$B^+ \to K^+\tau^\pm\ell^\mp$ analysis with semileptonic tagging in Belle & Belle II

This study is based on the Belle MC/Data.

Junaid Ur Rehman
12/11/2024

The Henryk Niewodniczański
Institute of Nuclear Physics
Polish Academy of Sciences

BELLE

Belle II

Outline

- Motivation
- Analysis approach
- Basic selections and BDT analysis
- Control modes analysis
  - $B^+ \rightarrow J/\psi\,(\mu^+\mu^-)\,K^+$
  - $B^+ \rightarrow \overline{D}^0\,(K^+\pi^-)\,\pi^+$
- Summary and outlook

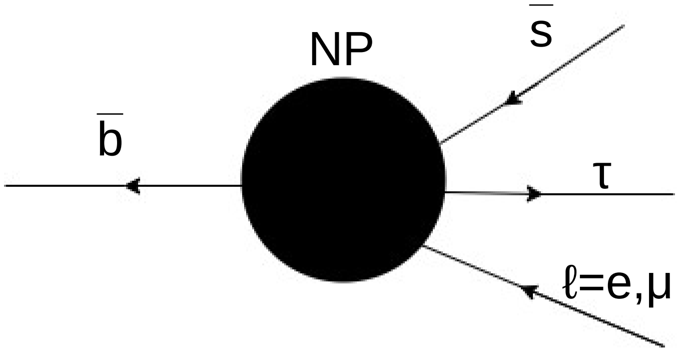
Motivation

- In the standard model (SM), $B^+ \to K^+ \tau^\pm \ell^\mp$ decays are not allowed.

- Third generational coupling and τ presence are sensitive to new physics (NP). Some NP models predict BR~10^{-7}.

- Current upper limits are $(0.59–2.45)\times10^{-5}$, measured by using hadronic tagging and mainly leptonic τ decays [1].

- We are trying to add additional statistics by including the semileptonic τ decays and using the semileptonic B_{tag}.



$$\mathcal{B}(B^+ \to K^+\tau^+\mu^-) < 0.59 \times 10^{-5}$$

$$\mathcal{B}(B^+ \to K^+\tau^+e^-) < 1.51 \times 10^{-5}$$

$$\mathcal{B}(B^+ \to K^+\tau^-\mu^+) < 2.45 \times 10^{-5}$$

$$\mathcal{B}(B^+ \to K^+\tau^-e^+) < 1.53 \times 10^{-5}$$

1. S. Watanuki et al.(Belle Collaboration),Search for the Lepton Flavor Violating Decays $B^+ \to K^+\tau^\pm\ell^\mp$ (ℓ = e,μ) at Belle PhysRevLett.130.261802

Analysis approach

- We are using the basic kinematic constraints of the experiment to reconstruct our decay.

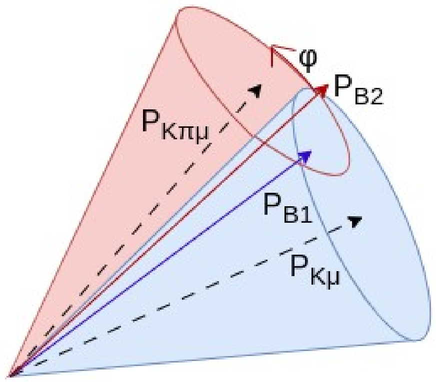
- Our complete decay has the following form

  $e^+e^- \rightarrow Y(4S) \rightarrow B^+ B^-$

  $B^+ \rightarrow K^+ \tau^- \mu^+$ (**B_{sig}**)

  $\tau^- \rightarrow \pi^- \nu_\tau$

  $B^- \rightarrow X\ell^-\nu_\ell$ (**B_{tag}**), $\ell = e,\mu$



- First we assume τ is missing and the missing momentum is constrained on a cone around $p_{k\mu}$ and after using $\tau^- \rightarrow \pi^-\nu_\tau$, we constrain the missing momentum around $P_{k\mu\pi}$ cone.

- Intersection of these two cones, provides us the B_{sig} momentum with two folds ambiguity (p_{B1},p_{B2}) and a discriminator variable sinφ.

Hadronic tagging

- To further reduce the background, we use information from tag side.

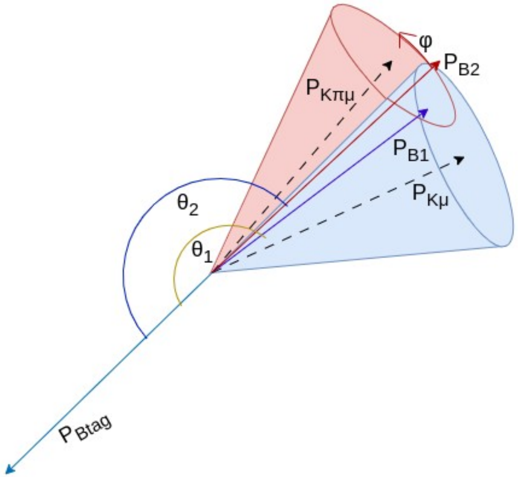
- In one of the preliminary studies [2], we used hadronic B-decays to reconstruct B_{tag}.

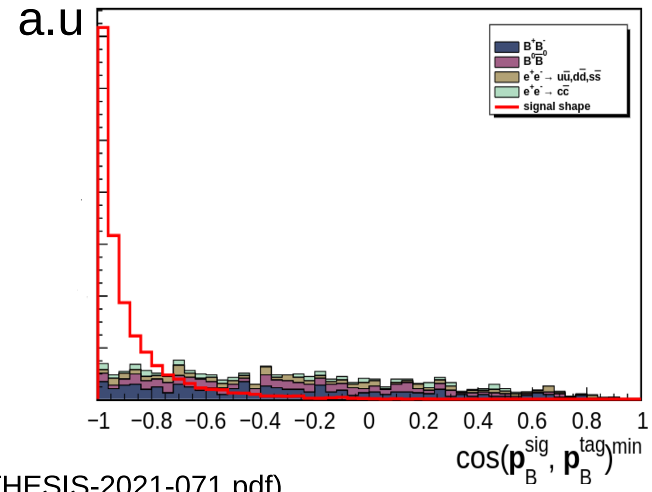
$$B^+ \rightarrow K^+ \tau^-(\rightarrow \pi^- \nu_\tau)\mu^+ \quad (\mathbf{B_{sig}})$$

$$B^- \rightarrow \text{hadrons} \quad (\mathbf{B_{tag}})$$

- We used the minimum of the cosine angles [min.($\cos\theta_1$, $\cos\theta_2$)], to suppress the background further.

- We were able to get B_{sig} reconstruction efficiency of 2.2×10^{-3} and corresponding UL~10^{-5}, which is quite optimistic to further work on this approach.





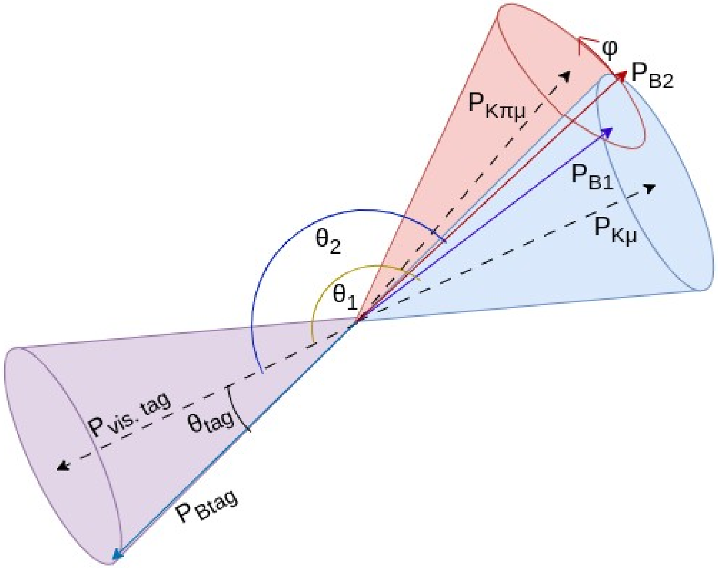
2. M. Kaleta, M.S thesis, Belle Collaboration. (https://docs.belle2.org/record/2549/files/BELLE2-MTHESIS-2021-071.pdf)

Semileptonic tagging

- We are using inclusive semileptonic (SL) tagging to reconstruct B_{tag} for further suppressing the background.

- We can constrain the missing momentum on B_{tag} side on a cone around $p_{vis.tag}$.

- We have the two sum of cosine angles, from which we pick the best one by using the following condition

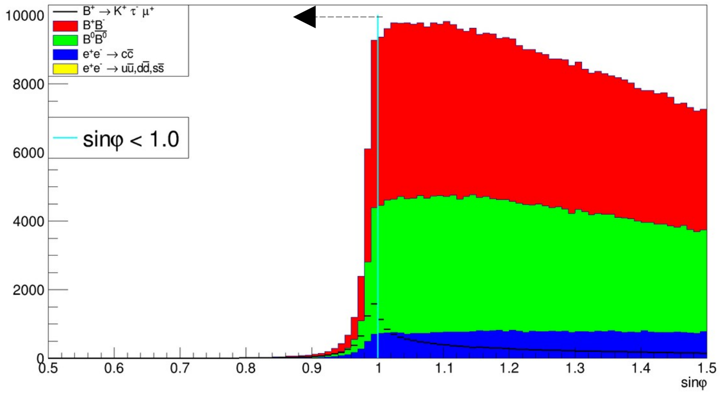


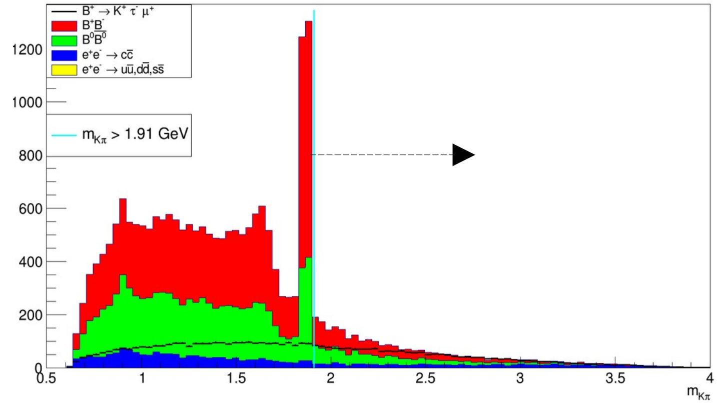
$$\Delta\cos\theta = \min|\cos\theta_{1,2} + \cos\theta_{tag}|$$

Signal side veto selections

- We are using the following veto selections on the signal side.

