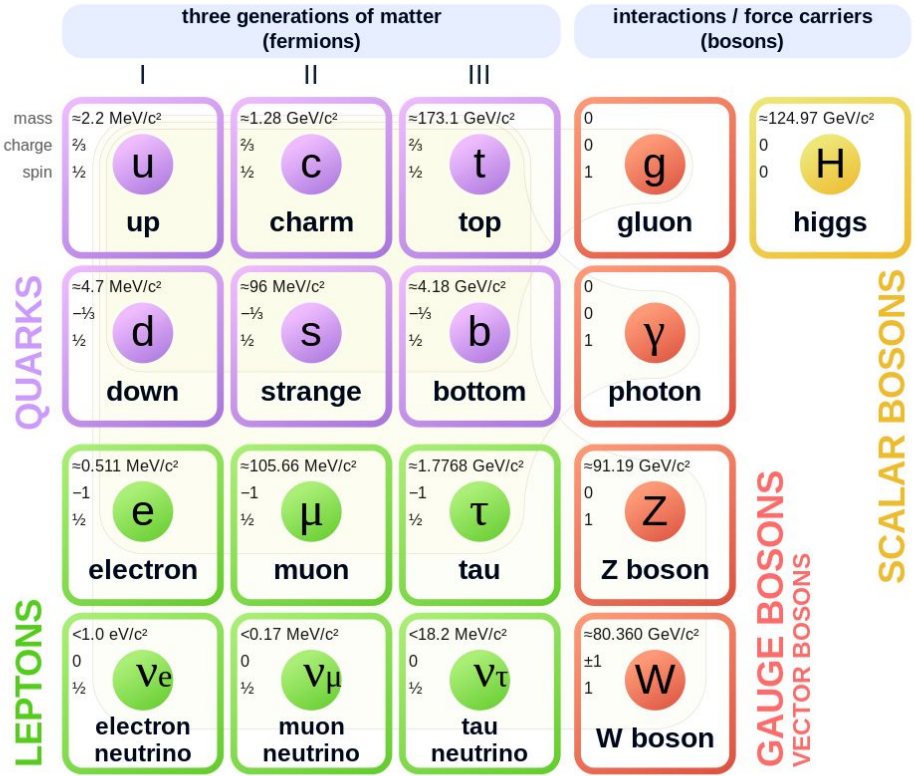
Evidence for $B^+ \to K^+ \bar{\nu}\nu$ decay at Belle II

