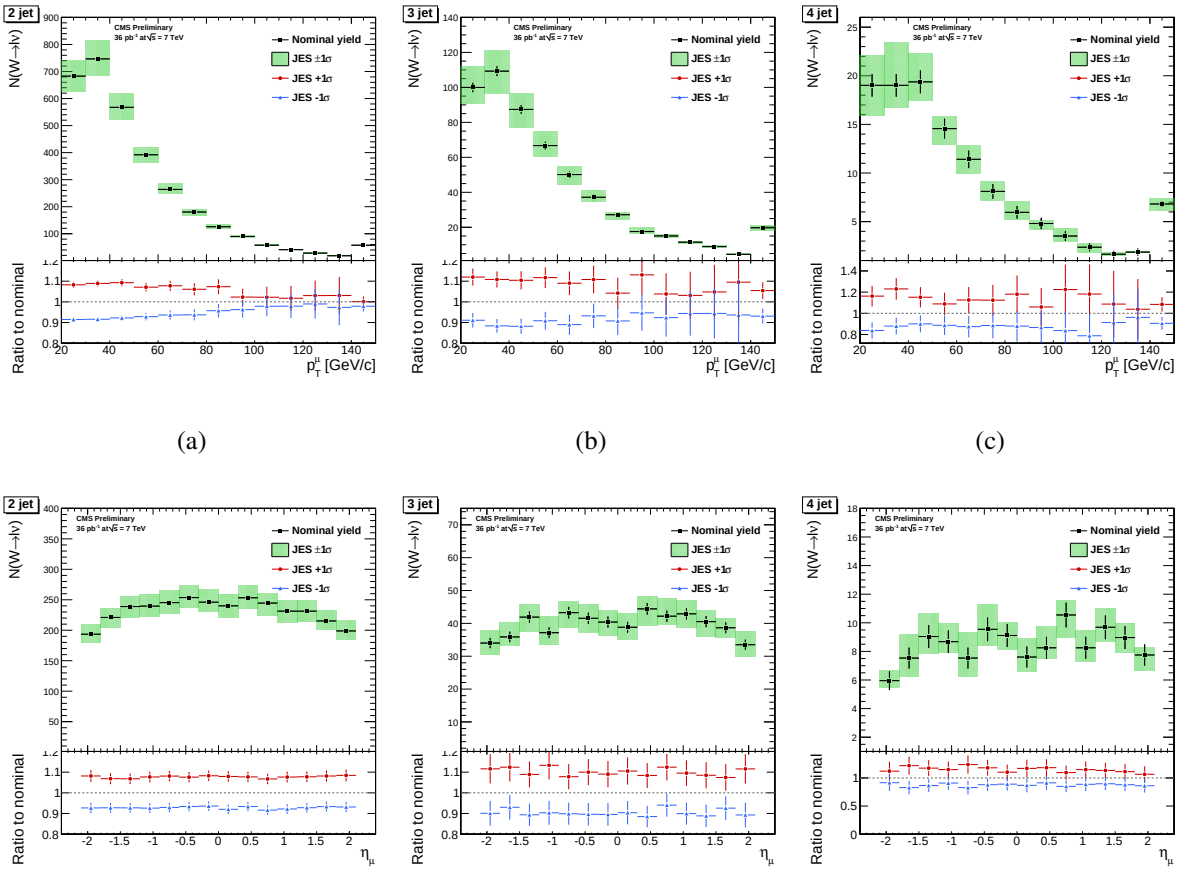


Figure 7.20: Jet multiplicity distribution for $W \to l\nu$ showing the uncertainty due to a 1σ shift in the JES. In the top plot, the green shaded area represents the 1σ spread. The bottom plot gives the ratio of a $+1\sigma$ JES shift (in red) and a -1σ shift (in blue) to the nominal, unshifted $W \to l\nu$ sample.



(a) (b) (c)

(d) (e) (f)

Figure 7.21: Muon p_{T} (top row) and muon η distributions (bottom row) of the $W \rightarrow l\nu$ sample for the 2, 3, and ≥ 4 jet bins showing the uncertainty due to a 1σ shift in the JES. In the top plots, the green shaded area represents the 1σ spread. The bottom plots give the ratio of a $+1\sigma$ JES shift (in red) and a -1σ shift (in blue) to the nominal, unshifted $W \rightarrow l\nu$ sample.

7.7.2 Factorization Scale Uncertainty

The value of the strong coupling constant α_s varies as a function of the momentum scale used in the simulation. In general, there are two scales that need to be considered: the factorization scale is used to evaluate α_s for the parton distribution function (PDF), while the renormalization scale is used to evaluate α_s at the various vertices in the Feynman diagram for the event being simulated. MADGRAPH chooses these scales to be identical for the hard scatter and denotes them as Q^2, where

$$Q^2 = m_t^2 + \sum p_{\mathrm{T}}^2(\text{jet}) \qquad \text{for } t\bar{t} \tag{7.2}$$

$$Q^2 = m_{W/Z}^2 + \sum p_{\mathrm{T}}^2(\text{jet}) \qquad \text{for } W \to l\nu \text{ or } Z/\gamma^* \to l^+l^- \tag{7.3}$$

The impact of uncertainty in this scale is estimated by varying Q^2 in each sample by factors of $1/2$ and 2 with respect to their default values. Because the ratio of the W and Z cross sections is constrained, these two processes are treated as being correlated, and their respective factorization scales are shifted either up or down simultaneously. The $t\bar{t}$ sample is considered to be uncorrelated, and so its scale is shifted independently and the corresponding systematic uncertainty is added in quadrature.

The variations in the event yields corresponding to the Q^2 shift are shown in Table 7.2. We see that varying this scale has a large impact on the number of $W \to l\nu$ events.

Shifted templates are prepared for each distribution used in the fit, for each MC sample. Figures 7.22 and 7.23 show the effect of shifting the Q^2 scale on the $t\bar{t}$ sample, while Figures 7.24 and 7.25 show the effect for the $W \to l\nu$ sample. Performing pseudo experiments as described above yields an uncertainty of $-6.95 \pm 0.15\%$ for the shift up

Q^2 scale systematic yield change				
Cut	Sys	$t\bar{t}$	$W \to l\nu$	$Z/\gamma^* \to l^+l^-$
2 jets	Scale +	+5.5%	-22%	-17%
	Scale -	-6.1%	+66%	+49%
3 jets	Scale +	+4.1%	-35%	-32%
	Scale -	-3%	+100%	+67%
$\geq$ 4 jets	Scale +	-5.1%	-44%	-41%
	Scale -	+9%	+96%	+74%

Table 7.2: Relative yield variation due to Q^2 scale shift up / down.

and $+11.65 \pm 0.15\%$ for the shift down of $t\bar{t}$. Correspondingly, we find an uncertainty of $-4.78 \pm 0.14\%$ for the shift up and $+4.82 \pm 0.17\%$ for the shift down of $W \to l\nu$.

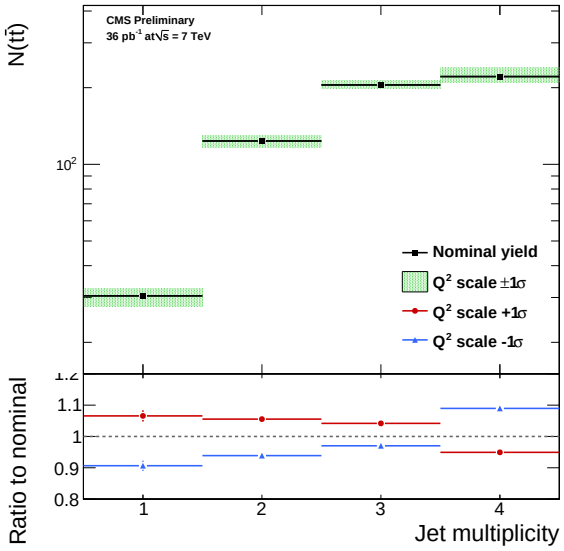
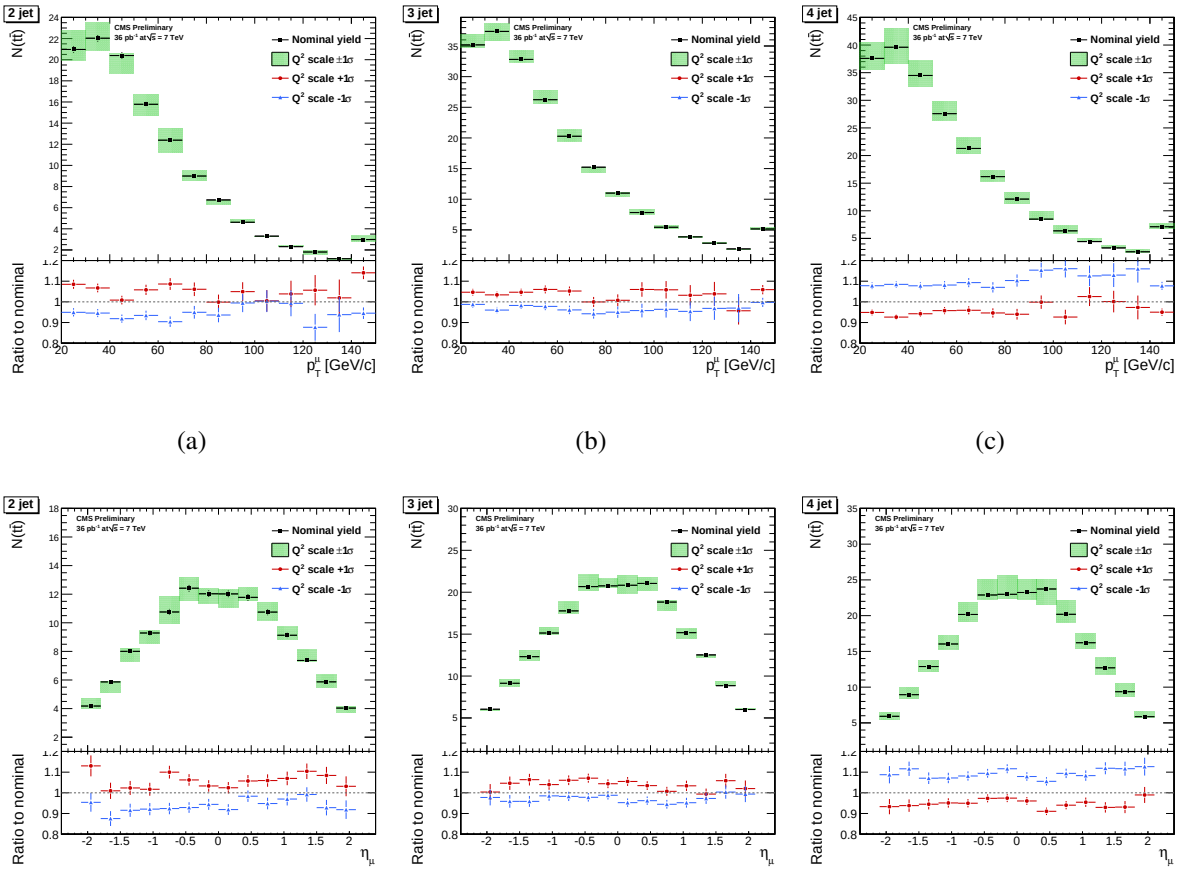


