ATLAS NOTE

ATLAS-CONF-2011-055

April 6, 2011

Measurement of Inclusive Two-Particle Angular Correlations in Proton-Proton Collisions at $\sqrt{s}$ = 900 GeV and 7 TeV

The ATLAS Collaboration

Abstract

Measurements of two-particle angular correlations in proton-proton collisions at centre-of-mass energies of 900 GeV and 7 TeV are presented. Correlations are measured for charged particles in the kinematic region $p_{\mathrm{T}} > 100$ MeV and $|\eta| < 2.5$. Collision events were recorded using a minimum bias trigger with the ATLAS detector at the LHC during 2009 and 2010. A complex correlation structure in $\Delta\eta$ and $\Delta\phi$ is observed at both energies. Results are compared to Pythia 8 and Phojet as well as the ATLAS MC09, DW and Perugia$_0$ tunes of Pythia 6.

1 Introduction

Several theoretical models have been proposed to describe the dynamics of multi-particle production in high-energy hadron collisions (e.g. string [1] and cluster [2] models) but so far they have all proven to be incomplete, since they provide a limited explanation of the emission of soft radiation. The study of correlations between final state particles is a powerful method for investigating the underlying mechanisms of particle production, as demonstrated by previous experiments [3, 4, 5]. Additionally, particle correlations can help to identify important dynamical information that can be incorporated in models or theories to gain a better and more global description of multi-particle production within soft interactions.

There are a large number of ways to explore correlations in multi-hadron final state environments including, but not limited to: angular correlations, forward-backward multiplicity correlations and charge correlations. This study focuses on the two-particle angular correlations in pseudorapidity, η, and azimuthal angle, ϕ, and explores its different components using data collected by the ATLAS Inner Detector [6].

The foreground correlation distribution F in η and ϕ between emitted particle pairs is given by

$$F(\Delta\eta, \Delta\phi) = \left\langle \frac{2}{N_{ch}(N_{ch}-1)} \sum_i \sum_{j\neq i} \delta_{\eta_i-\eta_j-\Delta\eta}\delta_{\phi_i-\phi_j-\Delta\phi} \right\rangle, \tag{1}$$

where the summation is over all charged particles in a single event, N_{ch} is the number of charged particles in that same single event and the average is taken over an ensemble of many events. The Kronecker delta functions, $\delta_{\eta_i-\eta_j-\Delta\eta}$ and $\delta_{\phi_i-\phi_j-\Delta\phi}$, select particle pairs with the correct $\Delta\eta$ and $\Delta\phi$ separation for $F(\Delta\eta, \Delta\phi)$. The factor of $2/N_{ch}(N_{ch}-1)$, the inverse of the total number of particle pairs in an event, accounts for the fact that a single particle in a high multiplicity event can be paired with more particles than an identical single particle in a low multiplicity event. Without this factor, a particle in a low multiplicity event would carry a lower weight than a particle in a higher multiplicity event because there are fewer possible pair combinations in the low multiplicity event.

There are two contributions to this two-particle density in $(\Delta\eta, \Delta\phi)$. The first way in which correlations can arise is if there are correlations between emissions in a single event; if an event already contains an emission at a given (η, ϕ), is it more or less likely that there will be another nearby emission?

However, since $\mathrm{d}N_{ch}/\mathrm{d}\eta$ is approximately flat in η, but the Inner Detector is limited to $|\eta| < 2.5$, phase-space alone dictates that the η separation between two particles, $\Delta\eta$, will favour smaller values. $\Delta\eta$ is peaked at $\Delta\eta = 0$ and falls approximately linearly to a maximum possible value of $\Delta\eta = 5$. This second contribution defines the background correlation distribution B, the $\Delta\eta$ dependence of which is given by

$$B(\Delta\eta) = \int_{-2.5}^{2.5}\int_{-2.5}^{2.5} \mathrm{d}\eta_1\mathrm{d}\eta_2\,\delta(\eta_1-\eta_2-\Delta\eta)\left.\frac{\mathrm{d}N_{ch}}{\mathrm{d}\eta}\right|_{\eta=\eta_1}\left.\frac{\mathrm{d}N_{ch}}{\mathrm{d}\eta}\right|_{\eta=\eta_2}, \tag{2}$$

where $\mathrm{d}N_{ch}/\mathrm{d}\eta$ is the average charged particle distribution in η. The $\Delta\phi$ dependence of B is defined in a similar way and is flatter than the $\Delta\eta$ dependence.

Labelling the foreground and background at a given charged particle multiplicity N_{ch} as $F(N_{ch}, \Delta\eta, \Delta\phi)$ and $B(N_{ch}, \Delta\eta, \Delta\phi)$ the correlation function $C(N_{ch}, \Delta\eta, \Delta\phi)$ can be constructed

$$C(N_{ch}, \Delta\eta, \Delta\phi) = F(N_{ch}, \Delta\eta, \Delta\phi) - B(N_{ch}, \Delta\eta, \Delta\phi). \tag{3}$$

Past analyses [3] have found that $(N_{ch}-1)\times C(N_{ch}, \Delta\eta, \Delta\phi)$ is approximately independent of the multiplicity. Averaging $(N_{ch}-1)\,C(N_{ch}, \Delta\eta, \Delta\phi)$ over all multiplicities and dividing out the background contribution gives the two-particle correlation function $R(\Delta\eta, \Delta\phi)$

$$R\left(\Delta\eta,\Delta\phi\right) = \frac{\langle\left(N_{ch}-1\right)F\left(N_{ch},\Delta\eta,\Delta\phi\right)\rangle_{ch}}{\langle B\left(N_{ch},\Delta\eta,\Delta\phi\right)\rangle_{ch}} - \frac{\langle\left(N_{ch}-1\right)B\left(N_{ch},\Delta\eta,\Delta\phi\right)\rangle_{ch}}{\langle B\left(N_{ch},\Delta\eta,\Delta\phi\right)\rangle_{ch}}, \tag{4}$$

where $\langle ... \rangle_{ch}$ indicates an average over contributions from all particle multiplicities. Since the $\mathrm{d}N_{ch}/\mathrm{d}\eta$ distribution, and hence the background, is not strongly dependent on the charged particle multiplicity, the factor of B approximately cancels from the second term on the right of Equation 4. The background *does* have some multiplicity dependence at low multiplicity due to the contribution from double diffraction, which favours higher $\Delta\eta$ values compared to the non-diffractive contribution. However, the effect of any track reconstruction inefficiency is larger in low multiplicity events where a small number of mis-reconstructed tracks has a proportionally larger impact on the event. Calculating a multiplicity independent background distribution using $\langle \mathrm{d}N_{ch}/\mathrm{d}\eta \rangle_{ch}$, which is already averaged over all multiplicities, therefore has the advantage of diluting the effect of the experimentally more troublesome lower multiplicity events. Such an approach reduces the sensitivity of the observable to diffraction, which is in any case not the motivation for this measurement. The final expression used for the inclusive two-particle correlation function is then

$$R\left(\Delta\eta,\Delta\phi\right) = \frac{\langle\left(N_{ch}-1\right)F\left(N_{ch},\Delta\eta,\Delta\phi\right)\rangle_{ch}}{B\left(\Delta\eta,\Delta\phi\right)} - \langle N_{ch}-1\rangle_{ch}\,. \tag{5}$$

In practice $\langle\left(N_{ch}-1\right)F\left(N_{ch},\Delta\eta,\Delta\phi\right)\rangle_{ch}$ is constructed by taking each pair of particles within a single event, calculating their η and ϕ separations and filling a two-dimensional histogram at those values using a weight of $2/N_{ch}$. This histogram is normalised by dividing each bin by the number of events entering the distribution. The background distribution is determined by taking pairs of un-related events and, for each particle in one event, the $\Delta\eta$ and $\Delta\phi$ values with each particle in the other event are calculated and used to fill another two-dimensional histogram. The background two-dimensional histogram is finalised by normalising it to unit integral.

The one-dimensional projections of the correlation function, R($\Delta\eta$) and R($\Delta\phi$), are calculated by first integrating separately the foreground and the background distributions and then taking the ratio between the two and normalising with the average track multiplicity.

This note presents the measurement of the inclusive two-particle angular correlation function for charged particles inside the kinematic region $p_{\mathrm{T}} > 100$ MeV and $|\eta| < 2.5$ in proton-proton collisions with centre-of-mass energies of 900 GeV and 7 TeV. A comparison of the data to different Monte Carlo generators and tunes is also performed. The note is organised as follows: the data samples and event selection are described in Section 2. Section 3 gives a brief description of the different Monte Carlo models used in this study. In Section 4 the data-driven procedure for unfolding the effect of the detection apparatus from the physics is presented. The effectiveness of the corrections method is subsequently demonstrated by a closure test that is based on Monte Carlo. The sources of systematic uncertainty affecting the analysis are listed in Section 5. Finally, the corrected two-particle angular correlation functions and a comparison to the Monte Carlo tunes are shown in Section 6.

2 Data Samples and Event Selection

This analysis follows the same selection criteria as the charged-particle multiplicity analysis at 900 GeV and 7 TeV [7]. Events were required:

- to have been triggered by the single-arm, single-counter level 1 Minimum Bias Trigger Scintillators (MBTS1),
- to contain at least one primary vertex,

- not to have a second vertex associated to more than four tracks (to remove events with more than one proton-proton interaction per bunch crossing),
- to contain at least two tracks with:
  - a transverse momentum $p_{\mathrm{T}} > 100$ MeV,
  - a pseudorapidity $|\eta| < 2.5$,
  - at least 1 hit in the B-Layer (innermost layer) of the Pixel detector, if expected,
  - at least 1 hit in any of the layers of the Pixel detector,
  - at least 2 ($p_{\mathrm{T}} > 100$ MeV), 4 ($p_{\mathrm{T}} > 200$ MeV) or 6 ($p_{\mathrm{T}} > 300$ MeV) hits in the SemiConductor Tracker (SCT),
  - a transverse impact parameter with respect to the Primary Vertex $|d_0^{PV}| < 1.5$ mm,
  - a longitudinal impact parameter with respect to the Primary Vertex $|z_0 \sin\theta^{PV}| < 1.5$ mm (where θ is the angle between the track and the z-axis),
  - a track-fit probability $\chi^2/ndof \geq 0.01$ if the track has a $p_{\mathrm{T}} > 10$ GeV.

