

Search for the rare

$B^+ \rightarrow K^+ \nu \bar{\nu}$ and $B^0 \rightarrow K^{*0} \nu \bar{\nu}$ decays

using an inclusive tagging method

at the Belle II experiment



PhD oral defense

Hamburg, September 26, 2022

Filippo Dattola



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG



Alle mie nonne.

Introduction

Flavour physics

Introduction

Flavour physics

What is flavour?

- Quarks and leptons in the Standard Model (SM) come in 6 *flavours*.
- Flavour physics studies their spectrum and transmutations in weak interactions.



Introduction

Flavour physics

What is flavour?

- Quarks and leptons in the Standard Model (SM) come in 6 *flavours*.
- Flavour physics studies their spectrum and transmutations in weak interactions.

Why flavour?



Introduction

Flavour physics

What is flavour?

- Quarks and leptons in the Standard Model (SM) come in 6 *flavours*.
- Flavour physics studies their spectrum and transmutations in weak interactions.



Why flavour?

- Flavour is at the **core** of the SM.

Introduction

Flavour physics

What is flavour?

- Quarks and leptons in the Standard Model (SM) come in 6 *flavours*.
- Flavour physics studies their spectrum and transmutations in weak interactions.



Why flavour?

- Flavour is at the **core** of the SM.

Introduction

Flavour physics

What is flavour?

- Quarks and leptons in the Standard Model (SM) come in 6 *flavours*.
- Flavour physics studies their spectrum and transmutations in weak interactions.



Why flavour?

- Flavour is at the **core of the SM**.
- Flavour is a portal to **new physics**: measurements of **suppressed flavour processes like $b \rightarrow s\nu\bar{\nu}$ are powerful probes of possible physics beyond the SM**.



Introduction

Flavour physics

What is flavour?

- Quarks and leptons in the Standard Model (SM) come in 6 *flavours*.
- Flavour physics studies their spectrum and transmutations in weak interactions.



Why flavour?

- Flavour is at the **core of the SM**.
- Flavour is a portal to **new physics**: measurements of **suppressed flavour processes like $b \rightarrow s\nu\bar{\nu}$ are powerful probes of possible physics beyond the SM**.

**Light scalar
dark matter?**
[Phys. Rev. D 2020, 101]



**New heavy bosons?
 Z' or leptoquarks?**

[Phys. Lett. B 2017, 768;
Phys. Rev. D 2018, 98]

Theory of $b \rightarrow s\nu\bar{\nu}$

The $b \rightarrow s$ transition

Theory of $b \rightarrow s\nu\bar{\nu}$

The $b \rightarrow s$ transition

Weak neutral currents are *flavour diagonal*

Theory of $b \rightarrow s\nu\bar{\nu}$

The $b \rightarrow s$ transition

Weak neutral currents are *flavour diagonal*

$$J_\mu^0 \sim (\bar{d}, \bar{s}, \bar{b})_L \gamma_\mu \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L \quad \longrightarrow \quad \text{flavour-changing neutral-currents
absent at tree-level in the SM}$$

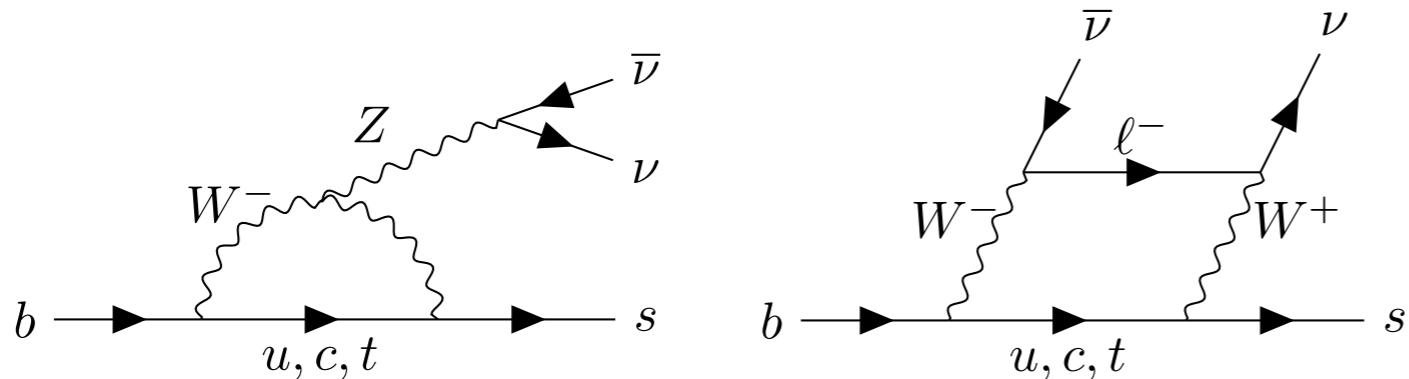
Theory of $b \rightarrow s\nu\bar{\nu}$

The $b \rightarrow s$ transition

Weak neutral currents are *flavour diagonal*

$$J_\mu^0 \sim (\bar{d}, \bar{s}, \bar{b})_L \gamma_\mu \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L \quad \longrightarrow \quad \text{flavour-changing neutral-currents
absent at tree-level in the SM}$$

$b \rightarrow s$ (e.g. $b \rightarrow s\nu\bar{\nu}$) only at **loop-level**



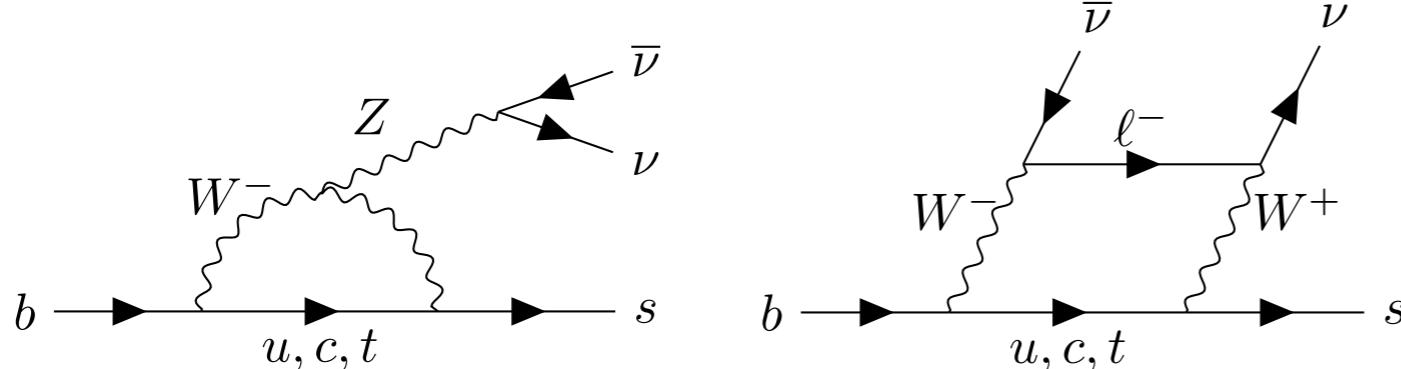
Theory of $b \rightarrow s\nu\bar{\nu}$

The $b \rightarrow s$ transition

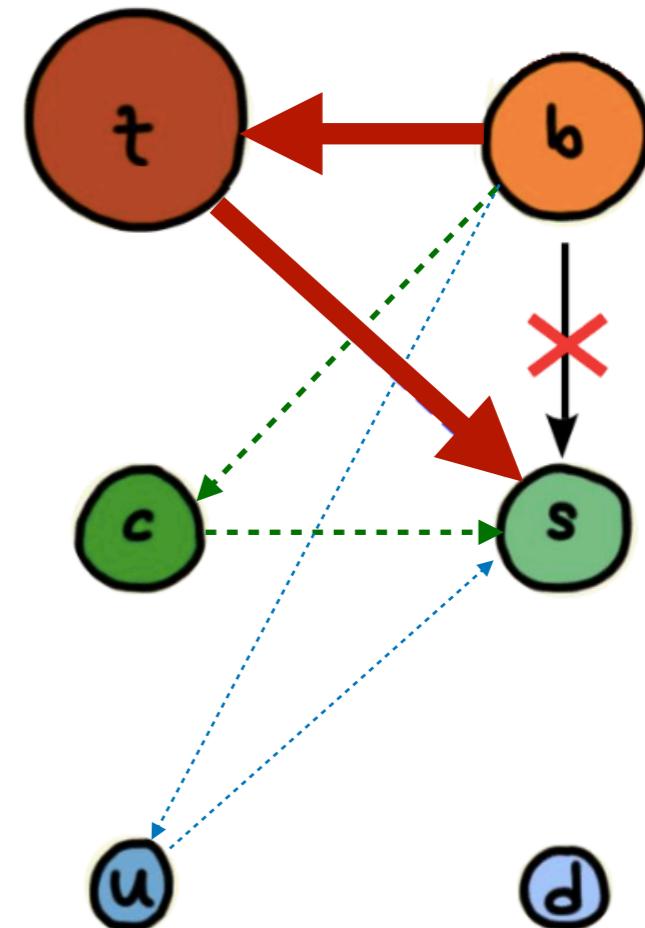
Weak neutral currents are *flavour diagonal*

$$J_\mu^0 \sim (\bar{d}, \bar{s}, \bar{b})_L \gamma_\mu \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L \quad \rightarrow \text{flavour-changing neutral-currents absent at tree-level in the SM}$$

$b \rightarrow s$ (e.g. $b \rightarrow s\nu\bar{\nu}$) only at **loop-level**



Dominated by **virtual top-quark exchange**



Theory of $b \rightarrow s\nu\bar{\nu}$

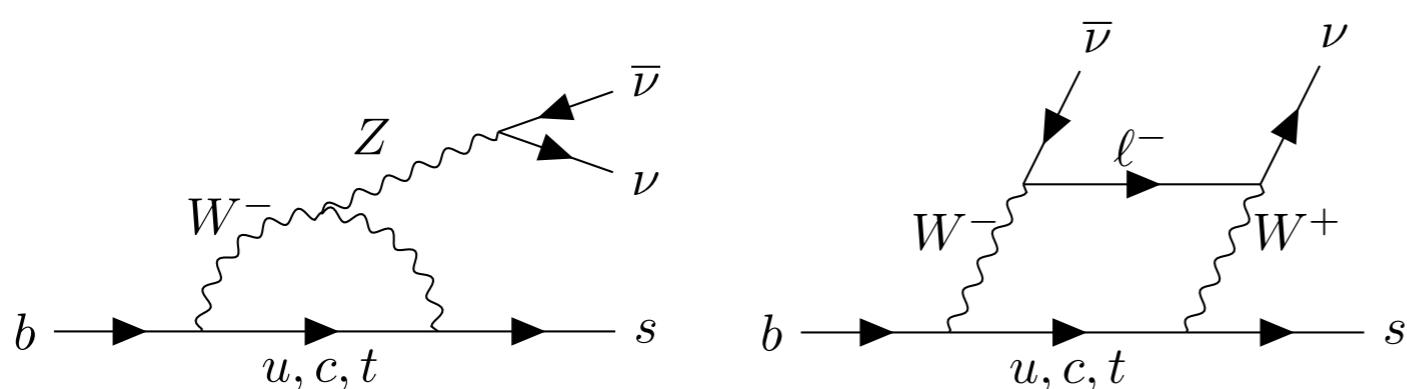
The $b \rightarrow s$ transition

Weak neutral currents are *flavour diagonal*

$$J_\mu^0 \sim (\bar{d}, \bar{s}, \bar{b})_L \gamma_\mu \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L \quad \xrightarrow{\text{orange arrow}} \quad \text{flavour-changing neutral-currents absent at tree-level in the SM}$$

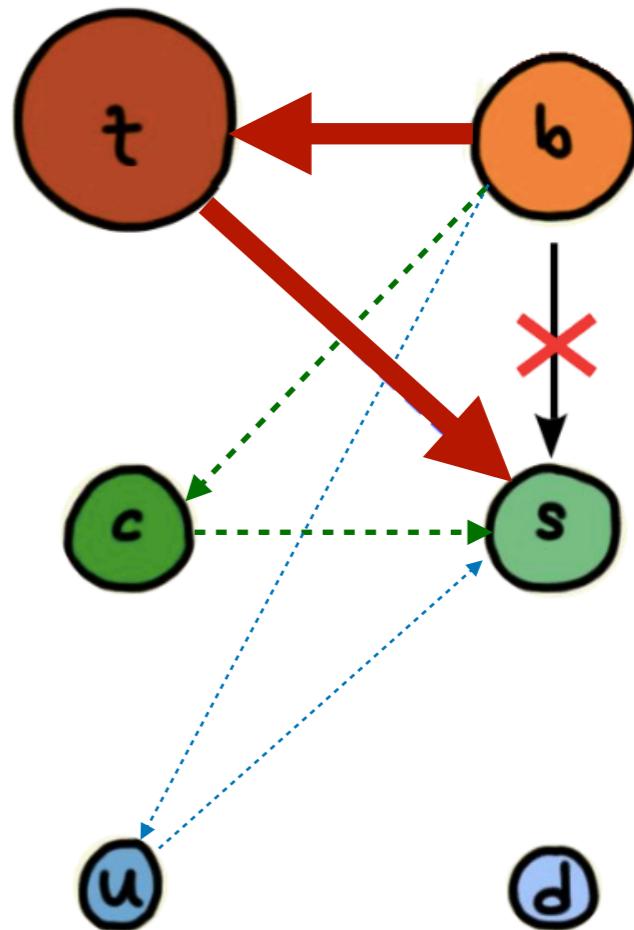
$$V_{\text{CKM}} \equiv \begin{pmatrix} d & s & b \\ u & \boxed{} & \boxed{} & \cdot \\ c & \boxed{} & \boxed{} & \cdot \\ t & \cdot & \cdot & \boxed{} \end{pmatrix}$$

$b \rightarrow s$ (e.g. $b \rightarrow s\nu\bar{\nu}$) only at **loop-level**



Dominated by **virtual top-quark exchange**

$$\text{amplitude} \sim \frac{1}{16\pi^2} \times V_{ts}^* V_{tb} \times F\left(\frac{m_t^2}{m_W^2}\right)$$





Theory of $b \rightarrow s\nu\bar{\nu}$

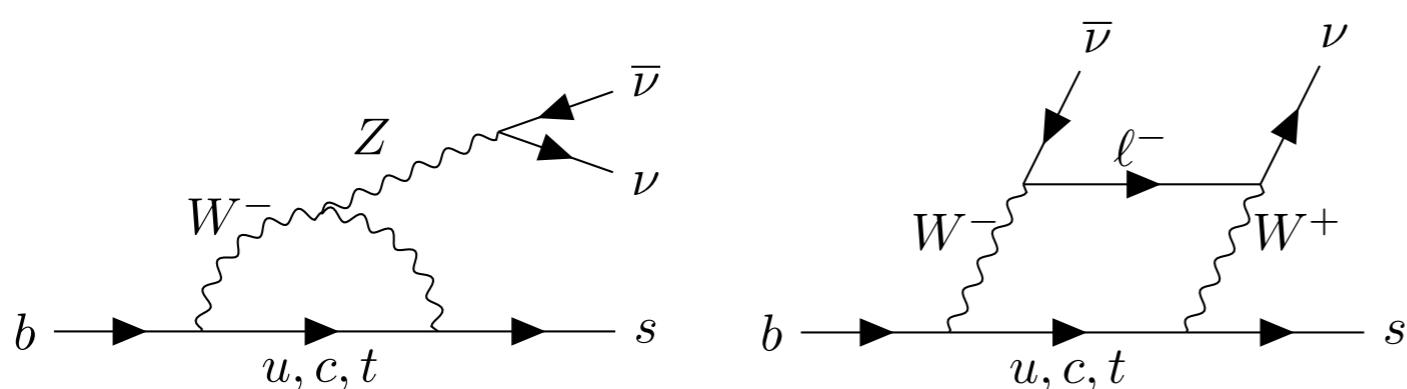
The $b \rightarrow s$ transition

Weak neutral currents are *flavour diagonal*

$$J_\mu^0 \sim (\bar{d}, \bar{s}, \bar{b})_L \gamma_\mu \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L \quad \xrightarrow{\text{orange arrow}} \quad \text{flavour-changing neutral-currents absent at tree-level in the SM}$$

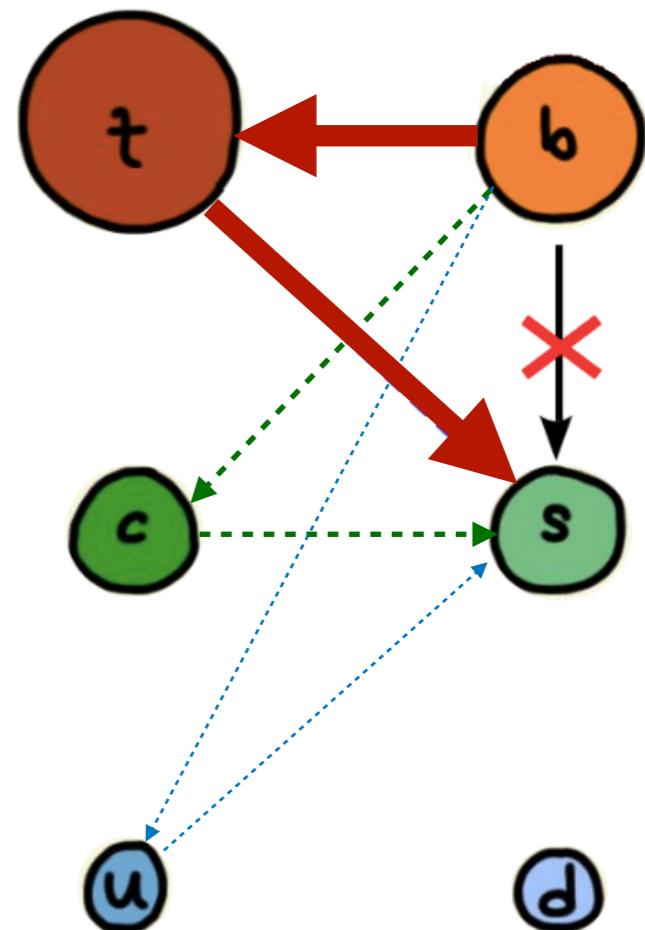
$$V_{\text{CKM}} \equiv \begin{pmatrix} d & s & b \\ u & \boxed{} & \boxed{} & \cdot \\ c & \boxed{} & \boxed{} & \cdot \\ t & \cdot & \cdot & \boxed{} \end{pmatrix}$$

$b \rightarrow s$ (e.g. $b \rightarrow s\nu\bar{\nu}$) only at **loop-level**



Dominated by **virtual top-quark exchange**

$$\text{amplitude} \sim \frac{1}{16\pi^2} \times V_{ts}^* V_{tb} \times F\left(\frac{m_t^2}{m_W^2}\right) \sim 2.5 \times 10^{-4} \rightarrow \text{rare transition}$$



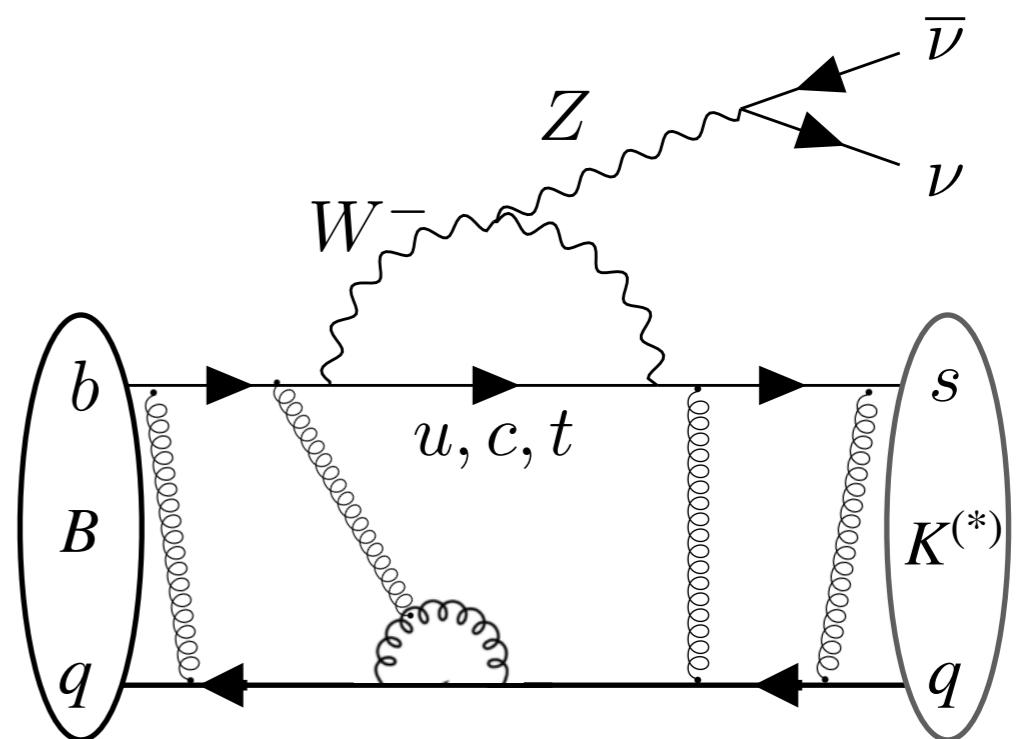
Theory of $b \rightarrow s\nu\bar{\nu}$

$B \rightarrow K^{(*)}\nu\bar{\nu}$ in the SM

Theory of $b \rightarrow s\nu\bar{\nu}$

$B \rightarrow K^{(*)}\nu\bar{\nu}$ in the SM

Confinement of quarks $\rightarrow b \rightarrow s\nu\bar{\nu}$ through $B \rightarrow K^{(*)}\nu\bar{\nu}$ decays

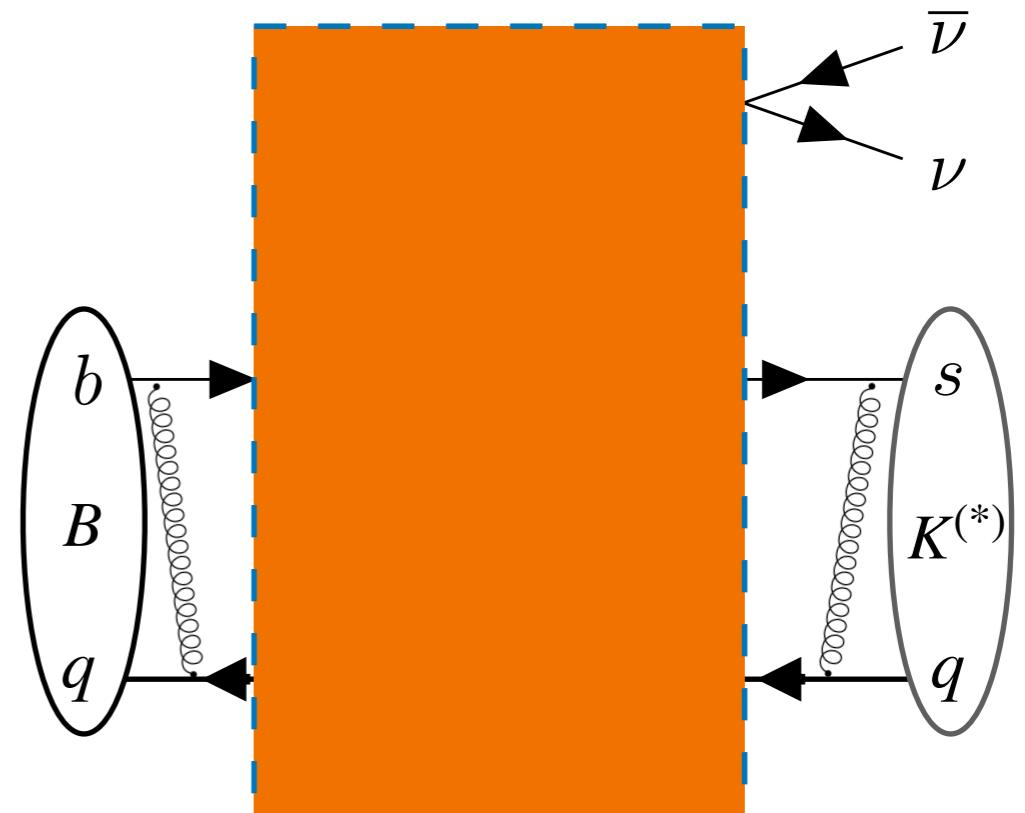


Theory of $b \rightarrow s\nu\bar{\nu}$

$B \rightarrow K^{(*)}\nu\bar{\nu}$ in the SM

Confinement of quarks $\rightarrow b \rightarrow s\nu\bar{\nu}$ through $B \rightarrow K^{(*)}\nu\bar{\nu}$ decays

At energy scale $\mu = \mathcal{O}(m_B)$ W, Z not resolved

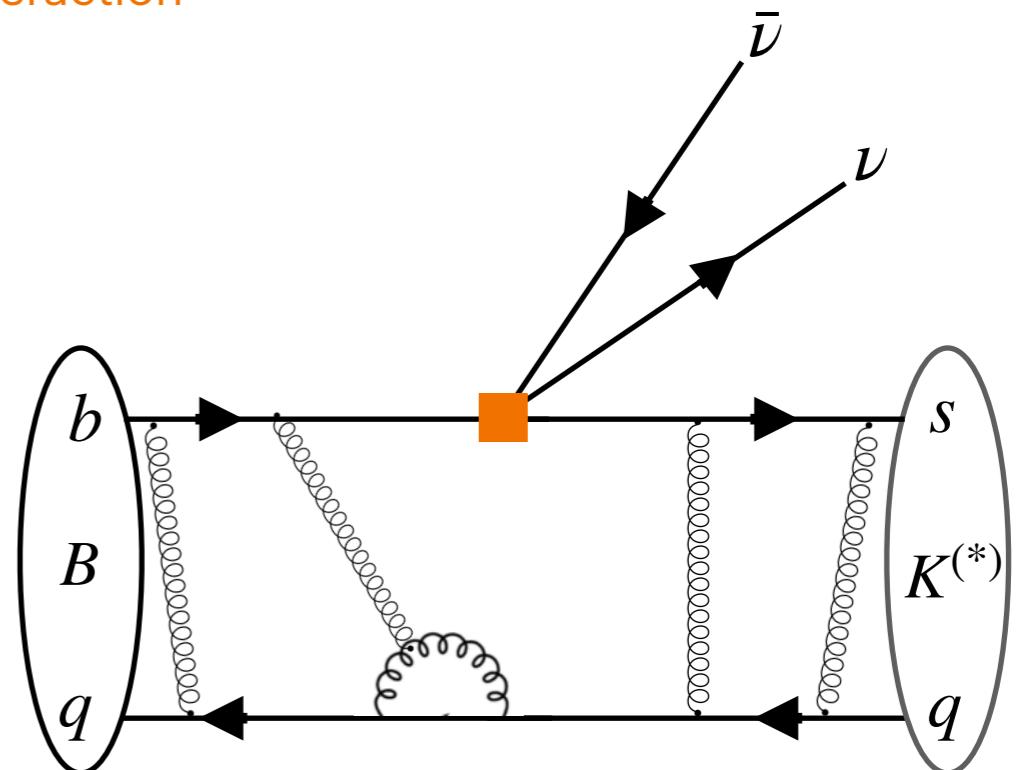


Theory of $b \rightarrow s\nu\bar{\nu}$

$B \rightarrow K^{(*)}\nu\bar{\nu}$ in the SM

Confinement of quarks $\rightarrow b \rightarrow s\nu\bar{\nu}$ through $B \rightarrow K^{(*)}\nu\bar{\nu}$ decays

At energy scale $\mu = \mathcal{O}(m_B)$ W, Z not resolved \sim point-like interaction



Theory of $b \rightarrow s\nu\bar{\nu}$

$B \rightarrow K^{(*)}\nu\bar{\nu}$ in the SM

Confinement of quarks $\rightarrow b \rightarrow s\nu\bar{\nu}$ through $B \rightarrow K^{(*)}\nu\bar{\nu}$ decays

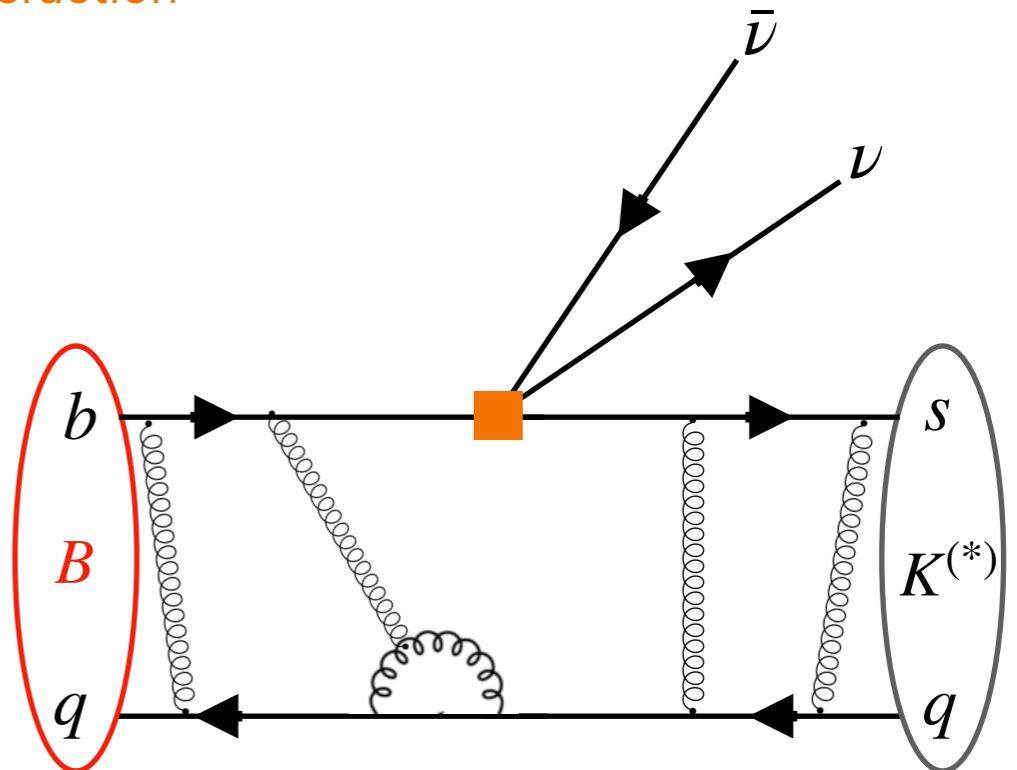
At energy scale $\mu = \mathcal{O}(m_B)$ W, Z not resolved \sim point-like interaction

Use an **effective theory**:

$$\mathcal{H}_{eff}^{SM} = -\frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* C_L^{SM} O_L + h.c.$$

\downarrow

$$\frac{2e^2}{16\pi^2} \left[\bar{s}_L \gamma_\mu b_L \right] \left[\bar{\nu}_L \gamma^\mu \nu_L \right]$$





Theory of $b \rightarrow s\nu\bar{\nu}$

$B \rightarrow K^{(*)}\nu\bar{\nu}$ in the SM

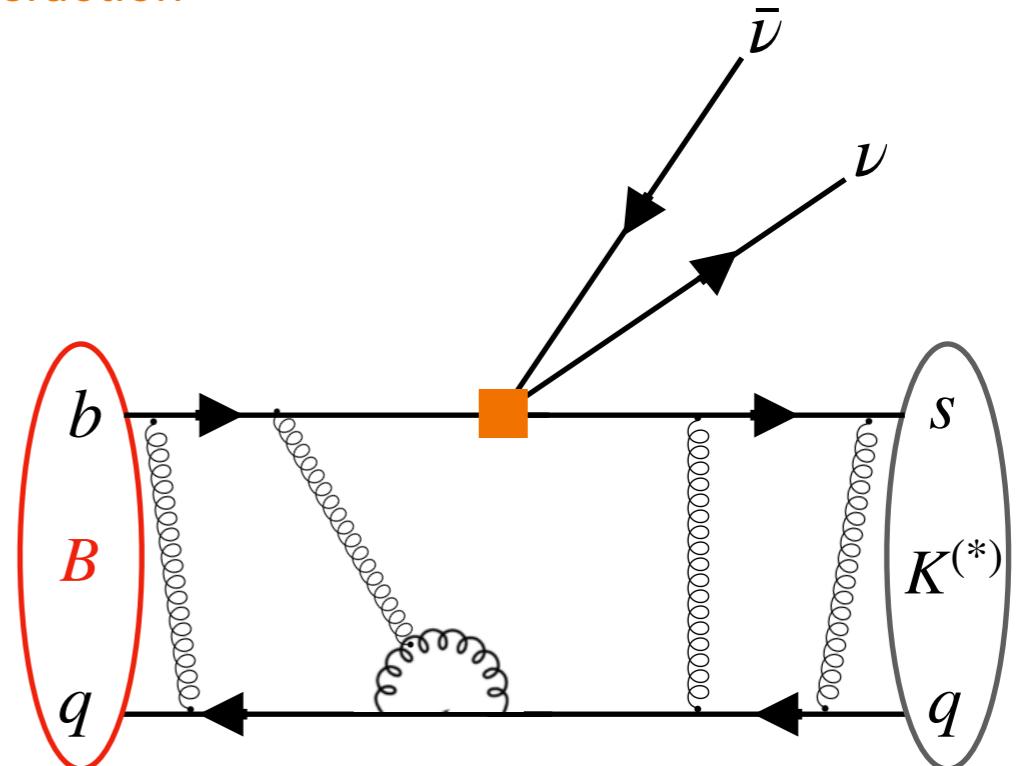
Confinement of quarks $\rightarrow b \rightarrow s\nu\bar{\nu}$ through $B \rightarrow K^{(*)}\nu\bar{\nu}$ decays

At energy scale $\mu = \mathcal{O}(m_B)$ W, Z not resolved \sim point-like interaction

Use an **effective theory**:

$$\mathcal{H}_{eff}^{SM} = -\frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* C_L^{SM} O_L + h.c.$$

$$\frac{2e^2}{16\pi^2} [\bar{s}_L \gamma_\mu b_L] [\bar{\nu}_L \gamma^\mu \nu_L]$$



Exact **factorisation**:

$$\langle K^{(*)}\nu\bar{\nu} | \mathcal{H}_{eff}^{SM} | B \rangle = -\frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{2e^2}{16\pi^2} C_L^{SM} \langle \nu\bar{\nu} | \bar{\nu}_L \gamma^\mu \nu_L | 0 \rangle \langle K^{(*)} | \bar{s}_L \gamma_\mu b_L | B \rangle$$

High-energy dynamics,
perturbation theory.

Hadronic form factors (FF),
non-perturbative contributions.

Theory of $b \rightarrow s\nu\bar{\nu}$

The $B \rightarrow K\nu\bar{\nu}$ decays

Theory of $b \rightarrow s\nu\bar{\nu}$

The $B \rightarrow K\nu\bar{\nu}$ decays

Pseudo-scalar ($J^P = 0^-$) $B^+(B^0)$ meson \longrightarrow pseudo-scalar $K^+(K^0)$ meson + neutrino pair.

Theory of $b \rightarrow s\nu\bar{\nu}$

The $B \rightarrow K\nu\bar{\nu}$ decays

Pseudo-scalar ($J^P = 0^-$) $B^+(B^0)$ meson \longrightarrow pseudo-scalar $K^+(K^0)$ meson + neutrino pair.

$$q^2 = (p_\nu + p_{\bar{\nu}})^2 \rightarrow$$

Theory of $b \rightarrow s\nu\bar{\nu}$

The $B \rightarrow K\nu\bar{\nu}$ decays

Pseudo-scalar ($J^P = 0^-$) $B^+(B^0)$ meson \longrightarrow pseudo-scalar $K^+(K^0)$ meson + neutrino pair.

[J. High Energ. Phys. 2015, 184]

$$q^2 = (p_\nu + p_{\bar{\nu}})^2 \rightarrow \frac{d\text{BR}(B \rightarrow K\nu\bar{\nu})_{\text{SM}}}{dq^2} = \tau_B 3 \left| V_{tb} V_{ts}^* \frac{G_F \alpha}{16\pi^2} \sqrt{\frac{m_B}{3\pi}} \right|^2 |C_L^{\text{SM}}|^2 \rho_K(q^2)$$

|||

- $\rho_K(q^2)$: $B \rightarrow K$ form factor.
- 3: neutrino flavours.
- τ_B, m_B : B-meson lifetime and mass.

$|N|^2$

Theory of $b \rightarrow s\nu\bar{\nu}$

The $B \rightarrow K\nu\bar{\nu}$ decays

Pseudo-scalar ($J^P = 0^-$) $B^+(B^0)$ meson \longrightarrow pseudo-scalar $K^+(K^0)$ meson + neutrino pair.

[J. High Energ. Phys. 2015, 184]

$$q^2 = (p_\nu + p_{\bar{\nu}})^2 \rightarrow \frac{d\text{BR}(B \rightarrow K\nu\bar{\nu})_{\text{SM}}}{dq^2} = \tau_B 3 \left| N \cdot C_L^{\text{SM}} \right|^2 \rho_K(q^2)$$

- $\rho_K(q^2)$: $B \rightarrow K$ form factor.
- 3: neutrino flavours.
- τ_B, m_B : B -meson lifetime and mass.



Theory of $b \rightarrow s\nu\bar{\nu}$

The $B \rightarrow K\nu\bar{\nu}$ decays

Pseudo-scalar ($J^P = 0^-$) $B^+(B^0)$ meson \longrightarrow pseudo-scalar $K^+(K^0)$ meson + neutrino pair.

[J. High Energ. Phys. 2015, 184]

$$q^2 = (p_\nu + p_{\bar{\nu}})^2 \rightarrow \frac{d\text{BR}(B \rightarrow K\nu\bar{\nu})_{\text{SM}}}{dq^2} = \tau_B 3 \left| N \cdot C_L^{\text{SM}} \right|^2 \rho_K(q^2)$$

- $\rho_K(q^2)$: $B \rightarrow K$ form factor.
- 3: neutrino flavours.
- τ_B , m_B : B -meson lifetime and mass.

Total branching fractions:

[Prog. Part. Nucl. Phys. 2017, 91]

$$\text{BR}(B^+ \rightarrow K^+\nu\bar{\nu})_{\text{SM}} = (4.6 \pm 0.5) \times 10^{-6}$$

$$\text{BR}(B^0 \rightarrow K^0\nu\bar{\nu})_{\text{SM}} = (4.3 \pm 0.5) \times 10^{-6}$$

Difference driven by B^+/B^0 lifetimes.

Uncertainty dominated by error on $B \rightarrow K$ form factor.

Theory of $b \rightarrow s\nu\bar{\nu}$

The $B \rightarrow K^*\nu\bar{\nu}$ decays

Theory of $b \rightarrow s\nu\bar{\nu}$

The $B \rightarrow K^*\nu\bar{\nu}$ decays

Pseudo-scalar ($J^P = 0^-$) $B^+(B^0)$ meson \longrightarrow vector ($J^P = 1^-$) $K^{*+}(K^{*0})$ meson + neutrino pair.

Theory of $b \rightarrow s\nu\bar{\nu}$

The $B \rightarrow K^*\nu\bar{\nu}$ decays

Pseudo-scalar ($J^P = 0^-$) $B^+(B^0)$ meson \longrightarrow vector ($J^P = 1^-$) $K^{*+}(K^{*0})$ meson + neutrino pair.

[J. High Energ. Phys. 2015, 184]

$$\frac{d\text{BR}(B \rightarrow K^*\nu\bar{\nu})_{\text{SM}}}{dq^2} = \tau_B 3 \left| N \cdot C_L^{\text{SM}} \right|^2 \left[\rho_{A_{12}}(q^2) + \rho_V(q^2) + \rho_{A_1}(q^2) \right]$$

- $\rho_{A_{12}}(q^2), \rho_V(q^2), \rho_{A_1}(q^2)$: $B \rightarrow K^*$ form factors.



Theory of $b \rightarrow s\nu\bar{\nu}$

The $B \rightarrow K^*\nu\bar{\nu}$ decays

Pseudo-scalar ($J^P = 0^-$) $B^+(B^0)$ meson \longrightarrow vector ($J^P = 1^-$) $K^{*+}(K^{*0})$ meson + neutrino pair.

[J. High Energ. Phys. 2015, 184]

$$\frac{d\text{BR}(B \rightarrow K^*\nu\bar{\nu})_{\text{SM}}}{dq^2} = \tau_B 3 \left| N \cdot C_L^{\text{SM}} \right|^2 \left[\rho_{A_{12}}(q^2) + \rho_V(q^2) + \rho_{A_1}(q^2) \right]$$

- $\rho_{A_{12}}(q^2), \rho_V(q^2), \rho_{A_1}(q^2)$: $B \rightarrow K^*$ form factors.

Total SM branching fractions:

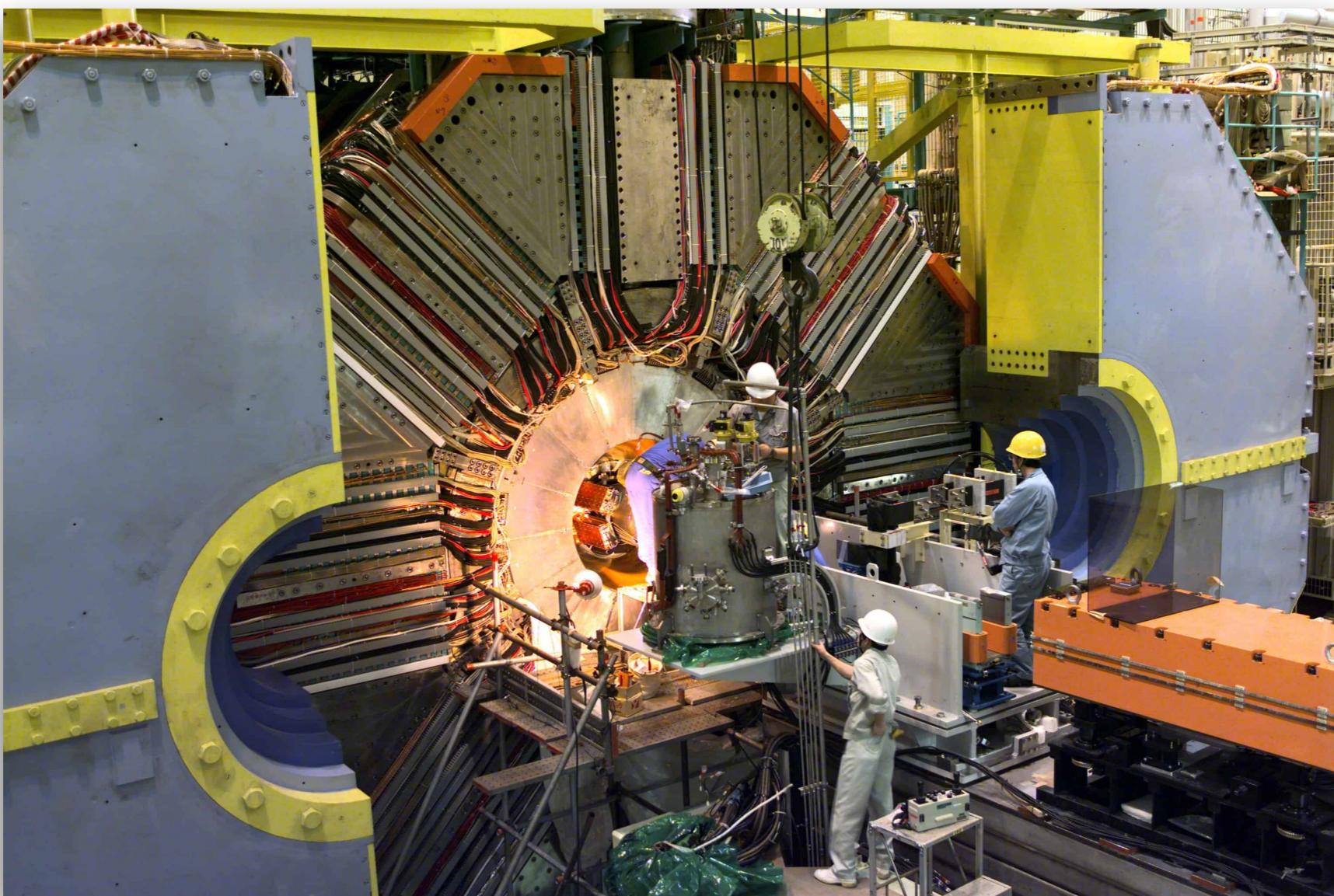
[Prog. Part. Nucl. Phys. 2017, 91]

$$\text{BR}(B^+ \rightarrow K^{*+}\nu\bar{\nu})_{\text{SM}} = (8.4 \pm 1.5) \times 10^{-6}$$

$$\text{BR}(B^0 \rightarrow K^{*0}\nu\bar{\nu})_{\text{SM}} = (7.8 \pm 1.4) \times 10^{-6}$$

Uncertainty dominated by error on $B \rightarrow K^*$ form factors.

Search at Belle II



The Belle II experiment

SuperKEKB



The Belle II experiment

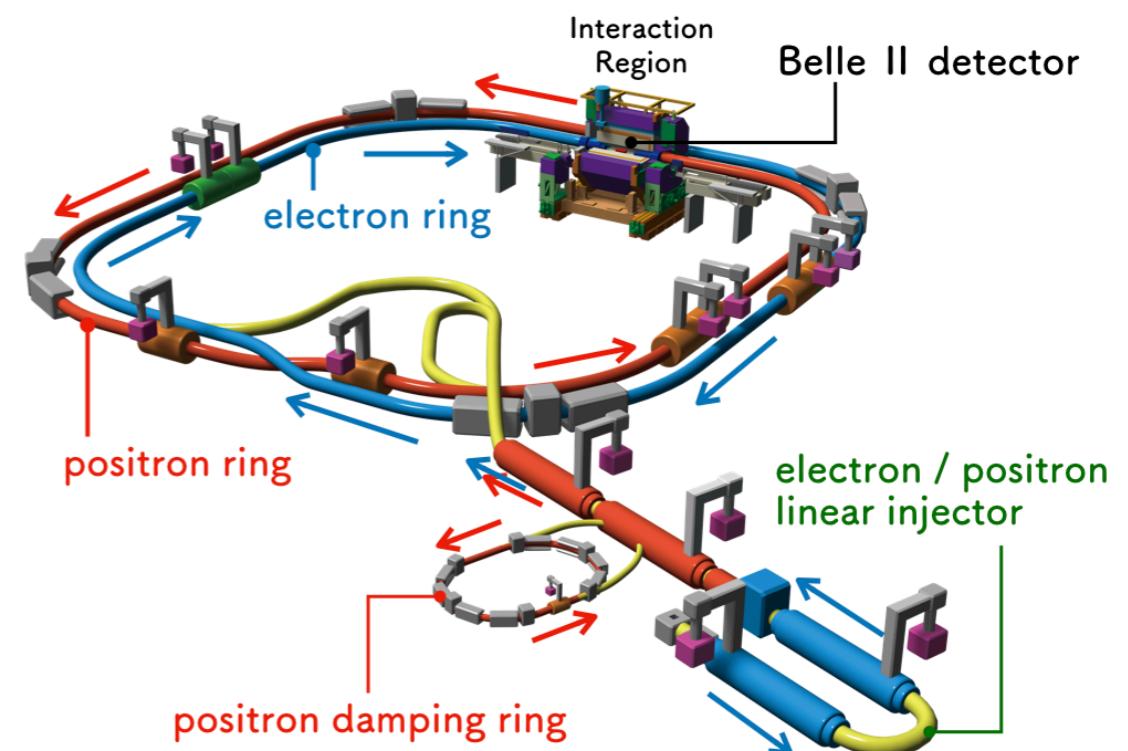
SuperKEKB

- Asymmetric-energy $e^+(4\text{ GeV}) e^-(7\text{ GeV})$ collider.
- Operates at $\sqrt{s} = 10.58\text{ GeV} \rightarrow \Upsilon(4S)$ resonance mass.

A (super) B -factory:

- $\sigma_{\Upsilon(4S)} \simeq 1.1\text{ nb}$
- $\text{BR}(\Upsilon(4S) \rightarrow B\bar{B}) \sim 100\%$

millions of B mesons at low background.

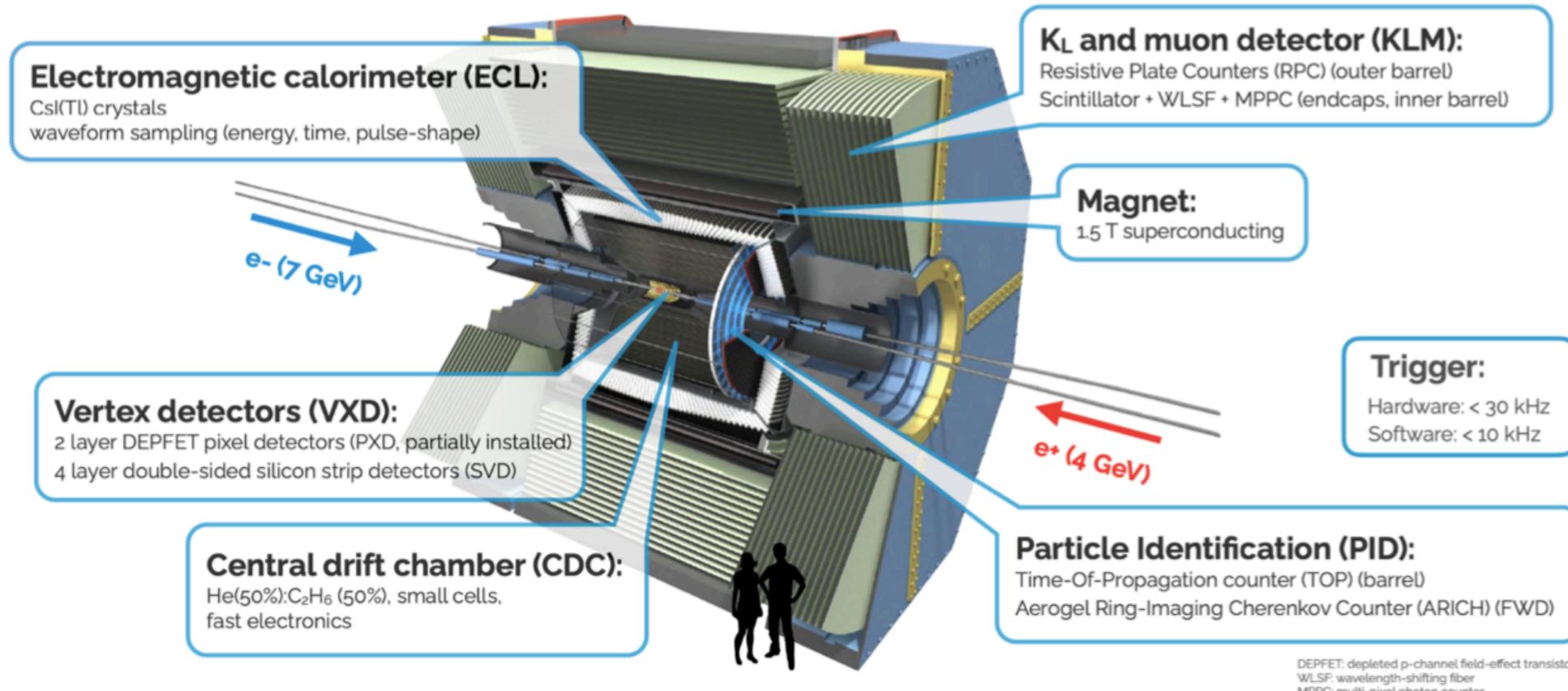


The Belle II experiment

Belle II detector

The Belle II experiment

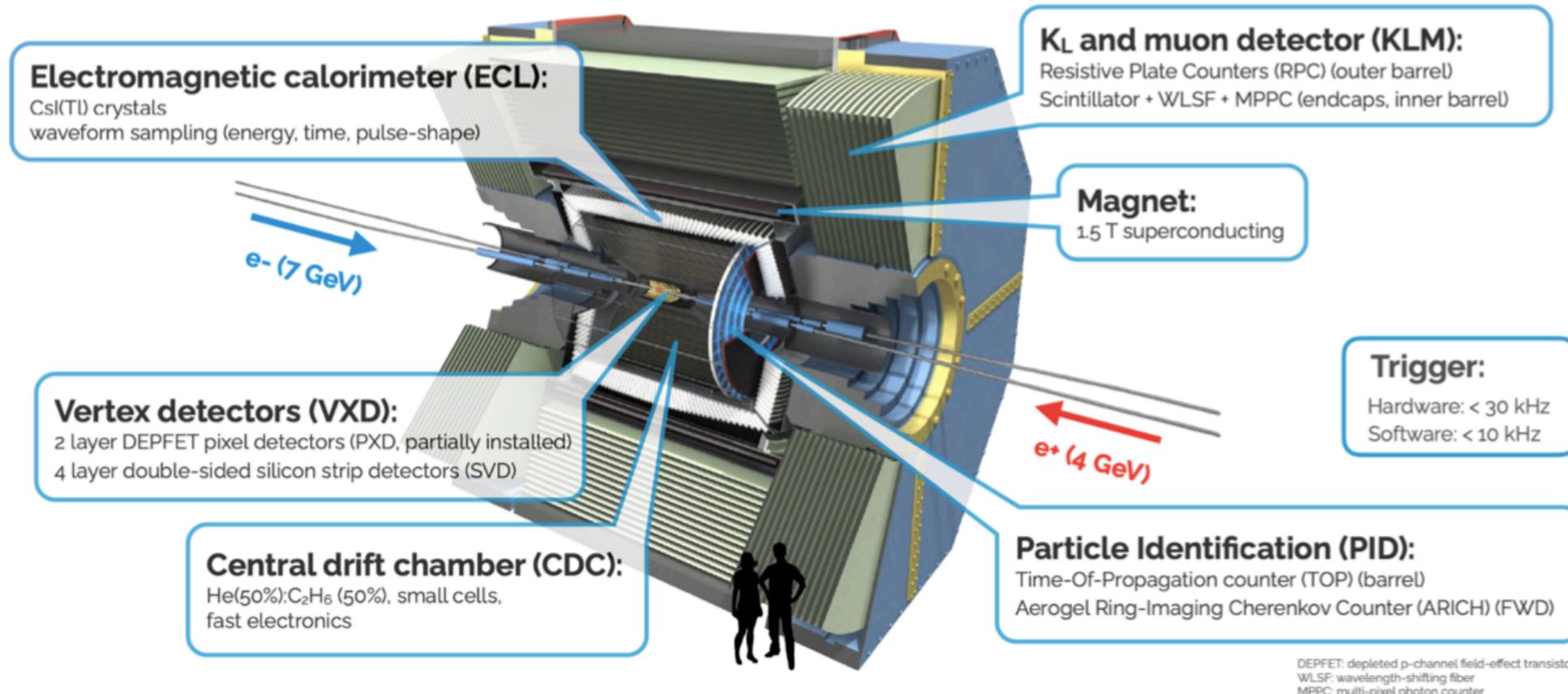
Belle II detector





The Belle II experiment

Belle II detector



- Suitable for $B \rightarrow K^{(*)}\nu\bar{\nu}$ searches:**
- **Nearly 4π coverage.**
 - **Clean environment** (~ 11 tracks per $\Upsilon(4S)$ event on average).
 - **Precise knowledge of the initial state.**

$B \rightarrow K^{(*)} \nu \bar{\nu}$ searches

Experimental review

$B \rightarrow K^{(*)} \nu \bar{\nu}$ searches

Experimental review

Challenge: 2 neutrinos in the final state
→ missing experimental signatures.

Common strategy: tagging

indirect info on B_{sig} from partner (non-signal) B_{tag}

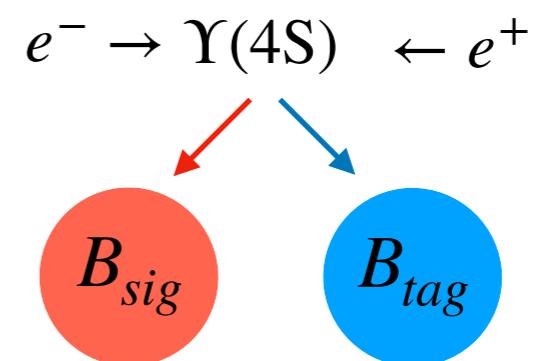
$B \rightarrow K^{(*)}\nu\bar{\nu}$ searches

Experimental review

Challenge: 2 neutrinos in the final state
→ missing experimental signatures.

Common strategy: tagging

indirect info on B_{sig} from partner (non-signal) B_{tag}



$B \rightarrow K^{(*)}\nu\bar{\nu}$ searches

Experimental review

Challenge: 2 neutrinos in the final state
→ missing experimental signatures.

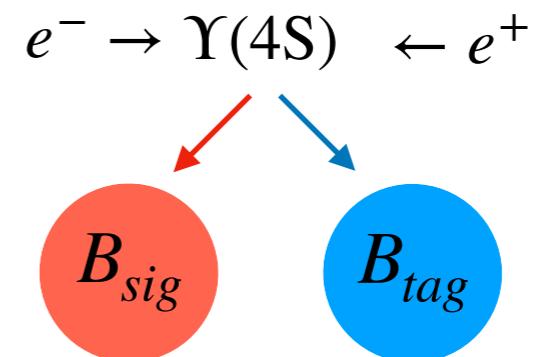
Common strategy: tagging

indirect info on B_{sig} from partner (non-signal) B_{tag}

Previous searches

Hadronic or/and semileptonic tagging.

Low tag reconstruction efficiency.



Tagging techniques		
	Purity	Efficiency
Inclusive $B \rightarrow \text{anything}$ $\epsilon \approx \mathcal{O}(100\%)$		
Semileptonic $B \rightarrow D^{(*)}\ell\nu_\ell$ $\epsilon \approx \mathcal{O}(1\%)$		
Hadronic $B \rightarrow \text{hadrons}$ $\epsilon \approx \mathcal{O}(0.1\%)$		



$B \rightarrow K^{(*)}\nu\bar{\nu}$ searches

Experimental review

Challenge: 2 neutrinos in the final state
 → missing experimental signatures.

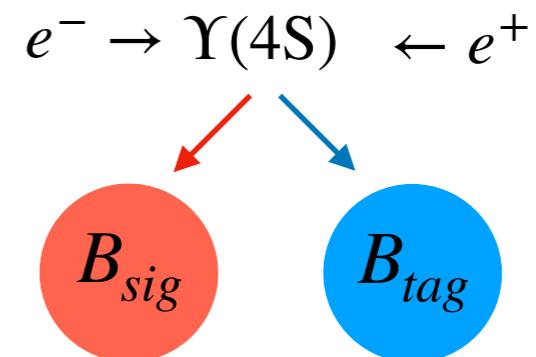
Common strategy: tagging

indirect info on B_{sig} from partner (non-signal) B_{tag}

Previous searches

Hadronic or/and semileptonic tagging.

Low tag reconstruction efficiency.



Tagging techniques		
Inclusive	Semileptonic	Hadronic
$B \rightarrow \text{anything}$	$B \rightarrow D^{(*)}\ell\nu_\ell$	$B \rightarrow \text{hadrons}$
$\epsilon \approx \mathcal{O}(100\%)$	$\epsilon \approx \mathcal{O}(1\%)$	$\epsilon \approx \mathcal{O}(0.1\%)$

Results: no statistically significant signal observed.

Measurement	Observed limits		Tagging	Data[fb ⁻¹]
	$B^+ \rightarrow K^+\nu\bar{\nu}$	$B^0 \rightarrow K^{*0}\nu\bar{\nu}$		
BaBar (2013) [Phys. Rev. D 2013, 87, 112005]	$< 1.6 \times 10^{-5}$	$< 12 \times 10^{-5}$	semileptonic + hadronic	429
Belle (2013) [Phys. Rev. D 2013, 87, 111103]	$< 5.5 \times 10^{-5}$	$< 5.5 \times 10^{-5}$	hadronic	711
Belle (2017) [Phys. Rev. D 2017, 87, 112005]	$< 1.9 \times 10^{-5}$	$< 1.8 \times 10^{-5}$	semileptonic	711

Search for $B^+ \rightarrow K^+ \nu \bar{\nu}$

with an inclusive tagging

at Belle II

Published in Physical Review Letters [Phys. Rev. Lett. 2021, 127]

Data samples

Data samples

Collision data

In terms of recorded integrated luminosity:

- **on-resonance** sample of 63 fb^{-1}
- **off-resonance** sample of 9 fb^{-1}

Data samples

Collision data

In terms of recorded integrated luminosity:

- **on-resonance** sample of 63 fb^{-1}
- **off-resonance** sample of 9 fb^{-1}

Simulated data



Data samples

Collision data

In terms of recorded integrated luminosity:

- **on-resonance** sample of 63 fb^{-1}
- **off-resonance** sample of 9 fb^{-1}

Simulated data

- **Signal Monte Carlo (MC):**

$e^+e^- \rightarrow \Upsilon(4S) \rightarrow B^+(\rightarrow K^+\nu\bar{\nu}) B^-$ events **following SM expectation.** [J. High Energ. Phys. 2015, 184]

- **Background MC:**

- generic $B^0\bar{B}^0$ and B^+B^- decays from the $\Upsilon(4S)$;
- continuum: $e^+e^- \rightarrow q\bar{q}$ ($q = u, d, s, c$) and $e^+e^- \rightarrow \tau^+\tau^-$ events.

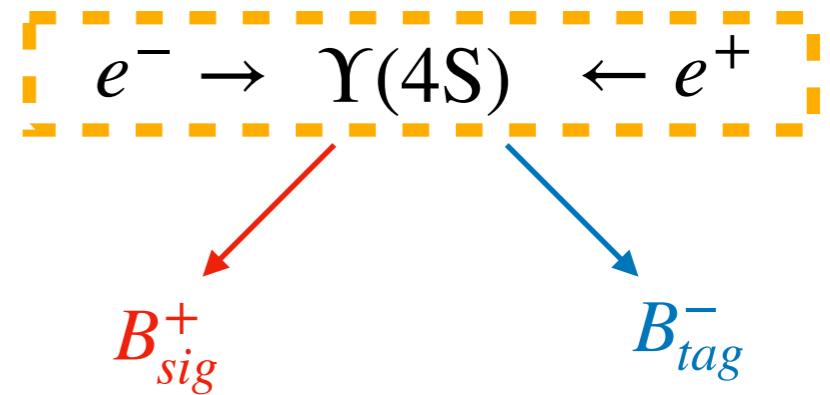
The inclusive tagging

Our method

The inclusive tagging

Our method

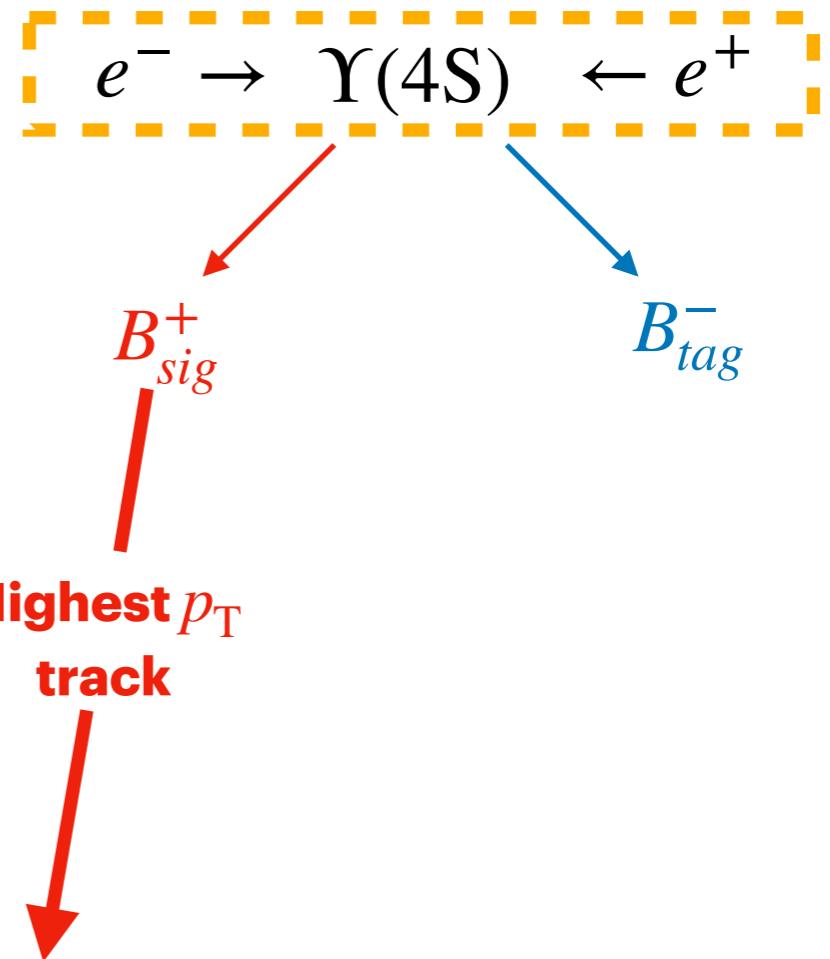
- **Step 1:** signal reconstruction.



The inclusive tagging

Our method

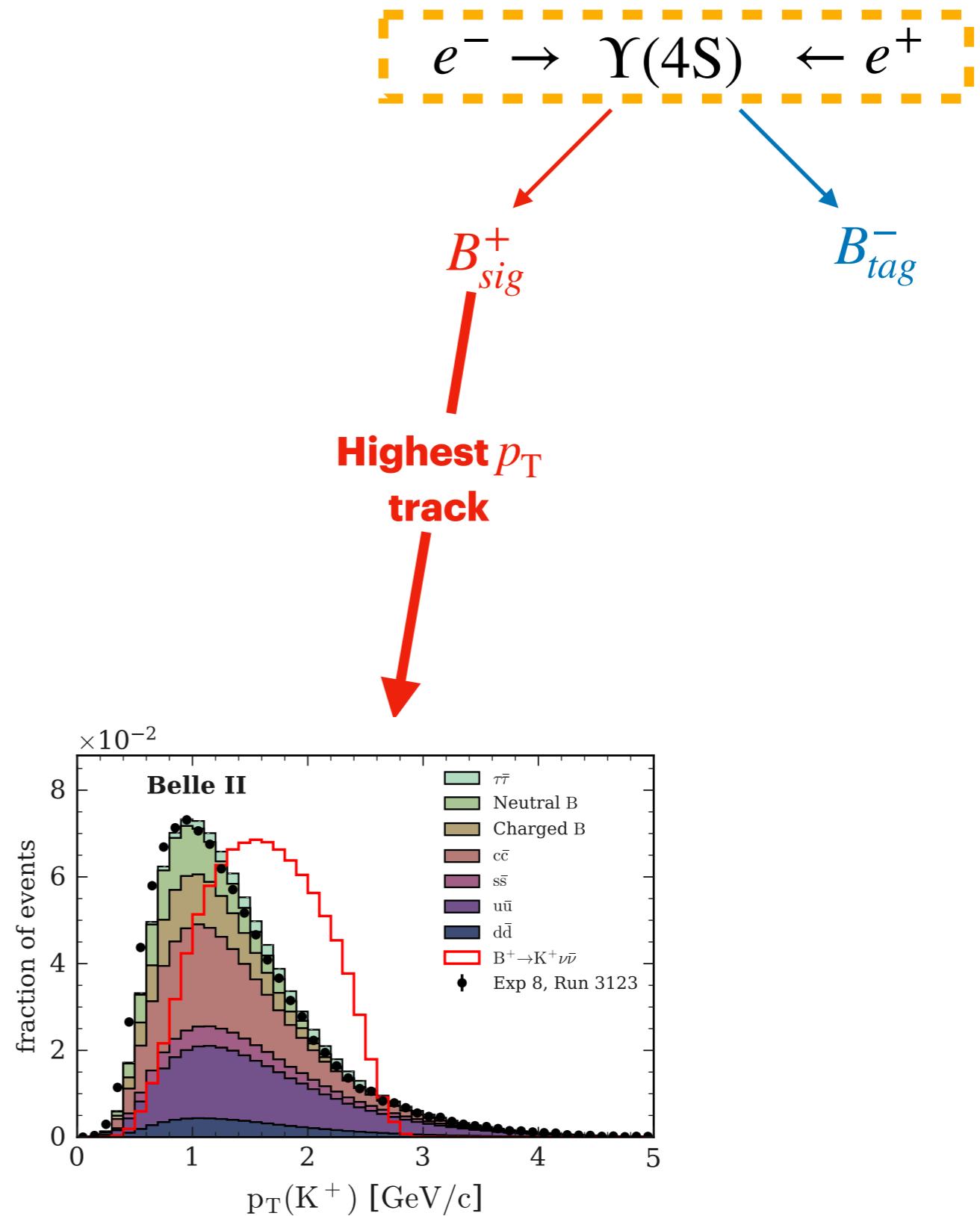
- **Step 1:** signal reconstruction.
Select highest p_T track



The inclusive tagging

Our method

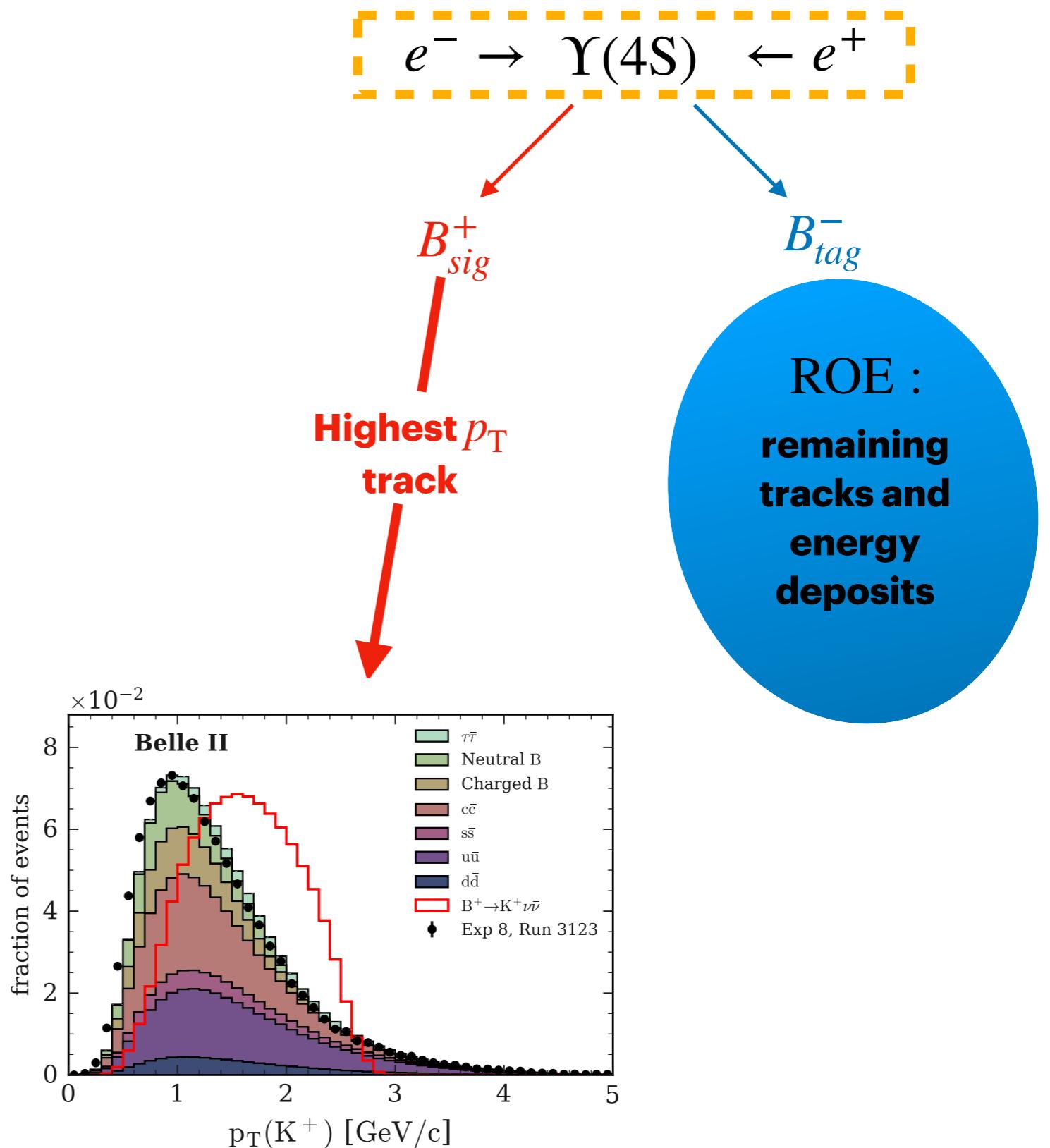
- **Step 1:** signal reconstruction.
 - Select highest p_T track
 - correct match $\sim 80\%$



The inclusive tagging

Our method

- **Step 1:** signal reconstruction.
Select highest p_T track
 \rightarrow correct match $\sim 80\%$
- **Step 2:** inclusive reconstruction
of the *rest of the event* (ROE).





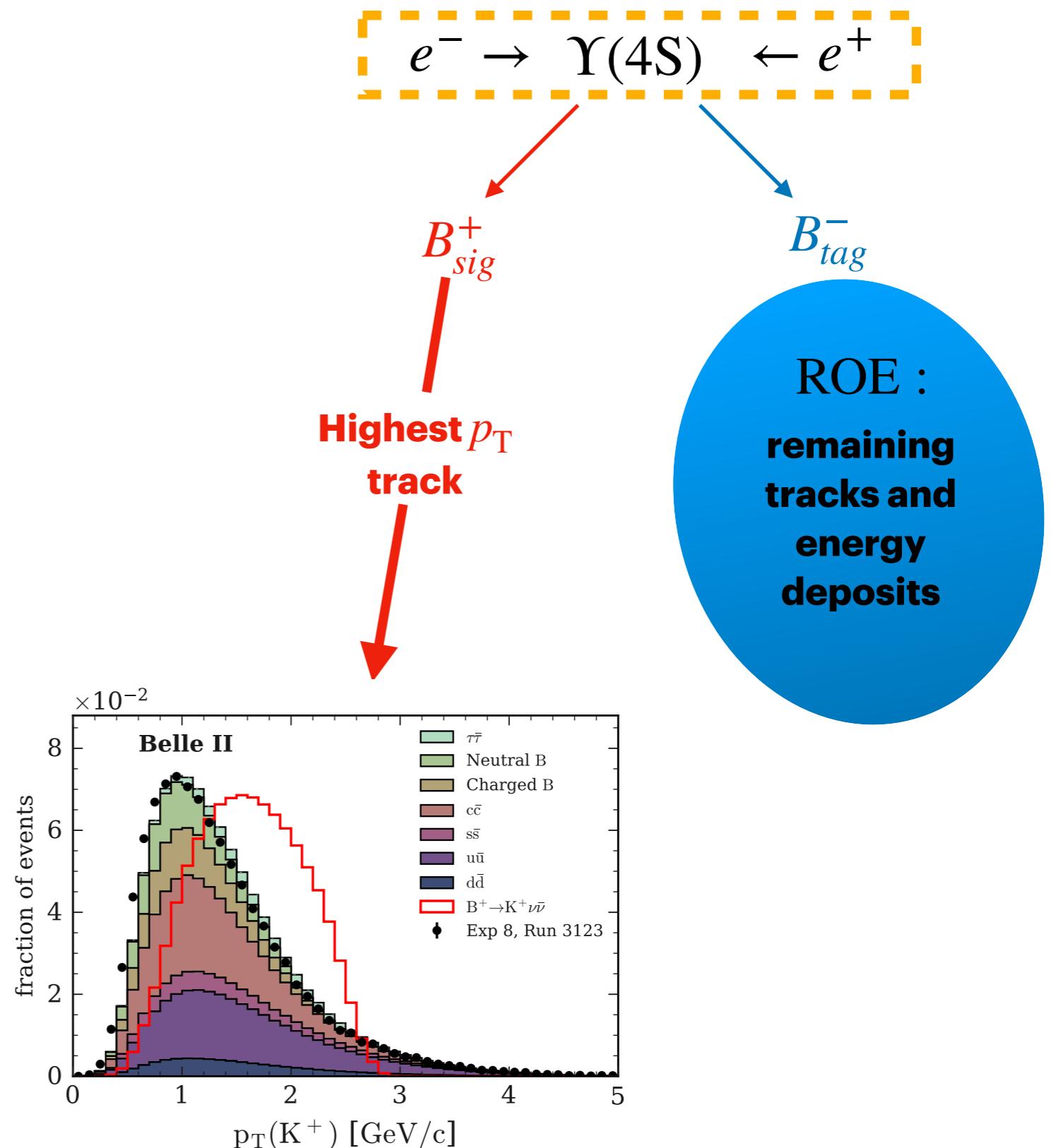
The inclusive tagging

Our method

- **Step 1:** signal reconstruction.
Select highest p_T track
→ correct match $\sim 80\%$
- **Step 2:** inclusive reconstruction
of the *rest of the event* (ROE).

Higher signal efficiency

(up to $\epsilon_{sig} \sim 4\%$ in the signal region)
but larger background (bkg).



Background suppression

Features of $B^+ \rightarrow K^+ \nu \bar{\nu}$

Background suppression

Features of $B^+ \rightarrow K^+ \nu \bar{\nu}$

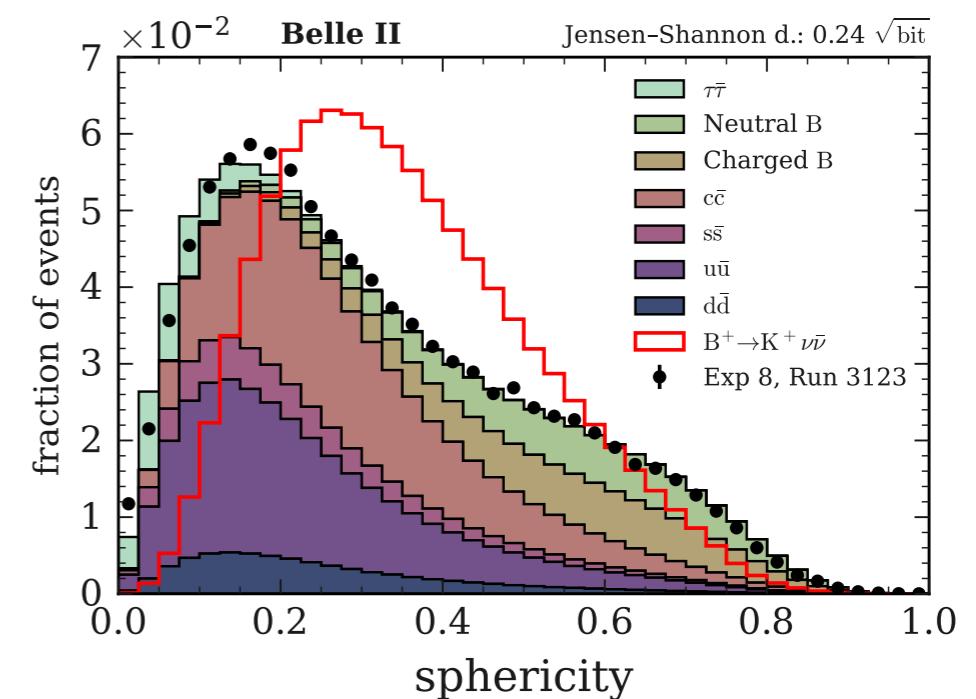
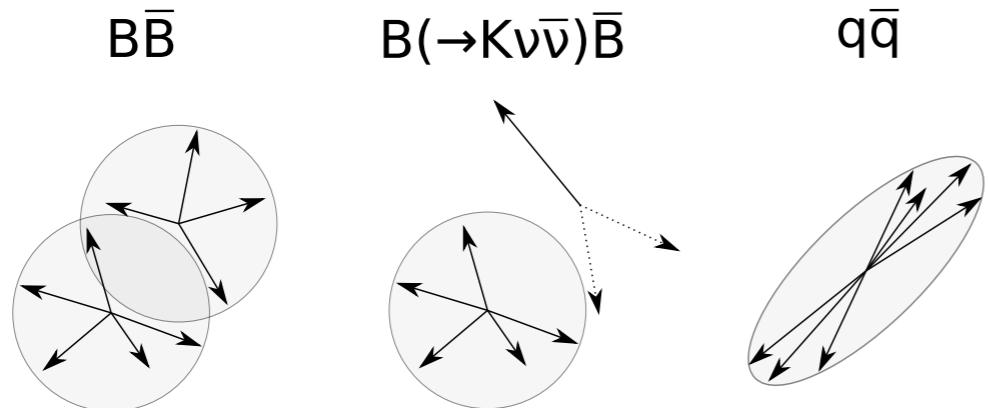
Variables capturing the characteristic features of $B^+ \rightarrow K^+ \nu \bar{\nu}$ events.

Background suppression

Features of $B^+ \rightarrow K^+ \nu \bar{\nu}$

Variables capturing the characteristic features of $B^+ \rightarrow K^+ \nu \bar{\nu}$ events.

