Highlights from Belle

Sen Jia (Fudan U.)
on behalf of the Belle Collaboration

全国第十八届重味物理与CP破坏研讨会(HFCPV2021)
2021年11月10-14日, 暨南大学

Content

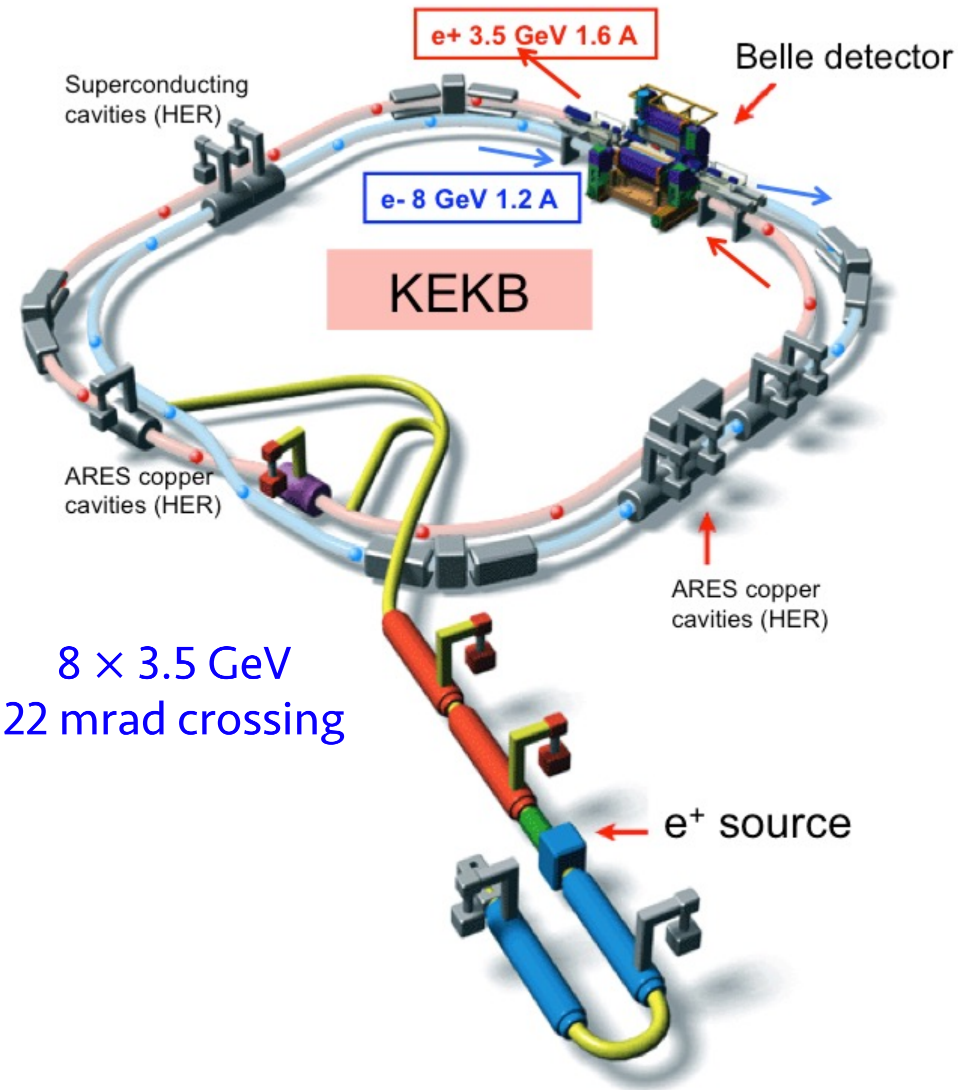
Recent results of B and B_s decays at Belle (this talk):

- Lepton flavor universality: R(D), R(D*)
- $B \to X_u \ell^+ \nu_\ell$ inclusive for $|V_{ub}|$ and differential branching fractions
- B and B_s decays: $B^0 \to K^{*0}\tau^+\tau^-$ and $B_s^0 \to D_s X$

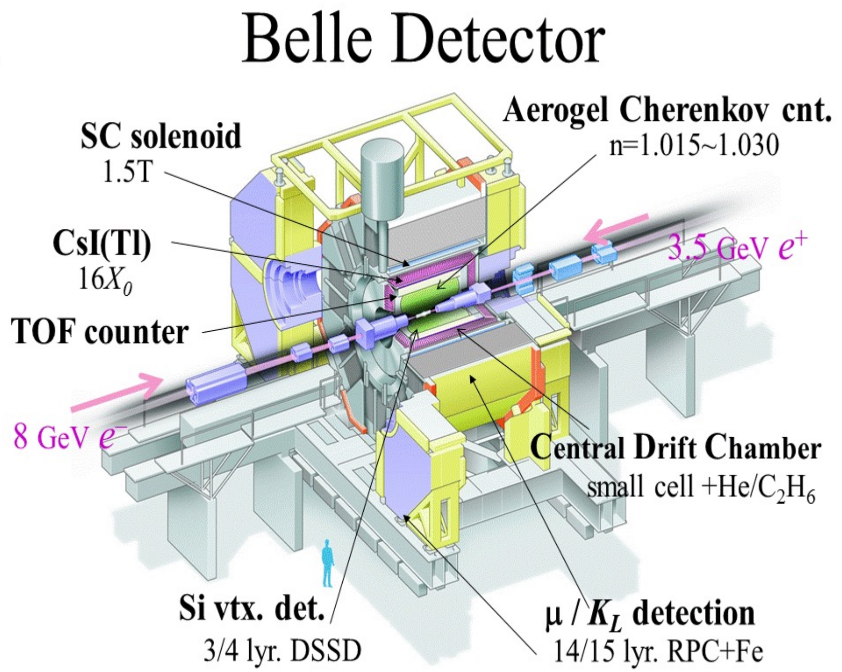
Other topics related to Belle (II)

- Charmed baryons at Belle from Y. B. Li, Nov 12 15:10
- Belle II prospects from X. Y. Zhou, Nov 13 14:30

Belle experiment and data samples



8 × 3.5 GeV
22 mrad crossing



Data taking: 1999 – 2010

On/off/Scan $\Upsilon(nS)$ peaks

Total luminosity: 980 fb^{-1}

772M $B\bar{B}$ events @$\Upsilon(4S)$

Lepton flavor universality: R(D), R(D*)

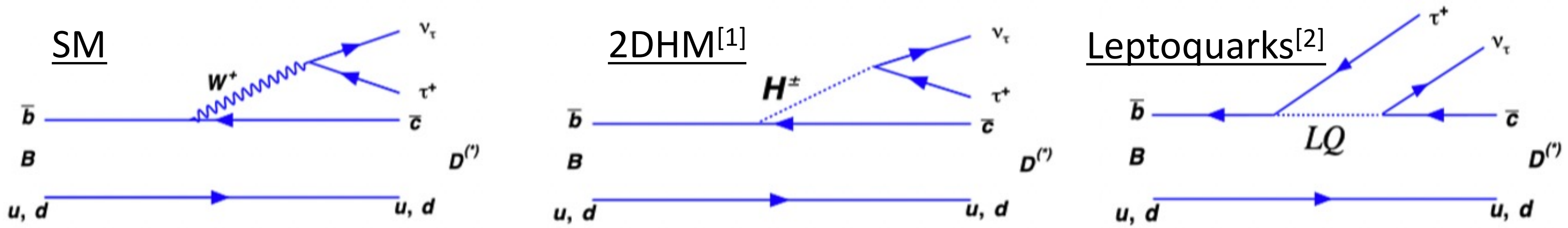
[PRL124, 161803 (2020)]

$\mathcal{R}(D^{(*)})$ in semitauonic B-decays

- Lepton flavor universality (LFU) enforces the equal coupling of the gauge bosons to the three lepton generations.

- Semitauonic decays such as $B \to D^{(*)}\tau\nu_\tau$ are more interesting, given the third generation of lepton family are involved in the transition.

- Sensitive to New Physics (NP): Two Higgs doublets, leptoquarks, etc.



- The presence of NP impact the experimentally observed LFU ratios:

$$\mathcal{R}(\mathcal{D}^{(*)}) \equiv \frac{\mathcal{B}(B \to D^{(*)}\tau\nu_\tau)}{\mathcal{B}(B \to D^{(*)}\ell\nu_\ell)}$$, where ℓ = e and μ

- Beyond SM particles can also alter the kinematic distributions of final state: τ lepton/$D^{(*)}$ meson polarization[3].

[1] Front. Phy. 80, 1 (2000) [2] Phys. Lett. B 191, 442 (1987); 448, 320(E) (1999) [3] PRD 97, 012004 (2018)

