W/Z+Jets Production Cross Section Ratio as a New Physics Search with the ATLAS Detector at CERN

by

James D. Pearce
B.Sc., McGill University, 2009

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Physics

W/Z+Jets Production Cross Section Ratio as a New Physics Search with the ATLAS Detector at CERN

by

James D. Pearce
B.Sc., McGill University, 2009

Supervisory Committee

Dr. Robert Kowalewski, Supervisor
(Department of Physics)

Dr. Richard Keeler, Departmental Member
(Department of Physics)

Dr. Colin Bradley, Outside Member
(Department of Mechanical Engineering)

Supervisory Committee

Dr. Robert Kowalewski, Supervisor
(Department of Physics)

Dr. Richard Keeler, Departmental Member
(Department of Physics)

Dr. Colin Bradley, Outside Member
(Department of Mechanical Engineering)

ABSTRACT

One of the dominant backgrounds in new physics searches at the Large Hadron Collider comes from the leptonic decays of Standard Model W and Z bosons recoiling off jets associated with the underlying event. The ratio of the W+jets and Z+jets cross sections, R_n, is predicted with high precision due to the similar masses and production mechanisms of the W and Z bosons. Any significant departures of R_n from predicted values would be an indication of new physics. This thesis studies a strategy to enhance the sensitivity of R_n to a specific type of signal. A measurement of the ratio R_n is presented, and its sensitivity to pair production of top quarks and leptoquarks is studied. Using a set of topology-discriminating variables, based upon calorimeter topoclusters, the sensitivity of R_n to top quark and leptoquark signals is enhanced using multivariate analysis techniques.

Table of Contents

List of Tables

List of Figures

ACKNOWLEDGEMENTS

I would like to give a special thanks to my supervisor, Bob Kowalewski, for his mentoring, patience and wisdom. I am greatly appreciative of his close involvement in this thesis and my continued studies. I would also like to thank Mathieu Plamondon, for making the time to help troubleshoot my buggy code and for showing me around CERN. And, of course, all of my fellow graduate students for their friendship, support, and always being around to offer a helping hand.

"After sleeping through a hundred million centuries we have finally opened our eyes on a sumptuous planet, sparkling with color, bountiful with life. Within decades we must close our eyes again. Isn't it a noble, an enlightened way of spending our brief time in the sun, to work at understanding the universe and how we have come to wake up in it?"

–Richard Dawkins

DEDICATION

This thesis is dedicated to my family. I love you all.

Chapter 1

Introduction

1.1 The Standard Model

The Standard Model (SM) of particles physics embodies our current understanding of all known elementary particles and their interactions. The SM is a theoretical framework that combines quantum chromodynamics (QCD) and the electroweak model into an internally consistent theory that incorporates the electromagnetic, weak and strong interactions. Over the past few decades the SM has been enormously successful in predicting experimental results. It has successfully predicted the existence of the weak neutral current, charm and top quarks as well as the W and Z bosons [1]. Additionally the consistency between theory and experiments tests radiative corrections and renormalization theory. Indeed, when combined with general relativity the SM accounts for almost all natural phenomena observed. The only remaining untested prediction of the SM is the cause of electroweak symmetry breaking; that is the proposed Higgs mechanism that gives mass to all fermions as well as the W and Z bosons has yet to be confirmed. However, despite all of the SM's triumphs, it has a number of known limitations, such as having 20 arbitrary parameters, not correctly accounting for neutrino oscillations, lacking of a viable dark matter candidate and failing to incorporate a quantum theory of gravitation. The SM's unprecedented accuracy juxtaposed to its known limitations implies that it may be part of, or incorporated into, a more comprehensive theory. Many theories *beyond* the Standard Model (BSM) have been developed to address these flaws and omissions. The main program of experimental high energy particle physics is to develop and conduct experiments that test both the SM and BSM theories.

Particle Type	Name	Label	Spin	Charge	Mass
Quarks	down	d	$\frac{1}{2}$	$-\frac{1}{3}$	3.5-6.0 MeV
	up	u		$\frac{2}{3}$	1.5-3.3 MeV
	strange	s		$-\frac{1}{3}$	104^{+26}_{-34} MeV
	charm	c		$\frac{2}{3}$	$1.27^{+0.07}_{-0.11}$ MeV
	bottom	b		$-\frac{1}{3}$	$4.20^{+0.17}_{-0.07}$GeV
	top	t		$\frac{2}{3}$	171.2 ± 2.1 GeV
Leptons	electron	e	$\frac{1}{2}$	-1	511 keV
	e-neutrino	ν_e		0	< 2 eV
	muon	μ		-1	106 MeV
	μ-neutrino	ν_μ		0	< 2 eV
	tau	τ		-1	1.77 GeV
	τ-neutrino	ν_τ		0	< 2 eV
Gauge Bosons	gluons	g	1	0	0
	photon	γ		0	0
	W-boson	$\mathrm{W}^{\pm}$		± 1	80.4 GeV
	Z-boson	Z^0		0	91.2 GeV
Scalar Boson	Higgs Boson	H^0	0	0	> 114 GeV

Table 1.1: Elementary particles of the Standard Model. Masses taken from [2]

1.1.1 Matter particles and force mediators

The fundamental particles that furnish the Standard Model are distinguished by the symmetries they observe. The most familiar presentation of the SM is given in Table 1.1 where particles are listed by their mass eigenstates – the eigenvalue of which is a readily measured observable. In addition to mass, particles are often identified by their quantum numbers that correspond to internal symmetries, such as electric charge or spin. For example, quarks and gluons carry a color charge, which is analogous to electric charge but with three distinct charges that are associated with the strong interaction. Quarks and leptons also carry weak isospin, with two distinct charges which are conserved in weak interactions. In addition particles may be distinguished by their space-time symmetries, e.g. each of the listed particles is also associated with an antiparticle partner, which carries the same quantum numbers, but have opposite charge(s) unless neutral.

In the mass basis there are 12 fermions (spin $= \frac{1}{2}$) and 5 bosons (spin $= 0, 1$) for a total of 17 fundamental particles that – with the exception of gravity – compose all known fields and matter in the universe. The gauge bosons are often described

as force carriers as they mediate interactions between particles. Photons mediate the electromagnetic interaction creating both attractive and repulsive forces between all particles that carry electric charge. Similarly the W and Z bosons mediate the weak interaction between particles that carry weak isospin while gluons mediate the strong interaction between particles that carry color charge. While photons are electrically neutral, the W and Z bosons as well as gluons, carry their own respective charges allowing for self-interactions. Fig. 1.1 illustrates how the gauge bosons interact with SM particles. Fermions, on the other hand, compose all known matter. Due to a phenomenon called color confinement quarks are perpetually bound to one another forming composite color-neutral particles called hadrons. The most common hadrons, protons and neutrons, are formed from the quark subset $\{u, d\}$. Exchange of gluons between protons and neutrons keeps nuclei bound together. Electrons then tend to form bound states with nuclei through the electromagnetic interaction through exchange of photons between the electron and nucleus. Thus this small subset of fermions, $\{e, u, d\}$ – three out of the known 12 – form the atoms and all the elements of the periodic table.

The Higgs boson plays a special role in the SM, it couples to particles with varying strength endowing them with a unique mass. Its existence would complete the SM explaining why photons and gluons are massless while the W and Z bosons are so heavy.

1.1.2 Gauge theories

The Standard Model is based upon the generalized theoretical framework of Quantum Field Theory (QFT) in which particles are treated as excitations of quantum oscillators of a corresponding field. Just as in classical field theory one can frame QFT using the Lagrangian formulation guided by the principle of least action. A gauge theory is a type of field theory in which the Lagrangian is invariant under transformations between possibles gauges. These gauge transformations are continuous transformations localized in space-time that together form a Lie group, which is referred to as the *symmetry group* of the theory.

The Standard Model and many of its extensions are gauge theories based on $SU(N)$ symmetries. These are the groups of **S**pecial (determinate equals unity) **U**nitary (each element has an inverse) $N \times N$ matrices. Naturally these $N \times N$ matrices operate on N component vectors or spinors, which are said to belong to the

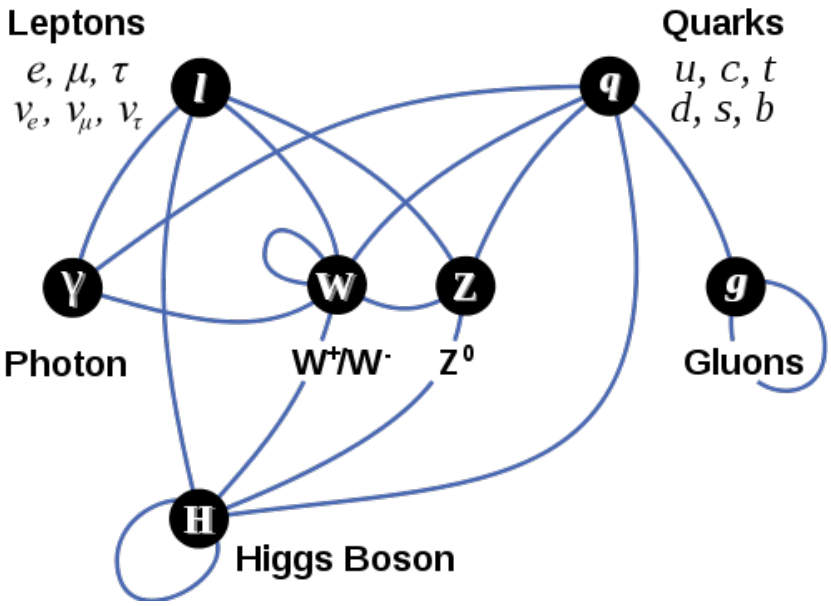

Figure 1.1: Standard Model particles and their interactions. Taken from [3].

group's *fundamental* representation. In this representation the full set of matrices that furnish $SU(N)$ can be constructed from a set of $N^2 - 1$, $N \times N$ traceless hermitian matrices called *generators*. Thus the generators form a basis of $SU(N)$ spanning a $N^2 - 1$ dimensional space in which the group elements live in.

One could argue that the fact that $SU(N)$ gauge theories are local in space-time is the most important aspect of the symmetry. Since all gauge theories are guaranteed to be renormalizable as a consequence of this locality [4]. And Noether's theorem [5] also tells us that there is a conserved charge that is attributed to every continuous symmetry group. This charge is mediated by the gauge fields (or in the quantized theory gauge bosons) allowing them to couple with fermions and spin-0 fields. All modern quantum field theories are based on some symmetry group that exhibits local gauge invariance as it seems to capture, or indeed require, the essence of particle interactions.

In a $SU(N)$ local gauge group the elements can be represented as unitary operators that are functions of space-time, $U(x)$. In gauge theories a spin $= 0$ or $\frac{1}{2}$ field, $\psi(x)$, transforms under the operation of $U(x)$ as [1]

$$\psi(x) \to \psi'(x) = U(x)\psi(x) = \psi(x)e^{i\vec{\gamma}(x)\cdot\vec{T}},$$

where $\vec{\gamma}(x)$ is a vector of arbitrary functions of space-time and $\vec{T}$ are the Lie generators of the symmetry group. In this equation tensor indices are suppressed, however it should be understood that spin-$\frac{1}{2}$ fields are represented as multiplets of Dirac spinors, spin-0 fields as multiplets of scalars and generators as hermitian matrices. Lagrangians that correspond to physical theories require kinetic terms involving the derivative operator ∂^μ. For a typical Lagrangian this term would break the gauge invariance since it operates on $\vec{\alpha}(x)$ in the exponential. To enforce gauge invariance one is required to add additional terms to the Lagrangian to cancel these symmetry breaking terms. It is conventional to absorb these extra terms into a redefinition of ∂^μ called the gauge covariant derivative:

$$\delta_{\alpha\beta}\partial^\mu \to D^\mu_{\alpha\beta} = \delta_{\alpha\beta}\partial^\mu + ig\vec{Q}^\mu \cdot \vec{T}_{\alpha\beta},$$

where α and β are multiplet indices of the fundamental representation, g is an arbitrary gauge coupling and $\vec{Q}^\mu$ are real vector gauge fields; one per Lie generator. These are spin = 0 massless[1] fields that when quantized can be identified with the gauge bosons of the theory. Hence by postulating gauge invariance of the Lagrangian we find the existence of gauge bosons is required. Indeed even gauge boson couplings to fermions are specified by the gauge symmetry.

For example, Quantum Electrodynamics (QED) is a gauge theory based upon a $U(1)$ gauge group, often written as $U(1)_{EM}$ to distinguish it, where the quanta of the vector potential field A^μ are identified as photons. The conserved charge is the familiar electric charge, which is mediated by the photon. The photon couples to other charged particles creating a conserved current. QED is the simplest example of a gauge theory since it is derived from the trivial Lie group $U(1)$. The SM has a more complex group structure it is the group product of the $SU(3)$ color group and the $SU(2) \times U(1)$ weak isospin and hyperchage group. Hence QCD and the electroweak model are combined to form the SM based on the $SU(3) \times SU(2) \times U(1)$ gauge symmetry.

1.1.3 Quantum Chromodynamics

Quantum Chromodynamics is the modern theory of the strong interaction based on the $SU(3)$ symmetry group. In QCD each of the different quark flavours, $q =$

[1] They are massless since $mQ_\mu Q^\mu$ terms are not gauge invariant since Q_μ transforms as $Q_\mu \to Q_\mu - \partial_\mu\gamma(x)$.

u, d, s, c, b, t, carry an additional quantum number called *color*, which is the conserved charge that results from the gauging of the $SU(3)$ symmetry group. The quark fields transform as a color triplet under the fundamental representation of $SU(3)$, where each quark flavour is described by a three component field q_α with $\alpha = 1, 2, 3$ being the color index. To distinguish the $SU(3)$ color gauge group from other possible $SU(3)$ groups it is labeled with a subscript "C", $SU(3)_C$. Since every particle of the SM has an antiparticle twin with opposite charge there exists an anticolor current as well. A state with one color index and one anticolor index of the same type will be color neutral. This is analogous to electromagnetism where a proton with anti-electric charge (positive) and a electron with electric charge (negative) form a neutral bound state. However in addition to this neutral state one could have the three different color charges combined in a single state which would also be color neutral. It is because of this that the name color was chosen, as when one combines light beams of the three primary colors a colorless white light is produced. This is why the three possible values of α are often labeled as *red, green* and *blue*.

$SU(3)_C$ gauge invariance of the QCD Lagrangian requires the introduction of the covariant derivative acting on the quark fields

$$D^\mu_{\alpha\beta} = \delta_{\alpha\beta}\partial^\mu - \frac{ig_s}{2}\vec{\lambda}_{\alpha\beta} \cdot \vec{G}^\mu,$$

where α and β are color indices and g_s is the strong coupling constant. $\vec{\lambda}$ are the Gell-Mann matrices which form a representation of the $SU(3)$ generators.

There are eight generators of $SU(3)_C$ and hence eight gauge bosons, $\vec{G}^\mu$, associated with QCD. These gauge bosons are the gluons, which carry two color indices, one color and one anti-color. This allows gluons to exchange color between quarks, mixing color indices. It also allows gluons to couple to one another lending to a rich and intricate phenomenology.

1.1.4 Electroweak model and spontaneous symmetry breaking

The Electroweak model is a gauge theory that unifies the electromagnetic and weak interaction based on the $U(1) \times SU(2)$ symmetry group. The conserved quantum numbers are *weak* isospin from gauging $SU(2)$ and hypercharge from gauging $U(1)$. Electric charge is given as a combination of weak isospin and hypercharge, thus uni-

fying the two interactions. Weak isospin current can transmute charged leptons into their associated neutrinos or different flavours of quarks into one another. Clearly to describe such an interaction the mass eigenstate basis is inconvenient. This type of weak interaction invites one to interpret leptons and their associated neutrinos as components of a single field, where say the electron and e-neutrino would transform together as a doublet, analogous to how quarks transform as a color triplet in QCD. A similar treatment needs to be applied to quark fields, however, they transmute between generations requiring quark states to be in a linear combination of one another. In addition $SU(2)$ transformations of the weak interaction are particular about the handedness of the field they operate on. So fermions are introduced into the electroweak model as "left-handed" (L) doublets and "right-handed" (R) singlets, where left and right handed fields are defined as $\psi^L = \frac{1}{2}(1-\gamma_5)\psi$ and $\psi^R = \frac{1}{2}(1+\gamma_5)\psi$ [6]. Therefore the fermions of the electroweak interaction are $\psi_i^L = (l_i, \nu_i)^L, (q_i^u, q_i'^d)^L$ and $\psi_i^R = l_i^R, q_i^R, q_i'^R$ where l labels charged leptons, ν neutrinos, q^u "up-type quarks" $\{u, c, t\}$ and q'^d is a linear combination of down-type $\{s, d, b\}$ quark states. i runs over the three lepton generations and quark flavours. There are no right-handed neutrino states here as in the electroweak model neutrinos are taken to be massless. The $SU(2)$ symmetry only acts on left-handed fields, giving it the subscript "L", $SU(2)_L$, while $U(1)$ acts on both left-handed and right-handed fields with hypercharge denoted $U(1)_Y$. The resulting covariant derivative for the $SU(2)_L \times U(1)_Y$ symmetry group is

$$D^\mu_{\alpha\beta} = \delta_{\alpha\beta}\partial^\mu - ig_Y\delta_{\alpha\beta}YB^\mu - \frac{ig_W}{2}\vec{\sigma}_{\alpha\beta}\cdot\vec{W}^\mu,$$

where g_Y and g_W are coupling constants and Y and $\vec{\sigma}$ are representations of the generators of $U(1)_Y$ and $SU(2)_L$ respectively. $\vec{W}_\mu = (W^-_\mu, W^0_\mu, W^+_\mu)$ and B_μ are the necessary gauge fields that need to be introduced to make the Lagrangian gauge invariant.

As stated earlier gauge bosons need to be massless for gauge invariance. However the W and Z bosons are observed to be massive particles. In fact, due to $SU(2)$ transformations only acting on left-handed states, none of the fermion fields are allowed mass terms in the Lagrangian [2]. This implies that the $SU(2)_L \times U(1)_Y$ symmetry is in fact not obeyed, or at least not at the low energy levels from which we observe nature. This observation is made consistent with the electroweak model by postulating that the $SU(2)_L \times U(1)_Y$ symmetry is *spontaneously broken* at a larger

[2]Since $m\psi\bar{\psi} = m(\bar{\psi}^L\psi^R + \bar{\psi}^R\psi^L)$ is clearly not invariant under $SU(2)_L$

energy scale. The simplest instrument that achieves this is the Higgs mechanism, which is introduced into the SM through a scalar spin-0 field that transforms as a doublet under $SU(2)_L \times U(1)_Y$. This field, Φ, known as the Higgs field, introduces a potential, $V = V(|\Phi|^2)$ into the Lagrangian. If, in the quantized theory, $|\Phi|^2$ has a non-vanishing vacuum expectation value a preferred direction in weak isospin plus hypercharge space is selected breaking the $SU(2)_L \times U(1)_Y$ symmetry to the $U(1)_{EM}$ symmetry of QED. The Higgs boson couples with the W and Z bosons as well as fermions to create mass terms in the Lagrangian. The different couplings between the Higgs field and the fermions determine the mass of the particle. The breaking of $SU(2)_L \times U(1)_Y$ induces a mixing of the $\vec{W}^\mu$ and B^μ gauge fields in their mass eigenstates. The result is one massless electrically-neutral gauge field, A_μ made from the linear combination $W^0_\mu \sin\theta_W + B_\mu \cos\theta_W$, one massive electrically-neutral gauge field Z^0_μ made from the linear combination $W^0_\mu \cos\theta_W - B_\mu \sin\theta_W$ and two massive electrically-charged gauge fields W^+_μ and W^-_μ. The quanta of these four gauge fields correspond to the observed gauge bosons of the SM. Strictly speaking the SM Higgs boson does not couple to neutrinos (or any left handed particles) and therefore cannot explain observed neutrino oscillations, as this requires neutrinos to have non-zero mass eigenstates. However minimal extensions to the SM, such as adding right handed neutrinos (Dirac mass) or combining left-handed neutrino with their complex conjugate (Majorana mass), allow the Higgs boson to couple to neutrinos giving them a non-zero mass.

1.2 Structure of hadrons

As stated above hadrons are composite particles composed of quarks held together by color charge which is exchanged between quarks by gluons. Gluons are appropriately named as they hold quarks in color-charge-neutral, or *colorless*, bound states. For example, color interactions between three quarks will form a color singlet baryon bound state by contraction of the anti-symmetric tensor $\epsilon_{\alpha\beta\gamma} q^\alpha q^\beta q^\gamma$ [7]. This state is colorless since the anti-symmetric tensor ϵ ensures that all three indices are different, resulting in a color-neutral state. The other possible color singlet sates are the antibaryon state, $\epsilon^{\alpha\beta\gamma} \bar{q_\alpha} \bar{q_\beta} \bar{q_\gamma}$, and the quark-antiquark meson state, $q^\alpha \bar{q_\alpha}$. Since gluons carry color charge themselves they can couple to one another, theoretically creating color-neutral bound states called "glueballs" – however such states have yet to be confirmed in nature. The quarks that make up color neutral configurations of

baryon and meson states are called *valence* quarks. In addition to valence quarks a fluctuating sea of virtual gluons and neutral $q\bar{q}$ pairs engulf the valence quarks within a hadron. These virtual partons are often ignored as they do not affect the quantum numbers of the hadron. However, in high energy collisions it is possible to scatter valance quarks off of virtual partons.

This sea of virtual gluons also plays a role in the unexpected strength of the strong force. Since gluons are massless one may expect that the force required to separate two quarks would scale as the inverse squared force law as with photons. However gluons, unlike photons, carry color charge themselves allowing exchanging gluons to induce a vacuum polarization in the virtual gluon sea surrounding valence quarks. This creates a string of gluons holding quarks together as if connected by a spring. Hence as the distance between quarks increases so does the strong force holding them together. This phenomena is called *color confinement*.

As a result of color confinement the strong force scales linearly with distance. Thus when scattering quarks in a hadron the quarks will resist separating from the hadron. Instead it is energetically favourable for the system to create new quark-antiquark pairs which may split to form a new bound state with the scattered quark creating a new hadron. This allows the string of gluons connecting the scattering quark and the incident quark to be broken into two. With high energy scattering this process can continue where the broken strings of gluons lead to a *jet* of hadrons. This process is often referred to as *fragmentation* or *hadronization*. At hadron colliders it is these jets that are observed rather than quarks directly.

1.2.1 Hadron scattering and parton distribution functions

Here we consider the scattering process of two incident hadrons A and B that produce an elementary particle c ($c =$ quark, lepton or W/Z boson) plus anything else X,

$$A + B \to c + X.$$

This is the process that one may observe at a hadron collider; however usually it is the subprocess of the hadron constituents, that is the partons, that one is interested in studying. Labelling the scattered partons from A and B as a and b respectively this process is

$$a + b \to c + X.$$

The momentum of the individual partons a and b will not be known. In hadron collider experiments all one knows is the momentum of the hadrons being collided and that this must be equal to the sum of the momenta of its parton constituents. However if the parton momentum density distribution, often called the *parton density function* (PDF), in the hadron is known one can integrate over all possible momenta. This leads to the convention of calculating cross sections with a parton's *fractional* momentum $x = p(\text{parton})/p(\text{hadron})$, where only the component of momentum along the beam axis is considered. The PDF of parton a of hadron A, $f_{a/A}$, is given as a function of the momentum fraction of a, x_a, and the momentum transfer of the process, Q. The cross section $\sigma(AB \to cX)$ may be obtained by multiplying the subprocess cross section $\sigma(ab \to cX)$ by $\mathrm{d}x_a f_{a/A}(x_a, Q^2)$ and $\mathrm{d}x_b f_{b/B}(x_b, Q^2)$, summing over parton and antipartion types a, b, integrating over x_a and x_b, and then averaging over the colors of a and b [4]. Thus the hadron process cross section is given by

$$\sigma(AB \to cX) = K \sum_{a,b} C_{ab} \int_0^1 \mathrm{d}x_a \int_0^1 \mathrm{d}x_b [f_{a/A}(x_a, Q^2) f_{b/B}(x_b, Q^2) + (A \leftrightarrow B)] \sigma(ab \to cX),$$

where C_{ab} are color averaging factors and K is a constant that may be necessary for perturbative corrections (K-factor).

It is not possible to calculate PDFs perturbatively due to non-perturbative QCD binding effects, instead they must be measured in the laboratory. PDFs for various values of Q^2 are extracted from large datasets from various groups worldwide. Some such datasets and collaborations are:

- **CTEQ** [8], The CTEQ Collaboration;
- **MRST** [9], A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne;
- **GRV** [10], M. Glck, E. Reya, and A. Vogt;
- **GJR** [11], M. Glck, P. Jimenez-Delgado, and E. Reya;
- **NNPDF** [12], the NNPDF Collaboration.

The function used to fit the PDF and the number of free parameters will depend upon the value of Q^2. In general, the total number of free parameters is quite large. For example the CTEQ6.6 PDF from the CTEQ Collaboration uses a total of 22 free parameters [13].

1.3 The Leptoquark as an indicator of beyond the Standard Model physics

The Standard Model, despite all of its success, cannot be the final theory of elementary particles and their interactions. For example, it fails to explain the striking similarities between quarks and leptons, such as the same number of generations, identical spins, and charge quantization in multiples of $e/3$. These similarities motivate BSM theories that predict the existence of leptoquarks (LQ), particles that couple both to leptons and quarks that carry color charge. Leptoquarks are predicted by a number of different theories which can be roughly categorized as follows:

- **Models that seek grand unification** [14] [15]. The $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge structure of the SM could easily arise from spontaneous symmetry breaking of a larger simple gauge structure such as $SU(5)$, $SO(10)$ or E_6 [1]. In these models leptons and quarks are placed in the same multiplets of the group's fundamental representation. Leptoquarks in these theories are associated with gauge bosons that mediate GUT interactions between leptons and quarks. Leptoquarks are also predicted in Pati-Salam unified theories based on the $SU(4)_c \times SU(2)_L \times SU(2)_R$ gauge structure, where leptons are identified as quarks of a fourth color. In such theories leptoquarks are introduced as spin-0 bosons that couple to the fermions. All these grand unifying symmetries would have to be broken at high energies to escape current detection at particle colliders.

- **Models that contain quark and lepton sub-structure** [16]. The similarities between quarks and leptons can alternatively be explained by postulating that both are composed of more fundamental particles often referred to as "preons". Preons are confined within quarks and leptons in an analogous way to how quarks are confined in hadrons through color confinement called *hypercolor* confinement. Such models may be consistent with GUTs in their gauge group structure and thus contain the same types of leptoquarks, the only difference being that they are composite particles.

- **Models of dynamical electroweak symmetry breaking** [17]. The cause of electroweak symmetry breaking has yet to be identified. Possible alternatives to the Higgs mechanism are so called "technicolor" theories, where a new

strong gauge is introduce in analogy to QCD. The conserved charge is called technicolor, which is carried by technifermions. Technicolor acts between technifermions to create bound states called technihadrons. One of the possible technipions can be associated with the Goldstone boson from electroweak symmetry breaking. Thus the dynamics of the technifermions is responsible for the spontaneous symmetry breaking of $SU(2)_L \times U(1)_Y$. In these theories color-triplet technipions are identified as leptoquarks which have Higgs like couplings, meaning they decay preferentially to third and second generations of quarks and leptons.

- **R-parity violating supersymmetry models** [18]. In supersymmetry the Poincaré group is extended to give each boson and fermion of the SM a "superpartner" with a spin differing by $\frac{1}{2}$. A discrete symmetry called R-parity is given to all particles, where SM particles have R=1 while their superpartners are assigned $R = -1$. If R-parity is conserved then superpartners cannot decay into SM particles. However in R-parity violating models such decays are possible allowing scalar quarks (squarks) to have Yukawa couplings to leptons. This squark-lepton interaction is associated with scalar leptoquarks.

