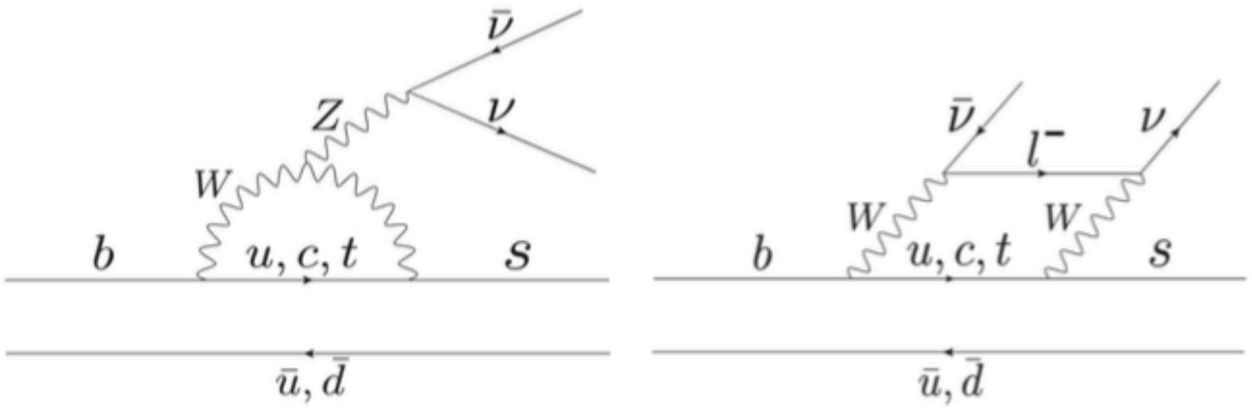
Search for $B^+ \to K^+ \nu \bar{\nu}$ decays with an inclusive tagging method at the Belle II experiment

55th Rencontres de Moriond
Electroweak Interactions & Unified Theories
March 24, 2021

Filippo Dattola on behalf of the Belle II Collaboration

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

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



In the Standard Model:

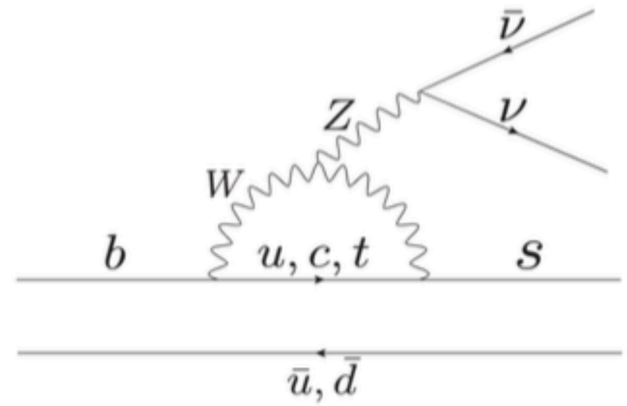
- $b \to s\nu\bar{\nu}$ **flavour-changing neutral-current transition;**
- occurs at the loop level, **suppressed** by the extended GIM mechanism;
- clean theoretical prediction:

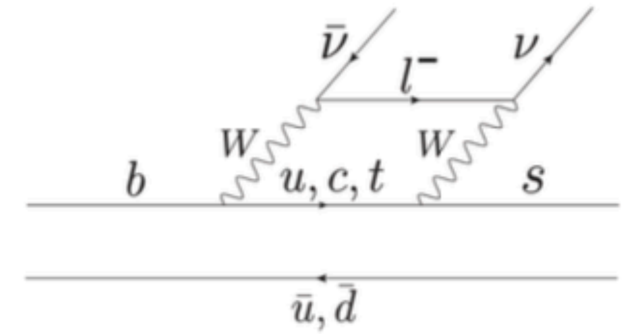
$$\boxed{\mathrm{BR}(\mathrm{B}^+ \to \mathrm{K}^+\nu\bar{\nu})_{\mathrm{SM}} = (4.6 \pm 0.5) \times 10^{-6}}$$

[W. Altmannshofer et al. The Belle II Physics Book. 2018.]

[David M. Straub et al. Rare B decays as tests of the Standard Model]

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





In the Standard Model:

- $b \to s\nu\bar{\nu}$ **flavour-changing neutral-current transition;**
- occurs at the loop level, **suppressed** by the extended GIM mechanism;
- clean theoretical prediction:

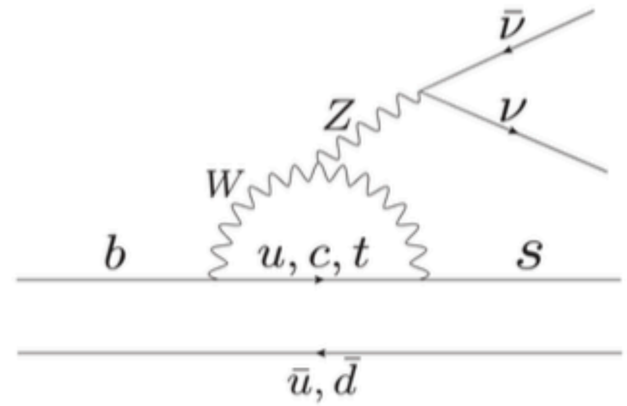
$$\boxed{\mathrm{BR}(\mathrm{B}^+ \to \mathrm{K}^+\nu\bar{\nu})_{\mathrm{SM}} = (4.6 \pm 0.5) \times 10^{-6}}$$

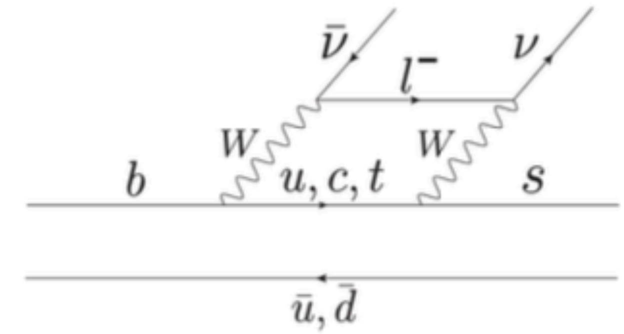
[W. Altmannshofer et al. The Belle II Physics Book. 2018.]

[David M. Straub et al. Rare B decays as tests of the Standard Model]

Optimal measurement to probe the SM and to constrain scenarios beyond it.

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





In the Standard Model:

- $b \to s\nu\bar{\nu}$ **flavour-changing neutral-current transition;**
- occurs at the loop level, **suppressed** by the extended GIM mechanism;
- clean theoretical prediction:

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

[W. Altmannshofer et al. The Belle II Physics Book. 2018.]

[David M. Straub et al. Rare B decays as tests of the Standard Model]

Optimal measurement to probe the SM and to constrain scenarios beyond it.

A challenging measurement:

- decay with **2 neutrinos** in the final state leaving **no signature in the detector;**
- **can be measured at *B* factories** because of the clean event environment and the well defined initial state.

SuperKEKB

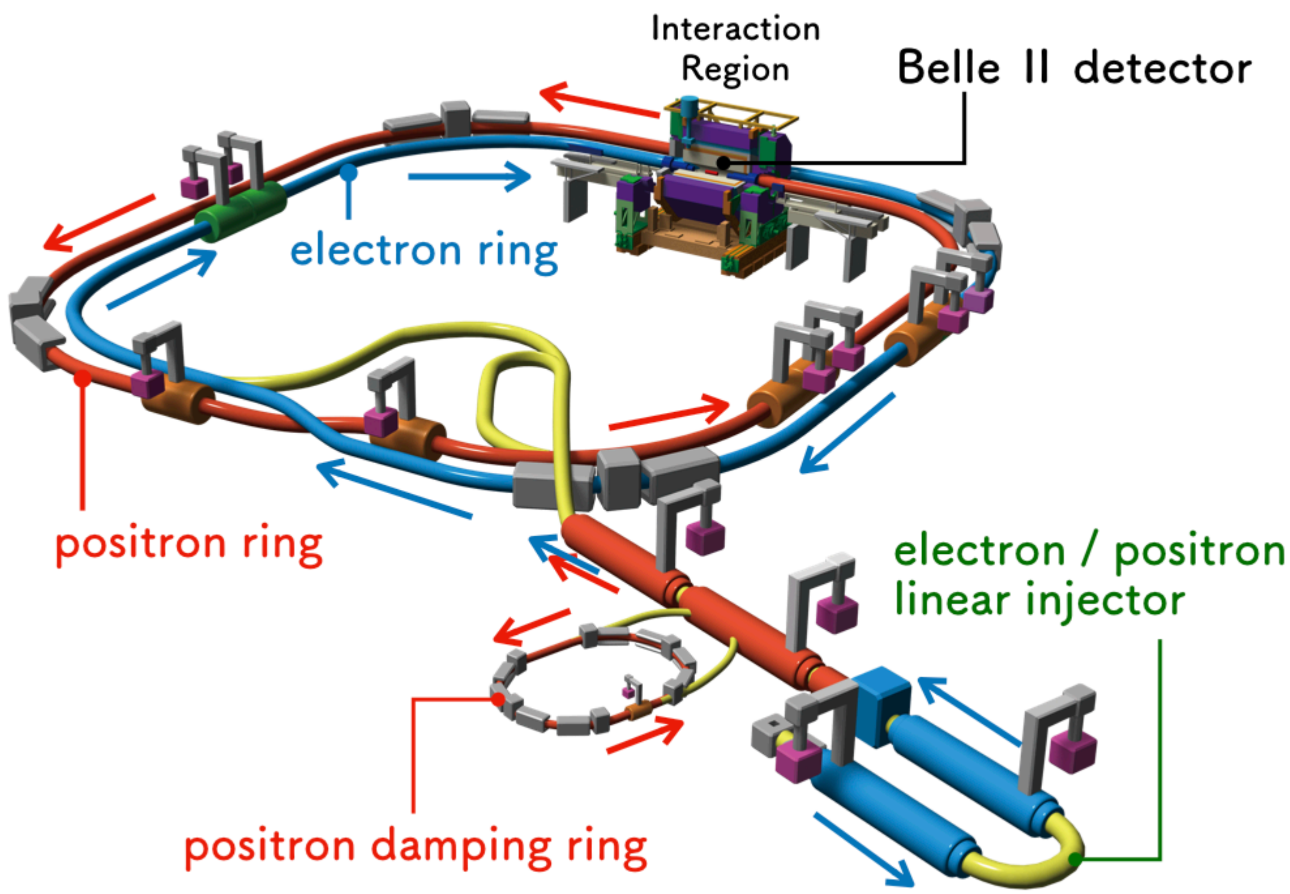
SuperKEKB

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

- Second generation B factory based on the **nanobeam scheme:** major upgrade of its predecessor KEKB.

- **World highest instantaneous luminosity:** 2.4×10^{34} cm^{-2} s^{-1} recorded in June 2020.

- Total integrated luminosity up to now (2021) ~ 90 fb^{-1}.

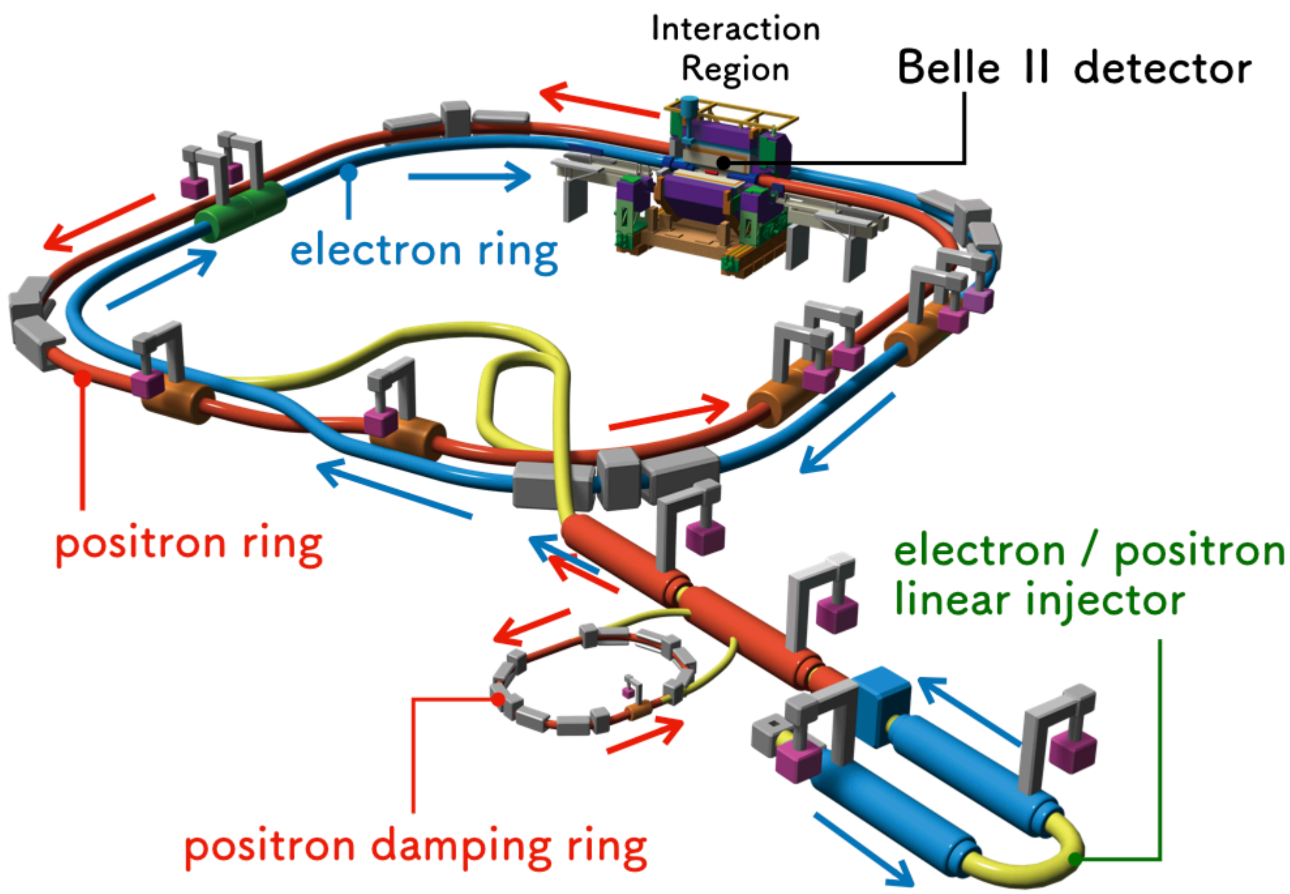


SuperKEKB

- Asymmetric-energy e^+e^- collider operating at $\sqrt{s} = 10.58$ GeV $\rightarrow \Upsilon(4S)$ resonance.
- Second generation B factory based on the **nanobeam scheme:** major upgrade of its predecessor KEKB.
- **World highest instantaneous luminosity:** 2.4×10^{34} cm^{-2} s^{-1} recorded in June 2020.
- Total integrated luminosity up to now (2021) ~ 90 fb^{-1}.
- **<u>For this study:</u>**

  - 63 fb^{-1} **of data collected at** $\sqrt{s} \rightarrow \Upsilon(4S)$ **resonance** ~ 68 **million** $B\bar{B}$ **pairs.**

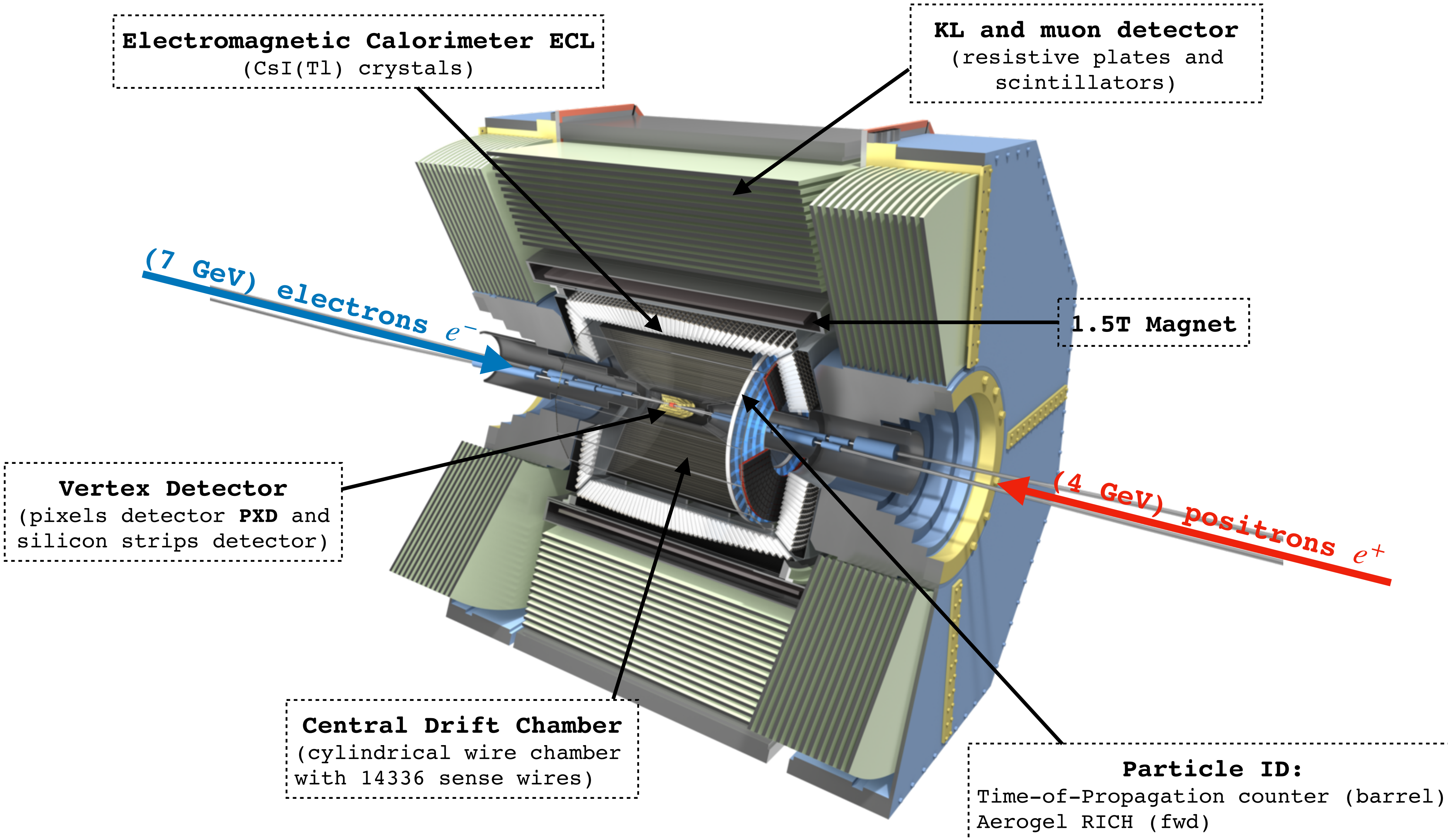
  - 9 fb^{-1} **of (off-resonance) data collected 60 MeV below the** $\Upsilon(4S)$ **resonance for background studies.**



The Belle II detector

The Belle II detector

New detector with respect to the predecessor Belle.



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_{tag}

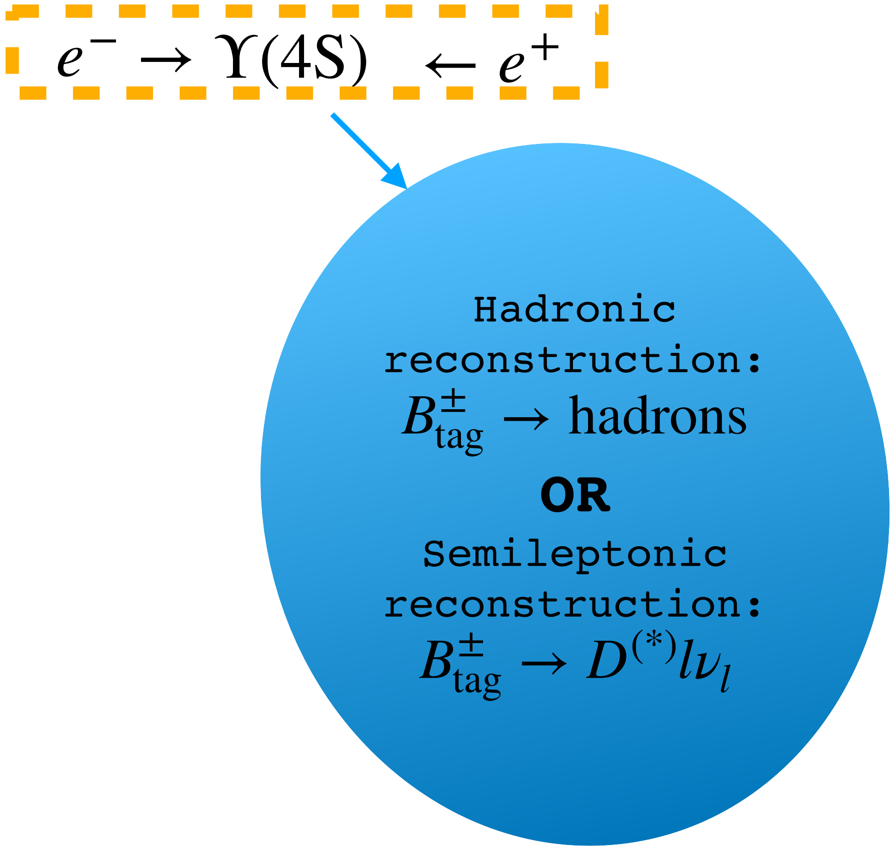
Previous searches for $B^+ \to K^+\nu\bar{\nu}$

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

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

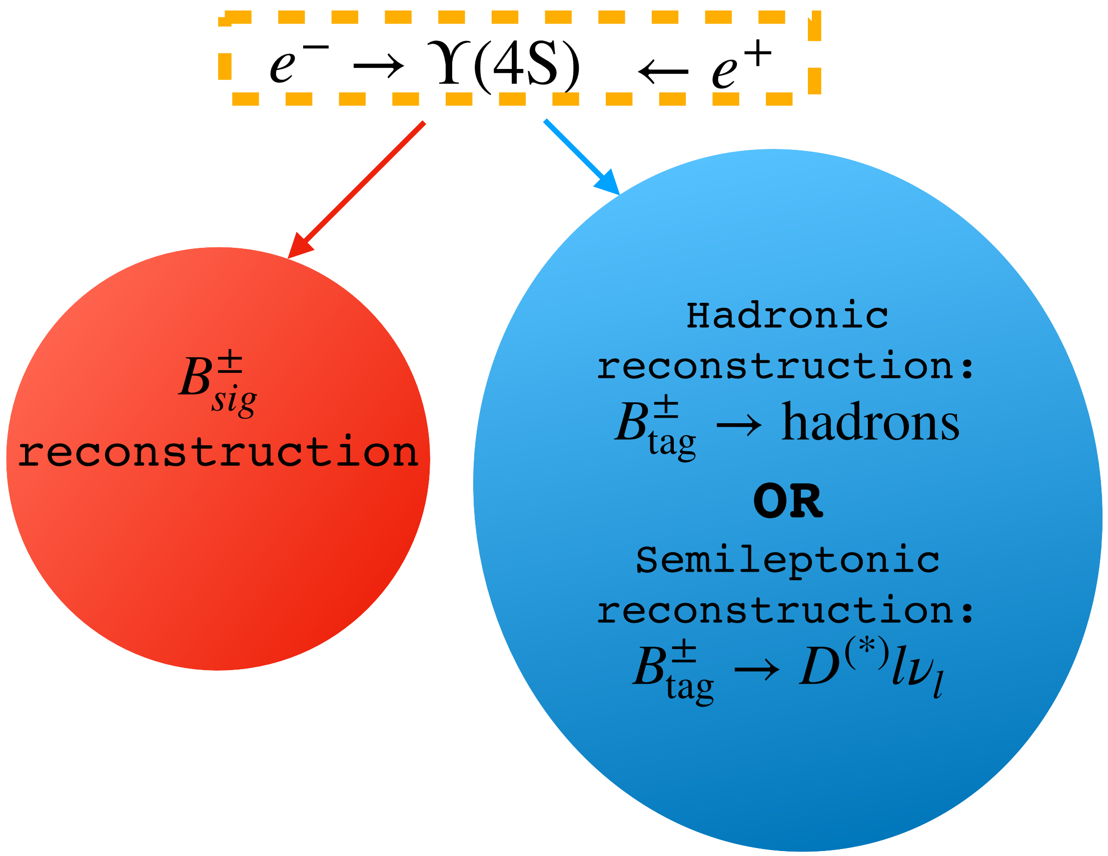
Previous searches for $B^+ \to K^+ \nu \bar{\nu}$

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



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

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



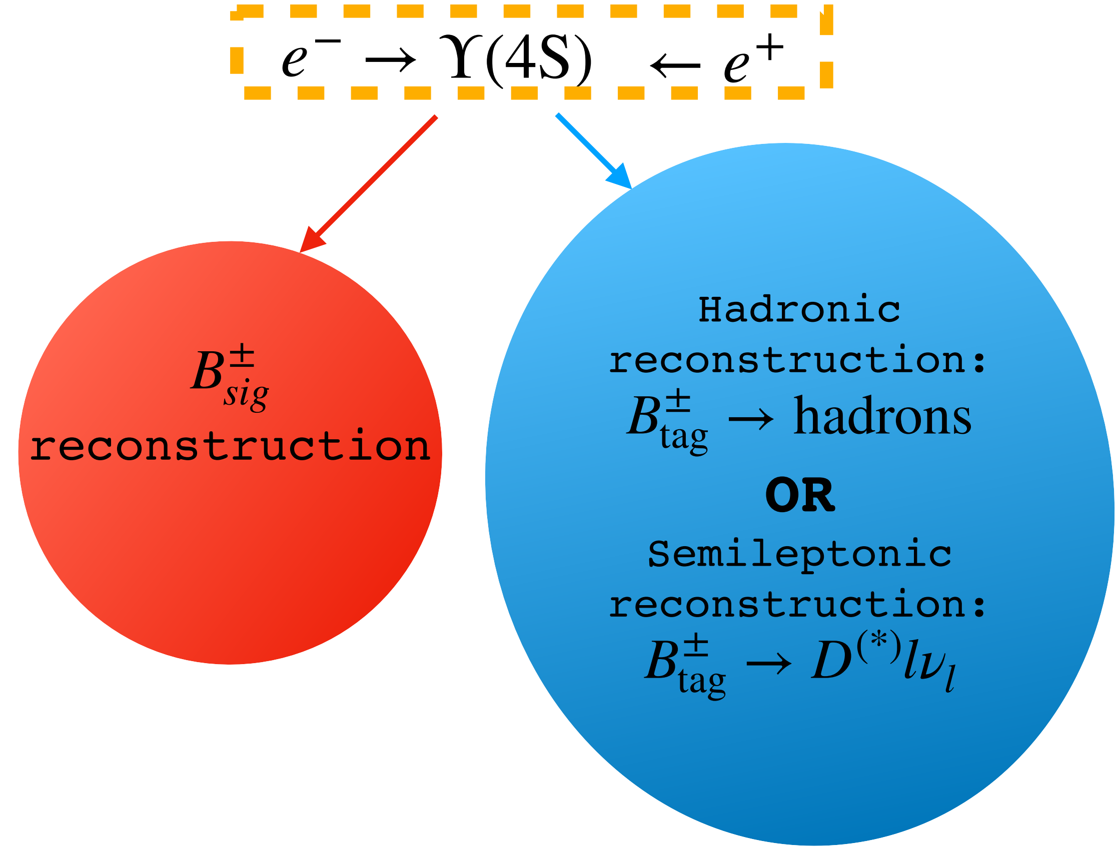
Previous searches for $B^+ \to K^+\nu\bar{\nu}$

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

$\downarrow$

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

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



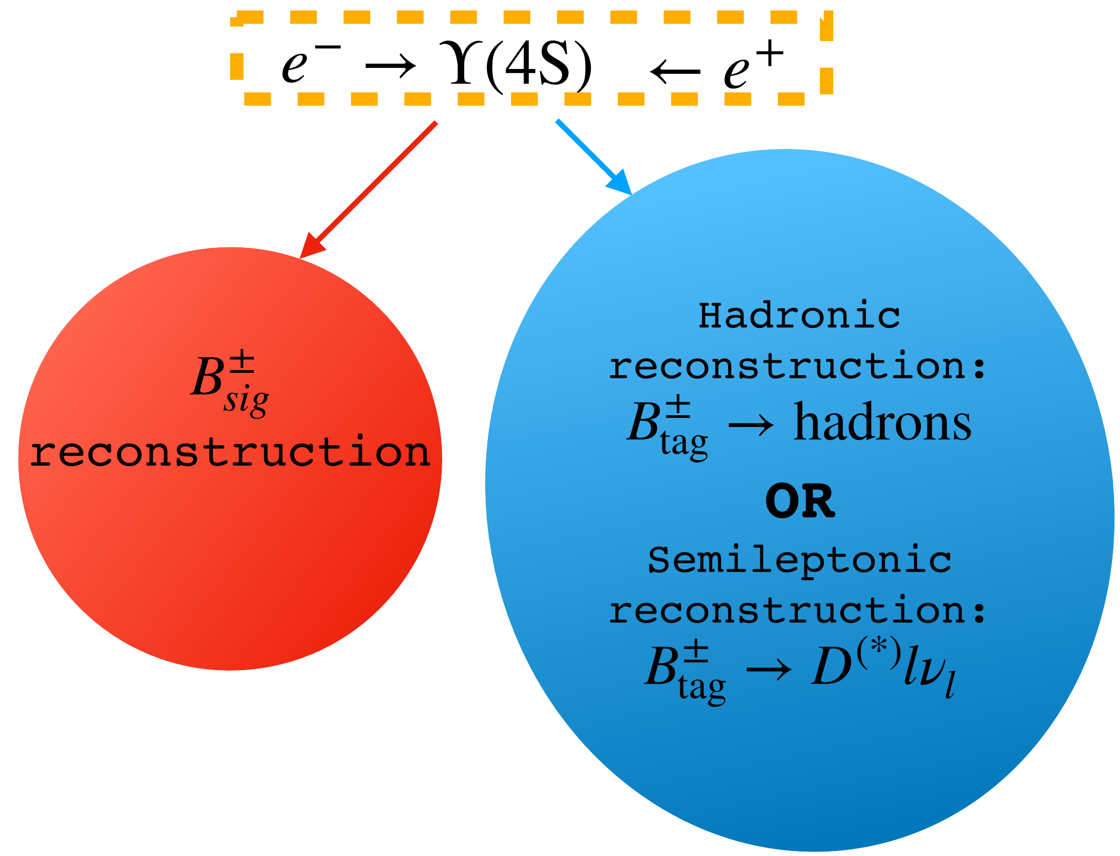
Previous searches for $B^+ \to K^+ \nu \bar{\nu}$

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

↓

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

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



Only 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 tag**	429
Belle	**2013**	$< 5.5 \times 10^{-5}$ **[Phys.Rev.D87,111103(R)]**	**Had tag**	711
Belle	**2017**	$< 1.9 \times 10^{-5}$ **[Phys.Rev.D96,091101(R)]**	**SL tag**	711

