

# First Year Transfer Report

James Keates  
University of Manchester

May 24, 2006

## **Abstract**

In this first year report I present my study of the simulation of  $W$ +jets events at the LHC in a range of parton shower and matrix element Monte Carlo Event Generators. This study looks specifically at the simulation of these events as a background to  $WW$  scattering. The tag jets in high pseudorapidity regions are a particular area of interest. Matrix element matching in tag jets is discussed. I also introduce future work on jet gap events from a theoretical point of view.

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
1.1	The Standard Model . . . . .	4
1.1.1	QCD and Confinement . . . . .	5
1.1.2	The Electroweak Sector and Symmetry Breaking . . . . .	5
1.1.3	Beyond the Standard Model . . . . .	6
1.2	The LHC and the ATLAS Experiment . . . . .	6
1.3	Monte Carlo Generators . . . . .	7
1.3.1	Hard Process . . . . .	8
1.3.2	Parton Shower . . . . .	8
1.3.3	Hadronisation . . . . .	10
1.3.4	Decays, Beam Remnants and Multiple Interactions . . . . .	12
1.3.5	Jet Finding . . . . .	12
1.3.6	Matrix Element Matching . . . . .	14
<b>2</b>	<b><math>WW</math> Scattering</b>	<b>15</b>
2.1	Motivation . . . . .	15
2.2	Previous Study of $WW$ Scattering . . . . .	16
2.2.1	Leptonic Cuts . . . . .	16
2.2.2	Hadronic Cuts and Subjet Analysis . . . . .	17
2.2.3	Tag Jet Analysis . . . . .	17
2.3	Background Simulation . . . . .	18
<b>3</b>	<b><math>W</math>+Jets Simulation Analysis</b>	<b>19</b>
3.1	Results . . . . .	20
3.2	Colour Coherence in Herwig and Pythia . . . . .	20
3.3	Comparison of Parton Shower and Matrix Element Programmes . . . . .	24
3.4	Matching Issues in Alpgen . . . . .	25

<b>4</b>	<b>Conclusions</b>	<b>28</b>
4.1	Future Plans . . . . .	28
4.1.1	$W$ +Jets Improvements . . . . .	28
4.1.2	Gaps Between Jets Processes in QCD . . . . .	29

# List of Figures

1.1	Diagram of $q \rightarrow qg$ branching showing momentum labels and a parton shower including all possible splittings . . . . .	9
1.2	Schematic diagram of string and cluster fragmentation schemes . . . . .	11
2.1	$WW$ Scattering and $W$ +jet Background . . . . .	16
3.1	$WW$ Scattering cut variables produced with Herwig . . . . .	20
3.2	$WW$ Scattering cut variables produced with Pythia . . . . .	21
3.3	$WW$ Scattering cut variables produced with Sherpa . . . . .	21
3.4	$WW$ Scattering cut variables produced with Alpgen + Herwig . . . . .	22
3.5	$WW$ Scattering cut variables produced with Alpgen + Pythia . . . . .	22
3.6	Pseudorapidity Distributions for tag jets produced in (a) Herwig, (b) Pythia and (c) Pythia without colour coherence . . . . .	23
3.7	Alpha Distributions for tag jets produced in Herwig (red), Pythia (blue) and Pythia without colour coherence (black) . . . . .	24
3.8	Pseudorapidity distribution produced in Sherpa. . . . .	25
3.9	Rapidity distribution of tag jets produced in Alpgen with $ \eta $ of the ME jets $< 2.5$ for (a) 1 jet, (b) 2 jets, (c) 3+ jets, (d) fully inclusive sample. . . . .	26
3.10	Rapidity distribution of tag jets produced in Sherpa with $ \eta $ of the ME jets $< 2.5$ . . . . .	26

# Chapter 1

## Introduction

At the Large Hadron Collider (LHC), the  $WW$  scattering process will provide a potential discovery channel for many interesting processes. I have studied the  $W$ +jets background to  $WW$  scattering using Monte Carlo event generators with different properties in order to see how well we can simulate this process. I also attempted to understand the physical reasons behind any differences seen in the simulated samples.

The rest of this section briefly reviews the Standard Model of particle physics, noting especially areas that will be of later interest in this report. The LHC, and specifically the ATLAS experiment are introduced, with a short description of the detectors. There is also an overview of the workings of Monte Carlo generators, of both the parton shower and matrix element types, and some discussion of the different approaches used.

### 1.1 The Standard Model

The Standard Model of particle physics consists of three gauge symmetry groups combined as  $SU(3)_C \times SU(2)_L \times U(1)_Y$  [1]. The strong interactions are contained in  $SU(3)_C$  and the electroweak interactions are in the symmetry group  $SU(2)_L \times U(1)_Y$ , where the symmetry is broken.

Within the Standard Model there are twelve fermionic particles with half integral spin which make up matter, six flavours of quark, three charged leptons and three massless neutrinos. A corresponding set of anti-particles with equal and opposite quantum numbers (e.g. charge or spin) also exist. The forces are mediated by gauge bosons, the gluon for the strong interactions and the  $W^+$ ,  $W^-$ ,  $Z$  and photon for the electroweak interactions. In order to give mass to the  $W$  and  $Z$  bosons, as observed at experiments, the electroweak symmetry

must be spontaneously broken. Within the standard model this is done using the Higgs mechanism, which requires an as yet undiscovered particle, known as the Higgs boson.

### 1.1.1 QCD and Confinement

The strong interaction is described in the Standard Model by the theory of quantum chromodynamics (QCD) [2]. The charges (similar to the electric charge in QED) of the strong force are called colours. The interactions occur through the exchange of gluons, which can also interact with themselves as they carry colour. There are three colours of quark (traditionally called red, green and blue or  $rgb$ ) and eight gluons.

QCD has a couple of important properties, most notably asymptotic freedom and confinement. Both properties originate from the running of the strong coupling,  $\alpha_s$ . The coupling strength grows with distance, which means at short range the forces between coloured objects are small and their interactions can be treated perturbatively. The partons in this regime (quarks and gluons) are described as asymptotically free.

At large distances (or alternatively low energies) the strength of the forces become large. Sufficient energy will build up between two partons moving apart that  $q\bar{q}$  pairs can be created. This shields the original partons from each other. As a result all coloured particles are confined to live within colourless hadrons, either baryons made of three quarks or mesons consisting of a quark-antiquark pair. When a quark or gluon is emitted in an interaction many particle pairs will be created. In experiments these are seen as jets of hadrons, which we associate with a parton in the original reaction.

### 1.1.2 The Electroweak Sector and Symmetry Breaking

The electroweak sector of the Standard model comes from the  $SU(2)_L \times U(1)_Y$  symmetry group, but this predicts massless vector bosons and we know from experiment that the  $W^\pm$  and  $Z$  bosons have mass.

Within the standard model the mass is a result of spontaneous symmetry breaking of the electroweak symmetry. This breaking is provided by the Higgs mechanism. In the Higgs mechanism a complex scalar field (the Higgs field) is introduced which interacts with particles to provide mass. It gives mass to the massive vector bosons and fermions. The hadron masses derive mainly however not from their constituent quark's Higgs mass but from other strong interactions.

An important prediction of introducing this new field is a massive particle associated with it, the famous Higgs boson. This is the last unobserved particle of the Standard Model and it is one of the major goals for experiments at the moment and in the near future to make a discovery of the Higgs.

### 1.1.3 Beyond the Standard Model

Although the Standard Model would seem to be complete if we observe the Higgs, there are still properties of it which remain unexplained. Therefore there have been many theoretical attempts to look 'Beyond the Standard Model' (BSM). This has produced a range of theories, many of which will be testable at the LHC (and the future International Linear Collider).

Probably the favoured BSM theory is supersymmetry, in which there is a symmetry between fermions and bosons, with every fermion gaining a new bosonic super-partner and vice versa. The simplest form of this is known as the MSSM, Minimal Supersymmetric Standard Model, but more complex theories also exist. In the MSSM there are five Higgs particles, some of them charged, so these Higgs will also be sought out at colliders, along with the aforementioned superpartners.

