

# CMS Draft Analysis Note

The content of this note is intended for CMS internal use and distribution only

2021/11/20

Archive Hash:

Archive Date:

## B meson $R_{AA}$ and Cross Section Ratios in $pp$ and PbPb Collisions at 5.02 TeV

Zhaozhong Shi<sup>1</sup>, Maria Faria<sup>2, 1</sup>, Ran Bi<sup>1</sup>, Nuno Leonardo<sup>2</sup>, and Yen-Jie Lee<sup>1</sup>

<sup>1</sup> Massachusetts Inst. of Technology

<sup>2</sup> Técnico Lisboa

### Abstract

Your abstract here. (One paragraph; max length roughly 200 words (longer will be truncated for EPJC submissions). Absolute max (at arXiv) is 1920 characters.)

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor:

"376"377"000Z"000h"000a"000o"000z"000h"000o"000n"000g"000"040"000S"000h"000i"000J"000i"000e"000"040"000L"000e"000e

PDFTitle:

"376"377"000B"000"040"000m"000e"000s"000o"000n"000"040"000R"000A"000A"000"040"000C"000M"000S

PDFSubject:

"376"377"000C"000M"000S"000,"000"040"000y"000o"000u"000r"000"040"000t"000o"000

PDFKeywords:

"376"377"000C"000M"000S"000,"000"040"000y"000o"000u"000r"000"040"000t"000o"000

Please also verify that the abstract does not use any user defined symbols



# Contents

1	1	Introduction . . . . .	2
2	2	Datasets and Event Selections . . . . .	3
3	2.1	Datasets . . . . .	3
4	2.2	MC samples . . . . .	3
5	2.2.1	MC reweighting . . . . .	3
6	3	$B^+$ meson reconstruction . . . . .	5
7	3.1	Muon and $J/\psi$ selection . . . . .	5
8	3.2	Track selection . . . . .	5
9	4	$B^+$ meson selection . . . . .	6
10	4.1	Cut optimization . . . . .	6
11	4.2	Fiducial region . . . . .	16
12	5	MC-Data Comparison . . . . .	18
13	5.1	B candidate properties comparison between data and MC . . . . .	18
14	6	Signal Extraction . . . . .	20
15	6.1	Closure test of the fitting procedure . . . . .	23
16	7	Acceptance and Efficiency . . . . .	25
17	8	Closure Test of the $B^+$ Corrected Yield . . . . .	26
18	8.1	Raw Yield Closure . . . . .	26
19	8.2	Efficiency Closure . . . . .	26
20	8.3	Corrected Yield Closure . . . . .	26
21	9	Results . . . . .	27
22	9.1	$p_T$ differential corrected yield in pp collisions at 5.02 TeV . . . . .	27
23	9.2	Rapidity differential corrected yield in pp collisions at 5.02 TeV . . . . .	28
24	9.3	Multiplicity differential corrected yield in pp collisions at 5.02 TeV . . . . .	28
25	9.4	$B_s^0/B^+$ ratio as a function of $p_T$ in pp collisions at 5.02 TeV . . . . .	28
26	9.5	$B_s^0/B^+$ ratio as a function of $y$ in pp collisions at 5.02 TeV . . . . .	28
27	9.6	$B_s^0/B^+$ ratio as a function of Multiplicity in pp collisions at 5.02 TeV . . . . .	28
28	9.7	B-meson $R_{AA}$ vs $p_T$ at 5.02 TeV . . . . .	28
29	9.8	B-meson $R_{AA}$ vs $y$ at 5.02 TeV . . . . .	28
30	9.9	B-meson $R_{AA}$ vs Multiplicity at 5.02 TeV . . . . .	28
31	10	Systematic uncertainties . . . . .	29
32	10.1	Summary table . . . . .	29
33	10.2	$T_{AA}$ and $N_{MB}$ . . . . .	31
34	10.3	Branching ratio . . . . .	31
35	10.4	Tracking efficiency . . . . .	31
36	10.5	Muon efficiency: Tag and Probe . . . . .	31
37	10.6	MC-Data Discrepancy . . . . .	34
38	10.7	$p_T$ shape: Bpt weight . . . . .	35
39	10.8	MC stats: Toy MC study . . . . .	37
40	10.9	Signal extraction: PDF variation . . . . .	38
41	A	Non-prompt $J/\psi$ background . . . . .	39
42			

DRAFT

## 2 Datasets and Event Selections

### 2.1 Datasets

This analysis is performed using the 2017 pp data at  $\sqrt{s_{NN}}=5.02$  TeV. The Primary dataset, triggers, and event selections we use are the same with  $B_s$  analysis ongoing parallel. For more information, please check the corresponding sections on  $B_s$  analysis note (CMS AN-19-055) [? ].

The pp dataset corresponds an integrated luminosity of  $1.5 \text{ nb}^{-1}$ . The analysis uses the dimuon particle datasets (*DoubleMu* PD) for pp collisions. A complete list of the used datasets can be found in Table ?? . The full name of the used datasets are as follows.

To extract pure collision events, several offline selections are applied to each event. Events are required to have at least one reconstructed primary vertex(PV). The primary vertex is formed by two or more associated tracks and is required to have a distance from the nominal interaction region of less than 15 cm along the beam axis and less than 0.15 cm in the transverse plane. For pp, an additional selection of hadronic collisions is applied by requiring a coincidence of at least 3 HF calorimeter towers, with more than 3 GeV of total energy, from the HF detectors on both sides of the interaction point. In addition to HF coincidence, a cluster compatibility filter is used in PbPb analysis. Details of these filters can be found in Section 2.2 of [? ].

### 2.2 MC samples

This section is identical to the  $B_s$  AN. The only difference is that we use different decay channels for reconstruction of B mesons. Here we use  $B^+ \rightarrow J/\psi K^+$  channel.

#### 2.2.1 MC reweighting

We use dedicated method to calculate the acceptance and efficiency correction factor, by taking average over data which is described in the section. ?? . The  $p_T$  weight we get here is not used for our nominal result, since the average over data already includes the information of data  $p_T$  shape. However, the  $p_T$  weight we obtain here is used to check the robustness of the method, and will be quoted as systematics uncertainties related with the method.

The  $p_T$  spectrum of the  $B^+$  MC was reweighted in order to match with the data. We first evaluated the raw yield of the data and MC without any  $p_T$  weight in  $p_T$  range 5-60GeV, and normalized them. Then we took the ratio of the two to get the first  $p_T$  weight distribution. This distribution is fitted with polynomials, and we found that the 4th order polynomial have the minimum  $\chi^2$  among the lowest order polynomials, thus we used this as our weight function.

After then, we examined the new raw yield of MC with the new weight, and compared the distributions to get a new weight functions. We iterated this twice, and we found that the first  $p_T$  weight obtained already shows good agreement with data. Thus we use that  $p_T$  weight as our weight function.

In Fig. ?? , the normalized distributions of data and MC raw yield and their ratios are presented. The second row (with 1st  $p_T$  weight) already shows reasonable agreement between data and MC. The red fit function on the top right plot is our nominal  $p_T$  weight function.

pp MC simulations are also reweighted in order to match the centrality distribution in data. On the left panel of Figure ?? , the centrality distribution of the MC simulation (red) is compared to the one in data (blue). On the right panel, MC is given the Ncoll weight and compared to the data. The unit (HiBin) on the x-axis corresponds 0.5% centrality.

In addition to the Bpt and centrality reweighting, it is known that the MB samples used for

87 embedding pp signal MC samples (with Cymal5Ev8 tune) has an offset in the primary vertex  
88 z position(PVz). This deviation is corrected by giving PVz weight. The procedure are shown in  
89 Fig. ???. The overall weight is given by the ratio between the two Gaussian functions. Note that  
90 this analysis is not sensitive to the absolute value of the PV position because the reconstruction  
91 of the  $B^+$  meson rely only on the relative distance between PV and  $B^+$  reconstructed vertex  
92 which is presented in the following section. We tried to estimate this effect by removing this  
93 re-weighting entirely and found that the difference is only 1.3%.

