Modelling Uncertainties in $t\bar{t}$ Production at the LHC

Jonathon Waters, University of Manchester, United Kingdom

September 3, 2013

Supervisors: Alexander Grohsjean and Judith Katzy, DESY, Hamburg

Abstract

Understanding the physics of $t\bar{t}$ production and decay is reliant on Monte Carlo simulations which is subject to various systematic uncertainties. This report studies the impact of not yet considered effects arising from the modelling of additional jet radiation. The importance of these variations is evaluated using the measurement of the jet gap fraction from the ATLAS collaboration.

Contents

1 Introduction

Discovered in 1995 by the collaborations at the CDF and D0 experiments, the top quark is the most massive elementary particle known to date. With the arrival of the LHC in 2008, a practical top quark 'factory' entered the scene. The large centre-of-mass energy of 7 resp. 8 TeV, yields a much higher cross-section for the production of top quark pairs, $t\bar{t}$, and hence much larger statistics. However, the LHC also comes with additional challenges. Two major factors are important. First the increased probability that a collision will produce several events, the so-called pile-up, second the additional radiation that may be emitted. Both of these make it difficult to resolve the event that is being attempted to be observed.
A dedicated measurement of the additional radiation in $t\bar{t}$ events was performed by the ATLAS collaboration [1] and is used in this report to quantify so far unconsidered systematic effects in the modelling of top pair events. The additional radiation can be described by either matrix element (ME) generator or by parton showering (PS). ME generators better describes hard radiation and PS better describes soft radiation so knowing which of these modelling methods best describes the observed events may yield a better understanding of the nature of these emissions. This report studies how both can best matched to describe the observable gap fraction and jet multiplicity in $t\bar{t}$ events.

2 Monte Carlo Simulation

In order to perform this analysis, Monte Carlo events were generated in two distinct steps. First the ME part of an event is generated using the leading order (LO) ME generator ALPGEN v2.1.4p5 with the CTEQ6L1 parton distribution function (PDF) [2]. Then, the event is showered using PYTHIA v6.427 with tune Perugia 2012 (MSTP370) [3].

One problem that can arise with this procedure is that an event with n additional partons generated with the ME generator can be mimicked by an event with $n-1$ ME partons and 1 additional PS parton. To prevent this sort of double-counting the so called MLM matching procedure is used [4].

There are several user defined inputs in ALPGEN. For this analysis the default configuration of the ATLAS collaboration was used [5]. The few exceptions are discussed in the following.
η_{Jmax}, the maximum value of the pseudorapidity of the additional jets, is set to 4.5. Originally this was set to 6, however because there is a cut on $|y|$ of 2.1[1], this means that far more events had to be generated in order to have statistically valid results. The two other options that are being changed are *iqopt*, which can take values 0, 1 (default) or 2, and p_{Tmin} which we are giving values of 10, 15 (default), 25, 35 GeV. p_{Tmin} defines the minimum transverse momentum of an additional light parton generated by ALP-

[1]Because the partons that are being studied are massless the value of $|y|$ and $|\eta|$ are here considered to be the same

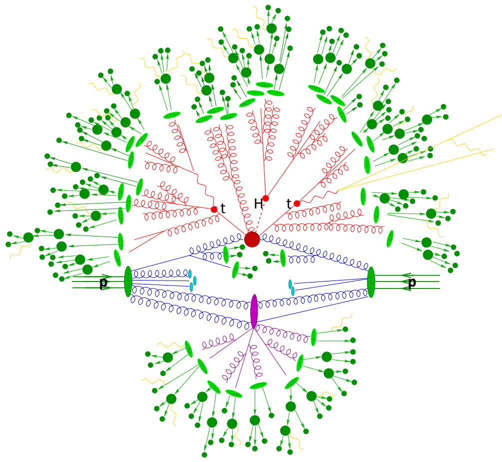

Figure 1: A representation of the production and decay of a $t\bar{t}$ pair at the LHC [6].