## 1.2 The LHC and the ATLAS Experiment

The Large Hadron Collider (LHC) is a proton proton collider experiment currently under construction at CERN, near Geneva in Switzerland. The protons will be collided at a centre of mass energy of  $\sqrt{s}=14$  TeV allowing new energy ranges to be explored for the first time. There are four detectors based at the LHC, ALICE, a heavy ion experiment, LHCb, which will study physics in the B-system and two general purpose experiments, CMS and ATLAS.

ATLAS (A Toroidal LHC ApparatuS) [3], is a cylindrical detector designed for detecting many possible products of a collision, but especially focused on the Higgs and BSM processes. It consists of an inner tracking detector, electromagnetic and hadronic calorimeters and muon chambers surrounding the others. Detectors are placed concentrically around the beam pipe, known as the barrel region and over the ends of the 'barrel', called the endcaps. This provides nearly complete coverage, however there is still an unavoidable gap along the line of the beam pipe.

The ATLAS inner detector consists of, from nearest the beam pipe outwards, the Pixel Detector, Semiconductor Tracker (SCT) and Transition Radiation

Tracker (TRT). These are enclosed in a solenoid magnet so momentum can be reconstructed from the curvature of the tracks.

The Pixel detector consists of silicon semiconductor arranged in pixels, and the SCT has strips of silicon. The passage of a charged particle causes ionisation in the silicon producing an electron-hole pair which drift in the electric field and cause a current which is detected. The TRT uses straw tubes which work similarly to silicon detectors. In the case of straw tubes it is a gas which is ionised, the electric field being provided by a wire inside the tube. Tracks are reconstructed from hits in these three detectors in order to measure both momentum and the vertex position.

Outside the inner detector are the calorimeters. First the electromagnetic liquid argon (LAr) calorimeter which detects the energy of electrons and photons. Outside this is the hadronic tile calorimeter, designed to detect energy in hadrons. A liquid argon calorimeter is used in the endcap regions for both electromagnetic and hadronic energy. In both cases layers of material designed to start showers are interspersed with active material to measure the energy. The idea is that all the energy of a particle is constrained within the calorimeters.

Finally come the muon chambers, which detect muons, the only standard model particles penetrating enough to be detected this far out. These are also enclosed in toroidal magnets in order to get a momentum measurement for the muons.

Neutrinos pass through the detector unseen and must be reconstructed by calculating the missing momentum in an event and assigning it to neutrinos. The uncertainty in the  $z$  component (defined to be parallel to the beam) of neutrino momentum gives rise to a two fold uncertainty in the reconstructed momentum of a leptonically decaying  $W$ .

### 1.3 Monte Carlo Generators

In order to simulate the interactions in a collider experiment we use random number based computer programmes known as Monte Carlo event generators. These work on three basic stages, a hard interaction calculated exactly from a matrix element and a parton shower stage which adds radiation down to some scale where the coloured partons must be hadronised into the particles we really see in a detector.

In this study I will look at two generators which include  $2 \rightarrow 2$  hard processes and then parton showers, Herwig [4] and Pythia [5], which I shall refer to as parton shower (PS) programmes. I will compare these to newer Matrix Element

(ME) generators which calculate the exact matrix elements for a high number of partons, often corresponding to a required number of jets. These must then also be parton showered and hadronised while being careful to avoid any double counting. The self contained programme Sherpa [6] was mainly used and some comparison was given with Alpgen [7] which links to either Herwig or Pythia to provide its parton shower stage.

### 1.3.1 Hard Process

Both Herwig and Pythia contain libraries of hard  $2 \rightarrow 2$  processes, containing information such as differential cross-sections and colour and spin information. These are important both in creating the desired particles and calculating cross-sections, but also play a role in the initialisation of the parton shower stage. The processes of interest in this report are  $q + g \rightarrow q + W$  and  $q + \bar{q} \rightarrow W + g$ . The extra jets required in this study are produced in the parton shower stage described below.

An important consideration is that of parton distribution functions (PDFs), which characterise the partonic nature of the initial hadrons, two protons at the LHC. The PDFs normally plot the distribution of  $x$ , effectively the momentum fraction of the hadron carried by a specific parton type against the scale at which the hadron is probed,  $Q^2$ . They cannot be derived from first principles so are produced from experimental data. It is possible however, to evolve the distributions to any required  $Q^2$  once it is measured at one value. The initial state parton shower links the hard subprocess with the PDFs of the beam particles.

For the matrix element programmes the required jets are produced by an exact tree-level matrix element calculation, up to a specified maximum number. In order to produce jets of hadrons these partonic states are showered, where the parton shower is allowed to add extra emissions. However in order to avoid double counting these emissions must not be resolvable as separate jets. Therefore it is necessary for some matching procedure to be implemented as described below.

### 1.3.2 Parton Shower

The parton shower is based on their being a probability of partons branching, for example a quark emitting a gluon, or gluon forming a  $q\bar{q}$  pair. The probability depends on the splitting functions,  $P_{a \rightarrow bc}$  listed below [2]

$$P_{g \rightarrow gg} = 2C_A \left[ \frac{1-z}{z} + \frac{z}{1-z} + z(1-z) \right] \quad (1.1)$$

$$P_{q \rightarrow qg} = C_F \frac{1+z^2}{1-z} \quad (1.2)$$

$$P_{g \rightarrow q\bar{q}} = T_R [z^2 + (1-z)^2] \quad (1.3)$$

where  $C_A=3$ ,  $C_F=4/3$  and  $T_R=1/2$  are colour factors derived from QCD group theory and  $z$  is the proportion of energy taken by the split parton, b. The kinematics are illustrated for  $q \rightarrow qg$  in figure 1.1.

It must also be taken into account that the parton can't have split at a higher scale, and this introduces a Sudakov form factor. This is an exponential term which describes the probability of the parton having evolved to the branching scale without a previous splitting. The exact definition depends on choice of evolution variables within a programme (t in equation 1.4 is the virtuality), but all are of the approximate form

$$\Delta(t) \equiv \exp \left[ - \int_{t_0}^t \frac{dt'}{t'} \int dz \frac{\alpha_s}{2\pi} P(z) \right] \quad (1.4)$$

where  $P(z)$  is the appropriate splitting function.

As a result an evolution down in scale occurs over a number of branchings until reaching some minimum scale at which non-perturbative effects become important and hadronisation occurs. This is normally of the order of 1 GeV.

A similar backward evolution occurs to recreate the initial state radiation. This starts from the incoming partons of the hard process and at each stage attempts to recreate which splitting the parton originated from. As the particles involved here have negative virtuality  $Q^2 = -m^2$  it is often called a space-like shower as opposed to the time-like evolution of the final state. This takes the hard parton interaction and relates it to the colliding protons, where it is combined with the PDFs discussed in the previous section. This initial state shower can also produce outgoing radiation, which often proves to be harder than the final state emissions.

The evolution variable chosen differs between the programmes. Pythia and Sherpa choose the virtuality  $Q^2$  whereas Herwig uses an angular variable  $\frac{E^2}{\xi_{bc}}$

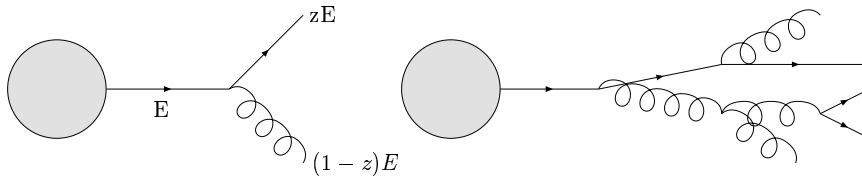


Figure 1.1: Diagram of  $q \rightarrow qg$  branching showing momentum labels and a parton shower including all possible splittings

where  $\xi_{bc} = \frac{(p_b \cdot p_c)}{(E_b E_c)}$  ( $\simeq \frac{1}{2} \theta_{bc}^2$  for small angles) in the final state and  $\theta_i$  the angle between the incoming hadron and emitted parton,  $i$ , in the space-like evolution. It is a property of the branching that the opening angle is always smaller than the previous splitting so Pythia imposes an additional veto on radiation at larger angles, whereas Herwig has this property automatically resulting from the choice of evolution variable.