The inclusive tagging

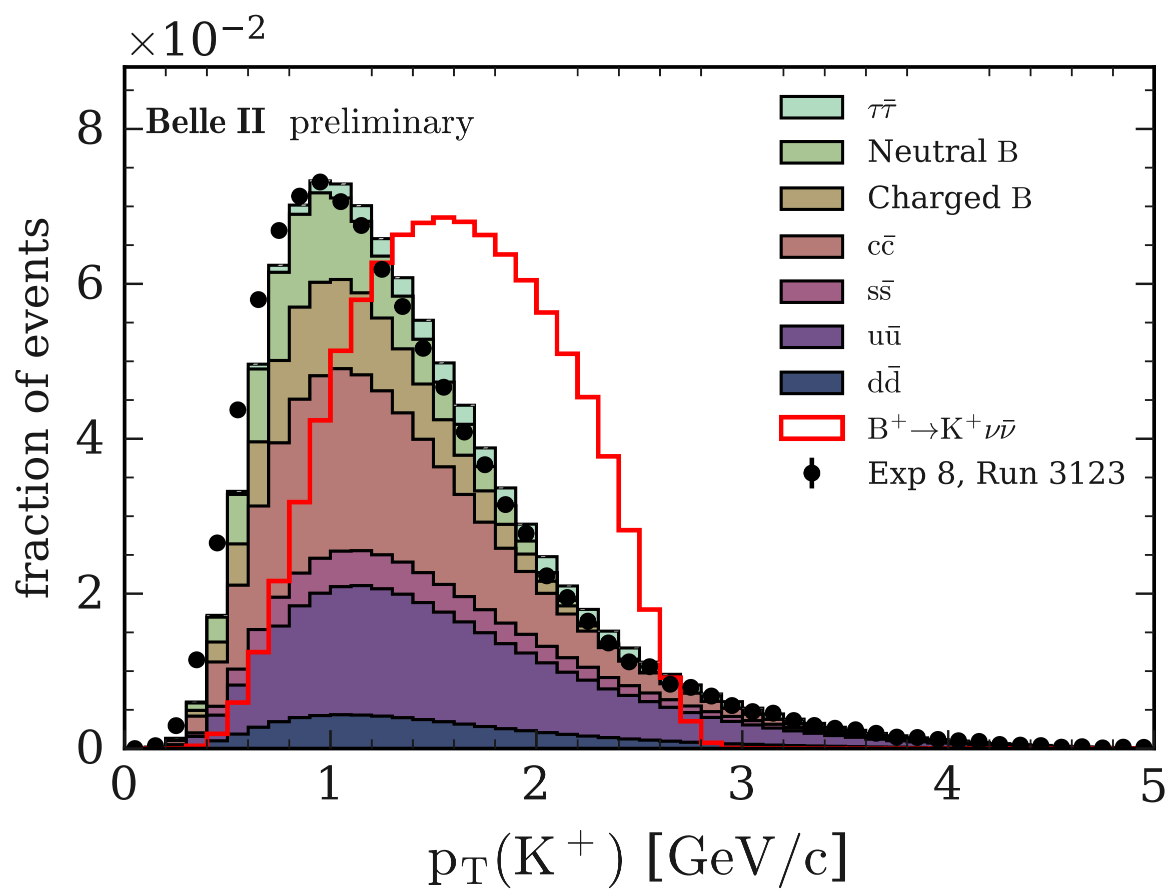
The inclusive tagging

The idea

The inclusive tagging

The idea

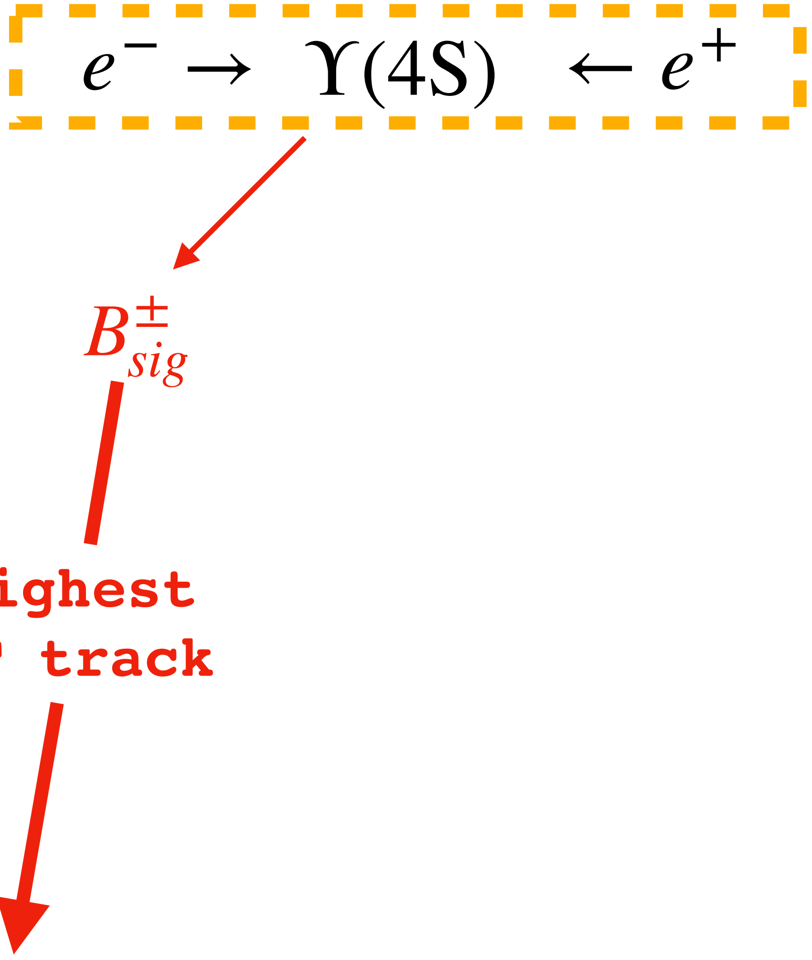
- **Signal reconstructed as the highest p_{T} track with at least 1 PXD hit** (correct match $\sim 80\,\%$)

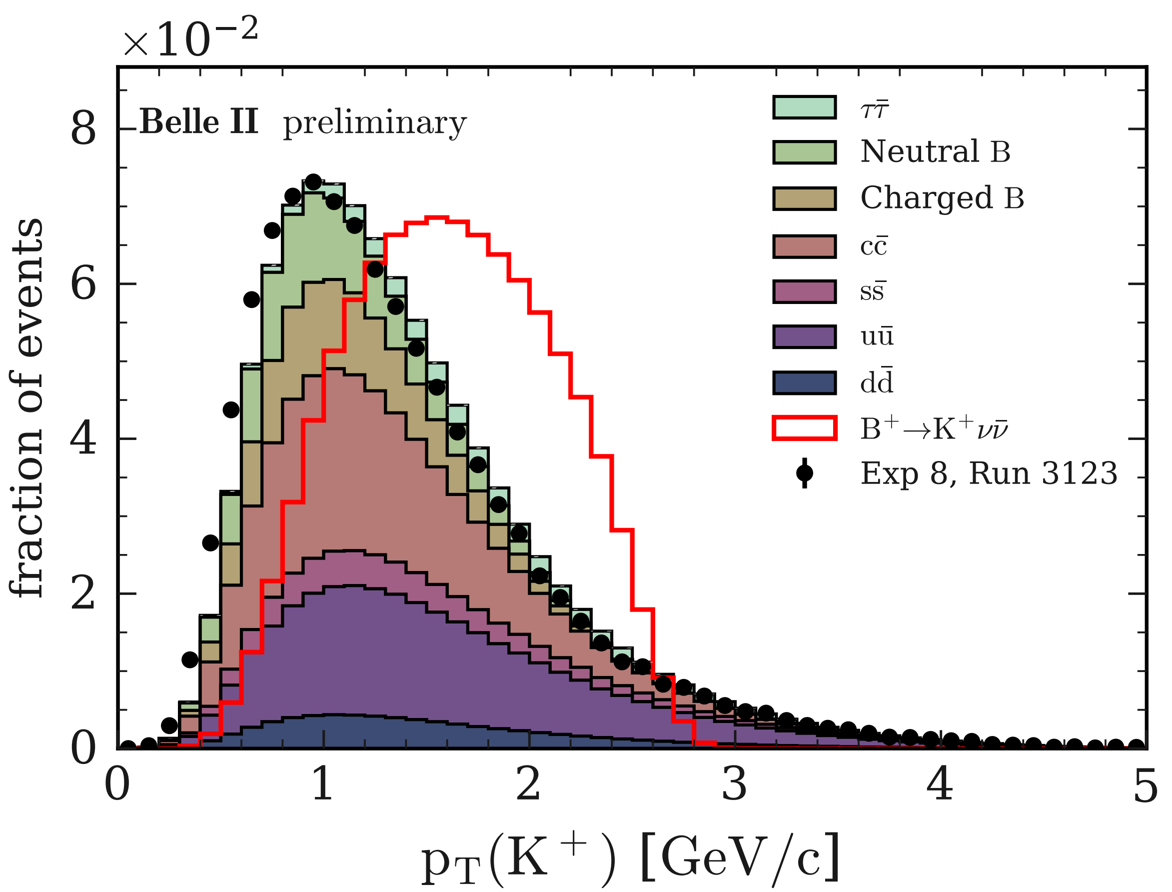


The inclusive tagging

The idea

- **Signal reconstructed as the highest p_T track with at least 1 PXD hit** (correct match $\sim 80\,\%$)

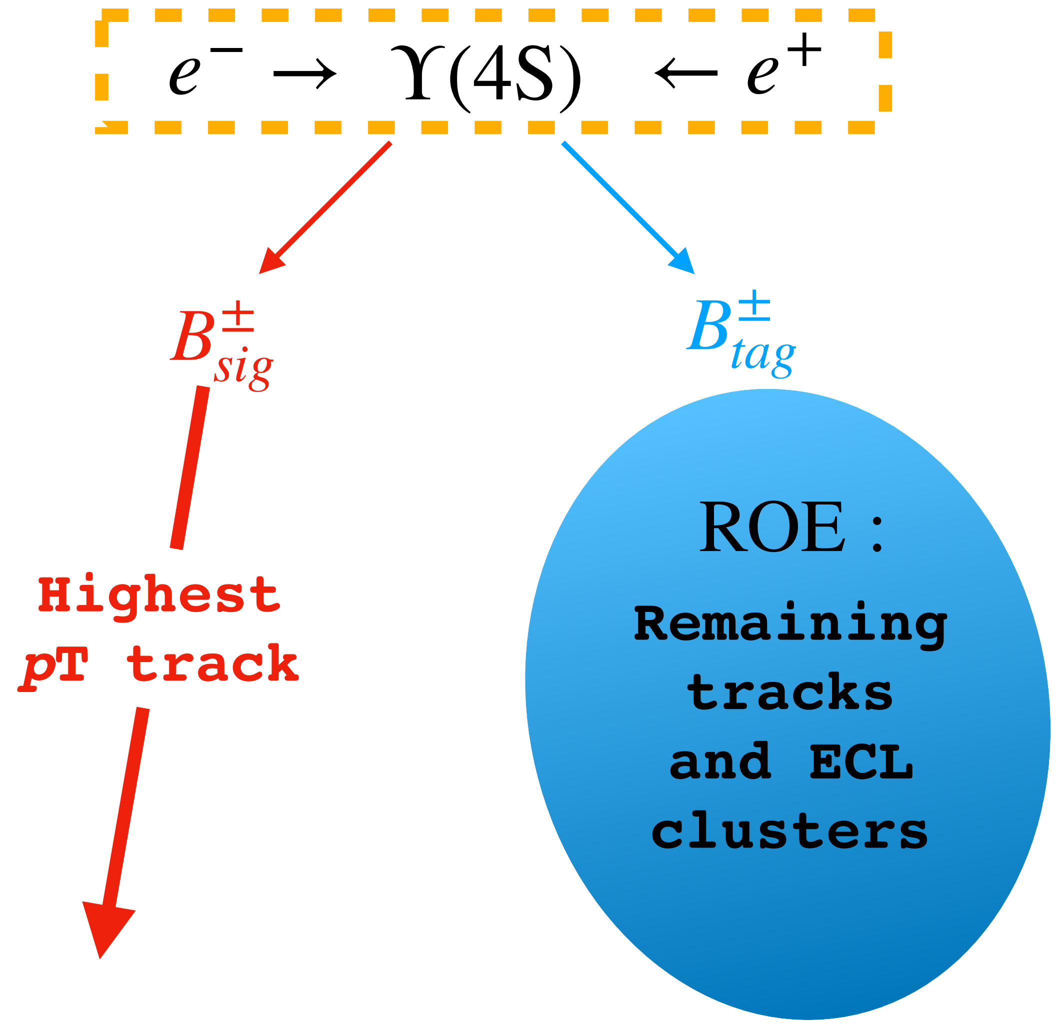


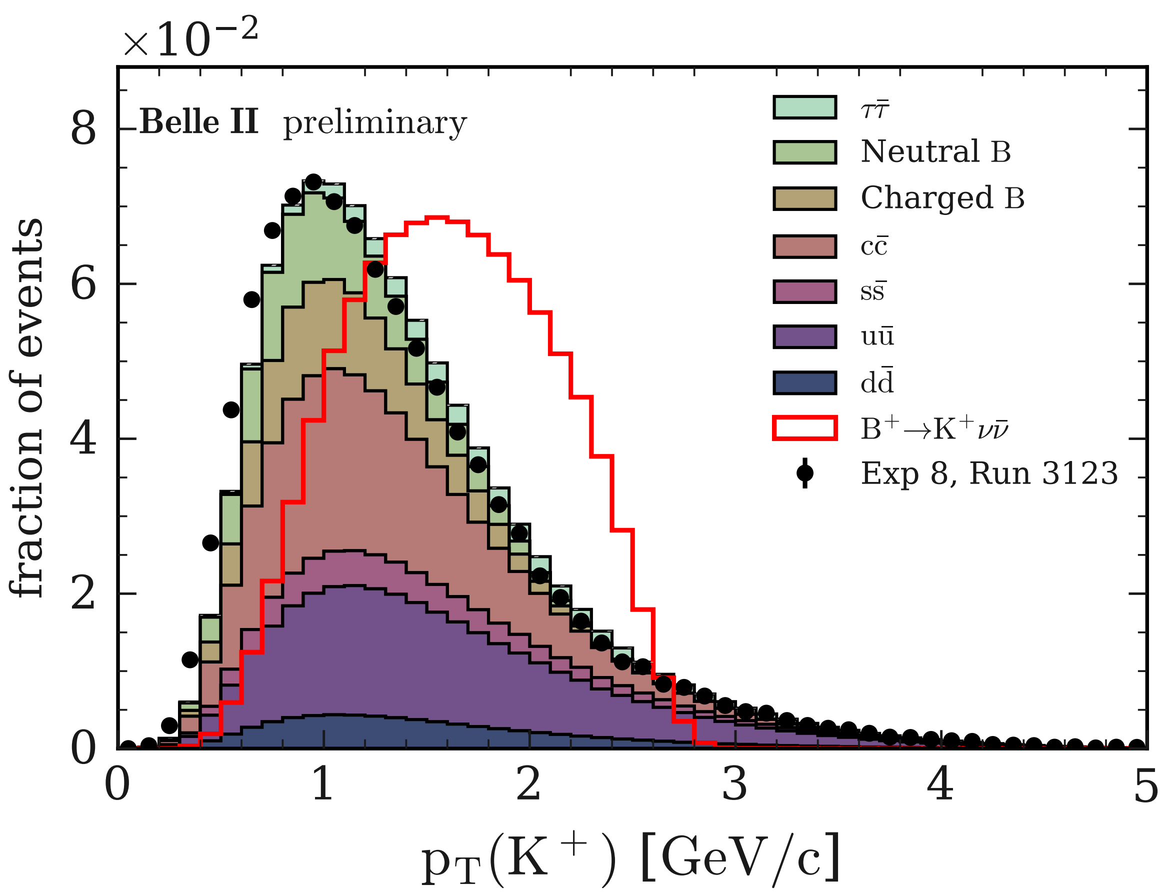


The inclusive tagging

The idea

- **Signal reconstructed as the highest p_T track with at least 1 PXD hit** (correct match $\sim 80\,\%$) **followed by inclusive reconstruction of the *rest of the event* (ROE).**



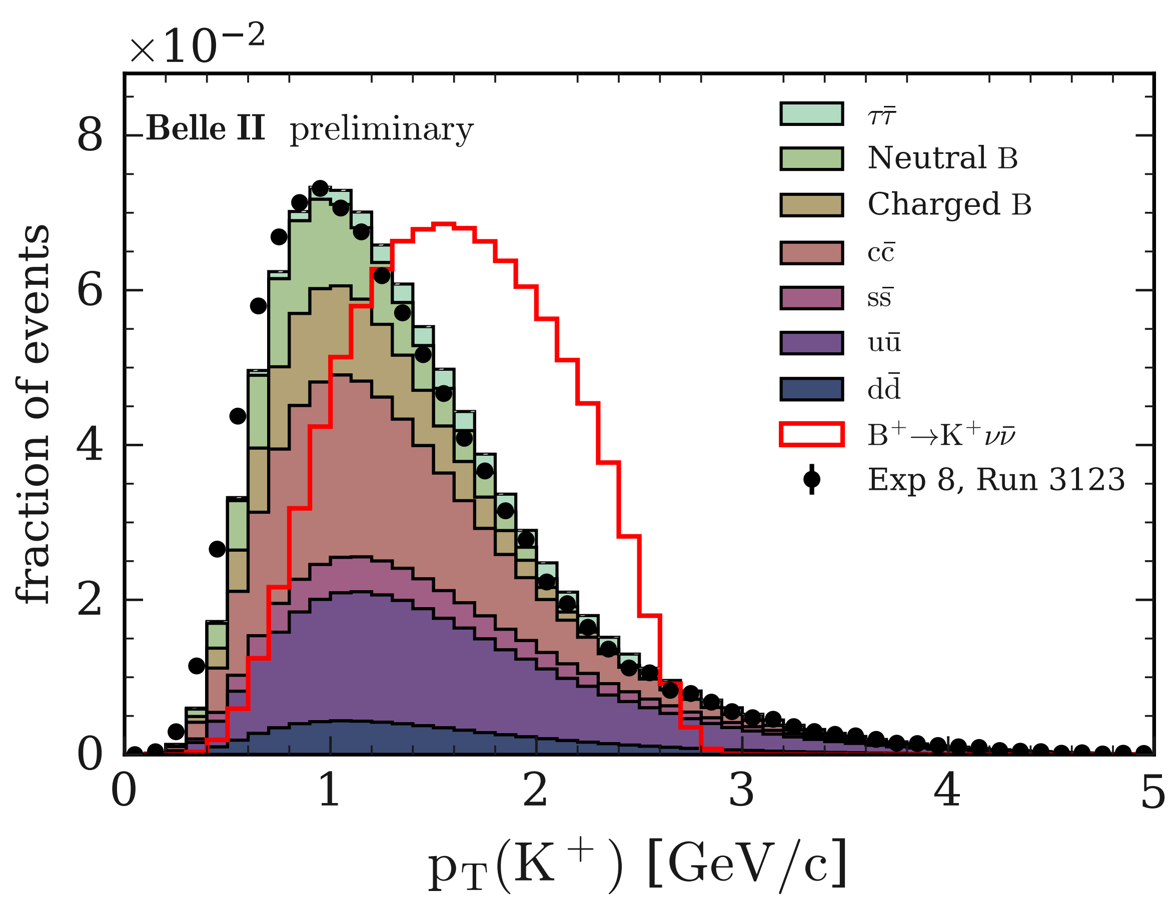


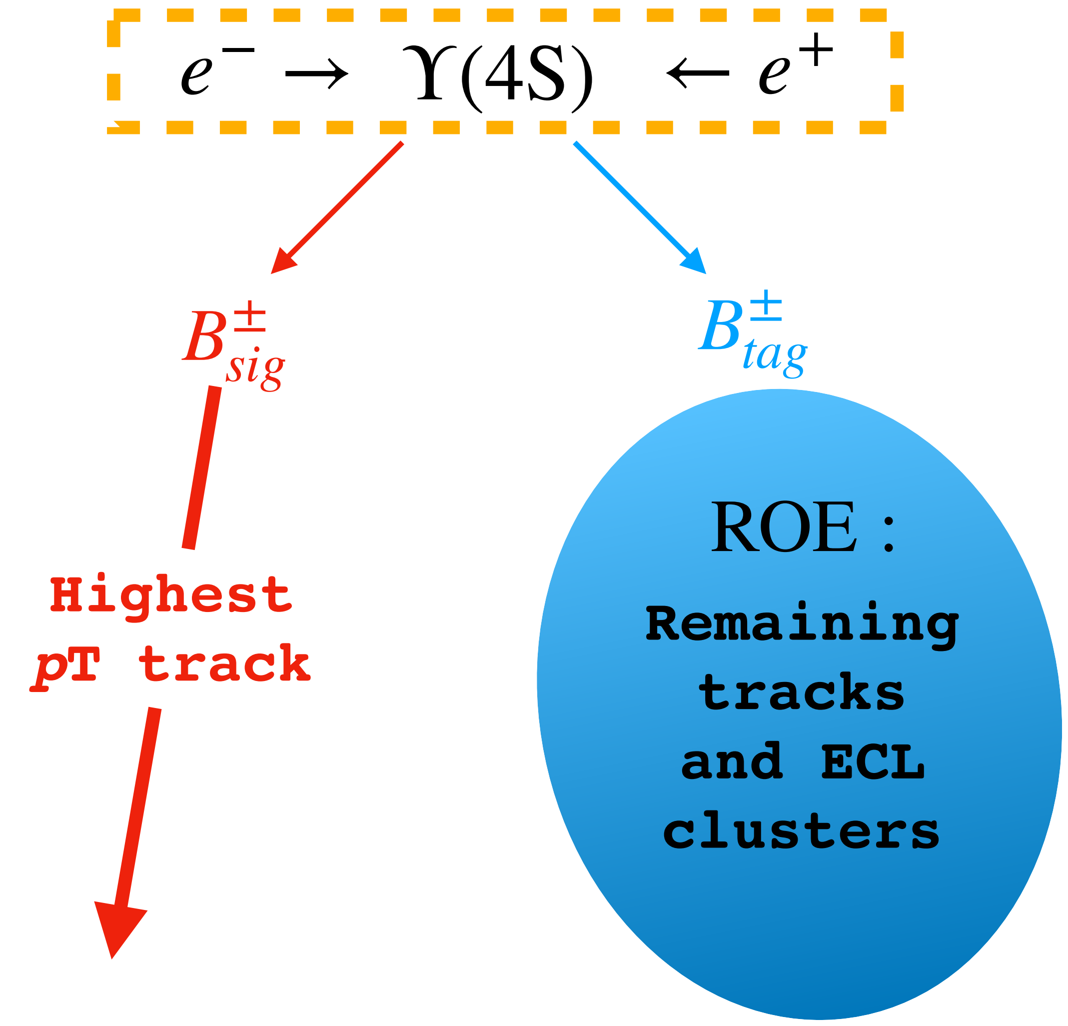
The inclusive tagging

The idea

- **Signal reconstructed as the highest p_T track with at least 1 PXD hit** (correct match $\sim 80\,\%$) **followed by inclusive reconstruction of the *rest of the event* (ROE).**

- **Higher signal efficiency** $\epsilon_{sig} \sim 4\,\%$ but larger background contributions from generic B decays and continuum production ($u\bar{u}$, $d\bar{d}$, $c\bar{c}$, $s\bar{s}$).





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

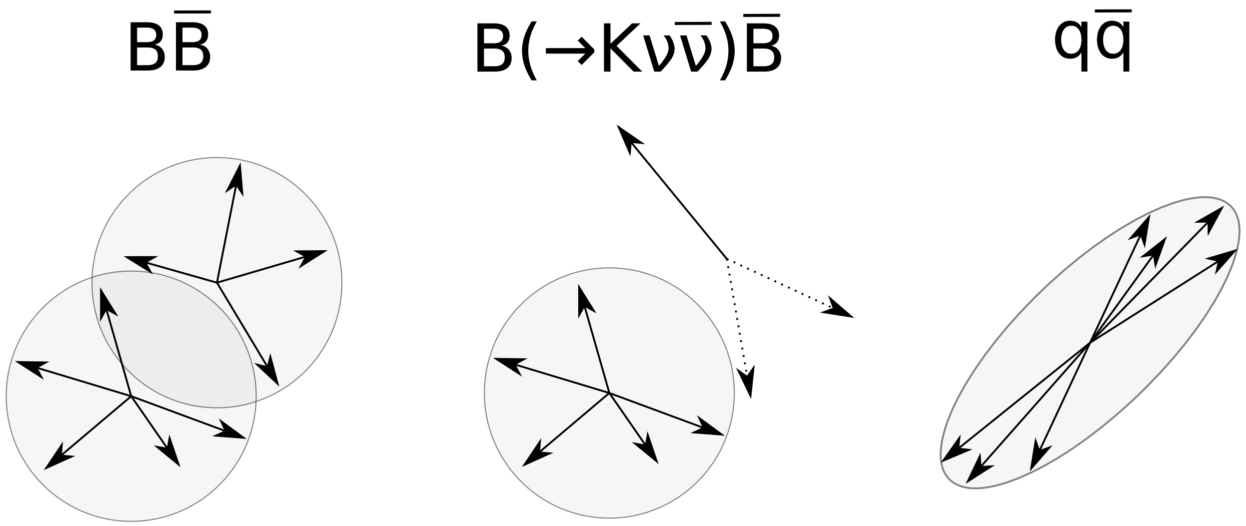
Features of $B^+ \to K^+\nu\bar{\nu}$

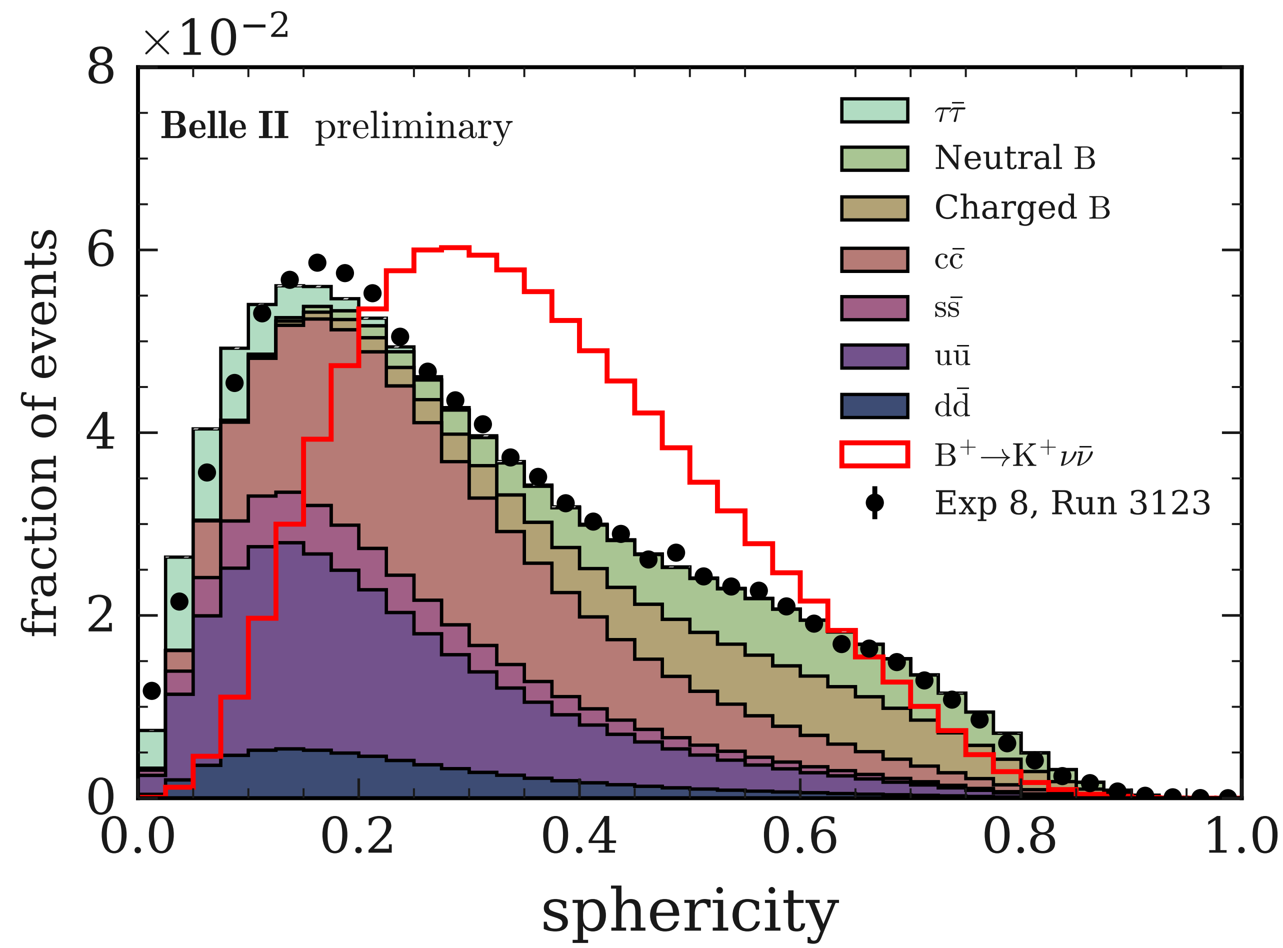
Signal identification exploiting topological features of $B^+ \to K^+\nu\bar{\nu}$

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:**

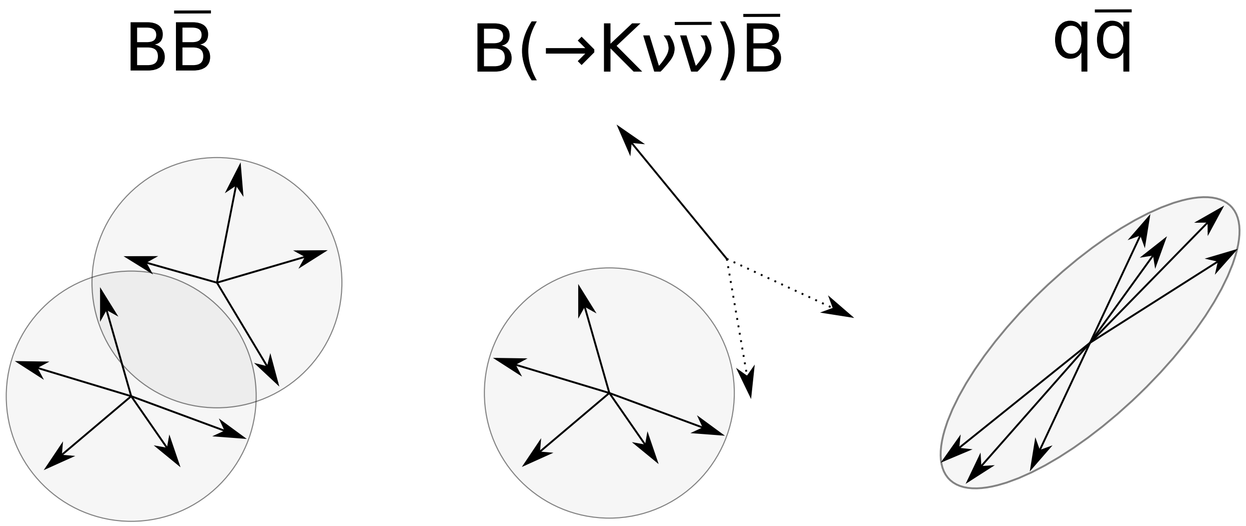


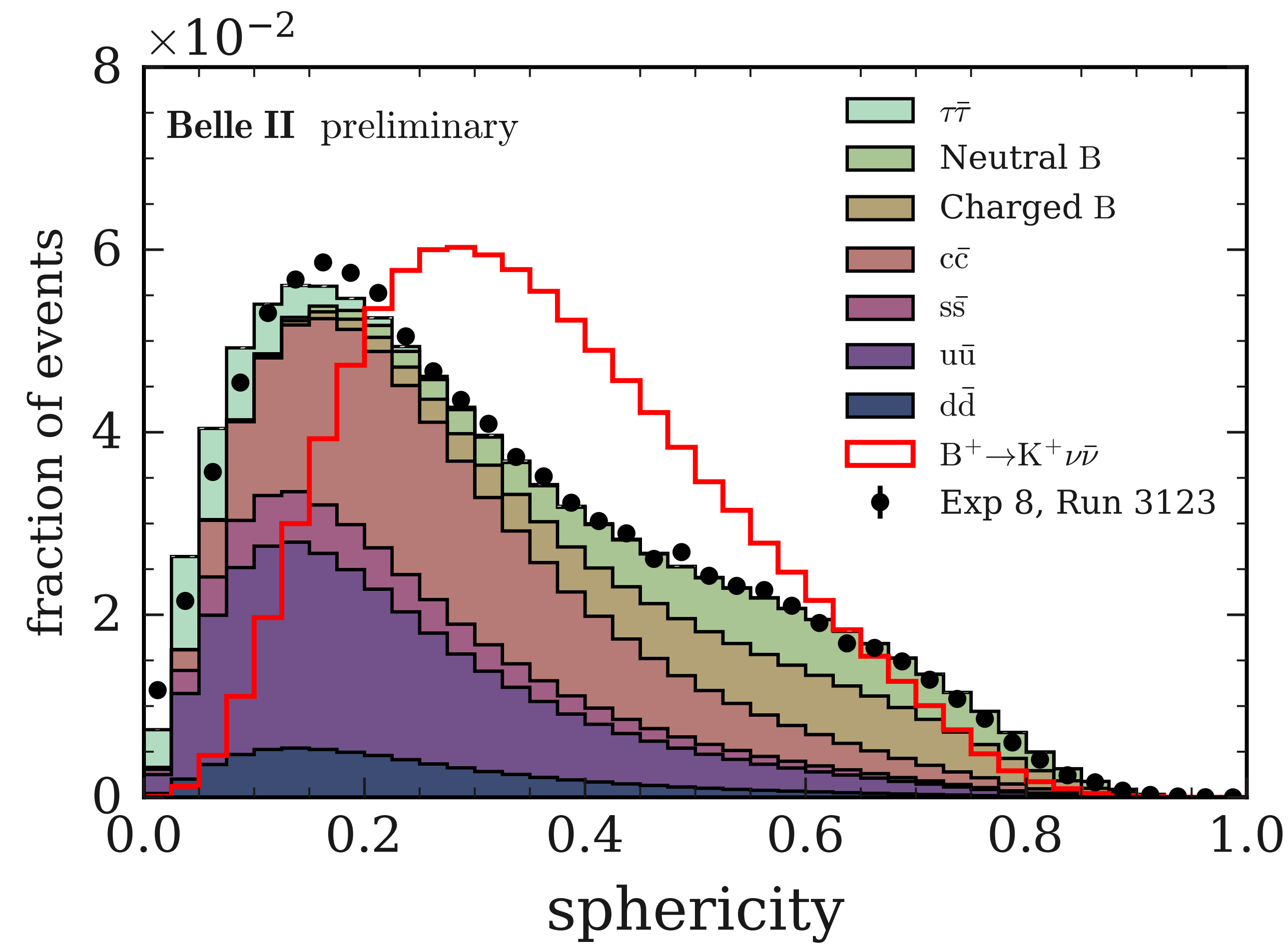


Features of $B^+ \to K^+\nu\bar{\nu}$

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.