- **Event-shape** variables



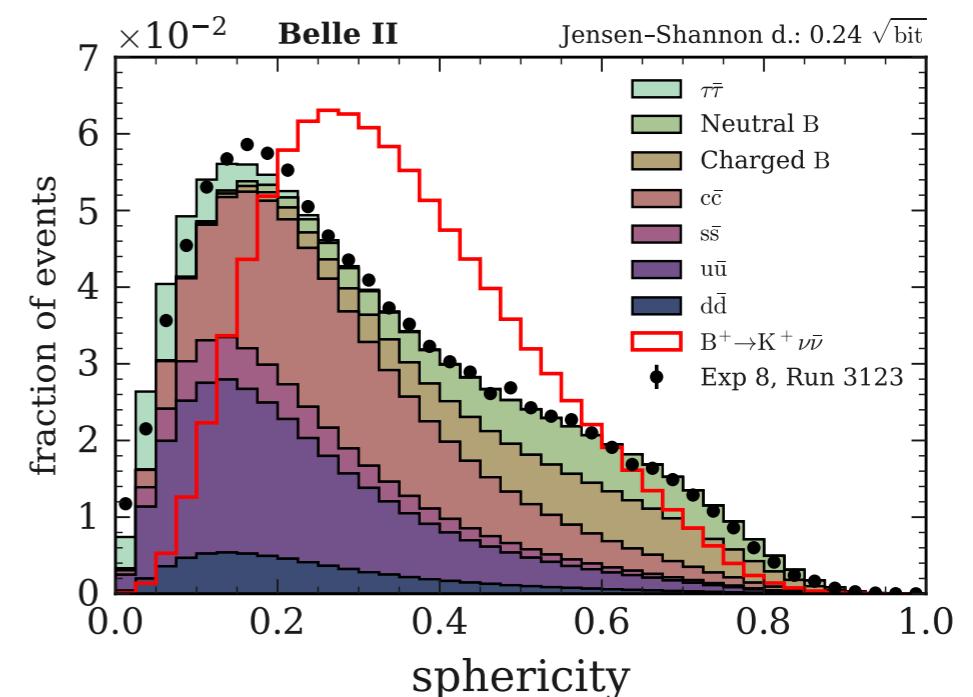
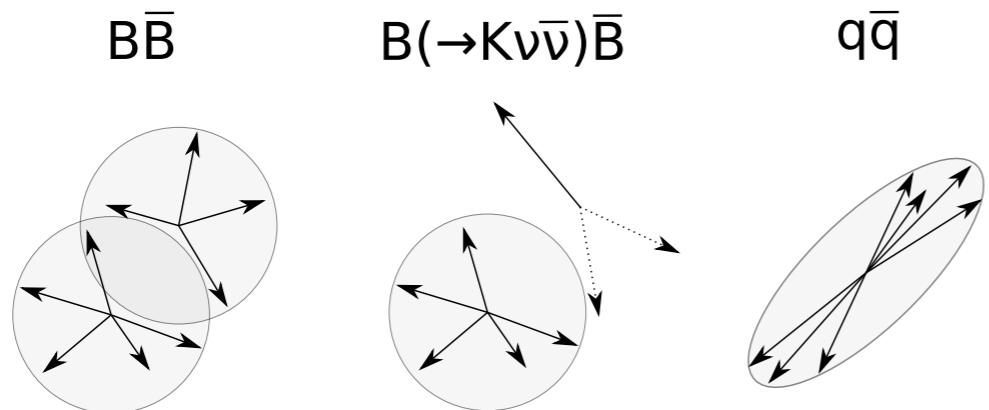


Background suppression

Features of $B^+ \rightarrow K^+ \nu \bar{\nu}$

Variables capturing the characteristic features of $B^+ \rightarrow K^+ \nu \bar{\nu}$ events.

- **Event-shape** variables



- Variables related to the **signal-kaon candidate**.
- Variables related to **ROE tracks, vertex, and energy deposits**.
- Variables **combining signal candidate and ROE**.

Multivariate classification

A classifier cascade

Multivariate classification

A classifier cascade

51 variables combined in **2 consecutive binary classifiers** (boosted decision trees): BDT_1 and BDT_2

Multivariate classification

A classifier cascade

51 variables combined in **2 consecutive binary classifiers** (boosted decision trees): BDT_1 and BDT_2 .

- train BDT_1 on simulated signal and bkg events:

Multivariate classification

A classifier cascade

51 variables combined in **2 consecutive binary classifiers** (boosted decision trees): BDT_1 and BDT_2 .

- train BDT_1 on simulated signal and bkg events:

INITIAL BACKGROUND SUPPRESSION

Multivariate classification

A classifier cascade

51 variables combined in **2 consecutive binary classifiers** (boosted decision trees): BDT_1 and BDT_2 .

- train BDT_1 on simulated signal and bkg events:

INITIAL BACKGROUND SUPPRESSION

- train BDT_2 - same features - **on events with $BDT_1 > 0.9$** reconstructed in simulated signal and bkg samples.

Multivariate classification

A classifier cascade

51 variables combined in **2 consecutive binary classifiers** (boosted decision trees): BDT_1 and BDT_2 .

- train BDT_1 on simulated signal and bkg events:

INITIAL BACKGROUND SUPPRESSION

- train BDT_2 - same features - **on events with $BDT_1 > 0.9$** [$\rightarrow 28\%$ efficiency, 0.02% purity]
reconstructed in simulated signal and bkg samples.

Multivariate classification

A classifier cascade

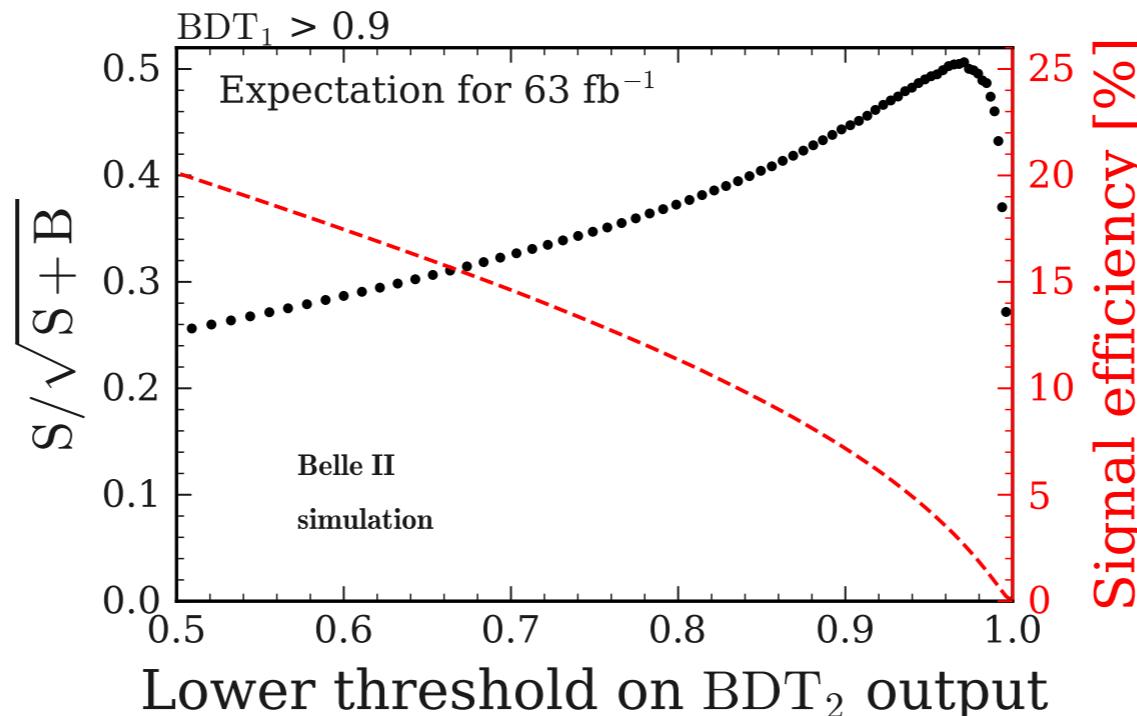
51 variables combined in **2 consecutive binary classifiers** (boosted decision trees): BDT_1 and BDT_2 .

- train BDT_1 on simulated signal and bkg events:

INITIAL BACKGROUND SUPPRESSION

- train BDT_2 - same features - **on events with $BDT_1 > 0.9$** [$\rightarrow 28\%$ efficiency, 0.02% purity]
reconstructed in simulated signal and bkg samples.

+10% EXPECTED SIGNIFICANCE: UP TO ~ 0.5





Multivariate classification

A classifier cascade

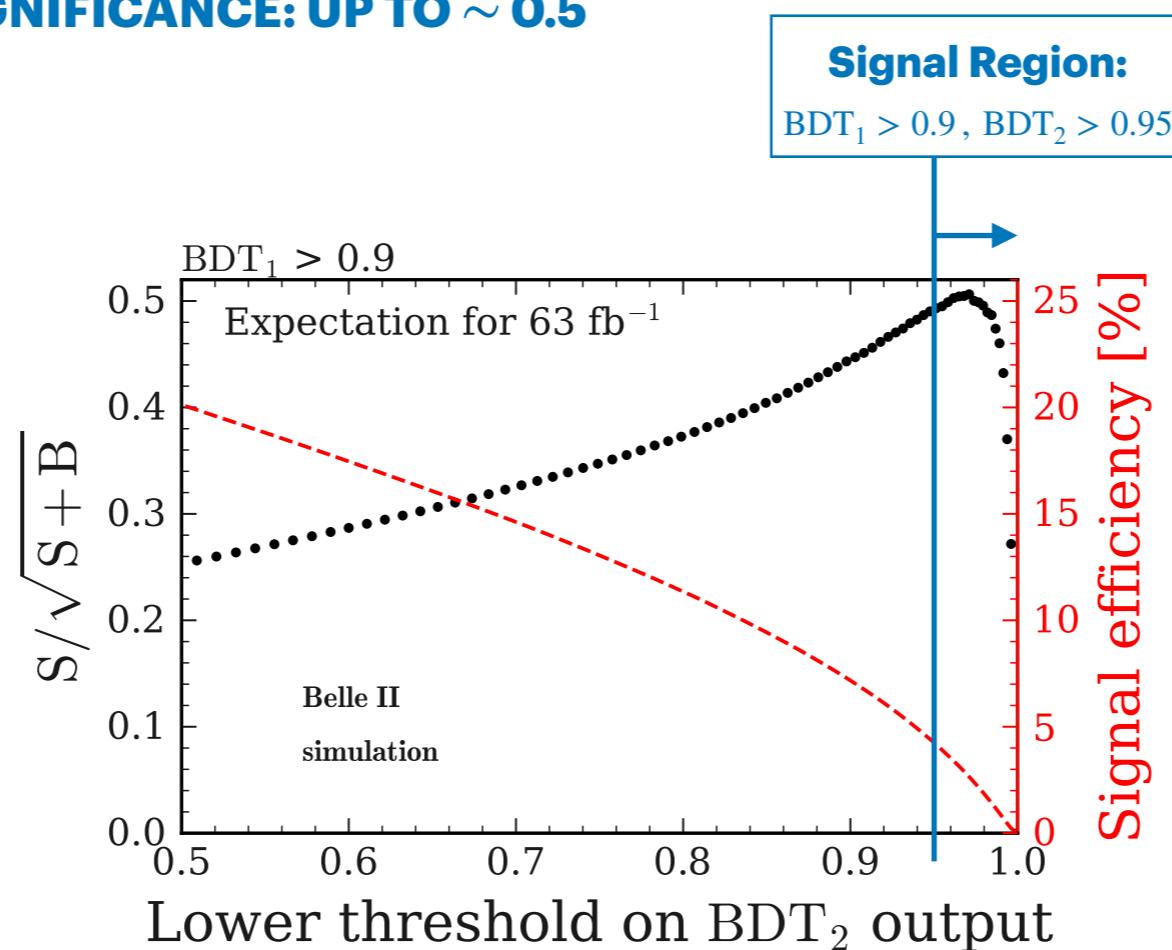
51 variables combined in **2 consecutive binary classifiers** (boosted decision trees): BDT_1 and BDT_2 .

- train BDT_1 on simulated signal and bkg events:

INITIAL BACKGROUND SUPPRESSION

- train BDT_2 - same features - **on events with $BDT_1 > 0.9$** [$\rightarrow 28\%$ efficiency, 0.02% purity]
reconstructed in simulated signal and bkg samples.

+10% EXPECTED SIGNIFICANCE: UP TO ~ 0.5



Signal and control regions

Optimised bin boundaries

Signal and control regions

Optimised bin boundaries

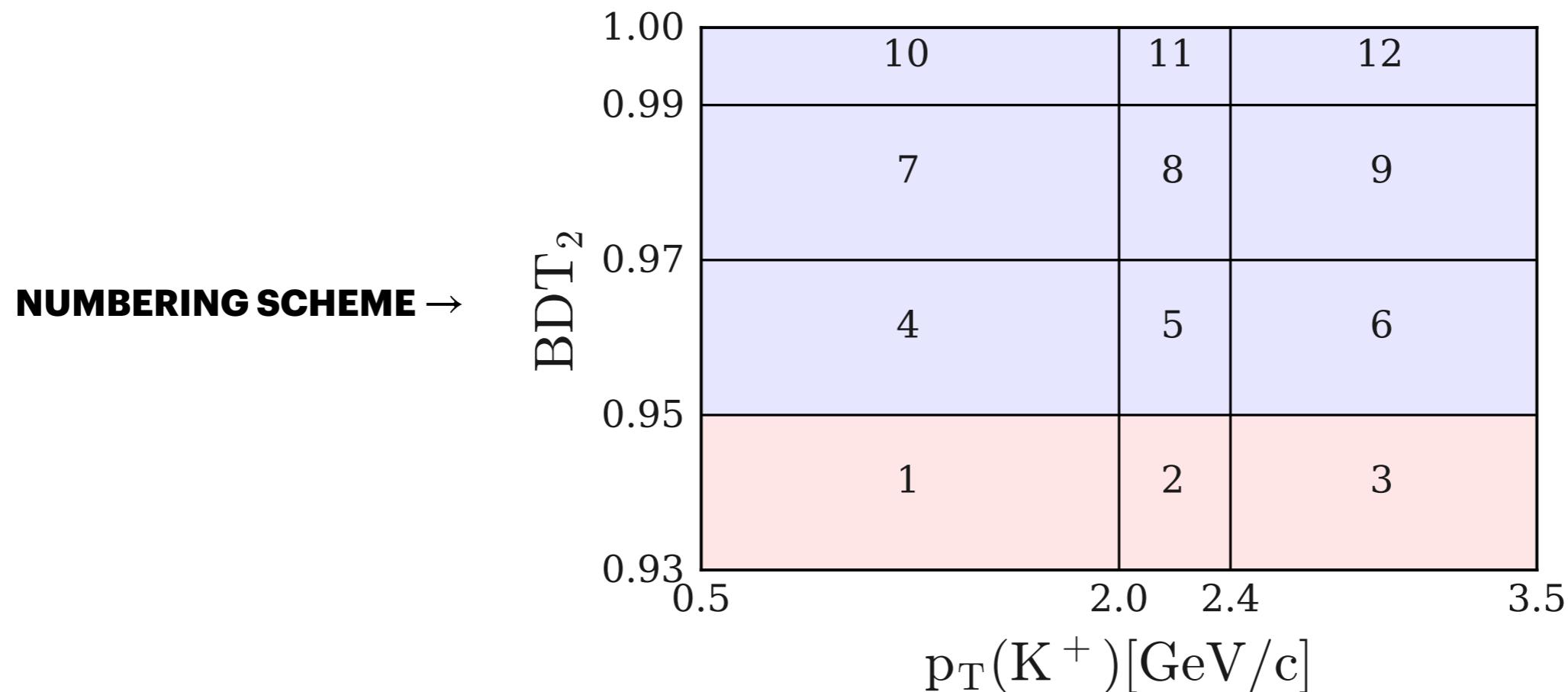
Goal: maximise separation of signal from background.



Signal and control regions

Optimised bin boundaries

Goal: maximise separation of signal from background.



Validation studies

Control channel: $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$

Validation studies

Control channel: $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$

Sizeable branching fraction and clean experimental signature.

Use it to validate the modelling of training variables
and classification outputs with on-resonance data.

Validation studies

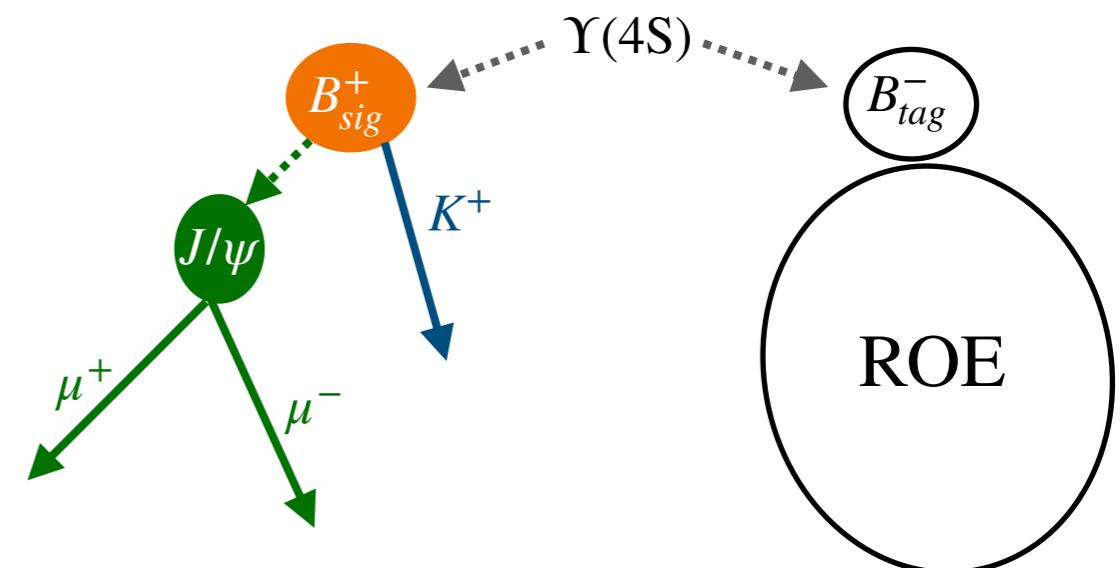
Control channel: $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$

Sizeable branching fraction and clean experimental signature.

Use it to validate the modelling of training variables
and classification outputs with on-resonance data.

The method

Identification of $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$ events



Validation studies

Control channel: $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$

Sizeable branching fraction and clean experimental signature.

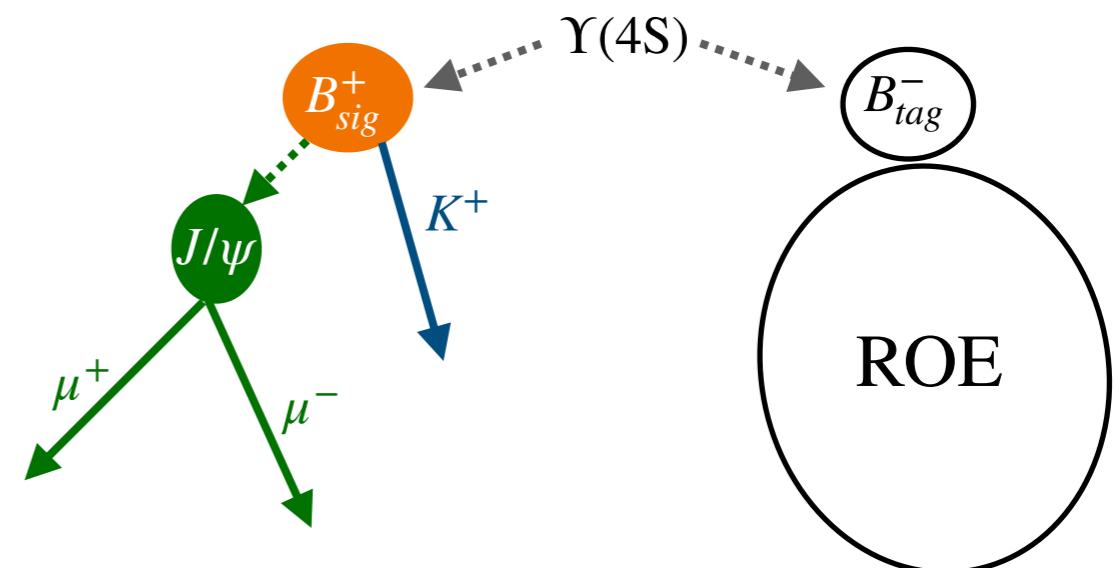
Use it to validate the modelling of training variables
and classification outputs with on-resonance data.

The method

Identification of $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$ events



Modification of the events to reproduce experimental
and kinematic features of $B^+ \rightarrow K^+ \nu \bar{\nu}$.



Validation studies

Control channel: $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$

Sizeable branching fraction and clean experimental signature.

Use it to validate the modelling of training variables
and classification outputs with on-resonance data.

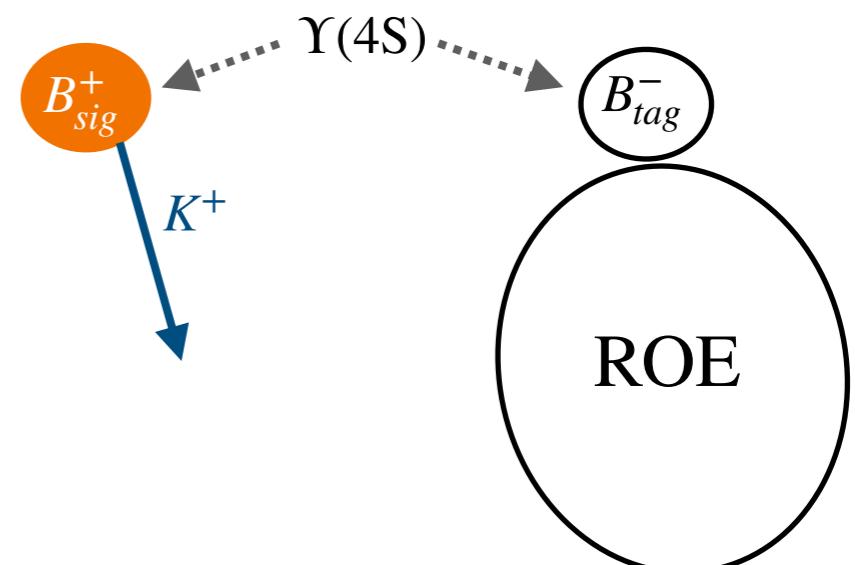
The method

Identification of $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$ events



Modification of the events to reproduce experimental
and kinematic features of $B^+ \rightarrow K^+ \nu \bar{\nu}$.

- Ignore $\mu^+ \mu^-$ tracks from J/ψ decay.



Validation studies

Control channel: $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$

Sizeable branching fraction and clean experimental signature.

Use it to validate the modelling of training variables and classification outputs with on-resonance data.

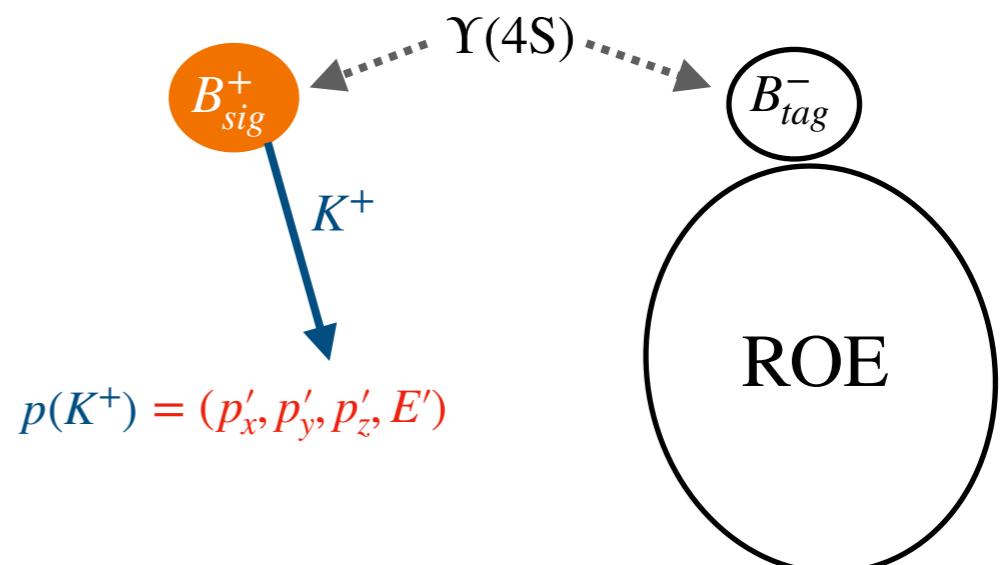
The method

Identification of $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$ events



Modification of the events to reproduce experimental and kinematic features of $B^+ \rightarrow K^+ \nu \bar{\nu}$.

- Ignore $\mu^+ \mu^-$ tracks from J/ψ decay.
- Correct **2-body** \rightarrow **3-body** kinematics.



Validation studies

Control channel: $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$

Sizeable branching fraction and clean experimental signature.

Use it to validate the modelling of training variables and classification outputs with on-resonance data.

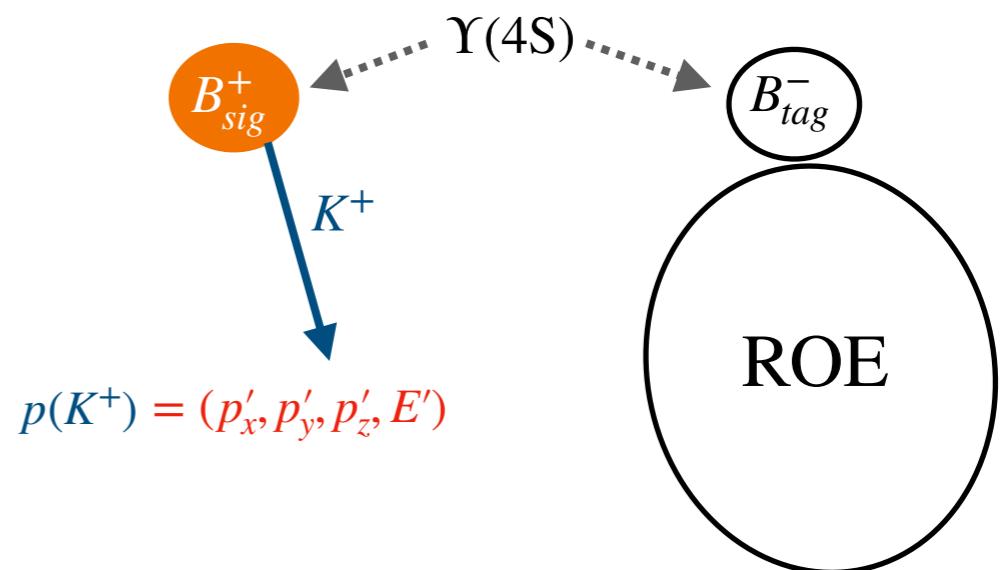
The method

Identification of $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$ events



Modification of the events to reproduce experimental and kinematic features of $B^+ \rightarrow K^+ \nu \bar{\nu}$.

- Ignore $\mu^+ \mu^-$ tracks from J/ψ decay.
- Correct **2-body** \rightarrow **3-body** kinematics:
replace $p(K^+)$ with $p'(K^+)$ from generated $B^+ \rightarrow K^+ \nu \bar{\nu}$.





Validation studies

Control channel: $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$

Sizeable branching fraction and clean experimental signature.

Use it to validate the modelling of training variables and classification outputs with on-resonance data.

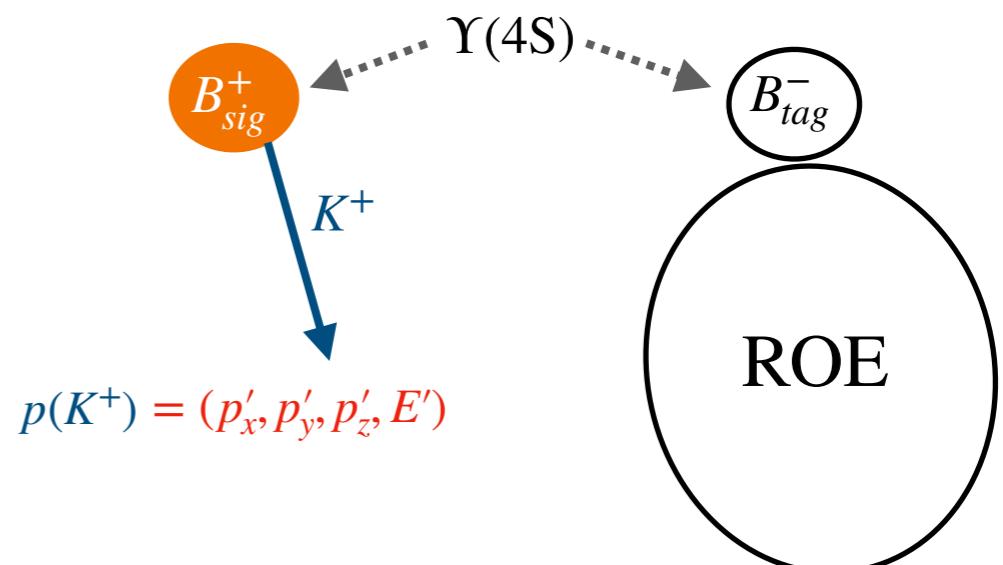
The method

Identification of $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$ events



Modification of the events to reproduce experimental and kinematic features of $B^+ \rightarrow K^+ \nu \bar{\nu}$.

- Ignore $\mu^+ \mu^-$ tracks from J/ψ decay.
- Correct **2-body** \rightarrow **3-body** kinematics:
replace $p(K^+)$ with $p'(K^+)$ from generated $B^+ \rightarrow K^+ \nu \bar{\nu}$.
- Reconstruct and **validate**.



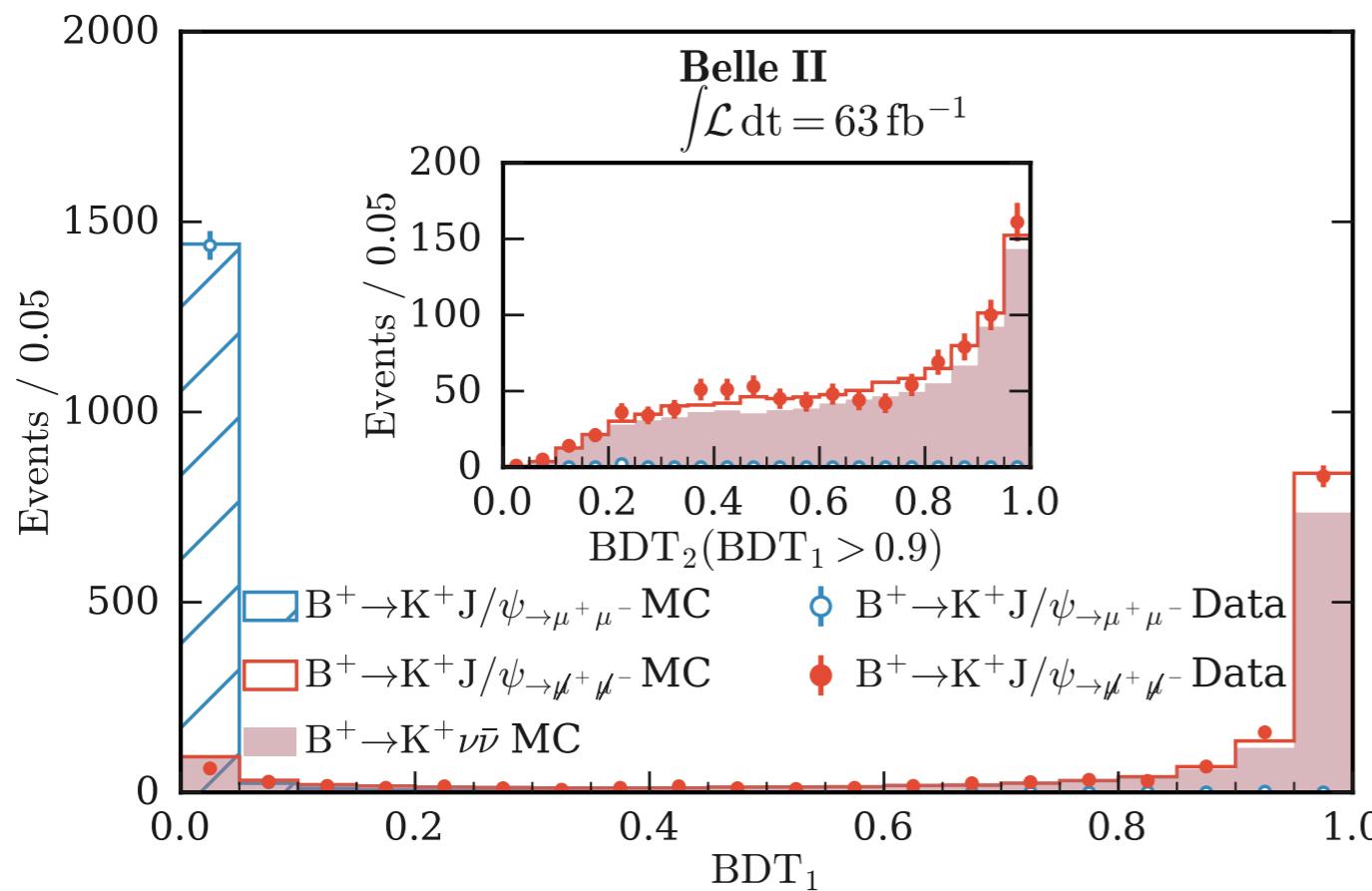
Validation studies

Control channel: $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$

Sizeable branching fraction and clean experimental signature.

Use it to validate the modelling of training variables
and classification outputs with on-resonance data.

The results



**Very good data-MC agreement for
BDT₁ and BDT₂ outputs.**



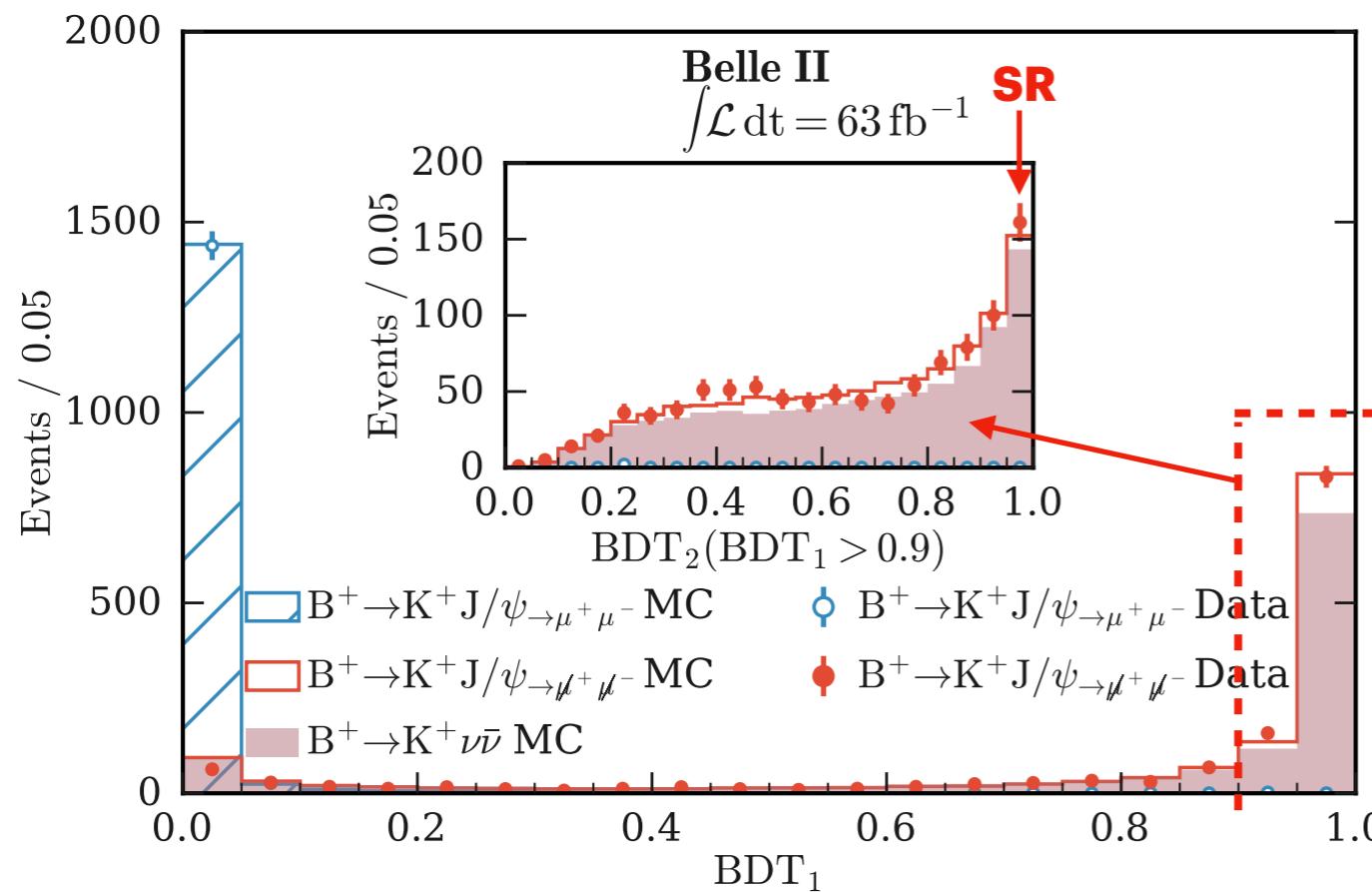
Validation studies

Control channel: $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$

Sizeable branching fraction and clean experimental signature.

Use it to validate the modelling of training variables and classification outputs with on-resonance data.

The results



Very good data-MC agreement for BDT₁ and BDT₂ outputs.

Validation studies

Off-resonance data

Validation studies

Off-resonance data

Check modelling of continuum background simulation.

Validation studies

Off-resonance data

Check modelling of continuum background simulation.

- Discrepancy in yields: $\text{data/MC} = 1.40 \pm 0.12$

Validation studies

Off-resonance data

Check modelling of continuum background simulation.

- Discrepancy in yields: $\text{data/MC} = 1.40 \pm 0.12 \rightarrow$ conservative **50% normalisation uncertainty** in the fit model **for each bkg yield.**

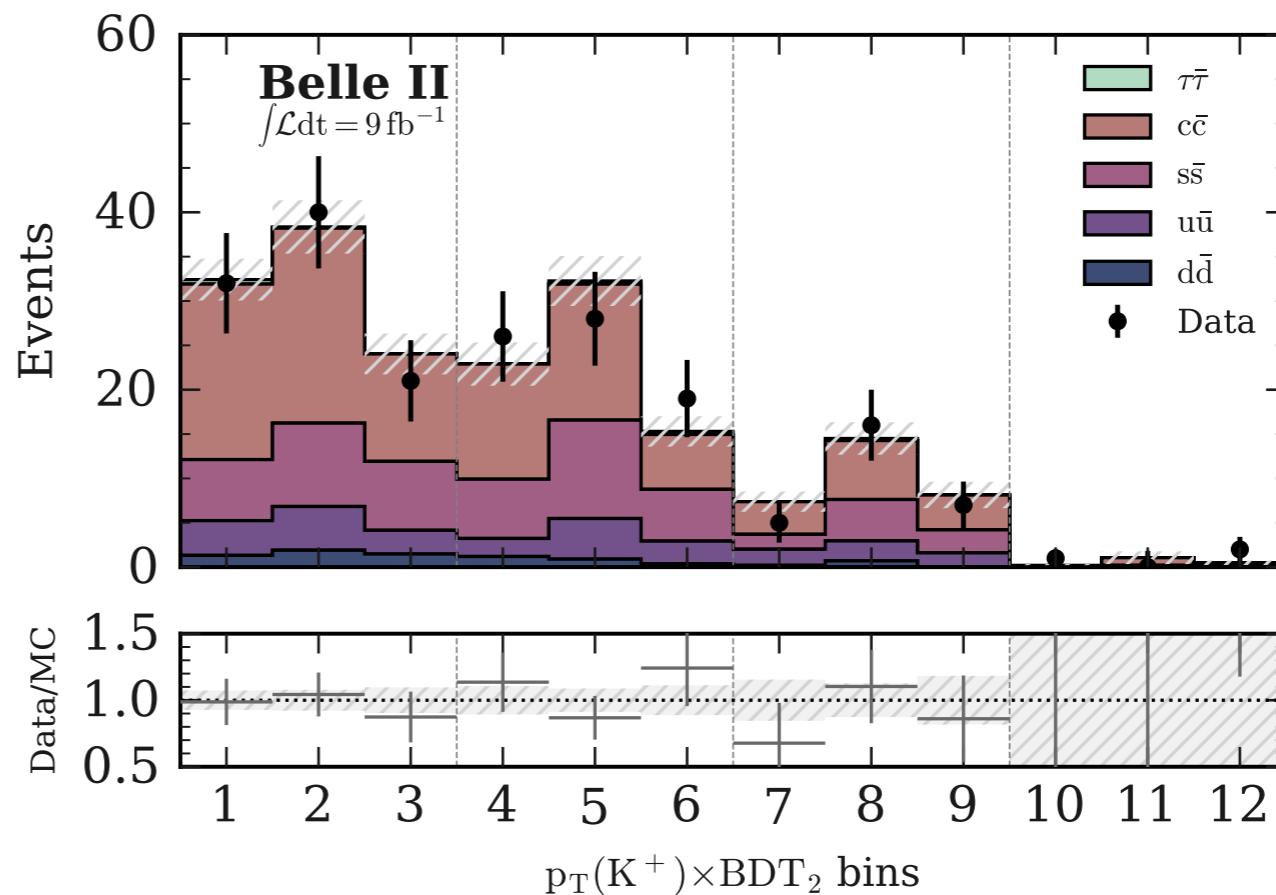
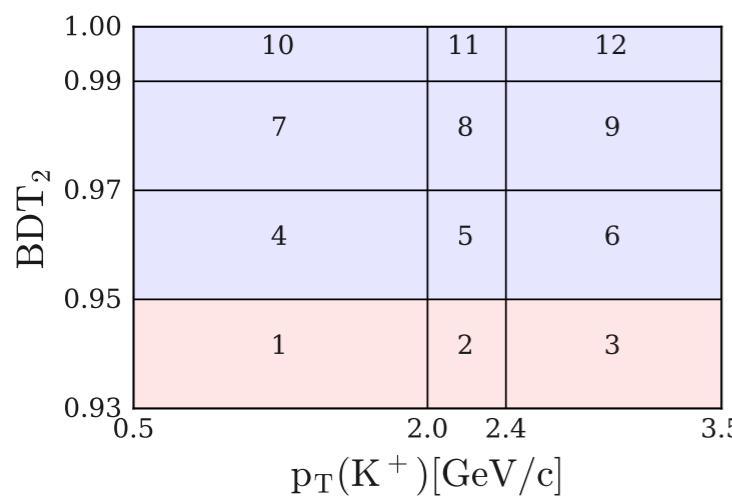


Validation studies

Off-resonance data

Check modelling of continuum background simulation.

- Discrepancy in yields: $\text{data/MC} = 1.40 \pm 0.12 \rightarrow \text{conservative } \mathbf{50\% \, normalisation \, uncertainty}$ in the fit model **for each bkg yield.**
- Good data-MC **shape agreement ✓**



Fitting procedure

Statistical model

Fitting procedure

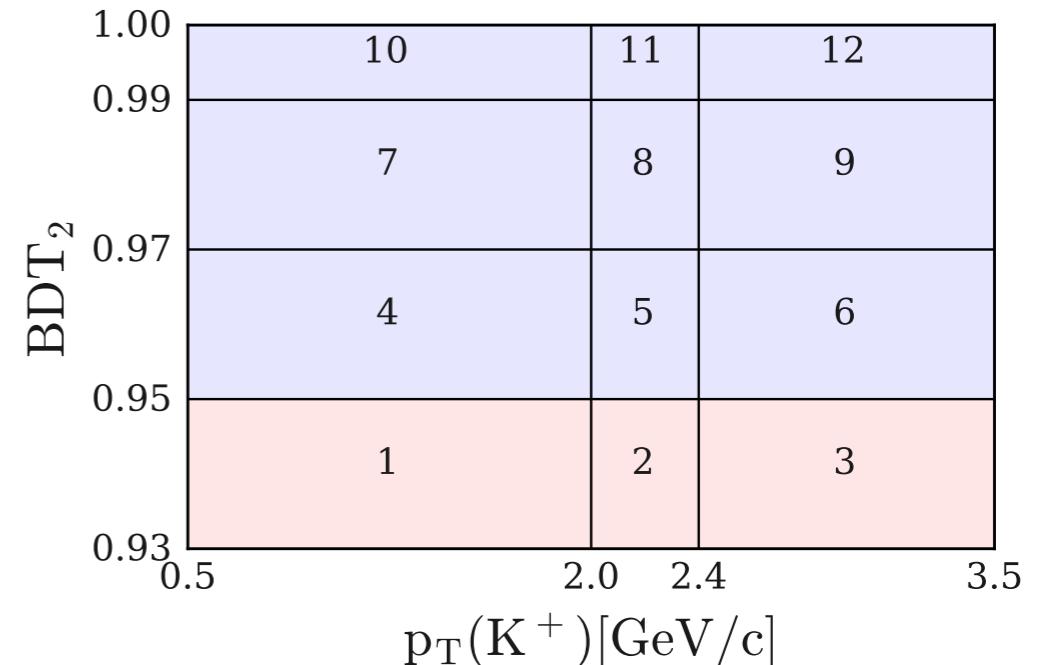
Statistical model

Extended binned maximum-likelihood fit: simultaneous fit of **on-resonance** and **off-resonance** samples.

Fitting procedure

Statistical model

Extended binned maximum-likelihood fit: simultaneous fit of **on-resonance** and **off-resonance** samples.



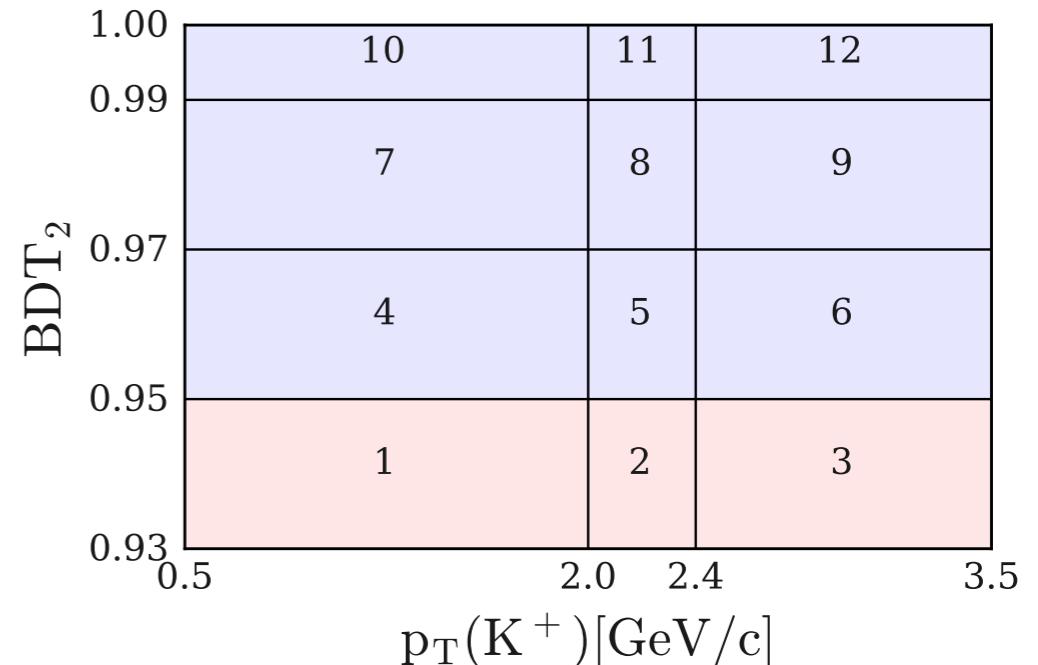
$\text{kaonID} > 0.9$ selection applied
(retains 62% of kaons, removes 97% of pions)

Fitting procedure

Statistical model

Extended binned maximum-likelihood fit: simultaneous fit of **on-resonance** and **off-resonance** samples.

Parametrised PDF based on template histograms.



kaonID > 0.9 selection applied
(retains 62% of kaons, removes 97% of pions)

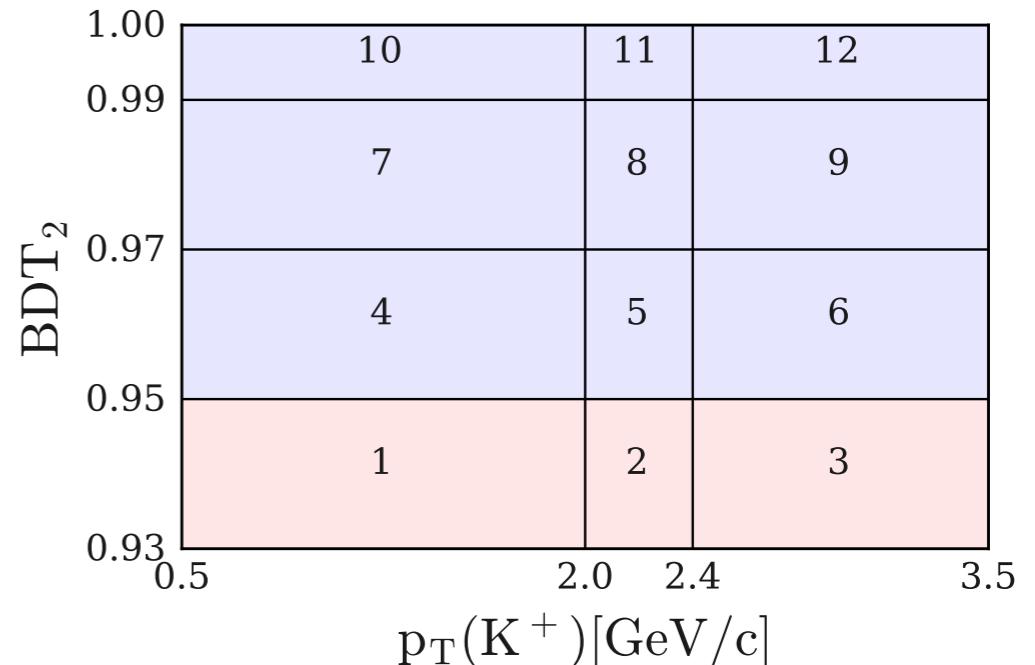
Fitting procedure

Statistical model

Extended binned maximum-likelihood fit: simultaneous fit of **on-resonance** and **off-resonance** samples.

Parametrised PDF based on template histograms.

- **1 parameter of interest** μ : $\mu = 1 \rightarrow \text{SM BR} = 4.6 \times 10^{-6}$



kaonID > 0.9 selection applied
(retains 62% of kaons, removes 97% of pions)

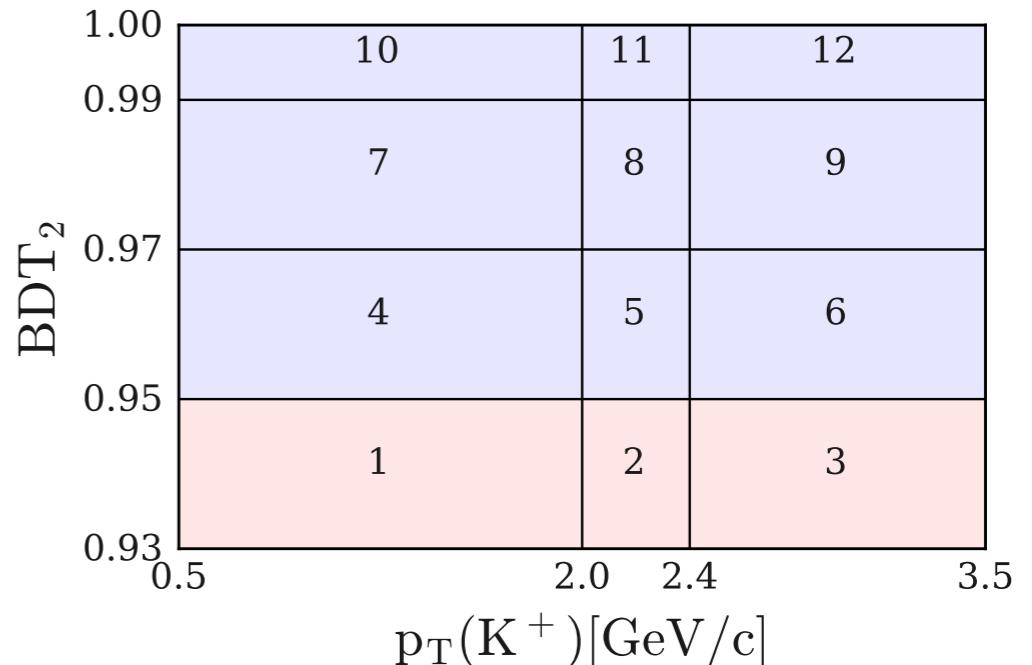
Fitting procedure

Statistical model

Extended binned maximum-likelihood fit: simultaneous fit of **on-resonance** and **off-resonance** samples.

Parametrised PDF based on template histograms.

- **1 parameter of interest** μ : $\mu = 1 \rightarrow \text{SM BR} = 4.6 \times 10^{-6}$
- **175 nuisance parameters** θ
accounting for MC stat. uncertainties
and **systematic uncertainties (19)** on



kaonID > 0.9 selection applied
(retains 62% of kaons, removes 97% of pions)

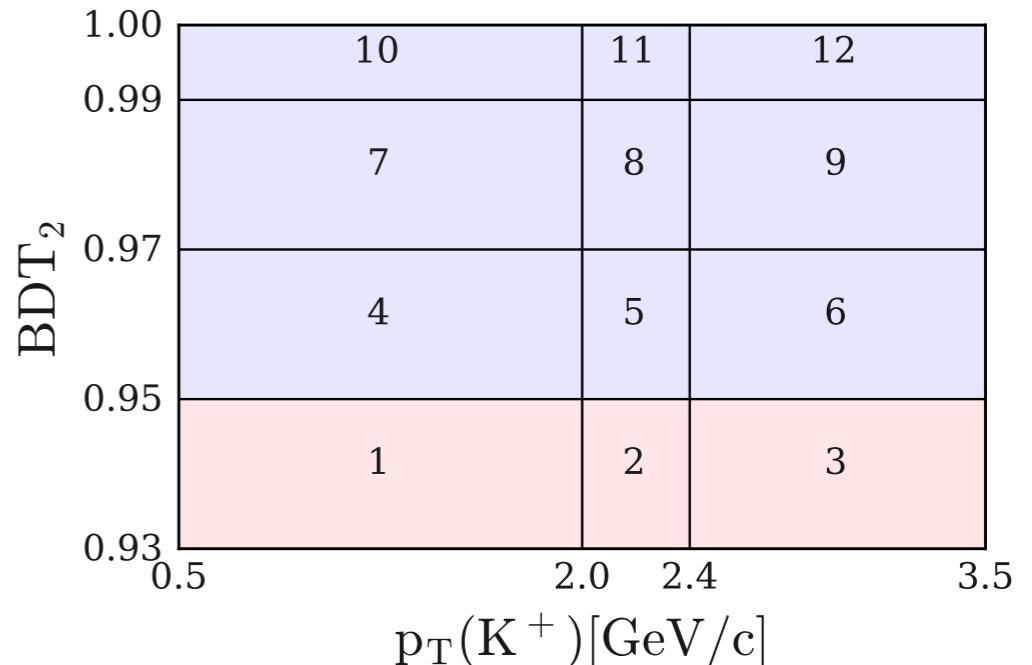
Fitting procedure

Statistical model

Extended binned maximum-likelihood fit: simultaneous fit of **on-resonance** and **off-resonance** samples.

Parametrised PDF based on template histograms.

- **1 parameter of interest** μ : $\mu = 1 \rightarrow \text{SM BR} = 4.6 \times 10^{-6}$
- **175 nuisance parameters** θ
accounting for MC stat. uncertainties
and **systematic uncertainties (19)** on
 - the normalisation of the 7 bkgs (50% each);
 - the BR of the leading B bkg decays;
 - the SM form factor;
 - PID correction weights;
 - tracking efficiency;
 - energy calibration of photon clusters;
 - energy calibration of ECL clusters not matched to photons.



kaonID > 0.9 selection applied
(retains 62% of kaons, removes 97% of pions)

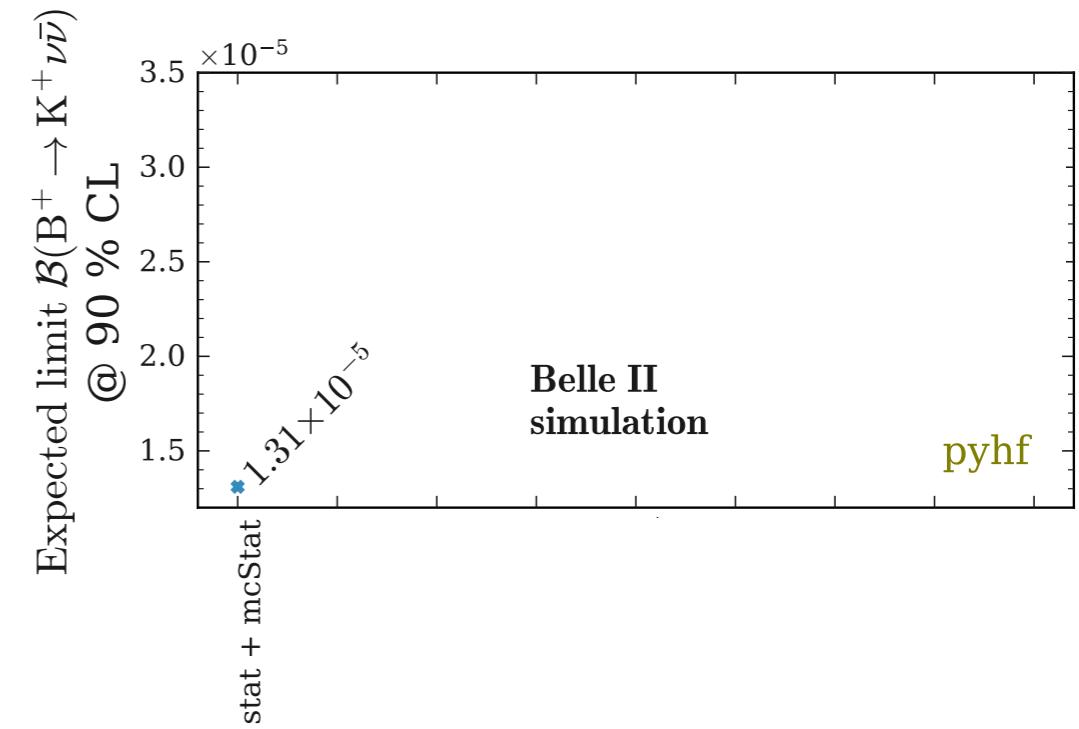
Fitting procedure

Statistical model

Extended binned maximum-likelihood fit: simultaneous fit of **on-resonance** and **off-resonance** samples.

Parametrised PDF based on template histograms.

- **1 parameter of interest** μ : $\mu = 1 \rightarrow \text{SM BR} = 4.6 \times 10^{-6}$
- **175 nuisance parameters** θ
accounting for MC stat. uncertainties
and **systematic uncertainties (19)** on
 - the normalisation of the 7 bkgs (50% each);
 - the BR of the leading B bkg decays;
 - the SM form factor;
 - PID correction weights;
 - tracking efficiency;
 - energy calibration of photon clusters;
 - energy calibration of ECL clusters not matched to photons.



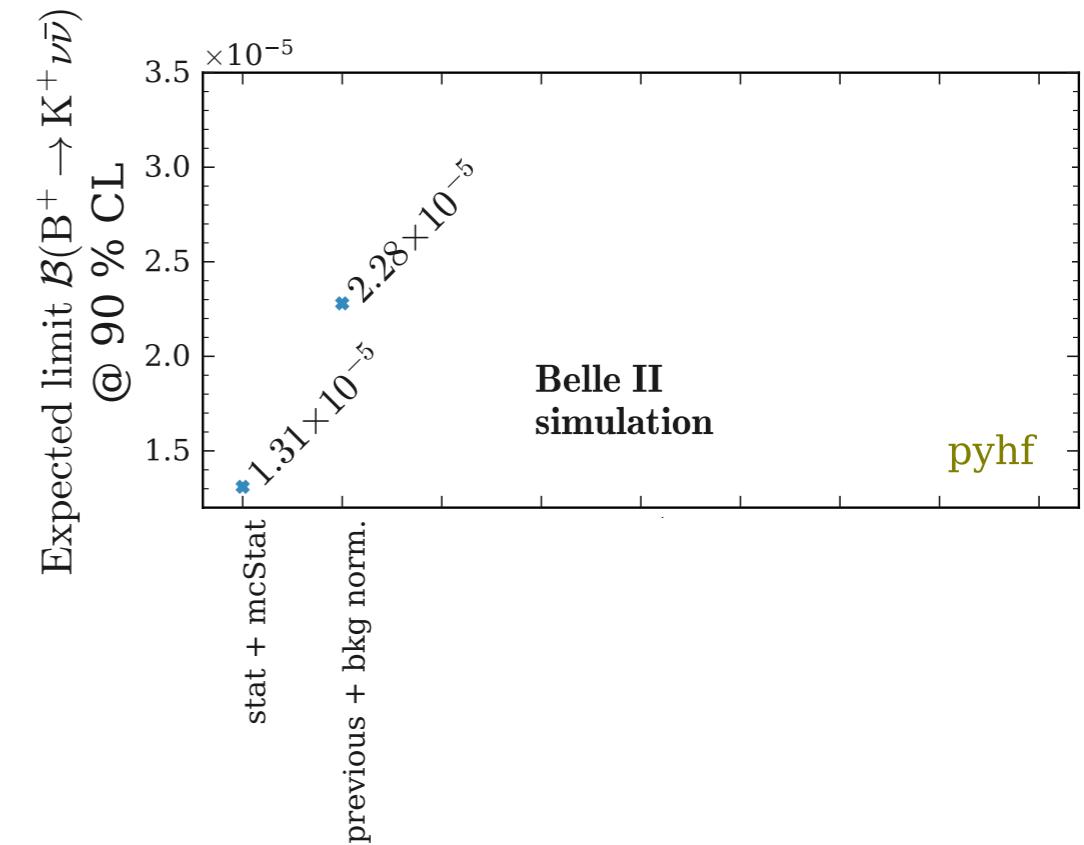
Fitting procedure

Statistical model

Extended binned maximum-likelihood fit: simultaneous fit of **on-resonance** and **off-resonance** samples.

Parametrised PDF based on template histograms.

- **1 parameter of interest** μ : $\mu = 1 \rightarrow \text{SM BR} = 4.6 \times 10^{-6}$
- **175 nuisance parameters** θ
accounting for MC stat. uncertainties
and **systematic uncertainties (19)** on
 - the normalisation of the 7 bkgs (50% each);
 - the BR of the leading B bkg decays;
 - the SM form factor;
 - PID correction weights;
 - tracking efficiency;
 - energy calibration of photon clusters;
 - energy calibration of ECL clusters not matched to photons.





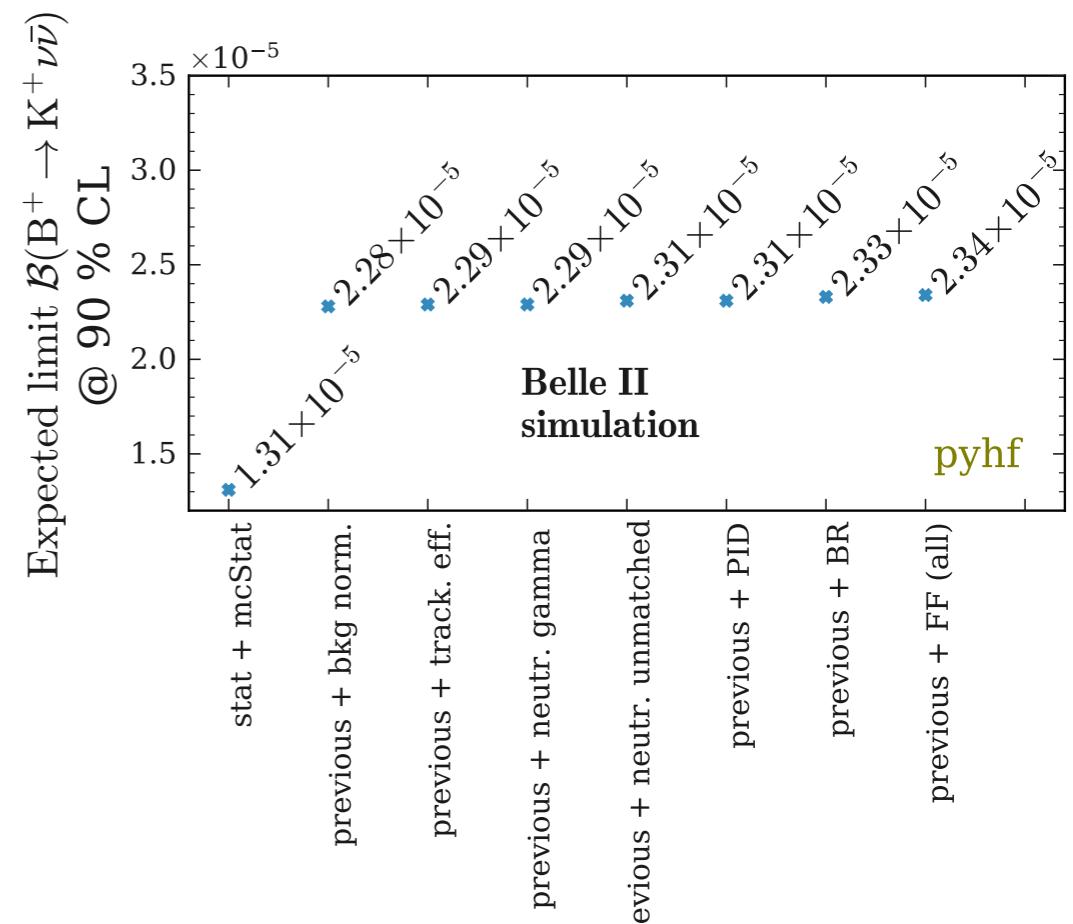
Fitting procedure

Statistical model

Extended binned maximum-likelihood fit: simultaneous fit of **on-resonance** and **off-resonance** samples.

Parametrised PDF based on template histograms.

- **1 parameter of interest** μ : $\mu = 1 \rightarrow \text{SM BR} = 4.6 \times 10^{-6}$
- **175 nuisance parameters** θ
accounting for MC stat. uncertainties
and **systematic uncertainties (19)** on
 - the **normalisation of the 7 bkg**s (50% each): **dominant**;
 - the BR of the leading B bkg decays;
 - the SM form factor;
 - PID correction weights;
 - tracking efficiency;
 - energy calibration of photon clusters;
 - energy calibration of ECL clusters not matched to photons.

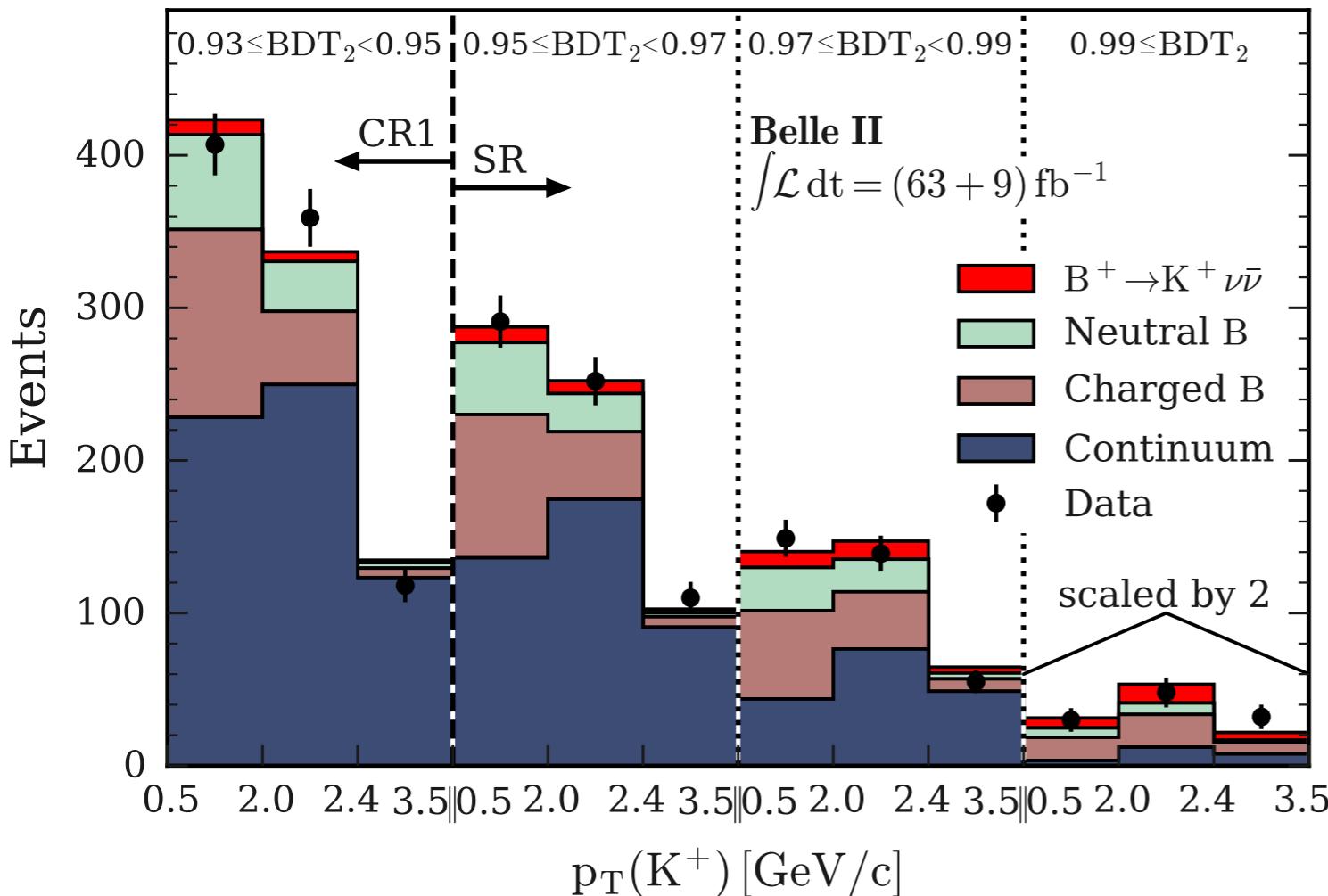


Fit results

Signal region unblinding

Fit results

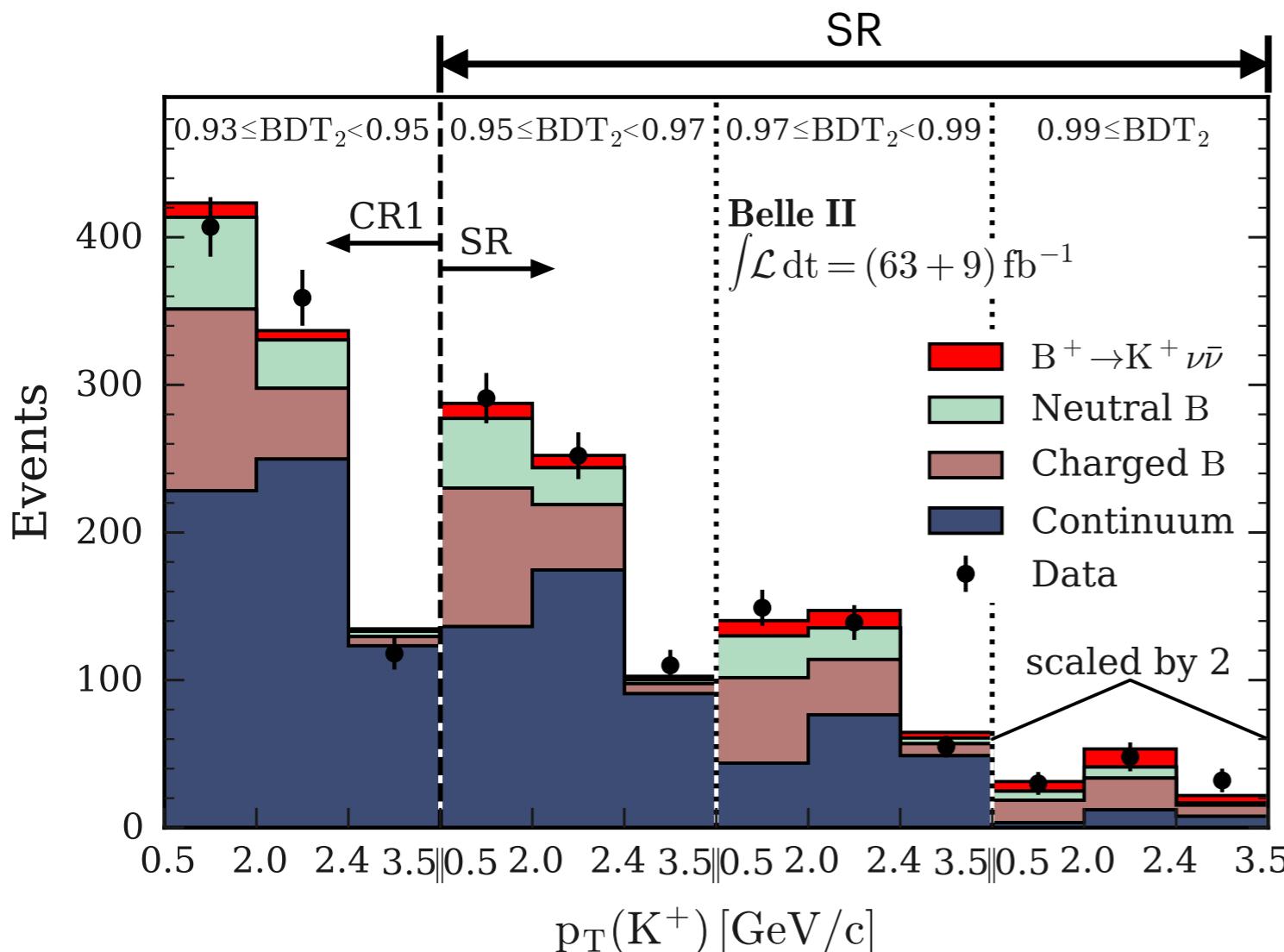
Signal region unblinding



- **Good agreement** between combined **post-fit yields and data**.

