AEPSHEP 2022
05-18 OCTOBER 2022,
Pyeongchang, SOUTH KOREA

Measurement of the Branching Ratio of $B^0_{(s)}$ mesons to $\mu^+\mu^-(\gamma)$ at LHCb

Flavour Physics

Phys. Rev. D105 (2022) 012010

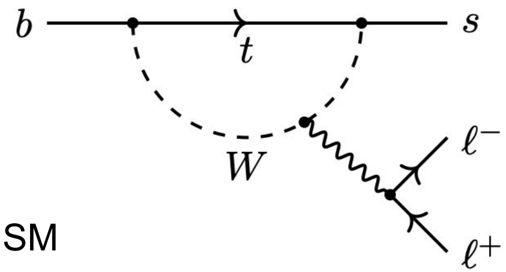
Phys. Rev. Lett. 128 (2022) 041801

Jayashri on behalf of Group B

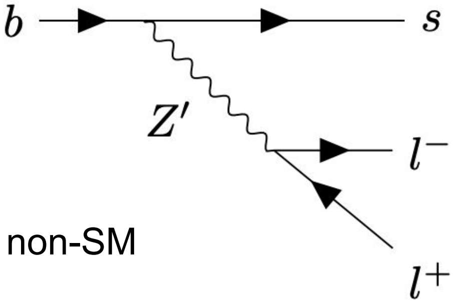
Group B: Quentin, Jiwoong, Laszlo, Lakshmi, Anežka, Aravind, Mike, Eszter, Ashish, Jinyoung, Francesco, Ritik, Madeeha, Jonas, Malinda

Introduction

- Flavour changing neutral currents (FCNC) forbidden at tree level in the Standard Model (SM)
- $b \rightarrow sll$ transition can only occur via loop → rare decays



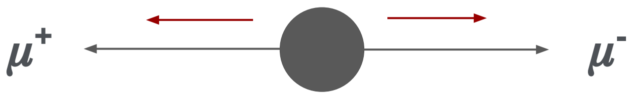
SM



non-SM

- Process could be mediated by **new virtual particles** → probe for New Physics (NP)
  - Z' gauge boson, leptoquarks, non-SM Higgs

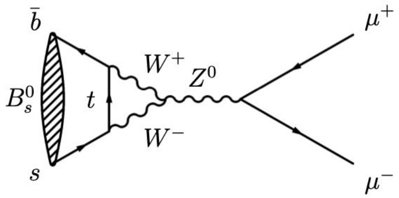
Indirect searches for NP using precision measurements of FCNC processes.



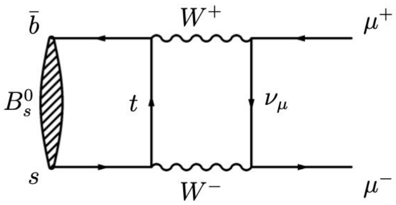
$B^0_{(s)} \rightarrow \mu^+\mu^-$ Decay

- $B^0_{(s)}$ decays to 2 muons: **loop** and **helicity suppressed**

Penguin



Box



- Effective field theory (EFT) gives SM prediction for branching ratios (BR)

$$\mathscr{B}(B^0_s \rightarrow \mu^+\mu^-)_{\mathrm{SM}} = (3.66 \pm 0.14) \times 10^{-9}$$

$$\mathscr{B}(B^0 \rightarrow \mu^+\mu^-)_{\mathrm{SM}} = (1.03 \pm 0.05) \times 10^{-10}$$

Phys. Rev. Lett. 112, 101801 (2014)

JHEP10(2019)232

- Previous measurement:
  - LHCb + ATLAS + CMS combined measurement at 95% CL, consistent with the SM within 2σ

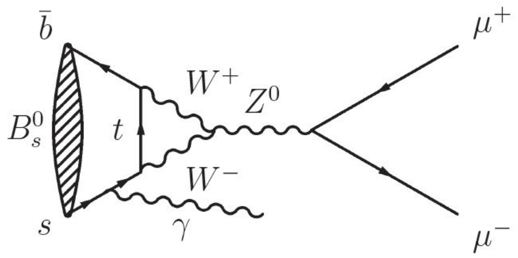
$$\mathcal{B}(B^0_s \rightarrow \mu^+\mu^-) = \left(2.69^{+0.37}_{-0.35}\right) \times 10^{-9}$$

$$\mathcal{B}(B^0 \rightarrow \mu^+\mu^-) < 1.9 \times 10^{-10}$$

ATLAS-CONF-2020-049

Analysis Goals and $B^0_s \to \mu^+\mu^-\gamma$

- Measure BRs for
  - $B^0 \to \mu^+\mu^-$
  - $B^0_s \to \mu^+\mu^-$ *
  - $B^0_s \to \mu^+\mu^-\gamma$

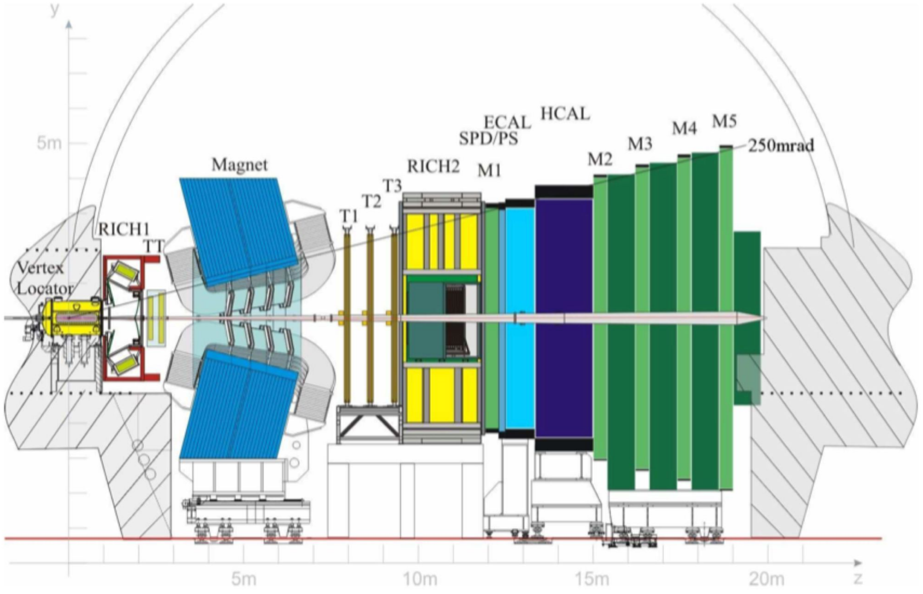


- B^0_s decays to 2 muons can be accompanied by the emission of photon**
  - Not helicity suppressed
- Sensitivity to initial-state radiation (ISR) from the valence quarks at high dimuon mass
  - SM BR prediction $O(10^{-10})$
  - Final-state radiation (FSR) included in B^0_s BR

*The paper also reports the measurement of the effective lifetime for $B^0_s \to \mu^+\mu^-$ which is not included in this talk.
**Note that B^0 ISR is negligible due to additional CKM suppression.

