

Chapter 10

Underlying Event Activity

In this chapter, measurements of K_S^0 , Λ , and $\bar{\Lambda}$ are reported as a function of underlying event activity classifiers R_T , $R_{T,\min}$, and $R_{T,\max}$. These observables quantify the magnitude of underlying event and are an experimental proxy of the n_{MPI} .

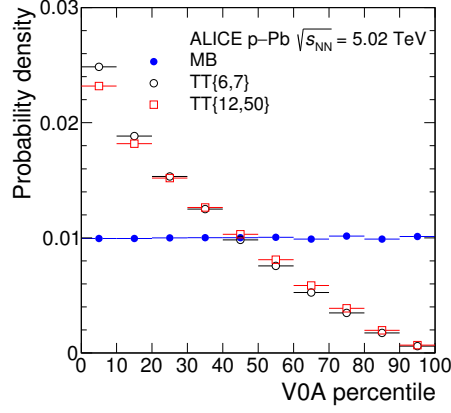
10.1 Motivation for studying event sub-structure

10.1.1 Jet pedestal: underlying event

As discussed in Chapter X, the underlying event, also known as jet pedestal in some literature, is composed of particles that are not directly related to the primary hard scattering and its related fragmentation. It can be studied to extract accurate information about the hard scattering process by subtracting it in precision measurements of jet properties. Moreover, since it is a manifestation of the proton substructure and the parton interactions, it can give us insight into the parton dynamics in the nonperturbative QCD region.

10.1.2 Hard process–multiplicity bias

Studying QGP phenomena in small systems as a function of event activity is challenging due to selection biases that arise when analyzing the data. It is known that selecting events with large momentum transfer leads to a bias towards higher multiplicities (and underlying event), and conversely, selecting events with higher multiplicities (and UE) enhances the hard processes. This bias can be understood in several



(a)

Figure 10.1: TBA

ways. Firstly, a hard process tends to occur with lower impact parameters, which in turn leads to higher particle multiplicities. Secondly, an event with n partonic interactions has n chances of containing a hard process. Lastly, harder processes fragment into more particles, further contributing in higher event activity. As an example, Figure 10.1 shows how the requirement of a high p_T track can skew the forward-rapidity centrality distribution to lower values (higher event activity), as observed in a result from ALICE.

10.1.3 Azimuthal regions and transverse activity

The selection bias of hard processes on UE becomes saturated at high p_T , where the impact parameter bias is fixed and stochastic effects become comparable. This saturation effect can be observed when studying particle production in three topological regions defined with respect to the highest momentum track, which serves as a proxy for the axis of the primary scattering process. The three regions are defined as follows:

1. Towards (also known as "Near"), where $|\phi - \phi^{\text{leading}}| < \frac{\pi}{3}$,
2. Away, where $|\phi - \phi^{\text{leading}}| > \frac{2\pi}{3}$, and
3. Transverse, where $\frac{\pi}{3} < |\phi - \phi^{\text{leading}}| < \frac{2\pi}{3}$.

Here, ϕ^{leading} is the azimuthal angle of the leading track. This definition is illustrated in Figure 10.2.

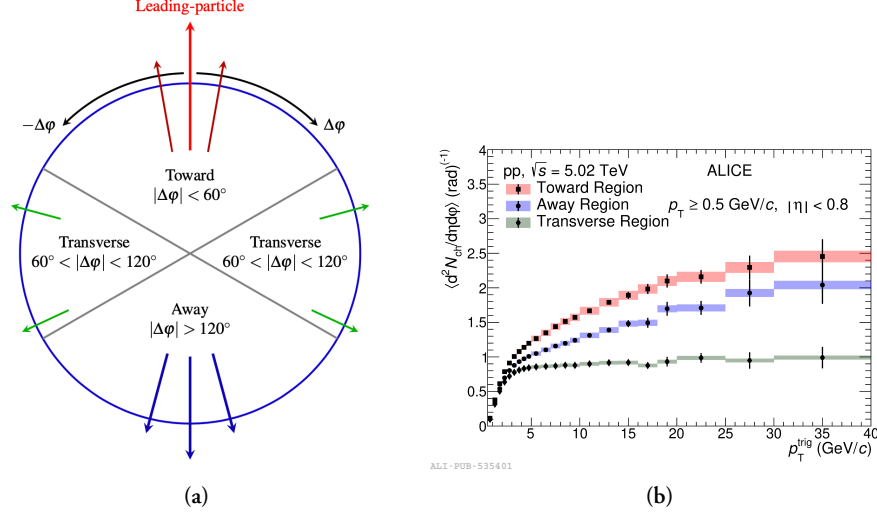


Figure 10.2: TBA

Studying particle multiplicity (or sum of their p_T) in these regions as a function of the transverse momentum of the leading track p_T^{leading} reveals that in the regions Towards and Away, the multiplicity continues to increase with the hardness of the primary process. These regions contain the leading and the recoil jet, respectively. In contrast, in the Transverse region, the multiplicity (further denoted as N_T in this thesis but $N_{\text{ch}}^{\text{trans}}$ is also used in cited literature) reaches a plateau at around 5 GeV/c. In this region, the underlying event becomes independent of the strength of the primary process, and the selection bias is minimized. Notably, this phenomenon is universal regardless of the system size or collision energy. As an example, measurements from ALICE are shown in Fig. 10.2.

10.2 R_T as an experimental observable

The magnitude of the underlying event can be measured using the self-normalized ratio:

$$R_T = \frac{N_T}{\langle N_T \rangle}, \quad (10.1)$$

which is often referred to as the underlying event activity, transverse activity, or relative transverse activity in various literature, and also in this thesis. This observable and its uses were suggested in Ref. [1].

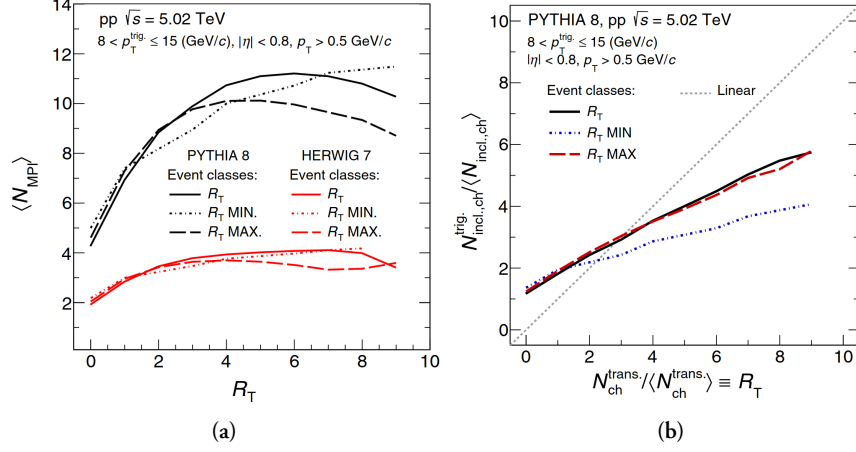


Figure 10.3: TBA

By applying R_T , two limits of events can be studied:

- $R_T \rightarrow 0$: the “ee” limit, where events with minimal UE are selected. These events are dominated by a single hard scattering and can be compared to LEP fragmentation models.
- $R_T \rightarrow \infty$: the “AA” limit, where events with very high transverse activity are selected, which can come from many MPIs and/or from transverse jets. These events may exhibit collective features similar to AA collisions.

10.2.1 Proxy to n_{MPI}

As could be intuitively expected, R_T serves as an experimental proxy for $\langle n_{\text{MPI}} \rangle$. Phenomenological models that incorporate MPIs provide an illustration of this relationship. As shown in Fig. 10.3, Pythia 8 predicts a strong dependence of $\langle n_{\text{MPI}} \rangle$ on R_T until $R_T \lesssim 5$. Similarly, Herwig 7 predicts a dependence until $R_T \lesssim 3$, albeit weaker. Pythia’s prediction for the relationship between R_T and the event multiplicity, which is affine, is also shown in the figure.

10.2.2 Extension to $R_{T,\text{min}}, R_{T,\text{max}}$

To separate the soft and hard components of the underlying event – namely, the MPIs from ISR/FSR – the definition of R_T can be extended. The two transverse

sub-regions can be further classified as Transverse-min or Transverse-max, based on which sub-region has fewer or more particles. Softer contributions from MPIs will enter both sub-regions, whereas the effects of harder radiation should be captured in the Transverse-max sub-region. This makes Transverse-min more sensitive to particle production from MPIs.

Analogously, the following underlying event activity classifiers can be defined:

$$R_{T,\min} = \frac{N_{T,\min}}{\langle N_{T,\min} \rangle} \quad , \quad (10.2)$$

$$R_{T,\max} = \frac{N_{T,\max}}{\langle N_{T,\max} \rangle} \quad , \quad (10.3)$$

where $N_{T,\min}$ and $N_{T,\max}$ are the particle multiplicities in the Transverse-min and Transverse-max sub-regions, respectively. This approach follows measurements developed at UE studies at Tevatron and have been suggested to use in searches for QGP phenomena in small systems based on investigations in phenomenological models? .

