UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

MEASUREMENT OF THE TOP QUARK POLE MASS

AND QCD RADIATION IN $t\bar{t}$ PAIR PRODUCTION

USING $e\mu$ EVENTS WITH b-TAGGED JETS

IN pp COLLISIONS AT $\sqrt{s} = 7 - 8$ TeV

WITH THE ATLAS DETECTOR AT THE LHC

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

CAROLYN M. BERTSCHE
Norman, Oklahoma
2015

Measurement of the Top Quark Pole Mass and QCD Radiation in $t\bar{t}$ Pair Production using $e\mu$ Events with b-tagged Jets in pp Collisions at $\sqrt{s} = 7 - 8$ TeV with the ATLAS Detector at the LHC

A Dissertation Approved for the
Homer L. Dodge Department of Physics and Astronomy

By

_____

Dr. Michael Strauss, Chair

_____

Dr. Braden Abbott

_____

Dr. Chung Kao

_____

Dr. Sheena Murphy

_____

Dr. Lance Lobban

To my initial state physics cluster at Taylor University.

Acknowledgements

While technically the interactions studied in this dissertation did occur in a vacuum, my efforts surely did not. I am deeply grateful for the support and encouragement of so many that have played integral roles in the interaction of my consciousness with the elegant world of physics, leading to this dissertation.

First of all, I would like to express my deep gratitude to my advisor, Dr. Mike Strauss, for his guidance, encouragement, patience, and holistic advice throughout graduate school, from our efficient meetings, to exploring Swiss chocolate and culture and solving conundrums even by email from 5000 miles away. You have given me both the freedom and space to develop my research independently, and the challenge and direction to which to rise, and for that I am always grateful.

I am also thankful for the High Energy Experimental Physics group at OU for encouraging and financially supporting my research, listening to my presentations and offering constructive and feedback. To my doctoral committee, thank you for listening to my talks, evaluating my progress and reviewing my writing with patience and good temper through the years. Thank you as well to the OU Physics department faculty and staff, for making my pursuit of physics education as robust and rewarding as it has been. I'll always carry with me the spirit of scientific curiosity and fascination communicated in lectures and meetings, with special thanks to Dr. Chung Kao, Dr. Ronald Kantowski, and Dr. Kim Milton—and I will remember your multi-hour exams with a special kind of scientific joy somewhere between enthralled interest and dread. But truly, thank you.

professors and scientists from various countries who have encouraged and honed my scientific endeavors along the way.

I cannot hope to properly thank the number of people who have influenced and enabled my thoroughly enjoyable physics trajectory. I am indebted to my professors and fellow students at Taylor University, for our first years of the physics world full of curiosity and enjoyment (and shared burden of rigorous scholarship). I would especially like to thank Dr. Ken Kiers, who believed in my potential to pursue science before I knew to believe in myself.

I owe a special debt of thanks to my first scientific (mathematics) role model, Irv Benson. Thank you for teaching, by showing, the beauty of elegant mathematics. I would like to express my thanks to Dave Allan for the opportunity to pursue my first experiments, and for the wonderful crew at Tecplot for an atmosphere of growth, camaraderie, and encouragement as I began to pursue graduate school.

And finally, my scientific endeavors and my life have been enriched and emboldened by the support and encouragement of my family and friends, without whose continual optimism and unwavering support across the miles, and occasional cajoling to come out of my physics bubble, my world would be gray indeed. Mom, Dad, David, Tiffany, Alex, Orchid, and the Bertsches, thank you. And at last, I am grateful to my husband, David Bertsche, for collaborating in physics and in life with me, and for your patience and encouragement with my many hours of research. From the first quantum mechanics lecture to the obscure speculation on the physical and metaphorical signatures of the universe and our places in it, thank you.

Table of Contents

List of Figures

List of Tables

Abstract

As heavy as an atom of gold but much smaller than a proton, the top quark is the most massive fundamental particle known in the universe and an integral part of current particle physics research. The characteristics of the top quark, such as its high mass—close to the scale of electroweak symmetry breaking—and its dominance in the running of the recently discovered Higgs boson's self-coupling, make the top quark a cornerstone of current research. Studies of the top quark are probes of new physics, essential tools to understand the Higgs boson, and valuable tests of the Standard Model. This thesis presents two measurements of top quark pair production events collected with the ATLAS Detector at the CERN Large Hadron Collider (LHC) in pp collisions at a center-of-mass energy $\sqrt{s} = 7 - 8$ TeV. Both measurements select $t\bar{t}$ events with an opposite-sign $e\mu$ pair and one or two b-tagged jets in the final state.

A study is first presented of the jet activity arising from quark and gluon radiation produced in association with top quark pairs, in events with an opposite-sign $e\mu$ pair as well as at least two b-tagged jets in the final state. It includes the complete 2012 ATLAS data sample of 20.3 fb^{-1} integrated luminosity of pp collision data at $\sqrt{s} = 8$ TeV. The fraction of events that does not contain additional jet activity in a central rapidity region is measured as a function of a) the minimum transverse momentum of any additional jet in the event, and b) the minimum scalar transverse momentum sum of all additional jets in the event, and the results given for four central rapidity regions and four regions of the invariant mass of the $e\mu bb$ system. Compensation for detector effects is applied to the experimental measurement and

the results compared at the particle level to simulations by several next-to-leading order (NLO) and leading order (LO) Monte Carlo generators. The resulting gap fraction measurements, in comparison with simulation, can be used to tune Monte Carlo modeling of quantum chromodynamic (QCD) radiation and reduce associated modeling uncertainties in ATLAS physics measurements involving top quark pair production, such as top quark measurements and Higgs boson measurements with top quark backgrounds.

Secondly, the top quark pole mass ($m_{\text{top}}^{\text{pole}}$) was extracted from top quark pair events having an opposite-sign $e\mu$ pair in the final state and one or two b-tagged jets. This measurement includes the complete 2012 ATLAS data sample, as well as the complete 2011 ATLAS data sample of 4.6 fb^{-1} integrated luminosity of pp collision data at $\sqrt{s} =$ 7 TeV. This is combined with the theoretical cross section prediction including QCD corrections at next-to-next-to-leading order (NNLO+NNLL), with three different sets of parton distribution functions. The results are:

$$m_{\text{t}}^{\text{pole}}(7 \text{ TeV}) = 171.40_{-2.56}^{+2.58} \text{ GeV} \tag{1}$$

$$m_{\text{t}}^{\text{pole}}(8 \text{ TeV}) = 174.10 \pm 2.64 \text{ GeV} \tag{2}$$

$$m_{\text{t}}^{\text{pole}}(7 - 8 \text{ TeV}) = 172.87_{-2.63}^{+2.51} \text{ GeV} \tag{3}$$

in which the uncertainties include data statistics, experimental systematic effects, the knowledge of integrated luminosity and LHC beam energy, as well as PDF, QCD scale, and strong coupling constant α_s uncertainties. An examination of consistency between the 7 and 8 TeV results found agreement within $\sim 1.6\sigma$.

Introduction

"We are the music-makers, and we are the dreamers of dreams."

– Arthur O'Shaughnessy

Background

As the world of the 1940's began to recover from the second world war, a small group of scientists and political figures in Europe and the United States came together with a vision to found an international scientific research center. This laboratory would both further fundamental research and unite nations, for the peaceful collaboration of scientific research. On 29th September 1954, twelve founding countries—including nations that had been at war a mere decade earlier—signed the convention establishing the existence of CERN, the *Centre Européen pour la Recherche Nucléaire* [1].

In the sixty years since its inception, CERN, centered in Geneva, Switzerland, has grown to include more than eighty nations from around the world, expanding the frontiers of particle physics research and cultivating a climate of international peace and stability. Astounding scientific advances have taken place at CERN, including the

1

invention of the world wide web[1], the establishment of the Large Hadron Collider, the discovery of the W [2, 3] and Z [4] particles, and the discovery of the Higgs boson, as well as searches to understand the mysteries of dark matter, dark energy, and the force of gravity. Particle accelerator technology developed in part at CERN has spread to become useful in many other areas of science and industry, such as cancer treatment, radiation therapy, materials science, medical research, forensics, and even optimization of the molecular structure of chocolate.[2]

The largest particle accelerator at CERN and in the world, the Large Hadron Collider (LHC), stretches in a circle twenty-seven kilometers in circumference, lying one hundred meters below the cow pastures, sunflower fields, majestic Alps, and picturesque villages on the border of France and Switzerland. In beam pipes cooled to temperatures colder than outer space, protons are accelerated to over 99.9999% of the speed of light and collided into each other in the center of four large particle detectors, which take sophisticated and complex "pictures" of the collisions for teams of scientists to analyze. The particles inside the accelerated protons, upon collision, are briefly transformed into a variety of less common particles as their energy and mass is transformed into different combinations of energy and mass, according to Einstein's famous equation, $E = mc^2$. (More generally, it is $E = \sqrt{m^2c^4 + p^2c^2}$.) These transformations give scientists brief glimpses into the fundamental nature of our universe.

---

[1] The supervisor of Tim Berners Lee, upon reviewing Lee's initial proposal of the HTTP protocol, commented prophetically at the top of the paper, "Vague, but interesting..."

[2] Cadbury UK used a collider in Australia to study the manufacturing of the molecular level of chocolate.

It is one of these particle detectors, the ATLAS[3] detector, that supplied the data analyzed in this dissertation. The ATLAS Detector is one of seven particle detectors at the LHC, and the largest experiment ever built—the size of a seven-story building, it is half the size of the Notre Dame cathedral of Paris, weighs about 7000 tons and took 16 years to be designed and built by scientists and engineers from around the world. More than 3000 physicists from 178 laboratories and universities from around the world, including 44 U.S. institutions, are involved in the ATLAS collaboration, analyzing the data collected by the ATLAS Detector.

The Top Quark

The Standard Model of particle physics, representing the structure of matter and its interactions as well as scientists understand it, describes fundamental particles and the laws by which they interact to form the visible universe. These include six particles called "quarks" that share common properties, including the most massive particle in the Standard Model, the top quark. Discovered in 1995 at Fermilab in Chicago, Illinois [5, 6], the top quark had been predicted to exist since 1977 with the inferred existence of the bottom quark from the discovery of the Upsilon particle (a particle composed of the bound state of a bottom and its antimatter partner, the anti-bottom quark) [7], but proved elusive to create and detect until the 1990's due to its high mass. The top quark is predicted to occur often in association with the Higgs boson, the second most massive particle of the Standard Model and the last one

---

[3]A Large Toroidal LHC Apparatus

discovered, in 2012 at CERN [8, 9]. Because the Higgs boson often occurs with the top quark, precise understanding of top quark production is crucial for understanding and measuring properties of the Higgs boson as well. In addition, the top quark cross section (the probability of top quark production relative to other particle productions for a given collision energy) is sensitive to the existence of physics beyond the Standard Model—measuring variations from prediction of top quark properties, such as mass, could point to new particles and test additional theories of particle interactions.

While the Standard Model is remarkably precise and self-consistent, with the recent discovery of the Higgs boson a fitting capstone, it does not encapsulate all of the physical universe. It does not include a description of the force of gravity, for example, or an explanation for indications of dark matter and dark energy. Especially in light of these physics frontiers, precise measurement and understanding of the top quark is crucial to current physics research.

The following thesis begins with an overview of the Standard Model of physics in Chapter 1 and the design of the ATLAS Detector at the LHC in Chapter 2. (The inclusion of irradiation effects in the ATLAS Pixel Detector is described in Appendix A.) The collection and selection of data and simulation samples are described in Chapters 3 and 4. Following these are included two analyses of top quark production: a measurement of QCD radiation in top quark production compared to simulation in Chapter 5, and an extraction of the top quark pole mass in the same decay channel in Chapter 6. Together these analyses comprise small but valuable advances in scientific understanding of the top quark.

Chapter 1

The Standard Model of Physics

"Nature uses only the longest threads to weave her patterns, so that each small piece
of her fabric reveals the organization of the entire tapestry."

– Richard Feynman

The quest to understand the fabric of the universe has led philosophers and scientists through diverse realms of thought, theory and experimental discovery over the past millennia. The precipice of current knowledge observes the universe to be comprised of minuscule particles—indivisible as far as science understands so far—and the interactions between these particles, collectively described by a theory known as the Standard Model (SM). This chapter briefly summarizes the fundamental concepts of the Standard Model, with emphasis on aspects of the particle studied in this thesis, the top quark. Section 1.1 describes the gauge theory and classification of the different particles and interactions. Section 1.2 describes how these particles interact in the context of the protons collided at the LHC. Section 1.3 discusses the post-collision combination of heavier particles into groupings called "jets", relevant for the gap

fraction measurement of Chapter 5. Section 1.4 describes a few key aspects of the top quark, and Section 1.5 discusses physics beyond the Standard Model. A full treatment of the Standard Model is beyond the scope of this work; for more thorough description, see works such as Ref. [10, 11].

1.1 The Standard Model

The Standard Model, developed in the 1960's and 1970's, is a relativistic quantum field theory that explains a physical framework for the universe, categorizing all matter as made up of elementary, point-like and structureless particles characterized by specific properties (such as mass and quantum numbers, including spin and angular momentum). These particles are observed to interact in four fundamental ways: electromagnetic interactions, described by quantum electrodynamics (QED), weak interactions, strong interactions described by quantum chromodynamics (QCD), and gravitational interactions. The theory that merges electromagnetic interactions with weak interactions is called electroweak theory (EW); the Standard Model is the combination of electroweak theory with strong interactions.

1.1.1 Gauge Theory

The formalism of the Standard Model is a relativistic quantum field theory, meaning that each fundamental particle is described by a quantum field [12, 13, 14, 15, 16]. All information about the fields and their interactions is encompassed in the Standard Model "Lagrangian density" $\mathcal{L}$, which is a scalar function of the quantum fields and

their four dimensional derivatives (x, y, z, t) [11]:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\overline{\psi}\slashed{D}\psi + \psi_i y_{ij}\psi_j \phi + \text{h.c.} + |D_\mu\phi|^2 - V(\phi) \tag{1.1}$$

where each term is defined as in Ref. [11]. The interaction between the fields, as well as the physical representation of particles, arises from requiring gauge invariance of the fermion fields: that is, the action $(\delta \int \mathcal{L} dt = 0)$ remains invariant under transformations of the fields in space-time, either globally (for all space) or locally. Each gauge transformation under which the Lagrangian is symmetric introduces a field corresponding to a gauge boson. These particles and interactions are described in the following section.

1.1.2 Particle Classifications

The SM particles can be classified overall into two distinct categories, called fermions (which have half-integer spin from Fermi-Dirac statistics) and bosons (having integer spins, described by Bose-Einstein statistics). In general, fermions make up matter, while bosons cause interactions between the fermions, acting as force carriers. Each fermion has a corresponding antiparticle, with opposite values of each quantum number (for example, electric charge and color, described following), except with the same mass.

Two varieties of fundamental fermions exist: leptons and quarks, both of which have properties making them apt "pairs" with another of their kind. The quarks include three "generations" or pairs of two (three isoweak doublets), with each pair including one quark with a charge of $+\frac{2}{3}e$ and the other $-\frac{1}{3}e$ (where e represents the

7

charge of the electron; each quark has a spin of $\frac{1}{2}$). The three generations are arranged by mass, with the two lowest mass quarks (the up u and down d quarks) joining the two lowest-mass leptons to make up most of the visible matter on earth. The pair of next two higher mass quarks are the charm (c) and strange (s) quarks, and the final generation and highest mass pair is made of the top (t) and bottom (b) quark. The top quark was the last and heaviest quark to be discovered, at the Tevatron in 1995, after having been postulated in 1977.

The leptons also exist in three generations (again, three isoweak doublets), increasing in mass with each generation, also having spin $\frac{1}{2}$, but with each generation much lower in mass than the quarks. Each pair consists of one lepton with charge $-1e$ (electron e, muon μ, tau τ) and one corresponding neutrally charged neutrino (ν_e, ν_μ, ν_τ). The lowest generation, the electron and electron neutrino, are the only stable leptons, meaning they do not decay to a lower generation, enabling them to make up visible matter as stated above.

Bosons provide interactions between the fermions and account for three of the four fundamental forces observed in nature: electromagnetic, weak, and strong forces. The fourth force gravity, is negligibly small between particles compared to the other three forces, and is not yet understood in the context of the Standard Model.[1] As summarized above, the bosons are described by gauge theory: the Lagrangian density is required to be invariant under local symmetry transformations, and the interaction associated with each symmetry is mediated by a spin-1 gauge boson. The Standard Model rests

---

[1]Gravitational forces become proportionally large, however, at high distances, such as cosmological interactions.

on the gauge symmetry of unitary gauge transformations: $SU_c(3) \otimes SU_L(2) \otimes U_y$.

The electromagnetic interaction is based on the symmetry group $U_y(1)$, formulating quantum electrodynamics (QED), mediated by the massless and chargeless photon γ. This can occur between any two particles carrying charge (Q). (Since photons are chargeless, they cannot interact with each other.)

The weak interaction is described by the $SU_L(2)$ symmetry group. The weak force interacts with all fermions by coupling to their weak isospin T eigenstate via three vector bosons, two of which are charged (W^+ and W^-) and one neutral (Z^0). The weak isospin eigenstates are not equal to the mass isospin eigenstates; they are related by a matrix formulated by Cabibbo, Kobayashi and Maskawa: the CKM matrix [17]. Fermions have weak isospin $T = \frac{1}{2}$, a vector with the third component of the form $T_z = \pm\frac{1}{2}$. In contrast, vector bosons have isospin $T = 1$ with $T_z = +1, -1, 0$ for W^+, W^- and Z^0, respectively. This weak interaction is weak (compared to other interactions) because of the large mass of the vector bosons involved ($m_W = 80.385 \pm 0.015$ GeV and $m_Z = 91.1876 \pm 0.0021$ GeV[2]), which are not predicted by gauge theory. Interactions through the W boson are called flavor changing because they change the flavor of the interacting particles. Interactions through the Z boson are called neutral current interactions.

The strong interaction is based on the symmetry group $SU_C(3)$ and described by quantum chromodynamics (QCD). Called the strong force, this is the strongest force at the particle level and occurs between quarks via a massless charge carrier called the

---

[2]For brevity, this thesis uses the units convention $h = c = 1$.

gluon g. Gluons are massless and chargeless but preferential in the way they interact with quarks in a way that has been named "color"; not a wavelength of light but a way to describe the property of gluons and other charge-carrying particles that determines how groupings occur. A gluon carries one color and one anti-color, and each quark (or anti-quark) carries one color (or anti-color). The gluons only cause groupings of quarks in ways that "cancel out" the colors: i.e., one color with its anti-color, or the three different colors together (analogous of red, blue and green), or the three different anti-colors together. Since the gluon carries one color and one anti-color, it can exist in eight different color states and can interact with other gluons. Because QCD interactions in top quark pair productions are studied in this thesis, they are discussed further in the following section.

The last fundamental building block in the Standard Model is the Higgs boson. This boson is the physical representation of the Brout-Englert-Higgs scalar field which permeates space and is responsible for the generation of mass in all other particles. The strength of the interaction with the Higgs field determines that particle's mass: the scalar Higgs field couples to different particles, giving them mass, as a result of spontaneous symmetry breaking. The Higgs potential is rotationally symmetric with a local maximum at zero and a rotationally symmetric local minimum, the vacuum expectation value (VEV), at which value the symmetry is thus spontaneously broken. Due to this shape, the Higgs potential is often described as a "Mexican hat" potential. The Standard Model predicts the interactions of the Higgs field with other particles, but did not predict the mass of the Higgs boson, which was observed at

CERN in 2012 [8, 9] and mass measured as of 2015 to be 125.09±0.21 (stat.) ±0.11 (syst.) GeV [18].

1.1.3 Quantum Chromodynamics

This section discusses Quantum Chromodynamics (QCD) in further detail due to its role in the the gap fraction measurement of Chapter 5. Quantum Chromodynamics (QCD) is the study of the strong force that occurs between quarks via the gluon, as well as between gluons. The interactions occur by preferences that are described as a "color" scheme, traditionally thought of as red (r, and anti-red), blue (b, and anti-blue), and green (g, and anti-green). The interactions only occur so that compound particles are always colorless, by combining red, blue and green (or anti-red, anti-blue and anti-green) or one color with its anti-color. This means that the color quantum number is not observed in nature, dictating that all interactions are invariant under "color interchange". This discrete symmetry is a special unitary group symmetry in three dimensions, $SU_C(3)$, which also allows for self-coupling between the gluons. Nine combinations of color and anti-color pairs are possible for the gluons, except that one of the states is a symmetric color singlet carrying no net color charge, so is decoupled. The remaining possible combinations of color that gluons can carry are:

$$r\bar{g}, \ r\bar{b}, \ g\bar{r}, \ g\bar{b}, \ b\bar{r}, \ b\bar{g}, \ \frac{1}{\sqrt{2}}(r\bar{r} - b\bar{b}), \ \frac{1}{\sqrt{6}}(r\bar{r} + b\bar{b} - 2g\bar{g}) \qquad (1.2)$$

"Color" is never a passively observable state, because quarks occur in nature bound by the strong force into color-neutral states called *hadrons*, rather than as free particles. Hadrons include either two quarks (mesons) or three (baryons) bound together with

total color charge of zero.

The color charge acts similarly to the electric charge between particles, in that it is mediated by a massless vector boson. However, unlike the photon which does not carry an electrical charge, the gluon does carry color charge, which allows it to couple to other gluons. (This type of theory that allows self-coupling is called non-Abelian). Because of this self-coupling, the strength of the potential as a function of distance is also different from the electric relation $\frac{1}{r}$; the quark-anti-quark potential can be described as:

$$V_s = -\frac{3}{4}\frac{\alpha_s}{r} + kr \tag{1.3}$$

in which α_s is a constant used to describe the coupling strength, which is on the order of unity (in contrast to the electroweak $\alpha \sim \frac{1}{137}$), except in the case of high-momentum collisions in which α_s decreases drastically to ~ 0.1.[3] The first term in V_s dominates at small distances, while the second term dominates at larger distances corresponding to quark confinement at large separations. It is because of this second term that separating one quark from another results in the creation of a new $q\bar{q}$ pair, called fragmentation. One physical theory for this second linear term creating a new pair proposes that the color field lines could be pulled together by the gluon-gluon interactions into a tube as stored energy (kr). When the tube is long enough, it is energetically preferential to break the tube and create two pairs with shorter tubes rather than one long tube, expected at a range of $k \sim 1$ GeV/fm.

---

[3] High momentum collisions really means high q^2 collisions, where q^2 is the square of the amount of momentum transferred in the collision, where $q = (p_i - p_f)$ for each particle, in which p_i and p_f are initial and final state momentum of the particle, respectively.

The interactions of QCD can be expressed as a Lagrangian, expressed in terms of the gluon field A_ν^a and gluon field tensor $F_{\mu\nu}^a$:

$$\mathcal{L}_{\mathrm{QCD}} = -\frac{1}{4}F^{\mu\nu,a}F_{\mu\nu}^a + \bar{q}(i\not{D} - m_q)q - g_s\bar{q}\gamma^\mu T^a q A_\mu^a \tag{1.4}$$

The first term is the kinetic term of the gluon field, where the gluon field tensor is described as:

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - g_s f^{abc} A_\mu^b A_\nu^c \tag{1.5}$$

The last term in the gluon field tensor accounts for the non-Abelian nature of the interactions, required by gauge invariance under local $SU_C(3)$ transformations of the gluon field. The second term of the QCD Lagrangian is the free Dirac field equation for quarks, and the third term is the color interaction current term between quarks and gluons.

The QCD coupling strength constant α_s can be written analogously to the QED coupling constant as $\alpha_s = \frac{g_s}{4\pi}$, in which α_s (and thus g_s) is a function of Q^2 $(= -q^2)$ such that the coupling strength decreases as Q^2 increases. This leads to asymptotic freedom, the phenomenon that at very high energies, quarks will behave like free particles! In fact this is a convenient interaction in particle colliders, where hadrons are accelerated so close to the speed of light that the inter-quark interactions may be neglected compared to hard scattering interactions, which are discussed in Section 1.3.

1.1.4 Charged Current Weak Interactions

The top quark decays about 98.6% of the time to a W boson and a b quark. Unlike electromagnetic and strong interactions, weak interactions, such as the charged current

weak interaction of a W decay, violate parity—meaning that the force interacts differently with left-handed and right-handed coordinate systems. Mathematically this invariance violation is described by the interaction equation:

$$-\frac{ig_w}{2\sqrt{2}}\gamma^\mu(1-\gamma^5) \tag{1.6}$$

in which g^w represents the weak coupling constant, γ^μ and γ^5 are Dirac matrices of the form $\mu \in 0, 1, 2, 3$; γ^μ is a vector coupling and $\gamma^\mu\gamma^5$ is an axial vector coupling, which causes the parity violation because its reflection includes a flip of the charge sign. Physically, this coupling means that the charged weak interaction acts on left-handed particles and right-handed anti-particles, where "handedness" is a property of fermions shown in Figure 1.1. "Right-handed" particles are those with parallel unit vectors of spin and momentum, and "left-handed" are those with anti-parallel spin and momentum directions. For massless particles, the handedness or chirality is intrinsic, since the relative momentum cannot change from one reference frame to another. For massive particles, the handedness, or helicity, can change between reference frames.

1.2 Partons

The fundamental particles of the Standard Model and their interactions make up all the matter observed so far in the universe. For example, the electrical charge of the proton is dictated by its constituent parts of two up quarks and one down quark, called the valence quarks of the proton. However, in the interactions of these valence quarks and the gluons binding them together can exist a "sea" of virtual quark-antiquark

14
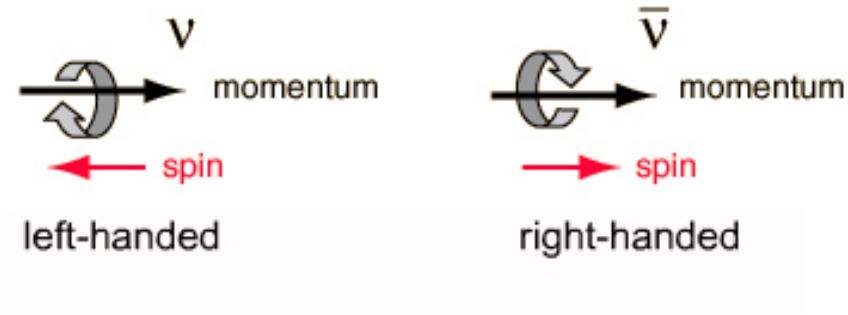
Figure 1.1: Helicity describes the relationship between a particle's spin and direction of movement. A particle is "right-handed" if its spin and momentum are parallel, and "left-handed" if anti-parallel [19].

pairs generated by the gluons, at ratios that vary according to the energy of the proton. The momentum distribution functions of these particles are called Parton Distribution Functions (PDF), where parton is the name given to all these particles that can exist inside the compound particle (such as the proton): quarks, anti-quarks, and gluons. Figure 1.2 shows an example of PDFs derived based on experiment [20].

1.3   Partons to Jet Formation

Hard scatter processes in proton-proton colliders such as at the LHC (described in Chapter 2) are collisions between partons. When two partons collide, fragmentation and hadronization of the parton follow, processes important for understanding what is detected in terms of the initial interaction in the detector. First, fragmentation occurs as the quarks, anti-quarks and gluons in collision lose energy by giving off gluons, which can decay into many additional quark-antiquark pairs. This continues to occur until
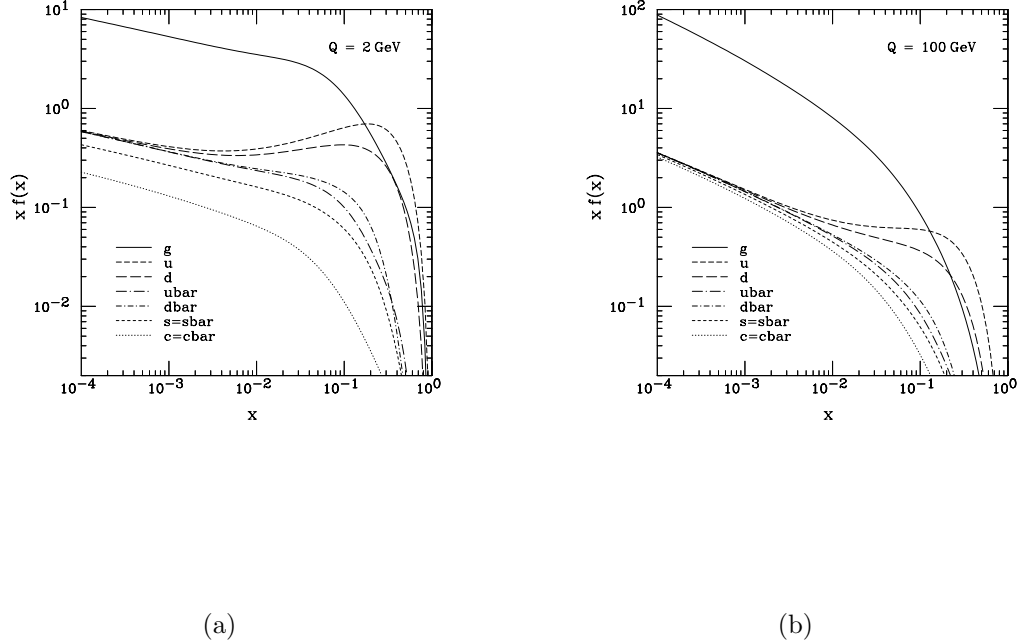
Figure 1.2: Overview of a simulated parton distribution at a) Q=2 GeV and b) Q=100 GeV [20].

the square of the virtual masses of the partons are on the order of the infrared cut-off scale t_0, on the order of 1 GeV². After this scale is reached, low-momentum-transfer long-distance interactions dominate, predominantly hadronisation: the quarks combine with other quarks to create hadrons. The hadrons can then decay into groups of particles called jets which are measured by the detector. The jets can be reconstructed from detector observations to aid in understanding the primary interaction.

1.4 The Top Quark

The existence of the top quark was predicted in 1977 by the Standard Model when the bottom quark was discovered at the Fermi National Accelerator Laboratory in Chicago, U.S.A. [7],[4] and discovered in 1995 in pp collisions at a center-of-mass energy of $\sqrt{s} = 1.8$ TeV by the CDF and $D\emptyset$ experiments at Fermilab [5, 6]. The top quark mass is not predicted by the Standard Model and has been measured with increasing precision since discovery; the latest combination of $D\emptyset$, CDF, ATLAS and CMS collaborations is $m_t = 173.34 \pm 0.27 \pm 0.71$ GeV [21]. The top quark's high mass dictates its high decay width of $\Gamma \propto m_t^3/m_W^3$, measured to be $2.0^{+0.47}_{-0.43}$ GeV with the short lifetime of $\tau \approx 3 \times 10^{-25}$ sec [22]. With its high mass and short lifetime, the top quark is the only quark that decays before forming hadrons, making it a prime candidate to test properties predicted by the Standard Model and to search for new physics.

Since quarks do not exist as free particles, as discussed above, their mass cannot be measured directly but only inferred from the measurement of hadronic observables that depend on the mass. The most common technique for measuring the top quark mass is by measuring the invariant mass of the decay products; however, this method must usually assume a top quark mass input for some of the MC simulations, creating a limitation in the mass measurement. The mass extraction described in this work avoids this limitation by instead comparing the dependence of the top quark pair cross

---

[4]The prediction followed the Standard Model description of quarks existing according to weak isospin doublets - the top quark was predicted as the weak isospin partner of the bottom quark.
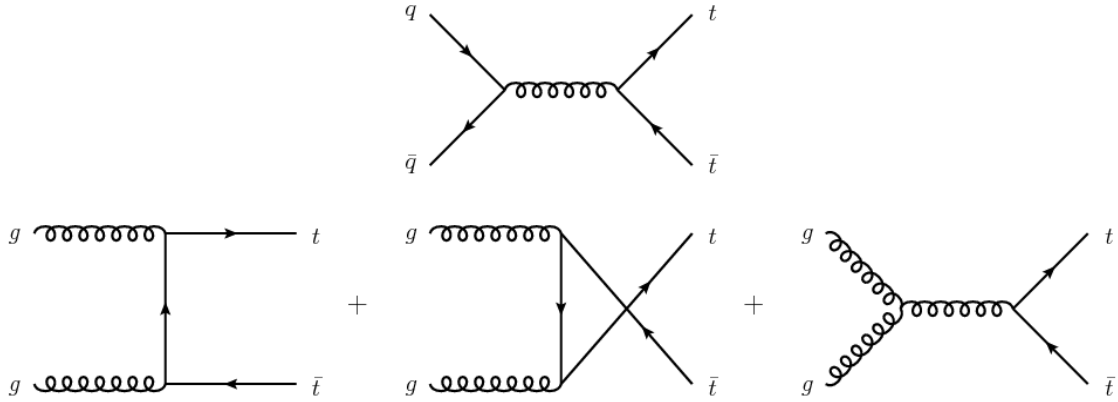
Figure 1.3: Processes by which the top quark is produced at particle colliders. The top interaction is representative of top quark pair production through $q\bar{q}$ pair annihilation. The bottom three interactions represent top quark pair production through gluon fusion.

section on the top quark mass in theory and in experiment.

1.4.1 Production of the Top Quark at the LHC

According to the interactions described above, the top quark can be created at particle accelerators by quark-antiquark interactions and gluon-gluon interactions, where the relative frequency of production mechanisms depends on the particles being collided and the center-of-mass energy of the collisions. At the Tevatron, about 85% of the $t\bar{t}$ pairs produced came through annihilation of $q\bar{q}$ pairs as shown in Figure 1.3. At the LHC, the proportions are approximately reversed: 80-90% of $t\bar{t}$ pairs are produced by gluon fusion, also shown in Figure 1.3 [23].

1.4.2 Decay of the Top Quark

The SM predicts that top quarks decay to a combination of a W boson and a b quark about 99.8% of the time. Because of this, decay signatures of the top quark pair are generally categorized by the ways in which the W boson decays.

1.5 The Horizon of the Standard Model

The Standard Model is an "effective theory", meaning that it is verifiable at least up to certain limits such as an energy level Λ, beyond which it has not yet been tested and may comprise part of a more fundamental model of the universe. It is clear that the Standard Model does not describe the whole of the universe, and current research strives to both test the predictions of the Standard Model and search for evidence for and understanding of physics beyond the Standard Model. For example, the Standard Model does not describe gravitational interactions; in addition, dark matter, dark energy, and the very small mass of neutrinos are not yet understood. The asymmetry of matter and anti-matter in the universe is also not yet explained—most observable matter is matter, and the Standard Model does predict an asymmetry due to CP violation (violation of charge conjugation and parity symmetry), but not to the extent that would explain the prevalence of matter in the observable universe.

Along with describing phenomena, another reason to search beyond the Standard Model is to attempt unification of the four fundamental forces of nature. Just as electricity and magnetism were previously thought to be separate forces that have now understood to have one source, it is hoped that electromagnetism could be unified

with the strong, weak, and gravitational forces. For example, one method proposed to do this is Supersymmetry [24]. In this theory, each fermion and each boson have a "partner" which has the same properties but is either a boson or a fermion, inversely, and has a heavier mass. A supersymmetric transformation would transfer a particle from a fermionic to bosonic state and vice versa. The supersymmetric operator Q would be an anti-commutating spinor such that:

$$Q|\text{Boson} >= |\text{Fermion} >, \quad Q|\text{Fermion} >= |\text{Boson} > \tag{1.7}$$

Since this operator Q interacts with fermions, it is thought to have angular momentum spin $\frac{1}{2}$, which requires that supersymmetry be a space-time symmetry.

Current particle physics research takes place in the context of the Standard Model, with each measurement such as those described in this thesis an attempt to both test the predictions of the Standard Model and search for any deviations from SM predictions that could indicate new physics. The following chapter describes the physical framework in which these measurements occurred.

Chapter 2

The ATLAS Detector at the LHC

"The one common misconception about the small black holes that may form at the Large Hadron Collider is that they would swallow the Earth...These black holes would be too small to consume any significant amount of matter."

– Frans Pretorius, Princeton

One hundred meters beneath the French and Swiss countrysides of vineyards, sunflower fields, cow pastures and villages hums the largest engine of scientific research ever built. The Large Hadron Collider, or LHC, is the latest iteration in a series of ever-larger circular colliders capable of accelerating particles ever-faster: the LHC accelerates protons and ions to speeds of more than 99.9999% of the speed of light and then steers them into high-energy collisions at four designated interaction points around the circular ring. The proton-proton (pp) collisions are comprised of collisions of the quarks and gluons inside the protons, causing a rich and complex array of high-energy particle interactions. Seven particle detectors, including the ATLAS Detector discussed here, take sophisticated "photographs" of these collisions in order

to analyze the particles and interactions. While a comprehensive description of the LHC and the ATLAS Detector are beyond the scope of this work, this chapter gives a brief overview of the key technologies and methods involved. Full details are available in each machine's Technical Design Report [25, 26].

2.1 The LHC

The LHC is a superconducting synchrotron, meaning that it uses superconducting electromagnetic technology to accelerate particles in a circular path. Built in the 26.7 kilometer tunnel of a previous collider at CERN, the LEP (Large Electron Positron) collider [27], the LHC is designed to collide protons at center-of-mass energies up to $\sqrt{s} = 14$ TeV at a rate of up to 40 million times per second. It lies at an average of 100 meters underground, varying from 50 meters to 150 meters with a slope of about 1.4% to ensure that the tunnel and each detector lie in solid rock while minimizing construction expense [27]. About 9600 magnets are used to optimize the trajectory of the particle beams—dipole magnets direct the protons in their near-circular path, quadrapole magnets are used to focus the beams of protons, and electromagnetic resonance chambers accelerate the particles and then compensate for energy loss to keep them at a constant energy. The beam pipes are maintained at a supercooled -271.3 degrees Celsius (1.9 K) to allow superfluid helium to be used for stable cooling of the superconducting magnets. The pressure in the beam pipes is kept at super-vacuum levels of 10^{-13} ATM, about ten times lower than the moon, to avoid collisions between gas molecules. In addition to accelerating protons, the LHC also accelerates

and collides lead ions for additional experiments, at up to 5.5 TeV per nucleon pair.

To isolate protons, electrons are stripped from hydrogen atoms using an electric field.[1] The protons are then accelerated through a series of smaller detectors. First, they are accelerated to about 50 MeV in the LINAC2[2] and then fed into the Proton Synchrotron Booster (PSB). They are boosted to about 1.4 GeV in the PSB and then channeled into the Proton Synchrotron, which accelerates them to about 25 GeV. From here, the protons are fed into the Super Proton Synchrotron (SPS), which accelerates them to about 450 GeV and then transfers them into the LHC. A diagram of these successive accelerators is included in Figure 2.1.

Half the protons are fed into the LHC clockwise and half counterclockwise, into two beam pipes, where it takes about twenty minutes for the protons to reach their peak energy of up to 7 TeV (in 2012, the peak energy was 4 GeV). At this point, the particles circumnavigate the LHC up to eleven thousand times per second. The protons in each beam are spaced out in up to 2808 groupings $7 - 14$ meters apart called "bunches".[3] About 30 cm long, the bunches have a transverse area of about 1 mm, but are squeezed into a much smaller transverse area at the interaction points (IP)—about 16 μm in diameter.[4] When the proton bunches are crossed at each interaction point, as illustrated in Figure 2.2, a mean of about 20 collisions occur between 200 billion

---

[1]The amount of hydrogen used is so small - about two nanograms per day of operation - that it would take the LHC about one million years to accelerate one gram of hydrogen.

[2]LINAC2 is a linear accelerator that began operation in 1978 and uses radiofrequency cavities to charge cylindrical conductors through which the protons pass. The protons are accelerated by the push and pull of the conductors.

[3]In 2012, each proton beam had 1374 bunches, with each bunch containing about 1.48×10^{11} protons. The energy of one proton at 1 TeV is about as much energy of motion as a flying mosquito - but in a space about a million million times smaller.

[4]For comparison, a human hair is about 50 μm in diameter.
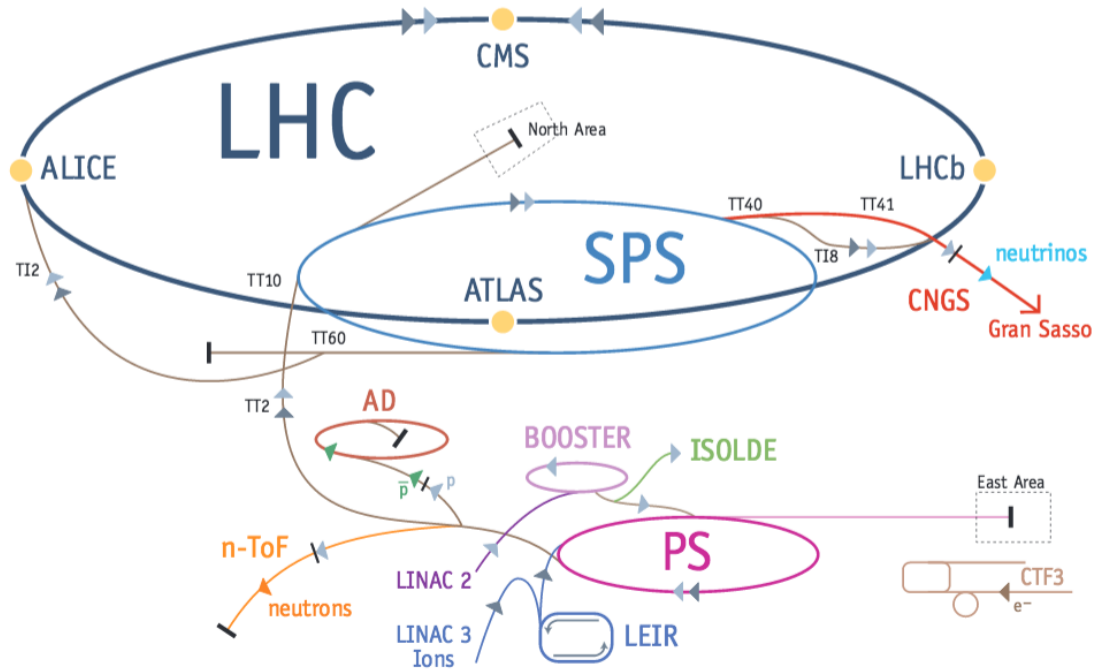
Figure 2.1: A schematic of the components of the LHC [28].

particles. The bunches cross every $25 - 50$ ns (in 2015, every 50 ns), which translates to $20 - 40$ million times per second and generates up to 600 million collisions per second [26].

The LHC magnet system engineered to accomplish this includes 1232 cryodipole magnets which produce a magnetic field of up to 8.33 T. These magnets use a dipole cold mass core, which runs at a temperature of 1.9 K, cooled by superfluid helium, and produces an electrical current of up to 11850 Amps.

Seven detectors spaced around the LHC record a fraction of these collisions at the four interaction points. The ATLAS (A Toroidal LHC ApparatuS) [29] and CMS (Compact Muon Solenoid) [30] Detectors serve as complementary general purpose detectors. The ALICE (A Large Ion Collider Experiment) detector [31], LHCb

Relative beam sizes around IP1 (Atlas) in collision
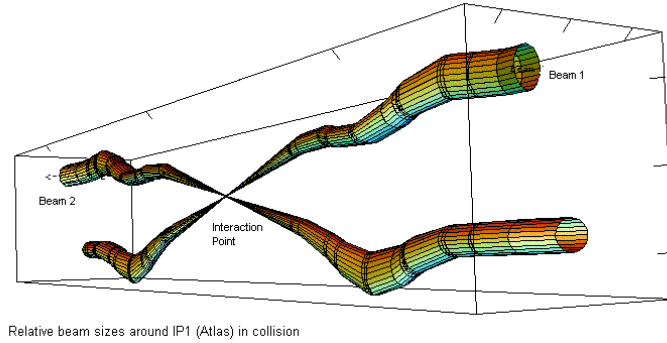
Figure 2.2: An illustration of the relative beam sizes and their manipulation around the interaction point (IP), at which the two beams of protons are collided [28].

(Large Hadron Collider Beauty) Detector [32], LHCf (Large Hadron Collider forward) Detector [33], TOTEM (TOTal Elastic and diffractive cross section Measurement) Detector [34], and MoEDAL (Monopole and Exotics Detector at the LHC) [35] each have more specialized areas of research. The ALICE experiment specializes in analyzing lead-ion collisions, studying the properties of quark-gluon plasma. The LHCb experiment specializes in collisions involving b quarks, and the LHCf experiment focuses on measuring particles produced traveling close to the beam pipe. TOTEM measures the cross section (effective size) of the proton, and MoEDAL is designed to search for magnetic monopoles and other highly-ionizing stable massive particles produced at the LHC. While the four largest experiments each have their own access point along the LHC, the TOTEM detector is installed near CMS and LHCf is installed near ATLAS; MoEDAL is installed in the LHCb cavern.

2.1.1 Luminosity

The number of a given type of events generated by the LHC every second is calculated by the probability of that event occurring, called the cross section, σ_{event}, and the luminosity L, defined as the number of particles that pass through a given area every second:

$$N_{\text{event}} = \sigma_{\text{event}} L \tag{2.1}$$

The machine luminosity for a Gaussian beam distribution is the result of several physical parameters [26]:

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F \tag{2.2}$$

in which N_b represents the number of particles per bunch and n_b the number of bunches per beam. The revolution frequency of the beam is given by f_{rev}, the relativistic gamma factor as γ_r, and the normalized transverse beam emittance ϵ_n. The factor β^* represents the beta function at the collision point, and F the geometric luminosity reduction factor due to the crossing angle at the interaction point, as illustrated in Figure 2.2:

$$F = 1/\sqrt{1 + \left(\frac{\theta_c \sigma_z}{2\sigma*}\right)^2} \tag{2.3}$$

The geometric luminosity reduction factor, assuming equal beam factors for both beams, depends on the crossing angle at the interaction point θ_c, the RMS bunch length σ_z, and the transverse RMS beam size at the interaction point σ^*.[5] From these relationships it is clear that the study of uncommon events at the LHC requires both high beam intensities and high beam energies.

---

[5]The design parameters for all variables in L and F are defined in Ref. [26].

2.1.2 The LHC Magnet System

To achieve the high luminosity goals of $L = 10^{34} \text{cm}^{-1}\text{s}^{-1}$ in the ATLAS and CMS experiments, the LHC uses two proton beams with separate vacuum chambers and magnet fields—the two beams use opposite magnet dipole systems to accelerate the beams in opposite directions. The beam pipes merge for approximately 130 meters around each interaction region (IR). Since the size of the LHC tunnel precludes the use of two rings of magnets, the LHC uses twin bore magnets, which have twin coil sets and beam channels controlled by the same mechanical structure and cooling system (cryostat). To provide peak beam energy of 7 TeV, the LHC magnets are designed for peak magnetic dipole field of 8.33 T through superconducting magnet technology.

2.2 The ATLAS Detector

The ATLAS Detector is one of the two general purpose detectors of the LHC. Along with the CMS detector, the ATLAS Detector was designed to cover a broad spectrum of particle physics research, including the search, discovery and investigation of the Higgs boson, tests of the Standard Model, and the search for SUSY and other physics beyond the Standard Model such as extra dimensions, dark matter, and dark energy. The ATLAS and CMS detectors have strategically different, complementary technological systems, with different magnetic field configurations for measuring the muons produced in collisions. The CMS detector uses a solenoid arrangement, in which the magnetic field is parallel to beam axis, in contrast to the air-core toroid magnets of the ATLAS Detector that create a magnetic field azimuthal to the beam axis. (The ATLAS

Detector magnet system also includes a solenoid magnet to provide the magnetic field for the inner tracking system, while CMS uses the same magnet system as for the muon detection.) The CMS and ATLAS magnet systems are the two largest of their kind, storing up to three GigiJoules of energy. In addition, the ATLAS and CMS detectors each use one of two electromagnetic calorimeter systems available during the design and construction phase that could perform in a high-radiation environment with enough electron and photon energy resolution to detect the decay of the Higgs boson to two photons. The CMS system uses dense lead tungstate scintillating crystals, while the ATLAS system uses a lead-liquid argon sampling calorimeter. In both technologies, electrons and positrons coming from electromagnetic showers excite atoms (in lead tungstate) or ionize atoms (in liquid argon) proportionally to the energy of the electrons and photons.

The ATLAS Detector weighs about 7000 tons, is forty-four meters long and twenty-five meters in diameter. Cylindrical in shape, it is comprised of smaller detectors arranged in onion-like layers progressing outward at different radii from the beam pipe, each layer optimized for a specific purpose. (Each layer of the detector also has an endcap part placed perpendicular to the axis.) As described in the following sections and shown in Figure 2.3, these systems consist of an inner detector to measure the trajectory and momentum of each charged particle, a solenoid superconducting magnet to bend charged particles in the inner detector, calorimeters to measure the energy carried by different particles, the air-core toroid magnet system to bend the muon paths, and a muon spectrometer to identify and measure the momenta of muons.
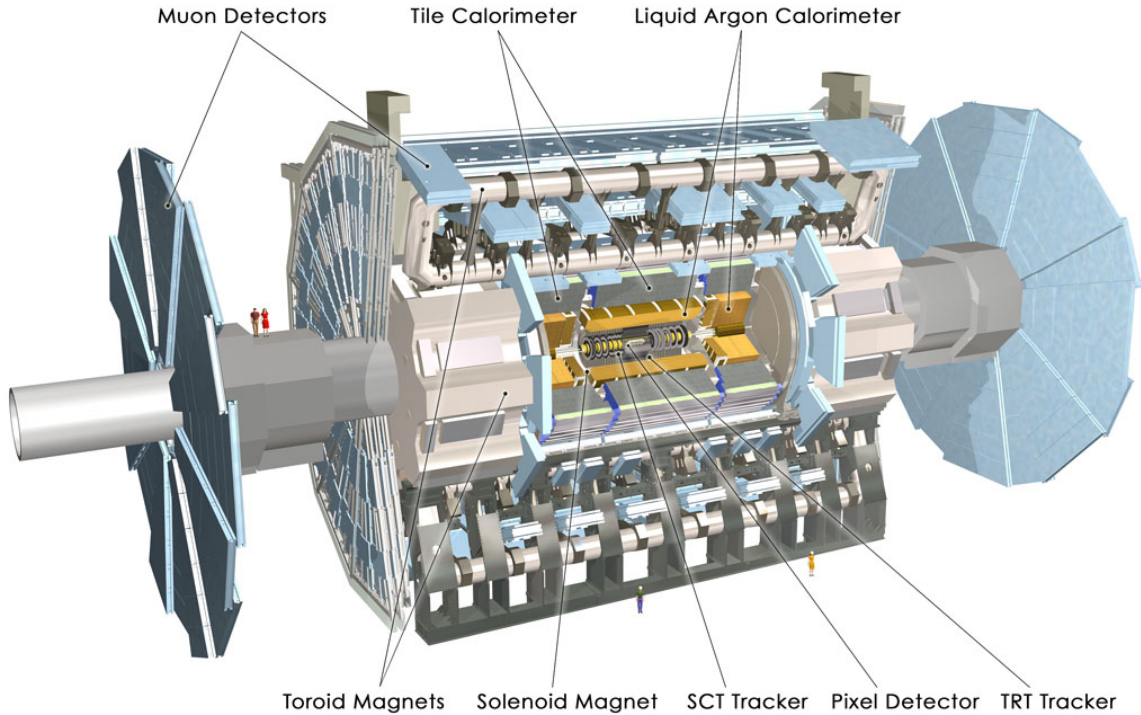
Figure 2.3: A cut-away image of the ATLAS Detector, with each major subdetector system labeled. The detector is 44 meters in length and 25 meters tall [29].

2.2.1 ATLAS Coordinate System

The ATLAS Detector is defined according to a right-handed Cartesian coordinate system as shown in Figure 2.4. The origin is defined to be at the nominal interaction point, with the beam pipe in the z direction and the xy plane transverse to the beam. The positive x axis points toward the center of the LHC ring, and the positive y direction approximately vertical (away from the earth's core).[6] The angle from the positive x axis is defined as ϕ, and the angle from the positive z axis as θ. The amount

---

[6]Due to a tilt in the LHC ring to minimize construction cost, the y axis is slightly displaced from vertical.

of energy, momentum and other properties in the xy plane are defined as transverse, such as transverse momentum $\vec{p}_\mathrm{T} = (p_x, p_y)$:

$$p_\mathrm{T} \equiv \sqrt{p_x^2 + p_y^2} \tag{2.4}$$

a convenient quantity since the initial state has zero (or nearly zero) transverse momentum. The transverse energy is also a useful quantity:

$$E_\mathrm{T} = E \sin\theta \tag{2.5}$$

The angular location is often measured by its pseudorapidity, which is useful because the pseudorapidity separation between two particles is Lorentz invariant:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \tag{2.6}$$

However, rapidity is a more physical quantity for massive objects:

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right) \tag{2.7}$$

The angular separation ΔR of two physics objects is sometimes used to differentiate object selection, such as is described in Chapter 4. This is done by a measure of their separation in η and ϕ, to be Lorentz invariant with respect to boosts in the z direction:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{2.8}$$

2.2.2 Physics Requirements

The design of the ATLAS Detector is carefully optimized for the goals of discovering and studying the Higgs boson, testing the Standard Model, and searching for new
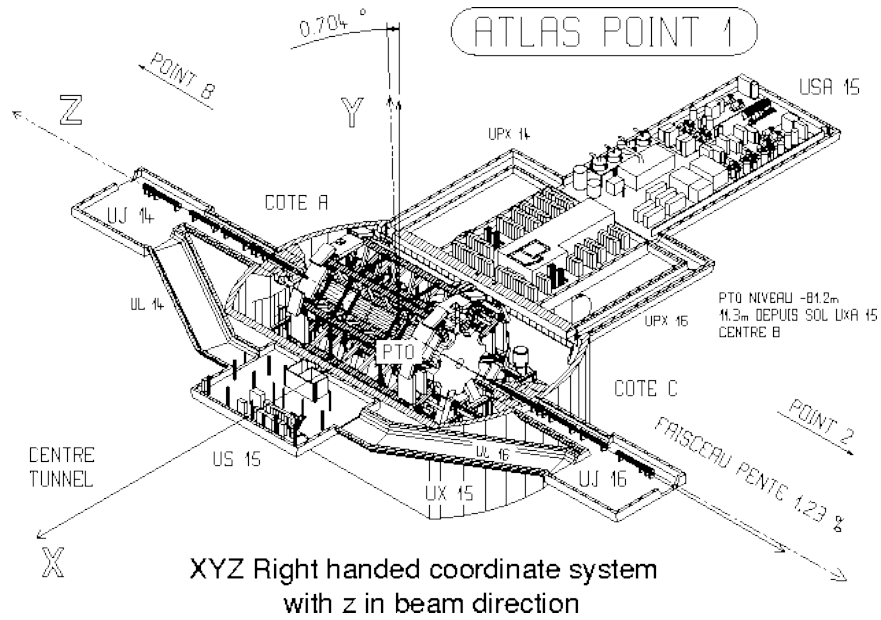
Figure 2.4: The ATLAS Detector uses a right-handed Cartesian coordinate system, with the z axis following the beam line and the xy plane is the plane transverse to the beam direction. The x axis points to the center of the LHC tunnel. The y axis is displaced from the vertical by $\sim 0.704°$ due to a slight tilt in the plane of the tunnel [36].

physics. The search for the Standard Model Higgs boson was used as a benchmark for system performance in the ATLAS Detector. For example, instrumental resolution should be able to resolve decay widths on the order of a few MeV (in the case of a low mass Higgs boson $m_H < 2m_Z$). Decay modes into hadrons require differentiation from QCD backgrounds. New physics could include heavy gauge bosons Z' and W' which could be accessible up to ~ 6 TeV, requiring high-resolution measurements and charge identification of energetic leptons. Theories of new physics processes also predict high p_T jet measurements. The decay of supersymmetric particles such as squarks and gluinos could include the presence of a lightest stable supersymmetric particle (LSP), requiring the ability to measure transverse missing energy $E_\mathrm{T}^\mathrm{miss}$ since the LSP would interact only weakly in the detector. Models of extra dimensions accessible at the LHC predict the emissions of gravitons escaping into extra dimensions, generating $E_\mathrm{T}^\mathrm{miss}$, and the formation of miniature black holes which could evaporate into diverse final states of a mix of jets, leptons, photons, neutrinos, W's and Z's.

These described processes are generally rare, requiring the high luminosity of the LHC. However, these requirements produce their own difficulties; for example, QCD jet production from the proton-proton collisions can dominate and must be differentiated; an average of 23 inelastic collisions also accompany each main collision. In all, the ATLAS Detector must have radiation-hard, fast electronics and sensors with high granularity; large acceptance angles; and high momentum resolution and reconstruction efficiency of charged particles, particularly in the inner detector region. Very good resolution of electromagnetic processes, complemented by full-coverage

hadronic calorimetry, good muon recognition and momentum resolution, and highly efficient triggering systems are all key to achieving these physics goals. The following sections discuss the specifics of each subsystem.

2.2.3   Inner Detector

The Inner Detector (ID) [29] lies closest to the beam pipe, and uses three different technologies to track charged particles leaving the collision point, measure their momenta, and determine the precise location of the primary collision vertex, as well as to identify jets coming from b quarks. Because of its proximity to the beam pipe, the ID also faces the challenge of absorbing the highest amount of radiation from the collisions, which can damage the detector and affect its operation. The three sections are arranged in layers as illustrated in Figure 2.5: a silicon pixel detector, a silicon microstrip detector (SCT), and a straw-tube transition radiation tracker (TRT).[7] All together, the ID surrounds the interaction point in the R-ϕ plane, and $|\eta| < 2.5$. Without pileup interactions, the ID can provide a resolution of up to 14 μm along the z axis and 18 μm in the transverse region for primary vertexes, and up to 0.7 mm radial resolution for secondary vertexes, depending on the η region. The ID can identify electron and pion tracks with p_{T} of approximately 5 GeV with efficiencies of 70-95%. The pixel detector [37] consists of advanced silicon technology containing over eighty million rectangular pixels capable of resolving particle positions to less than fourteen μm in the R-ϕ plane and 115 μm in z (barrel region) and R

---

[7]In Run Two of the LHC, an additional element was added closest to the beam pipe, the Insertable B-Layer (IBL).

(a)   (b)
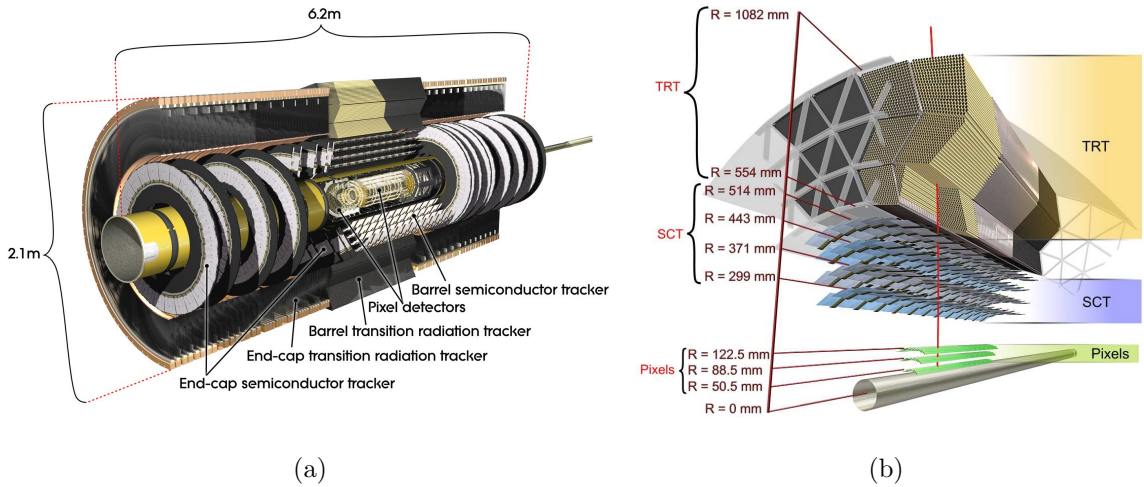
Figure 2.5: The inner detector system of the ATLAS Detector. Closest to the beam pipe, this system must withstand low temperatures and high irradiation [29].

(disks). The pixel detector is meant for precise track reconstruction in an environment of high particle density; so this part of the detector has the finest spatial resolution. In the barrel region, approximately 1500 modules are arranged from 5.05 to 12.25 cm from the beam pipe, covering the angular range $\eta < 1.7$; at the ends approximately 700 modules are arranged in disk layers, covering angular range $1.7 < \eta < 2.5$. Each pixel is 64.2 mm by 21 mm, containing 61440 pixel elements. Experimentally, the resolution depends on the momentum and incident angle of the particle; in 2011 the pixel detector operated at 99.8% efficiency [38] with a resolution of 15 μm in R-ϕ for tracks with $p_{\mathrm{T}} > 5$ GeV in 2010 pp collisions at $\sqrt{s} = 7$ TeV [39].

Although the pixel detector affords the highest precision, the prohibitively high cost of research, development and production necessitated diversification of technologies in the ID, keeping the pixel detector technology for the portion of the detector closest

to the beam pipe to optimize the need for precision. The semiconductor tracker (SCT) [29] uses layers of silicon microstrip sensors to track charged particles over an area of sixty square meters, with over six million readout strips. It is arranged in the same way as the pixel detector, with cylindrical layers in the barrel region and disks at the ends of the detector. It covers the entire R-ϕ space and $|\eta| < 2.5$ ($|\eta| < 1.4$ covered in the barrel region, and $1.4 < |\eta| < 2.5$ in the endcap region). The SCT is arranged from 29.9 to 51.4 cm distance from the beam pipe. The sensors are arranged as double-layered strips 80 μm x 6.4 cm so that a particle travels through at least eight layers (providing eight measurements per charged track). The SCT has a design resolution of 17 μm in the R-ϕ plane, 680 μm in z (barrel region) and R (disk region). It operated in 2011 at 99.6% efficiency [38].

The straw-tube transition radiation tracker (TRT) [29] is lowest resolution part of the ID but the most cost-effective, containing hundreds of thousands of "straws" tubes 4 mm in diameter filled with a Xenon-based gas mixture at high voltage with a wire down each tube axis. When a charged particle passes through a straw, it ionizes the gas, and then the voltage difference causes the freed electrons to travel to the wire and produce electrical pulses. From the timing of the pulses, the distance between the particle track and the wire can be calculated to within 0.17 mm. The Xenon also helps discriminate between hadrons and electrons. In addition, material between the tubes causes electrons passing through to produce x-rays, known as transition radiation, which can be used to help distinguish electrons from pions. The TRT is arranged 55.4 to 108.2 cm from the beam pipe within $|\eta| < 0.7$ (in the barrel region)

and $0.7 < \eta < 2.5$ (in the disk region). A typical charge leaves about 36 hits in the TRT. In the barrel region these tubes are 144 cm long arranged parallel to the beam axis in modules of 329-793 straws. In the disk region they are arranged radially, 37 cm long, as 18 wheel arrangements. Altogether the TRT has more than 350 thousand readout channels. In 2011 the TRT had a data-taking efficiency of 99.2% [38].

2.2.4 Magnet System: Solenoid Magnet

A superconducting solenoid magnet of diameter approximately 2.5 meters, length of 5.3 meters surrounds the inner detector, enclosing it in a 2 T magnetic field to be used to aid in measuring the momenta of charged particles. This magnet coil is made of aluminum, copper and niobium-titanium; it weighs 5 tons and contains 9 kilometers of superconducting wire coiled 1173 times and cooled by liquid helium. A current of 7.6 kAmps is run through the coil, producing a central field of 2 T (with a peak magnetic field of 2.6 T), which causes charged particles to bend as they move through the inner detector. The curvature of the tracks aids in measuring the momentum and electric charge of each particle.

2.2.5 Calorimeters

The next layer of the detector, the calorimeter, is dedicated to identifying and measuring the energy and direction of electrons, photons and jets, as well as missing transverse momentum. Along with the curvature of their path in the applied magnetic field, particles are identified by their interaction length (λ)[8] and nature of interactions

---

[8]The interaction length is the mean path length for a particle to lose $\frac{1}{e}$ of energy.

within the calorimeter. Since electrons and photons have shorter interaction lengths, the first section of the calorimeter is dedicated to these particles (the electromagnetic calorimeter). This is followed radially outward by the hadronic calorimeter. Both regions of the calorimeter cover nearly the full solid angle in the detector, $|\eta| < 4.9$, and are illustrated in Figure 2.6.

Electromagnetic Calorimeter The electromagnetic calorimeter captures and records the trajectory and energy of electrons and photons with closely spaced layers of lead and liquid argon (LAr) in the region $|\eta| < 4.9$. The electrons and photons shower and produce ions in the LAr, causing electrical charges to be read out. Electrons of higher than ~ 100 MeV lose energy through Bremsstrahlung radiation: they emit photons, and the photons decay through pair-production and are measured by the collection of charges at the Kapton electrodes. This decay process is collectively known as photon-electron cascades. The energy of the photons and electrons can be measured by the size of cascade they cause in the LAr. The lead layers are used to slow down and capture both electrons and photons.

The electromagnetic calorimeter is comprised of several sections to cover the η region: the barrel region extends through $|\eta| < 1.475$, made up of two half-cylinders with a small gap of 6 mm between them. Two end cap regions cover the region of $1.375 < |\eta| < 3.2$ by the combination of two wheels positioned at $1.375 < \eta < 2.5$ and $2.5 < \eta < 3.2$. Because LAr is used as the active medium, the electromagnetic calorimeter is operated at 88.5 K through the use of a cryostat system.

Hadronic Calorimeter   The hadronic calorimeter works on the same principles of absorption and showering, but the average nuclear interaction length determines the longitudinal length. As the interaction length of hadrons is much longer than that of electrons and photons (and to prevent hadrons from passing through into the muon spectrometer), the hadronic calorimeter is much thicker than the electromagnetic calorimeter, at nearly two meters (radial distance from the beam pipe of 2.28 to 4.25 meters). It covers $|\eta| < 4.9$ through the combination of three regions: the barrel region, or tile calorimeter, two endcap regions, and two forward regions. In the z direction, the hadronic calorimeter is 11 meters long in order to allow for 10 interaction lengths (λ) in every direction of η.

In the barrel region, the hadronic calorimeter absorbs particles that survive the electromagnetic calorimeter with alternating layers of steel and tiles of scintillating plastic that emit light when charged particles move through it. Photomultiplier tubes convert these light pulses from optical fibers to electronic signals, which are read out by electronics systems. This region covers $|\eta| < 1.7$ in three segmented barrel regions: one main region ($|\eta| < 1.0$) and two extended regions ($0.8 < |\eta| < 1.7$). Each region is composed of 64 modules arranged in three layers. The first two layers have a granularity of 0.1×0.1 ($\Delta\phi \times \Delta\eta$); the third 0.2×0.2.

Because the endcap and forward regions of the hadronic calorimeter encounter higher radiation, these regions also use LAr as scintillating material and are operated at 88.5 K with the cryostat system. Each endcap calorimeter is comprised of two wheels $1.5 < |\eta| < 3.2$, with copper plates for the absorbing material and LAr as the
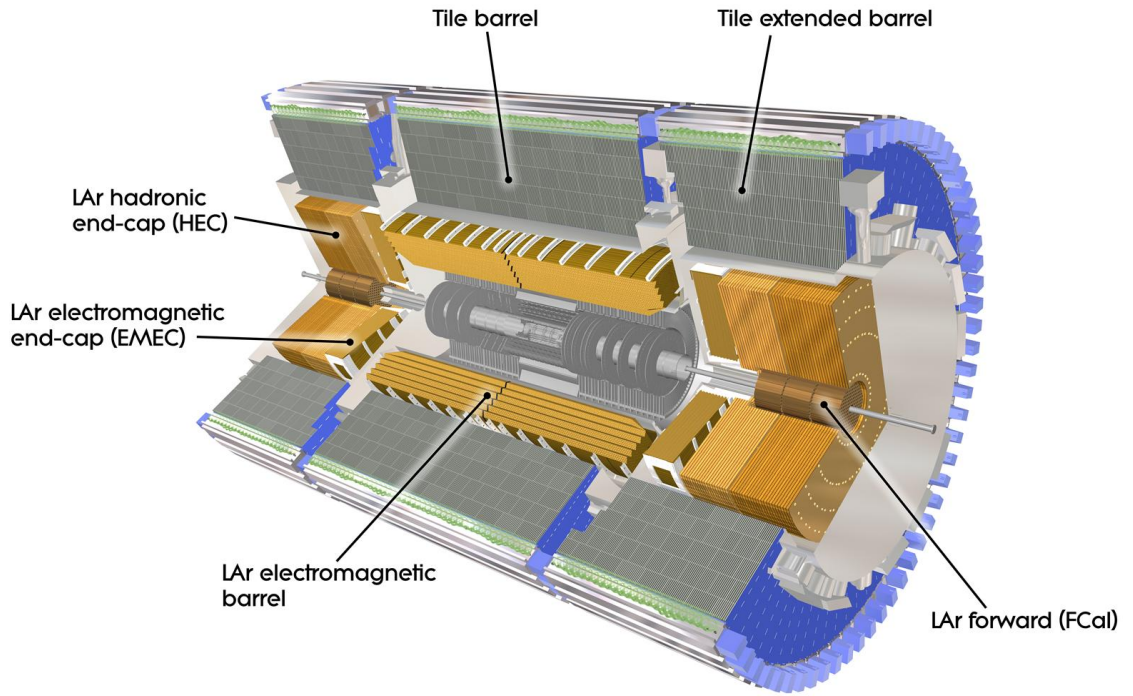
Figure 2.6: A cut-away image of the ATLAS calorimeter system [29].

scintillating material. Each wheel is made of 32 modules, with a 2 mm gap between each module. The first of the two wheels covers $1.5 < |\eta| < 2.5$ with 25 mm copper plates, enabling a granularity of 0.1×0.1. The second wheel uses 50 mm copper to cover $2.5 < |\eta| < 3.2$ with a granularity of 0.2×0.2.

The forward region calorimeters cover the region $3.1 < |\eta| < 4.9$ at a granularity of 0.2×0.2 using LAr. This region has one section of copper used for absorption and two layers of tungsten.
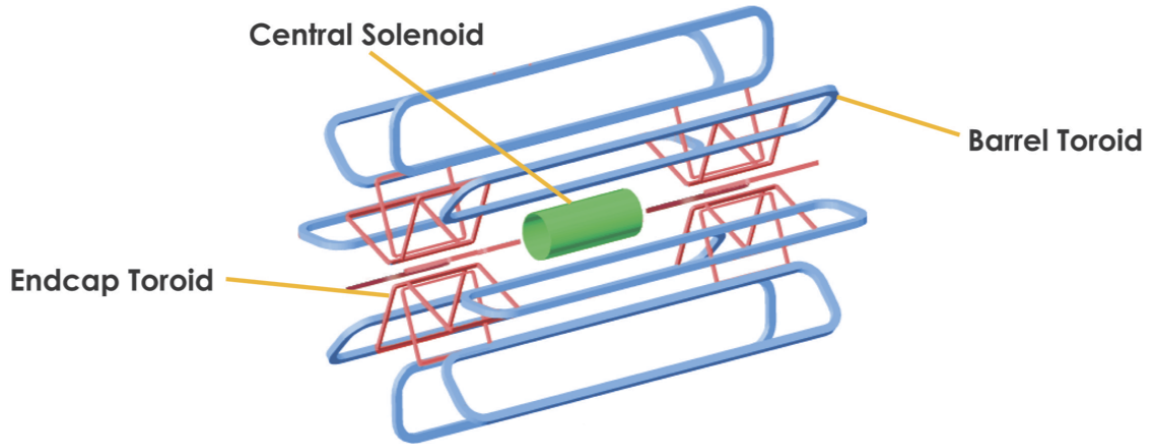
Figure 2.7: The ATLAS Detector superconducting magnet system. The solenoid magnet creates a magnetic field parallel to the z axis to track charged particles in the Inner Detector. The toroid magnets create a magnetic field azimuthal to the z axis to track muons [29].

2.2.6 Magnet System: Toroid Magnets

As stated above, the ATLAS Detector superconducting magnet system includes one solenoid magnet around the Inner Detector and three air-core toroid magnets around the muon spectrometer around the perimeter of the detector (one barrel toroid and two end-cap toroids). The toroid magnets are configured as in Figure 2.7 to create a magnetic field azimuthal to the beam pipe, perpendicular to the typical muon path. The magnetic field created by the toroid magnets bends the paths of muons (and any other charged particles' paths that are not stopped before the muon spectrometer) as they travel away from the interaction point, proportionately to the muon's mass and velocity for identification and momentum measurement. The toroid magnets, as the solenoid magnet, is comprised of aluminum, copper and niobium-titanium

(although with a higher aluminum composition). It contains about 80 kilometers of superconducting cable, comprising eight 25 meter long superconducting barrel coils in a cylindrical shape around the detector, as well as two endcap systems at either end of the detector. The magnets create a field of maximum 3.9 T in the barrel and 4.1 T in the disks. The endcap disk section is rotated 22.5 degrees relative to the barrel section to provide radial overlap and optimize bending powers.The barrel and endcap toroids have 8 coils (120 and 116 turns each), placed radially and symmetrically around the z axis, respectively.

The magnets are cooled with liquid helium at 4.5 K using a forced flow technology, as well as (in the case of the outer magnet system) cold helium pumps.

2.2.7   Muon Spectrometer

At least five meters from the interaction point, the muon spectrometer is the part of the detector farthest from the beam pipe, since muons travel through the entire detector, only weakly interacting in the calorimeter, while nearly all other particles except neutrinos are designed to be stopped before reaching the muon detector system. The muon spectrometer measures the charge and momentum of muons in the region $|\eta| < 2.7$ with a system of precision tracking chambers, in coordination with the toroid magnet system which bends the muons' path by the application of a magnetic field.

The spectrometer is made up of three cylindrical layers in the barrel region at 5, 7.5 and 10 meters from the beam pipe, and four large disk wheels on the ends of the detector at 7.4, 10.8, 14 and 21.5 meters. These are placed in and around

the toroid magnet coils and are made up of two different technologies. Most of the spectrometer is made up of Monitored Drift Tubes (MDT), gas-filled tubes with an anode on one end and a cathode on the other. When a muon travels through the gas, ionized electrons drift to the anode, initiating an electrical signal. This technology can determine the position of muons up to 80 μm. In the innermost layer of the disk region, where higher radiation backgrounds occur, a higher resolution technology is used called Cathode Strip Chambers (CSC). The CSC are thin arrays of parallel anode wires between narrow metal cathode strips and can measure muon positions with a design resolution of 40 μm in η and 5 mm in ϕ. To achieve these high resolutions in both MDT and CSC, alignment from data is augmented by twelve thousand optical alignment sensors that monitor position of each array and the deformation of the MDT chambers. In 2011, the muon spectrometer operated with an efficiency of 98.8-99.4%.

The ATLAS Detector also uses the presence of energetic muons in choosing some of the events to read out; to trigger these events, fast trigger chambers are interspersed in the muon system in the $|\eta| < 2.4$ region. Resistive Plate Chambers (RPC) are placed in the barrel region and Thin Gap Chambers (TGC) in the disk region. These chambers can be read out faster than the MDT and CSC and provide triggering capability for muons with p_{T} thresholds of $6 - 40$ GeV.

2.2.8   Triggering and Computing

The number of particle collisions per second at the center the ATLAS Detector is several orders of magnitude higher than the readout technology can record. In

addition, most of the collisions contain primarily particle interactions that are already well-understood; events with the potential to contribute new results are uncommon. Because of the combination of these factors, careful strategy is necessary to choose which collisions are recorded. This is handled through a three-tiered trigger system, in which certain properties of the collision are quickly and preferentially read out, and certain values of those properties "triggers" the recording of those collisions. Each tier of the trigger system refines the type of collisions that are recorded.

The first trigger ("Level 1 trigger") filters events based on detector information: high-momentum muons recorded in the muon spectrometer, high-momentum electrons and photons in the electromagnetic calorimeter, hadronic decays of τ's, and large amounts of missing transverse energy or large scalar sums of transverse energies recorded in the calorimeters are typically used as triggers. This reduces the number of events from about 1 GHz to less than 50 kHz in about 2.2 μs, or about 2000 times slower than the collision rate.

The second trigger ("Level 2") uses software restrictions to filter events that pass the Level 1 trigger based on variables such as the position, momentum, and total energy of recorded particles. To do this, high-granularity information is collected in particular regions of the detector (Regions of Interest, ROI) defined by the Level 1 trigger and the event reconstructed. This filter takes about 40 msec and reduces the event rate to about 5 kHz.

The third ("Level 3") trigger, or Event Filter (EF), filters events passing the Level 2 trigger using offline analysis procedures and more complex criteria to decide which

events are written out to permanent data storage in about 4 seconds, and the events filtered into corresponding physics streams. This stage filters the event rate to about 300 Hz.

The collisions selected by the triggering system are recorded and available for physics analysis, as described in the following chapters. See Appendix A for an highlight of the ATLAS Pixel Detector and the design challenges of irradiation damage throughout the lifetime of the detector.

Chapter 3

Data and Simulation Samples

"All of physics is either impossible or trivial.

It is impossible until you understand it, and then it becomes trivial."

– Ernest Rutherford

The two measurements described in the following chapters, the gap fraction measurement of Chapter 5 and top quark pole mass extraction of Chapter 6 (referred to for brevity as "gap fraction" and "mass extraction"), analyzed the complete pp collision data set recorded by the ATLAS Detector in 2011 and 2012, along with events generated by Monte Carlo (MC) simulations. This chapter describes the data sets and MC simulations.

The samples chosen closely follow those in the $2011 - 2012$ $e\mu+b$-tagged jets dilepton cross-section analysis [40], except that the gap fraction measurement uses the `NTupCommon` D3PD format [41, 42], to benefit from recent jet energy scale calibration improvements by the ATLAS `JetETMiss` group. Both data and MC samples were analyzed with the `AnalysisTop-1.8.0` software framework [43].

45

3.1 Data Samples

The gap fraction measurement was performed on the complete 2012 ATLAS data sample, and the pole mass extraction on the combination of the 2012 data sample and the 2011 ATLAS data sample. In the 2012 data sample, this includes data period designations A to L excluding periods F and K (which were not standard physics runs); the 2011 data sample includes periods B to M. Within these periods, events were selected that fulfill standard data quality requirements defined by the physics "All Good" run lists (GRL) [44, 45], which ensures that all the ATLAS detector systems crucial to reconstructing electrons, muons, jets, and other physics objects are fully functional. The data for the gap fraction measurement was accessed using the `NTupCommon` D3PD format [41], while the mass extraction was accessed with the `NTupTop` D3PD format. Both analyses used the processing tags recommended by the ATLAS top physics group [41].

Events were collected if they passed either a single electron or single muon trigger chain (corresponding to the `Egamma` and `Muons` data streams, respectively), with thresholds on these triggers set to be maximally efficient for leptons passing offline selections of $p_T > 25$ GeV. To avoid double counting events that appeared in both streams, events passing both triggers were *only* included from the `Egamma` stream. That is, events were accepted from the `Muons` stream only if they appeared *only* in the `Muons` stream. Luminosity blocks present in only one of the two data streams were removed, to prevent bias caused by losing events in the upstream reconstruction or data processing. This removed just 133 of the 98,446 luminosity blocks of the 2012
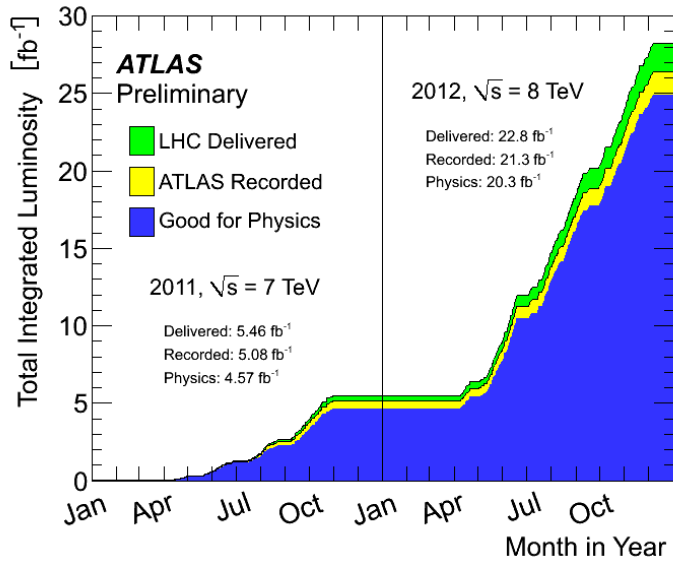
Figure 3.1: Integrated luminosity as a function of time delivered to the ATLAS Detector (green), collected by the ATLAS Detector (yellow), and certified as good-quality data (blue). The delivered luminosity includes luminosity during stable beam conditions for *pp* collisions [47, 48, 49].

data sample, corresponding to $\approx 0.2\%$ of the data; less than 0.1% of the 2011 data sample was removed for this reason.

The integrated luminosity of the full data sample as shown in Figure 3.1 was measured to be 20.28 fb^{-1} $\pm$ 2.8% in the 2012 data sample and 4.59 fb^{-1} $\pm$ 1.8% in the 2011 data sample [46].

3.2 Monte Carlo Simulation Samples

Both analyses included Monte Carlo simulations along with data measurements. In the gap fraction measurement, Monte Carlo event simulations were used to compare

with measurement from data at both particle and reconstruction level, to develop the correction factor, and to simulate event backgrounds and various systematic uncertainties. In the mass extraction, the underlying cross section measurement used Monte Carlo simulations to develop the analysis, compare with data, and evaluate signal and background efficiencies and uncertainties. In addition, Monte Carlo simulations were used in the mass extraction analysis to calculate the m_t dependence of the cross-section.

The samples were processed through the full ATLAS Geant4 [50] based detector simulation, with alternate samples through the `AtlFast2` [51] fast simulation. The simulated samples were re-weighted for agreement with data in a number of respects: the distributions of the average number of interactions per bunch crossing μ, and re-weighted on the order of unity to correspond to electron and muon reconstruction and trigger efficiencies, the width of the vertex distribution z, and b-tagging efficiencies. To vary the efficiencies for tagging b-jets in data, scale factors uncertainties derived from studies of b-jets containing muons using the $p_{\mathrm{T}}^{\mathrm{rel}}$ [52] and system-8 [53] methods were used. (The b-tagging scale factor was used in the gap fraction measurement but not in the mass extraction. Studies of $t\bar{t}$ events show that the b-tagging efficiencies are already well-modeled within uncertainties [54], and the uncertainties due to b-tagging are very small, as discussed in Chapter 5.)