Event though leptoquarks are predicted in many different ways – as gauge or scalar bosons, composite particles and technipions – they all share the same decay channels (although the branching fractions may differ among theories), namely $LQ \to lq$. Thus in terms of detection at particle colliders it is not necessary to make distinctions between the different types of leptoquarks. Detection of a leptoquark or a leptoquark pair ($LQL\bar{Q}$) would be a clear indication of BSM physics.

Chapter 2

The W/Z+jets production cross section ratio R_n

2.1 W/Z+jets ratio R_n definition

The cross section ratio R_n of W $\rightarrow \mu\nu + n$ jets over Z $\rightarrow \mu\mu + n$ jets is presented here as a cumulative distribution given as a function of an appropriate kinematic variable threshold k_T

$$R_n = \frac{Br_{W\rightarrow\mu\nu} \cdot \sigma_{W+njets}(k_T > x)}{Br_{Z\rightarrow\mu\mu} \cdot \sigma_{Z+njets}(k_T > x)}, \tag{2.1}$$

where x is a discrete value of k_T. If the same data sets are used in the W+jets and Z+jets analysis the luminosity, along with its associated uncertainty, will cancel in the ratio. In this case R_n reduces to the ratio of the *true* number of W+jets to Z+jets events

$$R_n = \frac{N_{W\rightarrow\mu\nu+njets}(k_T > x)}{N_{Z\rightarrow\mu\mu+njets}(k_T > x)}. \tag{2.2}$$

To find the number of W+jets and Z+jets events produced requires a carful study of detector acceptance and efficiencies in each channel; such a study is beyond the scope of this thesis. Instead in the following analysis R_n is approximated by measuring the *observed* number of events in each channel.

There are a number of choices for k_T, for example k_T could be defined as the sum of the transverse momentum of the jets $k_T = \sum p_T^{jets}$ or the sum of the transverse momentum of all final state objects $k_T = \sum p_T^{jets} + \sum p_T^{lep} + E_T^{miss}$. The exact definition of k_T may depend upon the analysis being conducted.

The advantages of measuring R_n as a function of threshold k_T are the increased

statistics in the lower k_T bins as well as reducing the statistical and systematic uncertainty due to k_T bin migration, which would affect the distribution differential in k_T. Thus, one should interpret R_n as successive cross section ratio measurements in kinematic regions of increasing k_T and decreasing phase space volume. Defining R_n as such – that is, a function of threshold k_T – creates statistical correlation between bins in the distribution of R_n, an effect of which is that large statistical fluctuation in a single bin will affect all preceding bins in a similar way.

One could define R_n just as easily in the electron channel, which would follow a very similar analysis. The motivation for choosing the muon channel in the following analysis is that muons deposit little energy in the calorimeters. This is convenient for the analysis since it makes use of calorimetry based variables designed to measure the energy flow of hadronic activity. An electron, which would deposit all of its energy in the calorimeter, would skew these variables in an undesirable way.

2.2 R_n measurement motivation

The main motivation for measuring R_n, rather than looking at each process individually, is the cancellation of many systematic uncertainties associated with the recoiling jets. These systematic uncertainties come from both theoretical models as well as experimental measurements. To a large extent theoretical uncertainties due to generator choices of renormalization and factorization scale, parton distribution functions, fragmentation and hadronization models will cancel. Also, to at least some degree, experimental uncertainties such as jet and cluster energy scale, jet and cluster resolution, pile-up contribution and luminosity will cancel in the ratio. Thus R_n provides for a more precise measurement than W+jets or Z+jets individually.

Since W+jets and Z+jets are important and often irreducible backgrounds to many new physics searches, R_n itself is sensitive to many of the same new physics signals. R_n is sensitive to certain particle final sates, namely any excess of muons, neutrinos and/or jets above SM predictions. New physics models such as Supersymmetry, Leptoquark and Technicolor models predict such an excess of final state particles [1]. However R_n is not dependent upon any parameters specific to these models such as the parent particle invariant mass. This independence from model specific parameters means a measurement of R_n can be used as a basis for a model-independent new physics search. Although, one should note, the sensitivity of R_n to new physics signals is reduced if the signal contributes events that pass both the

W+jets and Z+jets selection, as cancellation in the ratio can occur.

If such an excess of final state particles were observed it would be seen as a upward or downward deviation in R_n away from SM prediction. This deviation would have to be observed in the higher, unexplored, k_T region to be in agreement with previous measurements of vector boson in association with jets cross section measurements. Such an excess over SM predictions would not in itself tell us much about the type of new physics being observed. To understand the nature of the new physics signal R_n could be examined in a number of discriminating phase spaces based on event topologies. For example, it is known that most supersymmetric models have a more spherical event shape topology in the transverse plane (due to a resulting cascade of decay particles) than most SM processes [19]. Thus Transverse Sphericity and Thrust could be used as a discriminating variables, where making a cut on these variables would increase the sensitivity of R_n to supersymmetric models. Then as more cuts are made on additional discriminating variables the phase space volume in which R_n is defined will shrink and the dependence of the search on a specific model will increase. This can be seen as an evolution from a model-independent search to a model-specific search. As an alternative to a set of cuts on discriminating variables one could use the discriminating variables as inputs to a multivariate method to create a single optimized discriminant specific to a particular model.

2.3 Event topology discriminating variables

This section presents a set of five topology-discriminating variables that have been designed to measure the geometric distribution and energy flow of final state particles from a collision event. These variables, often referred to as *event shape variables*, can be calculated using cells, topoclusters or jets. In searches for signatures with large jet multiplicities such variables can be conveniently calculated using jets. However the signatures that contribute to R_n may have as little as one or two associated jets, in which case such jet-base variables are not well defined. Instead all topology-discriminating variables discussed below are calculated using topoclusters, as described in Section 3.3.2, which roughly correspond to the final state particles from a collision event. To avoid a bias from boosts all of the event shape variables defined below are boost invariant with respect to boosts along the beam axis (longitudinal direction).

2.3.1 Transverse sphericity S_T

Transverse sphericity, S_T, which is also known as circularity, is a measure of the isotropy of the event in the transverse plane. S_T is defined between 0 and 1 inclusive, $0 \leq S_T \leq 1$, where 0 corresponds to a back-to-back, or "pencil-like", di-jet event and 1 corresponds to a completely isotropic event. S_T can be a useful discriminating variable for distinguishing unusual event topologies, especially those that result from a cascade of decaying particles such as in supersymmetric models [19]. Transverse sphericity is defined as

$$S_T = \frac{2\lambda_2}{\lambda_1 + \lambda_2}, \qquad \lambda_1 > \lambda_2$$

where λ_1 and λ_2 are the eigenvalues of the 2×2 sphericity tensor

$$S_{ij} = \sum_{k}^{clusters} p_i^k p_j^k$$

with $i, j = x, y$ and k running over all selected topoclusters in the event.

2.3.2 Maximum transverse Fox-Wolfram moment C_{max}

The Fox-Wolfram moments make up a complete set of rotationally invariant observables that characterize the energy distribution of an event. These moments can be defined in the transverse plane to be made boost invariant. The Transverse Fox-Wolfram moments, C_l, are defined by

$$C_l = \sum_{i,j}^{clusters} \frac{p_T^i p_T^j}{\left(\sum p_T\right)^2} \cos[l(\phi_i - \phi_j)],$$

where in the following analysis l is calculated as an integer from 1 to 10 and i and j run over all topoclusters selected in the event. The C_l presented here are modified slightly from the definitions given in [20]. As C_l are defined here the cosines are weighed by the sum of the transverse momentum, $\sum p_T$, rather than the total energy of the event; which is neither boost invariant nor well modelled by MC.

C_l measures the rotational symmetry of an event in the transverse plane. For a back-to-back di-jet topology C_l will equal 1 for even l and 0 for odd l. For a 3 jet event with a perfect 3-fold rotational symmetry in the the transverse plane C_l will equal 1 for $l = 3n$, where n is an integer, and 0 for all other integer values of l. In

general for a l-fold rotationally symmetric event in the transverse plane $C_{l\times n} = 1$. For values of l that severely break this symmetry $C_{l\times n} \approx 0$.

The maximum Transverse Fox-Wolfram moment, C_{max}, is simply taken to be

$$C_{max} = \max_{l}(C_l).$$

Thus $C_{max} = 1$ if the event exhibits a l-fold rotational symmetry, where here $1 \leq l \leq 10$. In general C_{max} measure how 'close' an event is to having a l-fold rotational symmetry. Events with a large momentum imbalance in the transverse plane will have small values of C_{max}.

2.3.3 Transverse thrust

The transverse thrust axis is defined as the dominant direction of energy flow in ϕ. Transverse thrust then gives a measure of how much of the event, projected into the transverse plane, lies along this axis. The transverse thrust in an event is defined as

$$\text{Thrust} = \frac{\pi}{\pi - 2}\left(1 - \max_{\mathbf{n_T}} \frac{\sum_k |\vec{n_T} \cdot \vec{p_T}^k|}{\sum_k |\vec{p_T}^k|}\right),$$

where k runs over all selected topoclusters, $\vec{n_T}$ is the transverse thrust axis unit vector and $\vec{p_T}$ is the topocluster momentum vector in the transverse plane. Here the transverse thrust has been shifted and scaled to lie between 0 and 1, $0 \leq \text{Thrust} \leq 1$, where 0 corresponds to back-to-back di-jet events and 1 to events that have no defined transverse thrust axis because they are distributed isotropically in the transverse plane. In the limit of a perfectly isotropic or back-to-back event S_T and Thrust will have identical values: 1 and 0 respectively. Thus one can expect a fair amount of correlation between these two variables.

2.3.4 ΔR moment

The ΔR moment measures how distributed the p_T of an event is in ϕ-η space. The ΔR moment is defined by

$$\Delta R = \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \frac{p_T^i p_T^j}{(\sum p_T)^2}\Delta R_{ij},$$

where

$$\Delta R_{ij} = \sqrt{(\phi_i - \phi_j)^2 + (\eta_i - \eta_j)^2}$$

and n is the the number of selected topoclusters in the event. ΔR_{ij} measures the distance between two points in ϕ-η space. For an event where all of the p_T of the event is confined to a single point in ϕ-η space $\Delta R = 0$. For a back-to-back di-jet event with maximal values of η ($|\eta|= 4.5$) ΔR takes on its maximum value $\Delta R = 2.38$. For an event that is uniformly distributed in ϕ-η space $\Delta R \approx 1.83$.

Unlike the previous topology-discriminating variables ΔR is a 3-dimensional variable and is not restricted to the transverse plane. However ΔR is still invariant to longitudinal boosts when the momentum of the clusters is much greater than their mass, $|\mathbf{p}| \gg m$. In this case $\eta \approx y$, where y is the rapidity in the longitudinal direction. Rapidities transform under addition with collinear boosts, i.e. $y \to y + \epsilon$ for a longitudinal boost of ϵ. Hence $\Delta R_{ij} \to \Delta R_{ij}(\eta_i + \epsilon - [\eta_j + \epsilon]) = \Delta R_{ij}$ for $|\mathbf{p}_j| \gg m_j$ and $|\mathbf{p}_i| \gg m_i$ under such boosts.

2.3.5 Jet multiplicity N_{jets}

N_{jets} is simply the jet multiplicity or the number of reconstructed jets in an event. Section 3.3.3 describes the anti-k_T algorithm that is used to calculated jets from topoclusters. N_{jets} is usually interpreted to correspond to the number of partons created in a collision event. However, on the particle level, N_{jets} can be interpreted as the number of high-p_T groups of particles with similar trajectories. In both cases N_{jets} clearly can be used to discriminate topologies, but only in the latter sense can N_{jets} be classified as an event shape variable as defined above.

2.4 Analysis strategy

In this section an outline of the following analysis is presented. The analysis is designed to use the precision measurement R_n as a new physics search. It is demonstrated that the sensitivity of the search to a specific model can be improved upon by combining discriminating variables into an optimized multivariate discriminant. Making a cut on this discriminant is shown to enhance the sensitivity of the search to the specific physics signals.

- Apply preselection to data and Monte Carlo.

- For data only use lumi-blocks that pass the electroweak working group's good run list requirements, ensuring that the relevant ATLAS sub-detectors and systems were in stable operating conditions.
- Use the L1 trigger $\texttt{L1}_{\texttt{MU10}}$ on data and Monte Carlo to only select events with at least one high p_T (> 10 GeV) muon.

- Apply corrections to Monte Carlo simulations.
  - Perform necessary corrections on simulated data so that it reliably reproduces observed data.
  - Use data-driven methods when possible.
- Perform event selection on data and Monte Carlo.
  - Identify the relevant, well reconstructed signal events while rejecting backgrounds and poorly reconstructed events.
  - Reject soft and poorly reconstructed detector objects.
- Obtain a set of topology-discriminating variables.
  - Variables should offer good separation between signal and background.
  - Variables should be largely uncorrelated with one another.
- Use topology-discriminating variables as inputs into a multivariate classifier.
  - Train classifier with simulated data after relevant corrections have been applied.
  - Choose classifier method that offers the most discrimination power.
- Choose k_T such that signal and background are best separated into distinct phase spaces.
- Calculate observable R_n in signal phase space with selected events.
- Optimize a cut on the classifier response such that signal in R_n shows greatest enhancement with respect to background.

Chapter 3

The ATLAS Experiment

3.1 The Large Hadron Collider at CERN

The Large Hadron Collider (LHC) is currently the most powerful particle accelerator in the world. It is designed to collide beams of hadrons – protons or lead ions – at unprecedented energies; to open up a new era of discoveries at the energy and luminosity frontier. The LHC is installed 100 m below the Franco-Swiss border in a 27 km long tunnel that was formerly occupied by the LEP accelerator at the European Organization for Nuclear Research (CERN)[3]. The LHC is designed to accelerate two counter-rotating beams of protons with energies of up to 7 TeV and a peak luminosity of 10^{34} cm^{-2}s^{-1}. Even with the LHC currently operating at half the design energy with an instantaneous luminosity of about 2.5×10^{33} cm^{-2}s^{-1} [21] it continues to set world records.

Protons are accelerated through a succession of smaller accelerators before being injected into the LHC where they are boosted to a terminal velocity of 99.9999991% the speed of light [22]. Acceleration of hadrons in the LHC is achieved through the use of radio frequency accelerator cavities that are tuned to a frequency and field orientation that gives the protons a push forward through each cavity. This accelerating scheme necessitates that the proton beam is broken up into a series of bunches – currently 1380 per beam [21]. Proton bunches are directed and focused around the beam though a series of dipole and quadrupole magnets.

[3]As the story is told, when the name of CERN was to be changed from *Conseil Européen pour la Recherche Nucléaire* to *Organisation Européenne pour la Recherche Nucléaire* the acronym was to become the awkward OERN. However Heisenberg suggested to the former director of CERN, Lew Kowarski, that the acronym could "still be CERN even if the name is not."

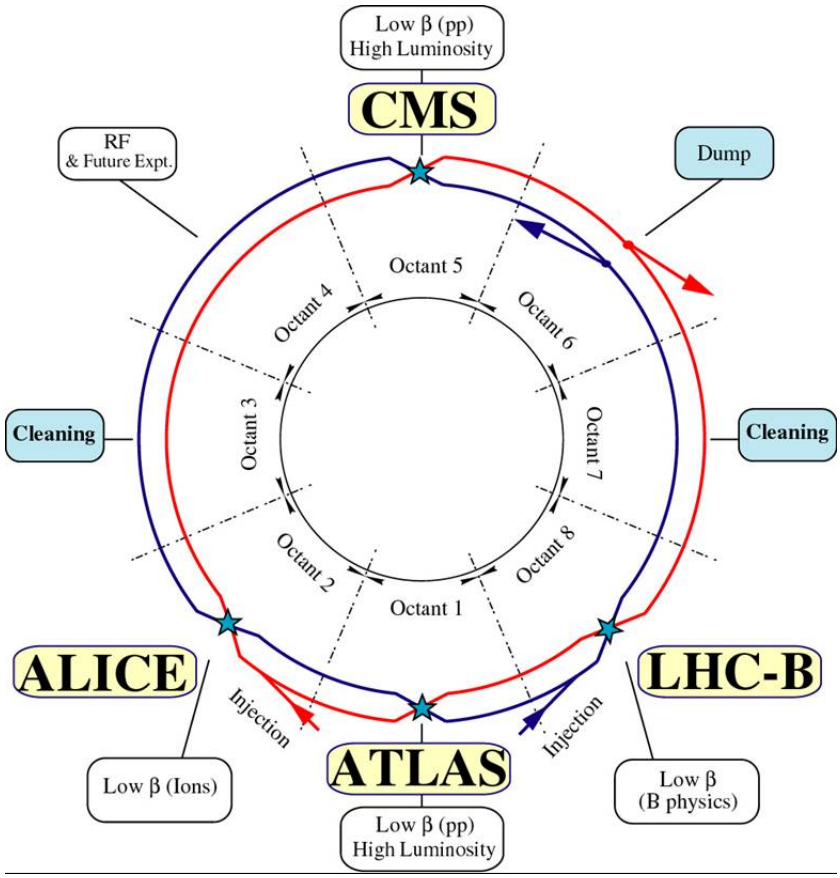

Figure 3.1: Schematic layout of the Large Hadron Collider. Figure from [23]

There are four main interaction points around the ring where the proton beams are squeezed and bunches are directed into one another. Four independent detectors are installed at these interaction points to record the resulting proton-proton collisions as show in Fig. 3.1. ALICE and LHCb are specialized experiments devoted to the study of heavy ion collisions and CP violation respectively. CMS and ATLAS are often referred to as "discovery machines" as they are general purpose experiments designed to be sensitive to a wide variety of known and undetected particles signatures.

3.2 The ATLAS Detector

The unprecedented energy and luminosity of the LHC provides for a rich physics potential of discoveries and precision measurements. With the LHC SM parameters can be measured at world leading accuracies and the discovery reach for new physical

phenomena is unrivalled. The ATLAS (**A T**oroidal **L**HC **A**pparatus**S**) detector is one the two (CMS being the other) general purpose detectors designed to exploit the full discovery potential of the LHC. The benchmark goal of the ATLAS collaboration is to discover the origin of spontaneous symmetry breaking in the electroweak sector of the SM. Since there are a number of possible mechanisms by which the electroweak gauge symmetry may be broken the ATLAS detector needs to be capable of measuring the broadest possible range of signals. For ATLAS to be capable of discerning such a wide range of new and possibly unexpected physics signals certain performance goals must be achieved [24]:

- Excellent calorimetry for electron and photon energy measurements and identification with full coverage for jet and E_T^{miss} reconstruction,
- Good muon momentum resolution, especially for high-p_T muons,
- Efficient particle tracking for high luminosity measurements,
- Full ϕ acceptance and large η coverage for all detector systems,
- High efficiency triggering at low-p_T thresholds.

In order to achieve these requirements, ATLAS is composed of a number of sub-detector systems that operate largely independently of one another. Fig. 3.2 displays an overview of the ATLAS detector with its labeled sub-detectors and components. The main sub-detectors and components of ATLAS can be divided into four systems:

- **Inner Detector** for measuring the trajectories and vertices of charged particles,
- **Calorimeter** for energy measurements and particle identification of electromagnetic and hadronic particles,
- **Muon spectrometer** for measuring the tracks of muons,
- **Magnet system** for bending the trajectories of charged particles providing momentum and charge measurements,
- **Trigger/DAQ** for quickly sorting through events, saving ones that are deemed to contain interesting physics based on a predefined set of selection criteria for offline analysis.

A brief overview of these systems, and their sub-systems, is provided in this section. For a more detailed description of these systems one is referred to [25].

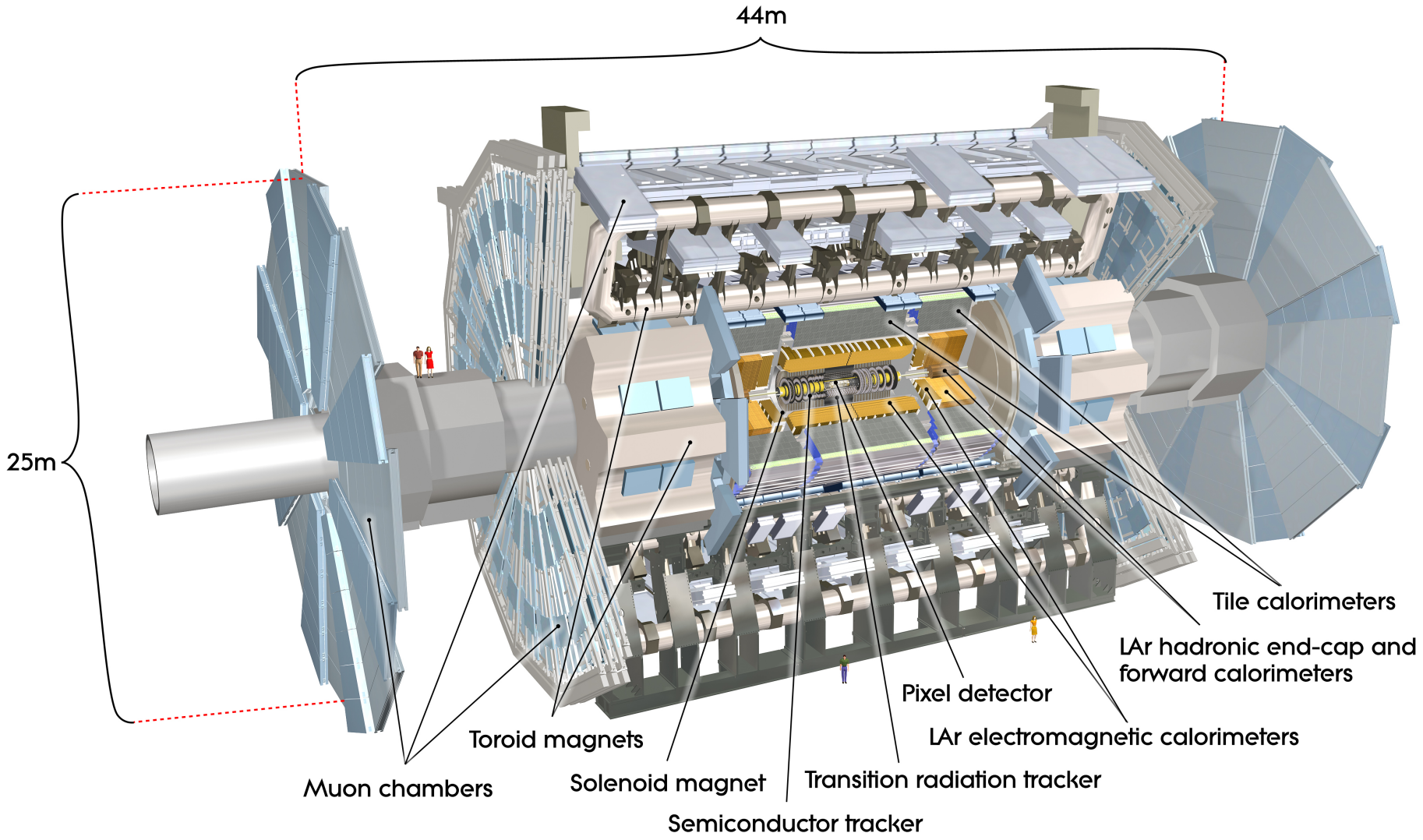

Figure 3.2: Cut-away view of the ATALS detector with its labeled sub-detectors. Image taken from [26]

3.2.1 Detector geometry, coordinate systems and nomenclature

The geometry, coordinate system and nomenclature used to describe the ATLAS detector and reconstructed detector objects is briefly described here. The geometry of the ATLAS detector is cylindrical with the origin defined to be the nominal interaction point (IP), where the counter-rotating proton beams are directed into one another. In Cartesian coordinates the z-axis, also referred to as the longitudinal direction, is defined to lie along the beam axis, while the x-y plane, often referred to as the transverse plane, is normal to the beam axis. The positive x-direction is defined to point towards the centre of the LHC ring from the IP while the positive y-direction points upward towards the surface of the earth. In cylindrical coordinates the azimuthal angle ϕ is defined in the transverse plane, measured around the z-axis while a radius coordinate, R, defines the radial distance from the z-axis. In spherical coordinates the additional angle θ is defined as the polar angle measured from the z-axis.

The longitudinal momentum of scattered particles at hadron colliders has a relatively large associated uncertainty and can vary significantly from event to event. This uncertainty is due to the fact that the initial momentum of the incident partons is unknown and that the ATLAS detector has a limited polar acceptance. The longitudinal rapidity of a particle, defined as $y = 1/2 \ln[(E + p_z)/(E - p_z)]$, is often used at hadron colliders since rapidities are additive under Lorentz boosts; hence the difference between two rapidities is boost invariant. The pseudorapidity, η, approximates rapidity in the massless limit and is defined with only the polar angle θ: $\eta = -\ln\tan(\theta/2)$. For high momentum particles where $m \ll |\mathbf{p}|$ this is a good approximation, and for this reason η has been adopted by ATLAS as the polar coordinate instead of θ. It is also often useful to use the distance ΔR in ϕ-η space between two points (ϕ_1, η_1) and (ϕ_2, η_2) which is defined as $\Delta R = \sqrt{(\eta_1 - \eta_2)^2 + (\phi_1 - \phi_2)^2}$.

The kinematic variables often used to describe particles such as momentum, $\mathbf{p}$, energy, E, and mass, m, are more conveniently defined in the transverse plane due to the aforementioned uncertainties. For example, the *transverse momentum*, $\mathbf{p_T}$, is simply the $\mathbf{p}$ projected into the transverse plane and the transverse energy, E_T, is defined by the projection $E_T = E \sin\theta$. E_T can again be projected along either the x-axis or y-axis in the transverse plane to define the components of the transverse energy vector, $\mathbf{E_T}$. By conservation of energy-momentum the vector sum running

over all scattered particle's $\mathbf{E_T}$ should be zero. However not all particles can be reconstructed with the ATLAS detector (e.g. neutrinos) and thus the measured sum $\mathbf{E_T}$ may not be zero. The magnitude of the sum of measured $\mathbf{E_T}$ is referred to as the *missing transverse energy*, E_T^{miss}, which is associated with scattered particles that escape detection.

3.2.2 Inner Detector

The ATLAS Inner Detector (ID) is designed to measure particle tracks from both primary and secondary vertices with excellent p_T resolution within a pseudorapidity range of $|\eta| < 2.5$. Fig. 3.3 shows a cut-away diagram of the ID with labeled sub-detectors and components. The ID has a cylindrical structure of length 3.51 m and a radius of 1.15 m and is composed of three independent sub-detectors. From the beamline outwards, these detectors are: a high-resolution silicon pixel detector with 3 layers, a silicon microstrip semiconductor tracker (SCT) detector with 4 double layers and a transition radiation tracker (TRT) composed of many layers of straw tubes filled with a Xe-based gas mixture. These three sub-detectors are placed in a central solenoid, which extends over a length of 5.3 m with a radius of 1.25 m and generates a 2 T magnetic field. For a more detailed discussion of the ATLAS ID see [27] [28].