Figure 7.22: Jet multiplicity distribution for $t\bar{t}$ showing the uncertainty due to a factor of 2 shift in the Q^2 scale. In the top plot, the green shaded area represents the factor of 2 spread. The bottom plot gives the ratio of a Q^2 scale shift up (in red) and a shift down (in blue) to the nominal, unshifted $t\bar{t}$ sample.



(a) (b) (c)

(d) (e) (f)

Figure 7.23: Muon p_{T} (top row) and muon η distributions (bottom row) of the $t\bar{t}$ sample for the 2, 3, and ≥ 4 jet bins showing the uncertainty due to a factor of 2 shift in the Q^2 scale. In the top plots, the green shaded area represents the factor of 2 spread. The bottom plots give the ratio of a Q^2 scale shift up (in red) and a shift down (in blue) to the nominal, unshifted $t\bar{t}$ sample.

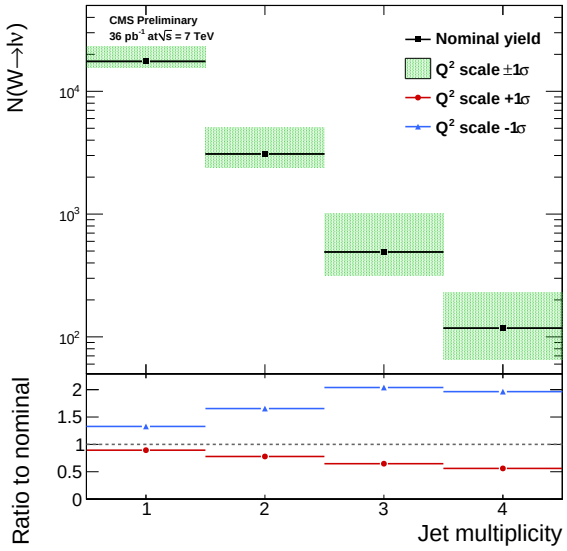


Figure 7.24: Jet multiplicity distribution for $W \rightarrow l\nu$ showing the uncertainty due to a factor of 2 shift in the Q^2 scale. In the top plot, the green shaded area represents the factor of 2 spread. The bottom plot gives the ratio of a Q^2 scale shift up (in red) and a shift down (in blue) to the nominal, unshifted $W \rightarrow l\nu$ sample.

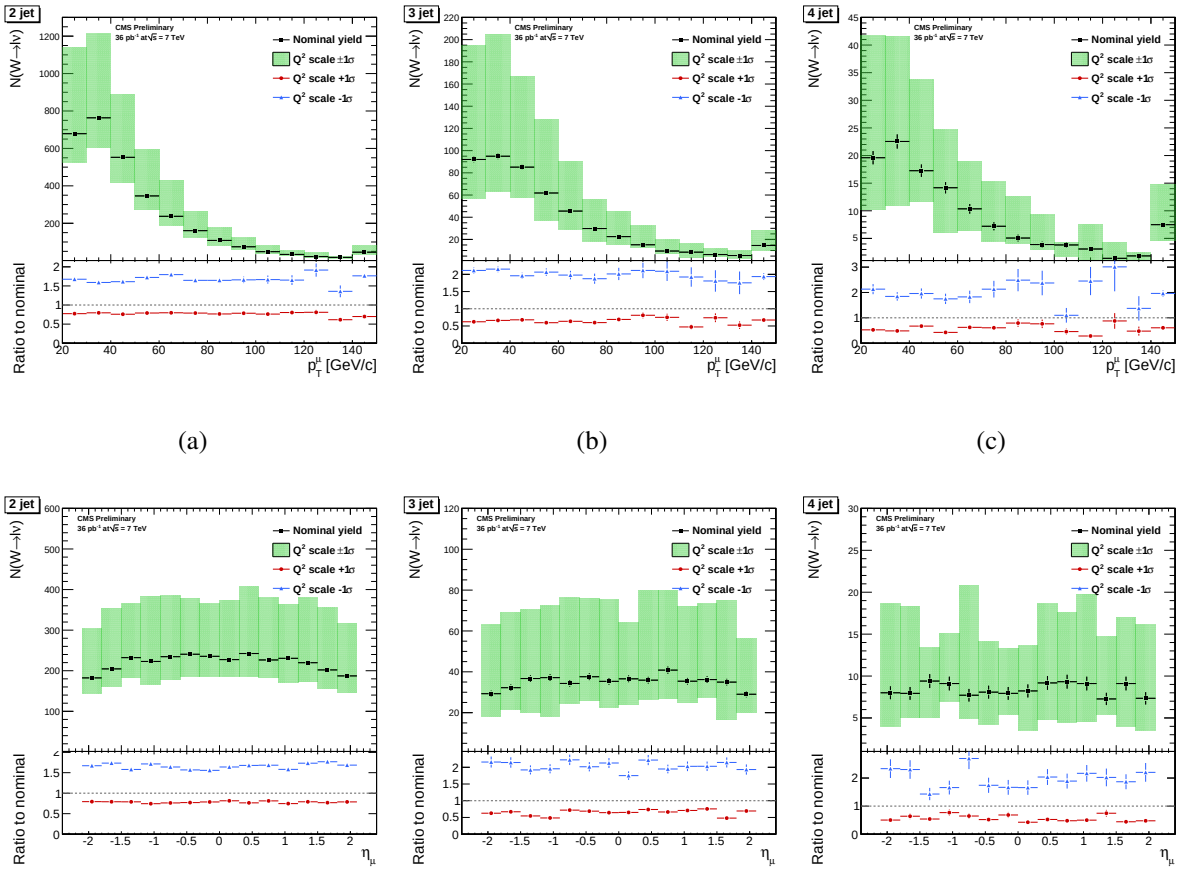


Figure 7.25: Muon p_{T} (top row) and muon η distributions (bottom row) of the $W \rightarrow l\nu$ sample for the 2, 3, and ≥ 4 jet bins showing the uncertainty due to a factor of 2 shift in the Q^2 scale. In the top plots, the green shaded area represents the factor of 2 spread. The bottom plots give the ratio of a Q^2 scale shift up (in red) and a shift down (in blue) to the nominal, unshifted $W \rightarrow l\nu$ sample.

Matching threshold systematic yield change				
Cut	Sys	$t\bar{t}$	$W \to l\nu$	$Z/\gamma^* \to l^+l^-$
2 jets	Match +	+1%	+4.5%	+0.59%
	Match -	-2.9%	-6.2%	-4.9%
3 jets	Match +	+1.6%	+5.4%	-4.6%
	Match -	-0.49%	+4.4%	-7.7%
$\geq$ 4 jets	Match +	-1.4%	-8.6%	-14%
	Match -	+1.6%	+1.3%	-5.4%

Table 7.3: Relative yield variation due to matching threshold shift up / down.

7.7.3 Parton Shower Matching Uncertainty

At the showering stage in the simulation, it is necessary to match individual matrix elements with corresponding parton showers. The effect of uncertainty in this matching threshold is evaluated by varying this threshold by factors of $1/2$ and 2 with respect to the nominal threshold. Analogously to the Q^2 scale variation, we vary the $t\bar{t}$ sample independently from the W and Z samples, and the uncertainties for both are then added in quadrature. The variations in the event yields corresponding to the matching threshold shift are shown in Table 7.2.

Shifted templates are prepared for each distribution used in the fit, for each MC sample. Figures 7.26 and 7.27 show the effect of shifting the matching threshold on the $t\bar{t}$ sample, while Figures 7.28 and 7.29 show the effect for the $W \to l\nu$ sample. Performing pseudo experiments as described above yields an uncertainty of $-2.01 \pm 0.15\%$ for the shift up and $+2.09 \pm 0.15\%$ for the shift down of $t\bar{t}$. Correspondingly, we find an uncertainty of $-8.25 \pm 0.15\%$ for the shift up and $+5.65 \pm 0.15\%$ for the shift down of $W \to l\nu$.

