Search for $B \to X_s \nu\bar{\nu}$ decay at Belle II experiment

Junewoo Park [A], Akimasa Ishikawa [B], Youngjoon Kwon [C], Yutaka Ushiroda [A,B],
The University of Tokyo [A], KEK [B], Yonsei University [C]

2025.02.22

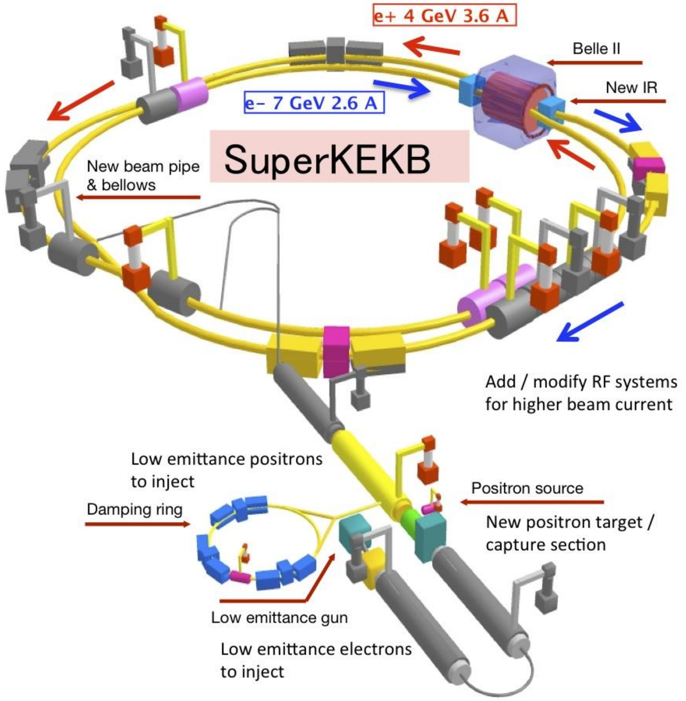
Flavor Physics Mini-Workshop 2025

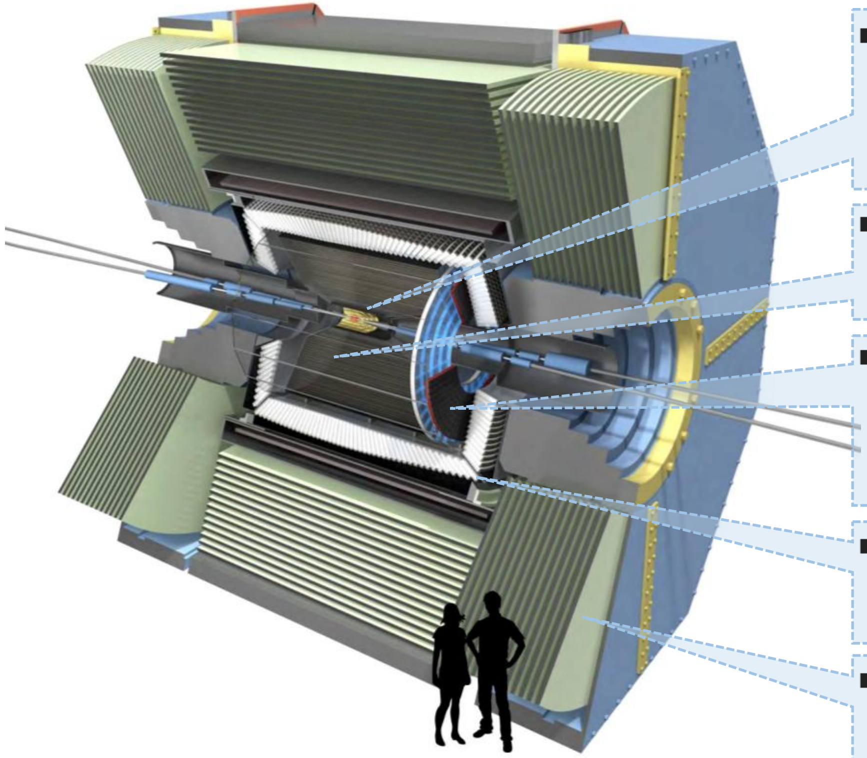
† covers $17^\circ < \theta < 150^\circ$

‡ covers $12.4^\circ < \theta < 155.1^\circ$

Belle II Experiment

- Electrons and positrons are accelerated up to $7\ \mathrm{GeV}$ and $4\ \mathrm{GeV}$ respectively by SuperKEKB accelerator
- Its energy correspond to the resonance of $\Upsilon(4S)$ which mainly decays into B meson pair
- Belle II detector consists of several sub-detector components



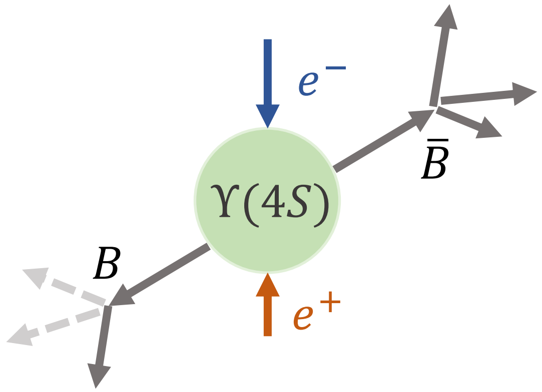


- For vertex detector
  - pixelated silicon sensors (PXD)
  - silicon strip sensors (SVD)
- For tracking
  - central drift chamber (CDC) †
- For particle identification
  - Aerogel Ring Imaging Cherenkov (ARICH)
  - Time of Propagation (TOP) Detector
- For electromagnetic shower energy detection
  - electromagnetic calorimeter (ECL) ‡
- For K_L^0 and μ detection
  - K_L^0 and muon detector (KLM)

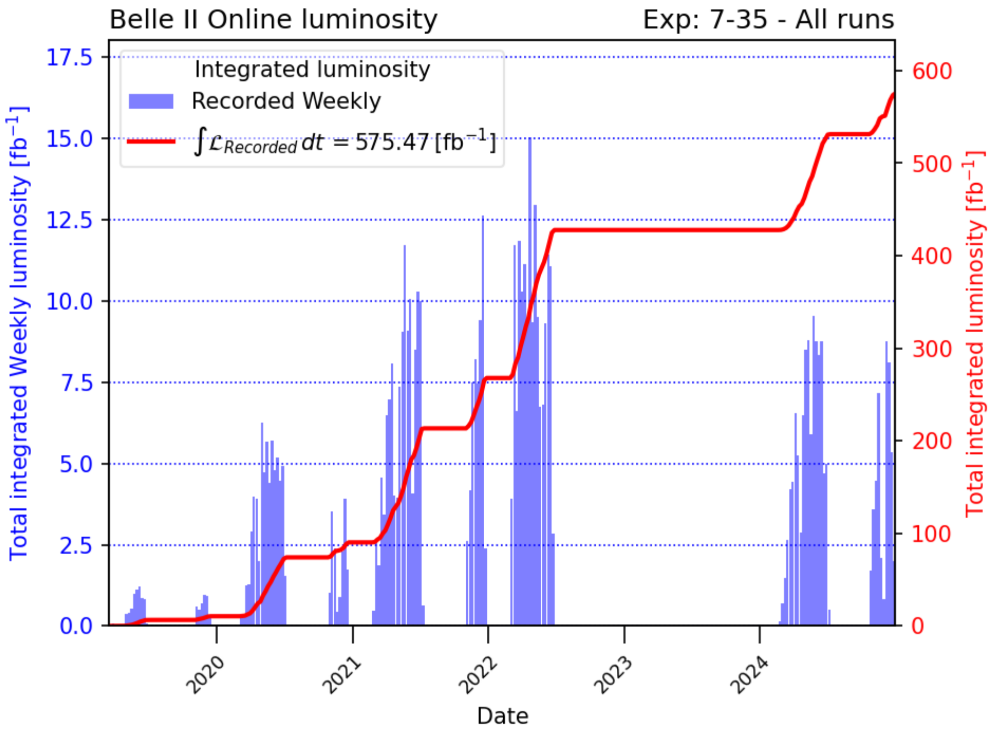
Belle II Experiment

- Entire information of initial states is known in Belle II experiment.

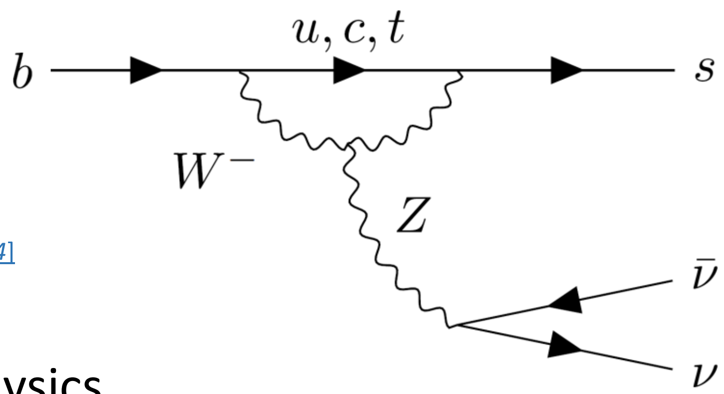
  → Belle II has an advantage on the decay modes with invisible particles, like neutrinos

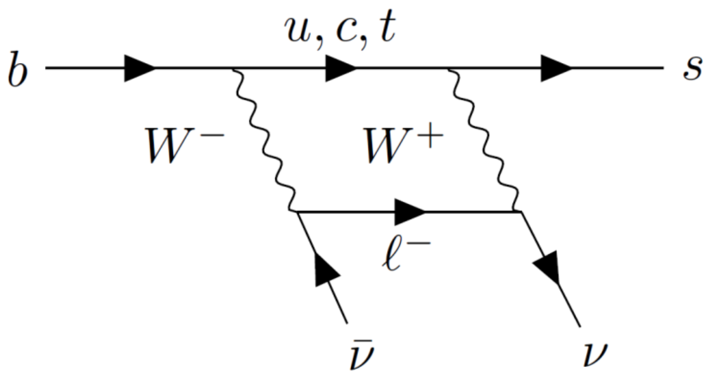


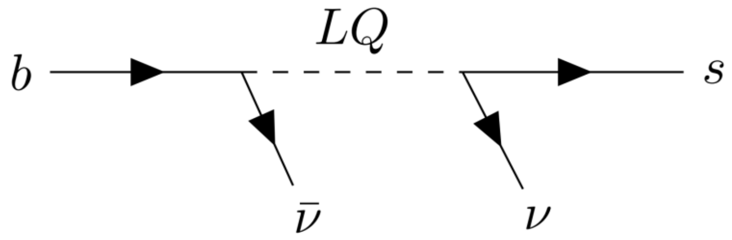
- The target integrated luminosity: $50\ ab^{-1}$
  - Currently, $575.5\ fb^{-1}$ data is recorded. The target sample is $364\ fb^{-1}$ $\Upsilon(4S)$ on-resonance data for this analysis



Motivation







- $B \to X_s \nu\bar{\nu}$ decay
  - Flavour-changing neutral currents process
  - $\mathrm{BR} = (2.9 \pm 0.3) \times 10^{-5}$ at SM [JHEP02(2015)184]

- This decay can give a clue for the new physics
  - leptoquark *[PhysRevD.95.035027]*
  - invisible light scalar *[Eur. Phys. J. C (2017) 77: 650]*
  - fermion dark matter *[arXiv:2405.06742]*

- Previous study
  - $UL(B \to X_s \nu\bar{\nu}) = 6.4 \times 10^{-4}$ (90% CL) by ALEPH † *[Eur.Phys.J.C 19 (2001) 213-227]*
    - There is no $B \to X_s \nu\bar{\nu}$ results from Belle or Belle II because $\nu\bar{\nu}$ makes it challenging
  - $BR(B^+ \to K^+ \nu\bar{\nu}) = (2.3 \pm 0.7) \times 10^{-5}$ by BELLE II ‡ *[PhysRevD.109.112006]*