DRAFT

### 3 $B^+$ meson reconstruction

In this section, the  $B^+$  meson reconstruction strategy is presented. The schematic diagram of the workflow is shown in Fig. 1, starting from muons and tracks to  $B^+$  meson candidates. Muon candidates and tracks are required to pass several quality selection criteria as described in Section 3.1 and Section 3.2.  $J/\psi$  candidates are reconstructed by vertexing muon pairs with opposite charge, using `KinematicConstrainedVertexFitter`. The  $B^+$  candidates are built by combining the  $J/\psi$  candidates with each of the selected tracks. Finally, a kinematic fit to the  $J/\psi$ - $K^+$  system is performed, forcing the mass of dimuon pair to be equal to the nominal  $J/\psi$  mass based on PDG [? ]. The selection criteria of the  $B^+$  candidates are described in Section 4.

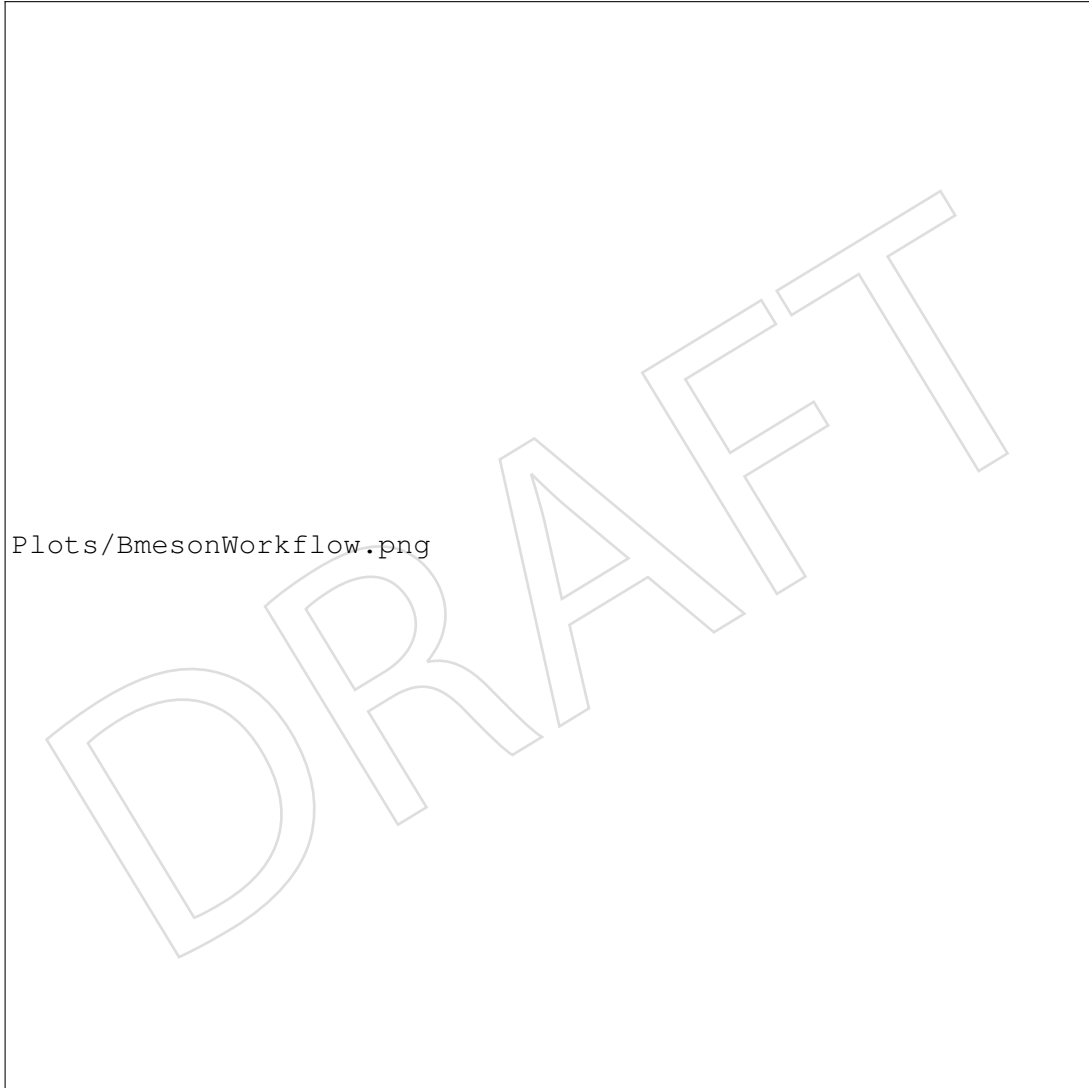


Figure 1: Schematic diagram of the B meson reconstruction workflow

#### 3.1 Muon and $J/\psi$ selection

We applied the same muon selection on  $B^+$  with  $B_s^0$  which is written exclusively on  $B_s^0$  AN [? ].

#### 3.2 Track selection

The track selection we applied are the same with  $B_s^0$  AN [? ].

## 4 $B^+$ meson selection

The selections described in the previous section (so-called prefilter) are not sufficient to distinguish signal from background, especially at low  $p_T$  range where the entire spectrum is dominated by combinatorial backgrounds which arises from random combination of muons and tracks. In order to see clear signals and reduce the uncertainties, several additional selection on the  $B^+$  decay topology were then applied. A cut optimization procedure was performed to decide the cut values. For more details, see Section 4.1.

### 4.1 Cut optimization

The goal of the optimization procedure is to maximize the statistical significance of the signals while keeping reasonably high signal efficiencies. The optimal cut that minimizes background efficiency for a specific signal efficiency is obtained by the TMVA(Toolkit for Multivariate Data Analysis with ROOT) [? ].

Boosted Decision Tree(BDT) is chosen to be the classification method in TMVA training for  $B^+$ . In addition, for comparison the CutsGA (Genetic algorithm-based rectangular cut decision) is also examined. We found that in the lowest  $p_T$  bin(5-7GeV), the prefilter itself does not filter out background sufficiently, which limits the performance of BDT method. In order to enhance the performance, we first found the optimal rectangular cuts from CutsGA method in this  $p_T$  bin. After then we used normalized decay length(dls3D) and  $\cos\theta$ (2D angle between  $B^+$  meson displacement and  $B^+$  meson momentum in transverse plane) cuts ( $dls3D > 12.0$   $\cos\theta > 0.95$ ; this is slightly looser than the optimal rectangular cuts, in order not to reject possible signals that BDT can reconstruct but CutsGA cannot) in our prefilter for BDT training. This enhanced the raw yield extraction and reduced the yield error. We also applied this strategy to the  $p_T$  7-10 GeV/c bin with the selection ( $dls3D > 9.0$   $\cos\theta > 0.92$ ) to enhance the performance. For other  $p_T$  bins, we just used the original prefilter for our BDT training.

The reconstructed candidates matched to the generated signal in MC sample are used as signal when training, while the reconstructed candidates in sideband ( $0.15 \text{ GeV}/c^2 < |M_{B^+} - M_{B^+}^{PDG}| < 0.25 \text{ GeV}/c^2$ ) of data sample are used as background.

The kinematic variable distributions of daughter tracks and  $B^+$  mesons before and after applying the prefilter is shown on Figure 2, 3, 4, 5





Figure 2: The normalized  $J/\psi \mu^-$  kinematic variable distributions before and after prefilter for  $p_T = 7 - 10$  GeV/c (left) and  $10 - 50$  GeV/c (right) are shown above.



Figure 3: The normalized  $J/\psi \mu^+$  kinematic variable distributions before and after prefilter for  $p_T = 7 - 10$  GeV/c (left) and 10 - 50 GeV/c (right) are shown above.



