# $b \to s\nu\bar{\nu}$ at Belle II

Anomalies and Precision in the Belle II Era
September 6-8, 2021

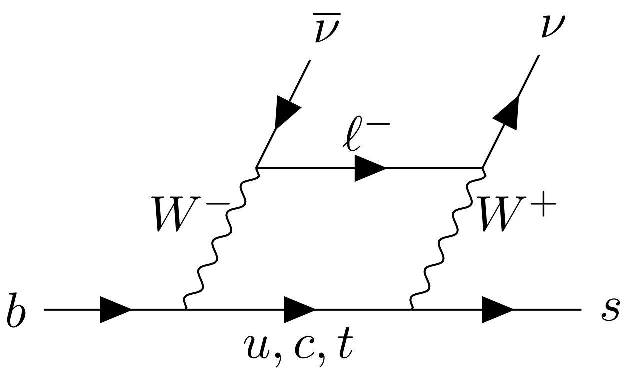
**Filippo Dattola**, on behalf of the Belle II Collaboration

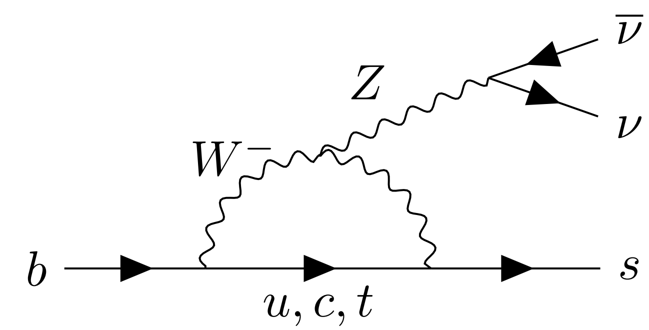
# $b \to s\nu\bar{\nu}$ transitions

# $b \to s\nu\bar{\nu}$ transitions

In the Standard Model (SM):

# $b \to s\nu\bar{\nu}$ transitions

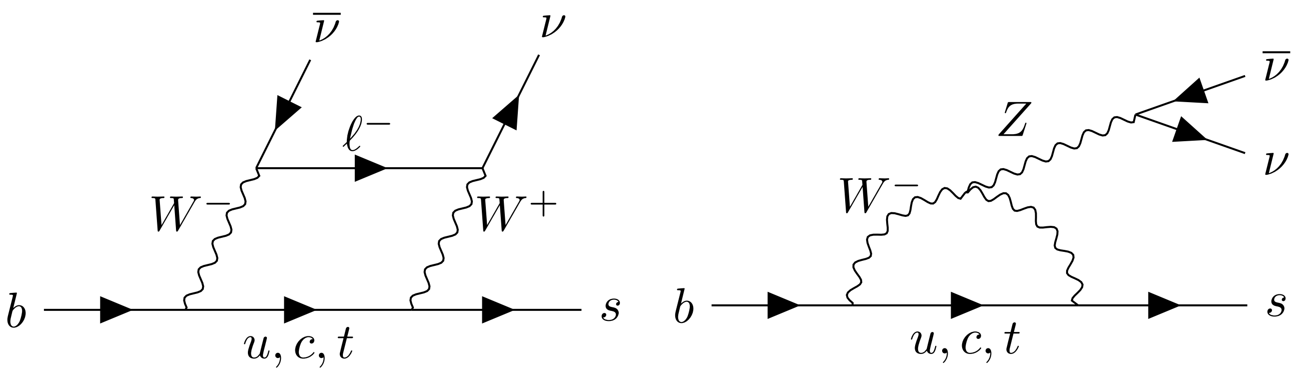



## In the Standard Model (SM):

- **flavour-changing neutral-current transitions (FCNCs);**
- can occur only at the loop level, **highly suppressed;**
- accurate theoretical predictions.

# $b \to s\nu\bar{\nu}$ transitions




## In the Standard Model (SM):

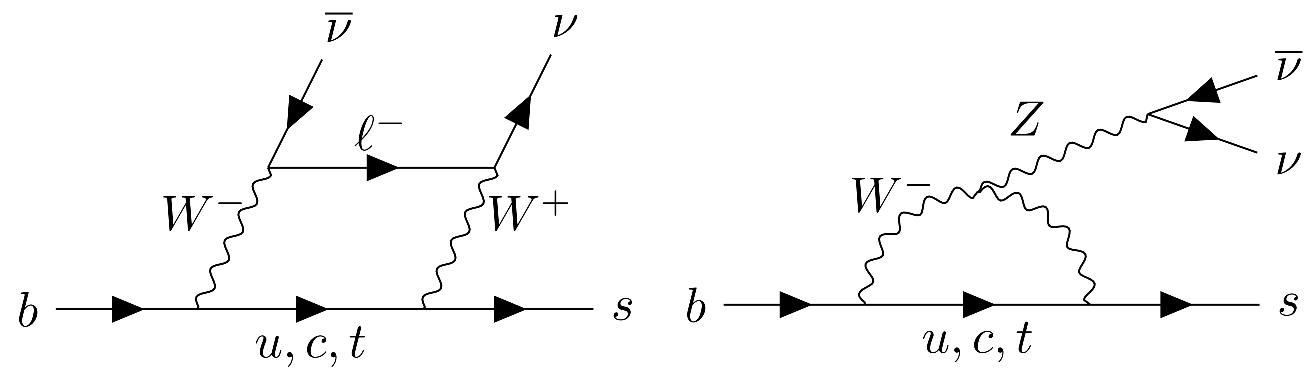
- **flavour-changing neutral-current transitions (FCNCs);**
- can occur only at the loop level, **highly suppressed;**
- accurate theoretical predictions.

**[T. Blake, et al., Prog. Part. Nucl. Phys. 92, 50 (2017)]**

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

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

# $b \to s\nu\bar{\nu}$ transitions




## In the Standard Model (SM):

- **flavour-changing neutral-current transitions (FCNCs);**
- can occur only at the loop level, **highly suppressed;**
- accurate theoretical predictions.

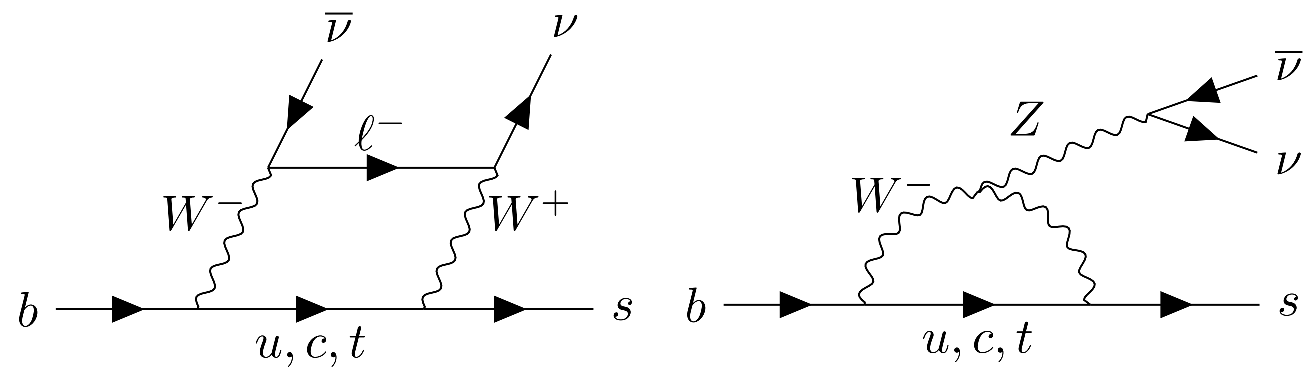
[T. Blake, et al., Prog. Part. Nucl. Phys. 92, 50 (2017)]

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

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

**Optimal measurements to probe the SM:**

# $b \to s\nu\bar{\nu}$ transitions




## In the Standard Model (SM):

- **flavour-changing neutral-current transitions (FCNCs);**
- can occur only at the loop level, **highly suppressed;**
- accurate theoretical predictions.

[T. Blake, et al., Prog. Part. Nucl. Phys. 92, 50 (2017)]

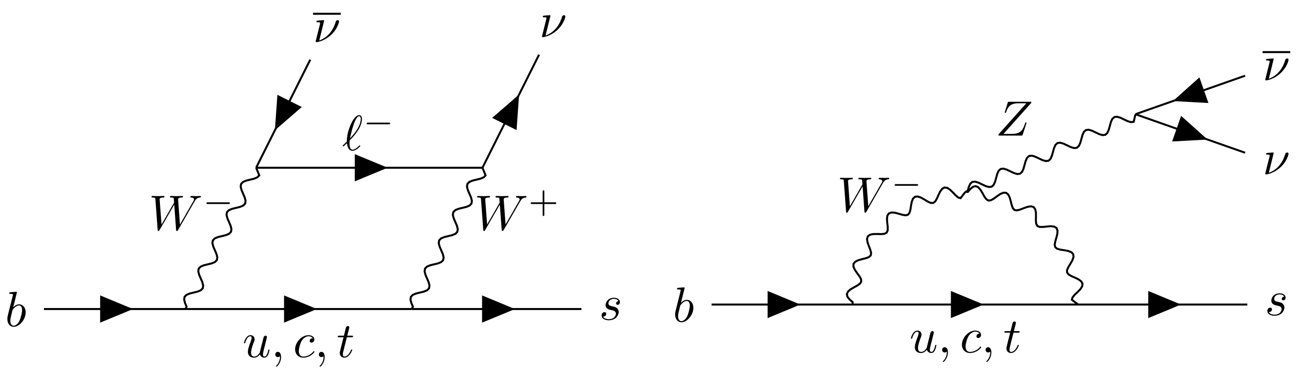
$$\mathrm{BR}(B \to K\nu\bar{\nu})_{SM} = (4.6 \pm 0.5) \times 10^{-6}$$

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

**Optimal measurements to probe the SM:**

not observed yet, possible enhancement from contributions beyond the SM.

# $b \to s\nu\bar{\nu}$ transitions




## In the Standard Model (SM):

- **flavour-changing neutral-current transitions (FCNCs);**
- can occur only at the loop level, **highly suppressed;**
- accurate theoretical predictions.

**[T. Blake, et al., Prog. Part. Nucl. Phys. 92, 50 (2017)]**

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

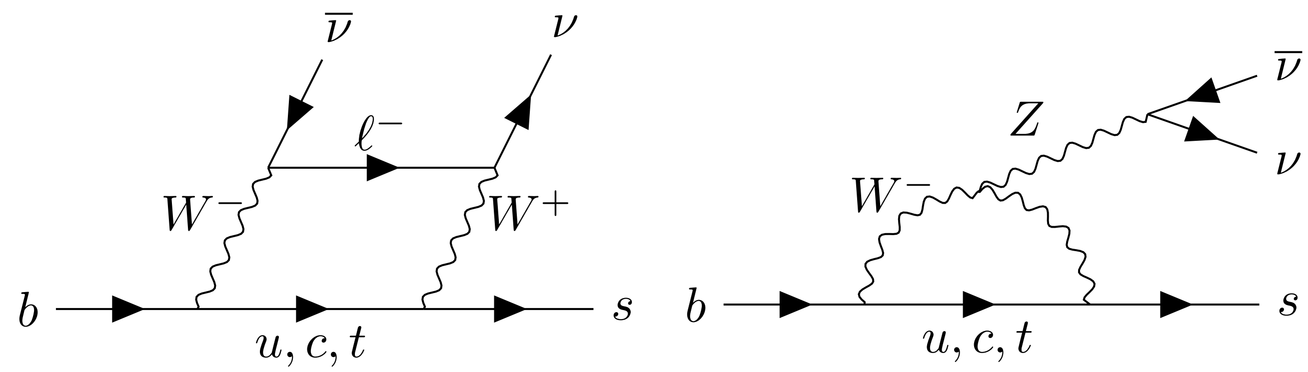
$$\mathrm{BR}(\mathrm{B} \to \mathrm{K}^{*}\nu\bar{\nu})_{\mathrm{SM}} = (8.4 \pm 1.5) \times 10^{-6}$$

**Optimal measurements to probe the SM:**

not observed yet, possible enhancement from contributions beyond the SM.

**Complementary to** $b \to s l^+ l^-$, where **tensions with the SM have been observed,** including the recent $R_K$ measurement by LHCb **[arXiv:2103.11769].**

# $b \to s\nu\bar{\nu}$ transitions




## In the Standard Model (SM):

- **flavour-changing neutral-current transitions (FCNCs);**
- can occur only at the loop level, **highly suppressed;**
- accurate theoretical predictions.

[T. Blake, et al., Prog. Part. Nucl. Phys. 92, 50 (2017)]

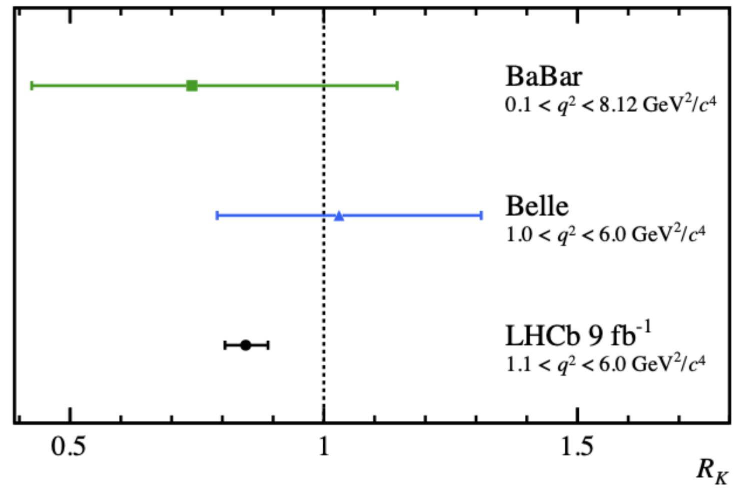
$$\mathrm{BR}(\mathrm{B} \to \mathrm{K}\nu\bar{\nu})_{\mathrm{SM}} = (4.6 \pm 0.5) \times 10^{-6}$$

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

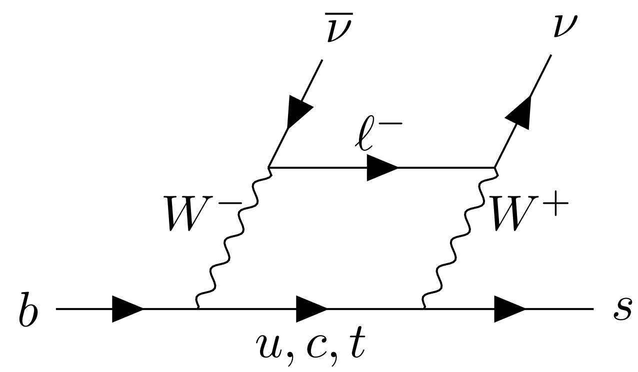
**Optimal measurements to probe the SM:**

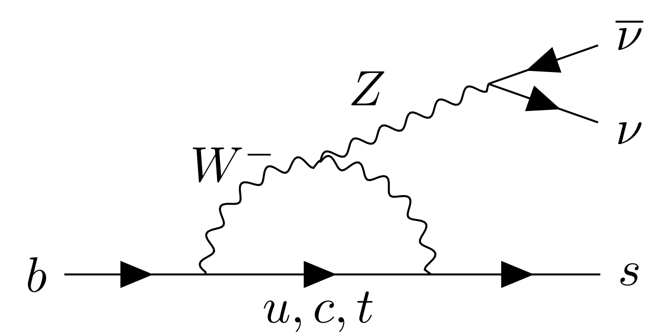
not observed yet, possible enhancement from contributions beyond the SM.

**Complementary to** $b \to sl^+l^-$, where **tensions with the SM have been observed,** including the recent $R_K$ measurement by LHCb **[arXiv:2103.11769].**



# $b \to s\nu\bar{\nu}$ transitions




## In the Standard Model (SM):

- **flavour-changing neutral-current transitions (FCNCs);**
- can occur only at the loop level, **highly suppressed;**
- accurate theoretical predictions.

**[T. Blake, et al., Prog. Part. Nucl. Phys. 92, 50 (2017)]**

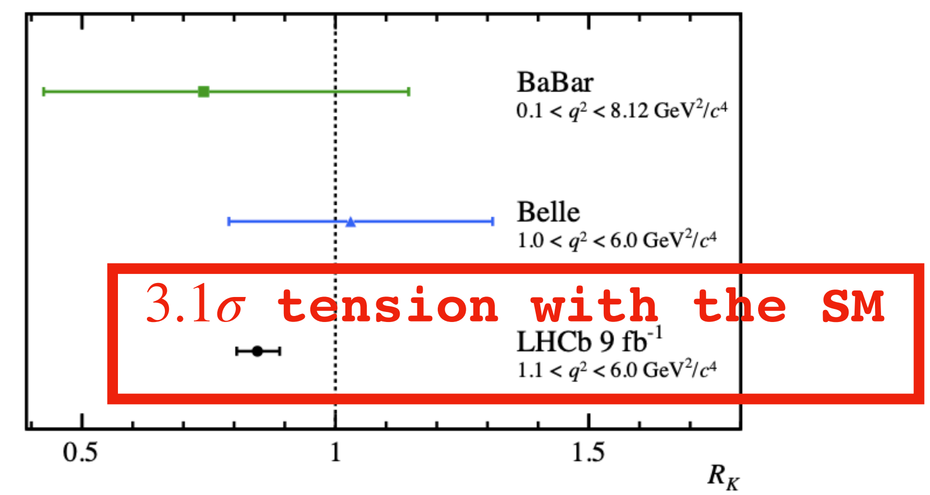
$$\mathrm{BR}(\mathrm{B} \to \mathrm{K}\nu\bar{\nu})_{\mathrm{SM}} = (4.6 \pm 0.5) \times 10^{-6}$$

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

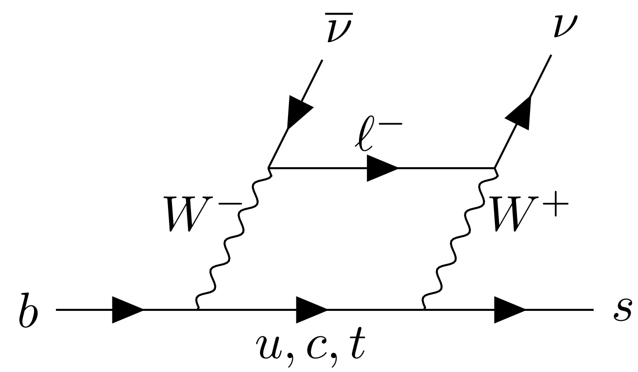
**Optimal measurements to probe the SM:**

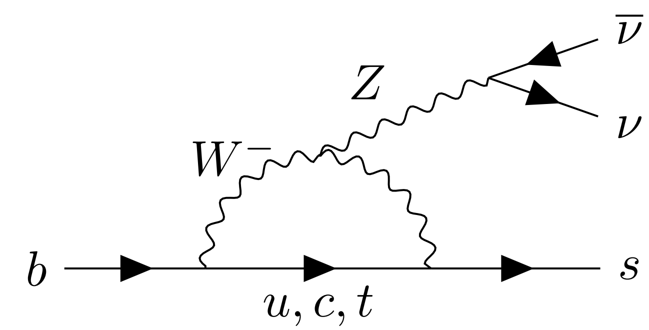
not observed yet, possible enhancement from contributions beyond the SM.

**Complementary to** $b \to sl^+l^-$, where **tensions with the SM have been observed,** including the recent $R_K$ measurement by LHCb **[arXiv:2103.11769].**



# $b \to s\nu\bar{\nu}$ transitions




## In the Standard Model (SM):

- **flavour-changing neutral-current transitions (FCNCs);**
- can occur only at the loop level, **highly suppressed;**
- accurate theoretical predictions.

**[T. Blake, et al., Prog. Part. Nucl. Phys. 92, 50 (2017)]**

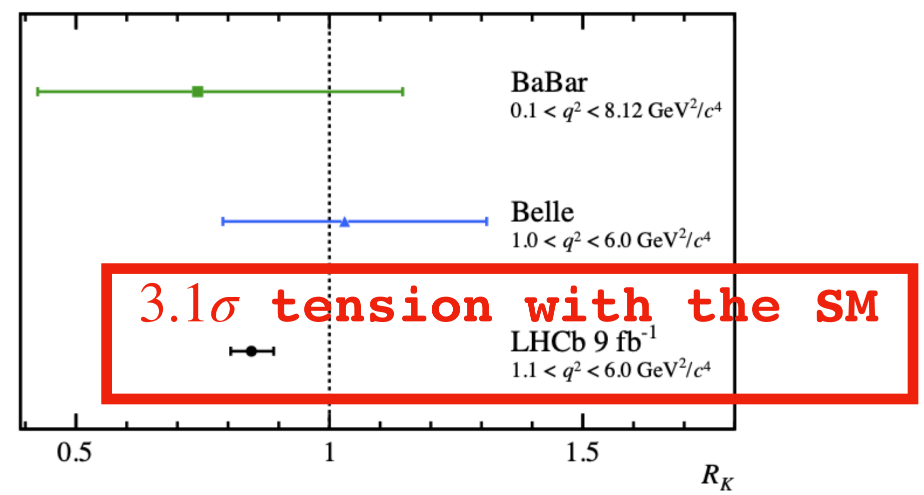
$$\mathrm{BR}(B \to K\nu\bar{\nu})_{\mathrm{SM}} = (4.6 \pm 0.5) \times 10^{-6}$$

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

**Optimal measurements to probe the SM:**

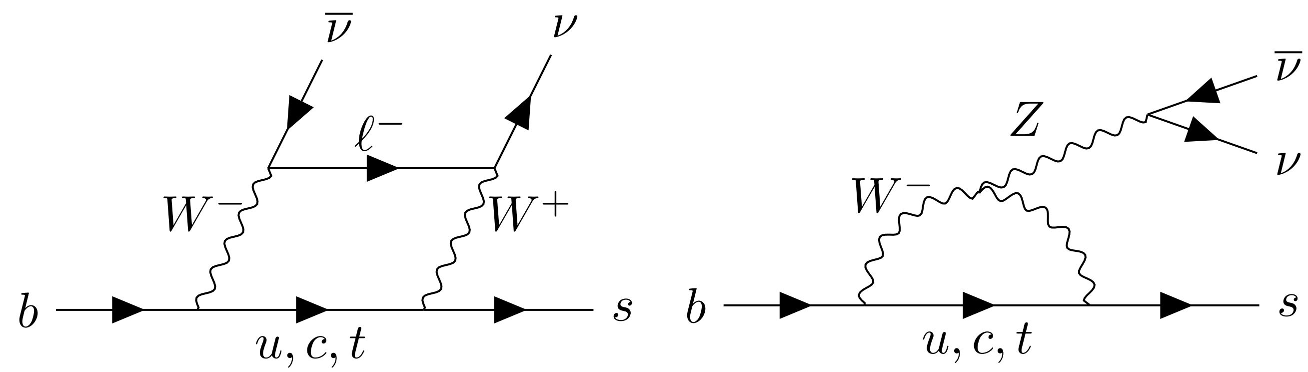
not observed yet, possible enhancement from contributions beyond the SM.

**Complementary to** $b \to sl^+l^-$, where **tensions with the SM have been observed,** including the recent $R_K$ measurement by LHCb **[arXiv:2103.11769].**




## To constrain scenarios beyond the SM:

# $b \to s\nu\bar{\nu}$ transitions




## In the Standard Model (SM):

- **flavour-changing neutral-current transitions (FCNCs);**
- can occur only at the loop level, **highly suppressed;**
- accurate theoretical predictions.

**[T. Blake, et al., Prog. Part. Nucl. Phys. 92, 50 (2017)]**

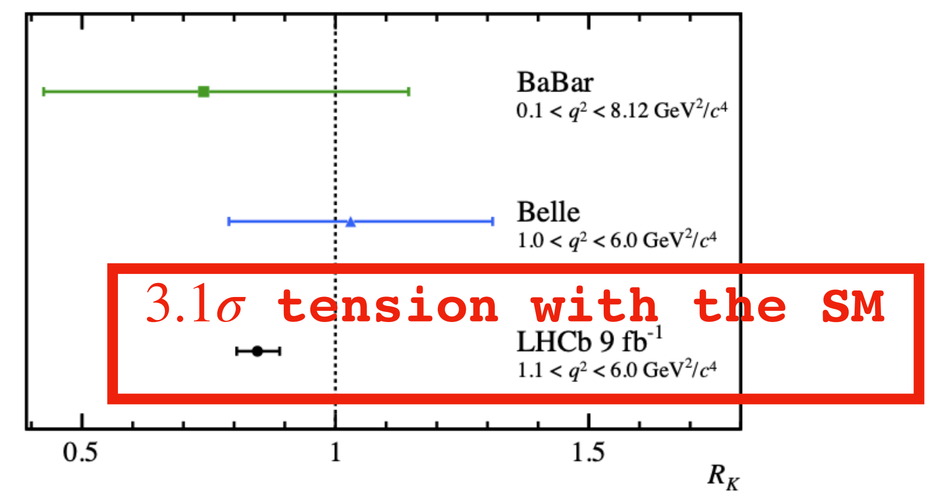
$$\mathrm{BR}(B \to K\nu\bar{\nu})_{\mathrm{SM}} = (4.6 \pm 0.5) \times 10^{-6}$$

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

**Optimal measurements to probe the SM:**

not observed yet, possible enhancement from contributions beyond the SM.

**Complementary to** $b \to s l^+ l^-$, where **tensions with the SM have been observed,** including the recent $R_K$ measurement by LHCb **[arXiv:2103.11769].**




## To constrain scenarios beyond the SM:

- **Dark matter [PRD 98, 055003 (2018)];**
- **Leptoquarks [PRD 102, 015023 (2020)];**
- **Axions [PRD 101, 095006 (2020)].**

# $b \to s\nu\bar{\nu}$ transitions

Challenging measurements:

# $b \to s\nu\bar{\nu}$ transitions

## Challenging measurements:

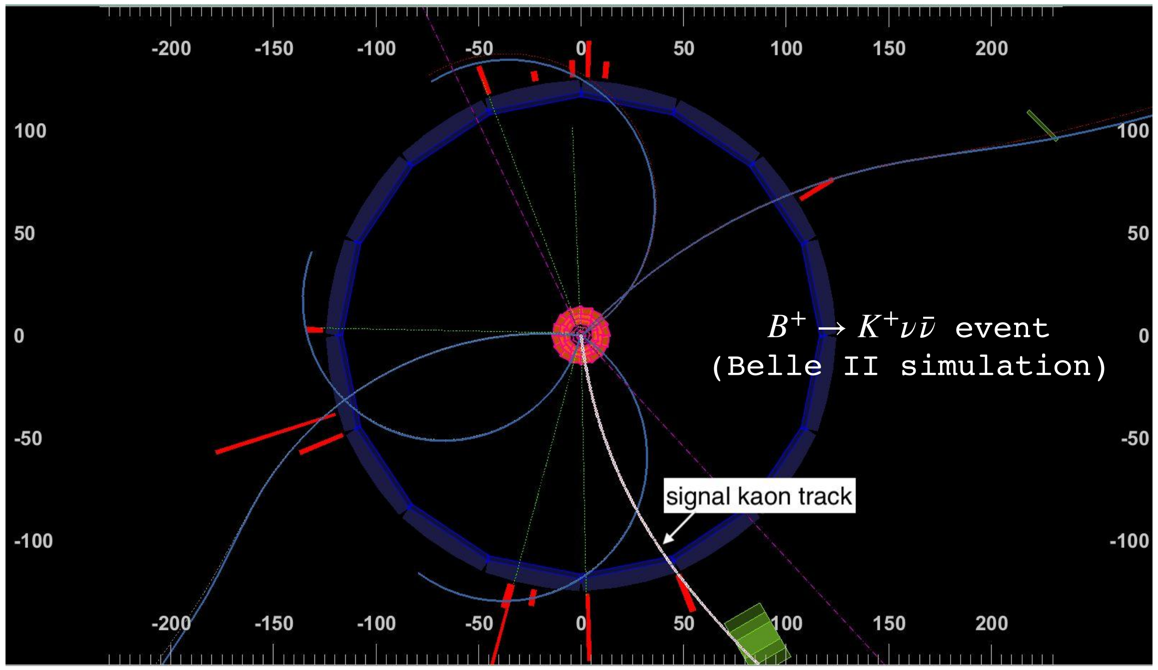
- decays with **2 neutrinos** in the final state leaving **no signature in the detector;**

# $b \to s\nu\bar{\nu}$ transitions

## Challenging measurements:

- decays with **2 neutrinos** in the final state leaving **no signature in the detector;**

- **can be measured at *B* factories like Belle II** because of the clean event environment and the well defined initial state.



# Search for $B^+ \to K^+\nu\bar{\nu}$ decays using an inclusive tagging method at Belle II

[arXiv:2104.12624] – submitted to PRL

# Previous searches for $B^+ \to K^+\nu\bar{\nu}$

# Previous searches for $B^+ \to K^+ \nu\bar{\nu}$

The **previous** studies all adopted an
**explicit reconstruction of the $B_{\text{tag}}$**

# Previous searches for $B^+ \to K^+ \nu \bar{\nu}$

$e^- \to \Upsilon(4S) \leftarrow e^+$

The **previous** studies all adopted an **explicit reconstruction of the $B_{\text{tag}}$**

# Previous searches for $B^+ \to K^+ \nu \bar{\nu}$

The **previous** studies all adopted an **explicit reconstruction of the $B_{\rm tag}$**

$e^- \to \Upsilon(4S) \leftarrow e^+$

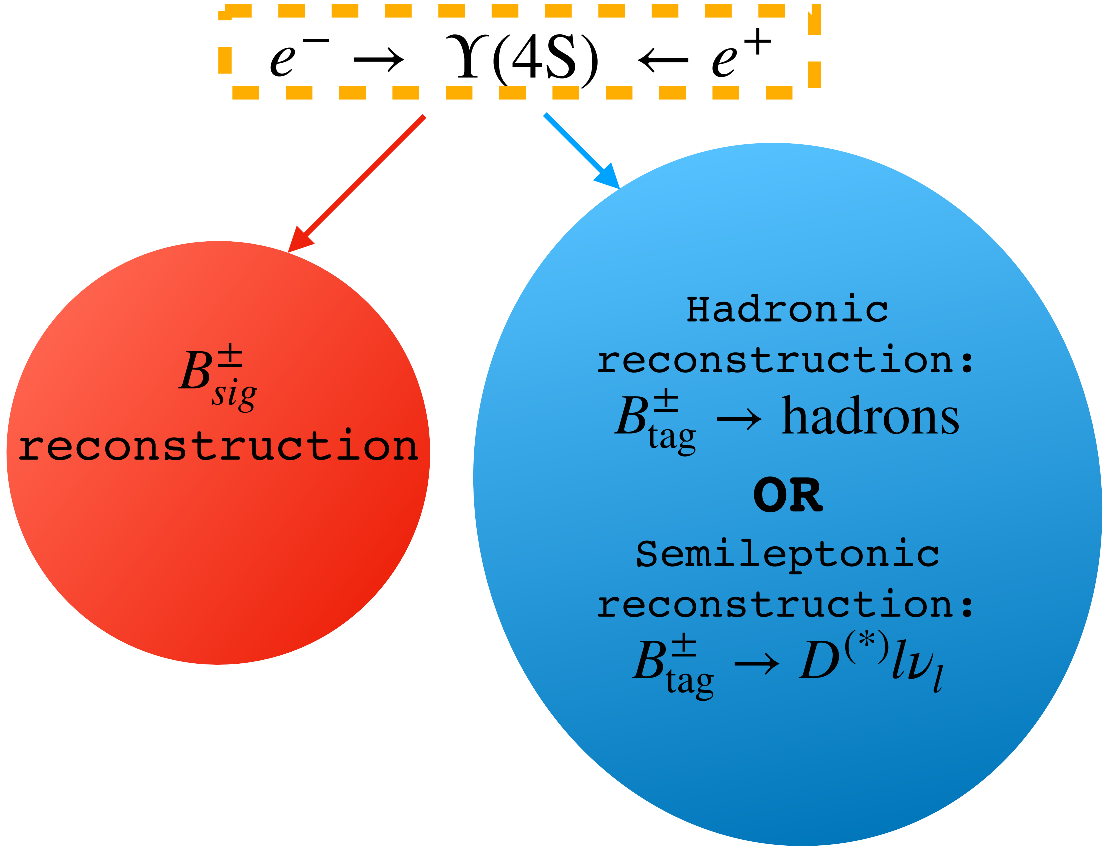
Hadronic reconstruction:
$B^{\pm}_{\rm tag} \to$ hadrons

**OR**

Semileptonic reconstruction:
$B^{\pm}_{\rm tag} \to D^{(*)} l \nu_l$

# Previous searches for $B^+ \to K^+ \nu\bar{\nu}$

The **previous** studies all adopted an **explicit reconstruction of the $B_{\text{tag}}$** **followed by the signal reconstruction.**



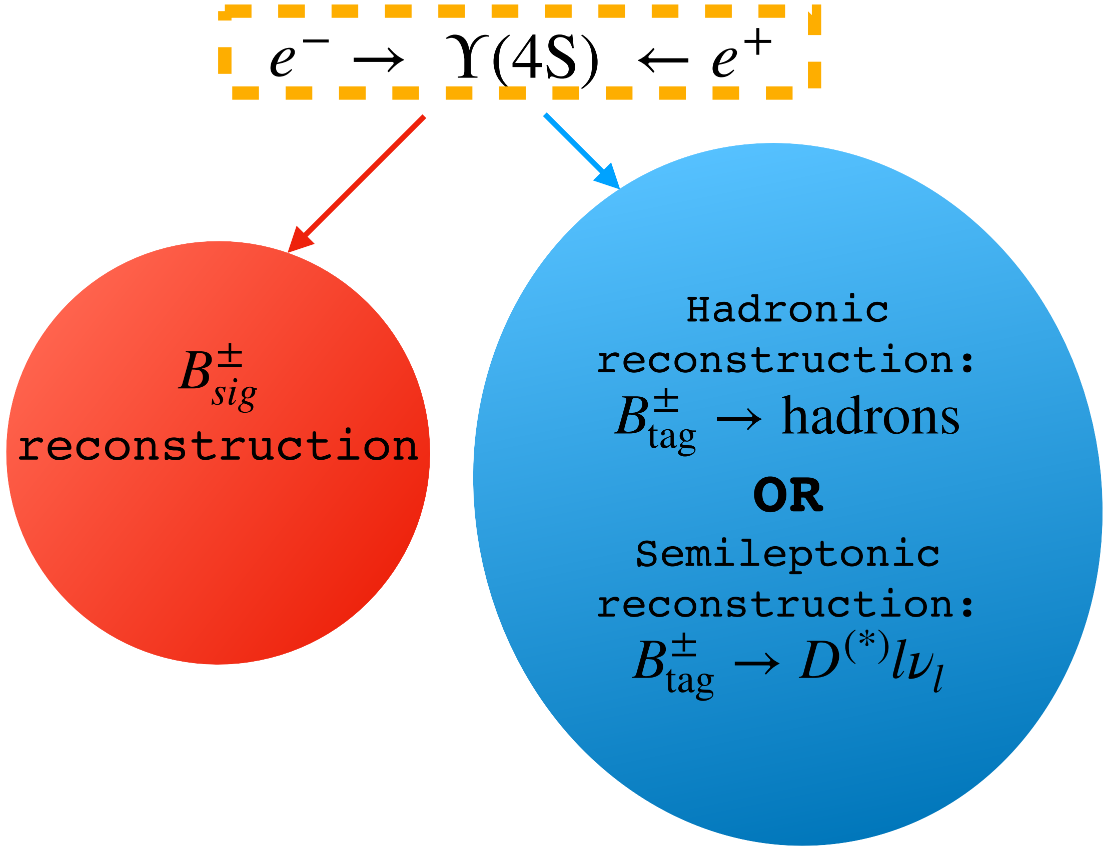
# Previous searches for $B^+ \to K^+ \nu\bar{\nu}$

The **previous** studies all adopted an **explicit reconstruction of the $B_{\text{tag}}$** **followed by the signal reconstruction.**

$\downarrow$

**Low reconstruction efficiency** because of the low tag-reconstruction efficiency:

- hadronic tagging $\epsilon_{sig} \cdot \epsilon_{tag} \sim 0.04\,\%$
- semileptonic tagging $\epsilon_{sig} \cdot \epsilon_{tag} \sim 0.2\,\%$



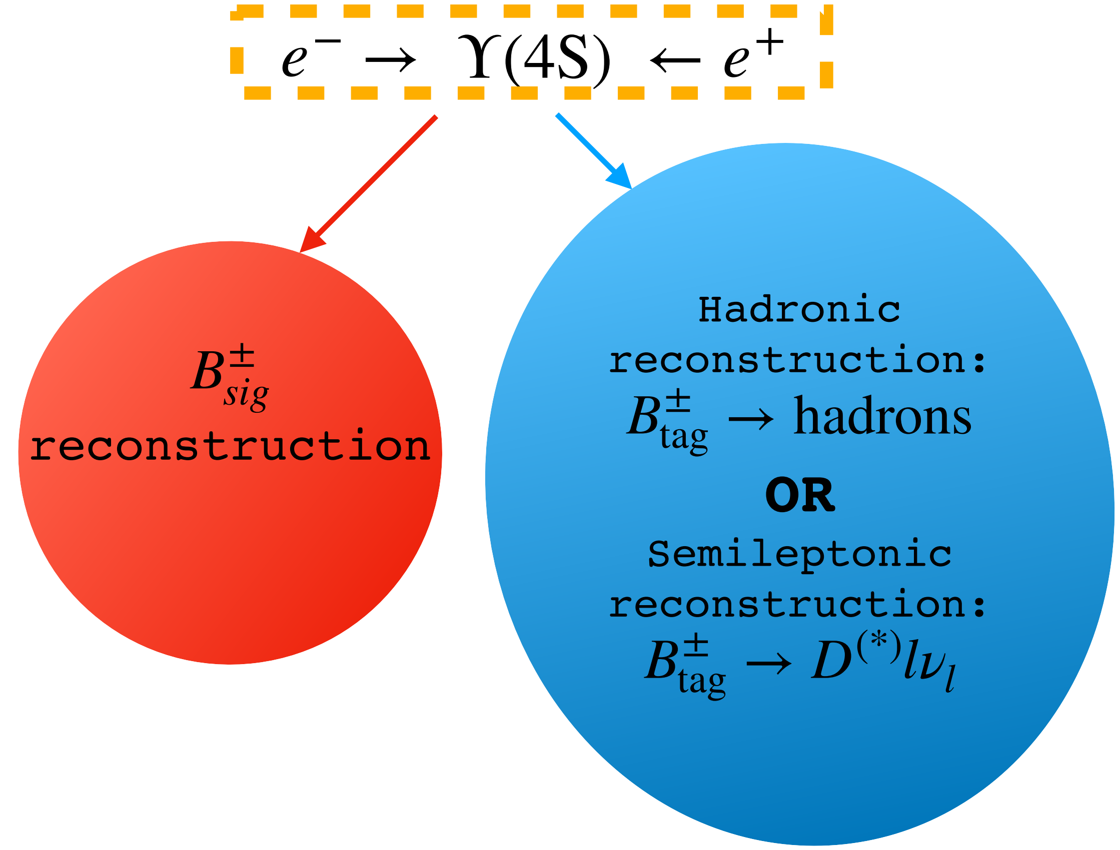
# Previous searches for $B^+ \to K^+ \nu\bar{\nu}$