According to Pythia 8, as shown in Fig. 10.3, $R_{T,\min}$ and $R_{T,\max}$ follow different relationships with $\langle n_{\text{MPI}} \rangle$. Whereas $\langle n_{\text{MPI}} \rangle$ starts falling as a function of $R_{T,\max}$ (due to the inclusion of mini-jets) at $R_{T,\max} \approx 5$, it continues rising as a function of $R_{T,\min}$ across the entire range. Furthermore, compared to R_T , $R_{T,\min}$ also shows some degree of decorrelation with event multiplicity.

Charged particle p_T spectra

Phenomenological models also reveal a significantly different evolution of transverse momentum spectra of inclusive charged particles based on $R_{T,\min}$ and $R_{T,\max}$, as shown in Fig. 10.4. For the highest reported ranges of $R_{T,\max}$ and R_T , a significant hardening of the spectrum is observed in both Pythia 8 and Herwig 7, similarly to multiplicity studies? , indicating a strong auto-correlation. In contrast, $R_{T,\min}$ exhibits a Cronin-like enhancement at intermediate p_T and a plateau at $p_T \gtrsim 6 \text{ GeV}/c$, even in the highest $R_{T,\min}$ bin.

10.2.3 Track and event selection

The event selection follows the same criteria as the $S_O^{(p_T=1.0)}$ measurement discussed in Section 9.1.4, which conform to the standard analysis of light flavour hadrons versus multiplicity in pp collisions conducted in ALICE. The $\text{INEL} > 0$ events, which require at least one hit in either V0A or V0C scintillators and a track reconstructed within $|\eta| < 1$, are used. The SPD is used for the reconstruction of primary vertex,

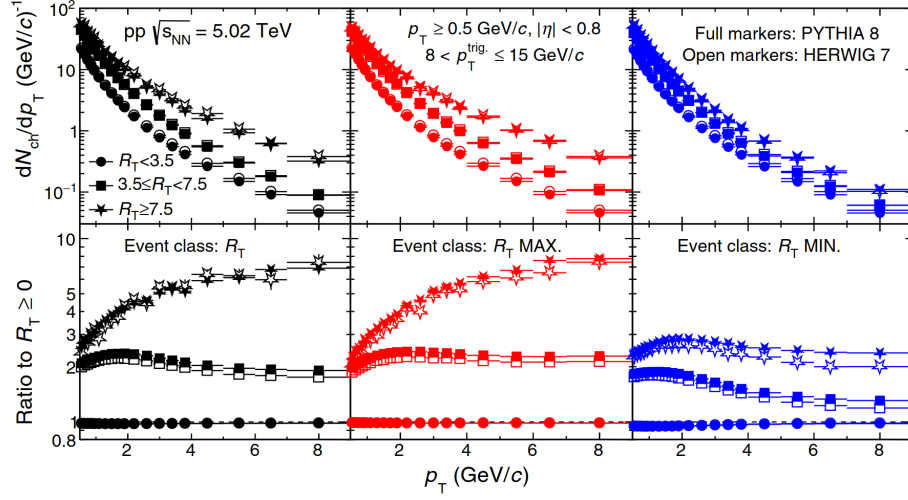


Figure 10.4: TBA

which is further required to have $|\Delta z| < 10 \text{ cm}$ to reject out-of-bunch pile-up. To remove in-bunch pile-up, events with multiple reconstructed vertices are excluded.

Events are required to have a leading track with reconstructed momentum $5 < p_T^{\text{leading}} < 40 \text{ GeV}/c^1$. These values were chosen to access the plateau in transverse activity and isolate the UE while retaining a large data sample. Maintaining a high momentum and spatial resolution of the leading track is crucial in this measurement. However, this can be compromised at high p_T when a significant portion of the track curvature can fall between two sectors of the TPC. To address this issue, geometrical cuts are used, as discussed in Section ??.

For both the leading particle as well as the particles entering N_T and R_T calculations, tracks are required to be within $|\eta| < 0.8$ and have $p_T > 0.15 \text{ GeV}/c$, and must satisfy the following:

1. “Hybrid tracks”, described in more detail in Section ??, are used for both leading and N_T tracks to ensure a high level of azimuthal acceptance uniformity. These tracks consist of high-quality “global track” requirements, including the SPD information, which leads to azimuthal non-uniformity, and “complementary track” cuts, a looser set requiring only ITS and TPC in cases where the first are not satisfied.
2. For the leading track, strict p_T -dependent DCA cuts are applied in the trans-

¹Note that p_T spectrum is falling very steeply, at an approximately exponential rate, making the upper bound negligibly restrictive compared to the lower bound.

verse direction ($|\text{DCA}_{xy}| < 0.0182 + \frac{0.0350}{p_T^{0.01}} \text{ cm}$, $p_T \in [\text{GeV}/c]$), to ensure good momentum resolution and that the track is a primary one.

3. For the N_T tracks, a DCA cut ($|\text{DCA}_{xy}| < 0.06 \text{ cm}$) is required to avoid biases in V^0 measurements, as explained in the text below.

10.2.4 R_T measurements of neutral particles vs. charged particles

The V^0 s are neutral particles and thus, they cannot be leading tracks nor enter N_T ($N_{T,\min}$, $N_{T,\max}$) and R_T ($R_{T,\min}$, $R_{T,\max}$) calculations. This has several implications:

1. V^0 s suffer much less from auto-correlation biases than $\pi/K/p$, which can be seen in azimuthal distributions and in K_S^0/K^\pm ratios. Requiring high/low N_T/R_T can lead to an increase/decrease of charged particles in the Transverse region due to selecting fluctuations in addition to the UE scaling. However, this effect is significantly smaller for neutral V^0 s. This behaviour is shown in Fig. 10.5. It is important to bear this caveat in mind when comparing p_T spectra and yields of $\pi/K/p$ and V^0 s.
2. While N_T is always at least 1 for $\pi/K/p$ in the Transverse region, for V^0 s it can be equal to 0. Similar logic applies to the Transverse-min/max sub-regions and $N_{T,\min}/N_{T,\max}$.
3. The maximum p_T measurable for $\pi/K/p$ in the Toward region is limited to $p_T < 5 \text{ GeV}/c$, where the trigger requirement leads to a trivial increase. For V^0 s, however, this limitation does not apply and their measured p_T range does not need to be restricted.
4. The charged daughters of V^0 s could sometimes enter N_T , leading to significant biases at low p_T in the Toward and Away regions of K_S^0/K^\pm ratios.

In this thesis, the behaviour described in the last point was rectified by making N_T track candidates and V^0 daughter tracks two disjunct sets. This was achieved by applying the $|\text{DCA}_{xy}| > 0.06 \text{ cm}$ cut, used in the V^0 reconstruction as discussed in Section ??, in opposite ways. This reduces the N_T track candidates by less than 5%. The effect of this solution can be seen in Fig. 10.6.

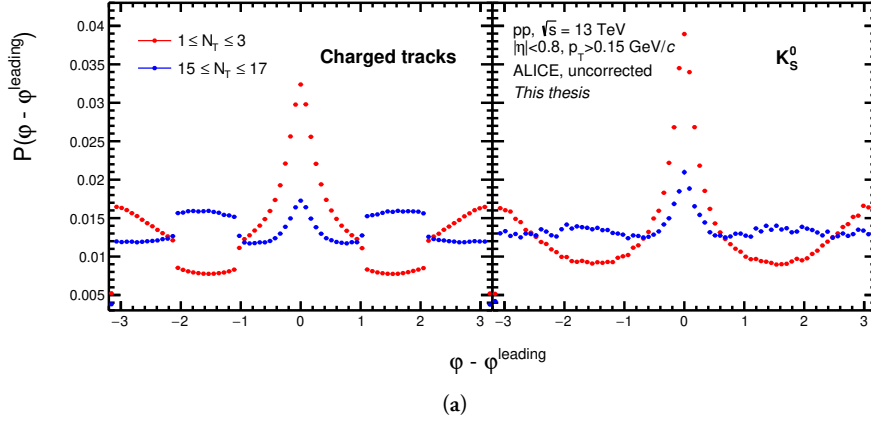


Figure 10.5: TBA, maybe move to chapter about tracks.

10.3 Bayesian unfolding procedure

The V^0 s measurements are conducted as a function of the number of measured tracks N_T^m within the detector acceptance. The measured multiplicity N_T^m includes a fraction of the true primary charged-particle multiplicity N_T^t not lost due to acceptance, efficiency, or track selection, as well as contributions from secondary particles or particles smeared into the measurement's kinematic acceptance due to detector resolution (i.e., from $p_T < 0.15$ GeV/c). These effects fluctuate on an event-by-event basis and thus there is no unique correlation between N_T^m and N_T^t . This means that events with true multiplicity N_T^t can be measured with different N_T^m , contributing to V^0 measurements in multiple N_T^m bins. Therefore, each spectrum contains particles from events with many true multiplicities N_T^t .

This thesis uses a Bayesian unfolding procedure, as discussed in Ref.?, to convert V^0 s measurements as a function of N_T^m into measurements as a function of N_T^t and thus correct for the mentioned effects.