† tag side is required to satisfy conditions about tracks and missing energy

‡ combined result for inclusive and hadronic tagging analysis

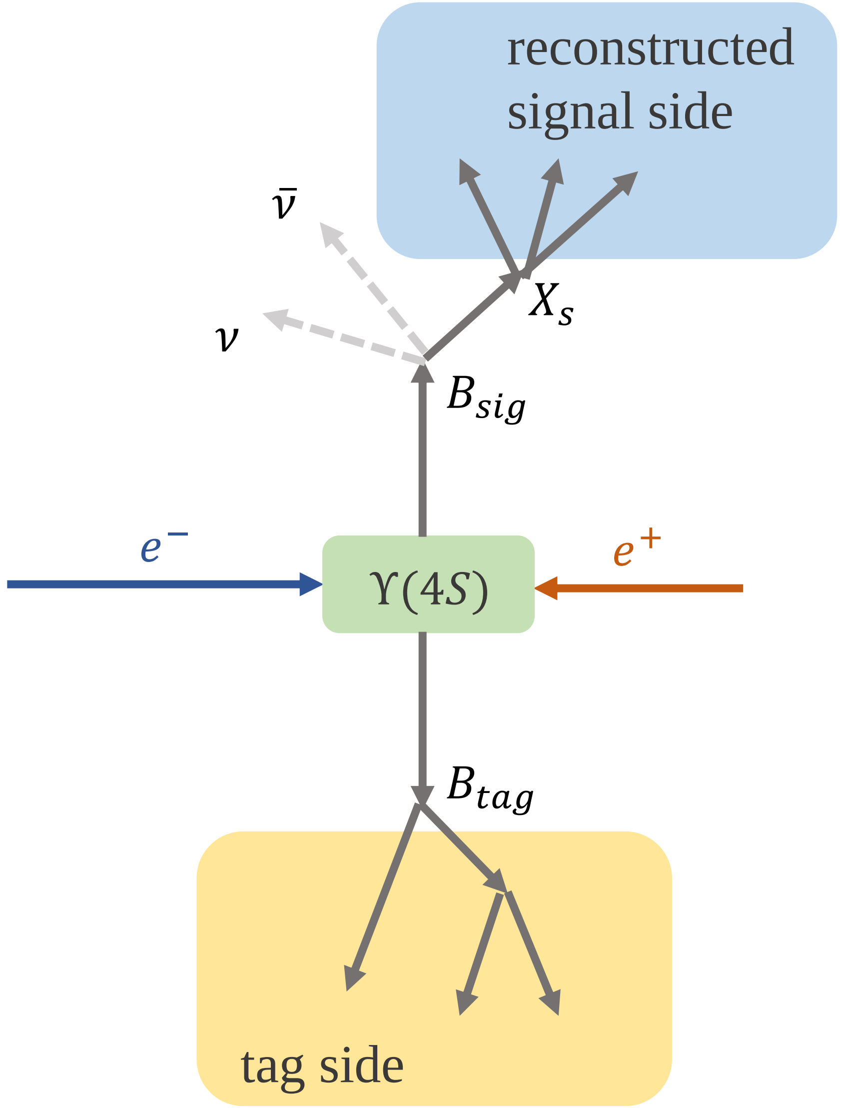
Analysis Strategy

- One side of B meson is hadronically reconstructed
  - Hierarchical reconstruction technique is used [Comput Softw Big Sci 3, 6 (2019)]
  - This B meson is called as tag side B meson, B_{tag}

- For signal side B meson, sum of exclusive method is used
  - 30 decay modes are reconstructed for X_s

	$B^0\bar{B}^0$			$B^\pm$		
K	K_S^0			$K^\pm$		
$K\pi$	$K^\pm\pi^\mp$	$K_S^0\pi^0$		$K^\pm\pi^0$	$K_S^0\pi^\pm$	
$K2\pi$	$K^\pm\pi^\mp\pi^0$	$K_S^0\pi^\pm\pi^\mp$	$K_S^0\pi^0\pi^0$	$K^\pm\pi^\mp\pi^\pm$	$K_S^0\pi^\pm\pi^0$	$K^\pm\pi^0\pi^0$
$K3\pi$	$K^\pm\pi^\mp\pi^\pm\pi^\mp$	$K_S^0\pi^\pm\pi^\mp\pi^0$	$K^\pm\pi^\mp\pi^0\pi^0$	$K^\pm\pi^\mp\pi^\pm\pi^0$	$K_S^0\pi^\pm\pi^\mp\pi^\pm$	$K_S^0\pi^\pm\pi^0\pi^0$
$K4\pi$	$K^\pm\pi^\mp\pi^\pm\pi^\mp\pi^0$	$K_S^0\pi^\pm\pi^\mp\pi^\pm\pi^\mp$	$K_S^0\pi^\pm\pi^\mp\pi^0\pi^0$	$K^\pm\pi^\mp\pi^\pm\pi^\mp\pi^\pm$	$K_S^0\pi^\pm\pi^\mp\pi^\pm\pi^0$	$K^\pm\pi^\mp\pi^\pm\pi^0\pi^0$
$3K$	$K^\pm K^\mp K_S^0$			$K^\pm K^\mp K^\pm$		
$3K\pi$	$K^\pm K^\mp K^\pm\pi^\mp$	$K^\pm K^\mp K_S^0\pi^0$		$K^\pm K^\mp K^\pm\pi^0$	$K_S^0 K^\pm K^\mp\pi^\pm$	

  - It covers ~92% of entire X_s sample, with assuming K^0 equally decays into K_S^0 or K_L^0



Event Generation

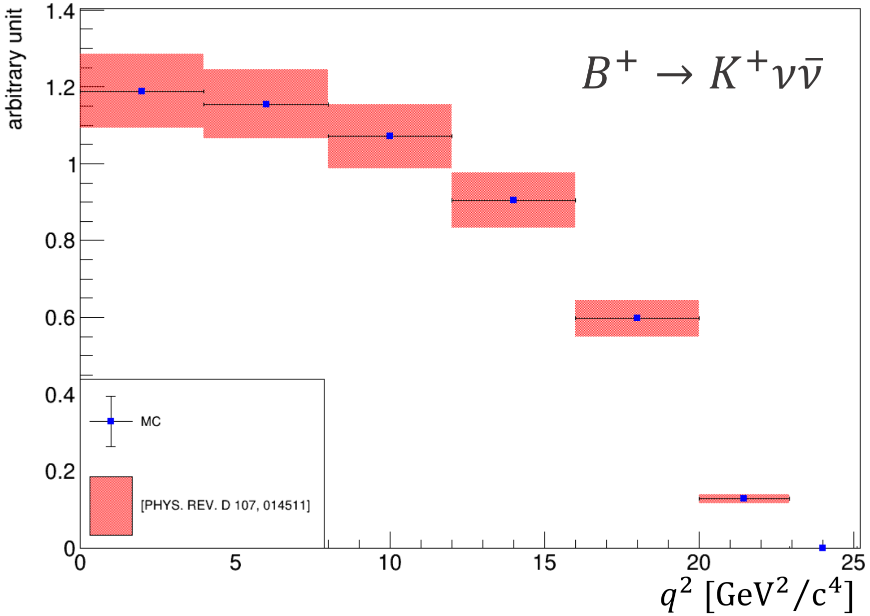
- We produce $B \to K\nu\bar{\nu}$, $B \to K^*\nu\bar{\nu}$, and non-resonant $B \to X_s\nu\bar{\nu}$ MC sample separately
- For $B \to K\nu\bar{\nu}$ and $B \to K^*\nu\bar{\nu}$ Monte-Carlo (MC) samples, the SM form factors are used

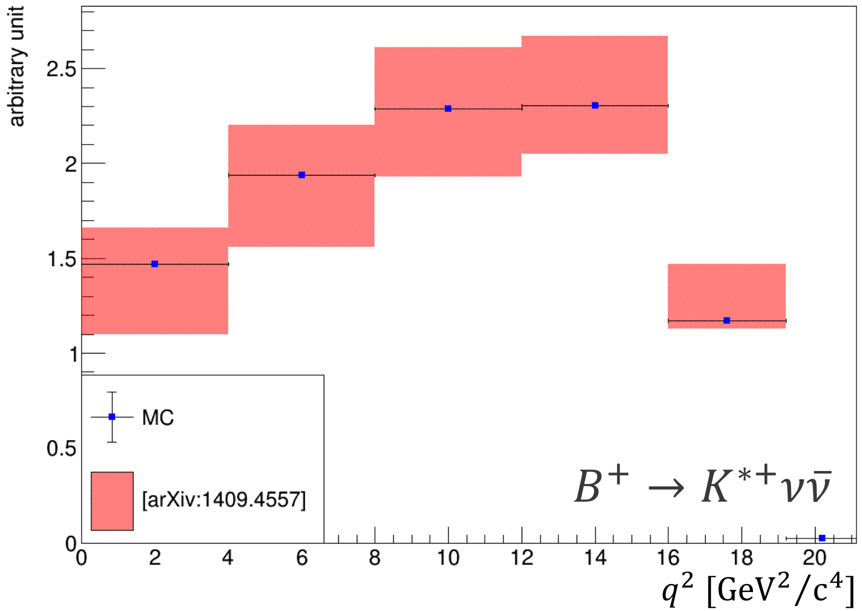
[Phys.Rev.D 107 (2023) 1, 014511] [JHEP02(2015)184] [JHEP02(2015)184]

Decay amplitude for $B \to K\nu\bar{\nu}$: $\mathcal{M}(b \to s\nu\bar{\nu}) \propto f_+(s)\left\{(p_B + p)_\mu - \frac{m_B^2 - m_K^2}{s} q_\mu\right\}(\bar{\nu}\gamma^\mu(1-\gamma_5)\nu)$

Decay amplitude for $B \to K^*\nu\bar{\nu}$: $\mathcal{M}(b \to s\nu\bar{\nu}) \propto \mathcal{T}_\mu\,(\bar{\nu}\gamma^\mu(1-\gamma_5)\nu)$

$$\mathcal{T}_\mu = (m_B + m_{K^*})A_1(s)\epsilon^*_\mu - A_2(q^2)\frac{\epsilon^* \cdot q}{m_B + m_{K^*}}(p + p_{K^*})_\mu + i\frac{2V(s)}{m_B + m_{K^*}}\epsilon_{\mu\nu\rho\sigma}\epsilon^{*\nu}p^\rho p^\sigma_{K^*}$$





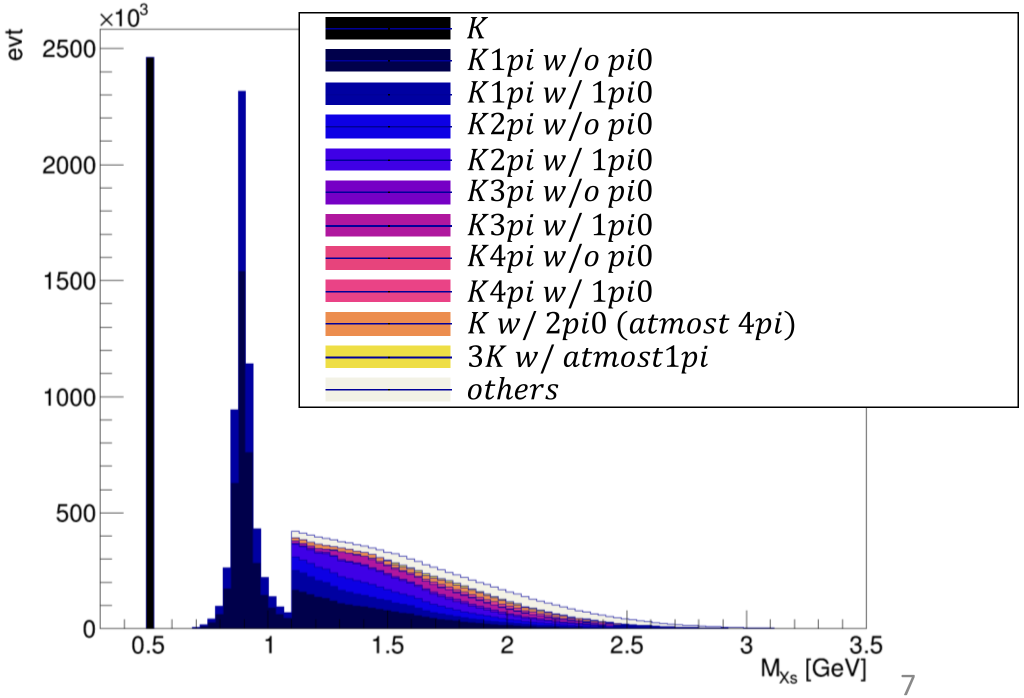
※ $q^2 = (p_\nu + p_{\bar{\nu}})^2$

Event Generation

- Non-resonant $B \to X_s \nu\bar{\nu}$ MC sample is produced by the following distribution [JHEP04(2009)022]

$$\frac{d\Gamma}{dq^2} \propto \sqrt{\lambda(1, \hat{m}_s^2, s_b)}\left[3s_b(1 + \hat{m}_s^2 - s_b - 4\hat{m}_s) + \lambda(1, \hat{m}_s^2, s_b)\right]$$

- To determine the mass of X_s, Fermi motion model is adopted [PhysRevD.55.4105]
  - In Fermi motion model, quarks have some velocity inside B meson
  - Also, $M_{X_s} > 1.1\ \mathrm{GeV/c^2}$ is required for non-resonant $X_s \nu\bar{\nu}$ MC sample
- Hadronization is done by PYTHIA



Signal Selection

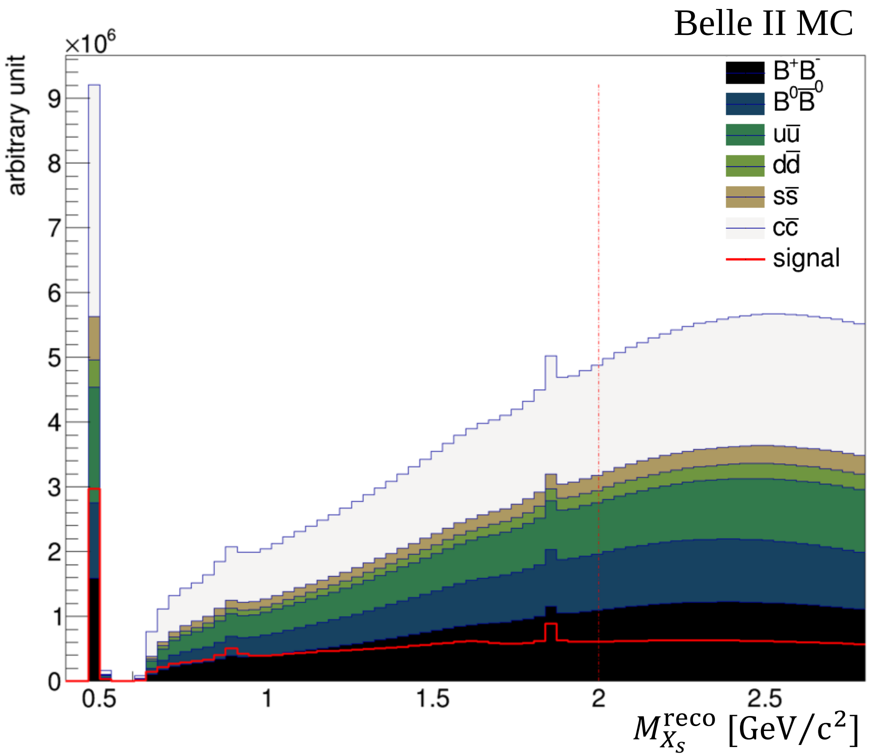
- Preselection

To reduce combinatorial backgrounds

- reconstructed X_s mass, $M_{X_s}^{\mathrm{reco}} < 2.0\ \mathrm{GeV/c^2}$

To reject backgrounds with additional particles, outside of $B_{tag}X_s$ candidates