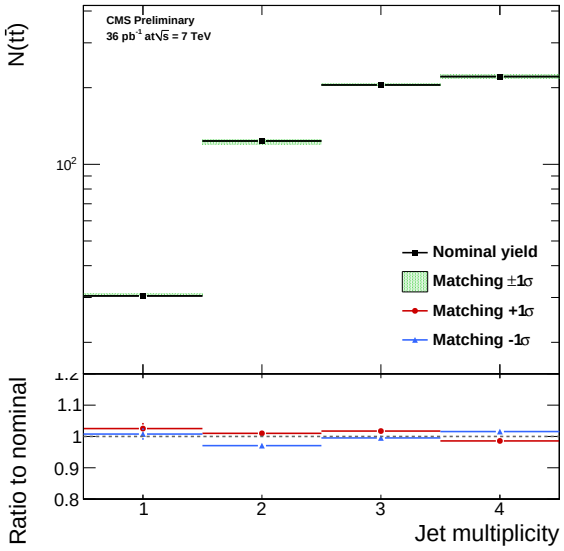
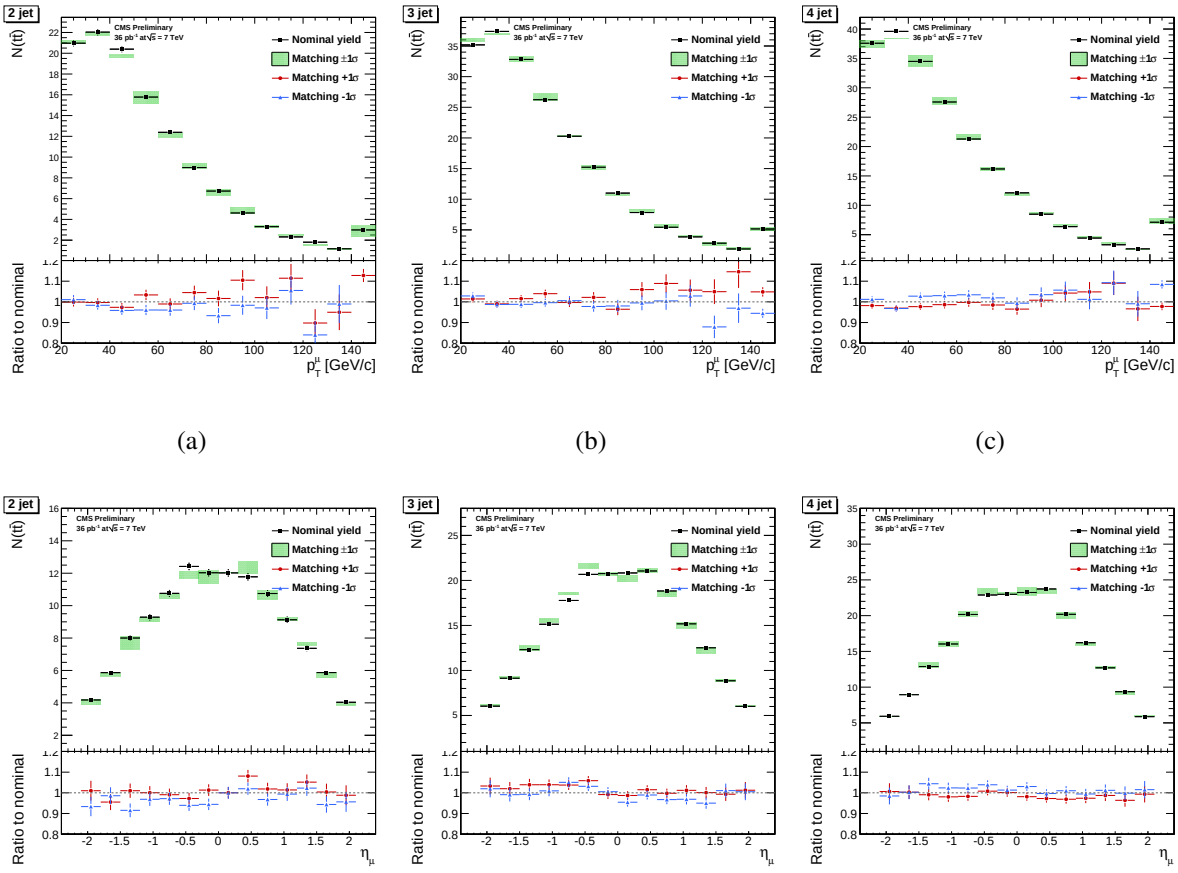


Figure 7.26: Jet multiplicity distribution for $t\bar{t}$ showing the uncertainty due to a factor of 2 shift in the matching threshold. In the top plot, the green shaded area represents the factor of 2 spread. The bottom plot gives the ratio of a matching shift up (in red) and a shift down (in blue) to the nominal, unshifted $t\bar{t}$ sample.



(a) (b) (c)

(d) (e) (f)

Figure 7.27: Muon p_{T} (top row) and muon η distributions (bottom row) of the $t\bar{t}$ sample for the 2, 3, and ≥ 4 jet bins showing the uncertainty due to a factor of 2 shift in the matching threshold. In the top plots, the green shaded area represents the factor of 2 spread. The bottom plots give the ratio of a matching shift up (in red) and a shift down (in blue) to the nominal, unshifted $t\bar{t}$ sample.

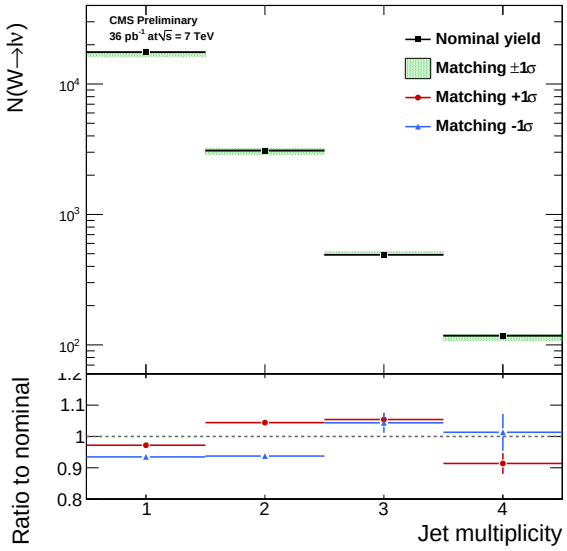


Figure 7.28: Jet multiplicity distribution for $W \rightarrow l\nu$ showing the uncertainty due to a factor of 2 shift in the matching threshold. In the top plot, the green shaded area represents the factor of 2 spread. The bottom plot gives the ratio of a matching shift up (in red) and a shift down (in blue) to the nominal, unshifted $W \rightarrow l\nu$ sample.

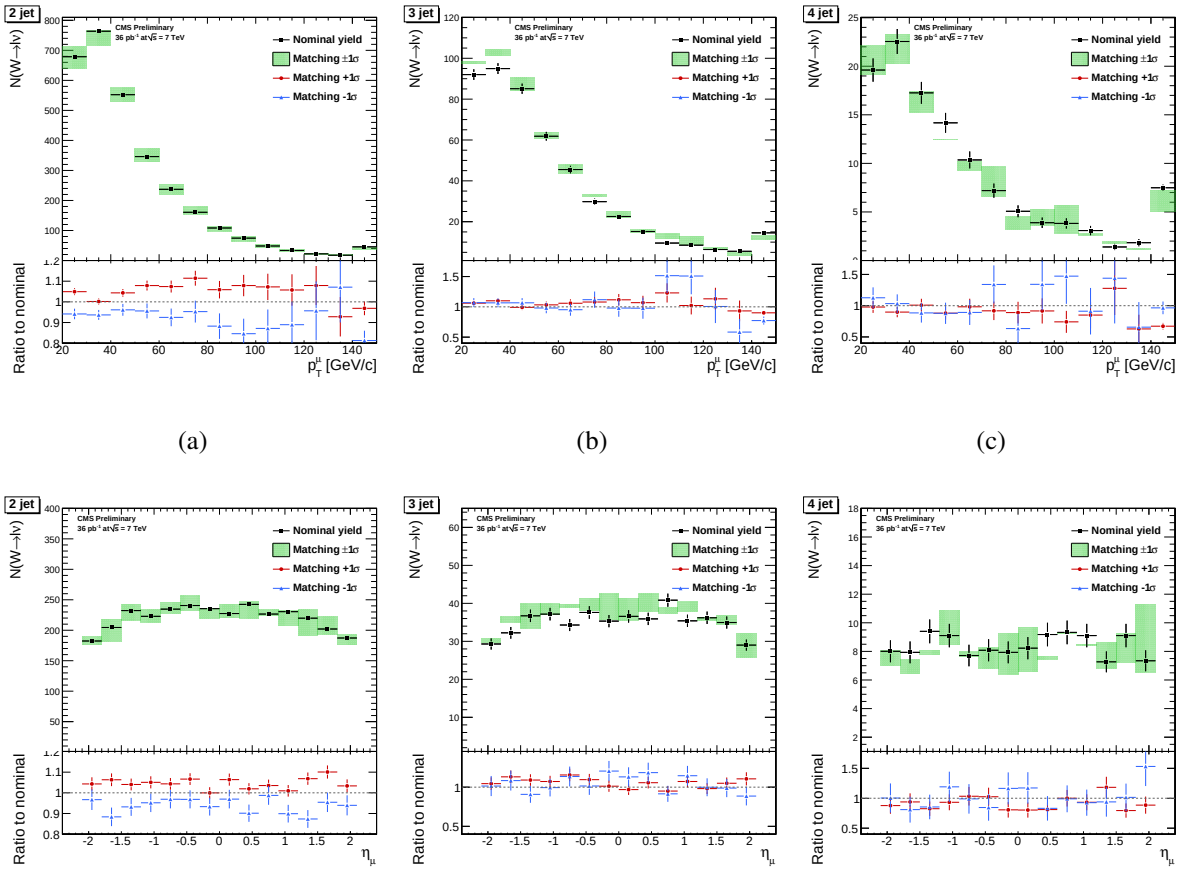


Figure 7.29: Muon p_{T} (top row) and muon η distributions (bottom row) of the $W \to l\nu$ sample for the 2, 3, and ≥ 4 jet bins showing the uncertainty due to a factor of 2 shift in the matching threshold. In the top plots, the green shaded area represents the factor of 2 spread. The bottom plots give the ratio of a matching shift up (in red) and a shift down (in blue) to the nominal, unshifted $W \to l\nu$ sample.

7.8 Other Sources of Uncertainty

There are a few additional sources of theoretical uncertainty that can have an impact on the overall systematic uncertainty. In particular, these arise from imperfect theoretical modeling of the PDFs of the colliding protons and the initial– and final-state radiation. In both cases the effect is expected to be very small compared to other sources of uncertainty, so a conservative upper limit is placed on them based on previous studies.