The **previous** studies all adopted an **explicit reconstruction of the $B_{\text{tag}}$** **followed by the signal reconstruction.**

$\downarrow$

**Low reconstruction efficiency** because of the low tag-reconstruction efficiency:

- hadronic tagging $\epsilon_{sig} \cdot \epsilon_{tag} \sim 0.04\,\%$
- semileptonic tagging $\epsilon_{sig} \cdot \epsilon_{tag} \sim 0.2\,\%$

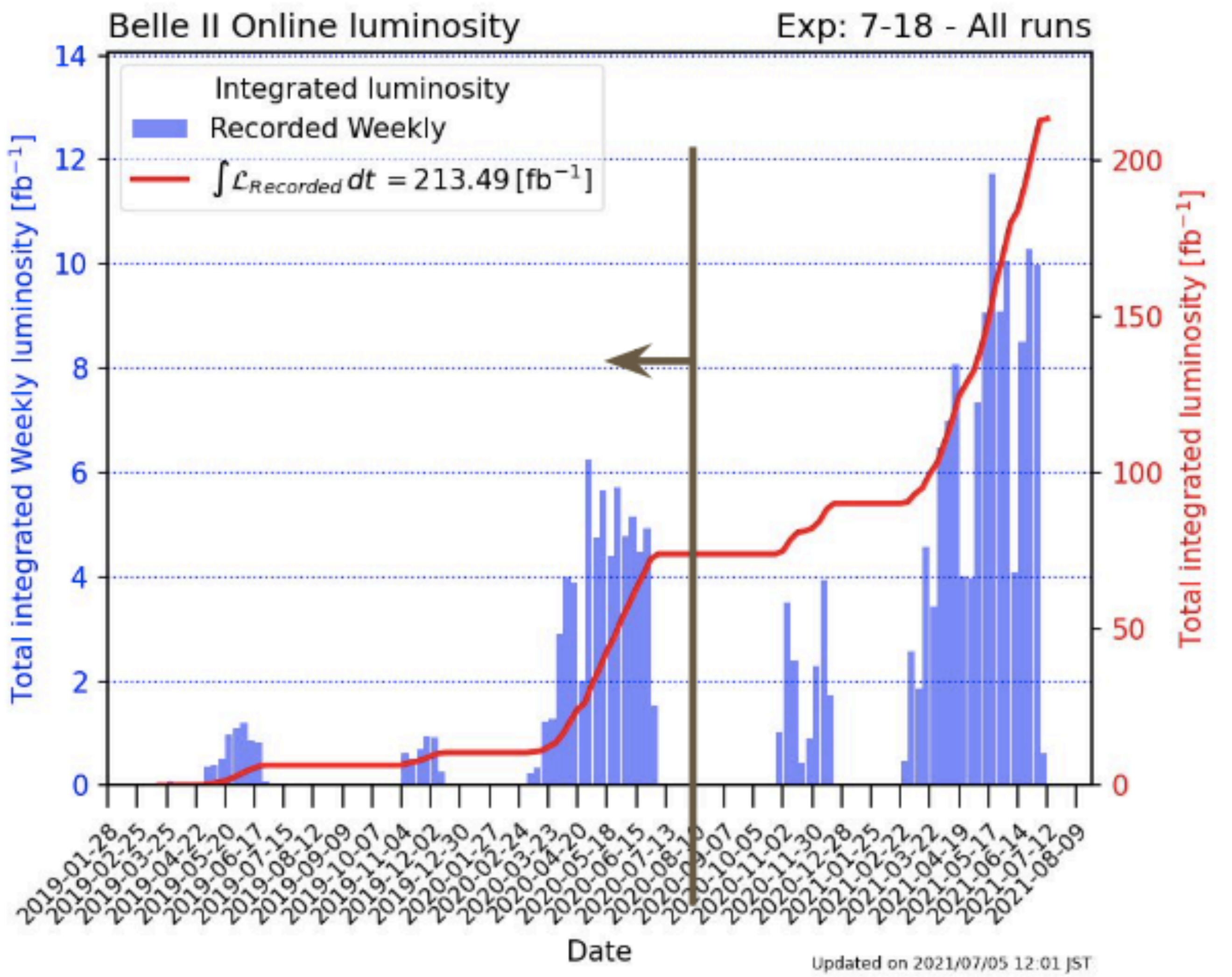



**Upper limits on the branching ratios were set:**

| Experiment | Year | Observed limit on $\text{BR}(B^+ \to K^+ \nu\bar{\nu})$ | Approach | Data $[\text{fb}^{-1}]$ |
|---|---|---|---|---|
| **BABAR** | **2013** | $< 1.6 \times 10^{-5}$ **[Phys.Rev.D87,112005]** | **SL + Had tagging** | 429 |
| **Belle** | **2013** | $< 5.5 \times 10^{-5}$ **[Phys.Rev.D87,111103(R)]** | **Had tagging** | 711 |
| **Belle** | **2017** | $< 1.9 \times 10^{-5}$ **[Phys.Rev.D96,091101(R)]** | **SL tagging** | 711 |

# Luminosity and data sample at Belle II

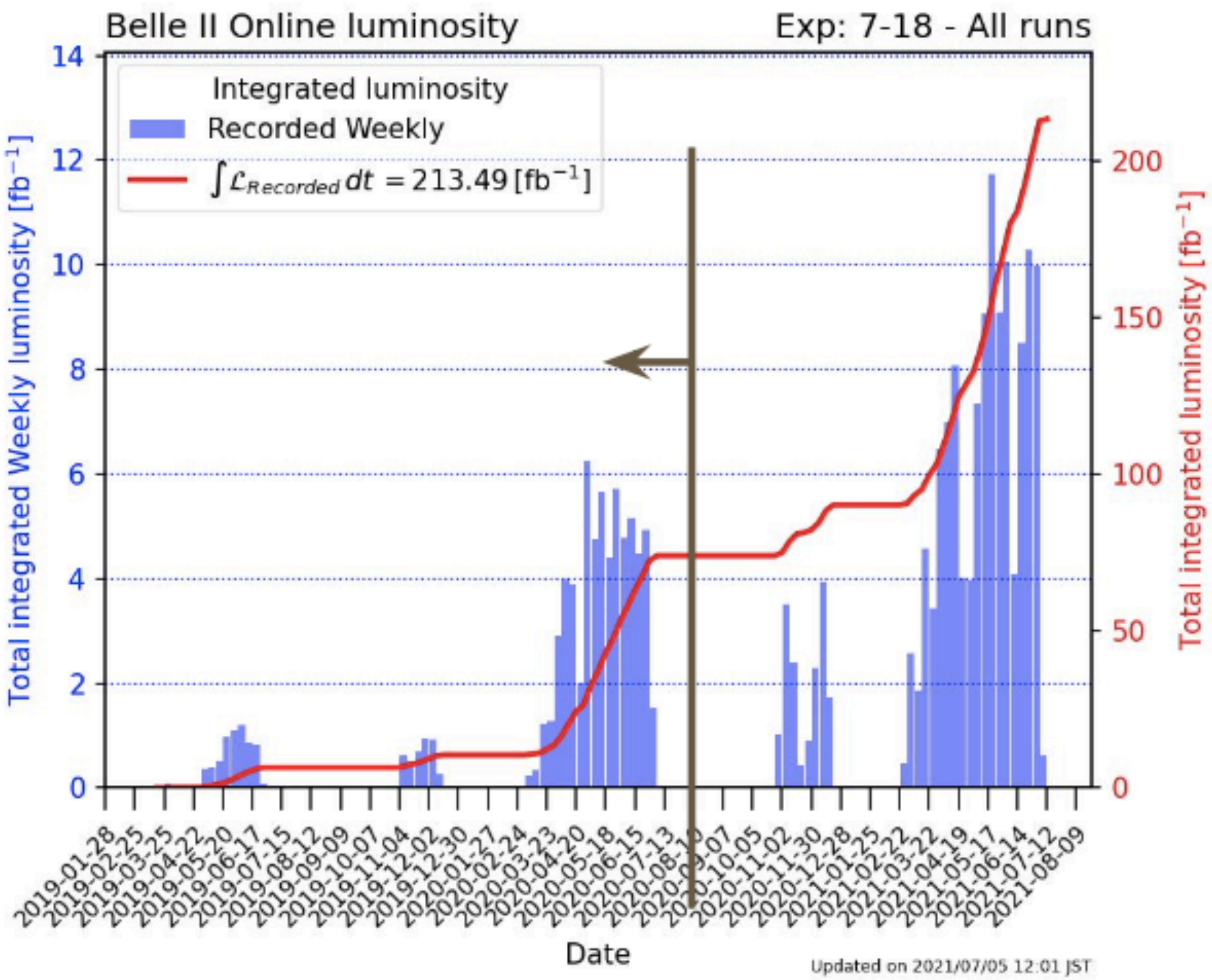
# Luminosity and data sample at Belle II




Phase III data

- **Data used in the analysis collected in 2019-2020(summer):**
  - $63\ \mathrm{fb}^{-1}$ **of data collected at** $\sqrt{s} \to \Upsilon(4S)$ **resonance** $\sim 68$ **million** $B\bar{B}$ **pairs.**
  - $9\ \mathrm{fb}^{-1}$ **of (off-resonance) data collected 60 MeV below the** $\Upsilon(4S)$ **resonance for background studies.**

# Luminosity and data sample at Belle II




Phase III data

- **Data used in the analysis collected in 2019-2020(summer):**
  - $63\ \mathrm{fb}^{-1}$ **of data collected at** $\sqrt{s} \to \Upsilon(4S)$ **resonance** $\sim 68$ **million** $B\bar{B}$ **pairs.**
  - $9\ \mathrm{fb}^{-1}$ **of (off-resonance) data collected 60 MeV below the** $\Upsilon(4S)$ **resonance for background studies.**
- $\sim 213\ \mathrm{fb}^{-1}$ of data collected before the summer 2021 shutdown.

# The inclusive tagging

# The inclusive tagging

## The idea

# The inclusive tagging

## The idea

- **Signal reconstructed as the highest $p_{\mathrm{T}}$ track** (correct match $\sim 80\,\%$ )

# The inclusive tagging

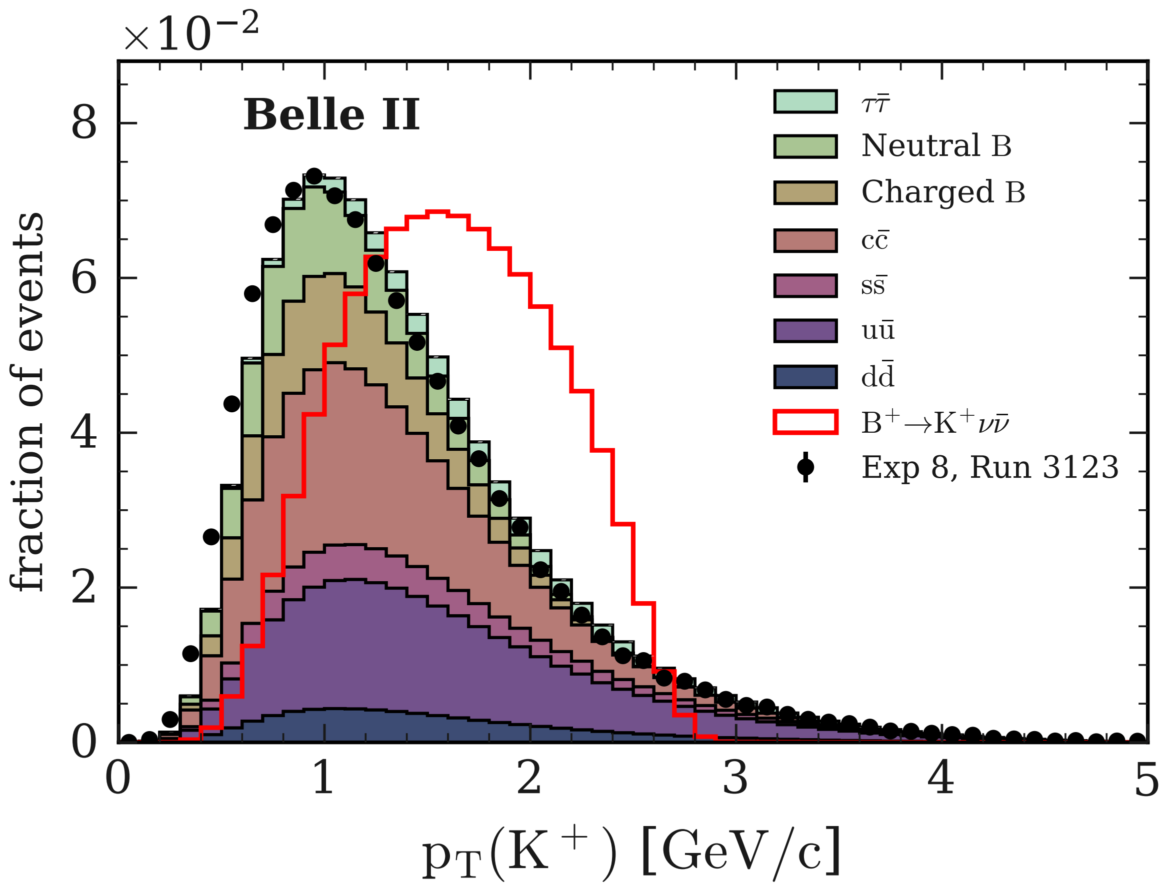
## The idea

- **Signal reconstructed as the highest** $p_{\mathrm{T}}$ **track** (correct match $\sim 80\,\%$ )

$$e^- \rightarrow \Upsilon(4S) \leftarrow e^+$$

$B^{\pm}_{sig}$

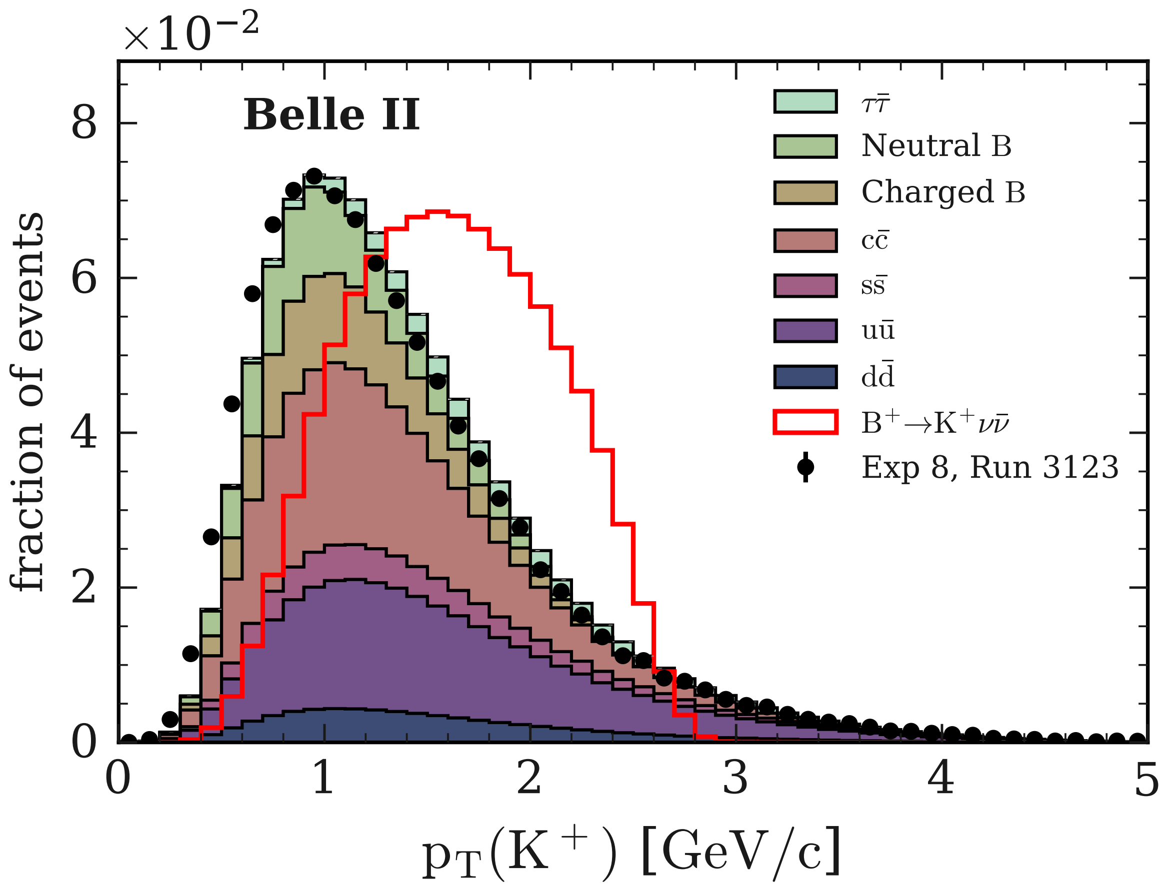
**Highest pT track**

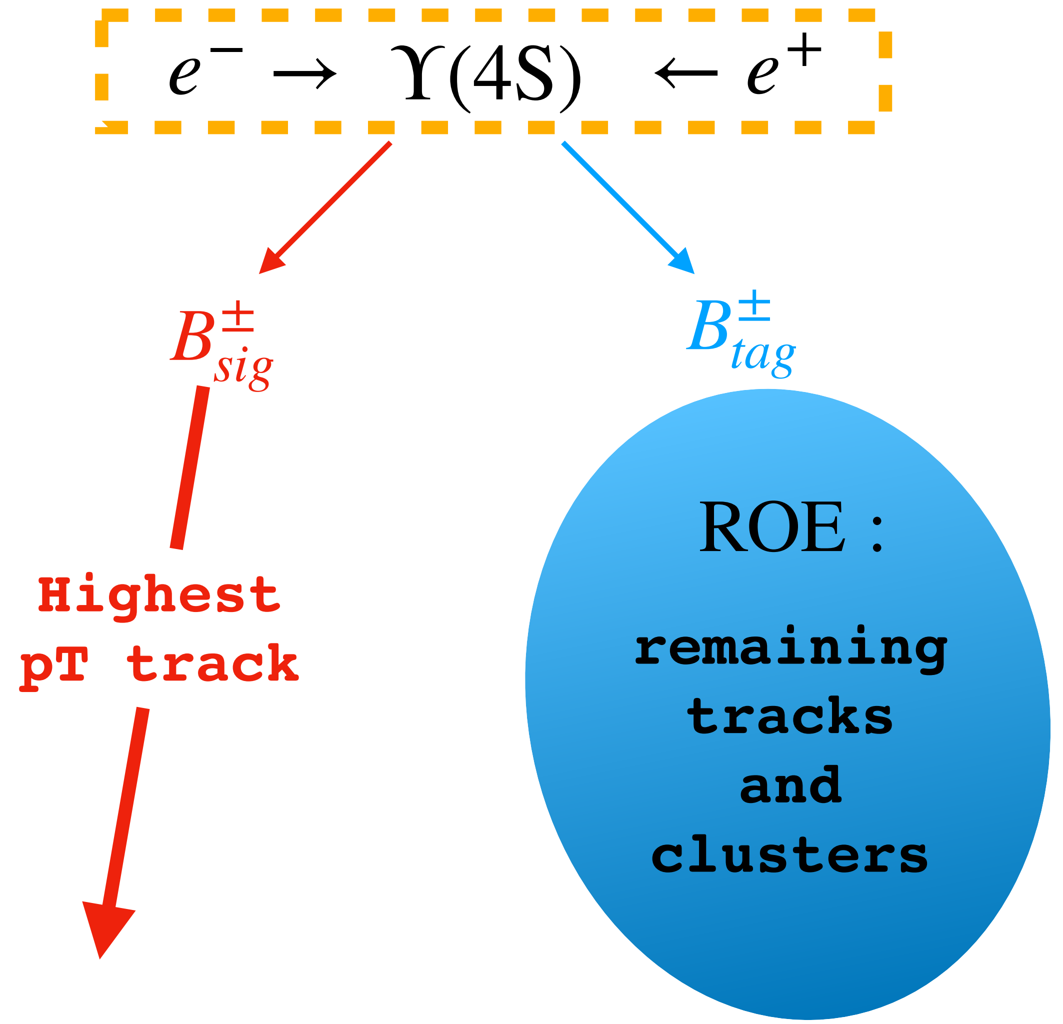


# The inclusive tagging

## The idea

- **Signal reconstructed as the highest $p_T$ track** (correct match $\sim 80\,\%$ ) **followed by inclusive reconstruction of the *rest of the event* (ROE).**

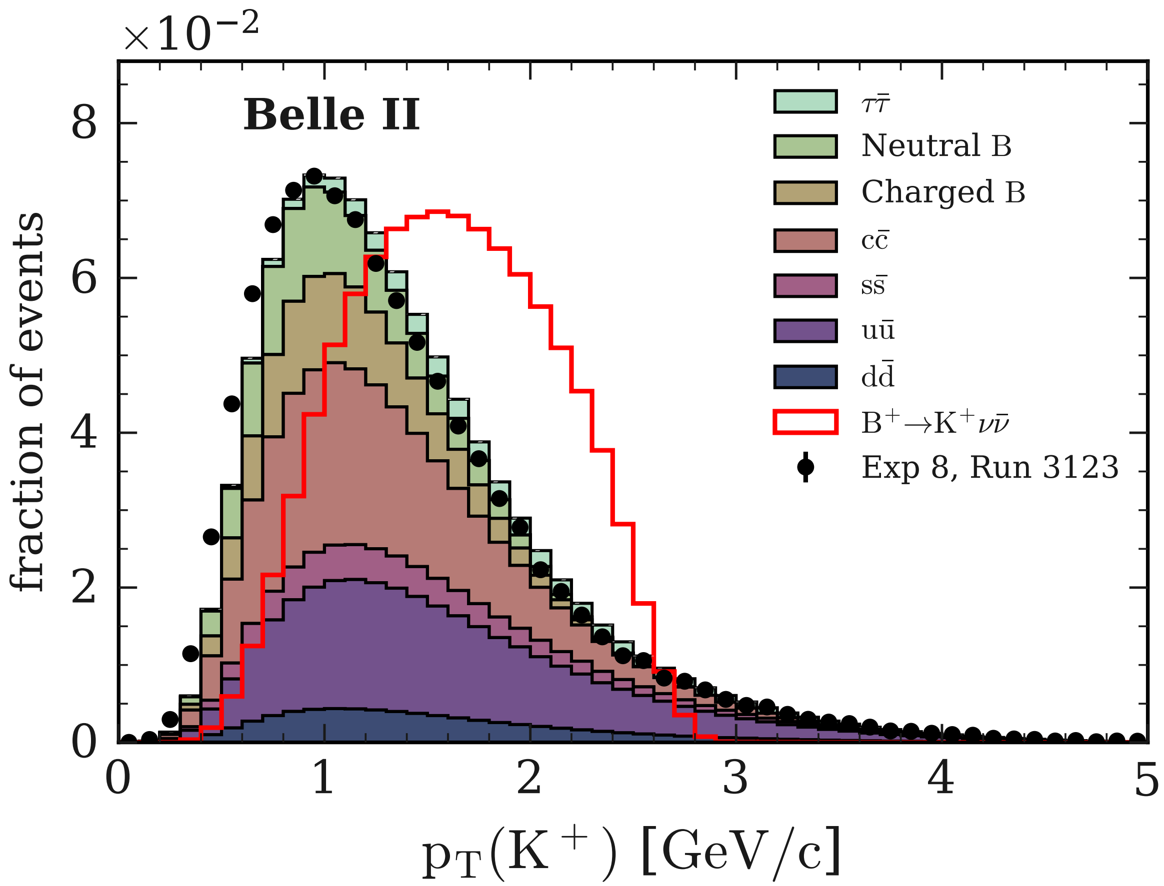


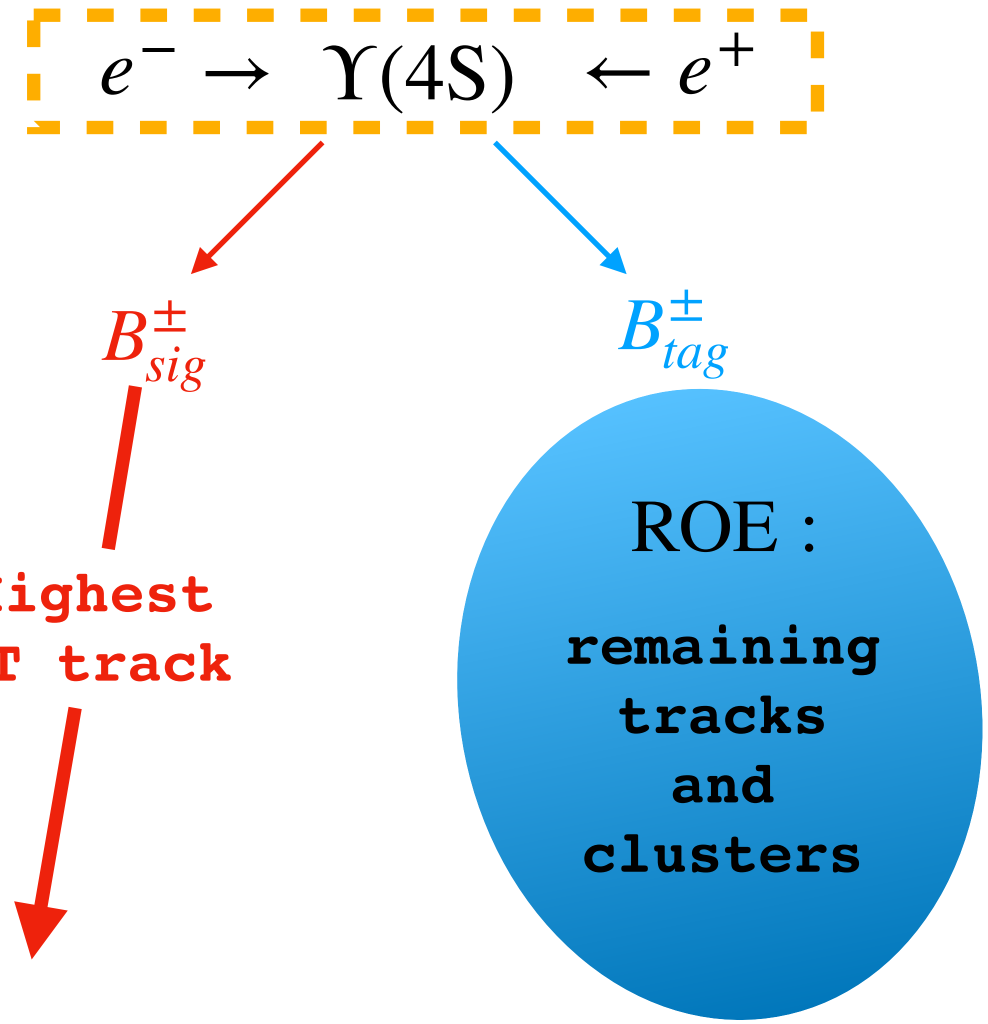



# The inclusive tagging

## The idea

- **Signal reconstructed as the highest $p_T$ track** (correct match $\sim 80\,\%$) **followed by inclusive reconstruction of the *rest of the event* (ROE).**



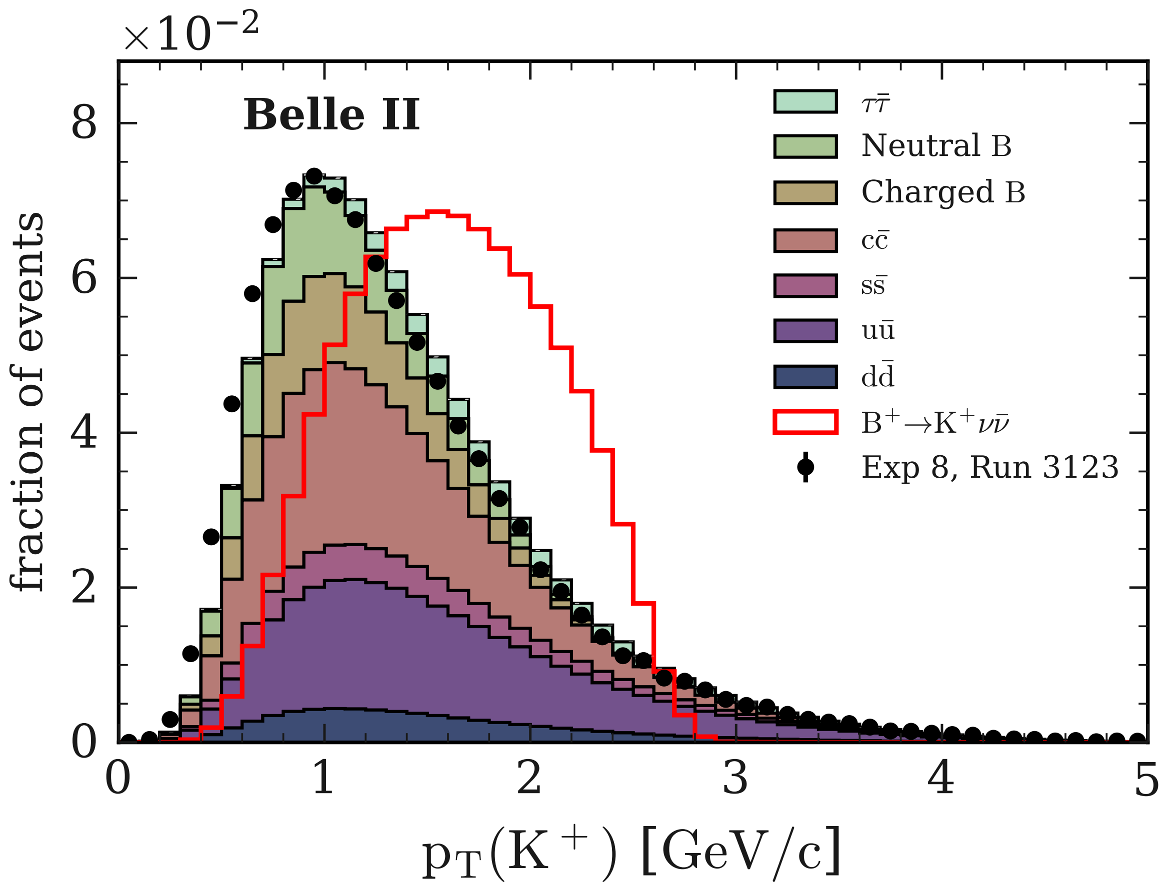




**For most background events the ROE consists of a wrong/too large combination of charged tracks/photons, while for the signal some objects can be missing.**

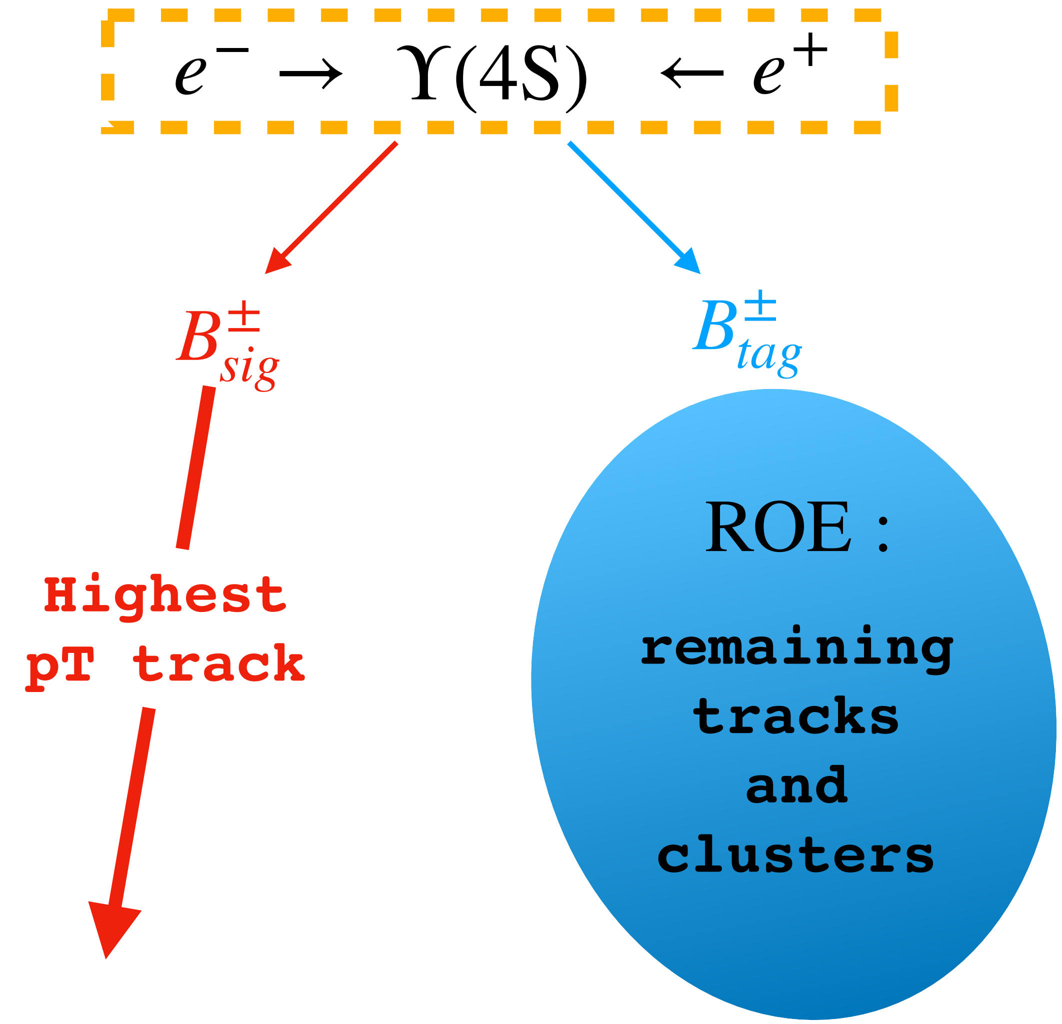
# The inclusive tagging

## The idea

- **Signal reconstructed as the highest $p_T$ track** (correct match $\sim 80\,\%$ ) **followed by 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 contributions from generic B decays and continuum ($u\bar{u}$, $d\bar{d}$, $c\bar{c}$, $s\bar{s}$).