The simulation samples chosen in both analyses were accessed through standard ATLAS top group D3PDs [55]. The gap fraction measurement samples were accessed with the `p1562` and `p1575` tags and processed with `AnalysisTop-1.8.0` software, using

the `TopFiducial` package for identifying particle level jets [56]. The mass extraction samples were accessed with the `p1400` and `p1401` (2012) and p937 (2011) tags and processed with `TopRootCore-14-00-01` (2012) and `TopRootCore-11-00-00-08` (2011) with some modifications in the lepton identification scale factors [57].

The list of generators and data set (DS) identification numbers used for $t\bar{t}$ and each background simulation sample in the gap fraction analysis is given in Table 3.1, and an exhaustive list of data set names given in Appendix B.3. The exhaustive list of samples used in the underlying cross section from which the pole mass was extracted is detailed thoroughly in [40]. In both analyses, the samples were normalized according to the recommended cross-section and K-factor values given in the `TopDataPreparation` package, tag `00-06-48` (gap fraction) and `00-05-09` (mass extraction).

Pair Production Simulation Samples

The primary simulation sample for both analyses was produced using POWHEG [58, 59] interfaced to PYTHIA6 [60] (PY) with the Perugia 2011C tune [61], and CT10 parton density functions (PDFs) [62], with the h_{damp} factor equal to infinity (DS 117050), referred to hereafter as POWHEG+PYTHIA. Additional simulation samples were produced with MC@NLO [63] interfaced to HERWIG [64] with JIMMY [65] for underlying event modeling, with the ATLAS AUET2 [66] and CT10 PDFs (DS 105200), and with POWHEG interfaced to HERWIG (DS 105860). Additional LO generators were used to generate samples such as ALPGEN [67] interfaced to HERWIG and JIMMY, with the CTEQ6L1 PDFs [20] (at $\sqrt{s} = 8$ TeV: DS $164440 - 164443$, 116108, and

49

116109; at $\sqrt{s} = 7$ TeV: DS $105890 - 105892$, $117897 - 117899$, $116108 - 9$), ALPGEN interfaced to PYTHIA (DS $201020 - 4$), and MADGRAPH interfaced to PYTHIA (DS 110872). In the gap fraction measurement, two additional variations generated by POWHEG interfaced to PYTHIA were included for comparison with $h_{\text{damp}} = m_t$, one using PYTHIA6 (DS 110404) and a second using PYTHIA8 (DS 117046).

While most of the simulation samples in the analysis are full simulations, fast-simulation (`fastsim`) versions of some samples (DS 117050, 110404, 117046, 105200, and 105860) are included in which interactions inside the calorimeter are generated using a parameterized model of the shower shape rather than full simulation [51]. These `fastsim` samples produce higher statistics for the same computing power than full-simulation (`fullsim`) samples, and were included for comparison, particularly for calculating the $t\bar{t}$ modeling uncertainty described in Section 5.5.5.

The effect of varying initial and final state radiation (ISR/FSR) was studied using two `fastsim` samples generated with the ACERMC+PYTHIA generator [68], and additionally in the gap fraction measurement with `CTEQ61` PDFs interfaced to PYTHIA 6, with two tunes based on `AUET2` to span the variations compatible with the ATLAS 2011 gap fraction analysis of additional jet activity in $t\bar{t}$ events at $\sqrt{s} = 7$ TeV [69] (DS 117209, 117210). Two additional sets of ISR/FSR radiation samples were included from the ALPGEN+PYTHIA generator (DS $201030 - 44$, $201040 - 201044$) and the MADGRAPH+PYTHIA generator (DS 110878, 110875).

A version of the MC@NLO+HERWIG sample containing only dileptonic $t\bar{t}$ decays (DS 110001 `fastsim`) was also used for evaluation of PDF uncertainties, to correspond

exactly to the event numbers in the PDF reweighting samples. This is discussed in Section 5.5.

The top quark mass was set to 172.5 GeV in all of the above-listed samples, and most of the simulated samples contained about five times as many events as data. In the mass extraction analysis, additional POWHEG+PYTHIA data sets used to vary the assumed mass were used: DS 117836, 117838, 117840, 117842, 117844 and 117846 for $m_t = 165$, 167.5, 170, 175, 177.5, and 180 GeV, respectively. (Both $\sqrt{s} = 7$ and 8 TeV use the same DS numbering.) The underlying cross-section measurement from which the top quark pole mass was extracted includes additional studies using Monte Carlo simulation samples, which are described thoroughly in Ref. [40].

Background Simulation Samples

Requiring both leptons and at least two b-tagged jets in the final state of the process $t\bar{t} \rightarrow W^{\pm}W^{\mp}bb \rightarrow e^{\pm}\nu e^{\mp}\nu bb$ eliminates most background events, producing a clean signal for investigating extra jet activity. The remaining backgrounds consist of two categories: events with two "prompt" leptons in which at least one originates from the decay of a Z or W boson that was part of the primary interaction, and events in which one or both of the leptons comes from a source other than a W or Z decay, categorized as "fake" or "misidentified" leptons. Each background was estimated from simulation (as listed in Table 3.2), with the normalization of the Z+jets simulation taken from data (following the method of the dilepton $t\bar{t}$ cross-section measurement at $\sqrt{s} = 7 - 8$ TeV [40]).

Events containing two prompt leptons include events in which a single top quark is produced along with a W boson (Wt), events in which a Z boson is produced along with associated jets (Z+jets) in which the Z boson decays leptonically, such as $Z \to \tau\tau\nu\nu \to e\mu\nu\nu\nu\nu$, and events producing two bosons (WW, WZ, ZZ) that both decay into leptons. The Wt background is the dominant background process in the analysis at approximately 2.9% of the total events, and is simulated using Powheg+Pythia with the CT10 PDFs and the Perugia P2011C tuning, using the "diagram removal" scheme (DS 110140) [70, 71]. An alternate sample using the "diagram subtraction" scheme was used for comparison in Section 5.5.7 (DS 110143) [70, 71].

Events producing a leptonically-decaying Z boson in conjunction with jets were modeled using Alpgen with CTEQ6L1 PDFs interfaced to Pythia 6 with the Perugia P2011C tune, including samples with $0 - 5$ additional light partons (DS $117650 - 117655$, $117660 - 117665$, $117670 - 117675$), samples with two charm quarks with additionally $0 - 3$ partons (DS $110805 - 110816$) and samples with two bottom quarks with $0 - 3$ additional partons (DS 110817-110828). As in the 2012 $t\bar{t}$ dilepton cross-section measurement, the heavy flavor overlap procedure (HFOR) [72] was used to avoid double-counting events in which the two charm quarks or two bottom quarks could be produced by either the matrix element or the parton showering. These samples include events with dilepton invariant mass of at least 60 GeV, with additional Alpgen+Herwig samples used for regions of dilepton invariant mass between 10 and 60 GeV (DS $146830 - 146835$, $146840 - 146845$, $146850 - 146855$). However,

the Z+jet background is subject to large theoretical uncertainties stemming from uncertainty in the Z+heavy flavour jet cross-sections. Because of this, the Z+jet events from simulation were scaled by the ratios of $Z \rightarrow ee$+2 b-jets and $Z \rightarrow \mu\mu$+2 b-jets measured in simulation and data, to account for mismodeling of heavy flavor jets produced with Z boson. This scaling factor is taken from the dilepton cross-section measurement [40] to be 1.13 ± 0.08.

Events producing two bosons were simulated with ALPGEN+HERWIG with up to three additional partons (DS $107100 - 107111$).

Background processes with at least one "fake" lepton come predominantly from $t\bar{t}$ decays in which only one of the top quark decays leptonically, and the second "fake" lepton comes from a source other than a W or Z decay. These "fake" leptons can come from a bottom or charm hadron, from an electron created by a photon interaction, a jet misidentified as a lepton, or a muon produced by a pion or kaon decay. As such, each simulation sample listed in Table 3.2 and the $t\bar{t}$ POWHEG+PYTHIA simulation was re-processed to select events that pass reconstruction-level selection but include at least one lepton not from a W or Z decay. Studies of this method done by Ref. [40] show that the numbers of fake lepton events with an opposite-sign $e^{\pm}\mu^{\mp}$ pair agrees with same-sign $e^{\pm}\mu^{\mp}$ pair results within a factor of two, and that same-sign event simulation agrees with the number of fake lepton events in data within approximately 20%. Since the results of this analysis find the same agreement between opposite and same-sign sample estimation of fake lepton contribution, as shown in Table 3.2, the fake lepton contribution was estimated from the opposite-sign sample. The scaling

using the data-to-simulation ratio of same-sign event counts was omitted since the opposite-sign estimation is well-within the background uncertainty of $\pm100\%$ of each background, discussed in Section 5.5.

Included in the category of mis-identified lepton backgrounds, additional near-negligible contributions come from t-channel single top production modeled with ACERMC interfaced to PYTHIA 6 (DS $117360 - 117363$) as well as events including a W boson and associated jets in which the W boson decays leptonically. These were simulated in the same way as the Z+jet events discussed previously, again using HFOR to remove overlap between the samples. The DS numbers and associated generators are listed in Table 3.2.

The mass extraction also required simulations of the Wt background with varying assumed m_t inputs. These were additional POWHEG+PYTHIA samples, with DS 110134, 110132, 110130, 110128, 110126, and 110124 for $m_t = 165$, 167.5, 170, 175, 177.5, and 180 GeV, respectively.

The following chapter discusses the selection of events from each data and simulation sample for the two analyses.

Sample Type	DS Number	Tag	Generator(s)
Signal Samples			
$t\bar{t}$ simulation	117050	p1562	POWHEG+PYTHIA
$t\bar{t}$ simulation	110404	p1575	POWHEG+PYTHIA ($h_{\mathrm{damp}} = m_t$)
$t\bar{t}$ simulation	117046	p1770	POWHEG+PYTHIA 8
$t\bar{t}$ simulation	105200, 110001	p1562	MC@NLO+HERWIG
$t\bar{t}$ simulation	105860	p1562	POWHEG+HERWIG
$t\bar{t}$ simulation	$164440 - 164443, 116108 - 116109$	p1562	ALPGEN+HERWIG
$t\bar{t}$ simulation	$201020 - 201024$	p1575	ALPGEN+PYTHIA
$t\bar{t}$ simulation	110872	p1575	MADGRAPH+PYTHIA
Radiation Variation			
$t\bar{t}$ + inc. ISR/FSR	117209	p1562	ACERMC+PYTHIA
$t\bar{t}$ + dec. ISR/FSR	117210	p1562	ACERMC+PYTHIA
$t\bar{t}$ + inc. ISR/FSR	$201030 - 201034$	p1575	ALPGEN+PYTHIA
$t\bar{t}$ + dec. ISR/FSR	$201040 - 201044$	p1575	ALPGEN+PYTHIA
$t\bar{t}$ + inc. ISR/FSR	110875	p1770	MADGRAPH+PYTHIA
$t\bar{t}$ + dec. ISR/FSR	110878	p1770	MADGRAPH+PYTHIA

Table 3.1: The $t\bar{t}$ simulation samples included in both analyses.

Sample Type	DS Number	Tag	Generator(s)
W+jets events			
$W \to e\nu$, $0-5$ additional partons	$117680 - 117685$	p1562	ALPGEN+PYTHIA
$W \to \mu\nu$, $0-5$ additional partons	$117690 - 117695$	p1562	ALPGEN+PYTHIA
$W \to \tau\nu$, $0-5$ additional partons	$117700 - 117705$	p1562	ALPGEN+PYTHIA
$Wb\bar{b}$, $0-3$ additional partons	$110801 - 110804$	p1562	ALPGEN+PYTHIA
$Wc\bar{c}$, $0-3$ additional partons	$126606 - 126609$	p1562	ALPGEN+PYTHIA
Wc, $0-4$ additional partons	$126601 - 126605$	p1562	ALPGEN+PYTHIA
Z+jets events			
$Z \to \tau\tau$, $0-5$ additional partons	$117670 - 117675$	p1562	ALPGEN+PYTHIA
$Zc\bar{c}$ ($Z \to \tau\tau$), $0-3$ additional partons	$110813 - 110816$	p1562	ALPGEN+PYTHIA
$Zb\bar{b}$ ($Z \to \tau\tau$), $0-3$ additional partons	$110825 - 110828$	p1562	ALPGEN+PYTHIA
low mass $Z \to \tau\tau$, $0-5$ additional partons	$146850 - 146855$	p1562	ALPGEN+HERWIG
Diboson events			
$WW \to \ell\ell\nu\nu$, $0-3$ additional partons	$107100 - 107103$	p1562	ALPGEN+HERWIG
$W \to (\text{incl.})Z \to \ell\ell$, $0-3$ additional partons	$107104 - 107107$	p1562	ALPGEN+HERWIG
$Z \to (\text{incl.})Z \to \ell\ell$, $0-3$ additional partons	$107108 - 107111$	p1562	ALPGEN+HERWIG
Single top events			
Wt (diagram subtraction scheme)	110143	p1562	POWHEG+PYTHIA
Wt (diagram removal scheme)	110140	p1575	POWHEG+PYTHIA
t channel single top with $W \to e\nu$	117360	p1562	ACERMC+PYTHIA
t channel single top with $W \to \mu\nu$	117361	p1562	ACERMC+PYTHIA
t channel single top with $W \to \tau\nu$	117362	p1562	ACERMC+PYTHIA

Table 3.2: The simulation samples from which estimations of background physics processes were made in both analyses. In addition, the POWHEG+PYTHIA sample was used to simulate background processes in which a $t\bar{t}$ pair is produced but only one top quark decays leptonically.

Chapter 4

Object and Event Selection

"Somewhere, something incredible is waiting to be known."

– Carl Sagan

The object and event selections of the following chapters follow the ATLAS top working group recommendations for $2011 - 2012$ data [73, 74], including specific recommendations for electrons [75] and muons [76], except with looser isolation cuts on electrons due to the clean final state. The choices were meant to optimize selection of well-defined $t\bar{t}$ events that decayed to an opposite-sign $e\mu$ pair, two neutrinos and b-tagged jets, closely following the definitions in the $2011 - 2012$ ATLAS $t\bar{t}$ dilepton cross-section measurement [40]. The reconstruction-level object selection for both analyses is outlined in Section 4.1, followed by the truth level selection for the gap fraction measurement in Section 4.2. The event selection at reconstructed level is discussed in Section 4.3 followed by the particle-level event selection for the gap fraction measurement in Section 4.4. Finally, additional considerations of b-jet ordering and jet matching in the gap fraction measurement are described in Sections 4.5 and 4.6,

respectively.

A precise definition of each selection level is necessary to delineate the stages of the analyses. "Reconstruction level" events or observables refers to those that have undergone reconstruction through the ATLAS software, sourced from both data and MC simulation. "Particle level" and "truth level" refer to events and observables from MC simulation in which direct knowledge is known about the objects' identity, rather than deduced from reconstruction. "Corrected-to-particle level" and correction for detector effects refer data events or observables that are corrected by comparison to MC simulation to be as equivalent as possible to the particle / truth level in MC simulation.

4.1   Reconstruction Level Object Selection

Selection of electrons, muons, jets, and the b-tagging of jets at the reconstruction level follow these constraints:

Electrons: Selected using the offline `tight++` identification, electrons were required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.47$, excluding the calorimeter transition region between the barrel and endcap calorimeters, $1.37 < |\eta| < 1.52$. The impact parameter of the muon longitudinally from the primary vertex was required to be within 2 mm, $z_0 < 2$ mm. Using the definition of angular distance ΔR given in Chapter 2, the calorimeter energy in a cone of radius $\Delta R < 0.2$ around the electron (excluding the electron's energy deposit) was required to be within `ETCone20`< 6 GeV, and the sum of p_{T} tracks in a cone of radius $\Delta R < 0.3$ (excluding the electron track) to be less than

6 GeV: `pTCone30`< 6 GeV. Additional kinematic-dependent cuts on these two isolation variables are included with the Egamma group tool `EisoTool2012` [77], corresponding to a 98% efficiency on choosing true prompt electrons. To prevent double-counting electron energy deposition as jets, jets within $\Delta R < 0.2$ of a reconstructed electron were removed. If the nearest surviving jet passing jet selection criteria was within $\Delta R < 0.4$ of the electron, the electron was discarded to ensure clean separation from nearby jet activity.

Muons: Muons reconstructed in both the muon spectrometer and inner detector were included by the `MuID` algorithm, with the requirement that they have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$, and to have the standard recommended qualifications of at least one pixel hit, at least 5 SCT hits, fewer than three holes in the pixel and SCT layers combined, and to fulfill the recommended ratio of outlier hits to central hits in the TRT:

$$n_{\mathrm{TRT\,Hits}} + n_{\mathrm{TRT\,Outliers}} > 5 \;\; \text{and} \;\; \frac{n_{\mathrm{TRT\,Hits}}}{n_{\mathrm{TRT\,Hits}} + n_{\mathrm{TRT\,Outliers}}} < 0.9 \qquad (4.1)$$

in the central region $0.1 < |\eta| < 0.9$. The muons were also required to satisfy $I^{\ell}_{\mathrm{mini}} < 0.05$, where the mini-isolation variable I^{ℓ}_{mini} represents the ratio of the sum of p_{T} of tracks in a variable-sized cone of radius $\Delta R = 10$ GeV$/p_{\mathrm{T}}(\mu)$ to the p_{T} of the muon $p_{\mathrm{T}}(\mu)$ [73, 74]. The impact parameter of the muon longitudinally from the primary vertex was required to be within 2 mm, $z_0 < 2$ mm. The muons were required to be separated cleanly from jets, such that any muon within $\Delta R < 0.4$ of a jet passing the reconstructed jet selection criteria was rejected, in order to reduce background from muons from heavy flavor decays inside jets. In 2011 data for the pole

mass extraction, standard fixed cone isolation cuts of `ETCone20`< 4 and `pTCone30`< 4 were applied instead of the mini-isolation variables, which were not available.

Jets: Jets were accepted if they had $p_\mathrm{T} > 25$ GeV and $|\eta| < 2.5$, after being reconstructed using the anti-k_t algorithm with radius parameter $R = 0.4$ [78, 79]. In 2012 data this was done with the local cluster reweighting (LCW) method, corrected for pileup effects using jet area and a residual correction derived from both the instantaneous luminosity and the number of vertices in the event [80]. In the 2011 data, no jet area correction was applied; clusters were calibrated at the electromagnetic scale. These methods follows the recommendations of the ATLAS Top working group. No jet energy corrections for semileptonic b-hadron decays were applied. Reconstructed jets within $\Delta R < 0.2$ of an electron were also removed, as discussed in the electron selection definitions, to prevent double-counting electron energy deposition as jets.

To reduce contribution from additional pp collisions, jets of $p_\mathrm{T} < 50$ GeV and $|\eta| < 2.4$ were also required to have $|\mathrm{JVF}| > 0.5$ (2012) and $|\mathrm{JVF}| > 0.75$ (2011), where JVF is measured as the ratio of the sum of p_T of tracks associated to the jet and the primary vertex, to the sum of the p_T of all tracks associated with the jet:

$$\mathrm{JVF} = \frac{\Sigma_i^{\text{tracks from vertex}} p_\mathrm{T}^i}{\Sigma_j^{\text{all tracks}} p_\mathrm{T}^j} \tag{4.2}$$

Jets with no associated tracks or in the region $|\eta| > 2.4$ were assigned a JVF value of -1 by convention and so accepted.

b-**tagging:** Jets were b-tagged with the MV1 algorithm [81, 52], which combines output from a number of algorithms into a multivariate discriminant variable with values between zero and one. With this metric, the closer an MV1 value is to unity,

the more likely that the jet came from a b quark. This analysis considered jets to be b-jets if their MV1 value was above 0.7892, which corresponds to approximately 70% efficiency of identifying b-jets.[1]

4.2 Particle Level Object Selection

Particle-level object selection was used in the gap fraction measurement to correct the data back to particle level and for comparison to data after correction. Particle-level objects selected in Monte Carlo simulations were processed using the `TopFiducial-00-00-09` package, and the cuts made to replicate the reconstruction level cuts as closely as possible. The `TopFiducial` package is used rather than the `AntiKt4Truth` definition to identify particle-level jets in the Monte Carlo simulations, since the `AntiKt4Truth` jets exclude muons and neutrinos, while the `TopFiducial` jet reconstructions include muons and neutrinos not originating from W and Z bosons [56]. The `TopFiducial` package optimizes correspondence with the jet energy scale defined by the ATLAS `JetETMiss` group, and standardizes the jet energy scale used across analyses.

Electrons: Electrons were required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$, to match reconstruction-level selection but without the constraint of the calorimeter region. They were required to be at least $\Delta R > 0.4$ from the nearest jet passing particle jet selection and to have come from a W boson. The four momenta of the leptons were "dressed" by adding the momentum of all radiated photons with a cone of $\Delta R < 0.1$

---

[1]The exact efficiency is a function of several variables, predominately p_{T}.

around the lepton, to account for final-state radiation from the lepton.

Muons: As at reconstruction level, muons were required to have $p_\text{T} > 25$ GeV and $|\eta| < 2.5$ and to be at least $\Delta R > 0.4$ from the nearest jet passing particle jet selection. They were required to have come from a W boson, and as the electron, the four momenta of the muons were "dressed" by adding the momentum of all radiated photons with a cone of $\Delta R < 0.1$ around it, to account for radiation from the lepton.

Jets: To match reconstruction level, jets were required to have $p_\text{T} > 25$ GeV and $|\eta| < 2.5$. The clustering of the jets was again calculated using the anti-k_T algorithm as defined by the `TopFiducial` software package.

b-**Jet Identification:** Particle-level jets were identified ("tagged") as coming from b quarks using the `TopFiducial-00-00-09` package definitions, which determines the particle-level flavor of jets by associating b-hadrons with $p_\text{T} > 5$ GeV to the jets using a ghost matching procedure [82]. Jets with an associated b-hadron are labeled as b-tagged jets.

4.3   Reconstruction Level Event Selection

Reconstruction level events from simulation and data were selected according to the following process. Events from data could be selected by either the electron trigger or the muon trigger (corresponding to the `Egamma` and `Muons` data streams), as discussed in Chapter 3. Events were required to have at least one primary vertex with at least five associated tracks, and to pass the ATLAS Top Group standard jet cleaning and quality cuts [83].

Events from data were vetoed if they were selected from the muon trigger but also had an electron trigger, or if they had no electron trigger. Events from data were also required to originate from a luminosity block present in both the electron and muon streams. Both simulated events and events from data were then required to pass cosmic event, $e\mu$ overlap, and jet cleaning cuts. They were then required to have exactly one electron and one muon of opposite charge, each lepton meeting the object selection requirements defined in Section 4.1. Finally, events were required to contain at least two b-tagged jets passing the jet selection requirements.

The resulting event counts at reconstruction level for the gap fraction measurement are shown in Table 4.1. The event counts in the underlying cross section from which the top quark pole mass was extracted are listed in Ref. [40].

This selection process followed the definitions described in the dilepton cross-section analysis [40], with two exceptions. First, the overlap cut for the two leptons was relaxed for this analysis to the ATLAS Top Group standard of $\Delta\Phi_{e,\mu} < 0.005$ and $\Delta\Theta_{e,\mu} < 0.005$, where Φ and Θ are the azimuthal and polar angles of the leptons, respectively. This restriction is used to exclude muons that underwent catastrophic bremsstrahlung/energy loss in the calorimeter, which can leave a large energy deposit and potentially be reconstructed as both an electron and a muon. Secondly, this analysis requires *at least* two b-tagged jets rather than *exactly* two b-tagged jets.

	$e\mu$	(%)	≥ 2 b-jets	(%)
Data	70854		12437	
Total simulation	66151	100.0	12410	100.0
$t\bar{t}$	40261	60.8	11947	96.3
Wt single top	3836	5.8	359	2.9
$Z(\to \tau\tau \to e\mu)$+jets	12825	19.4	6	0.1
Dibosons	8030	12.3	2	0.0
Misidentified leptons	1199	1.8	96	0.8

Table 4.1: Number of events with a final-state opposite-sign $e\mu$ pair, $e\mu$ with at least two b-tagged jets, and $e\mu$ with exactly two b-tagged jets in the 2012 $\sqrt{8}$ TeV pp collision data, compared with expected counts and statistical uncertainty from simulation. The simulation event counts are normalized to the data luminosity of 20.3 fb^{-1} and are broken down into contributions from $t\bar{t}$ events, single top events, Z+jet events, diboson events and mis-identified lepton events.

4.4   Particle Level Event Selection

In the gap fraction measurement, particle level events are selected according to the following process. Events are required to have exactly one electron and muon, each coming from a W boson (either directly or via a τ decay) and passing the p_T and $|\eta|$ cuts. The same overlap cuts are applied as at reconstruction level: first, muons within $\Delta R < 0.4$ of a jet passing the p_T and η selection are removed. Then, the closest jet to the electron is removed if it is within $\Delta R < 0.2$. Finally, electrons are removed if within $\Delta R < 0.4$ of a jet passing the p_T and η selection. Events were eliminated if the electron and muon share the same track, defined as $\Delta \Phi < 0.005$ and $\Delta \Theta < 0.005$. Finally, events were required to include at least two b-jets.

4.5   Additional b-Jets

In $e^{\pm}\mu^{\mp}bb$ events containing three or more b-tagged jets in the gap fraction measurement, it is important to explicitly define the method of choosing which two b-tagged jets to consider the top quark decay products. At both the particle and reconstruction level, the two b-tagged[2] jets with the highest p_T are considered to be from the top quarks, and any additional lower p_T b-tagged jets are eligible to be additional jets in the event. Selecting the two with highest MV1 value instead would change the resulting gap fraction by at most 0.7%: since less than 3% of $e^{\pm}\mu^{\mp}bb$ events have three or more b-tagged jets, and of these events, only about half of the b-tagged jets are

---

[2]Or, at particle level, b-jets. For simplicity, this section refers to particle-level b-jets and reconstruction level b-tagged jets both as b-tagged jets.

affected by the ordering choice, the ordering choice makes only a small change overall. For this analysis, it is advantageous to use p_T ordering to select the b-tagged jets in order to match the selection process at particle and reconstruction level as closely as possible. Figures showing the number of b-tagged jets in the events, and the difference in the gap fraction from simulation using the two ordering methods are included in Appendix B.1.

4.6   Jet Matching

Also in the gap fraction measurement, the process of jet matching using geometric ΔR criteria quantifies which reconstructed jets likely came from which particle-level jets and which reconstructed jets likely came instead from pileup or jet splitting. It is not used in the standard object and event selection in the gap fraction measurement, but is used to examine the resolution of jet momentum before and after reconstruction and to determine the effect of additional pp interactions and jet splitting. In the case of jet matching, all selection criteria apply as stated previously, except that the particle-level jets are allowed to have $p_T > 10$ GeV to allow for smearing across the cut boundary. This process is executed as follows:

- The geometric distance ΔR is first measured between each particle-level jet with $p_T > 10$ GeV, including b-jets, and the reconstruction-level b-tagged jets. If the closest reconstruction-level b-tagged jet is within $\Delta R < 0.4$ of a particle-level jet, that particle-level jet is considered matched to that reconstructed b-tagged jet.

66

- If a particle-level jet is not within $\Delta R < 0.4$ of a b-jet, it is then compared to each additional reconstructed jet. If the closest reconstructed jet is within $\Delta R < 0.4$ of the particle jet, the particle-level jet is matched to that reconstructed jet. Reconstructed jets with no match to a particle jet are referred to as "unmatched jets" for the purposes of understanding additional pp interaction and jet splitting effects.

- Each particle-level jet can be matched to only one reconstructed jet, and each reconstructed jet can be matched to only one particle-level jet. This prevents phenomena such as a two reconstruction level jets that split from one truth level jet both being matched to the particle level jet.

Using this process, a reconstructed b-tagged jet is matched to a particle-level b-jet approximately 97.3% of the time, matched to a particle-level light jet 2.6% of the time, and unmatched just 0.1% of the time. (These statistics are the same using both MV1 ordering and p_T ordering of the b-tagged jets.) Using p_T ordering, approximately 1.4% of all additional reconstructed jets fail to be matched to a particle-level jet in the central region $|y| < 2.1$ (approximately 1% using MV1 ordering). This percentage remains constant when the p_T threshold of the particle-level jets used for matching is decreased to 5 GeV; increasing the p_T threshold increases the percentage of unmatched jets to 2.2% with a p_T threshold of 15 GeV, and 4.5% with a p_T threshold of 20 GeV.

Chapter 5

Measurement of $t\bar{t}$ Pair Production with a Veto on Additional Central Jet Activity in pp Collisions at $\sqrt{s} = 8$ TeV

"It was not a biologist, a doctor, or an astronomer who invented the microscope or the telescope, but grinders of lenses, who as far as we know were motivated only by idle curiosity or the desire for amusement."

– Martin and Inge Goldstein, *How We Know*

As discussed in Chapter 1, measurements of the top quark play a key role in both understanding the Standard Model and searching for new physics. To make precision measurements for these purposes, understanding each aspect of top quark interactions is essential. When produced in colliders such as the LHC, however, the top quark is often accompanied by QCD radiation that is nontrivial to model. modelling of quark and gluon activity in top quark events with state-of-the-art Monte Carlo event generators has achieved moderate accuracy, but limitations of the modelling's agreement with data cause systematic uncertainties that have contributed significant uncertainty in recent measurements. A careful measurement of jet activity in $t\bar{t}$ pair

production events in data, and a comparison with the latest MC event generators, can enable improve understanding of QCD radiation in top quark events, tuning the QCD radiation modelling in MC generators, and reduction of the associated uncertainties. With these goals, this analysis presents a measurement of QCD radiation in $t\bar{t}$ events, using the $e\mu$ dilepton channel and correcting the result for detector effects.

This analysis measured events that included an $e\mu$ pair of opposite charge and at least two jets originating from b quarks, to select events including a $t\bar{t}$ decay to two leptonically decaying W bosons and two b quarks:

$$t\bar{t} \to W^{\pm}bW^{\mp}b \to e^{\pm}\nu b\mu^{\mp}\nu b \tag{5.1}$$

This decay channel can be measured with very little background process contributions, with any additional jets distinct from the b jets that comprise the signal.

The general method for the gap fraction analysis has been defined by previous analyses such as the ATLAS 2011 gap fraction measurement [69], which have been useful for generator tuning and comparison exercises [84]. This analysis improves upon the results of the ATLAS 2011 analysis with lower statistical and systematic uncertainties, particularly with lower uncertainties contributed by the jet energy scale (JES) and jet vertex fraction (JVF). This analysis also includes an additional examination of the gap fraction plots measured in regions of invariant mass of the $e^{\pm}\mu^{\mp}bb$ system, $m_{e\mu bb}$.

The analysis method is defined in Section 5.1. The resulting distributions of transverse momentum, rapidity and pseudorapidity are described in Section 5.2, followed by an investigation of the jet energy resolution in Section 5.3. The calculation

of correction for detector effects is detailed in Section 5.4, and Section 5.5 defines the contributing systematic uncertainties. Finally, Section 5.6 gives the analysis results and Section 5.7 the conclusion.

5.1 Analysis Method

The gap fraction measurement was made using the dilepton $e\mu$ decay channel for $t\bar{t}$ events, so that jet activity from quark and gluon radiation could be more clearly distinguished from the decay products of signal events. In these events, one top (or anti-top) quark decays as $t \to W^\pm b \to e^\pm \nu b$ and the other with opposite charge as $t \to W^\mp b \to \mu^\mp \nu b$. Selected events were required to have one electron, one muon of opposite charge, and at least two b-tagged jets in the final state, referred to as $e^\pm \mu^\mp bb$ events. Any additional jets in the selected events, after removing pileup and background as well as possible by the JVF cut and standard jet quality cuts, were considered to be QCD radiation.

The final state electron was required to have transverse energy $E_T > 25$ GeV within the pseudorapidity region of $|\eta| < 2.47$, with $1.37 < |\eta| < 1.52$ excluded, and the muon to have transverse momentum $p_T > 25$ GeV within $|\eta| < 2.5$. Jets were required to have $p_T > 25$ GeV and $|\eta| < 2.5$. Full object and event selection details are given in Chapter 4.

Contributing backgrounds to this analysis were dominated by single top events, with additional contributions from diboson production, events with misidentified leptons, and Z+jet events (in which the Z boson decays as $Z \to \tau\tau\nu\nu \to e\mu\nu\nu\nu\nu$).

These were estimated with simulation, as described in Section 3.2.

The QCD radiation present in $t\bar{t}$ events was examined by measuring the gap fraction, f_{gap}, which is defined as the fraction of $e^{\pm}\mu^{\mp}bb$ events that do not contain additional jet activity in a defined rapidity or mass region (referred to generally as the "veto region") above a given transverse momentum (p_{T}) threshold. The transverse momentum threshold is defined in two ways, and the gap fraction in two ways accordingly. In the first, the gap fraction is measured with respect to the highest p_{T} of any additional jet:

$$f_{\text{gap}}(Q_0) = \frac{n(Q_0)}{N_{t\bar{t}}} \tag{5.2}$$

$N_{t\bar{t}}$ represents the total number of $e^{\pm}\mu^{\mp}bb$ events, according to the selection requirements defined in Chapter 4. Q_0 represents the p_{T} threshold for any one additional jet in one of these events (with the lowest threshold at 25 GeV), so that $n(Q_0)$ represents the subset of $e^{\pm}\mu^{\mp}bb$ events with no additional jet having $p_{\text{T}} > Q_0$.

The second form of the gap fraction equation uses Q_{sum} as the threshold for the scalar p_{T} sum of all additional jets:

$$f_{\text{gap}}(Q_{\text{sum}}) = \frac{n(Q_{\text{sum}})}{N_{t\bar{t}}} \tag{5.3}$$

In this form, $n(Q_{\text{sum}})$ represents the subset of $e^{\pm}\mu^{\mp}bb$ events in which the scalar p_{T} sum of all additional jets in the event with $p_{\text{T}} > 25$ GeV is less than Q_{sum}.

Both of these forms of the gap fraction were measured for four rapidity region definitions and four invariant mass regions, discussed below.

The gap fraction measurement was then corrected for detector effects as described in Section 5.4 and compared to simulated results at the corrected-to-particle level.

Concluding the measurement at this corrected-to-particle level enables simpler detector-independent comparison with Monte Carlo event generators.

5.1.1 Rapidity Regions

The gap fraction was measured considering additional jet activity in four different rapidity regions. These regions were chosen to encompass the central rapidity region of the detector (defined as $|y| < 2.1$), with the region divisions corresponding to previous gap fraction analyses: $|y| < 0.8$, $0.8 < |y| < 1.5$, $1.5 < |y| < 2.1$, and $|y| < 2.1$. Jet rapidity was calculated using the ATLAS experiment's right-handed coordinate system, with the z-axis defined in the direction of the beam pipe, and cylindrical coordinates (r, ϕ) describing the transverse xy plane:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \tag{5.4}$$

in which $p_z = p_\mathrm{T} \sinh(\eta)$. The pseudorapidity η is calculated as:

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] \tag{5.5}$$

where θ is the polar angle of the particle direction measured from the positive z-axis, and the transverse momentum p_T is defined as the momentum perpendicular to the LHC beam axis, $p_\mathrm{T} = p \sin \theta$.

5.1.2 Invariant Mass Regions

This analysis extends previous gap fraction measurements by also including the measurement for four different regions of the invariant mass of the $e^\pm \mu^\mp b b$ system—that

is, the $t\bar{t}$ system except the undetected neutrinos. This can provide an additional nuance of understanding QCD radiation, since higher mass values of this system are created dominantly by quark-antiquark interactions, while lower mass values are created predominantly by gluon-gluon interactions. The $q\bar{q}$ and gg events have different probabilities of getting extra jet radiation, so the simulation of extra jets for different mass regions is influenced by these different probabilities.

The invariant mass of the $e^{\pm}\mu^{\mp}bb$ system is calculated by the standard formula:

$$m_{e\mu bb} = \sqrt{(E_{e^{\pm}} + E_{\mu^{\mp}} + E_{b1} + E_{b2})^2 - (\vec{p}_{e^{\pm}} + \vec{p}_{\mu^{\mp}} + \vec{p}_{b1} + \vec{p}_{b2})^2} \qquad (5.6)$$

Figure 5.1 shows the distribution of the invariant mass in $e^{\pm}\mu^{\mp}bb$ events from data and various Monte Carlo simulations defined in Chapter 3. The downward trend in the ratio of distribution from data and from simulation is inspected in Appendix B. The four invariant mass regions (below 300 GeV, $300 - 425$ GeV, $425 - 600$ GeV, and above 600 GeV) were chosen to maximize the number of events that fit in the same category at both particle level and reconstruction level, to minimize smearing between the regions. To choose these regions, the invariant mass of each event passing both particle and reconstruction level selection described in Section 4 in the POWHEG+PYTHIA simulation (defined in Chapter 3) was measured, as shown in Figure 5.2, and the bin edges adjusted to maximize the number of events fitting in the diagonal bins, as shown in Figure 5.3. The maximum amount of events fitting in the diagonal bins, for four bins, was $\geq 84\%$. The additional generators agree within 2.8% for all bins (except within 5.1% for the highest mass bin), documented in Appendix B.
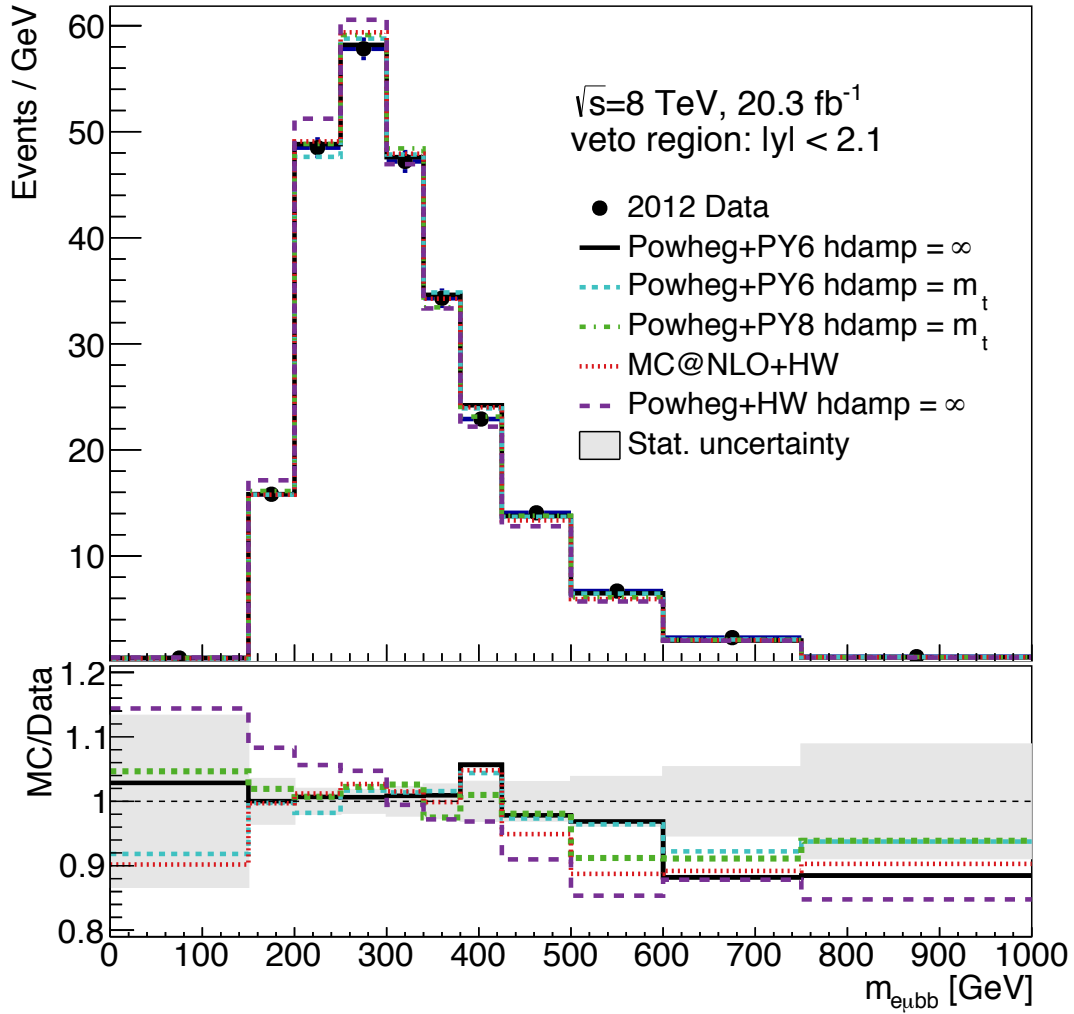
Figure 5.1: The distribution of mass of the $e^{\pm}\mu^{\mp}bb$ system at reconstruction level in data, compared to simulation by various NLO Monte Carlo generators. The grey shaded band represents the statistical uncertainty in data. The ratio plots show the ratio between the mass distribution in simulation compared to data.
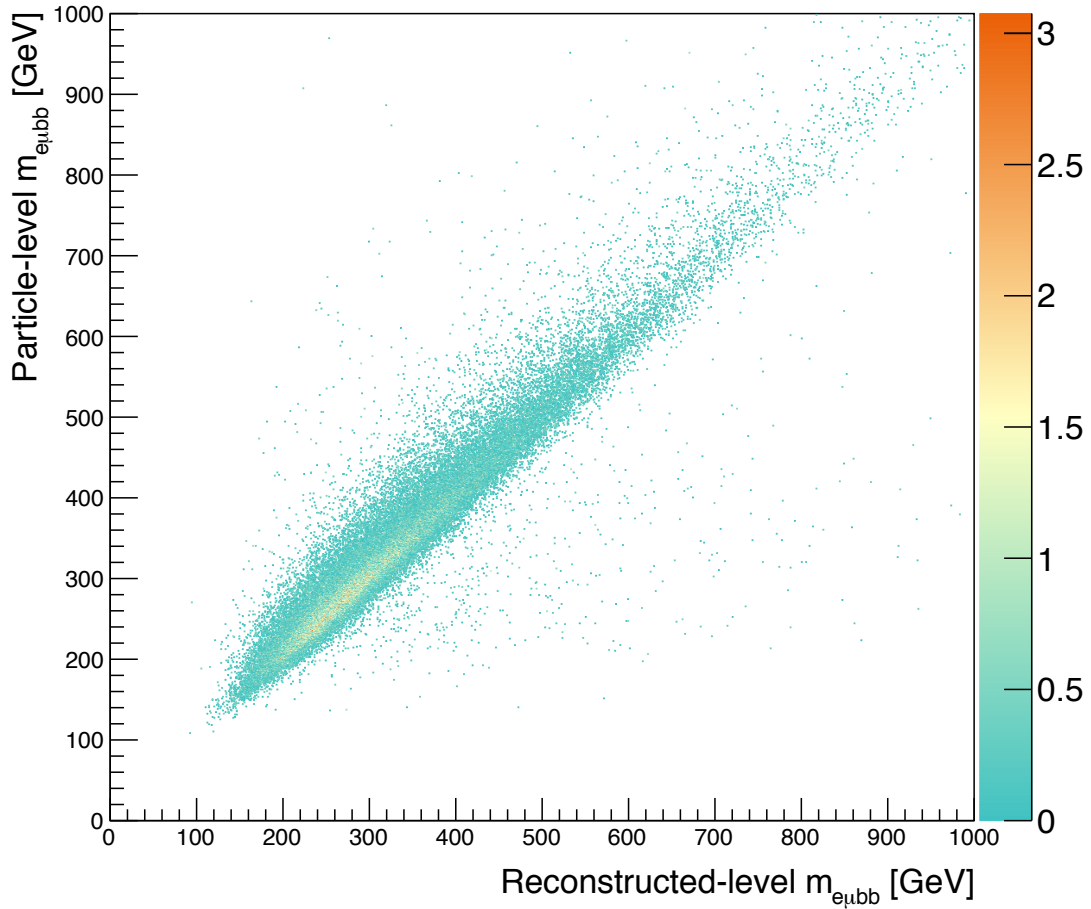
Figure 5.2: For all $e^{\pm}\mu^{\mp}bb$ events in the POWHEG+PYTHIA simulation that pass event selection at both the particle and reconstruction level, the invariant mass is constructed at both levels and the value plotted at the reconstruction level (x) versus the particle level (y).
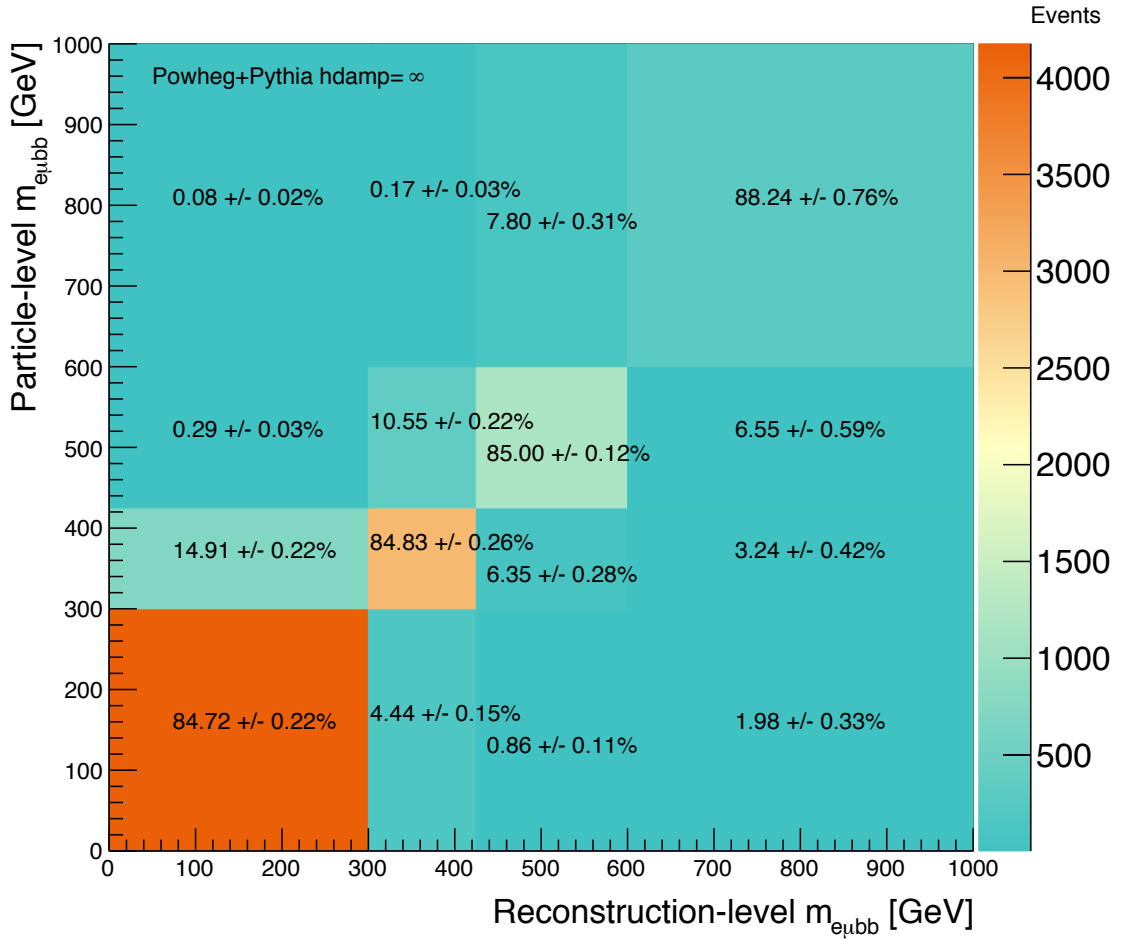
Figure 5.3: The bins chosen for the four invariant mass regions, such that the maximum possible fraction of $e^{\pm}\mu^{\mp}bb$ events that pass selection requirements fall in the same bin at both the particle and reconstruction level. The percentage in each bin shows the fraction of events at reconstruction level that fit into a given truth-level bin in the POWHEG+PYTHIA simulation (that is, the number of events in a given bin divided by the sum of the four bins in that column), along with the associated statistical uncertainty. The results from the additional Monte Carlo generators included in this analysis agree within 2.8% (except within 5.1% for the highest mass bin), included in Appendix B.

5.2 Detector-Level Distributions

To optimize comparison of detector-level distribution shapes, the number of signal+background $e^{\pm}\mu^{\mp}bb$ events from simulation was scaled to match the number of $e^{\pm}\mu^{\mp}bb$ events from data. This scaling factor is very close to unity, as can be seen in Table 4.1: $12437/11034.8 = 1.1271$. Both signal and background from simulation were multiplied by this factor. The scaled background from simulation was then subtracted from data, and the following figures show the detector-level results. Figures 5.4, 5.5, 5.6 and 5.7 show the number of additional jets present in reconstructed events, in data compared to expectation from simulation. Figure 5.8 shows the p_{T} for the leading additional jet in $e^{\pm}\mu^{\mp}bb$ events from data, compared to several Monte Carlo simulations, as a function of Q_0 for the four rapidity regions and displayed with the background from simulation. Figure 5.9 displays the same distribution but as a function of Q_{sum}, and Figures 5.10 and 5.11 display the same distributions but for the four invariant mass regions. Distributions of p_{T}, η and $|y|$ of the contributing electron and muon, b-tagged jets and additional jet activity from data and Monte Carlo simulations were compared to understand the shape and modelling of the data. The agreement of the distributions in the following figures confirms reasonable physics modelling. Figure 5.12 and 5.13 show the distributions of p_{T} and η of the two leptons in each signal event. Figure 5.14 shows the p_{T} and rapidity y distribution in the two b-tagged jets of each signal event. Figure 5.15 shows the p_{T} and $|y|$ rapidity distributions for the first leading additional jet, if any, in the events.

Finally, the resulting reconstructed-level gap fraction in data and simulation for

the four rapidity regions is shown in Figures 5.16 and 5.17 as a function of Q_0 and in Figures 5.18 and 5.19 as a function of Q_{sum}. Similarly, Figures 5.20, 5.21, 5.22 and 5.23 show the distributions for the four different invariant mass regions. Due to the "threshold" nature of the independent variable Q, the measurements for the different Q values are highly correlated, lending a cumulative nature to the distributions. (Each bin has a more stringent Q value cut than the bins to its left.) In each gap fraction plot, the left edge of each step in the ratio plot corresponds to the Q_0 or Q_{sum} threshold shown in the gap fraction plot above it.

The results from data generally agree with expectation from simulation within the statistical uncertainties. POWHEG+PYTHIA predicts more additional jets with higher p_T than data, and MC@NLO+HERWIG predicts fewer additional high p_T jets, with the exception of the rapidity region $1.5 < |y| < 2.1$.

(a)   (b)
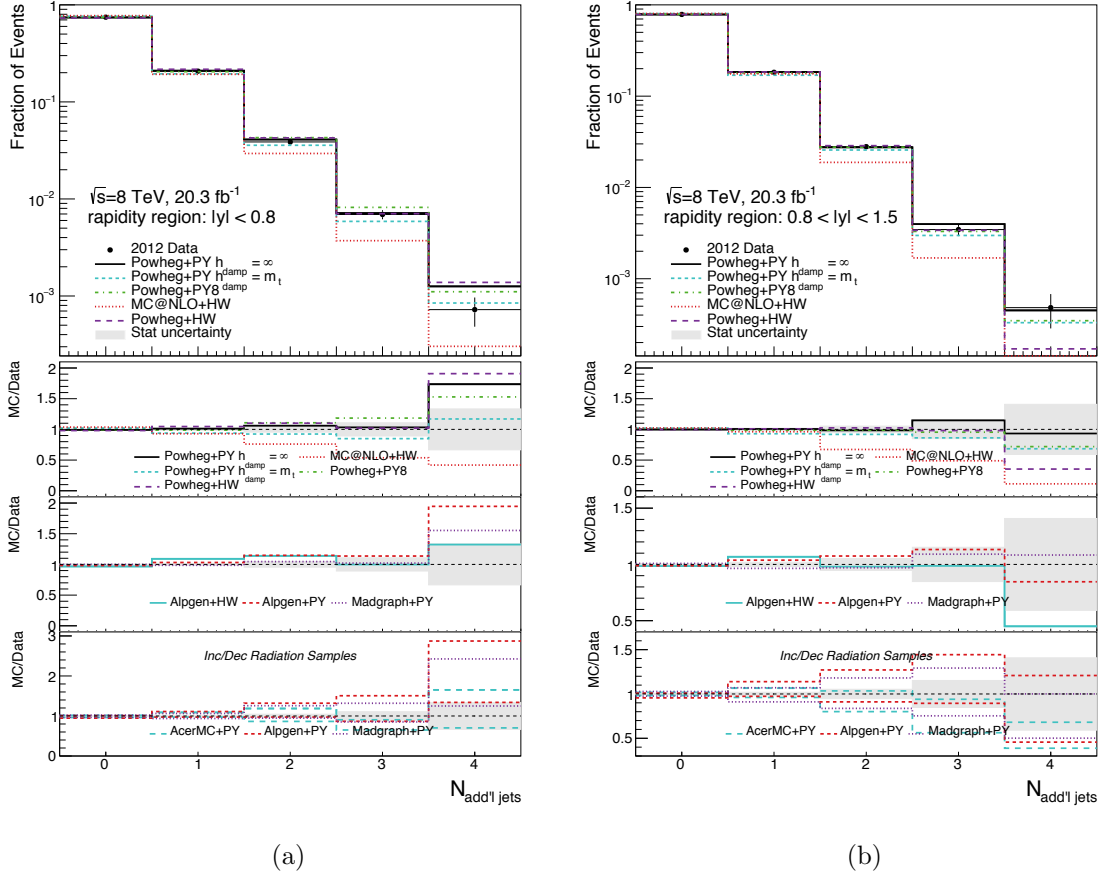
Figure 5.4: Multiplicity of additional jets within the given rapidity region in $e^{\pm}\mu^{\mp}bb$ events in data compared with expectation from Monte Carlo simulation, for regions a) $|y| < 0.8$ and b) $0.8 < |y| < 1.5$. The ratios show the multiplicity in simulation divided by the multiplicity in data; the statistical uncertainty in data is shown as a shaded band.
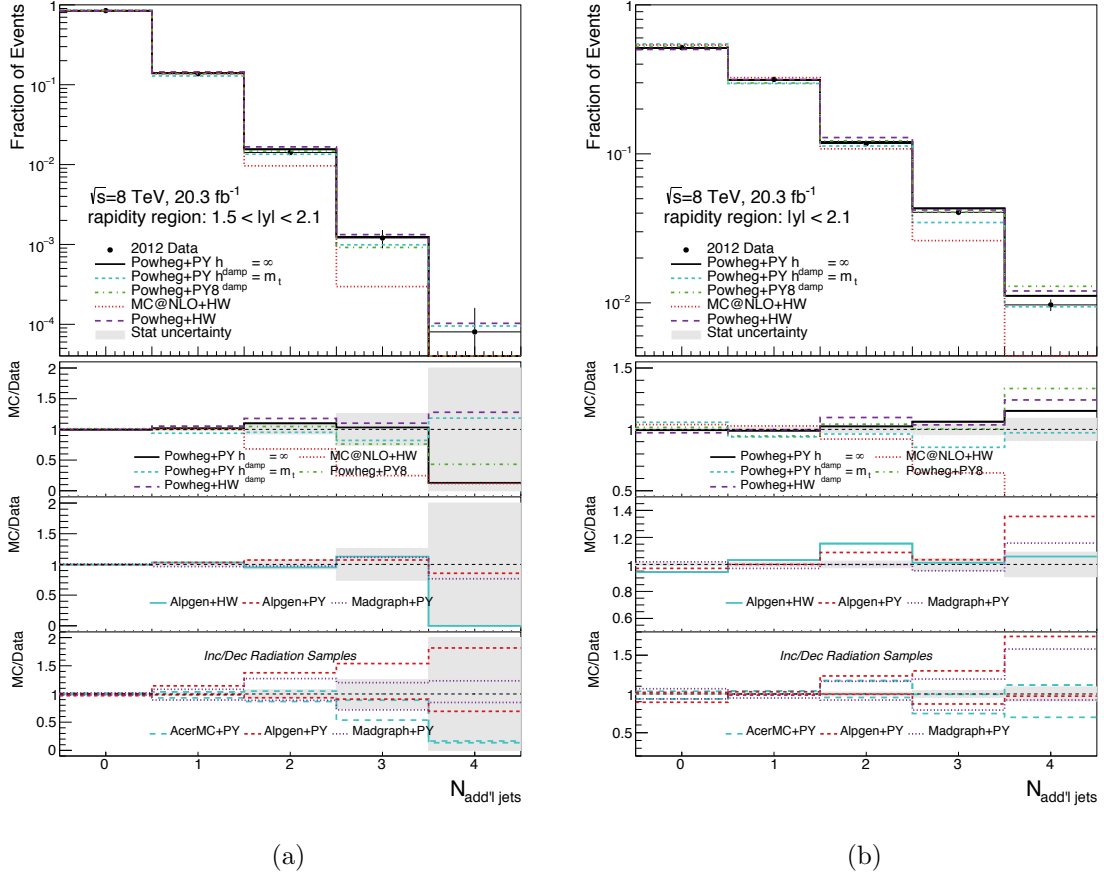
Figure 5.5: Multiplicity of additional jets within the given rapidity region in $e^{\pm}\mu^{\mp}bb$ events in data compared with expectation from Monte Carlo simulation, for the regions a) $1.5 < |y| < 2.1$ and b) $|y| < 2.1$. The ratios show the multiplicity in simulation divided by the multiplicity in data; the statistical uncertainty in data is shown as a shaded band.

(a)

(b)
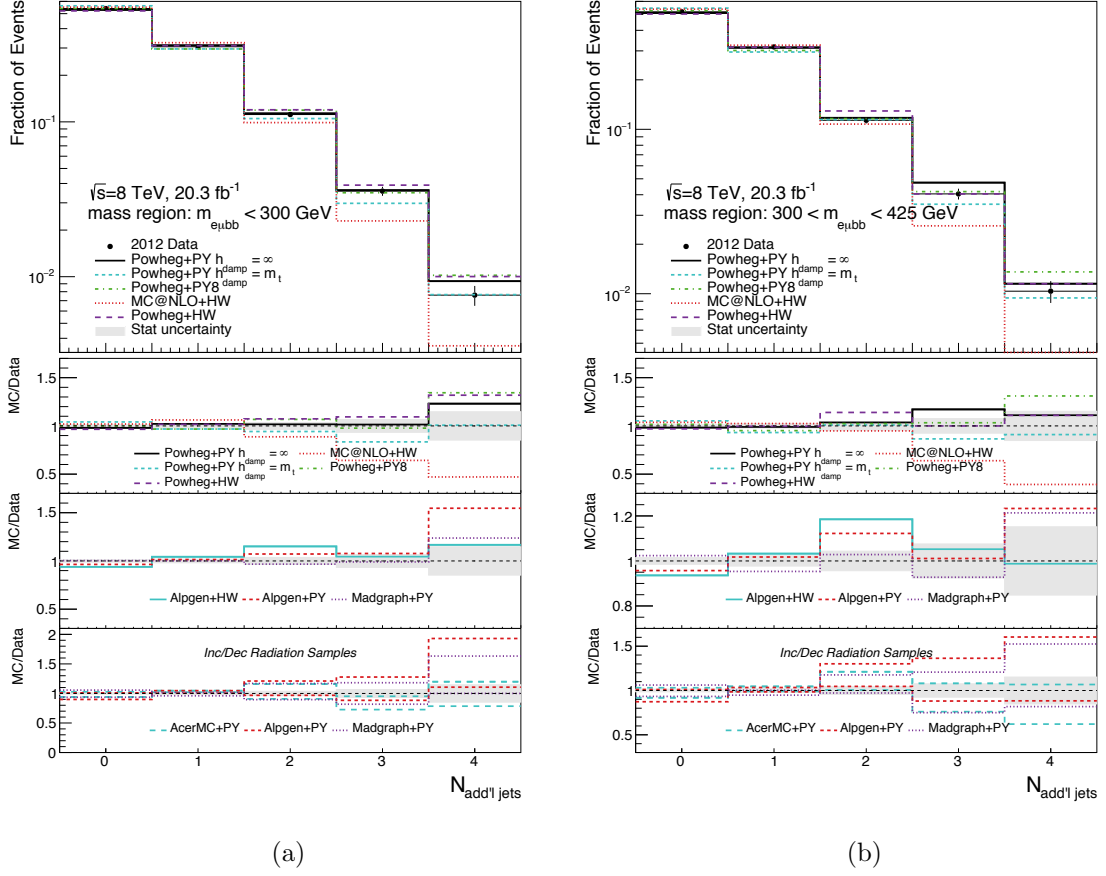
Figure 5.6: Multiplicity of additional jets in $e^{\pm}\mu^{\mp}bb$ events, in data and Monte Carlo simulations, for the $m_{e\mu bb}$ regions a) $m_{e\mu bb} < 300$ GeV and b) $300 < m_{e\mu bb} < 425$ GeV. The ratios show the multiplicity in simulation divided by the multiplicity in data; the statistical uncertainty in data is shown as a shaded band.
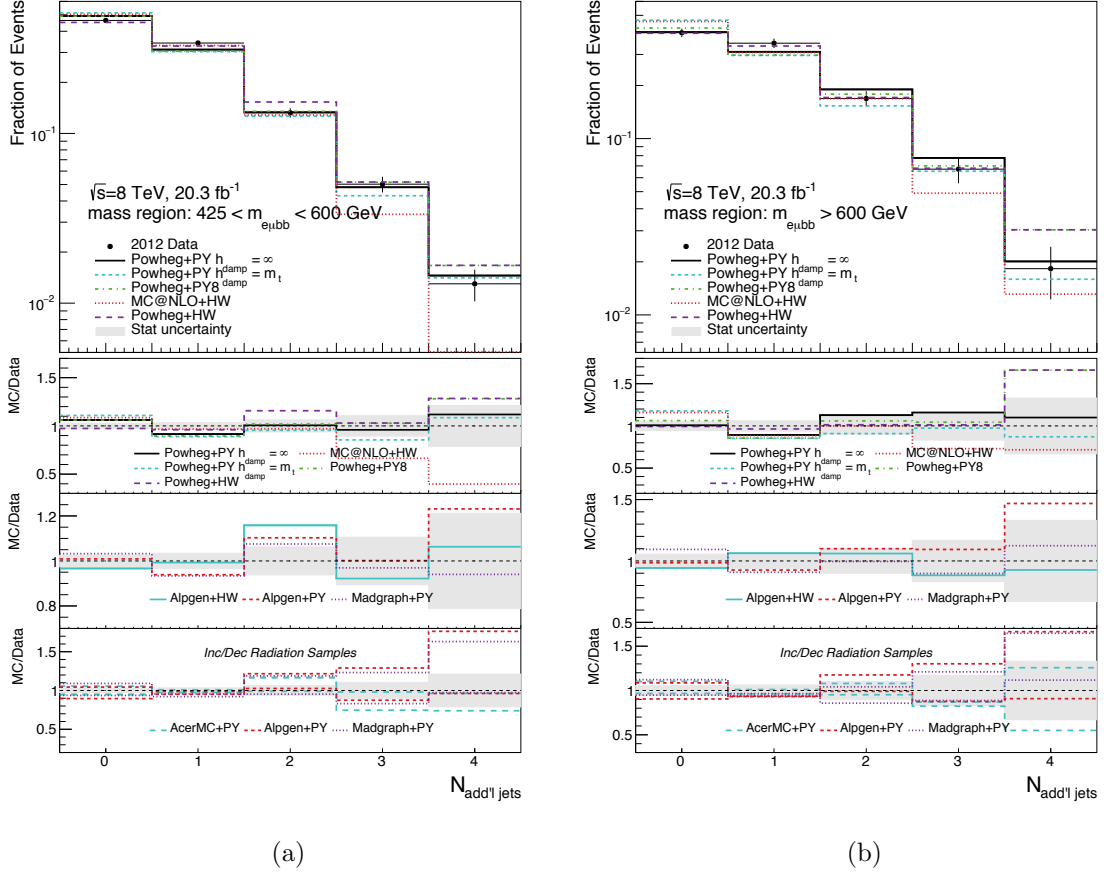
Figure 5.7: Multiplicity of additional jets in $e^{\pm}\mu^{\mp}bb$ events, in data and Monte Carlo simulations, for the $m_{e\mu bb}$ regions a) $425 < m_{e\mu bb} < 600$ GeV and b) $m_{e\mu bb} > 600$ GeV. The ratios show the multiplicity in simulation divided by the multiplicity in data; the statistical uncertainty in data is shown as a shaded band.
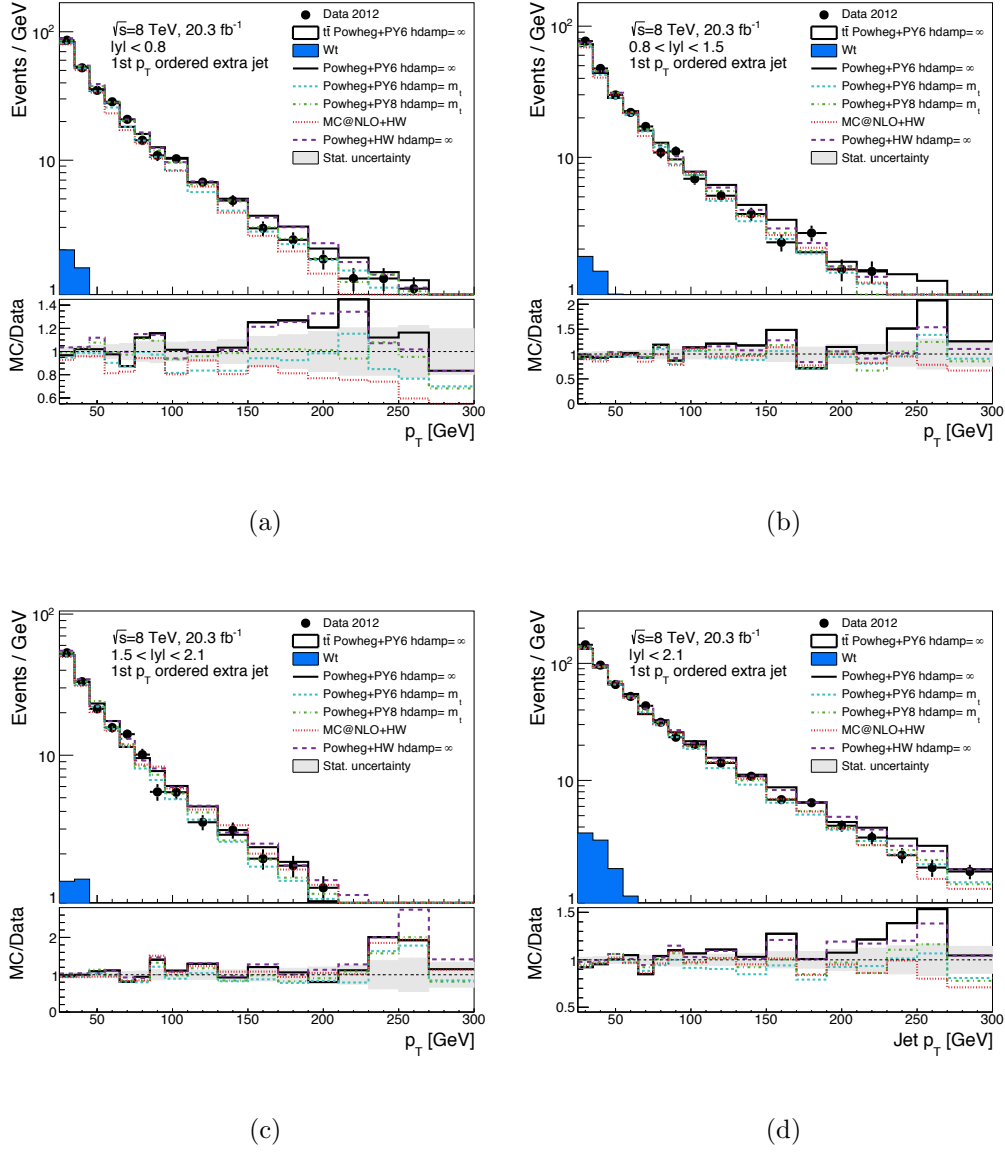
Figure 5.8: Leading additional jet p_{T} in the rapidity region a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$, as logarithmic plots, in $e^{\pm}\mu^{\mp}bb$ events in data compared to simulation (normalized to the number of $e^{\pm}\mu^{\mp}bb$ events in data). Background contributions are estimated from simulation. The single top background is shown, and the ratios show the events in simulation divided by the events in data.
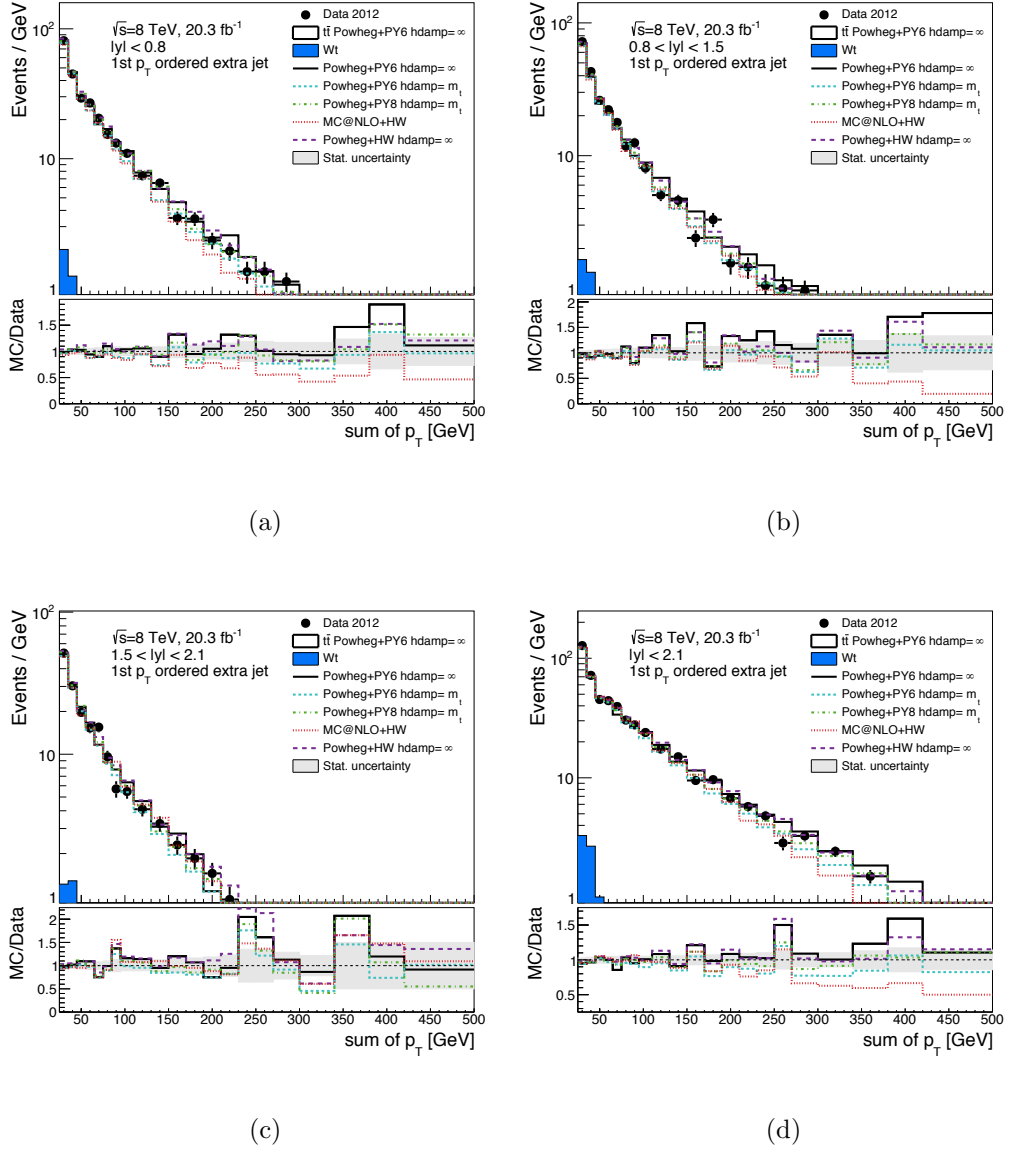
Figure 5.9: Total additional jet p_T in the rapidity region a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$ as logarithmic plots, in $e^\pm \mu^\mp bb$ events in data compared to MC simulation (normalized to the number of $e^\pm \mu^\mp bb$ events in data). Background contributions are estimated from simulation. The single top background is shown, and the ratios show the events in simulation divided by the events in data.
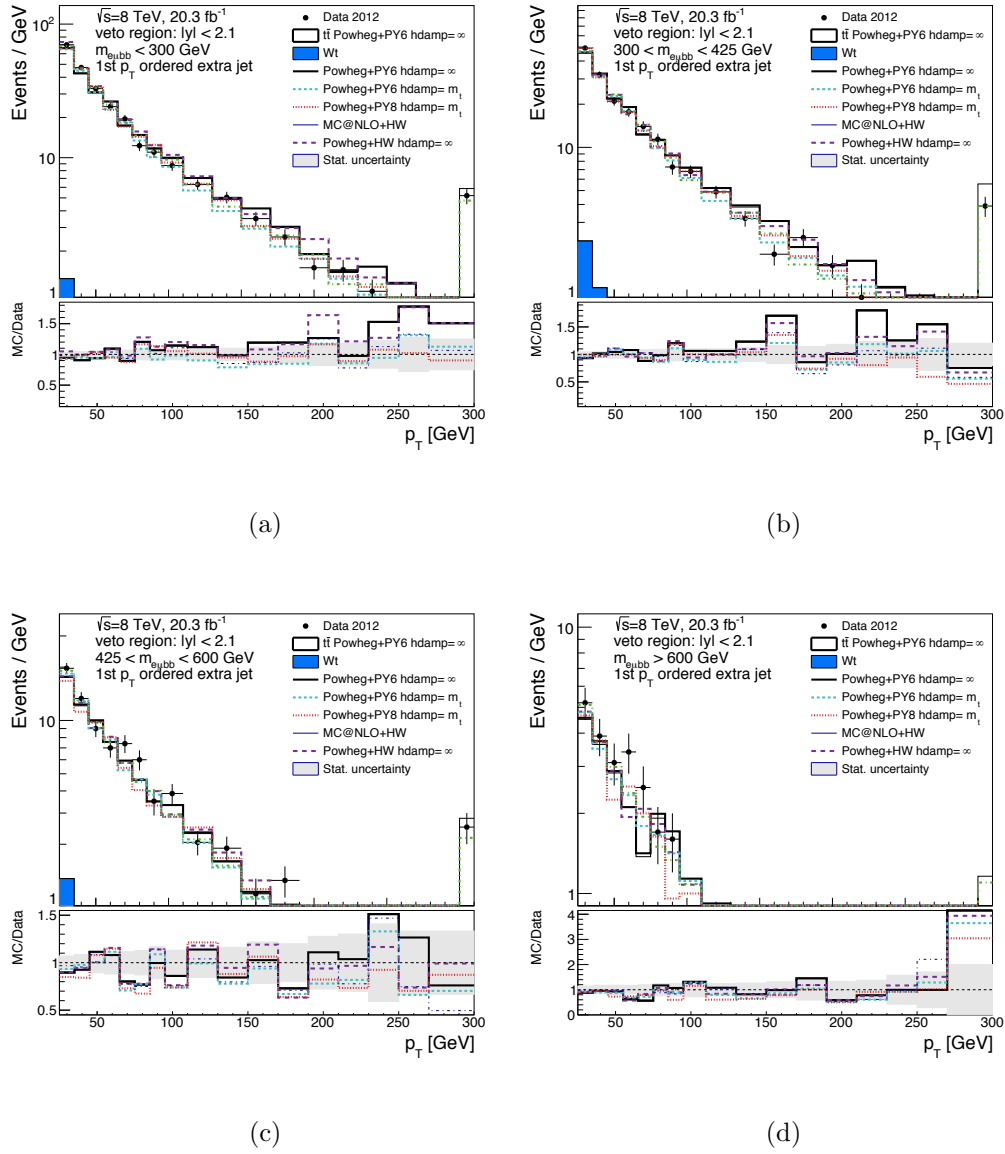
Figure 5.10: Leading additional jet p_T in the invariant mass region a) $m_{e\mu bb} <$ 300 GeV, b) $300 < m_{e\mu bb} < 425$ GeV, c) $425 < m_{e\mu bb} < 600$ GeV, and d) $m_{e\mu bb} >$ 600 GeV, as logarithmic plots, in $e^{\pm}\mu^{\mp}bb$ events in data compared to MC simulation (normalized to the number of $e^{\pm}\mu^{\mp}bb$ events in data). Background contributions are estimated from simulation. The ratios show the events in simulation divided by the events in data.
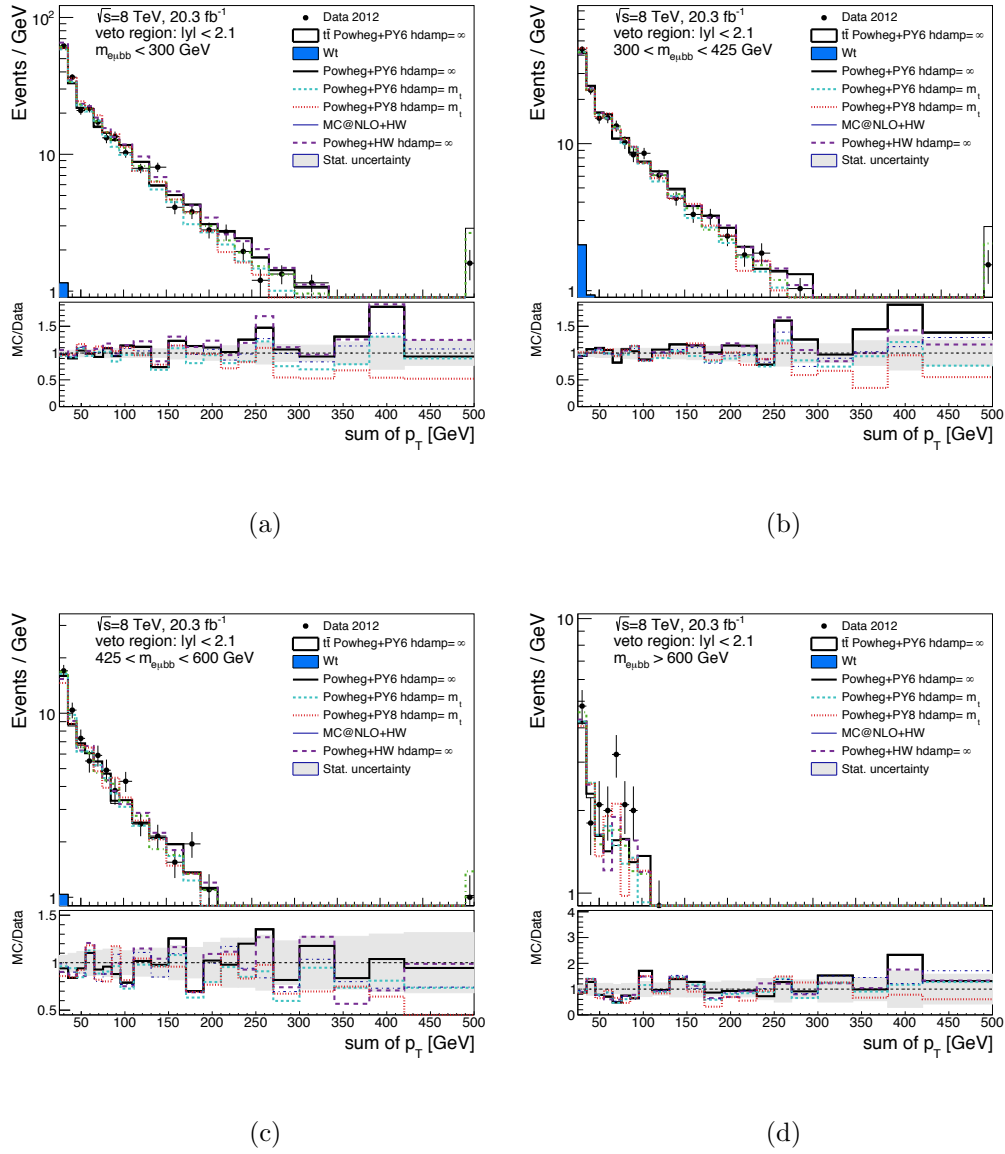
Figure 5.11: Total additional jet p_T in the invariant mass region a) $m_{e\mu bb} < 300$ GeV, b) $300 < m_{e\mu bb} < 425$ GeV, c) $425 < m_{e\mu bb} < 600$ GeV, and d) $m_{e\mu bb} > 600$ GeV, as logarithmic plots, in $e^{\pm}\mu^{\mp}bb$ events in data compared to MC simulation (normalized to the number of $e^{\pm}\mu^{\mp}bb$ events in data). Background contributions are estimated from simulation. The ratios show the events in simulation divided by the events in data.
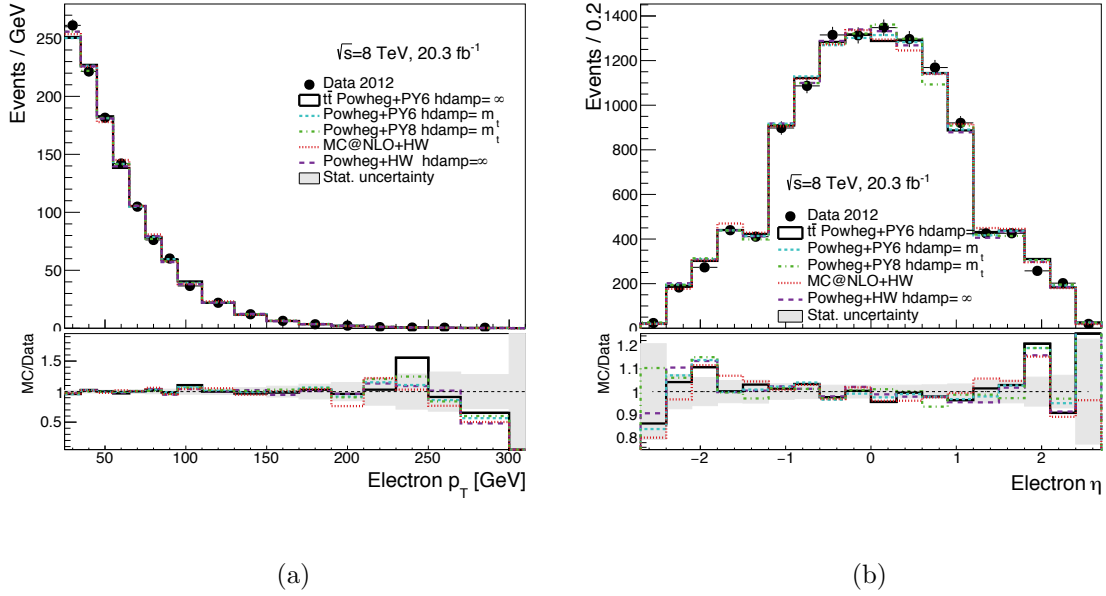
Figure 5.12: a) Electron p_T and b) η distributions from $e^\pm\mu^\mp bb$ data events with simulated background subtracted, compared to Monte Carlo simulation. The number of $e^\pm\mu^\mp bb$ events in simulation is normalized to the number of $e^\pm\mu^\mp bb$ events in data. The ratios show the events in simulation divided by the events in data.
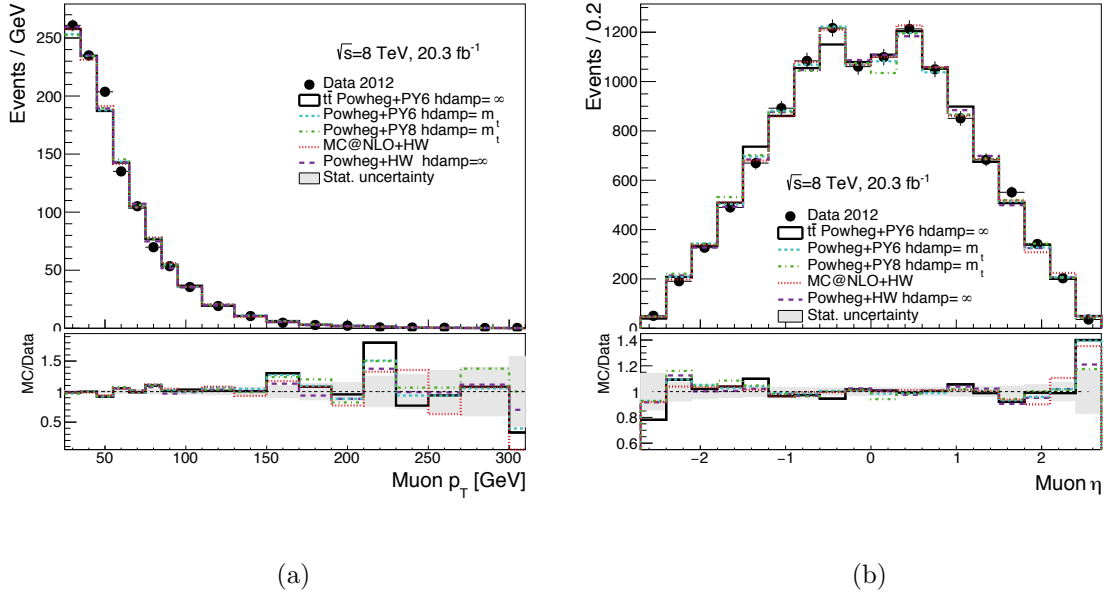
Figure 5.13: a) Muon p_T and b) η distributions from $e^{\pm}\mu^{\mp}bb$ data events with simulated background subtracted, compared to Monte Carlo simulation. The number of $e^{\pm}\mu^{\mp}bb$ events in simulation is normalized to the number of $e^{\pm}\mu^{\mp}bb$ events in data. The ratios show the events in simulation divided by the events in data.
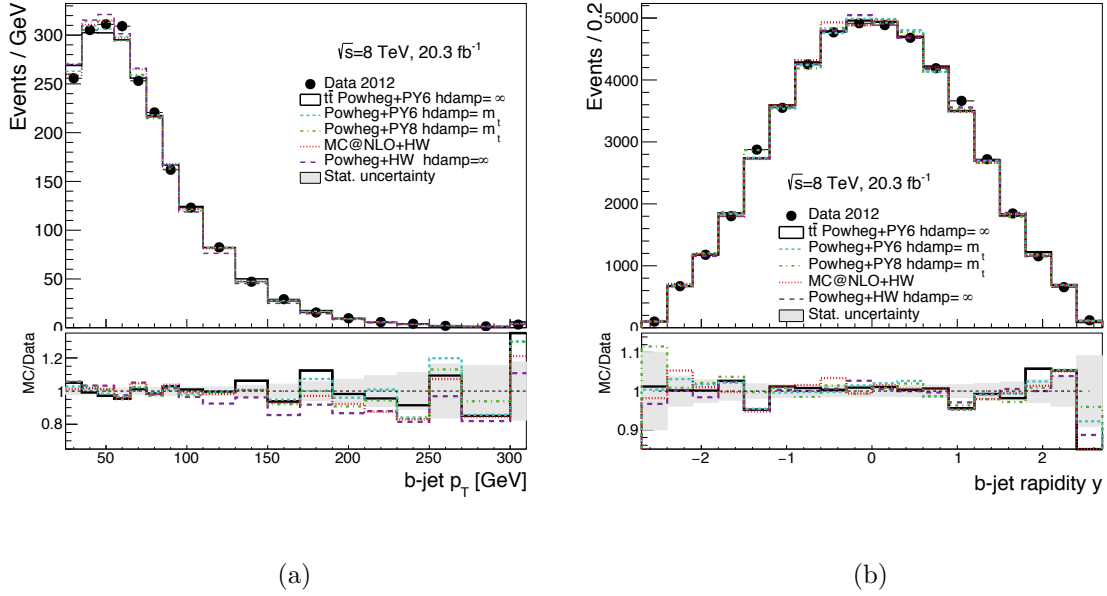
Figure 5.14: a) Jet p_{T} and b) y distributions for the two leading b-tagged jets in $e^{\pm}\mu^{\mp}bb$ data events with simulated background subtracted, compared to Monte Carlo simulation. The number of $e^{\pm}\mu^{\mp}bb$ events in simulation is normalized to the number of $e^{\pm}\mu^{\mp}bb$ events in data. The ratios show the events in simulation divided by the events in data.
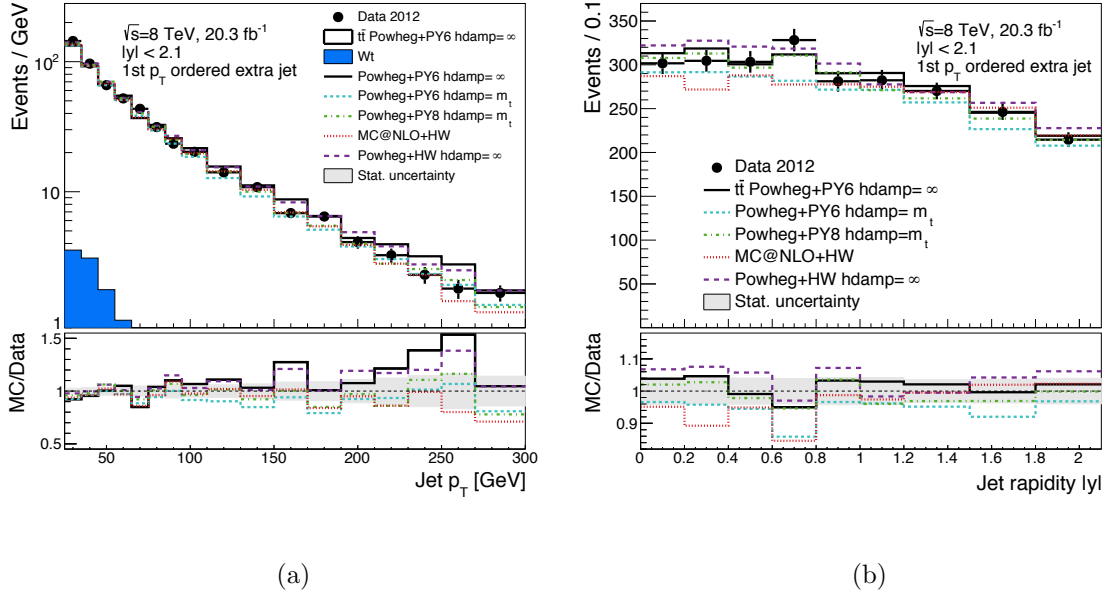
Figure 5.15: a) Jet p_T and b) $|y|$ distributions for the leading additional jet in $e^{\pm}\mu^{\mp}bb$ data events with simulated background subtracted, for $|y| < 2.1$, compared to several Monte Carlo simulations. The number of $e^{\pm}\mu^{\mp}bb$ events in simulation is normalized to the number of $e^{\pm}\mu^{\mp}bb$ events in data. The ratios show the events in simulation divided by the events in data.