Both Herwig and Pythia contain matrix element corrections to the parton showers for certain processes. The parton shower method only reproduces the soft and collinear regions, where branching is logarithmically enhanced. Although these regions are favoured this approach produces a 'dead cone' outside the angular ordered region, which it is necessary to populate to fully reproduce experimental effects. Therefore a correction is applied to the hardest emission in an event, using the order  $\alpha_S$  matrix element for that emission. The matrix element programmes automatically include this effect, and take it to more than just the one emission, so should be better at producing hard well separated jets.

### 1.3.3 Hadronisation

The hadronisation stage confines the coloured partons into the colour singlet hadrons seen in nature. Again Herwig and Pythia differ, Herwig uses cluster fragmentation whereas Pythia uses the Lund string model. Sherpa provides an interface to Pythia's string fragmentation routine. An alternative is independent fragmentation, where each parton is hadronised individually, which is available as an option in Pythia, but as I used the preferred string model I will leave any discussion of independent fragmentation. Figure 1.2 shows both cases for a simple colour singlet goes to  $q\bar{q}g$  process.

#### String Fragmentation

The Lund string model is based on a physical picture of what happens during confinement. A 'string' is produced between two outgoing, colour connected partons, representing the energy density of the colour field between the two. A linear potential is assumed for this string. The string then contains an energy given by this potential and taken from the partons. Gluons are included into this string picture as kinks in the string, carrying energy and momentum with a coupling twice that of a quark. This is due to gluons higher colour charge and is accurate in the large  $N_C$  limit, where  $N_C$  is the number of colours (known to be 3). This enters as  $\frac{1}{N_C^2}$  so is an approximation of the same order as  $\alpha_S$ . The string is represented in figure 1.2 by the green blobs joining a  $q\bar{q}$  pair with a

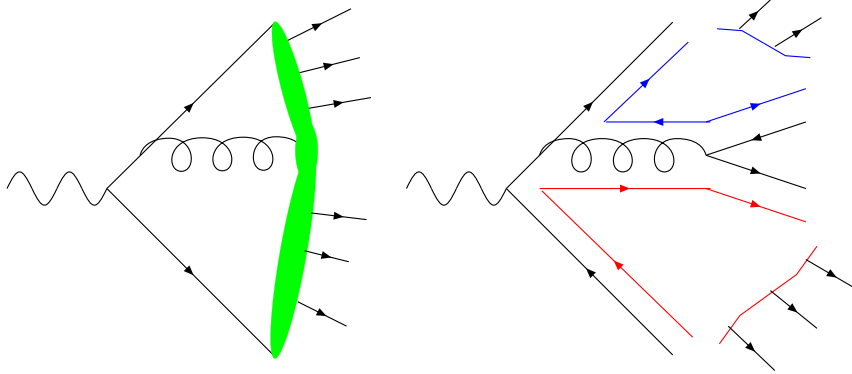


Figure 1.2: Schematic diagram of string and cluster fragmentation schemes

gluon adding a kink. The arrows coming off the string represent its subsequent fragmentation into hadrons.

In order to break the string a quark-antiquark pair are produced which tunnel quantum mechanically from the centre of the string to a classically allowed separation distance. Strings can continue to grow until there is insufficient energy remaining to form any more pairs. As well as these mesons, baryons can be produced by allowing a quark of a certain flavour to be replaced by an anti-diquark of the same flavour. Alternatively there is the popcorn model, in which the produced  $q\bar{q}$  pairs can be produced with any colour, not just that of the nearest particles on the string. This allows colour singlets containing  $rgb$  (baryons) to be formed.

### Cluster Fragmentation

Cluster fragmentation, as implemented in Herwig, works on the principle of pre-confinement. This is related to the QCD colour confinement property discussed previously. The colour structure of the event is fully recorded, with gluons assumed equivalent to a quark-antiquark pair with respect to their colour. This is accurate again in the large  $N_C$  limit.

After the shower has finished, gluons are split into quark-antiquark pairs. Clusters are formed by joining quarks and antiquarks according to the colour flow lines to form colour singlet states. This is represented on the right of figure 1.2 with red and blue zig-zag lines representing the fragmenting clusters.

These states are then decayed to two hadrons (if this isn't disallowed by the mass of the cluster). Lighter mass clusters are allowed to decay to one hadron, and heavier clusters are split by a more string like model, where the cluster undergoes fission iteratively until a sufficiently low mass is reached. The

flavours of the produced pairs and the types of hadrons they produce are chosen at random according to the phase space probabilities. Thus a spectrum of hadrons is produced.

### 1.3.4 Decays, Beam Remnants and Multiple Interactions

A few other points are important in Monte Carlo generators, and worth mentioning here though of less relevance to the work undertaken.

Heavy unstable particles will decay. These decays are treated according to experimentally known branching ratios and if they produce showers (rather than just the decays of final state hadrons) can be seen as secondary hard processes. Some heavy partons, such as the  $t$  quark always decay before hadronisation occurs. Effects in decays such as spin correlations are included in certain processes where they are known but not in all.

Spectators to the main process remain in the beam hadrons, colour connected to the main event and these beam remnants must be considered. Also there is a possibility for more than the two partons of interest in the hadrons to interact. However the possibility of these multiple interactions being hard is sufficiently low not to interfere majorly with the one interaction of interest. A smearing effect on jets can be produced for example.

Pythia contains a full multiple interaction model, whereas Herwig contains only a soft underlying event based on minimum bias events. It is possible to interface Herwig with a multiple interaction programme called Jimmy. I didn't consider the effects of multiple interactions in this study.

### 1.3.5 Jet Finding

When discussing the production of jets in event generators or real experiments it is important to note that all you really see are many stable hadrons and their recombination into jets is not unique. Various jet finding algorithms are used on the hadrons in order to assign them to particular jets. The high level of QCD activity and resulting high jet multiplicities make this particularly important at the LHC.

Jet finding is also important when matching jets to partons from the hard subprocess as described in more detail in section 1.3.6.

Two main methods exist for jet finding, cone algorithms [8] and  $k_{\perp}$  style algorithms [9]. The first is more geometrically based, whereas the second is physically inspired.

## Cone Algorithm

This works by selecting particles above a certain  $E_t$  as seed particles and drawing a cone around them in pseudorapidity-azimuthal angle ( $\eta$ - $\phi$ ) space. Any hadrons inside the cone are grouped with it and a new cone is drawn with its axis formed by the resulting 4-vector of the addition of the cone particles. This continues iteratively until a stable configuration is reached.

In order for this process to remain infra-red safe, after it is complete new cones must be seeded between the already complete cones in order to look for particles which may have been left behind previously. This ensures the procedure is not dependent on any soft radiation between two hard partons. The cone radius,  $R$ , is the most important adjustable parameter and can be optimised to find the best fit of jets to partons depending on the observable.

## $k_{\perp}$ Algorithm

The  $k_{\perp}$  or Durham Algorithm combines particles by defining some closeness parameter, which can be chosen but always reduces to the square of the relative  $k_{\perp}$  between two objects. The objects (stable particles) are grouped according to the smallness of this parameter and combined according to some recombination scheme which can also vary.

The closeness parameter is found between every pair of objects ( $d_{kl}$ ) as well as each object and the beam ( $d_{kB}$ ). The smallest of these  $d$  parameters is found. If  $d_{kl}$  is smallest the 2 objects are combined into a new object. If  $d_{kB}$  is the smallest the object is defined as a jet. This continues until all objects have been placed into jets. It is possible to scale this with the  $R$  parameter which multiplies  $d_{kB}$ , and affects the extent of jets similarly to the radius parameter in the cone algorithm.