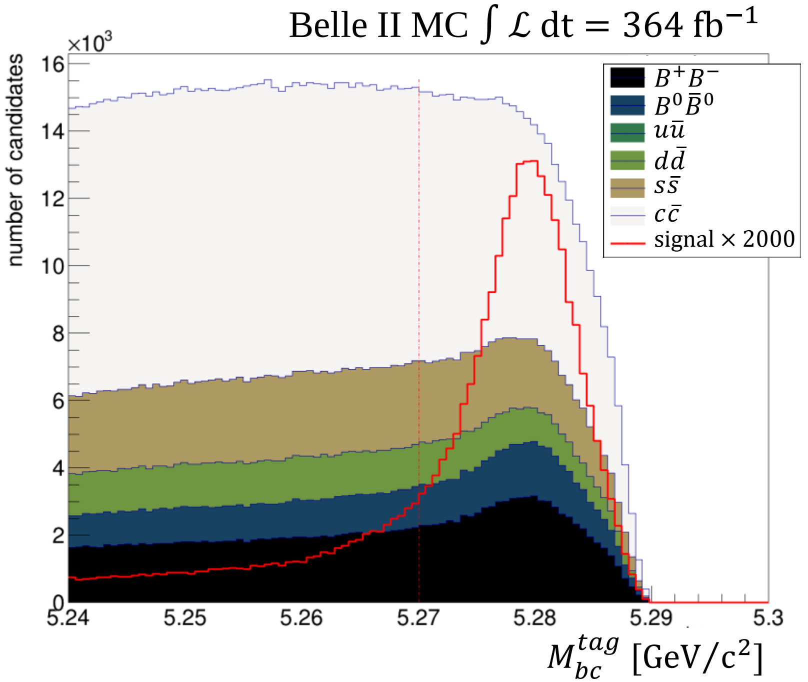
- the number of tracks ‡ = 0
- the number of π^0 candidates = 0
- the number of K_S^0 candidates = 0

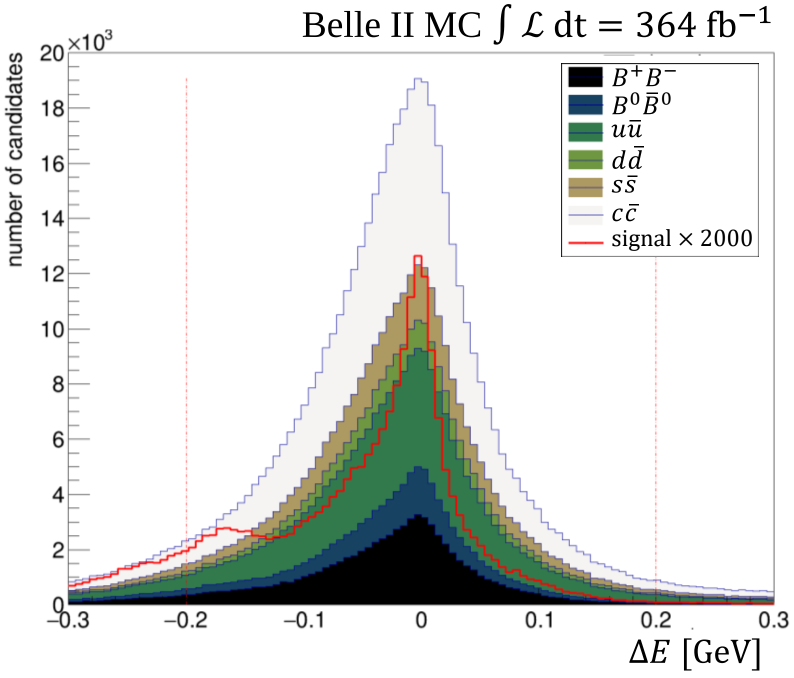


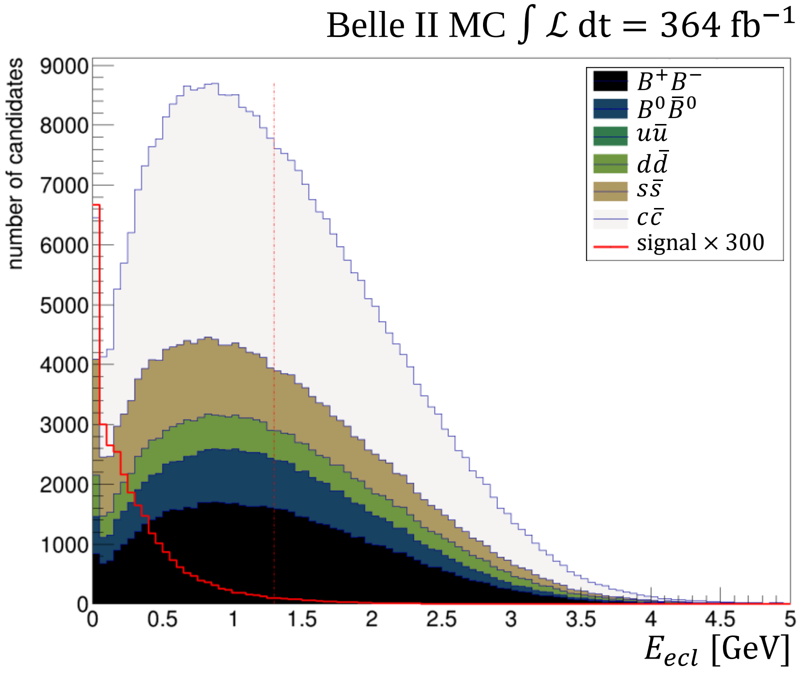
---

‡ $dr < 2$ cm and $|dz| < 4$ cm

Signal Selection







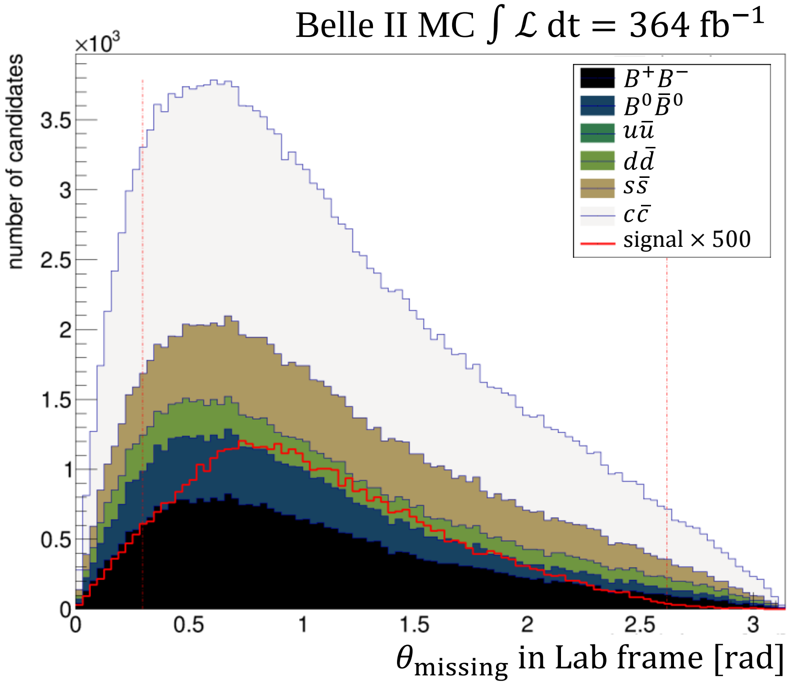
- Several selections are applied
  - $M_{bc}^{tag} > 5.27$ GeV/c^2
  - $|\Delta E^{tag}| < 0.2$ GeV
  - $E_{ecl} < 1.3$ GeV

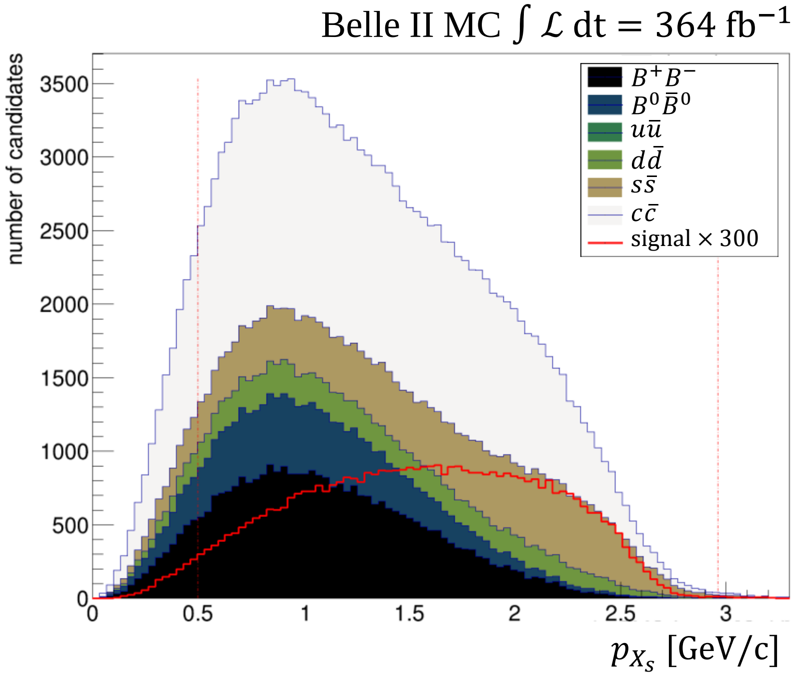
※ $M_{bc}^{tag} \equiv \sqrt{E_{beam}^2 - |p_B^2|}$

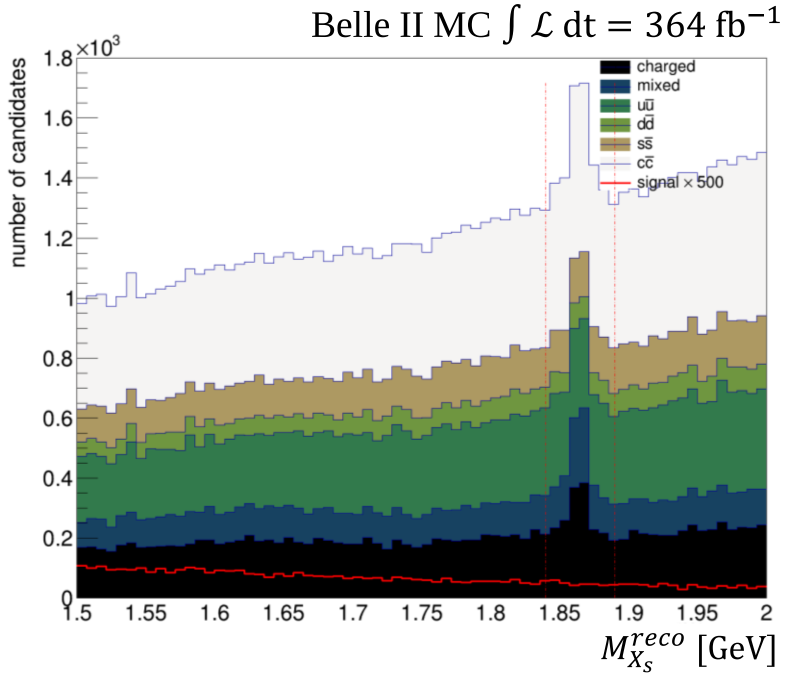
※ $\Delta E^{tag} \equiv E_B - E_{beam}$

※ E_{ecl}: remaining energy deposited in the calorimeter

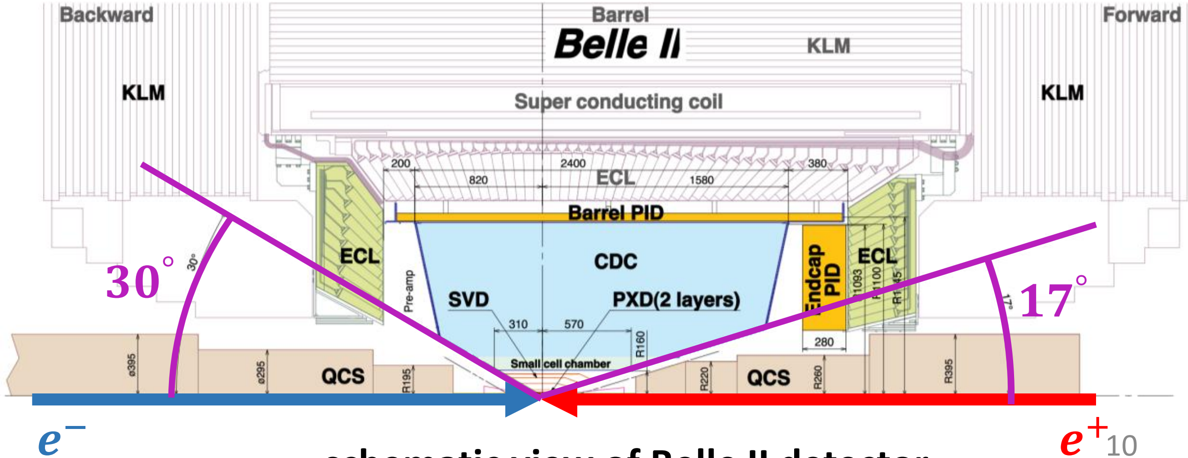
Signal Selection







- Several selections are applied
  - $17^\circ < \theta_{\mathrm{missing}} < 150^\circ$ **missing tracks can mimic $\nu\bar{\nu}$**
  - $0.5 < p_{X_s} < 2.96\ \mathrm{GeV/c}$
  - reject $1.84 < M_{X_s}^{reco} < 1.89\ \mathrm{GeV/c^2}$

