Heavy Quarks as Probes of the Primordial Plasma

Maria Faria, Artur Semião

Supervised by: Nuno Leonardo, Zhaozhong Shi

LIP Internship Program, Final Workshop, 2020.09.11

LABORATÓRIO DE INSTRUMENTAÇÃO
E FÍSICA EXPERIMENTAL DE PARTÍCULAS

TÉCNICO LISBOA

Ciências ULisboa

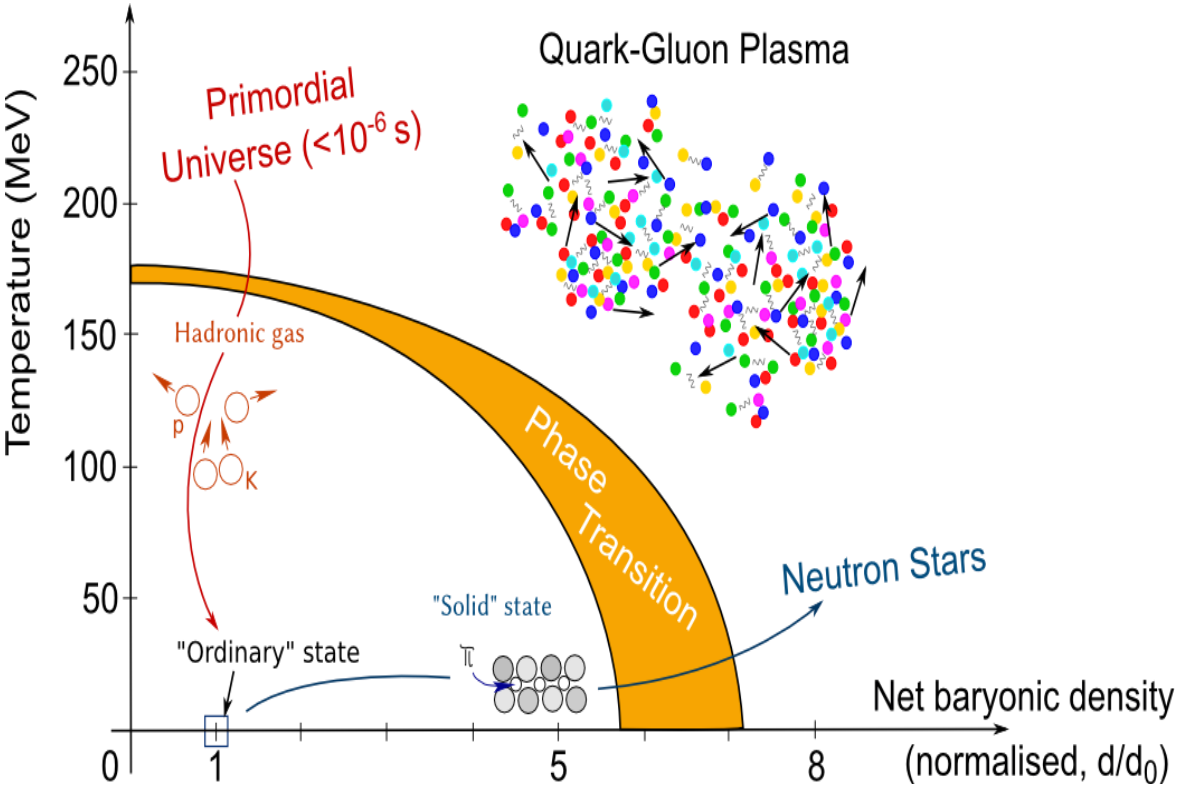
Introduction

Goal: To study the properties of the quark gluon plasma (QGP) using b-quark heavy mesons

- QGP is predicted to exist under extreme conditions of temperature and density (e.g. primordial Universe)

- Can be recreated in heavy ion collisions like Pb-Pb

- b-quarks are created at early stages and so “record” information about the QGP evolution



Introduction

- Study the hadronization of the b-quark by measuring the B_s^0(bs) cross section in pp collisions
- In this work the cross section of the B_s^0 is measured by reconstructing this meson from pp collision dataset through:

$$\boldsymbol{B_s^0}\ \textbf{decay channel:}\ B_s^0 \rightarrow J/\psi(\mu^+\mu^-)\phi(K^+K^-)$$

- Later on extending the study to the B^+ (bu) meson and comparing results with the Pb-Pb collision data will provide unique information about the properties of the QGP

Differential Cross Section

$$R_{AA} \propto \frac{\left(\frac{d\sigma}{dp_T}\right)_{PbPb}}{\left(\frac{d\sigma}{dp_T}\right)_{pp}}$$

$$\left(\frac{d\sigma}{dp_T}\right)_{pp} = \frac{1}{\epsilon B L} \frac{N_s}{\Delta p_T}$$

Cross Section Uncertainty:

Statistical
- Raw Yield

Systematic
- Raw Yield
- Efficiency
- Luminosity
- Branching Ratio

Efficiency (from MC)

Branching Ratio (PDG) $B = (31.3 \pm 2.3) \times 10^{-6}$

Integrated Luminosity $L = (302.3 \pm 2.3\%)\ fb^{-1}$

Raw Signal Yield (exctracted from fit)

Dataset and Pre-selection

Muon Tracks

$|P_T^{\mu}|$ > 3.5 GeV/c for $|\eta^{\mu}| < 1.2$

$> (5.47 - 1.89 \times |\eta^{\mu}|)$ GeV/c for $1.2 < |\eta^{\mu}| < 2.1$

> 1.5 GeV/c for $2.1 < |\eta^{\mu}| < 2.4$

Muons pixel + strip hits > 5, muon pixel > 0

Muon DCA xy < 0.3 cm - transverse impact parameter

Muon DCA z < 20 cm - longitudinal impact parameter

P(tracks muon from same decay vertex) > 1%

$|m(\mu^+\mu^-) - m(J/\psi)| < 0.15$ Gev/c²

Dataset: Colected by CMS

- pp collision data (from 2017)
- $\sqrt{s} = 5.02$ TeV
- Integrated luminosity L = 302.3 pb^{-1}.

MC: simulated by PYNTHIA + GEANT

High Purity Kaon Tracks

$|\eta^{K}|$ < 2.4

P_T^K < 1 GeV/c

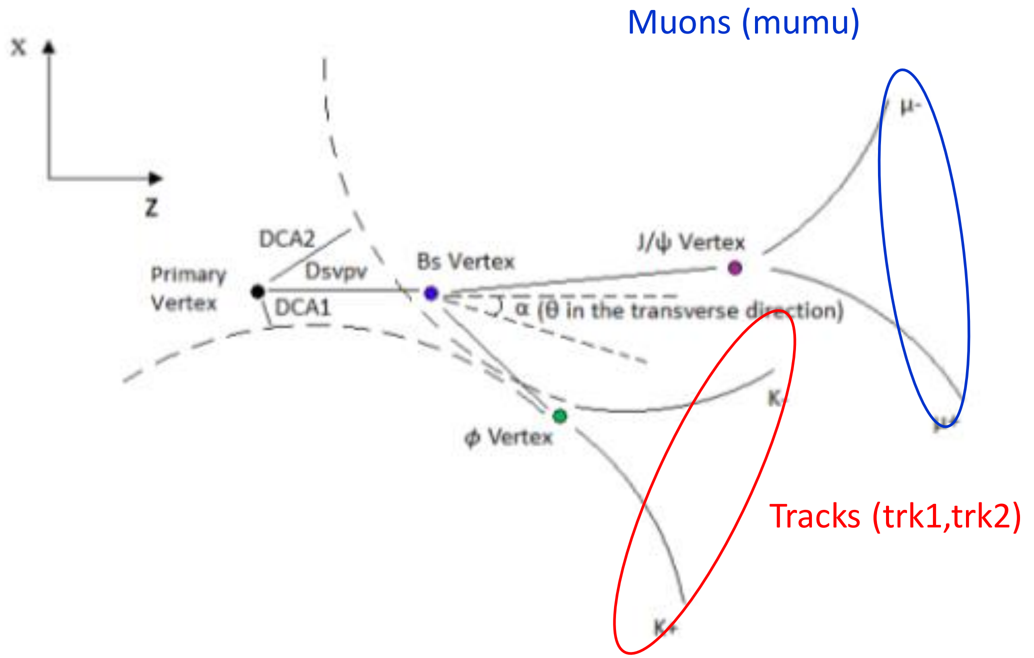
$\Delta P_T^K / P_T^K$ < 0.1

Kaon pixel + strip hits > 10

K χ2/d.o.f/total hits < 0.18

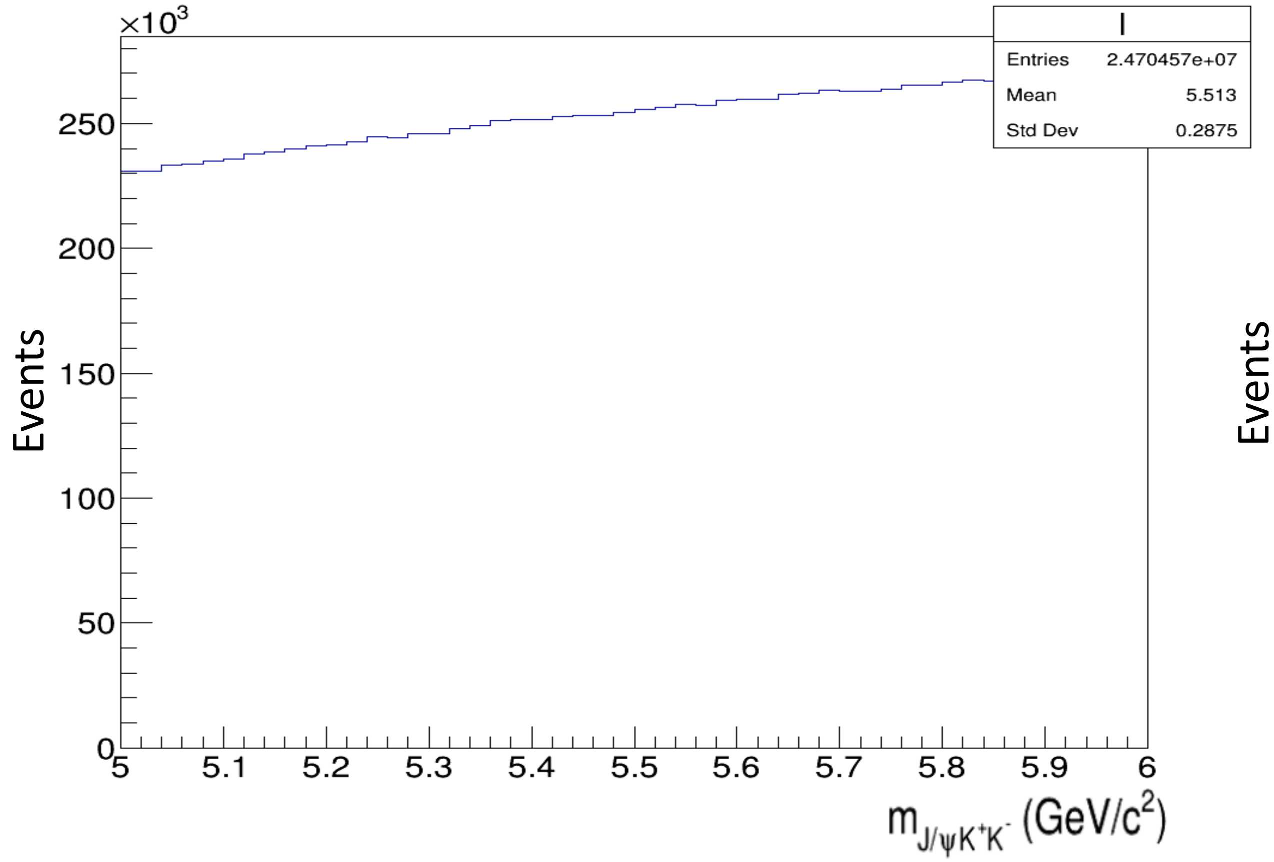
P(two kaon tracks from same decay vertex) > 0.05%

$|m(K^+K^-) - m(\phi)| < 0.015$ Gev/c²

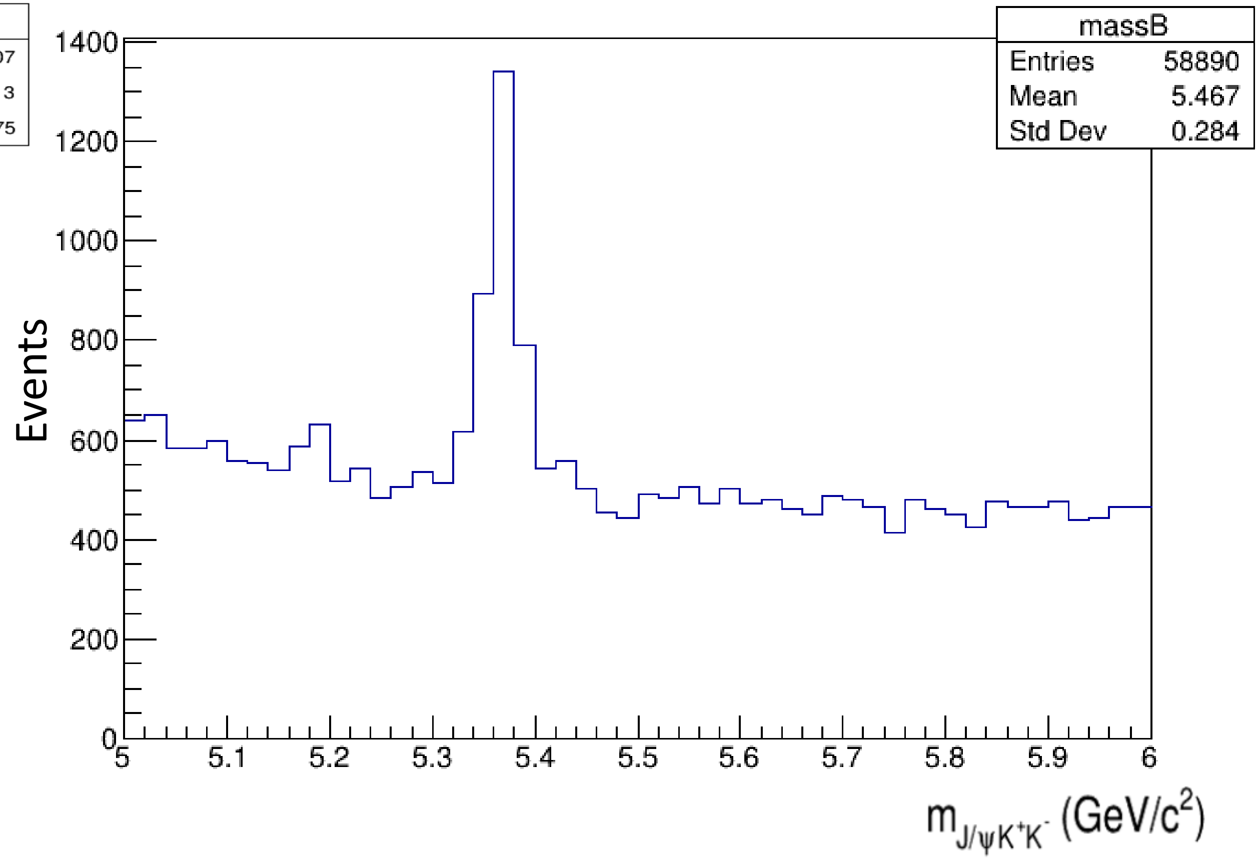


Bs candidates invariant mass $m(\mu^+\mu^-K^+K^-)$: Pre Selection

Data with no selection



Bs candidates after pre selection

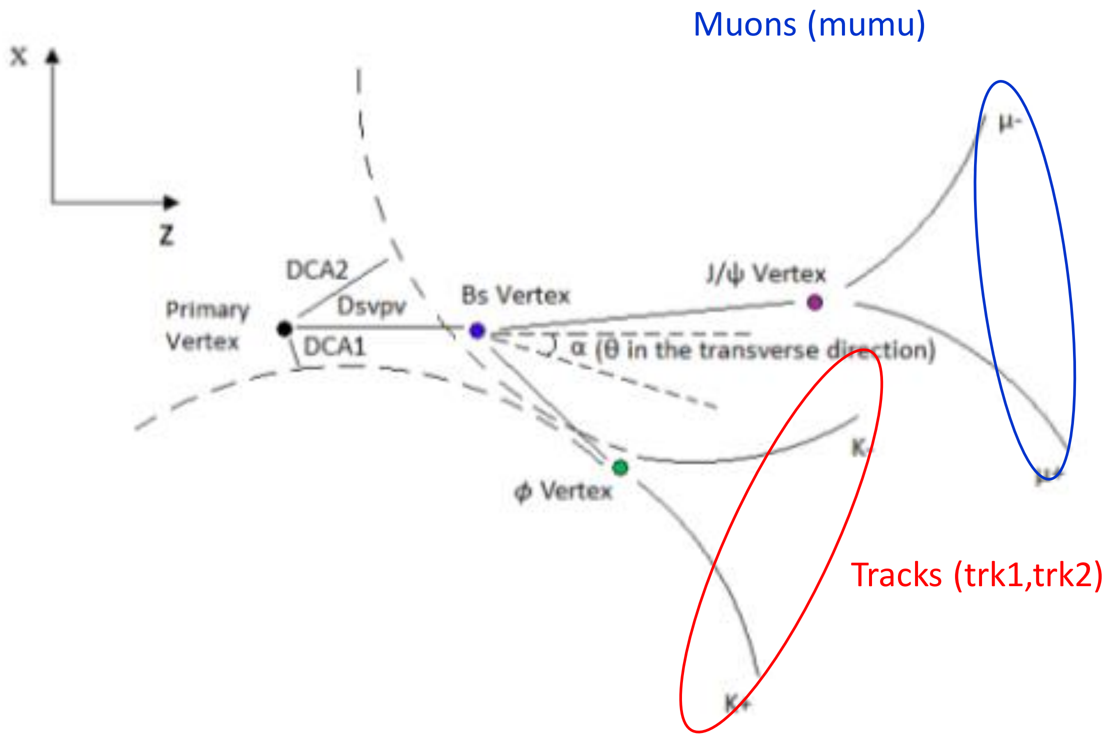


B_s^0 signal becomes visible after baseline selection

Optimised signal vs background selection using multivariate analysis (MVA)

Input variables used in the analysis for Bs

Btrk1Pt/Btrk2Pt Kaons Pt
Trk1DCAxy/Trk2DCAxy: Kaon tracks DCA PV xy: transverse impact parameter, devided by its error
Trk1DCAz/Trk2DCAz: Kaon tracks DCA PV z: longitudinal impact parameter, devided by its error
Bchi2cl: P(two kaon tracks from same decay vertex)
MassDis: $|m(K^+K^-) - m(\phi)|$
dls: Distance PV to SV, devided by its error
Balpha: angle between B_s^0 meson displacement and momentum
cos(Bdtheta): cosine of angle between B_s^0 displacement and mommentum in the transverse direction



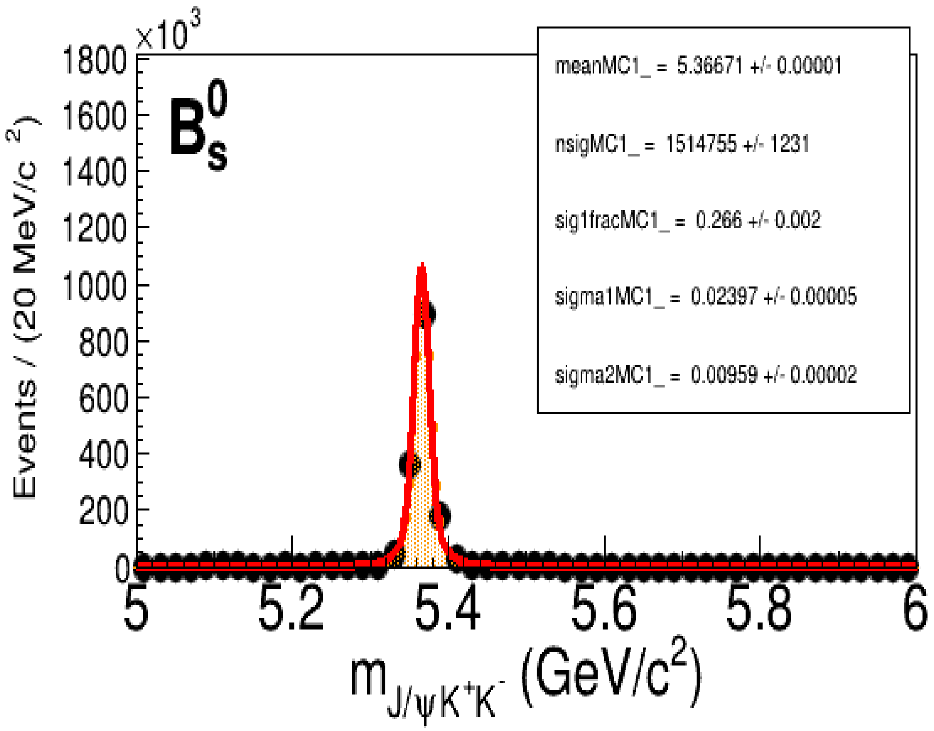
Training

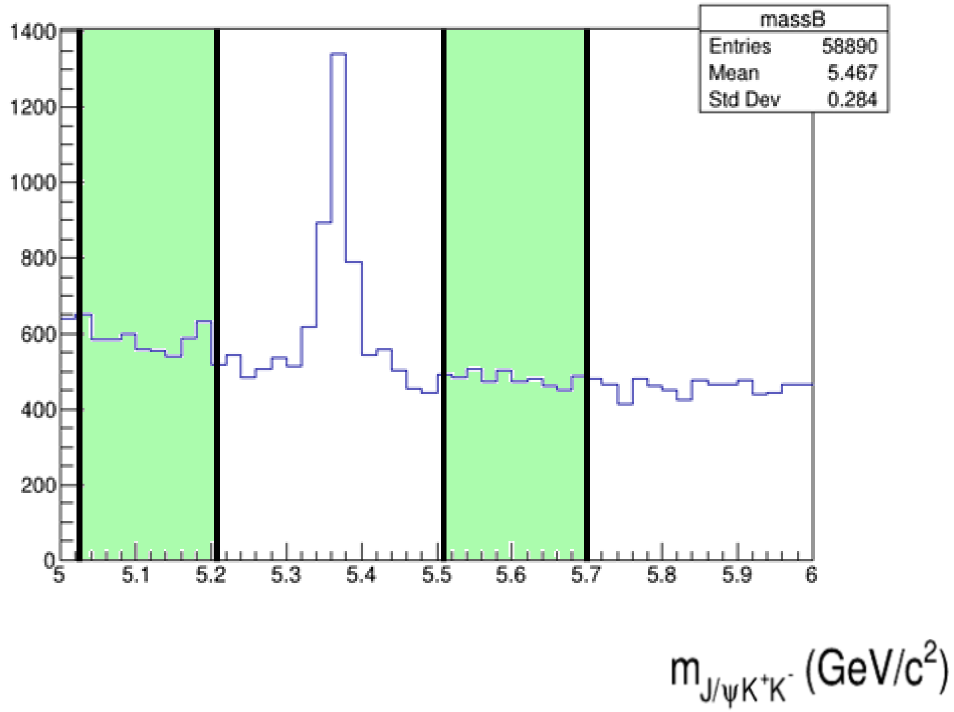
Classifiers: CutsGA and BDT

Signal events sample: MC candidates

Background events sample: Data sideband region candidates

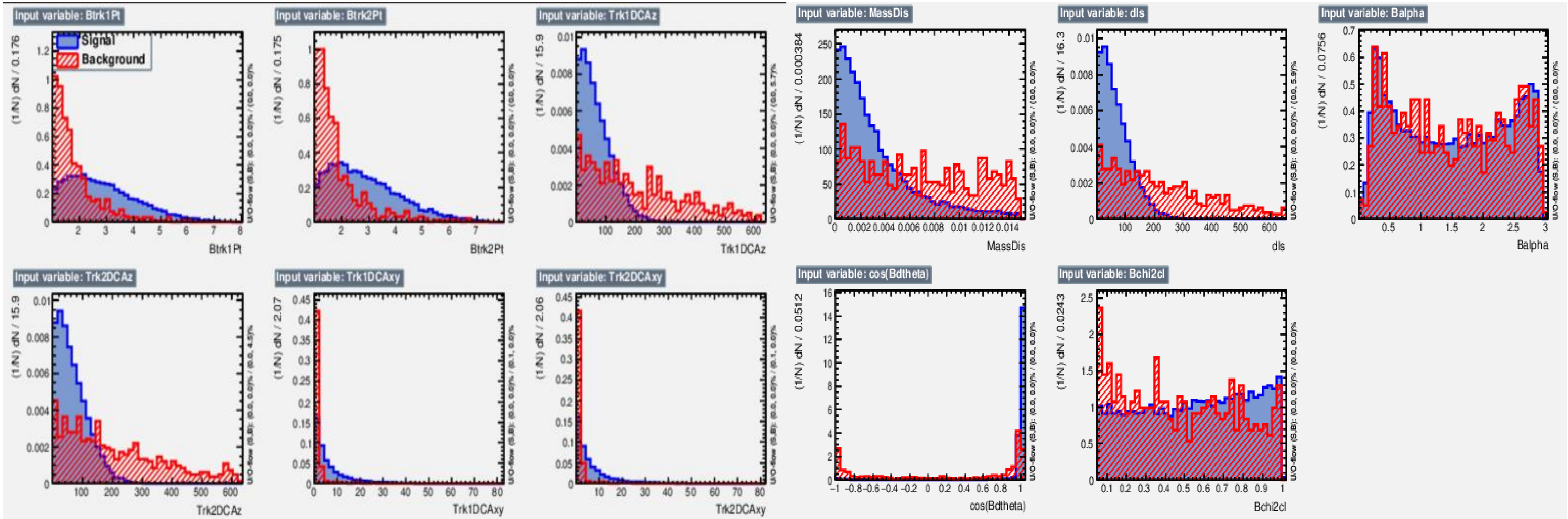
Training performed independently for 5 p_T bins: [4,7],[7,10]. {10,15], [15,20], [20,50] (GeV/c)





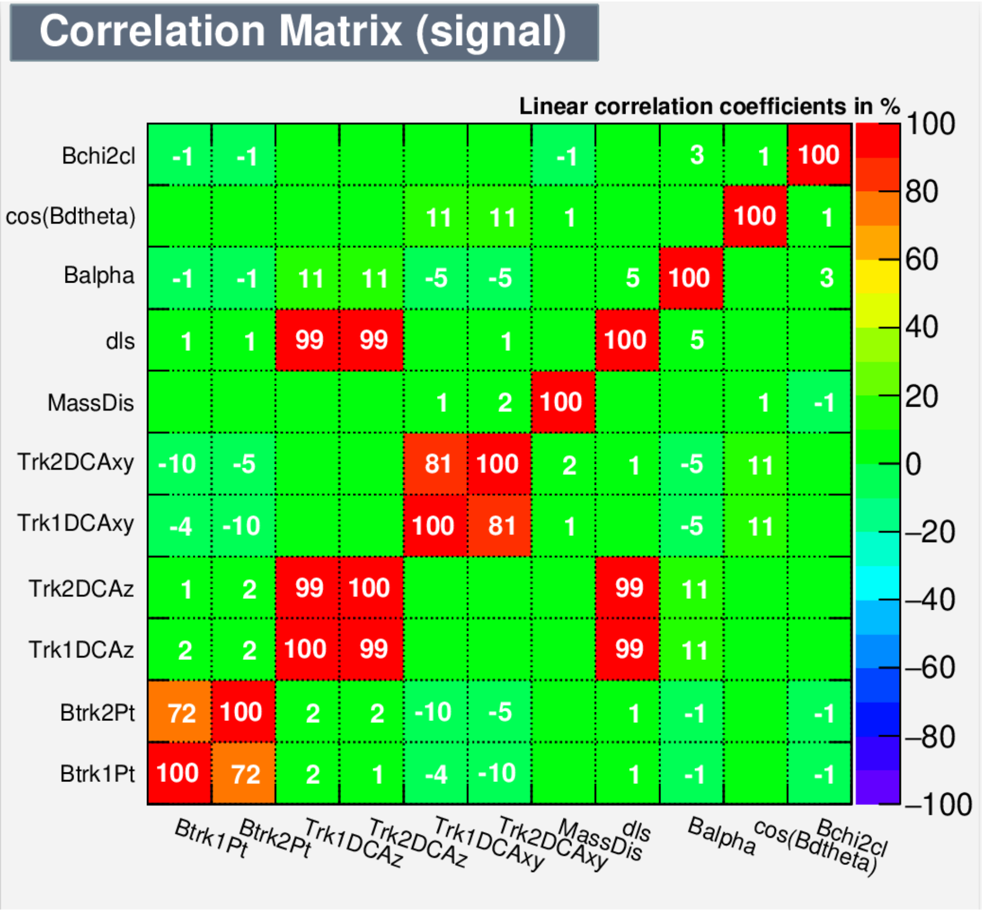
Input variables signal vs background

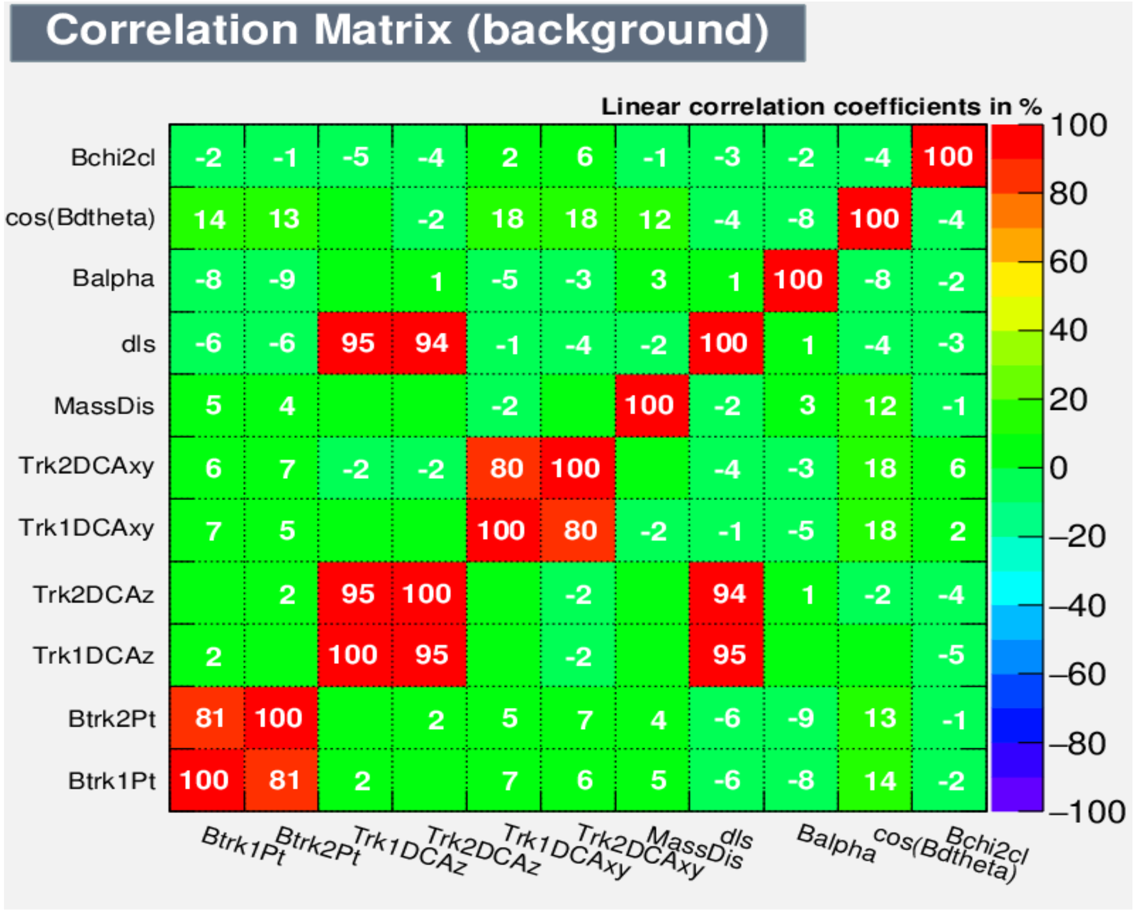
Distributions of input variables used in training samples for p_T bin [10-15] (Gev/c)