$\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ overview and status

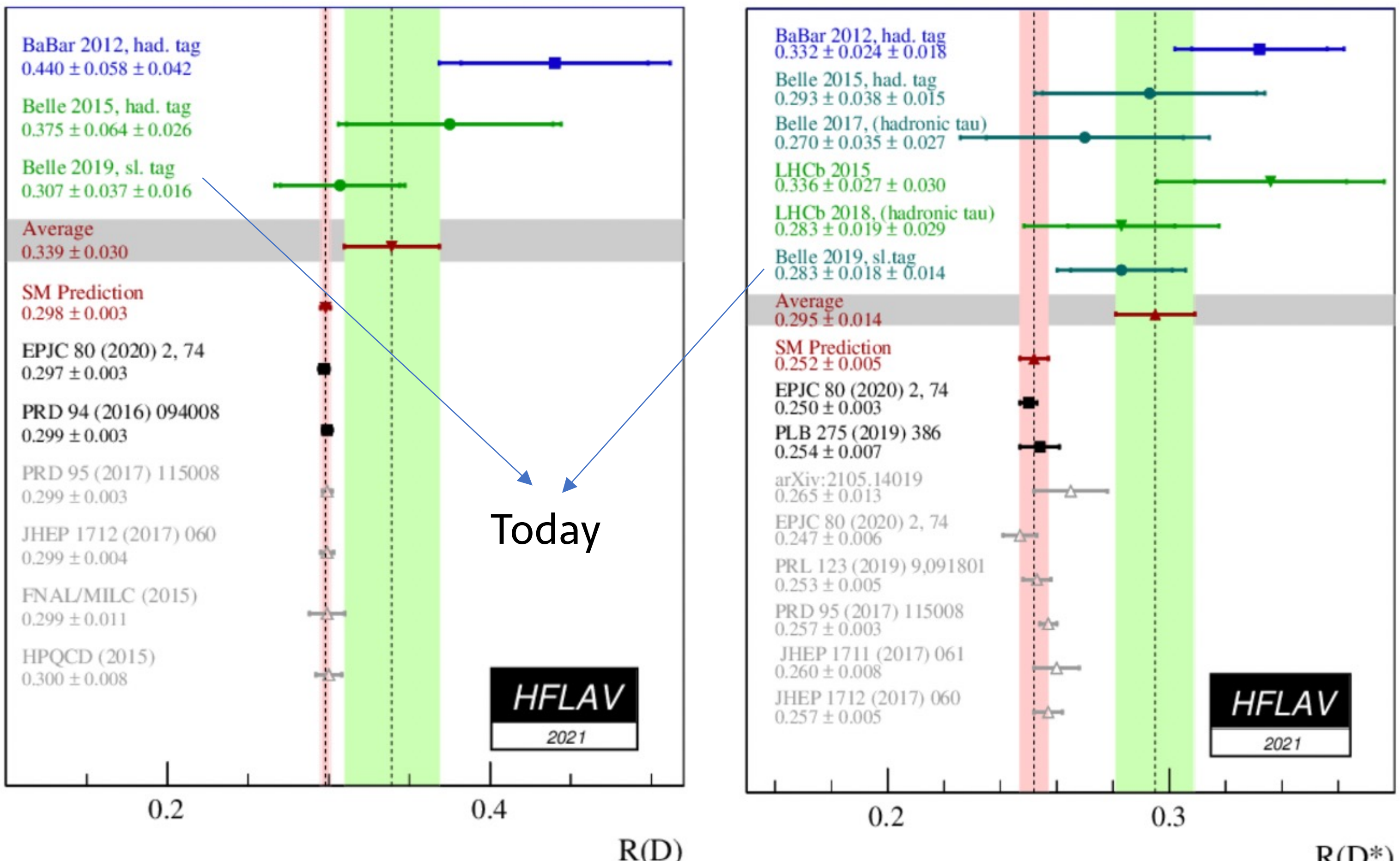
$$\mathcal{R}(D^{(*)}) \equiv \frac{\mathcal{B}(B \to D^{(*)}\tau\nu_\tau)}{\mathcal{B}(B \to D^{(*)}\ell\nu_\ell)}$$

- Common systematics will cancel out

Detection efficiency, theoretical uncertainty of form factor, and uncertainty of $|V_{cb}|$

- Predictions are theoretically clean

R(D) = 0.298±0.003 and R(D*) = 0.252±0.005

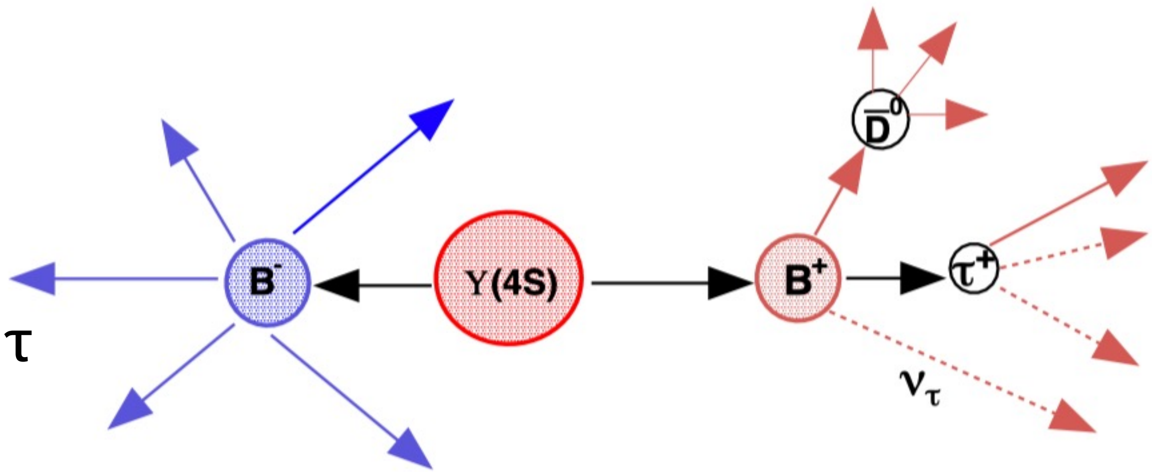


$B \to D^{(*)}\tau\nu$ reconstruction at *B* factories

- **Challenging with the presence of multiple neutrinos in the final states**

- Select the B_{sig} with

- a D(*) meson
- a charged daughter of τ



Tagging methods:

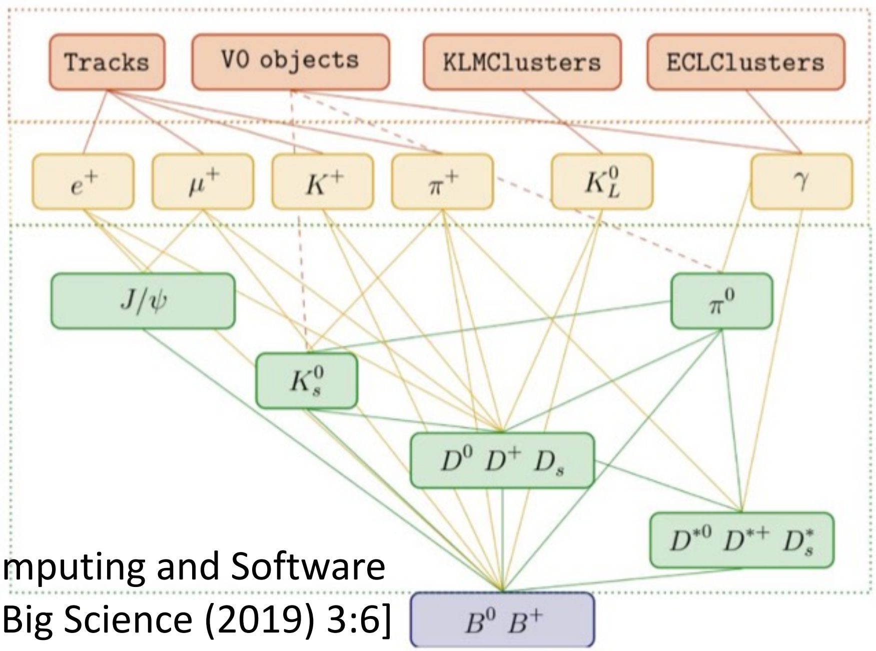
Hadronic Tag:
- Fully reconstructed in $B \to DX$ modes
- Tagging efficiency ~0.2%
- Low backgrounds

Semileptonic Tag:
- Reconstructed in $B \to D(*)l\nu$
- Tagging efficiency ~0.5%
- More backgrounds

Inclusive Tag:
- Reconstruct B tag with all particles except signal side
- Higher efficiency
- Need clean signal side final state

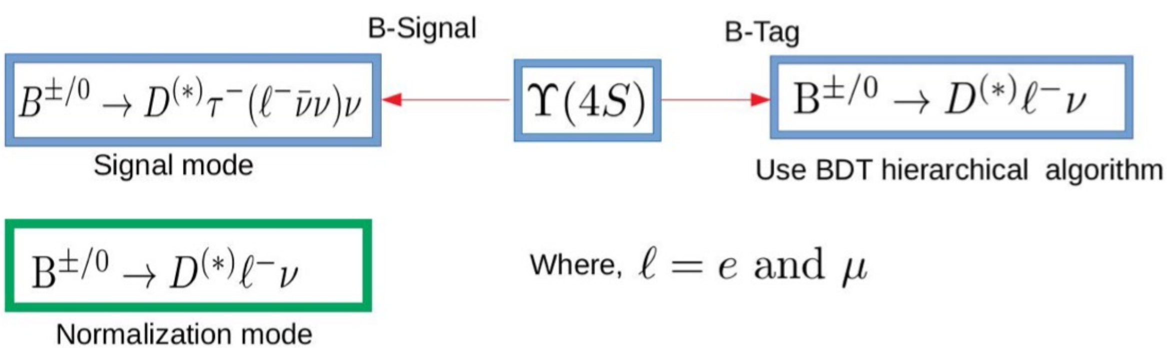
$\mathcal{R}(D^*)$ measurements at Belle with semileptonic tagging

New technique : Using Full Event Interpretation (FEI) tool developed in Belle II software framework.



[Computing and Software for Big Science (2019) 3:6]

- The B_{tag} is reconstructed using a hierarchical algorithm based on a multivariate analysis with Boosted-Decision Tree (BDT) in the $B \to D^{(*)}\tau\nu$ channel
- Better efficiency and enables to use more signal decay modes
- Simultaneous measurements of R(D) and R(D*)



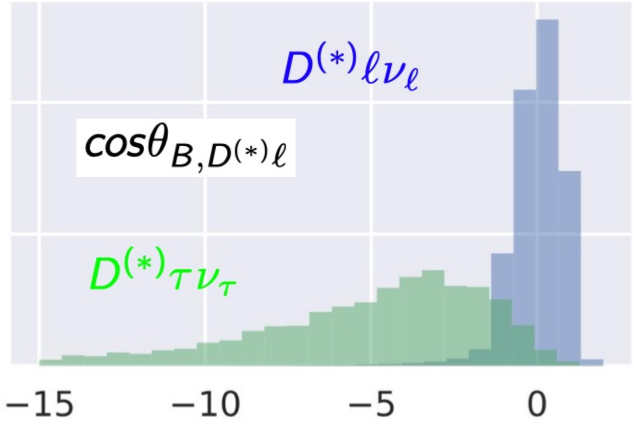
- 772×10^6 $B\bar{B}$ events
- Four-disjoint data samples: $D^+\ell^-$, $D^0\ell^-$, $D^{*+}\ell^-$. $D^{*0}\ell^-$
- B_{tag} and B_{sig} have opposite flavors

Selection criteria

- Separate $B_{tag} \to D^{(*)}l\nu$ and $B_{tag} \to D^{(*)}\tau\nu$ by cosine of the angle between the momentum of the B meson and $D^{(*)}l$ under the assumption that only one massless particle is not reconstructed:

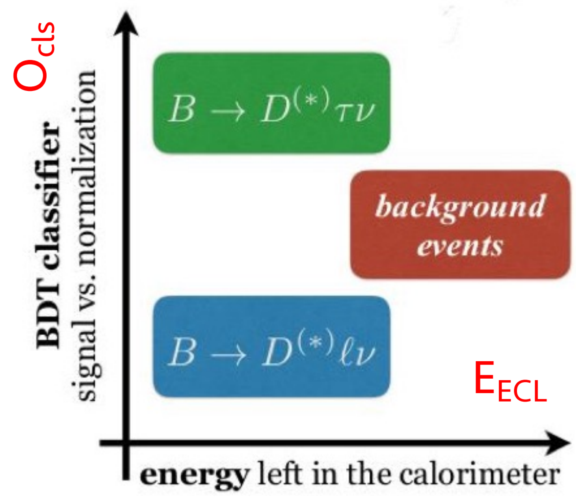
$$\boxed{\cos\theta_{B,D^{(*)}\ell} = \frac{2E_{beam}E_{D^{(*)}\ell} - m_B^2 - m_{D^{(*)}\ell}^2}{2|p_B||p_{D^{(*)}\ell}|}}$$

Correct B_{tag}: $-1 < \cos\theta_{B,D^{(*)}\ell} < 1$



- Separate signal and normalization modes by Boosted-Decision Tree (BDT) classifier output (Based on the XGBoost package)
  $\to$ Input variables: Visible energy, m_{miss}^2, $\cos\theta_{B,D^{(*)}l}$

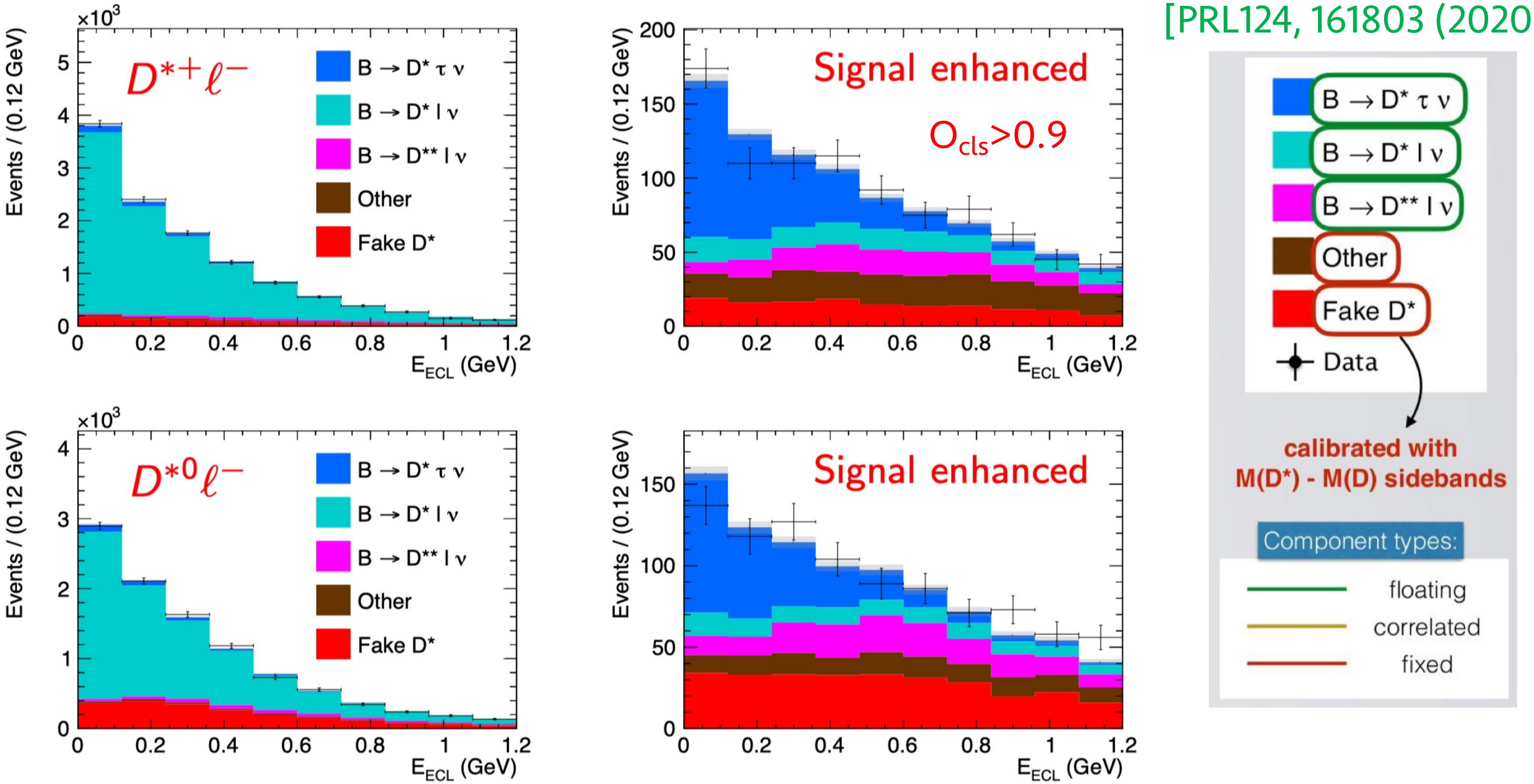
- Best candidates are selected in case of multiple B_{tag} (on highest tagging classifier output) and B_{sig} (on highest p-value of vertex fit of the charm daughters)



Fit results: D*ℓ sample

- We extract the yields of signal and normalization modes from a two-dimensional extended maximum-likelihood fit to the variables O_{cls} and E_{ECL}.
- The fit is performed simultaneously to the four $D^{(*)}\ell$ samples.

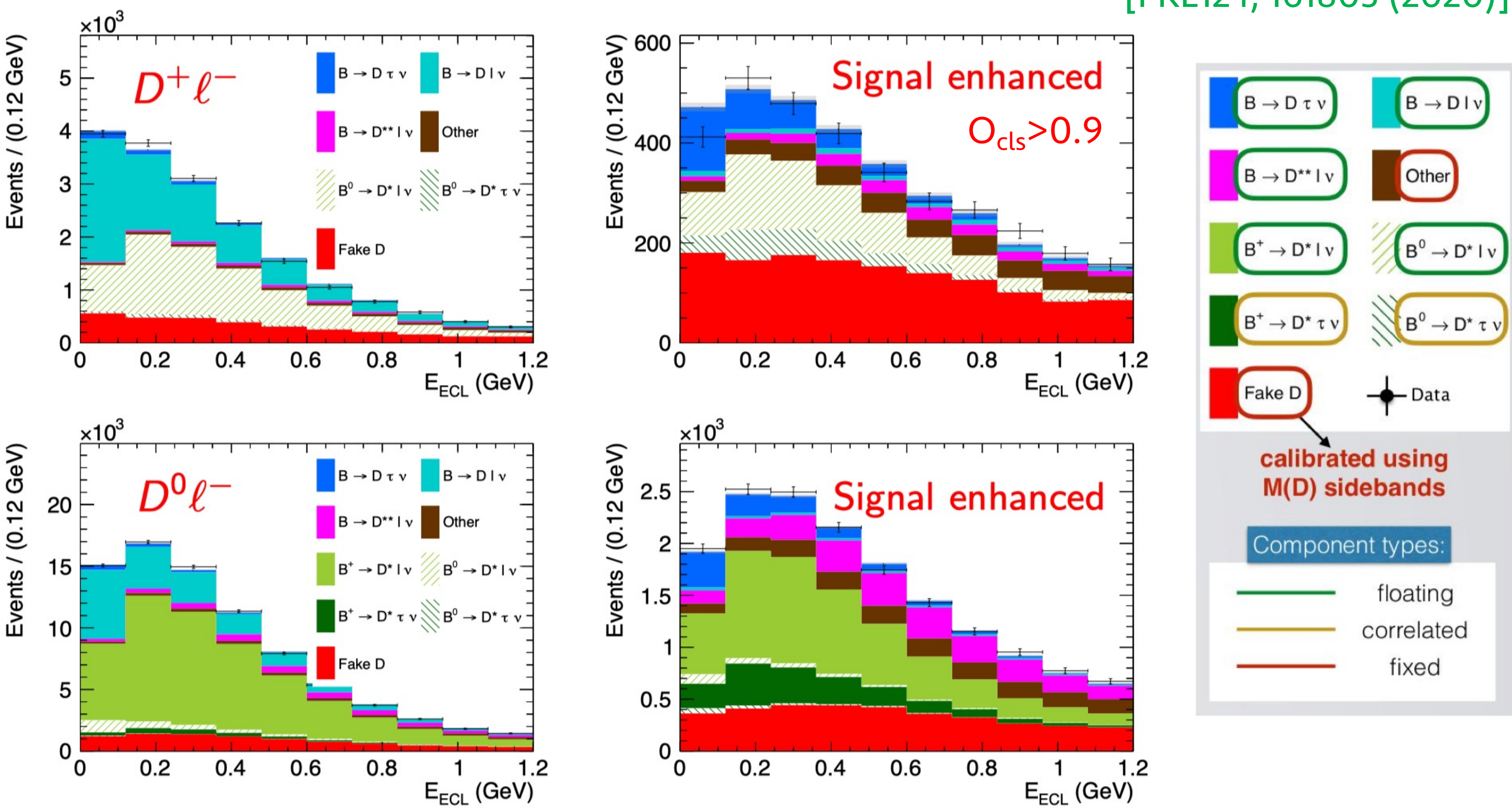
[PRL124, 161803 (2020)]



E_{ECL}: the sum of the energies of neutral clusters detected in the ECL that are not associated with any reconstructed particles.

Fit results: $D\ell$ sample

[PRL124, 161803 (2020)]



A large fraction of $B \to D^*\ell\nu$ and $B \to D^*\tau\nu$ decays from both B^0 and B^+ are reconstructed in the Dl samples.

