



Search for $B \rightarrow K\nu\bar{\nu}$ decays with a machine learning method at the Belle II experiment

Cyrille Praz

23.08.2022, DESY

Outline

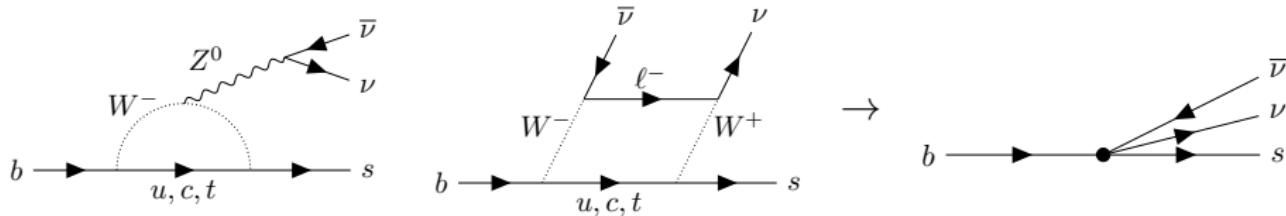
- 1 Theoretical motivation
- 2 The Belle II experiment
- 3 Search for $B^+ \rightarrow K^+ \nu \bar{\nu}$ decays
 - Introduction
 - Binary classification
 - Comparison between data and simulation
 - Result obtained with $(63 + 9) \text{ fb}^{-1}$ of data
- 4 Conclusion and outlook

Introduction

- The Standard Model (SM) describes the known elementary particles and their interactions.
- Studies of B -meson decays are interesting to test the predictions of the SM and potentially detect the presence of new physics effects.
- The decay $B \rightarrow K\nu\bar{\nu}$ is a weak process involving a $b \rightarrow s$ transition.

$B \rightarrow K\nu\bar{\nu}$ in the Standard Model

- $B \rightarrow K\nu\bar{\nu}$ is suppressed in the SM and has never been observed.
- Decay described by an effective Hamiltonian.



- $\mathcal{H}_{\text{eff}}^{\text{SM}} = -\frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* C_L^{\text{SM}} \mathcal{O}_L + \text{h.c.}$
- G_F : Fermi constant.
- V_{ij} : Elements of the Cabibbo-Kobayashi-Maskawa matrix.
- \mathcal{O}_L : effective vertex operator, $\mathcal{O}_L = \frac{e^2}{16\pi^2} (\bar{s}_L \gamma_\mu b_L)(\bar{\nu}_L \gamma^\mu \nu_L)$.
- C_L^{SM} : effective coupling constant (Wilson coefficient).

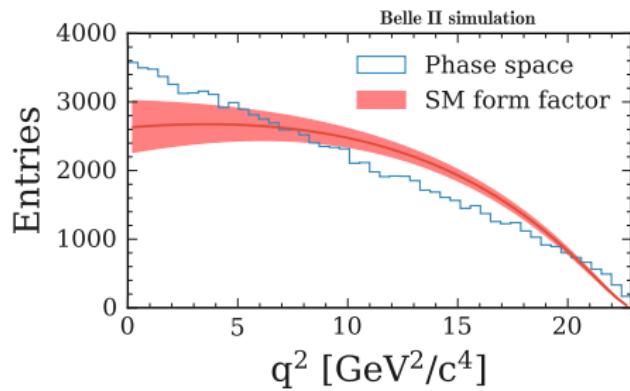
Branching fraction in the Standard Model

- $$\frac{d\text{Br}(B \rightarrow K\nu\bar{\nu})_{\text{SM}}}{dq^2} = 3 \tau_B \left| \frac{G_F \alpha}{16\pi^2} \sqrt{\frac{m_B^3}{3\pi}} V_{tb} V_{ts}^* C_L^{\text{SM}} f_+(q^2) \right|^2 \left(\frac{\lambda_K(q^2)}{m_B^4} \right)^{\frac{3}{2}}$$

- q^2 : Squared invariant mass of the two-neutrino system.
- τ_B, m_B : Lifetime and mass of the B meson.
- α : Electromagnetic coupling.
- $f_+(q^2)$: Hadronic form factor (q^2 -dependence of $\langle K | \bar{s}\gamma_\mu b | B \rangle$).
- $\lambda_K(q^2)$: Phase-space factor.

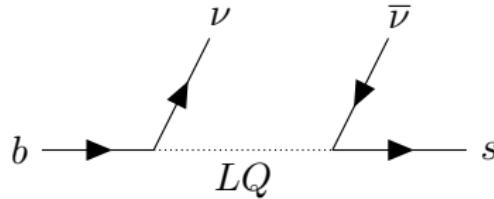
Branching fraction in the Standard Model

- $\text{Br}(B^+ \rightarrow K^+ \nu \bar{\nu})_{\text{SM}} = (4.6 \pm 0.5) \times 10^{-6}$. [Prog. Part. Nucl. Phys. 92, 50 (2017)]
- $\text{Br}(B^0 \rightarrow K^0 \nu \bar{\nu})_{\text{SM}} = (4.3 \pm 0.5) \times 10^{-6}$.
- 10% theoretical uncertainty mainly from $B \rightarrow K$ form factor.
- $B \rightarrow K$ form factor used for signal simulation. [J. High Energy Phys. 02, 184 (2015)]



New Physics (NP)

- Since ν_e , ν_μ and ν_τ contribute, $B \rightarrow K\nu\bar{\nu}$ is sensitive to potential lepton flavor universality violation.
- Complementary probe of NP scenarios proposed to explain anomalies observed in $b \rightarrow s\ell^+\ell^-$ transitions. [Phys. Lett. B 809, 135769 (2020)]
- Multiple models beyond the SM constrained by $\text{BR}(B \rightarrow K\nu\bar{\nu})$:
 - axions (dark matter candidate). [Phys. Rev. D 102, 015023 (2020)]
 - leptoquarks. [Phys. Rev. D 98, 055003 (2018)]



Outline

1 Theoretical motivation

2 The Belle II experiment

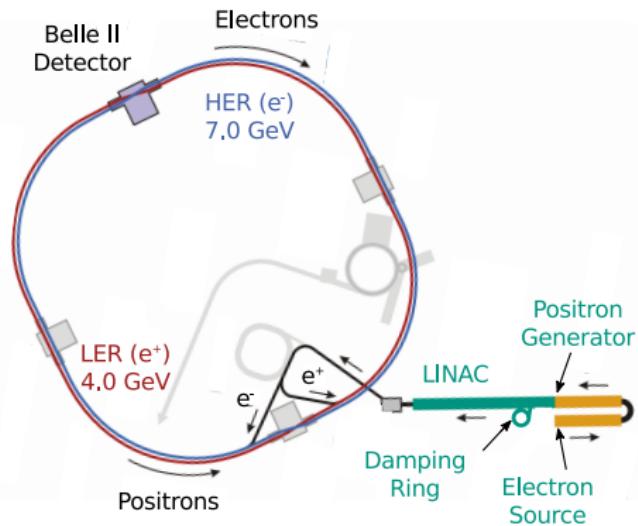
3 Search for $B^+ \rightarrow K^+ \nu \bar{\nu}$ decays

- Introduction
- Binary classification
- Comparison between data and simulation
- Result obtained with $(63 + 9) \text{ fb}^{-1}$ of data

4 Conclusion and outlook

The SuperKEKB accelerator

- e^+e^- collider in Tsukuba, Japan.
- $\sqrt{s} = 10.6 \text{ GeV} = m(\Upsilon(4S))$.
- $\text{BR}(\Upsilon(4S) \rightarrow B\bar{B}) > 96\%$.
- $\int_{25.03.2019}^{22.06.2022} L dt \approx 363 \text{ fb}^{-1}$.
 - This analysis: $189/63 \text{ fb}^{-1}$.
- Long-term target:
 - $\int L dt = 50 \text{ ab}^{-1}$.

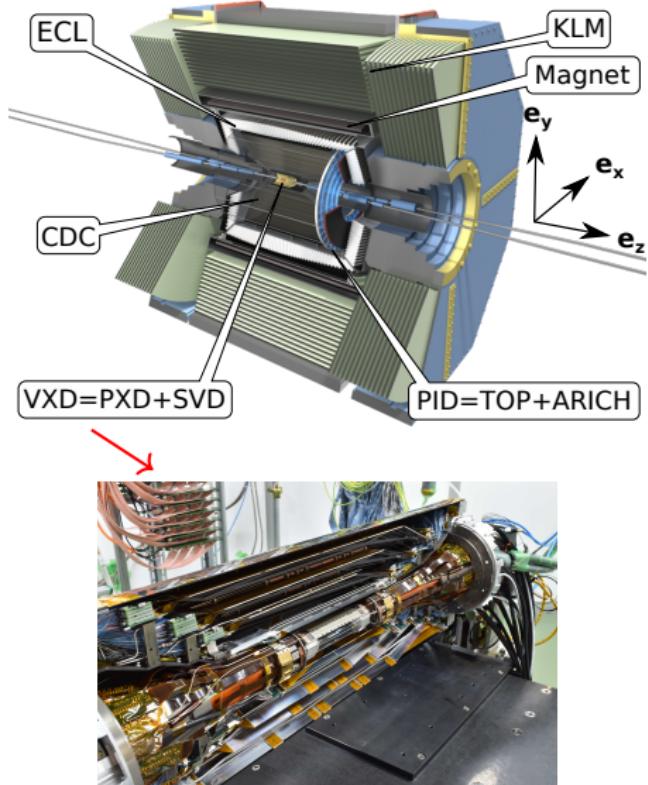


Main e^+e^- processes at $\sqrt{s} = 10.6 \text{ GeV}$

Process	Cross-section [nb]
$e^+e^- \rightarrow \gamma(4S)$	1.10
$e^+e^- \rightarrow u\bar{u}(\gamma)$	1.61
$e^+e^- \rightarrow d\bar{d}(\gamma)$	0.40
$e^+e^- \rightarrow s\bar{s}(\gamma)$	0.38
$e^+e^- \rightarrow c\bar{c}(\gamma)$	1.30
$e^+e^- \rightarrow \tau^+\tau^-(\gamma)$	0.92
$e^+e^- \rightarrow e^+e^-(\gamma)$	300
$e^+e^- \rightarrow e^+e^-e^+e^-$	39.7
$e^+e^- \rightarrow e^+e^-\mu^+\mu^-$	18.9
$e^+e^- \rightarrow \gamma\gamma(\gamma)$	4.99
$e^+e^- \rightarrow \mu^+\mu^-(\gamma)$	1.15

The Belle II detector

- Pixel Detector (PXD).
- Silicon Vertex Detector (SVD).
- Central Drift Chamber (CDC).
- Calorimeter (ECL).
- Aerogel Ring-Imaging Cherenkov (ARICH).
- Time-Of-Propagation (TOP) counter.
- K_L^0 and μ detection (KLM).