Correlation matrices

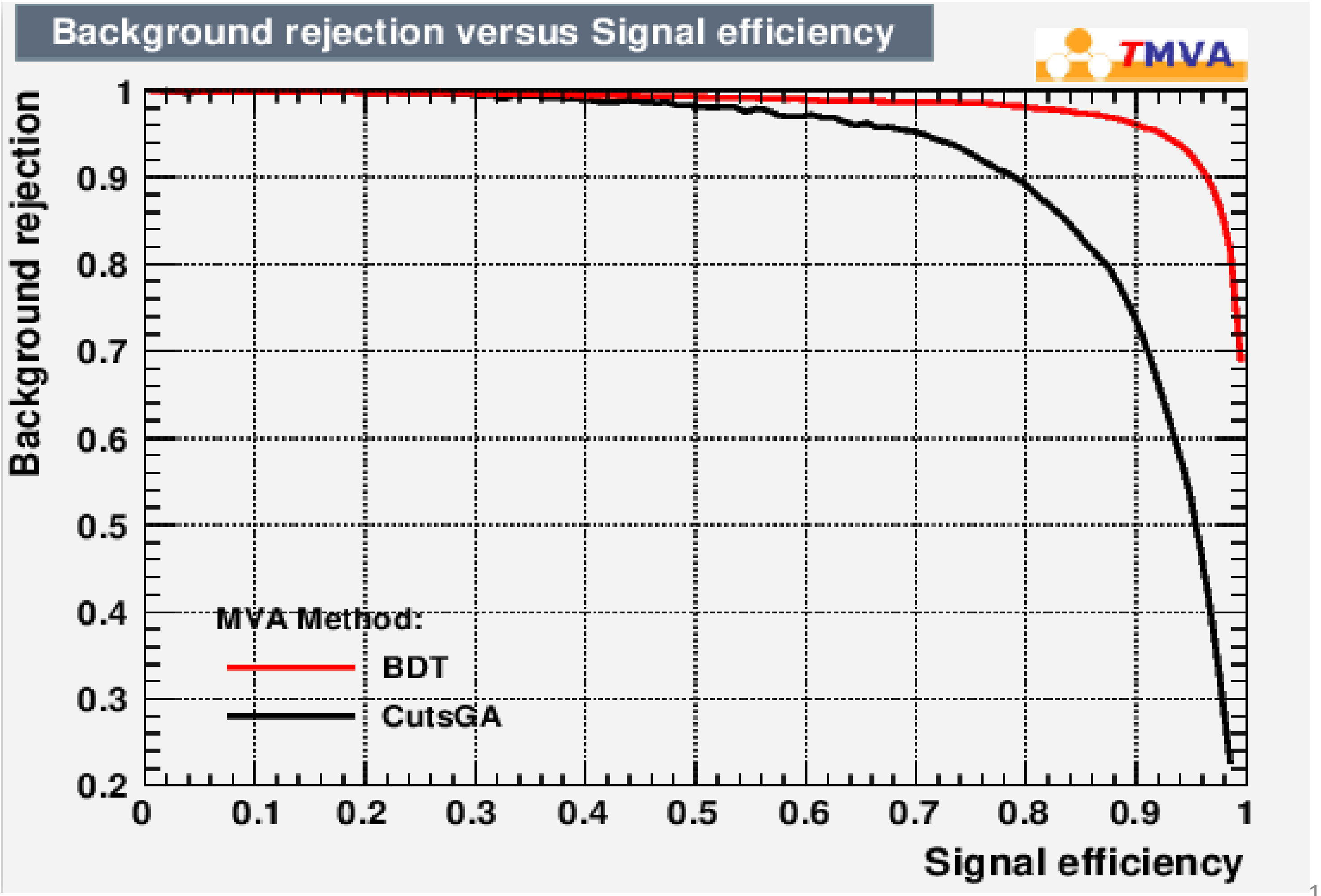
Correlation matrices between variables used in training for p_T bin [10-15] (Gev/c)





Classifier Performance

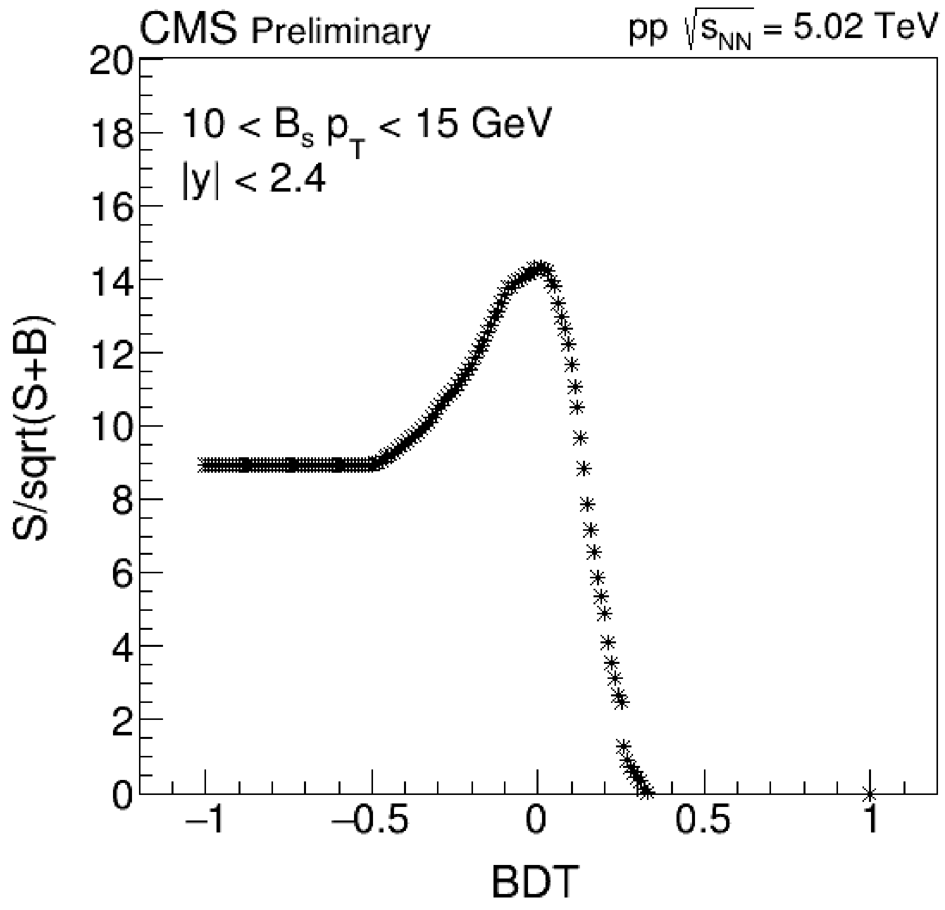
- Best classifier can be identified by the largest AUC (area under the curve) - BDT

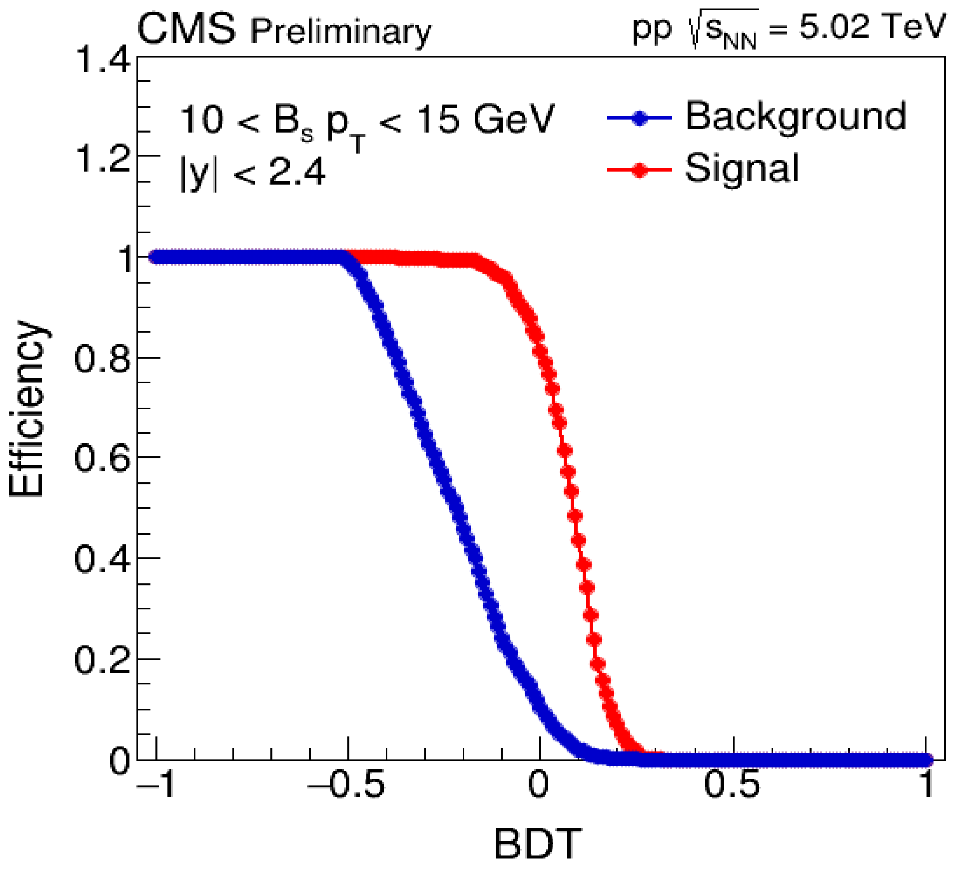


BDT optimal working point p_T [10,15] (Gev/c)

Maximize figure of merit FOM = $\frac{S}{\sqrt{S+B}}$

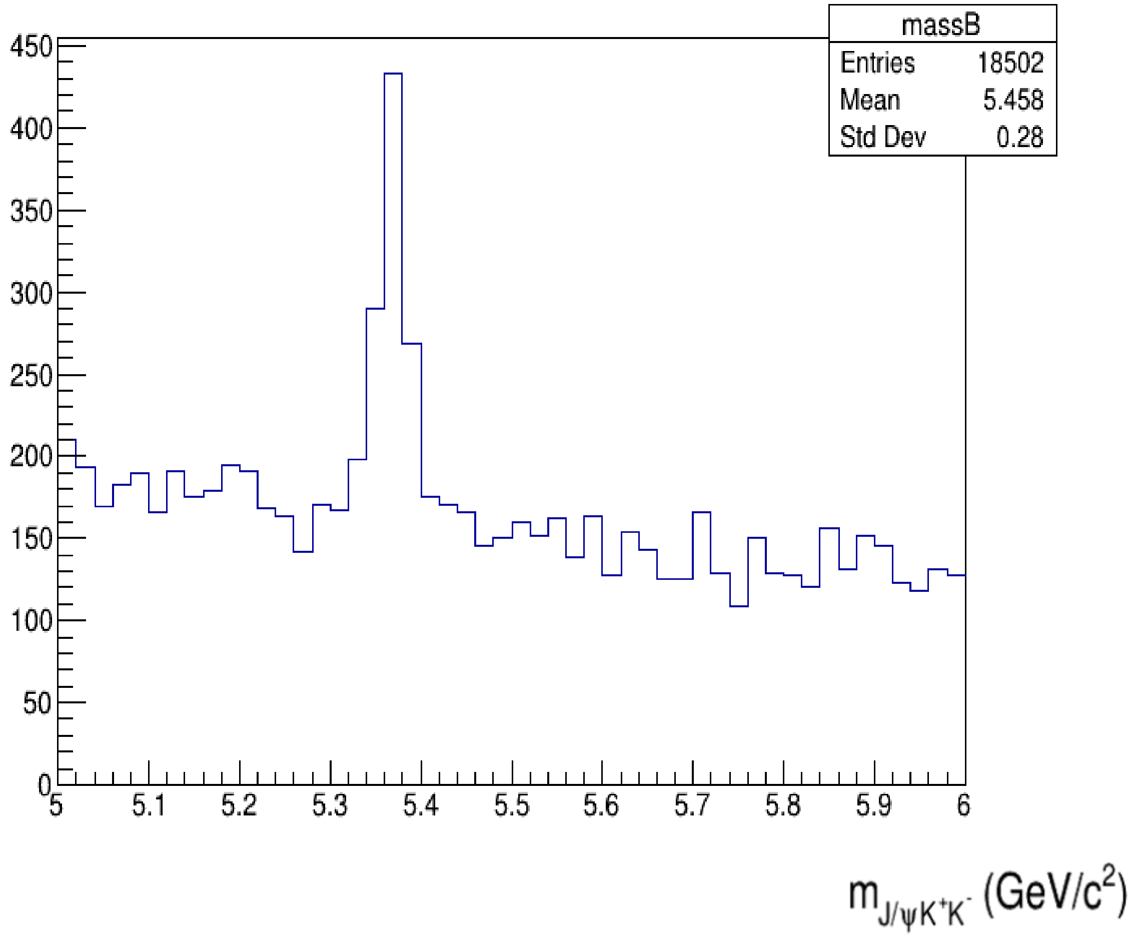
Max FOM = 14.31 for BDT score = 0.01 and effS = 0.79



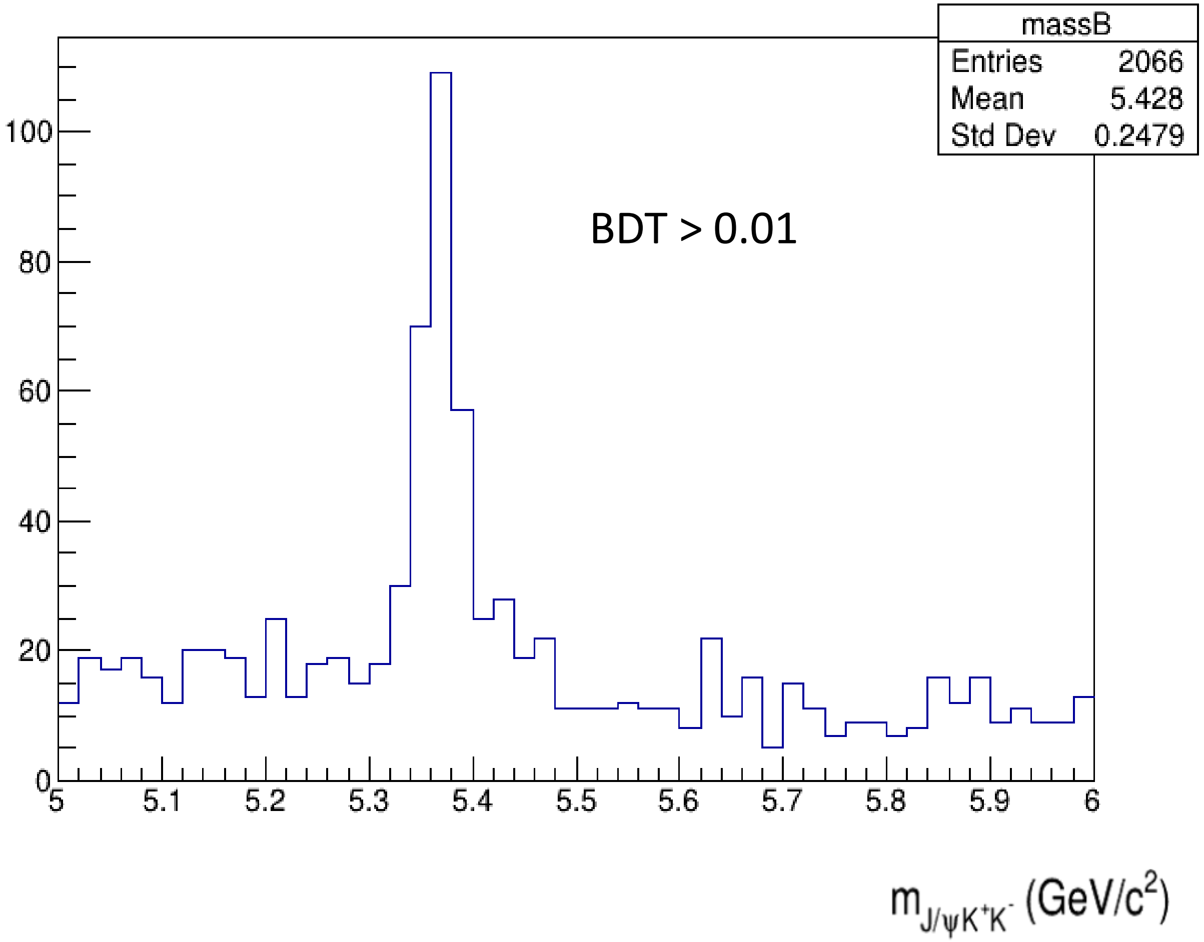


BDT optimal working point p_T [10,15] (Gev/c)

Bs candidates after pre selection p_T bin [10-15] (Gev/c)



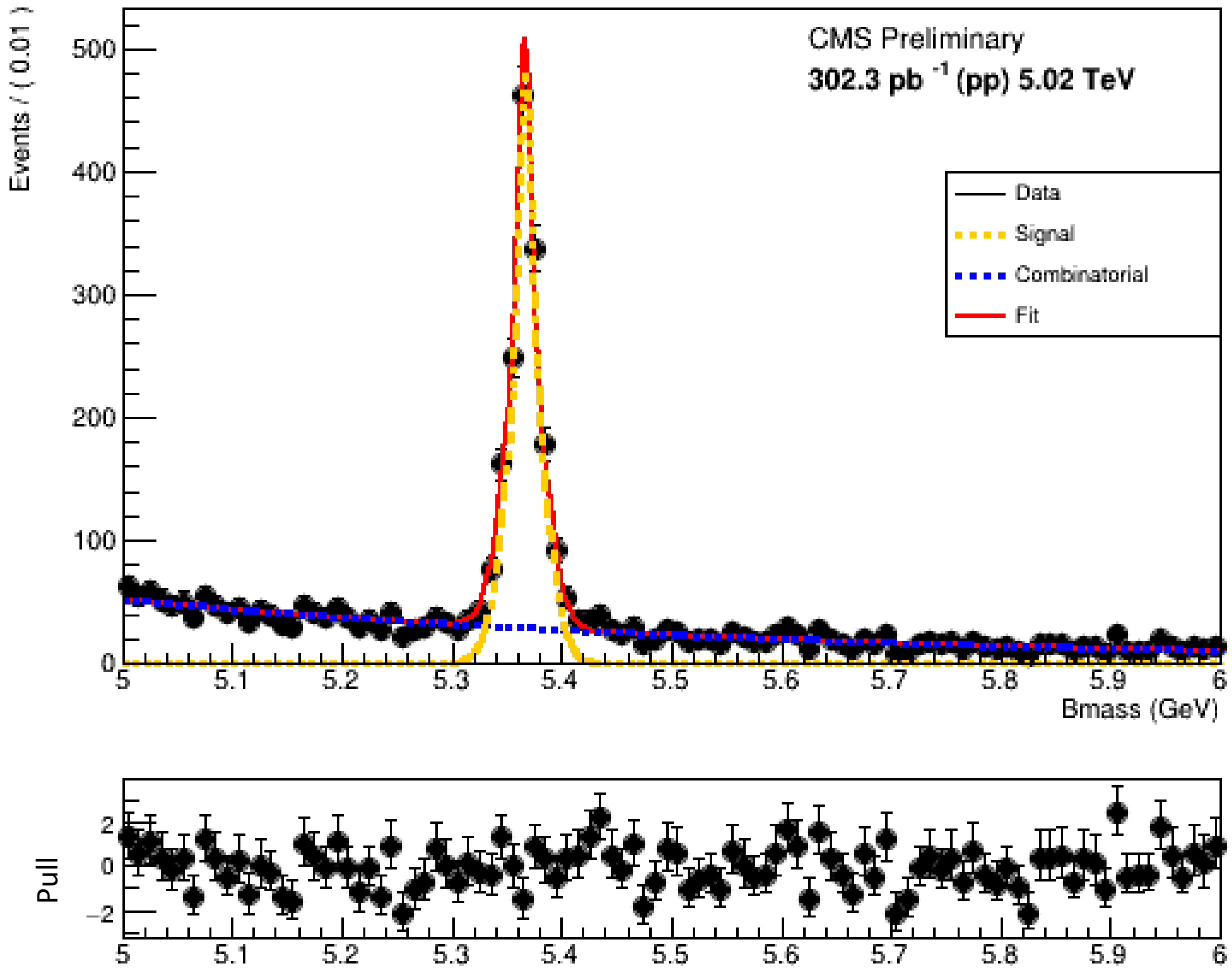
Bs candidates after optimal selection p_T bin [10-15] (Gev/c)



Likelihood Method & Signal Extraction

Likelihood Method

B_s^0 Invariant Mass Fit



Extended **Unbinned** Maximum Likelihood Fit

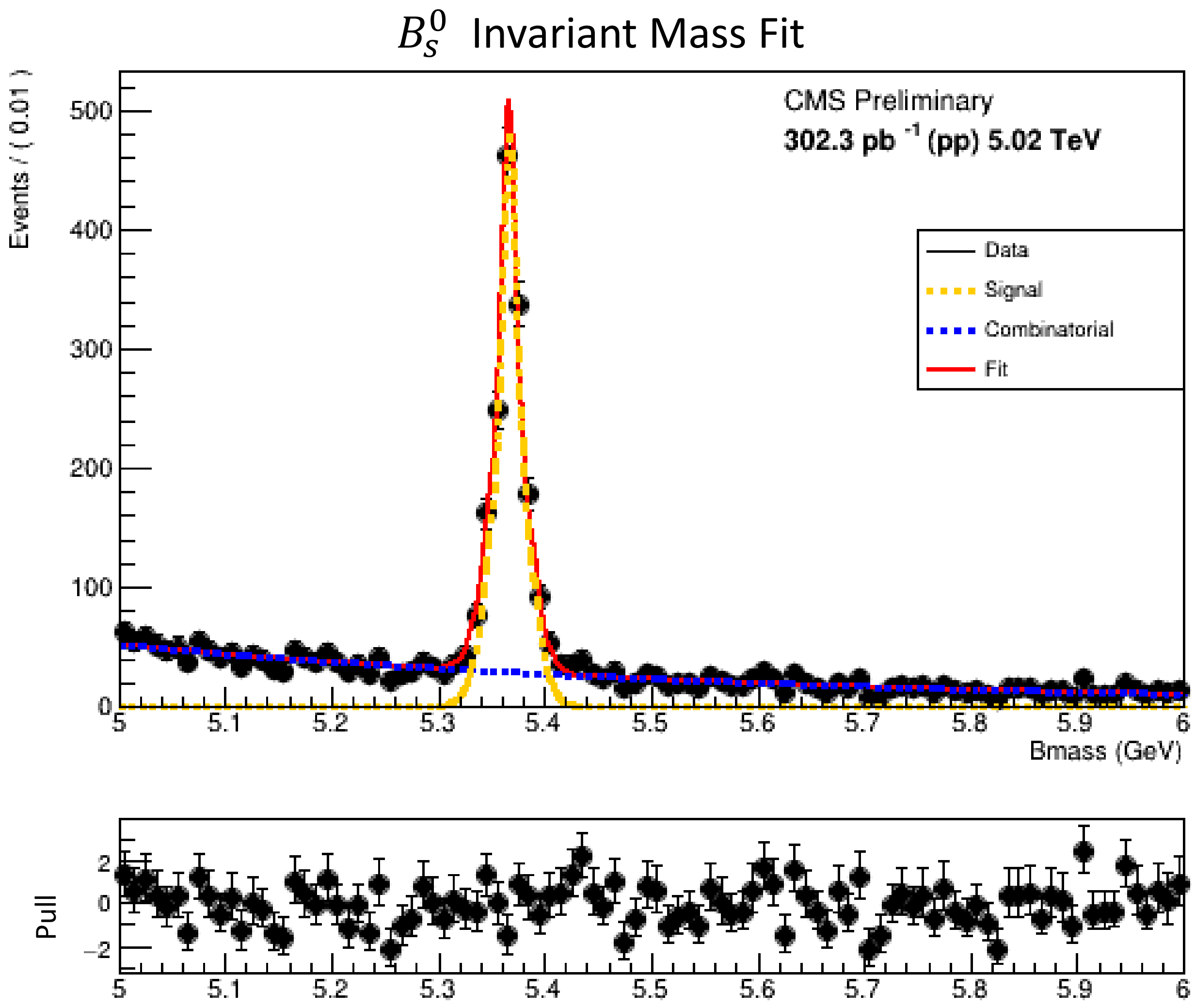
$$L(\{m_i\}, \boldsymbol{\lambda}) = \prod_{i=1}^{N_{obs}} l(m_i) \frac{e^{-N} N^{N_{obs}}}{N_{obs}!}$$

$$l(m_i) = N_S\, P_S(m_i; \mu, \sigma_1, \sigma_2) + N_{CB} P_{CB}(m_i; \lambda)$$

$$P_S = \alpha\, Gauss(\mu, \sigma_1) + (1 - \alpha) Gauss(\mu, \sigma_2)$$

$$P_{CB} = Exp(\lambda)$$

Likelihood Method



$$l(m_i) = N_S\, P_S(m_i; \mu, \sigma_1, \sigma_2) + N_{CB} P_{CB}(m_i; \lambda)$$

$$P_S = \alpha\, Gauss(\mu, \sigma_1) + (1 - \alpha) Gauss(\mu, \sigma_2)$$

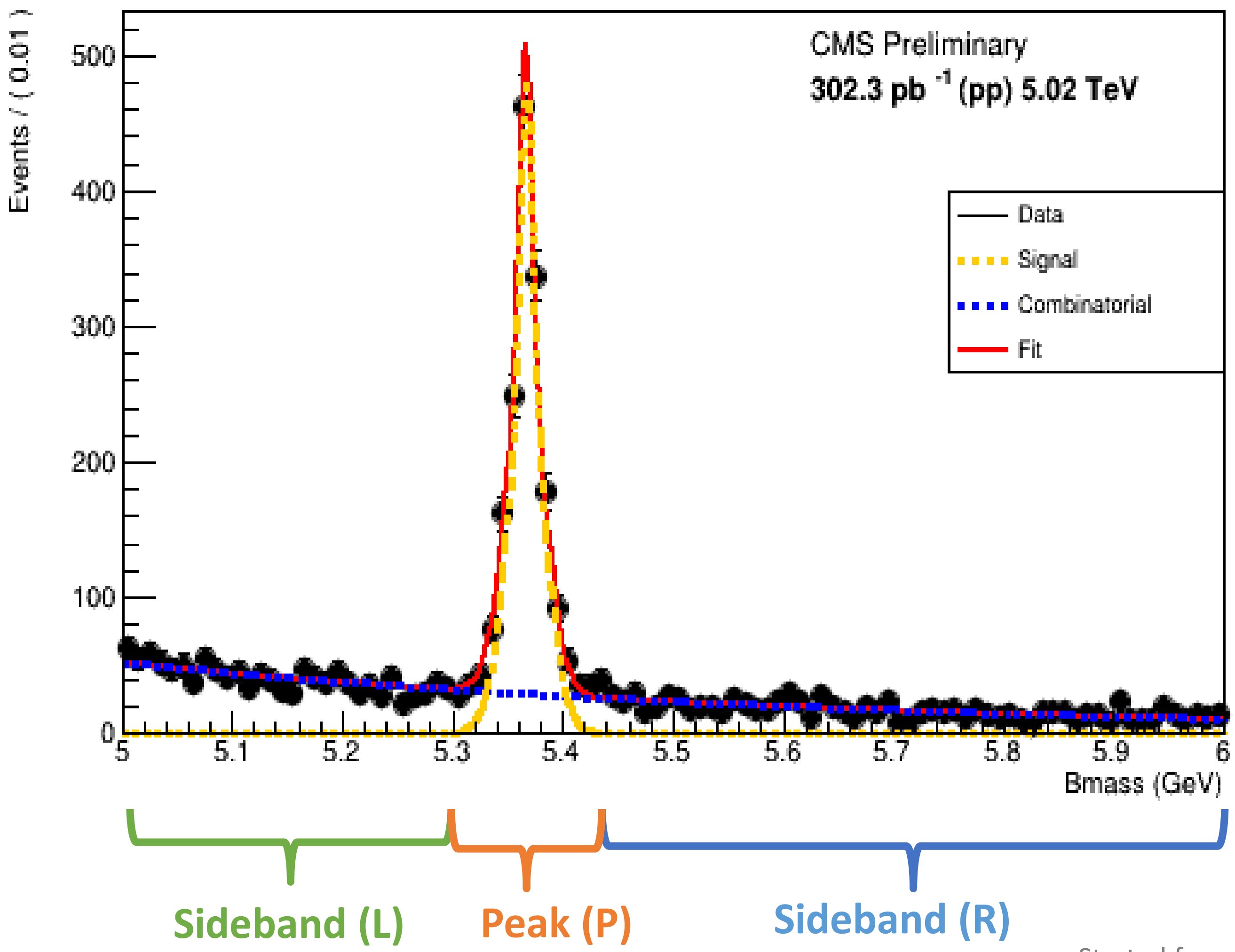
$$P_{CB} = Exp(\lambda)$$

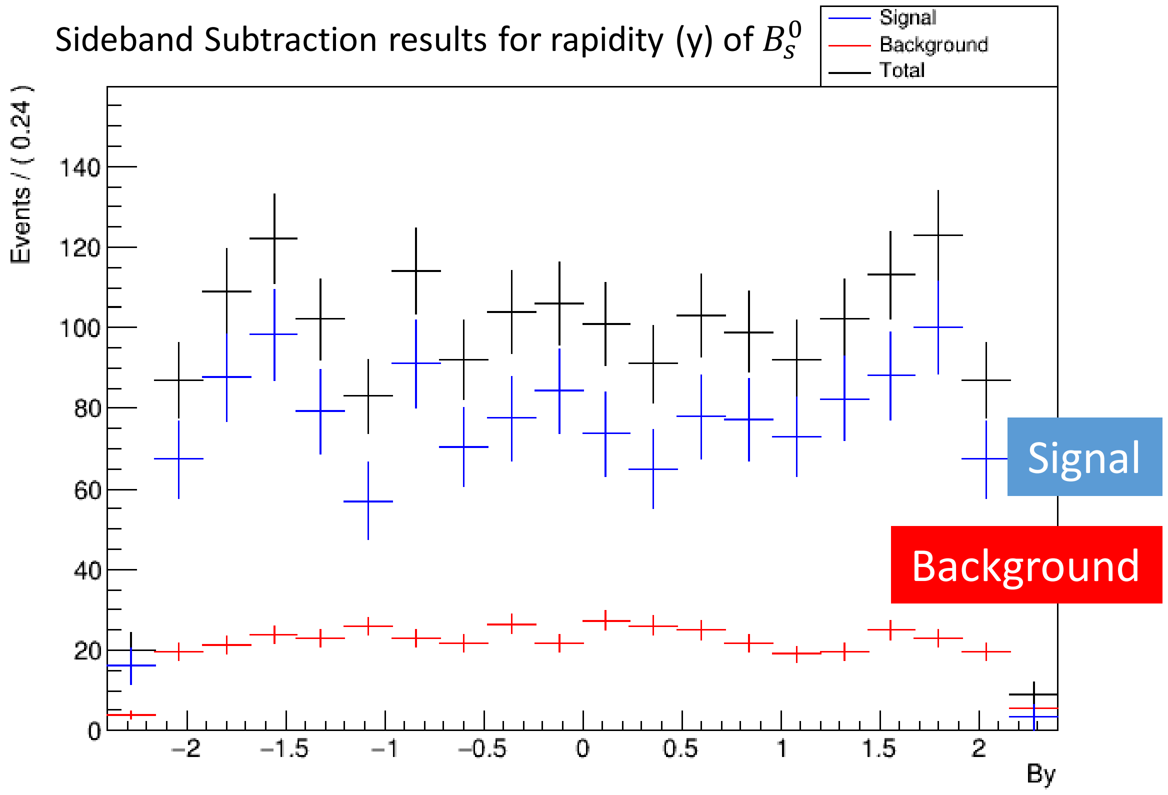
Coefficients	Value $\pm$ Statistical Uncertainty
α	0.776 ± 0.046
λ	-1.6065 ± 0.0726
μ	5.36677 ± 0.00041
σ_1	0.01775 ± 0.00091
σ_2	0.00539 ± 0.00073
N_S	1409 ± 42
N_{CB}	2582 ± 54