**For most background events the ROE consists of a wrong/too large combination of charged tracks/ photons, while for the signal some objects can be missing.**

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

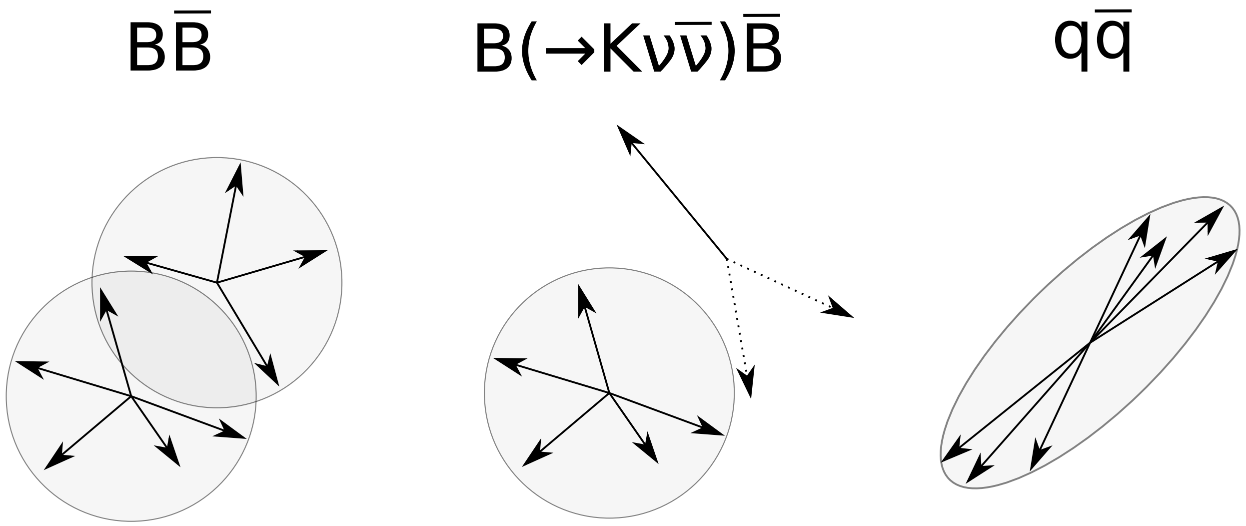
# Features of $B^+ \to K^+ \nu\bar{\nu}$

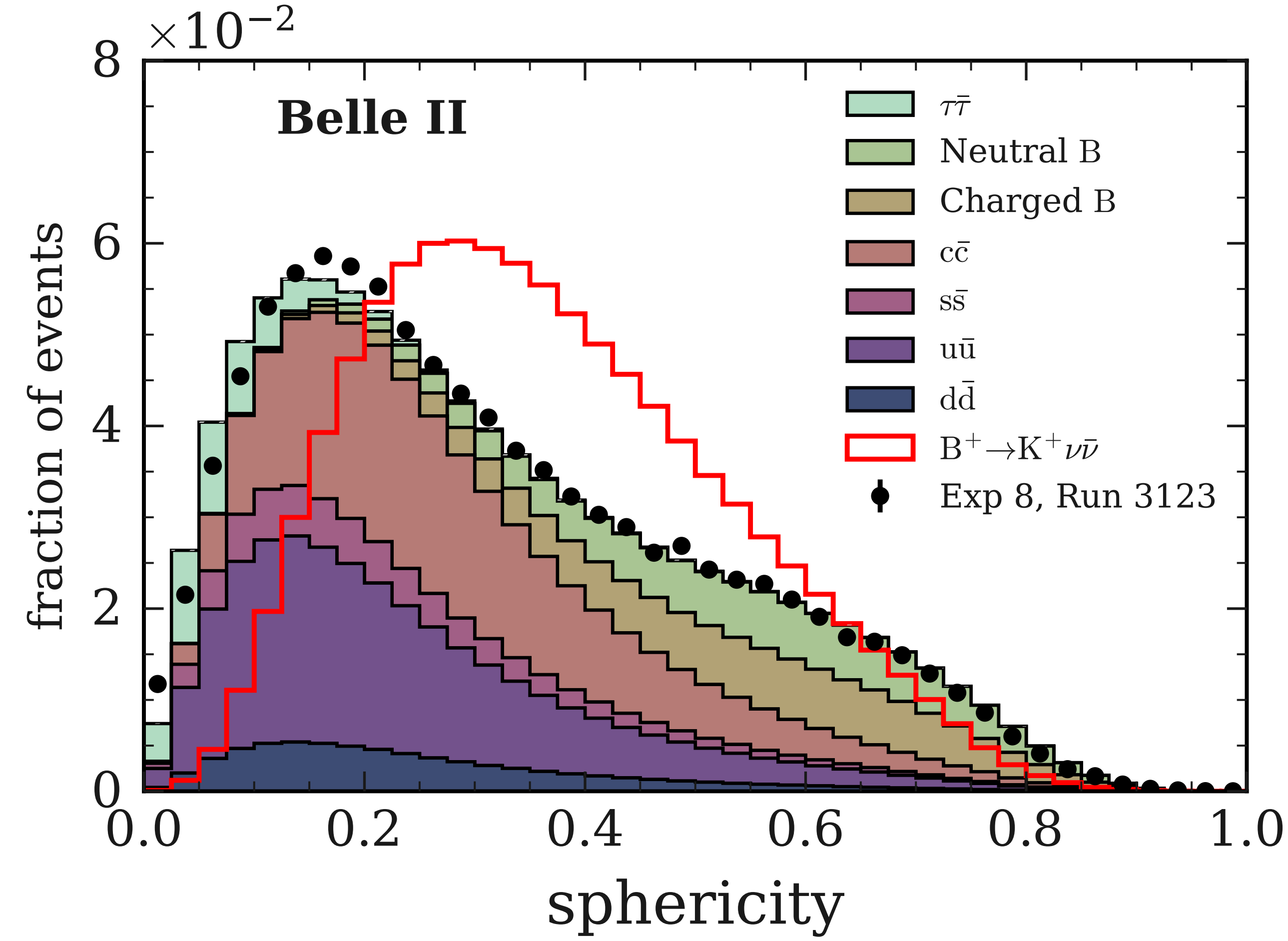
**Signal identification exploiting topological features of $B^+ \to K^+ \nu\bar{\nu}$**

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

**Signal identification exploiting topological features of $B^+ \to K^+ \nu\bar{\nu}$.**

- For example the **event shape:**

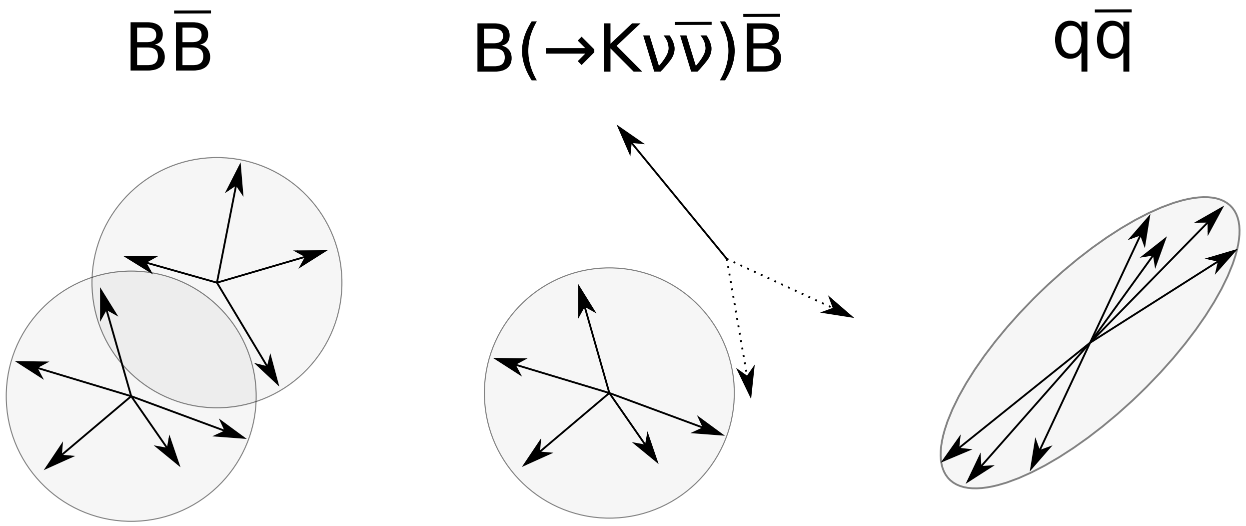


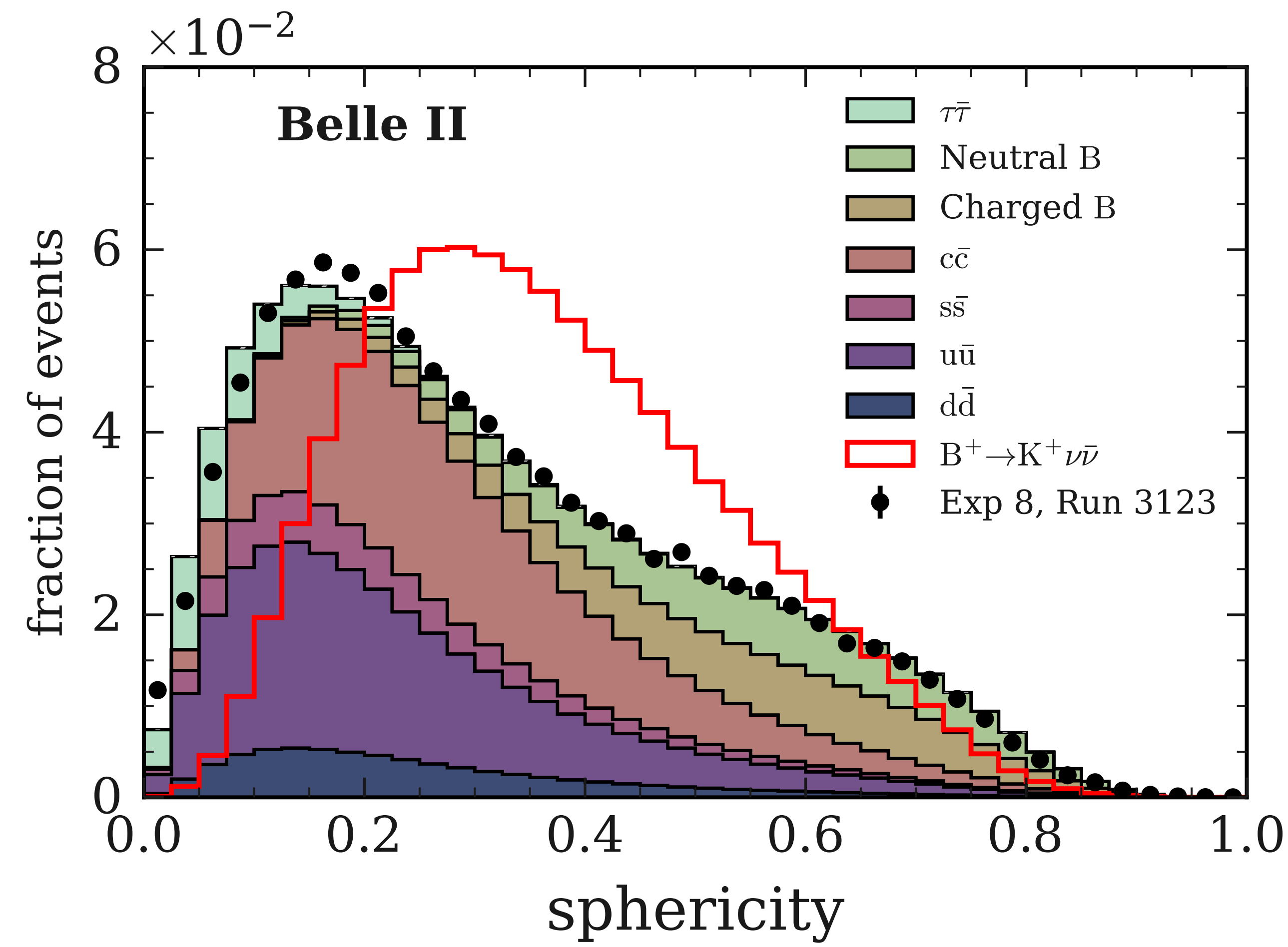



# Features of $B^+ \to K^+ \nu\bar{\nu}$

**Signal identification exploiting topological features of $B^+ \to K^+ \nu\bar{\nu}$.**

- For example the **event shape:**



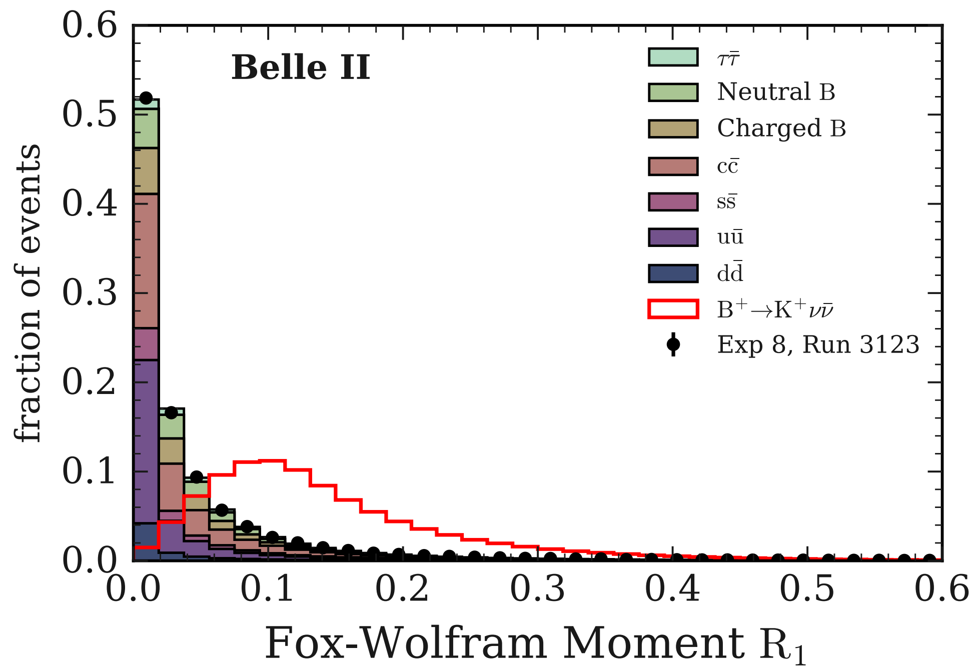




- But also:
  - other variables related to the event features;
  - variables related to the kinematics of the signal $K$ candidate;
  - variables related to the ROE;
  - variables related to the $D^0/D^+$ suppression.

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

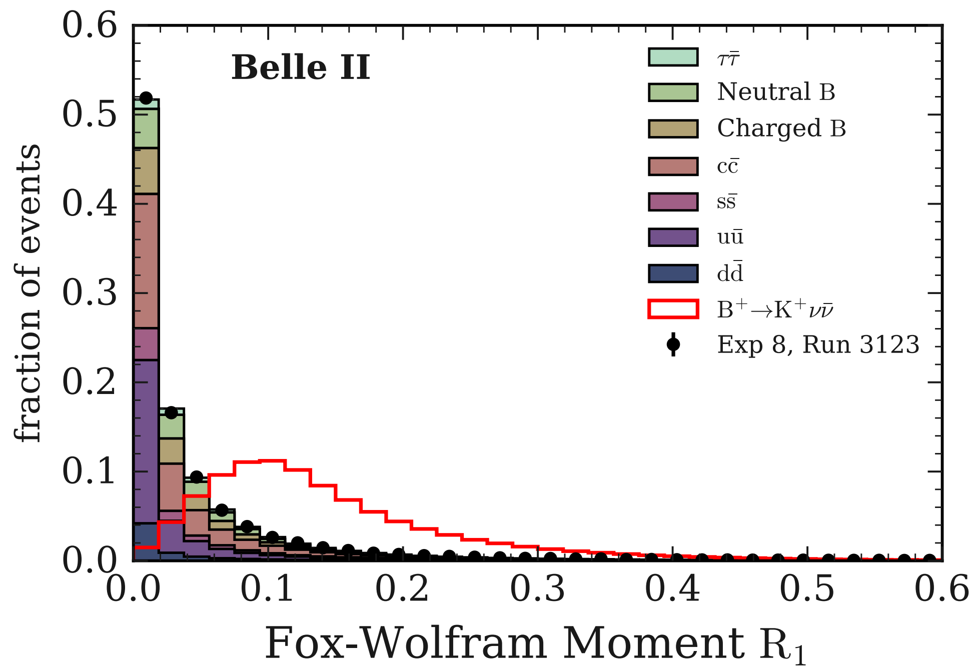
# Features of $B^+ \to K^+ \nu\bar{\nu}$

- Variables related to the event topology.

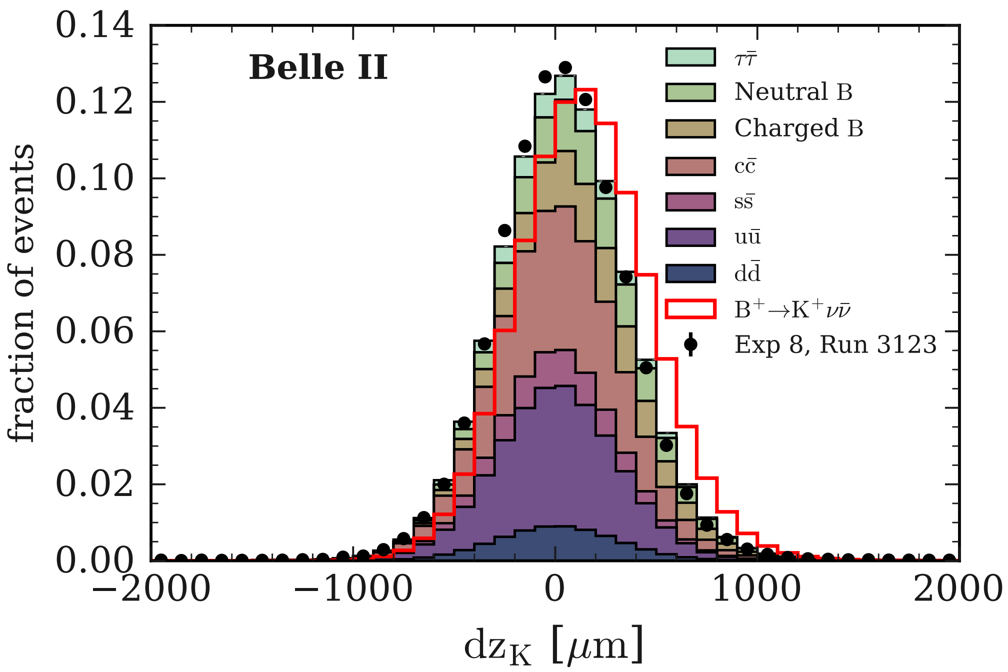


# Features of $B^+ \to K^+\nu\bar{\nu}$

- Variables related to the event topology.

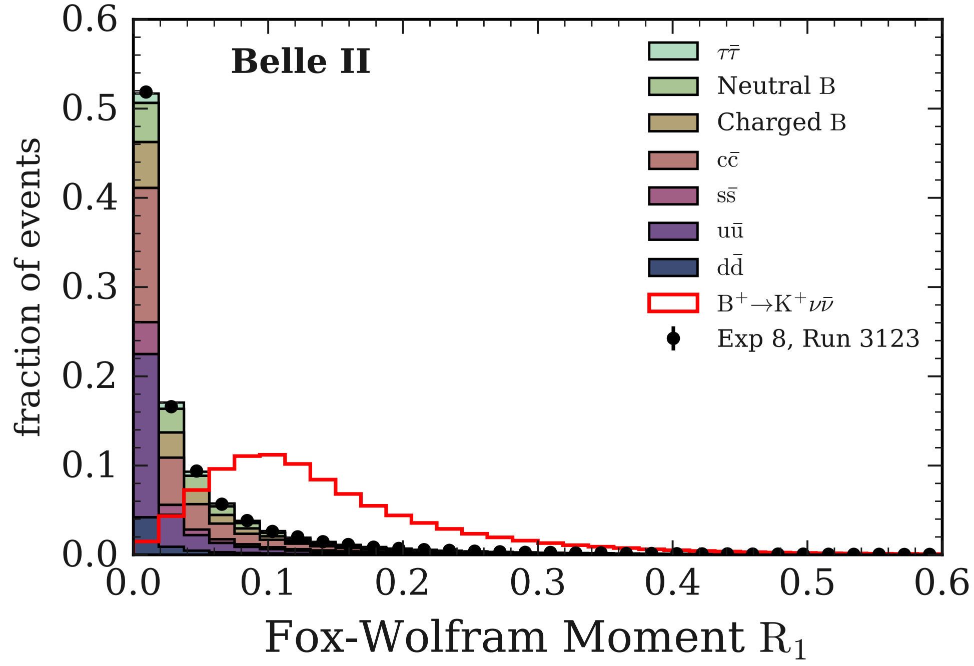



- Variables related to the signal $K^+$candidate.

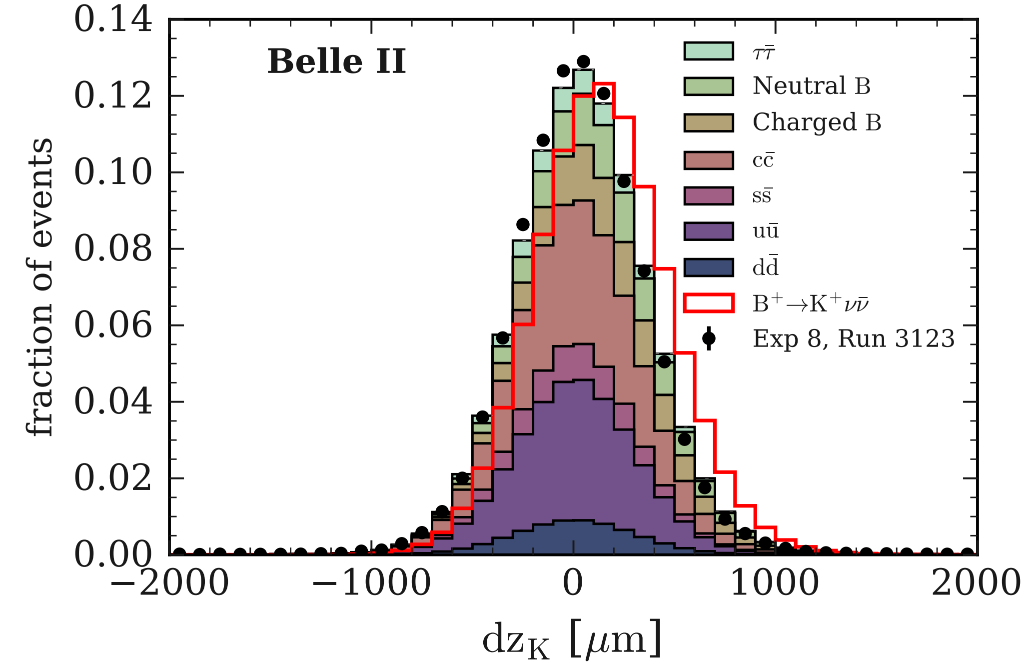


# Features of $B^+ \to K^+\nu\bar{\nu}$

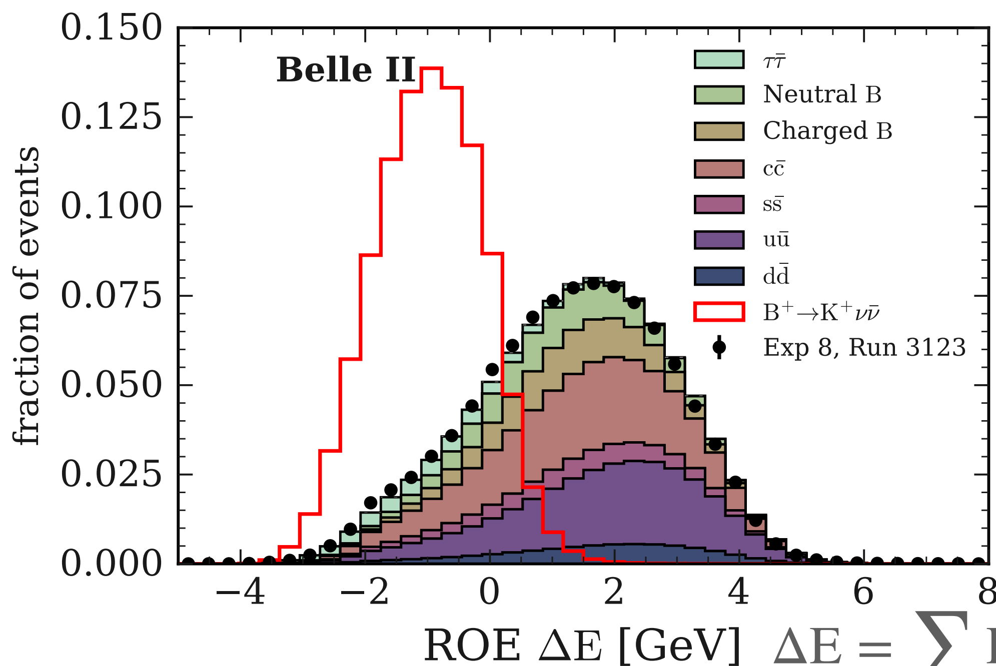
- Variables related to the event topology.




- Variables related to the signal $K^+$candidate.



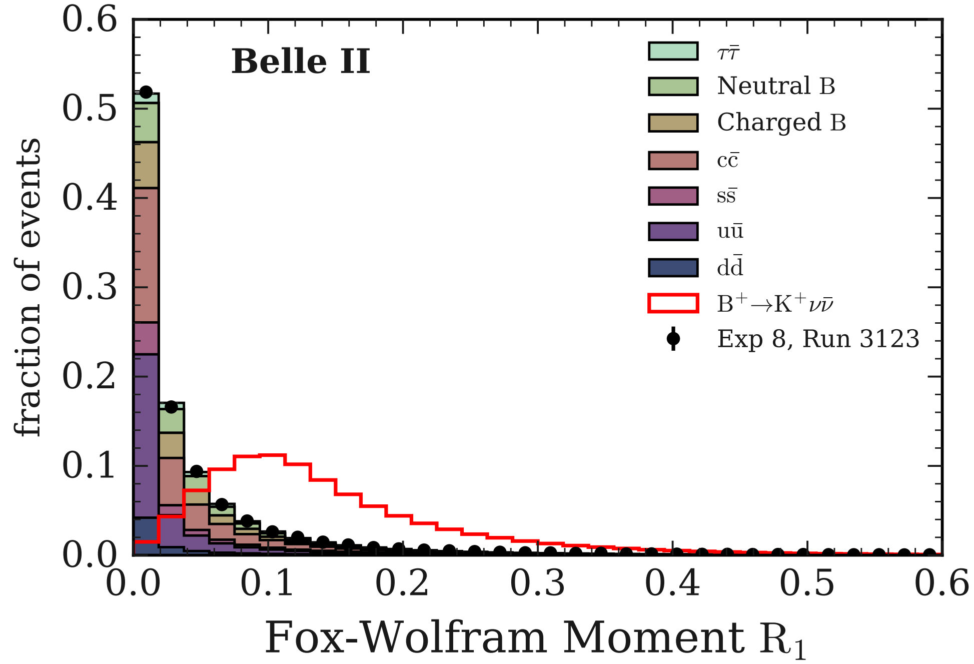

- Variables related to the ROE.



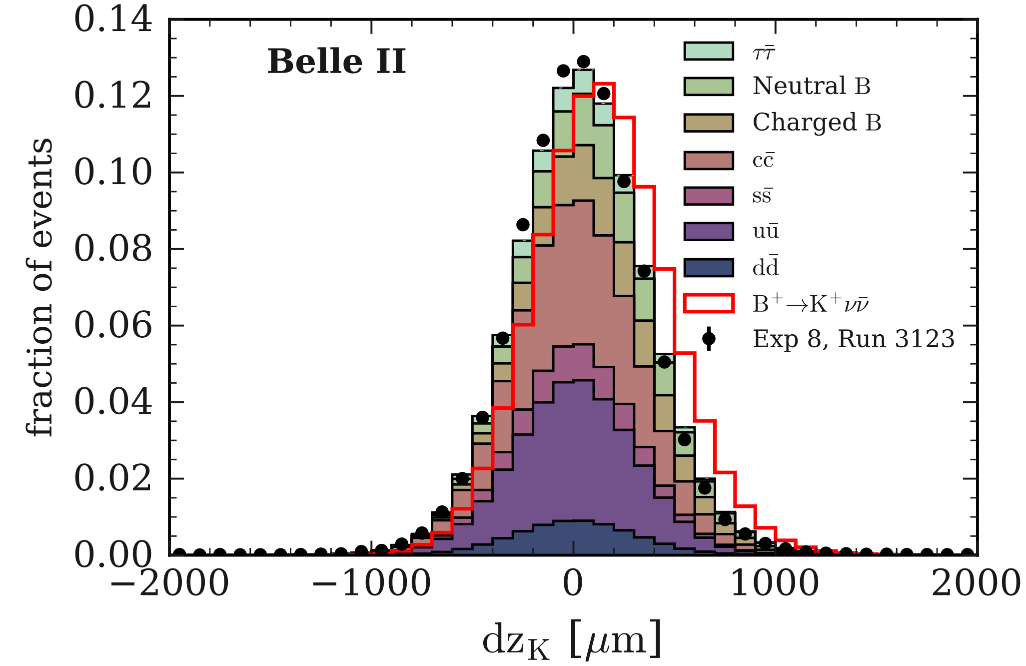

$\Delta E = \sum_i E_i^{CMS} - E_{beam}$

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

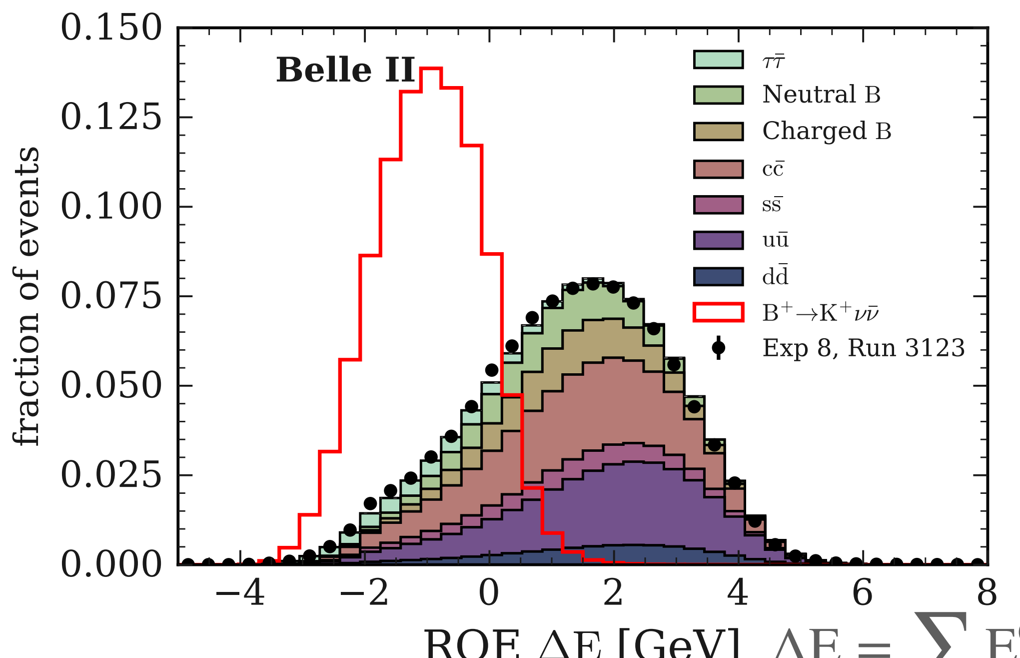
- Variables related to the event topology.




- Variables related to the signal $K^+$candidate.

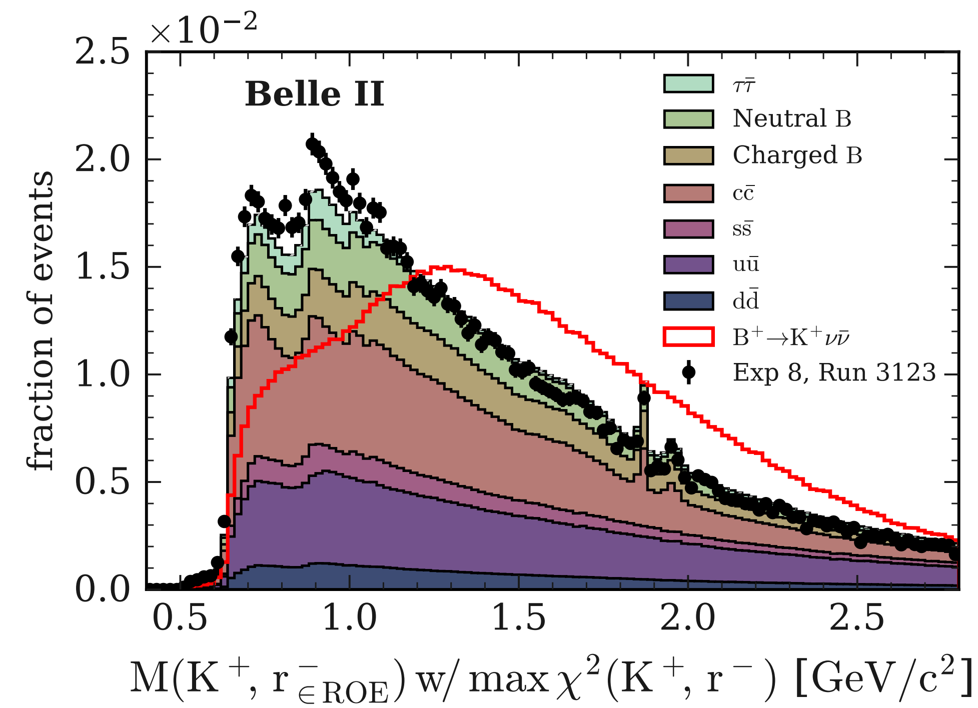



- Variables related to the ROE.




$$\Delta E = \sum_i E_i^{CMS} - E_{beam}$$

- Variables related to $D^0/D^+$ suppression.



# Multivariate classification

# Multivariate classification

51 well separating variables are used to train 2 consecutive binary classifiers (**FastBDT**) $\mathrm{BDT}_1$ and $\mathrm{BDT}_2$.

# Multivariate classification

51 well separating variables are used to train 2 consecutive binary classifiers (**FastBDT**) $\mathrm{BDT}_1$ and $\mathrm{BDT}_2$.

- train $\mathrm{BDT}_1$ on 1.6M signal events and 1.6M $\times\,(B^+B^-,\, B^0\bar{B}^0,\, u\bar{u},\, d\bar{d},\, c\bar{c},\, s\bar{s},\, \tau^+\tau^-)$ events:

# Multivariate classification

51 well separating variables are used to train 2 consecutive binary classifiers (**FastBDT**) $\mathrm{BDT}_1$ and $\mathrm{BDT}_2$.

- train $\mathrm{BDT}_1$ on 1.6M signal events and 1.6M $\times\,(B^+B^-,\, B^0\bar{B}^0,\, u\bar{u},\, d\bar{d},\, c\bar{c},\, s\bar{s},\, \tau^+\tau^-)$ events:

  **FIRST SIGNAL IDENTIFICATION**

# Multivariate classification

51 well separating variables are used to train 2 consecutive binary classifiers (**FastBDT**) $\mathrm{BDT}_1$ and $\mathrm{BDT}_2$.

- train $\mathrm{BDT}_1$ on 1.6M signal events and 1.6M $\times\,(B^+B^-,\, B^0\bar{B}^0,\, u\bar{u},\, d\bar{d},\, c\bar{c},\, s\bar{s},\, \tau^+\tau^-)$ events:

  **FIRST SIGNAL IDENTIFICATION**

- train $\mathrm{BDT}_2$ - same features - **on the events with** $\mathrm{BDT}_1 > 0.9$ among $100\ \mathrm{fb}^{-1}$ events of generic background and 1.6M events of signal:

# Multivariate classification

51 well separating variables are used to train 2 consecutive binary classifiers (**FastBDT**) $\mathrm{BDT}_1$ and $\mathrm{BDT}_2$.

- train $\mathrm{BDT}_1$ on 1.6M signal events and 1.6M $\times\,(B^+B^-,\, B^0\bar{B}^0,\, u\bar{u},\, d\bar{d},\, c\bar{c},\, s\bar{s},\, \tau^+\tau^-)$ events:

**FIRST SIGNAL IDENTIFICATION**

↓

- train $\mathrm{BDT}_2$ - same features - **on the events with** $\mathrm{BDT}_1 > 0.9$ among $100\ \mathrm{fb}^{-1}$ events of generic background and 1.6M events of signal:

**IMPROVEMENT OF SIGNAL SENSITIVITY**
**(+10%, up to ~50%)**
**IN THE HIGH PURITY REGION**
**(+35% purity at 4% signal eff.)**

# Multivariate classification

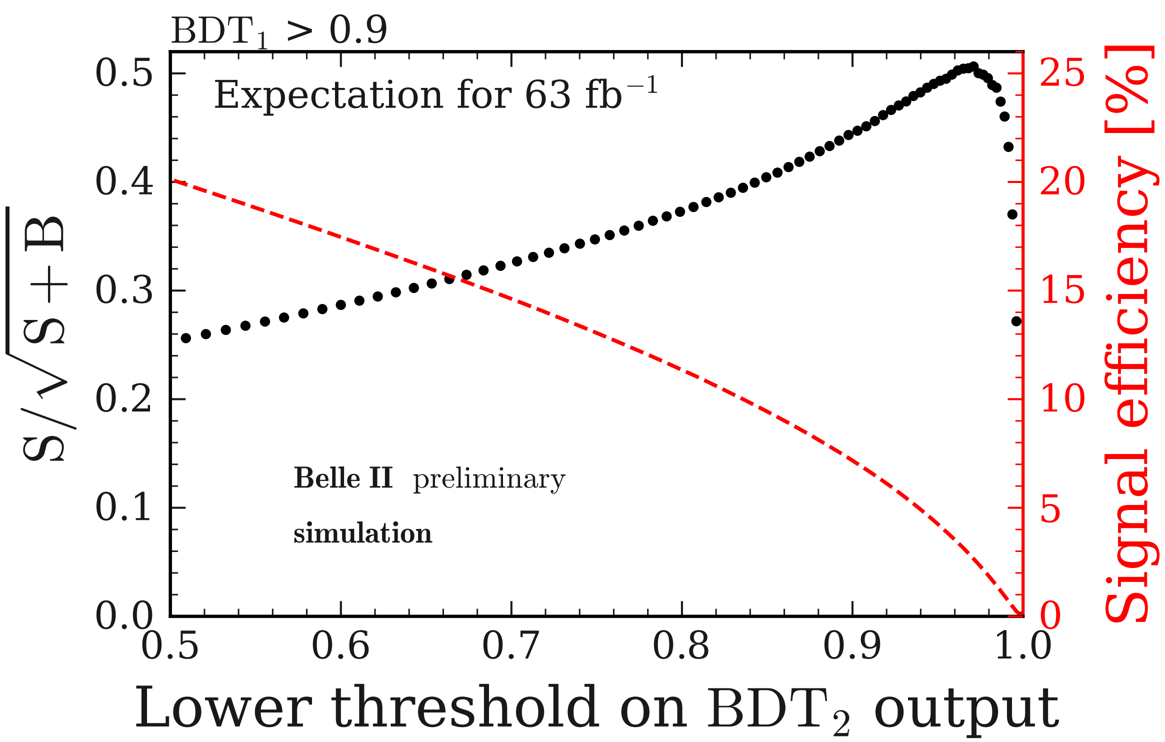
51 well separating variables are used to train 2 consecutive binary classifiers (**FastBDT**) $\mathrm{BDT}_1$ and $\mathrm{BDT}_2$.

- train $\mathrm{BDT}_1$ on 1.6M signal events and 1.6M $\times$ $(B^+B^-, B^0\bar{B}^0, u\bar{u}, d\bar{d}, c\bar{c}, s\bar{s}, \tau^+\tau^-)$ events:

**FIRST SIGNAL IDENTIFICATION**

- train $\mathrm{BDT}_2$ - same features - **on the events with** $\mathrm{BDT}_1 > 0.9$ among 100 $\mathrm{fb}^{-1}$ events of generic background and 1.6M events of signal:

**IMPROVEMENT OF SIGNAL SENSITIVITY (+10%, up to ~50%) IN THE HIGH PURITY REGION (+35% purity at 4% signal eff.)**



# Multivariate classification

51 well separating variables are used to train 2 consecutive binary classifiers (**FastBDT**) $\mathrm{BDT}_1$ and $\mathrm{BDT}_2$.

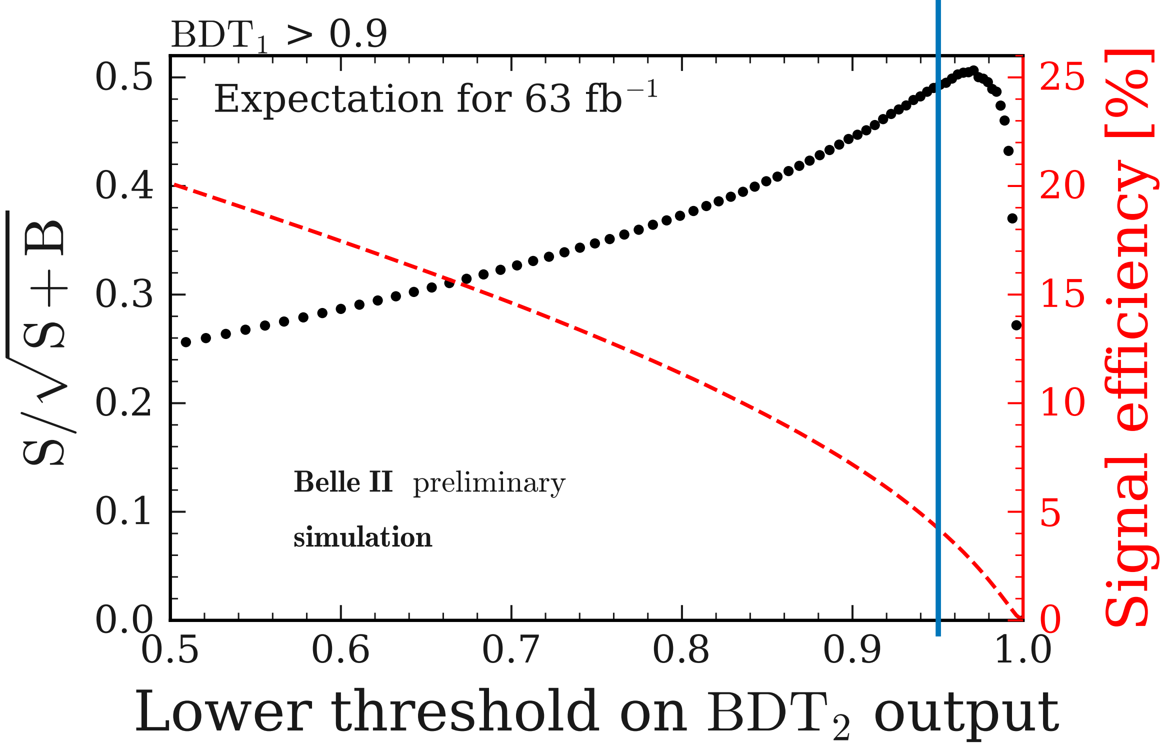
- train $\mathrm{BDT}_1$ on 1.6M signal events and 1.6M $\times (B^+B^-, B^0\bar{B}^0, u\bar{u}, d\bar{d}, c\bar{c}, s\bar{s}, \tau^+\tau^-)$ events:

**FIRST SIGNAL IDENTIFICATION**

- train $\mathrm{BDT}_2$ - same features - **on the events with** $\mathrm{BDT}_1 > 0.9$ among 100 $\mathrm{fb}^{-1}$ events of generic background and 1.6M events of signal:

**IMPROVEMENT OF SIGNAL SENSITIVITY (+10%, up to ~50%) IN THE HIGH PURITY REGION (+35% purity at 4% signal eff.)**

**Signal Region:**
$\mathrm{BDT}_1 > 0.9$ , $\mathrm{BDT}_2 > 0.95$



# Validation using $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$

# Validation using $B^+ \to K^+ J/\psi_{\to \mu^+\mu^-}$

Mode with large branching ratio characterised by clean experimental signature.