10.3.1 One-dimensional unfolding

The measured multiplicity distribution $n_{ev}(N_T^m)$ can be mathematically represented as the result of convolving (or “folding”) the true multiplicity distribution produced by the collisions, $n_{ev}(N_T^t)$, with the detector's response function. The response matrix S_{mt} , which represents the conditional probability $P(N_T^m | N_T^t)$ of an event with multiplicity N_T^t being measured with multiplicity N_T^m , can be obtained from MC simulations of the apparatus. Using this matrix, also shown in Fig. 10.9, $n_{ev}(N_T^m)$

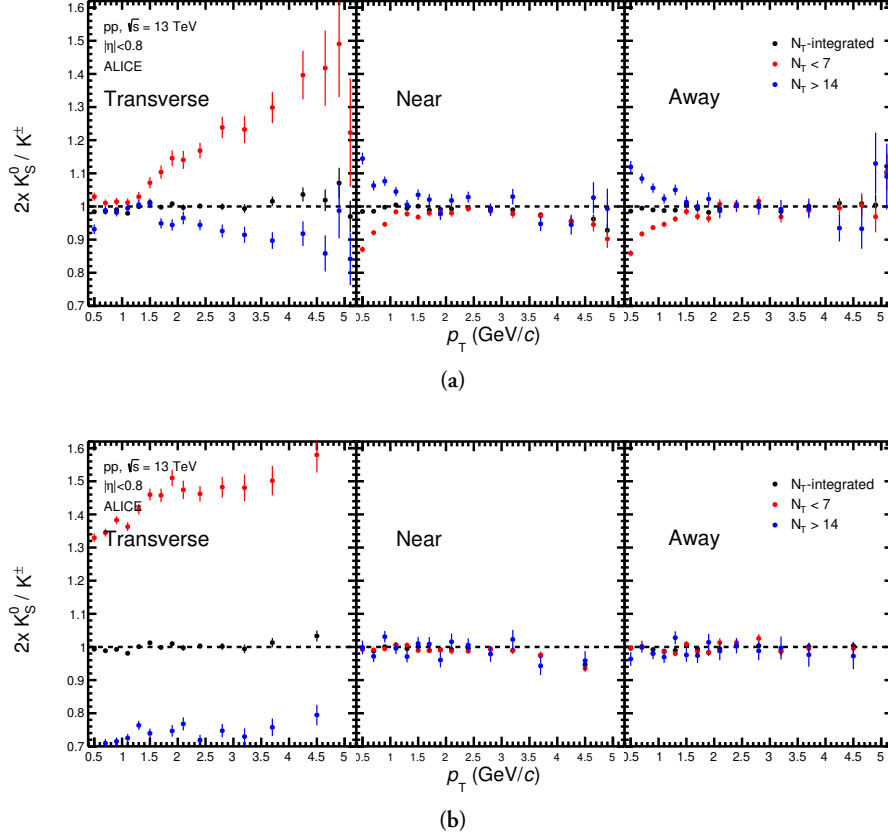


Figure 10.6: TBA

can be expressed in terms of $n_{\text{ev}}(N_T^t)$ as follows:

$$n_{\text{ev}}(N_T^m) = \sum_t S_{mt} \cdot n_{\text{ev}}(N_T^t) \quad , \quad (10.4)$$

To obtain the true multiplicity distribution from the measured distribution, the inverse of S_{mt} could be used, hypothetically, as shown below:

$$n_{\text{ev}}(N_T^t) = \sum_m S_{mt}^{-1} \cdot n_{\text{ev}}(N_T^m) \quad . \quad (10.5)$$

However, the inverse S_{mt}^{-1} may have multiple or zero solutions, making this approach unfeasible. Alternatively, S_{mt}^{-1} could be obtained directly from MC simulations, just like the detector response. However, this matrix would then strongly depend on the generated N_T^t distribution and be significantly model-dependent, as physics generators vary in their N_T^t predictions. In contrast, the detector response is mostly affected

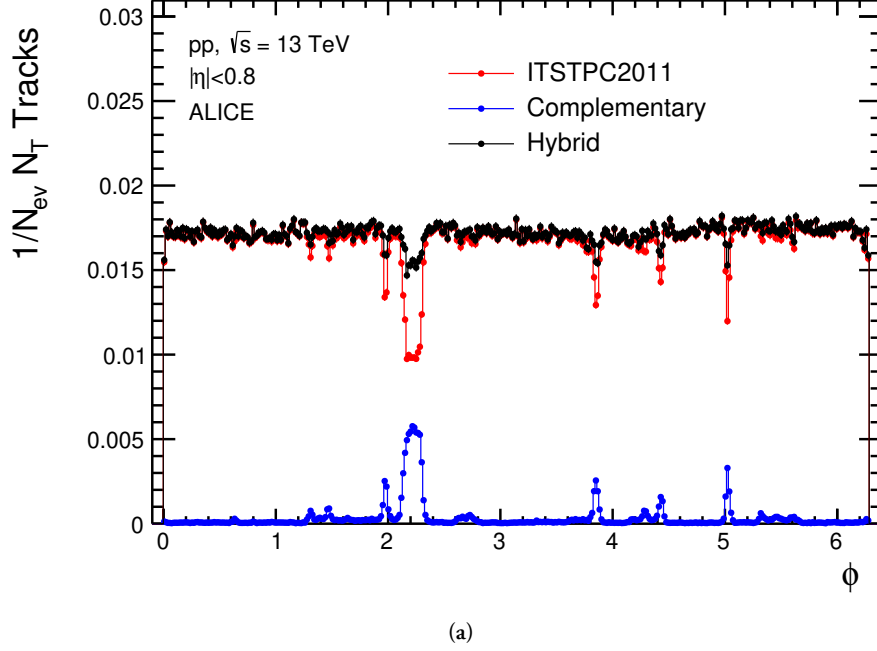


Figure 10.7: TBA, maybe move to chapter about tracks.

by the accuracy of the particle propagation simulations, which is a lot more understood. Therefore, an iterative numerical procedure based on Bayes' theorem is used to obtain the unfolding matrix M_{mt} , which represents the conditional probabilities $P(N_T^t | N_T^m)$.

In this application, Bayes' theorem can be expressed in terms of N_T^m and N_T^t as follows,

$$P(N_T^t | N_T^m) = \frac{P(N_T^m | N_T^t) P(N_T^t)}{P(N_T^m)} \quad , \quad (10.6)$$

where $P(N_T^t)$ and $P(N_T^m)$ are probability distributions for an event occurrence with N_T^t and N_T^m , respectively. Assuming that $P(N_T^t)$ is known, $P(N_T^m)$ can be calculated as follows:

$$P(N_T^m) = \sum_t P(N_T^m | N_T^t) P(N_T^t) \quad . \quad (10.7)$$

Therefore, using Eq. 10.6, the conditional probability in the unfolding matrix can be written as follows:

$$P(N_T^t | N_T^m) = \frac{P(N_T^m | N_T^t) P(N_T^t)}{\sum_{t'} P(N_T^m | N_T^{t'}) P(N_T^{t'})} \quad . \quad (10.8)$$

However, $P(N_T^t)$ (the “prior”) is initially unknown and must be arbitrarily chosen. The unfolding matrix can be calculated using this prior, and the unfolded distribution can be obtained as follows:

$$\hat{n}_{\text{ev}}(N_T^t) = \sum_m P(N_T^t | N_T^m) n_{\text{ev}}(N_T^m) \quad . \quad (10.9)$$

This unfolded multiplicity can subsequently be used to update the prior as follows:

$$\hat{P}(N_T^t) = \frac{\hat{n}_{\text{ev}}(N_T^t)}{\sum_{t'} \hat{n}_{\text{ev}}(N_T^{t'})} \quad , \quad (10.10)$$

starting a new iteration. The updated $\hat{P}(N_T^t)$ is closer to the true $P(N_T^t)$ than the initial guess because the arbitrarily chosen prior is constrained by the $n_{\text{ev}}(N_T^m)$ observable, which contains information about $P(N_T^t)$.

Multiple approaches can be taken to choose the prior: a uniform distribution, the N_T^t distribution generated by a model, or the N_T^m distribution acquired from data. In this thesis, the prior choice was found to not play a role.

TBA: Normalisation

The χ^2/ndf is calculated to determine the validity of the correction and the stopping point for the iterative process. It is calculated by comparing the N_T^t distribution – known a priori in the simulations – and the unfolded $\hat{n}_{\text{ev}}(N_T^t)$ distribution, where ndf refers to the number of data points in the distribution. The process is stopped when χ^2/ndf reaches a minimum value or the iterations take a maximum number of steps n_{iter} . This is imposed to avoid overfitting and overestimation of statistical uncertainties. The n_{iter} values are reported in Tab. 10.1. The entire iterative process is summarised in a diagram shown in Fig. 10.8.

The used response matrix, as well as the resulting unfolding matrix, can be seen in Fig. 10.9. The method still exhibits some degree of model dependence due to the generation of the response matrix. Previous studies in ALICE have compared the response matrix for N_T acquired from Pythia 8 and from EPOS LHC MC simulations, which revealed that the effect is less than 1%. This effect is taken into consideration as a source of systematic uncertainty.

Table 10.1: TBA.