Figure 4: The normalized  $K^+$  track, muons, and  $K^+$  track kinematic variable distributions before and after prefilter for  $p_T = 7 - 10$  GeV/c (left) and  $10 - 50$  GeV/c (right) are shown above.

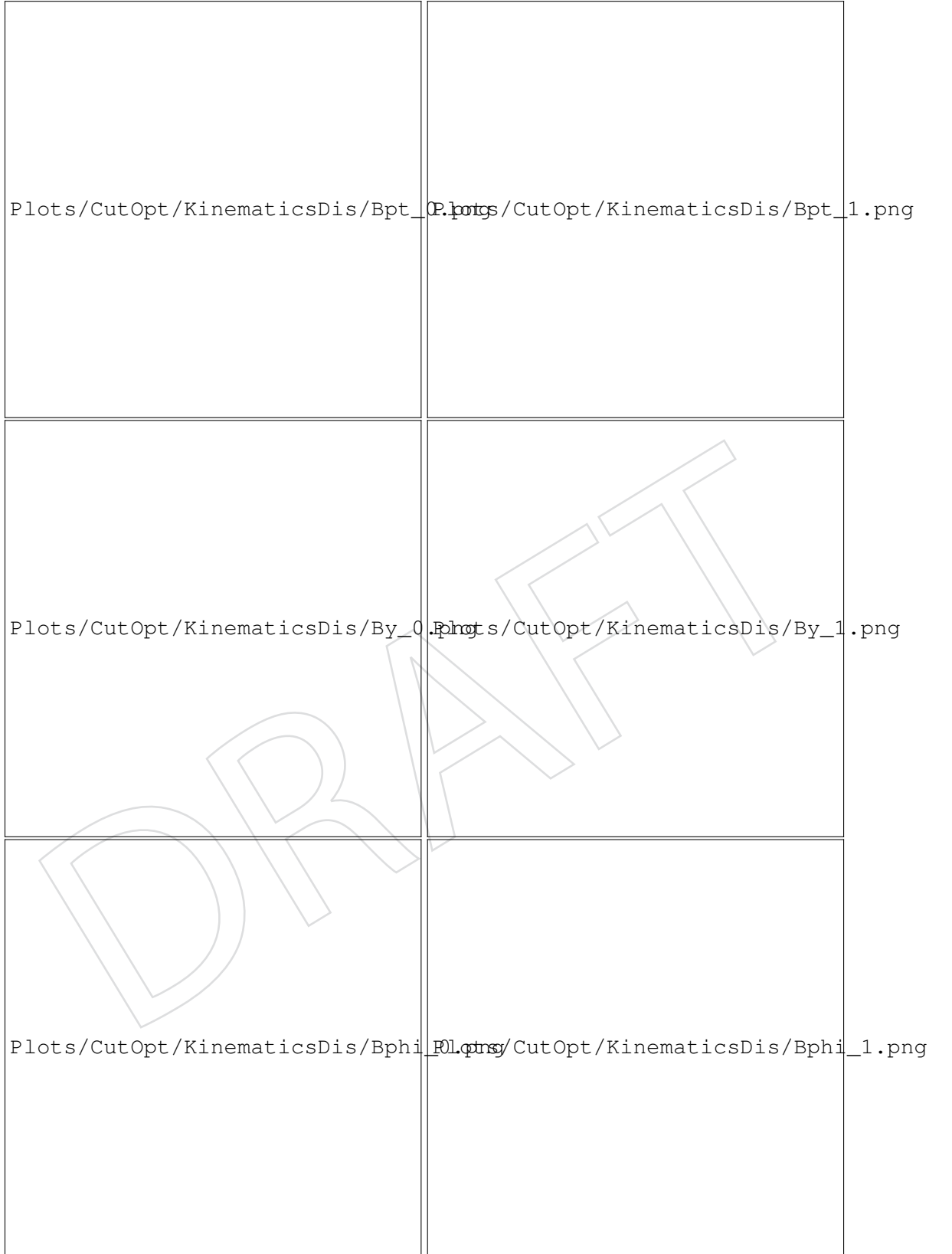


Figure 5: The normalized  $B^+$  kinematic variable distributions before and after prefilter for  $p_T = 7 - 10$  GeV/c (left) and  $10 - 50$  GeV/c (right) are shown above.

We can see that the prefilter cuts do not significantly change the distribution of the kinematic variable shapes and create potential bias. Therefore, these show that our prefilter cut is valid to use.

Then, we use the following six selection variables trained for PbPb:

- dls3D: Normalized decay length (SV PV distance), the distance between primary and  $B^+$  decay (secondary) vertex normalized by its uncertainty
- Balpha: The angle between  $B^+$  meson displacement and  $B^+$  meson momentum in 3D
- Btrk1Pt: Track  $p_T$ , the transverse momentum of the track
- Bchi2cl:  $\chi^2$  probability, the  $\chi^2$  probability of the secondary decay vertex fitting
- Btrk1Eta: The absolute value of the track pseudorapidity
- Btrk1Dxysig: Normalized track Dxy, the transverse distance between track and the primary vertex normalized by its uncertainty

The BDT training setup for all  $p_T$  bins are as follows: NTree = 850, MinNodeSize=2.5%, MaxDepth=3, BoostType=AdaBoost, AdaBoostBeta=0.5, UseBaggedBoost, BaggedSampleFraction=0.5, SeparationType=GiniIndex, nCuts=20

The optimal BDT cut values are defined as the numbers which maximize the figure of merit (statistical significance)  $S/\sqrt{S+B}$ . Here,  $S$  is the number of signal in signal region after applying optimal cuts, while  $B$  is the number of background in signal region after applying optimal cuts. Signal region is defined as  $|M_{B^+} - M_{B^+}^{PDG}| < 0.08 \text{ GeV}/c^2$ . For the lowest  $p_T$  bin, we used a new figure of merit  $S/\sqrt{\gamma_{\text{quantile}}(\alpha/2, S+B, 1)}$  (where,  $\alpha$  is chosen to be 1-0.6827) that is suitable for this low signal/background ratio.

- $S = S' \times (\text{signal optimal cut efficiency})$ , where  $S'$  is the number of signal in signal region before applying optimal cuts.
- $B = B' \times (\text{background optimal cut efficiency})$ , where  $B'$  is the number of background in signal region before applying optimal cuts.

$S'$  and  $B'$  are calculated by fitting the invariant mass plot with prefilter selection, with the same functional form used in the main analysis. However, at the 4 lowest  $p_T$  bins, the prefilter is not sufficient to reveal the signal. In those cases,  $S'$  is calculated by the expected number of signal from FONLL pp cross-section calculation multiplied by pre-filters efficiency, acceptance from MC, and expected  $R_{AA}$  value from previous measurement [? ].

Fig. 6 and Fig. 7 show the correlation matrices of the training, in the lowest and highest  $p_T$  bin, respectively.



Figure 6: TMVA training correlation matrices of PbPb in  $p_T$  5 - 7 GeV/c.



Figure 7: TMVA training correlation matrices of PbPb in  $p_T$  50 - 60 GeV/c.

Fig. 8 shows the overtraining check in the lowest and the highest  $p_T$  bin.



Figure 8: TMVA overtraining check of PbPb in  $p_T$  5-7 and 50 - 100 GeV/c, respectively

Fig. 9 shows the variable distribution comparison between signal and background, and the statistical significance curve in  $p_T$  15-20GeV. Some variables show prominently distinct distribution between signal and background.



Figure 9: TMVA training variable distribution of PbPb in  $p_T$  15 - 20 GeV/c.



The maximum significance point is selected as our nominal working point, as shown in Fig. 10. Tab. 1 shows the summary of the optimal selection criteria in different  $p_T$  intervals. All the other TMVA performance plots not listed in this section are on Appendix. ??.