$\sin\varphi < 1$
$m_{K\pi} > 1.91$ GeV
$\lvert m_{\mu\pi}-3.1\rvert > 0.05$ & $\lvert m_{\mu\pi}-3.69\rvert > 0.05$ GeV
$\lvert m_{\mu\mu}-3.1\rvert > 0.05$ & $\lvert m_{\mu\mu}-3.69\rvert > 0.05$ GeV





Sample size

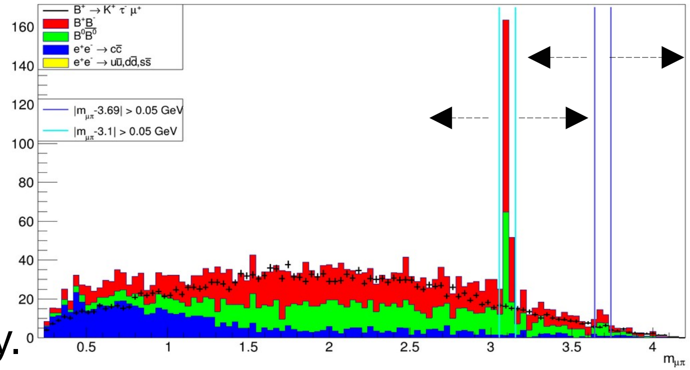
5.0 M **dedicated signal MC**

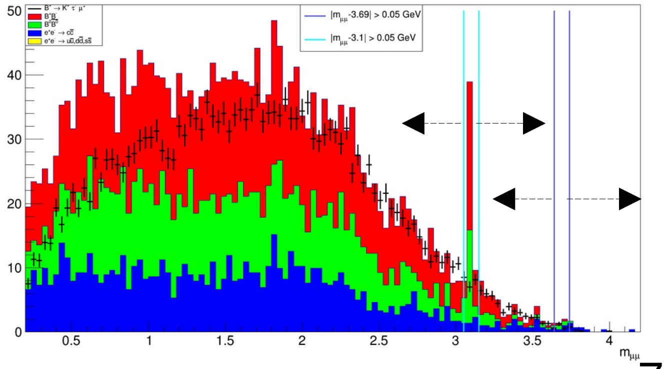
$B^+ \to K^+ \tau^- \mu^+$

$B^- \to$ generic

03 streams of generic MC

normalized to Belle luminosity.

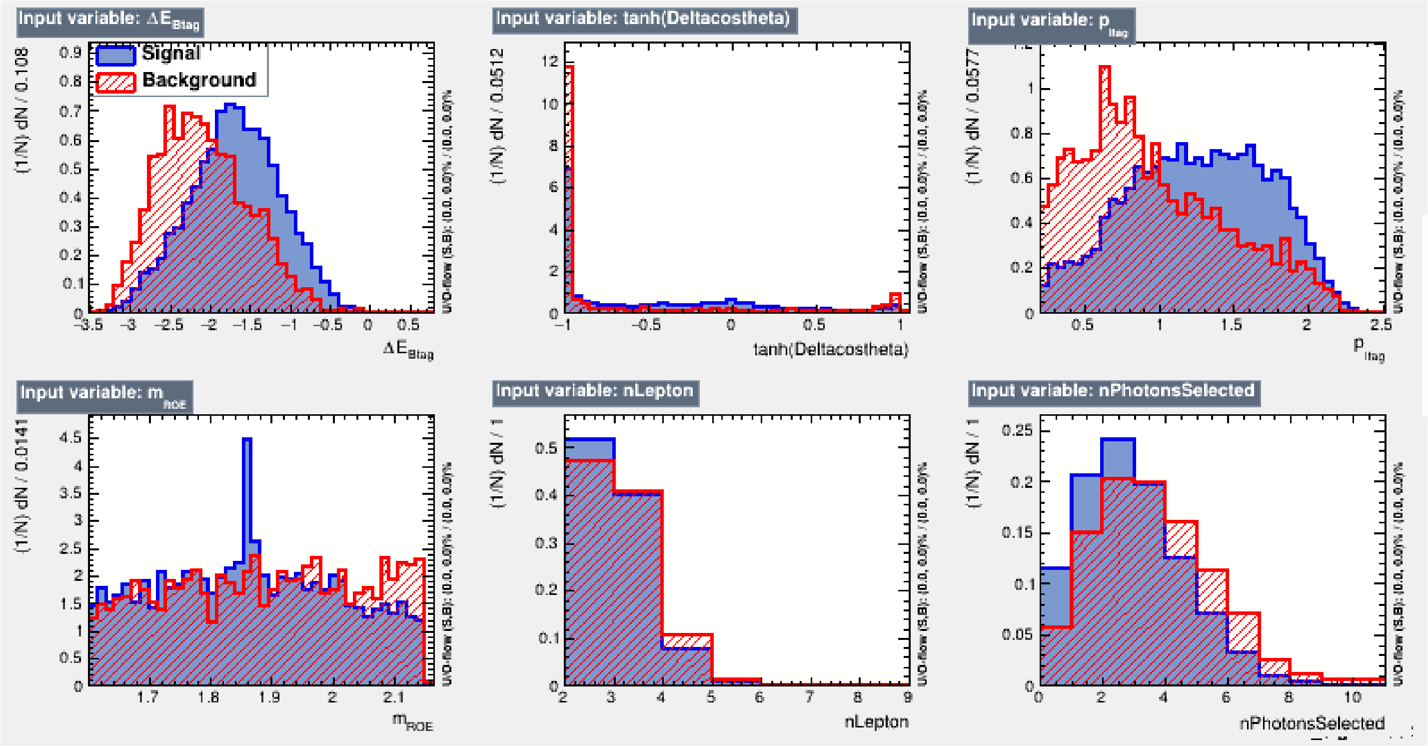




BDT training

- After applying the veto cuts, we are using six input variables in the BDT training.

Input variables = {ΔE_{Btag}, $\Delta\cos\theta$, p_{ltag} , m_{ROE}, nLepton, nPhotons}

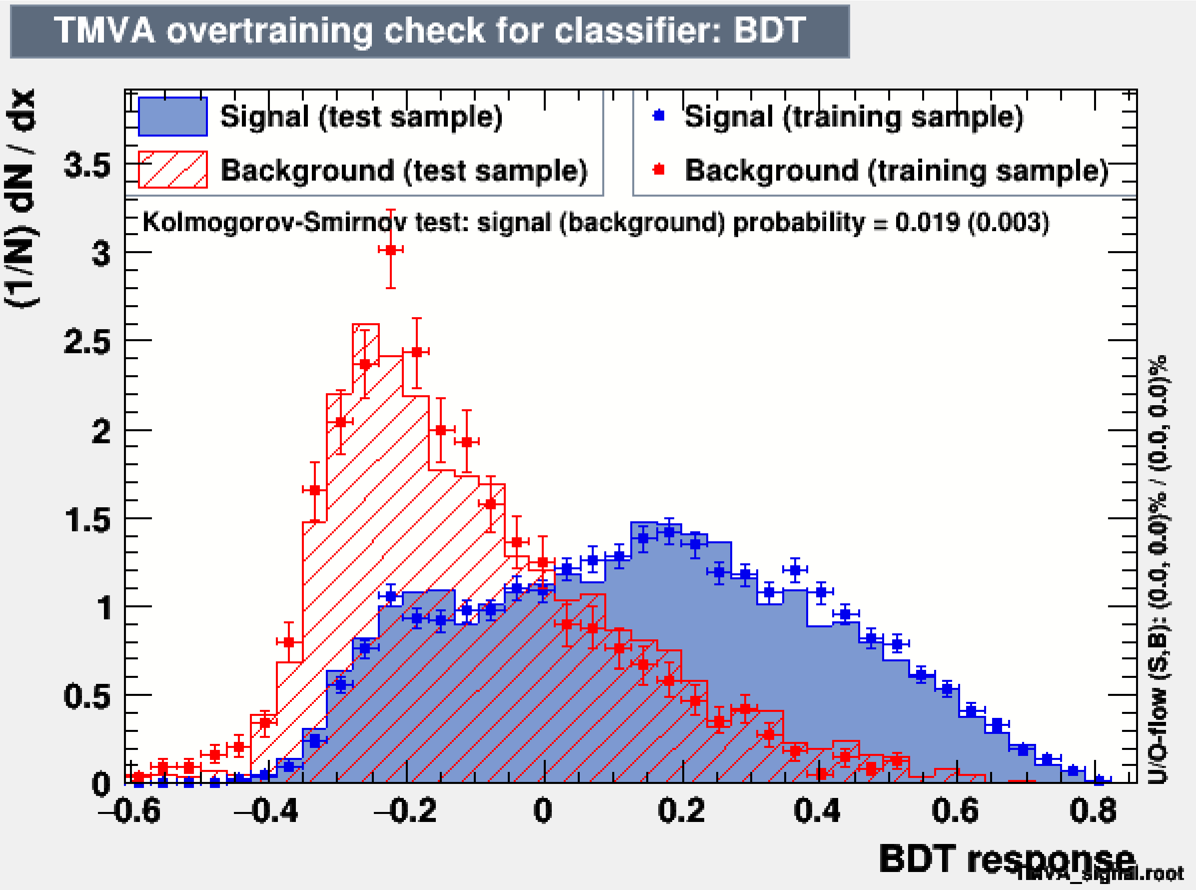


Variable importance

Ranking result (top variable is best ranked)