Multivariate classification

Multivariate classification

51 well separating variables are used to train 2 consecutive binary classifiers (**FastBDT**) BDT_1 and BDT_2.

Multivariate classification

51 well separating variables are used to train 2 consecutive binary classifiers (**FastBDT**) BDT_1 and BDT_2.

- train 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**) BDT_1 and BDT_2.

- train 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**) BDT_1 and BDT_2.

- train 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 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**) BDT_1 and BDT_2.

- train 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 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

Multivariate classification

51 well separating variables are used to train 2 consecutive binary classifiers (**FastBDT**) BDT_1 and BDT_2.

- train 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 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

Multivariate classification

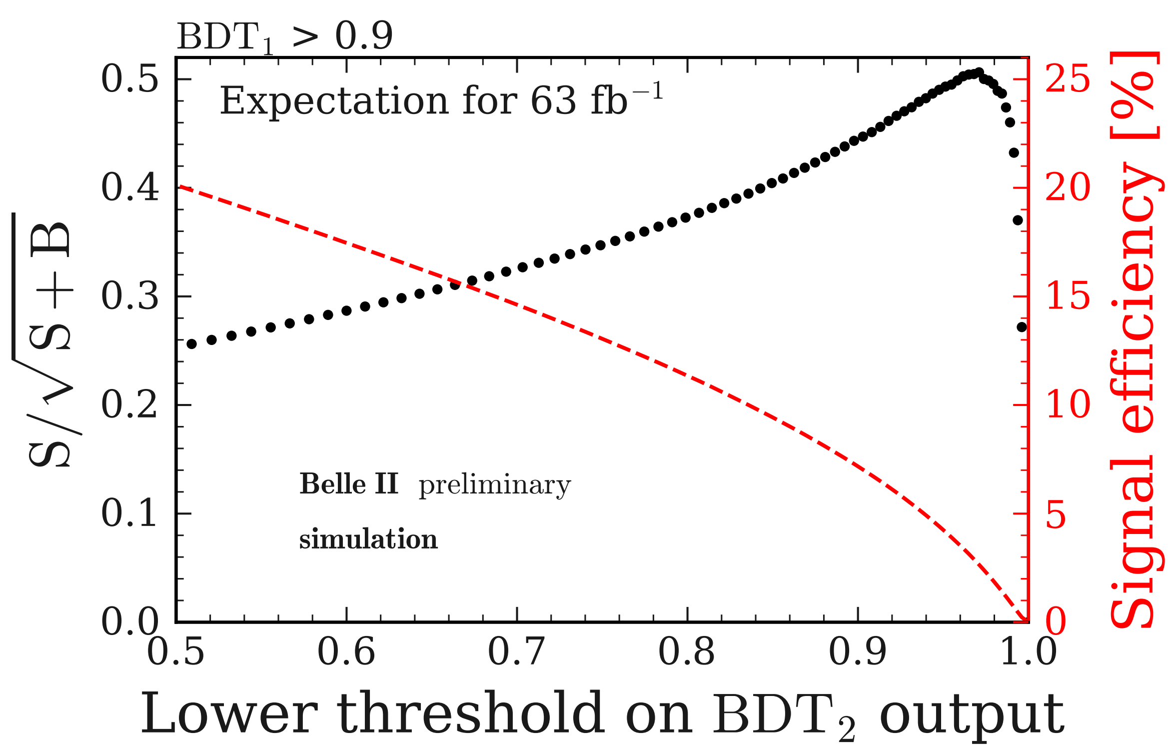
51 well separating variables are used to train 2 consecutive binary classifiers (**FastBDT**) BDT_1 and BDT_2.

- train 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 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



Multivariate classification

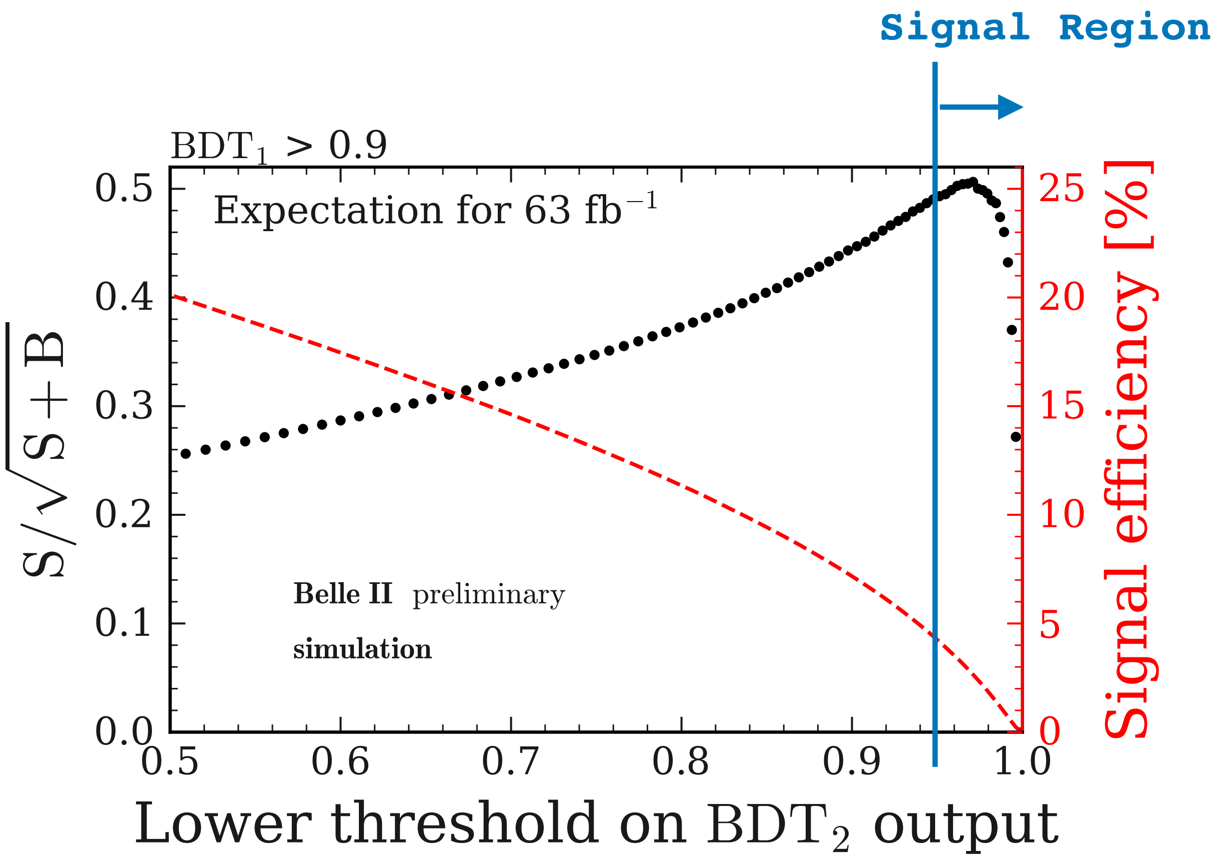
51 well separating variables are used to train 2 consecutive binary classifiers (**FastBDT**) BDT_1 and BDT_2.

- train 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 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:

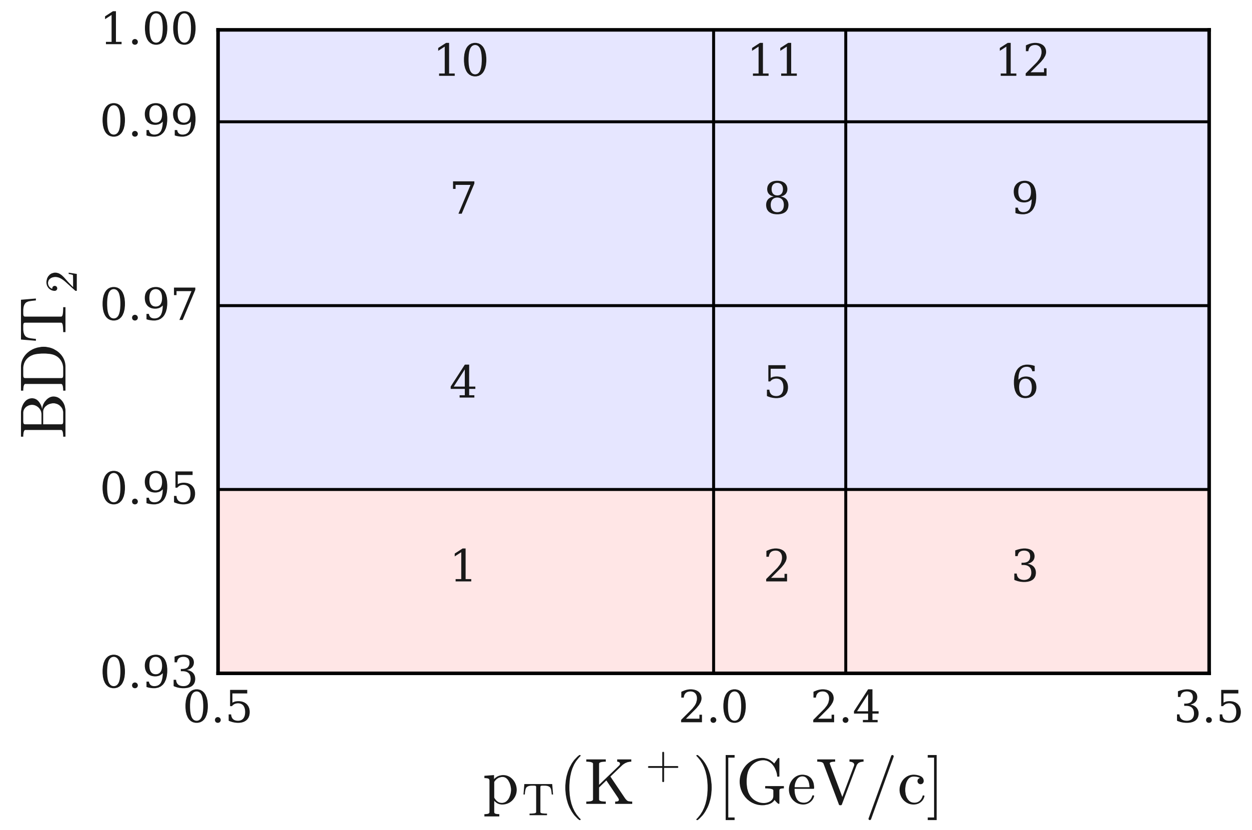
IMPROVEMENT OF SIGNAL SENSITIVITY (+10%, up to ~50%) IN THE HIGH PURITY REGION



Definition of the fit region

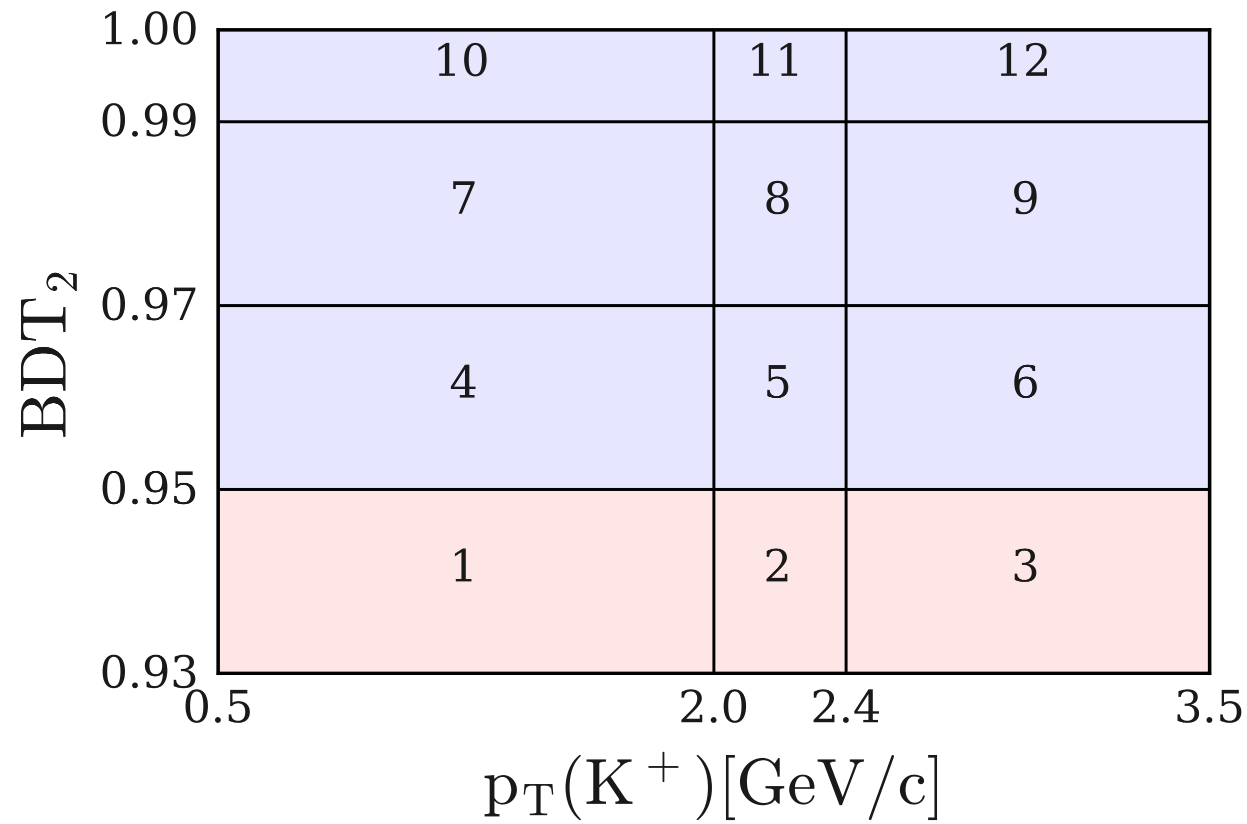
Definition of the fit region

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



Definition of the fit region

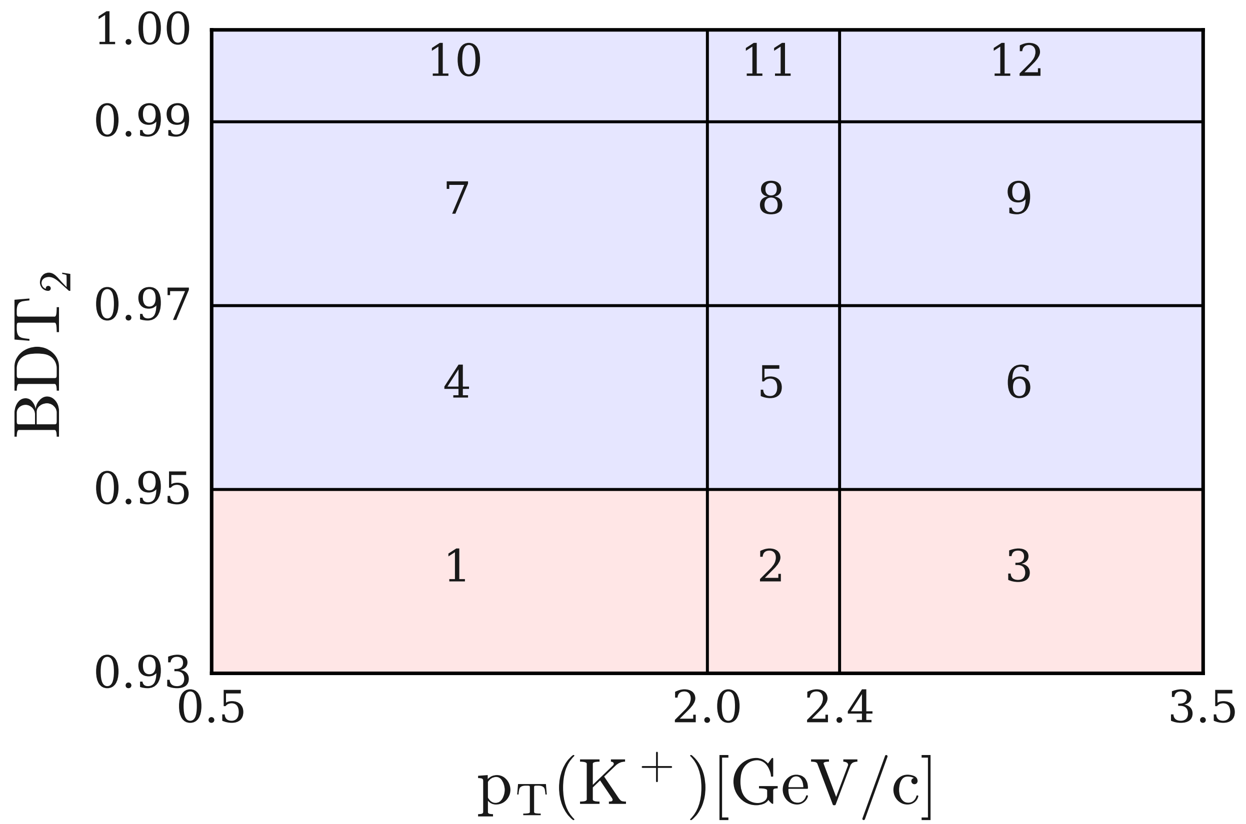
Optimised bin boundaries set up in the $\mathrm{p_T(K^+)} \times \mathrm{BDT_2}$ space:



Bins 4,5,6,7,8,9,10:

Definition of the fit region

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

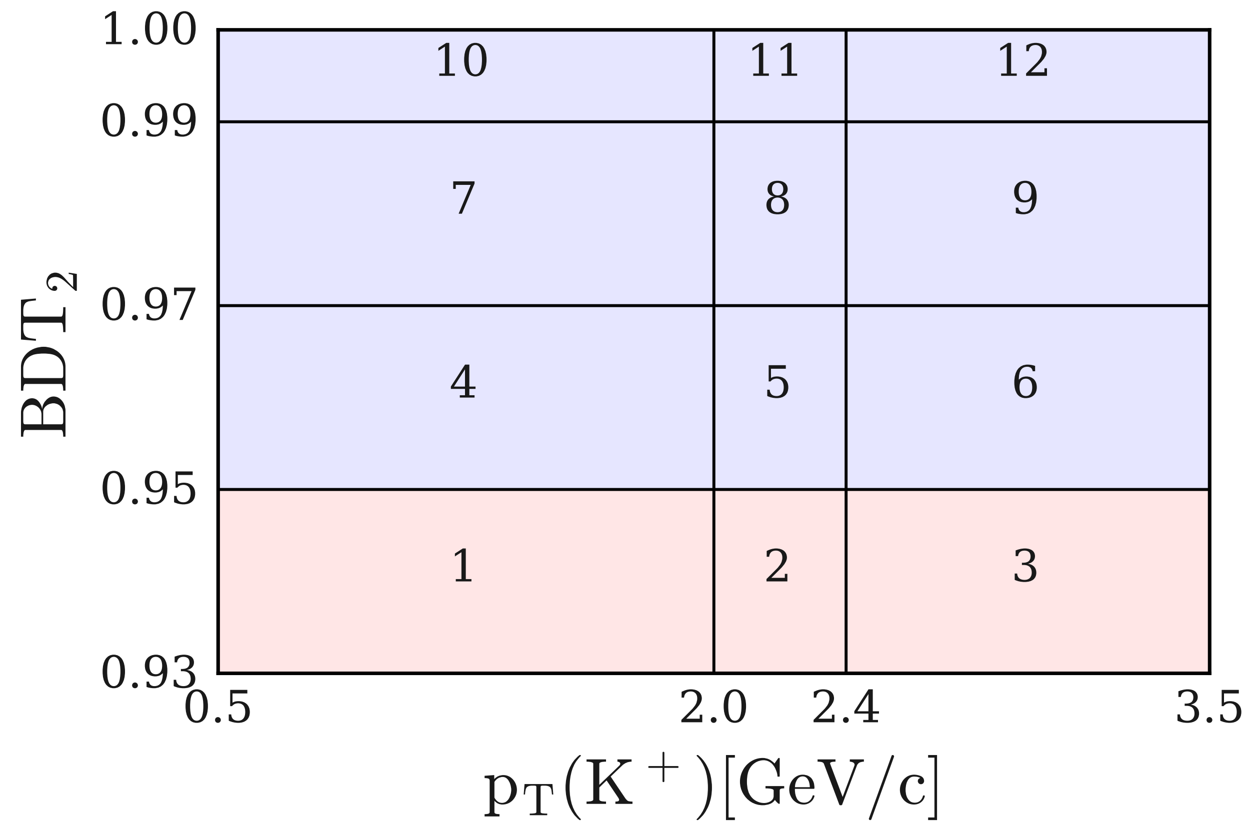


Bins 4,5,6,7,8,9,10:

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

Definition of the fit region

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

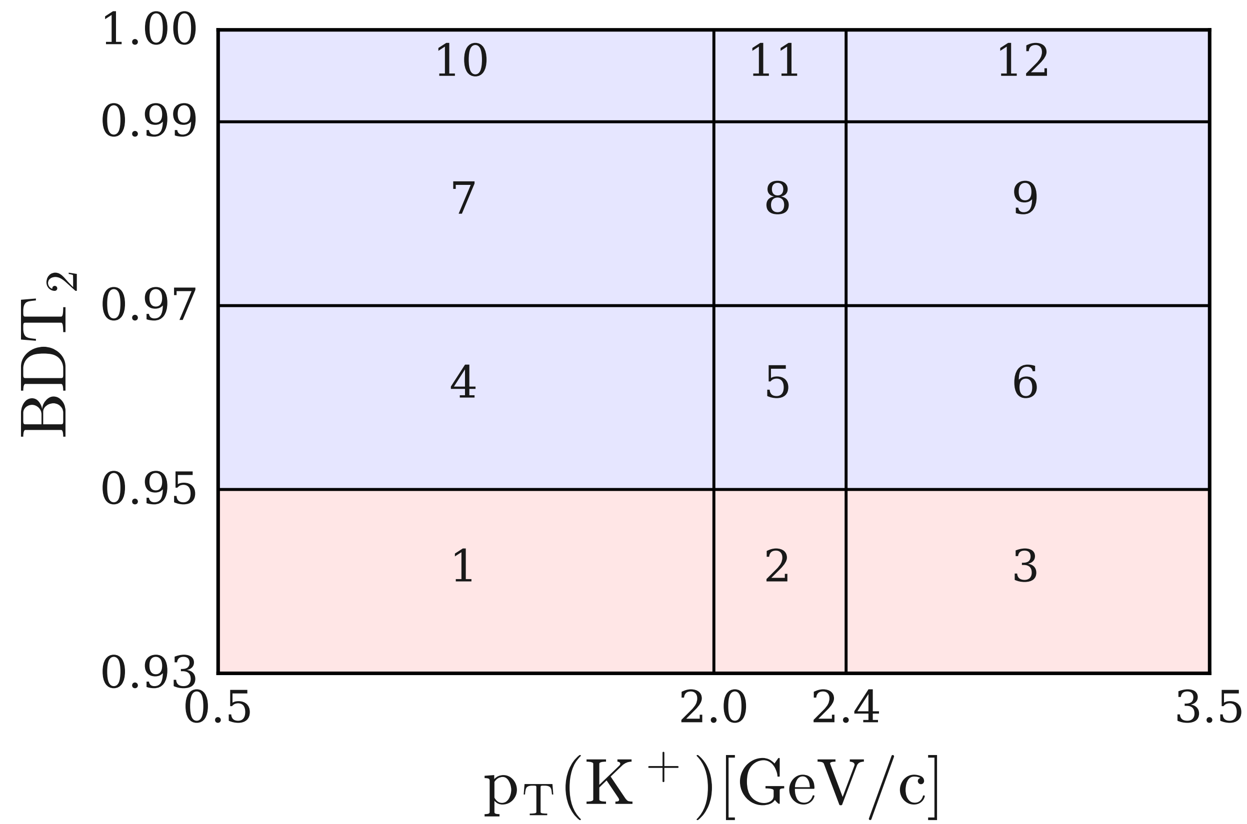


Bins 4,5,6,7,8,9,10:

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

Definition of the fit region

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



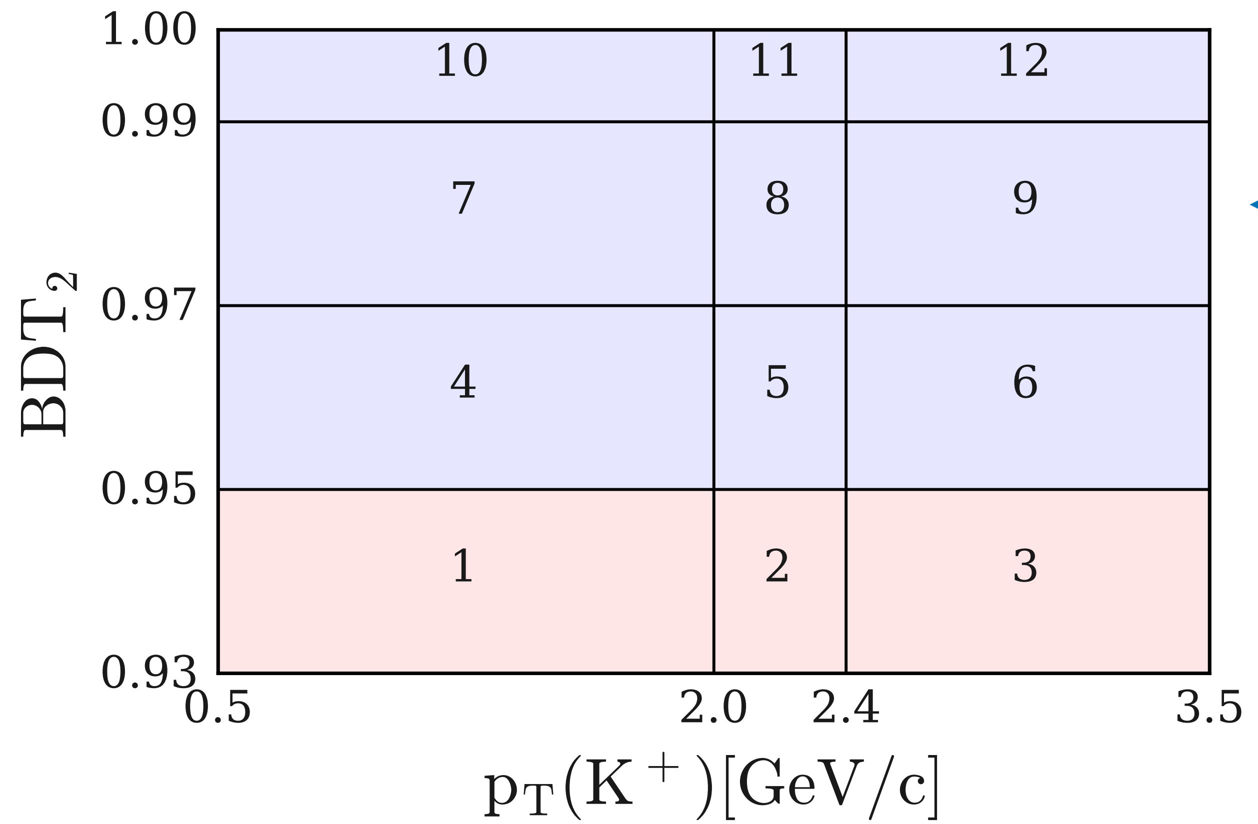
Bins 4,5,6,7,8,9,10:

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

Bins 1,2,3:

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:

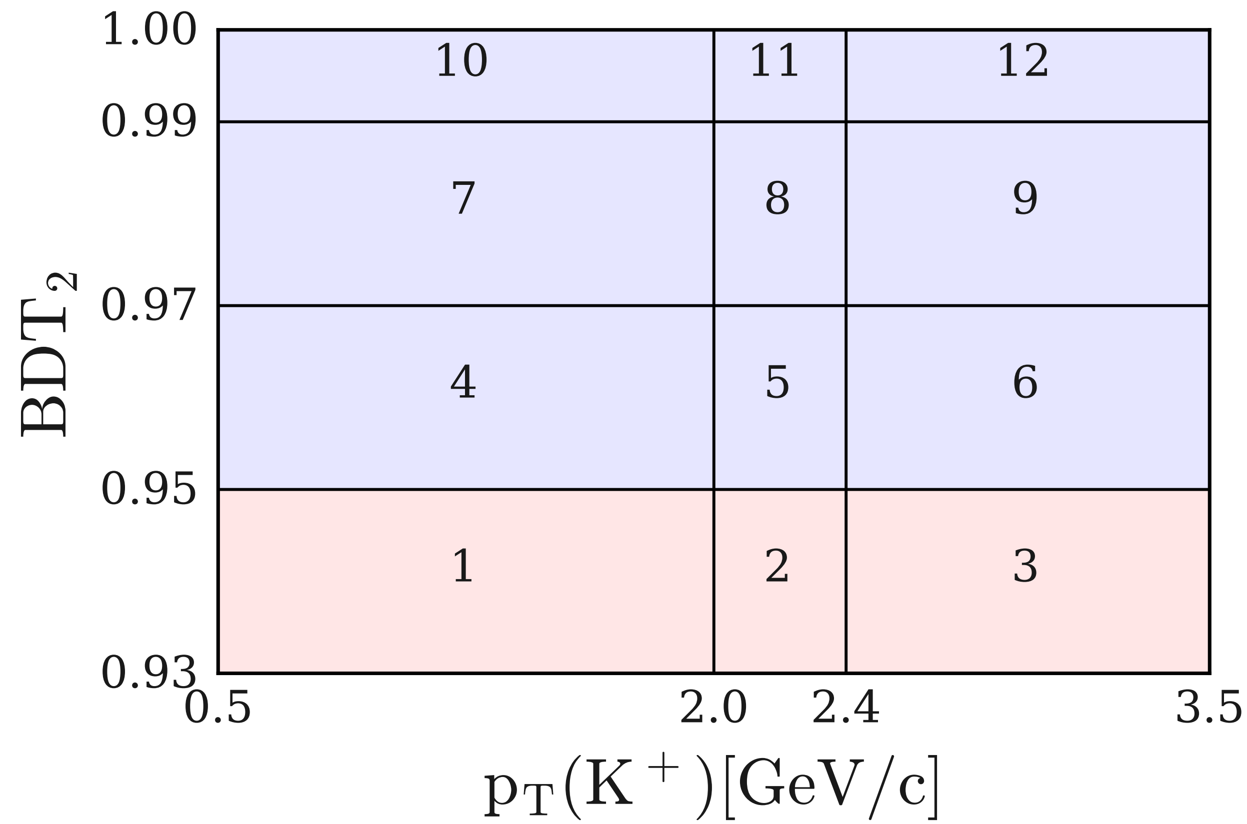
- **Signal Region (SR):** fit of data at the $\Upsilon(4S)$ resonance;
- **Control Region 2 (CR2):** fit of off-resonance data.

Bins 1,2,3:

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

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:

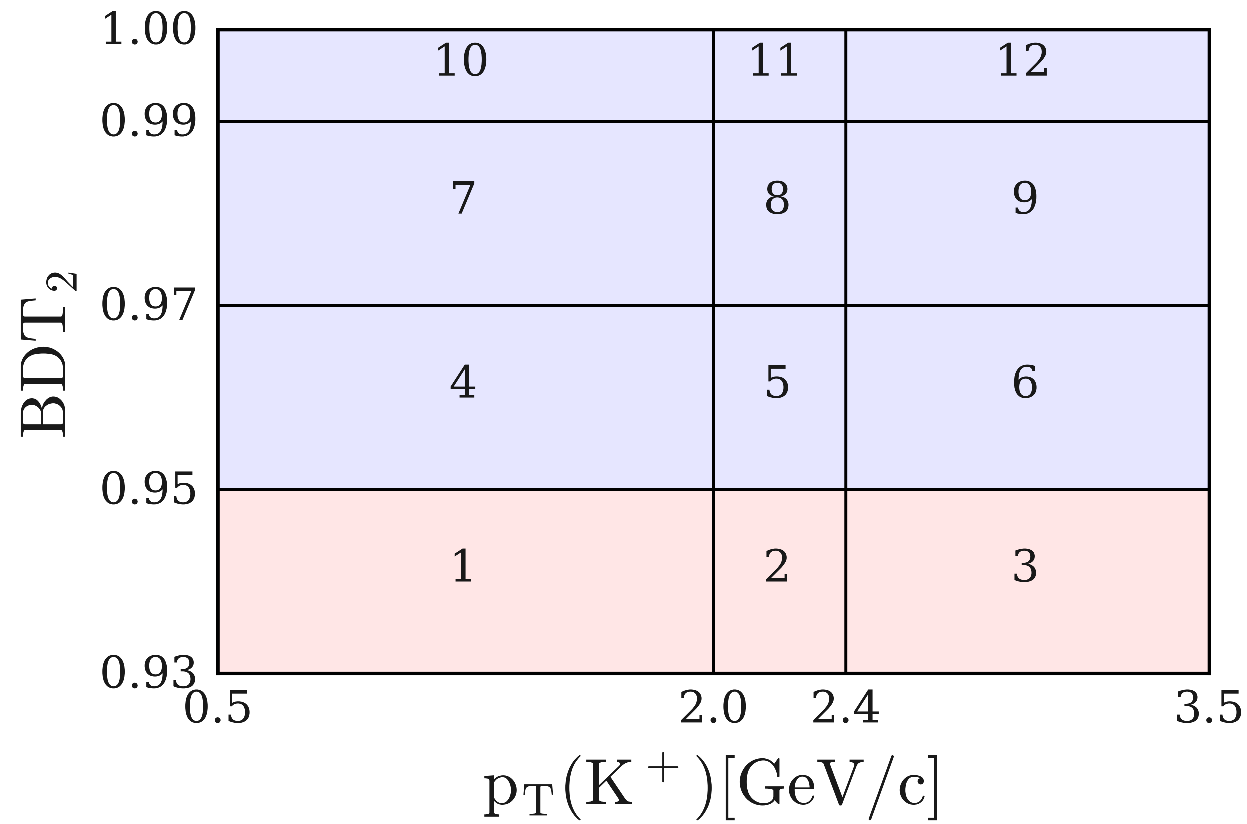
- **Signal Region (SR):** fit of data at the $\Upsilon(4S)$ resonance;
- **Control Region 2** (CR2)**:** fit of off-resonance data.

