

Influence of Anisotropic Flow on Jet Background Estimation

Govert Nijs 3477223

Supervisors:

Marco van Leeuwen & Redmer Bertens

June 21, 2013

Abstract

Anisotropic flow parameters are estimated on an event-by-event basis using methods with small biases. Soft jet backgrounds are estimated using k_t jets and modulated with the flow parameters. The quality of the modulation is assessed using random cones and δp_T^{RC} distributions, and the hard parton jet spectrum is obtained using anti- k_t jets.

Contents

1	Introduction	1
1.1	Events and their detection	2
2	Anisotropic flow	2
2.1	Events and tracks used	3
2.2	Event planes	5
2.3	Obtaining v_2 and v_3	7
2.4	Removing autocorrelations	9
2.5	Assessing quality of v_2 and v_3	11
2.6	Resolution correction	12
3	Jets	14
3.1	Jet Finder	14
3.2	Jets and tracks used	15
3.3	Background estimation	15
3.4	Random cones	15
3.5	Improving background estimation	17
3.6	Anti- k_t jet spectra	21
4	Conclusions and Outlook	21

1 Introduction

At extreme energy densities, it is expected that strongly interacting matter will undergo a phase transition to the so-called quark-gluon plasma (QGP). To study this new form of matter, for one month each year, the LHC collides lead ions. The

ALICE detector is optimized to study these collisions.

In this paper, we investigate the properties of the QGP using jets. Jets are a spray of particles originating from hard parton interactions. These hard interactions can be calculated in perturbative QCD, but as they travel through the QGP, they interact with it, causing them to lose momentum. This momentum loss is dependent on the density of the QGP, and the distance traveled within it. Since in many cases the QGP is not spherically symmetric (see Figure 1), we can relate the distance traveled by a jet to the direction it traveled in. If the QGP is elongated and the jet travels in the same direction, it is likely to have traveled a greater distance through the QGP than a jet in another direction. The jet spectrum without the presence of a QGP can be calculated from pQCD. Therefore a measurement of their spectrum after a passage through a QGP provides information about the QGP's properties. To do such an analysis, we need to quantify the anisotropy in the QGP, and we need to be able to separate the contribution to the momentum spectrum due to hard parton scattering (which we intend to study) from the soft contribution, which is caused by the thermal expansion of the QGP. Difficulties lie in two things: Firstly, we need to identify the jets. This is done by grouping the tracks with a so-called jet finding algorithm. The groups which contain tracks that originated

from a hard scattering process have to be identified. Secondly, what we measure is not the momentum of such a jet, but instead the momentum of the jet plus the soft background of the medium it travelled through. Therefore, to obtain the momentum of a jet, the soft background needs to be subtracted. In the next section, we will quantify the anisotropy of the QGP, and doing this also immediately helps us in modelling the soft background to the jets, something we will discuss in section 3.

1.1 Events and their detection

When two nuclei collide, they mostly pass through each other, leaving highly excited matter in their wake. This matter goes through a number of phases before it is detected (see also Figure 1):

- Pre-equilibrium: the system thermalizes to reach equilibrium.
- Equilibrium: This is the QGP we intend to study.
- Freeze-out: As the rate of interactions drops, the particles decouple, potentially decay, and fly away to form the final state particles of the collision, which can be measured in a detector.

In each phase, the collision system expands. This collective expansion is called flow. In the equilibrium phase the expansion can be described very well with a hydrodynamical model. Predictions for properties like anisotropic flow depend strongly on parameters of this model, such as its Equation of State (EoS) and viscosity. Additionally, properties like jet spectra depend on the energy loss of jets in the QGP, which is dependent on its characteristics. Hence both anisotropic flow and jet spectra can be used to study the QGP.

Two nuclei rarely collide head-on. Instead, at the moment of collision, the interaction region has the shape of a lens which is flattened along the beam axis.¹ The anisotropy of this region is determined by the so-called impact parameter, which unfortunately cannot be measured directly. We can, however, measure per event the number of charged particles. We then define an event's centrality as the percentage of events which are 'more central',

¹This is due to Lorentz contraction of the two highly relativistic nuclei.

i.e. the percentage of events with more charged particles than this one. Most measured properties of the event depend on its centrality.

The ALICE detector consists of several smaller detectors. Of these, we use the Inner Tracking System (ITS), Time Projection Chamber (TPC), V0A and V0C detectors. The ITS and TPC reconstruct individual tracks, computing momenta using track curvatures and the strength and direction of the magnetic field. The V0A and V0C detectors do not detect individual tracks, but instead detect energy deposition in a certain area in the (η, φ) -plane, where η denotes pseudorapidity,² and φ denotes the azimuthal angle. The TPC has the advantage of a higher resolution,³ while the V0s can be used to determine parameters in a way which is independent of the tracking system. A picture of ALICE is shown in Figure 2.

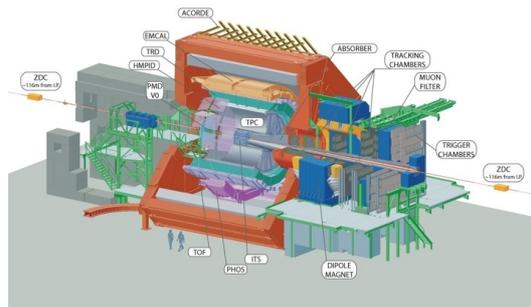


Figure 2: The ALICE detector.

2 Anisotropic flow

Immediately after a collision, the excited matter that has been formed has an anisotropic shape. This anisotropy is dependent on the geometry of the collision, i.e. the shape of the overlap region of the two nuclei, which is in turn dependent upon the collision's centrality. The initial anisotropies translate (in the hydrodynamical model) to different pressure gradients in different directions, which in turn cause anisotropies in the final measured state. However, local fluctuations in nucleon density of the nuclei before the collision also contribute

²We have $\eta = -\ln(\tan(\frac{\theta}{2}))$, with θ denoting the polar angle.

³A track can have a maximum of 159 hits in the TPC.

to the initial anisotropies. This leads to substantial variation between events of final state anisotropies.

To describe azimuthal angular dependence in the measured state, we write the particle density per unit azimuthal angle $dN/d\varphi$ as a Fourier expansion in φ .^[5] In this expansion, the first order Fourier coefficient equals zero, and fourth and higher order coefficients can be neglected at most desired precisions. The angular distribution then becomes

$$\frac{dN}{d\varphi} = v_0 \{1 + 2v_2 \cos[2(\varphi - \Psi_2)] + 2v_3 \cos[3(\varphi - \Psi_3)]\}. \quad (1)$$

We have written this formula in a way which makes explicit the convention to quote v_n as a percentage of v_0 . In the same way, we can describe the transverse momentum distribution as a function of φ , yielding the same formula as above but with $dp_T/d\varphi$ instead of $dN/d\varphi$. This characterizes only the φ -dependence of the collision. In principle, the various parameters could depend on pseudorapidity η and transverse momentum p_T . It is known experimentally that for the particles we will be using, there is no dependence on η .^[4] There is, however, a dependence on p_T . One can take this into account and make v_2 and v_3 in (1) dependent on p_T . In this case we talk about differential flow. In our study, we don't do this, so our parameters describe integrated flow.

This section first describes the events and tracks used, then the method for obtaining flow coefficients is discussed. Finally, biases are identified and when possible corrected.

2.1 Events and tracks used

In this study, we use data from 2010, and we use the following quality criteria:

- **Primary vertex:** We determine the z -coordinate of the primary vertex with the SPD (the innermost part of the ITS) and with the TPC. The two estimates are required to be within 5 mm of one another. Also, the z -coordinate must be within 10 cm of the nominal vertex.
- **Centrality:** Centrality is measured with the V0 detectors. This centrality is required to be within 5% of the centrality as measured with the TPC.

The first of these requirements is made because the TPC is less efficient for large values of $|\eta|$. We want the track recording efficiency to be uniform in the range where we use it so as not to create a bias in the results, and for events with a primary vertex close to the center, we know that the region of uniform efficiency is largest in η . The second requirement is made so that the uncertainty in the centrality is small. This is needed because most parameters to be determined are strongly dependent on centrality.

For each event, we only use certain tracks. This is again to make sure that we don't create biases because of nonuniform tracking efficiency. We use the following requirements on the tracks in our analysis of flow:⁴

- **Type of tracks:** We use hybrid tracks. These are tracks with at least 70 hits in the TPC, and in addition either with an SPD hit and an ITS refit, or with just an ITS refit constrained to the primary vertex, or with no refit at all.
- **p_T :** We require that $0.15 \text{ GeV} \leq p_T \leq 5 \text{ GeV}$.
- **η :** We require $|\eta| \leq 0.9$, and we want $|\eta - \eta_{\text{leading}}| \geq 0.2$, where η_{leading} is the pseudorapidity of the leading anti- k_t jet.

The design of the ALICE detector has limitations. This leaves some of the tracks being of better quality than others. This is the reasoning behind the conditions for hybrid tracks. Of the hybrid tracks, those with an SPD hit and an ITS refit have the best quality (i.e. best resolution in p_T). However, they have nonuniform acceptance in η and φ . This is why we also include the other two types of hybrid tracks. The determination of their p_T is less good due to the lack of an SPD hit, but including them does give us a uniform acceptance in a region of the (η, φ) -plane. The condition that $p_T \geq 0.15 \text{ GeV}$ is required because the p_T resolution of the detector is not as good in that energy region, and the $p_T \leq 5 \text{ GeV}$ condition is made because tracks with higher transverse momentum are likely to have originated from hard scatterings, which we don't want to affect our flow measurement. The final requirement of a track's pseudorapidity is made because the ITS and TPC have a

⁴Note that these requirements are only for flow measurements. In the jet analysis different conditions are used.

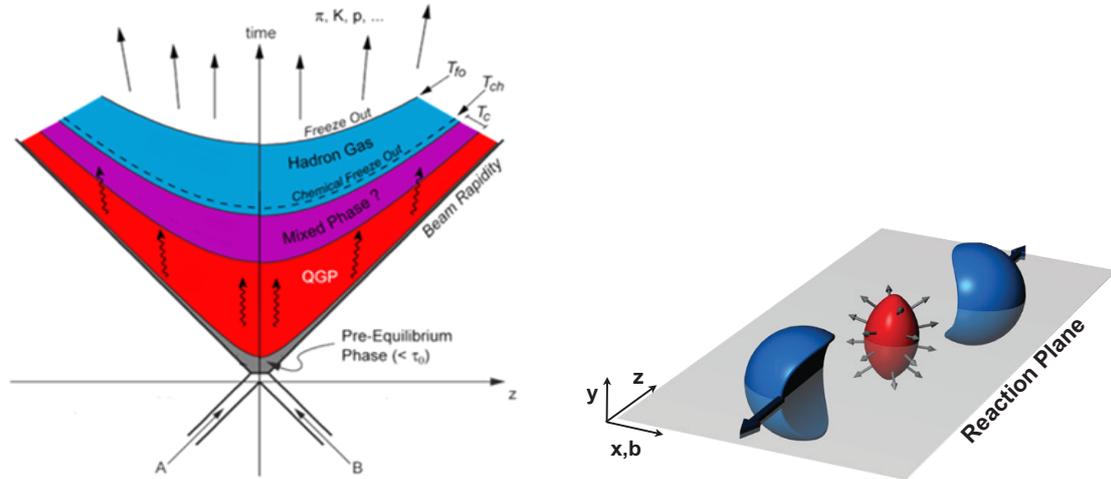


Figure 1: Left: The different phases the collision products are expected to go through. The QGP forms when thermal equilibrium is reached. As it cools, freeze-out occurs, with the possibility of a mixed phase preceding this. After freeze-out, all strongly interacting matter will have coalesced into colour-neutral composite particles. Finally, after some of the particles undergo radioactive decay, the final state is formed, which is detected in a detector. Right: Illustration of the geometry of a collision.