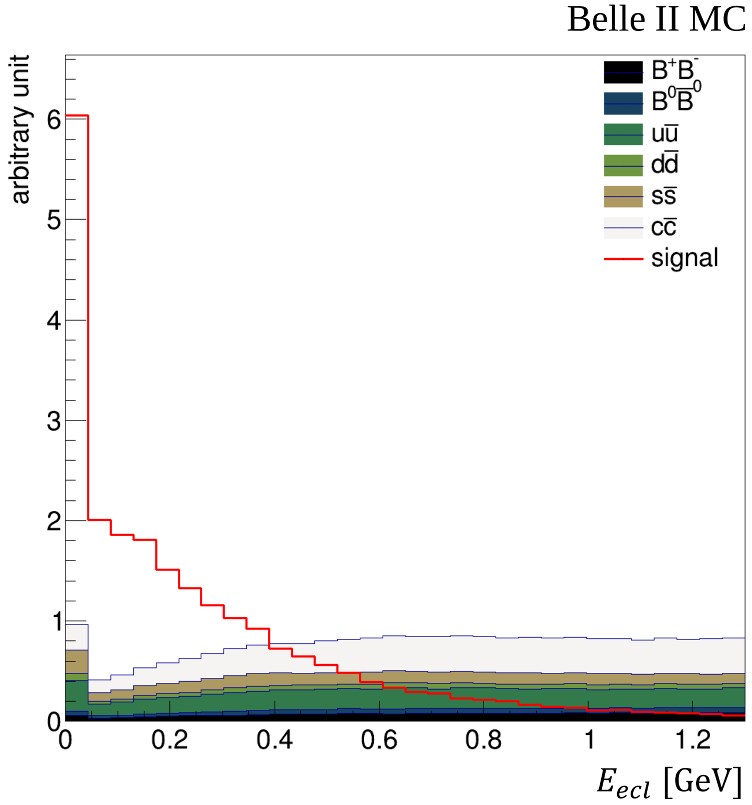


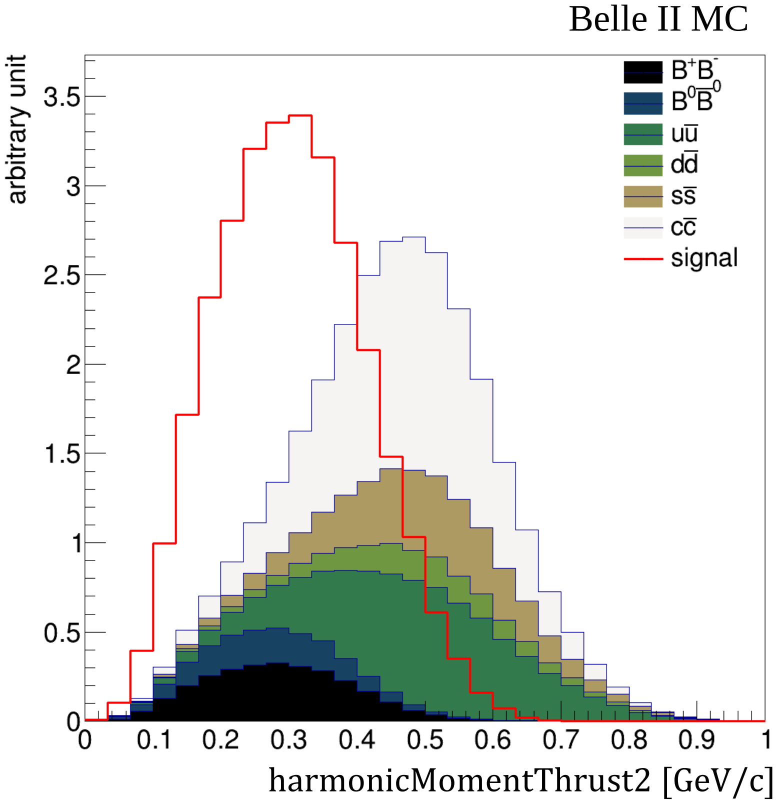
schematic view of Belle II detector

※ $M_{X_s}^{reco}$: mass of reconstructed X_s candidate

Signal Selection

- Boosted Decision Tree (BDT) is used as multivariate analysis (MVA) method
- 32 variables are used
  - Variables are selected based on discriminant power and goodness of data/MC agreement on sideband
  - Example of powerful variables: E_{ecl} and event shape variable *[Nucl. Phys. B 149, 413 (1979)]*



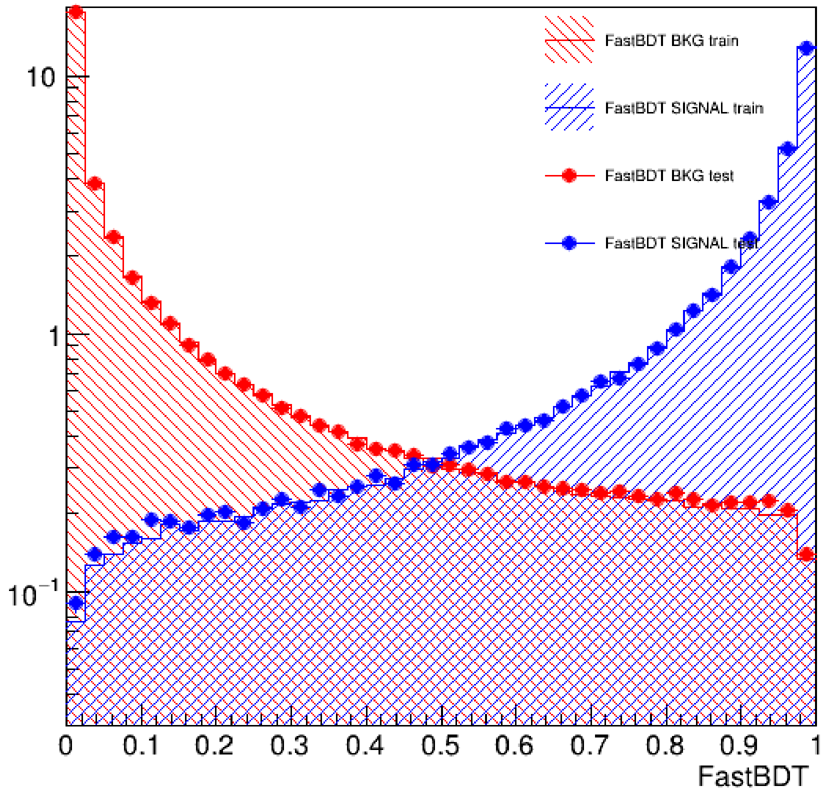


Signal Selection

- BDT has several hyperparameters
  - Grid search is done to achieve high performance and low overtraining

hyper parameter	tested values	selected value
nTrees	100, 500, 1000, 1500, 2000	2000
depth	2, 3, 4	3
shrinkage	0.05, 0.1, 0.15, 0.2	0.05
subsample	0.3, 0.4, 0.5, 0.6, 0.7	0.5
binning	6, 7, 8, 9	6

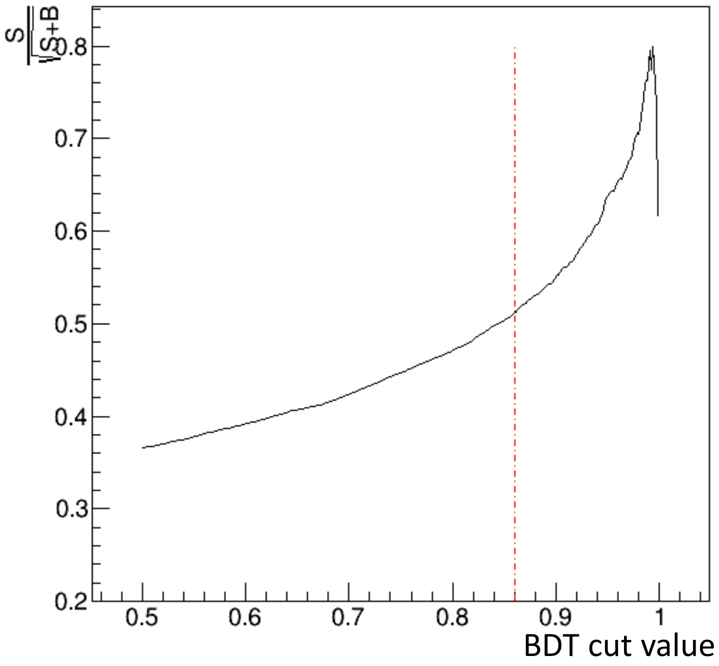
- AUC and BDT output is checked
  - AUC value for training sample: 0.966
  - AUC value for test sample: 0.964



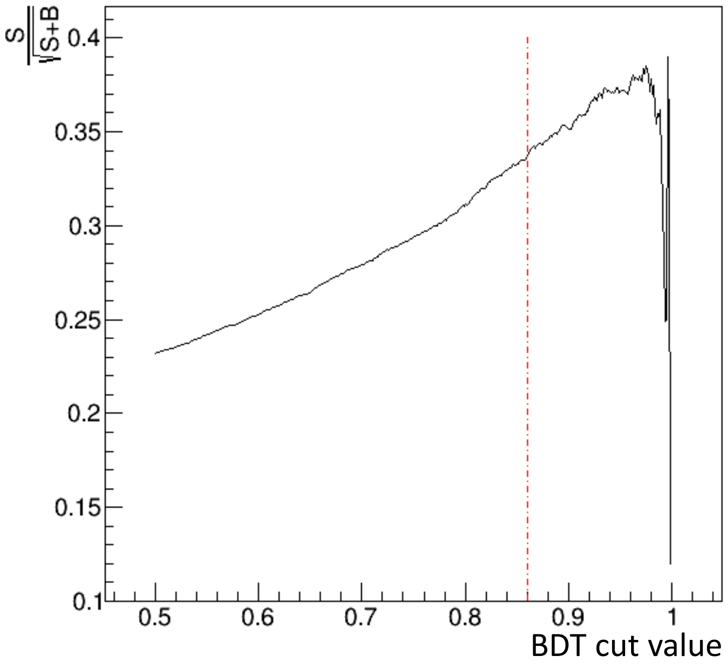
Signal Selection

- BDT output > 0.86 is applied
  - Because it is used as a fitting variable, deliberately loose cut is applied
  - Figure of merit $\left(= \frac{S}{\sqrt{S+B}}\right)$ is used to select this criteria

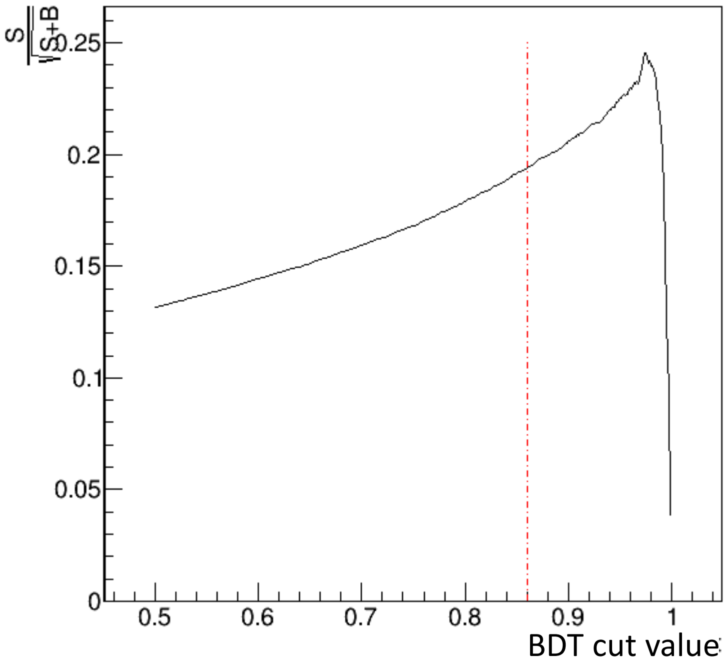
* $M_{X_s}^{\mathrm{reco}}$: mass of reconstructed X_s candidate