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.

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:

- **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.

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/ψ 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

$\bigoplus$

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

- Ignore the $\mu^+\mu^-$ from the selected J/ψ 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/ψ 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 $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$

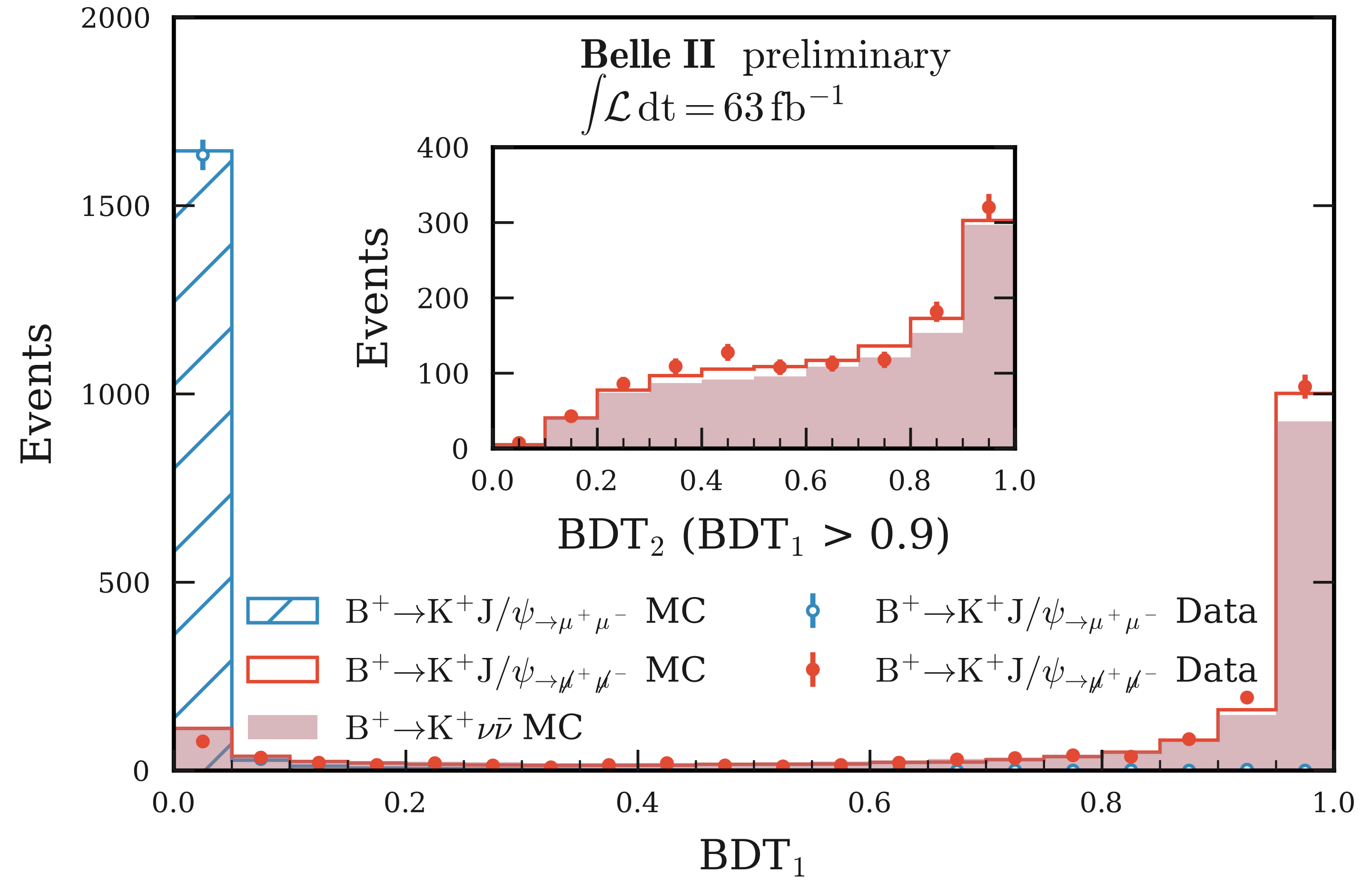
Mode with large branching ratio characterised by clean experimental signature.

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/ψ 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.

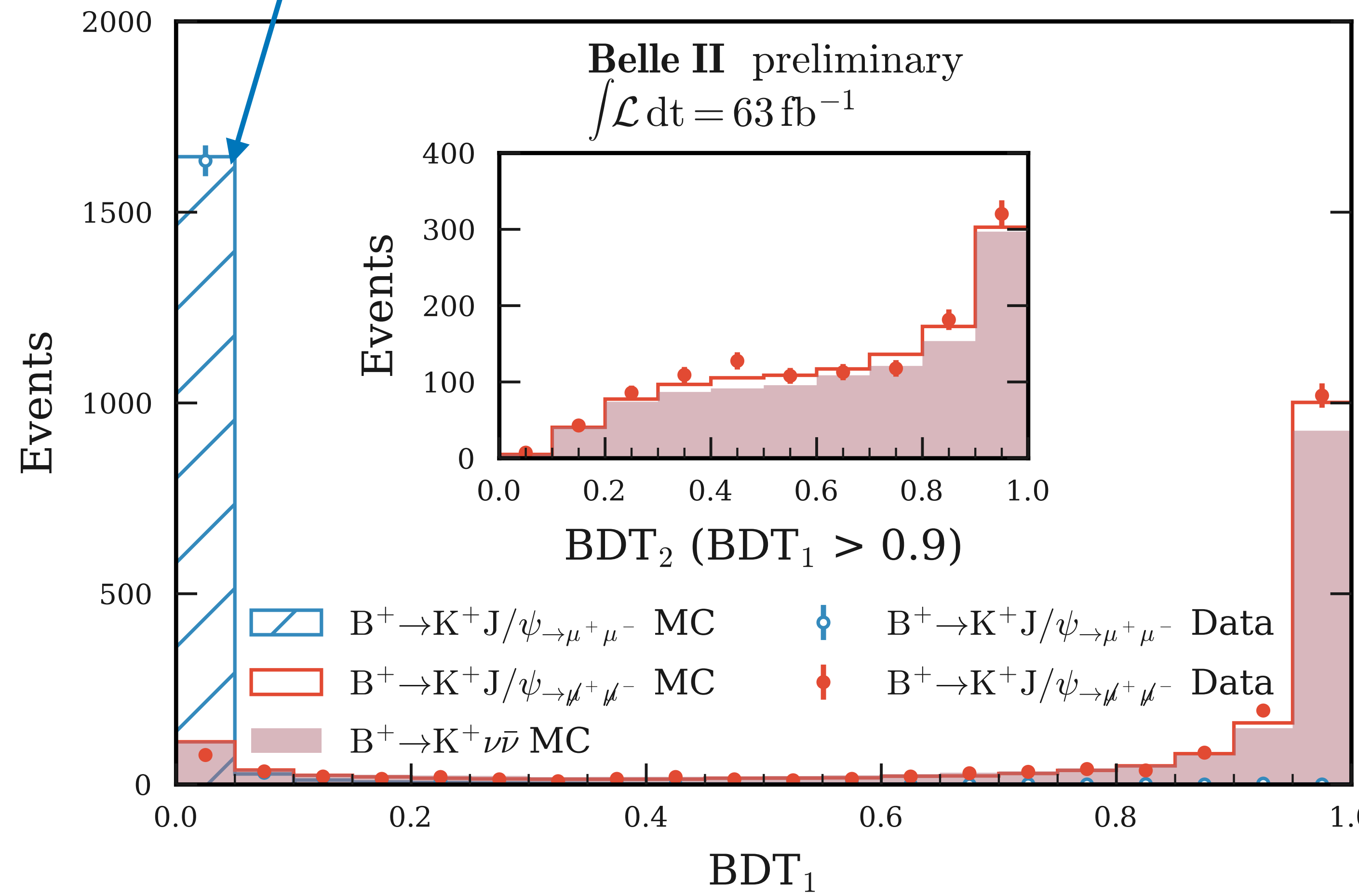
Identification of $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events

$\oplus$

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

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

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



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

Mode with large branching ratio characterised by clean experimental signature.

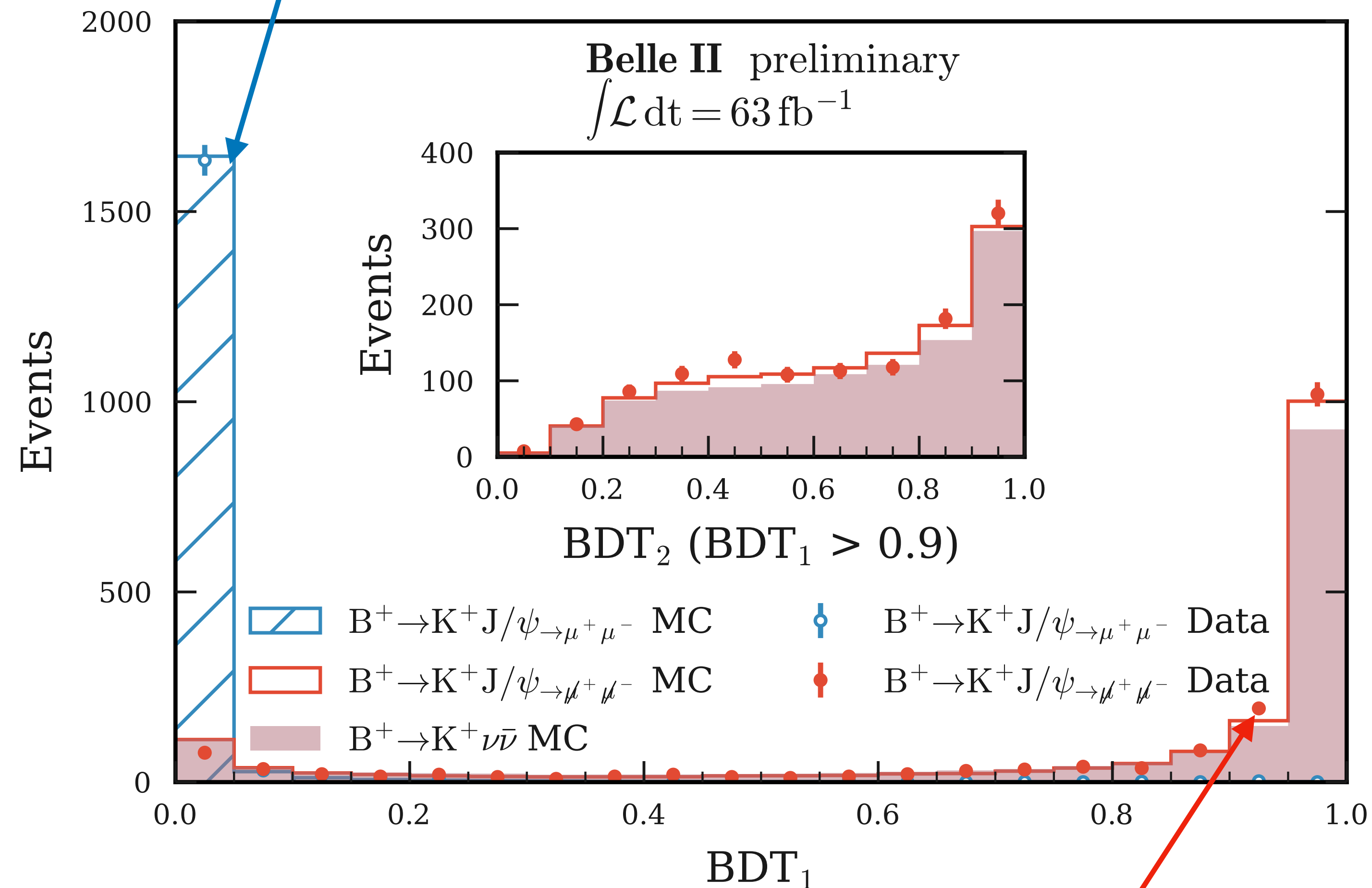
Identification of $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events

$\oplus$

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

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

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.**

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

Mode with large branching ratio characterised by clean experimental signature.

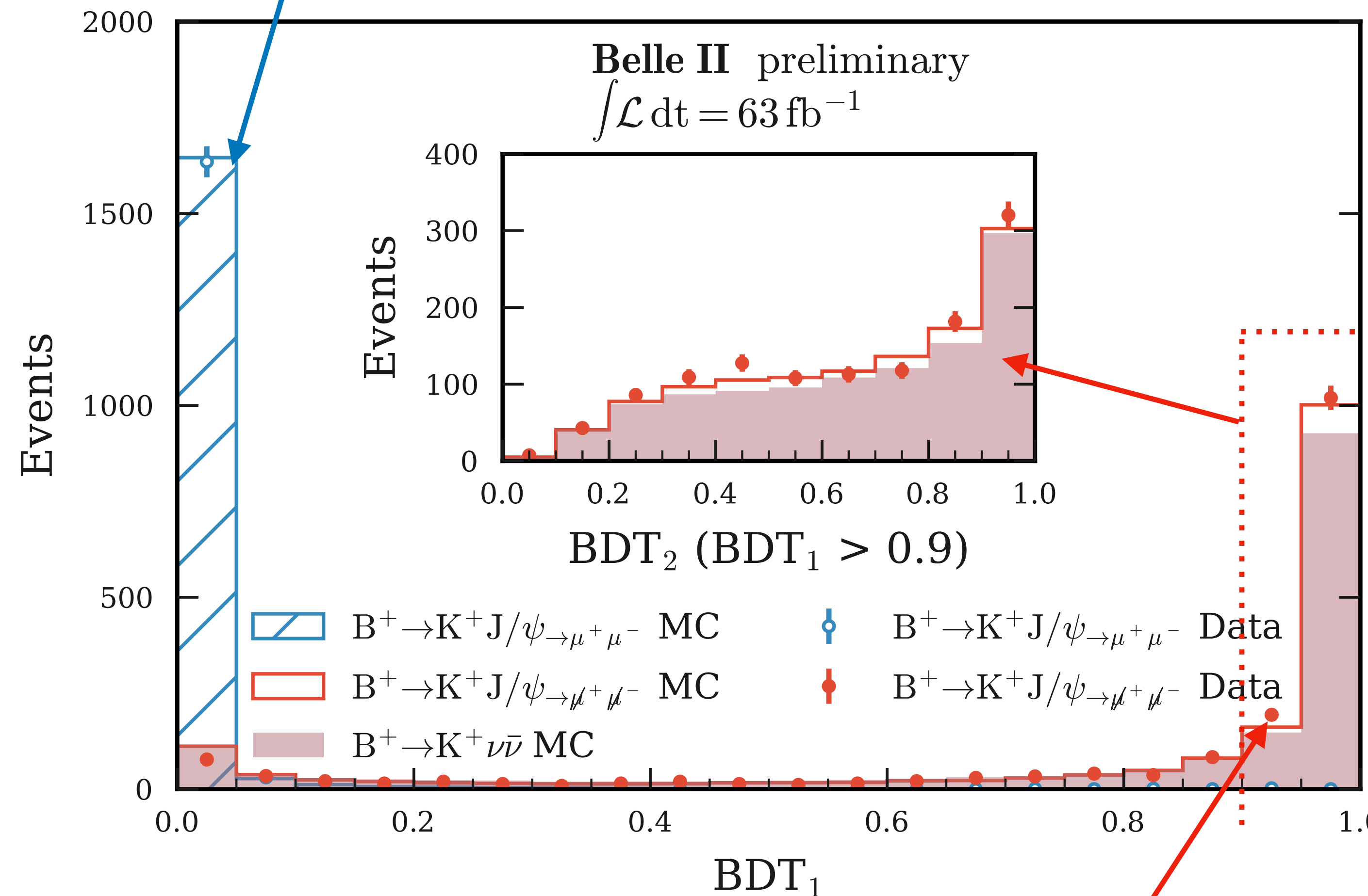
Identification of $B^+ \to K^+ J/\psi_{\to\mu^+\mu^-}$ events

$\oplus$

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

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

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.**

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

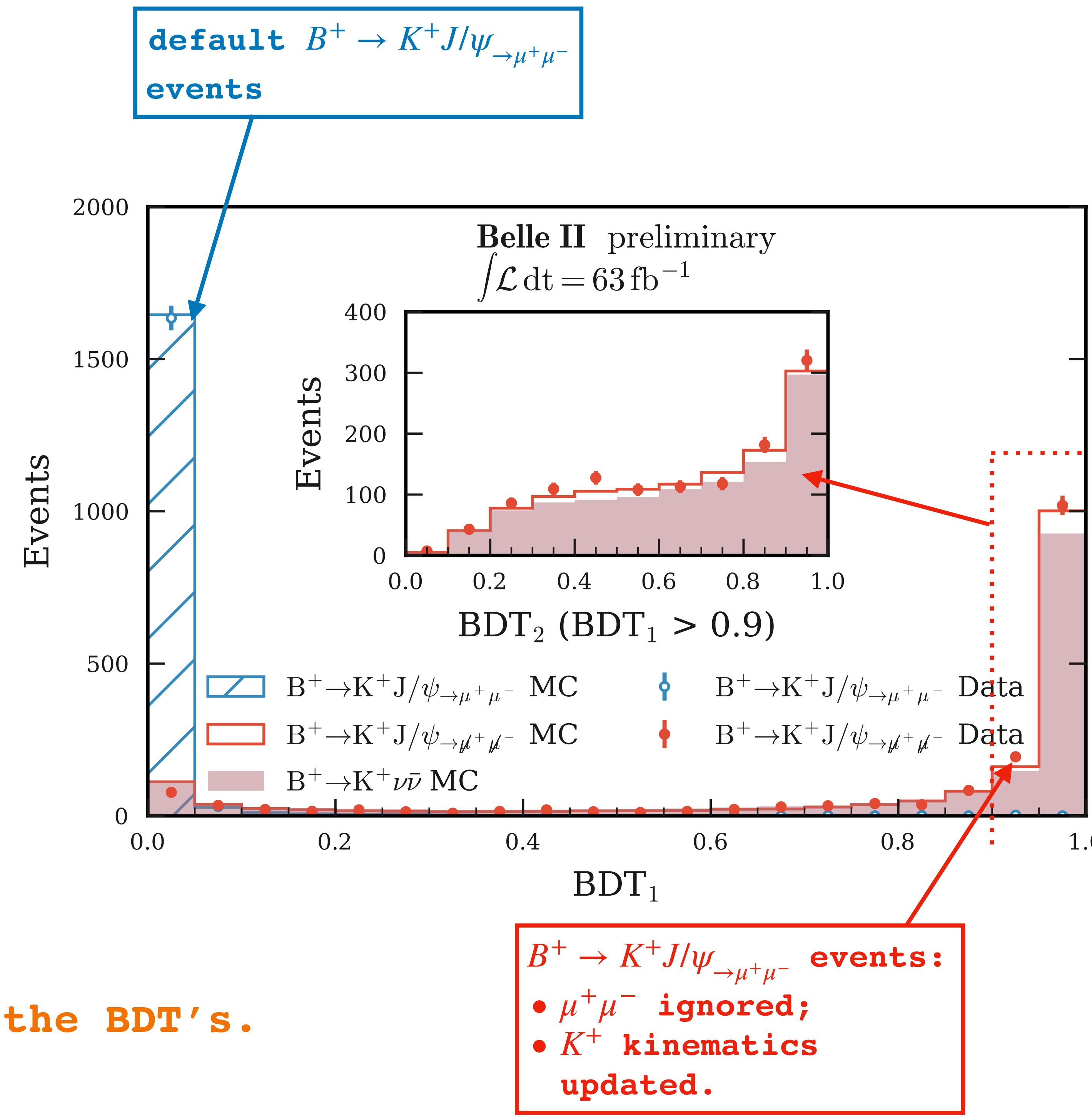
Mode with large branching ratio characterised by clean experimental signature.

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/ψ 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.



Excellent Data-MC agreement for the BDT's.

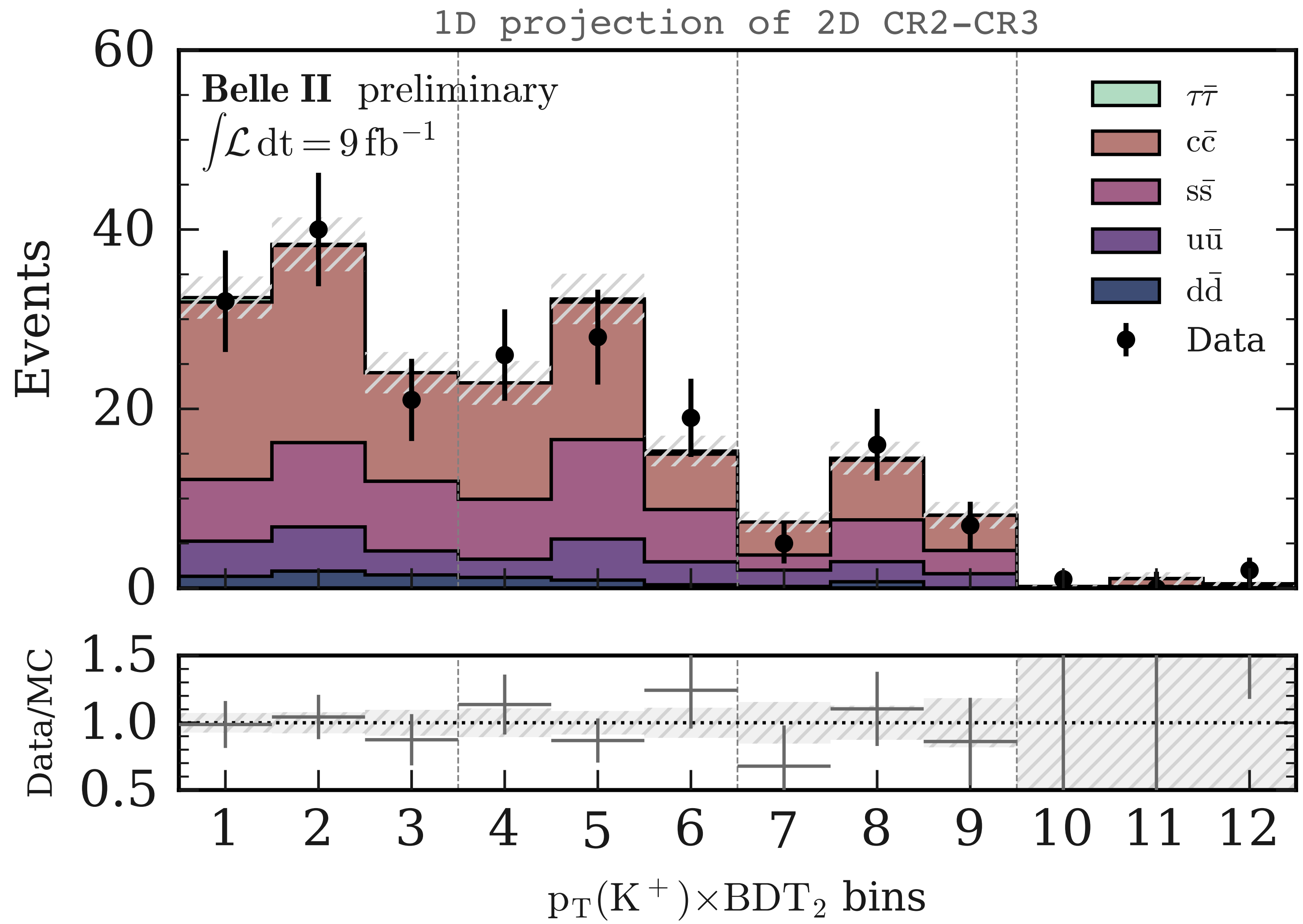
Validation using off-resonance data

Validation using off-resonance data

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

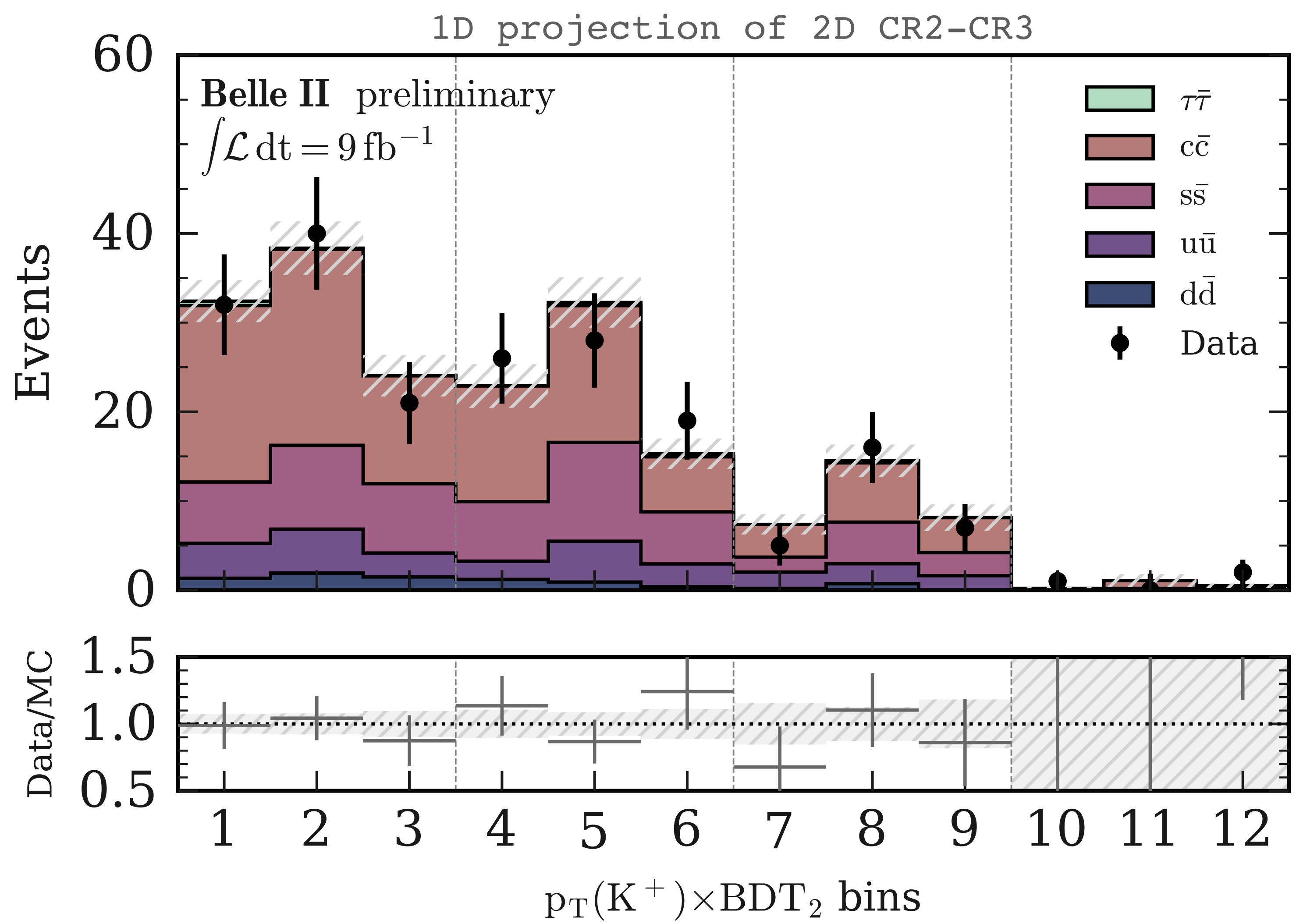
Validation using off-resonance data

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



Validation using off-resonance data

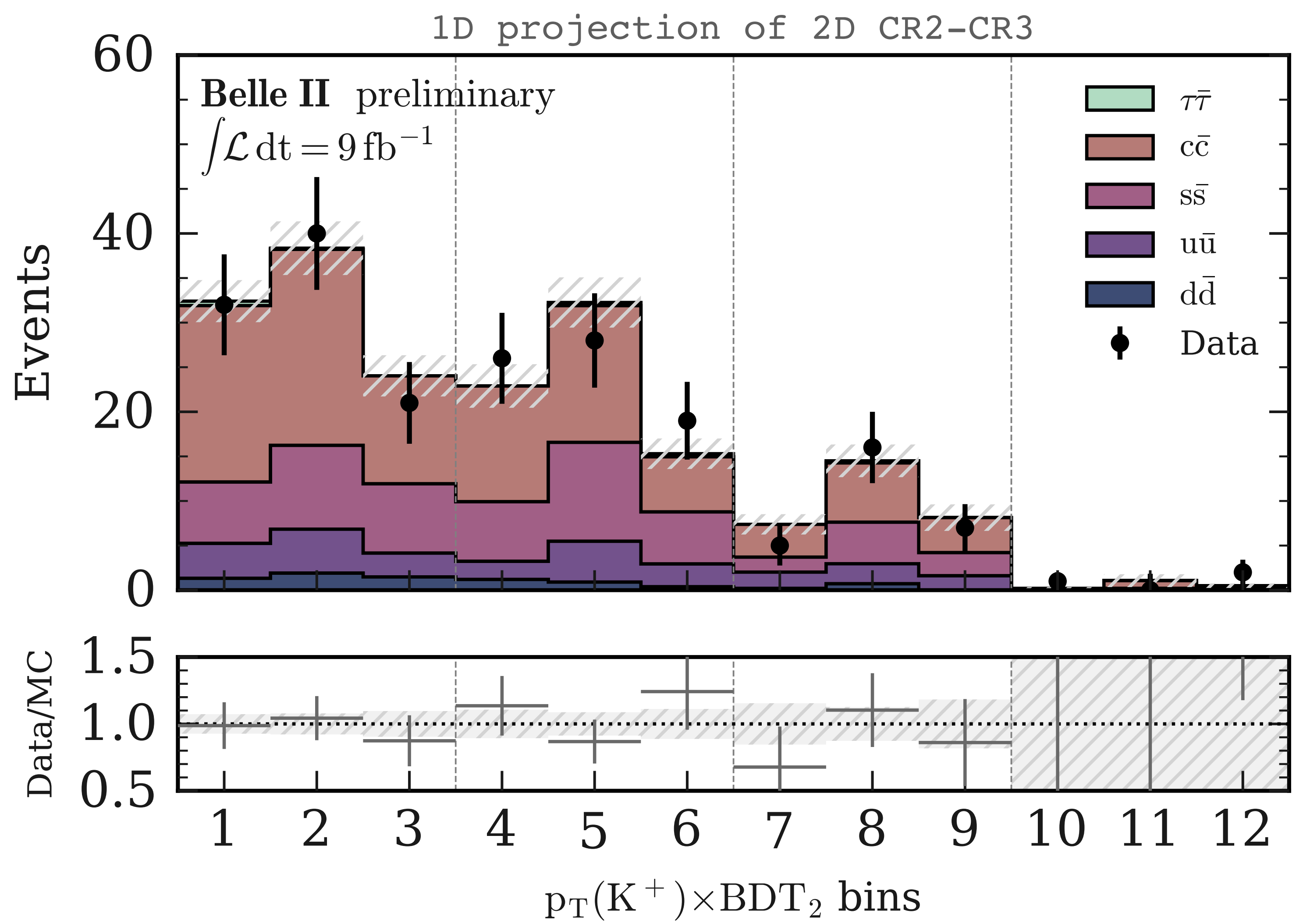
Investigation of the Data-MC agreement between simulated continuum and off-resonance data in CR2-CR3.