Rank	Variable	Variable Importance
1	#DeltaE_{Btag}	2.682e-01
2	tanh(Deltacostheta)	2.028e-01
3	m_{ROE}	1.670e-01
4	p_{ltag}	1.545e-01
5	nPhotonsSelected	1.303e-01
6	nLepton	7.719e-02

BDT response



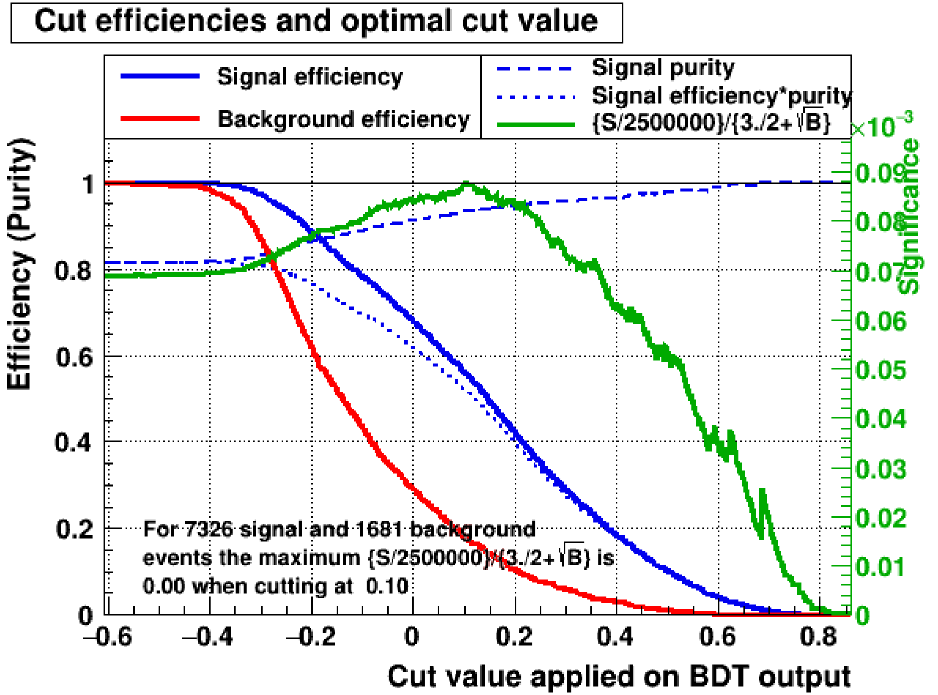
Hyper parameters of the BDT are yet to be optimized.

Punzi figure of merit

- For optimizing the BDT score, we are using the Punzi figure of merit.

$$FOM_{Punzi} = \frac{\epsilon(t)}{\frac{\alpha}{2} + \sqrt{B(t)}}$$

- $\varepsilon(t)$ = signal efficiency
- α = desired significance
- $B(t)$ = remaining background events

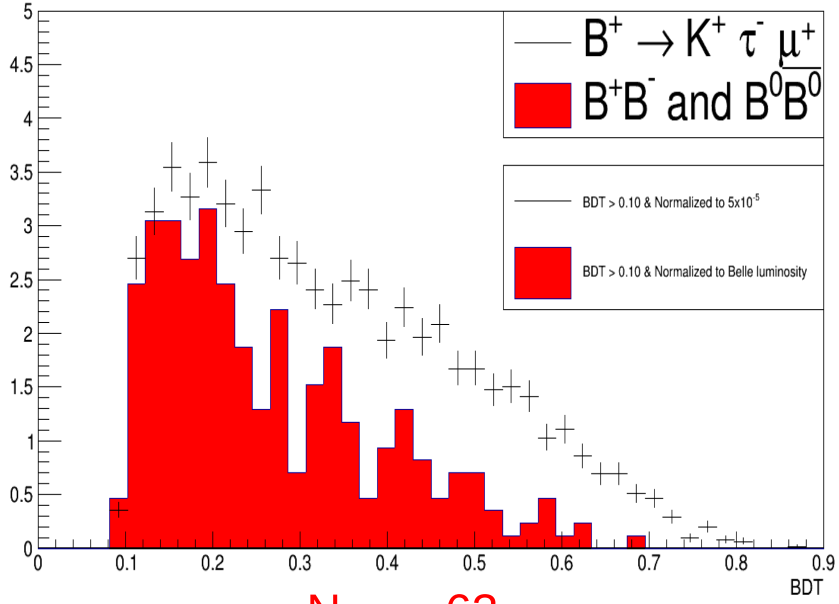


Optimal cut BDT > 0.10

Classifier	Input N_{sig}	Input N_{bg}	Optimal cut	FOM_{Punzi}	Final N_{sig}	Final N_{bg}	Signal eff.	Bg eff.
BDT	7326	1681	0.1041	8.8×10^{-5}	4070	290	0.5556	0.1725

BDT response after optimal cut

- We have normalized the signal to the BF of 5×10^{-5}.
- We are considering here only B^+B^- and $B^0\overline{B}^0$ background components as the other two are negligible.
- By using BDT, we have better S/B as compared to the simple selections.



N_{sig} = 63
N_{bg} = 53

Control mode $B^+ \to J/\psi(\mu^+\mu^-) K^+$

- To further check our results, we are using the following decay as our first control channel mode.

$B^+ \to J/\psi K^+$ (BF = 1.02×10^{-3})

$J/\psi \to \mu^+\mu^-$ (BF = 5.973 %)

- Topology of this decay is similar to our signal decay.

- We assume that one μ is missing, so that it can replicate our signal decay reconstruction.

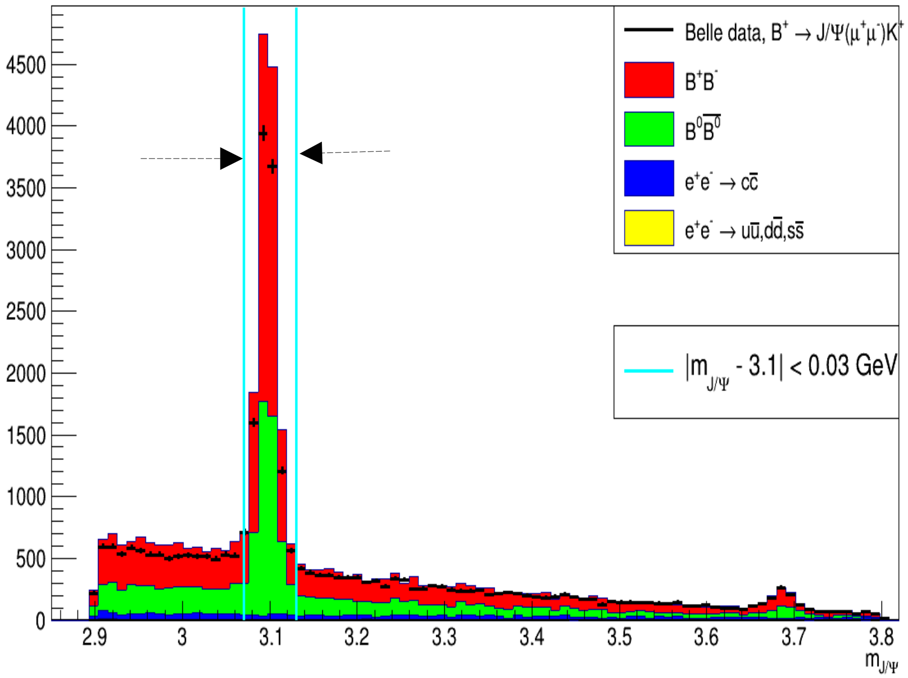
- Initial checks were performed on the dedicated MC.

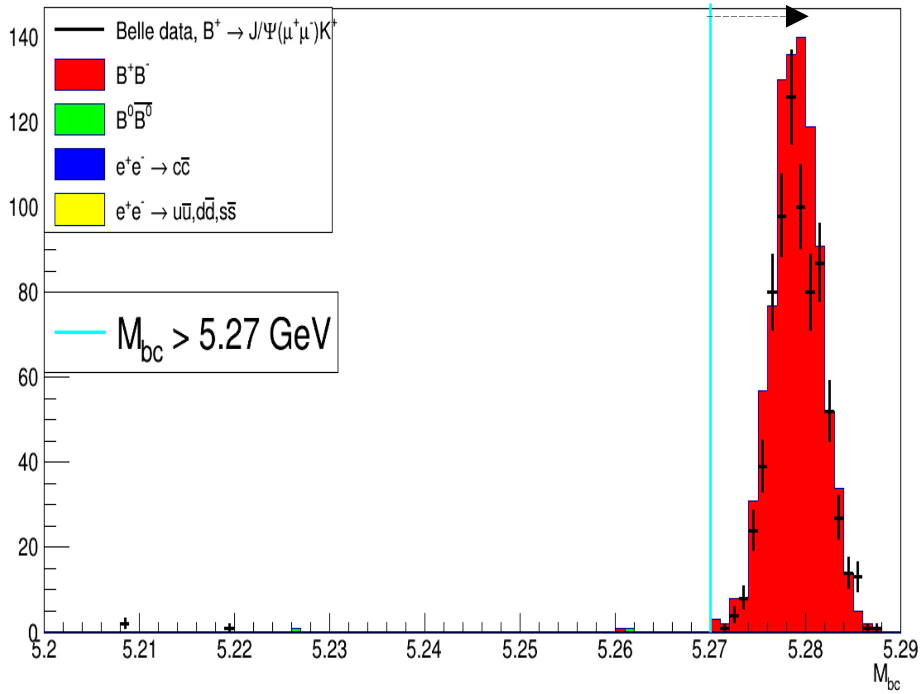
- Further checks are performed on the full Belle data set and one stream of generic MC.