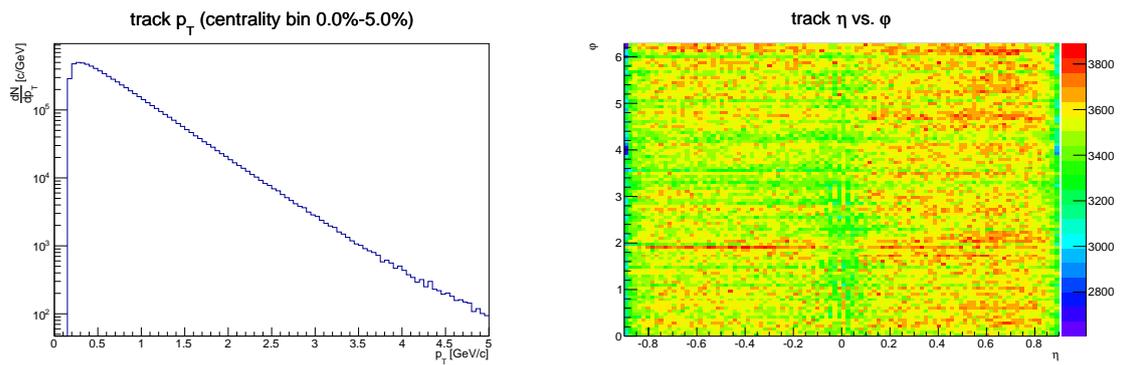


Figure 3: Left: Transverse momentum distribution for the 5% most central collisions. Right: (η, ϕ) -distribution for all tracks which pass track selection.

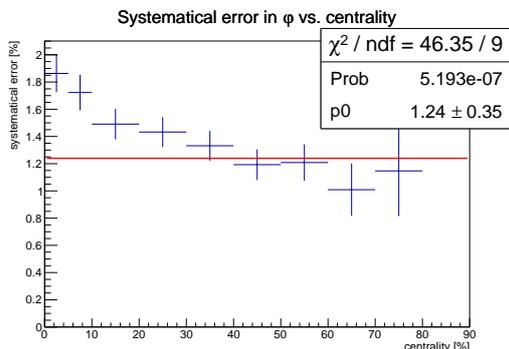


Figure 4: Fractional systematical error in the φ -distribution of tracks over a large number of events vs. centrality. The systematical effects appear larger for more central collisions.

finite size. This means that they are less efficient detecting tracks with $|\eta| > 0.9$.

Figure 3 is used to assess the quality of our tracks after the track selection. The left panel of Figure 3 shows the transverse momentum distribution of the most central collisions. We see no tracks with $p_T < 0.15$ GeV or $p_T > 5$ GeV, and the shape of the distribution appears consistent with earlier observations at ALICE. The right half shows the (η, φ) -distribution of the tracks which pass track selection. There are statistically significant deviations from a uniform distribution. The nonuniformity appears to be mostly in φ . To quantify this nonuniformity, the variance in the number of tracks $\bar{\sigma}^2$ for each of our centrality classes is measured.⁵ Subsequently, we calculate the expected variance in the number of tracks $\bar{\sigma}_{\text{ex}}^2$ assuming Poisson statistics. We then estimate the fractional systematical error as $\bar{\sigma}_{\text{sys}} = \sqrt{\bar{\sigma}^2 - \bar{\sigma}_{\text{ex}}^2}$.⁶ The result of this calculation can be seen in Figure 4. It shows that with the present amount of data ($3.8 \cdot 10^7$ events), the magnitude of the fractional systematical errors appears to be dependent on centrality. This is consistent with a nonuniformity in tracking efficiency which depends on track multiplicity.

⁵Throughout this article, the centrality bins are bounded by $\{0, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90\}$ unless indicated otherwise.

⁶Unless indicated otherwise, errors are propagated using the covariance matrix of the parameters, and the Jacobian of the function.

2.2 Event planes

The first step in the flow analysis is to calculate the event planes Ψ_2 and Ψ_3 . These can be seen as the second and third order symmetry planes of the event. First, the so-called Q -vector is calculated:[5]

$$Q_n = \sum_i w_i e^{in\varphi_i}, \quad (2)$$

where n is either 2 or 3, and w_i denote track weights. Next, the event planes are calculated with the atan2 function:

$$\Psi_n = \frac{1}{n} \text{atan2}(\text{Im } Q_n, \text{Re } Q_n).$$

For the ITS/TPC, the sum in (2) is taken over all tracks in the event that meet the requirements mentioned in the previous section. Weights are all set equal to 1. The result of this calculation can be seen in the upper left corner of figure 5. It is apparent that the distribution is uniform in both Ψ_2 and Ψ_3 . This is expected, since the event planes are dependent on the direction of the impact parameter, which is random, and uniformly distributed. This translates to uniform distributions for the event planes. We also see that Ψ_2 and Ψ_3 appear not to be correlated.

For the V0 detectors, this procedure gets changed slightly. The V0 detectors don't detect individual tracks. Instead, they each consist of 32 elements, which measure 'equalized multiplicity'. This quantity is proportional to the number of tracks passing through that particular detector element. Therefore, the sum in (2) is taken over all detector elements, and the weights are set equal to the equalized multiplicity. The result of this can be seen in the upper right corner of figure 5. The distribution appears far from flat, due to nonuniform efficiency of the V0 detector elements. This effect can be corrected with the following substitution:[6]

$$Q_n \rightarrow \frac{Q_n - \langle Q_n \rangle}{\sigma_{Q_n}},$$

where $\langle Q_n \rangle$ is the average of Q_n over a number of events with the same nonuniform efficiency, and σ_{Q_n} is the standard deviation of Q_n for a large number of events. As can be seen in the lower right corner of Figure 5, applying this 'recentering procedure' results in a flat distribution, which indicates an unbiased estimate.

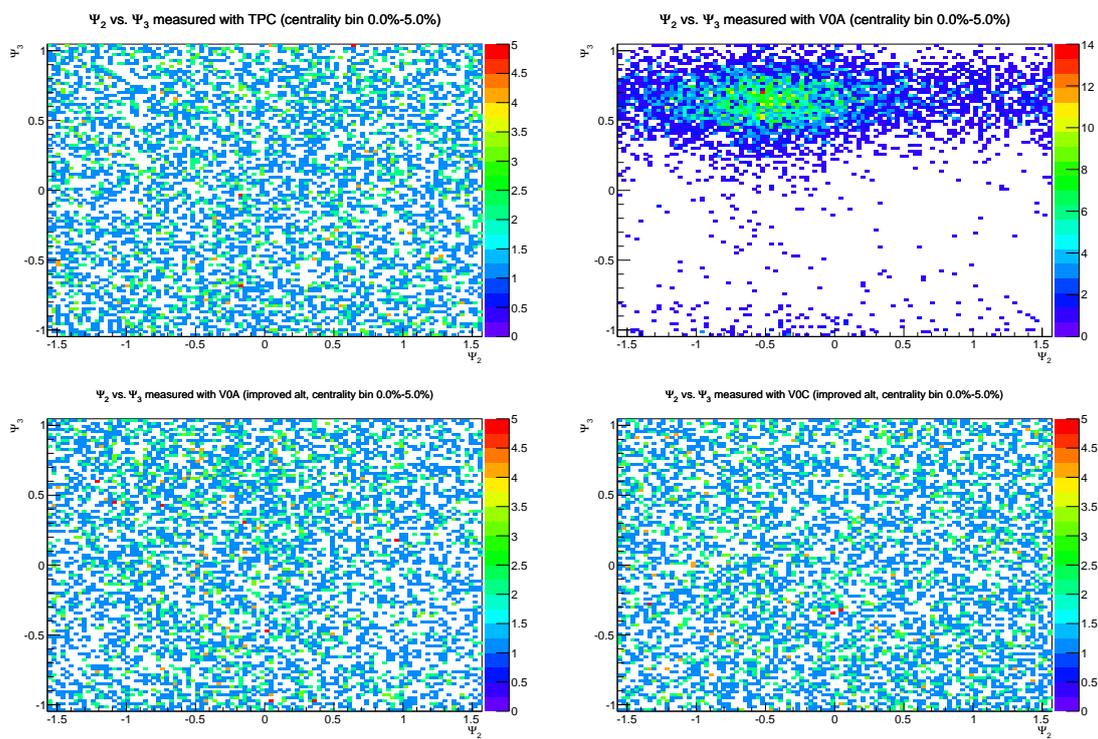


Figure 5: Upper left: Ψ_2 vs. Ψ_3 measured with the ITS/TPC. Upper right: The same measured with the VOA. Lower left: Same as upper right, but with VOA channels calibrated for 2010 data. Lower right: Same as lower left, but for V0C.

2.3 Obtaining v_2 and v_3

Now that the event planes have been obtained, we can determine v_2 and v_3 . This is first done using the event planes as measured by the ITS/TPC. A histogram of the azimuthal angle φ of the tracks in an event which pass track selection is first made. The number of bins is chosen close to the square root of the total number of tracks, to maximize the quality of the fits that will be obtained. This histogram can be made in two ways:

- Weigh each track with 1: This way we can obtain v_2 and v_3 for $dN/d\varphi$. In this case the number of tracks in each bin is assumed to have a Poisson distribution.
- Weigh each track with its transverse momentum: This enables us to determine v_2 and v_3 for $dp_T/d\varphi$. In this case there are two sources of statistical error on the total transverse momentum in a bin: variation in the average transverse momentum $\langle p_T \rangle$ (averaged over all tracks in a particular bin), and variation of the number of tracks in a bin (denoted by N). The error on $\langle p_T \rangle$ is calculated as the sample standard deviation of all tracks in an event divided by the square root of the number of tracks in a particular bin. The error on N is assumed to be Poisson. The total p_T in a bin is then equal to $N\langle p_T \rangle$, and thus the statistical error on the total p_T in a bin is equal to the quadratic sum of the fractional uncertainties of $\langle p_T \rangle$ and of N .⁷

Next, (1) is fitted to the histogram. An example of such a fit is shown in the upper left corner of Figure 6. In the upper right corner, we can see the results of a number of fits on events with centrality between 20% and 30%. Note that there does not appear to be any correlation between v_2 and v_3 . The middle panels show mean v_2 and v_3 for $dN/d\varphi$ fits versus centrality, and the bottom panels show the same for $dp_T/d\varphi$ fits.