Figure 5.16: The gap fraction in data at reconstruction level, compared to several Monte Carlo simulations, as a function of Q_0, for two of the four designated rapidity regions: a) $|y| < 0.8$ and b) $0.8 < |y| < 1.5$. The shaded band shows the statistical uncertainty, and the ratios show the gap fraction in simulation divided by the gap fraction from data. The left edge of each step in the ratio plots corresponds to the Q_0 threshold shown in the gap fraction plot above it.

Figure 5.17: The gap fraction in data at reconstruction level, compared to several Monte Carlo simulations, as a function of Q_0, for two of four designated rapidity regions: a) $1.5 < |y| < 2.1$ and b) $|y| < 2.1$. The shaded band shows the statistical uncertainty, and the ratios show the gap fraction in simulation divided by the gap fraction from data. The left edge of each step in the ratio plots corresponds to the Q_0 threshold shown in the gap fraction plot above it.

Figure 5.18: The gap fraction in data at reconstruction level, compared to several Monte Carlo simulations, as a function of Q_{sum}, again for the first two of four designated rapidity regions: a) $|y| < 0.8$ and b) $0.8 < |y| < 1.5$. The shaded band shows the statistical uncertainty, and the ratios show the events in simulation divided by the events in data. The left edge of each step in the ratio plots corresponds to the Q_{sum} threshold shown in the gap fraction plot above it.
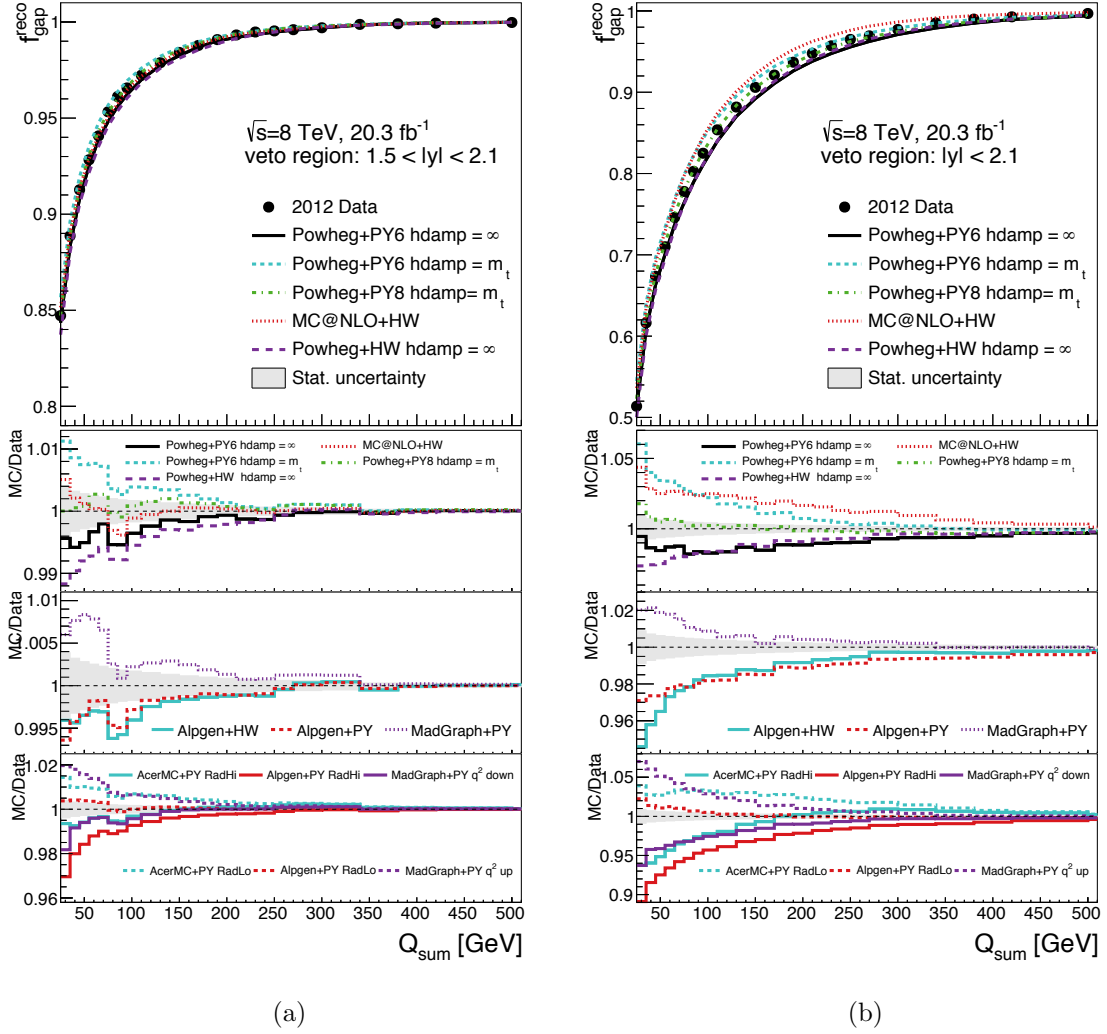
Figure 5.19: The gap fraction in data at reconstruction level, compared to several Monte Carlo simulations, as a function of Q_{sum}, again for the second two of four designated rapidity regions: a) $1.5 < |y| < 2.1$ and b) $|y| < 2.1$. The shaded band shows the statistical uncertainty, and the ratios show the events in simulation divided by the events in data. The left edge of each step in the ratio plots corresponds to the Q_{sum} threshold shown in the gap fraction plot above it.

(a)   (b)
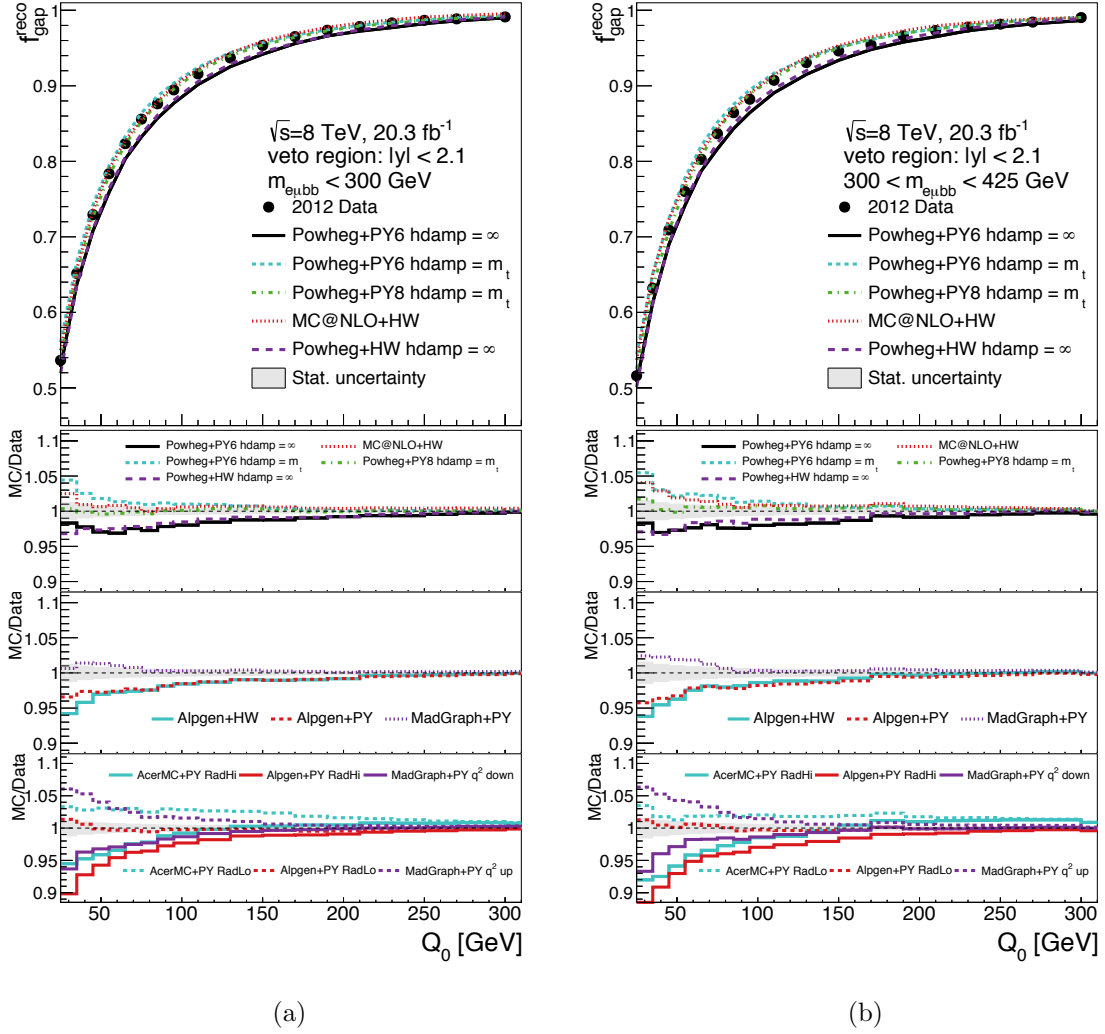
Figure 5.20: The gap fraction in data at reconstruction level, compared to several Monte Carlo simulations, as a function of Q_0, for two of the four invariant mass regions: a) $m_{e\mu bb} < 300$ GeVand b) $300 < m_{e\mu bb} < 425$ GeV. The shaded band shows statistical uncertainty, and the ratios show the events in simulation divided by the events in data. The left edge of each step in the ratio plots corresponds to the Q_0 threshold shown in the gap fraction plot above it.
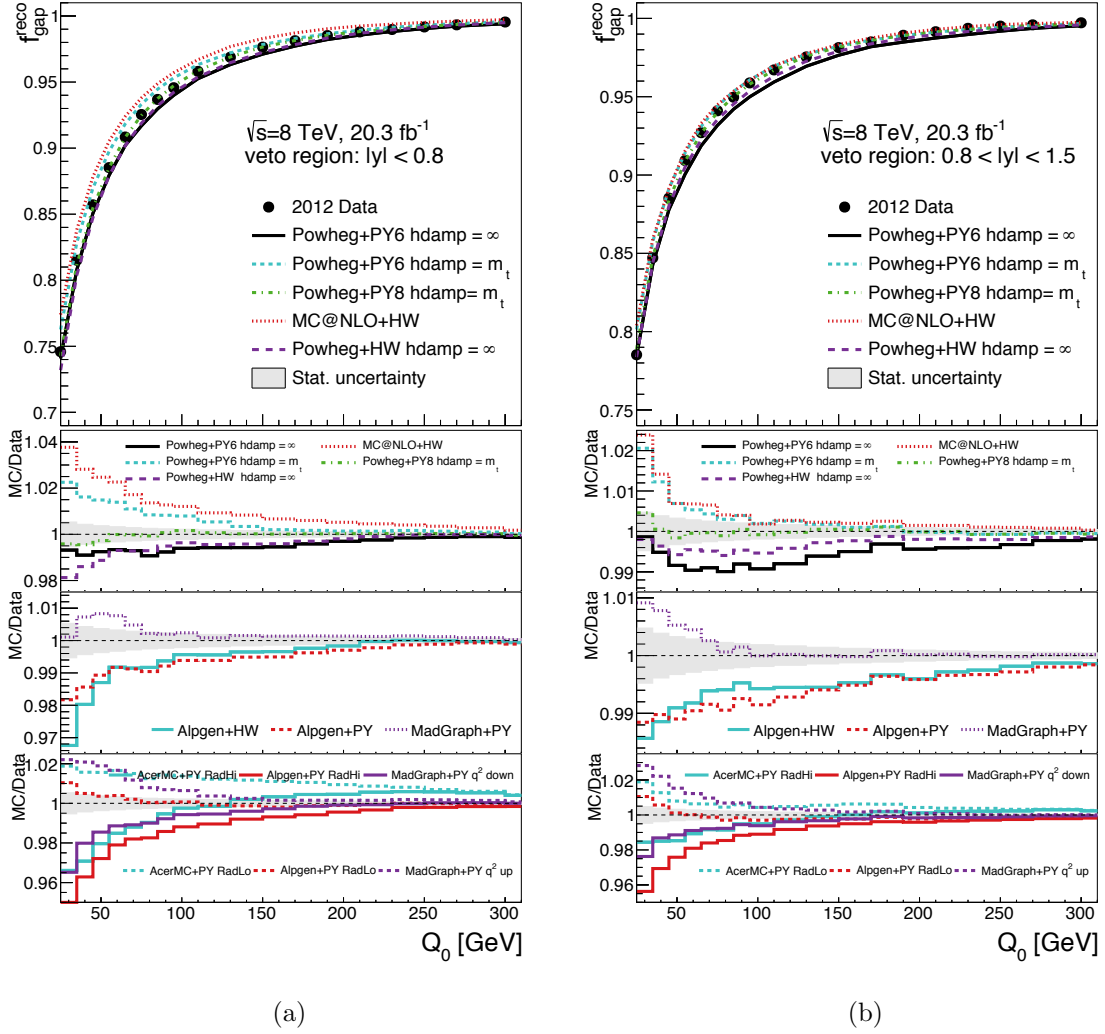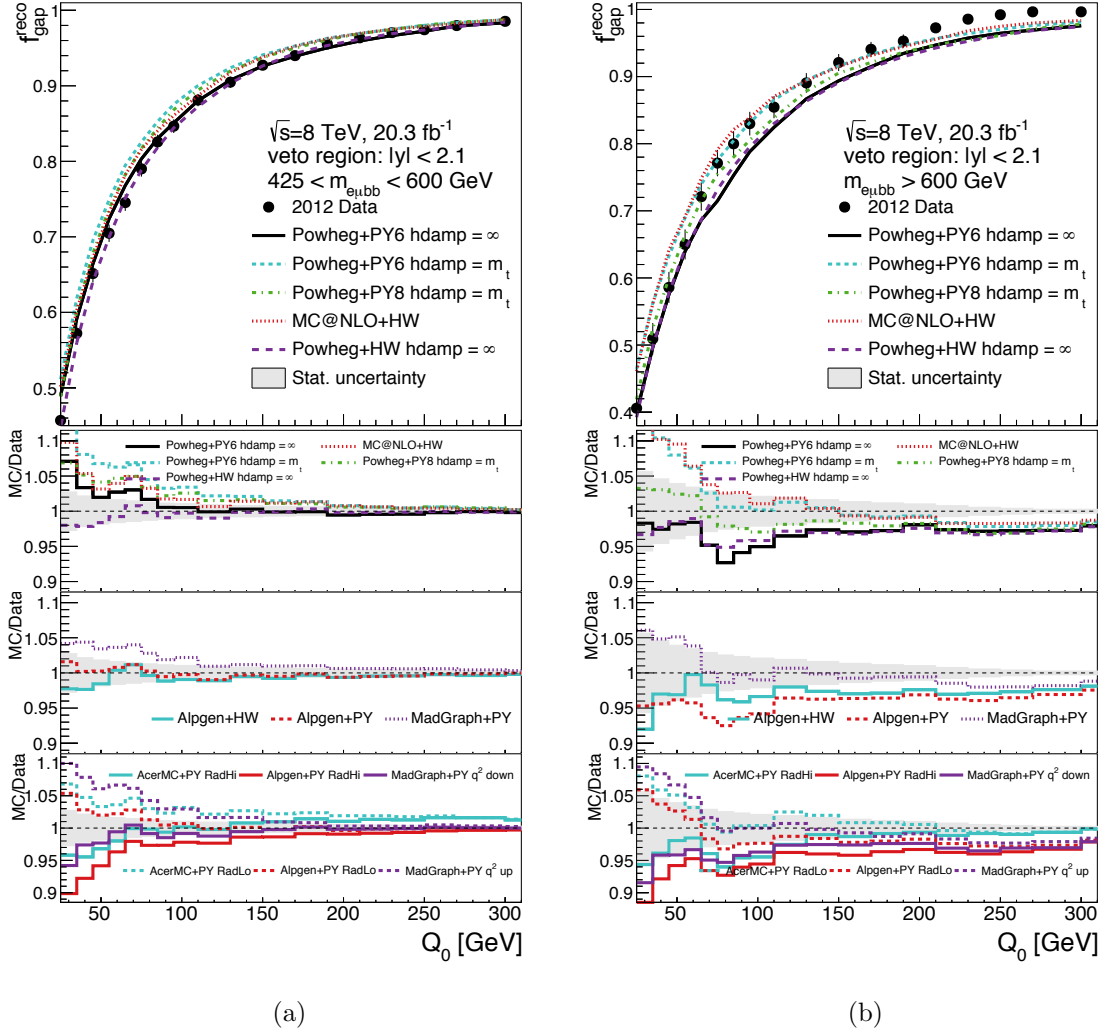
Figure 5.21: The gap fraction in data at reconstruction level, compared to several Monte Carlo simulations, as a function of Q_0, for two of the four invariant mass regions: a) $425 < m_{e\mu bb} < 600$ GeV and b) $m_{e\mu bb} > 600$ GeV. The shaded band shows statistical uncertainty, and the ratios show the events in simulation divided by the events in data. The left edge of each step in the ratio plots corresponds to the Q_0 threshold shown in the gap fraction plot above it.
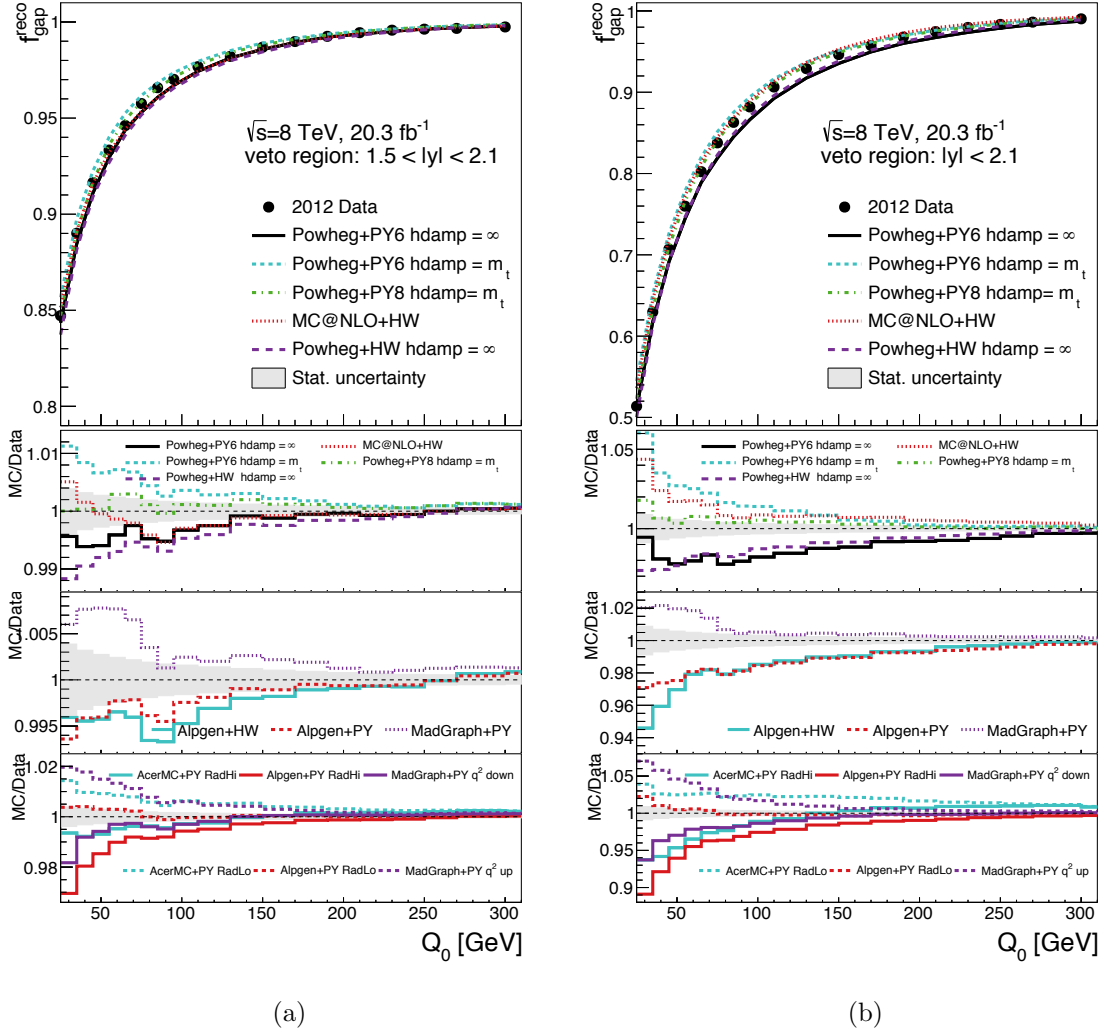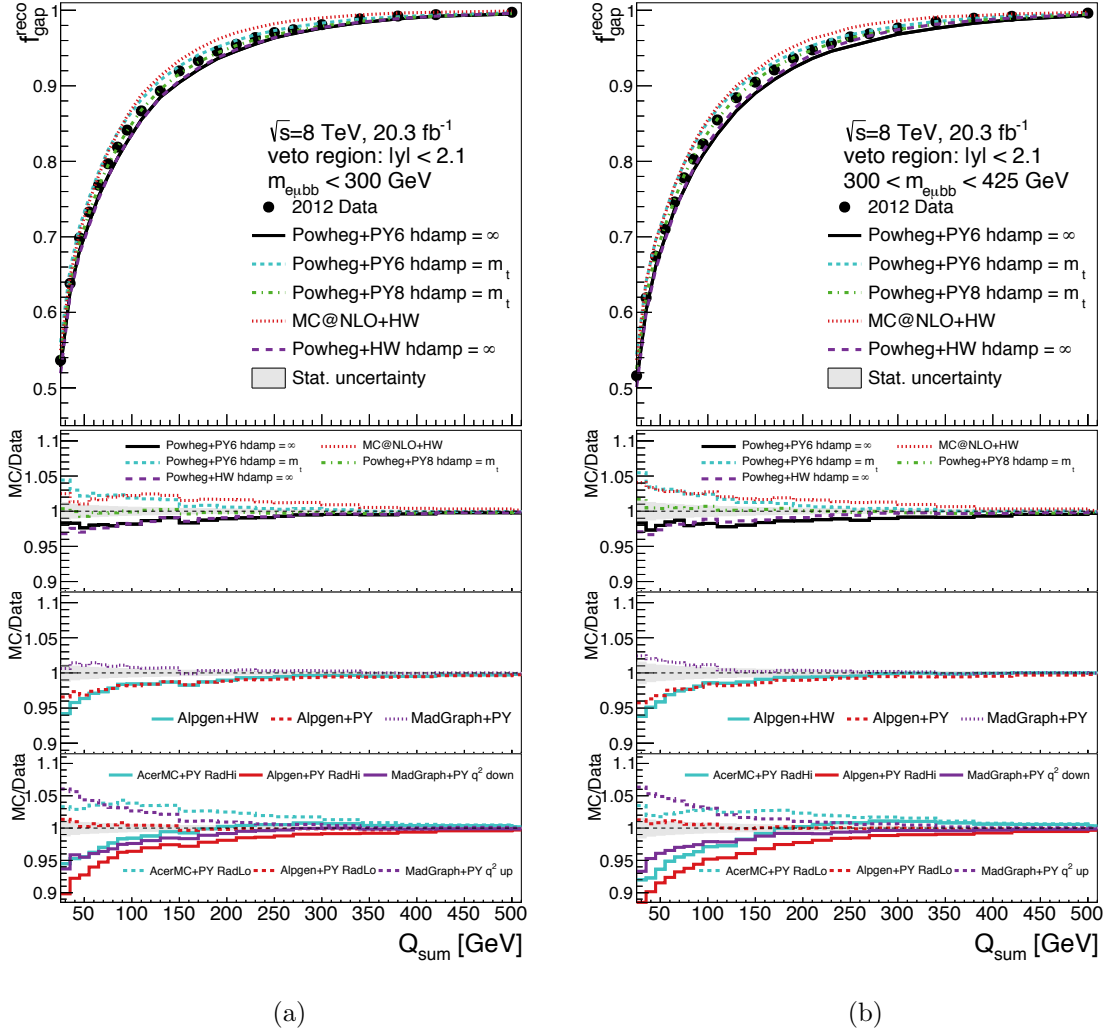
Figure 5.22: The gap fraction in data at reconstruction level, compared to several Monte Carlo simulations, as a function of Q_{sum}, for two of the four invariant mass regions: a) $m_{e\mu bb} < 300$ GeV and b) $300 < m_{e\mu bb} < 425$ GeV. The shaded band shows the statistical uncertainty, and the ratios show the events in simulation divided by the events in data. The left edge of each step in the ratio plots corresponds to the Q_{sum} threshold shown in the gap fraction plot above it.
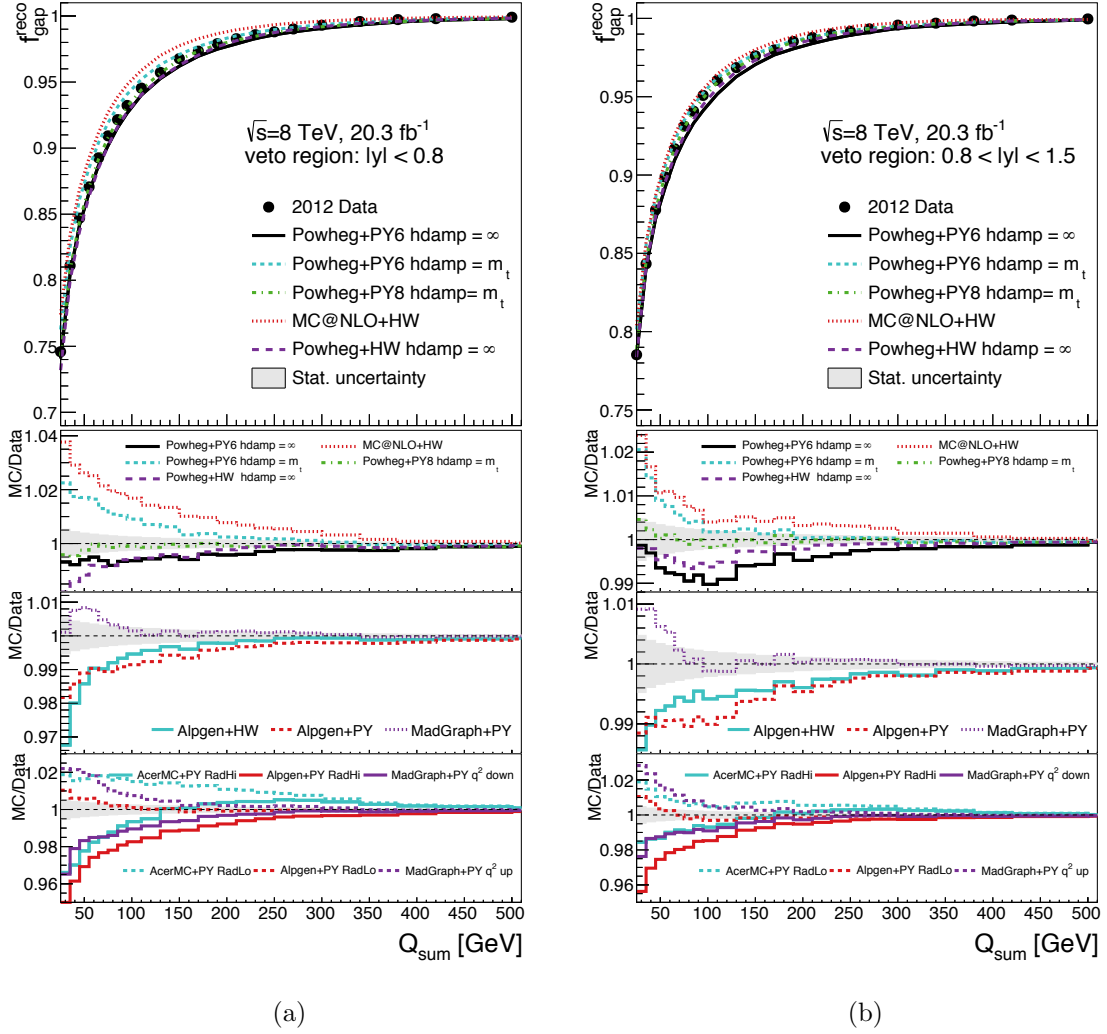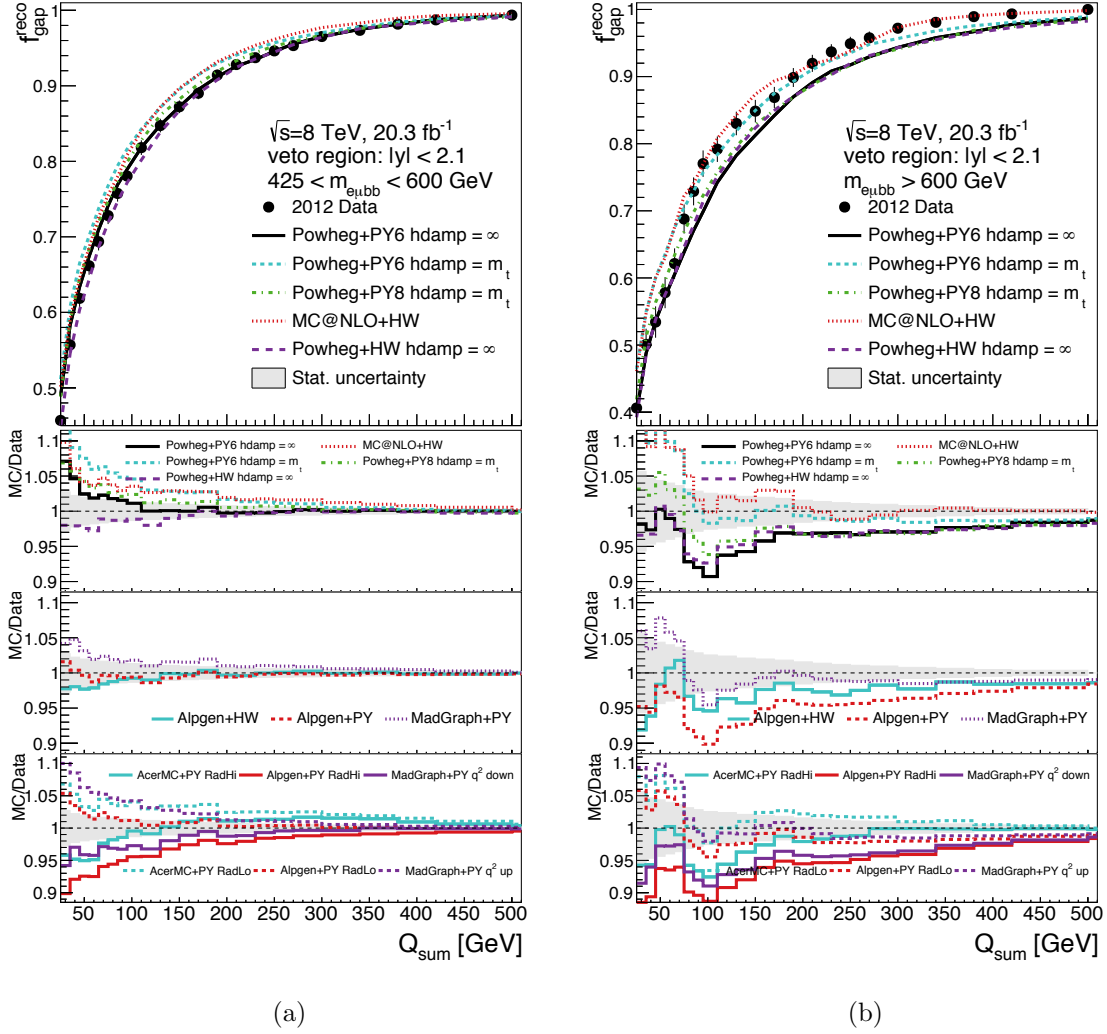
Figure 5.23: The gap fraction in data at reconstruction level, compared to several Monte Carlo simulations, as a function of Q_{sum}, for two of the four invariant mass regions: a) $425 < m_{e\mu bb} < 600$ GeV and b) $m_{e\mu bb} > 600$ GeV. The shaded band shows the statistical uncertainty, and the ratios show the events in simulation divided by the events in data. The left edge of each step in the ratio plots corresponds to the Q_{sum} threshold shown in the gap fraction plot above it.

5.3 Jet Transverse Momentum Resolution

The resolution of jet p_T at the particle and reconstruction level from simulation was used to choose commensurate threshold spacing in the gap fraction measurement.

First, to investigate the effect of reconstruction on jet p_T, the p_T of matched jets in $e^\pm\mu^\mp bb$ events was compared at particle and reconstruction level in the POWHEG+PYTHIA simulation (following $e^\pm\mu^\mp bb$ event selection described in Section 4.1, and jet matching requirements described in Section 4.6). Jets meeting the b-jet matching criteria are plotted in Figure 5.24a and additional jets in 5.24b, with the p_T at particle level juxtaposed with the p_T at reconstruction level. The particle level jets are allowed to have $p_T > 10$ GeV, which is illustrated in the plots' lower p_T threshold at particle level. A noticeable trend in the b-jet plot, shows a portion of the b-jets having higher p_T at particle level. This is expected, because weakly decaying b quarks produce neutrinos, which are included in the measurement of jet p_T at particle level with the `TopFiducial` package, but are not included at reconstruction level.

The RMS of differences in p_T between the particle and reconstruction level, in the POWHEG+PYTHIA simulation, is plotted for matched b-jets in Figure 5.25a and for matched additional jets in Figure 5.25b. Again, a trend of b-jets having lower p_T at reconstruction level occurs. The dashed line in Figure 5.25b shows the variable thresholds chosen for the gap fraction measurement, increasing with p_T^{part}. This choice of threshold spacing encompasses the majority of the spread of differences and matches those used in the 2011 gap fraction measurement. It is chosen to be:

$Q_0 = 25, 35, 45, 55, 65, 75, 85, 95, 110, 130, 150, 170, 190, 210, 230, 250, 270, 300$ GeV
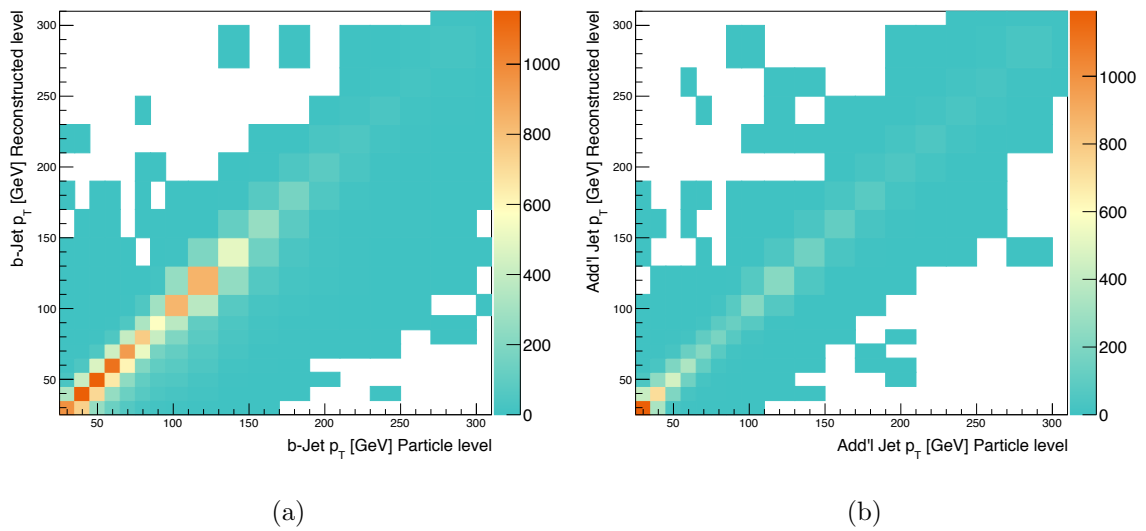
Figure 5.24: Jet energy resolution of a) all matched b-jets and b) all additional jets in $e^{\pm}\mu^{\mp}bb$ events, showing each particle level jet p_{T} plotted with its matched reconstruction level jet p_{T}.

$Q_{\text{sum}} = 25, 35, 45, 55, 65, 75, 85, 95, 110, 130, 150, 170, 190, 210, 230, 250,$

$270, 300, 340, 380, 420, 500$ GeV

(a)   (b)
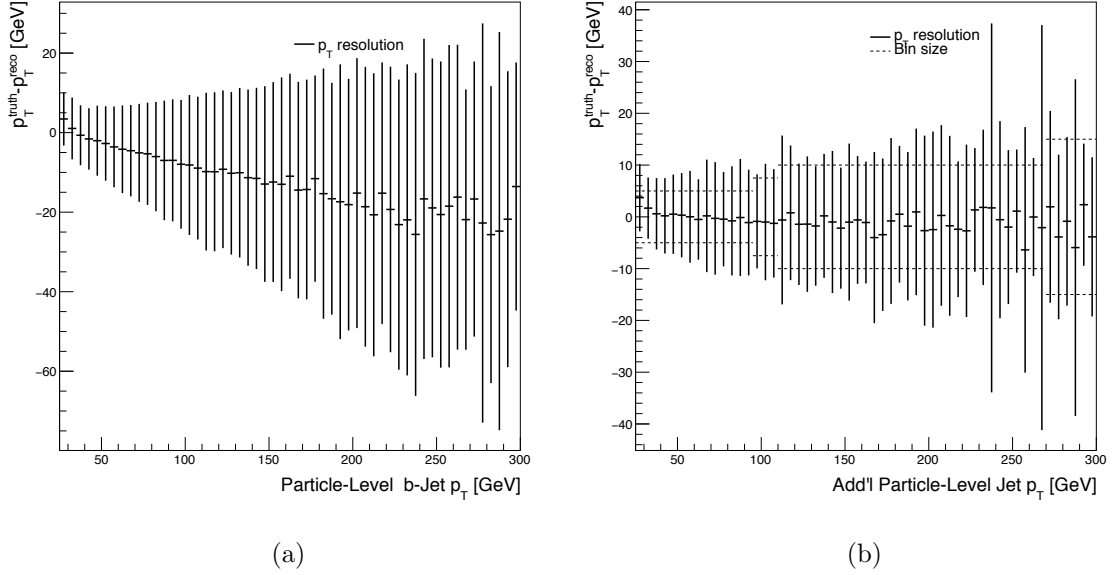
Figure 5.25: The difference in p_{T} of matched jets between the particle and reconstruction level in the POWHEG+PYTHIA sample, as a function of the particle-level jet p_{T}. The vertical line gives the spread of values and the horizontal line shows the mean value. In a), the difference in matched b-jet p_{T} is shown, verifying the downward shift of the p_{T} at reconstruction level discussed in the text. In b), the difference in additional matched jet p_{T} is shown, with dashed lines representing the width of the threshold spacing chosen for the gap fraction measurement.

5.4 Correction for Detector Effects

The gap fraction from data was corrected back to the particle level with a step-by-step correction factor built from simulation, to give a detector-independent form of the result for comparison to theoretical predictions and for tuning Monte Carlo generators. The correction factor $C[i]$ is calculated as the ratio of the truth level ("part") and reconstruction level ("reco") gap fraction from POWHEG+PYTHIA simulation for each Q_0 and Q_{sum} value:

$$C[i] = \frac{f_{\mathrm{gap}}^{\mathrm{part,MC}}[i]}{f_{\mathrm{gap}}^{\mathrm{reco,MC}}[i]} \tag{5.7}$$

In this formula, f_{gap} is calculated from simulation as in Equation 5.2, and i refers to the Q_0 or Q_{sum} threshold.

Figures 5.26 and 5.27 show the resulting correction factors from the POWHEG +PYTHIA sample along with additional Monte Carlo simulations for Q_0 and Q_{sum}, respectively, for the four rapidity regions, which are on the order of unity. For the four rapidity regions, the correction factors shown are generated from `fullsim` samples. Due to statistical limitations in the invariant mass regions, the correction factor for these regions is sourced from `fastsim` samples (as described in Section 3.2), and these are shown in Figures 5.28 and 5.29. The proximity of the correction factor to unity and the similarity between different Monte Carlo simulation results, particularly in the four rapidity regions, are evidence that the step-by-step correction factor should give nearly the same results as a full unfolding technique. The correction factors for the four mass regions are larger. In both the rapidity regions and the invariant mass

regions, the bins (thresholds) are highly correlated due to the cumulative nature of the correction factor. The uncertainty in the correction factor is discussed in Section 5.5.5.

A measurement of purity P and stability S in the correction factor was performed for the POWHEG+PYTHIA sample, where the purity and stability are defined in terms of the number of events contributing to the gap fraction measurement at the particle and reconstruction levels:

$$P(Q_0) = \frac{n_{\mathrm{gap}}^{\mathrm{reco\,\&\,part}}(Q_0)}{n_{\mathrm{gap}}^{\mathrm{reco}}(Q_0)} \tag{5.8}$$

$$S(Q_0) = \frac{n_{\mathrm{gap}}^{\mathrm{reco\,\&\,part}}(Q_0)}{n_{\mathrm{gap}}^{\mathrm{part}}(Q_0)} \tag{5.9}$$

In these definitions, $n_{\mathrm{gap}}^{\mathrm{reco\,\&\,part}}(Q_0)$ represents the number of events that pass both particle and reconstruction level event selections and have no additional jet with $p_{\mathrm{T}} > Q_0$ GeV at either the particle or reconstruction level in the selected veto region. Similarly, $n_{\mathrm{gap}}^{\mathrm{reco}}(Q_0)$ is the number of events that pass reconstruction level event selection and have no additional jet (at reconstruction level) with $p_{\mathrm{T}} > Q_0$ GeV within the selected veto region, and $n_{\mathrm{gap}}^{\mathrm{part}}(Q_0)$ is the number of events that pass particle level event selection and have no additional jet (at the particle level) with $p_{\mathrm{T}} > Q_0$ GeV within the selected veto region.

Figures 5.30 and 5.31 show the resulting purity and stability for each rapidity region, respectively. The purity and stability definitions of the invariant mass regions depend on whether the events passing particle and/or reconstruction level selection are required to have an invariant mass *within that invariant mass region* or only to have passed particle and/or reconstruction level selection in any invariant mass region. This means that requiring selection with an invariant mass in that region includes an
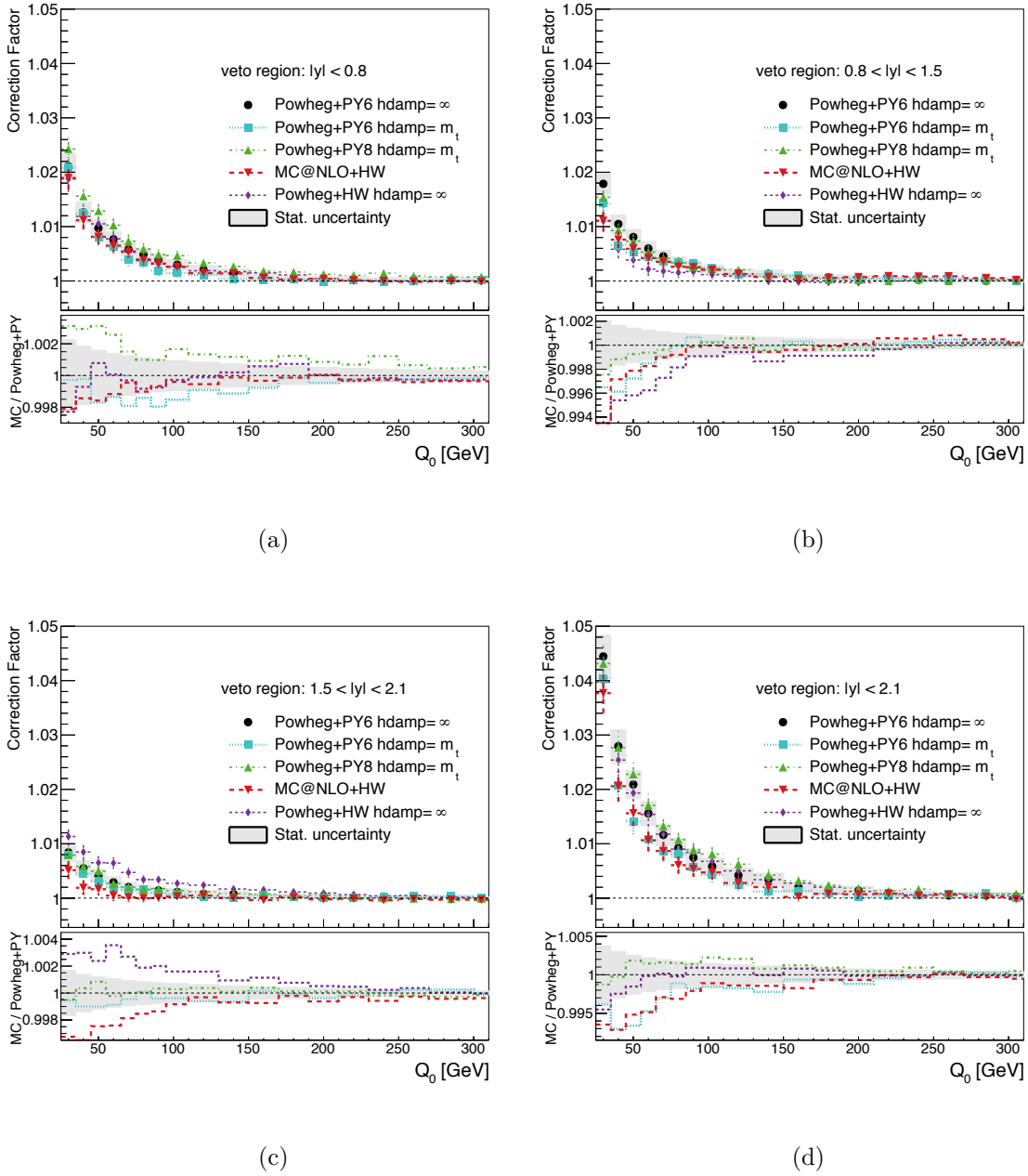
Figure 5.26: The correction factor for detector effects, built with the `fullsim` POWHEG+PYTHIA simulation, as a function of Q_0 for each of four rapidity regions: a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$. The error bars show the statistical uncertainties.
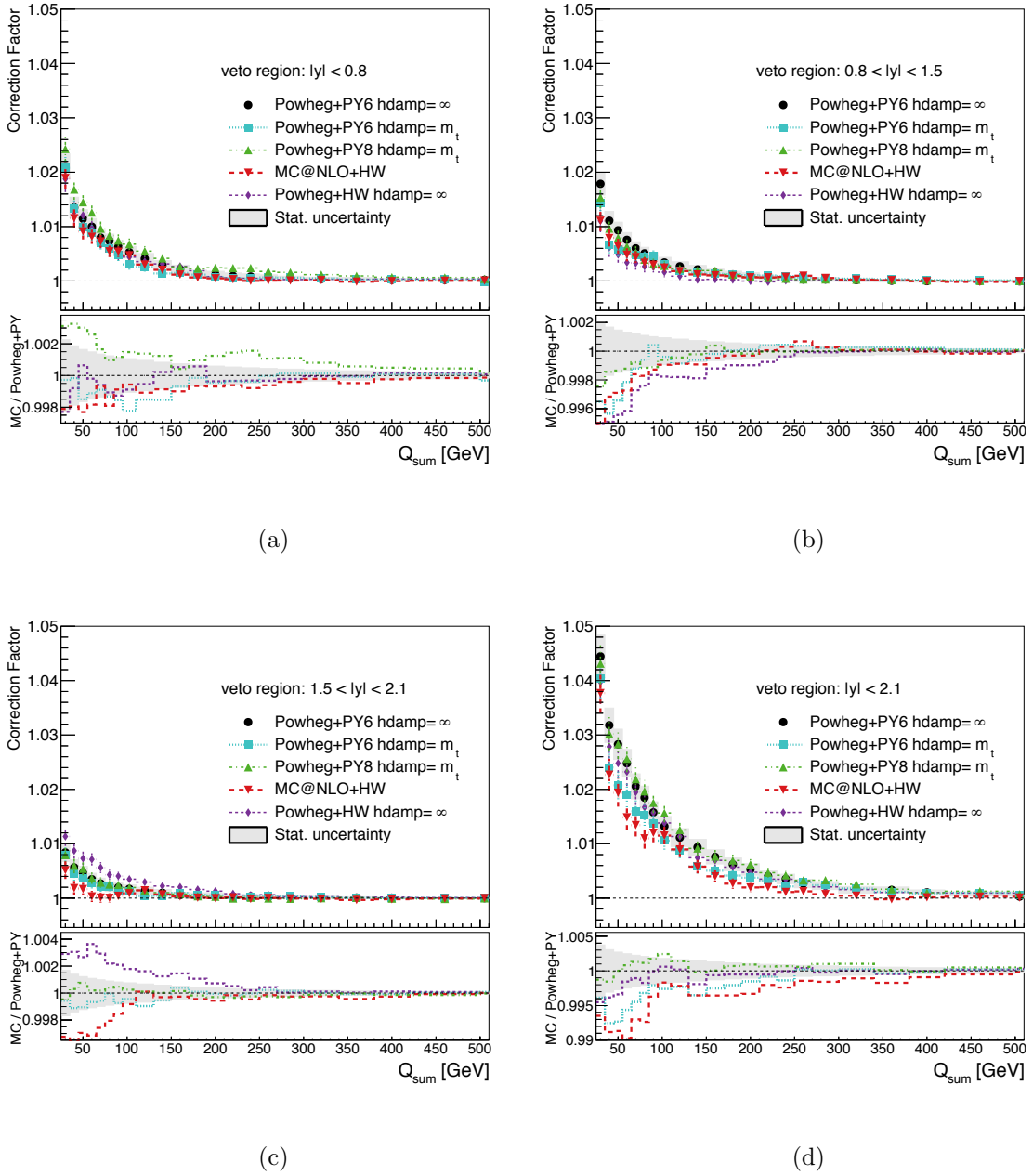
Figure 5.27: The correction factor for detector effects in the gap fraction, built with the `fullsim` POWHEG+PYTHIA simulation, as a function of Q_{sum} for each of four rapidity regions: a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$. The error bars show the statistical uncertainties.
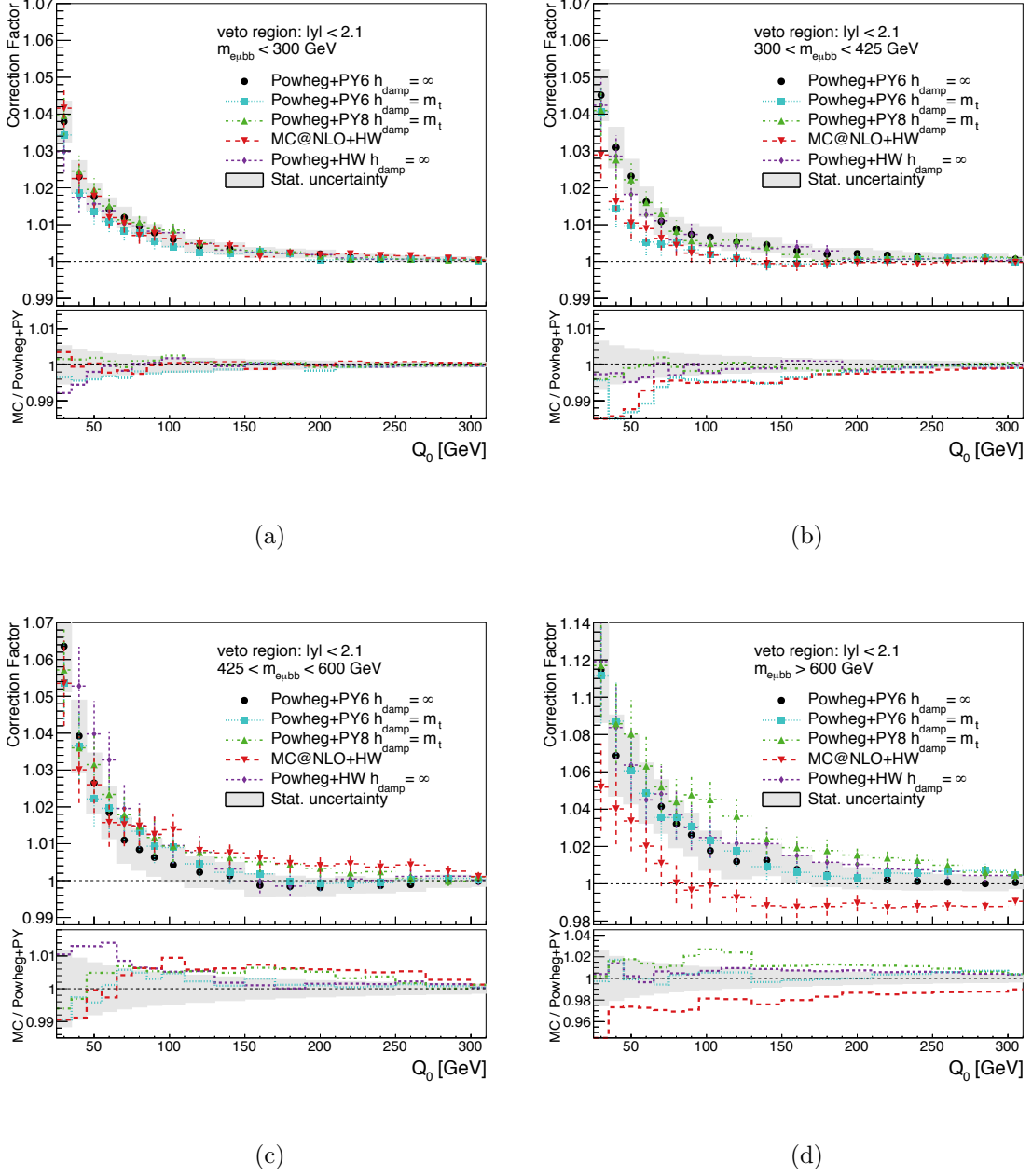
Figure 5.28: The correction factor for detector effects in the gap fraction, built with the fastsim POWHEG+PYTHIA simulation, as a function of Q_0 for each invariant mass region: a) $m_{e\mu bb} < 300$ GeV, b) $300 < m_{e\mu bb} < 425$ GeV, c) $425 < m_{e\mu bb} < 600$ GeV, and d) $m_{e\mu bb} > 600$ GeV. The error bars show the statistical uncertainties.
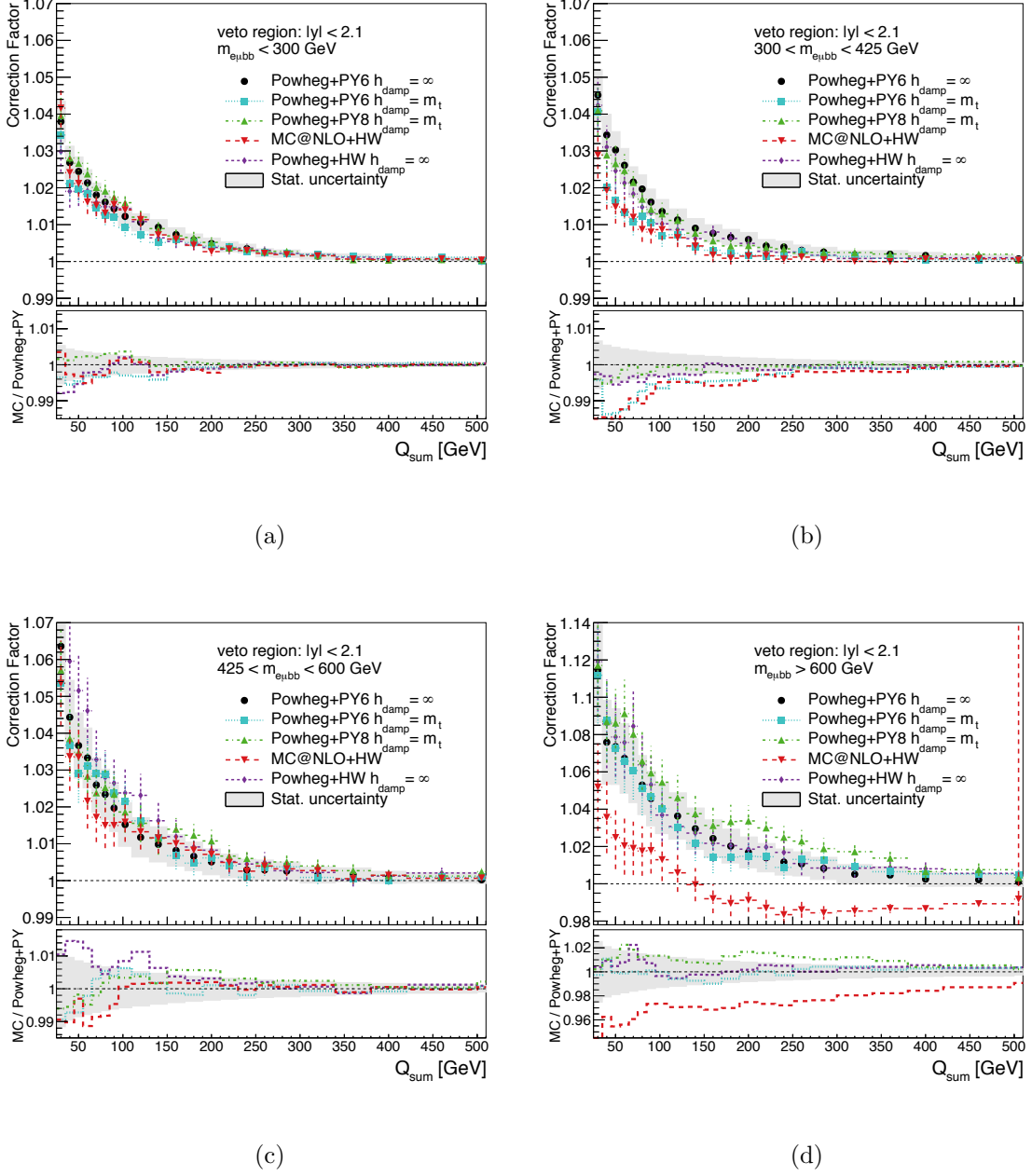
Figure 5.29: The correction factor for detector effects in the gap fraction, built with the `fastsim` Powheg+Pythia simulation, as a function of Q_{sum} for the four invariant mass regions: a) $m_{e\mu bb} < 300$ GeV, b) $300 < m_{e\mu bb} < 425$ GeV, c) $425 < m_{e\mu bb} < 600$ GeV, and d) $m_{e\mu bb} > 600$ GeV. The error bars show the statistical uncertainties.

efficiency of the mass reconstruction. The purity and stability for the invariant mass regions, therefore, are calculated and plotted both ways, as shown in Figures 5.32 and 5.33, respectively.

In general the purity should be as high as possible and both the purity and the stability should not vary much with Q_0, which is shown in the results. The purity measurement illustrates the effect of the resolution of lepton and jet energies and momentum, and the stability the effect of reconstruction efficiencies. The high purity in the rapidity regions ($> 80.3\%$) shows the small effect of jet and lepton resolutions at selection criteria boundaries. The stability in the rapidity regions ($> 19.1\%$) shows the effect of reconstruction efficiencies, which cancel in the ratio measurement and do not require explicit correction in the analysis. The consistency of the stability validates that event selection efficiencies do not depend strongly on Q_0. The lower values of both purity and stability for the invariant mass regions ($> 33.1\%$ and $> 6.3\%$, respectively) reflect the amount of migration between invariant mass regions, particularly at lower p_T values. Without requiring the events to pass particle and/or reconstruction level selection with an invariant mass *in that region*, the purity and stability of the invariant mass regions are much closer to that of the rapidity regions, as expected ($> 79.9\%$ and $> 19.3\%$, respectively).
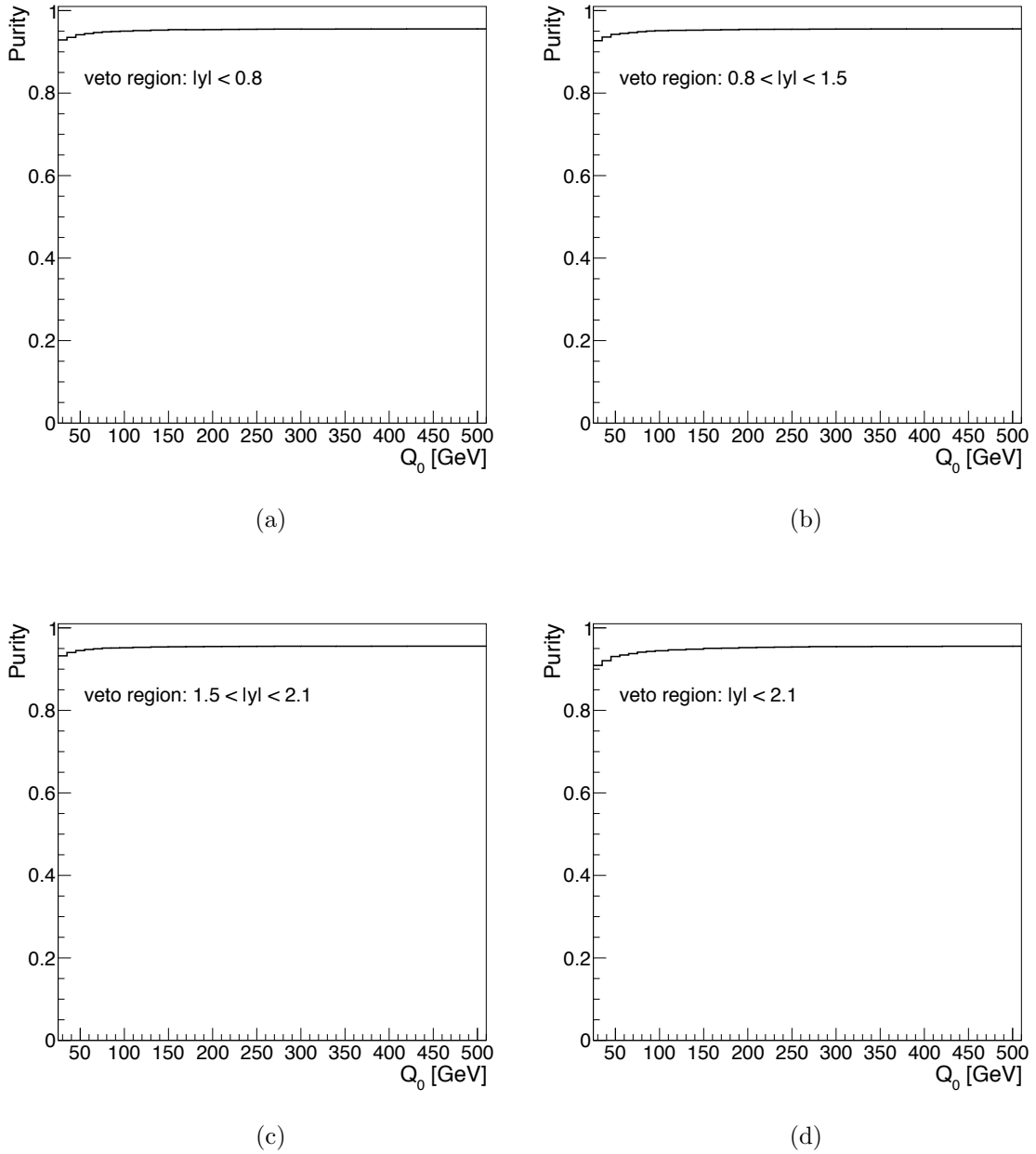
Figure 5.30: The purity $P(Q_0)$ of the gap fraction measurement, simulated with the POWHEG+PYTHIA sample for jets in the rapidity regions a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.
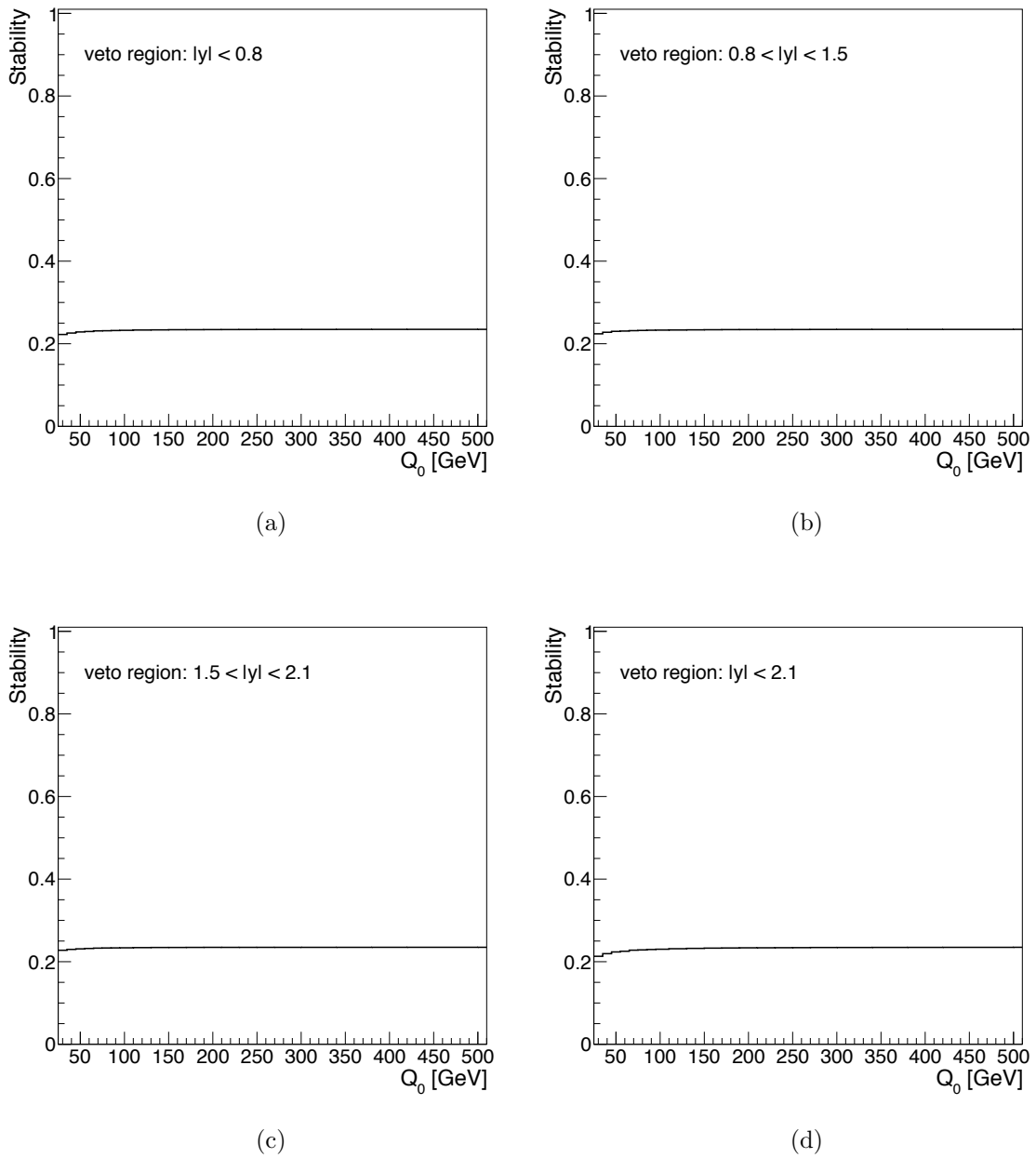
Figure 5.31: The stability $S(Q_0)$ of the gap fraction measurement, simulated with the POWHEG+PYTHIA sample for jets in the rapidity regions a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.

a)

b)

c)

d)
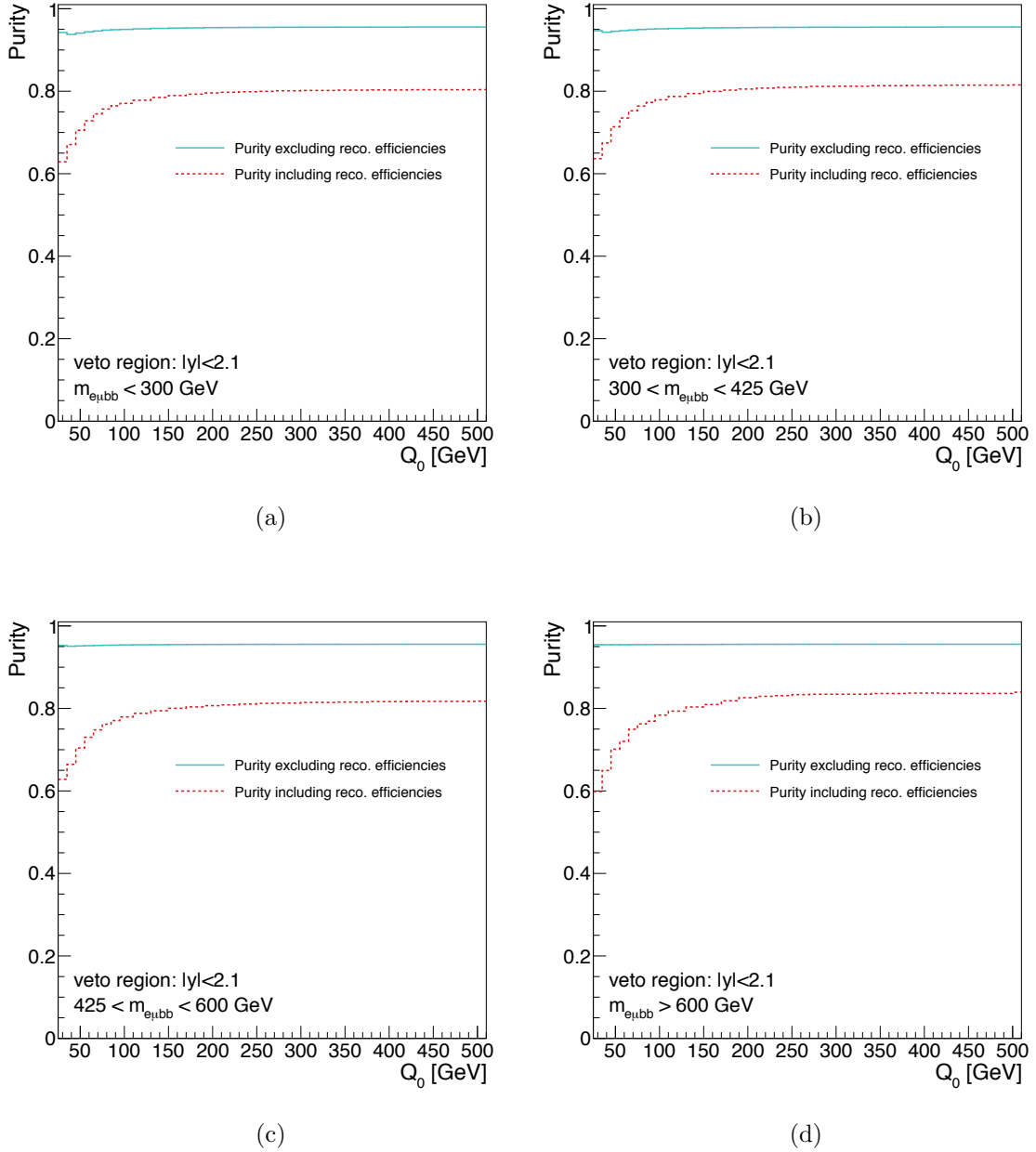
Figure 5.32: The purity $P(Q_0)$ of the gap fraction measurement, simulated with the POWHEG+PYTHIA sample for jets in the rapidity regions a) $m_{e\mu bb} < 300$ GeV, b) $300 < m_{e\mu bb} < 425$ GeV, c) $425 < m_{e\mu bb} < 600$ GeV, and d) $m_{e\mu bb} > 600$ GeV.

(a)

(b)

(c)

(d)
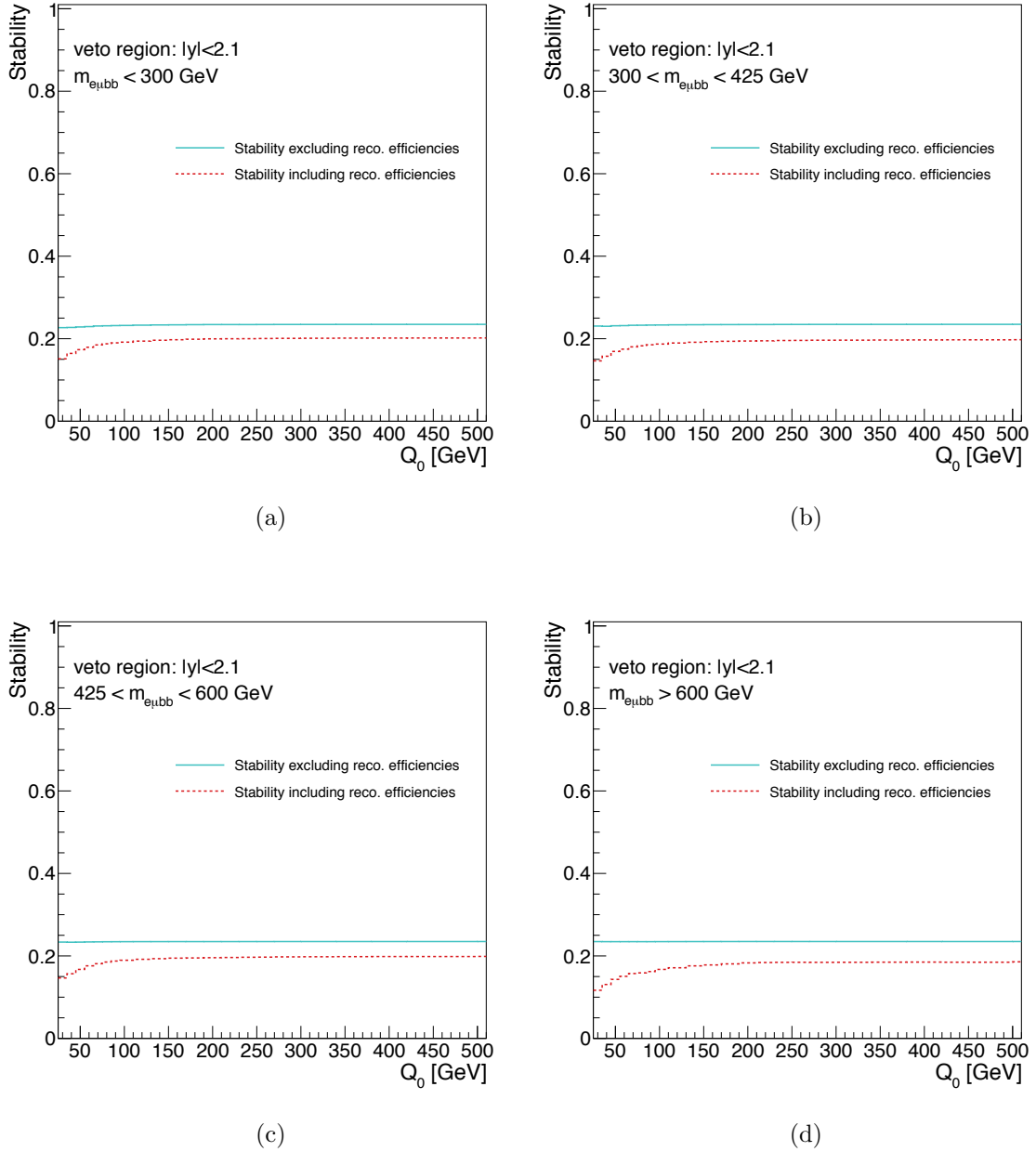
Figure 5.33: The stability $S(Q_0)$ of the gap fraction measurement, simulated with the POWHEG+PYTHIA sample for jets in the rapidity regions a) $m_{e\mu bb} < 300$ GeV, b) $300 < m_{e\mu bb} < 425$ GeV, c) $425 < m_{e\mu bb} < 600$ GeV, and d) $m_{e\mu bb} > 600$ GeV.