Selections for $B^+ \to J/\psi(\mu^+\mu^-) K^+$

-Additional variables (M_{bc} and ΔE) to use alongside other signal variables.

-Full Belle dataset
-01 Streams of generic MC

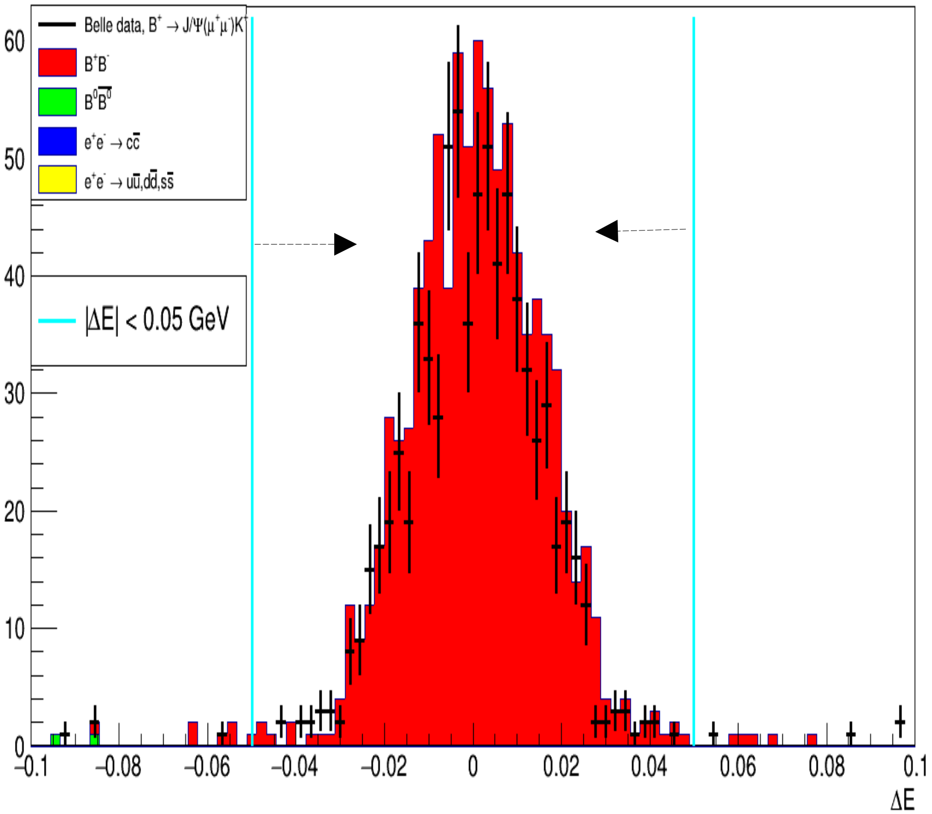




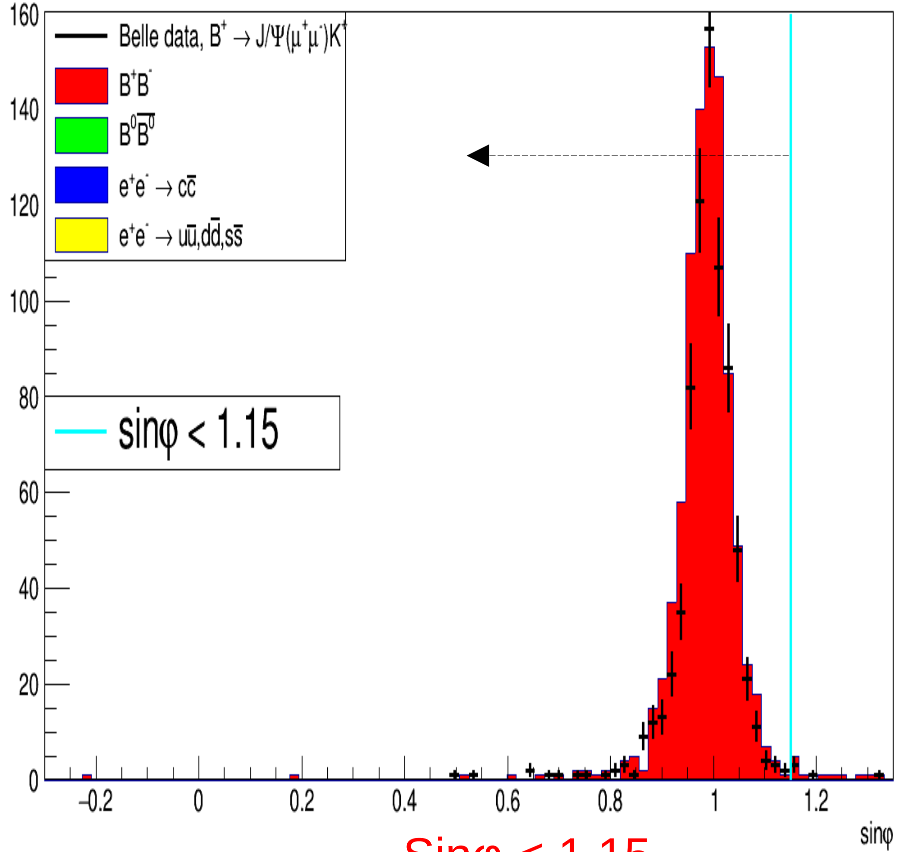
$|m_{J/\psi} - 3.1| < 0.03$ GeV

$M_{bc} > 5.27$ GeV

Selections for $B^+ \rightarrow J/\psi(\mu^+\mu^-) K^+$

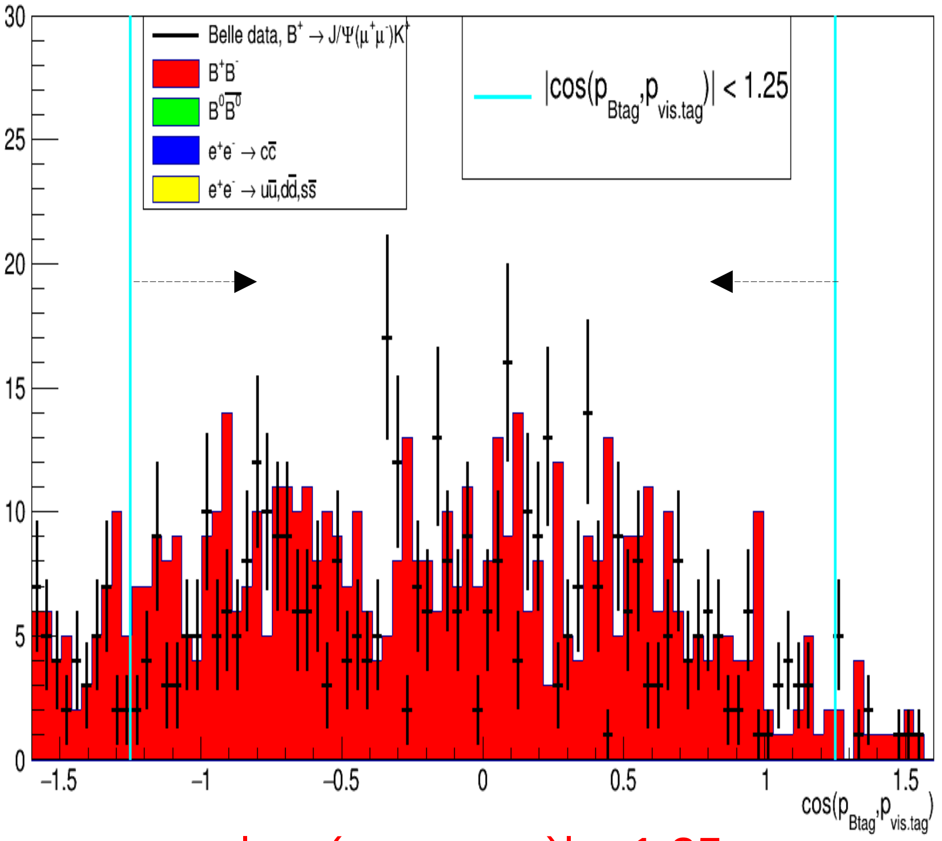


$|\Delta E| < 0.05$ GeV

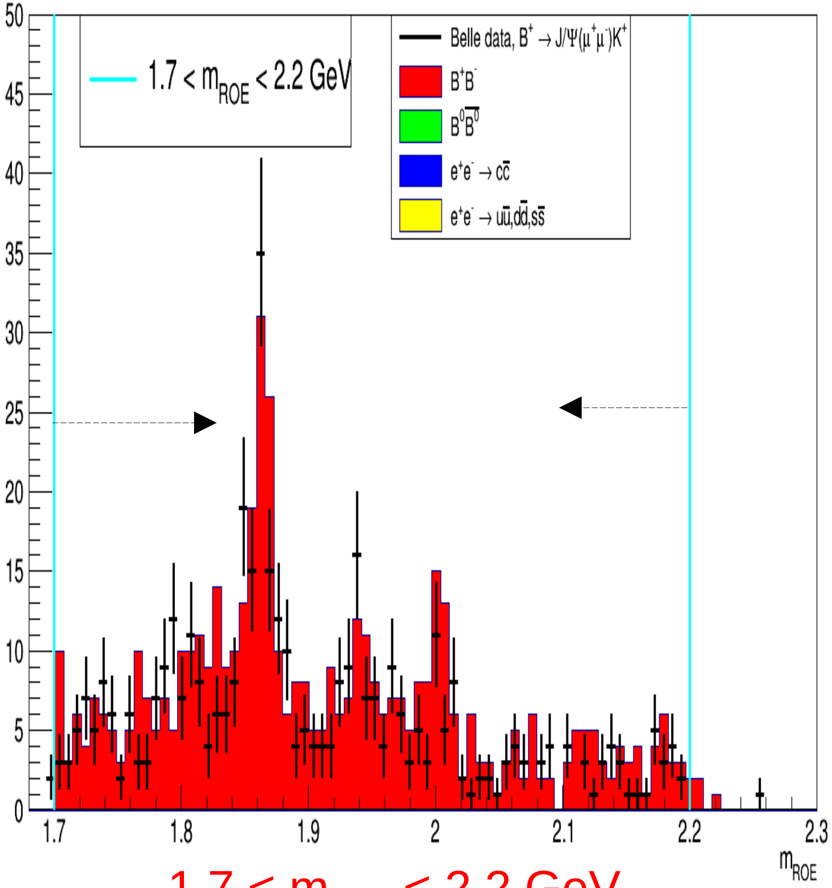


Sin$\varphi < 1.15$

Selections for $B^+ \to J/\psi(\mu^+\mu^-) K^+$

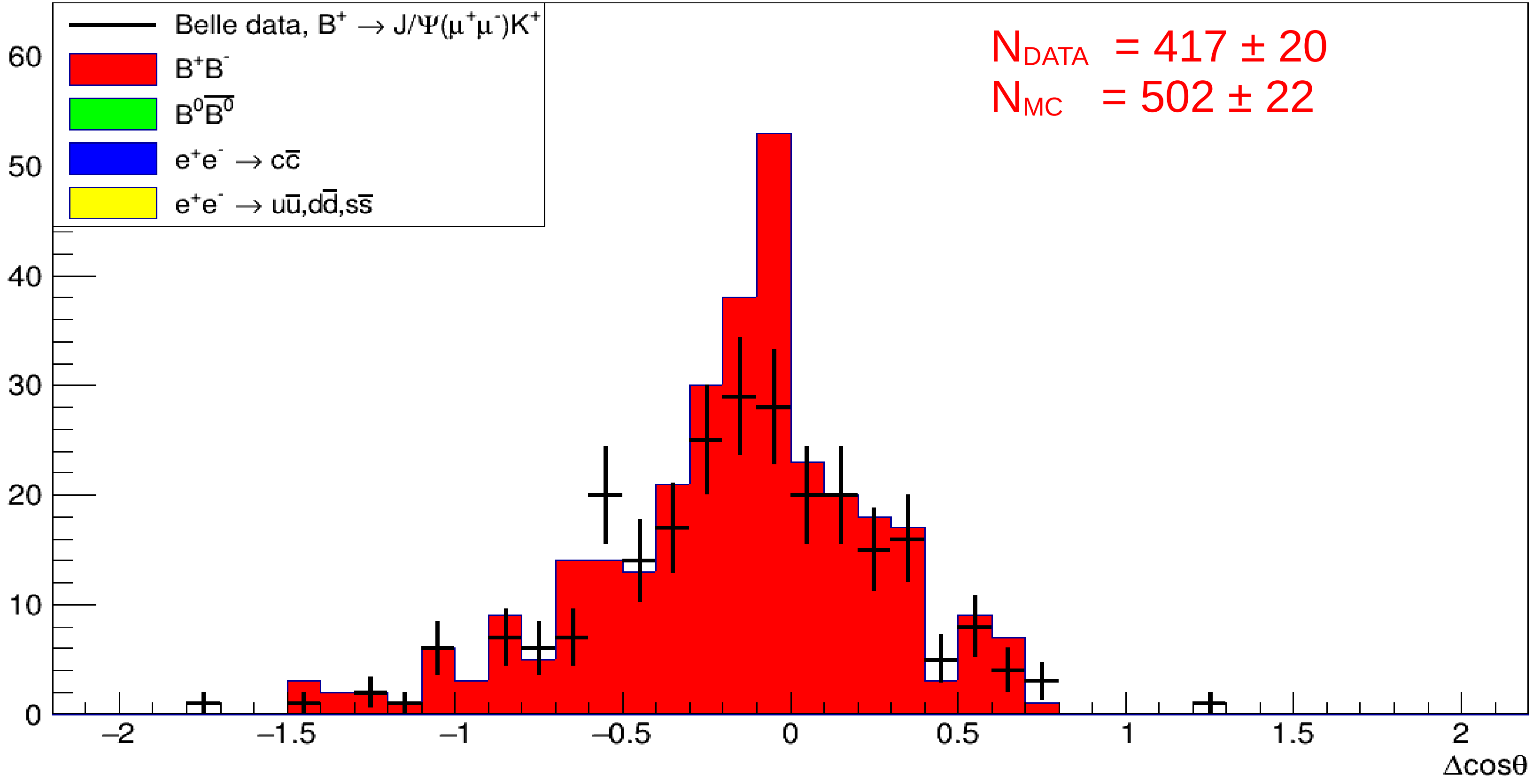


$|\cos(p_{Btag},p_{vis.tag})| < 1.25$



$1.7 < m_{ROE} < 2.2$ GeV