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

Validation using off-resonance data

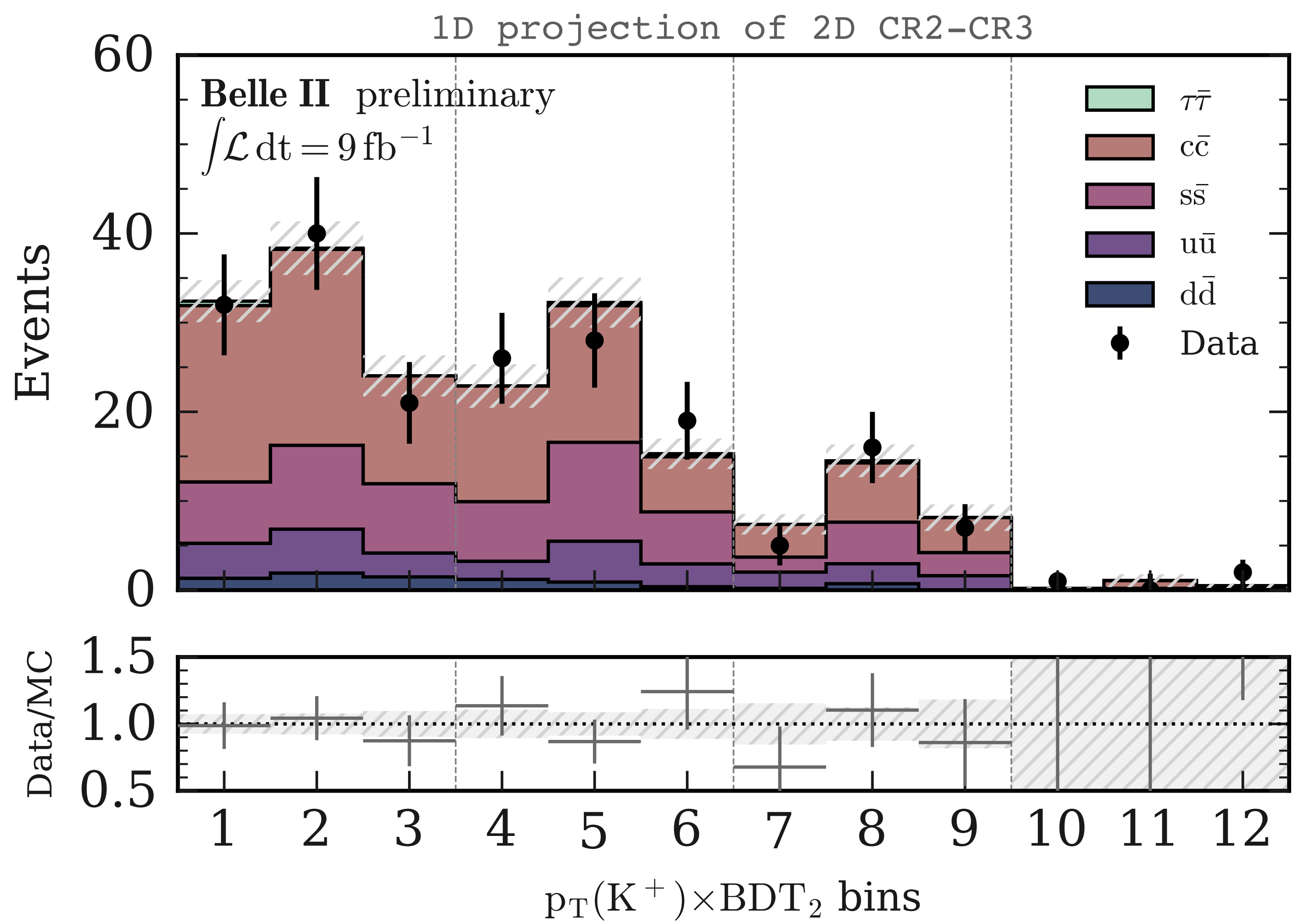
Investigation of the Data-MC agreement between simulated continuum and off-resonance data in CR2-CR3.



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

Fit procedure

Fit procedure

Statistical interpretation with pyhf [Heinrich, Lukas and Feickert, Matthew and Stark, Giordon. pyhf:v0.5.4.]

statistical model for multi-bin histogram-based analysis.

Fit procedure

Statistical interpretation with pyhf [Heinrich, Lukas and Feickert, Matthew and Stark, Giordon. pyhf:v0.5.4.]

statistical model for multi-bin histogram-based analysis.

Extended Maximum Likelihood Binned Fit:

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

η = parameter of interest
χ = nuisance parameters

Fit procedure

Statistical interpretation with pyhf [Heinrich, Lukas and Feickert, Matthew and Stark, Giordon. pyhf:v0.5.4.]

statistical model for multi-bin histogram-based analysis.

Extended Maximum Likelihood Binned Fit:

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

Simultaneous measurements of multiple regions

η = parameter of interest
χ = nuisance parameters

Fit procedure

Statistical interpretation with pyhf [Heinrich, Lukas and Feickert, Matthew and Stark, Giordon. pyhf:v0.5.4.]

statistical model for multi-bin histogram-based analysis.

Extended Maximum Likelihood Binned Fit:

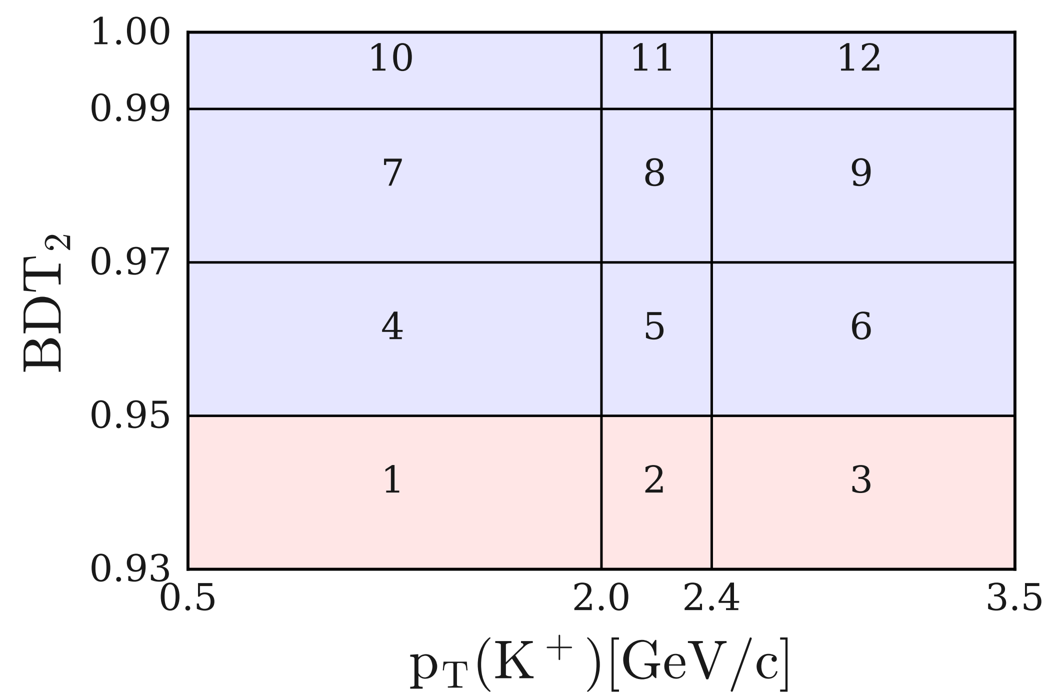
$$f(n, a \,|\, \eta, \chi) = \prod_{r \in \text{regions}} \prod_{b \in \text{bins}} \text{Pois}(n_{cb} \,|\, \nu_{cb}(\eta, \chi))$$

Simultaneous measurements of multiple regions

η = parameter of interest
χ = nuisance parameters

Regions = {SR, CR1, CR2, CR3}

Bins 1 to 12:



Fit procedure

Statistical interpretation with pyhf [Heinrich, Lukas and Feickert, Matthew and Stark, Giordon. pyhf:v0.5.4.]

statistical model for multi-bin histogram-based analysis.

Extended Maximum Likelihood Binned Fit:

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

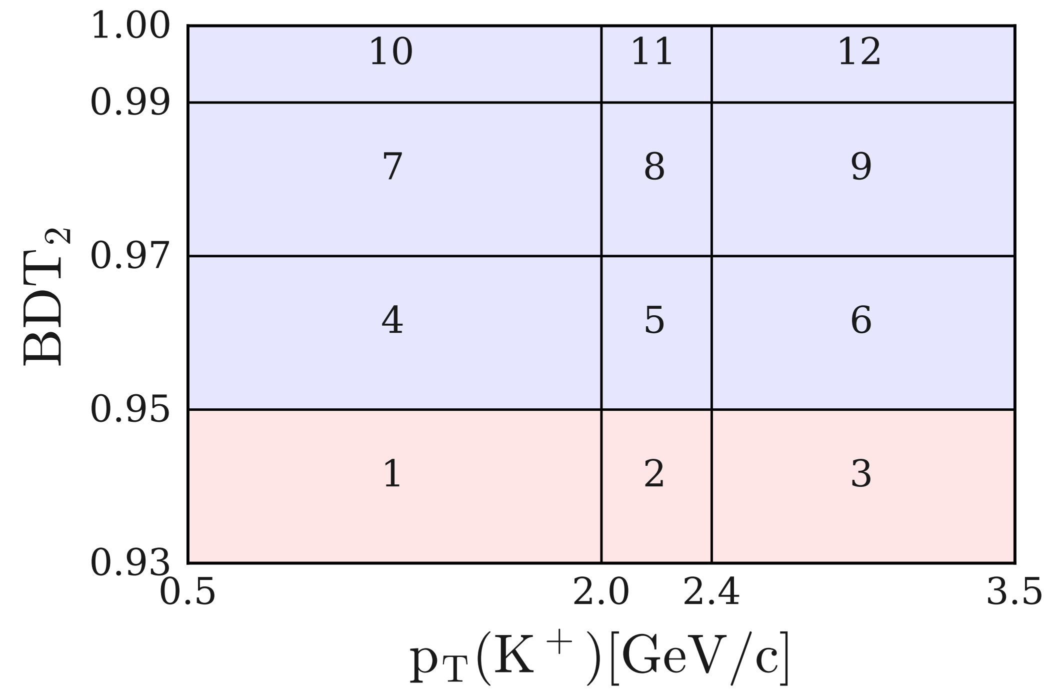
Simultaneous measurements of multiple regions

Constraints

η = parameter of interest
χ = nuisance parameters

Regions = {SR, CR1, CR2, CR3}

Bins 1 to 12:



Fit procedure

Statistical interpretation with pyhf [Heinrich, Lukas and Feickert, Matthew and Stark, Giordon. pyhf:v0.5.4.]

statistical model for multi-bin histogram-based analysis.

<u>Extended Maximum Likelihood Binned Fit:</u>

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

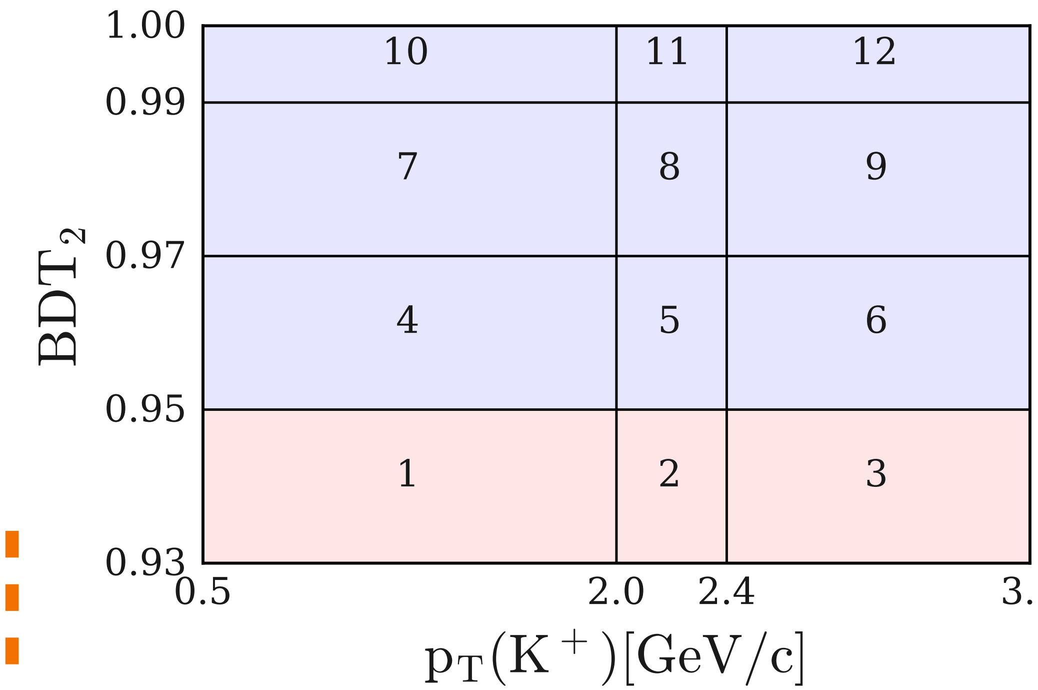
Simultaneous measurements of multiple regions

Constraints

η = parameter of interest
χ = nuisance parameters

Regions = {SR, CR1, CR2, CR3}

Bins 1 to 12:



- Systematic uncertainties (normalisations of bkg's yields, BR of the leading *B*-decays, PID correction, …) as (175) nuisance parameters.

- **1 parameter of interest: signal strength μ:** multiplicative factor with respect to the SM expectation.

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

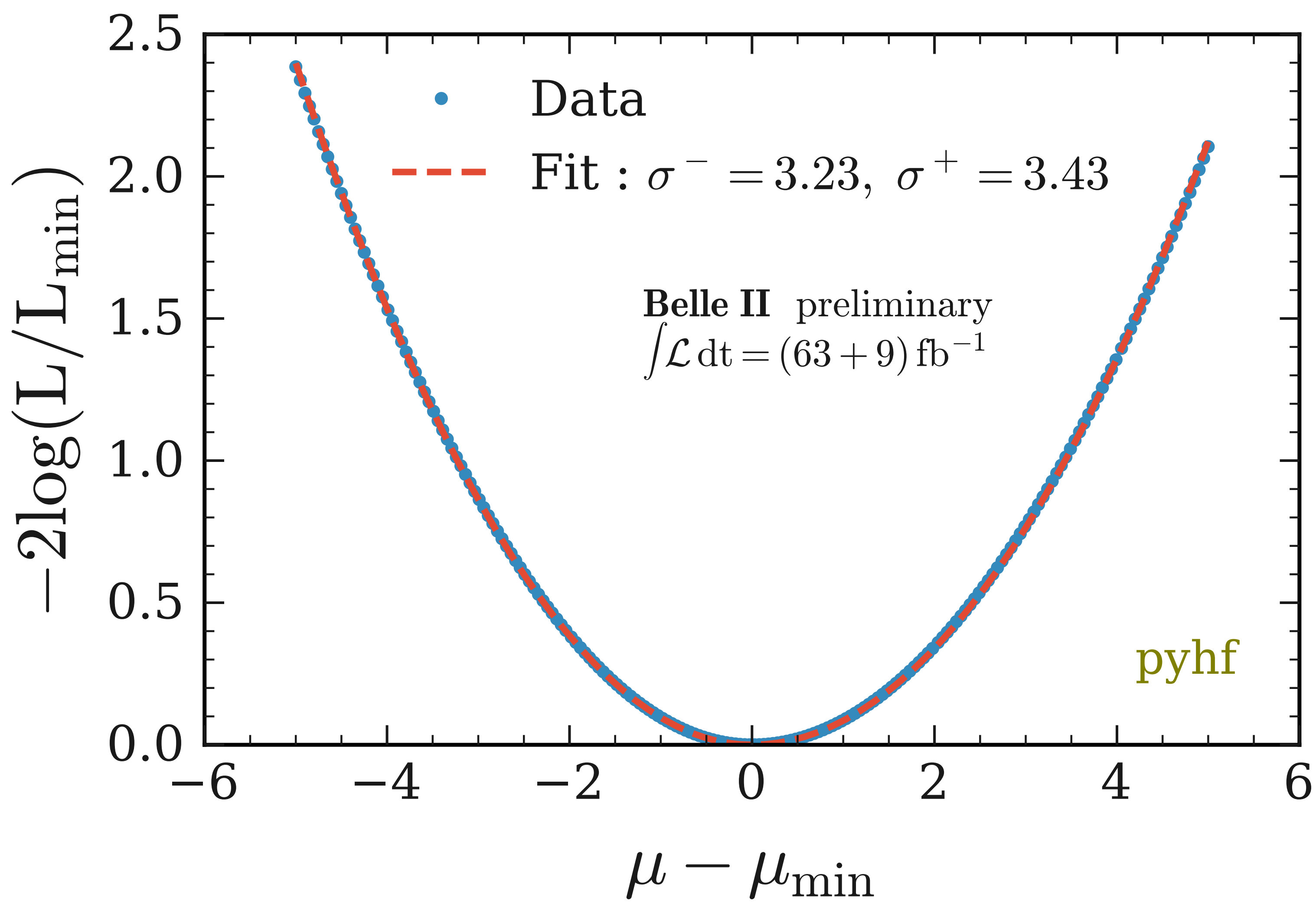
Fit to the Data

Fit to the Data

- Profile likelihood scan for the signal strength μ:

Fit to the Data

- Profile likelihood scan for the signal strength μ:



Asymmetric uncertainty on μ 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

Fit to the Data

- **Measured signal strength μ:**

Fit to the Data

- **Measured signal strength μ:**

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

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

Fit to the Data

- **Measured signal strength μ:**

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

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

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

Fit to the Data

- **Measured signal strength μ:**

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

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

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

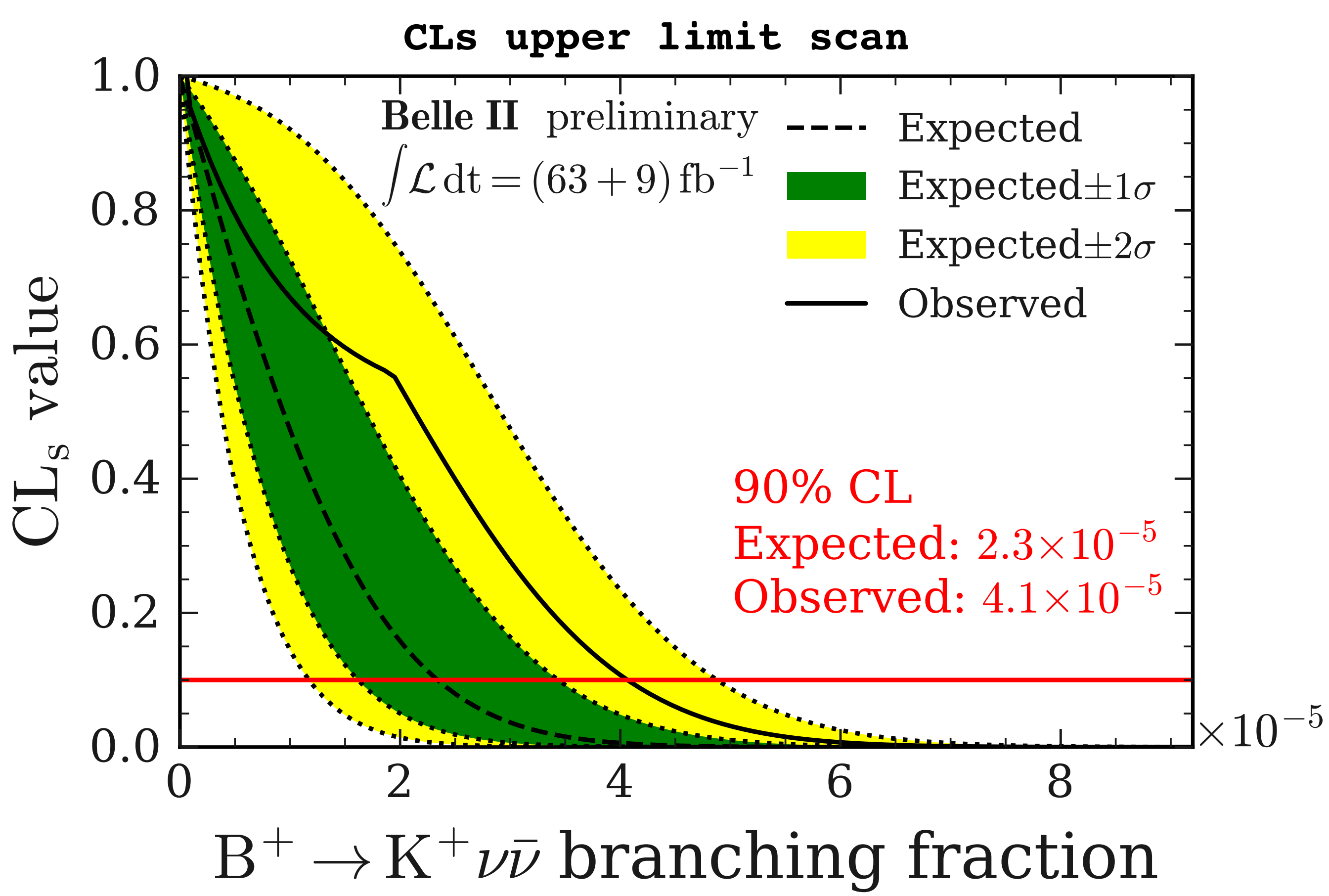
Fit to the Data

- **Measured signal strength μ:**

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

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

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



Fit to the Data

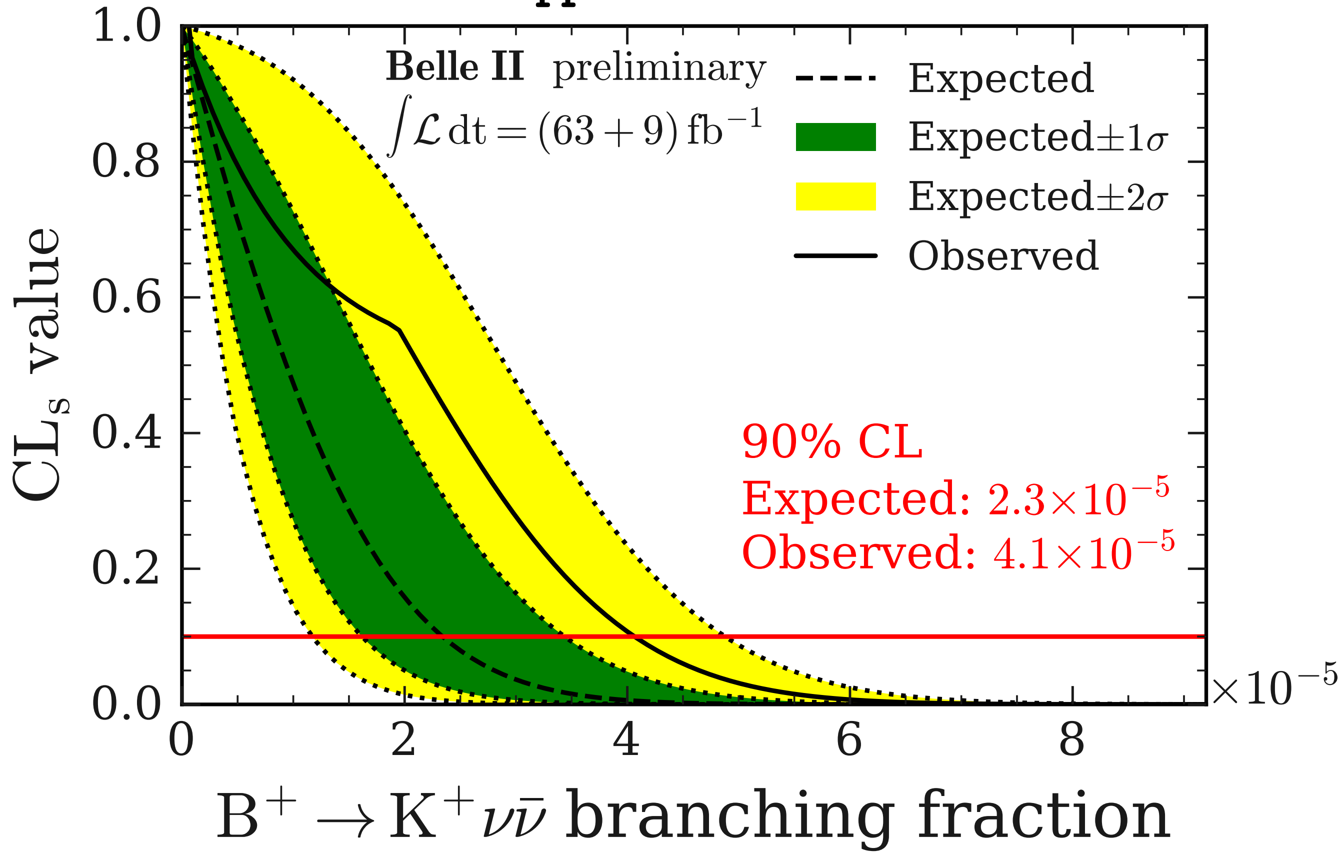
- **Measured signal strength μ:**

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

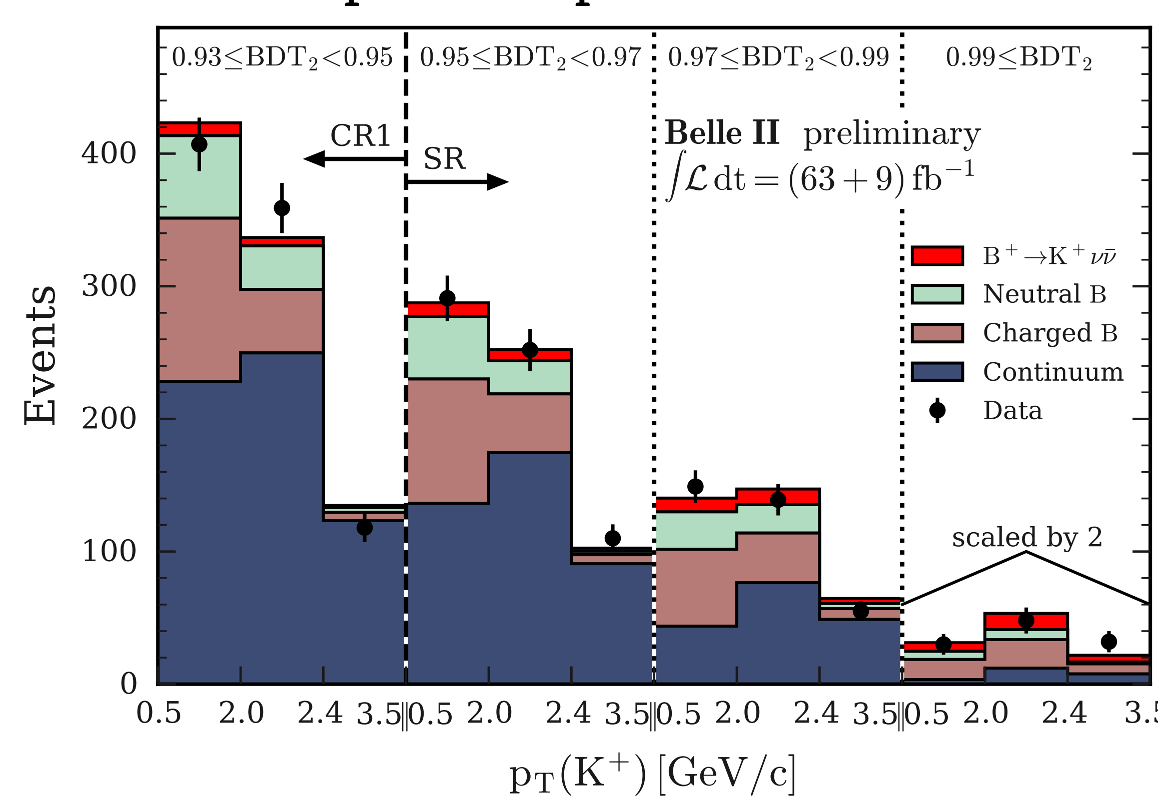
$$\text{BR}(B^+ \to K^+\nu\bar{\nu}) = 1.9^{+1.3}_{-1.3}(\text{stat})^{+0.8}_{-0.7}(\text{syst}) \times 10^{-5} = 1.9^{+1.6}_{-1.5} \times 10^{-5}$$

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

CLs upper limit scan



Data vs post-fit predictions in CR1 + SR



Measurement summary

Measurement summary

- This measurement represents the **first search for $B^+ \to K^+\nu\bar{\nu}$ performed with an inclusive tag.**

Measurement summary

- This measurement represents the **first search for $B^+ \to K^+\nu\bar{\nu}$ performed with an inclusive tag.**
- **No signal observed yet, but an observed upper limit on the branching ratio of 4.1×10^{-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 tag.**
- **No signal observed yet, but an observed upper limit on the branching ratio of 4.1×10^{-5} is set at the 90% CL.**
- With $63\ \mathrm{fb}^{-1}$ of $\Upsilon(4S)$ data recorded by the Belle II experiment, the inclusive tagging is **competitive with the previous searches despite the much lower integrated luminosity.**

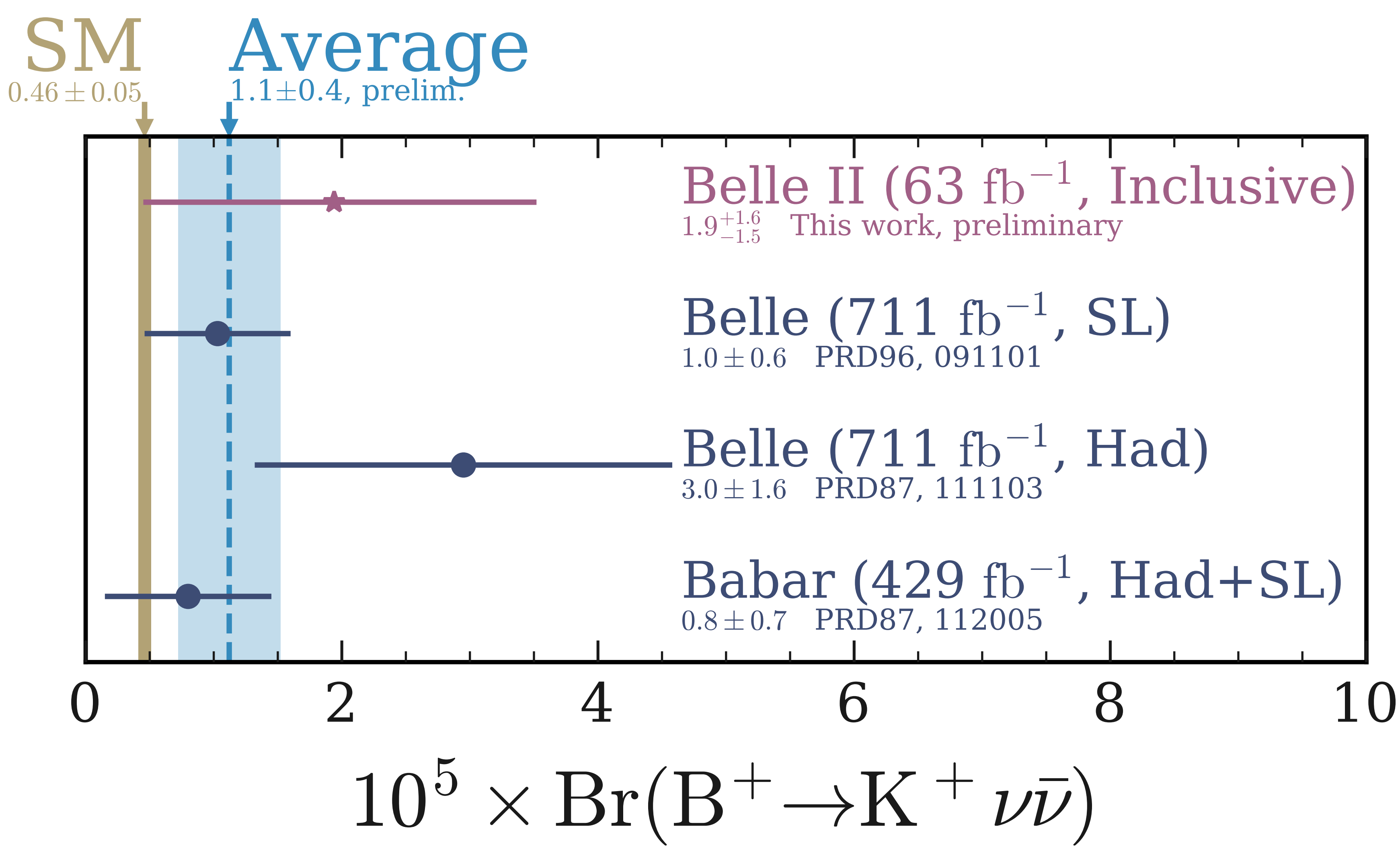
Experiment	Year	Observed limit on $\mathrm{BR}(B^+ \to K^+\nu\bar{\nu})$	Approach	Data $[\mathrm{fb}^{-1}]$
BABAR	**2013**	$< 1.6 \times 10^{-5}$ **[Phys.Rev.D87,112005]**	**SL + Had tag**	429
Belle	**2013**	$< 5.5 \times 10^{-5}$ **[Phys.Rev.D87,111103(R)]**	**Had tag**	711
Belle	**2017**	$< 1.9 \times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

Measurement summary

- This measurement represents the **first search for** $B^+ \to K^+ \nu\bar{\nu}$ **performed with an inclusive tag.**

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

- With $63\ \mathrm{fb}^{-1}$ of $\Upsilon(4S)$ data recorded by the Belle II experiment, the inclusive tagging is **competitive with the previous searches despite the much lower integrated luminosity.**



Supplemental material

Summary of the $B^+ \to K^+\nu\bar{\nu}$ searches

- Uncertainty on BR: Belle II vs Belle vs Babar.