Unfolding observable	N_T	$N_{T,\text{min}}$	$N_{T,\text{max}}$ I
n_{iter}	20 (max.)	10	18

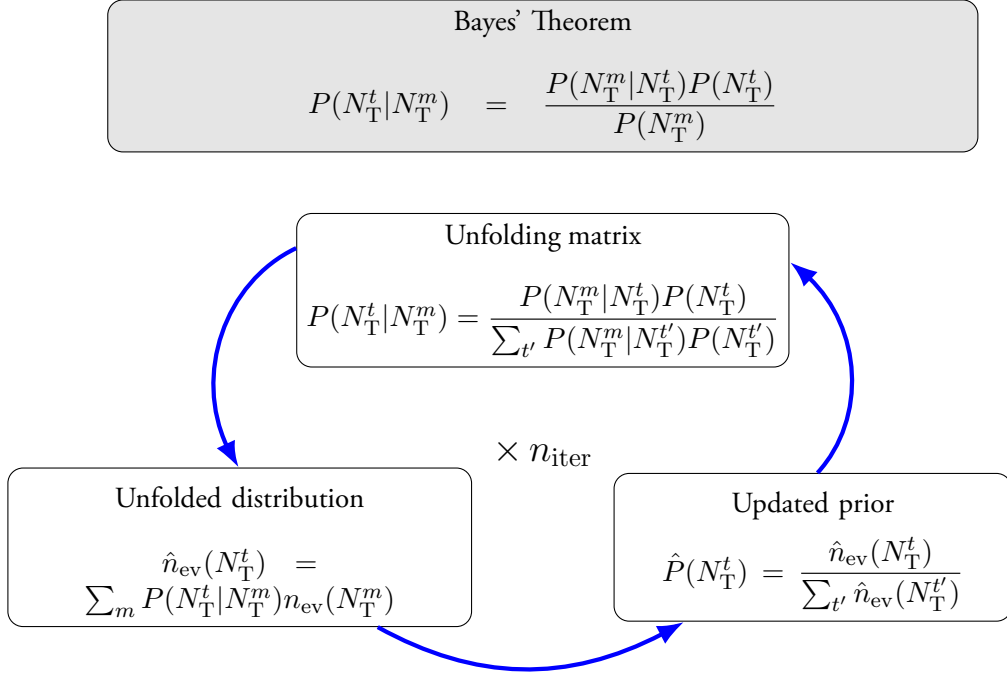


Figure 10.8: Diagram showing the Bayesian unfolding

10.3.2 Propagation of statistical uncertainties

TBA

10.3.3 Unfolding of K_S^0 , Λ , and $\bar{\Lambda}$ p_T spectra

TBA

10.4 R_T , $R_{T,\min}$, $R_{T,\max}$ distributions

The unfolded N_T , $N_{T,\min}$, and $N_{T,\max}$ distributions were self-normalised to obtain the R_T , $R_{T,\min}$, and $R_{T,\max}$ distributions, respectively. They are shown in Fig. 10.10 and compared with predictions from Pythia 8 (Monash tune and Ropes tune) as well as EPOS LHC. The different quantiles corresponding to the $R_T/R_{T,\min}/R_{T,\max}$ ranges used in this measurement are also highlighted. It should be noted that since $N_T, N_{T,\min}, N_{T,\max} \in \mathbb{N}_0$, the $R_T/R_{T,\min}/R_{T,\max}$ distributions are not continuous observables.

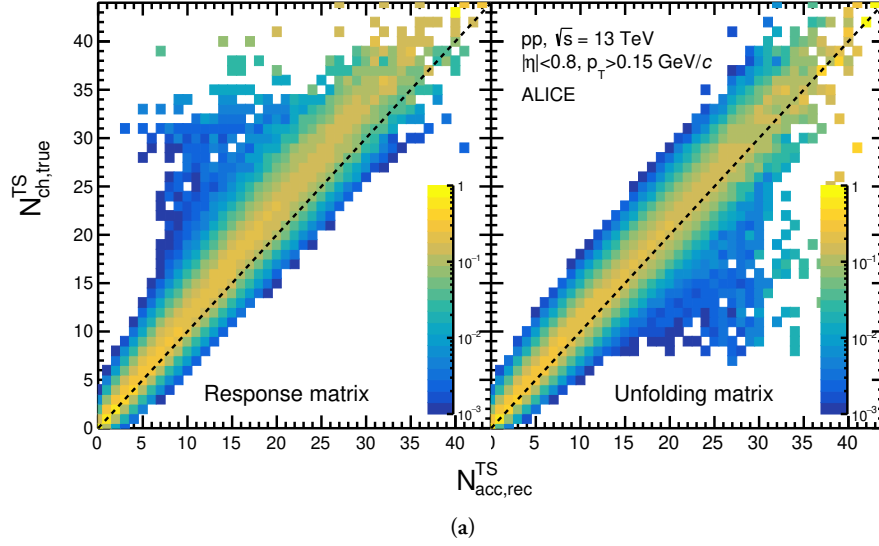


Figure 10.9: TBA

10.5 Systematic uncertainties

The systematic uncertainties on the p_T spectra were determined individually for each R_T interval and azimuthal region, following the procedures described in Section ???. They are reported in Fig. 10.11, Fig. ??, and Fig. ?? for the K_S^0 , Λ , and $\bar{\Lambda}$, respectively. As there are no reasons to believe they should differ, they are then also applied in the $R_{T,\min}$ and $R_{T,\max}$ measurements.

TBA comment on the largest ones

10.5.1 Uncertainties from the unfolding procedure

TBE

The deviations between the generated p_T spectra and the reconstructed, corrected, and unfolded p_T spectra were used to determine the systematic uncertainties associated with the unfolding procedure. To isolate the effect of unfolding from other reconstruction effects, the “non-closures” in each $R_T/R_{T,\min}/R_{T,\max}$ interval were divided by the non-closure in the $R_T/R_{T,\min}/R_{T,\max}$ -integrated bin.

The unfolding systematic uncertainties exhibited a large amount of correlation between K_S^0 and Λ . This correlation was expected, as the V^0 species should unfold in

similar patterns. Therefore, the systematic uncertainty on the baryon-to-meson ratio was also calculated independently to avoid these correlations and reduce the systematic uncertainty on those results.

Moreover, in the most extreme bins, the non-closures sometimes exhibited unrealistic deviations from unity due to limited statistics and fluctuations. To address this issue, a smoothing procedure was applied by fitting the resulting uncertainties with first- and second-order polynomials. The results are shown in Fig. 10.14.