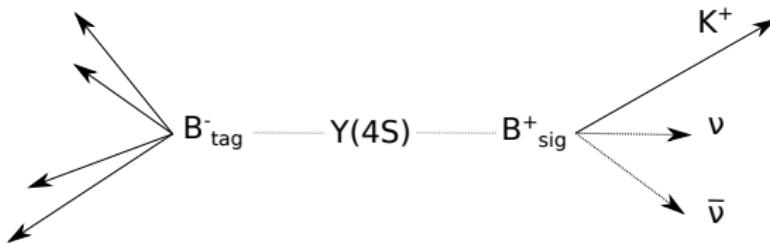


Outline

- 1 Theoretical motivation
- 2 The Belle II experiment
- 3 Search for $B^+ \rightarrow K^+ \nu \bar{\nu}$ decays
 - Introduction
 - Binary classification
 - Comparison between data and simulation
 - Result obtained with $(63 + 9) \text{ fb}^{-1}$ of data
- 4 Conclusion and outlook

B -meson tagging

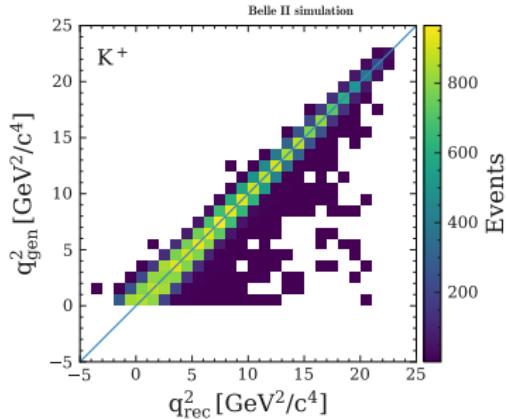
- Previous searches used tagging methods, where the second B -meson (B_{tag}) is reconstructed...
 - ...in a hadronic decay: $\varepsilon_{\text{sig}} = \mathcal{O}(0.04\%)$. [Phys. Rev. D 87, 112005 (2013) (Babar)]
 - ...in a semileptonic decay: $\varepsilon_{\text{sig}} = \mathcal{O}(0.2\%)$. [Phys. Rev. D 96, 091101 (2017) (Belle)]



- In the following, an inclusive tagging method is used.
 - No explicit reconstruction of the second B -meson.
 - Exploitation of distinctive topological properties of $B^+ \rightarrow K^+ \nu \bar{\nu}$.

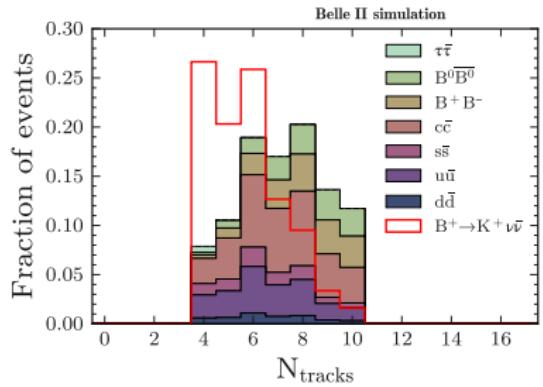
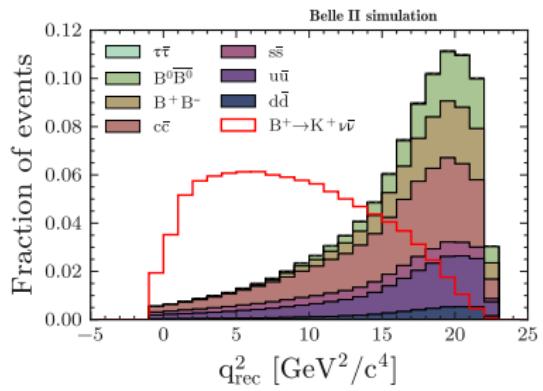
Reconstructed invariant mass of the two-neutrino system

- Invariant mass of the two-neutrino system not accessible.
 - $q^2 = (\mathbf{P}_B^* - \mathbf{P}_K^*)^2 = m_B^2 + m_K^2 - 2 E_B^* E_K^* + 2 \mathbf{p}_B^* \cdot \mathbf{p}_K^*$.
- Define a reconstructed invariant mass (approximation: $\mathbf{p}_B^* = \mathbf{0}$):
 - $q_{\text{rec}}^2 = m_B^2 + m_K^2 - 2 m_B E_K^*$.



Signal candidate selection and event pre-selection

- B^+ candidate with smallest q_{rec}^2 selected as signal candidate.
 - Correct candidate in >90% of the cases.
 - PID requirement to suppress pion background.
- Loose event pre-selection:
 - $4 \leq N_{\text{tracks}} \leq 10$.
 - $\theta(\mathbf{p}_{\text{miss}}) \in \text{CDC acceptance}$.
 - $E_{\text{visible}} > 4 \text{ GeV}$.
 - $q_{\text{rec}}^2 > -1 \text{ GeV}^2/\text{c}^4$.

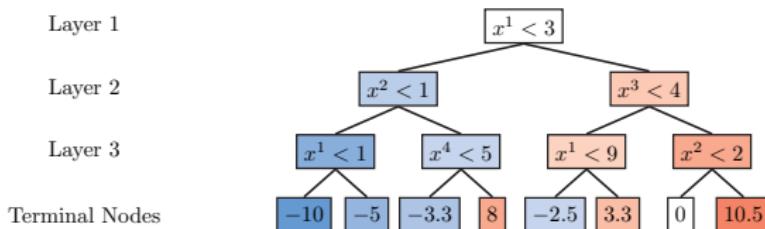


Outline

- 1 Theoretical motivation
- 2 The Belle II experiment
- 3 Search for $B^+ \rightarrow K^+ \nu \bar{\nu}$ decays
 - Introduction
 - **Binary classification**
 - Comparison between data and simulation
 - Result obtained with $(63 + 9) \text{ fb}^{-1}$ of data
- 4 Conclusion and outlook

Binary classification with boosted decision trees

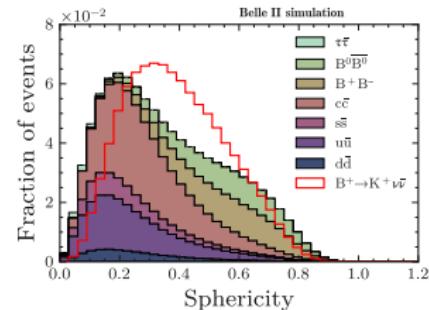
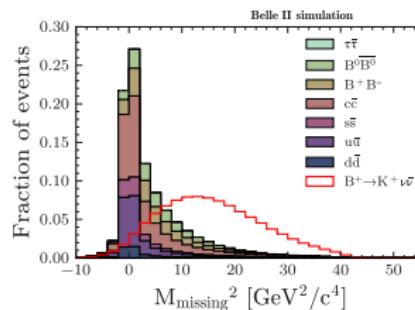
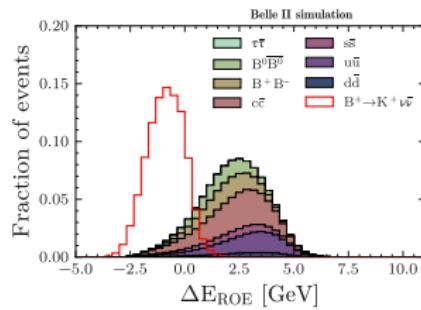
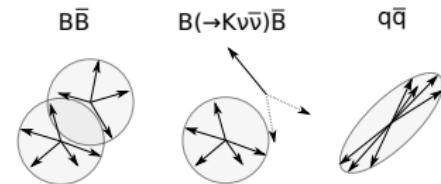
- Selecting $B^+ \rightarrow K^+ \nu \bar{\nu}$ decays is challenging.
 - Small branching fraction.
 - Neutrinos not detected.
- Boosted decision tree (BDT) for background suppression.
 - Trained with simulated signal and background events on a set of discriminative variables.
 - Combination of simple decision trees.



Variables

- Example of discriminative variables:

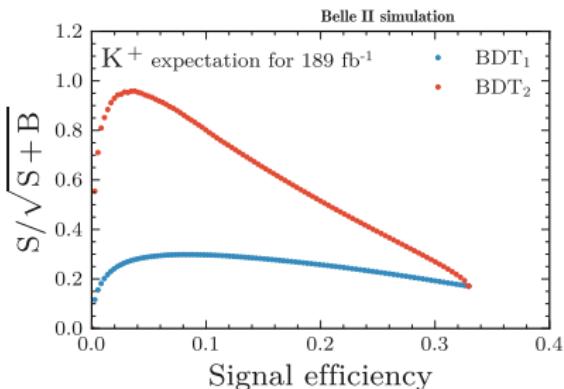
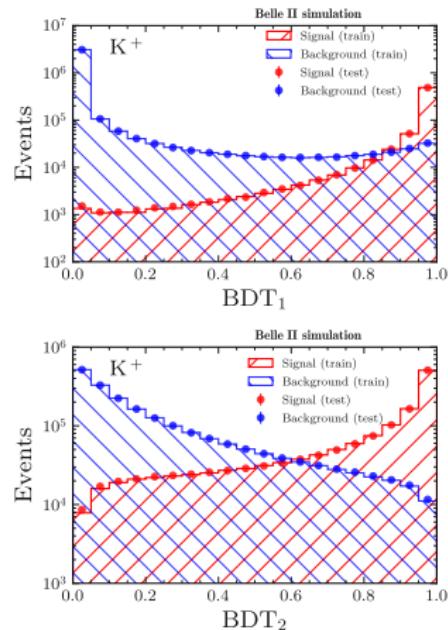
- Rest-of-the-event (ROE) variables.
- Missing energy, momentum.
- Event topology.



$$\Delta E_{\text{ROE}} \equiv E_{\text{ROE}} - \frac{\sqrt{s}}{2}$$

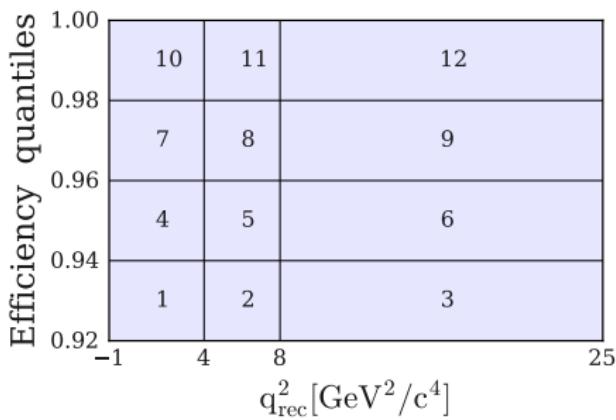
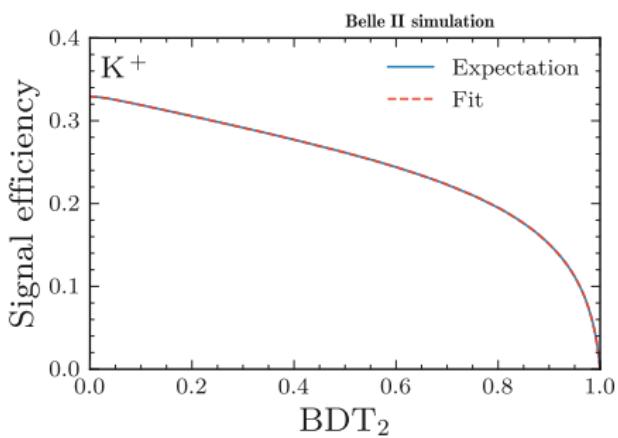
Event selection with two BDTs in series

- Train BDT₁ with 12 variables.
 - BDT₁ > 0.9 \implies signal retention of 85% / background rejection of 98%.
- Train BDT₂ with 35 variables on simulated events with BDT₁ > 0.9.