  - BR: measured branching ratio of $B^+ \to K^+\nu\bar{\nu}$;
  - σ : total symmetric uncertainty on the BR
  - L : integrated luminosity

BABAR 2013 - [Phys.Rev.D87,112005]
Belle 2013 - [Phys.Rev.D87,111103(R)]
Belle 2017 - [Phys.Rev.D96,091101(R)]

Experiment	Year	Approach	$L[\mathrm{fb}^{-1}]$	$\mathrm{BR}[\times 10^{-5}]$	$\sigma\,[\times 10^{-5}]$	$\sigma\sqrt{\frac{L}{L_{\mathrm{Belle2}}}}\,[\times 10^{-5}]$
BABAR(*)	**2013**	**SL + Had tag**	429	0.8	0.6	1.7
Belle(**)	**2013**	**Had tag**	711	3.0	1.6	5.5
Belle(**)	**2017**	**SL tag**	711	1.0	0.6	1.9
Belle II	**2021**	**Inclusive tag**	63	1.9	1.6	1.6

(*) Combined central value of $B^+ \to K^+\nu\bar{\nu}$ / $B^0 \to K^0\nu\bar{\nu}$
(**) Computed from $N_{\mathrm{sig}}/(\epsilon_{\mathrm{sig}} \cdot N_{B\bar{B}})$.

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

Scenarios beyond the SM → possible contribution of right-handed operators Q^l_R

$$\mathscr{H}_{eff.} = -\frac{4G_F}{\sqrt{2}} V_{tb}V^*_{ts} \sum_l (C^l_L Q^l_L + C^l_R Q^l_R) \quad \text{where} \quad Q^l_{L(R)} = \left(\bar{s}_{L(R)}\gamma_\mu b_{L(R)}\right)\left(\bar{\nu}^l_{L(R)}\gamma^\mu \nu^l_{L(R)}\right) \quad l = e, \mu, \tau$$

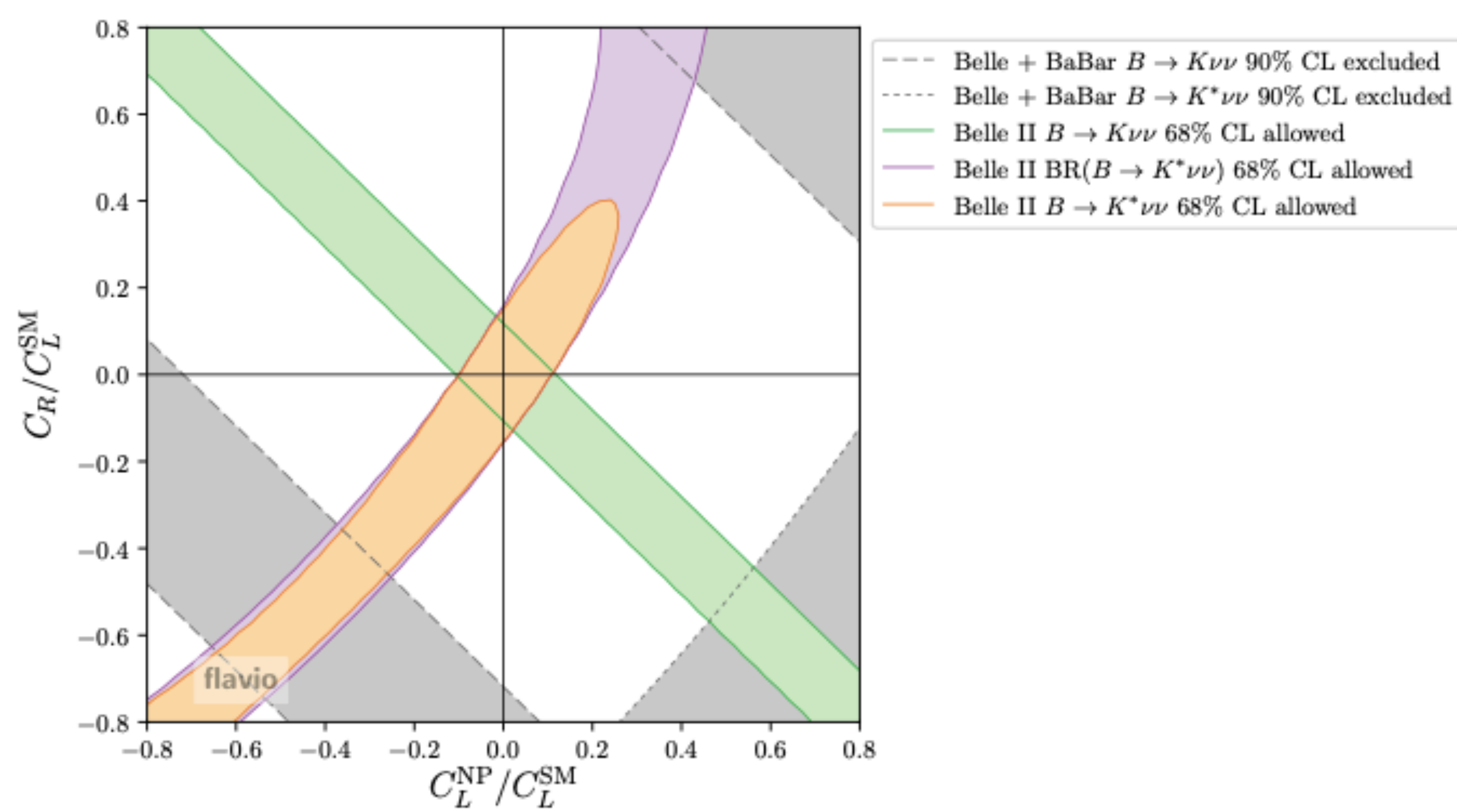
2 combinations of 6 Wilson Coefficients:

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

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

$$\epsilon_\ell = \frac{\sqrt{|C^\ell_L|^2 + |C^\ell_R|^2}}{|C^{\mathrm{SM}}_L|}\,,$$

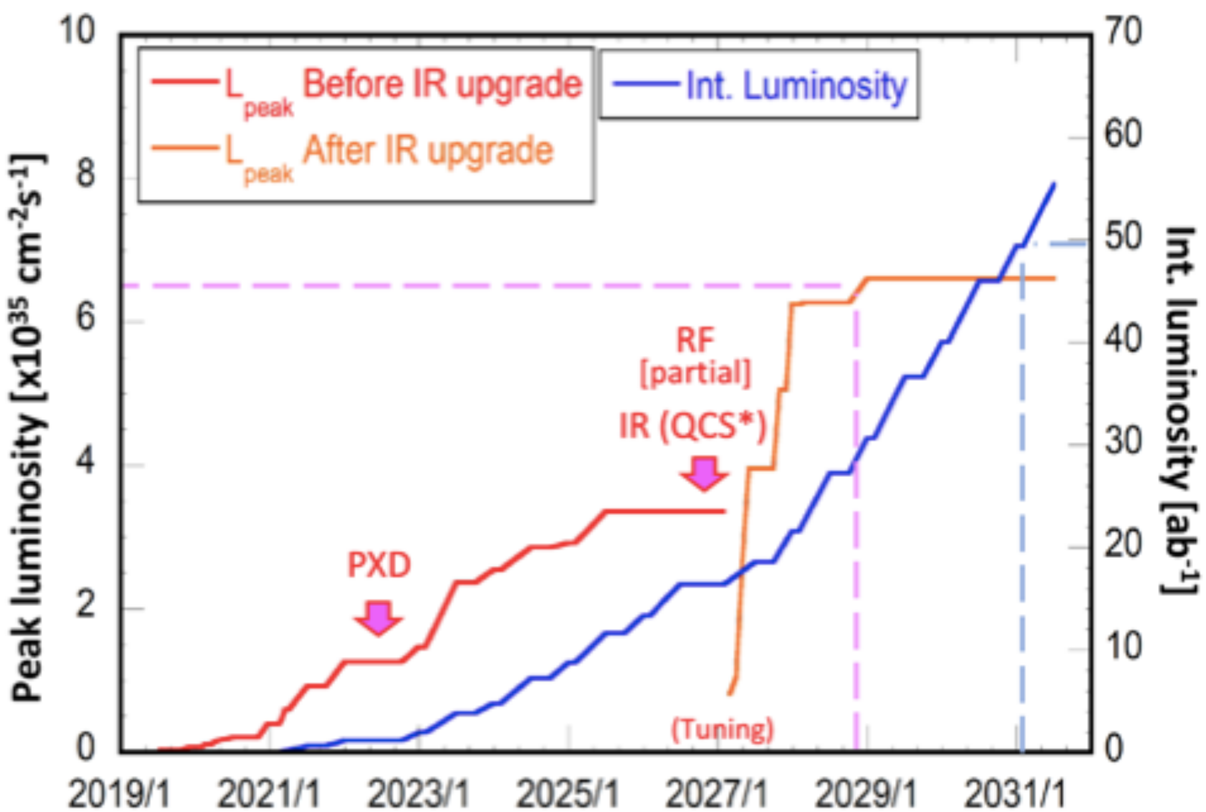
$$\eta_\ell = \frac{-\mathrm{Re}\left(C^\ell_L C^{\ell *}_R\right)}{|C^\ell_L|^2 + |C^\ell_R|^2}\,.$$



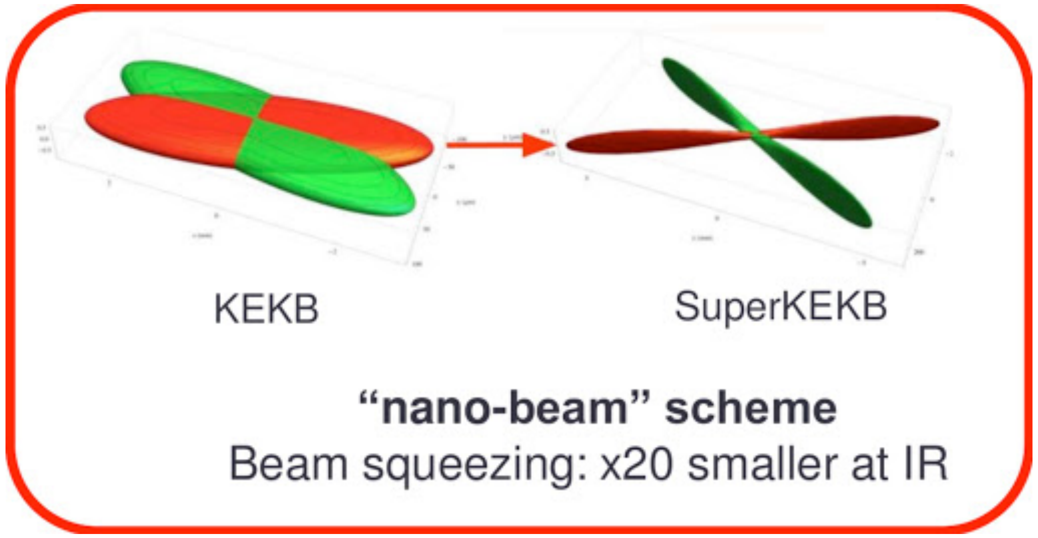
Constraint on new-physics contributions: Wilson coefficients C_L^{NP} and C_R normalised to the SM value of C_L (Belle II from expected 50 ab^{-1}).

SuperKEKB

- Peak luminosity projections:

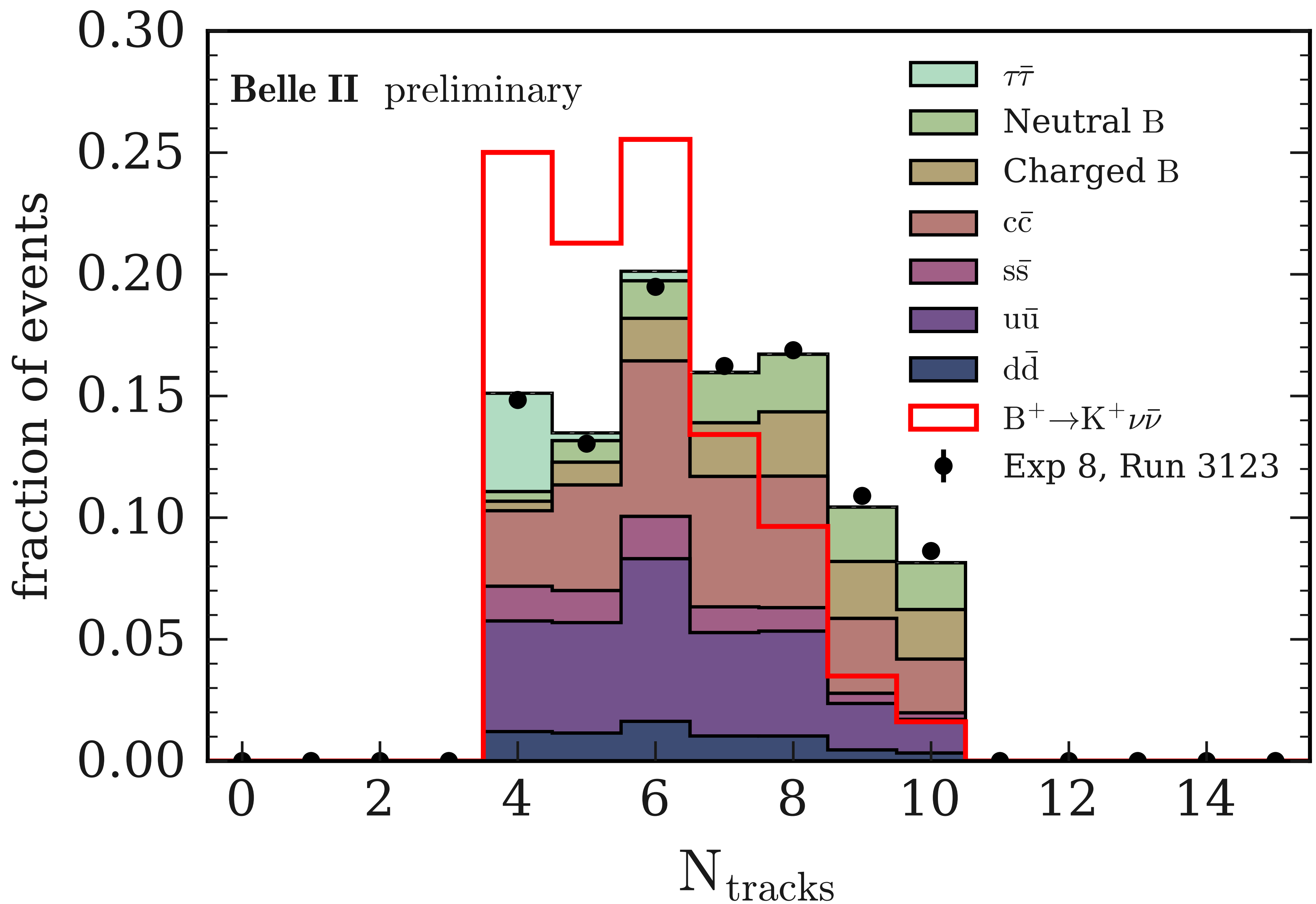


- Nano-beam scheme:



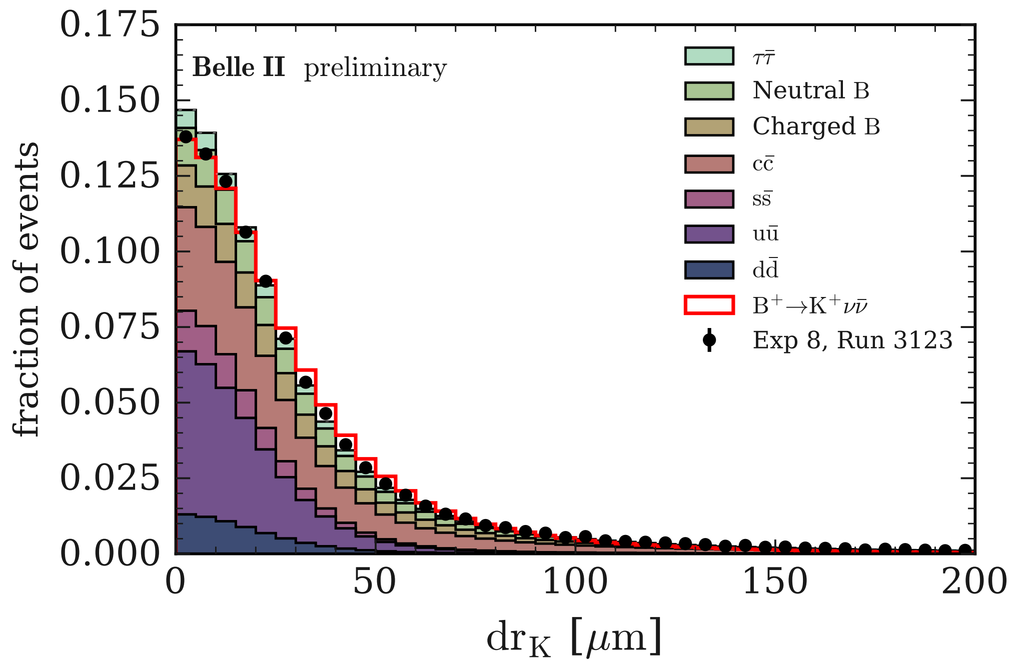
Features of $B^+ \to K^+\nu\bar{\nu}$

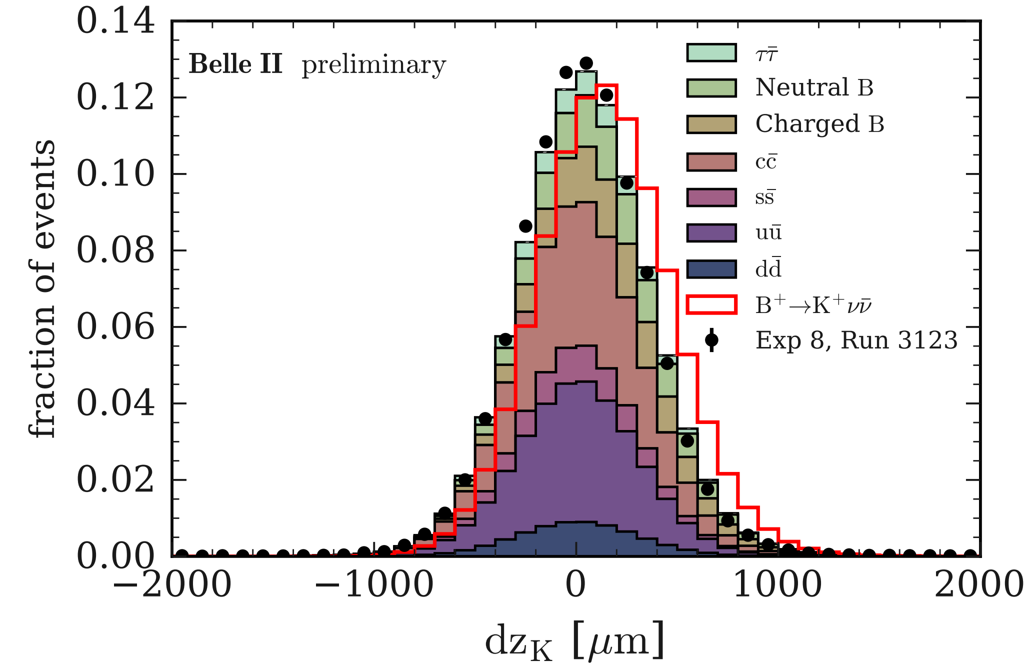
- Number of reconstructed tracks in the event.



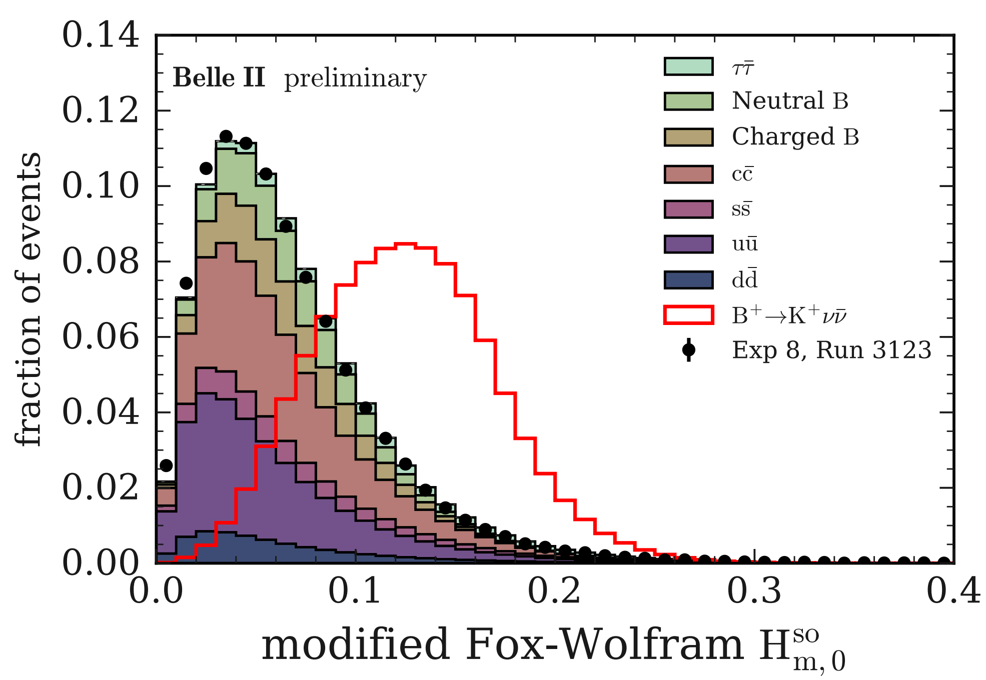
Features of $B^+ \to K^+\nu\bar{\nu}$

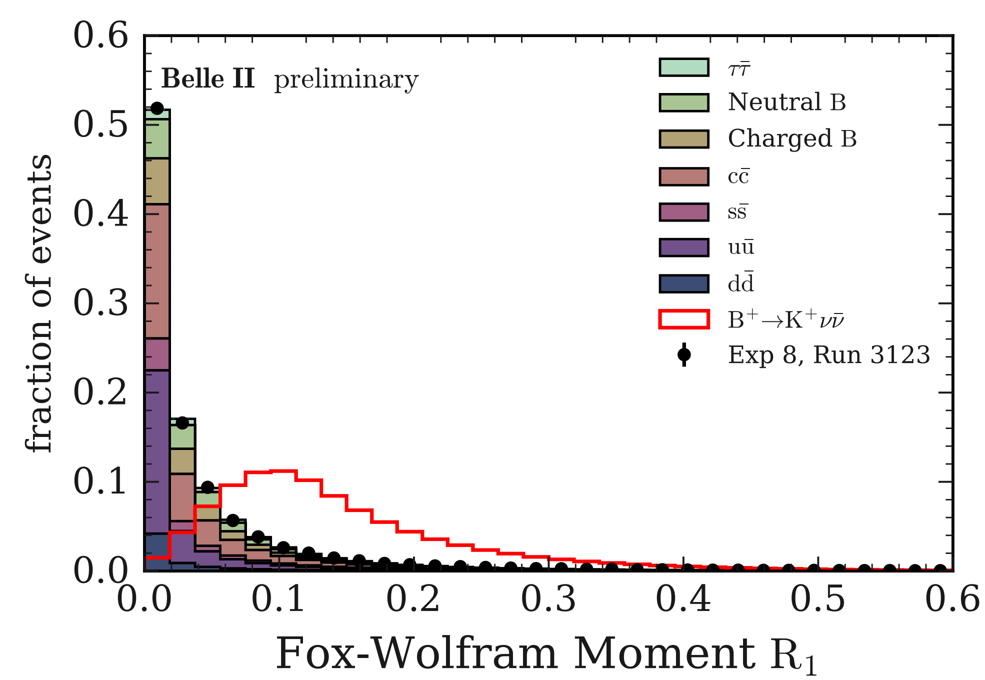
- Variables related to the signal K^+candidate.





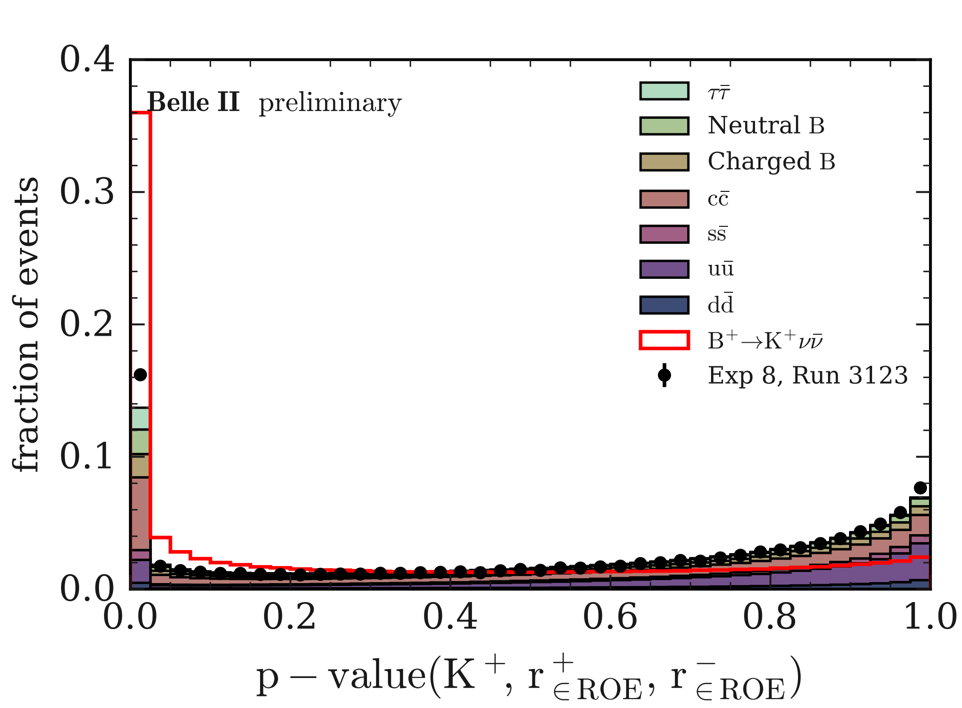
- Variables related to the event topology.

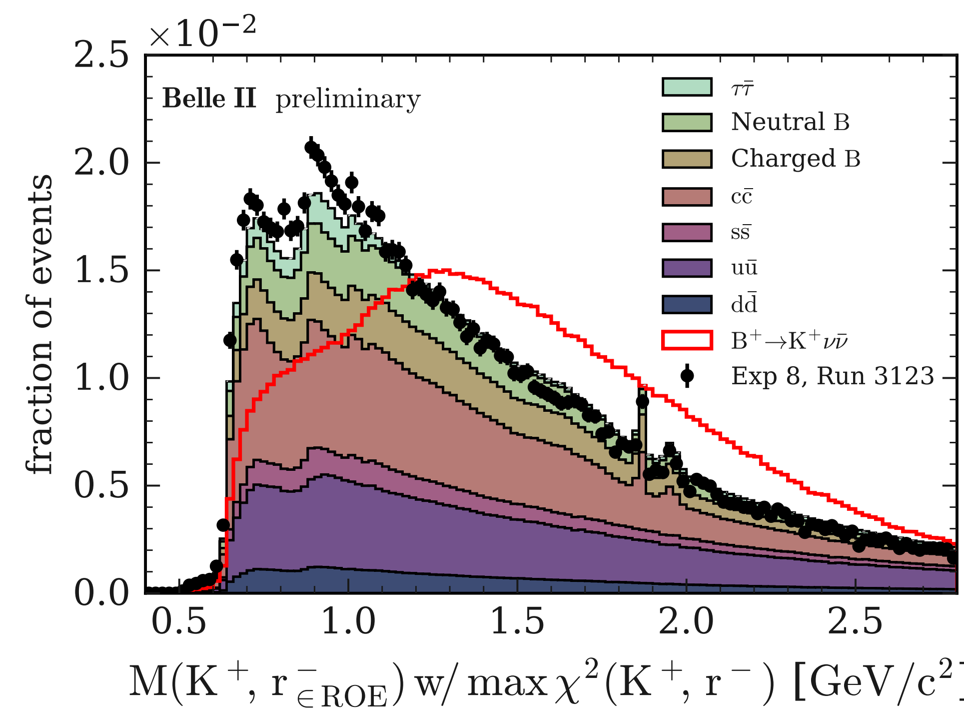




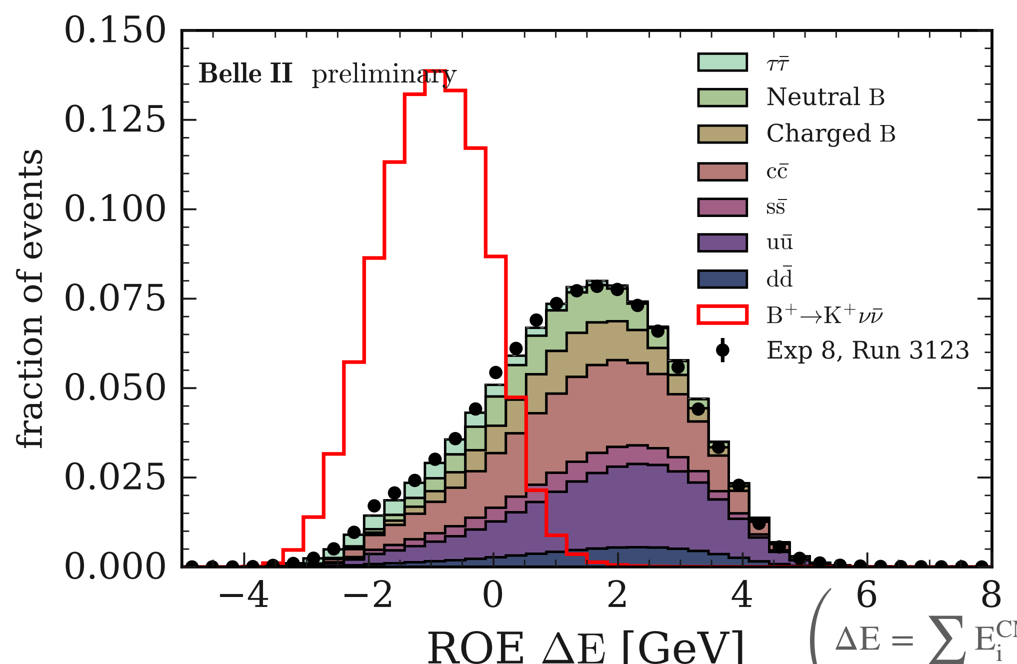
Features of $B^+ \to K^+\nu\bar{\nu}$

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

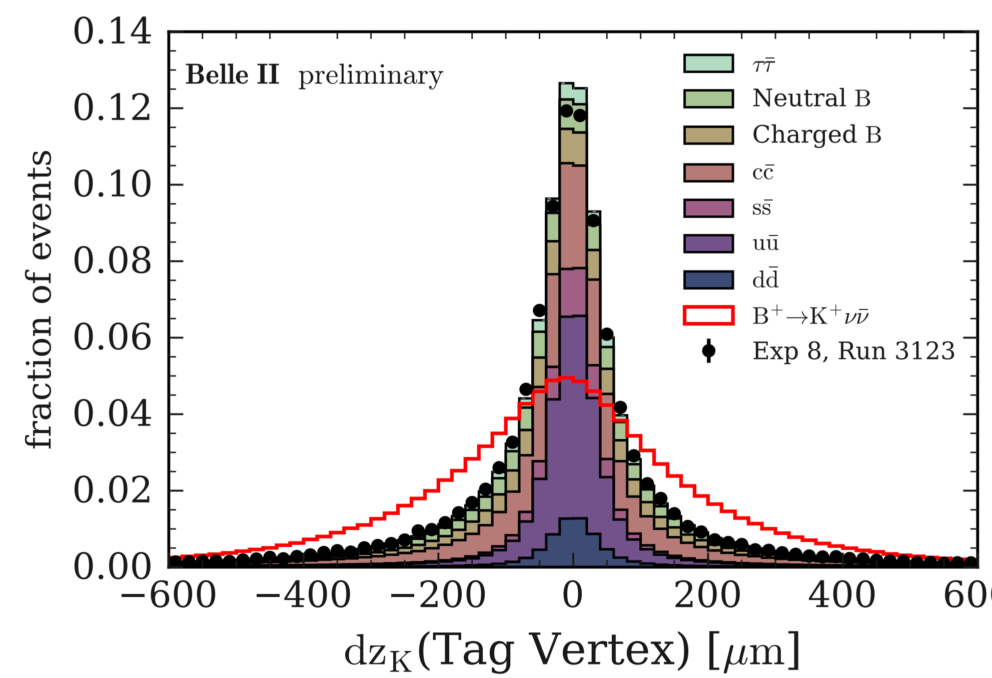




- Variables related to the ROE.

