

D^0 and D^* in $p + p$ at $\sqrt{s} = 500$ GeV Analysis Note

David Tlusty for the STAR collaboration, NPI ASCR, CTU Prague

August 23, 2017

Contents

1	Event Selection	4
2	Track Selection	7
2.1	Track Quality Cuts	7
2.2	Particle identification (PID)	10
2.2.1	D^0 PID Cuts	10
2.2.2	D^* PID Cuts	11
3	D^0 reconstruction	12
4	D^* reconstruction	17
5	Analysis of Simulated Data	18
5.1	TPC track reconstruction efficiency ε_R	21
5.2	Matching Efficiency, ε_M	22
5.3	Particle Identification Efficiency	28
6	Kinematical cuts efficiency	28
6.1	Trigger bias correction	30
7	Systematic Errors Evaluation	33
7.0.1	Double counting	34
7.1	Systematic Error Of Embedding	35
8	Results	37
A	Differential Invariant Cross Section	39
A.1	As a Hagedorn Shape Function	39
A.2	As a Levi Function	39
B	D^* Decay Kinematics	40
C	Monte-Carlo Toy Model	42
D	Treatment of Errors	44
D.1	Treatment of Errors in Efficiency Calculations	44
D.2	Treatment of Systematic Errors	45
D.3	Treatment of Correllated Errors	46

List of Figures

1	Number of events that passed cuts in the event chain	5
2	Correlation between z-position of the first TPC vertex and z-position of the VPD vertex from VPDMB trigger before (Upper Left Panel) and after (Upper Right Panel) 2x matching in Fast cut is applied. Lower Panel: Vertex z-position difference between the TPC vertex and the VPD vertex. The blue vertical lines indicate the cut window used to select the triggered events.	6
3	Charm quark fragmentation to D^0 and two main D^0 decay channels	7
4	Pion multiplicity as a function of the BBC Coincidence Rate for tracks with matching in fast detectors and for all TPC primary tracks	9
5	The scatter plot depicts $n\sigma_K^{1/\beta}$ as a function of the track momentum whose slices were projected along Y-axis and fitted by gaussian distributions. The red open circles represent second moments (σ) of those distributions that are fitted by power-law function f_{Res} and the dashed curve in the scatter plot is the power-law function f_{Pos} fitting the first moments (μ) of the gaussian distributions. The grey curves in the scatter plot represent PID cuts used to select kaons.	11
6	$n\sigma_\pi^{1/\beta}$ as a function of the track momentum. The black lines depict the TOF PID cut for pions. $1/\beta$ of pions is affected by coulomb re-scatterings at lower momenta so I used a wider range according to (2.4).	12
7	Upper Panel: Opposite-charged $K\pi$ invariant mass with the combinatorial background reconstructed by Like-Sign and Rotated Momentum techniques for all $K\pi$ pairs within $(1.0 < p_T < 2.0 \text{ GeV}/c)$ and $ y(K\pi) < 1$. The gray rectangle illustrates the zoom to D^0 mass window. Lower left panel: Opposite-charged $K\pi$ pairs invariant mass after Like-Sign background subtracted. Lower right panel: Opposite-charged $K\pi$ pairs invariant mass after Rotated-Momentum background subtracted.	14
8	Upper Panel: Opposite-charged $K\pi$ invariant mass with the combinatorial background reconstructed by Like-Sign and Rotated Momentum techniques for all $K\pi$ pairs within $(0.9 < p_T < 1.5 \text{ GeV}/c)$ and $ y(K\pi) < 1$. The gray rectangle illustrates the zoom to D^0 mass window. Lower left panel: Opposite-charged $K\pi$ pairs invariant mass after Like-Sign background subtracted. Lower right panel: Opposite-charged $K\pi$ pairs invariant mass after Rotated-Momentum background subtracted.	15
9	Upper Panel: Opposite-charged $K\pi$ invariant mass with the combinatorial background reconstructed by Like-Sign and Rotated Momentum techniques for all $K\pi$ pairs within $(1.5 < p_T < 2.0 \text{ GeV}/c)$ and $ y(K\pi) < 1$. The gray rectangle illustrates the zoom to D^0 mass window. Lower left panel: Opposite-charged $K\pi$ pairs invariant mass after Like-Sign background subtracted. Lower right panel: Opposite-charged $K\pi$ pairs invariant mass after Rotated-Momentum background subtracted.	16
10	$\cos(\theta^*)$ of the kaon in the CMS of the $K\pi$ pair	18
11	Upper panel: ΔM spectrum with combinatorial backgrounds reconstructed by Side-band and Wrong-sign technique. Green stars and the black line represent the ΔM with Side-band background subtracted and the gaussian fit to it respectively in the range from 0.14 to 0.16 GeV/c^2 . Lower panel: ΔM with Wrong-sign background subtracted and the gaussian fit to it in the range from 0.14 to 0.16 GeV/c^2 . The width of the gaussian function is fixed to be the same as in the case of the Side-band background subtracted.	19
12	ΔM spectra with combinatorial backgrounds reconstructed by Side-band and Wrong-sign techniques in $K^\mp\pi^\pm\pi^\pm$ p_T bins.	20

13	π^+, π^-, K^+, K^- track reconstruction efficiency. Contours reveal statistical uncertainties calculated as ??	21
14	TOF Matching efficiency as a function of the track p_T . Open red (closed blue) circles depict the efficiency calculated according to (5.3) from the production P11id for pions (kaons). The pp2pp_Production2009 trigger set was analyzed 1) by the same way with results depicted as green closed squares (kaons) and black thick dots (pions) 2) according to (5.4) with results depicted as dark red (kaons) and dark blue (pions) asterisks.	23
15	BEMC Matching efficiency as a function of the track p_T . Open (closed) black circles depict the efficiency calculated according to (5.3) from the production P11id for pions (kaons). The pp2pp_Production2009 trigger set was analyzed 1) by the same way with results depicted as red closed asterisks (kaons) and blue open asterisks (pions) 2) according to (5.4) with results depicted as brown closed asterisks (kaons) and green open stars (pions) asterisks.	24
16	The degree of correlation between phenomena T and B calculated according to (5.5). Red color is dedicated to pions while the black represents kaons.	25
17	The degree of correlation between phenomena B and T calculated according to (5.5). Red color is dedicated to pions while the black represents kaons.	25
18	The TOF and BEMC matching efficiency separately.	26
19	TOF and BEMC matching efficiency combined. Kaons with no TOF information and $p_T < 1.3 \text{ GeV}/c$ were rejected (see paragraph 2.2 , so that $\varepsilon_M = \varrho_T \varepsilon_T$ for kaons up to $p_T < 1.3 \text{ GeV}/c$	26
20	Daughter particle p_T resolution as a function of its p_T	28
21	Scatter plot of D^* transverse momentum t and reconstructed D^0 invariant mass M^{RC}	29
22	Left panel: The gaussian width of J/ψ peaks simulated by Monte-carlo Toy Model with $\rho(m, p_T)$ given by (6.1) and $A \times p_T$ where A was varied to found the minimum χ^2 , as shown in right panel. We used widths of real J/ψ peaks from analysis [4] as a benchmark. The widths of the real peaks are represented by red circles in left panel.	30
23	The mean (a) and width (b) of ΔM peak from Gaussian fit to minimum bias and BHT1 data (black and red circles respectively) compared with values from Toy Model simulations (green points).	31
24	Event counter for general events denoted as "MB" and events containing D^* mesons denoted as "Charmed". "fireBBC" and "FireVPD" columns represent events that initialized BBC and VPD trigger respectively. "good Vz" denotes events from which vertices with positive ranking were sucessfully reconstructed.	32
25	Trigger, Vertex reconstruction efficiency and Trigger Bias	33
26	Probability of a particle being in overlapping PID region	34
27	Probability of a particle being int overlapping PID region	35
28	Probability of a particle being in overlapping PID region	35
29	Systematic discrepancy between the data and the embedding	36
30	Kinematically possible values of the soft pion transverse momentum from the $D^* \rightarrow D^0 \pi_S$ decay	41
31	D^0 decay kinematics $g(t, p_T)$	42
32	D^* decay kinematics $g(t, p_T)$	43

The charm production is sensitive to early dynamics of the created system in RHIC heavy ion collisions. Dominant process of charm quark production at RHIC is believed to be initial gluon fusion which can be calculated in the perturbative QCD.

This text will present the first reconstruction of open charm meson D^0 via the weak decay to K and π mesons in the p+p collisions at midrapidity for $\sqrt{s} = 500$ GeV. The analysis is based on the large p+p minimum bias sample collected in RHIC year 2011 by the STAR detector. .

1 Event Selection

The data set used in this analysis was VPDMB trigger data (Trigger Id 320000, 320001, 320011, and 320021) taken in Run11 at $p + p \sqrt{s} = 500$ GeV. The VPDMB trigger was defined as $3944 < \text{VPD-TACdiff} < 4189$, $\text{VPD-EastADCsum-th} > 10$, and $\text{VPD-WestADCsum-th} > 10$ made Condition 1 to accept the an event. The FileCatalog options in getting the files for analysis are listed in the following:

- production = P11id
- trgsetupname = pp500_production_2011_noeemc || pp500_production_2011 || pp500_production_2011_long
- filetype = daq_reco_MuDst
- filename~physics
- sanity=1

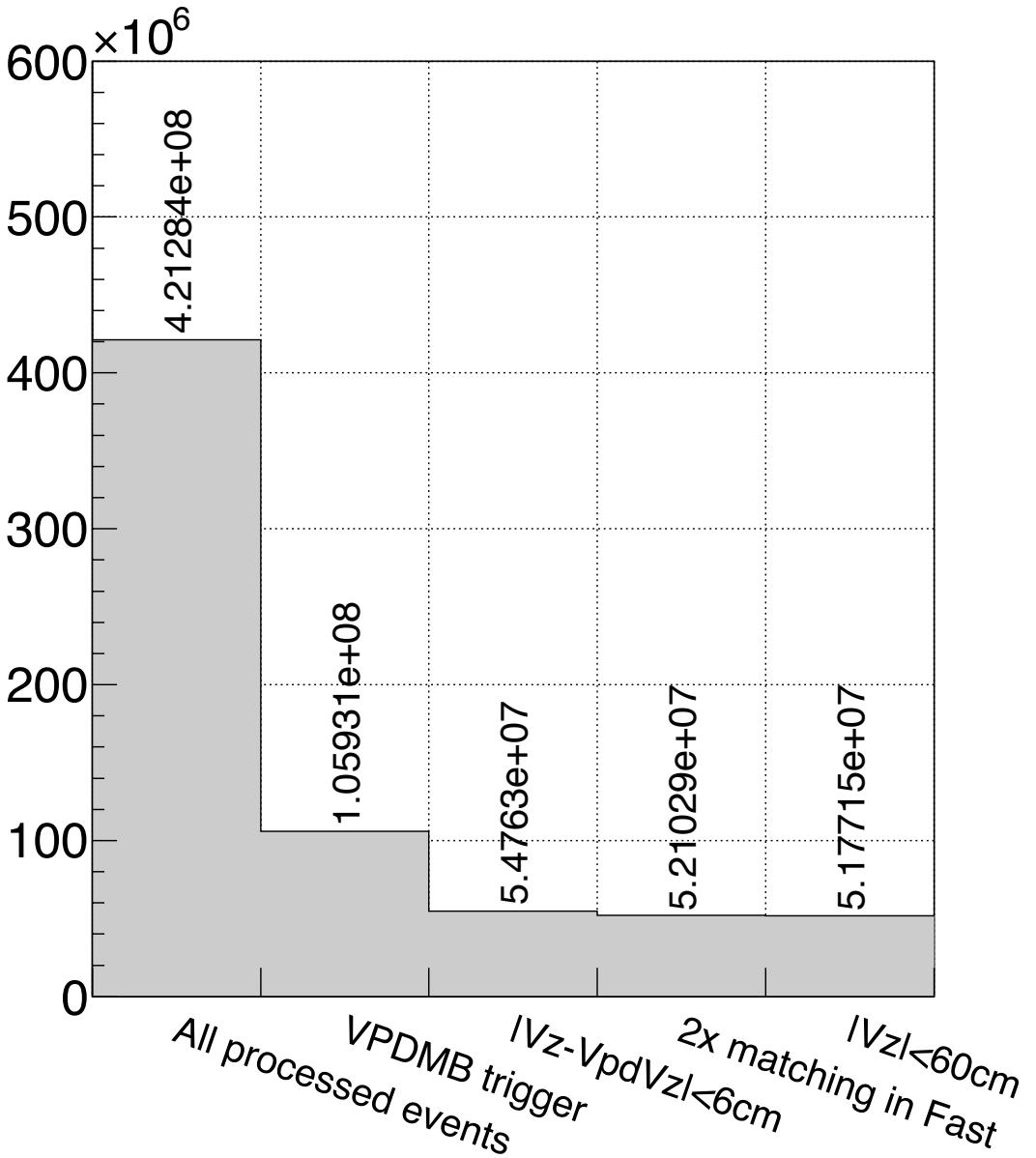


Figure 1: Number of events that passed cuts in the event chain

With these options I had 421M events that went into our analysis chain. Those events were processed through cuts that are listed below and their impact on statistics illustrated in Figure 1.

1. Events which passed the VPDMB trigger.
2. $|Vz_{VPD} - Vz[0]_{TPC}| < 6\text{ cm}$, where Vz_{VPD} is the z-position of vertex given by VPD and $Vz[0]_{TPC}$ is the z-position of the first vertex from the PPV vertex finder. The value 6 cm is derived from the VPD time resolution $\sim 80\text{ ps}$ giving a speed of light particle 2.4 cm track in one direction.
3. 2x matching in the "Fast" detectors (BEMC or TOF). The most of pile-up tracks were removed by the previous cut (see Figure 1), but there was a certain probability that some pile-up events would occur in the $|Vz_{VPD} - Vz[0]_{TPC}|$ during the one TPC readout; on the contrary, there's much lower probability that a pile-up would have any track with matching in BEMC or TOF. This is the same way as the PPV's Vertex Ranking is calculated, but PPV has not been implemented with TOF information in Run11 yet.

4. $|Vz[0]_{\text{TPC}}| < 60 \text{ cm}$

After those cuts applied, we had $N_{\text{evt}} = 51.2\text{M}$ of "good" events ready to be analyzed. To illustrate the impact of those cuts, Figure 2 depicts correlation between Vz_{VPD} and $Vz[0]_{\text{TPC}}$ as well as $|Vz_{\text{VPD}} - Vz[0]_{\text{TPC}}|$ both before and after the 2x matching in Fast is applied. One can observe significant reduction of events whose vertex's z-positions provided by TPC and VPD are not in correlation whereas the impact on the peak around zero in $|Vz_{\text{VPD}} - Vz[0]_{\text{TPC}}|$ is much smaller.

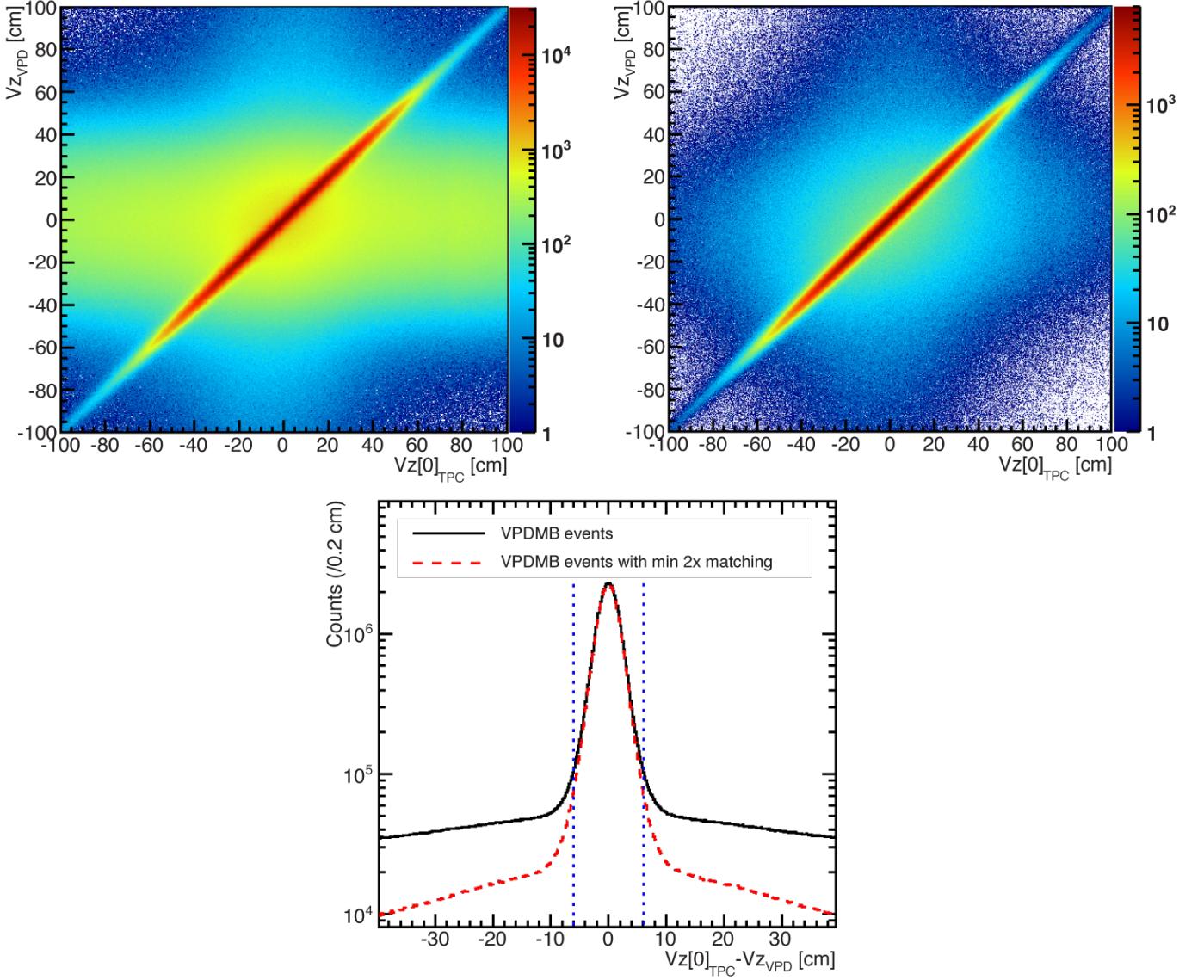


Figure 2: Correlation between z-position of the first TPC vertex and z-position of the VPD vertex from VPDMB trigger before (Upper Left Panel) and after (Upper Right Panel) 2x matching in Fast cut is applied. Lower Panel: Vertex z-position difference between the TPC vertex and the VPD vertex. The blue vertical lines indicate the cut window used to select the triggered events.

2 Track Selection

2.1 Track Quality Cuts

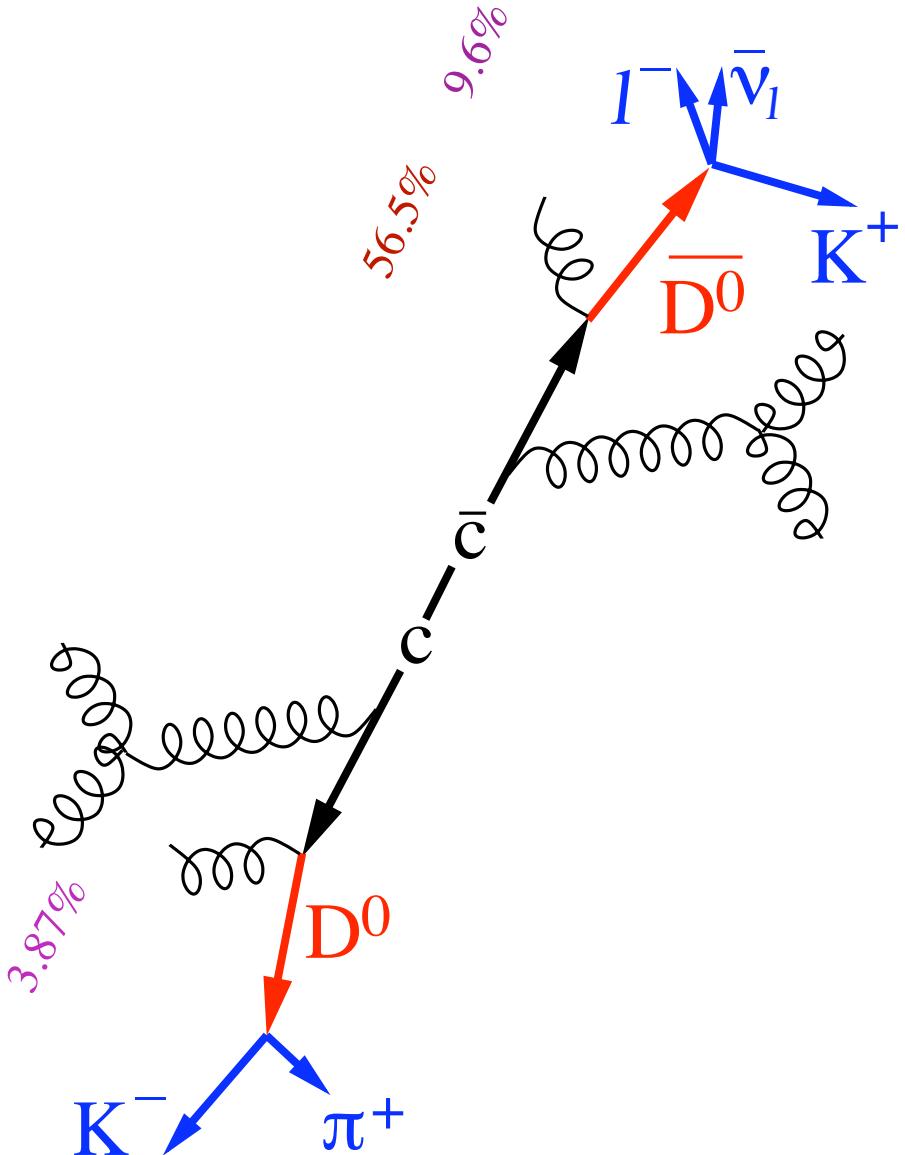


Figure 3: Charm quark fragmentation to D^0 and two main D^0 decay channels

This analysis is based on the direct D^0 invariant mass reconstruction via hadronic decay channel (see Figure 3) $D^0 \rightarrow K^-\pi^+$ with Branching ratio $\Gamma = 3.89\%$ [2]. To reconstruct D^* , one may exploit the softness of $D^{\pm*} \rightarrow D^0\pi_S^\pm$ decay ($p_{\pi_S}^* = 39.4$ MeV); combine pions with D^0 candidates ($K\pi$ pairs with $M(K\pi) \sim 1.865$ GeV/c 2 , and plot the difference $M(K\pi\pi_S) - M(K\pi)$ whose resolution is determined by mostly the π_S momentum resolution.