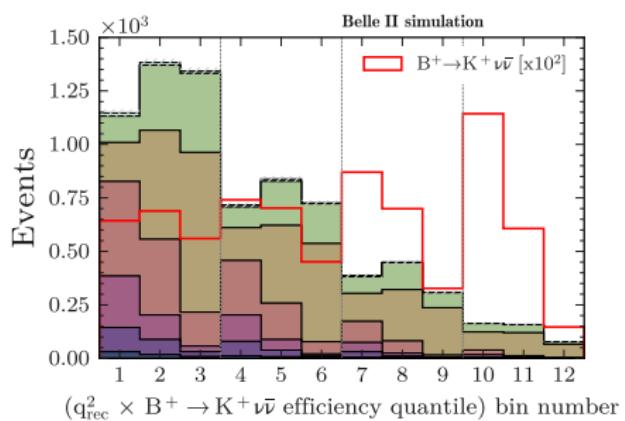
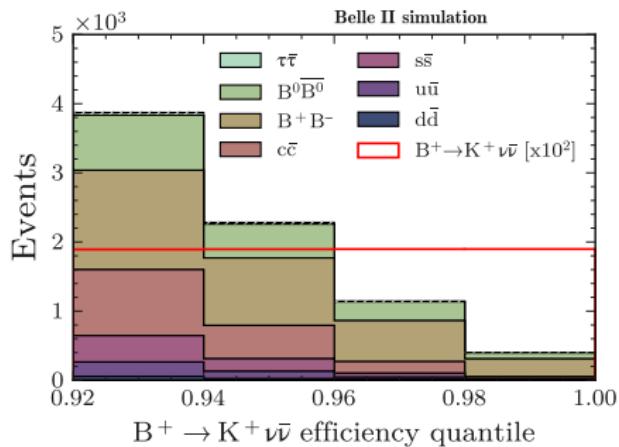
Signal search region

- Translate BDT_2 selection into signal efficiency quantile:
 - $\tilde{\varepsilon}_{\text{sig}} \equiv 1 - \varepsilon_{\text{sig}}(\text{BDT}_2)$.
- Define signal search region as $\tilde{\varepsilon}_{\text{sig}} > 0.92$.
 - 12 bins in the $\tilde{\varepsilon}_{\text{sig}} \times q_{\text{rec}}^2$ space.



Expected yields in the signal search region (189 fb^{-1})

- In signal search region, $e^+ e^- \rightarrow B^+ B^-$ background is dominant.
 - Background decay contributing the most: $B \rightarrow D^{(*)} \ell \nu$, with $\ell = e, \mu$.



Outline

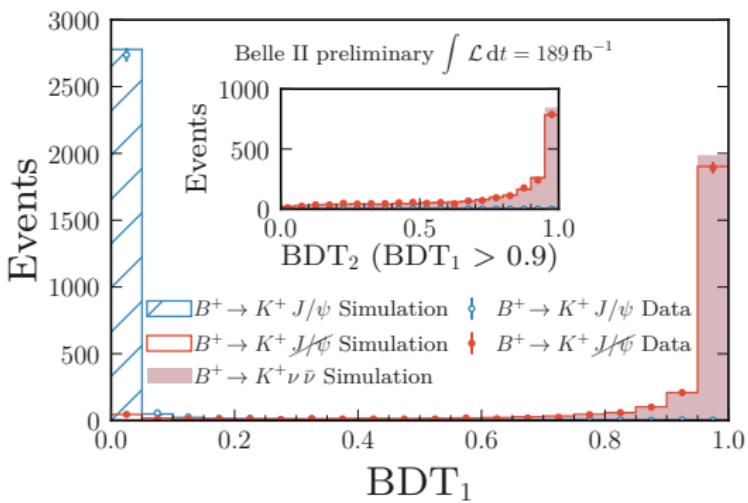
- 1 Theoretical motivation
- 2 The Belle II experiment
- 3 Search for $B^+ \rightarrow K^+ \nu \bar{\nu}$ decays
 - Introduction
 - Binary classification
 - Comparison between data and simulation
 - Validation channel
 - Off-resonance data and simulation
 - On-resonance data and simulation
 - Result obtained with $(63 + 9) \text{ fb}^{-1}$ of data
- 4 Conclusion and outlook

Data samples

- Data samples used for the comparison between data and simulation:
 - 189 fb^{-1} collected at $\sqrt{s} = m(\Upsilon(4S))$ ("on-resonance").
 - 18 fb^{-1} collected at $\sqrt{s} = m(\Upsilon(4S)) - 60 \text{ MeV}$ ("off-resonance").

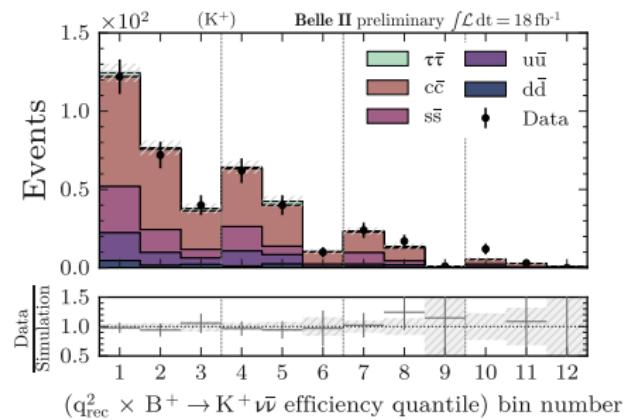
Validation channel: $B^+ \rightarrow K^+ J/\psi (\rightarrow \mu^+ \mu^-)$

- ① Select $B^+ \rightarrow K^+ J/\psi (\rightarrow \mu^+ \mu^-)$ decays in data and simulation.
- ② Remove the muons from the event and replace the kaon by a kaon sampled from simulated signal events.
- ③ Recalculate the classifier input variables for the modified events.
- ④ Examine the output of BDT_1 and BDT_2 .



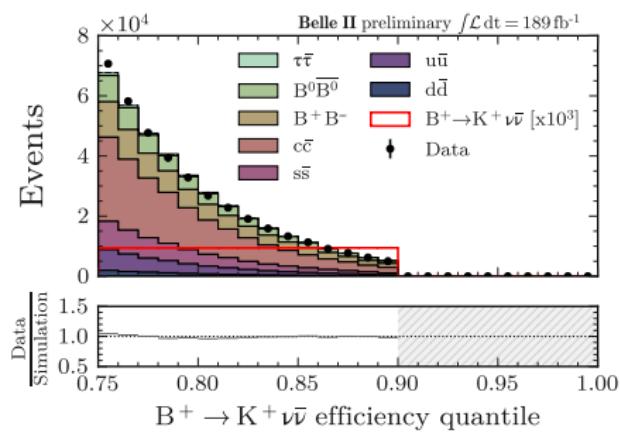
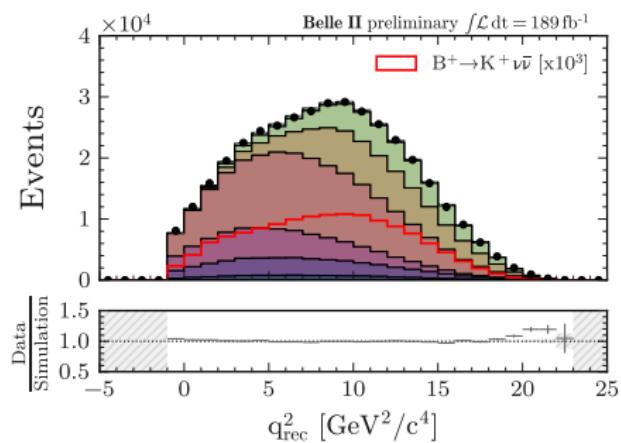
Off-resonance data and simulation

- Off-resonance data shape well modelled in the signal search region.
- Normalisation discrepancy factor $\frac{\text{data}}{\text{simulation}} = 1.5$ is observed.
 - Corrected on the figure, included as a systematic uncertainty.



On-resonance data and simulation

- On-resonance data and simulation in the $0.75 < \tilde{\varepsilon}_{\text{sig}} < 0.90$ sideband.
- Continuum simulation corrected for the normalisation discrepancy observed with off-resonance data.



Outline

- 1 Theoretical motivation
- 2 The Belle II experiment
- 3 Search for $B^+ \rightarrow K^+ \nu \bar{\nu}$ decays
 - Introduction
 - Binary classification
 - Comparison between data and simulation
 - Result obtained with $(63 + 9) \text{ fb}^{-1}$ of data
- 4 Conclusion and outlook

Data samples

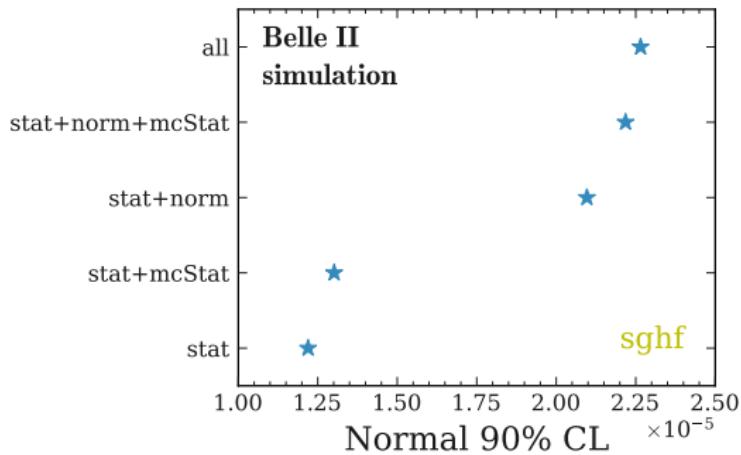
- Data samples used in the following:
 - 63 fb^{-1} collected at $\sqrt{s} = m(\Upsilon(4S))$ ("on-resonance").
 - 9 fb^{-1} collected at $\sqrt{s} = m(\Upsilon(4S)) - 60 \text{ MeV}$ ("off-resonance").
- Result published in Phys. Rev. Lett. 127 (2021) 181802.

Sources of systematic uncertainty

- Physics modelling:
 - Normalisation of each background source (50%).
 - Uncertainty on the branching fractions of B meson decays.
 - Uncertainty on the $B \rightarrow K$ form factor in signal simulation.
- Detector modelling:
 - Modelling of the track-finding efficiency.
 - Modelling of the measured energy of neutral particles.
 - Modelling of the PID selection efficiency.