Results

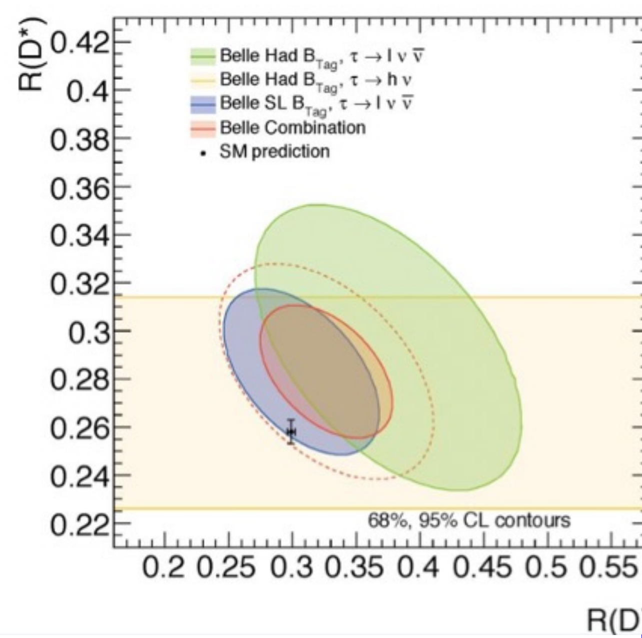
[PRL124, 161803 (2020)]

$$\mathcal{R}(\mathcal{D}^{(*)}) = \frac{\mathcal{B}(B\to D^{(*)}\tau^+\nu_\tau)}{\mathcal{B}(B\to D^{(*)}l^+\nu_l)} = \frac{1}{2\mathcal{B}(\tau^-\to\ell^-\bar{\nu}_\ell\nu_\tau)} \cdot \overset{\text{MC}}{\frac{\epsilon_{\rm sig}}{\epsilon_{norm}}} \cdot \frac{N_{\rm sig}}{N_{norm}}$$

$\boxed{\mathcal{R}(\mathcal{D}^*) = 0.283 \pm 0.018 \pm 0.014}$ Agrees with SM within 1.1σ

$\boxed{\mathcal{R}(\mathcal{D}) = 0.307 \pm 0.037 \pm 0.016}$ Agrees with SM within 0.2σ

- Most precise measurements of R(D) and R(D*) to date
- First R(D) measurement performed with a semileptonic tag
- R(D) − R(D*) Belle average is now within 1.6σ of the SM prediction
- Expected precisions at Belle II: (±2.0±2.5)% for R(D) and (±1.0±2.0)% for R(D*) [PTEP2019,123C01 (2019)]



$B \to X_u \ell^+ \nu_\ell$ inclusive for $|V_{ub}|$ and differential branching fractions

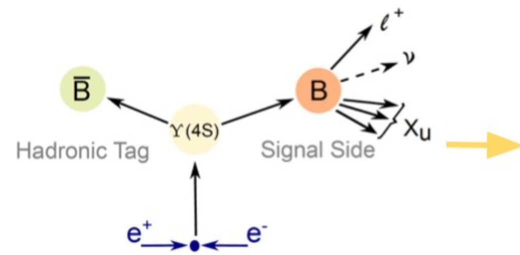
[PRD 104, 012008 (2021)]

[arXiv:2107.13855 (2021), accepted by PRL]

$B \rightarrow X_u \ell^+ \nu_\ell$ inclusive for $|V_{ub}|$

- Features of the analysis

- extend the probed $B \rightarrow X_u \ell^+ \nu_\ell$ phase space into regions dominated by $B \rightarrow X_c \ell^+ \nu_\ell$
- focus on integrated (over phase space) measurement of $\Delta\mathcal{B}$



Can fully assign each final state particle to either the tag or signal side

→ Allows to reconstruct X_u

- Main improvements compared previous work [PRL 104, 021801 (2010)]

- Using full Belle dataset of 711 fb^{-1}
- more efficient B tagging using NN ➔ effective increase of statistics (×1.8)
- improved modeling of $B \rightarrow X_c \ell^+ \nu_\ell$ background (e.g. for "gap modes") as well as signal

$B \rightarrow X_u \ell^+ \nu_\ell$

$\mathcal{B}$	Value B^+	Value B^0
$B \rightarrow X_u \ell^+ \nu_\ell$		
$B \rightarrow \pi \ell^+ \nu_\ell$	$(7.8 \pm 0.3) \times 10^{-5}$	$(1.5 \pm 0.06) \times 10^{-4}$
$B \rightarrow \eta \ell^+ \nu_\ell$	$(3.9 \pm 0.5) \times 10^{-5}$	...
$B \rightarrow \eta' \ell^+ \nu_\ell$	$(2.3 \pm 0.8) \times 10^{-5}$	...
$B \rightarrow \omega \ell^+ \nu_\ell$	$(1.2 \pm 0.1) \times 10^{-4}$	...
$B \rightarrow \rho \ell^+ \nu_\ell$	$(1.6 \pm 0.1) \times 10^{-4}$	$(2.9 \pm 0.2) \times 10^{-4}$
$B \rightarrow X_u \ell^+ \nu_\ell$	$(2.2 \pm 0.3) \times 10^{-3}$	$(2.0 \pm 0.3) \times 10^{-3}$

$B \rightarrow X_c \ell^+ \nu_\ell$

$B \rightarrow X_c \ell^+ \nu_\ell$		
$B \rightarrow D \ell^+ \nu_\ell$	$(2.5 \pm 0.1) \times 10^{-2}$	$(2.3 \pm 0.1) \times 10^{-2}$
$B \rightarrow D^* \ell^+ \nu_\ell$	$(5.4 \pm 0.1) \times 10^{-2}$	$(5.1 \pm 0.1) \times 10^{-2}$
$B \rightarrow D_0^* \ell^+ \nu_\ell$ $(\hookrightarrow D\pi)$	$(4.2 \pm 0.8) \times 10^{-3}$	$(3.9 \pm 0.7) \times 10^{-3}$
$B \rightarrow D_1^* \ell^+ \nu_\ell$ $(\hookrightarrow D^*\pi)$	$(4.2 \pm 0.8) \times 10^{-3}$	$(3.9 \pm 0.8) \times 10^{-3}$
$B \rightarrow D_1 \ell^+ \nu_\ell$ $(\hookrightarrow D^*\pi)$	$(4.2 \pm 0.3) \times 10^{-3}$	$(3.9 \pm 0.3) \times 10^{-3}$
$B \rightarrow D_2^* \ell^+ \nu_\ell$ $(\hookrightarrow D^*\pi)$	$(1.2 \pm 0.1) \times 10^{-3}$	$(1.1 \pm 0.1) \times 10^{-3}$
$B \rightarrow D_2^* \ell^+ \nu_\ell$ $(\hookrightarrow D\pi)$	$(1.8 \pm 0.2) \times 10^{-3}$	$(1.7 \pm 0.2) \times 10^{-3}$
$B \rightarrow D_1 \ell^+ \nu_\ell$ $(\hookrightarrow D\pi\pi)$	$(2.4 \pm 1.0) \times 10^{-3}$	$(2.3 \pm 0.9) \times 10^{-3}$
$B \rightarrow D\pi\pi \ell^+ \nu_\ell$	$(0.6 \pm 0.6) \times 10^{-3}$	$(0.6 \pm 0.6) \times 10^{-3}$
$B \rightarrow D^*\pi\pi \ell^+ \nu_\ell$	$(2.2 \pm 1.0) \times 10^{-3}$	$(2.0 \pm 1.0) \times 10^{-3}$
$B \rightarrow D\eta \ell^+ \nu_\ell$	$(4.0 \pm 4.0) \times 10^{-3}$	$(4.0 \pm 4.0) \times 10^{-3}$
$B \rightarrow D^*\eta \ell^+ \nu_\ell$	$(4.0 \pm 4.0) \times 10^{-3}$	$(4.0 \pm 4.0) \times 10^{-3}$
$B \rightarrow X_c \ell^+ \nu_\ell$	$(10.8 \pm 0.4) \times 10^{-2}$	$(10.1 \pm 0.4) \times 10^{-2}$

Background suppression using BDT

BDT traning ($B \to X_u \ell^+ \nu_\ell$ against $B \to X_c \ell^+ \nu_\ell$) using

- MM^2 (larger multiplicity of $X_c \to$ more broadening)
- D^* veto — use slow-π kinematics
- \# of kaons: $K^{\pm}$, K_S^0
- B_{sig} vertex fit
- Sum (charges)

Selection	$B \to X_u \ell^+ \nu_\ell$	$B \to X_c \ell^+ \nu_\ell$	Data
$M_{bc} > 5.27$ GeV	84.8%	83.8%	80.2%
$\mathcal{O}_{BDT} > 0.85$	18.5%	1.3%	1.6%
$\mathcal{O}_{BDT} > 0.83$	21.9%	1.7%	2.1%
$\mathcal{O}_{BDT} > 0.87$	14.5%	0.9%	1.1%

Reconstructed kinematic variables

- Hadronic system *X*:

$$p_X = \sum_i \left(\sqrt{m_\pi^2 + |\mathbf{p_i}|^2}, \mathbf{p_i}\right) + \sum_i (E_i, \mathbf{k_i})$$

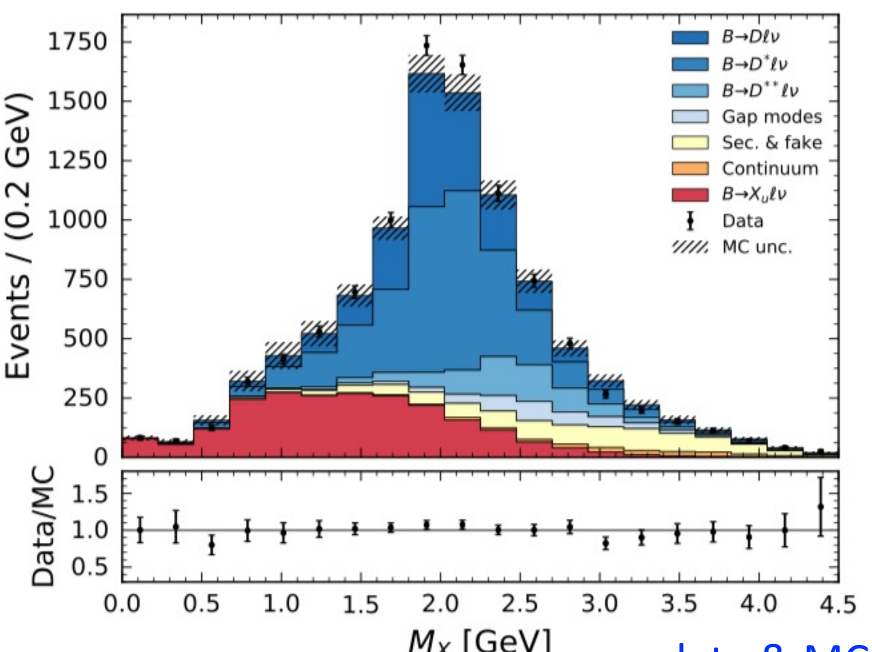
- Missing mass squared:

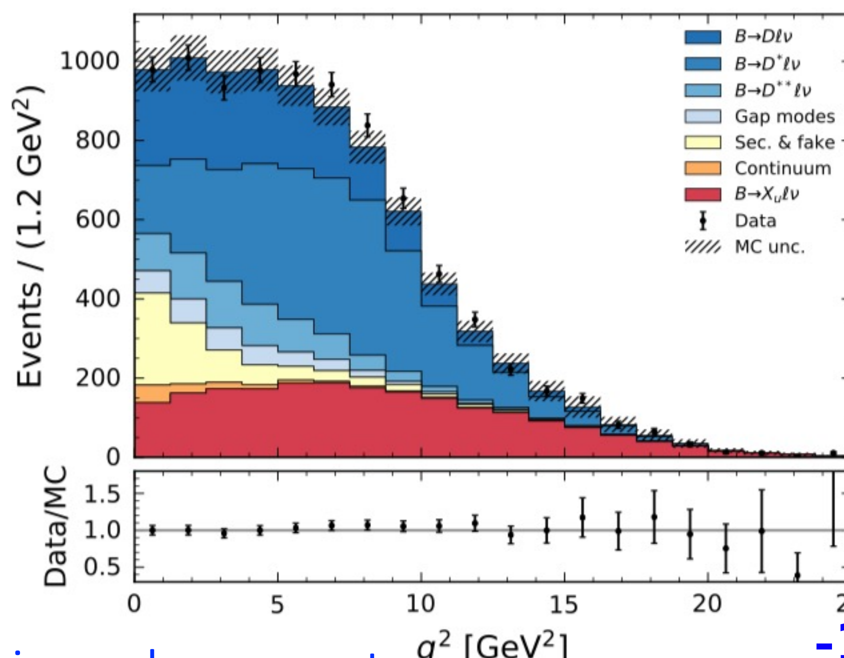
$$MM^2 = \left(P_{\Upsilon(4S)} - P_{tag} - P_X - P_\ell\right)^2$$

- Leptonic system:

$$q^2 = (P_B - P_X)^2 = (P_\ell + P_\nu)^2$$

Signal & background distributions after BDT:



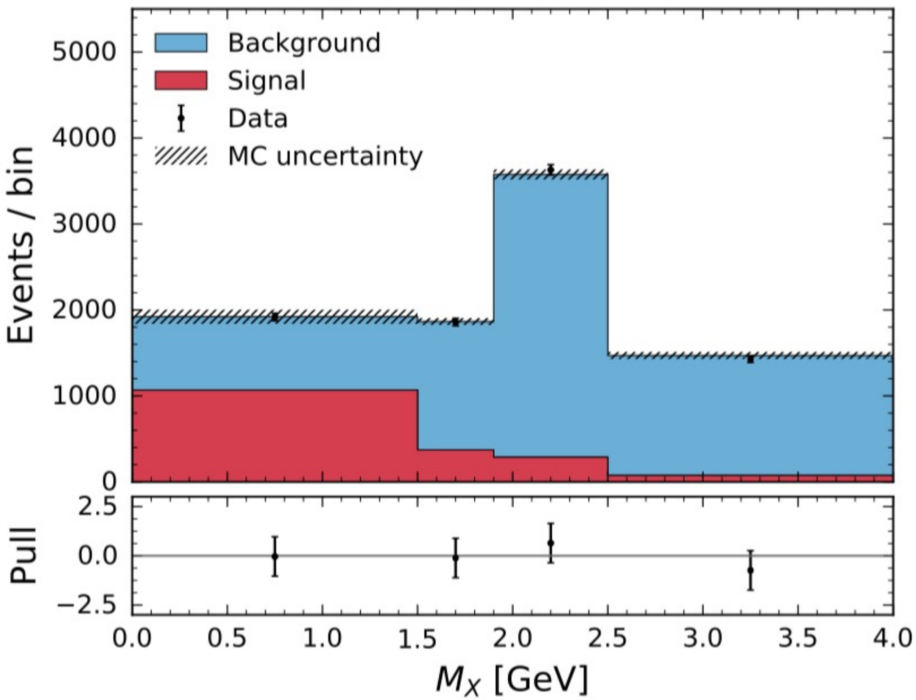


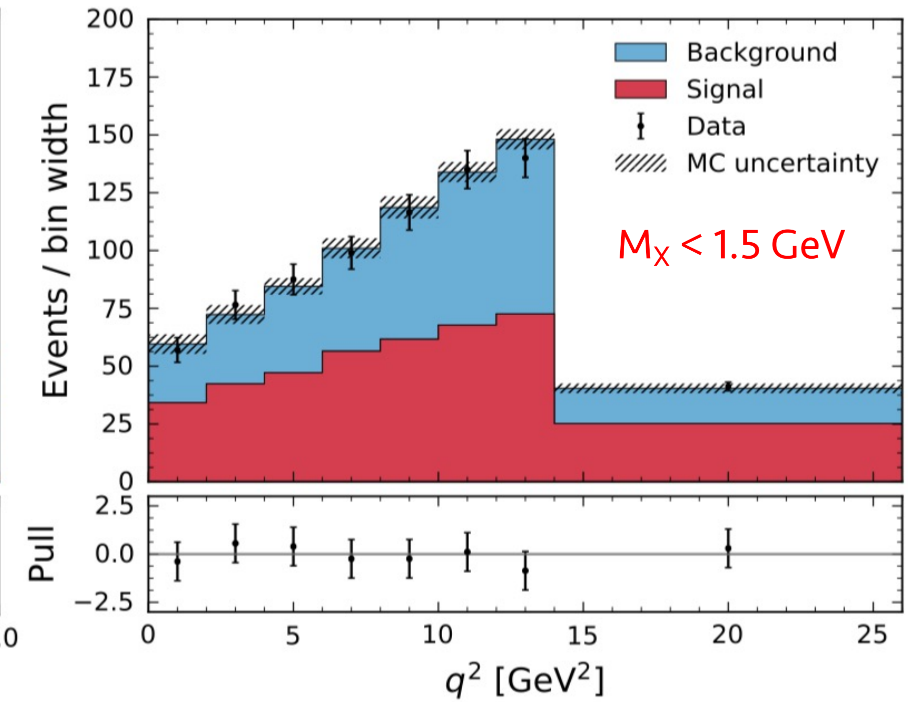
data & MC, in good agreement

Partial Branching Fractions of $B \to X_u \ell^+ \nu_\ell$

2D fit on (M_X, q^2) with $E_\ell^B > 1.0$ GeV:

- covers 86% of available signal phase space
- signal yields corrected by efficiency & acceptance





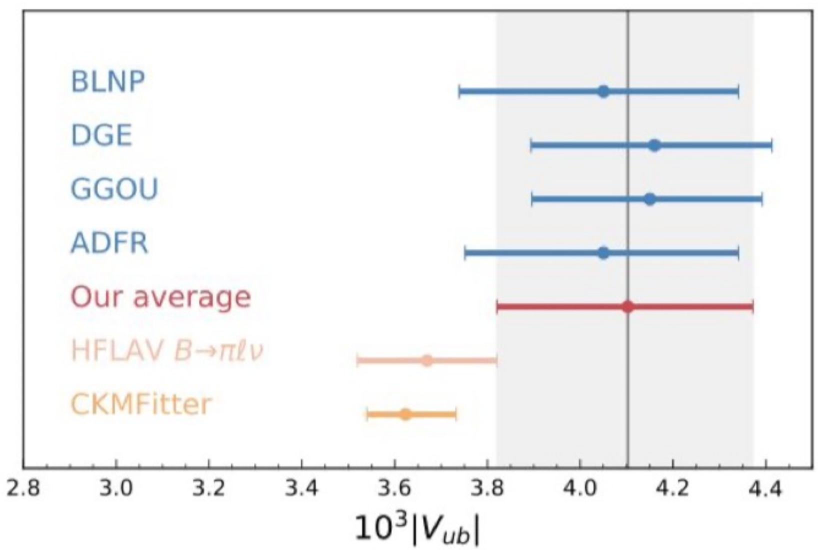
$$\Delta\mathcal{B}(E_\ell^B > 1\ \mathrm{GeV}) = (1.59 \pm 0.07_{stat} \pm 0.16_{sys}) \times 10^{-3}$$

- In good agreement with existing BaBar results [PRD 86, 032004 (2012)]
- Consistent within in 1.7σ compared to previous Belle results [PRL 104, 021801 (2010)]

Determination of $|V_{ub}|$

[PRD 104, 012008 (2021)]

- Using the relation $|V_{ub}| = \sqrt{\dfrac{\Delta\mathcal{B}(B \to X_u \ell^+ \nu_\ell)}{\tau_B \cdot \boxed{\Delta\Gamma(B \to X_u \ell^+ \nu_\ell)}}}$
- Based on four calculations of the decay rate