GEN. *iqopt* changes the functional form of the factorisation and renormalisation scale, Q^2. Selecting option 0 sets this factor to 1. Selecting option 1 sets this to the transverse mass of the the system which is defined as $\Sigma m_T^2 = \Sigma(p_T^2 + m^2)$ summed over the $t\bar{t}$ system and the additional ME generated light partons. Selecting the final option 2, sets this factor to $\hat{s}$ where $\hat{s} = x_1^2 \cdot x_2^2 \cdot s$. x_i is the momentum of the fraction of the initial parton i and s is the centre-of-mass energy. There is also an option $qfac$ which is a scaling factor for Q. This is set to 1 except for *iqopt* $= 0$ where it is set to the mass of the top quark[2].

ALPGEN along with PYTHIA employs MLM matching and is a step by step procedure which is run on a per event basis [4] to avoid double-counting as described above. The event generation happened as following:

- ALPGEN generates a $t\bar{t}$ event with up to four additional jets.
- Using PYTHIA and the renormalisation and factorisation scale Q, the partons are showered to produce their final jets.
- Every parton with a transverse momentum greater than p_{Tmin} is gathered into a cone jet with cone $R_{clus} = 0.7$ and a minimal transverse momentum of $E_{Tclus} = p_{Tmin} + 5$ GeV. Any event is only accepted if each of the cone jets can be matched to only one ME generated parton, by using $\Delta R < 1.5 * R_{clus}$ between the jets and the ME partons for light flavour partons.
- If generating at the highest multiplicity, PYTHIA will also accept an event with additional jets assuming that they have a lower transverse momentum then minimal transverse momentum of the matched jets. This allows the user to change how hard or soft QCD radiation is described by changing p_{Tmin}.

Specifically in the case of this analysis, the process was to first generate the ALPGEN results using various inputs at separate jet multiplicities of 0 through to 4, each multiplicity split into a number of jobs with a set number of events. The higher the multiplicity the larger the number of jobs required as it is more likely for an event to be rejected at high multiplicity. These results are then consolidated into single output files separated by their multiplicity and run through PYTHIA along with a RIVET analysis plugin [7]. The two analyses were ATLAS_2012_I1094568 and MC_MyAnalysis, which was written specifically for this project. Both analyses fill and output histograms of the desired observables, which are then merged using ROOT over all of the multiplicities. The final histograms are then overlaid for comparison.