Events in which the ATLAS Inner Detector was fully operational and the solenoid magnetic field was operating at the nominal field strength of 2 Tesla were used in this analysis. For the measurement at $\sqrt{s} = 900$ GeV, 357,523 events containing 4,532,663 tracks were selected and in the case of the $\sqrt{s} = 7$ TeV data sample, 10,066,072 events and 209,809,430 tracks passed the selection cuts. The corresponding integrated luminosities are approximately 7 $\mu\mathrm{b}^{-1}$ at 900 GeV and 190 $\mu\mathrm{b}^{-1}$ at 7 TeV.

3 Monte Carlo Models

In order to compare the present measurement to different soft QCD models, three different generators were considered: Pythia 6 [8], Pythia 8 [9] and Phojet [10].

The Pythia approach to soft interactions uses phenomenological adaptations of QCD in order to extend its physics description to the non-perturbative regime. Phojet is based on the Dual Parton Model (DPM), which describes soft-hadronic processes via Pomeron exchange and semi-hard processes by parton scattering.

Three tunes of Pythia 6.4.21 were compared to the 7 TeV data: DW [11], Perugia$_0$ [12] and MC09 [13]. The default tune of Pythia 8.130 using the CTEQ 5L parton density functions (PDF) and the Phojet 1.12.1.35 version (ATLAS default setup) were also used for comparison. For the 900 GeV data, the results are only compared to MC09.

MC09 is a tune of Pythia 6 using the MRST LO* PDF and was produced by the ATLAS collaboration in order to describe as well as possible a range of minimum bias and underlying event data from the Tevatron. Perugia$_0$ is also tuned to a wide range of Tevatron data, however it uses the CTEQ 5L PDF. Both MC09 and Perugia$_0$ use the newer p_{T} ordered shower and multi-parton interactions (MPI) model in which the MPI is interleaved with the initial state radiation. Tune DW is an older tune to CDF underlying event and Drell-Yan data that uses the older virtuality-ordered shower and non-interleaved MPI model.

Pythia 8 adds to the MPI model of Pythia 6 by also interleaving the final state radiation. Pythia 8 also includes an updated model for diffraction that allows harder colour singlet exchange, however at the time of production this had not been extensively tuned to data. Phojet contains separate hard and soft diffractive contributions and it has not been tuned to any recent experimental data.

All Monte Carlo samples used in the analysis are a mix of non-diffractive and diffractive components according to the generator cross-sections shown in Table 1. To exclude secondary particles such as the decay products of neutral kaons from the generator-level distributions, particles with a mean decay length ($c\tau$) greater than 10 mm were considered to be stable.

	Non-Diffractive	Single-Diffractive	Double-Diffractive	Central-Diffractive
Pythia (900GeV)	34.4 mb	11.8 mb	6.4 mb	—
Pythia (7TeV)	48.5 mb	13.7 mb	9.2 mb	—
Phojet (7TeV)	61.6 mb	10.7 mb	3.9 mb	1.3 mb

Table 1: Monte Carlo sub-process cross-sections at $\sqrt{s}$ = 900 GeV and 7 TeV. The values shown for Pythia were used for both Pythia 6 and Pythia 8.

4 Correction Procedure

For simplicity, the pseudorapidity correlation distribution R($\Delta\eta$) will serve as an example to illustrate the correction method, however, the extension to the two-dimensional distribution is straightforward. The uncorrected two-particle correlation distributions can be found in Appendix A.

4.1 Vertexing and Trigger Efficiencies

The trigger and vertex reconstruction efficiencies were calculated before a requirement on the vertex was made. Since these efficiencies depend on the track multiplicity, the total number of selected tracks was obtained by applying the selection criteria described in Section 2 with the exception of the d_0^{PV} and $z_0 \sin\theta^{PV}$ requirements. In their place a cut on the transverse impact parameter with respect to the beamspot position, $\left|d_0^{BS}\right|$, was made such that $\left|d_0^{BS}\right| < 1.8$ mm. Both the foreground and the multiplicity distributions are weighted with the inverse of these efficiencies. All of the efficiency distributions were taken directly from the charged-particle multiplicity analysis in [7].

No specific correction is made for tracks arising from the secondary decays of longer lived particles. The effect of secondary particles is discussed in Section 5.3.

4.2 Tracking Inefficiency Correction by Probabilistic Track Removal

The effect of mis-reconstructed tracks is corrected for using a data-driven method that requires knowledge only of the single track reconstruction efficiency as a function of track p_{T} and η (derived from Monte Carlo in [7]). The track-based observable (in this case the correlation function) is first computed using all reconstructed tracks satisfying the analysis cuts. The tracking efficiency, $E_i\left(p_{\mathrm{T}}, \eta\right)$, is determined for each track, t_i, and compared to a random number, r_i, generated uniformly between zero and one. Note that each track must have a unique random number. A subset of the original tracks is defined by taking all those tracks for which

$$E_i\left(p_{\mathrm{T}}, \eta\right) > r_i. \tag{6}$$

This is the same method that a Monte Carlo uses in order to generate random numbers according to a desired distribution. In this case the method is being used to *remove* tracks from the sample according to their tracking efficiency, i.e. tracks are lost in exactly the same way they are lost by the ATLAS detector. Having removed some tracks, the observable is computed again using only the remaining tracks. This defines one iteration of the track-removal procedure (the full set of uncorrected tracks is the zeroth iteration). A second iteration takes as input the tracks used in the first iteration and again uses a *new* set of random numbers to remove some tracks according to their tracking efficiency. Further iterations are carried out in the same way.

With this correction method it is possible to quantify the effect that detector inefficiencies have on an observable; each iteration corresponds to an additional application of the detector effect on the data. The -1 iteration corresponds to the observable if no detector effects are present. The value of (each bin of)

the observable can be plotted as a function of the known iteration number (0, 1, 2... N) and a function can be fitted to the resulting distribution. By extrapolating the fit to the -1 iteration an estimate of the true observable without detector effects can be made.

The fit and subsequent extrapolation use a third degree polynomial, with six track removal iterations in total used as the input. As with all correction procedures, this method requires high statistics in order to work properly. The number of track removal iterations is limited by statistics – there must be a sufficient number of events with sufficient tracks that some remain after all iterations. Six iterations is enough to make a fit, but is not so large that the last iteration has no tracks.

A third degree polynomial was chosen as it is the simplest type of function that can fit all of the bins of the observable. Both the foreground and the background distributions are track-counting variables and therefore decay approximately exponentially under each successive track removal iteration. The observable, O, should therefore behave something like

$$O(x) \simeq \frac{A + Be^{-x}}{C + De^{-x}}, \tag{7}$$

where x is the number of track removal iterations and A, B, C and D depend upon the precise details of the observable and the effect of the detector. Equation 7 asymptotically approaches A/C for large positive x and approaches B/D for large negative x. Therefore Equation 7 can in general describe an "S" like shape, the sign of the second derivative depending on x. A second degree polynomial is unable to fit such a function, whereas a third degree polynomial, which has two turning points, is able to produce a reasonable fit.

The effect of the track removal procedure on the data is shown in Figure 1. The black solid line in the two distributions corresponds to R($\Delta\eta$) in the 0th iteration, using all the selected tracks. The different markers correspond to that same distribution after each of the iterations. In Figures 2 and 3, examples of third degree polynomial fits done to central and outer bins are shown for both centre-of-mass energies.

4.3 Testing the Correction Procedure in Monte Carlo

In order to test the correction method, Monte Carlo studies were performed in which the distributions at the generated (or truth) level were recovered from a fully detector-simulated and reconstructed Monte Carlo sample. The following plots correspond to the MC09 Pythia tune only. After applying the event and track requirements listed in Section 2 a total of 790,877 events containing 8,746,908 tracks were selected for 900 GeV and 16,544,352 events and 291,723,405 tracks for these Monte Carlo tests at 7 TeV.

Figure 4 illustrates the third degree polynomial fits to two bins of the reconstructed Monte Carlo R($\Delta\eta$) distribution. Figure 5 shows the corrected Monte Carlo distributions at 900 GeV and 7 TeV compared to those at the truth or generated level. The two distributions are in good agreement with each other. The remaining differences will enter as a systematic uncertainty, as discussed in Section 5.3.

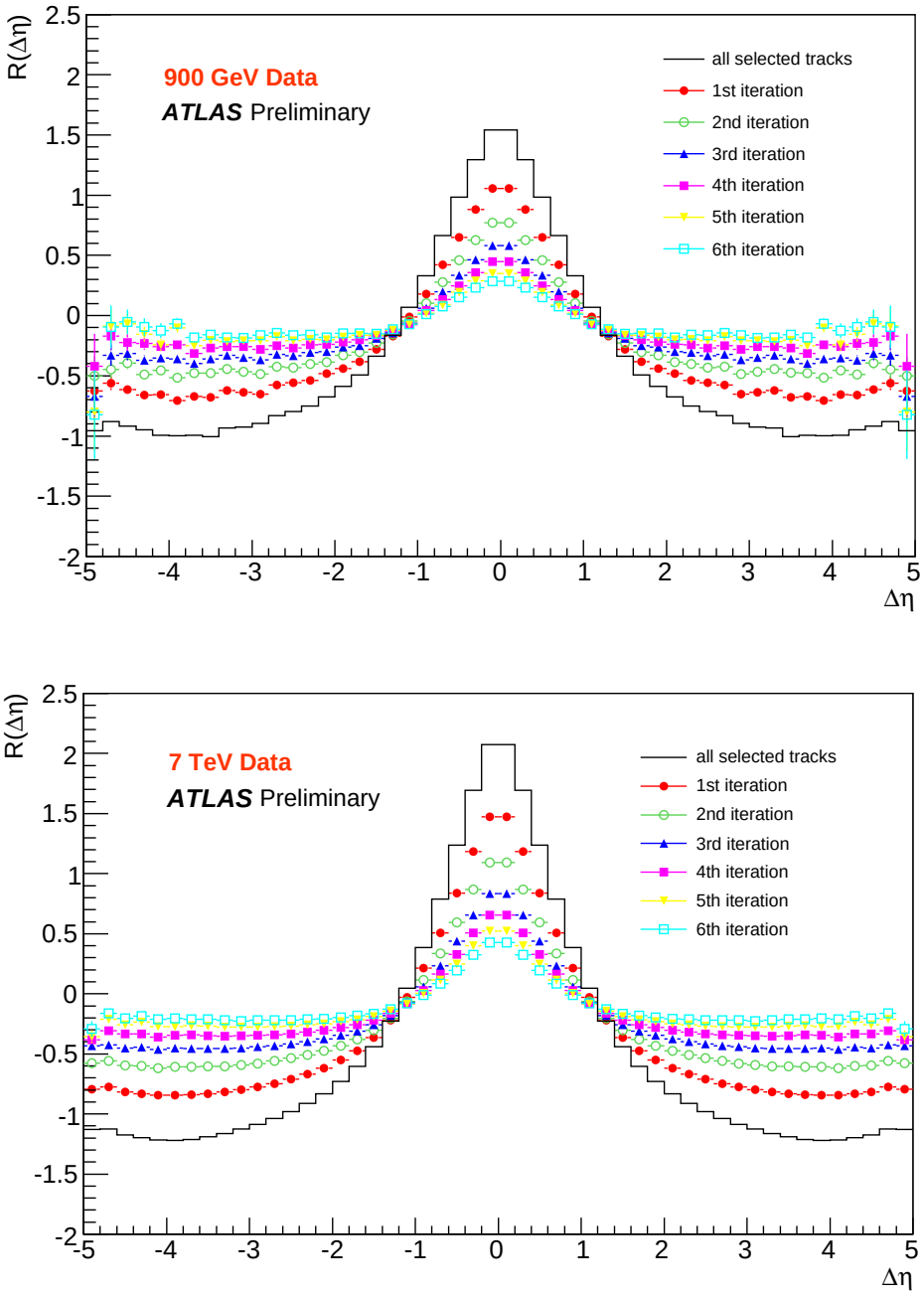

Figure 1: Two-particle pseudorapidity correlation function after each iteration in the probabilistic track removal correction method for 900 GeV (top) and 7 TeV (bottom). As the number of tracks gets reduced, the distributions flatten out.