Daughter particles were "good" TPC tracks identified through dE/dx provided by TPC and β provided by TOF. The "good" tracks are TPC primary tracks passed the tracks quality cuts listed below:

1. $0 < \text{StMuTrack->flag()} < 1000$,
2. number of TPC fit points > 20

3. $\frac{\text{number of TPC fit points}}{\text{number of max possible TPC fit points}} > 0.52$
4. Global DCA < 2cm
5. $p_T > 0.16 \text{ GeV}/c$
6. $|\eta| < 1$

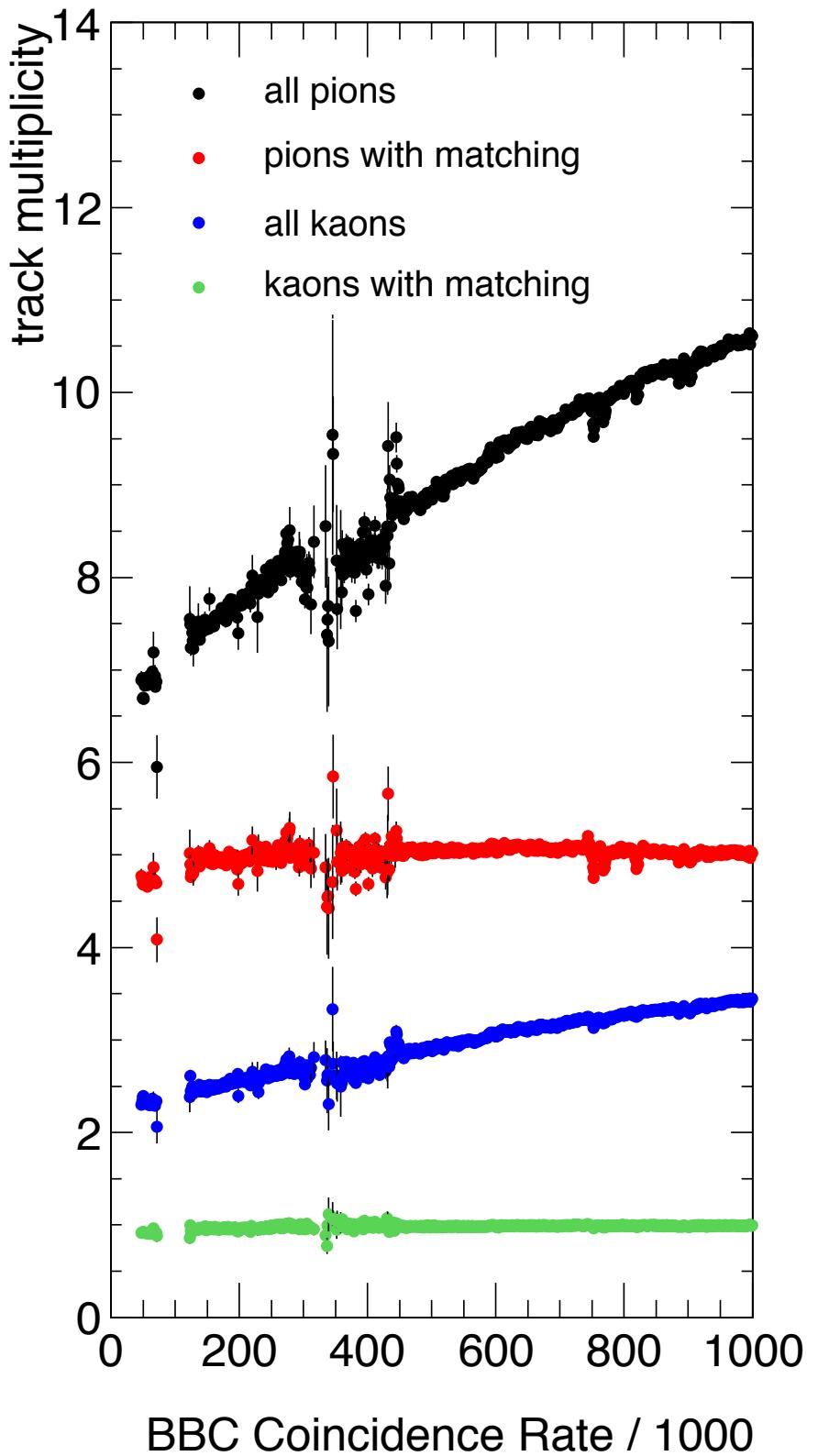


Figure 4: Pion multiplicity as a function of the BBC Coincidence Rate for tracks with matching in fast detectors and for all TPC primary tracks

The curvature radius R of a track that is able to reach TOF must be at least 1.05m. That corresponds to the track minimum $p_T = BRc = 0.158$ GeV/c, where B is the magnetic field and c the speed of light. All tracks with $p_T < 0.16$ GeV/c are useless for TOF as well as all tracks with $p_T < 0.18$ GeV/c for BEMC.

The typical BBCMB trigger scaler rate was 1.6 MHz which suggests at least 130 events in one TPC readout. So that some tracks might be assigned to a vertex originating from a different event. These tracks (let's call them pile-up tracks) would have a different timing and no matching in fast detectors as BEMC or TOF. To check the fraction of those tracks, I analyzed pion multiplicity as a function of the BBC coincidence (close and proportional to the event rate). As one can see in Figure 4, the track multiplicity is proportional to the event rate if the matching is not required and doesn't correspond to Physics. Yet the multiplicity of the tracks with matchings seems to be independent on the event rate. Regarding the matching efficiency, discussed in paragraph ?? (86% for pions and 83% for kaons), I can estimate that almost each second pion and 60% of kaons are pile-ups causing potentially 5x higher combinatorial background in D^0 and 10x higher combinatorial background in D^* analysis. However, the matching requirement made the D^* reconstruction efficiency dwindling at $D^* p_T < 3 \text{ GeV}/c$ so it's very important to set the track low p_T cut as low as possible together with the matching requirement (see paragraph B for details).

2.2 Particle identification (PID)

Particle identification (PID) for tracks was carried out with a combination of TPC dE/dx and the particle velocity β measurement from the barrel TOF detector.

Thus the normalized $dE/dx \left(n\sigma_X^{dE/dx} \right)$ and $1/\beta \left(n\sigma_X^{1/\beta} \right)$ were used to select daughter particle candidates. They are defined as follows:

$$n\sigma_X^{dE/dx} = \frac{\ln \frac{dE/dx}{\langle dE/dx \rangle_X}}{R^{dE/dx}}, \quad (2.1)$$

$$n\sigma_X^{1/\beta} = \frac{\frac{1}{\beta} - \sqrt{\frac{m_X^2}{p_X^2} + 1}}{R^{1/\beta}}, \quad (2.2)$$

where the X denotes expected values which are calculated with respect to one kind of particle species (π or K). The $\langle dE/dx \rangle_X$ is mean value of ionization loss in TPC given by the Bichsel function [1], $R^{dE/dx}$ and $R^{1/\beta}$ are corresponding resolutions. Different PID cuts were chosen for D^0 and D^* analysis.

2.2.1 D^0 PID Cuts

If a track had matching to TOF the $n\sigma_K^{1/\beta}$ and $n\sigma_\pi^{1/\beta}$ were calculated. At low track momentum (up to $0.4 \text{ GeV}/c$), the $R_K^{1/\beta}$ becomes appreciably larger due to Coulomb re-scatterings of the particle in TPC volume. The re-scatterings make Kaon fly along a longer path than the one parametrized by an ideal helix (which is used to calculate the path length) and reach the TOF later than expected, so $1/\beta$ is larger than expected. This phenomenon is clearly visible in the scatter plot of the Figure 5, where one can see not only a larger $1/\beta$ but also its worse resolution. Hence I decided to use a dynamic PID cut based on $n\sigma_K^{1/\beta}$ resolution which is depicted by the red open circles in Figure 5. Both $n\sigma_K^{1/\beta}$ resolution and position as a function of the kaon momentum is parametrized by power-law type functions marked as f_{Res} and f_{Pos} respectively. Kaons are then defined as tracks that fulfill the condition:

$$-2f_{\text{Res}} + f_{\text{Pos}} < n\sigma_K^{1/\beta} < \begin{cases} 2f_{\text{Res}} + f_{\text{Pos}}, & \text{in } D^0 \text{ analysis} \\ 3f_{\text{Res}} + f_{\text{Pos}}, & \text{in } D^* \text{ analysis} \end{cases} \quad (2.3)$$

$$\text{where } f_{\text{Res}} = 0.884 + \frac{0.0174}{(p + 0.0839)^{4.23}} \text{ and } f_{\text{Pos}} = 0.0316 + \frac{0.00137}{(p + 0.101)^{6.89}}.$$

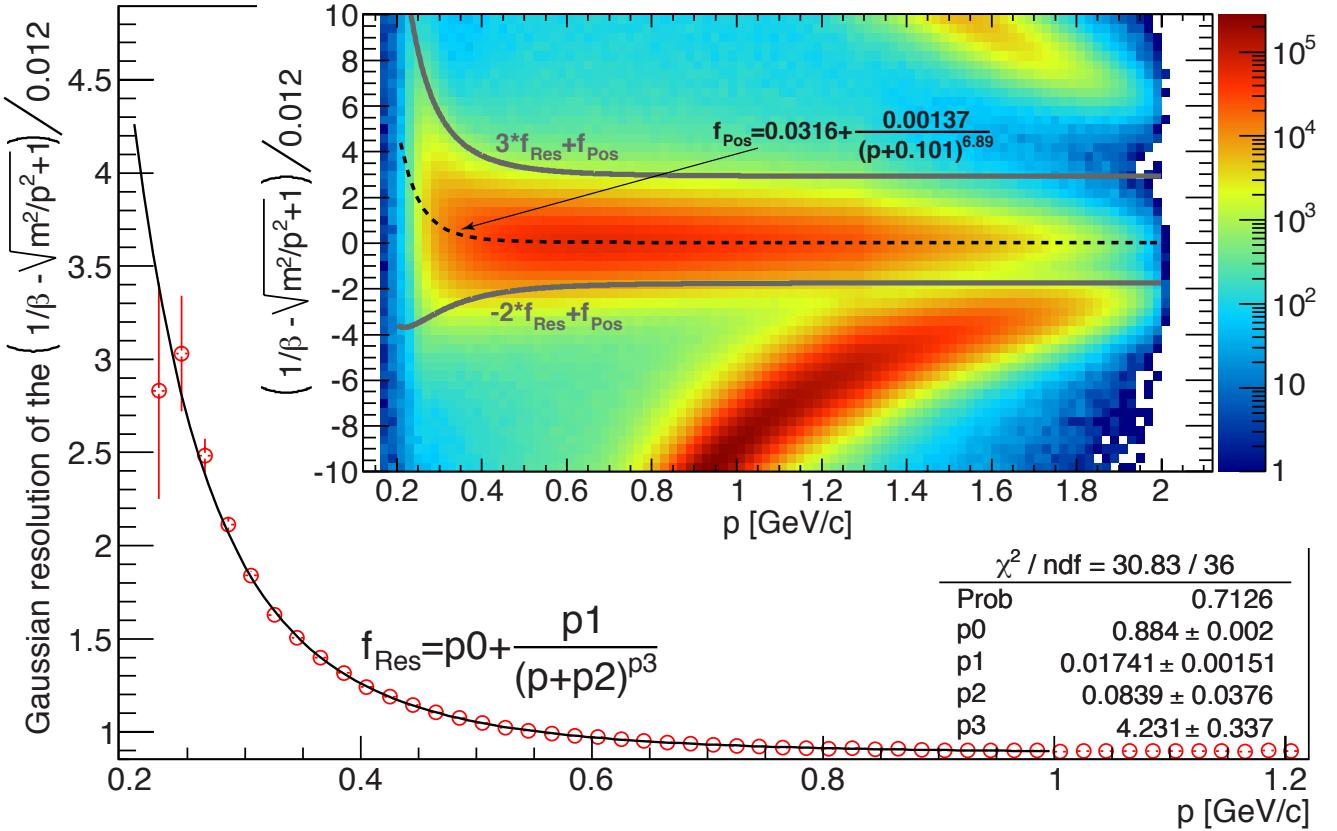


Figure 5: The scatter plot depicts $n\sigma_K^{1/\beta}$ as a function of the track momentum whose slices were projected along Y-axis and fitted by gaussian distributions. The red open circles represent second moments (σ) of those distributions that are fitted by power-law function f_{Res} and the dashed curve in the scatter plot is the power-law function f_{Pos} fitting the first moments (μ) of the gaussian distributions. The grey curves in the scatter plot represent PID cuts used to select kaons.

The $1/\beta$ of pions is much smaller at low momentum and hence relatively less affected by coulomb re-scatterings than this of kaons (the mean of the $n\sigma_\pi^{1/\beta}$ distribution is 0.1 with the $R_\pi^{1/\beta} = 0.011$, hence pions are defined as tracks that fulfill the condition:

$$-1.9 < n\sigma_\pi^{1/\beta} < 2.1$$

To enhance statistics, if a track with $p_T > 1.6$ GeV/c hadn't matching to TOF, but had matching to BEMC, I didn't reject it. Those tracks were identified as

- kaons if $-2 < n\sigma_K^{\text{d}E/\text{d}x} < 2$
- pions if $-2 < n\sigma_\pi^{\text{d}E/\text{d}x} < 2$

2.2.2 D^* PID Cuts

As shown in Appendix B, STAR can detect D^* with $p_T \gtrsim 2$ which decays into daughters whose momenta might be already beyond the region where TOF provides the best and unequivocal PID and such daughter particles need to be identified by TPC $\text{d}E/\text{d}x$.

Therefore the tracks were separated into three groups according to their momentum (each detector provides the best separation of kaons from pions at different track momentum):

- those with $p_T < 1.3$ GeV/c: TOF provides clear separation of kaons from pions so it was main PID tool in this track region. Kaons were defined as tracks that fulfilled the condition (2.3) and pions as those fulfilling the condition

$$|n\sigma_\pi^{1/\beta}| < \begin{cases} 6 - 2p, & \text{if } p < 1.5 \text{ GeV/c} \\ 3, & \text{if } p > 1.5 \text{ GeV/c} \end{cases} \quad \text{see Figure 6 depicting the cut} \quad (2.4)$$

which is designed to make sure all pions went into the analysis. I didn't reject tracks that had matching in BEMC only. If those passed $-3 < n\sigma_\pi^{dE/dx} < 3$ I counted them as pions.

- those with $1.3 < p_T < 2.07$ GeV/c: If a track had matching in TOF then TOF was used so Kaons were defined as tracks that fulfilled the condition (2.3) and pions as those fulfilling the $|n\sigma_\pi^{1/\beta}| < 3$. If a track had matching in BEMC and no matching in TOF then kaons were those passed $|n\sigma_K^{dE/dx}| < 2$ and pions $|n\sigma_\pi^{dE/dx}| < 3$.
- those with $p_T > 2.07$ GeV/c: TPC provides better, even if not fully clear, separation of kaons from pions. If a track had matching in any fast detector (BEMC or TOF) then kaons were those passed $|n\sigma_K^{dE/dx}| < 2$ and pions $|n\sigma_\pi^{dE/dx}| < 3$. There is a contribution from protons that might affect the raw yield of kaons (pion yield is less affected since pions dominate the particle production) so I didn't count the tracks that had matching in TOF and $n\sigma_K^{1/\beta} > 3$ as kaons.

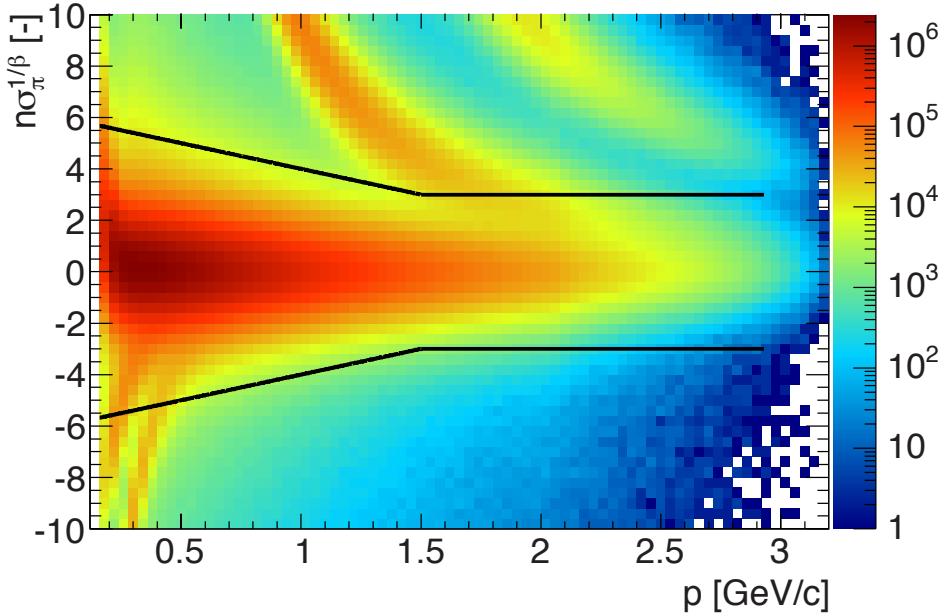


Figure 6: $n\sigma_\pi^{1/\beta}$ as a function of the track momentum. The black lines depict the TOF PID cut for pions. $1/\beta$ of pions is affected by coulomb re-scatterings at lower momenta so I used a wider range according to (2.4).

3 D^0 reconstruction

A combination of D^0 and $\overline{D^0}$ yield was analyzed ($M_{K-\pi^+} + M_{K^+\pi^-}$), because there has been not enough statistics to analyze the yields separately. Each kaon was combined with each pion of opposite

charge to form a D^0 candidate and its invariant mass was calculated. The spectrum of the invariant mass is shown in Figures 7, 8, and 9 for pairs with p_T between 1.0 and 2.0 GeV/c , 0.9 and 1.5 GeV/c , and 1.5 and 2.0 GeV/c respectively ; Let's note that the candidates whose rapidity exceeded the (-1,1) interval were rejected.

The combinatorial background, that constitutes the dominant part of the D^0 candidates invariant mass spectrum, was reconstructed by two independent techniques:

- **Like-Sign:** Pions are paired with the Kaons of the same charge. Then the geometric mean of the two subsets (the raw yield of positively $RY_{K^+\pi^+}$ and negatively $RY_{K^-\pi^-}$ charged pairs) is calculated by $2\sqrt{RY_{K^-\pi^-}RY_{K^+\pi^+}}$.
- **Rotated Momentum:** Each pion is paired with kaon with reversed 3-momenta. Track rotation technique is based on the assumption that by rotating one of the daughter track for 180 degree the decay kinematics is destroyed. Thus the distribution of a pair invariant mass with one track rotated is able to reproduce the random combinatorial background.

The combinatorial background reconstructed by either Like-Sign or Rotated Momentum technique was scaled to match the original Unlike-Sign spectrum of $K\pi$ pairs within the interval of the invariant mass to be between 1.7 and 1.8 GeV/c^2 . Figures 7, 8, and 9 displays excellent agreement of Unlike-sign spectrum with the scaled combinatorial background. Such agreement allowed to declare that both methods describe combinatorial background well and the background could be subtracted from Unlike-sign spectrum to extract the raw yield of D^0 meson. Results of the background subtractions are shown also in Figures 7, 8, and 9. From these subtractions one can see two lorentzian peaks corresponding to particles $K^{0*}(892)$ and $K^{2*}(1430)$ and gaussian peak corresponding to D^0 . However, those peaks sit on some residual background which needs to be addressed in raw yield calculations.

The raw yield was calculated by fitting the D^0 mass peak by gaussian function sitting on linear function that should describe the residual background with adequate precision. The fitting method was based on minimizing of the χ^2 between the data and the function. The results of the fit are shown both in Figures 7, 8, and 9 and Table 1.

p_T bin [GeV/c]		$1 < p_T < 2$	$.9 < p_T < 1.5$	$1.5 < p_T < 2$
Rotated Momentum	χ^2/ndf	27.4/33	27.4/33	25.5/33
	Raw Yield	4584 ± 1016	2296 ± 819	2479 ± 697
	Mean [Mev/c^2]	1865 ± 3	1862 ± 5	1863 ± 3
	Resolution [Mev/c^2]	11.2 ± 2.5	12.0 ± 3.6	9.2 ± 2.4
	Significance	4.51	2.8	3.6
Like Sign	χ^2/ndf	17.6/33	20.9/33	20.7/33
	Raw Yield	3564 ± 996	1947 ± 835	1711 ± 690
	Mean [Mev/c^2]	1866 ± 3	1867 ± 4	1863 ± 5
	Resolution [Mev/c^2]	11.0 ± 2.9	10.1 ± 4.6	10.7 ± 3.4
	Significance	3.58	2.33	2.5
Raw Yield Weighted Mean		4074 ± 711	2122 ± 585	2095 ± 490

Table 1: Results in p_T bins. If the $\chi^2/\text{n.d.f.}$ is greater than 1 the contribution to averaged raw yield is scaled by.

Simulation based on events reconstructed by GEANT from Monte-Carlo Pythia collisions show the width of the D^0 invariant mass peak to be 10 MeV/c^2 and the position at 1865 MeV/c^2 (this is discussed in paragraph ??). Results displayed in Table 1 are consistent with the results from the simulation. The statistical test of two hypotheses

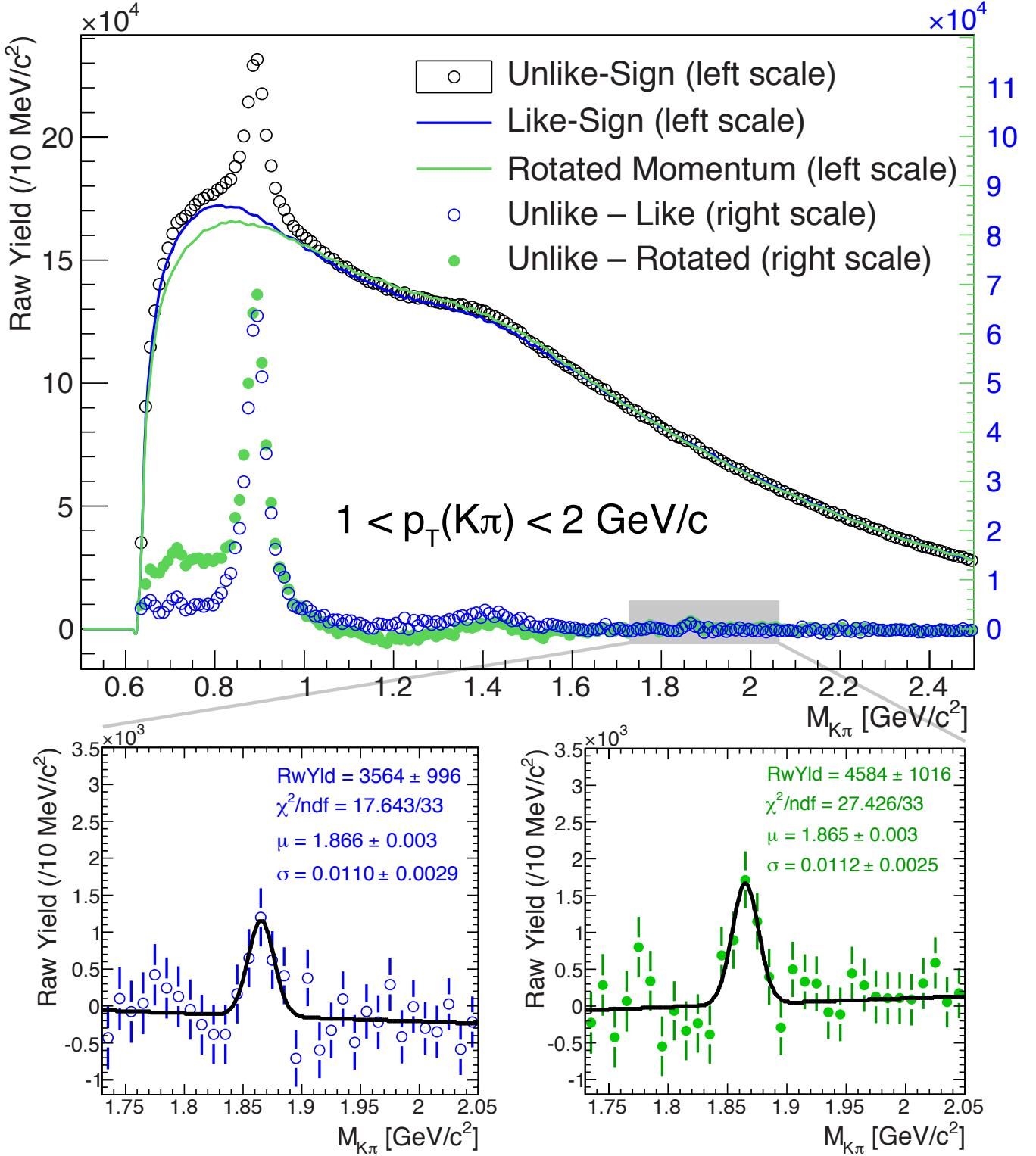


Figure 7: Upper Panel: Opposite-charged $K\pi$ invariant mass with the combinatorial background reconstructed by Like-Sign and Rotated Momentum techniques for all $K\pi$ pairs within $(1.0 < p_T < 2.0 \text{ GeV}/c)$ and $|y(K\pi)| < 1$. The gray rectangle illustrates the zoom to D^0 mass window. Lower left panel: Opposite-charged $K\pi$ pairs invariant mass after Like-Sign background subtracted. Lower right panel: Opposite-charged $K\pi$ pairs invariant mass after Rotated-Momentum background subtracted.