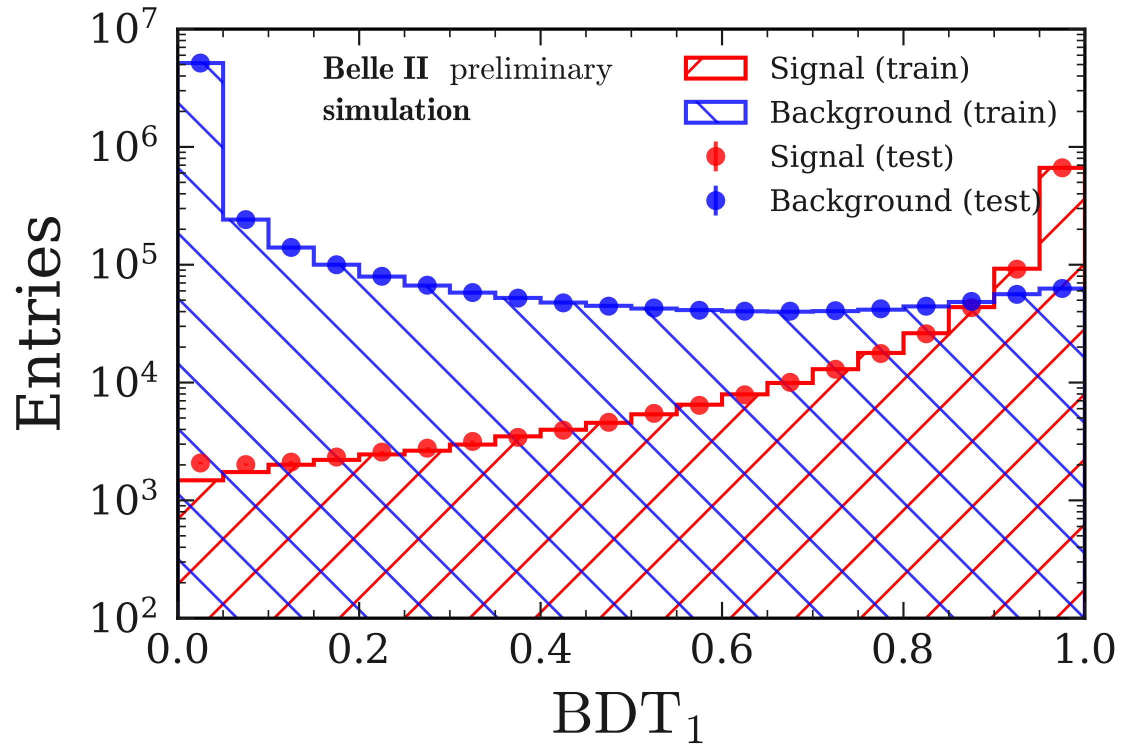


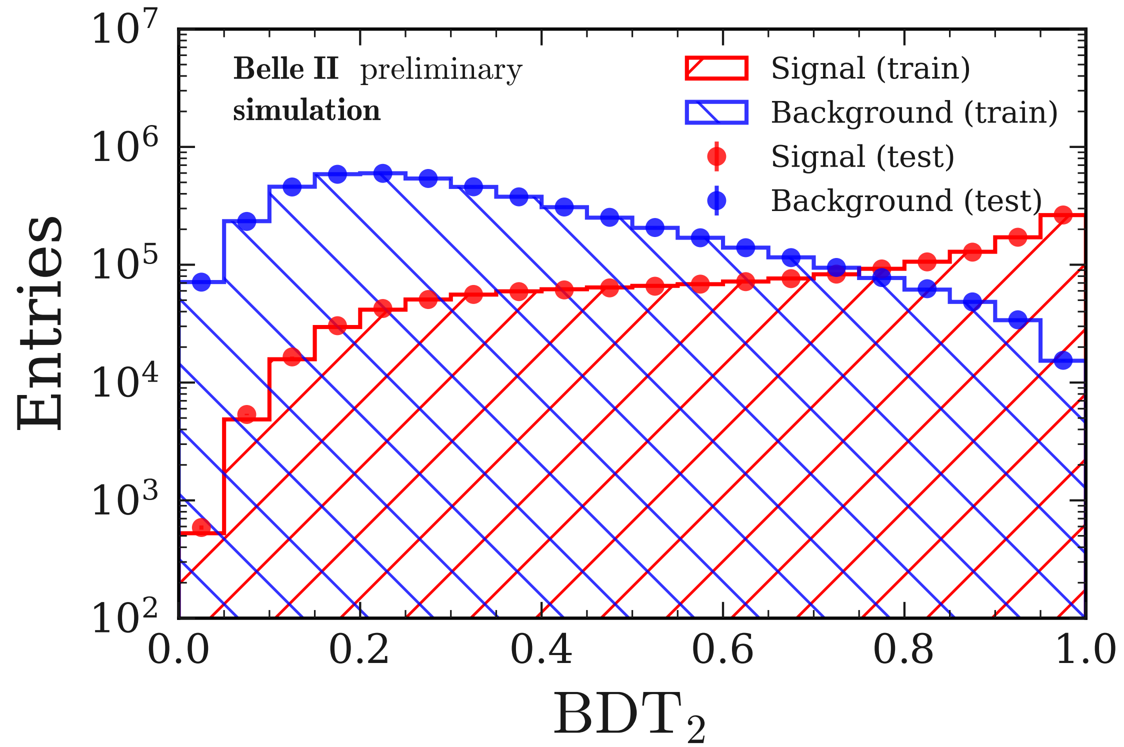
$$\left(\Delta E = \sum_i E_i^{CMS} - E_{beam}\right)$$



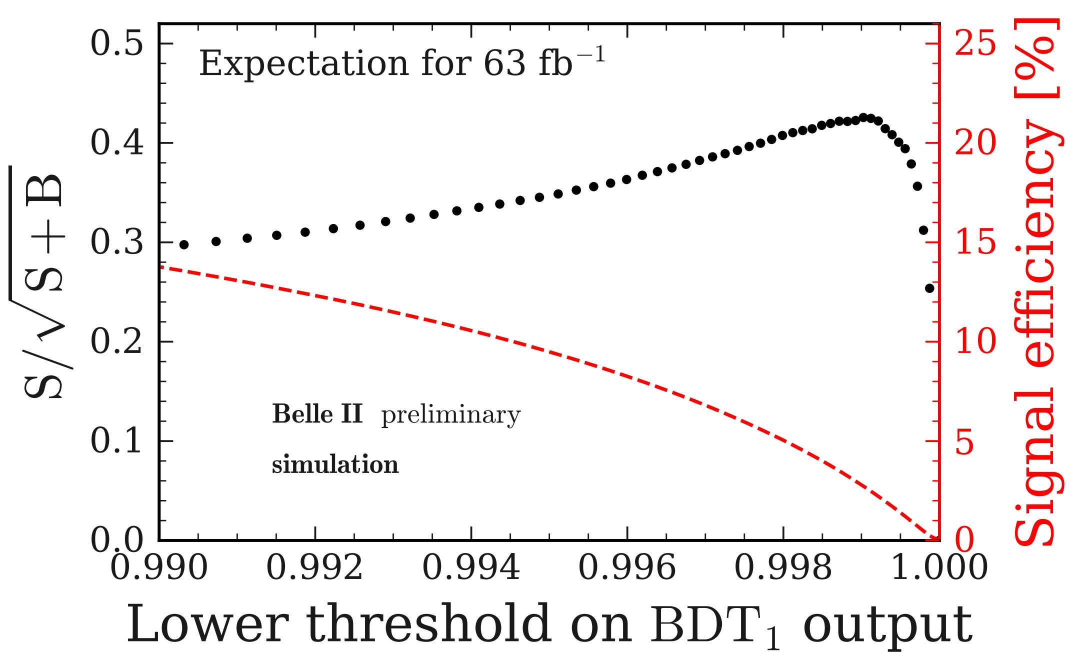
More on multivariate classification

- No overfitting observed neither for BDT1 nor for BDT2.





- Signal sensitivity of BDT1:



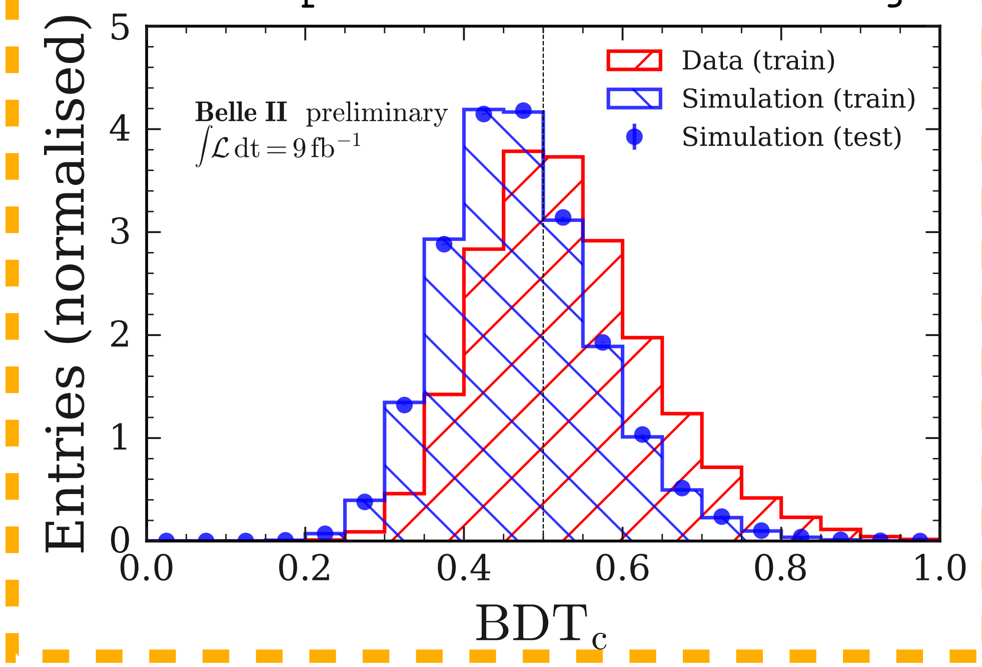
Reweighting of continuum MC

Discrepancies between simulated continuum and off-resonance data.

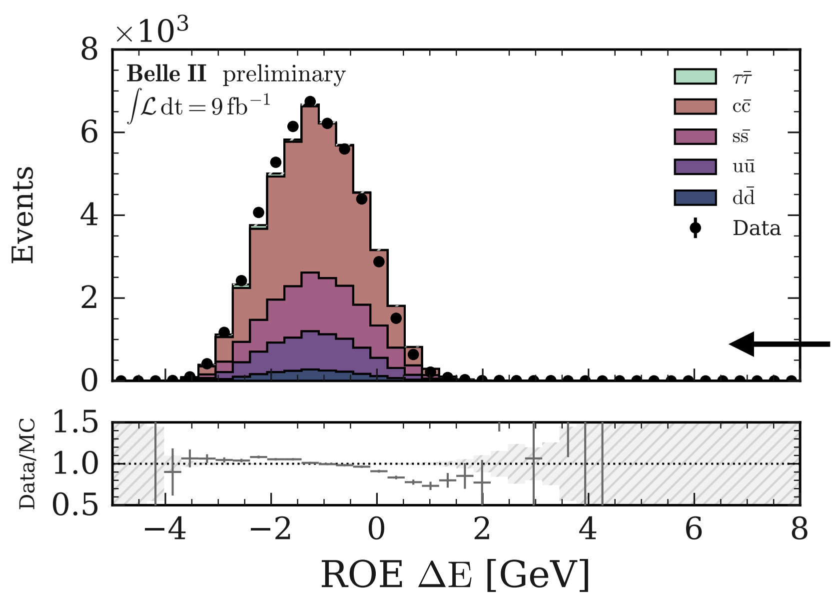
Data-driven correction by means of an additional fastBDT: $\mathrm{BDT_c}$.

- 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.



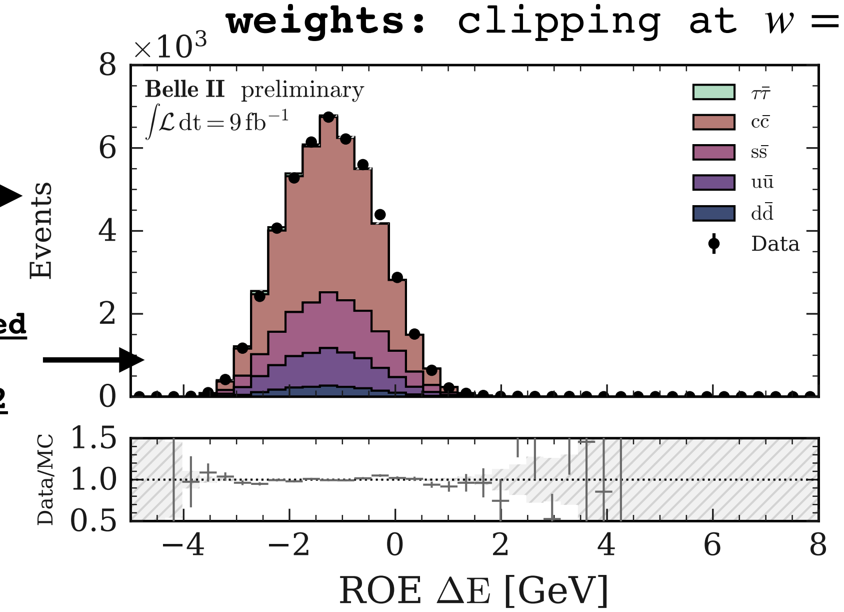
- Being p the $\mathrm{BDT_c}$ score, apply the **event weight** $p/(1-p) = \mathrm{P(Data-like)/P(MC-like)}$ **to correct the simulated continuum.**



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

Continuum MC yields scaled up to Data of normalisation ratio 1.22

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



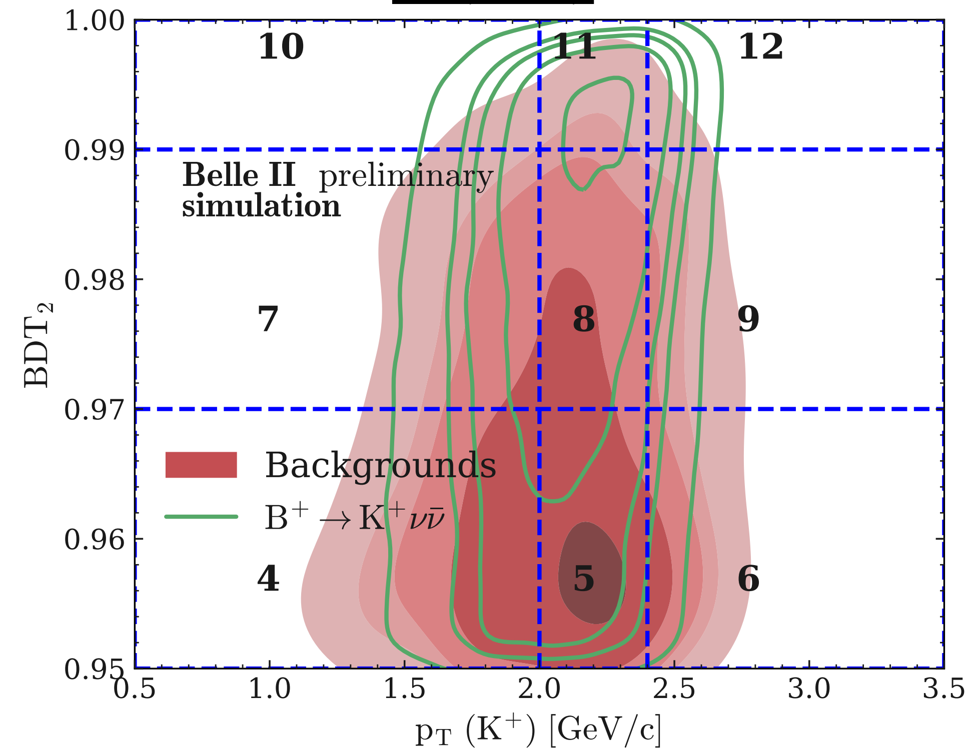
The fit region

- 1 signal region + 3 control regions.

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

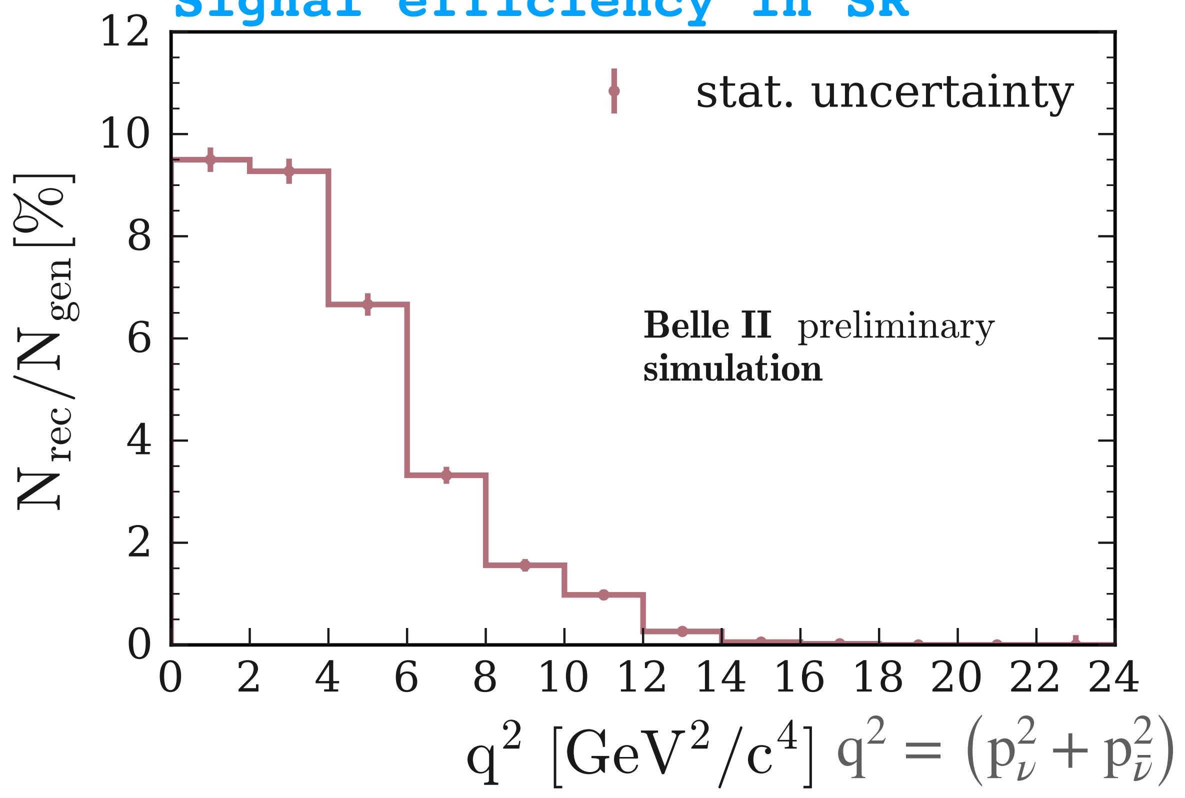
Region	2D Bin Boundary Definition	Physics Processes	$\sqrt{s}$
Signal Region (SR)	$p_T(K^+) \in [0.5, 2.0, 2.4, 3.5]$ GeV/c $BDT_2 \in [0.95, 0.97, 0.99, 1.0]$	signal + all backgrounds	$\Upsilon(4S)$
Control Region 1 (CR1)	$p_T(K^+) \in [0.5, 2.0, 2.4, 3.5]$ GeV/c $BDT_2 \in [0.93, 0.95]$	signal + all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

SR(CR2)

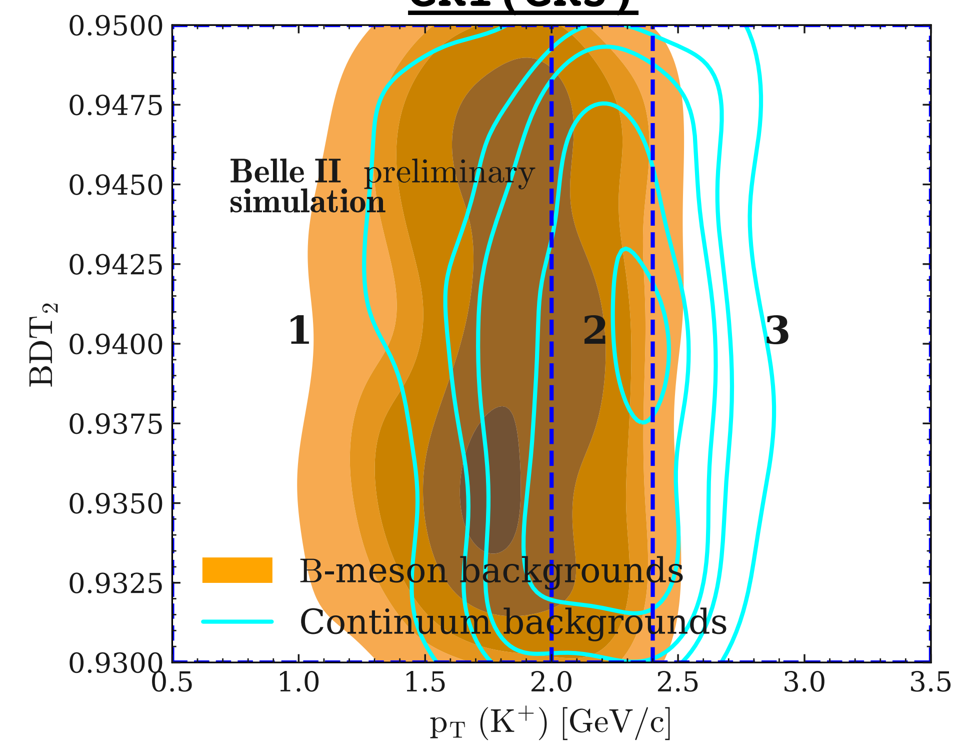


CR 1-2-3 to constrain bkg yields.

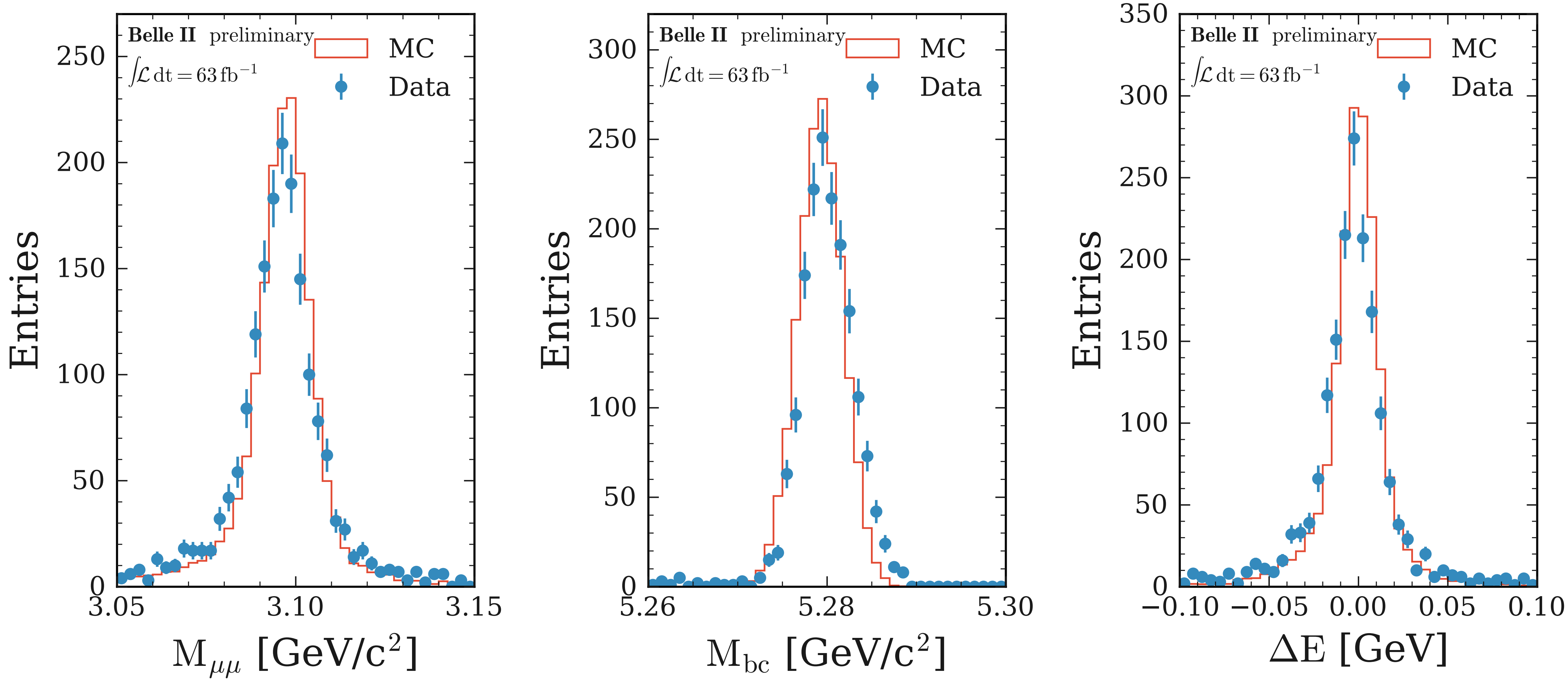
Signal efficiency in SR



CR1(CR3)



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

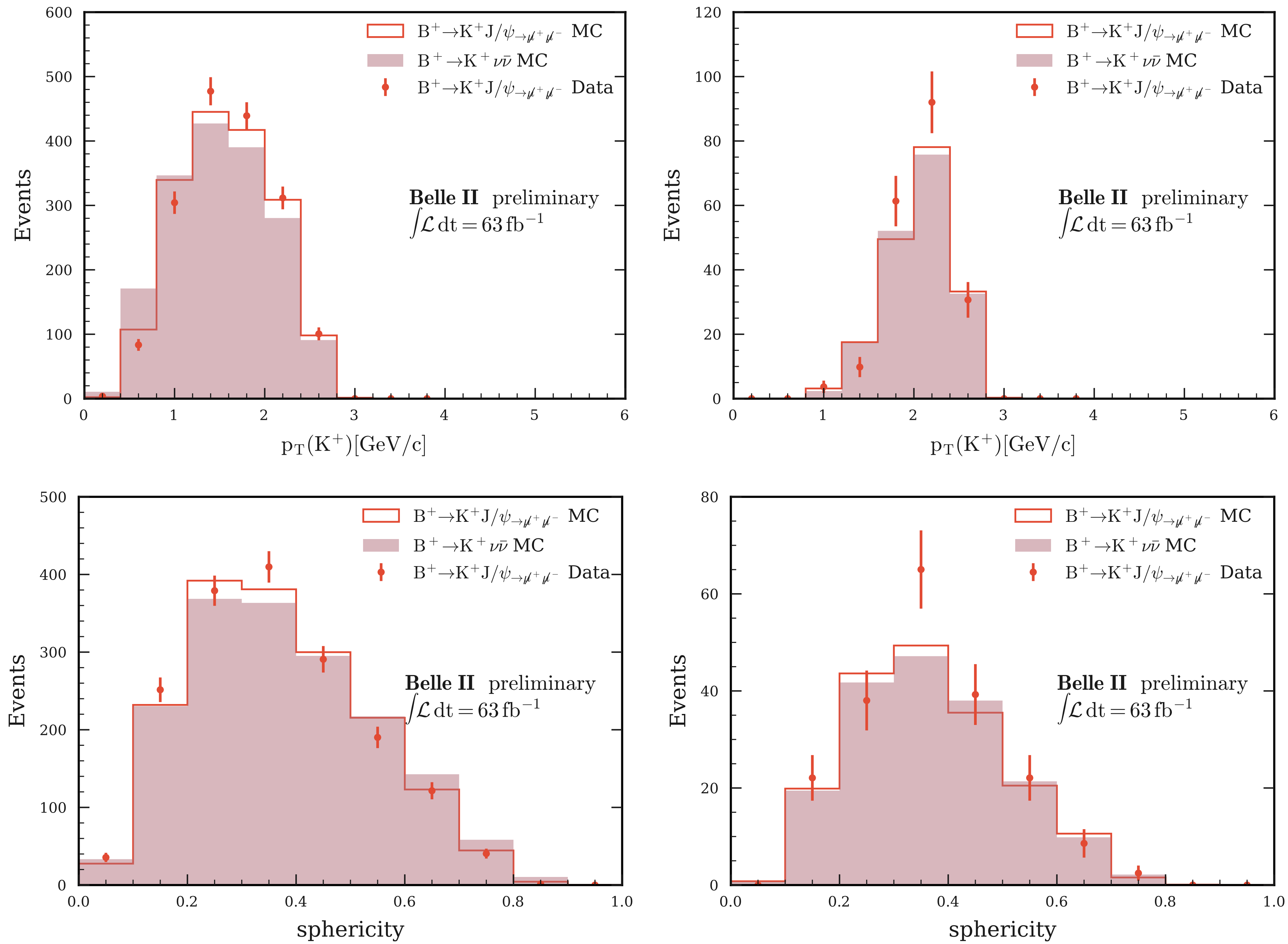


$$M_{bc} = \sqrt{E_{beam}^2 - (\vec{P}_B^{CMS})^2}$$

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

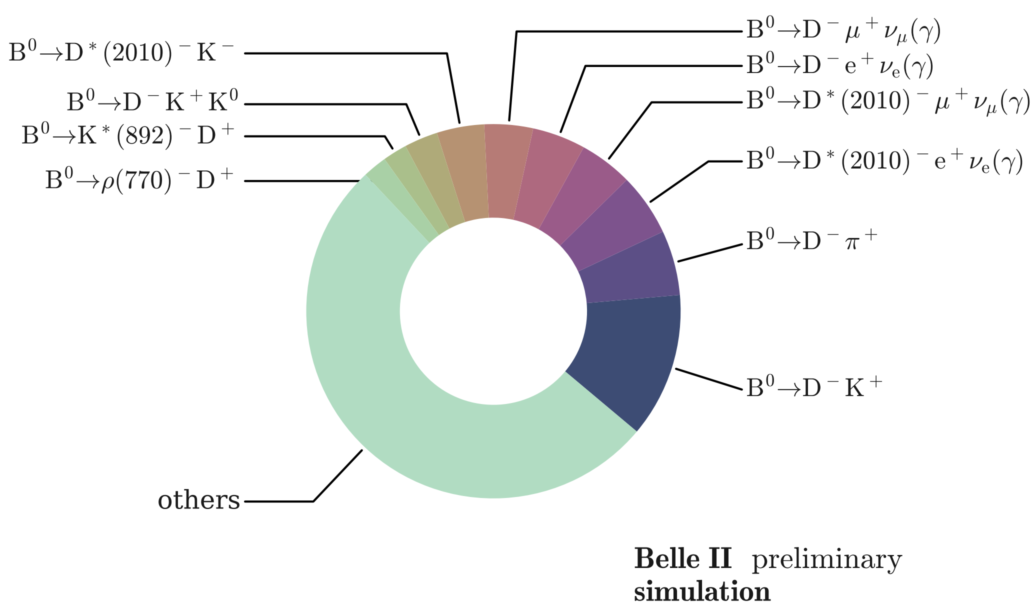
1720 data events from 63 fb^{-1} + bkg suppressed to percent level.