DESY pizza seminar,
27 Nov 2023, Sasha Glazov

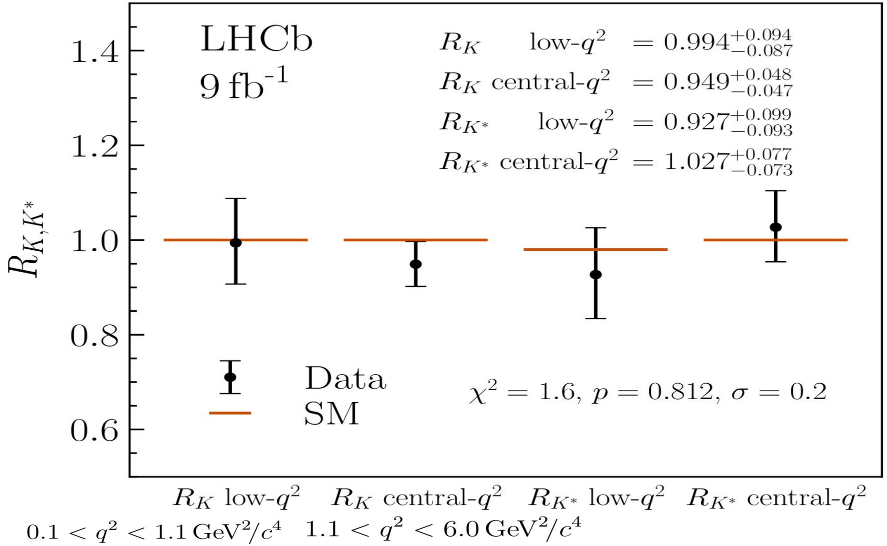
Standard Model and Flavour

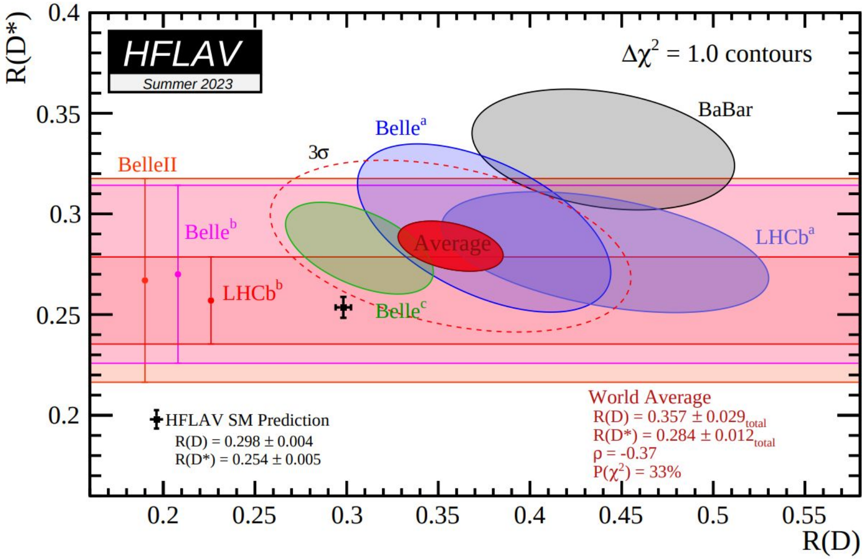


- Three flavours of quarks and leptons which in the standard model have identical couplings to the gauge bosons: lepton flavor universality
- Charged leptons and neutrinos couple to *W*, *γ*, *Z* bosons according to their *SU*(2)x*U*(1) charges.
- Any deviation from this pattern is a signal of new physics.

Flavour anomalies: some stay, some go

arXiv:2212.09153



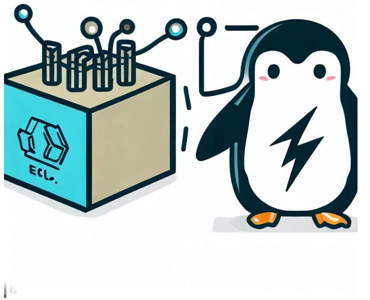


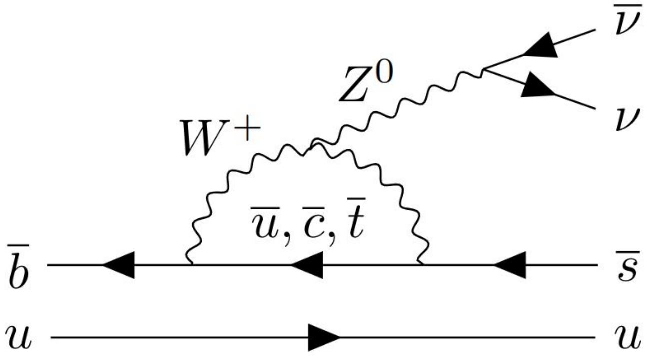
$$R_K = \frac{B(B^\pm \to K^\pm \mu^+ \mu^-)}{B(B^\pm \to K^\pm e^+ e^-)}$$

$$R(D(*)) = \frac{B(B \to D^{(*)} \tau^+ \nu)}{B(B \to D^{(*)} l^+ \nu)}$$

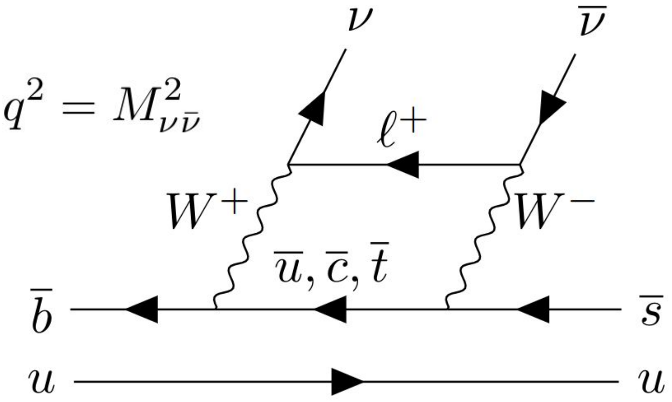
- Several searches for lepton universality violations, involving light (e, μ) and heavy (τ) leptons
- Tension for light leptons is resolved recently, while for τ-leptons it still remains