I used the above inclusive mode for jet finding in my study. The closeness parameter I used was  $\Delta R$  scheme, which uses  $p_t^2$  and the separation in  $\eta$ - $\phi$  space of the two particles,  $R_{kl}$ , similar to a cone algorithm ( $d_{kB} = p_{Tk}^2$  and  $d_{kl} = \min(p_{Tk}^2, p_{Tl}^2)R_{kl}^2$ ). I chose to recombine the objects in the  $E$  scheme, which simply adds the 4-momenta. I used the programme KtJet [10] for my work.

Another mode of running is called the subjet mode. This takes as an input all the particles in a jet, and combines them until the smallest separation is more than some cut off. Alternatively the separation scale at which  $N + 1$  jets are resolved into  $N$  jets can be found. The similar exclusive mode looks at jets in the whole event rather than the particles within one jet. The exclusive mode also either produces jets dependent on some input separation or outputs the

separation for a required number of jets. This separation scale is known as the  $y_{cut}$ .

### 1.3.6 Matrix Element Matching

Matrix element generators produce amplitudes for a set number of partons which it is assumed correspond to the jets seen in detectors. In order to produce the jets from these partons a parton shower stage is required but it is possible that this will produce extra jets if the radiation is sufficiently hard. For example one hard emission could give you a  $N + 1$  jet sample when you asked for  $N$  jets. It is therefore necessary to introduce a matching procedure to match the jets and partons exactly. Sherpa implements the Catani, Krauss, Kuhn, Webber (CKKW) procedure (explained for  $e^+e^-$  here [11] and extended to hadron collisions here [12]) whereas Alpgen uses the MLM procedure of Michelangelo Mangano [13].

The CKKW procedure produces the tree level matrix element, then clusters the partons according to an exclusive  $k_{\perp}$  algorithm in order to obtain the scales of each parton branching. These splitting scales are used as the inputs for  $\alpha_s$  and Sudakov factors which are used to reweight all the legs. The reweighted system should include the interference effects from the ME part and next to leading log (NLL) effects that are included in parton showers.

In addition to this reweighting stage a parton shower is allowed with emissions above a certain scale vetoed. This factorisation scale separates the hard ME calculation from the PS in order to avoid double counting. The parton splitting scales calculated previously are also used as inputs to start the parton showers at the appropriate scale.

The MLM prescription is a simpler method [13, 14]. A sample is produced containing  $N$  partons, these are parton showered then clustered into jets. In Alpgen jet finding is done using a quick cone algorithm, which is not fully infrared safe. The sample is retained if each original parton is uniquely identified with a jet. For the highest multiplicity sample, the parton shower is allowed to add further jets so all possible multiplicities can be produced. The samples for every  $N$  can then be combined, weighted according to their relative cross sections, in order to obtain a fully inclusive sample.

# Chapter 2

## $WW$ Scattering

### 2.1 Motivation

In the Standard Model without a Higgs, unitarity (probability conservation) breaks down in the electroweak sector at an energy of 1.2 TeV. Solving this requires the appearance of new resonances before this energy is reached, and these resonances will appear in a  $WW$  scattering process, shown in figure 2.1. As these energies will be reached at the LHC it is of interest to understand how the new resonances may appear and how it will be possible to detect them. It should be noted that within the Standard Model it is the Higgs boson, a scalar particle that plays this role.

In order to detect any Higgs or new physics particles through studying  $WW$  scattering it is necessary to have a good understanding of the backgrounds. Currently the only way to do this is through Monte Carlo programmes, but as these are only approximations we cannot be sure of their results. The extrapolation to LHC energies which have never been studied means there is no constraining experimental data available. An understanding of the uncertainties in the simulations can be gained by a comparison of different available programmes. A theoretical understanding of the differences may also help us to know which generators are more believable for different processes.

The largest background to  $WW$  scattering is  $W$ +jets [15]. I therefore attempted to compare and contrast the simulation of  $W$ +jets events at the LHC in both parton shower and matrix element generators. My study followed the previous work of Butterworth, Cox and Forshaw [15] done on  $WW$  events using the kinematic cuts described below. Of particular interest were the newer techniques presented, those of subjet analysis using the  $k_{\perp}$  algorithm and the

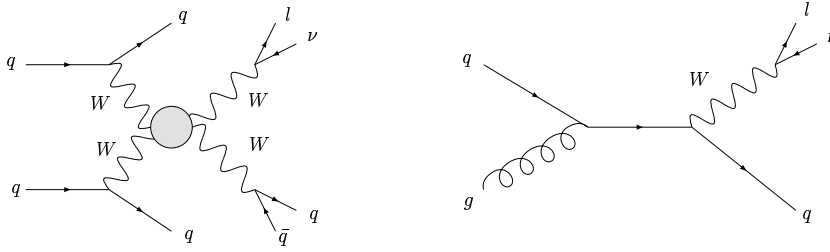


Figure 2.1:  $WW$  Scattering and  $W$ +jet Background

forward and backward tag jets.

## 2.2 Previous Study of $WW$ Scattering

For the study of Butterworth *et al* the semi-leptonic decay mode was chosen. The presence of a lepton is a good cut against the large QCD background expected at the LHC, where there could be more than 20 interactions in every bunch crossing (pile up). Although a fully leptonic decay mode would be therefore easier to detect above jet backgrounds it has a lower branching ratio. It is also impossible to reconstruct the mass of any new particle precisely because the two neutrinos can not be reconstructed from missing momentum.

The major backgrounds to this are  $t\bar{t}$  production with semi-leptonic decay of the resulting  $W$ s and  $W$ +jets with the  $W$  decaying leptonically and the jet misidentified as a  $W$ . In order to reduce these backgrounds a number of cuts are taken on both the leptonic and hadronic  $W$  candidates. A cut can also be taken on the quarks from which the  $W$ s originated, left in the exclusive  $WW$  process but absent in the backgrounds, through the tag jet method described below.

Jet finding is done using KtJet [10], a C++ implementation of the  $k_{\perp}$  algorithm, as in [15]. That is, the inclusive mode is used with an  $R$  parameter of 1 and the  $E$  recombination scheme. This programme is chosen in order to carry out the subjet analysis described below. A study of the effects of using a cone based algorithm as opposed to KtJet has been undertaken [16].

### 2.2.1 Leptonic Cuts

In [15] a number of variables relating to the leptonic  $W$  were studied including the transverse momentum and rapidity of the lepton, but the only cut taken is on the  $W$  reconstructed from the leptons and missing momentum. If more than

one lepton is present the highest  $p_T$  reconstructed  $W$  is chosen. A cut is taken requiring this  $W$  to have transverse momentum greater than 320 GeV.

### 2.2.2 Hadronic Cuts and Subjet Analysis

A transverse momentum cut of 320 GeV is also taken on the hadronic  $W$  candidate, defined to be the highest  $p_T$  jet in the event. In order to ensure that the jet does correspond to a  $W$  a cut on the invariant mass of the jet is taken around the  $W$  mass, it is required to be between 70 and 90 GeV.

The subjet analysis relies on a property of the  $k_\perp$  algorithm implemented in KtJet that allows you to find the scale at which the programme changes from resolving one jet to resolving it as two separate jets. If the jet is truly a  $W$ , this merging scale should be related to the  $W$  mass, this proves a good cut against the  $W$ +jets background. The quantity cut against is the logarithm of the product of the  $p_T$  of the  $W$  candidate and the square root of the  $y_{cut}$  variable discussed in section 1.3.5.

However as I was interested in the underlying causes of any differences in the generators as well as the size of the backgrounds, these cuts were only imposed up to the  $W$  mass cut. This allowed sufficient statistics for a comparison while imposing some  $WW$ -like character on my background.

### 2.2.3 Tag Jet Analysis

The tag jet analysis defines a tag jet as the highest  $E_T$  jet further in pseudo-rapidity forward than the most forward  $W$  candidate or further back than the most backward  $W$  candidate. In the signal process it is expected that these will be the quarks from which the two  $W$ s were emitted, and they will be far apart in rapidity.

In the background processes these tag jets are faked by other jets and prove to be much more central. As the  $WW$  process is colour neutral there will be limited gluon radiation in the central region when compared to the QCD backgrounds so this region will be depleted. In the simulations the fake tag jets can originate either from the parton shower or be produced directly in the matrix element generators.