Fit results

Signal region unblinding

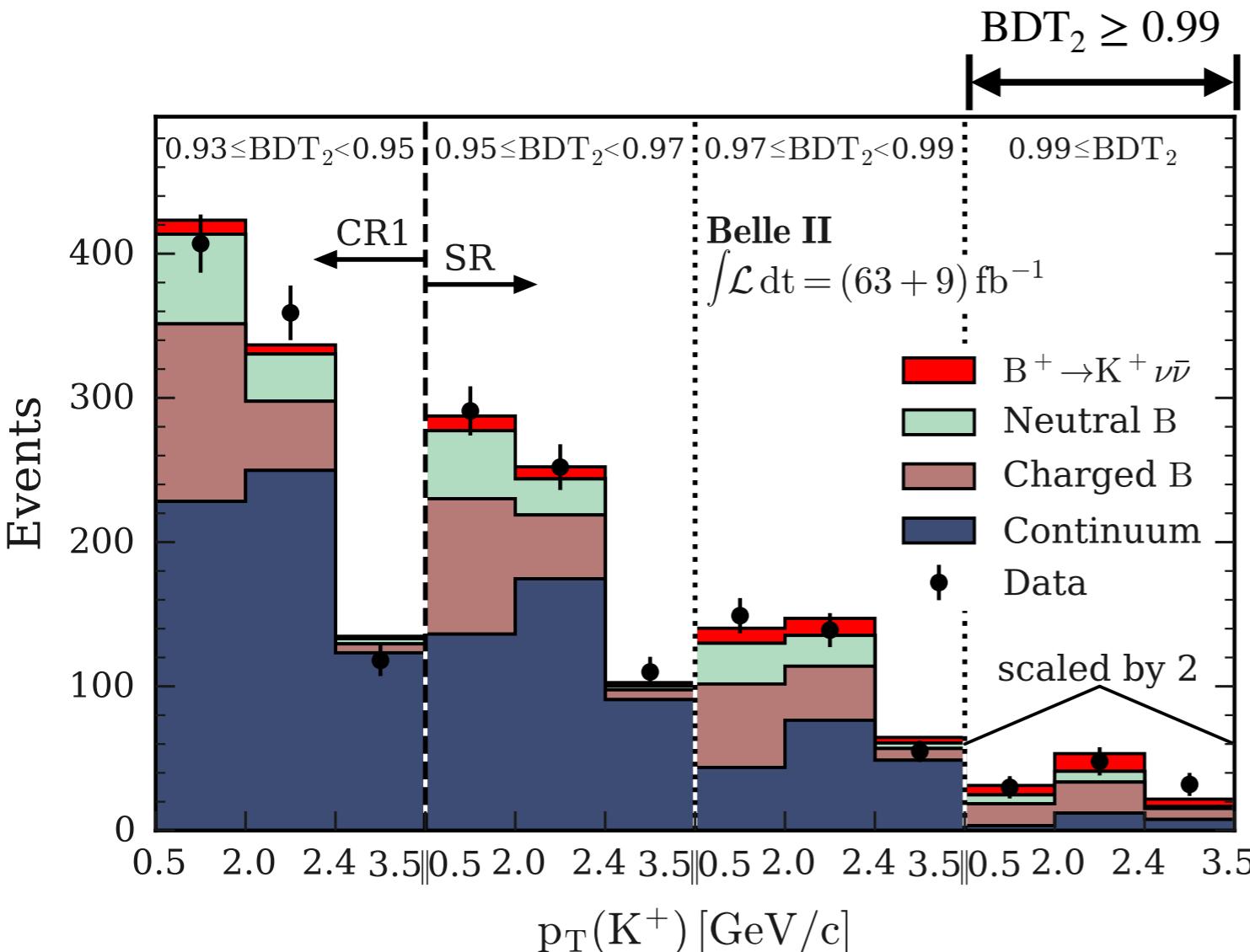


- **Good agreement** between combined **post-fit yields and data**.
- SR: **purity = 6%**
continuum = 59% of bkg



Fit results

Signal region unblinding



- **Good agreement** between combined **post-fit yields and data**.
- SR: **purity = 6%**
continuum = 59% of bkg
- BDT₂ ≥ 0.99 : **purity = 22%**
continuum = 28% of bkg

Fit results

Measured signal

Fit results

Measured signal

$$\rightarrow \mu = 4.2^{+3.4}_{-3.2} = 4.2^{+2.9}_{-2.8}(\text{stat})^{+1.8}_{-1.6}(\text{syst})$$

$$\rightarrow \text{BR}(B^+ \rightarrow K^+ \nu \bar{\nu}) = [1.9^{+1.6}_{-1.5}] \times 10^{-5} = [1.9^{+1.3}_{-1.3}(\text{stat})^{+0.8}_{-0.7}(\text{syst})] \times 10^{-5}$$

Fit results

Measured signal

$$\rightarrow \mu = 4.2^{+3.4}_{-3.2} = 4.2^{+2.9}_{-2.8}(\text{stat})^{+1.8}_{-1.6}(\text{syst})$$

$$\rightarrow \text{BR}(B^+ \rightarrow K^+ \nu \bar{\nu}) = [1.9^{+1.6}_{-1.5}] \times 10^{-5} = [1.9^{+1.3}_{-1.3}(\text{stat})^{+0.8}_{-0.7}(\text{syst})] \times 10^{-5}$$

- Compatible with:
 - **SM** expectation $\mu = 1$ at $\text{CL} = 1\sigma$
 - **bkg-only** hypothesis $\mu = 0$ at $\text{CL} = 1.3\sigma$



Fit results

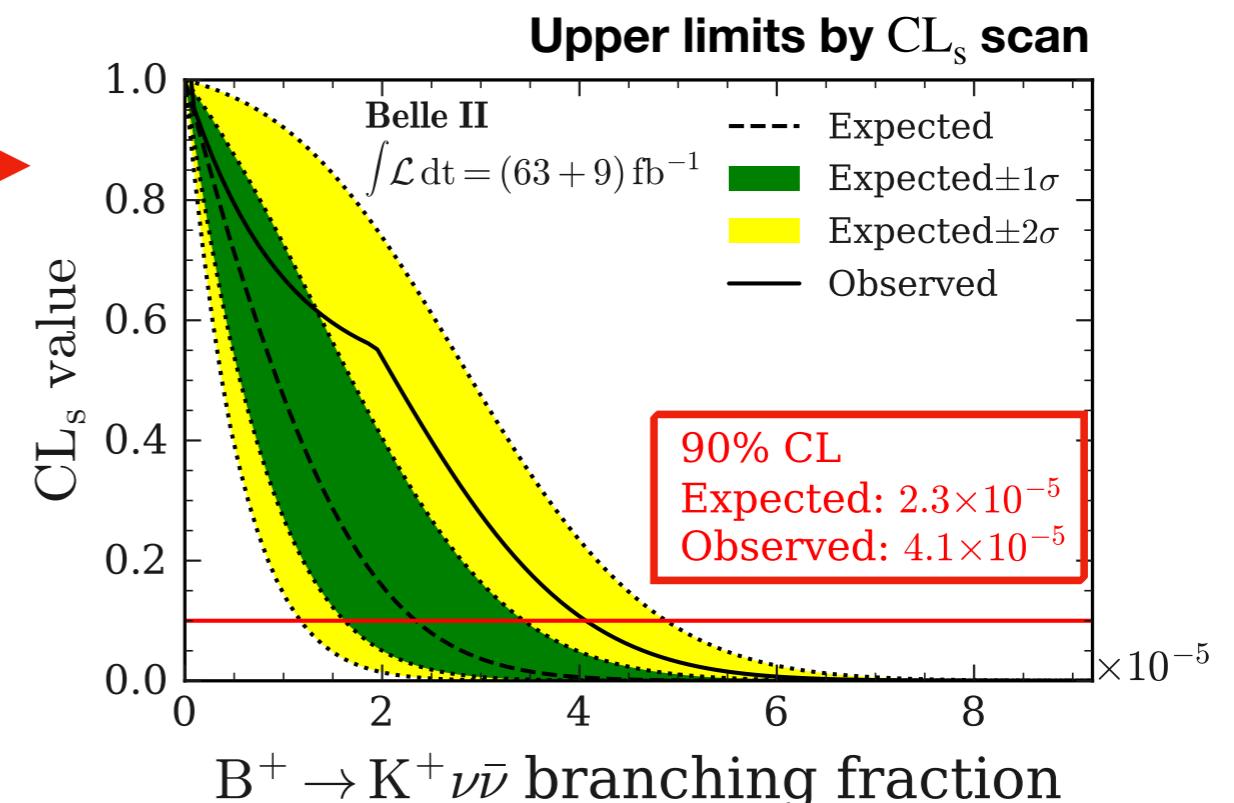
Measured signal

→ $\mu = 4.2^{+3.4}_{-3.2} = 4.2^{+2.9}_{-2.8}(\text{stat})^{+1.8}_{-1.6}(\text{syst})$

→ $\text{BR}(B^+ \rightarrow K^+ \nu \bar{\nu}) = [1.9^{+1.6}_{-1.5}] \times 10^{-5} = [1.9^{+1.3}_{-1.3}(\text{stat})^{+0.8}_{-0.7}(\text{syst})] \times 10^{-5}$

- Compatible with:

- **SM** expectation $\mu = 1$ at $\text{CL} = 1\sigma$
- **bkg-only** hypothesis $\mu = 0$ at $\text{CL} = 1.3\sigma$



Discussion of the results

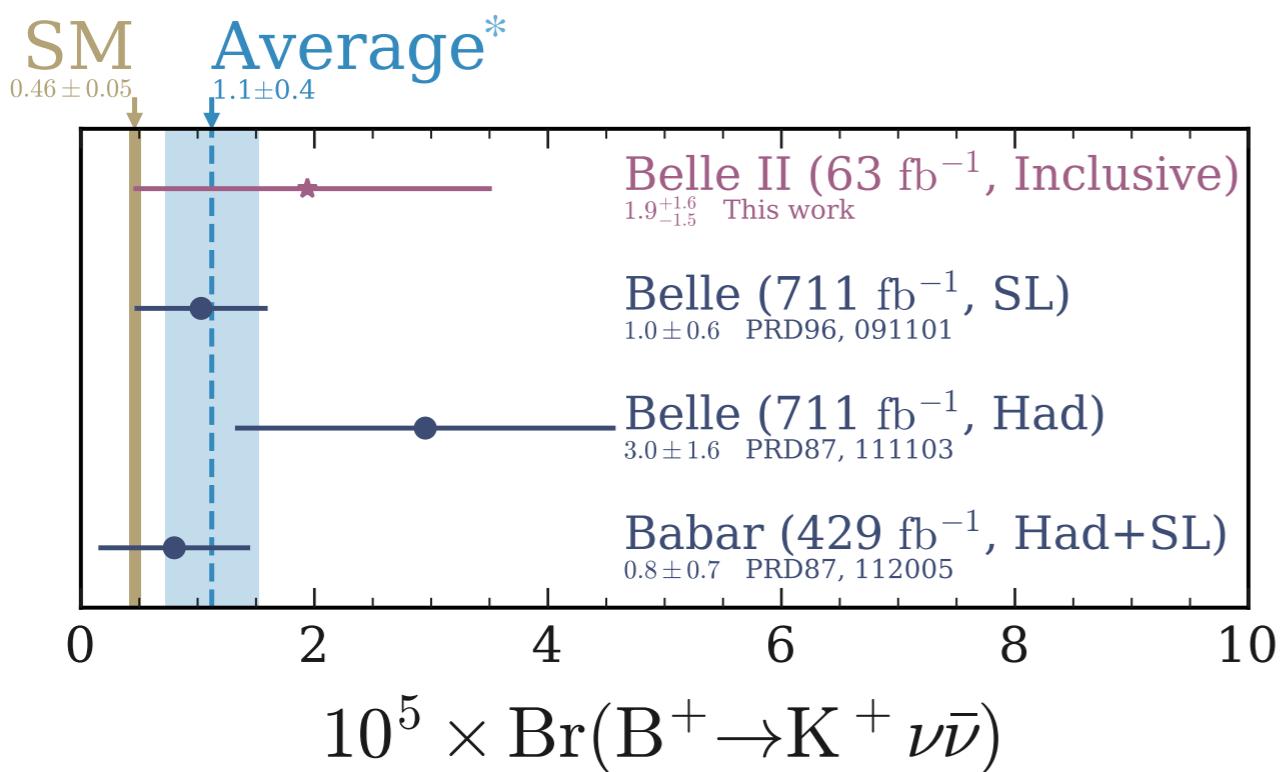
$$^* \frac{\sum_i w_i x_i}{\sum_i w_i} \pm \frac{1}{\left(\sum_i w_i\right)^{1/2}}, \quad w_i = 1/(\sigma_{x_i})^2$$

Comparison with previous measurements

$$^* \frac{\sum_i w_i x_i}{\sum_i w_i} \pm \frac{1}{\left(\sum_i w_i\right)^{1/2}}, \quad w_i = 1/(\sigma_{x_i})^2$$

Discussion of the results

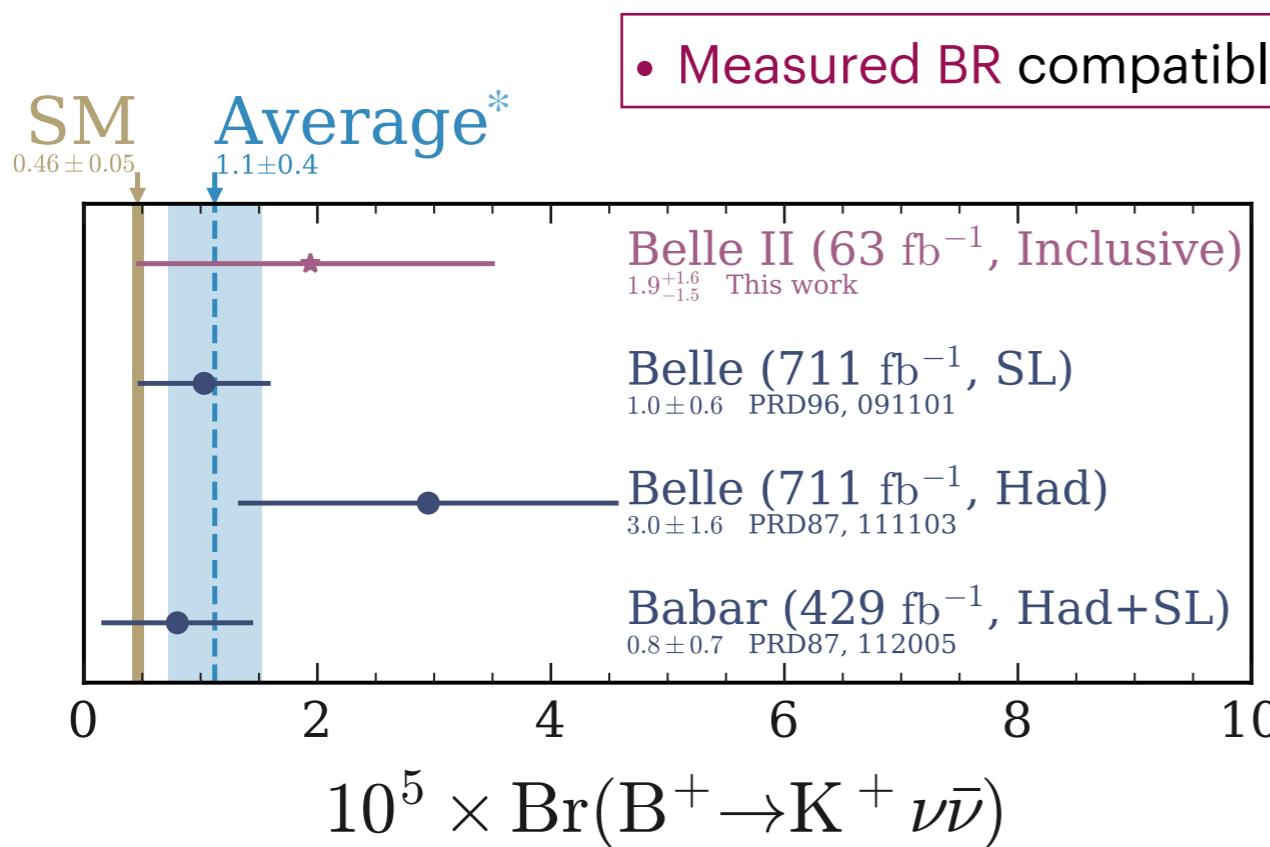
Comparison with previous measurements



$$^* \frac{\sum_i w_i x_i}{\sum_i w_i} \pm \frac{1}{\left(\sum_i w_i\right)^{1/2}}, \quad w_i = 1/(\sigma_{x_i})^2$$

Discussion of the results

Comparison with previous measurements

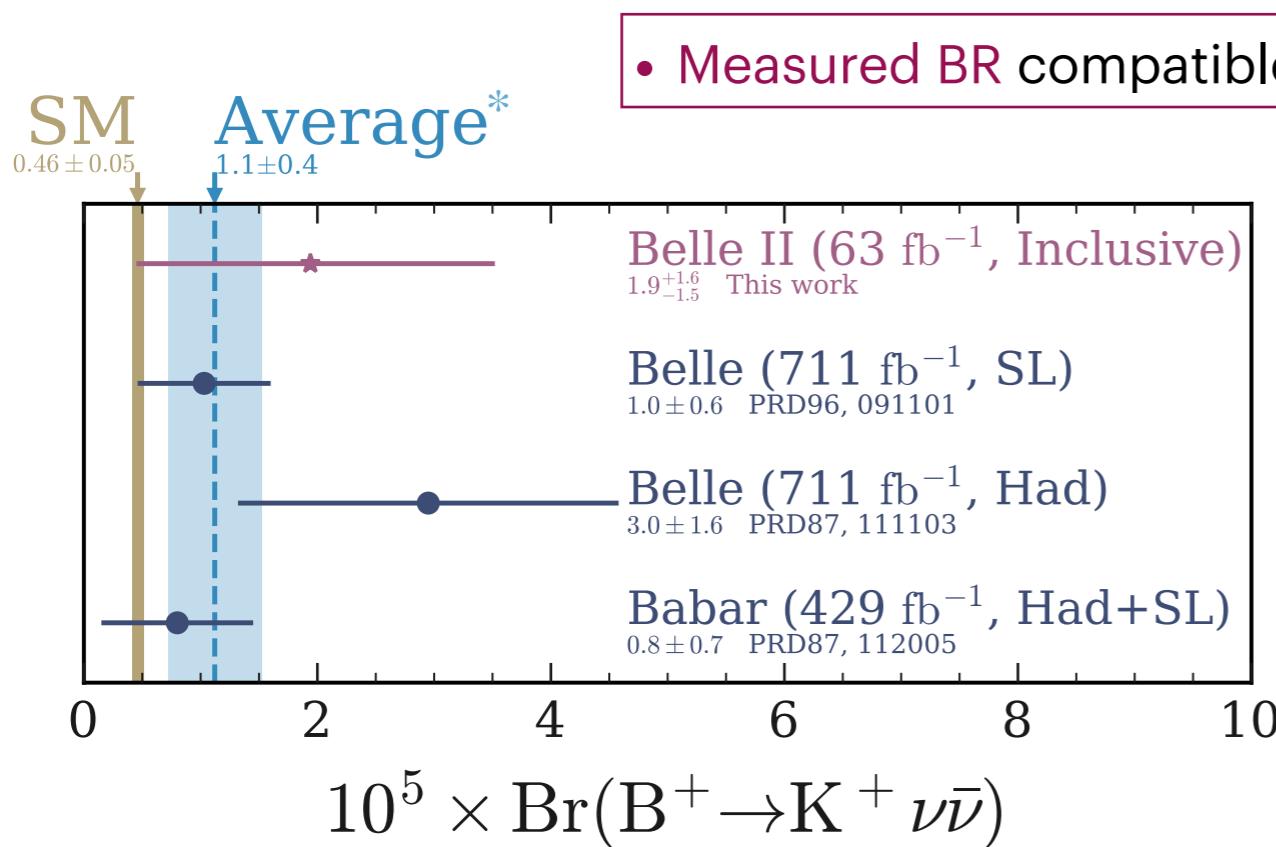


Using only 63 fb^{-1} competitive with previous Belle (BaBar) searches in ~ 11 (7) \times larger samples

$$^* \frac{\sum_i w_i x_i}{\sum_i w_i} \pm \frac{1}{\left(\sum_i w_i\right)^{1/2}}, \quad w_i = 1/(\sigma_{x_i})^2$$

Discussion of the results

Comparison with previous measurements



- Measured BR compatible with previous results.

- Average of the 4 measurements:
 - compatible with SM at CL = 1.6σ
 - central value above SM by factor 1.4

Using only 63 fb^{-1} competitive with previous Belle (BaBar) searches in ~ 11 (7) \times larger samples

Search for $B^0 \rightarrow K^{*0} \nu \bar{\nu}$
with an inclusive tagging
at Belle II

Data samples

Data samples

Collision data

Search optimised for:

- **on-resonance** sample of 189 fb^{-1}
- **off-resonance** sample of 18 fb^{-1}

Data samples

Collision data

Search optimised for:

- **on-resonance** sample of 189 fb^{-1}
- **off-resonance** sample of 18 fb^{-1}

Simulated data



Data samples

Collision data

Search optimised for:

- **on-resonance** sample of 189 fb^{-1}
- **off-resonance** sample of 18 fb^{-1}

Simulated data

- **Signal Monte Carlo (MC):** $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B^0(\rightarrow K^{*0}\nu\bar{\nu})\bar{B}^0$ **following SM expectation.**
 \downarrow
 $K^+\pi^-$
 $[\text{BR}(K^{*0} \rightarrow K^+\pi^-) = 2/3]$
[J. High Energ. Phys. 2015, 184]
[J. High Energ. Phys. 2016, 8]
- **Background MC:** - generic $B^0\bar{B}^0$ and B^+B^- decays from the $\Upsilon(4S)$;
- continuum: $e^+e^- \rightarrow q\bar{q}$ ($q = u, d, s, c$) and $e^+e^- \rightarrow \tau^+\tau^-$ events.

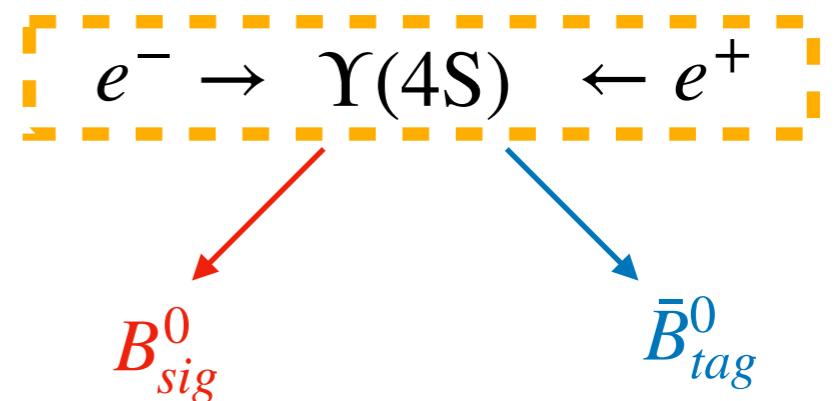
$B^0 \rightarrow K^{*0} \nu \bar{\nu}$ reconstruction

Signal selection

$B^0 \rightarrow K^{*0} \nu \bar{\nu}$ reconstruction

Signal selection

- **Step 1:** signal reconstruction.

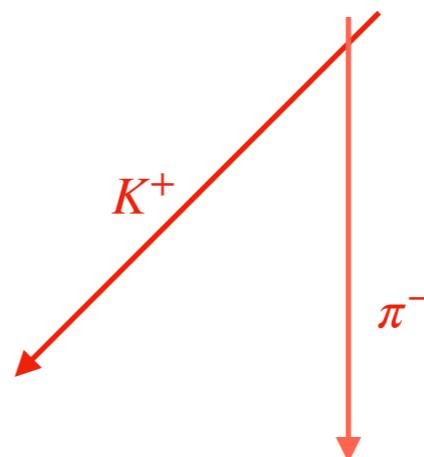
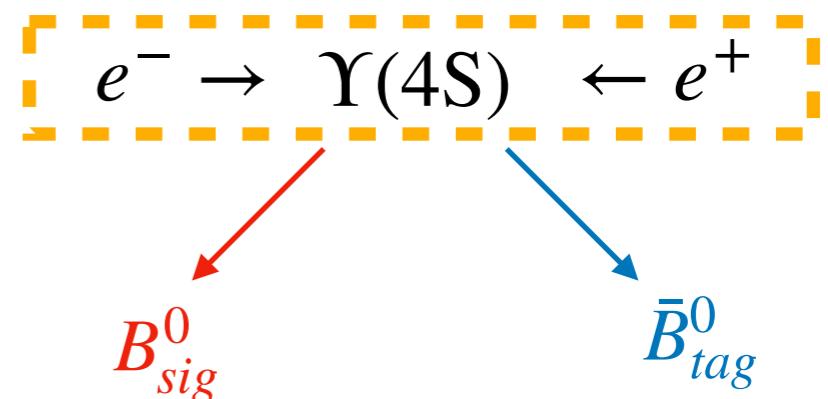


$B^0 \rightarrow K^{*0} \nu \bar{\nu}$ reconstruction

Signal selection

- **Step 1:** signal reconstruction.

Select candidate K^+ (kaonID > 0.9) and π^- tracks.



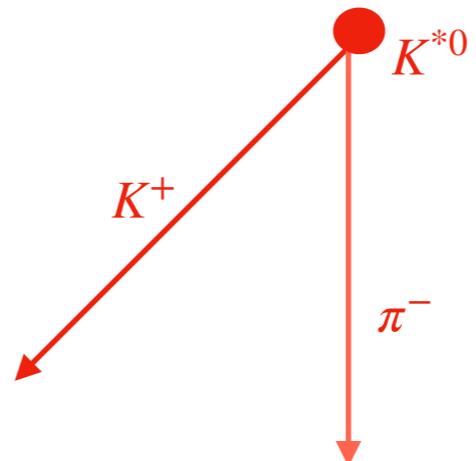
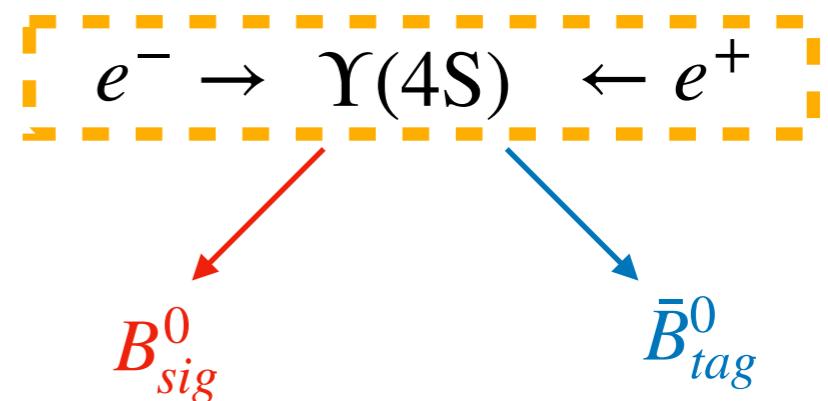
$B^0 \rightarrow K^{*0} \nu \bar{\nu}$ reconstruction

Signal selection

- **Step 1:** signal reconstruction.

Select candidate K^+ (kaonID > 0.9) and π^- tracks.

Fit to common vertex and select K^{*0} candidates



$B^0 \rightarrow K^{*0} \nu \bar{\nu}$ reconstruction

Signal selection

- **Step 1:** signal reconstruction.

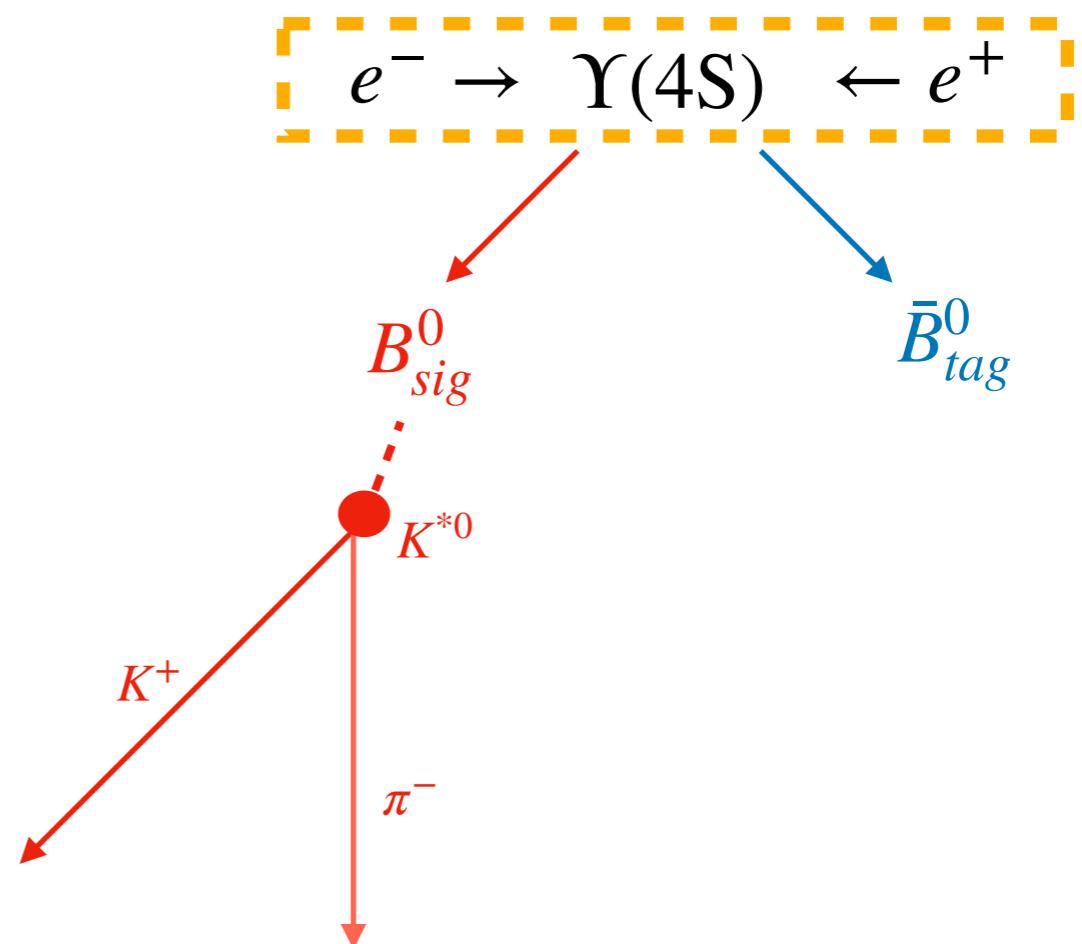
Select candidate K^+ (kaonID > 0.9) and π^- tracks.

Fit to common **vertex** and select K^{*0} candidates

Keep **only** candidate corresponding to **lowest**

$$q_{\text{rec}}^2 = M_{B^0}^2 + M_{K^{*0}}^2 - 2M_{B^0} E_{K^{*0}}$$

→ correct match $\sim 87\%$



$B^0 \rightarrow K^{*0} \nu \bar{\nu}$ reconstruction

Signal selection

- **Step 1:** signal reconstruction.

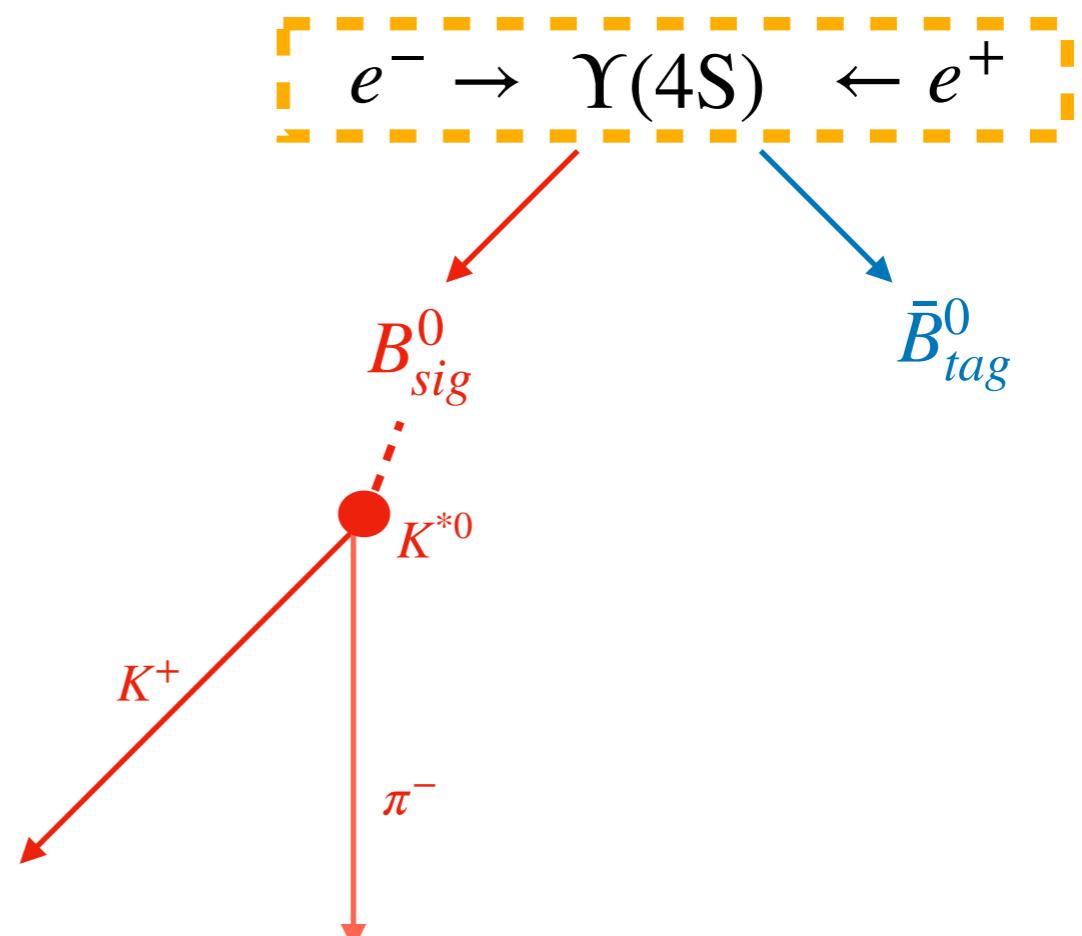
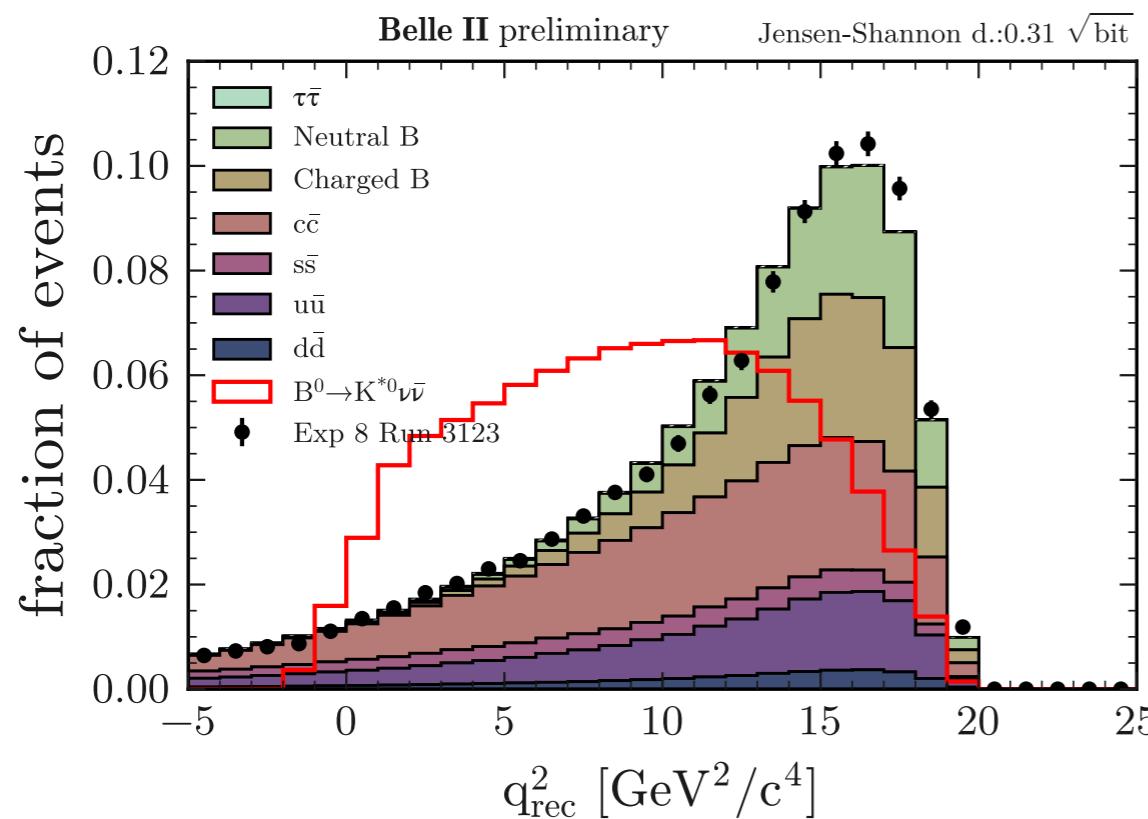
Select candidate K^+ (kaonID > 0.9) and π^- tracks.

Fit to common **vertex** and select K^{*0} candidates

Keep **only** candidate corresponding to **lowest**

$$q_{\text{rec}}^2 = M_{B^0}^2 + M_{K^{*0}}^2 - 2M_{B^0} E_{K^{*0}}$$

→ correct match $\sim 87\%$





$B^0 \rightarrow K^{*0} \nu \bar{\nu}$ reconstruction

Signal selection

- **Step 1:** signal reconstruction.

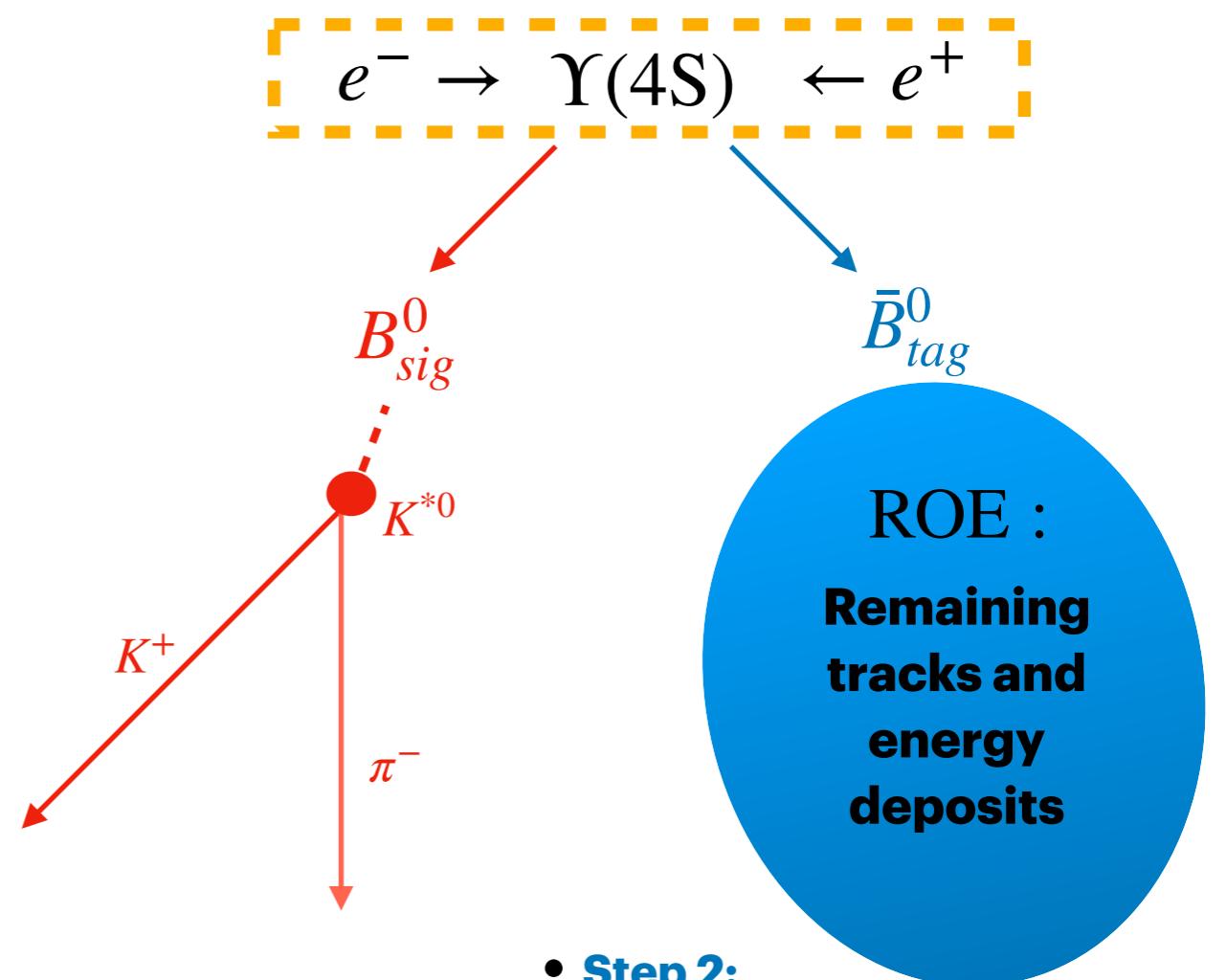
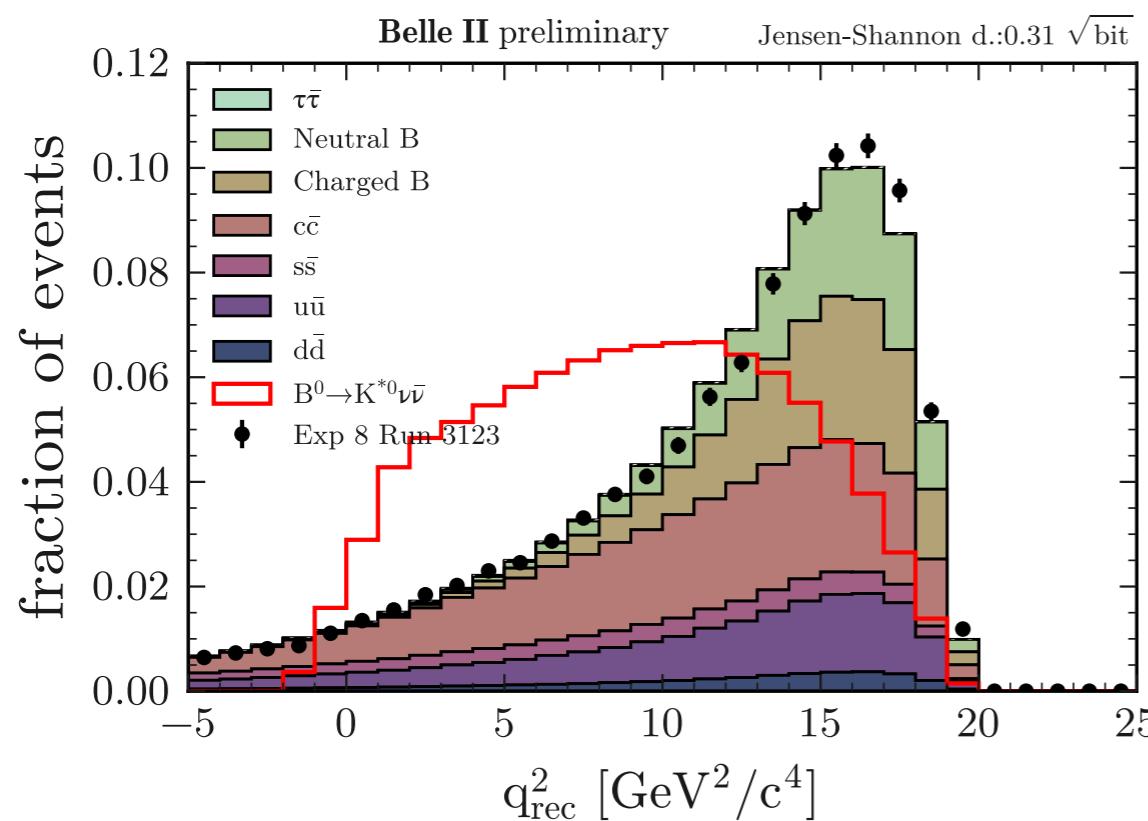
Select candidate K^+ (kaonID > 0.9) and π^- tracks.

Fit to common **vertex** and select K^{*0} candidates

Keep **only** candidate corresponding to **lowest**

$$q_{\text{rec}}^2 = M_{B^0}^2 + M_{K^{*0}}^2 - 2M_{B^0} E_{K^{*0}}$$

→ correct match $\sim 87\%$



- **Step 2:** inclusive reconstruction of the *rest of the event* (ROE).

Background suppression

Multivariate classification

Background suppression

Multivariate classification

Keep strategy of **2 consecutive binary classifiers.**

Background suppression

Multivariate classification

Keep strategy of **2 consecutive binary classifiers**.

- BDT_1 for **initial background suppression**

Background suppression

Multivariate classification

Keep strategy of **2 consecutive binary classifiers**.

- **BDT₁** for **initial background suppression**
 - Training variables: **event-shape** and **ROE kinematics**.
 - Trained on simulated signal and bkg samples.

Background suppression

Multivariate classification

Keep strategy of **2 consecutive binary classifiers**.

- **BDT₁** for **initial background suppression**
 - Training variables: **event-shape** and **ROE kinematics**.
 - Trained on simulated signal and bkg samples.
- **BDT₂** to **improve** background **suppression**

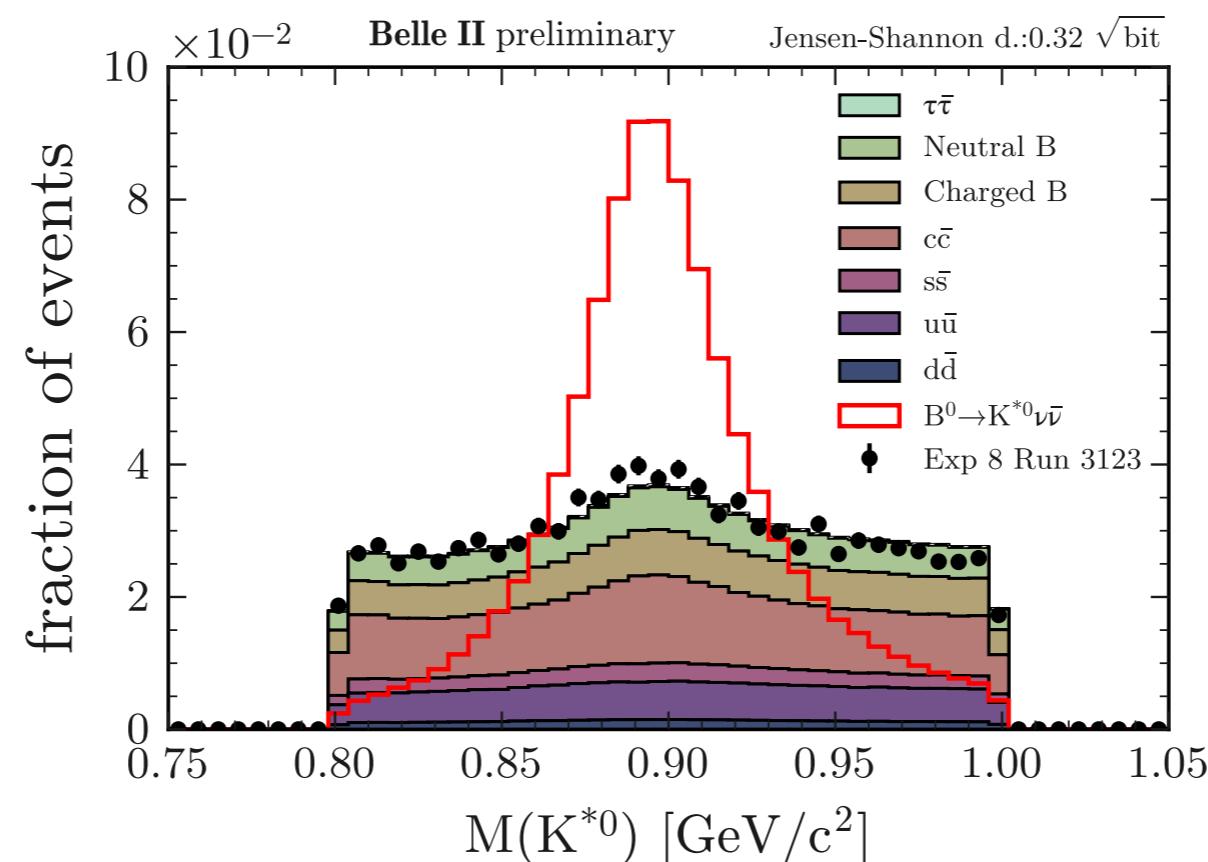


Background suppression

Multivariate classification

Keep strategy of **2 consecutive binary classifiers**.

- **BDT₁** for **initial background suppression**
 - Training variables: **event-shape** and **ROE kinematics**.
 - Trained on simulated signal and bkg samples.
- **BDT₂** to **improve background suppression**
 - Training variables: K^{*0} **specific**, event-based, ROE features, combining signal candidate and ROE.
 - Trained on [reconstructed events with BDT₁ > 0.9](#) in simulated signal and bkg samples.

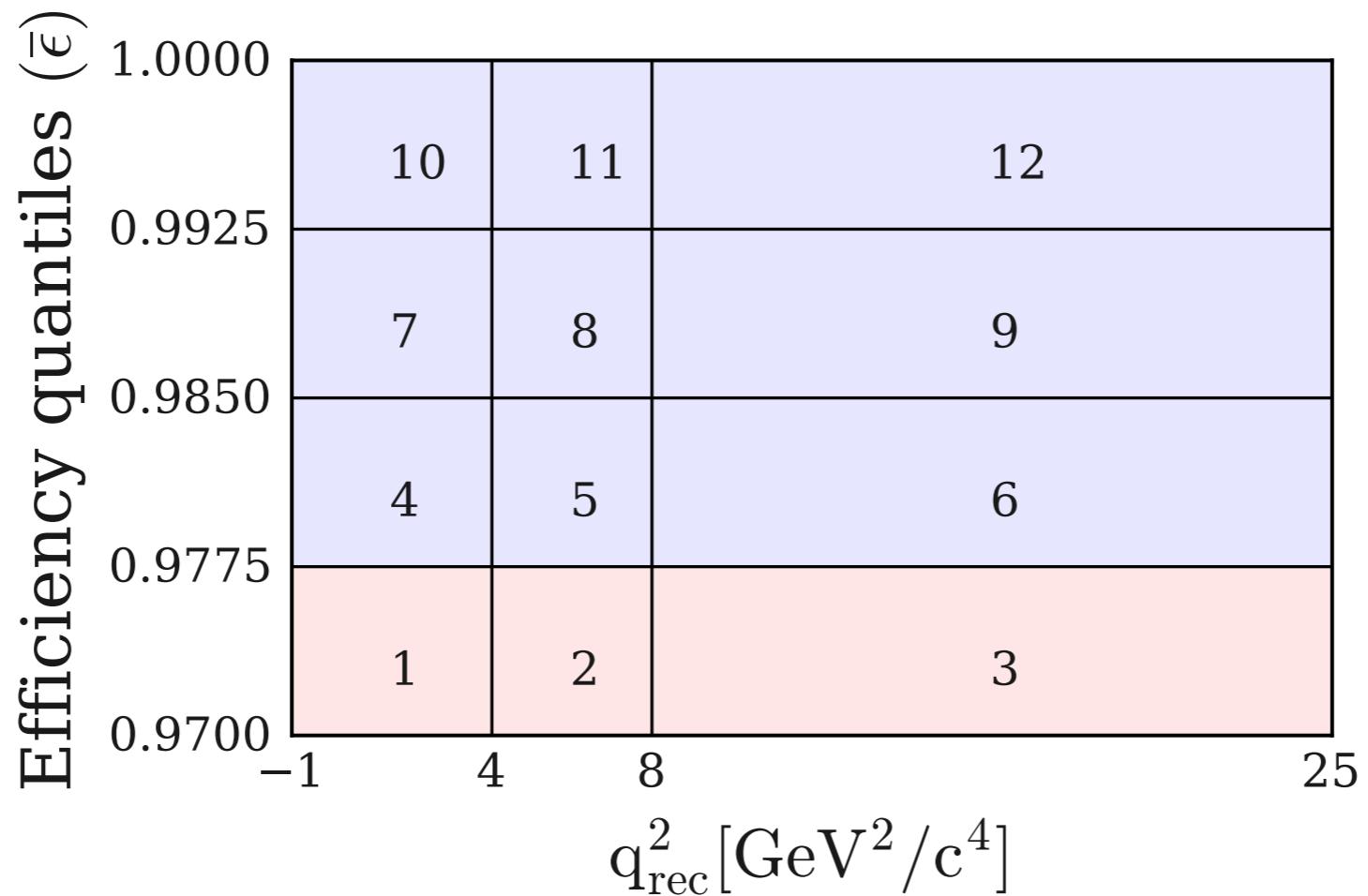


Signal and control regions

Definitions and bin boundaries

Signal and control regions

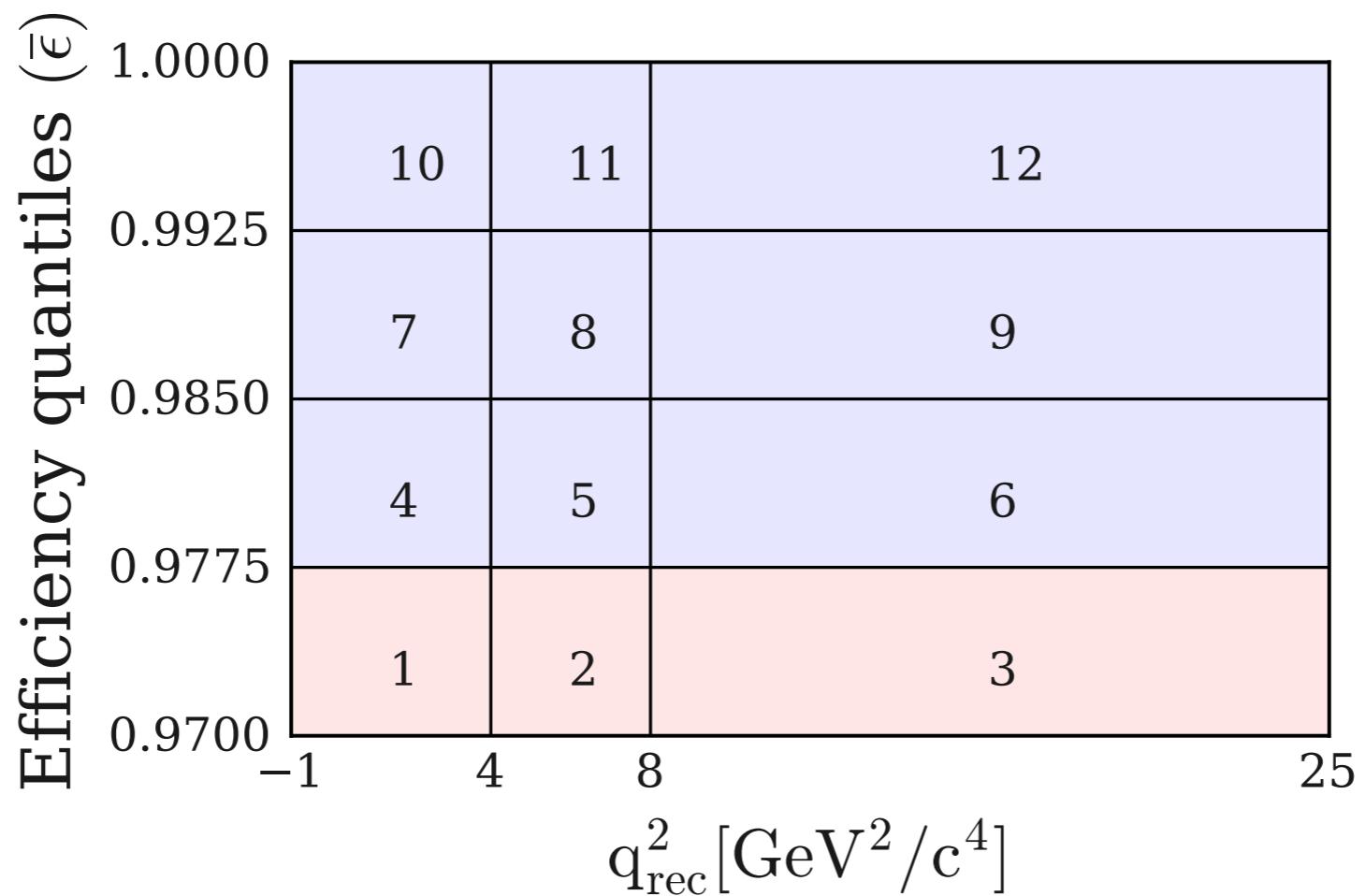
Definitions and bin boundaries



Signal and control regions

Definitions and bin boundaries

- q_{rec}^2 : 3 bins providing good separation, inspired by SM analysis [J. High Energ. Phys. 2015, 184]

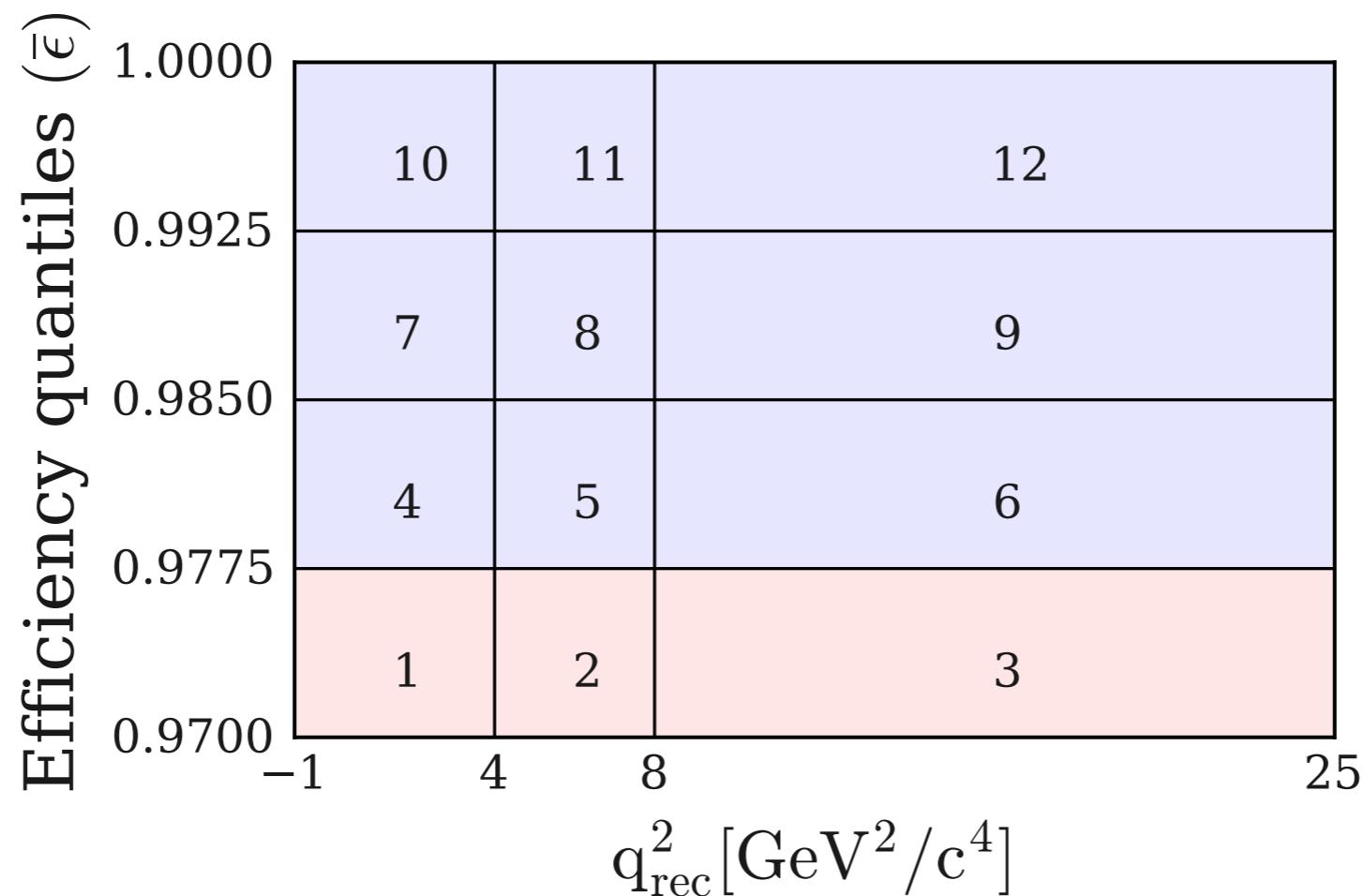




Signal and control regions

Definitions and bin boundaries

- q_{rec}^2 : 3 bins providing good separation, inspired by SM analysis [J. High Energ. Phys. 2015, 184]
- BDT₂ selection → signal efficiency ϵ and define efficiency quantile $\bar{\epsilon} = 1 - \epsilon$
4 bins with **highest signal significance** keeping **constant integrated efficiency** over q_{rec}^2 : $\epsilon = 0.75\%$



Validation studies

Control channel: $B^0 \rightarrow K^{*0} J/\psi \rightarrow \mu^+ \mu^-$

Validation studies

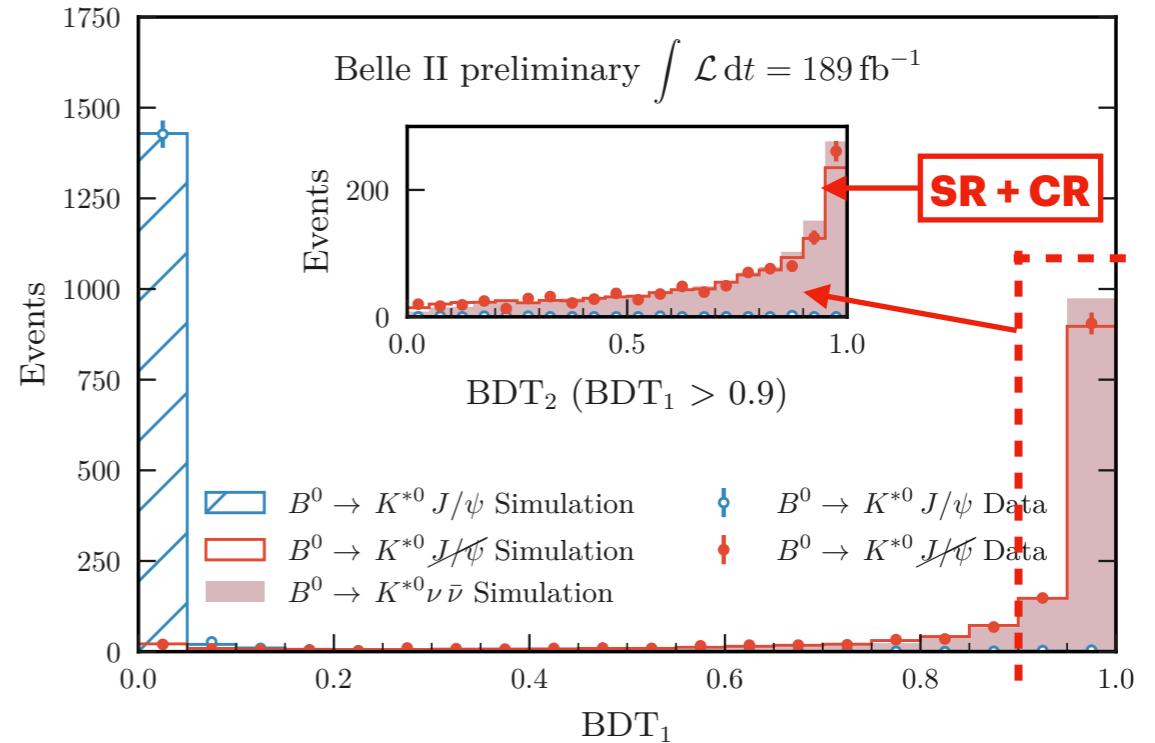
Control channel: $B^0 \rightarrow K^{*0} J/\psi \rightarrow \mu^+ \mu^-$

Extend the technique of the previous analysis
and validate BDT_1 and BDT_2 :

Validation studies

Control channel: $B^0 \rightarrow K^{*0} J/\psi \rightarrow \mu^+ \mu^-$

Extend the technique of the previous analysis
and validate BDT_1 and BDT_2 :
good overall data-MC agreement.

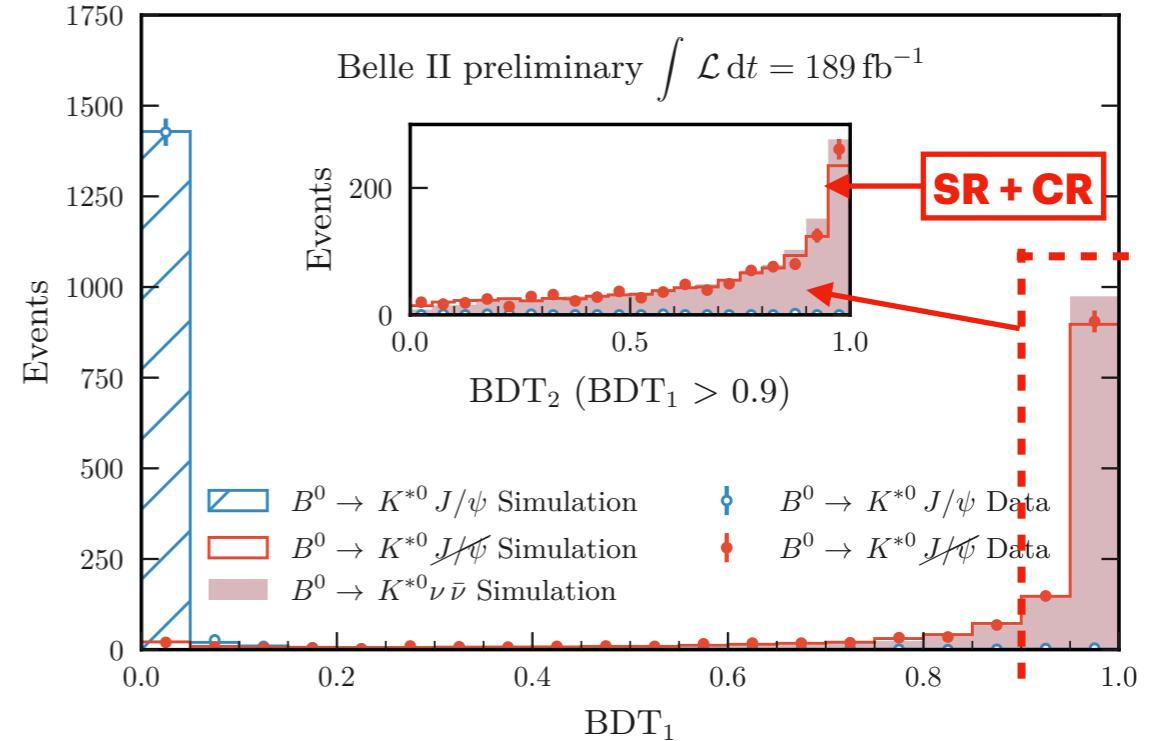


Validation studies

Control channel: $B^0 \rightarrow K^{*0} J/\psi \rightarrow \mu^+ \mu^-$

Extend the technique of the previous analysis
and validate BDT_1 and BDT_2 :
good overall data-MC agreement.

Off-resonance data



Check modelling of continuum simulation.

Validation studies

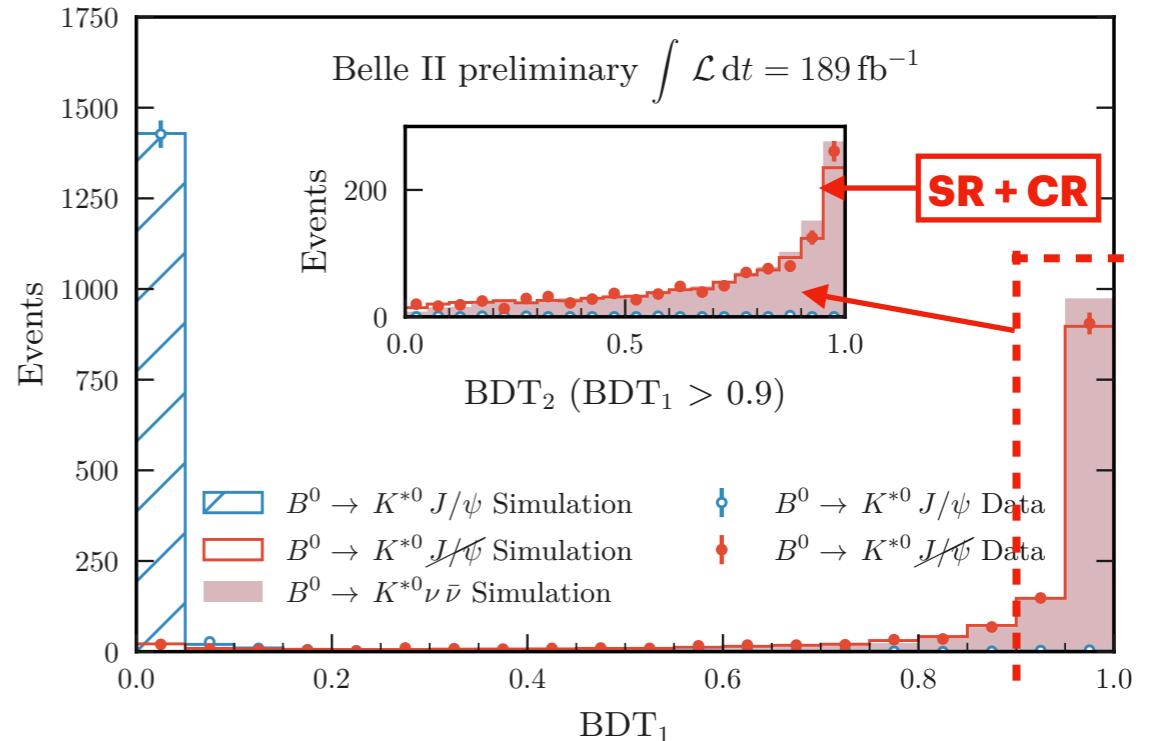
Control channel: $B^0 \rightarrow K^{*0} J/\psi \rightarrow \mu^+ \mu^-$

Extend the technique of the previous analysis
and validate BDT_1 and BDT_2 :
good overall data-MC agreement.

Off-resonance data

Check modelling of continuum simulation.

- Normalisation discrepancy:
 $\text{data/MC} = 1.41 \pm 0.11$



Validation studies

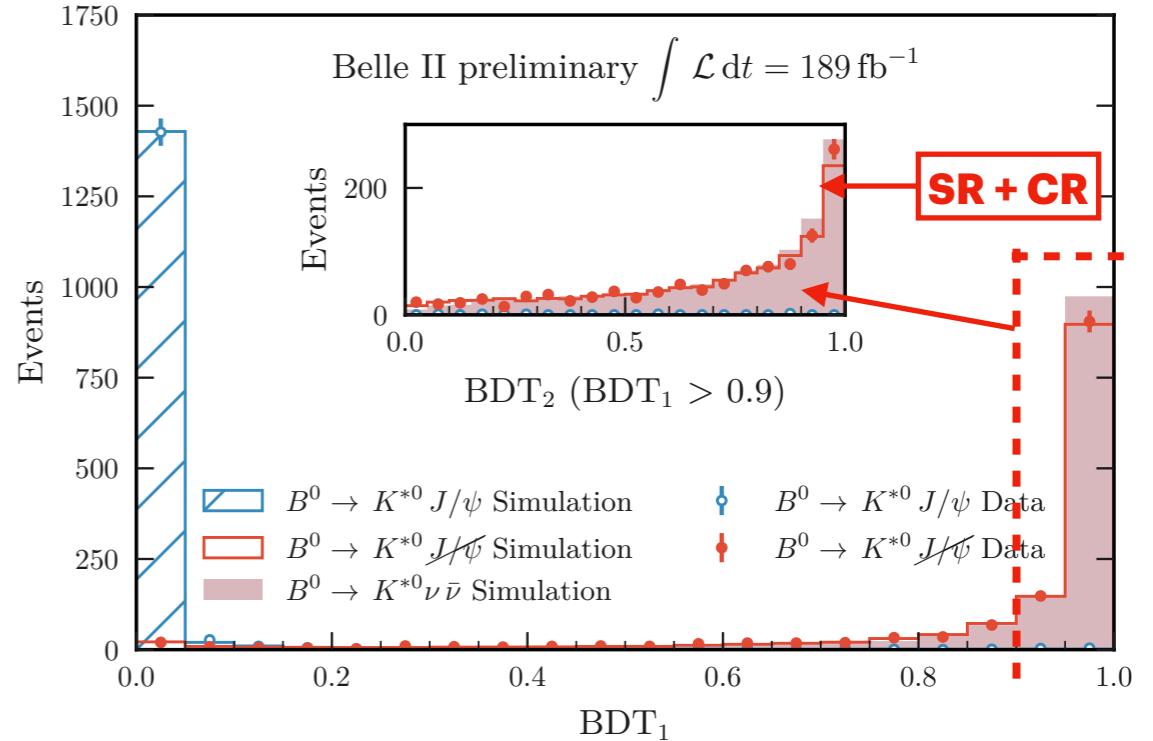
Control channel: $B^0 \rightarrow K^{*0} J/\psi \rightarrow \mu^+ \mu^-$

Extend the technique of the previous analysis
and validate BDT_1 and BDT_2 :
good overall data-MC agreement.

Off-resonance data

Check modelling of continuum simulation.

- Normalisation discrepancy:
 $\text{data/MC} = 1.41 \pm 0.11$
- Keep individual **50% normalisation uncertainty.**





Validation studies

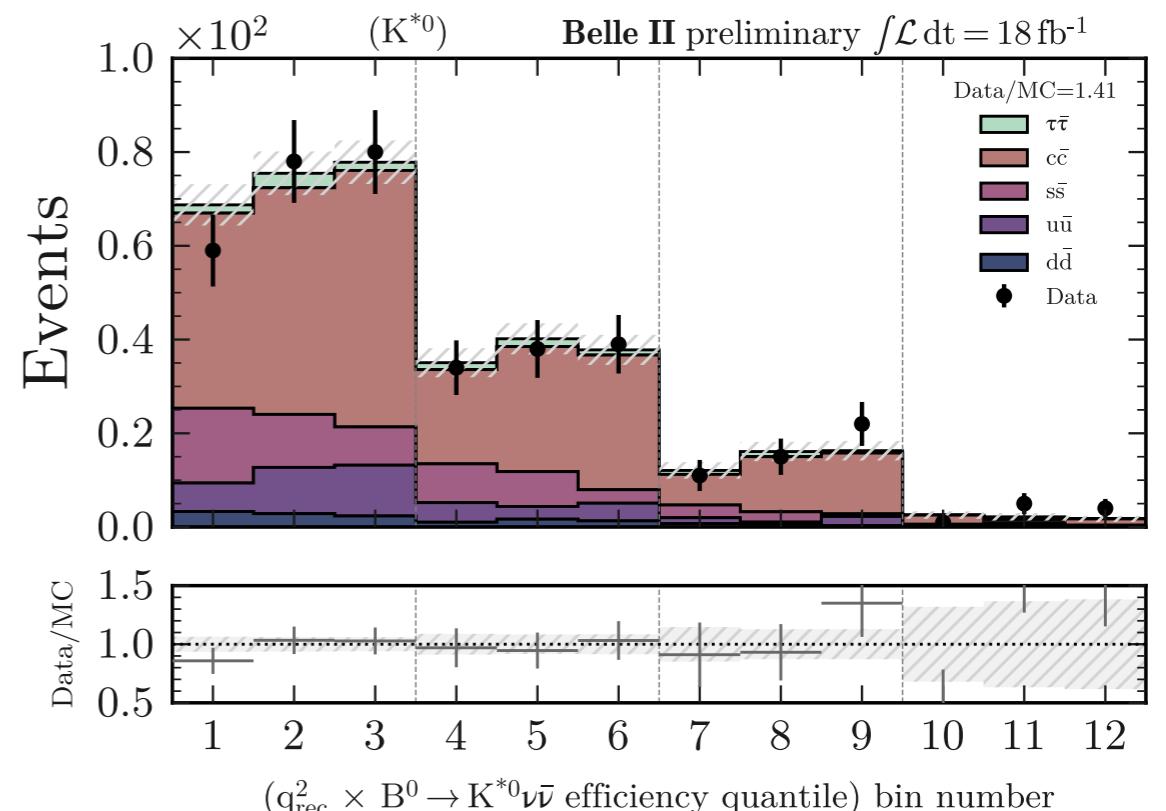
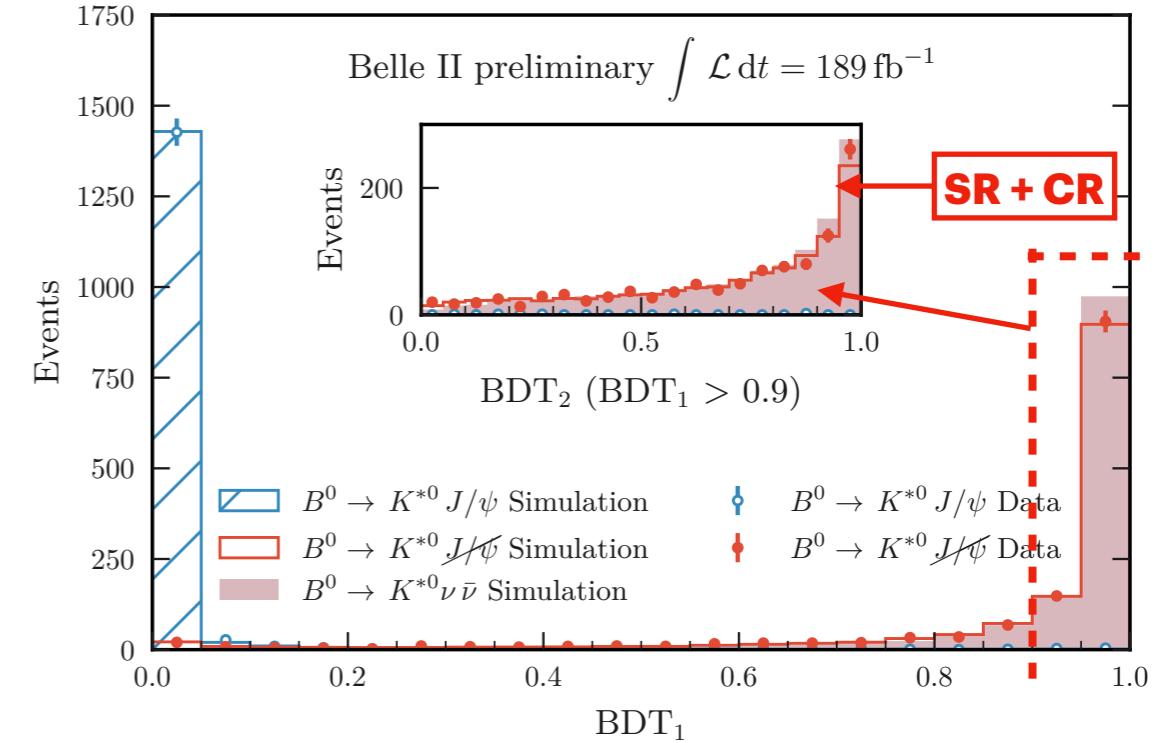
Control channel: $B^0 \rightarrow K^{*0} J/\psi \rightarrow \mu^+ \mu^-$

Extend the technique of the previous analysis
and validate BDT_1 and BDT_2 :
good overall data-MC agreement.

Off-resonance data

Check modelling of continuum simulation.

- Normalisation discrepancy:
 $\text{data/MC} = 1.41 \pm 0.11$
- Keep individual **50% normalisation uncertainty.**
- Good data-MC shape agreement. ✓



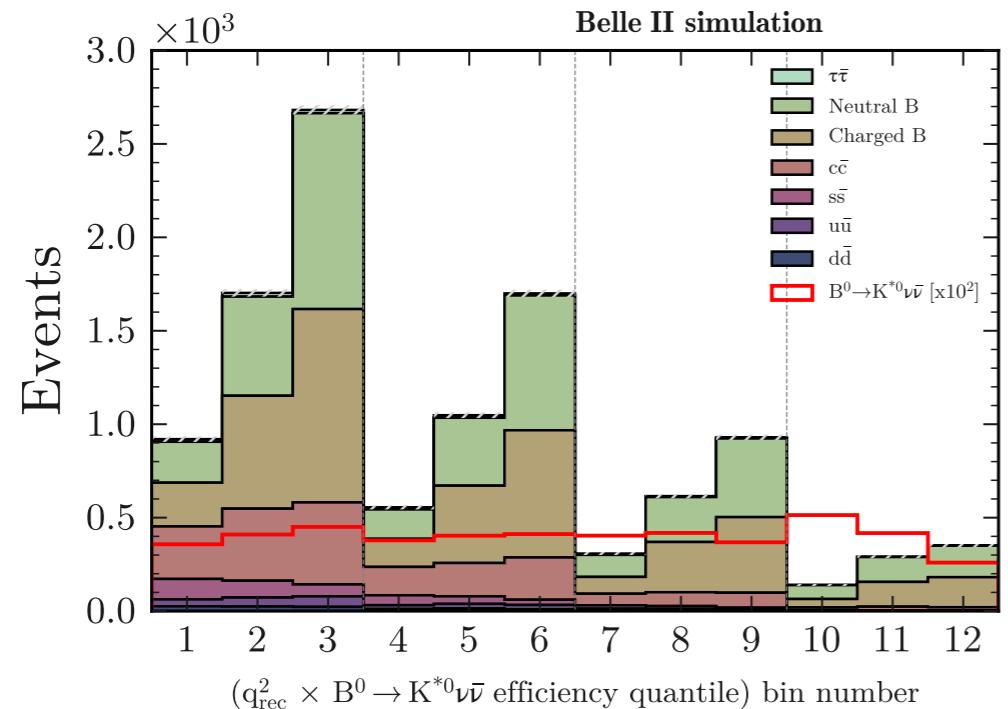
Statistical model

Overview

Statistical model

Overview

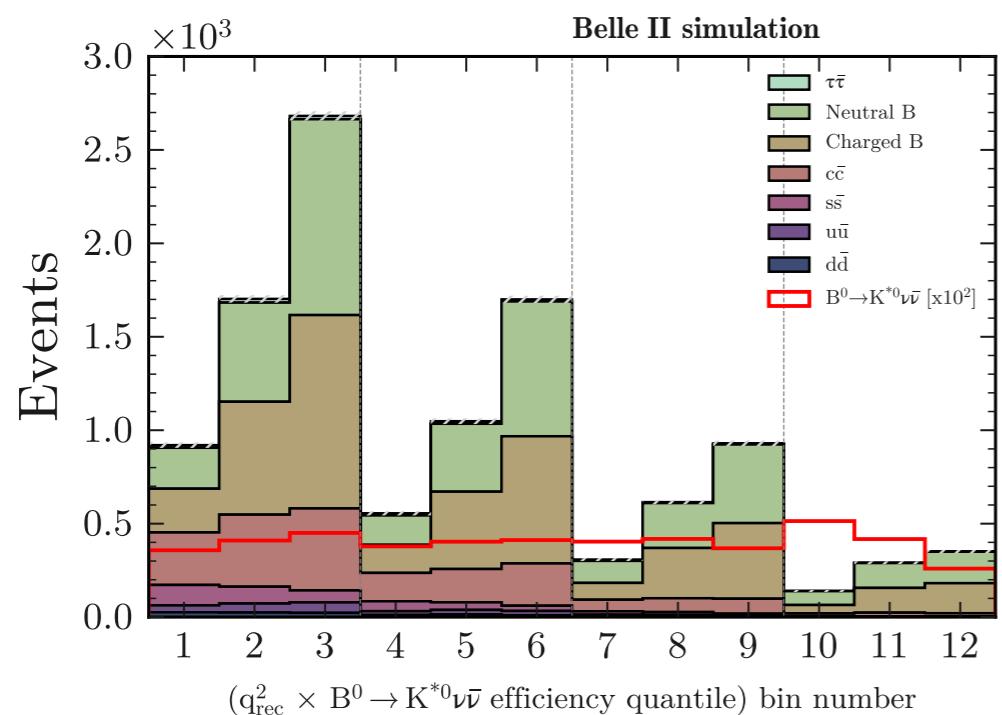
Signal and bkg templates from simulation.



Statistical model

Overview

Signal and bkg templates from simulation.

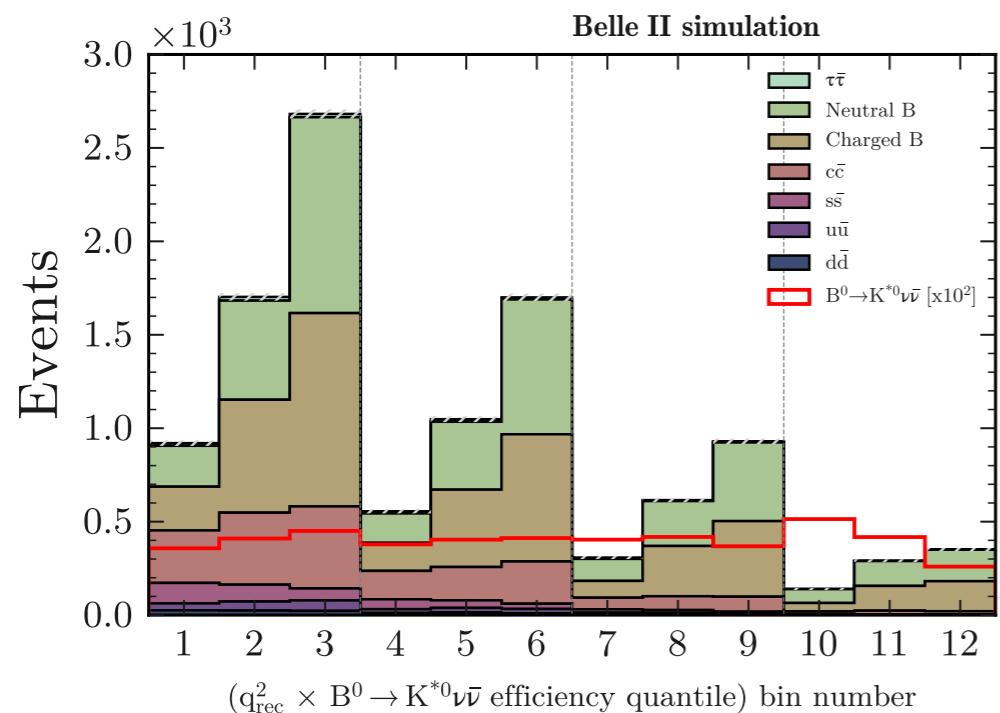


- **1 parameter of interest μ :**
 $\mu = 1 \rightarrow \text{SM BR} = 7.8 \times 10^{-6}$

Statistical model

Overview

Signal and bkg templates from simulation.

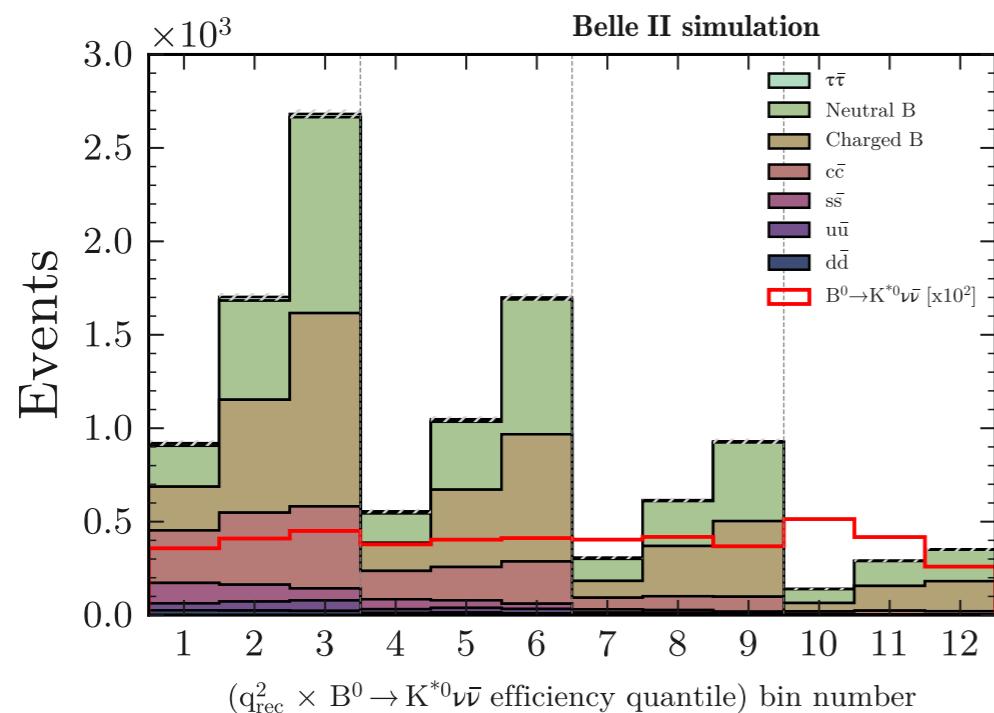


- **1 parameter of interest μ :**
 $\mu = 1 \rightarrow \text{SM BR} = 7.8 \times 10^{-6}$
- **190 nuisance parameters θ** accounting for MC stat. and **systematic uncertainties (34)**

Statistical model

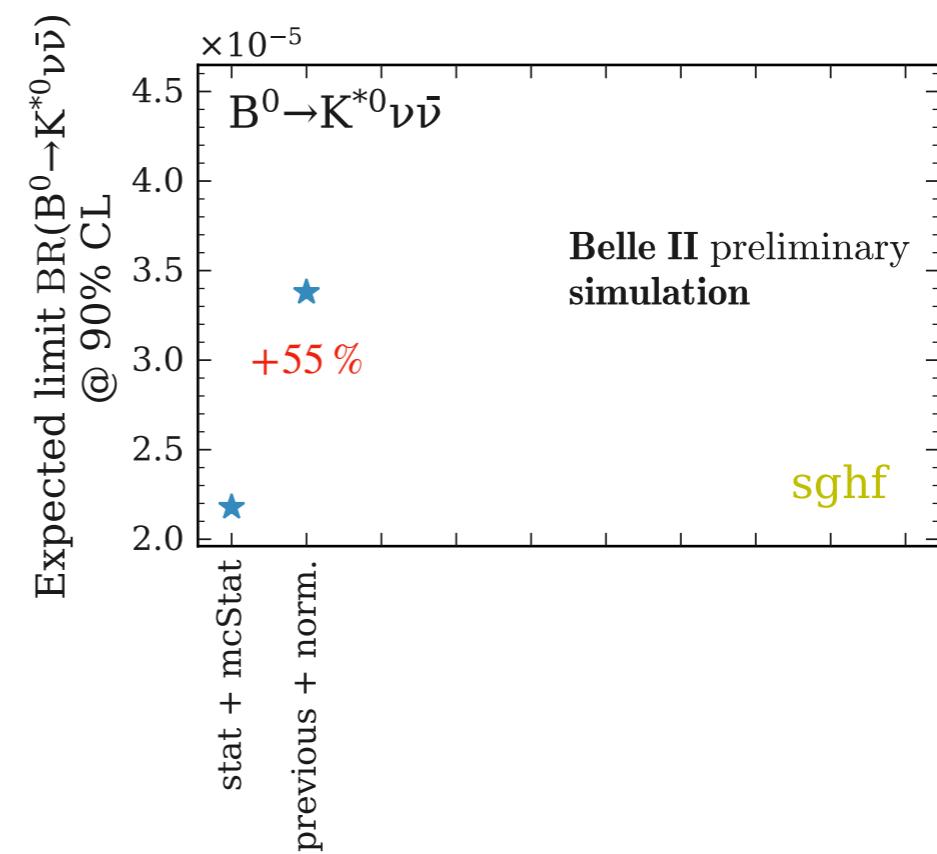
Overview

Signal and bkg templates from simulation.