$\Delta\cos\theta$ for $B^+ \to J/\psi(\mu^+\mu^-) K^+$



Control mode $B^+ \rightarrow \overline{D}^0(K^+\pi^-)\pi^+$

- To further check our results, we are using the following decay as our second control channel mode.

$B^+ \rightarrow \overline{D}^0 \pi^+$ (BF = 4.61×10^{-3})

$\overline{D}^0 \rightarrow K^+\pi^-$ (BF = 3.947 %)

- Topology of this decay is also similar to our signal decay.

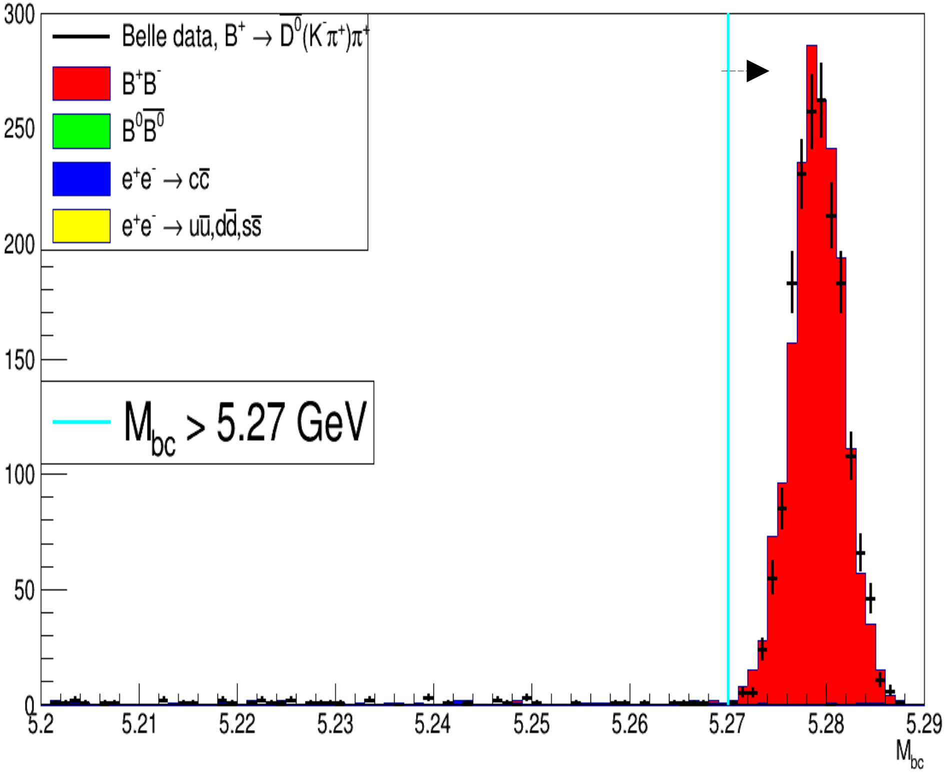
- We assume that π^- is missing, so that it can replicate our signal decay reconstruction.

- We have performed the initial checks on the dedicated MC.

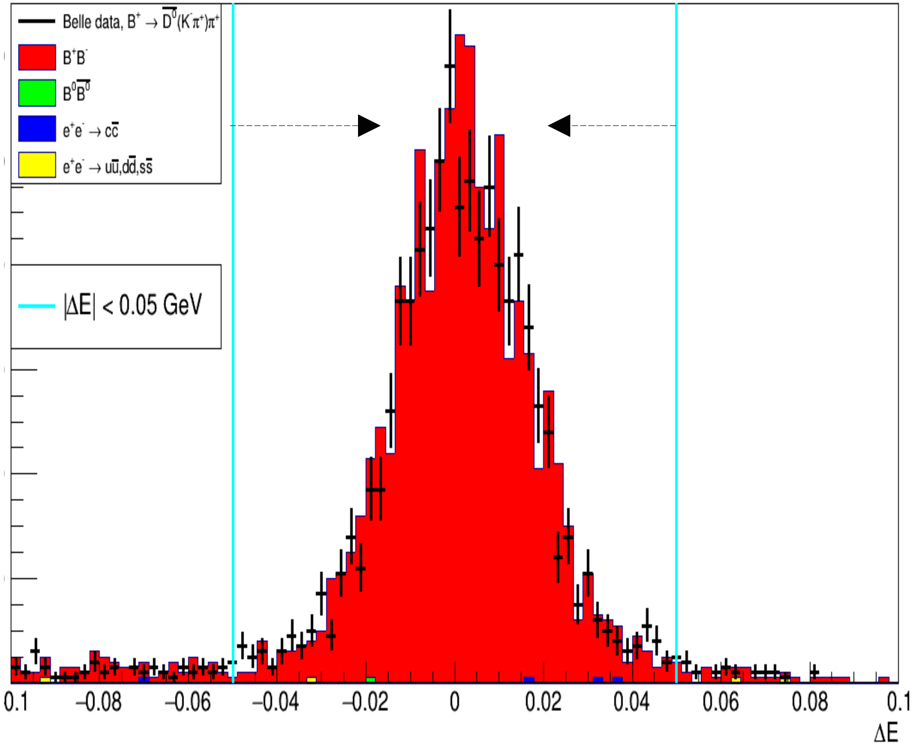
- We have also checked it on the Belle data set.