Motivation for $B^+ \to K^+ \nu\nu$





$$q^2 = M^2_{\nu\bar{\nu}}$$



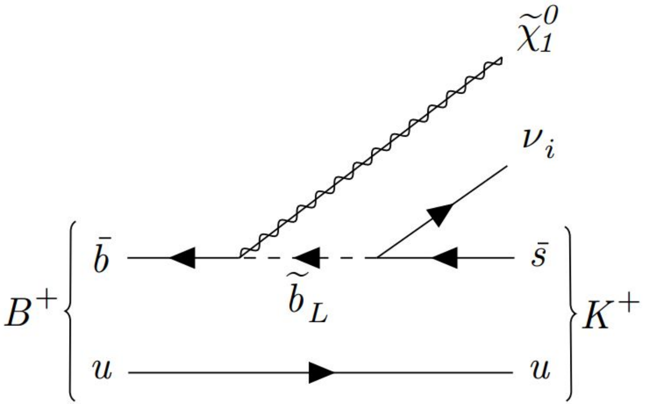
$$\frac{d\mathrm{BR}(B^+ \to K^+ \nu\bar{\nu})}{dq^2} = \tau_{B^+} 3 |N|^2 \frac{X_t^2}{s_w^4} \rho_K(q^2) \qquad N = V_{tb} V_{ts}^* \frac{G_F \alpha}{16\pi^2} \sqrt{\frac{m_B}{3\pi}}$$

- The $B^+ \to K^+ \nu\nu$ is a flavour-changing neutral current process which is known with high accuracy in the SM:

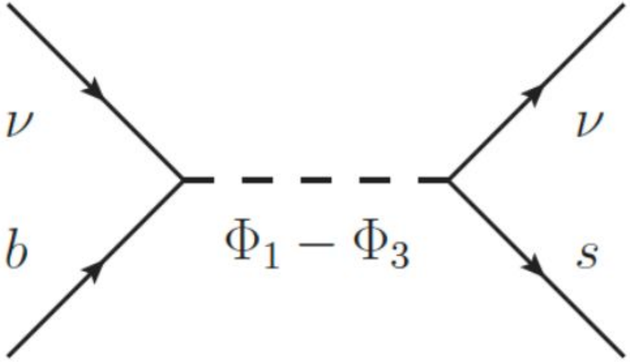
  $B(B^+ \to K^+ \nu\nu) = (5.6 \pm 0.4) \times 10^{-6}$ (arXiv:2207.13371)