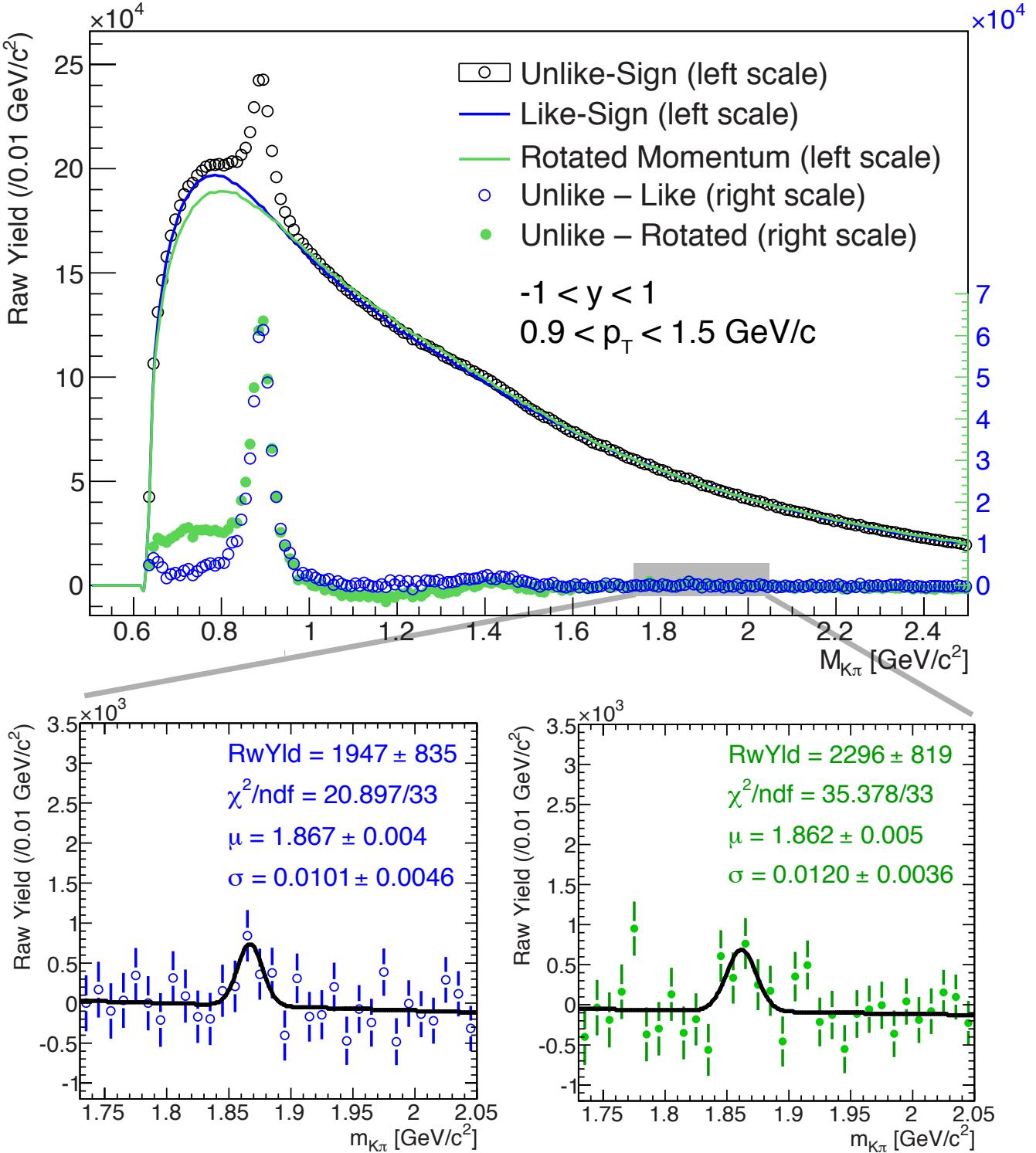


Figure 8: Upper Panel: Opposite-charged $K\pi$ invariant mass with the combinatorial background reconstructed by Like-Sign and Rotated Momentum techniques for all $K\pi$ pairs within $(0.9 < p_T < 1.5 \text{ GeV}/c)$ and $|y(K\pi)| < 1$. The gray rectangle illustrates the zoom to D^0 mass window. Lower left panel: Opposite-charged $K\pi$ pairs invariant mass after Like-Sign background subtracted. Lower right panel: Opposite-charged $K\pi$ pairs invariant mass after Rotated-Momentum background subtracted.

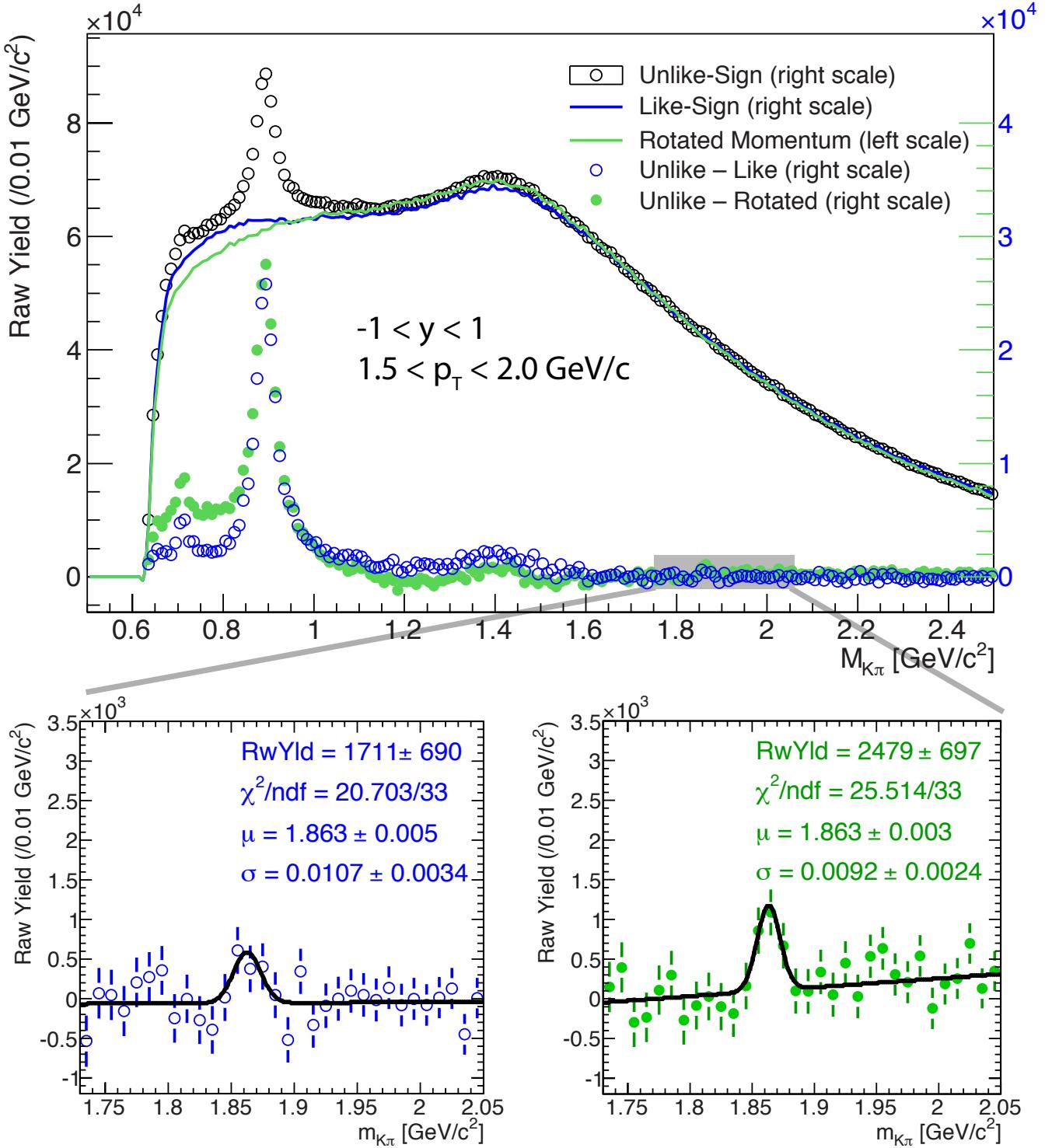


Figure 9: Upper Panel: Opposite-charged $K\pi$ invariant mass with the combinatorial background reconstructed by Like-Sign and Rotated Momentum techniques for all $K\pi$ pairs within $(1.5 < p_T < 2.0 \text{ GeV}/c)$ and $|y(K\pi)| < 1$. The gray rectangle illustrates the zoom to D^0 mass window. Lower left panel: Opposite-charged $K\pi$ pairs invariant mass after Like-Sign background subtracted. Lower right panel: Opposite-charged $K\pi$ pairs invariant mass after Rotated-Momentum background subtracted.

- H_0 : there's no D^0 peak and the observed signal is entirely a product of statistical fluctuations,
- H_1 : there's a D^0 signal described by gaussian function + linear residual background,

was made to evaluate the likability of the signal occurrence. The Table 2 displays both the p -values of the linear function fits and gaussian+linear function fits to the data in the critical region of the invariant mass to be between 1.72 and 2.1 GeV/c². As one can see, the result of the signal obtained

p_T range [GeV/c]	no signal		gaussian signal	
	Like-Sign	Rotated-Momentum	Like-Sign	Rotated-Momentum
$1 < p_T < 2$	47.5%	10.6%	98.7%	74.7%
$.9 < p_T < 1.5$	74.7%	12.5%	95%	35.7%
$1.5 < p_T < 2$	80.9%	13.3%	95.3%	82.1%

Table 2: The results of the χ^2 test of two hypothesis, one assuming no signal in the sample and the other gaussian signal in the same sample. It shows p -values of the linear function and linear+gaussian function fits to data in critical region of the invariant mass between 1.72 and 2.1 GeV/c².

by Rotated-momentum background subtraction rejects H_0 on the significance level $\alpha = 0.14$ for all p_T bins. In the case of the signal obtained by Like-sign background subtraction, the H_0 is entirely rejected on $\alpha = 0.81$ which is very weakly conclusive. However, H_1 has significantly higher p -value than H_0 for both methods and all p_T bins.

4 D^* reconstruction

The D^* meson undergoes a cascade decay

$$D^{*\pm} \xrightarrow[p^*=39 \text{ MeV}/c]{B.R.=67.7\%} D^0 \pi^\pm \xrightarrow{B.R.=3.89\%} K^\mp \pi^\pm \pi^\pm$$

$M_{D^*} - M_{D^0} = 145.43$ MeV/c², which is only slightly higher than the mass of the pion (139.57 MeV/c²). This gives opportunity to calculate the raw yield of the D^* meson as an area of the peak around 145.43 MeV/c² in $M_{K^\mp \pi^\pm \pi^\pm} - M_{K^\mp \pi^\pm}$ where $1.84 < M_{K^\mp \pi^\pm} < 1.89$ GeV/c². The kinematics of such decay described in paragraph B indicates that D^* mesons with $p_T \lesssim 1.6$ GeV/c are not detectable if the TOF is used to identify the daughter particles. It is hence reasonable to start D^* yield calculation at $p_T = 2$ GeV/c.

The first step was exactly the same as the one in D^0 reconstruction described in previous paragraph. If the invariant mass M_2 of the $K^\mp \pi^\pm$ pair lay between 1.84 and 1.89 GeV/c² and $\cos(\theta^*)$ of the kaon in the CMS frame of the $K\pi$ pair was smaller than 0.77, such a pair's four-vector was combined with pions to form a $K^\mp \pi^\pm \pi_S^\pm$ trinity¹ whose invariant mass M_3 was calculated and a histogram was filled with value $\Delta M \equiv M_3 - M_2$. The purpose of the $\cos(\theta^*)$ cut was to reduce the background coming from jets. As one can see in Figure 10, the background from jets becomes to be significant at higher p_T .

The combinatorial background was reconstructed by two independent techniques:

- **Wrong-sign:** In the trinity of daughter particles, π_S had opposite sign to π . Otherwise, the sequence of steps was exactly the same as in the previous paragraph.
- **Side-band:** The sequence of steps was exactly the same apart from M_2 had been lying between 1.7 and 1.8 or 1.92 and 2.02 GeV/c².

¹The symbol π_S denotes all pions which had different Track ID than the one in $K^\mp \pi^\pm$ pair.

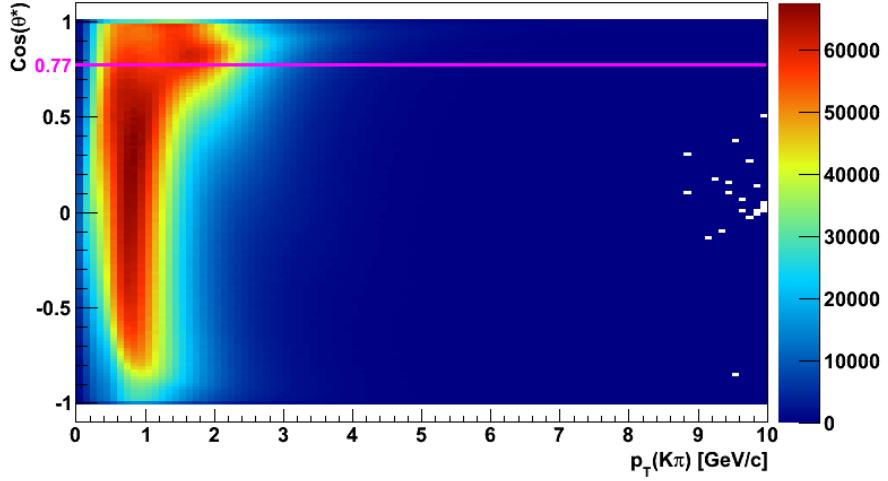


Figure 10: $\cos(\theta^*)$ of the kaon in the CMS of the $K\pi$ pair

The Wrong-sign background yield is “contaminated” by some real D^* signal whose Kaon and Pion daughters from D^0 decays are both mis-identified. The fraction of this over counting in the Wrong-sign background is estimated from a fast toy MC model described in paragraph 7.0.1. The Side-band background yield doesn’t suffer from this contamination so it was reasonable to use Side-band as the default method for combinatorial background reconstruction and the Wrong-sign (corrected on the over counting) as a source of systematic errors estimation.

Figure 11 depicts ΔM with both Wrong-sign and Side-band backgrounds which describe the combinatorial background so well that one doesn’t even need to deal with any residual background. Figure 12 depicts same plots in $K^\mp\pi^\pm\pi^\pm$ p_T bins. The D^* raw yield was calculated as an area of the Gaussian peak obtained by fitting into ΔM spectra after the Side-band background subtraction. Results of those fittings are summarized in Table 3.

p_T bin [GeV/c]	$2 < p_T < 3$	$3 < p_T < 4.2$	$4.2 < p_T < 5.5$	$5.5 < p_T < 8.0$
χ^2/ndf	3.16/5	2.10/5	4.15/5	10.78/6
Raw Yield	83.3 ± 16.9	75.0 ± 16.0	35.5 ± 7.7	16.9 ± 4.6
Mean [Mev/ c^2]	145.23 ± 0.06	145.62 ± 0.12	145.19 ± 0.11	145.37 ± 0.111
Resolution [Mev/ c^2]	0.25 ± 0.05	0.45 ± 0.09	0.45 ± 0.09	0.39 ± 0.08

Table 3: D^* raw yield results in p_T bins.

5 Analysis of Simulated Data

The efficiency and acceptance corrections include the following:

1. TPC track reconstruction efficiency with track quality cuts
2. Matching efficiency
3. Particle Identification Efficiency
4. Kinematical cuts efficiency

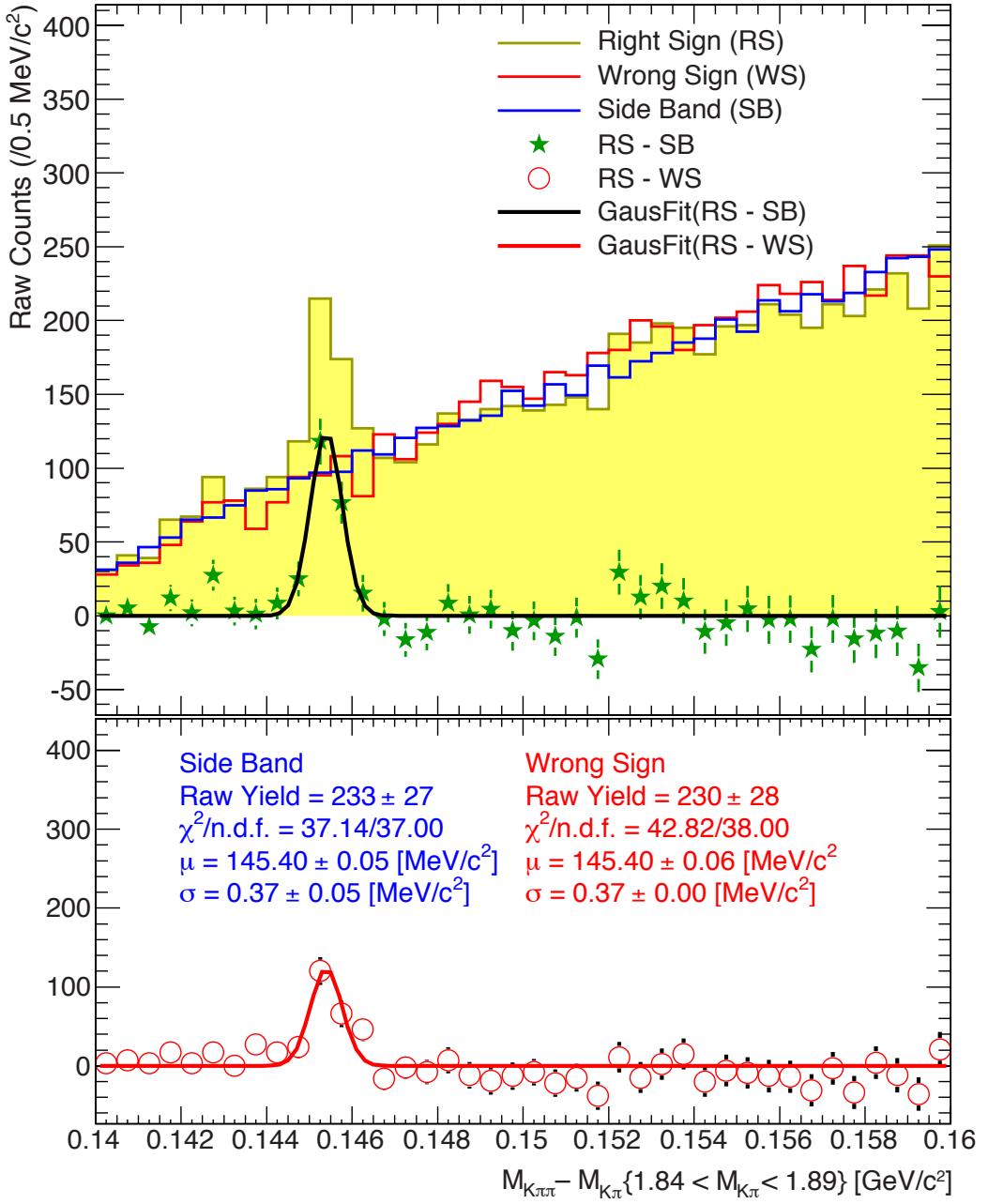


Figure 11: Upper panel: ΔM spectrum with combinatorial backgrounds reconstructed by Side-band and Wrong-sign technique. Green stars and the black line represent the ΔM with Side-band background subtracted and the gaussian fit to it respectively in the range from 0.14 to 0.16 GeV/c^2 . Lower panel: ΔM with Wrong-sign background subtracted and the gaussian fit to it in the range from 0.14 to 0.16 GeV/c^2 . The width of the gaussian function is fixed to be the same as in the case of the Side-band background subtracted.