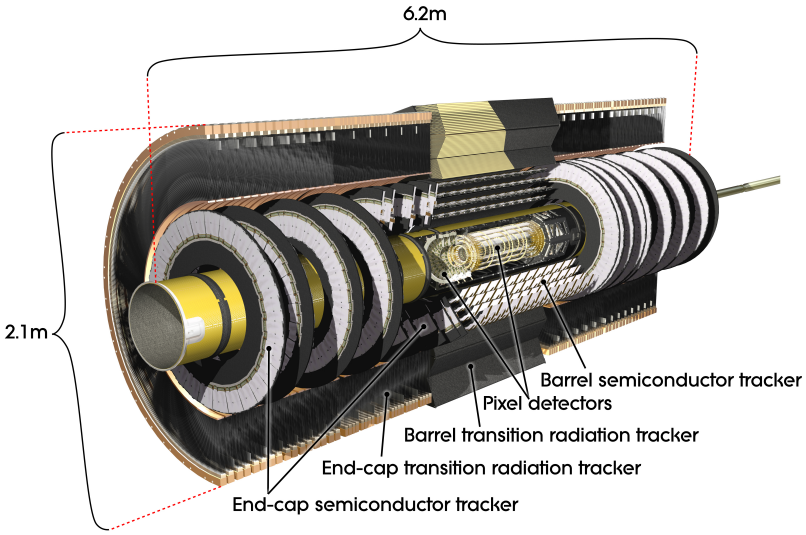

Figure 3.3: Cut-away diagram of the ATLAS Inner Detector with labeled sub-detectors and components. Image taken from [26]

Pixel detector

The Pixel Detector is the innermost sub-detector of the ID. It is composed of 3 "barrel" layers that wrap around the beampipe in concentric cylinders situated with radii, R, of 50.5 mm, 88.5 mm and 122.5 mm from the normal beam position and 3 layers of "end-caps" covering the ends of the barrels situated 494 mm, 580 mm and 650 mm from the collision point $z = 0$. The barrel and end-cap layers are covered by 13-31 million and 2.2 million identical silicon pixel sensors respectively. Each pixel sensor has an individual readout channel – approximately 80.1 million in total. The pixel layers can be segmented into $R\phi$ and z units, where all pixels are identical and have a size in $R\phi \times z$ of 50 $\times$400 μm^2. The intrinsic accuracies for each pixel sensor are 10 μm ($R\phi$) and 115 μm (z) in the barrel and 10 μm ($R\phi$) and 115 μm (R) in the end-caps. A typical track will transverse all three of these layers, leaving a hit in each. The pixel detector provides for the highest granularity around the vertex region to give the most precise measurements of the tracks and vertex positions possible. For a more detailed description of the pixel detector see [29].

SCT

The Semiconductor Tracker is similar to the pixel detector in that they are both made from similar silicon sensors. The SCT is made from pair of single-sided silicon micro-strip sensor connected end-to-end. Two of such pairs glued back-to-back form modules 126mm long. The total number of modules in the SCT is 4088 with approximately 6.3 million readouts. The two layers of silicon strips are designed to be slightly off parallel so that the z-coordinate of a particle transversing both layers can be measured by the slight difference in its R measurement. The intrinsic accuracies of the strips per module are 17 μm ($R\phi$) and 580 μm (z) in the barrel and 17 μm ($R\phi$) and 580 μm (R) in the end-caps. Like the pixel detector the SCT is wrapped around the beampipe in 4 concentric cylinders with 9 end-cap disks at each end. The barrel layers are situated at R-coordinates 284 mm, 355 mm, 427 mm and 498 mm, while the 9 end-cap disks have a $|z|$ position of 854-2720 mm. For a more detailed description of the SCT see [30] [31].

TRT

The Transition Radiation Tracker is the largest of the ID sub-detectors, mounted around the pixel and SCT detectors. The TRT is built from straw tubes of length

144 cm and of diameter 4 mm that run parallel to the beampipe in the barrel region. End-caps are composed of 37 cm long straws that extend radially outward in a disk at the end of the TRT barrel. There are 73 straw planes in the barrel that cover a radial region of 554< R <1082 mm and 160 in the end-caps covering a radial region of 617< R <1106 mm. The TRT can only provide $R\phi$ measurements with an intrinsic accuracy of 130 μm per straw. Typically a particle track transversing the TRT will leave 36 hits. For a more detailed description of the TRT see [32].

3.2.3 Calorimetry

High-energy electrons and photons form a cascade of particles when incident upon dense materials. This cascade, referred to as a *shower*, is the result of pair production of electron-positron pairs from photons as well as bremsstrahlung from charged particles. Such electromagnetic showers are characterized by their radiation length, X_0, and transverse profiles. An analogous type of showering occurs when high-energy hadrons, such as nucleons, pions and kaons, are incident upon dense material. The mechanism involved in *hadronic* showers is partially electromagnetic, since the particles are often charged, but also involves interactions with nuclei via the strong force, where inelastic hadron-nuclear interactions produce particle multiplication. The ATLAS calorimeter exploits particle showers by stopping these particles in a dense material sampling the energy of the shower and hence the incident particle. Additionally, by studying the penetration depth and shower spread the type of incident particle can often be identified.

The ATLAS calorimeter is used to measure the energy of electrons, photons and jets as well as the total missing transverse energy, E_T^{miss}, in an event. It is also one of the central systems used for triggering. The calorimeter can be separated into three parts, the electromagnetic (EM) calorimeter, the hadronic calorimeter and the forward calorimeter, which together cover a range of $|\eta|$ <4.9. Over the pseudorapidity range of $|\eta|$ <2.5, which matches that of the ID, the EM calorimeter has a fine granularity for precision measurements of electrons and photons. The rest of the calorimeter has a courser granularity, but is sufficient for measuring jets and E_T^{miss}.

Electromagnetic Calorimetry

The EM liquid argon (LAr) calorimeter consists of 1.5 mm thick triangular-wave sheets of lead stacked upon one another immersed in a bath of liquid argon. When

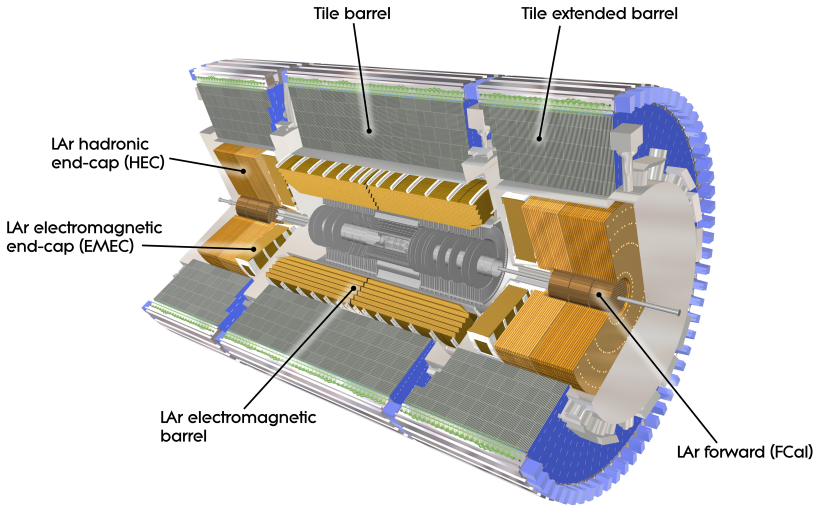

Figure 3.4: Cut-away view of the ATLAS calorimeter system with labeled sub-detectors ad components. Image taken from [26]

high-energy electrons and photons traverse these lead sheets they induce EM showers, converting their kinetic energies into lower energy shower electrons and photons. These shower electrons pass through the lead into 4 mm LAr filled gaps between sheets. When traversing the LAr the electrons knock out valence electrons from the LAr atoms the traversing electron encounters leaving a trail of electron-ion pairs in its wake. An electric field causes the displaced electrons to drift to readout electrode cells placed in the middle of the LAr gap. This motion of drifting electrons creates a current in an external circuit connected to the calorimeter. The number of produced shower electrons is proportional to the energy of the incident particle and thus the measured current from the calorimeter.

The EM calorimeter is divided into a barrel region, covering a pseudorapidity of $|\eta| < 1.475$, and two end-caps covering a pseudorapidity of $1.375 < |\eta| < 3.2$. The triangular-wave shape of the electrodes allows for continuous azimuthal coverage with minimal density variations. The total thickness of the EM calorimeter is greater than 22 X_0 in the barrel region and 24 X_0 in the end-caps – providing good resolution for high-energy jets. Electrons are identified by information from both the ID and shower shape. The high granularity of the EM calorimeter allows for reconstruction of the direction of the shower, allowing for discrimination of electrons and photons from secondary decays. A full description and technical specifications of the EM

calorimeter can be found in [33].

Hadronic Calorimeter

The hadronic tile calorimeter surrounds the EM calorimeter, absorbing the energies of hadrons that escape it. The energy of particles such as protons, neutrons, pions and kaons is transformed into showers of hadrons when passing through steel absorbers that face radially normal to the beam line. These absorbers are separated by scintillating plastic tiles which emit light in an amount proportional to the incident particles. Wave-length shifting fibre readouts that line the edges of the tiles feed these light signals into photomultiplier tubes that convert them into a current, via the photoelectric effect, which is measured through an external circuit connected to the calorimeter.

At large pseudorapidities, radiation from the proton beam becomes increasingly intense. The scintillating tiles can be damaged by excessive radiation exposure. For this reason, the hadronic end-caps are made from a similar sampling material as the EM calorimeters. The main differences being that the lead plates are replaced by 2.5 cm thick copper plates and the argon gap region in between the plates is increased to 8 mm.

The barrel and extended barrel regions of the hadronic calorimeter cover a pseudorapidity of $|\eta| < 1.7$, while the hadronic end-caps cover the barrel ends, which extend over a pseudorapidity of $1.5 < |\eta| < 3.2$. A full description and technical specifications of the hadronic calorimeter can be found in [34].

Forward Calorimeter

The forward calorimeters are designed to absorb and measure the energies of particles with pseudorapidities in the range of $3.1 < |\eta| < 4.9$. In order to absorb intense beam radiation, the forward calorimeters have a LAr active material where the gap between absorption plates is reduced to < 2 mm. The forward calorimeter is split into three 45 cm thick modules: one electromagnetic module and two hadronic modules. For the electromagnetic modules copper is used as the absorbing element while tungsten is used in the two hadronic modules, which is better suited for shorter absorption lengths. A full description of the forward calorimeters can be found in [35].

3.2.4 Muon Spectrometer

Muons are unique in that with sufficient energy they will escape both the EM and hadronic calorimeters since they have relatively large masses (about $\times 200$ the electron) and do not interact via the strong force. Since the power radiated by charged particles due to acceleration orthogonal (synchrotron radiation) and collinear (bremsstrahlung) to the particle's direction of motion scale as m^{-4} and m^{-6}, muons, unlike electrons, lose little energy through this mechanism. Instead high-momentum muons lose energy primarily through ionization. For this reason the muon spectrometer is made up of multiple gas filled chambers in which passing muons leave an ionization trail of electron-ion pairs. The free electrons drift to the closest anode under the influence of an applied electric field. This flow of electrons creates a current in an external circuit connected to the muon spectrometer.

Like the other sub-detectors the muon spectrometer is divided up into a barrel region and end-cap regions. In the barrel region there are three layers of muon chambers consisting of monitor drift tube (MDT) chambers for precision tracking and resistive plate chambers (RPC) for triggering. The muon chamber layers are arranged in the form of a cylinder surrounding the calorimeter covering a pseudorapidity of $|\eta| < 1$. There are also three layers of muon chambers in the end-cap region, which are arranged vertically in disks. The end-caps are made up of MDT chambers as well as thin gap chambers (TGC) which are used for triggering. At large pseudorapidities, with considerably higher particle flux, cathode strip chambers (CSC) are used for tracking, which offer greater radiation tolerances. The combined components of the muon spectrometer cover a pseudorapidity of $|\eta| < 2.7$ with the exception of a 300 mm gap at $\eta = 0$, needed for the passage of services for interior detector systems.

3.2.5 Magnet systems

Charged particles in a magnetic field are subject to the magnetic force component of the Lorentz force, causing them to accelerate in a direction perpendicular to the magnetic field and their direction of motion. For a constant magnetic field charged particles will be pulled from their paths into a helix of a constant radius proportional to their momentum. By inducing a large magnetic field of known field strength and by measuring a particle's radius of curvature, the Lorentz force can be exploited to make a momentum measurement. Additionally, particle charge can be measured by noting the direction of curvature, allowing discrimination between particles and

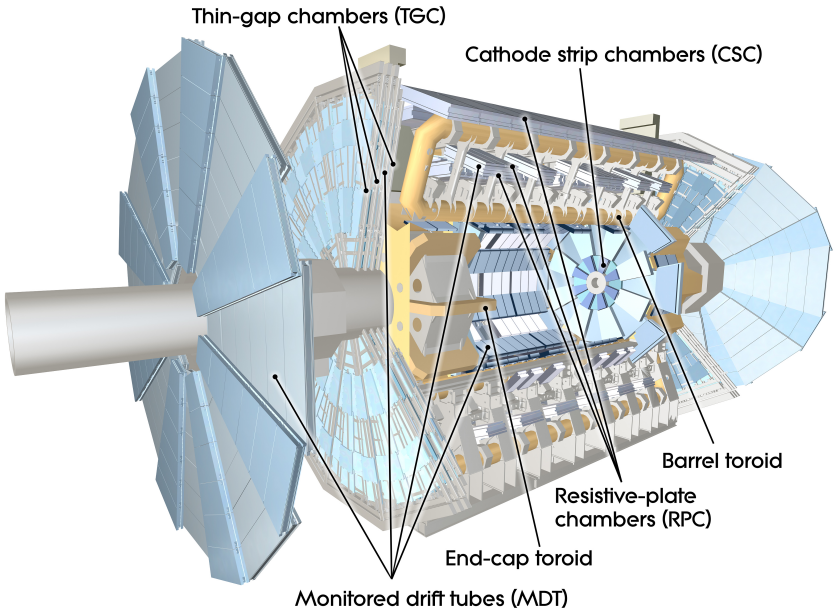

Figure 3.5: Cut-away view of the ATLAS muon spectrometer with labeled sub-detectors ad components. Image taken from [26]

antiparticles.

The ATLAS detector employs two magnet systems to measure the momentum and charge of reconstructed particles: a solenoid surrounding the ID and a system of three large toroids within the muon spectrometer. The central solenoid uses superconducting electromagnets to induce a 2 T magnetic field within the ID. The toroid system is composed of a barrel and two end-caps. There are eight barrel region toroidal coils that are arranged in a cylinder with an 8-fold symmetry around the calorimeter producing a magnetic field of 0.5 T for the barrel muon detectors. End-cap toroids are installed on both sides of the barrel toroid system producing a 1 T magnetic field within the detectors end-cap regions. Combined the magnet system provides a magnetic field of strength greater than 50 mT over a volume of approximately 12000 m^3. More details on the ATLAS magnet systems can be found in [36].

3.2.6 Trigger and data acquisition

Interesting processes produced by the LHC will typically have small cross sections, making them rare to observe. However, the collision rate at the LHC is high enough to produce these types of events in statistically meaningful quantities, but with ex-

tremely large background noise. Fast electronics and software are implemented to quickly sort through events keeping ones that are deemed to contain interesting physics, based on a predefined set of selection criteria, while discarding others.

The ATLAS trigger and data acquisition system (DAQ) is based on three levels of online selection: L1, L2 and the event filter. The trigger applies selection criteria at each level with increasing refinement while the DAQ receives and buffers the event data. Fig. 3.6 gives a conceptual overview of the trigger and DAQ system with labeled trigger levels and bunch crossing rates. The L1 trigger reduces the LHC bunch crossing rate (40 MHz at design luminosity) to 75 kHz by searching for high p_T leptons, photons, jets as well as large missing and total transverse energy. The L1 trigger uses information at reduced granularity from the RPC and TGC to identify high-p_T muons while electrons, photons, jets, τ-leptons, missing and total transverse energy are identified using the calorimeter. The selection criteria in the L1 trigger is often based on a p_T threshold, e.g. muon $p_T > 10$ GeV but can also require any combination of detector objects. The L1 trigger identifies Regions-of-Interests (RoIs) which are physical regions in ϕ-η space where a triggered on object maybe located. The RoI data includes information on the types of features identified and the criteria passed. If the event is selected by the L1 trigger the RoI data is read out from front-end electronics in readout divers (RODs) and then into readout buffers (ROBs) that feed event data to the L2 trigger. The L2 trigger is seeded by the L1 RoI data at full granularity and precision. The L2 triggers apply tighter selection criteria on the event data to reduce the trigger rate to approximately 3.5 kHz. If the event passes the L2 trigger all of the event data is transferred by the DAQ to the event filter – the last stage of event selection. The event filter reduces the event rate to roughly 200 Hz using algorithms similar to those used in offline environments. Events selected by the event filter are moved to permanent storage at the CERN computer centre where a typical event will occupy 1.3 MB of data. More information on the ATLAS trigger and DAQ systems can be found in [37].

3.3 Object reconstruction

In this section a brief outline is given on how detector level objects, such as muons, clusters and jets, are reconstructed. For a more detailed description of object reconstruction see [24].

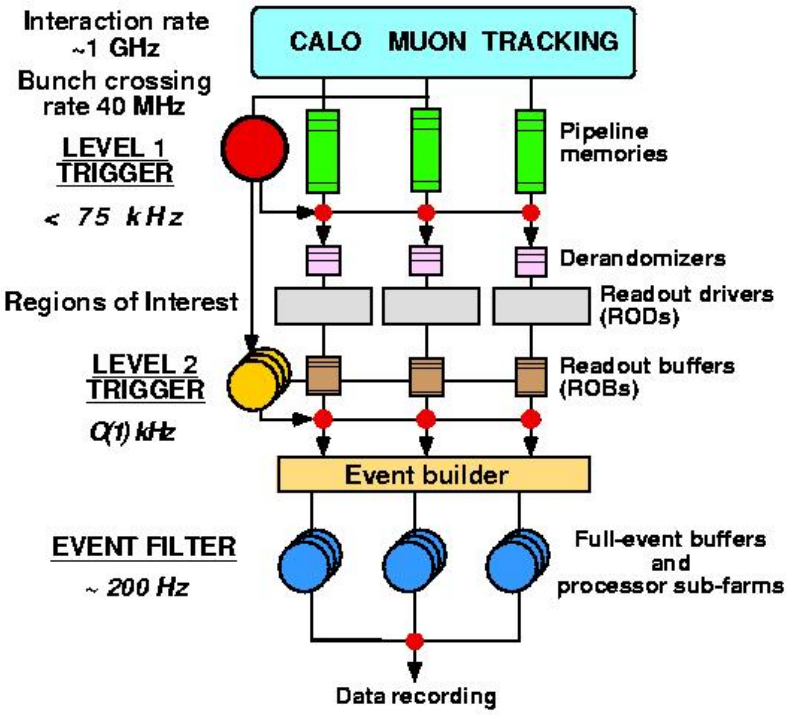

Figure 3.6: Block diagram of the ATLAS trigger and DAQ system.

3.3.1 Muon reconstruction

Muons are reconstructed from a combination of ID and muon spectrometer tracks. The ID provides good measurements of muon tracks at low and intermediate momenta, while the muon spectrometer gives more accurate measurements at momenta over 30 GeV [24]. The central solenoid and toroidal magnets bend the trajectories of transversing muons to allow for momentum measurements. Muon tracks are found by combining track segments, which are defined as straight lines in a single MDT or CSC station. Track candidates are built from segments found in the outer and middle stations of the muon spectrometer and extrapolated back through the magnetic field. When segments are found within the proximity of the track extrapolation they are added to the track candidate. The final track fitting algorithm takes into account the geometry and composition of the material transversed as well as magnetic field inhomogeneities. Using parameters found in the inner stations the track candidates found in the muon spectrometer are propagated back to the interaction point, correcting for energy loss in the calorimeter. ID tracks are matched with these propagated tracks

to form the combined tracks from which muon momenta is determined.

3.3.2 Topological cell clusters

Topological cell clusters attempt to reconstruct particle final states based on their 3-dimensional energy deposits in the calorimeter [24]. The cell clustering algorithm is a successive recombination algorithm, where cells that have an energy over a specific threshold are combined into a *topocluster*, often simply referred to as a cluster, through an iterative process. Cells with an absolute energy threshold above four standard deviations of the total noise, $|E_{cell}| > 4\sigma$, (electronics and pile-up) act as seeds for the clustering procedure. Cells neighbouring these seeds are collected together with the seed cell into a cluster. Then if any of the neighbouring cells have an absolute energy above two standard deviations, $|E_{cell}| > 2\sigma$, they are taken as secondary seeds, where again neighbouring cells are collected and added into the cluster. This process continues until all cluster perimeter cells have a $|E_{cell}| \leq 2\sigma$. At which point collected cell four-momenta are added together in a weighted sum, where the weights are found by calibration on the electromagnetic scale. The final four-momentum defines the final cluster four-momentum.

3.3.3 Jet reconstruction

As discussed in Section 1.2 partons are not observed directly due to color confinement, instead collimated sprays of hadrons called *jets* are observed. Jets are the result of the hadronization and fragmentation process from the hard scattered partons. To make accurate predictions from the parton-level to the hadron-level a well defined jet-finding procedure is essential.

The anti-k_t jet finding algorithm is a common jet finding procedure and is employed in the following analysis to reconstruct jets from topoclusters. The algorithm in its simplest from, i.e. not optimized for computation, can be defined as in [38]:

1. The "k_T distance", $d_{ij} = \min(1/p_{Ti}^2, 1/p_{Tj}^2)\frac{\Delta R_{ij}^2}{R}$, between clusters i and j is calculated along with the distance between cluster i and the beam axis $d_{iB} = 1/p_{Ti}^2$. Here R is a parameter of the anti-k_t algorithm that sets the size of the search cone.

2. The minimum of all d_{ij} and d_{iB} is found. If d_{min} is a d_{ij} then clusters i and j are merged, summing their four-momentum. If d_{iB} is the minimum distance

then the ith cluster is taken to be the reconstructed jet.

3. Steps 1 and 2 are repeated until all topoclusters are exhausted.

The ATLAS experiment has adopted the `FastJet` algorithm, which exploits geometric relations between clusters to reduce computation time.

3.4 Data samples

This section describes the data samples used in the following analysis. All datasets used are in D3PD format. D3PDs are (non-flat) n-tuples optimized for ROOT which contain reconstructed physics objects.

3.4.1 Run periods A-I

The data used for this analysis in both the W+jets and Z+jets channels was collected over the period from March 30, 2010 to October 29, 2010. The data collected during this period corresponds to run periods A through I, which constitutes the 2010 dataset. The total integrated luminosity of the 2010 data sample in the muon channel is 32.6 pb^{-1}. All data recored in this channel is triggered on the `L1_MU10` trigger, which requires a muon p_T threshold of 10 GeV. In a some of the 2010 runs the rate at which events passed this trigger was higher than what the DAQ had been allotted to record. To deal with this in a controlled way a fixed fraction of events that passed the `L1_MU10` trigger were discarded, a practice called prescaling. Thus when calculating the integrated luminosity for the `L1_MU10` trigger one has to take this prescaling into account. In the following analysis this is done by using the ATLAS iLumiCalc tool [39] where the integrated luminosity is corrected for prescaling. All luminosities are scaled by a factor of 0.9626 as recommended by [40]. There is a 3.4% uncertainty associated with the luminosity measurement, however this uncertainty completely cancels in R_n. Only runs where all parts of the detector are in stable operation are used. This is assured by using a common W and Z analysis good run list (GRL) that excludes data taking periods that may contain unstable operation conditions in the relevant detector components. The same GRL has been used in both the W+jets and Z+jets channel to ensure proper cancelation of the luminosity in calculation of R_n.

Run Period	Run Range	Number of Good LBs	Int. Luminosity (pb^{-1})
A-E	152166 – 161948	13357	1.238
F	162347 – 162882	2329	1.677
G	165591 – 166383	2194	5.597
H	166466 – 166964	1057	6.869
I	167575 – 167844	1283	18.20

Table 3.1: ATLAS runs used in analysis, with number of good lumi-blocks and their integrated luminosity

3.4.2 Monte Carlo simulation samples

Events simulated by Monte Carlo (MC) generators are an essential tool for any analysis in high energy physics. MC generators are used to make comparisons between data and predictions, obtain background estimates and efficiencies, understand detector performance and estimate feasibility of future analyses.

The MC data sets used in the following analysis are official, fully simulated, ATLAS datasets generated with a centre of mass energy of 7 TeV. Tables 3.3 and 3.4 list the simulated datasets used, with and without pile-up corrections respectively. The W/Z+jets production in the muon channel (W $\rightarrow \mu\nu$ and Z $\rightarrow \mu\mu$) is modelled using the ALPGEN [41] MC generator. Jet production is generated for up to five partons using the CTEQ6L1 [42] parton distribution function (PDF). The ALPGEN generator is interfaced with HERWIG [43] for showering and fragmentation, whereas JIMMY [44] is used to simulate the underlying event. The PYTHIA [45] event generator – which is equipped to simulate fragmentation and the underlying event – is used to model W$\rightarrow \tau\nu$, Z$\rightarrow \tau\tau$ and QCD multi-jet backgrounds as well as $LQLQ \rightarrow \mu\nu qq$ signatures. PYTHIA utilizes a modified leading-order PDF, MRST LO*. $t\bar{t}$ production is simulated with JIMMY interfaced to HERWIG and combined with next-to-leading-order calculations with MC@NLO [46] using the CTEQ6.6 PDF. The radiation of photons from leptons is simulated with PHOTOS [47] and decay of τ-leptons is handled by TAUOLA [48].

Samples simulated with pile-up in Table 3.3 are generated with minimum bias interactions on top of the hard scattering process. The number of minimum bias interactions follows a poisson distribution with a mean of approximately two [49]. In Section 4.2.2 a reweighing procedure that matches the number of primary vertices in an event between MC and data is discussed. These samples are used when comparing MC to data while samples listed in Table 3.4 are used for purely MC analyses.

Process	Approx.	$\sigma\times$ Br (nb)
$W \to \mu\nu$ + jets	NNLO	10.46
$W \to \tau\nu$ + jets	NNLO	10.46
$Z \to \mu\mu$ + jets	NNLO	1.069
$Z \to \tau\tau$ + jets	NNLO	1.069
$t\bar{t}$	NLO	0.16

Table 3.2: Cross sections estimated to NLO and NNLO approximation, used to scale MC samples listed in Table 3.3 and Table 3.4. Values are taken from [51]

Simulated data is scaled to cross sections listed in sample Tables 3.3 and 3.4, with the exceptions of the processes listed in Table 3.2 where they have been calculated to next-to-leading-order (NLO) and next-to-next-to-leading-order (NNLO).

On top of the simulation of scattering and hadronization it is also necessary to simulate the ATLAS detector's response to the final state particles for comparison of MC to data. Simulation of the passage of particles through detector matter is performed with GEANT4 [50] for all MC samples. The results of this simulation are then passed through a simulation of the trigger and object reconstruction using the same algorithms that are applied to data.