In addition to fitting on individual events, the averages of v_2 and v_3 can also be determined for the tracks in all events within a given centrality range

⁷It is assumed that N and $\langle p_T \rangle$ are uncorrelated, which is not strictly speaking true, but the correlation is small, and it can not be accurately measured on an event-by-event basis.

simultaneously. To do this, we need to take into account the effect that the event planes vary randomly between events (Figure 5). This can be done by making a histogram of $\varphi - \Psi_2$ of all tracks in all events. Since Ψ_2 and Ψ_3 are uncorrelated (Figure 5), the contribution from the third order Fourier coefficient in (1) will vanish, and we are left with the zeroth and second order terms. The same can be done for $\varphi - \Psi_3$. The result of this procedure for events with a centrality between 20% and 30% is shown in Figure 6, along with the other methods discussed in this section.

In the lower middle two panels of Figure 6, one can see that v_2 and v_3 increase with centrality, which is expected, since they are dependent on the anisotropy of the system, which increases with centrality. Also, v_2 measured with $dp_T/d\varphi$ is larger than v_2 measured with $dN/d\varphi$. This indicates that there are not just *more* particles emitted in-plane than out-of-plane, but also that these particles carry more momentum on average. Finally, for central events the v_n fitted per event do not differ significantly from those fitted to the sum of all events in a given centrality bin.

In the bottom two panels of Figure 6, the spreads of the v_n distributions are shown. These are calculated using the measured root mean squares of the measured v_n distributions (denoted by $\sigma(v_{n,m})$), and comparing them to the mean (over all events in a centrality bin) of the errors on v_n estimated in the fit procedure (denoted by $\sigma(v_{n,f})$). The spreads of the v_n distributions are then estimated as the root mean squares of the *true* v_n distributions, which are equal to $\sqrt{\sigma(v_{n,m})^2 - \sigma(v_{n,f})^2}$.⁸

For this procedure to work, the error estimation on v_n has to be accurate. To show that this is indeed the case, the mean of χ_{red}^2 is examined in Figure 7. The means are slightly smaller than one, indicating that the estimated errors are on average overestimated by a factor of $\sqrt{\chi_{\text{red}}^2}$, or approximately $\sqrt{0.9} \approx 0.95$. Since the overestimation is proportional to the inverse square root of the mean of χ_{red}^2 , the overestimation is small, and it is corrected by replacing $\sigma(v_{n,f}) \rightarrow \sigma(v_{n,f})\sqrt{\chi_{\text{red}}^2}$, where χ_{red}^2 is obtained from Figure 7.

Returning to the bottom two panels of Figure 6, it can be seen that for both $dN/d\varphi$ and $dp_T/d\varphi$,

⁸In the case that the argument of this square root is negative because of statistical fluctuations, no result is shown.

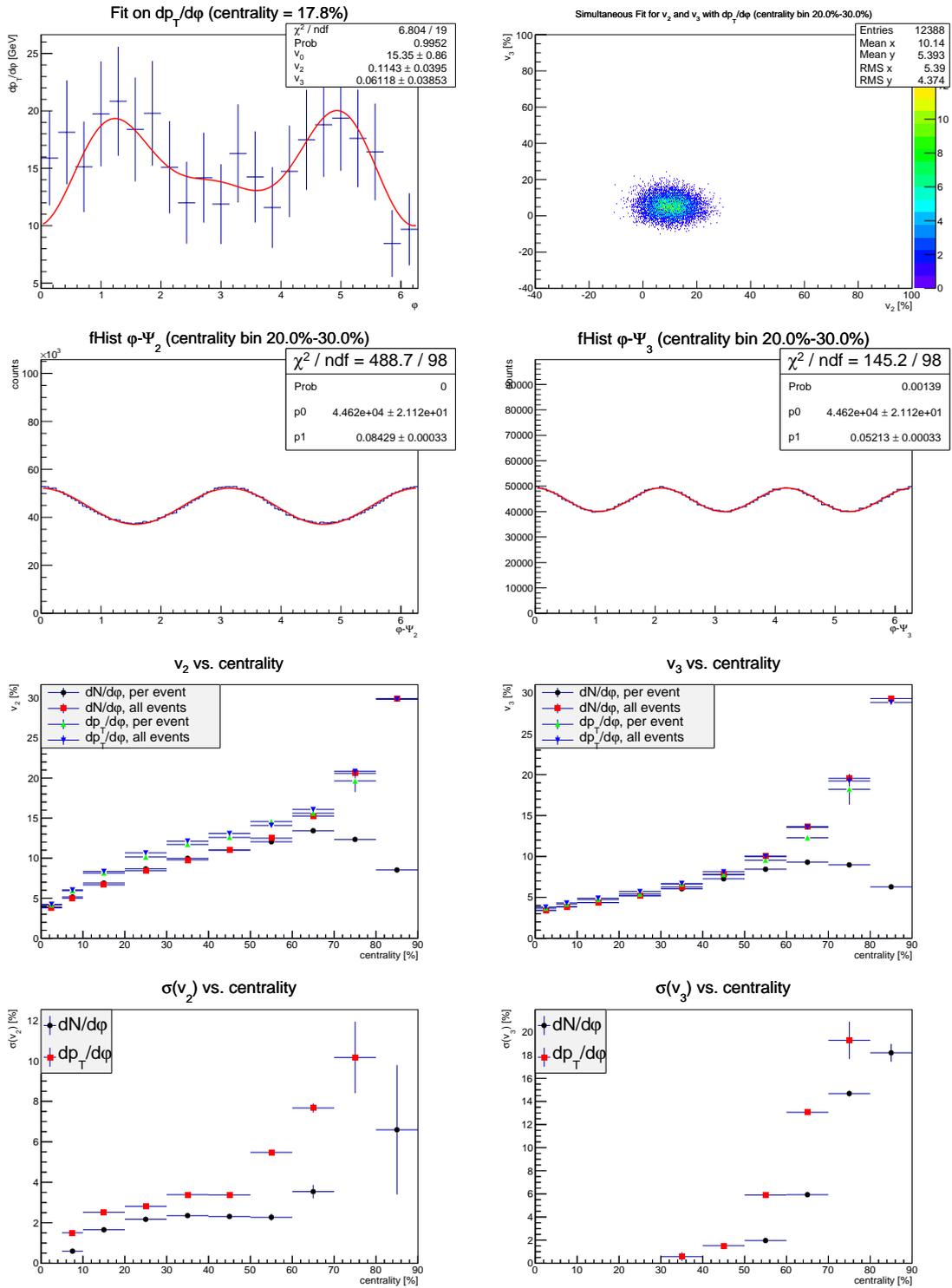


Figure 6: Upper left: An example of a fit to $dp_T/d\phi$. Upper right: Values of v_2 and v_3 from $dp_T/d\phi$ fits with centrality between 20% and 30%. Upper middle: $\phi - \Psi_2$ and $\phi - \Psi_3$ histograms without particle weights for events with centrality between 20% and 30%. Lower middle: Mean v_2 and v_3 for the different methods discussed in section 2.3. Bottom: Spread of v_2 and v_3 for both $dN/d\phi$ and $dp_T/d\phi$.

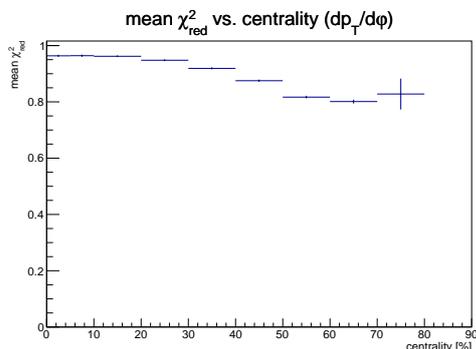


Figure 7: Mean χ_{red}^2 versus centrality.

the root mean squares of the true v_n distributions are small. This may indicate that the spread in v_n is small enough in any given centrality bin to approximate the v_n for each event by the mean v_n for its centrality bin.

2.4 Removing autocorrelations

If the v_n are determined using the same tracks which are already used to determine the Ψ_n , the v_n will be biased. This bias is called autocorrelation. The Ψ_n lie in the plane with the largest track density, so in events with a small number of tracks, statistical fluctuations determine the Ψ_n to a large degree. As a result of this, for these events, the v_n are overestimated, since the Ψ_n lie in the plane of the largest measured track density, and not in the plane of the largest track density of the underlying distribution. The way to solve this problem is to take away the cause: we need to use two different sets of data to determine respectively the symmetry planes and the flow coefficients.

One way to accomplish this is to split the ITS/TPC in two halves: a ‘left side’ ($\eta < 0$) and a ‘right side’ ($\eta > 0$). The v_n are fitted to one half, while using the symmetry planes obtained from the other half as fixed parameters. The two sets of flow coefficients are subsequently averaged to make optimal use of all the available data.

Another way to remove the autocorrelations is to determine the event planes using the V0 detectors instead of the ITS/TPC. The V0s are statistically independent from the tracks used in fits. They are obtained as described in section 2.2, and subsequently a fit is made. The results for this method

and the method described above are shown in the upper two panels of Figure 8.

There are clear differences with the results in Figure 6. The estimates of v_n are lower for all centralities, but the most visible change can be seen in peripheral collisions. This is where autocorrelations most impact figure 6. The different methods with the autocorrelations removed can not yet be compared to each other, since they are still biased. This bias will be discussed in section 2.6.

The last method used in this study is the method of cumulants. This involves calculating on an event-by-event basis the following quantities:⁹[1]

$$\begin{aligned}
 Q_{n,k} &= \sum_i^M (p_T)_i^k e^{in\varphi_i}, \\
 S_{p,k} &= \left[\sum_i^M (p_T)_i^k \right]^p, \\
 \mathcal{M}_{11} &= S_{2,1} - S_{1,2}, \\
 \langle 2 \rangle_n &= \frac{|Q_{n,1}|^2 - S_{1,2}}{\mathcal{M}_{11}}, \\
 \mathcal{M}_{1111} &= S_{4,1} - 6 \cdot S_{1,2}S_{2,1} + 8 \cdot S_{1,3}S_{1,1} \\
 &\quad + 3 \cdot S_{2,2} - 6 \cdot S_{1,4}, \\
 \langle 4 \rangle_n &= [|Q_{n,1}|^4 + |Q_{2n,2}|^2 \\
 &\quad - 2 \cdot \text{Re}(Q_{2n,2}Q_{n,1}^*Q_{n,1}^*) \\
 &\quad + 8 \cdot \text{Re}(Q_{n,3}Q_{n,1}^*) - 4 \cdot S_{1,2}|Q_{n,1}|^2 \\
 &\quad - 6 \cdot S_{1,4} + 2 \cdot S_{2,2}][\mathcal{M}_{1111}]^{-1}.
 \end{aligned}$$

These formulae implicitly do not have autocorrelations.[1] Next, the $\langle 2 \rangle_n$ are averaged with weights given by \mathcal{M}_{11} , over a large number of events. The result of this averaging is denoted by $\langle \langle 2 \rangle \rangle_n$. Similarly, the $\langle 4 \rangle_n$ are averaged with weights \mathcal{M}_{1111} to yield $\langle \langle 4 \rangle \rangle_n$. The flow coefficients can then be estimated using both these averages. The estimate from 2-particle cumulants is given by

$$v_n\{2\} = \sqrt{\langle \langle 2 \rangle \rangle_n},$$

and the estimate from 4-particle cumulants is given by

$$v_n\{4\} = \sqrt[4]{2 \cdot \langle \langle 2 \rangle \rangle_n^2 - \langle \langle 4 \rangle \rangle_n}.$$

⁹ M denotes the number of tracks in an event. Tracks used are the same as in the other methods described in this section.