# Validation using $B^+ \to K^+ J/\psi_{\to \mu^+\mu^-}$

Mode with large branching ratio characterised by clean experimental signature.

**Identification of $B^+ \to K^+ J/\psi_{\to \mu^+\mu^-}$ events**

# Validation using $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$

Mode with large branching ratio characterised by clean experimental signature.

**Identification of $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events**

$\oplus$

**Strategy to mimic reconstructed $B^+ \to K^+ \nu\bar{\nu}$ events.**

# Validation using $B^+ \to K^+ J/\psi_{\to \mu^+\mu^-}$

Mode with large branching ratio characterised by clean experimental signature.

**Identification of $B^+ \to K^+ J/\psi_{\to \mu^+\mu^-}$ events**

$\bigoplus$

**Strategy to mimic reconstructed $B^+ \to K^+ \nu\bar{\nu}$ events.**

- Ignore the $\mu^+\mu^-$ from the selected $J/\psi$ decay.

# Validation using $B^+ \to K^+ J/\psi_{\to \mu^+\mu^-}$

Mode with large branching ratio characterised by clean experimental signature.

**Identification of $B^+ \to K^+ J/\psi_{\to \mu^+\mu^-}$ events**

$\oplus$

**Strategy to mimic reconstructed $B^+ \to K^+ \nu\bar{\nu}$ events.**

- Ignore the $\mu^+\mu^-$from the selected $J/\psi$ decay.
- 2-body → 3-body kinematics: replace the 4-momentum of the $K^+$ with the generator-level 4-momentum taken from the $K^+$ in $B^+ \to K^+ \nu\bar{\nu}$.

# Validation using $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$

Mode with large branching ratio characterised by clean experimental signature.

**Identification of $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events**

$\bigoplus$

**Strategy to mimic reconstructed $B^+ \to K^+ \nu\bar{\nu}$ events.**

- Ignore the $\mu^+\mu^-$ from the selected $J/\psi$ decay.
- 2-body → 3-body kinematics: replace the 4-momentum of the $K^+$ with the generator-level 4-momentum taken from the $K^+$ in $B^+ \to K^+\nu\bar{\nu}$.
- Reconstruct the modified $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events with the inclusive tagging algorithm.

# Validation using $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$

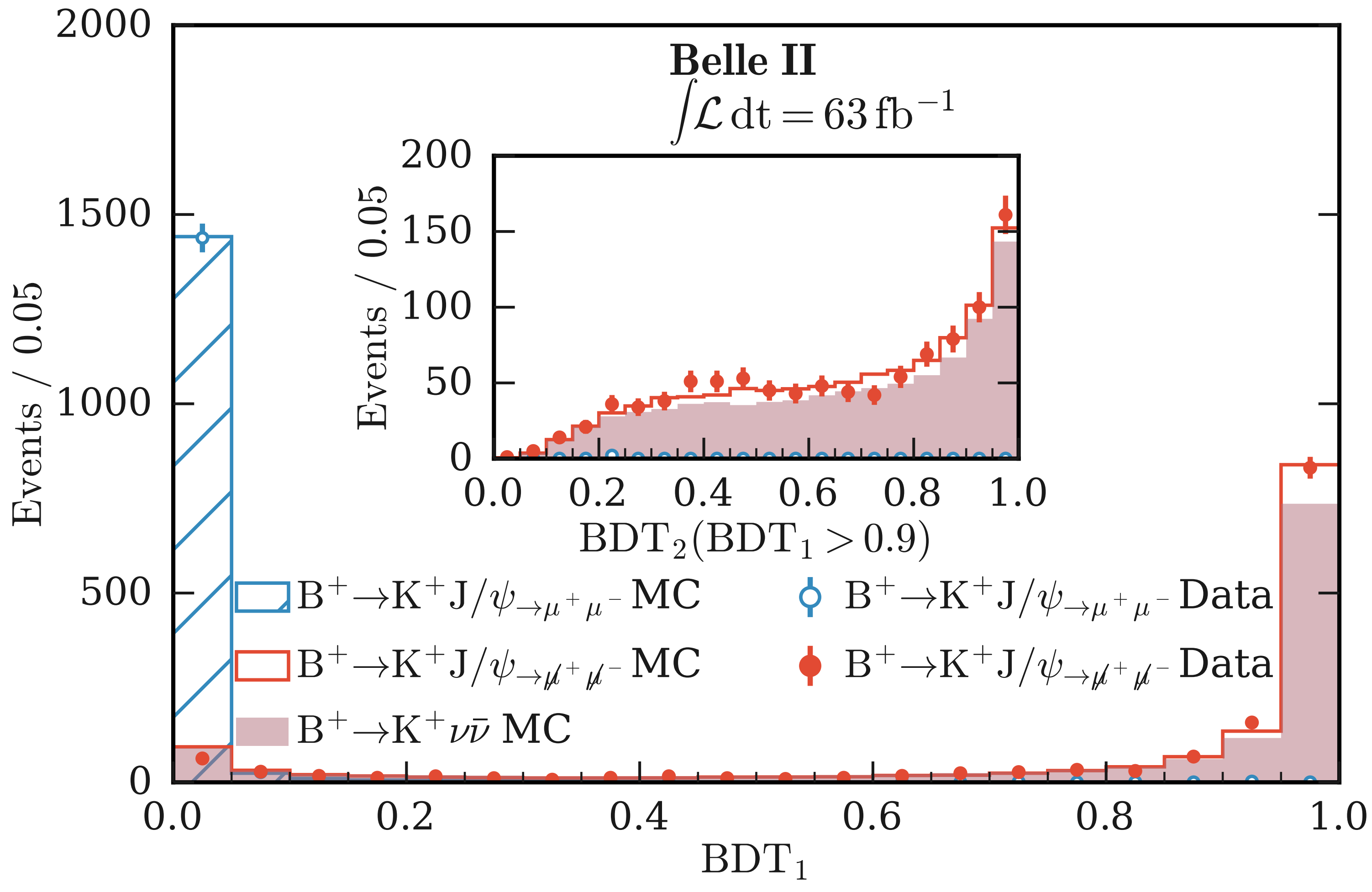
Mode with large branching ratio characterised by clean experimental signature.

**Identification of $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events**

$\oplus$

**Strategy to mimic reconstructed $B^+ \to K^+\nu\bar{\nu}$ events.**

- Ignore the $\mu^+\mu^-$from the selected $J/\psi$ decay.
- 2-body $\to$ 3-body kinematics: replace the 4-momentum of the $K^+$ with the generator-level 4-momentum taken from the $K^+$ in $B^+ \to K^+\nu\bar{\nu}$.
- Reconstruct the modified $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events with the inclusive tagging algorithm.



# Validation using $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$

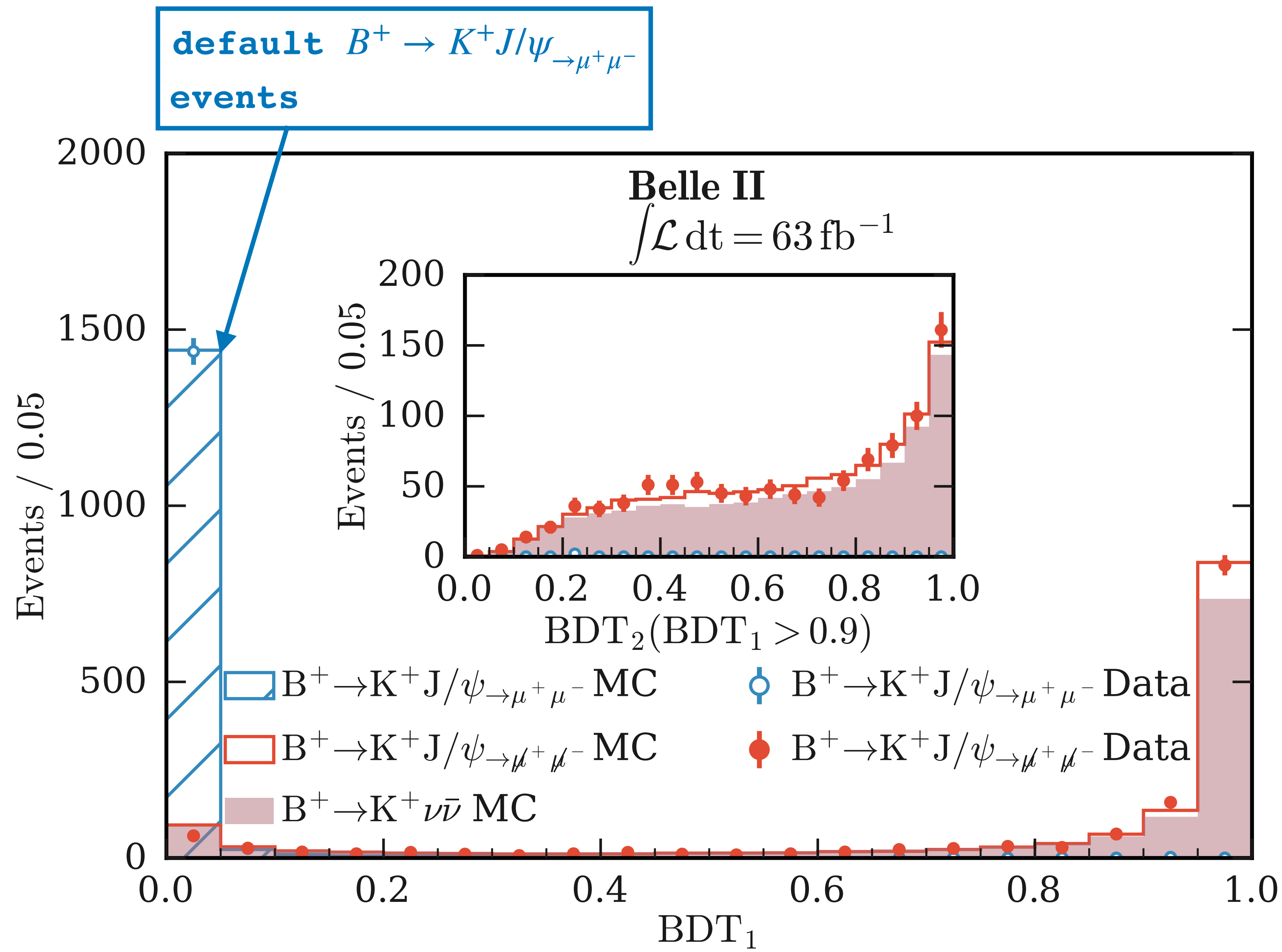
Mode with large branching ratio characterised by clean experimental signature.

**Identification of $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events**

$\oplus$

**Strategy to mimic reconstructed $B^+ \to K^+ \nu\bar{\nu}$ events.**

- Ignore the $\mu^+\mu^-$ from the selected $J/\psi$ decay.
- 2-body $\to$ 3-body kinematics: replace the 4-momentum of the $K^+$ with the generator-level 4-momentum taken from the $K^+$ in $B^+ \to K^+\nu\bar{\nu}$.
- Reconstruct the modified $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events with the inclusive tagging algorithm.



# Validation using $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$

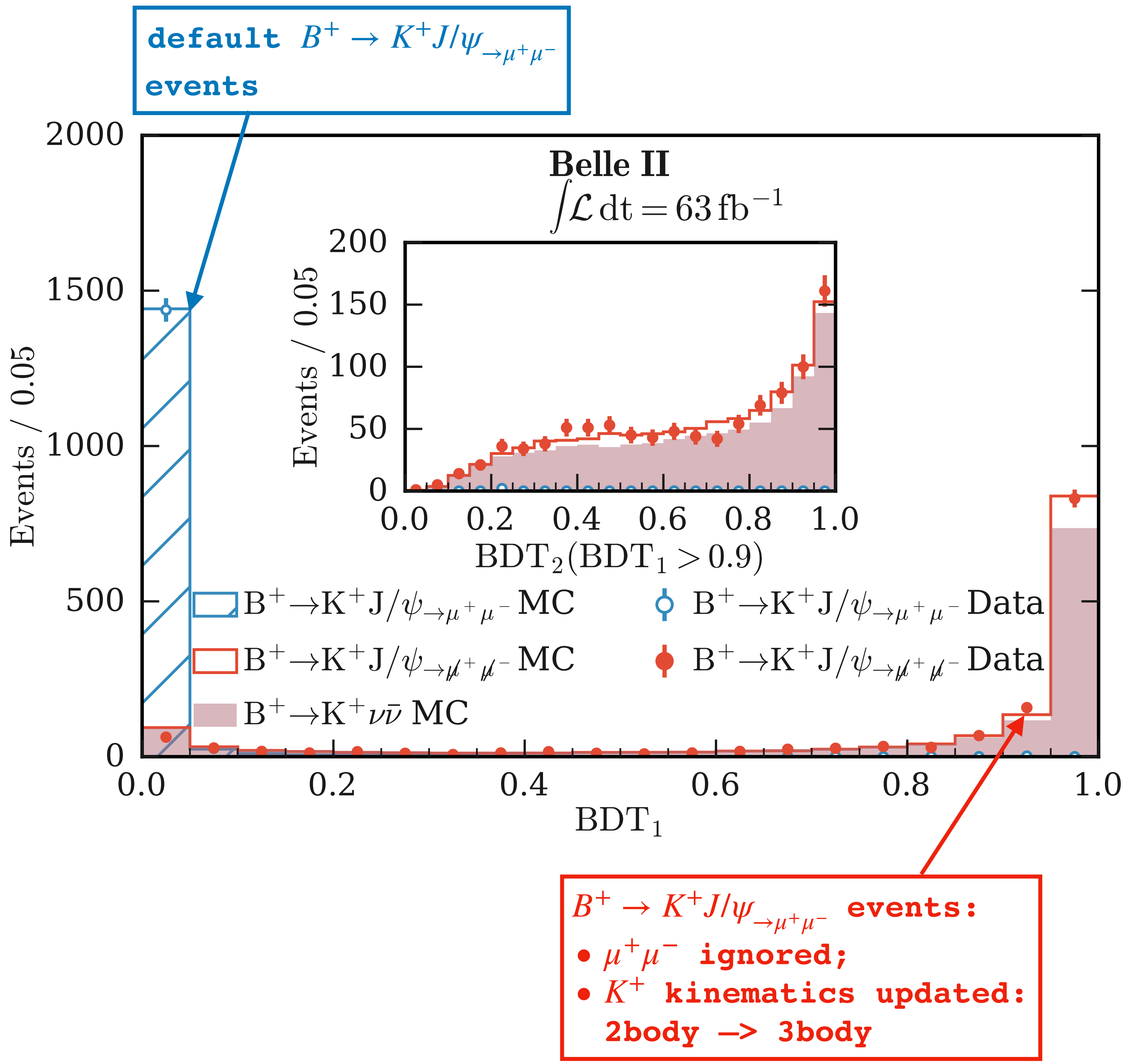
Mode with large branching ratio characterised by clean experimental signature.

**Identification of $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events**

$\oplus$

**Strategy to mimic reconstructed $B^+ \to K^+ \nu\bar{\nu}$ events.**

- Ignore the $\mu^+\mu^-$ from the selected $J/\psi$ decay.
- 2-body → 3-body kinematics: replace the 4-momentum of the $K^+$ with the generator-level 4-momentum taken from the $K^+$ in $B^+ \to K^+\nu\bar{\nu}$.
- Reconstruct the modified $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events with the inclusive tagging algorithm.



# Validation using $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$

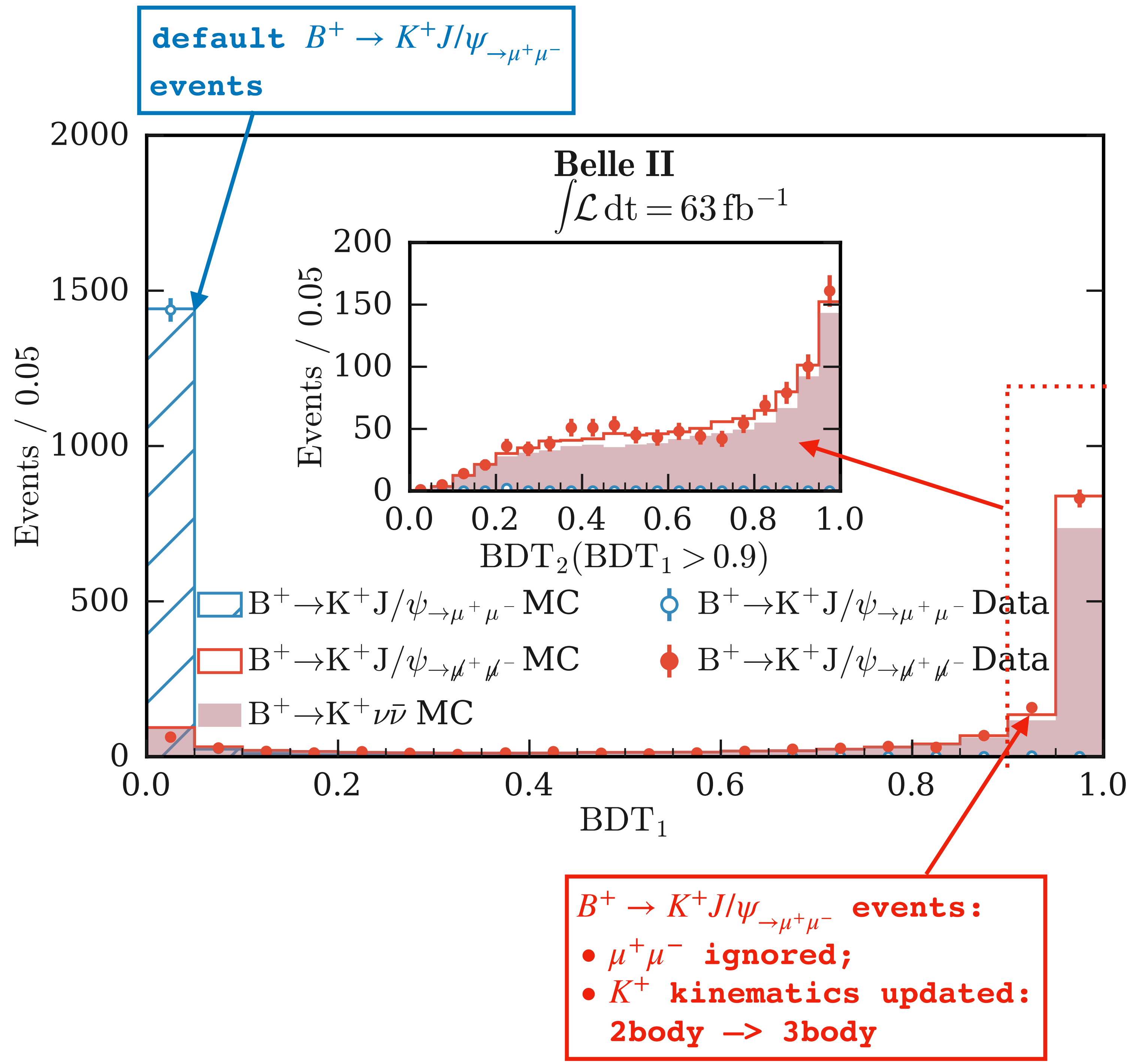
Mode with large branching ratio characterised by clean experimental signature.

**Identification of $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events**

$\oplus$

**Strategy to mimic reconstructed $B^+ \to K^+ \nu\bar{\nu}$ events.**

- Ignore the $\mu^+\mu^-$ from the selected $J/\psi$ decay.
- 2-body $\to$ 3-body kinematics: replace the 4-momentum of the $K^+$ with the generator-level 4-momentum taken from the $K^+$ in $B^+ \to K^+ \nu\bar{\nu}$.
- Reconstruct the modified $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events with the inclusive tagging algorithm.



# Validation using $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$

Mode with large branching ratio characterised by clean experimental signature.

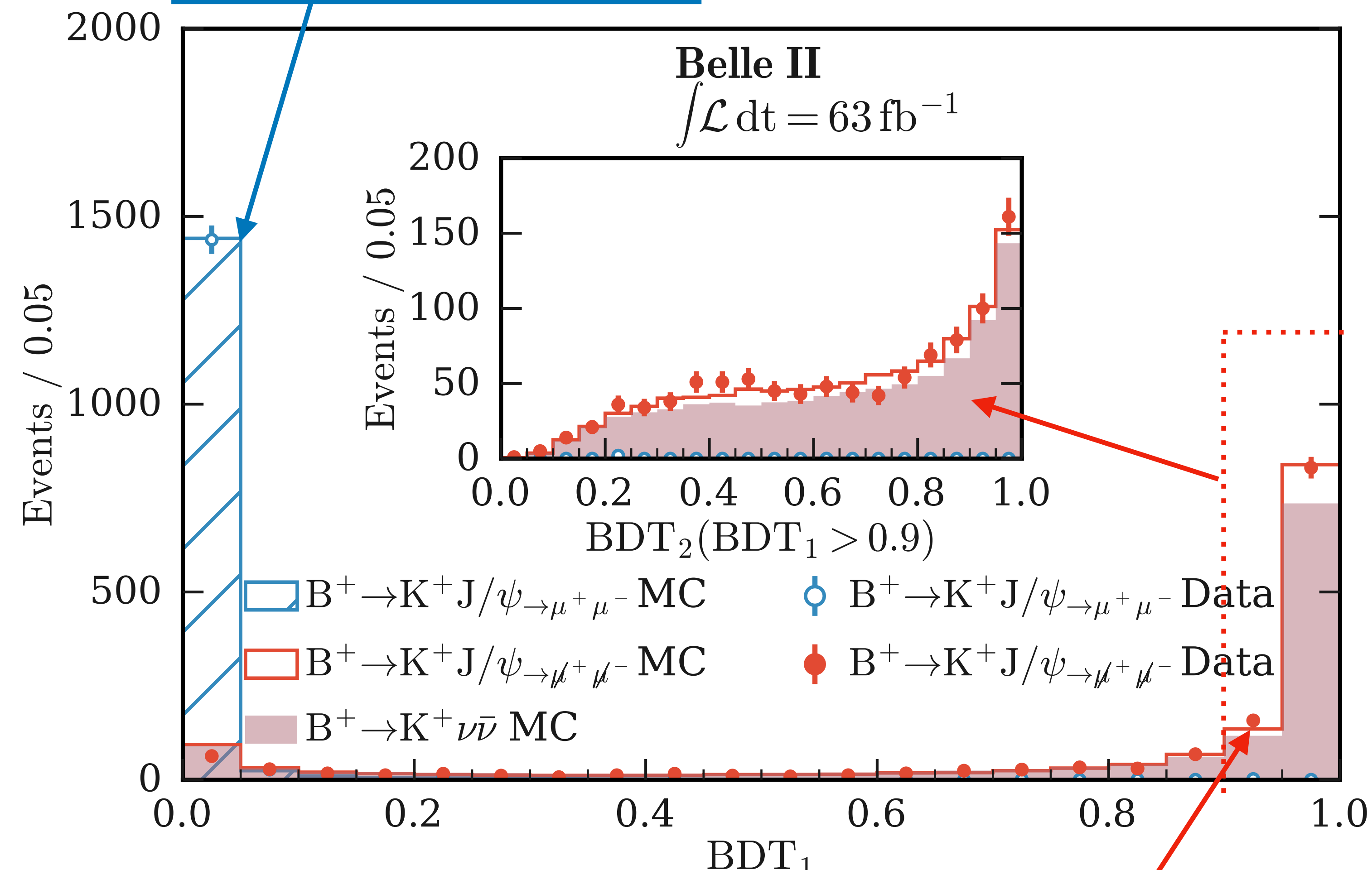
**Identification of $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events**

$\oplus$

**Strategy to mimic reconstructed $B^+ \to K^+ \nu\bar{\nu}$ events.**

- Ignore the $\mu^+\mu^-$ from the selected $J/\psi$ decay.
- 2-body $\to$ 3-body kinematics: replace the 4-momentum of the $K^+$ with the generator-level 4-momentum taken from the $K^+$ in $B^+ \to K^+\nu\bar{\nu}$.
- Reconstruct the modified $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events with the inclusive tagging algorithm.

**default $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events**




**$B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events:**
- **$\mu^+\mu^-$ ignored;**
- **$K^+$ kinematics updated: 2body –> 3body**

**Excellent Data-MC agreement for the BDT's.**

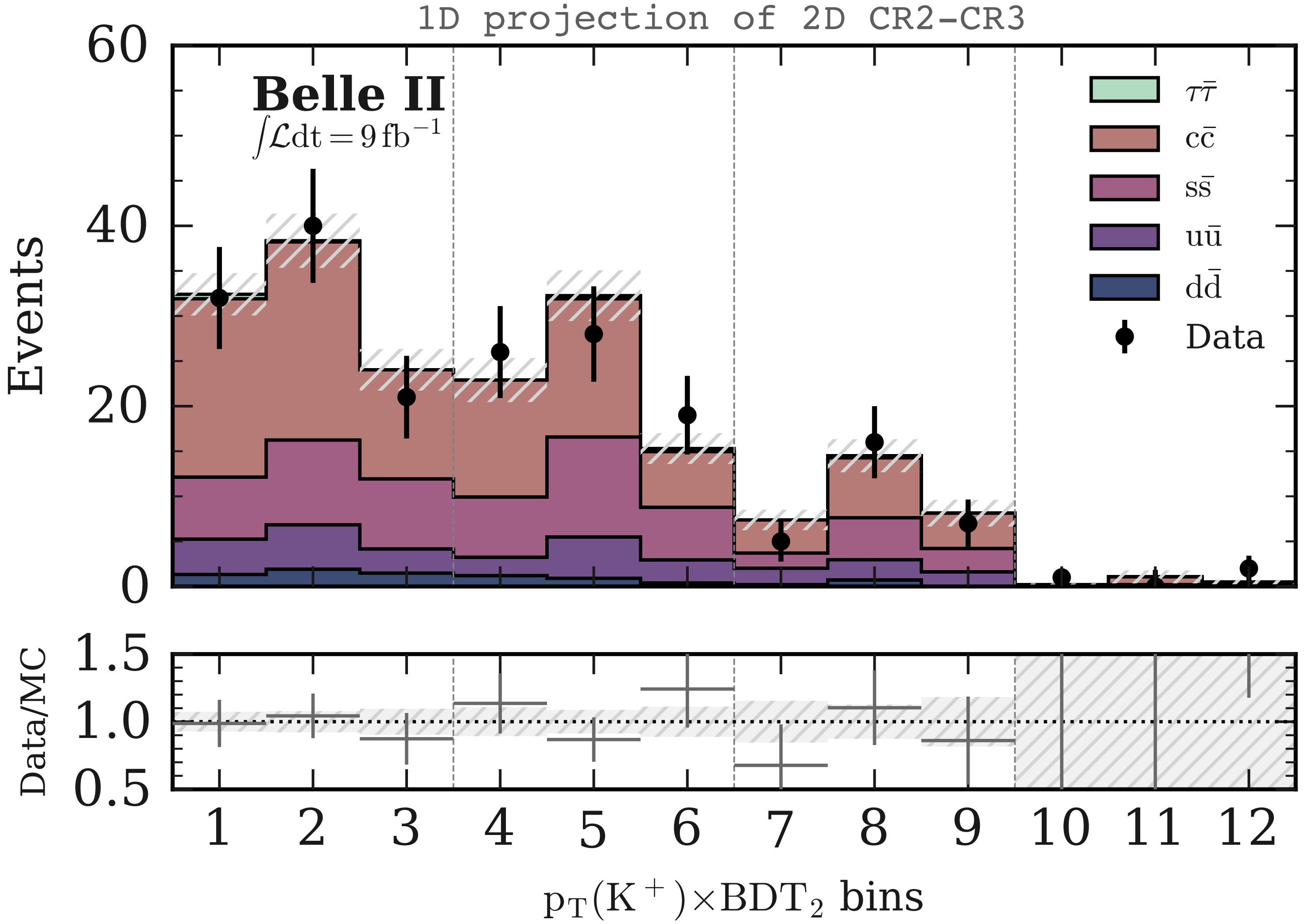
# Validation using off-resonance data

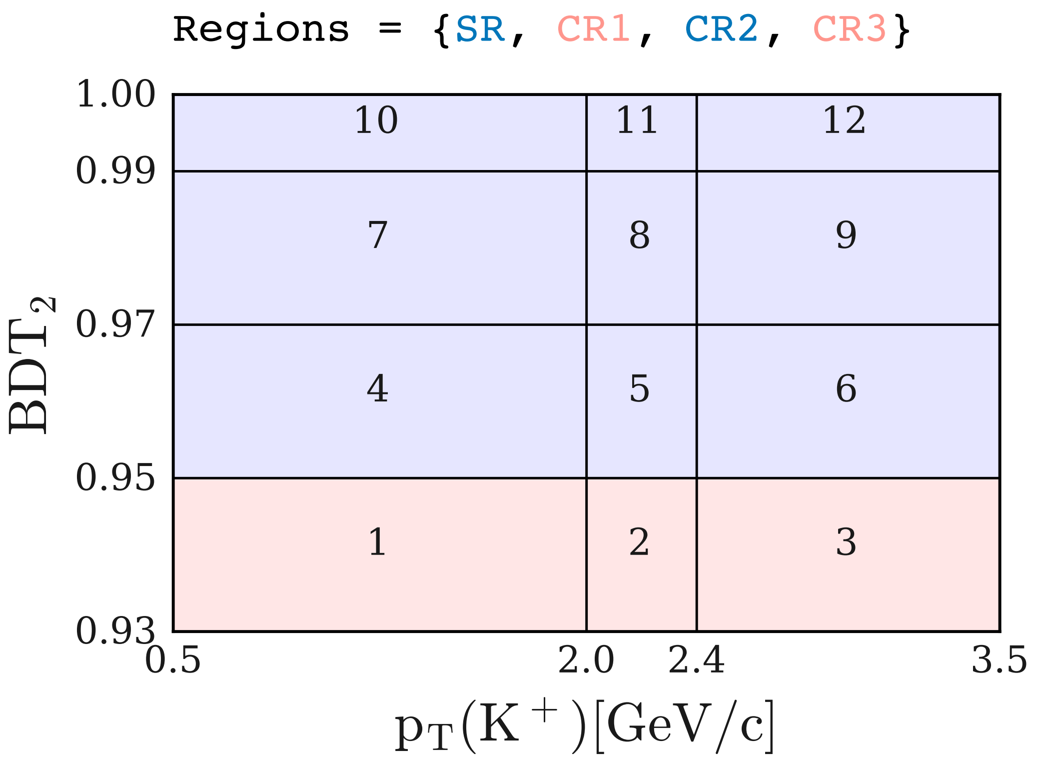
# Validation using off-resonance data

**Investigation of the Data-MC agreement between simulated continuum and off-resonance data in CR2-CR3.**

# Validation using off-resonance data

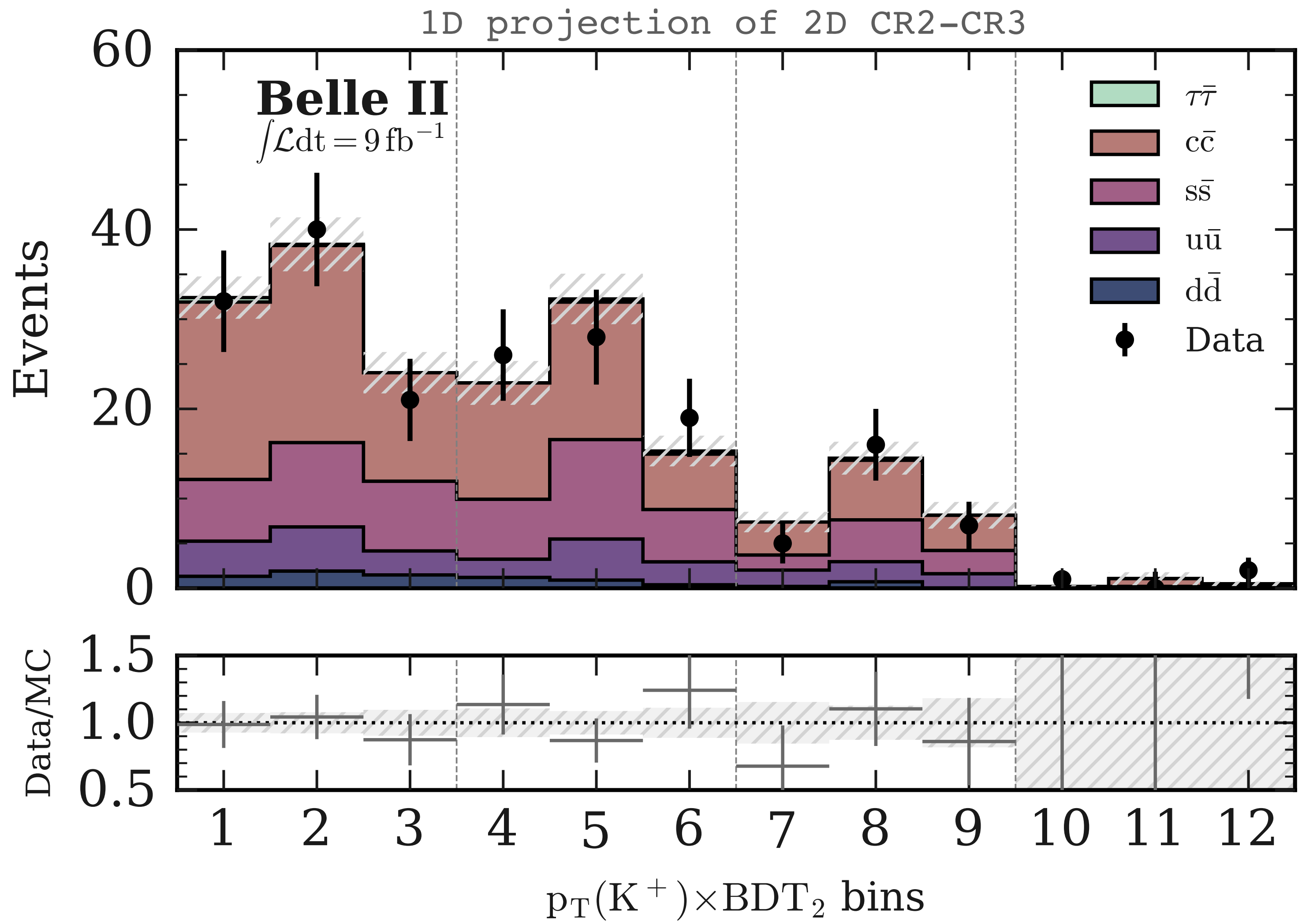
**Investigation of the Data-MC agreement between simulated continuum and off-resonance data in CR2-CR3.**