Process	Monte Carlo Sample	$\sigma\times$ Br (nb)	Filter eff.
W$\rightarrow \mu\nu$	mc09_7TeV_107690.AlpgenJimmyWmunuNp0_pt20.merge.AOD.e529_s765_s767_r1430_r1429	6.8711	1
	mc09_7TeV_107691.AlpgenJimmyWmunuNp1_pt20.merge.AOD.e529_s765_s767_r1430_r1429	1.2947	1
	mc09_7TeV_107692.AlpgenJimmyWmunuNp2_pt20.merge.AOD.e529_s765_s767_r1430_r1429	3.7608e-1	1
	mc09_7TeV_107693.AlpgenJimmyWmunuNp3_pt20.merge.AOD.e529_s765_s767_r1430_r1429	1.0072e-1	1
	mc09_7TeV_107694.AlpgenJimmyWmunuNp4_pt20.merge.AOD.e529_s765_s767_r1430_r1429	2.5993e-2	1
	mc09_7TeV_107695.AlpgenJimmyWmunuNp5_pt20.merge.AOD.e529_s765_s767_r1430_r1429	7.1300e-3	1
Z $\rightarrow \mu\mu$	mc09_7TeV_ 107660.AlpgenJimmyZmumuNp0_pt20.merge.AOD.e529_s765_s767_r1430_r1429	0.66379	1
	mc09_7TeV_107661.AlpgenJimmyZmumuNp1_pt20.merge.AOD.e529_s765_s767_r1430_r1429	0.13295	1
	mc09_7TeV_107662.AlpgenJimmyZmumuNp2_pt20.merge.AOD.e529_s765_s767_r1430_r1429	4.0375e-2	1
	mc09_7TeV_107663.AlpgenJimmyZmumuNp3_pt20.merge.AOD.e529_s765_s767_r1430_r1429	1.1161e-2	1
	mc09_7TeV_107664.AlpgenJimmyZmumuNp4_pt20.merge.AOD.e529_s765_s767_r1430_r1429	2.8987e-3	1
	mc09_7TeV_107665.AlpgenJimmyZmumuNp5_pt20.merge.AOD.e529_s765_s767_r1430_r1429	7.5662e-4	1
W $\rightarrow \tau\nu$	mc09_7TeV_107054.PythiaWtaunu_incl.recon.ESD.e514_s765_s767_r1425	8.9295	1
Z $\rightarrow \tau\tau$	mc09_7TeV_106052.PythiaZtautau.merge.AOD.e514_s765_s767_r1425	0.85402	1
WW	mc09_7TeV_105985.WW_Herwig.recon.AOD.e521_s765_s767_r1430	2.9603e-2	0.38842
WZ	mc09_7TeV_105987.WZ_Herwig.recon.AOD.e521_s765_s767_r1430	1.1229e-2	0.30998
ZZ	mc09_7TeV_105986.ZZ_Herwig.recon.AOD.e521_s765_s767_r1430	4.5988e-3	0.21232
QCD multi-jet	mc09_7TeV_109276.J0_pythia_jetjet_1muon.merge.AOD.e534_s765_s767_r1388_r1389	9.8605e6	7.93e-5
	mc09_7TeV_109277.J1_pythia_jetjet_1muon.merge.AOD.e534_s765_s767_r1388_r1389	6.7804e5	1.233e-3
	mc09_7TeV_109278.J2_pythia_jetjet_1muon.merge.AOD.e534_s765_s767_r1388_r1389	4.0968e4	5.4427e-3
	mc09_7TeV_109279.J3_pythia_jetjet_1muon.merge.AOD.e534_s765_s767_r1388_r1389	2.1929e3	1.305e-2
$t\bar{t}$	mc09_7TeV_105200.T1_McAtNlo_Jimmy.merge.AOD.e510_s765_s767_r1430_r1429	0.14412	1

Table 3.3: Monte Carlo simulation data samples with pile-up used in analysis with their production cross sections (taken from AMI) and their generator filter efficiencies.

Process	Monte Carlo Sample	$\sigma\times$ Br (nb)	Filter eff.
$W \to \mu\nu$	mc09_7TeV_107690.AlpgenJimmyWmunuNp0_pt20.merge.AOD.e511_s765_s767_r1302_r1306	6.8711	1
	mc09_7TeV_107691.AlpgenJimmyWmunuNp1_pt20.merge.AOD.e511_s765_s767_r1302_r1306	1.2947	1
	mc09_7TeV_107692.AlpgenJimmyWmunuNp2_pt20.merge.AOD.e511_s765_s767_r1302_r1306	3.7608e-1	1
	mc09_7TeV_107693.AlpgenJimmyWmunuNp3_pt20.merge.AOD.e511_s765_s767_r1302_r1306	1.0072e-1	1
	mc09_7TeV_107694.AlpgenJimmyWmunuNp4_pt20.merge.AOD.e511_s765_s767_r1302_r1306	2.5993e-2	1
	mc09_7TeV_107695.AlpgenJimmyWmunuNp5_pt20.merge.AOD.e511_s765_s767_r1302_r1306	7.1300e-3	1
$Z \to \mu\mu$	mc09_7TeV_107660.AlpgenJimmyZmumuNp0_pt20.merge.AOD.e529_s765_s767_r1302_r1306	0.66379	1
	mc09_7TeV_107661.AlpgenJimmyZmumuNp1_pt20.merge.AOD.e529_s765_s767_r1302_r1306	0.13295	1
	mc09_7TeV_107662.AlpgenJimmyZmumuNp2_pt20.merge.AOD.e529_s765_s767_r1302_r1306	4.0375e-2	1
	mc09_7TeV_107663.AlpgenJimmyZmumuNp3_pt20.merge.AOD.e529_s765_s767_r1302_r1306	1.1161e-2	1
	mc09_7TeV_107665.AlpgenJimmyZmumuNp5_pt20.merge.AOD.e529_s765_s767_r1302_r1306	2.8987e-3	1
	mc09_7TeV_107664.AlpgenJimmyZmumuNp4_pt20.merge.AOD.e529_s765_s767_r1302_r1306	7.5662e-4	1
$W \to \tau\nu$	mc09_7TeV_107054.PythiaWtaunu_incl.merge.AOD.e514_s765_s767_r1302_r1306	8.9295	1
QCD multi-jets	mc09_7TeV_109276.J0_pythia_jetjet_1muon.merge.AOD.e534_s765_s767_r1302_r1306	9.8605e6	7.93e-5
	mc09_7TeV_109277.J1_pythia_jetjet_1muon.merge.AOD.e534_s765_s767_r1302_r1306	6.7804e5	1.233e-3
	mc09_7TeV_109278.J2_pythia_jetjet_1muon.merge.AOD.e534_s765_s767_r1302_r1306	4.0968e4	5.4427e-3
	mc09_7TeV_109279.J3_pythia_jetjet_1muon.merge.AOD.e534_s765_s767_r1302_r1306	2.1929e3	1.305e-2
$t\bar{t}$	mc09_7TeV_105861.TTbar_PowHeg_Pythia.merge.AOD.e521_s765_s767_r1302_r1306	0.14412	1
$LQLQ \to \mu q\nu q$	mc09_7TeV_115179.Pythia_LQLQ_munuqq_600.merge.AOD.e579_s765_s767_r1302_r1306	3.8846e-6	1
	mc09_7TeV_115180.Pythia_LQLQ_munuqq_800.merge.AOD.e579_s765_s767_r1302_r1306	3.7572e-7	1
	mc09_7TeV_115181.Pythia_LQLQ_munuqq_1200.merge.AOD.e579_s765_s767_r1302_r1306	7.0955e-09	1

Table 3.4: Monte Carlo simulation data samples without pile-up used in analysis with their production cross sections (taken from AMI) and their generator filter efficiencies.

Chapter 4

Event selection and analysis

4.1 Preselection

The purpose of preselection is to reject events that do not pass certain data quality requirements. This selection is done at the event level, such that the entire event is rejected if it does not pass preselection. Preselection is the loosest of the offline selections, designed to quickly filter out uninteresting events. The preselection criteria used in the proceeding analysis is described in more detail below.

- **Good run list:** Lumi-blocks must pass certain data quality requirements to be placed on the good run list (GRL). The detector components relevant to the measurement (trigger, calorimetry, muon detector, inner detector, solenoid and toroid magnets) must be in stable operation and the LHC beam must be fully ramped and also in stable operating conditions. This corresponds to an ATLAS data quality tag of "Ready". Lumi-blocks that have been flagged due to anomalously high levels of calorimeter noise are also excluded from the good run list.

- **Primary vertex:** All events are required to have at least one reconstructed primary vertex to reduce the non-collision background. A vertex must have at least 3 reconstructed tracks to be considered a primary vertex.

- **Trigger:** The `L1_MU10` trigger is required for all events. This means that only events with at least one reconstructed muon with $p_T > 10$ GeV are selected.

- **Bad jet cleaning:** Certain types of calorimeter noise, such as noise bursts, are not modelled in Monte Carlo simulations. Lumi-blocks that suffer greatly

from this type of noise are excluded in the GRL. However, rejecting events with "bad" jets, as defined in [52], reduces this background on the event level. Only a small fraction of events are removed by this requirement, 0.05%.

4.2 Monte Carlo correction procedures

There are several quantities relevant to the proceeding analysis that are not properly simulated in MC and need to be corrected before event selection (although not pre-selection) can be performed. Since MC values for these quantities are unreliable a data-driven method is required; i.e. the correction to MC needs to be estimated from data. The need for these corrections arise from: uncertainties in theoretical calculations, quantities that scale with instantaneous luminosity, unknown detector object misclassification rates and detector resolution effects. Inevitably any correction for these effects will add some systematic uncertainty to the final measurement; however, a comprehensive study of these systematic uncertainties is beyond the scope of this thesis. In the case of one jet the systematic uncertainties on R_1 have already been studied in detail [53].

4.2.1 QCD background estimation

The main backgrounds in the $W \to \mu\nu$ channel consist of QCD multi-jets and leptonic decays of vector bosons ($W \to \tau\nu$, $Z \to \mu\mu$ and $Z \to \tau\tau$). At larger jet multiplicities (and higher $\sum p_T^{jet}$) $t\bar{t}$ production dominates the background. For the QCD multi-jet background the main mechanisms that produce a hard muon that passes the muon selection are the decays of heavy flavour mesons ($c\bar{c}$ and $b\bar{b}$) in jets. Smaller contributions to this background include pion and kaon decays and hadrons faking muons. In simulation very few of these events have the required E_T^{miss} to pass the W selection cuts and none of these events pass the Z invariant mass selection cut. However the QCD multi-jet cross section is large enough that this small fraction constitutes a non-negligible background in the W+jets channel.

For the QCD multi-jet background muon-filtered PYTHIA di-jet samples are used as listed in Table 3.3. Since the cross sections quoted for these samples have a large uncertainty and prompt fake muons are not simulated reliably a data-driven method is used to estimate the contribution of the QCD multi-jet background in the W+jets signal region. *Anti*-W selection cuts (E_T^{miss} <25 GeV and m_T <40 GeV)

Jet Multiplicity	0 jets	1 jet	2 jets	3jets	4 jets	≥5 jets
W+Jets note	0.560	0.623	0.519	–	–	–
E_T^{miss}	0.487	0.591	0.565	0.606	0.727	0.873
m_T	0.494	0.592	0.560	0.613	0.845	0.888
Average	0.4905	0.5915	0.5625	0.6095	0.786	0.8805

Table 4.1: Scale factors applied to QCD multi-jet background to estimate contribution in signal regions calculated in each jet multiplicity bin. Scale factors for jet multiplicity bins 0-2 are compared to the scale factors found in the W+jets cross section note.

define a control region where QCD multi-jets are the dominant processes. With this control region QCD multi-jet scaling factors are found by fitting the QCD multi-jet contribution to data while keeping signal and all other backgrounds fixed. The fit is performed by minimizing the χ^2 of the data and MC E_T^{miss} and m_T distributions in 20 bins for jet multiplicities 0-5. The results of the fit are summarized in table 4.1. These scaling factors are compared to those found in the W+jets cross section measurement note [51] with 1.3 pb^{-1} of integrated luminosity for E_T^{miss} distributions. The W+jets note uses a similar data-driven method to determine their scaling factors, however the control region does not make the anti-W selection cuts. Instead two templates are fitted, QCD multi-jet and leptonic backgrounds with signal, using the ROOT method TFractionFitter where the QCD multi-jet scaling factors are extracted from this fit. The two sets of QCD multi-jet scaling factors are consistent considering that the results of the fit depend on the control region chosen and the binning of the distributions to some degree. The fitting method used in this analysis is chosen because it is simpler and is not done with any of the events that are used to calculate and compare the final measurement of R_n to MC predictions.

The average of the scaling factors found for the E_T^{miss} and m_T distributions is used to estimate the QCD background in the W+jets signal region. Fig. 4.1 displays the E_T^{miss} and m_T distributions before and after the QCD multi-jet scaling factors are applied.

4.2.2 Re-weighting pile-up events in Monte Carlo

All MC distributions that are compared to data include in-time pile-up simulation. This means that the expected number of primary vertices in a collision is much closer to the average number of vertices seen in the full 2010 data set. However

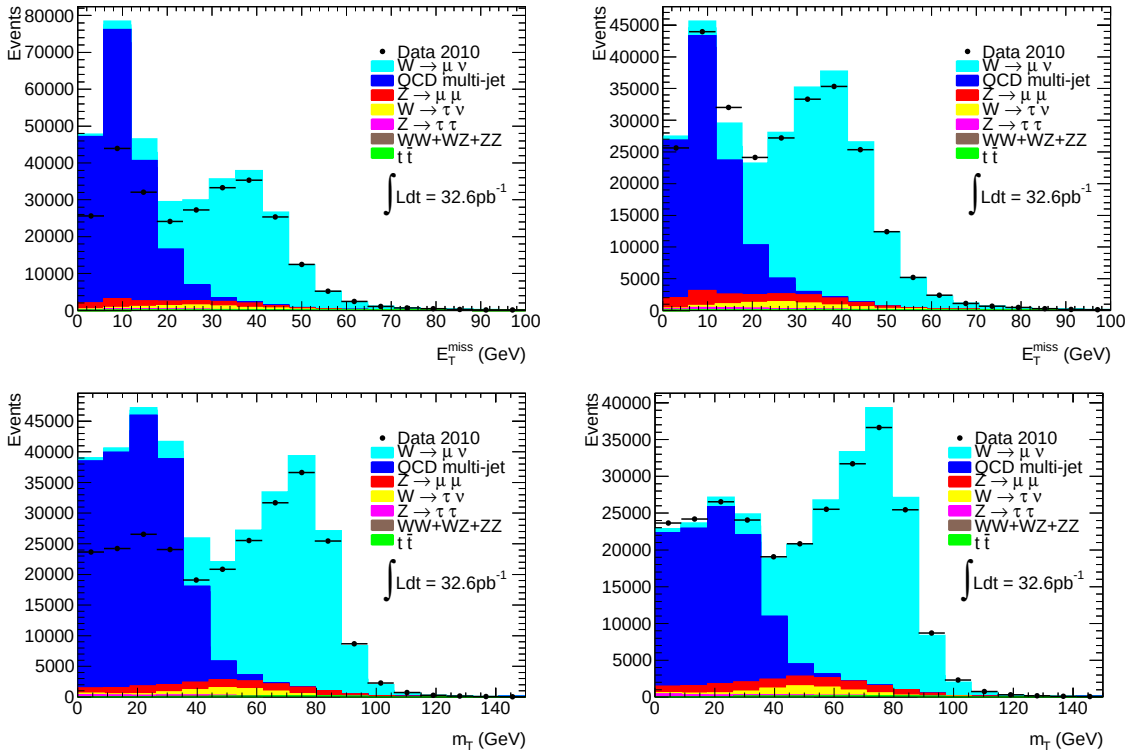

Figure 4.1: E_T^{miss} and m_T distributions before (left) and after (right) applying scaling factors to QCD multi-jet background. All other MC samples are scaled to 32.6 pb^{-1} using the cross sections give in Table 3.3.

Number of vertices	1	2	3	4	5
Vertex weight factor	1.632	1.096	0.8384	0.6904	0.6006
Number of vertices	6	7	8	9	10
Vertex weight factor	0.5213	0.4708	0.3056	0.2265	0.1140

Table 4.2: Vertex weight factors used to match the number of in-time pile-up events between data and MC.

pile-up changes as a function of instantaneous luminosity and can be quite different depending on the run period being considered. For this reason MC samples need to be weighted so that the distributions of the number of primary vertices matches that in data. The procedure used to calculate these vertex weight factors is that recommended by the ATLAS Standard Model Working Group [54].

- Select a data sample from a given stream
- Require that the runs and luminosity blocks in this data sample satisfy the common WZ & EWK good run list requirement
- Require that the event pass a given lepton trigger for a given data taking period
- Apply standard pre-selection to remove potential non-collision background events
- Compute the weights for each data taking period by calculating the primary vertex ratio with data to Monte Carlo distributions
    - a vertex must be reconstructed with at least 3 tracks
- Obtain the final set of weights (one for each number of primary vertices in the events) by taking the luminosity weighted average of the different weights in different data periods

The vertex weight factors are found after preseclection and muon selections, as described in Sections 4.1 and 4.3.1, but not before W, Z or jet selection. Table 4.2 shows the results of this procedure for 2010 run periods A-I. Fig. 4.2 compares the distributions of the number of primary vertices in the W+jets and Z+jets channels before and after the vertex weight factors have been applied. These weights are used for all MC samples in the proceeding analysis whenever MC is being compared to data.

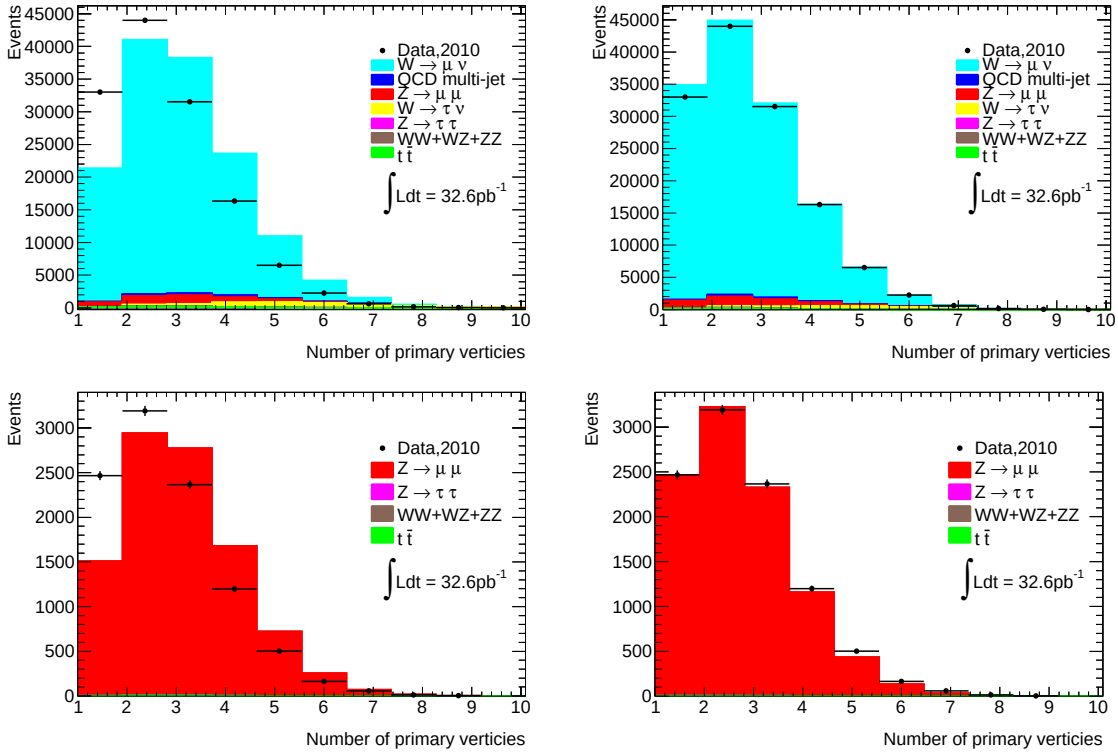

Figure 4.2: Distributions of the number of primary vertices before (left) and after (right) vertex weights are applied in the W $\rightarrow \mu\nu$ channel (top) and Z $\rightarrow \mu\mu$ channel (bottom).

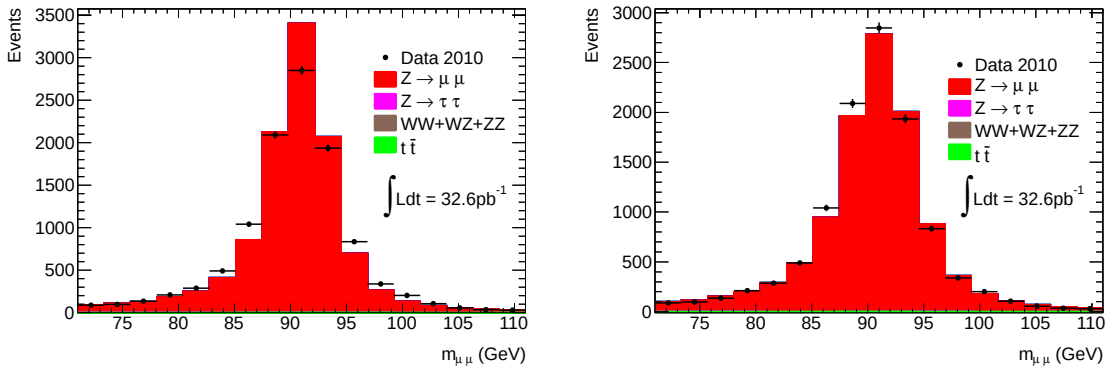

Figure 4.3: Z boson invariant mass distribution before and after invariant mass resolution smearing with jet multiplicities 0-3.

4.2.3 Z invariant mass resolution correction

The MC Z boson invariant mass, m_Z, distribution requires a resolution correction to match data due to limited muon p_T resolution. The correction procedure is often referred to as *smearing* as it has the affect of increasing the width of the distribution that it is being applied to. In the proceeding analysis a simple smearing scheme is adopted which is applied to the m_Z distribution itself, rather than smearing the muon p_T distributions and then reconstructing the Z boson invariant mass. In either case, it is the m_Z distribution that is being used to determine the level of smearing necessary.

The smeared m_Z distribution, $m_Z^{smeared}$, is obtained by applying a linear transformation to the original distribution that matches the standard deviation of the MC σ_{MC}, sample to data, σ_{data}

$$m_Z^{smeared} = am_Z + b, \tag{4.1}$$

where $a = \sigma_{MC}/\sigma_{data}$ and $b = -\mathrm{E}[m_Z](1-a)$. a is a scaling factor which smears the distribution while b translates the distribution back to its original mean. a is calculated in a 20 GeV window centred around the m_Z mean to best match the peaks of the distributions. Since a is calculated in a window the smearing function needs to be applied iteratively because after each iteration a number of the smeared values will fall outside this window. This smearing procedure is applied until the value of a converges between iterations. Fig. 4.3 shows the m_Z distributions before and after the smearing procedure is applied. It is found that best agreement between data and MC is achieved with 24% smearing on the m_Z distribution, which reduces the acceptance of the Z boson selection cuts by 4%.

4.3 W/Z+Jets event selection

The selection process is designed to identify the relevant, well reconstructed signal events while rejecting backgrounds and poorly reconstructed events. Each reconstructed object in the detector (such as a jet, muon or E_T^{miss}) is identified by passing a number of requirements specific to that object. Signal and background processes are simulated with MC. The luminosity of the simulated data is scaled to that of the collected data with the exception of the QCD multi-jet background which is scaled using the method outlined in Section 4.2.1. The efficiency of a selection cut is estimated by making the same selection with simulated data for a specific process. To compare results with Standard Model expectation all selection is done on data as well as MC with the exception of jet cleaning, since jet events that fail this requirement are not modelled in simulation.

4.3.1 Muon selection

Muon selection is performed on all muon candidates passing the trigger. The purpose of these selection requirements is to reject fake or poorly reconstructed muons while preferentially selecting hard muons that result from the decay of a heavy parent particle. These selection criteria are described in more detail below.

- **Combined tracks:** A muon candidate must be reconstructed with the combination of inner detector and muon spectrometer tracks. This reconstruction corresponds to the STACO collection of muons.

- **Primary vertex:** To check the consistency of a muon track with the selected primary vertex the absolute difference between the primary vertex z-coordinate and that of the muon's track extrapolated to the beam line must be less than 10 mm ($|z_\mu - z_{pv}| < 10$ mm).

- **Detector hits:** To ensure that the reconstructed track corresponds to a muon there are requirements on the number of sub-detector readings or "hits". In the inner detector muon tracks are required to have at least two hits in the pixel detector and at least six in the SCT detector. For tracks with $|\eta|$ <2.0 at least one hit in the TRT detector is also required.

- **Spectrometer p_T^μ:** The transverse momentum of the muon candidate in the muon spectrometer, p_T^{MS}, must be greater than 10 GeV to reduce the contribu-

tion of muons arising from π/K decay.

- **Consistent tracks:** To ensure consistency between the inner detector and muon spectrometer tracks the absolute difference between a muon candidate's transverse momentum reconstructed in the muon spectrometer and the inner detector, p_T^{ID}, must be less than $0.5p_T^{ID}$ ($|p_T^{ID} - p_T^{MS}|/p_T^{ID} < 0.5$).

- **Impact parameter:** To reduce non-collision backgrounds muons are required to have an impact parameter consistent with the area of the beam. The absolute value of the impact parameter, d_0, of the muon track relative to the primary vertex must be less than 0.1 mm ($|d_0|$ <0.1mm).

- **Track isolation:** To reduce multi-jet background muon candidate tracks are required to be isolated. This requires that the sum of the transverse momentum of all inner detector tracks within a cone of $\Delta R < 2.0$ of the muon be less than 1.8 GeV.

- **Muon p_T:** In this analysis only muons that originate from heavy parent particles, such as vector bosons, are of interest. Thus the muon's combined (inner detector and muon spectrometer) transverse momentum, p_T^{μ}, is required to be at least 20 GeV ($p_T^{\mu} > 20$ GeV).

- **Pseudorapidity:** For good reconstruction muon candidates are required to be in the detector's fiducial volume. Muon candidates must have an absolute value of pseudorapidity less than 2.4 ($|\eta^{\mu}|$ <2.4).

Table 4.3 lists the number of events in data that pass these selection cuts along with their efficiencies as estimated from MC. Agreement between data and MC for the muon variables $\sum p_T^{\mu}$ and $m_{\mu\mu}$ can be seen in Fig. 4.4. All of the signal and background samples listed in Table 3.3 are included. In both of these distributions a clear peak is seen at approximately 90 GeV corresponding to the Z boson mass.

4.3.2 Jet selection

Jet selection is designed to remove hadronic background and select only well reconstructed jets that come from the hard process. This includes removing jets that could arise from pile-up as well as jets faked by muons. All jets are reconstructed using the anti-k_T algorithm as described in Section 3.3.3. A jet's initial four-momentum is

Selection cut	Data 2010		Signal MC (%)	
	Number of events	Marginal Eff. (%)	Absolute Eff.	Marginal Eff.
Total with GRL	1.78356e8	—	—	—
Primary vertex	1.57287e8	88.18	99.76	99.76
Trigger	4.62437e7	29.40	63.14	63.29
Combined muons	2.33025e7	50.39	61.30	96.86
$\lvert z_\mu - z_{pv}\rvert < 10$ mm	2.27113e7	97.46	61.02	99.54
Detector hits	2.15095e7	94.71	60.41	99.00
$p_T^{MS} > 10$ GeV	7.47567e6	34.76	58.98	97.63
$\lvert p_T^{ID} - p_T^{MS}\rvert / p_T^{ID} < 0.5$	7.42537e6	99.33	58.66	99.46
$\lvert d_0\rvert < 0.1$ mm	5.61934e6	75.68	58.52	99.76
Track isolation	2.97599e6	52.96	57.67	98.54
$p_T^\mu > 20$ GeV	289422	9.725	49.55	85.92
$\lvert\eta^\mu\rvert < 2.4$	284755	98.39	48.60	98.08

Table 4.3: Number of events passing muon selection cuts with efficiency in percent of the cuts on W $\to \mu\nu$ + jets and Z $\to \mu\mu$ + jets ALPGEN Monte Carlo simulations. Marginal efficiency is given by the ratio of the number of events that passed the selection cut to the number that passed the previous cut while absolute efficiency efficiency is given by the number of events that passed the selection cut to the total number of events as calculated in MC.

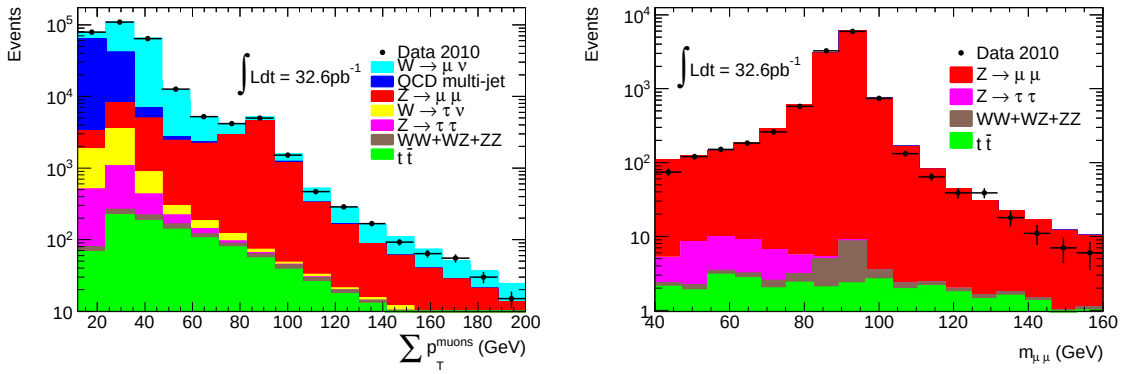

Figure 4.4: Muon kinematic variables: (left) the sum of the p_T of the reconstructed muons in an event, (right) the invariant mass distribution of two reconstructed muons. Points correspond to data, colored histograms MC.

taken to be the sum of its corresponding topoclusters taken at EM scale. Before jet selection the four-momentum is calibrated using the jet energy scale (JES) correction obtained from the Jet/EtMiss working group using a numerical inversion method [55]. More detail on the selection criteria is given below.