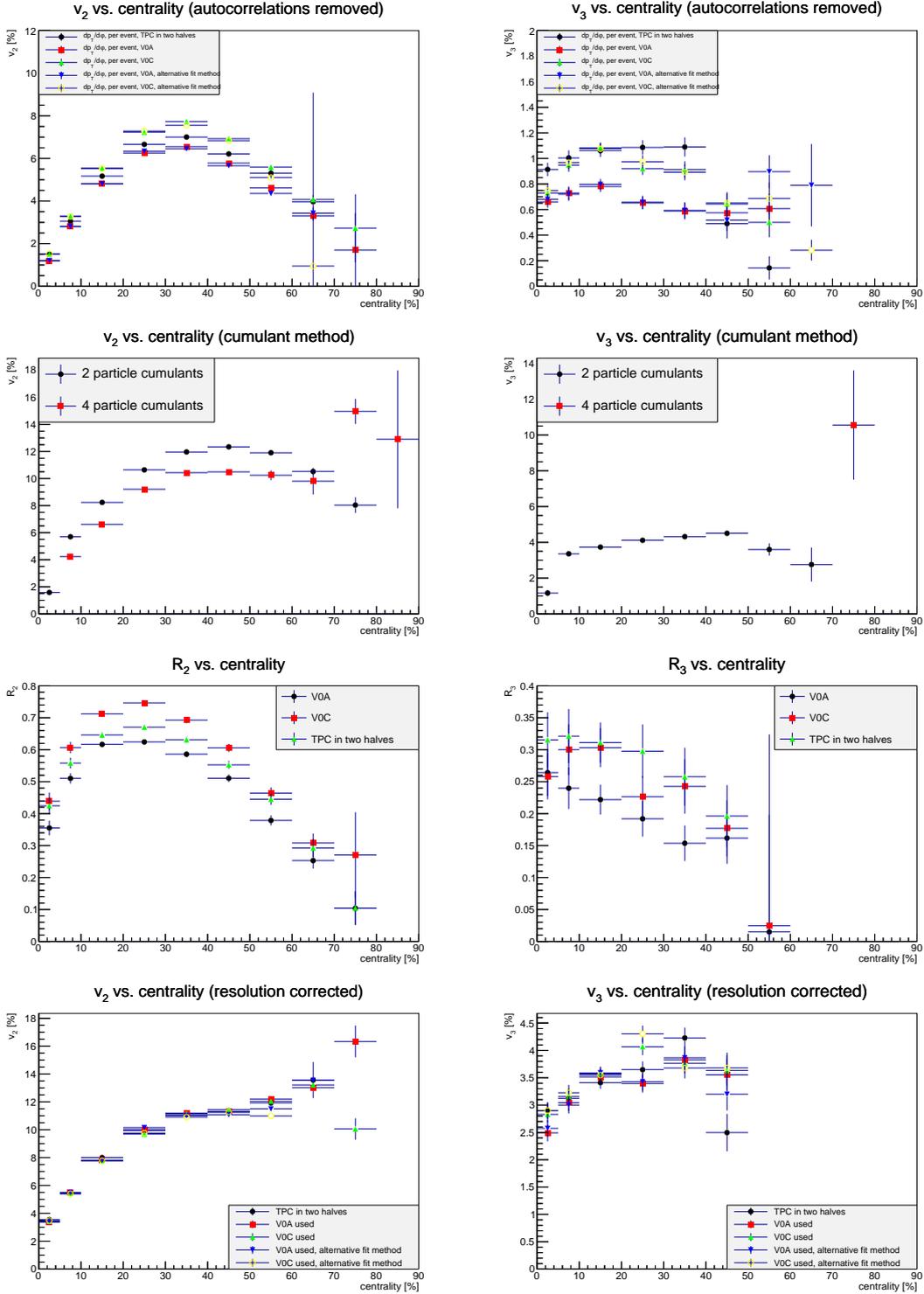


Figure 8: Top: Comparison of results with autocorrelations removed. For v_3 , the red and green points fall outside the plotted range for the most peripheral collisions. Upper middle: Comparison of results with autocorrelations removed by using cumulants. Lower middle: Resolution correction for three methods of determining the event planes. Bottom: Comparison of results for resolution corrected v_2 and v_3 with autocorrelations removed.

Flow coefficients measured with this method are shown in the upper middle panels of Figure 8.¹⁰ The estimates from the 2-particle cumulants are larger than those from the 4-particle cumulants. This is mostly due to nonflow effects like decays, but it is also known that both of these estimates have systematic biases. Figure 9 shows v_n as estimated in literature. Knowing that the true v_n lies

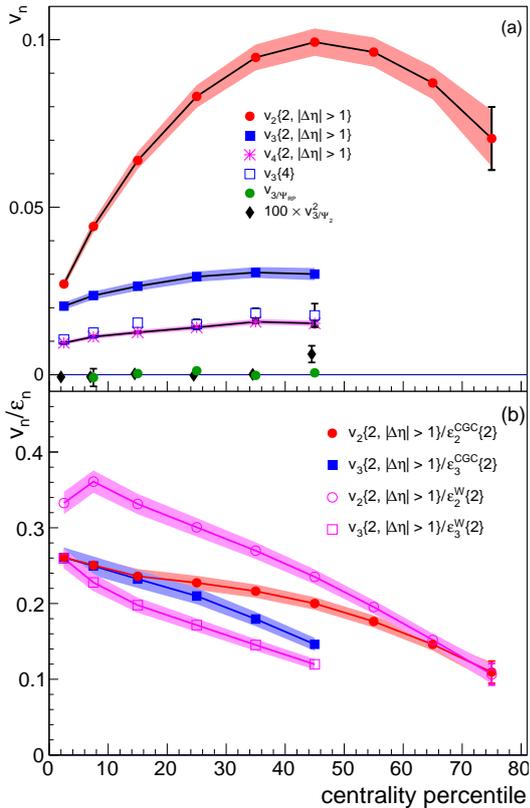


Figure 9:

somewhere between $v_n\{2\}$ and $v_n\{4\}$, the cumulant estimates agree well with literature.

2.5 Assessing quality of v_2 and v_3

To assess the quality of the fits, the p -value distribution of each fit is investigated. By the definition of the p -value, its distribution over all events should

¹⁰Absent data points indicate that the argument of one of the roots in the definitions of $v_n\{2\}$ and $v_n\{4\}$ is negative for that point.

be uniform.¹¹ This distribution is shown in Figure 10. For small p -values, the distribution is clearly

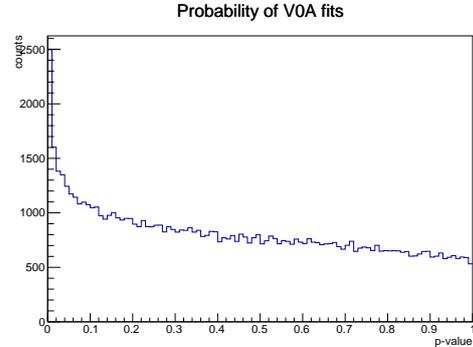


Figure 10: p -value of fits on $dp_T/d\varphi$. Autocorrelations have been removed by using the VOA detector for the determination of the event planes. Apart from the peak at small p , the distribution is uniform.

not uniform. This could be due to several things:

- An incomplete picture of the experimental error of the total p_T in a certain φ -bin: It is not uncommon for particles to undergo radioactive decay after the QGP freeze-out. Due to relativistic beaming, daughter particles will tend to have highly correlated flight directions. This makes them likely to end up in the same φ -bin, and causes the error on the number of tracks in a bin to deviate from its assumed Poissonian nature. Unfortunately, it is hard to measure how large this effect is.
- Equation (1) might not accurately describe the φ -distribution of the underlying event: It is known that (1) describes flow to a high degree of accuracy. However, events also contain jets, which are due to hard parton interactions. If the jets lose transverse momentum while travelling through the QGP, momentum conservation dictates that the soft background

¹¹Under the assumptions that the function fitted to the data correctly models the data, and that the errors of the data points are accurate, the p -value corresponds to the probability of a fit being 'more deviant' than the one observed. This means that when making a histogram of the distribution of such p -values, the fraction of entries smaller than a given value should equal that value itself. The only way in which this is possible is if the distribution is uniform.

to these jets should gain an equal amount of transverse momentum. This contribution does not follow equation (1), and the momentum exchange makes separating the two contributions very hard. Correcting for this effect is thus not easy.

Both effects described above can cause the nonuniformity seen in figure 10. At the time of writing, no method was known to estimate the magnitudes of these effects, which means that in this study there will not be an attempt to correct for them. What can however be determined is the effect they have on the determination of the flow coefficients. This is done by means of simulations.

Each track in a simulated event is given an azimuthal angle which has been randomly drawn in accordance to (1) with known flow parameters and event planes.¹² Next, the fit routine is repeated with this new event, which has an idealized distribution. Then Figure 10 can be replicated to show how the fit routine behaves in an ideal situation. This is shown in the upper left corner of Figure 11. There is still a small peak, so apparently, this fit routine is not perfect. To assess whether this impacts the determination of flow coefficients, the bias in v_n is studied. An event is simulated with known $v_{n,\text{in}}$, and the fitted $v_{n,\text{out}}$ are compared to $v_{n,\text{in}}$ by computing the mean of $v_{n,\text{out}} - v_{n,\text{in}}$. This bias as a function of $v_{n,\text{in}}$ is given in the middle panels of figure 11, for centralities between 20% and 30%. Next, this bias as a function of $v_{n,\text{in}}$ is modelled as a linear function with the constraint that the bias for $v_{n,\text{in}} = 0$ equals zero. The slopes of these linear functions are shown in the bottom panels of figure 11. It is clearly visible that there is a significant bias which has a magnitude of a few percent.

As a next step, nonflow is added in the form of the two effects described above. Each track is placed in its bin *twice* with a probability of 50%. This simulates possible systematic effects from decay. Also, 5 high p_T tracks (4 GeV each) are added to each event, randomly distributed in φ . As can be seen in the lower upper right panel of figure 11, the nonflow causes the p -values to become very much nonuniform. Looking at the bias in the determination in v_n in the bottom panels of 11, it is visi-

¹²To make sure the parameters are realistic, the parameters used are equal to their best estimates measured from the event.

ble that the bias has increased by a small amount. This would indicate that this amount and type of nonflow impacts the analysis, but not by a large amount, and since the p -values with this amount of nonflow are less uniformly distributed than the actual nonsimulated events, this gives some confidence that the actual events are not impacted by much either.