10.5.2 Uncorrelated uncertainties

TBA

10.6 Mean transverse momentum

TBA

10.7 Transverse momentum spectra

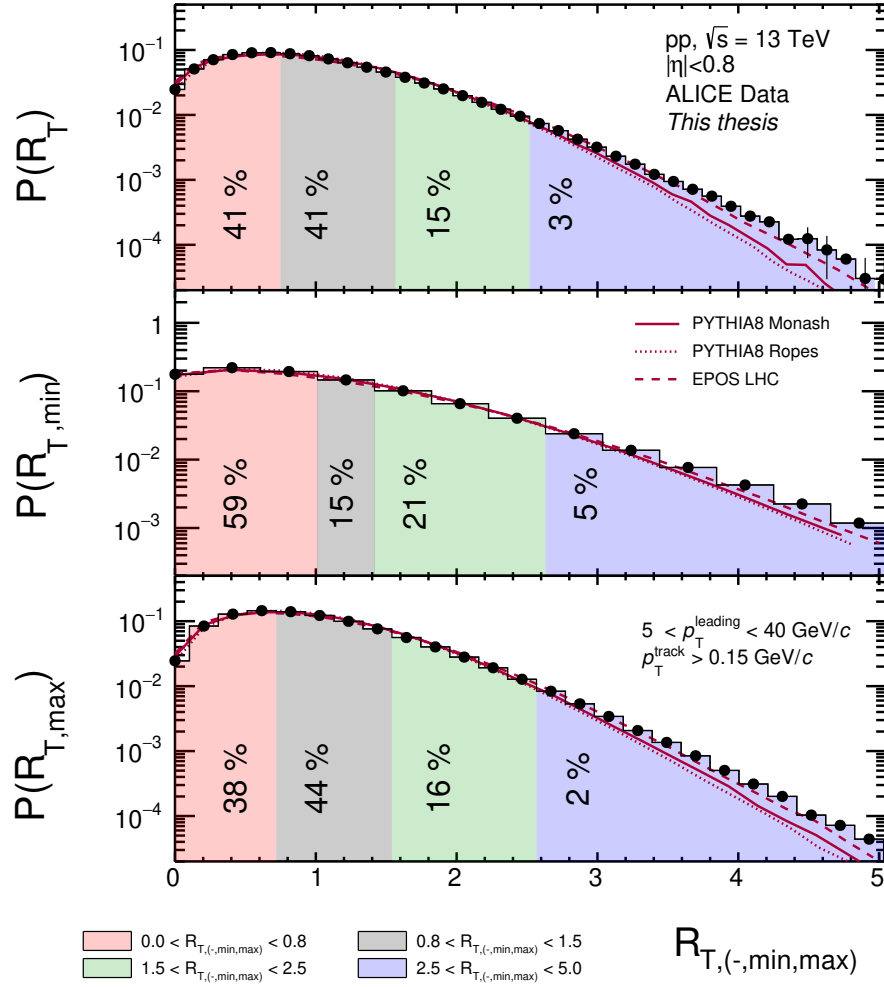
TBA

10.8 Baryon-to-meson ratio

TBA

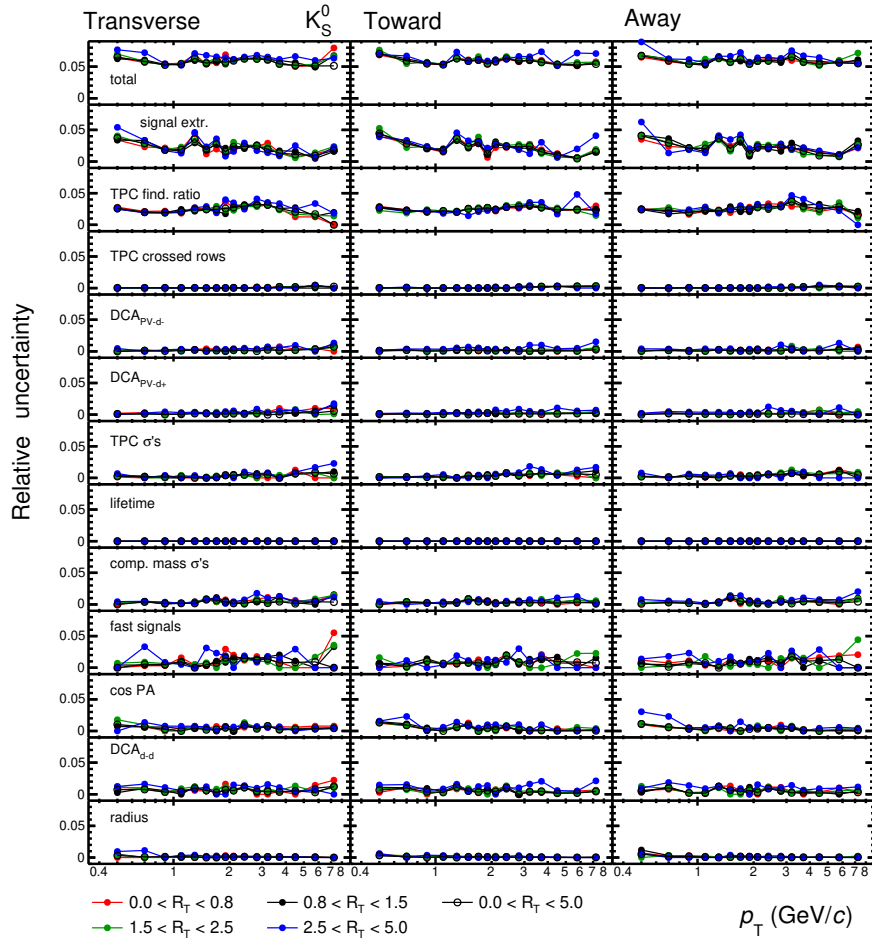
10.9 Integrated yields

TBA



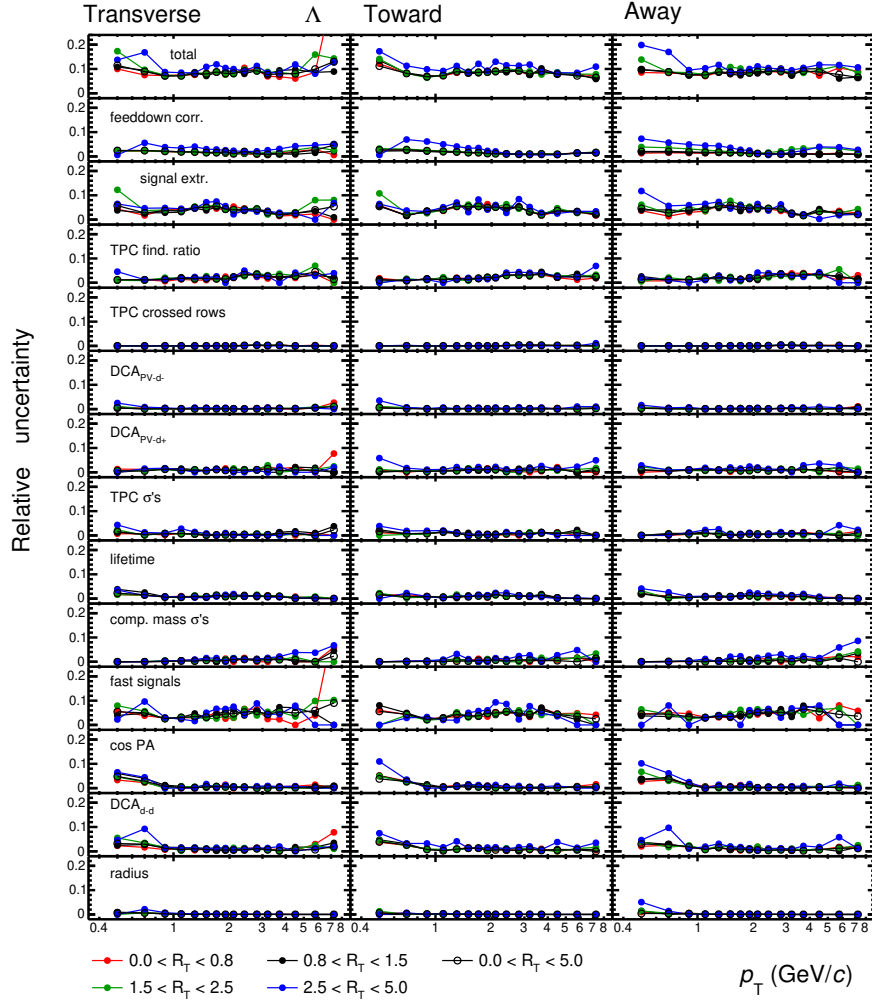
(a)

Figure 10.10: TBA



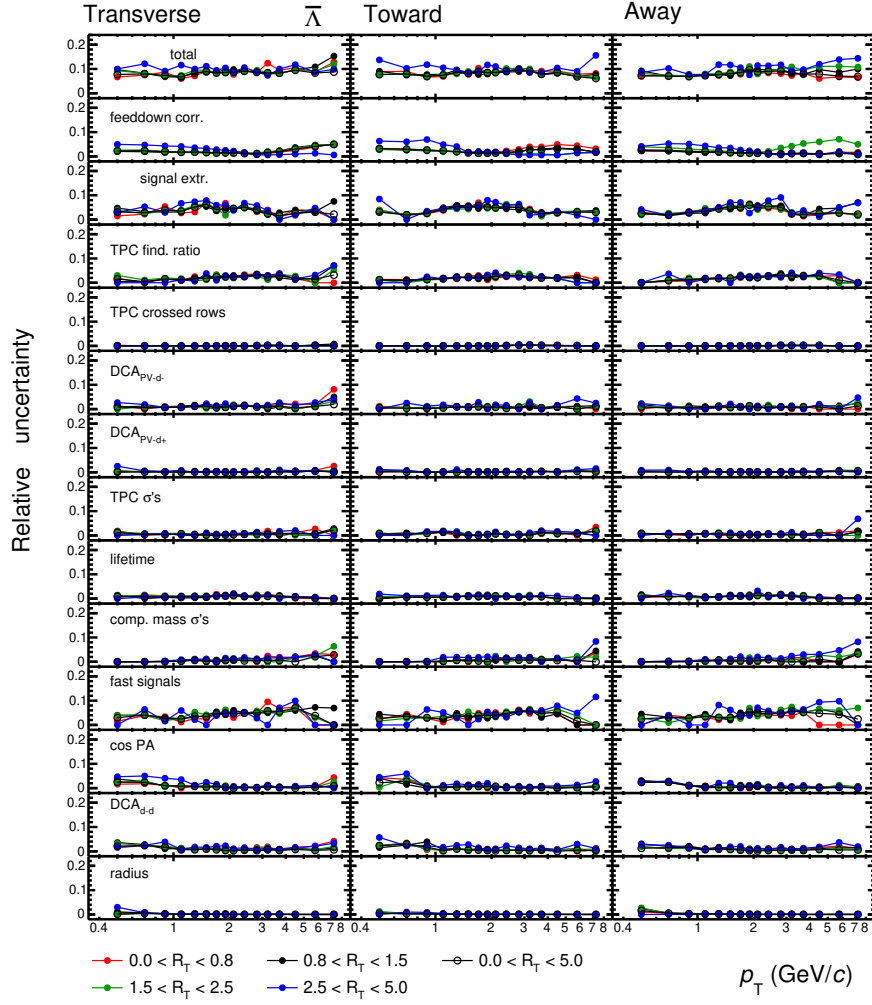
(a)

Figure 10.11: TBA.



(a)

Figure 10.12: TBA.



(a)

Figure 10.13: TBA.