LHCb Detector

- **Forward-spectrometer** specialized on *b* and *c* decays, located at the LHC interaction point 8
- **Vertex Locator at interaction region**
  - Only 7 mm distance to beam
  - Excellent impact parameter resolution to identify secondary vertices
  - Decay-time resolution < 50 fs (decay-time measurement)
- **Ring-Imaging-Cherenkov detectors** for particle identification (relevant for normalization channels)
- Tracking system with momentum resolution < 1.0% → **good mass resolution**
- Muon system to identify and trigger muons



JINST 3 (2008) S08005

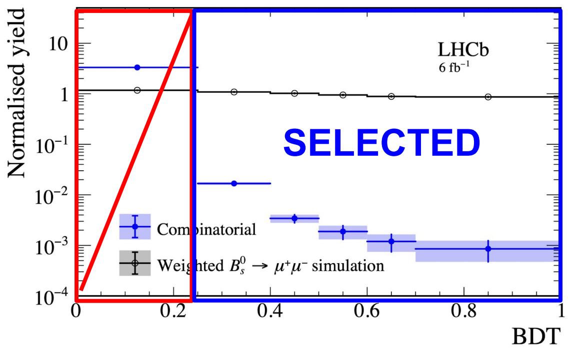
Signal Selection

Signal Region	Control Region (For normalization)	
$B^0_{(s)} \rightarrow \mu^+\mu^-(\gamma)$	$B^0_s \rightarrow hh$[2]	$B^+ \rightarrow J/\psi$[3] $K^+(\varphi)$
Trigger: Number of muon > 0	Triggered independently	# of muon > 0
Muon Selection • 0.25 < pT < 40 GeV • p < 500 GeV • Inconsistent with PV • Muon identification	Replacing muon selection to hadron selection	Same muon selection as signal
B^0_s Selection • Decay time < 13.25ps • pT> 0.5 GeV • Originated at least one PV • J/ψ veto[1] • 4900 ≤ M(μμ) ≤ 6000 MeV	Same but no J/ψ veto	• Same but M(μμ) close to J/ψ mass • Kaon inconsistent with PV

*Additional loose BDT based cut is applied to all regions

[1] J/ψ veto : |M(μμ) - M(J/ψ)| > 30 MeV [2] h = K or π [3] J/ψ decays to muon pair

Selection using boosted decision tree (BDT)

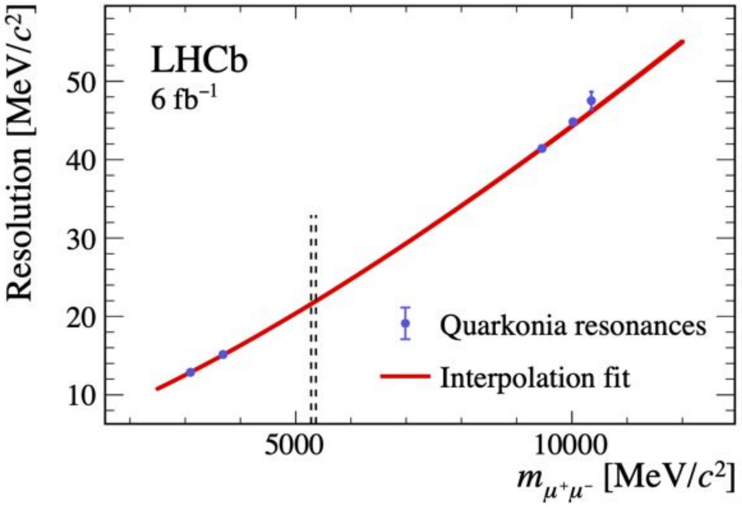


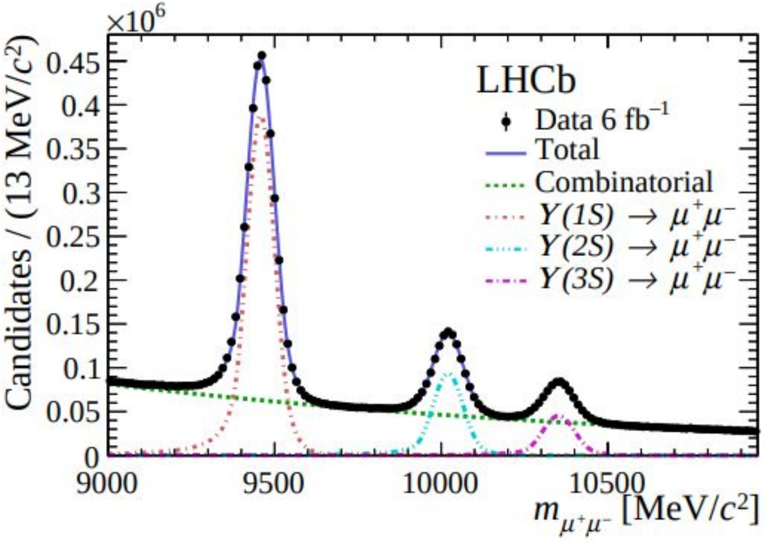
- Discriminating combinatorial background: *bbbar →μμ*
- Optimized using MC samples (*B^0_s→μμ vs bbbar →μμX*)
- Input features: topological variables*