Sideband Subtraction

$$V_{signal} = V_{peak} - r \times V_{sideband}$$

$$r = \frac{P}{L + R}$$





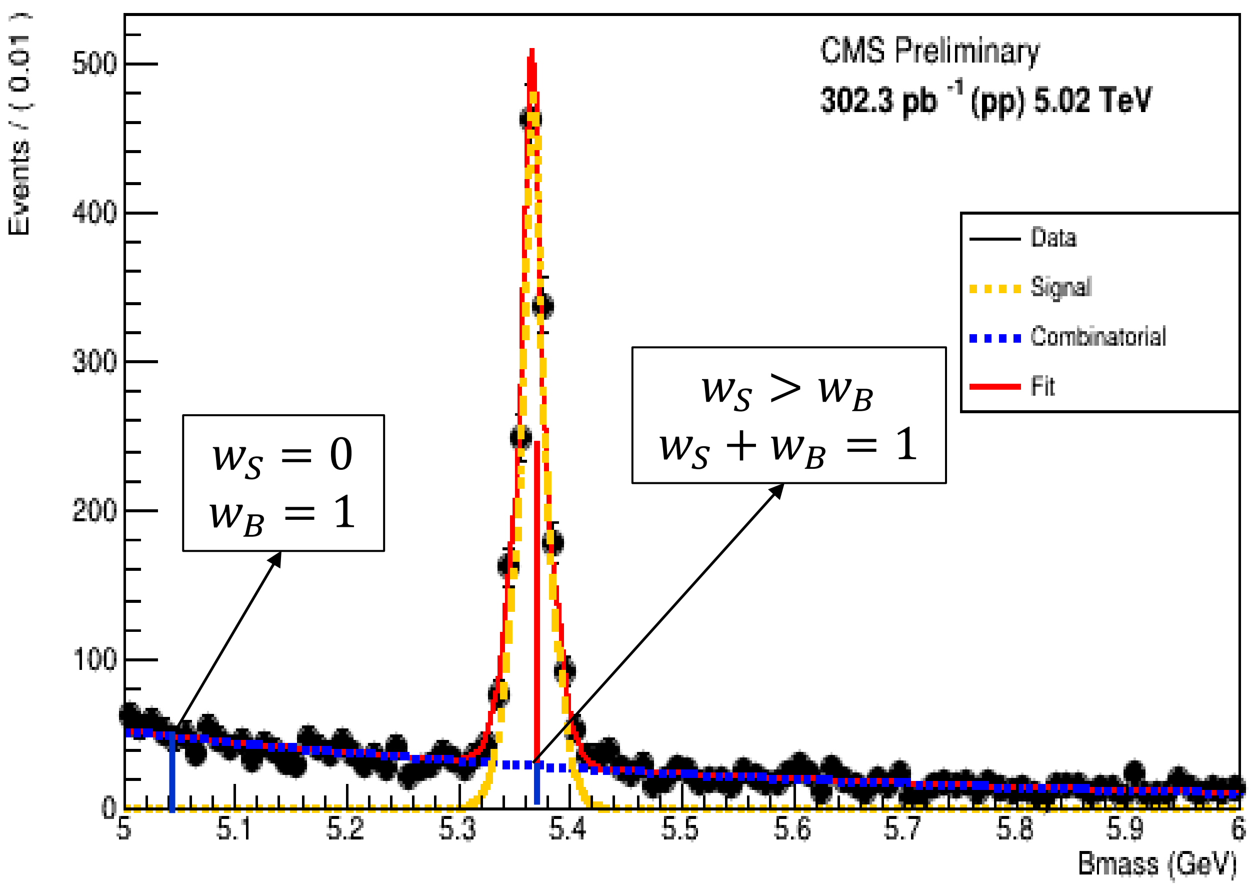
Started from baseline implementation for PbPb described in NA/2019-219 (A.Pardal, J.Gonçalves, et al.)

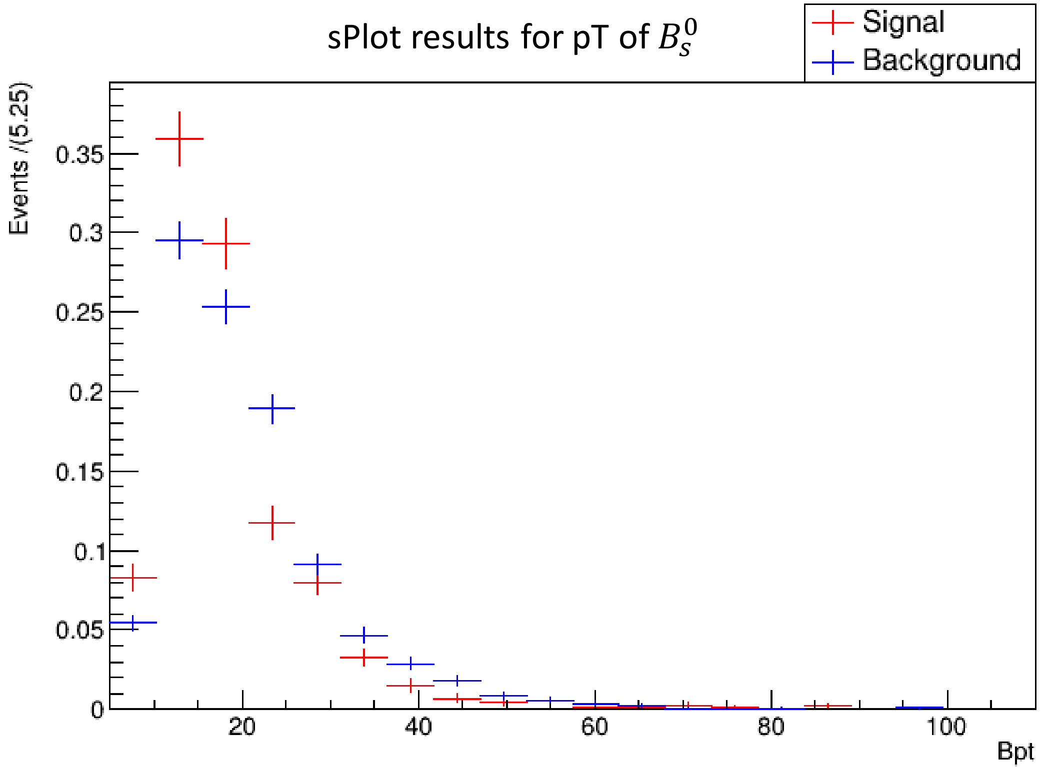
sPlot

Based on the likelihood fit to B_s^0 invariant mass

Attribute to each event two weights:

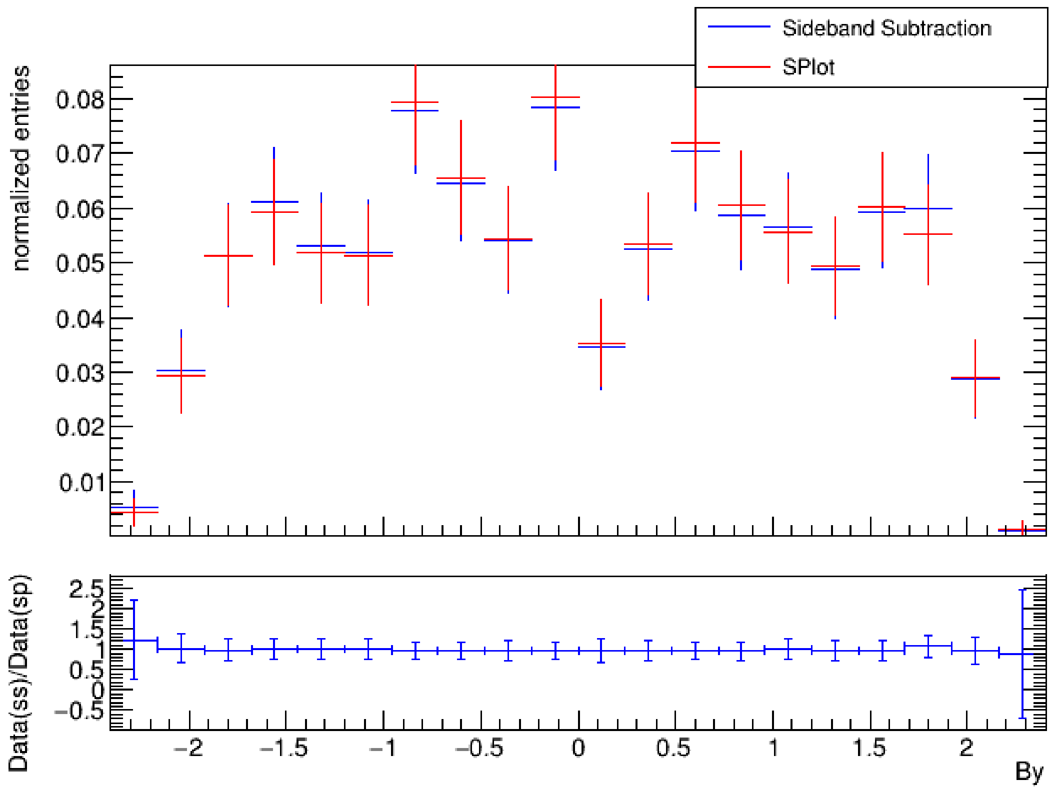
- w_S : probability of being signal
- w_B: probability of being background

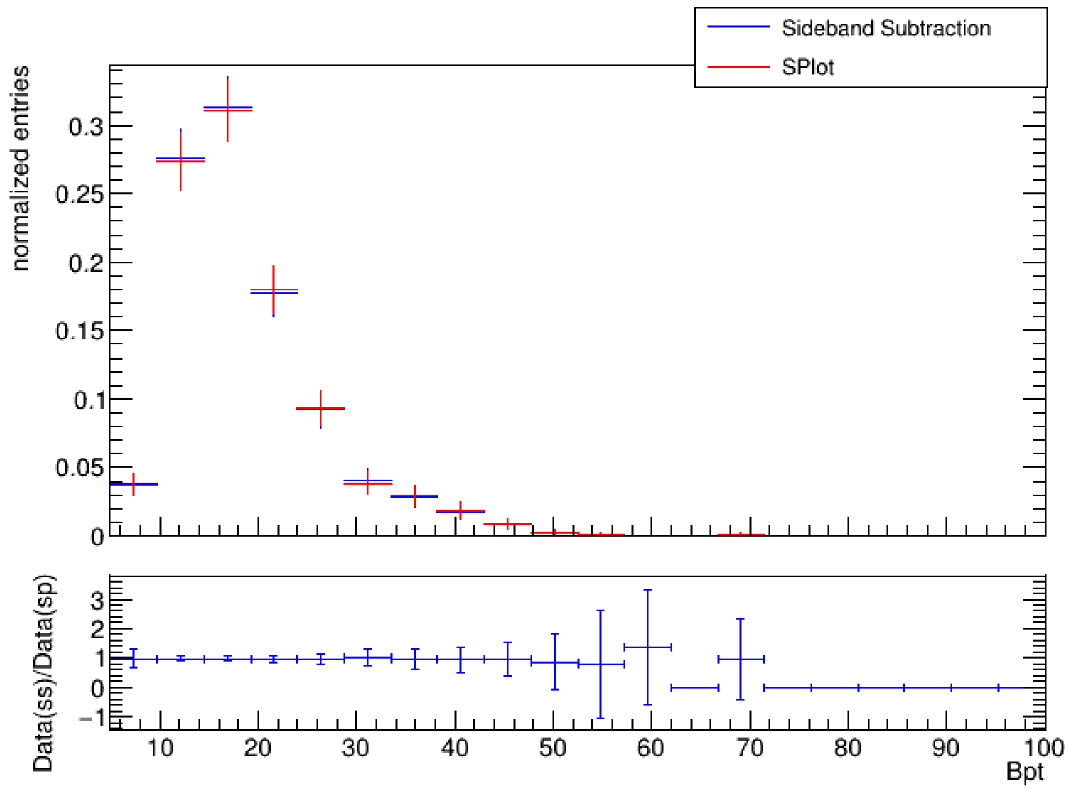




Started from baseline implementation for PbPb described in NA/2019-219 (A.Pardal, J.Gonçalves, et al.)

sPlot vs Sideband Subtraction

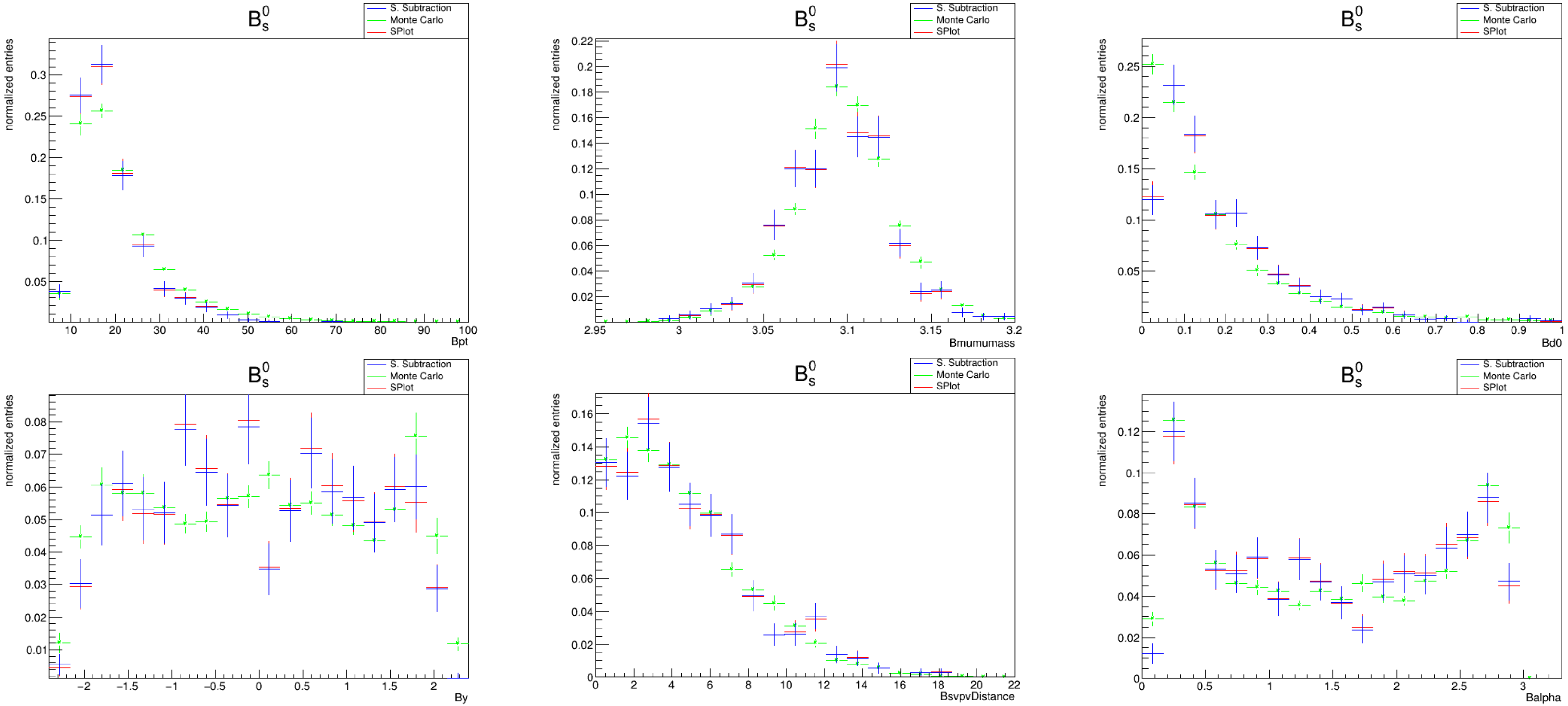




The two methods are found to agree well.

MC Validation

MC vs S.S. vs sPlot

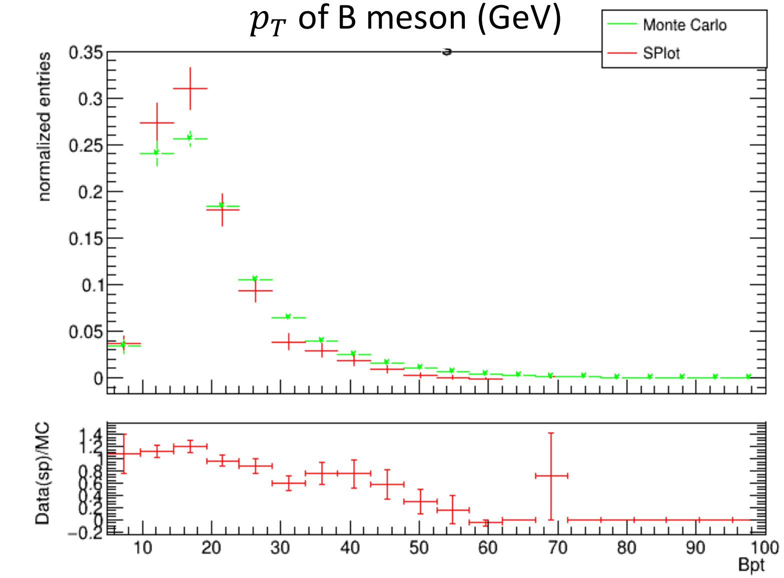


Some level of disagreement between data and MC can be observed for some variables.

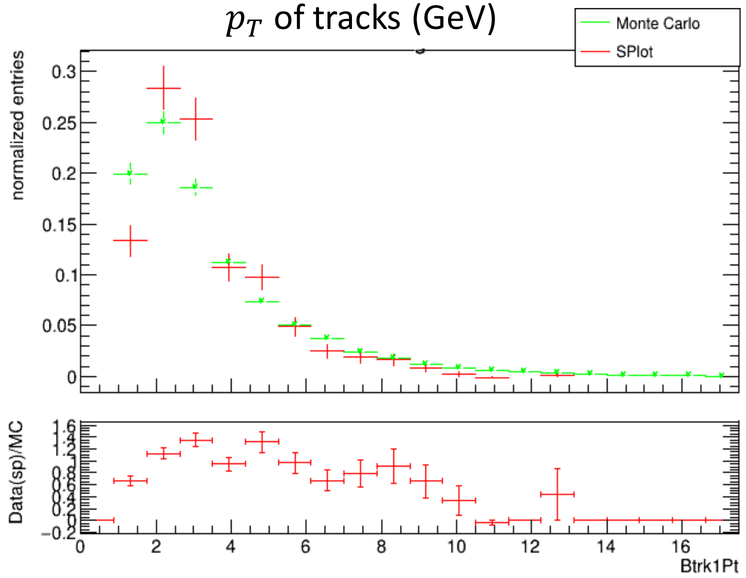
sPlot vs MC

Can use the Data/MC ratio of **Bpt** to <u>reweight</u> MC and see how it affects the data-MC agreement of the other variables

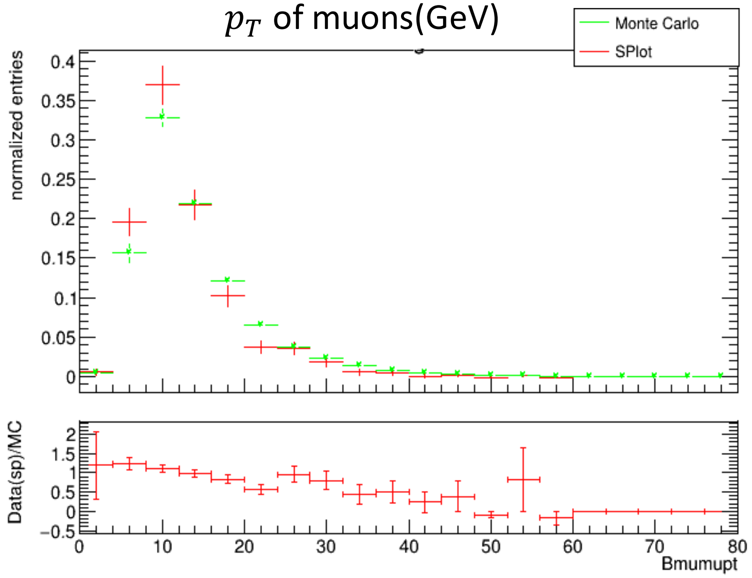
p_T of B meson (GeV)



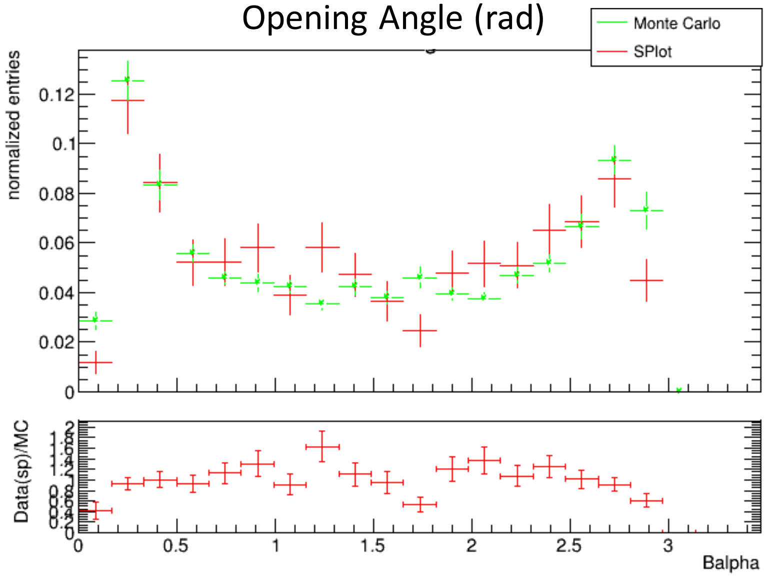
p_T of tracks (GeV)



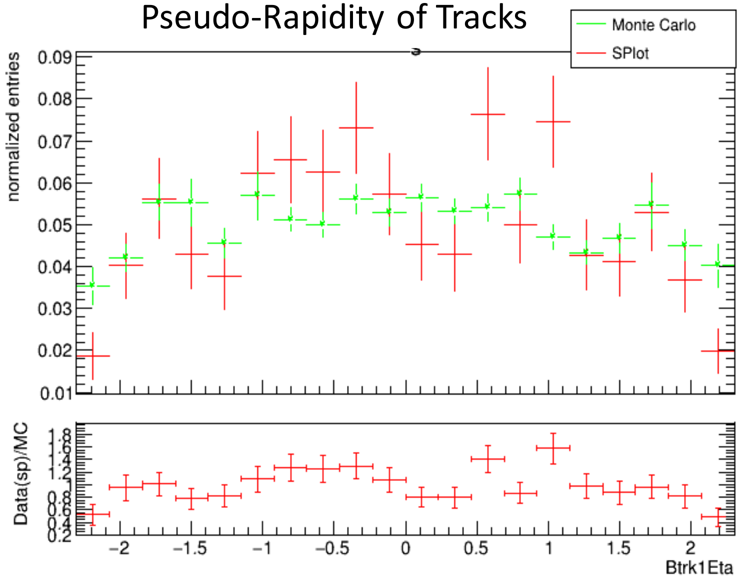
p_T of muons(GeV)



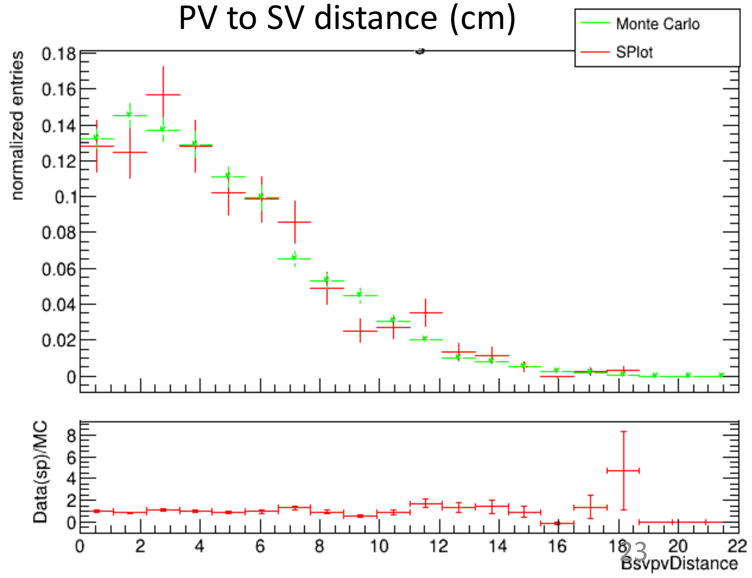
Opening Angle (rad)