In an attempt to improve upon this result, a slightly more sophisticated fit routine is used. The number of tracks per bin is rather small (< 20). This causes the discrete nature of the Poisson distribution to come into play. Describing errors with standard deviations is not particularly suited to discrete distributions, and can impact the outcome of a fit, especially if different bins hold a different number of tracks. Therefore, bin boundaries are chosen in a way such that the expected number of tracks in each bin (according to the fitted function) is the same. In detail, this means that the right boundary a of a bin i satisfies the following relation.

$$\int_0^a \frac{dp_T}{d\varphi} d\varphi = \frac{i}{N_{\text{bins}}} \int_0^{2\pi} \frac{dp_T}{d\varphi} d\varphi,$$

with N_{bins} the number of bins. Fitting (1) to this distribution would yield $v_2 \approx v_3 \approx 0$ since the distribution is by construction approximately flat. Therefore the function used for fitting has to be modified too. The solution is to use a piecewise defined function, equal within each bin to the integral of (1) over the bin interval. To evaluate the possible improvements of this fit method over the simple fit method, all simulations, with and without nonflow, are repeated and the results shown in the bottom panels of Figure 11.¹³ There *is* a slight reduction in the bias, but the bias is also more sensitive to nonflow compared to the the simple fit method. The alternative method has also been applied to non-simulated data, the results of which are shown in Figure 6.

2.6 Resolution correction

Even after removing autocorrelations, the method is not completely unbiased. This is because deviations from the true event plane tend to lower the

¹³The p -value distributions are so similar to the ones for the simple fit method that they are omitted for clarity.

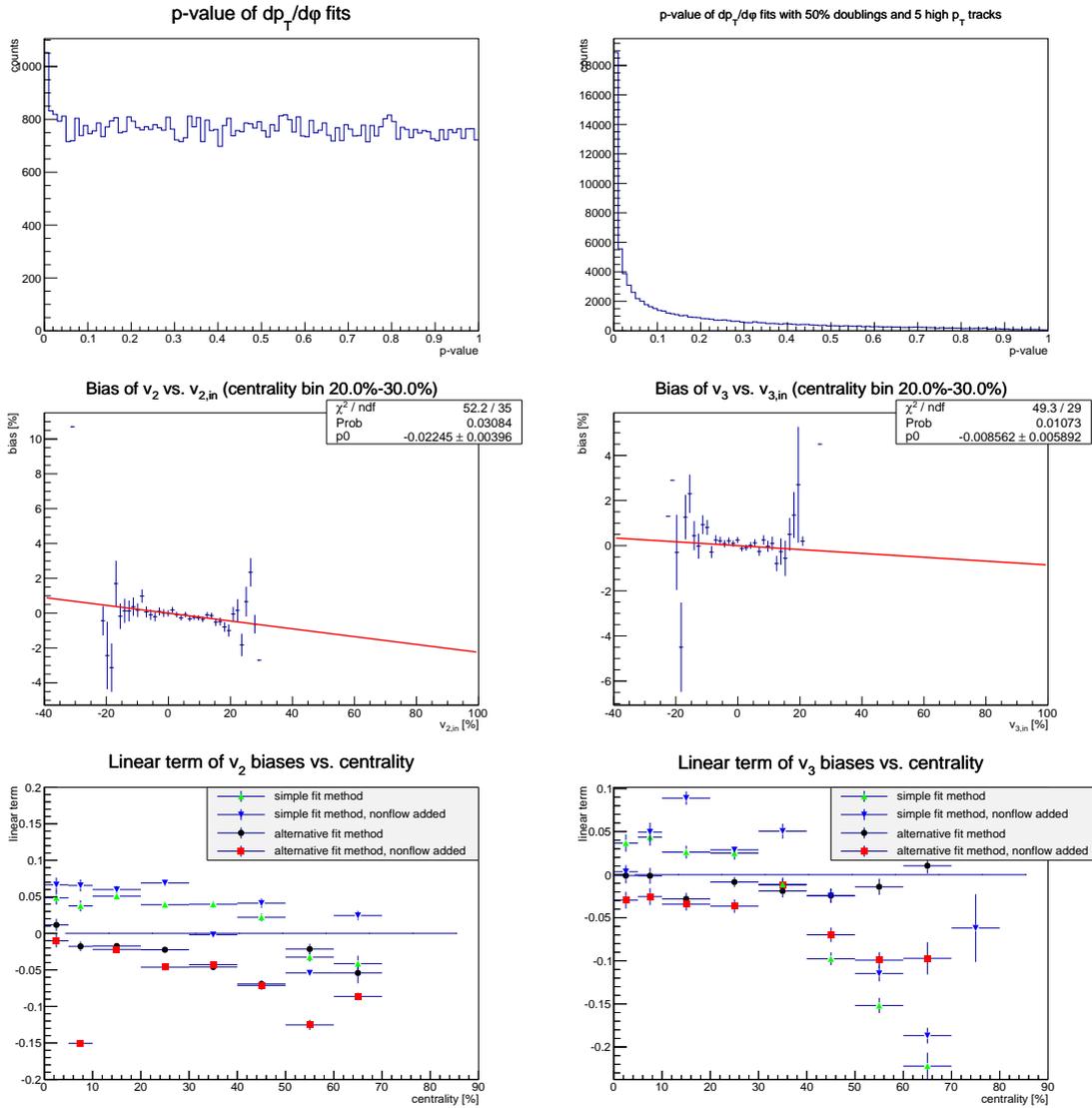


Figure 11: Top left: p -value distribution of simulated fits made with the simple fit method. Top right: p -value distribution of simulated fits with 50% doubly counted tracks and 5 high p_T tracks. Middle: Bias of v_2 and v_3 versus the known v_2 and v_3 values used to make the simulated event. Bottom: Slope of v_n bias as a function of known v_n versus centrality.

measured value for v_2 and v_3 . The underestimation can be corrected by dividing by the resolution correction. In our case, we estimate the event planes in three different ways, namely with the V0A, V0C, and with the TPC. Denoting the different event planes by $\Psi_{n,V0A}$, $\Psi_{n,V0C}$ and $\Psi_{n,TPC}$, and denoting the resolution corrections for each of these by $R_{n,V0A}$, $R_{n,V0C}$ and $R_{n,TPC}$, respectively, the resolution correction can be estimated by the following formulas:[5]

$$\begin{aligned} C_{n,V0A,V0C} &= \langle \cos(n(\Psi_{n,V0A} - \Psi_{n,V0C})) \rangle, \\ C_{n,V0A,TPC} &= \langle \cos(n(\Psi_{n,V0A} - \Psi_{n,TPC})) \rangle, \\ C_{n,V0C,TPC} &= \langle \cos(n(\Psi_{n,V0C} - \Psi_{n,TPC})) \rangle, \\ R_{n,V0A} &= \sqrt{\frac{C_{n,V0A,V0C} \cdot C_{n,V0A,TPC}}{C_{n,V0C,TPC}}}, \\ R_{n,V0C} &= \sqrt{\frac{C_{n,V0A,V0C} \cdot C_{n,V0C,TPC}}{C_{n,V0A,TPC}}}, \\ R_{n,TPC} &= \sqrt{\frac{C_{n,V0C,TPC} \cdot C_{n,V0A,TPC}}{C_{n,V0A,V0C}}}. \end{aligned}$$

The averages denote averages over a large number of events.

When calculating the resolution correction, care must be taken in the calculation of the errors. The three C_n are highly correlated, and this must be taken into account when determining the corrections. The results of this analysis are shown in the lower middle panels of Figure 8. In the lower panels of the same figure, the resolution corrected v_n are shown, which are obtained from the uncorrected v_n by dividing by the resolution correction for the method that is used to obtain the event planes. It can be seen that there are significant differences among the corrected v_n , which would indicate remaining biases in the analysis. Also, the resolution corrected v_n can be compared to literature. Comparing to 9, it can be seen that the resolution corrected v_n are significantly larger than literature. This supports the idea that there are biases remaining in the analysis. Also, theoretically, the resolution correction only accurately describes the underestimation of the true v_n if the cosines in the determination of C_n are approximately equal to one, which is not the case for peripheral collisions.

3 Jets

As was briefly mentioned before, jets are a spray of particles originating from hard parton interactions. They interact with the QGP as they travel through it. This interaction causes the jets to lose momentum and change in shape.

Jets are identified with the Jet Finder algorithm. It groups tracks together based on their direction, and these groups are then identified as jets. In this way, all tracks in an event are (as far as the analysis is concerned) part of a jet, but not all of these jets originated from hard scatterings. This section discusses the process in which the jets are identified, how the soft background to the hard interactions is modelled, and the resulting spectra for the hard interactions.

The actual recovery of the hard scattering p_T spectra, which will not be done in this study, is achieved by the so-called unfolding algorithm. As was mentioned before, the measured jet spectrum has two major contributions: the one due to hard scatterings, and the one due to flow. The width of the flow contribution causes the measured spectrum to have a shape which is different from the hard scattering spectrum which we intend to study. The unfolding algorithm attempts to reverse this, and recover the true spectrum. For this, it requires the width of the soft background spectrum as a parameter.

This section will discuss various ways to model the soft background spectrum, describe methods to assess the quality of such models, and finally present the jet spectrum to which the unfolding algorithm can be applied in future studies.

3.1 Jet Finder

The Jet Finder is an algorithm to identify jets. In this study, two main varieties are used: the k_t , and the anti- k_t algorithms. First, the k_t algorithm is discussed. The algorithm iterates the following steps:[2]

1. For each pair of tracks (i, j) , compute the k_t distance $d_{ij} = \min(k_{ti}^2, k_{tj}^2) \cdot ((\eta_i - \eta_j)^2 + (\varphi_i - \varphi_j)^2) / R^2$, with k_t the transverse momentum of the track, and R is a parameter known as the jet radius.¹⁴ In this

¹⁴For the distance in φ between the tracks, the smallest

analysis, a jet radius of 0.4 is used. Also, compute the beam distance $d_{iB} = k_{ti}^2$.

2. Find the minimum of $\{d_{ij}\} \cap \{d_{iB}\}$. If this minimum is an element of $\{d_{ij}\}$, the two particles are merged into one, provided that $(\eta_i - \eta_j)^2 + (\varphi_i - \varphi_j)^2 \geq 1$. The four-momentum of the resulting ‘particle’ is defined as the sum of its constituents. If the minimum distance is an element of $\{d_{iB}\}$, the corresponding track is defined to be a final jet, and is removed from the list.

The iteration is stopped when there are no tracks left which fulfil the criteria to be merged, and the remaining tracks are defined to be final jets. For the anti- k_t algorithm, we replace d_{ij} by $d_{ij} = \min(k_{ti}^{-2}, k_{tj}^{-2}) \cdot ((\eta_i - \eta_j)^2 + (\varphi_i - \varphi_j)^2) / R^2$, and d_{iB} by $d_{iB} = k_{ti}^{-2}$. This means that the k_t algorithm starts clustering with the low- p_T tracks, and the anti- k_t starts with the high- p_T tracks.

The transverse momentum, pseudorapidity and azimuthal angle of jets are determined by adding the 4-momenta of a jet’s constituent tracks. Another property, known as area, is calculated by adding so-called ghost particles to the event. These are randomly distributed in $\eta \in (-0.9, 0.9)$ and $\varphi \in [0, 2\pi)$, with a uniform distribution. Their momenta are as small as possible, i.e. 10^{-100} GeV. When the Jet Finder groups tracks into jets, these ghost particles do not influence the p_T , η and φ of the final jets, and the number of ghost particles associated to a particular jet is taken as a measure proportional to its area. The last jet property used in this analysis is ρ , energy density, which is defined as a jet’s transverse momentum divided by its area.