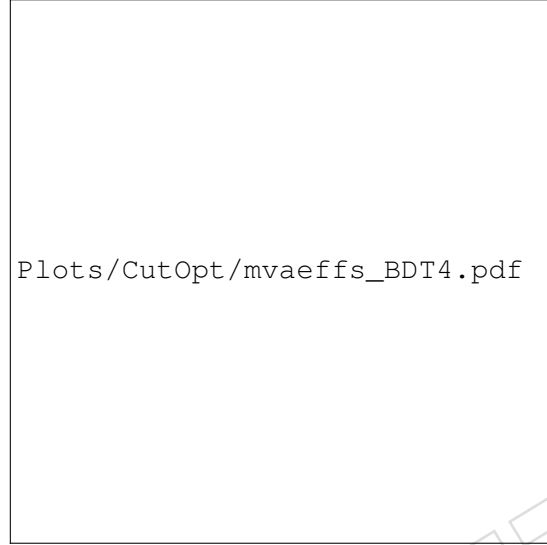


Figure 10: TMVA significance curve of PbPb in  $p_T$  15 - 20 GeV/c.

$p_T$ (GeV/c)	5-7	7-10	10-15	15-20	20-30	30-40	40-50	50-60
BDT	>0.02	>0.03	>0.09	>0.07	>0.10	>0.16	>0.20	>0.27

Table 1: Summary table of selection criteria in different  $p_T$  intervals in PbPb collisions

## 4.2 Fiducial region

As we go to low  $B^+ p_T$  where the daughter particles also have predominantly low  $p_T$ , it becomes improbable to detect muons in low rapidity ranges due to the limited acceptance of the muons. In other words, the signal and background candidates of  $B^+$  are largely confined to forward rapidity region at low  $B^+ p_T$ . Fig 11 shows theoretically possible inverse of acceptance, selection efficiency, and total efficiency vs  $B^+ p_T$  and  $|y|$  2D distribution with only single muon and track selections but without BDT selections.



Figure 11: The two dimensional maps of the inverse of the acceptance efficiency (left), selection efficiency (middle), and total efficiency (right) vs  $B^+ p_T$  and rapidity in MC for the centrality range of 0 - 90% without applying BDT cuts are shown above.

Next, we quantitatively examine the 2D maps of candidate distributions in both data and MC after applying the optimal BDT selections, which are shown in Fig. 12.

It is found that  $B^+$  acceptance is better than  $B_s^0$ , presumably because of less number of tracks needed to reconstruct B meson, which allows broader phase space that the daughter particles can reside. Since this analysis is mainly for the comparison with  $B_s^0$ , we adopted more strict fiducial region ( $5 < p_T < 10\text{GeV}$ ,  $1.5 < |y_{lab}| < 2.4$ ) that is defined using  $B_s^0$  2D map.



Figure 12:  $B^+$   $p_T$  and rapidity two dimensional map of candidates that passed the full analysis cut in data (Left) and MC (right).

## 5 MC-Data Comparison

### 5.1 B candidate properties comparison between data and MC

Potential differences between Data and MC distributions of the variables used for the selection of the  $B^+$  signal, can introduce biases in the efficiency corrections. In this section, the distributions of the  $B^+$  selection variables in PbPb data and MC samples are compared.

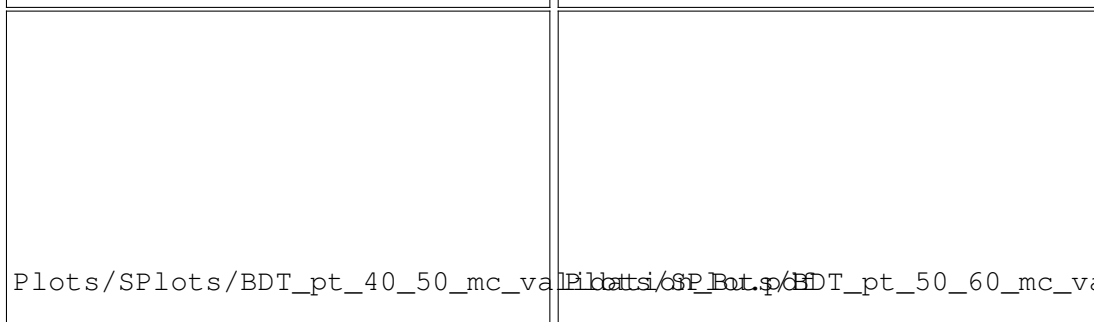
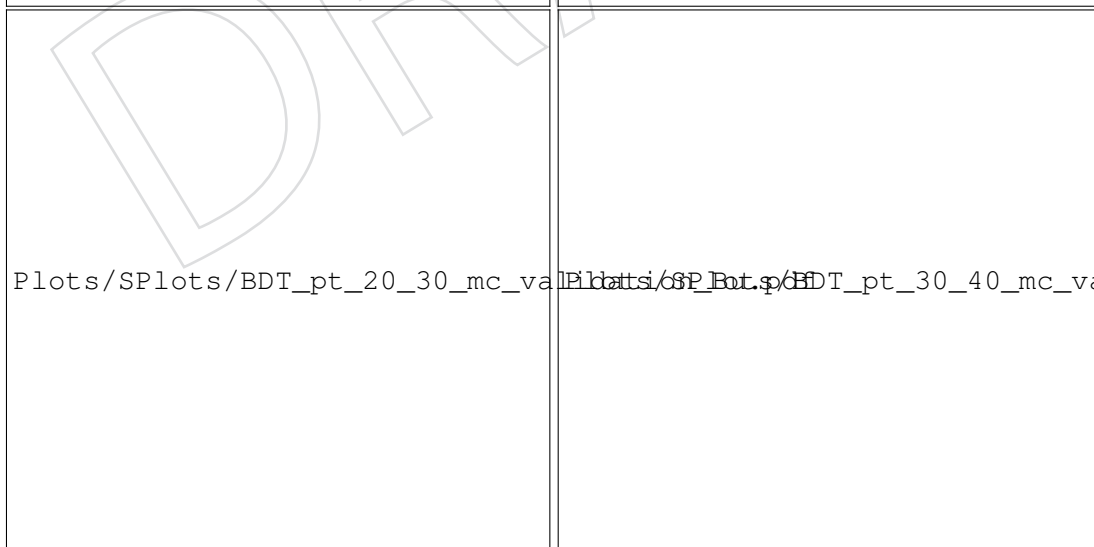
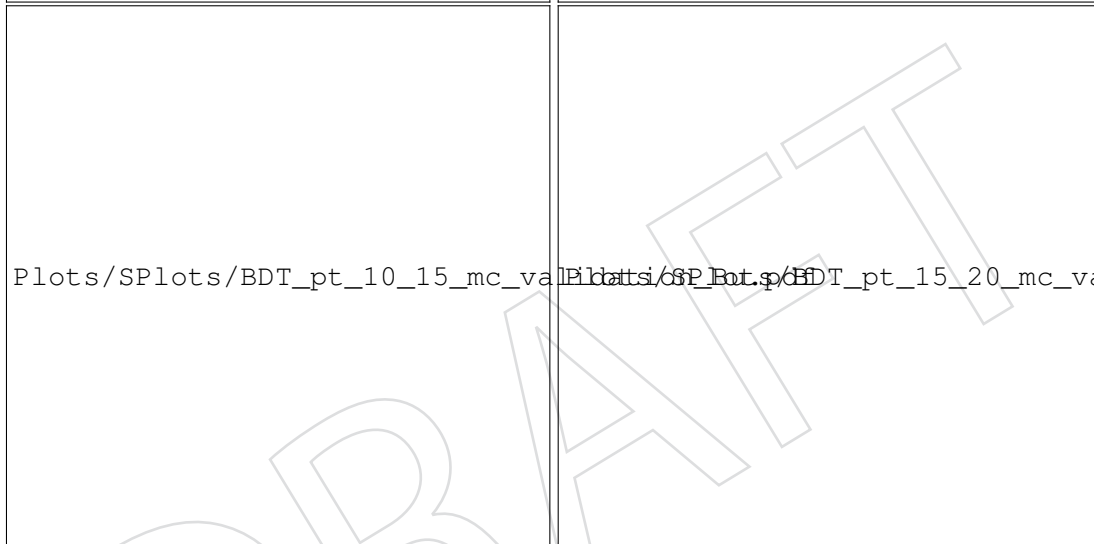
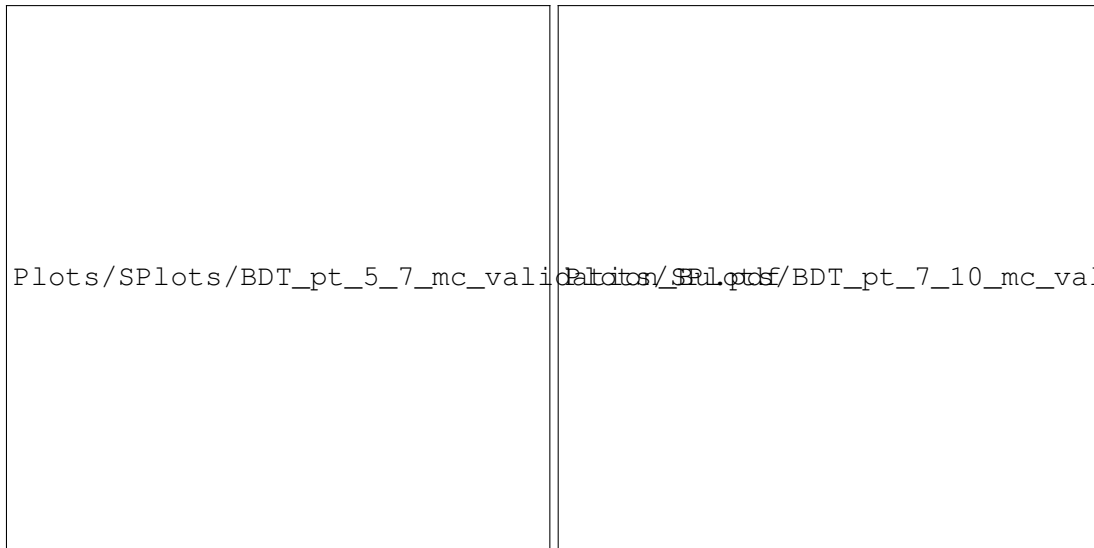
The Splot method is used for our analysis. This is a likelihood-based method by which we reweight the data using the unbinned fit result. The weights are added to the dataset based on model and yield extraction variables. Each event has two weights: probability of belonging to the signal given its mass, probability of belonging to the background given its mass. The Splot class gives us the distributions of our variables for a given species (signal or background). The advantage of using this method is that we use the full dataset for the comparison in contrast to the sideband subtraction method where one should select the investigation range of signal and background. Furthermore, we use likelihood to describe events' behavior in contrast to the potential misidentification of signal events in background region which might occur in sideband subtraction method. For more details on general description of Splot method applied to this analysis, please refer to the section 4 and 5 of the analysis note CMS AN-19-219( [? ]).