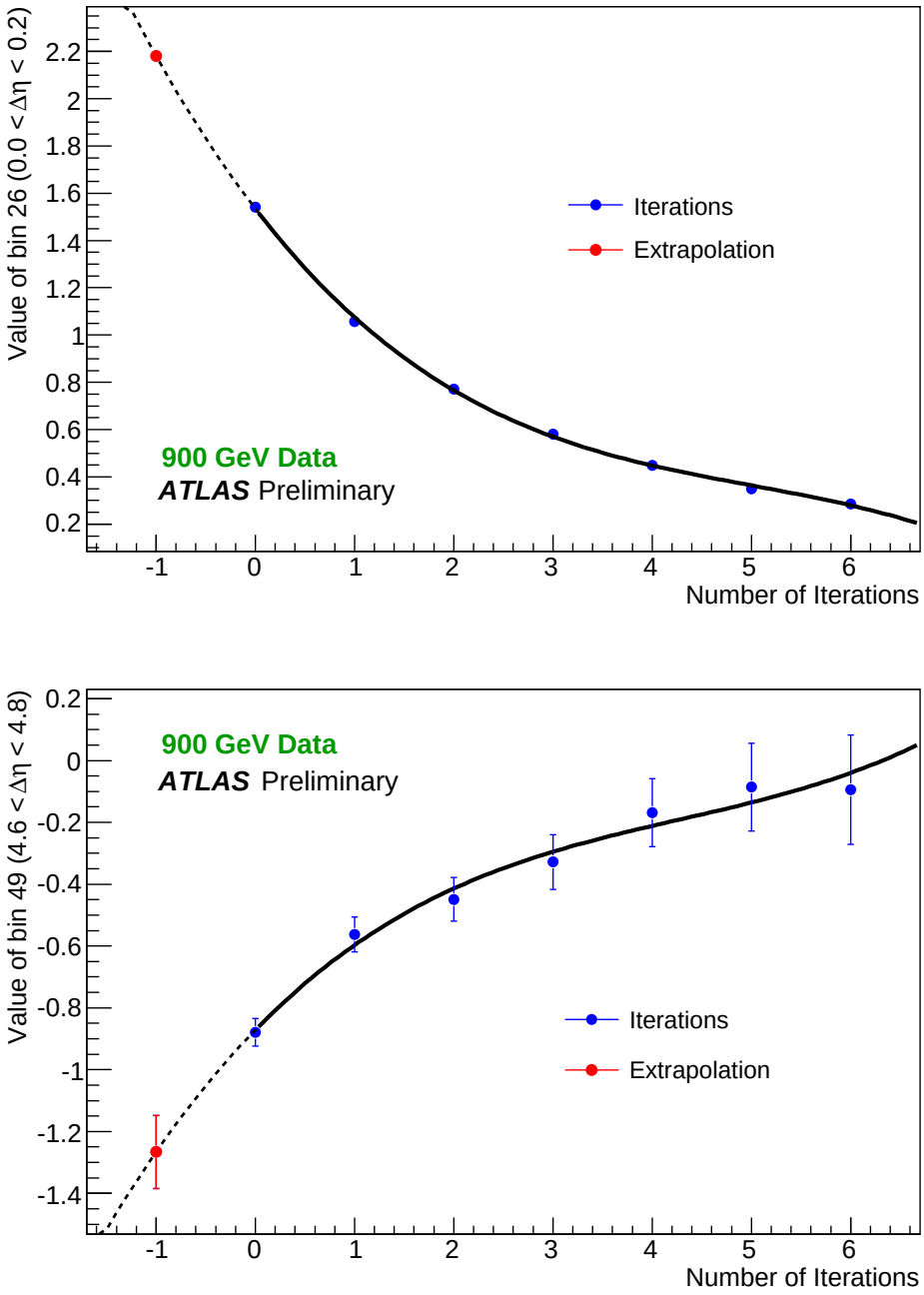

Figure 2: Example of a third degree polynomial fit to the values of R($\Delta\eta$) in a central (top) and an outer (bottom) bin for 900 GeV data. The N= -1 point, extrapolated from the fit to the N=0 to N=6 points, is included to show the stability of the fit. The error bars of N= -1 represent the error in extrapolation. The statistical uncertainties in the top plot are smaller than or of the same size as the markers.

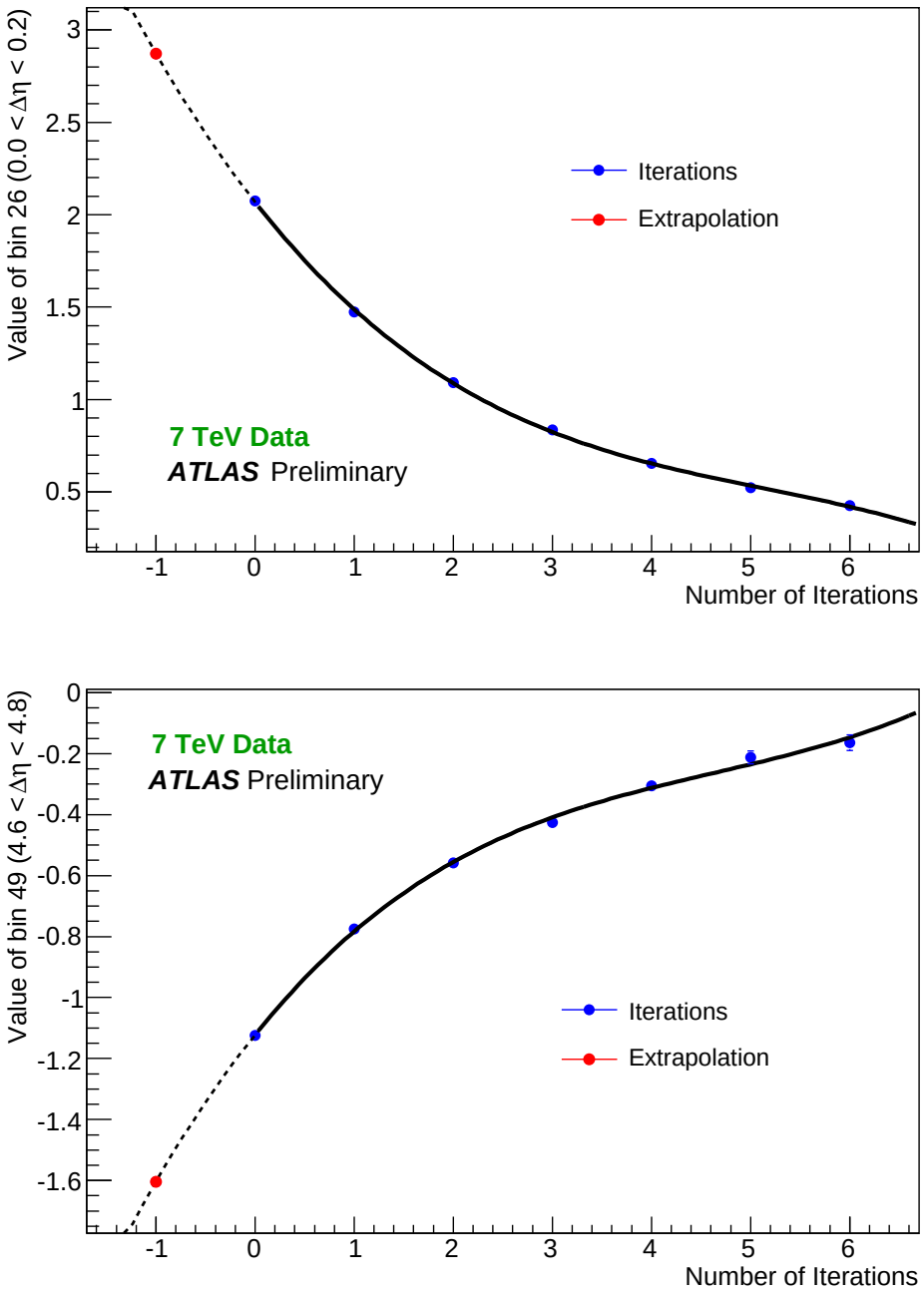

Figure 3: Example of a third degree polynomial fit to the values of R($\Delta\eta$) in a central (top) and an outer (bottom) bin for 7 TeV data. The N= -1 point, extrapolated from the fit to the N=0 to N=6 points, is included to show the stability of the fit. The error bars of N= -1 represent the error in extrapolation. The statistical uncertainties are smaller than or of the size of the markers.