- **Dominant:** normalisation of the 7 bkgs (50% each)

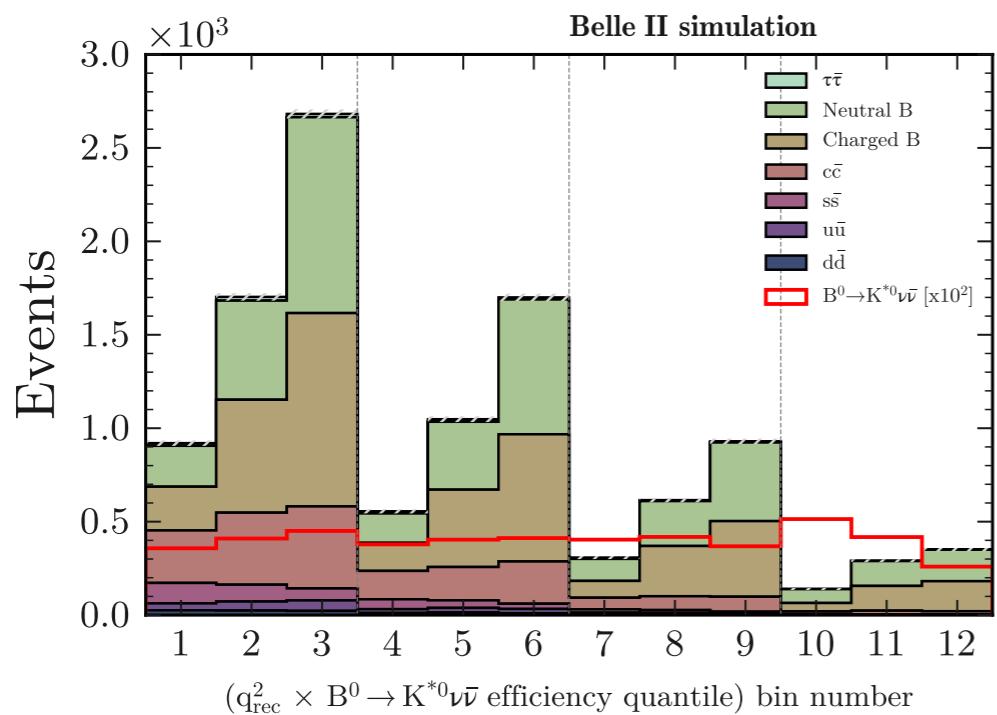
- **1 parameter of interest μ :**
 $\mu = 1 \rightarrow \text{SM BR} = 7.8 \times 10^{-6}$
- **190 nuisance parameters θ** accounting for MC stat. and **systematic uncertainties (34)**



Statistical model

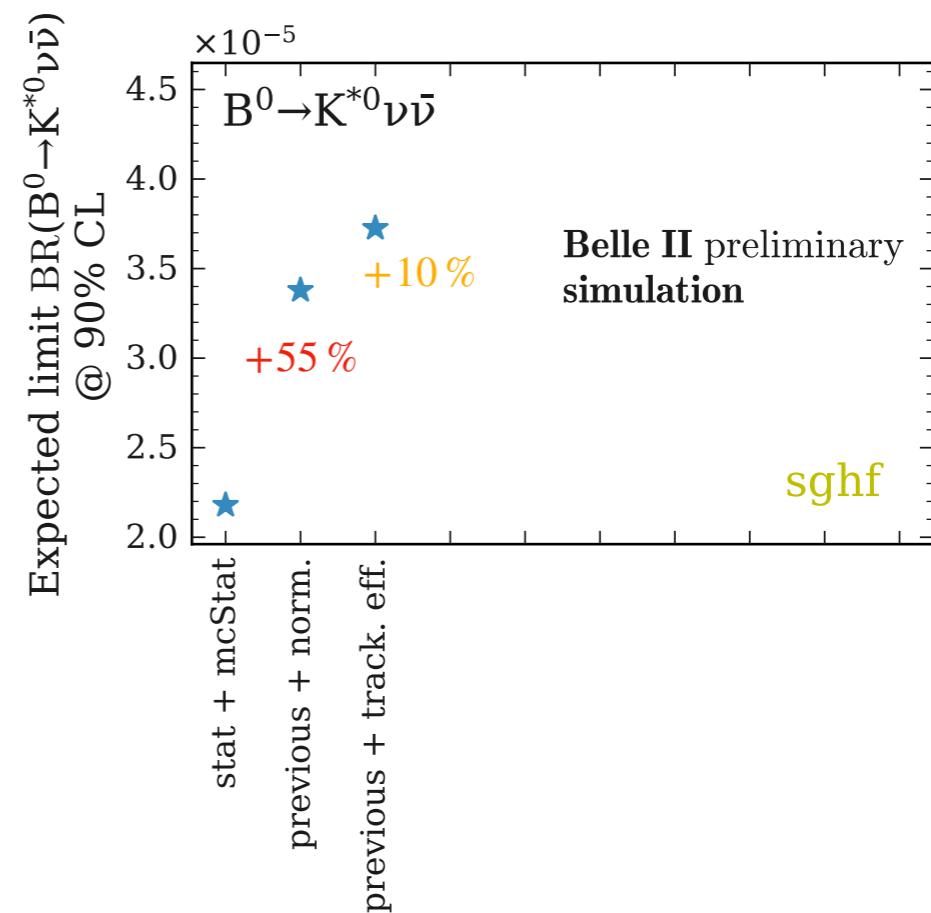
Overview

Signal and bkg templates from simulation.



- **Dominant:** normalisation of the 7 bkgs (50% each).
- **Sizeable:**
 - tracking efficiency

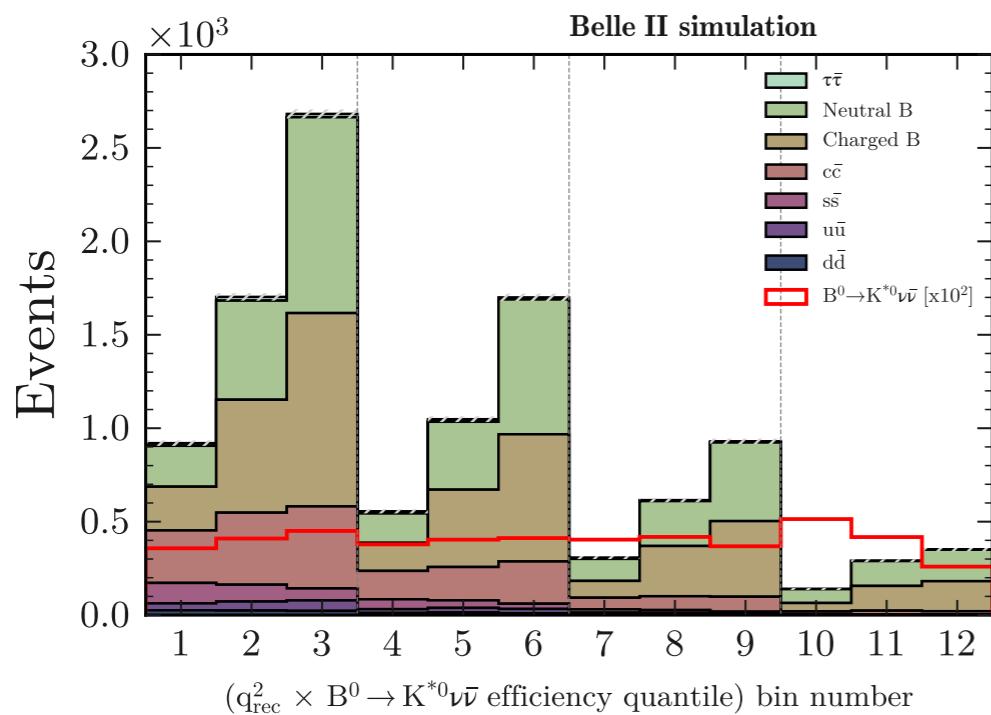
- **1 parameter of interest μ :**
$$\mu = 1 \rightarrow \text{SM BR} = 7.8 \times 10^{-6}$$
- **190 nuisance parameters θ** accounting for MC stat. and **systematic uncertainties (34)**



Statistical model

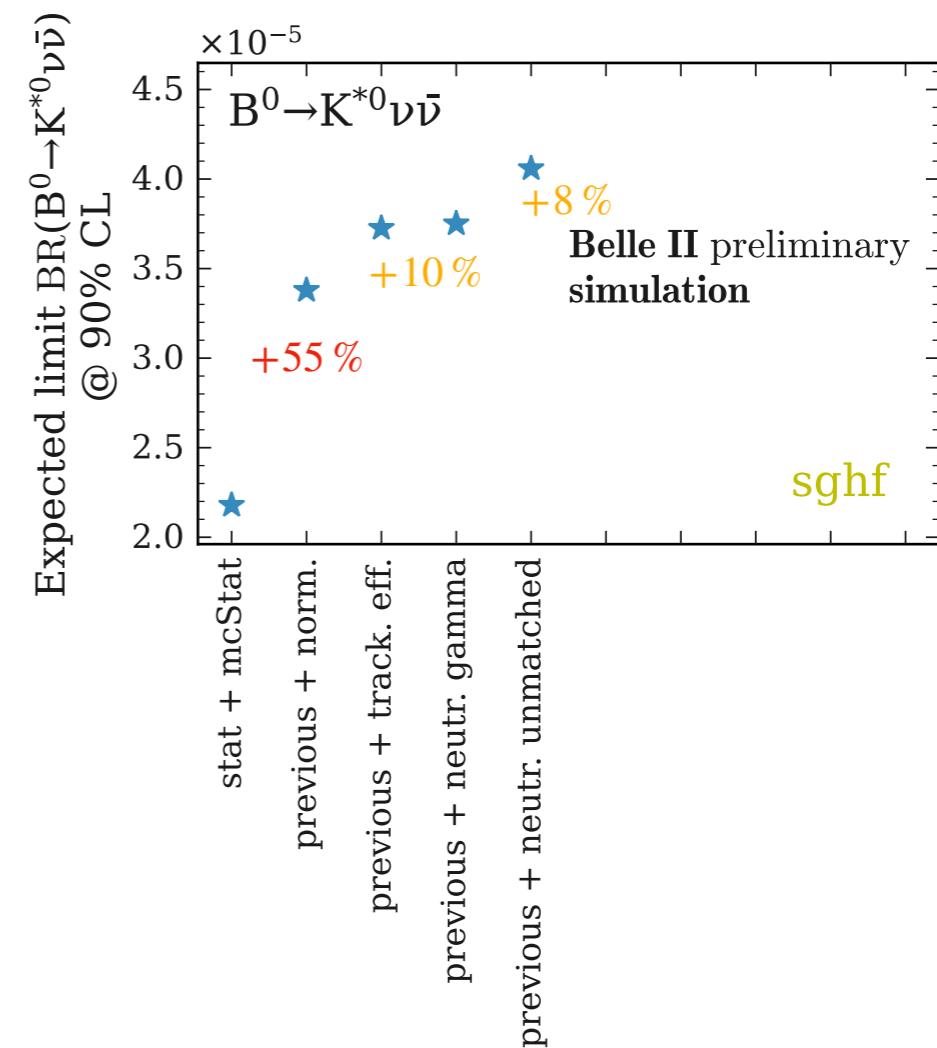
Overview

Signal and bkg templates from simulation.



- **Dominant:** normalisation of the 7 bkgs (50% each).
- **Sizeable:**
 - tracking efficiency
 - energy calibration of ECL clusters not matched to photons.

- **1 parameter of interest μ :**
$$\mu = 1 \rightarrow \text{SM BR} = 7.8 \times 10^{-6}$$
- **190 nuisance parameters θ** accounting for MC stat. and **systematic uncertainties (34)**

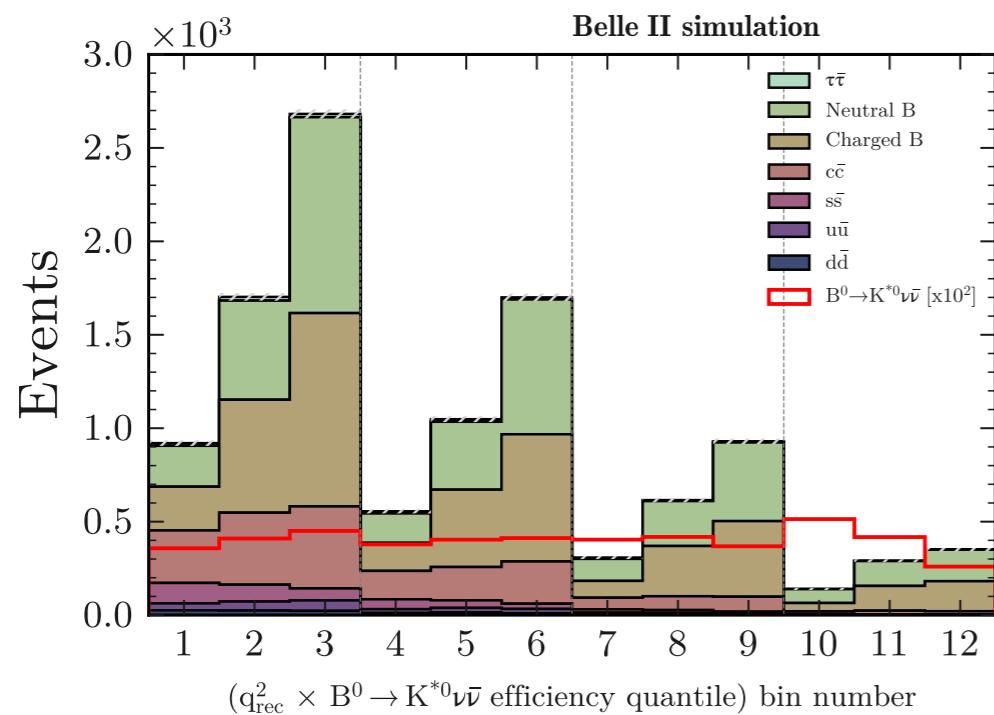




Statistical model

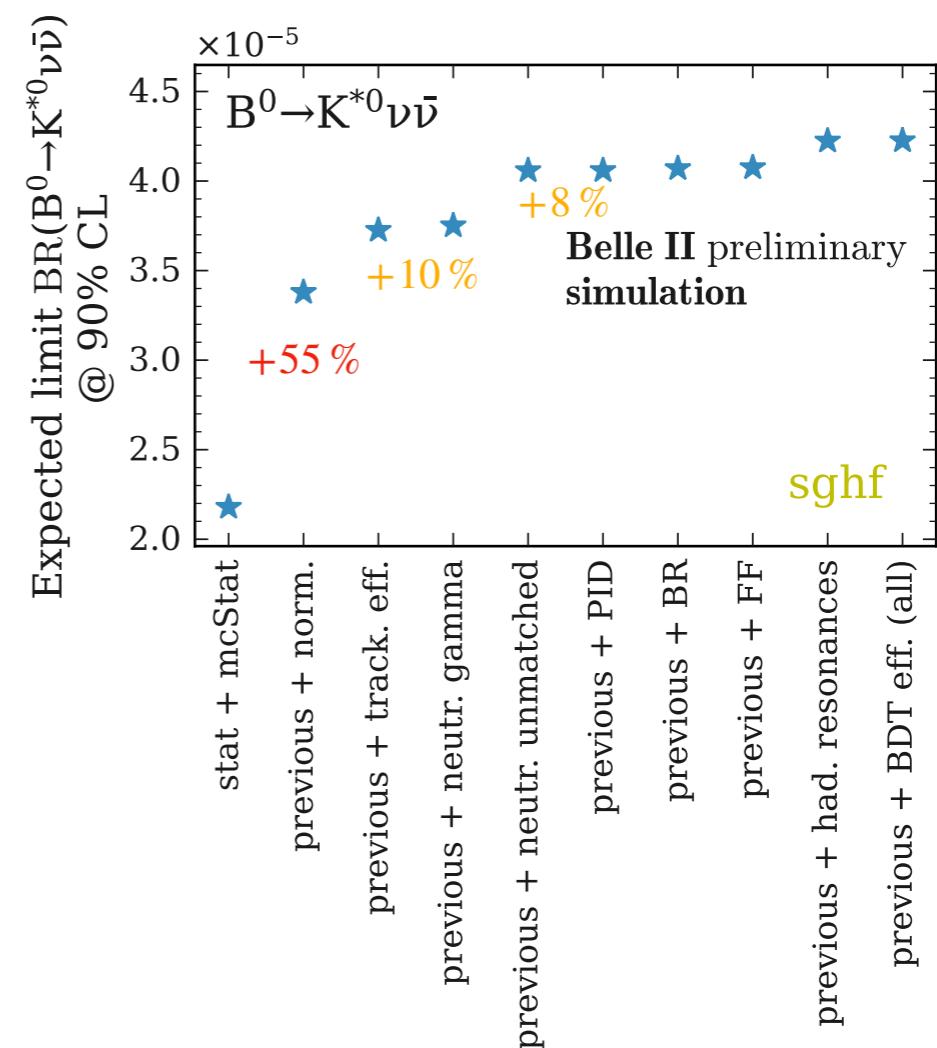
Overview

Signal and bkg templates from simulation.



- **Dominant**: normalisation of the 7 bkgs (50% each).
- **Sizeable**:
 - tracking efficiency
 - energy calibration of ECL clusters not matched to photons.

- **1 parameter of interest μ :**
 $\mu = 1 \rightarrow \text{SM BR} = 7.8 \times 10^{-6}$
- **190 nuisance parameters θ** accounting for MC stat. and **systematic uncertainties (34)**



Results

Results



Results based on MC expectations: analysis still under internal review.

Results



Results based on MC expectations: analysis still under internal review.

Expected upper limit on $\text{BR}(B^0 \rightarrow K^{*0} \nu \bar{\nu})$ at 90 % CL $\rightarrow 4.4 \times 10^{-5}$ (CL_s)

Results



Results based on MC expectations: analysis still under internal review.

Expected upper limit on $\text{BR}(B^0 \rightarrow K^{*0} \nu \bar{\nu})$ at **90 % CL** $\rightarrow 4.4 \times 10^{-5}$ (CL_s)

In the ballpark of previous results.

Results



Results based on MC expectations: analysis still under internal review.

Expected upper limit on $\text{BR}(B^0 \rightarrow K^{*0} \nu \bar{\nu})$ at 90 % CL $\rightarrow 4.4 \times 10^{-5}$ (CL_s)

In the ballpark of previous results.

Expected uncertainty on the BR: $\sigma_{\text{BR}} = 2.6 \times 10^{-5}$

Results



Results based on MC expectations: analysis still under internal review.

Expected upper limit on $\text{BR}(B^0 \rightarrow K^{*0} \nu \bar{\nu})$ at 90 % CL $\rightarrow 4.4 \times 10^{-5}$ (CL_s)

In the ballpark of previous results.

Expected uncertainty on the BR: $\sigma_{\text{BR}} = 2.6 \times 10^{-5}$

Precision comparison (assuming $\sigma_{\text{BR}} \sim 1/\sqrt{L}$)



Results



Results based on MC expectations: analysis still under internal review.

Expected upper limit on $\text{BR}(B^0 \rightarrow K^{*0} \nu \bar{\nu})$ at 90 % CL $\rightarrow 4.4 \times 10^{-5}$ (CL_s)

In the ballpark of previous results.

Expected uncertainty on the BR: $\sigma_{\text{BR}} = 2.6 \times 10^{-5}$

Precision comparison (assuming $\sigma_{\text{BR}} \sim 1/\sqrt{L}$)

Measurement	$L (\text{fb}^{-1})$	$\sigma_{\text{BR}} \times 10^5$	$\sigma_{\text{BR}}^{\text{scaled}} \times 10^5$
BaBar, hadronic and semileptonic taggings [124]	429	2.8	4.2
Belle, hadronic tagging [123]	711	6.2	12.0
Belle, semileptonic tagging [126]	711	1.1	2.1
Belle II, inclusive tagging (preliminary)	189	2.6	2.6

[123: Phys. Rev. D 2013, 87, 111103]

[124: Phys. Rev. D 2013, 87, 112005]

[126: Phys. Rev. D 2017, 87, 112005]

Inclusive tagging performs:

- $3.6 \times$ better than Belle had.
- 60 % better than BaBar had. + SL

Conclusions

Conclusions

- First search for $B^+ \rightarrow K^+\nu\bar{\nu}$ with an inclusive tagging method:
result **competitive** with previous measurements and **room for improvement**.
Important step towards future discovery.

Conclusions

- First search for $B^+ \rightarrow K^+\nu\bar{\nu}$ with an inclusive tagging method:
result **competitive** with previous measurements and **room for improvement**.
Important step towards future discovery.
- First search for $B^0 \rightarrow K^{*0}\nu\bar{\nu}$ with an inclusive tagging method:
measurement in a **good state, on the way to finalisation** and approval.
Encouraging prospects.

Conclusions

- First search for $B^+ \rightarrow K^+\nu\bar{\nu}$ with an inclusive tagging method:
result **competitive** with previous measurements and **room for improvement**.
Important step towards future discovery.
- First search for $B^0 \rightarrow K^{*0}\nu\bar{\nu}$ with an inclusive tagging method:
measurement in a **good state, on the way to finalisation** and approval.
Encouraging prospects.



...a challenging, great journey!

More to discuss...

Backup

- FCNC transitions
- $B \rightarrow K^{(*)}$ form factors
- Regimes of QCD
- $B \rightarrow K^* \nu \bar{\nu}$ angular analysis
- Non-factorisable effects in $B \rightarrow K^{(*)} l^+ l^-$
- The B anomalies
- Model-independent analysis of non-SM effects
- Z' boson
- Leptoquarks
- Dark scalar in $B \rightarrow K^{(*)} E$
- Explicit tagging methods
- SuperKEKB components
- SuperKEKB nanobeam
- Production cross-sections
- The Belle II detector
- Particle identification
- Generators
- Event-shape variables
- B-counting
- Statistical tools and methods
- Search for $B^+ \rightarrow K^+ \nu \bar{\nu}$
- Search for $B^0 \rightarrow K^{*0} \nu \bar{\nu}$
- Weights correction

Theory of $b \rightarrow s\nu\bar{\nu}$



FCNC transitions

[B. Grinstein. Lectures on Flavor Physics and CP Violation, Jan. 2017]

Suppression mechanism in $b \rightarrow s$:

$$\mathcal{M} \sim \frac{g^2}{16\pi^2} \sum_{i=u,c,t} V_{ib} V_{is}^* F\left(\frac{m_i^2}{m_W^2}\right) \xrightarrow{m_i \ll m_W} \mathcal{M} \sim \frac{g^2}{16\pi^2} \left[F\left(\frac{m_t^2}{m_W^2}\right) V_{tb} V_{ts}^* + F(0) \sum_{i=u,c} V_{ib} V_{is}^* \frac{m_i^2}{m_W^2} + \text{h.c.} \right]$$

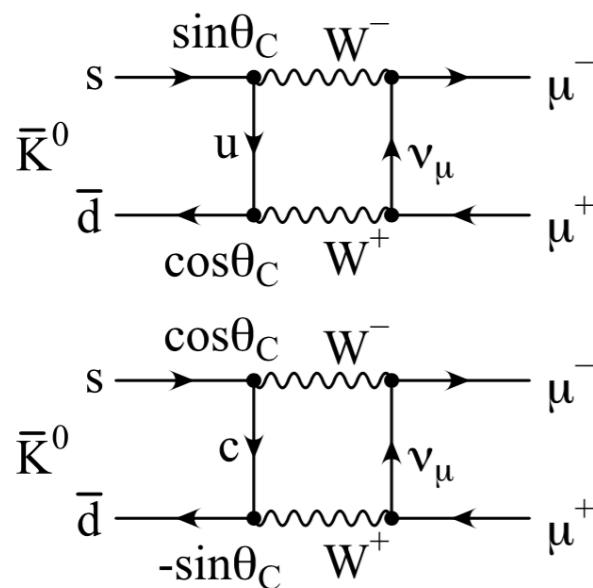
$F(x) \equiv$ function resulting from explicit calculation of the loop integral

$F(0) = 0$ and F increases with m_i up to $\mathcal{O}(1)$ for $m_i = m_t$

negligible

$$\mathcal{M} \sim \frac{g^2}{16\pi^2} V_{tb} V_{ts}^* F\left(\frac{m_t^2}{m_W^2}\right) \text{ transition dominated by virtual top-quark exchange}$$

Mechanism first proposed by Glashow, Iliopoulos and Maiani in 1970 (**GIM mechanism**) to explain suppression of $K^0 \bar{K}^0$ mixing and $K_L^0 \rightarrow \mu^+ \mu^-$ when only u, d, s were known.



Exploiting unitarity they found: $\mathcal{M} \propto \frac{m_c^2 - m_u^2}{m_W^2}$

Smallness set by mass-squared difference of the virtual u - c quarks exchanged.

$m_c \gg m_u$ prevents a complete cancellation of the interfering box diagrams.



Theory of $b \rightarrow s\nu\bar{\nu}$

$B \rightarrow K^{(*)}$ form factors

Hadronic form factors evaluated by means of **non-perturbative methods**:

- **Lattice QCD at large q^2** , good accuracy;
- **Light Cone Sum Rules (LCSR) at low q^2** ($\lesssim 15 \text{ GeV}^2/c^4$), accuracy $\sim 10\%$.
Approximate method: QCD vacuum-hadron correlation functions.

Parametrise the form factor with a **series expansion** and **interpolate lattice QCD and LCSR results**.

$$F(q^2) = \frac{1}{1 - q^2/m_+^2} \sum_k \alpha_k [z(q^2)]^k$$

- $m_+ = m_B + 0.046 \text{ GeV}/c^2$ (resonance mass)
- $\alpha_k \rightarrow$ real parameters to estimate: $(\alpha_0, \alpha_1, \alpha_2)$ after truncation.
- $z(t) = \frac{\sqrt{t_+ - q^2} - \sqrt{t_+ - t_0}}{\sqrt{t_+ - q^2} + \sqrt{t_+ - t_0}}$ $t_0 = t_+ \left(1 - \sqrt{1 - t_-/t_+} \right)$
- $t_\pm = (m_B \pm m_{K^{(*)}})^2$

- rescaled $B \rightarrow K$ FF: $\rho_K(q^2) = \frac{\lambda_K^{3/2}(q^2)}{m_B^4} [f_+^K(q^2)]^2$

- rescaled $B \rightarrow K^*$ FFs: $\rho_V(q^2) = \frac{2q^2 \lambda_{K^*}^{3/2}(q^2)}{(m_B + m_{K^*})^2 m_B^4} [V(q^2)]$, $\rho_{A_{12}}(q^2) = \frac{64m_{K^*}^2 \lambda_{K^*}^{1/2}(q^2)}{m_B^2} [A_{12}(q^2)]$

$$\rho_{A_1}(q^2) = \frac{2q^2 \lambda_{K^*}^{1/2}(q^2)(m_B + m_{K^*})^2}{m_B^4} [A_1(q^2)]$$

$$\lambda_{K^{(*)}} \equiv \lambda(m_B^2, m_{K^{(*)}}^2, q^2)$$

2nd order polynomial

$(\alpha_0, \alpha_1, \alpha_2)_{f_+}$
[J. High Energ. Phys. 2015, 184]

$(\alpha_0, \alpha_1, \alpha_2)_V$ $(\alpha_0, \alpha_1, \alpha_2)_{A_{12}}$
 $(\alpha_0, \alpha_1, \alpha_2)_{A_1}$
[J. High Energ. Phys. 2016, 98]



Theory of $b \rightarrow s\nu\bar{\nu}$

$B \rightarrow K^*\nu\bar{\nu}$: angular analysis

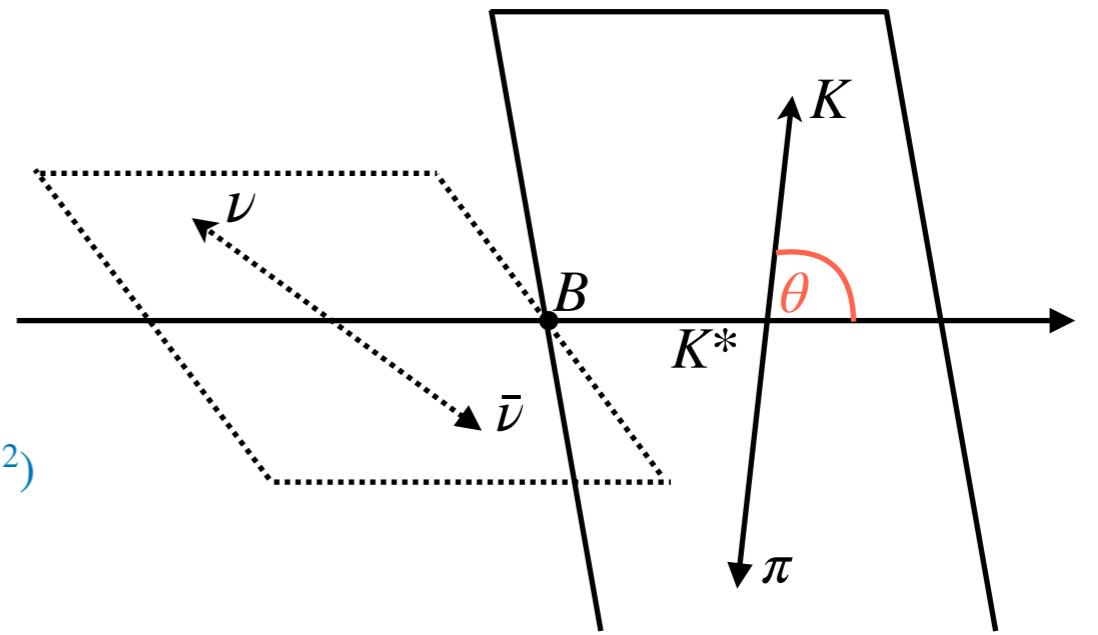
Pseudo-scalar ($J^P = 0^-$) $B^+(B^0)$ meson \rightarrow vector ($J^P = 1^-$) $K^{*+}(K^{*0})$ meson and neutrino pair.

$$\frac{d^2\text{BR}(B \rightarrow K^*\nu\bar{\nu})_{\text{SM}}}{dq^2 d\cos\theta} = \frac{3}{4} \frac{d\text{BR}_T}{dq^2} \sin^2\theta + \frac{3}{2} \frac{d\text{BR}_L}{dq^2} \cos^2\theta$$

K^* polarisation (spin orientation)
determines $K\pi$ angular distribution

$$\frac{d\text{BR}_L}{dq^2} = \tau_B 3 \left| N \cdot C_L^{\text{SM}} \right|^2 \rho_{A_1}(q^2)$$

$$\frac{d\text{BR}_T}{dq^2} = \tau_B 3 \left| N \cdot C_L^{\text{SM}} \right|^2 [\rho_{A_{12}}(q^2) + \rho_V(q^2)]$$





Theory of $b \rightarrow s\nu\bar{\nu}$

$B \rightarrow K^*\nu\bar{\nu}$: angular analysis

Pseudo-scalar ($J^P = 0^-$) $B^+(B^0)$ meson \rightarrow vector ($J^P = 1^-$) $K^{*+}(K^{*0})$ meson and neutrino pair.

Polarisation measurement

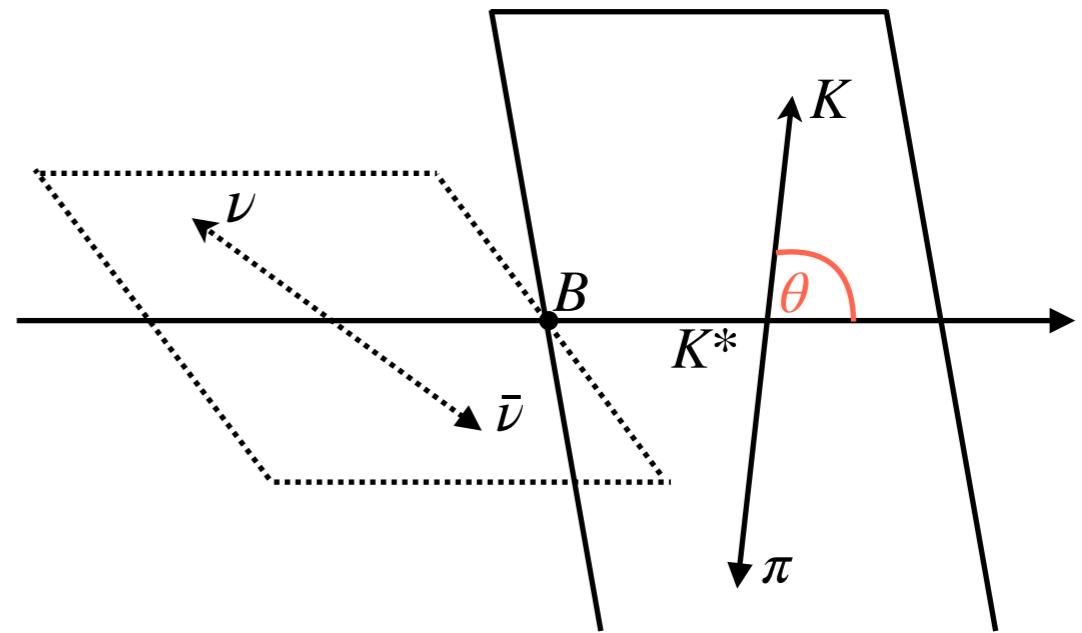
Extract polarisation fraction

$$F_{L(T)} = \frac{d\text{BR}_{L(T)}/dq^2}{d\text{BR}/dq^2}, \quad F_{L(T)} = 1 - F_{T(L)}$$

from **fit to the angular distribution** of kaon

$$\frac{d\text{BR}}{d\cos\theta} = \frac{3}{2}F_L \cos^2\theta + \frac{3}{4}(1 - F_L)(1 - \cos^2\theta)$$

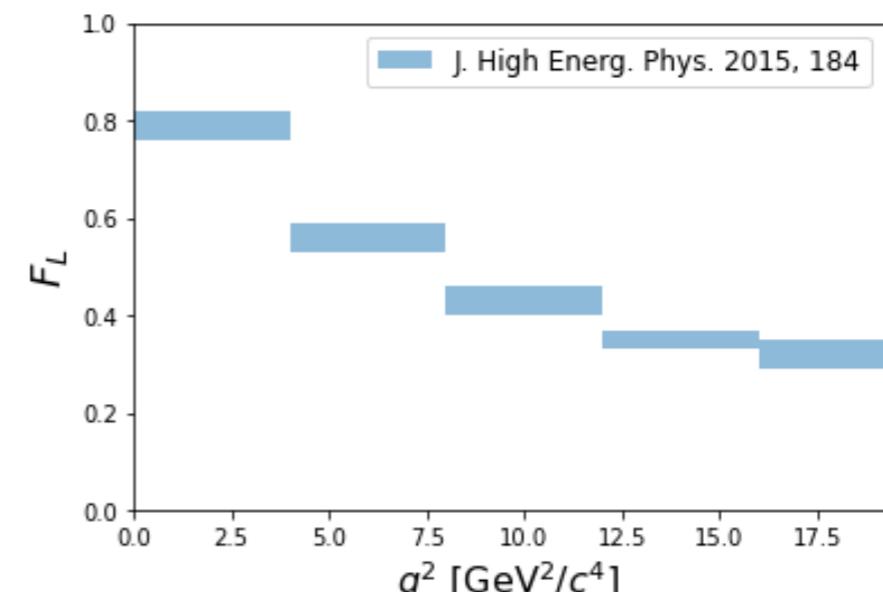
K^* polarisation (spin orientation)
determines $K\pi$ angular distribution



Clean observable: uncertainties due to form factors
and CKM elements factorise in the ratio

SM predictions from $B \rightarrow K^*\nu\bar{\nu}$ analysis: $F_L^{\text{SM}} = 0.47 \pm 0.03$

[J. High Energ. Phys. 2015, 184]

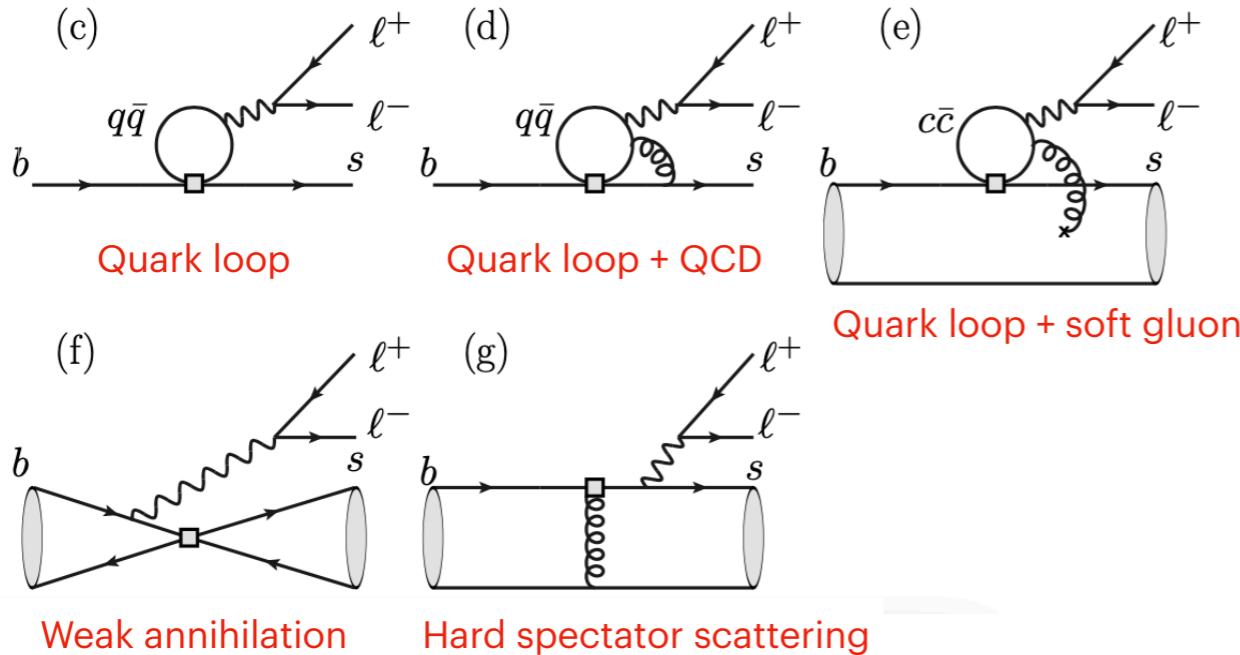


Non-factorisable effects in $B \rightarrow K^{(*)} l^+ l^-$



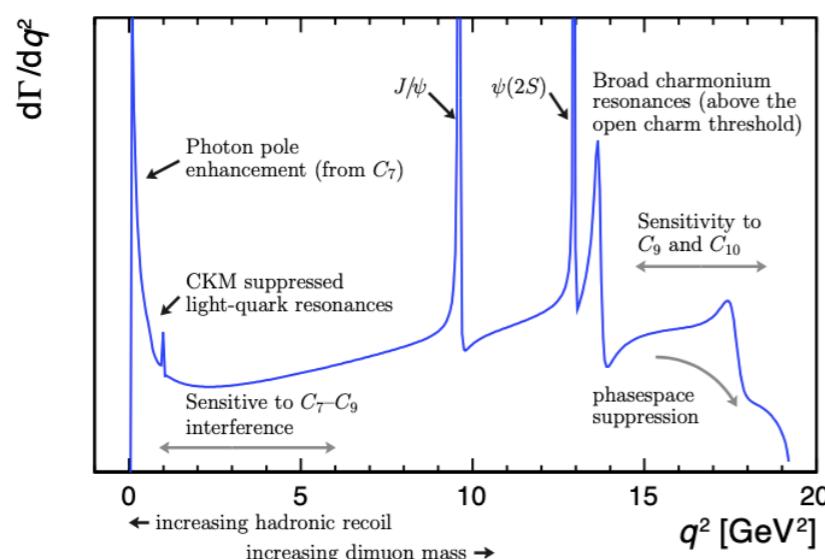
γ exchange results in non-factorisable contributions: major source of uncertainty in predictions.

Dominant non-factorisable effects from 4-quark operators: (c), (d), (e).



Size and methods to compute them depend on q^2 .

- $q^2 \lesssim 1 \text{ GeV}^2$: dominant contribution from (c), CKM-suppressed light-quark resonances ρ, ω, ϕ etc.
- Low q^2 region: (f), (g) and (d); **major uncertainty from (e), soft-gluon correction to charm loop** estimated in LCSR.
- Intermediate q^2 region: dominant contribution from (c), the **charm loop goes on shell (J/ψ and $\psi(2S)$ resonances)** \rightarrow **decays become non-leptonic**. Vetoed regions in experimental analyses.
- $q^2 > 15 \text{ GeV}^2$: main contribution from (c), broad $c\bar{c}$ resonances.



The B anomalies



Since 2013 measurements of semileptonic B decays are *in tension* with the SM.

Experimental results point to a **possible violation** of **Lepton Flavour Universality**.

Different behaviour (beside kinematical effects) of **different lepton species** seems to be observed in

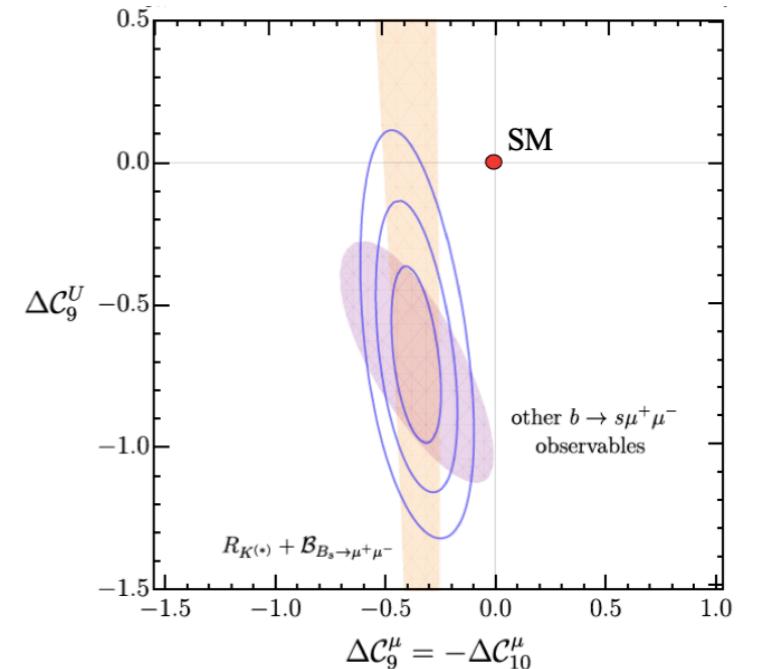
- $b \rightarrow sl^+l^-$ (neutral currents) $\longrightarrow \mu$ vs e

List of **observables exhibiting anomalies** (=deviations from SM predictions):

- P'_5 anomaly [$B \rightarrow K^*\mu^+\mu^-$ angular distribution] $\rightarrow \sim 2.5\sigma$ in some q^2 bins.
- smallness of all $B \rightarrow H_s \mu^+\mu^-$ decay rates [$H_s = K, K^*, \phi$ (from B_s)] $\rightarrow \sim 3\sigma$ at low q^2 .
- LFU ratios (μ vs e) in $B \rightarrow K^*l^+l^-$ and $B \rightarrow Kl^+l^-$ \rightarrow from $\sim 2.1\sigma$ to $\sim 3.1\sigma$.
- smallness of BR($B_s \rightarrow \mu^+\mu^-$) $\rightarrow \sim 2.3\sigma$ deviation.

[combined ATLAS+CMS+LHCb 2021]

[J. High Energ. Phys. 2021, 50]



Deficit of the muon modes.

Fit including all observables.

Significance of **BSM** hypothesis **vs SM** $> 5\sigma$

$$\Delta C_i^\mu = C_i^\mu - C_i^e$$

$$\mathcal{O}_{10}^l = (\bar{s}_L \gamma_\mu b_L)(\bar{l} \gamma^\mu \gamma^5 l)$$

$$\mathcal{O}_9^l = (\bar{s}_L \gamma_\mu b_L)(\bar{l} \gamma^\mu l)$$

The B anomalies



Since 2013 measurements of semileptonic B decays are *in tension* with the SM.

Experimental results point to a **possible violation** of **Lepton Flavour Universality**.

Different behaviour (beside kinematical effects) of **different lepton species** seems to be observed in

- $b \rightarrow c l \bar{\nu}$ (charged currents) $\longrightarrow \tau$ vs light leptons (μ, e)

$$R(X_c) \equiv \frac{\text{BR}(B \rightarrow X_c \tau^- \bar{\nu})}{\text{BR}(B \rightarrow X_c l^- \bar{\nu})} \quad X_c = D^+, D^{*+}$$

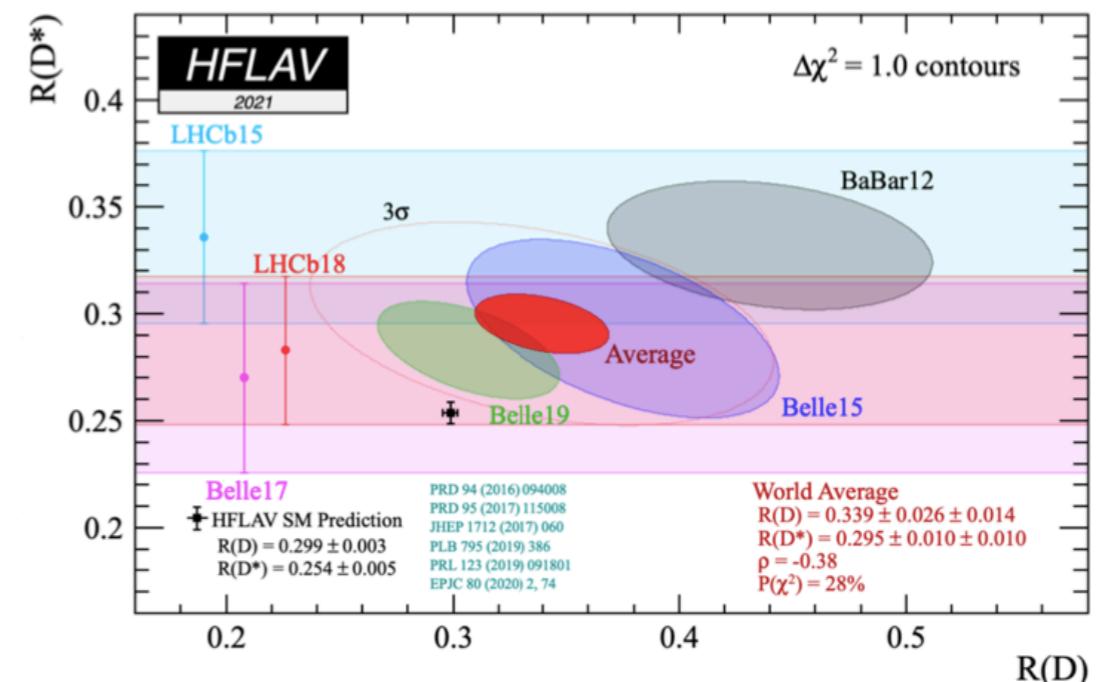
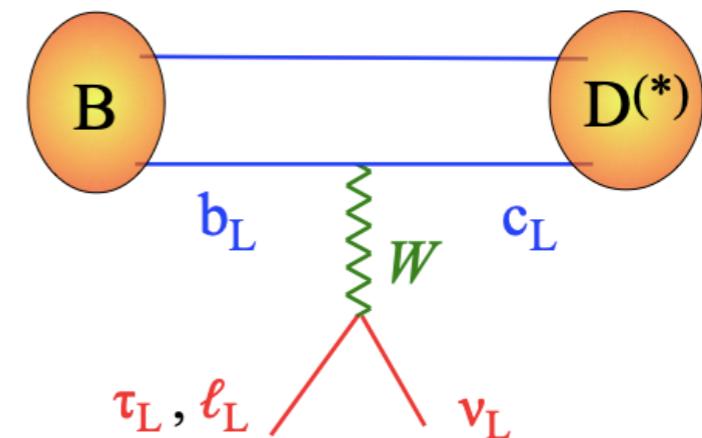
$R(D)_{\text{SM}} = 0.299 \pm 0.003$ [Phys. Rev. D 2016, 94]

$R(D^*)_{\text{SM}} = 0.258 \pm 0.005$ [J. High Energ. Phys. 2017, 61]

(not unity due to phase-space of τ)

Measured by BaBar, Belle and LHCb:

combined $\sim 3.1\sigma$ deviation from the SM





Theory of $b \rightarrow s\nu\bar{\nu}$

Model-independent analysis of non-SM effects

- 1) Simple extension of the SM in a **low-energy effective theory** [J. High Energ. Phys. 2015, 184]

- preserve LFU (same NP- ν couplings for the 3 flavours)
- assume $M_{\text{NP}} > m_B$

$$\mathcal{H}_{\text{eff.}} = -\frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* (C_L O_L + \textcolor{blue}{C_R} O_R + \text{h.c.})$$

$O_R = \frac{e^2}{16\pi^2} [\bar{s}\gamma_\mu P_R b][\bar{\nu}\gamma^\mu(1-\gamma^5)\nu]$ accounts for **BSM right-handed interactions** via $P_R = \frac{1+\gamma^5}{2}$

Wilson Coefficients C_L, C_R used to define the **test variables** $\epsilon = \frac{\sqrt{|C_L|^2 + |C_R|^2}}{|C_L^{\text{SM}}|}$ and $\eta = \frac{-\text{Re}(C_L C_R^*)}{|C_L|^2 + |C_R|^2}$

- $\epsilon = 1, \eta = 0 \rightarrow$ only SM interactions
- $\epsilon \neq 1, \eta \neq 0 \rightarrow$ signals right-handed currents

evaluated in fits to experimental measurements.

$$\mathcal{R}_K^\nu = \frac{\text{BR}(B \rightarrow K\nu\bar{\nu})}{\text{BR}(B \rightarrow K\nu\bar{\nu})_{\text{SM}}} = (1 - 2\eta)\epsilon^2, \quad \mathcal{R}_{K^*}^\nu = \frac{\text{BR}(B \rightarrow K^*\nu\bar{\nu})}{\text{BR}(B \rightarrow K^*\nu\bar{\nu})_{\text{SM}}} = (1 + k_\eta\eta)\epsilon^2, \quad \mathcal{R}_{F_L} = \frac{F_L}{F_L^{\text{SM}}} = \frac{1 + 2\eta}{1 + k_\eta\eta},$$

$k_\eta \equiv$ parameter dependent on FFs ratio

Theory of $b \rightarrow s\nu\bar{\nu}$



Model-independent analysis of non-SM effects

2) **SM Effective Field Theory** (SM-EFT) [Nucl. Phys. B 1986, 268; J. High Energy Phys. 2010, 10]

- correlate $b \rightarrow s\nu\bar{\nu}$ and $b \rightarrow sl^+l^-$ via $SU(2)_L$ symmetry in lepton sector
- expansion in a basis of **dimension-6 operators** Q_i invariant under $SU(2)_L \times U(1)_Y$
- integrate out only **NP above the electroweak scale** Λ_{EW}

$\mathbf{q}_L \equiv$ left-handed quark doublets
 $\mathbf{l}_L \equiv$ left-handed lepton doublets
 $\Phi \equiv$ Higgs doublet

↷ $c_i \equiv c_i(\Lambda_{EW})$ Wilson Coefficients

$$\mathcal{L}^{(6)} = \sum_i \frac{c_i}{\Lambda^2} Q_i$$

↓

$\Lambda \equiv$ NP scale

Rescaling (after EWSB) to low energy

$$\mathcal{H}_{eff} = -\frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{e^2}{16\pi^2} \sum_i C_i O_i$$

which includes

$$O_9^{(')} = (\bar{s}\gamma_\mu P_{L(R)} b)(\bar{l}\gamma^\mu l)$$

$$O_{10}^{(')} = (\bar{s}\gamma_\mu P_{L(R)} b)(\bar{l}\gamma^\mu \gamma_5 l)$$

$$O_L = (\bar{s}\gamma_\mu P_L b)(\bar{\nu}\gamma^\mu(1-\gamma_5)\nu)$$

$$O_R = (\bar{s}\gamma_\mu P_R b)(\bar{\nu}\gamma^\mu(1-\gamma_5)\nu)$$

Operators contributing to $b \rightarrow s\nu\bar{\nu}$ and $b \rightarrow sl^+l^-$

$$Q_{\Phi q}^{(1)} = i(\bar{\mathbf{q}}_L \gamma_\mu \mathbf{q}_L) \Phi^\dagger \mathcal{D}^\mu \Phi$$

$$Q_{\Phi q}^{(3)} = i(\bar{\mathbf{q}}_L \gamma_\mu \sigma^a \mathbf{q}_L) \Phi^\dagger \mathcal{D}^\mu \sigma_a \Phi$$

$$Q_{\Phi d} = i(\bar{\mathbf{d}}_R \gamma_\mu \mathbf{d}_R) \Phi^\dagger \mathcal{D}^\mu \Phi$$

... and only to $b \rightarrow sl^+l^-$

$$Q_{ql}^{(1)} = (\bar{\mathbf{q}}_L \gamma_\mu \mathbf{q}_L)(\bar{\mathbf{l}}_L \gamma^\mu \mathbf{l}_L)$$

$$Q_{ql}^{(3)} = (\bar{\mathbf{q}}_L \gamma_\mu \sigma^a \mathbf{q}_L)(\bar{\mathbf{l}}_L \gamma^\mu \sigma_a \mathbf{l}_L)$$

$$Q_{dl} = (\bar{\mathbf{d}}_L \gamma_\mu \mathbf{d}_L)(\bar{\mathbf{l}}_L \gamma^\mu \mathbf{l}_L)$$

$$Q_{de} = (\bar{\mathbf{d}}_R \gamma_\mu \mathbf{d}_R)(\bar{\mathbf{e}}_R \gamma^\mu \mathbf{e}_R)$$

$$Q_{qe} = (\bar{\mathbf{q}}_L \gamma_\mu \mathbf{q}_L)(\bar{\mathbf{e}}_R \gamma^\mu \mathbf{e}_R)$$

$$C_L^{\text{NP}} = C_L^{\text{SM}} + \tilde{c}_{ql}^{(1)} - \tilde{c}_{ql}^{(3)} + \tilde{c}_Z$$

$$C_9^{\text{NP}} = C_9^{\text{SM}} + \tilde{c}_{qe} + \tilde{c}_{ql}^{(1)} + \tilde{c}_{ql}^3 - \xi \tilde{c}_Z \quad C'_9 = \tilde{c}_{de} + \tilde{c}_{dl} - \xi \tilde{c}'_Z$$

$$C_{10}^{\text{NP}} = C_{10}^{\text{SM}} + \tilde{c}_{qe} - \tilde{c}_{ql}^{(1)} - \tilde{c}_{ql}^3 + \xi \tilde{c}_Z \quad C'_{10} = \tilde{c}_{de} - \tilde{c}_{dl} + \xi \tilde{c}'_Z$$

Estimated in global fits to measured observables

EWSB: electroweak symmetry breaking

Theory of $b \rightarrow s\nu\bar{\nu}$



Model-independent analysis of non-SM effects

2) **SM Effective Field Theory** (SM-EFT) [Nucl. Phys. B 1986, 268; J. High Energy Phys. 2010, 10]

- correlate $b \rightarrow s\nu\bar{\nu}$ and $b \rightarrow sl^+l^-$ via $SU(2)_L$ symmetry in lepton sector
- expansion in a basis of **dimension-6 operators** Q_i invariant under $SU(2)_L \times U(1)_Y$
- integrate out only **NP above the electroweak scale** Λ_{EW}

Recent analysis [Phys. Rev. D 2022, 105]

fits $b \rightarrow sl^+l^-$ measurements and checks implications on $b \rightarrow s\nu\bar{\nu}$ in SM-EFT.

4 scenarios provide good compatibility with $b \rightarrow sl^+l^-$ anomalies but would correspond to $> 3\sigma$ deviations from the SM prediction for $B \rightarrow K^{(*)}\nu\bar{\nu}$

SM-EFT couplings	$BR(B \rightarrow K\nu\bar{\nu}) \times 10^{-6}$	$BR(B \rightarrow K^*\nu\bar{\nu}) \times 10^{-6}$
SM	4.006 ± 0.261	9.331 ± 0.744
$(\tilde{c}_{ql}^{(3)}, \tilde{c}'_Z)$	5.485 ± 0.358	3.197 ± 0.223
$(\tilde{c}_Z, \tilde{c}'_Z)$	3.260 ± 0.213	2.141 ± 0.146
$(\tilde{c}_{ql}^{(1)} + \tilde{c}_{ql}^{(3)}, \tilde{c}_Z)$	7.419 ± 0.484	17.284 ± 1.378
$(\tilde{c}_{ql}^{(1)} + \tilde{c}_{ql}^{(3)}, \tilde{c}'_Z)$	2.319 ± 0.151	12.857 ± 1.111



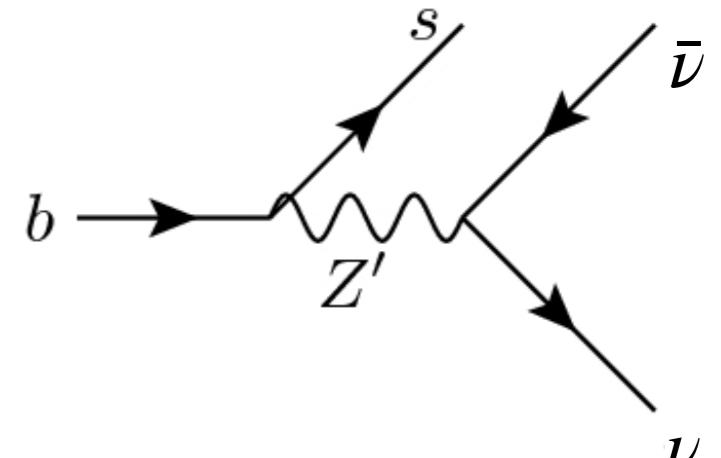
Theory of $b \rightarrow s\nu\bar{\nu}$

Models with new mediators

Z' boson

- Massive, neutral, vector gauge-boson from $U(1)$ **extensions of the SM**
- $U(1)$ can originate from larger spontaneously broken symmetry:
e.g. $E_6 \rightarrow SO(10) \times U(1)_\psi \rightarrow SU(5) \times U(1)_\chi \times U(1)_\psi \rightarrow SM \times U(1)_\beta$
- Predicted by GUTs, $L_\mu - L_\tau$ model, $B - L$ models.

$M_{Z'}$ from $\mathcal{O}(10 \text{ GeV})$ to several TeV depending on the model.



Can mediate **FCNC transitions at tree level**

Interaction of **generic Z'** , transforming as $SU(2)_L$ singlet, with SM fermions [J. High. Energy Phys. 2015, 184]

$$\mathcal{L}_{Z'} = \bar{\mathbf{f}}_i \gamma^\mu \left[\Delta_L^{ij}(Z') P_L + \Delta_R^{ij}(Z') P_R \right] \mathbf{f}_j Z'_\mu \quad (SU(2)_L \rightarrow \Delta_L^{\nu\bar{\nu}} = \Delta_L^{ll})$$

Tree-level contributions to $b \rightarrow s\nu\bar{\nu}$ and $b \rightarrow sl^+l^-$

depend on **Z' -fermion couplings ($\Delta_{L(R)}^{ij}$)** and **Z' mass:**

$$\tilde{c}_{ql}^{(1)} = -\frac{\Delta_L^{sb}\Delta_L^{ll}}{V_{tb}V_{ts}^*} \left[\frac{5 \text{ TeV}}{M_{Z'}} \right]^2$$

$$\tilde{c}_{dl} = -\frac{\Delta_R^{sb}\Delta_L^{ll}}{V_{tb}V_{ts}^*} \left[\frac{5 \text{ TeV}}{M_{Z'}} \right]^2$$

$$\tilde{c}_{qe} = -\frac{\Delta_L^{sb}\Delta_R^{ll}}{V_{tb}V_{ts}^*} \left[\frac{5 \text{ TeV}}{M_{Z'}} \right]^2$$

$$\tilde{c}_{de} = -\frac{\Delta_R^{sb}\Delta_R^{ll}}{V_{tb}V_{ts}^*} \left[\frac{5 \text{ TeV}}{M_{Z'}} \right]^2$$

↓
low-energy

only LH couplings: $C_9^{\text{NP}} = -C_{10}^{\text{NP}} = \tilde{c}_{ql}^{(1)}$

LH + RH couplings: $C_9^{\text{NP}} = -C_{10}^{\text{NP}} = \tilde{c}_{ql}^{(1)}$ $C'_9 = -C'_{10} = \tilde{c}_{dl}$



Theory of $b \rightarrow s\nu\bar{\nu}$

Models with new mediators

Z' boson

- Massive, neutral, vector gauge-boson from $U(1)$ **extensions of the SM**
- $U(1)$ can originate from larger spontaneously broken symmetry:
e.g. $E_6 \rightarrow SO(10) \times U(1)_\psi \rightarrow SU(5) \times U(1)_\chi \times U(1)_\psi \rightarrow SM \times U(1)_\beta$
- Predicted by GUTs, $L_\mu - L_\tau$ model, $B - L$ models.

Can mediate **FCNC transitions at tree level**

Generic Z'

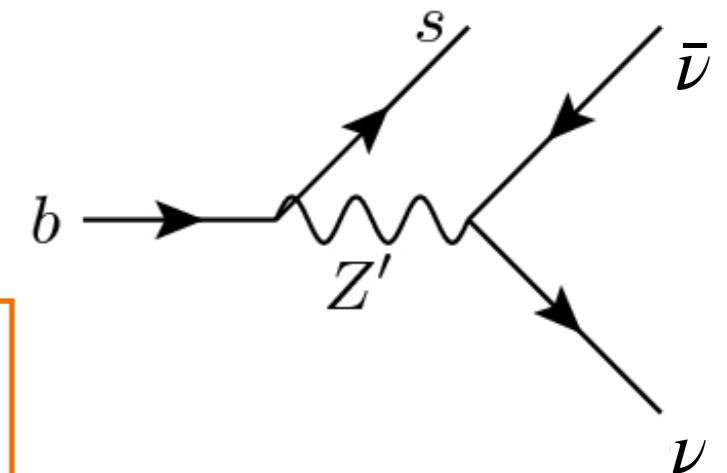
In recent analysis [Phys. Rev. D 2021, 104] **good fit** to $b \rightarrow sl^+l^-$ anomalies obtained via $C_9^{\text{NP}} = -C_{10}^{\text{NP}} = -0.41^{+0.07}_{-0.07}$ [Eur. Phys. J. C 2021, 81] for

$$\Delta_L^{sb} = (8.5 \pm 6.4) \times 10^{-3}, \Delta_L^{\mu\mu} = 2.00 \pm 0.95 \text{ with } M_{Z'} = 5 \text{ TeV}/c^2$$

Constrained parameter space compatible with current R_K^ν data but **prediction** is $R_K^\nu = 1.05 \pm 0.03$, **close to SM expectation**.

Turning on also Δ_R^{sb} \rightarrow prediction: $R_K^\nu \approx 1.1$

$M_{Z'}$ from $\mathcal{O}(10 \text{ GeV})$ to several TeV depending on the model.



Current experimental average:

$$\mathcal{R}_K^\nu = \frac{\text{BR}(B \rightarrow K\nu\bar{\nu})}{\text{BR}(B \rightarrow K\nu\bar{\nu})_{\text{SM}}} = 2.4 \pm 0.9$$

Theory of $b \rightarrow s\nu\bar{\nu}$

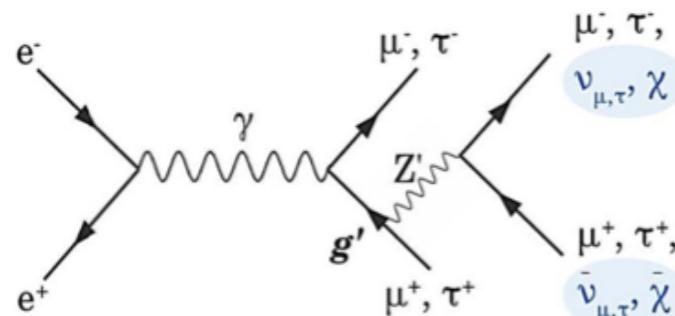


Models with new mediators

Z' boson

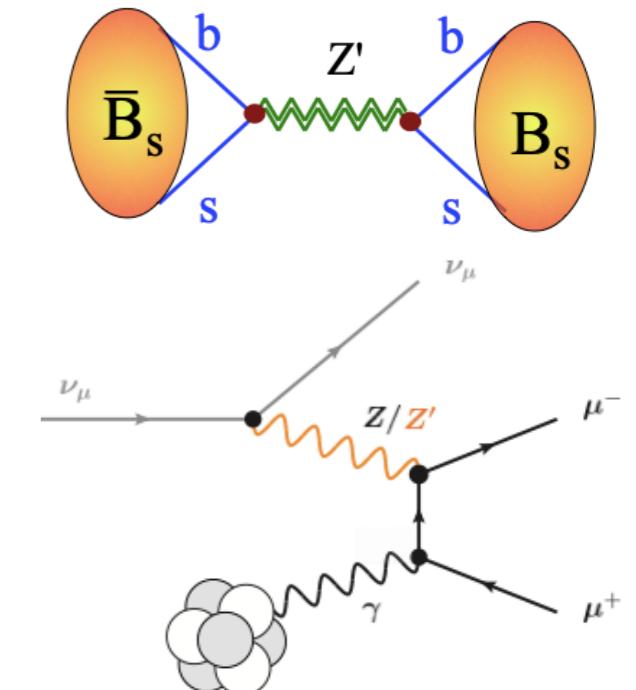
Low-energy:

- bounds from **4-quark processes** (e.g., $B_S^0 - \bar{B}_S^0$ oscillation) and $\tau \rightarrow l\nu\bar{\nu}$ where a Z' would contribute at the tree level.
- **neutrino-trident** in the Coulomb field of a heavy nucleus, e.g. $\nu_\mu \rightarrow \nu_\mu \mu^+ \mu^-$
- $(L_\mu - L_\tau)$ searches for $e^+e^- \rightarrow \mu^+\mu^-(\tau^+\tau^-)Z'$, with $Z' \rightarrow l, \nu, X$



Z' signature: peak in the distribution of the invariant mass of the system recoiling against a $\mu^+\mu^-(\tau^+\tau^-)$ pair.

Belle II: UL on Z' coupling for $M_{Z'} \leq 6 \text{ GeV}/c^2$ [Phys. Rev. Lett. 2020, 124]



At LHC:

Z' coupling to quarks in s-channel \rightarrow **resonance** in the invariant mass distribution of decay products.

A possible process at LHC is $pp \rightarrow Z'X \rightarrow f\bar{f}X$:

- Searches for $Z' \rightarrow e^+e^-(\mu^+\mu^-)$ by ATLAS and CMS: lower mass limit of $4.9 \text{ TeV}/c^2$ [Phys. Lett. B 2019, 796]
- Searches for $Z' \rightarrow b\bar{b}$ by ATLAS and CMS: lower mass limit of $2.7 \text{ TeV}/c^2$ [J. High Energy. Phys. 2020, 145]



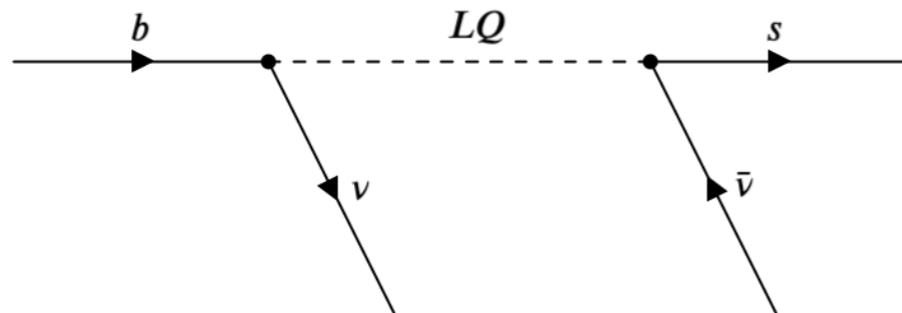
Theory of $b \rightarrow s\nu\bar{\nu}$

Models with new mediators

Leptoquarks (LQ)

- Hypothetical **bosons carrying both baryon number and lepton number.**
- Predicted in extensions of the SM unifying quark and leptons: Pati Salam model, $SU(5)$, $SO(10)$ GUTs.
- Can be **scalar** (spin-0) **or vector** (spin-1).

Can mediate $b \rightarrow s\nu\bar{\nu}$ and $b \rightarrow sl^+l^-$ **at tree (or loop) level — depending on the model.**



Expected mass is $\mathcal{O}(\text{TeV})$.

Good LQ candidates mediating $b \rightarrow s\nu\bar{\nu}$ and $b \rightarrow sl^+l^-$ at tree level and accommodating the current experimental data are [Phys. Rev. D 2021, 104]:

- $S_3(\bar{3}, 3, 1/3)$ scalar \longrightarrow interaction term: $+\bar{q}_L^c Y_{S_3} i\sigma_2 \boldsymbol{\sigma} \cdot \mathbf{S}_3 \mathbf{l}_L$
 $SU(3)_C \times SU(2)_L \times U(1)_Y$

Y_{S_3} : Yukawa matrix in flavour space

$f^c \equiv C\bar{f}$: charge conjugated

σ : pauli matrices

High-energy:
$$\frac{[\tilde{c}_{ql}^{(1)}]^{ij\alpha\beta}}{\Lambda^2} = 3 \frac{[\tilde{c}_{ql}^{(3)}]^{ij\alpha\beta}}{\Lambda^2} = \frac{3 Y_{S_3}^{j\beta} Y_{S_3}^{*i\alpha}}{4 M_{S_3}^2}$$

 (SM-EFT)

Low-energy: $C_L^{\text{NP}} = \tilde{c}_{ql}^{(1)} - \tilde{c}_{ql}^{(3)}, \quad C_9^{\text{NP}} = -C_{10}^{\text{NP}} = \tilde{c}_{ql}^{(1)} + \tilde{c}_{ql}^{(3)}$



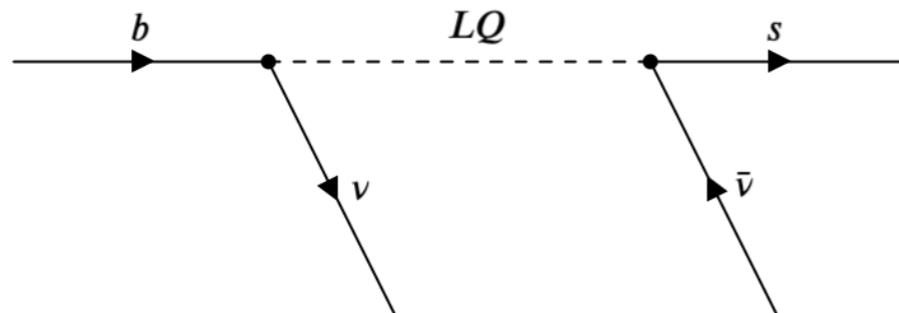
Theory of $b \rightarrow s\nu\bar{\nu}$

Models with new mediators

Leptoquarks (LQ)

- Hypothetical **bosons carrying both baryon number and lepton number.**
- Predicted in extensions of the SM unifying quark and leptons: Pati Salam model, $SU(5)$, $SO(10)$ GUTs.
- Can be **scalar** (spin-0) **or vector** (spin-1).

Can mediate $b \rightarrow s\nu\bar{\nu}$ and $b \rightarrow sl^+l^-$ **at tree (or loop) level — depending on the model.**



Expected mass is $\mathcal{O}(\text{TeV})$.

Good LQ candidates mediating $b \rightarrow s\nu\bar{\nu}$ and $b \rightarrow sl^+l^-$ at tree level and accommodating the current experimental data are [Phys. Rev. D 2021, 104]:

- $S_3(\bar{3}, 3, 1/3)$ scalar \longrightarrow interaction term: $+\bar{q}_L^c Y_{S_3} i\sigma_2 \boldsymbol{\sigma} \cdot \mathbf{S}_3 \mathbf{l}_L$
 $SU(3)_C \times SU(2)_L \times U(1)_Y$

Y_{S_3} : Yukawa matrix in flavour space

$f^c \equiv C\bar{f}$: charge conjugated

σ : pauli matrices

With $M_{S_3} = 2 \text{ TeV}/c^2$: $b \rightarrow sl^+l^-$ anomalies and current

$R_K^\nu = 2.4 \pm 0.9$ explained by turning on $Y_{S_3}^{22}, Y_{S_3}^{23}, Y_{S_3}^{32}, Y_{S_3}^{33}$.



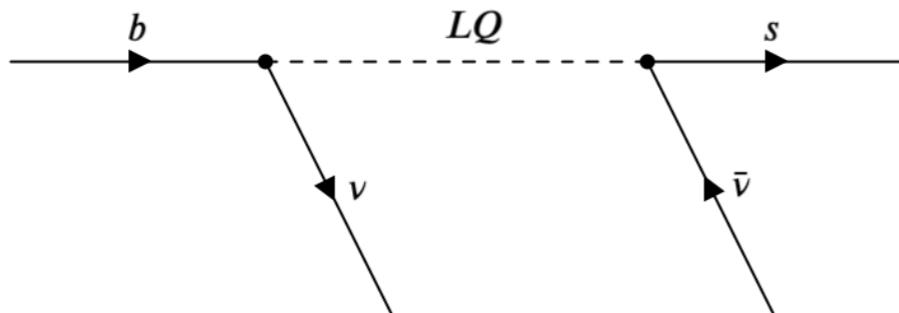
Theory of $b \rightarrow s\nu\bar{\nu}$

Models with new mediators

Leptoquarks (LQ)

- Hypothetical **bosons carrying both baryon number and lepton number.**
- Predicted in extensions of the SM unifying quark and leptons: Pati Salam model, $SU(5)$, $SO(10)$ GUTs.
- Can be **scalar** (spin-0) **or vector** (spin-1).

Can mediate $b \rightarrow s\nu\bar{\nu}$ and $b \rightarrow sl^+l^-$ **at tree (or loop) level — depending on the model.**



Expected mass is $\mathcal{O}(\text{TeV})$.

Good LQ candidates mediating $b \rightarrow s\nu\bar{\nu}$ and $b \rightarrow sl^+l^-$ at tree level and accommodating the current experimental data are [Phys. Rev. D 2021, 104]:

- $U_3^\mu(3, 3, 2/3)$ vector \longrightarrow interaction term: $+\bar{q}_L \gamma^\mu \sigma^a Y_{U_3} l_L U_{3\mu}^a$

Y_{S_3} : Yukawa matrix in flavour space

$f^c \equiv C\bar{f}$: charge conjugated

σ : pauli matrices



High-energy:
$$\frac{[\tilde{c}_{ql}^{(1)}]^{ij\alpha\beta}}{\Lambda^2} = -3 \frac{[\tilde{c}_{ql}^{(3)}]^{ij\alpha\beta}}{\Lambda^2} = -\frac{3Y_{U_3}^{j\beta} Y_{U_3}^{*i\alpha}}{2M_{U_3}^2}$$

(SM-EFT)

Low-energy: $C_L^{\text{NP}} = \tilde{c}_{ql}^{(1)} - \tilde{c}_{ql}^{(3)}$, $C_9^{\text{NP}} = -C_{10}^{\text{NP}} = \tilde{c}_{ql}^{(1)} + \tilde{c}_{ql}^{(3)}$



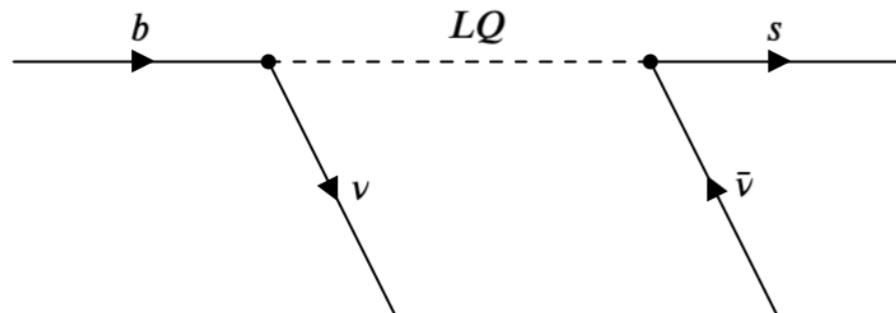
Theory of $b \rightarrow s\nu\bar{\nu}$

Models with new mediators

Leptoquarks (LQ)

- Hypothetical **bosons carrying both baryon number and lepton number.**
- Predicted in extensions of the SM unifying quark and leptons: Pati Salam model, $SU(5)$, $SO(10)$ GUTs.
- Can be **scalar** (spin-0) **or vector** (spin-1).

Can mediate $b \rightarrow s\nu\bar{\nu}$ and $b \rightarrow sl^+l^-$ **at tree (or loop) level — depending on the model.**



Expected mass is $\mathcal{O}(\text{TeV})$.

Good LQ candidates mediating $b \rightarrow s\nu\bar{\nu}$ and $b \rightarrow sl^+l^-$ at tree level and accommodating the current experimental data are [Phys. Rev. D 2021, 104]:

- $U_3^\mu(3, 3, 2/3)$ vector \longrightarrow interaction term: $+\bar{q}_L \gamma^\mu \sigma^a Y_{U_3} l_L U_{3\mu}^a$
 $SU(3)_C \times SU(2)_L \times U(1)_Y$

Y_{S_3} : Yukawa matrix in flavour space
 $f^c \equiv C\bar{f}$: charge conjugated
 σ : pauli matrices

With $M_{U_3} = 2 \text{ TeV}/c^2$: enhance $B^+ \rightarrow K^+ \nu\bar{\nu}$ decay rate and explain

$b \rightarrow sl^+l^-$ anomalies within $\pm 1\sigma$ by turning on $Y_{U_3}^{22}, Y_{U_3}^{23}, Y_{U_3}^{32}, Y_{U_3}^{33}$.

Theory of $b \rightarrow s\nu\bar{\nu}$



Models with new mediators

Leptoquarks (LQ)

- Hypothetical **bosons carrying both baryon number and lepton number.**
- Predicted in extensions of the SM unifying quark and leptons: Pati Salam model, $SU(5)$, $SO(10)$ GUTs.
- Can be **scalar** (spin-0) **or vector** (spin-1).

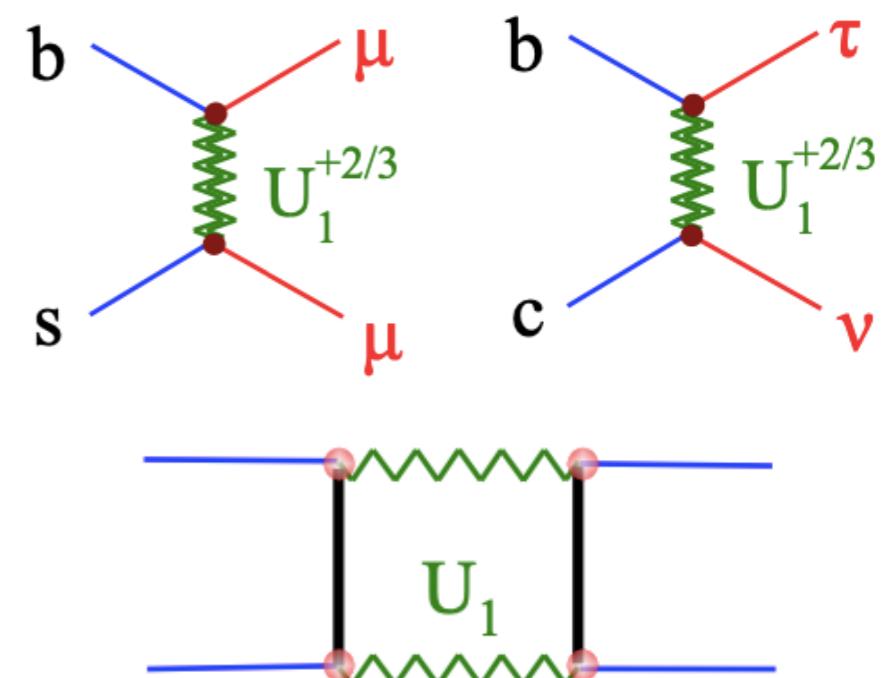
Can mediate $b \rightarrow s\nu\bar{\nu}$ and $b \rightarrow sl^+l^-$ **at tree (or loop) level — depending on the model.**

One of the most favoured candidates is :

- $U_1^\mu(3, 1, 2/3)$ vector:

$$SU(3)_C \times SU(2)_L \times U(1)_Y$$

- can explain $b \rightarrow sl^+l^-$ and $b \rightarrow cl\nu$ anomalies at tree level
[Phys. Rev. D 2021, 104]
- contribution to $b \rightarrow s\nu\bar{\nu}$ at loop level providing up to 20% enhancement
[J. High Energ. Phys. 2021, 50]
- **UV completion** provided by Pati-Salam and 4321 models: massive U_1 from $SU(4)$ breaking.