*Phys. Rev. Lett. 128 (2022) 041801

Signal Calibration: Mass Shape

- Signal described with a double-sided Crystal-Ball function
- **Mean** of $B^0_s \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ signal peaks calibrated with $B^0_s \rightarrow K^+K^-$ and $B^0 \rightarrow K^+\pi^-$ data
- **Resolution** interpolated from mass fits to *cc* and *bb* resonances
- **Tail parameters** extracted from simulation





Relative BR Measurement

$$N_{sig} = N_{Tot} f_{sig} B_{sig} \epsilon_{sig}$$

- BR based on
  - Total number of events N_{Tot}
  - Signal yield N_{sig}
  - Selection efficiency ϵ_{sig}
  - Hadronization probability f_{sig}
- **Large uncertainties cancel out in the ratio, e.g., $\sigma(pp \to bb)$ and luminosity**

$$\frac{\mathcal{B}_{sig}}{\mathcal{B}_{norm}} = \frac{f_{norm}}{f_{sig}} \frac{\varepsilon_{norm}}{\varepsilon_{sig}} \frac{N_{sig}}{N_{norm}}$$

Normalization

$$\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-(\gamma)) = \frac{f_{\rm norm}}{f_{\rm sig}} \frac{\varepsilon_{\rm norm}}{\varepsilon_{\rm sig}} \frac{N_{\rm sig}}{N_{\rm norm}} \mathcal{B}_{\rm norm} = \alpha_{\rm sig} N_{\rm sig}$$

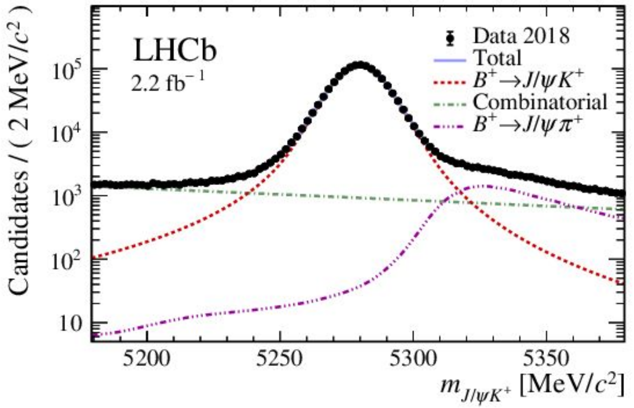
α_{sig} → single event sensitivity

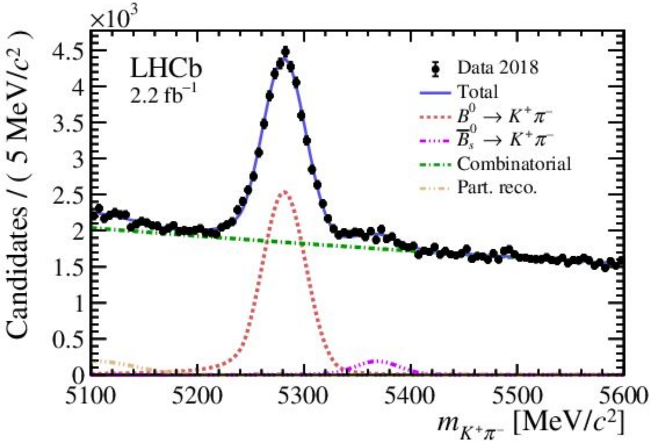
- BRs measured relative to well known normalization channels with similar selections
    - $B^+ \to J/\psi(\to\mu^+\mu^-)\ K^+$
    - $B_0 \to K^+\ \pi^-$
- Extraction of normalization channel yields via maximum likelihood fit
- **Efficiencies calculated for signal and normalization channels**
- External input for hadronization ratios f_s/f_d (f_s/f_u) → large systematic

$$f_s/f_d\ (13\,\text{TeV}) = 0.254 \pm 0.008$$

YIELD

Period	$B^+ \to J/\psi K^+$	$B^0 \to K^+\pi^-$
Run 1	$(11.259 \pm 0.015) \times 10^5$	$(14.05 \pm 0.26) \times 10^3$
Run 2	$(36.072 \pm 0.026) \times 10^5$	$(7.99 \pm 0.10) \times 10^4$





Backgrounds

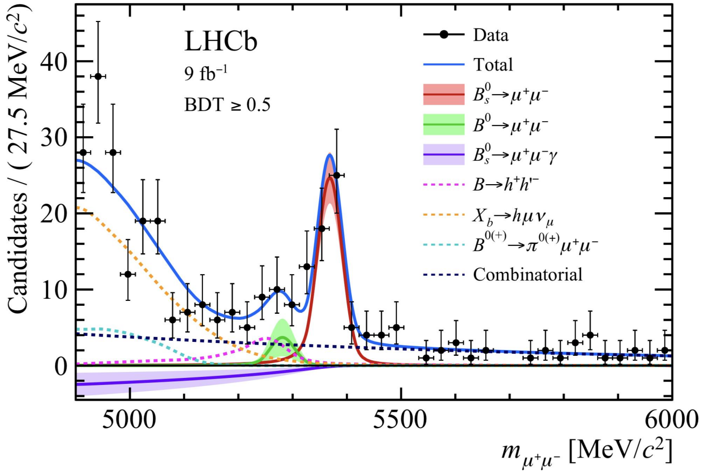
Physical

- One or more misidentified final state particles (protons, pions, and kaons)
- Use stringent PID requirement to suppress it

- One particle is not reconstructed
  - $B^{(0)+} \rightarrow \pi^{(0)+} \mu^+ \mu^-$
- One hadron is misidentified as muon
  - $X_b \rightarrow h \mu \nu_\mu$
- Two hadrons are misidentified as muons
  - $B^0_{(S)} \rightarrow h^+ h'^-$
- **Dominant in low-mass region**

Combinatorial

- Random combination of two muons not originating from a common decay
- Suppress by forming *B* candidate using topology differences and relative isolation of the muon tracks
- **Dominant in high-mass region**