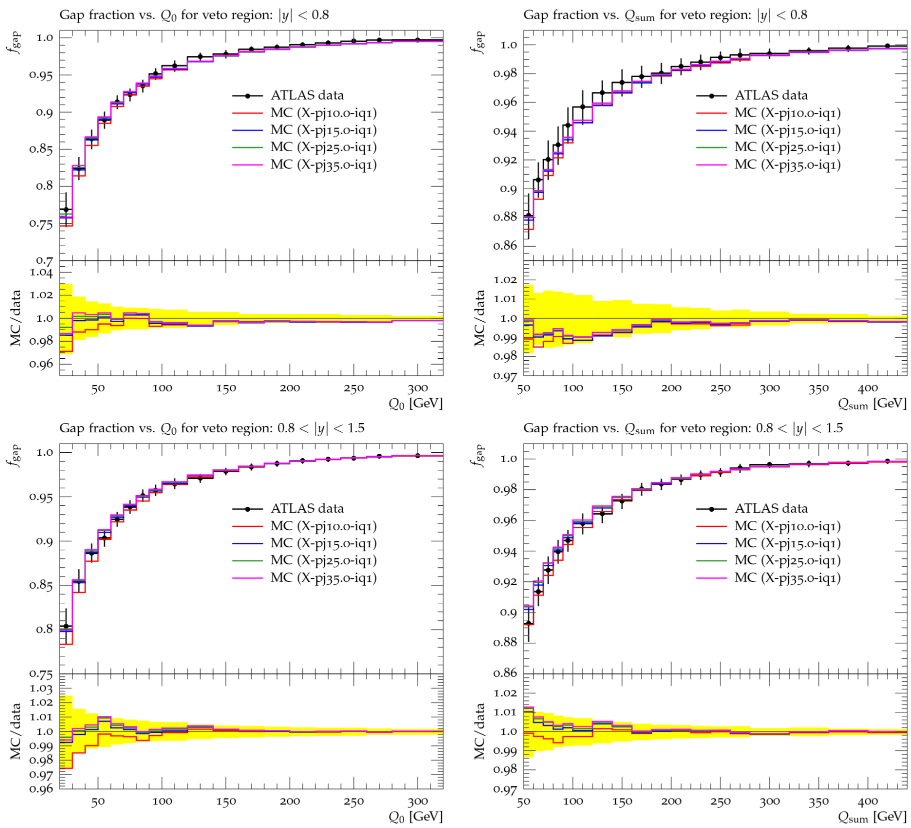

Figure 2: The gap fraction as a function of Q_0 and Q_{sum} is compared for 4 different values of p_{Tmin} as well as the measured gap fraction from the ATLAS detector for the first two rapidity regions. The real data is displayed as black circles with the statistical uncertainty and the Monte Carlo generated data is shown by coloured lines; red for $p_{Tmin} = 10$ GeV, blue for 15 GeV, green for 25 GeV, purple for 35 GeV. The yellow band in the ratio plot indicates the statistical uncertainty of the measured data.

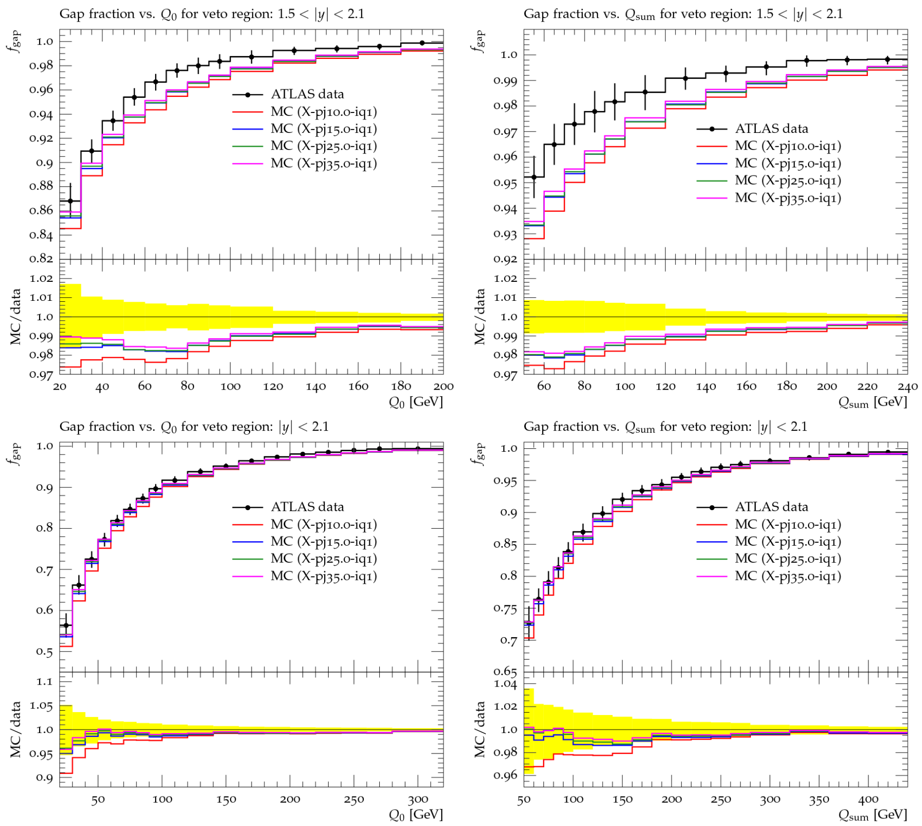

Figure 3: The gap fraction as a function of Q_0 and Q_{sum} is compared for 4 different values of p_{Tmin} as well as the measured gap fraction from the ATLAS detector for the final rapidity region as well as the total region. The results are displayed in the same way as in Figure 2.