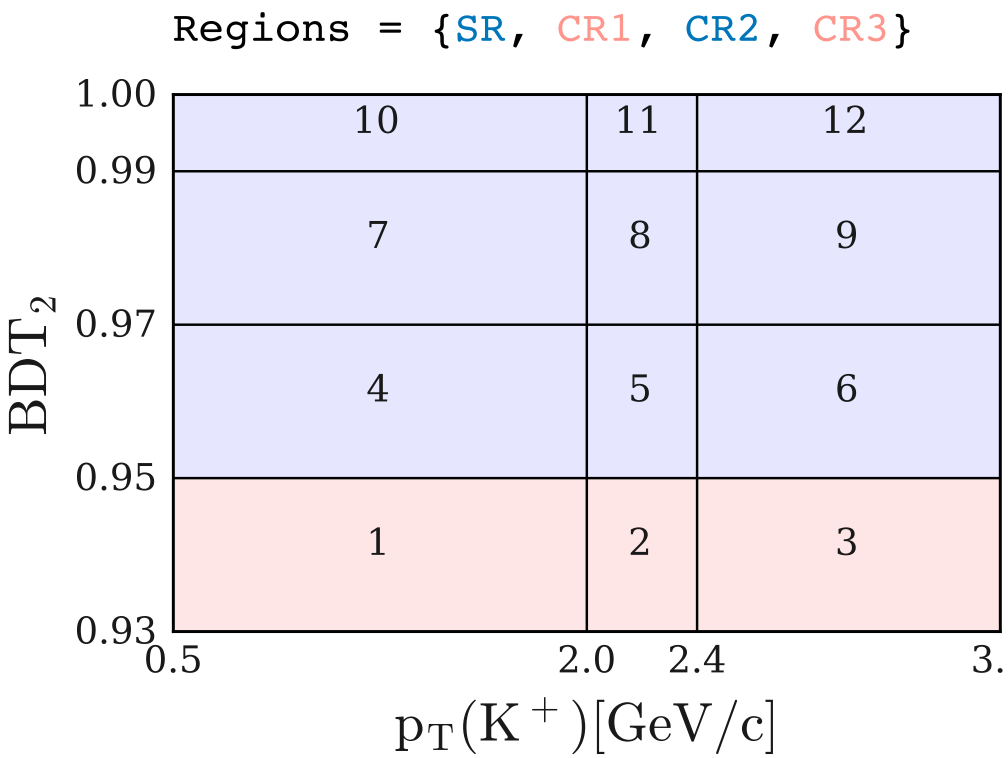


# Validation using off-resonance data

**Investigation of the Data-MC agreement between simulated continuum and off-resonance data in CR2-CR3.**

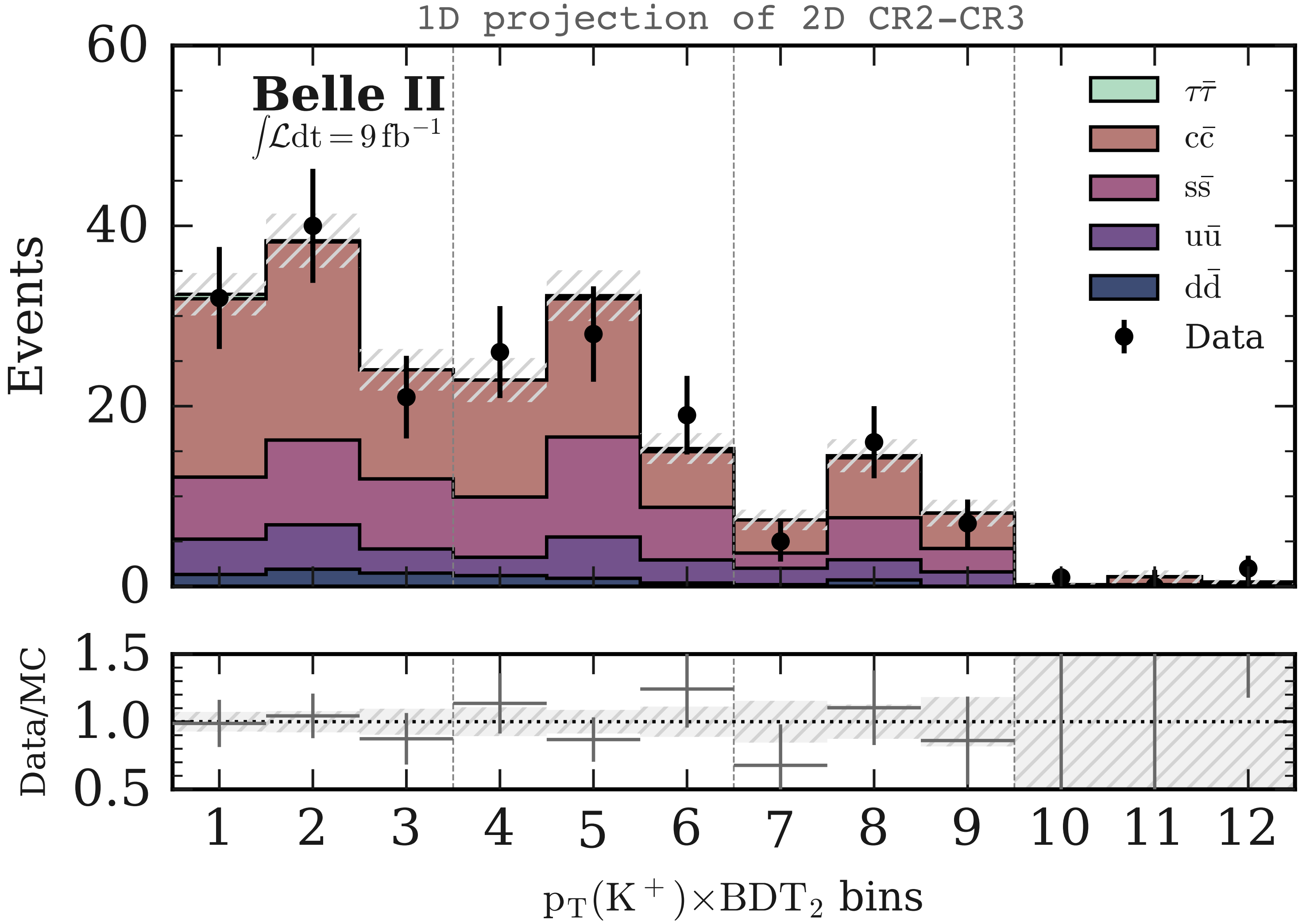


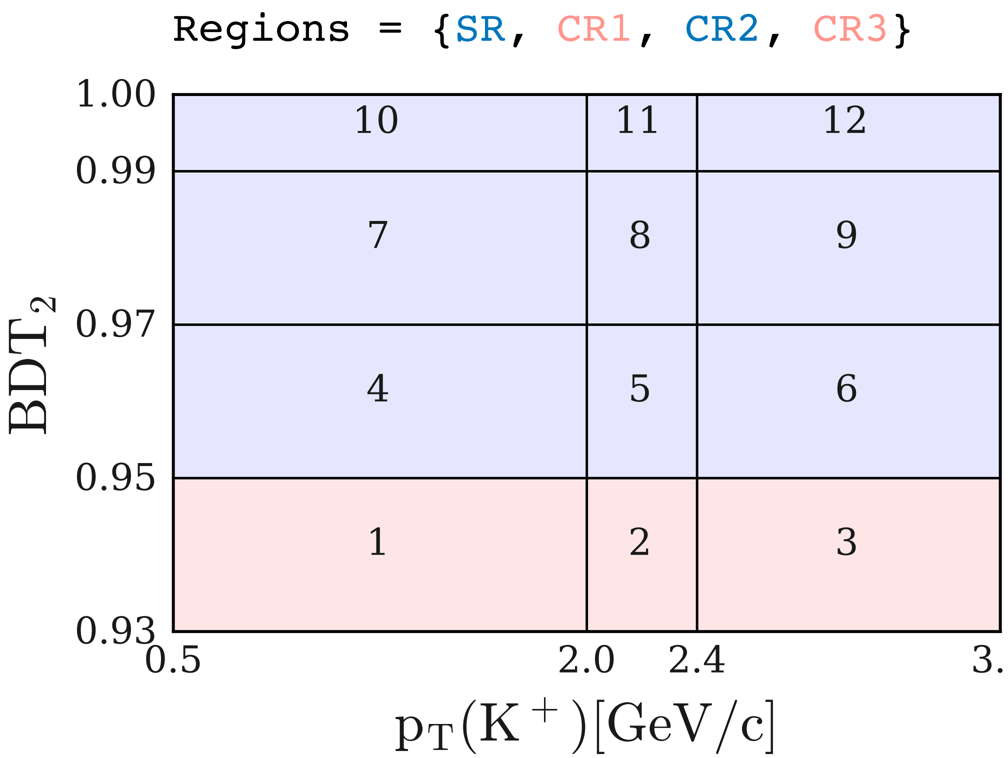




- **Very good Data-MC shape agreement.**

# Validation using off-resonance data

**Investigation of the Data-MC agreement between simulated continuum and off-resonance data in CR2-CR3.**

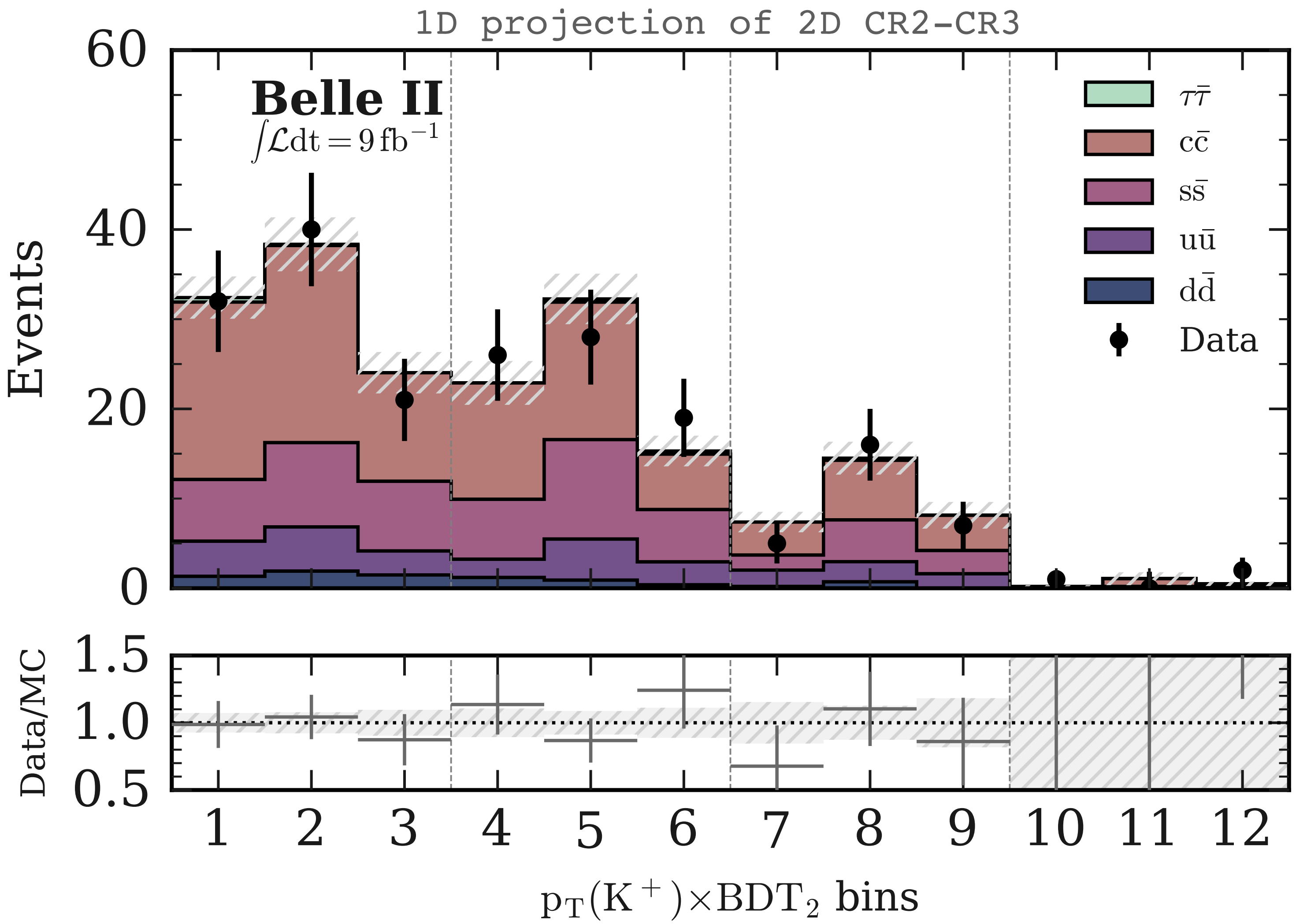


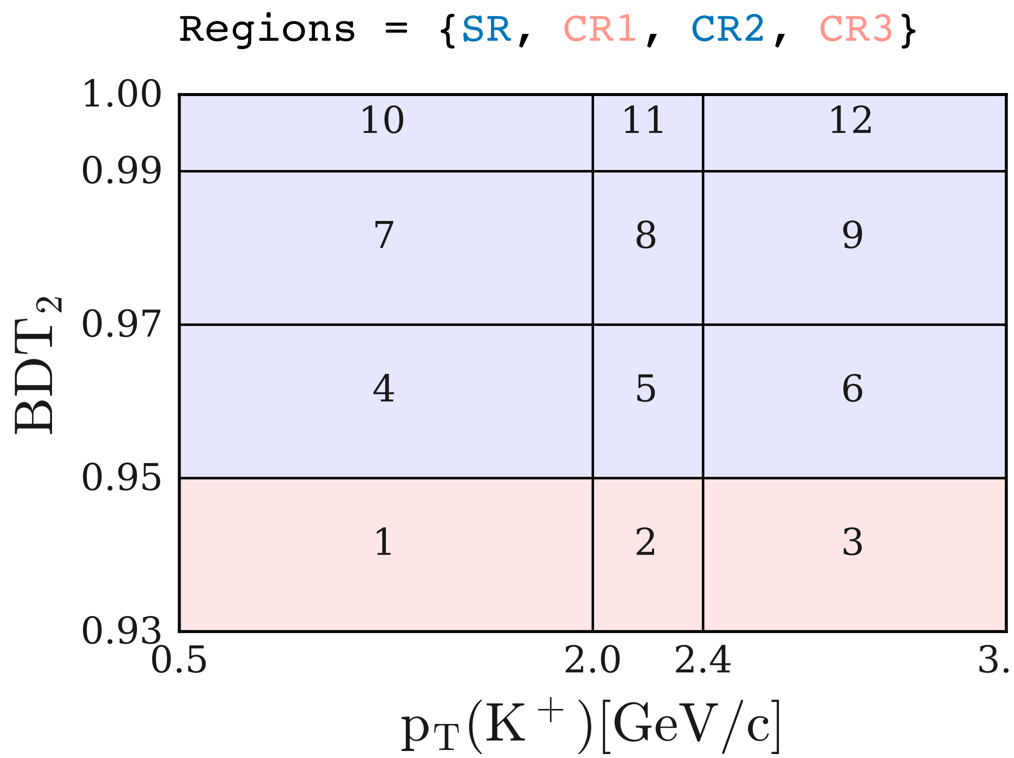




- **Very good Data-MC shape agreement.**
- **But discrepancy in yields:** data/simulation = $1.4 \pm 0.1$.

# Validation using off-resonance data

**Investigation of the Data-MC agreement between simulated continuum and off-resonance data in CR2-CR3.**

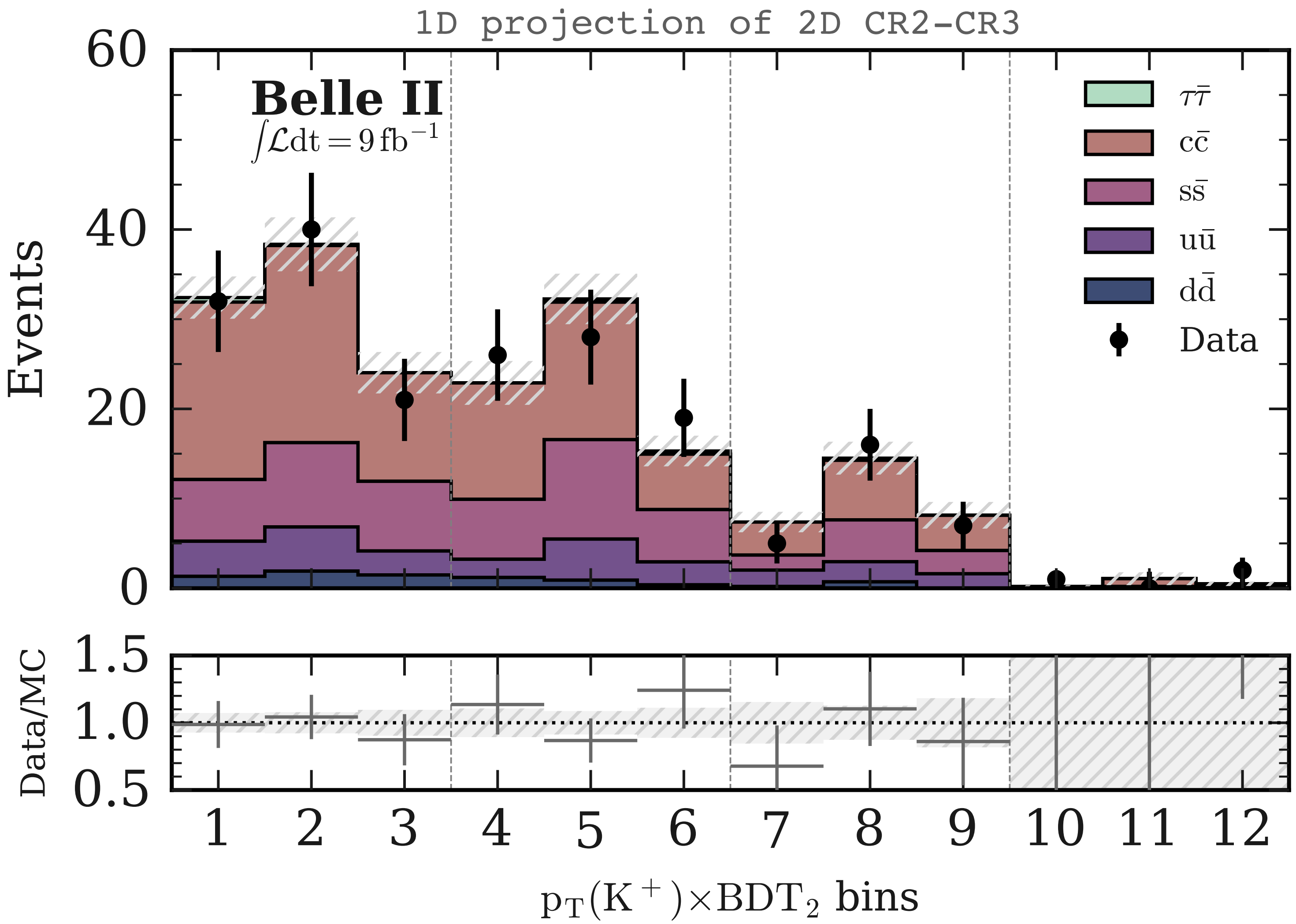


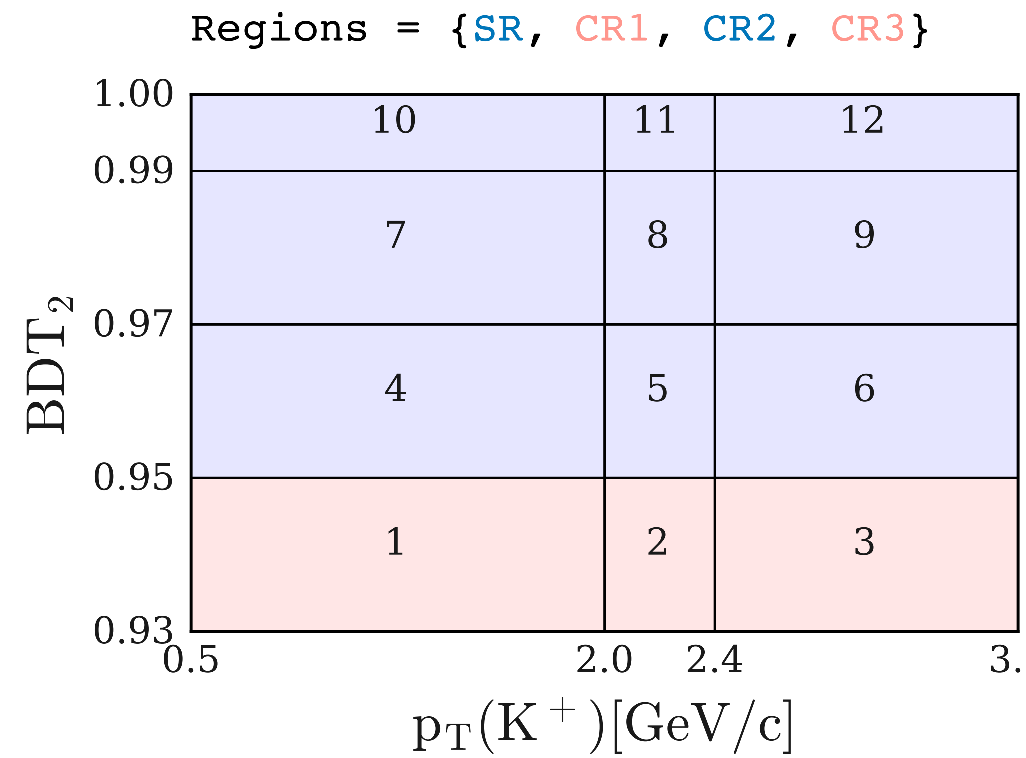




- **Very good Data-MC shape agreement.**
- **But discrepancy in yields:** data/simulation = $1.4 \pm 0.1$.
- Introduction of **conservative 50% normalisation uncertainty in the fit** for each bkg yield individually

# Validation using off-resonance data

**Investigation of the Data-MC agreement between simulated continuum and off-resonance data in CR2-CR3.**




- **Very good Data-MC shape agreement.**
- **But discrepancy in yields:** data/simulation = $1.4 \pm 0.1$.
- Introduction of **conservative 50% normalisation uncertainty in the fit** for each bkg yield individually ⟵ **LEADING SYSTEMATIC UNCERTAINTY.**

# Fit procedure

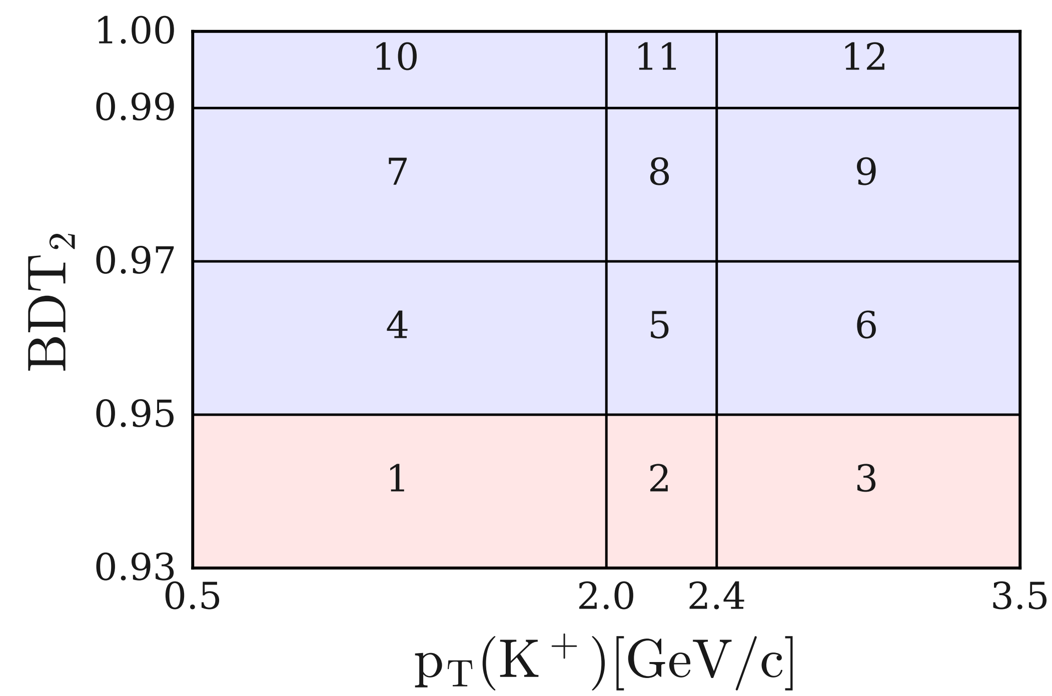
# Fit procedure

Extended Maximum Likelihood Binned Fit:

# Fit procedure

### Extended Maximum Likelihood Binned Fit:

Regions = {SR, CR1, CR2, CR3}



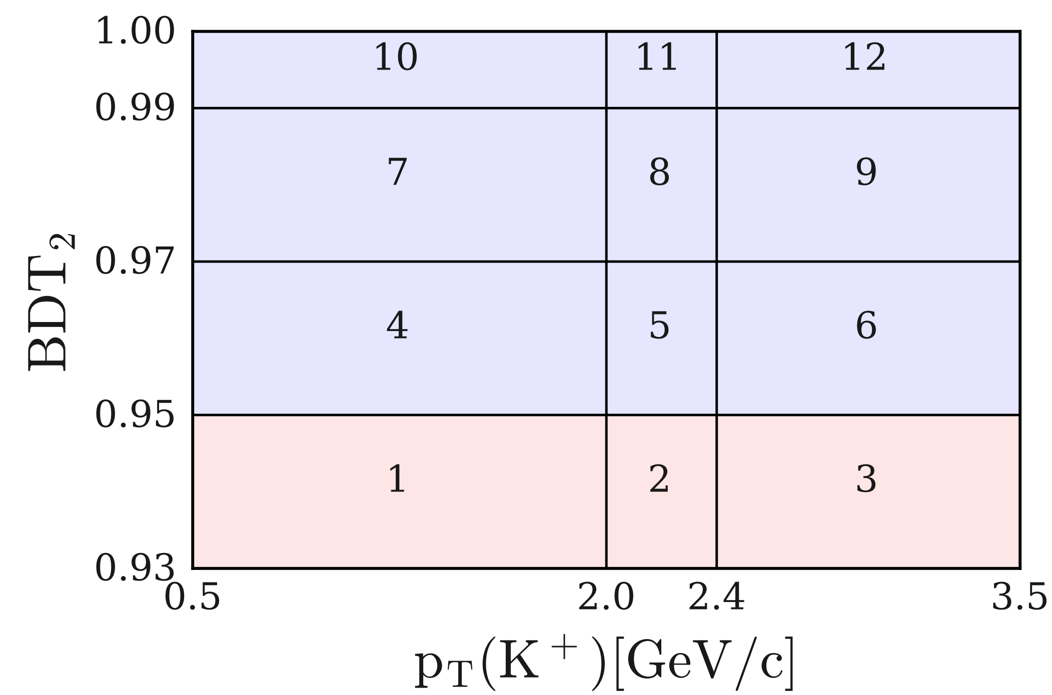
# Fit procedure

## Extended Maximum Likelihood Binned Fit:

$f(n, a \,|\, \eta, \chi) =$

$\eta$ = parameter of interest
$\chi$ = nuisance parameters

Regions = {SR, CR1, CR2, CR3}



# Fit procedure

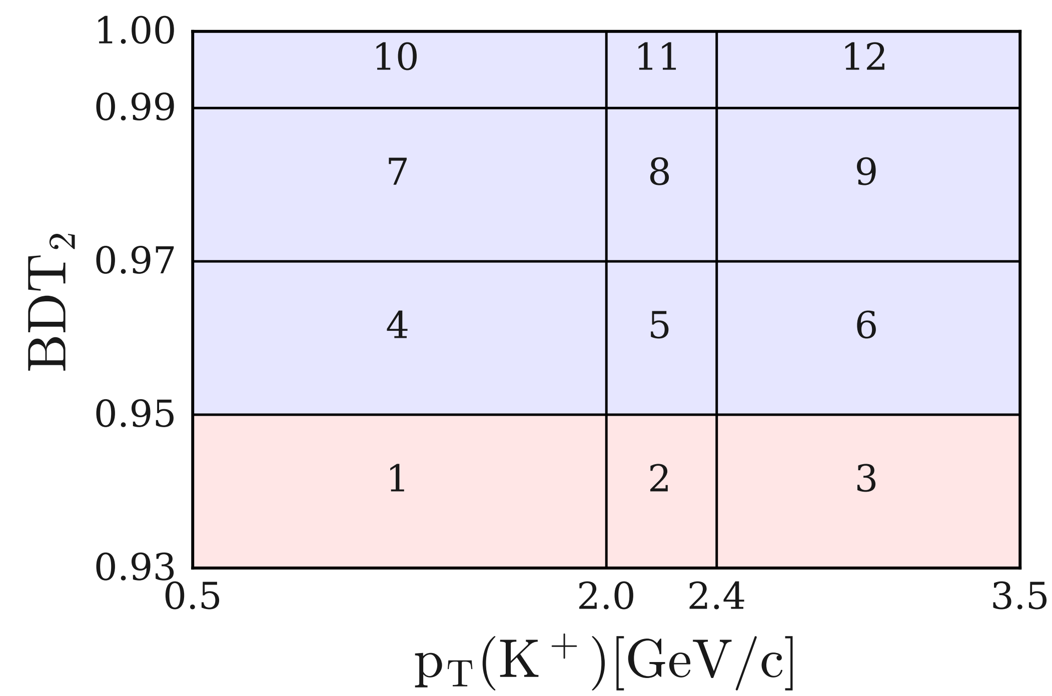
## Extended Maximum Likelihood Binned Fit:

$$f(n, a \,|\, \eta, \chi) = \prod_{r \in \text{regions}} \prod_{b \in \text{bins}} \text{Pois}(n_{rb} \,|\, \nu_{rb}(\eta, \chi))$$

Simultaneous measurements of multiple regions

$\eta$ = parameter of interest
$\chi$ = nuisance parameters

Regions = {SR, CR1, CR2, CR3}



# Fit procedure

## Extended Maximum Likelihood Binned Fit:

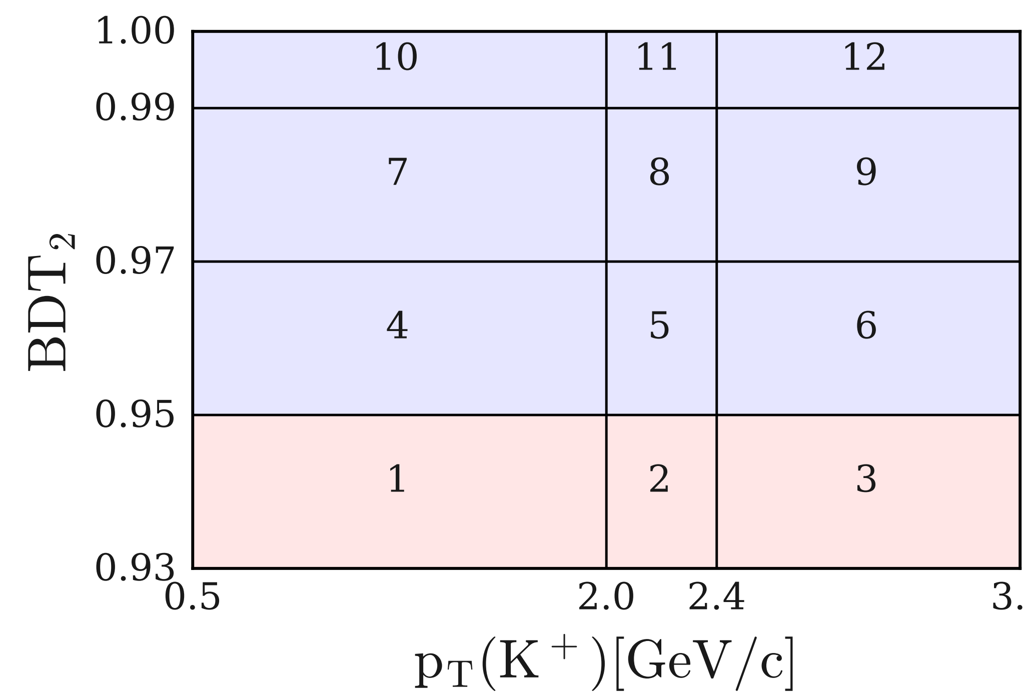
$$f(n, a \mid \eta, \chi) = \prod_{r \in \text{regions}} \prod_{b \in \text{bins}} \text{Pois}(n_{rb} \mid \nu_{rb}(\eta, \chi)) \prod_{\chi} c_{\chi}(a_{\chi} \mid \chi)$$

Simultaneous measurements of multiple regions

Constraints

$\eta$ = parameter of interest
$\chi$ = nuisance parameters

Regions = {SR, CR1, CR2, CR3}



# Fit procedure

## Extended Maximum Likelihood Binned Fit:

$$f(n, a \mid \eta, \chi) = \prod_{r \in \text{regions}} \prod_{b \in \text{bins}} \text{Pois}(n_{rb} \mid \nu_{rb}(\eta, \chi)) \prod_{\chi} c_\chi(a_\chi \mid \chi)$$

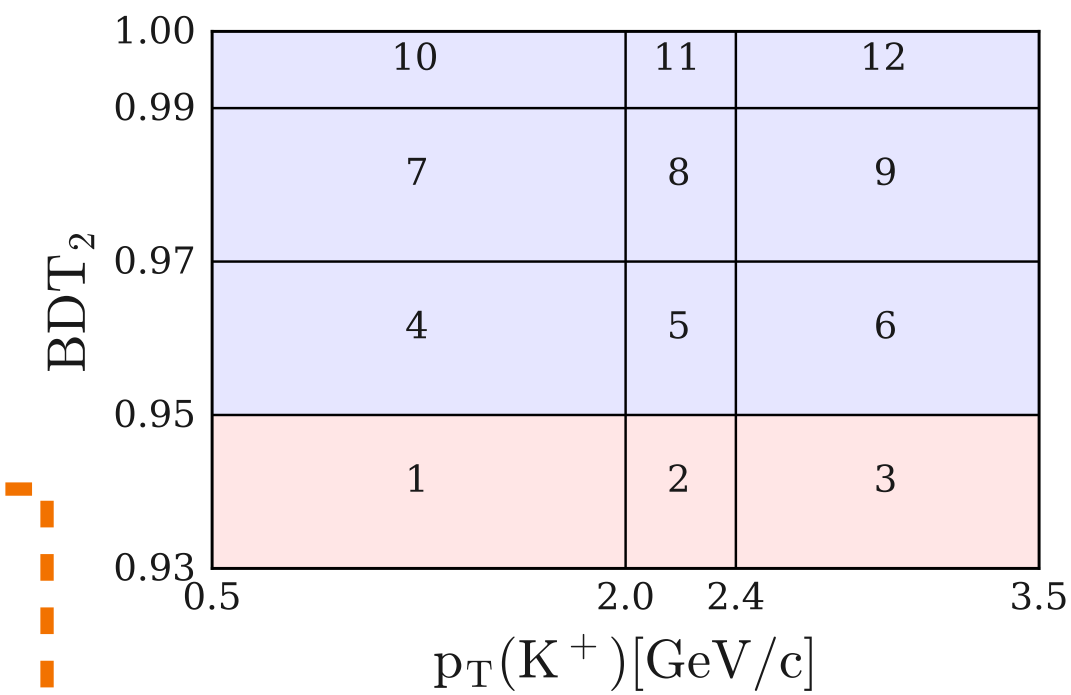
Simultaneous measurements of multiple regions

Constraints

$\eta$ = parameter of interest
$\chi$ = nuisance parameters

- Templates for bkg's and signal yields from simulation.
- **Systematic uncertainties** (normalisations of bkg's yields, BR of the leading *B*-decays, PID correction, …) as (175) **nuisance parameters: event count modifiers.**
- **1 parameter of interest: signal strength μ:** multiplicative factor with respect to the SM expectation.

Regions = {SR, CR1, CR2, CR3}




$\mu = 1 \;\rightarrow$ SM BF $= 4.6 \times 10^{-6}$

# Fit to the Data

# Fit to the Data

- **Measured signal strength** $\mu$

# Fit to the Data

- **Measured signal strength $\mu$**

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

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

# Fit to the Data

- **Measured signal strength $\mu$**

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

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

**- Consistent with the SM expectation ($\mu = 1$) at CL = $1\sigma$.**

# Fit to the Data

- **Measured signal strength $\mu$**

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

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

- **Consistent with the SM expectation ($\mu = 1$) at CL = $1\sigma$.**
- **Consistent with the bkg-only hypothesis ($\mu = 0$) at CL = $1.3\sigma$.**

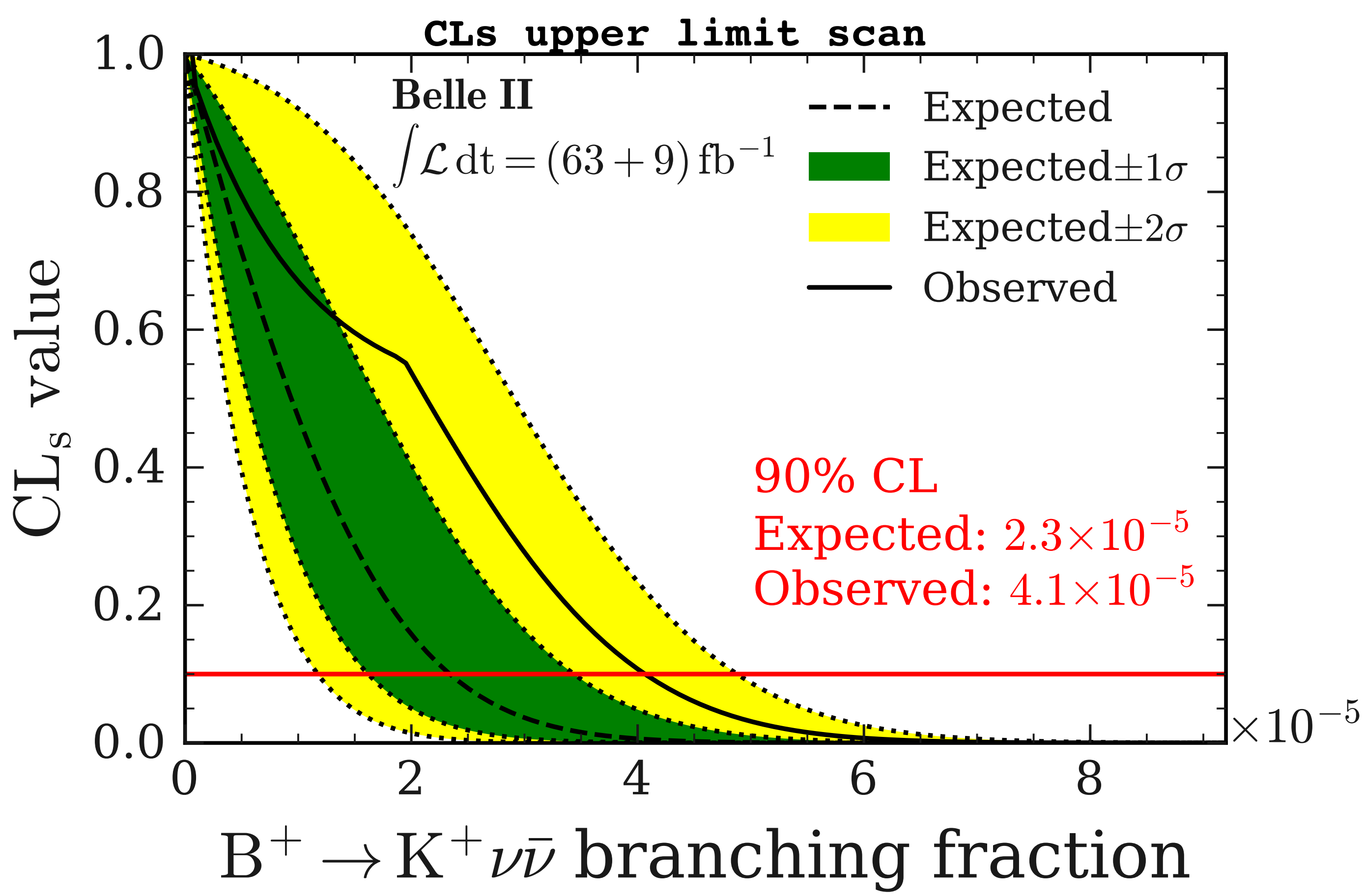
# Fit to the Data

- Measured signal strength $\mu$

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

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

- Consistent with the SM expectation ($\mu = 1$) at CL = $1\sigma$.
- Consistent with the bkg-only hypothesis ($\mu = 0$) at CL = $1.3\sigma$.