Pseudo-Rapidity of Tracks

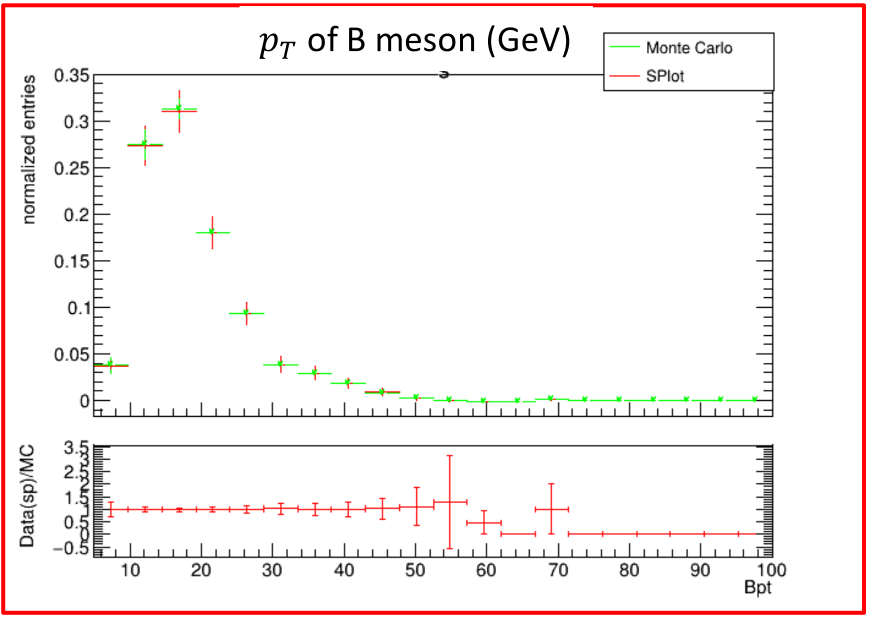


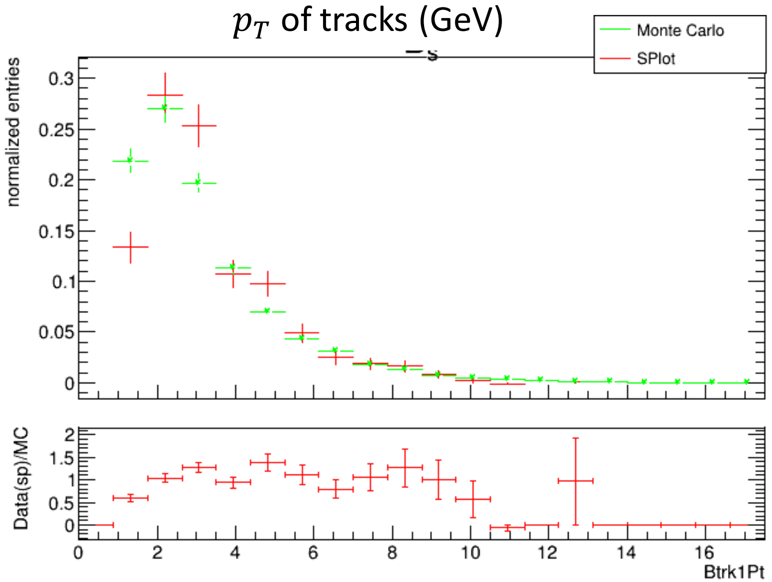
PV to SV distance (cm)

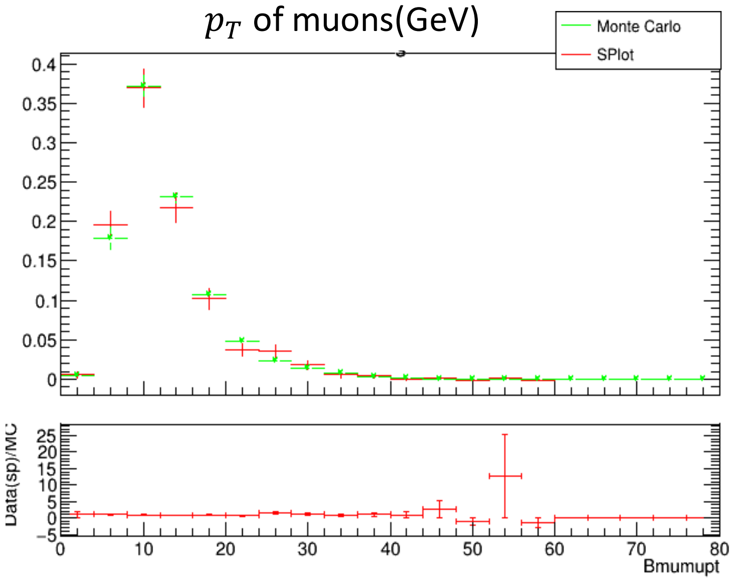


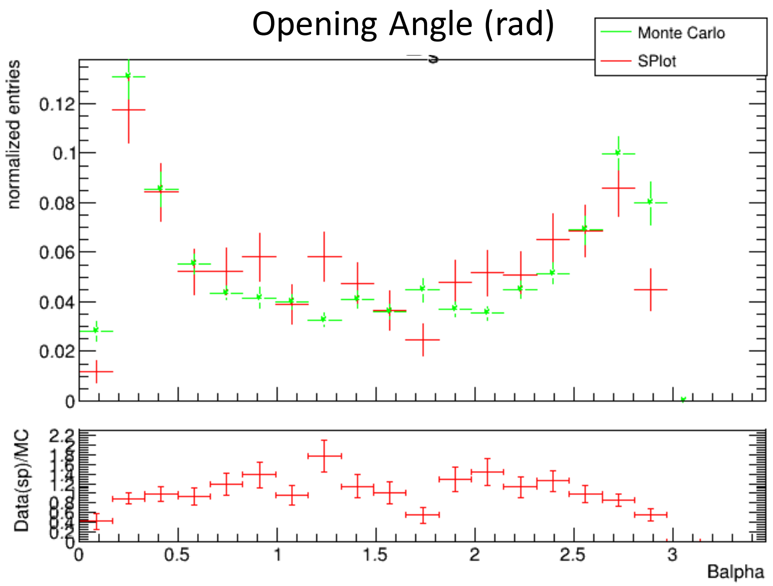
Reweighted MC

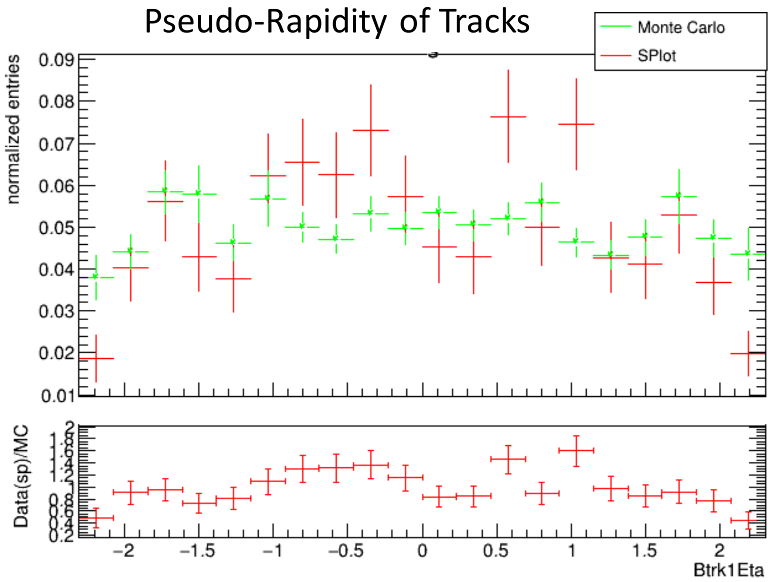
Remaining disagreement can be propagated as systematic uncertainty

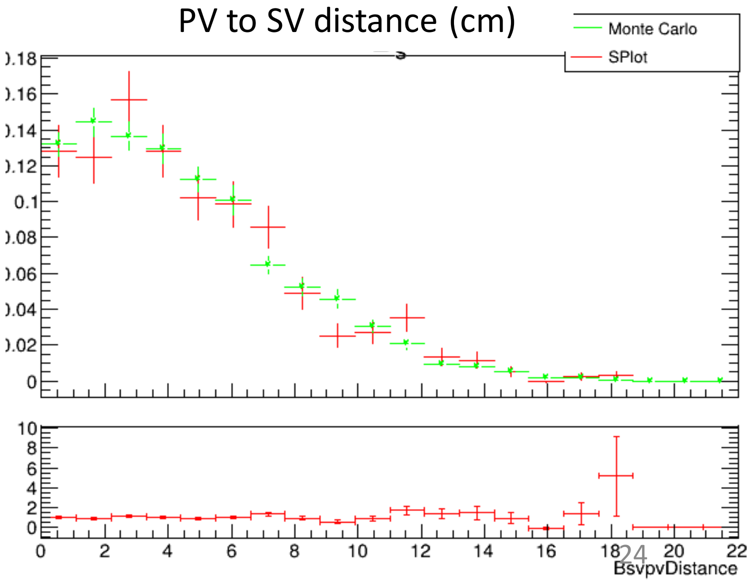












Differential Cross Section and Systematic Uncertainties

Efficiency

Efficiency: determined from MC simulations, measures how much signal is not reconstructed or rejected by selection cuts.

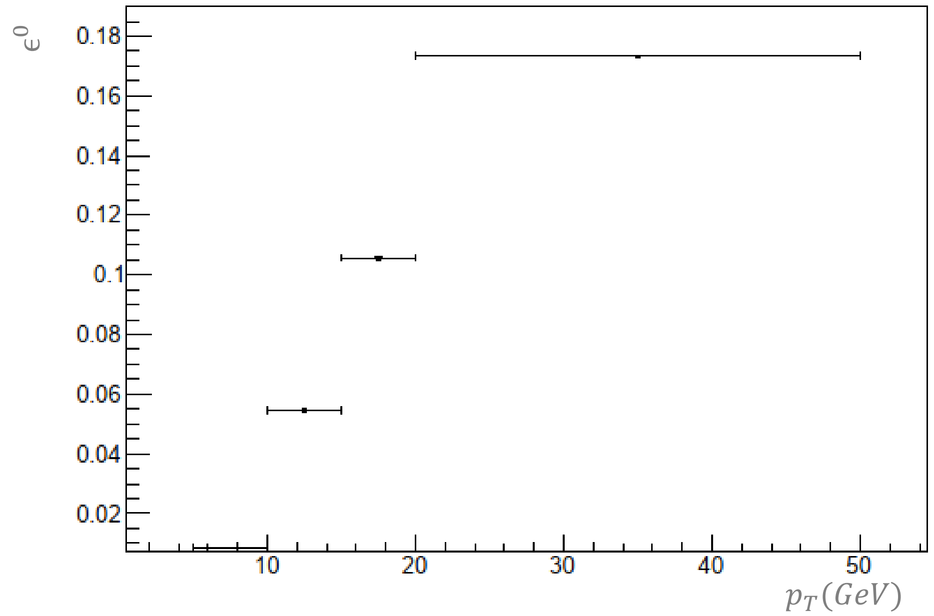
$$\epsilon = \frac{N_{after\ cuts}}{N_{before\ cuts}}$$

Efficiency Systematic Uncertainty: quantifies data-MC disagreement

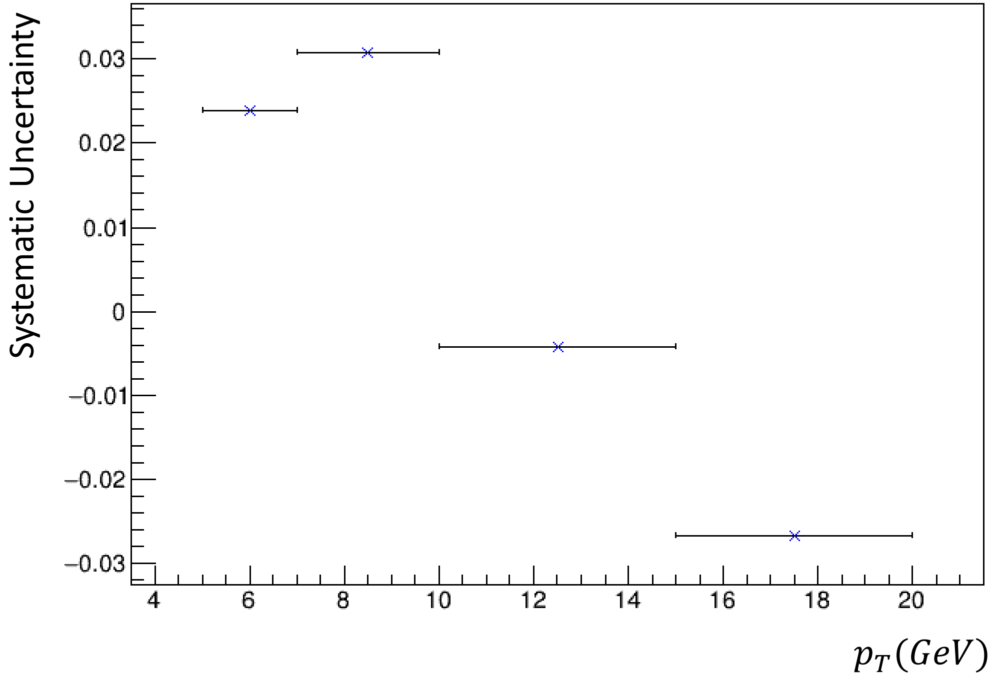
ϵ^0: without data/MC weights
ϵ^1: with data/MC weights

$$\Delta = \frac{\epsilon^1 - \epsilon^0}{\epsilon^0}$$

Nominal Efficiency ϵ^0



BDT efficiency systematic error



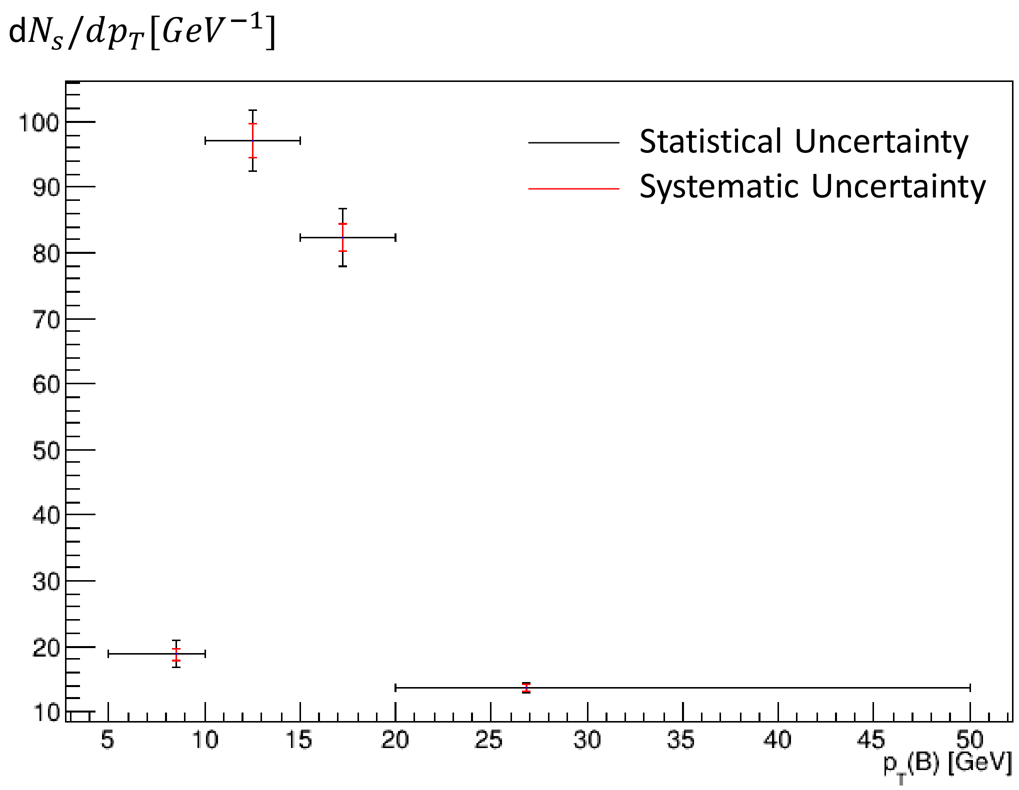
Normalized Signal Yield

1. Divide p_T range into several bins
2. Exctract signal yield from fit (N_s)
3. Place abcissa at p_T **mean**, evaluated with sPlot
4. Normalize raw yield to p_T bin width ($N_s / \Delta p_T$)

Statistical Uncertainty: from fit (σ_{stat})

Systematic Uncertainty: compare signal yield results obtained with diferente PDF models with model used in EUML (nominal)

$$\sigma_{syst\_raw} = \sqrt{\sigma^2_{bkg\_poly} + \sigma^2_{fit\_range} + \sigma^2_{sig\_1gauss}}$$



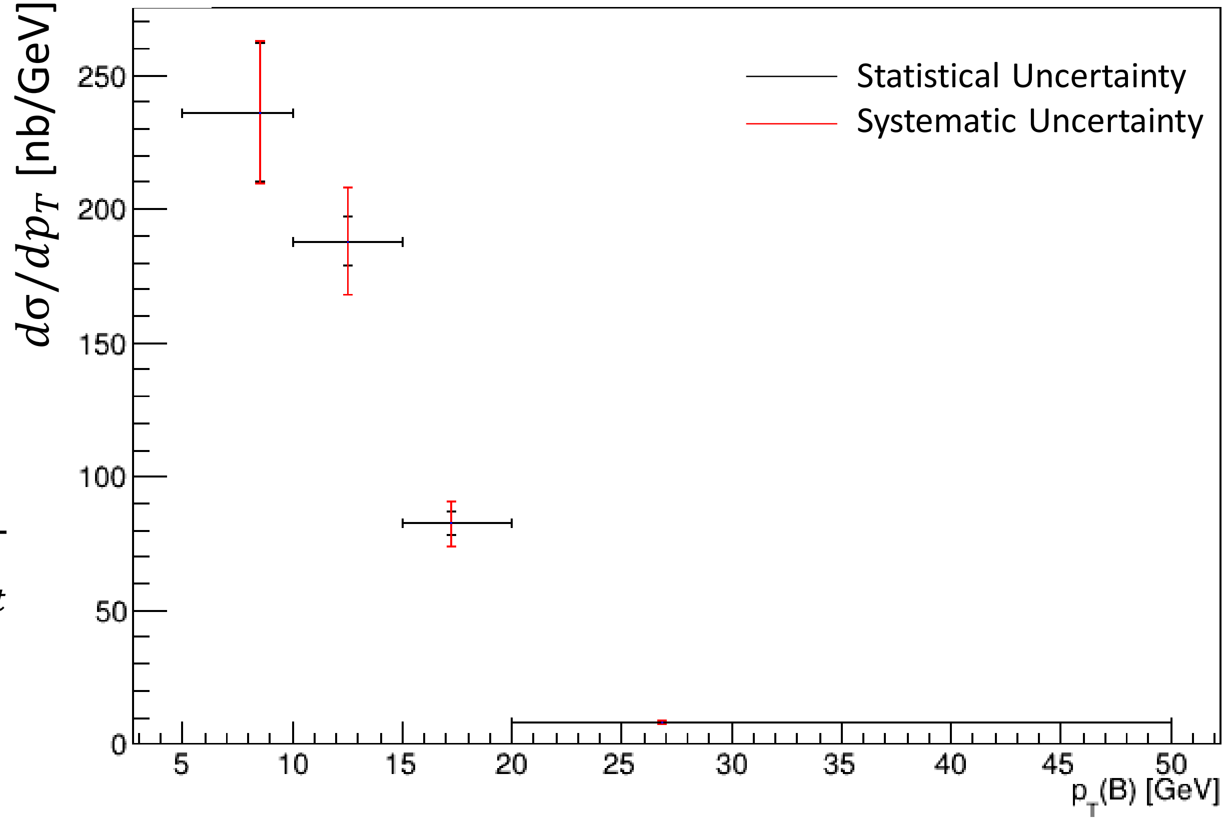
Nominal : sig (2 gaussians) + bkg (exponential)
Bkg_Poly : bkg (1st order polynomial)
Fit_Range : excludes left sideband region from fit
Signal_1Gauss : sig (1 gaussian)

Cross Section

$\sigma_{lumi\_syst} = 2.3\%$

$\sigma_{branch\_syst} = 9.6\%$

$$\sigma_{syst} = \sqrt{\sigma^2_{lumi\_syst} + \sigma^2_{branch\_syst} + \sigma^2_{raw\_syst} + \sigma^2_{eff\_syst}}$$



p_T (GeV)	$d\sigma/dp_T$ (nb/GeV)	σ_{stat} (nb/GeV)	σ_{syst} (nb/GeV)
5-10	236	26	27
10-15	188	9.2	21
15-20	82.5	4.4	8.4
20-50	8.28	4.5	9.0

Conclusions

- We have studied the production of B_s^0 mesons at LHC with CMS experiment
- Signal vs background discrimination was optimized with ML (BDT)
- Signal yield extracted from data with likelihood procedure
- Detector efficiency was estimated with MC, validated with sPlot and sideband subtraction
- Systematic uncertainties were estimated
- B_s^0 production cross section in pp collisions at 5 TeV was measured

Next Steps

- Extend study to B^+ meson (work ongoing)
- Measure the relative production B_s^0/B^+ (fragmentation fraction)
- Measure R_{AA} (nuclear modification factor), comparing PbPb with pp

Overall study probes QCD, both in vacuum (pp) and in QGP media (PbPb)

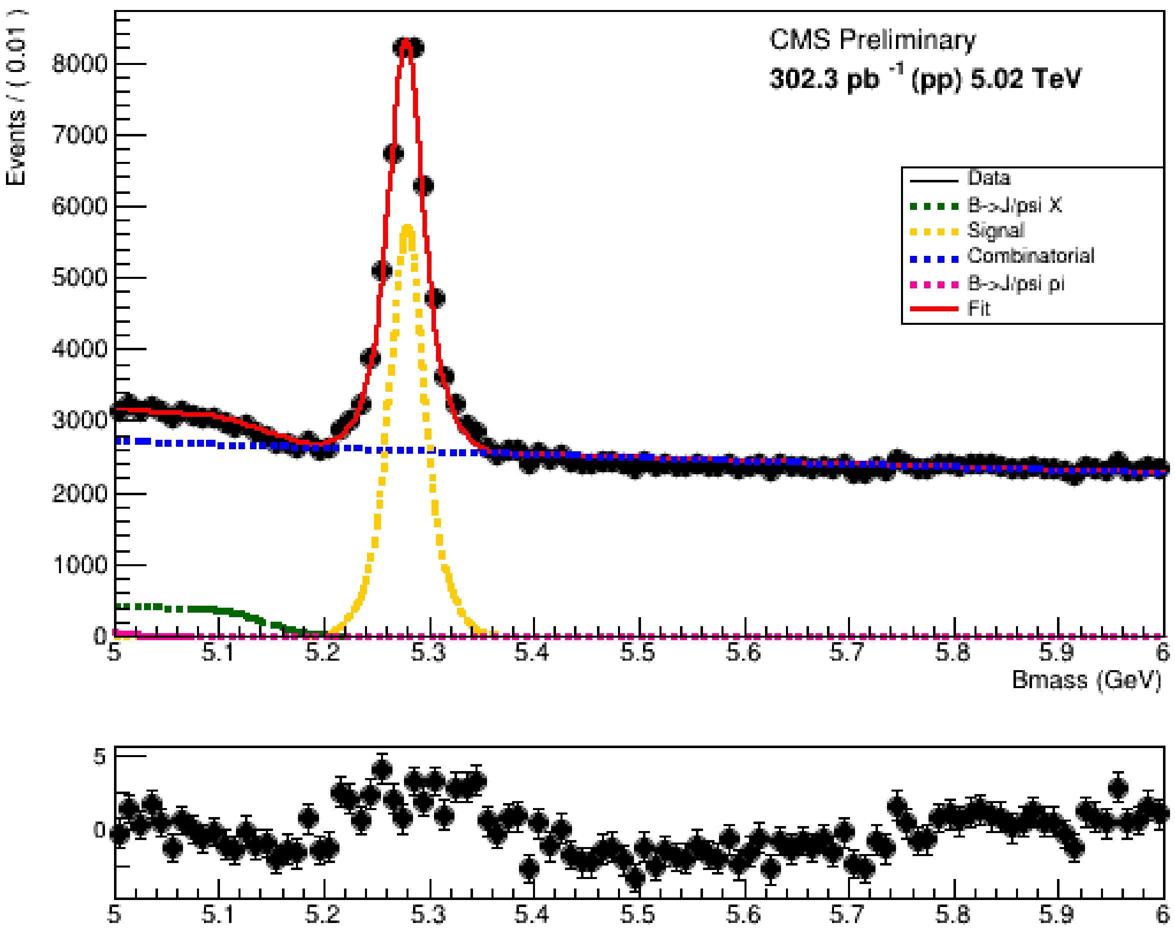
Questions?

Backup

Likelihood Method - B^+result

$$l(m_i) = N_S\, P_S(m_i; \mu, \sigma_1, \sigma_2) + N_{CB} P_{CB}(m_i; \lambda) + \boldsymbol{N_{erf} P_{erf}(m_i)} + \boldsymbol{N_{jpp}(fixed)}$$

B^+Invariant Mass Fit



$$P_S = \alpha\, Gauss(\mu, \sigma_1) + (1 - \alpha) Gauss(\mu, \sigma_2)$$

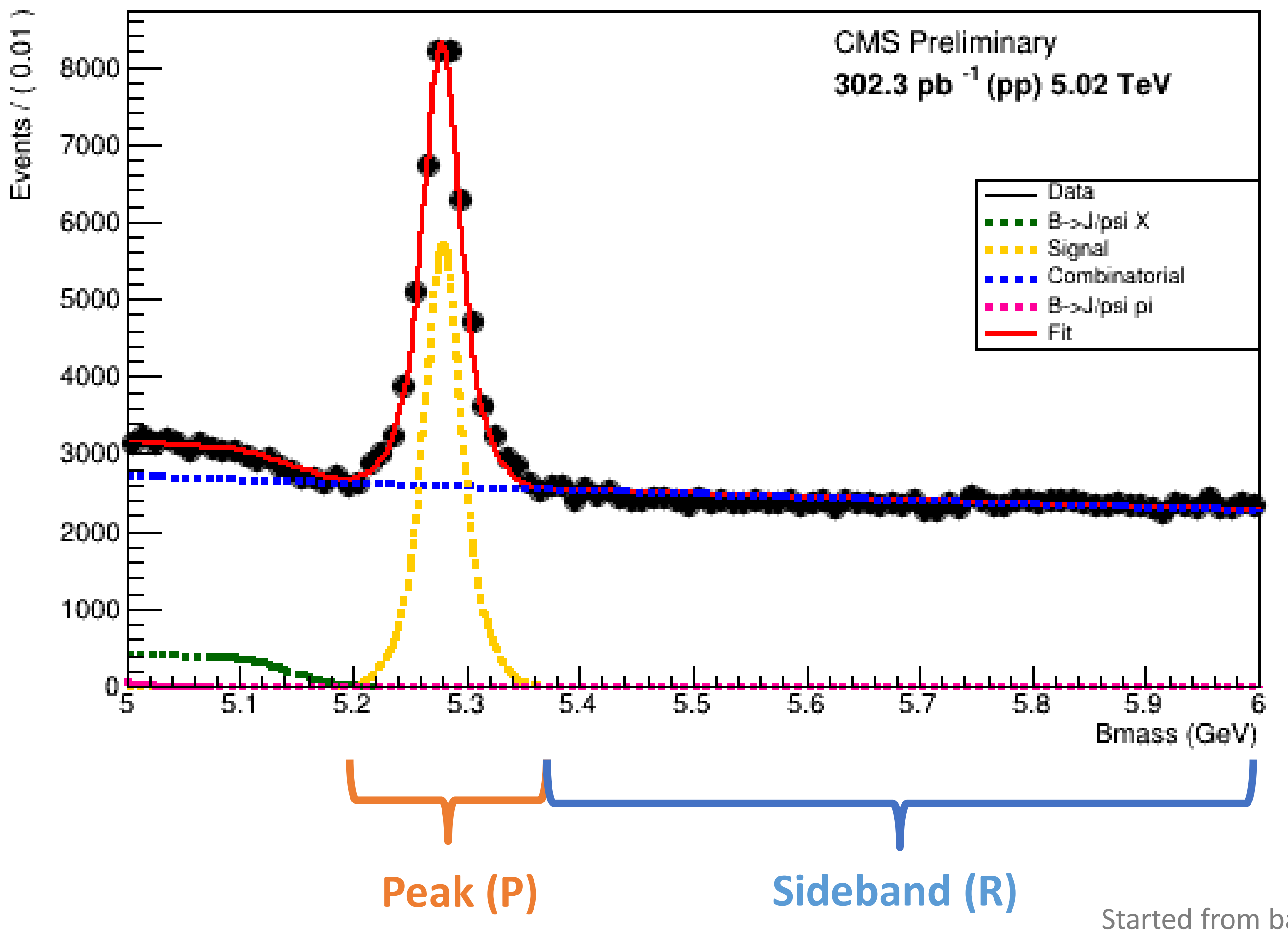
$$P_{CB} = Exp(\lambda) \qquad P_{erf} = Erf \qquad N_{erf} = f_{erf} N_S$$

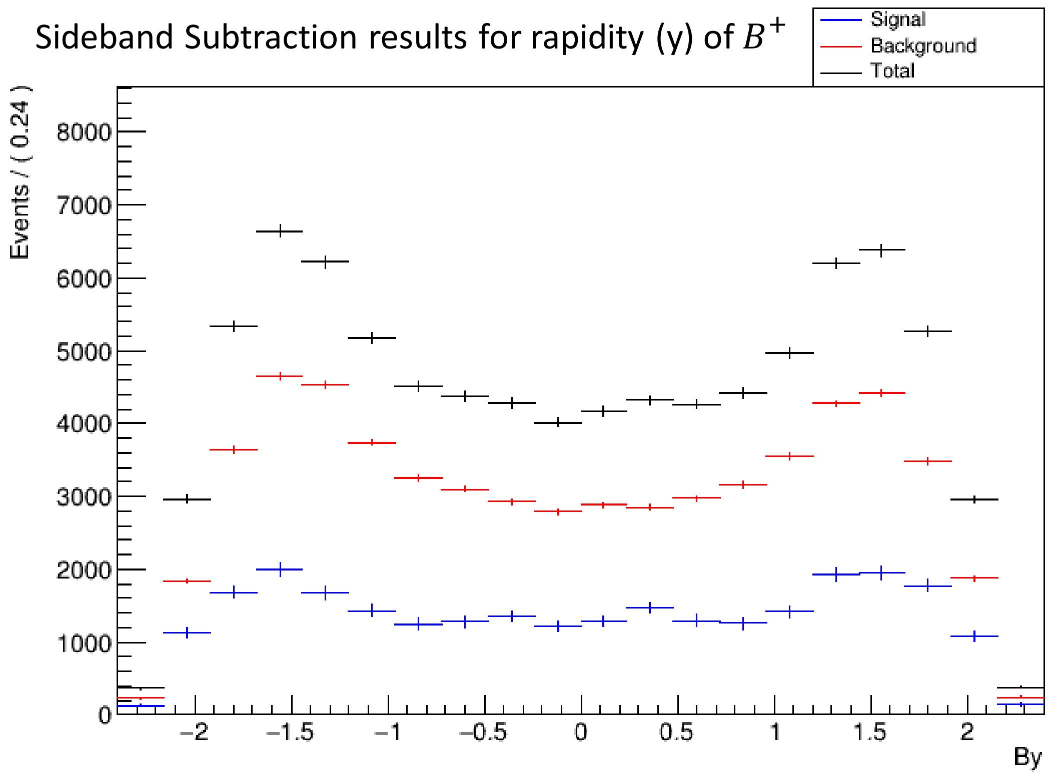
Coefficients	Value $\pm$ Statistical Uncertainty
α	0.551 ± 0.045
λ	-0.17937 ± 0.01204
μ	5.27896 ± 0.00019
σ_1	0.02667 ± 0.00110
σ_2	0.01351 ± 0.00053
N_S	26.450 ± 333
N_{CB}	248.846 ± 794
f_{erf}	0.2197 ± 0.0138