Selections for $B^+ \rightarrow \overline{D}^0(K^+\pi^-)\pi^+$

-Full Belle dataset
-01 Streams of generic MC

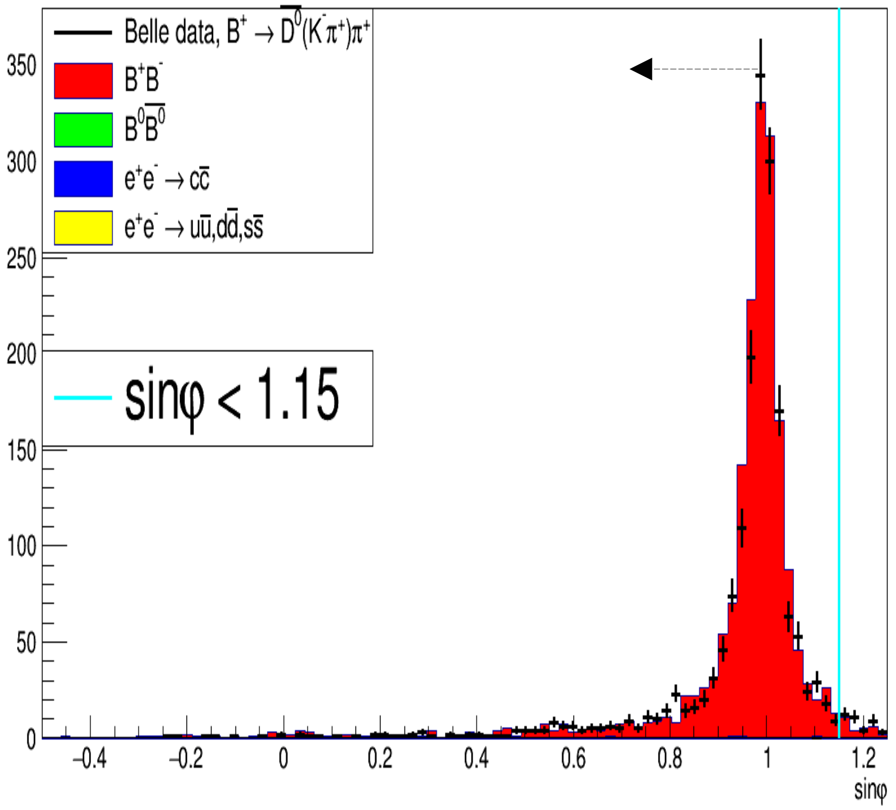


M_{bc} > 5.27 GeV

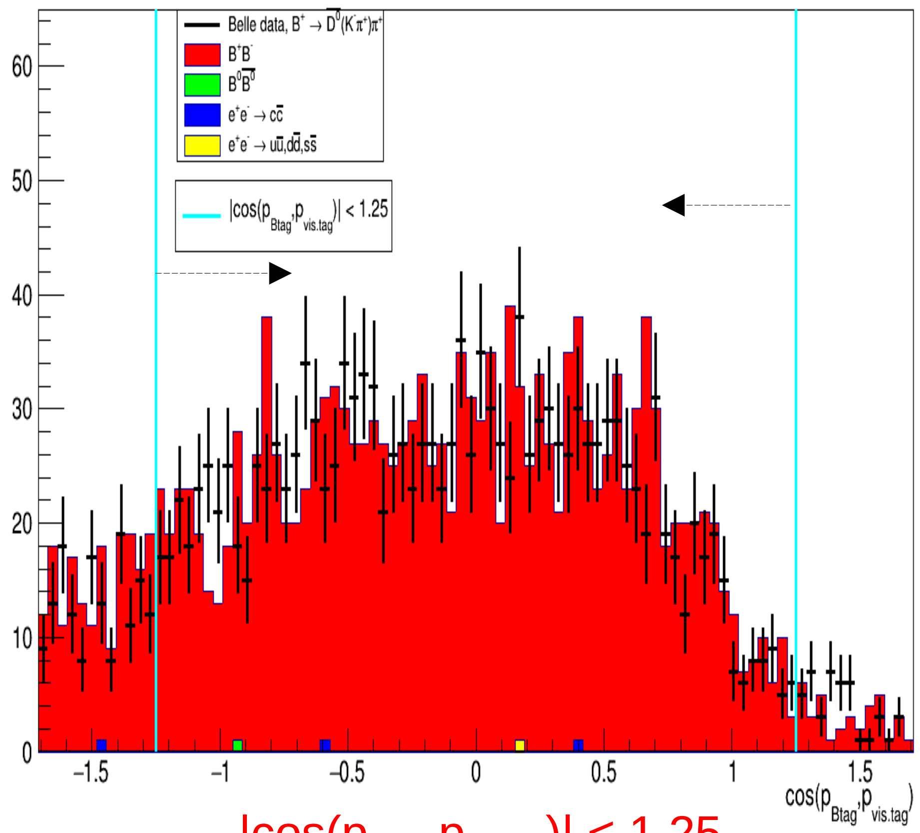


$|\Delta E|$ < 0.05 GeV

Selections for $B^+ \rightarrow \overline{D}^0(K^+\pi^-)\pi^+$

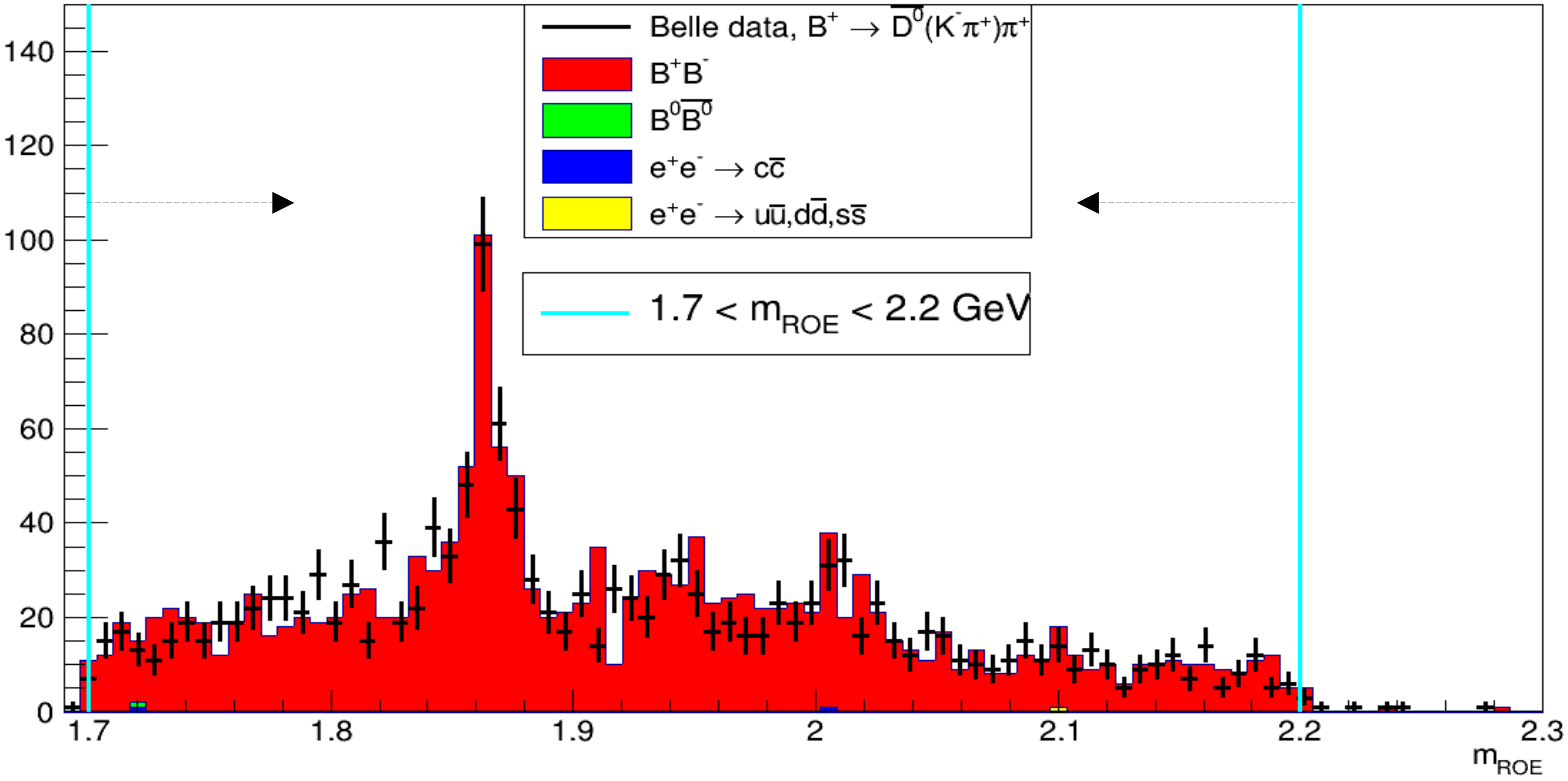


Sinφ < 1.15



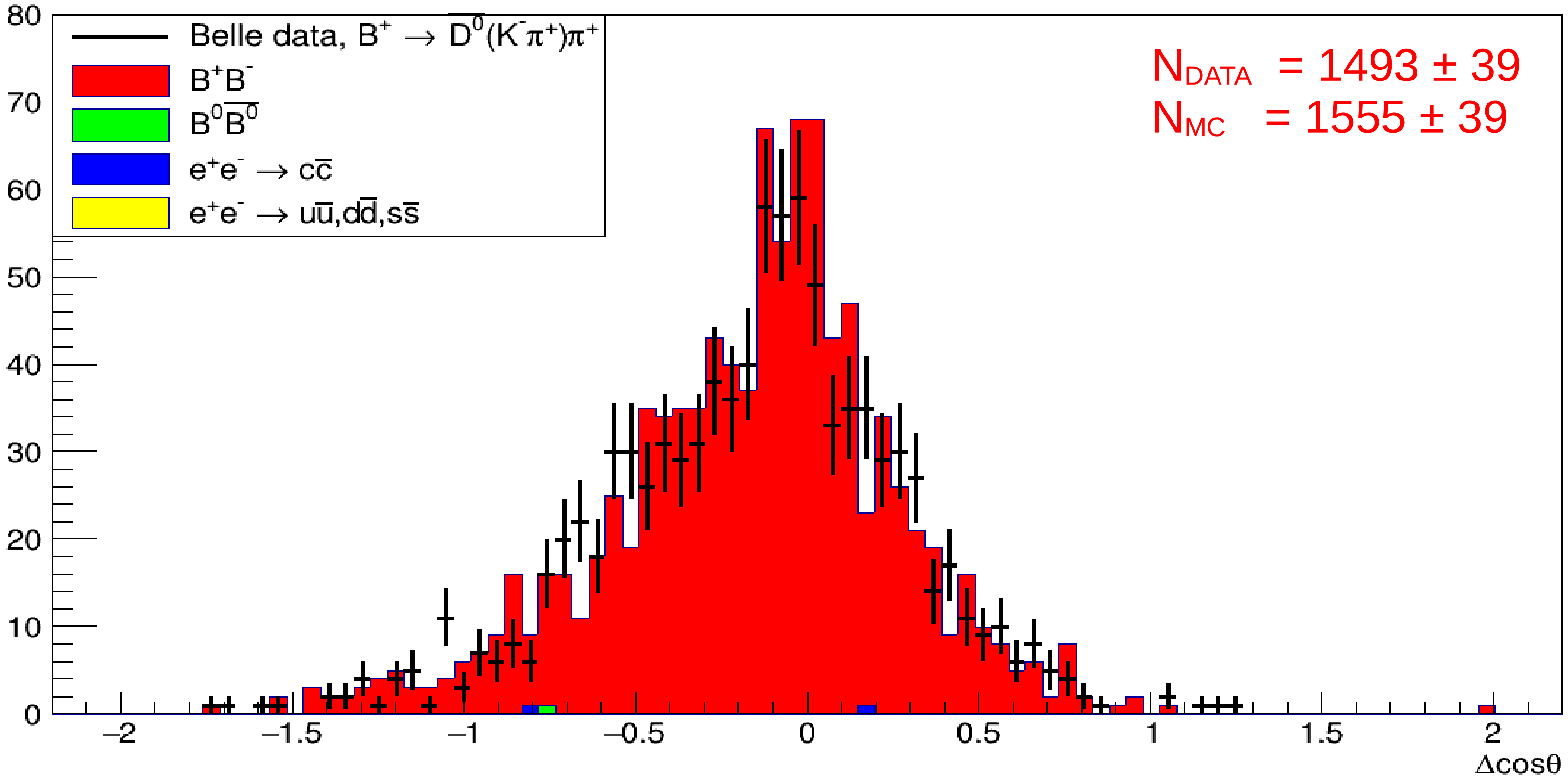
$|\cos(p_{Btag},p_{vis.tag})| < 1.25$

Selections for $B^+ \to \overline{D}^0(K^+\pi^-)\pi^+$



$1.7 < m_{ROE} < 2.2$ GeV

$\Delta\cos\theta$ for $B^+ \rightarrow \overline{D}^0(K^+\pi^-)\pi^+$



Summary and Outlook

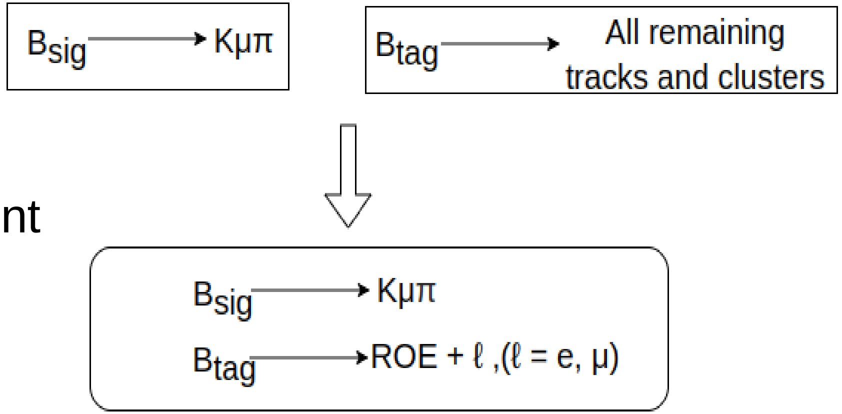
- By using BDT, we have got better signal to background ratio.

- In the preliminary control channel analysis, there is a reasonable agreement of shape between data and MC.

- Next we will work to include the other decay modes.

Backup

Reconstruction methodology