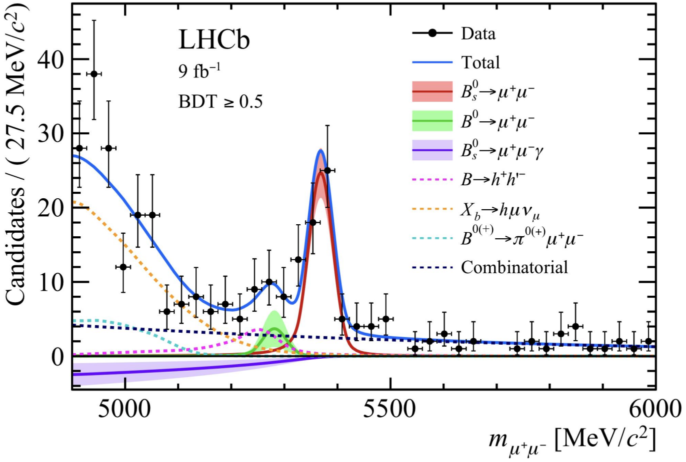
Measurement of Signal BR

Signal yields extracted by **unbinned extended maximum-likelihood fit** with constraint background components

Resulting BRs

$$\begin{aligned}\mathcal{B}(B_s^0 \to \mu^+\mu^-) &= \left(3.09\,^{+0.46}_{-0.43}\,^{+0.15}_{-0.11}\right)\times 10^{-9},\\ \mathcal{B}(B^0 \to \mu^+\mu^-) &= \left(1.20\,^{+0.83}_{-0.74} \pm 0.14\right)\times 10^{-10},\\ \mathcal{B}(B_s^0 \to \mu^+\mu^-\gamma) &= (-2.5 \pm 1.4 \pm 0.8)\times 10^{-9} \text{ with } m_{\mu\mu} > 4.9\,\mathrm{GeV}/c^2.\end{aligned}$$

- Stat. by repeating the fit with all nuisance parameters fixed
- Syst. by subtracting in quadrature the stat. uncertainties from the total



Main contribution to the systematic uncertainties

$\mathcal{B}(B_s^0 \to \mu^+\mu^-)$: f_s/f_d (3%)

$\mathcal{B}(B^0 \to \mu^+\mu^-)$: $B^0_s \to h^+h'^-$ and semileptonic b-hadron background (9%)

Limit Setting of Signal BR

Significance

- More than 10σ for $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$
- 1.7σ for $\mathcal{B}(B^0 \to \mu^+\mu^-)$
- 1.5σ for $\mathcal{B}(B_s^0 \to \mu^+\mu^-\gamma)$

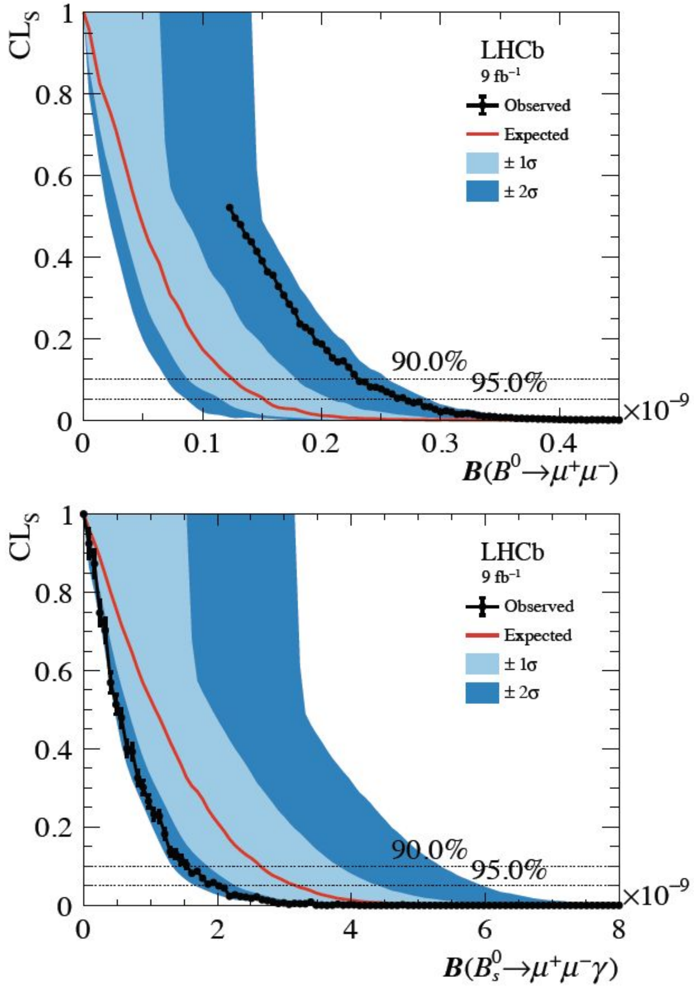
*No evidence → **"Limit"***

CL$_s$ method with a one-sided test statistic is used for determining upper limits.

Limit on the BRs

$$\mathcal{B}(B^0 \to \mu^+\mu^-) \quad < \quad 2.3\,(2.6) \times 10^{-10},$$
$$\mathcal{B}(B_s^0 \to \mu^+\mu^-\gamma) \quad < \quad 1.5\,(2.0) \times 10^{-9} \text{ with } m_{\mu\mu} > 4.9\,\mathrm{GeV}/c^2,$$

at 90% (95%) CL.



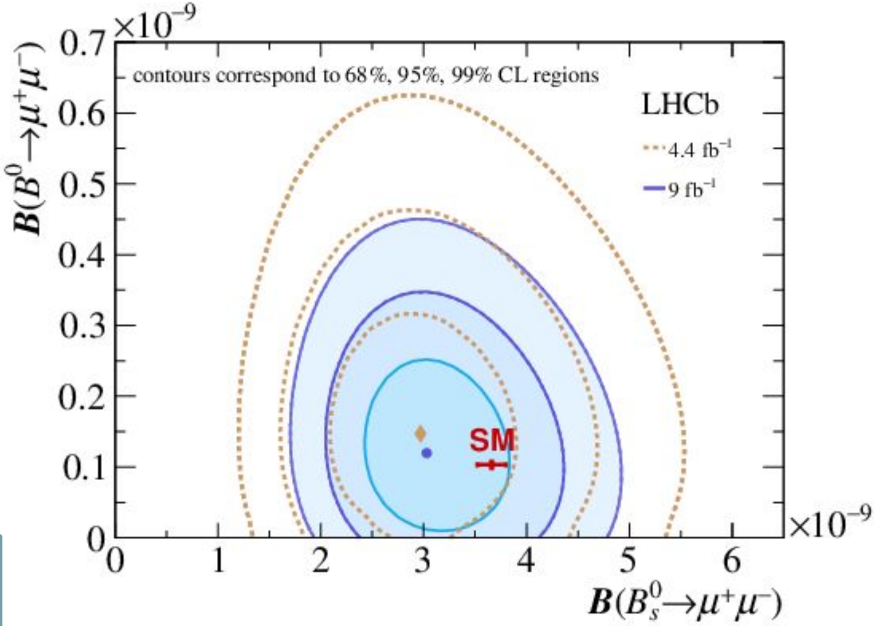
Limit Setting of Signal BR

Significance

- More than 10σ for $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$
- 1.7σ for $\mathcal{B}(B^0 \to \mu^+\mu^-)$
- 1.5σ for $\mathcal{B}(B_s^0 \to \mu^+\mu^-\gamma)$