5.5 Systematic Uncertainties

Because the gap fraction measurement is a ratio, many of the systematic uncertainties that typically contribute to a $t\bar{t}$ measurement cancel in this case. These include systematic uncertainties in the $t\bar{t}$ production cross-section, luminosity, lepton efficiencies, momentum scale, momentum resolution and reconstruction efficiency, b-jet energy scale and tagging efficiency, and trigger efficiencies.

Instead, the prominent systematic uncertainties on the gap fraction are those that affect the jet activity: systematic uncertainty in the jet energy scale, jet reconstruction efficiency and the effect of additional interactions. Additional uncertainties include uncertainty in $t\bar{t}$ modelling, b-tagging, PDF modelling, and simulation of background.

Most systematic uncertainties in the gap fraction measurement are estimated by varying the corresponding part x of the Monte Carlo simulation, which results in a variation of $f_{\text{reco}}^{\text{MC},x}$ and thus a variation in the correction factor, $\sigma_{x,C[i]}$. The corresponding uncertainty in the particle-level gap fraction measurement follows for each step i of Q_0 and Q_{sum}, using the shorthand $f_{\text{gap}}^{\text{reco}} = f_{\text{reco}}$ and $f_{\text{gap}}^{\text{part}} = f_{\text{part}}$:

$$\sigma_x \left(f_{\text{part}}^{\text{data}}[i] \right) = \sigma_{x,C[i]} \cdot f_{\text{reco}}^{\text{data}}[i] \tag{5.10}$$

where the uncertainty in the correction factor due to each factor x is defined as:[1]

$$\sigma_{x,C[i]} = \frac{f_{\text{part}}^{\text{MC},\,x}[i]}{f_{\text{reco}}^{\text{MC},\,x}[i]} - \frac{f_{\text{part}}^{\text{MC}}[i]}{f_{\text{reco}}^{\text{MC}}[i]} \tag{5.11}$$

Some systematic uncertainties, such as the jet energy scale uncertainty in Section 5.5.1, include uncertainty from several different factors j simulated separately. In

---

[1]In Equation 5.11, the truth ("part") level gap fraction from MC simulation does not always change, but it is described here with the x designation since occasionally it does change.

this case, the uncertainty from each factor is estimated separately and then added in quadrature:

$$\sigma_{x,C[i]} = \sqrt{\sum_j \left[\frac{f_{\text{part}}^{\text{MC},j}[i]}{f_{\text{reco}}^{\text{MC},j}[i]} - \frac{f_{\text{part}}^{\text{MC}}[i]}{f_{\text{reco}}^{\text{MC}}[i]} \right]^2} \tag{5.12}$$

Finally, the systematic uncertainties in the following sections are shown as fractional uncertainties:

$$\frac{\sigma_x \left(f_{\text{part}}^{\text{data}}[i] \right)}{f_{\text{part}}^{\text{data}}[i]} = \sigma_{x,C[i]} \cdot \frac{f_{\text{reco}}^{\text{data}}[i]}{f_{\text{part}}^{\text{data}}[i]} = \sigma_{x,C[i]} \cdot \frac{f_{\text{reco}}^{\text{data}}[i]}{C[i] \cdot f_{\text{reco}}^{\text{data}}[i]} = \frac{\sigma_{x,C[i]}}{C[i]} \tag{5.13}$$

All uncertainties are shown as a function of Q_0 for the four rapidity regions. For brevity, only selected uncertainties with additional significant nuances are also shown as a function of Q_{sum} and/or for the invariant mass regions. Summary plots showing all uncertainties as functions of each Q_0 and Q_{sum} for the four rapidity regions and the four invariant mass regions are included in Section .

5.5.1 Uncertainty in the Jet Energy Scale

Uncertainty in the jet energy scale (JES) was simulated with the `JetUncertainties` package using the `MultiJESEnergyScaleProvider` [85]. This package estimates uncertainty in the calibration method, uncertainty from nearby jets, uncertainty due to different response of quark/gluon initiated jets, and uncertainty from additional soft activity from multiple proton-proton interactions. The `AnalysisTop-1.8.0` code uses this package to reproduce the $t\bar{t} \to W^{\pm}W^{\mp}bb$ simulation from the POWHEG+PYTHIA sample for $\pm 1\sigma$ JES uncertainty variations due to each contributing factor. The event selections were then applied to each variation, the gap fractions calculated, and the resulting correction factor variations added in quadrature as defined in Section

5.5 to find the total $\pm 1\sigma$ fractional uncertainty due to the jet energy scale. The uncertainty benefited from improved jet energy scale calibration made available by the `JetETMiss` group with the package `JetUncertainties-00-08-25`; Figure 5.34 shows the resulting uncertainty of $\leq 2.5\%$ in contrast with the uncertainty without the improved `JetETMiss` calibration (`JetUncertainties-00-08-07`), a relative decrease in uncertainty of up to 85%.

5.5.2   Uncertainty from Additional pp Interactions and Jet Splitting

Additional pp collisions and jets that are split during the reconstruction process can contribute both extra jets and extra jet energy to $t\bar{t}$ events, introducing an uncertainty in the gap fraction. The effect of the extra jet energy contributions is included in the jet energy scale (JES) systematic in Section 5.5.1, and the effect of extra jets is discussed here.

First, the effect of additional jets is mitigated by the application of the JVF cut (defined in Section 5.5.3). As an investigation into the effect of the JVF cut, the gap fraction was measured with the POWHEG+PYTHIA simulation sample as a function of the average number of pp interactions, $<\mu>$, for selected Q_0 values and with various JVF cuts. With no JVF cut, the gap fraction should decrease as $<\mu>$ increases due to the inclusion of increased number of jets from pileup. Including a JVF cut should decrease the number of pileup jets, especially at higher $<\mu>$ values which would presumably include more additional pp interactions, and thus increase the gap fraction. The results with various JVF cuts confirm these trends, as shown in Figure 5.35. To
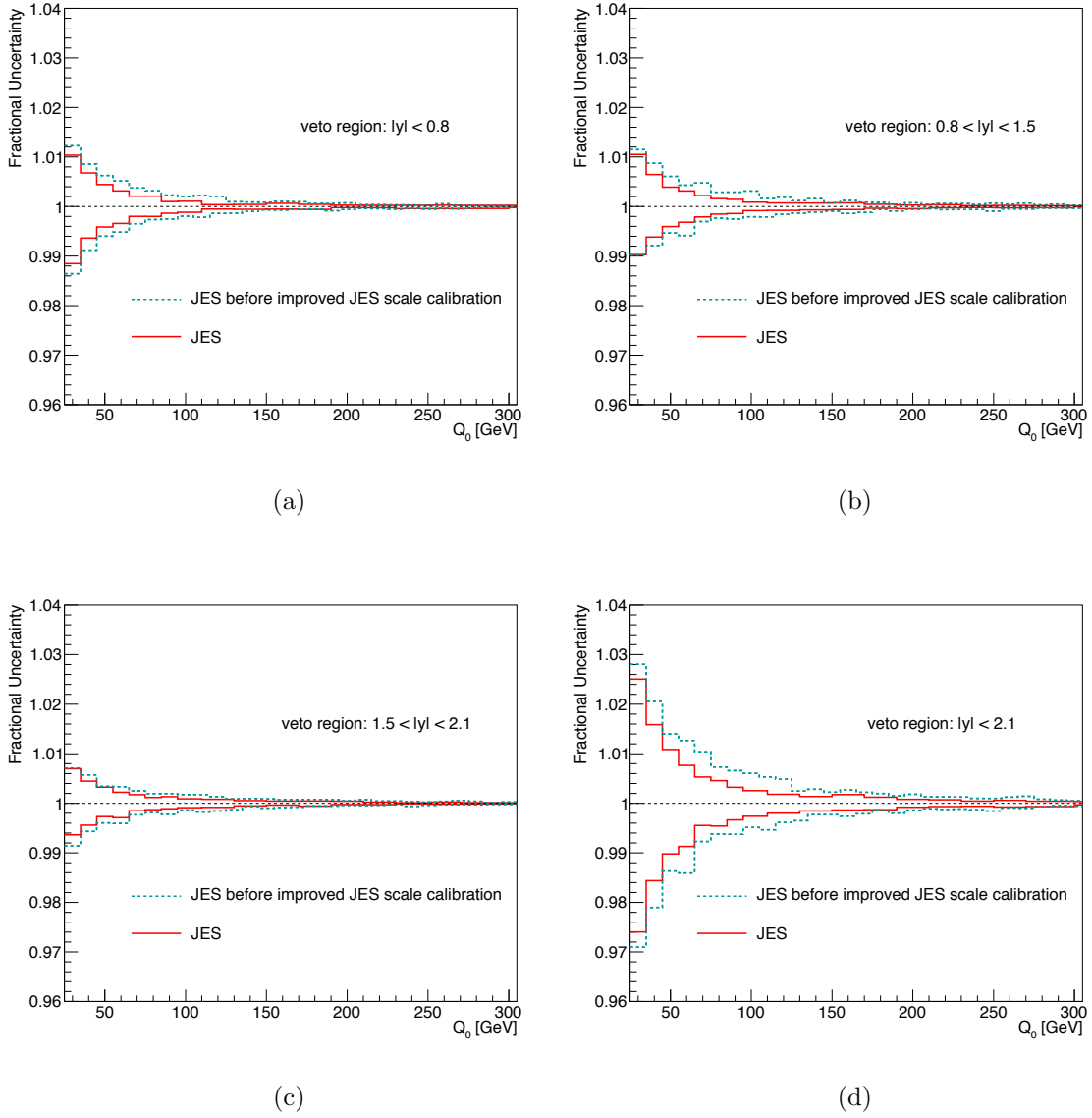
Figure 5.34: Fractional uncertainty in the gap fraction due to jet energy scale as a function of Q_0 for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$. The resulting fractional uncertainty using improved jet energy scale calibration from the `JetETMiss` group is displayed in comparison to the uncertainty without that improved handling.

(a)　　　　　　　　　　　　　　　　(b)

(c)　　　　　　　　　　　　　　　　(d)
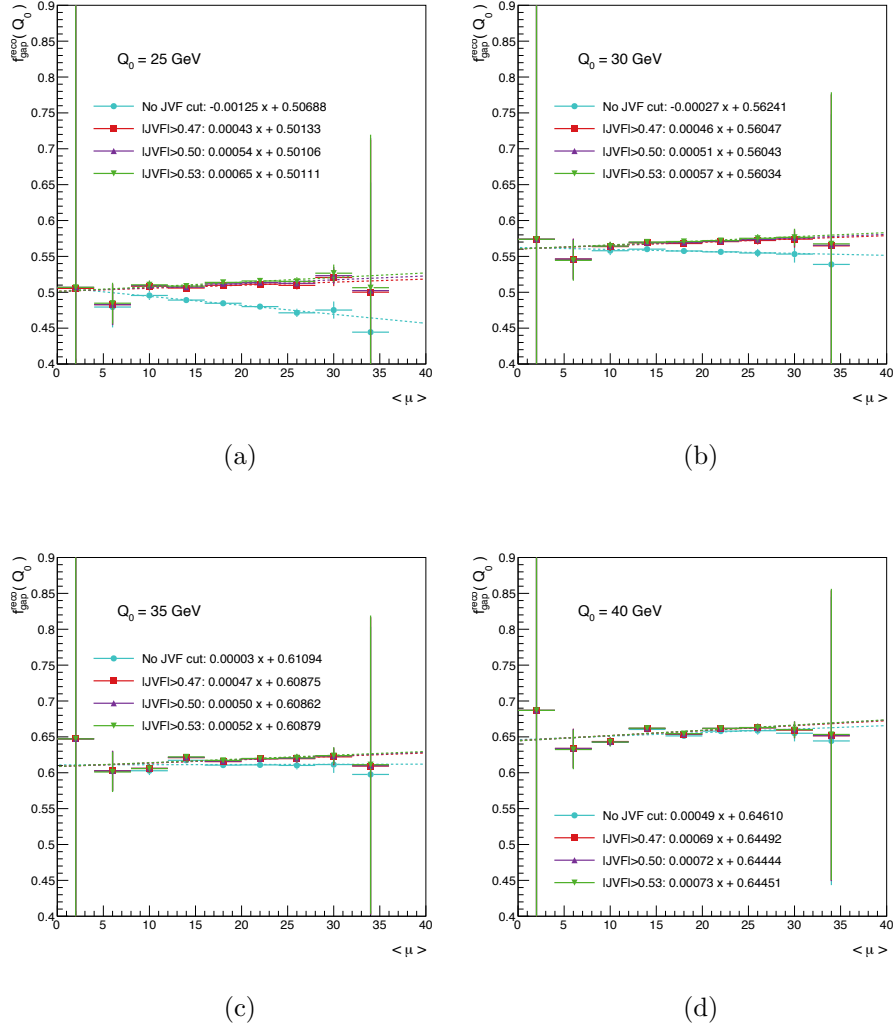
Figure 5.35: The reconstruction-level gap fraction for $|y| < 2.1$ and various Q_0 values as a function of the average number of pp interactions, $<\mu>$, in the POWHEG+PYTHIA simulation. The plots show the gap fraction for a JVF cut of 0.5, the typical value used for the ATLAS 2012 data sample, together with the variations of the cut used for the uncertainty estimation. The dashed lines show the linear fits to the points described in the legend.

understand the efficiency of the JVF cut, the process of jet matching described in Section 4.6 was used with the POWHEG+PYTHIA sample, and ΔR calculated between each reconstructed additional jet and the closest particle-level additional jet. Since pileup jets are not included at the particle level, a reconstructed jet close to a particle-level jet is much more likely to be from a hard-scattering event, while a reconstructed jet far from any particle-level jet (*i.e.*, with a large ΔR value between it and the closest particle-level jet) is likely from pileup. Figure 5.36 shows the ΔR value calculated between each low-energy reconstructed additional jet ($25 < p_\mathrm{T} < 30$ GeV) and the closest particle-level additional jet from the same event, where the particle-level jet was allowed to have transverse momentum of $p_\mathrm{T} > 10$ GeV to avoid resolution issues around the 25 GeV p_T cut on reconstructed jets. The distribution was measured with JVF cuts of 0.1 and 0.50, and the efficiency of the JVF cut as a function of ΔR was calculated as the fraction of jets in a particular ΔR bin which pass the JVF cut. The minimum efficiency for a JVF cut of 0.50 within $\Delta R < 0.4$ is 61%.

To estimate the remaining effect of unmatched jets in the central rapidity region ($|y| < 2.1$), due to both pileup and jet splitting, the gap fraction was recalculated using the POWHEG+PYTHIA simulation, including only "matched" reconstruction-level extra jets, according to the matching definition given in Section 4.6. Compared to the gap fraction without the jet matching restriction, the fractional change accounts for the full effect of unmatched jets, shown in Figure 5.37. This effect is small ($\leq 1.5\%$) and not dominant compared to other systematic uncertainties such as that from the jet energy scale and jet resolution.

(a)   (b)
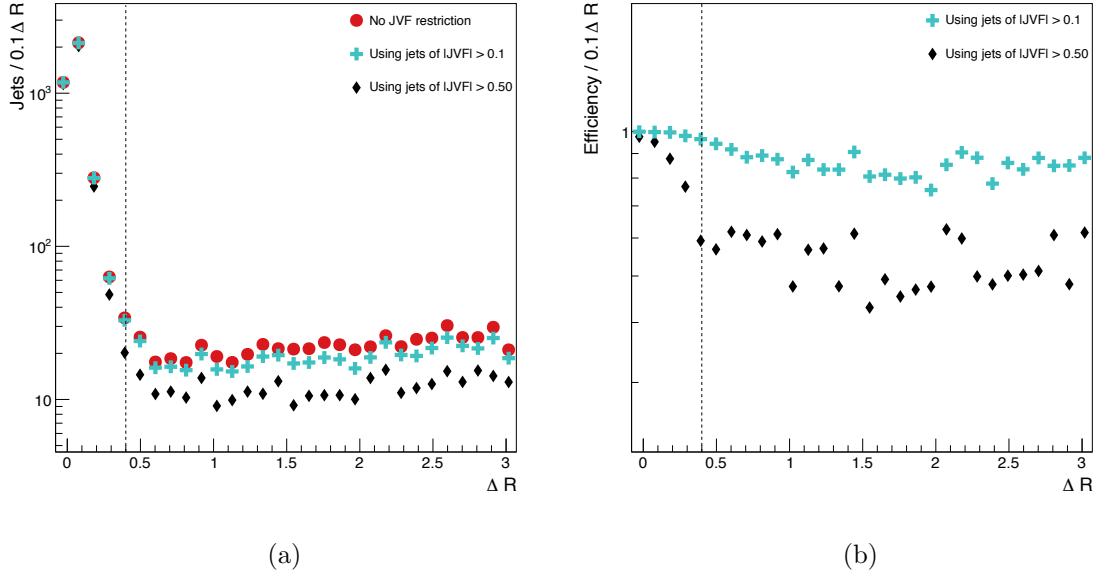
Figure 5.36: Studies of the distance between each reconstructed additional jet and the closest particle-level additional jet, using the POWHEG+PYTHIA sample, showing a) the distance ΔR in θ-ϕ space between each reconstructed jet and its nearest particle-level jet originating from the $t\bar{t}$ event, and b) the corresponding efficiency of the JVF cut as a function of the ΔR measurement used in a). The dashed line shows boundary of $\Delta R < 0.4$ used to consider the jets "matched".
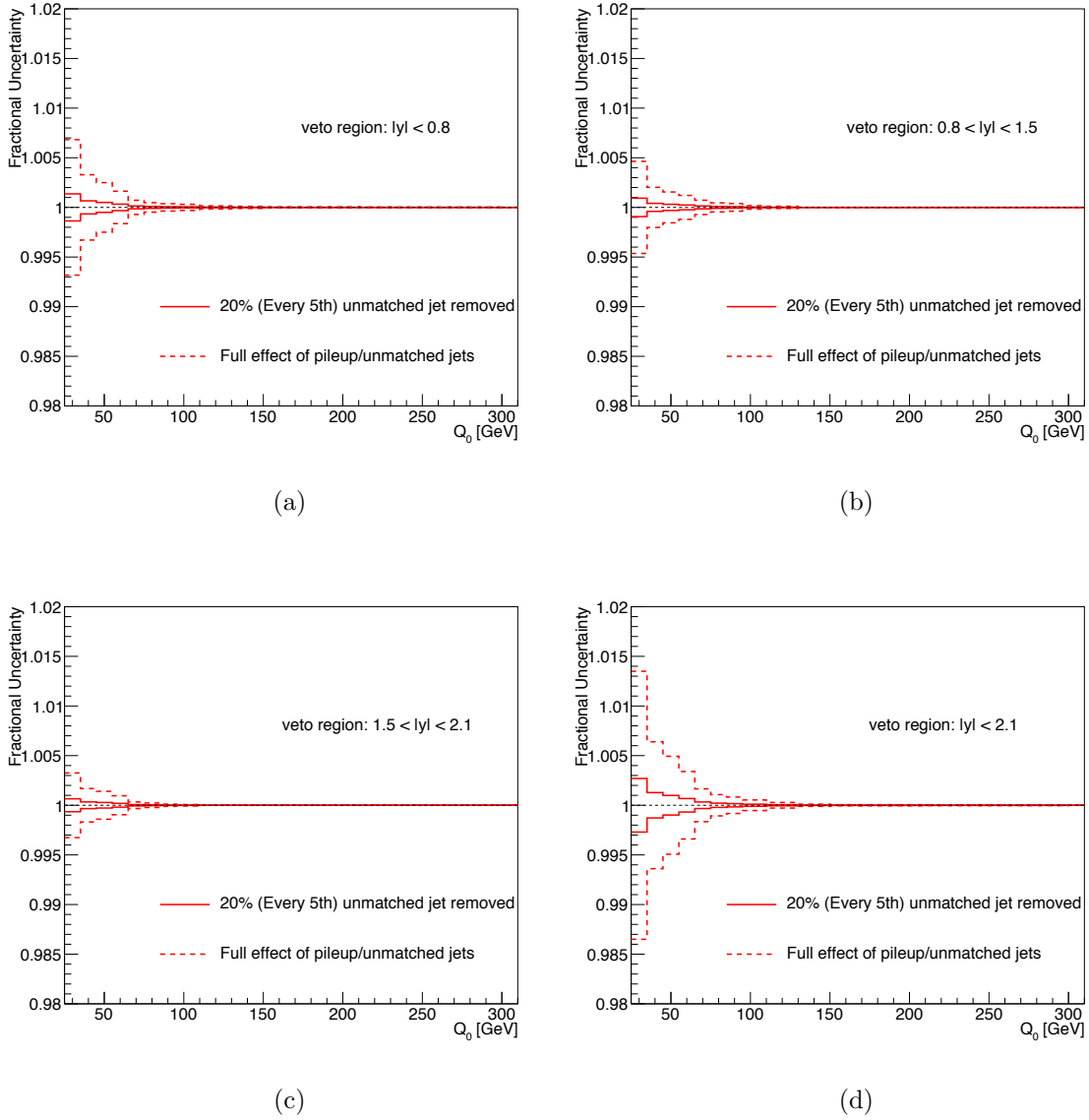
Figure 5.37: Estimation of the full effect of including all unmatched jets (dashed line) in the gap fraction measurement, and the estimated systematic uncertainty in modelling this effect, 20% of the full effect (solid line). The uncertainty is shown as a function of Q_0 for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.

A study of how well the Powheg+Pythia simulation models pileup in data, using zero-bias events, found that the simulation agrees with data in the central rapidity region to a precision of $< 20\%$ in Section 6.1 of Ref. [86]. This analysis compared results and found the same overestimation, shown in Figure 5.38. Therefore the uncertainty due to unmatched jets is taken to be 20% of the full effect.

5.5.3 Uncertainty from the Jet Vertex Fraction

The use of the jet vertex fraction requirement (JVF), as defined in Section 4.1 and discussed as a means of reducing pileup collisions in Section 5.5.2 also introduces an uncertainty into the gap fraction, because simulation of the JVF efficiency for additional jets from QCD radiation may not completely match the JVF efficiency in data.

To estimate this uncertainty, the JVF cut value is varied in simulation but *not* in data. Thus the uncertainty is the difference between the correction factor built with the default JVF cut and the correction factor built using the JVF cut variations. Figure 5.39 gives the result due to varying the JVF cut to $|JVF| < 0.47$ and $|JVF| < 0.53$, according to the ATLAS recommendation [87], as a function of Q_0 for the four rapidity regions.

5.5.4 Uncertainty from b-Tagging

Uncertainty in the method of tagging jets as originating from b quarks is estimated from simulation by rerunning the Powheg+Pythia simulation with b-tagging scale factors
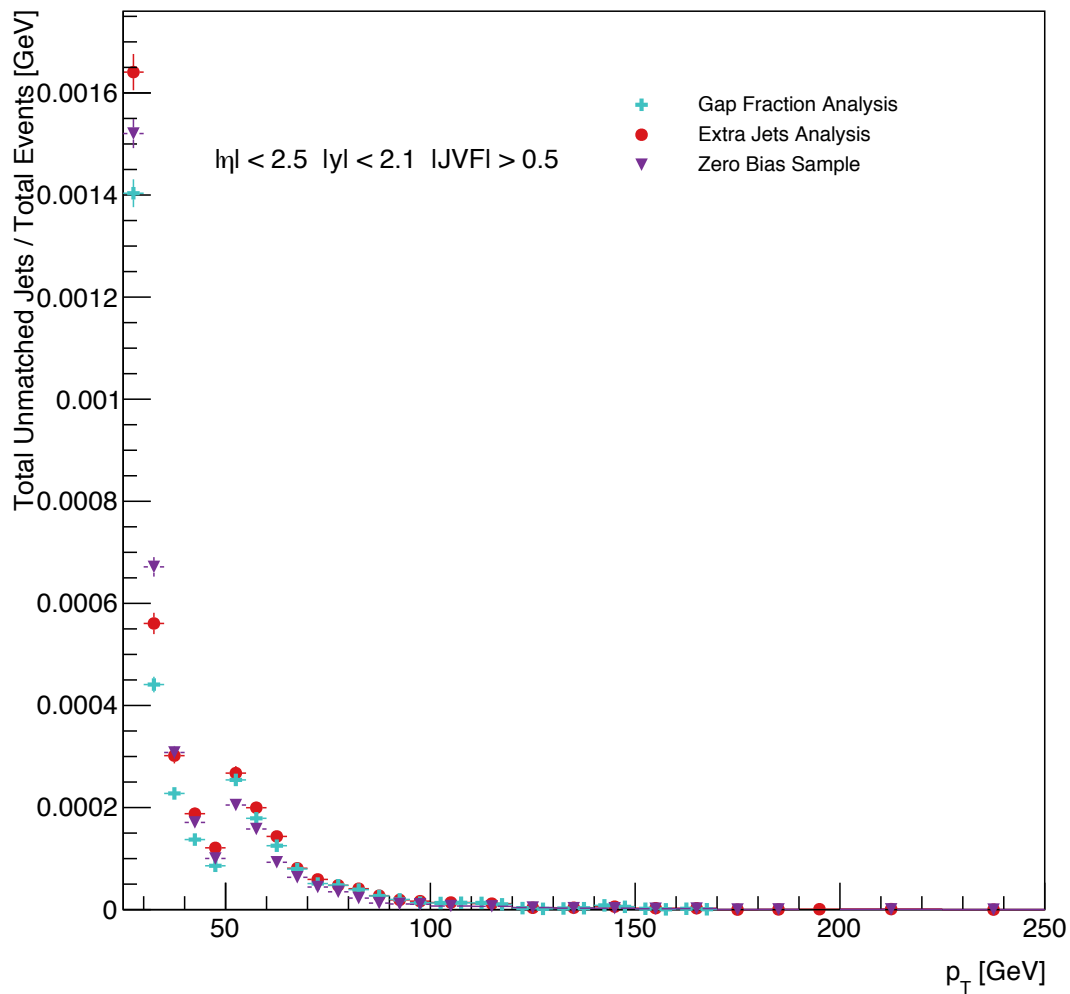
Figure 5.38: A measurement of the p_T spectrum of unmatched jets from a zerio-bias sample described in Ref. [86] compared with simulation in Ref. [86] ("Extra Jet Analysis") and in this analysis ("Gap Fraction Analysis").
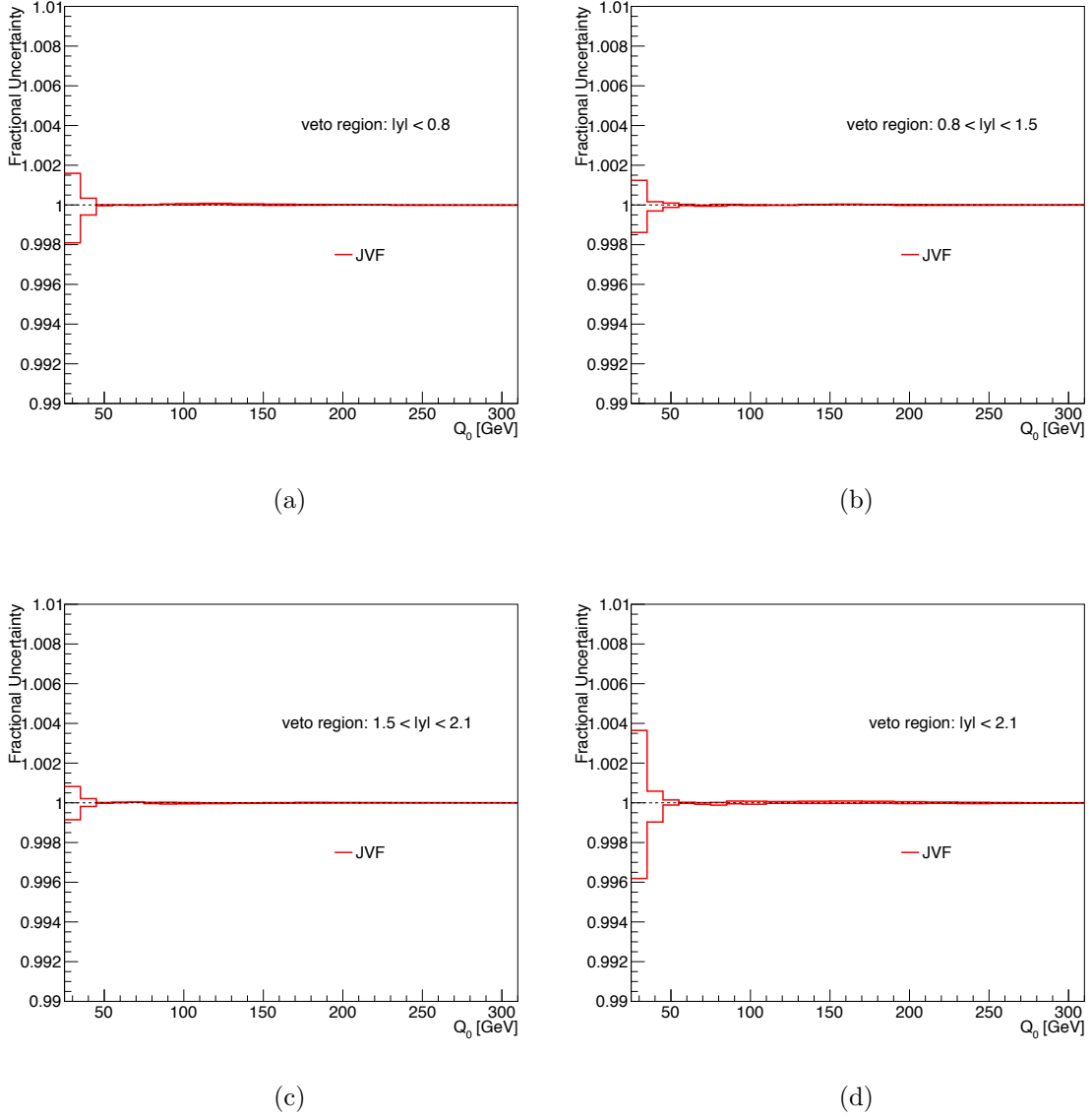
Figure 5.39: Uncertainty from the jet vertex fraction (JVF) restriction on the gap fraction calculation as a function of Q_0 for the four rapidity regions: a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.

applied at $\pm 1\sigma$ variations. The correction factors were calculated with the application of each scale factor, `BTAGSFUP`, `CTAGSFUP`, `MISTAGSFUP` and associated `DOWN` deviations. (No attempt was made to split each of these scale factors into components.) The resulting differences in the correction factor were used to recalculate the uncertainty as described in Section 5.5. The result is small at $< 0.5\%$, as shown in Figure 5.40.

5.5.5 Uncertainty in $t\bar{t}$ modelling

The step-by-step correction procedure discussed in Section 5.4 introduces an uncertainty due to the physics modelling of the event generator. This uncertainty is calculated by varying the NLO generator, the parton showering generator, and the initial/final state radiation (ISR/FSR), and adding the differences in quadrature with the statistical uncertainty from the main generator, POWHEG+PYTHIA. Care was taken not to compare `fastsim` with `fullsim` samples. The size of the correction factors built from the different generators is given in Section 5.4 in Figures 5.26, 5.27, 5.28 and 5.29.

Calculating the effect of varying the initial and final state radiation (ISR/FSR) was done by finding half the difference in the correction factors of the increased and decreased radiation samples of the ACERMC+PYTHIA generator as shown in Figure 5.41, of the ALPGEN+PYTHIA generator in Figure 5.42, and of the MAD-GRAPH+PYTHIA generator in Figure 5.43. Since all three ISR/FSR samples are considered state-of-the-art, the most recent ALPGEN+PYTHIA generator result was chosen as the ISR/FSR contribution to the uncertainty in $t\bar{t}$ modelling. The effect of
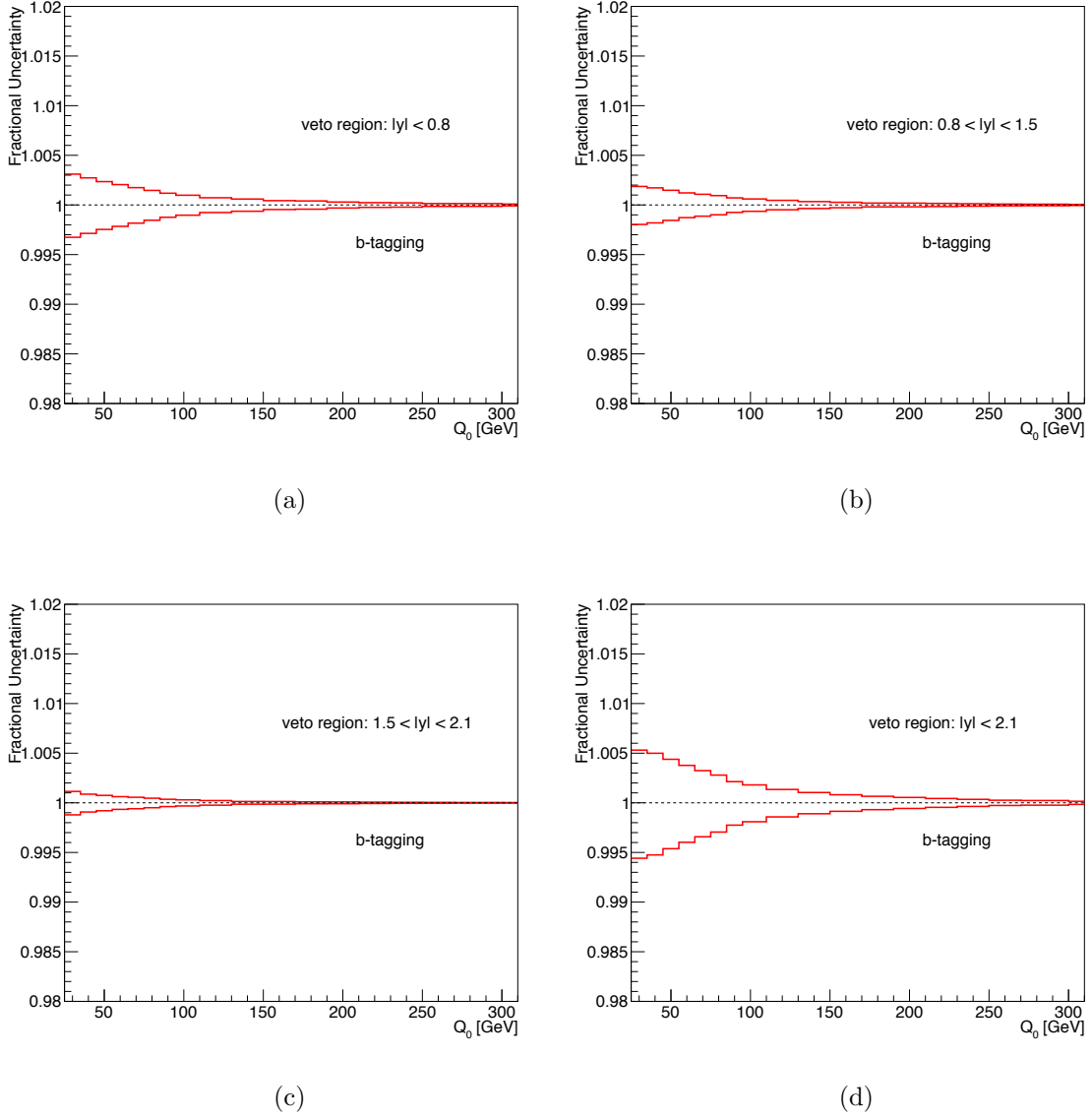
Figure 5.40: Uncertainty in the gap fraction due to b-tagging effects as a function of Q_0. The results are shown for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.
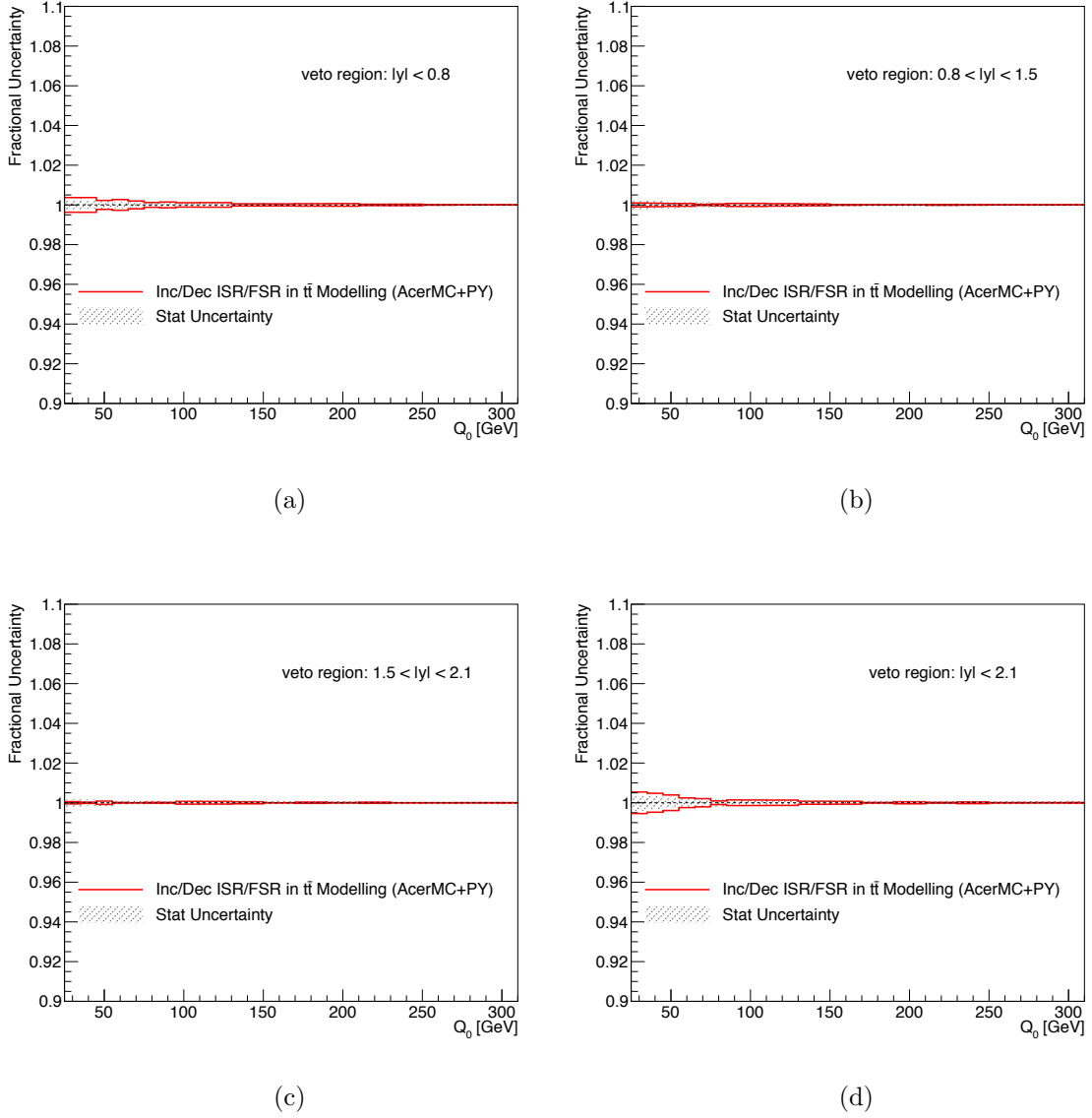
Figure 5.41: Half the difference between correction factors built from the AC-ERMC+PYTHIA increased and decreased radiation samples, for the four rapidity regions as a function of Q_0, as an aspect of the systematic uncertainty in $t\bar{t}$ modelling. The shaded band shows the statistical uncertainty in POWHEG+PYTHIA. The results shown are for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.
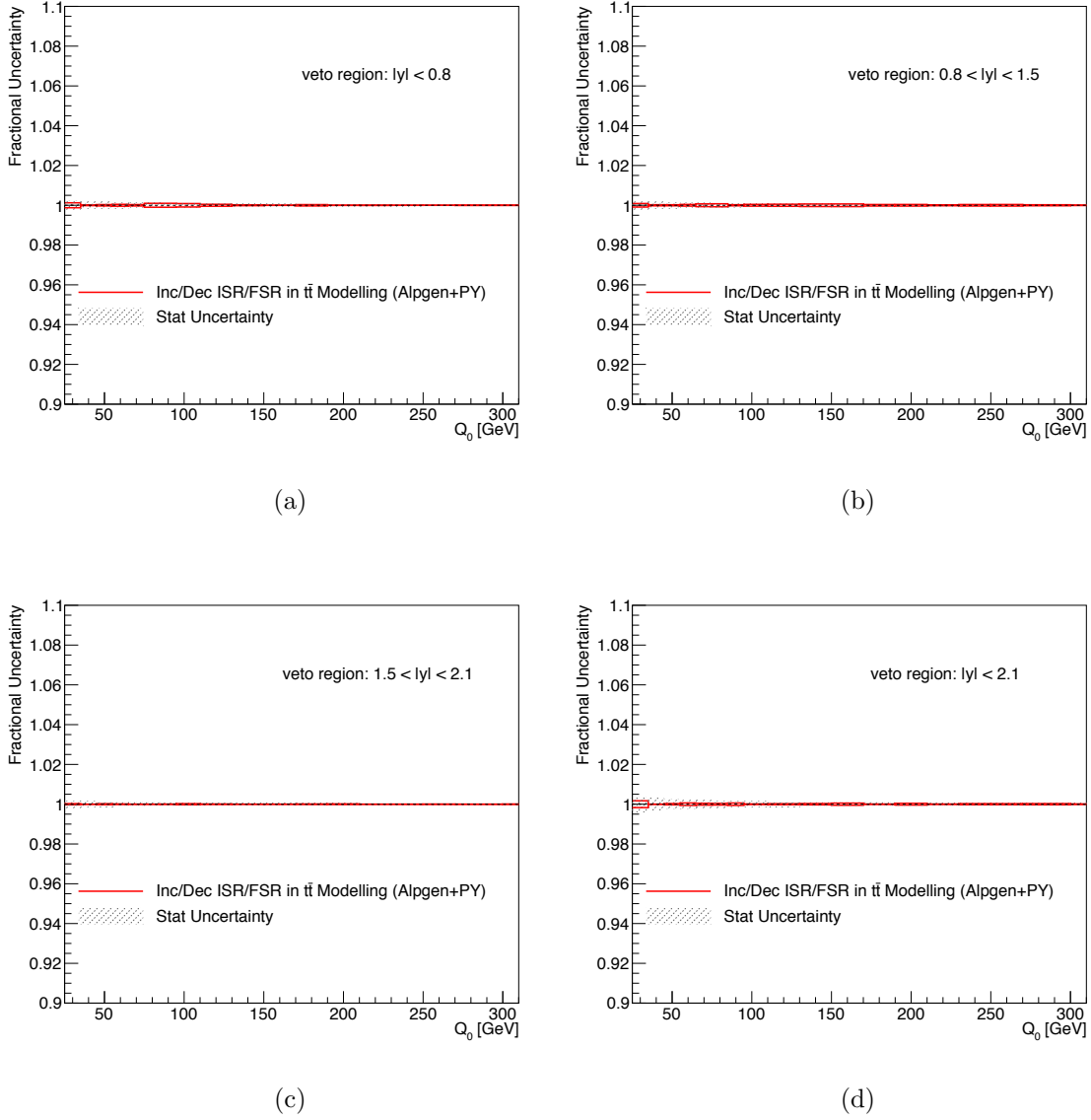
Figure 5.42: Half the difference between correction factors built from the ALP-GEN+PYTHIA increased and decreased radiation samples, for the four rapidity regions as a function of Q_0, as an aspect of the systematic uncertainty in $t\bar{t}$ modelling. The shaded band shows the statistical uncertainty in POWHEG+PYTHIA. The results shown are for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.
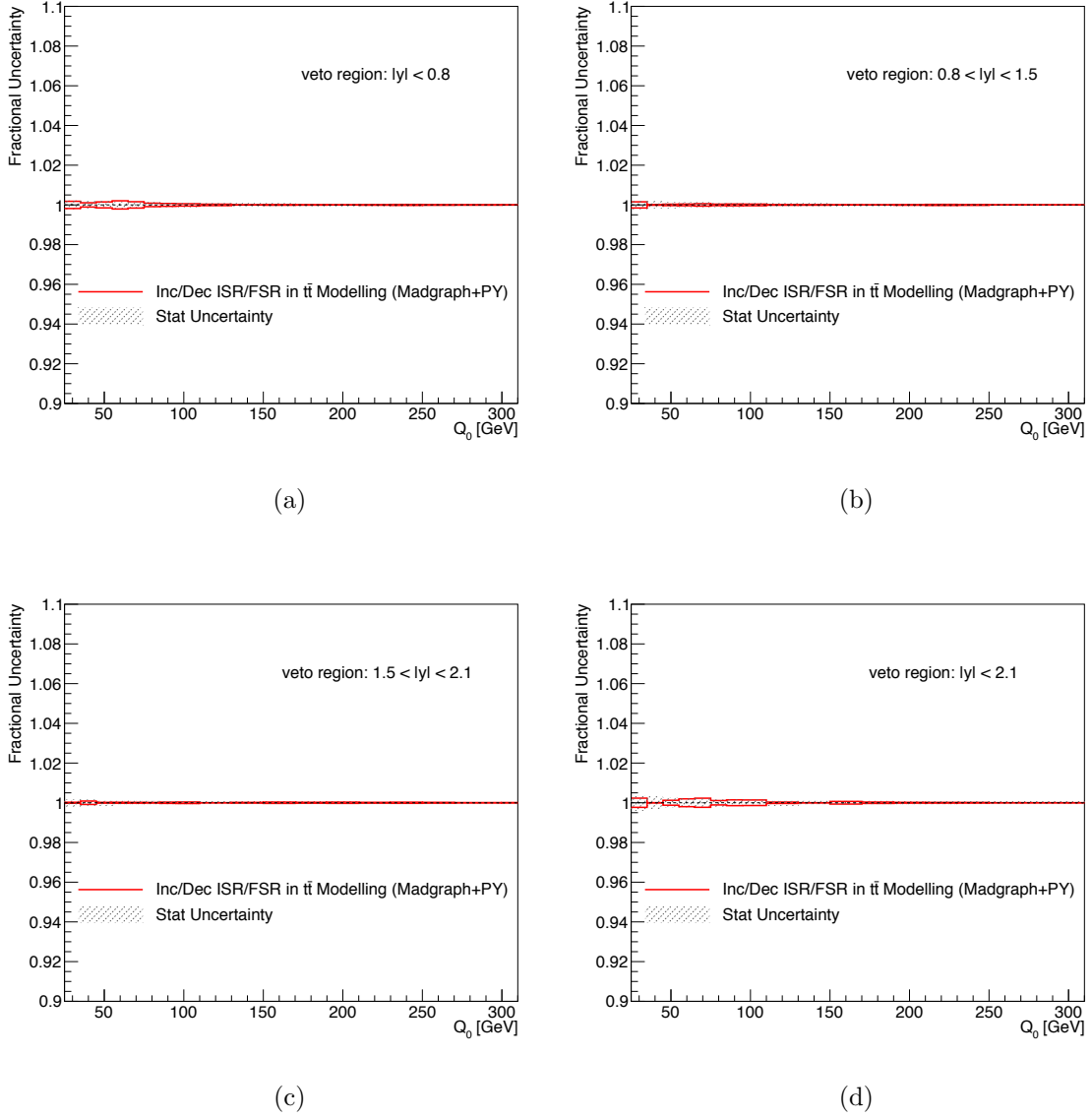
Figure 5.43: Half the difference between correction factors built from the MAD-GRAPH+PYTHIA increased and decreased radiation samples, for the four rapidity regions as a function of Q_0, as an aspect of the systematic uncertainty in $t\bar{t}$ modelling. The shaded band shows the statistical uncertainty in POWHEG+PYTHIA. The results shown are for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.

the NLO generator was measured as the difference in the correction factor built from POWHEG+HERWIG compared with that built from MC@NLO+HERWIG; the effect of the parton showering was found as the difference in the correction factors built from POWHEG+PYTHIA and POWHEG+HERWIG. These differences are included in Figure 5.44 and 5.45, along with the difference in correction factors due to ISR/FSR modelling discussed above, and the difference in the correction factors built using $h_{\mathrm{damp}} = m_t$ and $h_{\mathrm{damp}} = \infty$, using POWHEG+PYTHIA and POWHEG+PYTHIA ($h_{\mathrm{damp}} = m_t$). The total uncertainty due to $t\bar{t}$ modelling was found by adding in quadrature the uncertainty due to the NLO generator, the parton showering, and the initial/final state radiation, along with the statistical uncertainty from the POWHEG+PYTHIA generator (neglecting any of the three components for a given veto region if they are smaller than the statistical uncertainty at every point in that veto region). Figure 5.46 shows the result for the four rapidity regions as a function of Q_0, and Figure 5.47 for the four mass regions as a function of Q_0.

5.5.6 Comparison with Undressed Leptons

This analysis uses "dressed" charged leptons at the particle level, which includes the momentum of all photons radiated within a cone of $\Delta R < 0.1$ around around the lepton in calculating the lepton's four-vector momentum to account for final-state radiation from the lepton. A comparison was done by re-measuring the gap fraction using "bare" leptons for the measurement to exclude these radiated photons. The difference was found to be negligible, as shown in Figure 5.48 (still included as a
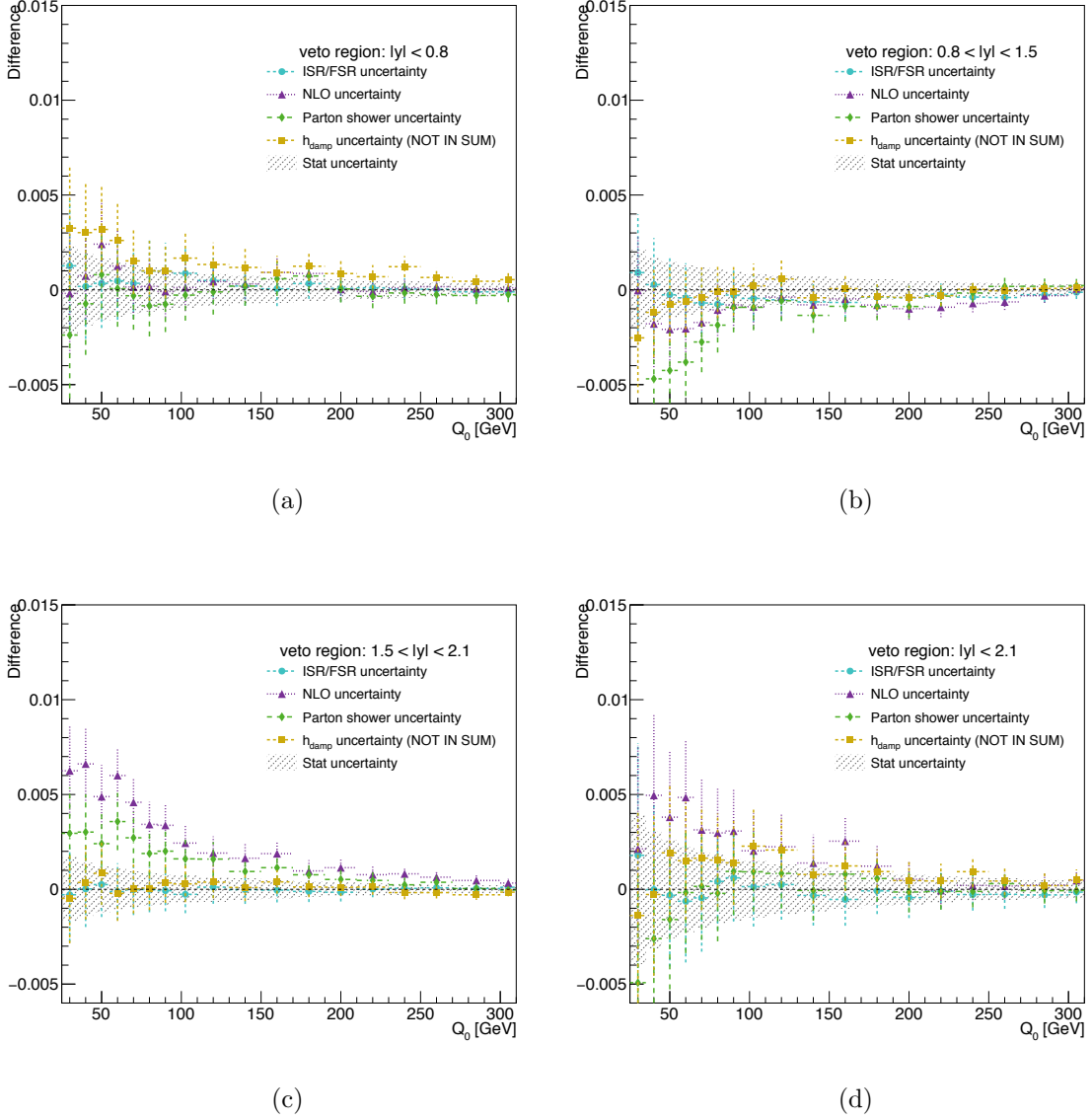
Figure 5.44: The absolute difference between correction factors built with two different NLO generators, two different parton showering generators, increased and decreased ISR/FSR, and two different h_{damp} definitions, m_t and ∞, along with the statistical uncertainty of each combination and the statistical uncertainty of the standard correction factor (built with POWHEG+PYTHIA). The results shown are for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.
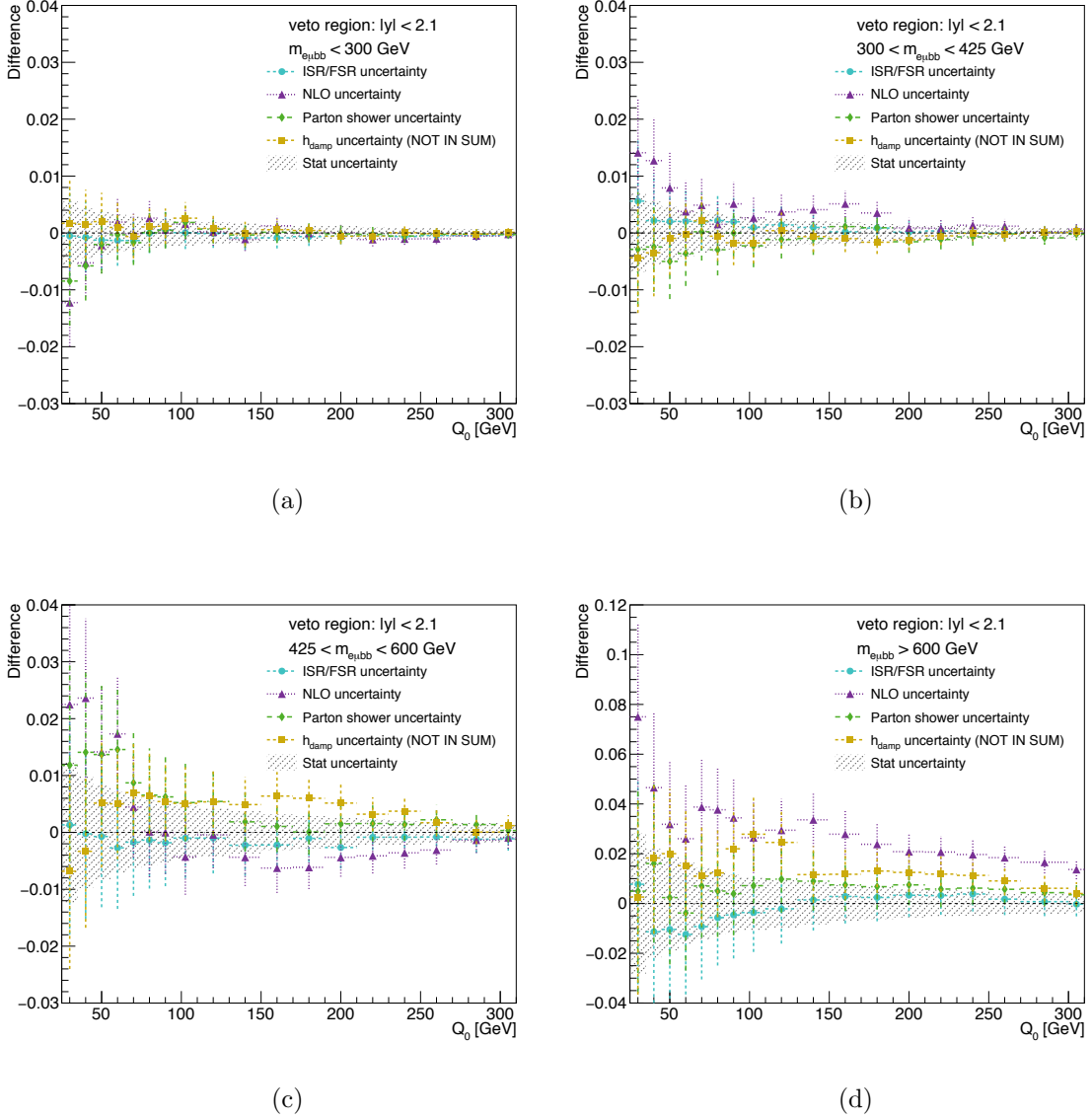
Figure 5.45: The absolute difference between correction factors built with two different NLO generators, two different parton showering generators, increased and decreased ISR/FSR, and two different h_{damp} definitions, m_t and ∞, along with the statistical uncertainty of each combination and the statistical uncertainty of the standard correction factor (built with POWHEG+PYTHIA). The results shown are for a) $m_{e\mu bb} < 300$ GeV, b) $300 < m_{e\mu bb} < 425$ GeV, c) $425 < m_{e\mu bb} < 600$ GeV, and d) $m_{e\mu bb} > 600$ GeV.

131
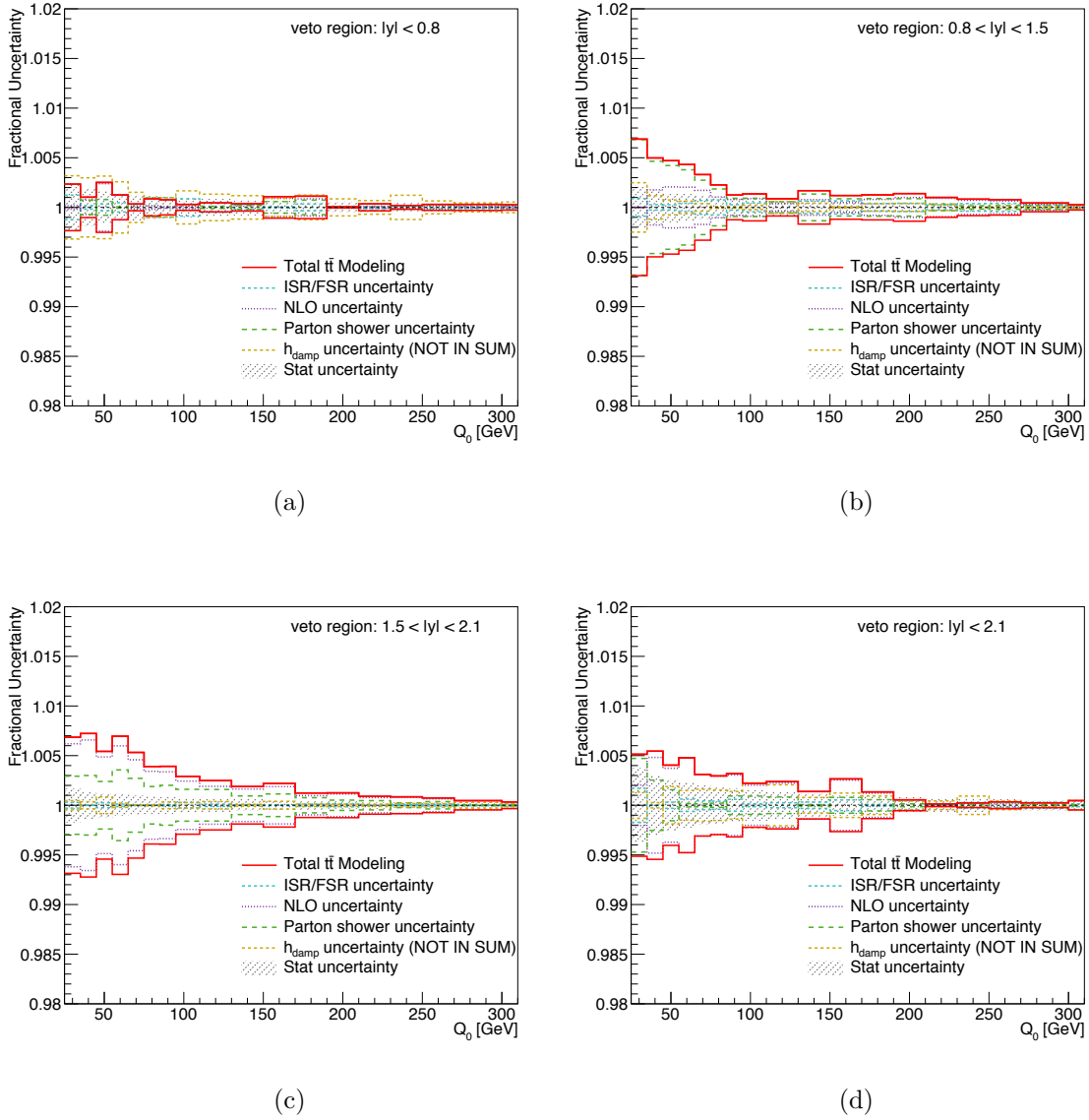
Figure 5.46: Uncertainty in the gap fraction due to $t\bar{t}$ modelling for the four rapidity regions as a function of Q_0, from each contributing aspect and the quadrature sum. The shaded band shows the statistical uncertainty in POWHEG+PYTHIA. The results shown are for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.
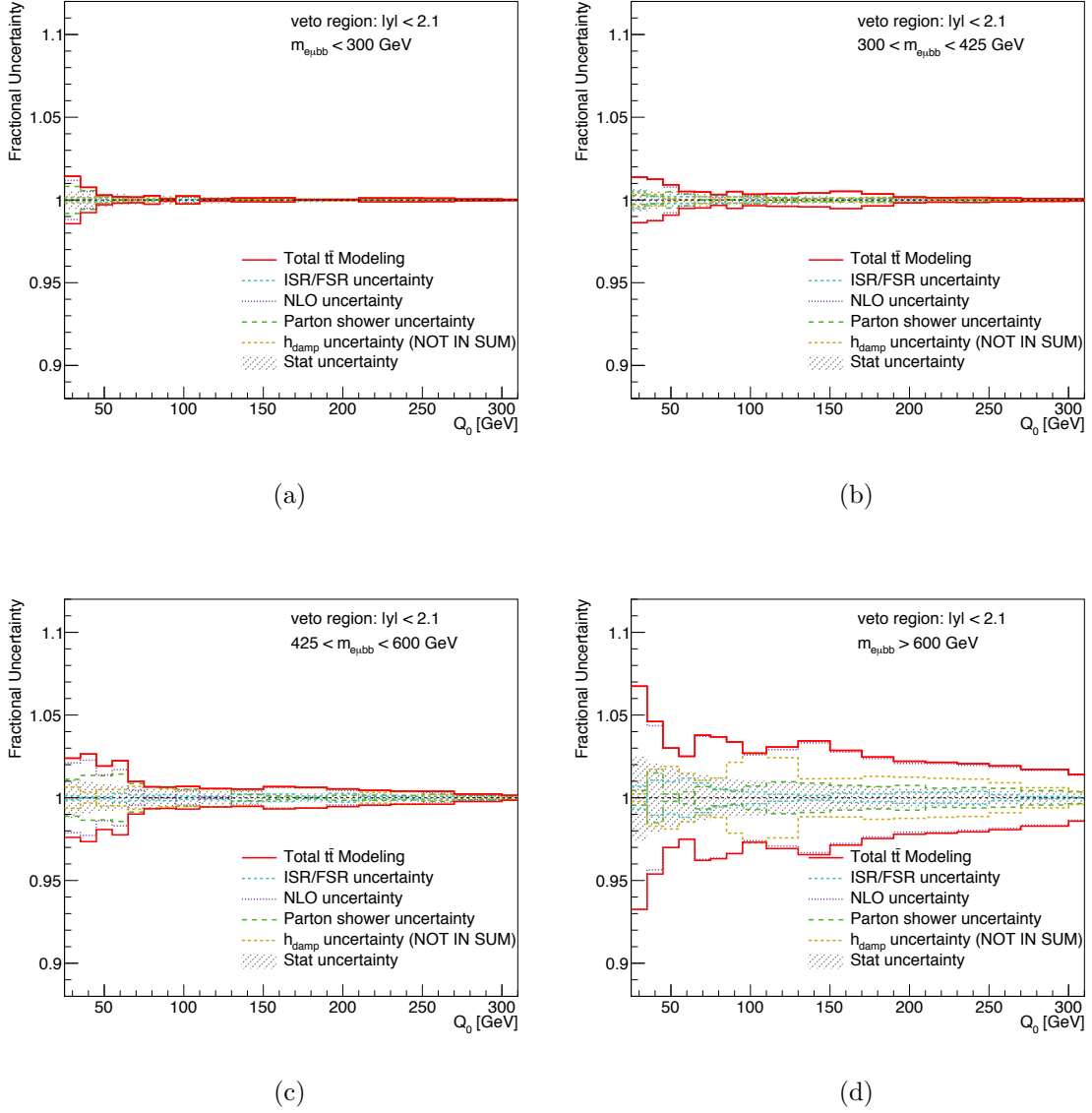
Figure 5.47: Uncertainty in the gap fraction due to $t\bar{t}$ modelling as a function of Q_0 for the four mass regions, from each contributing aspect and the quadrature sum. The shaded band shows the statistical uncertainty in POWHEG+PYTHIA. The results are shown for a) $m_{e\mu bb} < 300$ GeV, b) $300 < m_{e\mu bb} < 425$ GeV, c) $425 < m_{e\mu bb} < 600$ GeV, and d) $m_{e\mu bb} > 600$ GeV.

systematic uncertainty).

5.5.7   Uncertainty in Background Contributions

The amount of each contributing background (single top events, events with mis-identified leptons, diboson events and Z+jet events) was estimated using simulation as described in Section 3.2. To estimate the associated uncertainty, the amount of each estimated background was increased and decreased by 100% before the Monte Carlo simulation is scaled to data as described in Section 3.2, the estimated backgrounds subtracted from data, and the gap fraction from data recalculated accordingly. Since only the gap fraction from data, and not the correction factor, changes with the variation of subtracted background, the estimated uncertainty follows as:

$$\frac{\sigma_x\left(f_{\text{part}}^{\text{data}}[i]\right)}{f_{\text{part}}^{\text{data}}[i]} = \frac{C[i] \cdot \sigma_x\left(f_{\text{reco}}^{\text{data}}[i]\right)}{f_{\text{part}}^{\text{data}}[i]} = \frac{C[i] \cdot \sigma_x\left(f_{\text{reco}}^{\text{data}}[i]\right)}{C[i] \cdot f_{\text{reco}}^{\text{data}}[i]} = \frac{\sigma_x\left(f_{\text{reco}}^{\text{data}}[i]\right)}{f_{\text{reco}}^{\text{data}}[i]} \tag{5.14}$$

$Wt+b$ events can interfere with $t\bar{t}$ events having the final state $WbWb$ [88, 71], and the generators' handling of this interference by either a "diagram removal" (DR) or "diagram subtraction" (DS) scheme adds an additional uncertainty in the Wt background. (The choice of schema can also affect the number of mis-identified leptons slightly.) This effect was studied by comparing the gap fraction from data with the background subtracted using the DS scheme and the background subtracted using the DR scheme. The results, shown in Figure 5.49 for the rapidity regions and Figure 5.50 for the mass regions, are the same size or smaller than the uncertainty from the amount of background discussed above for each rapidity and mass region.

The total uncertainty due to background, calculated as the quadrature sum of

(a)   (b)

(c)   (d)

Figure 5.48: Uncertainty in the gap fraction due to using dressed charged leptons, as a function of Q_0. The results are shown for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.

the difference due to the Wt scheme choice and the differences from varying each background by $\pm 100\%$, are shown in Figure 5.51 as a function of Q_0 for the four rapidity regions. The uncertainty from backgrounds for the four mass regions is approximately commensurate in size, with the exception of $\leq 5\%$ in the highest mass region.

5.5.8 Uncertainty in Lepton Efficiencies and Resolution

Variations of the POWHEG+PYTHIA simulation were calculated for $\pm 1\sigma$ variations in selecting and triggering on the electrons and muons using the `AnalysisTop-1.8.0` software. The event selection and correction factor were rebuilt with these variations and the fractional differences in the correction factor added in quadrature as in Section 5.5 as the systematic uncertainty. This resulting uncertainty was negligible, as shown in Figure 5.52 (again, still included as a systematic uncertainty).
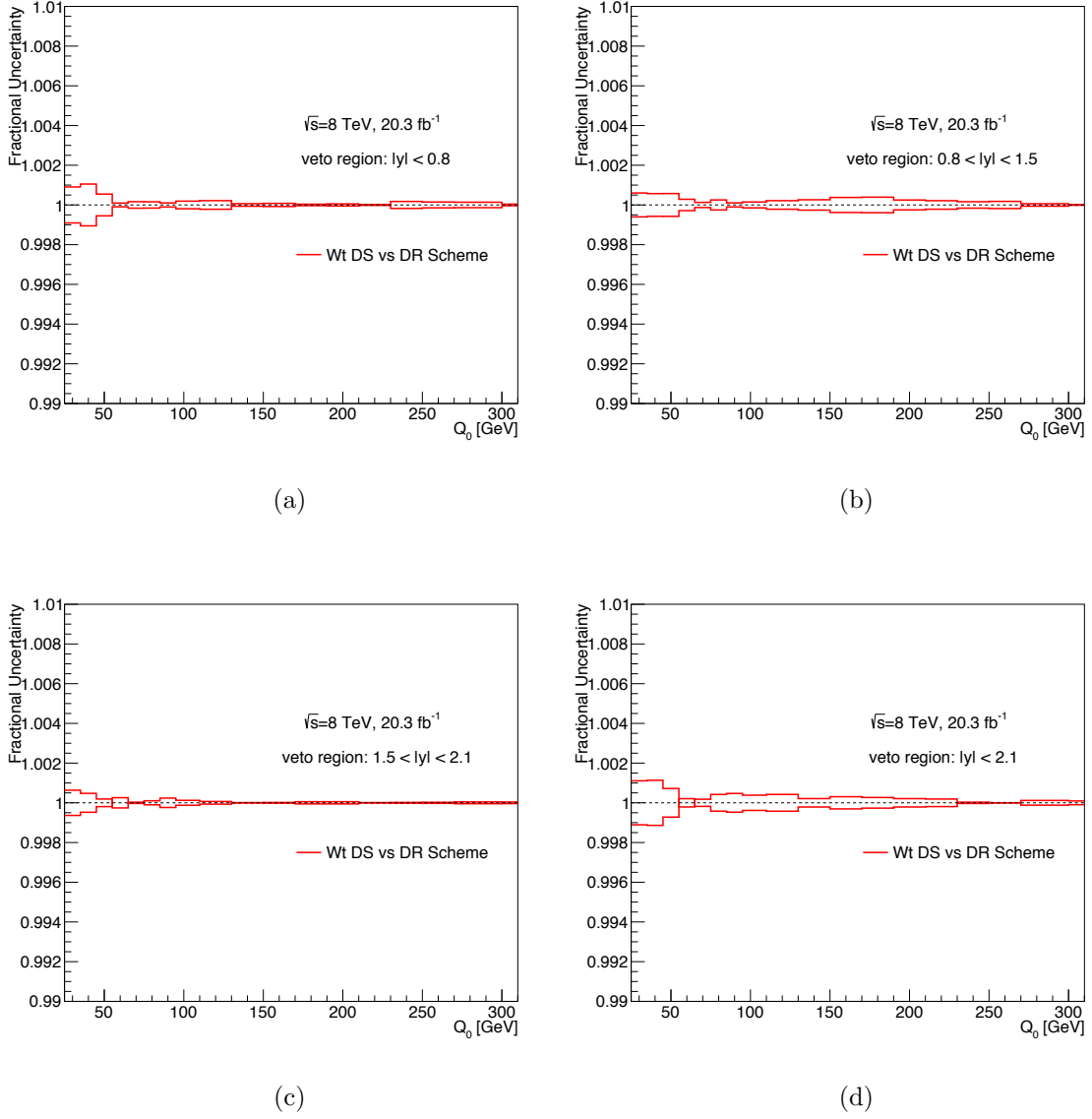
Figure 5.49: Uncertainty in the gap fraction due to the choice of schema used to model $Wt+b$ event interference with $t\bar{t}$ events, as a function of Q_0. The results are shown for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.
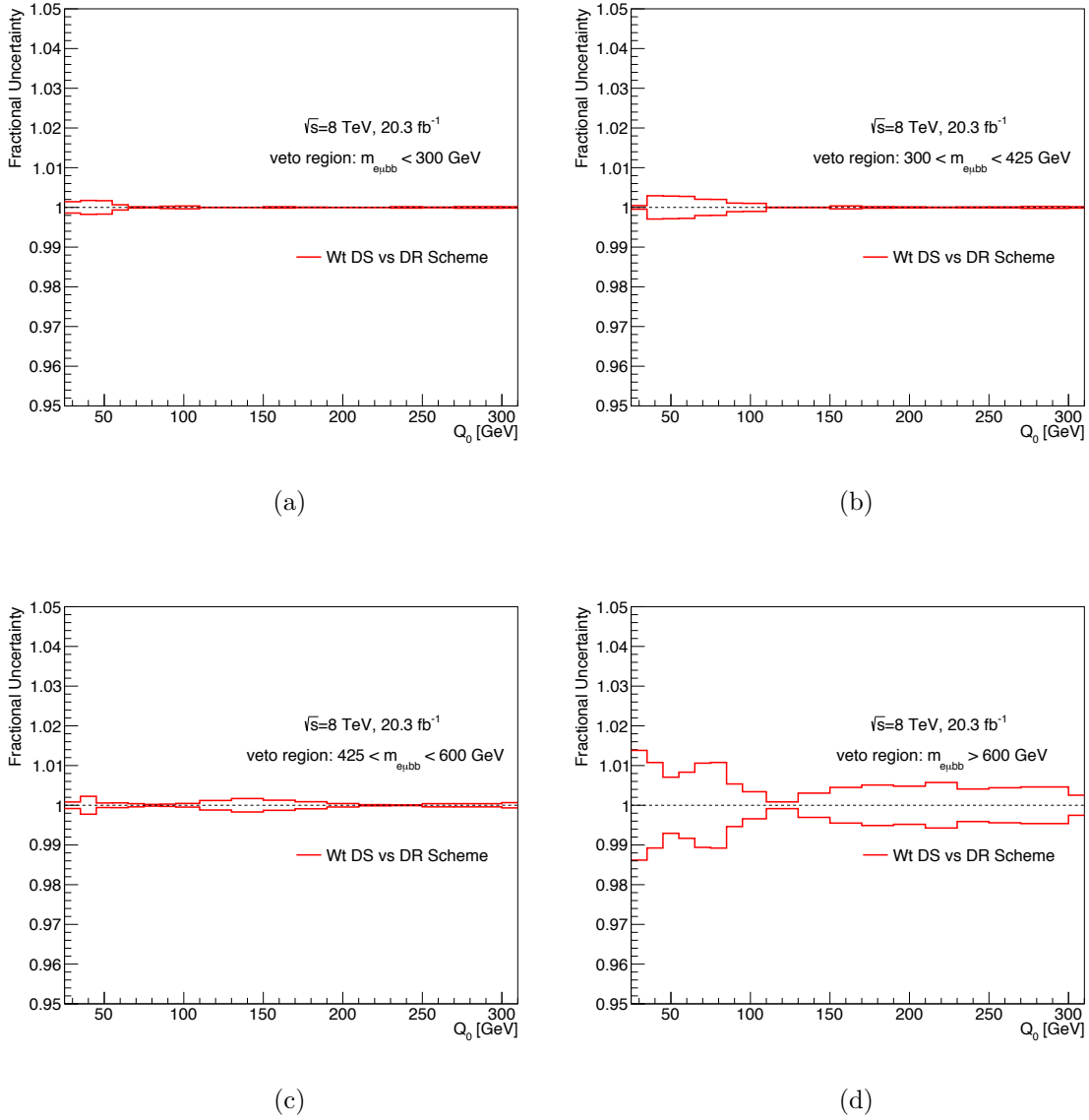
Figure 5.50: Uncertainty in the gap fraction due to the choice of schema used to model $Wt+b$ event interference with $t\bar{t}$ events, as a function of Q_0. The results are shown for a) $m_{e\mu bb} < 300$ GeV, b) $300 < m_{e\mu bb} < 425$ GeV, c) $425 < m_{e\mu bb} < 600$ GeV, and d) $m_{e\mu bb} > 600$ GeV.
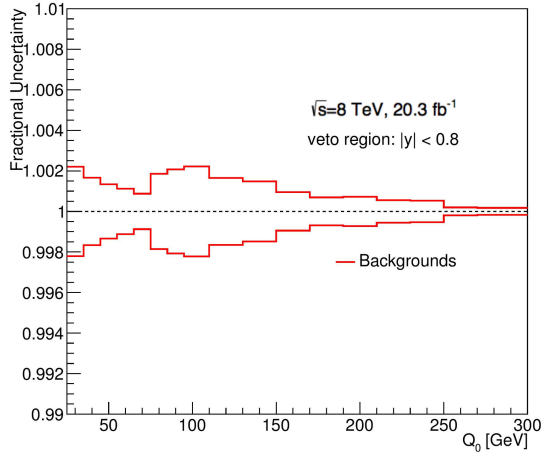
(a)
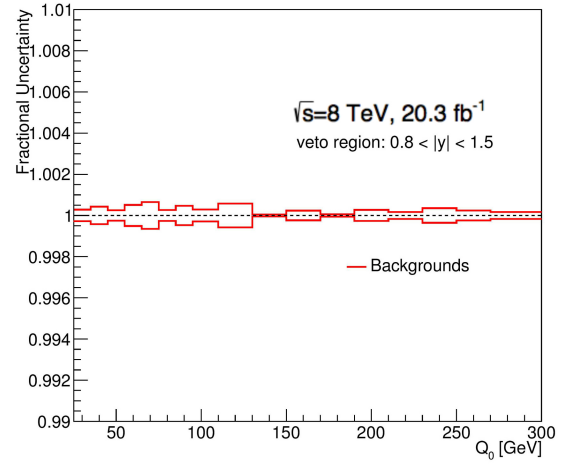
(b)
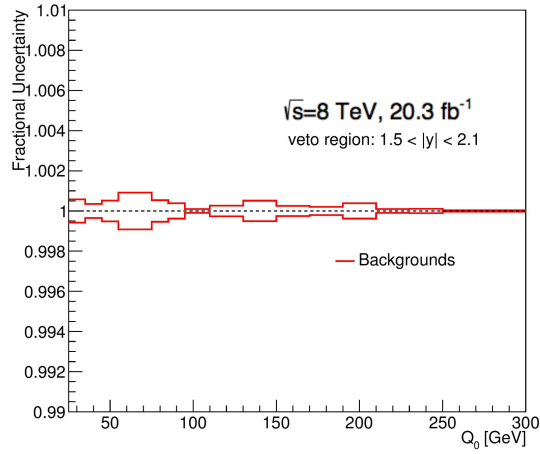
(c)
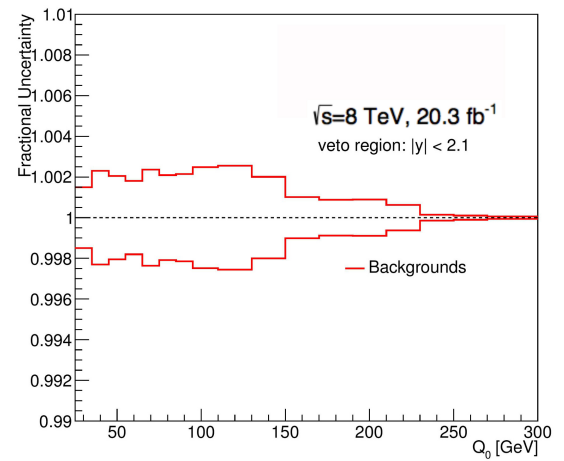
(d)

Figure 5.51: Uncertainty in the gap fraction due to background estimations, as a function of Q_0. The results are shown for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.