3 Observables

3.1 Gap Fraction

To quantify the studied variations, the RIVET analysis routine ATLAS_2012_I1094568 was used [7].The gap fraction of a process is defined as $f(Q_0) = n(Q_0)/N$, where $n(Q_0)$ is the subset of these events, within the selected rapidity region, that do not contain an additional jet with transverse momentum above a given value Q_0, and N is the total number of selected events. This observable is useful as there is a strong dependence on the number of additional jets that are produced by the PS process. In order to observe dependence of this variable on more than just to overall data, the gap fraction is presented in 4 rapidity intervals; $|y| <0.8$, $0.8 <|y| <1.5$, $1.5 <|y| <2.1$ and $|y| <2.1$ [1].

In order to probe deeper than just the leading additional jet, the gap fraction is also measured as a function of Q_{sum} $f(Q_{sum}) = n(Q_{sum})/N$, where Q_{sum} is the sum of the transverse momenta of all jets in the considered region. This variable is sensitive to

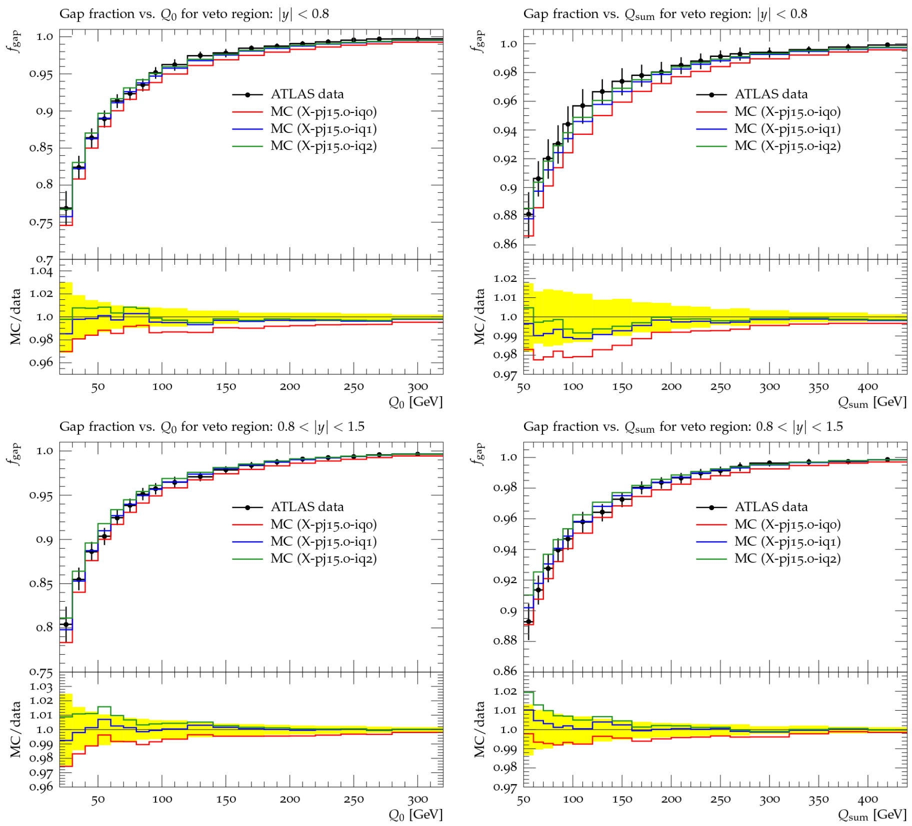

Figure 4: The gap fraction as a function of Q_0 and Q_{sum} is compared for 3 different values of *iqopt* as well as the measured gap fraction from the ATLAS detector for the first two rapidity regions. The real data is displayed as black circles with the statistical uncertainty and the Monte Carlo generated data is shown by coloured lines; red for $iqopt = 0$, blue for 1, green for 2. The yellow band in the ratio plot indicates the statistical uncertainty of the measured data.

all hard jet emissions whereas the gap fraction of Q_0 is mostly sensitive to leading jet emissions [1].

The measurement of the jet gap fraction performed at the ATLAS experiment, is based on the dilepton $t\bar{t}$ decay channel, where both W bosons, from $t \rightarrow Wb$, decay into a lepton and the corresponding neutrino. Thus 2 b quark jets and 2 leptons are required, where only electron or muon leptons are considered. This measurement uses the dilepton channel because any other channel would result in additional decay jets, which would be hard to distinguish from the additional jets that are being studied. For full details on the cuts used in this analysis see the ATLAS collaboration paper on the jet veto analysis [1].