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

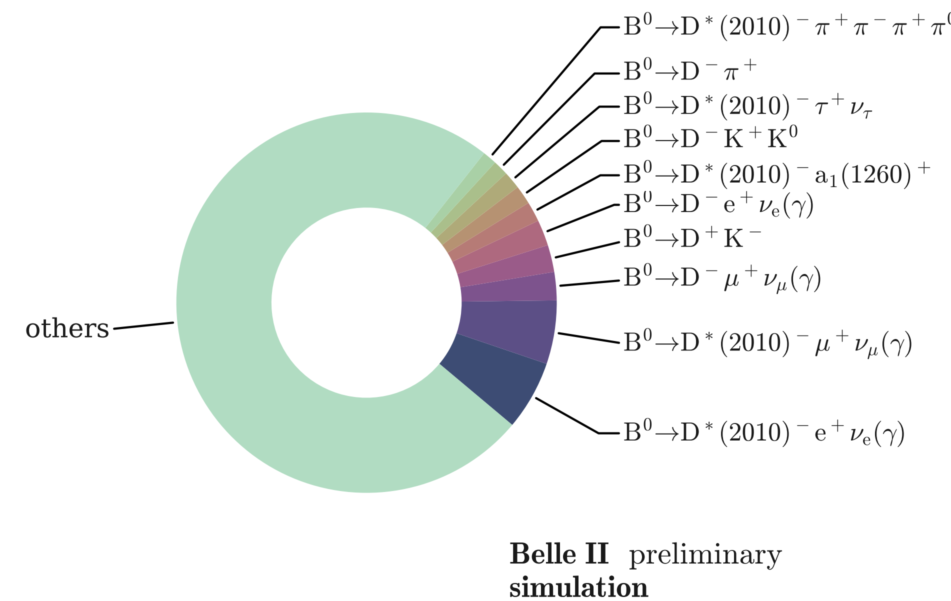


Background composition in the fit region

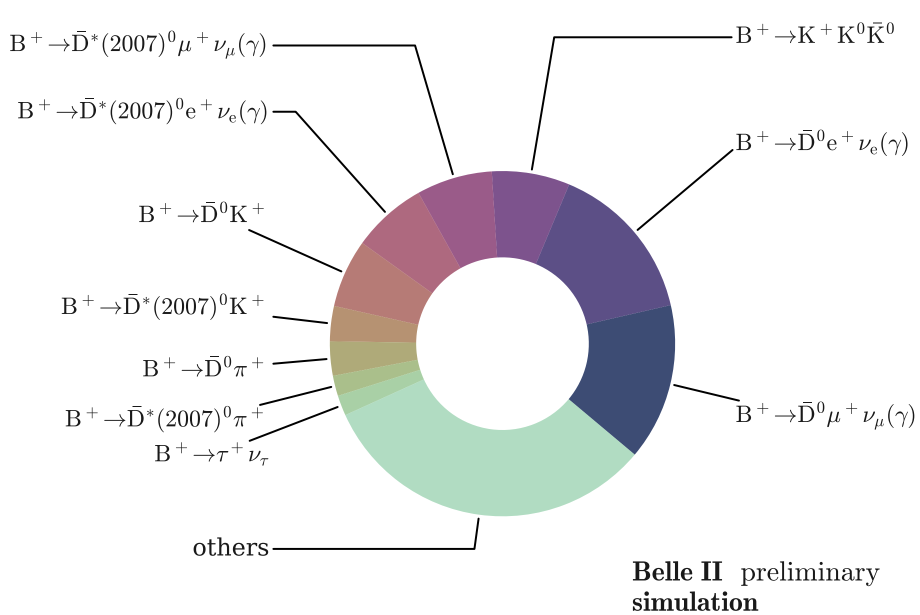
- $B^0\bar{B}^0$ signal side:



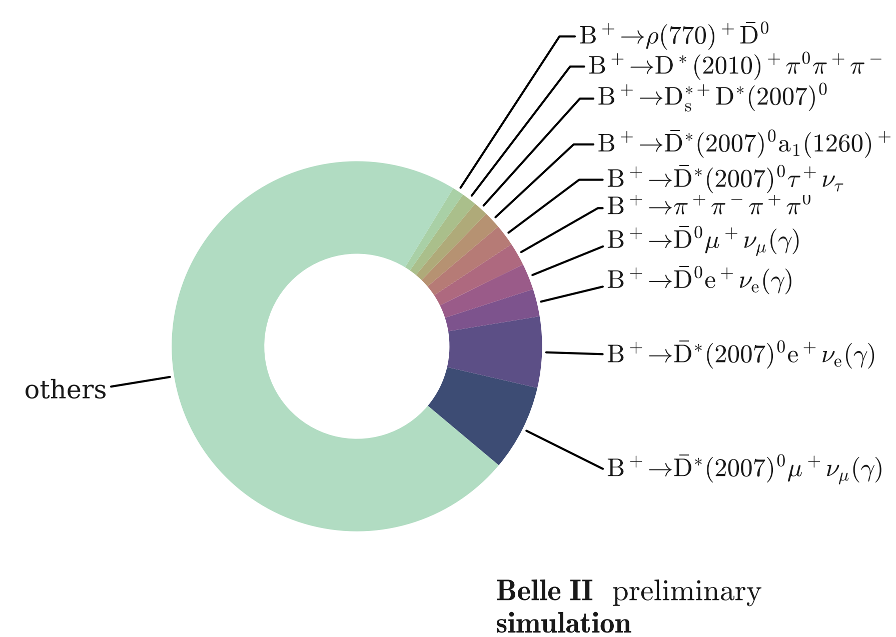
- $B^0\bar{B}^0$ tag side:



- B^+B^- signal side:

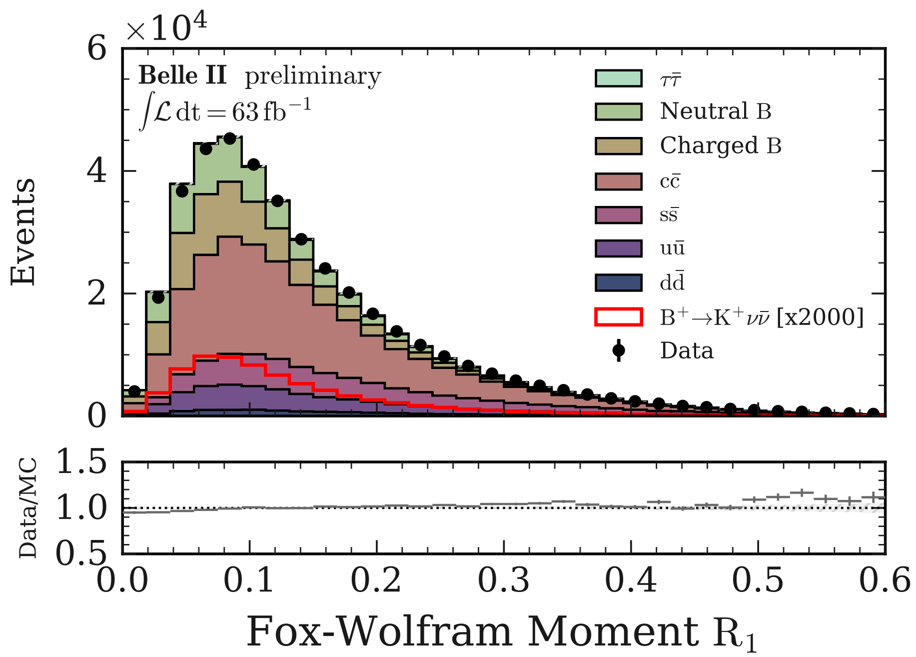


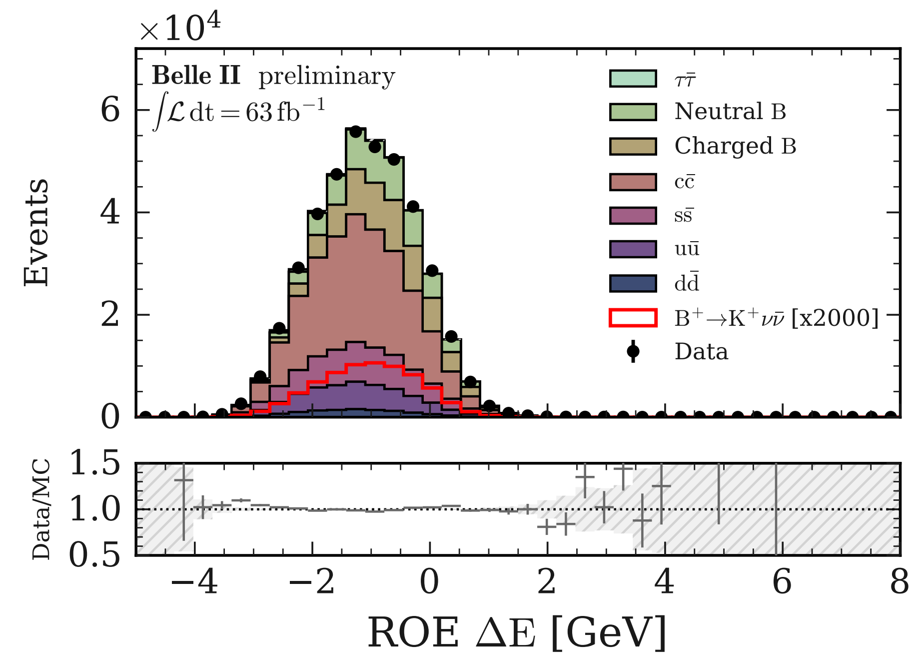
- $B^+\bar{B}^-$ tag side:

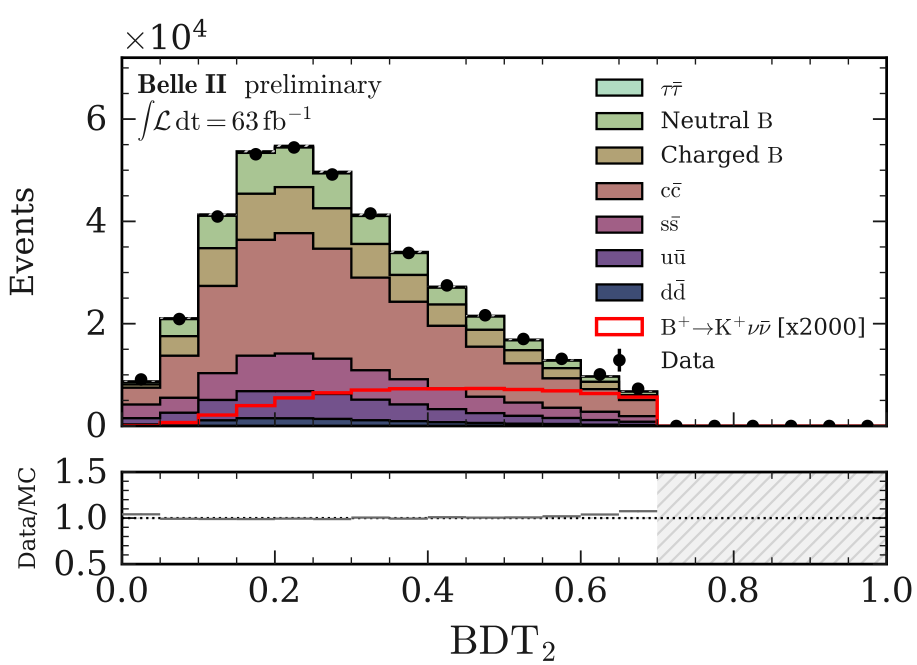


Validation in the BDT sideband

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



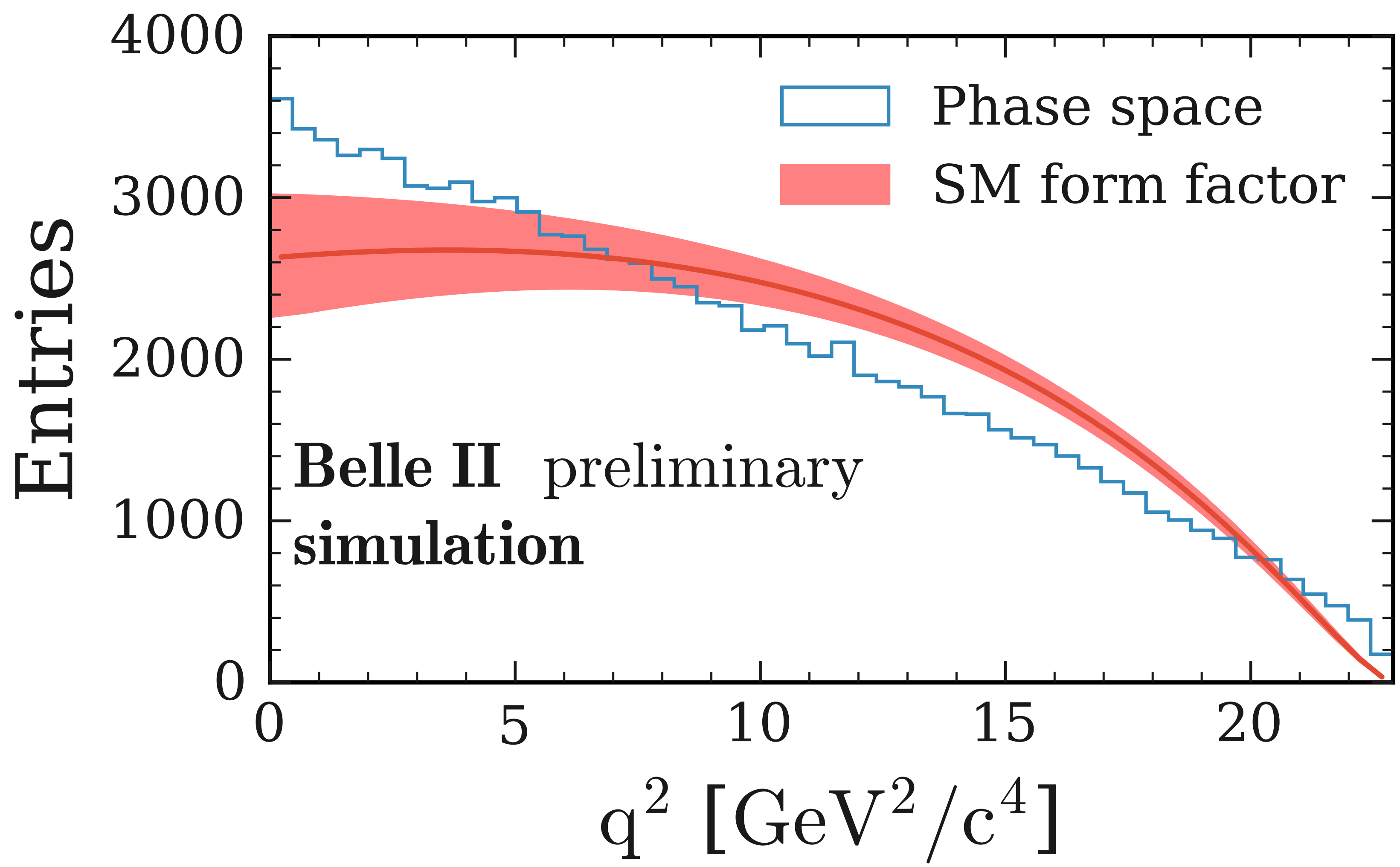




- **Only if the continuum background is scaled by a factor of 1.22 as obtained from the comparison with off-resonance data, the data/MC ratio is then 1.00 in the moderate BDT sideband.**

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:

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

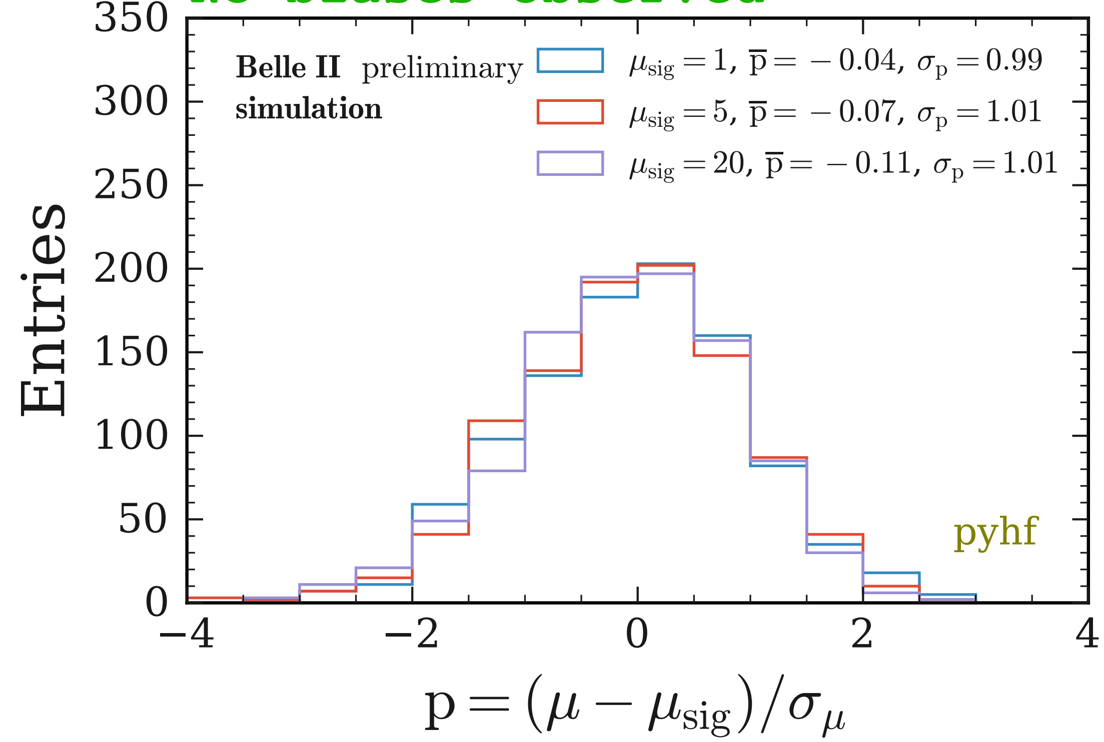
Description	Modification	Constraint Term c_χ	Input
Uncorrelated Shape	$\kappa_{scb}(\gamma_b) = \gamma_b$	$\prod_b \text{Pois}\left(r_b = \sigma_b^{-2} \middle\vert\, \rho_b = \sigma_b^{-2}\gamma_b\right)$	σ_b
Correlated Shape	$\Delta_{scb}(\alpha) = f_p\left(\alpha \middle\vert\, \Delta_{scb,\alpha=-1}, \Delta_{scb,\alpha=1}\right)$	$\text{Gaus}\,(a = 0 \vert\, \alpha, \sigma = 1)$	$\Delta_{scb,\alpha=\pm1}$
Normalisation Unc.	$\kappa_{scb}(\alpha) = g_p\left(\alpha \middle\vert\, \kappa_{scb,\alpha=-1}, \kappa_{scb,\alpha=1}\right)$	$\text{Gaus}\,(a = 0 \vert\, \alpha, \sigma = 1)$	$\kappa_{scb,\alpha=\pm1}$
MC Stat. Uncertainty	$\kappa_{scb}(\gamma_b) = \gamma_b$	$\prod_b \text{Gaus}\left(a_{\gamma_b} = 1 \middle\vert\, \gamma_b, \delta_b\right)$	$\delta_b^2 = \sum_s \delta_{sb}^2$
Luminosity	$\kappa_{scb}(\lambda) = \lambda$	$\text{Gaus}\,(l = \lambda_0 \vert\, \lambda, \sigma_\lambda)$	$\lambda_0, \sigma_\lambda$
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 $\left(\mu_{\mathrm{fit}} - \mu_{\mathrm{inj.}}\right)/\sigma_\mu$

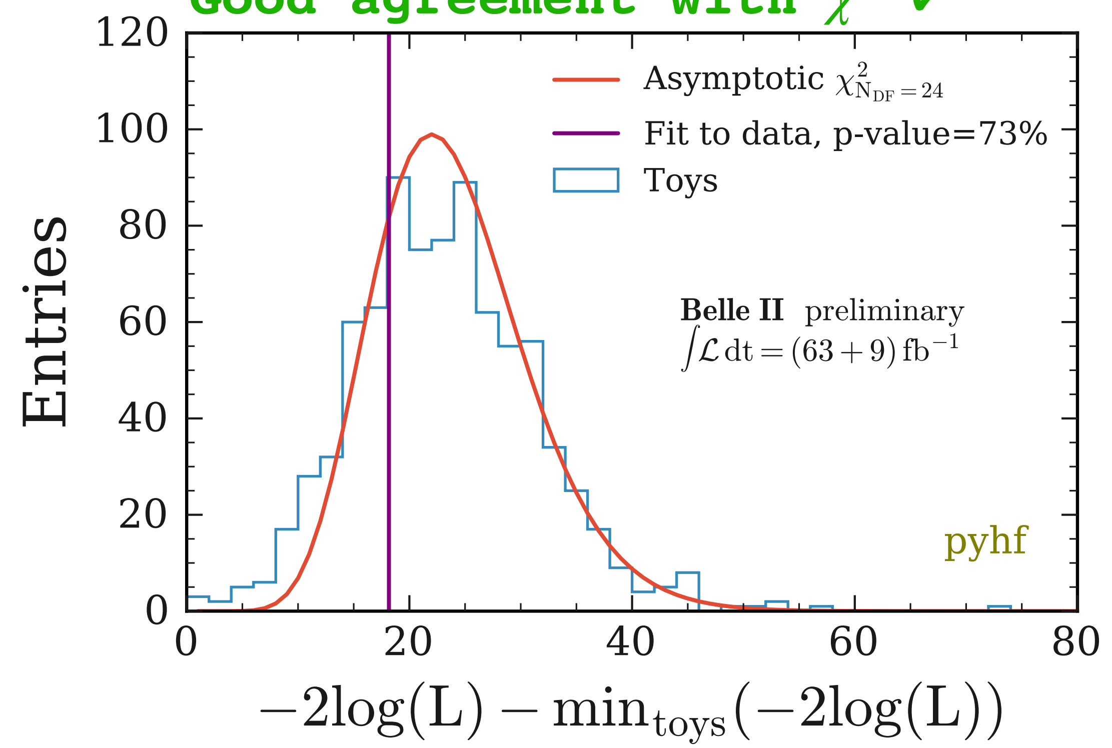
No biases observed ✔



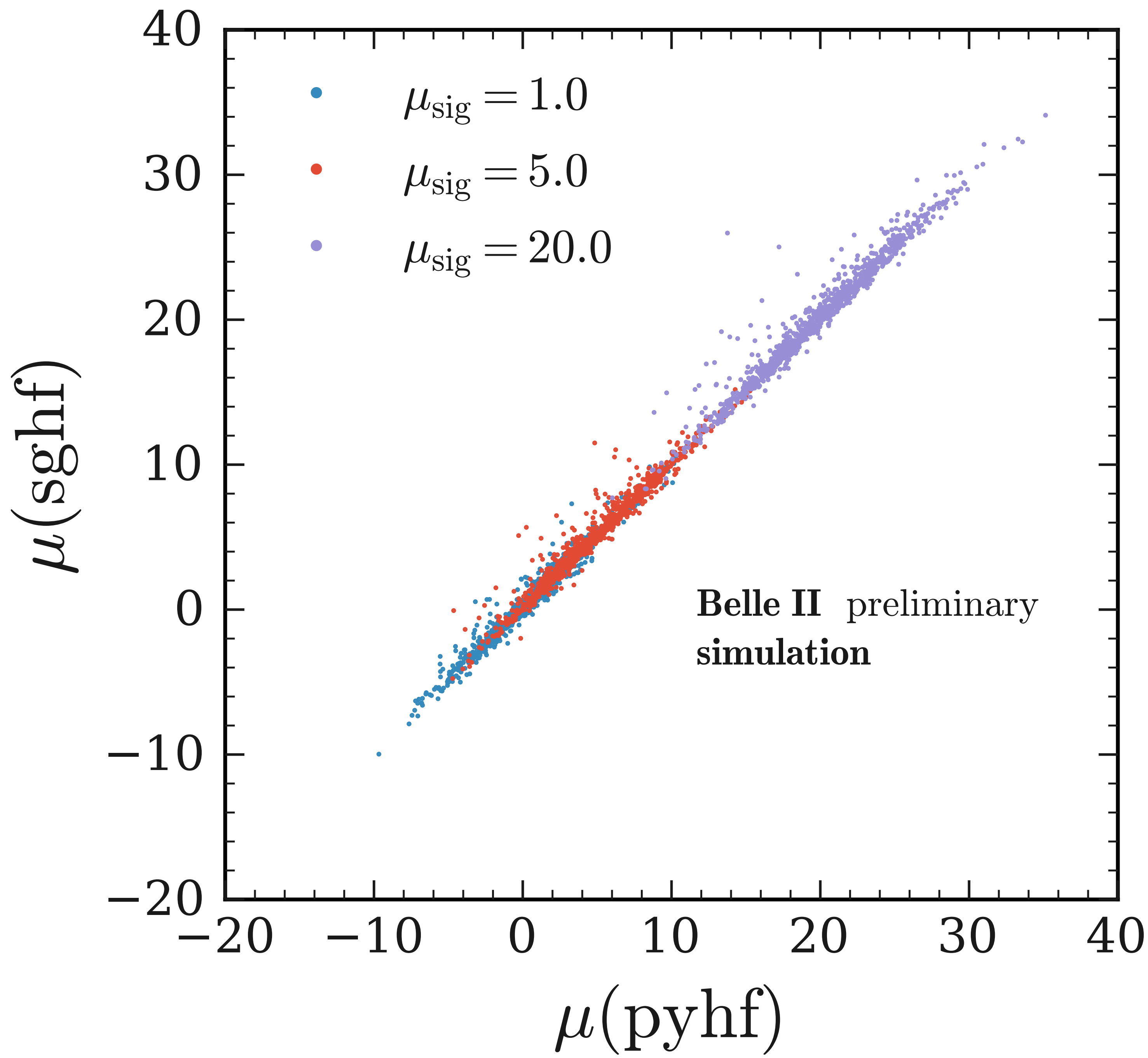
- **Test the fit quality:**

Check the p-value of the fit on observations

High p-value ✔
Good agreement with χ^2 ✔

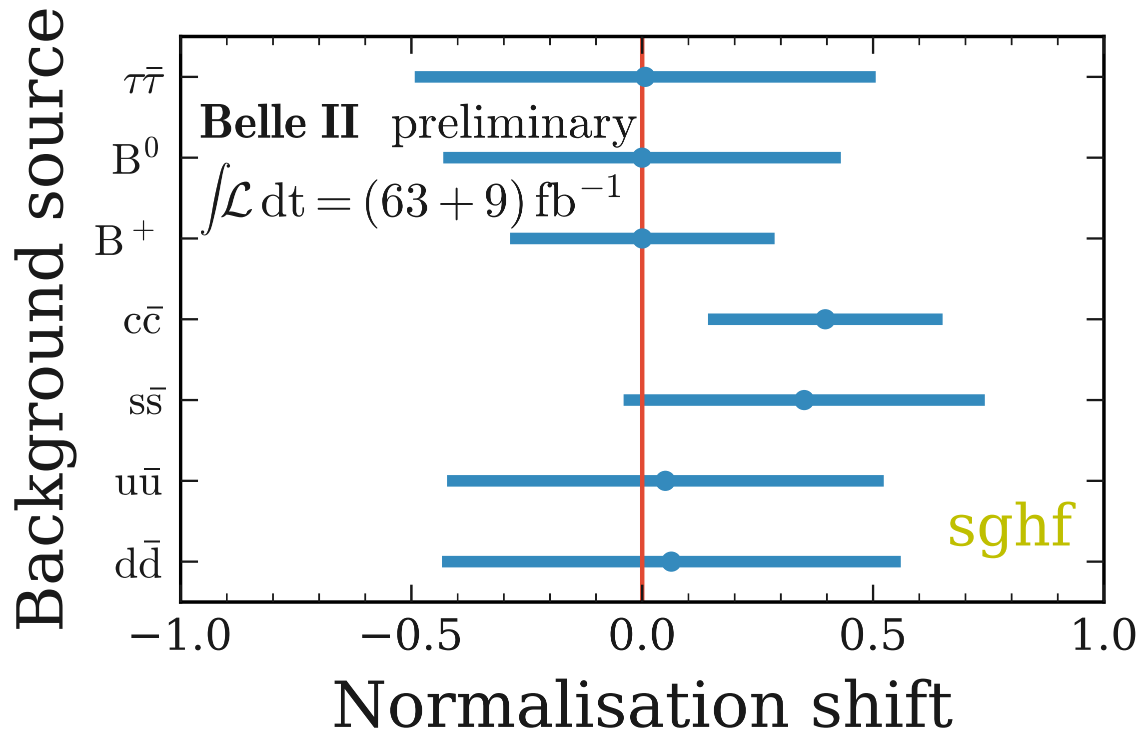


Cross validation of PyHf with a simplified Gaussian model



Fit to the Data

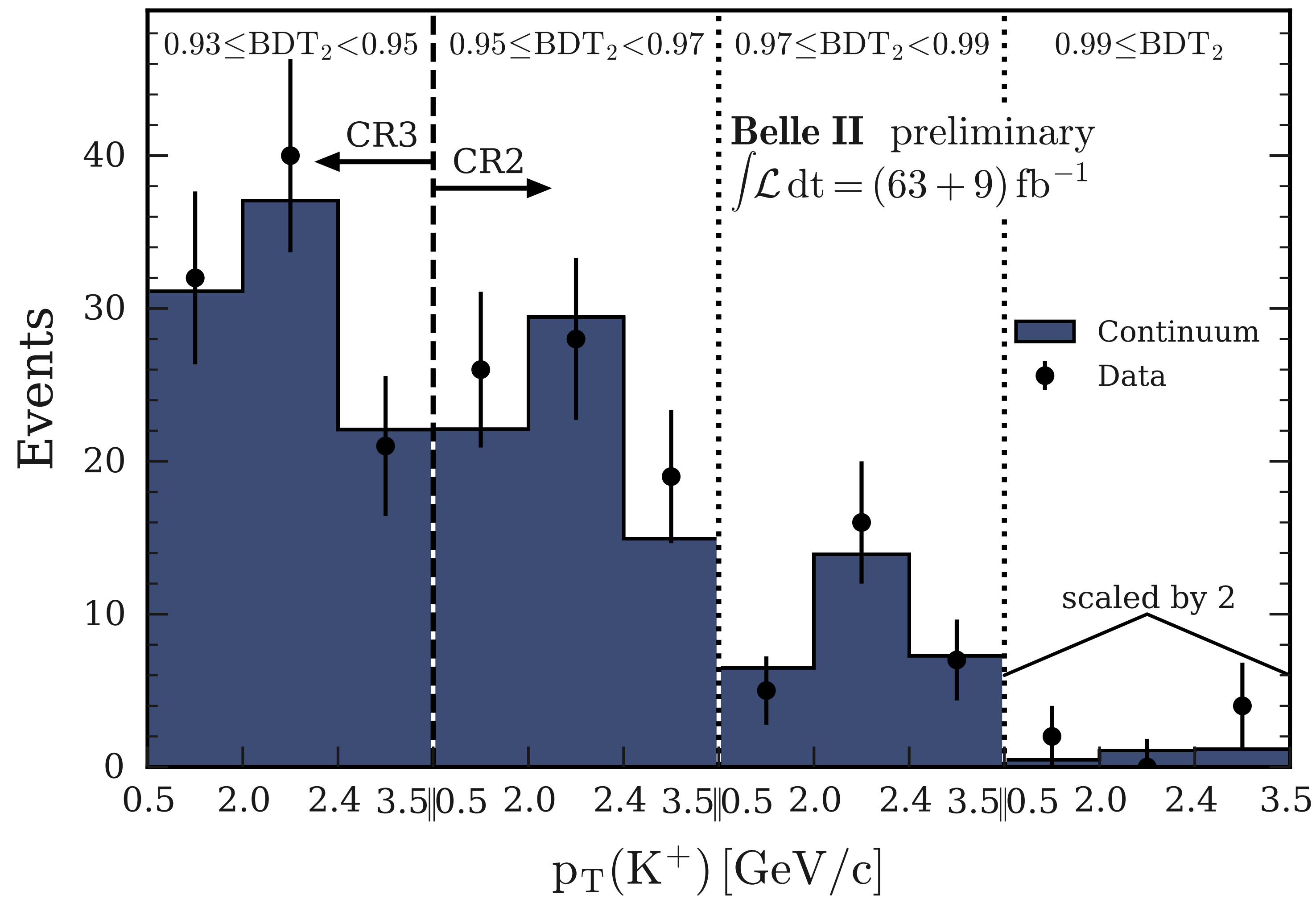
- Post-fit shifts of the bkg's normalisations.



- 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.**

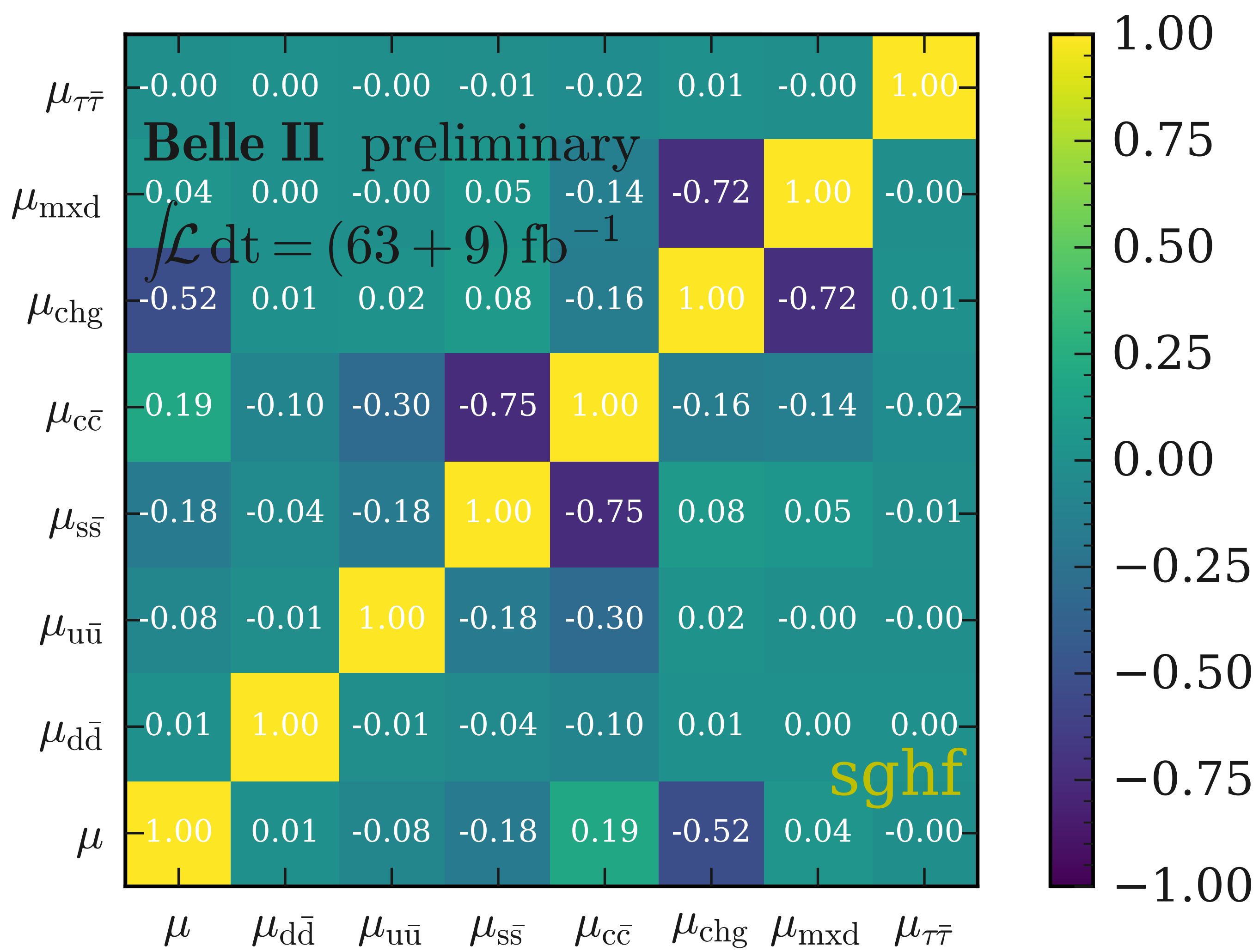
Fit to the Data

- 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