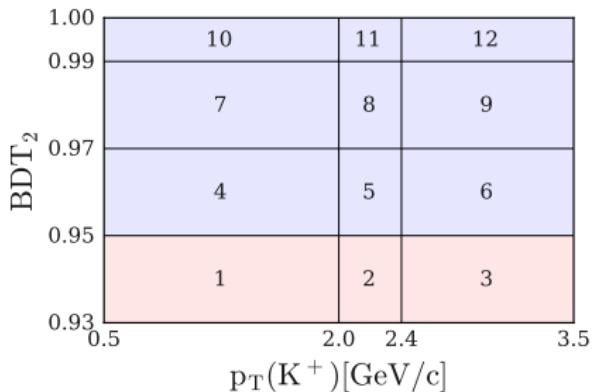
Expected upper limit for $(63 + 9) \text{ fb}^{-1}$ of data

- Examine how the expected upper limit on $\text{Br}(B^+ \rightarrow K^+ \nu \bar{\nu})$ increases when including sources of systematic uncertainties:



Statistical model

- Binned likelihood defined in the $\text{BDT}_2 \times p_T(K^+) \times \sqrt{s}$ space.

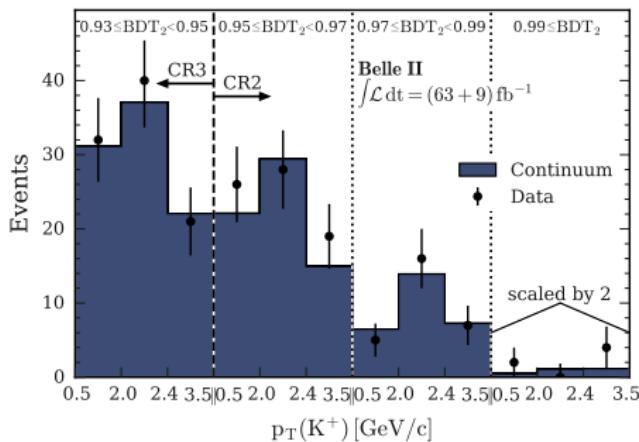


- $\mathcal{L}(\mu, \theta | n_1, \dots, n_{24}) = \frac{1}{Z} \prod_{b=1}^{24} \text{Pois}(n_b | \nu_b(\mu, \theta)) p(\theta)$.
 - n_1, \dots, n_{24} (ν_1, \dots, ν_{24}) events are observed (expected).
 - Signal strength $\mu = \frac{\text{Br}(B^+ \rightarrow K^+ \nu \bar{\nu})}{\text{Br}(B^+ \rightarrow K^+ \nu \bar{\nu})_{\text{SM}}}$
 - Nuisance parameters θ to include systematic uncertainties via event-count modifiers.

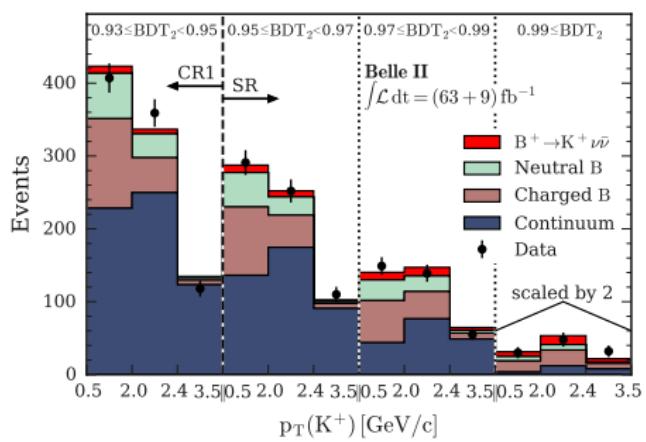
Result of the fit to $(63 + 9) \text{ fb}^{-1}$ of data

- $\mu = 4.2^{+3.4}_{-3.2} = 4.2^{+2.9}_{-2.8} (\text{stat})^{+1.8}_{-1.6} (\text{syst})$.
- $\text{BR}(B^+ \rightarrow K^+ \nu \bar{\nu}) = [1.9^{+1.6}_{-1.5}] \times 10^{-5} = [1.9^{+1.3}_{-1.3} (\text{stat})^{+0.8}_{-0.7} (\text{syst})] \times 10^{-5}$.

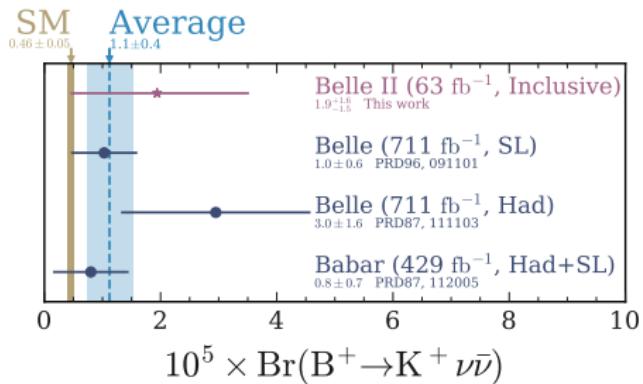
Off-resonance.



On-resonance.



Comparison with previous results



- Assuming uncertainty scaling as $1/\sqrt{L}$, inclusive tagging is better than hadronic and semileptonic tagging methods.

Experiment	$L [\text{fb}^{-1}]$	Method	$\sigma_{\text{Br}} [10^{-5}]$	$\sigma_{\text{Br}} \cdot \sqrt{L} [10^{-4} \sqrt{\text{fb}^{-1}}]$	Ref.
Belle	711	Hadronic	1.63	4.36	Phys. Rev. D 87, 111103 (2013)
Belle	711	Semileptonic	0.57	1.51	Phys. Rev. D 96, 091101 (2017)
Babar	429	Combined	0.65	1.35	Phys. Rev. D 87, 112005 (2013)
Belle II	63	Inclusive	1.55	1.23	Phys. Rev. Lett. 127, 181802 (2021)

Expected upper limit for 189 fb^{-1} of data

- Expected upper limit with 189 fb^{-1} of data.
 - $\text{Br}(B^+ \rightarrow K^+ \nu \bar{\nu}) < 1.0 \times 10^{-5}$ (90% C.L.).
- Current best upper limit obtained by Babar with 429 fb^{-1} of data.
 - $\text{Br}(B^+ \rightarrow K^+ \nu \bar{\nu}) < 1.6 \times 10^{-5}$ (90% C.L.).

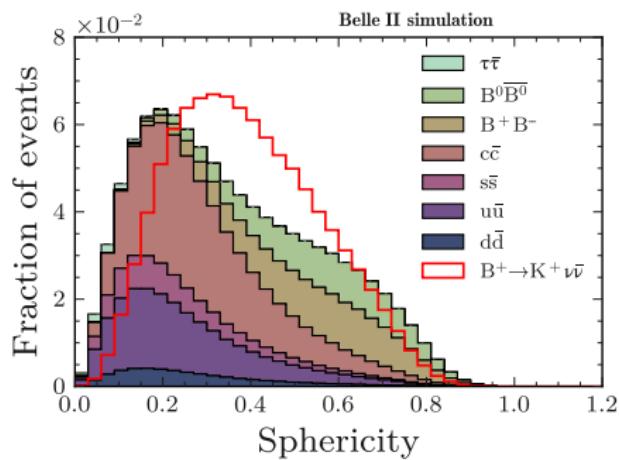
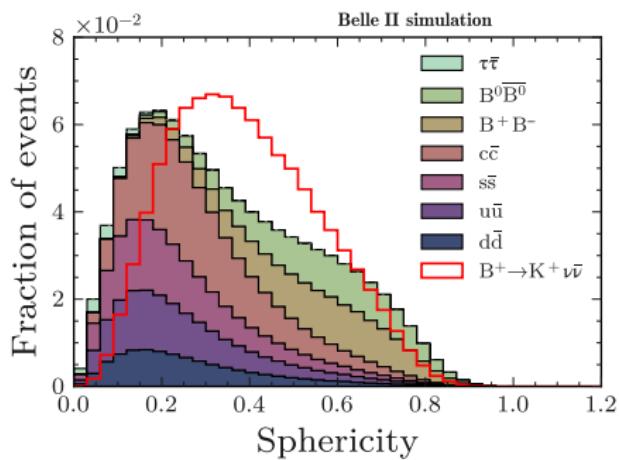
[Phys. Rev. D **87**, 112005 (2013) (Babar)]

Conclusion and outlook

- First search for $B \rightarrow K\nu\bar{\nu}$ decays with an inclusive tagging method.
- Result published with $(63 + 9) \text{ fb}^{-1}$ of Belle II data.
 - [Phys. Rev. Lett. 127 \(2021\) 181802](#).
- The inclusive tagging method provides a better sensitivity per integrated luminosity than the tagging methods of previous searches.
- The *expected* upper limit on $\text{Br}(B^+ \rightarrow K^+\nu\bar{\nu})$ with $(189 + 18) \text{ fb}^{-1}$ is better than the current best upper limit.
- The success of the inclusive tagging method opens new opportunities for the search of rare decays with neutrinos in the final state.

Thank you for your attention.

Bug in background category weighting



Transition between two generations of quarks

- Transition between two generations of quarks possible because the weak eigenstates are linear combinations of the mass eigenstates:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

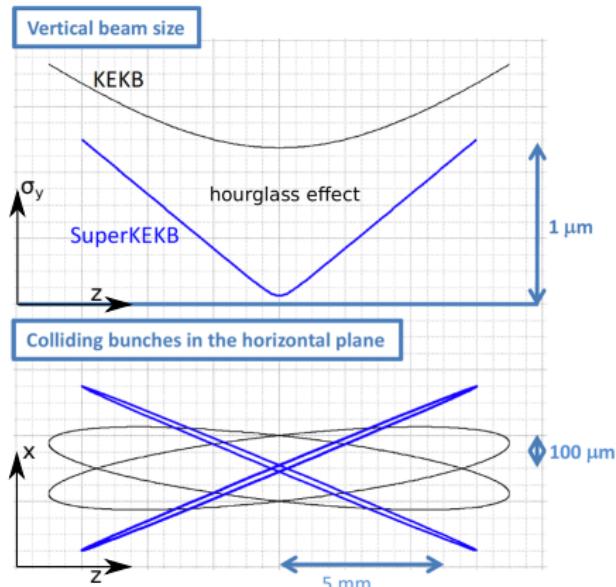
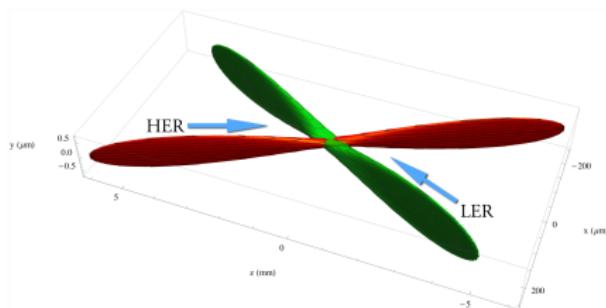
$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} \approx \begin{pmatrix} 0.974 & 0.226 & 0.004 \\ 0.226 & 0.973 & 0.041 \\ 0.009 & 0.040 & 0.999 \end{pmatrix}$$

Experimental anomalies in $b \rightarrow s\ell^+\ell^-$

- $R_H \equiv \frac{\text{Br}(B \rightarrow H\mu^+\mu^-)}{\text{Br}(B \rightarrow He^+e^-)}$.
- In 2017, LHCb reported a value of $R_{K^{*0}}$ 2.3σ smaller than the SM prediction.
- In 2020, LHCb reported the result of a fit to the angular variables of $B^0 \rightarrow K^{*0}\mu^+\mu^-$ decays showing a 3.3σ deviation from the predicted SM value of a Wilson coefficient specific to $B \rightarrow K^{(*)}\ell^+\ell^-$ decays called C_9 .
- In 2021, LHCb reported the result of a fit to the angular variables of $B^+ \rightarrow K^{*+}\mu^+\mu^-$ decays showing a 3.1σ deviation from the predicted SM value of the same Wilson coefficient C_9 .
- In 2022, LHCb reported a value of R_{K^+} 3.1σ smaller than the SM prediction.