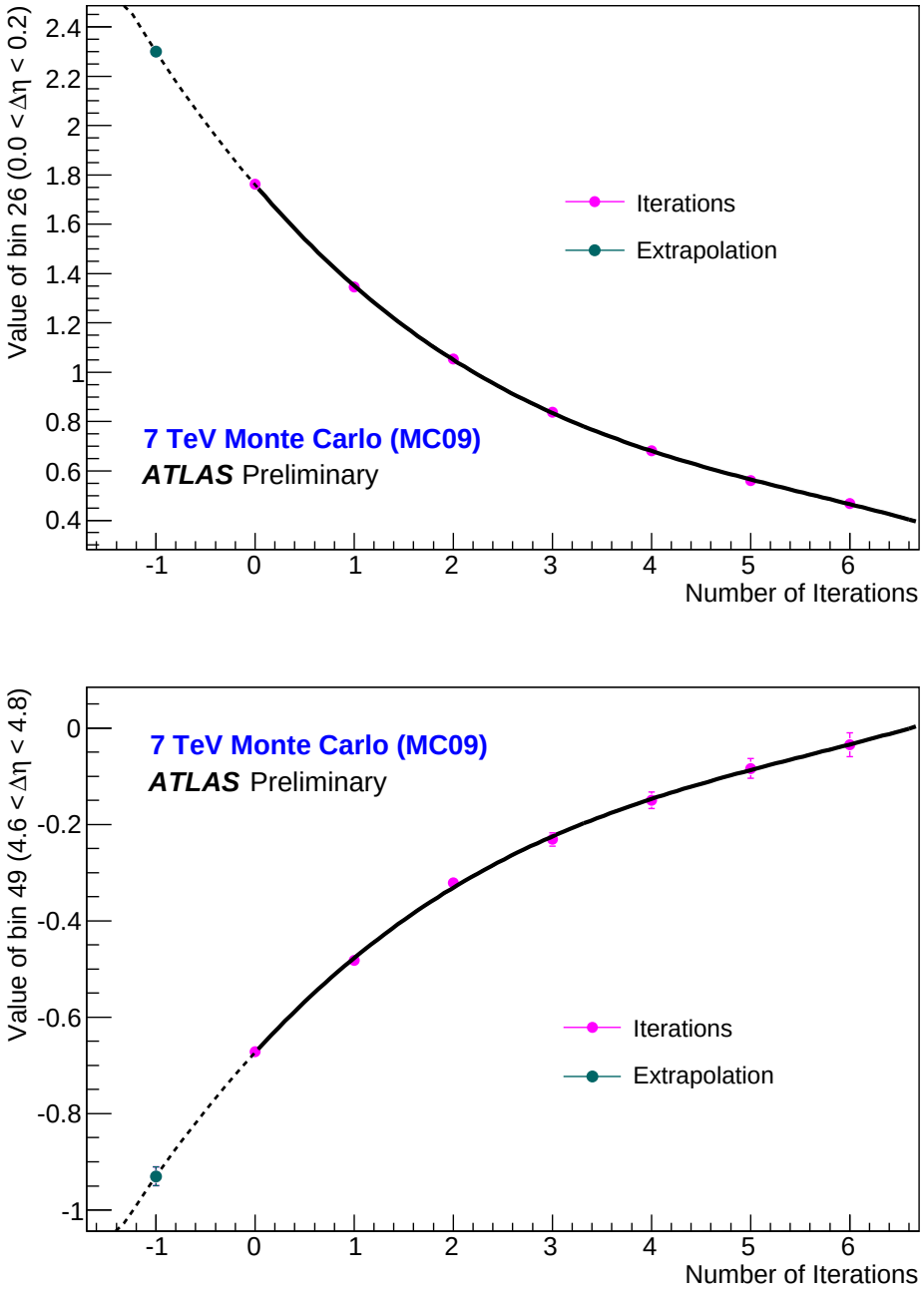

Figure 4: Third-degree polynomial fit to the values of the Monte Carlo R($\Delta\eta$) distribution in a central (top) and outer (bottom) bin. The N= -1 point, extrapolated from the fit to the N=0 to N=6 points, is included to show the stability of the fit. The error bars of N= -1 represent the error in extrapolation. The statistical uncertainties in the top plot and for some points in the bottom plot are smaller than or of the same size as the markers.

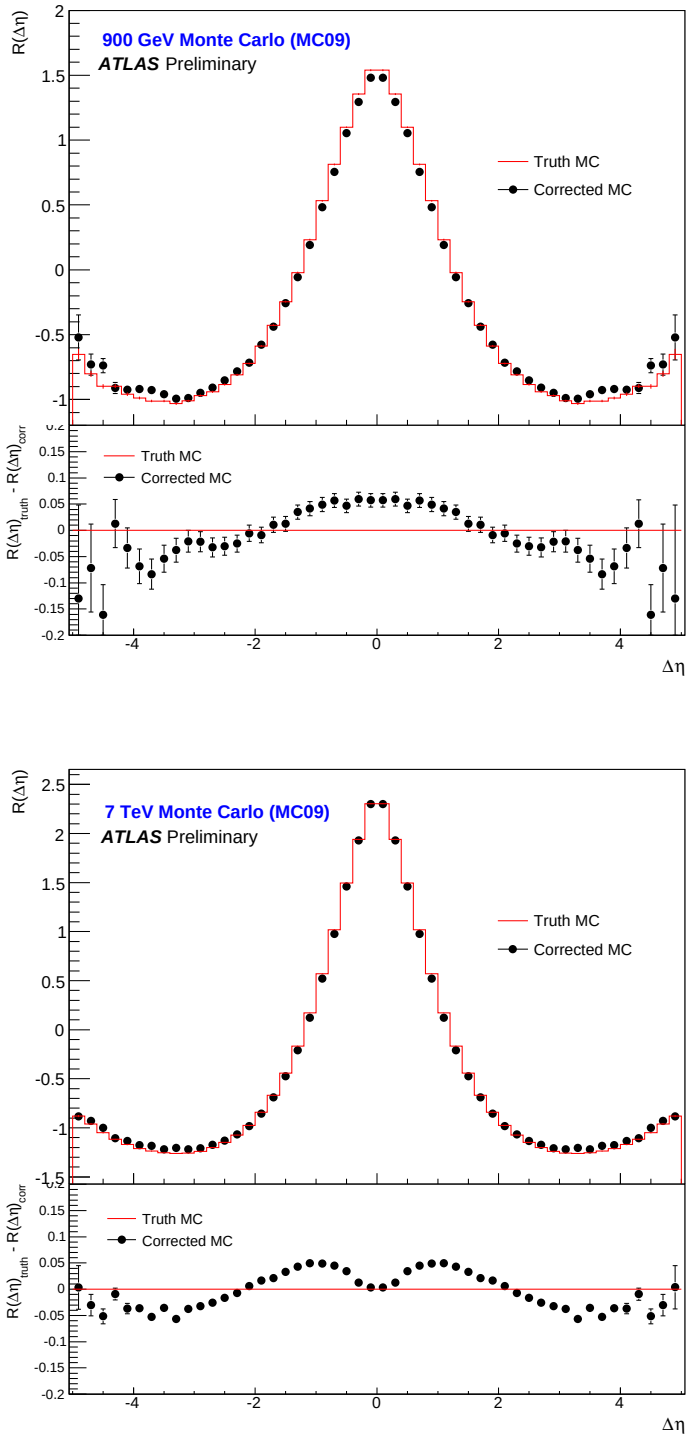

Figure 5: Comparison between the generated and corrected Monte Carlo R($\Delta\eta$) distributions at 900 GeV (top) and 7 TeV (bottom). Good closure is obtained with this correction method. The absolute difference between the truth and corrected MC is plotted at the bottom of the distributions and a maximum value of ~ 0.05 is observed at 7 TeV.

5 Statistical and Systematic Uncertainties

This section describes the sources of statistical and systematic uncertainty identified for this analysis. The most significant contributions to the total uncertainty are associated to the track reconstruction efficiency and to the correction procedure (arising from the small differences between the generated and corrected Monte Carlo distributions). Exactly the same techniques were applied for both the 1D and 2D distributions.

5.1 Extrapolation to N$= -1$

From the fits done to extract the corrected distribution in the probabilistic track removal method, a third-degree polynomial equation is obtained and each of its parameters carries a statistical error. The statistical error in the corrected value will be the result of propagating these statistical uncertainties when evaluating the equation at N $= -1$. As a further test of how well the extrapolation method works, and of whether a polynomial of degree three is a suitable choice, a fit was made to only the iterations for which N ≥ 1 and extrapolated back to N $= 0$. This extrapolated estimate for the N $= 0$ point was found to be close to the known true point. We interpret this test to be a test of whether the corrections method works, rather than a measure of the uncertainty in the extrapolation. As such, it is commensurate with the non-closure systematic assigned and discussed in section 5.3.

5.2 Uncertainties on the Efficiencies

To evaluate the systematic uncertainty in the two-particle correlation distribution associated to uncertainties in the trigger, vertex and track reconstruction efficiencies from data, the complete analysis was repeated varying each of these quantities accordingly, one at a time. These systematics are considered to be the same for both energies and a more detailed description of these uncertainties can be found in [7].

The systematic uncertainties associated to the tracking efficiency come from different sources: material effects, track selection, p_{T} resolution and badly measured high-p_{T} tracks. The uncertainty due to the material description in Monte Carlo increases with increasing $|\eta|$ and decreases as the track p_{T} grows, reaching a maximum of 15% in the range $2.3 < |\eta| < 2.5$ and $100 < p_{\mathrm{T}} < 150$ MeV. A variation of 1% is assigned to all bins due to the track selection and an additional 5% systematic uncertainty is given to the first p_{T} bin to account for the different momentum resolution between data and simulation. If the track has a p_{T} greater than 10 GeV, two additional contributions need to be considered: a 10% uncertainty must be added to cover the different efficiencies of the track-fit χ^2 cut between Monte Carlo and data and an uncertainty due to mis-measured tracks depending on both the η and the p_{T} of the track. All of these uncertainties are added in quadrature.

In the case of the trigger and vertex efficiencies, the systematic uncertainties are of the order of 1% in the first multiplicity bin, decreasing rapidly as the multiplicity increases.

5.3 Uncertainty Due to Lack of Closure in Monte Carlo

The small discrepancies between the generated and corrected distributions in Monte Carlo (non-closure) are considered to be a systematic uncertainty.

It is possible to use the distributions in Figure 5 to define a relative fractional non-closure uncertainty as a function of $\Delta\eta$ and then apply that relative uncertainty to the data distribution. However, at some $\Delta\eta$ values close to $\Delta\eta = 1.5$ the relative uncertainty derived from Monte Carlo is large simply because the model predicts a small value of R for that $\Delta\eta$. Applying such a relative uncertainty to the data would result in a spuriously large non-closure estimate because data and Monte Carlo do not agree on the values of $\Delta\eta$ at which R is close to zero.