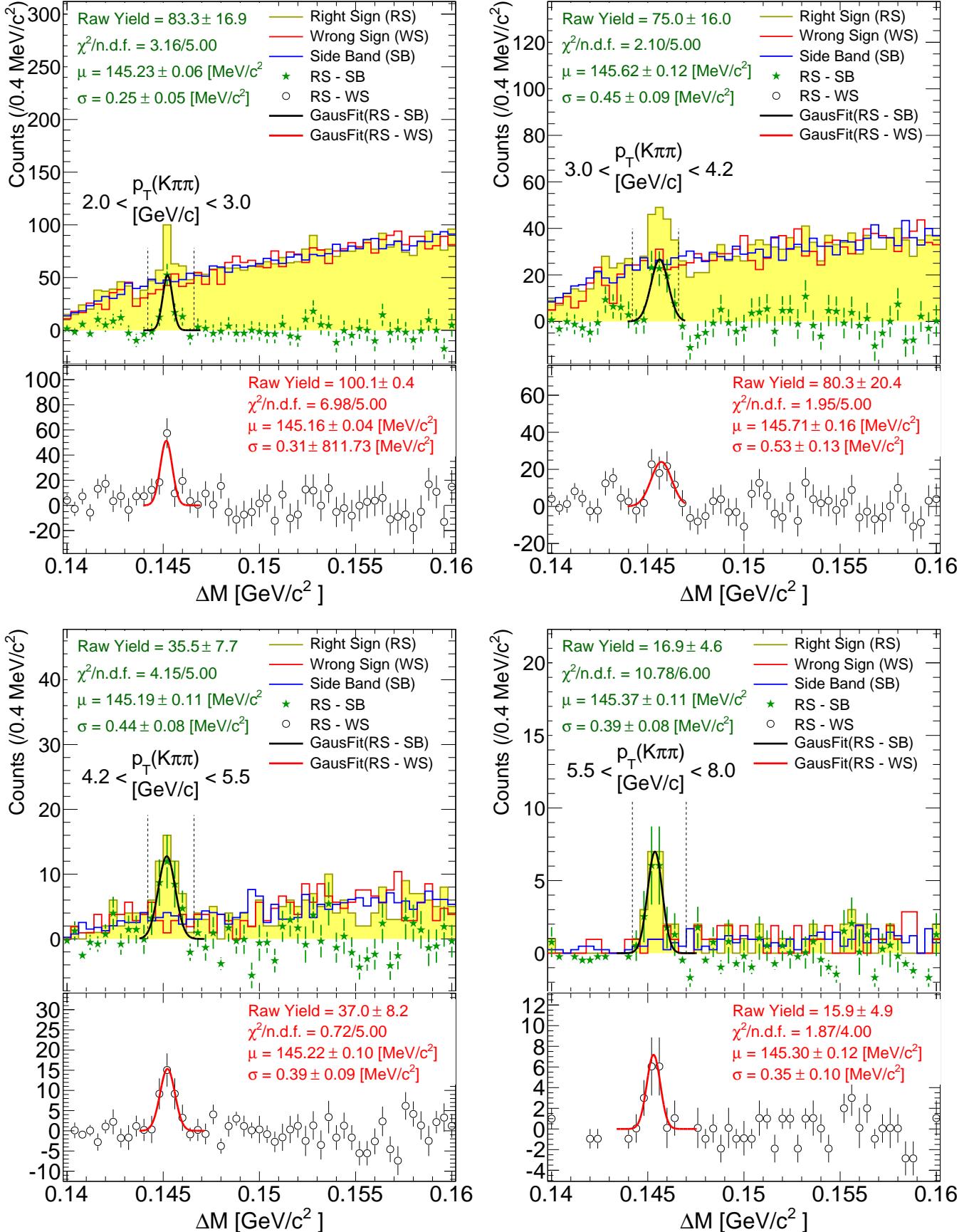


Figure 12: ΔM spectra with combinatorial backgrounds reconstructed by Side-band and Wrong-sign techniques in $K^\pm\pi^\pm$ p_T bins.

5.1 TPC track reconstruction efficiency ε_R

K^-, K^+, π^-, π^+ Monte-Carlo (MC) particles were separately embedded into real events, each type with following setup:

- 50000 p+p 500GeV VPDMB trigger events
- Production Library: P11id
- Geometry: y2011
- Particles per event: 5
- $-50 < V_z < 50$ cm
- $0 < \phi < 6.39$ in radian, flat distribution
- $-1 < \eta < 1$, flat distribution
- $0 < p_T < 8\text{GeV}/c$, flat distribution

The embedding is realized through GEANT simulation of the STAR detector response on those MC particles, so that simulated TPC hits are included in the real hits and such enriched events are then reconstructed through the same reconstruction chain as the real data events. This allows an analyst to track those MC particles down and count how many of them were properly reconstructed and how many were lost in order to calculate reconstruction efficiency. Figure 13 displays ε_R for π^+, π^-, K^+, K^- separately revealing that charge of the particle doesn't affect the reconstruction efficiency at all.

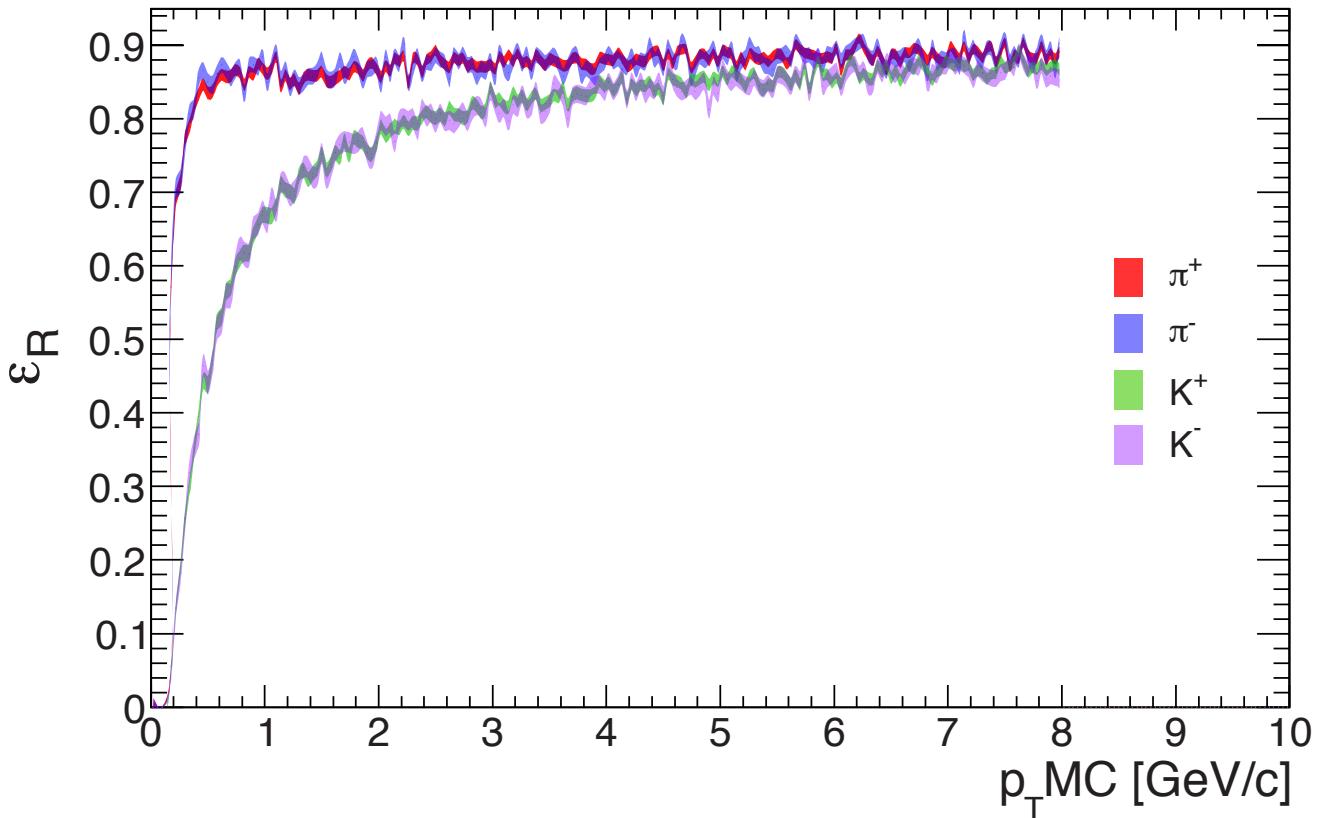


Figure 13: π^+, π^-, K^+, K^- track reconstruction efficiency. Contours reveal statistical uncertainties calculated as ??

5.2 Matching Efficiency, ε_M

The efficiency of matching of the TPC projected tracks to fast detectors (TOF, BEMC) was calculated from the real data. One cannot simply compare number of matched track with the number of all tracks because of a significant amount of pile-up tracks which don't match hits in the fast detectors. A possible solution is using two fast detectors.

Let's consider $P(T)$ and $P(B)$ as a probability of matching in TOF (phenomenon T) and BEMC (phenomenon B) respectively. Then

$$P(T \cup B) = P(T) + P(B) - P(T \cap B). \quad (5.1)$$

If $P(B) > 0$, the $P(T \cap B)$ is possible to express according to Kolmogorov definition as

$$P(T \cap B) = P(T|B)P(B), \quad (5.2)$$

where $P(T|B)$ is the conditional probability of T given B .

The TOF and BEMC matching efficiency as a function of the track p_T , ε_T and ε_B respectively, was then calculated as

$$\varepsilon_T(p_T) = \frac{h\left(p_T; |n\sigma_{K(\pi)}^{dE/dx}| < 2 \wedge B \wedge T\right)}{h\left(p_T; |n\sigma_{K(\pi)}^{dE/dx}| < 2 \wedge B\right)} \quad \varepsilon_B(p_T) = \frac{h\left(p_T; |n\sigma_{K(\pi)}^{1/\beta}| < 2 \wedge T \wedge B\right)}{h\left(p_T; |n\sigma_{K(\pi)}^{1/\beta}| < 2 \wedge T\right)}, \quad (5.3)$$

where $h(p_T; x)$ denotes p_T histogram of the tracks fulfilling the condition x . ε_T is shown in Figure 14 as open and closed circles for pions and kaons respectively. Let's note that $\varepsilon_T(p_T) = P(T|B)$ and $\varepsilon_B(p_T) = P(B|T)$.

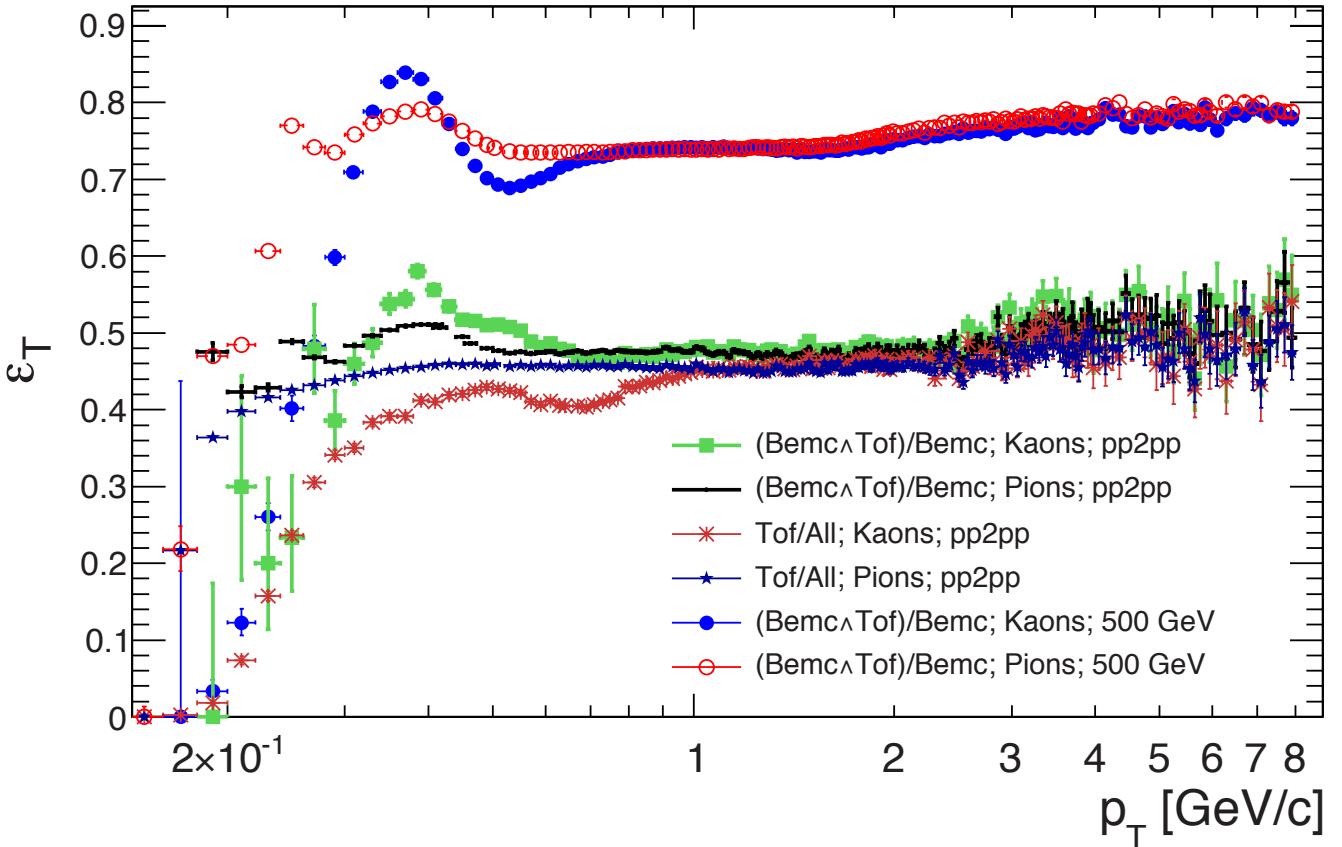


Figure 14: TOF Matching efficiency as a function of the track p_T . Open red (closed blue) circles depict the efficiency calculated according to (5.3) from the production P11id for pions (kaons). The pp2pp_Production2009 trigger set was analyzed 1) by the same way with results depicted as green closed squares (kaons) and black thick dots (pions) 2) according to (5.4) with results depicted as dark red (kaons) and dark blue (pions) asterisks.

As a matter of fact, phenomena T and B are not entirely independent and one must check the degree of their mutual correlation. To do so, there are low-luminosity data called **pp2pp st_physics**, trigger set **pp2pp_Production2009**; let's call them **pp2pp data** in the following text. Events in this trigger set don't contain pile-up events/tracks, so that $\varepsilon_T, \varepsilon_B$ can be calculated as

$$\varepsilon_T^{\text{pp2pp}}(p_T) = \frac{h\left(p_T; |n\sigma_{K(\pi)}^{\text{d}E/\text{d}x}| < 2 \wedge T\right)}{h\left(p_T; |n\sigma_{K(\pi)}^{\text{d}E/\text{d}x}| < 2\right)} \quad \varepsilon_B^{\text{pp2pp}}(p_T) = \frac{h\left(p_T; |n\sigma_{K(\pi)}^{1/\beta}| < 2 \wedge B\right)}{h\left(p_T; |n\sigma_{K(\pi)}^{1/\beta}| < 2\right)}, \quad (5.4)$$

respectively. By comparing the matching efficiency calculated according to (5.4) and the one calculated according to (5.3), it is possible to estimate the degree of correlation between phenomena T and B :

$$\varrho_T = \frac{\varepsilon_T^{\text{pp2pp}}}{\varepsilon_T} \quad \varrho_B = \frac{\varepsilon_B^{\text{pp2pp}}}{\varepsilon_B} \quad (5.5)$$

Figure 14 displays the TOF matching efficiency obtained from pp2pp data, tagged with the "pp2pp" term in the legend. Compared to the efficiency obtained from P11id production (tagged with the "500 GeV" term in the legend), the lower ε_T value is caused by lower TOF acceptance; the TOF detector was not fully installed during Run 9. Further the in the legend of Figure 14, there are terms "(Bemc \wedge Tof)/Bemc" and "Tof/All" denoting the efficiency calculated according to formulas (5.3) and (5.4) respectively. One can see that the efficiency calculated by (5.3) is higher than the other

one and has a "bump" structure at p_T around 0.4 GeV/c (compare the green squares with dark red stars for kaons and black thick dots with dark blue stars for pions) indicating the correlation of phenomena T and B . The Figure 15 displays the same as Figure 14 for BEMC Matching Efficiency ε_B . Figures 16 and 17 display ϱ_T and ϱ_B respectively. As one can see from the plots, unsurprisingly $\varrho_T \simeq \varrho_B$, which confirms the mutual correlation of the phenomena T and B . Let's use simple ϱ in further text. Figure 18 shows corrected TOF and BEMC matching efficiency separately both for pions and kaons while Figure 19 shows it combined. Kaons with no TOF information and $p_T < 1.3$ GeV/c were rejected (see paragraph 2.2), so that $\varepsilon_M = \varrho \varepsilon_T$ for kaons up to $p_T < 1.3$ GeV/c whereas

$$\varepsilon_M = \varrho \varepsilon_{T \cup B} = \varrho (\varepsilon_T + \varepsilon_B - \varepsilon_T \varepsilon_B) \quad (5.6)$$

for all other cases. Let's note that this is the case of D^* analysis. PID cuts in D^0 analysis were different (see paragraph 2.2).

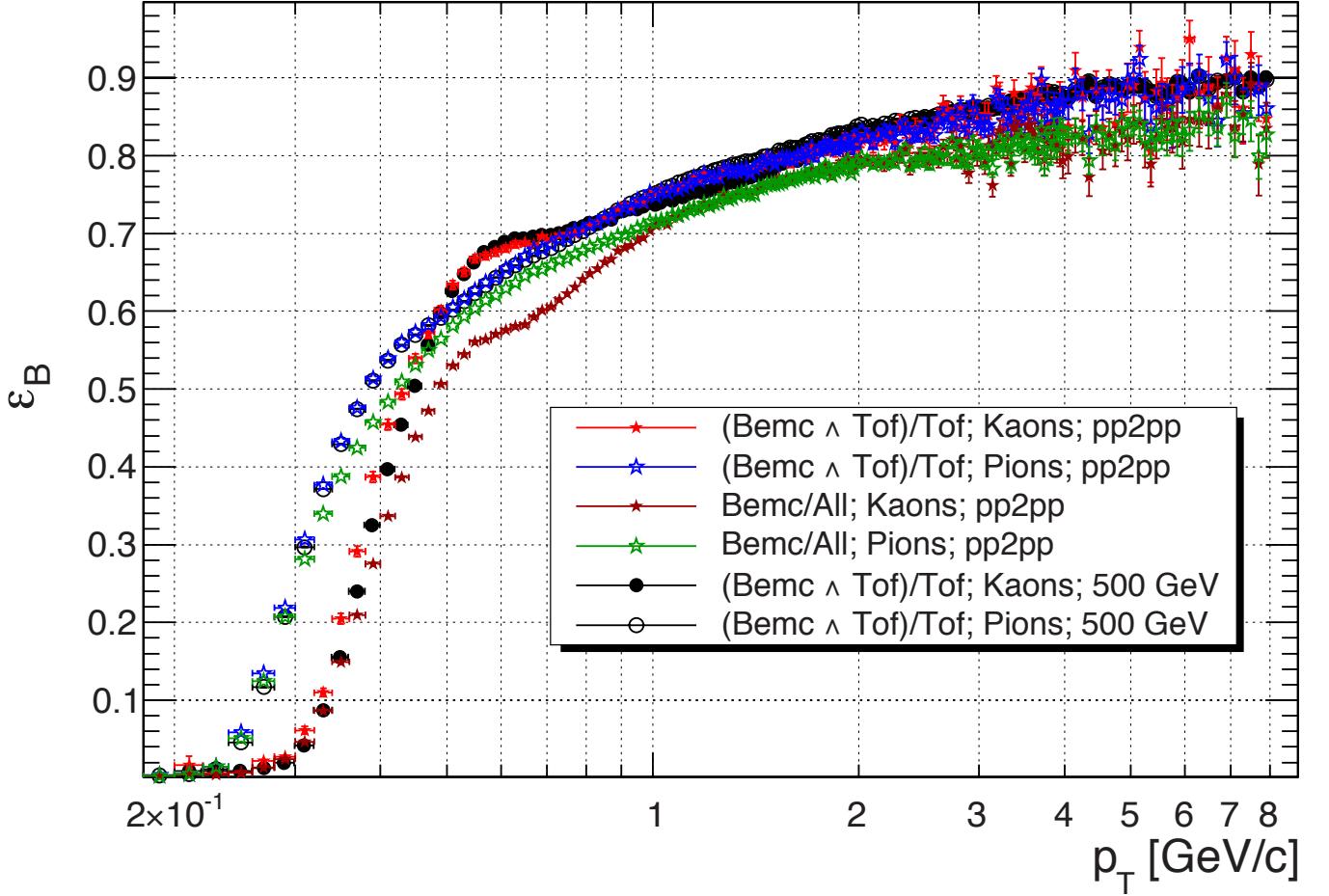


Figure 15: BEMC Matching efficiency as a function of the track p_T . Open (closed) black circles depict the efficiency calculated according to (5.3) from the production P11id for pions (kaons). The pp2pp_Production2009 trigger set was analyzed 1) by the same way with results depicted as red closed asterisks (kaons) and blue open asterisks (pions) 2) according to (5.4) with results depicted as brown closed asterisks (kaons) and green open stars (pions) asterisks.

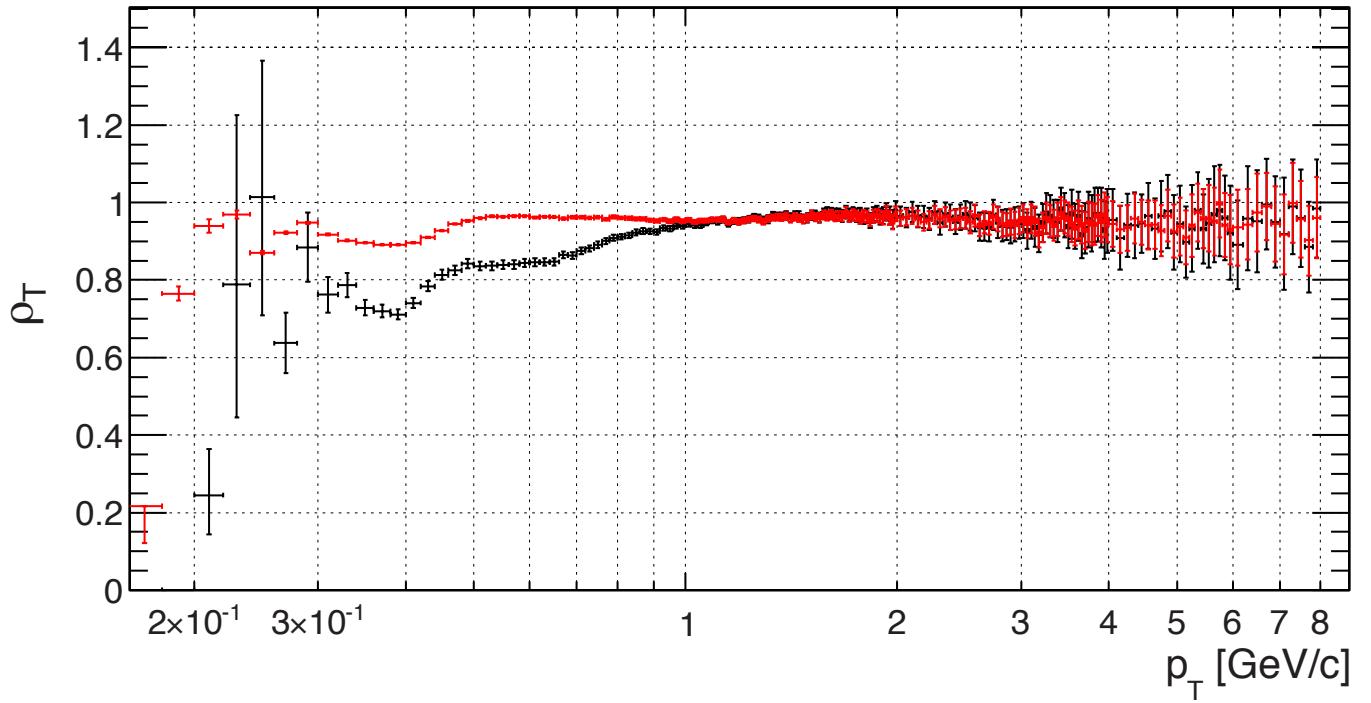


Figure 16: The degree of correlation between phenomena T and B calculated according to (5.5). Red color is dedicated to pions while the black represents kaons.