Nano-beam scheme (idea from Pantaleo Raimondi)

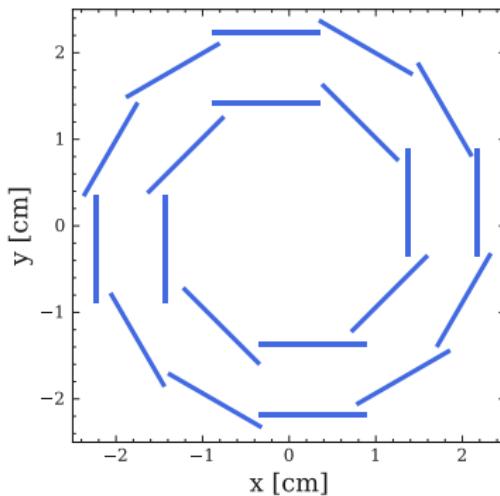
- Goal: $\beta_y^* = 0.3 \text{ mm}$.
- Hourglass effect limited if $\sigma_z^{\text{eff}} < \beta_y^*$.
- Half crossing angle:
 - $\phi_x \approx 40 \text{ mrad}$.
- Nominal beam spot parameters:
 - $\sigma_x \approx 10 \mu\text{m}$.
 - $\sigma_z^{\text{eff}} = \frac{\sigma_x}{\sin \phi_x} \approx 0.25 \text{ mm}$.
 - $\sigma_y \approx 50 \text{ nm}$.



[BELLE2-TALK-CONF-2018-142]
[1809.01958]

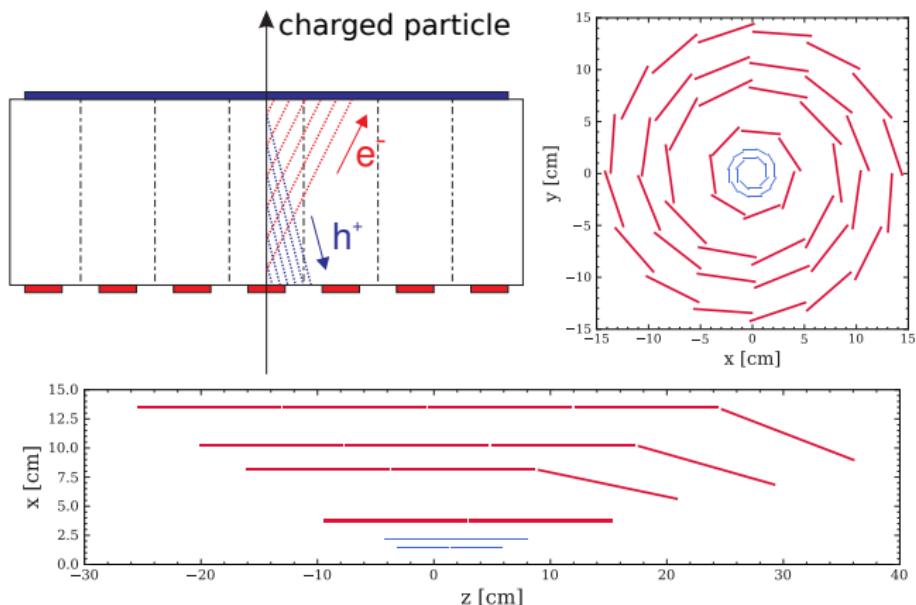
Pixel Detector (PXD)

- Based on depleted p-channel field-effect transistors (DEPFETs).
- 40 modules of 768×250 DEPFET pixels.
- Impact-parameter resolution of approximately $12 \mu\text{m}$.



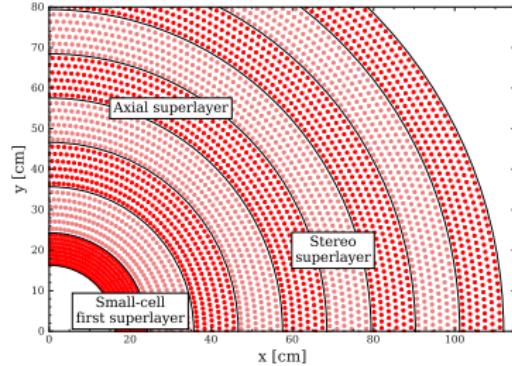
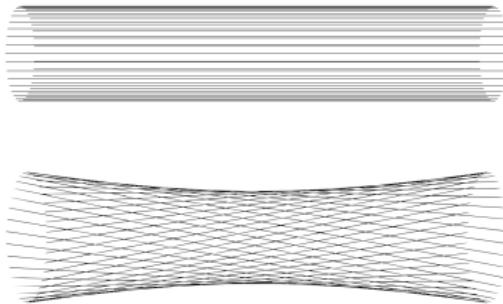
Silicon Vertex Detector (SVD)

- Based on double-sided silicon strips.
- p -side and n -side strips perpendicular to determine coordinates of charged particles.



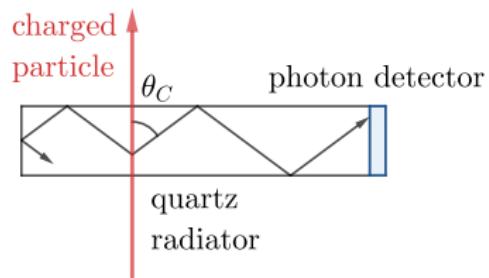
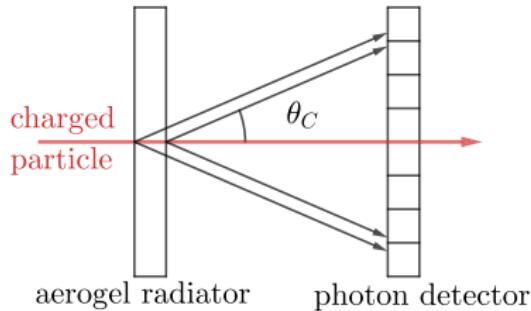
Central Drift Chamber (CDC)

- 56 layers of wires with axial or skewed orientation.
- Electron released during ionisation of a gas mixture cause avalanches of electrons.
- $p_T = |q|B\rho$ with a relative resolution of the order of 0.1%.

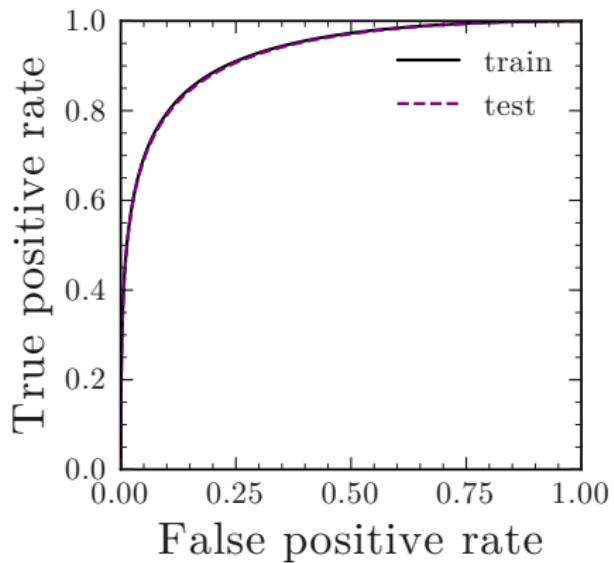
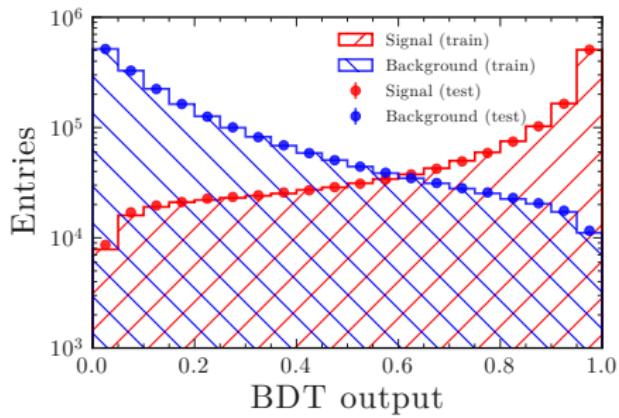


Particle identification

- Two detectors use Cherenkov radiations for particle identification.
 - Aerogel Ring-Imaging Cherenkov (ARICH).
 - Time-Of-Propagation (TOP) counter.
- Particle identification from $\beta = 1/(n \cos \theta_C)$ and $m = p/(\gamma\beta)$.

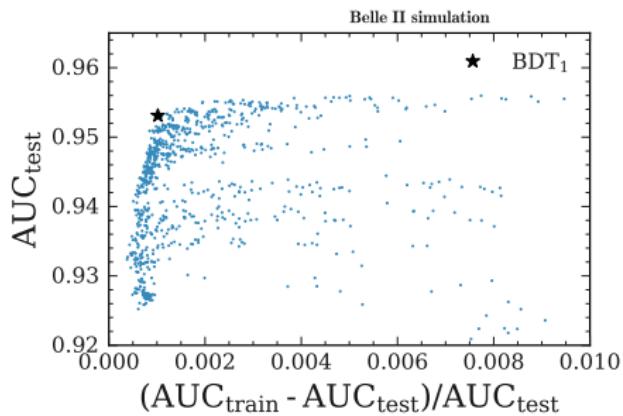
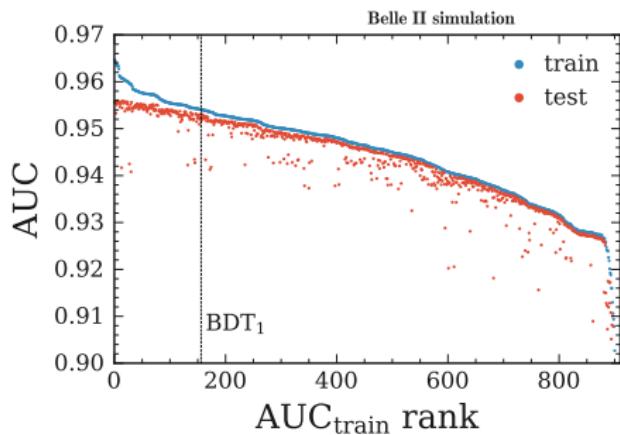


Example of ROC curve



Choice of BDT parameters

Parameter	Tested values	BDT_1	BDT_2
Number of trees	[200, 500, 1000, 2000]	2000	2000
Tree depth	[2, 3, 4, 5, 6]	2	3
Learning rate	[0.05, 0.1, 0.2]	0.2	0.2
Sampling rate	[0.5, 0.8, 1.0]	0.5	0.5
Number of equal-frequency bins	$[2^4, 2^6, 2^8, 2^{10}, 2^{12}]$	2^8	2^8