*No evidence → **"Limit"***

CL$_s$ method with a one-sided test statistic is used for determining upper limits.



In agreement with the SM predictions

Limit on the BRs

$$\mathcal{B}(B^0 \to \mu^+\mu^-) < 2.3\,(2.6) \times 10^{-10},$$
$$\mathcal{B}(B_s^0 \to \mu^+\mu^-\gamma) < 1.5\,(2.0) \times 10^{-9} \text{ with } m_{\mu\mu} > 4.9\,\mathrm{GeV}/c^2,$$

at 90% (95%) CL.

Conclusion

- Rare decays of $B^0_{(s)}$ to two muons → powerful probe for NP
- Most precise single experiment measurement
  - BR of $B^0_s \to \mu^+\mu^-$ measured (> 10σ significance)
  - $$\mathcal{B}(B^0_s \to \mu^+\mu^-) = \left(3.09\,^{+\,0.46}_{-\,0.43}\,^{+\,0.15}_{-\,0.11}\right) \times 10^{-9},$$
  - Improved limits set for $B^0 \to \mu^+\mu^-$ and $B^0_s \to \mu^+\mu^-\gamma$
  - $$\mathcal{B}(B^0 \to \mu^+\mu^-) = \left(1.20\,^{+\,0.83}_{-\,0.74} \pm 0.14\right) \times 10^{-10},$$ $$\mathcal{B}(B^0_s \to \mu^+\mu^-\gamma) = (-2.5 \pm 1.4 \pm 0.8) \times 10^{-9} \text{ with } m_{\mu\mu} > 4.9\,\text{GeV}/c^2.$$


ALPENSIA
Convention Centre
AEPSHEP 2022
Asia-Europe-Pacific School of High Energy Physics
2022. 10. 05 ~ 2022. 10. 18
Pyeongchang, SOUTH KOREA

Back-up

BR Calculation

Time integrated BR:

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = \left[\frac{1 + A_{\Delta\Gamma_s}^{\mu\mu} y_s}{1 - y_s^2}\right] \mathcal{B}(B_s^0 \to \mu^+\mu^-)_{t=0}$$

where

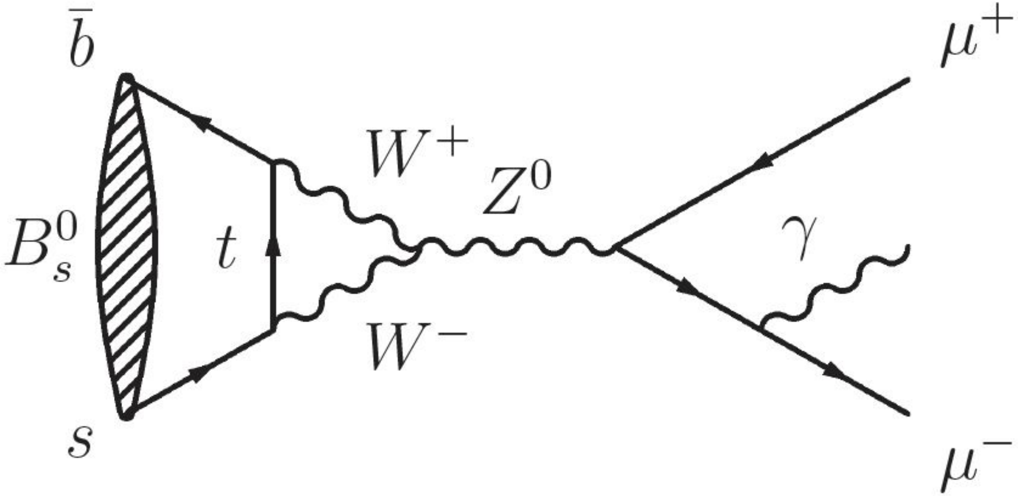
$$y_s \equiv \Delta\Gamma_s/(2\Gamma_s) = 0.065 \pm 0.003$$

$$A_{\Delta\Gamma_s}^{\mu\mu} \equiv -2\Re(\lambda)/(1 + |\lambda|^2), \text{ with } \lambda = (q/p)(A(\overline{B}_s^0 \to \mu^+\mu^-)/A(B_s^0 \to \mu^+\mu^-))$$