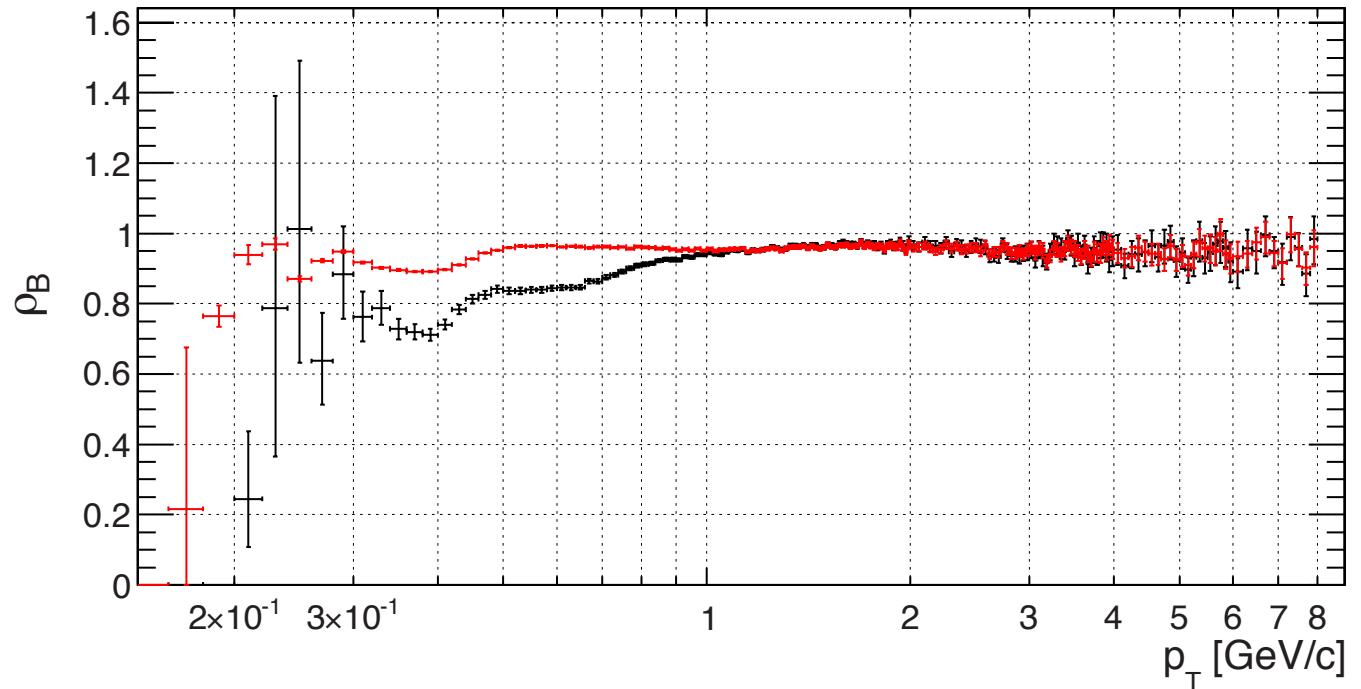


Figure 17: The degree of correlation between phenomena B and T calculated according to (5.5). Red color is dedicated to pions while the black represents kaons.

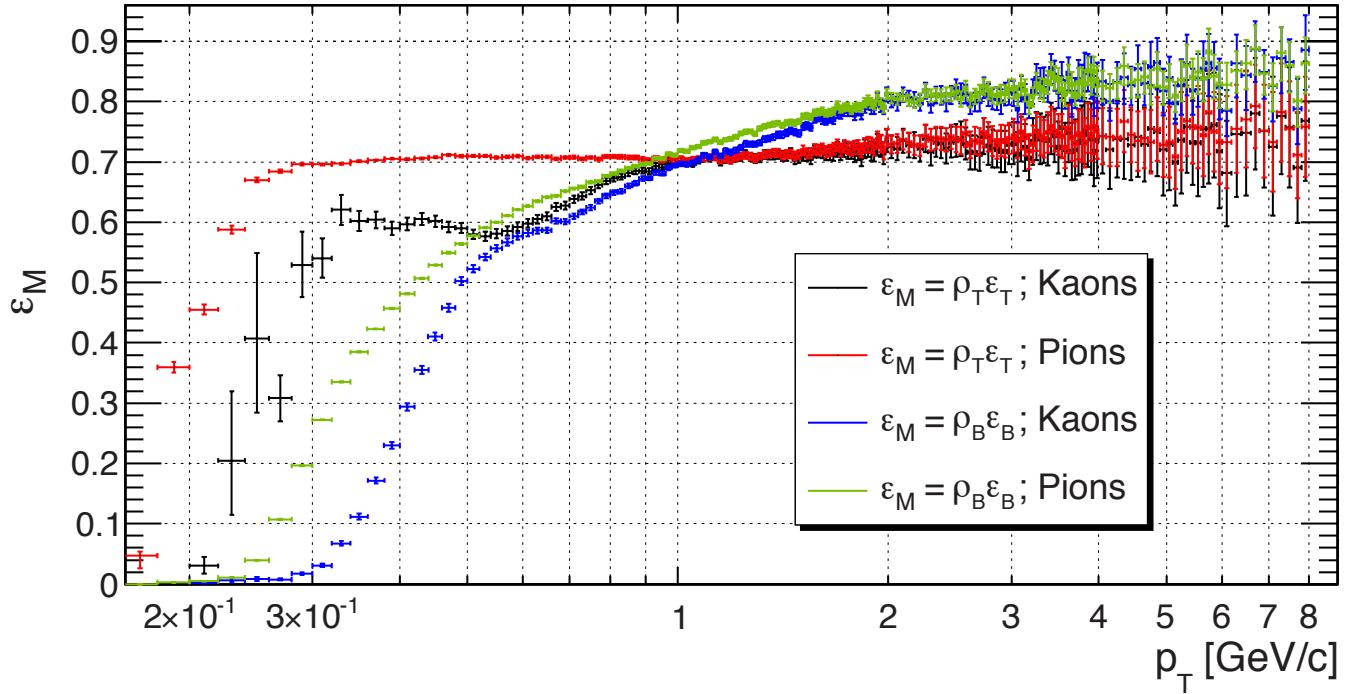


Figure 18: The TOF and BEMC matching efficiency separately.

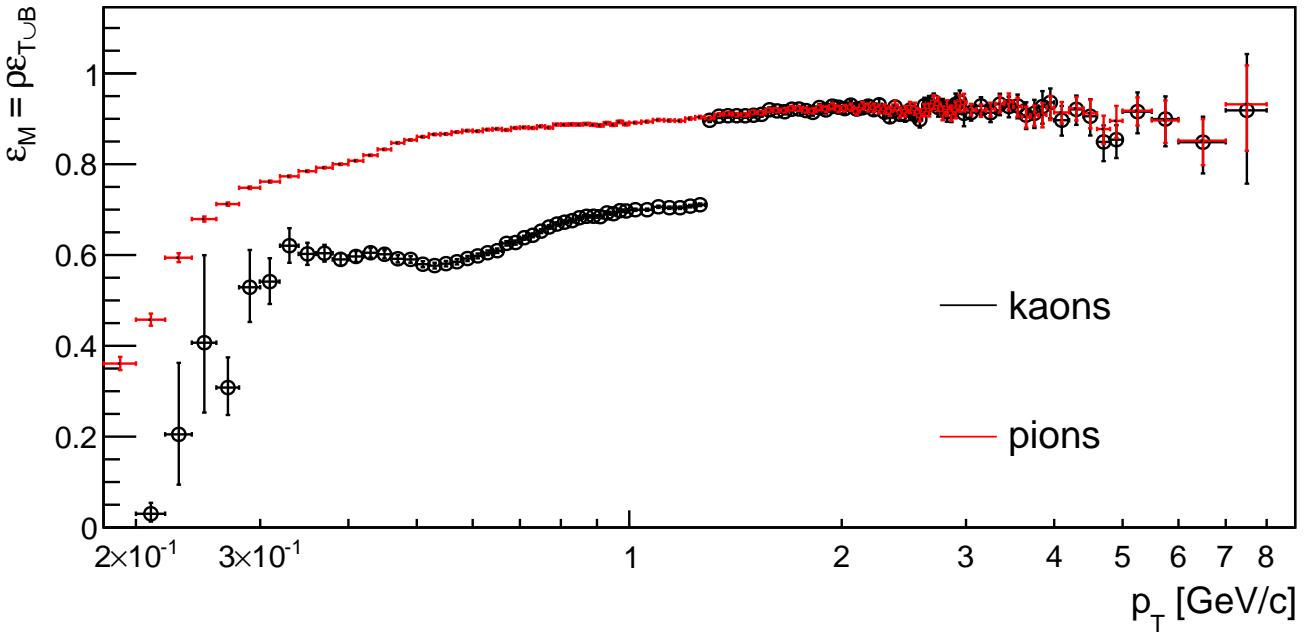


Figure 19: TOF and BEMC matching efficiency combined. Kaons with no TOF information and $p_T < 1.3 \text{ GeV}/c$ were rejected (see paragraph 2.2 , so that $\varepsilon_M = \varrho_T \varepsilon_T$ for kaons up to $p_T < 1.3 \text{ GeV}/c$.

The expected value of the efficiency ξ in given D^0/D^* transverse momentum bin (a, b) was calculated as the inner product of the single track efficiency $\varepsilon(p_T)$ and the track p_T distribution $G^{(a,b)}(p_T)$

for D^0/D^* $p_T \in (a, b)$:

$$\xi \equiv E^{(a,b)}(\varepsilon(p_T)) = \int_0^b \varepsilon(p_T) G^{(a,b)}(p_T) dp_T, \quad (5.7)$$

$$G^{(a,b)}(p_T) \equiv \int_a^b g(t, p_T) dt, \quad (5.8)$$

where t denotes p_T of a mother particle. $g(t, p_T)$ for all mother and daughter particles are shown in Figures 31 and 32. The results are summarized in Table 4.

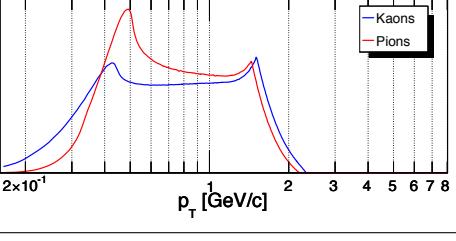
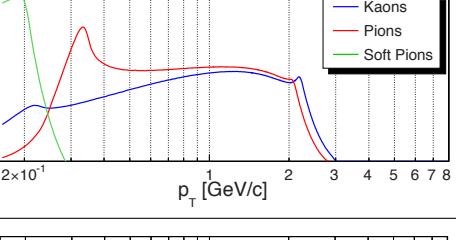
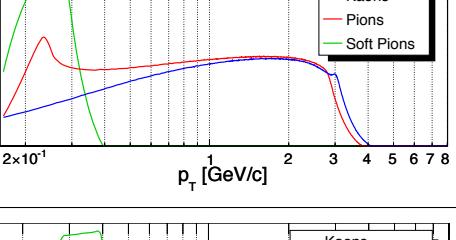
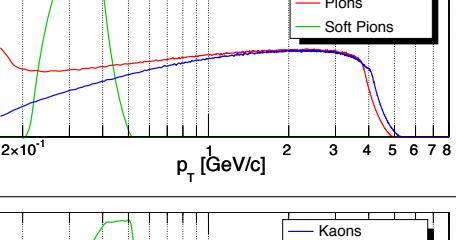
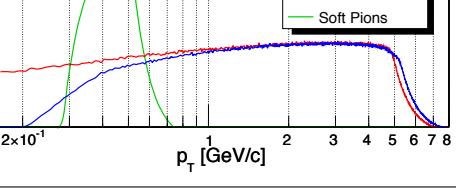
(a, b) [GeV/c]	$G^{(a,b)}(p_T)$	$\xi_R [\%]$			$\xi_M [\%]$		
		K	π	π_S	K	π	π_S
(1.0,2.0)		62.78 +1.48 -1.50	85.26 +1.14 -1.20		70.15 +0.94 -0.90	73.07 +0.72 -0.69	
(2.0,3.0)		66.00 +1.15 -1.16	85.20 +0.95 -0.98	37.70 +0.98 -1.00	75.02 +0.99 -0.94	88.18 +0.43 -0.43	22.18 +2.22 -1.36
(3.0,4.2)		70.81 +1.04 -1.05	85.33 +0.88 -0.92	68.82 +1.45 -1.48	79.17 +1.19 -1.19	88.29 +0.72 -0.72	57.69 +1.34 -0.97
(4.2,5.5)		74.62 +0.90 -0.91	85.35 +0.77 -0.79	78.47 +1.32 -1.37	81.83 +1.60 -1.68	88.45 +1.10 -1.12	75.92 +0.28 -0.28
(5.5,8.0)		78.04 +0.75 -0.76	85.85 +0.64 -0.66	82.85 +1.22 -1.27	83.28 +2.16 -2.40	88.66 +1.60 -1.71	82.48 +0.19 -0.19

Table 4: D^* raw yield results in p_T bins. Let's note that $\xi = \xi^K \times \xi^\pi \times \xi^{\pi_S}$.

5.3 Particle Identification Efficiency

PID cuts were dynamic following the PID resolution. The cut for Pions corresponds to gaussian area between $-3\sigma, +3\sigma$, i.e. 99.6%. The cut for Kaons is split into 3 p_T regions:

- $0.2 < p_T < 1.3 \text{ GeV}/c$: corresponds to gaussian area between $-2\sigma, +3\sigma$, i.e. 97.6%.
- $1.3 < p_T < 2.07 \text{ GeV}/c$: 74 % of kaons candidates identified by TOF, while 26 % by TPC with PID efficiency corresponding to gaussian area between $-2\sigma, +2\sigma$, i.e. 95.45%. The overall PID efficiency in this region was then $0.74 \times 97.6 + 0.26 \times 95.45 = 97\%$
- $p_T > 2.07 \text{ GeV}/c$: corresponds to gaussian area between $-2\sigma, +2\sigma$, i.e. 95.45%.

6 Kinematical cuts efficiency

The reconstructed D^0 invariant mass peak width depends almost exclusively on the daughter particle p_T resolution. When carrying the D^* analysis out the invariant mass of $K\pi\pi$ triplet was calculated only if the $K\pi$ pair invariant mass had been between 1.84 and 1.89 GeV/c^2 and $\cos(\theta^*)$ of the kaon in the CMS frame of the $K\pi$ pair was smaller than 0.77. These cuts introduce a loss in the D^* yield. This section describes the simulation to evaluate the loss.

One can calculate the daughter particle p_T resolution from embedding as a gaussian sigma of the relative difference between reconstructed p_T^{RC} and Monte-carlo p_T^{MC} transverse momentum: $\sigma \left(\frac{p_T^{\text{RC}} - p_T^{\text{MC}}}{p_T^{\text{MC}}} \right)$. Figure 20 shows hereby calculated p_T resolution for Kaons and Pions. The resolution was parametrized by function

$$\rho(m, p_T) \equiv \sqrt{c^2 p_T^2 + \frac{d^2 m^2}{p_T^2} + d^2}, \quad (6.1)$$

where c, d are free parameters obtained by a fit into $\sigma \left(\frac{p_T^{\text{RC}} - p_T^{\text{MC}}}{p_T^{\text{MC}}} \right)$ and m is the mass of the particle. The values of c, d are also displayed in Figure 20.

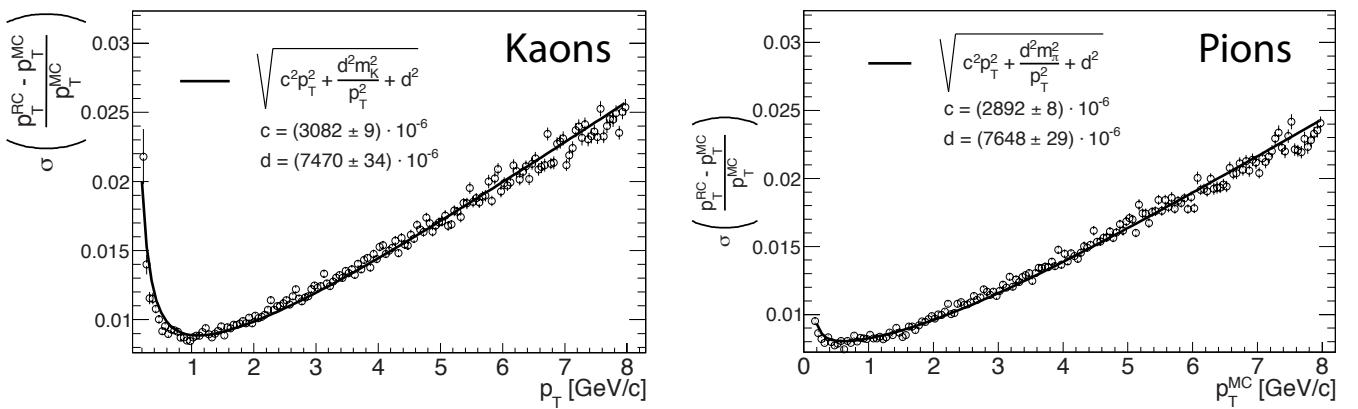


Figure 20: Daughter particle p_T resolution as a function of its p_T

The Monte-carlo Toy model described in paragraph C was used to produce the daughter particles transverse momenta which were smeared according to a gaussian function with the sigma parameter given by $\rho(m, p_T)$. Hereby smeared particles underwent the same cuts on p_T and η as those in the data analysis. $g(t, p_T)$ defined in paragraph C describes the physical D^* kinematics. Figure

21 shows scatter plot, represented by the function $g^{\text{RC}}(t, M^{\text{RC}})$, of D^* transverse momentum t and reconstructed D^0 invariant mass M^{RC} . The kinematical cut efficiency $\varepsilon_{\text{Mass}}$ was calculated for given $D^* p_T$ bin (a, b) as

$$\varepsilon_{\text{Mass}} \equiv \frac{\int_a^{1.84} \int_b^\infty g^{\text{RC}}(t, M^{\text{RC}}) dt dM^{\text{RC}}}{\int_0^{1.89} \int_a^b g^{\text{RC}}(t, M^{\text{RC}}) dt dM^{\text{RC}}} \quad (6.2)$$

and the kinematical cut efficiency of $\cos(\theta^*)$, ε_θ , was the same for all $D^* p_T$ at 0.885.

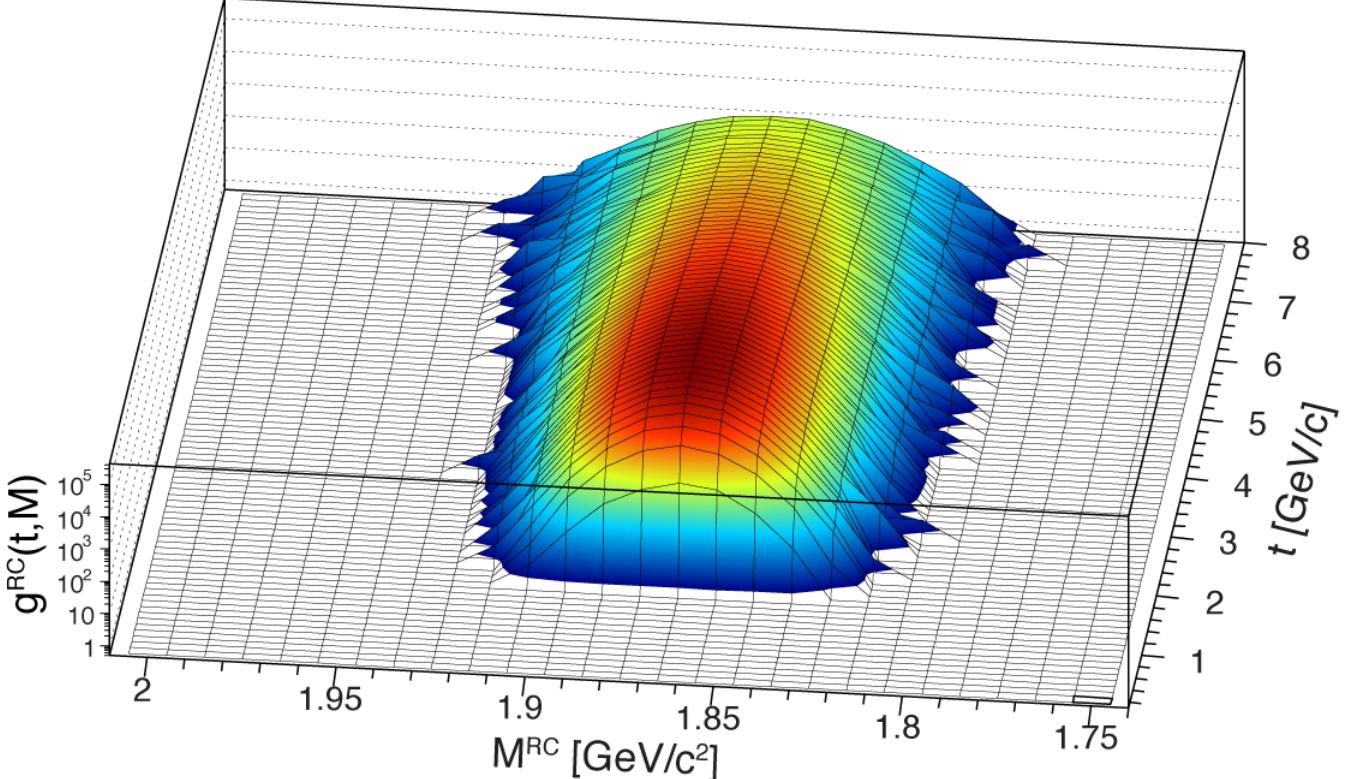


Figure 21: Scatter plot of D^* transverse momentum t and reconstructed D^0 invariant mass M^{RC}

However, we found that $\Delta M = M_{K^\mp\pi^\pm\pi_S^\pm} - M_{K^\mp\pi^\pm}$ peaks were wider than those predicted by our Toy Model if used smearing by $\rho(m, p_T)$ only. It led us to assume that the real momentum resolution was worse than the embedded data analysis suggested. To find the real p_T resolution of primary tracks, we implemented the technique used in Barbara Trzeciak's PhD thesis where an additional smearing function $A \times p_T$ in a Toy Model was implemented (the daughter particle p_T was smeared by $\rho(m, p_T) + Ap_T$, where m was the electron mass) so that the e^+e^- invariant mass resolution (black points and black curve in the left panel of Figure 22) matched the mass resolution of J/ψ peaks (red closed circles in the left panel of Figure 22). The simulation was done various times with a different value of A and χ^2 between the data and simulation was always calculated. The χ^2 values are shown in the right panel of Figure 22 fitted by sixth-degree polynomial whose minimum χ^2_{\min} was get at $A = 4.95 \times 10^{-3}$. The result of the simulation with this value of A is shown black points in the left panel of Figure 22. Blue points in the same plot are results of $A = 5.11 \times 10^{-3}$ and $A = 4.78 \times 10^{-3}$ corresponding to $\chi^2_{\min} \pm 1$.