The k_t and the anti- k_t algorithm have different properties which make them suited for different applications. The k_t algorithm produces jets with areas approximately equal to πR^2 , while the shapes of these jets can be irregular. The anti- k_t algorithm produces near circular jets, but these jets vary more strongly in size. The k_t jets are used to model the soft background, while the anti- k_t jets are used as signal jets.

distance along the circle is taken.

3.2 Jets and tracks used

The track selections used in the jet analysis are similar to the ones used for the flow analysis, with the change that the requirement that a track’s transverse momentum be smaller than 5 GeV is dropped. This requirement was made because such tracks are likely to have originated from hard scatterings. For flow measurements, this means these tracks should be excluded, but for jets, these are the tracks of interest, so they are *not* excluded here.

Figure 12 shows that the (η, φ) -distribution of both k_t and anti- k_t jets is not uniform. Near $|\eta| = 0.9$, the Jet Finder groups tracks into a large number of jets. As can be seen in figure 12, these jets near have a significantly different ρ from the other jets. For this reason, the requirement is made that $|\eta| < 0.5$ for both k_t and anti- k_t jets.

For anti- k_t jets, an additional area cut is used. Anti- k_t jets are used to study hard parton interactions, and these interactions are expected not to yield jets with small areas. Different values can be taken for the threshold for this area cut. More on this in section 3.6.

3.3 Background estimation

The next step in the analysis is to model the soft background of each event. This is done with k_t jets, since they rarely have a very small area. Jets with a small area have few constituent particles, and therefore their ρ varies more than jets with a larger area. In the following analysis, the variation of ρ greatly impacts the estimation of the background. This makes k_t jets suited for estimating the soft background. As a first approximation, this background $\rho(\varphi)$ is assumed to be a constant which is equal to the median of ρ taken over all k_t jets in an event. The two jets with the highest transverse momentum are ignored for this median, since these are probably due to hard parton interactions.

3.4 Random cones

Random cones are used to assess the quality of a particular method of background estimation. A random cone is a circular area in the (η, φ) -plane, with a radius equal to R , and with a randomly chosen η and φ . This area is denoted by A_{RC} . Per event 20 random cones are created, and for each

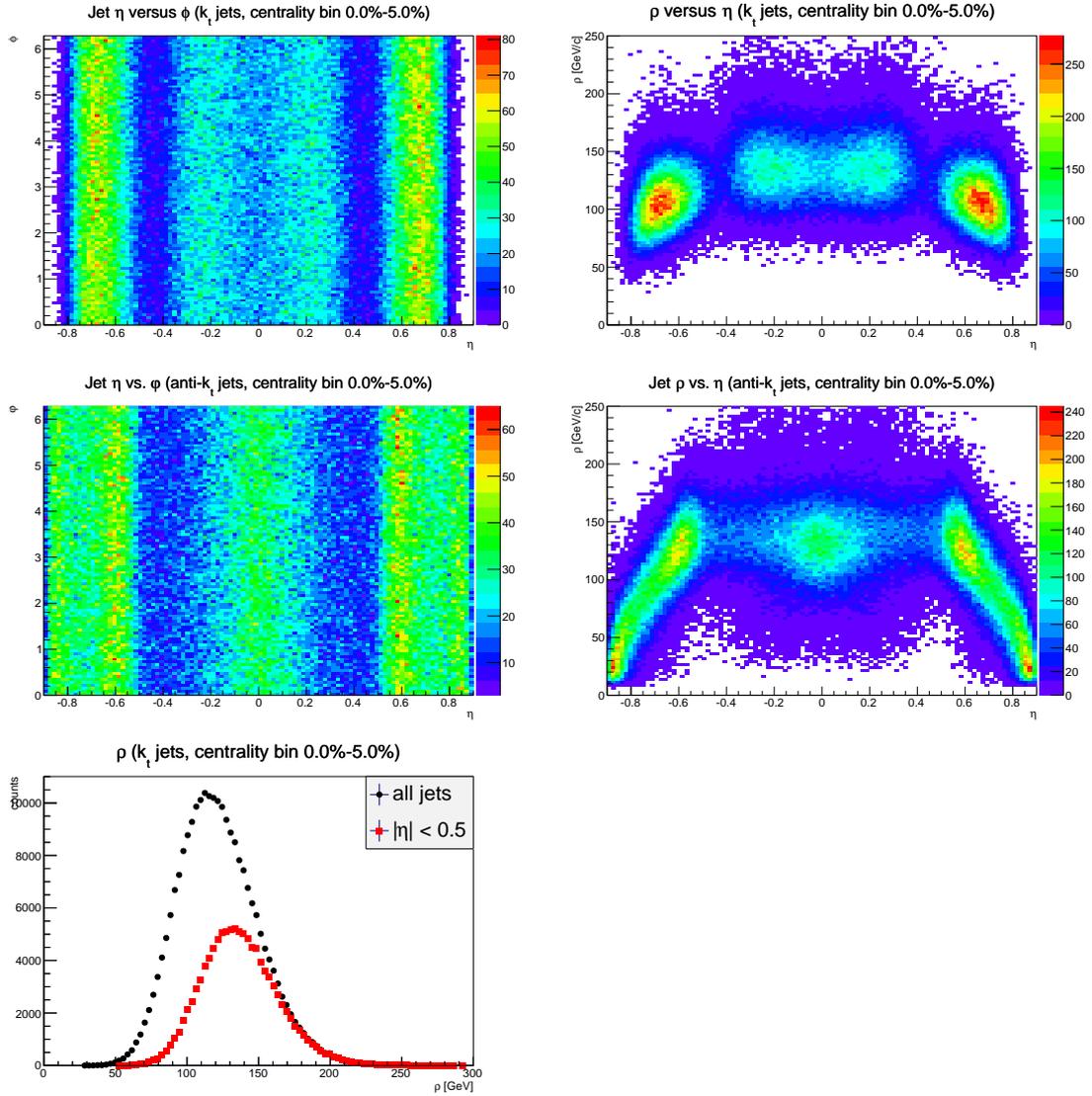


Figure 12: Upper left: Distribution of k_t jets in the (η, ϕ) -plane. It can be seen that near the η -edges, there is a high concentration, followed by a low concentration. The distribution becomes more uniform for $|\eta| < 0.5$. Upper right: ρ versus η for k_t jets. It can be seen that the jets near the edges have a significantly lower ρ than jets nearer to the center. The boundary appears to be close to $|\eta| = 0.5$. Middle left: Distribution of anti- k_t jets in the (η, ϕ) -plane. Middle right: ρ versus η for anti- k_t jets. Bottom: Distribution of ρ for k_t jets, both with and without a cutoff at $|\eta| = 0.5$.

random cone, the quantity $\delta p_T^{\text{RC}}[3]$

$$\delta p_T^{\text{RC}} = \left(\sum_{\text{all tracks}} p_T \right) - \int_{A_{\text{RC}}} \rho(\varphi) d\eta d\varphi. \quad (3)$$

is computed. For the approximation described above, the integral is just equal to $\pi R^2 \cdot \text{median}\{\rho\}$, where the median excludes the two leading jets. The distribution of the results of this calculation are shown in Figure 13.

These distributions are used to assess the quality of the background estimation. The first property used is the mean of the background after background subtraction. This is defined as the mean of a gaussian fitted to the left side ($[\mu - 3\sigma, \mu + 0.5\sigma]$) of the distribution, since this part is expected not to contain hard parton contributions. Since this quantity represents the mean of the background after the background has been subtracted, this quantity should be close to zero. The deviation from zero gives a measure of the quality of the estimation. The second property that is used is the width of the background, which quantifies the uncertainty in the background that was subtracted. This width is defined as the root mean square of the distribution, since the unfolding algorithm takes the shape of the *entire* δp_T^{RC} distribution as a parameter, and the natural definition for a width is then the root mean square of the entire distribution. The unfolding algorithm works better as the width gets smaller. Therefore, this also gives a measure of quality. Both properties can be viewed in Figure 13.

3.5 Improving background estimation

To improve estimation of the v_n contribution to the soft background, several things are done. Firstly, the random cones whose center lies within a distance R of the leading anti- k_t jet are ignored. This is done because random cones are meant to study the quality of the estimation of the soft background, and random cones near the leading anti- k_t jet are likely to contain tracks from hard interactions. Figure 14 shows how this changes the means and standard deviations of the gaussians fitted to the δp_T^{RC} distributions. The means are not expected to change when a different fit method is used. At the

moment of writing, it is not understood why this is not the case.

Also shown in figure 14 is the difference between random cones which are in-plane ($\cos(2(\varphi_{\text{RC}} - \Psi_2)) > 0.8$) and random cones which are out-of-plane ($\cos(2(\varphi_{\text{RC}} - \Psi_2)) < -0.8$). Cones which are in-plane have a higher than average transverse momentum, while out-of-plane cones have a lower than average transverse momentum. This can clearly be seen in the middle two panels of Figure 14. This is due to anisotropic flow, and it leads to suspect that the width of the δp_T^{RC} distributions might be lowered by taking anisotropic flow into account when estimating the background. To do this, $\rho(\varphi)$ is estimated as follows:

$$\rho(\varphi) = \text{median}\{\rho\} \{1 + 2v_2 \cos[2(\varphi - \Psi_2)] + 2v_3 \cos[3(\varphi - \Psi_3)]\},$$

where v_2 and v_3 are estimated using p_T -weighted fits. The δp_T^{RC} distributions are then calculated with (3). However, integral in (3) has to be taken over a disc, which leads to intractable integrals. To solve this, two solutions are proposed:

- Approximate the cosines in $\rho(\varphi)$ by 5-th order Taylor polynomials. This should lead to only minor differences, since the cosines do not vary much over the domain over which is integrated.
- Replace the circular integration area by a square one with sides equal to the circle's diameter. The resulting integral is then multiplied by a factor to take into account the difference in area between a disc and a square.

Both methods have been investigated, and the results are shown in the bottom panels of Figure 14 and the upper panels of Figure 15. In the bottom right panel of Figure 14, the width using a modulated background is also compared to both in-plane, out-of-plane and nonmodulated widths. For central collisions, the modulated background width is similar to the out-of-plane width. The in-plane widths appear closer to the nonmodulated widths. The reason for this is that the in-plane distribution, having higher track density, will also have a larger width than the out-of-plane distribution, due to the Poisson nature of the track density. This last figure also summarizes the methods applied thus far. It can be seen in Figure 15 that the two modulation methods give similar means and widths.