$0.0 < M_{X_s}^{\mathrm{reco}} < 0.6\ \mathrm{GeV/c^2}$



$0.6 < M_{X_s}^{\mathrm{reco}} < 1.0\,\mathrm{GeV/c^2}$

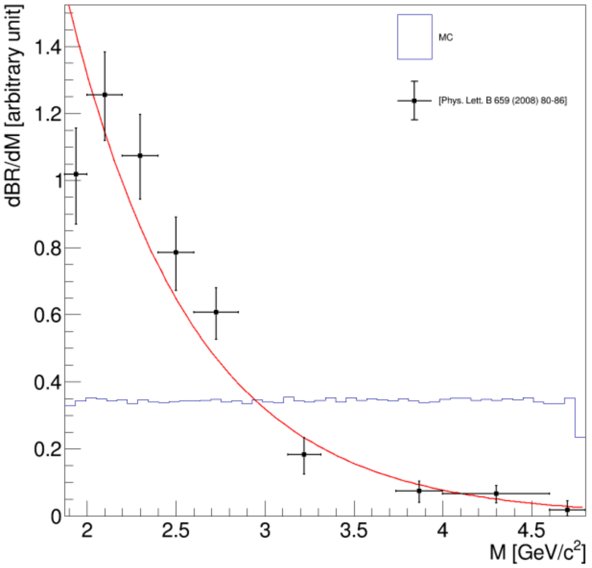


$1.0 < M_{X_s}^{\mathrm{reco}} < 2.0\,\mathrm{GeV/c^2}$

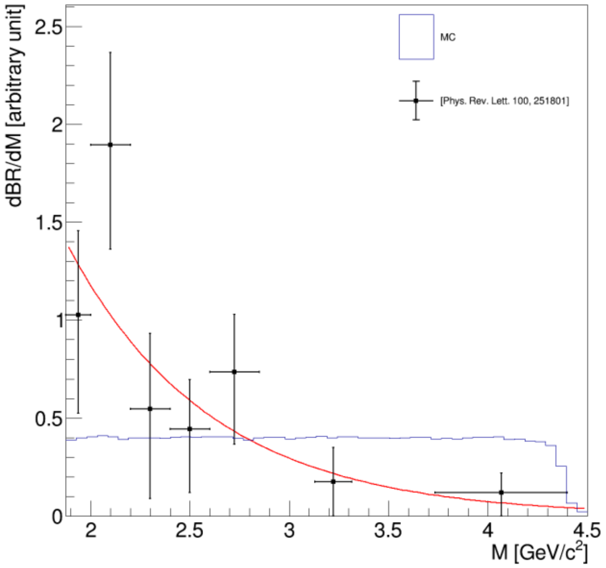
- Finally, signal efficiency is about 0.11%

Corrections - $B \to K^{(*)} n\bar{n}$

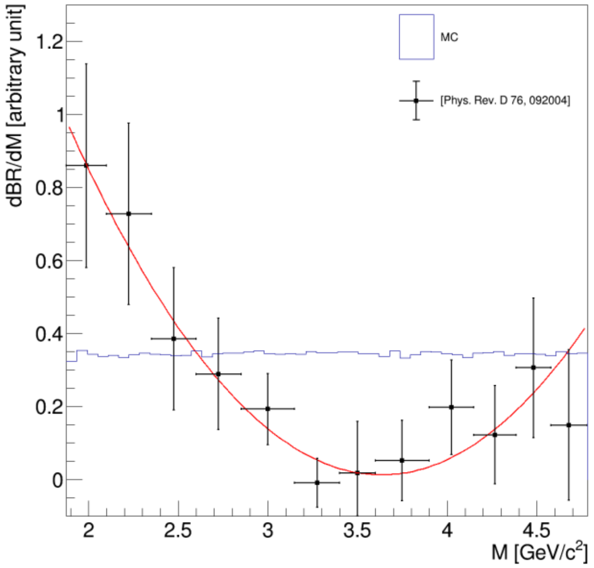
- $B \to K^{(*)} n\bar{n}$ mismodeling correction
  - $B \to K^{(*)} n\bar{n}$ MC sample is not modeled well, even though it can mimic $B \to K^{(*)} \nu\bar{\nu}$
  - Produce $B \to K^{(*)} n\bar{n}$ MC with flat $M_{n\bar{n}}$ distribution and reweight it!
  - use $B \to K^{(*)} p\bar{p}$ study result to obtain $M_{n\bar{n}}$ distribution for $B \to K^{(*)} n\bar{n}$ decay
  - use exponential function or 2nd order polynomial to get a form of $M_{n\bar{n}}$ distribution



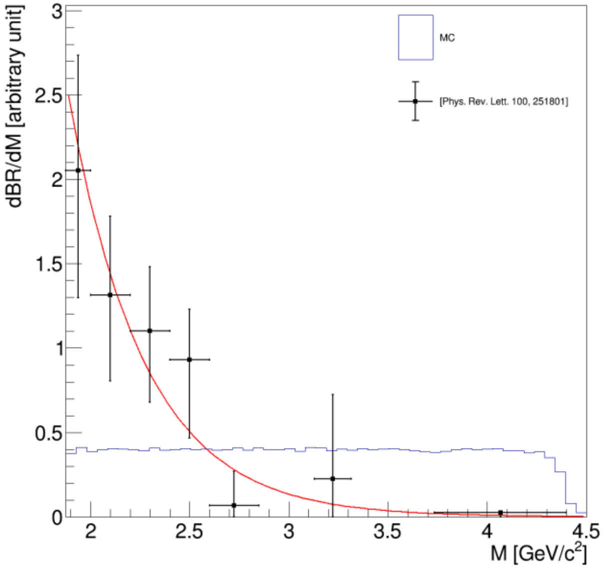
$B^+ \to K^+ p\bar{p}$



$B^+ \to K^{*+} p\bar{p}$



$B^0 \to K^0 p\bar{p}$



$B^+ \to K^{*0} p\bar{p}$

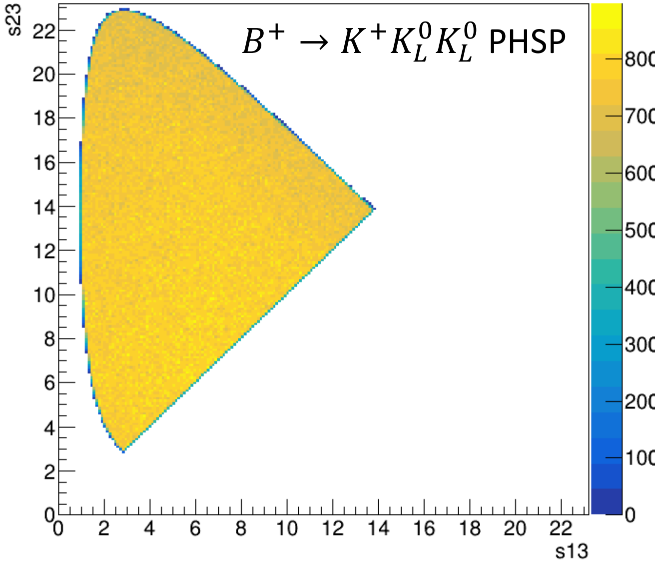
Corrections - $B \to K$

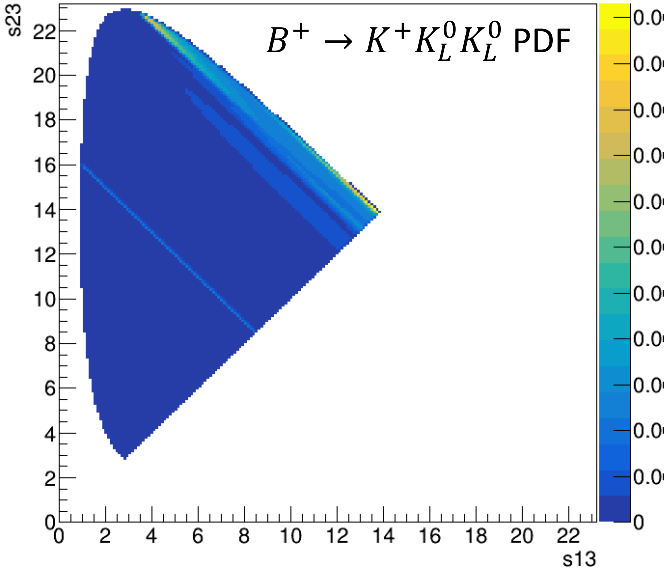
- $B \to KK_L^0K_L^0$ mismodeling correction
  - $B \to KK_L^0K_L^0$ MC sample is not modeled well
  - need to correct Dalitz plot
  - use $B \to KK_S^0K_S^0$ study result to obtain PDF

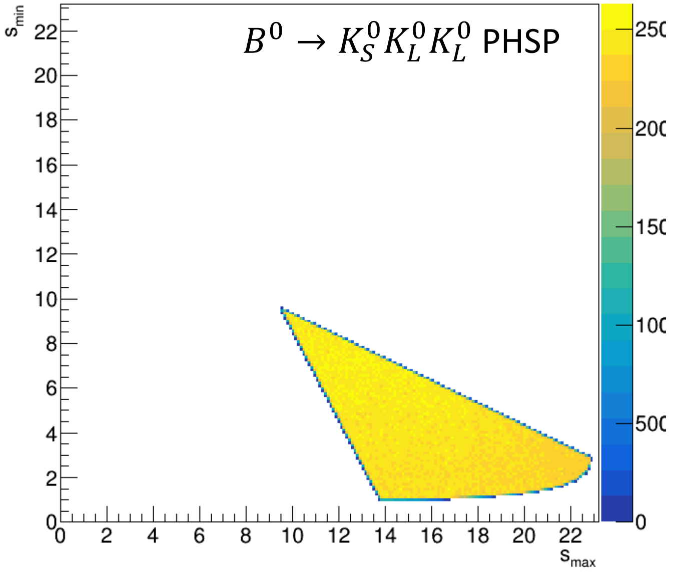
[Physical Review D 85.11 (2012): 112010]
[Physical Review D 85.5 (2012): 054023]

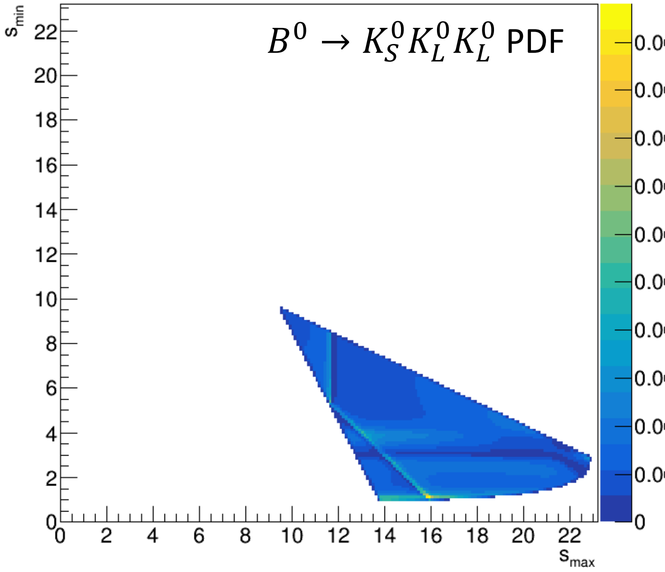
$$F_j^L\left(s_{12}, s_{23}\right) = R_j\left(m\right) X_L\left(|\vec{p}^{\,*}|\, r'\right) T_j\left(L, \vec{p}, \vec{q}\right)$$

$$F_j^L\left(s_{\min}, s_{\max}, L\right) = R_j\left(m\right) X_L\left(|\vec{p}^{\,*}|\, r'\right) X_L\left(|\vec{q}|\, r\right) T_j\left(L, \vec{p}, \vec{q}\right)$$









Off-resonance sample

- Off-resonance
  - Off-resonance samples are analyzed for the correction and validation

- M_{bc}^{tag} and ΔE^{tag} are calculated as follows, when analyze the off-resonance sample:

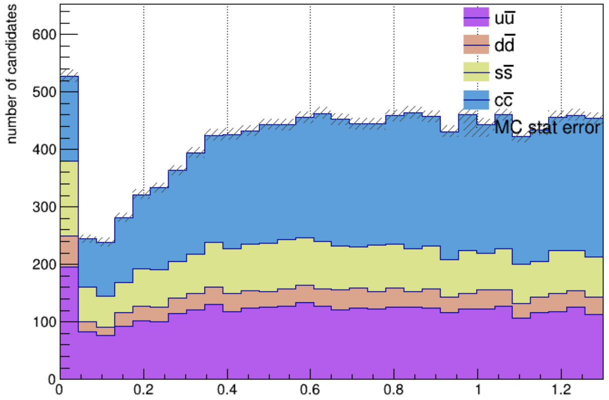
$$M_{bc,corrected}^{tag} = \sqrt{\left(E_{beam}^{off}\right)^2 - (p_B^*)^2} + \left(E_{beam}^{on} - E_{beam}^{off}\right)$$

$$\Delta E^{tag} = E_B^* - E_{beam}^{off}$$

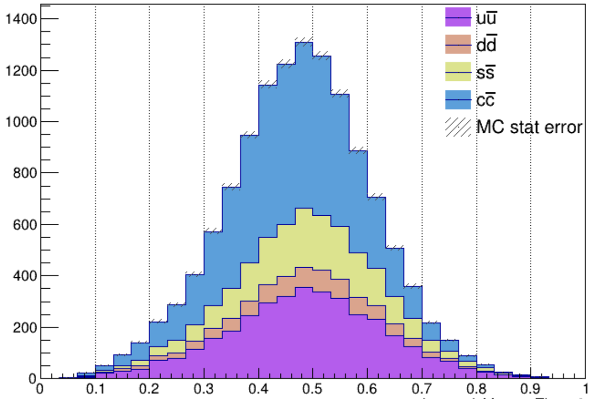
  - this correction is applied to use the same M_{bc} and ΔE cut