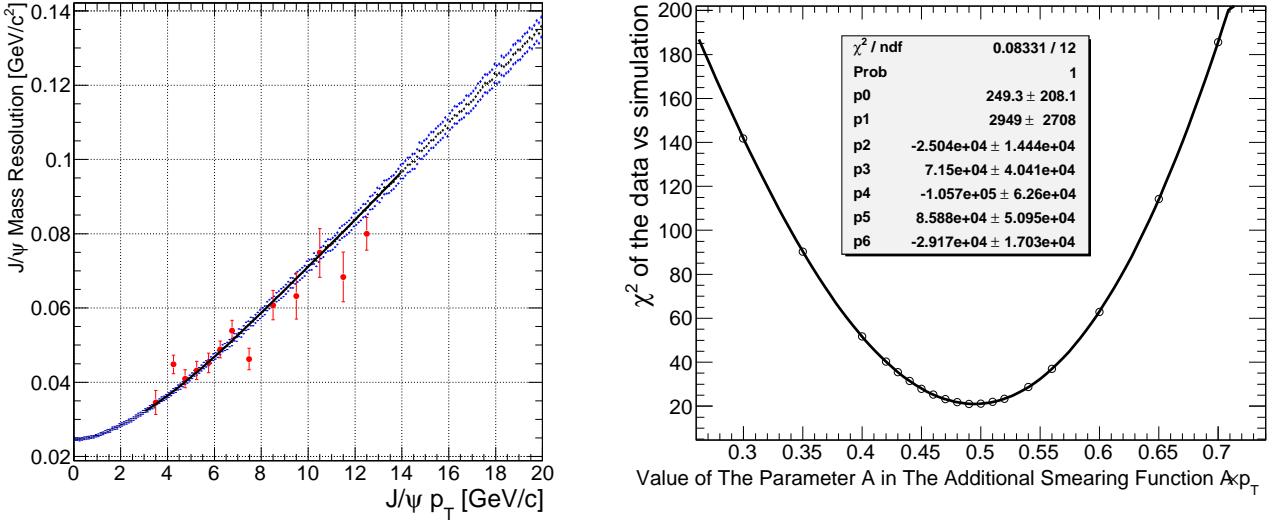


Figure 22: Left panel: The gaussian width of J/ψ peaks simulated by Monte-carlo Toy Model with $\rho(m, p_T)$ given by (6.1) and $A \times p_T$ where A was varied to found the minimum χ^2 , as shown in right panel. We used widths of real J/ψ peaks from analysis [4] as a benchmark. The widths of the real peaks are represented by red circles in left panel.

We simulated ΔM with the $\rho(m, p_T)$ and A (m was the mass of pion or kaon, the parameters c and d were only slightly different for pions from kaons so we expect they aren't different for electrons as well). The mean and sigma of the simulated ΔM are shown in left and right panels respectively of Figure 23. With the additional smearing involved the data and the simulations are in good agreement.

The numerical results are summarized in Table 5

(a, b) [GeV/c]	(2.0,3.0)	(3.0,4.2)	(4.2,5.5)	(5.5,8.0)
$\varepsilon_{\text{Mass}}$ [%]	94.03	89.49	82.22	69.90

Table 5: Kinematical efficiency on D^0 mass

6.1 Trigger bias correction

Heavy quarks are produced during initial hard scatterings creating high p_T particles penetrating easier into calorimeters which makes a higher probability of reconstruction of a vertex. This introduces bias skewed towards events containing charmed particles. Such bias was calculated as a ratio

$$\beta(t) \equiv \frac{\epsilon_{Vpd}\epsilon_{Vtx}}{\xi_{Vpd}(t)\xi_{Vtx}(t)}, \quad (6.3)$$

where ϵ_{Vpd} is the VPD trigger efficiency, ϵ_{Vtx} is the efficiency vertex reconstruction, $\xi_{Vpd}(t)$ is the VPD trigger efficiency for events containing D^* mesons, $\xi_{Vtx}(t)$ is the vertex reconstruction efficiency for events containing D^* mesons, and t a transverse momentum of D^* meson.

The PYTHIA version 6.205 with minimum processes selected and with the CDF TuneA settings was used as the event generator in GEANT to simulate events which were then reconstructed the

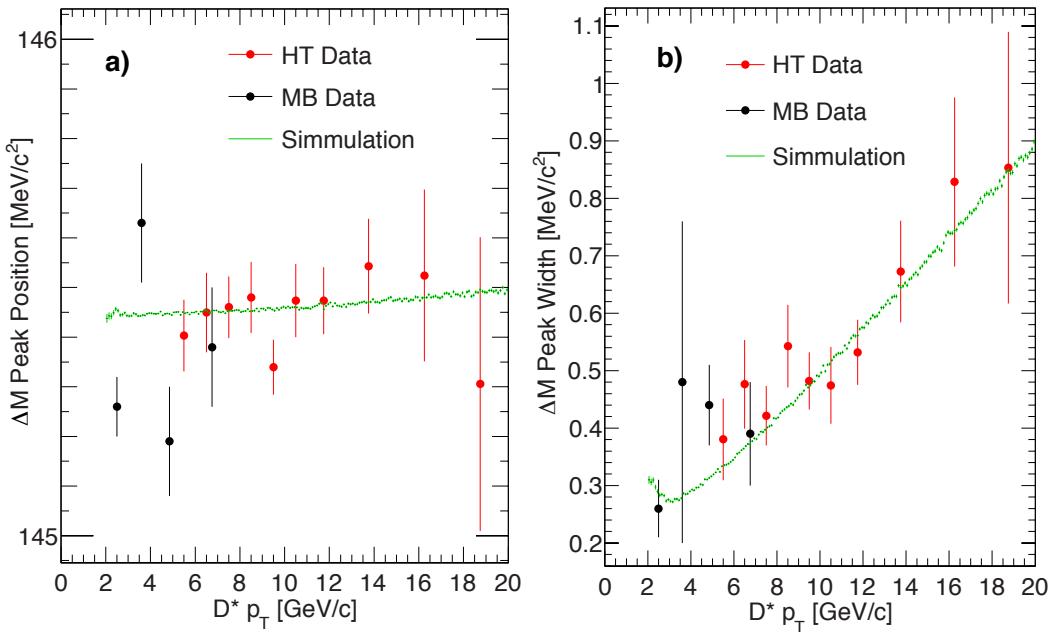


Figure 23: The mean (a) and width (b) of ΔM peak from Gaussian fit to minimum bias and BHT1 data (black and red circles respectively) compared with values from Toy Model simulations (green points).

same way as real events had been. Those reconstructed events were analyzed in order to get BBC and VPD trigger response and vertex. Figure 24 shows an event counter for general events denoted as "MB" and events containing D^* mesons denoted as "Charmed". One can see that the main discrepancy between "MB" and "Charmed" events lies in vertex reconstruction efficiency. Let's note that events containing charmed particles comprise actually high- p_T particles more likely to leave some signal in EMC, which is necessary to give an event positive ranking. Furthermore, D^0 itself decays into particles e^\pm, η which very likely directly or through their decay products leave signal in EMC. Figure 25 then depicts $\xi_{\text{Vpd}}(t)\xi_{\text{Vtx}}(t)$ together with the Trigger bias $\beta(t)$ calculated according to (6.3), where $\epsilon_{\text{Vpd}}\epsilon_{\text{Vtx}}$ was found to be 38.82%, and Table 6 summarizes results for D^0/D^* transverse momentum bins.

$t \in (a, b) [\text{GeV}/c]$	(1.0,2.0)	(2.0,3.0)	(3.0,4.2)	(4.2,5.5)	(5.5,8.0)
$\beta(t)$	0.696 ± 0.003	0.652 ± 0.003	0.630 ± 0.005	0.629 ± 0.008	0.636 ± 0.013

Table 6: Trigger Bias

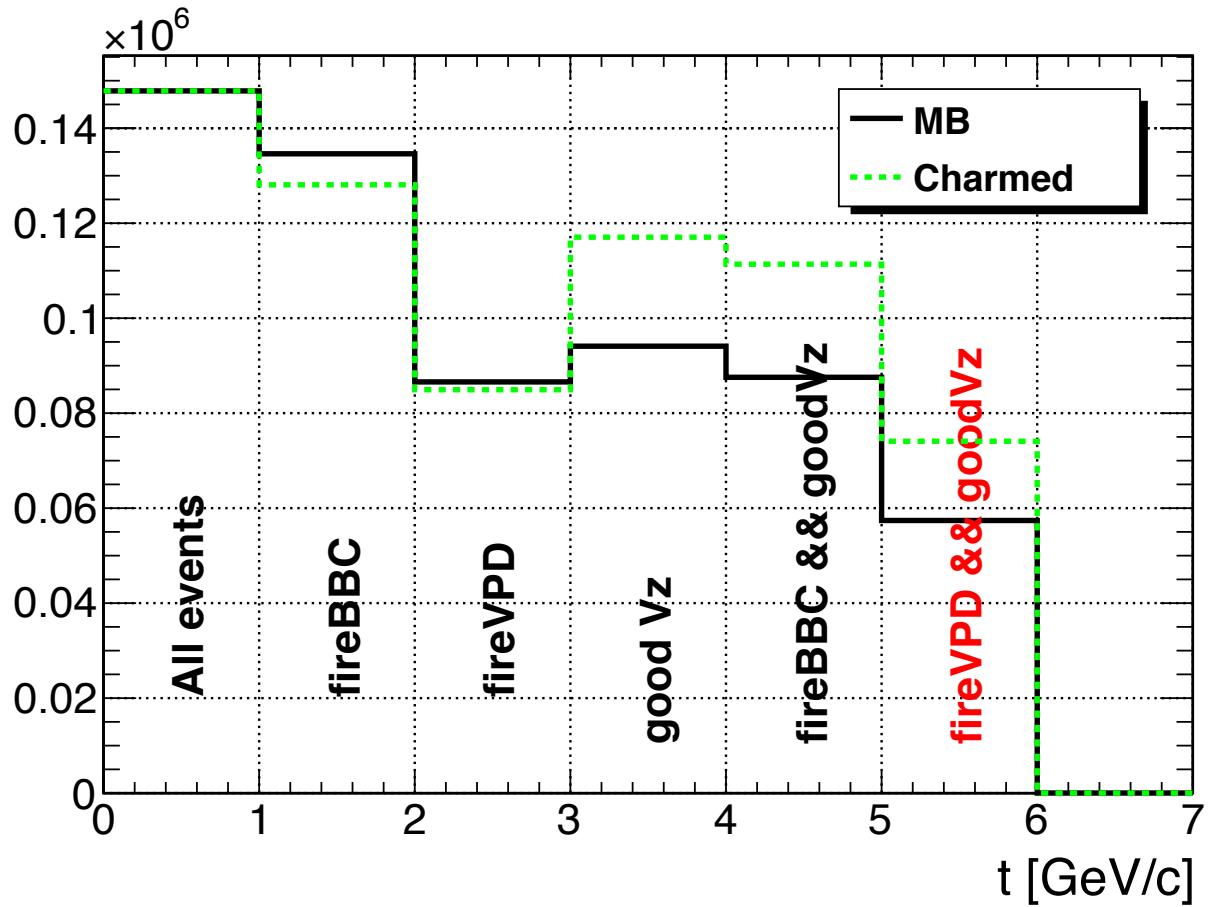


Figure 24: Event counter for general events denoted as "MB" and events containing D^* mesons denoted as "Charmed". "fireBBC" and "FireVPD" columns represent events that initialized BBC and VPD trigger respectively. "good Vz" denotes events from which vertices with positive ranking were successfully reconstructed.

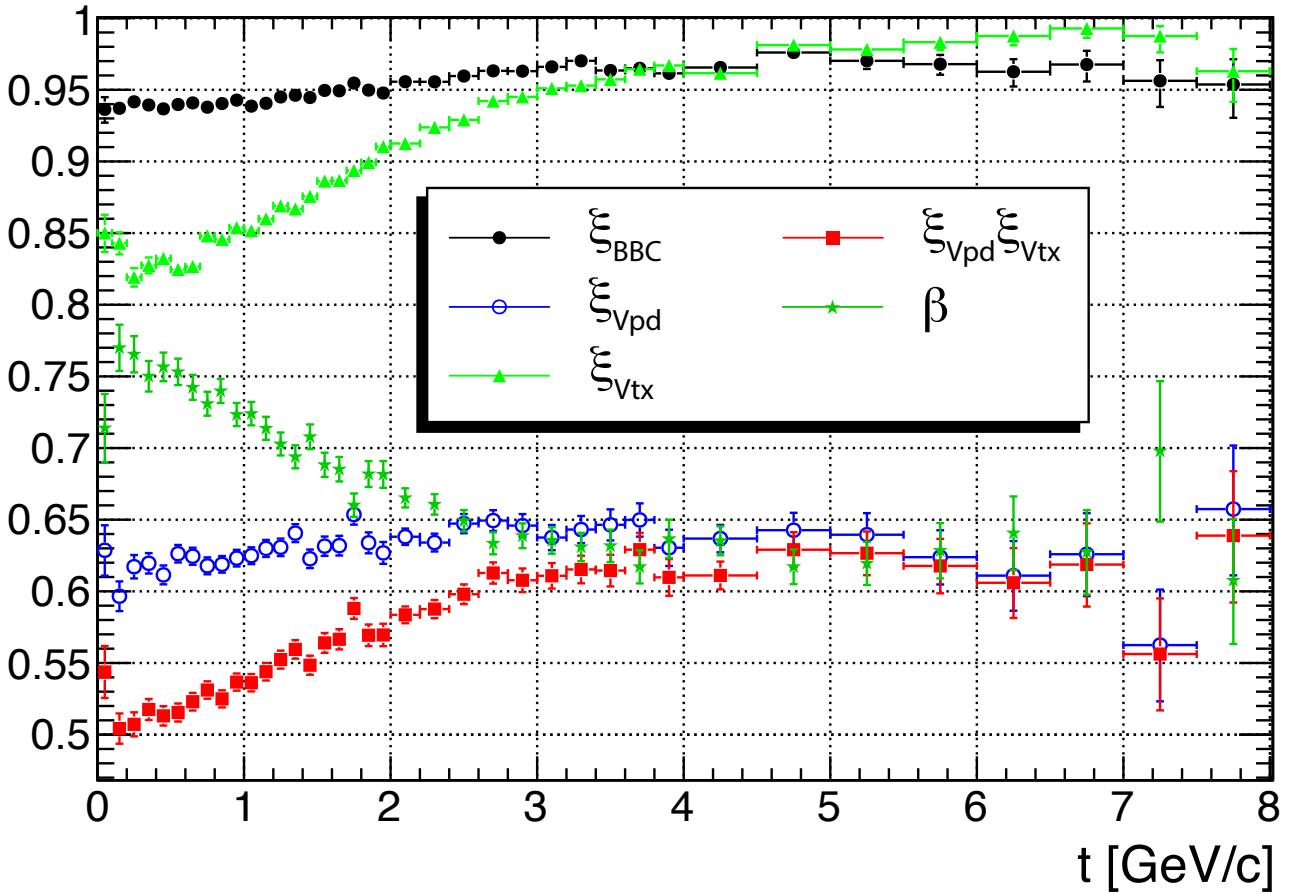


Figure 25: Trigger, Vertex reconstruction efficiency and Trigger Bias

7 Systematic Errors Evaluation

The bin-to-bin systematic errors have two main source: uncertainty on the raw yield determination and uncertainty on the efficiency determination. Systematic errors on the raw yield determination are described in detail in the analysis note of HT data.

There were also several minor sources which were found to have either minor or negligible impact, namely:

- propagation of statistical errors from the TPC track reconstruction and matching efficiency evaluation. This uncertainty was found to be between 2 and 4.6 % which had a minor impact on the total bin-by-bin systematic errors.
- systematic uncertainty from the BHT1 trigger efficiency evaluation found to be between 0.8% (p_T points around 9 GeV/c) and 4.5% (upper error at the highest p_T point) and 5.7% (upper error at the p_T point between 5 and 6 GeV/c). This error ticks up the upper systematic error of the 5-6 GeV/c point from 11.6% to 13.1%. The highest p_T has large systematic uncertainty on the yield determination (31.7%) so the 4.5% doesn't make a difference. Let's note that this error affects only the HT data.
- error from the trigger or vertex reconstruction bias determination which was estimated in the previous STAR measurement in p+p collisions at 200 GeV to be 5.2%.

- impact from TPC reconstructed pileup events on the trigger or vertex reconstruction bias, which was also estimated in the previous STAR measurement to be 4%.

7.0.1 Double counting

The D^0 is reconstructed via hadronic decay channel: $D^0 \rightarrow K^- + \pi^+$. If K^- is mis-identified as a π^- and π^+ is mis-identified as a K^+ , the $(K^+\pi^-)$ combination will contribute in the $\overline{D^0}$ reconstruction, and if it falls into the $\overline{D^0}$ mass selection window, this D^0 signal will be then also counted as a $\overline{D^0}$ signal. This is called double counting.

For D^* reconstruction via $D^\pm \rightarrow D^0\pi_S^\pm \rightarrow K^\mp\pi^\pm\pi_S^\pm$, the situation is different. If K^\mp is mis-identified as π^\mp and π^\pm as K^\pm , then the combination becomes $K^\pm\pi^\mp\pi_S^\pm$ which doesn't contribute to the ΔM peak, but into the Wrong-sign background. Hence the raw yield obtained by the Wrong-sign background subtraction is undercounted.

Once again, the Monte-carlo Toy model (par C) was employed to produce $g^{\text{RC}}(t, M^{\text{RC}})$ (see paragraph 6, particularly Figure 21). Furthermore, if both Kaon and Pion happened to be in the overlapping PID region, an additional histogram $g^{\text{RC}}(t, M^{\text{Mis}})$ was filled. M^{Mis} is the invariant mass of mis-identified Kaon (actually a Pion) and Pion (actually a Kaon). This so called mis-identified Kaon/Pion was simply created by replacement of the rest mass in the first particle's four-momentum with the second particle's rest mass and vice versa. The fraction of the over counting in the Wrong-sign background was calculated then as:

$$f_{\text{OC}}(t) \equiv \frac{\int_{1.84}^{1.89} g^{\text{RC}}(t, M^{\text{Mis}}) dM^{\text{Mis}}}{\int_{1.84}^{1.89} g^{\text{RC}}(t, M^{\text{RC}}) dM^{\text{RC}}} \quad (7.1)$$

The probabilities of a Kaon and a Pion in the overlapping PID region are shown in Figure 26, $g^{\text{RC}}(t, M^{\text{RC}})$ and $g^{\text{RC}}(t, M^{\text{Mis}})$ in Figure 27, and $f_{\text{OC}}(t)$ in Figure 28.

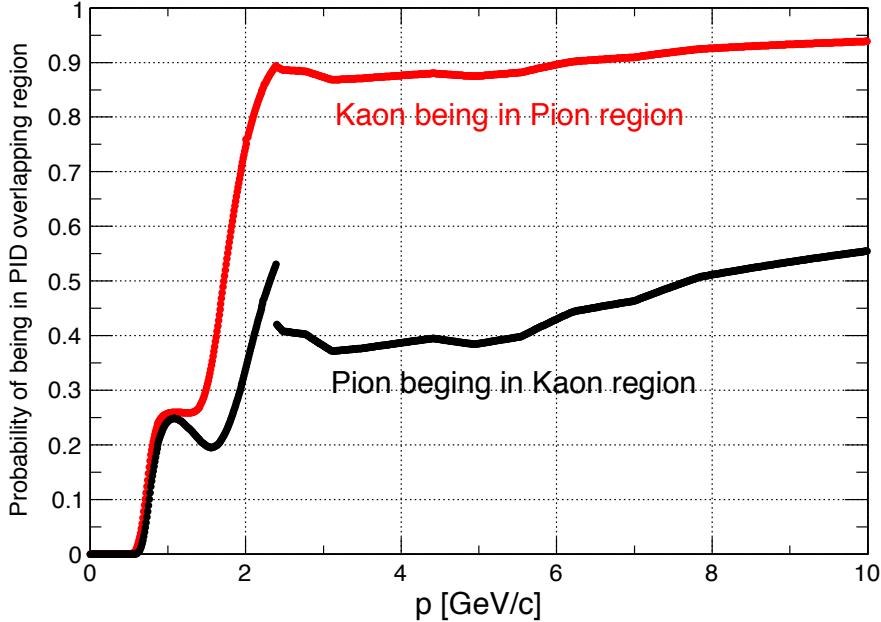


Figure 26: Probability of a particle being in overlapping PID region

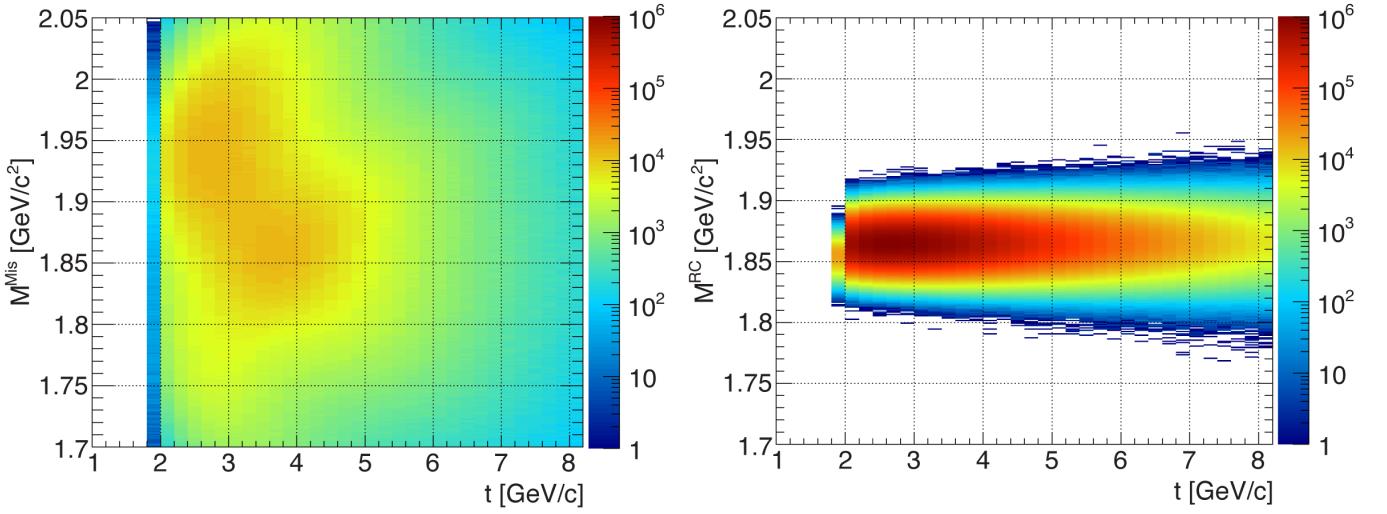


Figure 27: Probability of a particle being int overlapping PID region

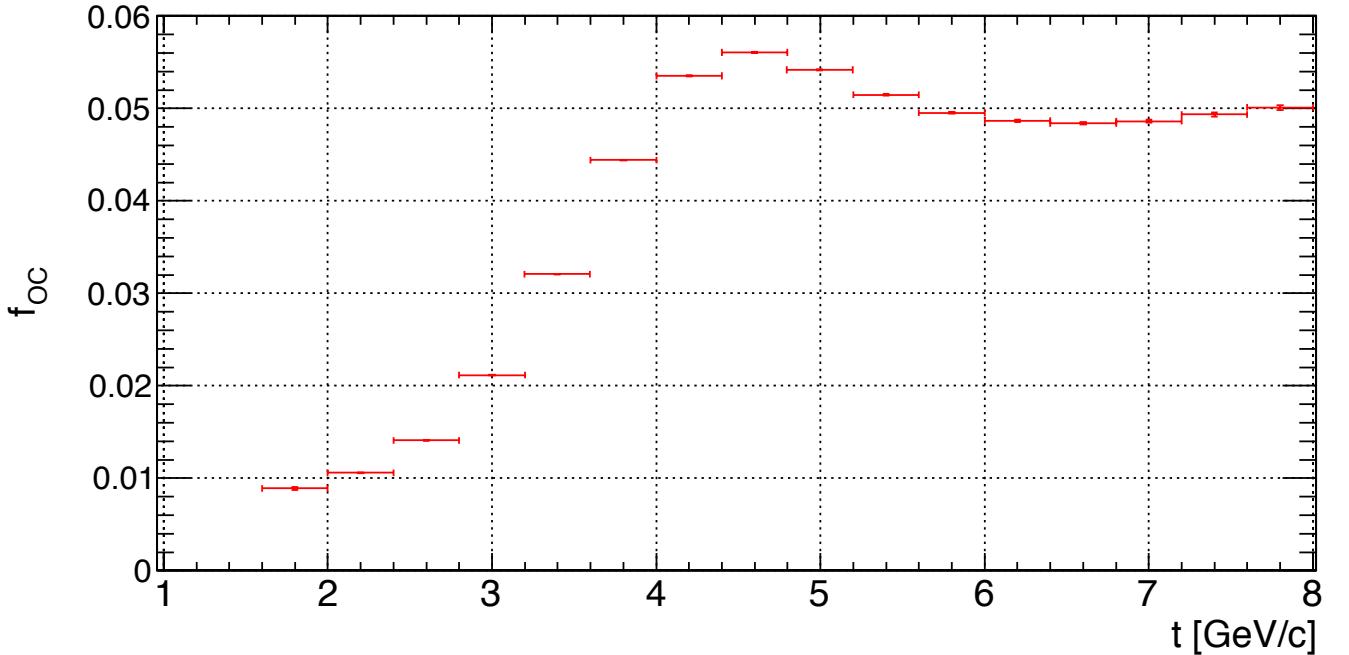


Figure 28: Probability of a particle being in overlapping PID region

7.1 Systematic Error Of Embedding

The absolute (“true”) reconstruction efficiency in data is unknown, so it is calculated from embedding, which is an approximation. The degree of how the basic track quality distributions (DCA, NHITS) from embedding matches those from the data represents how good is the approximation. And the difference between them is considered as systematic uncertainty due to this approximation.