- We are using B2BII module for this analysis.

- We are right now working only on Belle environment and Belle II will be added later.

- We reconstruct B_{sig} by combining K,π and μ tracks.

$$B^+ \to K^+ \tau^-(\to \pi^- \nu_\tau)\mu^+$$

$$B^- \to X \ell^- \nu_\ell$$

- By combining B_{sig} and B_{tag}, we form an Υ(4S) candidate.

Particles selection

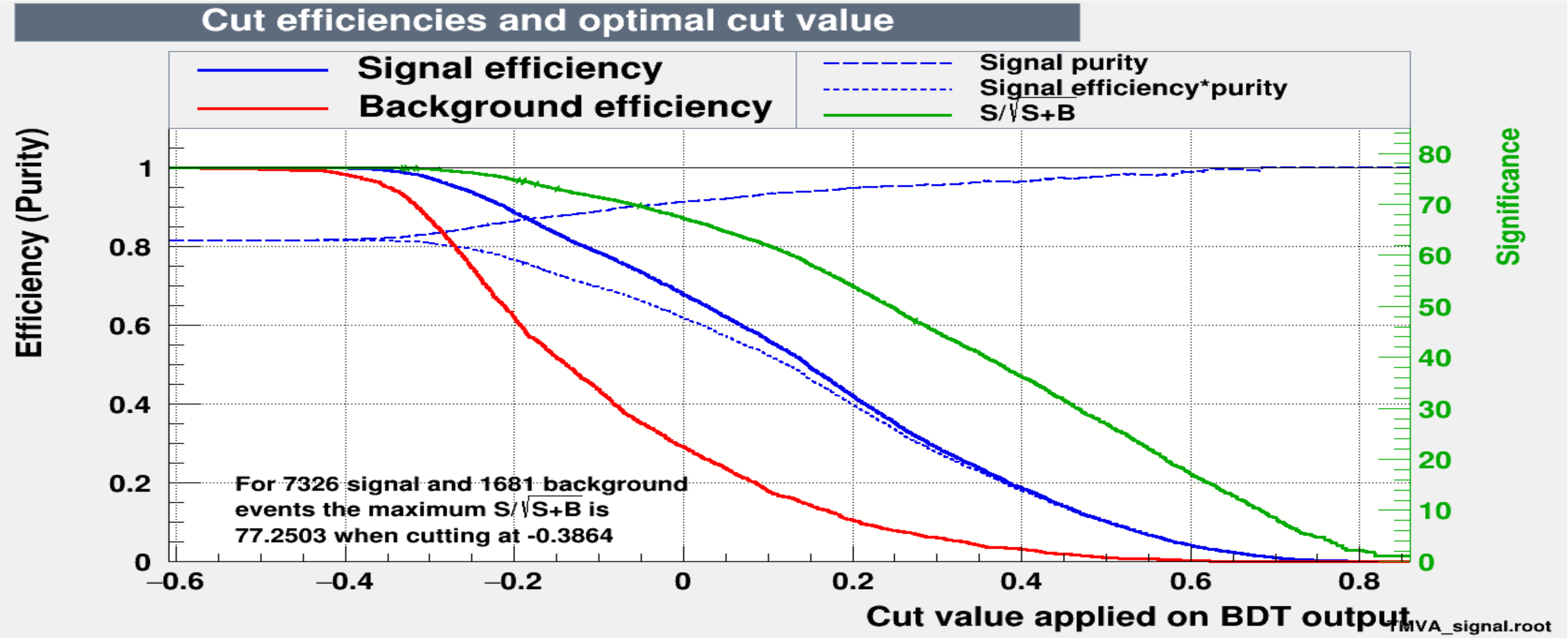
- **e^- selection:** $d_0 < 1$ cm, $|z_0| < 4$ cm, $p > 0.05$ GeV, eIDBelle >0.6, muIDBelle <0.98, atcPIDBelle(3,0) <0.98
- **μ^- selection:** $d_0 < 1$ cm, $|z_0| < 4$ cm, $p > 0.05$ GeV, muIDBelle >0.6, eIDBelle <0.98, atcPIDBelle(3,1) <0.98
- **K^- selection:** $d_0 < 1$ cm, $|z_0| < 4$ cm, $p > 0.05$ GeV, muIDBelle <0.98, eIDBelle <0.98, atcPIDBelle(3,2) >0.6
- **π^+ selection:** $d_0 < 1$ cm, $|z_0| < 4$ cm, $p > 0.05$ GeV, atcPIDBelle(3,2) <0.6
- **π^0 selection:** $0.08 < m_{\pi 0} < 0.18$ GeV