Off-resonance sample

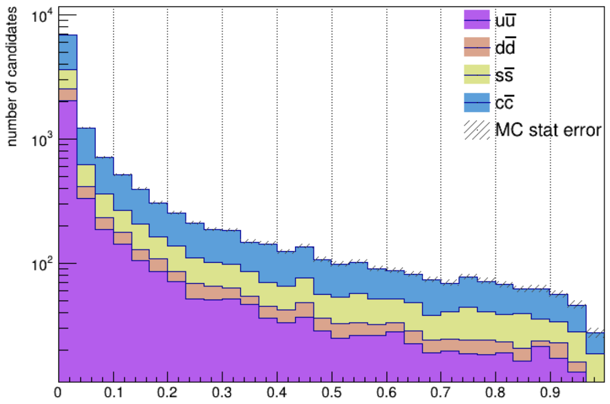
- Off-resonance
  - Off-resonance samples are analyzed



E_{ecl}



Event shape variable

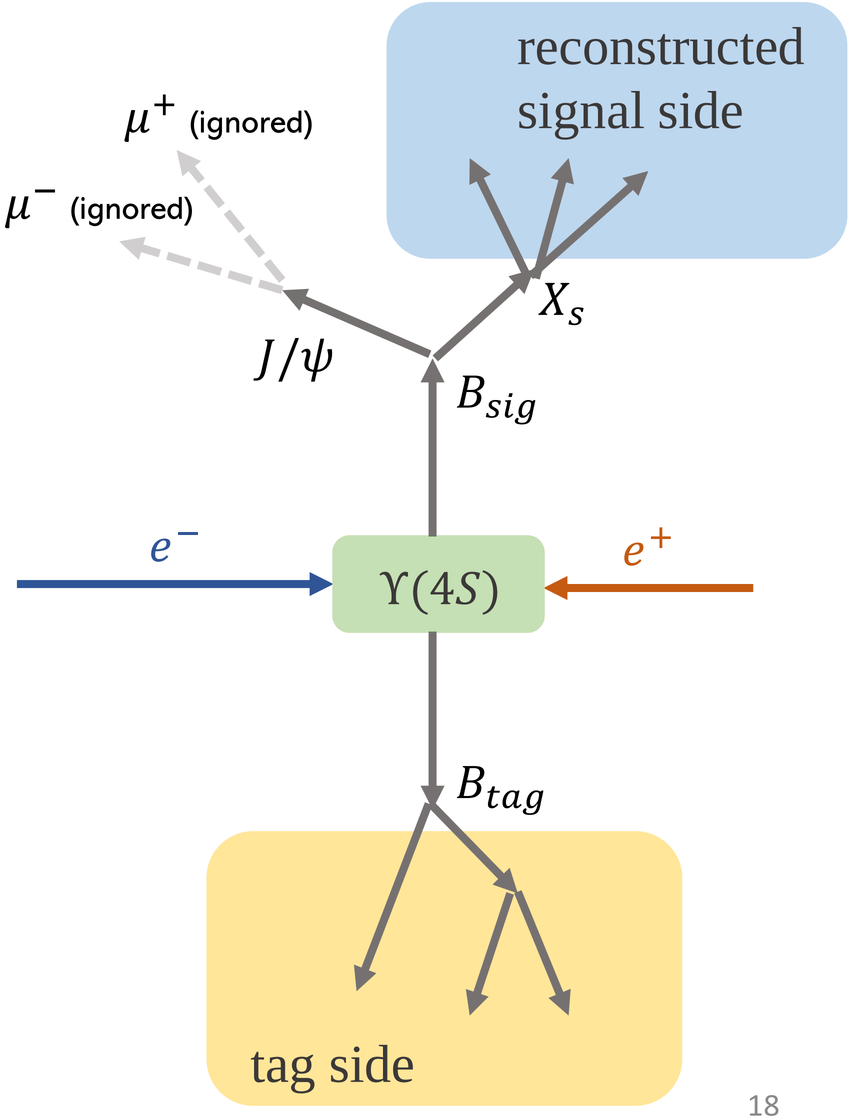


BDT output

- This result is used to correct the continuum MC sample in on-resonance
  - Overall normalization factor is obtained for each $M_{X_s}^{reco}$ region
- Also, systematic uncertainty from the shape of variables is estimated by off-resonance sample

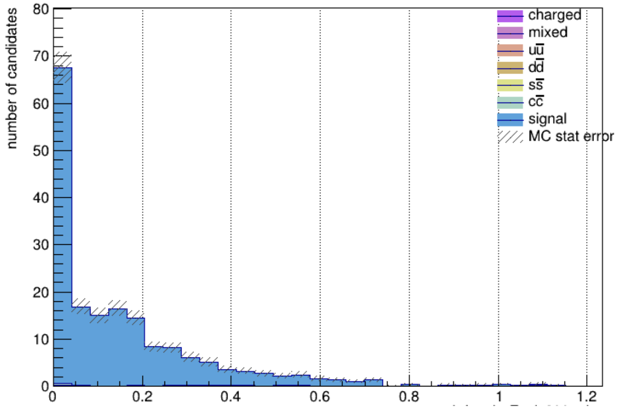
Control Channel

- $B \to X_s J/\psi$ analysis is also done
  - Analysis method is almost same
  - after reconstruction of $B \to X_s J/\psi$, ignore μ

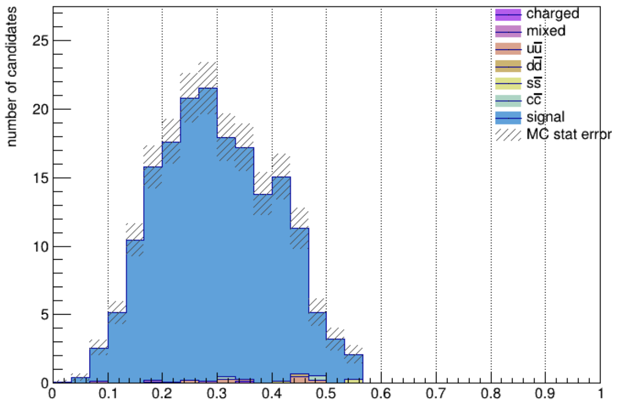


Control Channel

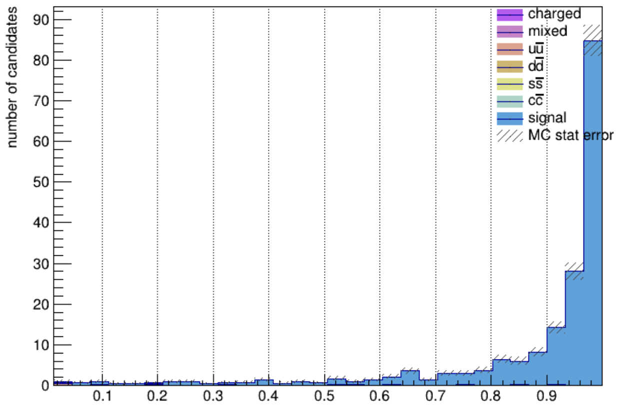
- $B \rightarrow X_s J/\psi$ analysis is also done



E_{ecl}



harmonicMomentThrust2



FBDT

- This result is used to obtain correction/systematic uncertainty
  - BDT Efficiency correction
  - Systematic uncertainty from the Efficiency for B_{tag}

Systematics

- Several systematics are estimated

some major sources:

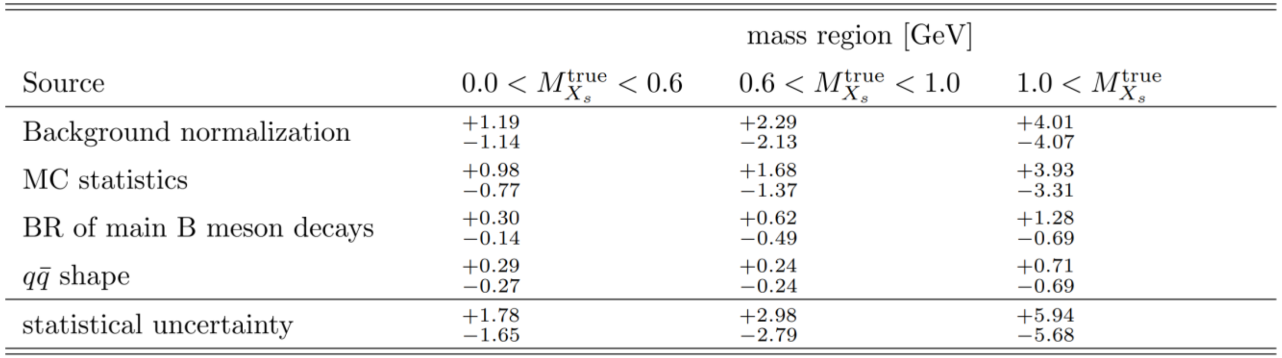
Background normalization
- apply $\pm 20\%$ uncertainties on each backgrounds
- motivated by the data/MC of sideband

MC Statistics
- comes from statistical uncertainty of MC sample

BR of main B meson decays
- uncertainty of BR which is obtained from PDG
- for major B decays, uncertainties are applied

$q\bar{q}$ **background shape**
- estimated by training another BDT with 60 MeV lower E_{beam} sample

σ_μ at each $M_{X_s}^{\text{true}}$ region

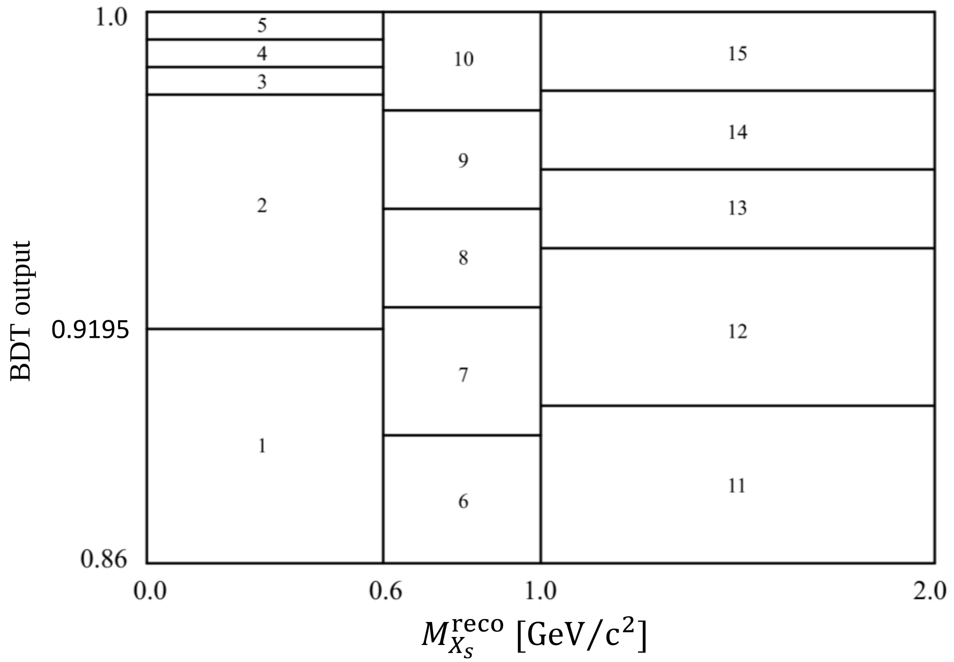


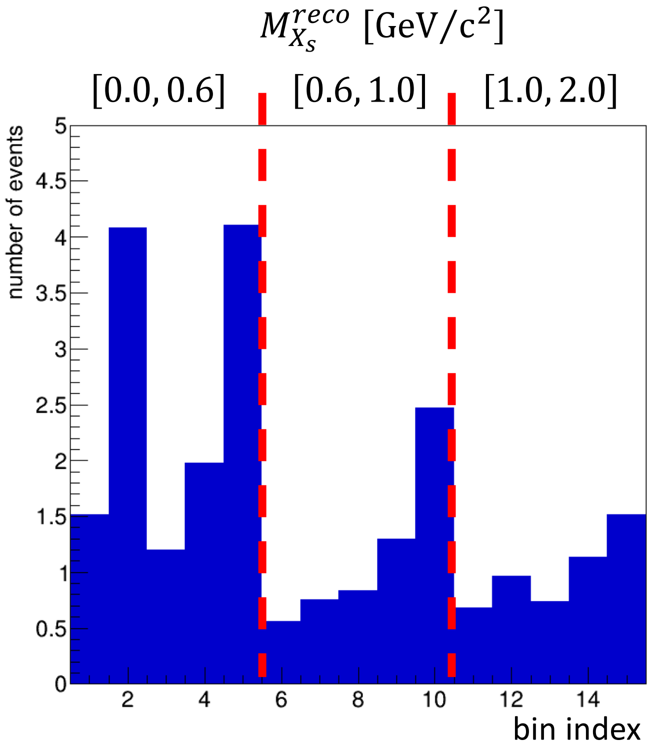
	mass region [GeV]		
Source	$0.0 < M_{X_s}^{\text{true}} < 0.6$	$0.6 < M_{X_s}^{\text{true}} < 1.0$	$1.0 < M_{X_s}^{\text{true}}$
Background normalization	$^{+1.19}_{-1.14}$	$^{+2.29}_{-2.13}$	$^{+4.01}_{-4.07}$
MC statistics	$^{+0.98}_{-0.77}$	$^{+1.68}_{-1.37}$	$^{+3.93}_{-3.31}$
BR of main B meson decays	$^{+0.30}_{-0.14}$	$^{+0.62}_{-0.49}$	$^{+1.28}_{-0.69}$
$q\bar{q}$ shape	$^{+0.29}_{-0.27}$	$^{+0.24}_{-0.24}$	$^{+0.71}_{-0.69}$
statistical uncertainty	$^{+1.78}_{-1.65}$	$^{+2.98}_{-2.79}$	$^{+5.94}_{-5.68}$