After the tag jets are identified it is required that there is one in each region (forward and backward) with  $p_T > 20$  GeV and  $E > 300$  GeV.

## 2.3 Background Simulation

The backgrounds in [15] were mainly done using Pythia, with Herwig used as a cross check.  $W$ +jets and  $t\bar{t}$  samples were produced with the minimum  $p_T$  of the hard scatter set to be above 250 GeV for the  $W$ +jets and 300 GeV for  $t\bar{t}$ , in order to have a reasonable chance of passing the cuts. Table 2 in [15] shows that  $W$ +jets is the larger background by an order of magnitude. It is also worth noting that despite predicting a smaller production cross-section, Herwig predicts larger backgrounds after cuts are taken.

## Chapter 3

# $W$ +Jets Simulation Analysis

My simulation of the  $W$ +jets background was undertaken using the standard  $W$ +jets process codes in Herwig and Pythia. As in [15] a minimum  $p_T$  of 250 GeV was imposed on the central system. In addition the  $W$  was forced to decay leptonically. Multiple interactions were turned off in both generators.

Both Alpgen and Sherpa were run with  $W$ +3 jets as the highest multiplicity sample. This should be capable of providing the hadronic  $W$  candidate (hardest jet) and two tag jets. It has also been seen that the parton shower programmes are good at adding one extra jet to the matrix element number so this should be sufficient for my study. The  $W$  was again required to decay leptonically.

The jets in Alpgen were produced with a minimum  $p_T$  of 20 GeV and required to be separated by cone radius  $R$  of at least 0.7. These parameters were also used in the MLM matching.

For Sherpa the  $W$  was required to have a  $p_T$  over 300 GeV. The  $y_{cut}$  parameter used in the CKKW matching scheme was set to 100 GeV. Although a cut of around 20 GeV is standard, it is expected that as the jets recoil against the  $W$  this scale should be sufficient to separate them properly, while also improving the efficiency of the programme massively [20]. As a check a sample was also produced using  $y_{cut}=80$  GeV and  $y_{cut}=20$  GeV. The results were stable under this change, so the more efficient 100 GeV was used to create the graphs seen in this report. The only slight trend was to more central results which I took into account in my conclusions.

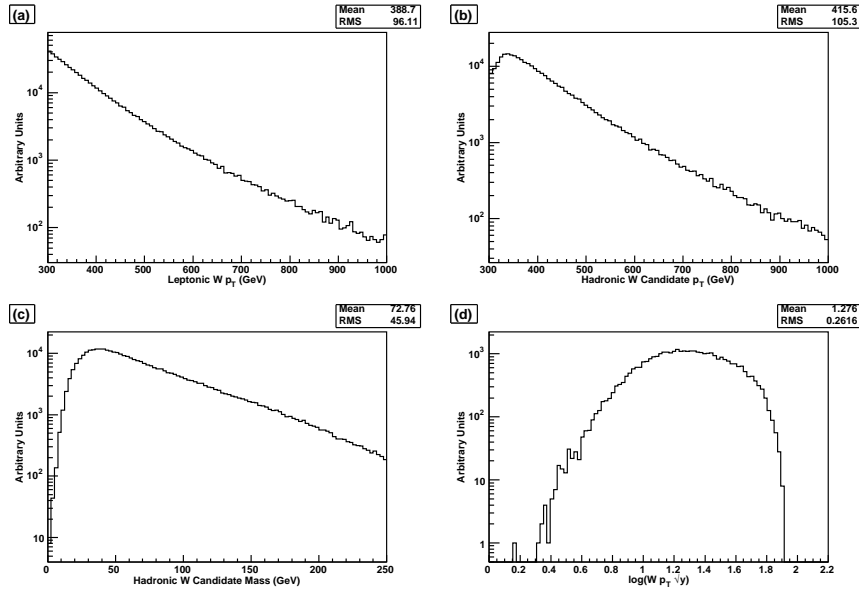


Figure 3.1:  $WW$  Scattering cut variables produced with Herwig

### 3.1 Results

The variables cut against in the study of Butterworth *et al.* (as described in section 2.2) were plotted for all the programmes (figs. 3.1 to 3.5). As can be seen from the plots the distribution shapes are all very similar. The first real differences appear in the tag jet distributions (note these are plotted after the hadronic mass cut, further cuts were taken in [15]).

### 3.2 Colour Coherence in Herwig and Pythia

The rapidity distributions of the tag jets are shown in 3.6. Comparing Herwig (a) and Pythia (b) we see that Herwig predicts a wider spread in rapidity with a noticeable central gap. The distribution in Pythia is much flatter in the central region.

A possible explanation for this difference between Herwig and Pythia is the implementation of colour coherence effects. Gluon radiation occurs from accelerated colour charges, rather like Bremsstrahlung off electric charges. The colour flow lines act like antennae producing the radiation, and interference can form regions in which radiation is suppressed. This colour coherence is responsible for properties such as the angular ordering mentioned in section 1.3.2.