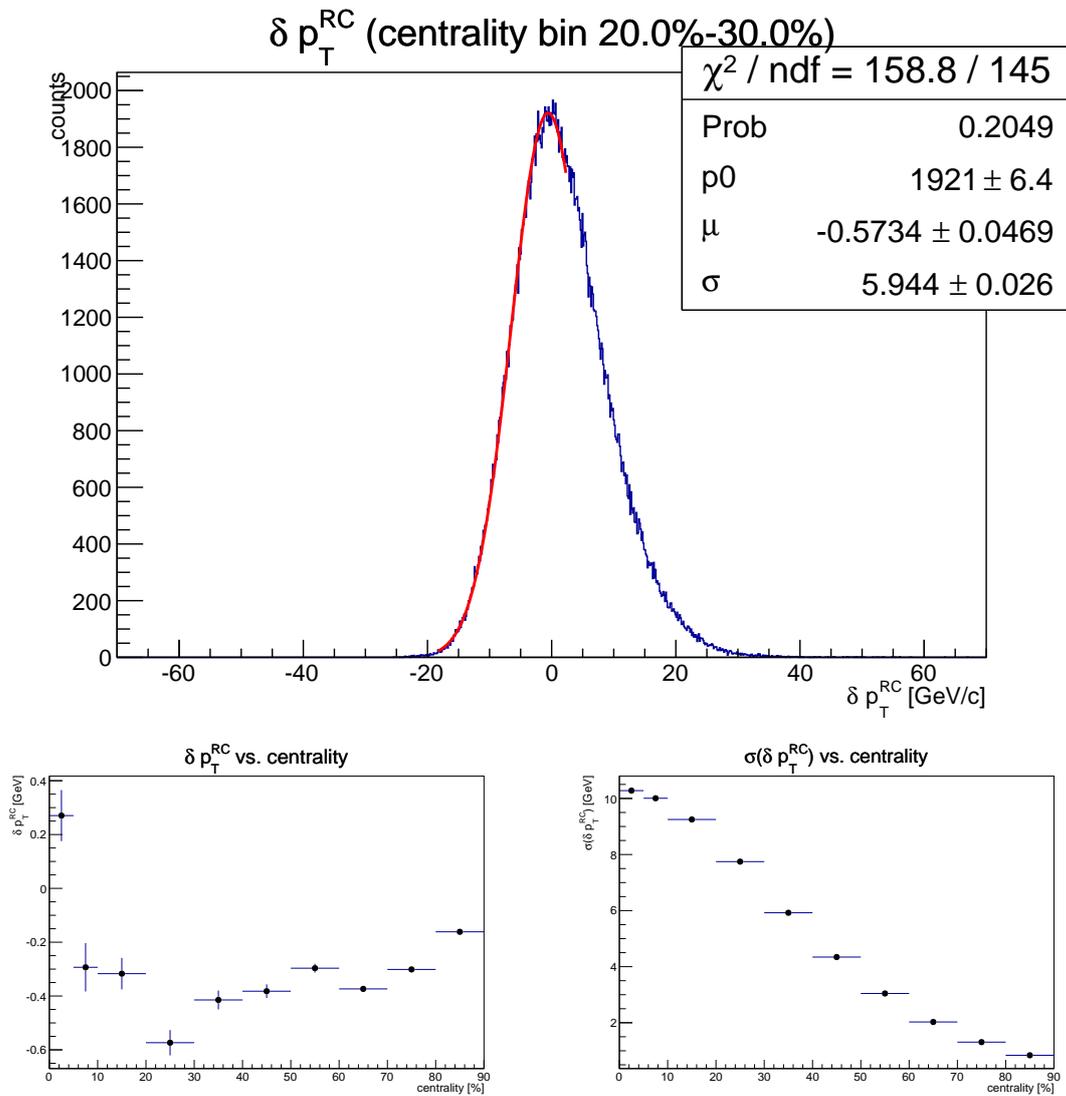


Figure 13: Top: δp_T^{RC} distribution for events with a centrality between 20% and 30%. Lower left: Means of gaussians fitted to the left side of δp_T^{RC} distributions for different centrality classes. Lower right: Root mean squares of the same δp_T^{RC} distributions.

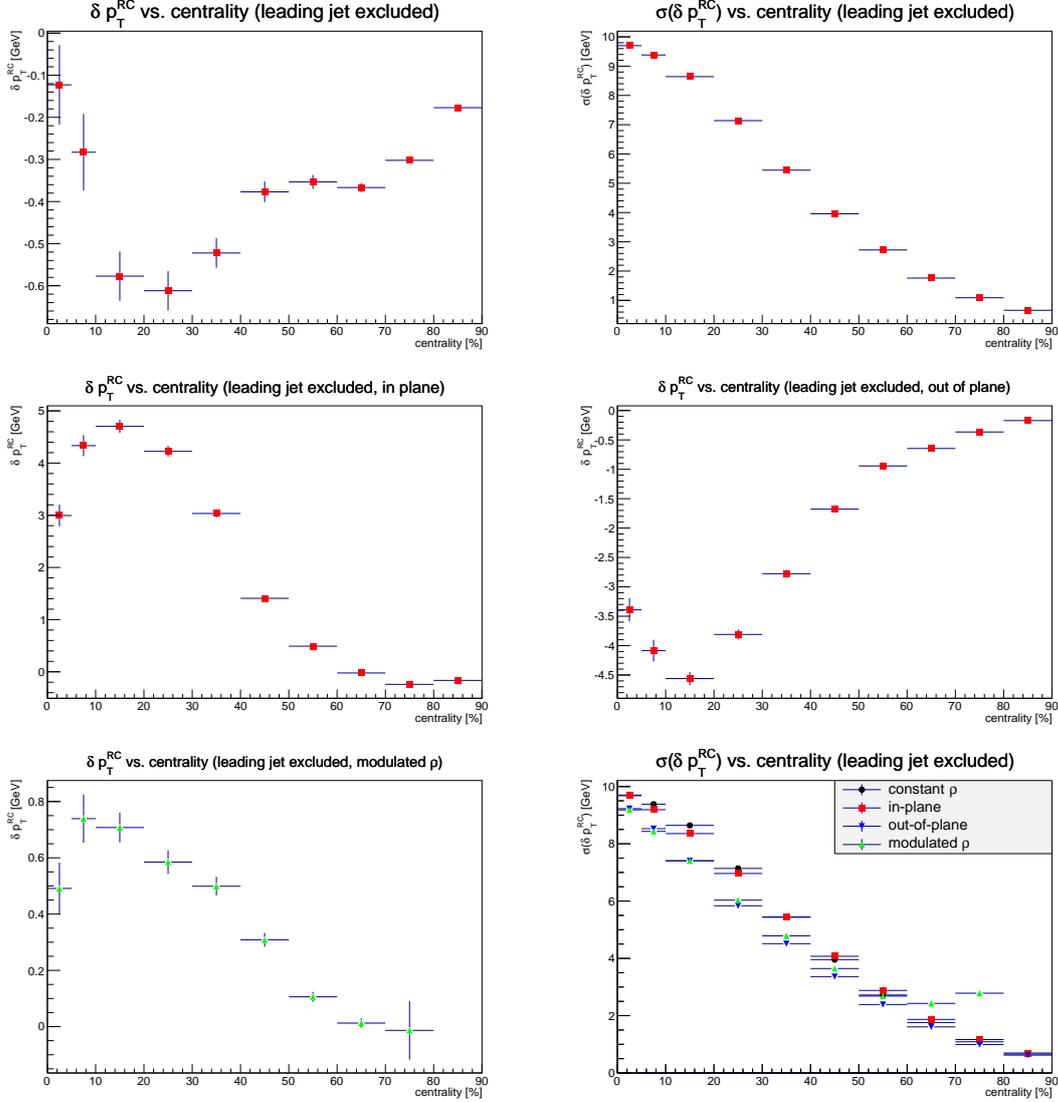


Figure 14: Upper left: means of gaussians fitted to δp_T^{RC} distributions, where cones within a distance R from the leading k_t jet have been ignored. Upper right: root mean squares of δp_T^{RC} distributions with only cones further than R removed from the leading jet. Middle left: means of gaussians fitted to δp_T^{RC} distributions (without those around the leading jet), where only in-plane cones have been used. Middle right: same as upper middle left, but now with out-of-plane cones. Bottom left: means of gaussians, now fitted to δp_T^{RC} distributions where anisotropic flow has been taken into account. The integration has been done over the circular area. Bottom right: root mean squares of the same distribution as used in the lower left panel, with the in-plane root mean squares of the distributions without a modulation for ρ added for comparison.

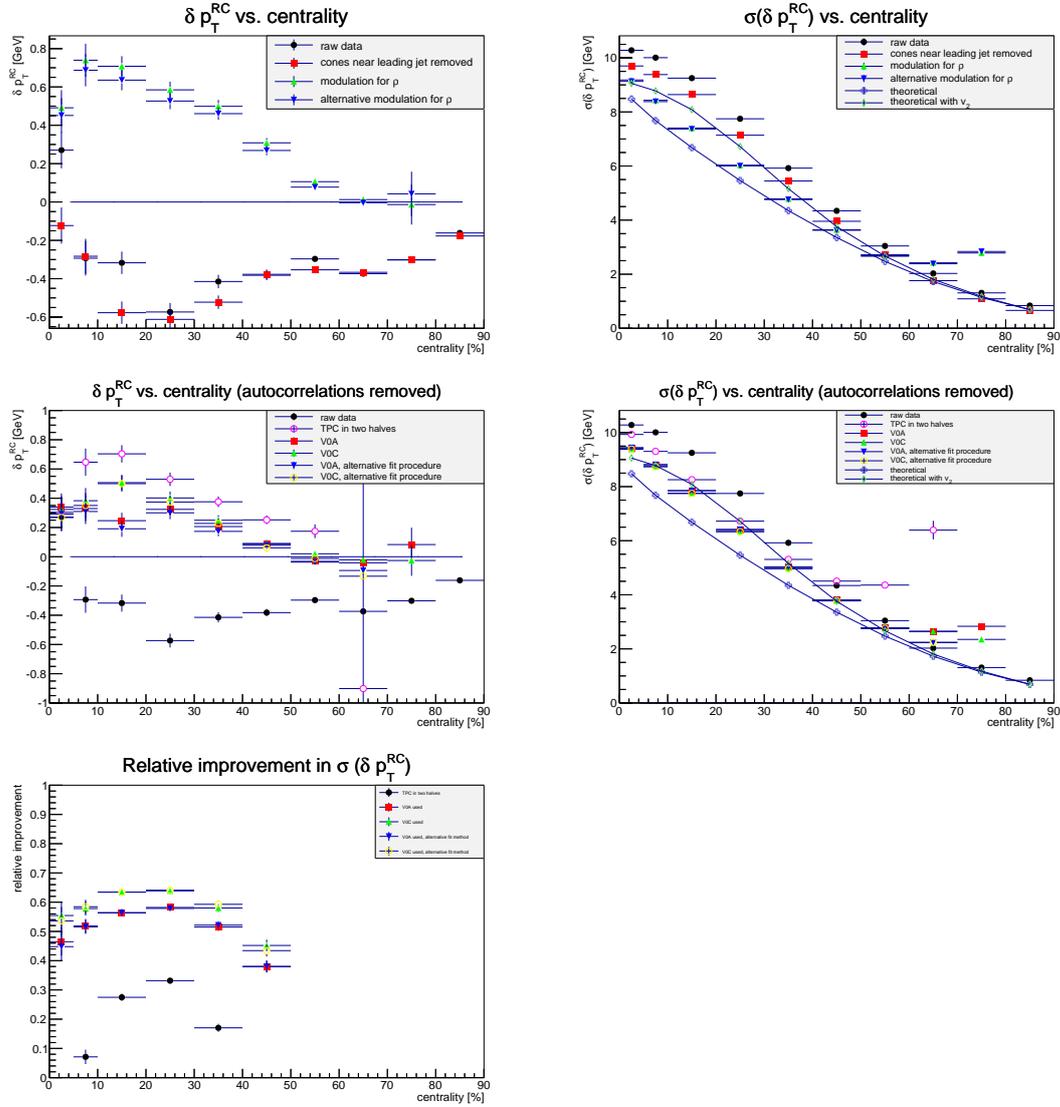


Figure 15: Upper left: means of gaussians fitted to δp_T^{RC} distributions for different background estimations. It can be seen that the two methods for integration over a random cone yield approximately the same result. Upper right: root mean squares of δp_T^{RC} distributions for different background estimations. Also shown is the theoretical width of the spectrum assuming no correlations between tracks. It can be seen that the methods which take anisotropic flow into account have a significantly lower width for central collisions, but that this estimate worsens as the centrality increases. Middle: Same as above, but now for estimation methods which have been corrected for autocorrelations. It can be seen that all methods yield large widths above a certain centrality, but that the methods using the V0 detectors appear to remain good for a larger range of centralities. Bottom: Difference in width caused by modulation as a fraction of theoretical difference.