Input variables for BDT_1

Variable	$B^+ \rightarrow K^+ \nu \bar{\nu}$	$B^0 \rightarrow K_S^0 \nu \bar{\nu}$
ΔE_{ROE}	0.62682	0.59926
Modified Fox-Wolfram $H_{m,2}^{so}$	0.08731	0.09382
p_{ROE}	0.03517	0.03869
Modified Fox-Wolfram $H_{m,4}^{so}$	0.02372	0.01397
Modified Fox-Wolfram R_0^{oo}	0.02003	0.00876
Modified Fox-Wolfram R_2^{oo}	0.01988	0.02958
$\theta(p_{\text{ROE}})$	0.01813	0.01584
Harmonic Moment B_0	0.01732	0.02434
$\cos(\text{thrust}_B, \text{thrust}_{\text{ROE}})$	0.01371	0.00954
Fox-Wolfram Moment R_1	0.01072	0.03725
$\cos(\theta(\text{thrust}))$	0.00953	0.00992
Harmonic Moment B_2	0.00648	0.01390

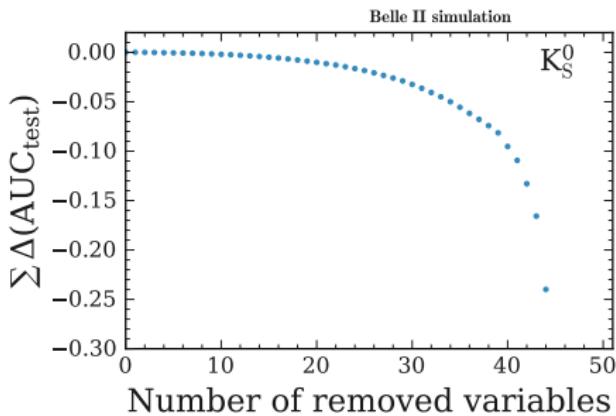
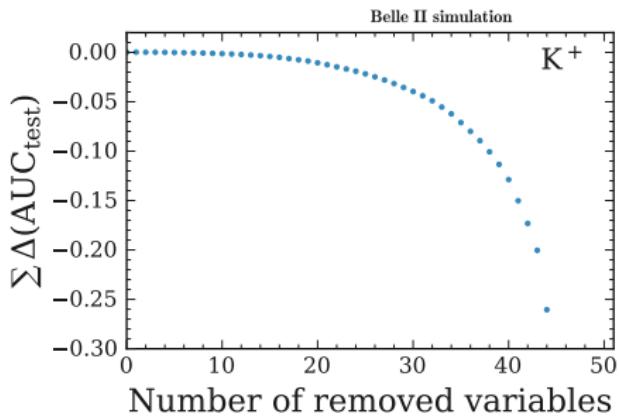
Input variables for BDT_2 ($B^+ \rightarrow K^+ \nu \bar{\nu}$)

Variable	$\Delta(\text{AUC})$	Variable	$\Delta(\text{AUC})$
Modified Fox-Wolfram $H_{m,2}^{so}$	-0.53286	$dx(\text{Tag Vertex})$	-0.00212
$\cos(\text{thrust}_B, \text{thrust}_{\text{ROE}})$	-0.11077	p-value(D^0)	-0.00198
Median(p-value(D^0))	-0.06013	$dz(\text{Tag Vertex})$	-0.00176
p-value(Tag Vertex)	-0.02718	Fox-Wolfram Moment R_1	-0.00126
$M(D^0)$	-0.02297	$\cos(\text{thrust}_B, z)$	-0.00125
$\theta(p_{\text{missing}})$	-0.02144	$\cos(\theta(\text{thrust}))$	-0.00121
ΔE_{ROE}	-0.01536	Modified Fox-Wolfram R_0^{so}	-0.00088
Modified Fox-Wolfram $R_2^{so}(D^0)$	-0.01272	VariancerOE(ρ_T)	-0.00088
p-value(D^+)	-0.01123	Fox-Wolfram Moment R_3	-0.00058
N_{lepton}	-0.00946	$dr(K^+, \text{Tag Vertex})$	-0.00057
Total charge squared	-0.00889	Harmonic Moment B_0	-0.00045
$dz(K^+, \text{Tag Vertex})$	-0.00881	$dy(\text{Tag Vertex})$	-0.00036
$dr(K^+)$	-0.00688	$dr(D^0)$	-0.00033
$\theta(p_{\text{ROE}})$	-0.00641	$M(\text{ROE})$	-0.00032
Modified Fox-Wolfram $H_{n,2}^{so}$	-0.00497	thrust _{ROE}	-0.00019
Modified Fox-Wolfram $H_{c,2}^{so}$	-0.00436	$dz(D^0)$	-0.00019
Modified Fox-Wolfram $H_{m,4}^{so}$	-0.00409	$dz(K^+)$	-0.00018
ρ_{ROE}	-0.00398	$dz(D^+)$	-0.00012
M_{missing}^2	-0.00354	$N_{\text{tracks}} + N_\gamma$	-0.00010
Sphericity	-0.00333	Harmonic Moment B_2	-0.00004
Fox-Wolfram Moment R_2	-0.00299	$\phi(K^+)$	-0.00003
N_{tracks}	-0.00263	Modified Fox-Wolfram $H_{m,0}^{so}$	+0.00003
N_γ	-0.00223	Thrust	+0.00013
$dr(D^+)$	-0.00220		

Input variables for BDT_2 ($B^0 \rightarrow K_S^0 \nu \bar{\nu}$)

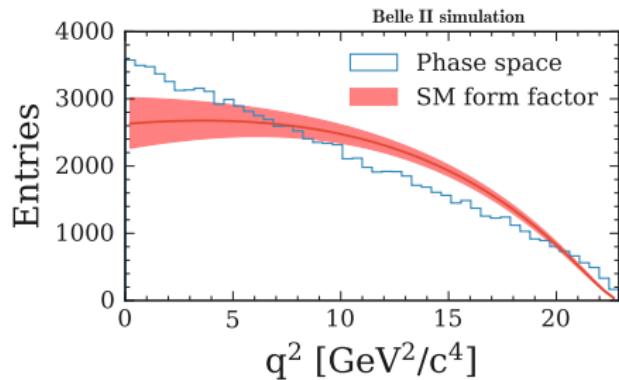
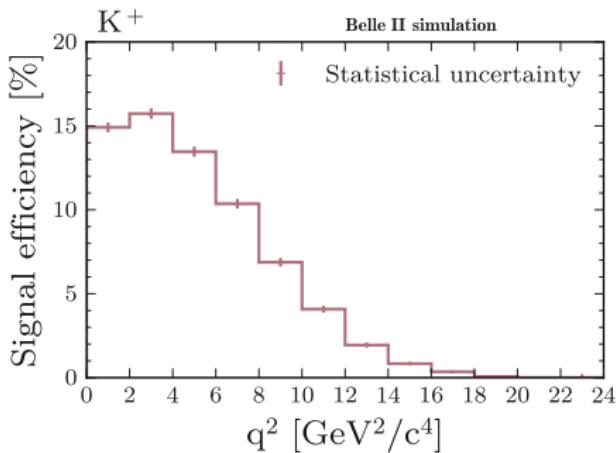
Variable	$\Delta(\text{AUC})$	Variable	$\Delta(\text{AUC})$
Modified Fox-Wolfram $H_{m,2}^{so}$	-0.68504	$M(K_S^0)$	-0.00137
$\cos(\text{thrust}_B, \text{thrust}_{\text{ROE}})$	-0.07418	$\cos(\text{thrust}_B, z)$	-0.00135
$\theta(p_{\text{missing}})$	-0.03280	M_{missing}^2	-0.00121
$\cos(p_{K_S^0}, \text{line(IP, } K_S^0 \text{ vertex))}$	-0.02363	$dr(p_{K_S^0})$	-0.00113
Fox-Wolfram Moment R_2	-0.01403	$dz(\text{Tag Vertex})$	-0.00106
Modified Fox-Wolfram R_2^{oo}	-0.01384	$\cos(\theta(\text{thrust}))$	-0.00095
ΔE_{ROE}	-0.00724	Fox-Wolfram Moment R_3	-0.00080
N_{lepton}	-0.00636	Modified Fox-Wolfram $H_{m,0}^{so}$	-0.00071
Modified Fox-Wolfram $H_{m,4}^{so}$	-0.00613	$M(D^0)$	-0.00068
$\theta(p_{\text{ROE}})$	-0.00612	Variance $_{\text{ROE}}(p_T)$	-0.00061
$dz(p_{D^0})$	-0.00560	$dz(p_{K_S^0})$	-0.00054
Modified Fox-Wolfram $H_{c,2}^{so}$	-0.00503	Fox-Wolfram Moment R_1	-0.00042
p_{ROE}	-0.00436	$M(\bar{D}^+)$	-0.00040
p-value(Tag Vertex)	-0.00428	$dr(p_{K_S^0}, \text{ Tag Vertex})$	-0.00034
Modified Fox-Wolfram $H_{n,2}^{so}$	-0.00386	Modified Fox-Wolfram R_0^{oo}	-0.00030
N_{tracks}	-0.00360	$M(\text{ROE})$	-0.00022
N_γ	-0.00290	$N_{\text{tracks}} + N_\gamma$	-0.00021
p-value(D^+)	-0.00268	$dr(p_{D^0})$	-0.00019
Harmonic Moment B_0	-0.00257	$dy(\text{Tag Vertex})$	-0.00018
$dz(p_{K_S^0}, \text{ Tag Vertex})$	-0.00228	$dz(p_{D^+})$	-0.00011
Total charge squared	-0.00203	Thrust	-0.00007
$dx(\text{Tag Vertex})$	-0.00188	Harmonic Moment B_2	-0.00002
Sphericity	-0.00171	$\text{thrust}_{\text{ROE}}$	+0.00008

Input variable selection



Signal selection efficiency in the signal search region

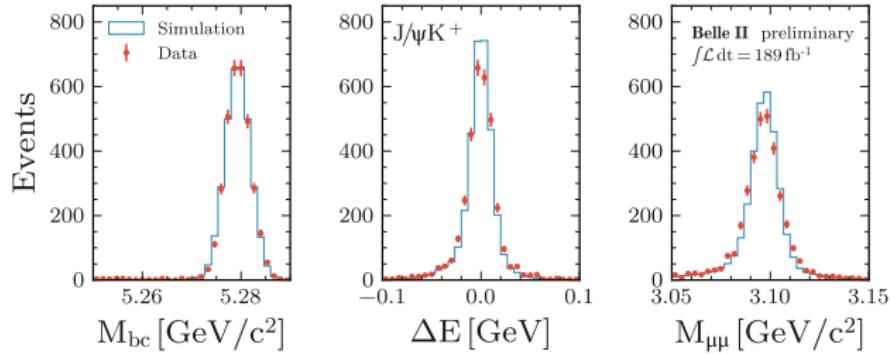
- In the signal search region, the signal efficiency is 15% for $q^2 \approx 0$ and drops to zero for $q^2 > 18 \text{ GeV}^2/c^4$.
 - Sensitive to potential light dark matter candidates.