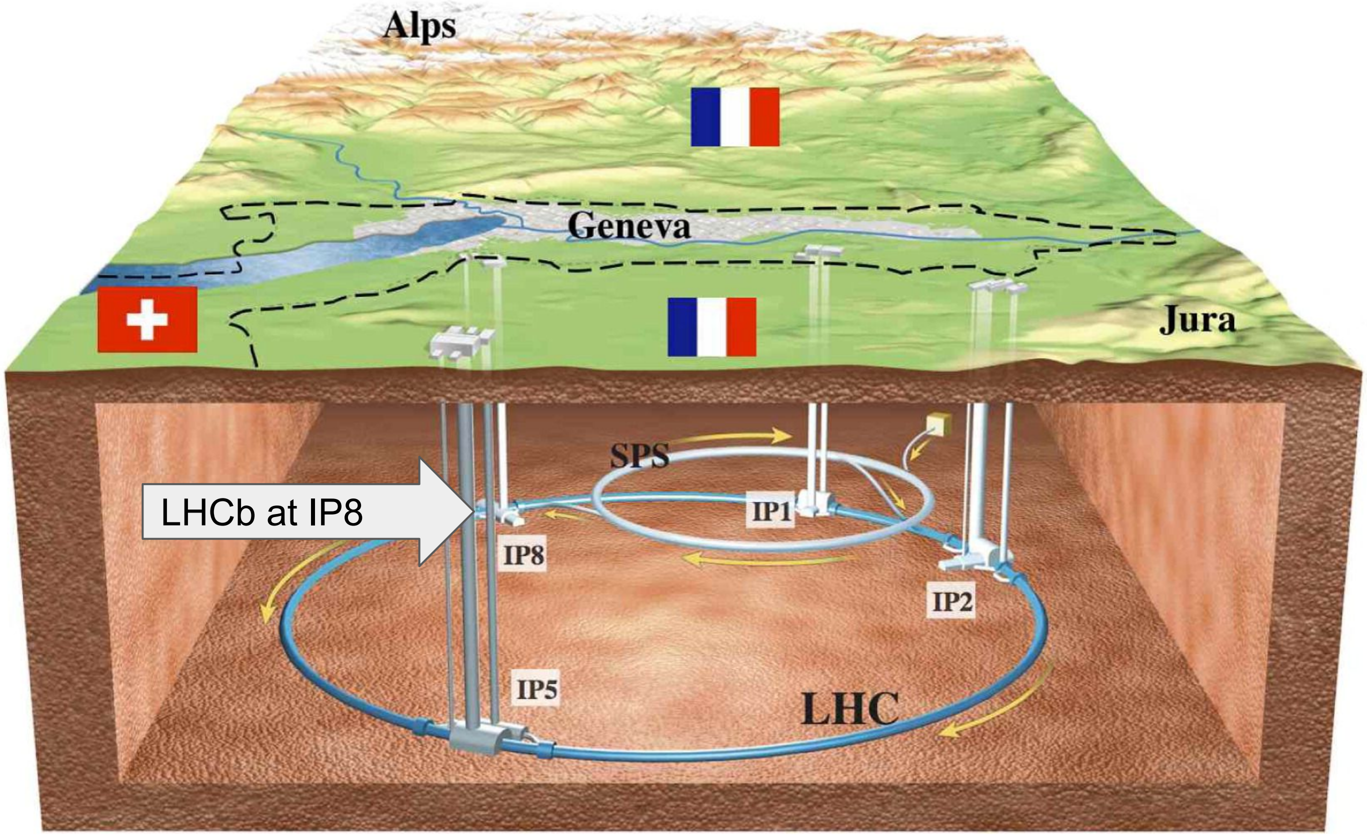
For the SM, $A_{\Delta\Gamma_s}^{\mu\mu}$ = 1 as only CP odd eigenstate of $B^0_{(s)}$ decays two muons $\longrightarrow$ $\mathscr{B}(B_s^0 \to \mu^+\mu^-)_{SM} = (3.66 \pm 0.14) \times 10^{-9}$
In presence of NP, it can assume any value from -1 to +1.

Final-State Radiation (FSR)

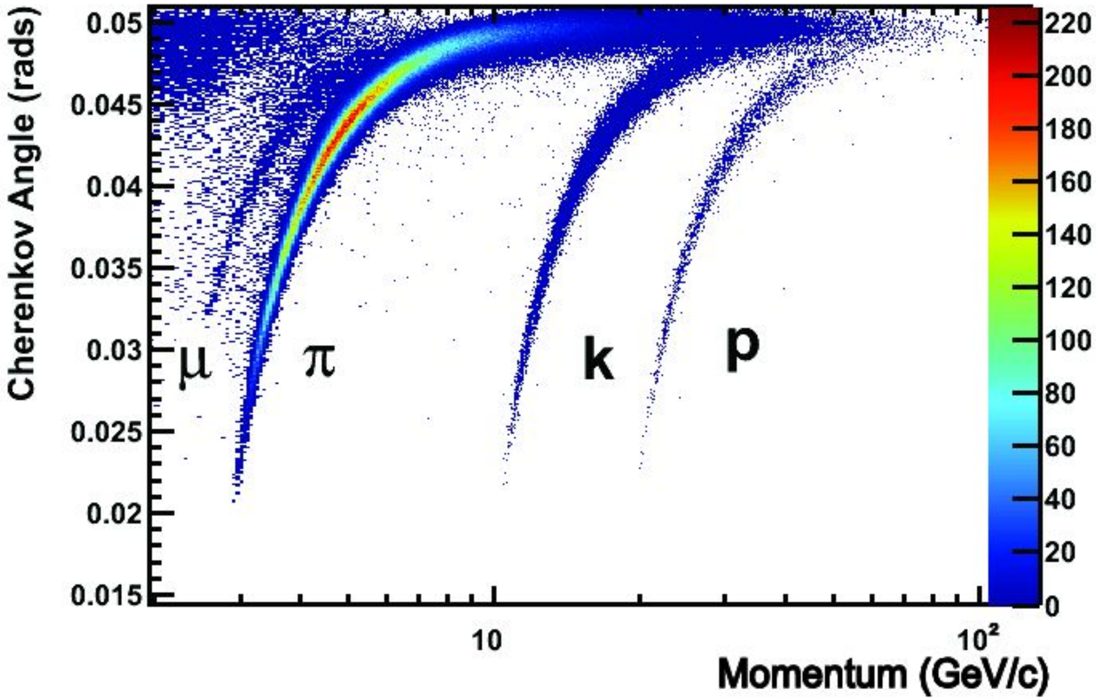
- Included experimentally in $B^0_s \to \mu^+\mu^-$ in the description of the radiative tail in its mass distribution



Detector



Detector

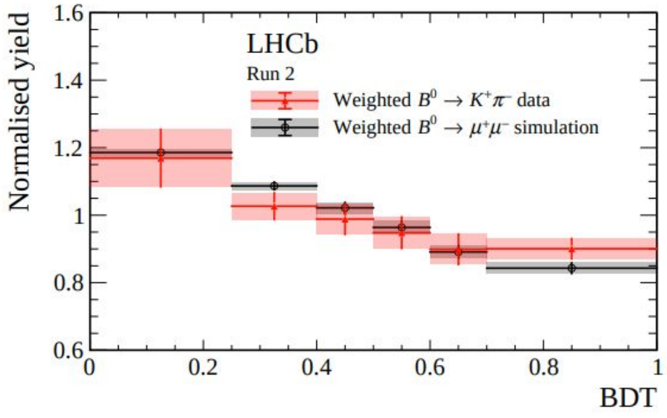

Cherenkov Angle (rads)
0.05
0.045
0.04
0.035
0.03
0.025
0.02
0.015
μ
π
k
p
10
10²
Momentum (GeV/c)
220
200
180
160
140
120
100
80
60
40
20
0