- Extensions beyond SM may lead to significant rate increase

New Physics predictions



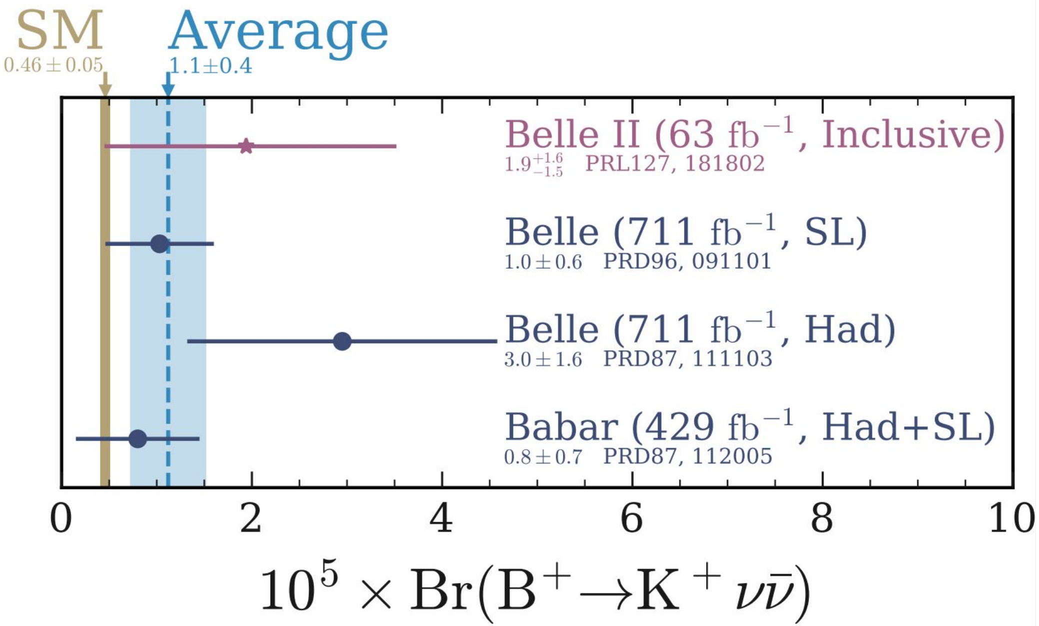
arXiv:2309.03727



The process can have significant enhancement in several new physics models. These include (scalar) leptoquarks, axions, dark-sector mediators. Recent papers discuss also R-parity violating SUSY, sterile neutrinos.

Interesting interplay of $B^+ \to K^+ \nu\nu$ and $B \to K^* \nu\nu$ channels. Some common explanations of $R(D^{(*)})$, muon g-2 anomalies

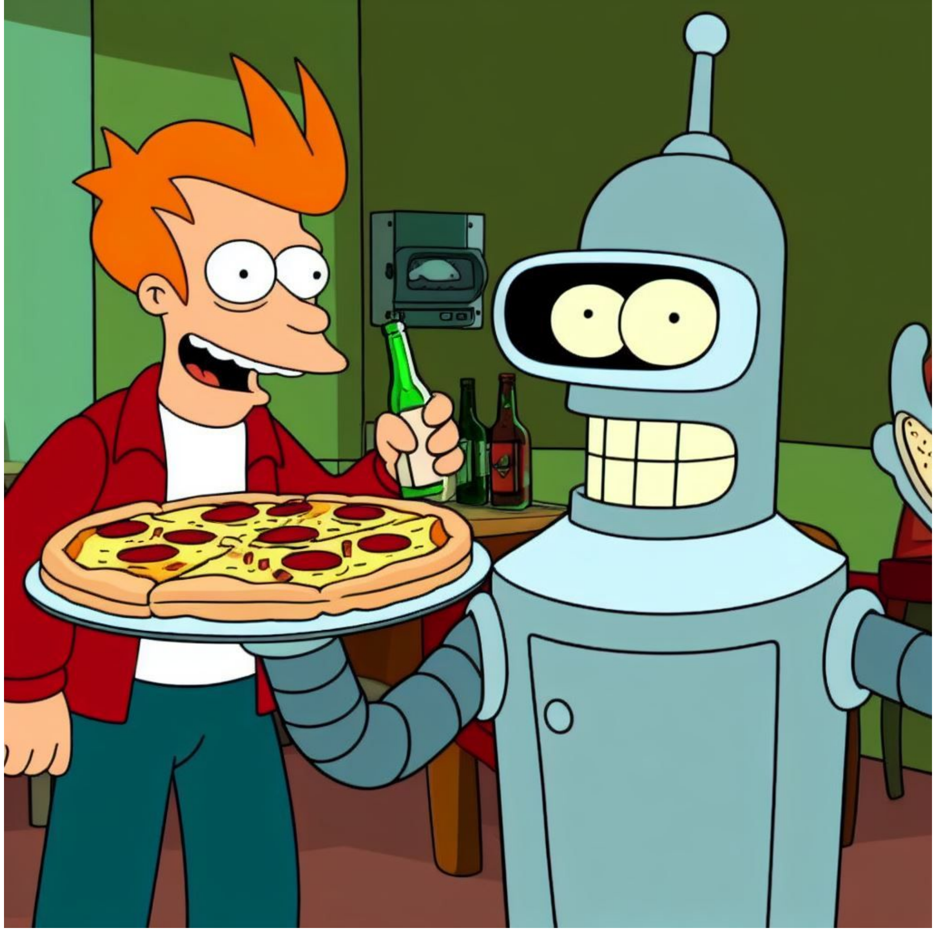
Experimental status before the latest Belle II result