To show the correlations between the BDT variables to the  $B^+$  invariant mass and validate our Splot techniques approach, we first make the correlation matrices for BDT variables vs  $B^+$  invariant mass Bmass for data and MC as follows in Figure 13:



Figure 13: The correlation matrices in data (left) and MC (right) are shown above.

Fig. 14 shows the data and MC comparison results based on Splot method. Here, we focus on BDT values that are directly used in our signal extraction and related to MC distribution validation, rather than the variables themselves used in BDT training. In a wide ranges of BDT, the two distributions show good agreement. We only focus on the region where BDT is greater than the working point and is smaller than the maximum value that candidates have.



## 6 Signal Extraction

The method and the fit model for signal and background we applied are mostly the same with  $B_s$  AN. Please refer to the corresponding section in [? ].

However, there is one major difference between  $B_s$  and  $B^+$  yield extraction. From the inclusive  $J/\psi$  sample study, we have found that there is a clear and sizable contribution from non-prompt  $J/\psi$  candidates that are fed down from the  $B^+$  signal in our region of interest. We model this component with an error function and a Gaussian. The error function component mostly comes from 4-prong B meson decay (e.g.  $B^+ \rightarrow J/\psi K^*(892)^+$ ) that are partially reconstructed as  $B^+ \rightarrow J/\psi K^+$  (one track is lost) and can form peaking structures for values of the invariant mass below  $\approx 5.20 \text{ GeV}/c^2$ . The Gaussian contribution mainly comes from  $B^+ \rightarrow J/\psi \pi^+$  where pion is misidentified as kaon.

The ratio of the height of error function and the Gaussian is fixed by MC simulation. However, the Gaussian component is shown to be relatively negligible compared to the error function, and even more to the signal double Gaussian of our nominal channel. More details on the non-prompt  $J/\psi$  study can be found in Appendix Sec. A.

Note that we do not discriminate  $B^+$  from  $B^-$  in raw yield extraction process. Therefore, the invariant mass plots shown below contains both  $B^+$  and  $B^-$  signal and backgrounds. This double counting is properly canceled in the corrected yield calculation process later.

In Fig. 15 and Fig. 16, the invariant mass spectra and their pull distributions obtained in the  $p_T$  intervals in the PbPb analyses are presented.

We also make the comparison between the symmetric raw yield error and their asymmetric upper and lower yield error using the RooFit framework on the unbinned fit. The table is shown below in Table 2:

Table 2: The comparison between RooFit and unbinned fit framework.

Centrality	$p_T$ (GeV/c)	Raw Yield Error	RawYield Error Up	Raw Yield Error Down
0 - 90%	7 - 10	10.88	11.20	10.66
0 - 90%	10 - 15	20.68	21.05	20.35
0 - 90%	15 - 20	17.86	18.21	17.56
0 - 90%	20 - 50	19.84	20.19	19.55
0 - 30%	10 - 50	27.72	28.01	27.45
30 - 90%	10 - 50	19.53	19.92	19.20
0 - 90%	10 - 50	33.78	34.13	33.47



Figure 15: Invariant mass distribution of  $B^+$  candidates obtained in PbPb collisions in  $p_T$  intervals in the transverse momentum range from 7-50 GeV/c and Centrality 0-90%.



Figure 16: Invariant mass distribution of  $B^+$  candidates obtained in PbPb collisions in centrality intervals 0 - 30%, 30 - 90%, and 0 - 90% in the transverse momentum range from 10 - 50 GeV/c.



## 6.1 Closure test of the fitting procedure

In order to validate the yield extraction procedure, we generate 5000 toy MC for the fit and make the pull distribution. Then we perform the Gaussian fits to the pull distribution to obtain the mean and width. The results are shown in Fig 17 and Fig 18:



Figure 17: The pull distribution and the Gaussian fits for 0 - 90% at 7 - 10, 10 - 15, 15 - 20, 20 - 50 are shown respectively above.



Figure 18: The pull distribution and the Gaussian fits for 0 - 90% at 7 - 10, 10 - 15, 15 - 20, 20 - 50 are shown respectfully above.

248 We can see that all  $p_T$  and centrality bins have zero mean and unit width from the Gaussian fits  
 249 to the pull distribution. This validate the closure of our fits to extract the  $B^+$  raw yield.

---

## 250 7 Acceptance and Efficiency

DRAFT

251 **8 Closure Test of the  $B^+$  Corrected Yield**

252 **8.1 Raw Yield Closure**

253 **8.2 Efficiency Closure**

254 **8.3 Corrected Yield Closure**

DRAFT

## 9 Results

### 9.1 $p_T$ differential corrected yield in pp collisions at 5.02 TeV

The method we applied is the same with  $B_s$  AN [? ].