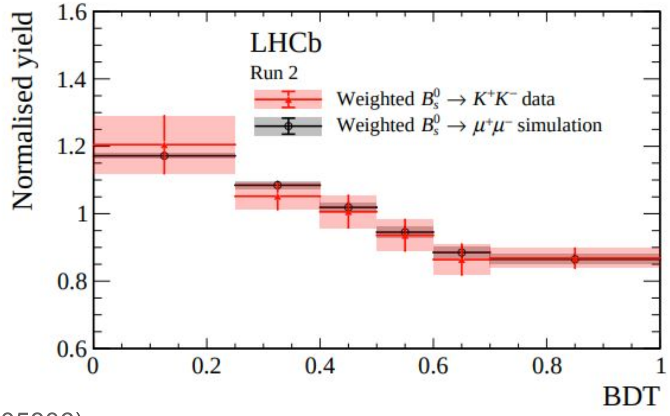
Signal Calibration: BDT

- Events categorised into 6 BDT bins
- Calibration of signal BDT output done using data-corrected simulation*
- Corrections applied to:
  - B meson kinematics and detector occupancy
  - PID and trigger efficiency
  - BDT-lifetime correlations in $B^0_s \to \mu^+\mu^-$
- Fraction of signal decays in each BDT region:

$$f_{\mathrm{BDT},i} = f^{\mu\mu}_{\mathrm{sim},i} \cdot f^{\mu\mu}_{\mathrm{PID},i} \cdot f^{\mu\mu}_{\mathrm{trig},i} \ (\cdot\ k_i)$$

- Cross-checked on $B^0 \to K^+\pi^-$ and $B^0_s \to K^+K^-$ data
  - Same selection as for the signal, except for trigger and PID requirements





* Reweighting method based on BDT: Extracting w = data/MC using BDT (arXiv:1608.05806)

BDT Calibration

- **Correction to B meson kinematics:**
  - Boosted decision tree classifier trained to align the data and the simulation
  - pT, η, and the χ^2_{IP} of the B candidate used as input variables
  - Weights obtained applied to all simulation samples used for calibration and normalisation
- **Correction from event occupancy (measured as the number of tracks in the event):**
  - Weights determined by comparing the relative number of $B^+ \to J/\psi K^+$ decays in background-subtracted data and simulated samples
- **Corrections to BDT-lifetime correlation for $B^0_s \to \mu^+\mu^-$ decays:**
  - Simulation generated using the mean B^0_s lifetime; but effective lifetime may vary between that of light and the heavy mass eigenstates.
  - Correction evaluated for $A^{\mu\mu}_{\Delta\Gamma s} = -1, 0$ and 1
  - Simulated candidates weighted according to:

$$\omega_j = \frac{\tau_{\mathrm{gen}}}{\tau_{\mu^+\mu^-}} e^{-t_j\left(1/\tau_{\mu^+\mu^-} - 1/\tau_{\mathrm{gen}}\right)}$$

  - Correction factor applied to each BDT region:

$$k_i = \sum_{j=1}^{N_i} \frac{\omega_j}{N_i}$$

Normalization

- Branching ratios measured relative to well known normalization channels
  - $B^+ \to Jpsi(\to\mu^+\mu^-)\ K^+$ [Br. frc. = (6.02 ± 0.17) × 10^{-5}]
  - $B_0 \to K^+ \pi^-$ [Br. frc. = (1.96 ± 0.05) × 10^{-5}]
- Similar selections, slightly modified (Do we want to give details?)
- Extraction of normalization channel yields via maximum likelihood fit
- External input for hadronization ratios fs/fd (fs/fu) -> large systematic
- Carefully estimated reconstruction and selection efficiencies

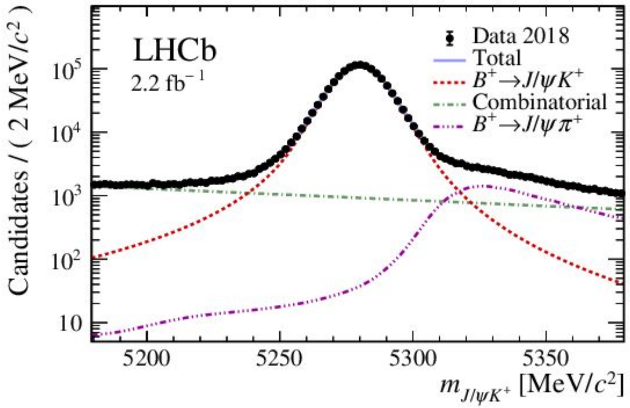
$$\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-(\gamma)) = \frac{f_{\rm norm}}{f_{\rm sig}} \frac{\varepsilon_{\rm norm}}{\varepsilon_{\rm sig}} \frac{N_{\rm sig}}{N_{\rm norm}} \mathcal{B}_{\rm norm} = \alpha_{\rm sig} N_{\rm sig}$$

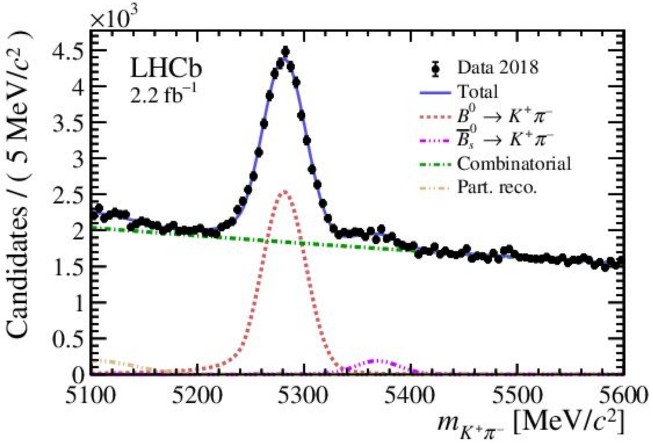
- α_{sig} → single event sensitivity

$$f_s/f_d\ (13\,\mathrm{TeV}) = 0.254 \pm 0.008$$

YIELD

Period	$B^+ \to J/\psi K^+$	$B^0 \to K^+\pi^-$
Run 1	$(11.259 \pm 0.015) \times 10^5$	$(14.05 \pm 0.26) \times 10^3$
Run 2	$(36.072 \pm 0.026) \times 10^5$	$(7.99 \pm 0.10) \times 10^4$