Theory of $b \rightarrow s\nu\bar{\nu}$



Models with new mediators

Leptoquarks (LQ)

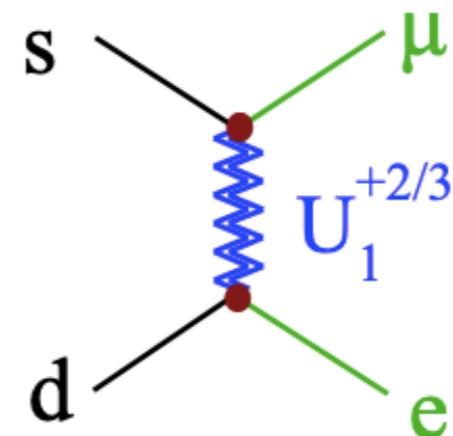
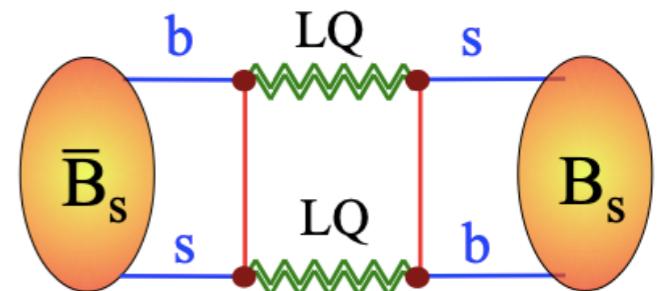
- Hypothetical **bosons carrying both baryon number and lepton number.**
- Predicted in extensions of the SM unifying quark and leptons: Pati Salam model, $SU(5)$, $SO(10)$ GUTs.
- Can be **scalar** (spin-0) **or vector** (spin-1).

Can mediate $b \rightarrow s\nu\bar{\nu}$ and $b \rightarrow sl^+l^-$ **at tree (or loop) level — depending on the model.**

Advantages: would contribute to $B_S^0 - \bar{B}_S^0$ oscillation and $\tau \rightarrow l\nu\bar{\nu}$ only at loop level;

Indirect low-energy constraints from:

- proton stability;
- upper limits on LFV kaon decays, such as $K_L \rightarrow \mu e$.



Theory of $b \rightarrow s\nu\bar{\nu}$



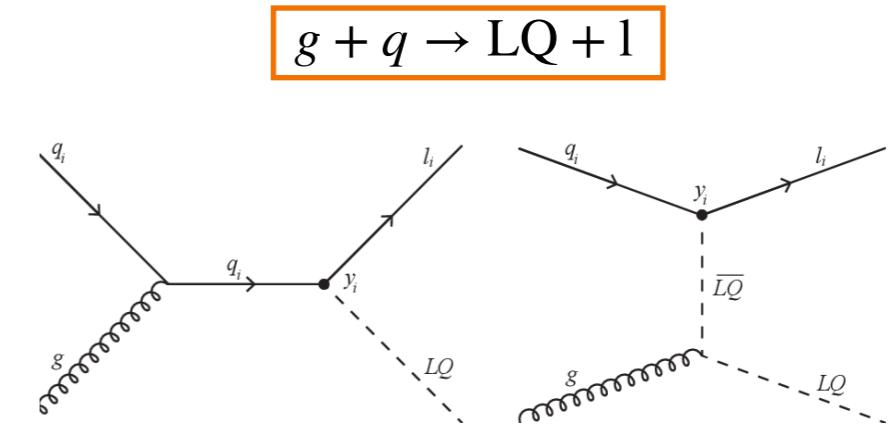
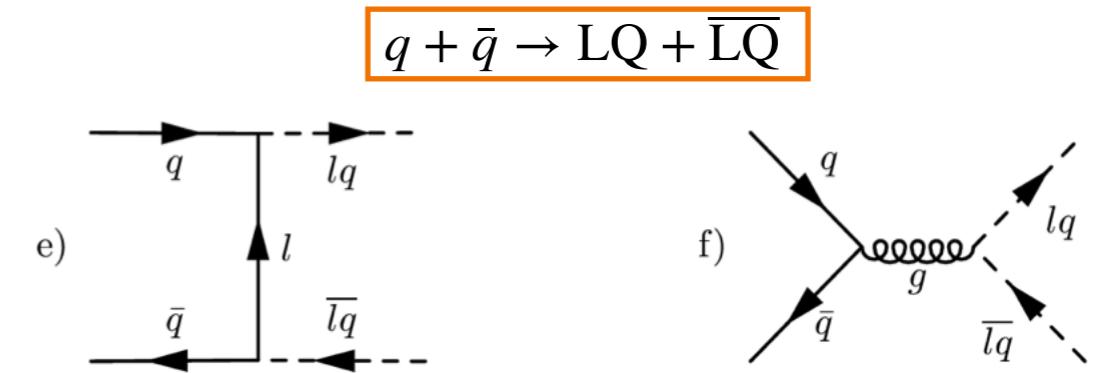
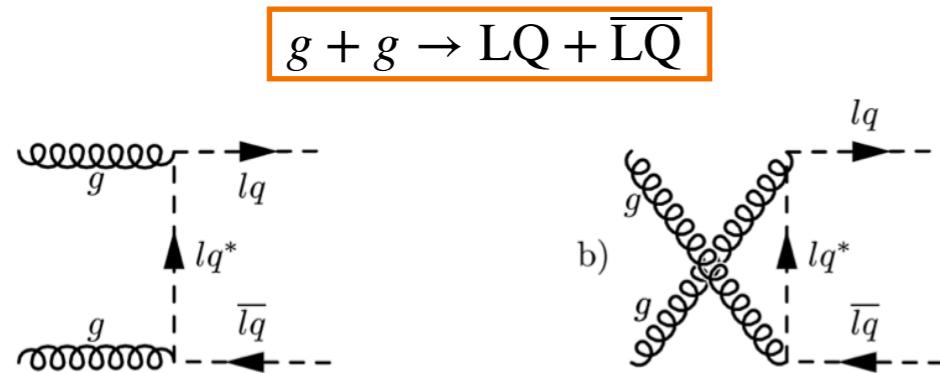
Models with new mediators

Leptoquarks (LQ)

Constraints from LHC:

limits on leptoquarks states through limits on the **pair- and single-production** cross sections, $\sigma \equiv \sigma(M_{LQ})$.

Signatures include high p_T charged leptons, high E_T jets and large missing transverse energy.





Theory of $b \rightarrow s\nu\bar{\nu}$

Models with new mediators

Leptoquarks (LQ)

Constraints from LHC:

- mass > $1270 \text{ GeV}/c^2$ ($1290 \text{ GeV}/c^2$) for **first-generation scalar** by CMS (ATLAS)
[Phys. Rev. D 2019, 5; Eur. Phys. J. C 2019, 79]

- mass > $1285 \text{ GeV}/c^2$ ($1230 \text{ GeV}/c^2$) for **second-generation scalar** by CMS (ATLAS)
[Phys. Rev. D 2019, 3; Eur. Phys. J. C 2019, 79]

- **third-generation scalar:**

LQ $\rightarrow t\tau$, mass > 900 (1430) GeV/c^2 CMS (ATLAS)

LQ $\rightarrow b\tau$, mass > $1020 \text{ GeV}/c^2$ (CMS)

[Eur. Phys. J. C 2018, 78; JHEP 2019, 03; JHEP 2021, 06]

- **vector:** most stringent constraints from CMS search for LQ pair production with LQ $\rightarrow q\nu$
Mass > $1530 \text{ GeV}/c^2$ for vector LQ decaying in $t\nu_\tau$ [Phys. Rev. D 2018, 3]



Theory of $b \rightarrow s\nu\bar{\nu}$

Models with new mediators

Leptoquarks (LQ)

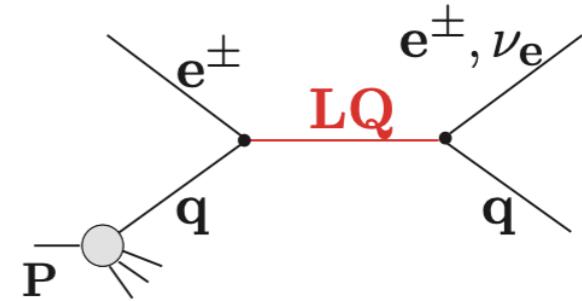
Constraints from HERA ($\sqrt{s} \simeq 320$ GeV): **searches for first generation LQ in ep collisions.**

Production-cross section dependent on **Yukawa coupling** to electron and electron neutrino, λ .

ZEUS and H1 probed:

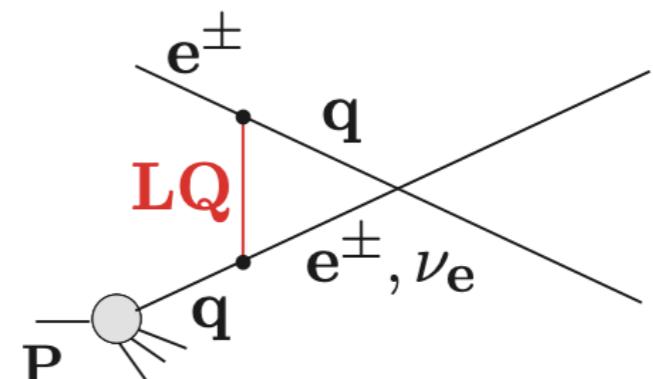
- **direct production via electron-quark fusion**

(possible if $M_{\text{LQ}} < \sqrt{s}$) resulting in peaks in lepton-jets invariant mass distribution;



- **virtual LQ exchange** (still possible if $M_{\text{LQ}} < \sqrt{s}$)

resulting in deviation from the SM prediction in final-state lepton-jets invariant mass spectrum.



For $\lambda = 0.1$, $M_{\text{LQ}} > 290 \text{ GeV}/c^2$ (ZEUS) [Phys. Rev. D 2003, 68]

For $\lambda = 0.3$, $M_{\text{LQ}} > 800 \text{ GeV}/c^2$ (H1) [Phys. Lett. B 2011, 704]

Theory of $b \rightarrow s\nu\bar{\nu}$



Dark matter as a source of missing energy

$B \rightarrow K^{(*)}\nu\bar{\nu}$ have same experimental signature of generic $B \rightarrow K^{(*)}E$ processes: $K^{(*)} + \text{missing energy}$.

Dark matter (DM) is a well motivated BSM source of missing energy and momentum.



- Barely interacts with baryonic matter and radiation, except through gravity.
- $\sim 85\%$ of the matter in the Universe from astrophysical observations.
- **Various evidences:** e.g. high **revolution speeds** of stars at large radii in galaxies; **velocities** of galaxies in clusters; **X-ray emission** in galaxy clusters; **weak gravitational lensing** of background galaxies by clusters (DM halo).

Dark matter scenario explorable at Belle II: **resonant production of dark scalar** in $B \rightarrow K^{(*)}E$ decays.



- $B \rightarrow K^{(*)}S(\rightarrow \chi\bar{\chi})$.
- mass in the GeV range.
- decay into **pair of dark fermions**.



Theory of $b \rightarrow s\nu\bar{\nu}$

Dark matter as a source of missing energy

Dark scalars in $B \rightarrow K^{(*)}\ell\ell$ decays [Phys. Rev. D 2020, 101]

Extend SM with a real scalar field ϕ and a Dirac fermion χ , singlets under $SU(3)_c$ and $SU(2)_L$.

- ϕ : mediates a new interaction between Higgs field and χ
- χ : stable dark-matter candidate not mixing with SM neutrinos

After **EWSB**: dark scalar S + observed Higgs.

S can **decay into** pair of **dark fermions** ($\chi\bar{\chi}$) **or SM leptons** ($l\bar{l}$): $\Gamma_S = s_\theta^2 \Gamma_{\text{SM}} + c_\theta^2 \Gamma_{\chi\bar{\chi}}$

\uparrow \uparrow
sin and cos of mixing angle θ

$$\text{BR}(S \rightarrow l\bar{l}) = \frac{s_\theta^2 \Gamma_{l\bar{l}}}{\Gamma_S} = \frac{m_l^2 s_\theta^2}{8\pi v^2} \frac{m_S}{\Gamma_S} \left(1 - \frac{4m_l^2}{m_S^2}\right)^{3/2}$$

$v \equiv$ Higgs vev

$$\text{BR}(S \rightarrow \chi\bar{\chi}) = \frac{c_\theta^2 \Gamma_{\chi\bar{\chi}}}{\Gamma_S} = \frac{y_\chi^2 c_\theta^2}{8\pi} \frac{m_S}{\Gamma_S} \left(1 - \frac{4m_\chi^2}{m_S^2}\right)^{3/2}$$

y_χ : Yukawa couplings of the Higgs with dark fermions χ

Visible decays dominate for $m_S < 2m_\chi$

Invisible decays dominate for $m_S > 2m_\chi$ and $y_\chi c_\theta > m_l s_\theta / v$

Fulfilled even for weak y_χ couplings since $m_l/v \ll 1$

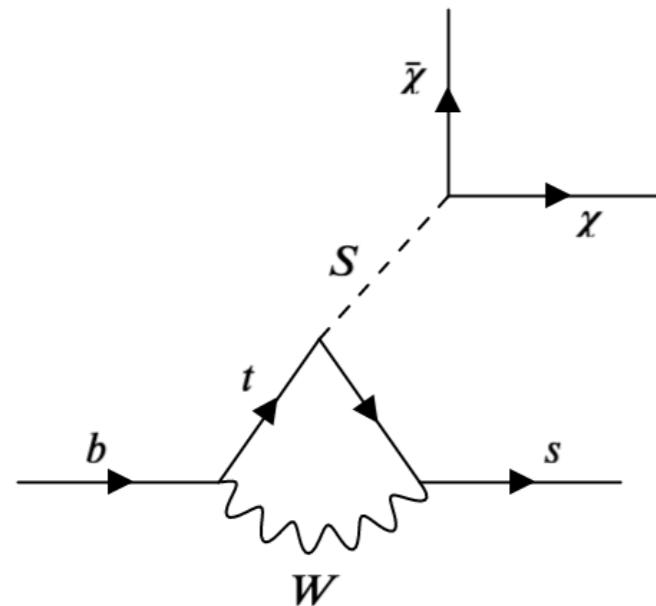
Theory of $b \rightarrow s\nu\bar{\nu}$



Dark matter as a source of missing energy

Dark scalars in $B \rightarrow K^{(*)}\chi\bar{\chi}$ decays [Phys. Rev. D 2020, 101]

The production of dark scalars in $B \rightarrow K^{(*)}S$ decays is **loop-induced**



[J. High Energ. Phys. 2012, 03; 2019, 11]

$$\text{BR}(B^+ \rightarrow K^+ S) = \frac{\sqrt{2} G_F |C_{bs}|^2}{64\pi \Gamma_B m_B^3} \frac{(m_b + m_s)^2}{(m_b - m_s)^2} f_0^2(m_S^2) \times (m_B^2 - m_K^2) [(m_B^2 - m_K^2 - m_S^2)^2 - 4m_K^2 m_S^2]^{1/2}$$

Γ_B : total B decay width

$f_0(m_S^2)$: FF computed at $q^2 = m_S^2$

C_{bs} : Wilson Coefficient of the effective interaction $\frac{C_{bq}}{v} (m_b \bar{q}_L b_R + m_q \bar{q}_R b_L) S$

$$S \rightarrow \chi\bar{\chi} \text{ through } \text{narrow resonance} \rightarrow \text{BR}(B^+ \rightarrow K^+ \chi\bar{\chi}) = \text{BR}(B^+ \rightarrow K^+ S) \text{BR}(S \rightarrow \chi\bar{\chi}) \propto s_\theta^2 \frac{c_\theta^2 \Gamma_{\chi\bar{\chi}}}{\Gamma_S}$$

Experimental challenges

- Measurements of $B \rightarrow K^{(*)}\nu\bar{\nu}$ not easy to recast: 3-body vs 2-body decay.
- Need to optimise sensitivity to $B \rightarrow K^{(*)}S(\rightarrow \chi\bar{\chi})$ in dedicated search.

Tagging methods



Algorithms

Semi-Exclusive B reconstruction (SER) at BaBar [Eur. Phys. J. C 2014, 74]

1) Reconstruct a $D^{(*)}$ meson

D^0 decays	D^+ decays	D^{*+} decays	D^{*0} decays	D_s^+ decays
$K^-\pi^+$	$K^-\pi^+\pi^-$	$\bar{D}^0\pi^+$	$D^0\pi^0$	$\phi\pi^0$
$K^-\pi^+\pi^0$	$K^-\pi^+\pi^-\pi^0$	$D^+\pi^0$	$D^0\gamma$	$K_s^0K^+$
$K^-3\pi$	$K_s^0\pi^+$			
K^-K^+	$K_s^0\pi^+\pi^0$			
$K_s^0\pi^0$	$K_s^0\pi^+\pi^+\pi^-$			
$K_s^0\pi^+\pi^-$	$K^+K^-\pi^+$			
$K_s^0\pi^+\pi^-\pi^0$	$K^+K^-\pi^+\pi^0$			
$\pi^-\pi^+$				
$\pi^-\pi^+\pi^0$				

2) Combine the $D^{(*)}$ meson to reconstruct a B meson

- **hadronic tagging:** combination with charmless hadrons

$$B_{\text{tag}} \rightarrow D_{\text{seed}} Y^\pm \text{ with } Y^\pm = n_1\pi^\pm + n_2K^\pm + n_3\pi^0 + n_4K_S^0 \quad n_1 \in [1,5], n_2 \in [0,2] \\ n_3 \in [0,2], n_4 \in [0,1]$$

- **semileptonic tagging:** combination with a charged lepton

$$B_{\text{tag}} \rightarrow D_{\text{seed}} Y^\pm \text{ with } Y^\pm = e^\pm, \mu^\pm$$

3) Well reconstructed B_{tag} candidates selected using the variables $\Delta E = E_B^* - \sqrt{s}/2$ and $M_{bc} = \sqrt{s/4 - p_B^{*2}}$

Tagging methods



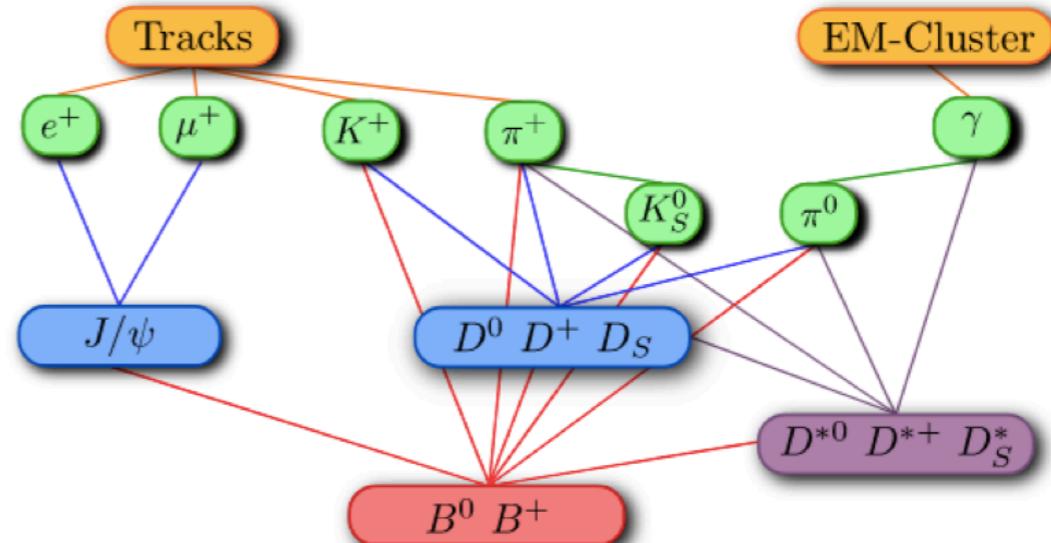
Algorithms

Full Reconstruction (FR) at Belle [Nucl. Instrum. Methods Phys. Res. A 2011, 654]

Hierarchical algorithm starting from reconstructed tracks and clusters.

Full-reconstruction in **4 stages**

stage	particles
1	tracks, K_S , γ , π^0
2	$D_{(s)}^\pm$, D^0 , and J/ψ mesons
3	$D_{(s)}^{*\pm}$ and D^{*0} mesons
4	B^\pm and B^0 mesons



At each stage **neural-networks trained on MC samples** simulating $\Upsilon(4S)$ events are used **to classify particle candidates**.

In the final stage B_{tag} **candidates** are reconstructed and **classified based on a list of decay modes with sizeable branching fractions**.

B^+		B^0	
mode	BR	mode	BR
$B^+ \rightarrow \bar{D}^0 \pi^+$	0.484%	$B^0 \rightarrow D^- \pi^+$	0.268%
$B^+ \rightarrow \bar{D}^0 \pi^+ \pi^0$	1.340%	$B^0 \rightarrow D^- \pi^+ \pi^0$	0.760%
$B^+ \rightarrow \bar{D}^0 \pi^+ \pi^+ \pi^-$	1.100%	$B^0 \rightarrow D^- \pi^+ \pi^+ \pi^-$	0.800%
$B^+ \rightarrow D_s^+ \bar{D}^0$	1.000%	$B^0 \rightarrow \bar{D}^0 \pi^0$	0.026%
$B^+ \rightarrow \bar{D}^{0*} \pi^+$	0.519%	$B^0 \rightarrow D_s^+ D^-$	0.720%
$B^+ \rightarrow \bar{D}^{0*} \pi^+ \pi^0$	0.980%	$B^0 \rightarrow D^{*-} \pi^+$	0.276%
$B^+ \rightarrow \bar{D}^{0*} \pi^+ \pi^+ \pi^-$	1.030%	$B^0 \rightarrow D^{*-} \pi^+ \pi^0$	1.500%
$B^+ \rightarrow \bar{D}^{0*} \pi^+ \pi^+ \pi^- \pi^0$	1.800%	$B^0 \rightarrow D^{*-} \pi^+ \pi^+ \pi^-$	0.700%
$B^+ \rightarrow D_s^{+*} \bar{D}^0$	0.760%	$B^0 \rightarrow D^{*-} \pi^+ \pi^+ \pi^- \pi^0$	1.760%
$B^+ \rightarrow D_s^{+*} \bar{D}^{0*}$	0.820%	$B^0 \rightarrow D_s^{+*} D^-$	0.740%
$B^+ \rightarrow D_s^{+*} \bar{D}^{0*}$	1.710%	$B^0 \rightarrow D_s^{+*} D^{*-}$	0.800%
$B^+ \rightarrow \bar{D}^0 K^+$	0.037%	$B^0 \rightarrow D_s^{+*} D^{*-}$	1.770%
$B^+ \rightarrow D^- \pi^+ \pi^+$	0.107%	$B^0 \rightarrow J/\psi K_S^0$	0.087%
$B^+ \rightarrow J/\psi K^+$	0.101%	$B^0 \rightarrow J/\psi K^+ \pi^-$	0.120%
$B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$	0.107%	$B^0 \rightarrow J/\psi K_S^0 \pi^+ \pi^-$	0.100%
$B^+ \rightarrow J/\psi K^+ \pi^0$	0.047%		
$B^+ \rightarrow J/\psi K_S^0 \pi^+$	0.094%		

Tagging methods



Algorithms

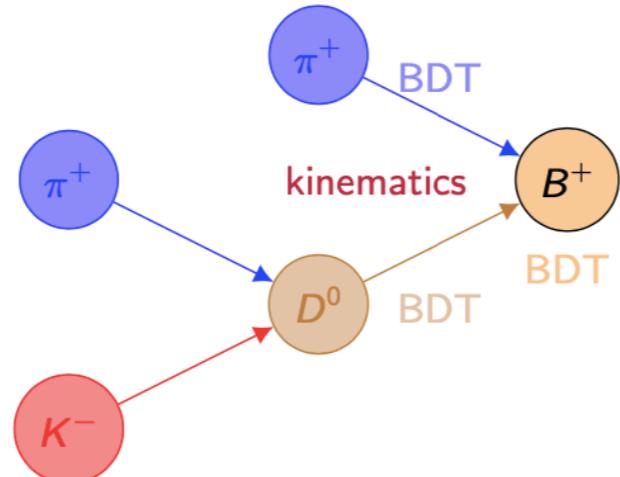
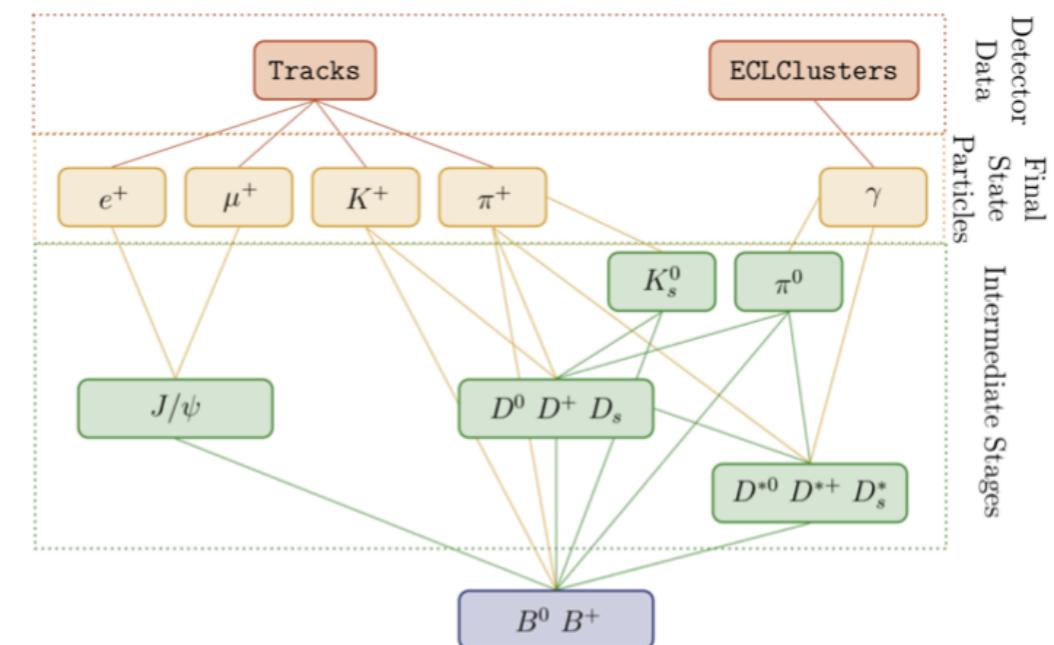
Full Event Interpretation (FEI) at Belle II [Nucl. Instrum. Methods Phys. Res. A 2011, 654]

Improves the FR **hierarchical approach** to tagging:

- At each stage one BDT to classify particle candidates;
- $\mathcal{O}(200)$ decay channels tagged;
- $\mathcal{O}(10k)$ decay chains reconstructed in 6 consecutive stages.

The algorithm

- Particle candidates from tracks and clusters.
- Stable particles combined into intermediate particles;
- **Intermediate and stable particles** combined to form a **B candidate**;

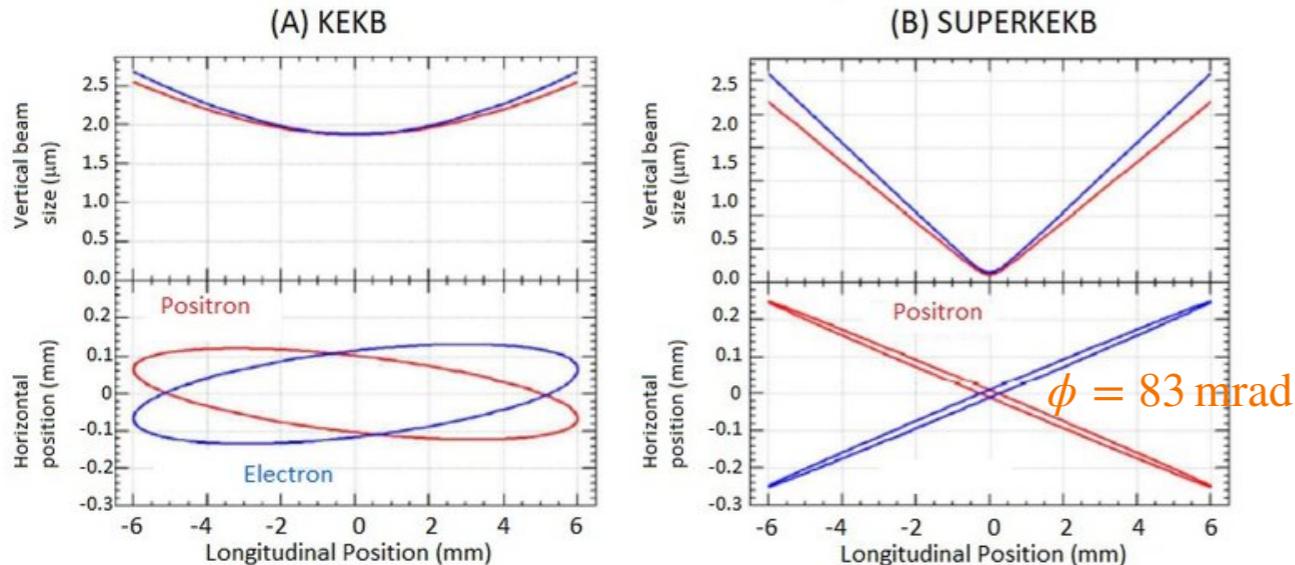


The Belle II experiment



SuperKEKB

Nanobeam: reduce beam sizes $\sigma_{x,y}^*$ at the IP and maximise the crossing angle ϕ .



$$\text{Large crossing angle } \phi \rightarrow \beta_y^* \geq \frac{\sigma_x^*}{\phi} \sim 300 \mu\text{m}$$

$I_{e^\pm} \equiv \text{beam currents}$
 $\xi_y^{e^\pm} \equiv \text{beam-beam parameter}$
 $R \equiv \text{geometric parameters}$

$$L = \frac{\gamma_{e^\pm}}{2er_e} \left(1 + \frac{\sigma_y^*}{\sigma_x^*} \right) \left(\frac{I_{e^\pm} \xi_y^{e^\pm}}{\beta_y^*} \right) \left(\frac{R_L}{R_{\xi_y}} \right)$$

$\beta_{x,y}^*$ related to the beam sizes $\sigma_{x,y}^*$ at the IP:

$$\sigma_{x,y}^* = \sqrt{\epsilon \beta_{x,y}^*} \quad \text{where } \epsilon \equiv \text{beam emittance}$$

		KEKB		SuperKEKB		units
		LER	HER	LER	HER	
Beam energy	E_b	3.5	8	4	7.007	GeV
Beam crossing angle	ϕ	22		83		mrad
β function @ IP	β_x^*/β_y	1200/5.9		32/0.27	25/0.30	mm
Beam current	I_b	1.64	1.19	3.6	2.6	A
Luminosity	L	2.1×10^{34}		8×10^{35}		$\text{cm}^{-2}\text{s}^{-1}$

X 20

X 2

X 40

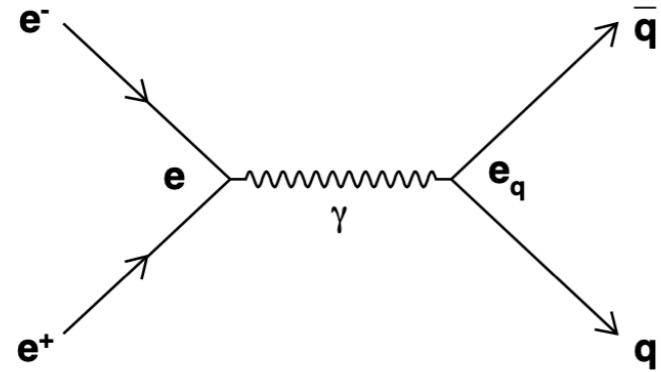
The Belle II experiment



Production cross sections

At $\sqrt{s} = 10.58 \text{ GeV}$ **photon exchange dominant.**

Same tree-level diagram for $e^+e^- \rightarrow \mu^+\mu^-, \tau^+\tau^-q\bar{q}$



$$\frac{d\sigma_{\mu\mu}}{d\Omega} = \frac{\alpha^2}{4s}(1 + \cos^2\theta), \quad \frac{d\sigma_{q\bar{q}}}{d\Omega} = \frac{\alpha^2}{4s} \cdot N_C \cdot Q_i^2(1 + \cos^2\theta)$$

$$\frac{d\sigma_{\tau\tau}}{d\Omega} = \frac{\alpha^2}{4s} \cdot \beta \cdot (2 - \beta^2 \sin^2\theta)$$

$N_C \equiv$ color factor

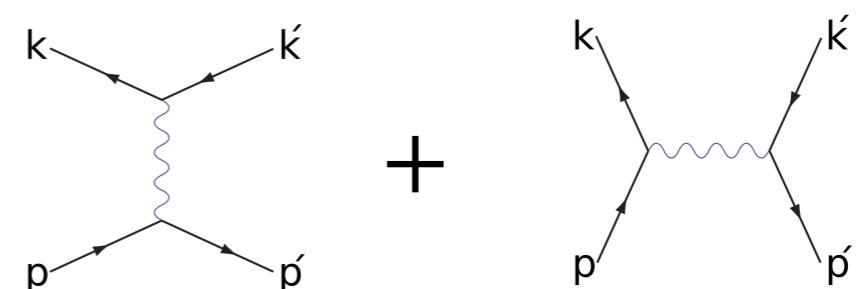
$Q_i = +2/3, -1/3$

$\beta = v/c$

Bhabha ($e^+e^- \rightarrow e^+e^-$):

Main processes at $\sqrt{s} = 10.58 \text{ GeV}$

Physics process	Cross section [nb]
$\Upsilon(4S)$	1.110 ± 0.008
$u\bar{u}(\gamma)$	1.61
$d\bar{d}(\gamma)$	0.40
$s\bar{s}(\gamma)$	0.38
$c\bar{c}(\gamma)$	1.30
$e^+e^-(\gamma)$	$300 \pm 3 \text{ (MC stat.)}$
$\gamma\gamma(\gamma)$	$4.99 \pm 0.005 \text{ (MC stat.)}$
$\mu^+\mu^-(\gamma)$	1.148
$\tau^+\tau^-(\gamma)$	0.919
$\nu\bar{\nu}(\gamma)$	0.25×10^{-3}
$e^+e^-e^+e^-$	$39.7 \pm 0.1 \text{ (MC stat.)}$
$e^+e^-\mu^+\mu^-$	$18.9 \pm 0.1 \text{ (MC stat.)}$



$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{2s} \left[\frac{t^2}{s^2} + \frac{s^2}{t^2} + u^2 \left(\frac{1}{s} + \frac{1}{t} \right)^2 \right]$$

$$s = (k+p)^2 = (k'+p')^2 \approx 2k \cdot p \approx 2k' \cdot p'$$

$$t = (k-k')^2 = (p-p')^2 \approx -2k \cdot k' \approx -2p \cdot p'$$

$$u = (k-p')^2 = (p-k')^2 \approx -2k \cdot p' \approx -2k' \cdot p$$

The Belle II experiment

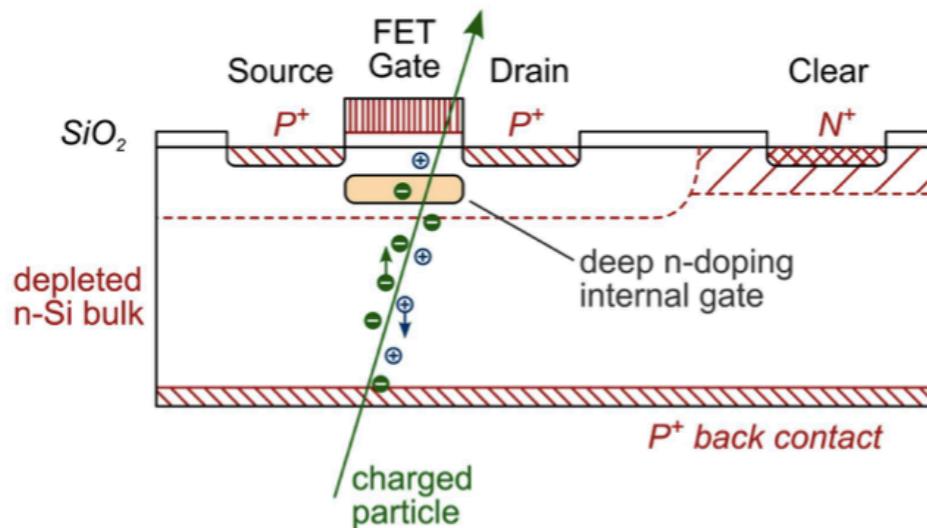


The Pixel Detector (PXD)

Innermost component of vertex detector

- 40 Si pixel sensors in 2 cylindrical layers ($r_1 = 1.4\text{ cm}$, $r_2 = 2.2\text{ cm}$) around the beam pipe.
- Angular acceptance: $17^\circ < \theta < 150^\circ$

DEPFET pixels: depleted p-channel field effect transistors

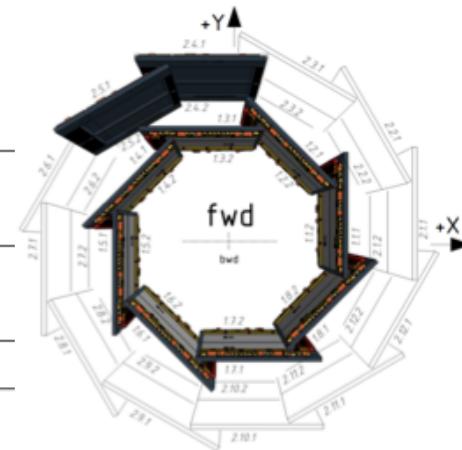


FET on top of **depleted Si** bulk: **gate** voltage controls current between **source** and **drain** terminals.

Internal gate (n-implant below FET gate) **collects electrons created by ionisation altering FET current.**

The **signal** is measured and indicates which pixel was hit.

Layer	Radius (mm)	Ladders	Sensors	Pixels/Sensor $u \times v$
1	14	8	16	250×768
2	22	12	24	250×768
Sum		20	40	7 680 000



PXD roles

Fundamental for **B** (and D) vertexing

PXD hits used in **offline track fit**: improved impact parameter resolution and bkg rejection.

Performance:

- Hit efficiency $> 98\%$;
- Average spatial resolution $\sim 14\mu\text{m}$.

[Nucl. Instrum. Methods Phys. Res. A 2021, 987]

DEPFET advantages:

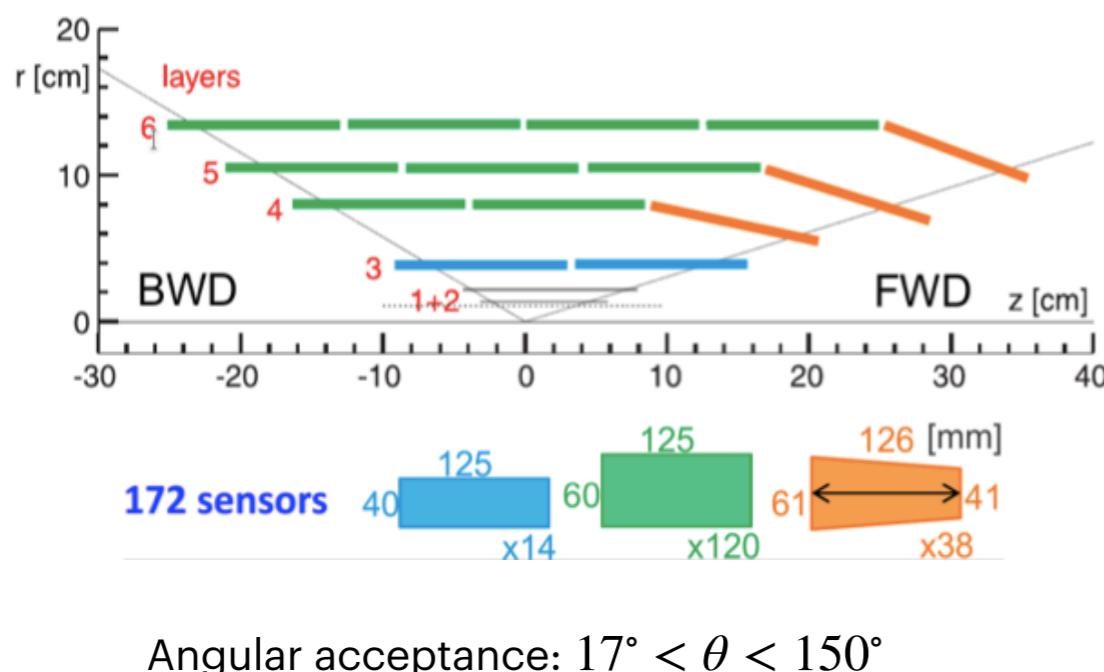
- fast charge collection $\mathcal{O}(\text{ns})$;
- internal amplification;
- extremely thin sensors: $75\mu\text{m}$ only!

The Belle II experiment



The Silicon Vertex Detector (SVD)

- **4 layers of double-sided silicon-strip sensors:** 2D spatial info.
- Individual readout on each silicon sensor: fast ~ 50 ns.
- Trapezoidal shape for the outer 3 layers (optimise track incident angle).



Layer	Radius (mm)	Ladders	Sensors	Strips/Sensor
				u, v
3	39	7	14	768, 768
4	80	10	30	768, 512
5	104	12	48	768, 512
6	135	16	80	768, 512
Sum		35	172	132 096, 91 648

SVD roles

- **Vertexing:** extrapolate tracks to the PXD;
- **Standalone tracking** for low p_T tracks;
- **PID** via dE/dx measurements.
- Precise **vertexing of K_S^0**

Performance:

- average hit efficiency $> 99.5\%$;
- spatial resolution from $18 \mu\text{m}$ to $35 \mu\text{m}$.

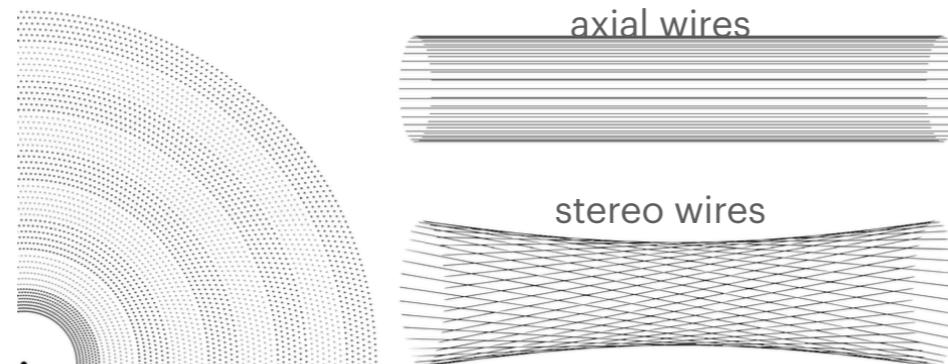
[JPS Conf. Proc. 2021, 34]



The Belle II experiment

The Central Drift Chamber (CDC)

- He – C₂H₆ 50:50 gas mixture.
- ~ 50 000 sense and field wires.
- Sense wires arranged in layers and adjacent layers make a *superlayer*.
- Superlayers alternate axial (A) and stereo (U,V) orientations to reconstruct full 3D trajectory.



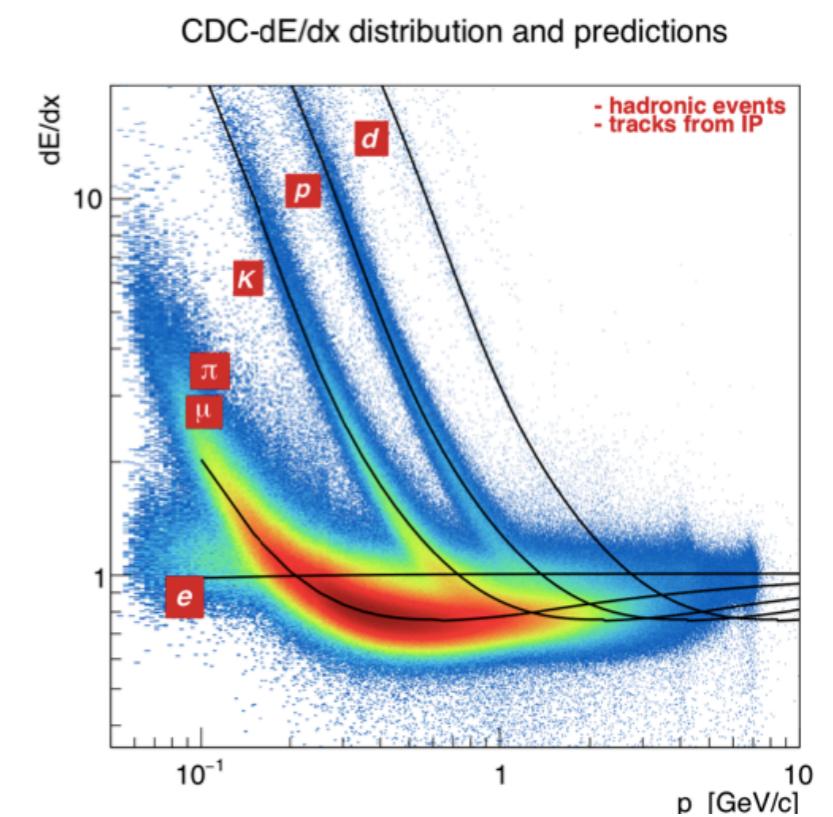
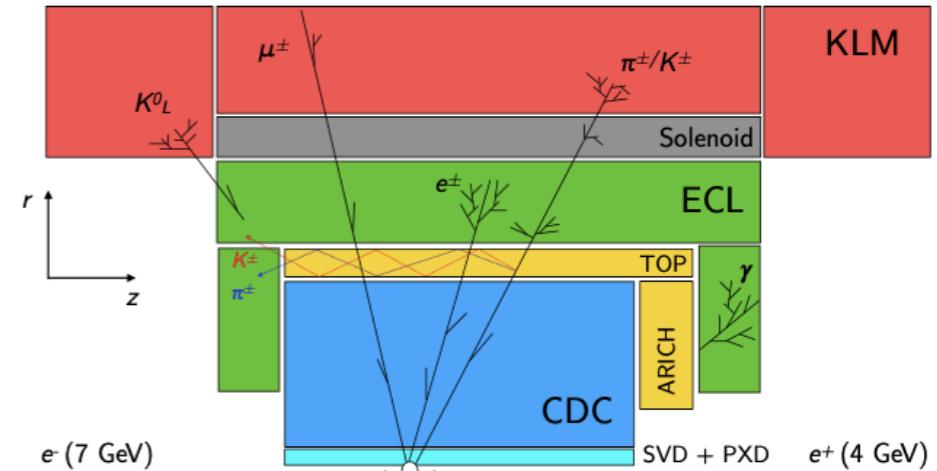
Angular acceptance: $17^\circ < \theta < 150^\circ$

CDC roles:

- **Tracking**
 - reconstruct 3D helix track of charged particles;
 - momentum measurement;
- **PID via energy-loss (dE/dx) measurement;**
- **trigger signals** for charged particles.

Performance: • p_T resolution $\approx 0.4\%$;

• spatial resolution $\approx 120 \mu\text{m}$ [Comput. Phys. Commun. 2021, 259]



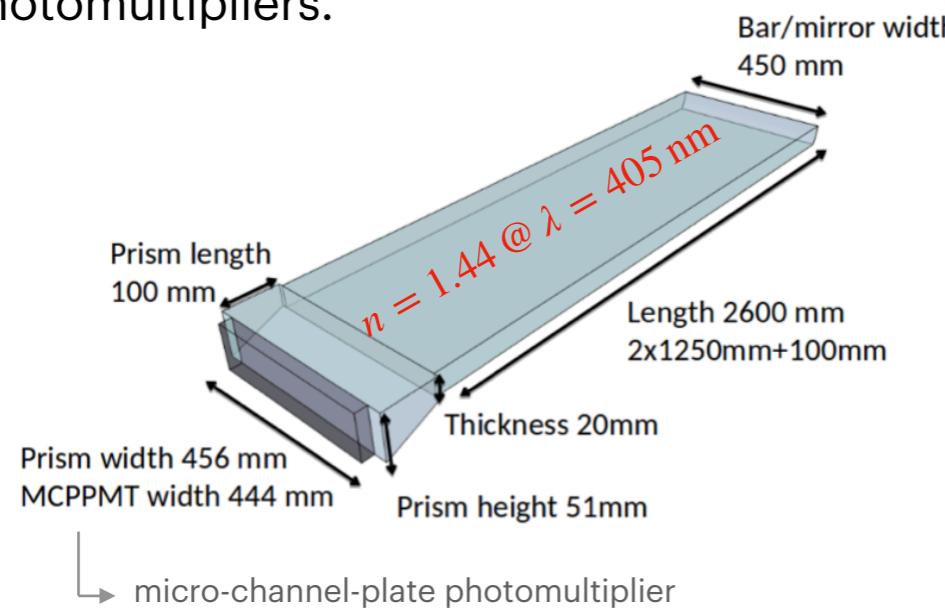
The Belle II experiment



The Time-Of-Propagation Counter (TOP)

Barrel PID system:

16 modules around the CDC, quartz bars coupled with photomultipliers.

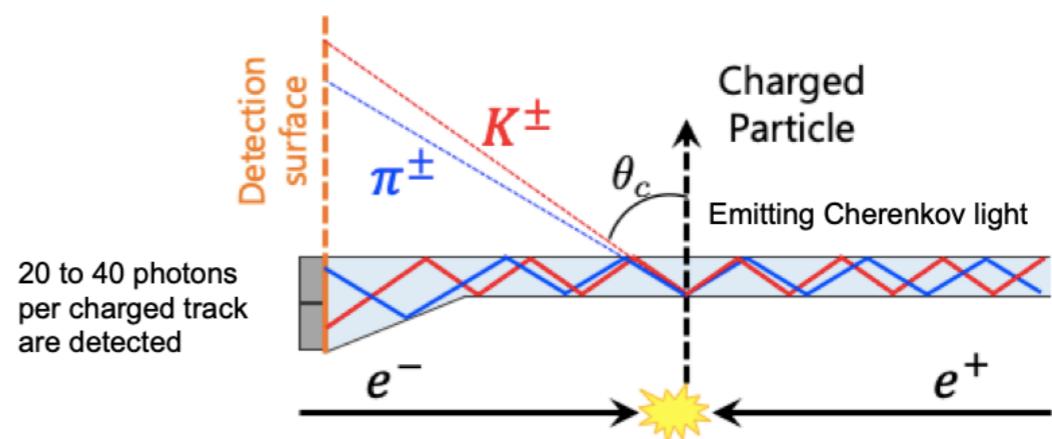
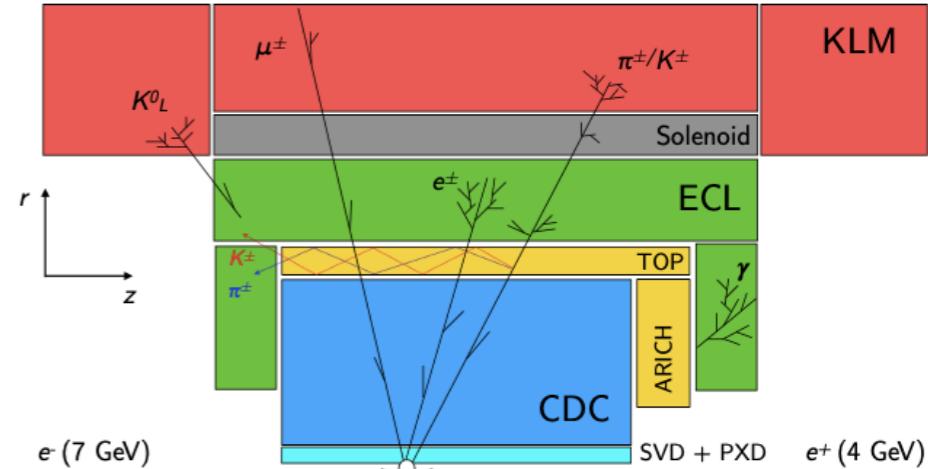


TOP roles

- PID info, mostly K^\pm/π^\pm separation.

Exploit arrival time of Cherenkov photons:

compare hit-time distribution with expected distribution for e, μ, π, K, p, d hypotheses.



$$\cos \theta_c = \frac{1}{n\beta} = \frac{\sqrt{m^2 + p^2}}{np}$$

$$p = 3 \text{ GeV } \pi/K \\ \Delta\theta_c \simeq 0.6^\circ \Rightarrow \Delta t = \mathcal{O}(100) \text{ ps}$$

Performance:

- 85% kaon ID efficiency at 10% pion mis-ID rate.
[Proc. Sci. 2022, 398]

The Belle II experiment



The Aerogel Ring-Imaging Cherenkov Counter (ARICH)

Forward end-cap PID system:

2 layers of aerogel radiator combined with a system of photo detectors.

ARICH roles

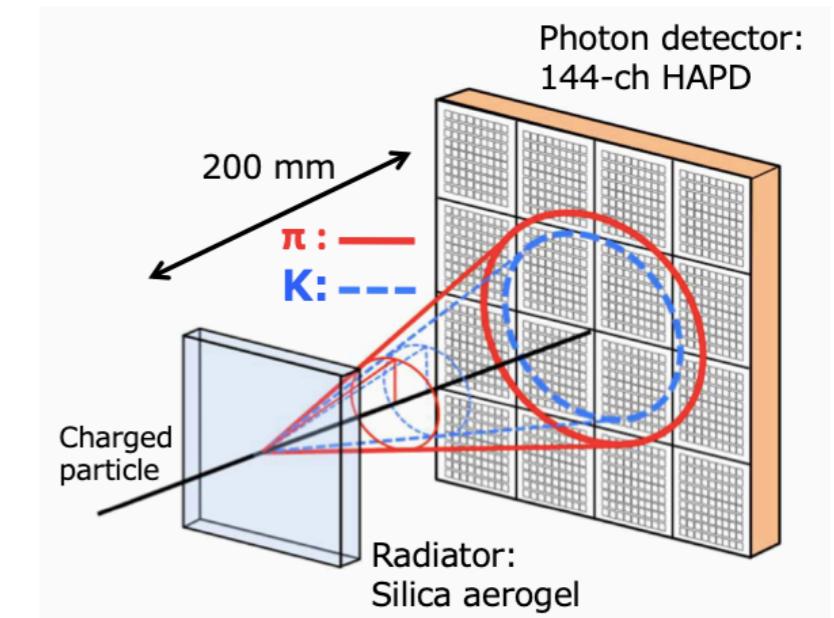
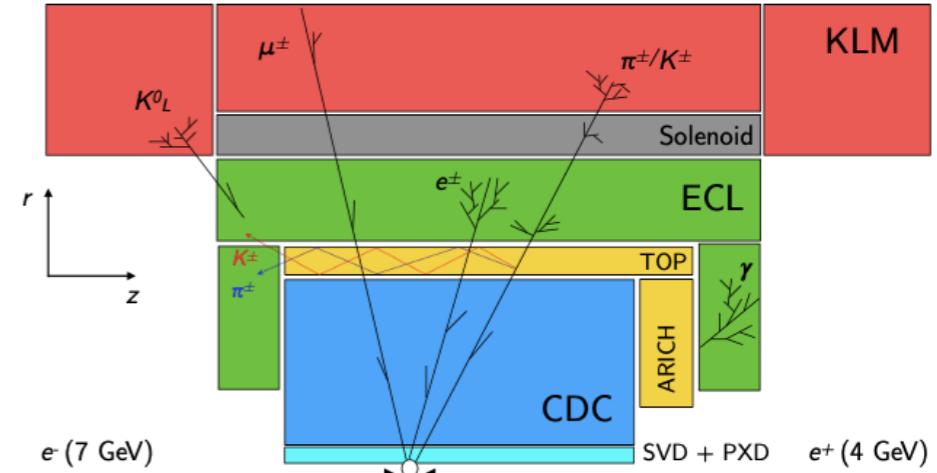
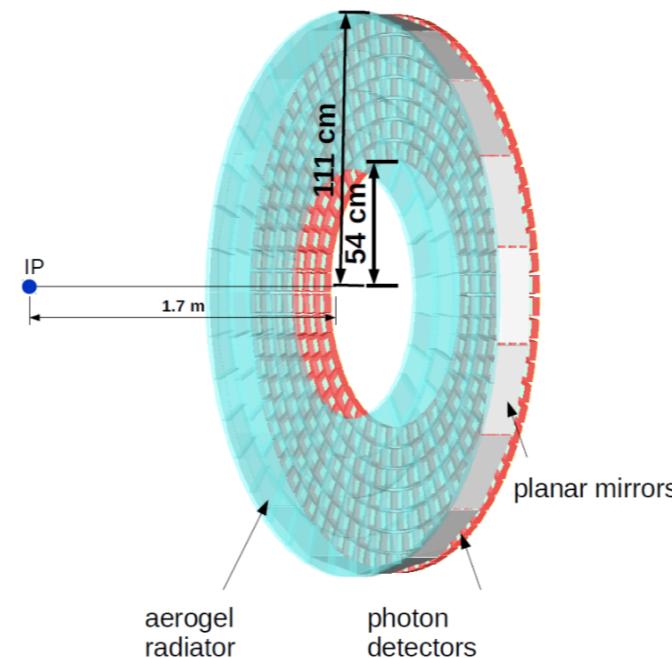
- K^\pm/π^\pm separation ($\pi/\mu, e$ below $\sim 1 \text{ GeV}/c$)

Exploit emission angle of Cherenkov photons:
compare photon hit-map with expected
distributions for different particle hypotheses.

Performance:

- up to 93% kaon ID efficiency
at 10% pion mis-ID rate.

[Nucl. Instrum. Meth. A 2014, 766]



The Belle II experiment

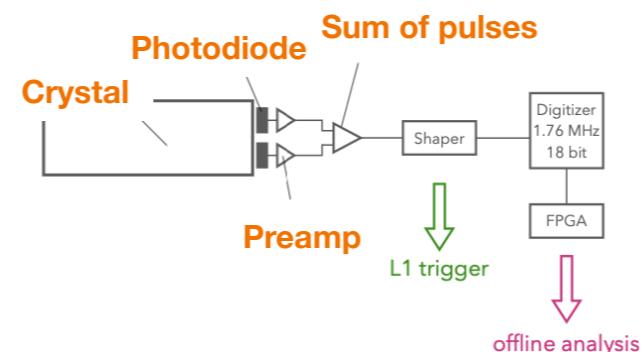


The Electromagnetic Calorimeter (ECL)

3 m long barrel section ($r_i = 1.25$ m) + fwd and bwd endcaps

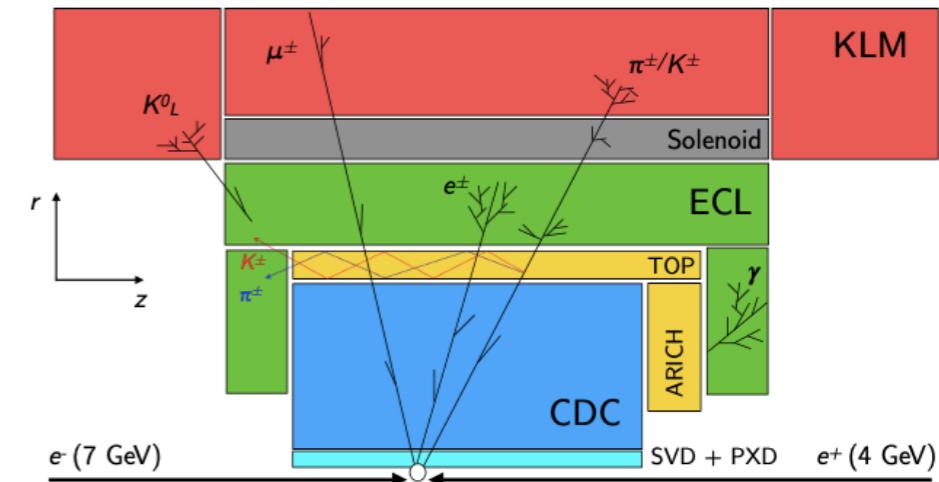
- 8736 crystals
- $\sim 5 \times 5 \times 30 \text{ cm}^3 \text{ CsI(Tl)} \rightarrow 16.1 X_0$
- Crystals, photodiodes and preamplifiers reused from Belle experiment
- New shapers and digitizers/FPGAs

Light readout:



ECL roles

- Photon (γ) detection (e.g. via lateral shower shape).
- Determination of γ energy and angular coordinates.
- Info for K_L^0 ID (e.g. cluster shape, distance from closest track, energy deposition).
- Lepton ID: electron/hadron separation via $E_{\text{cluster}}/p_{\text{track}}$ and identification of low-momentum muons (out of KLM).
- Generation of trigger signals.



Angular acceptance: $12.4^\circ < \theta < 155.1^\circ$

Performance:

- photon-energy resolution (σ_E/E) from $\sim 2.5\%$ at 100 MeV to $\sim 1.7\%$ at 5 GeV.

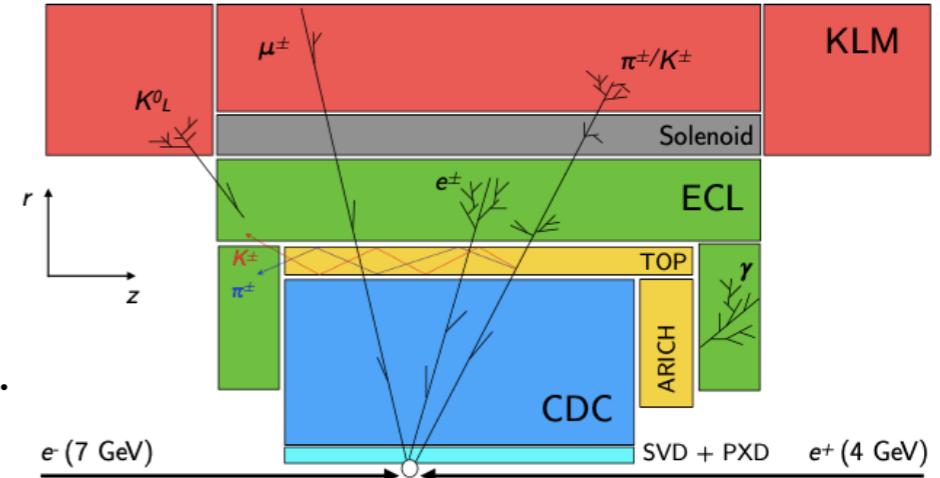
[J. Phys. Conf. Ser. 2015, 587]



The Belle II experiment

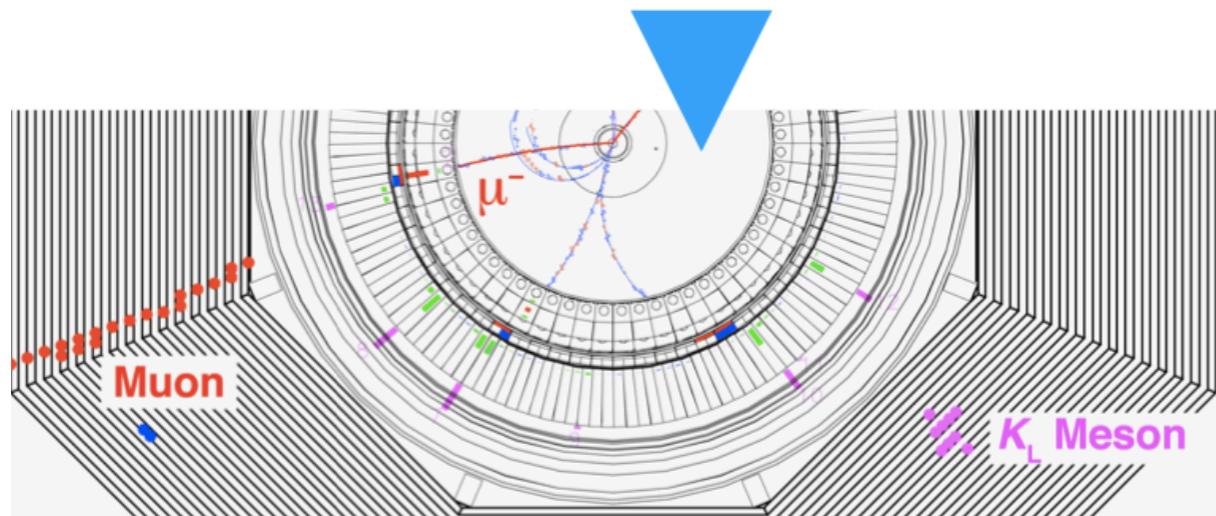
K_L^0 and μ detector (KLM)

- Located outside of the solenoid.
- Made of **iron plates** (for hadronic shower, $\sim 3.9\lambda_I$)
+ **detector modules**:
 - Pairs of Resistive Plate Chambers (outer barrel);
 - Scintillator strips with Si photomultiplier (inner barrel, end caps).
- Angular acceptance: $20^\circ < \theta < 155^\circ$.



KLM roles

- K_L^0 ID info: distance between neutral cluster and closest track, cluster timing, n. of hits in innermost layers.
- μ ID info: hits aligned with CDC track, penetration depth (separation from charged hadrons).



Performance:

- muon ID efficiency up to 89 % at $p > 1 \text{ GeV}/c$, hadron fake rate of $\sim 1.3 \%$;
- K_L^0 ID efficiency increases linearly with p up to 80 % (plateau) at $p \sim 3 \text{ GeV}/c$.

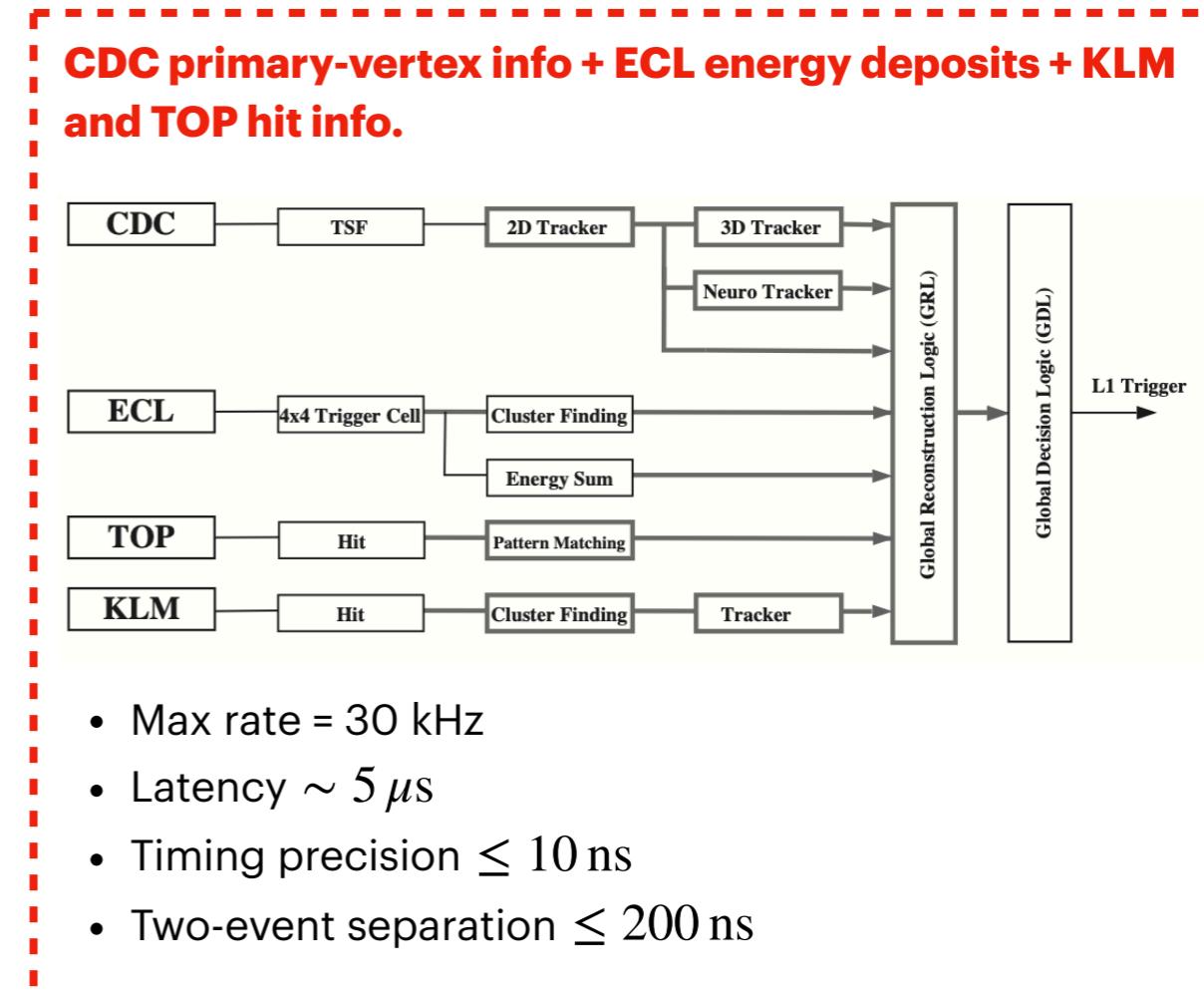
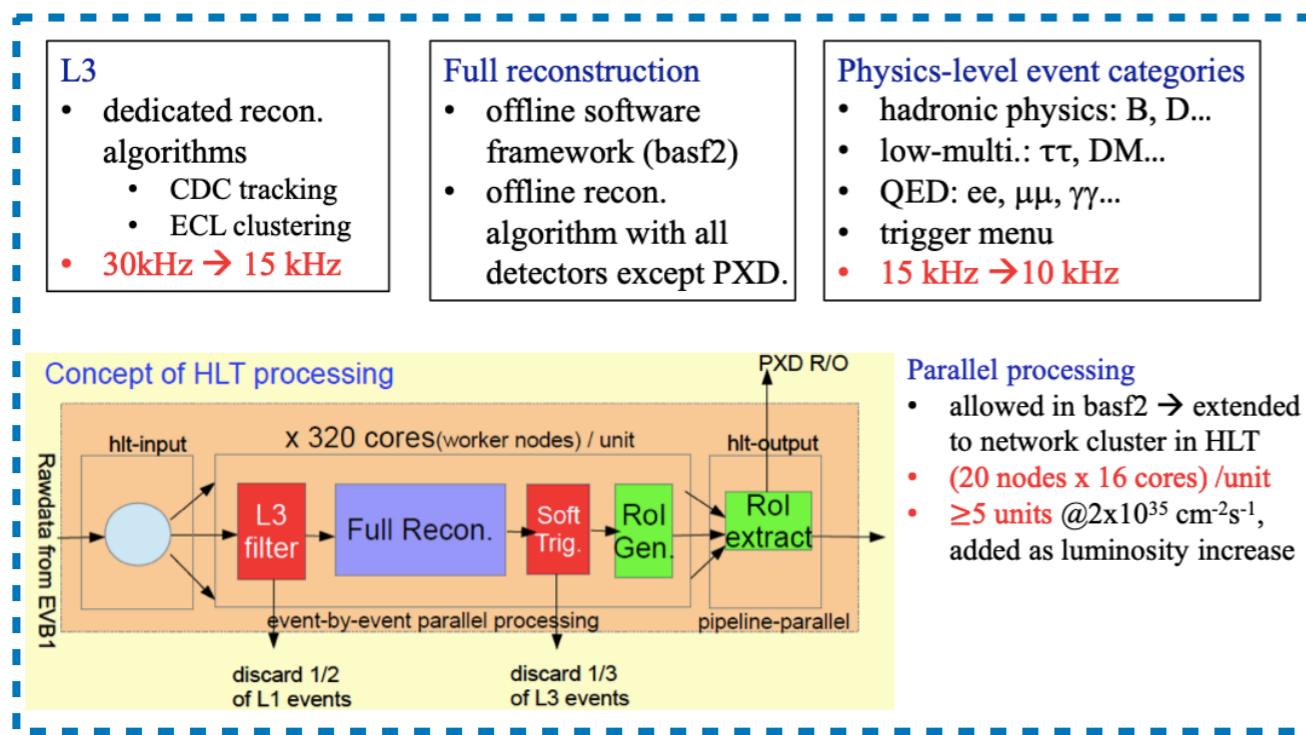
[Belle II TDR, 2010]

The Belle II experiment



Trigger system

- >99.9% efficiency for B and D physics.
- High efficiency for low-multiplicity events (e.g., leptonic τ decays, dark-sector searches...).
- Hardware + software trigger:
 - **Level 1 (L1)**, hardware based
 - **High Level trigger (HLT)**, software based.



The Belle II experiment



Charged particle identification

- **Charged hadrons:** TOP and ARICH info, dE/dx info from SVD and CDC.
- **Charged leptons:** primary info from ECL for e^\pm , from KLM for μ^\pm .

→ **PID variables**

PID variables are **likelihood** based:

detector likelihood \mathcal{L}_i^{det} for hypotheses $i = e^+, \mu^+, \pi^+, K^+, p^+, d^+$ → global likelihood $\mathcal{L}_i = \prod_{det} \mathcal{L}_i^{det}$

PID variables as **likelihood ratios**:

- global ratios $PID_i = \frac{\mathcal{L}_i}{\sum_{j=e,\mu,\pi,K,p,d} \mathcal{L}_j}$
- binary ratios $PID(i|j) = \frac{\mathcal{L}_i}{\mathcal{L}_i + L_j}$

PID selections characterised by:

- efficiency = (number of i particles identified as i) / (number of i particles)
- misID rate = (number of i particles identified as j) / (number of i particles)



data/MC correction weights:
efficiency ratios in $p, \cos\theta$
from control channel

The Belle II experiment



Charged particle identification

- **Charged hadrons:** TOP and ARICH info, dE/dx info from SVD and CDC.
- **Charged leptons:** primary info from ECL for e^\pm , from KLM for μ^\pm .

→ **PID variables**

PID variables are **likelihood** based:

detector likelihood \mathcal{L}_i^{det} for hypotheses $i = e^+, \mu^+, \pi^+, K^+, p^+, d^+$ → global likelihood $\mathcal{L}_i = \prod_{det} \mathcal{L}_i^{det}$

PID variables as **likelihood ratios**:

$$\text{kaonID} = \text{PID}_K = \frac{\mathcal{L}_K}{\sum_{j=e,\mu,\pi,K,p,d} \mathcal{L}_j}$$

data/MC correction weights from $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$:

- misID rate: π^+ mis-identified as K^+

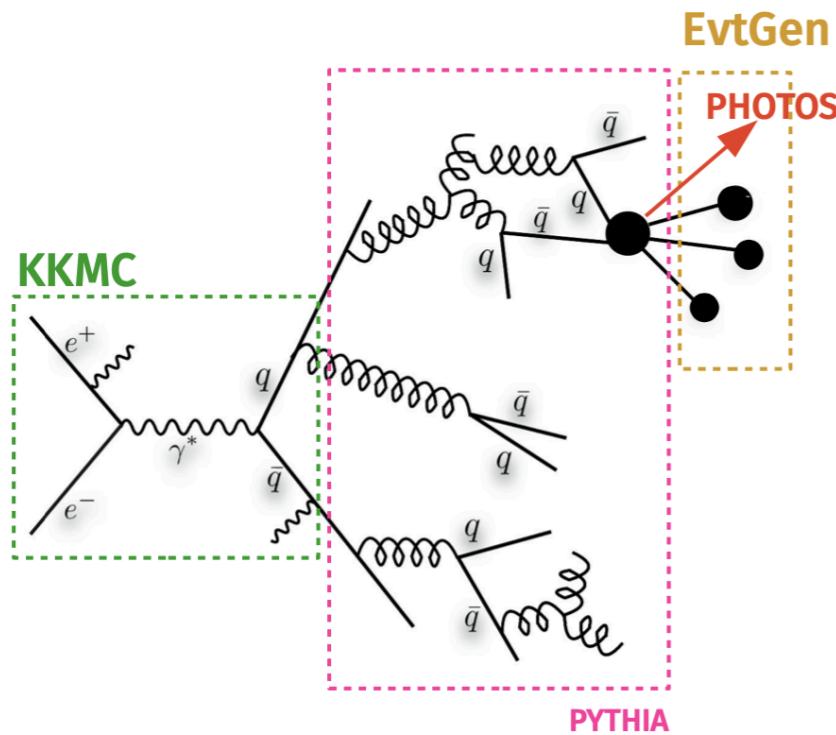
- slow pions to tag D^0 ;
- tagged D^0 to identify kaons and pions;
- efficiencies in $p, \cos\theta$ bins from fits to D^0 invariant mass distributions.

Generators



PYTHIA

Models parton shower and fragmentation into observable hadrons.



Implements Lund Model:

phenomenological model **parametrising QCD dynamics of hadronisation.**

Parameters tunable by comparison with experimental data.

Probabilistic production of new flavours and sharing of energy and momentum among decay products.

Relies on **fragmentation functions** quantifying **probability of parton to fragment into a hadron carrying a fraction z of the parton energy/momentum.**

→ Lund symmetric fragmentation functions

$$f(z) = \frac{1}{z} (1 - z)^{\text{StringZ:aLund}} e^{-\text{StringZ:bLund} m_T^2/z}$$

→ parameters modifying **s-quark, diquark and c-quark fragmentation**

→ the total width in the fragmentation process

$$\langle p_T^2 \rangle_{\text{hadron}} = 2 \text{StringPT : sigma}$$

Credit: A. Rostomyan

→ parameters to describe the non-Gaussian tail in transverse momentum distribution

Belle measurement [Phys. Rev. D 2020, 101] of charged di-hadron cross-sections ($e^+e^- \rightarrow h_1h_2X$) as function of fractional energy or momentum, z , finds **data-MC discrepancy of $\sim 40\%$ for kaon pairs at low z .**

Background suppression



Event-shape variables

- **Thrust:** $T(\vec{v}_T) = \max_{|\vec{v}_T|=1} \frac{\sum_i |\vec{v}_T \cdot \vec{p}_i|}{\sum_i |\vec{p}_i|}$

- **Harmonic moments:** coefficients of event expansion in spherical harmonics around an arbitrary axis (e.g., thrust axis \vec{v}_T).

$$B_l = \sum_i \frac{|\vec{p}_i|}{\sqrt{s}} P_l(\cos\alpha_i)$$

P_l \equiv Legendre polynomial

- **Fox-Wolfram (FW) moments:** rotationally invariant observables capturing distributions of momentum and energy flows.

$$H_l = \frac{4\pi}{2l+1} \sum_{m=-l}^{+l} \left| \sum_i Y_l^m(\Omega_i) \frac{|\vec{p}_i|}{\sqrt{s}} \right|^2 = \sum_{i,j} \frac{|\vec{p}_i| |\vec{p}_j|}{s} P_l(\cos\phi_{ij})$$

Normalised FW moments: $R_k = H_k / H_0$

- **Sphericity:** $S^{\alpha,\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |\vec{p}_i|^2} \rightarrow$ eigenvalues: $\lambda_1, \lambda_2, \lambda_3$

sphericity scalar: $S = \frac{3}{2}(\lambda_2 + \lambda_3)$, with $\lambda_2, \lambda_3 < \lambda_1$
 $S \simeq 0 \rightarrow$ jet-like events
 $S \simeq 1 \rightarrow$ spherical symmetry

- **Kakuno-Super-Fox-Wolfram (KSFW) moments:** distinguish signal B (s), ROE particles (o) and missing momentum (m)

Linear: $H_{i,l}^{so} = \sum_j C_{lj} \frac{|\vec{p}_j| P_l(\cos\theta_{Kj})}{2(\sqrt{s} - E_K^*)^2}$

$i = c \rightarrow \mathbf{j = charged ROE}$ $i = n \rightarrow \mathbf{j = neutral ROE}$ $i = m \rightarrow \mathbf{j = missing momentum}$

$l = \text{even} \rightarrow C_{lj} = 1$

$l = \text{odd} \rightarrow C_{lj} = 0$ if $i = m, n$, otherwise charge product j-k

Quadratic:

$$R_l^{oo} = \sum_i \sum_j C_{lij} \frac{|\vec{p}_i| |\vec{p}_j| P_l(\cos\theta_{ij})}{2(\sqrt{s} - E_K^*)^2}$$

$l = \text{even} \rightarrow C_{lij} = 1$

$l = \text{odd} \rightarrow C_{lij} = 0$ if $i, j = n$ otherwise i-j charge product



Data samples

B counting method

Number of $B\bar{B}$ estimated as:

$$N_{B\bar{B}} = \frac{(N_{had}^{on-res} - R_{lumi} \times N_{had}^{off-res} \times \kappa)}{\epsilon_{B\bar{B}}}$$

- N_{had}^{on-res} : N of selected hadronic events in on-resonance data;
- $R_{lumi} \times N_{had}^{off-res} \times \kappa$: estimated N of non- $B\bar{B}$ events in on-resonance data
 - R_{lumi} is the luminosity ratio $L_{on-res}/L_{off-res}$
 - $N_{had}^{off-res}$ is the number of hadronic events selected in off-resonance data
 - $\kappa = \sum_i \epsilon_i \sigma_i / \sum_i \epsilon'_i \sigma'_i$ where $\epsilon_i(\epsilon'_i)$ and $\sigma_i(\sigma'_i)$ are efficiencies and cross section of non- $B\bar{B}$ processes at $\Upsilon(4S)$ ($\Upsilon(4S) - 60$ MeV)
- $\epsilon_{B\bar{B}}$: hadronic-selection efficiency for $B\bar{B}$ events.