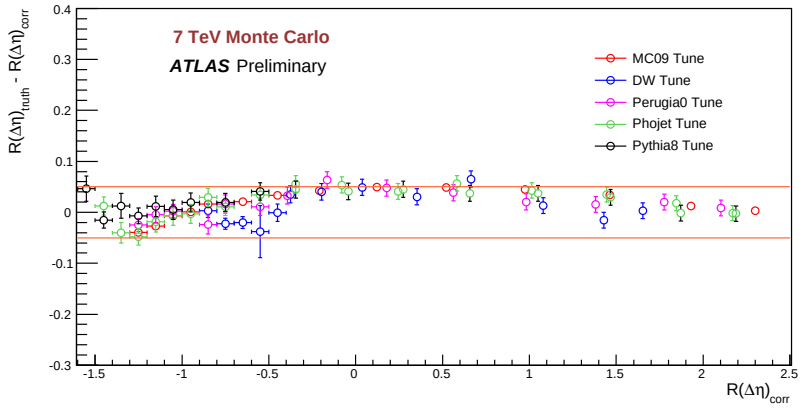

Figure 6: Difference between the generated and corrected two-particle pseudorapidity distributions as a function of $R(\Delta\eta)_{corr}$ for each Monte Carlo tune at 7 TeV. Bins with similar values of $R(\Delta\eta)_{corr}$ have been merged to reduce the statistical fluctuations and better illustrate the behaviour of the models.

The absolute difference between corrected Monte Carlo and Monte Carlo truth is therefore used as a systematic uncertainty resulting from the small non-closure of the correction procedure. The absolute difference between corrected Monte Carlo and truth is shown as a function of the corrected value of R in Figure 6 for all of the different Monte Carlo models used. It would be possible to use these distributions to derive a non-closure estimate for data as a function of the observed values of R. However, since the absolute amount of non-closure is not strongly dependent on either the Monte Carlo model or the value of R and is generally in the region of or less than 0.05, our lack of confidence in the closure of the corrections method can be expressed by a constant systematic uncertainty of 0.05 in each bin.

The non-closure of the Monte Carlo samples is, if not entirely then at least partly due to the effect of secondary decays of longer-lived particles. A long-lived particle whose species has a mean lifetime greater than $c\tau = 10$ mm can occasionally decay inside the tracking cuts of $|d_0|$ <1.5 mm and $|z_0 \sin(\theta)|$ <1.5 mm. Such a particle contributes to the uncorrected distributions, but is not present in the generator truth level distributions. Since the Monte Carlo generators generally overestimate the production of such secondaries compared to data [7], the resulting non-closure ought to be larger for Monte Carlo than for data. The non-closure systematic uncertainty therefore encompasses any uncertainty resulting from the secondary decays of longer lived particles. In order to fully correct for the effect of secondary vertices, the longer-lived neutral particles and their decay products would have to be identified in Monte Carlo, and their contribution to the correlations distribution determined. Given that the effect of secondaries is small and included in the non-closure uncertainty, and that correcting for the effect would require quite an involved procedure dependent on Monte Carlo, no particular correction is applied to account for secondary decays of longer-lived particles.

In addition to the non-closure due to secondaries, there may be other model-dependent effects, which would make a different contribution to the (unknown) non-closure in data. If such model-dependent effects are significant then it would be expected that the amount of non-closure shown by the different models in Figure 6 would vary model-to-model. In fact, the different models show a broadly similar amount of non-closure, and there are no models that show a consistently higher or lower amount of non-closure; the distributions from the different models are interleaved. Any model-dependent effects would therefore have to act in such a way as to increase the amount of non-closure at some values of R whilst decreasing it at others in order that all the models appear broadly similar in their amount of non-closure. It therefore seems unlikely that the non-closure amount in data could be higher for all data points than the

value of 0.05 estimated from Monte Carlo. Further, as will be seen in Figure 8, the difference between the correlations distributions from the different models (in particular Pythia 6 tune DW and Pythia 8) is larger than the difference between a given model and the data. The variation between the models is greater than the difference between any model and the data, yet all of the models show a similar amount of non-closure.

6 Results

The corrected inclusive two-particle correlation functions for data and Monte Carlo are shown in Figure 7 for 900 GeV and 7 TeV. The total uncertainties of these 2D plots are shown in Figures 15 and 16 in Appendix B. At both energies a complex structure can be observed across the full $\Delta\eta$ and $\Delta\phi$ range. Several components can be identified, for example, a sharp peak around (0,0) which can be associated to particles coming from the same high-p_{T} cluster, e.g. jets (near-side correlations), a ridge extending across the whole $\Delta\eta$ range around $\Delta\phi \sim \pi$ coming from, for instance, back-to-back jets (away-side correlations) and, a gaussian-shaped distribution across all $\Delta\phi$ with a width of $|\Delta\eta| \sim 2$ related to resonances, string fragmentations, low-p_{T} clusters, etc (short-range correlations). Similar structures are observed in the Pythia tune MC09, however the strength of the correlation seen in the data is not reproduced by the model.

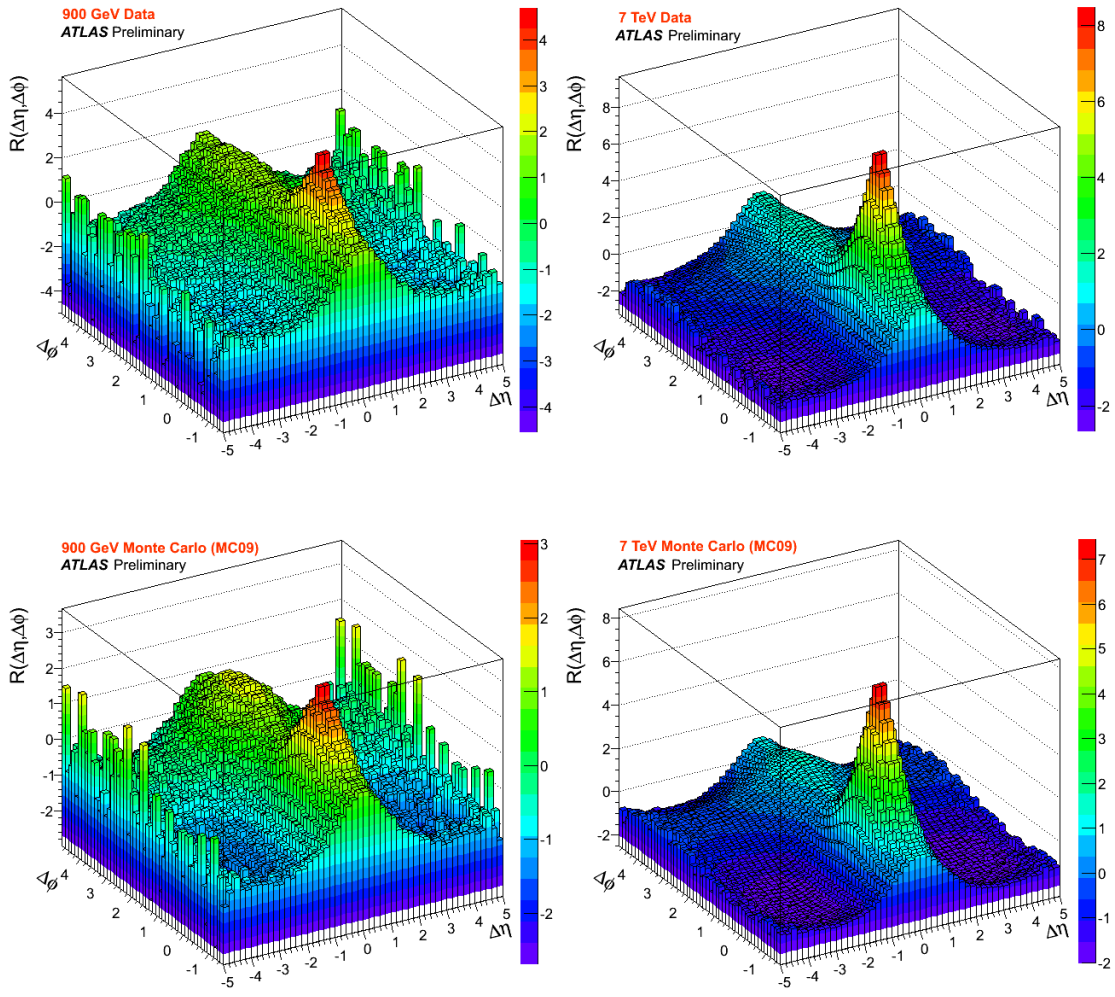

Figure 7: Corrected two-particle correlation distribution functions in $\Delta\eta$ and $\Delta\phi$ for 900 GeV (left) and 7 TeV (right) for data (top) and Monte Carlo (bottom). By construction, these plots are symmetric around $\Delta\eta = 0$ and $\Delta\phi$ is plotted from $-\frac{\pi}{2}$ to $\frac{3\pi}{2}$ to avoid splitting the away-side region.

6.1 Projections in $\Delta\eta$ and $\Delta\phi$

Projections of the two-dimensional correlation function along both $\Delta\eta$ and $\Delta\phi$ reveal more details of the structure of the correlations. These projected distributions were corrected in the same way as the two-dimensional distributions using the method described in Section 4. Likewise, the calculation of the total systematic uncertainty follows the techniques presented in Section 5. Note that these projections are obtained by first integrating the foreground and background two-dimensional distributions separately and then taking the ratio of the two and normalising by the average track multiplicity.

The corrected pseudo-rapidity correlation distributions are shown in Figure 8, obtained by integrating $\Delta\phi$ from $\{0 : \pi\}$. The blue dots correspond to the corrected values, the blue error bars are statistical only (coming from the extrapolation to N= -1) and the green bands correspond to the total uncertainty in the bin, obtained by adding in quadrature the uncertainties described in Section 5. The data distributions are compared to the different Monte Carlo tunes, and it can be seen that the strength of the correlation is not reproduced by any of the models, Pythia 8 being the closest.

Additionally, the pseudorapidity correlation function was calculated by integrating $\Delta\phi$ over the limited ranges $\left\{0 : \frac{\pi}{2}\right\}$ and $\left\{\frac{\pi}{2} : \pi\right\}$. By focusing on the particle pairs in the near-side (Figure 9), the distributions are dominated by the peak at (0,0) so they are narrower and higher, while in the case of pairs in the away-side (Figure 10), the distributions are flatter and wider. The data points, shown in blue along with their statistical and systematic uncertainties, are compared to the different Monte Carlo tunes. The differences between data and simulation can be seen in both the peak and the tails and are slightly more pronounced in the near-side distributions; in the case of the away-side distributions, the Pythia 8 tune at 7 TeV is the closest to the data.