The results are compared in Figure 2 and 3 separately for Q_0 and Q_{sum}. In general, when we vary the value of p_{Tmin}, we find that all settings except for $p_{Tmin} = 10$ GeV behave quite similar within the given uncertainty. Except for the very forward region,

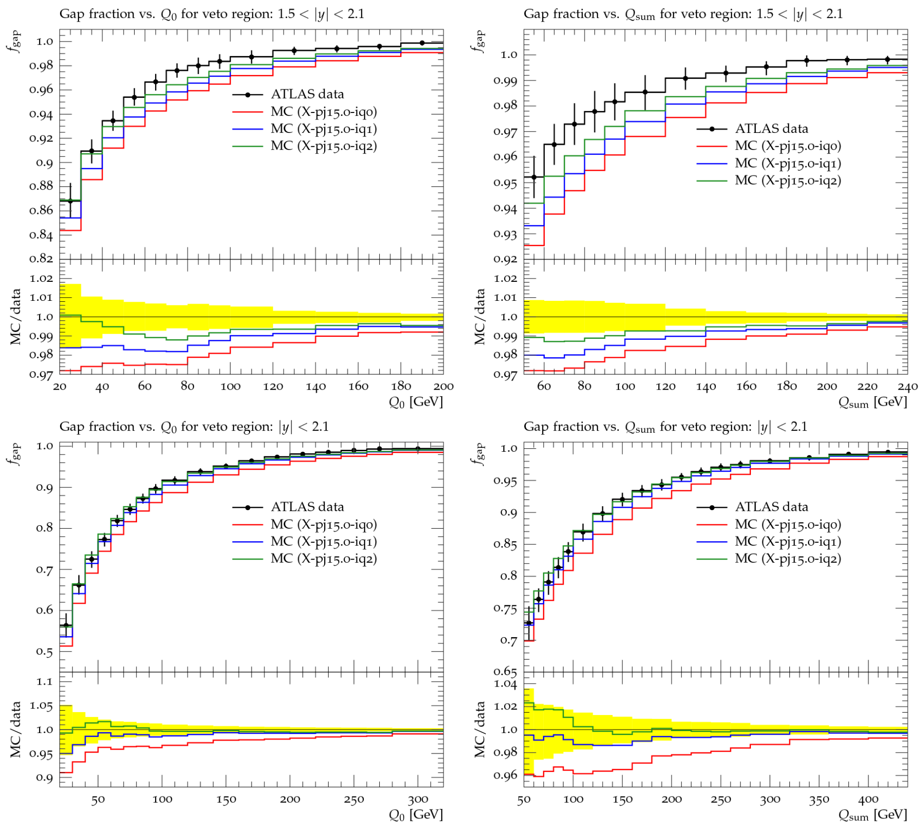

Figure 5: The gap fraction as a function of Q_0 and Q_{sum} is compared for 4 different values of *iqopt* as well as the measured gap fraction from the ATLAS detector for the final rapidity region as well as the total region. The results are displayed in the same way as in Figure 4.

where all settings are below the data, i.e. have a higher jet activity, a good description of the data is obtained. When the minimal transverse parton momentum is lowered to 10 GeV, the jet gap fraction is in general too low. While still ok within the uncertainty for the central region, a large discrepancy is observed for the very forward region. This behaviour can be understood as the setting allows a larger phase space for the ME generation. In summary however, the systematic effect from the variation of the matching scale can be considered small.

Figures 4 and 5, show similar histograms as those in 2 and 3, but this time comparing the results of using different functional forms for the renormalisation and factorisation scale by varying *iqopt*. In general, the jet activity decreases, when decreasing α_s and increasing Q^2. Thus the softest jet spectrum is obtained for $\hat{s}$. While in the central region all settings are okay with the given uncertainties, discrepancies are observed in the very forward region and the inclusive results for the lowest scale. The best performance is obtained when using $\hat{s}$. This might suggest a change of the default ATLAS scale choice for ALPGEN top pair production.