Estimated numbers of $B\bar{B}$ pairs in the collision-data samples used for the two analyses:

- $63 \text{ fb}^{-1} \rightarrow N_{B\bar{B}} = [68.21 \pm 0.06(\text{stat}) \pm 0.78(\text{syst})] \times 10^6$
- $189 \text{ fb}^{-1} \rightarrow N_{B\bar{B}} = [197.2 \pm 5.7] \times 10^6$



Statistical tools and methods

Fit tools



HistFactory and pyhf

Parametrised PDFs based on template histograms: statistical model built on simultaneous measurement of disjoint binned distributions (*channels*) of event counts \mathbf{n}

$$f(\mathbf{n}, \mathbf{a} | \boldsymbol{\eta}, \boldsymbol{\chi}) = \prod_{c \in \text{channels}} \prod_{b \in \text{bins}} \text{Pois}(n_{cb} | \nu_{cb}(\boldsymbol{\eta}, \boldsymbol{\chi})) \prod_{\chi \in \boldsymbol{\chi}} c_\chi(a_\chi | \chi)$$

- $\nu_{cb} \equiv$ expected event rate
- $\boldsymbol{\eta} \equiv$ free parameters
- $\boldsymbol{\chi} \equiv$ constrained parameters
- $c_\chi(a_\chi | \chi) \equiv$ 1D constraints built on auxiliary data a_χ

$$\nu_{cb} = \sum_{s \in \text{samples}} \nu_{scb}(\boldsymbol{\eta}, \boldsymbol{\chi}) = \sum_{s \in \text{samples}} \left(\prod_{\kappa \in \boldsymbol{\kappa}} \kappa_{scb}(\boldsymbol{\eta}, \boldsymbol{\chi}) \right) \left(\nu_{scb}^0(\boldsymbol{\eta}, \boldsymbol{\chi}) + \sum_{\Delta \in \boldsymbol{\Delta}} \Delta_{scb}(\boldsymbol{\eta}, \boldsymbol{\chi}) \right)$$

- samples \equiv physics processes
- $\nu_{scb}^0 \equiv$ nominal rates
- $\kappa_{scb}(\boldsymbol{\eta} | \boldsymbol{\chi}) \equiv$ multiplicative rate modifier
- $\Delta_{scb}(\boldsymbol{\eta} | \boldsymbol{\chi}) \equiv$ additive rate modifier
- $\kappa_{scb}(\alpha)$ change total number of events per sample but **keep shape invariant across bins.**
- $\Delta_{scb}(\alpha)$ produce correlated shape modifications in the various samples. **One common parameter shared among all the samples.**

Description	Modifier	Constraint term c_χ	Input
Uncorrelated shape	$\kappa_{scb}(\gamma_b) = \gamma_b$	$\prod_b \text{Pois}(r_b = \sigma_b^{-2} \rho_b = \sigma_b^{-2} \gamma_b)$	σ_b
Correlated shape	$\Delta_{scb}(\alpha) = f_p(\alpha \Delta_{scb, \alpha=-1}, \Delta_{scb, \alpha=+1})$	$\text{Gaus}(a = 0 \alpha, \sigma)$	$\Delta_{scb, \alpha=\pm 1}$
Normalisation uncertainty	$\kappa_{scb}(\alpha) = g_p(\alpha \kappa_{scb, \alpha=-1}, \kappa_{scb, \alpha=+1})$	$\text{Gaus}(a = 0 \alpha, \sigma)$	$\kappa_{scb, \alpha=\pm 1}$
MC stat. uncertainty	$\kappa_{scb}(\gamma_b) = \gamma_b$	$\prod_b \text{Gaus}(a_{\gamma_b} = 1 \gamma_b, \delta_b)$	$\delta_b^2 = \sum_s \delta_{sb}^2$
Luminosity	$\kappa_{scb}(\lambda) = \lambda$	$\text{Gaus}(l = \lambda_0 \lambda, \sigma_\lambda)$	$\lambda_0, \sigma_\lambda$
Normalisation	$\kappa_{scb}(\mu_b) = \mu_b$		
Data-driven shape	$\kappa_{scb}(\gamma_b) = \gamma_b$		

- $\sigma_b \equiv$ event-rate relative uncertainties
- $\delta_b \equiv$ event-rate uncertainties divided by total nominal rate
- $f_p(\alpha | \Delta_{scb, \alpha=-1}, \Delta_{scb, \alpha=+1})$ and $g_p(\alpha | \kappa_{scb, \alpha=-1}, \kappa_{scb, \alpha=+1})$ are **interpolating functions** where $\kappa_{scb, \alpha=\pm 1}$ and $\Delta_{scb, \alpha=\pm 1}$ represent $\pm 1\sigma$ variations of nominal expectations.

$$\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\theta}) = \mathcal{L}(\boldsymbol{\eta}, \boldsymbol{\chi})$$

$$-2\ln\lambda_p(\boldsymbol{\mu}) = -2\ln \frac{\mathcal{L}(\boldsymbol{\mu}, \hat{\boldsymbol{\theta}})}{\mathcal{L}(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\theta}})}$$

Fit tools



Correlated shape modifications

Captured by **bin-covariance matrix S** → **Goal:** decompose S into **vectors of orthogonal bin variations.**

- **Diagonalisation:** $S = VDV^{-1}$ - $V \equiv$ matrix of eigenvectors
 - $D \equiv$ diagonal matrix of eigenvalues $\lambda_i, i \in \text{bins}$
- **Matrix of principal components:** $U = \Lambda V^T$ - $\Lambda \equiv$ diagonal matrix of singular values $\sqrt{\lambda_i}$
- **Approximation:** $S \approx u^T u + \text{diag}(\sigma_1^2, \dots, \sigma_n^2)$
 - u matrix of the major principal components of U (largest $\sqrt{\lambda_i}$)
 - $\sigma_i^2 \equiv$ uncorrelated bin-variation: sum in quadrature of entries of a pruned principal component + MC stat. uncertainty.

HistFactory interpolates **variations in each u vector** to implement **correlated shape modifiers each depending on a continuous nuisance parameter.**

Fit tools



Simplified Gaussian model

The framework assumes independent bin counts **following Gaussian distributions with mean equal to the expected rates and Gaussian nuisance parameters.**

→ $\chi^2_c(\boldsymbol{\eta}, \boldsymbol{\chi}) = -2\ln\mathcal{L}_c(\boldsymbol{\eta}, \boldsymbol{\chi}) + \text{constant}$ for each channel.

$$\chi^2(\boldsymbol{\mu}, \boldsymbol{\theta}) = \sum_{b \in \text{bins}} \left(\frac{n_b - \sum_{s \in \text{samples}} M_b^s(\mu_s + \sum_{e \in \text{unc.}} \Gamma_{sb}^e \theta_e))}{\sigma_b} \right)^2 + \sum_{s \in \text{samples}} p_s^\mu (\mu_s - 1)^2 + \sum_{e \in \text{unc.}} p_e^\theta \theta_e^2$$

- n_b : observed counts in a bin.
- M_b^s : nominal expectations in a bin b for each sample s .
- Γ_{sb}^e : quantifies influence of uncertainty source e on a sample s in a bin b .
- μ_s : normalisation parameter of each sample s (unconstrained signal strength for signal).
- θ_e : nuisance parameter for a source of uncertainty e .
- p^μ, p_e^β : priors, by default set to unity (to $\mathcal{O}(10^{-4})$ for the signal strength).
- σ_b : total uncertainty in a bin b , corresponds to $\sqrt{n_b}$.

Upper limits



The CL_s method

Conservative frequentist approach: mitigate exclusion of signal models (μ) due to low experimental sensitivity.

Definition:

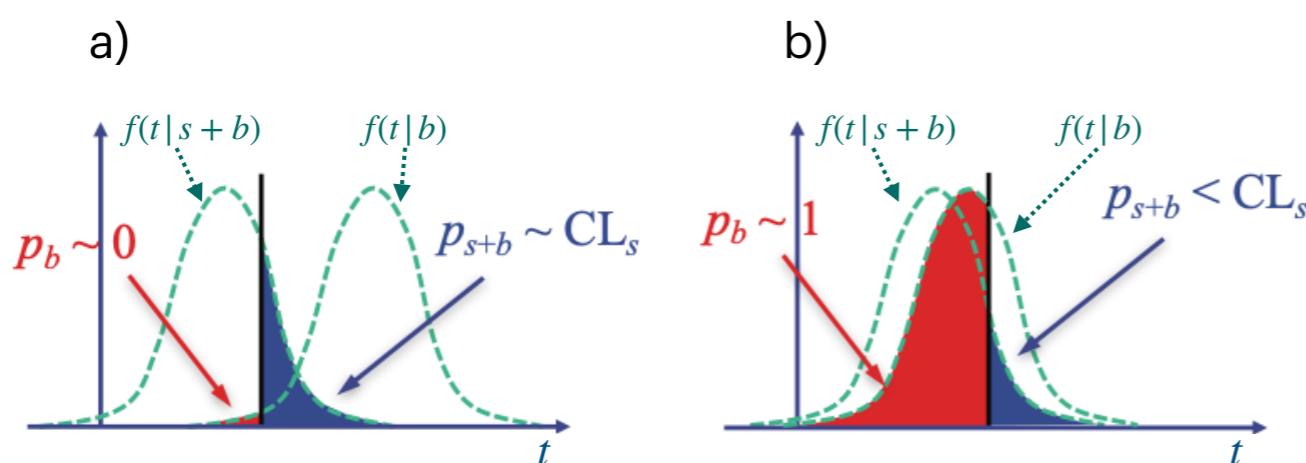
$$\text{CL}_s = \frac{\text{CL}_{s+b}}{\text{CL}_b} = \frac{p_{s+b}}{1 - p_b}$$

- $\text{CL}_{s+b} = p_{s+b} = P(t \geq t_{obs} | s + b) = \int_{t_{obs}}^{\infty} f(t | s + b) dt = 1 - F(t_{obs} | \mu)$
- $\text{CL}_b = 1 - p_b = P(t \geq t_{obs} | b) = 1 - \int_{-\infty}^{t_{obs}} f(t | b) dt = 1 - F(t_{obs} | 0)$

[test statistic t based on profile likelihood ratio $\mathcal{L}(\mu, \hat{\theta}) / \mathcal{L}(\hat{\mu}, \hat{\theta})$]

Signal model μ rejected if $\text{CL}_s \leq \alpha$ (corresponding $\text{CL} = 1 - \alpha$)

CL_s always larger than simple p_{s+b} : $1 - p_b \leq 1 \rightarrow \text{CL}_s \geq p_{s+b} \rightarrow \text{CL}_s$ conservative



- a) $f(t | s + b), f(t | b)$ **well separated** $\rightarrow \text{CL}_s \simeq p_{s+b}$
- b) **low sensitivity to signal:** $f(t | s + b), f(t | b)$ largely overlap $\rightarrow p_{s+b}, 1 - p_b$ small and $\text{CL}_s \leq \alpha$ not satisfied (signal model not rejected).



Upper limits

CL_s in pyhf

Test statistics

$$q_\mu = \begin{cases} -2\ln\lambda_p(\mu) & \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu. \end{cases}$$

$$\tilde{q}(\mu) = \begin{cases} -2\ln\tilde{\lambda}_p(\mu) & \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu. \end{cases} = \begin{cases} -2\ln\frac{\mathcal{L}(\mu, \hat{\theta}(\mu))}{\mathcal{L}(0, \hat{\theta}(0))} & \hat{\mu} < 0, \\ -2\ln\frac{\mathcal{L}(\mu, \hat{\theta})}{\mathcal{L}(\hat{\mu}, \hat{\theta})} & 0 \leq \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu. \end{cases}$$

Toys

$$\text{CL}_s = \frac{N(q \geq q_{obs} | s + b)}{N(q \geq q_{obs} | b)}$$

Asymptotic formulae

$$q_\mu = \begin{cases} \frac{(\mu - \hat{\mu})^2}{\sigma^2} & \hat{\mu} < \mu, \\ 0 & \hat{\mu} > \mu, \end{cases}$$

$$F(q_\mu | \mu) = \Phi(\sqrt{q_\mu}) ,$$

$$\tilde{q}(\mu) = \begin{cases} \frac{\mu^2}{\sigma^2} - \frac{2\mu\hat{\mu}}{\sigma^2} & \hat{\mu} < 0, \\ \frac{(\mu - \hat{\mu})^2}{\sigma^2} & 0 \leq \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu, \end{cases}$$

$$F(\tilde{q}_\mu | \mu) = \begin{cases} \Phi(\sqrt{\tilde{q}_\mu}) & 0 < \tilde{q}(\mu) \leq \mu^2/\sigma^2, \\ \Phi\left(\frac{\tilde{q}_\mu + \mu^2/\sigma^2}{2\mu/\sigma}\right) & \tilde{q}(\mu) > \mu^2/\sigma^2, \end{cases}$$

Upper limits



Limits with sghf

sghf estimators have **Gaussian distributions** → upper limit on μ in terms of **Gaussian significance Z** .

For a chosen $CL = 1 - \alpha$:

$$\mu_{\text{up}} = Z \cdot \sigma_\mu = \Phi^{-1}(1 - \alpha) \cdot \sigma_\mu$$

where $\sigma_\mu \equiv$ estimated uncertainty on μ .

$\Phi \equiv$ standard Gaussian cumulative distribution

If $CL = 90\%$:

$$Z = \Phi^{-1}(0.90) = 1.645 \rightarrow \mu_{\text{up}} = 1.645 \cdot \sigma_\mu$$



Search for $B^+ \rightarrow K^+ \nu \bar{\nu}$

with an inclusive tagging

at Belle II

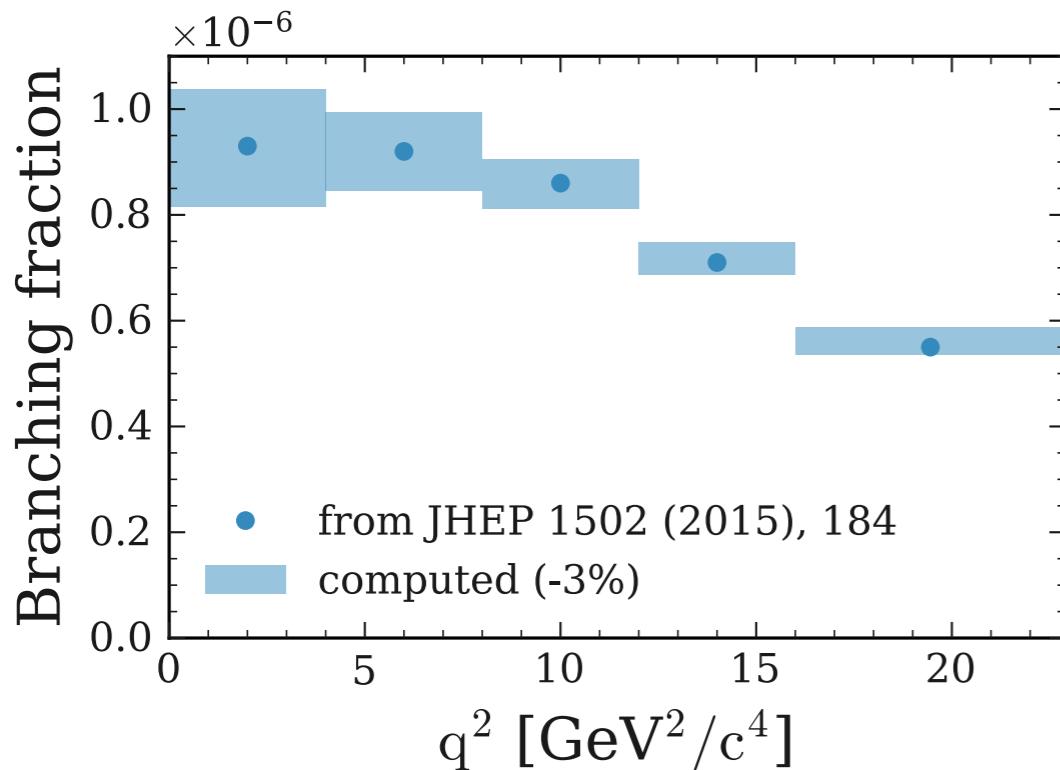
Simulated signal



Reweighting procedure

Validation

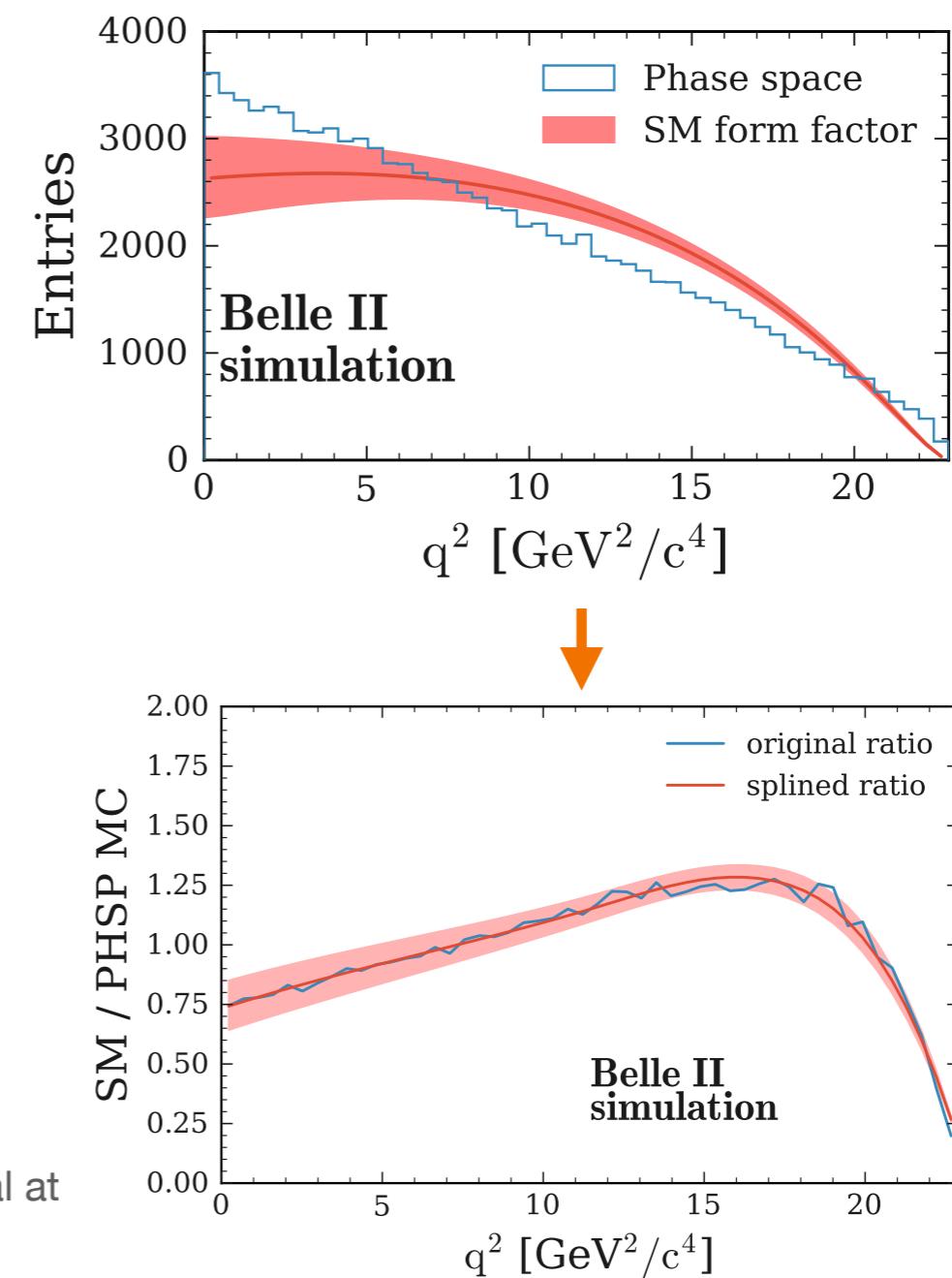
Computation based on FF estimate from fit to Lattice QCD and LCSR results and input parameters from [J. High Energ. Phys. 2015, 184](#)



Spline interpolation: piecewise polynomial continuous over equispaced points.

Prevents oscillation at edges of the interval at higher degrees.

Computation of the weights:



$B^+ \rightarrow K^+ \nu \bar{\nu}$ reconstruction



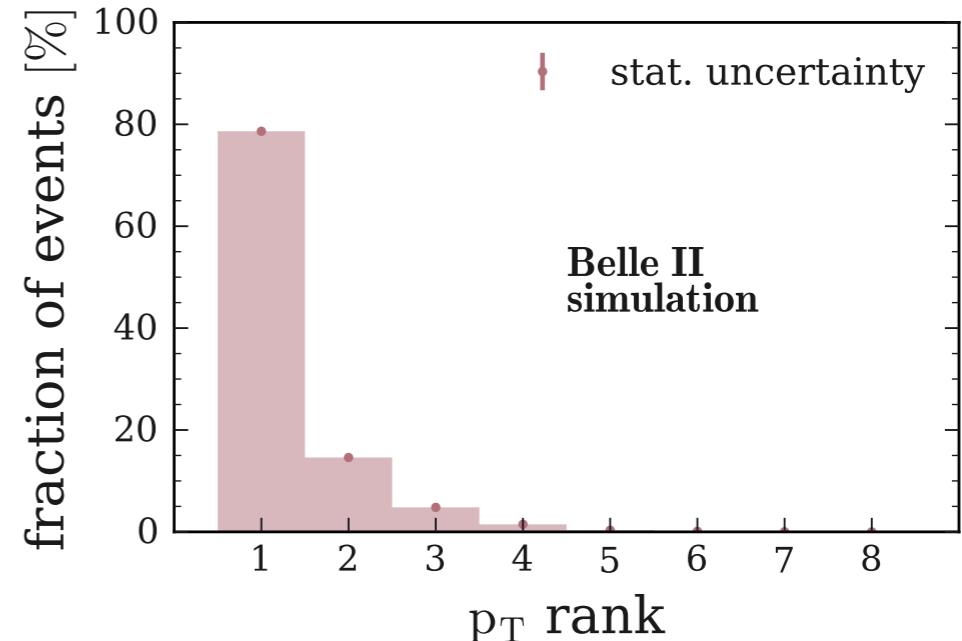
Object selection

- **Track cleanup:** $p_T > 0.1 \text{ GeV}/c$,
 $E < 5.5 \text{ GeV}$,
 $|dz| < 3 \text{ cm}$,
 $dr < 0.5 \text{ cm}$,
 $17^\circ < \theta < 150^\circ$,
- **Photon cleanup:** $0.1 \text{ GeV} < E_{\text{ECL}} < 5.5 \text{ GeV}$,
 $17^\circ < \theta < 150^\circ$,

+ ≥ 1 PXD hit
for signal track

Signal selection

Signal candidate matched to a true signal kaon in the acceptance as a function of p_T rank



Event selection

- $4 < \text{nTracks} < 10$
- $17^\circ < \theta(\vec{p}_{\text{miss.}}) < 160^\circ$
- $E_{\text{CMS}} > 4.0 \text{ GeV}$

Good suppression of low-multiplicity background like $\gamma\gamma \rightarrow 4\pi$.

Selection stage	Signal efficiency (%)
Object selection	89.6 ± 0.3
Signal selection	67.4 ± 0.5
Event selection	57.4 ± 0.5

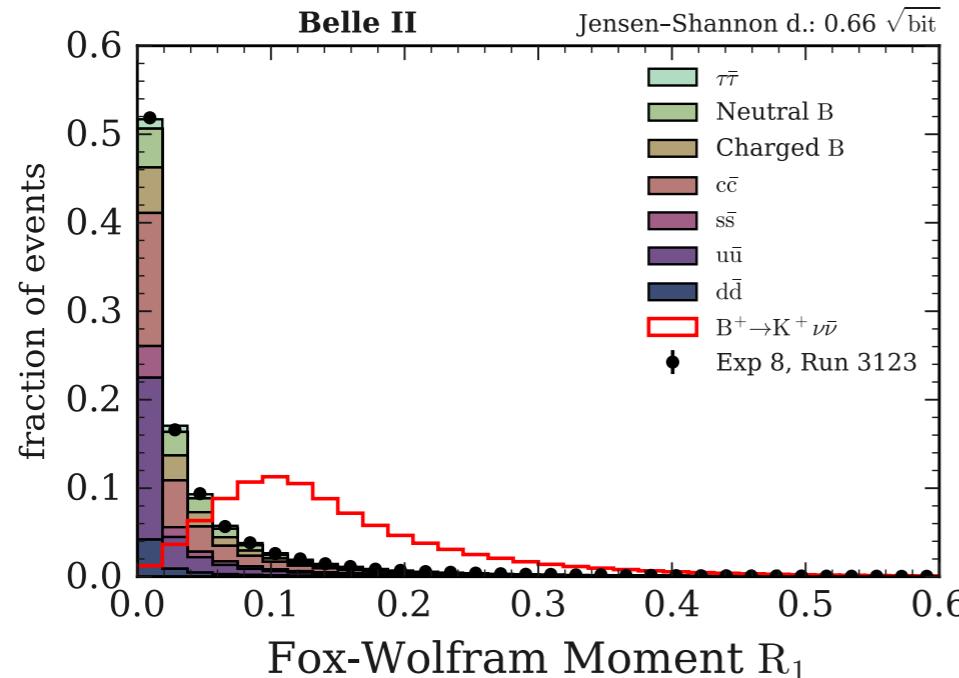
Keep high signal efficiency in the initial stage of the analysis

Background suppression

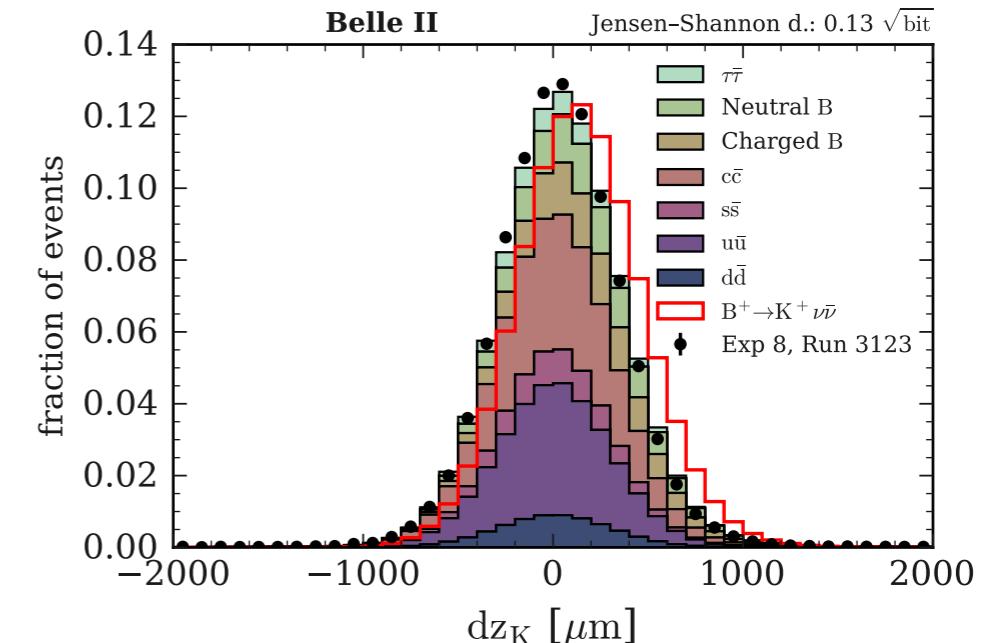


Features of $B^+ \rightarrow K^+ \nu \bar{\nu}$

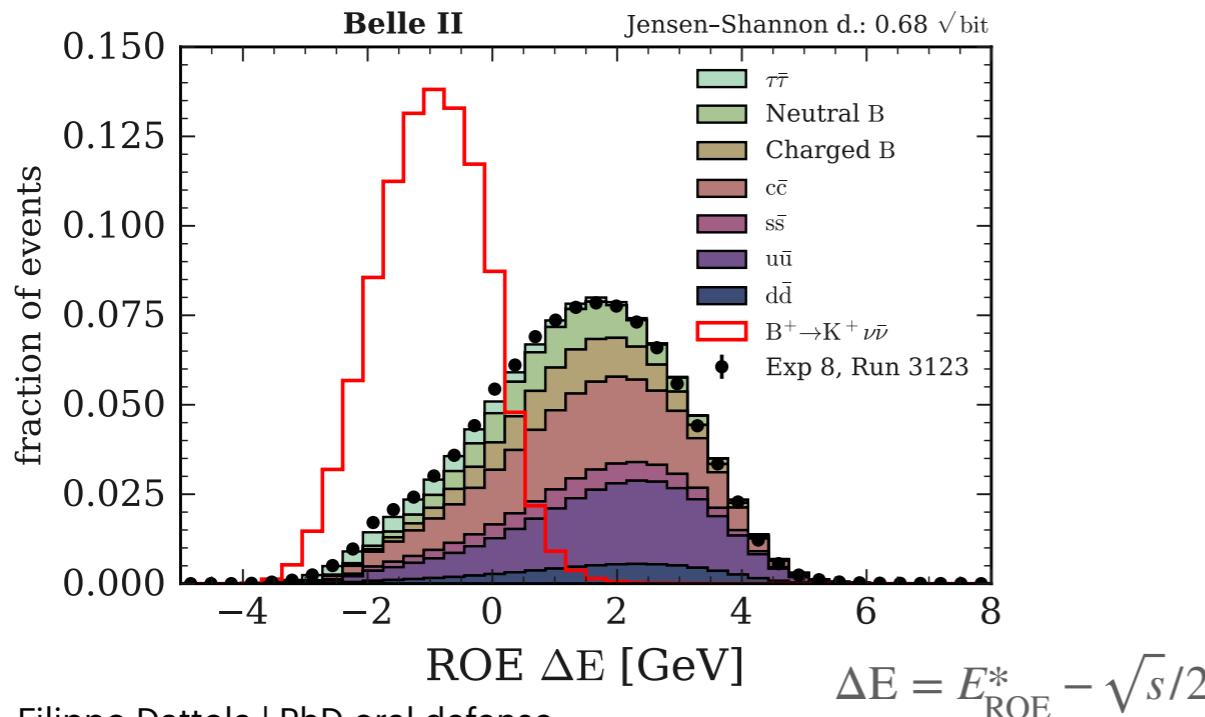
Event-shape variables



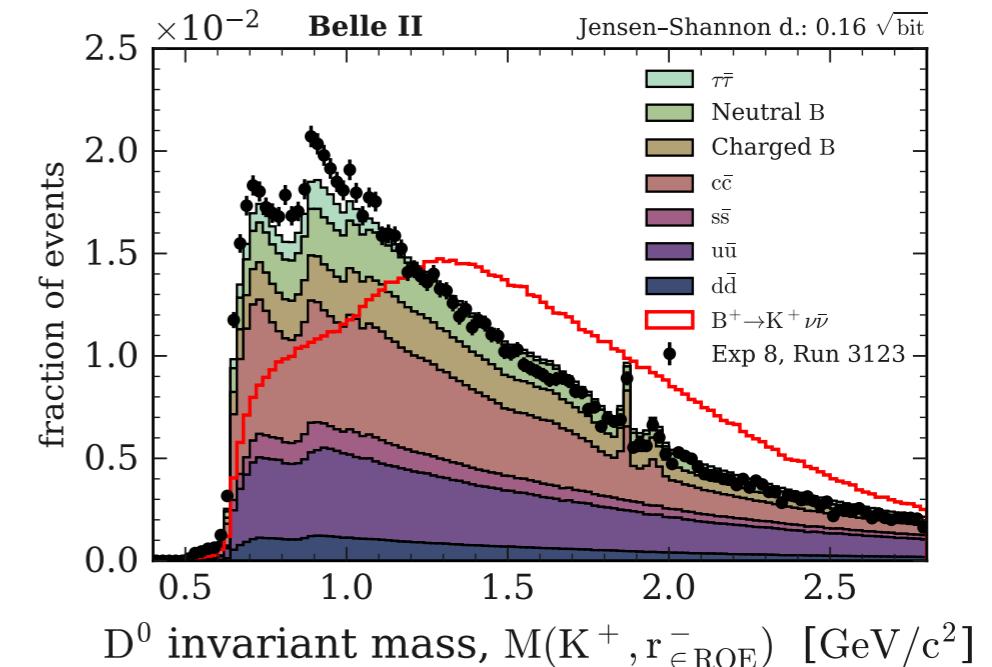
Variables related to the **signal-kaon candidate**



Variables related to the **ROE**



Variables related to **D-background suppression**



Multivariate classification



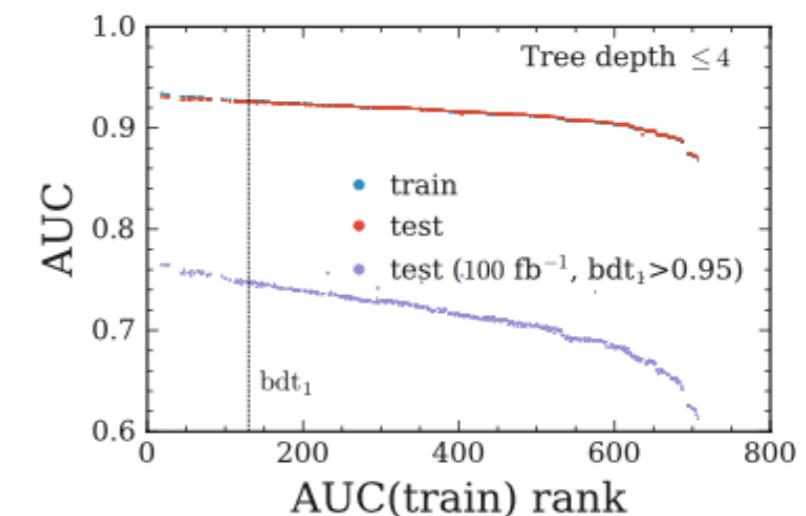
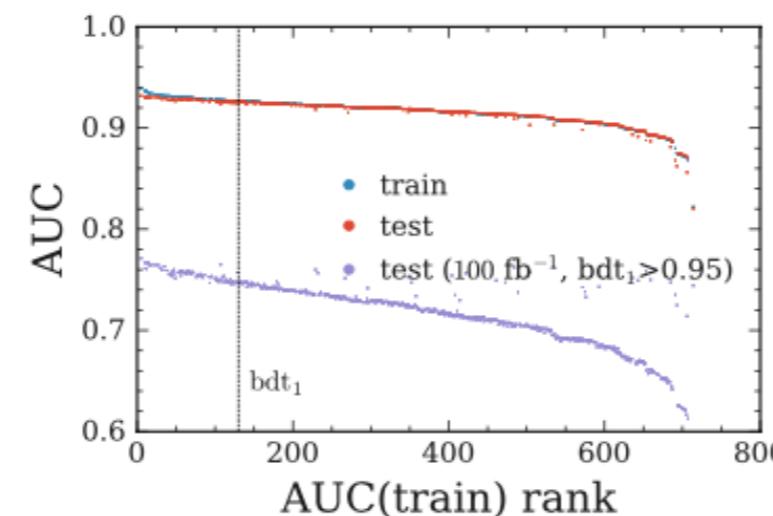
- **BDT₁ and BDT₂ classification model [FastBDT]**

Hyperparameter	Value
Number of trees	2000
Tree depth	4
Shrinkage	0.1
Sampling rate	0.8
Number of equal-frequency bins	16

- **Grid-search optimisation:**

Hyperparameter	basf2 name	Tested values	Chosen value
n_T	m_nTrees	[200, 500, 1000, 2000]	2000
d_T	m_nLevels	[2, 3, 4, 5, 6]	4
n_C	m_nCuts	[3, 4, 5, 6, 7, 8]	4
η	m_shrinkage	[0.05, 0.1, 0.2]	0.1
α	m_randRatio	[0.5, 0.8]	0.8

evaluation based on area under ROC curve



Multivariate classification



- Feature importances

- BDT_1 :

Training variable	Importance score
First normalised Fox-Wolfram moment R_1 computed in the CMS	0.323
ΔE of the ROE	0.167
Linear Fox-Wolfram moment $H_{m,2}^{so}$ computed in the CMS	0.071
Zeroth-order harmonic moment B_0 with respect to the thrust axis in the CMS	0.061
Normalised Fox-Wolfram moment R_1 computed in the CMS	0.051
Magnitude of the ROE momentum	0.038
Normalised Fox-Wolfram moment R_2 computed in the CMS	0.028
Squared missing invariant mass in the event	0.024
Invariant mass of the ROE	0.024
Linear Fox-Wolfram $H_{m,4}^{so}$ moment computed in the CMS	0.023

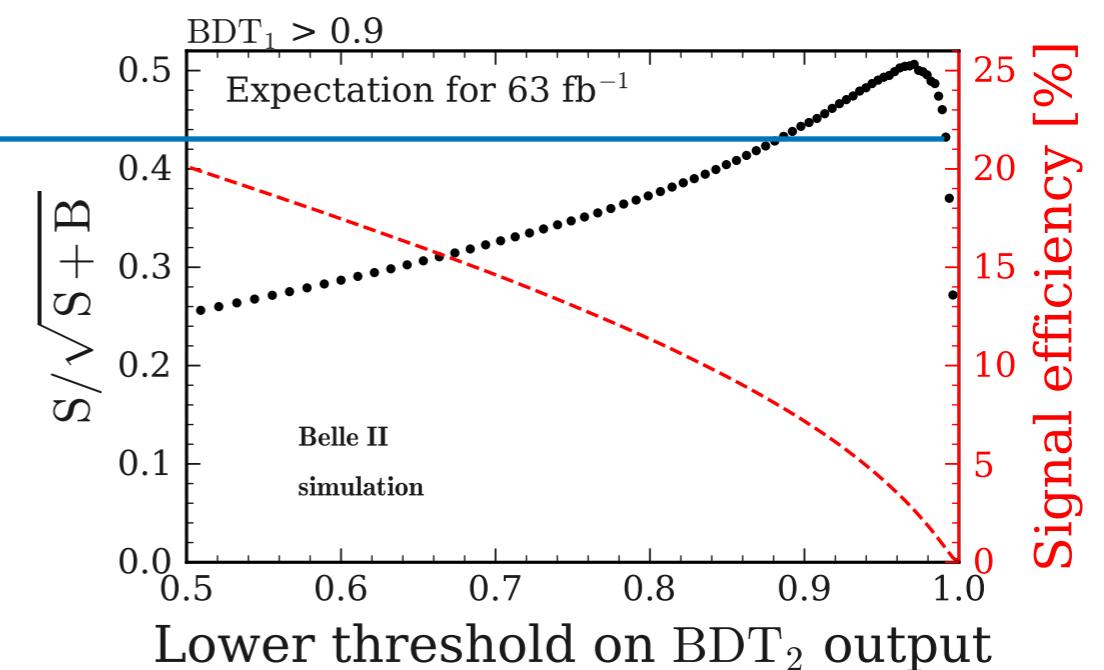
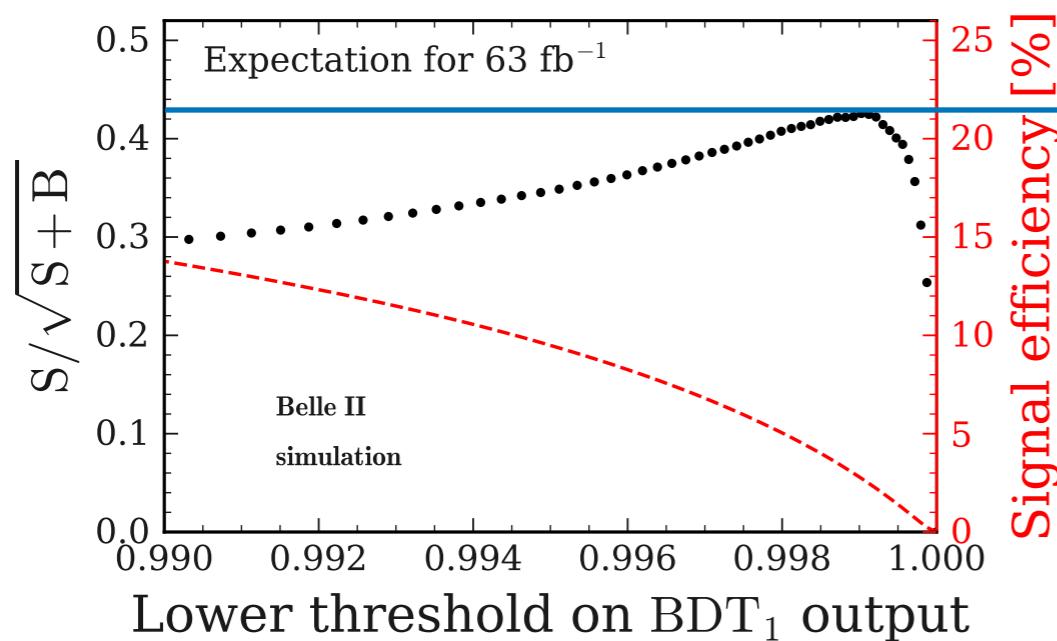
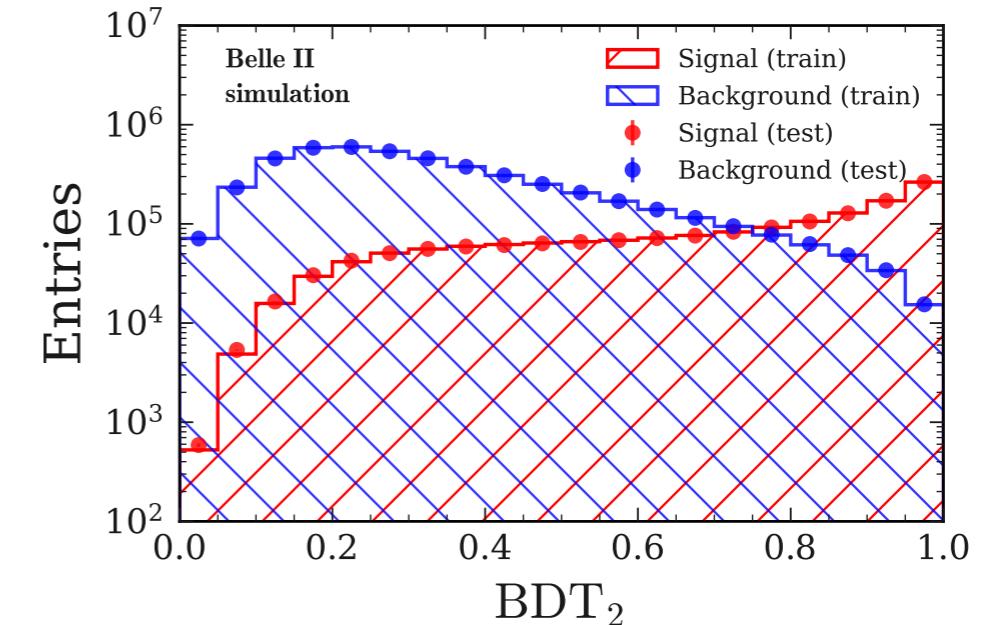
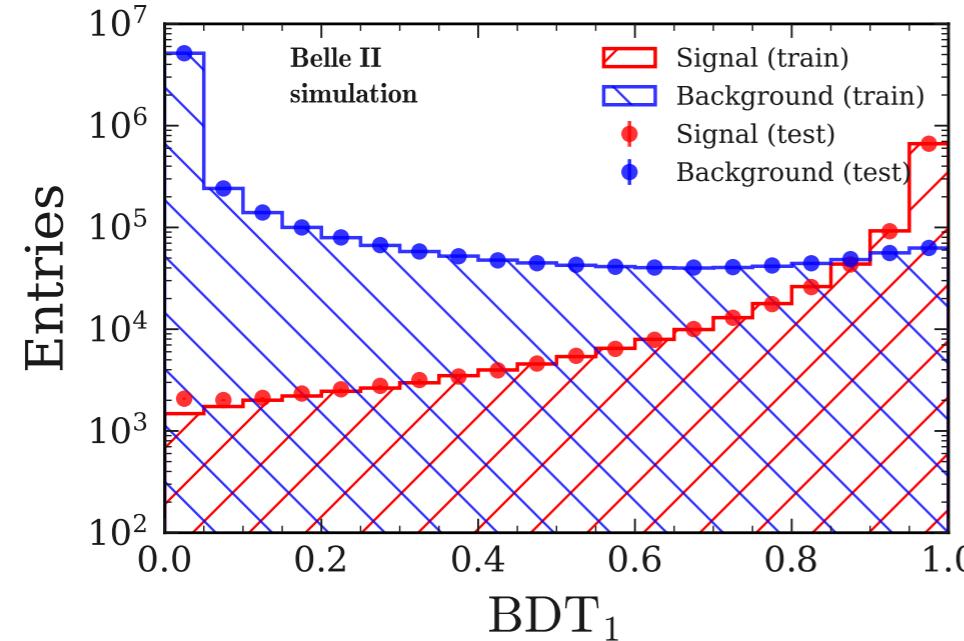
- BDT_2 :

Training variable	Importance score
Linear Fox-Wolfram moment $H_{m,2}^{so}$ computed in the CMS	0.118
ΔE of the ROE	0.053
p -value of the ROE vertex fit	0.050
Event sphericity computed in the CMS	0.039
Linear Fox-Wolfram moment $H_{c,2}^{so}$ computed in the CMS	0.039
Linear Fox-Wolfram moment $H_{m,4}^{so}$ computed in the CMS	0.035
Number of charged lepton candidates in the event	0.034
dz of the signal-kaon candidate track from the ROE vertex	0.034
χ^2 -probability of the vertex fit to the best D^+ candidate	0.033
Median χ^2 -probability of the vertex fits to all the D^0 candidates	0.033

Multivariate classification



- Performance



BDT₂ on top of BDT₁:
+35 % purity at 4 % efficiency.

Validation studies



Control channel: $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$

Mode with **sizeable branching fraction** characterised by **clean experimental signature**.

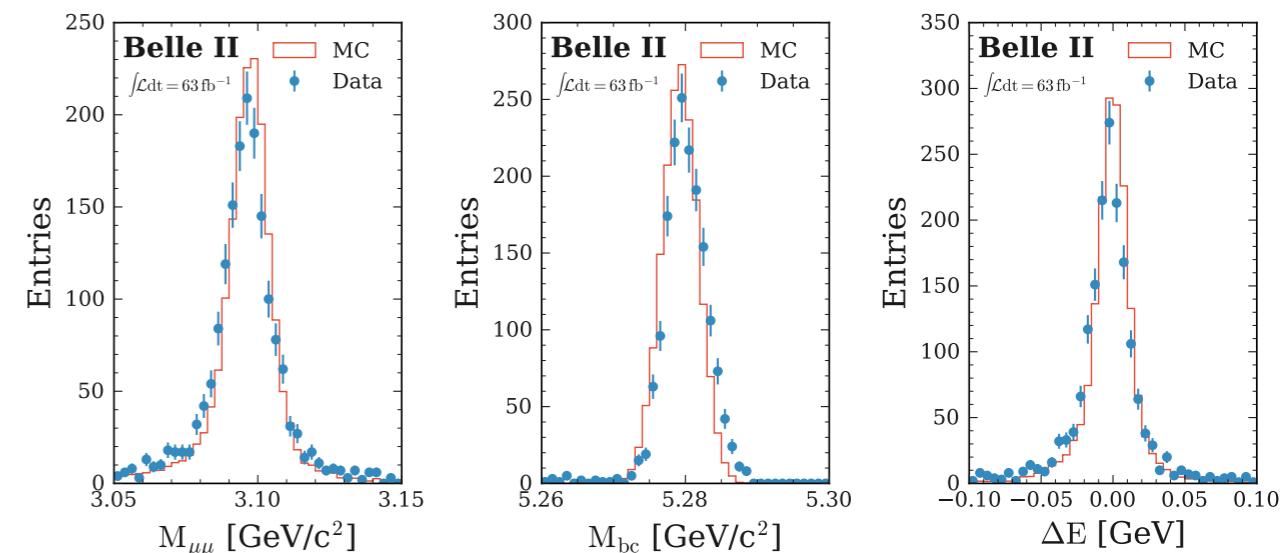
The method

Identification of $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$ events

- Candidate muons: track cleanup + muonID > 0.5
- Candidate J/ψ : pair of selected muons with $|M_{J/\psi}^{\text{PDG}} - M_{\mu^+\mu^-}| < 50 \text{ MeV}/c^2$
- Candidate K^+ : highest-pT track in the event + kaonID > 0.1 and muonID < 0.5
- Candidate B^+_{sig} : combine K^+ and J/ψ candidates + require $M_{bc} > 5.25 \text{ GeV}/c^2$ and $|\Delta E| < 100 \text{ MeV}$
- **1720** $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$ **candidates** in 1722 events reconstructed in data.
- **46% selection efficiency** estimated in MC simulating the decay.
- **Background level < 5%**.

$$\begin{aligned}\text{BR}(B^+ \rightarrow K^+ J/\psi) &= (1.020 \pm 0.019) \times 10^{-3} \\ \text{BR}(J/\psi \rightarrow \mu^+ \mu^-) &= (5.961 \pm 0.033) \times 10^{-2}\end{aligned}$$

[Prog. Theor. Exp. Phys 2022, 083C01]



$$M_{bc} = \sqrt{s/4 - p_B^{*2}}$$

$$\Delta E = E_B^* - \sqrt{s}/2$$

Validation studies



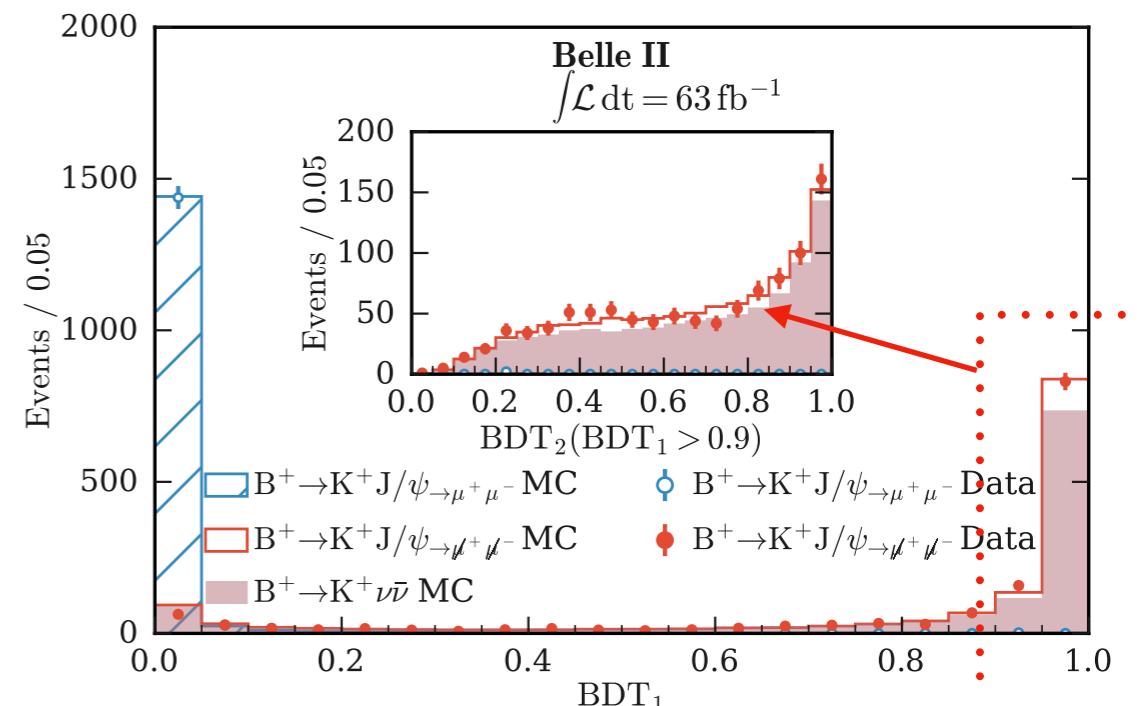
Control channel: $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$

- Inclusive-tagging selection efficiency:
 - $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$ in data $\rightarrow (81.5 \pm 0.9)\%$
 - $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$ in MC $\rightarrow (83.8 \pm 0.2)\%$
- Selection efficiency at $BDT_1 > 0.9$ and $BDT_2 > 0.95$,
region of highest signal sensitivity:
 - $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$ in data $\rightarrow (9.36 \pm 0.71)\%$
 - $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$ in MC $\rightarrow (8.85 \pm 0.13)\%$
- K-S test data vs MC for $BDT_2(BDT_1 > 0.9)$:
p-value = 23%

$$\text{BR}(B^+ \rightarrow K^+ J/\psi) = (1.020 \pm 0.019) \times 10^{-3}$$

$$\text{BR}(J/\psi \rightarrow \mu^+ \mu^-) = (5.961 \pm 0.033) \times 10^{-2}$$

[Prog. Theor. Exp. Phys 2022, 083C01]



Validation studies



Reweighting of the continuum simulation

Data-driven technique relying on a **binary classifier**:

- enhance the data-MC agreement by **event reweighting of the MC**.

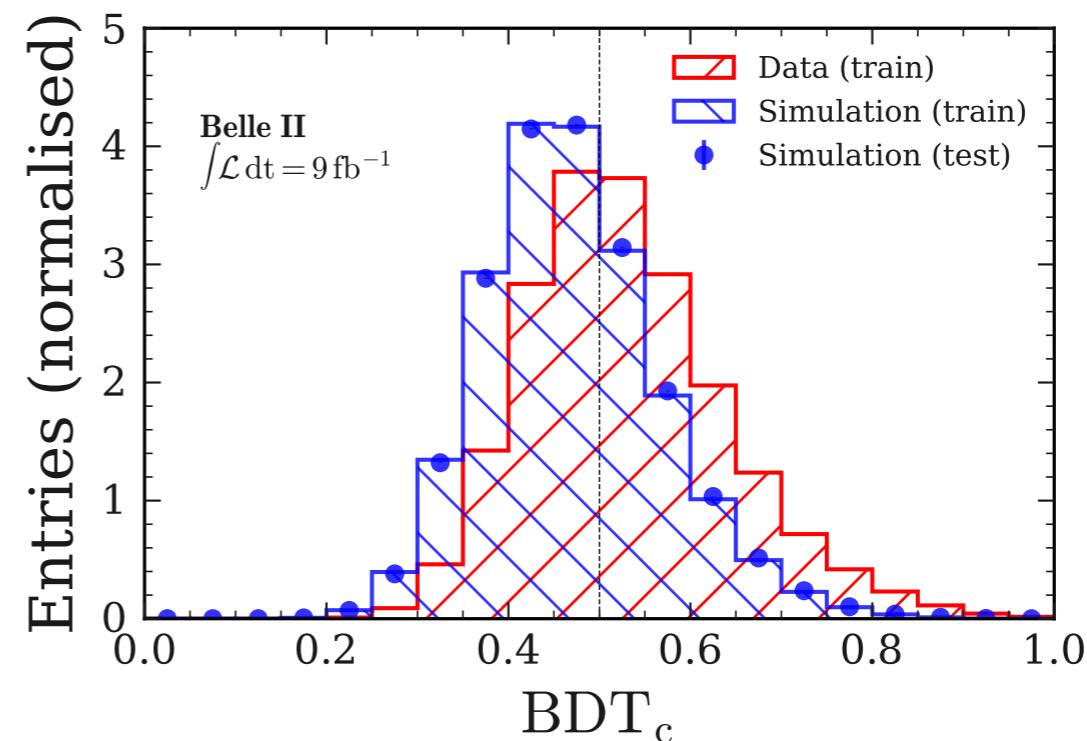
Training samples

- Simulated events of the continuum bkg with $BDT_1 > 0.9$ reconstructed in MC sample of 100 fb^{-1} .
- Events with $BDT_1 > 0.9$ reconstructed in the off-resonance data sample of 9 fb^{-1} .

The classifier

BDT_C implements same model of BDT_1 and BDT_2 : trained with the same 51 input variables using simulated (off-resonance data) events as background (signal) class.

Reliable performance, **no overfitting**.



Validation studies



Reweighting of the continuum simulation

Data-driven technique relying on a **binary classifier**:

- enhance the data-MC agreement by **event reweighting of the MC**.

Training samples

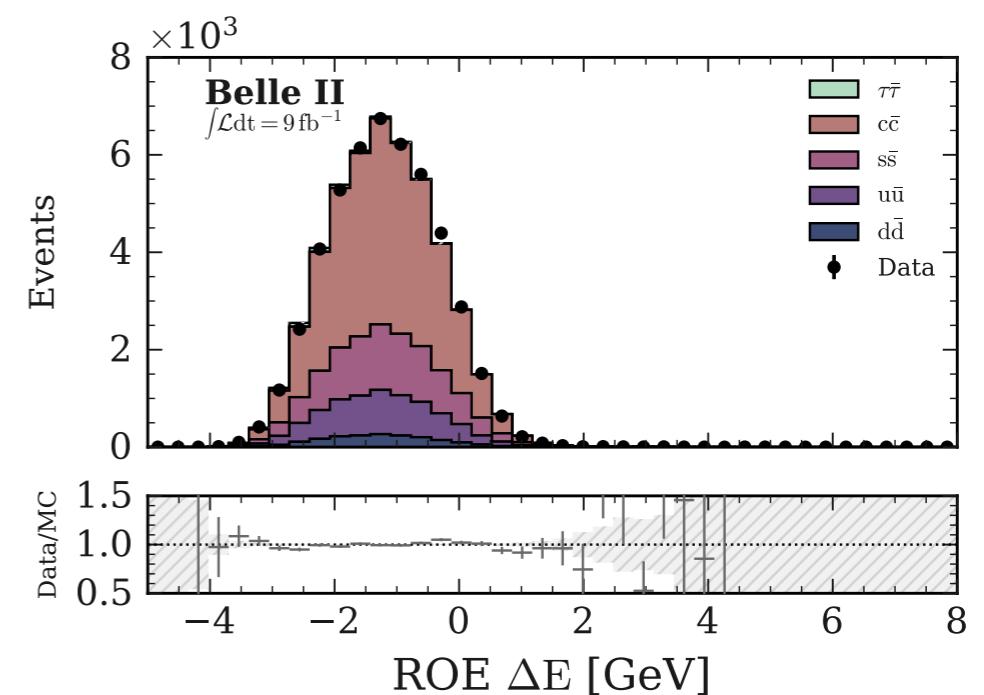
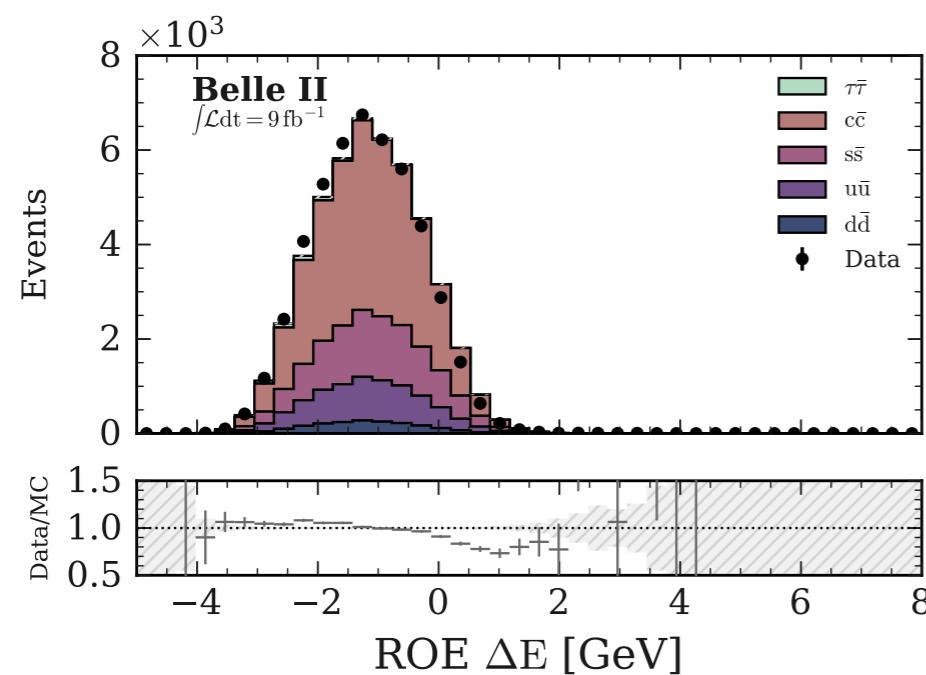
- Simulated events of the continuum bkg with $BDT_1 > 0.9$ reconstructed in MC sample of 100 fb^{-1} .
- Events with $BDT_1 > 0.9$ reconstructed in the off-resonance data sample of 9 fb^{-1} .

The classifier

BDT_C implements same model of BDT_1 and BDT_2 : trained with the same 51 input variables using simulated (off-resonance data) events as background (signal) class.

The event weight

Output p of $BDT_C \rightarrow$ weight $p/(1-p)$ assigned to simulation (\sim likelihood ratio $L(\text{data})/L(\text{MC})$).



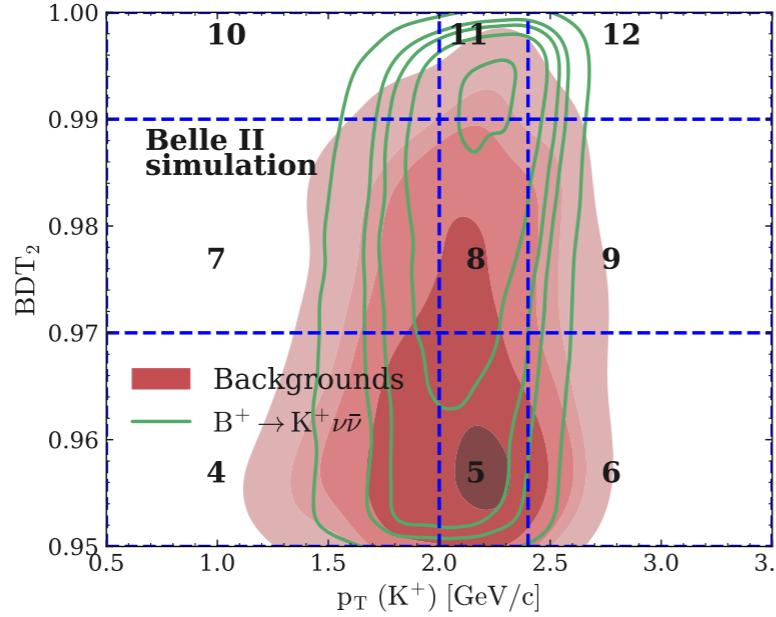
Definition of the fit regions



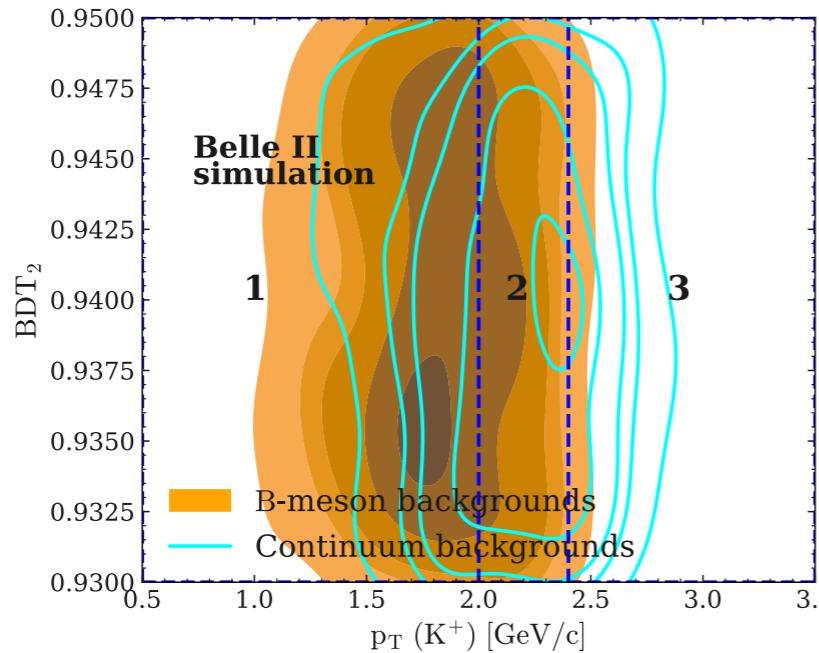
Optimised bin boundaries

Goal: maximise the separation of signal of signal from background.

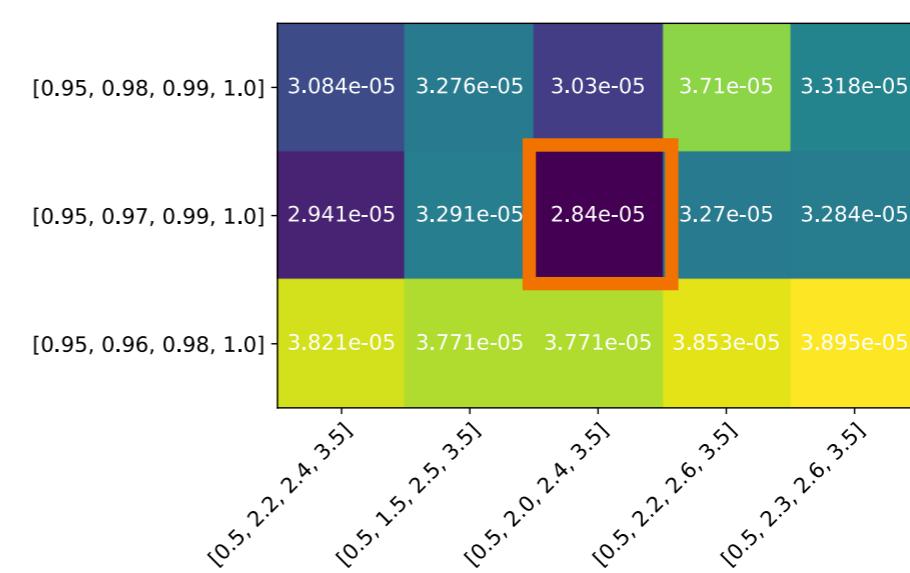
**Signal region
(and CR2)**



**CR1
(and CR3)**



**Belle II
simulation**



CL_S sensitivity scan
(MC stat. + norm)

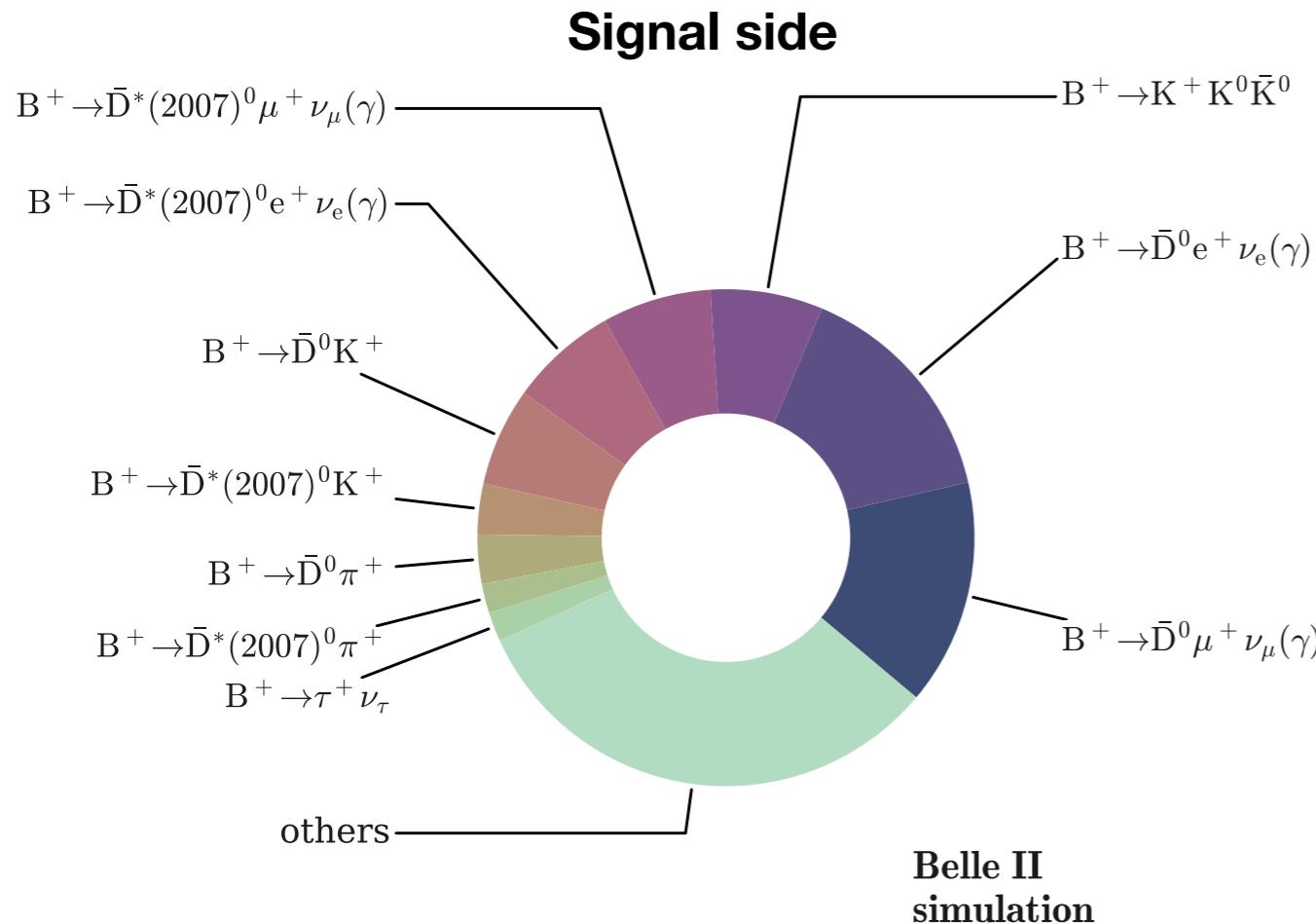
Contours: from innermost to outermost contain 5, 40, 65, 75, 90 % of the reconstructed events.

Background composition



Charged B background

Major background decays in the reconstructed events with $\text{BDT}_1 > 0.9$ and $\text{BDT}_2 > 0.93$



Inclusive D^0 modes

- $\text{BR}(D^0 \rightarrow K^- X) \simeq 55\%$
- $\text{BR}(D^0 \rightarrow K^+ X) \simeq 3\%$

Inclusive B^+ modes

- $\text{BR}(B^+ \rightarrow \bar{D}^0 X) \simeq 79\%$
- $\text{BR}(B^+ \rightarrow D^0 X) \simeq 9\%$

Specific exclusive B^+ modes

- $\text{BR}(B^+ \rightarrow \bar{D}^0 l^+ \nu_l) \simeq 2\%$
- $\text{BR}(B^+ \rightarrow \bar{D}^*(2007)^0 l^+ \nu_l) \simeq 6\%$
- $\text{BR}(B^+ \rightarrow \bar{D}^{(*)0} \pi^+) \simeq 5 \times 10^{-3}$
- $\text{BR}(B^+ \rightarrow \bar{D}^{(*)0} K^+) \simeq 4 \times 10^{-4}$
- $\text{BR}(B^+ \rightarrow \tau^+ \nu_\tau) \simeq 10^{-4}$
- $\text{BR}(B^+ \rightarrow K^+ K^0 \bar{K}^0) \simeq 10^{-6}$

Major exclusive D^0 modes

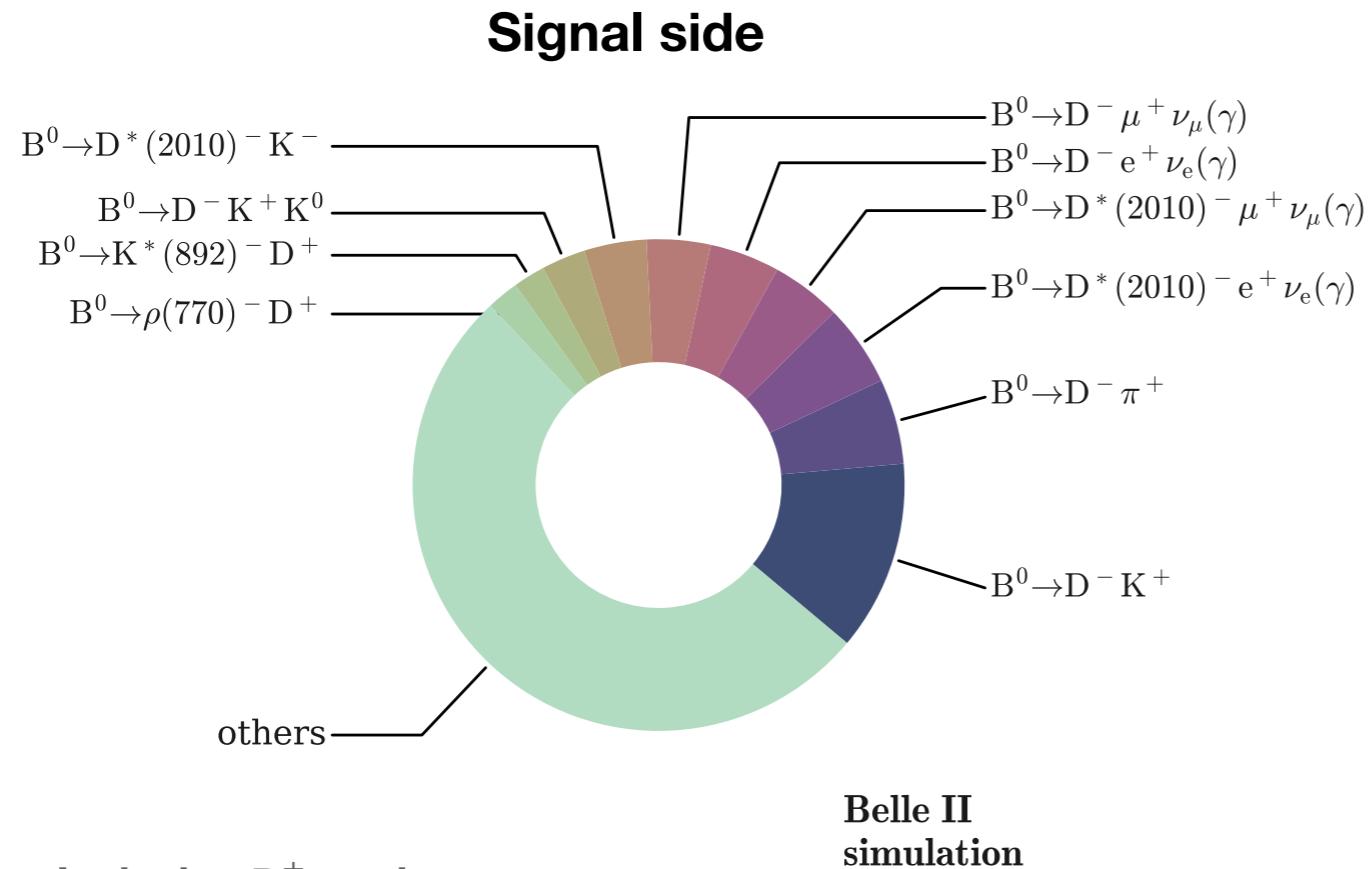
- $\text{BR}(D^0 \rightarrow K^- l^+ \nu_l) \simeq 7\%$
- $\text{BR}(D^0 \rightarrow K^*(892)^- l^+ \nu_l) \simeq 4\%$
- $\text{BR}(D^0 \rightarrow K^- \pi^+) \simeq 4\%$
- $\text{BR}(D^0 \rightarrow K^- \pi^+ \pi^0) \simeq 14\%$
- $\text{BR}(D^0 \rightarrow K^- \rho^+) \simeq 11\%$

Background composition



Neutral B background

Major background decays in the reconstructed events with $\text{BDT}_1 > 0.9$ and $\text{BDT}_2 > 0.93$



Inclusive D^\pm modes

- $\text{BR}(D^\pm \rightarrow K^\pm X) \simeq 26\%$
- $\text{BR}(D^\pm \rightarrow K^\mp X) \simeq 6\%$

D^* modes

- $\text{BR}(D^*(2010)^+ \rightarrow D^0 \pi^+) \simeq 68\%$
- $\text{BR}(D^*(2010)^+ \rightarrow D^+ \pi^0) \simeq 31\%$
- $\text{BR}(D^*(2010)^+ \rightarrow D^+ \gamma) \simeq 1\%$

Inclusive B^0 modes

- $\text{BR}(B^0 \rightarrow D^- X) \simeq 37\%$

Specific exclusive B^0 modes

- $\text{BR}(B^0 \rightarrow D^- l^+ \nu_l) \simeq 5\%$
- $\text{BR}(B^0 \rightarrow \bar{D}^*(2010)^- l^+ \nu_l) \simeq 5\%$
- $\text{BR}(B^0 \rightarrow D^- \pi^+) \simeq 2 \times 10^{-3}$
- $\text{BR}(B^0 \rightarrow D^- \rho^+) \simeq 8 \times 10^{-3}$
- $\text{BR}(B^0 \rightarrow D^{(*)-} K^+) \simeq 2 \times 10^{-4}$

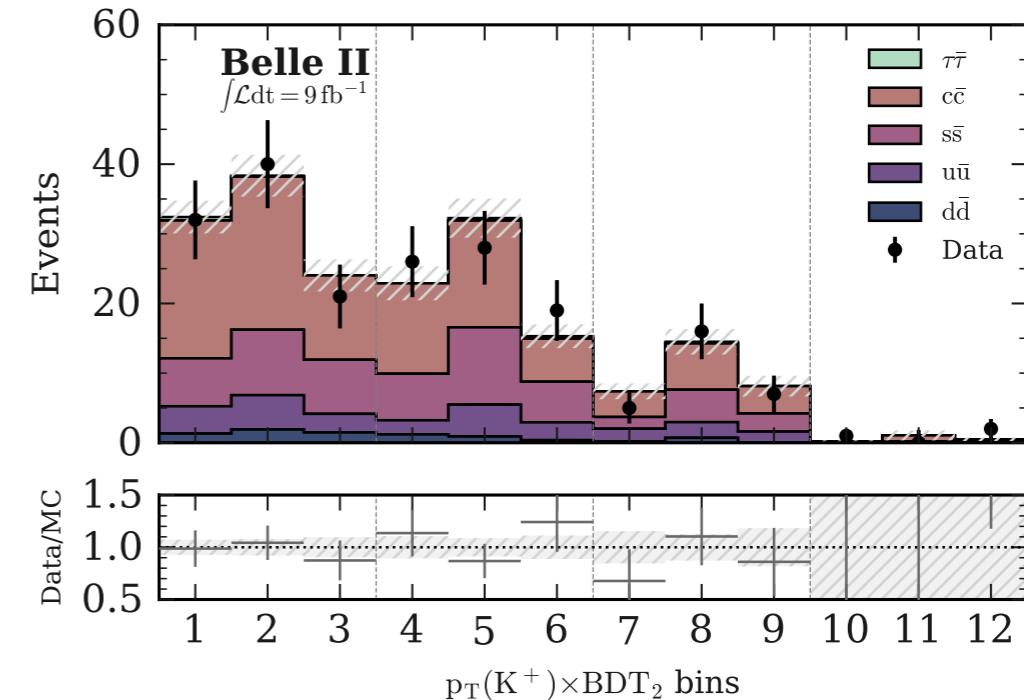
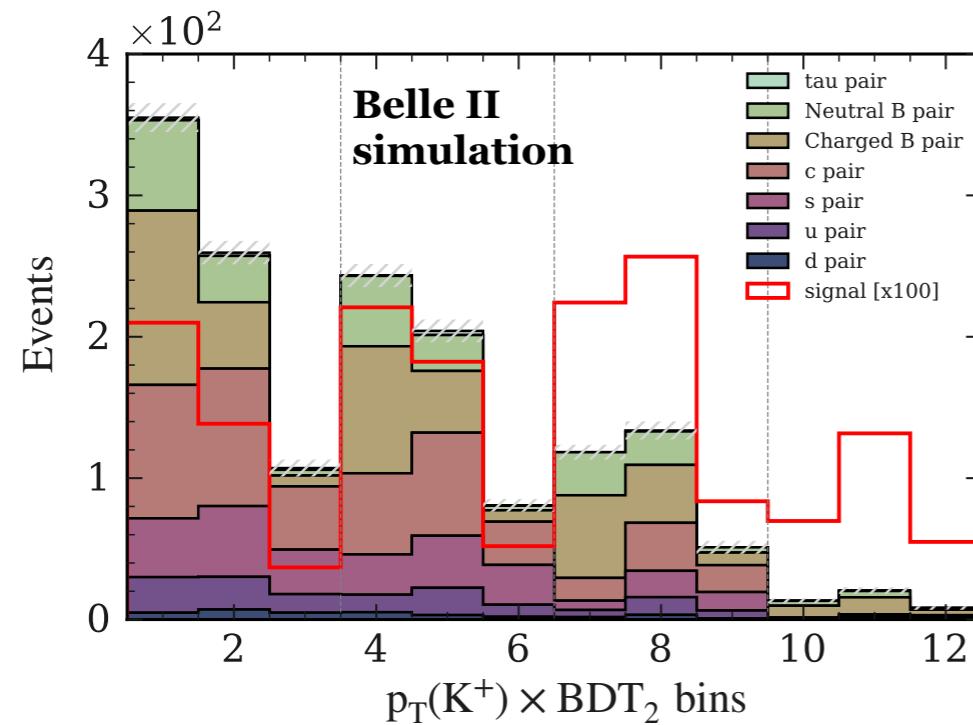
Major exclusive D^\pm modes

- $\text{BR}(D^+ \rightarrow \bar{K}^0 l^+ \nu_l) \simeq 20\%$
- $\text{BR}(D^+ \rightarrow \bar{K}^*(892)^0 l^+ \nu_l) \simeq 10\%$
- $\text{BR}(D^+ \rightarrow K^- \pi^+ \pi^+) \simeq 9\%$
- $\text{BR}(D^+ \rightarrow K_L^0 \pi^+) \simeq 1.5\%$

Fitting procedure



Statistical model



SR: charged B background dominant.