As discussed in Chapter 4, the production cross section for any process at the LHC depends on the PDFs used to describe the colliding protons. Because these are not known with infinite precision, it is necessary to assign an uncertainty to the final measurement based on the impact of changing the PDFs. Based on the effect of this variation in comparable measurements, we assign an uncertainty of $\pm 3.0\%$ to the final result [65].

Initial– and final-state radiation (ISR/FSR) refers to the radiation of high-momentum gluons from the initial or final particles in the hard scattering process. These gluons can produce additional jets in the event, which can affect event kinematics. From comparable measurements, we expect this variation to contribute an uncertainty of $\pm 4.6\%$ to the final result [65].

7.9 Fake Rate Determination

The second part of the analysis, using events with same-sign dileptons to search for new physics signatures, relies on the presence of the additional lepton to control backgrounds. Because contributions from SM processes with real same-sign dileptons are so small, the dominant background to the search comes from events where either one or two

jets mimic the lepton signature. These can come from multiple sources, including jets reconstructed as electrons, real leptons inside a jet due to heavy-flavor decays, electrons from unidentified photon conversions or muons from meson decays in-flight. For this analysis we refer to all such leptons as "fake leptons" to distinguish them from the signal leptons coming from vector bosons, SUSY particles, etc, labeled "prompt leptons".

7.9.1 QCD-Enriched Selection for Fake Rate

Contributions from fake leptons are estimated directly from data using the *fake rate* method [68, 69]. First, a QCD-enriched sample is selected from data. Within this sample we define two types of leptons: "tight" leptons, defined as leptons passing all standard analysis requirements, and "loose" leptons, for which certain selection criteria have been loosened or removed. Because this sample is very pure in QCD, both sets are predominantly populated by fake leptons. The fake rate ϵ_{FR} is then defined, as a function of both p_{T} and η, as the probability that a loose lepton candidate also passes the tight selection requirements.

The definition of a loose electron is identical to the nominal same-sign electron selection defined in Section 6.4, with the following exceptions:

- All ID requirements ($\sigma_{i\eta i\eta}$, $\Delta\phi$, $\Delta\eta$ and H/E) are removed.
- The relative isolation requirement is loosened from $I_{\text{rel}} < 0.1$ to $I_{\text{rel}} < 0.5$.
- The selection for the impact parameter $d_0(\text{bsp})$ is removed.

The definition of a loose muon is identical to the nominal same-sign muon selection defined in Section 6.3, with the following exceptions:

- The global fit requirement is loosened from $\chi^2/n_{\text{dof}} < 10.0$ to $\chi^2/n_{\text{dof}} < 50.0$.

- The relative isolation requirement is loosened from $I_{\text{rel}} < 0.1$ to $I_{\text{rel}} < 0.75$.
- The selection for the impact parameter is loosened from $d_0(\text{bsp}) < 200\mu m$ to $d_0(\text{bsp}) < 2mm$.

It is not possible to use jet-triggered data to measure the fake rate, because low-p_{T} jet triggers are heavily prescaled in the early data. Instead, we employ lepton triggers and impose an additional offline jet requirement in place of a jet trigger. If there is only one jet passing all selection requirements in the event, then we consider only leptons well separated from it, with $\Delta R(l, \text{jet}) > 1.0$. We further require that each lepton be matched to a corresponding trigger object with the requirement $\Delta R(l, \text{HLT}) < 0.4$. This ensures that the choice of lepton trigger does not bias the lepton fake rate.

To correctly calculate the fake rate from QCD-enriched data, we must remove contamination due to prompt leptons coming from a W or Z decay. Thus, the sample for the fake rate estimation uses the following requirements:

- $\not{E}_T < 20\,\text{GeV}$ and $M_T(W) < 25\,\text{GeV}/c^2$, to remove contributions from the decay $W \to l\nu$
- $|M_{ll} - M_Z| > 20\,\text{GeV}/c^2$ for any events with two opposite-sign leptons, to remove contributions from the decay $Z/\gamma^* \to l^+l^-$

Even with these criteria, the fake rate may be biased by prompt leptons at high p_{T}. To eliminate this effect, the fake rate is determined only up to $p_{\text{T}} = 35\,\text{GeV}/c$, and a constant value is used above this.

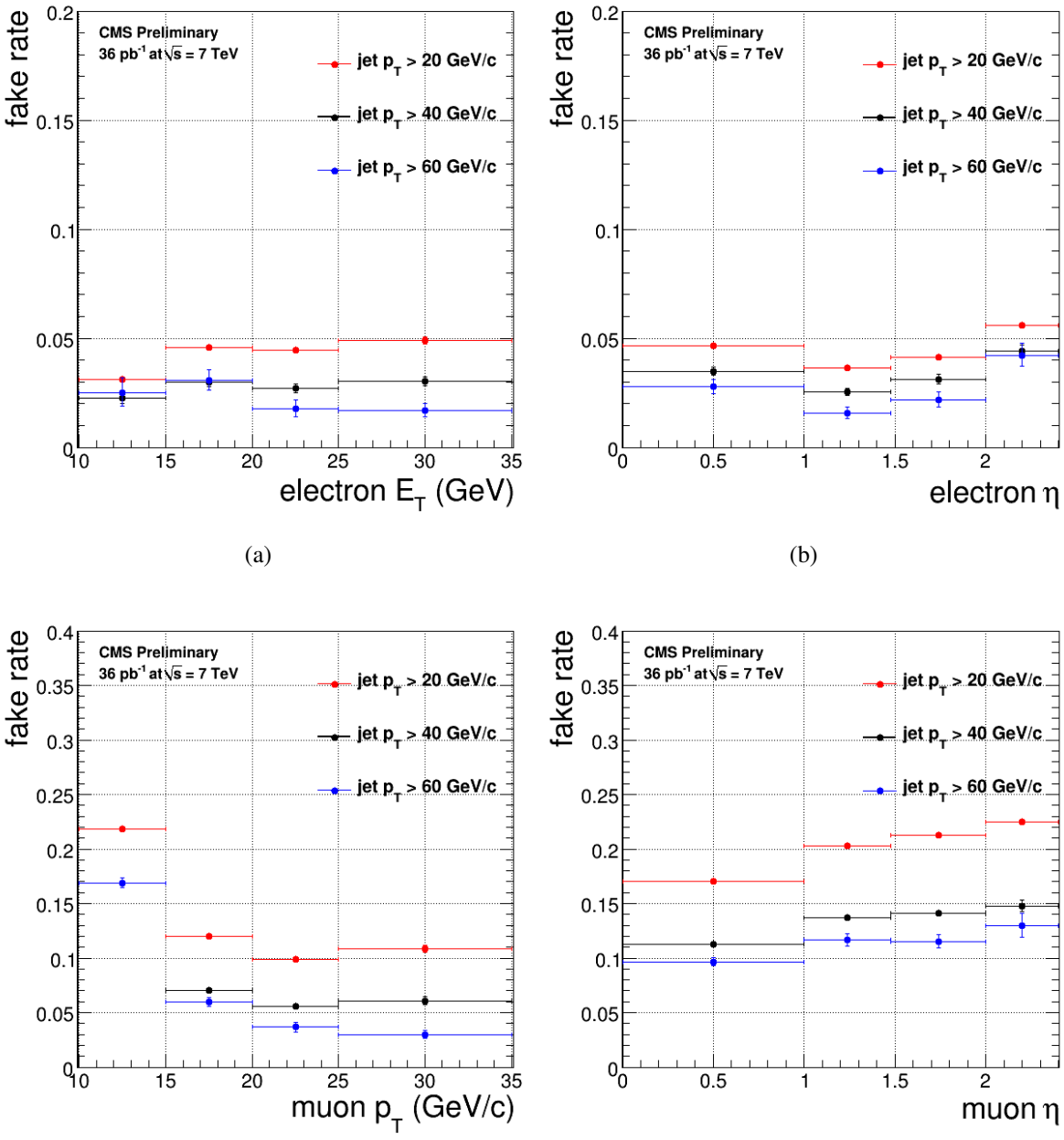
Figure 7.30 shows the fake rate (the ratio of tight to loose leptons) as a function of lepton p_{T} and η. Rates are shown for multiple jet p_{T} thresholds to illustrate the dependence

of fake rate on jet p_{T}.

7.9.2 Application of Fake Rate to Same-Sign Analysis

This fake rate is then used to estimate backgrounds with fake leptons. Events are selected using all selection requirements, with the exception of lepton selection. Backgrounds with one real and one fake lepton are estimated by requiring the event to have one tight lepton, and one lepton passing the loose selection but failing the tight selection (that is, one lepton which is loose but not tight). For backgrounds with two fake leptons, we require both leptons in the event to be loose but not tight.

For the single-fake case, each event is weighted by a factor $\epsilon_{FR}/(1-\epsilon_{FR})$, For the double-fake case, each event is weighted by two factors $\epsilon_{FR1}/(1-\epsilon_{FR1}) \cdot \epsilon_{FR2}/(1-\epsilon_{FR2})$, where ϵ_{FR1} and ϵ_{FR2} are the fake rates for each lepton. The sum of these two weights (for the single– and double-fake cases) gives the total background prediction.



(a) (b)

(c) (d)

Figure 7.30: Fake rates ϵ_{FR} for electrons (top) and muons (bottom) as a function of p_{T} (left) and η (right). Starting from the electron or muon datasets, the rates are computed from the QCD-enriched region defined in the text with different requirements on the minimum jet p_{T} threshold.

Chapter 8

Results

The "paradox" is only a conflict between reality and your feeling of what reality "ought to be."

— Richard Feynman

8.1 Measurement of Top Cross Section

The $t\bar{t}$ production cross section is extracted from data using a simultaneous template fit to the muon p_{T} and η distributions in the 3 and ≥ 4 jet bins, as discussed in Chapter 7. The sources of statistical and systematic uncertainties have been accounted for via large numbers of pseudoexperiments with shapes taken from the Monte Carlo templates.