139
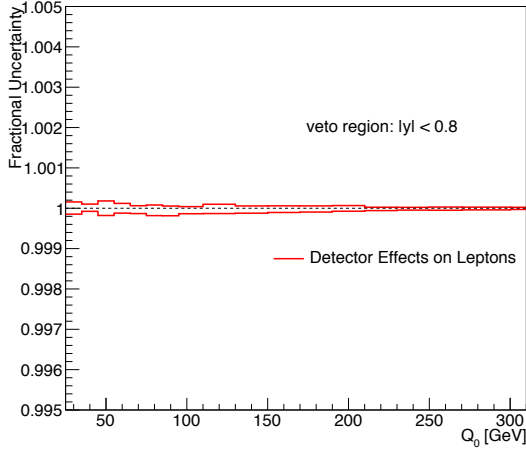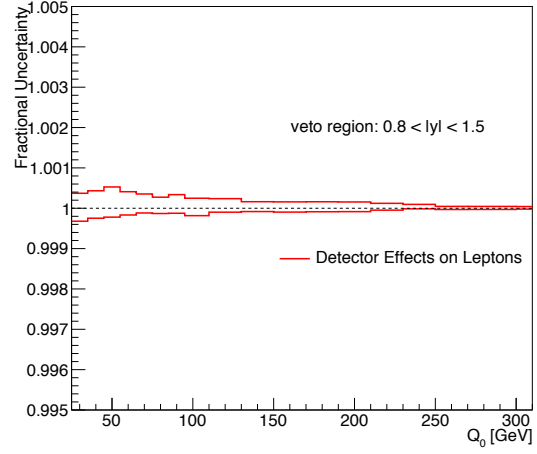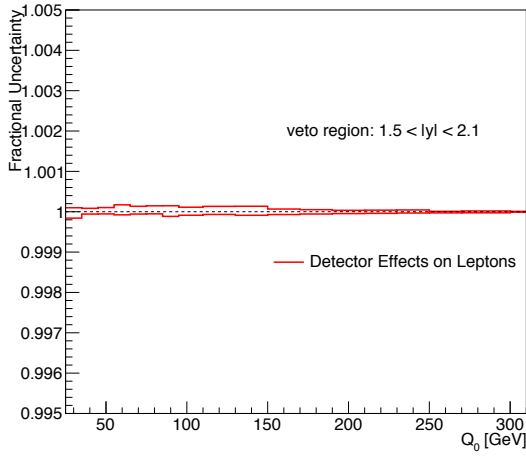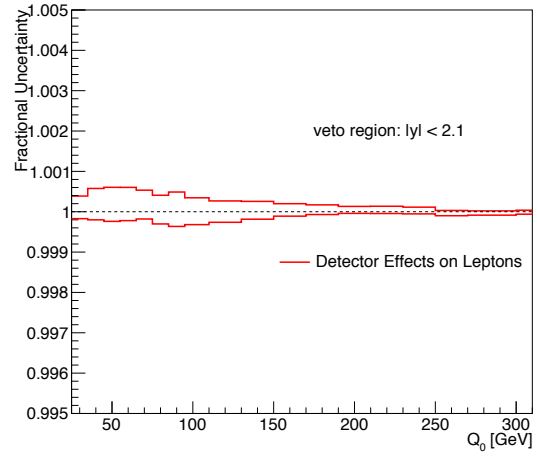
Figure 5.52: Uncertainty in the gap fraction due to lepton selection efficiency and resolution, as a function of Q_0. The results are shown for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.

5.5.9 Uncertainty in Jet Resolution, Efficiency, and Cleaning

The uncertainties due to jet energy resolution, the efficiency of jet selection, and the impact of jet cleaning were calculated from the POWHEG+PYTHIA sample as in Section 5.5. The variations in the correction factor were found varying the resolution and jet selection and turning off the jet cleaning. The differences added in quadrature as in Section 5.5 give the total contributing uncertainty, as shown in Figure 5.53. Each of the three factors of resolution, efficiency, and cleaning are shown separately in Figures 5.54, 5.55, and 5.56, showing that most of the uncertainty comes from the jet energy resolution.
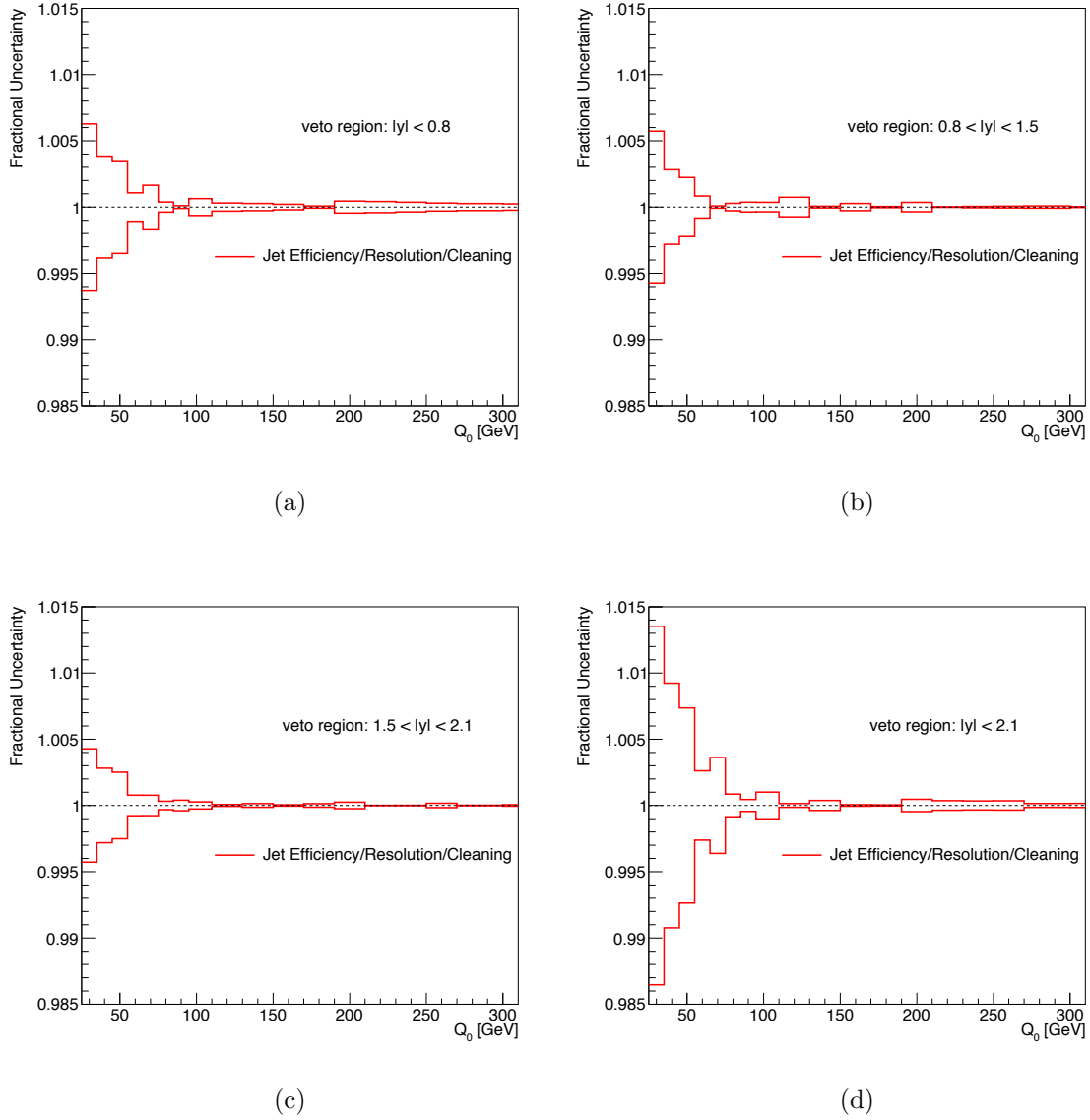
Figure 5.53: Uncertainty in the gap fraction due to jet energy resolution, efficiency and cleaning, as a function of Q_0. The results are shown for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.
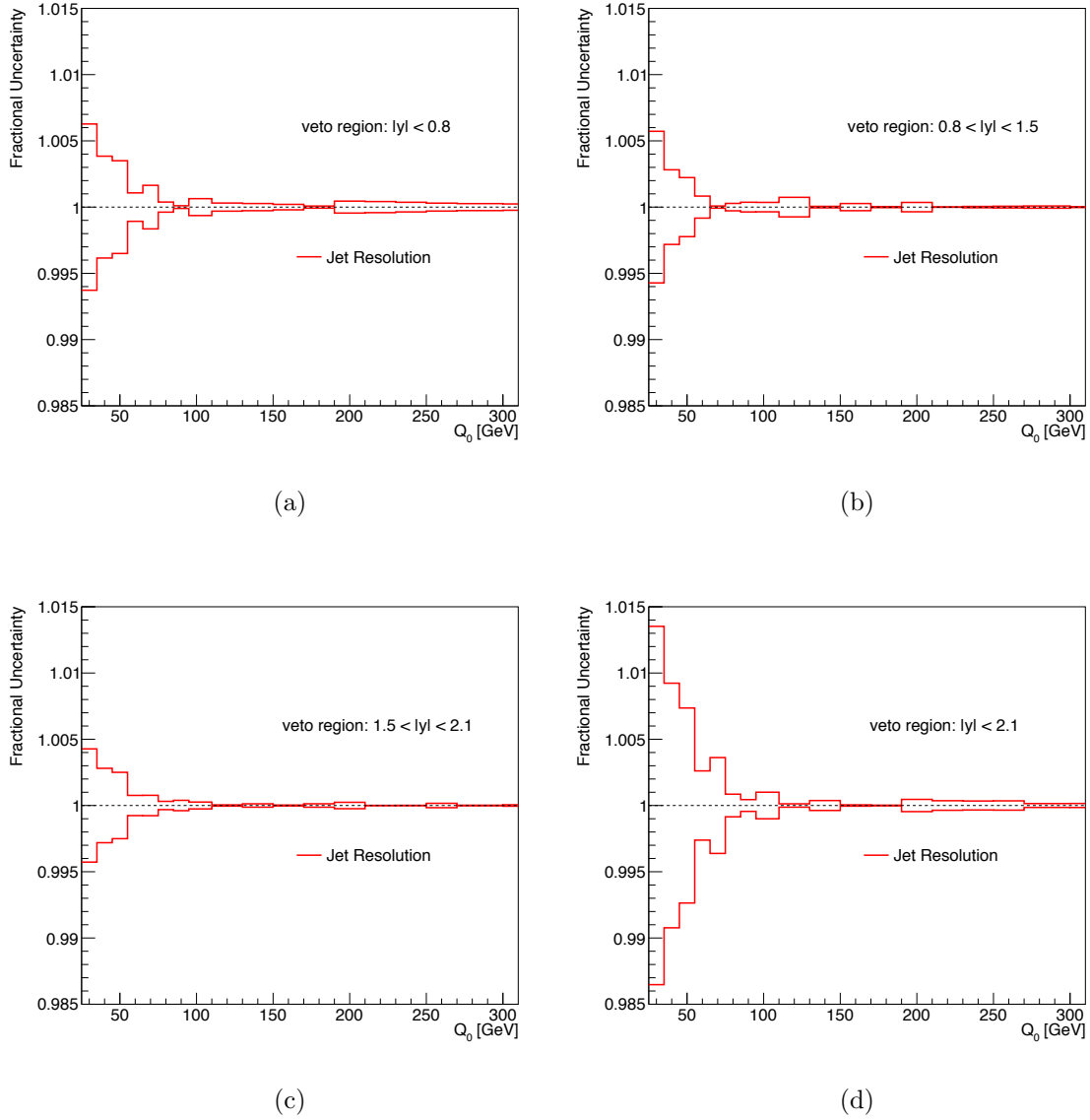
Figure 5.54: Uncertainty in the gap fraction due to jet energy resolution, as a function of Q_0. The results are shown for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.

(a)

(b)

(c)

(d)
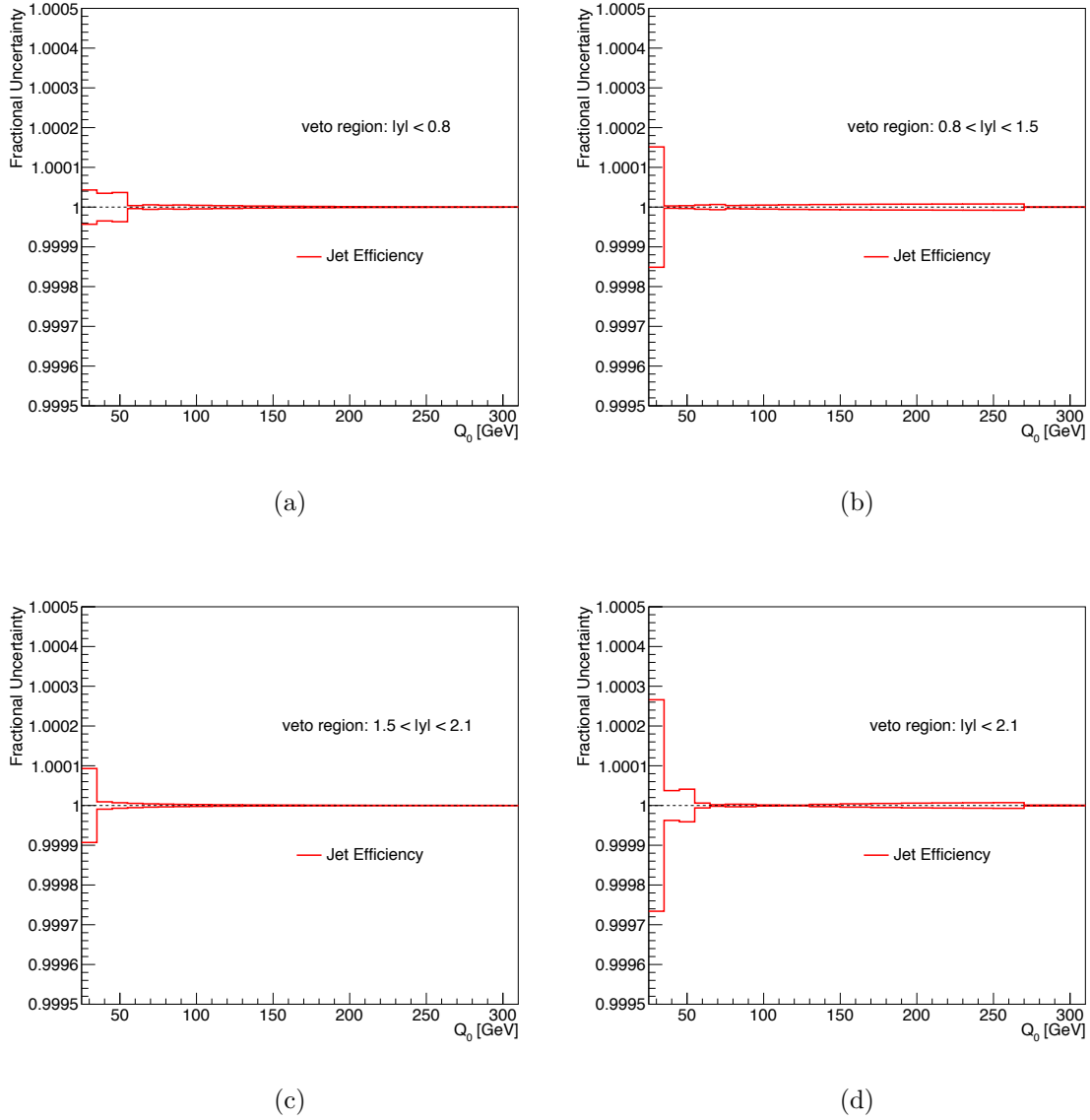
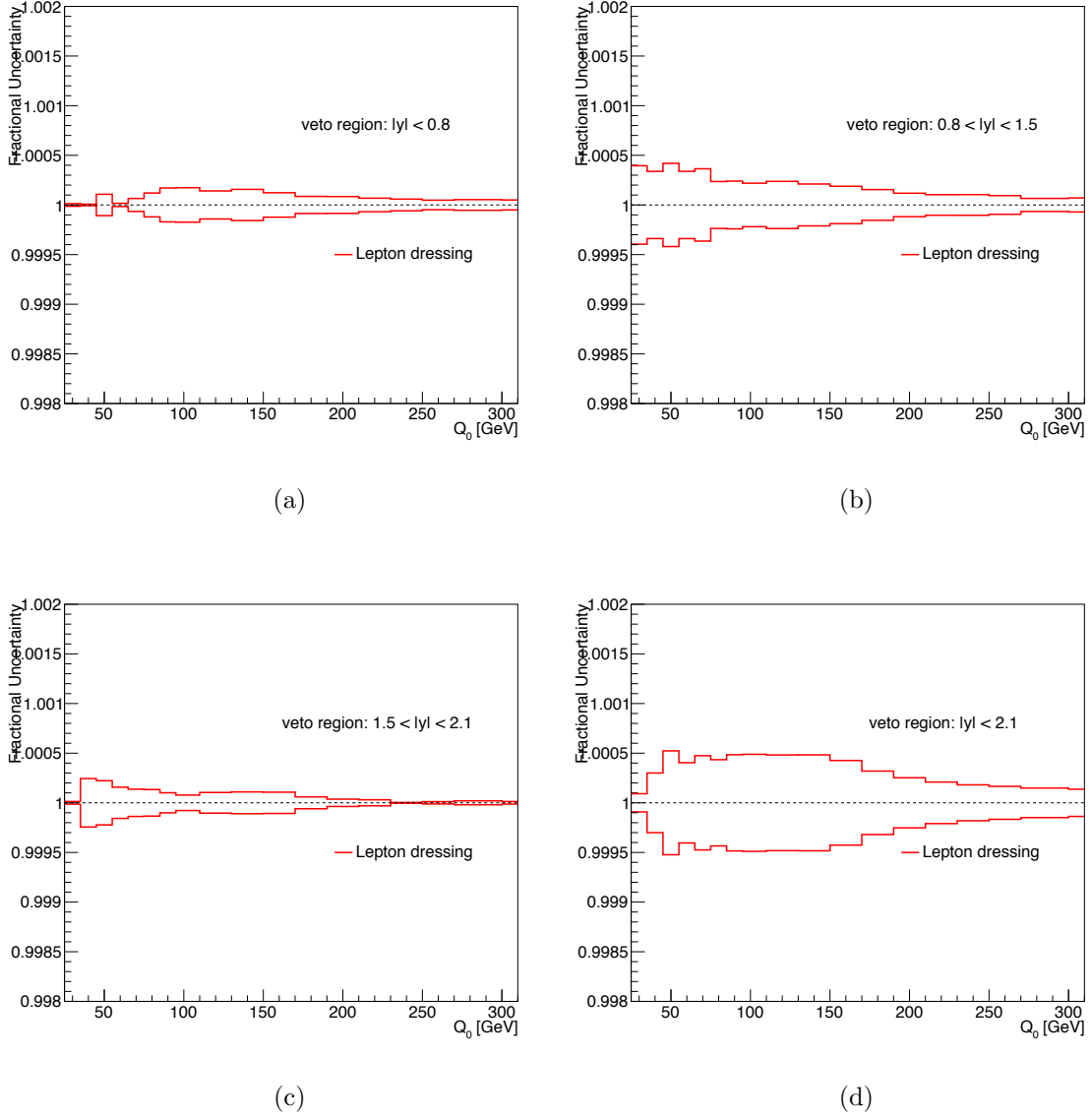Figure 5.55: Uncertainty in the gap fraction due to jet selection efficiency, as a function of Q_0. The results are shown for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.
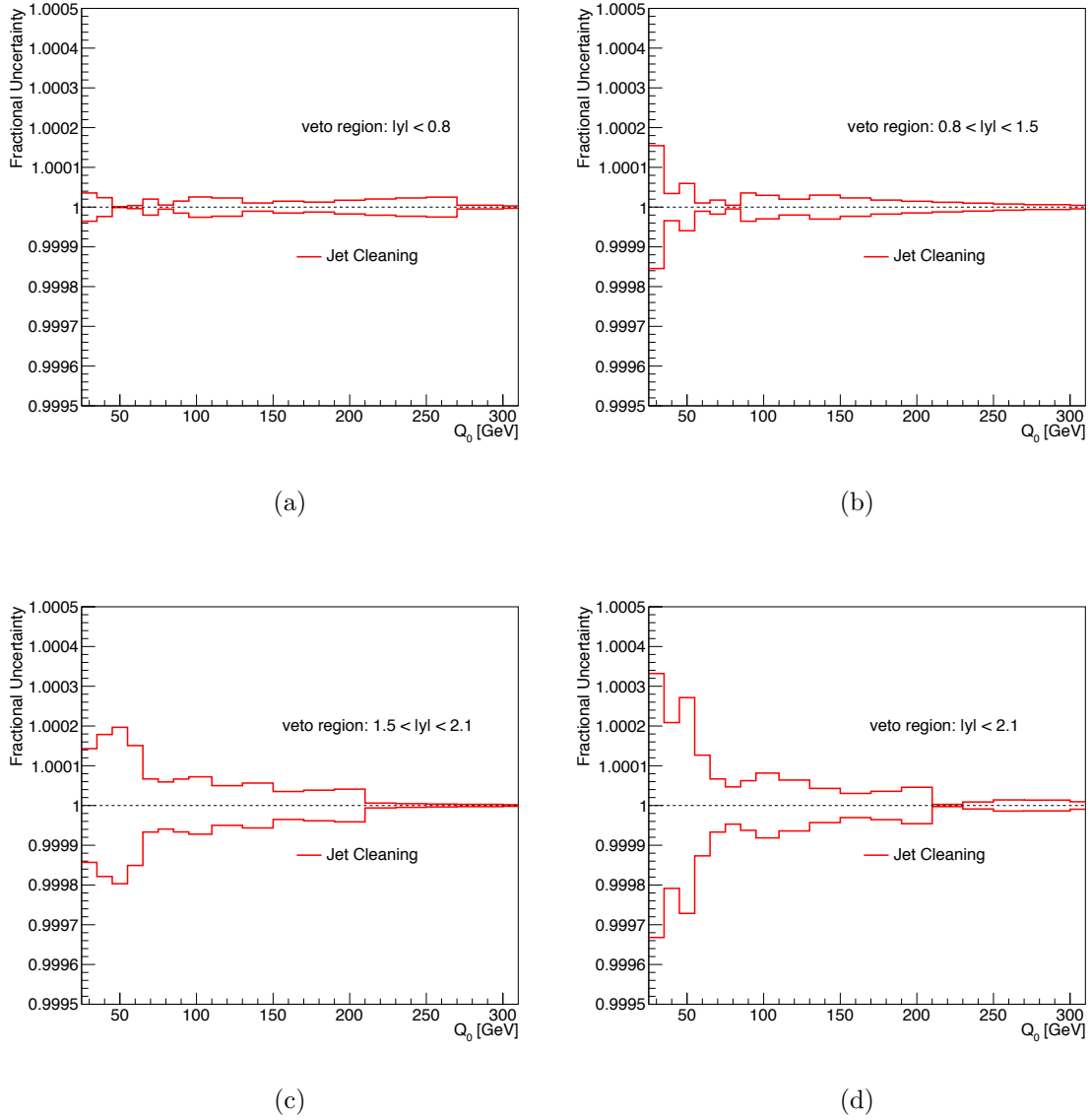
Figure 5.56: Uncertainty in the gap fraction due to impact of jet cleaning, as a function of Q_0. The results are shown for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.

5.5.10 Uncertainty in PDF modelling

Uncertainty in Monte Carlo simulation of the parton distribution functions (PDFs) affects the lepton and jet properties such as p_{T} and η, adding an additional uncertainty into the gap fraction measurement. This uncertainty was evaluated using the top group standard prescription [89], based on the PDF4LHC recommendations [90]. The method prescribes reweighting each simulated event according to the x and Q^2 of each colliding parton to correspond to each PDF variation. Using the MC@NLO+HERWIG $t\bar{t}$ Monte Carlo sample (DS 110001, `atlfast`), each event was reweighted for each variation of each PDF set at both the particle and reconstruction level. The fractional difference in the correction factor between each weighted sample and the original CT10 sample was found, and the differences combined according to the following formulas:

- **CT10:** The CT10 PDF [62] has 52 variations, corresponding to 26 $+1\sigma$ and 26 -1σ uncertainty components. The total uncertainty Δ is calculated with the symmetric Hessian formula: $\Delta = \frac{1}{2}\sqrt{\Sigma_i(\delta_i^+ - \delta_i^-)^2}$, in which $\delta_i^{\pm}$ corresponds to the fractional change in the correction factor for one of up/down variations. Because this result gives a 90% confidence level (CL), it was scaled down by 1.645 to correspond to 68% CL uncertainty.

- **MSTW:** The MSTW2088nlo68cl PDF [91, 92] includes 40 variations, corresponding to 20 $+1\sigma$ and 20 -1σ uncertainty components. The uncertainty is found with the symmetric Hessian formula, by calculating the quadrature sum of the $+1\sigma$ variations, Δ^+, and the separate quadrature sum of the -1σ variations,

Δ^-.

- **NNPDF:** The NNPDF 2.3 PDF [93] includes 100 variations. The RMS of fractional changes in the correction factor due to the variations is taken as $\pm 1\sigma$ uncertainty relative to the NNPDF central value. The magnitude of these $\pm 1\sigma$ variations is then set relative to the CT10 central value.

The total uncertainty due to each PDF is shown for the four rapidity regions as a function of Q_0 in Figure 5.57, along with the central value for each of the three PDF sets. According to the prescribed method, the total PDF uncertainty is taken as half the envelope that encompasses all three PDF set variations.
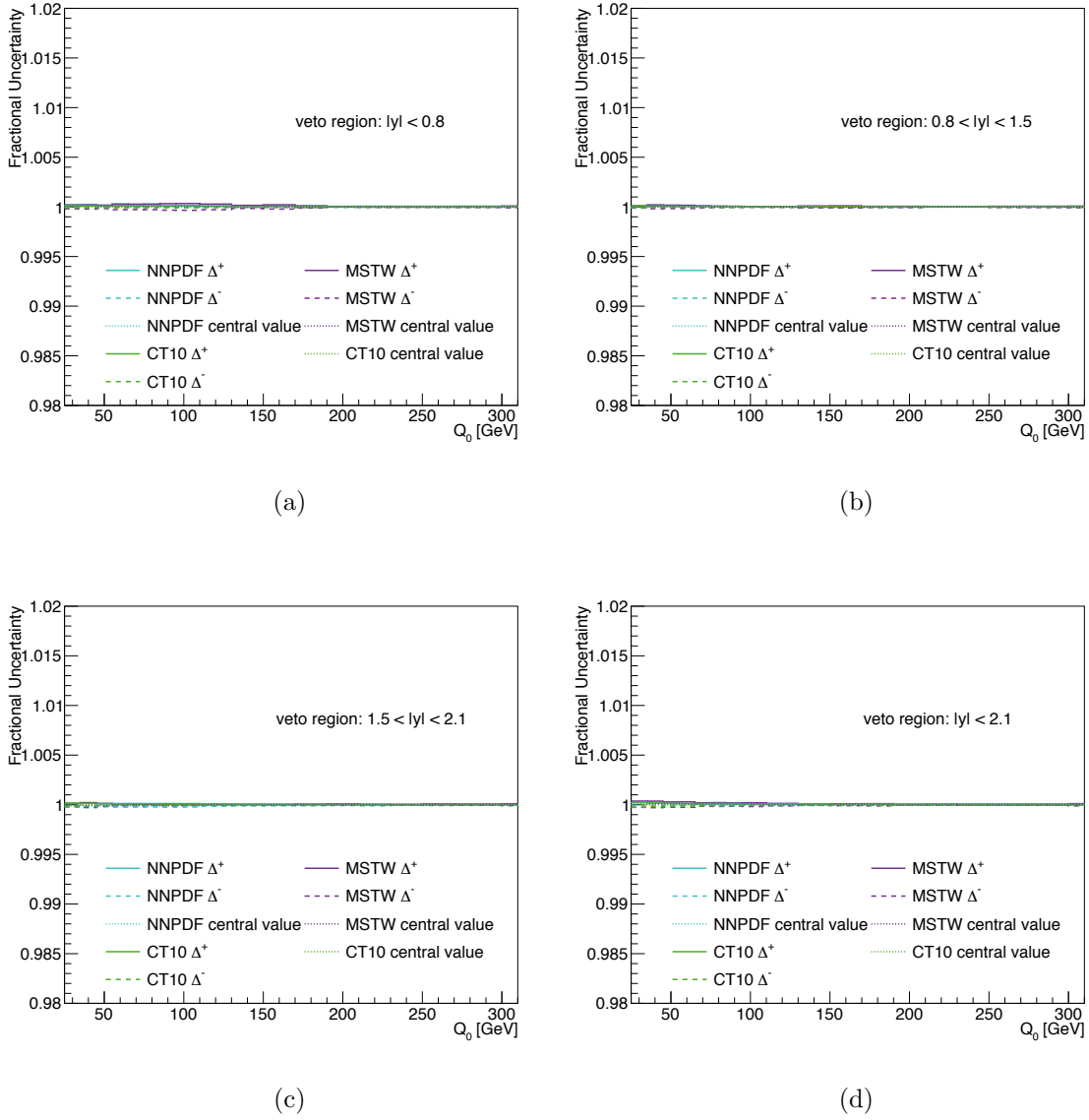
Figure 5.57: Uncertainty in the gap fraction due to PDF modelling, as a function of Q_0, for the three PDF sets of CT10, MSTW, and NNPDF. The central values for each PDF set are also shown, all in comparison to the MC@NLO+HERWIG CT10 sample. The results are shown for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.

5.5.11 Summary of Uncertainties

Each of the systematic uncertainties was added in quadrature to find the total systematic uncertainty. These are shown in Figures 5.58 and 5.59 for Q_0 and Q_{sum}, respectively, along with the statistical uncertainty, for the four rapidity regions. Figures 5.60 and 5.61 show the distributions for the four invariant mass regions. The uncertainties for $Q_0 = 30$, 40, 50, 100 and 150 GeV for the rapidity region $|y| < 2.1$ are given in Table 5.1 compared to estimated uncertainties in the 2011 analysis [69]. For the four rapidity regions, the jet-related uncertainties dominate, as expected, along with uncertainty in the correction factor at higher p_{T} values. The JES and JVF uncertainties are lower than those measured by the previous analysis using 2011 data [69], probably due predominately to the improved jet energy scale calibration from the `JetETMiss` group. In the invariant mass regions, the uncertainty in $t\bar{t}$ modelling dominates.
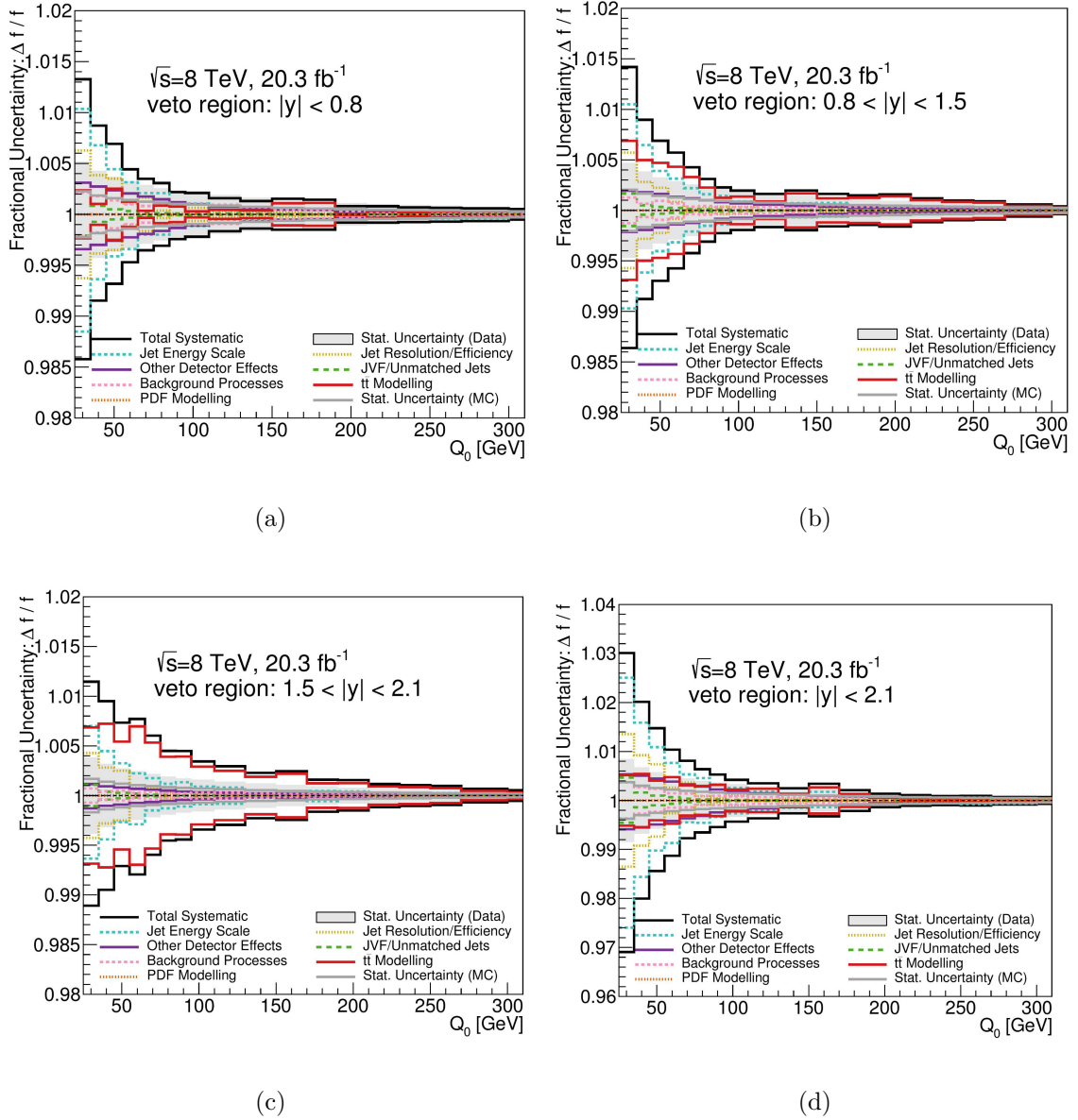
Figure 5.58: Summary of systematic uncertainties in the gap fraction measurement, as a function of Q_0, for the four rapidity regions. The "Total Systematic" line shown is the quadrature sum of the individual uncertainties. The results are shown for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.
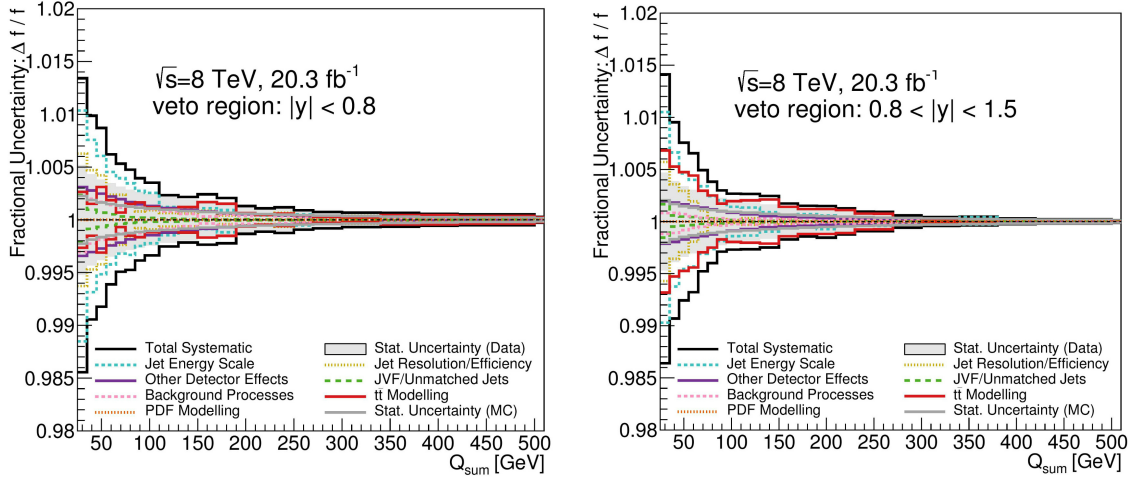
(a)

(b)
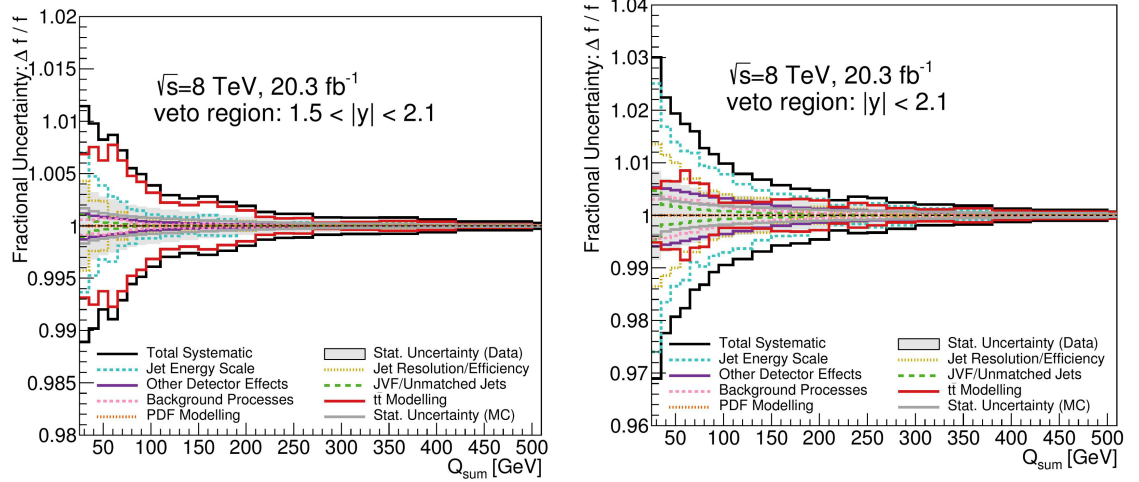
(c)

(d)

Figure 5.59: Summary of systematic uncertainties in the gap fraction measurement, as a function of Q_{sum}, for the four rapidity regions. The "Total Systematic" line shown is the quadrature sum of the individual uncertainties. The results are shown for a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.
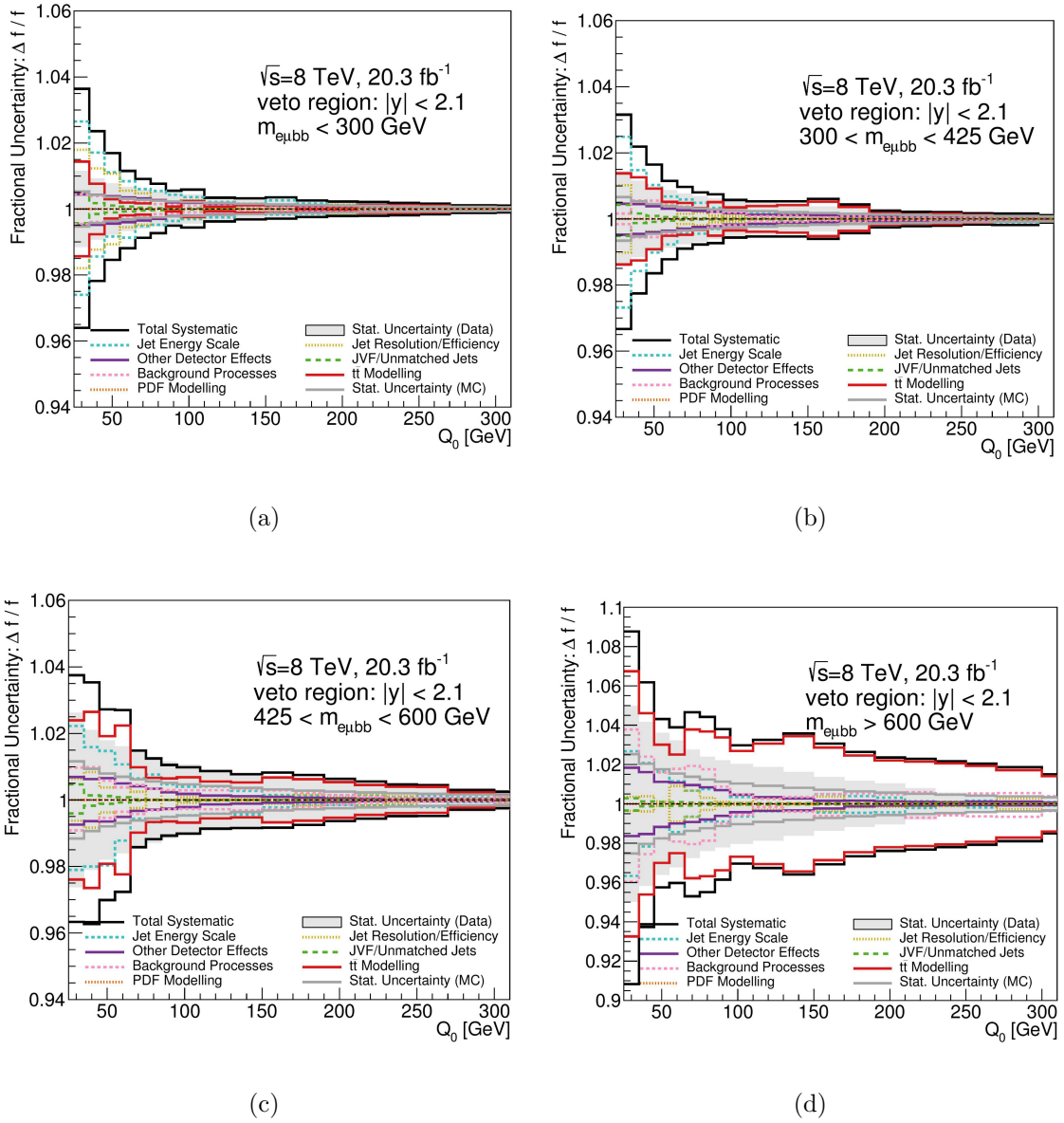
Figure 5.60: Summary of systematic uncertainties in the gap fraction measurement, as a function of Q_0, for the four invariant mass regions. The "Total Systematic" line shown is the quadrature sum of the individual uncertainties. The results are shown for a) $m_{e\mu bb} < 300$ GeV, b) $300 < m_{e\mu bb} < 425$ GeV, c) $425 < m_{e\mu bb} < 600$ GeV, and d) $m_{e\mu bb} > 600$ GeV.
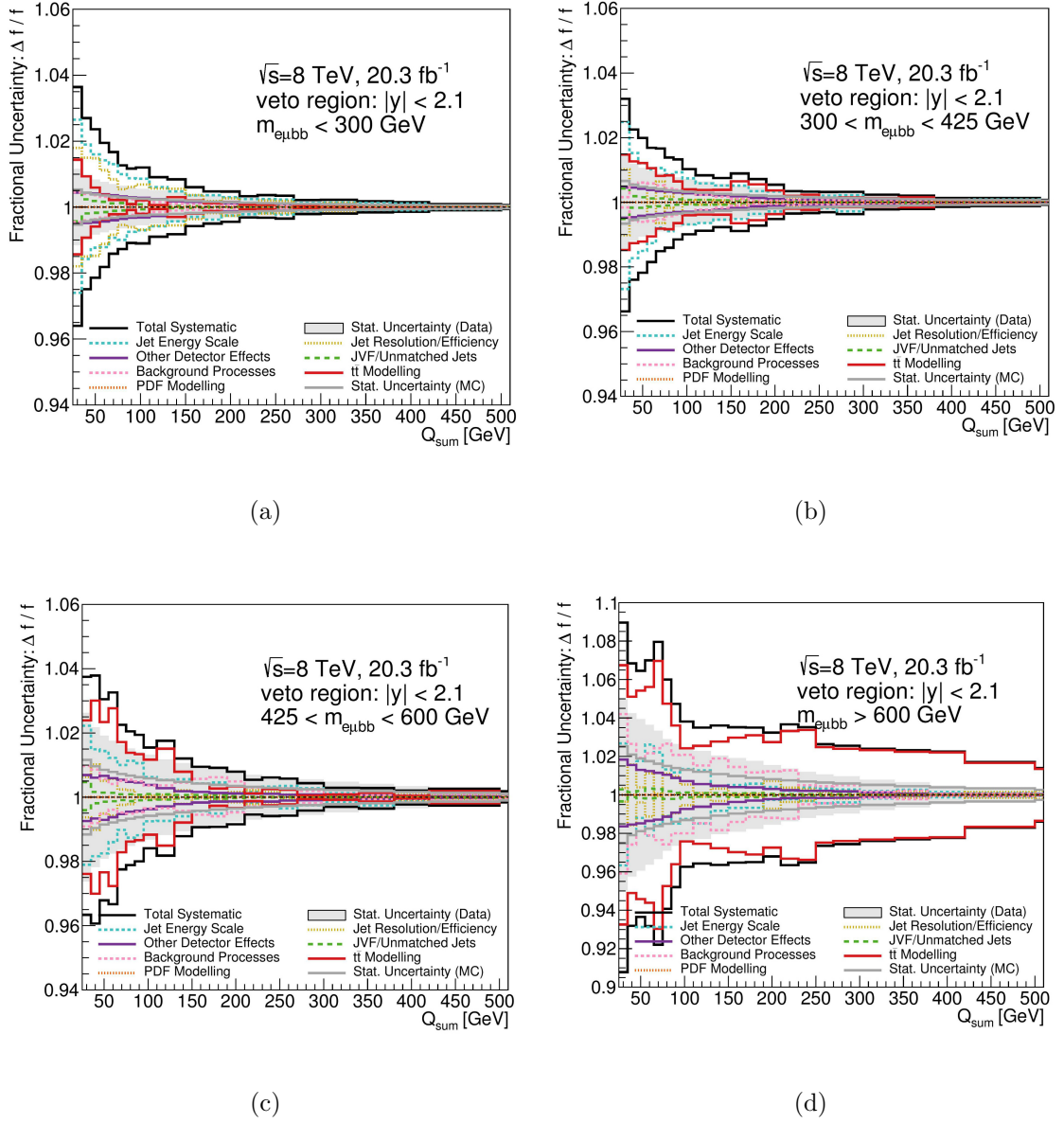
Figure 5.61: Summary of systematic uncertainties in the gap fraction measurement, as a function of Q_{sum}, for the four invariant mass regions. The "Total Systematic" line shown is the quadrature sum of the individual uncertainties. The results are shown for a) $m_{e\mu bb} < 300$ GeV, b) $300 < m_{e\mu bb} < 425$ GeV, c) $425 < m_{e\mu bb} < 600$ GeV, and d) $m_{e\mu bb} > 600$ GeV.

Uncertainty $|y| < 2.1$ (%)	Q_0=25 GeV		Q_0=35 GeV		Q_0=45 GeV		Q_0=95 GeV		Q_0=150 GeV	
	2011	2012	2011	2012	2011	2012	2011	2012	2011	2012
Jet Energy Scale, Res, Eff	3.4	2.85	2.0	1.84	1.1	1.31	0.6	0.275	0.2	0.177
Detector Effects	0.8	0.533	0.7	0.504	0.6	0.446	0.2	0.19	0.1	0.0939
Background Processes	0.8	0.113	0.9	0.224	0.8	0.2	0.7	0.0693	<0.1	0.0234
JVF/Unmatched Jets	2.2	0.467	0.4	0.16	0.2	0.0993	0.1	0.0133	<0.1	0.00879
$t\bar{t}$ Modelling	2.0	0.514	2.1	0.545	1.6	0.404	0.6	0.222	0.4	0.265
PDF Uncertainty	–	<0.1	–	<0.1	–	<0.1	–	<0.1	–	<0.1
Systematic Total	4.8	2.99	3.0	2.01	2.3	1.47	1.0	0.418	0.5	0.339
Statistical (Data)	2.0	0.847	1.8	0.679	1.5	0.574	0.9	0.331	0.7	0.217
Statistical (MC)	–	0.374	–	0.299	–	0.253	–	0.147	–	0.0983

Table 5.1: Fractional systematic and statistical uncertainties in the gap fraction from data corrected to particle level, for $|y| < 2.1$, $Q_0 = 25$, 35, 45, 95, and 150 GeV. These are shown compared to estimates from graphical results of the fractional uncertainties in the 2011 gap fraction analysis [69].

5.6 Results

Figures 5.62, 5.63, 5.64 and 5.65 show the resulting gap fraction from data corrected to particle level with the Monte Carlo simulations as a function of Q_0 and Q_{sum} for the four rapidity regions. The gap fraction plot in each figure shows the results from data and the four NLO Monte Carlo simulations, and the ratio plots show three multi-leg LO Monte Carlo simulations and three sets of simulation samples with varied levels of initial and final state radiation (ISR/FSR). The left edge of each step in the ratio plots corresponds to the Q_0 or Q_{sum} threshold plotted in the gap fraction plot above it.

Similarly, Figures 5.66, 5.67, 5.68, and 5.69 show the gap fraction from data corrected to particle level for the four invariant mass regions as a function of Q_0 and Q_{sum}. Figure 5.70 gives an alternative presentation of the gap fraction for the different mass regions for four Q_0 values.

The correlation between Q thresholds was calculated in matrix form for both statistical and systematic contributions, and the combination used to calculate the χ^2 values measuring the level of each generator's agreement with data, for both Q_0 and Q_{sum} for each rapidity region and invariant mass region. First, the statistical covariance matrices for the gap fraction from data were built by fluctuating the p_{T} distributions used in the gap fraction (as well as the number of events without additional jets with $p_{\mathrm{T}} > 25$ GeV) over a Poisson random number distribution 1000 times. Next, the systematic covariance matrices were built by creating the covariance

155

matrix of each symmeterized component of each systematic uncertainty separately,[2] and then adding the covariance matrices for all the components together. Finally, the combination of statistical and systematic correlations were built as the matrix sum of statistical and systematic covariance matrices, divided by the diagonal terms. The correlation matrices combining statistical and systematic sources are given in Figures 5.71 and 5.72 for the rapidity regions and Figures 5.73 and 5.74 for the mass regions.

The agreement of each generator's prediction of the gap fraction with data is a straightforward matrix calculation of χ^2 with the correlation matrices:

$$\chi^2 = M^T C^{-1} M \tag{5.15}$$

where for each region, M represents a matrix of dimensions $(1 \times n_Q)$ of the difference between the gap fraction from data and the gap fraction from simulation for each Q threshold. C represents the correlation matrix of dimensions $(n_Q \times n_Q)$, which was inverted with `Python`. The χ^2 values are given in Table 5.6 for the rapidity regions and Table 5.7 for the invariant mass regions.

Tables 5.2 and 5.3 give the gap fraction measurement for selected Q values compared to prediction by NLO generators, with combined systematic and statistical correlations between the Q values. Similarly, Tables 5.4 and 5.5 give the gap fraction measurement in the invariant mass regions for selected Q values compared to prediction by NLO generators, with combined systematic and statistical correlations between the Q values. The complete tables of gap fraction values from data and corresponding uncertainties

---

[2]Two exceptions to this method were made - for the uncertainty in PDF modelling, the envelope was chosen, following the PDF4LHC prescription. For the uncertainty from b-tagging, no attempt was made to break the three scale factors into their components.

will be submitted to HEPData for public access [94].

In both rapidity regions and invariant mass regions, the results show the POWHEG+PYTHIA 8 simulation in the best agreement with data, followed by the POWHEG+PYTHIA simulation. The sample generated by POWHEG+PYTHIA ($h_{\mathrm{damp}} = m_t$) correlates noticeably less closely, as well as the MC@NLO+HERWIG sample, especially in the central-most rapidity region. None of the generators is a close match for every region; particularly around the 100 GeV region most generators diverge somewhat from data measurement.

The ISR/FSR radiation variation samples generated from the 2011 gap fraction in Ref. [69] were intended to span $\pm 1\sigma$ uncertainty of jet activity; this result shows that the 2011 tuning was conservative and can be reduced for future analyses. In general the invariant mass regions follow similar trends as in the rapidity regions, although with lower statistical precision. The gap fraction measurement from data falls fully within the spread of ISR/FSR variations in the $\sqrt{s} = 7$ TeV results.
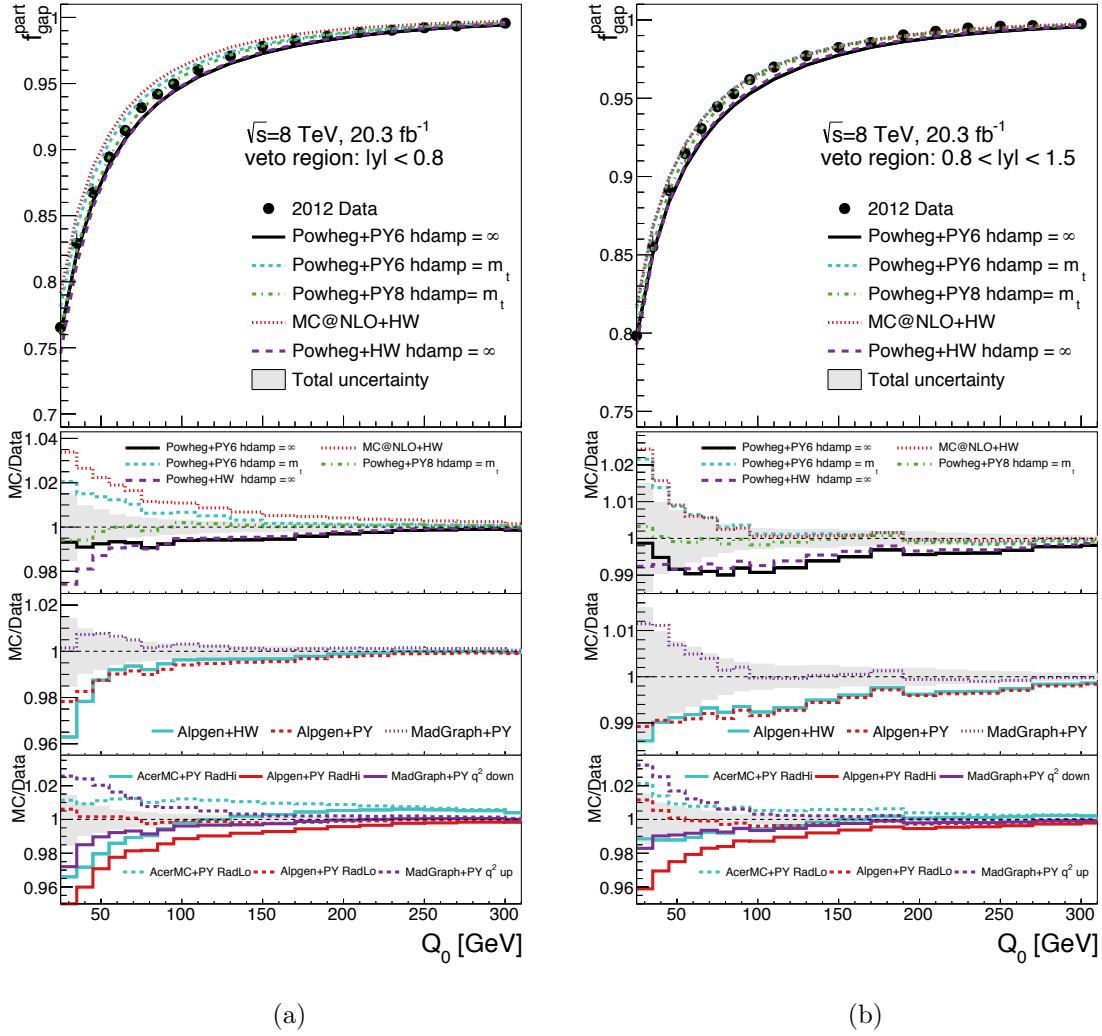
Figure 5.62: The gap fraction plots at particle level. The gap fraction plot from data is shown corrected back to the particle level and compared to various Monte Carlo simulations as a function of Q_0 for two of the four rapidity regions of a) $|y| < 0.8$ and b) $0.8 < |y| < 1.5$. The ratios show the gap fraction from simulation divided by the gap fraction from data, with the left edge of each step corresponding to each Q_0 threshold. The shaded band and error bars on the data show the combination of statistical and systematic uncertainty in data.
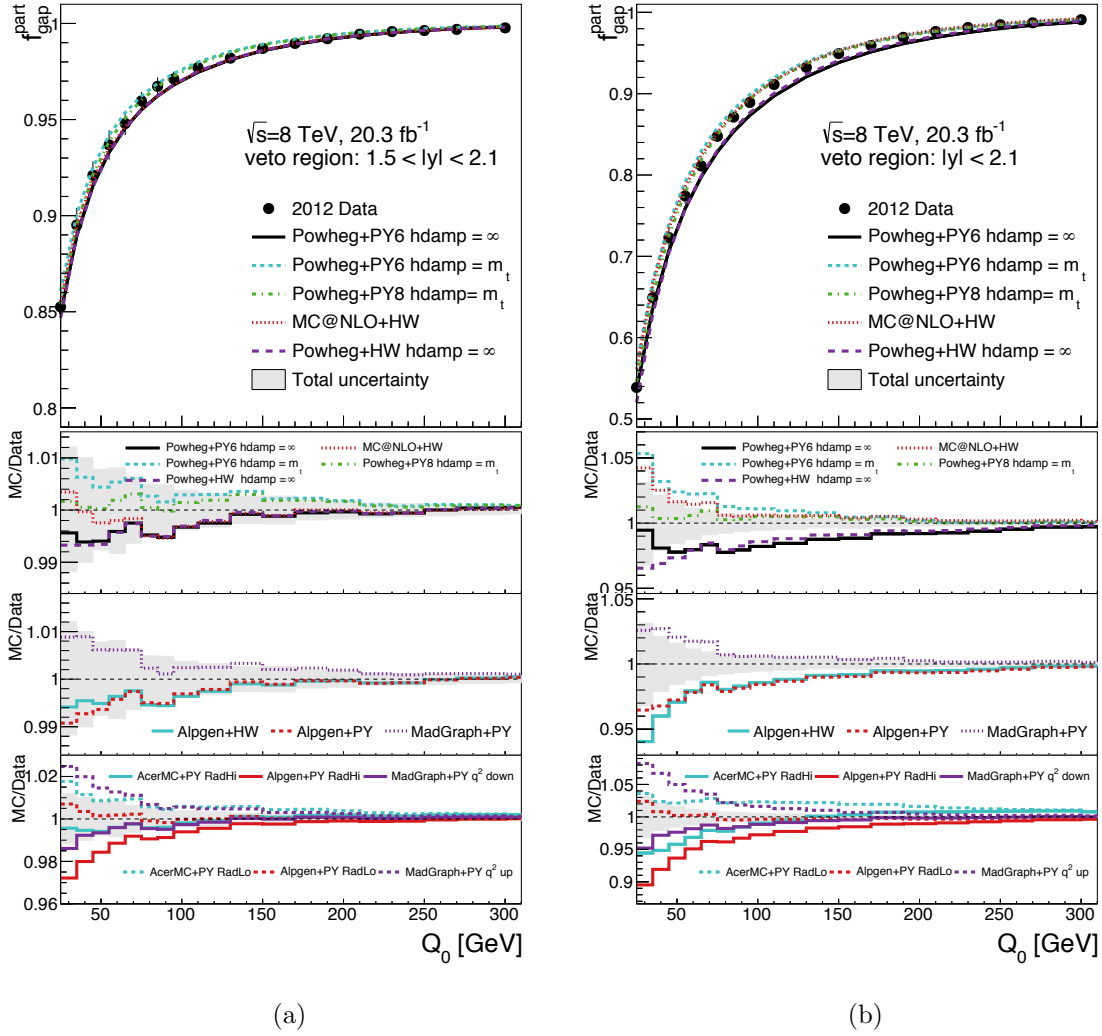
Figure 5.63: The gap fraction plots at particle level. The gap fraction plot from data is shown corrected back to the particle level and compared to various Monte Carlo simulations as a function of Q_0 for two of the four rapidity regions of a) $1.5 < |y| < 2.1$ and b) $|y| < 2.1$. The ratios show the gap fraction from simulation divided by the gap fraction from data, with the left edge of each step corresponding to each Q_0 threshold. The shaded band and error bars on the data show the combination of statistical and systematic uncertainty in data.
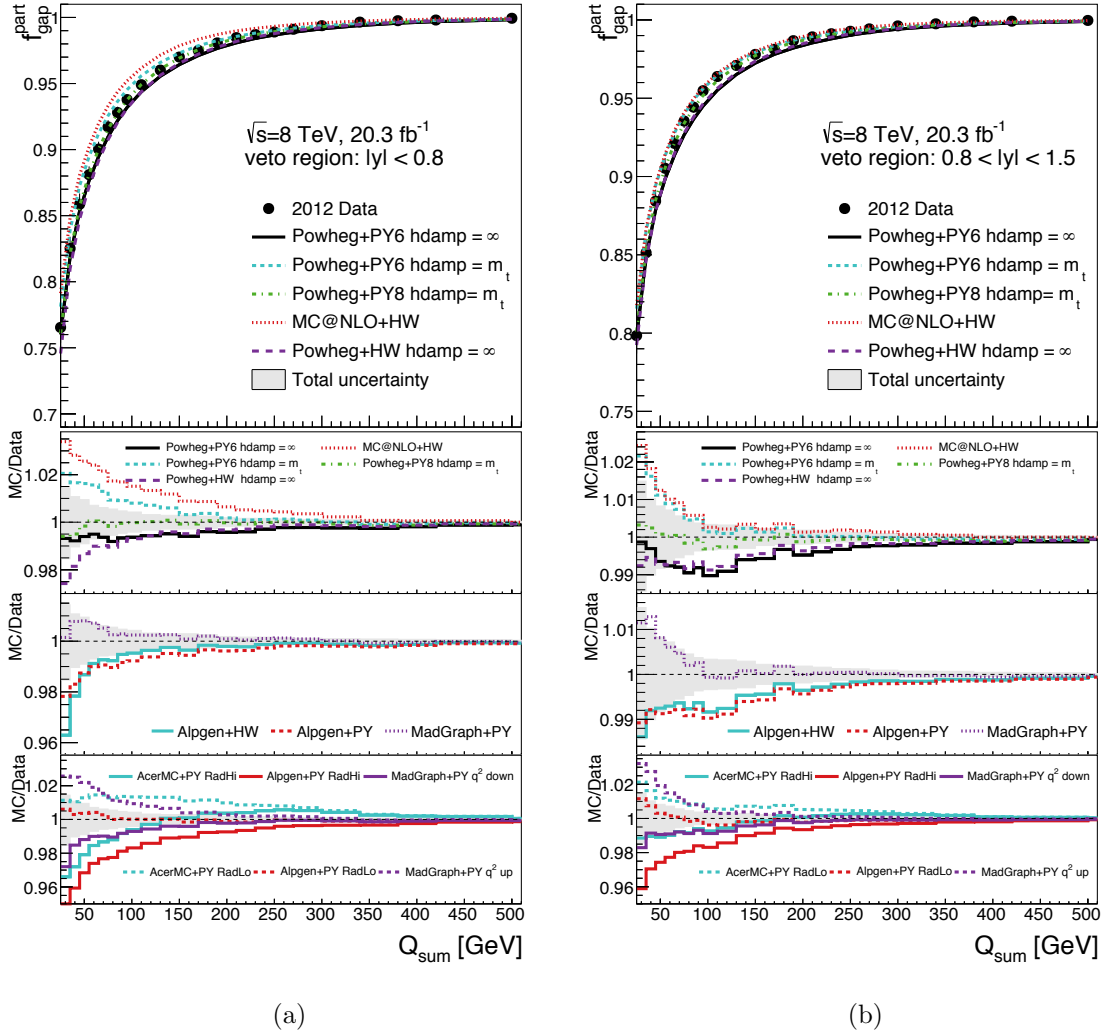
Figure 5.64: The gap fraction plots at particle level. The gap fraction plot from data is shown corrected back to the particle level and compared to various Monte Carlo simulations as a function of Q_{sum} for two of the four rapidity regions of a) $|y| < 0.8$ and b) $0.8 < |y| < 1.5$. The ratios show the gap fraction from simulation divided by the gap fraction from data, with the left edge of each step corresponding to each Q_{sum} threshold. The shaded band and error bars on the data show the combination of statistical and systematic uncertainty in data.
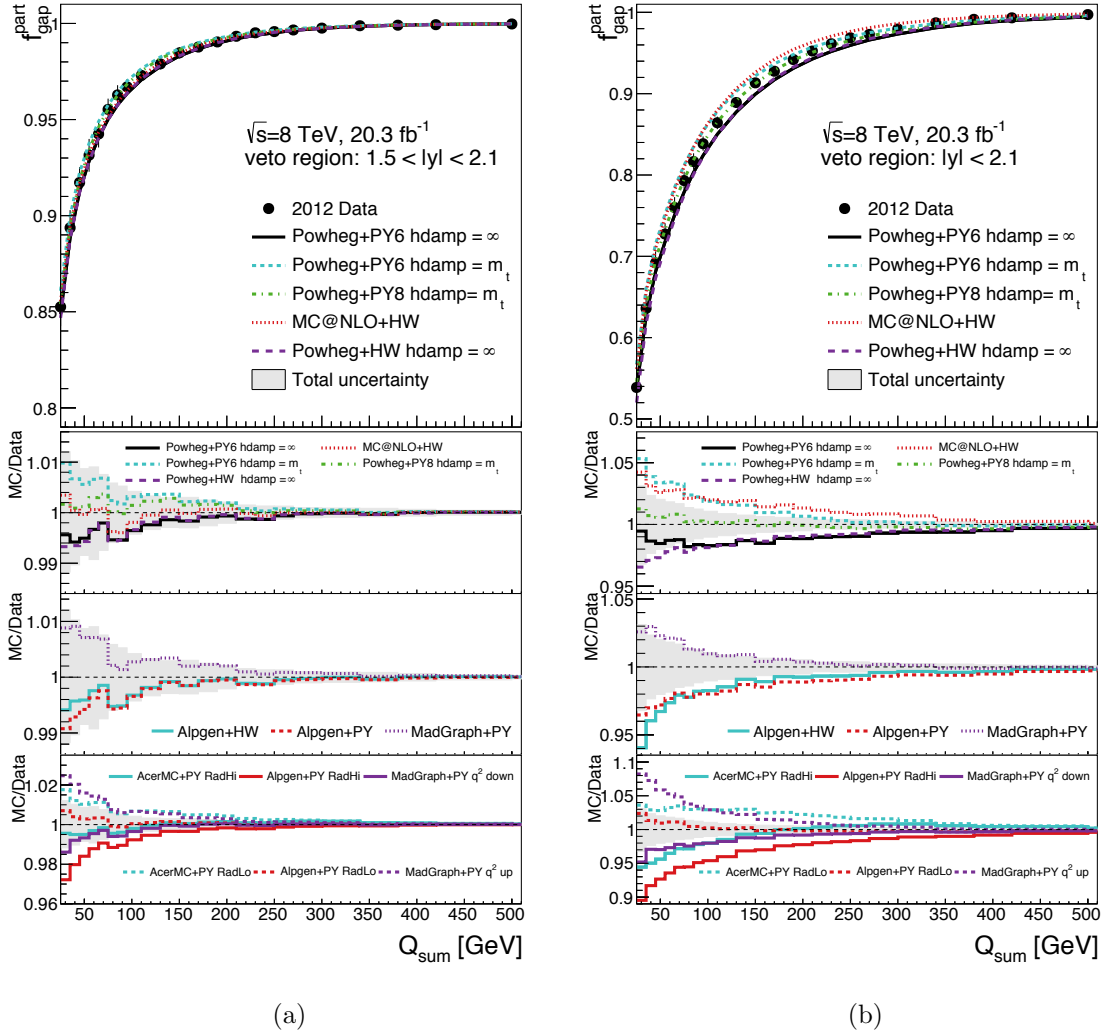
Figure 5.65: The gap fraction plots at particle level. The gap fraction plot from data is shown corrected back to the particle level and compared to various Monte Carlo simulations as a function of Q_{sum} for two of the four rapidity regions of a) $1.5 < |y| < 2.1$ and b) $|y| < 2.1$. The ratios show the gap fraction from simulation divided by the gap fraction from data, with the left edge of each step corresponding to each Q_{sum} threshold. The shaded band and error bars on the data show the combination of statistical and systematic uncertainty in data.
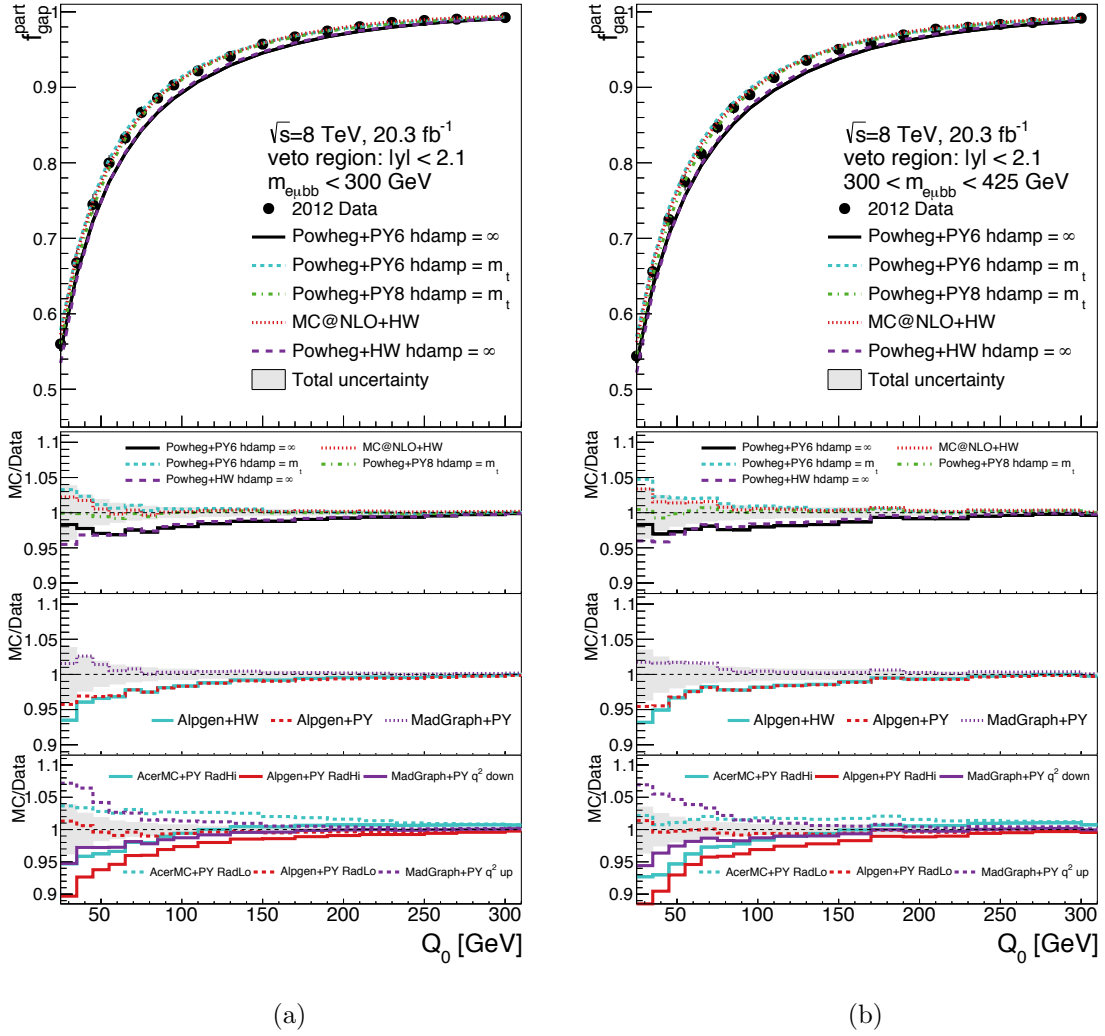
Figure 5.66: The gap fraction plots at particle level. The gap fraction plot from data is shown corrected back to the particle level and compared to various Monte Carlo simulations as a function of Q_0 for two of the four invariant mass regions of a) $m_{e\mu bb} < 300$ GeVand b) $300 < m_{e\mu bb} < 425$ GeV. The ratios show the gap fraction from simulation divided by the gap fraction from data, with the left edge of each step corresponding to each Q_0 threshold. The shaded band and error bars on the data show the combination of statistical and systematic uncertainty in data.
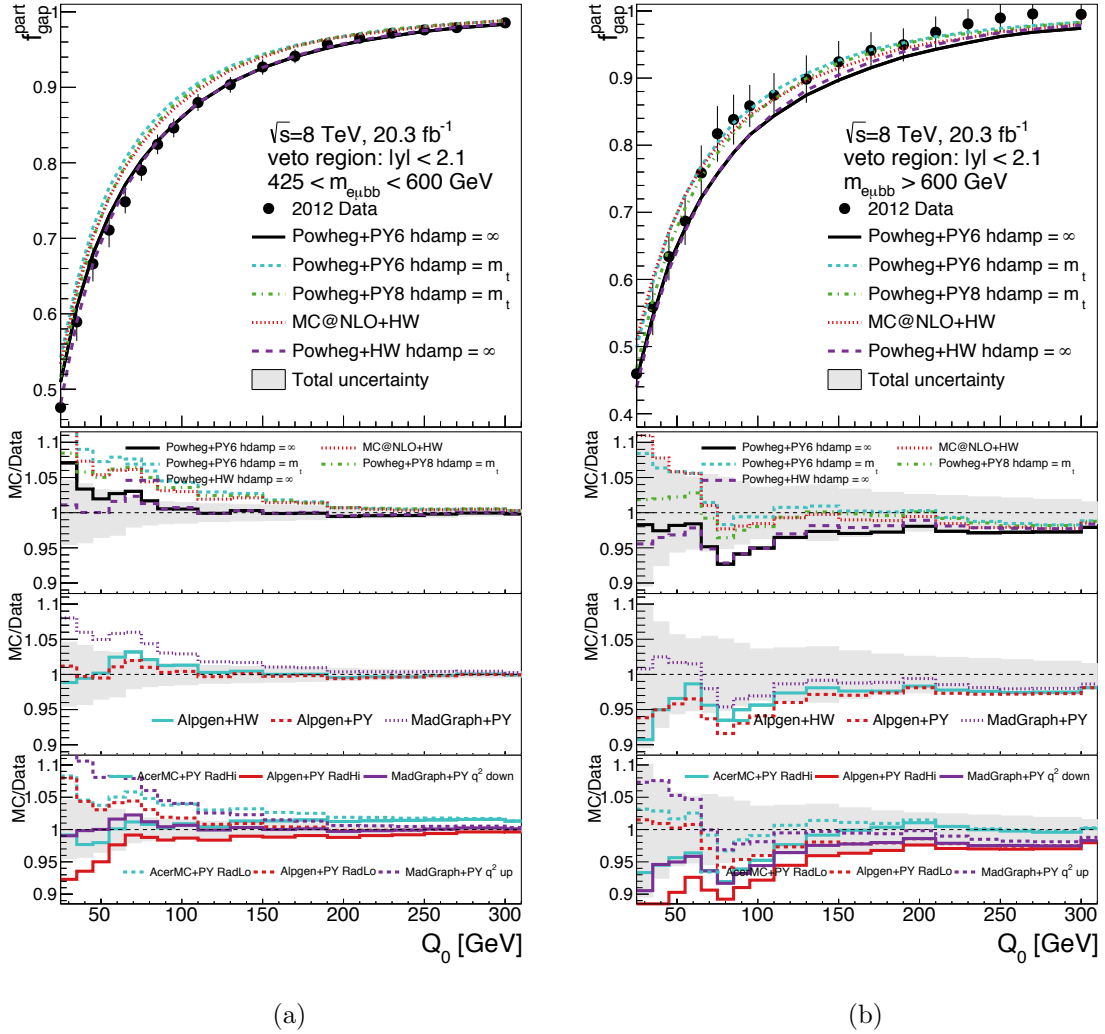
Figure 5.67: The gap fraction plots at particle level. The gap fraction plot from data is shown corrected back to the particle level and compared to various Monte Carlo simulations as a function of Q_0 for two of the four invariant mass regions of a) $425 < m_{e\mu bb} < 600$ GeV and b) $m_{e\mu bb} > 600$ GeV. The ratios show the gap fraction from simulation divided by the gap fraction from data, with the left edge of each step corresponding to each Q_0 threshold. The shaded band and error bars on the data show the combination of statistical and systematic uncertainty in data.
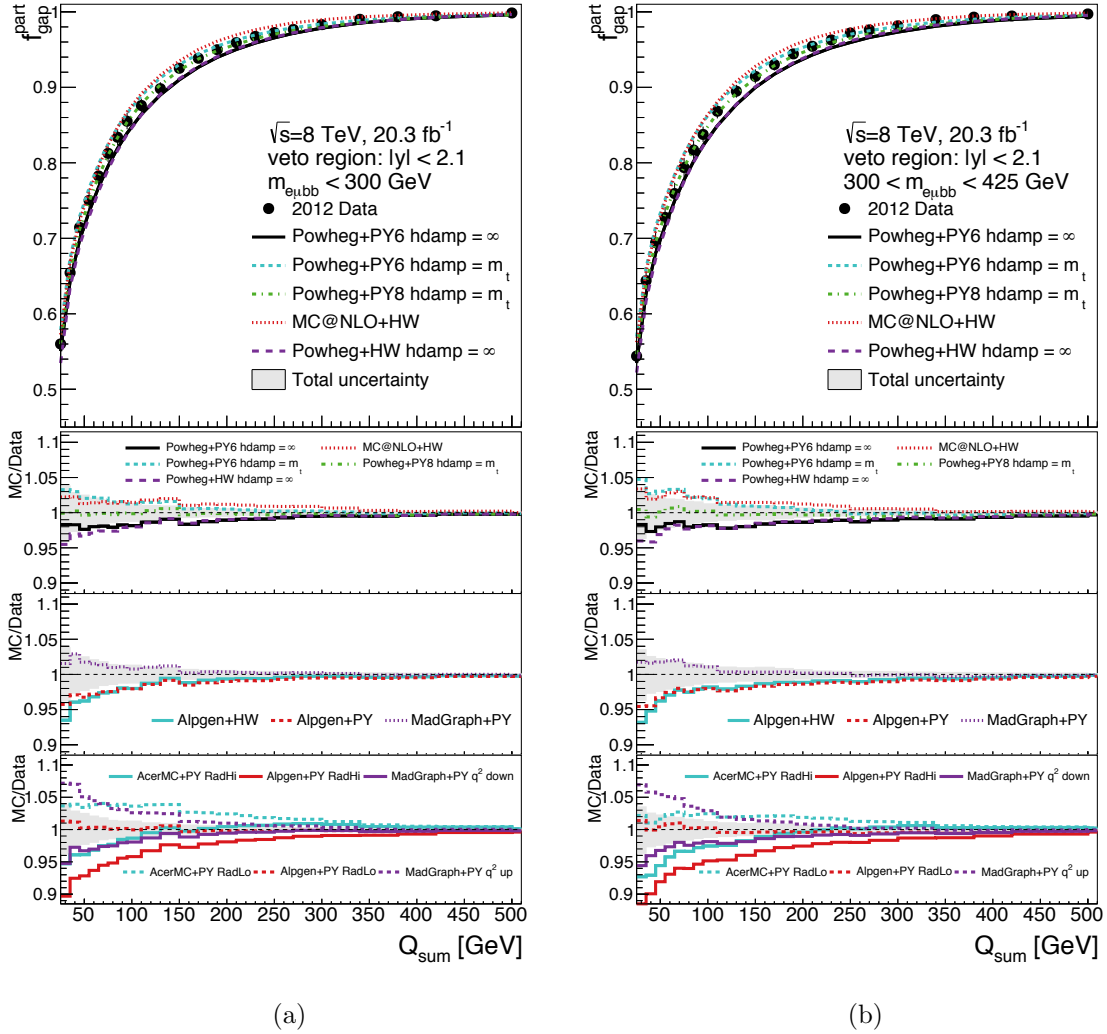
Figure 5.68: The gap fraction plots at particle level. The gap fraction plot from data is shown corrected back to the particle level and compared to various Monte Carlo simulations as a function of Q_{sum} for two of the four invariant mass regions of a) $m_{e\mu bb} < 300$ GeVand b) $300 < m_{e\mu bb} < 425$ GeV. The ratios show the gap fraction from simulation divided by the gap fraction from data, with the left edge of each step corresponding to each Q_{sum} threshold. The shaded band and error bars on the data show the combination of statistical and systematic uncertainty in data.
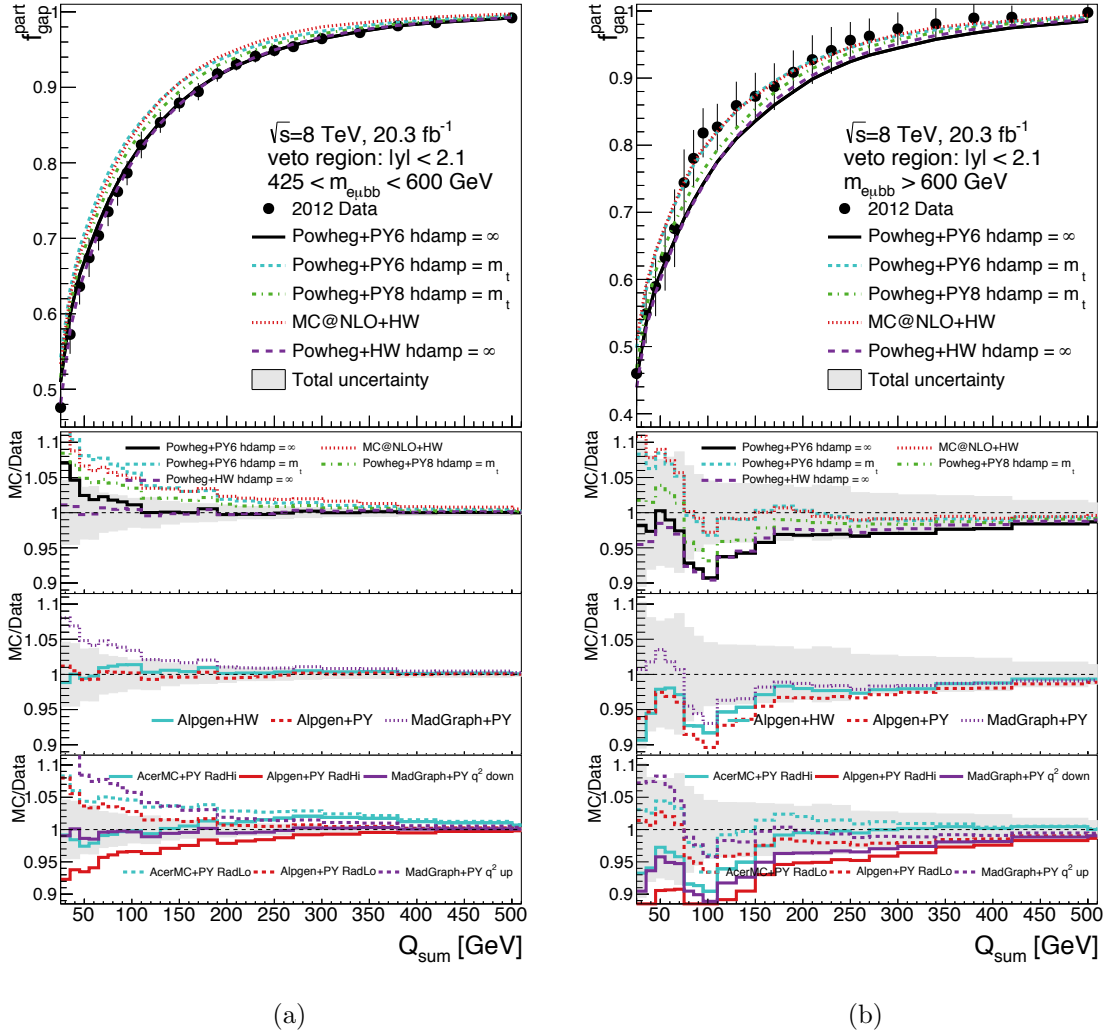
Figure 5.69: The gap fraction plots at particle level. The gap fraction plot from data is shown corrected back to the particle level and compared to various Monte Carlo simulations as a function of Q_{sum} for two of the four invariant mass regions of a) $425 < m_{e\mu bb} < 600$ GeV and b) $m_{e\mu bb} > 600$ GeV. The ratios show the gap fraction from simulation divided by the gap fraction from data, with the left edge of each step corresponding to each Q_{sum} threshold. The shaded band and error bars on the data show the combination of statistical and systematic uncertainty in data.
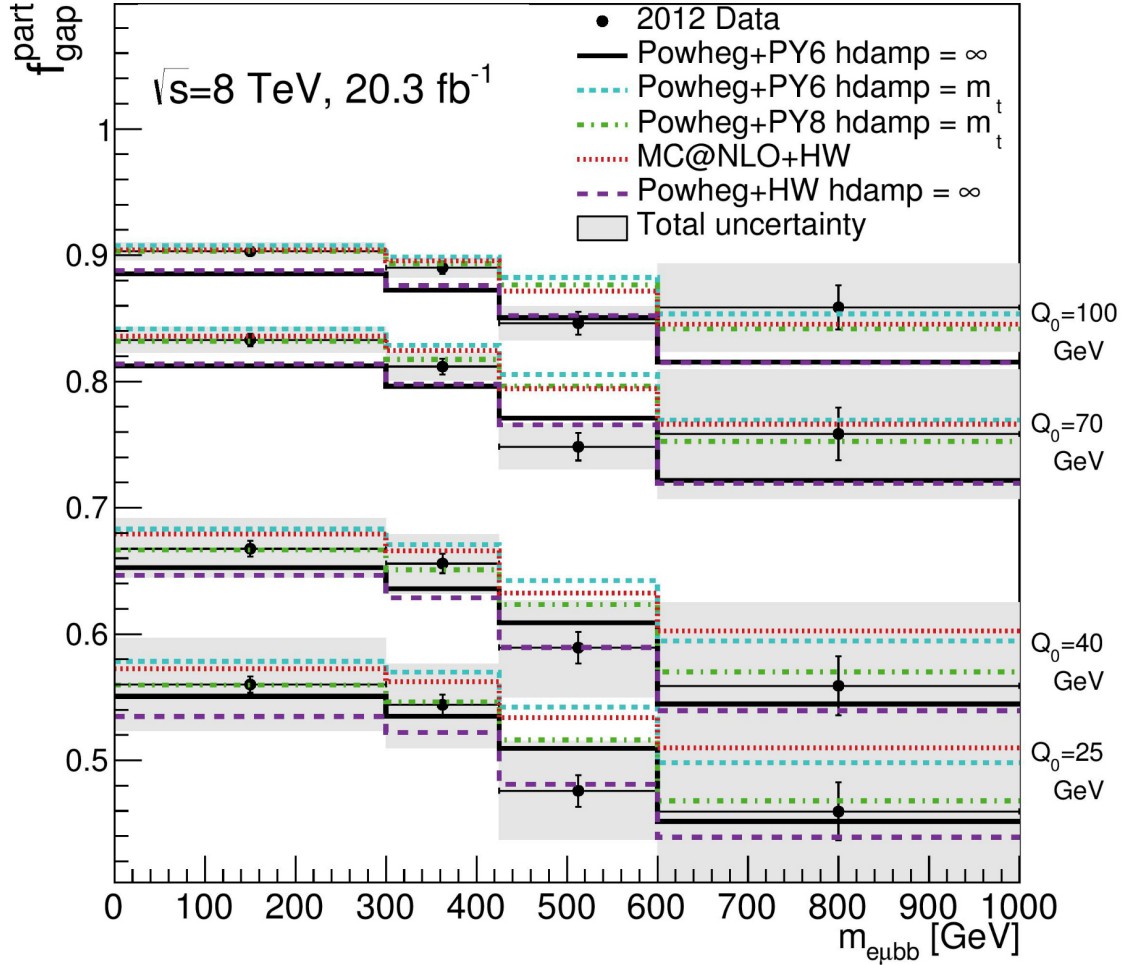
Figure 5.70: The gap fraction measurement from data compared to several Monte Carlo generators, by invariant mass region, for four different Q_0 values. The grey shaded band shows the combination of statistical and systematic uncertainty in event counts from data.
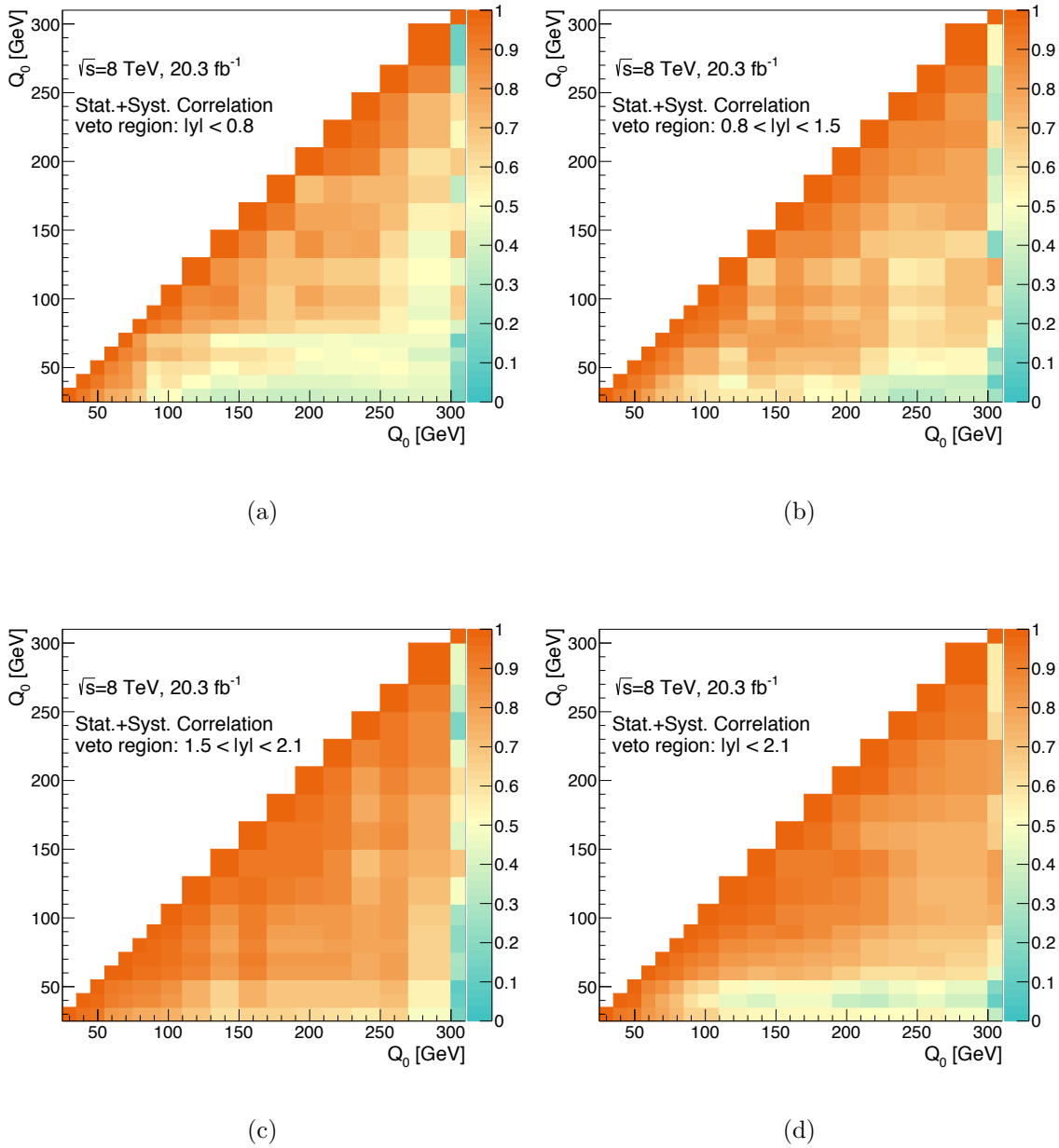
Figure 5.71: The correlation matrices from both systematic and statistical sources for the gap fraction measurement at different levels of Q_0 for the four rapidity regions of a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.
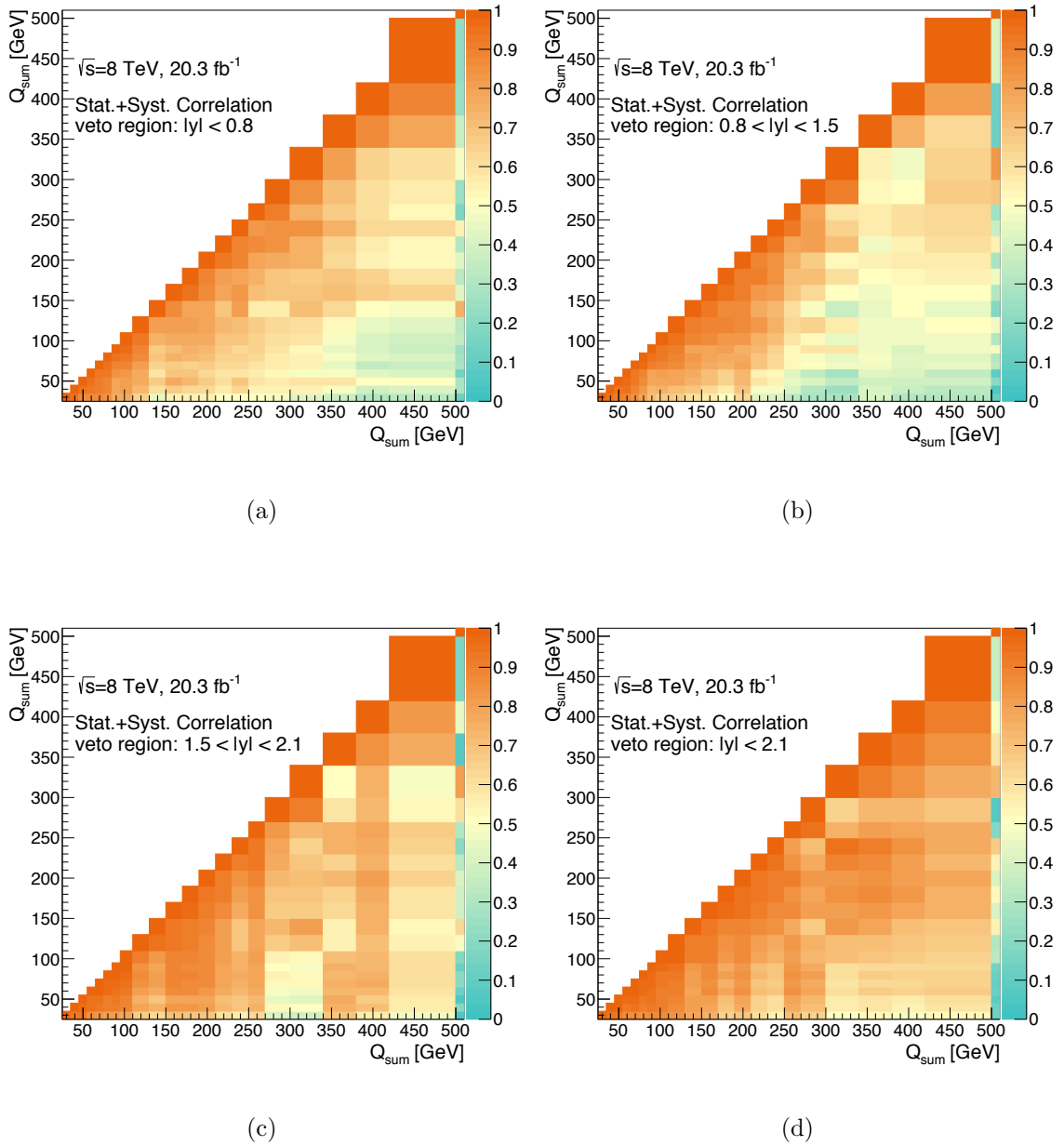
Figure 5.72: The correlation matrices from both systematic and statistical sources for the gap fraction measurement at different levels of Q_{sum} for the four rapidity regions of a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.
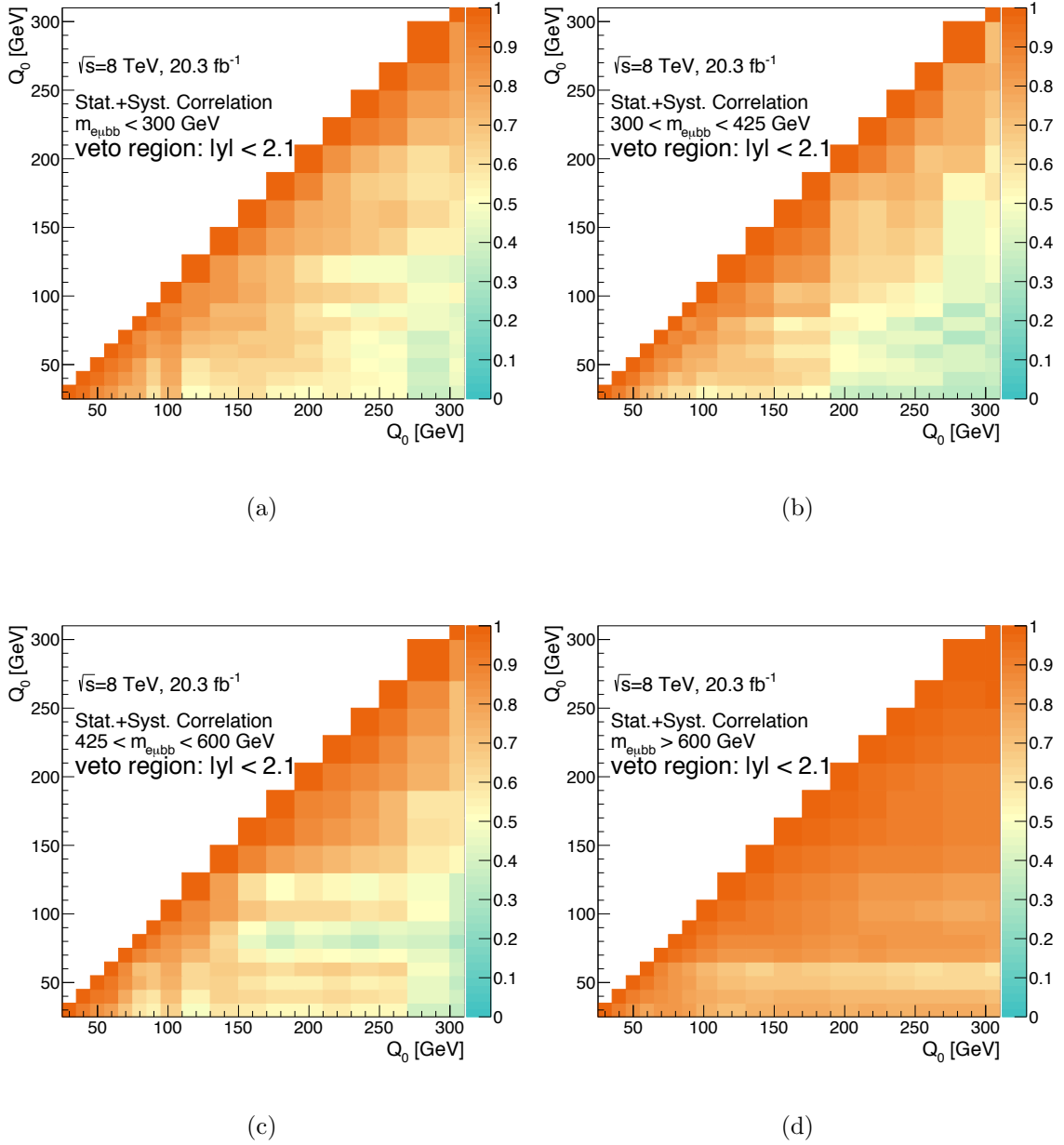
Figure 5.73: The correlation matrices from both systematic and statistical sources for the gap fraction measurement at different levels of Q_0 for the four invariant mass regions of a) $m_{e\mu bb} < 300$ GeV, b) $300 < m_{e\mu bb} < 425$ GeV, c) $425 < m_{e\mu bb} < 600$ GeV, and d) $m_{e\mu bb} > 600$ GeV.
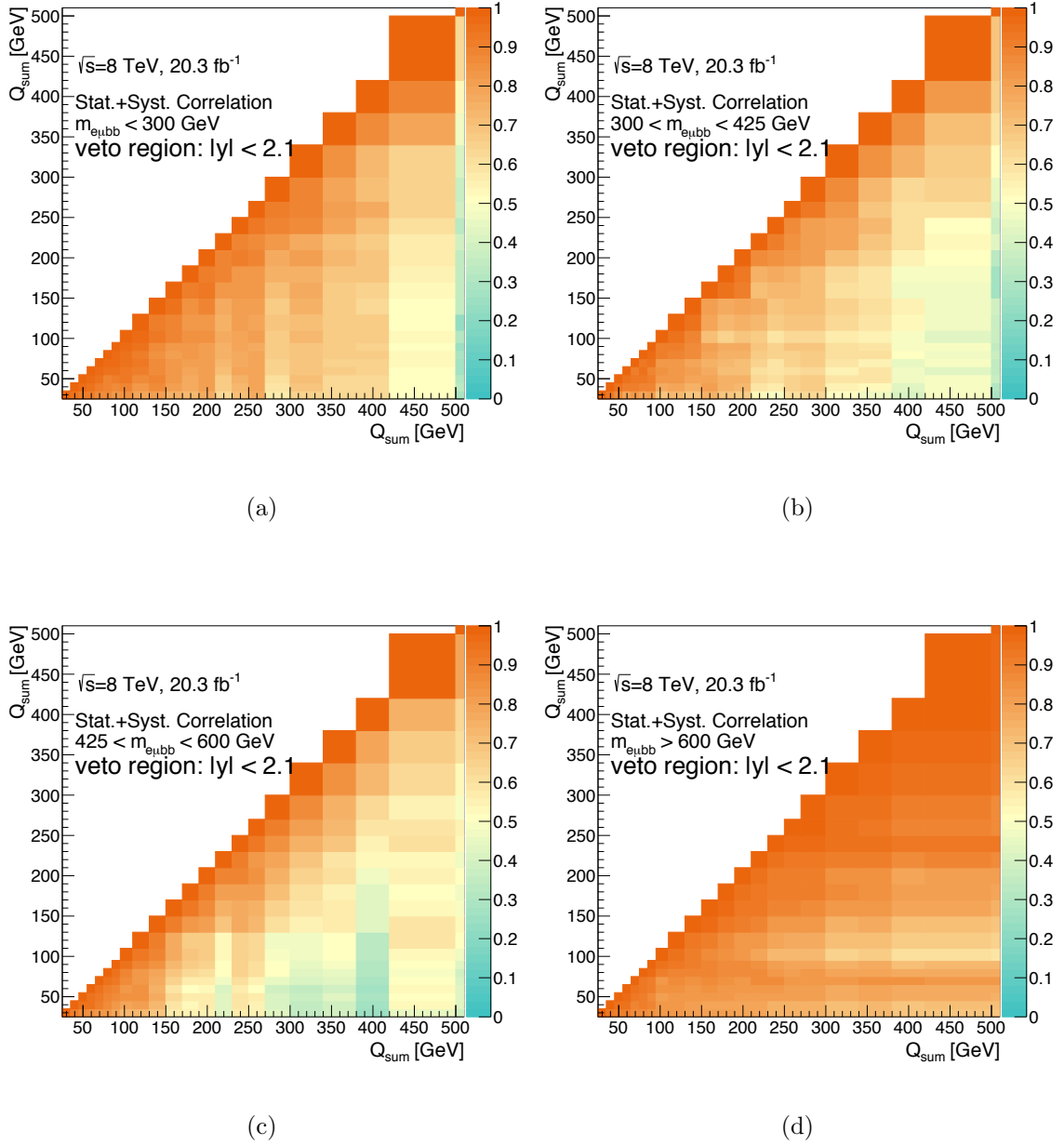
Figure 5.74: The correlation matrices from both systematic and statistical sources for the gap fraction measurement at different levels of Q_{sum} for the four invariant mass regions of a) $m_{e\mu bb} < 300$ GeV, b) $300 < m_{e\mu bb} < 425$ GeV, c) $425 < m_{e\mu bb} < 600$ GeV, and d) $m_{e\mu bb} > 600$ GeV.