CR1: combined continuum dominant
($e^+e^- \rightarrow c\bar{c}$ largest contribution).

CR2 and CR3: dominant contribution from $e^+e^- \rightarrow c\bar{c}$.

Sample	Channels	Expected yields
signal	SR, CR1	14, 4
charged B	SR, CR1	263, 174
neutral B	SR, CR1	142, 102
$c\bar{c}$	SR, CR1, CR2, CR3	228, 249, 34, 38
$d\bar{d}$	SR, CR1, CR2, CR3	15, 16, 2, 3
$s\bar{s}$	SR, CR1, CR2, CR3	129, 121, 23, 17
$u\bar{u}$	SR, CR1, CR2, CR3	62, 57, 11, 8
$\tau^+\tau^-$	SR, CR1, CR2, CR3	3, 5, 1, 0

Systematic uncertainties



BR of the leading B -background decays

Only central values of the BR used in simulation → evaluate shape variations due to σ_{BR} (from PDG).

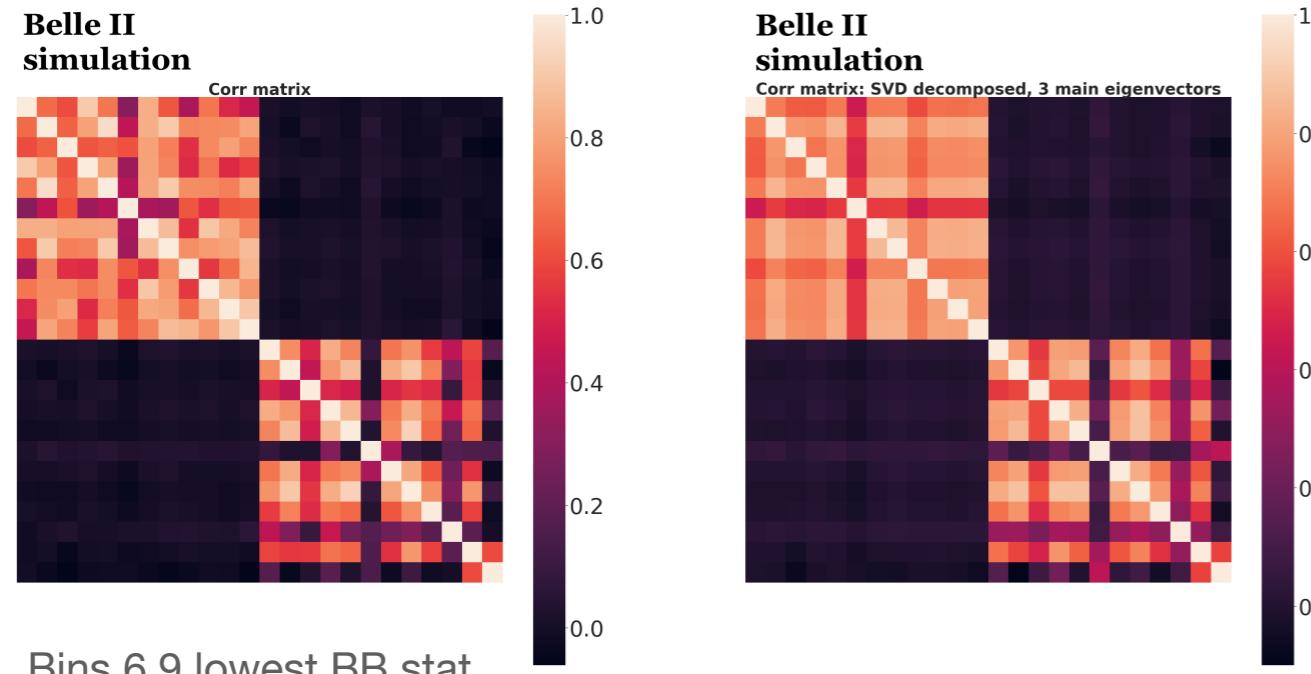
Samples: charged B , neutral B from 300 fb^{-1} sample of generic bkg.

Fractions of covered decays: $\sim 80\%$ of charged B , $\sim 62\%$ of neutral B .

1000 MC toys: weighted replicas of reconstructed B samples in the fit regions.

Event weights: $w_{\text{BR}} = (\text{BR} + \Delta)/\text{BR}$ with Δ from $N(0, \sigma_{\text{BR}})$

Compute bin-covariance matrix $\text{cov}(S_i, S_k)$ from samples of weighted bin counts $S_i = \{n_i\}_{j=1}^{1000}$ ($i = 1, \dots, 24$).



- Determine 3 main principal components.
 - Split among charged ($i = 1, \dots, 12$) and neutral ($i = 13, \dots, 24$) samples.
- ↓
- 3 orthogonal vectors** of systematic bin variations per sample, coupled to **3 nuisance parameters for correlated shape variation**.

Systematic uncertainties



SM form factor

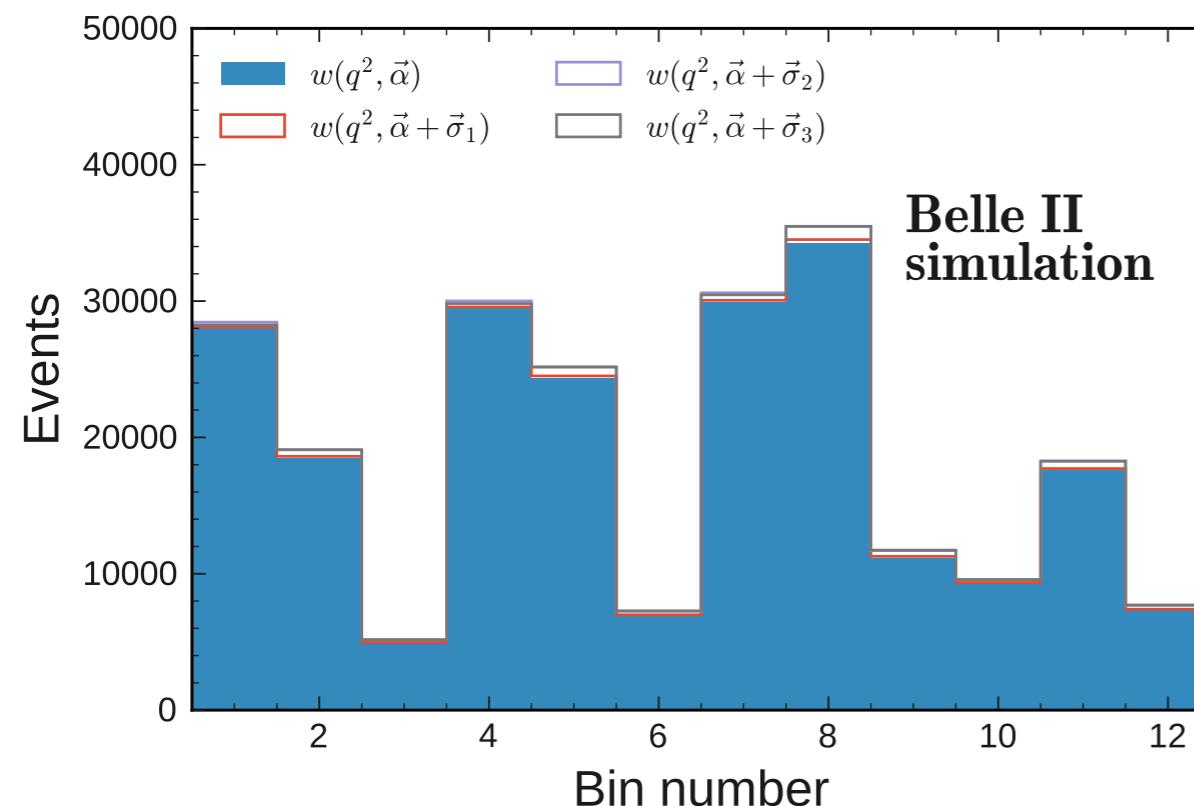
Depends on three real parameters estimated in fit to Lattice QCD and LCSR results:

- central values $\vec{\alpha} \equiv (\alpha_0, \alpha_1, \alpha_2)$;
- 3x3 correlation matrix $\text{corr}(\alpha_i, \alpha_j)$.

[J. High Energ. Phys. 2015, 184]

Decompose equivalent covariance matrix $\text{cov}(\alpha_i, \alpha_j)$ to the 3 principal components $\vec{\sigma}_i$ and derive

3 modified form factors based on $\vec{\alpha} + \vec{\sigma}_i$



- Compare the 3 modified signal distributions in CR1 + SR to the expected distribution to derive **3 independent vectors of bin-count variations**.
- Couple the 3 vectors to **3 nuisance parameters for correlated shape variation**.

Systematic uncertainties



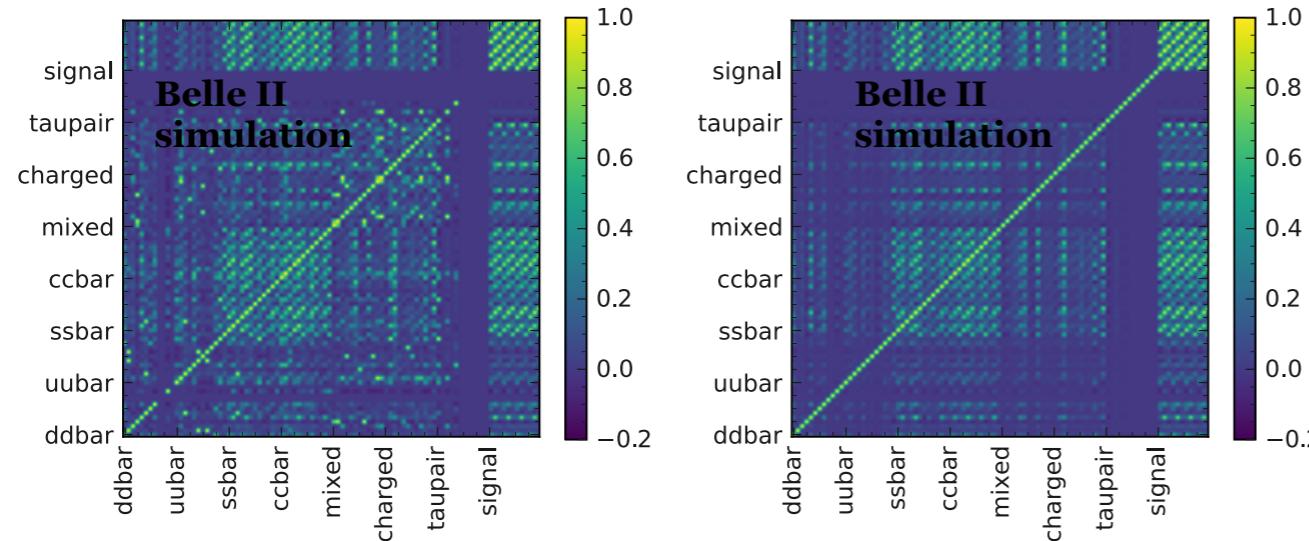
PID correction

Data/MC correction weights determined in **specific $(p, \cos\theta)$ bins** → related uncertainties can introduce correlated shape variation.

Samples: 300 fb^{-1} of generic bkg + 4×10^6 simulated signal events.

500 MC toys: weighted replicas of reconstructed samples in the fit regions.

- **In each replica: vary PID correction weights** in $(p, \cos\theta)$ bins (using lognormal with width equal to the weight uncertainty) **and compute weighted signal and bkg histograms.**
- Compute bin-covariance matrix $\text{cov}(S_{ij}, S_{mn})$ from samples of weighted bin counts $S_{ij} = \{n_{ij}\}_{k=1}^{500}$ ($i \in \text{samples}, j \in \text{bins}$)



- Determine 3 main principal components and split them among samples.



3 orthogonal vectors of systematic bin variations per sample, coupled to **3 nuisance parameters for correlated shape variation.**

Systematic uncertainties

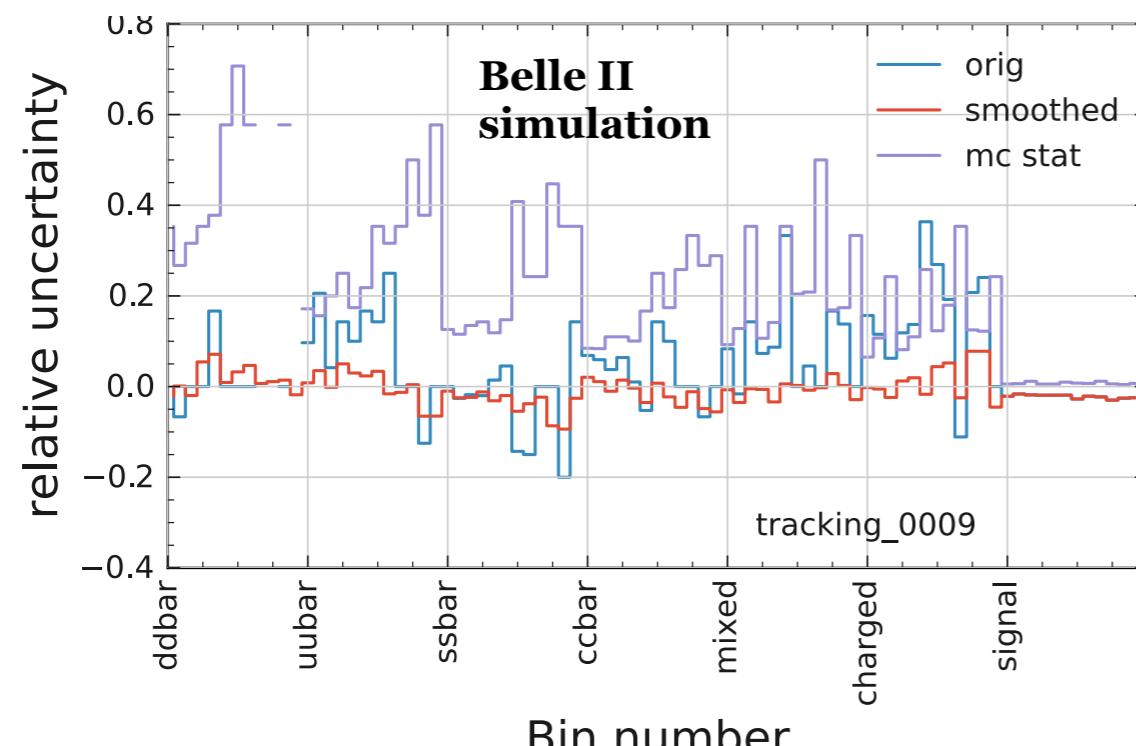


Tracking efficiency

Estimated uncertainty is 0.9% [BELLE2-NOTE-PL-2020-014] → simulate 0.9% probability to lose a track.

Samples: 100 fb^{-1} of generic bkg + 4×10^6 simulated signal events.

- Compute signal and bkg histograms in the fit regions with **tracking inefficiency turned on**.
- **Smooth reference and modified bkg histograms** using Gaussian KDE:
avoid large fluctuations due to limited size of bkg samples.
- Evaluate **relative systematic uncertainties** in the bins as **modified/reference counts - 1**.



1 vector of systematic bin variations per sample, coupled to **1 common nuisance parameter for correlated shape variation**.

$\tau^+\tau^-$ excluded: minor bkg.

Systematic uncertainties

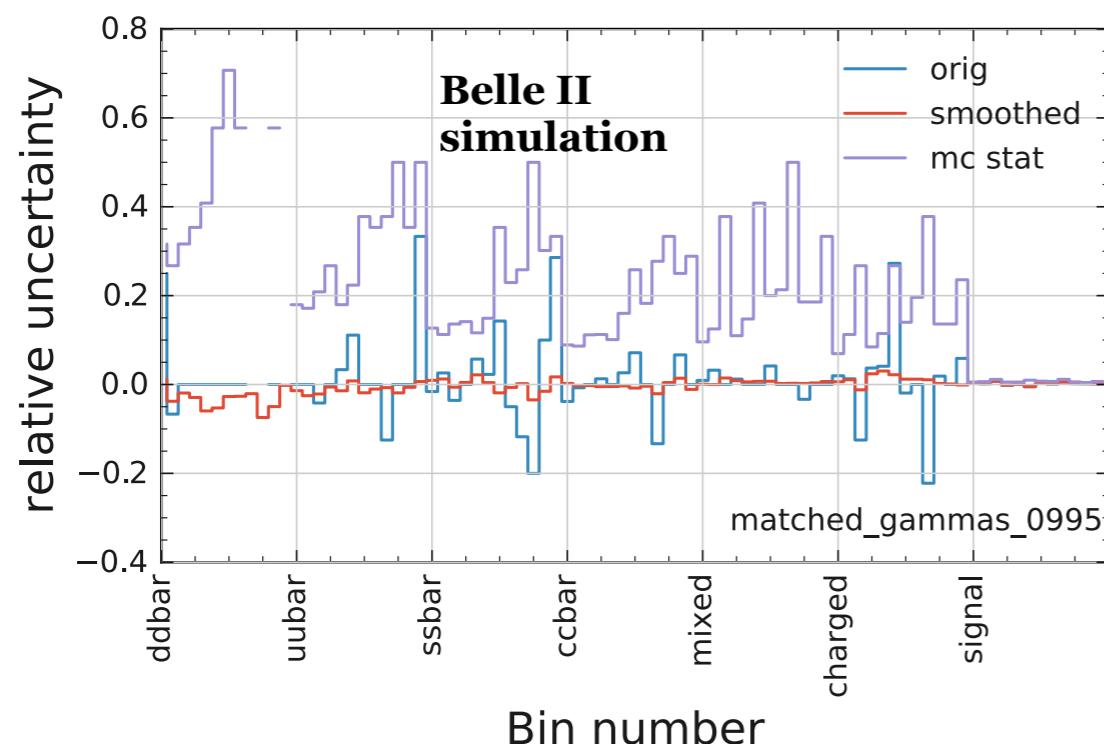


Energy calibration of photon clusters

Estimated energy uncertainty is 0.5% [BELLE2-NOTE-PL-2020-014] → scale down by 0.5% the energy of ECL clusters matched to photons in simulation.

Samples: 100 fb^{-1} of generic bkg + 4×10^6 simulated signal events.

- Compute signal and bkg histograms in the fit regions with **photon energy scaled down**.
- **Smooth reference and modified bkg histograms** using Gaussian KDE:
avoid large fluctuations due to limited size of bkg samples.
- Evaluate **relative systematic uncertainties** in the bins as **modified/reference counts - 1**.



1 vector of systematic bin variations per sample, coupled to **1 common nuisance parameter for correlated shape variation**.

$\tau^+\tau^-$ excluded: minor bkg.

Systematic uncertainties



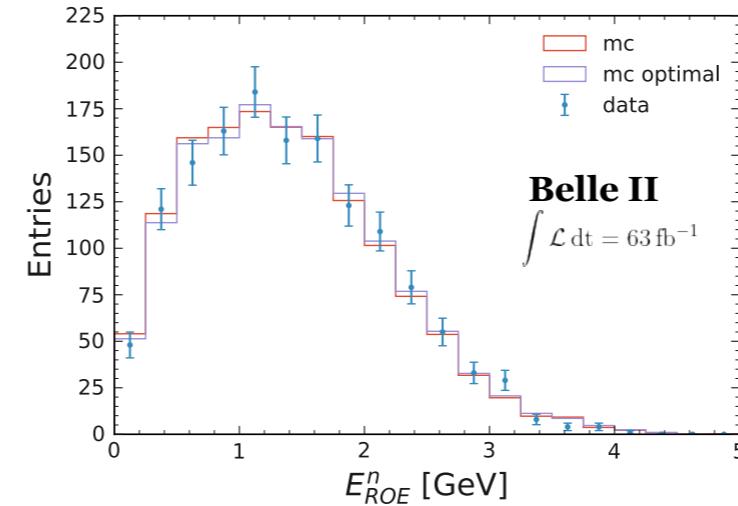
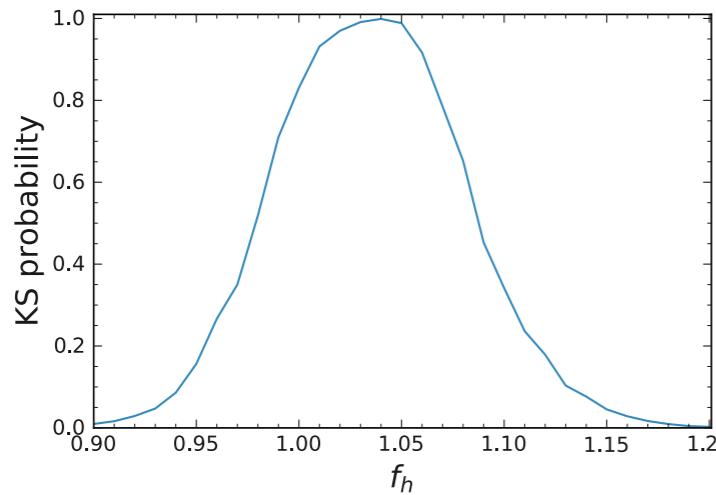
Energy calibration of ECL clusters not matched to photons

- Clusters mostly produced by **neutral hadrons**.
- Uncertainty can be $> 0.5\%$ → use $B^+ \rightarrow K^+ J/\psi \rightarrow \mu^+ \mu^-$ to estimate it.

Decompose ROE energy in the assumption of accurate determination of photon energy:

$$E_{\text{ROE}}^n(f_h) = \sum_i E_i^\gamma + f_h \sum_j E_j^n$$

- $i \in$ ECL clusters matched to photons.
- $j \in$ ECL clusters **not matched** to photons.
- $f_h \equiv$ scale factor quantifying accuracy of energy calibration.



Max K-S probability for $f_h \simeq 1.05$, but FWHM $\sim 10\%$ → assume 10% systematic uncertainty on the energy of the unmatched clusters.

Systematic uncertainties

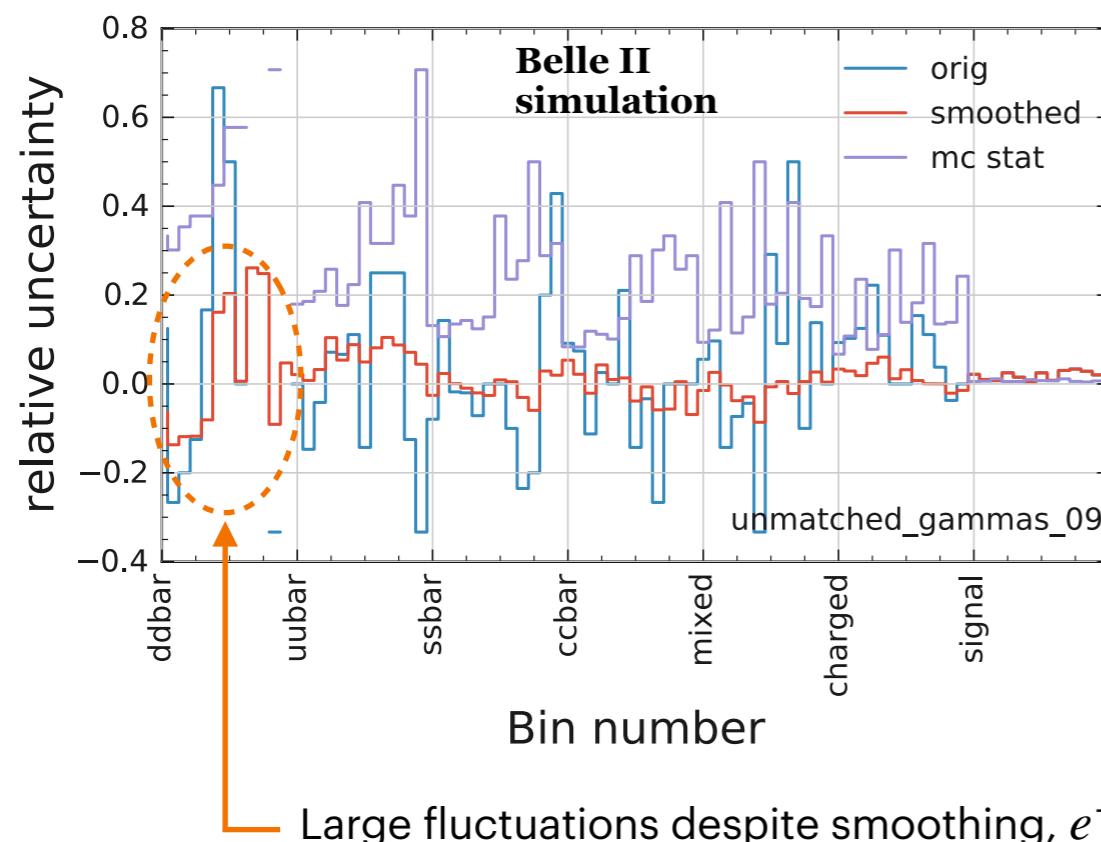


Energy calibration of ECL clusters not matched to photons

Scale down by 10% the energy of ECL clusters not matched to photons in simulation.

Samples: 100 fb^{-1} of generic bkg + 4×10^6 simulated signal events.

- Compute signal and bkg histograms in the fit regions with **energy of unmatched clusters scaled down**.
- **Smooth reference and modified bkg histograms** using Gaussian KDE:
avoid large fluctuations due to limited size of bkg samples.
- Evaluate **relative systematic uncertainties** in the bins as **modified/reference counts - 1**.



1 vector of systematic bin variations per sample, coupled to **1 common nuisance parameter for correlated shape variation**.

$\tau^+\tau^-$ excluded: minor bkg.

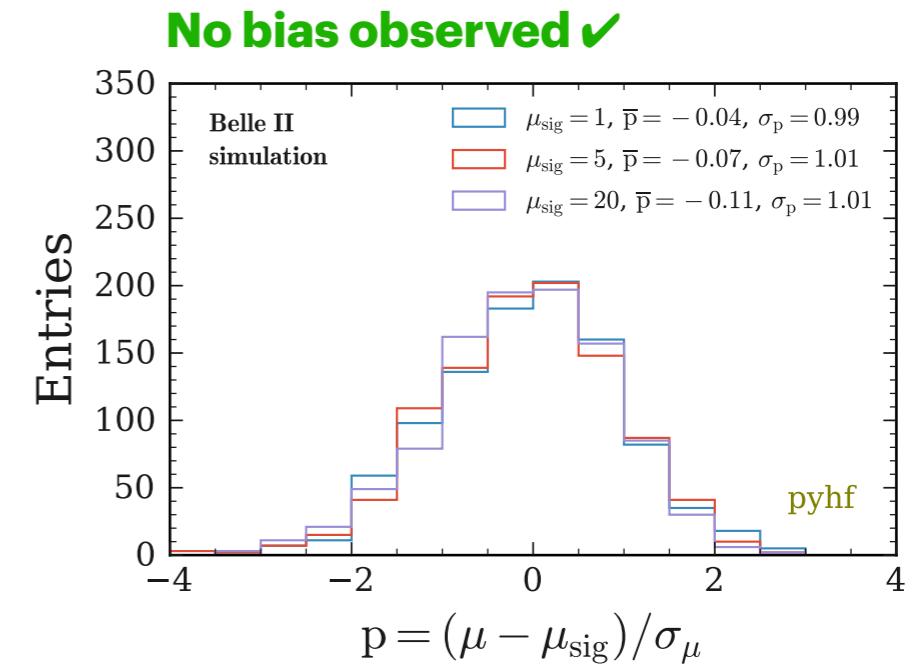
Fit validation



Signal-injection study

Verify no bias in signal extraction

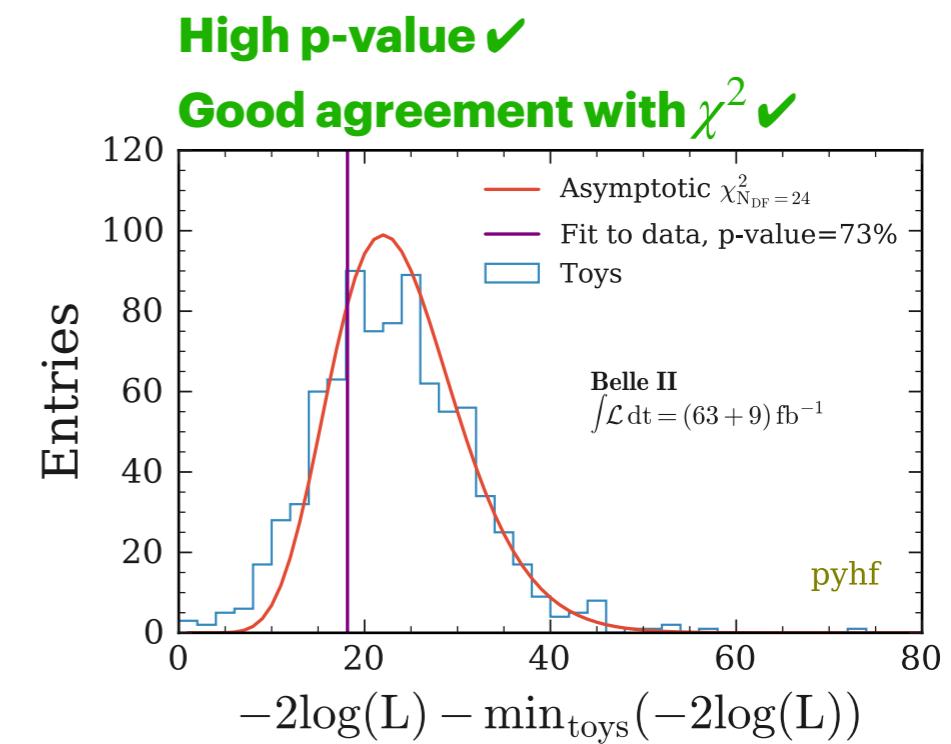
- 3 signal strength hypotheses tested: $\mu_{\text{sig}} = 1.0, 5.0, 20.0$
- 3 ensembles of 1k MC toys:
 - Poisson fluctuations of expected counts;
 - Gaussian systematic fluctuations.



Test of fit quality

Blinded test of data-model compatibility: p-value of fit to data.

- Fits to toys of pseudo-observations centred on expectations.
- Ensemble of MC toys:
 - $n_{\text{toy}} = n_{\text{exp}} + \Delta n_{\text{obs}}$ where $\Delta n_{\text{obs}} = n_{\text{obs}} - n'_{\text{obs}}$ with n'_{obs} from Poisson fluctuations.
 - Gaussian systematic fluctuations.



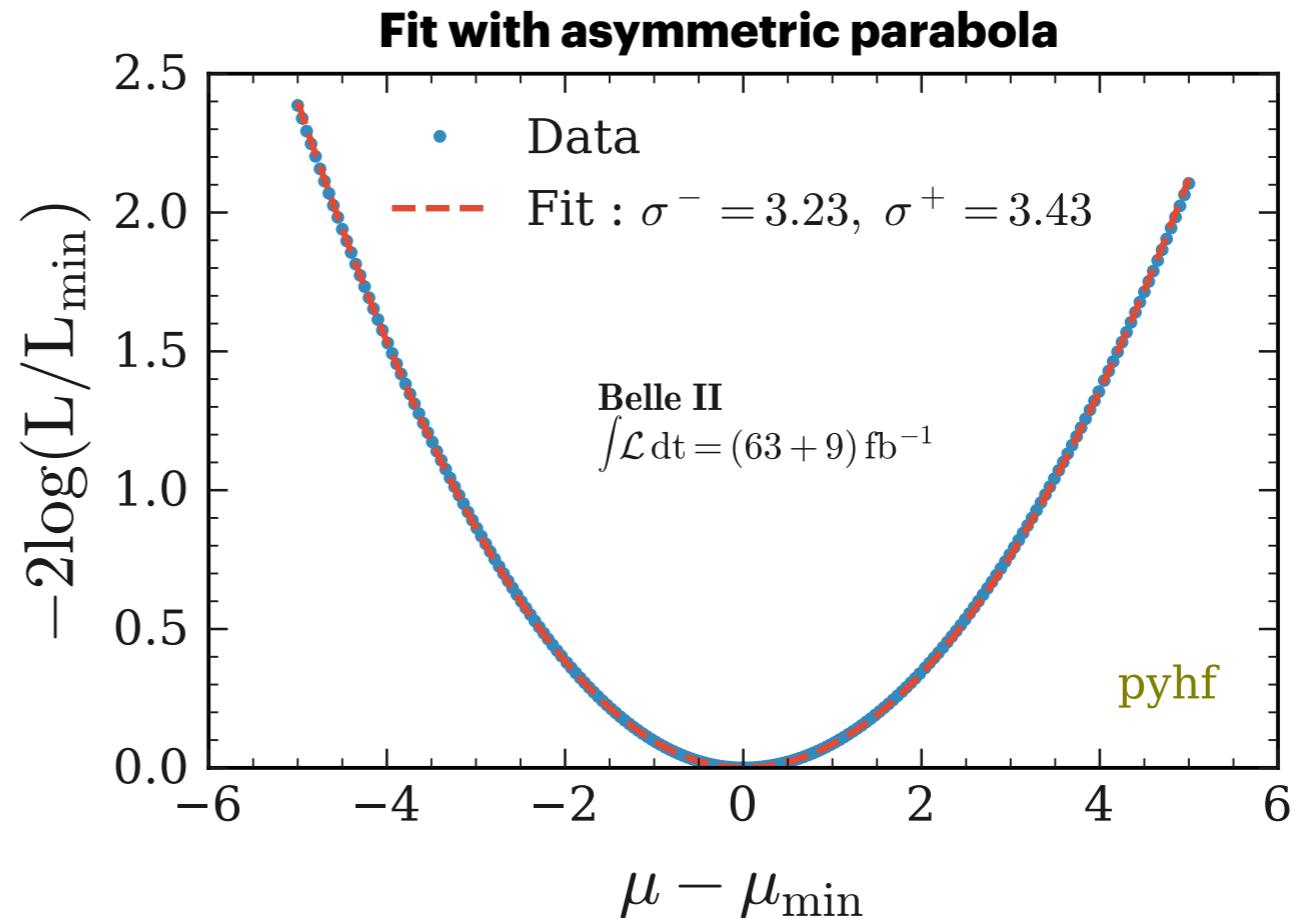
Fit results



Profile likelihood scan

Wilk's theorem:

$$-2\ln\lambda_p(\mu) = \frac{(\mu - \hat{\mu})^2}{\sigma_\mu^2} + \mathcal{O}(1/\sqrt{N})$$

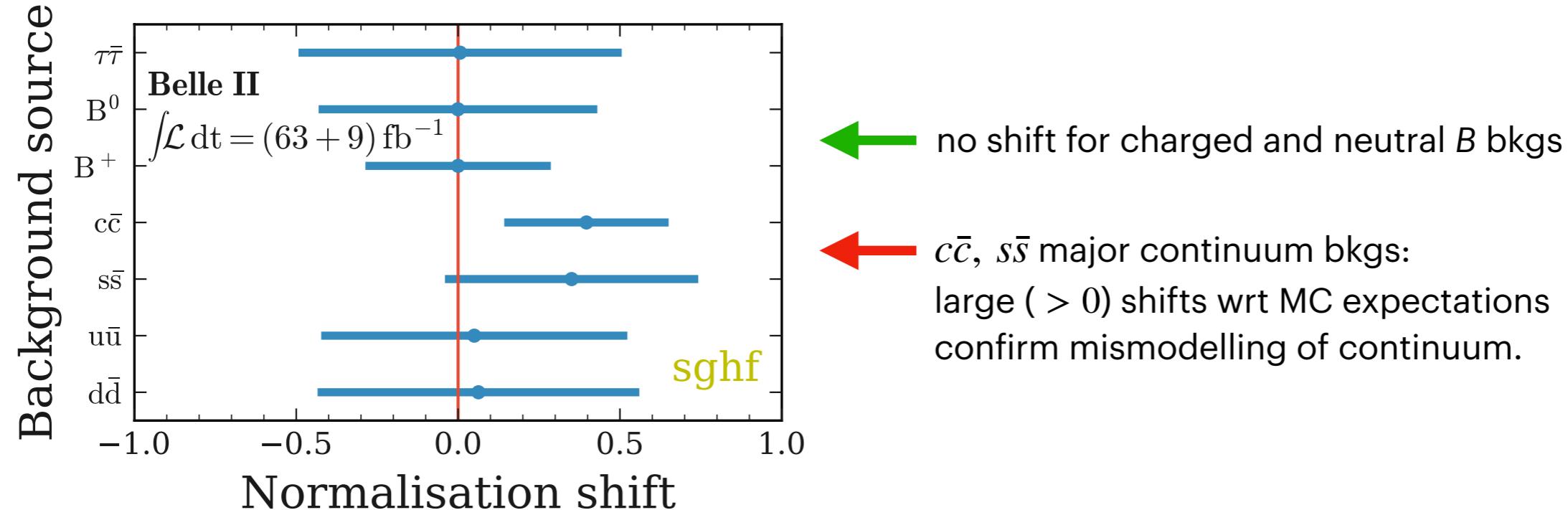


Fit results



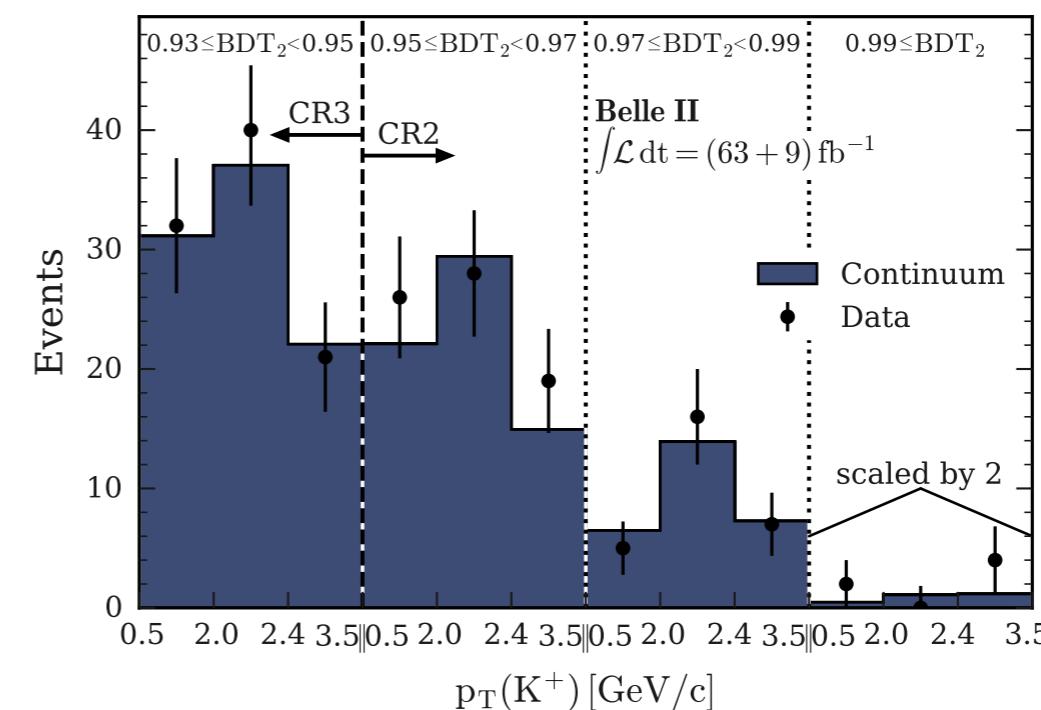
Post-fit shifts

Normalisation modifiers for the 7 background categories



Post-fit yields

Combined continuum vs off-resonance data



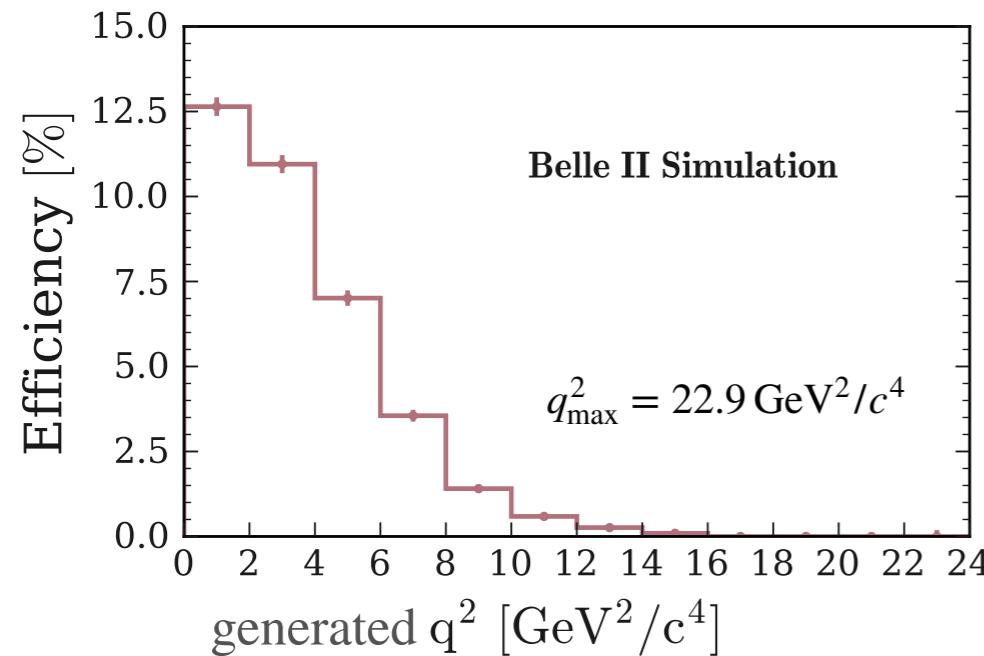
Discussion of the results



Performance of the inclusive tagging

[123: Phys. Rev. D 2013, 87, 111103]
[124: Phys. Rev. D 2013, 87, 112005]
[126: Phys. Rev. D 2017, 87, 112005]

Signal efficiency in SR:



Performance of the inclusive tagging in terms of precision (assuming $\sigma_{\text{BR}} \sim 1/\sqrt{L}$)

Measurement	$L (\text{fb}^{-1})$	$\sigma_{\text{BR}} \times 10^5$	$\sigma_{\text{scaled}} \times 10^5$
BaBar, hadronic and semileptonic taggings [124]	429	0.65	1.70
Belle, hadronic tagging [123]	711	1.63	5.47
Belle, semileptonic tagging [126]	711	0.57	1.91
Belle II, inclusive tagging (this work)	63	1.55	1.55

Inclusive tagging performs:

- $2.5 \times$ better than Belle had.
- 20 % better than Belle SL.
- 10 % better than BaBar had. + SL

Outlook

Projections of $\sigma_{\text{BR}}/\text{BR}_{\text{SM}}$ in the assumption that $\sigma_{\text{BR}} = 1.55 \times 10^{-5}$ scales as $1/\sqrt{L}$:

$L(\text{ab}^{-1})$	Baseline scenario	Improved scenario
1	0.85	0.69
5	0.38	0.31
10	0.27	0.22
50	0.12	0.10

- $L = 10 \text{ ab}^{-1}$ sensitivity to SM signal rate at 3σ (5σ) level in the baseline (improved) scenario.
- SM accuracy reachable with 50 ab^{-1} .

Baseline = no improvements

Improved = 50% increase in signal eff. at same bkg level



Search for $B^0 \rightarrow K^{*0} \nu \bar{\nu}$

with an inclusive tagging

at Belle II

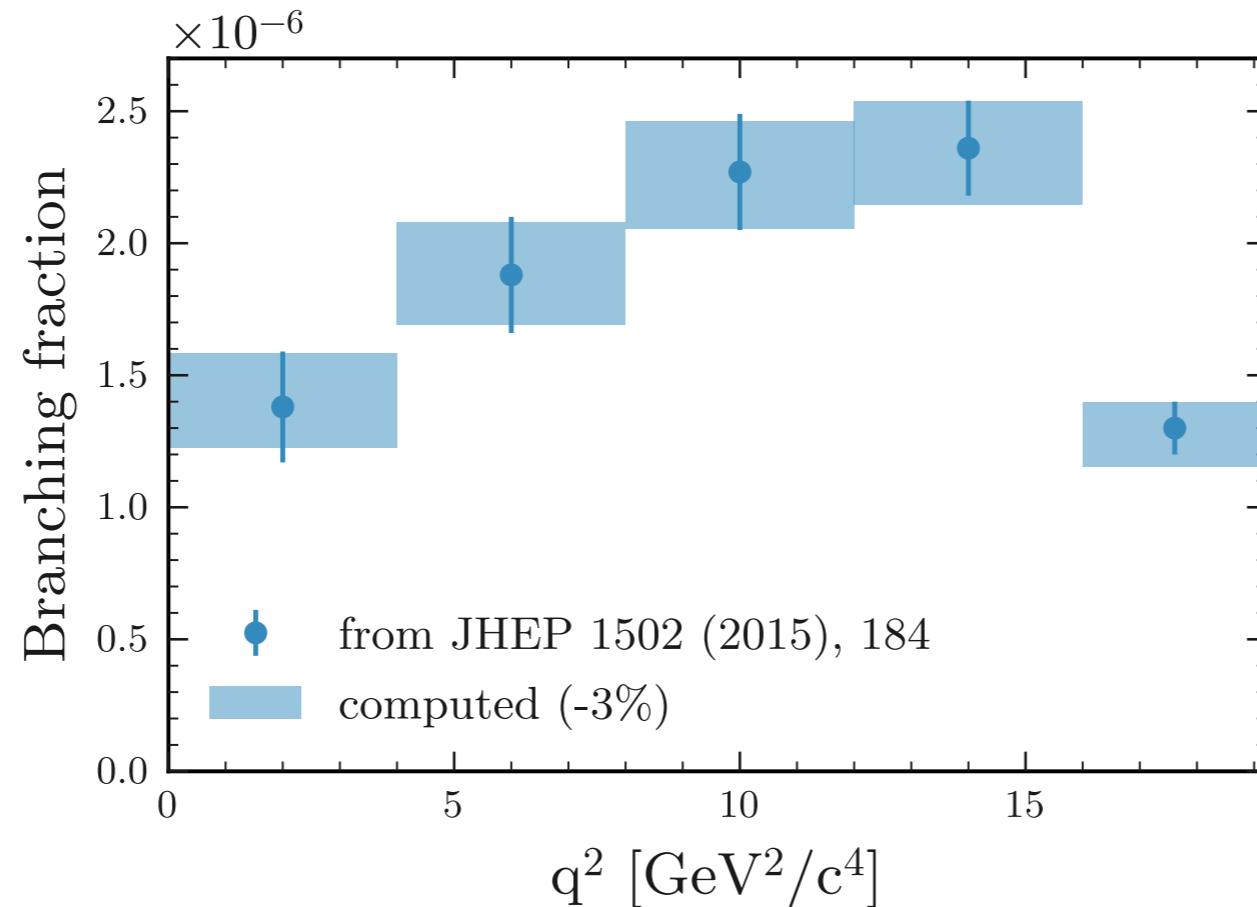
Simulated signal



Reweighting procedure

Validation

Computation based on FFs estimate from fit to Lattice QCD and LCSR results in [J. High Energ. Phys. 2016, 8](#) and input parameters from [J. High Energ. Phys. 2015, 184](#)

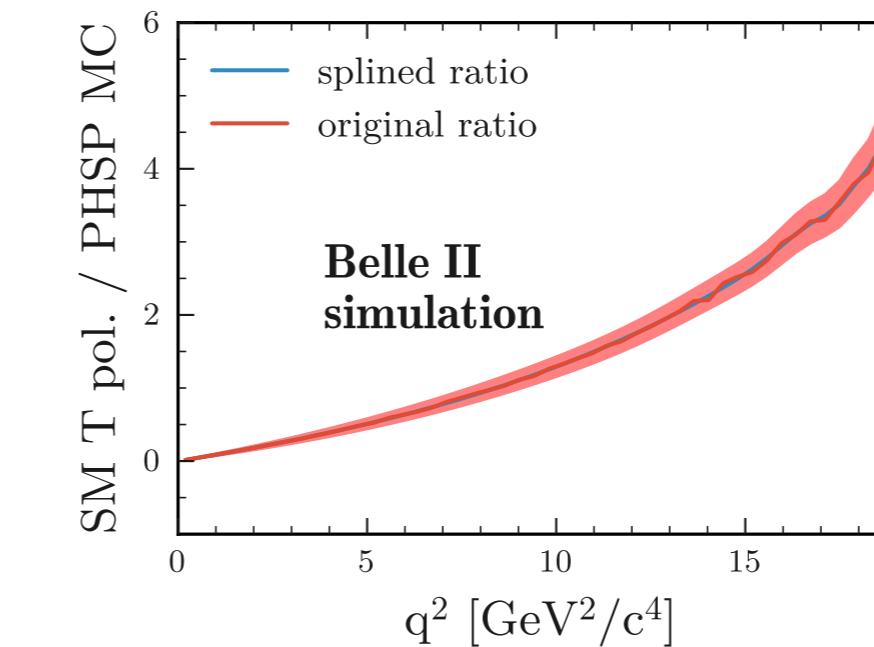
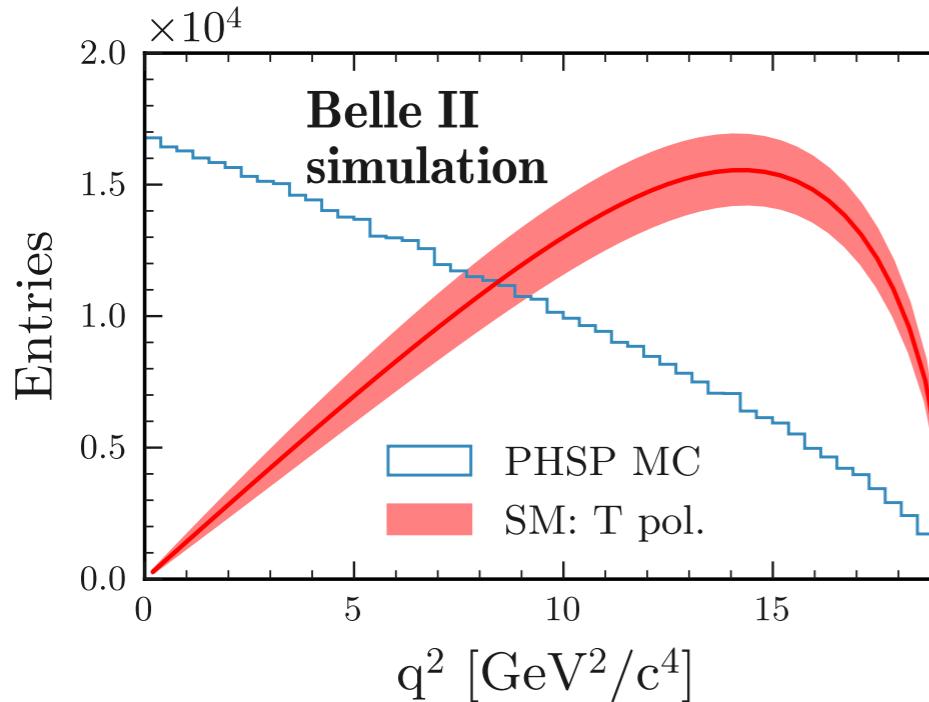
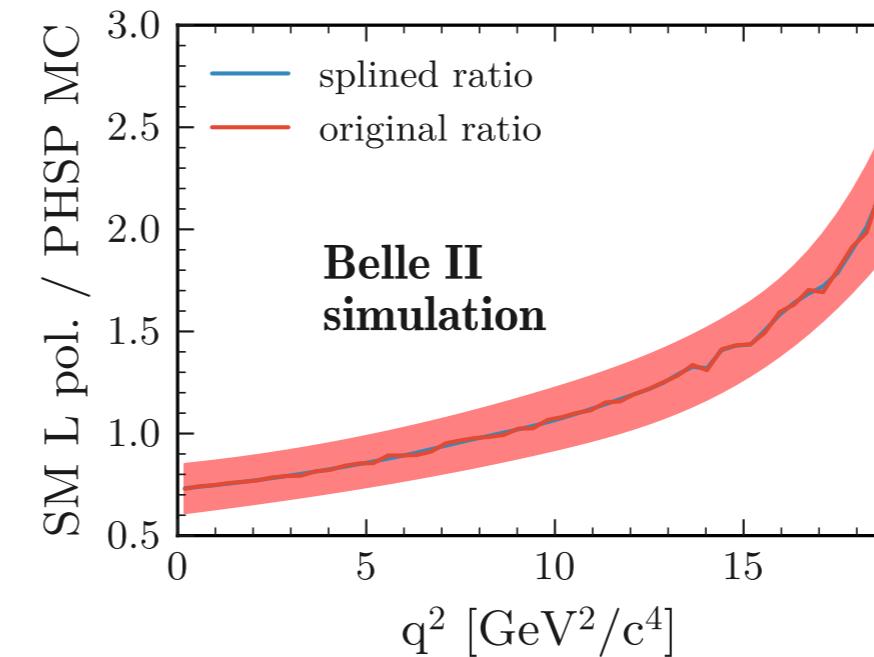
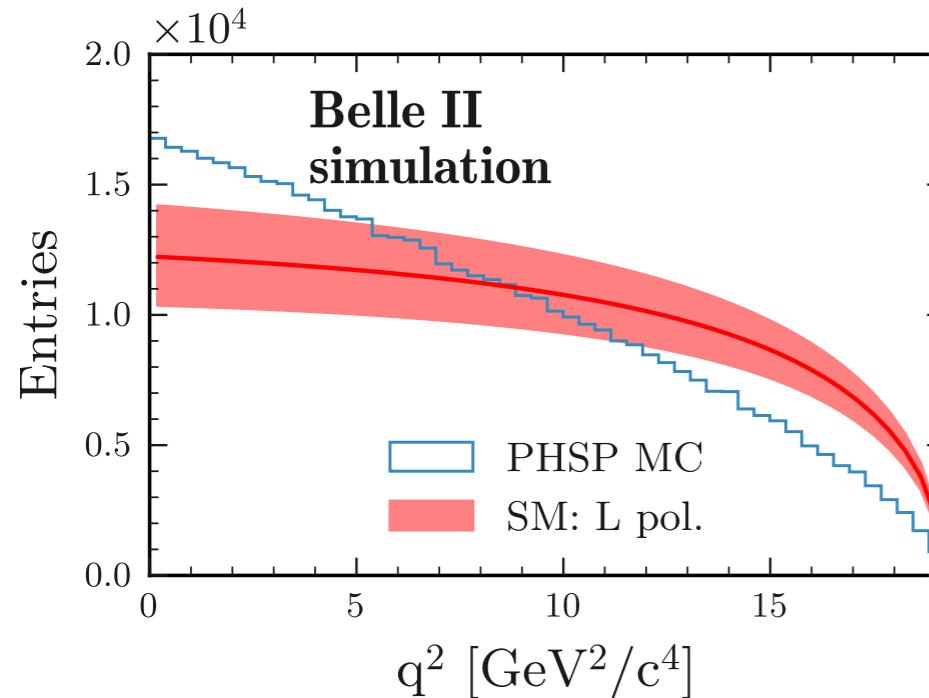




Simulated signal

Reweighting procedure

Computation of the weights:



Spline interpolation: piecewise polynomial continuous over equispaced points.

Prevents oscillation at edges of the interval at higher degrees.

$$\text{spl}_L(q^2)$$

$$\text{spl}_T(q^2)$$

Simulated signal

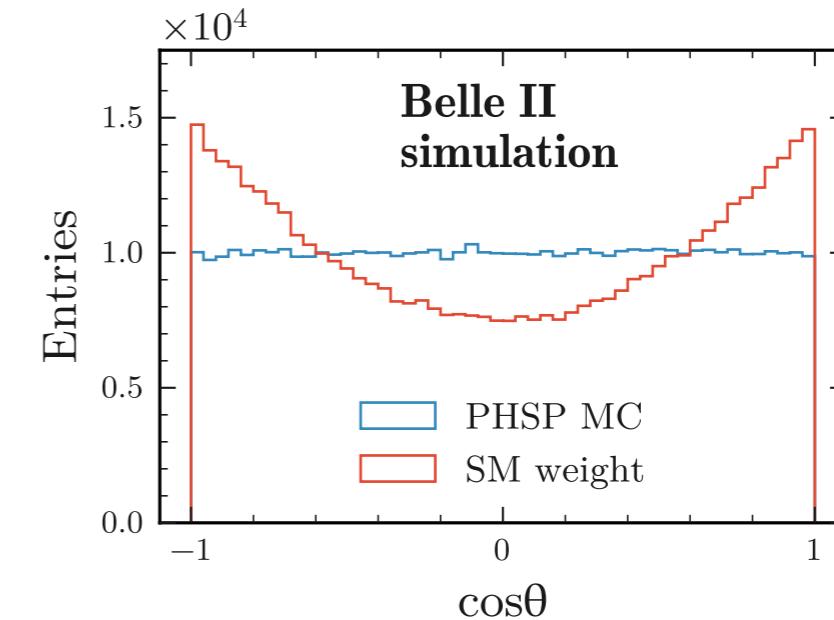
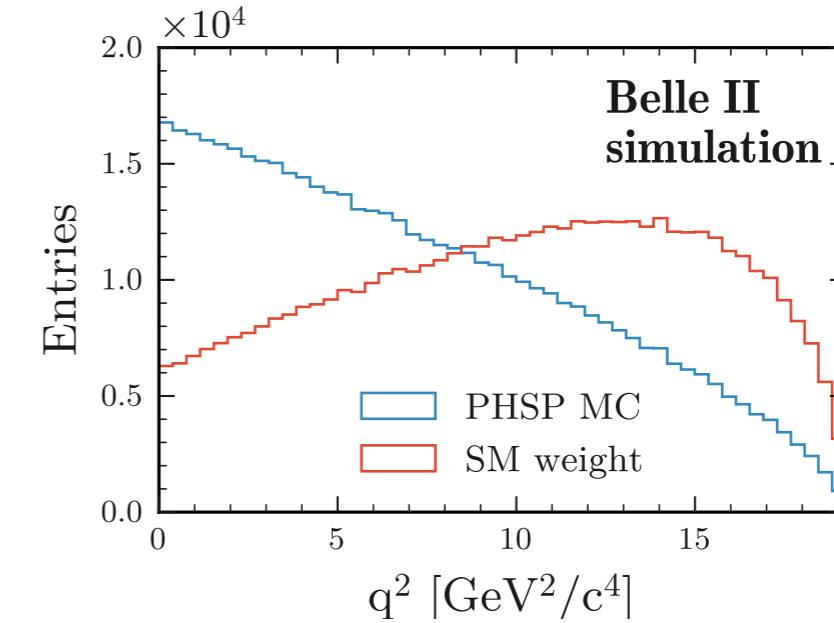
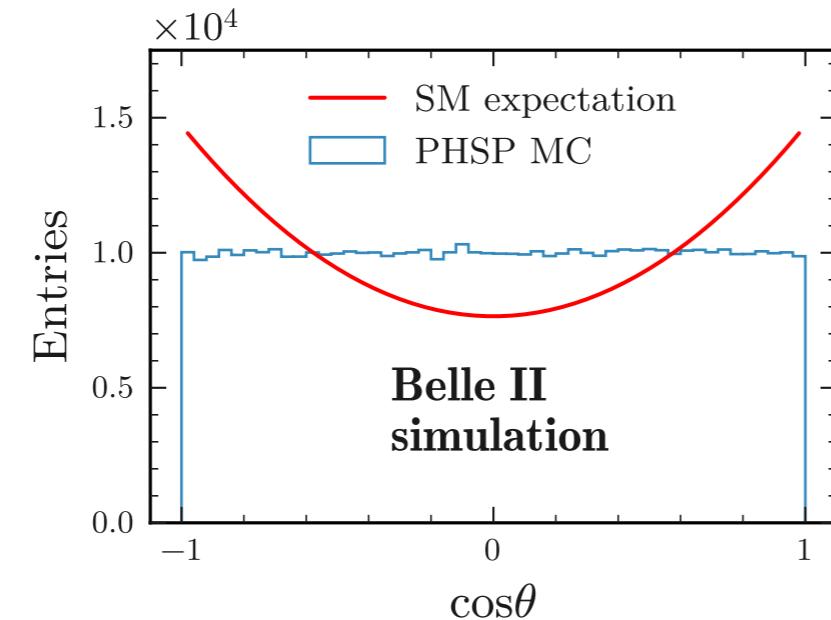
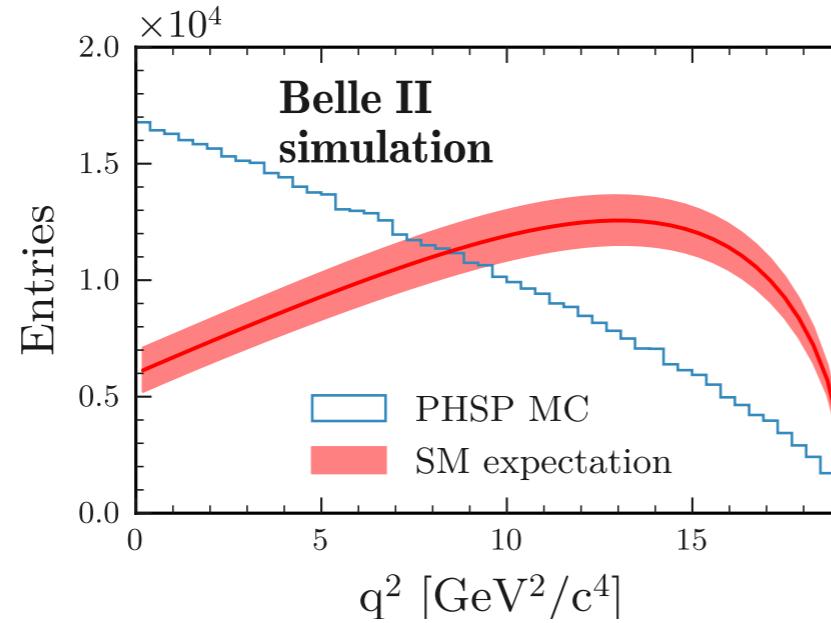


Reweighting procedure

Weights:

$$w(q^2, \cos\theta) = \frac{3}{4} \text{spl}_T(q^2)(1 - \cos^2\theta) + \frac{3}{2} \text{spl}_L(q^2)\cos^2\theta$$

Closure test:



$B^0 \rightarrow K^{*0} \nu \bar{\nu}$ reconstruction



Object selection

- **Track cleanup:**

$p_T > 0.1 \text{ GeV}/c,$
 $E < 5.5 \text{ GeV},$
 $|dz| < 3 \text{ cm},$
 $dr < 0.5 \text{ cm},$
 $17^\circ < \theta < 150^\circ,$

- **Photon cleanup:**

$0.1 \text{ GeV} < E_{\text{ECL}} < 5.5 \text{ GeV},$
 $17^\circ < \theta < 150^\circ,$

Event selection

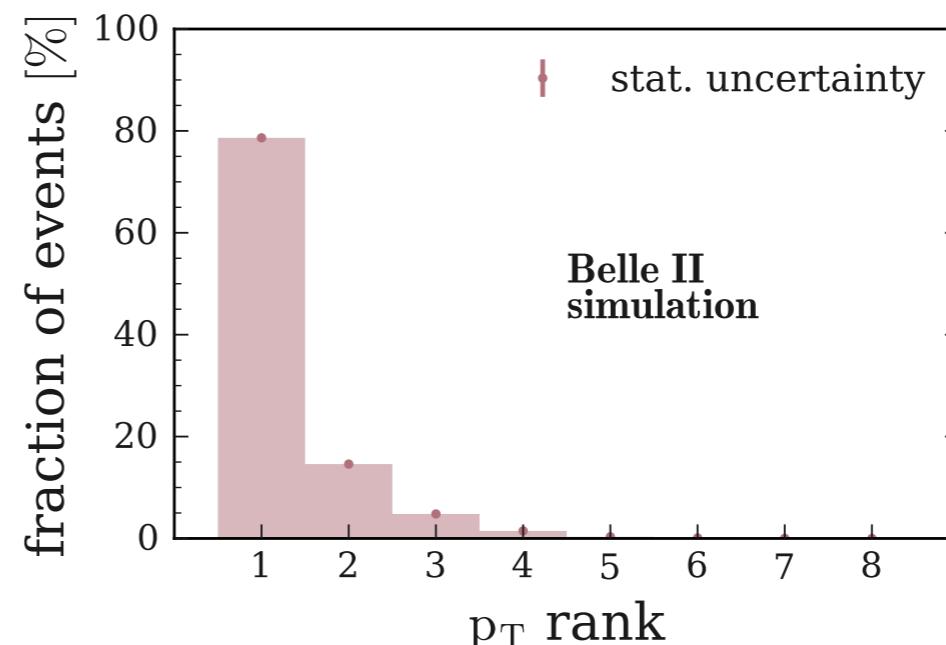
- $4 < \text{nTracks} < 11$
- $17^\circ < \theta(\vec{p}_{\text{miss.}}) < 160^\circ$
- $E_{\text{CMS}} > 4.0 \text{ GeV}$

Good suppression of low-multiplicity background like $\gamma\gamma \rightarrow 4\pi$.

Signal selection

Cut in the signal candidate selection	Signal efficiency (%)
$K^+ (\text{nPXDHits} > 0 \text{ and } \text{nCDCHits} > 20) \text{ and } \pi^- (\text{nPXDHits} > 0)$	39.4 ± 0.4
$K^+ (\text{kaonID} > 0.9)$	28.4 ± 0.4
$0.8 \text{ GeV}/c^2 < M(K^{*0}) < 1.0 \text{ GeV}/c^2$	25.2 ± 0.3
$p_{\text{fit}} > 0.1\%$	25.2 ± 0.3
$dr_{K^{*0}} < 0.1 \text{ cm}$	25.1 ± 0.3

Signal K^{*0} candidate matched to a true signal K^{*0} in the acceptance as a function of q_{rec}^2 rank



Selection stage	Signal efficiency (%)
Object selection	43.0 ± 0.4
Signal selection	21.8 ± 0.3
Event selection	19.3 ± 0.3

$B^0 \rightarrow K^{*0} \nu \bar{\nu}$ reconstruction



Object selection

- **Track cleanup:**

$p_T > 0.1 \text{ GeV}/c,$

$E < 5.5 \text{ GeV},$

$|dz| < 3 \text{ cm},$

$dr < 0.5 \text{ cm},$

$17^\circ < \theta < 150^\circ,$

- **Photon cleanup:**

$0.1 \text{ GeV} < E_{\text{ECL}} < 5.5 \text{ GeV},$

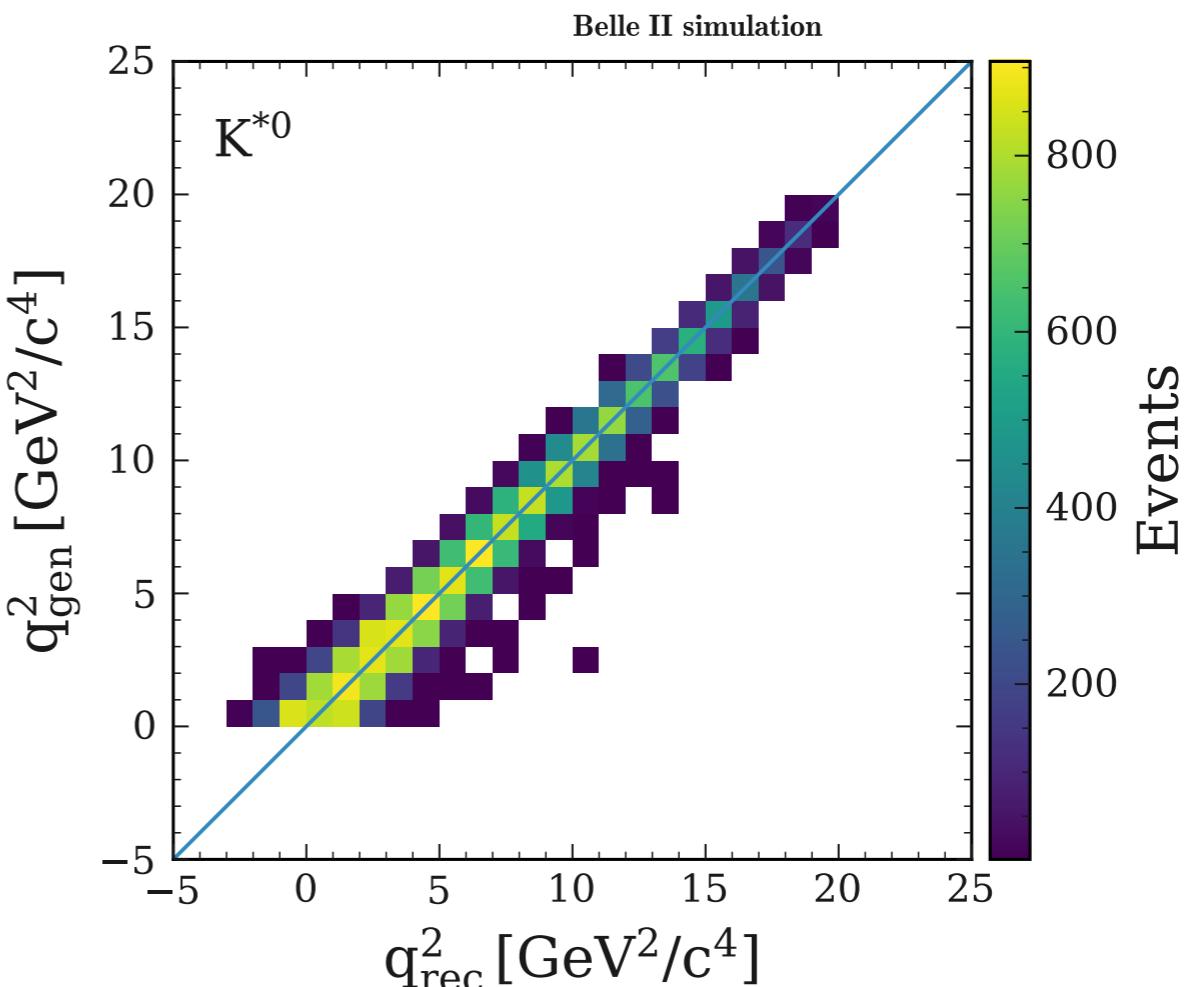
$17^\circ < \theta < 150^\circ,$

Event selection

- $4 < \text{nTracks} < 11$
- $17^\circ < \theta(\vec{p}_{\text{miss.}}) < 160^\circ$
- $E_{\text{CMS}} > 4.0 \text{ GeV}$

Good suppression of low-multiplicity background like $\gamma\gamma \rightarrow 4\pi$.

Signal selection



$$q_{\text{rec}}^2 = M_{B^0}^2 + M_{K^{*0}}^2 - 2M_{B^0} E_{K^{*0}}$$

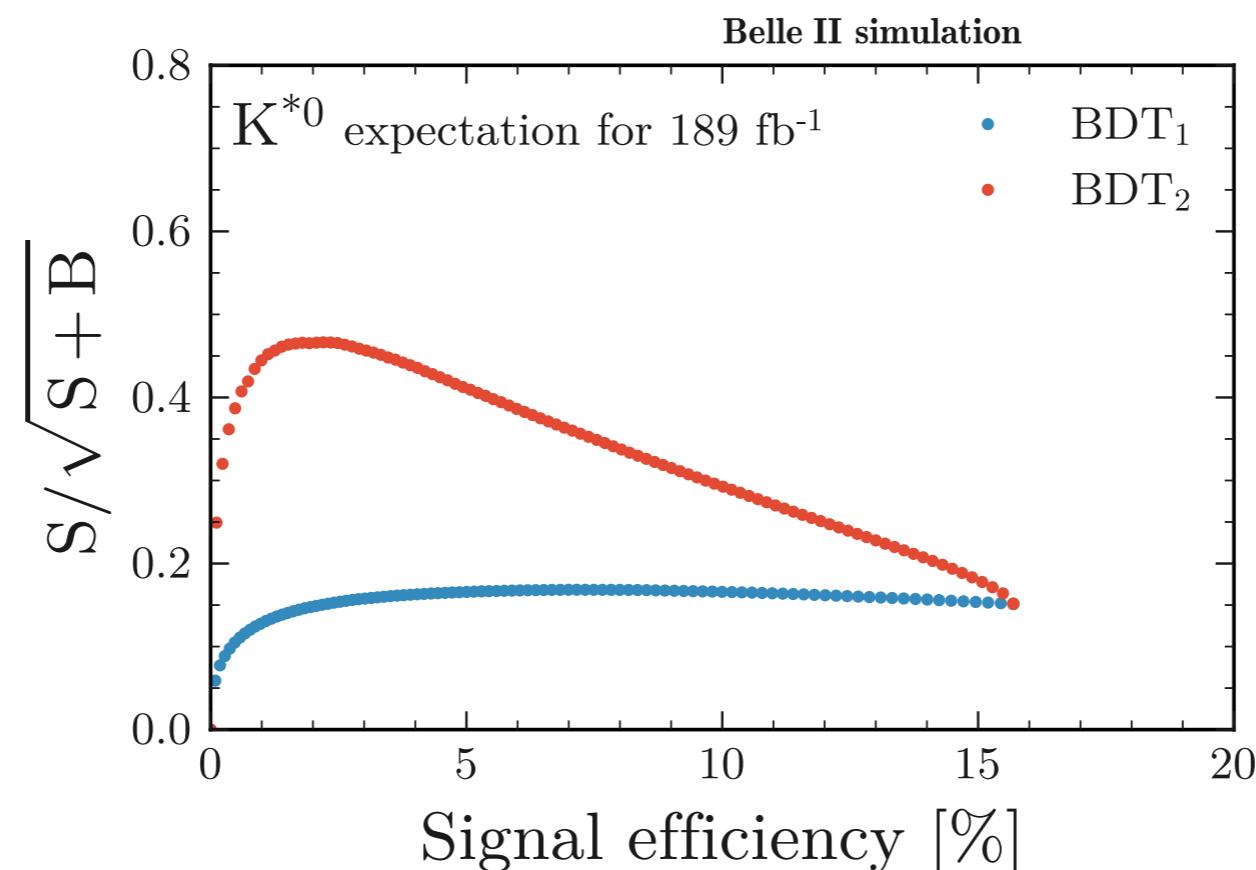
Assume B^0 at rest in CMS: $\sqrt{s} - 2M_B c^2 \simeq 20 \text{ MeV}$

Background suppression



Multivariate classification

- BDT_1 for initial background suppression:
 - **12** training variables: **event-shape** and **ROE kinematics**.
- BDT_2 to improve background suppression:
 - **37** training variables: K^{*0} **specific**, event-based, ROE features, combined signal and ROE.
 - Trained on reconstructed events with $\text{BDT}_1 > 0.9$ in simulated signal and bkg samples.
- **Performance**

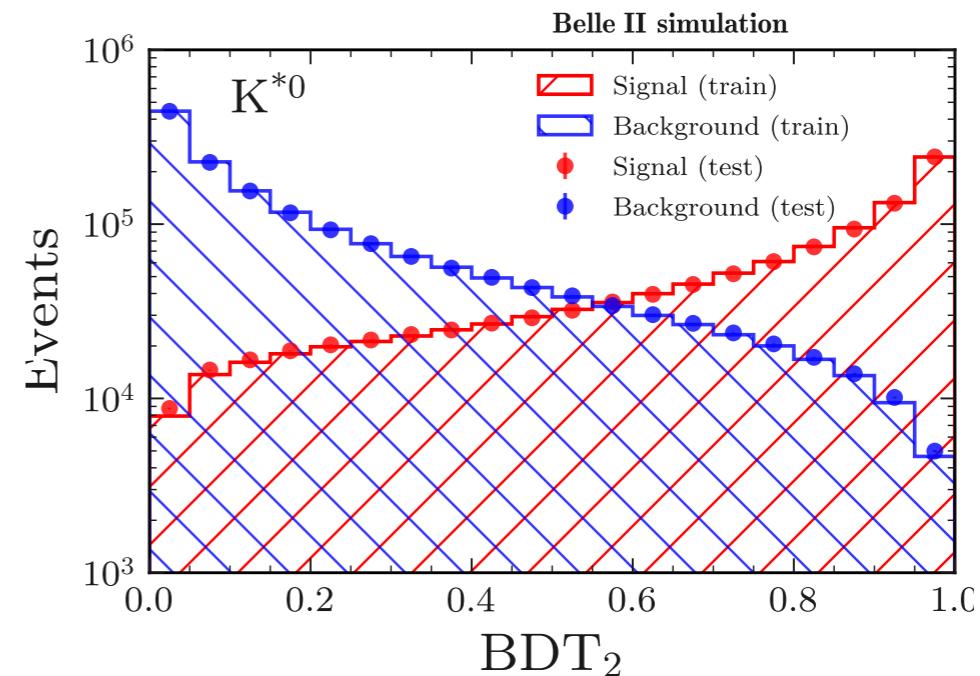
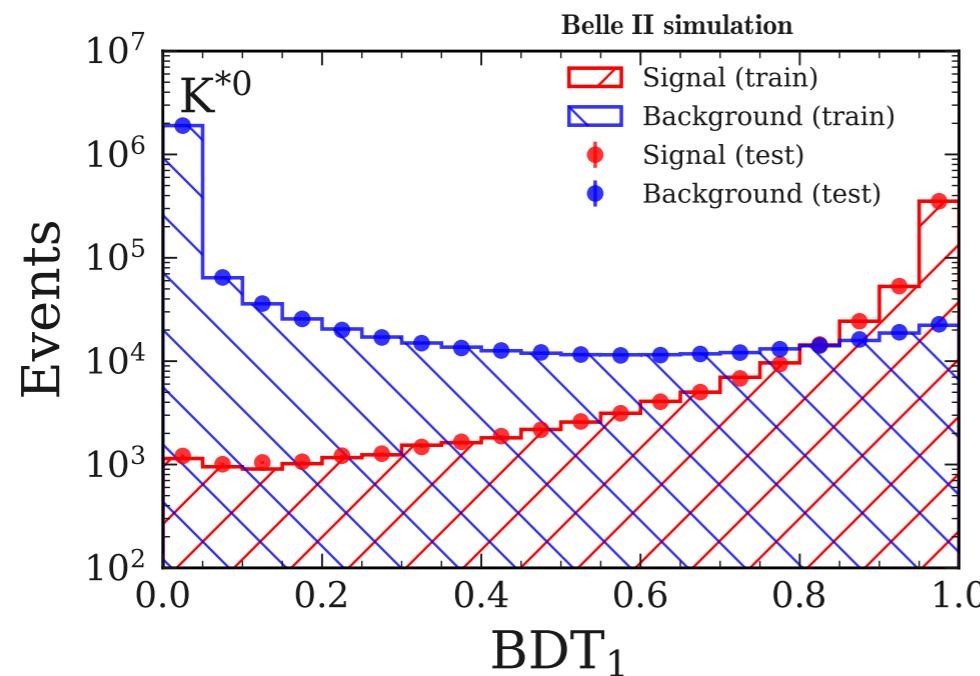


Background suppression



Multivariate classification

- BDT₁ for initial background suppression:
 - **12** training variables: **event-shape** and **ROE kinematics**.
- BDT₂ to improve background suppression:
 - **37** training variables: K^{*0} **specific**, event-based, ROE features, combined signal and ROE.
 - Trained on [reconstructed events with BDT₁ > 0.9](#) in simulated signal and bkg samples.
- **Performance**



Multivariate classification



- **BDT₁ [FastBDT] and BDT₂ [XGBoost] models**

Hyperparameter	Value
Number of trees	2000
Tree depth	2 for BDT ₁ , 3 for BDT ₂
Shrinkage	0.2
Sampling rate	0.5
Number of equal-frequency bins	256

- **Feature importances:**

- BDT₁

Training variable	Importance score
ΔE of the ROE	0.350
First normalised Fox-Wolfram moment R_1 computed in the CMS	0.210
Linear KSFW moment $H_{m,2}^{so}$ computed in the CMS	0.143
Harmonic moment B_0 computed in the CMS	0.060
Magnitude of the total ROE momentum	0.052
Quadratic KSFW moment R_0^{oo} computed in the CMS	0.046
Linear KSFW moment $H_{m,4}^{so}$ computed in the CMS	0.039
Harmonic moment B_2 computed in the CMS	0.034
Quadratic KSFW moment R_2^{oo} computed in the CMS	0.027
Cosine of the angle between the signal- B thrust axis and the ROE thrust axis	0.018
Cosine of the polar-angle thrust-axis	0.010
Polar angle of the total ROE momentum	0.009

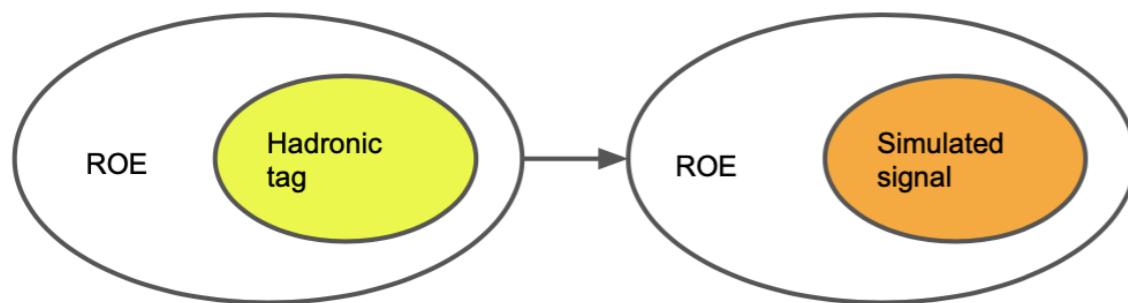
- BDT₂

Training variable	Importance score
Invariant mass of the signal K^{*0} candidate	0.236
χ^2 -probability of the vertex fit to the best D^+ candidate	0.083
dr of the signal K^{*0} candidate	0.067
Number of photon candidates in the event	0.049
Linear Fox-Wolfram moment $H_{n,2}^{so}$ computed in the CMS	0.046
Polar angle of the missing momentum vector \vec{p}_{miss}	0.043
Longitudinal distance between the POCA of the track pair of the best D^0	0.037
Linear Fox-Wolfram moment $H_{c,2}^{so}$ computed in the CMS	0.035
Number of charged lepton candidates in the event	0.034
Longitudinal distance between the fitted ROE vertex and the signal K^{*0} vertex	0.033

Signal embedding

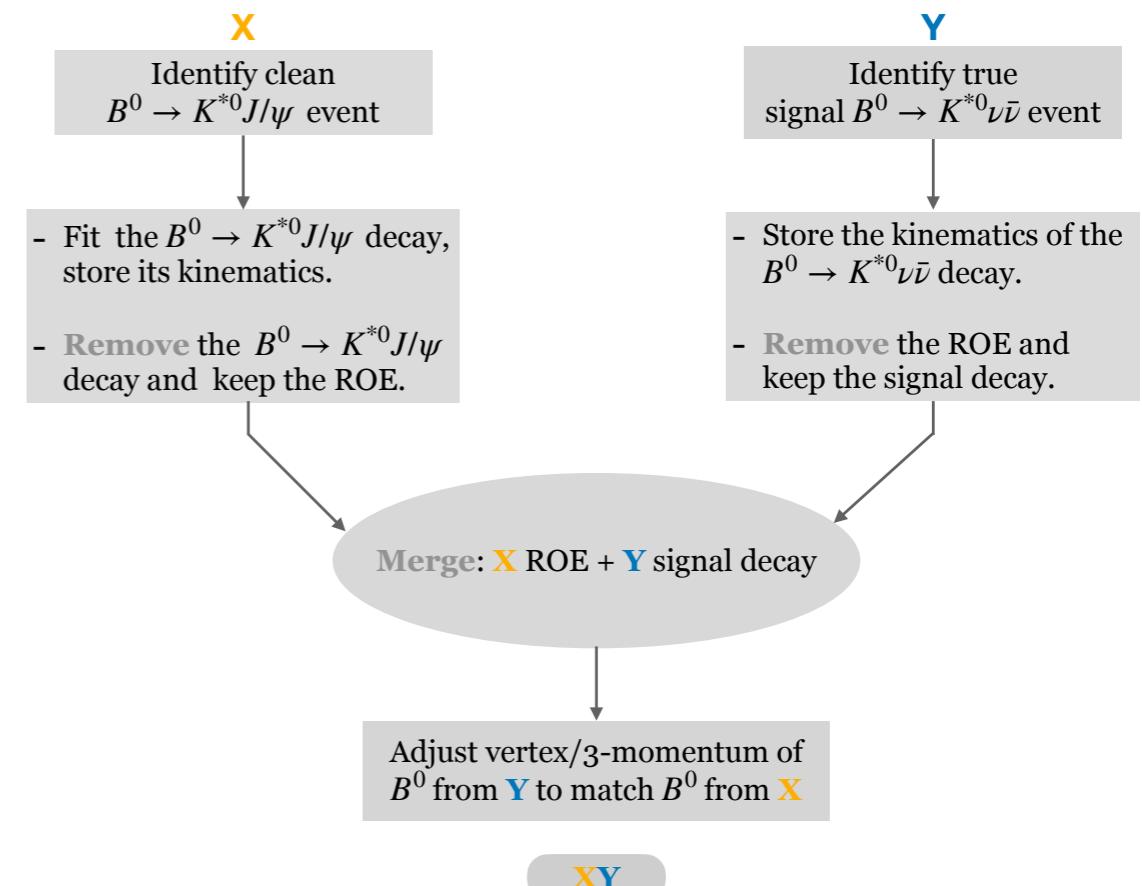


Idea



- Identify B decay by a clean **hadronic tag** (e.g. $B^0 \rightarrow K^{*0}J/\psi$).
- Remove the **hadronic tag** from the event.
- Replace it with a **signal B decay**.