Similarly, the azimuthal correlation function was studied by integrating $\Delta\eta$ over two different ranges, (0-2) and (2-5). The first projection, shown in Figure 11, focuses on the short-range correlations, including the central peak, and as such displays an "M" shape, with peaks at both $\Delta\phi = 0$ (more pronounced in the 7 TeV data) and $\Delta\phi = \pi$. Again, the data are compared to the different Monte Carlo models and none of them reproduces the observed distributions despite the agreement in some sections of the plots. The second $\Delta\phi$ projection avoids the central peak to study the underlying structure of the long-range correlations. As shown in Figure 12, at $\Delta\phi = 0$ the distribution reaches a minimum while the away-side exhibits the same ridge structure. All of the Monte Carlo models lie above the data points. This type of $\Delta\phi$ projection was recently studied in high-multiplicity events by the CMS collaboration and a ridge-like structure was observed at $\Delta\phi = 0$ in a specific range of the pairs' transverse momentum [14]. In the case of this inclusive analysis, no such structure is present.

The relatively larger statistical and systematic uncertainties in the azimuthal correlation function are associated to the non-closure in Monte Carlo (slightly worse than for the pseudorapidity distributions) and to the decrease in statistics when dividing the sample in two different ranges of integration. The 900 GeV distributions are the most affected.

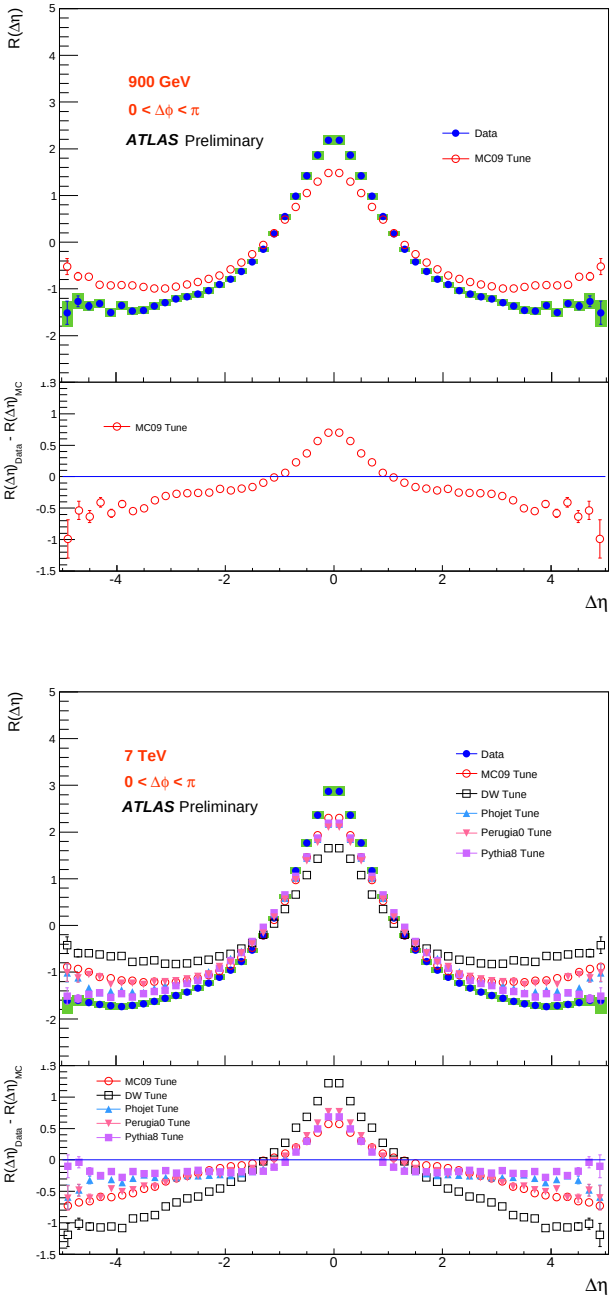

Figure 8: Corrected two-particle pseudorapidity distributions at 900 GeV (top) and 7 TeV (bottom) compared to different corrected Monte Carlo tunes. In data, the green bands correspond to the total uncertainties (statistical and systematic, added in quadrature) and the blue lines are the statistical errors calculated from the extrapolation to N= -1. These plots are symmetric around $\Delta\eta = 0$ by construction.

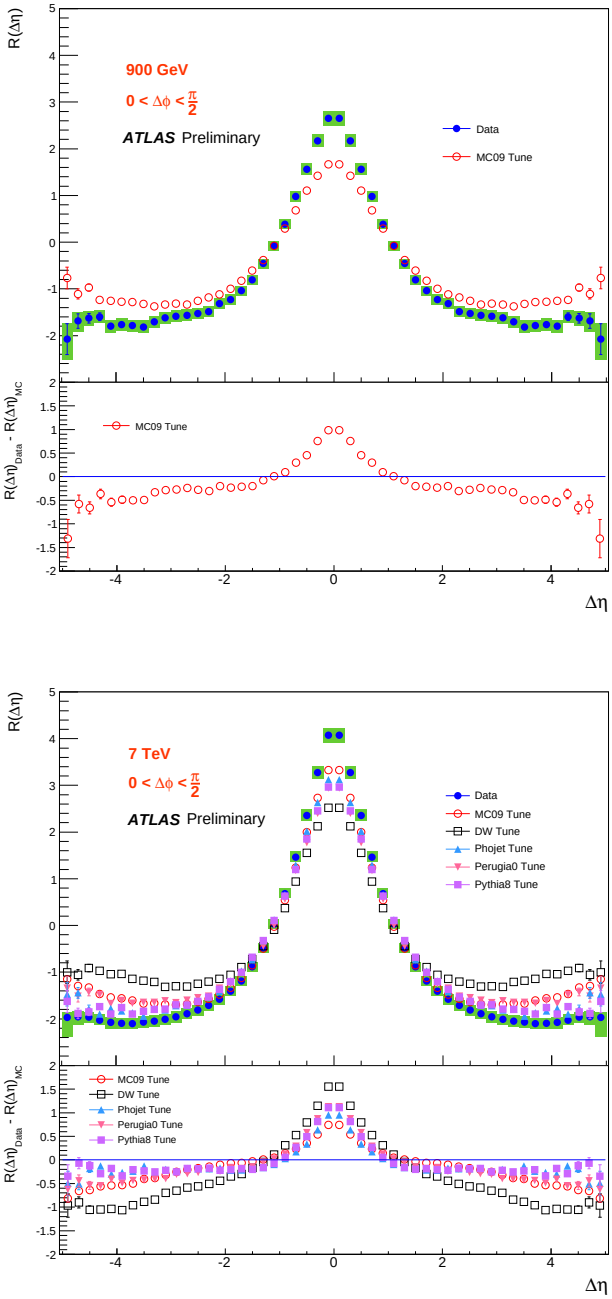

Figure 9: Two-particle pseudorapidity correlation distributions obtained by integrating over the $\Delta\phi$ range from 0 to $\frac{\pi}{2}$ for data and the different Monte Carlo tunes at 900 GeV (top) and 7 TeV (bottom). In data, the green bands correspond to the total uncertainties (statistical and systematic, added in quadrature) and the blue lines are the statistical errors calculated from the extrapolation to N= -1.

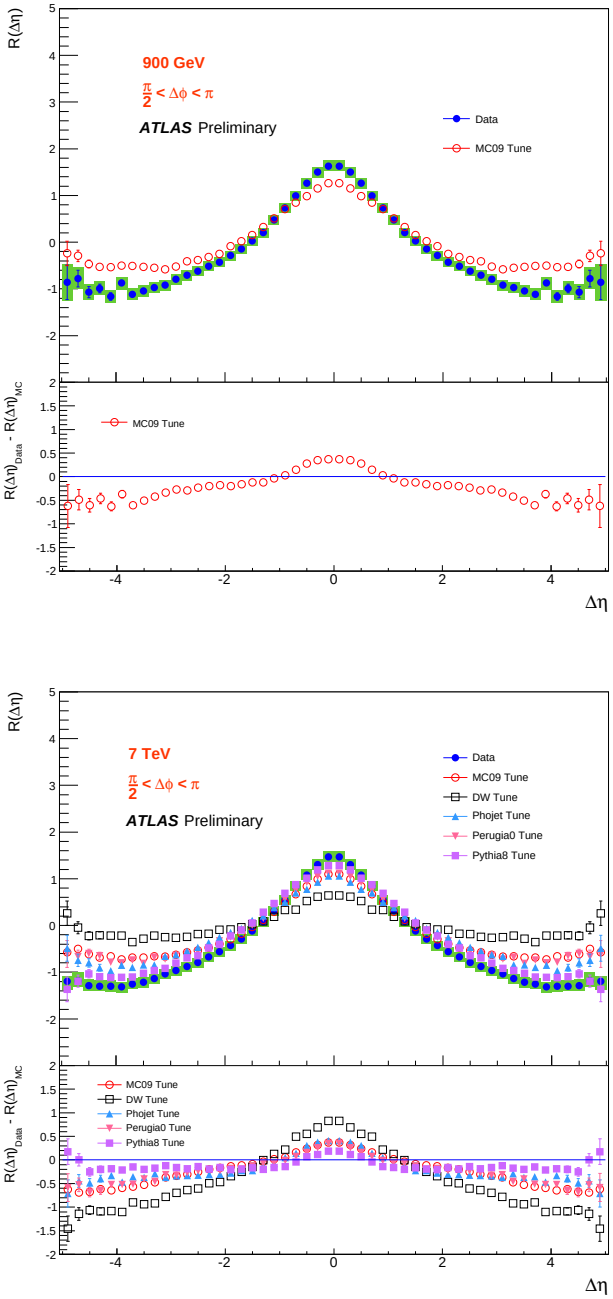

Figure 10: Two-particle pseudorapidity correlation distributions obtained by integrating over the $\Delta\phi$ range from $\frac{\pi}{2}$ to π for data and the different Monte Carlo tunes at 900 GeV (top) and 7 TeV (bottom). In data, the green bands correspond to the total uncertainties (statistical and systematic, added in quadrature) and the blue lines are the statistical errors calculated from the extrapolation to N= -1.