Sideband Subtraction

$$V_{signal} = V_{peak} - rV_{sideband}$$

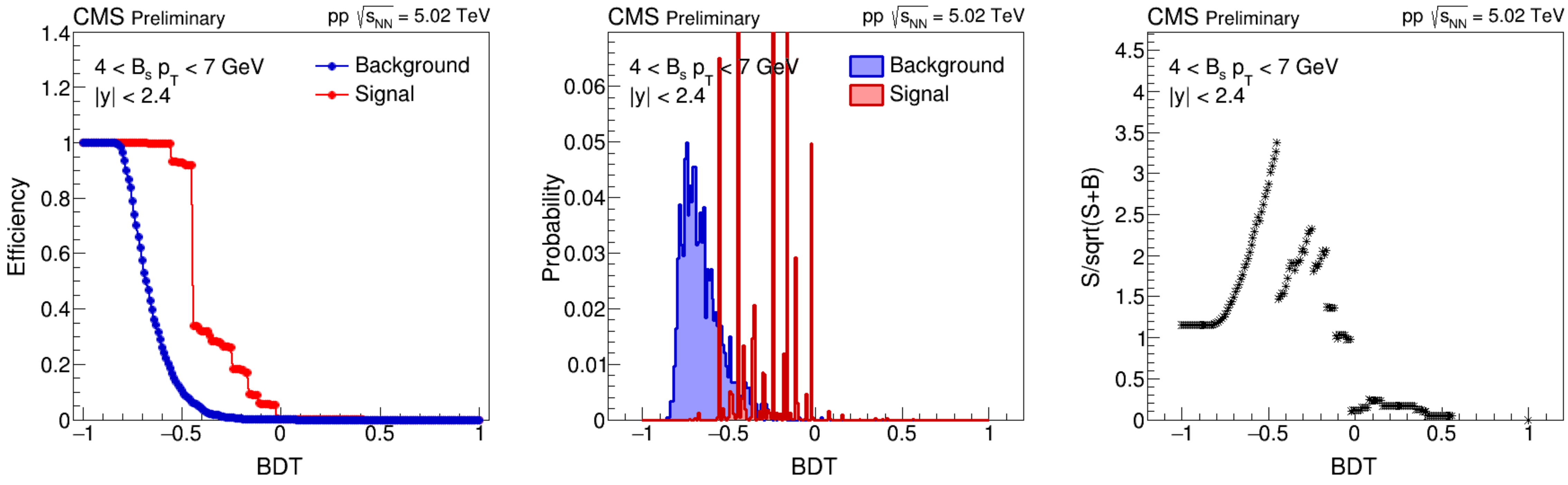
$$r = \frac{P}{R}$$



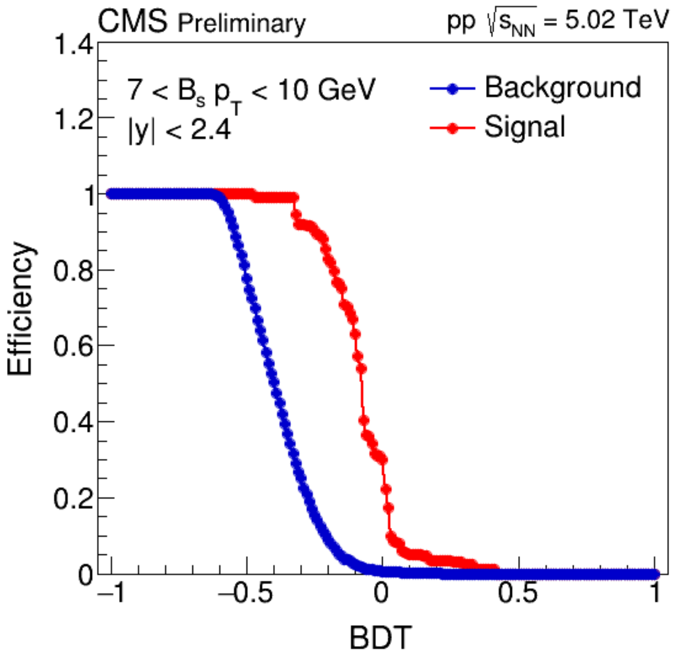


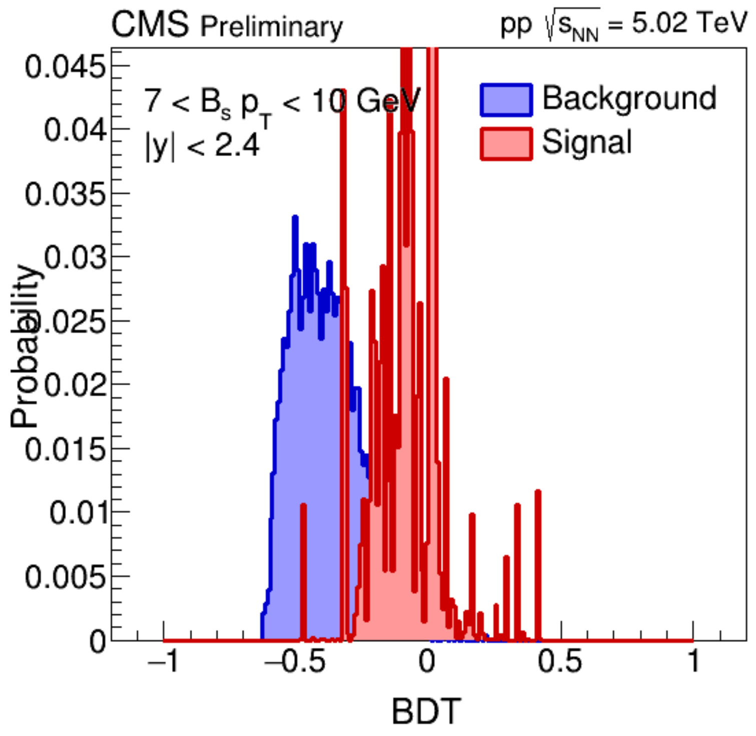
Started from baseline implementation for PbPb described in NA/2019-219 (A.Pardal, J.Gonçalves, et al.)

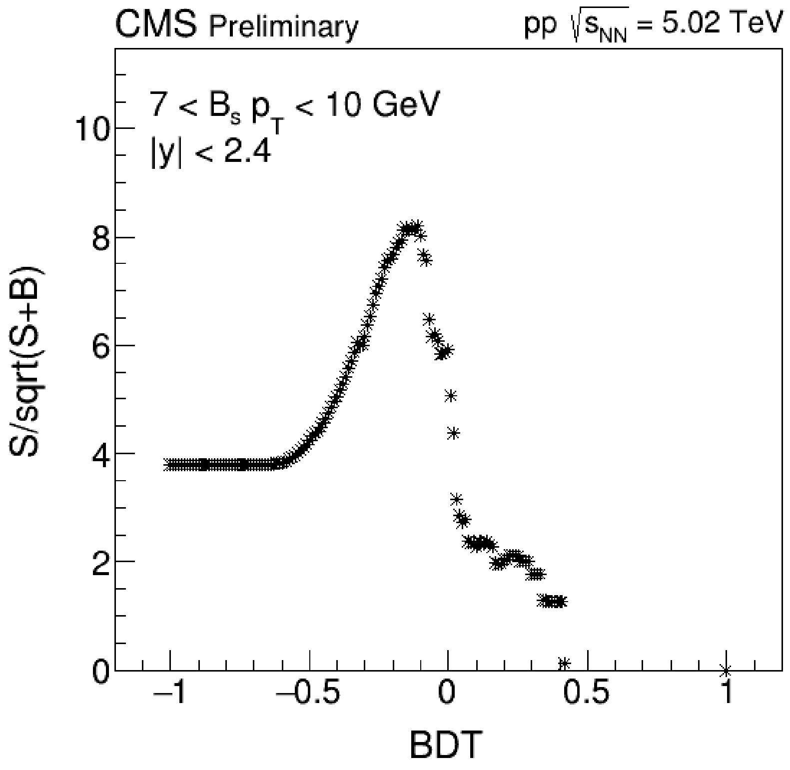
BDT training Pt Bin [4,7]



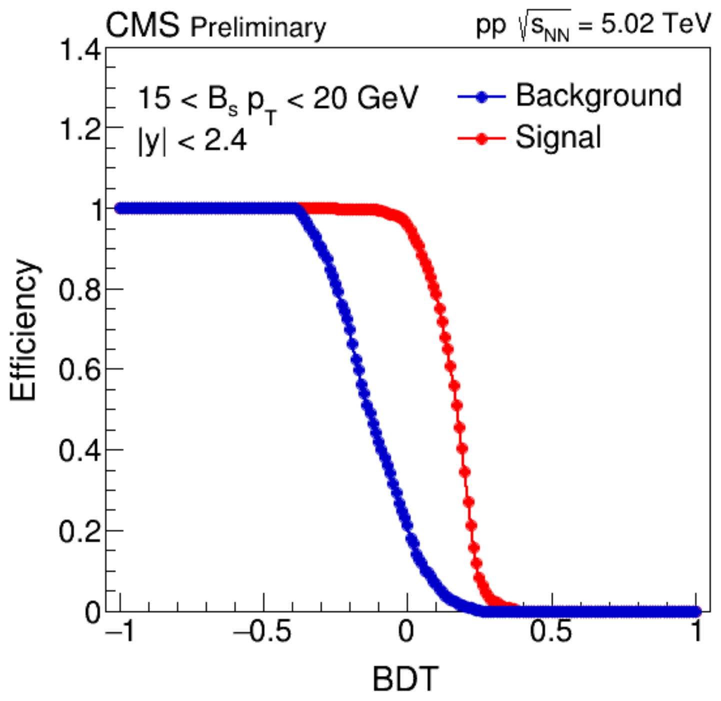
BDT training Pt Bin [7,10]

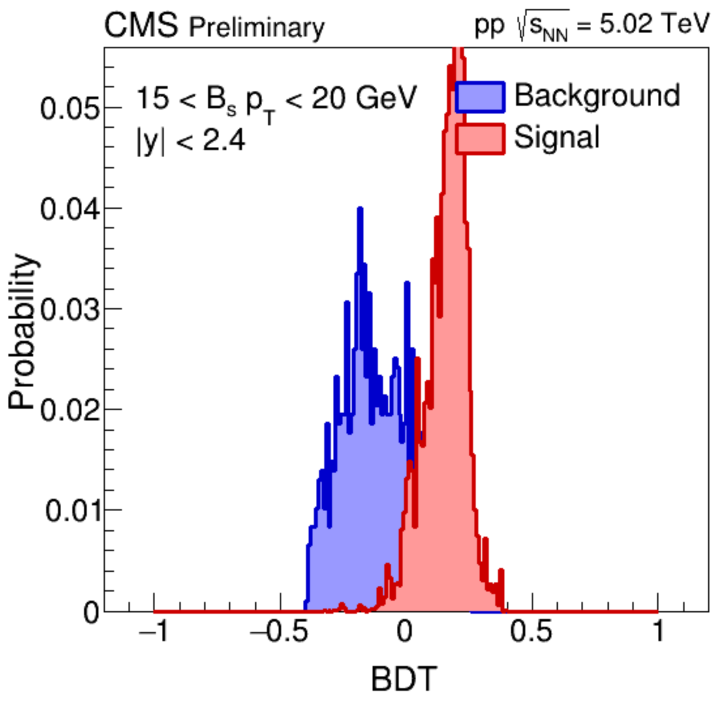


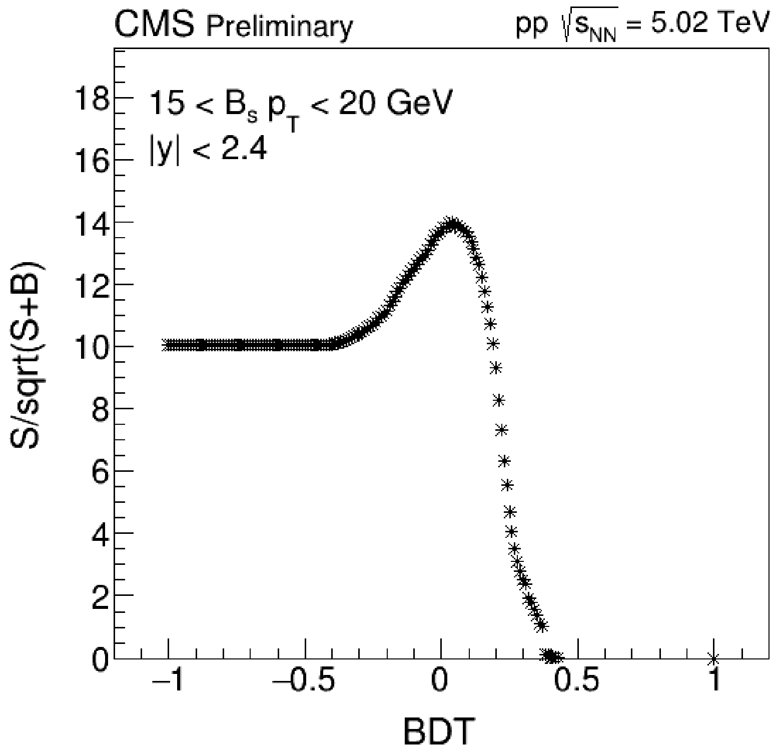




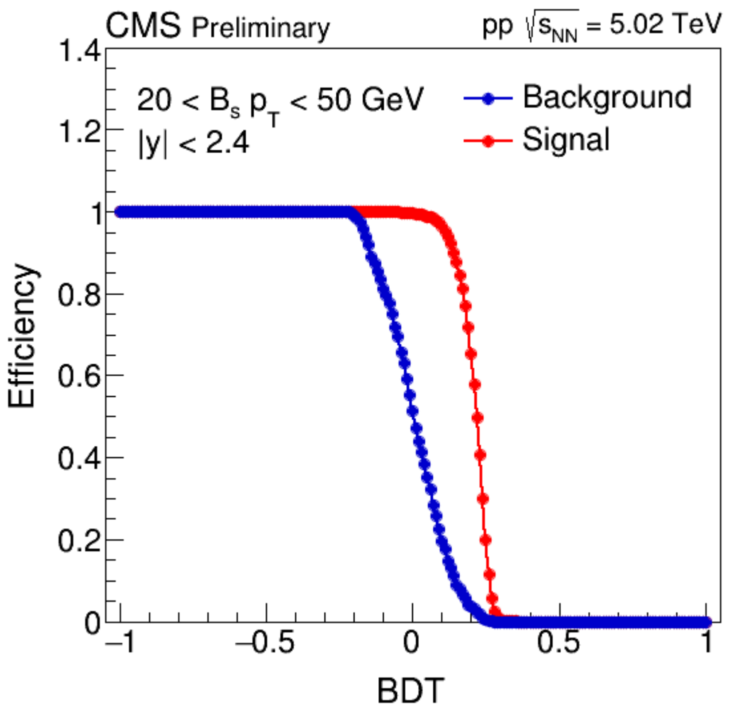
BDT training Pt Bin [15,20]

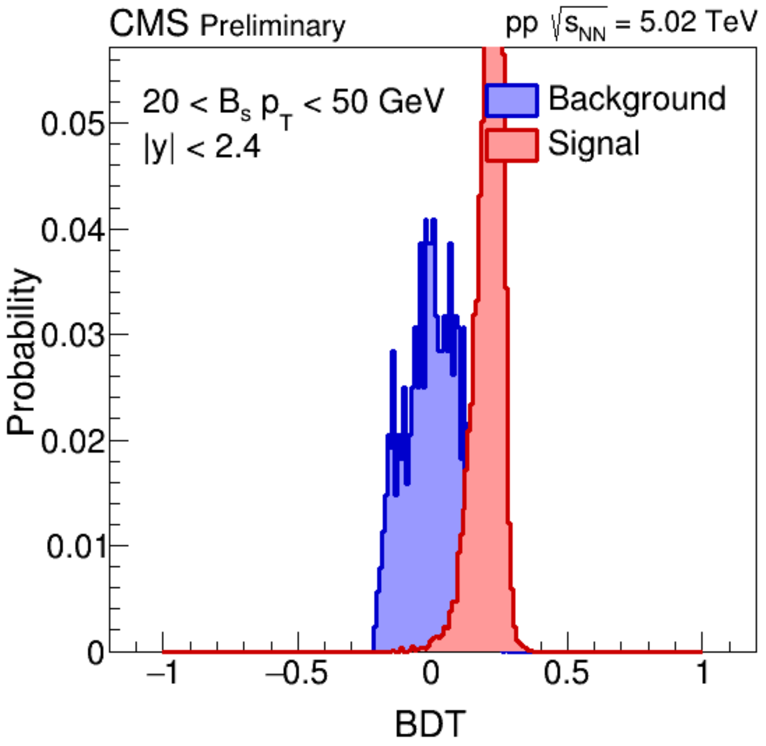


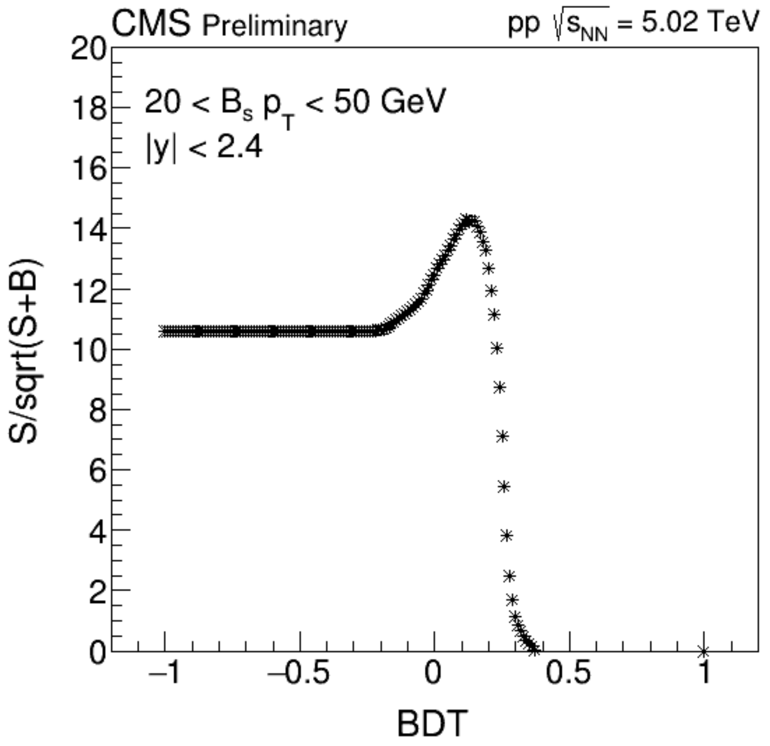




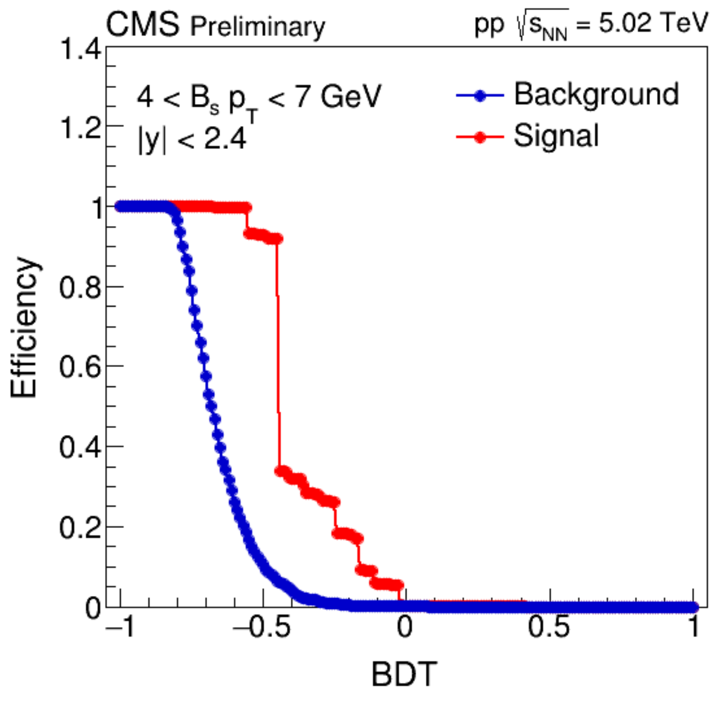
BDT training Pt Bin [20,50]

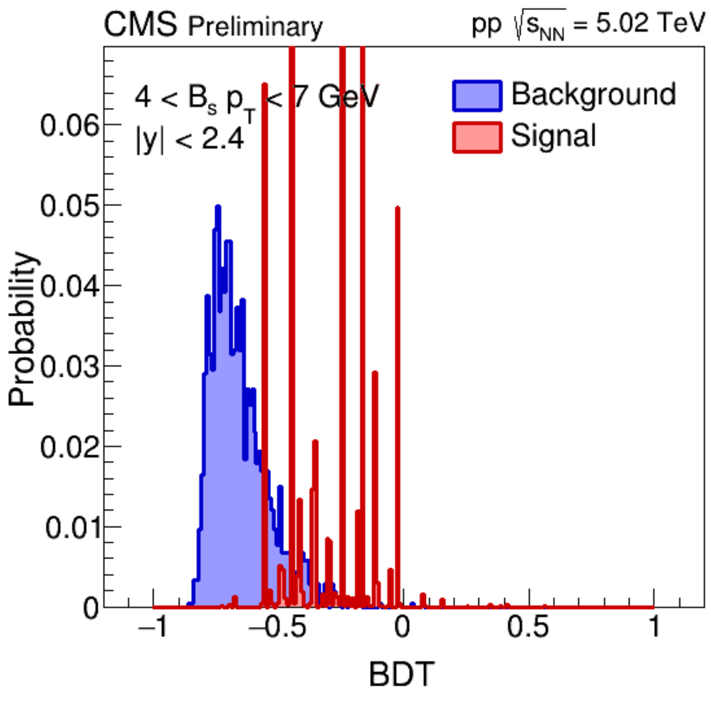


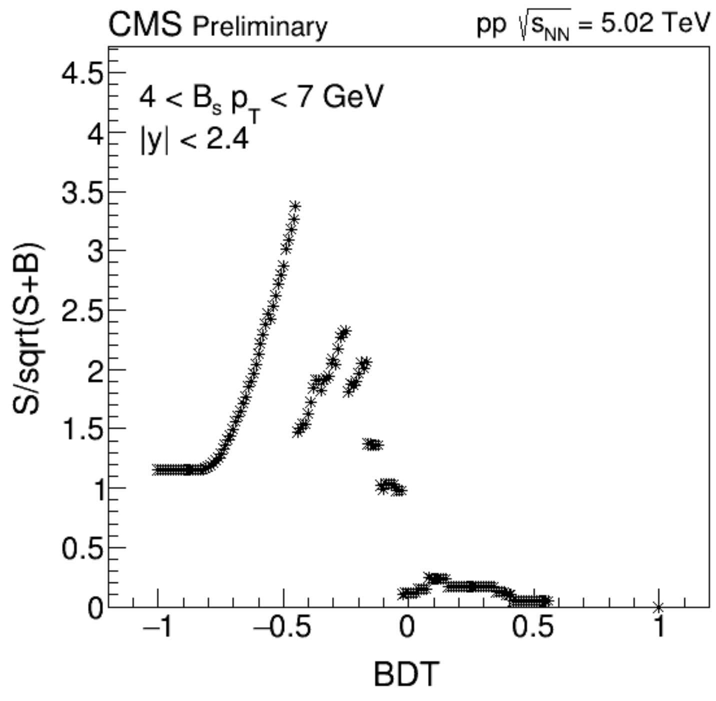


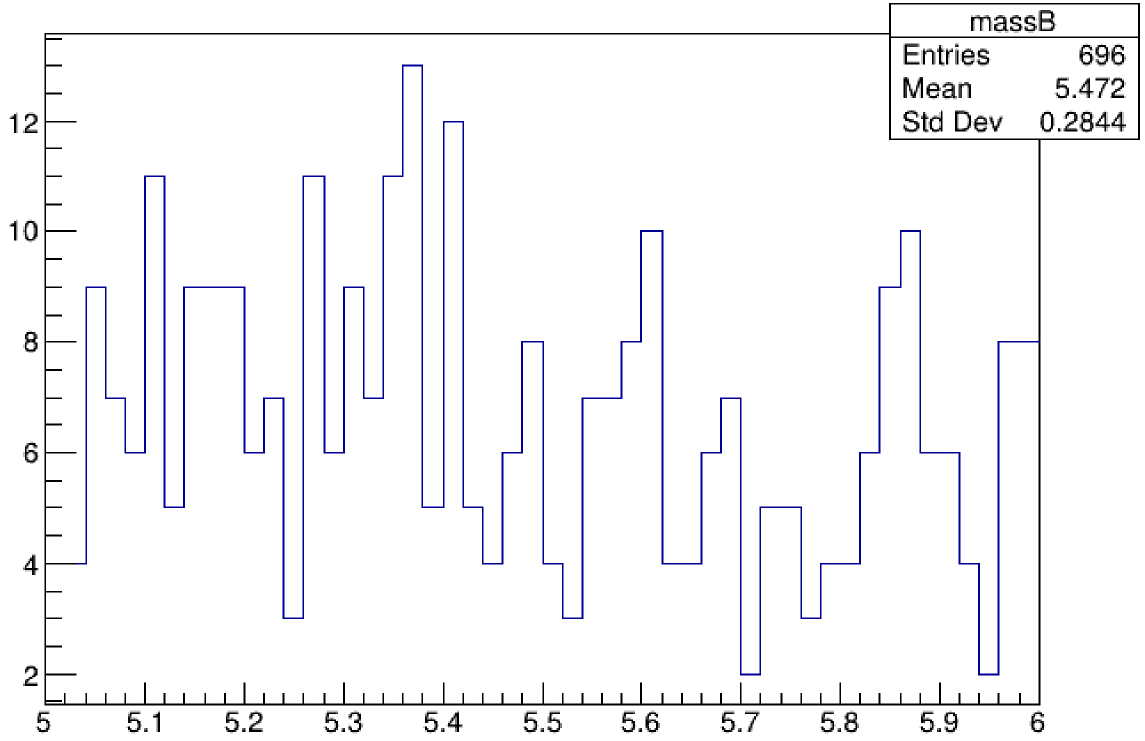


BDT training Pt Bin [4,7]