For ROE only

- **Photons selection:** goodBelleGamma==1 and pybdt_bb>0.3 and pybdt_fp>0.3

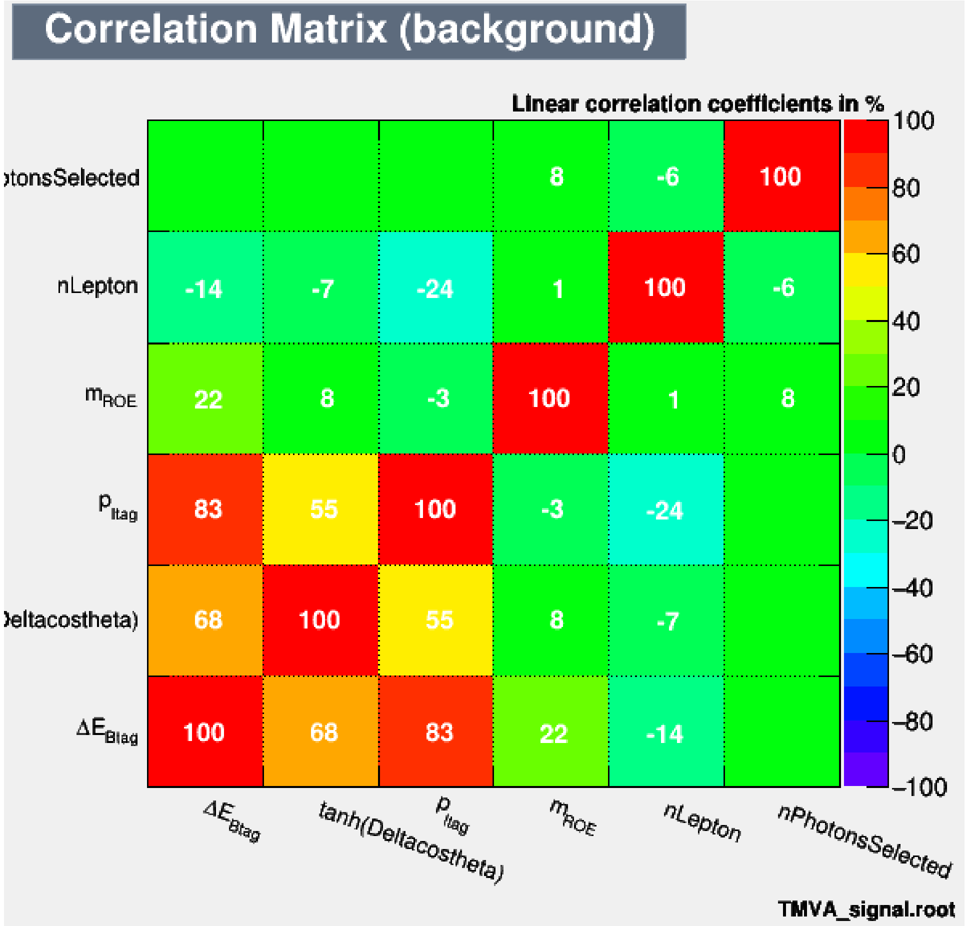
MVA Photon cuts

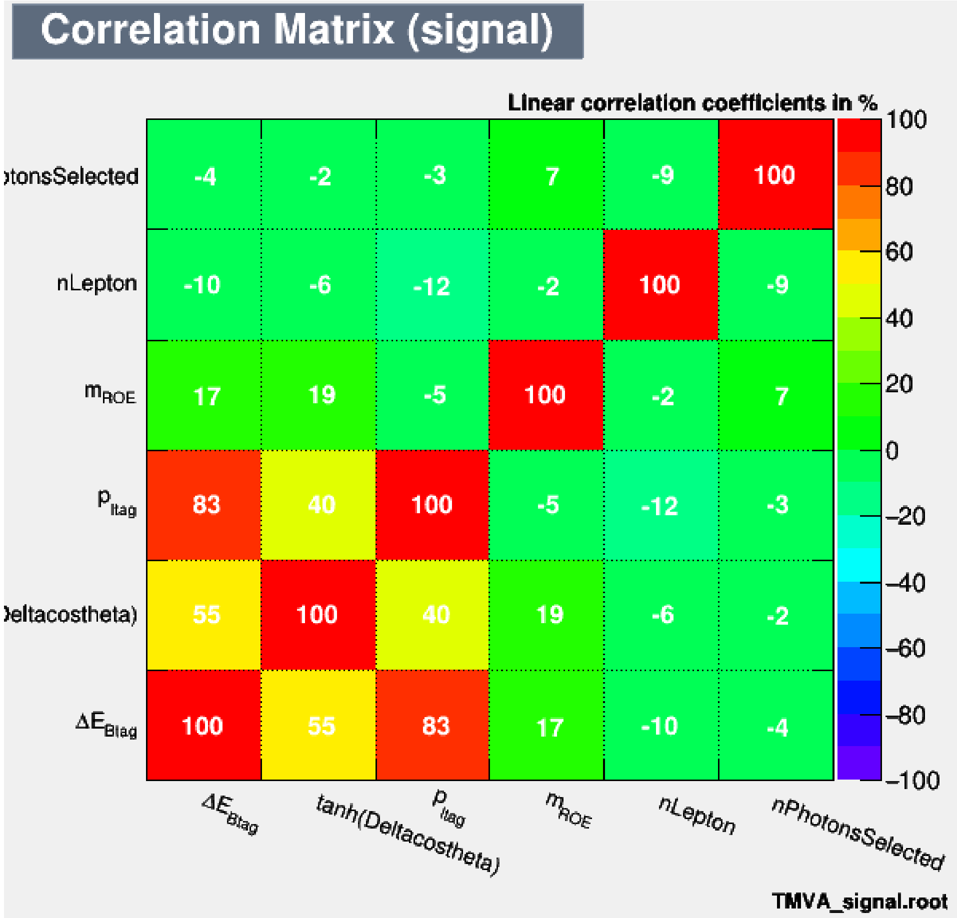
Figure of merit



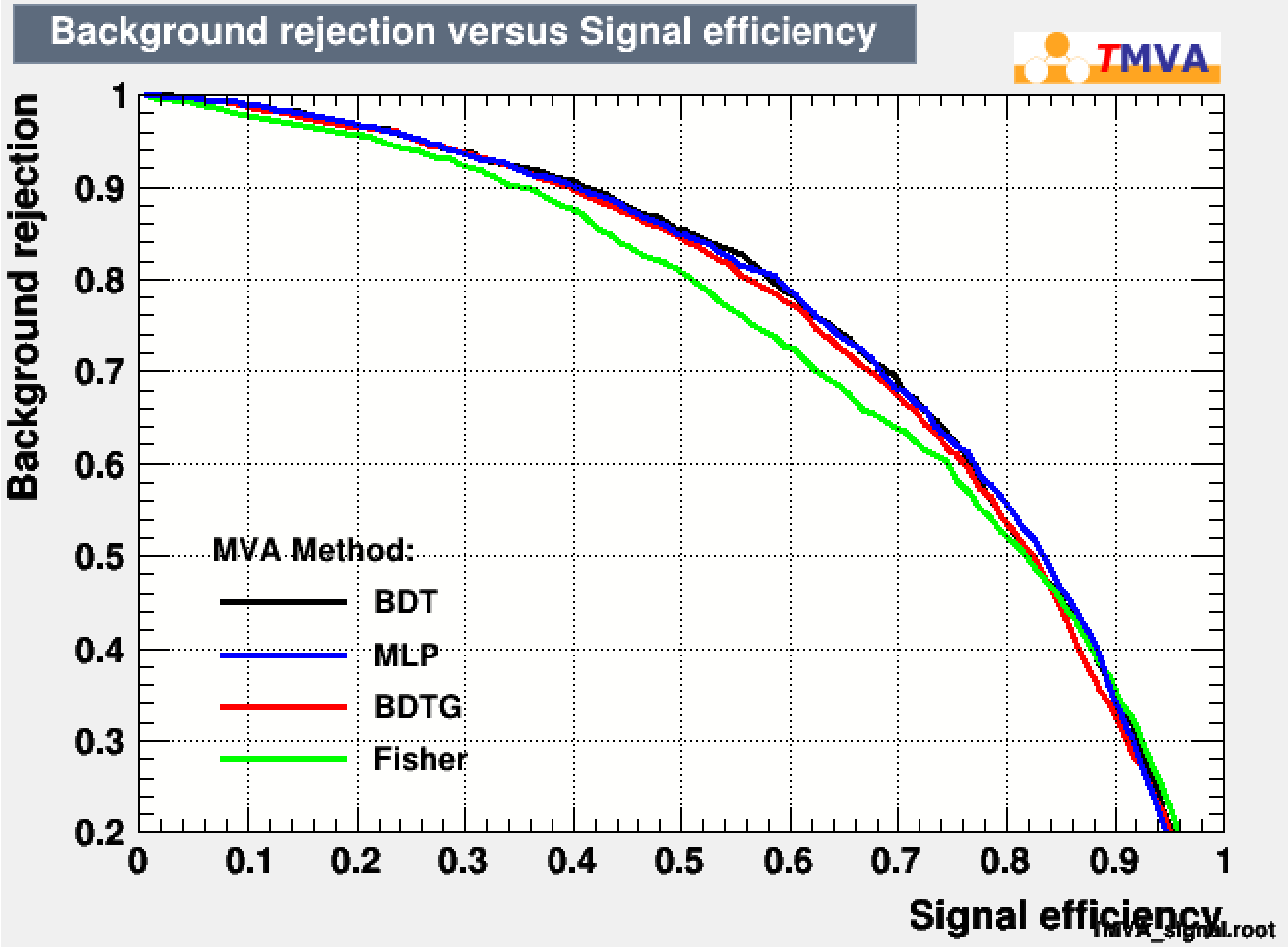
Classifier	(#signal,	#backgr.)	Optimal-cut	S/sqrt(S+B)	NSig	NBkg	EffSig	EffBkg
BDT:	(7326,	1681)	-0.3864	77.2503	7307	1640	0.9974	0.9756
BDTG:	(7326,	1681)	-0.8996	77.21	7326	1677	1	0.9976
Fisher:	(7326,	1681)	-0.6159	77.2212	7323	1670	0.9996	0.9935
MLP:	(7326,	1681)	0.0729	77.2322	7322	1666	0.9995	0.9911

Correlation matrices





ROC



DataSet Name:	MVA Method:		ROC-integ
dataset	BDT	:	0.756
dataset	MLP	:	0.756
dataset	BDTG	:	0.748
dataset	Fisher	:	0.732