In Fig. 19, Number of MinBias events-normalized, TAA-normalized, efficiency-corrected and  $p_T$ -width divided yield (let's call this "normalized corrected yield") of  $B^+$  mesons in the centrality 0-90% and rapidity region  $|y_{lab}| < 2.4$  (Caveat: Note the fiducial region  $1.5 < |y_{lab}| < 2.4$  in  $B^+$   $p_T$  7 - 10GeV/c) in PbPb collisions at  $\sqrt{s} = 5.02$  TeV is presented. The boxes around the data points represent the total systematic uncertainties.

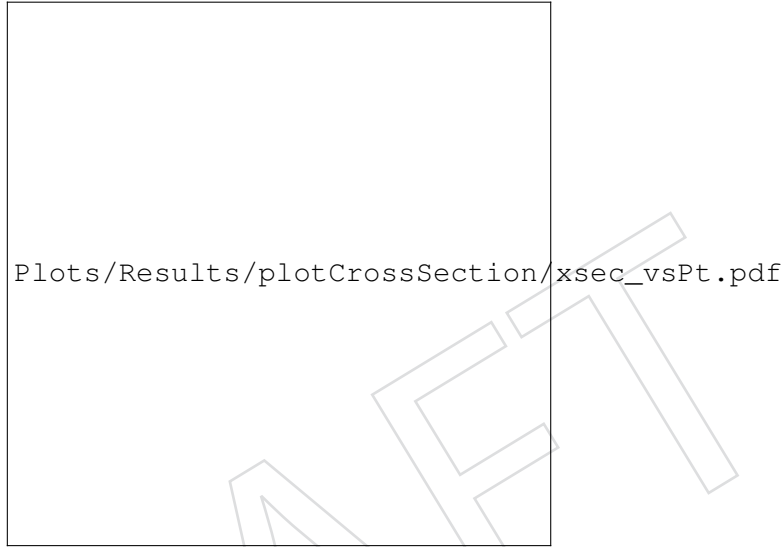


Figure 19:  $B^+$  normalized corrected yield vs.  $p_T$  in the inclusive centrality range 0-90%. (Left)  $B^+$  analysis  $p_T$  bin (Right)  $B_s^0$  analysis  $p_T$  bin. Note the fiducial region  $p_T$  5-10GeV and  $1.5 < |y_{lab}| < 2.4$

In addition, we provide the normalized corrected yield for inclusive  $p_T$  10-50 GeV/c in 0 - 30%, 30 - 90%, and 0 -90% centrality bins in Fig. 20.

The normalized corrected yields and their uncertainties for various kinematic ranges are summarized in Table. 3 and Table. 4. For more details on systematic uncertainties, see Section. 10.

Table 3:  $B^+$  normalized corrected yields and the uncertainties in various  $p_T$  ranges within centrality 0-90%. All the corrected yield values are in units of pb\*c/GeV. Note that the global uncertainties are not included here on this table, which are sub 4% level. Note also the fiducial region  $p_T$  5-10GeV and  $1.5 < |y_{lab}| < 2.4$  and the same as the  $B_s$   $p_T$  binning is reported.

$p_T$ (GeV/c)	Center (GeV/c)	Corr. yield (pb GeV <sup>-1</sup> c)	Stat. up	Stat. down	Syst. up	Syst. down
(7,10)	8.73	331984	15.8%	15.1%	13.6%	13.6%
(10,15)	12.4	298062	7.60%	8.77%	19.1%	19.1%
(15,20)	17.2	66567.7	6.54%	6.66%	7.51%	7.51%
(20,50)	27.3	7730.12	6.64%	5.49%	6.02%	6.02%

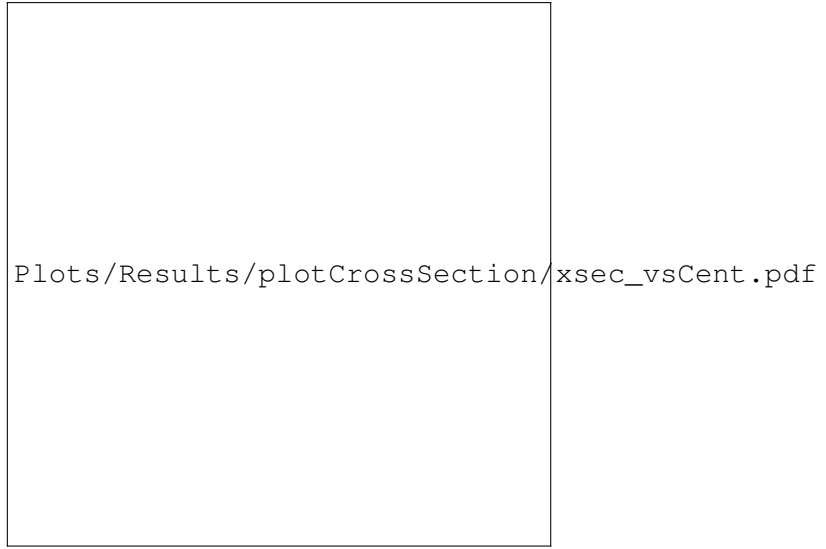


Figure 20:  $B^+$  normalized corrected yield vs.  $\langle N_{part} \rangle$  in inclusive  $p_T$ . 10-50GeV is shown above.

Table 4: Summary table of  $B^+$  corrected yield for different centrality bins

Centrality	Corr. yield (pb)	stat. up (+)	stat. down (-)	syst.. up (+)	syst. down (-)
0% - 30%	1918800	5.81%	5.54%	16.6%	16.6%
30% - 90%	2481820	8.22%	6.97%	13.6%	13.6%
0 - 90%	2092040	5.06%	4.54%	15.2%	15.2%

## 9.2 Rapidity differential corrected yield in pp collisions at 5.02 TeV

## 9.3 Multiplicity differential corrected yield in pp collisions at 5.02 TeV

## 9.4 $B_s^0/B^+$ ratio as a function of $p_T$ in pp collisions at 5.02 TeV

## 9.5 $B_s^0/B^+$ ratio as a function of $y$ in pp collisions at 5.02 TeV

## 9.6 $B_s^0/B^+$ ratio as a function of Multiplicity in pp collisions at 5.02 TeV

## 9.7 B-meson $R_{AA}$ vs $p_T$ at 5.02 TeV

## 9.8 B-meson $R_{AA}$ vs $y$ at 5.02 TeV

## 9.9 B-meson $R_{AA}$ vs Multiplicity at 5.02 TeV

## 10 Systematic uncertainties

### 10.1 Summary table

Below are the summary tables of various systematic uncertainties. For detailed description of each systematics source, please refer to the subsequent subsections.