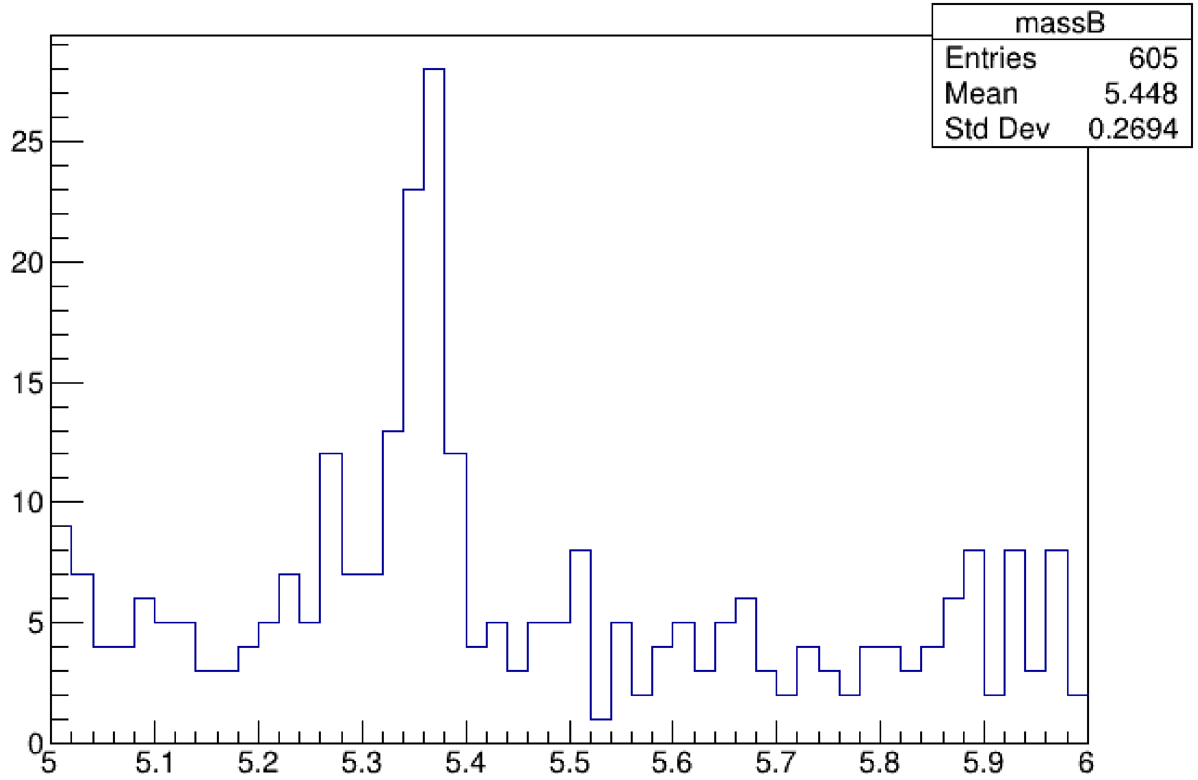






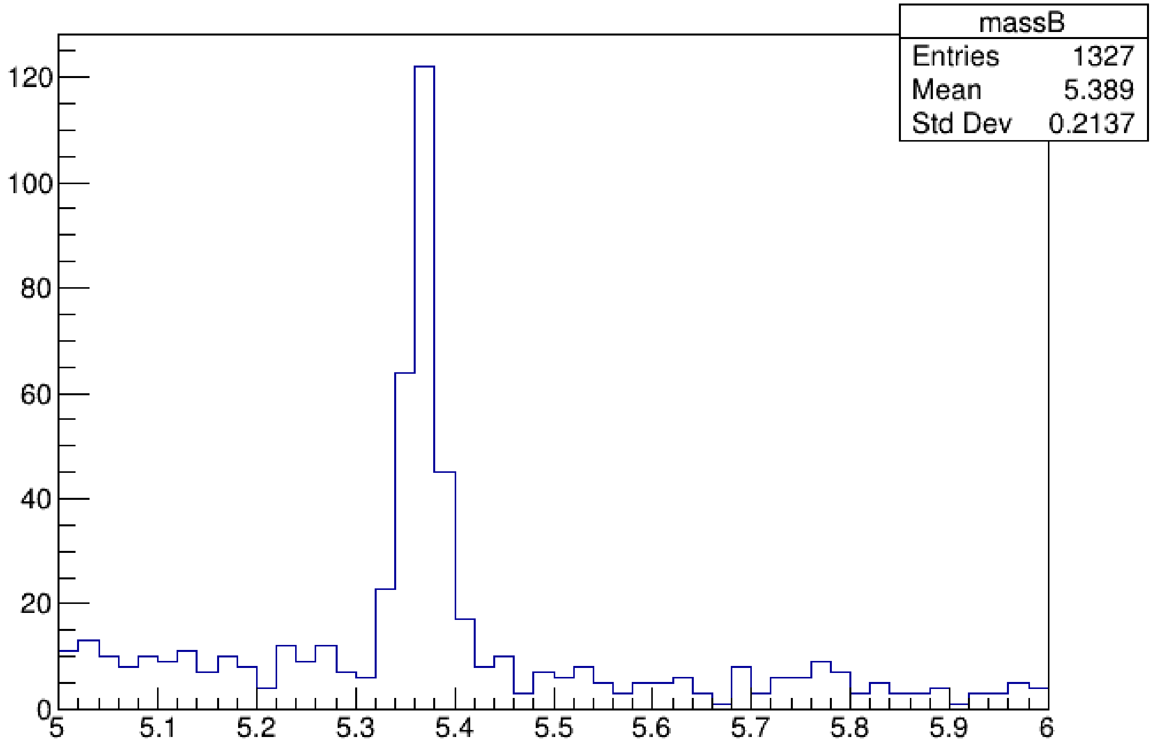


Bs mass reconstruction after BDT optimal working point cut in pt range 4-7 Gev/c

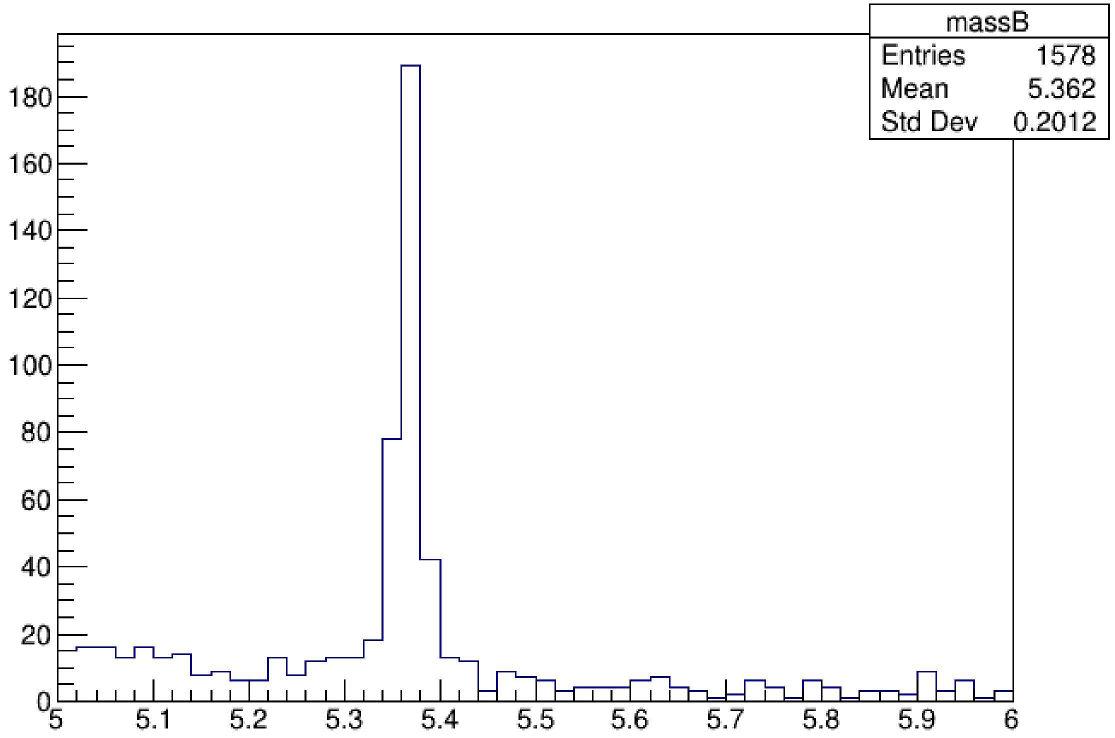


Bs mass reconstruction after BDT optimal working point cut in pt range 7-10 Gev/c

Backup

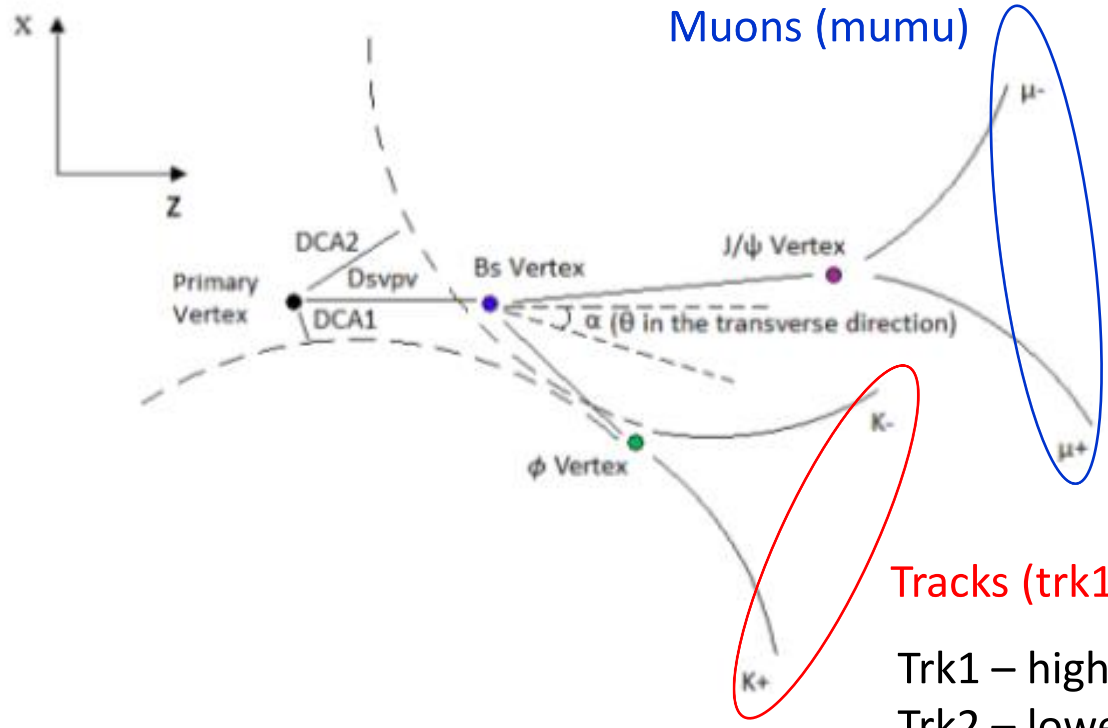


Bs mass reconstruction after BDT optimal working point cut in pt range 15-20 Gev/c



Bs mass reconstruction after BDT optimal working point cut in pt range 20-50 Gev/c

(Some) Variables used in the analysis



Trk1 – higher pT
Trk2 – lower pT

Bpt – Transverse momentum of B meson
Bmass – Invariant mass of B meson

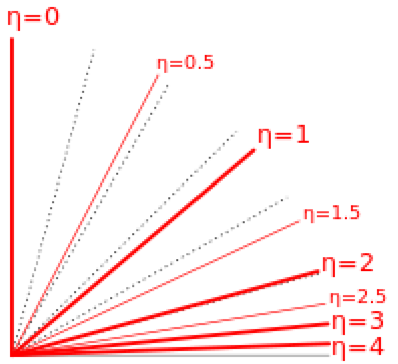
Rapidity (By)

$$y = \frac{1}{2} ln\left(\frac{E + p_z c}{E - p_z c}\right)$$

$y \rightarrow 0$: particle directed in XY plane
$y \rightarrow +\infty$: particle along beam axis (positive z)
$y \rightarrow -\infty$: particle along beam axis (negative z)

Pseudo-Rapidity (eta)

$$\eta = -ln\left(tan\left(\frac{\theta}{2}\right)\right)$$



Dsvpv – distance between PV and SV (3D)

DCA – Distance of Closest Approach between kaon's /muon's trajectory and PV
Dxy, dxy – xy componente of DCA (kaon,muon)
Dz,dxy – z componente of DCA (kaon,muon)

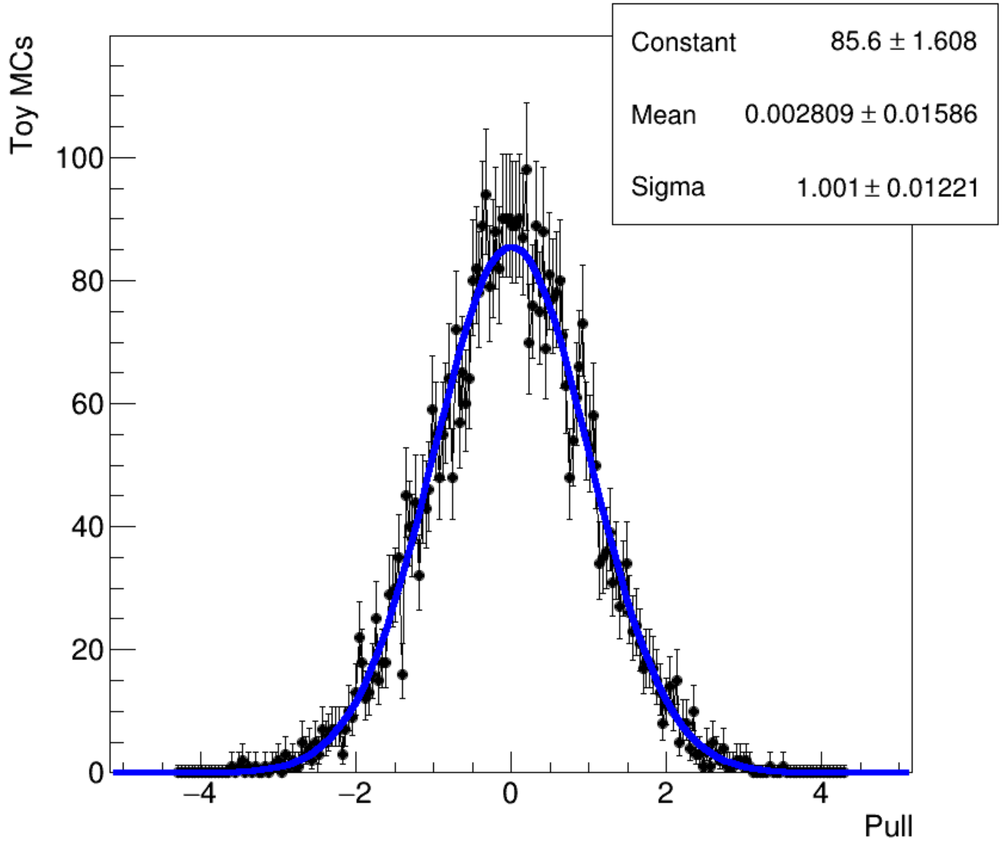
Fit Validation

- Fit is unbiased if it gives the correct value for N_s
- Test this by generating 5000 pseudo-experiments (toy MC)
- Fit generated pseudo-data with the same function that was used in fitting the data

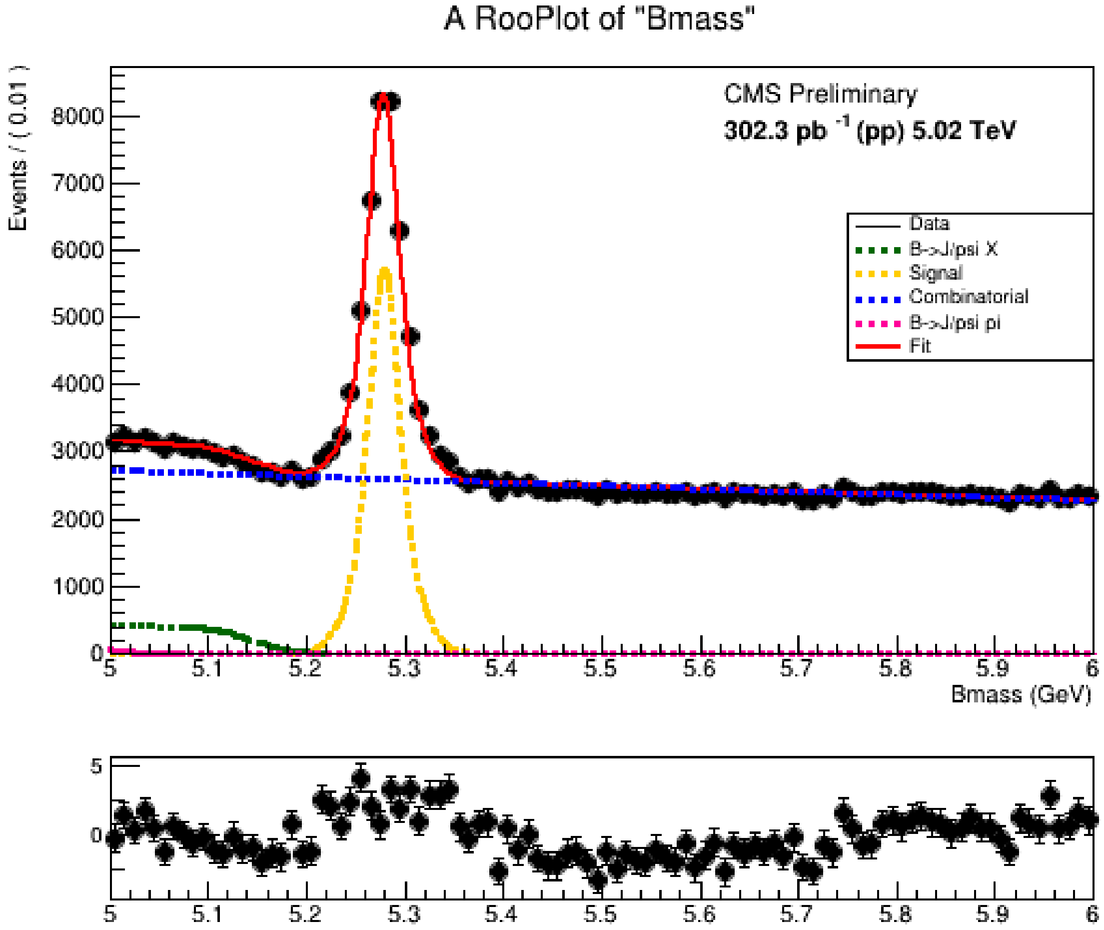
$$Pull = \frac{N_i - N_s}{\sigma_i}$$

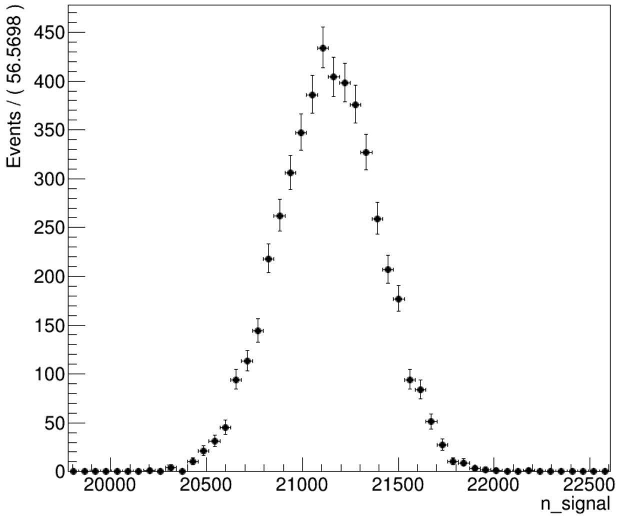
N_i: signal yield of pseudo-data i
N_s: signal yield of data
σ_i: uncertainty on N_i

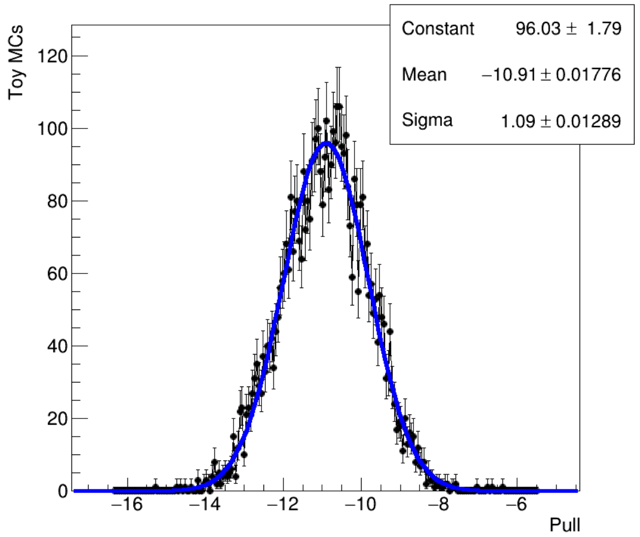
Expected: Pull distribution mean close to zero and sigma close to 1



Fit Validation - B^+ (5 < pt < 100 GeV)

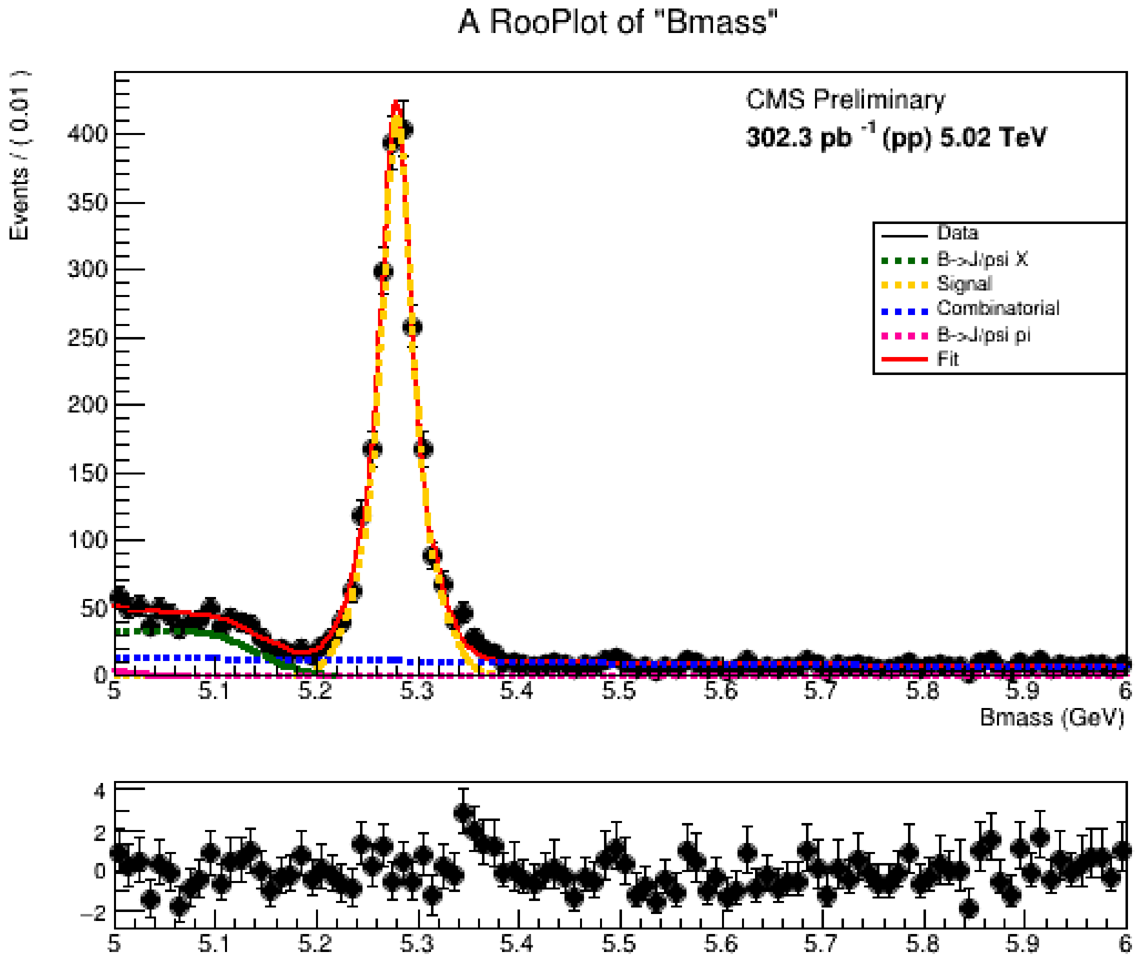


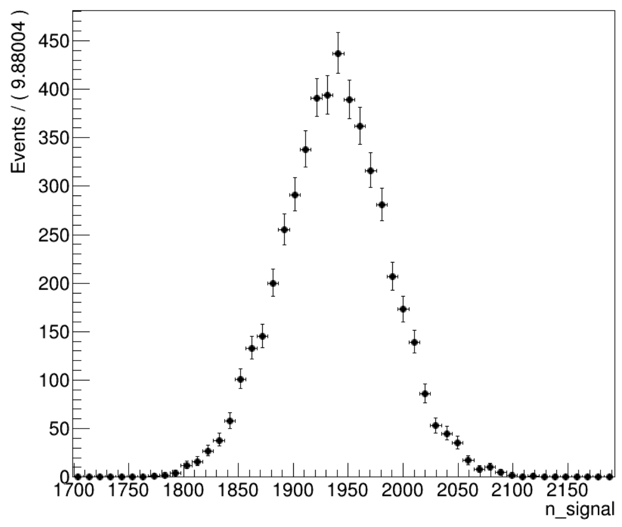


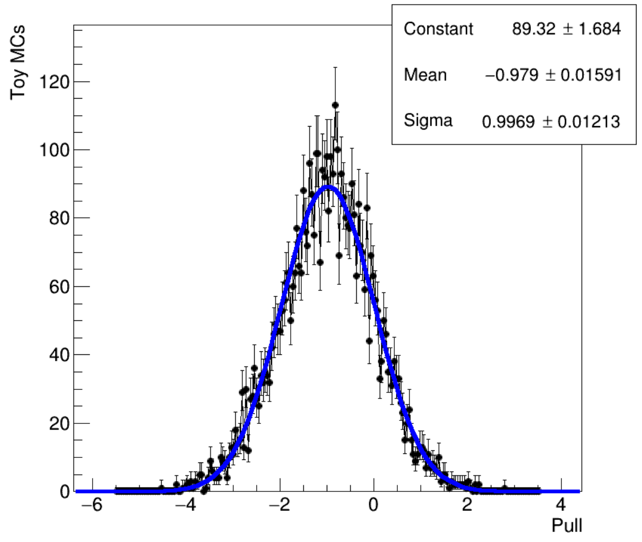


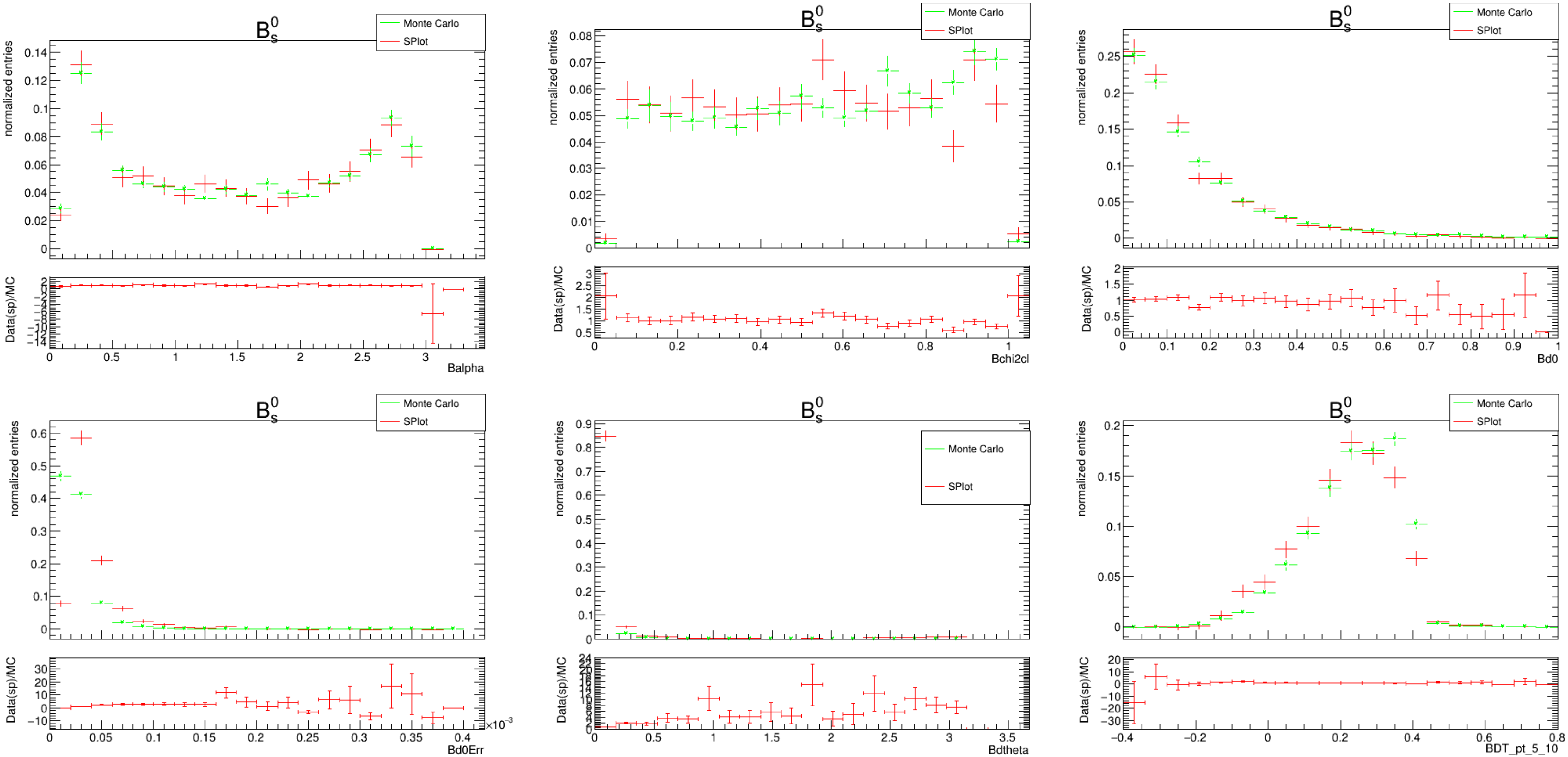
$$N_S = 26.450 \pm 333$$

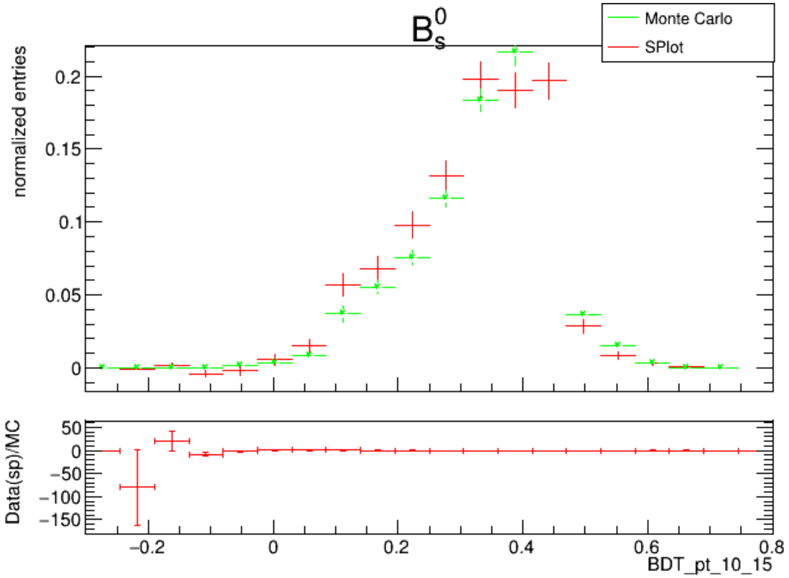
Fit Validation - B^+ (5 < pt < 50 GeV)