Validation channel: $B^+ \rightarrow K^+ J/\psi (\rightarrow \mu^+ \mu^-)$

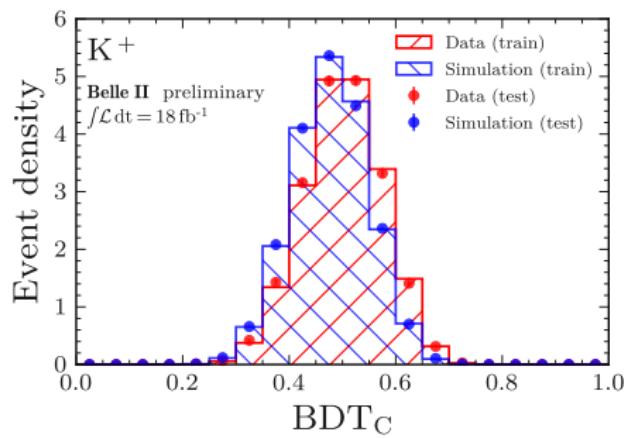
① Select $B^+ \rightarrow K^+ J/\psi (\rightarrow \mu^+ \mu^-)$ decays in data and simulation.

- $|M_{\mu\mu} - M_{J/\psi}^{\text{PDG}}| < 0.05 \text{ GeV}/c^2$
- $|\Delta E| \equiv \left|E_B^* - \frac{\sqrt{s}}{2}\right| < 0.1 \text{ GeV}$
- $M_{bc} \equiv \sqrt{\left(\frac{\sqrt{s}}{2c^2}\right)^2 - \left(\frac{p_B^*}{c}\right)^2} > 5.25 \text{ GeV}/c^2$



Correction of continuum background mis-modelling

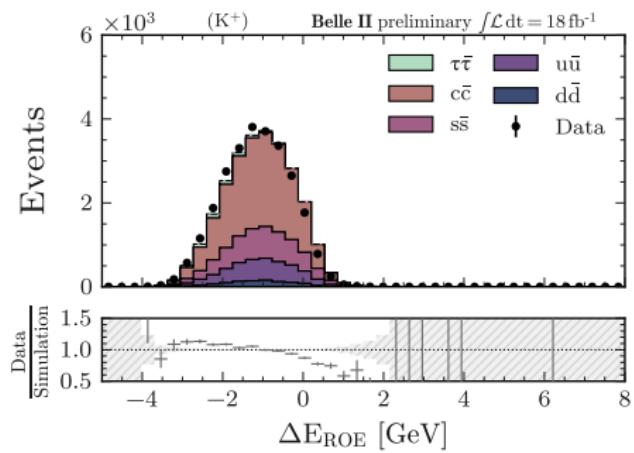
- Data-driven method to correct mis-modelling of off-resonance data.
 - ➊ Train a binary classifier (BDT_c) to distinguish **simulation vs data**.
 - ➋ Given the output of BDT_c , the simulated events are weighted according to $BDT_c/(1 - BDT_c)$.



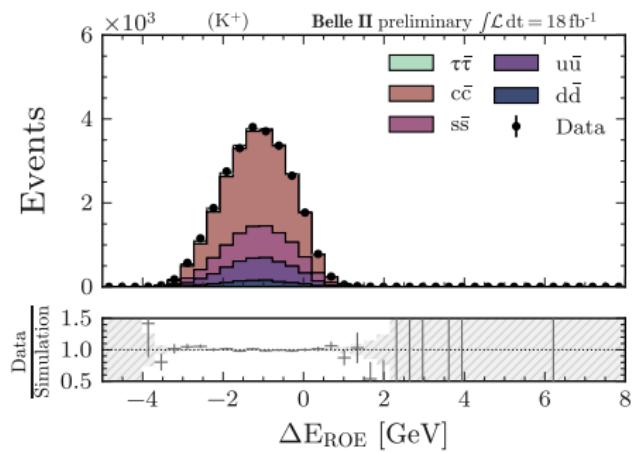
Correction of continuum background mis-modelling

- Agreement between off-resonance data and simulation improved.

Before reweighting:



After reweighting:

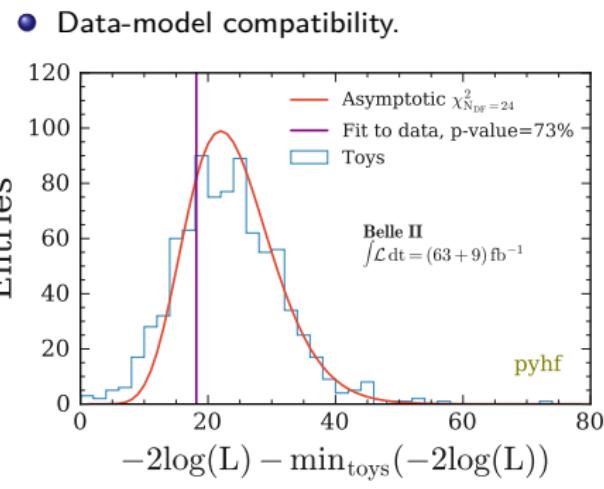
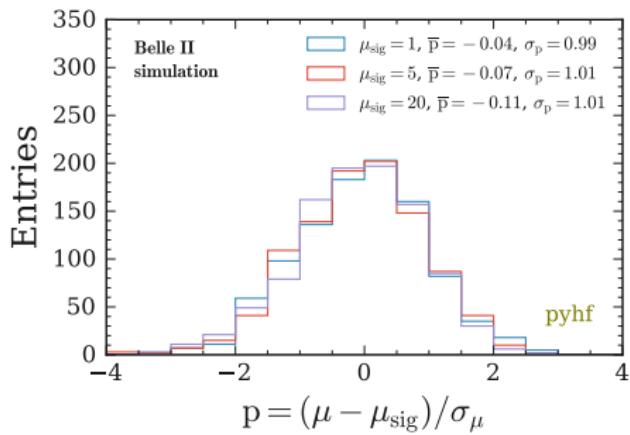


Statistical model

- {samples} = {Signal, $u\bar{u}$, $d\bar{d}$, $c\bar{c}$, $s\bar{s}$, $\tau\bar{\tau}$, $B^0\bar{B}^0$, B^+B^- }.
- 24 bins where n_1, \dots, n_{24} (ν_1, \dots, ν_{24}) events are observed (expected).
- Fit parameters:
 - Signal strength $\mu = \text{Br}(B \rightarrow K\nu\bar{\nu})/\text{Br}(B \rightarrow K\nu\bar{\nu})_{\text{SM}}$.
 - Nuisance parameters $\boldsymbol{\theta} = (\mu_{u\bar{u}}, \mu_{d\bar{d}}, \dots, \mu_{B^+B^-}, \theta_8, \dots, \theta_N)^T$ to include systematic uncertainties via event-count modifiers.
- Likelihood function = product of Poisson probability density functions combining the information from the 24 bins.
 - $\mathcal{L}(\mu, \boldsymbol{\theta} | n_1, \dots, n_{24}) = \frac{1}{Z} \prod_{b=1}^{24} \text{Pois}(n_b | \nu_b(\mu, \boldsymbol{\theta})) p(\boldsymbol{\theta})$.
 - $\nu_b(\mu, \boldsymbol{\theta}) = \sum_{s \in \{\text{samples}\}} \mu_s (\nu_{bs}^0 + \Delta_{bs}(\boldsymbol{\theta}))$.

Fit validation

- Toys generated for the simulated data set.
 - Poisson statistical fluctuations.
 - Gaussian systematic fluctuations.
- Signal injection study, $\mu_{\text{sig}} \in \{1, 5, 20\}$.

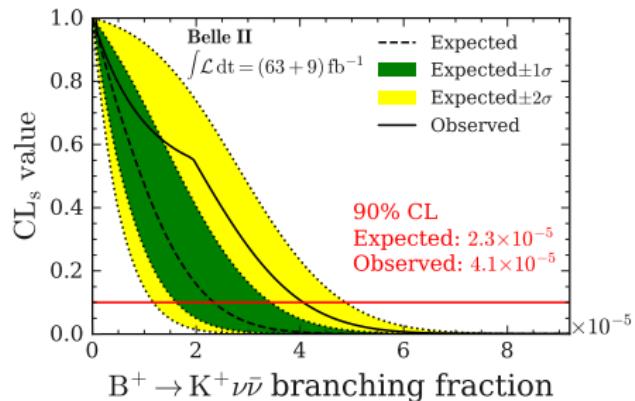


Upper limit determination with the CL_s method

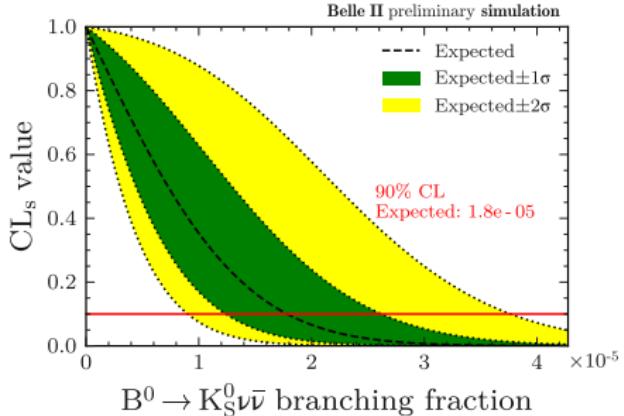
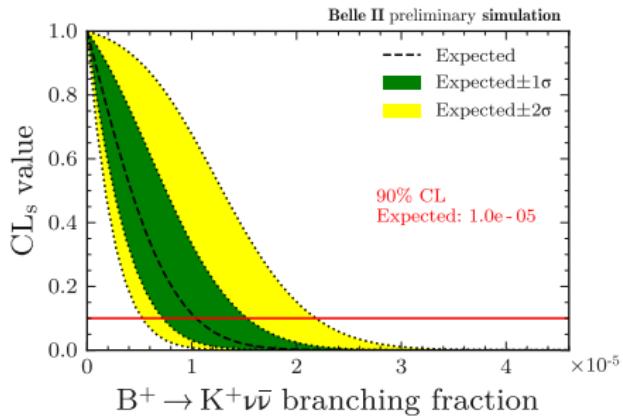
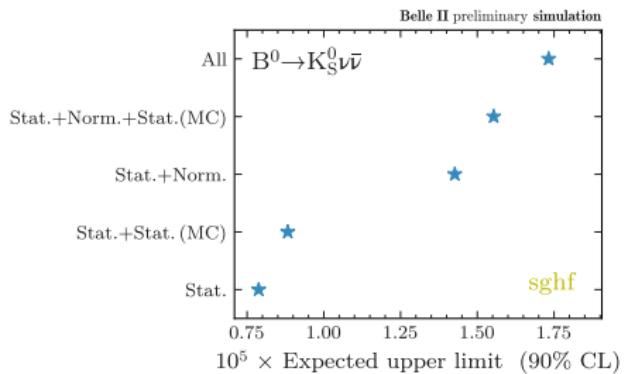
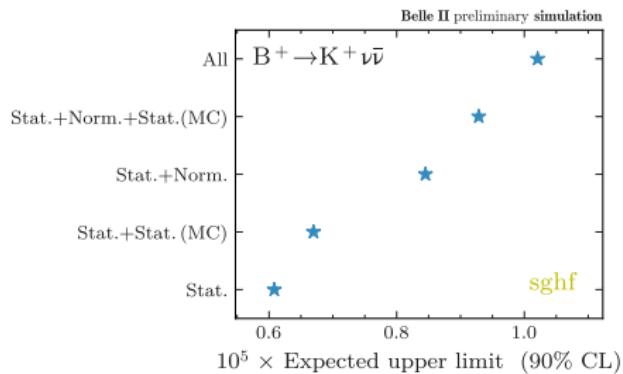
- Likelihood ratio: $\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\theta} | n_1, \dots, n_{N_b})}{\mathcal{L}(\hat{\mu}, \hat{\theta} | n_1, \dots, n_{N_b})}$.
- Test statistic: $q_\mu = \begin{cases} -2 \ln \lambda(\mu) & \text{if } \mu \geq \hat{\mu}, \\ 0 & \text{otherwise.} \end{cases}$
- p -value (level of agreement between data and hypothesised μ):
 - $p_{s+b} = P(q_\mu > q_{\mu, \text{obs}} | \mu) = \int_{q_{\mu, \text{obs}}}^{\infty} p(q_\mu | \mu) dq_\mu.$
- p -value (background-only hypothesis):
 - $p_b = P(q_\mu > q_{\mu, \text{obs}} | 0) = \int_{q_{\mu, \text{obs}}}^{\infty} p(q_\mu | 0) dq_\mu.$
- CL_s ratio: $\text{CL}_s = \frac{p_{s+b}}{p_b}.$
- 90% CL upper limit on μ : largest value of μ such that $\text{CL}_s \geq 0.1$.