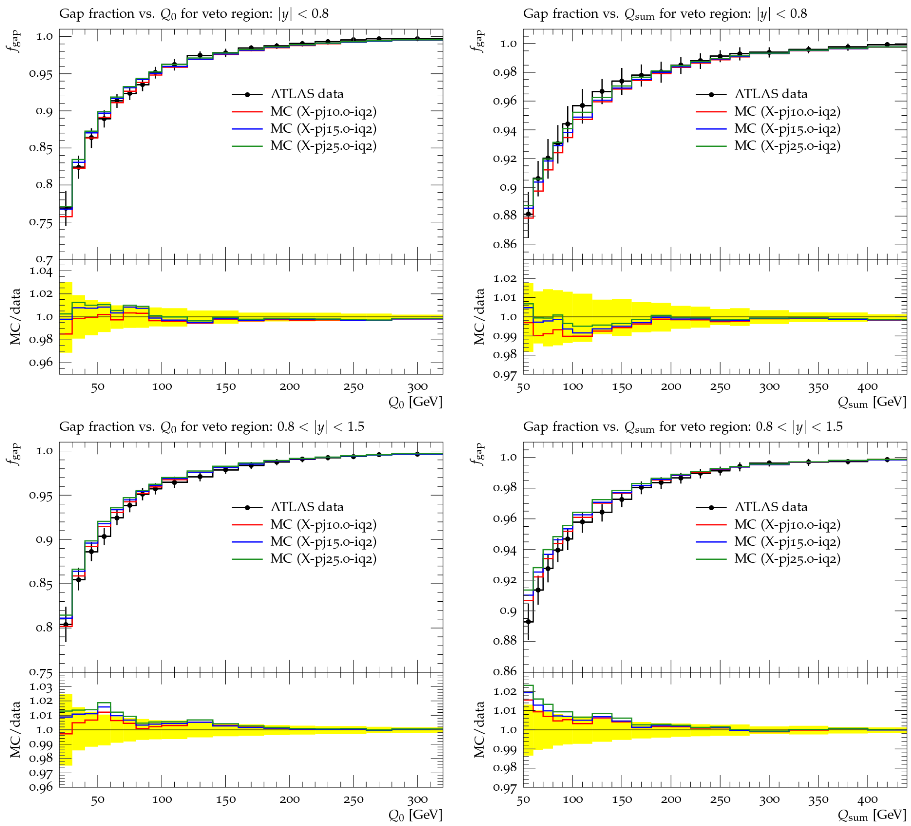

Figure 6: The gap fraction as a function of Q_0 and Q_{sum} is compared for 3 different values of p_{Tmin}, using $iqopt = 2$, as well as the measured gap fraction from the ATLAS detector for the first two rapidity regions. The real data is displayed as black circles with the statistical uncertainty and the Monte Carlo generated data is shown by coloured lines; red for $p_{Tmin} = 10$ GeV, blue for 15 GeV, green for 25 GeV. The yellow band in the ratio plot indicates the statistical uncertainty of the measured data.

Using the $iqopt = 2$ option the relationship for p_{Tmin} was again checked. This time, only two other settings for p_{Tmin} were chosen along with the default of 15 GeV: 10 GeV and 25 GeV. The results of these variations are shown in figures 6 and 7. As with the default option of *iqopt* there is very little change when increasing the value of p_{Tmin}, but a much larger change when decreasing it. All previous findings can be found consistently here.

3.2 Multivariable Analysis

In addition to the gap fraction, several other observables may have a dependence on the varied options. In order to study them, a separate RIVET plugin based on the ATLAS_2012_I1094568 analysis was written and named MC_MyAnalysis. It uses the