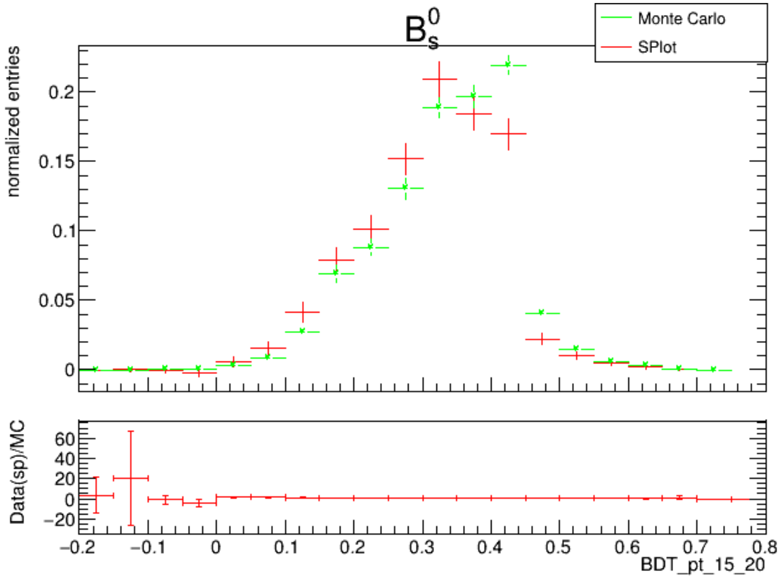


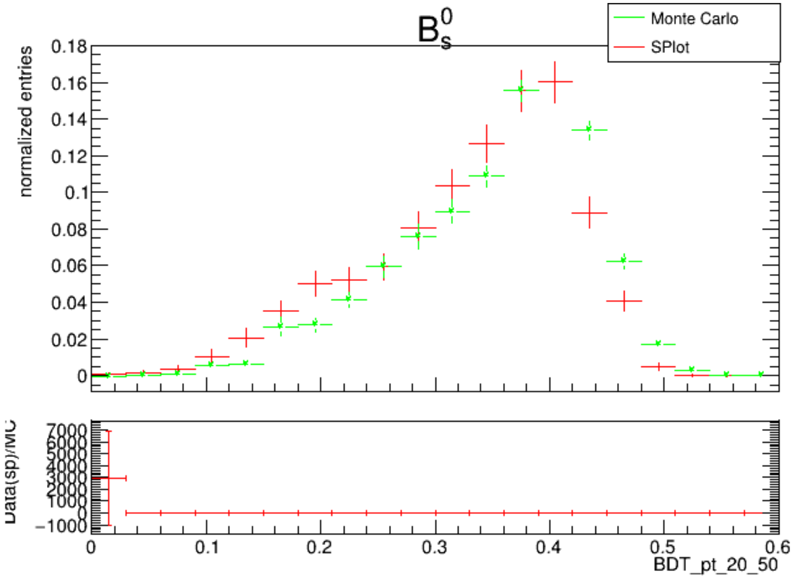


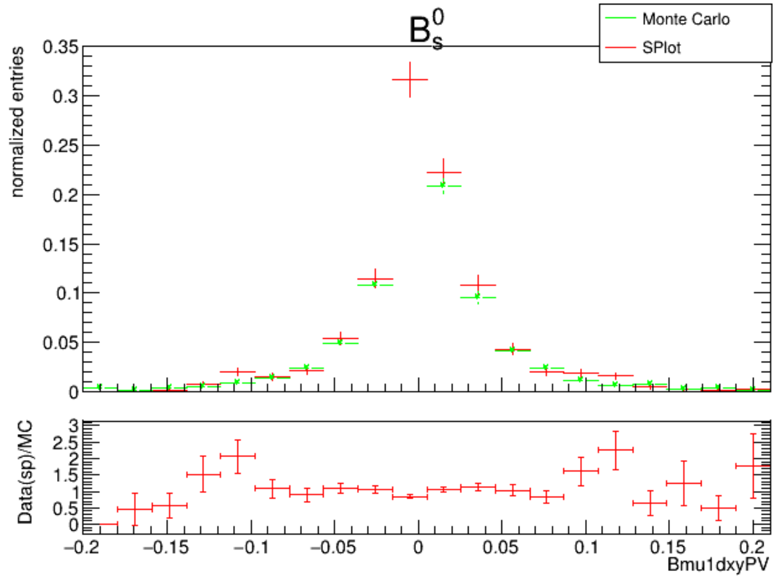


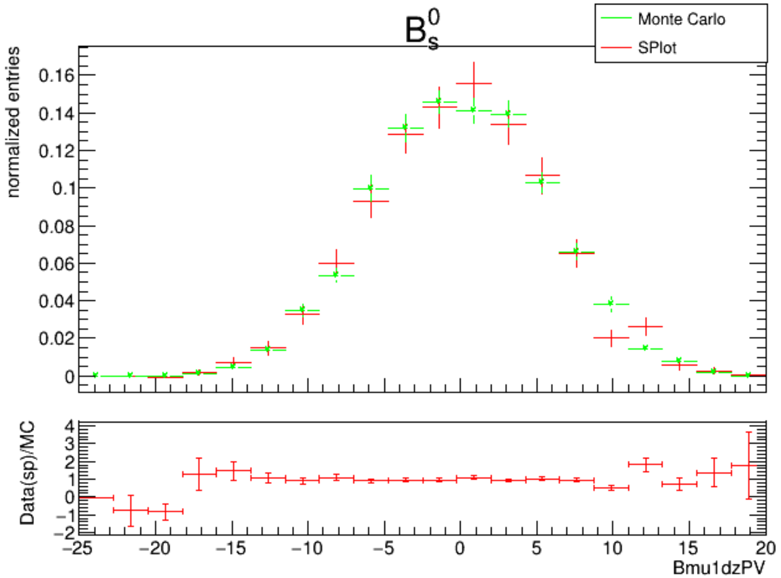

B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Balpha
Bchi2cl
Bd0
Bd0Err
Bdtheta
BDT_pt_5_10

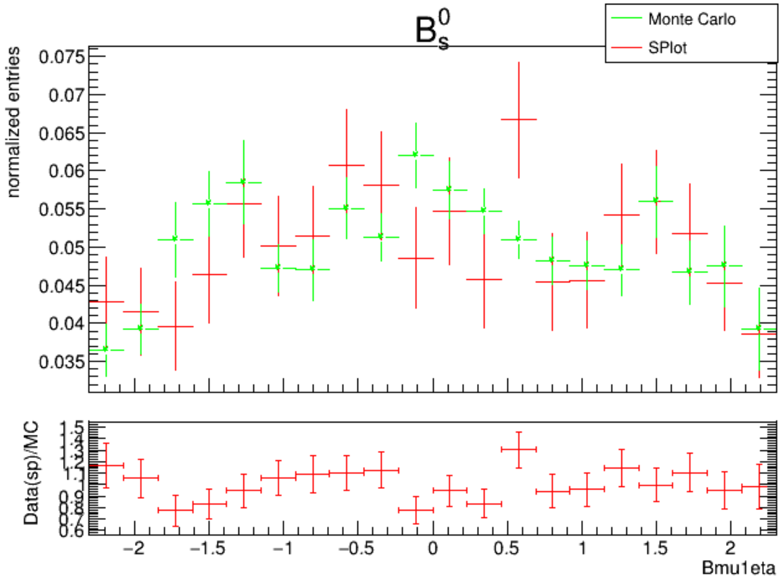

B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
BDT_pt_10_15

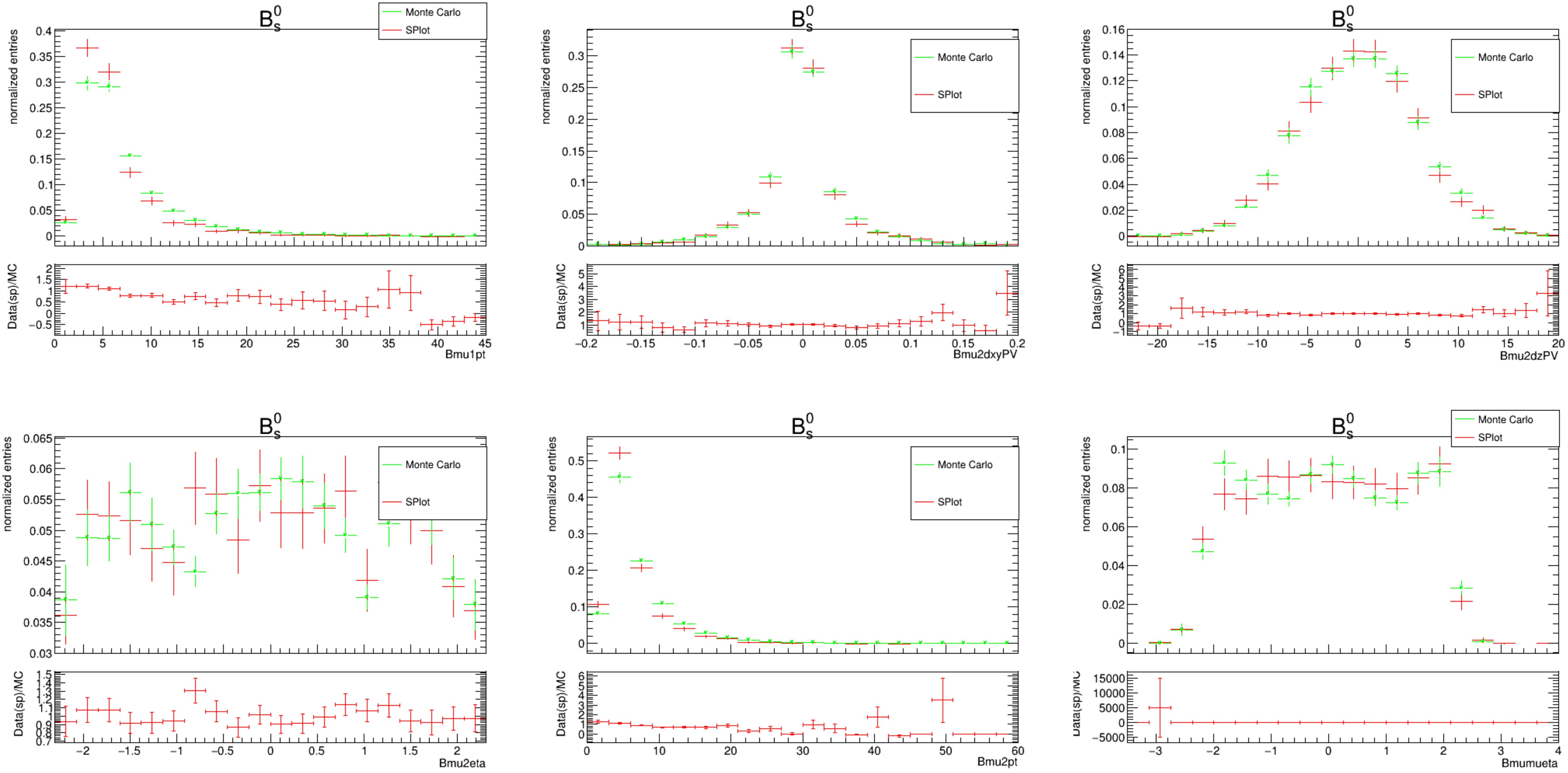

B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
BDT_pt_15_20


B_s^0
Monte Carlo
SPlot
normalized entries
BDT_pt_20_50

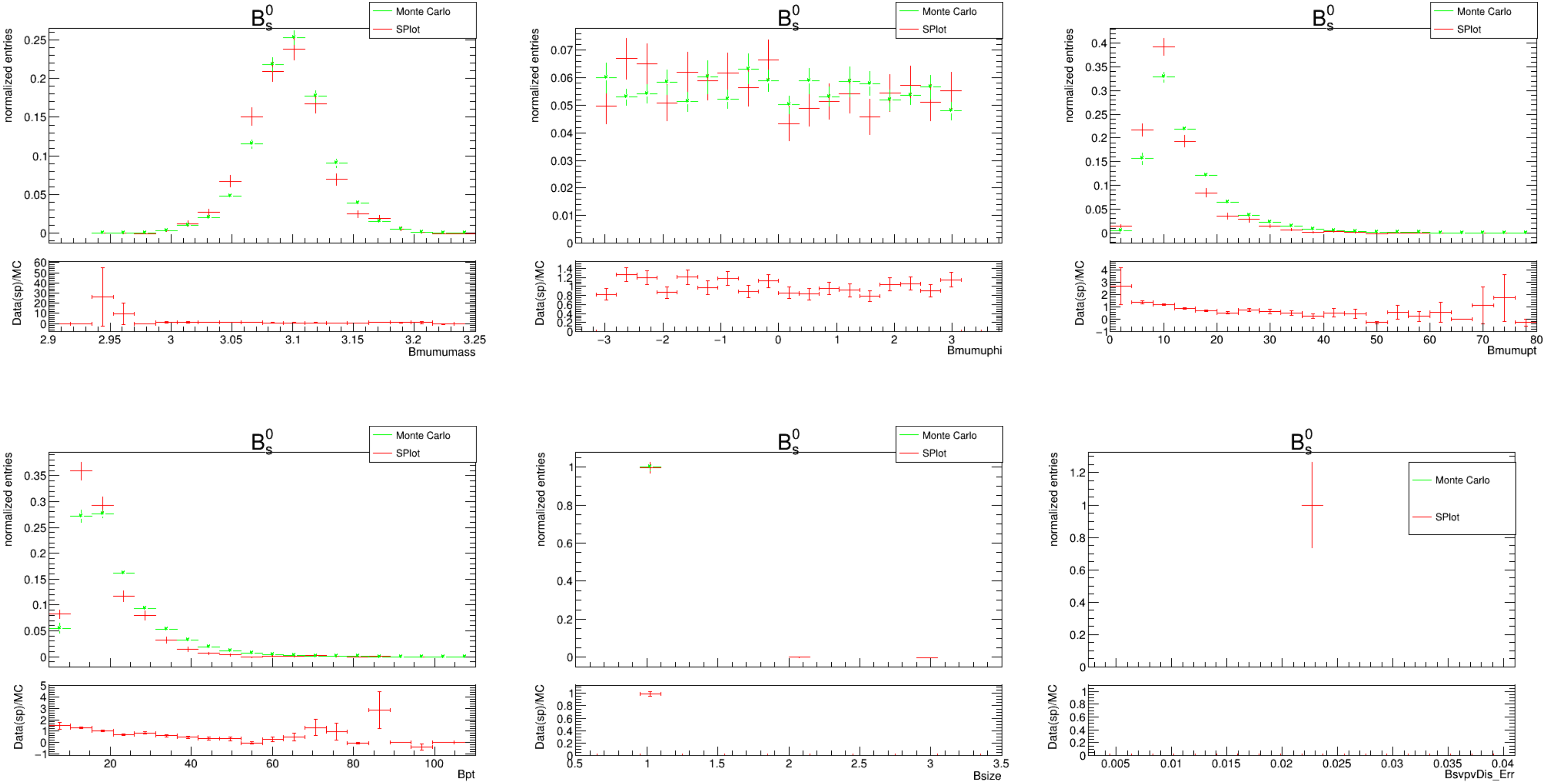

B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Bmu1dxyPV


B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Bmu1dzPV

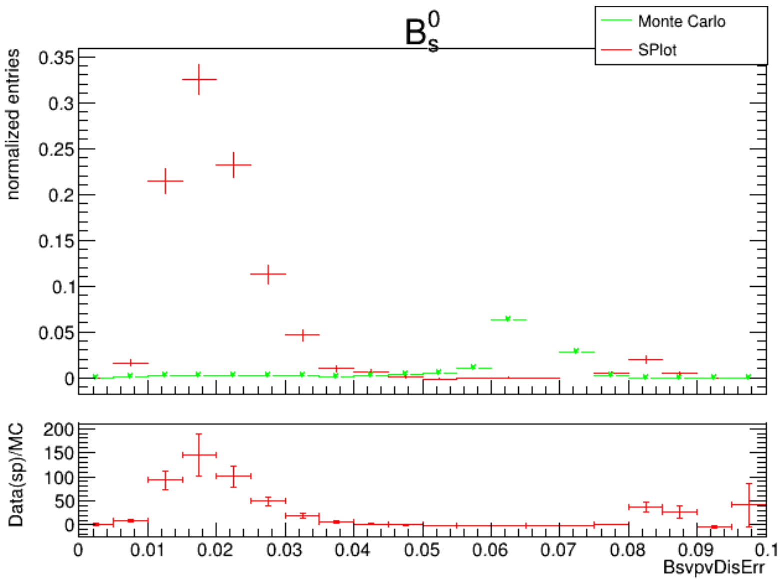

B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Bmu1eta



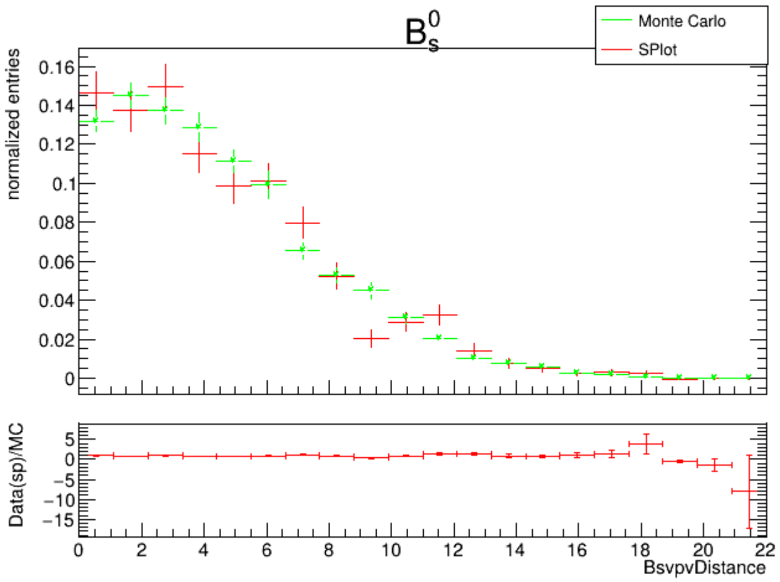
B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Bmu1pt
Bmu2dxyPV
Bmu2dzPV
Bmu2eta
Bmu2pt
Bmumueta



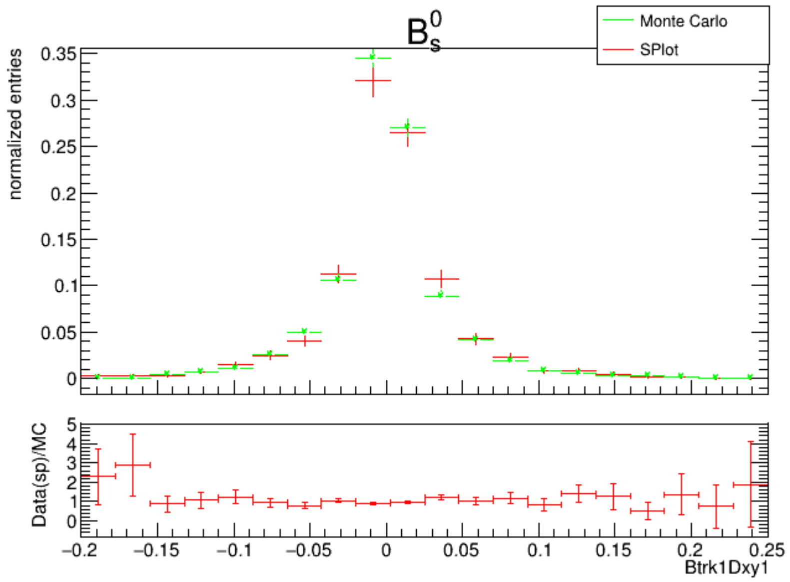
B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Bmumumass
Bmumuphi
Bmumupt
Bpt
Bsize
BsvpvDis_Err



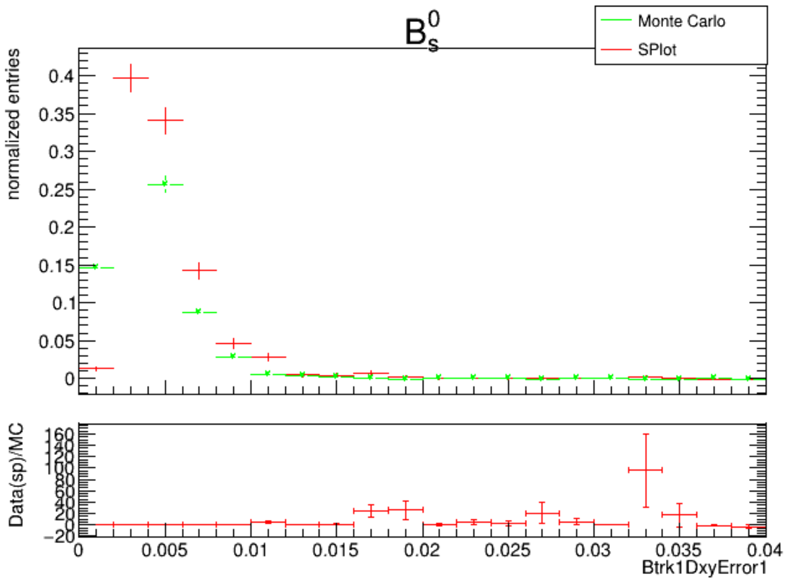
B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
BsvpvDisErr



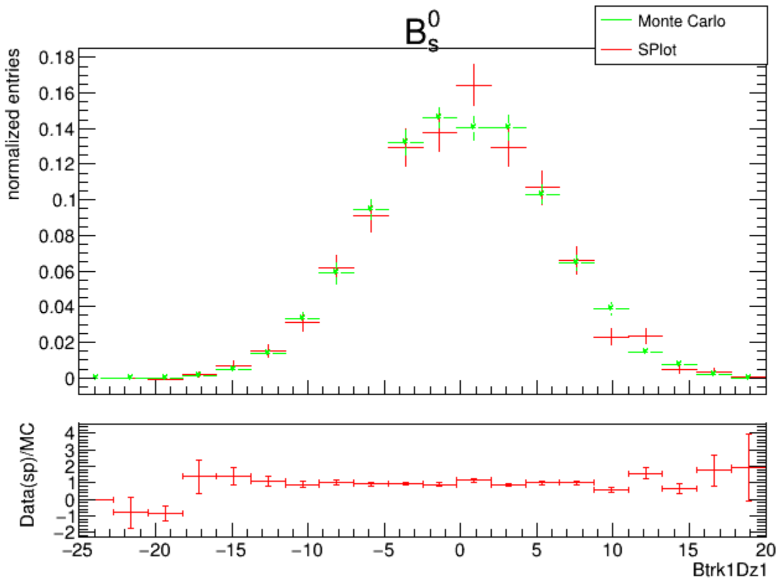
B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
BsvpvDistance



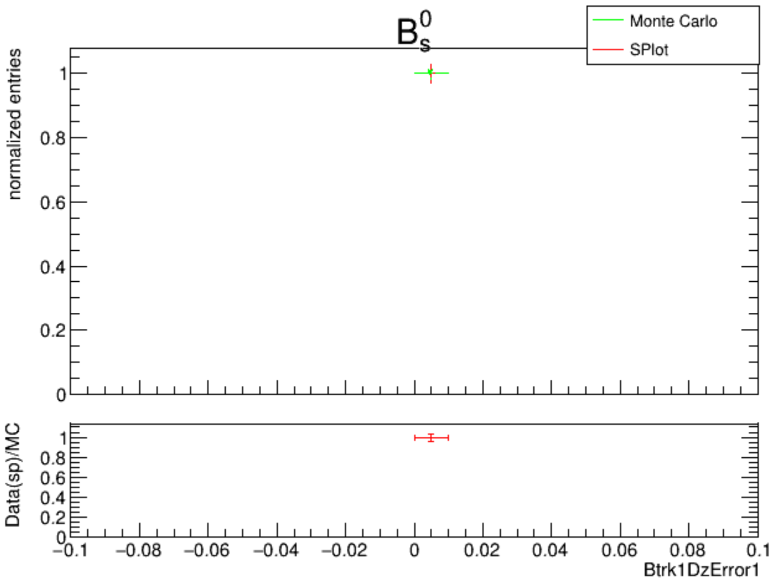
B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Btrk1Dxy1



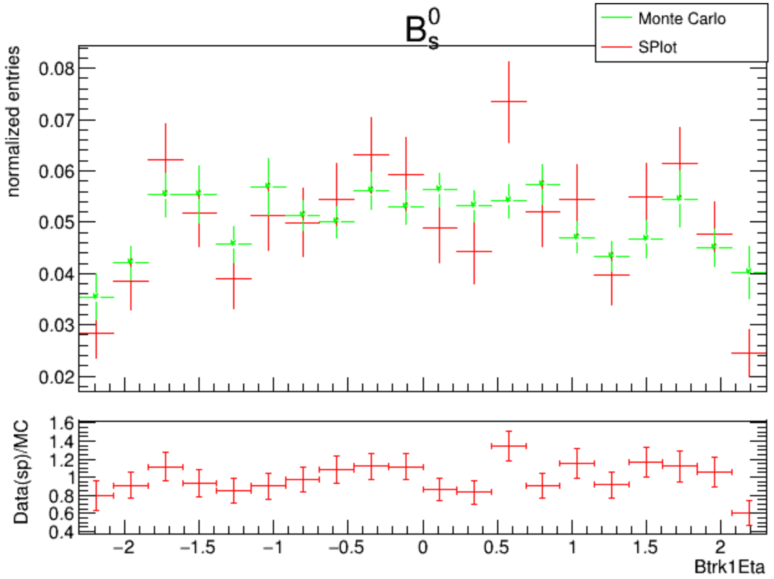
B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Btrk1DxyError1



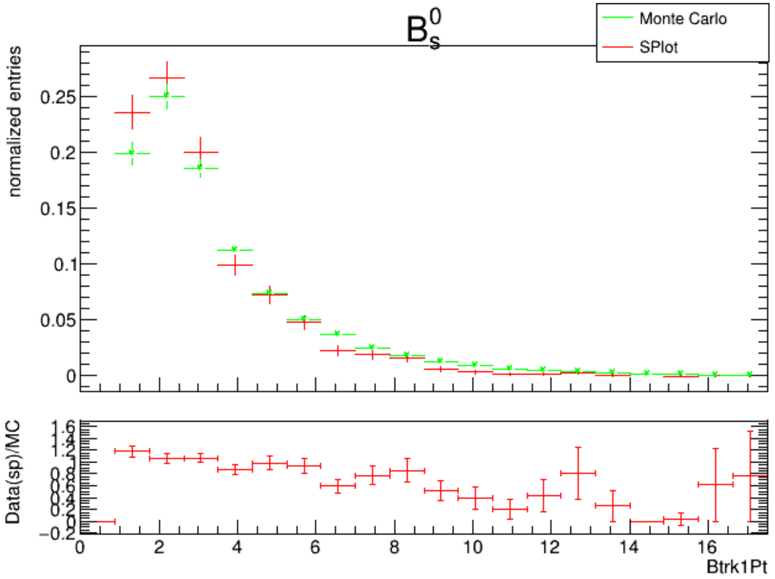
B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Btrk1Dz1



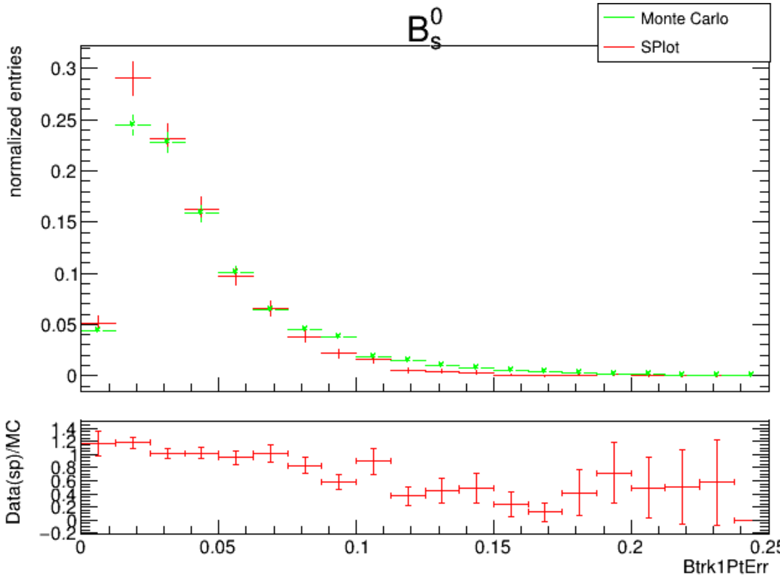
B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Btrk1DzError1



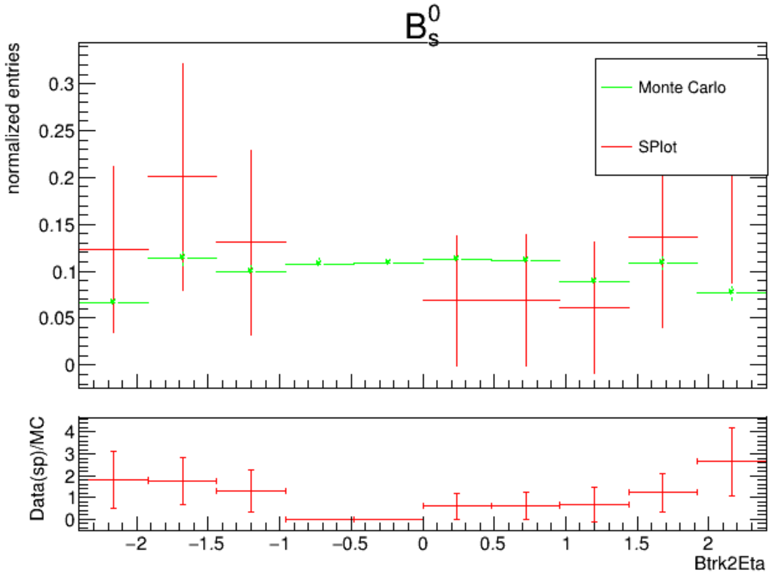
B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Btrk1Eta



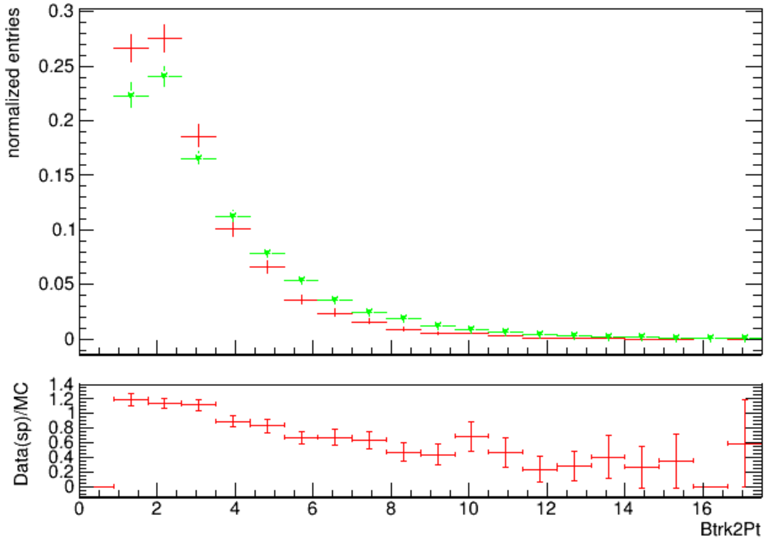
B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Btrk1Pt