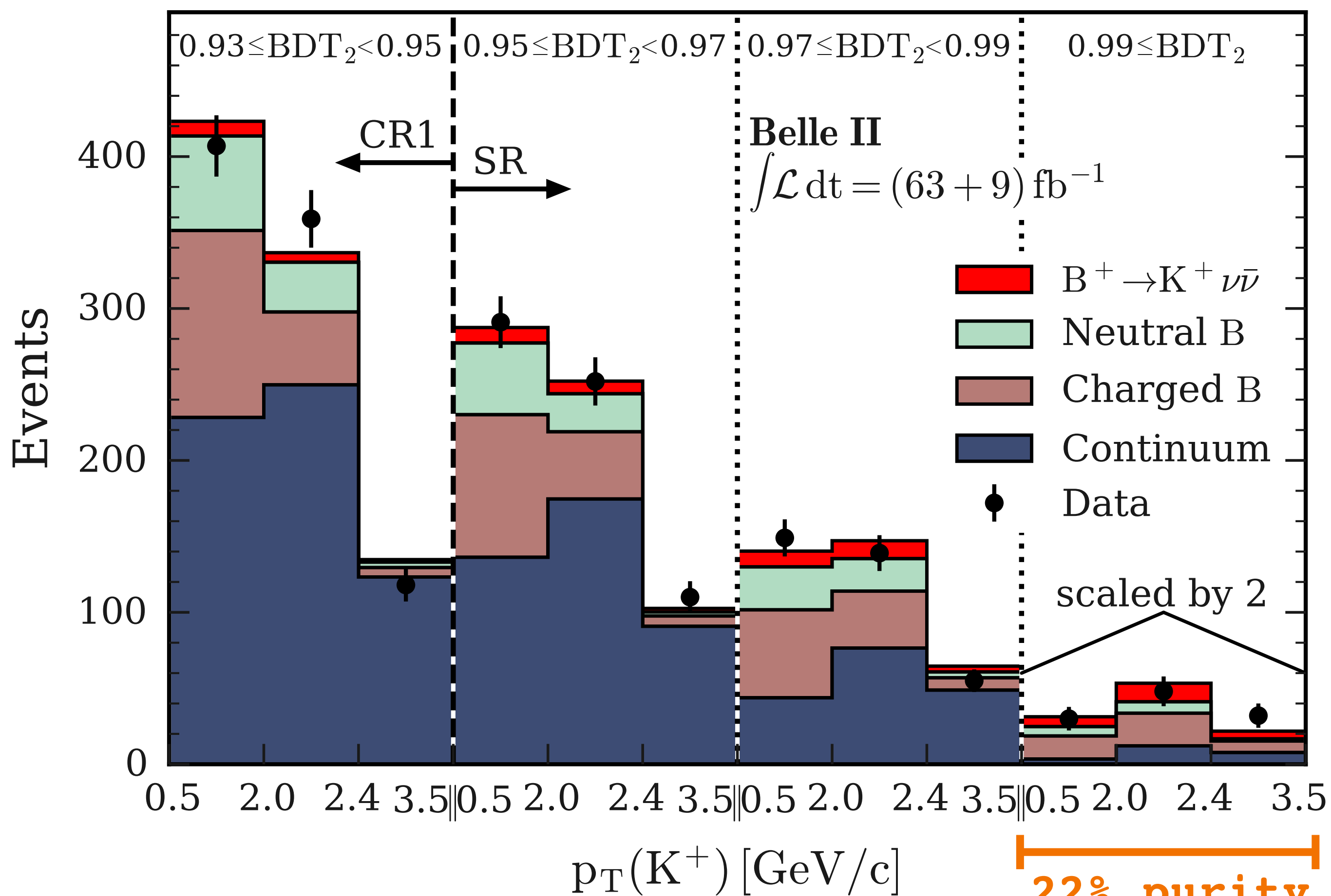
# Fit to the Data

- Measured signal strength $\mu$

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

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

- Data vs post-fit predictions in CR1 + SR




22% purity

6% signal purity in the SR

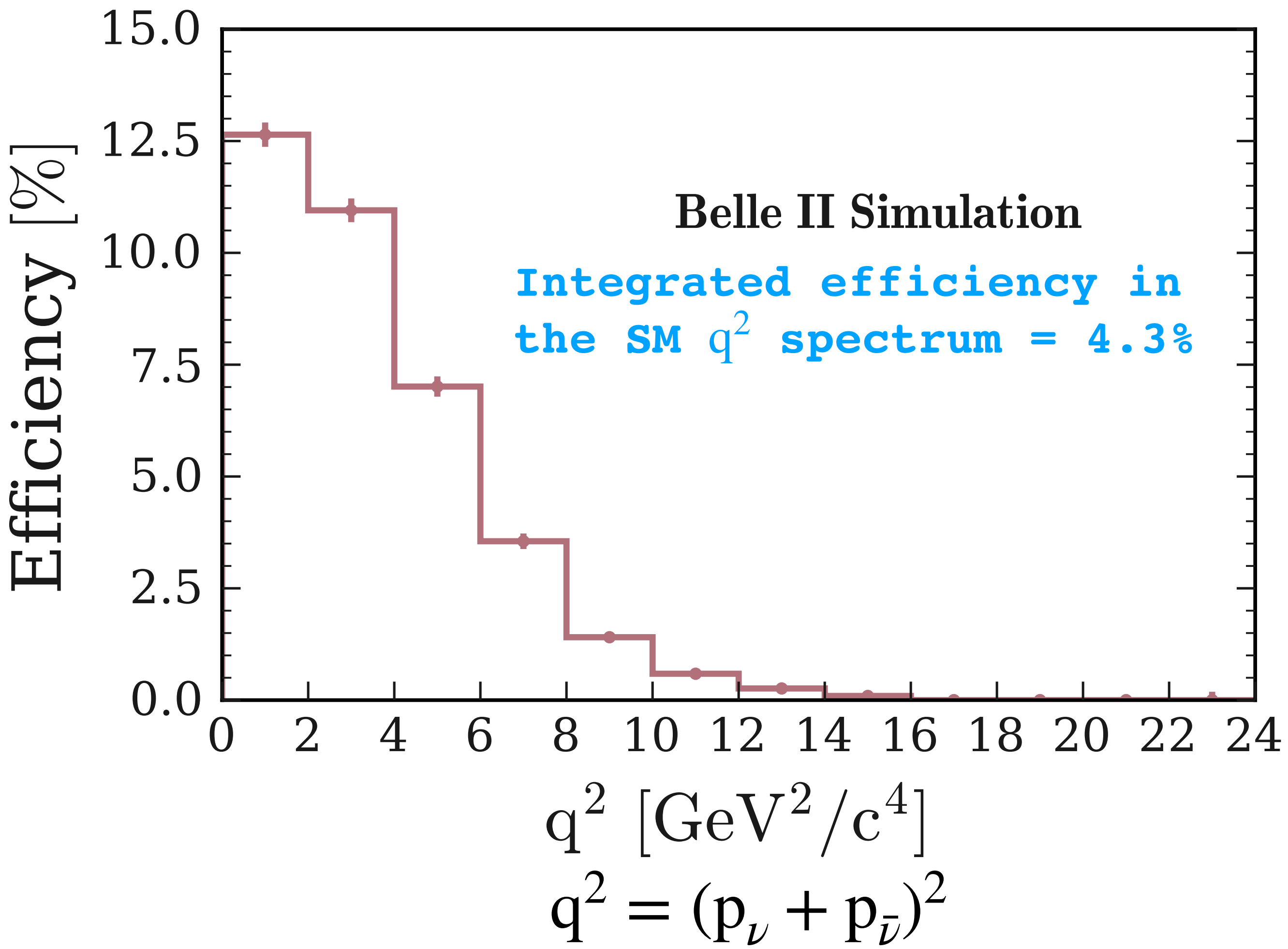
# Fit to the Data

- **Measured signal strength $\mu$**

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

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

- **Signal efficiency (in the signal region SR)**



# Measurement summary…

# Measurement summary…

- This measurement represents the **first search for $B^+ \to K^+ \nu\bar{\nu}$ performed with an inclusive tagging and the first measurement using Belle II in its nominal configuration.**

# Measurement summary…

- This measurement represents the **first search for $B^+ \to K^+ \nu \bar{\nu}$ performed with an inclusive tagging and the first measurement using Belle II in its nominal configuration.**

- No signal yet, but an observed **upper limit on the branching ratio of $4.1 \times 10^{-5}$ is set at the 90% CL.**

# Measurement summary…

- This measurement represents the **first search for $B^+ \to K^+\nu\bar{\nu}$ performed with an inclusive tagging and the first measurement using Belle II in its nominal configuration.**

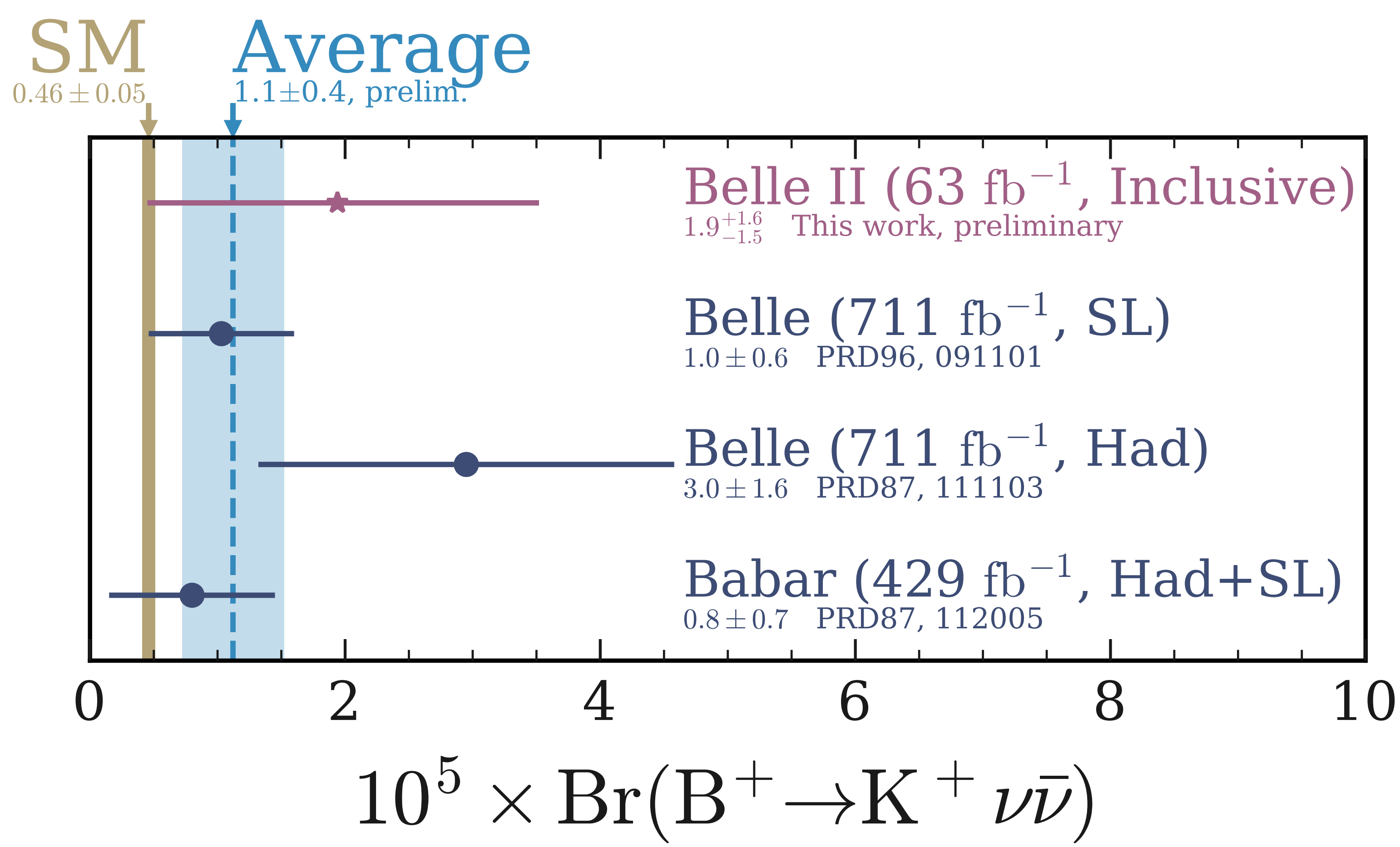
- No signal yet, but an **observed upper limit on the branching ratio of $4.1 \times 10^{-5}$ is set at the 90% CL.**

| Experiment | Year | Observed limit on BR$(B^+ \to K^+\nu\bar{\nu})$ | Approach | Data $[\mathrm{fb}^{-1}]$ |
|---|---|---|---|---|
| **BABAR** | **2013** | $< 1.6 \times 10^{-5}$ **[Phys.Rev.D87,112005]** | **SL + Had tagging** | 429 |
| **Belle** | **2013** | $< 5.5 \times 10^{-5}$ **[Phys.Rev.D87,111103(R)]** | **Had tagging** | 711 |
| **Belle** | **2017** | $< 1.9 \times 10^{-5}$ **[Phys.Rev.D96,091101(R)]** | **SL tagging** | 711 |
| **Belle II** | **2021** | $< 4.1 \times 10^{-5}$ | **Inclusive tagging** | 63 |

# Measurement summary…

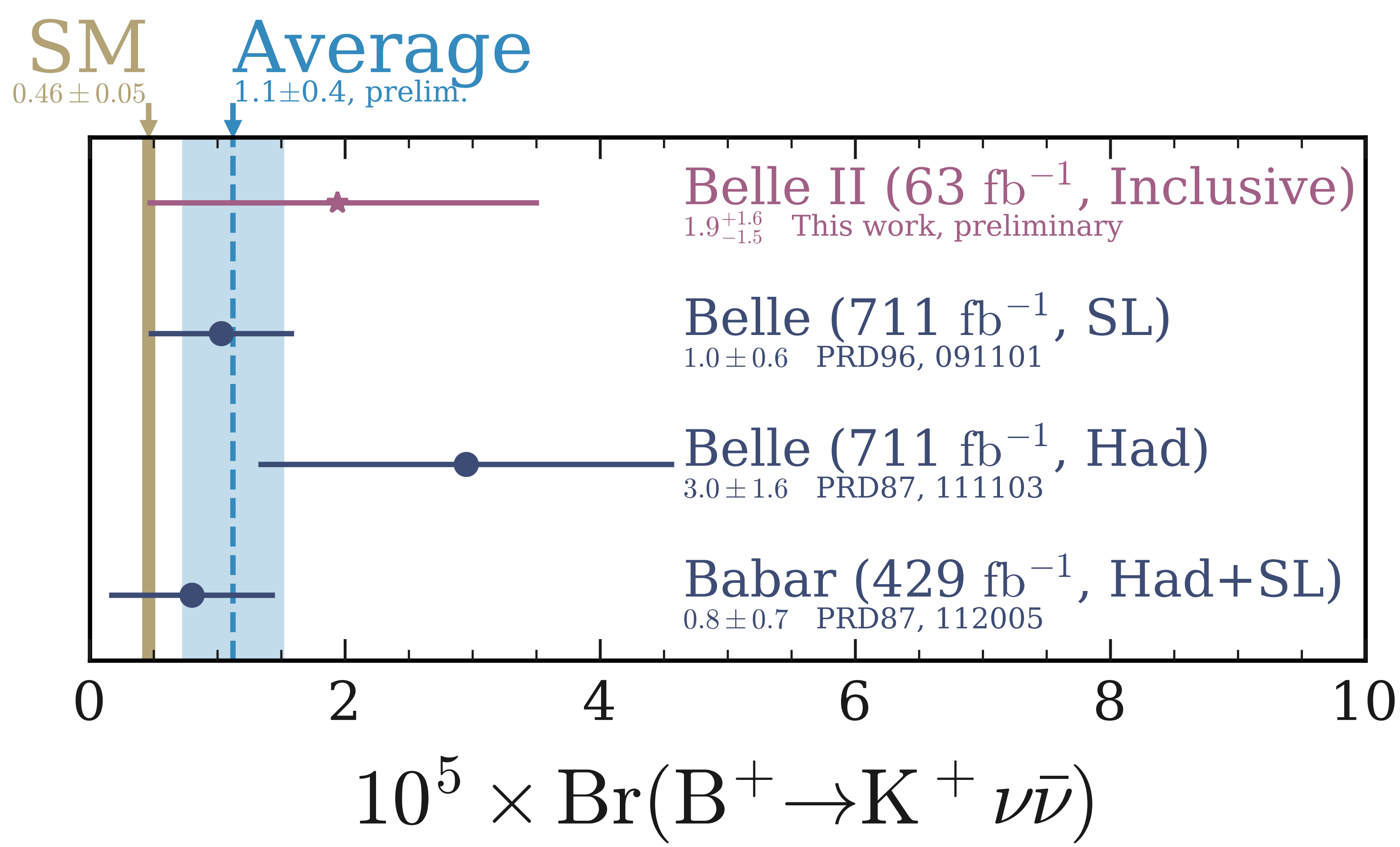
- This measurement represents the **first search for $B^+ \to K^+\nu\bar{\nu}$ performed with an inclusive tagging and the first measurement using Belle II in its nominal configuration.**

- No signal yet, but an **observed upper limit on the branching ratio of $4.1 \times 10^{-5}$ is set at the 90% CL.**



# Measurement summary…

- This measurement represents the **first search for $B^+ \to K^+\nu\bar{\nu}$ performed with an inclusive tagging and the first measurement using Belle II in its nominal configuration.**
- No signal yet, but an **observed upper limit on the branching ratio of $4.1 \times 10^{-5}$ is set at the 90% CL.**
- When converted to the same luminosity, **the Belle II inclusive tagging performs 10-20% better than the semileptonic tagging and a factor 3.5 better than the hadronic tagging.**



# …and future perspectives

# …and future perspectives

- **More data:** $3\times$ larger sample ready to be analysed.

# …and future perspectives

- **More data:** $3\times$ larger sample ready to be analysed.
- **More input variables:** e.g. $K_L$ ID.

# …and future perspectives

- **More data:** $3\times$ larger sample ready to be analysed.
- **More input variables:** e.g. $K_L$ ID.
- **More channels:** extension of the technique to $B^0 \to K_S^0 \nu\bar{\nu}$ and $B \to K^* \nu\bar{\nu}$.

# …and future perspectives

- **More data:** $3\times$ larger sample ready to be analysed.
- **More input variables:** e.g. $K_L$ ID.
- **More channels:** extension of the technique to $B^0 \to K_S^0 \nu\bar{\nu}$ and $B \to K^* \nu\bar{\nu}$.
- **Reduction of systematics**: improvement of the continuum modelling.

# …and future perspectives

- **More data:** $3\times$ larger sample ready to be analysed.
- **More input variables:** e.g. $K_L$ ID.
- **More channels:** extension of the technique to $B^0 \to K_S^0 \nu\bar{\nu}$ and $B \to K^* \nu\bar{\nu}$.
- **Reduction of systematics**: improvement of the continuum modelling.
- **Improvement of the multivariate classification:** possible mixed NN and BDT architecture.

# …and future perspectives

- **More data:** $3\times$ larger sample ready to be analysed.
- **More input variables:** e.g. $K_L$ ID.
- **More channels:** extension of the technique to $B^0 \to K_S^0 \nu\bar{\nu}$ and $B \to K^* \nu\bar{\nu}$.
- **Reduction of systematics**: improvement of the continuum modelling.
- **Improvement of the multivariate classification:** possible mixed NN and BDT architecture.

# …and future perspectives

- **More data:** $3\times$ larger sample ready to be analysed.
- **More input variables:** e.g. $K_L$ ID.
- **More channels:** extension of the technique to $B^0 \to K^0_S\nu\bar{\nu}$ and $B \to K^*\nu\bar{\nu}$.
- **Reduction of systematics**: improvement of the continuum modelling.
- **Improvement of the multivariate classification:** possible mixed NN and BDT architecture.


$10^5 \times \sigma_{\rm BR}$ uncertainty for next analyses, assuming 25% improvement + 40% $K^0_S$

| | $63\ \mathrm{fb}^{-1}$ (arXiv:2104.12624) | $197\ \mathrm{fb}^{-1}$ (Summer 2021 – current lumi) | $450\ \mathrm{fb}^{-1}$ (Summer 2022 – expected) | $(450+700)\ \mathrm{fb}^{-1}$ (+ Belle I sample) |
|---|---|---|---|---|
| $\sigma_{BR}(K^+)$ | 1.55 | 0.78 | 0.52 | **0.32** |
| $\sigma_{BR}(K^+ + K^0_S)$ | - | 0.68 | 0.45 | **0.28** |

Preliminary

# Supplemental material

# The $B^+ \to K^+ \nu\bar{\nu}$ decay

**Scenarios beyond the SM → possible contribution of right-handed operators $Q_R^l$**

$$\mathcal{H}_{eff.} = -\frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_l (C_L^l Q_L^l + C_R^l Q_R^l) \quad \text{where} \quad Q_{L(R)}^l = \left(\bar{s}_{L(R)}\gamma_\mu b_{L(R)}\right)\left(\bar{\nu}_{L(R)}^l \gamma^\mu \nu_{L(R)}^l\right) \quad l = e, \mu, \tau$$

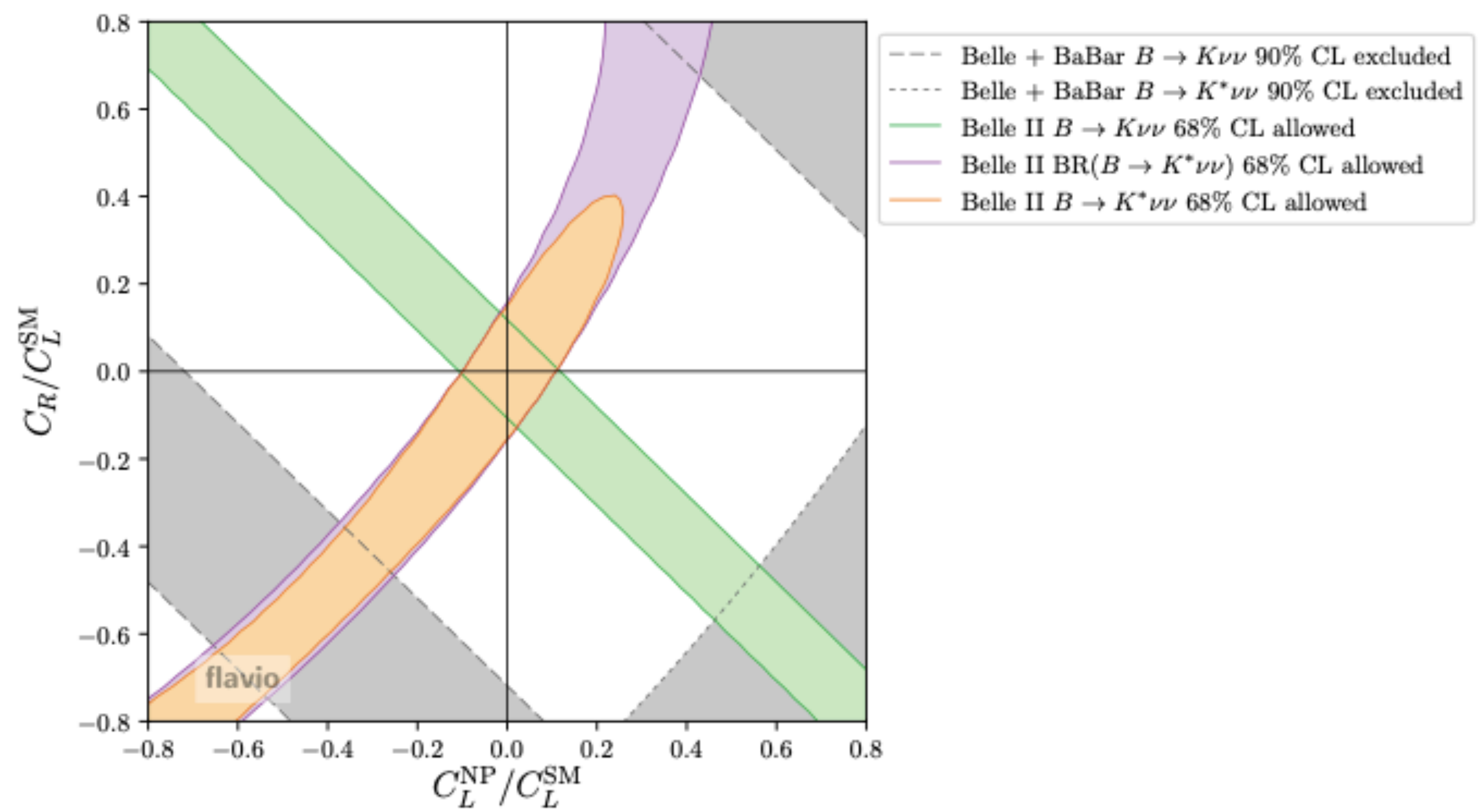
2 combinations of 6 Wilson Coefficients:

$$\frac{\mathrm{Br}(B \to K\nu\bar{\nu})}{\mathrm{Br}(B \to K\nu\bar{\nu})_{\mathrm{SM}}} = \frac{1}{3}\sum_\ell (1 - 2\,\eta_\ell)\,\epsilon_\ell^2\,,$$

$$\frac{\mathrm{Br}(B \to K^*\nu\bar{\nu})}{\mathrm{Br}(B \to K^*\nu\bar{\nu})_{\mathrm{SM}}} = \frac{1}{3}\sum_\ell (1 + \kappa_\eta \eta_\ell)\,\epsilon_\ell^2\,,$$

$$\epsilon_\ell = \frac{\sqrt{|C_L^\ell|^2 + |C_R^\ell|^2}}{|C_L^{\mathrm{SM}}|}\,,$$

$$\eta_\ell = \frac{-\mathrm{Re}\left(C_L^\ell C_R^{\ell *}\right)}{|C_L^\ell|^2 + |C_R^\ell|^2}\,.$$

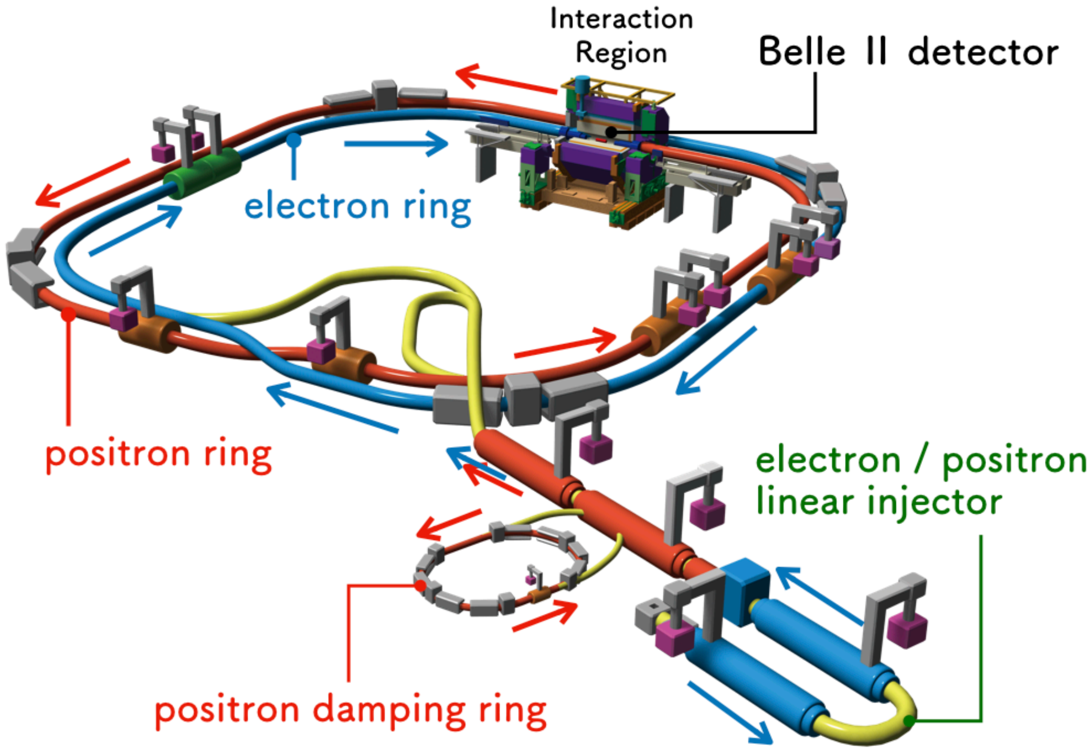



**Constraint on new-physics contributions:** Wilson coefficients $C_L^{NP}$ and $C_R$ normalised to the SM value of $C_L$ (Belle II from expected 50 $ab^{-1}$).

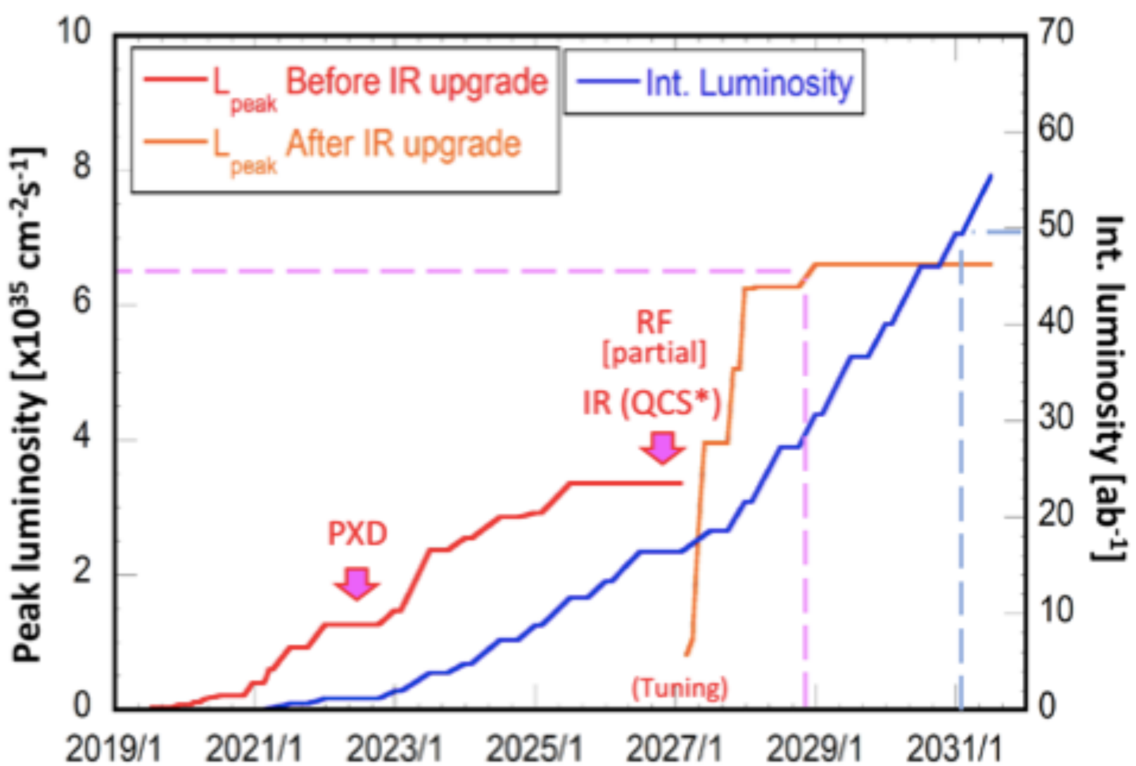
**Significant increase in the $B \to K^{(*)}\nu\bar{\nu}$ decay BR can be accommodated in models describing CC and NC anomalies with leptoquarks. [arXiv:2107.01080v2]**

# SuperKEKB

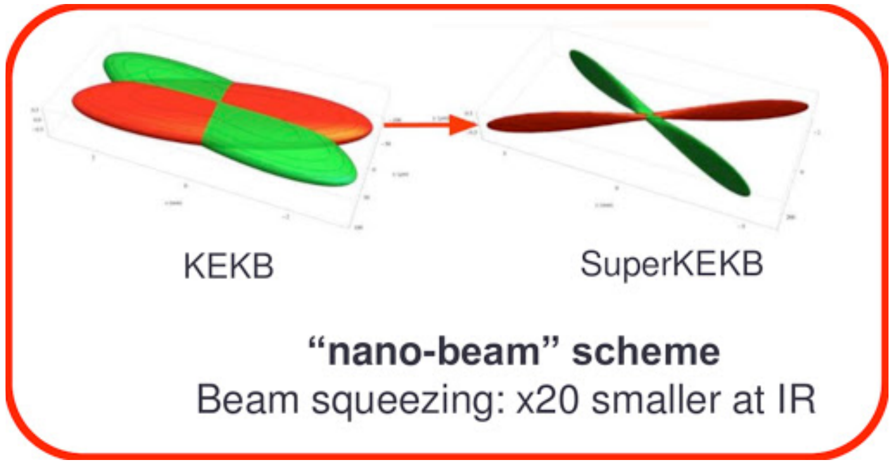
- Asymmetric-energy $e^+e^-$ collider operating at $\sqrt{s} = 10.58$ GeV $\to \Upsilon(4S)$ resonance.
- Second generation B factory based on the **nanobeam scheme:** major upgrade of its predecessor KEKB.
- **World highest instantaneous luminosity:** $3.12 \times 10^{34}$ cm$^{-2}$ s$^{-1}$.




- Peak luminosity projections:

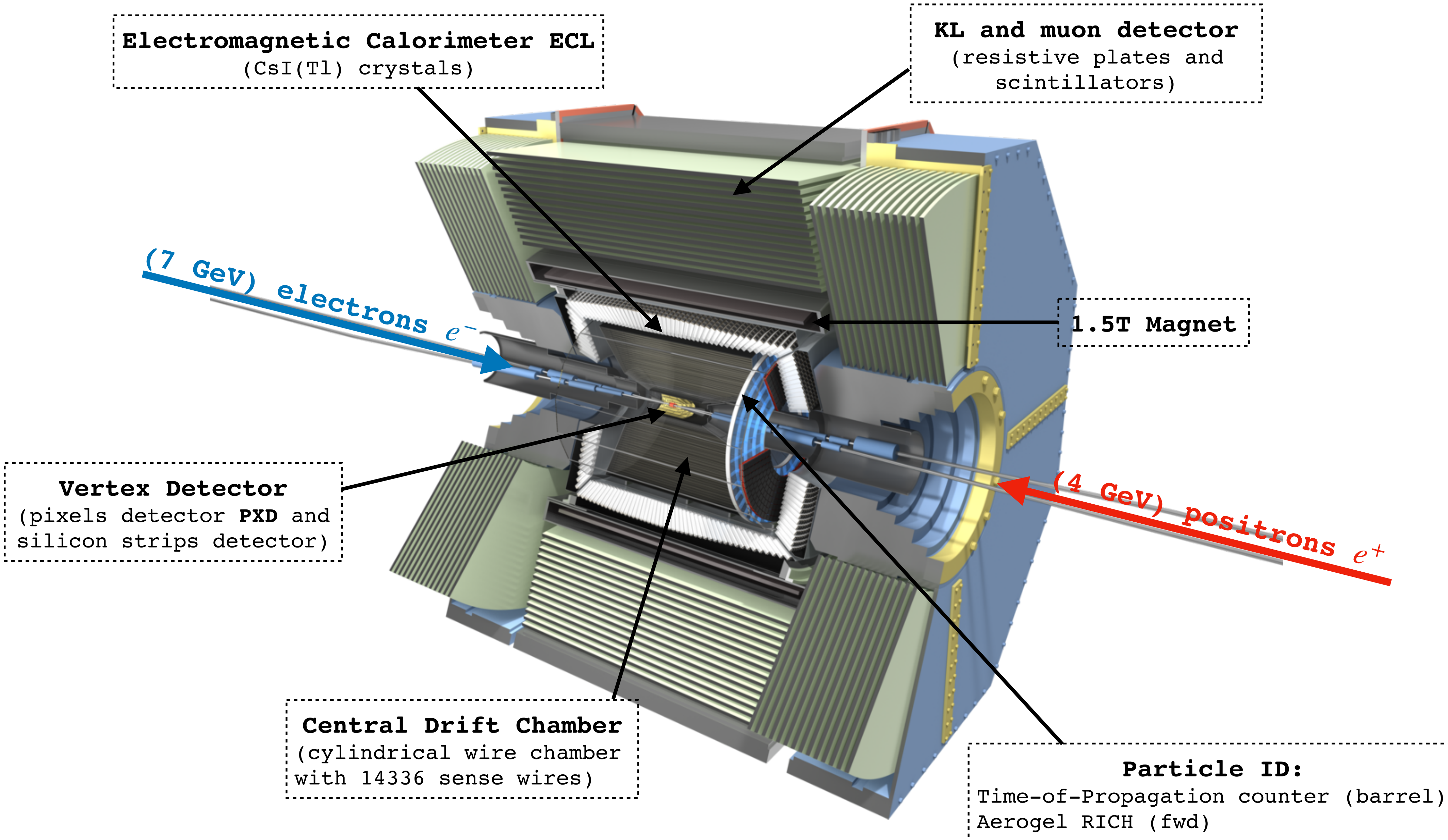



- Nano-beam scheme:



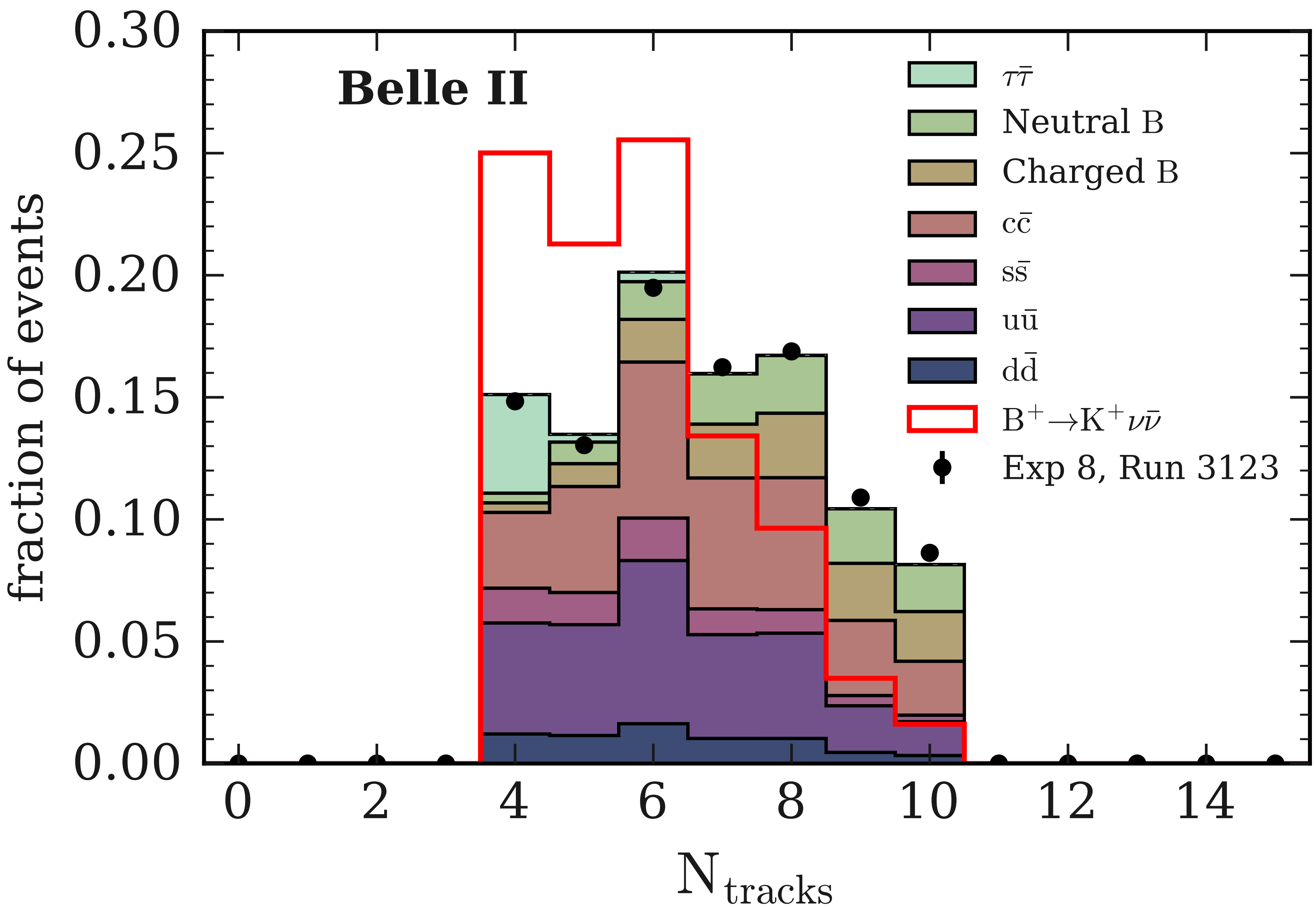
# The Belle II detector

New detector with respect to the predecessor Belle.