- **Jet p_T:** To reduce multi-jet background jet candidates are required to have a transverse momentum greater than 20 GeV ($p_T^{jet} > 20$ GeV).
- **Pseudorapidity:** Jet candidates are required to have a absolute pseudorapidity value less than 2.8 ($|\eta^{jet}| <2.8$).
- **Ugly jets:** All "ugly" jets as defined in [55] are removed as jet candidates. An Ugly jet is defined as having a TileGap3 energy fraction greater than 0.5 or an energy fraction in dead cells receiving a large correction (`BCH_CORR_CELL>0.5`). Ugly jets are not modelled in Monte Carlo simulations hence this selection is only made on data.
- **Pile-up:** Some extra jets may be counted due to additional proton-proton interactions in the same bunch crossing (in-time pile-up). To remove jets due to pile-up the absolute value of the jet vertex fraction (JVF), a measure of the probability that the jet originated from the primary vertex, is required to be greater than 0.75 (|JVF| >0.75) as recommended in [56] [57].
- **Jet-muon overlap:** To ensure muons do not fake jets, all jet candidates must be isolated from the selected muon(s). Specifically, all jet candidates with a ΔR with respect to the muon ($\Delta R(\mu, jet) = \sqrt{(\eta^{\mu} - \eta^{jet})^2 + (\phi^{\mu} - \phi^{jet})^2}$) less than 0.5 are excluded ($\Delta R(\mu, jet) >0.5$).

Table 4.4 lists the number of events in data that pass these selection cuts along with their efficiencies. Fig. 4.5 displays jet multiplicities, $\sum p_T^{jets}$, leading jet p_T and second leading jet p_T for all jets that pass this selection criteria. Agreement between data and MC is quite reasonable. The largest source of background to the W/Z+Jets signal is the QCD multi-jet background in the lower p_T region while $t\bar{t}$ dominates the higher p_T regions.

4.3.3 W+jet selection

Selection in the W channel is designed to identify W bosons based on the decay signature of one muon and a large missing transverse energy associated with a neu-

Selection cut	Data 2010		Signal MC (%)	
	Number of Jets	Marginal Eff. (%)	Absolute Eff.	Marginal Eff.
Preselection/muon cuts	1.37217e6	—	—	—
$p_T^{Jet} > 20$ GeV	216171	15.75	8.68	—
$\|\eta^{Jet}\| < 2.8$	193603	89.56	7.86	90.55
Ugly jets	193548	99.97	—	—
JVF < 0.75	185294	95.74	6.72	85.50
Jet-muon overlap	152245	82.16	6.34	94.35

Table 4.4: Number of events passing jet selection cuts with efficiency in percent of the cuts on W $\rightarrow \mu\nu$ + jets and Z $\rightarrow \mu\mu$ + jets ALPGEN Monte Carlo simulations. Marginal efficiency is given by the ratio of the number of events that passed the selection cut to the number that passed the previous cut while absolute efficiency efficiency is given by the number of events that passed the selection cut to the total number of events as calculated in MC.

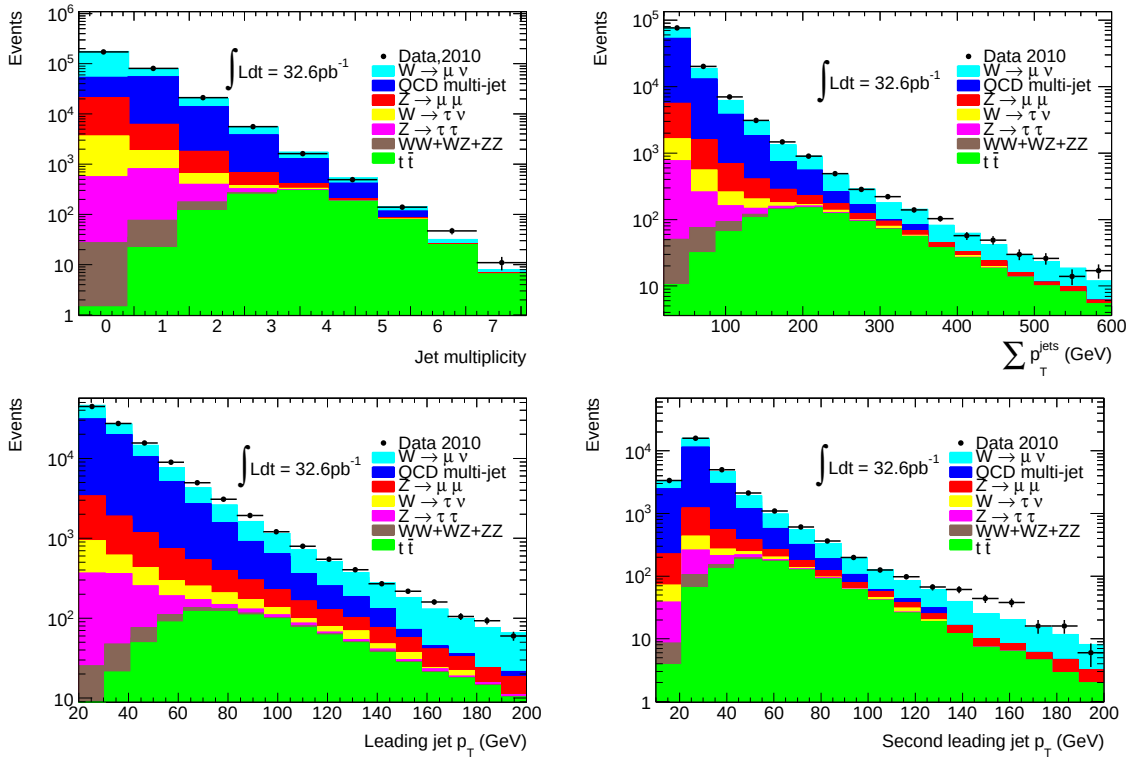

Figure 4.5: Jet variable distributions before W or Z boson selection: (top left) jet multiplicities, (top right) the sum of the p_T of all reconstructed jets in an event, (bottom left) leading jet p_T in an event, (bottom right) second leading jet p_T in an event.

Selection cut	Data 2010		Signal MC (%)	
	Number of events	Marginal Eff. (%)	Absolute Eff.	Marginal Eff.
One μ	272760	95.79	46.58	95.84
$E_T^{miss} > 25$ GeV	139634	51.19	38.67	83.02
$m_T > 40$ GeV	135796	97.25	38.34	99.15
$N_{Jets} \geq 1$	29295	21.57	7.626	19.89
$N_{Jets} \geq 2$	7593	25.92	1.929	25.30
$N_{Jets} \geq 3$	2404	31.66	0.4759	24.67

Table 4.5: Number of events passing W boson selection cuts with efficiency in percent of the cuts on W $\rightarrow \mu\nu$ + jets ALPGEN Monte Carlo simulations. Marginal efficiency is given by the ratio of the number of events that passed the selection cut to the number that passed the previous cut while absolute efficiency efficiency is given by the number of events that passed the selection cut to the total number of events as calculated in MC.

trino. This channel can be decomposed into sub-channels based on associated jet multiplicities. Details on W+jet selection are given below.

- **Single muon:** To veto Z bosons with a large missing energy W boson candidates require exactly one muon passing the muon selection criteria.

- **Missing E_T:** Neutrinos cannot be measured directly with the ATLAS detector. Their presence in an event is inferred by a large missing transverse energy, E_T^{miss}. Thus W boson candidates are required to have a missing transverse energy greater than 25 GeV ($E_T^{miss} > 25$ GeV).

- **Transverse mass:** Since the momentum of the neutrino is only reconstructed in the transverse plane (E_T^{miss}) it is not possible to reconstruct the invariant mass of the W boson. Instead the invariant mass of the W boson projected onto the transverse plane is reconstructed. Here the transverse mass is defined as

$$m_T = \sqrt{2p_T^{\mu} E_T^{miss}(1 - \cos[\phi^{\mu} - \phi^{miss}])}, \tag{4.2}$$

  where ϕ^{μ} and ϕ^{miss} are the ϕ-coordinates of the reconstructed muon and missing transverse energy vector respectively. W boson candidates are required to have a transverse mass greater than 40 GeV ($m_T > 40$ GeV).

Table 4.5 lists the number of events that pass the W+jets selection in data as well as the efficiencies of these selection cuts. Fig. 4.6 and Fig.4.7 shows the E_T^{miss} and m_T

Selection cut	Data 2010		Signal MC (%)	
	Number of events	Marginal Eff. (%)	Absolute Eff.	Marginal Eff.
$\mu^+\mu^-$	11942	—	34.51	—
71< m_Z <121 GeV	10755	90.06	31.75	92.00
$N_{Jets} \geq 1$	2538	23.60	7.307	23.01
$N_{Jets} \geq 2$	754	29.71	1.916	26.22
$N_{Jets} \geq 3$	200	26.53	0.4977	25.98

Table 4.6: Number of events passing Z boson selection cuts with efficiency in percent of the cuts on Z $\rightarrow \mu\mu$ + jets ALPGEN Monte Carlo simulations. Marginal efficiency is given by the ratio of the number of events that passed the selection cut to the number that passed the previous cut while absolute efficiency efficiency is given by the number of events that passed the selection cut to the total number of events as calculated in MC.

distributions before and after the E_T^{miss} and m_T selection cuts with jet multiplicities 0-3. The multi-jet QCD background clearly dominates in the $E_T^{miss} < 25$ GeV region. With these two cuts 98.6% of the multi-jet QCD background in the W channel is eliminated. Overall agreement between data and MC is good. MC over estimates data in the 0 jet bin around the E_T^{miss} and m_T distribution peaks in the signal region. this discrepancy is not understood however it is also seen in the W+jets cross section note [51] with 1.3 nb^{-1} of integrated luminosity.

4.3.4 Z+jets selection

Selection in the Z channel is designed to identify Z bosons based on the decay signature of two oppositely charge muons. The invariant mass of the two muons is reconstructed and required to be consistent with the mass of the Z boson. This channel can be decomposed into sub-channels based on associated jet multiplicities. Details on the selection are given below.

- **Opposite-charge μ:** Since the Z boson is neutral, the two muons associated with its decay must be oppositely charged. Therefore Z boson candidates require two muons, both passing the muon selection criteria, with opposite charge.

- **Invariant mass:** The Z boson has a distinct invariant mass peak at 91.2 GeV. For this reason the invariant mass of a Z boson candidate must fall within a 40 GeV window centred around 91 GeV (71 GeV < m_Z <111 GeV).

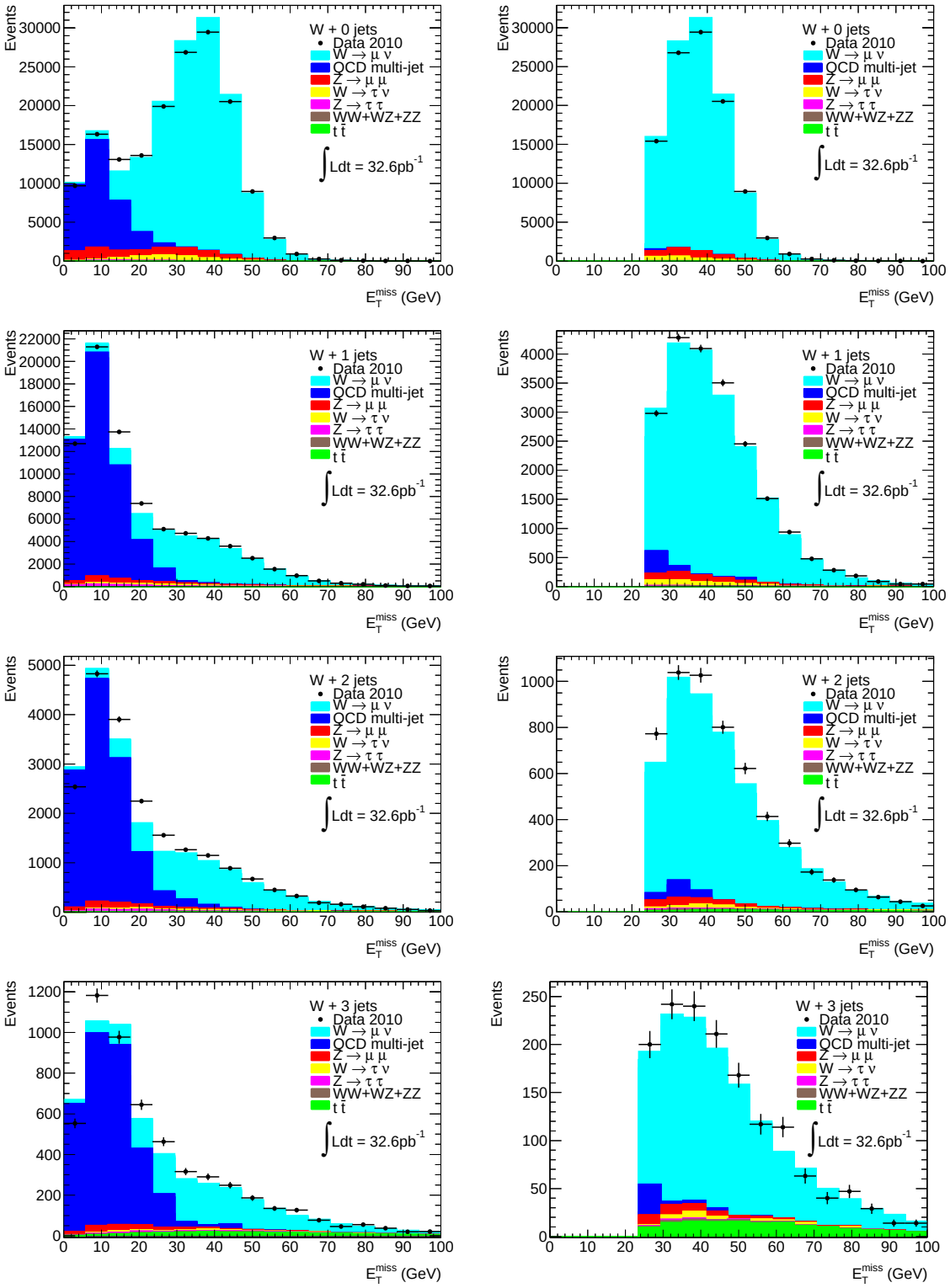

Figure 4.6: Missing transverse energy distributions before (left) and after (right) E_T^{miss} and m_T selection cuts with jet multiplicities 0-3.

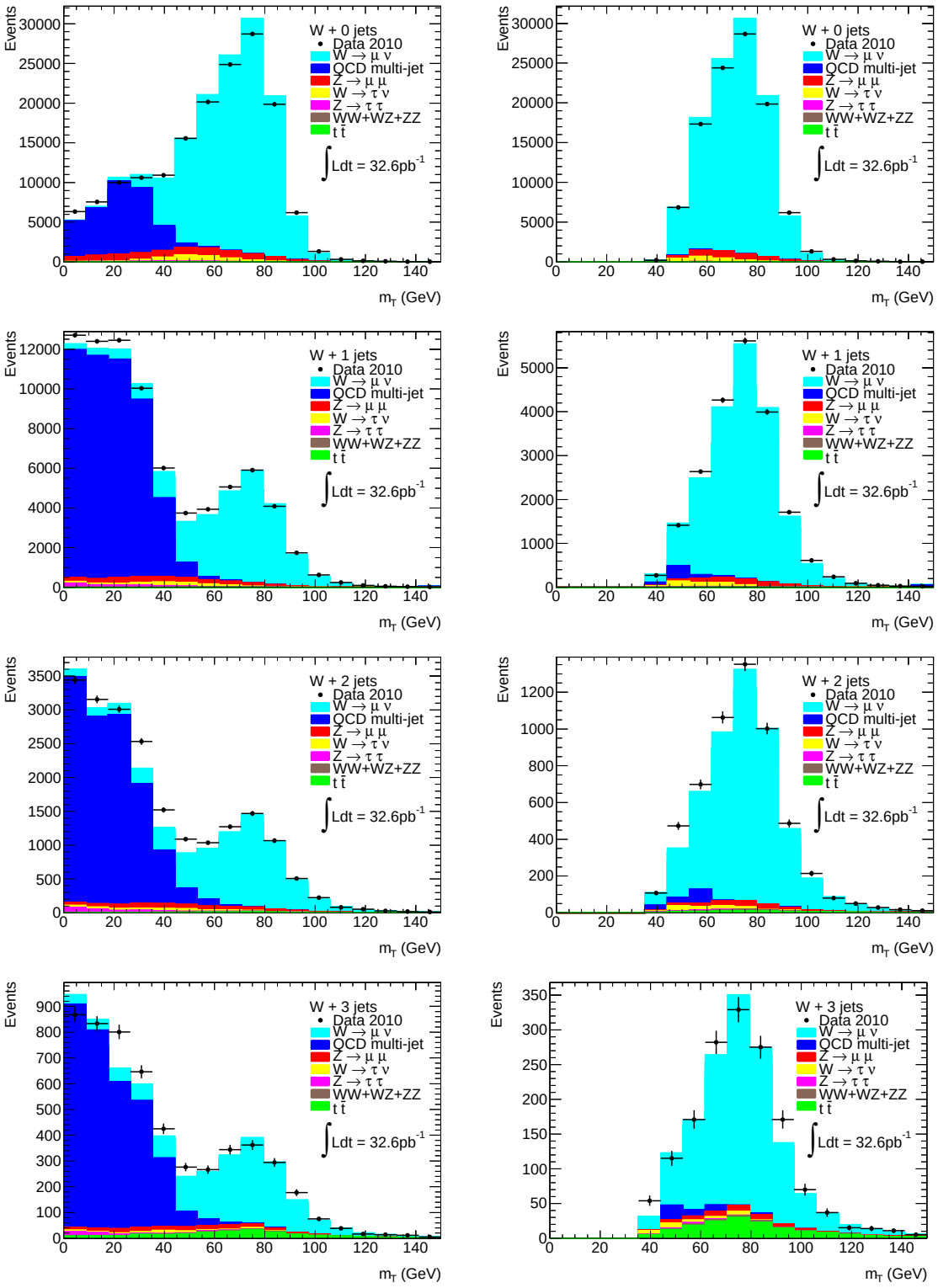

Figure 4.7: Transverse mass distributions before (left) and after (right) E_T^{miss} and m_T selection cuts with jet multiplicities 0-3.

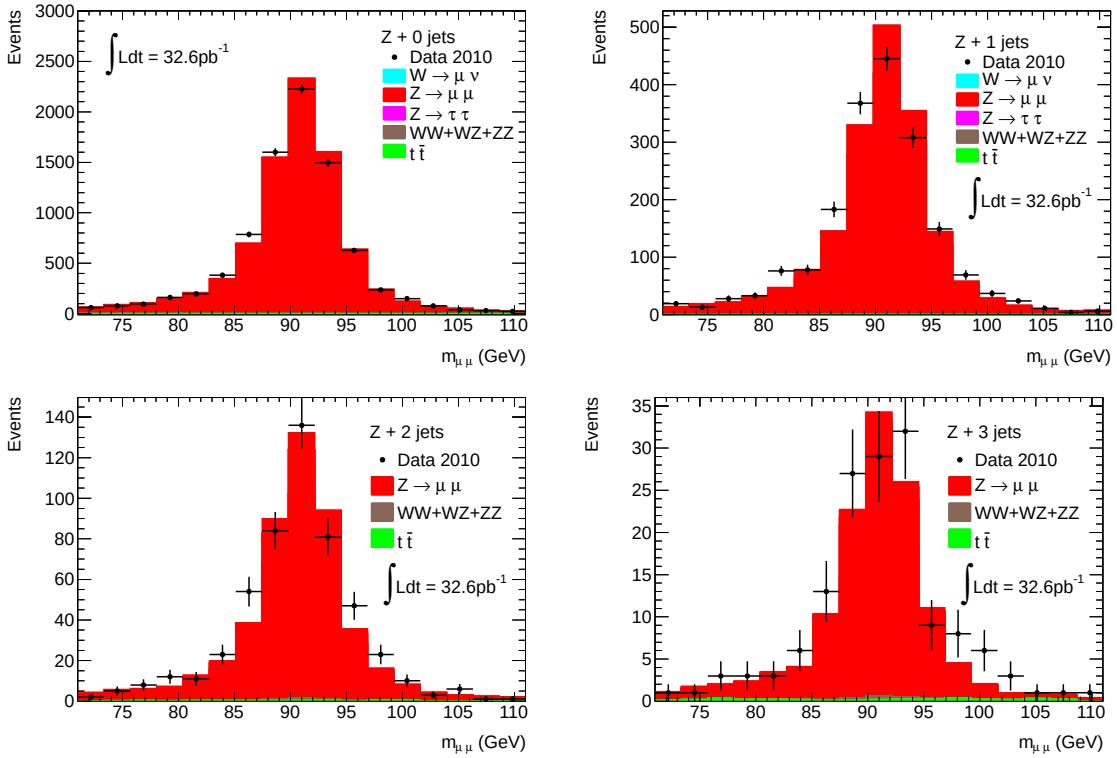

Figure 4.8: Z boson invariant mass distribution with jet multiplicites 0-3. Results are shown after smear procedure outlined in Section 4.2.3

Table 4.6 list the number of events in data that pass the Z channel selection cuts. Fig.4.8 shows the Z boson invariant mass distribution reconstructed from the muons, $m_{\mu\mu}$, with jet multiplicities 0-3. These distributions are shown after the smearing procedure outlined in Section 4.2.3 is applied.

4.4 R_n analysis

This section presents R_n as a function of k_T threshold measured with the 2010 data set. Fig. 4.9 shows R_n calculated as a function of the sum of the p_T of the jets. Prediction curve (red dashed line) is calculated as the sum of the MC samples listed in Table 3.3 after corrections and selection of Sections 4.2 and 4.3. The error bars on data are purely statistical and therefore underestimate the total uncertainty. The sharp increase in the size of the error bars with increasing k_T indicate limited statistics in the Z+jets measurement.

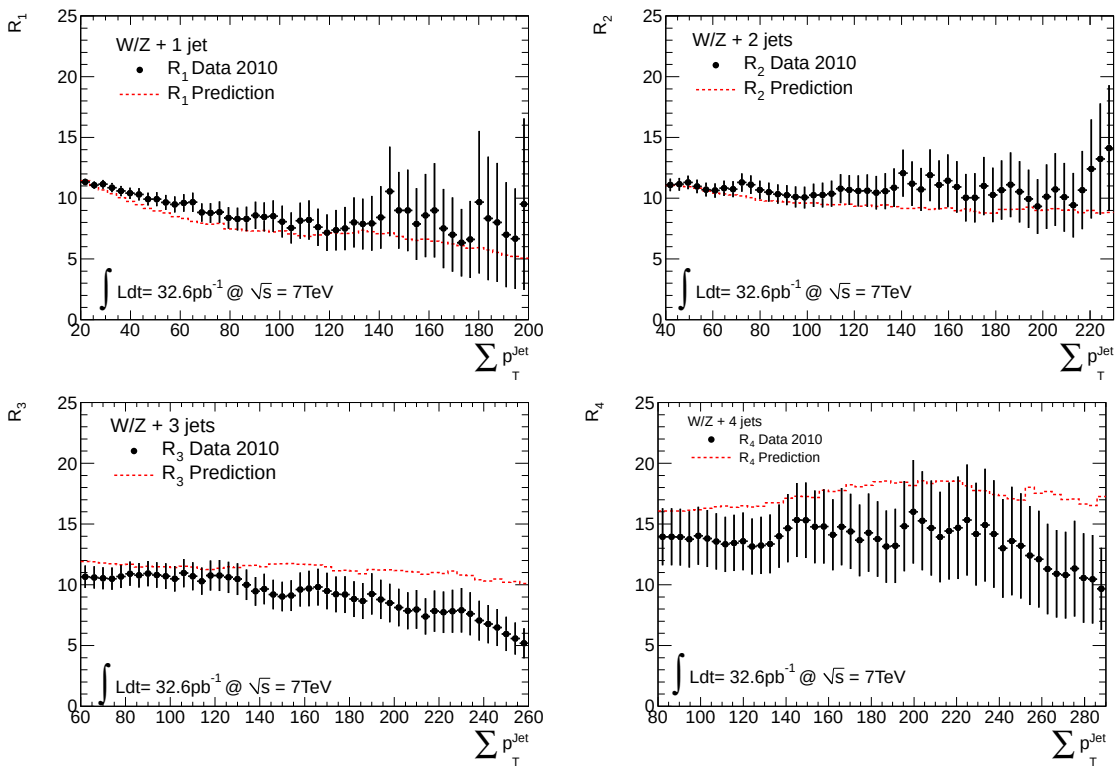

Figure 4.9: W/Z+jets ratio R_n presented as a function of $k_T = \sum p_T^{jet}$ for jet multiplicities 1-4. Points correspond to data and dashed line MC with all backgrounds and signals as listed in Table 3.3. Error bars are purely statistical.

MC prediction agrees reasonably well with data. For larger values of jet multiplicity, such as $n \geq 4$, R_n displays a systematic shift to larger values. This effect is due to the $t\bar{t}$ background in the W+jets channel, which is not found in the Z+jets channel. This extra background in the numerator of R_n is the cause of this upward shift. The discrepancy between data and prediction seen in R_3 and R_4 is not fully understood. One possible source of this discrepancy could be poor modelling due to the uncertainty associated with leading-order parton density function calculations used in ALPGEN signal samples. This possibility is supported by the observation that the discrepancy between data and MC systematically increases with the number of jets, which corresponds to the number of partons on the generator level.

Chapter 5

Multivariate analysis

5.1 Machine learning algorithms

In general, machine learning involves a program that is capable of changing its structure or manipulating data in such a manner that its expected future performance improves [58]. Learning algorithms of this type are useful in pattern recognition, or *classification*, and prediction of complex systems. The multi-dimensional cuts and neural network multivariate methods discussed below are examples of such machine learning algorithms. These algorithms need to be calculated from some set of inputs and evaluated based on specified criteria before a response can be determined. The learning paradigm used by these multivariate methods is known as *supervised learning* since the algorithm learns from a set of training examples, each of which is made up of inputs and a desired output(s) or *target*. The supervised learning algorithm uses the training examples to infer a function that maps the inputs to the target. This function is found by minimizing the error, or extremizing the *cost function*, between the model output and the target. This is done by continuously adjusting the function parameters through an iterative process called *training*. The resulting function that maps the inputs to the target with the least 'cost', as defined by the cost function, is called the *classifier*. The difference between this type of learning and "unsupervised" learning is that the cost function will implicitly contain prior knowledge of the problem in order to evaluate the correct mapping, i.e. the desired result must already be known beforehand. Applications in high energy physics usually involve training a classifier on simulated data with a signal and background target of 1 and 0 respectively. Some classifiers, such as a Fisher discriminant, are optimized for in-

put variables that have mostly linear correlations while others, like artificial neural networks, perform better on variables with non-linear correlations . In the following analysis the topology-discriminating variables are used as inputs to a classifier, the response of which is used as a discriminant.

5.1.1 Multi-dimensional cuts

The simplest and most common classifier is a set of rectangular cuts that maximizes signal efficiency and background rejection. This classifier returns a binary response; an event is classified as either signal or background. In the following analysis the optimal set of cuts on the input variables is found for a given signal efficiency using a Monte Carlo sampling method. Training consists of generating a large sample of random cuts on the input variables, where signal efficiencies and background rejections are calculated for each set of cuts with the training sample. Signal efficiencies are finely binned and their respective background rejections are compared for each bin. The set of cuts that has the largest background rejection for a given efficiency bin is retained while the others are rejected.