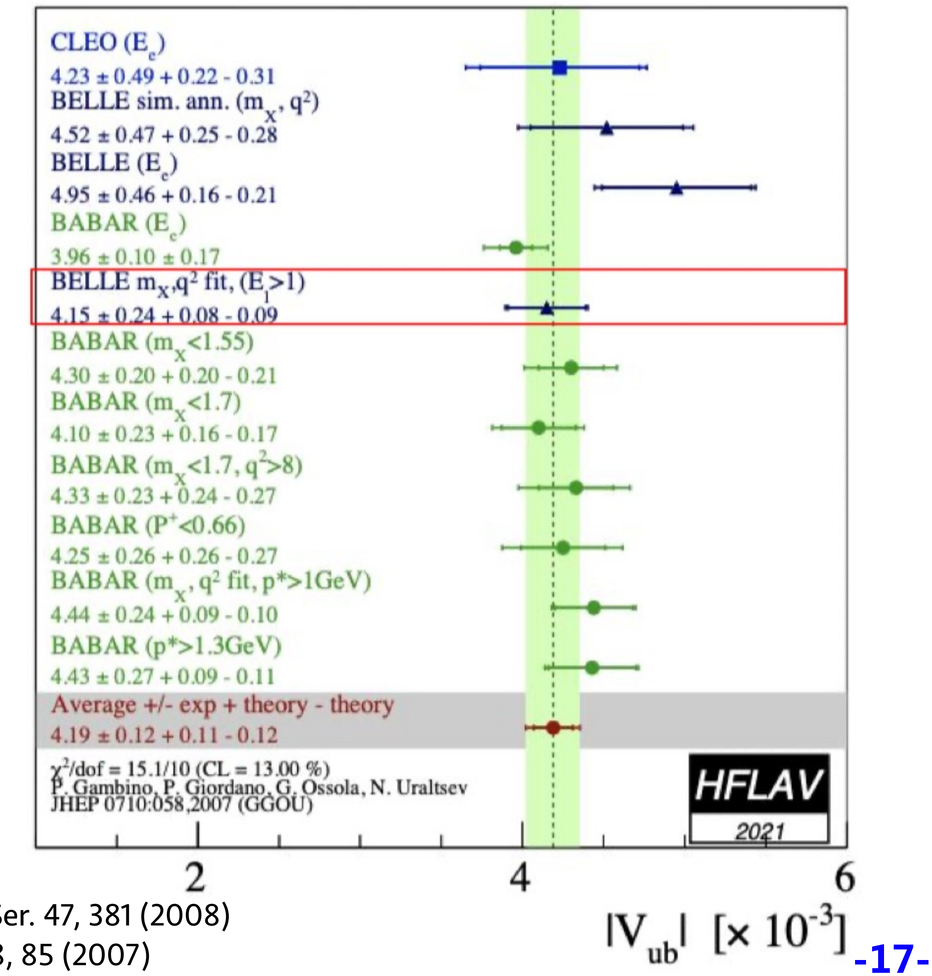
Our average:

$|V_{ub}| = (4.10 \pm 0.09_{stat} \pm 0.22_{sys} \pm 0.15_{theo}) \times 10^{-3}$

compatible with excl. and CKM expectation

within **1.3σ** and **1.6σ** respectively

Comparions based on GGOU



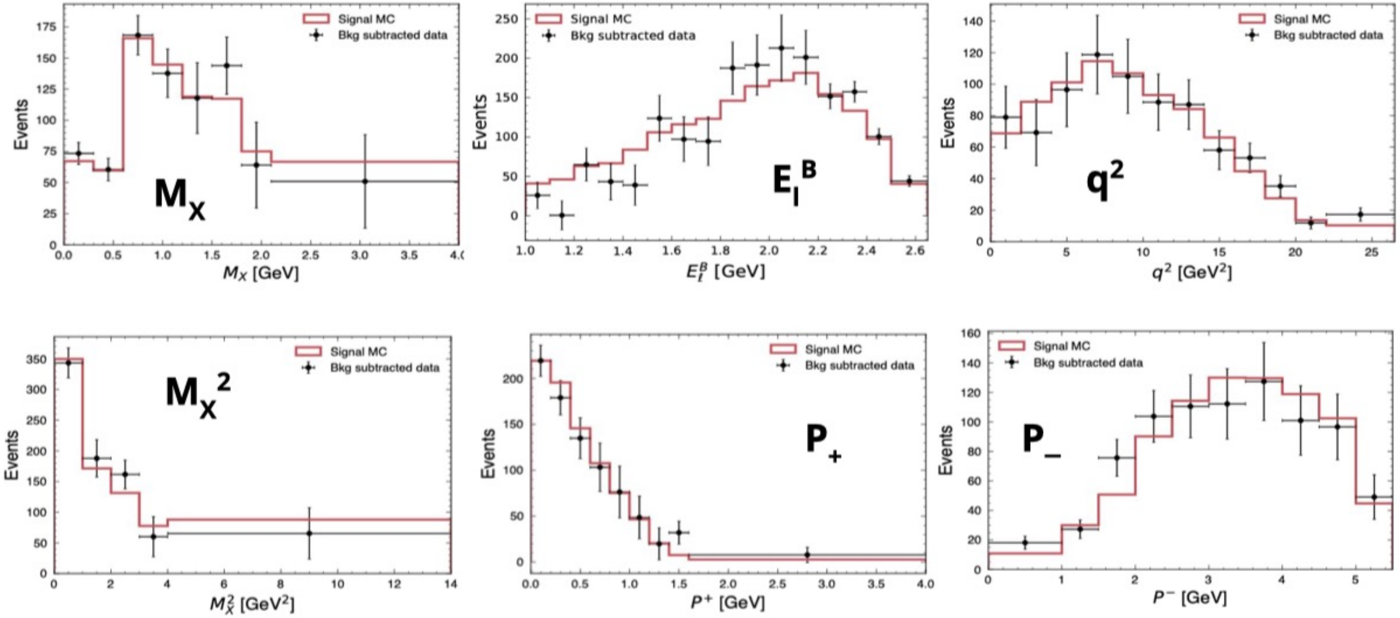
BLNP: PRD 72, 073006 (2005) DGE: JHEP 01, 097 (2006), Frascati Phys. Ser. 47, 381 (2008)
GGOU: JHEP 10, 058 (2007) ADFR: EPJC 59, 831 (2009), Nucl. Phys. B 768, 85 (2007)

$B \to X_u \ell^+ \nu_\ell$ differential BF

- Measure kinematic variables in the phase space of $E_\ell^B > 1.0$ GeV

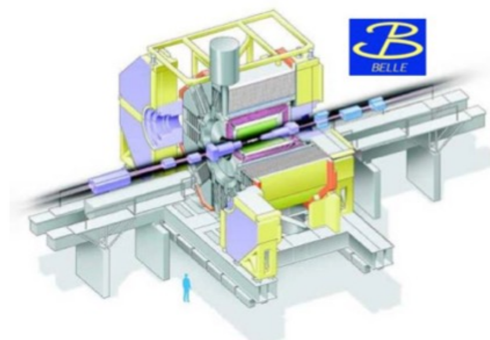
  q^2, E_l^B, M_X, M_X^2, P_+, $P\_$ (light-cone momenta: $P_\pm = E_X \mp |p_X|$)

- Basic selection & reconstruction follows integrated partial BF analysis [PRD 104, 012008 (2021)]
- Additional selections on $|E_{miss} - P_{miss}| < 0.1$ GeV and reconstructed $M_X < 2.4$ GeV to improve resolution and reduce background shape uncertainty
- Background subtraction via M_X fit

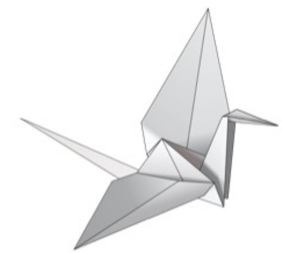


Unfolding for differential BF

***X*: true distribution**

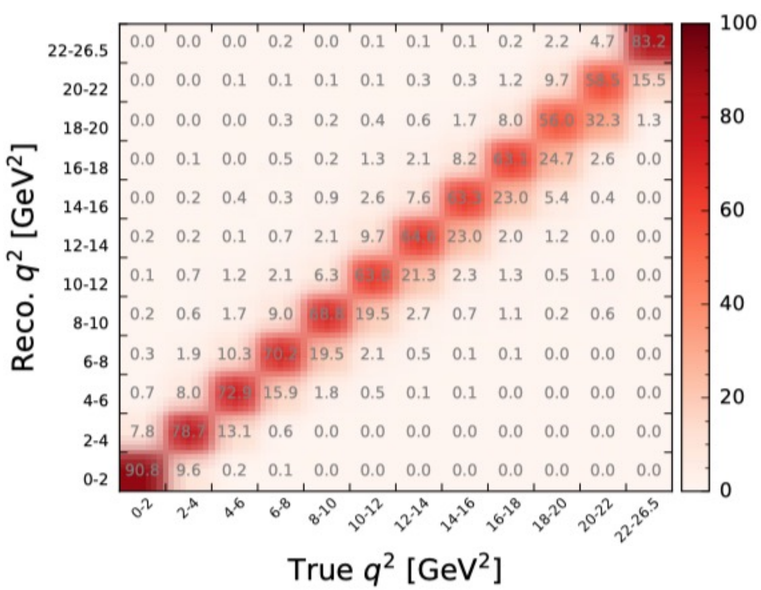


***M*: detector response**



***Y*: measured distribution**

- The migration matrix M as a representation of detector response
- M(i, j) = prob(%) to observe an event in bin i if it has a true value in bin j



MX* = *Y

Direct solution for X is

X* = *M*$^{-1}$*Y

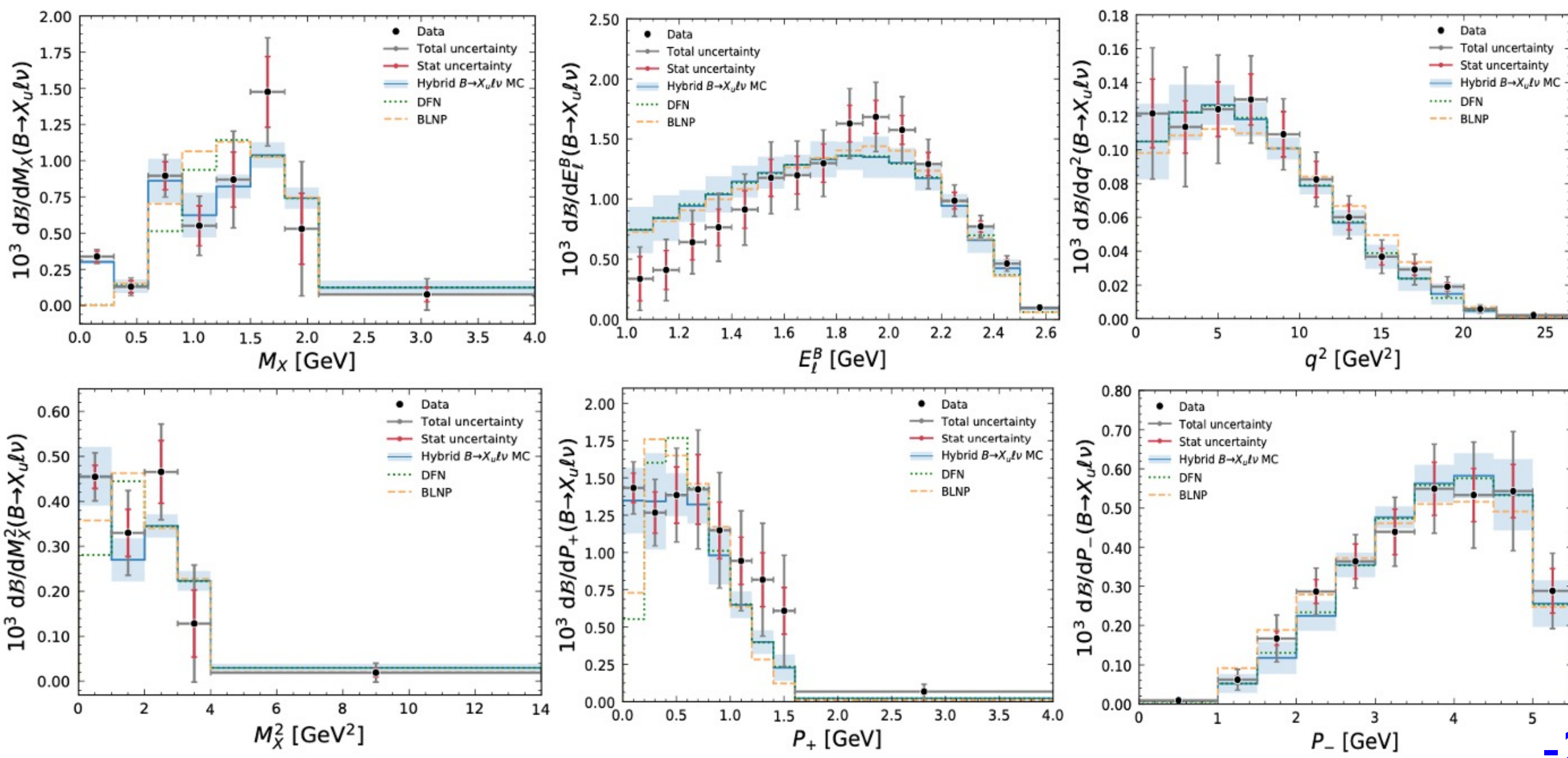
Singular-Value-Decomposition (SVD) [NIMA 372:469(1996)] is applied for unfolding in this analysis.