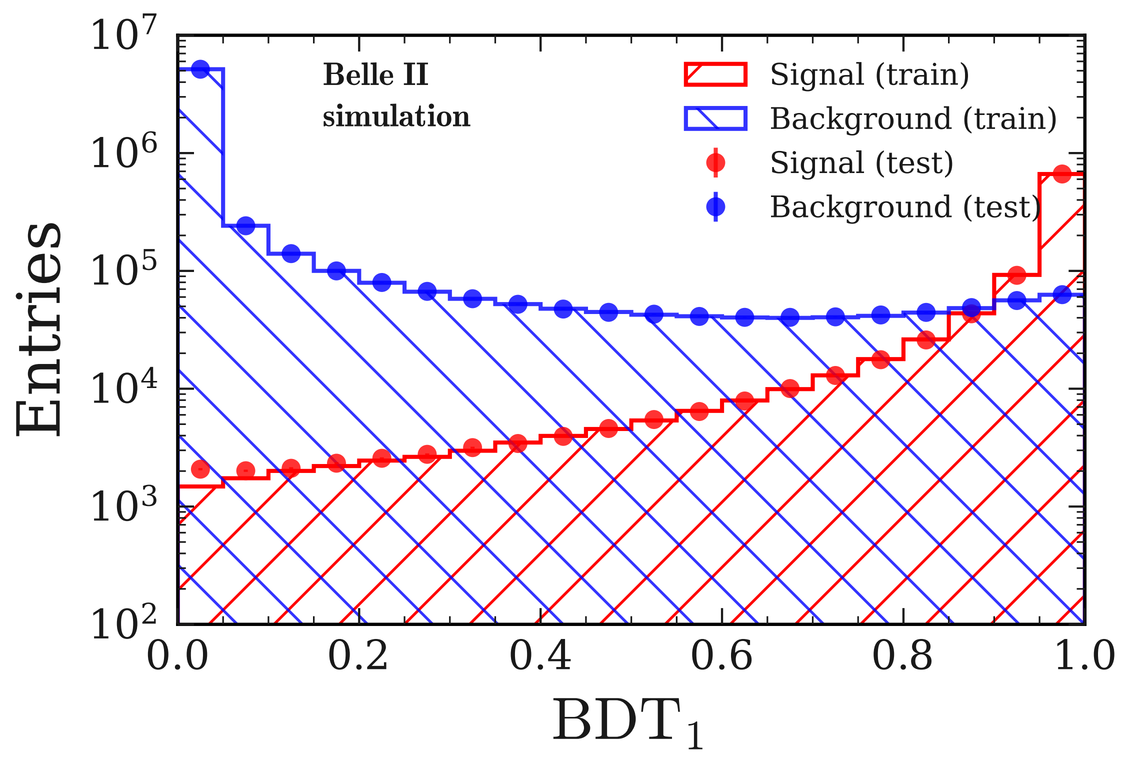
# Features of $B^+ \to K^+ \nu\bar{\nu}$

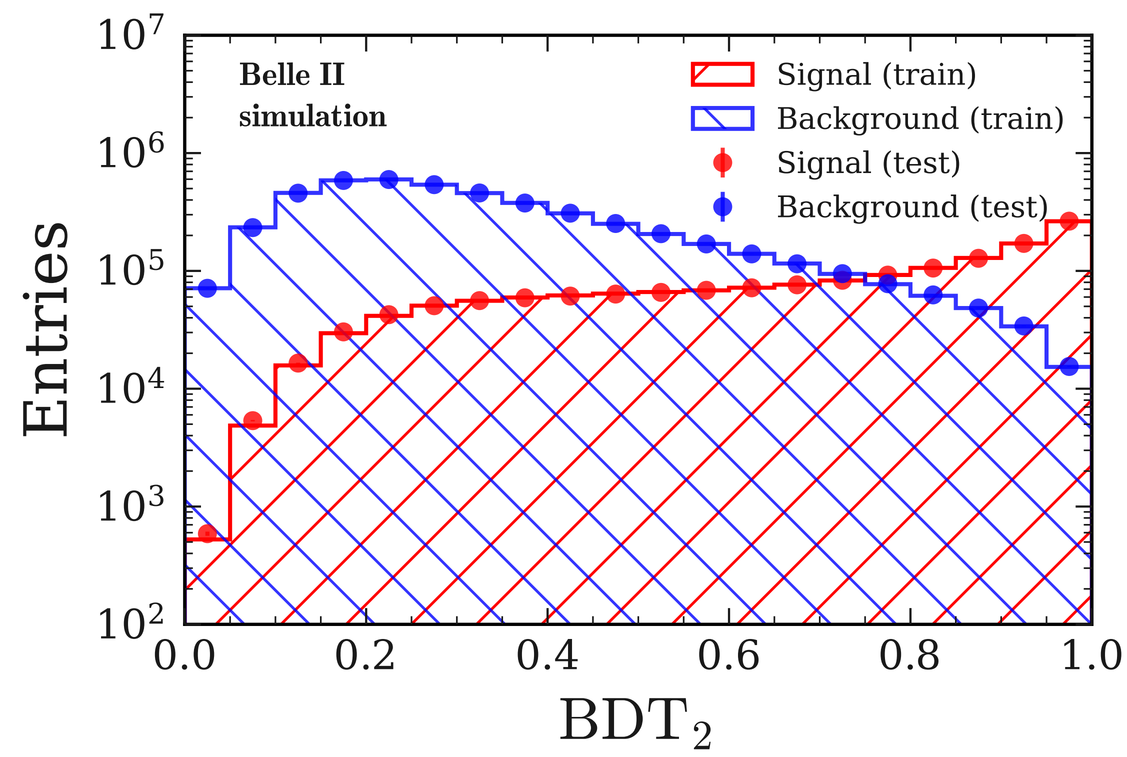
- Number of reconstructed tracks in the event.



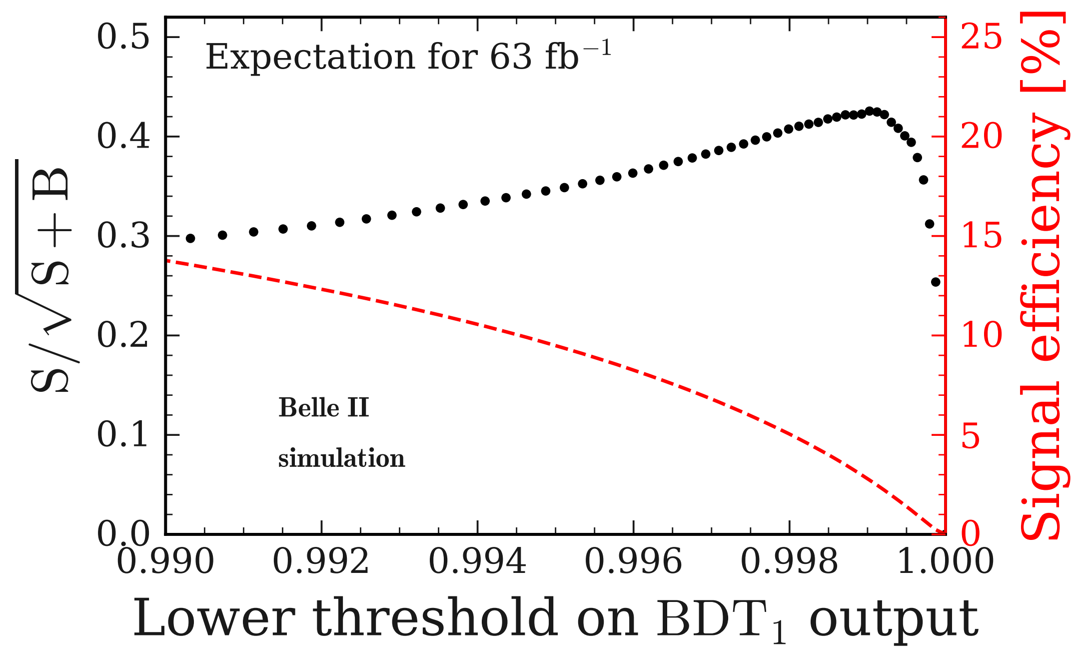
# More on multivariate classification

- No overfitting observed neither for BDT1 nor for BDT2.




- Signal sensitivity of BDT1:



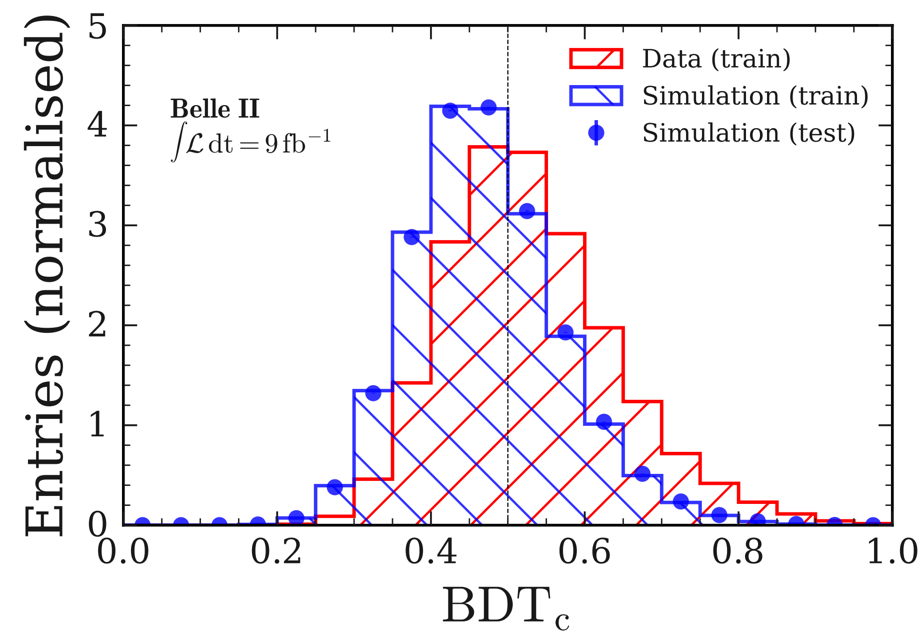
# Reweighting of continuum MC

Discrepancies between simulated continuum and off-resonance data.

**Data-driven correction** by means of an additional fastBDT: $\mathrm{BDT_c}$.

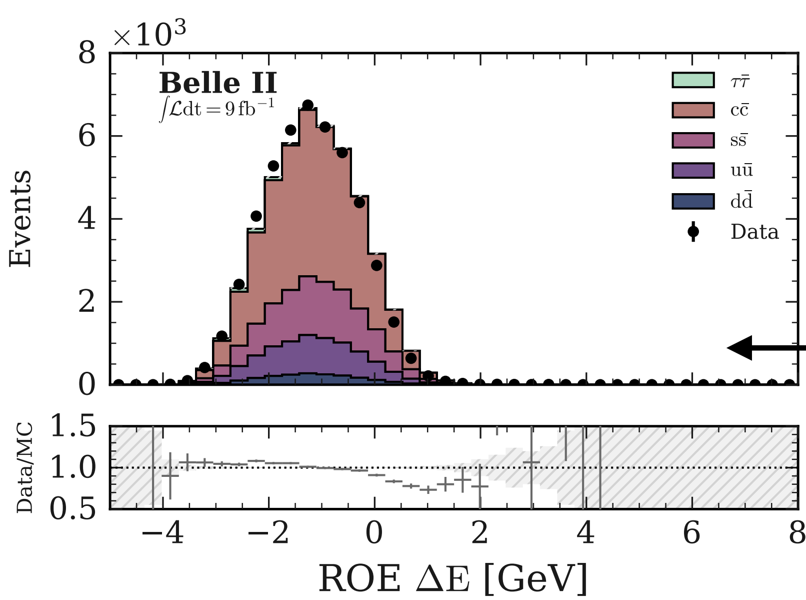
- Select simulated continuum ($100\ \mathrm{fb}^{-1}$) with $\mathrm{BDT_1} > 0.9$;
- Select off-resonance data ($9\ \mathrm{fb}^{-1}$) with $\mathrm{BDT_1} > 0.9$;
- Train $\mathrm{BDT_c}$ with the set of 51 variables using data as signal and simulation as bkg;

**Before reweighting:** no perfect overlap at 0.5 → mismodelling.



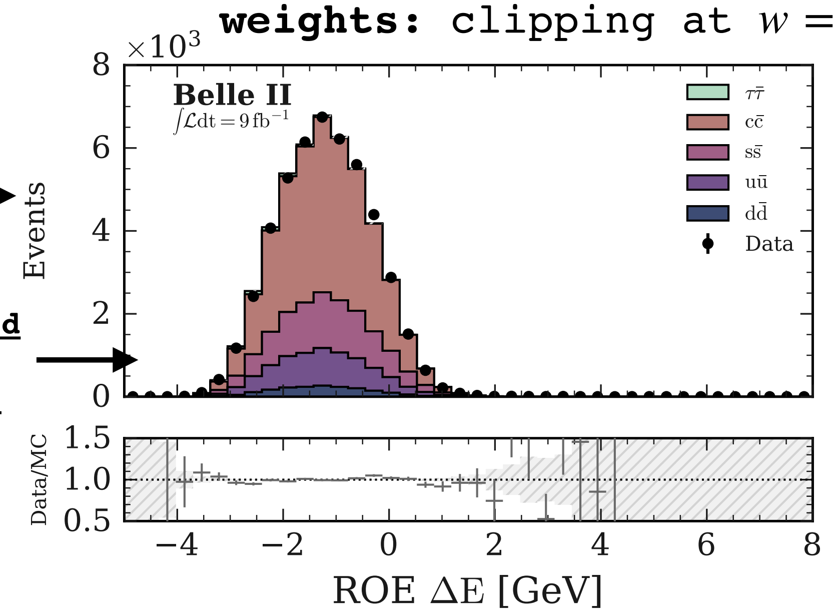

- Being $p$ the $\mathrm{BDT_c}$ score, apply the **event weight** $p/(1-p) = \mathrm{P(Data-like)/P(MC-like)}$ **to correct the simulated continuum.**

**Protection against large weights:** clipping at $w = 10$



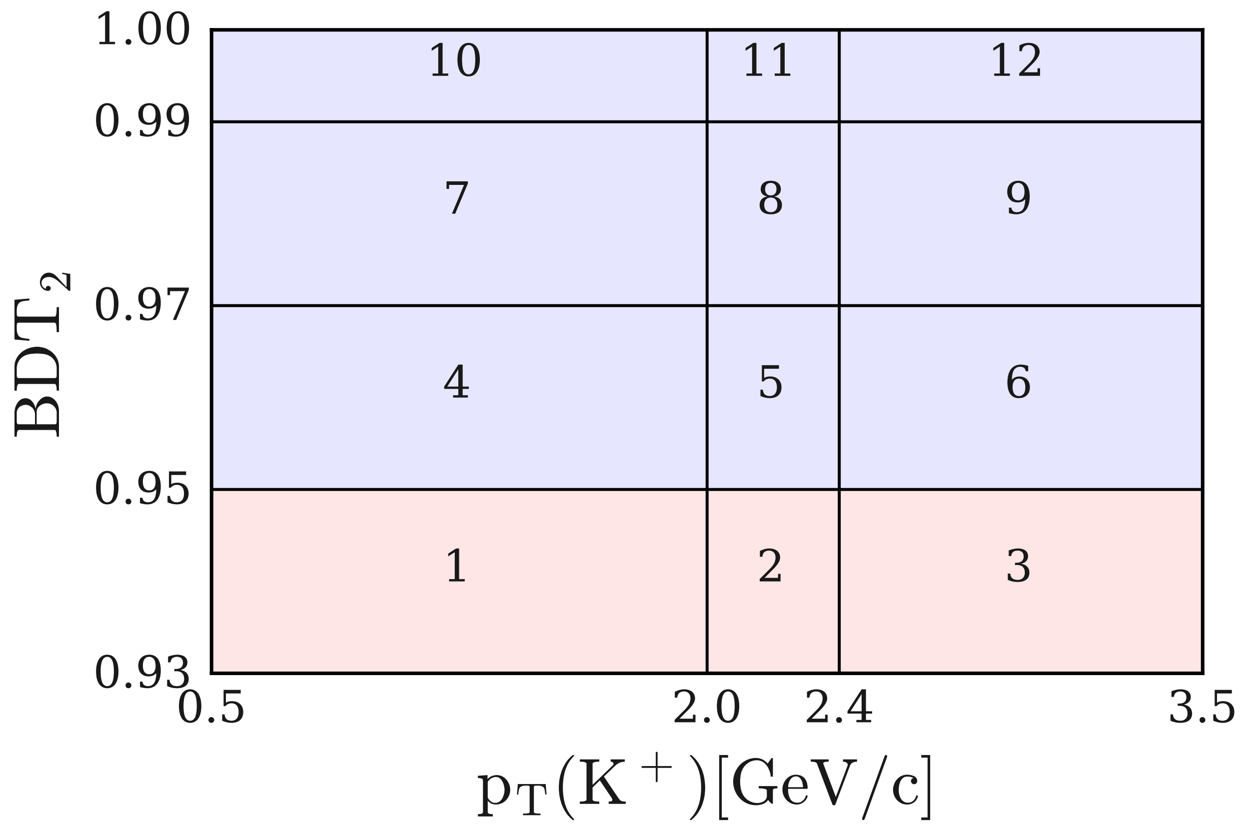

$$w_{event} = \frac{\mathrm{P(Data-like)}}{\mathrm{P(MC-like)}}$$

**Continuum MC yields scaled up to Data of normalisation ratio 1.22**



# Definition of the fit region

Optimised bin boundaries set up in the $p_T(K^+) \times BDT_2$ space:




**Bins 4,5,6,7,8,9,10,11,12:**

- **Signal Region (SR):** fit of data at the $\Upsilon(4S)$ resonance;
- **Control Region 2** (CR2)**:** fit of off-resonance data.

**Bins 1,2,3:**

- **Control Region 1 (CR1):** fit of data at the $\Upsilon(4S)$ resonance;
- **Control Region 3 (CR3):** fit of off-resonance data.

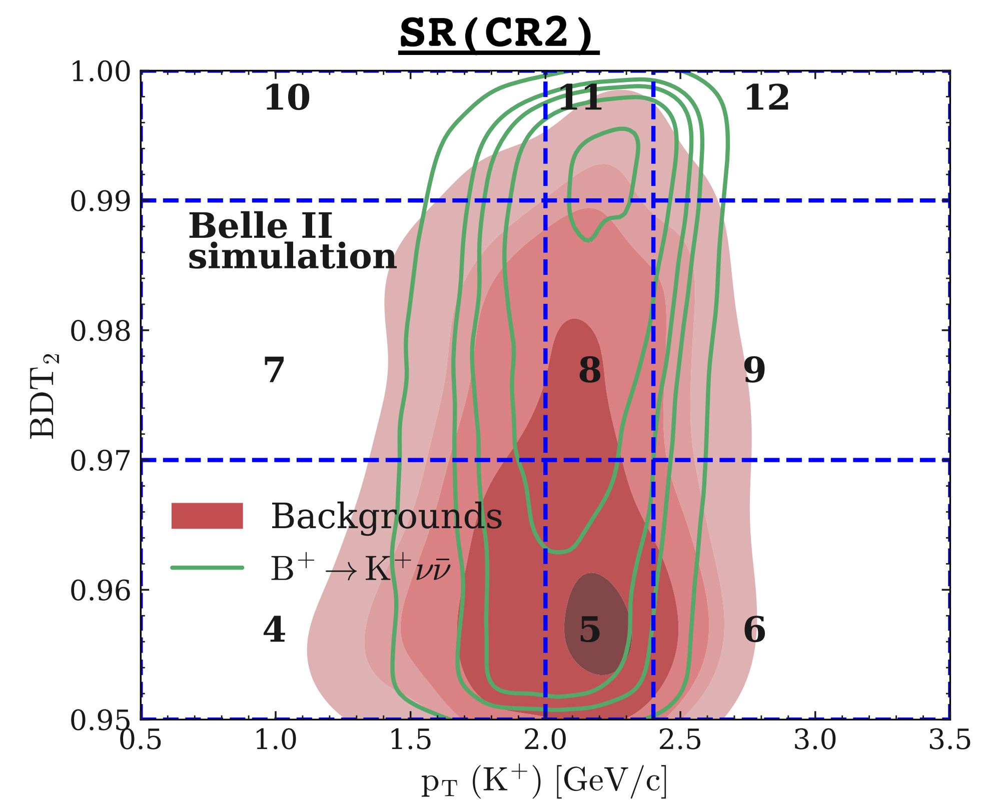
**Control Region 1-2-3 to constrain bkg's yields.**

# The fit region

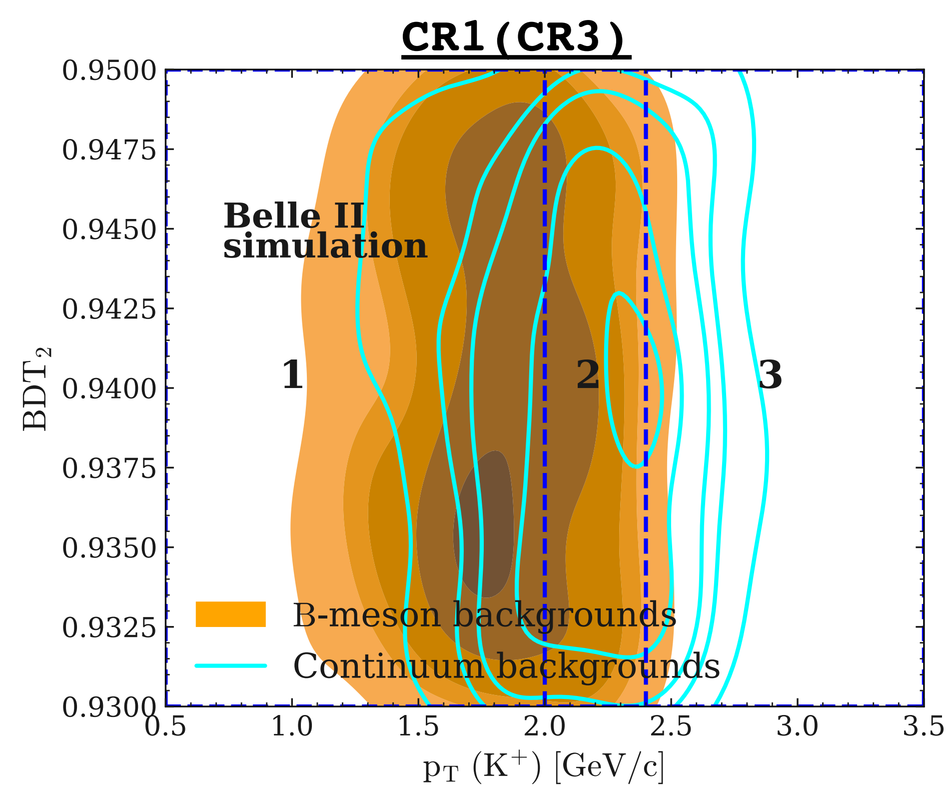
- 1 signal region + 3 control regions.

**Bin boundaries in the SR** specifically optimised by minimisation of the expected upper limit on the $BR(B^+ \to K^+ \nu\bar{\nu})$.

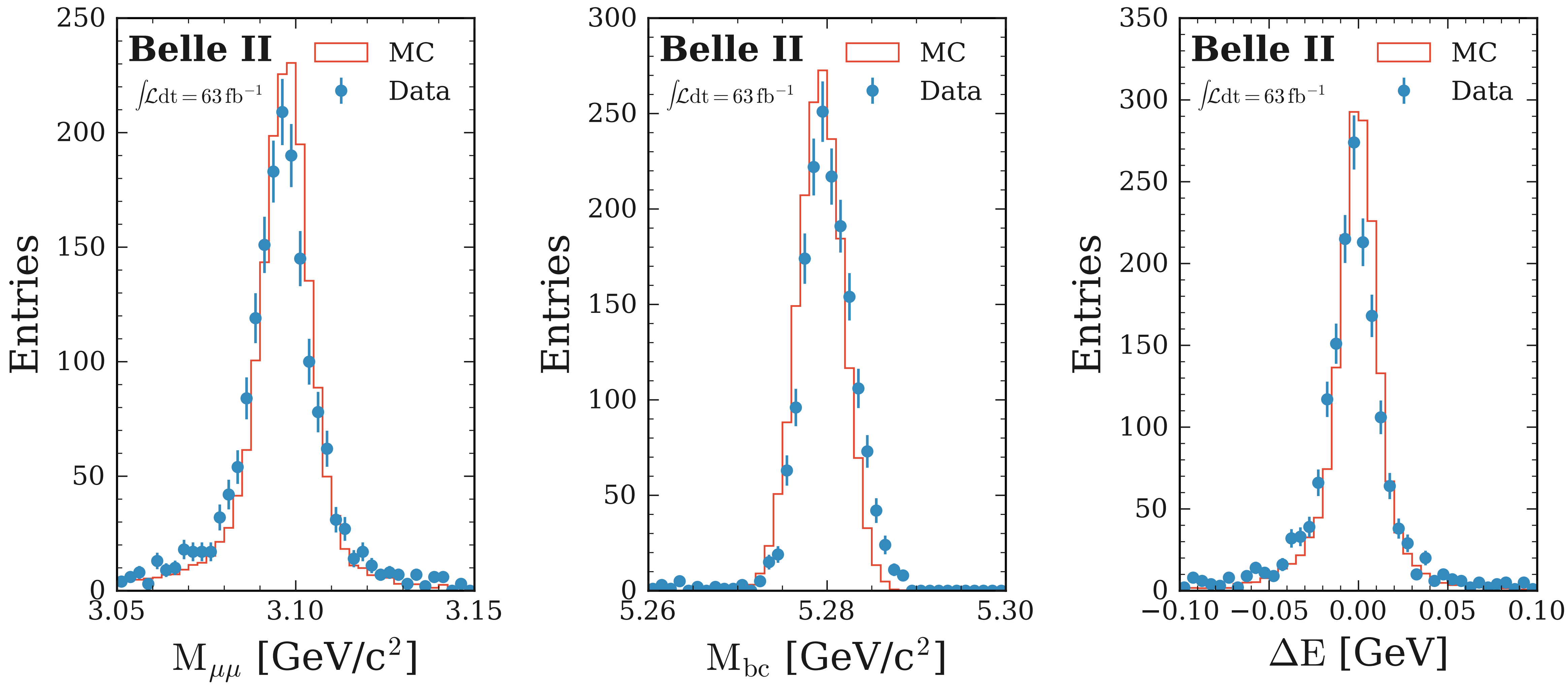
| Region | 2D Bin Boundary Definition | Physics Processes | $\sqrt{s}$ |
|---|---|---|---|
| Signal Region (SR) | $p_T(K^+) \in [0.5, 2.0, 2.4, 3.5]$ GeV/c<br>$BDT_2 \in [0.95, 0.97, 0.99, 1.0]$ | signal + all backgrounds | $\Upsilon(4S)$ |
| Control Region 1 (CR1) | $p_T(K^+) \in [0.5, 2.0, 2.4, 3.5]$ GeV/c<br>$BDT_2 \in [0.93, 0.95]$ | signal + all backgrounds | $\Upsilon(4S)$ |
| Control Region 2 (CR2) | $p_T(K^+) \in [0.5, 2.0, 2.4, 3.5]$ GeV/c<br>$BDT_2 \in [0.95, 0.97, 0.99, 1.0]$ | continuum backgrounds | off-resonance ($-60$ MeV/$c^2$) |
| Control Region 3 (CR3) | $p_T(K^+) \in [0.5, 2.0, 2.4, 3.5]$ GeV/c<br>$BDT_2 \in [0.93, 0.95]$ | continuum backgrounds | off-resonance ($-60$ MeV/$c^2$) |




**CR 1-2-3 to constrain bkg yields.**



# Identification of $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events

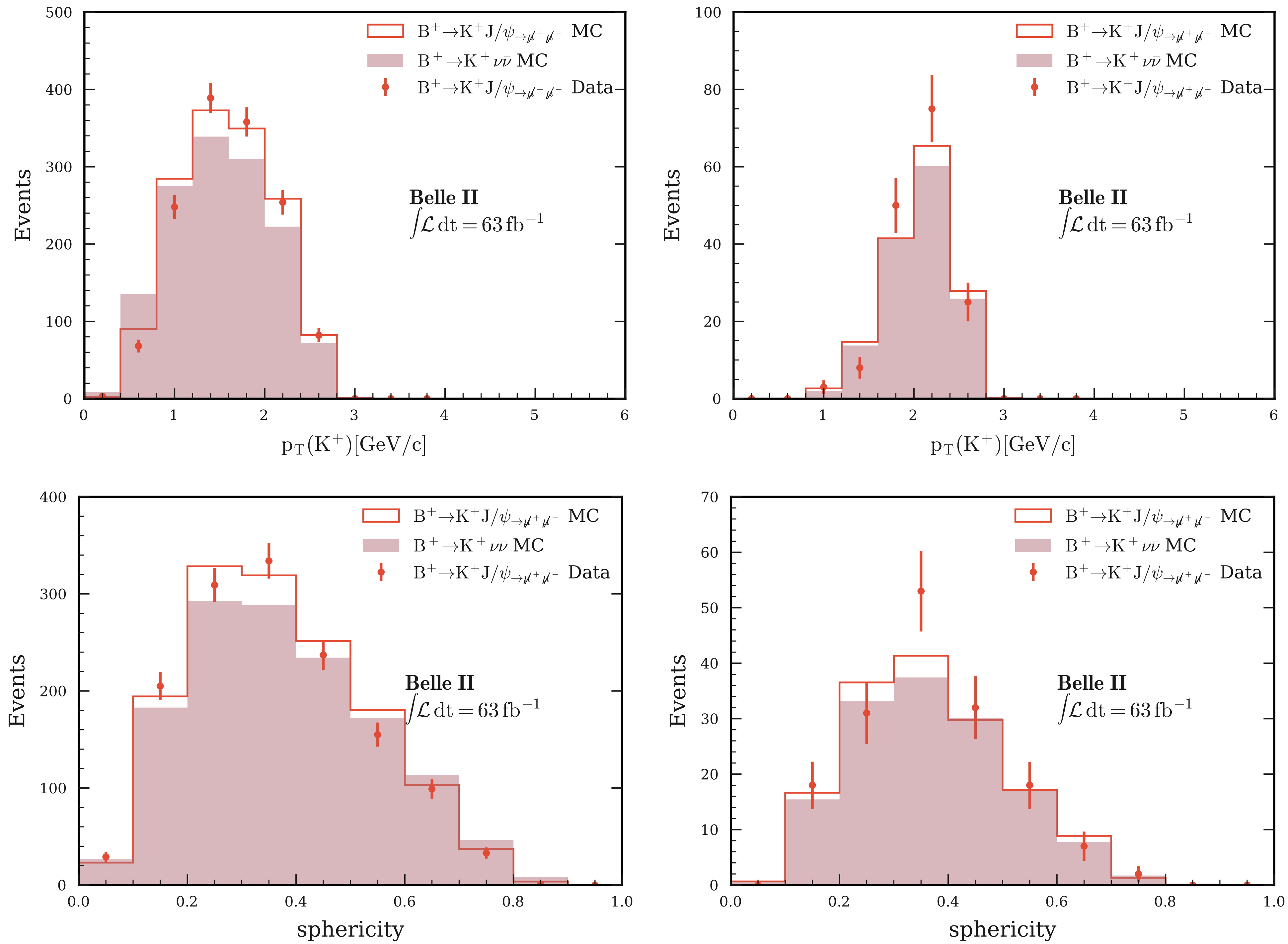



$$M_{bc} = \sqrt{E_{beam}^2 - (\vec{P}_B^{CMS})^2}$$

$$\Delta E = \sum_i E_i^{CMS} - E_{beam}$$

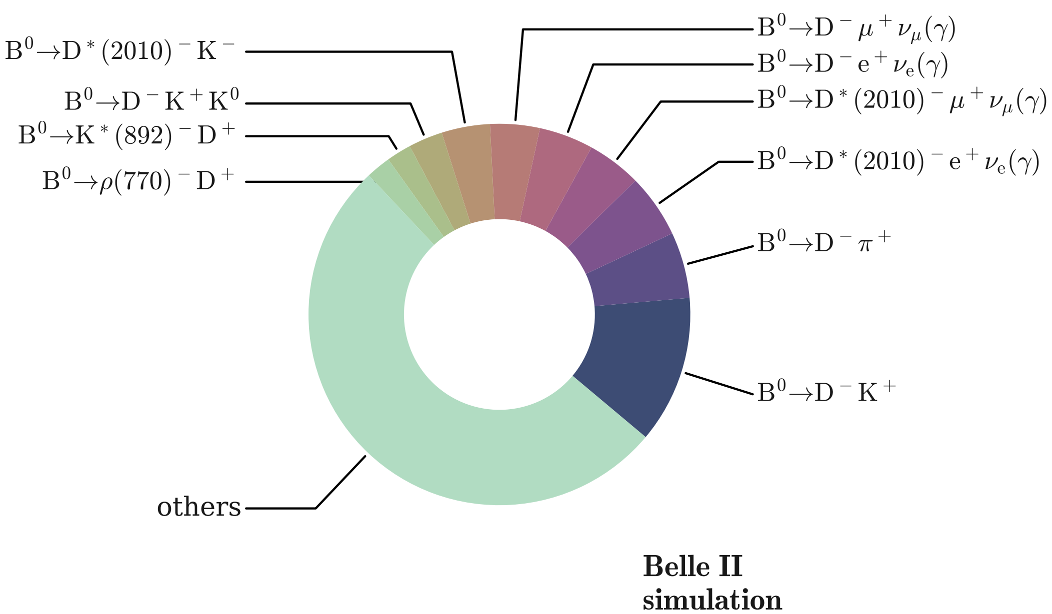
1720 data events from 63 $fb^{-1}$ + bkg suppressed to percent level.