5.1.2 Artificial Neural Networks

Generally speaking an Artificial Neural Network (ANN) is a collection of interconnected nodes which each produce a certain response from a given set of input signals [59]. Originally ANN's were designed as a model of biological neural networks, which is where they inherit their name. However, most modern ANN's, which are usually referred to simply as neural networks, are used as non-linear statistical data modelling tools for classification and regression problems. By far the most common of these types of neural networks are multi-layer perceptrons (MLP). MLP's consist of multiple layers of nodes called *neurons* in a directed graph that maps a set of input data onto a single (classification) or set (regression) of outputs. The neurons are associated with a *neuron response function* that maps the inputs of that neuron to an output, where these outputs act as inputs for the next neuron layer. Each layer of neurons is fully connected by weights that determine the strength of the connection between neurons. There is no limit to the number of layers in a MLP; however there must be at least two layers (input and an output layer). All other layers are referred to as *hidden* layers since their states are usually not know to the user. MLP's use a supervised learning paradigm called *backpropagation* for training that utilizes the

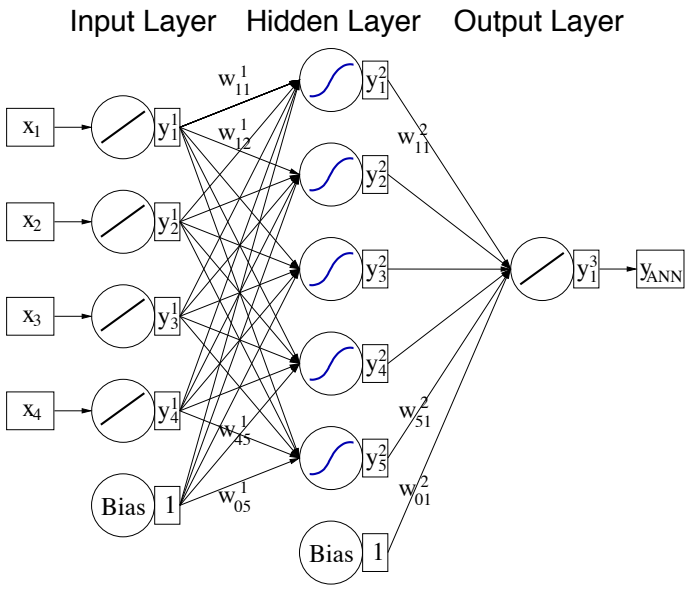

Figure 5.1: Network diagram of a multilayer perceptron with one hidden layer, taken from [60].

method of steepest decent for minimization of the cost function. Fig. 5.1 show a network diagram for a MLP with one hidden layer.

Neuron response function

The neuron response function, α, of neuron j in a given layer l maps the neuron inputs i onto the neuron output $y_j^{(l)}$. The response function can be decomposed into a $\mathbb{R}^n \mapsto \mathbb{R}$ *synapse function* β,

$$\beta = \sum_i y_i^{(l)} w_{ij}^{(l)}, \tag{5.1}$$

where $w_{ij}^{(l)}$ are the inter-neuron connection weights, and a $\mathbb{R} \mapsto \mathbb{R}$ *neuron activation function* $\gamma(\beta)$

$$\gamma(\beta) = \tanh(\beta). \tag{5.2}$$

Equations 5.1 and 5.2 make up one choice of a possible neuron response function, $\alpha = \gamma(\beta)$, for other possibilities see [59]. With the neuron response function defined the neuron output $y_j^{(l)}$ can be given in terms of the inter-neuron connection weights

and the responses from the neurons in the previous layer

$$y_j^{(l)} = \gamma_j(\beta) = \tanh\left(\sum_i y_i^{(l-1)} w_{ij}^{(l-1)}\right),$$

where l is a hidden layer. For the input and output layers the activation function is usually taken to be a linear function of β.

Backpropagation

Backpropagation uses the method of steepest decent to adjust the inter-neuron connection weights such that the cost function is minimized. Here the cost function, Φ, which measures agreement between the neural network response, y_{ANN}, and the desired target value y_{targ}, is defined by

$$\Phi = \frac{1}{2}(y_{ANN} - y_{targ})^2.$$

The learning is said to be done online as the cost function is evaluated for each training event, rather than evaluating it as the sum over all events. Φ is minimized by starting with a random set of weights, $\vec{w}$, and then adjusted by moving a small distance in w-space in the direction where Φ decreases most rapidly

$$\vec{w} \to \vec{w} - \xi \nabla_{\vec{w}} \Phi,$$

where ξ is a real positive value that determines the step size. The weights are adjusted by starting with the last hidden layer propagating the results backwards to the first; hence the term *backpropagation*. This process of adjusting the weight is continued through to the weights connecting the first hidden layer to the input layer. Once all weights have been adjusted the cost function is evaluated again and the process is repeated until Φ is minimized.

Single hidden layer MLP network architecture

This section demonstrates how neural networks can be used to solve classification problems in high energy physics. Given a simulated sample of signal and background events and a set of discriminating variables x_i a MLP neural network can be trained to separate data into signal and background events. In theory there is no limit to

the number of hidden layers a MLP can have; in practice computation time makes only a few hidden layers practical. Computationally it is usually more efficient to increase the number of neurons instead of layers. In fact any continuous function can be approximated using a single hidden layer MLP given a sufficient number of neurons [59]. In the following analysis only single hidden layer MLP's are used where the number of neurons in the hidden layer is chosen to be 5 plus the number of input variables. Bias nodes, with weights $w_{0j}^{(l)}$ and $\alpha = 1$, in the first two layers add a constant offset to the neuron response function allowing it to be shifted left-or-right making the MLP more flexible. Fig. 5.1 is an example of such a network with four input variables.

For a MLP with n input variables and a neuron response function $\alpha = \beta$ for the input and output layers the neural network response y_{ANN} is given by

$$y_{ANN} = w_{01}^{(2)} + \sum_{j=1}^{n+5} y_j^{(2)} w_{j1}^{(2)} = w_{01}^{(2)} + \sum_{j=1}^{n+5} \tanh\left(w_{0j}^{(1)} + \sum_{i=1}^{n} x_i w_{ij}^{(1)} \right) w_{j1}^{(2)}. \qquad (5.3)$$

In the first iteration y_{ANN} is calculated with a random set of weights. Then the weights are adjusted using the backpropagation method. Weights $w_{j1}^{(2)}$ are adjusted by an amount

$$\Delta w_{j1}^{(2)} = -\xi \frac{\partial \Phi}{\partial w_{j1}^{(2)}} = -\xi (y_{ANN} - y_{targ}) \cdot y_j^{(2)},$$

where y_{targ} is taken to be 0 for background and 1 for signal events. After this the second set of weights $w_{ij}^{(1)}$ are adjusted by the amount

$$\Delta w_{ij}^{(1)} = -\xi \frac{\partial \Phi}{\partial w_{ij}^{(1)}} = -\xi (y_{NN} - y_{targ}) w_{j1}^{(2)} \cdot (1 - (y_j^{(2)})^2) \cdot x_i,$$

where the relation $\tanh' x = 1 - \tanh^2 x$ is used. The new weights, $\vec{w} + \Delta\vec{w}$, are used to recalculate y_{ANN} with Equation 5.3 and Φ is re-evaluated. Then process is repeated until Φ converges between iterations, where the converged value is taken to be the minimum.

Variable ranking

After training of the MLP neural network it may be informative to rank the input variables based on some measure of their importance in forming the neural network response. One such measure is the sum of the squared weights of the inter-neuron

connections between the input layer and the first hidden layer. I_i, the importance of the ith input variable is defined as

$$I_i = \mu_i \sum_{j=1}^{n} \left(w_{ij}^{(1)} \right)^2, \tag{5.4}$$

where μ_i is the sample mean of the input variable x_i.

5.1.3 Fisher discriminant

A Fisher discriminant is a linear combination of input variables that maximizes the separation in F-space of two output targets for input variables with linear correlations only. Fisher discriminants can be thought of as linear neural networks with zero hidden layers – although, strictly speaking, they do not meet the definition of a machine learning algorithm given in Section 5.1. The Fisher discriminant F is given by [61]

$$F = w_0 + \sum_{i=1}^{n} w_i x_i,$$

where w_i are the weights for the connections between the input layer and output layer of neurons. The cost function is given by the separation, S, between signal and background in F-space defined by

$$S = \frac{(\mu_F^S - \mu_F^B)^2}{(\sigma_F^S)^2 + (\sigma_F^B)^2} \tag{5.5}$$

where $\mu_F^{S(B)}$ and $\sigma_F^{S(B)}$ are the mean and standard deviation of the Fisher outputs for a signal (background) sample. For two distributions defined in exclusive non-overlapping domains $S = 1$ and for identical distributions $S = 0$. It can be shown [62] that S is maximized by the choice of weights

$$w_i = \sum_{j=1}^{n} (V^S + V^B)_{ij}^{-1} (\mu_j^S - \mu_j^B),$$

where $V^{S(B)}$ is the covariance matrix and $\mu_j^{S(B)}$ is the mean of the input variable x_i for the signal (background) sample.

By only taking into account linear correlations between input variables it is possible to extremize the cost function analytically; hence no iterations or learning is

necessary. This leads to a much shorter computation time compared to a MLP neural network or multi-dimensional cuts.

In Chapter 6 the performance of the above classifiers is evaluated in two separate applications. The set of topology-discriminating variables are used as inputs and the classifiers are trained and tested on two separate sets of simulated data.

Chapter 6

Signal extraction

In this section the analysis strategy outlined in Section 2.4 is applied to the measurement of R_n in two parallel analyses. The first analysis enhances a $t\bar{t}$ signal in $R_{n\leq 4}$ relative to the remaining SM background using the 2010 data set. The second analysis enhances a $LQL\bar{Q}$ signal in R_n with respect to a SM background using MC simulation. In Section 6.1 a common set of topology-discriminating variables are chosen to be used as inputs to multivariate classifiers for both analyses. Then in Sections 6.2 and 6.3 multivariate classifiers are trained using MC simulation and their respective responses are evaluated based on discriminating power. The kinematic threshold variable k_T is chosen such that signal and background are found in distinct phase spaces. A single cut is then made on the chosen classifiers response and R_n is calculating in the remaining k_T phase space. The result is then compared to R_n before the cut to determine if the signal has been enhanced.

6.1 Topology-discriminating variables

The set of topology-discriminating variables used in both the $t\bar{t}$ and $LQLQ$ analysis are: N_{jets}, ΔR moment, C_{max}, S_T and Transverse Thrust, which are defined in Section 2.3. Fig. 6.1 displays these variables calculated with the 2010 data set, shown after preselection, muon selection and jet selection as well as MC corrections from Section 4. Variable distributions are simulated well in MC, showing good agreement between data and MC for all variables. However variable distribution shapes may be sensitive to other considerations that affect the topoclusters in an event from which the variables are calculated. Two such things to consider are pile-up, as discussed in

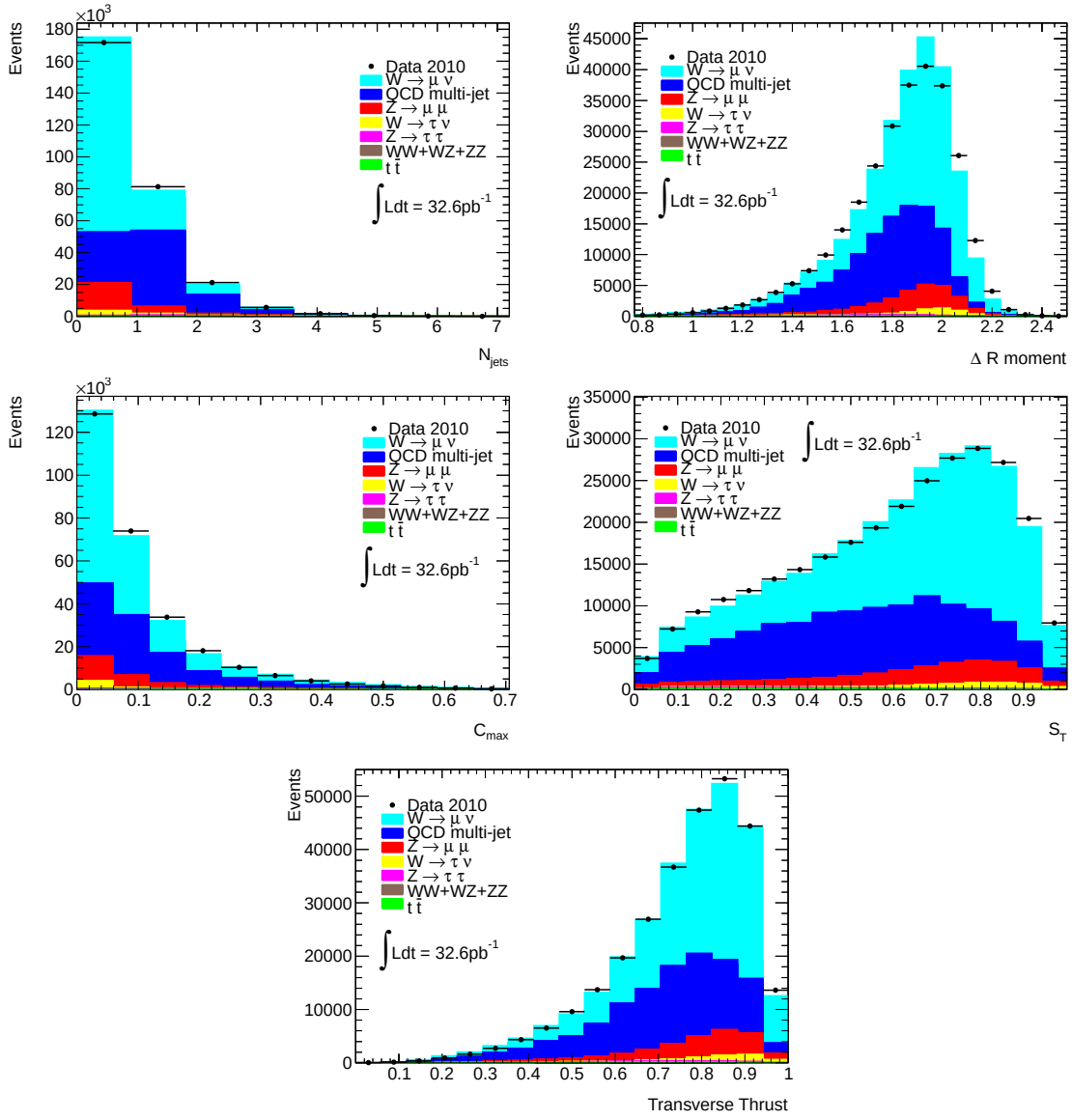

Figure 6.1: Topology-discriminating variable distributions: N_{jets} , ΔR moment, C_{max}, S_T and Transverse Thrust. Points correspond to 2010 data and coloured histograms to MC scaled to 2010 data integrated luminosity. Preselection, muon selection and jet selection as well as MC corrections from Section 4 have been applied to distributions.

the Appendix, and the total transverse momentum measured in the calorimeter.

6.1.1 Sensitivity of topology-discriminating variables to $\sum p_T^{clust}$

As the total energy of an event increases the relative production rates of the different processes can change. For example the relative production rate of $t\bar{t}$ will increase with increasing total energy while the QCD multi-jet background will decrease. This can have significant effects on the shape of the topology-discriminating variable distributions. Fig 6.2 illustrates this with the ΔR moment distributions shown in increasing bins of $\sum p_T^{clust}$. The shape, mean value and width of the ΔR moment distribution dramatically changes with increasing $\sum p_T^{clust}$. This suggests that the separation and discriminating power of the ΔR moment is dependent upon the $\sum p_T^{clust}$ region that it is calculated in. This effect is common to all of the topology-discriminating variables and for irreducible backgrounds, and is ultimately unavoidable since it is the relative contributions from different process topologies that causes it. Similar figures for the other four variables can be found in Appendix A.2. When these variables are used as inputs to a multivariate classifier this effect could result in less-than-optimal discriminating power of the classifier response in certain $\sum p_T^{clust}$ regions. In the next section a binned MLP neural network classifier is discussed where multiple neural networks are computed in distinct $\sum p_T^{clust}$ regions to accommodate this effect.

6.2 $t\bar{t}$ analysis

With an integrated luminosity of 32.6 pb^{-1} the reach of any new physics search would be too limited to detect any beyond-Standard-Model (BSM) processes. However, the analysis strategy and signal-extraction techniques that are being developed in this thesis can be tested by analysing a known SM process that can act as a proxy to a BSM process. For this purpose top quark pair production in the 2010 data set is used as a proxy for leptoquark pair production ($LQ\bar{LQ}$). Both processes involve production of two massive particles at, or near, threshold that rapidly decay in both leptonic and hadronic channels. The decay products contribute to the irreducible background in R_n in the W+jets channel. The current exclusion limit placed on second generation leptoquarks puts the leptoquark mass at over 2.4 times that of the top quark's [63]. Thus, the decay products of the leptoquark would be found in a higher k_T region than those of the top quark. However the topology of the two processes is similar,

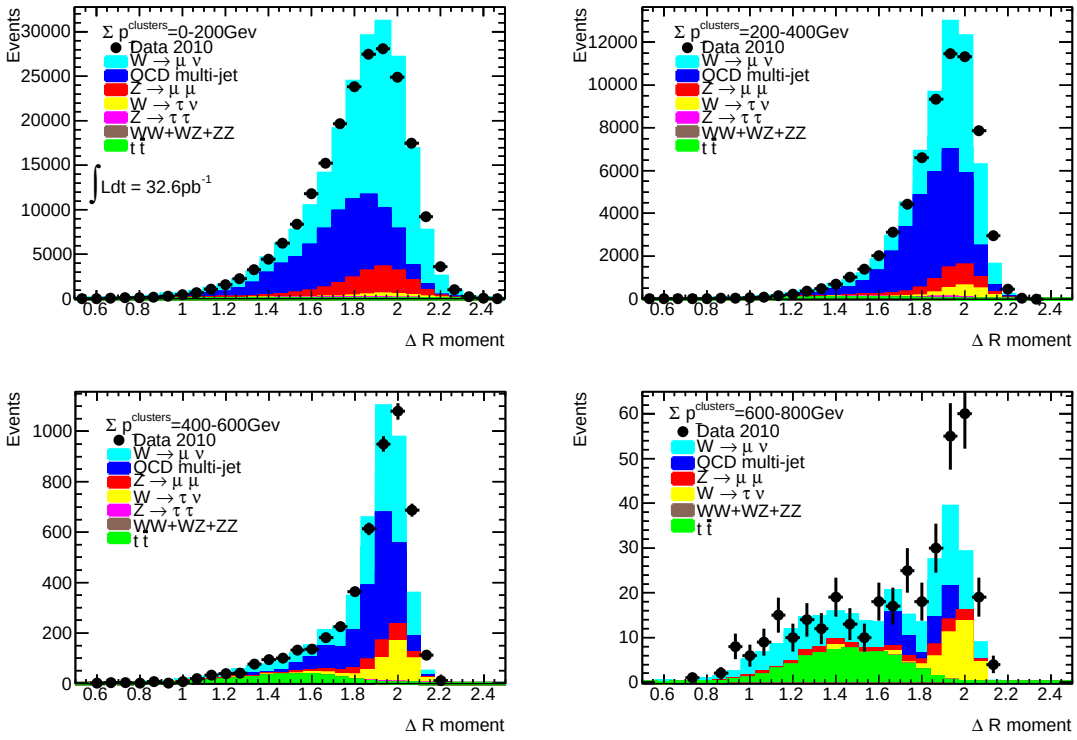

Figure 6.2: ΔR distribution calculated in increasing bins of $\sum p_T^{clust}$. For further details see caption of Fig. 6.1.

which allows the two analyses to be conducted in an analogous way.

In this section an analysis with the 2010 data set is presented where the sensitivity of $R_{n\leq 4}$ to $t\bar{t}$ production is enhanced relative to background using the set of topology-discriminating variables discussed in the last section. After W+jets selection W$\rightarrow$ $\mu\nu$ is the dominant process with contributions from W$\rightarrow$ $\tau\nu$, Z$\rightarrow$ $\mu\mu$, Z$\rightarrow$ $\tau\tau$, $t\bar{t}$ and diboson production (WW, ZZ, WZ). In this section three different optimized multivariate discriminants are constructed with the topology-discriminating variables as inputs by taking $t\bar{t}$ as the signal to be extracted from remaining background. The analysis is restricted to 1-4 jets since the denominator of $R_{n\geq 5}$, Z + $n \geq 5$ jets, is too statistically limited. The analysis is conducted after all corrections and selection outlined in Section 4.2 and Section 4.3.

Variable	N_{jets}	ΔR	C_{max}	S_T	Thrust
Separation	0.517	0.433	0.314	0.177	0.149

Table 6.1: Separation of topology-based discriminating variables between $t\bar{t}$ signal and SM-$t\bar{t}$ background in descending order of separation.

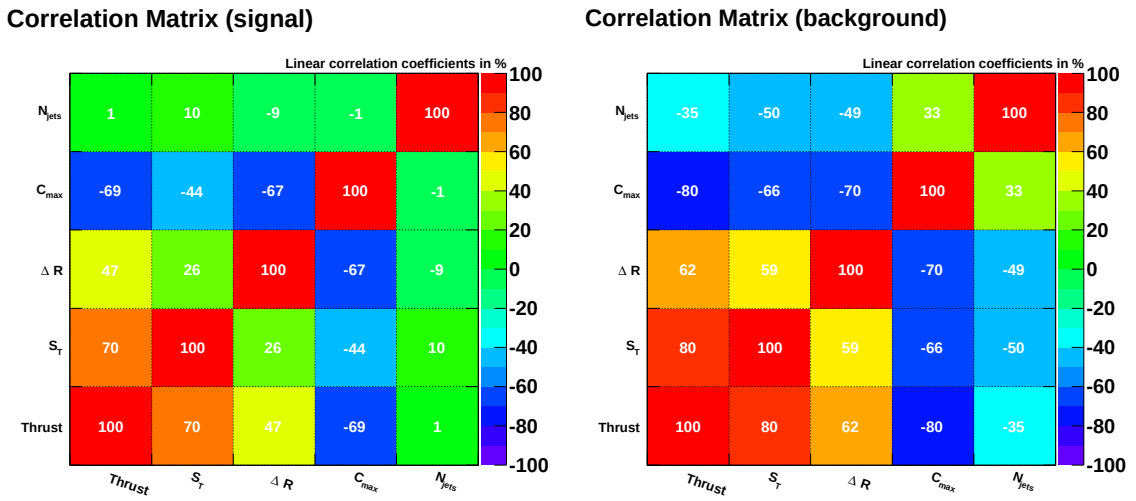

Figure 6.3: Linear correlation matrices for topology-based discriminating variables for $t\bar{t}$ signal (left) and SM-$t\bar{t}$ background (right).

6.2.1 Multivariate analysis with $t\bar{t}$ signal

The shapes of the multivariate input variables are compared in Fig 6.4 where signal and background have been normalized to equal area. Using equation 5.5 the separation between signal and background for each variable is calculated and shown in Table 6.1. In general the variables with the largest separation will provide the most discriminating power for any given multivariate classifier. It is also important that the input variables are largely uncorrelated since highly correlated variables will not add any additional information for the classifier. The linear correlation matrices for signal and background are show in Fig 6.3. N_{jets} is seen to provide the most separation and the least correlation among the variables, which would suggest that it the most powerful discriminating variable among the set.

To enhance the sensitivity of $R_{n\leq 4}$ to $t\bar{t}$ production three different multivariate classifiers are calculated with the set of topology-discriminating variables as inputs: multidimensional cuts, fisher discriminant and a MLP neural network. There are

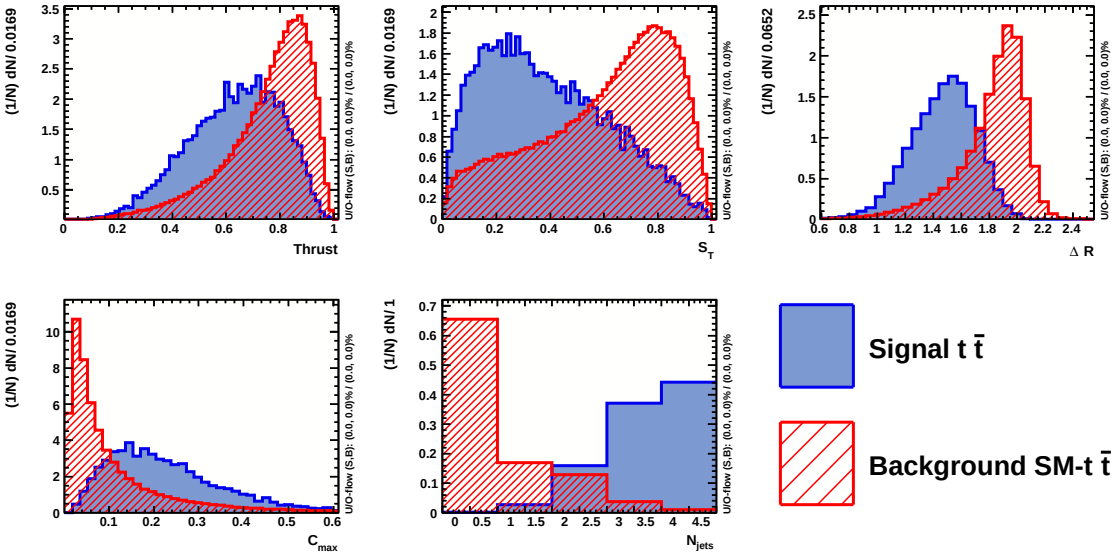

Figure 6.4: Comparison of topology-based discriminating variables distributions for $t\bar{t}$ signal (blue) and SM-$t\bar{t}$ background (red). Signal and background have been normalized to equal area.

Rank	$\sum p_T^{clust}$ bin				
	<200 GeV	200-400 GeV	400-600 GeV	600-800 GeV	>800 GeV
1	C_{max}	C_{max}	C_{max}	N_{jets}	N_{jets}
2	ΔR	ΔR	ΔR	S_T	S_T
3	N_{jets}	Thrust	Thrust	C_{max}	ΔR
4	Thrust	N_{jets}	N_{jets}	ΔR	C_{max}
5	S_T	S_T	S_T	Thrust	Thrust

Table 6.2: Topology-discriminating variables ranked in order of most important (1) to least important (5) in construction of MLP neural network in different $\sum p_T^{clust}$ bins.

many other classifiers that are often used in high energy physics that could be tested as well. However, these three classifiers are among the most common and are chosen to be representative of the three classes: classical selection cuts, linear discriminants and non-linear discriminants. The classifiers are trained and evaluated with the ROOT native C++ library TMVA [60]. In training half of the simulated events are used for training and the other half for testing and evaluation. Fig. 6.5 compares the background rejection vs signal efficiency curves for the classifier outputs. Out of the three classifiers the neural network offers the best background rejections for a given

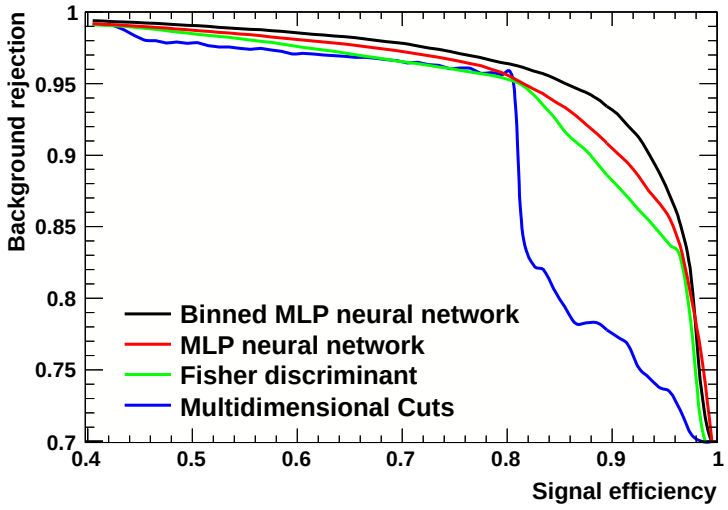

Figure 6.5: Comparison of background rejection vs signal efficiency curves for the classifier outputs multidimensional cuts, Fisher discriminant and MLP neural network. MC simulation is trained with $t\bar{t}$ as signal and SM-$t\bar{t}$ as background using TMVA

signal efficiency with the Fisher discriminant only slightly under performing it. The signal-background separation for the Fisher discriminant and neural network are 0.716 and 0.718 respectively. In contrast multidimensional cuts significantly under-performs both of these classifiers for signal efficiencies greater than 0.8.