The last improvement that will be made concerns the removal of autocorrelations, which cause a positive bias in the determination of v_2 and v_3 , and therefore cause the anisotropic effects in the soft background to be overcompensated. To correct this, flow coefficients are used which have minimal autocorrelations (see section 2.4). The results are shown in figure 15. In the plots of $\sigma(\delta p_T^{\text{RC}})$, the theoretical width of the δp_T^{RC} spectrum is also shown. It is given by the following formula:[3]

$$\sigma(\delta p_T^{\text{RC}}) = \sqrt{N_A \cdot \sigma^2(p_T) + N_A \cdot \langle p_T \rangle^2},$$

where N_A denotes the expected number of tracks within a cone. This formula assumes an uncorrelated track distribution for each event, with $v_n = 0$. To incorporate v_2 , an extra term is added:

$$\sigma(\delta p_T^{\text{RC}}) = \sqrt{N_A \cdot \sigma^2(p_T) + (N_A + 2v_2^2 N_A^2) \cdot \langle p_T \rangle^2}.$$

As can be seen in Figure 15, the theoretical width which includes v_2 is smaller than that of the non-modulated δp_T^{RC} distribution. This is probably due to nonflow, which would mean that the widths of the modulated distributions are probably also affected by nonflow. Assuming that both the modulated and the nonmodulated distributions are affected by the same amount, it is still possible to compare the decrease in width when using a φ -modulated background to the theoretical decrease. This is done by looking at the decrease in width caused by modulating the background as a fraction of the theoretical width.¹⁵ The result of this computation is shown in the bottom panel of Figure 15. It can be seen that the V0C detector with the simple fit method causes the largest increase. This is possibly due to the high resolution of this detector. It can also be seen that methods which result in a larger estimate of v_n generally achieve a smaller δp_T^{RC} width. The cause of this and possible uses for it is a subject of further study.

3.6 Anti- k_t jet spectra

Finally, the anti- k_t jet p_T spectra are studied. Recall that the anti- k_t algorithm produces circular jets. This is well-suited for studying hard parton

¹⁵For the unmodulated background, the δp_T^{RC} distribution is used that *does not* include random cones near the leading jet.

scatterings, but as a byproduct this property also causes many tracks to be clustered into very small jets, which have no physical meaning. Therefore, when studying anti- k_t spectra, an area cutoff is applied. In this study, two different cutoffs were used: one which selects all jets with area larger than 0.2, and one which selects all jets with area larger than 0.4. Subsequently, the background estimated with the integral in (3) is subtracted. The integration domain is taken to be a disc with area equal to the jet area. The results for different centralities can be seen in figure 16. Near $p_T = 0$, one can see that there are many jets with small areas, which is expected, since these jets consist of fewer particles. At high p_T , the area cuts don't influence the spectrum at all.

4 Conclusions and Outlook

As was shown in detail in section 2, the methods to estimate v_n have a significant but small bias. This means that this method is a reliable way to estimate flow parameters on an event by event basis. Earlier methods, such as using cumulants, can only estimate averages of these flow parameters over many events, which makes such estimates unsuitable for estimating soft jet backgrounds. It can also be seen, however, that the method is less accurate for more peripheral collisions, where there are fewer tracks per event.

Another conclusion is that $dp_T/d\varphi$ flow is larger than $dN/d\varphi$ flow. Since the two estimates yield different results, the two different estimates might also yield different results when applied to jet background estimation. In this study, $dp_T/d\varphi$ estimates have been used for jet backgrounds, but the debate as to whether this is indeed correct is still unsettled.

The calculation of the *true* widths of the v_n distributions makes yet another way to estimate jet backgrounds possible. Since the widths appear to be small, it might be possible to estimate the v_n per event as the average v_n over many event. This way, methods which normally do not work on an event-by-event basis could still be used for background estimation. Some such methods have the advantage of being less biased.

In section 3, it was shown how the soft background to hard parton scatterings can be estimated, and that this estimation improves if the

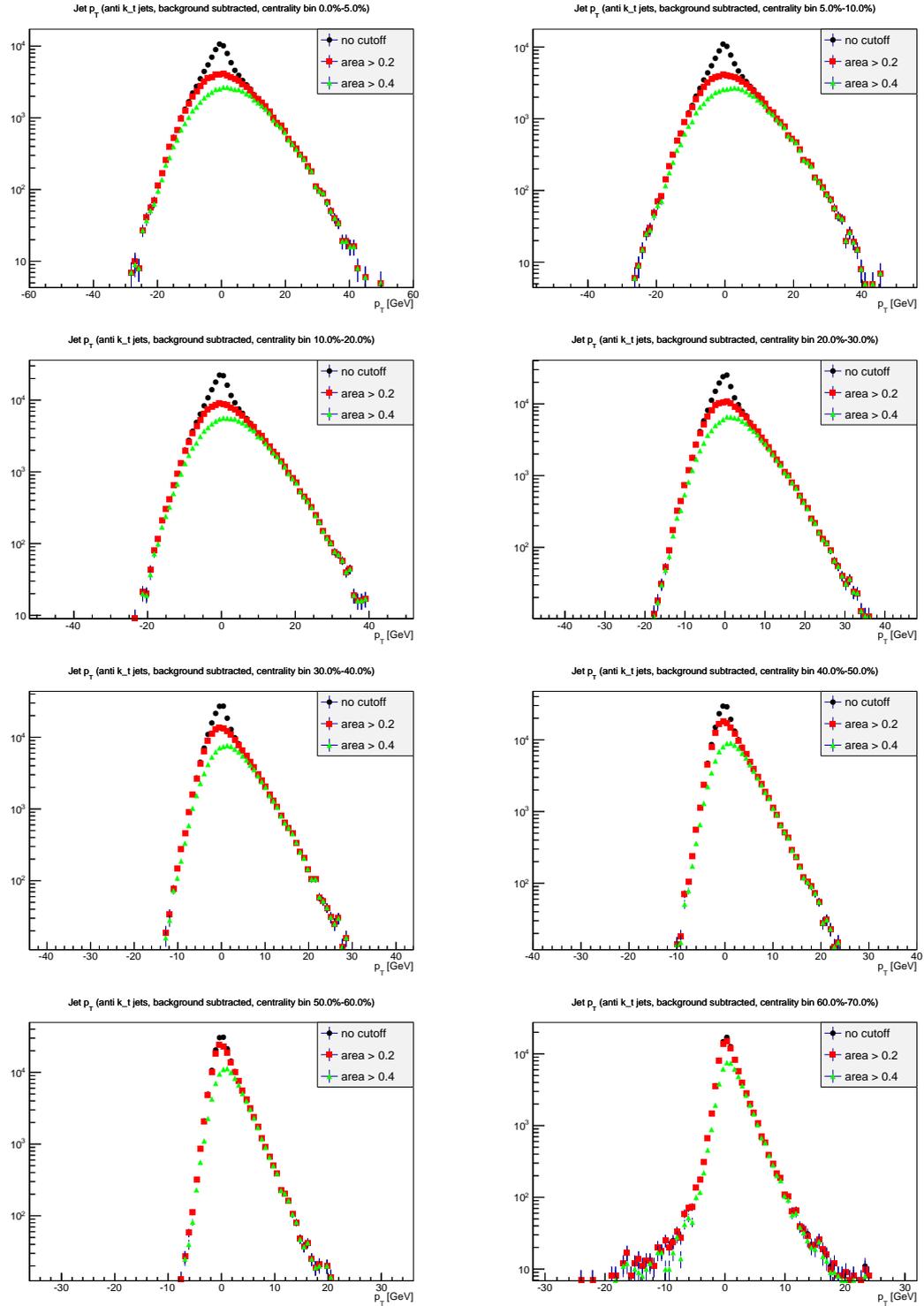


Figure 16: Anti k_t jet p_T spectra for different centralities. The high ends of the spectra are unaffected by the area cutoffs.

anisotropies in the flow are taken into account. However, the means of the background after background subtraction differ significantly from zero, which indicates a systematic effect present in this method. Also, for more peripheral collisions, taking v_2 and v_3 into account resulted in wider δp_T^{RC} distributions. This could be taken into account in future analyses by either improving the method to make it more accurate, or by applying the flow correction only to events which are central enough for the flow correction to improve the results.

This paper covers just the first steps in a jet quenching analysis. Next steps would be to run the unfolding algorithm on the anti- k_t jet spectra that were obtained, to obtain the hard part of this spectrum. Also, the jet spectra could be studied as a function of $\varphi - \Psi_2$, to study in-plane versus out-of-plane differences. Another interesting study would be to investigate the influence of biases in the determination of v_n on jet spectra after unfolding. If such biases are small enough, they might be ignored, but if they are large, a way must be found to either avoid them, or to quantify them and correct for them. The results of these steps should further indicate the way forward to a more complete understanding of the Quark Gluon Plasma.

References

- [1] Ante Bilandzic. *Anisotropic Flow Measurements in ALICE at the Large Hadron Collider*. PhD thesis, Universiteit Utrecht, 2012.
- [2] Matteo Cacciari and Gavin P. Salam. Dispelling the N^3 myth for the k_t jet-finder. *Physics Letters B*, August 2006.
- [3] The ALICE Collaboration. Measurement of event background fluctuations for charged particle jet reconstruction in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV. *Journal Of High Energy Physics*, 2012, March 2012.
- [4] Alexander Colliander Hansen. Pseudorapidity dependence of elliptic flow in Pb+Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV with ALICE. Master’s thesis, University of Copenhagen, August 2011.
- [5] Arthur M. Poskanzer and Sergei A. Voloshin. Methods for analyzing anisotropic flow in relativistic nuclear collisions. *Physical Review C*, May 1998.
- [6] Ilya Selyuzhenkov and Sergei A. Voloshin. Effects of non-uniform acceptance in anisotropic flow measurement. *Physical Review C*, March 2008.