170

	f(Q_0) (%)						
Q_0	Data	POWHEG	POWHEG+PY6	POWHEG+PY8	MC@NLO	POWHEG	ρ_j^i
[GeV]	±(stat.)±(syst.)	+PY6	$h_{\mathrm{damp}} = m_t$	$h_{\mathrm{damp}} = m_t$	+HW	+HW	(stat.+syst.)
veto region: $\|y\| < 0.8$							
25	$76.5\pm0.4^{+1.0}_{-1.1}$	76.0±0.2	78.1±0.2	76.1±0.2	79.1±0.2	74.6±0.2	$\rho_{75}^{25} = 0.66$
75	$93.2\pm0.2^{+0.3}_{-0.3}$	92.3±0.1	93.8±0.1	93.0±0.1	94.3±0.1	92.3±0.1	$\rho_{150}^{75} = 0.69$
150	$97.8\pm0.1^{+0.1}_{-0.1}$	97.3±0.1	98.0±0.1	97.8±0.1	98.3±0.1	97.4±0.1	$\rho_{25}^{150} = 0.40$
veto region: $0.8 < \|y\| < 1.5$							
25	$79.8\pm0.4^{+1.1}_{-1.1}$	79.7±0.2	81.6±0.2	80.1±0.2	81.8±0.2	79.2±0.2	$\rho_{75}^{25} = 0.70$
75	$94.5\pm0.2^{+0.3}_{-0.3}$	93.5±0.1	94.7±0.1	94.3±0.1	94.7±0.1	93.7±0.1	$\rho_{150}^{75} = 0.83$
150	$98.2\pm0.1^{+0.2}_{-0.1}$	97.8±0.1	98.3±0.1	98.3±0.1	98.3±0.1	97.9±0.1	$\rho_{25}^{150} = 0.60$
veto region: $1.5 < \|y\| < 2.1$							
25	$85.3\pm0.3^{+1.0}_{-0.9}$	84.9±0.2	86.1±0.2	85.4±0.2	85.5±0.2	84.7±0.2	$\rho_{75}^{25} = 0.78$
75	$96.0\pm0.2^{+0.4}_{-0.4}$	95.5±0.1	96.2±0.1	96.0±0.1	95.5±0.1	95.5±0.1	$\rho_{150}^{75} = 0.89$
150	$98.7\pm0.1^{+0.2}_{-0.2}$	98.6±0.1	98.9±0.0	98.9±0.0	98.6±0.1	98.6±0.1	$\rho_{25}^{150} = 0.66$
veto region: $\|y\| < 2.1$							
25	$53.9\pm0.5^{+1.6}_{-1.7}$	53.6±0.2	56.7±0.2	54.5±0.2	56.2±0.2	52.0±0.2	$\rho_{75}^{25} = 0.73$
75	$84.8\pm0.3^{+0.5}_{-0.6}$	82.9±0.2	85.8±0.2	85.0±0.2	85.3±0.2	83.0±0.2	$\rho_{150}^{75} = 0.83$
150	$94.9\pm0.2^{+0.3}_{-0.3}$	93.8±0.1	95.4±0.1	95.2±0.1	95.3±0.1	94.1±0.1	$\rho_{25}^{150} = 0.52$

Table 5.2: The measured gap fraction values $f(Q_0)$ for the four rapidity regions, for the Q_0 values of 25, 75 and 100 GeV in data compared to expectation from NLO generator simulations. The statistical and systematic uncertainties are given rounded to the nearest 0.1%. The combination of statistical and systematic correlations between measurements at $Q_0 = i$ and $Q_0 = j$ is given as ρ_j^i.

Q_{sum}	\multicolumn{6}{c}{f(Q_{sum}) (%)}	ρ^i_j
	Data	POWHEG	POWHEG+PY6	POWHEG+PY8	MC@NLO	POWHEG	
[GeV]	$\pm$(stat.)$\pm$(syst.)	+PY6	$h_{\mathrm{damp}}=m_t$	$h_{\mathrm{damp}}=m_t$	+HW	+HW	(stat.+syst.)
veto region: $|y| < 0.8$							
55	$88.1\pm0.3^{+0.5}_{-0.5}$	87.7 ± 0.2	89.5 ± 0.1	88.1 ± 0.2	90.2 ± 0.1	87.2 ± 0.2	$\rho^{55}_{150}=0.70$
150	$97.0\pm0.2^{+0.2}_{-0.2}$	96.4 ± 0.1	97.3 ± 0.1	96.9 ± 0.1	97.9 ± 0.1	96.5 ± 0.1	$\rho^{150}_{300}=0.70$
300	$99.4\pm0.1^{+0.1}_{-0.1}$	99.1 ± 0.0	99.4 ± 0.0	99.3 ± 0.0	99.6 ± 0.0	99.2 ± 0.0	$\rho^{300}_{55}=0.50$
veto region: $0.8 < |y| < 1.5$							
55	$90.5\pm0.3^{+0.6}_{-0.6}$	89.8 ± 0.1	91.3 ± 0.1	90.5 ± 0.1	91.4 ± 0.1	89.9 ± 0.1	$\rho^{55}_{150}=0.70$
150	$97.8\pm0.1^{+0.2}_{-0.2}$	97.2 ± 0.1	97.9 ± 0.1	97.7 ± 0.1	98.0 ± 0.1	97.4 ± 0.1	$\rho^{150}_{300}=0.50$
300	$99.6\pm0.1^{+0.0}_{-0.0}$	99.4 ± 0.0	99.6 ± 0.0	99.5 ± 0.0	99.7 ± 0.0	99.5 ± 0.0	$\rho^{300}_{55}=0.26$
veto region: $1.5 < |y| < 2.1$							
55	$93.1\pm0.2^{+0.8}_{-0.8}$	92.8 ± 0.1	93.8 ± 0.1	93.4 ± 0.1	93.2 ± 0.1	92.8 ± 0.1	$\rho^{55}_{150}=0.88$
150	$98.5\pm0.1^{+0.3}_{-0.3}$	98.3 ± 0.1	98.7 ± 0.1	98.6 ± 0.1	98.5 ± 0.1	98.3 ± 0.1	$\rho^{150}_{300}=0.67$
300	$99.8\pm0.0^{+0.1}_{-0.1}$	99.7 ± 0.0	99.8 ± 0.0	99.8 ± 0.0	99.8 ± 0.0	99.7 ± 0.0	$\rho^{300}_{55}=0.48$
veto region: $|y| < 2.1$							
55	$72.7\pm0.4^{+1.3}_{-1.3}$	71.8 ± 0.2	75.2 ± 0.2	73.3 ± 0.2	74.7 ± 0.2	71.1 ± 0.2	$\rho^{55}_{150}=0.87$
150	$91.3\pm0.3^{+0.5}_{-0.5}$	89.9 ± 0.1	92.2 ± 0.1	91.2 ± 0.1	92.6 ± 0.1	90.1 ± 0.1	$\rho^{150}_{300}=0.81$
300	$98.0\pm0.1^{+0.2}_{-0.2}$	97.3 ± 0.1	98.1 ± 0.1	97.7 ± 0.1	98.6 ± 0.1	97.5 ± 0.1	$\rho^{300}_{55}=0.62$

Table 5.3: The measured gap fraction values $f(Q_{\mathrm{sum}})$ for the four rapidity regions, for the Q_{sum} values of 55, 150, and 300 GeV in data compared to expectation from NLO generator simulations. The statistical and systematic uncertainties are given rounded to the nearest 0.1%. The combination of statistical and systematic correlations between measurements at $Q_{\mathrm{sum}} = i$ and $Q_{\mathrm{sum}} = j$ is given as ρ^i_j.

| Q_0 | | f(Q_0) (%) | | | | | ρ_j^i |
| | Data | POWHEG | POWHEG+PY6 | POWHEG+PY8 | MC@NLO | POWHEG | |
[GeV]	±(stat.)±(syst.)	+PY6	$h_{\mathrm{damp}} = m_t$	$h_{\mathrm{damp}} = m_t$	+HW	+HW	(stat.+syst.)
veto region: $|y| < 2.1$, $m_{e\mu bb} < 300$ GeV							
25	$56.0\pm0.6^{+2.0}_{-2.0}$	55.1 ± 0.3	57.8 ± 0.3	56.0 ± 0.3	57.3 ± 0.3	53.5 ± 0.3	$\rho_{75}^{25} = 0.74$
75	$86.7\pm0.5^{+0.7}_{-0.6}$	84.3 ± 0.3	86.9 ± 0.2	86.2 ± 0.2	86.4 ± 0.2	84.4 ± 0.2	$\rho_{150}^{75} = 0.76$
150	$95.7\pm0.3^{+0.3}_{-0.3}$	94.5 ± 0.2	95.9 ± 0.1	95.7 ± 0.1	95.9 ± 0.1	94.7 ± 0.2	$\rho_{25}^{150} = 0.60$
veto region: $|y| < 2.1$, $300 < m_{e\mu bb} < 425$ GeV							
25	$54.4\pm0.8^{+1.7}_{-1.8}$	53.5 ± 0.4	57.0 ± 0.4	54.6 ± 0.4	56.2 ± 0.4	52.2 ± 0.4	$\rho_{75}^{25} = 0.60$
75	$84.7\pm0.6^{+0.7}_{-0.7}$	82.7 ± 0.3	85.8 ± 0.3	84.9 ± 0.3	85.4 ± 0.3	83.0 ± 0.3	$\rho_{150}^{75} = 0.54$
150	$95.0\pm0.4^{+0.6}_{-0.6}$	93.8 ± 0.2	95.4 ± 0.2	95.2 ± 0.2	95.4 ± 0.2	94.1 ± 0.2	$\rho_{25}^{150} = 0.55$
veto region: $|y| < 2.1$, $425 < m_{e\mu bb} < 600$ GeV							
25	$47.6\pm1.3^{+1.8}_{-1.7}$	51.0 ± 0.7	54.2 ± 0.7	51.6 ± 0.6	53.4 ± 0.6	48.1 ± 0.7	$\rho_{75}^{25} = 0.58$
75	$79.0\pm1.0^{+1.0}_{-0.9}$	80.3 ± 0.5	83.7 ± 0.5	82.9 ± 0.5	82.5 ± 0.5	80.0 ± 0.5	$\rho_{150}^{75} = 0.45$
150	$92.7\pm0.7^{+0.8}_{-0.8}$	92.6 ± 0.3	94.4 ± 0.3	94.2 ± 0.3	94.0 ± 0.3	92.6 ± 0.3	$\rho_{25}^{150} = 0.57$
veto region: $|y| < 2.1$, $m_{e\mu bb} > 600$ GeV							
25	$45.9\pm2.3^{+4.0}_{-4.2}$	45.2 ± 1.2	49.8 ± 1.2	46.8 ± 1.2	51.0 ± 1.2	43.9 ± 1.2	$\rho_{75}^{25} = 0.84$
75	$81.7\pm2.0^{+3.6}_{-3.7}$	75.7 ± 1.0	80.3 ± 1.0	78.8 ± 1.0	79.8 ± 1.0	75.9 ± 1.1	$\rho_{150}^{75} = 0.87$
150	$92.4\pm1.3^{+2.8}_{-2.8}$	89.7 ± 0.7	92.6 ± 0.6	92.3 ± 0.6	91.5 ± 0.7	90.5 ± 0.7	$\rho_{25}^{150} = 0.76$

Table 5.4: The measured gap fraction values $f(Q_0)$ for the four invariant mass regions, for the Q_0 values of 25, 75 and 100 GeV in data compared to expectation from NLO generator simulations. The statistical and systematic uncertainties are given rounded to the nearest 0.1%. The combination of statistical and systematic correlations between measurements at $Q_0 = i$ and $Q_0 = j$ is given as ρ_j^i.

Q_{sum}			f(Q_{sum}) (%)				ρ_j^i
	Data	POWHEG	POWHEG+PY6	POWHEG+PY8	MC@NLO	POWHEG	
[GeV]	$\pm$(stat.)$\pm$(syst.)	+PY6	$h_{\mathrm{damp}}=m_t$	$h_{\mathrm{damp}}=m_t$	+HW	+HW	(stat.+syst.)
veto region: $\|y\| < 2.1$, $m_{e\mu bb} < 300$ GeV							
55	$75.0\pm0.6^{+1.5}_{-1.4}$	73.5 ± 0.3	76.7 ± 0.3	74.9 ± 0.3	76.1 ± 0.3	72.8 ± 0.3	$\rho^{55}_{150}=0.83$
150	$92.5\pm0.4^{+0.6}_{-0.5}$	91.0 ± 0.2	93.0 ± 0.2	92.2 ± 0.2	93.5 ± 0.2	91.1 ± 0.2	$\rho^{150}_{300}=0.76$
300	$98.3\pm0.2^{+0.2}_{-0.2}$	97.8 ± 0.1	98.4 ± 0.1	98.1 ± 0.1	98.9 ± 0.1	97.9 ± 0.1	$\rho^{300}_{55}=0.70$
veto region: $\|y\| < 2.1$, $300 < m_{e\mu bb} < 425$ GeV							
55	$72.8\pm0.7^{+1.2}_{-1.3}$	71.7 ± 0.4	75.2 ± 0.4	73.2 ± 0.4	74.8 ± 0.3	71.1 ± 0.4	$\rho^{55}_{150}=0.77$
150	$91.4\pm0.5^{+0.8}_{-0.8}$	90.0 ± 0.2	92.2 ± 0.2	91.2 ± 0.2	92.8 ± 0.2	90.1 ± 0.2	$\rho^{150}_{300}=0.62$
300	$98.1\pm0.2^{+0.2}_{-0.2}$	97.3 ± 0.1	98.1 ± 0.1	97.7 ± 0.1	98.6 ± 0.1	97.5 ± 0.1	$\rho^{300}_{55}=0.56$
veto region: $\|y\| < 2.1$, $425 < m_{e\mu bb} < 600$ GeV							
55	$67.4\pm1.2^{+2.2}_{-2.2}$	68.7 ± 0.6	72.5 ± 0.6	70.3 ± 0.6	71.5 ± 0.6	67.3 ± 0.6	$\rho^{55}_{150}=0.58$
150	$87.9\pm0.8^{+0.8}_{-0.8}$	87.9 ± 0.4	90.6 ± 0.4	89.5 ± 0.4	90.6 ± 0.4	87.7 ± 0.4	$\rho^{150}_{300}=0.58$
300	$96.4\pm0.5^{+0.3}_{-0.3}$	96.4 ± 0.2	97.4 ± 0.2	97.1 ± 0.2	98.0 ± 0.2	96.6 ± 0.2	$\rho^{300}_{55}=0.35$
veto region: $\|y\| < 2.1$, $m_{e\mu bb} > 600$ GeV							
55	$63.2\pm2.3^{+4.4}_{-4.3}$	62.6 ± 1.2	67.6 ± 1.1	65.0 ± 1.1	67.9 ± 1.1	61.9 ± 1.2	$\rho^{55}_{150}=0.84$
150	$87.3\pm1.7^{+3.1}_{-3.1}$	83.6 ± 0.9	87.6 ± 0.8	85.4 ± 0.8	87.3 ± 0.8	84.1 ± 0.9	$\rho^{150}_{300}=0.83$
300	$97.3\pm0.8^{+2.3}_{-2.3}$	94.5 ± 0.6	96.3 ± 0.5	95.7 ± 0.5	96.5 ± 0.5	95.1 ± 0.5	$\rho^{300}_{55}=0.77$

Table 5.5: The measured gap fraction values $f(Q_{\mathrm{sum}})$ for the four invariant mass regions, for the Q_{sum} values of 55, 150, and 300 GeV in data compared to expectation from NLO generator simulations. The statistical and systematic uncertainties are given rounded to the nearest 0.1%. The combination of statistical and systematic correlations between measurements at $Q_{\mathrm{sum}} = i$ and $Q_{\mathrm{sum}} = j$ is given as ρ_j^i.

Veto Region	POWHEG +PY6	POWHEG+PY6 $h_{\mathrm{damp}} = m_t$	POWHEG+PY8 $h_{\mathrm{damp}} = m_t$	MC@NLO +HW	POWHEG +HW
Q_0, $|y| < 0.8$	19.4 / 0.37	15.1 / 0.66	12.4 / 0.83	27.2 / 0.08	20.6 / 0.30
Q_0, $0.8 < |y| < 1.5$	37.8 / 0.00	22.1 / 0.23	17.4 / 0.49	23.0 / 0.19	25.8 / 0.11
Q_0, $1.5 < |y| < 2.1$	32.7 / 0.02	31.2 / 0.03	29.7 / 0.04	35.6 / 0.01	33.4 / 0.01
Q_0, $|y| < 2.1$	43.5 / 0.00	31.7 / 0.02	22.2 / 0.22	28.1 / 0.06	29.5 / 0.04
Q_{sum}, $|y| < 0.8$	35.1 / 0.04	32.3 / 0.07	27.0 / 0.21	40.5 / 0.01	31.0 / 0.10
Q_{sum}, $0.8 < |y| < 1.5$	54.4 / 0.00	37.9 / 0.02	32.7 / 0.07	39.2 / 0.01	40.8 / 0.01
Q_{sum}, $1.5 < |y| < 2.1$	41.1 / 0.01	36.2 / 0.03	32.7 / 0.07	47.1 / 0.00	44.4 / 0.00
Q_{sum}, $|y| < 2.1$	48.1 / 0.00	47.7 / 0.00	40.4 / 0.01	56.2 / 0.00	34.4 / 0.04

Table 5.6: The χ^2 / p-values for each simulation's prediction of the gap fraction compared to data at the corrected-to-particle level. The calculations are shown for each of the four veto regions, as a function of Q_0 (18 degrees of freedom) and Q_{sum} (22 degrees of freedom).

Mass Region [GeV] $|y| < 2.1$	Powheg +PY6	Powheg+PY6 $h_{\mathrm{damp}} = m_t$	Powheg+PY8 $h_{\mathrm{damp}} = m_t$	MC@NLO +HW	Powheg +HW
Q_0, $m_{e\mu bb} < 300$ GeV	30.9 / 0.03	17.5 / 0.49	14.1 / 0.72	20.7 / 0.29	23.0 / 0.19
Q_0, $300 < m_{e\mu bb} < 425$ GeV	43.2 / 0.00	27.9 / 0.06	24.8 / 0.13	26.8 / 0.08	32.8/ 0.02
Q_0, $425 < m_{e\mu bb} < 600$ GeV	20.2 / 0.32	36.2 / 0.01	25.8 / 0.10	30.3 / 0.03	14.1/ 0.72
Q_0, $m_{e\mu bb} > 600$ GeV	19.6 / 0.36	22.2 / 0.22	20.8 / 0.29	29.4 / 0.04	19.3/ 0.37
Q_{sum}, $m_{e\mu bb} < 300$ GeV	47.7 / 0.00	46.5 / 0.00	46.5 / 0.00	50.1 / 0.00	43.2 / 0.00
Q_{sum}, $300 < m_{e\mu bb} < 425$ GeV	34.5 / 0.04	28.0 / 0.18	27.4 / 0.20	28.9 / 0.15	26.7 / 0.22
Q_{sum}, $425 < m_{e\mu bb} < 600$ GeV	16.9 / 0.77	32.5 / 0.07	18.2 / 0.69	31.9 / 0.08	12.3 / 0.95
Q_{sum}, $m_{e\mu bb} > 600$ GeV	46.6 / 0.00	39.4 / 0.01	43.7 / 0.00	30.1 / 0.12	43.3 / 0.00

Table 5.7: The χ^2 / p-values for each simulation's prediction of the gap fraction compared to data at the corrected-to-particle level. The calculations are shown for the four invariant mass regions, as a function of each Q_0 (18 degrees of freedom) and Q_{sum} (22 degrees of freedom).

5.7   Conclusion

A precise measurement of jet activity in $t\bar{t}$ events as a gap fraction has been done as described in this chapter, using $t\bar{t}$ pair production events from the 2012 pp collision data sample collected by the ATLAS Detector at the LHC at $\sqrt{s} = 8$ TeV. The analysis corrected the data for detector effects and then compared the resulting gap fraction plots to various Monte Carlo event simulators. Given the small systematic and statistical uncertainty on the gap fraction, the results can be used to constrain the uncertainty in quark and gluon radiation modelling and aid in future tuning of QCD radiation simulation.

Chapter 6

Extraction of the Top Quark Pole Mass from $t\bar{t}$ Pair-Production Events in pp Collisions at $\sqrt{s} = 7 - 8$ TeV using NNLO+NNLL Predictions

"I would just like to point out that if the top quark mass is a little higher,

the universe will eventually die, but if it's a little lower, it will be stable.

So it kind of matters."

– Michael Peskin, on the relationship of m_t^{pole} and m_H

Direct measurement of the top quark mass m_t typically relies on the relation between m_t and an experimental observable, such as a reconstructed invariant mass, compared to results from events generated by Monte Carlo simulations as a function of m_t. However, a full theoretical definition of the top quark mass depends on the renormalization scheme beyond the leading order of QCD corrections, while the Monte Carlo inputs do not use a well-defined renormalization scheme – the Monte Carlo generators use matrix elements at leading (LO) or next-to-leading order (NLO) and simulate higher orders with parton showering. This lack of defined renormalization

contributes to uncertainty in the definition of the top quark mass being measured.

To measure the top quark mass according to a clearly defined renormalization scheme, this study compares the experimental cross section measurement of a top quark pair and its mass dependence, $\sigma_{t\bar{t}}(m_t)$, with three theoretically predicted $t\bar{t}$ cross sections (corresponding to three different parton distribution functions, PDFs), renormalized as the pole ("on-shell") mass $m_{\text{top}}^{\text{pole}}$ and computed at the next-to-next-to leading order (NNLO) of QCD corrections and including next-to-next-to-leading-logrithmic (NNLL) soft-gluon resummations. The experimental cross section chosen is nearly ideal for the purpose, since the $t\bar{t}$ cross section measurement using events with an opposite-sign $e\mu$ pair and b-tagged jets in the final state is nearly independent of the top mass definition; it depends on m_t through only two quantities, derived from Monte Carlo simulation: the efficiency for the selection of opposite-sign $e\mu$ events and the level of single top Wt background [95].

The following sections discuss the experimental cross section measurement (Section 6.1) and the predicted cross section (Section 6.2), followed by the probabilistic mass extraction in Section 6.3, cross-checks in Section 6.4, and a summary of the results in Section 6.5. The results have been published in the *European Physical Journal C* [40] and featured in the CERN Courier magazine, September 2014 edition [96].

6.1 Measured Cross Section

The following section describes the experimental cross section measurement $\sigma_{t\bar{t}}$ as a function of m_t, using the data and simulation samples described in Chapter 3 and

object and event selection from Chapter 4.

6.1.1 Central Mass Point Measurement

The measured $\sigma_{t\bar{t}}$ used for the mass extraction was done in the dileptonic decay channel by the ATLAS Collaboration using events including an opposite-sign $e\mu$ pair and one or more b-tagged jets in the final state, with the data and simulation samples described in Chapter 3 and object selection described in Chapter 4. The cross section $\sigma_{t\bar{t}}$ was solved in Ref. [40] simultaneously with the combined jet/b-tagging efficiency ϵ_b based on the equations:

$$N_1 = L\sigma_{t\bar{t}}\epsilon_{e\mu}2\epsilon_b(1 - C_b\epsilon_b) + N_1^{\text{bkg}} \tag{6.1}$$

$$N_2 = L\sigma_{t\bar{t}}\epsilon_{e\mu}C_b\epsilon_b^2 + N_2^{\text{bkg}} \tag{6.2}$$

where N_1 and N_2 represent the number of expected $e\mu$ events with 1 and 2 b-tagged jets, respectively, L is the integrated luminosity of the sample, $\sigma_{t\bar{t}}$ represents the $t\bar{t}$ production cross section, and $\epsilon_{e\mu}$ the efficiency for a $t\bar{t}$ event to pass the $e\mu$ selection. The ϵ_b variable represents the probability of a b jet from a top quark decay to have a trajectory within the acceptance of the detector, be reconstructed as a b-tagged jet and pass the object selection. The probability of two b jets to meet these criteria would naively be ϵ_b^2, but due to small physical and detector correlations is defined as $C_b = \epsilon_{bb}/\epsilon_b^2$. Assuming $m_t = 172.5$ GeV and solving Equation 6.2 simultaneously for $\sigma_{t\bar{t}}$ and ϵ_b, the observed cross section is 182.90 ± 7.11 pb at 7 TeV and 242.43 ± 10.34 pb at 8 TeV [95]. The uncertainties include data statistics, internal systematic effects, and knowledge limitations in integrated luminosity and LHC beam energy, resulting

in an overall uncertainty of 3.9% at 7 TeV and 4.3% at 8 TeV [95].

The cross section measurement assumes $m_t = 172.5$ GeV for the calculations from Monte Carlo simulations of the efficiency of selecting $e\mu$ events $\epsilon_{e\mu}$ and the single top Wt background. The efficiency ϵ_b and the other backgrounds do not depend on m_t; for example, the fake lepton backgrounds use same-sign $e\mu$ events from data with b-tagged jets instead of m_t-dependent Monte Carlo simulations.

6.1.2 Additional Mass Points

To find the dependence of the measured cross section on the assumed value of m_t, the analysis was re-done for six additional mass points: $m_t = 165, 167.5, 170, 175, 177.5$, and 180 GeV, with the MC simulation samples defined in Chapter 3. With the alternate $t\bar{t}$ samples, values of $\epsilon_{e\mu}$ and associated statistical uncertainty were obtained for each mass point. With the alternate Wt samples, a count of background Wt events with one and two b-tagged jets, represented as $N_{1,Wt\,\mathrm{bkg}}$ and $N_{2,Wt\,\mathrm{bkg}}$ respectively, was obtained for each mass point.

With these variables for each mass point, the sensitivity of the cross section $\sigma_{t\bar{t}}$ and associated uncertainty due to the efficiency $\epsilon_{e\mu}$ can be solved by following the process in Ref [40]. Equation 6.2 was solved by minimization with a likelihood calculation. The likelihood L of obtaining N_1 and N_2 given observed event counts N_1^{obs} and N_2^{obs} is written as the product of two Poisson likelihoods, and $-\ln L$ minimized using a Minuit [97] software program to obtain a fitted cross section $\sigma_{t\bar{t}}$. The derivative of the fitted cross section with respect to different parameters gives the dependence of

the fitted cross section on those parameters:

$$\frac{d\sigma_{t\bar{t}}\epsilon_{e\mu}}{d\epsilon_{e\mu}\sigma_{t\bar{t}}} = 1 \tag{6.3}$$

$$\frac{d\sigma_{t\bar{t}}C_b}{dC_b\sigma_{t\bar{t}}} = -1 \tag{6.4}$$

where, at 8 TeV for example, $\sigma_{t\bar{t}} = 242.1$ GeV and $\epsilon_{e\mu} = 0.81$ at the central mass point.

The background event counts $N_{1,Wt\,\mathrm{bkg}}$ and $N_{2,Wt\,\mathrm{bkg}}$, after being scaled to correspond with the luminosity at the central mass point, were combined with the other, m_t-independent, backgrounds to find $N_{1,\mathrm{bkg}}$ and $N_{2,\mathrm{bkg}}$, the total number of Wt events with one and two b-tagged jets, respectively, from all backgrounds. The sensitivity of the cross section due to the background event counts is similarly described from the fitted cross section with respect to N_1^{bkg} and N_2^{bkg}:

$$S_1^{\mathrm{bkg}} = \frac{d\sigma_{t\bar{t}}N_1^{\mathrm{bkg}}}{dN_1^{\mathrm{bkg}}\sigma_{t\bar{t}}} \tag{6.5}$$

$$S_2^{\mathrm{bkg}} = \frac{d\sigma_{t\bar{t}}N_2^{\mathrm{bkg}}}{dN_2^{\mathrm{bkg}}\sigma_{t\bar{t}}} \tag{6.6}$$

where $S_1^{\mathrm{bkg}} = -0.128$ and $S_2^{\mathrm{bkg}} = -0.0036$ from simulation at the central mass point at 8 TeV [40]. To find the uncertainty in the change in Wt background, the sample weight (due to the varying number of events in the different Monte Carlo samples) is combined with the scaled number of events $N_{1,Wt\,\mathrm{bkg}}$ and $N_{2,Wt\,\mathrm{bkg}}$:

$$dN_{i,Wt\,\mathrm{bkg}} = \sqrt{\frac{N_{i,Wt\,\mathrm{bkg}}}{\mathrm{sample\,weight}}} \tag{6.7}$$

where i corresponds to 1 and 2 for the two background event counts. The results added in quadrature give the total uncertainty in the change in Wt background.

(a)   (b)
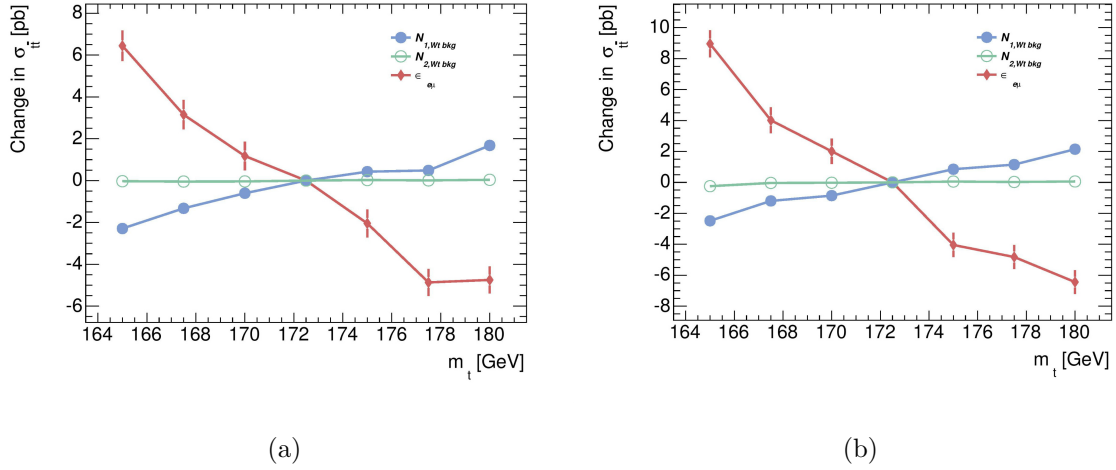
Figure 6.1: Contributions to the change in cross section measurement as a function of the assumed value of m_t, where $N_{1,Wt\,\mathrm{bkg}}$ and $N_{2,Wt\,\mathrm{bkg}}$ refer to the Wt background event counts including 1 and 2 b-tagged jets, respectively. Shown for a) 7 TeV and b) 8 TeV.

Figure 6.1 and Tables 6.1 and 6.2 show the resulting cross section measurement as a function of assumed m_t as a result of varying $\epsilon_{e\mu}$ and Wt backgrounds at 7 and 8 TeV. The combination of each source of variation gives the total change in cross section for each mass point.

6.1.3   Fitting and Mass Dependence

To model the mass dependence of the measurement, a polynomial fit was then applied to the experimental cross section data points, shown in Figure 6.2. The fit uses the same polynomial form as the theory curves (Equation 6.11), with $m_t^{\mathrm{ref}} = 172.5$ GeV and including only the uncorrelated statistical uncertainties (shown in Tables 6.1

183

Mass (GeV)	$\Delta\sigma_{t\bar{t}}(e\mu$ efficiency) (pb)	$\Delta\sigma_{t\bar{t}}(Wt$ background) (pb)	Combined $\Delta\sigma_{t\bar{t}}$ (pb)
165.0	6.44±0.73	-2.32±0.04	4.12±0.73
167.5	3.15±0.71	-1.37±0.04	1.78±0.71
170.0	1.18±0.70	-0.64±0.04	0.53±0.69
175.0	-2.05±0.67	0.45±0.04	-1.59±0.67
177.5	-4.87±0.65	0.49±0.04	-4.37±0.65
180.0	-4.75±0.65	1.72±0.04	-3.03±0.65

Table 6.1: Contributions to the change in cross section measurement at 7 TeV, with associated uncertainties from Monte Carlo statistics, due to the change in $e\mu$ efficiency ($\epsilon_{e\mu}$) and the Wt background estimations.

and 6.2). Table 6.3 gives the resulting fit parameters. The mass dependence can be expressed as the derivative of the fit equation:

$$\frac{\mathrm{d}\sigma_{t\bar{t}}(m_t)}{\mathrm{d}m_t} = \sigma_{t\bar{t}}(m_t^{\mathrm{ref}})\frac{m_t^{\mathrm{ref}\,2}}{m_t^5}\left(4(m_t^{\mathrm{ref}})^2(a_1 - a_2 - 1) - 3m_t^{\mathrm{ref}}m_t(a_1 - 2a_2) - 2a_2 m_t^2\right)$$

(6.8)

At the central mass point $m_t = 172.5$ GeV at $\sqrt{s} = 8$ TeV, this simplifies to:

$$\frac{\mathrm{d}\sigma_{t\bar{t}}(m_t)}{\mathrm{d}m_t}\bigg|_{m_t = m_t^{\mathrm{ref}}} = \frac{\sigma_{t\bar{t}}(m_t^{\mathrm{ref}})}{m_t^{\mathrm{ref}}}(a_1 - 4) = -0.69 \pm 0.06 \frac{\mathrm{pb}}{\mathrm{GeV}} = -0.28 \pm 0.03 \frac{\%}{\mathrm{GeV}}$$

(6.9)

To calculate the uncertainty in the fit σ_{fit}, the standard error propagation was used for a function of correlated variables, a_1 and $\sigma_{t\bar{t}}^{\mathrm{ref}} = \sigma_{t\bar{t}}(m_t^{\mathrm{ref}})$, where $f(a_1, \sigma_{t\bar{t}}^{\mathrm{ref}}) = \frac{\mathrm{d}\sigma_{t\bar{t}}(m_t)}{\mathrm{d}m_t}$ [98] (the a_2 dependence cancels out at the central mass point):

$$\sigma_{\mathrm{fit}}^2 = \left(\frac{df(a_1, \sigma_{t\bar{t}}^{\mathrm{ref}})}{d\sigma_{t\bar{t}}^{\mathrm{ref}}}\right)^2 \sigma_{t\bar{t}}^{\mathrm{ref}\,2} + \left(\frac{df(a_1, \sigma_{t\bar{t}}^{\mathrm{ref}})}{da_1}\right)^2 \sigma_{a_1}^2 + 2\frac{df(a_1, \sigma_{t\bar{t}}^{\mathrm{ref}})}{d\sigma_{t\bar{t}}^{\mathrm{ref}}}\frac{df(a_1, \sigma_{t\bar{t}}^{\mathrm{ref}})}{da_1}\mathrm{cov}(a_1, \sigma_{t\bar{t}}^{\mathrm{ref}})$$

(6.10)

Mass (GeV)	$\Delta\sigma_{t\bar{t}}(e\mu \text{ efficiency})$ (pb)	$\Delta\sigma_{t\bar{t}}(Wt \text{ background})$ (pb)	Combined $\Delta\sigma_{t\bar{t}}$ (pb)
165.0	8.96±0.87	-2.51±0.02	6.45±0.87
167.5	4.01±0.84	-1.23±0.02	2.78±0.84
170.0	2.01±0.82	-0.88±0.02	1.13±0.82
175.0	-4.04±0.78	0.90±0.02	-3.14±0.78
177.5	-4.83±0.78	1.18±0.02	-3.65±0.78
180.0	-6.44±0.77	2.20±0.02	-4.24±0.77

Table 6.2: Contributions to the change in cross section measurement at 8 TeV, with associated uncertainties from Monte Carlo statistics, due to the change in $e\mu$ efficiency ($\epsilon_{e\mu}$) and the Wt background estimations.

At $\sqrt{s} = 7$ TeV, by the same calculation, the mass dependence is $-0.28 \pm 0.03 \frac{\%}{\text{GeV}}$.

$\sqrt{s}$ GeV	a_1	a_2	$\sigma_{t\bar{t}}(m_t^{\text{ref}})$
7	3.51	3.79	182.89
8	3.51	5.42	242.43

Table 6.3: Parameters used to fit the cross section measurement to mass points at 7 and 8 TeV, using $m_t^{\text{ref}} = 172.5$ GeV, according to the polynomial given in Equation 6.11.

6.2 Predicted Cross Section

The theoretical prediction of the $t\bar{t}$ cross section for pp collisions used in this analysis is $\sigma_{t\bar{t}} = 177^{+10}_{-11}$ pb at a center-of-mass energy of $\sqrt{s} = 7$ TeV and $\sigma_{t\bar{t}} = 253^{+13}_{-15}$ pb at $\sqrt{s} = $
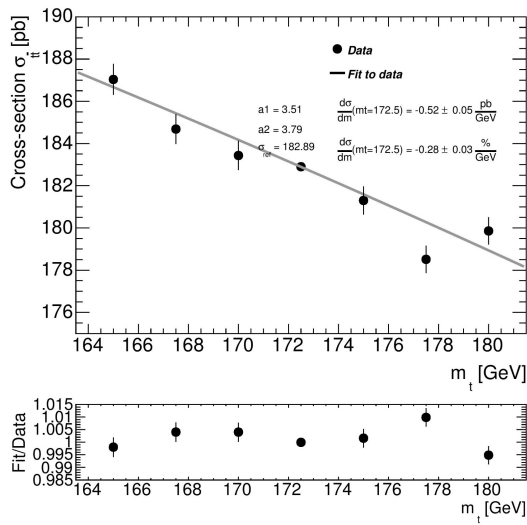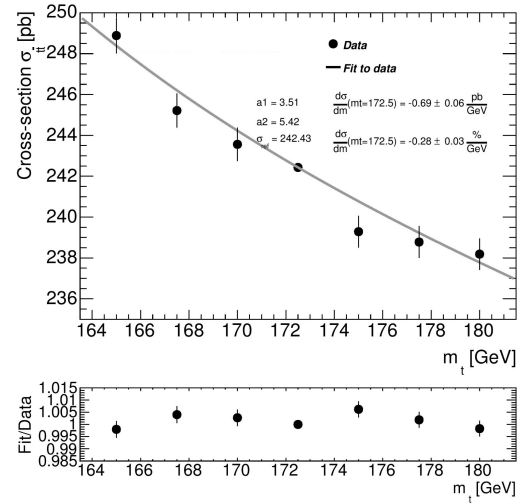
Figure 6.2: Polynomial fit to the experimental data points at a) 7 TeV and b) 8 TeV, with associated uncorrelated statistical uncertainty. The accompanying lower plots show the ratio between the fit and the data points.

8 TeV for $m_t = 172.5$ GeV. It has been calculated at NNLO order in QCD including resummation of NNLL soft gluon terms with top++2.0 [99, 100, 101, 102, 23, 103]. The PDF and α_S uncertainties were calculated using the PDF4LHC prescription [104] with the MSTW2008 68% CL NNLO [105, 106], CT10 NNLO [62, 107] and NNPDF2.3 5f FFN [108] PDF sets, added in quadrature to the scale uncertainty. The NNLO+NNLL value is about 3% larger than the exact NNLO prediction, as implemented in Hathor 1.5 [109]. These computations use pole mass renormalization, defining this extraction as that of the pole mass.

These theoretical results can be used to predict the cross section for any top mass. Doing these calculations for various masses and fitting the results with the polynomial given in Equation 6.11, using $m_t^{\text{ref}} = 172.5$ GeV, gives a relationship for the theoretical dependence of $\sigma_{t\bar{t}}(m_t)$ [110].

$$\sigma_{t\bar{t}}(m_t) = \sigma_{t\bar{t}}(m_t^{\text{ref}}) \left(\frac{m_t^{\text{ref}}}{m_t} \right)^4 \left[1 + a_1 \frac{m_t - m_t^{\text{ref}}}{m_t^{\text{ref}}} + a_2 \left(\frac{m_t - m_t^{\text{ref}}}{m_t^{\text{ref}}} \right)^2 \right] \quad (6.11)$$

Tables 6.4 and 6.5 show the parameters corresponding to the polynomial fit for three different NNLO PDF sets for comparison, MSTW 2008 NNLO, CT10 NNLO, and NNPDF 2.3 (NNLO 5f FFN). The scale and PDF+α_s uncertainties for each PDF set were calculated at 68% confidence level and fit with the same polynomial (not shown in tables), where α_s represents the strong coupling constant [110].

Figure 6.3 shows the resulting $\sigma_{t\bar{t}}$ for each PDF and center-of-mass energy for 7 (6.3a) and 8 TeV (6.3b), as a function of m_t^{pole}. Figure 6.4 shows the spread of uncertainties in the $\sigma_{t\bar{t}}$ results for 7 (6.4a) and 8 TeV (6.4b) (not shown in Table 6.4 and Table 6.5), with the scale and PDF uncertainties added quadratically. The NNPDF

PDF	α_s	α_s uncertainty	a_1	a_2	$\sigma_{t\bar{t}}(m_t^{\mathrm{ref}})$ (pb)
MSTW 2008 NNLO	0.1176	± 0.0014	-1.2149	0.8746	176.2278
CT10 NNLO	0.1180	± 0.0012	-1.2114	0.7513	176.6800
NNPDF2.3 (NNLO 5f FFN)	0.1180	± 0.0012	-1.2674	0.8901	176.9559

Table 6.4: Parameters for theoretical NNLO+NNLL predictions from MSTW 2008 NNLO [105, 106], CT10 NNLO [62, 107], and NNPDF2.3 (NNLO 5f FFN) [108] at $\sqrt{s} = 7$ TeV [110].

prediction's lower uncertainty is likely due to the higher number of included data sets in that PDF. The MSTW prediction extracts α_s simultaneously in the PDF fit and uses the associated uncertainty, while CT10 and NNPDF use the (slightly lower) PDG uncertainty definition for α_s [111].

This analysis method assumes that although the top quark mass used by the Monte Carlo simulations m_t does not require a defined renormalization scheme, it is equivalent to the pole mass, m_t^{pole}. If m_t and m_t^{pole} differ, the experimental curve of the cross section as a function of m_t would shift relative to the theoretical curves shown in Figure 6.3. Because of this, it is advantageous that the slope of the experimental curve is so small. For example, using the polynomial fit to experimental data points, shifting the experimental curve by 5 GeV on either side of 172.5 GeV shifts the experimental cross section by 2.5 ± 0.3 pb at 7 TeV and 3.8 ± 0.3 pb at 8 TeV. At 172.5 GeV, the theoretical curves have a slope of $5 - 7$ pb/GeV, so this would cause the extracted mass value to shift less than 1 GeV, well within the experimental and theoretical uncertainties.

PDF	α_s	α_s uncertainty	a_1	a_2	$\sigma_{t\bar{t}}(m_t^{\text{ref}})$ (pb)
MSTW 2008 NNLO	0.1176	±0.0014	-1.0926	0.7079	251.6590
CT10 NNLO	0.1180	±0.0012	-1.0895	0.8099	252.1426
NNPDF 2.3 (NNLO 5f FFN)	0.1180	±0.0012	-1.1397	0.7279	254.0692

Table 6.5: Parameters for theoretical NNLO+NNLL predictions from MSTW 2008 NNLO [105, 106], CT10 NNLO [62, 107], and NNPDF2.3 (NNLO 5f FFN) [108] at $\sqrt{s} = 8$ TeV [110].

(a)   (b)
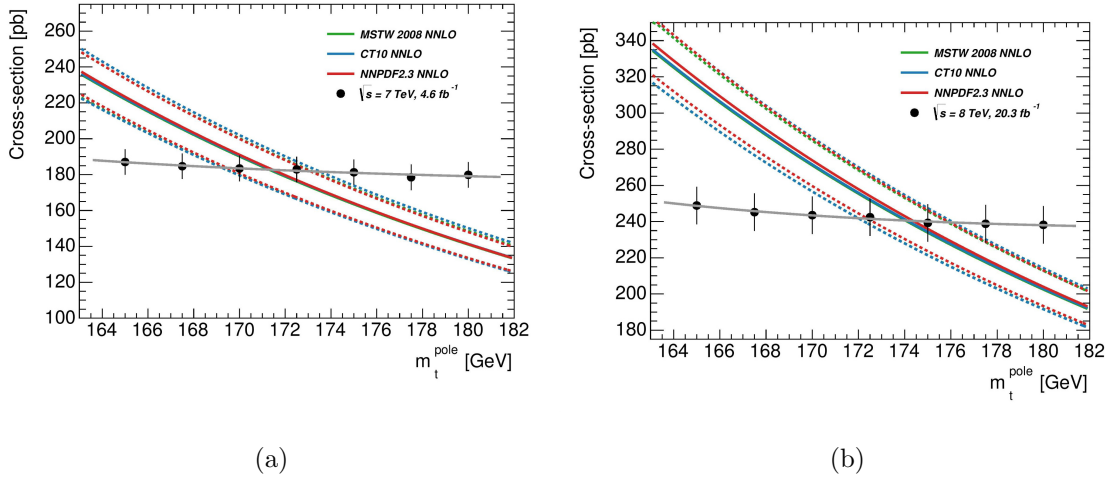
Figure 6.3: Predicted NNLO+NNLL cross sections compared with the measured cross sections from Figure 6.2, as functions of the top quark pole mass m_t^{pole} for a) 7 TeV and b) 8 TeV. The three different NNLO PDF sets are indicated by the three differently colored lines. The dashed lines represent the associated uncertainties at 68% confidence level (CL). The plots also show the experimental data points and accompanying polynomial fit from Equation 6.11; the experimental error bars include all sources of uncertainty given in Ref. [95] as well as uncorrelated statistical uncertainty.
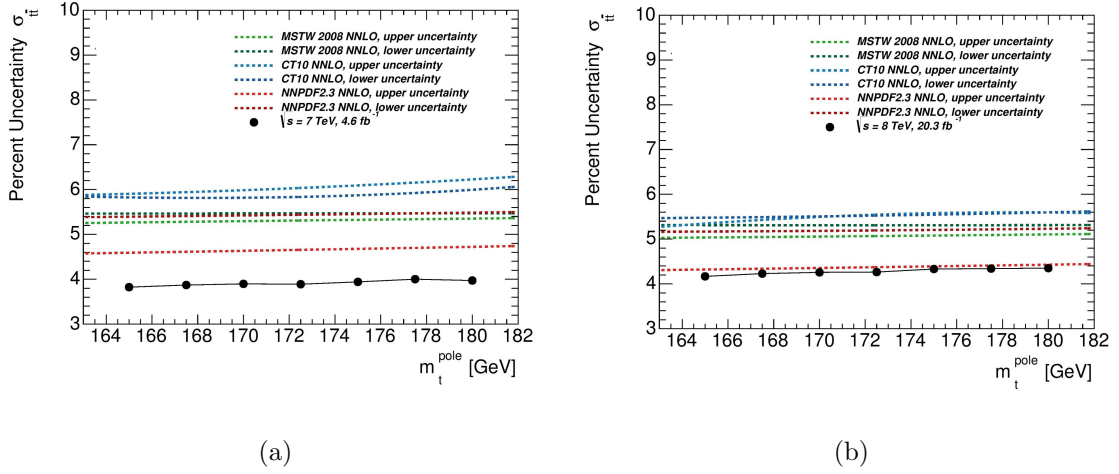
Figure 6.4: Associated uncertainties for the predicted and measured cross sections $\sigma_{t\bar{t}}$, by percentage as a function of m_t^{pole} at a) 7 TeV and b) 8 TeV. The relative uncertainty on the measured cross section σ_{tt} includes all sources of uncertainty given in Ref. [95] as well as uncorrelated statistical uncertainty from mass variation, and is assumed to be independent of the assumed m_t.

6.3 Maximum Likelihood Fitting

6.3.1 Method

To extract $m_{\mathrm{t}}^{\mathrm{pole}}$ from the combination of experimental (exp) and theoretical (th) results, including associated uncertainties, this study uses a combined Bayesian likelihood fit constructed as shown in Equation 6.12:

$$L(m_t) \propto \int G\left(\sigma_{t\bar{t}} \,|\, \sigma_{t\bar{t}}^{\mathrm{exp}}(m_t), \rho_{\mathrm{exp}}\right) \cdot G\left(\sigma_{t\bar{t}} \,|\, \sigma_{t\bar{t}}^{\mathrm{th}}(m_t), \rho_{\mathrm{th}}(m_t)\right) d\sigma_{t\bar{t}} \qquad (6.12)$$

In this equation:

- $G\left(\sigma_{t\bar{t}} \,|\, \sigma_{t\bar{t}}^{\mathrm{exp}}(m_t), \rho_{\mathrm{exp}}\right)$ represents the probability density function of the experimental results, constructed using a Gaussian likelihood function centered on $\sigma_{t\bar{t}}^{\mathrm{exp}}(m_t)$ and half the experimental uncertainty envelope, ρ_{exp}, as the width.

- $G\left(\sigma_{t\bar{t}} \,|\, \sigma_{t\bar{t}}^{\mathrm{th}}(m_t), \rho_{\mathrm{th}}(m_t)\right)$ represents the probability density function of the theoretical results. Since the theoretical predictions have asymmetrical uncertainties, this function is constructed as an asymmetrical Gaussian likelihood function centered on $\sigma_{t\bar{t}}^{\mathrm{th}}(m_t)$ and using either the upper or lower uncertainty, as shown below, as the width of the Gaussian $\rho_{\mathrm{th}}(m_t)$. This asymmetrical Gaussian is normalized so that the function is continuous at the mean and the area under the entire curve integrates to unity, by the expression:

$$G\left(\sigma_{t\bar{t}} \,|\, \sigma_{t\bar{t}}^{\mathrm{th}}(m_t), \rho_{\mathrm{th}}(m_t)\right) = \frac{2}{\sqrt{2\pi}(\rho_{\mathrm{th}}^{\mathrm{up}}(m_t) + \rho_{\mathrm{th}}^{\mathrm{dn}}(m_t))} \exp\left[-\frac{1}{2}\left(\frac{\sigma_{t\bar{t}} - \sigma_{t\bar{t}}^{\mathrm{th}}}{\rho_{\mathrm{th}}(\mathrm{m_t})}\right)^2\right] \quad (6.13)$$

where $\rho_{\mathrm{th}}(m_t) = \rho_{\mathrm{th}}^{\mathrm{up}}(m_t)$ for $\sigma_{t\bar{t}} > \sigma_{t\bar{t}}^{\mathrm{th}}(m_t)$, and $\rho_{\mathrm{th}}(m_t) = \rho_{\mathrm{th}}^{\mathrm{dn}}(m_t)$ for $\sigma_{t\bar{t}} < \sigma_{t\bar{t}}^{\mathrm{th}}(m_t)$.

Figures 6.5a and 6.5b show the probability density functions for theory and experiment

(a) (b)
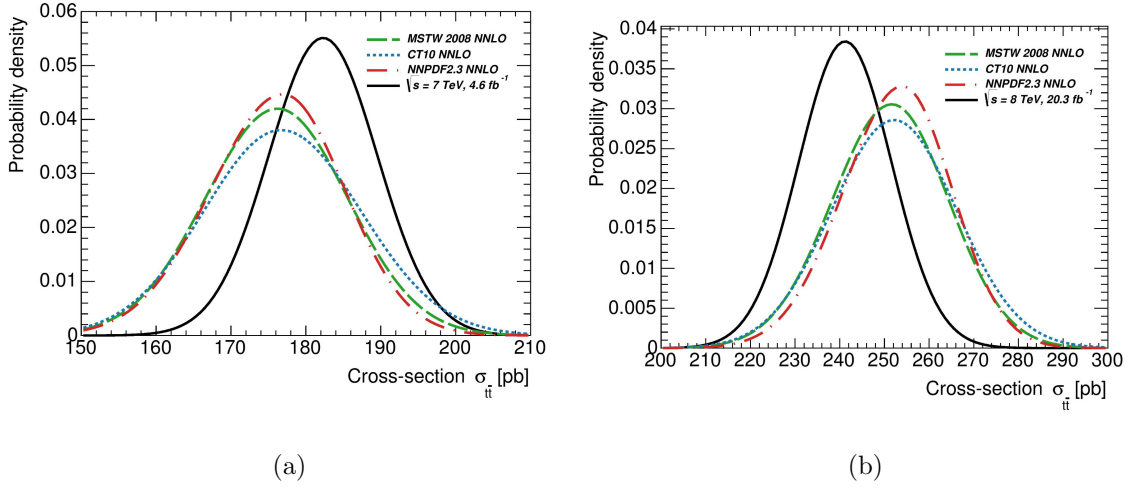
Figure 6.5: Gaussian probability distributions for the predicted (NNLO+NNLL) and measured cross section at a) 7 TeV and b) 8 TeV, with $m_t = 172.5$ GeV.

at the central mass point, $m_t = 172.5$ GeV.

The likelihood is maximized at the minimum value of $-\ln L(m_t)$, with the uncertainty given by $\ln L(\hat{m}_t) + 0.5$, where $\hat{m}_t$ represents the m_t corresponding to maximum likelihood. Figures 6.6a and 6.6b show the results for 7 and 8 TeV, respectively, using the three theoretical curves discussed in Section 6.2.

$$\text{(a)} \qquad\qquad\qquad\qquad\qquad \text{(b)}$$
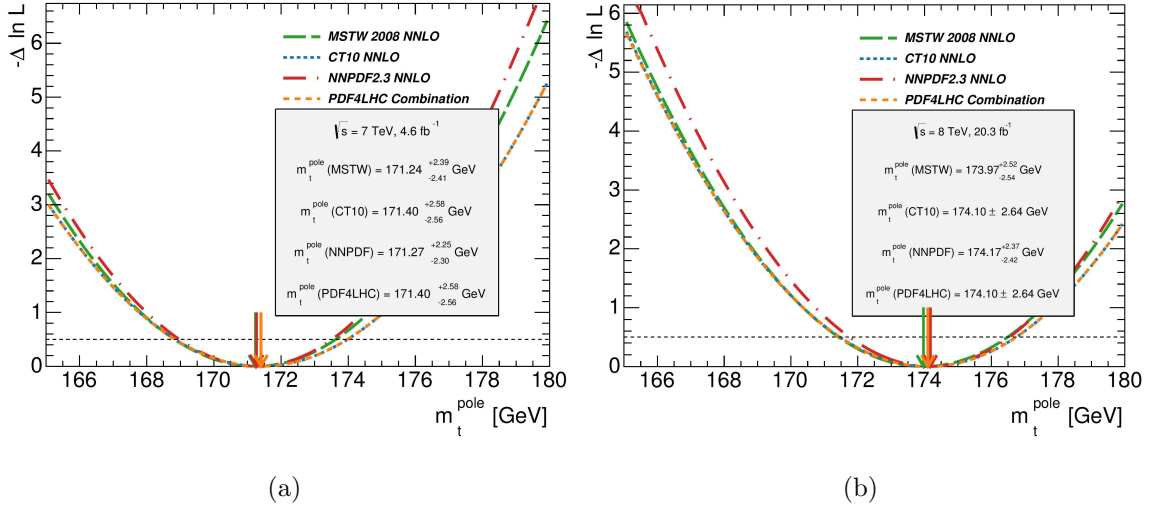
Figure 6.6: Extracted top pole mass m_t^{pole} for the three predicted NNLO+NNLL solutions and the PDF4LHC solution with the measured results, at a) 7 TeV and b) 8 TeV. The correspondingly colored dashed lines represent the spread of uncertainties.

Given the three theoretical PDFs used, the result with full uncertainty is:

$$m_{\text{top}}^{\text{pole}}(7 \text{ TeV}) = 171.24^{+2.39}_{-2.41} \text{ GeV (MSTW)} \tag{6.14}$$

$$171.40^{+2.58}_{-2.56} \text{ GeV (CT10)} \tag{6.15}$$

$$171.27^{+2.25}_{-2.30} \text{ GeV (NNPDF)} \tag{6.16}$$

$$m_{\text{top}}^{\text{pole}}(8 \text{ TeV}) = 173.97^{+2.52}_{-2.54} \text{ GeV (MSTW)} \tag{6.17}$$

$$174.10 \pm 2.64 \text{ GeV (CT10)} \tag{6.18}$$

$$174.17^{+2.37}_{-2.42} \text{ GeV (NNPDF)} \tag{6.19}$$

193

6.3.2 PDF4LHC Combined Result

The results from the three PDFs can be combined by following the PDF4LHC convention of including the entire uncertainty envelope comprised by the three PDFs, retaining the CT10 PDF as the central value [104]. This method finds the lowest lower uncertainty and highest upper uncertainty for each PDF. The difference between this value and the CT10 central value is used as the lower and upper theoretical uncertainty, respectively, according to the asymmetrical Gaussian method described above:

$$m_{\text{top}}^{\text{pole}}(7 \text{ TeV}) = 171.40^{+2.58}_{-2.56} \text{ GeV} \tag{6.20}$$

$$m_{\text{top}}^{\text{pole}}(8 \text{ TeV}) = 174.10 \pm 2.64 \text{ GeV} \tag{6.21}$$

It happens that this matches the CT10 result, since the CT10 PDF has the lowest and highest uncertainty values of the three PDFs at almost all points.

6.3.3 Combination of 7 and 8 TeV Results

The combination of 7 and 8 TeV results includes consideration of the correlated and uncorrelated uncertainties on the experimental and theoretical cross sections. The theoretical uncertainty is considered completely correlated. The experimental uncertainty includes a combination of uncorrelated and correlated components. Uncertainty from data statistics and knowledge of the luminosity are considered uncorrelated, while uncertainty from experimental systematic effects and LHC beam energy are considered

correlated:

$$\sigma_{t\bar{t}}\,(7\ \mathrm{TeV}) = 182.90 \pm 4.75\,(\mathrm{uncorr})\ \pm 5.34\,(\mathrm{corr})\,\mathrm{pb} \tag{6.22}$$

$$\sigma_{t\bar{t}}\,(8\ \mathrm{TeV}) = 242.43 \pm 7.69\,(\mathrm{uncorr})\ \pm 6.92\,(\mathrm{corr})\,\mathrm{pb} \tag{6.23}$$

The combined likelihood was found by doing a likelihood fitting at 7 and 8 TeV simultaneously and including nuisance parameters p_0 and p_1 for correlated theoretical and correlated experimental error, respectively. This prescription also uses the asymmetrical Bayesian method:

$$L \propto \int G\left(\sigma_{t\bar{t}}^{7\ \mathrm{TeV}} | \sigma_{t\bar{t}}^{\mathrm{exp},7\ \mathrm{TeV}}(m_t), \rho_{\mathrm{uncorr}}^{\mathrm{exp},7\ \mathrm{TeV}}\right) G\left(\sigma_{t\bar{t}}^{7\ \mathrm{TeV}} | \sigma_{t\bar{t}}^{\mathrm{th},7\ \mathrm{TeV}}(m_t), \rho_{\mathrm{uncorr}}^{\mathrm{th},7\ \mathrm{TeV}}\right) d\sigma_{t\bar{t}}^{7\ \mathrm{TeV}}$$
$$\tag{6.24}$$

$$\cdot \int G\left(\sigma_{t\bar{t}}^{8\ \mathrm{TeV}} | \sigma_{t\bar{t}}^{\mathrm{exp},8\ \mathrm{TeV}}(m_t), \rho_{\mathrm{uncorr}}^{\mathrm{exp},8\ \mathrm{TeV}}\right) G\left(\sigma_{t\bar{t}}^{8\ \mathrm{TeV}} | \sigma_{t\bar{t}}^{\mathrm{th},8\ \mathrm{TeV}}(m_t), \rho_{\mathrm{uncorr}}^{\mathrm{th},8\ \mathrm{TeV}}\right) d\sigma_{t\bar{t}}^{8\ \mathrm{TeV}}$$
$$\tag{6.25}$$

$$\cdot\, G(p_0|0,1) \cdot G(p_1|0,1) \tag{6.26}$$

Each cross section includes the correlated error:

$$\sigma_{t\bar{t}}^{\mathrm{th,x}} = \sigma_{t\bar{t}}^{\mathrm{th,x,orig}} + \beta^{\mathrm{th,x}} p_0 \tag{6.27}$$

$$\sigma_{t\bar{t}}^{\mathrm{exp,x}} = \sigma_{t\bar{t}}^{\mathrm{exp,x,orig}} + \beta^{\mathrm{exp,x}} p_1 \tag{6.28}$$

where $\beta^{\mathrm{th,x}}$ represents the correlated theoretical uncertainty at x TeV (assuming all theoretical uncertainty is correlated), and $\beta^{\mathrm{exp,x}}$ the correlated experimental uncertainty at x TeV. The ρ variable represents the uncorrelated experimental uncertainty.

Since all theoretical uncertainty is assumed to be correlated, $\rho_{\mathrm{uncorr}}^{\mathrm{th,x}} = 0$, the

195

theoretical Gaussian function becomes a delta function, simplifying the equation:

$$L \propto G\left(\sigma_{t\bar{t}}^{\mathrm{th},7\ \mathrm{TeV}}(m_t)|\sigma_{t\bar{t}}^{\mathrm{exp},7\ \mathrm{TeV}}(m_t), \rho_{\mathrm{uncorr}}^{\mathrm{exp},7\ \mathrm{TeV}}\right) G\left(\sigma_{t\bar{t}}^{\mathrm{th},8\ \mathrm{TeV}}(m_t)|\sigma_{t\bar{t}}^{\mathrm{exp},8\ \mathrm{TeV}}(m_t), \rho_{\mathrm{uncorr}}^{\mathrm{exp},8\ \mathrm{TeV}}\right)$$

(6.29)

$$\cdot\, G(p_0|0,1)G(p_1|0,1)$$

(6.30)

Asymmetrical errors are again used for the correlated theoretical uncertainty, such that $\beta^{\mathrm{th}} = \rho_{\mathrm{th}}^{\mathrm{upper}}(m_t)$ for $p_0 > 0$ and $\beta^{\mathrm{th}} = \rho_{\mathrm{th}}^{\mathrm{lower}}(m_t)$ for $p_0 < 0$.