- No significant signal
- Best upper limit from BaBar

Br($B^+ \to K^+ \nu\nu$)< 1.6 x10^{-5} (90% cl)

- Limits for K^* channels too
- Recent measurement from Belle II exploiting a new "inclusive" reconstruction method which shows good sensitivity for small data samples.



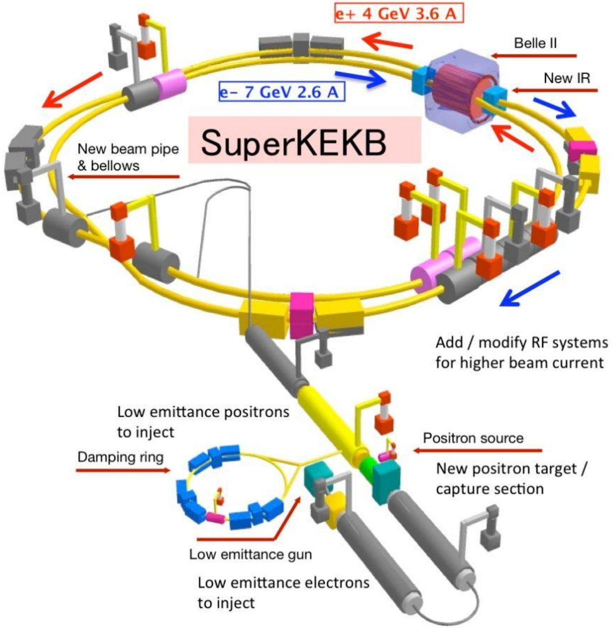
The paper appeared on the arXiv **today**: https://arxiv.org/abs/2311.14647

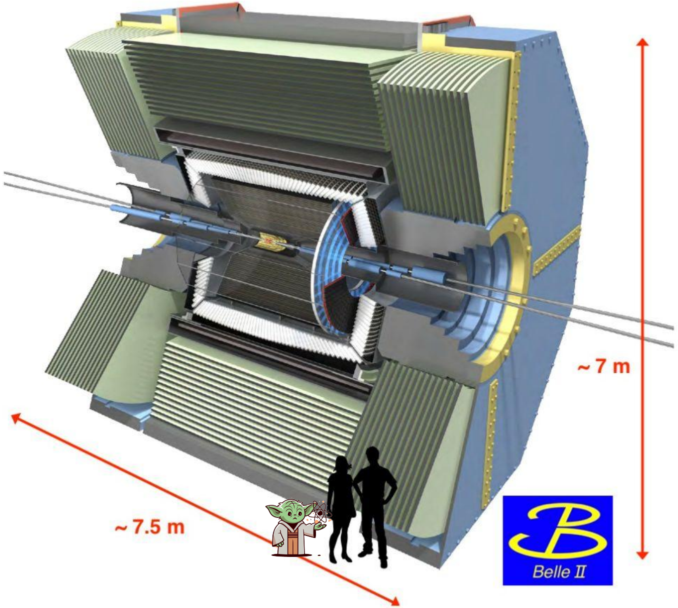
Belle II and SuperKEKB

Collected at $\Upsilon(4S)$: in total:	362 fb^{-1}, about 0.4×10^9 BB 424 fb^{-1}

SuperKEKB:

- e^+e^- collider with energies 4 GeV and 7 GeV operating around ***Υ(4S)*** resonance
- Achieved world-record peak luminosity of **4.7×10^{34} cm^{-2} s^{-1}**