It’s difficult to calculate the absolute efficiency of the Number of TPC Hit Points (Nhits) or DCA from the data, because of the contamination of Pile-up and ”ghost” tracks, so the relative efficiencies $\varepsilon_{\text{DCA}}^{(\text{rel})}, \varepsilon_{\text{Nhits}}^{(\text{rel})}$ was used instead:

$$\varepsilon_{\text{DCA}}^{(\text{rel})}(p_T) \equiv \frac{\int_0^1 h(p_T, r) dr}{\int_0^3 h(p_T, r) dr}, \quad \varepsilon_{\text{Nhits}}^{(\text{rel})}(p_T) \equiv \frac{\int_{25}^{45} h(p_T, n) dn}{\int_{15}^{45} h(p_T, n) dn}, \quad (7.2)$$

where $h(p_T, r), h(p_T, n)$ are DCA, Nhits distributions respectively at given p_T . Systematic discrepancy at given p_T was then:

$$\delta_{\text{DCA}}(p_T) \equiv \frac{\varepsilon_{\text{DCA}}^{(\text{rel})}(p_T) \text{ from data}}{\varepsilon_{\text{DCA}}^{(\text{rel})}(p_T) \text{ from embedding}}, \quad \delta_{\text{Nhits}}(p_T) \equiv \frac{\varepsilon_{\text{Nhits}}^{(\text{rel})}(p_T) \text{ from data}}{\varepsilon_{\text{Nhits}}^{(\text{rel})}(p_T) \text{ from embedding}}, \quad (7.3)$$

Figure 29 depicts $\delta_{\text{DCA}}(p_T)$ and $\delta_{\text{Nhits}}(p_T)$.

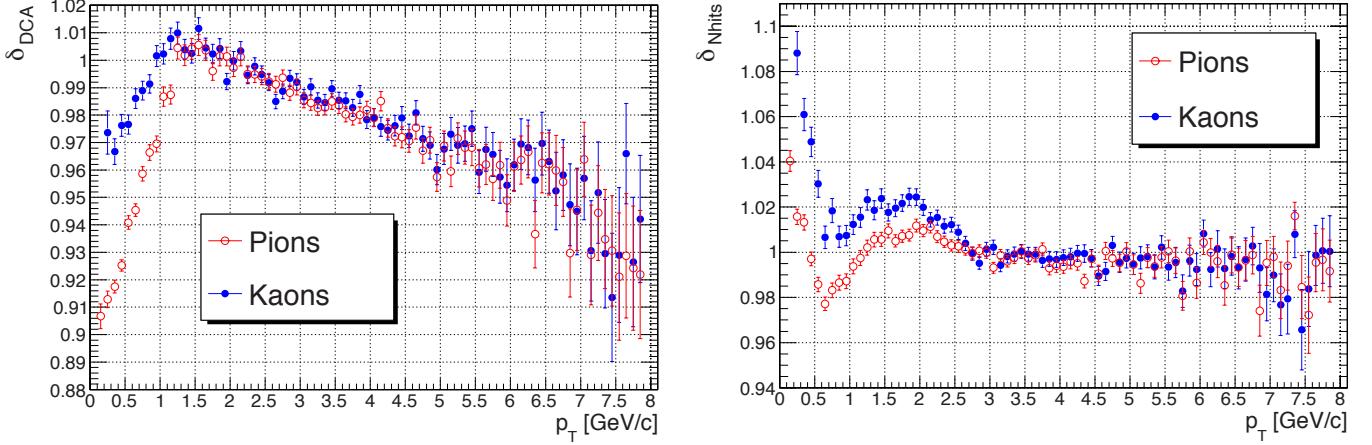


Figure 29: Systematic discrepancy between the data and the embedding

The expected value of the systematic uncertainty from embedding ς in given D^0/D^* transverse momentum bin (a, b) was calculated analogously with the reconstruction/matching efficiency calculated according to (5.7):

$$\varsigma \equiv E^{(a,b)}(|1 - \delta(p_T)|) = \int_0^8 |1 - \delta(p_T)| G^{(a,b)}(p_T) dp_T, \quad (7.4)$$

where $G^{(a,b)}(p_T)$ is defined by (5.8). Results are summarized in Table 7. With the assumption of an ideal correlation among daughter particles's systematic uncertainties, the systematic error of the yield could have been directly calculated as

$$\varsigma = \sqrt{\left(\sum_{i=1}^{\#\text{daughters}} \varsigma_{\text{Nhits}}^i \right)^2 + \left(\sum_{i=1}^{\#\text{daughters}} \varsigma_{\text{DCA}}^i \right)^2}$$

and is listed in the last column in Table 7.

(a, b) [GeV/c]	$\varsigma_{\text{Nhits}} [\%]$				$\varsigma_{\text{DCA}} [\%]$				$\varsigma [\%]$
	K	π	π_S	Sum	K	π	π_S	Sum	
(1.0,2.0)	2.36	0.95		3.32	1.06	3.07		4.13	5.30
(2.0,3.0)	2.14	0.94	3.34	6.42	0.88	2.51	9.16	12.55	14.10
(3.0,4.2)	1.65	0.80	2.00	4.44	0.90	2.08	8.75	11.73	12.54
(4.2,5.5)	1.27	0.70	1.21	3.17	1.13	2.01	8.25	11.39	11.83
(5.5,8.0)	1.01	0.65	1.01	2.67	1.62	2.27	7.22	11.10	11.42

Table 7: Systematic errors from discrepancy between embedded and experimental data

The contribution from matching efficiency uncertainty was found to be negligible (below 3% of ς).

8 Results

The differential invariant cross section $E \frac{d^3\sigma}{dp^3}$ of an averaged single charm quark ($\frac{c+\bar{c}}{2}$) production at mid-rapidity was calculated according to equation

$$E \frac{d^3\sigma}{dp^3} \Big|_{y=0} = \frac{d^2\sigma}{2\pi p_T dp_T dy} \Big|_{y=0} = \frac{1}{2} \frac{1}{2\pi} \frac{\sigma^{\text{NSD}} \beta}{N f_c \Gamma} \frac{Y}{p_T \Delta p_T \Delta y} \frac{1}{\varepsilon}, \quad (8.1)$$

where σ^{NSD} is inelastic non-singly diffractive $p + p$ cross section, β is the trigger bias (paragraph 6.1), N is the total number of events entered the analysis, f_c represents the ratio of a charm quark hadronizing to an open charm meson, Γ denotes the branching ratio of a decay, and Y is the raw yield in a p_T bin of the width Δp_T within the rapidity window $\Delta y = 2$. ε is further the combined efficiency of the experiment including the track reconstruction efficiency ε_R (see paragraph 5.1), the track matching efficiency ε_M (see paragraph 5.2), the particle identification efficiency ε_{PID} (see paragraph 5.3), and the kinematical cuts efficiencies $\varepsilon_{\text{Mass}}$ and ε_θ (see paragraph 6). Results are summarized in Table 8.

σ^{NSD} [mb]	34 ± 2.7				
N	$51\,771\,500$				
f_c [%]	56.5 ± 3.2	22.4 ± 2.8			
Γ [%]	3.89	2.63			
β	0.696	0.652	0.630	0.629	0.636
p_T [GeV/c]	1.5	2.5	3.6	4.85	6.75
Δp_T [GeV/c]	1.0	1.0	1.2	1.3	2.5
Δy	2				
Y	4064 ± 1006	83.9 ± 18.1	82.5 ± 17.0	35.4 ± 7.7	16.8 ± 4.6
ε [%]	25.9	4.9	14.7	20.9	21.5
$E \frac{d^3\sigma}{dp^3} \Big _{y=0}$ [nb]	8606 ± 1502	2005 ± 406.9	333.1 ± 71.05	76.16 ± 16.52	13.31 ± 3.622
ς [%]	5.3	14.1	12.54	11.83	11.42

Table 8: Results in p_T bins. Let's note those results are not corrected on bin widths. ς doesn't represent all systematic uncertainties.

Let's note that the p_T values shown in Table 8 were chosen arbitrarily as bin centers since the real values hadn't been known. To know exact p_T values, one must know the exact shape of the p_T and that had also been unknown. This problem can be solved iteratively. Let's define the new transverse momentum

$$p_T^{(i+1)} = \mathcal{F}^{(i)-1} \left(\int_a^b \mathcal{F}^{(i)}(p_T) dp_T \right), \quad (8.2)$$

where $\mathcal{F}^{(i)} \equiv p_T f(p_T)$ with parameters obtained from the $f(p_T)$ fit into $\frac{d^2\sigma^{(i)}}{2\pi p_T^{(i)} dp_T dy} \Big|_{y=0}$ calculated according to (8.1). $f(p_T)$ is the power law function either of Hagedorn's (A.4) or Lévy's (A.2) shape. $\frac{d^2\sigma^{(0)}}{2\pi p_T^{(0)} dp_T dy} \Big|_{y=0}$ and $p_T^{(0)}$ were set to have values from Table 8. After the third iteration the results

became to be very stable. Results are shown in the Analysis note of the HT data.

A Differential Invariant Cross Section

A.1 As a Hagedorn Shape Function

Hard scattering amplitudes follow a power-law function giving us assumption of open charm invariant cross section power-law behavior. Long time ago Hagedorn proposed the QCD inspired empirical formula describing the data of the invariant cross section of hadrons as a function of p_T over a wide range [6]:

$$\frac{d^2\sigma}{p_T dp_T dy} = A \left(1 + \frac{p_T}{p_0}\right)^{-n} \rightarrow \begin{cases} \exp\left(-\frac{np_T}{p_0}\right) & \text{for } p_T \rightarrow 0 \\ \left(\frac{p_0}{p_T}\right)^n & \text{for } p_T \rightarrow \infty \end{cases} \quad (\text{A.1})$$

where A, p_0, n are arbitrary constants. This function has indeed become a purely exponential function for small p_T and a purely power law function for large p_T ¹. The mean p_T becomes:

$$\langle p_T \rangle = \frac{A \int_{-\infty}^{\infty} \frac{d\sigma}{dy} dy \int_0^{\infty} p_T^2 \left(1 + \frac{p_T}{p_0}\right)^{-n}}{A \int_{-\infty}^{\infty} \frac{d\sigma}{dy} dy \int_0^{\infty} p_T \left(1 + \frac{p_T}{p_0}\right)^{-n}} = \frac{2p_0}{n-3}$$

The normalization constant A can be obtained from the relation:

$$d\sigma/dy = A \int_0^{\infty} p_T \left(1 + \frac{2p_T}{p_0}\right)^{-n} = \frac{Ap_0^2}{(n-1)(n-2)}$$

so we get

$$A = 4 \frac{d\sigma}{dy} \frac{(n-1)(n-2)}{\langle p_T \rangle^2 (n-3)^2}, \quad p_0 = \frac{\langle p_T \rangle}{2} (n-3)$$

and $c\bar{c}$ invariant cross section can be represented by

$$\frac{d^2\sigma^{c\bar{c}}}{2\pi p_T dp_T dy} = \frac{2}{\pi} \frac{d\sigma^{c\bar{c}}}{dy} \frac{(n-1)(n-2)}{\langle p_T \rangle^2 (n-3)^2} \left(1 + \frac{2p_T}{\langle p_T \rangle (n-3)}\right)^{-n} \quad (\text{A.2})$$

with three free parameters $d\sigma^{c\bar{c}}/dy, \langle p_T \rangle, n$ need to be obtained from the least square fit of the real corrected data points.

A.2 As a Levi Function

As an alternative to Hagedorn (A.1) formula, one can use a different approach based on the Tsallis statistics [7] to fit particle spectra. The Tsallis distribution was derived from a generalized form of the

¹It is widely known from experimental data that, as expected from pQCD calculations [5], a pure power law shape successfully describes the high p_T region of particle spectra. At low p_T , suggests a thermal interpretation in which the bulk of the produced particles are emitted by a system in thermal equilibrium with a Boltzmann-Gibbs statistical description of their spectra:

$$E \frac{d^3\sigma}{dp^3} = Ae^{-E/T}$$

where A is a normalization factor and E is the particle energy. At mid-rapidity one can replace E by $m_T = \sqrt{p_T^2 + m_0^2}$, where m_0 is the particle rest mass.

Boltzmann-Gibbs entropy. However, there are other origins discussed in recent days [8] suggesting for example hard collisions approach [9]. The distribution could be written in the form:

$$\frac{d^2\sigma}{2\pi p_T dp_T dy} = C_n \left(1 + \frac{\sqrt{p_T^2 + m_0^2}}{nT} \right)^{-n} \quad (\text{A.3})$$

where C_n is the normalization constant, n the power and T an inverse slope parameter.

To calculate C_n , one must integrate (A.3) over p_T

$$C_n = \frac{1}{\int_0^\infty p_T \left(1 + \frac{\sqrt{p_T^2 + m_0^2}}{nT} \right)^{-n} dp_T} = \frac{1}{\frac{(nT + m)^{1-n}}{(nT)^{-n}} \frac{m(n-1) + nT}{(n-1)(n-2)}} \quad (\text{A.3})$$

to get the (A.3) into the form appropriate to fitting:

$$\frac{d^2\sigma^{c\bar{c}}}{2\pi p_T dp_T dy} = \frac{1}{2\pi} \frac{d\sigma^{c\bar{c}}}{dy} \frac{(n-1)(n-2)}{(nT + m_0)[m_0(n-1) + nT]} \left(\frac{nT + \sqrt{p_T^2 + m_0^2}}{nT + m_0} \right)^{-n} \quad (\text{A.4})$$

with three free fitting parameters $d\sigma^{c\bar{c}}/dy, n, T$.

B D^* Decay Kinematics

Consider the decay of a D^* of mass M in its CMS frame. Then its 4-momentum is $P^* = (M, 0, 0, 0)$. Denote the 4-momenta of the daughter particles in CMS by p_1^* (for D^0) and p_2^* (for π_S): $p_1^* = (E_1^*, \mathbf{p}_1^*)$, $p_2^* = (E_2^*, \mathbf{p}_2^*)$. The 4-momentum conservation requires $P^* = p_1^* + p_2^*$ and hence $\mathbf{p}_2^* = -\mathbf{p}_1^*$. We can therefore omit the subscript on the daughter particle momenta and hence energy conservation takes on the form $E_1 + E_2 = \sqrt{m_1^2 + p^{*2}} + \sqrt{m_2^2 + p^{*2}} = M$. Solving this equation for p^* we get

$$p^* = \frac{1}{2M} \sqrt{[M^2 + (m_1 - m_2)^2][M^2 - (m_1 - m_2)^2]} = 39.397 \text{ MeV}, \quad (\text{B.1})$$

where $M = 2010.28$ MeV, $m_1 = 1864.84$ MeV, and $m_2 = 139.57$ MeV.

If the D^* moves in one direction with velocity β the basic Lorentz transformation equations in Cartesian form give the daughter particle 4-momentum components the following relations:

$$p_T = p_\perp^*, \text{ transverse component}^1 \quad (\text{B.2a})$$

$$p_\parallel = \gamma p_\parallel^* + \gamma \beta E_\parallel^*, \text{ longitudinal component}^1 \quad (\text{B.2b})$$

$$E = \gamma \beta p_\parallel^* + \gamma E_\parallel^*, \text{ energy component} \quad (\text{B.2c})$$

Let's note that all variables are considered to be in natural units.

More specifically, if the D^* transverse momentum is P then the soft pion's transverse momentum² p will be between

$$\begin{aligned} \gamma \beta E_\pi^* - \gamma p^* &\leq p \leq \gamma \beta E_\pi^* + \gamma p^*, \quad \beta = \frac{P}{E}, \quad \gamma = \frac{E}{M}, \quad \gamma \beta = \frac{P}{M} \\ \frac{P}{M} \sqrt{p^{*2} + m_\pi^2} - \frac{\sqrt{P^2 + M^2}}{M} p^* &\leq p \leq \frac{P}{M} \sqrt{p^{*2} + m_\pi^2} + \frac{\sqrt{P^2 + M^2}}{M} p^* \end{aligned} \quad (\text{B.3})$$

To illustrate this inequality, we plotted the two maximum and minimum curves shown in Figure 30. If $P < \frac{M}{m_\pi^2} p^*$ then the π_S 3-momenta has negative sign, but the Figure 30 shows absolute values.

The Magenta horizontal dashed line illustrates the track p_T cut corresponding to the minimum track p_T value for the track to be able to reach the TOF detector.

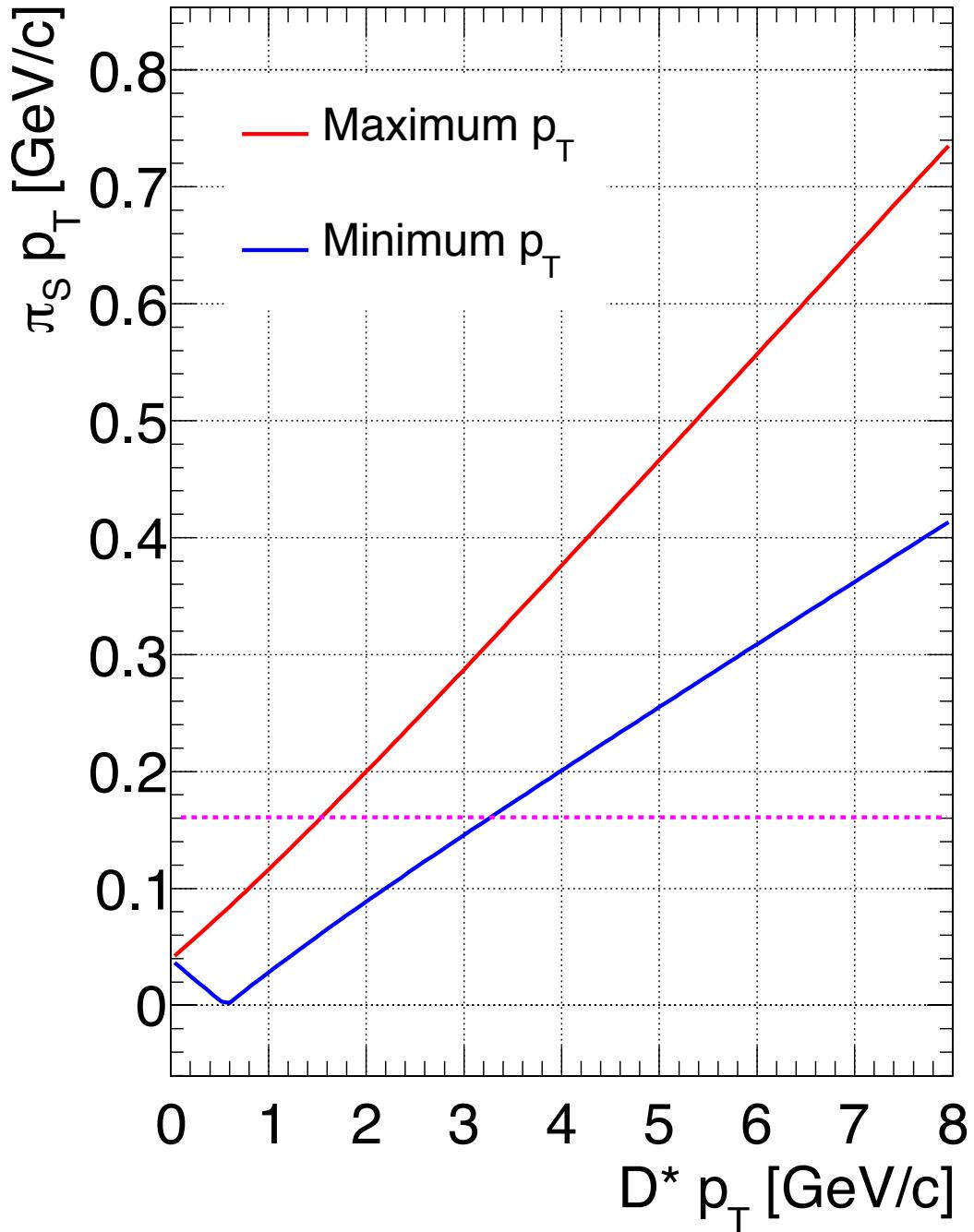


Figure 30: Kinematically possible values of the soft pion transverse momentum from the $D^* \rightarrow D^0 \pi_S$ decay

¹transverse or longitudinal relative to the momentum vector in the LAB frame

² $p \equiv p_T$ in LAB

C Monte-Carlo Toy Model

In order to get the exact daughter particle p_T distribution for any given mother particle p_T , the Monte-Carlo (MC) mother particle (D^0 or D^*) was generated by the ROOT TRandom3 MC generator with p_T distributed according to $p_T \times (A.4)$, $\frac{d\sigma^{c\bar{c}}}{dy} = 0.19, T = 0.12, n = 8.37$, and azimuthal angle ϕ and rapidity y were distributed uniformly from $(0, 2\pi)$ and $(-1, 1)$ respectively. Such MC particle was let to decay isotropically and daughter particles were boosted then in the direction of the mother particle. Figures 31 and 32 show results of this simulation in the form of $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ function $g(t, p_T)$, where t denotes a transverse momentum of D^* and p_T a transverse momentum of the respective daughter particle.