Using the ROOT [112] implementation of the MINUIT minimization function and error analysis code [97] to find the minimum of $-\ln(L)$, including consideration of nuisance parameters, the result is:

$$m_{\mathrm{t}}^{\mathrm{pole}}(7-8\ \mathrm{TeV},\ \mathrm{MSTW}) = 172.53^{+2.36}_{-2.44}\ \mathrm{GeV}$$

(6.31)

$$m_{\mathrm{t}}^{\mathrm{pole}}(7-8\ \mathrm{TeV},\ \mathrm{CT10}) = 172.88^{+2.50}_{-2.64}\ \mathrm{GeV}$$

(6.32)

$$m_{\mathrm{t}}^{\mathrm{pole}}(7-8\ \mathrm{TeV},\ \mathrm{NNPDF}) = 172.86^{+2.11}_{-2.38}\ \mathrm{GeV}$$

(6.33)

$$m_{\mathrm{t}}^{\mathrm{pole}}(7-8\ \mathrm{TeV},\ \mathrm{PDF4LHC}) = 172.87^{+2.51}_{-2.63}\ \mathrm{GeV}$$

(6.34)

As expected, the combined result lies between the 7 and 8 TeV results. The uncertainty does not decrease substantially in the combined result, since much of the error is correlated.

Figure 6.7 shows the results at 7 and 8 TeV. The following section includes an examination of the consistency between the separate 7 and 8 TeV results.
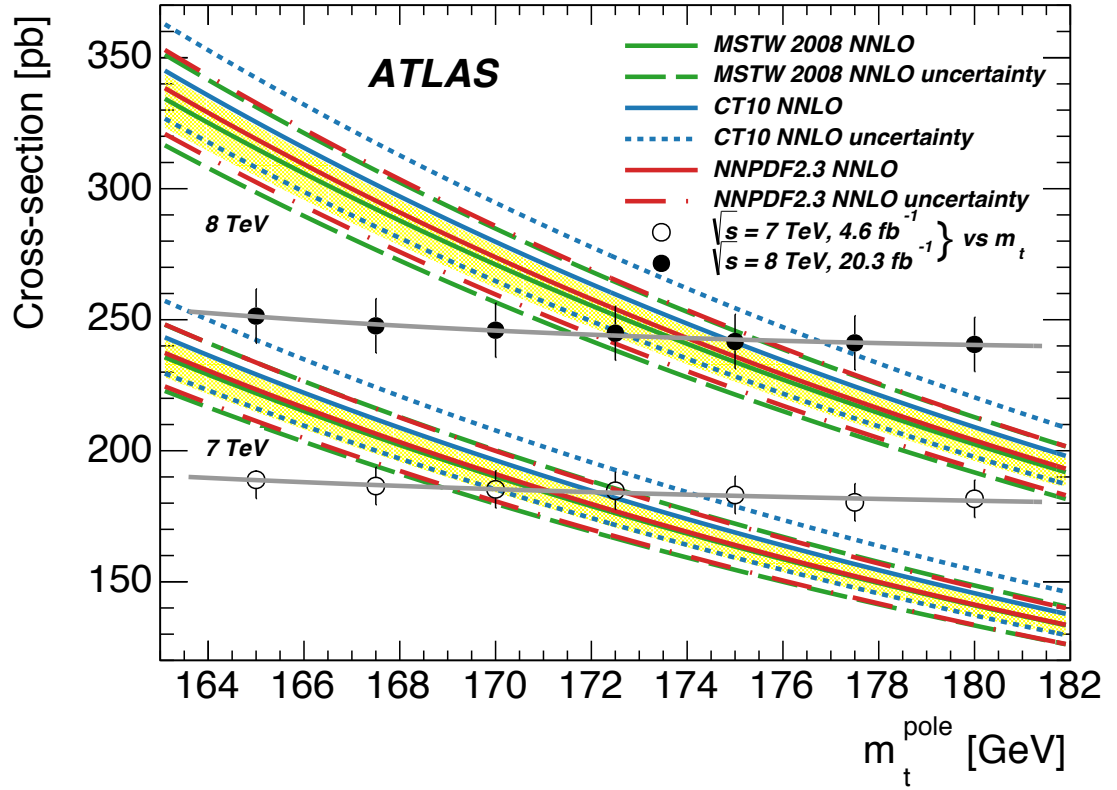
Figure 6.7: Predicted NNLO+NNLL cross sections compared with the measured cross sections from Figure 6.2, as functions of the top quark pole mass m_t^{pole} for 7 TeV and 8 TeV. The three NNLO PDF sets are indicated by the three differently colored lines, the dashed lines represent the associated uncertainties at 68% confidence level (CL), and the shaded band indicates the PDF uncertainty range. The plots also show the experimental data points and accompanying polynomial fit from Equation 6.11; the experimental error bars include all sources of uncertainty given in Ref. [95] as well as uncorrelated statistical uncertainty.

6.4 Cross-Checks

This section includes several cross-checks of the results, including an examination of uncertainties and consistency between the 7 and 8 TeV results, a comparison with a frequentist approach for the mass extraction, comparison with previous ATLAS data collected at $\sqrt{s} = 7$ TeV, and comparison with CMS data collected at 7 TeV.

6.4.1 Examination of Uncertainties

To compare the role of different sources of uncertainty, the extraction was redone considering only the experimental uncertainty and then only the theoretical uncertainty, both shown in Table 6.6. As expected, the PDF and scale uncertainties in the predicted cross sections are larger contributors to the total uncertainties.

Table 6.6 also shows the results of an examination of consistency between the 7 and 8 TeV results, in which the extraction was redone including only the uncorrelated experimental uncertainties. The central value changes by up to 0.16 GeV in this case, since the theoretical uncertainty is a function of the mass $m_\mathrm{t}^\mathrm{pole}$. The inclusion of the theoretical uncertainty affects the central value.

Considering only uncorrelated sources of uncertainty, the 7 and 8 TeV results agree

within $\sim 1.6\sigma$:

$$\Delta(\text{7-8 TeV MSTW}) = (173.90 - 171.15) \pm \sqrt{1.19^2 + 1.00^2} = 2.75 \pm 1.55 \text{ GeV}$$

$$(6.35)$$

$$\Delta(\text{7-8 TeV CT10}) = (173.98 - 171.24) \pm \sqrt{1.19^2 + 1.00^2} = 2.74 \pm 1.55 \text{ GeV}$$

$$(6.36)$$

$$\Delta(\text{7-8 TeV NNPDF}) = (174.24 - 171.31) \pm \sqrt{1.18^2 + 0.99^2} = 2.93 \pm 1.54 \text{ GeV}$$

$$(6.37)$$

	$m_{\text{t}}^{\text{pole}}$ (MSTW) GeV	$m_{\text{t}}^{\text{pole}}$ (CT10) GeV	$m_{\text{t}}^{\text{pole}}$ (NNPDF) GeV
7TeV: only σ_{th}	$171.26^{+1.88}_{-2.01}$	$171.38^{+2.14}_{-2.15}$	$171.40^{+1.64}_{-1.95}$
7TeV: only σ_{expt}	$171.15^{+1.44}_{-1.36}$	$171.24^{+1.44}_{-1.37}$	$171.31^{+1.43}_{-1.35}$
7TeV: only $\sigma_{\text{uncor expt}}$	$171.15^{+1.00}_{-0.96}$	$171.24^{+1.00}_{-0.96}$	$171.31^{+0.99}_{-0.95}$
8TeV: only σ_{th}	$174.00^{+1.86}_{-2.03}$	$174.10^{+2.05}_{-2.13}$	$174.32^{+1.60}_{-1.94}$
8TeV: only σ_{expt}	$173.90^{+1.66}_{-1.57}$	$173.98^{+1.66}_{-1.57}$	$174.24^{+1.66}_{-1.55}$
8TeV: only $\sigma_{\text{uncor expt}}$	$173.90^{+1.24}_{-1.19}$	$173.98^{+1.24}_{-1.19}$	$174.24^{+1.23}_{-1.18}$

Table 6.6: The top quark pole mass $m_{\text{t}}^{\text{pole}}$ extracted at 7 and 8 TeV, considering only theoretical uncertainty σ_{th}, only experimental uncertainty σ_{expt}, and only uncorrelated experimental uncertainty $\sigma_{\text{uncor expt}}$.

6.4.2 Frequentist Likelihood Formulation

The mass extraction was also done with the following frequentist approach:

$$L(m_t) \propto G\left(\sigma_{t\bar{t}}^{\text{th}}(m_t) \,|\, \sigma_{t\bar{t}}^{\text{exp}}(m_t), \rho_{\text{tot}}\right) \tag{6.38}$$

In this method, the probability density function is built as an asymmetric Gaussian combination of the theoretical and experimental cross sections and uncertainties, according to the following form:

$$G\left(\sigma_{t\bar{t}}^{\text{th}} \,|\, \sigma_{t\bar{t}}^{\text{exp}}(m_t), \rho_{\text{tot}}\right) = \frac{2}{\sqrt{2\pi}(\rho_{\text{tot}}^{\text{upper}} + \rho_{\text{tot}}^{\text{lower}})} \exp\left(-\frac{1}{2}\left(\frac{\sigma_{t\bar{t}} - \sigma_{t\bar{t}}^{\text{th}}}{\rho_{\text{tot}}}\right)^2\right) \tag{6.39}$$

where $\rho_{\text{tot}} = \rho_{\text{tot}}^{\text{upper}}$ for $\sigma_{t\bar{t}}^{\text{th}}(m_t) < \sigma_{t\bar{t}}^{\text{exp}}(m_t)$, and $\rho_{\text{tot}} = \rho_{\text{tot}}^{\text{lower}}$ for $\sigma_{t\bar{t}}^{\text{th}}(m_t) > \sigma_{t\bar{t}}^{\text{exp}}(m_t)$. The theoretical uncertainties are held constant such that $\rho_{\text{th}}(m_t) = \rho_{\text{th}}(m_t = 172.5 \text{ GeV})$. The total uncertainties used are the quadratic sum of theoretical and experimental uncertainties:

$$\rho_{\text{tot}}^{\text{upper}} = \sqrt{(\rho_{\text{exp}})^2 + (\rho_{\text{th}}^{\text{upper}})^2} \tag{6.40}$$

and likewise for the lower theoretical uncertainty. This was done for each of the three theoretical PDF prescriptions, as well as the PDF4LHC method. Using the PDF4LHC method, the results are:

$$m_t^{\text{pole}}(7 \text{ TeV, Frequentist}) = 171.24_{-2.41}^{+2.67} \text{ GeV} \tag{6.41}$$

$$m_t^{\text{pole}}(8 \text{ TeV, Frequentist}) = 173.98_{-2.59}^{+2.85} \text{ GeV} \tag{6.42}$$

The results using each of the three PDF predictions and the PDF4LHC method all agree with the asymmetric Bayesian method within 0.16 GeV.

6.4.3 Comparison with ATLAS 2011 Results

The mass extraction was also done using the same theoretical predictions but with previously published ATLAS data at $\sqrt{s} = 7$ TeV [113]. The previous 7 TeV result measured $186.3^{+24.0}_{-22.6}$ pb cross section at $m_t = 172.5$ GeV, with eight additional data points and an estimated NNLO+NNLL solution to extract a top pole mass of $166.4^{+7.8}_{-7.3}$ GeV, as shown in Figure 6.8. The full NNLO+NNLL prediction with these data points and polynomial fit results in an extracted mass $m_{\mathrm{t}}^{\mathrm{pole}}$ several GeV larger than the previous result, with smaller uncertainties. This difference is due primarily to the previous low theoretical estimate of the cross section $\sigma_{t\bar{t}}$, the previous overestimate of the correlating theoretical uncertainties (the scale and PDF uncertainties were added linearly rather than in quadrature), and higher experimental uncertainty:

$$m_{\mathrm{top}}^{\mathrm{pole}}(\text{ATLAS 2011 Comparison}) = 168.53^{+3.25}_{-3.33} \text{ GeV (MSTW)} \tag{6.43}$$

$$168.75^{+3.44}_{-3.51} \text{ GeV (CT10)} \tag{6.44}$$

$$168.58^{+3.09}_{-3.18} \text{ GeV (NNPDF)} \tag{6.45}$$

The experimental result in this case also has a higher dependence on the assumed m_t. By using the same method described in Section 6.1.2, the mass dependence is more than twice the magnitude of the mass dependence of the analysis used in this note:

$$\left. \frac{\mathrm{d}\sigma_{t\bar{t}}(m_t)}{\mathrm{d}m_t} \right|_{m_t = m_t^{\mathrm{ref}}} = \frac{\sigma_{t\bar{t}}(m_t^{\mathrm{ref}})}{m_t^{\mathrm{ref}}} (a_1 - 4) = -1.51 \frac{\mathrm{pb}}{\mathrm{GeV}} = -0.78 \frac{\%}{\mathrm{GeV}} \tag{6.46}$$
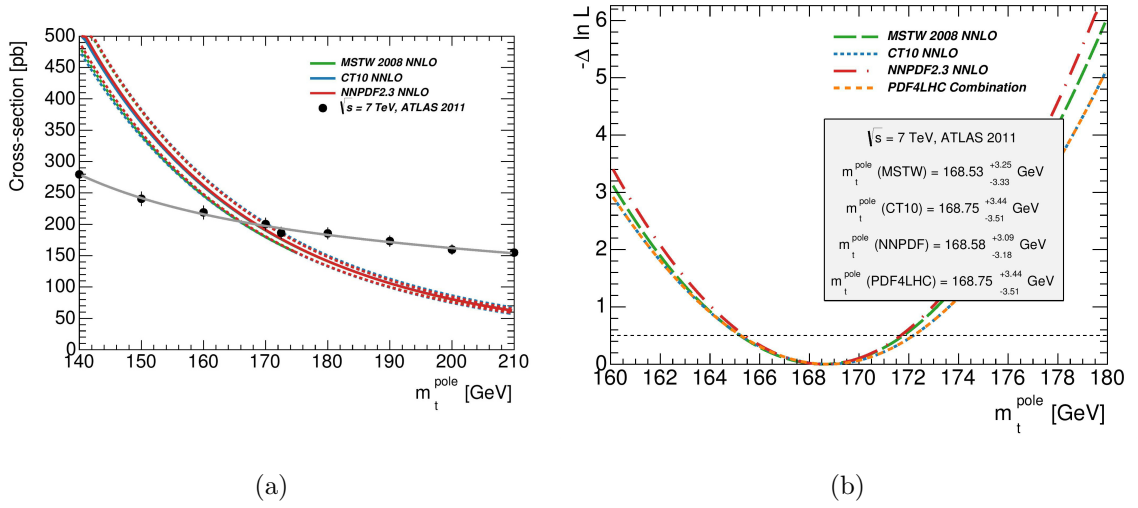
Figure 6.8: A comparison a) of the three predicted NNLO+NNLL solutions with previous ATLAS data at 7 TeV, and b) the accompanying mass extraction. The correspondingly colored dashed lines represent the spread of uncertainties.

6.4.4   Comparison with 2013 CMS Results

Finally, this analysis was done using recently published CMS data [30] with the three predicted NNLO+NNLL solutions, shown in Figures 6.9a and 6.9b. The CMS analysis used the NNPDF2.3 NNLO+NNLL solution (NNPDF2.3, the third NNLO PDF set used by this analysis), and extracted a top pole mass of $176.7^{+3.8}_{-3.4}$ GeV. Extracting the top pole mass using the CMS data, the experimental polynomial fits given in this analysis, and the NNLO+NNLL theoretically predicted cross sections used in this
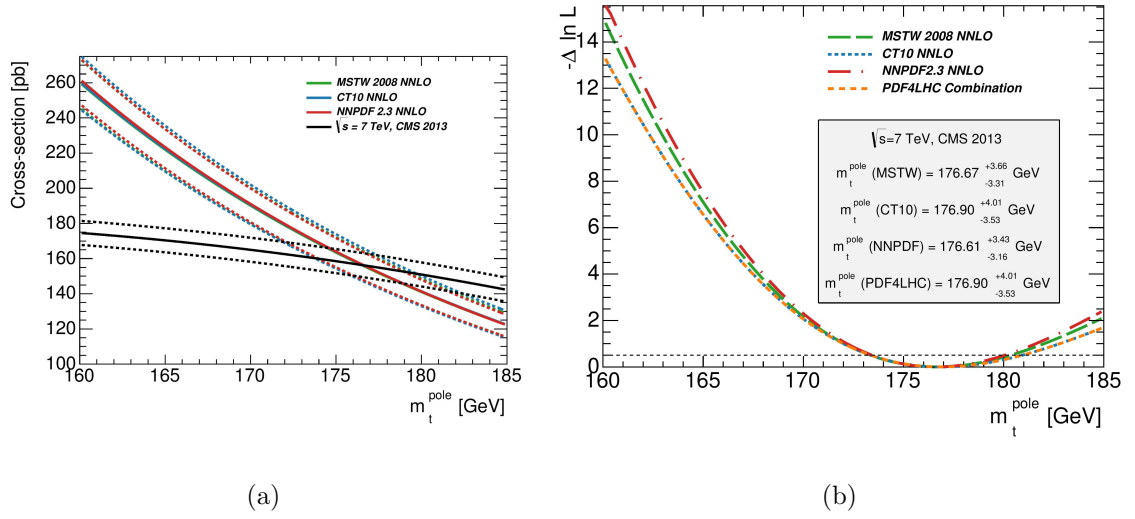
Figure 6.9: A comparison a) of the three predicted NNLO+NNLL solutions with previous CMS data [30] at 7 TeV, and b) the accompanying mass extraction. In the comparison a), the black line represents a fit to the measured cross section, and the black dashed lines the correlating uncertainties. The correspondingly colored dashed lines represent the spread of uncertainties in a) the predicted cross sections and b) maximized likelihood.

analysis gives consistent results:

$$m_{\text{top}}^{\text{pole}}(\text{CMS 2013 Comparison}) = 176.67^{+3.66}_{-3.31} \text{ GeV (MSTW)} \quad (6.47)$$

$$176.90^{+4.01}_{-3.53} \text{ GeV (CT10)} \quad (6.48)$$

$$176.61^{+3.43}_{-3.16} \text{ GeV (NNPDF)} \quad (6.49)$$

6.5 Conclusion

This chapter presents the extracted top quark pole mass m_t^{pole} using experimental results with minimal m_t dependence and theoretical curves that include QCD corrections at the NNLO order for the first time, as well as soft-gluon resummation at NNLL accuracy. These results complement direct measurements of the top quark mass by using a specific renormalization scheme and decreasing mass dependence in the experimental result. Following the PDF4LHC combination recommendation, the result is:

$$m_{\text{top}}^{\text{pole}}(7 \text{ TeV}) = 171.40^{+2.58}_{-2.56} \text{ GeV} \tag{6.50}$$

$$m_{\text{top}}^{\text{pole}}(8 \text{ TeV}) = 174.10 \pm 2.64 \text{ GeV} \tag{6.51}$$

$$m_{\text{t}}^{\text{pole}}(7\text{-}8 \text{ TeV}) = 172.87^{+2.51}_{-2.63} \text{ GeV} \tag{6.52}$$

As explained in Section 6.2, the method considers the top quark mass input in the Monte Carlo simulations to be equivalent to the top quark pole mass, although the Monte Carlo simulations do not use a specified renormalization scheme. These results are consistent with the latest world average [111] and are more precise than previous extractions of the top quark pole mass, as shown in Figure 6.10. They are also consistent with theoretical projections using electroweak fitting and including the impact of the Standard Model Higgs boson [114, 115, 116].

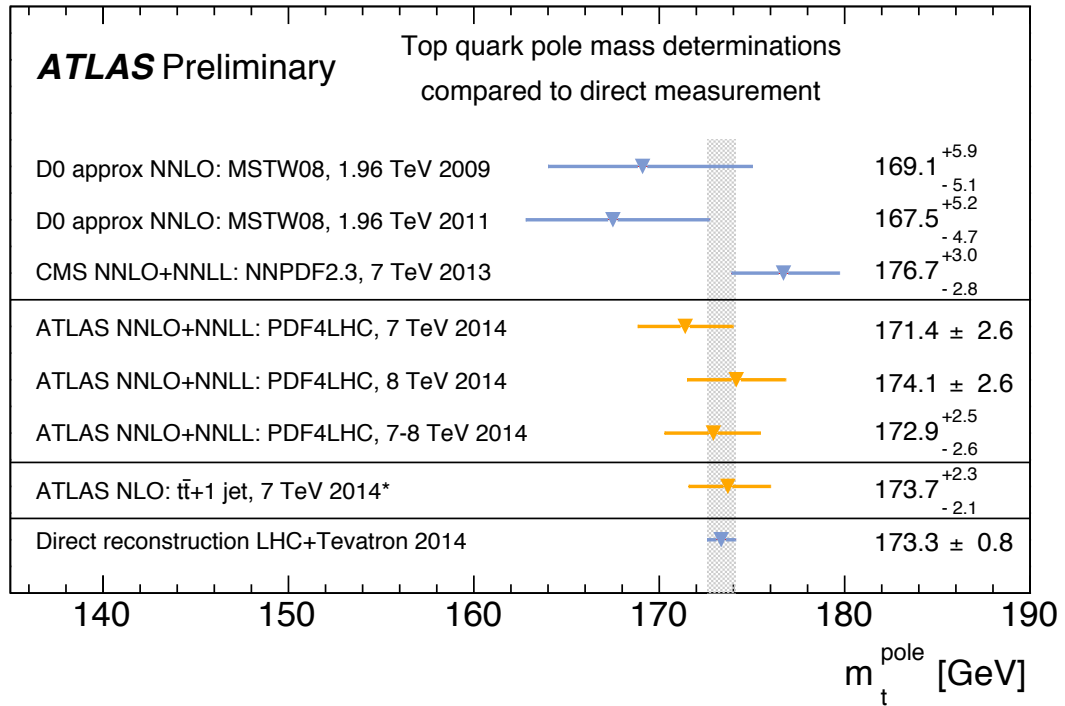Figure 6.10: The results of this extraction, compared with previous top quark pole mass extractions from ATLAS [113], CMS [30], D0 [117, 118], as well as a preliminary result from ATLAS [119] and the Particle Data Group (PDG) 2013 average of measurements using direct reconstruction of $t\bar{t}$ events [111]. The 2014 LHC+Tevatron combined result from direct reconstruction [21] agrees closely with the PDG measurement.

Chapter 7

Epilogue

"As we conquer peak after peak we see in front of us regions full of interest and beauty, but we do not see our goal, we do not see the horizon; in the distance tower still higher peaks, which will yield to those who ascend them still wider prospects, and deepen the feeling, the truth of which is emphasized by every advance in science, that 'Great are the Works of the Lord'.
— J. J. Thomson, Nobel Prize winner, discoverer of the electron

Two precision measurements of the top quark in pair production $t\bar{t}$ events using the ATLAS Detector at the LHC have been presented in this work. Both measurements leverage the clean signal of selecting events with an opposite-sign pair of an electron and muon, and b-tagged jets, to achieve low uncertainties in the result. Both measurements provide additional understanding of the top quark and its properties in the environment of a proton-proton collider, and by extension of interactions with the Higgs boson.

The first analysis measured the amount of jet activity, and thus QCD radiation, in $t\bar{t}$ pair-production events from the 2012 pp collision data sample collected by the ATLAS Detector. The gap fraction method results in plots of the fraction of events without qualifying QCD radiation as a function of transverse momentum p_{T} of the additional jet activity. These plots were corrected for detector effects to enable

detector-independent comparison to simulation, and subsequently compared to the simulations by state-of-the-art Monte Carlo event generators. The small size of the systematic and statistical uncertainties in the measurement enables it to be used to constrain uncertainty in quark and gluon radiation modeling in other ATLAS measurements, as well as to aid in future tuning of QCD radiation simulation in Monte Carlo generators.

The second analysis presented a measurement of the top quark pole mass $m_{\mathrm{t}}^{\mathrm{pole}}$ using a combination of experimental measurement and theoretical prediction of $t\bar{t}$ pair-production events. The clean signal and low dependence of the experimental result on the top quark mass led to the most precise pole mass measurement yet at the time of publication, and has since been followed by a very similar measurement by the CMS Collaboration [120].

The first complete physics run of the LHC in 2011-2012 at center-of-mass energies of $\sqrt{s} = 7 - 8$ TeV has provided an unprecedented opportunity to study the top quark in a wealth of top quark pair-production events. The second physics run scheduled for 2015-2018 at $\sqrt{s} = 13 - 14$ TeV is expected to increase the concentration of top quark pair-production events by nearly fourfold [23]. Leveraging precision results such as these from the first run is valuable for maximizing future measurements at the LHC to test the Standard Model, measure properties of the recently discovered Higgs boson, and look for new physics beyond the Standard Model.

References

[1] S. Boutas, *CERN Turns 60, Celebrates Peaceful Collaboration for Science*, CERN, 2014, `http://home.web.cern.ch/about/updates/2014/09/cern-turns-60-celebrates-peaceful-collaboration-science`. (document)

[2] UA1 Collaboration, G. Arnison et al., *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV*, Phys. Lett. **B122** (1983) 103–116. (document)

[3] UA2 Collaboration, M. Banner et al., *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN $\bar{p}p$ Collider*, Phys. Lett. **B122** (1983) 476–485. (document)

[4] UA2 Collaboration, P. Bagnaia et al., *Evidence for $Z_0 \to e^+e^-$ at the CERN $\bar{p}p$ Collider*, Phys. Lett. **B129** (1983) 130–140. (document)

[5] CDF Collaboration, *Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74** (1995) 2626–2631. (document), 1.4

[6] D0 Collaboration, S. Abachi et al., *Observation of the Top Quark*, Phys. Rev. Lett. **74** (1995) 2632–2637, `arXiv:hep-ex/9503003 [hep-ex]`. (document), 1.4

[7] S. W. Herb et al., *Observation of a Dimuon Resonance at 9.5 GeV in 400 GeV Proton-Nucleus Collisions*, Phys. Rev. Lett. **39** (1977) 252–255. (document), 1.4

[8] CMS Collaboration, S. Chatrchyan et al., *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. **B716** (2012) 30–61, `arXiv:1207.7235 [hep-ex]`. (document), 1.1.2

[9] ATLAS Collaboration, *Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC*, Phys. Lett. **B716** (2013) 1–29, `arXiv:1207.7214 [hep-ex]`. (document), 1.1.2

[10] D. Griffiths, *Introduction to Elementary Particles*,. 1

[11] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory*,. 1, 1.1.1, 1.1.1

[12] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579–588. 1.1.1

[13] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266. 1.1.1

[14] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367–377. 1.1.1

[15] M. Y. Han and Y. Nambu, *Three-Triplet Model with Double* SU(3) *Symmetry*, Phys. Rev. **139** (1965) B1006–B1010. 1.1.1

[16] O. W. Greenberg, *Spin and Unitary-Spin Independence in a Paraquark Model of Baryons and Mesons*, Phys. Rev. Lett. **13** (1964) 598–602. 1.1.1

[17] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652–657. 1.1.2

[18] ATLAS and CMS Collaborations, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, Phys. Rev. Lett. **114** (2015) 191803, `arXiv:1503.07589 [hep-ex]`. 1.1.2

[19] Georgia State University, *HyperPhysics*, `http://hyperphysics.phy-astr.gsu.edu/hbase/particles/neutrino3.html`. 1.1

[20] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, et al., *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, JHEP **0207** (2002) 012, `arXiv:hep-ph/0201195 [hep-ph]`. 1.2, 1.2, 3.2

[21] ATLAS, CDF, CMS, DØ Collaborations, *First Combination of Tevatron and LHC Measurements of the Top-Quark Mass*, `arXiv:1403.4427 [hep-ex]`. 1.4, 6.10

[22] DØ Collaboration, *An Improved Determination of the Width of the Top Quark*, Phys. Rev. **D85** (2012) 091104, `arXiv:1201.4156 [hep-ex]`. 1.4

[23] M. Czakon, P. Fiedler, and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through* $O(\alpha_S^4)$, Phys. Rev. Lett. **110** (2013) 252004, `arXiv:1303.6254 [hep-ph]`. 1.4.1, 6.2, 7

[24] S. P. Martin, *A Supersymmetry Primer*, `arXiv:hep-ph/9709356 [hep-ph]`, [Adv. Ser. Direct. High Energy Phys.18,1(1998)]. 1.5

[25] ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report 1*,. 2

[26] O. S. Bruning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock. 2, 2.1, 2.1.1, 5

[27] S. Myers, *The LEP Collider from Design to Approval and Commissioning*,. 2.1

[28] CERN, *LHC Machine Outreach*, CERN, 2015,
http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach. 2.1,
2.2

[29] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron
Collider*, Journal of Instrumentation**3** no. 08, S08003. 2.1, 2.3, 2.2.3, 2.5, 2.2.3,
2.6, 2.7

[30] CMS Collaboration, *The CMS Experiment at the CERN LHC*, Journal of
Instrumentation **3** no. 08, (2008) S08004. 2.1, 6.4.4, 6.9, 6.10

[31] ALICE Collaboration, *The ALICE Experiment at the CERN LHC*, Journal of
Instrumentation **3** no. 08, (2008) S08002. 2.1

[32] LHCb Collaboration, *The LHCb Detector at the CERN LHC*, Journal of
Instrumentation **3** no. 08, (2008) S08005. 2.1

[33] LHCf Collaboration, *The LHCf Detector at the CERN Large Hadron Collider*,
Journal of Instrumentation **3** no. 08, (2008) S08006. 2.1

[34] TOTEM Collaboration, *The TOTEM Experiment at the CERN Large Hadron
Collider*, Journal of Instrumentation **3** no. 08, (2008) S08007. 2.1

[35] MoEDAL Collaboration, *The Physics Programme of the MoEDAL Experiment
at the LHC*, Int. J. Mod. Phys. **A29** (2014) 1430050, arXiv:1405.7662
[hep-ph]. 2.1

[36] M. Schott and M. Dunford, *Review of Single Vector Boson Production in pp
Collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. **C74** (2014) 2916, arXiv:1405.1160
[hep-ex]. 2.4

[37] ATLAS Collaboration, *ATLAS Pixel Detector Electronics and Sensors*, Journal
of Instrumentation**3** no. 07, P07007. 2.2.3, A.2, A.2.1, A.3, A.4

[38] ATLAS Collaboration, *Data Quality Information for 2011 Data*,
https://atlas.web.cern.ch/Atlas/GROUPS/DATAPREPARATION/
PublicPlots/DQ/atlas-dq-eff-pp2011-AtoM-.png.2011. 2.2.3

[39] ATLAS Collaboration, *RMS of residuals as a function of track incident angle
(Collision data)*, https://twiki.cern.ch/twiki/pub/AtlasPublic/
PixelPublicResults/res_RMS_phi_repro_MC.eps.Mar.2011. 2.2.3

[40] ATLAS Collaboration, *Measurement of the $t\bar{t}$ Production Cross-Section using
$e\mu$ Events with b-Tagged Jets in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the
ATLAS Detector*, Eur. Phys. J. C **74** (2014) 3109, arXiv:1406.5375
[hep-ex]. 3, 3.2, 4, 4.3, 4.3, 6, 6.1.1, 6.1.2, 6.1.2

[41] *ATLAS Top Data 2012 NTupCommon DPDs*, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopData12NtupCommon. 3, 3.1

[42] *ATLAS Top MC 2012 DiTop Samples*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopMC12DiTopSamples. 3

[43] *ATLAS AnalysisTop Framework*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/AnalysisTop. 3

[44] *ATLAS Good Run List*, data12_8TeV.periodAllYear_DetStatus-v61-pro14-02_DQDefects-00-01-00_PHYS_StandardGRL_All_Good.xml. 3.1

[45] *ATLAS Good Run List Recommendations*, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopGRLs. 3.1

[46] ATLAS Collaboration, *Improved Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector at the LHC*, Eur. Phys. J. **C73** no. 8, (2013) 2518, arXiv:1302.4393 [hep-ex]. 3.1

[47] ATLAS Collaboration, *Improved Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector at the LHC*, Eur.Phys.J. **C73** no. 8, (2013) 2518, arXiv:1302.4393 [hep-ex]. 3.1

[48] T. Luminosity Group, *Preliminary Luminosity Determination in pp Collisions at $\sqrt{s} = 8$ TeV using the ATLAS Detector in 2012*, Tech. Rep. ATL-COM-LUM-2012-013, CERN, Geneva, Nov, 2012. https://cds.cern.ch/record/1494059. 3.1

[49] *Luminosity for Physics, Revision 11*, https://twiki.cern.ch/twiki/bin/view/Atlas/LuminosityForPhysics?rev=11. 3.1

[50] S. Agostinelli et al., *Geant4— a Simulation Toolkit,*. 3.2

[51] C. ATLAS, M. Beckingham, M. Duehrssen, E. Schmidt, M. Shapiro, M. Venturi, J. Virzi, I. Vivarelli, M. Werner, S. Yamamoto, and T. Yamanaka, *The Simulation Principle and Performance of the ATLAS Fast Calorimeter Simulation FastCaloSim*, Tech. Rep. ATL-PHYS-PUB-2010-013, CERN, Geneva, Oct, 2010. https://cds.cern.ch/record/1300517. 3.2

[52] ATLAS Collaboration, *Measurement of the b-Tag Efficiency in a Sample of Jets Containing Muons with 5 fb^{-1} of Data from the ATLAS Detector*, ATLAS-CONF-2012-043, 2012, http://cdsweb.cern.ch/record/1435197. 3.2, 4.1

[53] ATLAS Collaboration, *b-Jet Tagging Efficiency Calibration using the System8 Method*, ATLAS-CONF-2011-143, 2011,
http://cdsweb.cern.ch/record/1386703. 3.2

[54] Hawkings, R., *Measurement of the b-Tagging Efficiency of the MV1 Algorithm in pp Collisions at $\sqrt{s} = 8$ TeV using $e\mu$ Dilepton $t\bar{t}$ Events*,
ATL-COM-PHYS-2013-381, 2013, https://cds.cern.ch/record/1537694.
3.2

[55] *ATLAS Top D3PD Production*, https://twiki.cern.ch/twiki/bin/
viewauth/AtlasProtected/TopD3PDProduction. 3.2

[56] *Fiducial Top-Quark Measurements*,
https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopFiducial.
3.2, 4.2

[57] *Top Root Core Releases*, https:
//twiki.cern.ch/twiki/bin/view/AtlasProtected/TopRootCoreRelease.
3.2

[58] P. Nason, *A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms*, JHEP **0411** (2004) 040, arXiv:hep-ph/0409146 [hep-ph]. 3.2

[59] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD Computations with Parton Shower Simulations: the POWHEG Method*, JHEP **0711** (2007) 070,
arXiv:0709.2092 [hep-ph]. 3.2

[60] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*,
JHEP **0605** (2006) 026, arXiv:hep-ph/0603175 [hep-ph]. 3.2

[61] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, Phys.Rev.
D82 (2010) 074018, arXiv:1005.3457 [hep-ph]. 3.2

[62] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, et al., *New Parton Distributions for Collider Physics*, Phys.Rev. **D82** (2010) 074024,
arXiv:1007.2241 [hep-ph]. 3.2, 5.5.10, 6.2, 6.4, 6.5

[63] S. Frixione and B. R. Webber, *Matching NLO QCD Computations and Parton Shower Simulations*, JHEP **0206** (2002) 029, arXiv:hep-ph/0204244
[hep-ph]. 3.2

[64] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al., *HERWIG 6: An Event Generator for Hadron Emission Reactions with Interfering Gluons (Including Supersymmetric Processes)*, JHEP **0101** (2001) 010,
arXiv:hep-ph/0011363 [hep-ph]. 3.2

[65] J. Butterworth, J. R. Forshaw, and M. Seymour, *Multiparton Interactions in Photoproduction at HERA*, Z.Phys. **C72** (1996) 637–646, `arXiv:hep-ph/9601371 [hep-ph]`. 3.2

[66] ATLAS Collaboration, *New ATLAS Event Generator Tunes to 2010 Data*, ATL-PHYS-PUB-2011-008, 2011, `http://cds.cern.ch/record/1345343`. 3.2

[67] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a Generator for Hard Multiparton Processes in Hadronic Collisions*, JHEP **0307** (2003) 001, `arXiv:hep-ph/0206293 [hep-ph]`. 3.2

[68] E. Kersevan, B. P.; Richter-Was, *The Monte Carlo Event Generator AcerMC Version 2.0 with Interfaces to PYTHIA 6.2 and HERWIG 6.5*, Eur. Phys. J. C **55** (2004) 553, `0405247 [hep-ph]`. 3.2

[69] ATLAS Collaboration, *Measurement of $t\bar{t}$ Production with a Veto on Additional Central Jet Activity in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector*, Eur. Phys. J. **C72** (2012) 2043, `arXiv:1203.5015 [hep-ex]`. 3.2, 5, 5.5.11, 5.1, 5.6

[70] S. Frixione, E. Laenen, P. Motylinski, B. R. Webber, and C. D. White, *Single-Top Hadroproduction in Association with a W Boson*, JHEP **07** (2008) 029, `arXiv:0805.3067 [hep-ph]`. 3.2

[71] E. Re, *Single-Top Wt-Channel Production Matched with Parton Showers using the POWHEG Method*, Eur.Phys.J. **C71** (2011) 1547, `arXiv:1009.2450 [hep-ph]`. 3.2, 5.5.7

[72] *Heavy Flavour Overlap Removal Procedure*, `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/HforTool`. 3.2

[73] B. Acharya et al., *Object Selection and Calibration, Background Estimations and MC Samples for Top Quark Analyses using the full 2012 Data Set*, ATLAS-COM-PHYS-2013-1016, 2014, `https://cds.cern.ch/record/1563201`. 4, 4.1

[74] B. Acharya et al., *Object Selection and Calibration, Background Estimations and MC Samples for the Autumn 2012 Top Quark Analyses with 2012 Data*, ATLAS-COM-PHYS-2013-088, 2014, `https://cds.cern.ch/record/1509562`. 4, 4.1

[75] ATLAS Collaboration, *Electron Reconstruction and Identification Efficiency Measurements with the ATLAS Detector using the 2011 LHC Proton–Proton Collision Data*, Eur. Phys. J. C **74** (2014) 2941, `arXiv:1404.2240 [hep-ex]`. 4

[76] ATLAS Collaboration, *Measurement of the Muon Reconstruction Performance of the ATLAS Detector using 2011 and 2012 LHC Proton–Proton Collision Data*, Eur. Phys. J. C **74** (2014) 3130, `arXiv:1407.3935 [hep-ex]`. 4

213

[77] *Electron Isolation Tool from ATLAS `egammaAnalysisUtils` Package,.* 4.1

[78] G. P. Cacciari, M.; Salam, *Dispelling the N^3 Myth for the k_t Jet Finder*, Phys. Lett. B **B641** (2006) 57–61, arXiv:0512210 [hep-ph]. 4.1

[79] G. P. Cacciari, M.; Salam, *The Anti-k_t Jet Clustering Algorithm*, JHEP **0804** (2008) 063, arXiv:0802.1189 [hep-ph]. 4.1

[80] *Pile-up Subtraction and Suppression for Jets in ATLAS*, Tech. Rep. ATLAS-CONF-2013-083, CERN, Geneva, Aug, 2013. https://cds.cern.ch/record/1570994. 4.1

[81] ATLAS Collaboration, *Commissioning of the ATLAS High-Performance b-Tagging Algorithms in the 7 TeV Collision Data*, ATL-PHYS-INT-2013-023, 2013, https://cds.cern.ch/record/1627637. 4.1

[82] M. Cacciari, G. P. Salam, and G. Soyez, *The Catchment Area of Jets*, JHEP **0804** (2008) 005, arXiv:0802.1188 [hep-ph]. 4.2

[83] *ATLAS Top Common Recommended Object Selection*, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopCommonObjects. 4.3

[84] ATLAS Collaboration, *Comparison of Monte Carlo Generator Predictions for Gap Fraction and Jet Multiplicity Observables in $t\bar{t}$ Events*, ATL-PHYS-PUB-2014-005, 2014, http://cdsweb.cern.ch/record/1703034. 5

[85] *ATLAS Jet Energy Scale Recommendations*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetUncertainties2012Final. 5.5.1

[86] Brosamer, J; Shapiro, M; et al., *Measurement of Jets Produced in Top Quark Events using eμ Final State with 2 b-Tagged Jets in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector*, ATL-COM-PHYS-2015-041, 2015, https://cds.cern.ch/record/1983211. 5.5.2, 5.38

[87] *ATLAS Systematic Uncertainties Recommendations*, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopSystematicUncertainties. 5.5.3

[88] C. D. White, S. Frixione, E. Laenen, and F. Maltoni, *Isolating Wt Production at the LHC*, JHEP **0911** (2009) 074, arXiv:0908.0631 [hep-ph]. 5.5.7

[89] *ATLAS Top Group PDF Uncertainty Recommendations*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopPdfUncertainty. 5.5.10

[90] M. Botje et al., *The PDF4LHC Working Group Interim Recommendations*, arXiv:1101.0538 [hep-ph]. 5.5.10

[91] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton Distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph]. 5.5.10

[92] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Uncertainties on α_S in Global PDF Analyses and Implications for Predicted Hadronic Cross Sections*, Eur. Phys. J. **C64** (2009) 653–680, arXiv:0905.3531 [hep-ph]. 5.5.10

[93] R. D. Ball et al., *Parton Distributions with LHC Data*, Nucl. Phys. **B867** (2013) 244–289, arXiv:1207.1303 [hep-ph]. 5.5.10

[94] *HEPData*, http://hepdata.cedar.ac.uk/. 5.6

[95] R. Hawkings, *Measurement of the $t\bar{t}$ Production Cross-section in pp Collisions at $\sqrt{s} = 7 - 8$ TeV using $e\mu$ Events with b-Tagged Jets*, ATL-COM-PHYS-2014-017, 2014, http://cdsweb.cern.ch/record/1642806. 6, 6.1.1, 6.3, 6.4, 6.7

[96] CERN, *Precise Measurements of Top-Quark Production*, CERN Courier V54 (2014) 7, 2014, http://cerncourier.com/cws/article/cern/58191. 6

[97] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10** (1975) 343–367. 6.1.2, 6.3.3

[98] R. J. Barlow, *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences*,. 6.1.3

[99] M. Cacciari, M. Czakon, M. Mangano, A. Mitov, and P. Nason, *Top-Pair Production at Hadron Colliders with Next-to-Next-to-Leading Logarithmic Soft-Gluon Resummation*, Phys. Lett. **B710** (2012) 612–622, arXiv:1111.5869 [hep-ph]. 6.2

[100] P. Bärnreuther, M. Czakon, and A. Mitov, *Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t} + X$*, Phys. Rev. Lett. **109** (2012) 132001, arXiv:1204.5201 [hep-ph]. 6.2

[101] M. Czakon and A. Mitov, *NNLO Corrections to Top-Pair Production at Hadron Colliders: the All-Fermionic Scattering Channels*, JHEP **12** (2012) 054, arXiv:1207.0236 [hep-ph]. 6.2

[102] M. Czakon and A. Mitov, *NNLO Corrections to Top Pair Production at Hadron Colliders: the Quark-Gluon Reaction*, JHEP **01** (2013) 080, arXiv:1210.6832 [hep-ph]. 6.2

[103] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, Comput. Phys. Commun. **185** (2014) 2930, arXiv:1112.5675 [hep-ph]. 6.2

[104] M. Botje et al., *The PDF4LHC Working Group Interim Recommendations*, arXiv:1101.0538 [hep-ph]. 6.2, 6.3.2

[105] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton Distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph]. 6.2, 6.4, 6.5

[106] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Uncertainties on α_S in Global PDF Analyses and Implications for Predicted Hadronic Cross Sections*, Eur. Phys. J. **C64** (2009) 653–680, arXiv:0905.3531 [hep-ph]. 6.2, 6.4, 6.5

[107] J. Gao, M. Guzzi, J. Huston, H.-L. Lai, Z. Li, P. Nadolsky, J. Pumplin, D. Stump, and C. P. Yuan, *CT10 Next-to-Next-to-Leading Order Global Analysis of QCD*, Phys. Rev. **D89** no. 3, (2014) 033009, arXiv:1302.6246 [hep-ph]. 6.2, 6.4, 6.5

[108] R. D. Ball et al., *Parton Distributions with LHC Data*, Nucl. Phys. **B867** (2013) 244–289, arXiv:1207.1303 [hep-ph]. 6.2, 6.4, 6.5

[109] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, and M. Wiedermann, *HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR*, Comput. Phys. Commun. **182** (2011) 1034–1046, arXiv:1007.1327 [hep-ph]. 6.2

[110] J. Ferrando and D Wendland, *Reference $t\bar{t}$ Production Cross Sections for Use in ATLAS Analyses*, ATL-COM-PHYS-2014-112, 2014. 6.2, 6.2, 6.4, 6.5

[111] J. B. et al., *2013 Review of Particle Physics*, Phys. Rev. D **86** (2012 and 2013 partial update for the 2014 edition) 010001. 6.2, 6.5, 6.10

[112] R. Brun and F. Rademakers, *ROOT: An Object Oriented Data Analysis Framework*, Nucl. Instrum. Meth. **A389** (1997) 81–86. 6.3.3

[113] ATLAS Collaboration, *Determination of the Top-Quark Mass from the $t\bar{t}$ Cross Section Measurement in pp Collisions at $\sqrt{s} = 7$ TeV with the ATLAS Detector*, ATLAS-CONF-2011-054, 2011, http://cdsweb.cern.ch/record/1342551. 6.4.3, 6.10

[114] J. M. Campbell and R. K. Ellis, *Top-Quark Processes at NLO in Production and Decay*, J. Phys. **G42** no. 1, (2015) 015005, arXiv:1204.1513 [hep-ph]. 6.5

[115] O. Eberhardt, G. Herbert, H. Lacker, A. Lenz, A. Menzel, U. Nierste, and M. Wiebusch, *Impact of a Higgs Boson at a Mass of 126 GeV on the Standard Model with Three and Four Fermion Generations*, Phys. Rev. Lett. **109** (2012) 241802, arXiv:1209.1101 [hep-ph]. 6.5

216

[116] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Kennedy, R. Kogler, K. Moenig, M. Schott, and J. Stelzer, *The Electroweak Fit of the Standard Model after the Discovery of a New Boson at the LHC*, Eur. Phys. J. **C72** (2012) 2205, arXiv:1209.2716 [hep-ph]. 6.5

[117] DØ Collaboration, *Determination of the Pole and MSbar Masses of the Top Quark from the $t\bar{t}$ Cross Section*, Phys. Lett. **B703** (2011) 422–427, arXiv:1104.2887 [hep-ex]. 6.10

[118] DØ Collaboration, *Combination of $t\bar{t}$ Cross Section Measurements and Constraints on the Mass of the Top Quark and its Decays into Charged Higgs Bosons*, Phys. Rev. **D80** (2009) 071102, arXiv:0903.5525 [hep-ex]. 6.10

[119] ATLAS Collaboration, *Measurement of the Top-Quark Mass using the Template Method in pp Collisions at $\sqrt{s} = 7$ TeV with the ATLAS Detector*, ATLAS-CONF-2011-033, 2011, http://cdsweb.cern.ch/record/1337783. 6.10

[120] CMS Collaboration, *Measurement of the $t\bar{t}$ Production Cross Section in the $e\mu$ Channel in pp Collisions at $\sqrt{s} = 7$ and 8 TeV*, Tech. Rep. CMS-PAS-TOP-13-004, CERN, Geneva, 2015. http://cds.cern.ch/record/2048642. 7

[121] C. Deluca et. al., *Radiation Damage Simulation in the ATLAS Pixel Detector*, ATL-COM-INDET-2011-011, 2011. A

[122] ATLAS Collaboration, *Radiation Background Simulations: Utopia*, https://twiki.cern.ch/twiki/bin/viewauth/Atlas/ \RadiationBackgroundSimulationsUtopia#IBL_studies. A.1, A.1

[123] M. Capeans, G. Darbo, K. Einsweiller, M. Elsing, T. Flick, M. Garcia-Sciveres, C. Gemme, H. Pernegger, O. Rohne, and R. Vuillermet, *ATLAS Insertable B-Layer Technical Design Report*, Tech. Rep. CERN-LHCC-2010-013. ATLAS-TDR-19, CERN, Geneva, Sep, 2010. http://cds.cern.ch/record/1291633. A.1, A.1

[124] ATLAS Collaboration, *Radiation Background Simulations*, https://twiki.cern.ch/twiki/bin/viewauth/Atlas/ \RadiationBackgroundSimulations#RZ_maps_of_fluence_and_dose_AN1. A.1, A.2, A.1

[125] M. Moll, *Radiation Damage in Silicon Particle Detectors: Microscopic Defects and Macroscopic Properties*, DESY, 1999. A.2.1, A.2.2

[126] G. Kramberger, V. Cindro, I. Mandić, M. Mikuž, and M. Zavrtanik, *Determination of Effective Trapping Times for Electrons and Holes in Irradiated Silicon*, Nucl. Instr. and Meth. **A476** (2002) 645–651. A.2.3, A.4.5, A.5.2

217

[127] M. B. e. a. J. Weingarten, S. Altenheiner, *Planar Pixel Sensors for the ATLAS Upgrade: Beam Tests Results*, Journal of Instrumentation**7** no. 10, P10028. A.2.3

[128] Z. He, *Review of the ShockleyRamo Theorem and its Application in Semiconductor Gamma-Ray Detectors*, Nuclear Instruments and Methods in Physics Research **463** no. A, (2001) 250–267. A.3.1

[129] W. Shockley, *Electrons and Holes in Semiconductors*,. A.3.2

[130] J. R. Haynes and W. C. Westphal, *The Drift Mobility of Electrons in Silicon*, Phys. Rev. **85** (1952) 680. A.3.2

[131] T. Lari, *Lorentz Angle Variation with Electric Field for ATLAS Silicon Detectors*,. A.3.2

[132] S. M. Sze, *Physics of Semiconductor Devices*,. A.3.3, 4

[133] A. Einstein, *Über die von der Molekularkinetischen Theorie der Wärme Geforderte Bewegung von in Ruhenden Flüssigkeiten Suspendierten Teilchen*, Annalen der Physik**8** . A.3.3

[134] V. Eremin, E. Verbitskaya, and Z. Li, *The Origin of Double Peak Electric Field Distribution in Heavily Irradiated Silicon Detectors*, Nucl. Instrum. Meth. **A476** (2002) 556–564. A.5, A.3.4, A.4.1

[135] V. Chiochia, M. Swartz, D. Bortoletto, L. Cremaldi, S. Cucciarelli, A. Dorokhov, C. Hormann, D. Kim, M. Konecki, D. Kotlinski, K. Prokofiev, C. Regenfus, T. Rohe, D. A. Sanders, S. Son, and T. Speer, *Simulation of Heavily Irradiated Silicon Pixel Sensors and Comparison With Test Beam Measurements*, IEEE Transactions on Nuclear Science **52** (2005) 1067–1075, physics/0411143. A.5, A.3.4, A.4.1, A.4.2, A.7, A.8, A.9

[136] V. Chiochia et al., *A Double Junction Model of Irradiated Silicon Pixel Sensors for LHC*, Nucl. Instrum. Meth. **A568** (2006) 51–55, arXiv:physics/0506228 [physics.ins-det]. A.5, A.3.4, A.4.1, A.4.2, A.7, A.8, A.9

[137] G. Kramberger, *Electric Field and Space Charge in Neutron Irradiated n^+p Sensors*, https://indico.cern.ch/getFile.py/access?contribId=7&sessionId=1&resId=0&materialId=slides&confId=148833. A.3.4

[138] Henrich, B. et. al., *Depth Profile of Signal Charge Collected in Heavily Irradiated Silicon Pixels*, CMS, 1997, http://cmsdoc.cern.ch/doc/notes/docs/NOTE1997_021. A.3.4

[139] Z. Li, *Modeling and Simulation of Neutron Induced Changes and Temperature Annealing of N_{eff} and Changes in Resistivity in High Resistivity Silicon Detectors*, Nucl. Instr. and Meth. **A342** (1994) 105–118. A.3.5

[140] Silvaco, Inc. A.4

[141] D. Pennicard, C. Fleta, R. Bates, V. O'Shea, C. Parkes, G. Pellegrini, and N. Tartoni, *Simulations of Radiation-Damaged 3D Detectors for the Super-LHC*, Nucl. Instrum. Meth. **A592** (2008) 16–25. A.4.1

[142] M. Petasecca, F. Moscatelli, D. Passeri, and G. U. Pignatel, *Numerical Simulation of Radiation Damage Effects in p-type and n-type FZ Silicon Detectors*, IEEE Trans. Nucl. Sci. **53** (2006) 2971–2976. A.4.1

[143] V. Eremin, *Status and Program of the RD50 Simulation Working Group*, https://indico.cern.ch/getFile.py/access?contribId=1&sessionId=0&resId=1&materialId=slides&confId=237859. A.4.4

[144] G. Alimonti, A. Andreazza, M. Citterio, T. Lari, C. Meroni, F. Ragusa, and C. Troncon, *A Study of Charge Trapping in Irradiated Silicon with Test Beam Data*, Tech. Rep. ATL-INDET-2003-014, CERN, Geneva, Aug, 2003. http://cds.cern.ch/record/685542. A.5.2

[145] C. Jacoboni, C. Canali, G. Ottaviani, and A. A. Quaranta, *A Review of Some Charge Transport Properties of Silicon*, Solid State Electronics **20** (1977) 77–89. A.5.2, A.6, A.15

[146] Kirnas, G. et. al. Phys. Stat. Sol. **A23** (1974) 123. A.6

[147] Landolt-Börnstein, *Numerical Data and Functional Relationships in Science and Technology*,. A.6

[148] ATLAS Collaboration, *The ATLAS IBL Upgrade: IBL Test Beam June 2011*, https://twiki.cern.ch/twiki/bin/view/Atlas/IBLTBJune2011. A.6.2

Appendix A

Irradiation Simulation
in the ATLAS Pixel Detector

"The most exciting phrase to hear in science, the one that heralds new discoveries,

is not Eureka! (I found it!) but rather, 'Hmm…That's funny…' "

– Isaac Asimov

Because the ATLAS Pixel Detector is situated closest to the interaction point to track particles resulting from pp collisions with maximal accuracy, it also receives the most particles radiating outward from the collision, which over time modify the response of the Pixel Detector. Designed to perform well under conditions of moderate radiation damage (up to fluence of 10^{15}(1 MeV) n_{eq}/cm^2), the Pixel Detector software has not, through the $2009 - 2012$ run, included compensation for irradiation. However, as fluence will exceed this threshold in successive operational phases, accounting for irradiation effects is increasingly crucial for maintaining realistic signal formation. The following solution uses Technology Computer Aided Design (TCAD) simulations for the electric field and Ramo potential, along with calculations such as the projected

220

time the charge travels before being trapped, to obtain a reasonable estimate of the radiation damage effect. Simulations are verified with test beam data using the AllPix simulation package.

The following appendix defines fluence thresholds at the LHC in Section A.1 and the expected effects in Section A.2, as well as the key components of signal formation in Section A.3. The TCAD simulation technique and results are described next in Section A.4 followed by the implementation in the ATHENA digitizer in Section A.5. Section A.6 discusses the first round of validation testing using AllPix software and test beam data, and Section A.7 gives the conclusion.

This project benefited greatly from the previous work of Caroline Deluca, Beniamino de Girolamo, Stephen Gibson, Giacinto Piacquadio, John Stupak, and Dmitri Tsybychev, as well as their ATLAS Internal Note, "Radiation Damage Simulation in the ATLAS Pixel Detector" [121].

A.1   Expected Radiation Fluences

As the LHC operates for a second run (scheduled for 2015-2018) and subsequently a third run, fluences (Φ) will accumulate in the Pixel Detector up to an estimated 3.3×10^{15} and 10×10^{15} (1 MeV) n_{eq}/cm^2, shown in Figure A.1 and Figure A.2, respectively.

The fluence projected to be accumulated by the end of Run 2, corresponding to an integrated luminosity of $\mathcal{L} = 550$ fb^{-1} at 84.5 mb, was calculated with the PHOJET pp event generator and FLUKA 2008 particle transport code [122, 123]. Figure A.1a shows
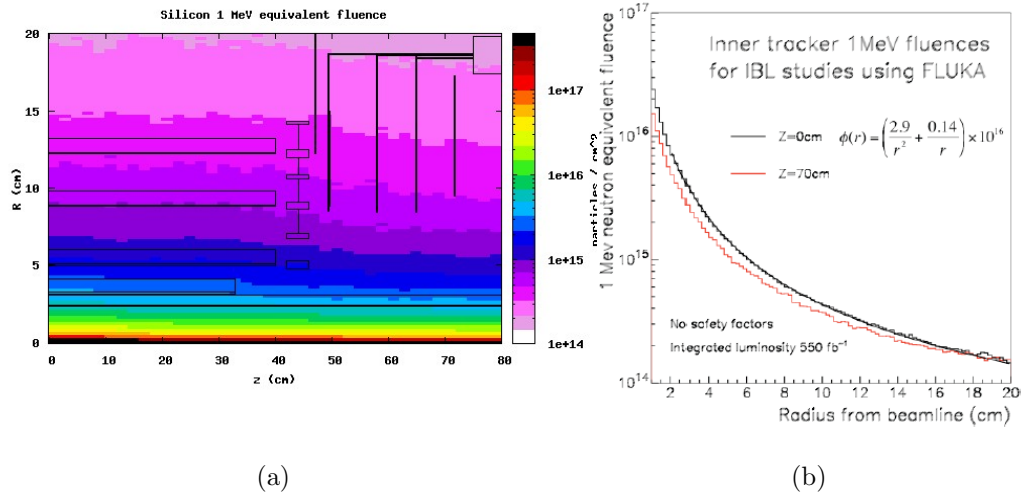
Figure A.1: Projected fluence accumulated during Run 2 of the LHC. Fluences of 1 MeV neutron equivalent silicon damage are normalized to an integrated luminosity of 550 fb^{-1} [122, 123].

the result as a function of distance from the interaction point, both radially (R) and along the beam axis (Z). Figure A.1b illustrates the expected fluence after $\mathcal{L} = 550$ fb^{-1} as a function of radial distance from the interaction point for two distances along the beam axis ($z = 0$ cm and $z = 70$ cm). For the radial distance $r = 3.1$ cm, corresponding to the minimum radius of the Insertable B-Layer (IBL) sensors [123], the 1 MeV neutron fluence is estimated to be $\Phi = 3.3 \times 10^{15}$ n$_{\text{eq}}$/cm^2. The fluence accumulated by the end of Run 3 was simulated with Pythia8 event simulations, normalized to an integrated luminosity of $\mathcal{L} = 3000$ fb^{-1} for 79.310 mb [124]. The results are presented in Figure A.2a as a function of the distance from the interaction point, again along the beam axis (Z) and radially (R). Figure A.2b shows the expected integrated fluence at the end of the Run 3 at $z = 0$ as a function of radial distance from
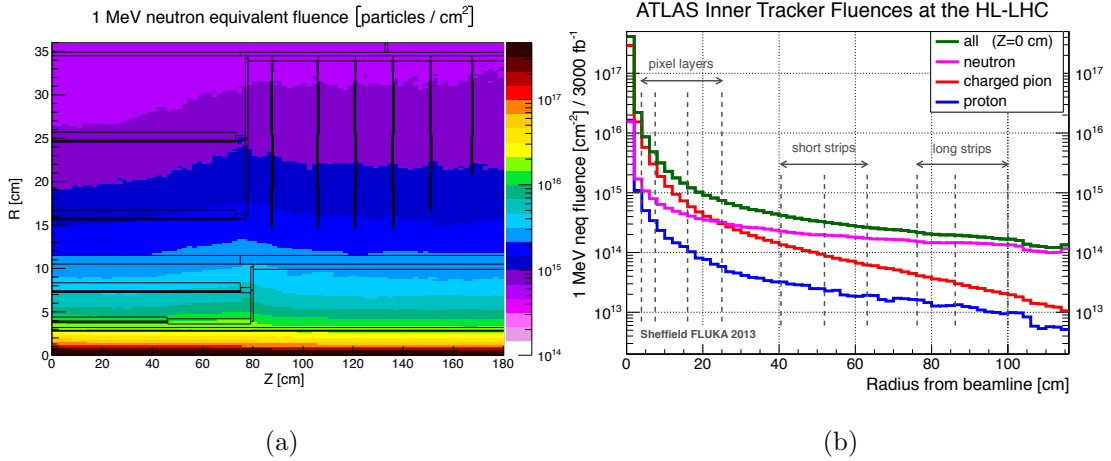
Figure A.2: Projected fluence accumulated over the course of Run 3 of the LHC. Fluences of 1 MeV neutron equivalent silicon damage are normalized to an integrated luminosity of 3000 fb^{-1} [124].

the interaction point; the individual contribution to the total fluence from neutrons, charged pions and protons is also presented. In the innermost pixel layer, the expected fluence at $z = 0$ after 3000 fb^{-1} is in excess of 1×10^{16} (1 MeV) n_{eq}/cm^2. The fluence predictably decreases as radial distance from the interaction point increases: At a distance of $r = 20$ cm, the fluence is about 1^{15} n_{eq}/cm^2. Table A.1 lists the expected fluence by the end of Run 3 for the different pixel barrel layers and disks. The expected fluence varies by a factor of 10 between the innermost and outermost layer.

A.2 Irradiation Effects

Fluence of these magnitudes causes two main effects in silicon sensors: **ionizing energy loss**, in which damage of the Silicon Oxide and SiO$_2$-Si interface occurs as

Location	R (cm)	Z (cm)	Φ (1 MeV n_{eq} 10^{14} cm^{-2})
layer 1	3.9	45.5	132.9
layer 2	7.8	74.5	50.8
layer 3	15.5	72.0	17.7
layer 4	25.0	72.0	9.3
disk 1	15.05	87.7	17.0
disk 5	17.1	150.9	13.2
disk 6	21.3	167.5	10.3

Table A.1: Table of expected fluence in the ATLAS Pixel Detector, by radial (R) and beam axis (Z) distance from the interaction point, by the end of Run 3 [124].

electrons escape from the oxide and create short-circuits, and **non-ionizing energy loss**, in which the sensor bulk is damaged. The non-ionizing defects affect the pixel sensors more critically, since the insulation in the sensor structure limits ionizing energy loss [37]. The following sections discuss the three primary sensor bulk effects: effective doping concentration changes, leakage current increases, and charge trapping.

A.2.1   Effective Doping Concentration

ATLAS pixel sensors consist of an array of diodes on a high-resistivity n-type bulk, with an asymmetric depletion region at the p^+-n junction that operates in reverse bias voltage across the sensor bulk [37]. Defects create states in the forbidden silicon energy band gap (such as described in Ref. [125]). When occupied, these states change the effective doping concentration N_{eff}[1], as shown in Figure A.3. Radiation causes these changes through the removal of donor impurities and the creation of acceptor impurities.

Because the ATLAS pixel sensors use n-type wafers, radiation first decreases N_{eff}, both in absolute and relative value, causing a decrease in the depletion voltage. After a certain fluence threshold is reached, the sensor bulk type changes, as illustrated schematically in Figure A.4a and A.4b, after which the depletion voltage and N_{eff} increase. The depletion zone grows from the n-side after type inversion, which enables signal formation even if the bulk is not fully depleted.

ATLAS pixel sensors undergo type inversion at an estimated integrated luminosity

---

[1]N_{eff} is defined as positive (negative) for n-(p-)bulk material
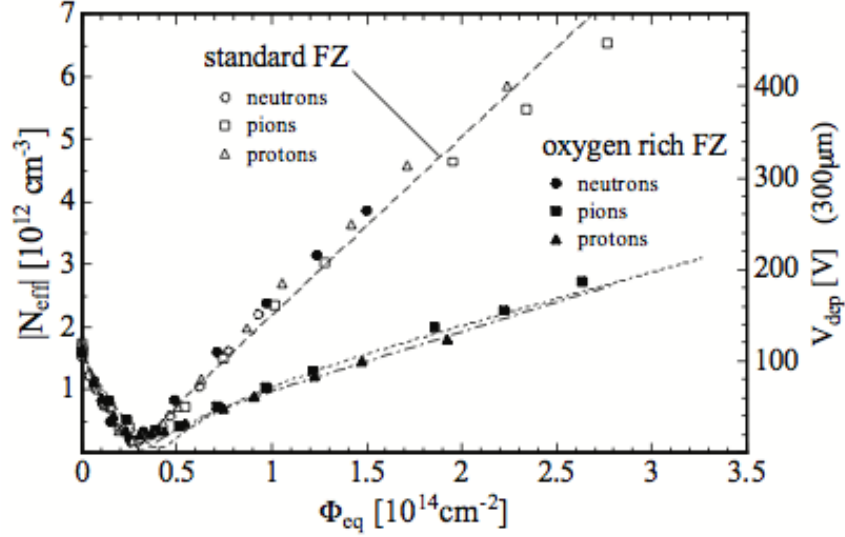
Figure A.3: Effective doping concentration and depletion voltage in the pixel sensors, as a function of fluence, for standard and oxegenated silicon. ATLAS uses oxegenated silicon to mitigate radiation effects [37].

of 10 fb^{-1}. After this, the voltage required to fully deplete the sensor increases gradually over the detector's operational lifetime.

A.2.2 Leakage Current

Irradiation creates generation and recombination centers in the silicon. Because of this, the leakage current through the sensor bulk increases linearly with fluence, in turn increasing electronic noise and power consumption in the detector, and thus changing in the signal-to-noise ratio in the detector.

Fortunately, this increase of leakage current can be somewhat mitigated by controlling the temperature of the sensors' environment, since the leakage current depends
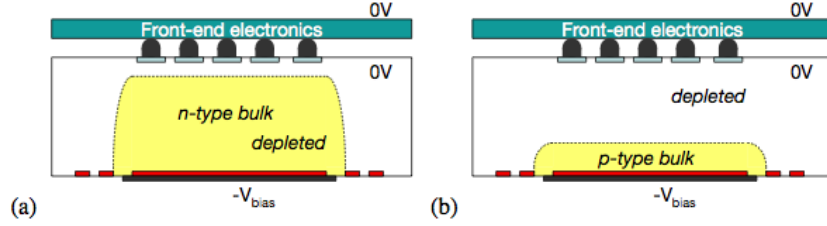
Figure A.4: Depletion zones in the pixel sensors a) before and b) after type inversion [37].

on the temperature according to the following equation [125]:

$$I_{\text{leak}}(T) \approx T^2 e^{-E_g/2k_B T} \tag{A.1}$$

where k_B represents the Boltzmann constant, T the absolute temperature of the sensor environment, and E_g the energy of the gap for silicon.

A.2.3   Charge Trapping

Impurities that develop in the sensors due to irradiation can also prevent some charge carriers from traveling to the electrode for collection. The effective trapping time, $\tau_{e,h}$, which gives the average time that electrons and holes travel through the sensor before being trapped, decreases proportionally with fluence. The total collected charge therefore decreases as fluence increases, according to the relationship $Q(t_{\text{coll}}) \propto e^{-t_{\text{coll}}/\tau}$, where t_{coll} represents the time the charge would take to drift to the collecting electrode with no trapping. The decrease in collected charge affects tracking resolution and b-tagging efficiencies.

Following Ref. [126], a calculation for the trapping time $\tau_{e(h)}$ of electrons (holes) can be given for relevant fluence thresholds. If saturation velocity is assumed for

227

Φ (10^{14} 1MeV n_{eq}/cm^2)	τ_e (ns)	τ_h (ns)	λ_e (μm)	λ_h (μm)
1	24	16	2400	1600
10	2.4	1.6	240	160
100	0.24	0.16	240	16

Table A.2: Table of expected trapping time and distance for electrons and holes.

charge carriers, the average trapping distance λ can be estimated as:

$$\lambda_{e,h} = v_{e,h}^{sat} \cdot \tau_{e(h)} \tag{A.2}$$

Table A.2 lists estimates of the trapping time and distance for electrons and holes for several fluence thresholds. For a sensor of thickness w and a trapping distance of λ the expected charge collection efficiency (CCE) can be estimated as:

$$\text{CCE} = \frac{\lambda}{w} \left[1 - e^{-\frac{w}{\lambda}} \right] \tag{A.3}$$

Based on these estimates, the CCE can be derived for some fluence and sensor thickness values as in Table A.3. The calculations given in Table A.3 do not include the signal induced by trapped charges, so it should be regarded as a lower bound for expected signal; recent results indicate that more charge than expected is observed after fluence thresholds are reached comparable to those expected by the end of Run 3 [127].

A.3   Signal Formation

This section summarizes the signal formation process and the effects of radiation damage on that process given the sensor changes discussed in Section A.2.

228

Φ (10^{14} 1MeV n_{eq}/cm^2)	w (μm)	CCE (%)
1	250	95
5	200	82
5	250	78
10	200	67
10	250	62
50	200	23
50	250	19
100	100	23
100	150	15
100	200	11
100	250	9

Table A.3: Table of expected charge collection efficiency (CCE) for electrons as a function of fluence Φ and bulk thickness w.

A.3.1 Signal Formation and Induced Signal

A charged particle traversing the silicon sensor bulk produces electron holes pairs with a most probable value (MPV) of 80 $\frac{\text{pairs}}{\mu\text{m}}$. The created charge carriers drift toward the sensor electrodes (the electrons toward the pixel electrode, the holes toward the backside electrode) under the influence of the electric field present in the depleted region. This movement of the charge carriers in the electric field induces signal on the readout electrodes, described by the Shockley-Ramo theorem [128]. The theorem states that the charge Q and current i on an electrode induced by a moving point charge q depend on the instantaneous velocity $\vec{v}$ of charge q, and the electric potential V_w and electric field $\vec{E}_w$ that exist due to q at instantaneous position $\vec{r}$:

$$Q \;=\; -qV_w(\vec{r}) \tag{A.4}$$

$$i \;=\; q\vec{v} \cdot \vec{E}_w(\vec{r}) \tag{A.5}$$

given the selected electrode is held at unitary potential, all other electrodes at zero potential, and all charges removed.

The sum of induced currents gives the total instantaneous current $I(t)$:

$$I(t) = \sum i(t) = \sum q\vec{E}_w(\vec{r}) \cdot \vec{v}_{e,h}(t,\vec{r}) \tag{A.6}$$

where the carrier drift velocity is the product of the drift electric field $\vec{E}(\vec{r})$ with the carrier mobility $\mu_{e,h}$:

$$\vec{v}_{e,h}(t,\vec{r}) = \mu_{e,h}(\vec{E},T)\vec{E}(\vec{r}) \tag{A.7}$$

As shown in Equation A.7, the carrier mobility $\mu_{e,h}$ depends on the electric field, the temperature, and the nature of the carrier (see Section A.4.5). Conveniently for the

digitizer calculations, the calculation of the induced signal for each charge carrier is completely independent from all the other carriers (see Section A.5).

A carrier q that either completes its path to the electrode or is trapped, at position $\vec{r}_f$ relative to its initial position $\vec{r}_i$, induces total charge Q given by the Shockley-Ramo Theorem. $\vec{V}_w(\vec{r})$ represents the weighting ("Ramo") potential at $\vec{r}$:

$$Q = -q\left[V_w(\vec{r}_f) - V_w(\vec{r}_i)\right] \tag{A.8}$$

Both trapped electrons and trapped holes reduce the final signal amplitude, as shown in Equation A.8.

A.3.2 Charge Drift

As charges travel through the electromagnetic field, they are affected by the Lorentz force:

$$F_L = q\left[\vec{E}(\vec{r}) + \vec{v}_{e,h}(t,\vec{r}) \times \vec{B}\right] \tag{A.9}$$

This force is at a maximum in the Pixel Detector, since the electric and magnetic fields are perpendicular. The charges drift in the direction defined by Lorentz angle θ_L, calculated from the magnetic field B, the drift mobility $\mu_{e,h}$, and the Hall factor r [129]:

$$\tan\theta_L = \mu_{H(e,h)}B = r\mu_{e,h}B \tag{A.10}$$

The Hall factor includes dependence on the scattering mechanism details of charge carriers in the material, as well as temperature. Table A.6 in Section A.5 shows the parameters for silicon.

This drift angle changes as fluence, bias voltage, depletion voltage and temperature change, since the electric field (and thus the mobility) depend on these factors [130, 131]. Thus a careful inclusion of this effect is necessary to maintain accuracy in cluster size, tracking resolution, and other derived quantities.

A.3.3 Charge Diffusion

The charges also diffuse as they move through the pixel cell, due to the influence of the Lorentz force:

$$r_{\text{dif}} = L_d \sqrt{\frac{\cot \theta_L}{0.3}} \tag{A.11}$$

where the diffusion length L_d represents the distance the carriers diffuse before being annihilated [132]:

$$L_d = \sqrt{D_{e,h} \cdot t} \tag{A.12}$$

as derived from the diffusion coefficient $D_{e,h}$ given by the Einstein relation [133]:

$$D_{e,h} = \frac{k_B T}{e} \mu_{e,h} \tag{A.13}$$

In these equations, k_B represents the Boltzmann constant, T the absolute temperature, $\mu_{e,h}$ the drift mobility, e elemental charge, and t the time the charge travels before being trapped.

A.3.4 Modification of the Electric Field

The radiation-induced states in the forbidden gap affect the electric field in the pixel cells by altering the electric field distribution in the bulk. Since many variables used in
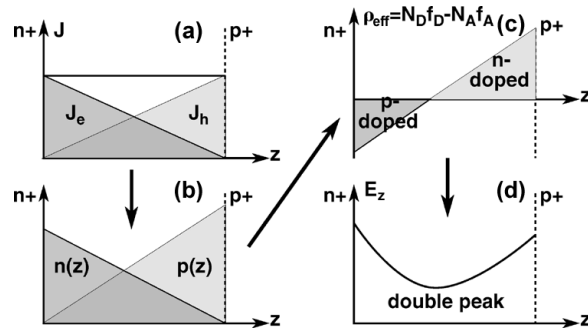
Figure A.5: An illustrative sketch of the EVL model [134] for a reverse biased device [135, 136].

signal formation calculations depend on the electric field (see Section A.4.5), including radiation damage effects requires a careful parameterization of the electric field in the pixels.

Investigation of the electric field profile in silicon sensor bulks have shown that for certain materials, the electric field is no longer linear with the bulk depth after irradiation (see, for example, Refs. [134, 135, 136, 137]). In these materials, the irradiation causes two junctions, one at the pixel side and one at the opposite side, as first explained by Ref. [134] (shown schematically in Figure A.5). As electrons and holes drift toward the electrode and opposite side, respectively, a) the current density becomes non-constant along the bulk, causing b) variation in carrier distribution. The carriers can get trapped while drifting, which translates into c) non-uniform space charge region with a linear dependence on the bulk depth; this gives rise to d) an electric field with two maxima at the two edges of the sensor and a minima somewhere in the middle of the bulk.

These effects can be measured with beam tests, using the "grazing angle" technique [138], and reproduced in TCAD simulations by taking into account some effective states (typically 2 to 4) in the forbidden silicon gap. See Section A.4 for additional detail.

A.3.5 Constancy of the Ramo Potential

Radiation-induced defects do not cause modification of the Ramo potential [139]. Even if part of the sensor's volume is not depleted after irradiation, the bulk resistivity remains close enough to the intrinsic value ($\mathcal{O}(10^5)$ Ω·cm) that it remains non-conductive. Hence the Ramo potential is the same as for a non-irradiated full depleted sensor.

A.4 TCAD Simulation

TCAD simulations of the electric field profile along the bulk, the Ramo potential and the trapping constant were run using Silvaco tools [140]. Silvaco TCAD uses a complete set of physical models for semiconductor device simulation, including models for concentration dependent mobility, field dependent mobility, bandgap narrowing, concentration dependent lifetime, trap-assisted and Auger recombination. Silvaco TCAD also allows custom input of the Oxide fixed charge density (with surface density $Q_{\mathrm{ox}} = 10^{11}/\mathrm{cm}^2$ before irradiation and $Q_{\mathrm{ox}} = 3 \times 10^{12}/\mathrm{cm}^2$ after), generation-recombination lifetimes, and surface recombination velocity according to measured IV and CV characteristics of modules. The following sections summarize simulation of
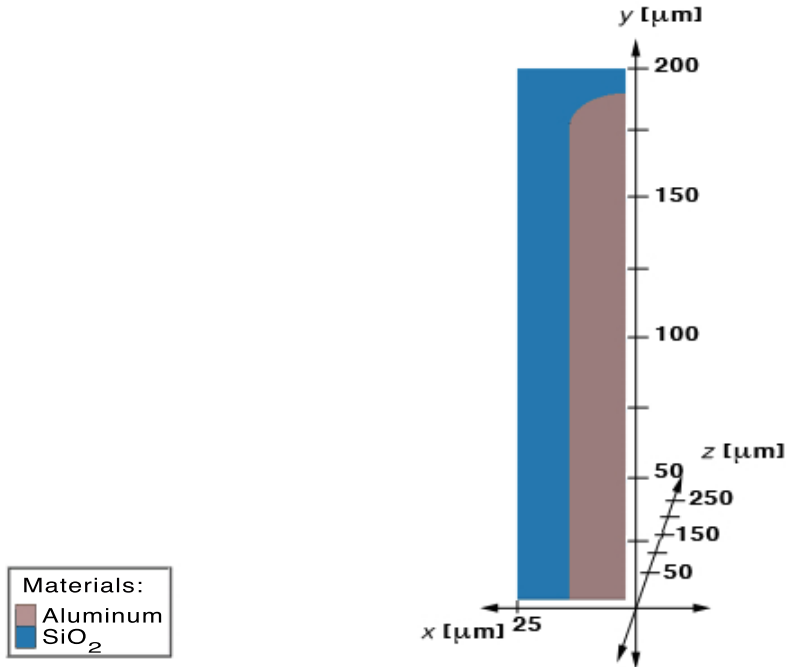
Figure A.6: Structure used to extract the electric field profile.

the electric field, Ramo potential, and charge trapping in the pixel sensors.

A.4.1 Electric Field Simulation

TCAD simulations can parameterize the electric field profile along the bulk as a function of fluence Φ, temperature T and material properties. Due to the symmetries of the electric field and pixel cells, three-dimensional TCAD simulations of one-quarter of a standard ATLAS pixel cell (50 x 400 x 250 μm) simulate the electric field for the entire cell, as shown in Figure A.6.

TCAD simulations can reproduce the radiation-induced modifications to a silicon sensor by the addition of extra effective states in the forbidden gap. Several models exist, including EVL [134], Chiochia [135, 136], Pennicard [141], and Petasecca [142].

Type	Energy (eV)	$\sigma_e(\text{cm}^2)$	$\sigma_h(\text{cm}^2)$	$\eta(\text{cm}^{-1})$
A	E_C -0.42	9.5×10^{-15}	9.5×10^{-14}	1.613
A	E_C -0.46	5.0×10^{-15}	5.0×10^{-14}	0.9
D	E_V +0.36	3.23×10^{-13}	3.23×10^{-14}	0.9

Table A.4: Relevant parameters for acceptors (A) and donor (D) deep levels in the bandgap, describing the radiation damage.

These models include some states in the forbidden gap (2 in EVL and Chiochia, 3 in Pennicard and Petasecca, 4 in other models). Each state is characterized by a type (acceptor (A) or donor (D)), an energy activation value E, an introduction rate η (the ratio between defect density and fluence) and by electron and hole capture cross section $\sigma_{e,h}$. For example, Table A.4 shows the Pennicard model 3 state parameters, where $E_{C,V}$ indicates the conductive and valence band energies, respectively. The model is intended for float zone (FZ) p-type sensors after proton irradiation and was validated with data. It is a good candidate for HL-LHC pixel sensors.

Two models, EVL and Chiochia, can simulate sensors built on diffusive oxygenated float zone (DOFZ) n-type wafers. This project used the Chiochia model, since it accepts values from test beam data of DOFZ n-on-n sensors with 24 GeV protons with the "grazing angle" technique. Table A.5 gives the model parameter values. The variables N_A and N_D represent acceptor and donor concentrations, respectively. The simulations were run at -10 degrees C as in the original model, and the structure

simulated at the three different fluence thresholds reported in Table A.5.

Φ	N_A	N_D	$\sigma_e^{A/D}$	σ_h^A	σ_h^D
$(\text{n}_{\text{eq}}/\text{cm}^2)$	(cm^{-3})	(cm^{-3})	(cm^2)	(cm^2)	(cm^2)
0.5×10^{14}	0.19×10^{15}	0.25×10^{15}	6.60×10^{-15}	1.65×10^{-15}	6.60×10^{-15}
2×10^{14}	0.68×10^{15}	1.0×10^{15}	6.60×10^{-15}	1.65×10^{-15}	6.60×10^{-15}
5.9×10^{14}	1.60×10^{15}	4.0×10^{15}	6.60×10^{-15}	1.65×10^{-15}	1.65×10^{-15}

Table A.5: Chiochia double trap model parameters extracted from the fit to the data.