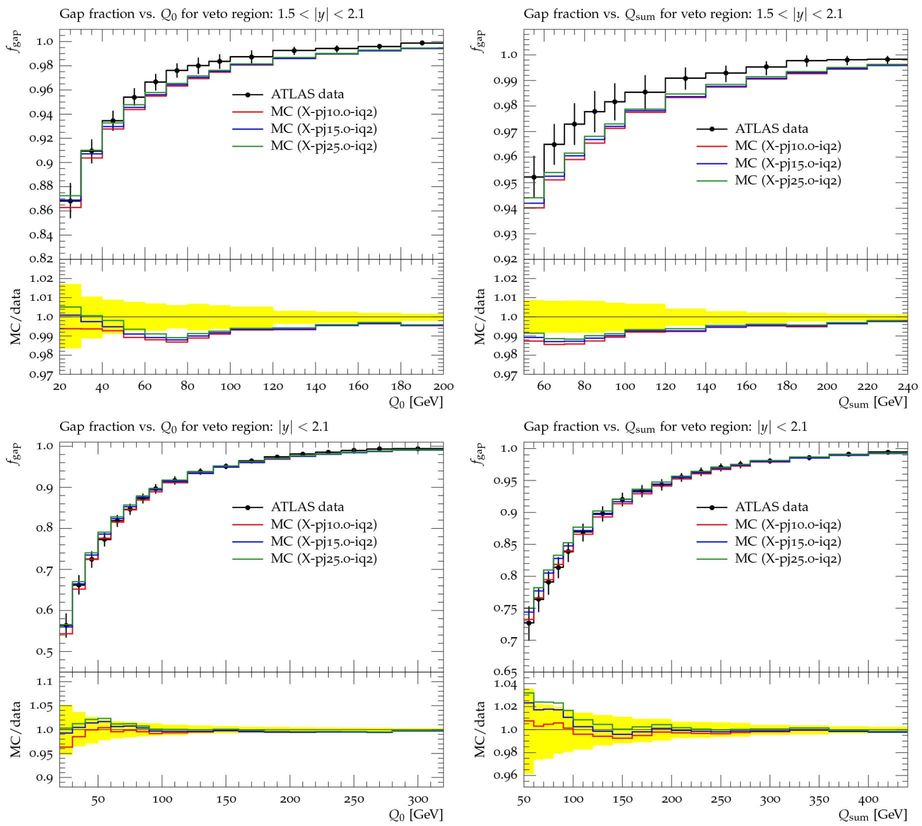

Figure 7: The gap fraction as a function of Q_0 and Q_{sum} is compared for 3 different values of p_{Tmin}, using $iqopt = 2$, as well as the measured gap fraction from the ATLAS detector for the final rapidity region as well as the total region. The results are displayed in the same way as in Figure 6.

same cuts on the events as the original analysis but includes the b quark jets in the final results.

The variables that are investigated are; the total number of jets that are produced, the transverse momentum of the lead jet, the transverse momentum of the sub leading jet and the the value of ΔR between these two jets. These variables allow us to see the effect that differing these options has on the physical processes that are being modelled and hence how the hardness of the radiation that is emitted may vary these physical observables.

The results of this additional analysis are shown in Figures 8 and 9. The distributions of ΔR don't show any significant pattern of deviation from the default. This is expected as there should be no dependence on the phase space positioning of the jets based on the nature of the additional radiation produced. The distribution of the number of jets per event, also shown in Figure 8, is also as expected. The figure shows that there are fewer jets produced overall by $iqopt = 2$ than the default, i.e a higher frequency of events with 2 or 3 jets and a lower frequency of events with 4 or more jets. This is consistent

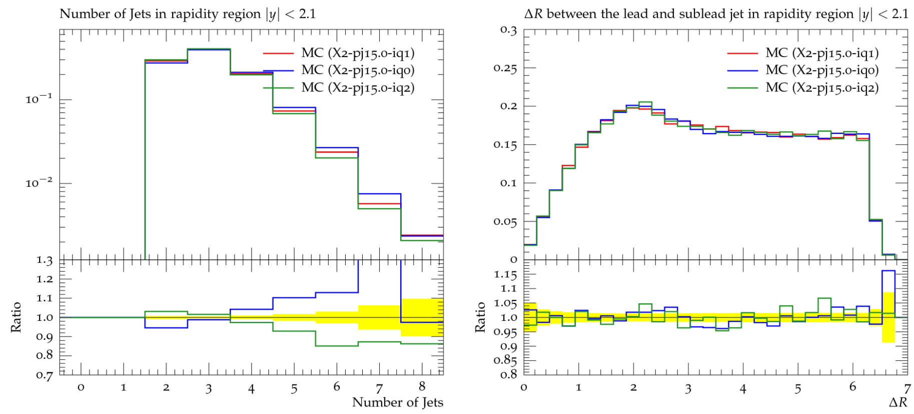

Figure 8: Histograms of a) the number of jets per event, and b) the value of ΔR between the lead and sublead jets of the event. Each of these were taken over the total rapidity region. In the 8 jets bin is also included all events with greater than 8 jets. The red line for each shows the result for using $iqopt = 1$, blue shows 0 and green shows 2. The ratio plot show the ratio between the results and the default setting of 1. The yellow band is the statistical uncertainty of this result.