Table 8.1 summarizes the uncertainties from the sources considered in Chapter 7. The largest contribution comes from the jet energy scale uncertainty, with other significant contributions coming from the Q^2 scale and the matching threshold uncertainties. The total systematic uncertainty, determined via pseudoexperiments, from all sources considered in this analysis is found to be $^{+21.2}_{-17.1}\%$, where the contributions from each source are added in quadrature. The statistical uncertainty, also determined from pseudoexperiments, is $\pm 7.6\%$. The behavior of the systematic and statistical uncertainties is likewise assumed to be uncor-

Summary of Uncertainties	
Source	Relative uncertainty
JES	$+12.8/-10.7\%$
Q^2 scale up/down ($t\bar{t}$)	$+11.7/-7.0\%$
Q^2 scale up/down (W/Z)	$\pm 4.8\%$
Matching up/down ($t\bar{t}$)	$+2.1/-2.0\%$
Matching up/down (W/Z)	$\pm 8.2\%$
Pile up	$\pm 4.2\%$
PDF	$\pm 3.0\%$
ISR / FSR	$\pm 4.6\%$
μ-trigger/ID/iso scale	$\pm 3.0\%$
Total systematic	$+21.2/-17.7\%$
Statistical	$\pm 7.6\%$
Total uncertainty (stat + syst)	$+22.6/-19.3\%$
Lumi	$\pm 4.0\%$

Table 8.1: Summary of uncertainties from various sources.

related, such that they can be added in quadrature to obtain the overall uncertainty of $^{+22.6}_{-19.3}\%$ (excluding the luminosity uncertainty, which is quoted separately).

The parameters used in the fitting procedure are the production cross sections of each process considered (*e.g.* $\sigma(t\bar{t})$, $\sigma(W \rightarrow l\nu)$, etc.). The ratios between these fitted cross sections and the corresponding theoretical cross sections represent the factors by which the Monte Carlo samples should be scaled to represent the data. In order to obtain the true factor it is also necessary to include the acceptance for both data and Monte Carlo:

$$s = \frac{\sigma_{\text{fit}}}{\sigma_{\text{th}}} \cdot \frac{\epsilon_{\text{data}}}{\epsilon_{\text{MC}}} \tag{8.1}$$

The ratio $\epsilon_{\text{data}}/\epsilon_{\text{MC}}$ was determined in Section 6.7 to be $0.995 \times 0.970 = 0.965$.

Table 8.2 shows the values of these scale factors, including the trigger, ID and isolation efficiency corrections. As was seen in Section 7.1, the results of Monte Carlo simulation underestimate the contributions of the processes found in data. This is also reflected in the

Sample	Scale factor
$t\bar{t}$	1.01 ± 0.08
$W \to l\nu$	1.11 ± 0.07
$Z/\gamma^* \to l^+l^-$	1.16 ± 0.33
Single-top	1.04 ± 0.30
QCD	2.47 ± 0.43

Table 8.2: Scale factors extracted from the fit. These scale factors have been adjusted for the results of the muon trigger, ID and isolation efficiency measurement.

table, where we find the scale factors for all processes in the fit are greater than one. In some cases, particularly for $t\bar{t}$, the scale factors are consistent with unity, indicating that this process is well modeled in Monte Carlo, and requires little scaling to match the data.

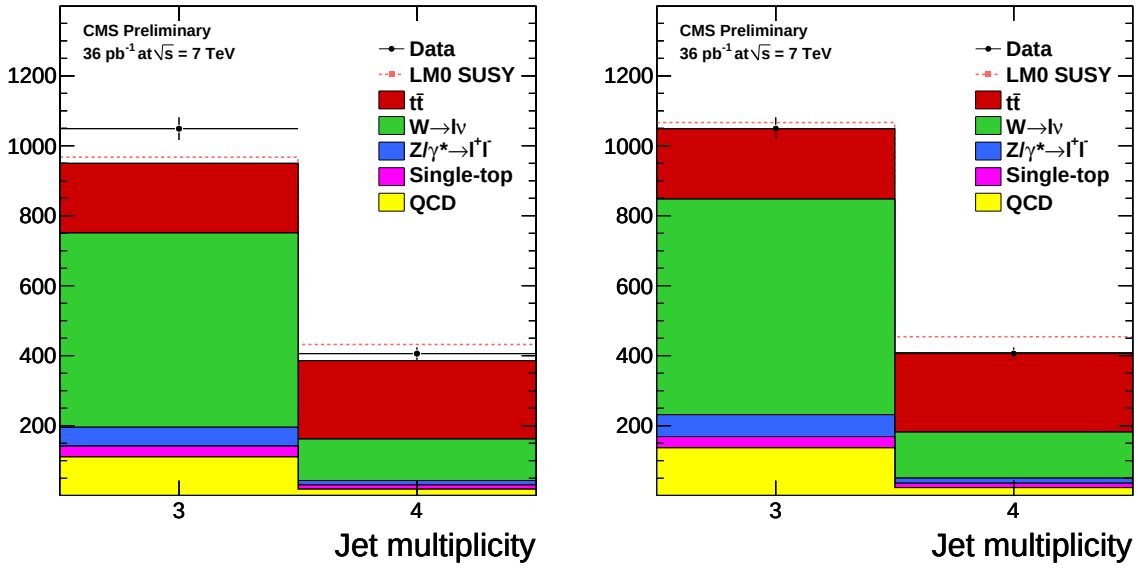
The total cross section for $t\bar{t}$ production, measured in the muon + jets channel in this analysis, is therefore

$$\sigma(t\bar{t}) = 159.1 \pm 12.1(\text{stat})^{+33.8}_{-28.2}(\text{syst}) \pm 6.4(\text{lumi})\text{pb} \tag{8.2}$$

with the statistical, systematic and luminosity uncertainties are quoted separately. The luminosity is determined separately [63], and multiple sources of uncertainty are taken into account. Added in quadrature, the total statistical and systematic uncertainty for this measurement is $^{+35.9}_{-30.7}$pb. This is in agreement with the theoretical prediction based on NLO calculations in Monte Carlo.

The scale factors calculated in Table 8.2 can be applied to our original kinematic distributions, permitting us to compare the data to the Monte Carlo scaled by the results of the fits. In Figure 8.1 we see the number of events in the 3 and ≥ 4 jet bins compared between data and MC. Contributions from the SUSY parameter point LM0 are also shown to give a representative estimate of the comparative size of possible new physics signals.

Figure 8.1(a) compares data to the original MC samples, scaled to the luminosity based on NLO calculations. Figure 8.1(b) shows the data compared to the MC scaled by the results from the fit. The overall normalization is significantly improved in the later case.



(a) (b)

Figure 8.1: Jet multiplicity distribution for data compared to Monte Carlo (a) normalized to the NLO cross section and (b) using the results from the fit. The SUSY parameter point LM0 is also shown for comparison.

The distributions used in the simultaneous fit, muon p_{T} and muon η, are shown in Figure 8.2 with the MC scaled to the results of the fit. Similarly, Figure 8.3 shows the H_{T} and $\not\!\!E_T$ distributions with the MC scaled to the fit. The normalization between data and MC should be compared to that shown in Figures 7.1 and 7.4.