A.4.2 Electric Field Results

Figures A.7, A.8 and A.9 show the z-component of the electric field E_z across the bulk for the fluences $\Phi = 0.5 \times 10^{14}$, 2×10^{14} and 6×10^{14} $\text{n}_{\text{eq}}/\text{cm}^2$, respectively. The figures compare published models (from Ref. [135, 136]) to this project's TCAD simulation for each fluence, at different bias voltages. This project's electric field simulations should measure slightly higher than the published models, since the CMS pixels sensors used in Ref. [135, 136] were 285 μm thick while current ATLAS sensors are only 250 μm thick. Generally, a higher electric field is expected when the bias voltage is applied over a shorter distance (as in the ATLAS sensors). Considering this, the CMS models and the TCAD simulation results agree well. The Chiochia model predicts an electric field with two maxima, appearing already at what are usually considered low fluences. This module was used to evaluate the trapping time, as described in Section A.4.5, and then validated against beam test data using AllPix, as discussed in Section A.6.
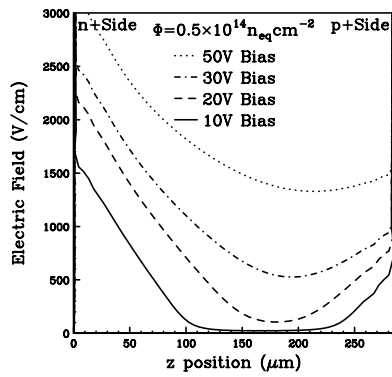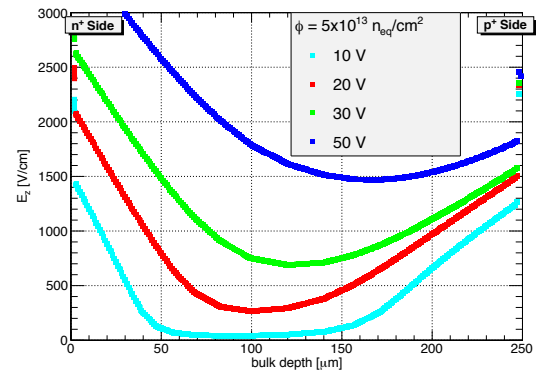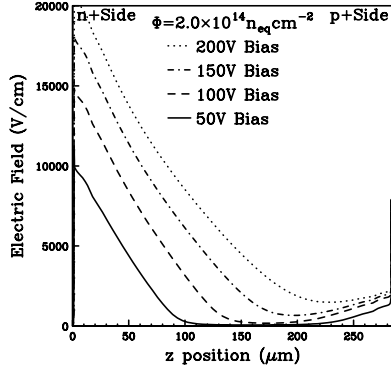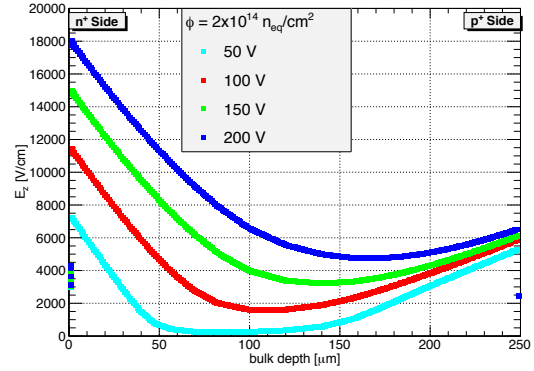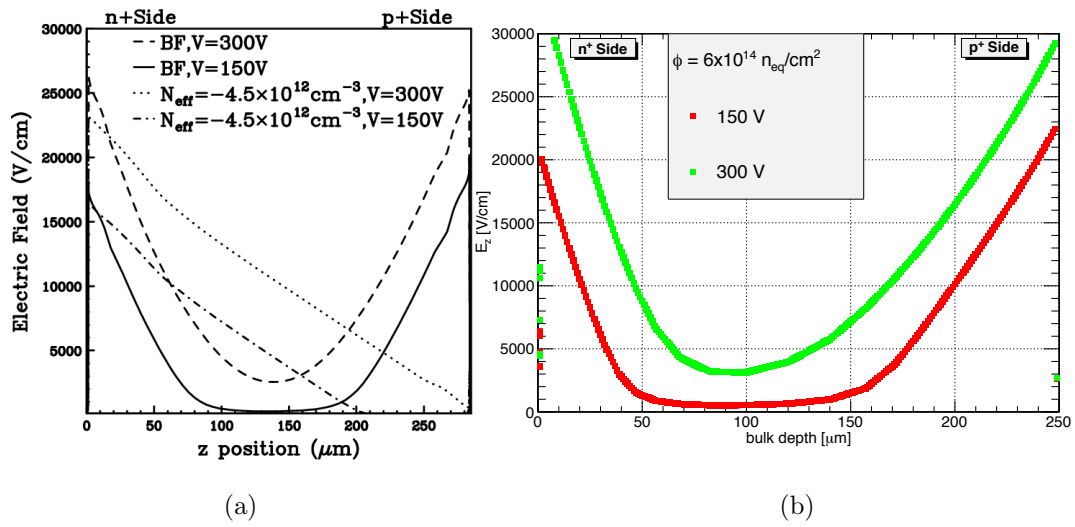
Figure A.7: The z-component of the simulated electric field, plotted as a function of the bulk depth, after a simulated fluence of $\Phi = 0.5 \times 10^{14} n_{eq}/cm^2$. The field profiles for 10, 20, 30 and 50 V are shown a) in the Chiochia model results [135, 136] and b) this project.

(a)   (b)

Figure A.8: The z-component of the simulated electric field, plotted as a function of z, after a simulated fluence of $\Phi = 2 \times 10^{14} \mathrm{n_{eq}/cm^2}$. The field profiles for 50, 100, 150 and 200 V are shown in a) the Chiochia model [135, 136] and b) this project.

A.4.3 Ramo Potential Simulations

Due to symmetries in the pixel cells, three-dimensional TCAD simulations for a structure consisting of one-quarter of a 3×3 matrix of standard ATLAS pixel cells sufficiently simulates the Ramo potential, shown in Figure A.10. To create the simulation, all materials (including implants) except the electrodes (aluminum) are removed from the model, as prescribed by the Ramo theorem. This leaves a structure of a "vacuum" bulk and five aluminum electrodes. The "PX1" electrode (see Figure A.10) is then raised to unitary voltage, keeping all others at ground, and the resulting potential distribution $V_w(\vec{r})$ extracted. Because of symmetries in the pixel structure, only this configuration is necessary; shifting the position $\vec{r}$ to the center of

239

Figure A.9: The z-component of the simulated electric field, plotted as a function of z, after a simulated fluence of $\Phi = 6 \times 10^{14}\,\mathrm{n_{eq}/cm^2}$. The field profiles for 150 and 300 V are shown a) in the Chiochia model [135, 136] and b) this project.
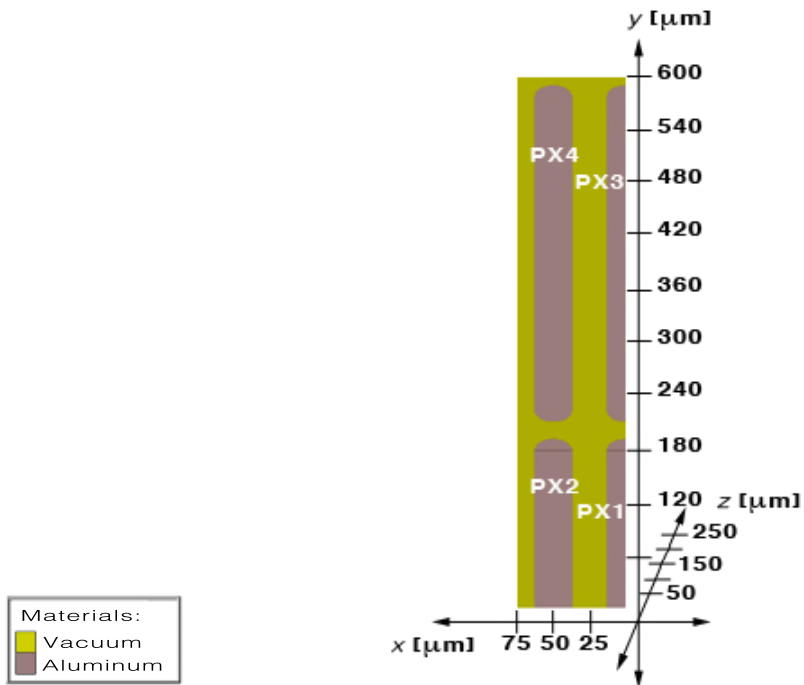
Figure A.10: Structure used to extract the Ramo potential. With dimensions of 75 x 600 x 250 μm, it represents one-quarter of a 3×3 matrix of standard ATLAS pixel cells.
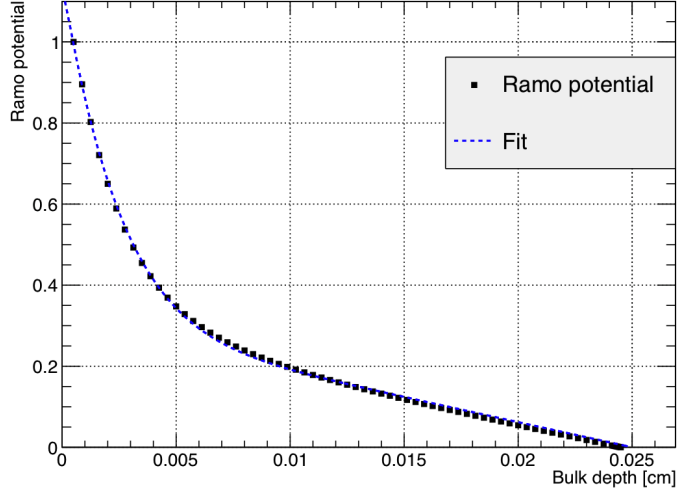
Figure A.11: Ramo potential along the sensor bulk below the pixel under evaluation, displayed as the simulation results (squares) and a fit to the simulation (dashed line).

the pixel under examination gives the Ramo potential in each of the other pixels, as in Equation A.14:

$$V_w^{(i)}(\vec{r}) = V_w^{(1)}(\vec{r} - \vec{r}_C^{(i)}) \equiv V_w(\vec{r} - \vec{r}_C^{(i)}) \tag{A.14}$$

where $V_w^{(i)}(\vec{r})$ represents the Ramo potential for "PXi" and $\vec{r}_C^{(i)}$ remains its geometrical center. As stated, only the Ramo potential for "PX1" $V_w^{(1)}(\vec{r})$ needs to be calculated, which for brevity is indicated as $V_w(\vec{r})$.

A.4.4   Ramo Potential Results

Figure A.11 illustrates the Ramo potential along the bulk depth in a region below the pixel under evaluation. As shown, the Ramo potential decreases rapidly initially (halving within the first 30 μm) but then decays slowly toward zero. A parametrization for the weighting electric field, found from the Ramo potential, is shown in literature [143].

For the case under examination:

$$E_w(z) = \frac{K_1}{D} + K_2 \left(\frac{z}{D}\right)^a \tag{A.15}$$

where D is the bulk thickness, and K_1, K_2 and a are constants determined from data. The simulated and parameterized results agree well, as shown in Figure A.11.

A.4.5 Charge Trapping

While charge trapping TCAD simulations have not yet been built, the following section describes the framework by which the charge trapping can be estimated by TCAD simulation.

The probable time before an electron or hole is trapped, trapping time $\tau_{e,h}$, is inversely proportional to the fluence Φ, as in Equation A.16:

$$\tau_{e,h}^{-1} = \beta_{e,h}\Phi \tag{A.16}$$

The transient current technique (TCT) can be used to extract the $\beta_{e,h}$ coefficient that relates the trapping time $\tau_{e,h}$ to the fluence Φ [126]. This method simulates a laser beam entering parallel to the sensor surface at different fluences and bias voltages beyond full depletion.[2] The induced current follows the formula given in Equation A.6. In the one-dimensional diode approximation, the weighting field is simply $1/D$, where D represents the bulk thickness, and the total instantaneous current $I(t)$ is:

$$I(t)_{e,h} = qN_{e,h}(t)\frac{1}{D}v_{e,h}(t) \tag{A.17}$$

---

[2]The value of bias voltage at full depletion is evaluated by a CV analysis.

where q represents the charge, $N_{e,h}(t)$ the number of drifting electrons and holes, respectively, and $v_{e,h}(t)$ the drift velocity. The amount of the drifting charge decreases with time due to trapping as:

$$N_{e,h}(t) = N_{e,h}(0)e^{\frac{-t}{\tau_{\text{eff}_{e,h}}}} \tag{A.18}$$

where $N_{e,h}(0)$ is the number of generated electron and hole pairs and $1/\tau_{\text{eff}_{e,h}}$ the effective trapping probability. The effective trapping time can be determined by observing the behavior of the current integral at voltages above full depletion. From Equation A.17 and A.18, the measured current including trapping $I_c(t)$ can be calculated from the measured current without trapping, $I_m(t)$ with a single exponential to compensate for trapping:

$$I_c(t) = I_m(t) \exp\left(\frac{t - t_0}{\tau_{\text{tr}}}\right) \tag{A.19}$$

Here t_0 is the carrier injection time (in this case, the start of the laser pulse). If τ_{tr} in Equation A.19 represents the correct effective trapping time then the integral over time is equal for all voltages above full depletion.

A.5   ATHENA Implementation

This section describes the implementation of irradiation damage calculations in the ATHENA digitizer framework.

A.5.1 Structure

The ATLAS pixel digitizer uses simulation software to calculate the relationship between hits in the pixel sensors and expected output from the sensors, as illustrated in Figure A.12. To begin the process, a physics generator program is used to simulate
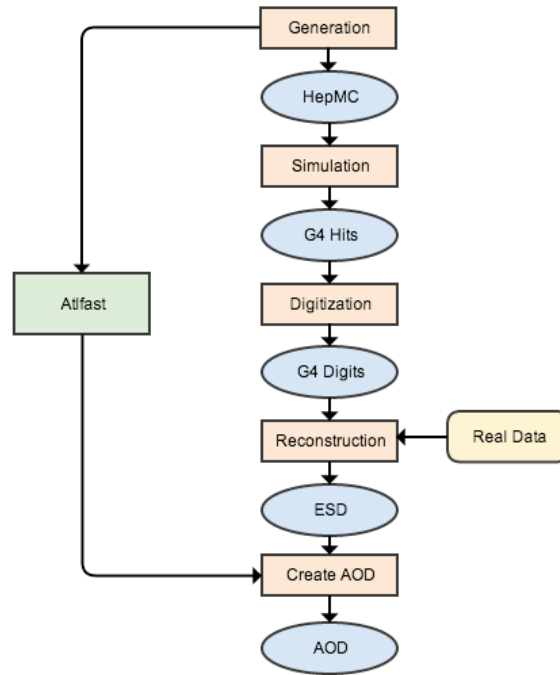

Figure A.12: Schematic representation of the digitizer's Monte Carlo simulation and its use with data.

the appropriate physics process. The stable particles simulated by that process comprise the input into Geant, a detailed detector simulation. Geant outputs the energy loss per step for the process, called "hits". The digitizer uses these hits as input and simulates signal formation, called "digits", producing them in the same form as data. These digits are then used for event reconstruction in combination with data.

The digitizer divides the Geant hits into steps, and then divides each step into sub-steps (referred to here as charges) representing each electron/hole released in the sensor.[3] The digitizer calculates each charge's drift to the electrode plane, including lateral diffusion due to the Lorentz force as discussed previously. The current digitizer assumes that all charges reach the electrode plane, generating signal only in the electrode that each one reaches.

Within the digitizer, inclusion of the discussed radiation effects has been added to the `17.3.4.8` branch of the ATHENA Pixel Digitization software package (`InnerDetector /InDetDigitization /PixelDigitization`) and ATHENA Pixel Conditions Services package (`InnerDetector /InDetCondtions /PixelConditionsServices`). This code is intended to be merged into the main packages after completion of testing, with options to include charge trapping and/or signal induction effects due to the Ramo potential. According to the Digitization package standards, the code has been added as different services accessible from any class.

The Pixel Digitization package includes the main radiation damage code that controls charge trapping, Ramo potential, and other effects, in the `PixelDistortions SimpleSurfaceChargesGenerator` class. This code calls the `ChargeTrappingSvc` and `RamoPotentialSVC` classes in the `PixelConditionsServices` package to calculate the appropriate variables, through the `RadDamageSummarySvc` service. Figure A.13 illustrates the process. Figure A.14 illustrates the logic of including the charge trapping

---

[3]The digitizer calculates each sub-step twice, once as a charge and once as a hole, to include the nontrivial effect of holes in the sensors.
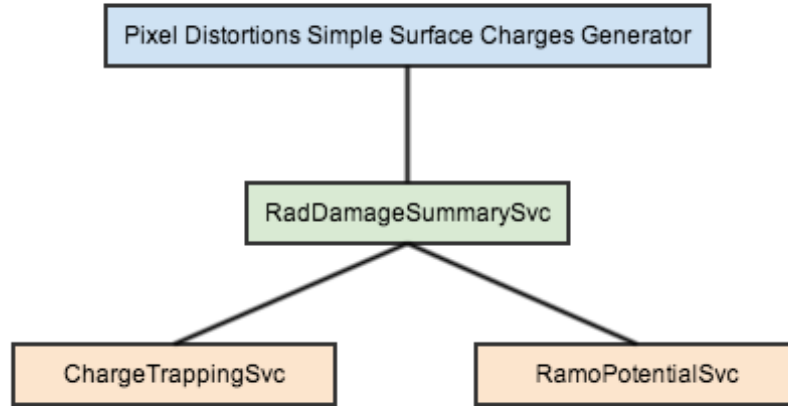
Figure A.13: The structure of radiation damage additions to the `PixelDigitization` package within the ATHENA framework.

and/or Ramo potential effects. The charge trapping and Ramo potential calculations can each be turned on or off in the package's job options, by setting `doRamo` and `doTrapping` to True or False.

A.5.2 Charge Trapping

The digitizer assumes that the probability of a charge being trapped in the sensor increases exponentially the farther that it travels in the z direction through the sensor, according to the following equation:

$$P_{\text{trap}}(z) = 1 - e^{-t/\tau_{e,h}} = 1 - e^{-z/\lambda_{e,h}} \tag{A.20}$$

where t represents the drift time of the charge, τ the trapping constant, and λ the mean free path of the charge through the sensor. The mean free path depends on the
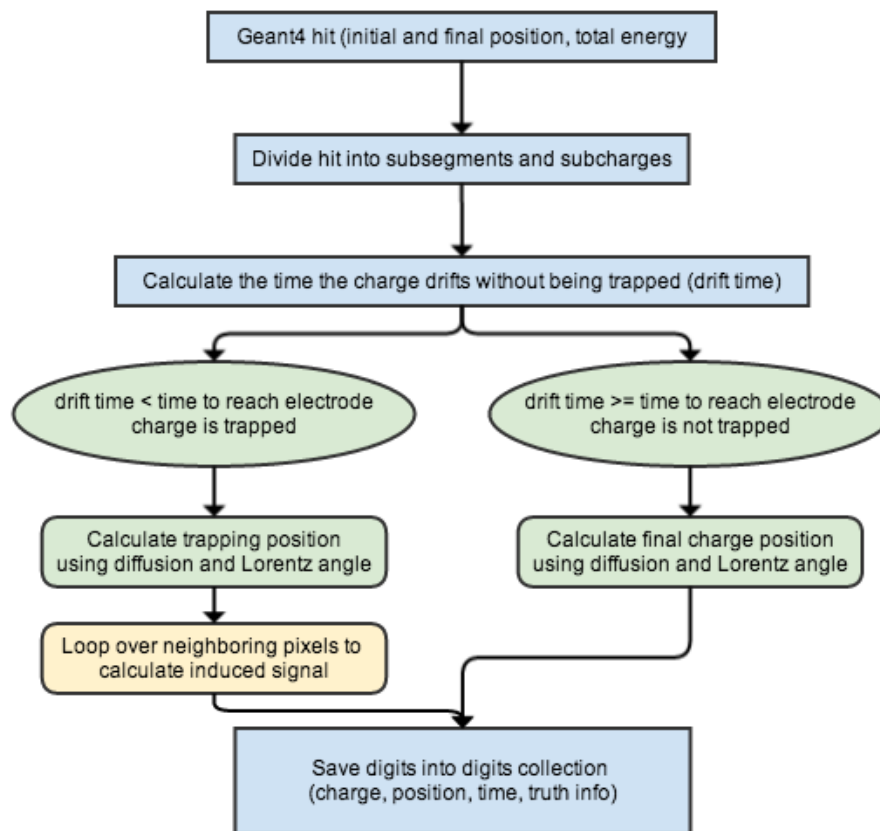
Figure A.14: Flowchart of the digitizer process with charge trapping (green) and/or Ramo potential (yellow) inclusions.

drift velocity v_d and the trapping constant:

$$\lambda_{e,h} = \int v_{e,h} \, d\tau_{e,h} \tag{A.21}$$

with the drift velocity dependent on the mobility and electric field :

$$v_{e,h} = \mu_{e,h} E_{e,h} \tag{A.22}$$

$$\lambda_{e,h} = \int \mu_{e,h} \, E_{e,h} \, d\tau_{e,h} \tag{A.23}$$

In these equations, the trapping constant τ is calculated as a function of fluence, Φ, and a constant of the material, β, as discussed previously:

$$\tau_{e,h}^{-1} = \beta \Phi \tag{A.24}$$

The variable β depends on the detector's annealing history. The digitizer uses the average value from measurements of the 2001 CERN test beams , $\beta = 3 \cdot 10^{-16}$ cm^2/ns [144]. Mobility of the charge carriers depends on the electric field and temperature, in addition to several values that depend on the material, according to the following equation[4] [145]:

$$\mu_{e,h} = \frac{v_s/E_c}{\left(1 + \left(\frac{E}{E_c}\right)^{\beta}\right)^{1/\beta}} \tag{A.25}$$

where the variables used are defined in Table A.6.

Although hole mobility is lower, the effect of hole trapping is not negligible and is included in the digitizer's charge trapping feature.

---

[4]This equation is only valid while $E < E_c$. At electric fields higher than the critical value E_c, the velocity stops increasing and additional interpretations are required [132].

	Electrons	Holes
v_s (cm/s)	$1.53 \cdot 10^9 \cdot T^{-0.87}$	$1.62 \cdot 10^8 \cdot T^{-0.52}$
E_c (V/cm)	$1.01 \cdot T^{1.55}$	$1.24 \cdot T^{1.68}$
β	$2.57 \cdot 10^{-2} \cdot T^{0.66}$	$0.46 \cdot T^{0.17}$
r	$1.13 + 0.0008 \cdot (T - 273)$	$0.72 - 0.0005 \cdot (T - 273)$

Table A.6: Parameterizations in silicon for mobility [145] and the Hall scattering factor [146, 147].

After trapping, the charge has a small probability of being re-released. However, this case considers that probability negligible, since typical de-trapping times greatly exceed the bunch crossing times at the LHC (BCT = 25 ns), so once trapped, the charge is lost for measurement [126].

When possible, the digitizer references the ATHENA Conditions database for input variables; otherwise it uses customized values.[5] With these input values, the digitizer first references the TCAD electric field simulation[6] and then uses the electric field value at the initial position to calculate mobility μ at that position, as shown in Figure A.15.

To appropriately randomize the chances of a charge being trapped, the drift time

---

[5]Future additions could include a database for the radiation damage variables, and/or configurable variables via a job options file.

[6]The TCAD simulation chosen is customized based on chosen fluence. In the future, it would be more robust to enable a dynamic relationship between the set fluence and the resulting mappings used.
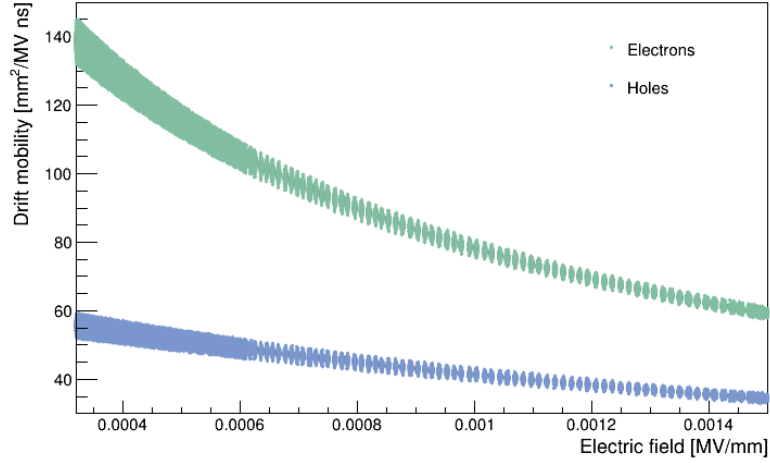
Figure A.15: Comparison of mobility and electric field at initial hit positions. These results are consistent with published data [145].

t is solved from Equation A.20, given a random number between 0 and 1, u:

$$t = -\tau \ln u \tag{A.26}$$

The charge is trapped if this drift time is less than the time it would take for that charge to reach the electrode plane, $t_{\text{electrode}}$[7]:

$$t_{\text{electrode}} = \int_{z_i}^{z_f} \frac{1}{\mu_{e,h}(z) E_{e,h}(z)} dz \tag{A.27}$$

This expression for time to the electrode requires numerical solution of the integral, since the electric field and mobility (and thus the velocity) change as the charge moves through the sensor. A Reimann sum is used to approximate the integral and thus produce a map of time to the electrode as a function of initial depth in the sensor. This map needs to be produced only once per fluence. One example map for electrons

---

[7]Solved numerically with a parameterized fit for the electric field.

(a)   (b)
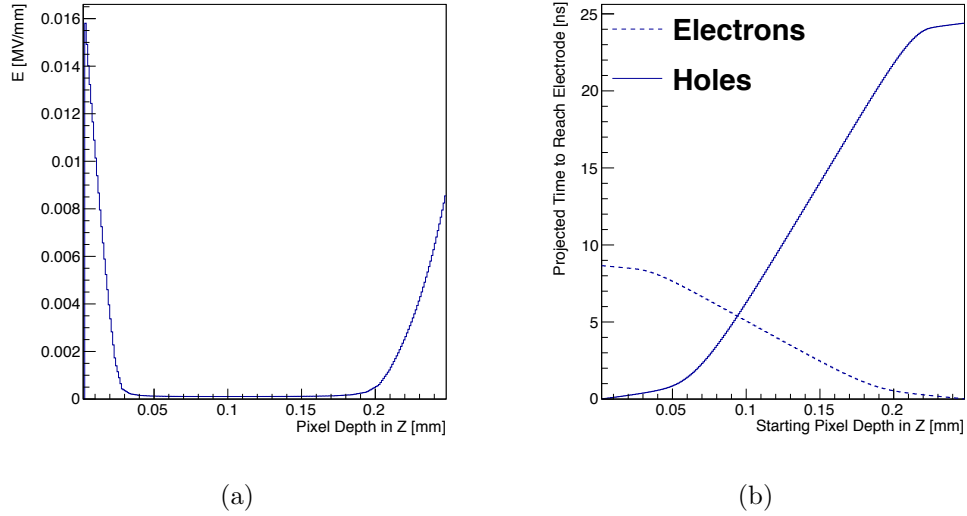
Figure A.16: The a). electric field magnitude and b). time to electrode mapping using Reimann integration for electrons and holes for the electric field shown on the left, which corresponds to 3.8×10^{15} $n_{\mathrm{eq}}/\mathrm{cm}^2$, $V_{\mathrm{bias}} = 400$ V.

and holes is shown in the right plot of Figure A.16 for 3.8×10^{15} $n_{\mathrm{eq}}/\mathrm{cm}^2$, $V_{\mathrm{bias}} = 400$ V.

The digitizer uses the amount of untrapped charge to calculate the total charge gathered at the electrode, and pushes the results through the Ramo potential calculation, if included. Figure A.17 illustrates the structure of the charge trapping contributions.

A.5.3 Ramo Potential

The Ramo potential due to a trapped charge is also included when the `doRamo` option is turned on. In this case, the digitizer uses the 3-dimensional position of the charge
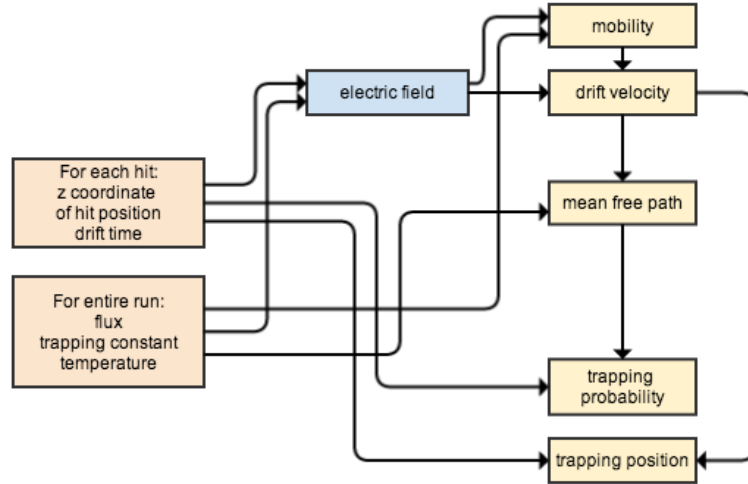
Figure A.17: Structure of the charge trapping services in the ATHENA framework.

and reads the corresponding Ramo potential from the TCAD simulation, included as a 3D ROOT histogram.

A.6 Validation and Testing

This section discusses initial stages of testing and validation of the radiation damage inclusions into the ATHENA pixel digitizer.

A.6.1 Cross-Checks of the $t_{\text{electrode}}$ Calculation

The electric field and time to reach electrode simulated in this project were simulated and shown to be reasonable, as shown in the following figures. Validation and testing efforts continue.
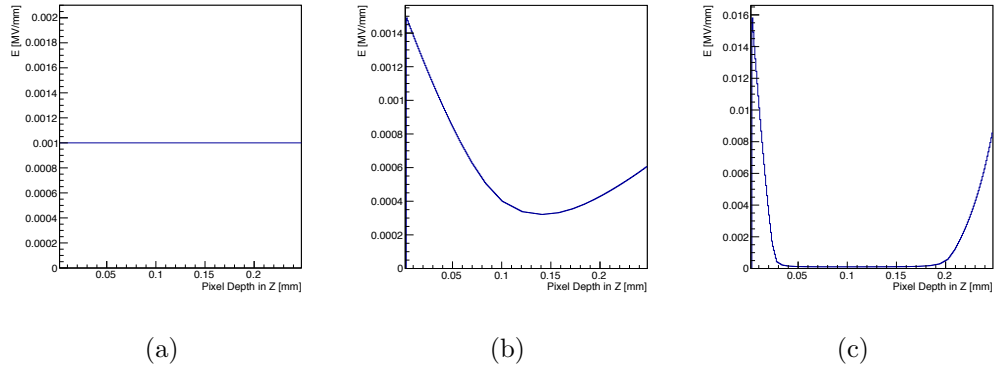
Figure A.18: The a). idealized integration time for a uniform field. b). 1×10^{14} $n_{\text{eq}}/\text{cm}^2$, $V_{\text{bias}} = 200$ V and c). 3.8×10^{15} $n_{\text{eq}}/\text{cm}^2$, $V_{\text{bias}} = 400$ V
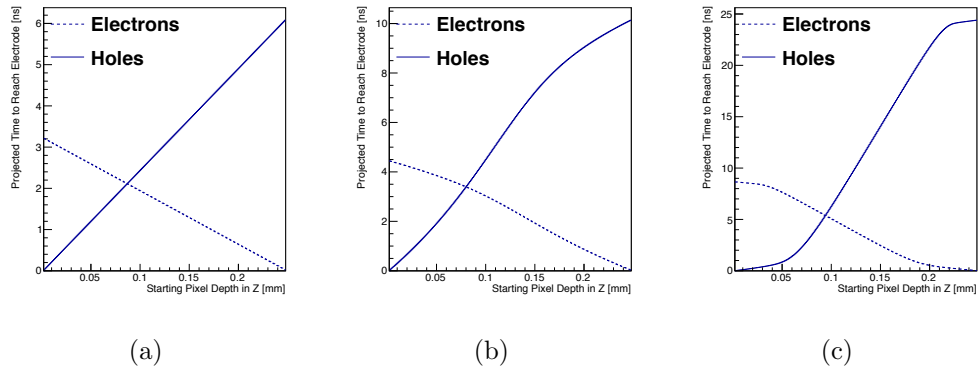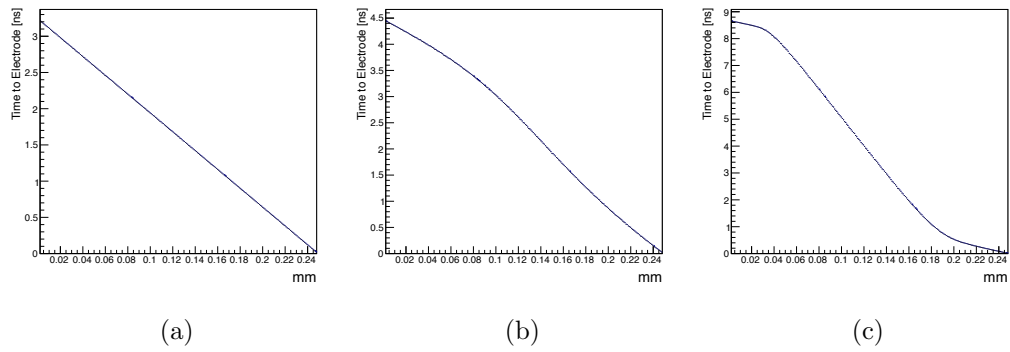

Figure A.19: a). Idealized. b). 1×10^{14} $n_{\text{eq}}/\text{cm}^2$, $V_{\text{bias}} = 200$ V c). 3.8×10^{15} $n_{\text{eq}}/\text{cm}^2$, $V_{\text{bias}} = 400$ V
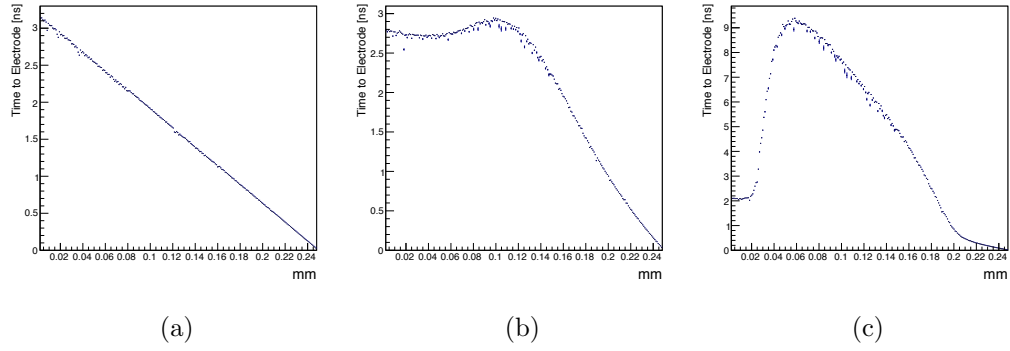
Figure A.20: a). Idealized. b). 1×10^{14} $n_{\mathrm{eq}}/\mathrm{cm}^2$, $V_{\mathrm{bias}} = 200$ V c). 3.8×10^{15} $n_{\mathrm{eq}}/\mathrm{cm}^2$,

$V_{\mathrm{bias}} = 400$ V

Figure A.21: a). Idealized. b). 1×10^{14} $n_{\mathrm{eq}}/\mathrm{cm}^2$, $V_{\mathrm{bias}} = 200$ V c). 3.8×10^{15} $n_{\mathrm{eq}}/\mathrm{cm}^2$,

$V_{\mathrm{bias}} = 400$ V

A.6.2   Test Beam and AllPix

Simulation using AllPix is in the process of being compared to data collected from a CERN test beam in June of 2011, run with the CERN SPS H8 beamline [148]. Figure A.22 shows some representative distributions from the reconstructed tracks. The data from the test beam uses a PPS 200 SE unirradiated sensor with a threshold of 1600e tuned to 8 ToT at 10,000e. Validation efforts are necessary and continuing.

Sample 1

ID ID 40, which is DUT3 from Batch1

Sensor planar SE 200um

Radiation zero fluence

Orientation rotated 15 degrees in phi

Bias Voltage 100V

Sample 2

ID ID 61, which is DUT1 from Batches 3 and 4

Sensor planar SE 200um

Radiation fluence of 6e15 (protons)

Orientation rotated 15 degrees in ϕ

Bias Voltage 600, 800, and 1000V

Sample 3
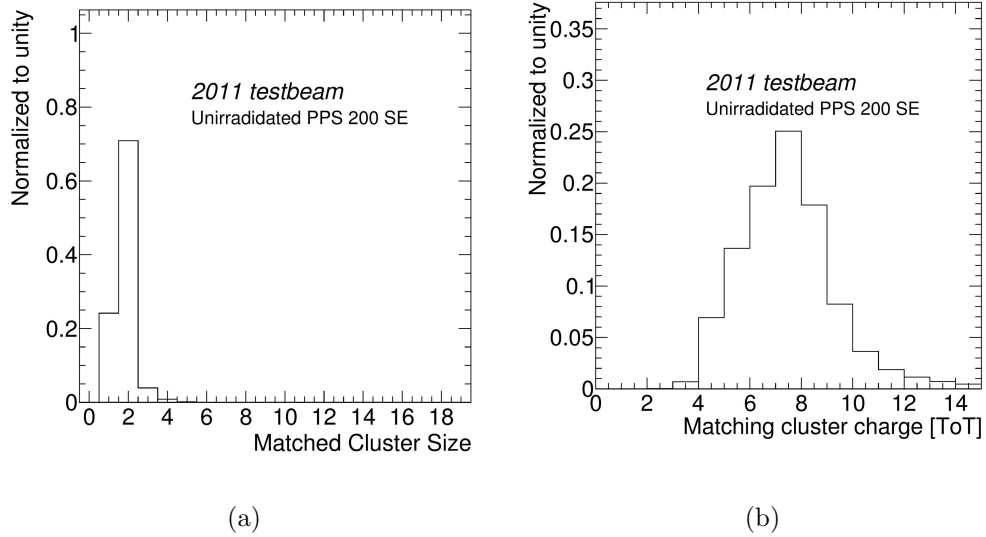
Figure A.22: The left plot shows the matched cluster size and the right plot shows the time over threshold for these clusters using 2011 CERN test beam data.

ID LUB2, which is DUT4 from Batches 3 and 4

Sensor planar SE 250um

Radiation fluence of 3.8e15 (neutrons)

Orientation rotated 15 degrees in ϕ

Bias Voltage 400, 600, 800, 1000V

A.6.3 Monitoring Measurements

In addition to the tests described above, the variables used in the radiation damage services, such as Lorentz angle, leakage current, operating temperature, fluence, and depletion/bias voltages, should be monitored for ongoing accuracy, and the digitizer

updated if necessary. Since the digitizer uses customized values of some of these variables, corresponding updates may be necessary.

A.7 Conclusion

Services to simulate the effects of radiation have been added to the ATHENA pixel digitizer. The current radiation damage services implement TCAD simulations for the parameterized electric field and Ramo potential, dependent on bias voltage and fluence. Validation and testing phases are a continuing effort. Additional improvements could include building dynamic relationships between customized variables such as bias voltage and fluence, with the mappings and derived quantities, as well as simulation of the increase in the electronic noise to signal ratio in the pixels.

Appendix B

Gap Fraction Additional Studies

"Setting goals requires gazing way out at the horizon of your life. But once you set your course, most of the time your awareness should be on the trail under your feet."

– Lauren Fleshman, Runner's World

The following appendix includes several peripheral studies done for the gap fraction measurement, as well as a comprehensive list of data set containers for the simulation samples considered in the gap fraction analysis.

B.1   Ordering of b-Jets

As discussed in Section 4.1, events containing three or more b-tagged jets are selected using the two highest p_T b-jets as coming from the $t\bar{t}$ event, at both reconstruction and (in simulation) at truth level. This affects less than 3% of $e\mu bb$ events, as shown in Figure B.1. Figures B.2 and B.3 show the difference in the gap fraction from simulation when p_T ordering is used versus MV1 ordering at reconstruction level,

which is less than 0.7% using the POWHEG+PYTHIA `fullsim` simulation (the `fastsim` POWHEG+PYTHIA simulation agrees within 0.09%). The difference is calculated as the difference in the gap fraction by the formula:

$$\Delta(f_{\text{gap}}^{\text{part}}) = \frac{f_{\text{gap}}^{\text{part}}(p_{\text{T}}) - f_{\text{gap}}^{\text{part}}(\text{MV1})}{f_{\text{gap}}^{\text{part}}(p_{\text{T}})} \tag{B.1}$$

B.2   JES Bug Fix

This analysis included a small bug fix in the jet energy calibration tools used in data reconstruction, which previously underestimated the JES in data by up to 1.5% for anti-k_T ($\Delta R < 0.4$) jets. The following Figures B.4 and B.5 illustrate the small difference the bug fix made in the gap fraction from data in the rapidity regions ($< 0.47\%$) and Figures B.6 and B.7 in the invariant mass veto regions ($< 1.35\%$, predominantly in the highest mass region).
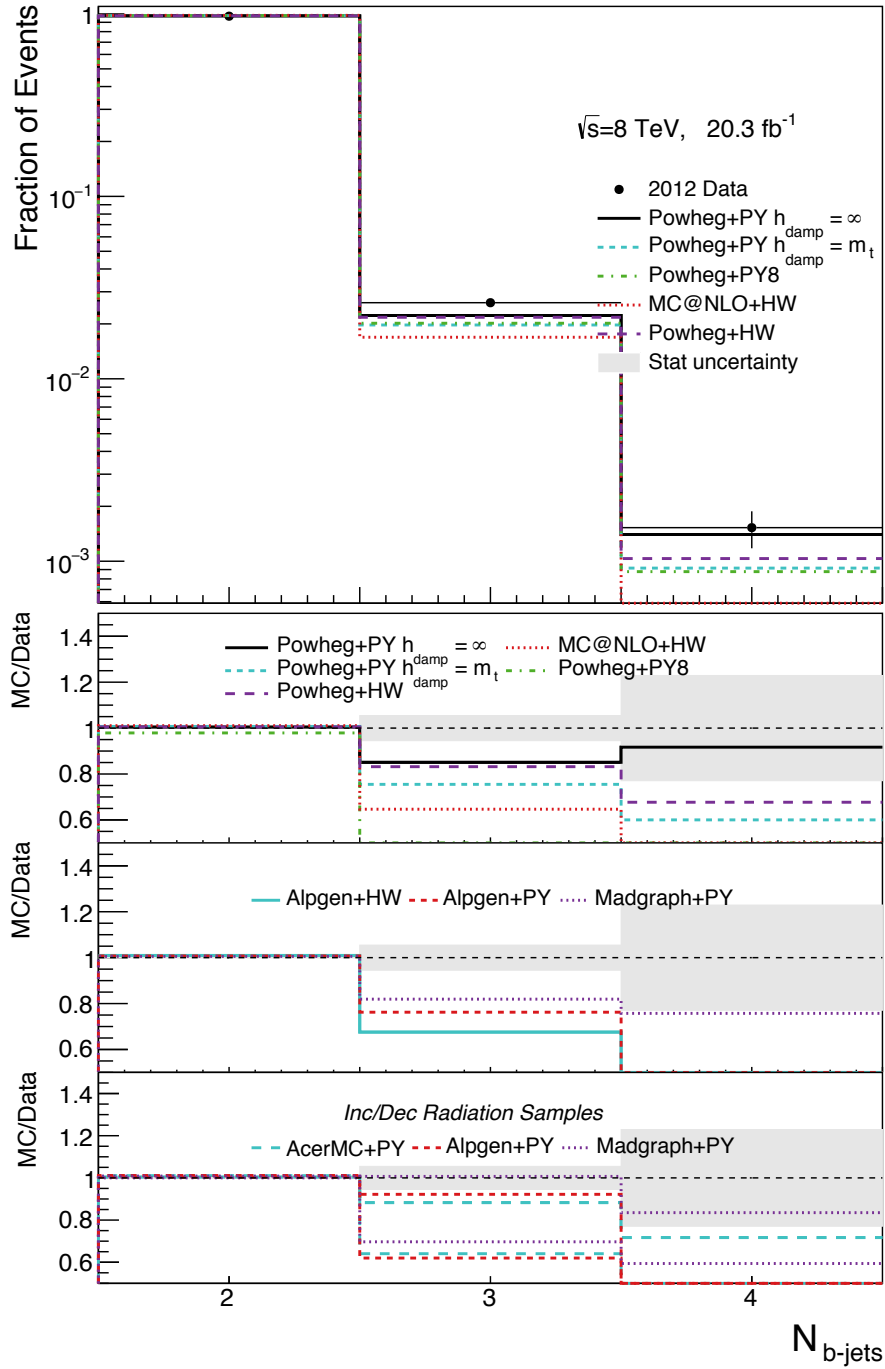
Figure B.1: The number of *b*-jets in *eμbb* events. Less than 3% of the events have 3 or more *b*-jets, requiring the choice of p_T or MV1 ordering to select the *b*-jets used in event selection.
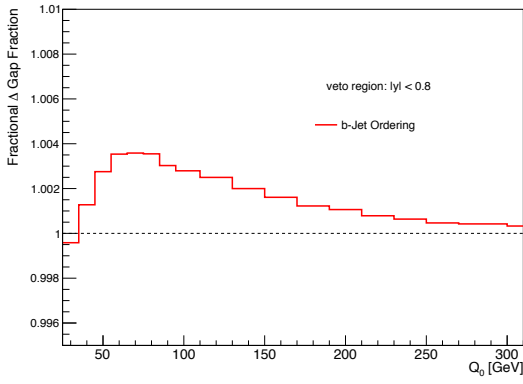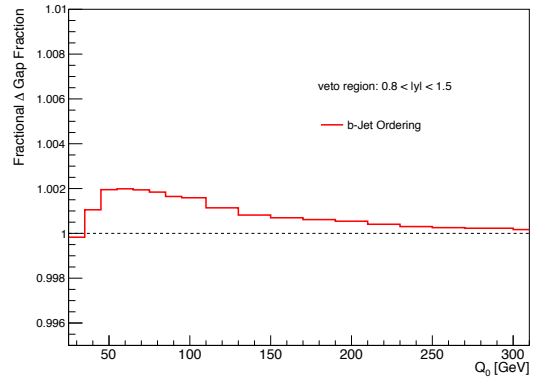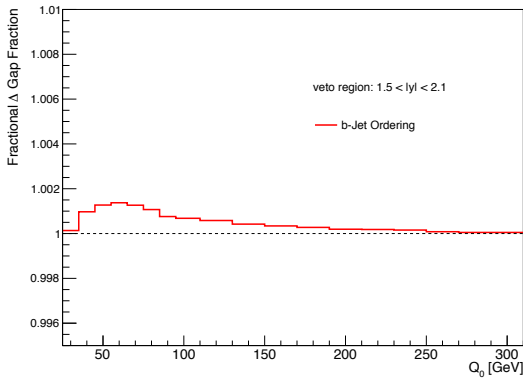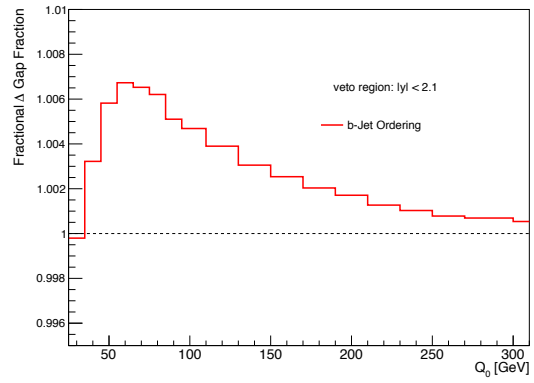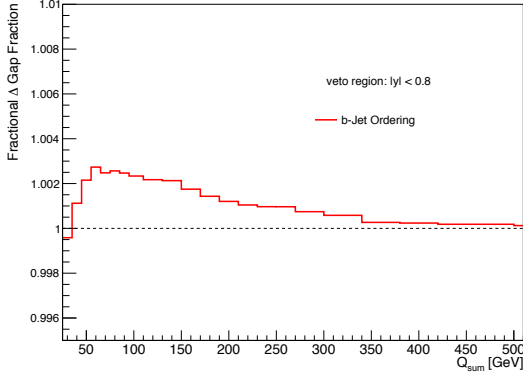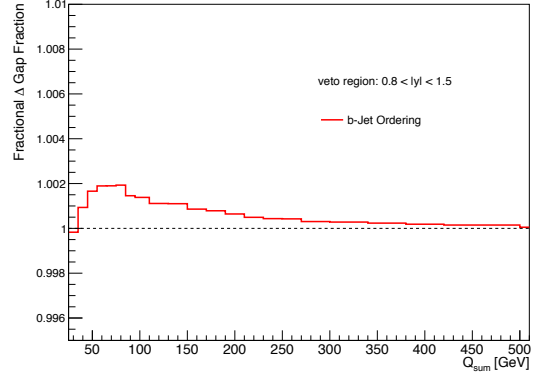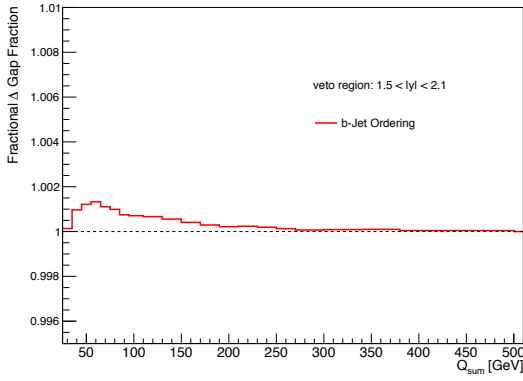
Figure B.2: Difference in the gap fraction, from data corrected to particle level, as a function of Q_0, when b-jets are selected as coming from the $t\bar{t}$ decay, by p_{T} or MV1 value.
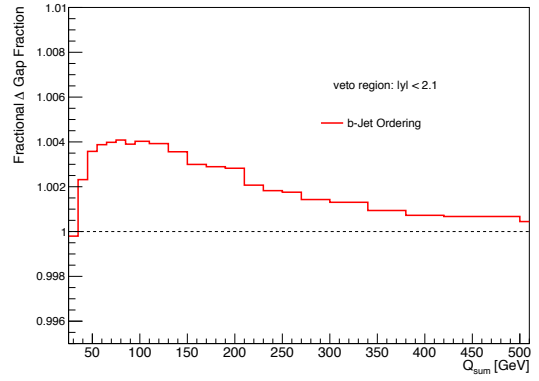
Figure B.3: Difference in the gap fraction, from data corrected to particle level, as a function of Q_{sum}, when b-jets are selected as coming from the $t\bar{t}$ decay, by p_{T} or MV1 value.
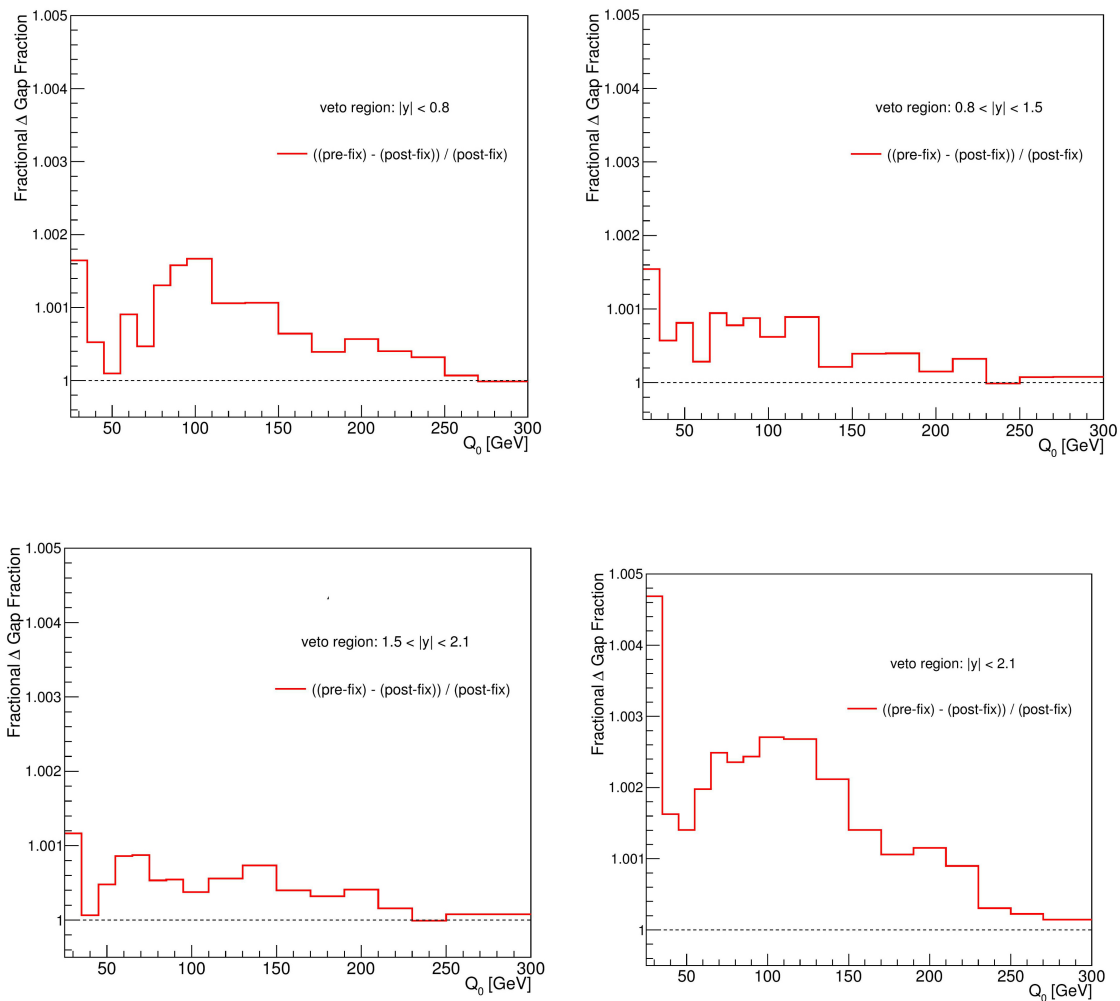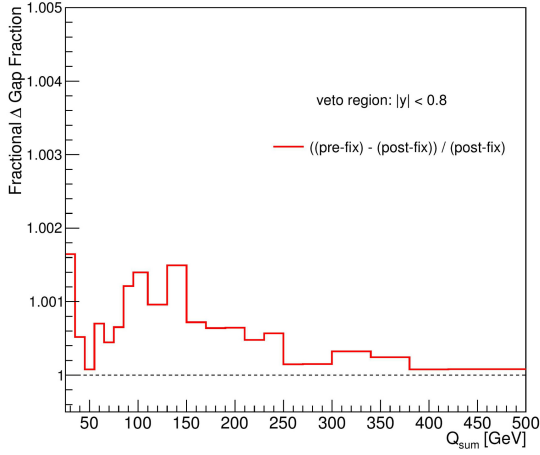
Figure B.4: Difference in the gap fraction, from data at reconstruction level, before and after the bug fix in the jet energy scale. Shown as a function of Q_0 for the four rapidity regions: a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.
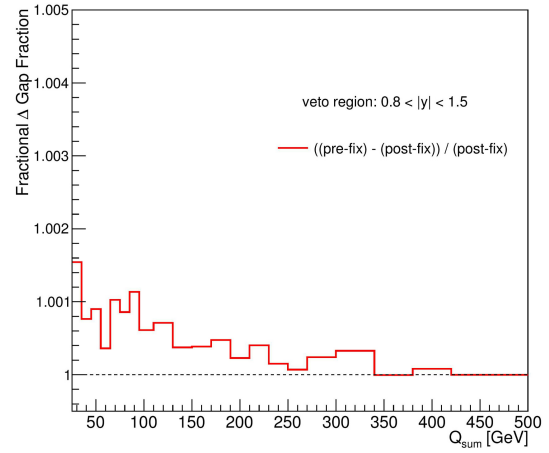
(a)

(b)

(c)

(d)

Figure B.5: Difference in the gap fraction, from data at reconstruction level, before and after the bug fix in the jet energy scale. Shown as a function of Q_{sum} for the four rapidity regions: a) $|y| < 0.8$, b) $0.8 < |y| < 1.5$, c) $1.5 < |y| < 2.1$, and d) $|y| < 2.1$.
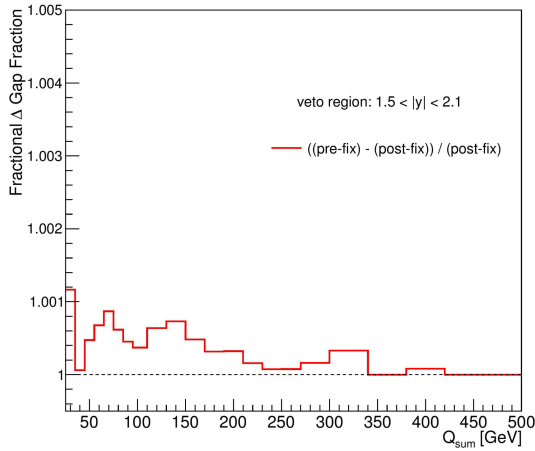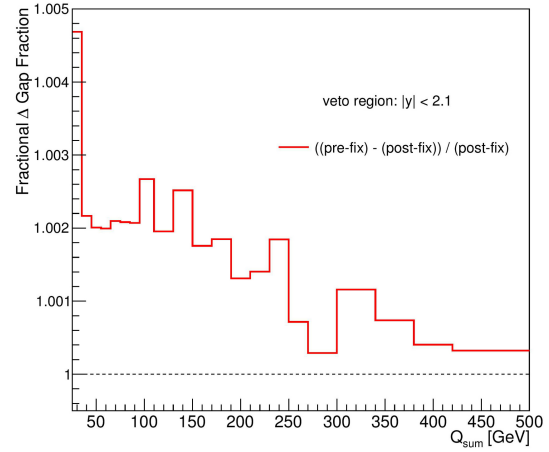
Figure B.6: Difference in the gap fraction, from data at reconstruction level, before and after the bug fix in the jet energy scale. Shown as a function of Q_0 for the four invariant mass regions: a) $m_{e\mu bb} < 300$ GeV, b) $300 < m_{e\mu bb} < 425$ GeV, c) $425 < m_{e\mu bb} < 600$ GeV, and d) $m_{e\mu bb} > 600$ GeV.
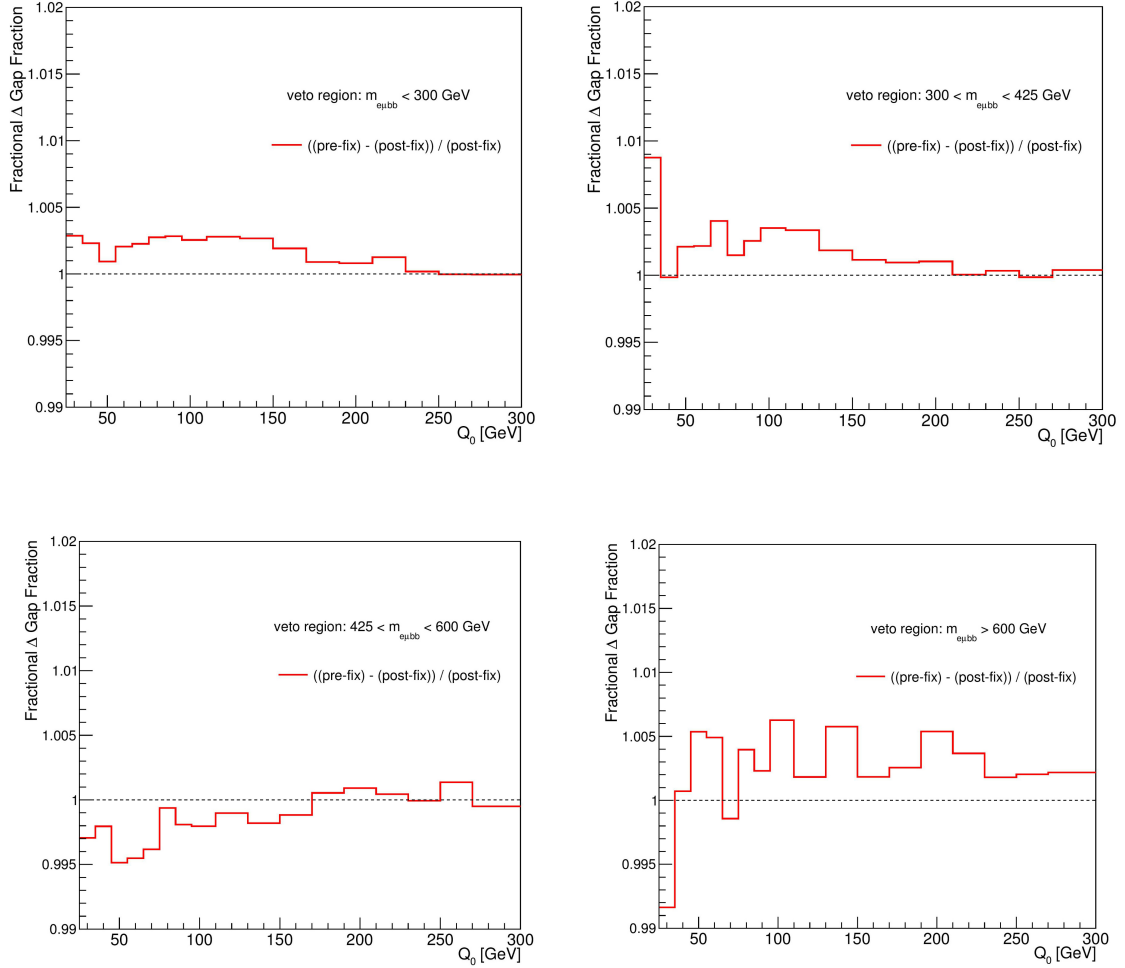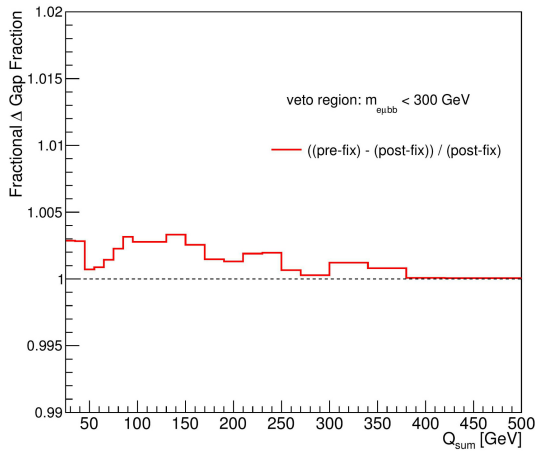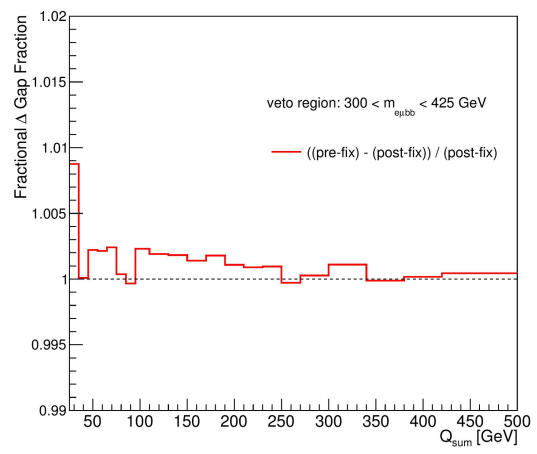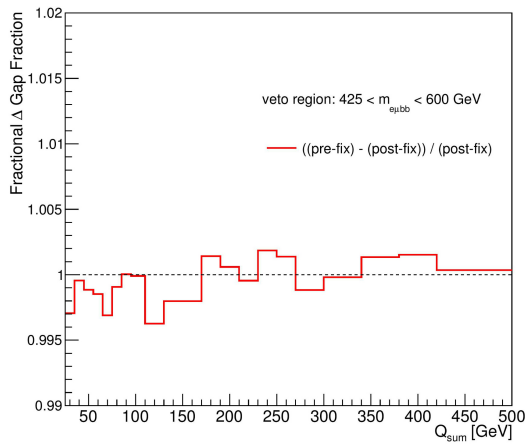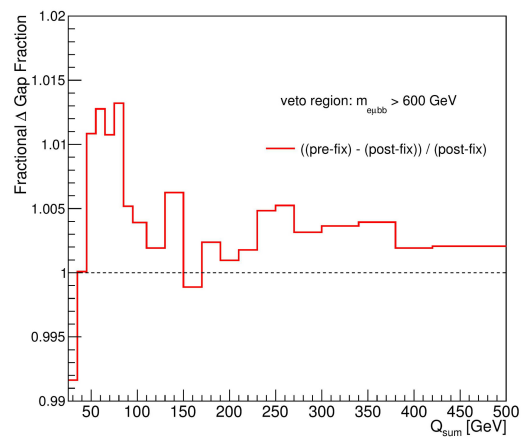
Figure B.7: Difference in the gap fraction, from data at reconstruction level, before and after the bug fix in the jet energy scale. Shown as a function of Q_{sum} for the four invariant mass regions: a) $m_{e\mu bb} < 300$ GeV, b) $300 < m_{e\mu bb} < 425$ GeV, c) $425 < m_{e\mu bb} < 600$ GeV, and d) $m_{e\mu bb} > 600$ GeV.

B.3 Data Set References

The following is an exhaustive list of the Monte Carlo simulation sample data set (DS) containers used in the gap fraction analysis, corresponding to the definitions given in Tables 3.1 and 3.2. The mass extraction analysis used corresponding data sets within the `NTupTop` D3PD framework.

<div align="center">

W+Jet Samples (FullSim)

</div>

Data Set	Data Set Container
117680	mc12_8TeV.117680.AlpgenPythia_P2011C_WenuNp0.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
117681	mc12_8TeV.117681.AlpgenPythia_P2011C_WenuNp1.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
117682	mc12_8TeV.117682.AlpgenPythia_P2011C_WenuNp2.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
117683	mc12_8TeV.117683.AlpgenPythia_P2011C_WenuNp3.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
117684	mc12_8TeV.117684.AlpgenPythia_P2011C_WenuNp4.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
117685	mc12_8TeV.117685.AlpgenPythia_P2011C_WenuNp5.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
117690	mc12_8TeV.117690.AlpgenPythia_P2011C_WmunuNp0.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
117691	mc12_8TeV.117691.AlpgenPythia_P2011C_WmunuNp1.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
117692	mc12_8TeV.117692.AlpgenPythia_P2011C_WmunuNp2.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
117693	mc12_8TeV.117693.AlpgenPythia_P2011C_WmunuNp3.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
117694	mc12_8TeV.117694.AlpgenPythia_P2011C_WmunuNp4.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
117695	mc12_8TeV.117695.AlpgenPythia_P2011C_WmunuNp5.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
117700	mc12_8TeV.117700.AlpgenPythia_P2011C_WtaunuNp0.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
117701	mc12_8TeV.117701.AlpgenPythia_P2011C_WtaunuNp1.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
117702	mc12_8TeV.117702.AlpgenPythia_P2011C_WtaunuNp2.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
117703	mc12_8TeV.117703.AlpgenPythia_P2011C_WtaunuNp3.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
117704	mc12_8TeV.117704.AlpgenPythia_P2011C_WtaunuNp4.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
117705	mc12_8TeV.117705.AlpgenPythia_P2011C_WtaunuNp5.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
110801	mc12_8TeV.110801.AlpgenPythia_P2011C_WbbNp0.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
110802	mc12_8TeV.110802.AlpgenPythia_P2011C_WbbNp1.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
110803	mc12_8TeV.110803.AlpgenPythia_P2011C_WbbNp2.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
110804	mc12_8TeV.110804.AlpgenPythia_P2011C_WbbNp3.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
126606	mc12_8TeV.126606.AlpgenPythia_P2011C_WccNp0.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
126607	mc12_8TeV.126607.AlpgenPythia_P2011C_WccNp1.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
126608	mc12_8TeV.126608.AlpgenPythia_P2011C_WccNp2.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
126609	mc12_8TeV.126609.AlpgenPythia_P2011C_WccNp3.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
126601	mc12_8TeV.126601.AlpgenPythia_P2011C_WcNp0.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
126602	mc12_8TeV.126602.AlpgenPythia_P2011C_WcNp1.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
126603	mc12_8TeV.126603.AlpgenPythia_P2011C_WcNp2.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
126604	mc12_8TeV.126604.AlpgenPythia_P2011C_WcNp3.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/
126605	mc12_8TeV.126605.AlpgenPythia_P2011C_WcNp4.merge.NTUP_COMMON.e1477_s1499_s1504_r3658_r3549_p1562/

Z+Jet Samples (FullSim)

Data Set	Data Set Container
117670	mc12_8TeV.117670.AlpgenPythia_P2011C_ZtautauNp0.merge.NTUP_COMMON.e1711_s1581_s1586_r3658_r3549_p1562/
117671	mc12_8TeV.117671.AlpgenPythia_P2011C_ZtautauNp1.merge.NTUP_COMMON.e1711_s1581_s1586_r3658_r3549_p1562/
117672	mc12_8TeV.117672.AlpgenPythia_P2011C_ZtautauNp2.merge.NTUP_COMMON.e1711_s1581_s1586_r3658_r3549_p1562/
117673	mc12_8TeV.117673.AlpgenPythia_P2011C_ZtautauNp3.merge.NTUP_COMMON.e1711_s1581_s1586_r3658_r3549_p1562/
117674	mc12_8TeV.117674.AlpgenPythia_P2011C_ZtautauNp4.merge.NTUP_COMMON.e1711_s1581_s1586_r3658_r3549_p1562/
117675	mc12_8TeV.117675.AlpgenPythia_P2011C_ZtautauNp5.merge.NTUP_COMMON.e1711_s1581_s1586_r3658_r3549_p1562/
110813	mc12_8TeV.110813.AlpgenPythia_P2011C_ZtautauccNp0.merge.NTUP_COMMON.e1711_s1581_s1586_r3658_r3549_p1562/
110814	mc12_8TeV.110814.AlpgenPythia_P2011C_ZtautauccNp1.merge.NTUP_COMMON.e1711_s1581_s1586_r3658_r3549_p1562/
110815	mc12_8TeV.110815.AlpgenPythia_P2011C_ZtautauccNp2.merge.NTUP_COMMON.e1711_s1581_s1586_r3658_r3549_p1562/
110816	mc12_8TeV.110816.AlpgenPythia_P2011C_ZtautauccNp3.merge.NTUP_COMMON.e1711_s1581_s1586_r3658_r3549_p1562/
110825	mc12_8TeV.110825.AlpgenPythia_P2011C_ZtautaubbNp0.merge.NTUP_COMMON.e1711_s1581_s1586_r3658_r3549_p1562/
110826	mc12_8TeV.110826.AlpgenPythia_P2011C_ZtautaubbNp1.merge.NTUP_COMMON.e1711_s1581_s1586_r3658_r3549_p1562/
110827	mc12_8TeV.110827.AlpgenPythia_P2011C_ZtautaubbNp2.merge.NTUP_COMMON.e1711_s1581_s1586_r3658_r3549_p1562/
110828	mc12_8TeV.110828.AlpgenPythia_P2011C_ZtautaubbNp3.merge.NTUP_COMMON.e1711_s1581_s1586_r3658_r3549_p1562/
146850	mc12_8TeV.146850.AlpgenJimmy_Auto_AUET2CTEQ6L1_ZtautauNp0Excl_Mll10to60.merge.NTUP_COMMON.e1600_s1499_s1504_r3658_r3549_p1562/
146851	mc12_8TeV.146851.AlpgenJimmy_Auto_AUET2CTEQ6L1_ZtautauNp1Excl_Mll10to60.merge.NTUP_COMMON.e1600_s1499_s1504_r3658_r3549_p1562/
146852	mc12_8TeV.146852.AlpgenJimmy_Auto_AUET2CTEQ6L1_ZtautauNp2Excl_Mll10to60.merge.NTUP_COMMON.e1600_s1499_s1504_r3658_r3549_p1562/
146853	mc12_8TeV.146853.AlpgenJimmy_Auto_AUET2CTEQ6L1_ZtautauNp3Excl_Mll10to60.merge.NTUP_COMMON.e1600_s1499_s1504_r3658_r3549_p1562/
146854	mc12_8TeV.146854.AlpgenJimmy_Auto_AUET2CTEQ6L1_ZtautauNp4Excl_Mll10to60.merge.NTUP_COMMON.e1600_s1499_s1504_r3658_r3549_p1562/
146855	mc12_8TeV.146855.AlpgenJimmy_Auto_AUET2CTEQ6L1_ZtautauNp5Excl_Mll10to60.merge.NTUP_COMMON.e1600_s1499_s1504_r3658_r3549_p1562/

Diboson Samples (FullSim)

Data Set	Data Set Container
107100	mc12_8TeV.107100.AlpgenJimmy_AUET2CTEQ6L1_WWlnulnuNp0.merge.NTUP_COMMON.e1596_s1499_s1504_r3658_r3549_p1562/
107101	mc12_8TeV.107101.AlpgenJimmy_AUET2CTEQ6L1_WWlnulnuNp1.merge.NTUP_COMMON.e1596_s1499_s1504_r3658_r3549_p1562/
107102	mc12_8TeV.107102.AlpgenJimmy_AUET2CTEQ6L1_WWlnulnuNp2.merge.NTUP_COMMON.e1596_s1499_s1504_r3658_r3549_p1562/
107103	mc12_8TeV.107103.AlpgenJimmy_AUET2CTEQ6L1_WWlnulnuNp3.merge.NTUP_COMMON.e1596_s1499_s1504_r3658_r3549_p1562/
107104	mc12_8TeV.107104.AlpgenJimmy_AUET2CTEQ6L1_WZincllNp0.merge.NTUP_COMMON.e1596_s1499_s1504_r3658_r3549_p1562/
107105	mc12_8TeV.107105.AlpgenJimmy_AUET2CTEQ6L1_WZincllNp1.merge.NTUP_COMMON.e1596_s1499_s1504_r3658_r3549_p1562/
107106	mc12_8TeV.107106.AlpgenJimmy_AUET2CTEQ6L1_WZincllNp2.merge.NTUP_COMMON.e1596_s1499_s1504_r3658_r3549_p1562/
107107	mc12_8TeV.107107.AlpgenJimmy_AUET2CTEQ6L1_WZincllNp3.merge.NTUP_COMMON.e1596_s1499_s1504_r3658_r3549_p1562/
107108	mc12_8TeV.107108.AlpgenJimmy_AUET2CTEQ6L1_ZZincllNp0.merge.NTUP_COMMON.e1596_s1499_s1504_r3658_r3549_p1562/
107109	mc12_8TeV.107109.AlpgenJimmy_AUET2CTEQ6L1_ZZincllNp1.merge.NTUP_COMMON.e1596_s1499_s1504_r3658_r3549_p1562/
107110	mc12_8TeV.1071010.AlpgenJimmy_AUET2CTEQ6L1_ZZincllNp2.merge.NTUP_COMMON.e1596_s1499_s1504_r3658_r3549_p1562/
107111	mc12_8TeV.1071011.AlpgenJimmy_AUET2CTEQ6L1_ZZincllNp3.merge.NTUP_COMMON.e1596_s1499_s1504_r3658_r3549_p1562/

Single Top Samples (FullSim)

Data Set	Data Set Container
110140	mc12_8TeV.110140.PowhegPythia_P2011C_st_Wtchan_incl_DR.merge.NTUP_COMMON.e1743_s1581_s1586_r3925_r3549_p1575/
110143	mc12_8TeV.110143.PowhegPythia_P2011C_st_Wtchan_dilepton_DS.merge.NTUP_COMMON.e1743_s1581_s1586_r3658_r3549_p1562/
117360	mc12_8TeV.117360.AcerMCPythia_AUET2BCTEQ6L1_singletop_tchan_e.merge.NTUP_COMMON.e1346_s1499_s1504_r3658_r3549_p1562/
117361	mc12_8TeV.117361.AcerMCPythia_AUET2BCTEQ6L1_singletop_tchan_mu.merge.NTUP_COMMON.e1346_s1499_s1504_r3658_r3549_p1562/
117362	mc12_8TeV.117362.AcerMCPythia_AUET2BCTEQ6L1_singletop_tchan_tau.merge.NTUP_COMMON.e1346_s1499_s1504_r3658_r3549_p1562/

$t\bar{t}$ Samples

Data Set	Sim	Data Set Container
117050	Full	mc12_8TeV.117050.PowhegPythia_P2011C_ttbar.merge.NTUP_COMMON.e1728_s1581_s1586_r3658_r3549_p1562/
117050	Fast	mc12_8TeV.117050.PowhegPythia_P2011C_ttbar.merge.NTUP_COMMON.e1727_a188_a171_r3549_p1575/
110404	Full	mc12_8TeV.110404.PowhegPythia_P2011C_ttbar_hdamp172p5_nonallhad.merge.NTUP_COMMON.e3151_s1773_s1776_r4485_r4540_p1575/
110404	Fast	mc12_8TeV.110404.PowhegPythia_P2011C_ttbar_hdamp172p5_nonallhad.merge.NTUP_COMMON.e3151_a220_a205_r4540_p1575/
105200	Full	mc12_8TeV.105200.McAtNloJimmy_CT10_ttbar_LeptonFilter.merge.NTUP_COMMON.e1513_s1499_s1504_r3658_r3549_p1562/
105200	Fast	mc12_8TeV.105200.McAtNloJimmy_CT10_ttbar_LeptonFilter.merge.NTUP_COMMON.e1513_a159_a171_r3549_p1575/
110001	Fast	mc12_8TeV.110001.McAtNloJimmy_CT10_ttbar_dilepton.merge.NTUP_COMMON.e1576_s1499_s1504_r3658_r3549_p1562/
105860	Fast	mc12_8TeV.105860.PowhegJimmy_AUET2CT10_ttbar_LeptonFilter.merge.NTUP_COMMON.e1576_a159_a171_r3549_p1562/
117046	Fast	mc12_8TeV.117046.PowhegPythia8_A14_ttbar_noallhad.merge.NTUP_COMMON.e3615_a220_a263_a264_r4540_p1770/
164440	Full	mc12_8TeV.164440.AlpgenJimmy_AUET2CTEQ6L1_ttbarlnlnNp0_baseline.merge.NTUP_COMMON.e1527_s1499_s1504_r3658_r3549_p1562/
164441	Full	mc12_8TeV.164441.AlpgenJimmy_AUET2CTEQ6L1_ttbarlnlnNp1_baseline.merge.NTUP_COMMON.e1527_s1499_s1504_r3658_r3549_p1562/
164442	Full	mc12_8TeV.164442.AlpgenJimmy_AUET2CTEQ6L1_ttbarlnlnNp2_baseline.merge.NTUP_COMMON.e1527_s1499_s1504_r3658_r3549_p1562/
164443	Full	mc12_8TeV.164443.AlpgenJimmy_AUET2CTEQ6L1_ttbarlnlnNp3_baseline.merge.NTUP_COMMON.e1527_s1499_s1504_r3658_r3549_p1562/
116108	Fast	mc12_8TeV.116108.AlpgenJimmy_AUET2CTEQ6L1_ttbbincl.merge.NTUP_COMMON.e1533_a159_a171_r3549_p1562/
116109	Fast	mc12_8TeV.116109.AlpgenJimmy_AUET2CTEQ6L1_ttbbincl.merge.NTUP_COMMON.e1533_a159_a171_r3549_p1562/
201020	Full	mc12_8TeV.201020.AlpgenAutoPythia_P2012_ttbar_lnlnNp0.merge.NTUP_COMMON.e2356_s1581_s1586_r3925_r4540_p1575/
201021	Full	mc12_8TeV.201021.AlpgenAutoPythia_P2012_ttbar_lnlnNp1.merge.NTUP_COMMON.e2356_s1581_s1586_r3925_r4540_p1575/
201022	Full	mc12_8TeV.201022.AlpgenAutoPythia_P2012_ttbar_lnlnNp2.merge.NTUP_COMMON.e2356_s1581_s1586_r3925_r4540_p1575/
201023	Full	mc12_8TeV.201023.AlpgenAutoPythia_P2012_ttbar_lnlnNp3.merge.NTUP_COMMON.e2356_s1581_s1586_r3925_r4540_p1575/
201024	Full	mc12_8TeV.201024.AlpgenAutoPythia_P2012_ttbar_lnlnNp4.merge.NTUP_COMMON.e2356_s1581_s1586_r3925_r4540_p1575/

$t\bar{t}$ Samples, continued

Data Set	Sim	Data Set Container
110872	Fast	mc12_8TeV.110872.MadGraphCT10Pythia_CT10_P2011C_ttbar_dilepton.merge.NTUP_COMMON .e2609_a188_a222_r3549_p1575/
110875	Fast	mc12_8TeV.110875.MadgraphPythia_CT10_P2012radLo_ttbarQ2up_dilepton.merge.NTUP_COMMON .e3293_a220_a205_r4540_p1770/
110878	Fast	mc12_8TeV.110878.MadgraphPythia_CT10_P2012radHi_ttbarQ2down_dilepton.merge.NTUP_COMMON .e3293_a220_a205_r4540_p1770/
117209	Fast	mc12_8TeV.117209.AcerMCPythia_AUET2BCTEQ6L1_MorePS_ttbar_noallhad.merge.NTUP_COMMON .e1378_a159_a171_r3549_p1562/
117210	Fast	mc12_8TeV.117210.AcerMCPythia_AUET2BCTEQ6L1_LessPS_ttbar_noallhad.merge.NTUP_COMMON .e1378_a159_a171_r3549_p1562/
201030	Fast	mc12_8TeV.201030.AlpgenAutoPythia_P2012_radHi_ttbar_ktfac0p5_lnlnNp0.merge.NTUP_COMMON .e2356_a188_a222_r4540_p1575/
201031	Fast	mc12_8TeV.201031.AlpgenAutoPythia_P2012_radHi_ttbar_ktfac0p5_lnlnNp1.merge.NTUP_COMMON .e2356_a188_a222_r4540_p1575/
201032	Fast	mc12_8TeV.201032.AlpgenAutoPythia_P2012_radHi_ttbar_ktfac0p5_lnlnNp2.merge.NTUP_COMMON .e2356_a188_a222_r4540_p1575/
201033	Fast	mc12_8TeV.201033.AlpgenAutoPythia_P2012_radHi_ttbar_ktfac0p5_lnlnNp3.merge.NTUP_COMMON .e2356_a188_a222_r4540_p1575/
201034	Fast	mc12_8TeV.201034.AlpgenAutoPythia_P2012_radHi_ttbar_ktfac0p5_lnlnNp4.merge.NTUP_COMMON .e2356_a188_a222_r4540_p1575/
201040	Fast	mc12_8TeV.201040.AlpgenAutoPythia_P2012_radLo_ttbar_ktfac2p0_lnlnNp0.merge.NTUP_COMMON .e2356_a188_a222_r4540_p1575/
201041	Fast	mc12_8TeV.201041.AlpgenAutoPythia_P2012_radLo_ttbar_ktfac2p0_lnlnNp1.merge.NTUP_COMMON .e2356_a188_a222_r4540_p1575/
201042	Fast	mc12_8TeV.201042.AlpgenAutoPythia_P2012_radLo_ttbar_ktfac2p0_lnlnNp2.merge.NTUP_COMMON .e2356_a188_a222_r4540_p1575/
201043	Fast	mc12_8TeV.201043.AlpgenAutoPythia_P2012_radLo_ttbar_ktfac2p0_lnlnNp3.merge.NTUP_COMMON .e2356_a188_a222_r4540_p1575/
201044	Fast	mc12_8TeV.201044.AlpgenAutoPythia_P2012_radLo_ttbar_ktfac2p0_lnlnNp4.merge.NTUP_COMMON .e2356_a188_a222_r4540_p1575/