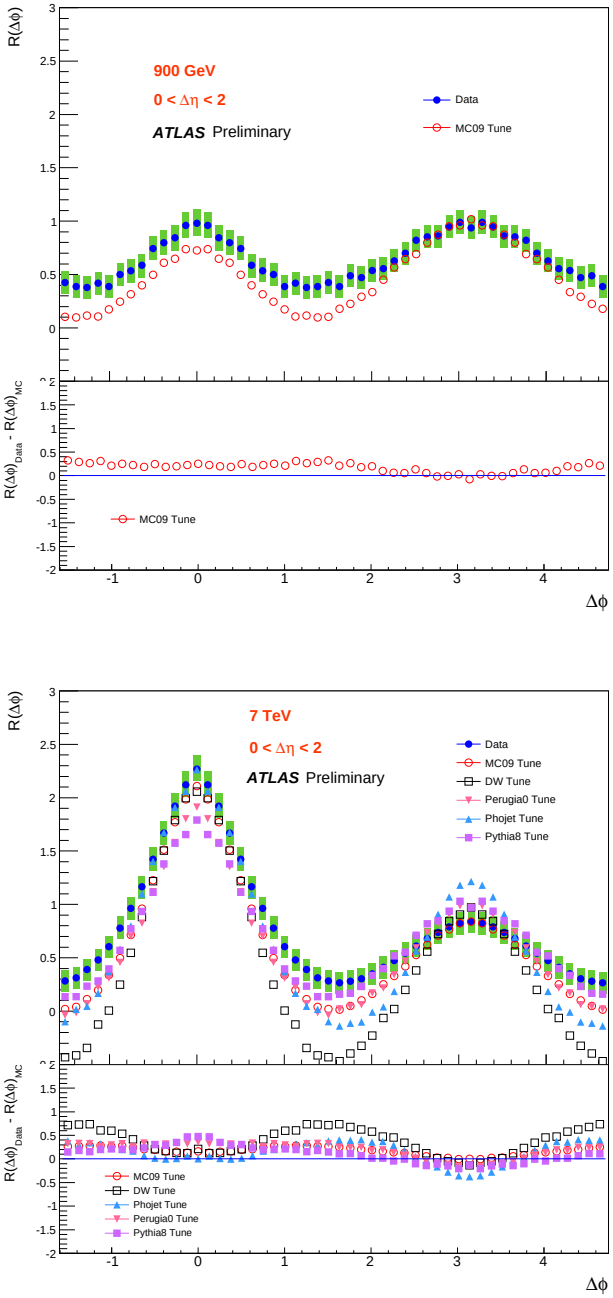

Figure 11: Two-particle azimuthal correlation distributions obtained by integrating over the $\Delta\eta$ range from 0 to 2 for data and the different Monte Carlo tunes at 900 GeV (top) and 7 TeV (bottom). In data, the green bands correspond to the total uncertainties (statistical and systematic, added in quadrature) and the blue lines are the statistical errors calculated from the extrapolation to N= -1.

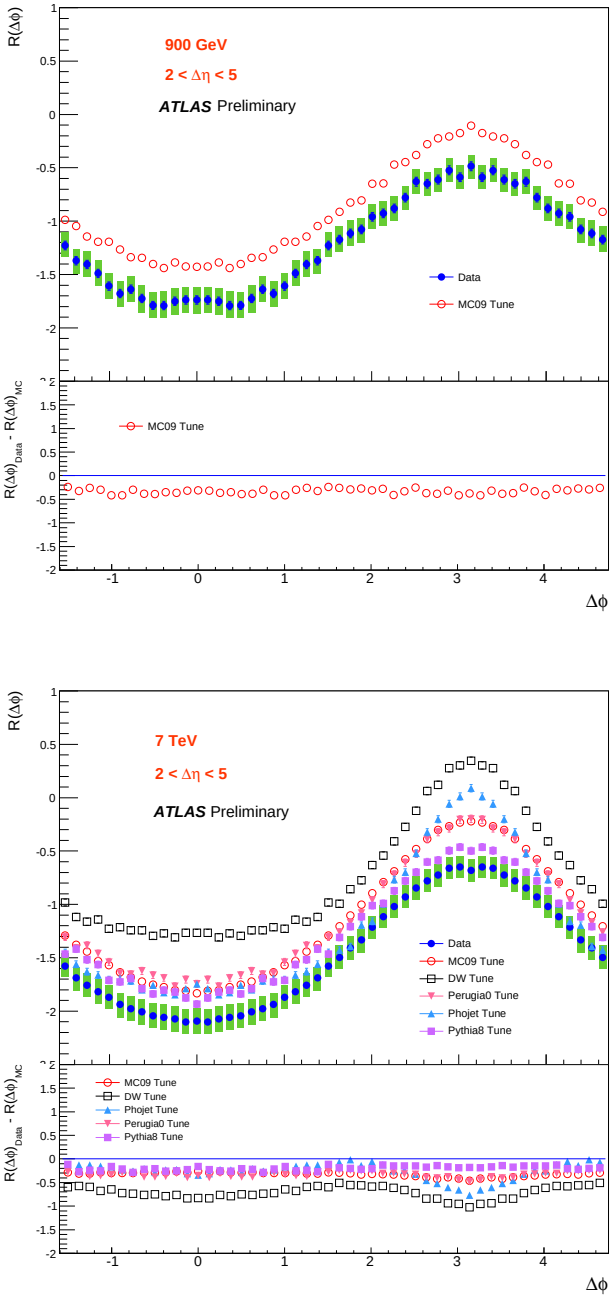

Figure 12: Two-particle azimuthal correlation distributions obtained by integrating over the $\Delta\eta$ range from 2 to 5 for data and the different Monte Carlo tunes at 900 GeV (top) and 7 TeV (bottom). In data, the green bands correspond to the total uncertainties (statistical and systematic, added in quadrature) and the blue lines are the statistical errors calculated from the extrapolation to N= -1.

7 Summary and Conclusions

The two-particle angular correlation functions in $\Delta\eta$ and $\Delta\phi$ have been measured for p_{T} inclusive minimum bias events in proton-proton collisions at two different energies, 900 GeV and 7 TeV. The structure of the correlation was explored in more detail by projecting the two-dimensional distribution into both $\Delta\eta$ and $\Delta\phi$. All of these results have been compared to Monte Carlo samples which show a similar complex structure in $\Delta\eta$ and $\Delta\phi$ but fail to reproduce the strength of the correlations seen in data.

References

[1] B. Andersson et al., *A model for low-p_T hadronic reactions with generalizations to hadron-nucleus and nucleus-nucleus collisions*, Nucl. Phys **B281** (1987) 289.

[2] A. Morel and G. Plaut, *How do clusters look like in semi-inclusive cross sections?*, Nucl. Phys. **B78** (1974) 541–551.

[3] K. Eggert et al., *Angular Correlations Between the Charged Particles Produced in pp Collisions at ISR Energies*, Nucl. Phys. **B86** (1975) 201–215.

[4] UA5 Collaboration, R. E. Ansorge et al., *Charged particle correlations in anti-p p collisions at c.m. energies of 200, 546 and 900 GeV*, Z. Phys. **C37** (1988) 191–203.

[5] PHOBOS Collaboration, B. Alver et al., *Cluster properties from two-particle angular correlations in p+p collisions at $\sqrt{s}$ = 200 and 410 GeV*, Phys. Rev. **C75** (2007) 054913.

[6] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[7] ATLAS Collaboration, G. Aad et al., *Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC*, accepted by New J. Phys. , arXiv:1012.5104v1.

[8] T. Sjostrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[9] T. Sjostrand, S. Mrenna, and P. Skands, *Brief Introduction to Pythia 8.1*, Comput. Phys. Comm. **178** (2008) , arXiv:0710.3820.

[10] R. Engel, *Photoproduction within the two component dual parton model: Amplitudes and cross-sections*, Z. Phys. **C66** (1995) 203–214.

[11] M. Albrow et al., *Tevatron-for-LHC Report of the QCD Working Group*, arXiv:hep-ph/0610012.

[12] P. Skands, *The Perugia Tunes*, arXiv:0905.3418 [hep-ph].

[13] ATLAS Collaboration, *ATLAS Monte Carlo Tunes for MC09*, ATL-PHYS-PUB-2010-002.

[14] CMS Collaboration, V. Khachatryan et al., *Observation of Long-Range, Near-Side Angular Correlations in Proton-Proton Collisions at the LHC*, JHEP **09** (2010) 091.

Appendices

A Uncorrected Two-particle Correlation Distributions

The correlation distributions in $\Delta\eta$ and $\Delta\phi$ before applying the correction method are shown in Figure 13 and the uncorrected pseudorapidity correlation distributions are presented in Figure 14.

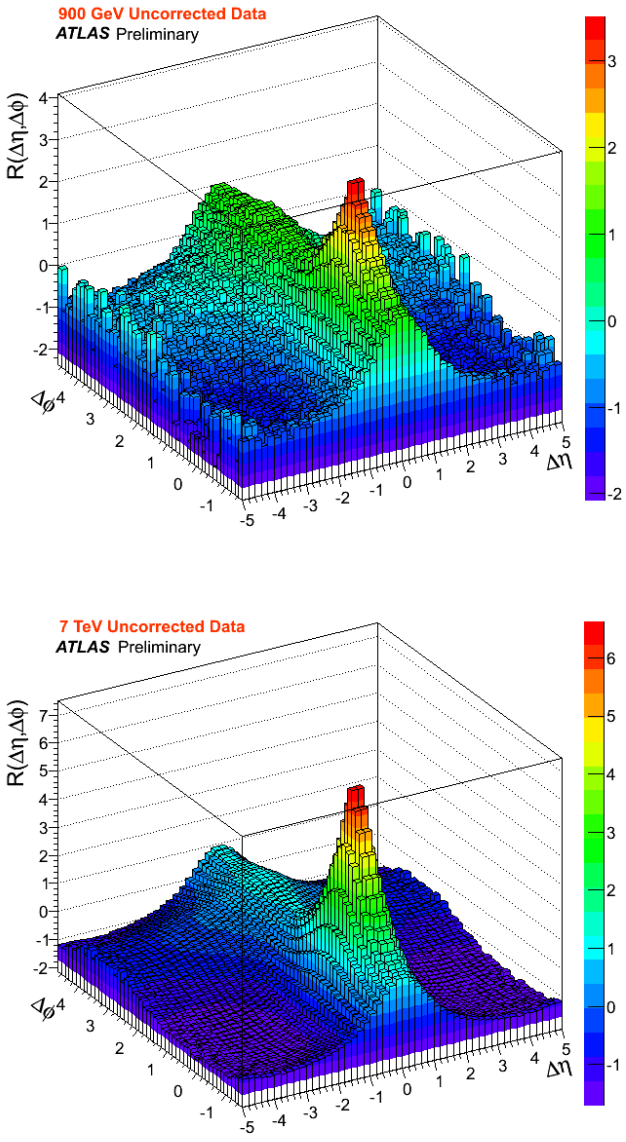

Figure 13: Uncorrected two-particle correlation functions in $\Delta\eta$ and $\Delta\phi$ at 900 GeV (top) and 7 TeV (bottom). These plots are symmetric around $\Delta\eta = 0$ by construction.