The performance of the neural network can be improved upon by calculating multiple classifiers in bins of $\sum p_T^{clust}$. In Section 6.1.1 the dependence of the topology-discriminating variable distributions on $\sum p_T^{clust}$ is discussed and it is seen that these variables can have dramatically different distributions in different $\sum p_T^{clust}$ regions. This could lead to different signal-background separation for a given variable depending upon the $\sum p_T^{clust}$ region it is calculated in. In order to optimize discrimination of the neural network classifier a *binned* MLP neural network is trained in five separate $\sum p_T^{clust}$ bins, which correspond to a rough-grained classification of distinct $\sum p_T^{clust}$ regions. Table 6.2 ranks the importance of each variable in construction of the neural network output in a given $\sum p_T^{clust}$ bins as calculated with Equation 5.4. The order of ranking for the variables is different in all but two bins, which demonstrates the need for a binned classifier for optimal separation. In Fig. 6.5 the binned MLP neural network can be seen to out perform its unbinned counter-part making it the most powerful classifier. The signal-background separation for the output of the binned

neural network is 0.738. This is a significant improvement in separation over any of the individual topology-discriminating variables since N_{jets}, which offers the most separation in the set, only give a separation of 0.517.

6.2.2 $t\bar{t}$ signal enhanced $R_{n\leq 4}$

$R_{n\leq 4}$ is presented in this section where the kinematic threshold variable is taken to be the sum of the transverse momentum of the jets in the event, $k_T = \sum p_T^{jets}$. This choice of k_T offers good kinematic discrimination since the top quark decays predominantly in the hadronic channel. Fig. 6.6 compares $R_{n\leq 4}$ with and without a cut on the binned neural network response. The green curve corresponds to SM prediction without a $t\bar{t}$ signal while the red curves corresponds to SM prediction (with $t\bar{t}$ signal). The cut on the binned neural network is applied to the numerator of $R_{n\leq 4}$, or in the W$\rightarrow \mu\nu$ + jets channel while no cuts are applied to the denominator since the $t\bar{t}$ signal is not present in the Z$\rightarrow \mu\mu$ + jets channel. The cut on the binned neural network enhances the significance ($S/\sqrt{S+B}$) of the $t\bar{t}$ signal in $R_{n\leq 4}$ from 1.59 to 4.54, or by a factor of 2.86. This shows significant improvement on the original sensitivity of the $R_{n\leq 4}$ search, making a more compelling argument for the existence of top quark pair production.

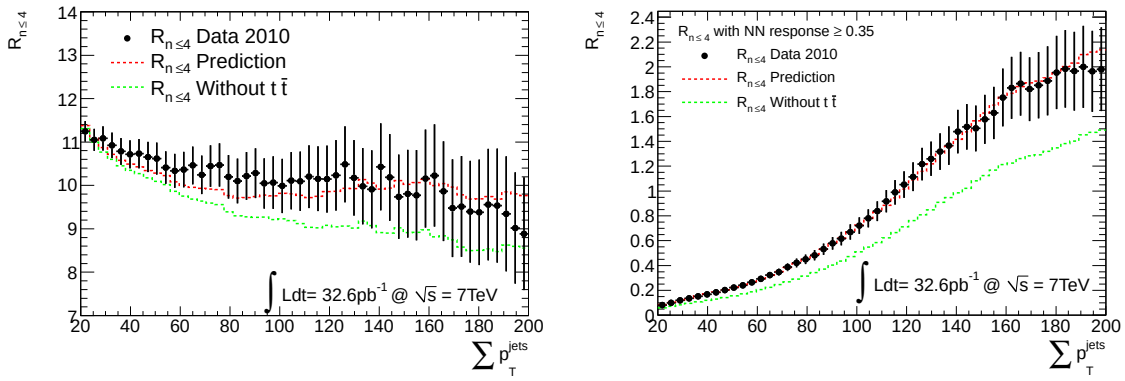

Figure 6.6: R_n with (right) and without (left) binned neural network response cut. Points correspond to 2010 data dashed curves correspond to MC SM prediction with (red) and without (green) $t\bar{t}$ signal.

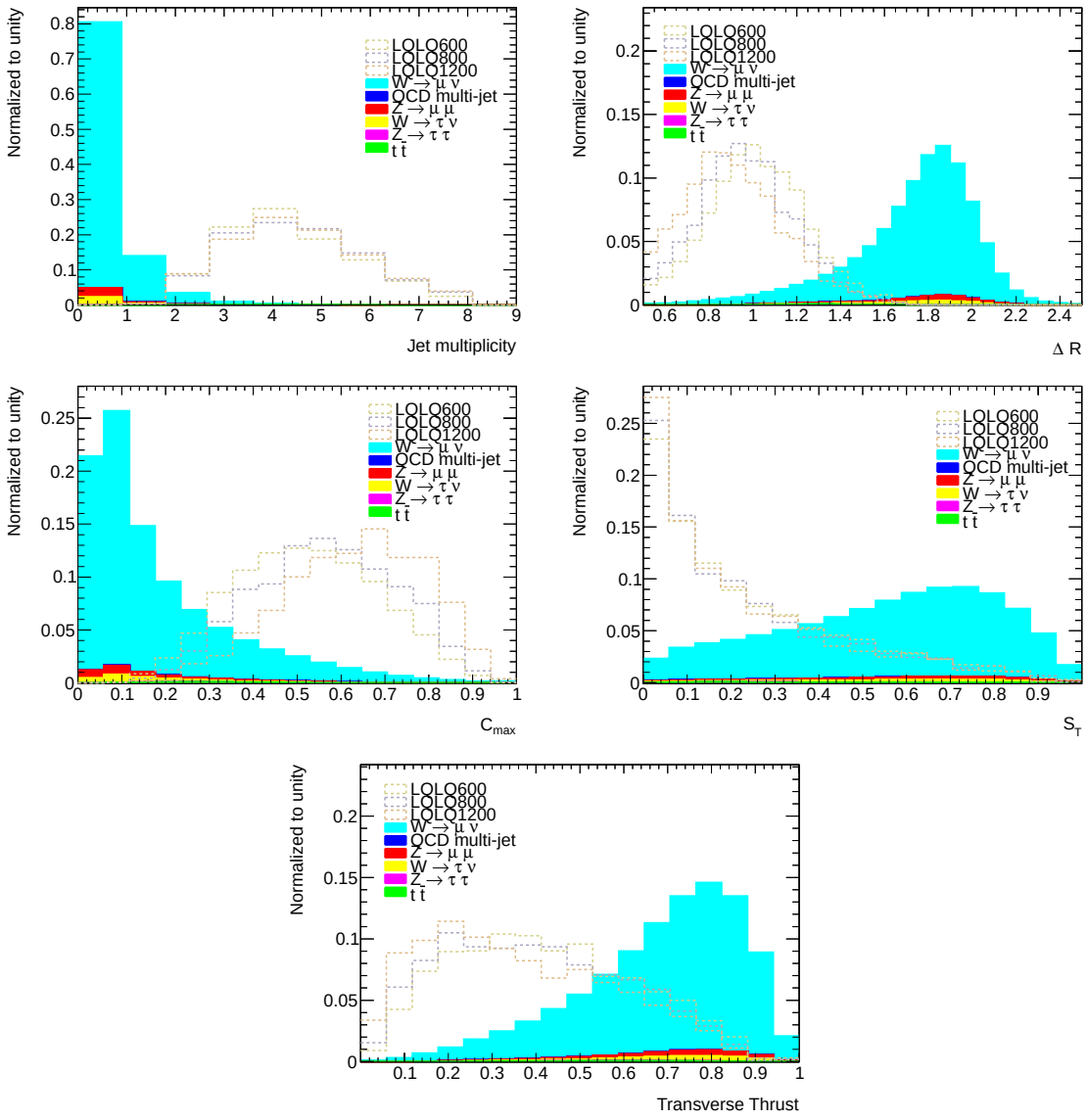

Figure 6.7: Topology-based variables: N_{jets} , ΔR moment, C_{max}, S_T and Transverse Thrust. Dashed lines correspond to leptoquark pair production signal with varying leptoquark mass and coloured histograms to SM background. Signal and background are normalized to unity.

Variable	N_{jets}	ΔR	C_{max}	Thrust	S_T
Separation	0.614	0.595	0.552	0.285	0.239

Table 6.3: Separation of topology-based discriminating variables between $LQ\bar{L}\bar{Q}$ signal and SM background in decending order of separation.

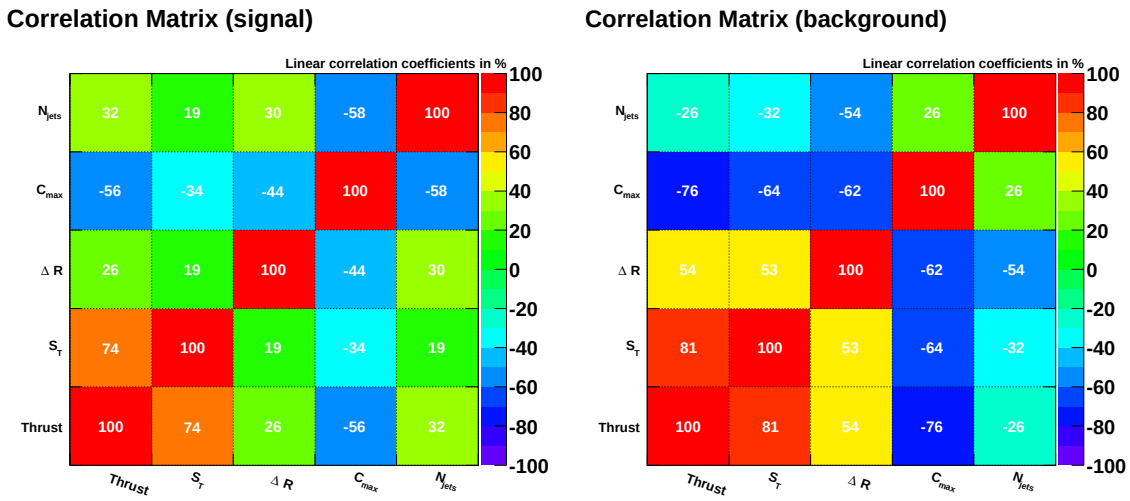

Figure 6.8: Linear correlation matrices for topology-based discriminating variables for $LQ\bar{L}\bar{Q}$ signal (left) and SM background (right).

6.3 Leptoquark analysis

In this section an analysis to enhance a $LQ\bar{L}\bar{Q} \to \mu\nu qq$ signal over SM background in R_n is presented. Current limits on second generation leptoquark searches exclude leptoquarks with Standard Model coupling below 422 GeV [63]. In this analysis three different values for leptoquark mass, m_{LQ}, are studied: $m_{LQ} = 600$ GeV, $m_{LQ} = 800$ GeV and $m_{LQ} = 1200$ GeV. MC data sets used with production cross sections are given in table 3.4. The small cross sections for leptoquark production requires a large integrated luminosity to extract a signal from the SM background. For this reason distributions are scaled to an integrated luminosity of 30 fb^{-1}.

6.3.1 Leptoquark multivariate analysis

Shapes of the multivariate input variables are compared in Fig. 6.7 where distributions have been normalized to unity. MC samples used for both background and signal can be found in Table 3.4. As one might expect leptoquark distributions show similar

trends to $t\bar{t}$ distributions. However leptoquark distributions benefit from greater separation, as can be seen in Table 6.3. Additionally Fig. 6.8 shows the linear correlation matrices for leptoquark signal and SM background. From these figures N_{jets} and ΔR can be seen to have the most separation and least correlation. In contrast to this Thrust and S_T appear to be have the least separation and the highest correlation.

Just as in the $t\bar{t}$ analysis three distinct classifiers are trained and compared along with a binned MLP neural network. Classifiers are trained on the $m_{LQ} = 800$ GeV sample but responses are evaluated on all three $LQL\bar{Q}$ samples. Fig. 6.9 compares the background rejection vs signal efficiency of the responses of the four multivariate methods. The ordering of discrimination power is the same as in the $t\bar{t}$ analysis with multidimensional cuts giving the least discrimination and the binned neural network giving the most. In this case there is a very significant improvement in signal efficiency for a given background rejection for the binned neural network over its unbinned counterpart. This would suggest that the separation of the set of input variables, and their relative ranking in importance, differs significantly in some of the $\sum p_T^{clust}$ bins. Table 6.4 shows the relative rankings of the input variables in importance as defined by equation 5.4. In the $\sum p_T^{clust} < 200$ bin N_{jets} and ΔR are ranked the highest while in the higher $\sum p_T^{clust}$ bins S_T is ranked higher and ΔR drops to one of the lowest ranked variables. It is due to this shift in relative importance of the input variables in the construction of the neural network that binning gives such a significant improvement in discrimination power. A possible explanation for this shift is that $t\bar{t}$ becomes the dominant background as one goes to higher $\sum p_T^{clust}$ regions, which, as we've seen in the previous section, has a distinct topology from W+jets. Thus the background from which the $LQL\bar{Q}$ signal needs to be discriminated from is quite different in the $t\bar{t}$ dominated bins. The binned neural network gives a signal-to-background separation of 0.893 where, in comparison, N_{jet} gives a separation of 0.614.

6.3.2 $LQL\bar{Q}$ signal enhanced R_n

R_n is presented in this section where the kinematic threshold is taken to be the sum of the transverse momentum of the leading and second leading jet, the missing transverse energy and the leading muon's transverse momentum, $k_T = p_T^{leading-jet} + p_T^{second-jet} + E_T^{miss} + p_T^{\mu}$. This choice of k_T offers excellent kinematic discrimination since the leptoquark decay channel being studied here is $LQL\bar{Q} \to \mu\nu qq$, where, at the

Rank	$\sum p_T^{clust}$ bin				
	<200 GeV	200-400 GeV	400-600 GeV	600-800 GeV	>800 GeV
1	N_{jets}	N_{jets}	S_T	S_T	S_T
2	ΔR	S_T	N_{jets}	N_{jets}	N_{jets}
3	C_{max}	C_{max}	C_{max}	C_{max}	Thrust
4	Thrust	ΔR	Thrust	Thrust	ΔR
5	S_T	Thrust	ΔR	ΔR	C_{max}

Table 6.4: classifier input variables ranked in order of most important (1) to least important (5) in construction of MLP neural network in different $\sum p_T^{clust}$ bins.

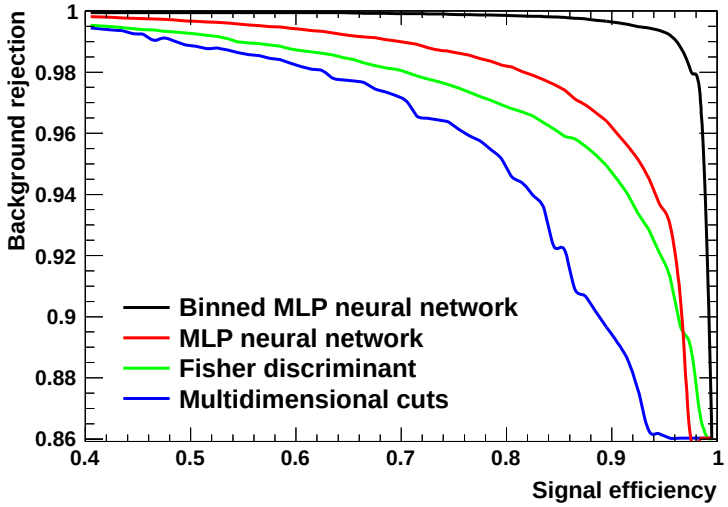

Figure 6.9: Comparison of background rejection vs signal efficiency curves for the classifier outputs multidimensional cuts, Fisher discriminant and MLP neural network. MC simulation is trained with $LQLQ$ as signal and SM as background using TMVA

detector level, the two quarks correspond to the two hardest jets in the event and the E_T^{miss} to the undetected neutrino. Since the leptoquark samples used have production cross sections that are 10^6-10^9 times smaller than W+jets, SM background needs to be largely eliminated to resolve the $LQL\bar{Q}$ signal. For this purpose the transverse mass, m_T, cut made in the W+jets selection is increased to $m_T > 200$ GeV to reject SM background. Fig. 6.10 shows R_n before and after a cut on the binned neural network response for the numerator, N_W, of R_n, where MC has been scaled to an integrated luminosity of 30 fb^{-1} to make the leptoquark signal discernible. Dashed

lines correspond to SM plus a leptoquark signal with the three different choices of m_{LQ} while the points correspond to SM prediction alone. Since leptoquarks must have large masses compared to SM particles signal resolution in R_n is optimal in the k_T range of 500-700 GeV, where Z+jets is still statistically significant. Table 6.5 shows the significance ($S/\sqrt{S+B}$) for the three leptoquark samples, before and after the binned neural network cut, where signal and background have been calculated after W+jets selection. The cut offers substantial improvement on the sensitivity of the R_n search from which the significance is enhanced by a factor of 7.99, 15.6 and 24.3 for m_{LQ} = 600, 800 and 1200 GeV signals respectively.

$LQL\bar{Q}$ sample mass (GeV)	Before NN cut	After NN cut
m_{LQ} =600	0.148	1.18
m_{LQ} =800	1.38e-2	0.216
m_{LQ} =1200	2.32e-4	5.65e-3

Table 6.5: $LQL\bar{Q}$ significance before and after NN cut. Signal and background are calculated after W+jets selection

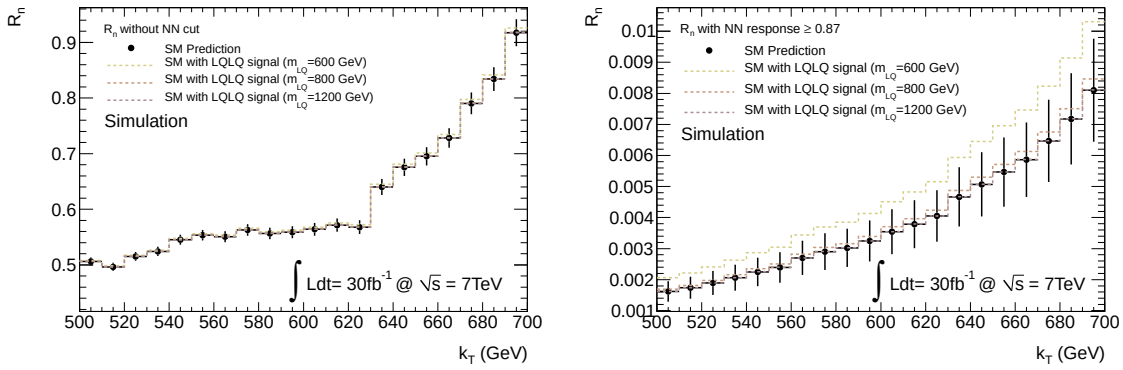

Figure 6.10: R_n with (right) and without (left) binned neural network response cut. Points correspond to SM prediction and dashed lines to SM with LQLQ signal. MC has been scaled to an integrated luminosity of 30 fb^{-1}.

6.4 Effects of signal enhancement on systematic uncertainties

Even though systematic uncertainties have not been studied in this analysis it is worth considering what types of effects the signal enhancement will have on the precision with which R_n can be measured. As discussed in Section 2.2 to extract any signal in R_n requires an excess over the expected events in the numerator or denominator of R_n; which would be measured as an upward or downward deviation in R_n as a function of k_T. Methodology aside, enhancement of any signal in R_n will mean making selection cuts in either the numerator or denominator placing them in different phase spaces. In this case it is not clear to what extent cancellation of the systematic uncertainties listed in Section 2.2 will occur; however an overall decrease in cancellation is expected.

In general, one would not expect much in the way of cancellation of generator uncertainties, as the signal may be derived from a very different process than the W/Z+jets background. Thus PDFs, hadronization models and renormalization and factorization scale may vary. Indeed, different types of generators used by the ATLAS Collaboration specialize in simulating specific physical processes.

Jet energy scale and resolution uncertainties can cause bin migration in the numerator or denominator of R_n. If R_n is constant as a function of k_T bin migration will only affect the measurement by a linear transformation in k_T, which would not create any signal-like effects in R_n. However to enhance a signal in R_n requires background in the numerator or denominator to be removed. This leads to an upward or downward deviation in R_n, where bin migration could possibly enhance or partially remove this signal.

Uncertainties related to corrections that are applied as a constant scaling factor to all datasets, such as the luminosity scaling, will still cancel in R_n. Also uncertainties associated with quantities that are uncorrelated with the classifier, i.e. have similar distributions both before and after the cut, will still be expected to cancel.

By imposing cuts on R_n, which are designed to enhance a specific signal, the analysis moves from a model-independent search to a model-dependent search. In doing so the precision in which R_n can be measured is traded for signal significance. This thesis has explored the extent to which R_n can be made model-dependent; however, in such an extreme case R_n would not be expected to be more sensitive than a dedicated search.

Chapter 7

Conclusion

In this thesis a measurement of R_n is presented and its usefulness as a probe for new physics signals is explored. As a demonstration of the method, pair production of top quarks and leptoquarks is enhanced in R_n using optimized discriminants derived from a set of topology-discriminating variables.

A measurement of R_n as a function of total jet p_T in the 1,2,3 and 4 jet multiplicity bins was made using 32.6 pb^{-1} of collected ATLAS data from the 2010 runs. The measurement agrees reasonably well with prediction for all four jet multiplicities despite the fact that systematic uncertainties have not been evaluated.

A common set of topology-discriminating variables is chosen, based on signal-background separation and minimal correlation. Performance of the multi-dimensional cuts, fisher and neural network classifiers are evaluated and compared. The neural network outperforms the other two classifiers for both analyses and is even further improved by determining multiple responses, calculated in one of five non-overlapping sum p_T cluster regions. An optimal cut on the binned neural network response in the $t\bar{t}$ and $LQL\bar{Q}$ analysis enhanced signal significance by a factor of 2.86 and 7.99-24.3 (depending on leptoquark mass) respectively. However, the improved signal sensitivity in R_n comes at the price of reduced precision. Since a comprehensive study of systematic uncertainties is beyond the scope of this thesis it is not clear to what extent the precision in which R_n can be measured is degraded.

Bibliography

[1] P. Langacker, "The Standard Model and Beyond," *CRC press.*

[2] Particle Data Group, "Review of Particle Physics," *Physics Letters B Vol 667/1-5 (2008) 1-1340.*

[3] Wikimedia commons *http://commons.wikimedia.org/wiki.*

[4] V. Barger, R. Phillips, "Collider Physics," *Westview Press.*

[5] E. Noether, "Invariante variationsprobleme," *Nachr. v. d. Ges. d. Wiss. zu Gattingen, (235- 257), 1918.*

[6] V. Barger, R. Phillips, "Collider Physics," *p.40, Westview Press.*

[7] H. Georgi, "Lie algebras in particle physics: from isospin to unified theories," *Benjamin/Cummings Pub. Co.*

[8] The CTEQ Collabortaion *http://www.phys.psu.edu/ cteq/.*

[9] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne *http://durpdg.dur.ac.uk/hepdata/mrs.html.*

[10] M. Glck, E. Reya, and A. Vogt *http://zebu.uoregon.edu/ parton/partonGRV.html.*

[11] M. Glck, P. Jimenez-Delgado, and E. Reya *http://durpdg.dur.ac.uk/hepdata/grv.html.*

[12] The NNPDF Collaboration *http://sophia.ecm.ub.es/nnpdf/nnpdf_pdfsets.htm.*

[13] Hung-Liang Lai, Marco Guzzi, Joey Huston, Zhao Li, Pavel M. Nadolsky, Jon Pumplin, and C.-P. Yuan, "New parton distributions for collider physics," *Phys.Rev. D82 (2010) 074024.*

[14] H. Georgi and S.L. Glashow, "Unity of all elementary particle forces," *Phys. Rev. Lett. 32: 438 (1974).*

[15] J. Pati and A. Salam, "Lepton number as the fourth color," *Phys. Rev. D10 (1974), 275.*

[16] B. Schrempp and F.Schrempp, "Light leptoquarks," *Phys Lett. 153B (1985) 101.*

[17] S. Dimopoulos and L. Susskind, "Weak Interactions with Lepton-Hadron Symmetry," *Nucl. Phys. B91 (1981) 370.*

[18] P. Langacker, "The Standard Model and Beyond, p.505-506," *CRC press.*

[19] The ATLAS Collaboration, "Expected Performance of the ATLAS Experiment Detector, Trigger and Physics, Volume III," *CERN-OPEN-2008-020.*

[20] Geoffrey C. Fox, Stephen Wolfram, "Event shapes in e^+e^- annihilation," *Nuclear Physics B149 (1979) 413-496.*

[21] The LHC Collaboration, "LHC main website," *http://lhc.web.cern.ch/lhc/.*

[22] CERN Communication Group, "CERN faq LHC the guide," *CERN-Brochure-2008-001-Eng.*

[23] The LHC Collaboration, "LHC public outreach website," *http://lhc-machine-outreach.web.cern.ch.*

[24] The ATLAS Collaboration, "Detector and physics performance technical design report," *CERN/LHCC/99-015.*

[25] The ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider," *2008 JINST 3 S08003.*

[26] The ATLAS Collaboration, "ATLAS Website," *http://www.atlas.ch/photos/index.html.*

[27] The ATLAS Collaboration, "ATLAS Inner detector: technical design report. 1," *CERN-LHCC-97-016, http://cdsweb.cern.ch/record/331063.*

[28] The ATLAS Collaboration, "ATLAS Inner detector: technical design report. 2," *CERN-LHCC-97-017, http://cdsweb.cern.ch/record/331064.*

[29] The ATLAS Collaboration, "ATLAS pixel detector: technical design report," *CERN-LHCC-98-013, http://cdsweb.cern.ch/record/381263.*

[30] A. A. et al., "The barrel modules of the ATLAS semiconductor tracker," *Nucl. Instrum. Meth. A 568 (2006) 642.*

[31] A. A. et al., "The ATLAS semiconductor tracker end-cap module," *Nucl. Instrum. Meth. A 575 (2007) 353.*

[32] E. A. et al., "The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance," *2008 JINST 3 P02013.*

[33] The ATLAS Collaboration, "Liquid argon calorimeter technical design report," *CERN-LHCC-96-041.*

[34] The ATLAS Collaboration, "Tile calorimeter technical design report," *CERN-LHCC-96-042.*

[35] The ATLAS Collaboration, "The ATLAS forward calorimeters," *2008 JINST 3 P02010.*

[36] The ATLAS Collaboration, "Magnet system technical design report," *CERN-LHCC-97-018, http://cdsweb.cern.ch/record/338080.*

[37] The ATLAS Collaboration, "The ATLAS Data Acquisition and Trigger: concept, design and status," *doi:10.1016/j.physletb.2003.10.071.*

[38] Matteo Cacciari, Gavin P. Salam and Gregory Soyez, "The anti-kt jet clustering algorithm," *JHEP04(2008)063 doi:10.1088/1126-6708/2008/04/063.*

[39] The ATLAS Collaboration, "Good Run Lists," *http://atlasdqm.web.cern.ch/atlasdqm/grlgen/StandardModel/.*