Differential spectra of $B \to X_u \ell^+ \nu_\ell$

- Convert unfolded yield to $\Delta\mathcal{B}$ in each bin considering efficiency & acceptance
- All MC shapes are normalized to $\Delta\mathcal{B} = 1.59 \times 10^{-3}$

[arXiv:2107.13855 (2021)]

- Differential braching fractions ($E_l^B > 1$ GeV) are measured for the first time
- Necessary input for future model-independent determinations of $|V_{ub}|$



DFN: JHEP 06, 017 (1999) BLNP: PRD 72, 073006 (2005)

$B^0 \rightarrow K^{*0}\tau^+\tau^-$ and $B_s^0 \rightarrow D_s X$

[arXiv:2110.03871 (2021), submitted to PRD]
[arXiv:2106.11265 (2021), submitted to PRD]

Measurement of $B^0 \to K^{*0}\tau^+\tau^-$

Motivation:

- The decay of $B^0 \to K^{*0}\tau^+\tau^-$ is highly suppressed in the SM and can only proceed via a flavor-changing neutral current, with a predicted branching fraction of order $\mathcal{O}(10^{-7})$.

- Compared with electron and muon modes, the decay is expected to be more sensitive to new physics in a model which has a coupling proportionate to the particle mass [1-5].

- The decay $B^+ \to K^+\tau^+\tau^-$ has been conducted by BaBar setting an upper limit $Br(B^+ \to K^+\tau^+\tau^-) < 2.25\times10^{-3}$ at 90% C.L. [6].

Reconstructions:

- B_{tag}: 489 exclusive B meson decays. $\ln(\mathcal{O}_{NB}) > -7$, $|\Delta E| < 0.06$ GeV, and $5.275 < M_{bc} < 5.290$ GeV/c^2 ($\varepsilon = 0.24\%$).

$$\Delta E = E_{\mathrm{Btag}} - E_{\mathrm{cm}}/2$$

- B_{sig}: $B^0 \to K^{*0}\tau^+\tau^-$

$$M_{\mathrm{bc}} = \sqrt{(E_{\mathrm{cm}}/2)^2/c^4 - |\vec{p}_{\mathrm{Btag}}|^2/c^2}$$

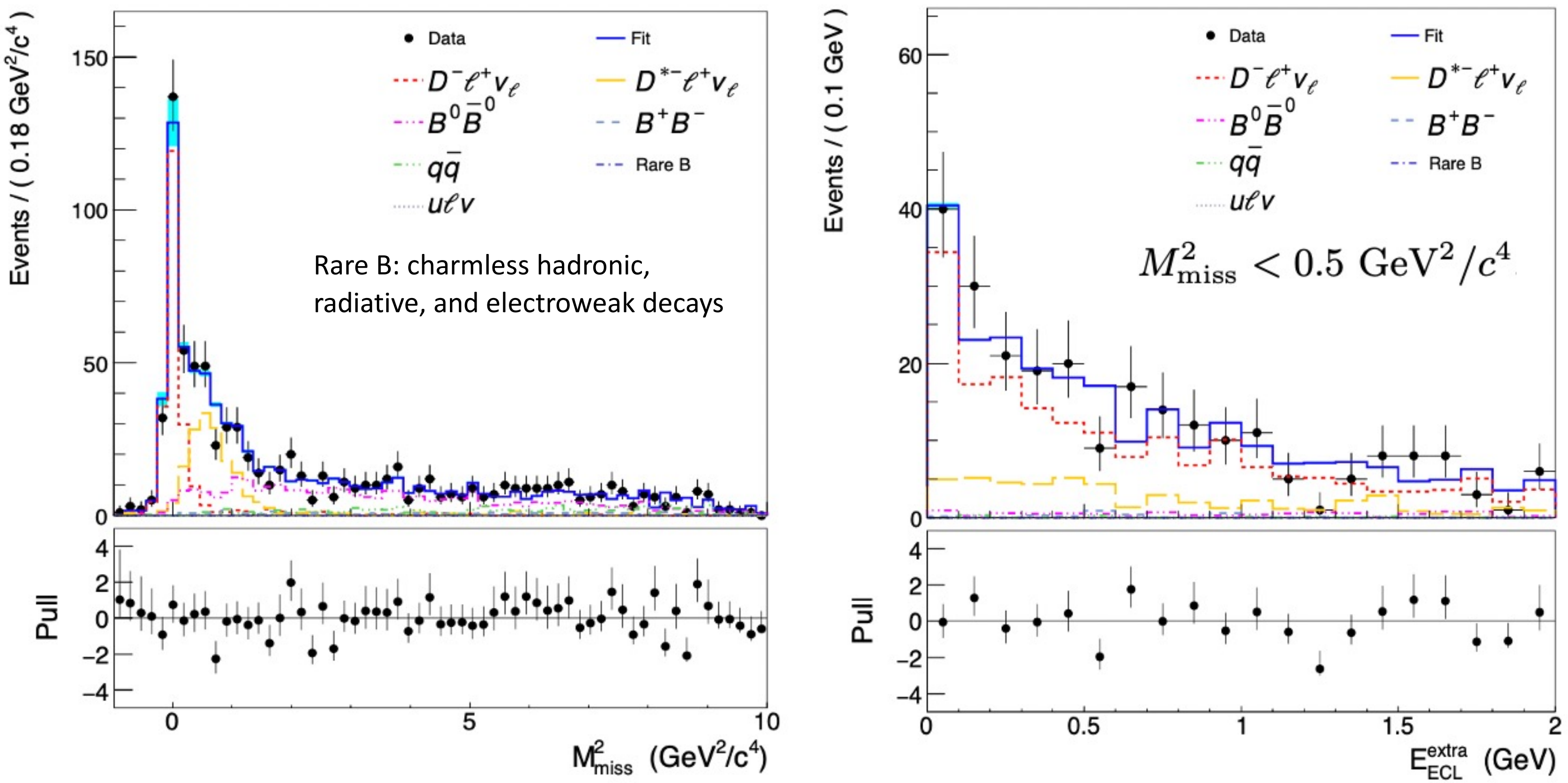
- $K^{*0} \to K^+\pi^-$; $\tau^- \to e^-\bar{\nu}_e\nu_\tau$, $\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$, and $\tau^- \to \pi^-\nu$

Data sample:
711 fb^{-1}, 772×10^6 $B\bar{B}$ pairs

[1] EPJC 77, 701 (2017) [2] PRL 120, 181802 (2018) [3] JHEP10, 184 (2015)
[4] JHEP 09, 40 (2017) [5] J. Phys. G 28, 307 (2002) [6] PRL 118, 031802 (2017)