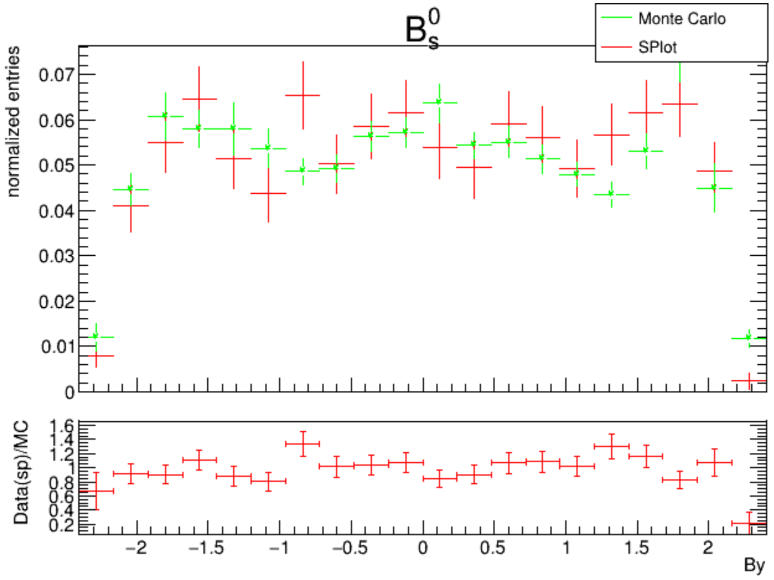
B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Btrk1PtErr



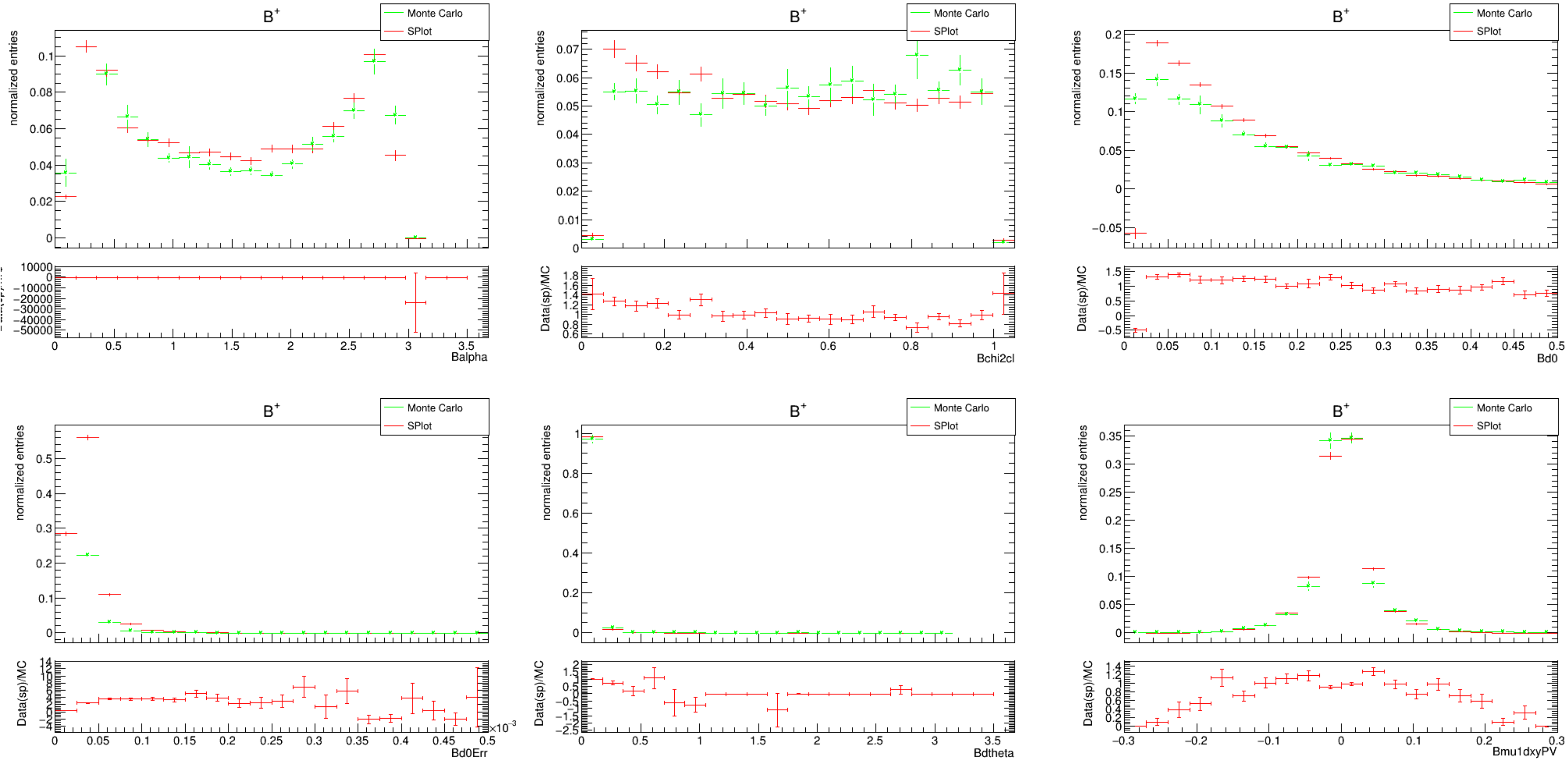
B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Btrk2Eta

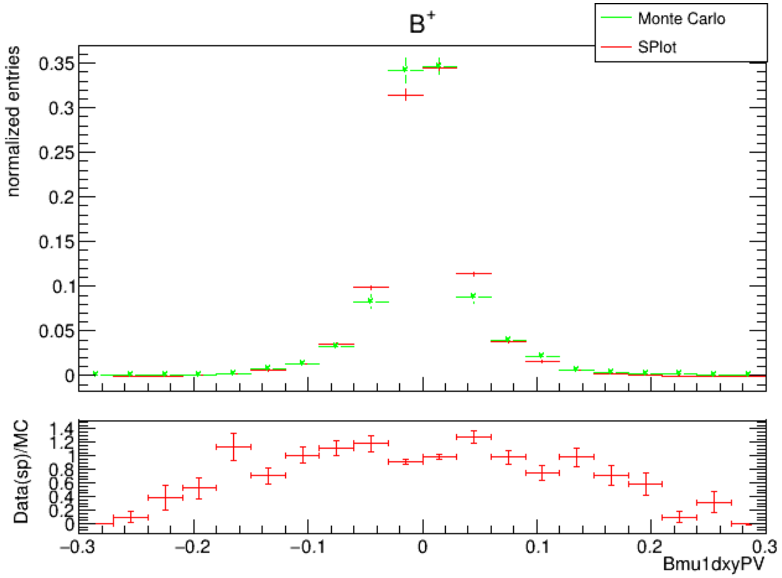


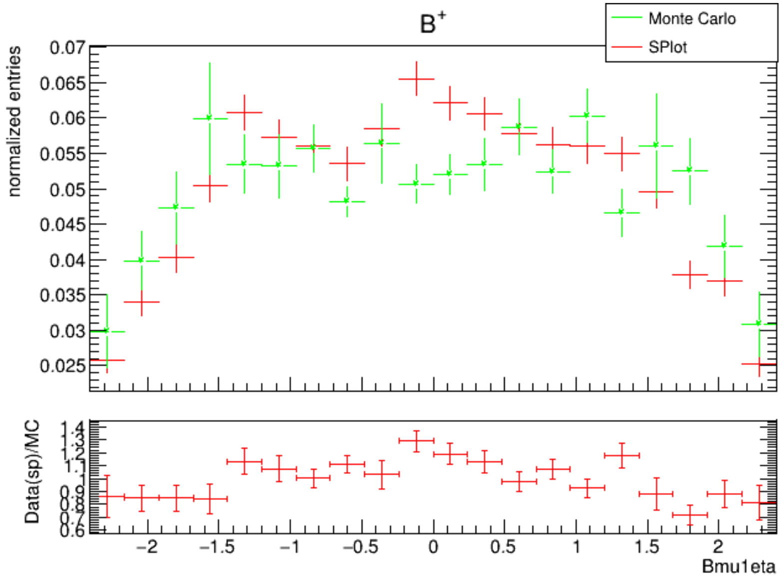
normalized entries
Data(sp)/MC
Btrk2Pt

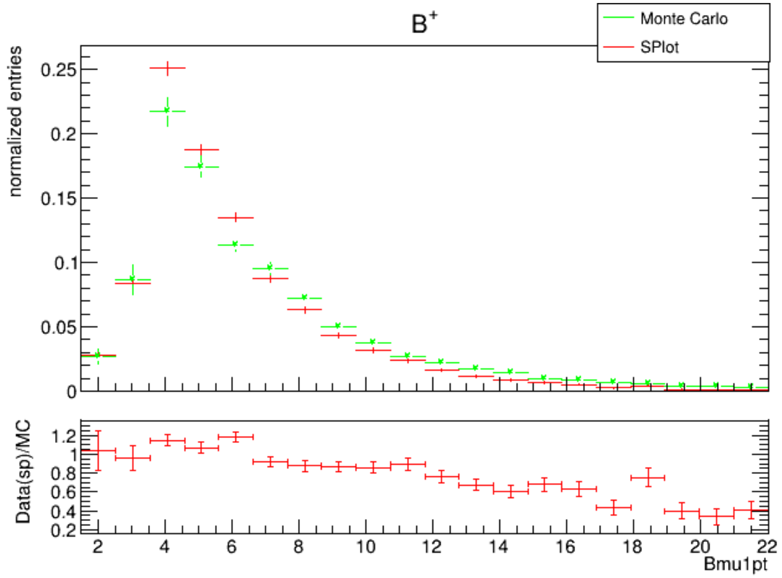


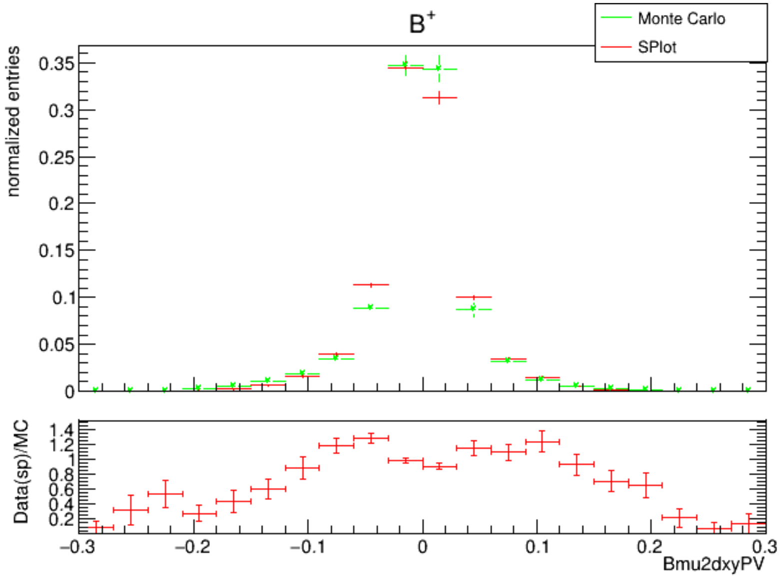
B_s^0
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
By

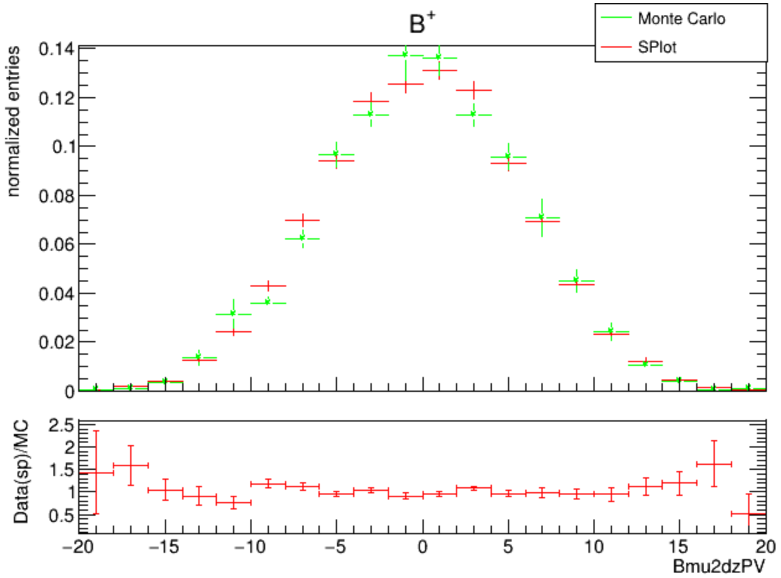

B+
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Balpha
Bchi2cl
Bd0
Bd0Err
Bdtheta
Bmu1dxyPV

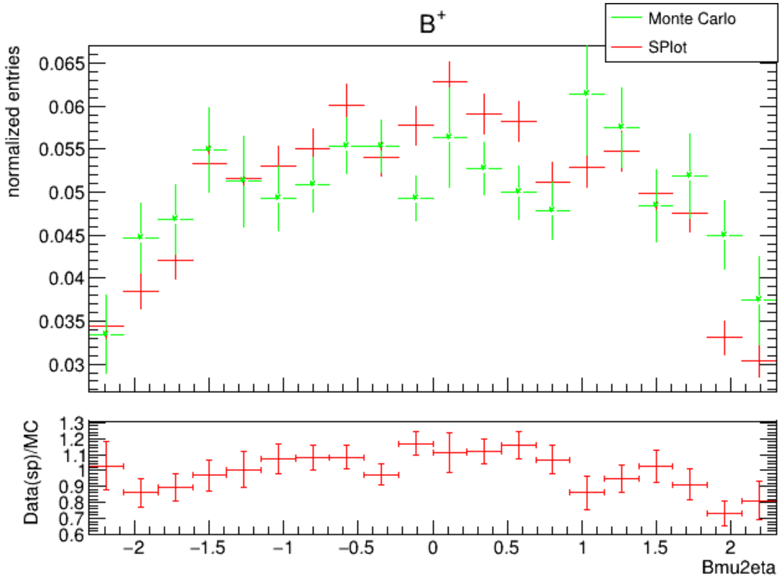

B+
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Bmu1dxyPV

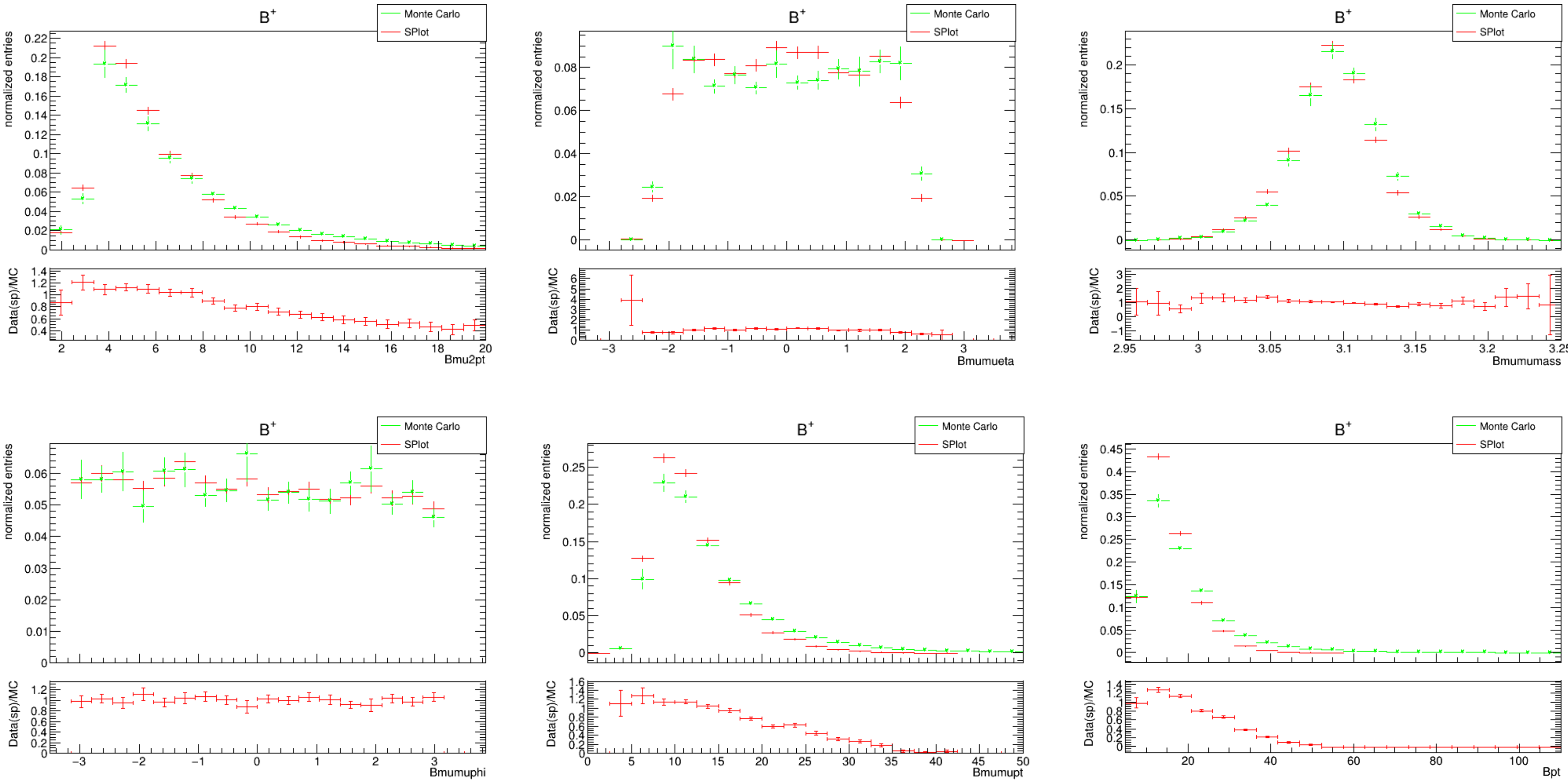

B+
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Bmu1eta

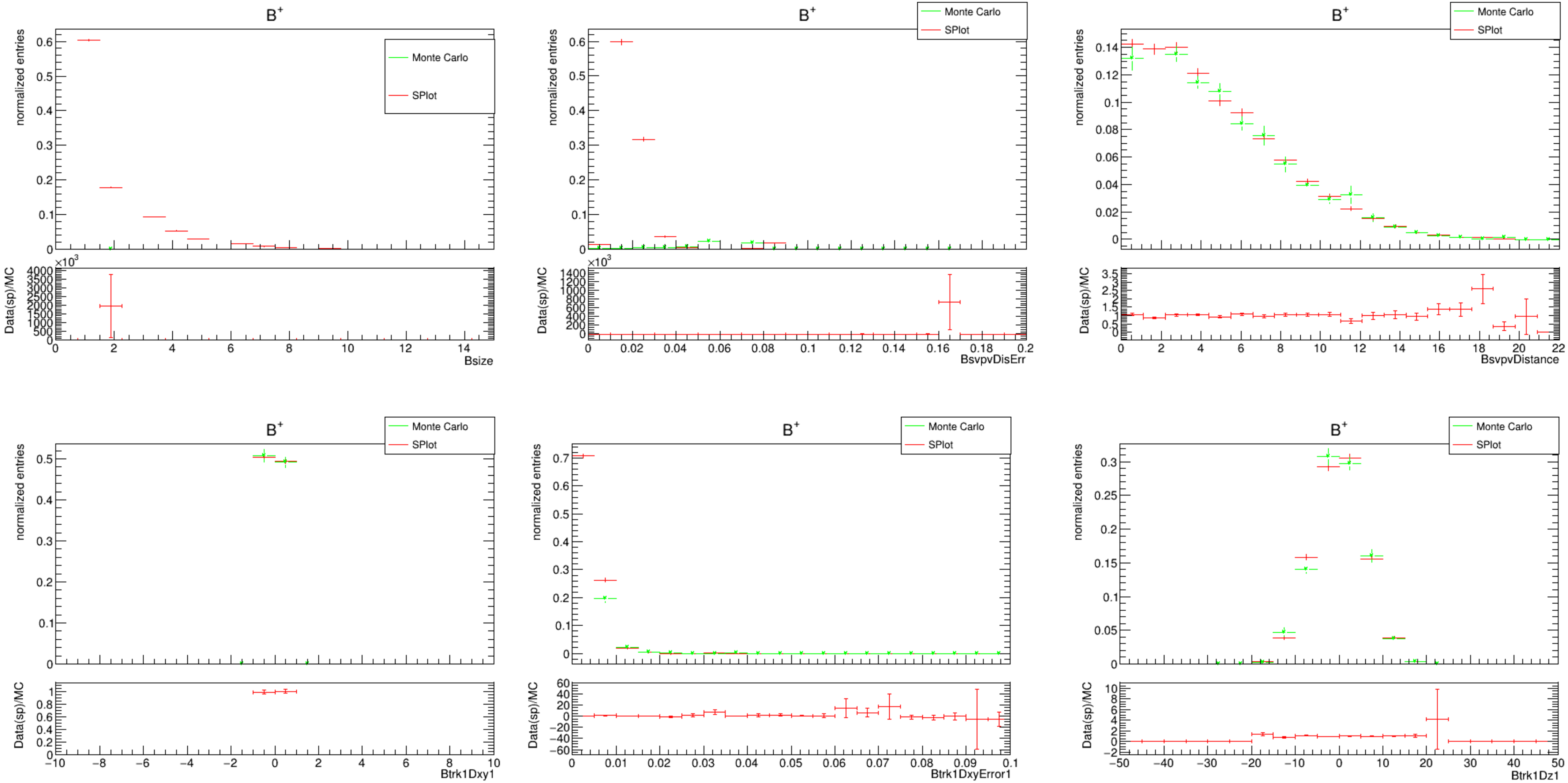

B+
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Bmu1pt


B+
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Bmu2dxyPV

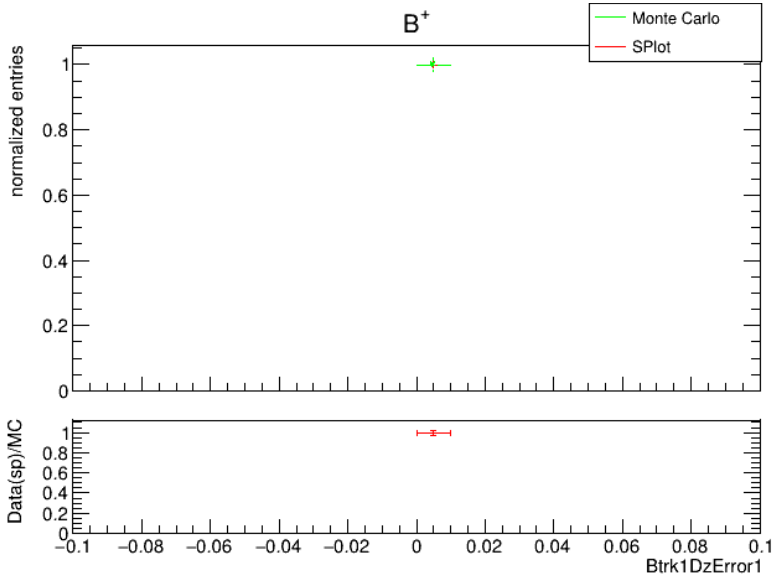

B+
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Bmu2dzPV

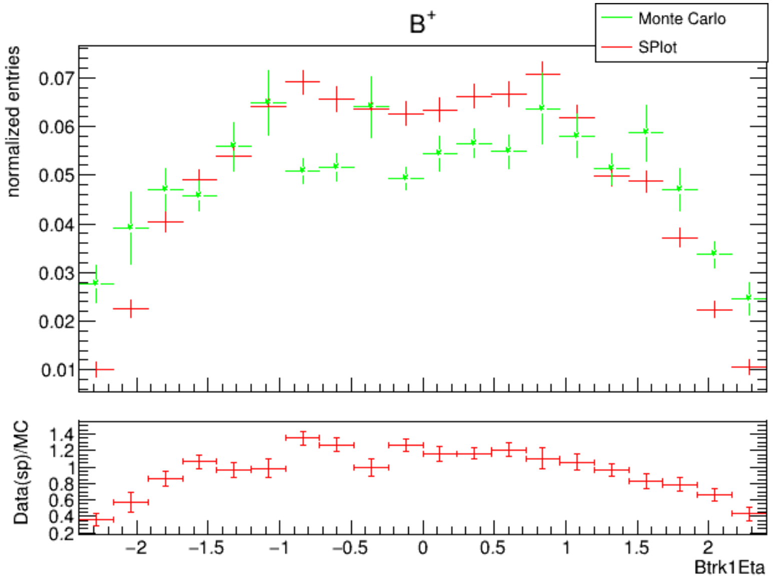

B+
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Bmu2eta

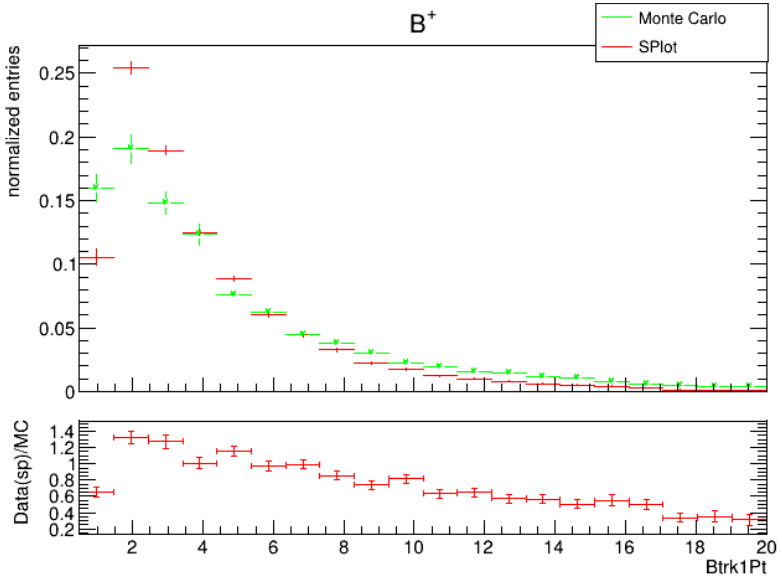

B+
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Bmu2pt
Bmumueta
Bmumumass
Bmumuphi
Bmumupt
Bpt

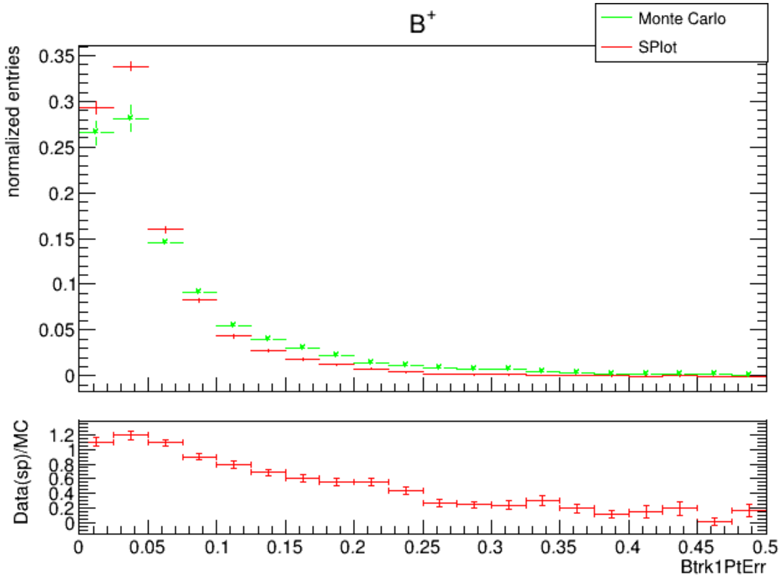


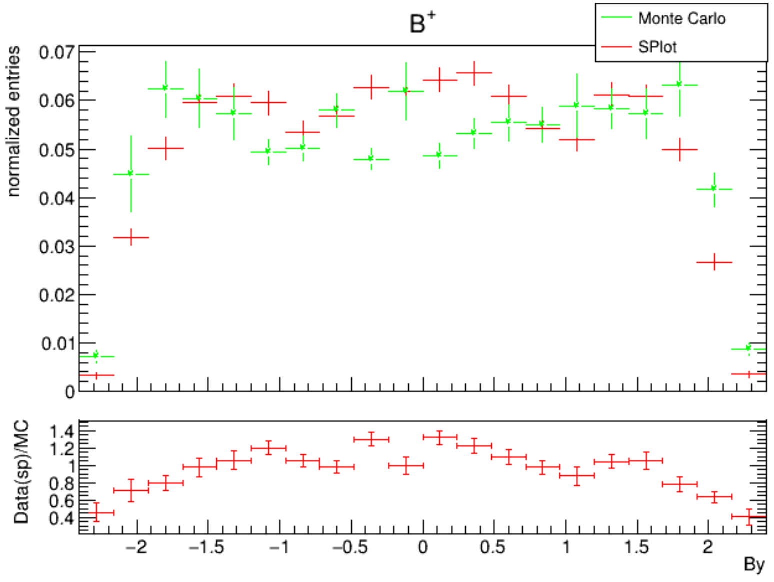
B+
normalized entries
Data(sp)/MC
Monte Carlo
SPlot
Bsize
BsvpvDisErr
BsvpvDistance
Btrk1Dxy1
Btrk1DxyError1
Btrk1Dz1


B+
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Btrk1DzError1


B+
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Btrk1Eta


B+
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Btrk1Pt


B+
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
Btrk1PtErr


B+
Monte Carlo
SPlot
normalized entries
Data(sp)/MC
By