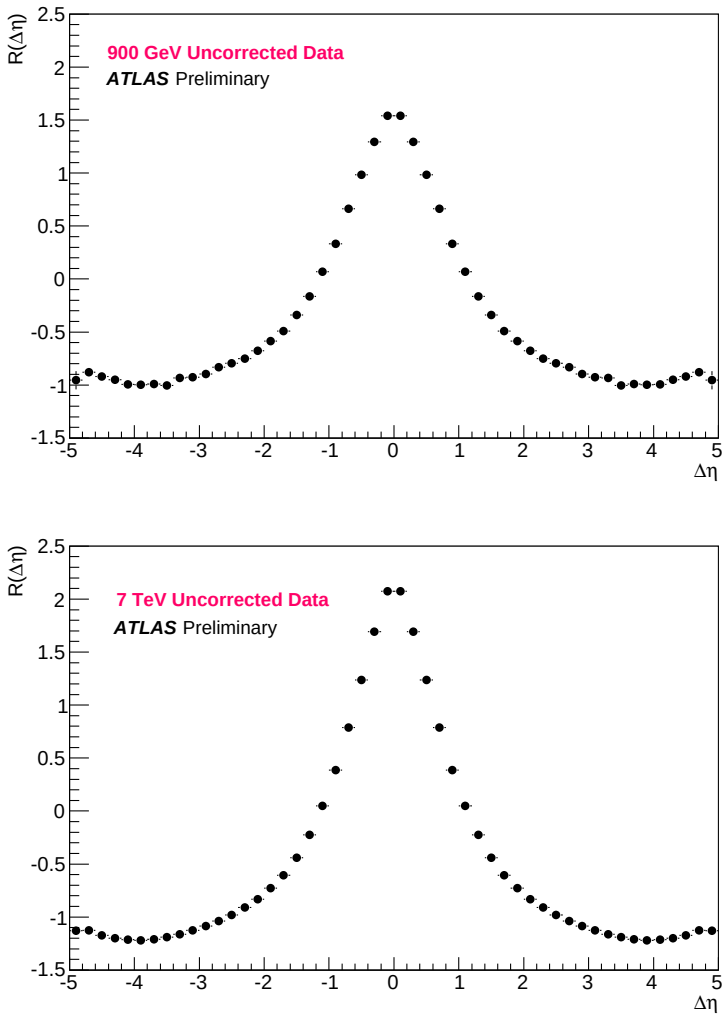

Figure 14: Uncorrected two-particle correlation functions in $\Delta\eta$ at 900 GeV (top) and 7 TeV (bottom). These plots are symmetric around $\Delta\eta = 0$ by construction.

B Total Uncertainties for the 2D Correlation Functions

The total uncertainty in each bin of $R(\Delta\eta, \Delta\phi)$ for the 900 GeV and 7 TeV data are shown in Figures 15 and 16, respectively.

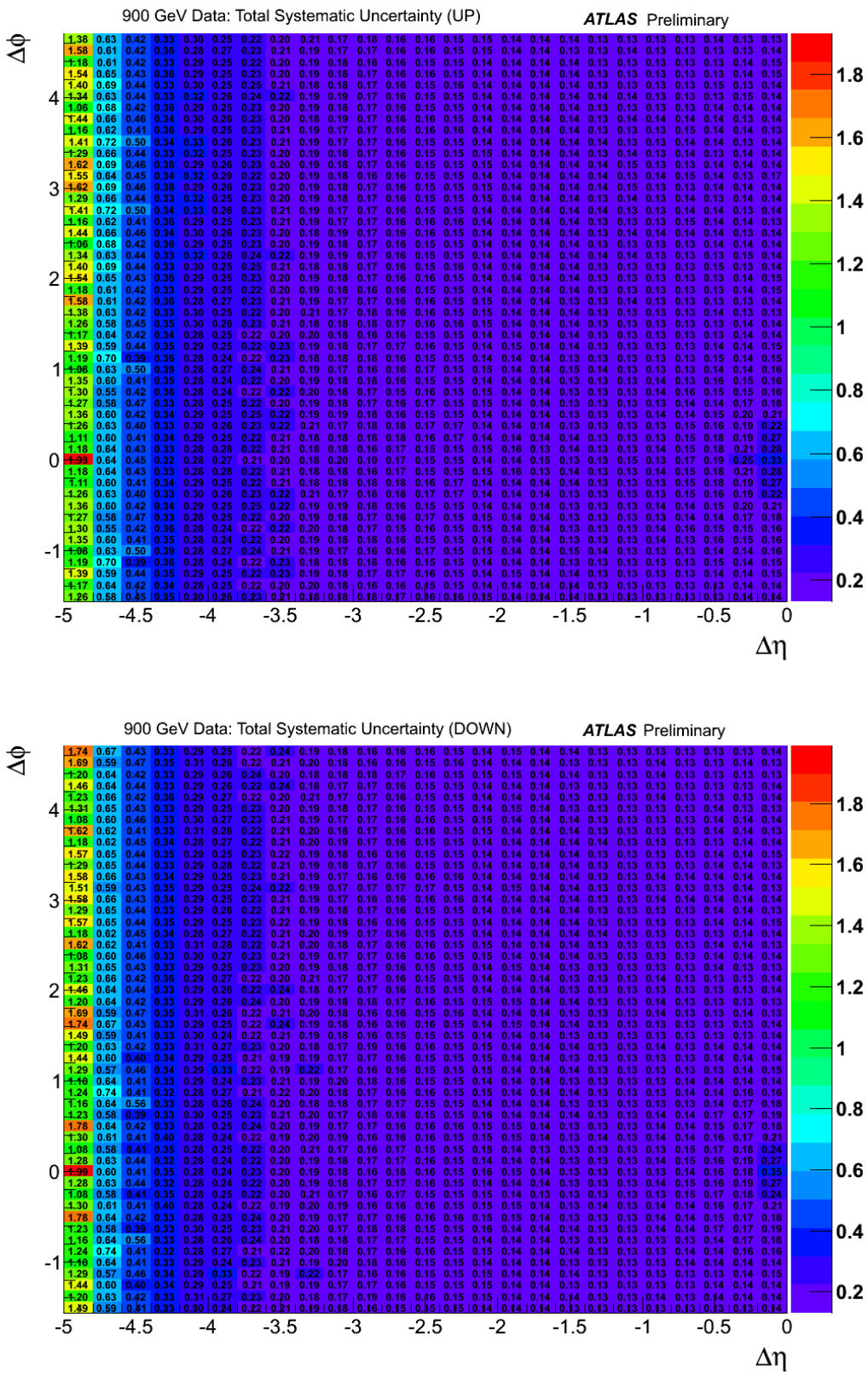

Figure 15: Total systematic uncertainty for each bin in the two-dimensional correlation function for 900 GeV data. Since the distributions are symmetric, only the $\Delta\eta$ range from -5 to 0 is shown.

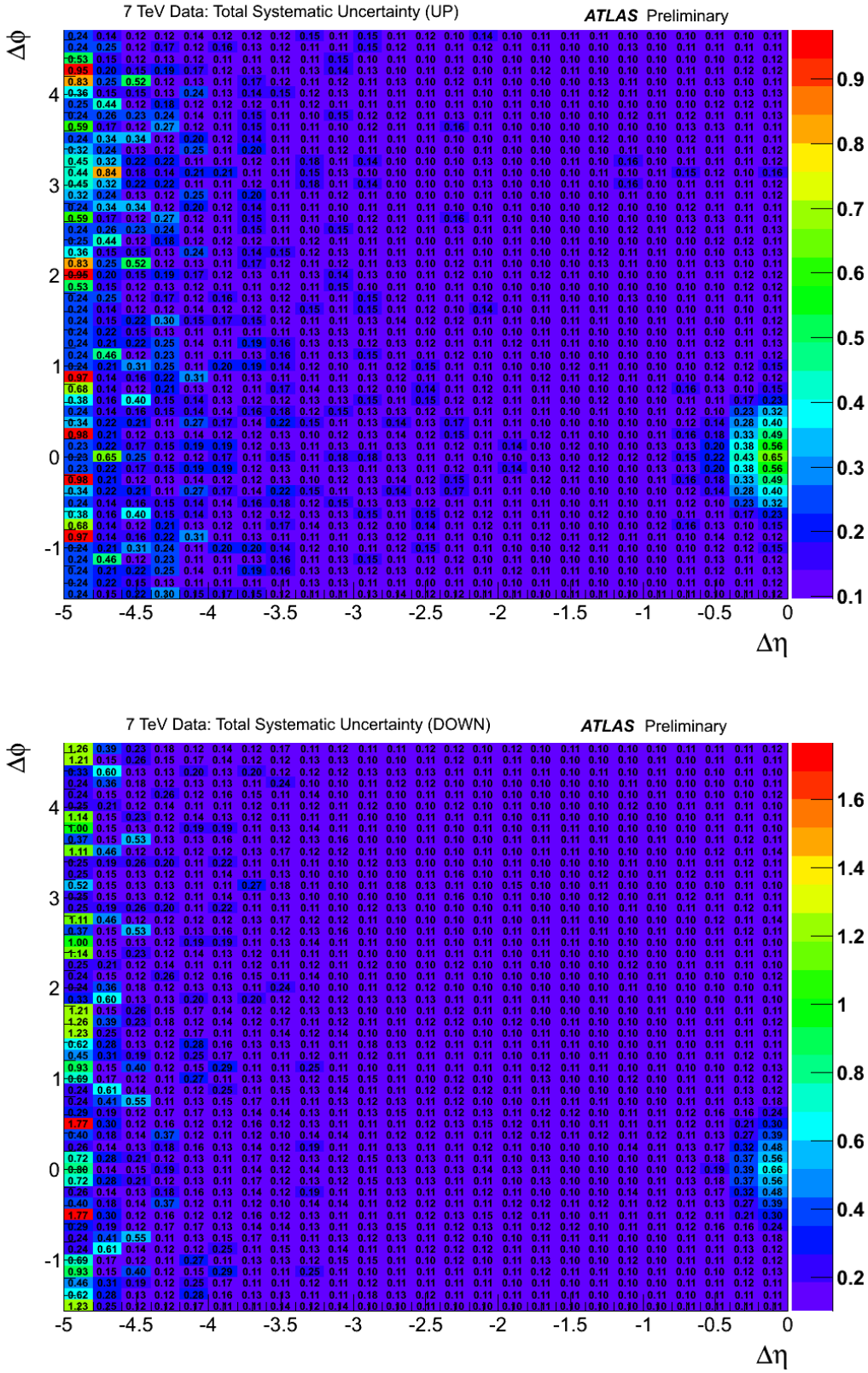

Figure 16: Total systematic uncertainty for each bin in the two-dimensional correlation function for 7 TeV data. Since the distributions are symmetric, only the $\Delta\eta$ range from -5 to 0 is shown.