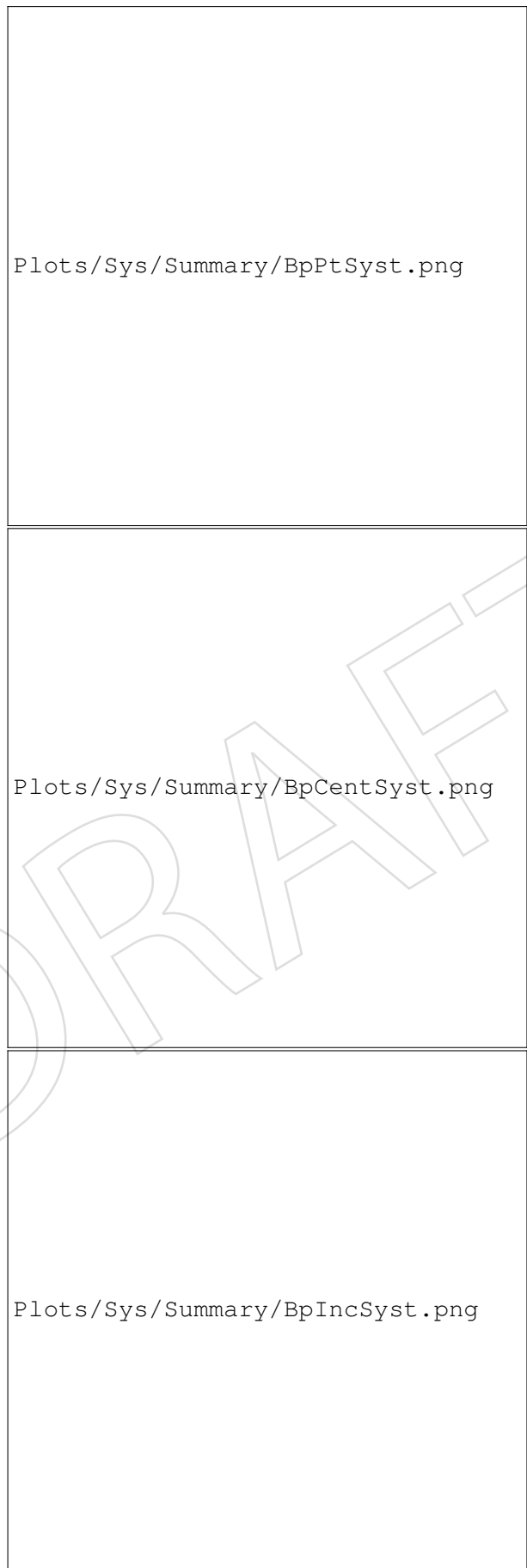
Table 5: Summary of systematic uncertainties from each factor in  $B^+$  PbPb analysis for corrected yield vs  $p_T$ . All the values are shown in percentage.

Factors	(7,10)	(10,15)	(15,20)	(20,50)
Hadron tracking efficiency	5%	5%	5%	5%
Data-MC Discrepancy	4.17%	15.25%	3.01%	1.65%
$p_T$ shape	0.162%	0.211%	0.010%	0.008%
PDF variation background	4.46%	2.67%	2.74%	2.36%
PDF variation signal	0.117%	0.546%	0.576%	1.03%
TnP Systematics	6.12%	9.36%	3.28%	0.34%
MC stat.	9.22%	3.36%	1.92%	1.35%
Total	13.59%	19.08%	7.51%	6.02%
$N_{MB}$ events	1.26%	1.26%	1.26%	1.26%
$T_{AA}$	2.2%	2.2%	2.2%	2.2%
Branching fractions	2.9%	2.9%	2.9%	2.9%
Global Systematics	3.85%	3.85%	3.85%	3.85%

Table 6: Summary of systematic uncertainties from each factor in  $B^+$  PbPb analysis for corrected yield vs centrality. All the values are shown in percentage.

Factors	0 - 30 %	30 % - 90%	0 - 90 %
Hadron tracking efficiency	5%	5%	5%
Data-MC Discrepancy	13.28%	8.49%	11.51%
$p_T$ shape	0.170%	0.106%	0.154%
PDF variation background	0.412%	1.13%	0.427%
PDF variation signal	2.50%	2.57%	2.60%
TnP Systematics	7.20%	7.85%	7.43%
MC stat.	3.37%	2.26%	2.49%
$T_{AA}$	2.0%	3.6%	2.2%
$N_{MB}$ events	1.26%	1.26%	1.26%
Total	16.63%	13.65%	15.24%
Branching fractions	2.92%	2.92%	2.92%
Global Systematics +)	2.92%	2.92%	2.92%
Global Systematics -)	2.92%	2.92%	2.92%

We also plot the summary plots for the table as follows in Fig 21



Plots/Sys/Summary/BpPtSyst.png

Plots/Sys/Summary/BpCentSyst.png

Plots/Sys/Summary/BpIncSyst.png

Figure 21: The plots summarizing the systematic uncertainties for  $p_T$  bins and centrality bins are shown above.



## 10.2 $T_{AA}$ and $N_{MB}$

We use the same uncertainties for  $T_{AA}$  (Nuclear overlap function) and  $N_{MB}$  (Number of Min-Bias events) listed on the B<sub>s</sub> AN [? ].

## 10.3 Branching ratio

The systematic uncertainty on the branching ratio of the decay  $B^+ \rightarrow J/\psi K^+$ , with  $J/\psi \rightarrow \mu^+ \mu^-$ , is calculated by adding in quadrature the uncertainties on each sub-channel. The resulting uncertainty for the full decay chain is 2.8% [? ]. This is global to all  $p_T$  and centrality selections in our analysis.

## 10.4 Tracking efficiency

The current standard value of tracking efficiency uncertainty for one track is 5% for now, suggested in the link <https://twiki.cern.ch/twiki/bin/viewauth/CMS/HITracking2018PbPb>. This is global to all  $p_T$  and centrality selections in our analysis. This number may be updated later.

## 10.5 Muon efficiency: Tag and Probe

DRAFT

295 The difference between the nominal and varied values are quoted as our systematics, and they  
 296 are shown in Figure 22 and Figure 23 and Table. 5 and Table. 6.



Figure 22: The upper bound and lower bound systematic uncertainties in  $\langle \frac{1}{\alpha \times \epsilon} \rangle$  vs  $p_T$  with total tag and probe correction are shown above.



Figure 23: The upper bound and lower bound systematic uncertainties in  $\langle \frac{1}{\alpha \times \epsilon} \rangle$  vs  $p_T$  with total tag and probe correction are shown above.

297 According to the studies, we calculate the systematic uncertainties due to tag and probe scale  
 298 factor. Table ?? and table ?? summarize the tag and probe systematic uncertainties results for  
 299  $p_T$  and centrality from our studies

300 However, according to the Muon POG, due to the issues in the trigger tag and probe scale  
 301 factor, we have also conduct the difference between the efficiencies with and without any tag  
 302 and probe scale factor correction applied. Our results for  $p_T$  and centrality are shown in Figure  
 303 ??:

304 The final summary of tag and probe systematic results are shown on Table ?? and ??.

DRAFT

## 305 10.6 MC-Data Discrepancy

DRAFT

**10.7  $p_T$  shape: Bpt weight**

The potential difference in  $p_T$  distributions in data and MC entails difference in calculation of efficiency correction. The  $B^+$   $p_T$  distributions of MC can be modified by Bpt weight (shown in 2.2.1) to have closer distribution with data. The Exponential and Polynomial weight function are given by:

DRAFT

311 We quote the percentage deviation of B  $p_T$  weighted efficiency correction factor  $\langle 1/(acc \times eff) \rangle$   
312 from the nominal without B  $p_T$  weight as the systematic uncertainties. Table. ?? and Table. ??  
313 show the systematics for differential  $p_T$  and inclusive  $p_T$ , respectively. Fig. ?? and Fig. ?? shows  
314 the comparison plots for differential  $p_T$  and inclusive  $p_T$ , respectively. Note that the systematics  
315 are very small, thus the differences may not be distinguished prominently on the plots.

316 We can see that the  $p_T$  shape systematic uncertainties on the efficiency correction have been  
317 reduced to negligible using the  $\langle \frac{1}{a \times e} \rangle$  approach.

DRAFT

## 10.8 MC stats: Toy MC study

The statistical uncertainties of  $\langle \frac{1}{\alpha \times \epsilon} \rangle$  in MC are examined by toy MC study. From the nominal 2D map, we generated 10000 toy MCs for each rapidity  $p_T$  bin. The toy 2D maps are then propagated to the  $\langle \frac{1}{\alpha \times \epsilon} \rangle$  data-average calculation. The distribution of data-averages are drawn in each analysis  $p_T$  and centrality bin, and the RMS deviation of the distribution (supposedly Gaussian) is compared to the nominal value. The ratio between RMS and the nominal value is quoted as systematics related to MC stats.

Table. ?? and Table. ?? shows the systematics for differential  $p_T$  and inclusive  $p_T$ , respectively. Fig. ?? and Fig. ?? shows the toy MC  $1/(\text{acceptance} \times \text{efficiency})$  distributions for differential  $p_T$  and inclusive  $p_T$ , respectively. The blue markers are toy MC distribution, the red lines are the nominal correction factors in the main analysis.