Control channel $B^0 \to D^-(\to K^{*0}\pi^-)\ell^+\nu$

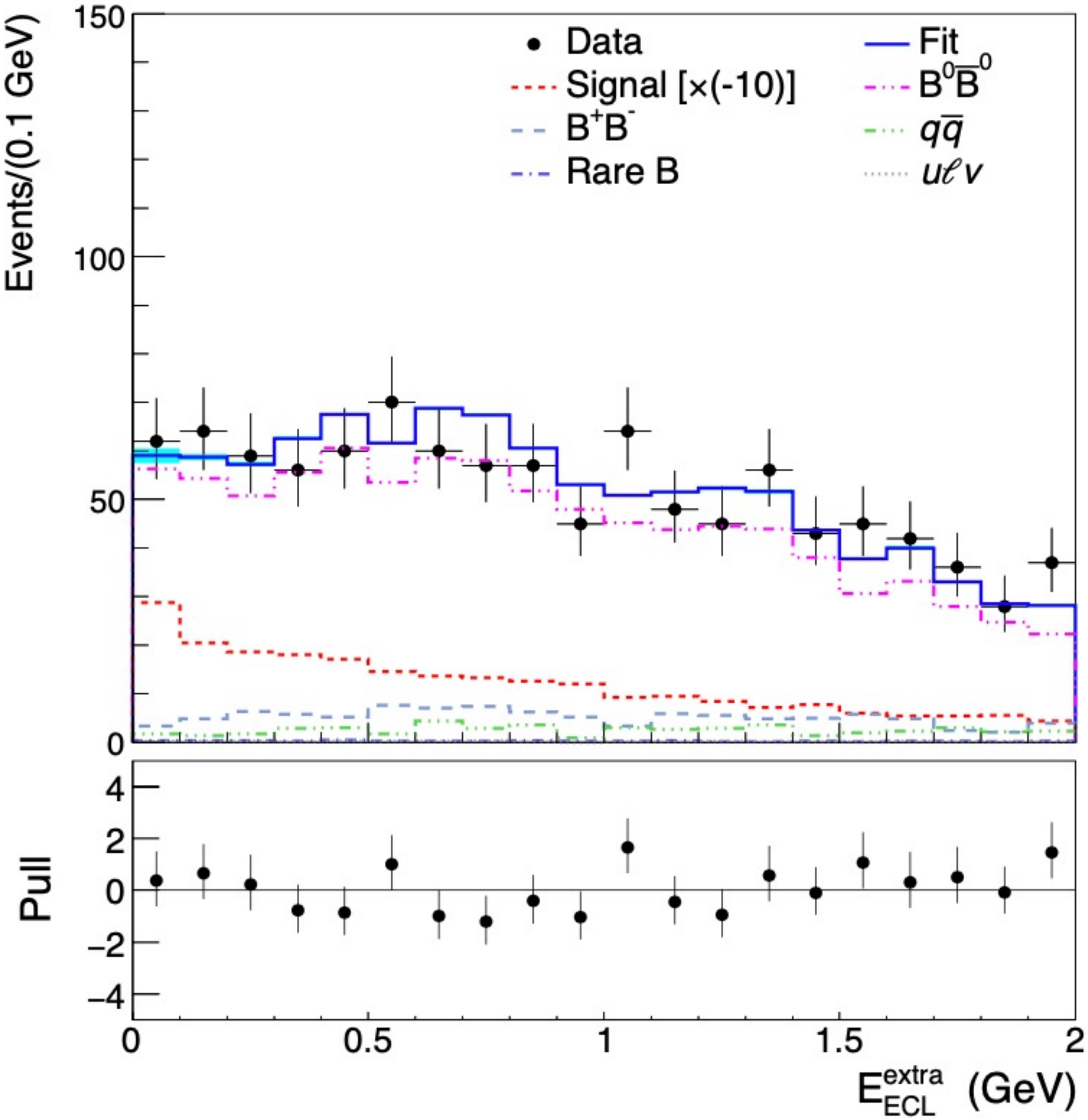
- Dominant backgrounds: such events in $B^0 \to D^-(\to K^{*0}\pi^-)\ell^+\nu$ have the same final-state particles as signal events.
- No requirements of $M(K^{*0}\pi^-)$ and M^2_{miss}



The branching fraction measured by fitting to $E^{\mathrm{extra}}_{\mathrm{ECL}}$ is (2.26±0.17)% and to M^2_{miss} is (2.19±0.15)%, respectively, which are in good agreement with the world average of (2.31±0.10)%.

$\mathcal{B}(B^0 \to K^{*0}\tau^+\tau^-)$

[arXiv:2110.03871 (2021)]



Optimized by the Punzi figure of merit [arXiv:physics/0308063.]

Signal Mode	$M_{K^{*0}\pi^-}$ (GeV/c^2)	M^2_{miss} $(\mathrm{GeV}^2/\mathrm{c}^4)$
$K^{*0}e^+e^-$	> 1.4	> 3.2
$K^{*0}e^\mp\mu^\pm$	> 1.4	> 1.6
$K^{*0}\mu^+\mu^-$	> 1.6	> 1.6
$K^{*0}\pi^\mp e^\pm$	> 1.4	> 2.0
$K^{*0}\pi^\mp\mu^\pm$	> 1.4	> 2.0
$K^{*0}\pi^+\pi^-$	> 1.5	< 9

- In signal region of $E_{\mathrm{ECL}}^{\mathrm{extra}} <$ 0.2 GeV: N_{sig} = -4.9 ± 6.0 and N_{bkg} = 122.4 ± 4.9

- The first experimental upper limit at 90% C.L. on $B^0 \to K^{*0}\tau^+\tau^-$ is 2.0×10^{-3}.

Measurement of $B_s^0 \to D_s X$

Motivation:

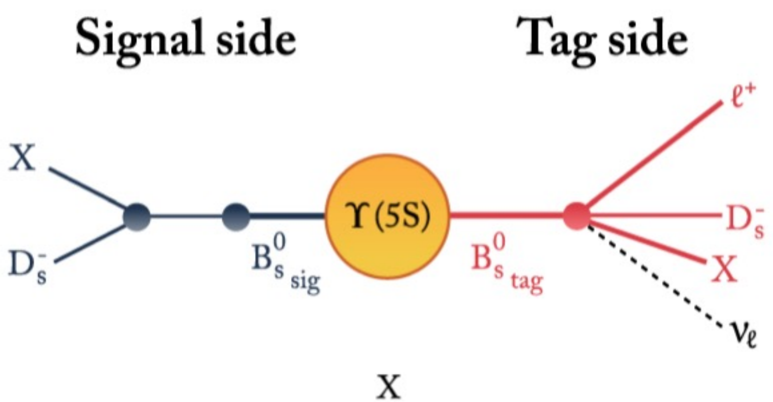
- Determine f_s (the fraction of Υ(5S) events containing B_s-meson pairs)

$$\mathcal{B}(\Upsilon(5S) \to D_x X)/2 = f_s \cdot \mathcal{B}(B_s \to D_x X) + f_q \cdot \mathcal{B}(B \to D_x X)$$

- B_s properties can provide important information on CKM matrix parameters.

Reconstruction:

The semileptonic tagging method:



Tag Channel	Signal Channel	Efficiency (%)
$\phi\pi$	$\phi\{K^+K^-\}\pi$	26.1 ± 0.5
	$K_S^0\{\pi^+\pi^-\}K$	38.5 ± 0.6
	$K^{*0}\{K^\pm\pi^\mp\}K$	24.6 ± 0.5
K_S^0K	$\phi\{K^+K^-\}\pi$	27.6 ± 0.5
	$K_S^0\{\pi^+\pi^-\}K$	37.8 ± 0.6
	$K^{*0}\{K^\pm\pi^\mp\}K$	24.6 ± 0.4

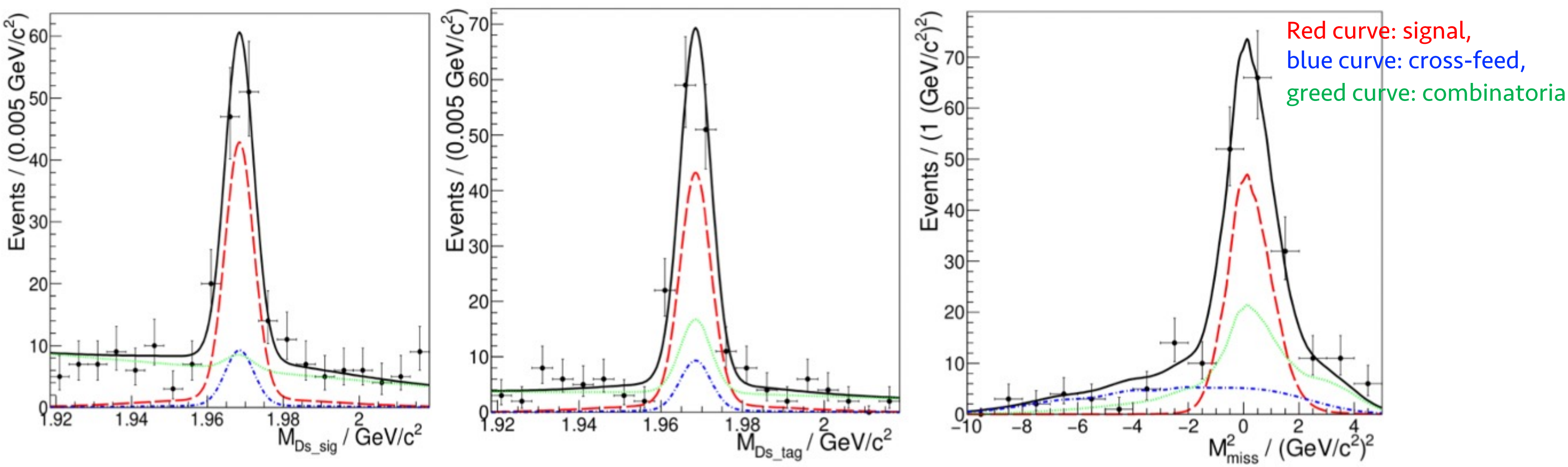
$$M_{\rm miss}^2 = (\sqrt{s}/2 - \delta E - E_{D\ell}^*)^2 - (p_{D\ell}^*)^2$$

Tag side: The number of B^0 tags is obtained by fitting $M_{\rm miss}^2$ and M_{Ds} distributions. Three categories: correct tags, incorrect B tags where the tag-lepton is combined with the signal-side D_s (cross-feed), and other incorrect tags (combinatorial backgrounds).

$\mathcal{B}(B_s^0 \to D_s X)$ and f_s

[arXiv:2106.11265 (2021)]

Signal side: The number of signal-side D_s obtained by performing a 3-dimensional binned maximum likelihood fit to the M^2_{miss}, $M^{tag}_{D_s}$, and $M^{sig}_{D_s}$ distributions.



$$\mathcal{B}(B_s^0 \to D_s X) = [61.6 \pm 5.3(\text{syst.}) \pm 2.1(\text{syst.})]\%$$

$$f_s = 0.278 \pm 0.028(\text{stat.}) \pm 0.035(\text{syst.})$$

- The first direct measurement of the $B_s \to D_s X$ inclusive branching fraction.
- $\mathcal{B}(B_s^0 \to D_s X)$ substantially lower than the world average but consistent within its large uncertainties.

Summary and conclusion

- Belle recently provided the most precise measurements of $R(D^{(*)})$, is now within 1.6σ of the SM prediction.

- The first measurement of several differential kinematic distributions in $B \to X_u \ell^+ \nu$, which enables a robust determination of $|V_{ub}|$.

- The first experimental upper limit at 90% C.L. on $\mathcal{B}(B^0 \to K^{*0}\tau^+\tau^-)$ and the most precise measurements of $\mathcal{B}(B_s^0 \to D_s X)$ and f_s are given.

Thanks for your attention!

$\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ overview and status

$$\mathcal{R}(\mathcal{D}^{(*)}) \equiv \frac{\mathcal{B}(B \to D^{(*)}\tau\nu_\tau)}{\mathcal{B}(B \to D^{(*)}\ell\nu_\ell)}$$

- Common systematics will cancel out

Detection efficiency, theoretical uncertainty of form factor, and uncertainty of $|V_{cb}|$

- Predictions are theoretically clean

R(D) = 0.298±0.003 and R(D*) = 0.252±0.005