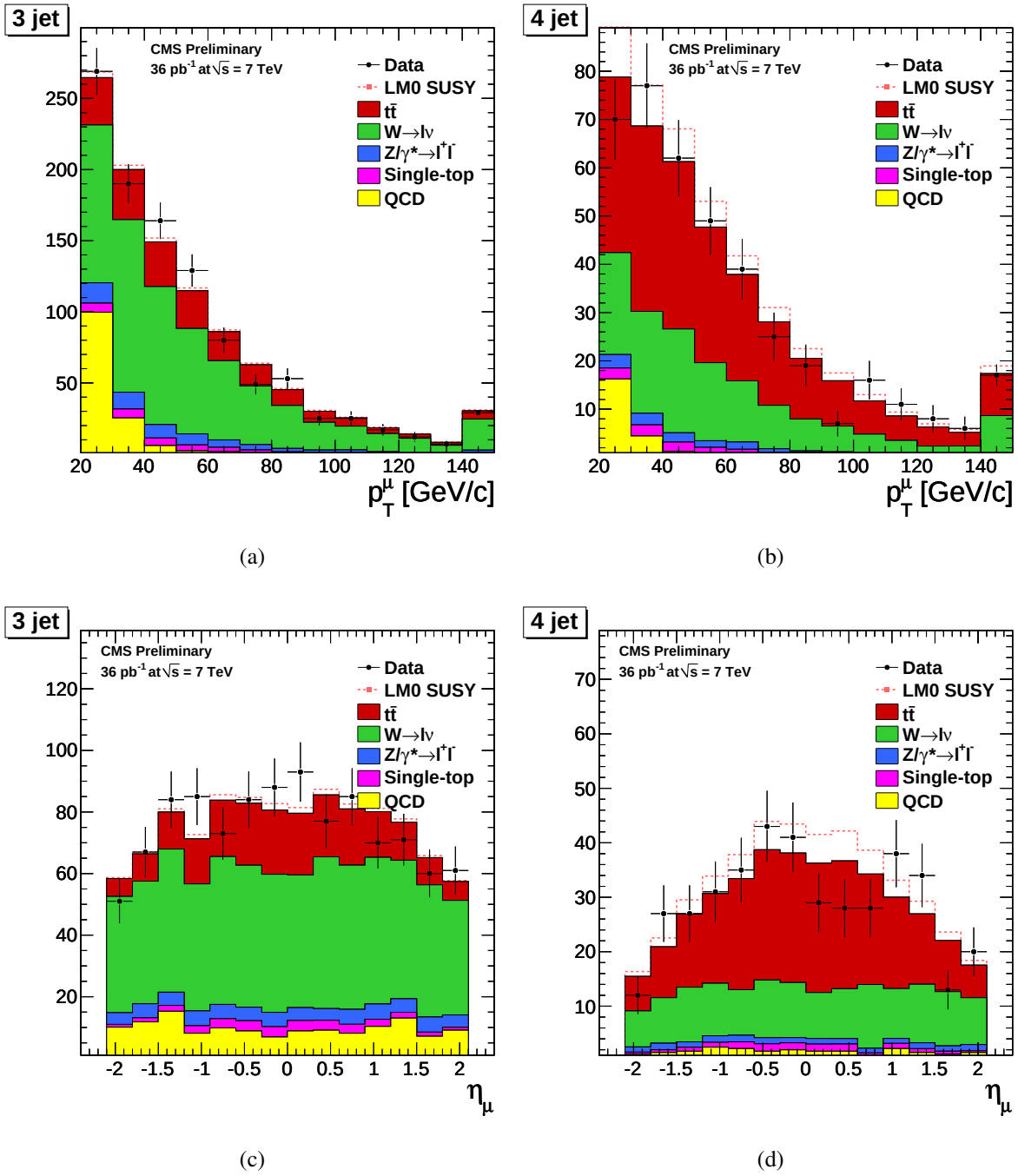
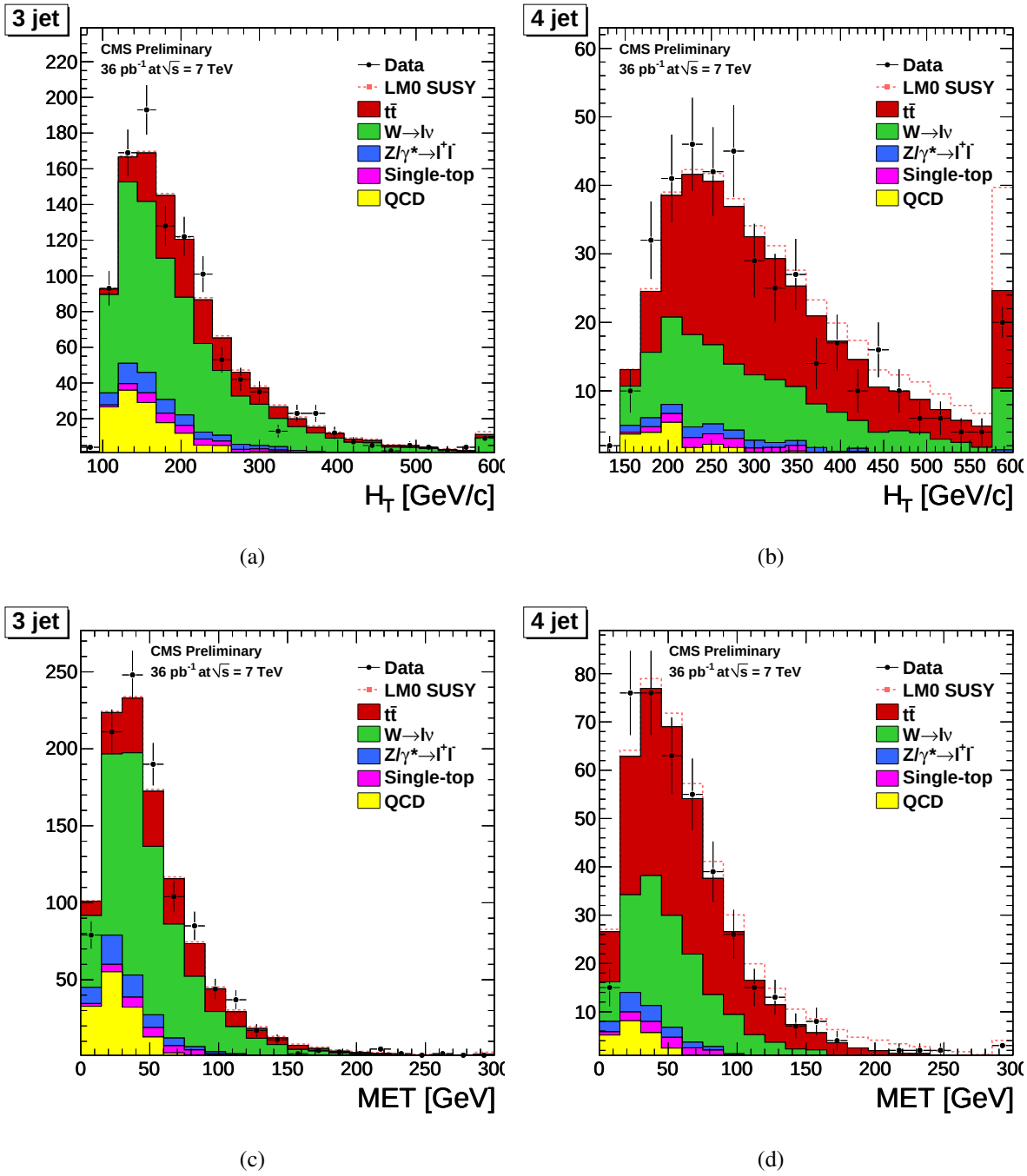


Figure 8.2: Distributions of muon p_{T} (top) and muon η (bottom) for data compared to Monte Carlo normalized to the fit results for events with 3 jets (left) and ≥ 4 jets (right). The SUSY parameter point LM0 is also shown for comparison.



(a) (b)

(c) (d)

Figure 8.3: Distributions of H_{T} (top) and $\not\!\!E_T$ (bottom) for data compared to Monte Carlo normalized to the fit results for events with 3 jets (left) and $\geq$ 4 jets (right). The SUSY parameter point LM0 is also shown for comparison.

8.2 Same Sign Dilepton Events

The results of the same-sign dilepton search in all three selection regions are summarized in Table 8.3. Each region is divided into the three same-sign dilepton channels considered in this analysis: ee, $e\mu$ and $\mu\mu$. In each channel predictions for SM contributions obtained from the fake rate method described in Section 7.9 are shown for both the data and the Monte Carlo simulation, along with the number of events observed in data. In all cases, the MC simulated samples are scaled based on the results of Table 8.2.

Table 8.4 summarizes the composition of backgrounds from events with one or two fake leptons. for all three search regions presented in Table 8.3. In the $\not\!\!E_T > 80\,\text{GeV}$ and the $H_\text{T} > 200\,\text{GeV}$ regions the entire contribution comes from the lepton plus jets background where the second lepton candidate is a fake from a jet. The low-p_T lepton analysis additionally has a small but non-negligible background contribution to the same-sign dilepton signal from events with two fake leptons, of order 20%. All estimates of backgrounds from events with fake leptons were obtained directly from data via the fake rate method, as described in Section 7.9.

In total, there are two events in data passing the selection criteria of Chapter 6. In Table 8.3, the same ee event passes all three selections. This was event number 156279004 in run 148822, taken on October 24, 2010. In addition, the $H_\text{T} > 200\,\text{GeV}$ region has one $e\mu$ event: 149934571 in run 147755, taken on October 12, 2010. Figure 8.4 shows event displays for each of these events. The ee event shows two electrons of charge -1 with $E_\text{T} = 77\,\text{GeV}$ and $E_\text{T} = 22\,\text{GeV}$ accompanied by five jets above the $p_\text{T} = 30\,\text{GeV}/c$ threshold and $\not\!\!E_T = 109\,\text{GeV}$. The $e\mu$ event has one muon and one well isolated electron of

Search region	ee	$e\mu$	$\mu\mu$	total	95% CL UL yield
Lepton trigger					
$E_T^{\text{miss}} > 80$ GeV					
MC	0.09	0.28	0.20	0.58	
predicted BG	0.07 ± 0.04	0.64 ± 0.17	0.43 ± 0.22	1.14 ± 0.25	
observed	1	0	0	1	4.2
$H_{\text{T}} > 200$ GeV					
MC	0.26	0.47	0.11	0.84	
predicted BG	0.22 ± 0.08	0.79 ± 0.18	0.54 ± 0.19	1.55 ± 0.26	
observed	1	1	0	2	5.2
H_{T} trigger					
Low-p_T					
MC	0.16	0.18	0.05	0.39	
predicted BG	0.15 ± 0.02	0.33 ± 0.08	0.05 ± 0.03	0.52 ± 0.07	
observed	1	0	0	1	4.5

Table 8.3: Observed and estimated event yields for all search regions. The rows labeled "predicted BG" refer to the sum of the data-driven estimates of the fake lepton contributions (from Section 7.9). The rows labeled "MC" refer to the SM background as predicted from the simulation alone. Rows labeled "observed" show the actual number of events seen in data. The last column (95% CL UL yield) represents observed upper limits on event yields from new physics.

	$\not{E}_T > 80$ GeV			$H_{\text{T}} > 200$ GeV			Low-p_{T}		
	1 fake l	2 fake l	total BG	1 fake l	2 fake l	total BG	1 fake l	2 fake l	total BG
ee	0.07	0.0	0.07	0.22	0.0	0.22	0.12	0.03	0.15
$e\mu$	0.64	0.0	0.64	0.79	0.0	0.79	0.31	0.02	0.33
$\mu\mu$	0.43	0.0	0.43	0.54	0.0	0.54	0.0	0.05	0.05
total	1.14	0.0	1.14	1.55	0.0	1.55	0.43	0.09	0.52

Table 8.4: A summary of the expected number of background events with one or two fake leptons for the three search regions.

charge $+1$. The muon has a $p_{\mathrm{T}} = 32\,\mathrm{GeV}/c$ and the electron has $E_{\mathrm{T}} = 33\,\mathrm{GeV}$. In addition, the event contains three jets above threshold and $\not\!\!E_T = 49\,\mathrm{GeV}$.