[40] The ATLAS Collaboration, "Updated Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector," *ATLAS-CONF-2011-011.*

[41] The ATLAS Collaboration, "Alpgen FAQ," *https://twiki.cern.ch/twiki/bin/view/Main/AlpgenFAQ.*

[42] P. M. Nadolsky et al., "Implications of CTEQ global analysis for collider observables," *Phys. Rev. D78 (2008) 013004.*

[43] G. Corcella et al., "HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes)," *JHEP 0101:010,2001.*

[44] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, "Multiparton interactions in photoproduction at HERA," *Z. Phys. C72 (1996) 637.*

[45] A. Sherstnev and R. Thorne, "Parton Distributions for LO Generators," *Eur. Phys. J. C55 (2008) 553575, arXiv:0711.2473 [hep-ph].*

[46] S. Frixione and B.R. Webber, "Matching NLO QCD computations and parton shower simulations," *JHEP 0206 (2002) 029 [hep-ph/0204244].*

[47] P. Golonka and Z. Was, "PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays," *Eur. Phys. J. C45 (2006) 97.*

[48] N. Davidson et al., "Universal interface of TAUOLA technical and physics documentation," *arXiv:1009.2338 [hep-ph].*

[49] The ATLAS Collaboration, "Performance of primary vertex reconstruction in proton-proton collisions at s = 7 TeV in the ATLAS experiment," *ATLAS-CONF-2010-069.*

[50] S. Agostinelli et al., "GEANT4 - a simulation toolkit," *Nucl. Inst. and Meth A506 (2003) 250.*

[51] The ATLAS Collaboration, "Measurement of the cross-section for jets produced in association with a W-boson in *pp* collisions at $\sqrt{7}$ TeV," *Phys.Lett.B698:325-345,2011.*

[52] The ATLAS Collaboration, "How to Clean Jets," *https://twiki.cern.ch/twiki/bin/view/AtlasProtected/HowToCleanJets.*

[53] The ATLAS Collaboration, "A measurement of the ratio of the W and Z cross sections with exactly one associated jet in pp collisions at sqrt(s) = 7 TeV with ATLAS," *arXiv:1108.4908v1.*

[54] The ATLAS Collaboration, "Pile-up reweighting procedure," *https://twiki.cern.ch/twiki/bin/view/AtlasProtected/WZElectroweakCommonTopics.*

[55] D. Lopez Mateos, E. W. Hughes, and A. Schwartzman, "A Simple pT - and h-Dependent Monte Carlo-Based Jet Calibration," *ATL-PHYS-INT-2009-077.*

[56] The ATLAS Collaboration, "Jet selection based on the jet-vertex fraction (JVF)," *https://twiki.cern.ch/twiki/bin/view/Atlas Protected/JetQualityAndSelectionForPileup.*

[57] The ATLAS Collaboration, "Jet energy scale and its systematic uncertainty for jets produced in proton-proton collisions at $\sqrt{s} = 7$ TeV and measured with the ATLAS detector," *ATLAS-CONF-2010-056.*

[58] T. Mitchell, "Machine Learning," *McGraw Hill, (1997).*

[59] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss, "TMVA 4 Users Guide, Section 8.10," *CERN-OPEN-2007-007.*

[60] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss, "TMVA 4 Users Guide," *CERN-OPEN-2007-007.*

[61] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss, "TMVA 4 Users Guide, Section 8.7," *CERN-OPEN-2007-007.*

[62] R. Fisher, "The Use of Multiple Measurements in Taxonomic Problems," *Annals of Eugenics, 7, p. 179–188.*

[63] The ATLAS Collaboration, "Search for pair production of first or second generation leptoquarks in protonproton," *CERN-PH-EP-2011-050.*

Appendix A

Appendix

A.1 Sensitivity of topology-discriminating variables to pile-up

In-time pile-up refers to the situation where there are multiple proton-to-proton interactions in a single bunch crossing, causing the event to have multiple primary vertices. The number of primary vertices per bunch crossing is Poisson-distributed with a mean determined by the beam parameters. The amount of pile-up in an event can be estimated by the number of reconstructed primary vertices. However, it should be noted that heavier particles that decay inside the inner detector, such as τ leptons, can also have vertices that are reconstructed as primary vertices; thus the number of primary vertices can differ from the number of proton-to-proton interactions. Pile-up in an event can have a significant impact on the shape of the topology-discriminating variables. The number of topoclusters in an event will increase with pile-up adding to the number of topoclusters used to calculate the variable. If the extra topoclusters are mostly contained in a certain fiducial region the topology of the event can be dramatically altered. Certain variables are more sensitive to pile-up effects than others, especially those that are a function of the relative position of topoclusters to one another (or are a function of $\Delta\phi$ and/or $\Delta\eta$) such as ΔR and C_{max}. Fig. A.1 displays how the ΔR moment distribution changes as a function of the number of primary vertices, N_{vtx} . Similar figures for variables N_{jets}, C_{max}, S_T and Thrust can be found in this appendix. In the limit of infinite pile-up all signal would lost under the pile-up energy deposits in the calorimeter. Thus pile-up has the effect of reducing the discriminating power of the topology-discriminating variables. To model pile-up

accurately in MC it is important to use the vertex weight factors found in Table 4.2 when comparing data to MC.

A.2 Additional figures for sensitivity of topology-discriminating variables to $\sum p_T^{clust}$ and pileup

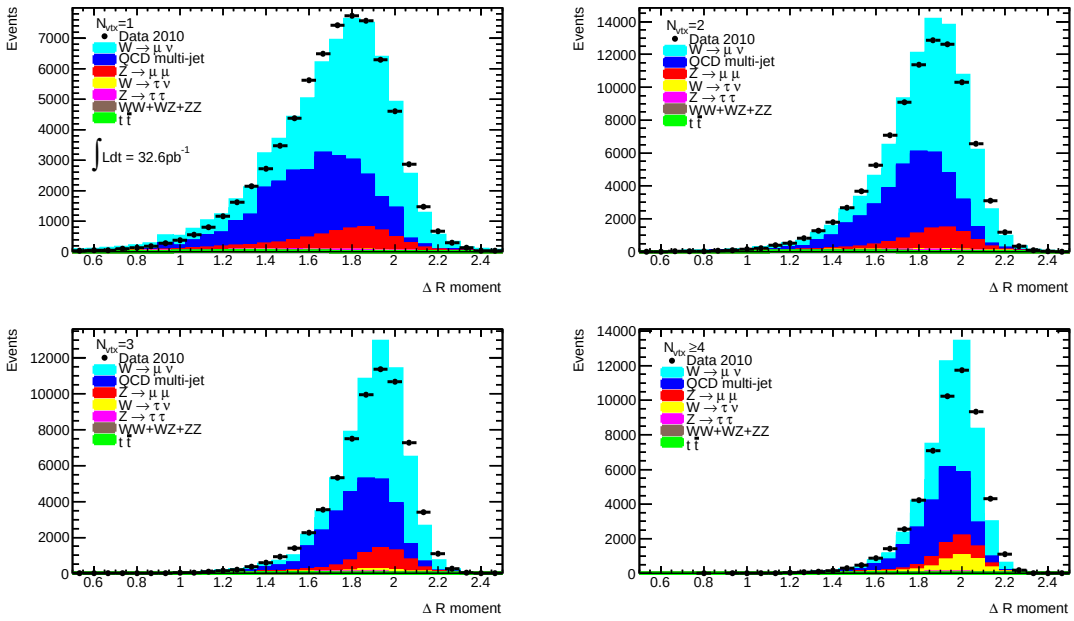

Figure A.1: ΔR distribution calculated in increasing bins of number of primary vertices (N_{vtx}). For further details see caption of Fig. 6.1.

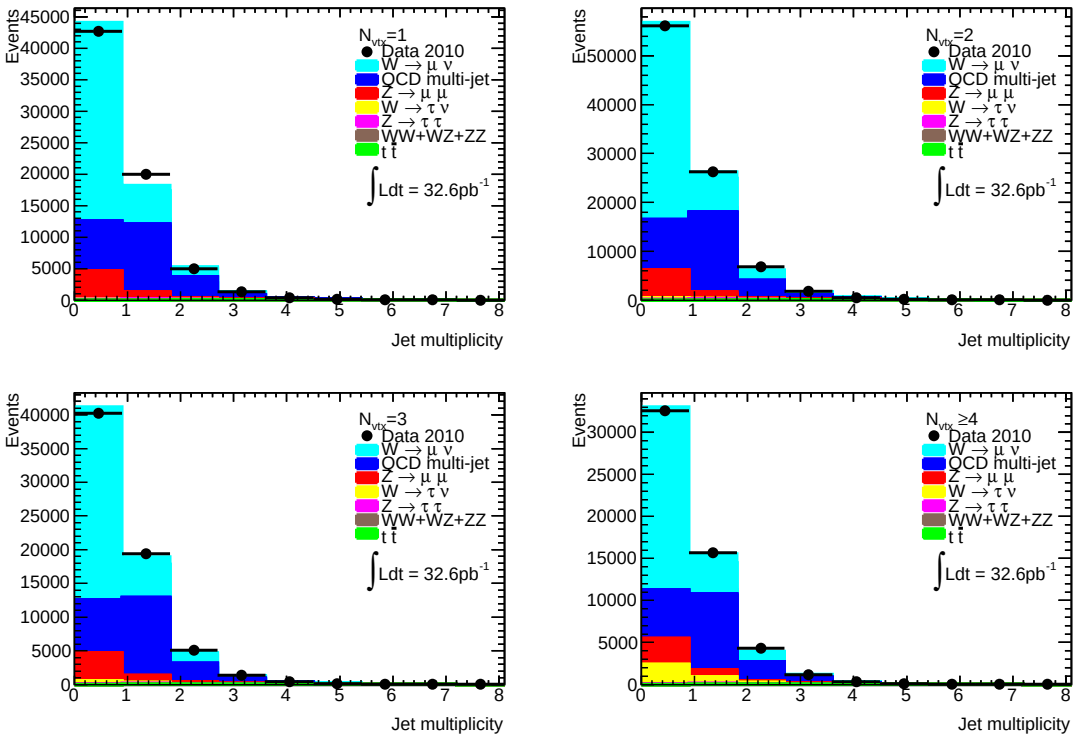

Figure A.2: N_{jets} distribution calculated in bins of number of primary vertices (N_{vtx}). For further details see caption of Fig. 6.1.

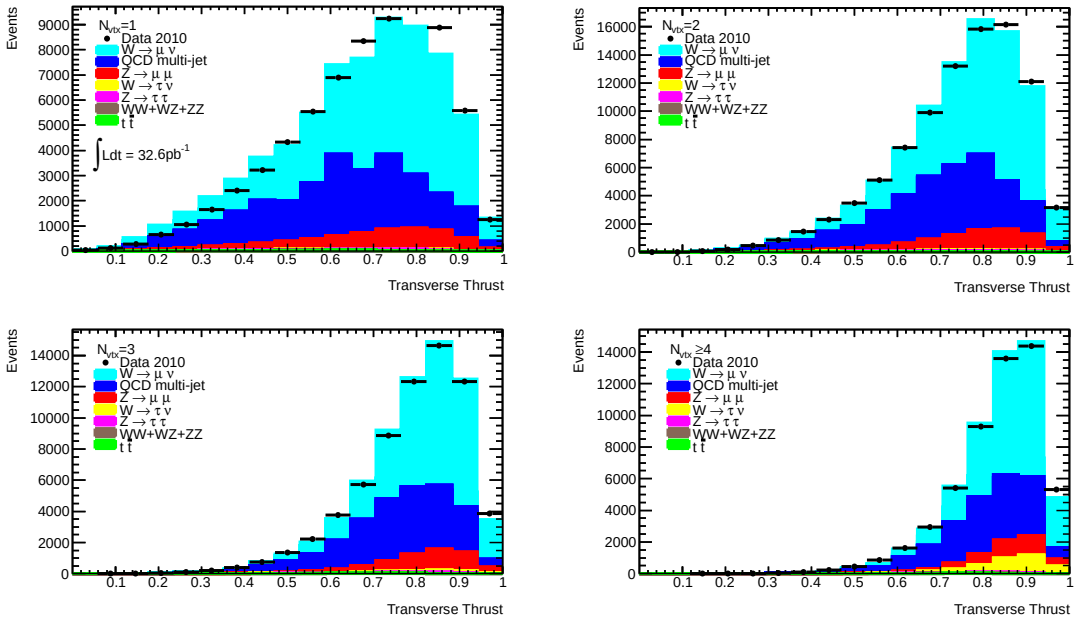

Figure A.3: Thrust distribution calculated in bins of number of primary vertices (N_{vtx}). For further details see caption of Fig. 6.1.

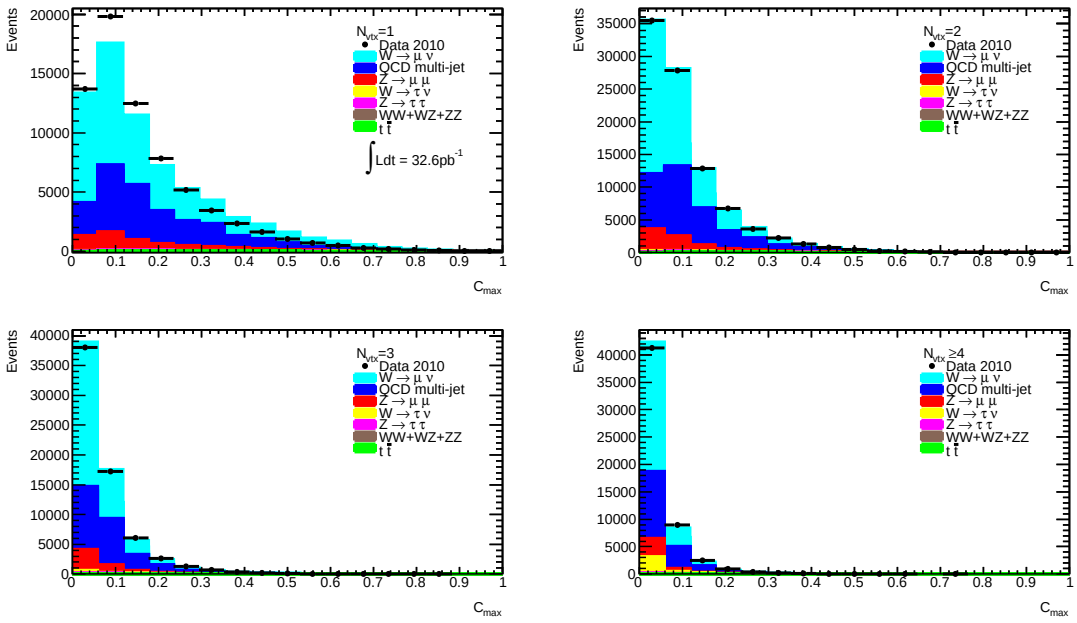

Figure A.4: C_{max} distribution calculated in bins of number of primary vertices (N_{vtx}). For further details see caption of Fig. 6.1.

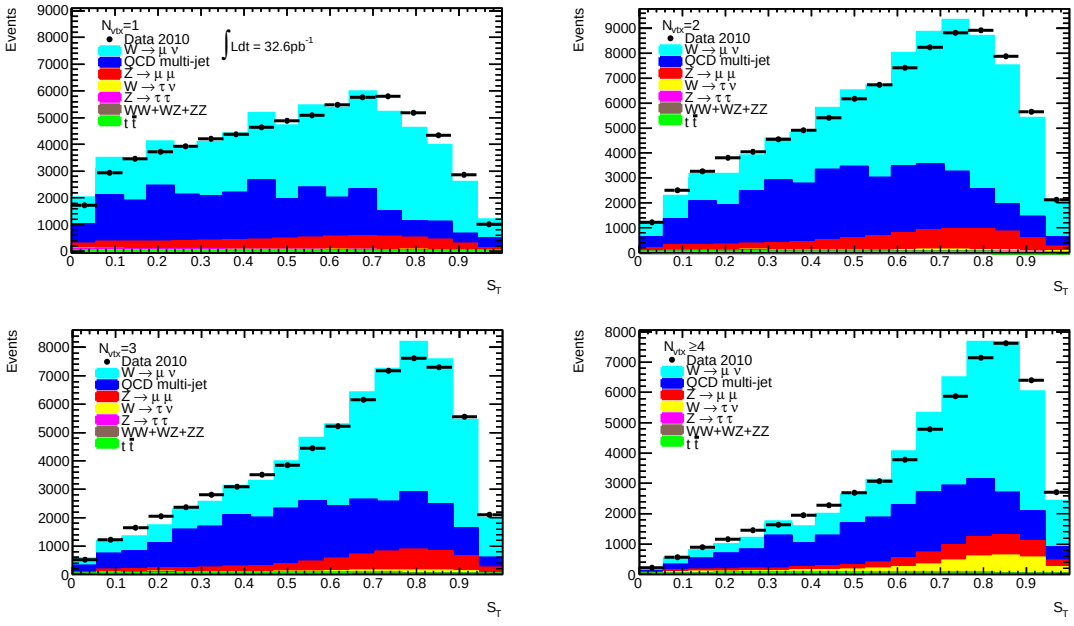

Figure A.5: S_T distribution calculated in bins of number of primary vertices (N_{vtx}). For further details see caption of Fig. 6.1.

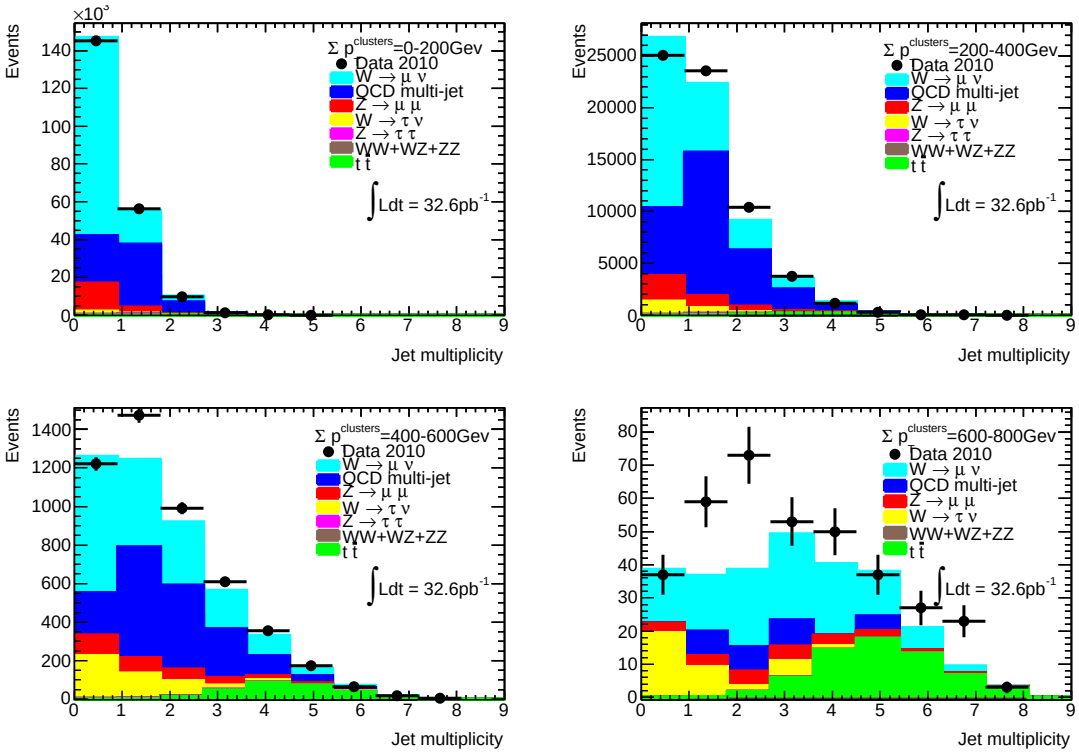

Figure A.6: N_{jets} distribution calculated in bins of increasing $\sum p_T^{clust}$. For further details see caption of Fig. 6.1.

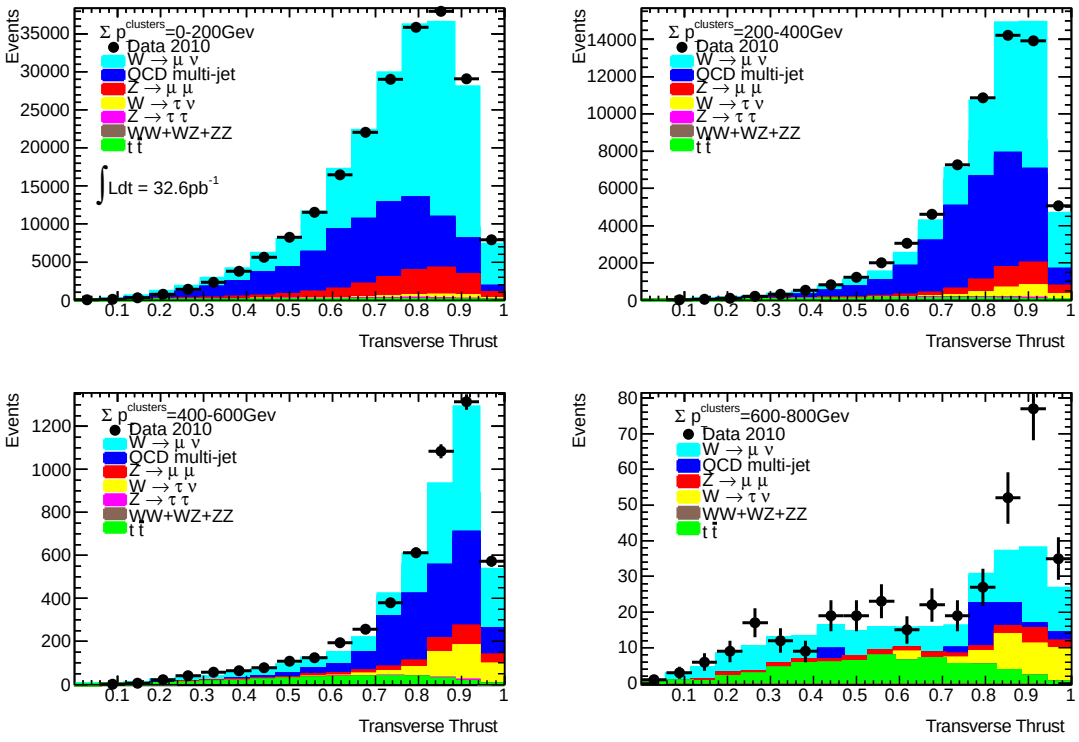

Figure A.7: Thrust distribution calculated in bins of number of increasing $\sum p_T^{clust}$. For further details see caption of Fig. 6.1.

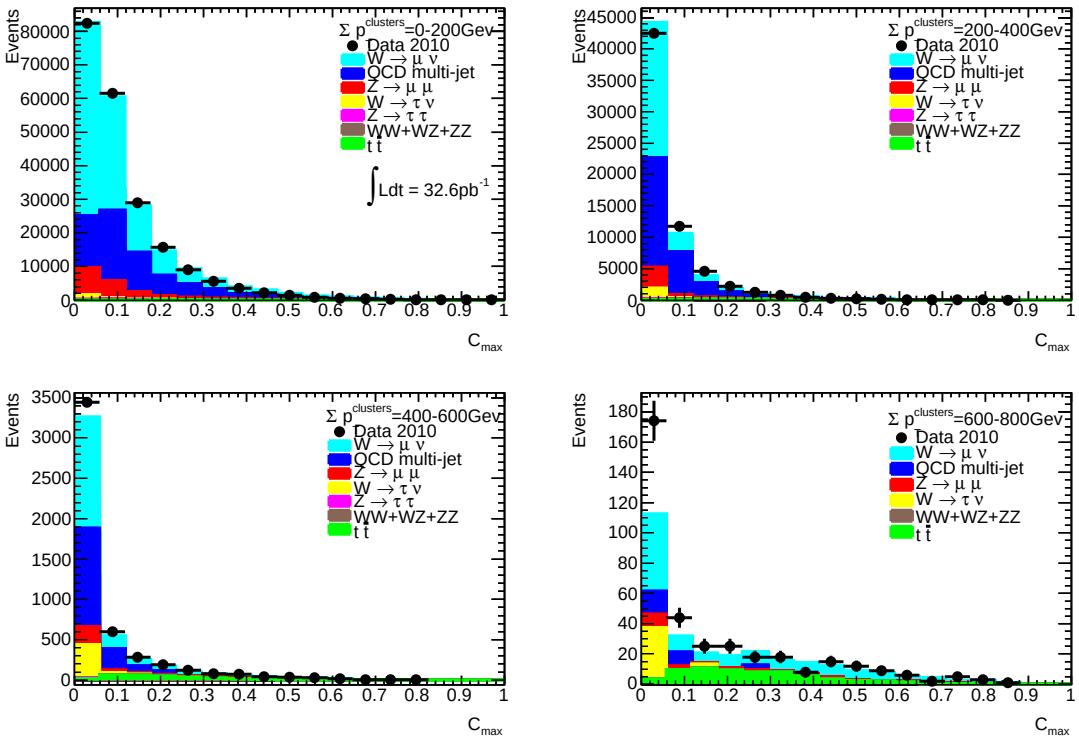

Figure A.8: C_{max} distribution calculated in bins of increasing $\sum p_T^{clust}$. For further details see caption of Fig. 6.1.

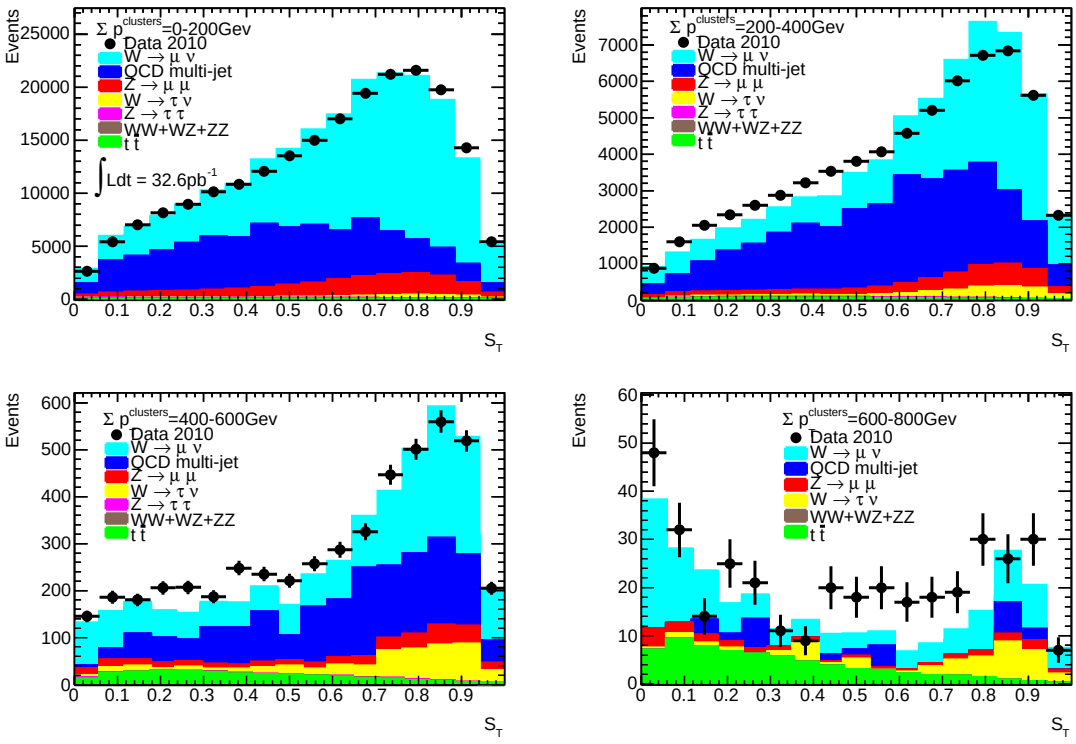

Figure A.9: S_T distribution calculated in bins of increasing $\sum p_T^{clust}$. For further details see caption of Fig. 6.1.