with the higher gap fraction that is observed in Figures 4 and 5 for this option and indicates that the additional radiation emitted is softer. The results of $iqopt = 0$ have the opposite distribution, with more overall jets and indications of harder radiation. This is also consistent with the results of the previous analysis.

Figure 9 shows the distribution of the transverse momentum of both the leading and subleading jets of each event. The lead jet transverse momentum shows a higher frequency of events with low p_T for $iqopt = 2$ than the default and a corresponding lower frequency of high p_T jets. This also is consistent with the results for the gap fractions of these options. Setting $iqopt$ equal to 0 once again shows the opposite trend. These results are also present in the subleading jet p_T, however they are far less clear at high p_T values.

4 Conclusions

The jet gap fractions as a function of Q_0 and Q_{sum} in four rapidity intervals is used in this analysis to study systematic uncertainties in top pair production. The ATLAS measurement of the jet gap fraction is compared to several Monte Carlo based simulations with different settings of ALPGEN and PYTHIA.

It was found that there is only a small systematic uncertainty arising form the variation of p_{Tmin} on these results, mostly resulting from a value lower than the default. Changing the functional form of the normalisation and factorisation scale did show a large systematic uncertainty present due to the different settings. Improvements in

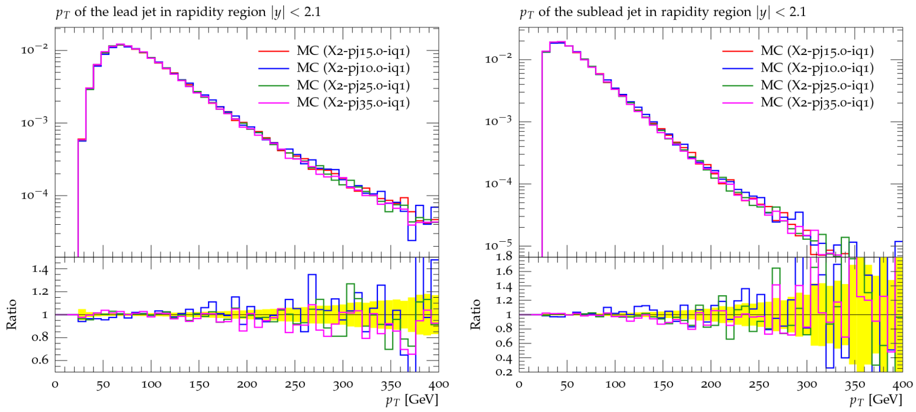

Figure 9: Histograms of transverse momentum of the lead jet and the sublead jet in the total rapidity region. The results are plotted in the same way as in Figure 8.

modelling top pair production can be made by using $\hat{s}$ as scale rather than the current default.

References

[1] ATLAS Collarboration, Eur.Phys.J. C72 (2012) 2043, arXiv:1203.5015v2.

[2] M.L. Mangano *et al.*, JHEP 0307 (2003) 001, arXiv:hep-ph/0206293.

[3] T. Sjöstrand, S. Mrenna and P. Skands, JHEP 05 (2006) 026, arXiv:hep-ph/0603175.

[4] D. Geerts, Nikhef, (18/04/2013), https://indico.cern.ch/getFile.py/access?contribId=2&resId=0&materialId=slides&confId=247811.

[5] ATLAS Collaboration, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/McProductionCommonParametersMC12#ALPGEN.

[6] SHERPA, https://sherpa.hepforge.org/trac/wiki.

[7] A. Buckley *et al.*, arXiv:1003.0694v8.