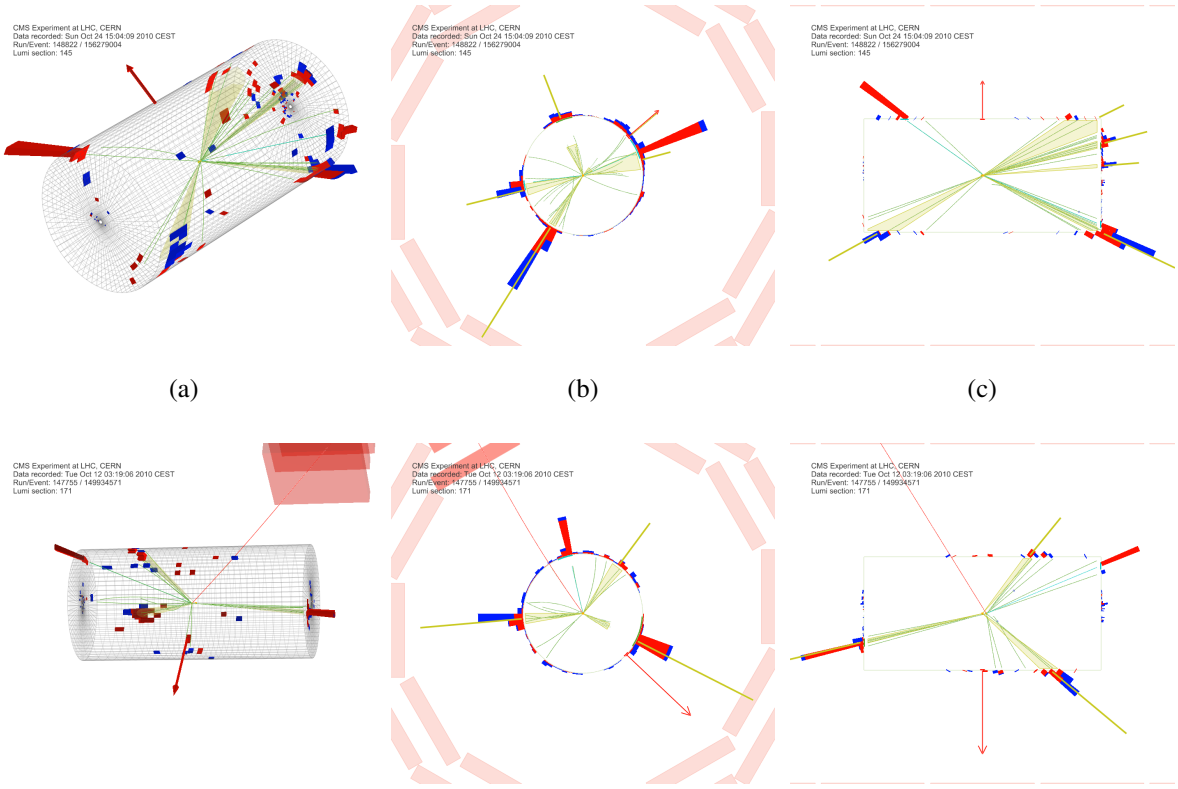


Figure 8.4: Event displays showing the two events passing all selection criteria in data. On the top is the event with two same-sign electrons, and on the bottom is the event with a same-sign electron-muon pair. In both cases a 3D view is shown along with a view in ρ-ϕ space and a view in ρ-Z space.

There is no evidence of an excess in the event yield over the expected number of observed events, and we therefore set a 95% confidence level upper limit on the event yield due to new physics using a Bayesian method [9] with a flat prior on the signal strength. Taking the example parameter point LM0 for the mSUGRA model of SUSY, we find yields

of 5.0, 6.0 and 5.7 events for the $\not\!\!E_T > 80\,\text{GeV}$, $H_\text{T} > 200\,\text{GeV}$ and low-p_T lepton regions, respectively. These yields are based on the NLO cross section for this parameter point. We can therefore exclude LM0 at the 95% confidence level. The signal region yields and SM background contributions for all search regions are summarized in Figure 8.5, as well as the 95% CL upper limit.

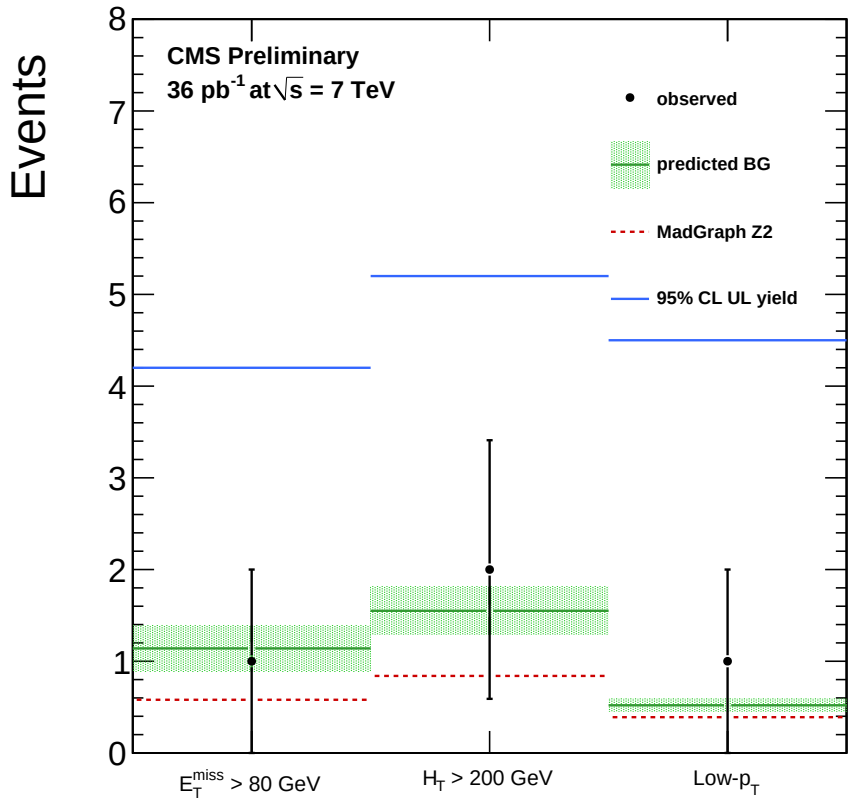


Figure 8.5: Summary of the observed number of data events and the expected number of background events from Standard Model sources obtained from Monte Carlo simulation (in red) and the fake rate method (in green). Also shown is the 95% CL upper limit determined from the Bayesian method.

Chapter 9

Conclusions

An expert is a person who avoids the small errors while sweeping on to the grand fallacy.

— Steven Weinberg

9.1 Summary

A complete analysis is presented yielding a measurement of the cross section for top-quark production in parallel with a search for new physics with same-sign dilepton events. The analysis is performed at the LHC with the CMS detector using proton-proton collisions at a center-of-mass energy of 7 TeV. The data sample used corresponds to a total integrated luminosity of $36\,\mathrm{pb}^{-1}$ and represents the complete set of validated data taken in 2010. Events taken from this data are reconstructed with CMS software as described in Chapter 5 and the results are compared with theoretical models.

The top cross section measurement is based on the reconstruction of a final state consisting of one isolated, high transverse-momentum muon and hadronic jets. After the event selection described in Chapter 6, multiple kinematic distributions were used in a simultaneous binned likelihood template fit, and the results were used to determine the cross sections

of the processes involved. After efficiency corrections, the final measured cross section for the muon + jets channel is found to be $\sigma(t\bar{t}) = 159.1^{+35.9}_{-30.7}(\text{stat + syst}) \pm 6.4(\text{lumi})\text{pb}$. This result is in excellent agreement with current theoretical values (e.g. [70, 71, 72, 73, 74]), which are also in agreement with each other, within quoted uncertainties. For example, [72] computes a theoretical top pair production cross section of $\sigma_{\text{NNLO approx}}(t\bar{t}) = 163^{+7}_{-5}(\text{scale}) \pm 9(\text{PDF})\text{pb}$. This result also agrees with the earlier CMS measurement of the cross section in the lepton + jets channel [65] of $\sigma(t\bar{t}) = 173^{+39}_{-32}(\text{stat + syst})\text{pb}$, as well as the ATLAS measurement in the combined dilepton and lepton + jets channels [75] of $\sigma(t\bar{t}) = 145 \pm 31(\text{stat})^{+42}_{-27}(\text{syst})\text{pb}$, but has a smaller uncertainty than both of these previous results.

The new physics search uses events with same-sign dileptons and includes final states with ee, $e\mu$ and $\mu\mu$. In addition, multiple trigger strategies are employed, and multiple corresponding search regions are examined in order to cover the widest-possible phase space. In all search regions, and for all channels, the dominant contributions involved events with one fake or non-isolated lepton. Estimates for both these SM backgrounds and those coming from events with two fake leptons were determined directly from data using the "fake rate" method described in Chapter 7. No evidence was seen for an excess over the SM prediction, and we have set 95% CL upper limits on the total number of signal events expected in $36\,\text{pb}^{-1}$. These limits ranged from 4.2 to 5.2 events, depending on the signal region. These results, along with signal efficiencies (obtained from Ref [64]) can be used to examine a wide variety of new physics models.

9.2 Outlook

The prospects for both parts of this analysis, the top cross section measurement and the SUSY search, are considerably enhanced by the tremendous rate at which data is being

taken at the LHC. At the time of this writing, an integrated luminosity of $1.2\,\text{fb}^{-1}$ has been recorded by CMS. Future studies with this enlarged data sample will make possible higher-precision measurements of the top cross section, and will be particularly useful for the study of $t\bar{t}+4$ or more jets, which currently suffers from limited statistics. In addition, events with two well-isolated same-sign leptons are very rare in the Standard Model, and hence searches for new physics in this channel stand to benefit considerably from higher luminosity studies. This may make it possible to observe potential signals from SUSY or other new physics in these very rare events in the coming years.

Bibliography

[1] Weinberg, S. *The Quantum Theory of Fields*. Cambridge University Press., 1995.

[2] F. Halzen, A. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley & Sons, Inc., 1984.

[3] V. Barger, R. Phillips. *Collider Physics, Updated Edition*. Westview Press, 1991.

[4] H. David Politzer. Asymptotic Freedom: An Approach to Strong Interactions. *Phys. Rept.*, 14:129–180, 1974.

[5] F. Abe et al. Observation of top quark production in anti-p p collisions. *Phys. Rev. Lett.*, 74:2626–2631, 1995.

[6] S. Abachi et al. Observation of the top quark. *Phys. Rev. Lett.*, 74:2632–2637, 1995.

[7] Joseph R. Incandela, Arnulf Quadt, Wolfgang Wagner, and Daniel Wicke. Status and prospects of top-quark physics. *Progress in Particle and Nuclear Physics*, 63(2):239–292, 2009.

[8] Lyndon Evans and Philip Bryant. LHC Machine. *Journal of Instrumentation*, 3(08):S08001, 2008.

[9] Particle Data Group Collaboration. Review of Particle Physics. *J. Phys.*, G 37, 2010.