※ $M_{X_s}^{\text{true}}$: mass of X_s recorded by MC information

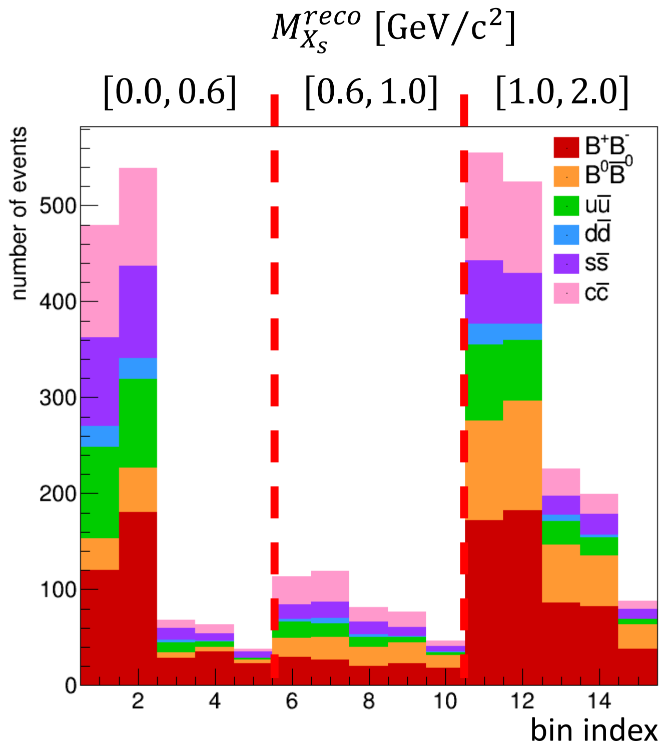
binning and fitting range

- BDT value is mapped into “bin index”





Signal $(\int \mathcal{L}\, dt = 364\, fb^{-1})$



Background $(\int \mathcal{L}\, dt = 364\, fb^{-1})$

- Probability density function is constructed from each bins

$$\mathcal{P} = \underbrace{\prod_{b\in\text{bins}} \text{Pois}(n_b|\nu_b)}_{\text{Poisson distribution for each bin/channel}} \cdot \underbrace{\prod_{p} f_p(a_p|\alpha_p)}_{\text{Constraint term for systematic uncertainty (nuisance parameters)}}$$

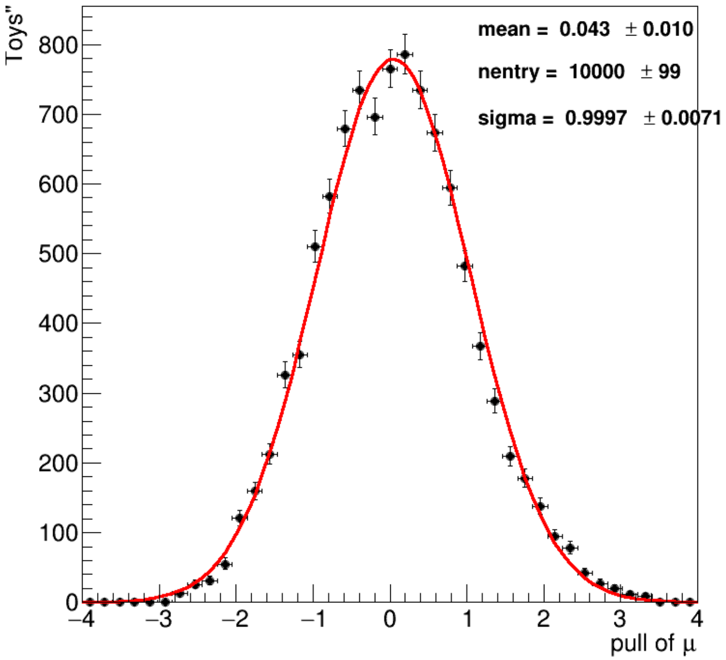
- extended maximum likelihood fit is done
- Fitting parameter μ (signal strength):

a factor relative to the SM expectation

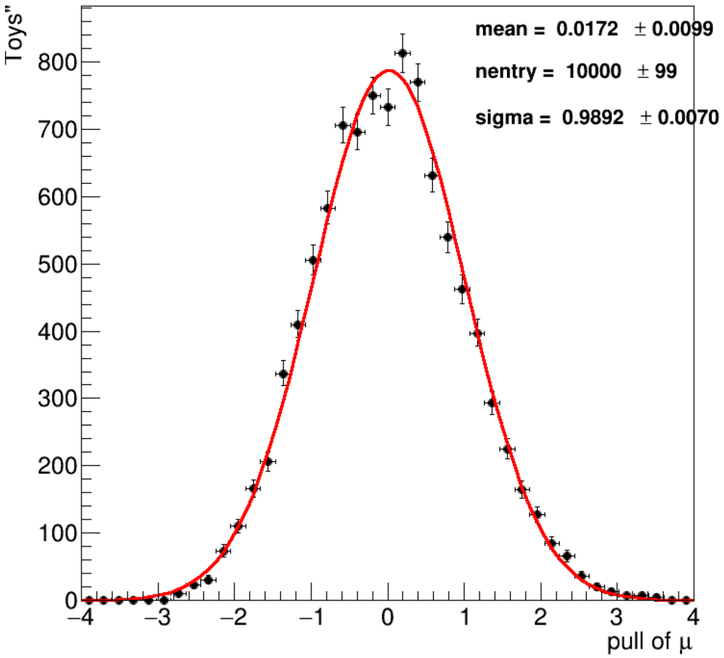
Toy MC study

- Toy MC study is done
  - Generate 10000 toy MC sample with fluctuating nuisance parameter
  - Fit and obtain the MINOS asymmetric error for each toys
  - Pull is calculated from the fitting result:

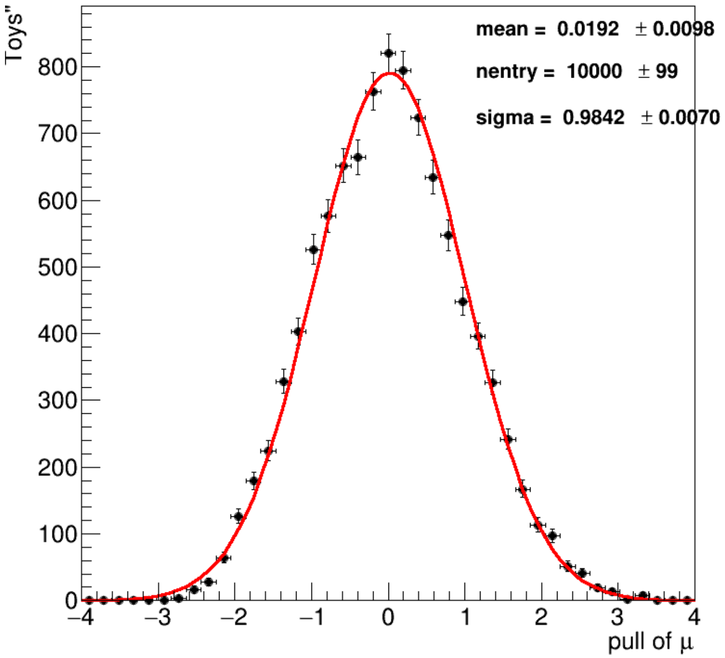
$$\text{pull} = \begin{cases} \dfrac{(\text{true value})-(\text{fit result})}{(\text{positive MINOS error})} & \text{if } (\text{fit result}) \leq (\text{true value}) \\ \dfrac{(\text{fit result})-(\text{true value})}{(\text{negative MINOS error})} & \text{otherwise,} \end{cases}$$



$0.0 < M_{X_s}^{\text{true}} < 0.6$ GeV



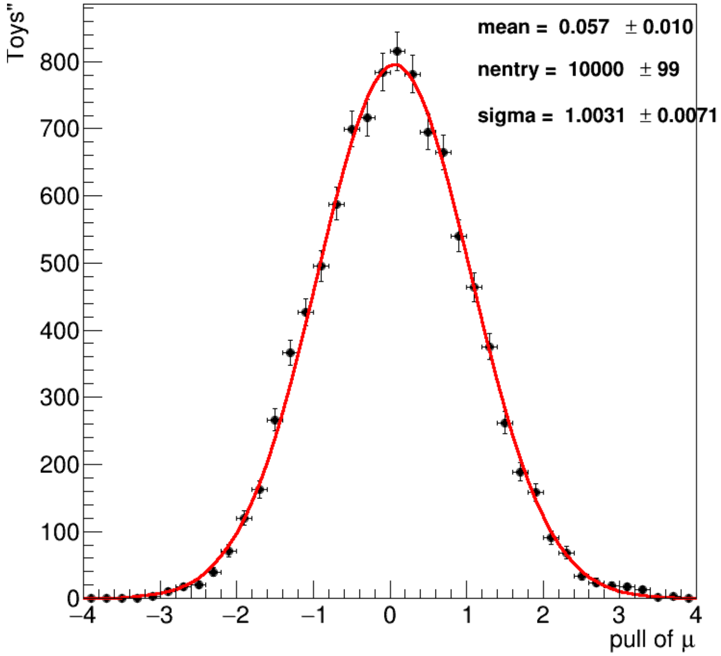
$0.6 < M_{X_s}^{\text{true}} < 1.0$ GeV



$1.0 \text{ GeV} < M_{X_s}^{\text{true}}$

Toy MC study

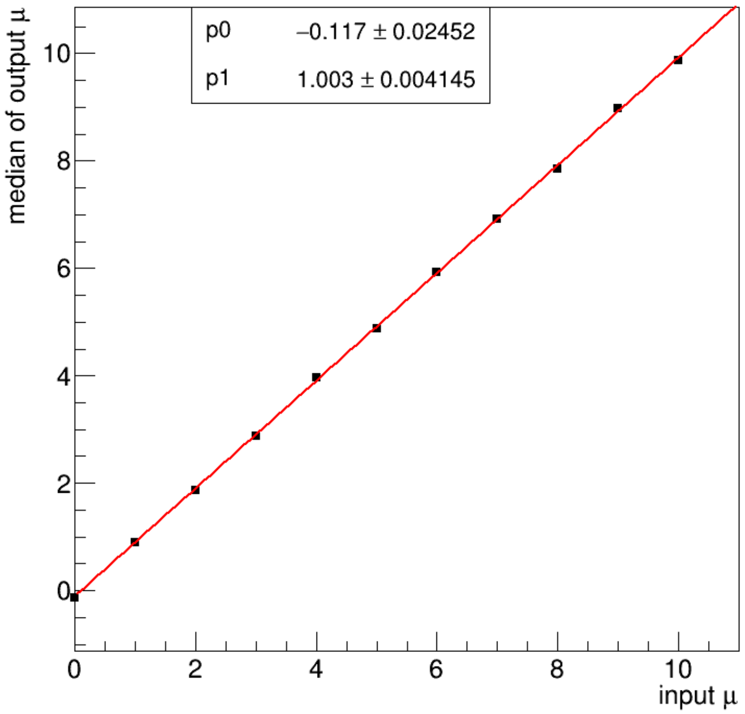
- Toy MC study is done
  - Signal strength through all $M_{X_s}^{\mathrm{true}}$ can be easily obtained: $\mu = \frac{BR_1 \times \mu_1 + BR_2 \times \mu_2 + BR_3 \times \mu_3}{BR}$
  - Pull is calculated with the signal strength through all $M_{X_s}^{\mathrm{true}}$ region