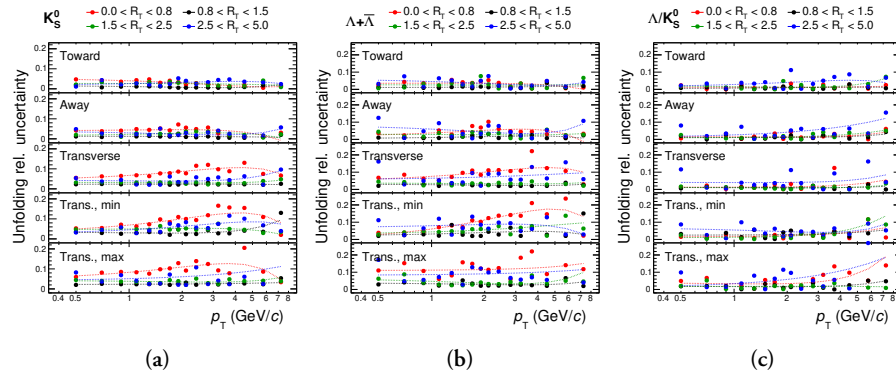
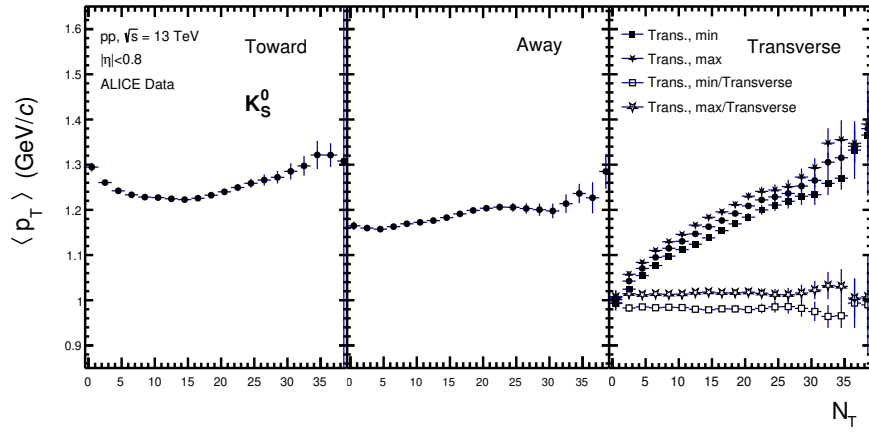
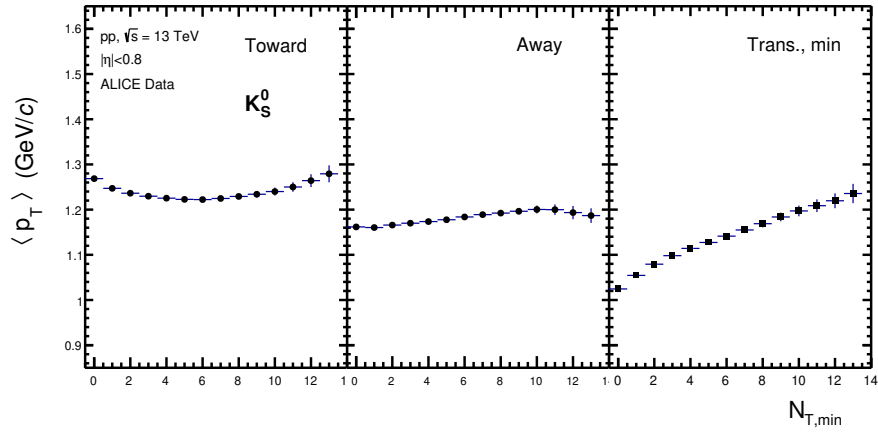


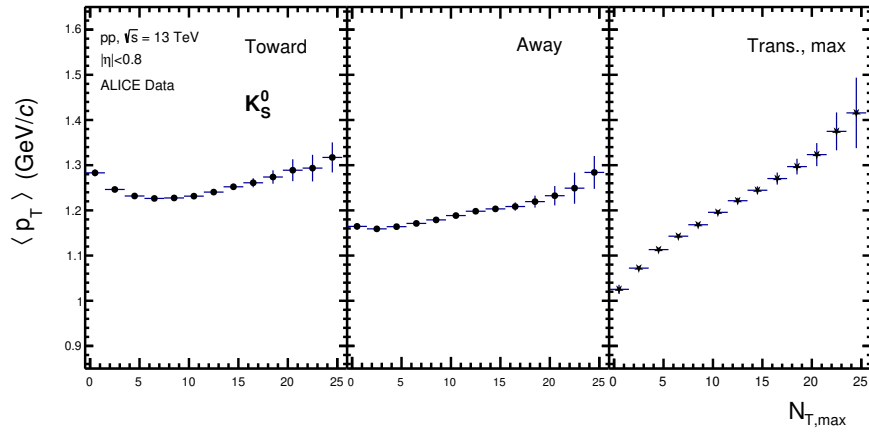
Figure 10.14: TBA.



(a)

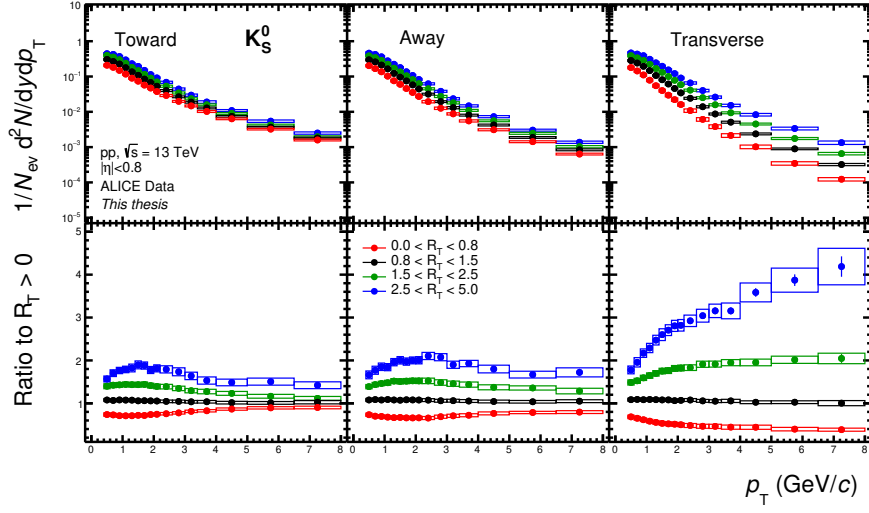


(b)

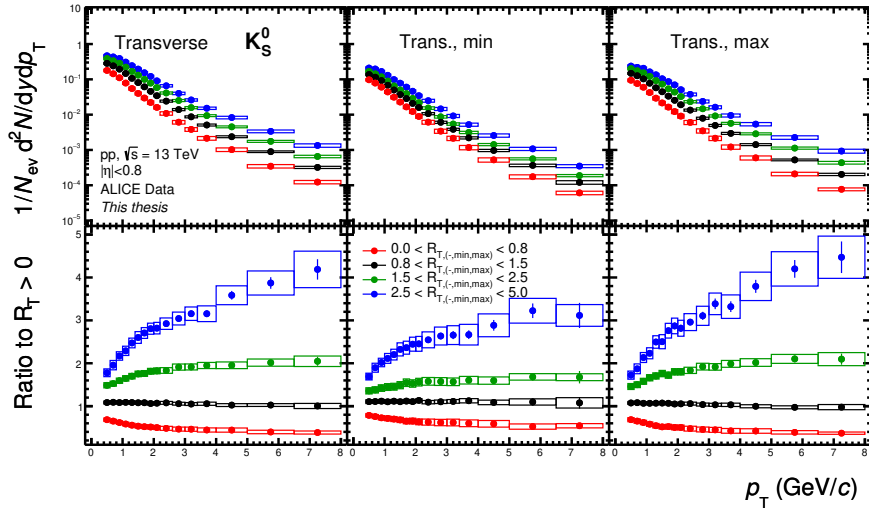


(c)

Figure 10.15: TBA.

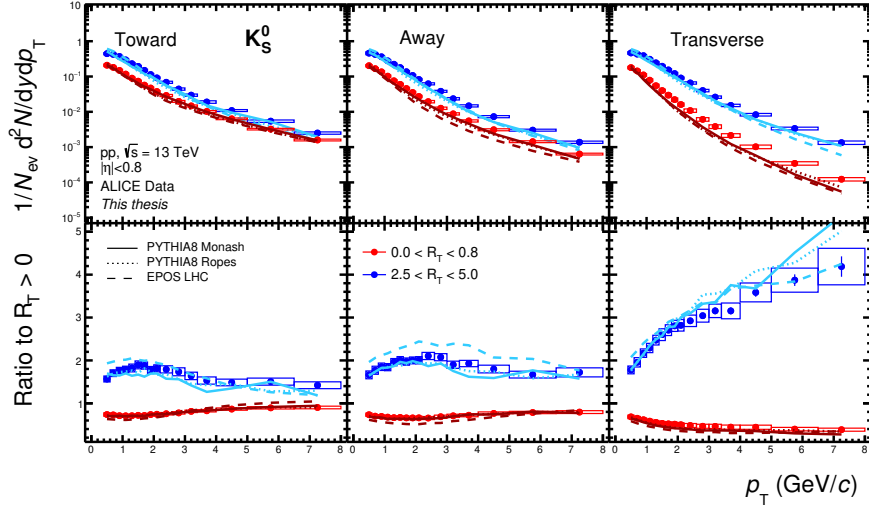


(a)

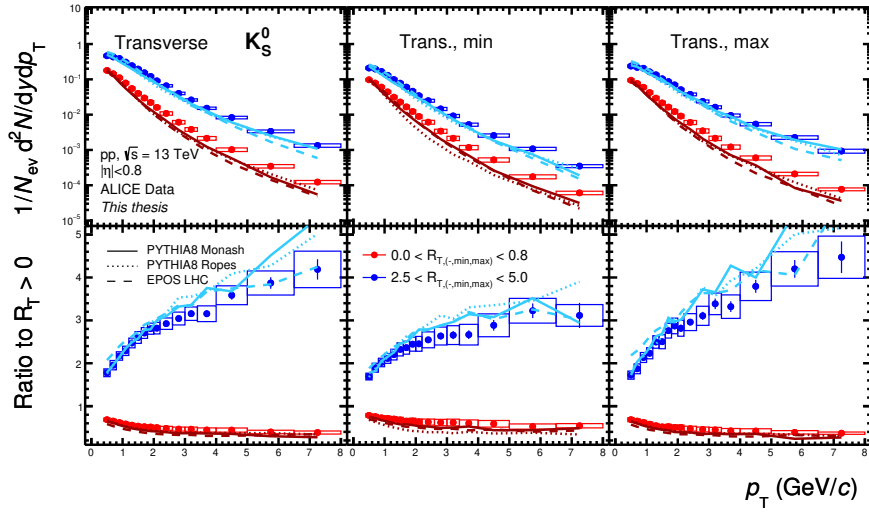


(b)

Figure 10.16: The measured and fully corrected $S_O^{(p_T=1.0)}$ distributions for both (a) $N_{\text{SPD}_{\text{TkIds}}}$ 0–1%, (b) 0–10% and (c) V0M 0–1% . The curves represent different model prediction, where the shaded area represents the statistical uncertainty of the models.