[10] Stephen P. Martin. A Supersymmetry Primer. 1997.

[11] Marc Davis, George Efstathiou, Carlos S. Frenk, and Simon D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. *Astrophys. J.*, 292:371–394, 1985.

[12] NASA, CFHT, CXC, U. Victoria, A. Mahdavi, et al. Dark matter, February 2005.

[13] Michael Peskin and Daniel Schroeder. *An Introduction to Quantum Field Theory*. Westview Press, 1995.

[14] E. Witten. Dynamical breaking of supersymmetry. *Nuclear Physics B*, 188:513–554, October 1981.

[15] Sidney Coleman and Jeffrey Mandula. All possible symmetries of the s matrix. *Phys. Rev.*, 159(5):1251–1256, Jul 1967.

[16] Nir Polonski. *Supersymmetry: Structure and Phenomena: Extensions of the Standard Model*. Springer, 2001.

[17] Z. Kunszt and F. Zwirner. Testing the Higgs sector of the minimal supersymmetric standard model at large hadron colliders. *Nucl. Phys.*, B385:3–75, 1992.

[18] Rohini M. Godbole, Probir Roy, and Xerxes Tata. Tau signals of R-parity breaking at LEP-200. *Nucl. Phys.*, B401:67–92, 1993.

[19] Gautam Bhattacharyya. R-parity-violating supersymmetric Yukawa couplings: A mini-review. *Nucl. Phys. Proc. Suppl.*, 52A:83–88, 1997.

[20] Gianfranco Bertone, Dan Hooper, and Joseph Silk. Particle dark matter: evidence, candidates and constraints. *Physics Reports*, 405(5-6):279–390, 2005.

[21] Makoto Kobayashi and Toshihide Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[22] Riccardo Barbieri and L. J. Hall. Signals for supersymmetric unification. *Phys. Lett.*, B338:212–218, 1994.

[23] Riccardo Barbieri, Lawrence J. Hall, and Alessandro Strumia. Violations of lepton flavor and CP in supersymmetric unified theories. *Nucl. Phys.*, B445:219–251, 1995.

[24] R. Michael Barnett, John F. Gunion, and Howard E. Haber. Discovering supersymmetry with like-sign dileptons. *Physics Letters B*, 315(3-4):349–354, 1993.

[25] Manoranjan Guchait and D. P. Roy. Like-sign dilepton signature for gluino production at the cern lhc including top quark and higgs boson effects. *Phys. Rev. D*, 52(1):133–141, Jul 1995.

[26] Howard Baer, Chih hao Chen, Frank Paige, and Xerxes Tata. Signals for minimal supergravity at the cern large hadron collider. ii. multilepton channels. *Phys. Rev. D*, 53(11):6241–6264, Jun 1996.

[27] A. et al. Abulencia. Inclusive search for new physics with like-sign dilepton events in $p\bar{p}$ collisions at $s = 1.96tev$. *Phys. Rev. Lett.*, 98(22):221803, Jun 2007.

[28] The D0 Collaboration. Search for the associated production of charginos and neutralinos in the like sign dimuon channel. *D0 Conference Note*, (5126), 2006.

[29] Friedemann Brandt. Lectures on supergravity. *Fortsch. Phys.*, 50:1126–1172, 2002.

[30] T. Aaltonen et al. Measurements of the top-quark mass using charged particle tracking. *Phys. Rev.*, D81:032002, 2010.

[31] Victor Mukhamedovich Abazov et al. Measurement of the top quark pair production cross section in the lepton+jets channel in proton-antiproton collisions at $\sqrt{s}$=1.96 TeV. 2011.

[32] CMS Collaboration. CMS Luminosity - Public Results. https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults2010, 2010.

[33] CMS Collaboration. The CMS Experiment at the CERN LHC. *Journal of Instrumentation*, 3:S08004, 2008.

[34] CMS Collaboration. Performance and operation of the CMS electromagnetic calorimeter. *Journal of Instrumentation*, 5:3010–+, March 2010.

[35] CMS Collaboration. Performance of the CMS hadron calorimeter with cosmic ray muons and LHC beam data. *Journal of Instrumentation*, 5:3012–+, March 2010.

[36] N Metropolis and S Ulam. The Monte Carlo Method. *Journal of the American Statistical Association*, 44(247):335–341, 1949.

[37] M. A. Dobbs et al. Les Houches guidebook to Monte Carlo generators for hadron collider physics. 2004.

[38] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne. Physical gluons and high E(T) jets. *Phys. Lett.*, B604:61–68, 2004.

[39] Daniel Stump et al. Inclusive jet production, parton distributions, and the search for new physics. *JHEP*, 10:046, 2003.

[40] Torbjörn Sjöstrand. A model for initial state parton showers. *Physics Letters B*, 157(4):321–325, 1985.

[41] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand. Parton fragmentation and string dynamics. *Physics Reports*, 97:31–145, July 1983.

[42] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05:026, 2006.

[43] G. Corcella et al. HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes). *JHEP*, 01:010, 2001.

[44] Fabio Maltoni and Tim Stelzer. MadEvent: Automatic event generation with MadGraph. *JHEP*, 02:027, 2003.

[45] Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, Roberto Pittau, and Antonio D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 07:001, 2003.

[46] The CMS Collaboration. Measurement of the Underlying Event Activity at the LHC with $\sqrt{s} = 7\,\text{TeV}$ and Comparison with $\sqrt{s} = 0.9\,\text{TeV}$. *PAS-QCD-10-010*, 2010.

[47] Peter Z. Skands. Tuning Monte Carlo generators: The Perugia tunes. *Phys. Rev. D*, 82(7):074018, Oct 2010.

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

[49] CMS Collaboration. Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. *CMS-PAS*, PFT-09-001, Apr 2009.

[50] CMS Collaboration. CMS tracking performance results from early LHC operation. *The European Physical Journal C - Particles and Fields*, 70:1165–1192, 2010. 10.1140/epjc/s10052-010-1491-3.

[51] R Frühwirth, Wolfgang Waltenberger, and Pascal Vanlaer. Adaptive Vertex Fitting. Technical Report CMS-NOTE-2007-008. CERN-CMS-NOTE-2007-008, CERN, Geneva, Mar 2007.

[52] CMS Collaboration. Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV. *CMS-PAS*, PFT-10-002, 2010.

[53] CMS Collaboration. Electron Reconstruction and Identification at $\sqrt{s} = 7$ TeV. *CMS-PAS*, EGM-10-004, 2010.

[54] CMS collaboration collaboration collaboration collaboration. Cms physics technical design report volume i: Detector performance and software. 2006.

[55] Wolfgang Adam, R Frühwirth, Are Strandlie, and T Todor. Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC. Technical Report CMS-NOTE-2005-001. CERN-CMS-NOTE-2005-001, CERN, Geneva, Jan 2005.

[56] S. Baffioni et al. Electron reconstruction in CMS. *Eur. Phys. J.*, C49:1099–1116, 2007.

[57] CMS Collaboration. Performance of muon identification in pp collisions at $\sqrt{s} = 7$ TeV. *CMS-PAS*, MUO-10-002, 2010.

[58] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063, 2008.

[59] Stephen D. Ellis and Davison E. Soper. Successive combination jet algorithm for hadron collisions. *Phys. Rev.*, D48:3160–3166, 1993.

[60] CMS Collaboration. Jet Performance in pp Collisions at 7 TeV. *CMS-PAS*, JME-10-003, 2010.

[61] CMS Collaboration. Jet Energy Corrections determination at 7 TeV. *CMS-PAS*, JME-10-010, 2010.

[62] CMS Collaboration. Missing Transverse Energy Performance in Minimum-Bias and Jet Events from Proton-Proton Collisions at $\sqrt{s} = 7\,\text{TeV}$. *CMS-PAS*, JME-10-004, 2010.

[63] The CMS Collaboration. Measurement of CMS Luminosity. *PAS-EWK-10-004*, 2010.

[64] Serguei Chatrchyan et al. Search for new physics with same-sign isolated dilepton events with jets and missing transverse energy at the LHC. *JHEP*, 06:077, 2011.

[65] Serguei Chatrchyan et al. Measurement of the Top-antitop Production Cross Section in pp Collisions at sqrt(s)=7 TeV using the Kinematic Properties of Events with Leptons and Jets. 2011.

[66] G. L. Kane, Chris Kolda, Leszek Roszkowski, and James D. Wells. Study of constrained minimal supersymmetry. *Phys. Rev. D*, 49(11):6173–6210, Jun 1994.

[67] The CMS Collaboration. Determination of the Jet Energy Scale in CMS with pp Collisions at sqrts = 7 TeV. *PAS-JME-10-010*, 2010.

[68] V. Khachatryan et. al. et. al. First measurement of the cross section for top-quark pair production in proton-proton collisions at. *Physics Letters B*, 695(5):424–443, 2011.

[69] The CMS Collaboration. Performance of Methods for Data-Driven Background Estimation in SUSY Searches. *PAS-SUS-10-001*, 2010.

[70] M. Aliev et al. – HATHOR – HAdronic Top and Heavy quarks crOss section calculatoR. *Comput. Phys. Commun.*, 182:1034–1046, 2011.

[71] Ulrich Langenfeld, Sven-Olaf Moch, and Peter Uwer. Measuring the running Top quark mass. 2010.

[72] Nikolaos Kidonakis. Next-to-next-to-leading soft-gluon corrections for the top quark cross section and transverse momentum distribution. *Phys. Rev.*, D82:114030, 2010.

[73] John M. Campbell and R. Keith Ellis. Radiative corrections to Z b anti-b production. *Phys. Rev.*, D62:114012, 2000.

[74] R. Kleiss and W. J. Stirling. Top quark production at hadron colliders: Some useful formulae. *Zeitschrift für Physik C Particles and Fields*, 40:419–423, 1988. 10.1007/BF01548856.

[75] Georges Aad et al. Measurement of the top quark-pair production cross section with ATLAS in pp collisions at $\sqrt{s} = 7\,\text{TeV}$. *Eur. Phys. J.*, C71:1577, 2011.