DRAFT

## 10.9 Signal extraction: PDF variation

Here we quote the numbers. For plots with more details, see Appendix. ??.

As discussed in detail in Sec. 6, the central value of the raw yields were extracted using a fit function of Double Gaussian signal and exponential combinatorial background. The non-prompt component was modeled with an error function and a Gaussian peak. The systematic uncertainty on the signal extraction was evaluated by varying the functions used to model the various components:

- Model the signal with a triple Gaussian function.
- Release the constraint on the width of the signal double Gaussian (fixed in the default fit to the MC extracted values). In this case, a scaling factor ( $a$ ) between the MC widths is left as a free parameter in order to account for possible differences between the resolution in data and MC.
- Fixed the mean of the signal double Gaussian.
- Consider 1st, 2nd, and 3rd order polynomial for combinatorial background.

The detailed description on the release of the constraints on the widths of the signal double Gaussian is as follows. The full signal model used in the fit to data is:

$$\alpha \frac{1}{a\sigma_1\sqrt{2\pi}} e^{-\frac{1}{2} \frac{(B_{mass}-\mu)^2}{(a\sigma_1)^2}} + (1-\alpha) \frac{1}{a\sigma_2\sqrt{2\pi}} e^{-\frac{1}{2} \frac{(B_{mass}-\mu)^2}{(a\sigma_2)^2}}, \quad (1)$$

where  $a$  is the resolution scaling factor (the same for both gaussians) which describes the possible discrepancy between data and MC fit,  $\alpha$  is the relative proportion between the gaussians,  $\sigma_1$  and  $\sigma_2$  are the Gaussians' widths that are directly derived from MC fit, and  $\mu$  is the mean shared by both Gaussians.

In order to examine the potential systematic difference in data and MC signal fit, we are required to define a moderate variation range of the scaling factor in order not to introduce statistical fluctuations in our estimation. To achieve this, we first performed fit by letting the scaling factor float around in individual  $p_T$  bins and inclusive  $p_T$  bin (Figure ??). The parameter values from the best fit are summarised in Table ??.

We observed that in several individual  $p_T$  bins, the deviations from unity (value for nominal fit) are sizable. In mid  $p_T$  bins (10-15-20-30 GeV) where statistics are comparably large, the deviations are small. Whereas in low  $p_T$  bins (5-7-10 GeV) and high  $p_T$  bins (30-40-50-60 GeV) where statistics are small, the deviations are relatively large. On the other hand, the optimal scaling factors of individual  $p_T$  bins agree with that of inclusive  $p_T$  bin within a significance of  $2\sigma$ . From this observation, the sizable differences can be considered to mainly come from statistical limitation of each small  $p_T$  ranges. In addition, the  $2\sigma$  difference can be considered to be the statistical uncertainties of scaling factors of individual  $p_T$  bins.

To conclude, we can claim that the scaling factor of the inclusive  $p_T$  ( $a = 1.09 \pm 0.04$ ) is a representative of the scaling factor for all  $p_T$  bins. Since that factor is a optimal fit parameter, we can take 10% variation from the nominal value (unity) as our signal PDF variation range where the fits are good enough and has reasonably low statistical uncertainties, which is desirable for systematic uncertainties estimation. Here we used jargons increased and decreased width which refers to  $a = 1.1$  and  $0.9$ , respectively.



## A Non-prompt $J/\psi$ background

In the  $B^+$  invariant mass spectrum, there are potential background feed-down sources coming from other B meson decays that can form peaking structures in the region of interest, and need to be properly subtracted in order not to bias the yield extraction procedure. In order to estimate these components, we processed the inclusive B meson MC sample with the nominal  $B^+$  channel workflow, and vetoed the candidates that are matched to a genuine  $B^+$  signal. The resulting B candidate mass spectrum in the inclusive  $p_T$  range (5-100 GeV/c) is shown in Fig. 24 for PbPb MC samples.

It is clear that these sources create a peaking structure in the region of  $M_{\text{inv}} < 5.20 \text{ GeV}/c^2$ . This structure can be nicely fit with an error function as done previously in B proton-proton analyses [? ]. In addition, there is a minor peak on the right shoulder ( $\approx 5.34 \text{ GeV}/c^2$ ) of the nominal signal ( $\approx 5.28 \text{ GeV}/c^2$ ), and this can be fit with an Gaussian function. There is additional combinatorial background which is fitted with a linear function. This contribution is absorbed in the total combinatorial background of our nominal channel of the main analysis. As described in details in Sec. 6, the shape of the Non-prompt function is used as template in the fit extraction procedure.

Further MC studies were done in order to identify the different channels that give rise to the non-prompt peaking structure in the  $B^+$  invariant mass spectrum. Few main processes were identified:

- 4-body  $B^+$  decays which occur via resonant decay channels e.g.  $B^+ \rightarrow J/\psi K^*(892)^+$ . In these cases, we distinguish the kaons coming from the  $K^*(892)^+$  decays as coming from a signal  $B^+ \rightarrow J/\psi K^+$  decay.
- 4-body  $B^0$  decays channels e.g.  $B^0 \rightarrow J/\psi K^*(892)^0$ .
- $B^+ \rightarrow J/\psi \pi^+$  decays in which we misidentified the  $\pi^+$  as a  $K^+$ .

The different contributions in PbPb are presented in Fig. 25. The contribution from  $B^+ \rightarrow J/\psi \pi$  clearly form a peaking structure on the right shoulder of the nominal decay channel  $B^+ \rightarrow J/\psi K^+$ . However, the overall magnitude of this component is tiny compared to the other two sources, and negligible compared to the nominal signal. As a consequence, we can barely see the contribution of this peaking structure in the invariant mass plot of  $B^+$  nominal channel.



Figure 24:  $B^+$  candidate mass spectrum obtained in inclusive B meson MC production after vetoing the contribution of genuine  $B^+$  signal candidates in PbPb.

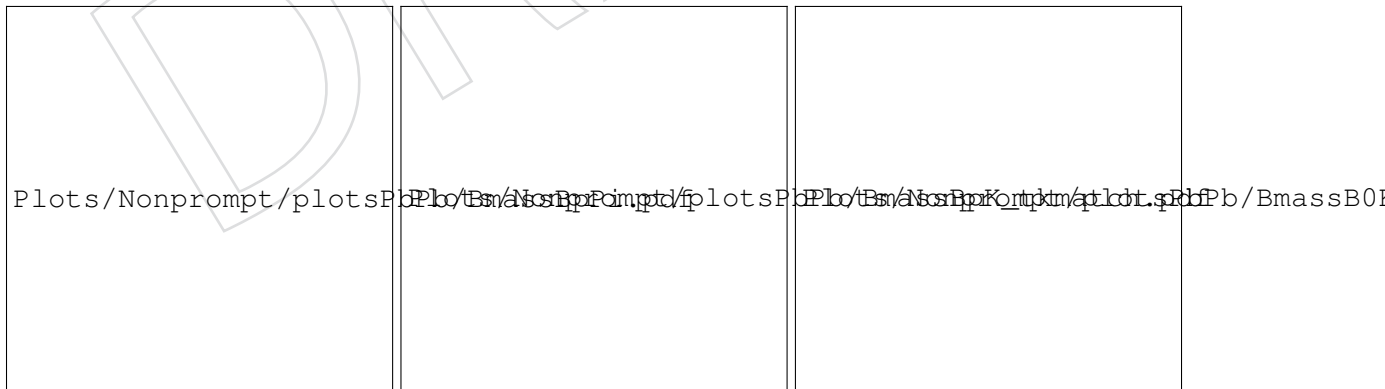


Figure 25: Peaking background contribution from  $B^+ \rightarrow J/\psi \pi$  and from K resonant decay channels of  $B^0$  and  $B^+$  in PbPb MC.

396 **Acknowledgments**

397 **References**

DRAFT