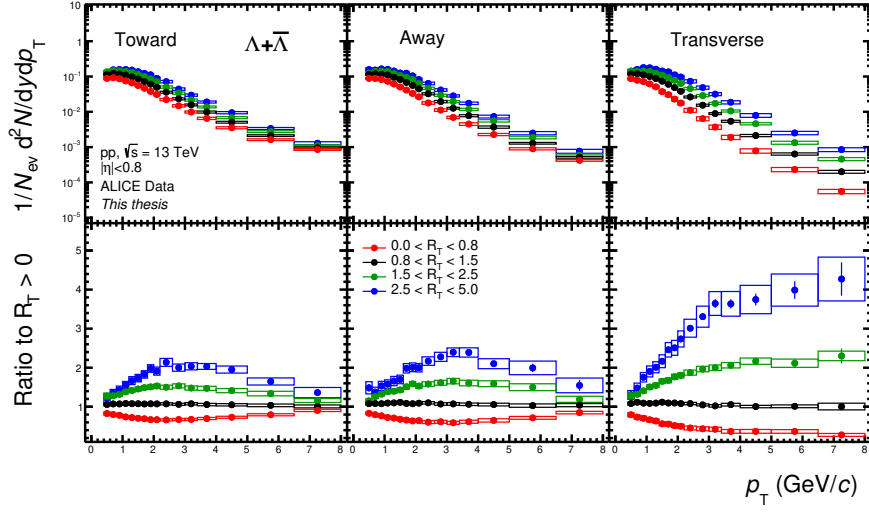


(a)

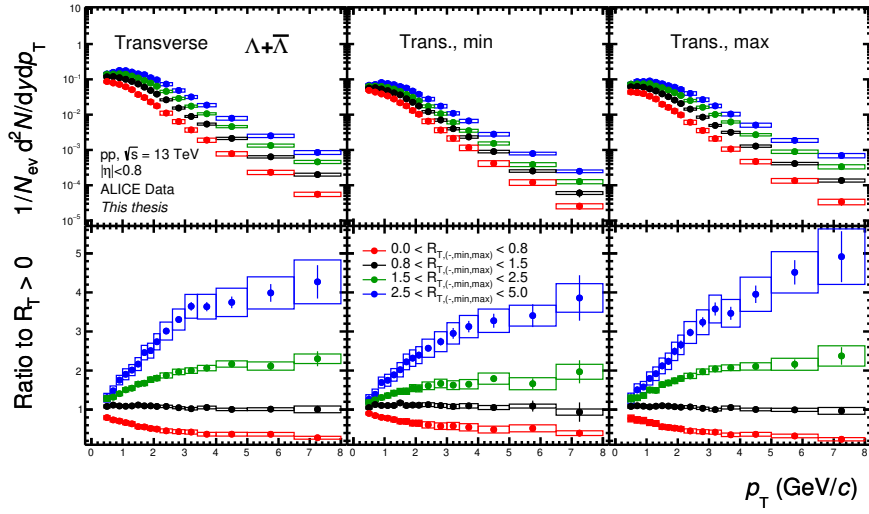


(b)

Figure 10.17: The measured and fully corrected $S_O^{(p_T=1.0)}$ distributions for both (a) $N_{\text{SPD}_{\text{TrkIds}}}$ 0–1%, (b) 0–10% and (c) V0M 0–1% . The curves represent different model prediction, where the shaded area represents the statistical uncertainty of the models.

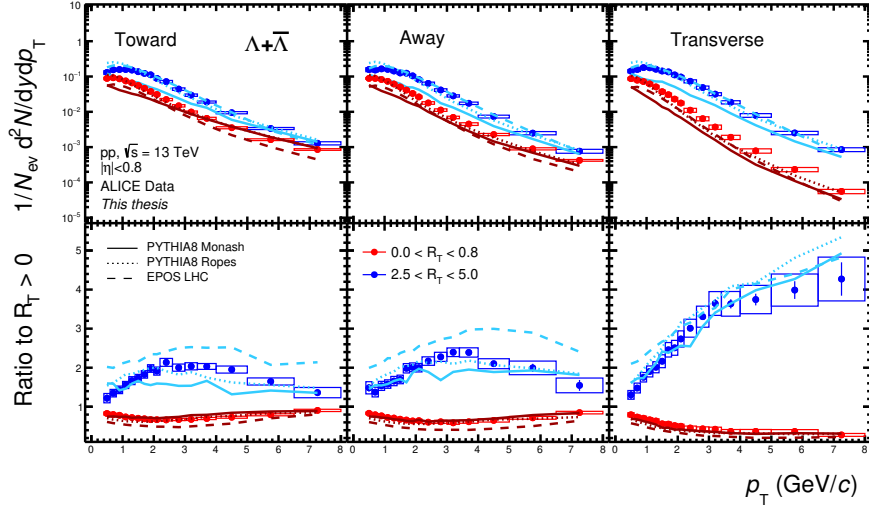


(a)

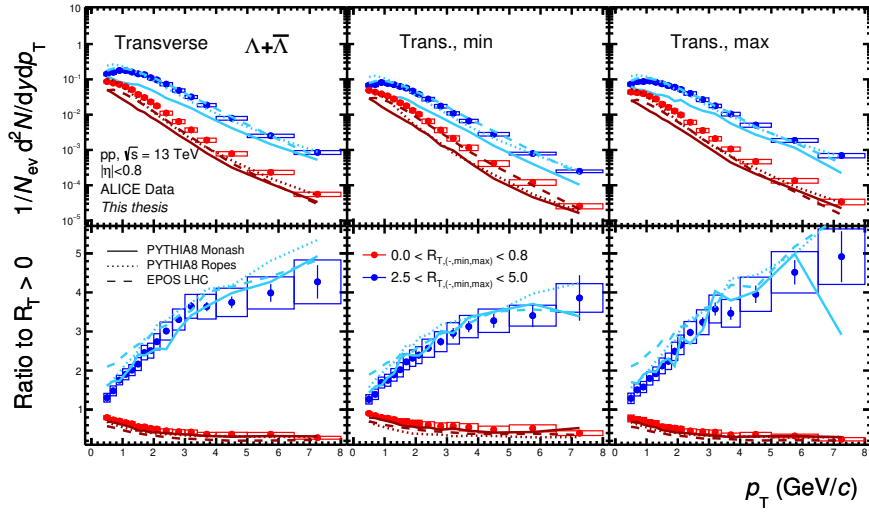


(b)

Figure 10.18: The measured and fully corrected $S_O^{(p_T=1.0)}$ distributions for both (a) $N_{\text{SPD}_{\text{TkIds}}}$ 0–1%, (b) 0–10% and (c) V0M 0–1% . The curves represent different model prediction, where the shaded area represents the statistical uncertainty of the models.

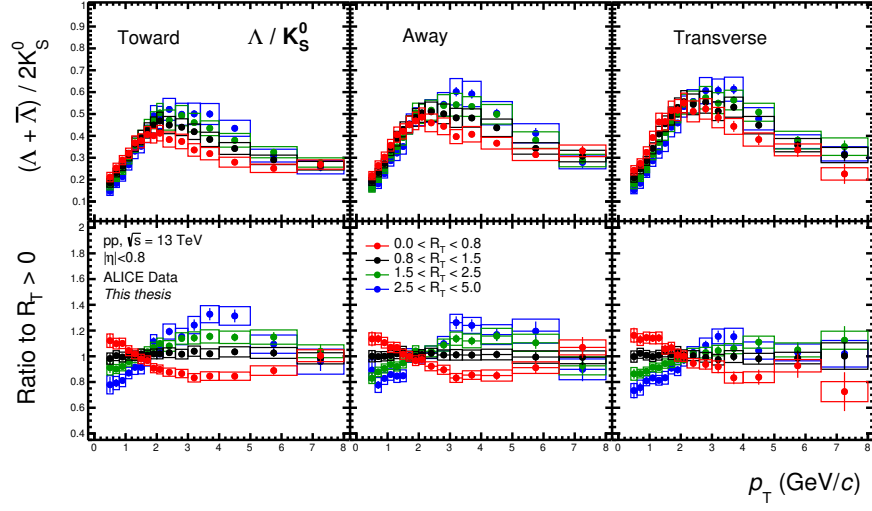


(a)