Upper limit on $\text{Br}(B^+ \rightarrow K^+\nu\bar{\nu})$ with $(63 + 9) \text{ fb}^{-1}$ of data

- Upper on the branching fraction determined using the CL_s method.
- $\text{Br}(B^+ \rightarrow K^+\nu\bar{\nu}) < 4.1 \times 10^{-5}$ @ 90% C.L.

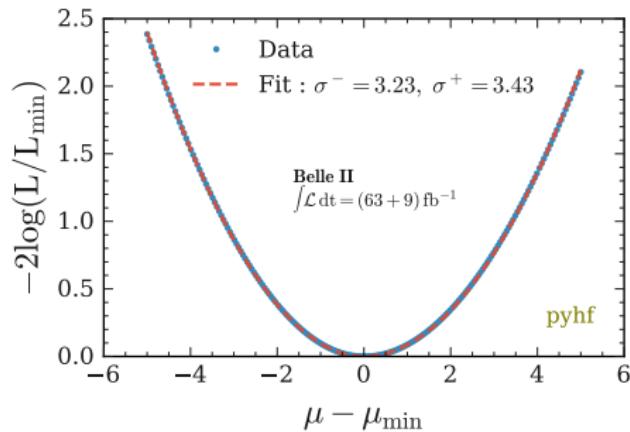


Expected upper limit with $(189 + 18) \text{ fb}^{-1}$ of data



Uncertainty on μ with $(63 + 9) \text{ fb}^{-1}$ of data

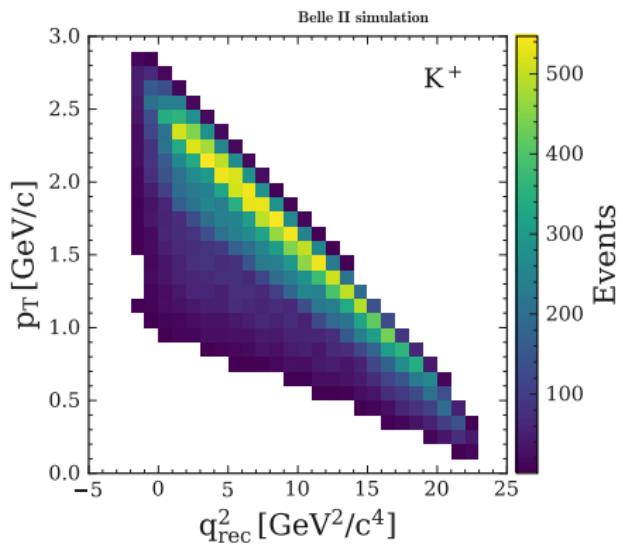
- $\mu = 4.2^{+3.4}_{-3.2} = 4.2^{+2.9}_{-2.8}(\text{stat})^{+1.8}_{-1.6}(\text{syst})$.
- Total uncertainty on μ : profile likelihood scan, fitting the model with fixed values of μ while keeping the other fit parameters free.



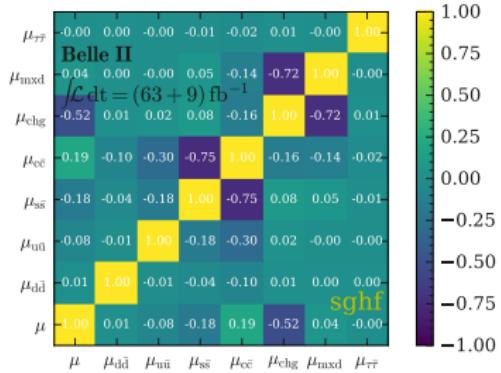
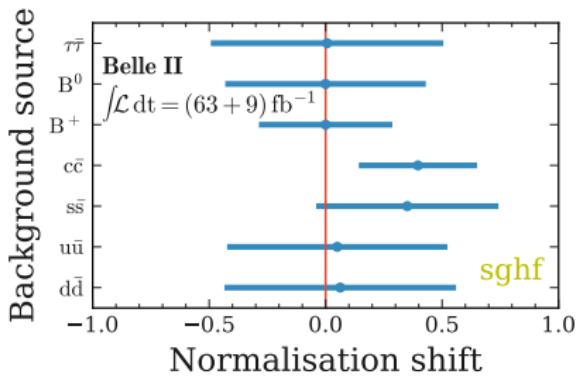
Upper limit on $\text{Br}(B \rightarrow K\nu\bar{\nu})$

- Upper on the branching fraction determined using the CL_s method.
- Reminder:
 - $\text{Br}(B^+ \rightarrow K^+\nu\bar{\nu})_{\text{SM}} = (4.6 \pm 0.5) \times 10^{-6}$. [Prog. Part. Nucl. Phys. 92, 50 (2017)]
 - $\text{Br}(B^0 \rightarrow K^0\nu\bar{\nu})_{\text{SM}} = (4.3 \pm 0.5) \times 10^{-6}$.

Experiment	Year	$L [\text{fb}^{-1}]$	Method	Mode	$\varepsilon_{\text{sig}} [\%]$	Limit at 90% CL	Ref.
Babar	2013	429	HAD	K^+	0.04	$< 3.7 \times 10^{-5}$	Phys. Rev. D 87, 112005 (2013)
				K^0	0.01	$< 8.1 \times 10^{-5}$	Phys. Rev. D 87, 112005 (2013)
Babar	2013	429	COM	K^+	-	$< \mathbf{1.6 \times 10^{-5}}$	Phys. Rev. D 87, 112005 (2013)
				K^0	-	$< 4.9 \times 10^{-5}$	Phys. Rev. D 87, 112005 (2013)
Belle	2013	711	HAD	K^+	0.06	$< 5.5 \times 10^{-5}$	Phys. Rev. D 87, 111103 (2013)
				K^0	0.004	$< 19 \times 10^{-5}$	Phys. Rev. D 87, 111103 (2013)
Belle	2017	711	SL	K^+	0.2	$< 1.9 \times 10^{-5}$	Phys. Rev. D 96, 091101 (2017)
				K^0	0.05	$< \mathbf{2.6 \times 10^{-5}}$	Phys. Rev. D 96, 091101 (2017)
Belle II	2021	63	INC	K^+	4.0	$< \mathbf{4.1 \times 10^{-5}}$	Phys. Rev. Lett. 127, 181802 (2021)
Belle II	2022	189	INC	K^+	4.0	$< 1.0 \times 10^{-5}$	(expected)
				K^0	4.0	$< 3.6 \times 10^{-5}$	(expected)

p_T vs q_{rec}^2 

Post-fit normalisation parameters



B^+ background in the signal search region

$\tilde{\varepsilon}_{\text{sig}}(B^+ \rightarrow K^+ \nu \bar{\nu}) > 0.92$	Fraction [%]	$\tilde{\varepsilon}_{\text{sig}}(B^+ \rightarrow K^+ \nu \bar{\nu}) > 0.98$	Fraction [%]
$B^\pm \rightarrow D^0 \mu^\pm \nu_\mu$	13.4	$B^\pm \rightarrow D^0 \mu^\pm \nu_\mu$	13.0
$B^\pm \rightarrow D^*(2007)^0 \mu^\pm \nu_\mu$	10.0	$B^\pm \rightarrow D^0 e^\pm \nu_e$	8.2
$B^\pm \rightarrow D^0 e^\pm \nu_e$	10.0	$B^\pm \rightarrow K^\pm K^0 K^0$	6.7
$B^\pm \rightarrow D^*(2007)^0 e^\pm \nu_e$	7.2	$B^\pm \rightarrow D^*(2007)^0 \mu^\pm \nu_\mu$	6.3
$B^\pm \rightarrow D^0 e^\pm \nu_e \gamma$	4.6	$B^\pm \rightarrow D^0 K^\pm$	4.4
$B^\pm \rightarrow K^\pm K^0 K^0$	3.6	$B^\pm \rightarrow D^0 e^\pm \nu_e \gamma$	4.1
$B^\pm \rightarrow D^*(2007)^0 e^\pm \nu_e \gamma$	3.4	$B^\pm \rightarrow D^*(2007)^0 e^\pm \nu_e$	3.9
$B^\pm \rightarrow D^0 K^\pm$	2.9	$B^\pm \rightarrow D^*(2007)^0 K^\pm$	3.7
$B^\pm \rightarrow D^*(2007)^0 K^\pm$	2.6	$B^\pm \rightarrow D^*(2007)^0 e^\pm \nu_e \gamma$	3.6
$B^\pm \rightarrow \eta_c(1S) K^\pm$	2.4	$B^\pm \rightarrow \eta_c(1S) K^\pm$	3.3
$B^\pm \rightarrow D^0 K^\pm K^0$	1.7	$B^\pm \rightarrow D^0 \pi^\pm$	3.2
$B^\pm \rightarrow D^0 \pi^\pm$	1.6	$B^\pm \rightarrow \tau^\pm \nu_\tau$	2.0
$B^\pm \rightarrow \rho(770)^\pm D^0$	1.5	$B^\pm \rightarrow K^0 K^\pm$	1.7
$B^\pm \rightarrow D^0 \tau^\pm \nu_\tau$	1.5	$B^\pm \rightarrow \rho(770)^\pm D^0$	1.6
$B^\pm \rightarrow J/\psi(1S) K^\pm$	1.4	$B^\pm \rightarrow nnK^\pm$	1.4