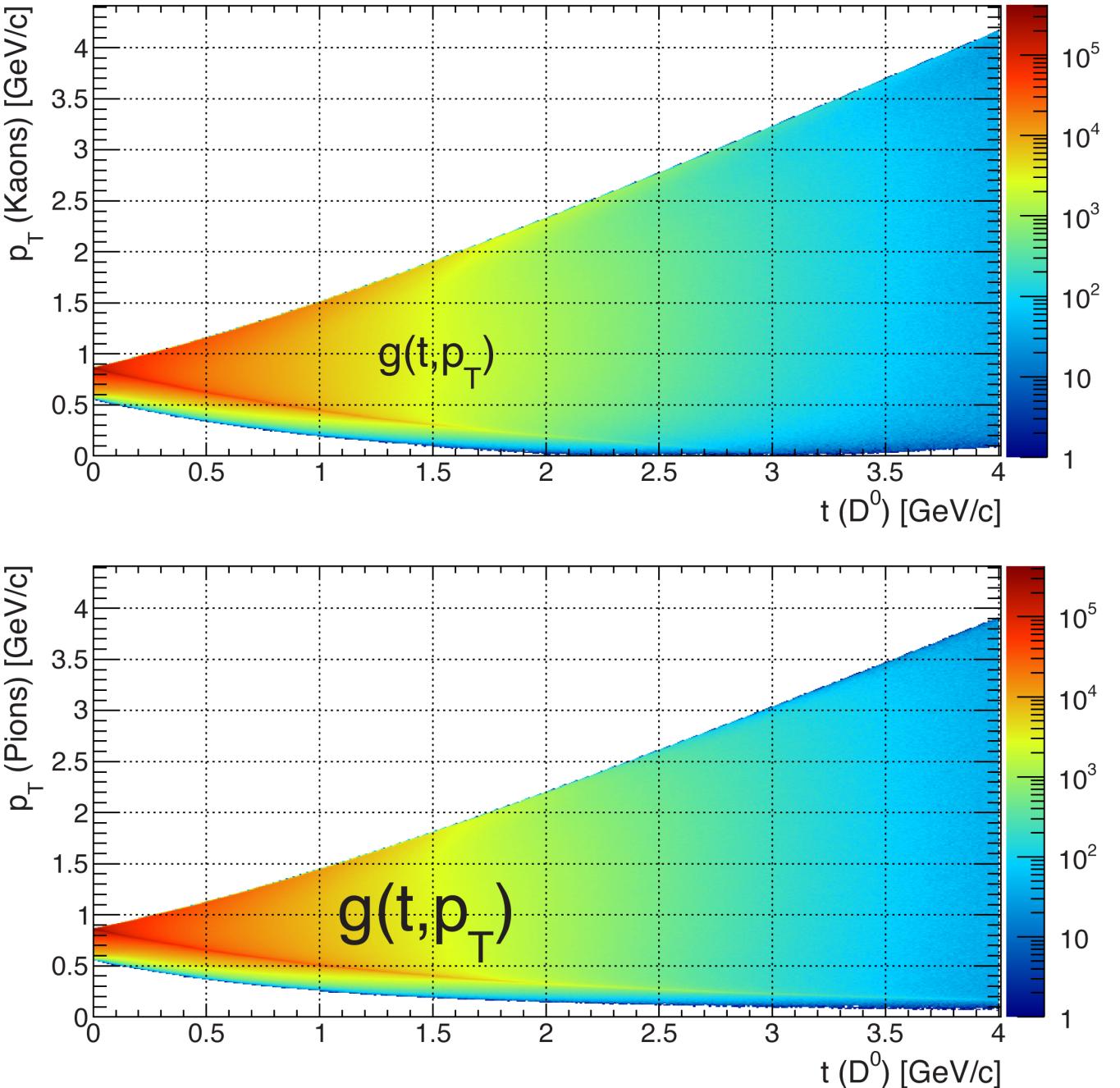


Figure 31: D^0 decay kinematics $g(t, p_T)$

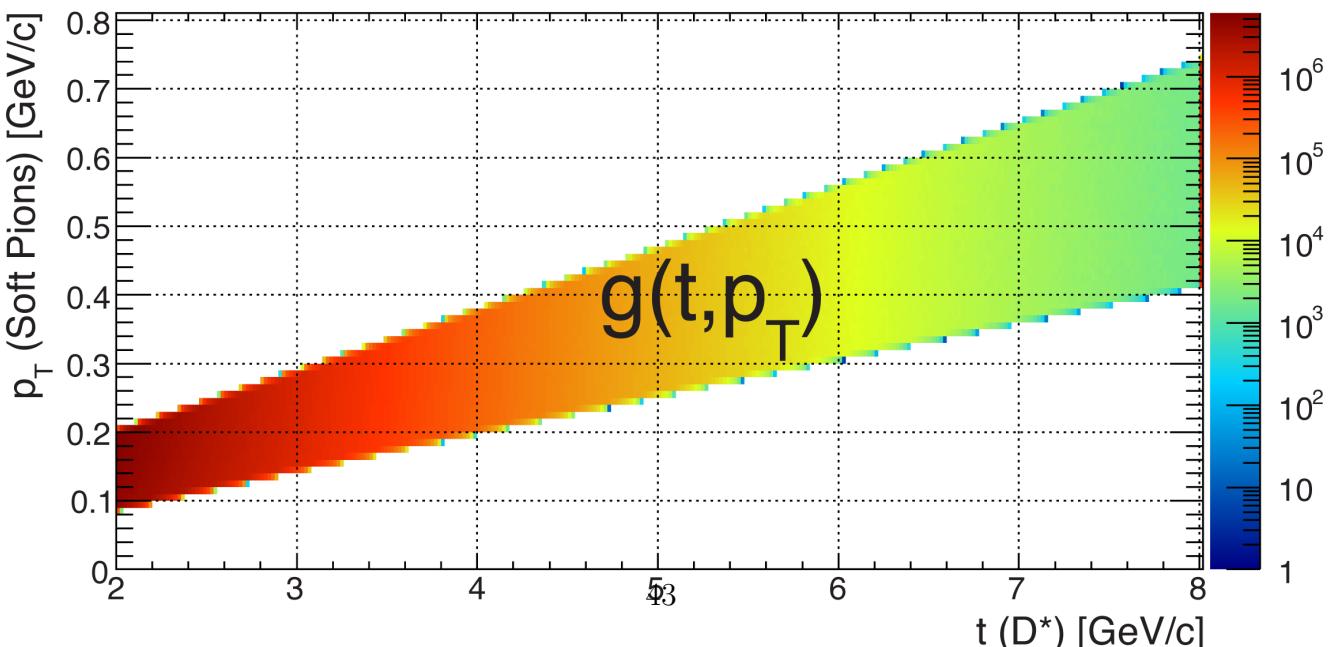
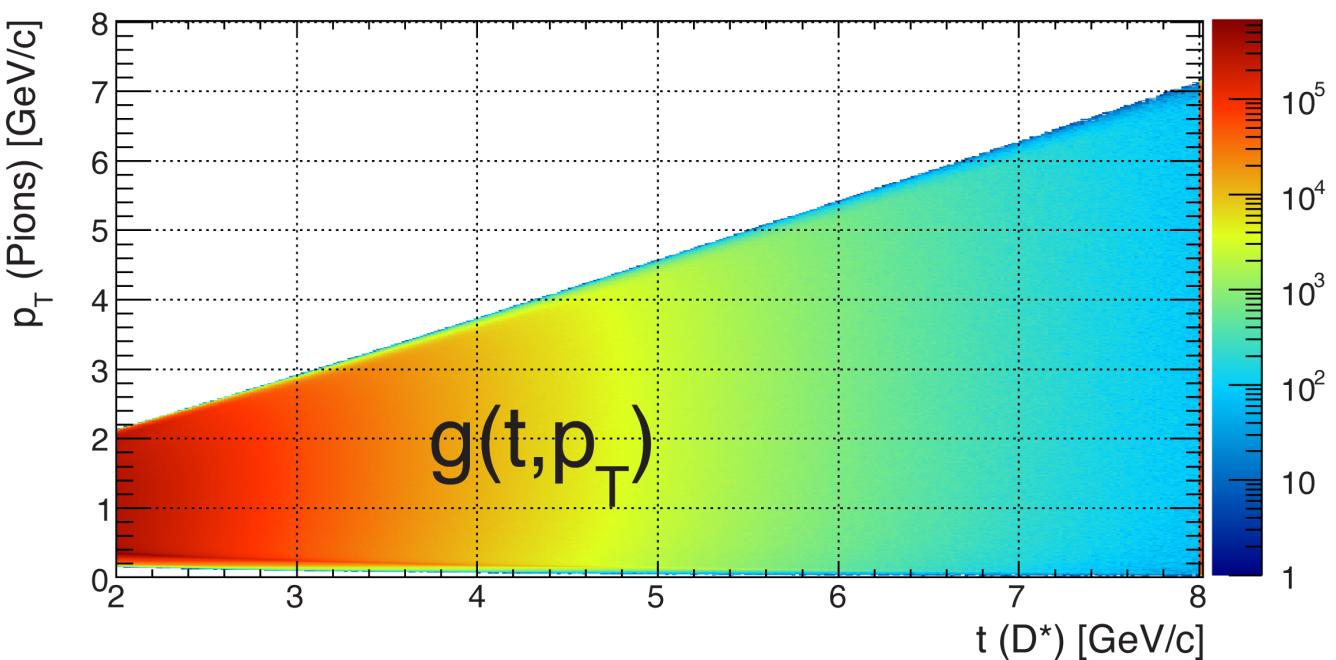
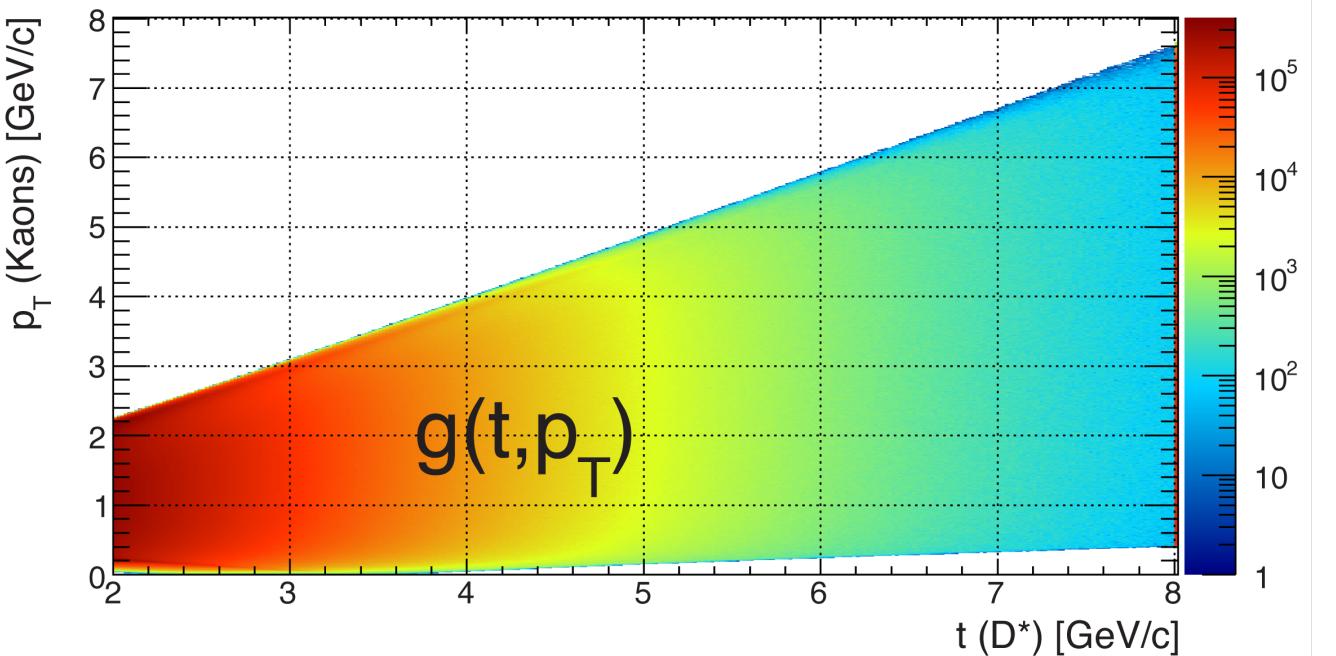


Figure 32: D^* decay kinematics $g(t, p_T)$

D Treatment of Errors

D.1 Treatment of Errors in Efficiency Calculations

An Application of a cut can be considered a Binomial process. $P(k; \varepsilon, n)$ defined as

$$P(k; \varepsilon, n) = \binom{n}{k} \varepsilon^k (1 - \varepsilon)^{n-k} \quad (\text{D.1})$$

denotes the probability that k events will pass the cut, given the conditions that the true efficiency is ε and that there are n events in the sample. The probability density function $P(\varepsilon; k, n)$, which gives the probability function of ε for a given n and k , can be determined by Bayesian theorem:

$$P(\varepsilon; k, n) = \frac{P(k; \varepsilon, n)P(\varepsilon; n)}{C}, \quad (\text{D.2})$$

where C is a constant to be determined by normalization, and $P(\varepsilon; n)$ is the probability assigned for the true efficiency before the data are considered. There is no reason to favor one value of the efficiency over another, so it's reasonable to take

$$P(\varepsilon; n) = \begin{cases} 1 & \text{if } 0 \leq \varepsilon \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

independent of n .

$$\begin{aligned} \int_{-\infty}^{+\infty} P(\varepsilon; k, n) d\varepsilon &= \frac{1}{C} \binom{n}{k} \int_0^1 \varepsilon^k (1 - \varepsilon)^{n-k} d\varepsilon = 1 \\ \int_0^1 \varepsilon^k (1 - \varepsilon)^{n-k} d\varepsilon &= \frac{\Gamma(k+1)\Gamma(n-k+1)}{\Gamma(n+2)} = \frac{k!(n-k)!}{(n+1)!} \\ \frac{1}{C} \frac{n!}{(n-k)!k!} \frac{k!(n-k)!}{(n+1)!} &\Rightarrow C = \frac{1}{n+1} \end{aligned}$$

The final efficiency probability density function thus reads:

$$P(\varepsilon; k, n) = (n+1) \binom{n}{k} \varepsilon^k (1 - \varepsilon)^{n-k} \quad (\text{D.3})$$

The mean value of D.3:

$$\bar{\varepsilon} = \int_0^1 \varepsilon P(\varepsilon; k, n) d\varepsilon = \frac{(n+1)!}{k!(n-k)!} \int_0^1 \varepsilon^{k+1} (1 - \varepsilon)^{n-k} d\varepsilon = \frac{k+1}{n+2}, \quad (\text{D.4})$$

the mode of D.3:

$$\frac{dP}{d\varepsilon} = 0 \Rightarrow \varepsilon = \frac{k}{n}, \quad (\text{D.5})$$

and variance of D.3:

$$\begin{aligned} V(\varepsilon) &= \int_0^1 (\varepsilon - \bar{\varepsilon})^2 P(\varepsilon; k, n) d\varepsilon = \int_0^1 \varepsilon^2 P(\varepsilon; k, n) d\varepsilon - \bar{\varepsilon}^2 \\ &= \frac{(k+1)(k+2)}{(n+2)(n+3)} - \frac{(k+1)^2}{(n+2)^2}. \end{aligned} \quad (\text{D.6})$$

$$V(\varepsilon)|_{k=0,n} = \frac{n+1}{(n+2)^2(n+3)} > 0$$

$$\lim_{n \rightarrow \infty} V(\varepsilon) = 1/n^2$$

To find the shortest interval $[\alpha, \beta]$ which contains probability content $\lambda = 0.683$ (corresponds to $\pm 1\sigma$ interval of a Normal distribution), the interval $\beta - \alpha$ must be minimized, subject to the constraint

$$\int_{\alpha}^{\beta} P(\varepsilon; k, n) d\varepsilon = \lambda$$

A formal solution can be found using the method of Lagrange multipliers. The solution for α and β is found by the simultaneous solutions of the nonlinear equations:

$$\begin{aligned} \mathcal{G} + \rho\alpha^k(1-\alpha)^{n-k} &= 0 \\ \mathcal{G} + \rho\beta^k(1-\beta)^{n-k} &= 0 \\ \mathcal{B}_{\beta}(k+1, n-k+1) - \mathcal{B}_{\alpha}(k+1, n-k+1) &= \lambda\mathcal{G} \end{aligned}$$

where $\mathcal{G} = \frac{\Gamma(k+1)\Gamma(n-k+1)}{\Gamma(n+2)}$, ρ is a Lagrange multiplier introduced for the constraint, and $\mathcal{B}_x(k+1, n-k+1)$ is the incomplete Beta function, defined by

$$\mathcal{B}_x(k+1, n-k+1) = \int_0^x t^k(1-t)^{n-k} dt.$$

The calculation of λ is implemented in the ROOT routine `TGraphAsymmErrors::BayesDivide`.

D.2 Treatment of Systematic Errors

Suppose a yield is estimated from a data sample by two different methods, giving estimates Y_1 and Y_2 , with statistical errors σ_1 and σ_2 . It is required to find error on the difference

$$\Delta = Y_1 - Y_2 \tag{D.7}$$

and thus to establish whether the estimates agree within one standard deviation. The error on the difference is given by

$$\sigma_{\Delta}^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2 \tag{D.8}$$

where ρ is the correlation between the two estimates. If two estimates are, at any rate to some extent, independent, they can be weighted and combined to give an improved estimate that is more efficient than either one separately. Let the joint estimator be

$$Y = wY_1 + (1-w)Y_2 \tag{D.9}$$

where w is some chosen weight. Then the error on Y is given by

$$\sigma^2 = w^2\sigma_1^2 + (1-w)^2\sigma_2^2 + 2w(1-w)\rho\sigma_1\sigma_2. \tag{D.10}$$

Setting the derivative with respect to w to zero to find the minimum of σ gives

$$w = \frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \quad 1-w = \frac{\sigma_1^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \tag{D.11}$$

and the resulting minimum error on Y is given by

$$\sigma = \frac{\sigma_1^2 \sigma_2^2 (1 - \rho^2)}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \quad (\text{D.12})$$

(If the results are uncorrelated this gives the familiar form for the best combination of estimates with different errors).

However no estimator can be more efficient than the Minimum Variance Bound (MVB)

$$\sigma_0 = \frac{1}{N \int \left(\frac{d \ln P(x; Y)}{dY} \right)^2 P(x; Y) dx} \quad (\text{D.13})$$

where $P(x; Y)$ is the probability density function for the distribution being measured. Thus

$$\frac{\sigma_1^2 \sigma_2^2 (1 - \rho^2)}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \geq \sigma_0 \quad (\text{D.14})$$

which gives an expression for ρ which is satisfied only between the two roots of the quadratic

$$\rho = \frac{\sigma_0^2 \pm \sqrt{(\sigma_1^2 - \sigma_0^2)(\sigma_2^2 - \sigma_0^2)}}{\sigma_1\sigma_2} \quad (\text{D.15})$$

The maximum value for σ_Δ comes from the minimum value for ρ , and vice versa

$$\begin{aligned} \sigma_\Delta^{max} &= \sqrt{(\sigma_1^2 - \sigma_0^2)} + \sqrt{(\sigma_2^2 - \sigma_0^2)} \\ \sigma_\Delta^{min} &= \left| \sqrt{(\sigma_1^2 - \sigma_0^2)} - \sqrt{(\sigma_2^2 - \sigma_0^2)} \right| \end{aligned} \quad (\text{D.16})$$

D.3 Treatment of Correllated Errors

Suppose n measurements x_i . Let's define the covariance matrix \mathbb{V} , whose terms are given by

$$\mathbb{V}_{ij} = E[(x_i - \mu_i)(x_j - \mu_j)] = E[x_i x_j] - \mu_i \mu_j$$

where E denotes expected value and μ mean value ($\mu \equiv E(x)$). The diagonal members of \mathbb{V} are the variances, e.g., $\mathbb{V}_{ii} = \sigma_i^2$. The correlation coefficient r_{ij} is given by

$$r_{ij} = \frac{\mathbb{V}_{ij}}{\sigma_i \sigma_j}.$$

The best linear estimator accounting for all measurements can be defined generally as

$$\bar{x} \equiv \sum_i w_i x_i \quad (\text{D.17})$$

where the weights w_i must be found. Assuming that the x_i each provide an unbiased estimate and requiring that the estimator \bar{x} be unbiased gives the condition $\sum_i w_i = 1$. Writing $\bar{x} - E(\bar{x}) = \sum_i w_i(x_i - E(x_i))$ and using the definition of the standard deviation, one can obtain the relation for the standard deviation of \bar{x} :

$$\sigma_{\bar{x}}^2 = \sum_{ij} w_i w_j \mathbb{V}_{ij} \quad (\text{D.18})$$

To minimize $\sigma_{\bar{x}}^2$ subject to condition $\sum_i w_i = 1$, the Lagrange multiplier technique can be used. Writing

$$\sigma_{\bar{x}}^2 = \sum_{ij} w_i w_j \mathbb{V}_{ij} + \lambda \left(\sum_i w_i - 1 \right) \quad (\text{D.19})$$

and setting the derivative of $\sigma_{\bar{x}}^2$ with respect to the w_i and λ to zero, one gets the solution²

$$w_i = \frac{\sum_k \mathbb{V}_{ik}^{-1}}{\sum_{jk} \mathbb{V}_{jk}^{-1}} \quad (\text{D.20})$$

²Actually, \mathbb{V}^{-1} does not have to exist for the weights to be determined. Such singular covariance matrices arise when only a fully correlated error is present. The simplest way to deal with this situation numerically is to add a small term ϵ to the diagonal elements of \mathbb{V} , calculate \mathbb{V}^{-1} , and neglect ϵ afterwards.

References

- [1] Y. Xu *et al.*, Nuclear Instruments and Methods in Physics Research Section A, Volume 614, Issue 1, 21 February 2010, Pages 28-33
- [2] K. Nakamura *et al.*, J. Phys. G **37**, 075021 (2010)
- [3] M.Cacciari, P.Nason and R.Vogt, Phys.Rev.Lett. **95**, 122001 (2005)
- [4] Barbara Trzeciak's PhD Thesis, https://drupal.star.bnl.gov/STAR/files/JpsiPolarization2009STAR_phdThesis_official.pdf
- [5] D. de Florian and R. Sassot, Phys. Rev. D 69, 074028 (2004).
- [6] Riv. Nuovo Cim., 6N10 (1984) 1-50 CERN-TH-3684.
- [7] C. Tsallis, J. Stat. Phys. 52, 479 (1988)
- [8] G. Wilk and Z. Włodarczyk, Eur. Phys. J. A40, 299
- [9] G. Wilk and Ch. Wong, Phys. Rev. D 87, 114007 (2013)
- [10] Vogt R. et al., PoS ConfinementX 203, (2012)