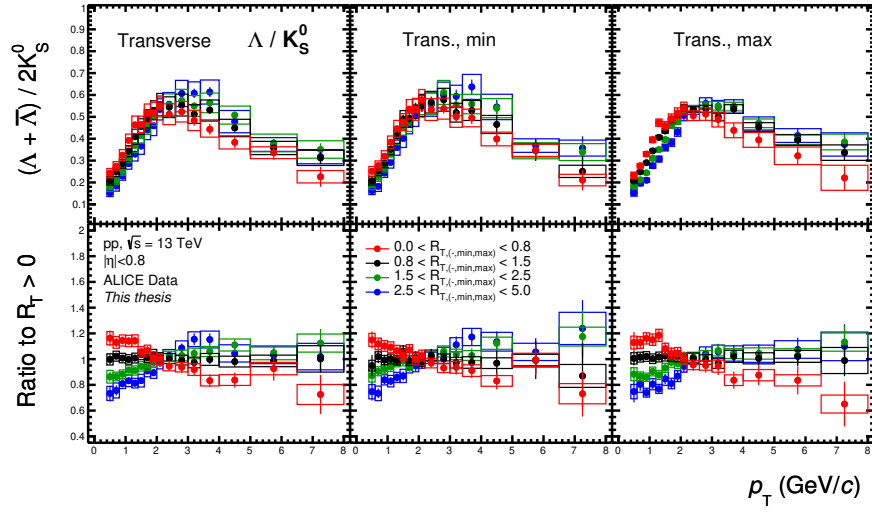


(b)

Figure 10.19: The measured and fully corrected $S_O^{(p_T=1.0)}$ distributions for both (a) $N_{\text{SPD-TkIds}}$ 0–1%, (b) 0–10% and (c) V0M 0–1% . The curves represent different model prediction, where the shaded area represents the statistical uncertainty of the models.

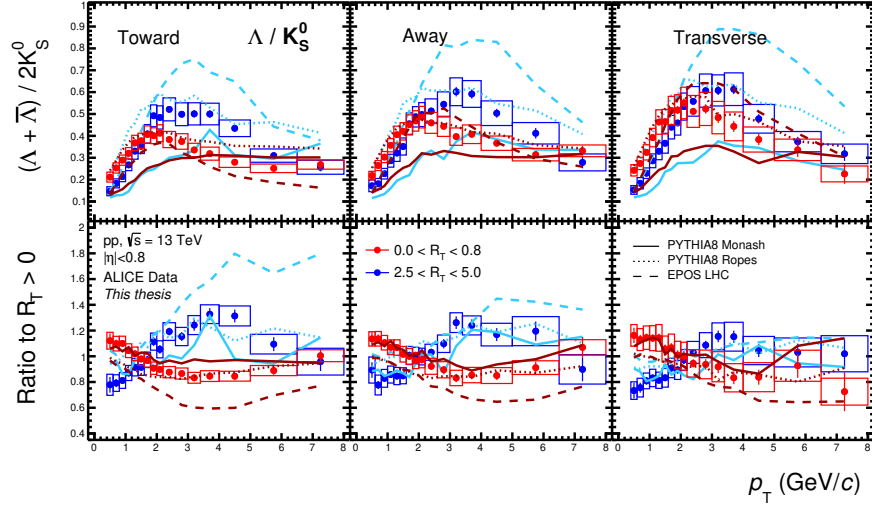


(a)

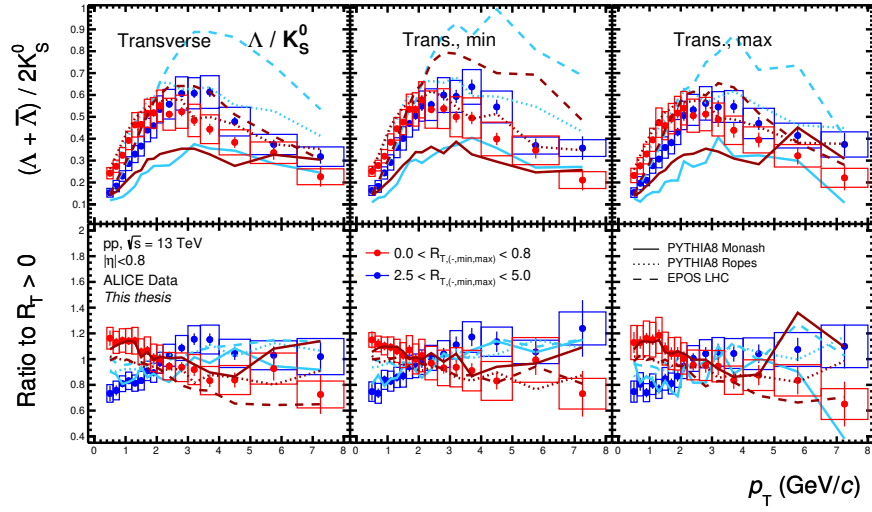


(b)

Figure 10.20: The measured and fully corrected $S_O^{(p_T=1.0)}$ distributions for both (a) $N_{\text{SPD}_{\text{Tidks}}}$ 0–1%, (b) 0–10% and (c) V0M 0–1% . The curves represent different model prediction, where the shaded area represents the statistical uncertainty of the models.

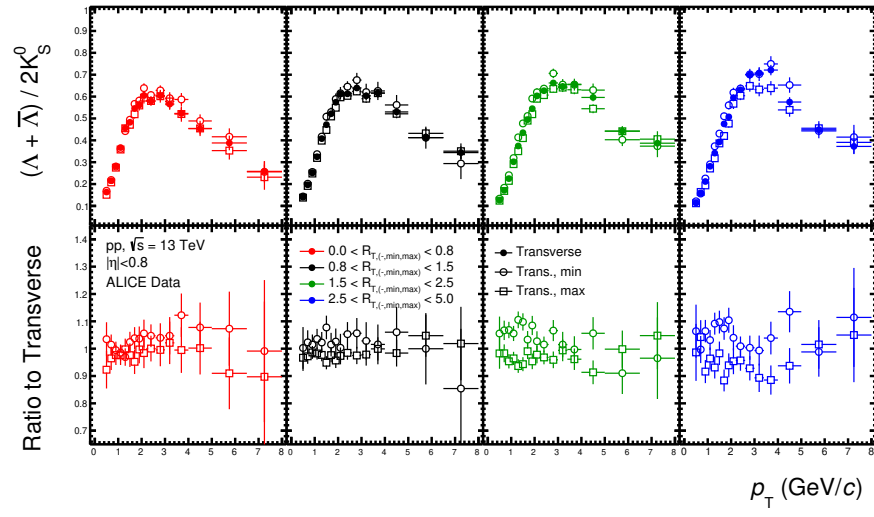


(a)



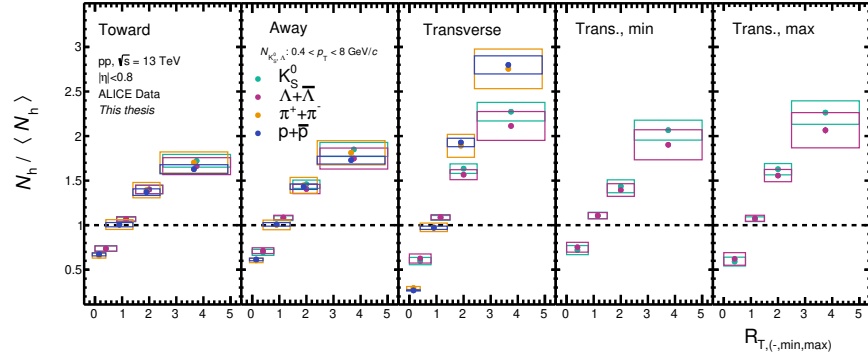
(b)

Figure 10.21: The measured and fully corrected $S_O^{(p_T=1.0)}$ distributions for both (a) $N_{\text{SPD-}T_{\text{Rdts}}}$ 0–1%, (b) 0–10% and (c) V0M 0–1% . The curves represent different model prediction, where the shaded area represents the statistical uncertainty of the models.

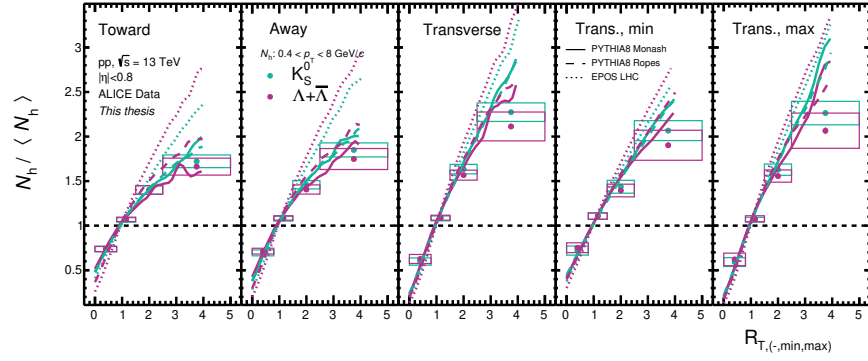


(a)

Figure 10.22: TBA.



(a)



(b)

Figure 10.23: The measured and fully corrected $S_0^{(p_T=1.0)}$ distributions for both (a) $N_{\text{SPD}_{\text{TkIds}}}$ 0–1%, (b) 0–10% and (c) VOM 0–1% . The curves represent different model prediction, where the shaded area represents the statistical uncertainty of the models.