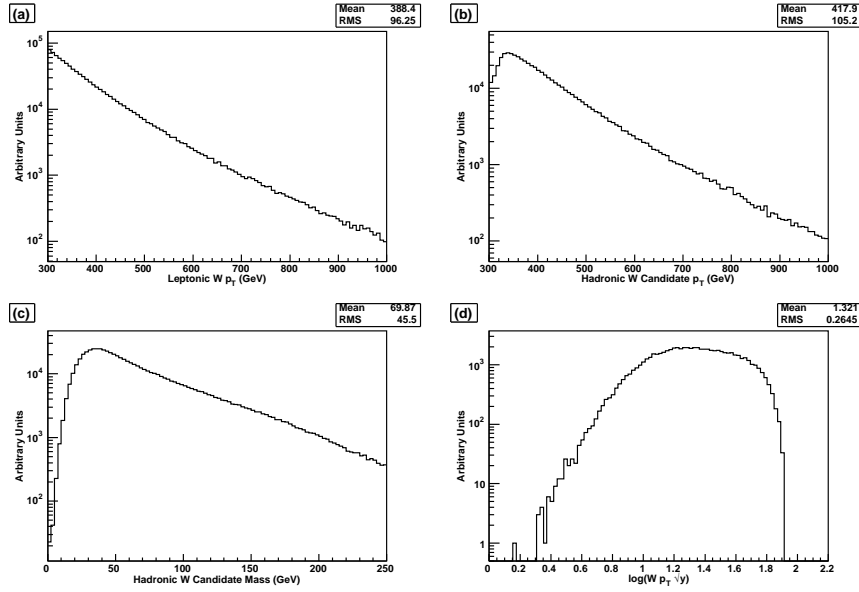


Figure 3.2:  $WW$  Scattering cut variables produced with Pythia

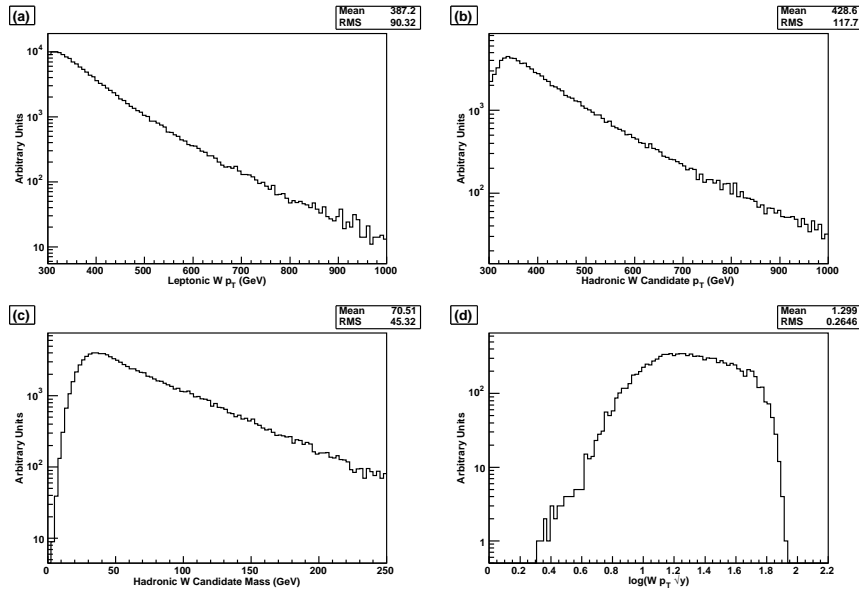


Figure 3.3:  $WW$  Scattering cut variables produced with Sherpa

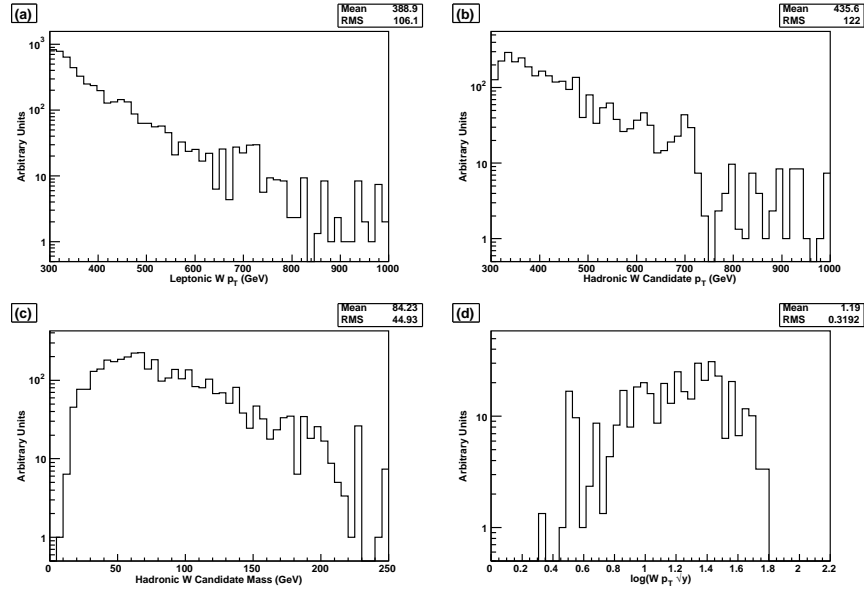


Figure 3.4:  $WW$  Scattering cut variables produced with Alpgen + Herwig

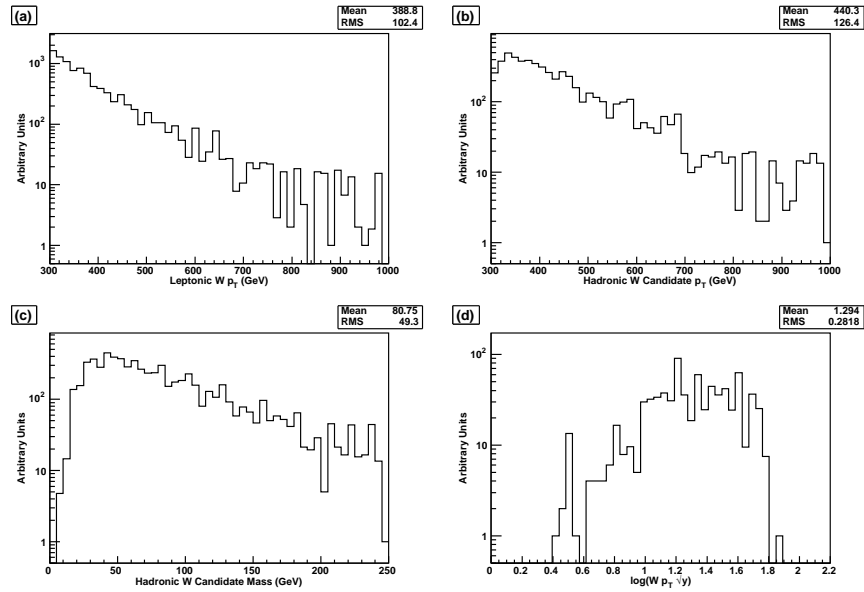


Figure 3.5:  $WW$  Scattering cut variables produced with Alpgen + Pythia

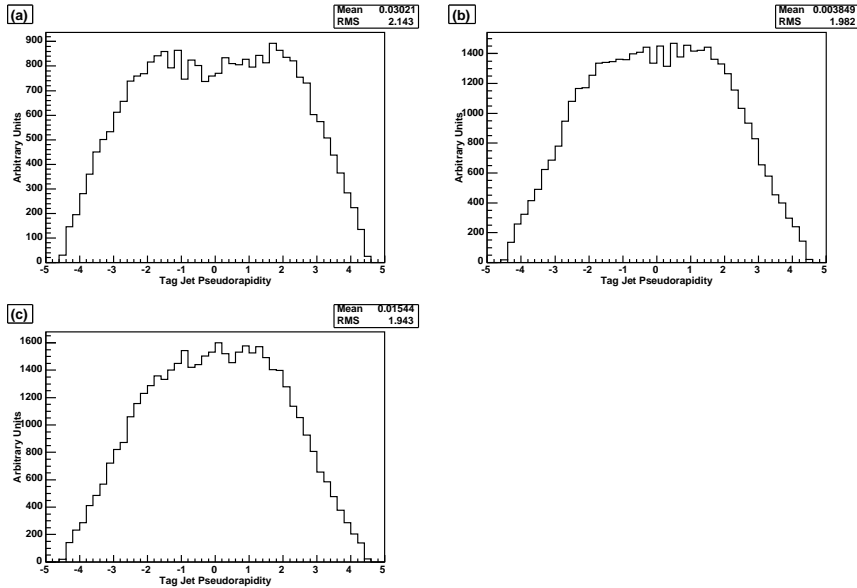


Figure 3.6: Pseudorapidity Distributions for tag jets produced in (a) Herwig, (b) Pythia and (c) Pythia without colour coherence

A previous study by the CDF collaboration [17] on the rapidity distribution of the third hardest jet in multijet events found this to be an important consideration. As a result partial colour coherence was introduced to Pythia, in the form of colour flow between the initial and final state being properly implemented. However Pythia still contains coherence effects to a lesser extent than Herwig’s full implementation. The D0 collaboration later produced similar findings for both multijet events [18] and  $W$ +jets events [19].

It is possible to turn off the colour coherence effects in Pythia ( $\text{MSTJ}(50)=0$  and  $\text{MSTP}(67)=0$ ). I produced a sample of events using these settings and analysed it using an identical method, the results are shown in figure 3.6 (c). It can be seen that though a small effect, this does change in the expected direction, i.e. Pythia without colour coherence is more round centrally than ordinary Pythia. It is therefore possible that Herwig’s full use of colour coherence causes the observed differences in this distribution.

In order both to check that the correct switches in Pythia had been used and to more clearly demonstrate the partial colour effects in Pythia, another colour sensitive variable was plotted. Each tag jet was compared to the highest  $p_T$  jet (the hadronic  $W$  candidate) and the relative distribution in  $\eta$ - $\phi$  space studied. The chosen variable is  $\alpha = \tan^{-1}\left(\text{sgn}(\eta)\frac{\Delta\eta}{\Delta\phi}\right)$ , the polar angle in this space.

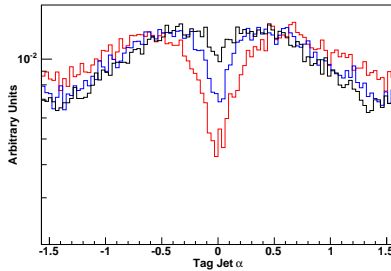


Figure 3.7: Alpha Distributions for tag jets produced in Herwig (red), Pythia (blue) and Pythia without colour coherence (black)

Figure 3.7 shows the unit normalised distributions for Herwig and Pythia, both with and without colour coherence. The difference is clearly visible.