Workflow



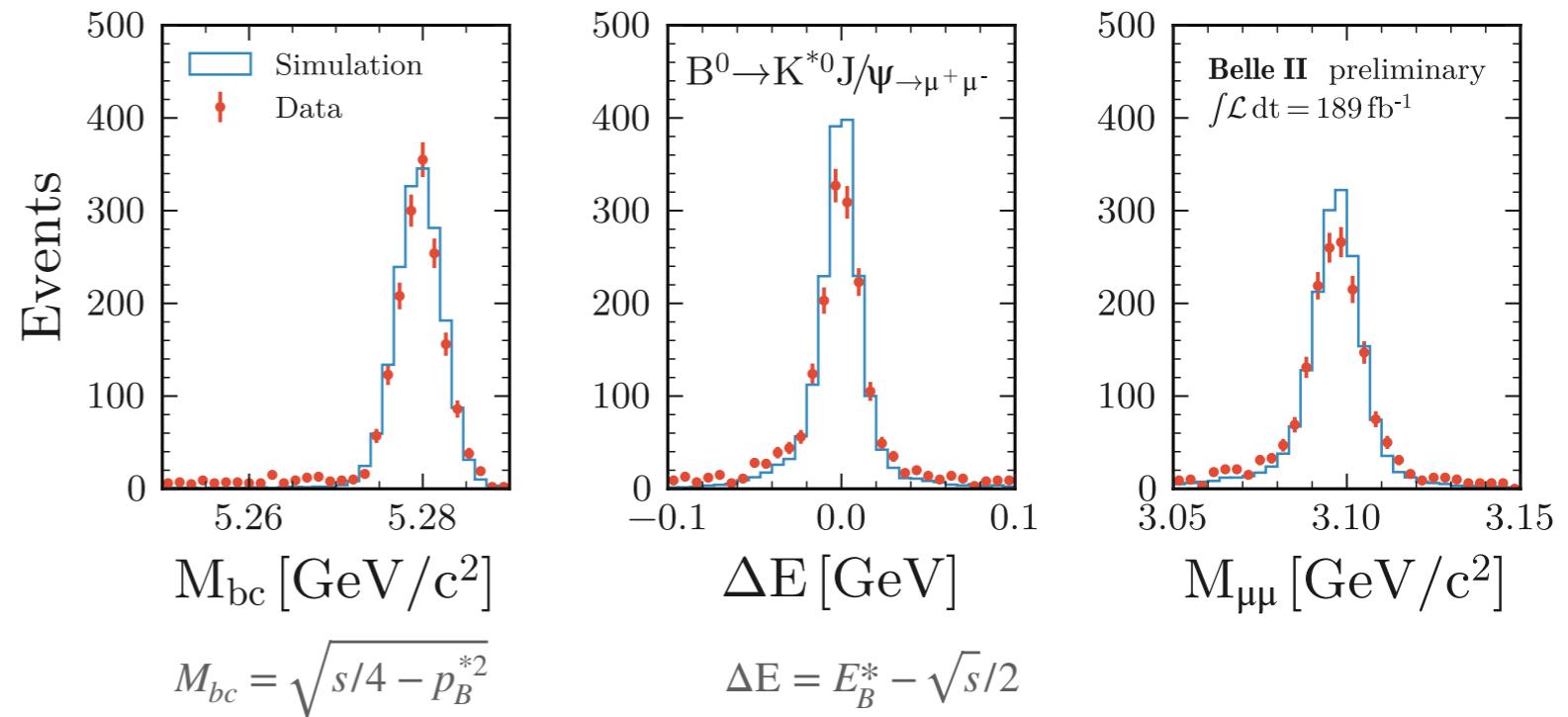
Validation studies



Control channel: $B^0 \rightarrow K^{*0} J/\psi \rightarrow \mu^+ \mu^-$

Identification of $B^0 \rightarrow K^{*0} J/\psi \rightarrow \mu^+ \mu^-$ events

- Candidate muons: track cleanup + $n_{\text{PXDHits}} > 0$ + muonID > 0.5
- Candidate J/ψ : pair of selected muons with $|M_{J/\psi}^{\text{PDG}} - M_{\mu^+ \mu^-}| < 50 \text{ MeV}/c^2$
- Candidate K^{*0} : default selection + muonID < 0.5 for the 2 tracks
- Candidate B_{sig}^0 : combine K^{*0} and J/ψ candidates + require $M_{bc} > 5.25 \text{ GeV}/c^2$ and $|\Delta E| < 100 \text{ MeV}$
- **1855** $B^0 \rightarrow K^{*0} J/\psi \rightarrow \mu^+ \mu^-$ events reconstructed in data.
- **11% selection efficiency** estimated in $B^0 \rightarrow K^{*0}(\rightarrow K^+ \pi^-) J/\psi(\rightarrow l^+ l^-)$ MC
- **Background level** $\mathcal{O}(5\%)$



$$\begin{aligned}\text{BR}(B^0 \rightarrow K^{*0} J/\psi) &= (1.27 \pm 0.05) \times 10^{-3} \\ \text{BR}(J/\psi \rightarrow \mu^+ \mu^-) &= (5.961 \pm 0.033) \times 10^{-2} \\ [\text{Prog. Theor. Exp. Phys } 2022, 083C01]\end{aligned}$$

Validation studies



Control channel: $B^0 \rightarrow K^{*0} J/\psi \rightarrow \mu^+ \mu^-$

- Inclusive-tagging selection efficiency:

- $B^0 \rightarrow K^{*0} J/\psi \rightarrow \mu^+ \mu^-$ in data $\rightarrow (80.9 \pm 1.0)\%$
- $B^0 \rightarrow K^{*0} J/\psi \rightarrow \mu^+ \mu^-$ in MC $\rightarrow (80.4 \pm 0.3)\%$

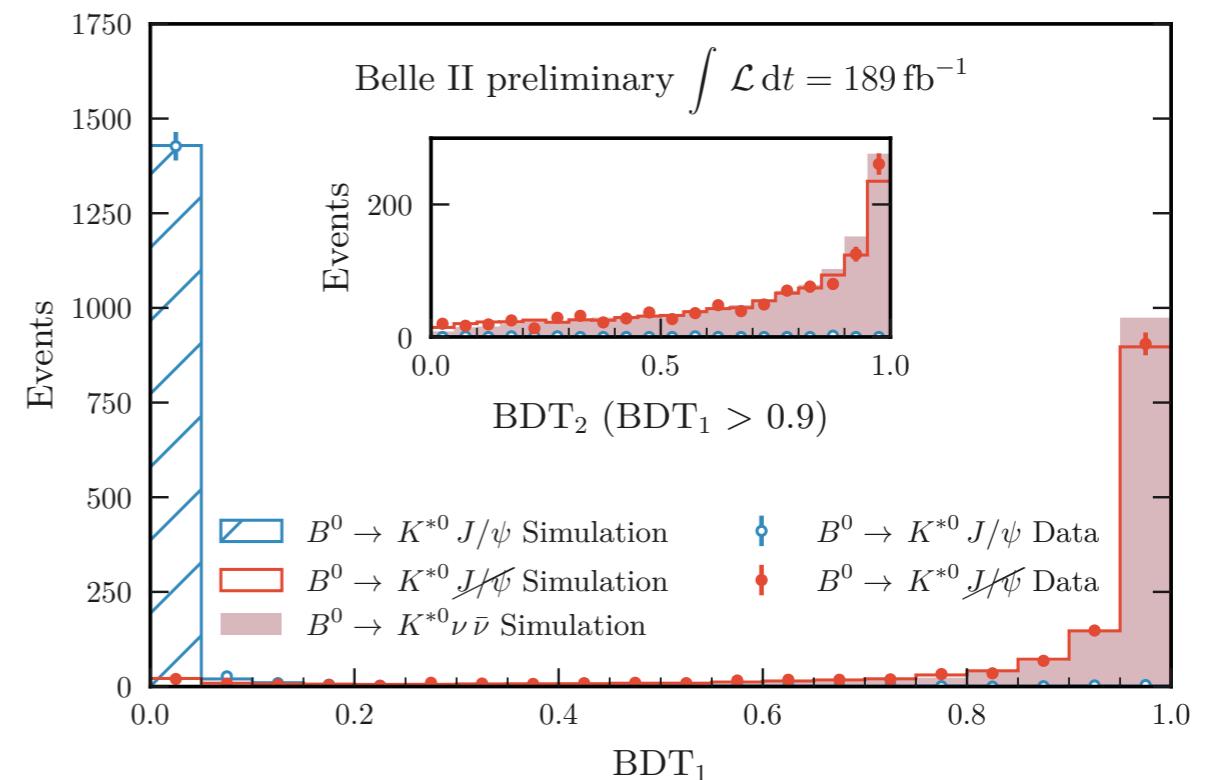
- Selection efficiency at $BDT_1 > 0.9$ and $BDT_2 > 0.94$,
($\epsilon = < 3\%$) region of highest signal sensitivity:

- $B^0 \rightarrow K^{*0} J/\psi$ in data $\rightarrow (20.9 \pm 1.1)\%$
- $B^0 \rightarrow K^{*0} J/\psi$ in MC $\rightarrow (19.2 \pm 0.3)\%$
- Data/MC efficiency ratio $= 1.09 \pm 0.06$
 \rightarrow introduce 8% systematic uncertainty.

- K-S test data vs MC for $BDT_2(BDT_1 > 0.9)$:

p-value = 51%

$$\begin{aligned} BR(B^0 \rightarrow K^{*0} J/\psi) &= (1.27 \pm 0.05) \times 10^{-3} \\ BR(J/\psi \rightarrow \mu^+ \mu^-) &= (5.961 \pm 0.033) \times 10^{-2} \\ &\text{[Prog. Theor. Exp. Phys 2022, 083C01]} \end{aligned}$$



Validation studies



Reweighting of the continuum simulation

Data-driven technique relying on a **binary classifier**:

- enhance the data-MC agreement by **event reweighting of the MC**.

Training samples

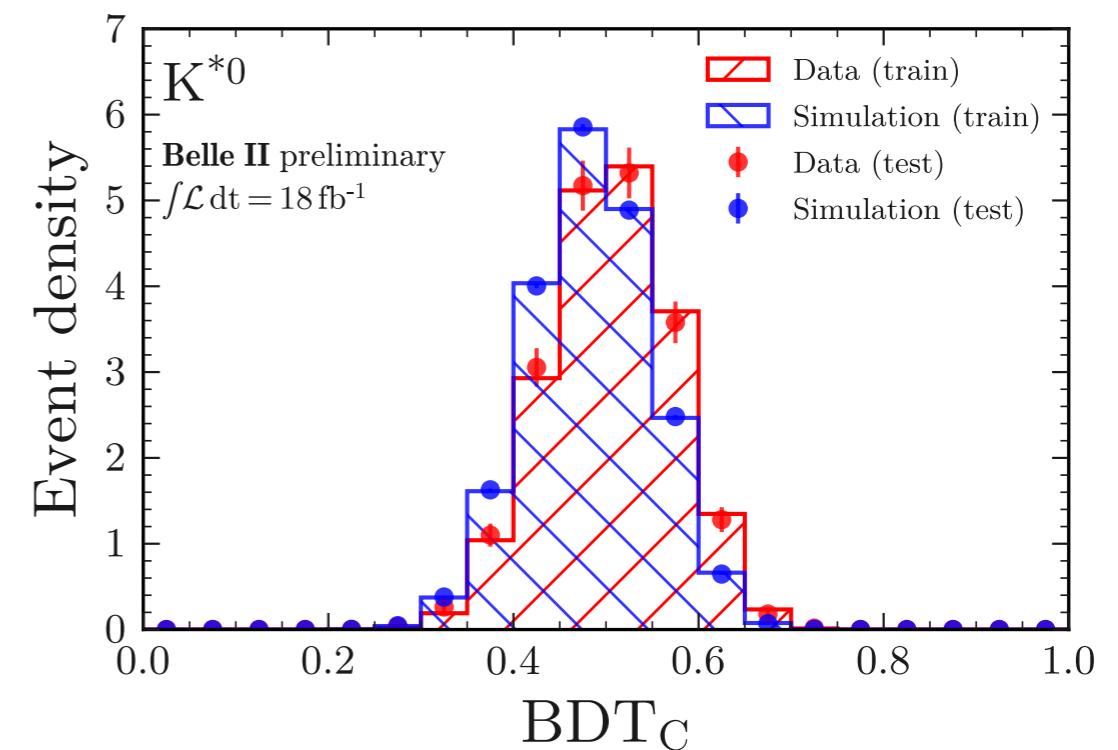
- Simulated events of the continuum bkg with $BDT_1 > 0.9$ reconstructed in MC sample of 100 fb^{-1} .
- Events with $BDT_1 > 0.9$ reconstructed in the off-resonance data sample of 18 fb^{-1} .

The classifier

BDT_C [XGBoost] trained with BDT_1 , BDT_2 variables + BDT_2 output + q_{rec}^2 , using simulated (off-resonance data) events as background (signal) class.

Model optimised against overfitting.

Hyperparameter	Value
Number of trees	2000
Tree depth	1
Shrinkage	0.01
Sampling rate	0.01
Number of equal-frequency bins	256



Validation studies



Reweighting of the continuum simulation

Data-driven technique relying on a **binary classifier**:

- enhance the data-MC agreement by **event reweighting of the MC**.

Training samples

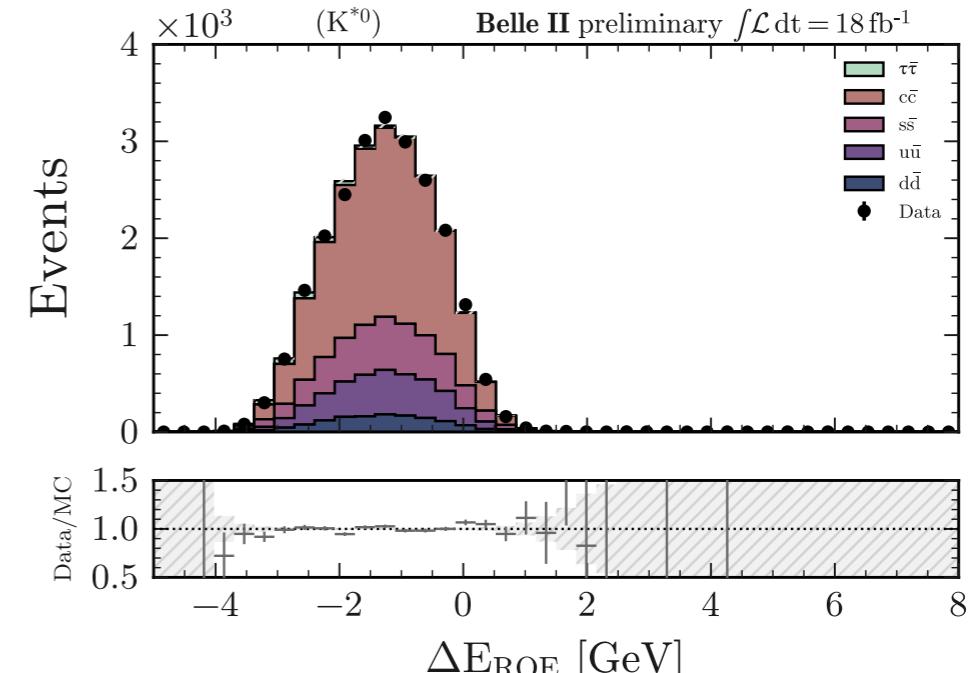
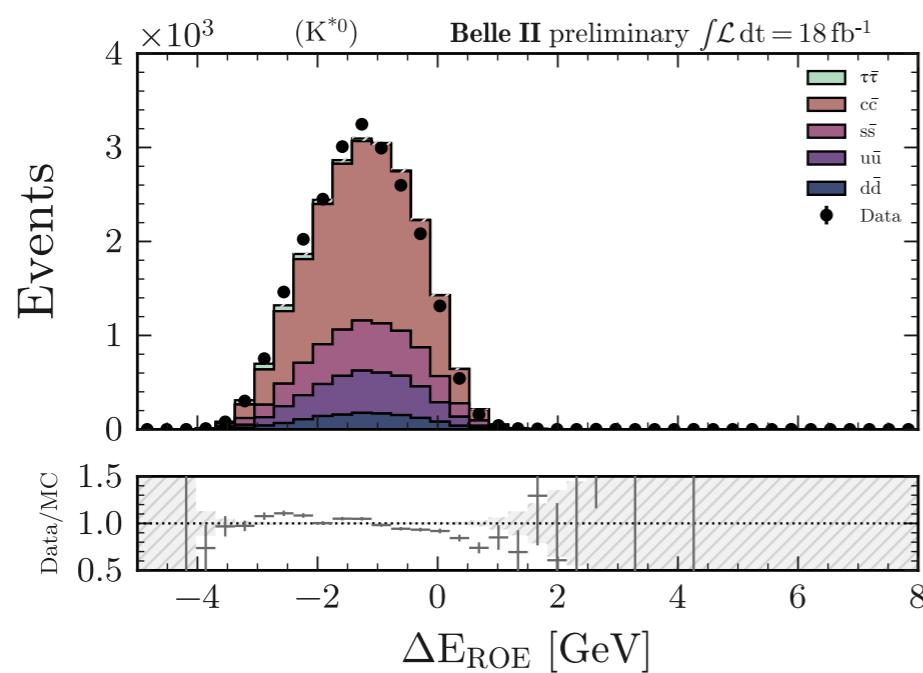
- Simulated events of the continuum bkg with $BDT_1 > 0.9$ reconstructed in MC sample of 100 fb^{-1} .
- Events with $BDT_1 > 0.9$ reconstructed in the off-resonance data sample of 18 fb^{-1} .

The classifier

BDT_C [XGBoost] trained with BDT_1 , BDT_2 variables + BDT_2 output + q_{rec}^2 , using simulated (off-resonance data) events as background (signal) class.

The event weight

Output p of $BDT_c \rightarrow$ weight $p/(1-p)$ assigned to simulation.

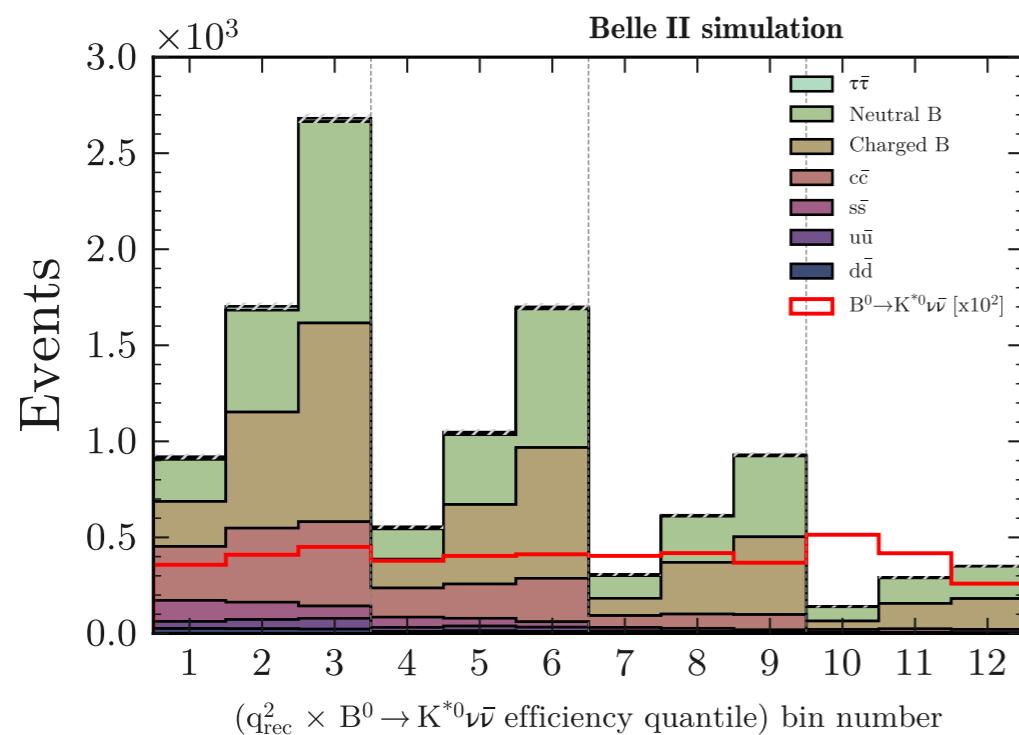


Statistical model



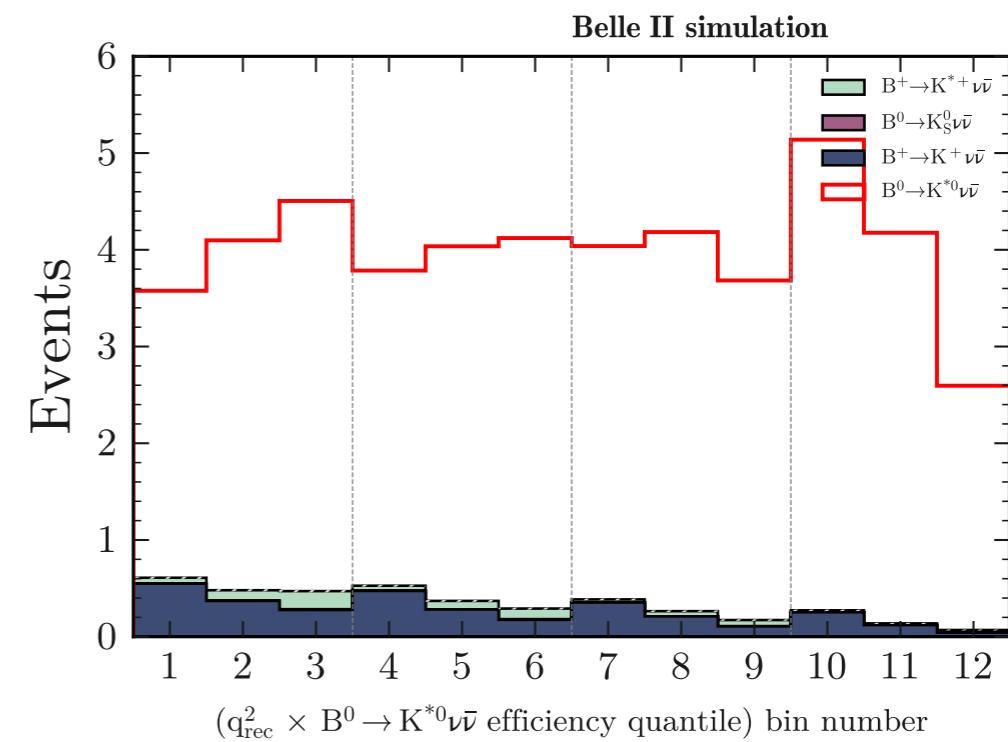
Expectations

- **CR1 + SR**



- sum of expected signal yields **constant** in the $\bar{\epsilon}$ bins ✓
- **dominant bkgds:** charged and neutral B
- **continuum** almost **negligible** in bins where signal sensitivity is highest (**10, 11, 12**)

Sample	Channels	Expected yields
signal	SR, CR1	36, 12
charged B	SR, CR1	2345, 1873
neutral B	SR, CR1	2397, 1797
$c\bar{c}$	SR, CR1, CR2, CR3	824, 1105, 78, 102
$d\bar{d}$	SR, CR1, CR2, CR3	46, 73, 4, 6
$s\bar{s}$	SR, CR1, CR2, CR3	170, 264, 17, 25
$u\bar{u}$	SR, CR1, CR2, CR3	98, 141, 11, 19
$\tau^+ \tau^-$	SR, CR1, CR2, CR3	58, 52, 5, 5



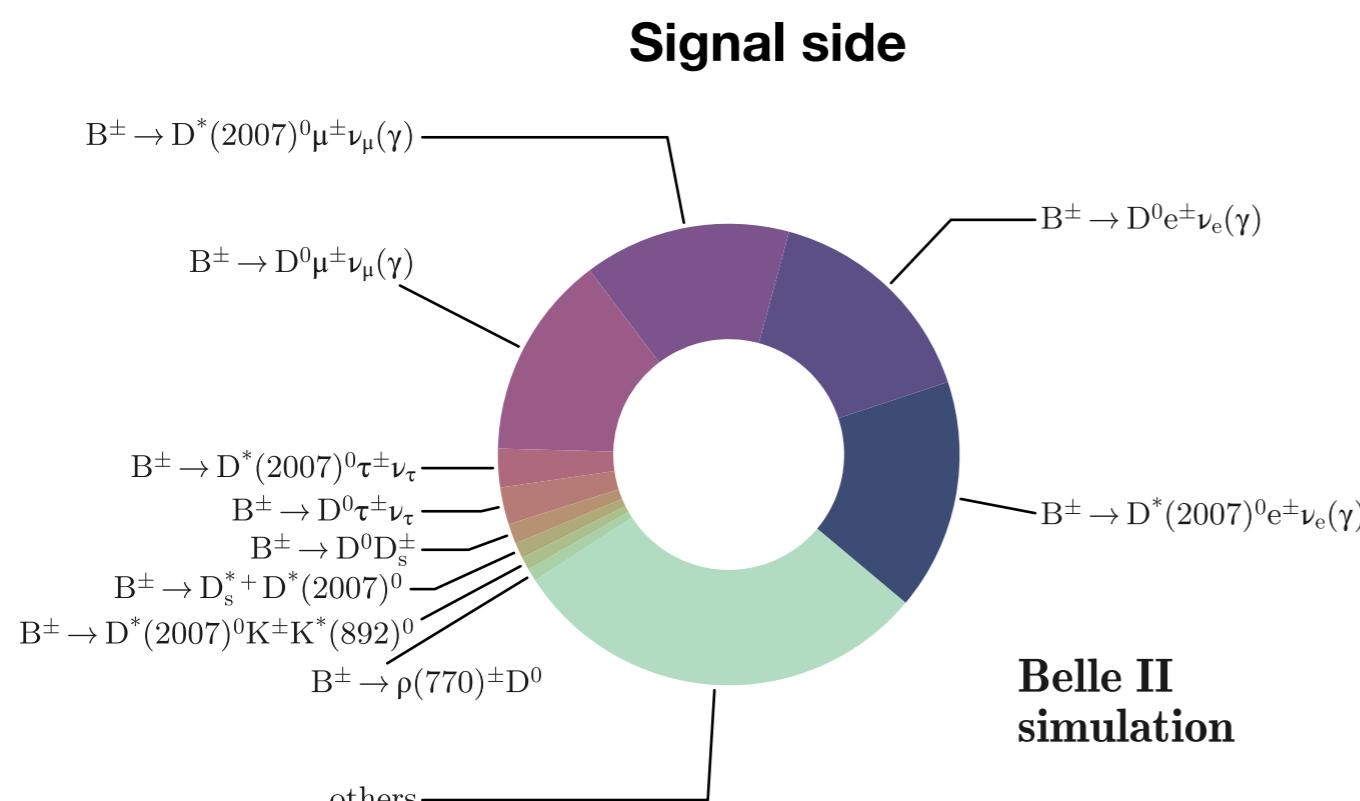
Combined **crossfeed** $\sim 10\%$ of $B^0 \rightarrow K^{*0} \nu \bar{\nu}$

Background composition



Charged B background

Major background decays in the reconstructed events in SR and CR1.



Inclusive D^0 modes

- $\text{BR}(D^0 \rightarrow K^- X) \simeq 55\%$
- $\text{BR}(D^0 \rightarrow K^+ X) \simeq 3\%$

Inclusive B^+ modes

- $\text{BR}(B^+ \rightarrow \bar{D}^0 X) \simeq 79\%$
- $\text{BR}(B^+ \rightarrow D^0 X) \simeq 9\%$

Specific exclusive B^+ modes

- $\text{BR}(B^+ \rightarrow \bar{D}^0 l^+ \nu_l) \simeq 2\%$
- $\text{BR}(B^+ \rightarrow \bar{D}^*(2007)^0 l^+ \nu_l) \simeq 6\%$
- $\text{BR}(B^+ \rightarrow \bar{D}^0 \tau^+ \nu_\tau) \simeq 8 \times 10^{-3}$
- $\text{BR}(B^+ \rightarrow \bar{D}^*(2007)^0 \tau^+ \nu_\tau) \simeq 2\%$
- $\text{BR}(B^+ \rightarrow \bar{D}^0 D_s^+) \simeq 9 \times 10^{-3}$
- $\text{BR}(B^+ \rightarrow D_s^{*+} \bar{D}^*(2007)^0) \simeq 2\%$
- $\text{BR}(B^+ \rightarrow \bar{D}^*(2007)^0 K^+ \bar{K}^*(892)^0) \simeq 1.5 \times 10^{-3}$
- $\text{BR}(B^+ \rightarrow \bar{D}^0 \rho^+) \simeq 1\%$

Major exclusive D^0 modes

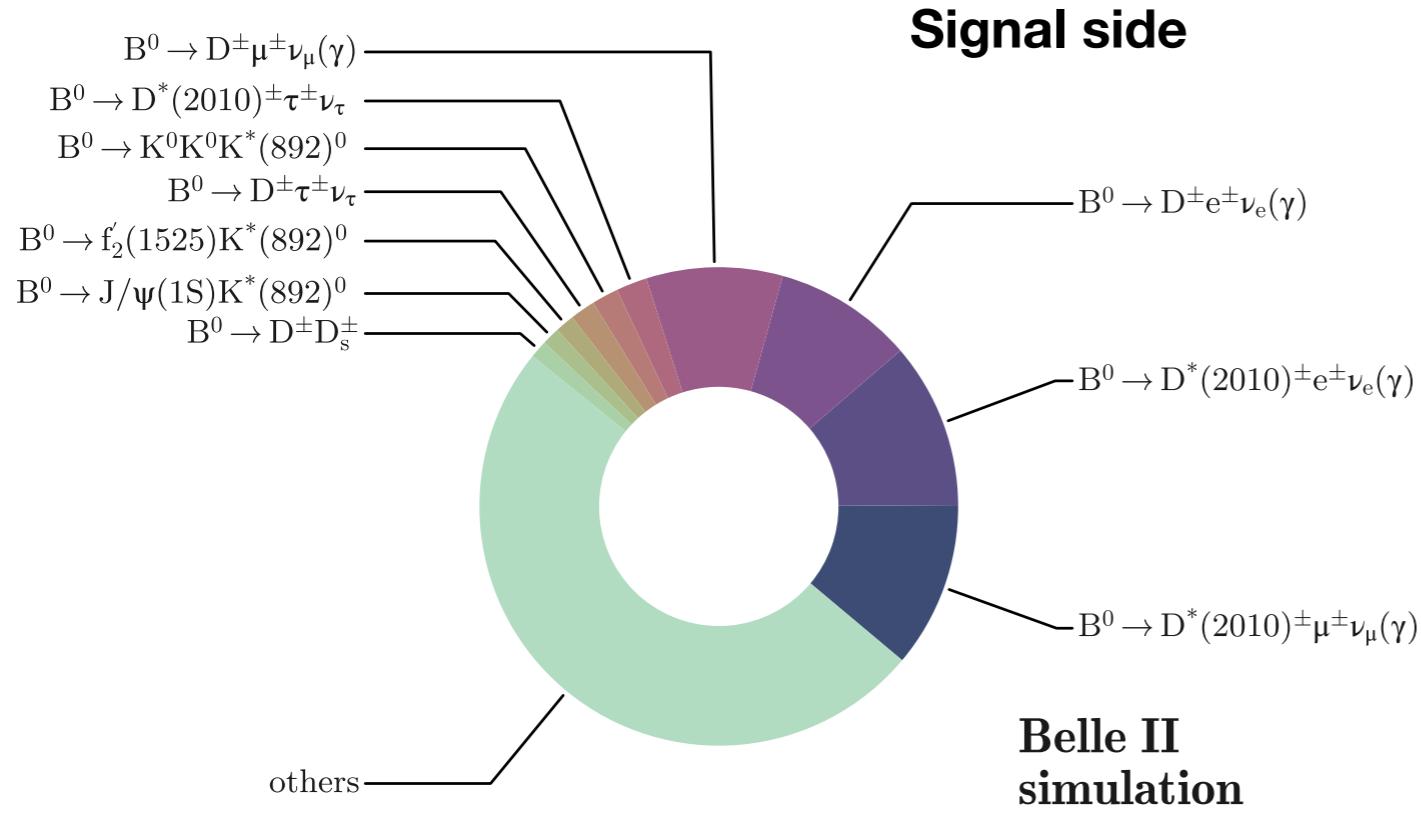
- $\text{BR}(D^0 \rightarrow K^- l^+ \nu_l) \simeq 7\%$
- $\text{BR}(D^0 \rightarrow K^*(892)^- l^+ \nu_l) \simeq 4\%$
- $\text{BR}(D^0 \rightarrow K^- \pi^+) \simeq 4\%$
- $\text{BR}(D^0 \rightarrow K^- \pi^+ \pi^0) \simeq 14\%$
- $\text{BR}(D^0 \rightarrow K^- \rho^+) \simeq 11\%$

Background composition



Neutral B background

Major background decays in the reconstructed events in SR and CR1.



Inclusive D^\pm modes

- $\text{BR}(D^\pm \rightarrow K^\pm X) \simeq 26\%$
- $\text{BR}(D^\pm \rightarrow K^\mp X) \simeq 6\%$

D^* modes

- $\text{BR}(D^*(2010)^+ \rightarrow D^0 \pi^+) \simeq 68\%$
- $\text{BR}(D^*(2010)^+ \rightarrow D^+ \pi^0) \simeq 31\%$
- $\text{BR}(D^*(2010)^+ \rightarrow D^+ \gamma) \simeq 1\%$

Inclusive B^0 modes

- $\text{BR}(B^0 \rightarrow D^- X) \simeq 37\%$

Specific exclusive B^0 modes

- $\text{BR}(B^0 \rightarrow D^- l^+ \nu_l) \simeq 5\%$
- $\text{BR}(B^0 \rightarrow \bar{D}^*(2010)^- l^+ \nu_l) \simeq 5\%$
- $\text{BR}(B^0 \rightarrow D^- \tau^+ \nu_\tau) \simeq 1\%$
- $\text{BR}(B^0 \rightarrow \bar{D}^*(2010)^- \tau^+ \nu_\tau) \simeq 2\%$
- $\text{BR}(B^0 \rightarrow J/\psi K^{*0}(892)) \simeq 1.3 \times 10^{-3}$
 $\text{BR}(J/\psi \rightarrow n\bar{n}) = 2 \times 10^{-3}$
- $\text{BR}(B^0 \rightarrow D^+ D_s^-) \simeq 7 \times 10^{-3}$

Major exclusive D^\pm modes

- $\text{BR}(D^+ \rightarrow \bar{K}^0 l^+ \nu_l) \simeq 20\%$
- $\text{BR}(D^+ \rightarrow \bar{K}^*(892)^0 l^+ \nu_l) \simeq 10\%$
- $\text{BR}(D^+ \rightarrow K^- \pi^+ \pi^+) \simeq 9\%$
- $\text{BR}(D^+ \rightarrow K_L^0 \pi^+) \simeq 1.5\%$

Systematic uncertainties



BR of the leading B -background decays

Only central values of the BR used in simulation → evaluate shape variations due to σ_{BR} (from PDG).

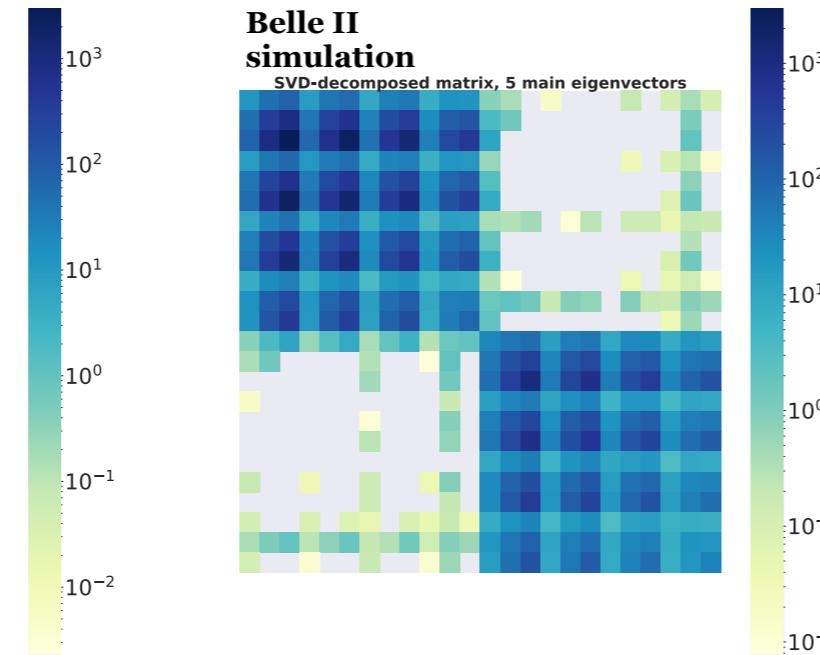
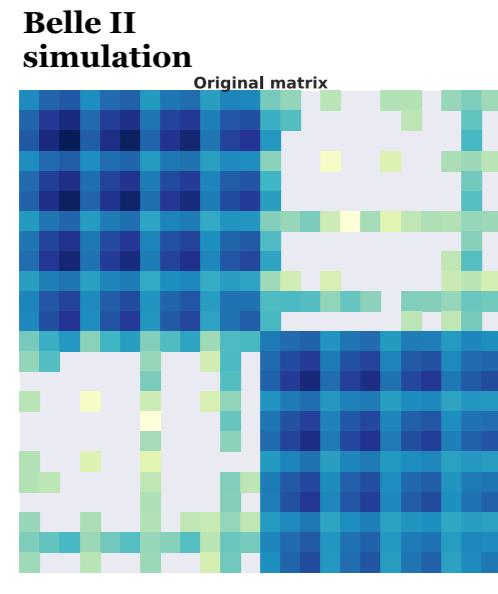
Samples: charged B , neutral B from 800 fb^{-1} sample of generic bkg.

Fractions of covered decays: $\sim 76\%$ of charged B , $\sim 58\%$ of neutral B .

1000 MC toys: weighted replicas of reconstructed B samples in the fit regions.

Event weights: $w_{\text{BR}} = (\text{BR} + \Delta)/\text{BR}$ with Δ from $N(0, \sigma_{\text{BR}})$

Compute bin-covariance matrix $\text{cov}(S_i, S_k)$ from samples of weighted bin counts $S_i = \{n_i\}_{j=1}^{1000}$ ($i = 1, \dots, 24$).



Bins 1, 4, 7, 10 low stat.

- Determine 5 main principal components.
 - Split among charged ($i = 1, \dots, 12$) and neutral ($i = 13, \dots, 24$) samples.
- ↓
- 5 orthogonal vectors** of systematic bin variations per sample, coupled to **3 nuisance parameters for correlated shape variation**.

Systematic uncertainties



BR of the leading B -background decays

Only central values of the BR used in simulation → evaluate shape variations due to σ_{BR} (from PDG).

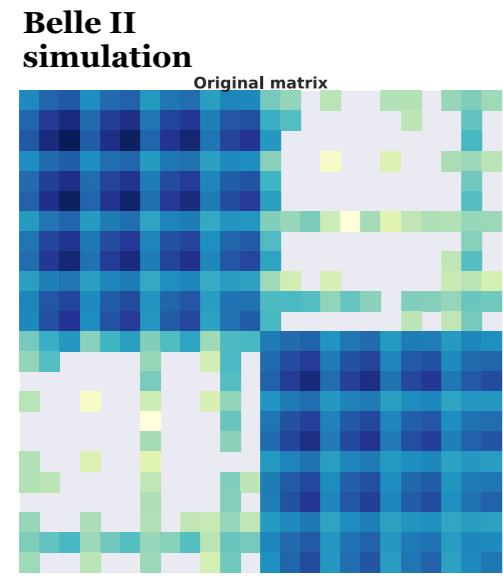
Samples: charged B , neutral B from 800 fb^{-1} sample of generic bkg.

Fractions of covered decays: $\sim 76\%$ of charged B , $\sim 58\%$ of neutral B .

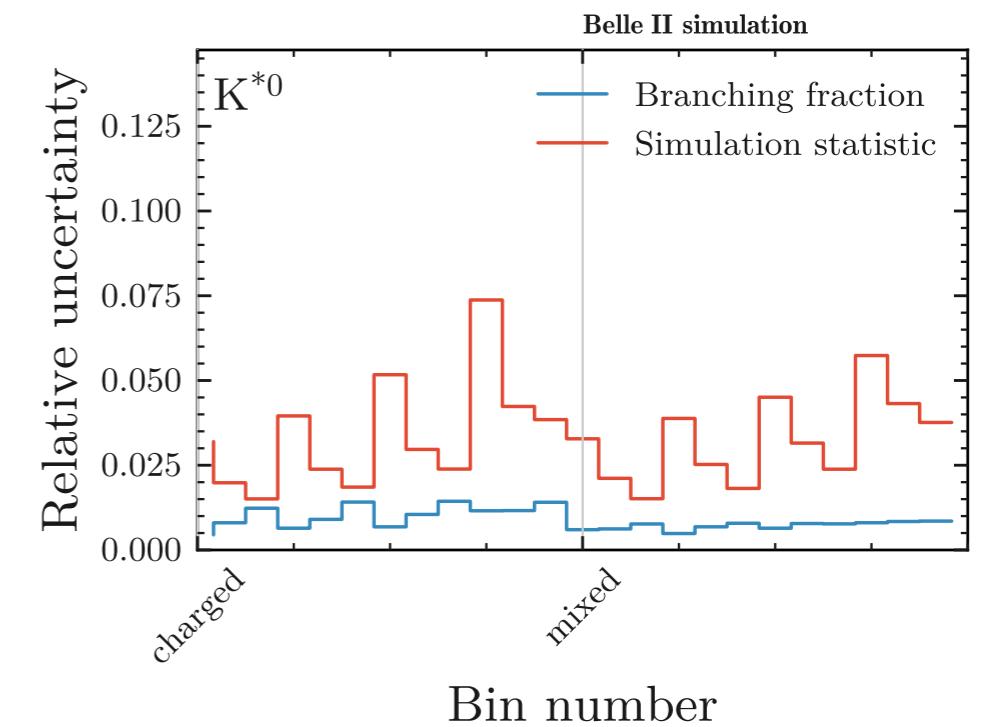
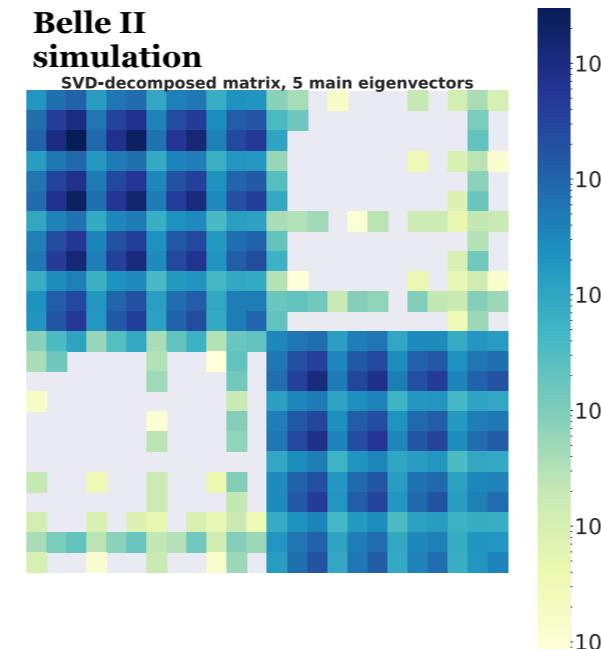
1000 MC toys: weighted replicas of reconstructed B samples in the fit regions.

Event weights: $w_{\text{BR}} = (\text{BR} + \Delta)/\text{BR}$ with Δ from $N(0, \sigma_{\text{BR}})$

Compute bin-covariance matrix $\text{cov}(S_i, S_k)$ from samples of weighted bin counts $S_i = \{n_i\}_{j=1}^{1000}$ ($i = 1, \dots, 24$).



Bins 1, 4, 7, 10 low stat.



Systematic uncertainties



SM form factors

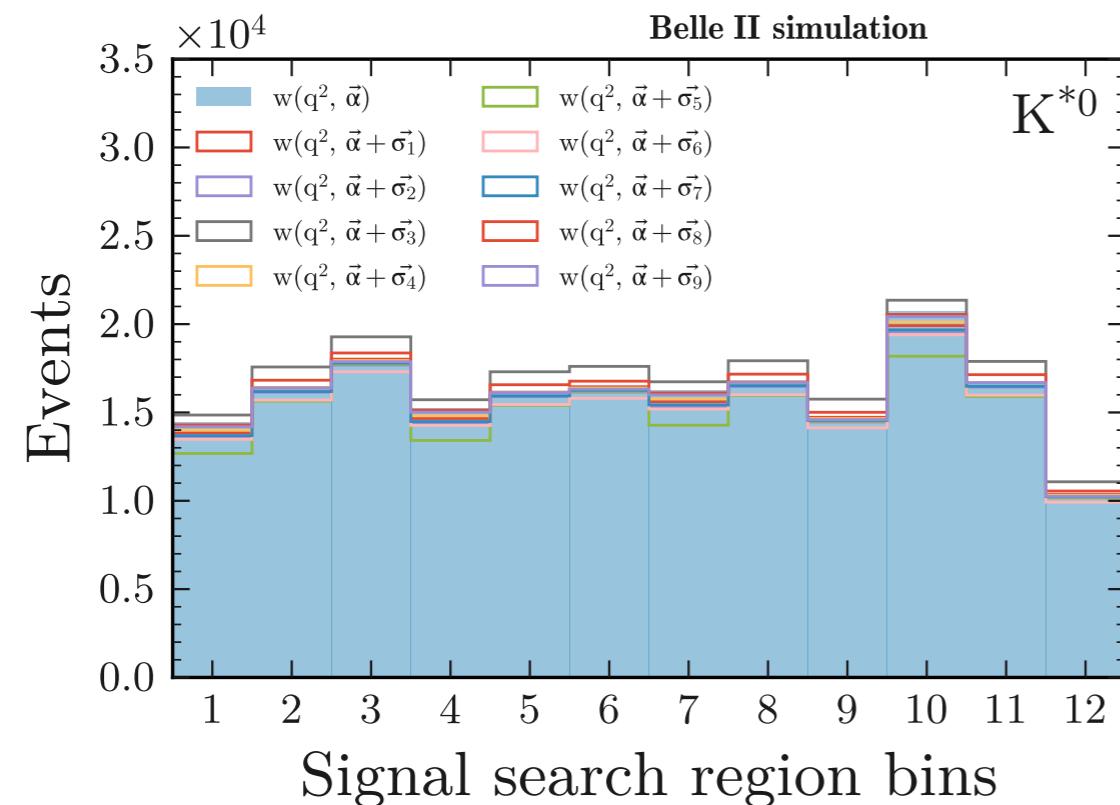
Each form factor depends on three real parameters estimated in fit to Lattice QCD and LCSR results:

- central values $\vec{\alpha} \equiv (\alpha_0, \alpha_1, \alpha_2)_i$, $i = A_1, A_{12}, V$;
- 9×9 covariance matrix $\text{cov}(\alpha_i, \alpha_j)$.

[J. High Energ. Phys. 2016, 8]

Decompose $\text{cov}(\alpha_i, \alpha_j)$ to the 9 principal components $\vec{\sigma}_i$.

Derive **9 modified form factors** based on $\vec{\alpha} + \vec{\sigma}_i$



- Compare the 9 modified signal distributions in CR1 + SR to the expected distribution to derive **9 independent vectors of bin-count variations**.
- Couple the 9 vectors to **9 nuisance parameters for correlated shape variation**.

Systematic uncertainties



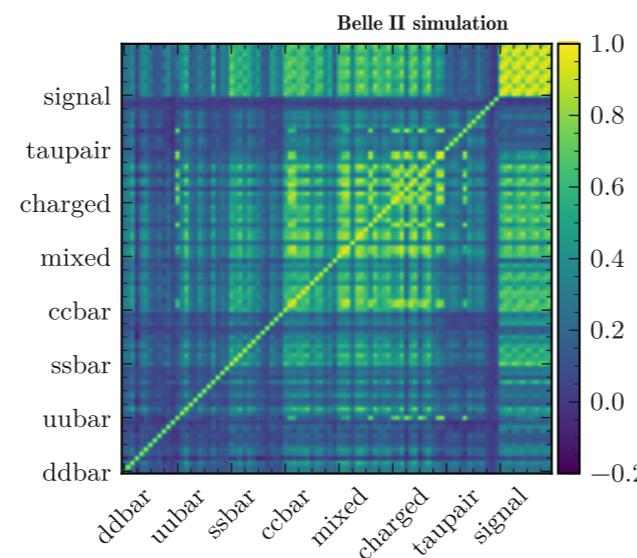
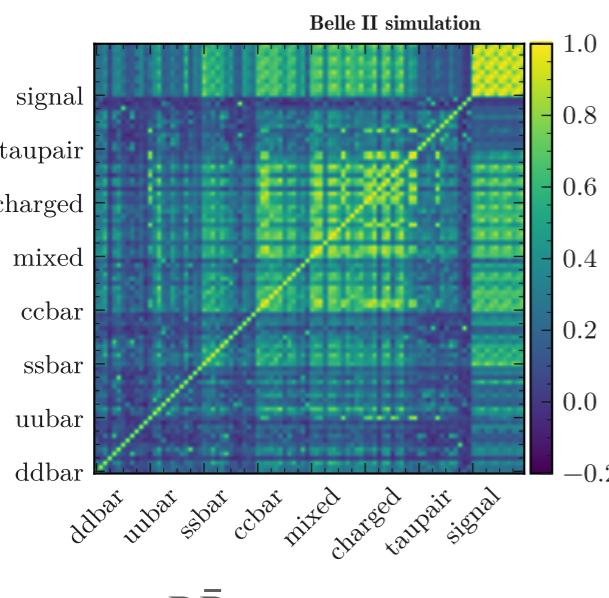
PID correction

Data/MC correction weights determined in **specific $(p, \cos\theta)$ bins** → related uncertainties can introduce correlated shape variation.

Samples: 800 fb^{-1} of generic bkg + 4×10^6 simulated signal events.

500 MC toys: weighted replicas of reconstructed samples in the fit regions.

- **In each replica: vary PID correction weights** in $(p, \cos\theta)$ bins (using lognormal with width equal to the weight uncertainty) **and compute weighted signal and bkg histograms.**
- Compute bin-covariance matrix $\text{cov}(S_{ij}, S_{mn})$ from samples of weighted bin counts $S_{ij} = \{n_{ij}\}_{k=1}^{500}$ ($i \in \text{samples}, j \in \text{bins}$)



- Determine 7 main principal components and split them among samples.



7 orthogonal vectors of systematic bin variations per sample, coupled to **7 nuisance parameters for correlated shape variation.**

Systematic uncertainties



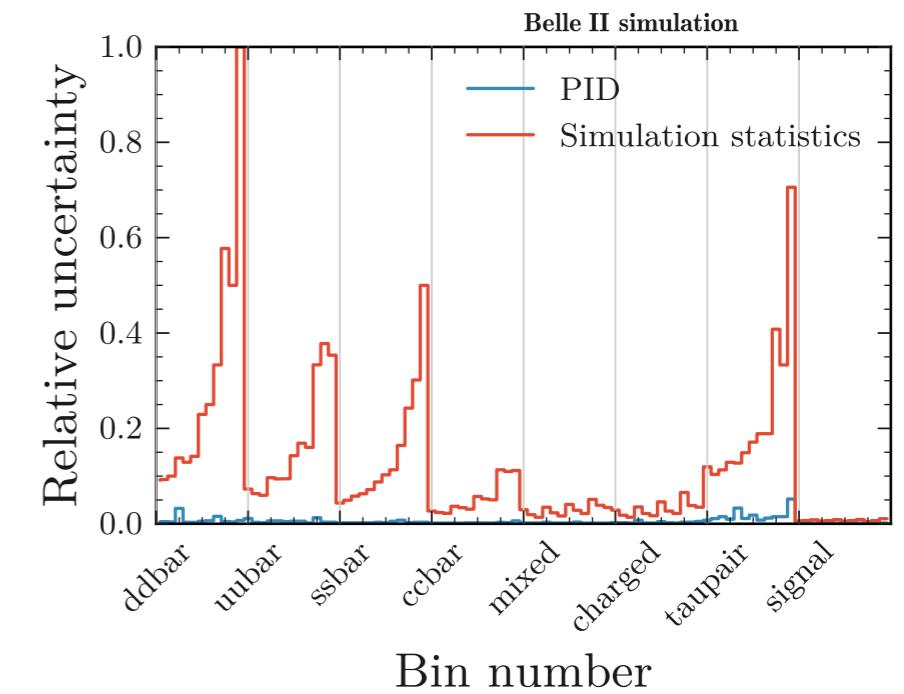
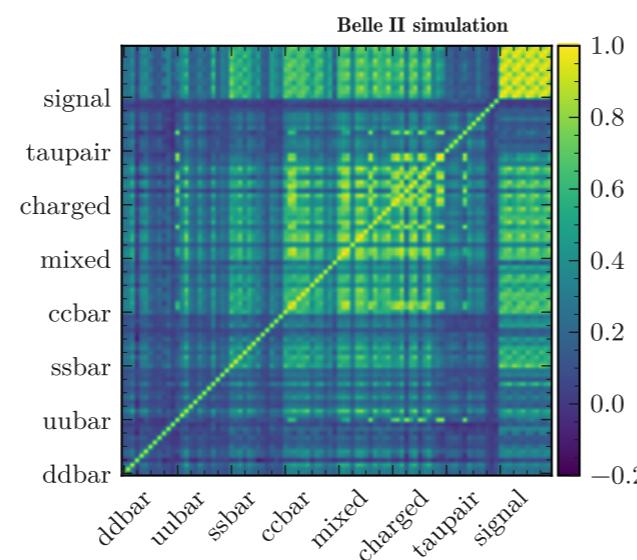
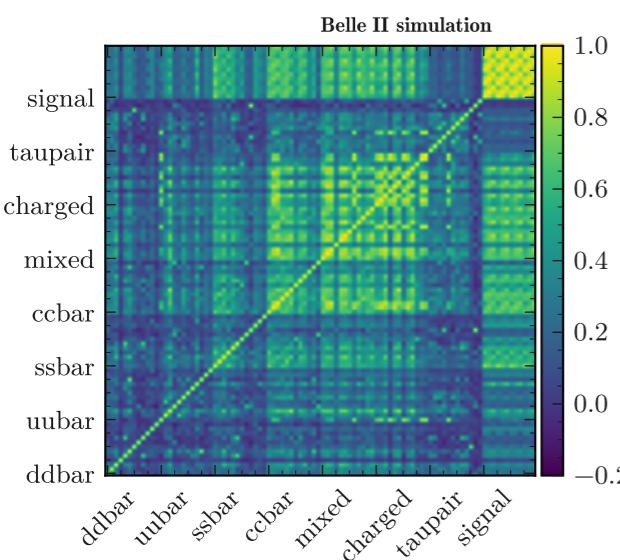
PID correction

Data/MC correction weights determined in **specific $(p, \cos\theta)$ bins** → related uncertainties can introduce correlated shape variation.

Samples: 800 fb^{-1} of generic bkg + 4×10^6 simulated signal events.

500 MC toys: weighted replicas of reconstructed samples in the fit regions.

- **In each replica: vary PID correction weights** in $(p, \cos\theta)$ bins (using lognormal with width equal to the weight uncertainty) **and compute weighted signal and bkg histograms.**
- Compute bin-covariance matrix $\text{cov}(S_{ij}, S_{mn})$ from samples of weighted bin counts $S_{ij} = \{n_{ij}\}_{k=1}^{500}$ ($i \in \text{samples}, j \in \text{bins}$)



Systematic uncertainties

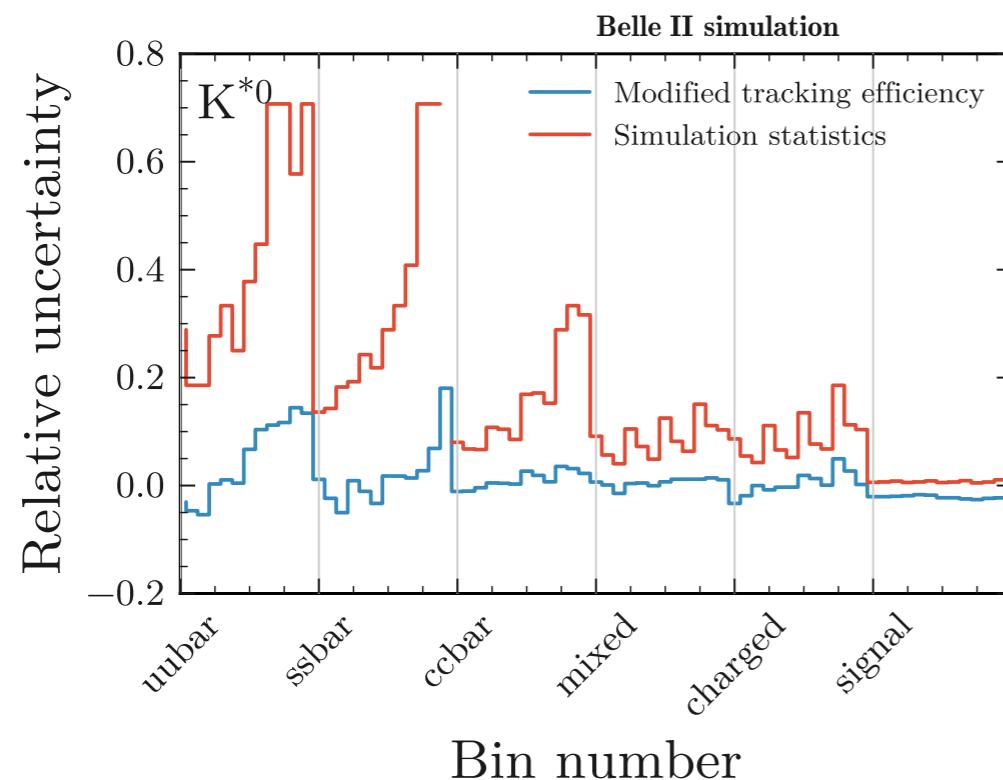


Tracking efficiency

Estimated uncertainty is 0.9% [BELLE2-NOTE-PL-2020-014] → simulate 0.9% probability to lose a track.

Samples: 100 fb^{-1} of generic bkg + 4×10^6 simulated signal events.

- Compute signal and bkg histograms in the fit regions with **tracking inefficiency turned on**.
- **Smooth reference and modified bkg histograms** using Gaussian KDE:
avoid large fluctuations due to limited size of bkg samples.
- Evaluate **relative systematic uncertainties** in the bins as **modified/reference counts - 1**.



1 vector of systematic bin variations per sample, coupled to **1 common nuisance parameter for correlated shape variation**.

$\tau^+\tau^-$ and $d\bar{d}$ excluded: minor bkg.

Systematic uncertainties

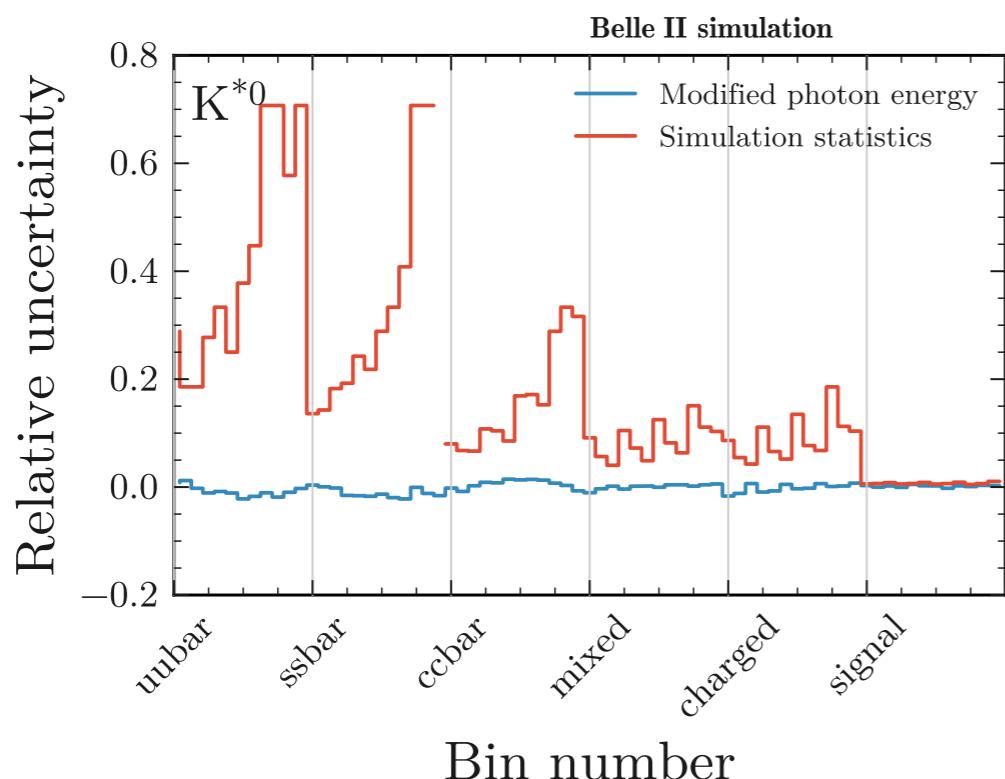


Energy calibration of photon clusters

Estimated energy uncertainty is 0.5% [BELLE2-NOTE-PL-2020-014] → scale down by 0.5% the energy of ECL clusters matched to photons in simulation.

Samples: 100 fb^{-1} of generic bkg + 4×10^6 simulated signal events.

- Compute signal and bkg histograms in the fit regions with **photon energy scaled down**.
- **Smooth reference and modified bkg histograms** using Gaussian KDE:
avoid large fluctuations due to limited size of bkg samples.
- Evaluate **relative systematic uncertainties** in the bins as **modified/reference counts - 1**.



1 vector of systematic bin variations per sample, coupled to **1 common nuisance parameter for correlated shape variation**.

$\tau^+\tau^-$ and $d\bar{d}$ excluded: minor bkg.

Systematic uncertainties

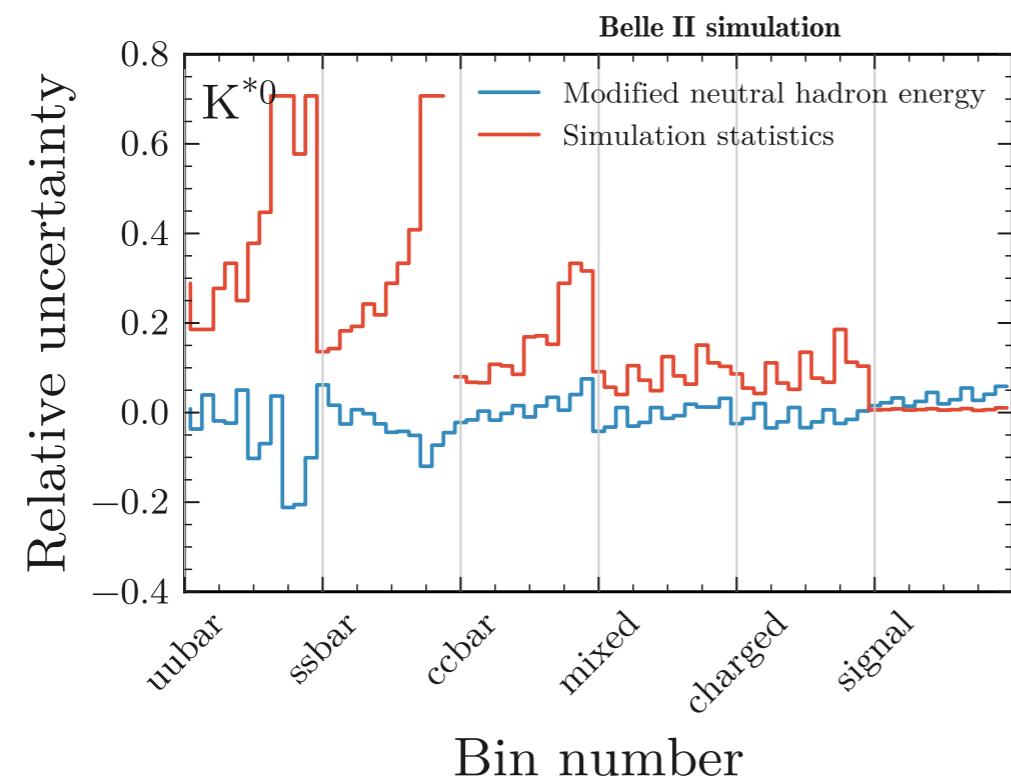


Energy calibration of ECL clusters not matched to photons

Scale down by 10% the energy of ECL clusters not matched to photons in simulation.

Samples: 100 fb^{-1} of generic bkg + 4×10^6 simulated signal events.

- Compute signal and bkg histograms in the fit regions with **energy of unmatched clusters scaled down**.
- **Smooth reference and modified bkg histograms** using Gaussian KDE:
avoid large fluctuations due to limited size of bkg samples.
- Evaluate **relative systematic uncertainties** in the bins as **modified/reference counts - 1**.



1 vector of systematic bin variations per sample, coupled to **1 common nuisance parameter for correlated shape variation**.

$\tau^+\tau^-$ and $d\bar{d}$ excluded: minor bkg.

Signal: systematic uncertainty ($< 5\%$) slightly larger than stat.

Systematic uncertainties



Mismodelling of hadronic resonances

Partial mismodelling of wide hadronic resonances generated by fragmentation in the MC simulation: wide resonances generated by **KKMC+PYTHIA** are produced with wrong masses.

Correction weights: if true primary K^{*0} (mcMother $\equiv Z^0$) fix mass distribution using weight originating from convolution of Breit-Wigner PDFs.

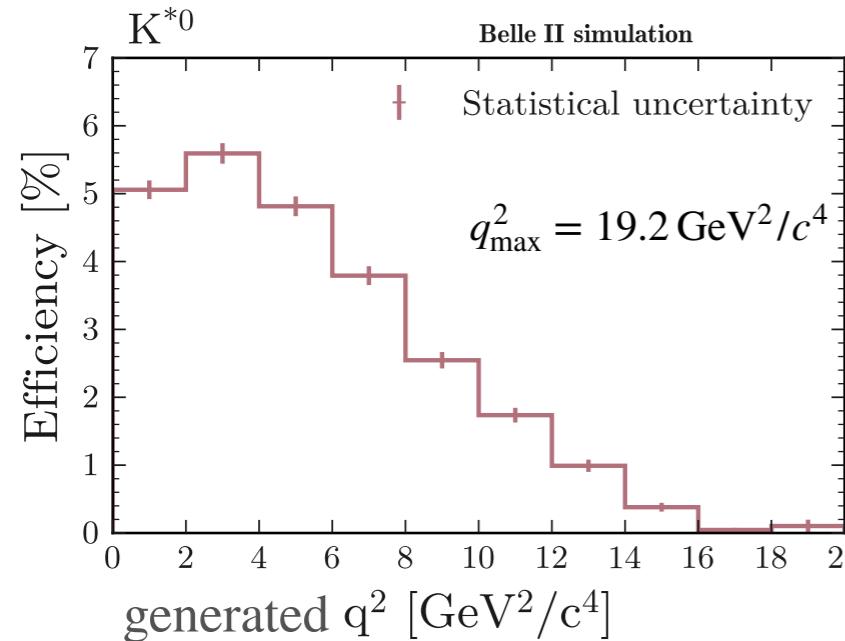
In the statistical model: **1 common nuisance parameter for correlated shape modification** coupled to the systematic variations determined by comparing expectations with and without weights for each bkg sample.

Discussion of the results



Performance of the inclusive tagging

Signal efficiency in SR:



Measurement	Tagging	Signal efficiency
BaBar (2013), Ref. [124]	hadronic + semileptonic	$1.3 \times 10^{-4} + 6.9 \times 10^{-4}$
Belle (2013), Ref. [123]	hadronic	1.44×10^{-4}
Belle (2017), Ref. [126]	semileptonic	0.51×10^{-3}
Belle II (2022), this work	inclusive	2.25×10^{-2}

Highly efficient at low q^2 : region of a possible dark-scalar narrow resonance [Phys. Rev. D 2020, 101]

Outlook

Projections of $\sigma_{\text{BR}}/\text{BR}_{\text{SM}}$ in the assumption that **expected** $\sigma_{\text{BR}} = 2.6 \times 10^{-5}$ scales as $1/\sqrt{L}$:

$L(\text{ab}^{-1})$	Baseline scenario	Improved scenario
1	1.45	1.18
5	0.65	0.53
10	0.46	0.37
50	0.20	0.17

- $L = 10 \text{ ab}^{-1}$ sensitivity to SM signal rate at 3σ level in the improved scenario.
- SM accuracy reachable with 50 ab^{-1} .

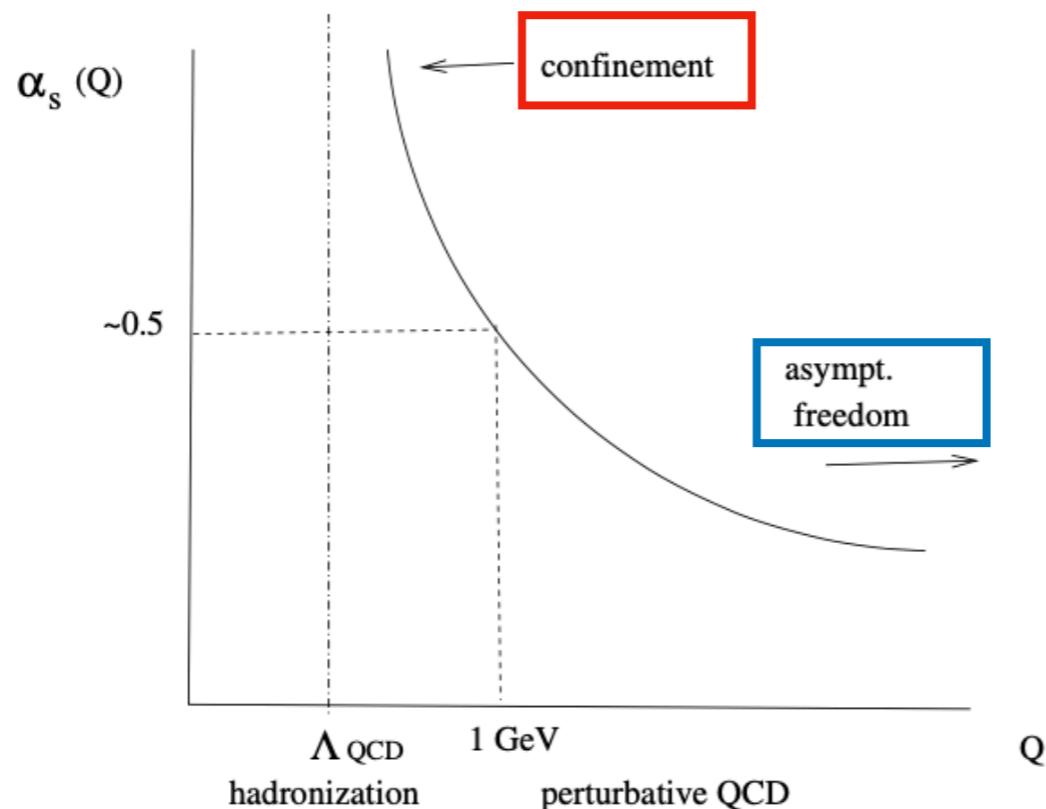
Baseline = no improvements

Improved = 50% increase in signal eff. at same bkg level

Regimes of QCD

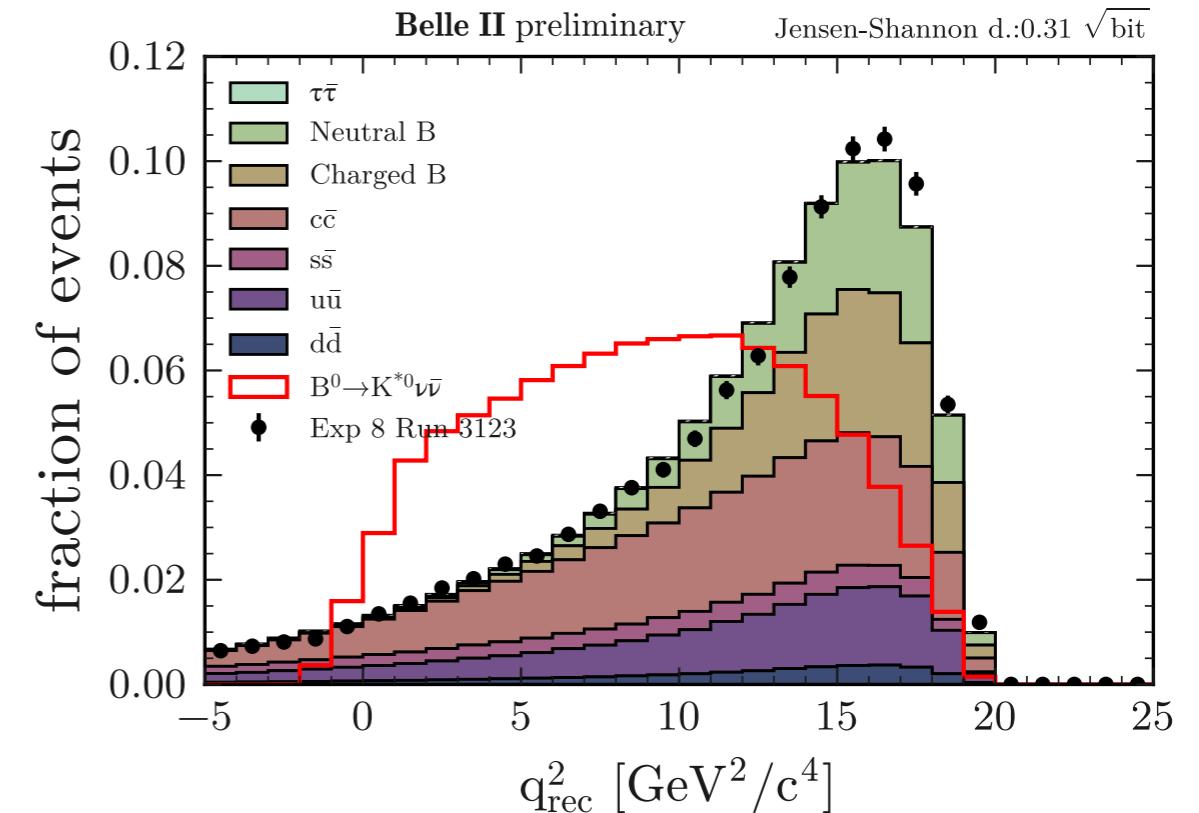
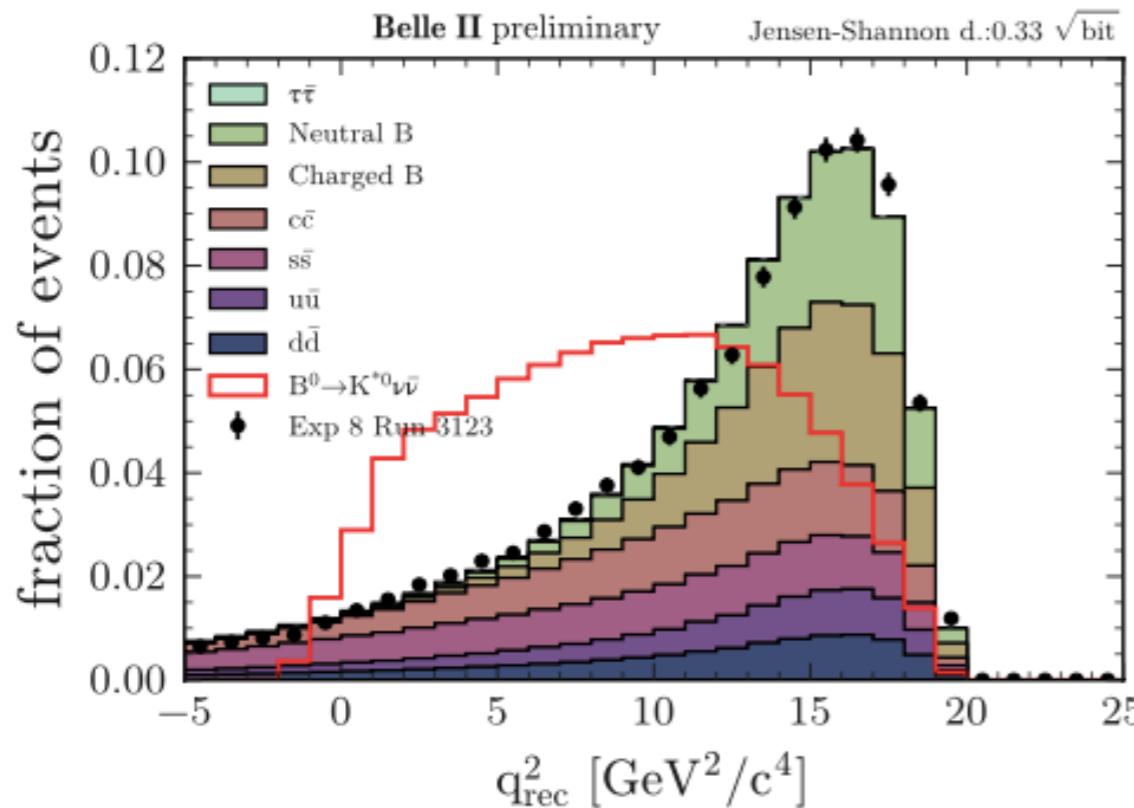


- $\alpha_s(Q) \ll 1$ at energy/momenta scales $Q \geq 1 \text{ GeV}$
- $\alpha_s(Q) \rightarrow \infty$ at $\Lambda_{\text{QCD}} \sim 200\text{--}300 \text{ MeV}$
- at $Q \lesssim \Lambda_{\text{QCD}}$ **nonperturbative regime of QCD**
- at $Q \sim \Lambda_{\text{QCD}}$ **confinement of quarks/gluons, hadronization**



e.g., $Z \rightarrow \bar{q}q \rightarrow \text{jets}$,
 $\alpha_s(M_Z) = 0.1185 \pm 0.0006$,
determined from experiment!

Weights correction





SuperKEKB

Components

- e^- source = **photo-cathode radio-frequency gun**:
photo-cathode illuminated by laser converting photons into e^- by photoelectric effect + radio-frequency (RF) cavities to accelerate e^- to MeV in few cm → LINAC up to GeV via RF sources producing alternating voltages between which the particle accelerates.
- e^+ source: electrons impinging on tungsten target radiate a photon which does e^+e^- pair production → LINAC + damping ring to reduce emittance by radiation damping = a combination of synchrotron radiation induced by bending magnets to reduce momentum and energy gain in radio-frequency cavities.