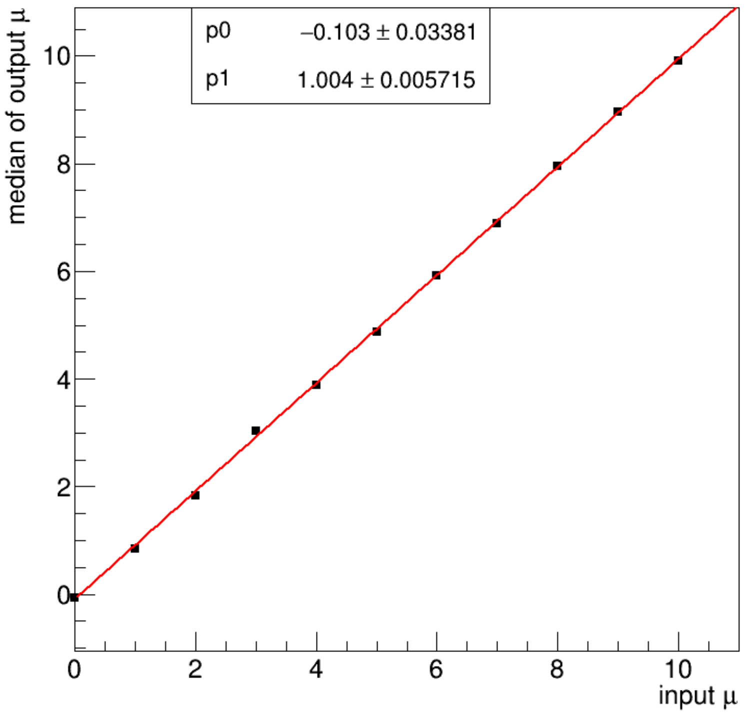
all $M_{X_s}^{\mathrm{true}}$

Linearity Test

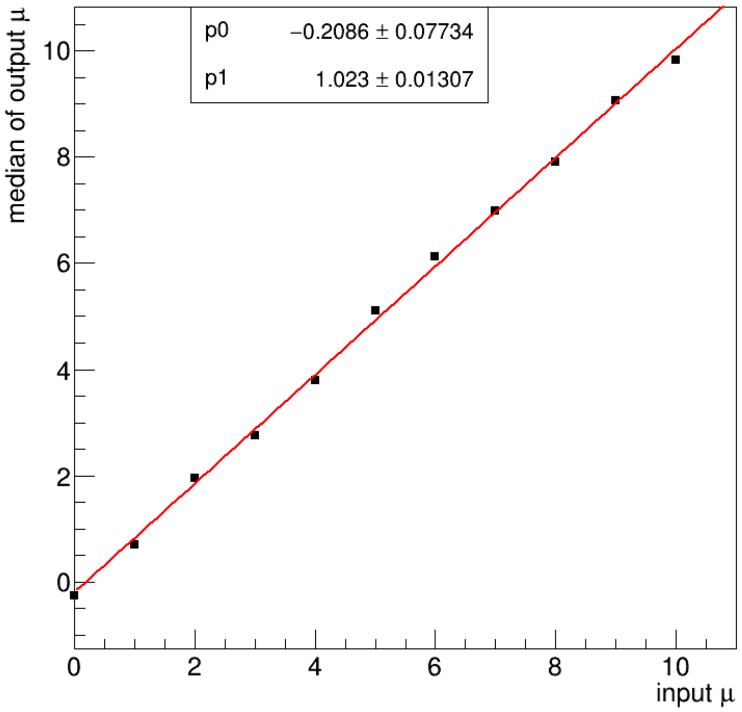
- Linearity is done
  - do toy MC study for different input μ values
  - check the output μ distributions and median value is selected, because error is asymmetric
  - Some bias can be found. This fitter bias is included in systematic uncertainty



$0.0 < M_{X_s}^{\text{true}} < 0.6$ GeV



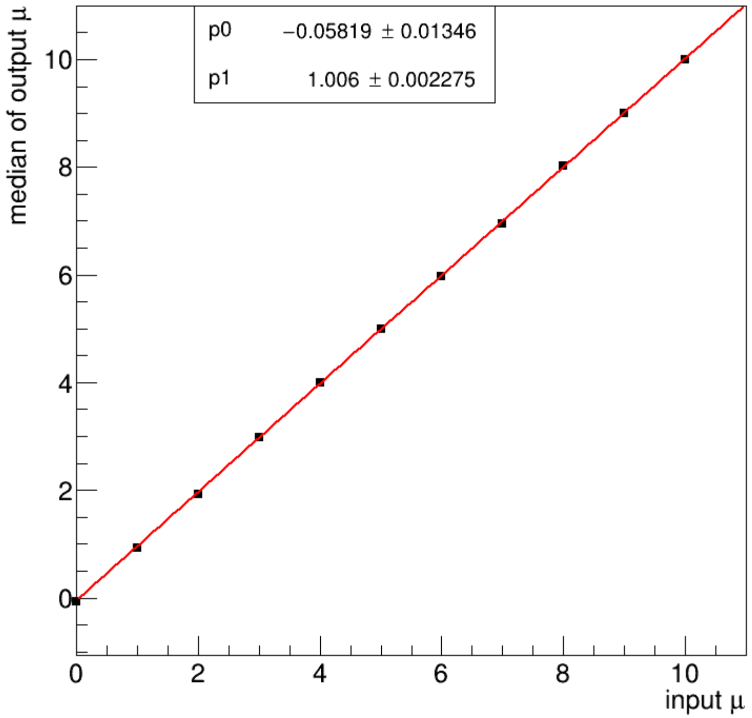
$0.6 < M_{X_s}^{\text{true}} < 1.0$ GeV



1.0 GeV $< M_{X_s}^{\text{true}}$

Linearity Test

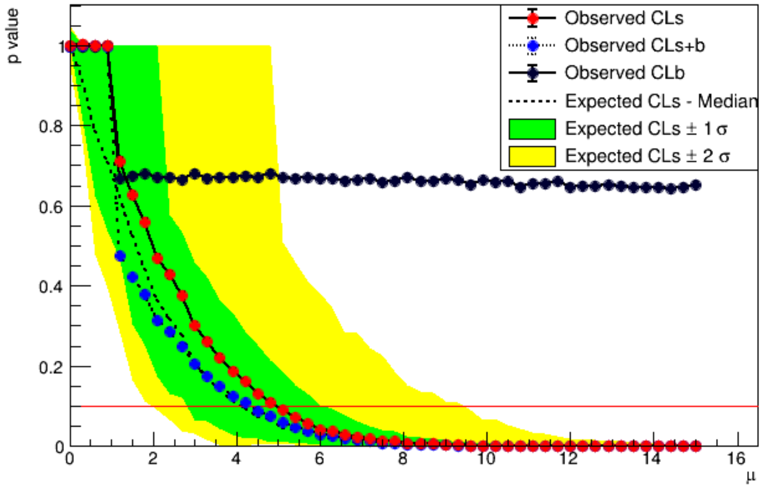
- Linearity is done
  - Signal strength through all $M_{X_s}^{\mathrm{true}}$ can be easily obtained: $\mu = \frac{BR_1 \times \mu_1 + BR_2 \times \mu_2 + BR_3 \times \mu_3}{BR}$
  - The linearity test is done with the signal strength through all $M_{X_s}^{\mathrm{true}}$ region
  - Some bias can be found. This fitter bias is included in systematic uncertainty



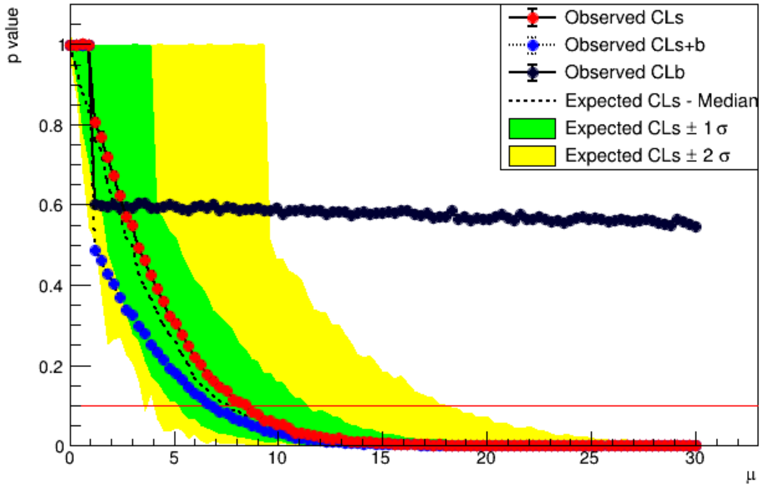
$0.6 < M_{X_s}^{\mathrm{true}} < 1.0\ \mathrm{GeV}$

Upper Limit

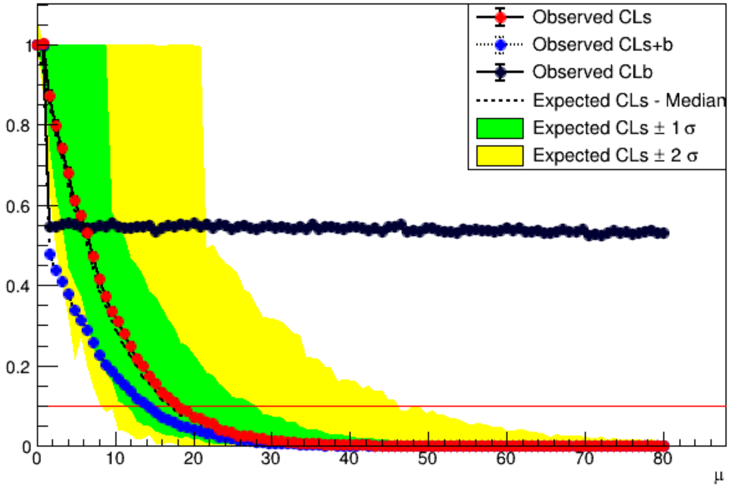
- Upper limit of branching ratio is calculated by CLs method (with $364\ fb^{-1}$)
  - MC sample is used to calculate the upper limit of the branching ratio



$0.0 < M_{X_s}^{\text{true}} < 0.6\ \text{GeV}/\text{c}^2$



$0.6 < M_{X_s}^{\text{true}} < 1.0\ \text{GeV}/\text{c}^2$



$1.0\ \text{GeV}/\text{c}^2 < M_{X_s}^{\text{true}}$

$$UL(\mu) = \begin{cases} 5.0 & \left(0.0 < M_{X_s}^{\text{true}} < 0.6\ \ \text{GeV}/\text{c}^2\right), \\ 8.5 & \left(0.6 < M_{X_s}^{\text{true}} < 1.0\ \ \text{GeV}/\text{c}^2\right), \\ 18.1 & \left(1.0\,\text{GeV}/\text{c}^2 < M_{X_s}^{\text{true}}\right). \end{cases}$$

at 90% confidence level

convert into the BR

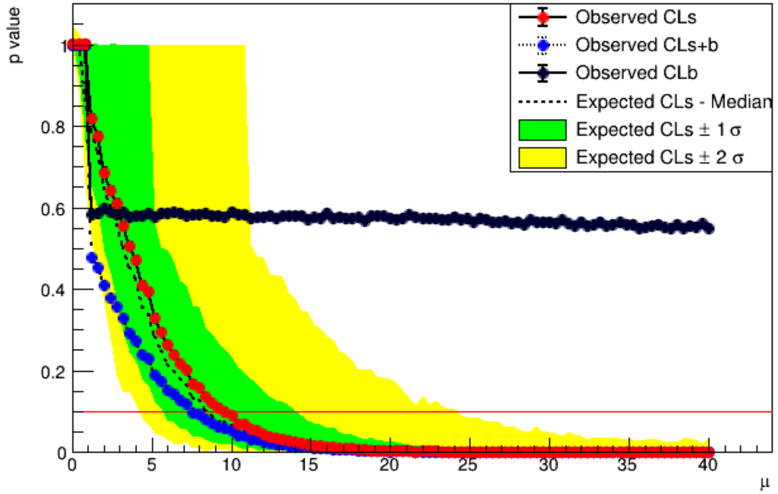
$$UL(B \rightarrow X_s \nu \bar{\nu}) = \begin{cases} 2.4 \times 10^{-5} & \left(0.0 < M_{X_s}^{\text{true}} < 0.6\ \ \text{GeV}/\text{c}^2\right), \\ 7.2 \times 10^{-5} & \left(0.6 < M_{X_s}^{\text{true}} < 1.0\ \ \text{GeV}/\text{c}^2\right), \\ 28.3 \times 10^{-5} & \left(1.0\,\text{GeV}/\text{c}^2 < M_{X_s}^{\text{true}}\right). \end{cases}$$

at 90% confidence level

※ $M_{X_s}^{\text{true}}$: mass of X_s recorded by MC information

Upper Limit

- Upper limit of branching ratio is calculated by CLs method (with $364\ fb^{-1}$)
  - MC sample is used to calculate the upper limit of the branching ratio



all $M_{X_s}^{\mathrm{true}}$

$UL(\mu) = 9.6 \quad \left(\mathrm{all}\ M_{X_s}^{\mathrm{true}}\ \mathrm{region}\right)$

at 90% confidence level

convert into the BR

$UL(B \to X_s \nu\bar{\nu}) = 27.9 \times 10^{-5} \quad \left(\mathrm{all}\ M_{X_s}^{\mathrm{true}}\ \mathrm{region}\right)$

at 90% confidence level

※ $M_{X_s}^{\mathrm{true}}$: mass of X_s recorded by MC information

Conclusion

- $B \to X_s \nu \bar{\nu}$ is interesting decay because it can give a clue for several new physics
    - leptoquark *[PhysRevD.95.035027]*
    - invisible light scalar *[Eur. Phys. J. C (2017) 77: 650]*
    - fermion dark matter *[arXiv:2405.06742]*
- Several selection, including MVA, are used to suppress background
- Background normalization and MC statistic are dominant systematic sources
- Most of analysis procedures are done
    - With $364\ fb^{-1}$, MC expectation is

$$UL(B \to X_s \nu \bar{\nu}) = \begin{cases} 2.4 \times 10^{-5} & \left(0.0 < M_{X_s}^{\mathrm{true}} < 0.6\ \ \mathrm{GeV/c^2}\right), \\ 7.2 \times 10^{-5} & \left(0.6 < M_{X_s}^{\mathrm{true}} < 1.0\ \ \mathrm{GeV/c^2}\right), \\ 28.3 \times 10^{-5} & \left(1.0\,\mathrm{GeV/c^2} < M_{X_s}^{\mathrm{true}}\right). \end{cases}$$

$$UL(B \to X_s \nu \bar{\nu}) = 27.9 \times 10^{-5} \quad \left(\text{all } M_{X_s}^{\mathrm{true}} \text{ region}\right)$$

at 90% confidence level

Backup