### 3.3 Comparison of Parton Shower and Matrix Element Programmes

As expected, for the variables relating to the  $W$  and hardest jet, the output of matrix element and parton shower programmes was similar (figures 3.1 to 3.5). For one jet both types of programme effectively work in the same way. Although different production cuts were used with the different programmes as previously explained, once analysis cuts are taken the samples are looking at the same regions.

The differences would be expected studying the tag jets, which must be produced in the parton shower stage of Herwig or Pythia, but can be produced directly from the matrix element in Sherpa. The ME programmes should provide a more reliable simulation and help to differentiate between Herwig and Pythia. Unfortunately I had insufficient statistics from Alpgen to provide a full tag jet sample for comparison.

Sherpa shows a central dip in the tag jet pseudorapidity distribution (figure 3.8) as seen in Herwig. This is expected as the coherence effects should be included automatically in the matrix element calculation. However this gap is less than Herwig's and in the tail regions where the cuts are taken Sherpa is in fact much more like Pythia. There was a slight trend towards even more central results as the  $y_{cut}$  parameter was reduced, which would again tend towards a more Pythia like situation.

Alpgen run with the jets restricted to a central rapidity region (see discussion

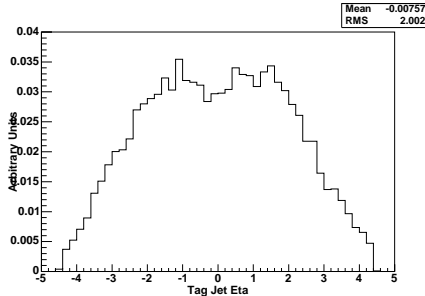


Figure 3.8: Pseudorapidity distribution produced in Sherpa.

below) still exhibited a central dip (figure 3.9) despite this restriction which suggests this feature would still be present in a sample covering the whole of phase space.

### 3.4 Matching Issues in Alpgen

When Alpgen is run with a phase space restriction on the hard partons, this should correspond to the hard jets observed in the event. However the tag jets probe an area of phase space in which potential problems in the matching scheme can start to show up.

Restricting the matrix element jets to a central region  $|\eta| < 2.5$  (Alpgen default), with a veto on parton shower emissions above  $p_T=20$  GeV should stop any tag jets being produced. However I found that the parton shower produced tag jets in the regions outside  $|\eta| = 2.5$ , with obvious dips at  $|\eta| \simeq 2.8$ . This is shown in figure 3.9, where Herwig has been used for the parton shower. Parton showering with Pythia showed similar effects.

This is caused by the differences between the jet finding used in the production and in my analysis. The vetoing of the parton shower jets is done using a cone algorithm, whereas my analysis used KtJet. The different methods in both identifying a jets contents and combining them into a jet, mean that a jet produced with  $p_T$  just under 20 GeV could be pushed above 20 GeV when found using the  $k_\perp$  algorithm. I used the  $E$  scheme in KtJet (4-vector addition) whereas the cone algorithm uses an  $E_T$  wieghted recombination scheme.

Centrally produced soft PS jets can also be absorbed into the hard jets which tend to populate the central region. However in the regions far out in rapidity soft radiation can't be part of a hard central jet, so it is possible to appear as a tag jet.

Comparing this to Sherpa run with the same cut on pseudorapidity (figure

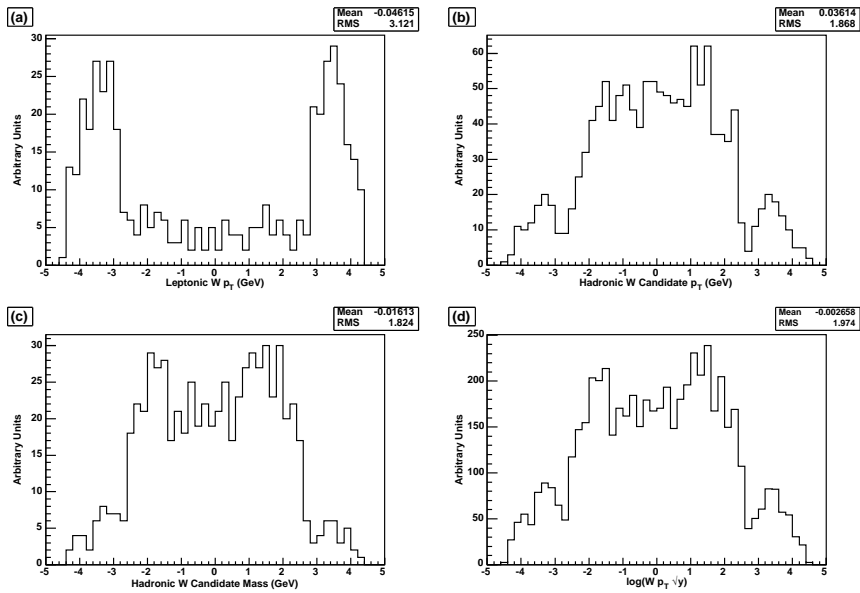


Figure 3.9: Rapidity distribution of tag jets produced in Alpgen with  $|\eta|$  of the ME jets  $< 2.5$  for (a) 1 jet, (b) 2 jets, (c) 3+ jets, (d) fully inclusive sample.

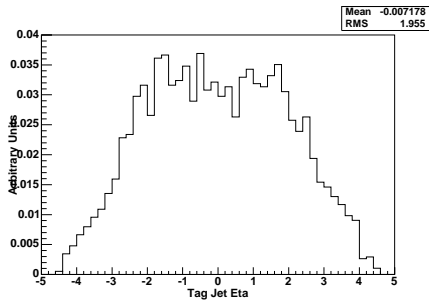


Figure 3.10: Rapidity distribution of tag jets produced in Sherpa with  $|\eta|$  of the ME jets  $< 2.5$ .

3.10) we do not see the same effects. The  $k_{\perp}$  algorithm is used for both the matching and the analysis, although it is run in different modes. It seems that the CKKW prescription produces a smoother covering of the phase space. The matching between the ME and PS produced jets hasn't introduced any unphysical effects as in Alpgen.

The unwanted events in Alpgen are rare compared to the number of events produced and few have both forward and backward jets, both with  $E > 300$  GeV. However it still indicates that care must be taken to cover all the phase space required properly when using the matrix element - parton shower matching method. A different choice of matching parameters, especially the Alpgen  $p_T$  veto, would probably produce better results.

# Chapter 4

## Conclusions

From my study of simulating  $W$ +jets events I conclude that there are slight differences in the distributions predicted using different Monte Carlo generators, but these differences are only small and not a cause for major concern. Although it is expected that ME generators will be more reliable than the traditional PS programmes, for the case of  $W$ +jets backgrounds to the  $WW$  scattering study of Butterworth *et al* the results of both programme types act to confirm each other.

The most interesting differences occurred in the far forward and backward regions of rapidity, studying the tag jets. Herwig's full implementation of colour coherence causes it to predict wider tag jet rapidity distributions than Pythia. However when the tag jets are produced by the leading order matrix element calculations in Sherpa they correspond more closely with the predictions of Pythia.

I also looked at the matching between the matrix element programmes and parton showers in the small corner of phase space probed by the tag jets. I conclude that Sherpa's CKKW method matches more smoothly than Alpgen's MLM method in this particular case. However I believe this could be improved by choosing the matching parameters better to suit the variables of interest.

### 4.1 Future Plans

#### 4.1.1 $W$ +Jets Improvements

Two main effects in simulating events have been left out of my study. In order to extend it multiple interactions and detector simulation could be introduced.

Multiple interactions, or the underlying event, are a result of beam remnant

interactions (see section 1.3.4). These would be expected to have the effect of filling any rapidity gap with soft radiation making it more difficult to detect. I would expect this to reduce the efficiency of the tag jet cuts and smear out many of the differences between the different generators. A study of the effects of different multiple interaction scenarios in [15] concluded the signal to background ratio would be reduced by at least a factor of two.

In order for the Monte Carlo programmes to be compared to real data they must be fed through some simulation of the physical detector in order to produce the same signals as will really be seen. The ATLAS experiment has both a full detector simulation and an alternative fast simulation ATLFAST [21]. Again the inclusion of detector effects would be expected to smear any results when compared to our 'perfect detector'. This should again mean that the differences between generators are less important, but of course the effects need to be properly tested. Sherpa is not currently implemented within the ATLAS framework.