Efficiencies

For signal only

$$\varepsilon_{\mathrm{norm(sig)}} = \varepsilon_{\mathrm{RecSel}} \cdot \varepsilon_{\mathrm{PID}} \cdot \varepsilon_{\mathrm{Trig}} \left(\cdot\ \varepsilon_{\mathrm{BDT}}\right)$$

- Acceptance and reconstruction efficiency
  - Obtained from MC simulation with correction factors applied for simulation modelling
- PID: Particle Identification Efficiency
  - Obtained from high purity sample from data
    - μ from $J/\psi \rightarrow \mu\mu$
    - $K^{\pm}$ from $D^0 \rightarrow K^{\mp} \pi^{\pm}$
    - p from $\Lambda^0 \rightarrow p\pi^-$ and $\Lambda^+_c \rightarrow p K \pi$
- Trigger efficiency

Efficiencies

$\varepsilon_{\mathrm{RecSel}}$		
	Run 1	Run 2
$B_s^0 \to \mu^+\mu^-$	0.0602 ± 0.0003	0.0640 ± 0.0004
$B^0 \to \mu^+\mu^-$	0.0594 ± 0.0003	0.0635 ± 0.0004
$B_s^0 \to \mu^+\mu^-\gamma$	0.0508 ± 0.0003	0.0546 ± 0.0004
$B^0 \to K^+\pi^-$	0.0462 ± 0.0007	0.0500 ± 0.0006
$B^+ \to J/\psi K^+$	0.0290 ± 0.0003	0.0305 ± 0.0003

$\varepsilon_{\mathrm{Trig}}$		
	Run 1	Run 2
$B_s^0 \to \mu^+\mu^-$	$0.9579 \pm 0.0033 \pm 0.0164$	$0.9712 \pm 0.0014 \pm 0.0093$
$B^0 \to \mu^+\mu^-$	$0.9570 \pm 0.0032 \pm 0.0176$	$0.9708 \pm 0.0014 \pm 0.0097$
$B_s^0 \to \mu^+\mu^-\gamma$	$0.9538 \pm 0.0032 \pm 0.0195$	$0.9694 \pm 0.0013 \pm 0.0111$
$B^0 \to K^+\pi^-$	$0.0433 \pm 0.0002 \pm 0.0016$	$0.0727 \pm 0.0002 \pm 0.0020$
$B^+ \to J/\psi K^+$	$0.8810 \pm 0.0040 \pm 0.0080$	$0.9033 \pm 0.0016 \pm 0.0089$

$\varepsilon_{\mathrm{PID}}$		
	Run 1	Run 2
$B_s^0 \to \mu^+\mu^-$	$0.8580 \pm 0.0006 \pm 0.0053$	$0.8822 \pm 0.0003 \pm 0.0039$
$B^0 \to \mu^+\mu^-$	$0.8518 \pm 0.0007 \pm 0.0063$	$0.8759 \pm 0.0004 \pm 0.0046$
$B_s^0 \to \mu^+\mu^-\gamma$	$0.8487 \pm 0.0006 \pm 0.0088$	$0.8785 \pm 0.0003 \pm 0.0064$
$B^0 \to K^+\pi^-$	$0.4741 \pm 0.0049 \pm 0.0010$	$0.5004 \pm 0.0027 \pm 0.0012$
$B^+ \to J/\psi K^+$	1.0096 ± 0.0005	1.00260 ± 0.00018

$\varepsilon_{\mathrm{BDT}}$		
	Run 1	Run 2
$B_s^0 \to \mu^+\mu^-$	0.723 ± 0.006	0.7071 ± 0.0026
$B^0 \to \mu^+\mu^-$	0.720 ± 0.006	0.7036 ± 0.0027
$B_s^0 \to \mu^+\mu^-\gamma$	0.656 ± 0.007	0.6531 ± 0.0035

Cross-Check on the Fragmentation Fraction

To cross-check the ratio of the B_s^0 and B^+ fragmentation fractions and its stability over the data taking, the ratio of $B^+ \to J/\psi K^+$ and $B_s^0 \to J/\psi \phi$ efficiency-corrected yields,

$$\mathcal{R} = \frac{N_{B_s^0 \to J/\psi \phi}}{N_{B^+ \to J/\psi K^+}} \frac{\varepsilon_{B^+ \to J/\psi K^+}}{\varepsilon_{B_s^0 \to J/\psi \phi}} = \frac{f_s}{f_u} \frac{\mathcal{B}(B_s^0 \to J/\psi \phi)}{\mathcal{B}(B^+ \to J/\psi K^+)}, \quad (12)$$

is also measured, following a similar approach to Ref. [82]. The ratios are found to be similar to those measured in Ref. [82], although the two methods explore different kinematic regions. A dependence on the centre-of-mass energy is seen and found to be consistent with Ref. [82] and the combined analysis of Ref. [61], justifying the use of different f_s/f_d values for the Run 1 and Run 2 data samples.

Reweighting with BDT

$$\text{multiplier}_{\text{bin}} = \frac{w_{\text{bin, RD}}}{w_{\text{bin, MC}}}$$

$$\chi^2 = \sum_{\text{leaf}} \frac{(w_{\text{leaf, MC}} - w_{\text{leaf, RD}})^2}{w_{\text{leaf, MC}} + w_{\text{leaf, RD}}}.$$

Basic weighting method

- Counting histogram
- Limitation: reweighting of one variable often bring disagreement in others

Reweighting with ML

- Define loss function χ^2 as a difference of Data-MC
- Find the region that has maximum χ^2 and minimize the difference

Measurement of Signal BR