# Results of the validation on $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$

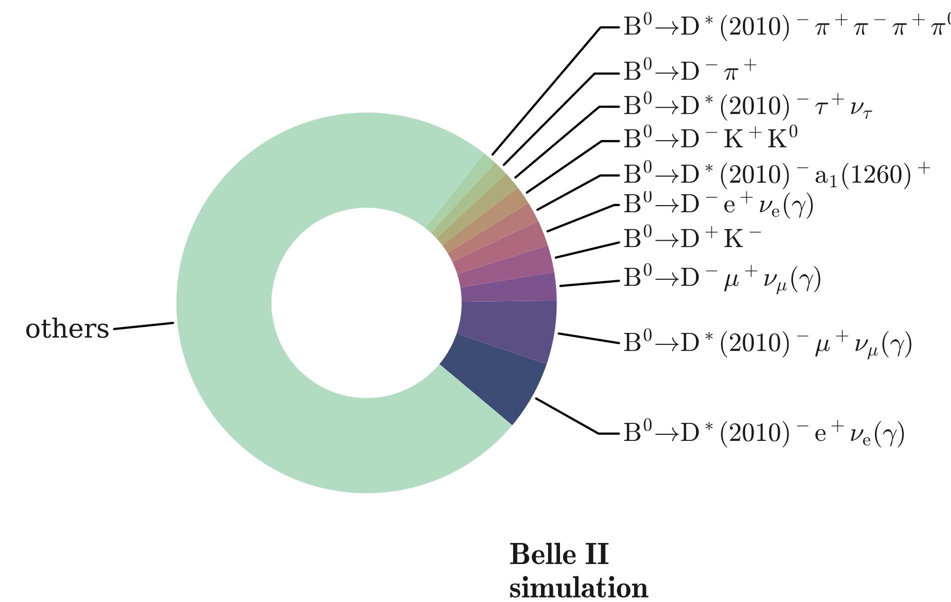


# Background composition in the fit region

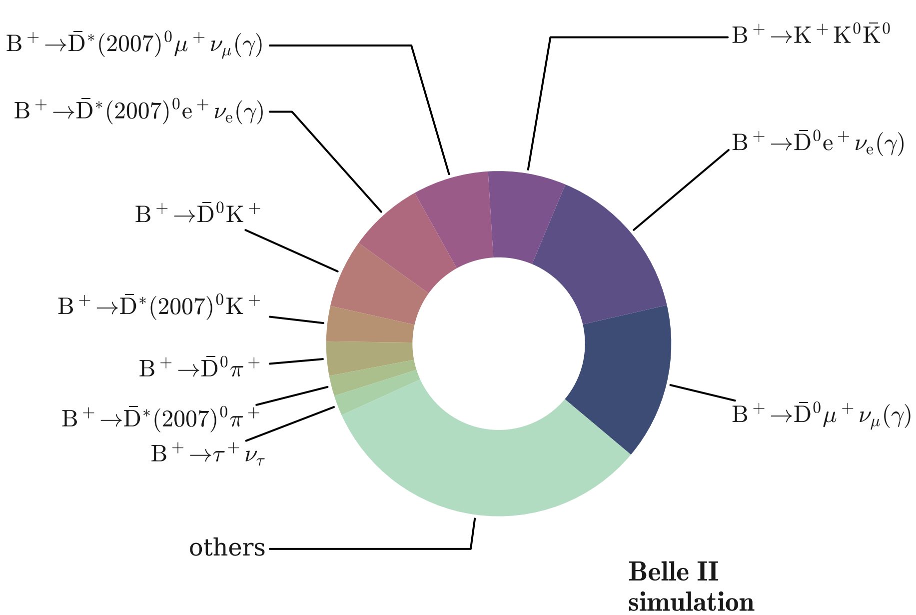
- $B^0\bar{B}^0$ signal side:



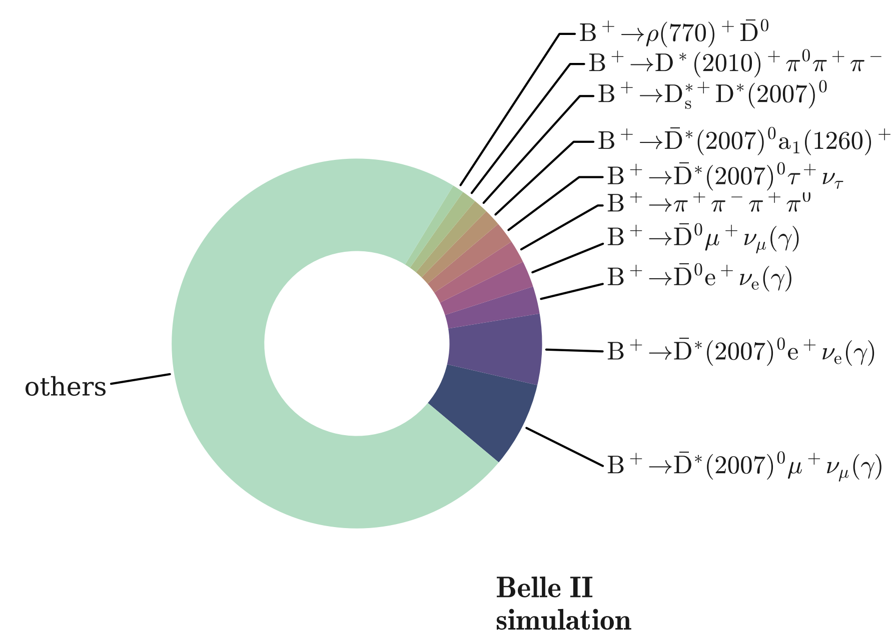

- $B^0\bar{B}^0$ tag side:




- $B^+B^-$ signal side:

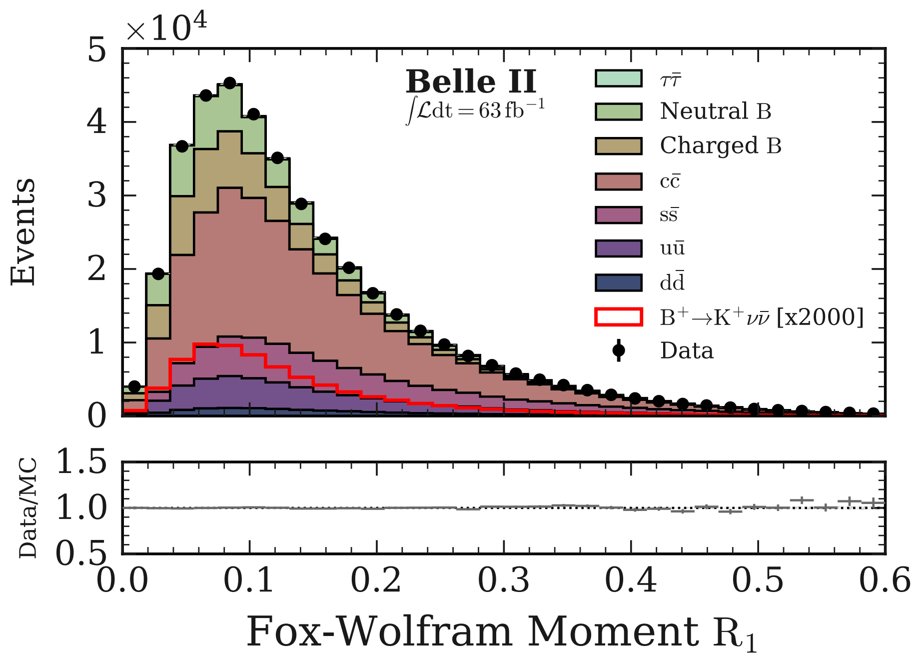


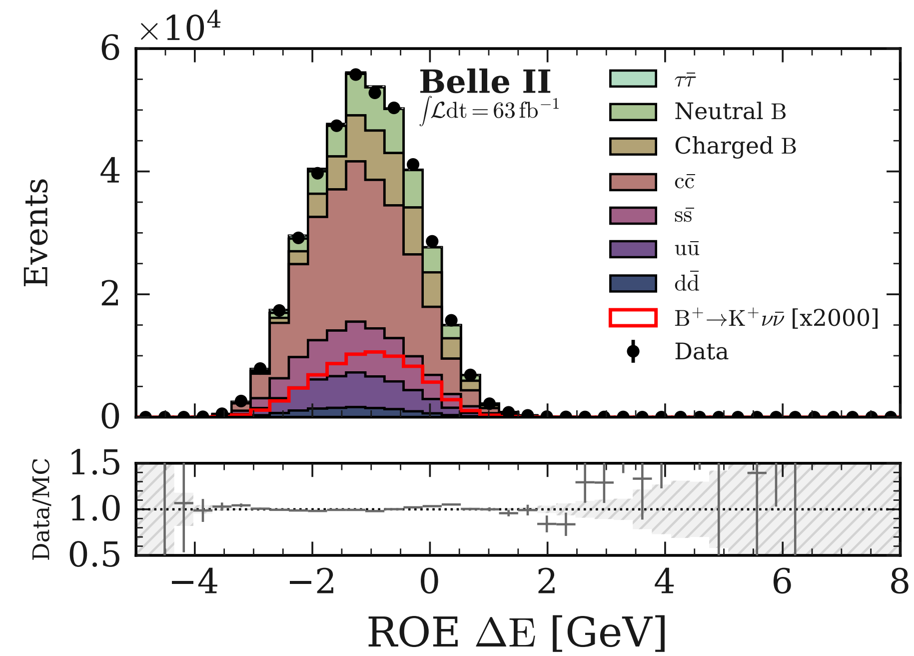

- $B^+\bar{B}^-$ tag side:

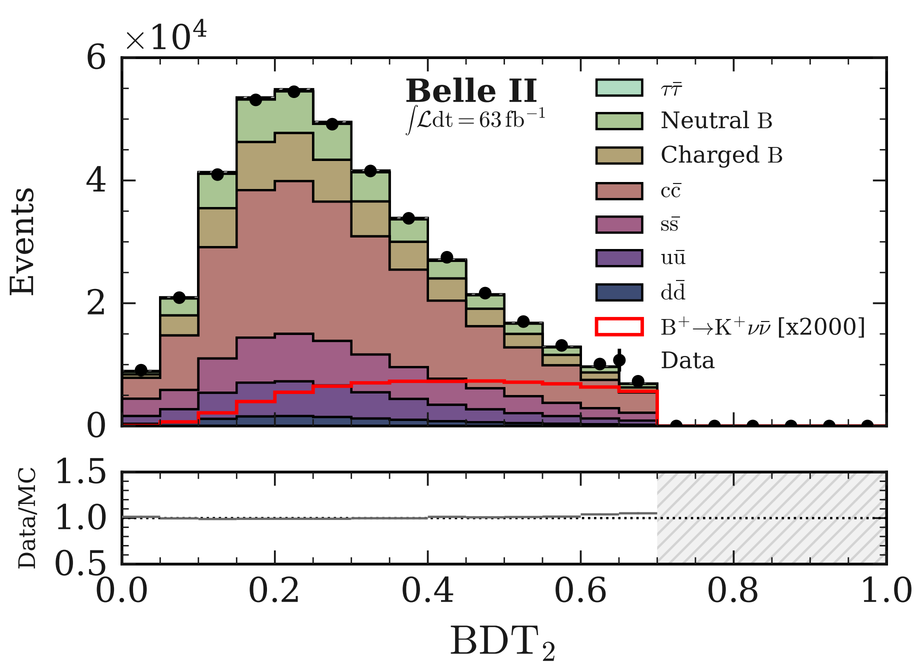


# Validation in the BDT sideband

- Agreement between $\Upsilon(4S)$ on-resonance data and simulation in the sideband $0.9 < \mathrm{BDT}_1 < 0.99$ and $\mathrm{BDT}_2 < 0.7$:



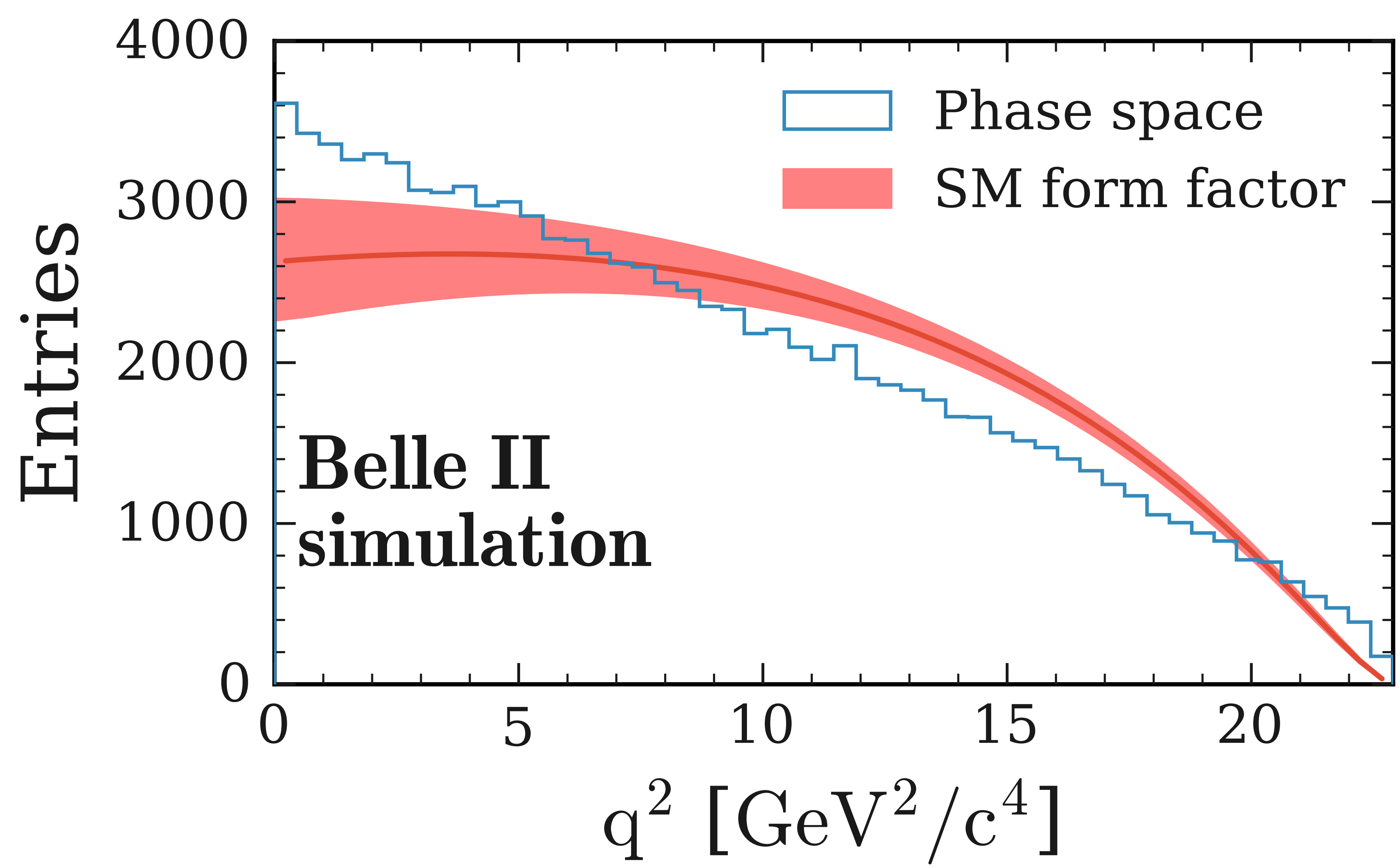




- **Only if the continuum background is scaled by a factor of 1.22 as obtained from the comparison with off-resonance data, the data/MC ratio is then 1.00 in the moderate BDT sideband.**

# SM form factor vs q2

- $q^2$ spectrum from PHSP simulation compared to the SM form factor from **[J. High Energ. Phys. 2015, 184 (2015)]** as a function of $q^2$.



# Fit procedure

- pyhf modifiers and constraints:

$$\nu_{cb}(\phi) = \sum_{s \in \text{ samples}} \nu_{scb}(\eta, \chi) = \sum_{s \in \text{ samples}} \underbrace{\left(\prod_{\kappa \in \kappa} \kappa_{scb}(\eta, \chi)\right)}_{\text{multiplicative modifiers}} \left(\nu^0_{scb}(\eta, \chi) + \underbrace{\sum_{\Delta \in \Delta} \Delta_{scb}(\eta, \chi)}_{\text{additive modifiers}}\right).$$

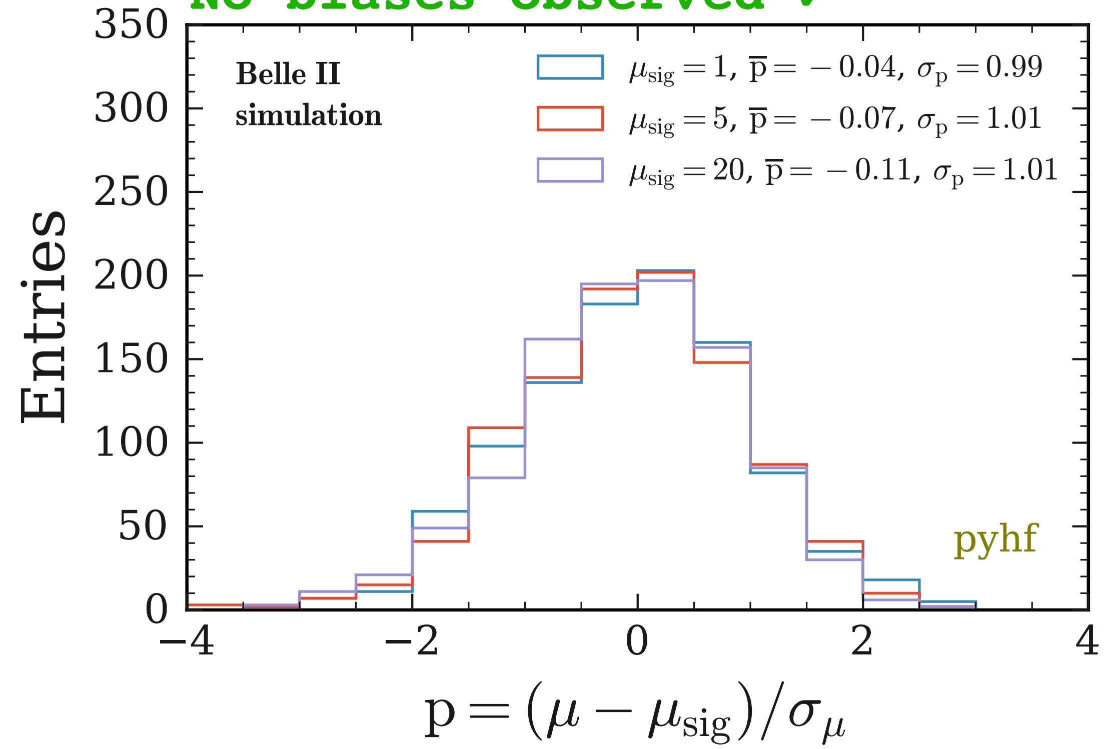
| Description | Modification | Constraint Term $c_\chi$ | Input |
|---|---|---|---|
| Uncorrelated Shape | $\kappa_{scb}(\gamma_b) = \gamma_b$ | $\prod_b \text{Pois}\left(r_b = \sigma_b^{-2} \middle\| \rho_b = \sigma_b^{-2}\gamma_b\right)$ | $\sigma_b$ |
| Correlated Shape | $\Delta_{scb}(\alpha) = f_p\left(\alpha \middle\| \Delta_{scb,\alpha=-1}, \Delta_{scb,\alpha=1}\right)$ | $\text{Gaus}\left(a = 0 \middle\| \alpha, \sigma = 1\right)$ | $\Delta_{scb,\alpha=\pm 1}$ |
| Normalisation Unc. | $\kappa_{scb}(\alpha) = g_p\left(\alpha \middle\| \kappa_{scb,\alpha=-1}, \kappa_{scb,\alpha=1}\right)$ | $\text{Gaus}\left(a = 0 \middle\| \alpha, \sigma = 1\right)$ | $\kappa_{scb,\alpha=\pm 1}$ |
| MC Stat. Uncertainty | $\kappa_{scb}(\gamma_b) = \gamma_b$ | $\prod_b \text{Gaus}\left(a_{\gamma_b} = 1 \middle\| \gamma_b, \delta_b\right)$ | $\delta_b^2 = \sum_s \delta_{sb}^2$ |
| Luminosity | $\kappa_{scb}(\lambda) = \lambda$ | $\text{Gaus}\left(l = \lambda_0 \middle\| \lambda, \sigma_\lambda\right)$ | $\lambda_0, \sigma_\lambda$ |
| Normalisation | $\kappa_{scb}(\mu_b) = \mu_b$ | | |
| Data-driven Shape | $\kappa_{scb}(\gamma_b) = \gamma_b$ | | |

# Fit validation

- **Test with injected signal:**

Check the pulls $\left(\mu_{\text{fit}} - \mu_{\text{inj.}}\right)/\sigma_{\mu}$

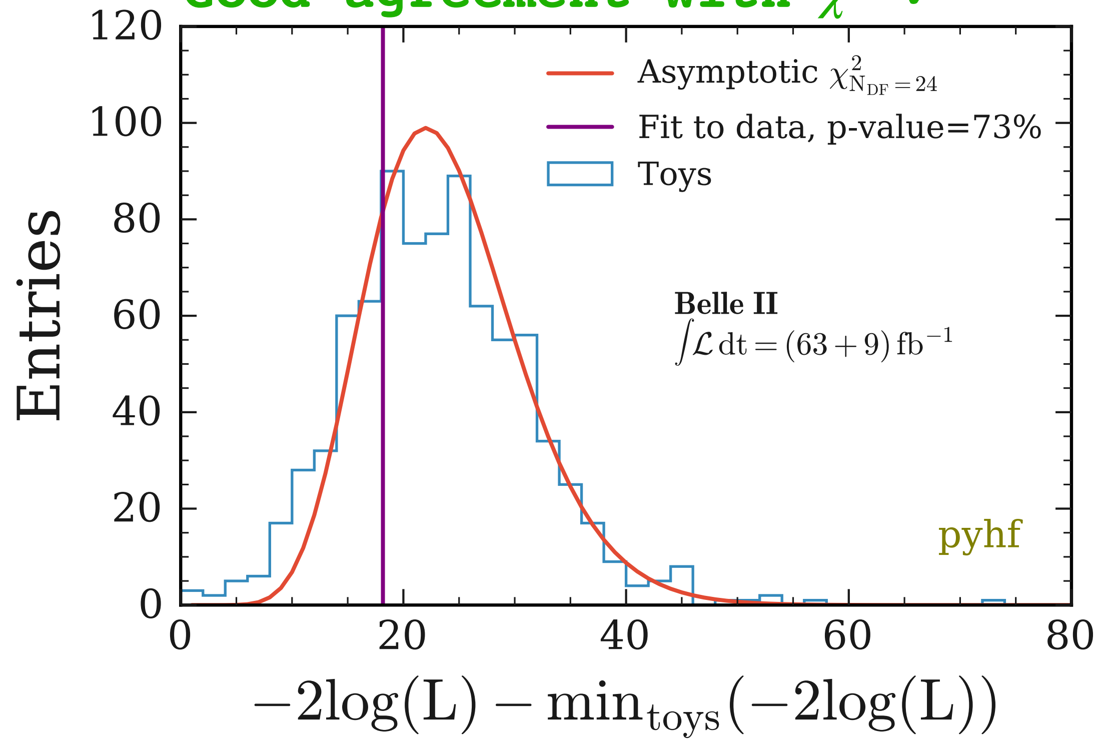
No biases observed ✔



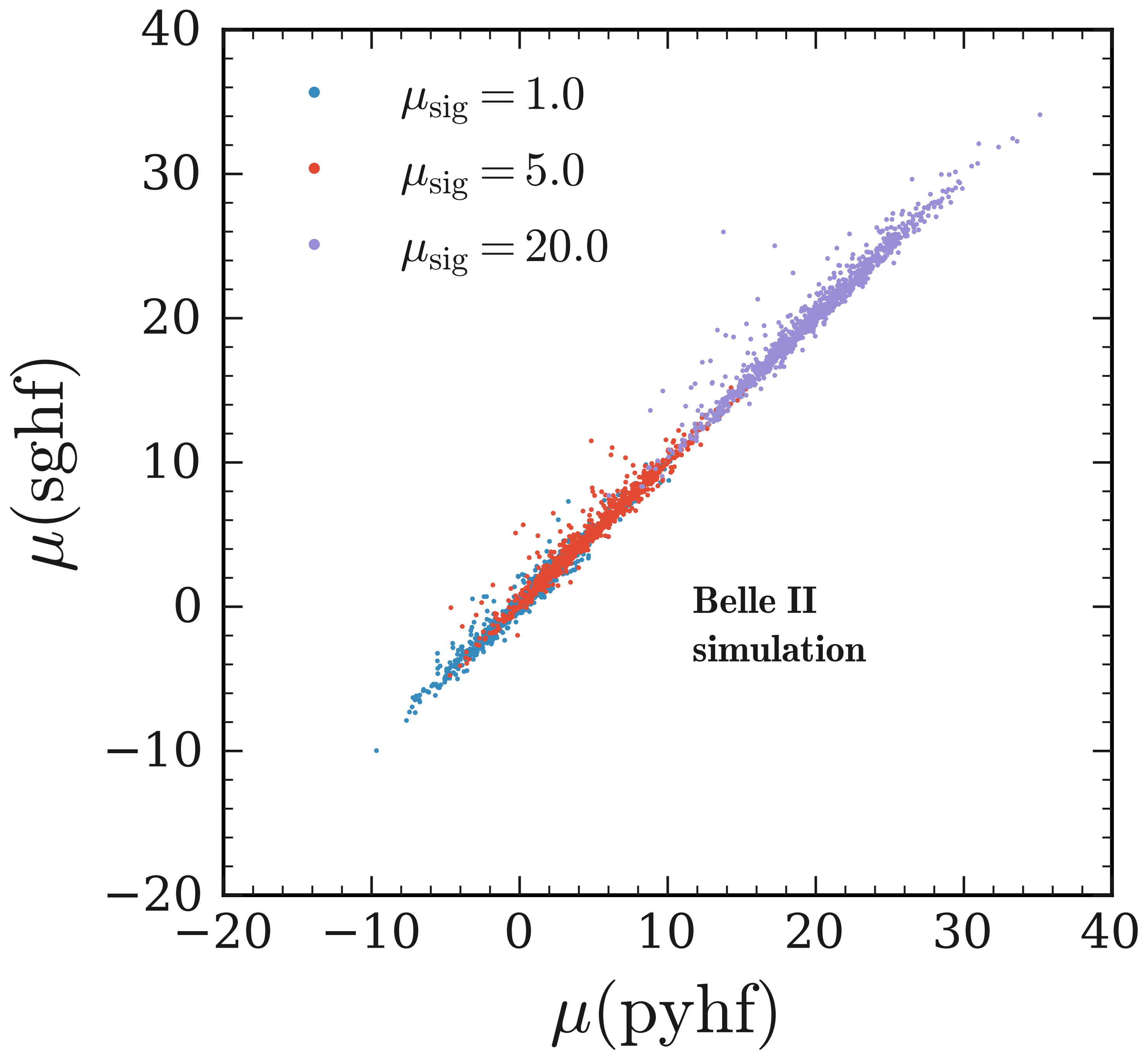

- **Test the fit quality:**

Check the p-value of the fit on observations

High p-value ✔
Good agreement with $\chi^2$ ✔

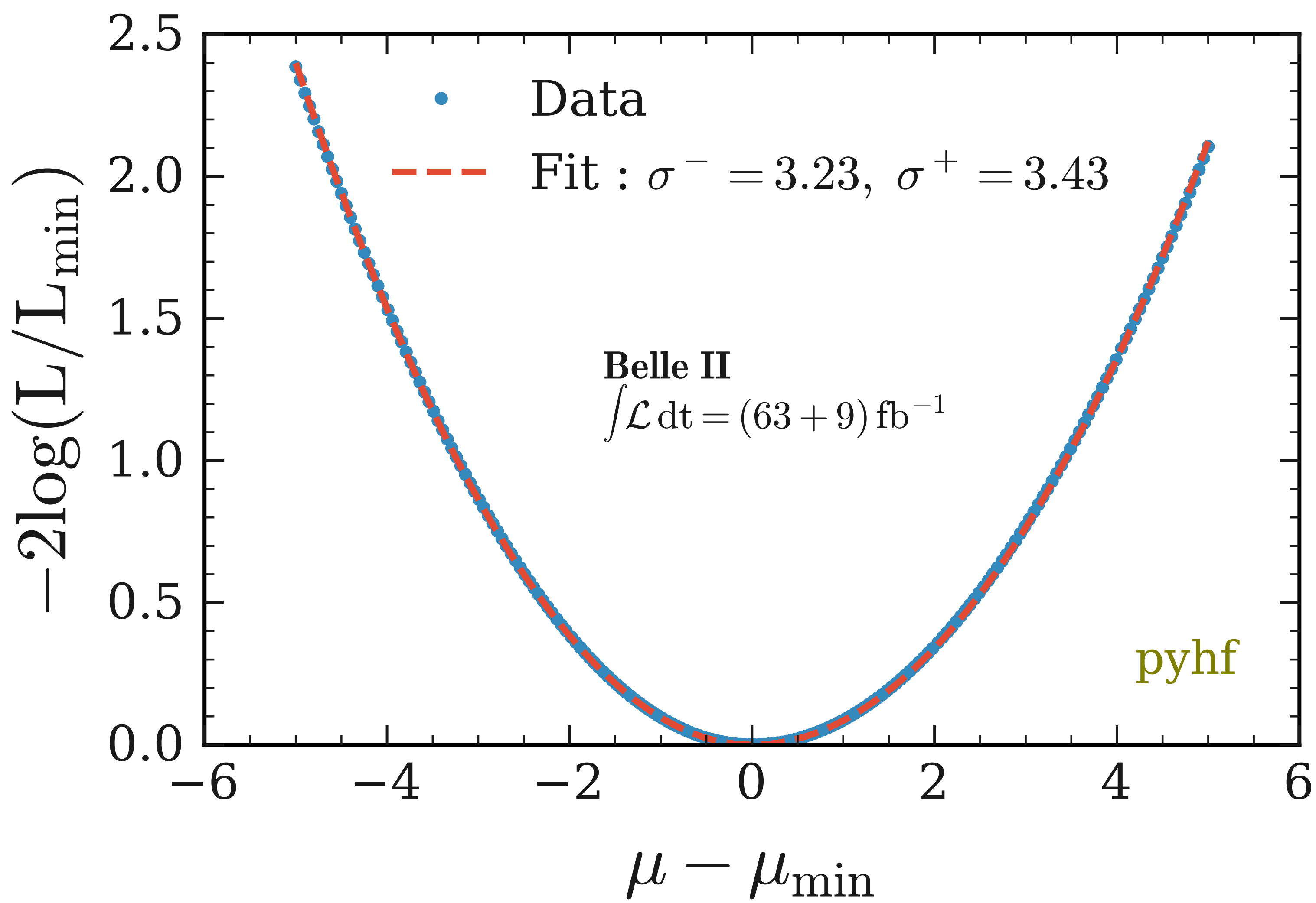


# Cross validation of PyHf with a simplified Gaussian model



# Fit to the Data

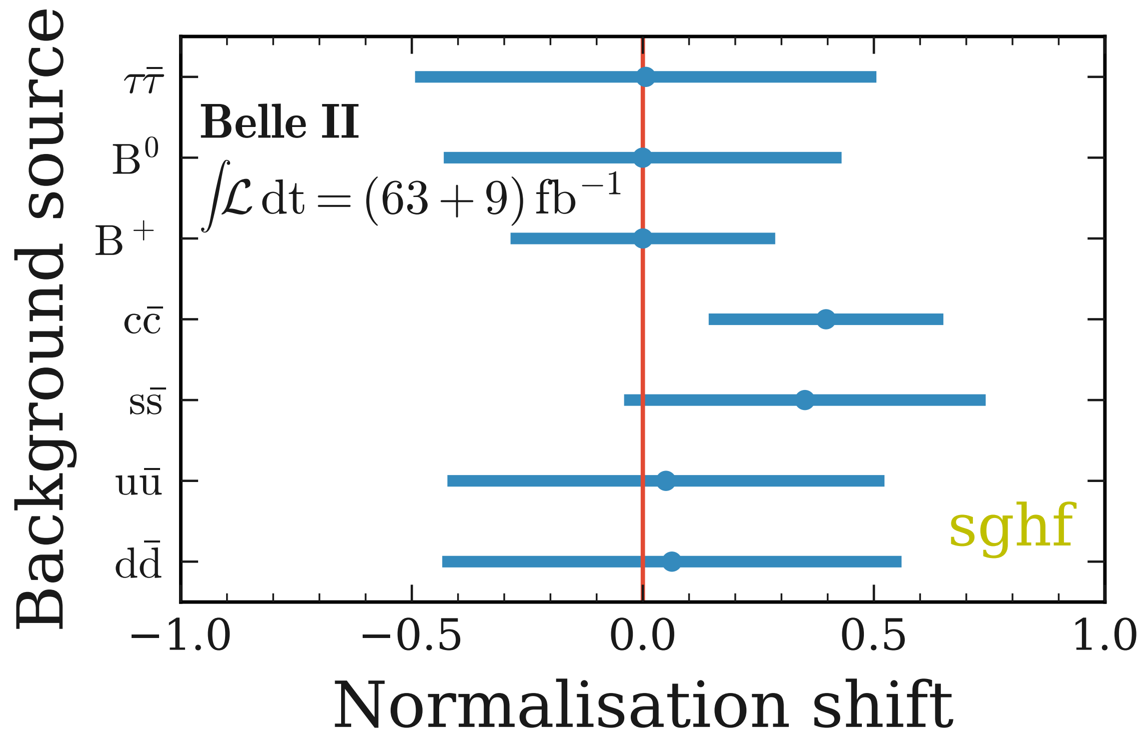
- Profile likelihood scan for the signal strength $\mu$:




**Asymmetric uncertainty on** $\mu$ estimated by fitting the scanned points with an asymmetric parabola $f(x) = (x/\sigma^-)^2$ for $x < 0$ and $f(x) = (x/\sigma^+)^2$ for $x > 0$.

# Fit to the Data

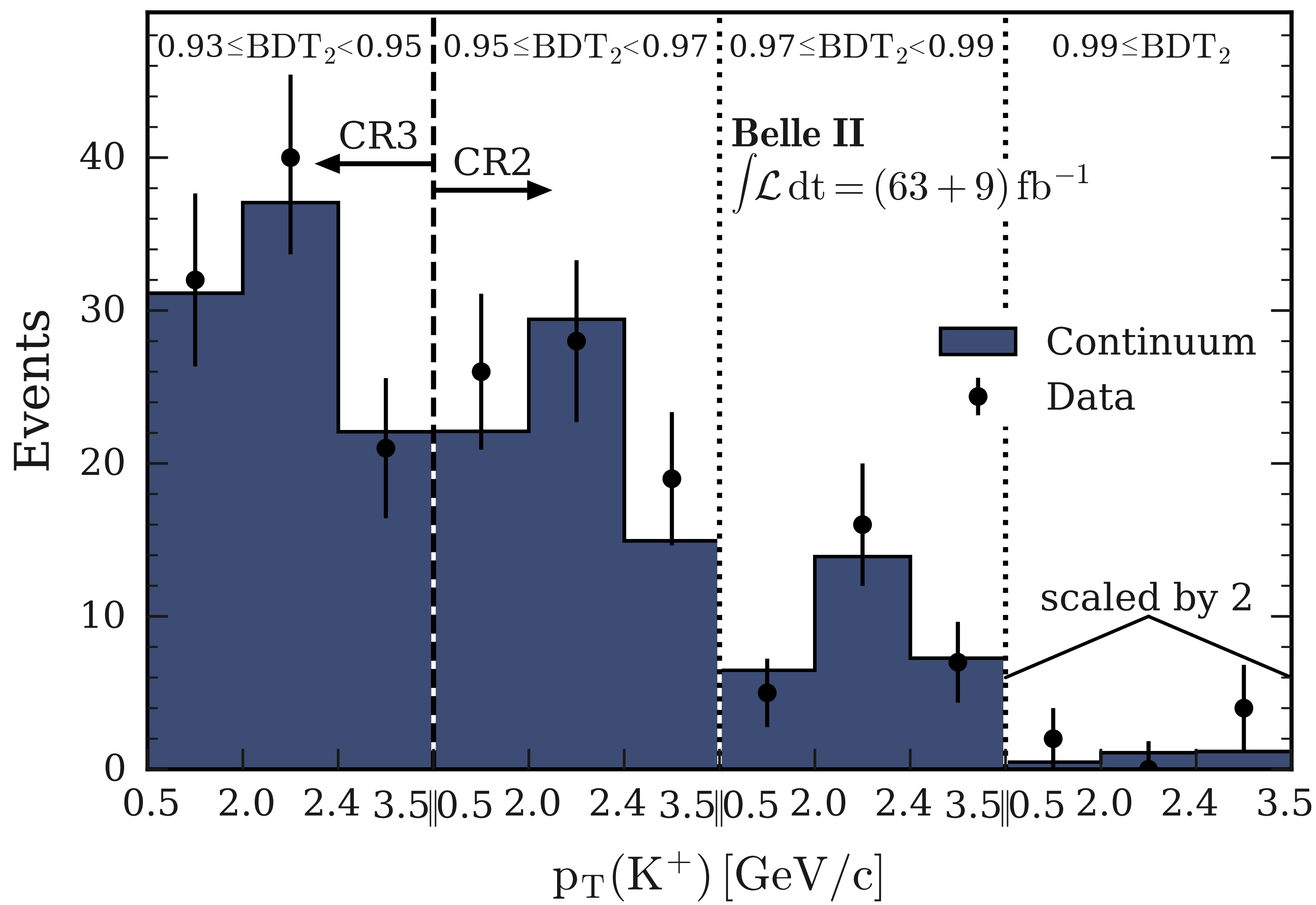
- Post-fit shifts of the bkg's normalisations.




- 50% pre-fit uncertainty attached to each of the bkg's normalisations.
- No post-fit shift wrt to expectations for $B^+B^-$ and $B^0\bar{B}^0$ that are the larger bkg's.
- Post-fit shift of $\sim 1\sigma$ wrt to the expectations for some continuum sources $(c\bar{c}, s\bar{s})$ consistent **with the observed Data-MC normalisation discrepancy.**

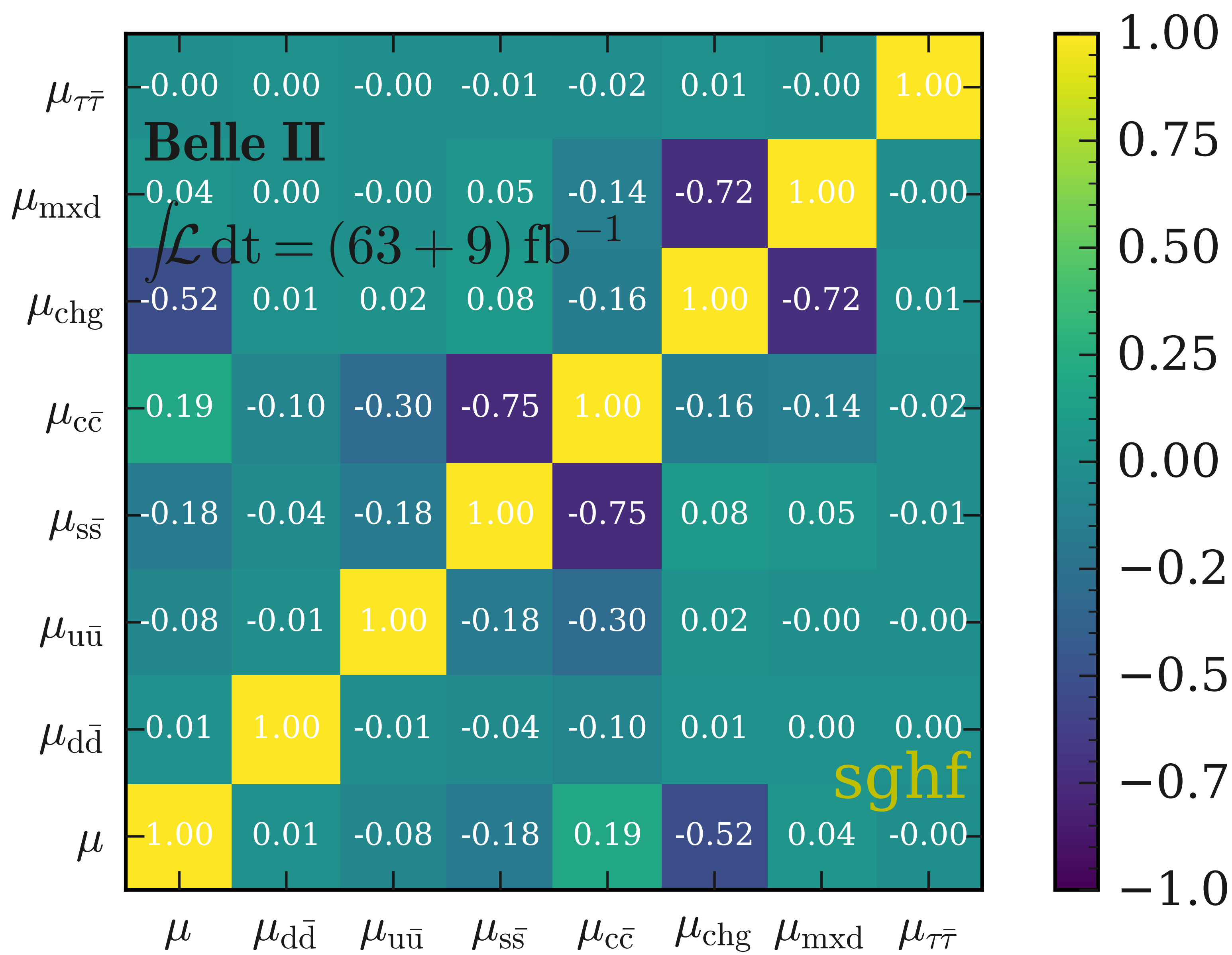
# Fit to the Data

- Post-fit predictions for continuum vs off-resonance data.



# Fit to the Data

- Correlation of post-fit shifts of the bkg's normalisations.



# Limit vs uncertainties