It is worth noting that the study of the  $WW$  signal process in reference [15] didn't include detector simulation, although Pythia using ATLFAST has been studied in [16]. Full simulation of the  $WW$  scenarios considered in [15] and [16] and backgrounds is currently being studied in Manchester.

An obvious improvement would be an increase in the statistics available, especially from the ME programmes. This would just be a matter of time and would allow a better comparison of the distribution shapes as well as a meaningful quantitative look at background sizes. A more thorough tuning Sherpa's  $y_{cut}$  parameter so as to find the best pay off between the stability of the results and efficiency could be done. It would also be interesting to see if Alpgen's efficiency could be increased in a similar way. Tuning the matching parameters in both Alpgen and Sherpa to investigate further the matching problems I found could also be done.

#### 4.1.2 Gaps Between Jets Processes in QCD

Processes in which jets are produced with a rapidity gap between them are of practical and theoretical interest. As mentioned in section 2.2.3 colour singlet events are expected to have less central radiation than QCD events. In order to understand processes such as  $WW$  scattering it is necessary to have an understanding of the probabilities of events occurring with no radiation between the jets.

In reality it is impossible to measure an absence of radiation completely due to detector uncertainties and it is also impossible to calculate perturbatively the

cross sections for these processes, so in practice an upper limit,  $Q_0$  on radiation into the gap is imposed. This leads to large logarithms,  $\ln(\frac{Q}{Q_0})$ , which must be resummed to all orders to understand the gap cross sections.

To see this consider the simple example of  $e^+e^- \rightarrow q\bar{q}$ . If one gluon is emitted from the quarks, then without any requirement of a gap the real and virtual gluon emissions will exactly cancel (Bloch-Nordsieck Theorem) as they are divergent. Note also that the virtual gluon contributes an imaginary 'coulomb gluon' term which cancels with the complex conjugate. When we impose a gap in rapidity, of size  $Y$ , no real gluons are allowed within the gap, with energy above  $Q_0$ . Therefore virtual gluons in the gap with energy above  $Q_0$  have nothing to cancel against. In order to calculate the gap cross section you need consider only these virtual gluons. This leads to

$$\sigma_{gap} = \sigma_0 \left( 1 - C_F \frac{\alpha_s}{\pi} \int_{Q_0}^Q \frac{dk_{\perp}}{k_{\perp}} \int_{-Y/2}^{Y/2} dy \right) \quad (4.1)$$

$$= \sigma_0 \left( 1 - C_F \frac{\alpha_s}{\pi} \ln\left(\frac{Q}{Q_0}\right) Y \right) \quad (4.2)$$

where  $Y$  is the size of the gap and  $Q$  is some energy describing the hard 2→2 process. We assume the emitted gluon is soft compared to this. In order to include the possibility of more virtual gluons in the gap we resum this equation to all orders, which simply exponentiates this factor giving

$$\sigma_{gap} = \sigma_0 \exp\left(-C_F \frac{\alpha_s}{\pi} \ln\left(\frac{Q}{Q_0}\right) Y\right) \quad (4.3)$$

. Extending this to QCD processes such as  $qq \rightarrow qq$  introduces a matrix that acts on the vectors of different colour states.

However, Dasgupta and Salam [22] discovered that it is not as simple as that. A gluon (real or virtual) outside the gap can emit another virtual gluon which can be inside the gap, again with nothing to cancel against. This leads to a new class of logarithms appearing known as non-global logarithms, as their appearance is due to the non-global nature of the observable (i.e. the phase space restriction of requiring a gap). Further investigation of this by Forshaw, Kyrieleis and Seymour [23] in  $qq \rightarrow qq$  processes, recently revealed that some of these new logarithms, appearing at order  $\alpha_s^4$ , are in fact super-leading. This means the logs have a higher power than the previously thought leading logarithms and are formally more important. The origin of these terms is the coulomb gluons mentioned above.

There is therefore much interesting work to be done on this area and it will form the basis of much of my future work. In particular my aim is to implement

an order-by-order equivalent of this calculation into a computer programme in order to check these new results.

# Bibliography

- [1] M. E. Peskin, D. V. Schroeder, *An Introduction to Quantum Field Theory*, (Westview Press, Oxford, 1995).
- [2] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and Collider Physics*, (Cambridge University Press, Cambridge, 1996).
- [3] ATLAS Collaboration, ATLAS Detector and Physics Performance Technical Design Report (1999) [CERN/LHCC/99-14]
- [4] HERWIG 6.5, G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour and B. R. Webber, *JHEP* **0101** (2001) 010 [arXiv:hep-ph/0011363, hep-ph/0210213].
- [5] PYTHIA 6.3, T. Sjöstrand, P. Edén, C. Friberg, L. Lönnblad, G. Miu, S. Mrenna and E. Norrbin, *Comput. Phys. Commun.* **135** (2001) 238.
- [6] SHERPA 1.0, A. Schalicke and F. Krauss, *JHEP* **0507** (2005) 018 [arXiv:hep-ph/050328].
- [7] ALPGEN 2.0, M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. Polosa, *JHEP* **0307** (2003) 001; [arXiv:hep-ph/0206293].
- [8] G. Sterman and S. Weinberg, *Phys. Rev. Lett.* **39** (1977) 1436; F. Abe *et al.* [CDF Collaboration], *Phys. Rev. D* **45** (1992) 1448; S. D. Ellis, Z. Kunszt and D. E. Soper, *Phys. Rev. Lett.* **69** (1992) 3615 [arXiv:hep-ph/9208249].
- [9] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock and B. R. Webber, *Phys. Lett. B* **269** (1991) 432; S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, *Nucl. Phys. B* **406** (1993) 187.
- [10] J.M. Butterworth, J. P. Couchman, B. E. Cox and B. M. Waugh, *Comput. Phys. Commun* **153** (2003) 85.
- [11] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, *JHEP* **0111** (2001) 063 [arXiv:hep-ph/0109231].

- [12] F. Krauss, JHEP **08** (2002) 015, [arXiv:hep-ph/0205283].
- [13] M. L. Mangano, *Exploring theoretical systematics in the ME-to-shower MC merging for multijet process*, E-Proceedings of Matrix Element/Monte Carlo Tuning Working Group, Fermilab (2002).
- [14] S. Mrenna and P. Richardson, JHEP **0405** (2004) 040 [arXiv:hep-ph/0312274].
- [15] J. M. Butterworth, B. E. Cox and J. R. Forshaw, Phys. Rev. D **65** (2002) 096014 [arXiv:hep-ph/0201098].
- [16] S. E. Allwood, PhD Thesis (to be submitted) (2006); S. E. Allwood, J. M. Butterworth, B. E. Cox and F. K. Loebinger, Unpublished (2005) [[https://www.hep.ucl.ac.uk/twiki/bin/view/Main/AtlasWW/atlas\\_ww.ps](https://www.hep.ucl.ac.uk/twiki/bin/view/Main/AtlasWW/atlas_ww.ps)].
- [17] F. Abe *et al.* [CDF Collaboration], Phys. Rev. D **50** (1994) 5562.
- [18] B. Abbott *et al.* [D0 Collaboration], Phys. Lett. B **414** (1997) 419 [arXiv:hep-ex/9706012].
- [19] B. Abbott *et al.* [D0 Collaboration], Phys. Lett. B **464** (1999) 145 [arXiv:hep-ex/9908017].
- [20] J. Winter, private communication.
- [21] ATLFast 2.0, E. Richter-Was, D. Froidevaux and L. Poggioli, ATLAS Internal Note, (1998) [ATL-PHYS-98-131].
- [22] M. Dasgupta and G. P. Salam, JHEP **0203** (2002) 017 [arXiv:hep-ph/0203009].
- [23] J. R. Forshaw, A. Kyrieleis, M. H. Seymour [arXiv:hep-ph/0604094].