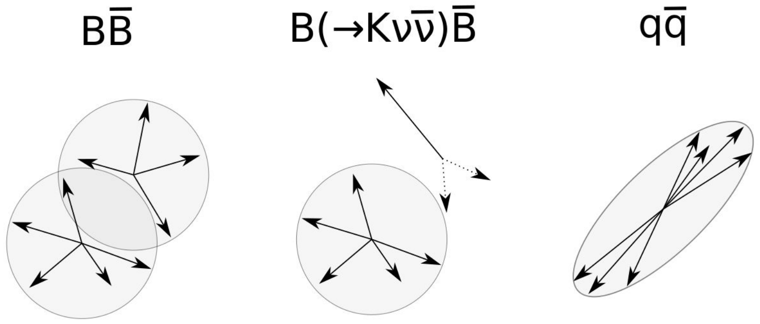


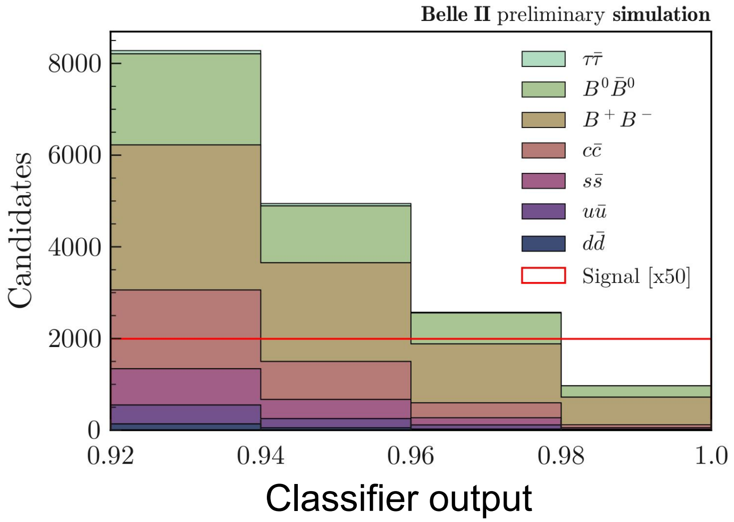


Belle II:

- Nearly 4π detector
- Tracking, PID, and photon reconstruction capabilities
- Similar performance for electrons and muons
- Well-suited to measure decays with missing energy, π^0 in the final state, inclusive measurements

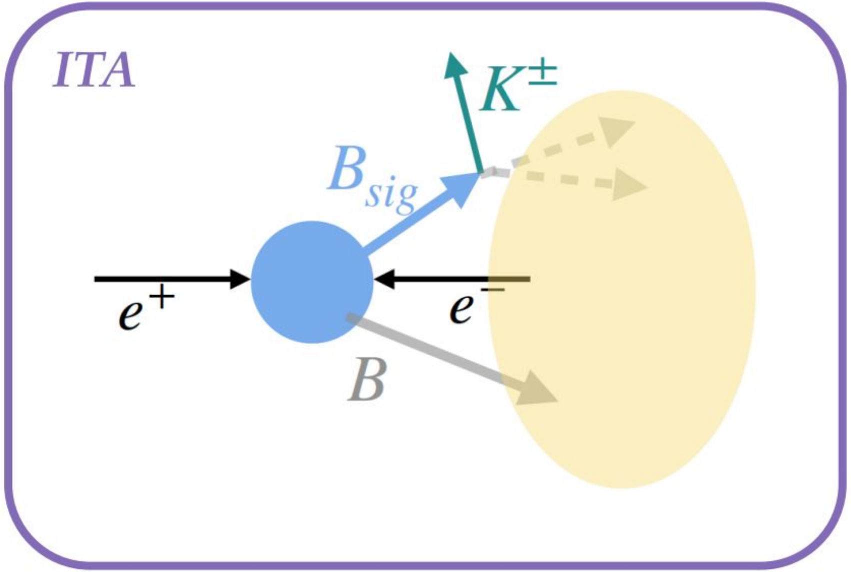
Analysis strategy





- Two analyses: more sensitive **inclusive** (total efficiency: **8%**) and conventional **hadronic** tagging (total efficiency: **0.4%**)
- Use event properties to suppress background with multiple variables combined
- Use classifier output as (one of) the fit variable(s), use **simulation** for signal and background templates
- Use multiple control channels to validate simulation with data

Inclusive tag



Reconstruction and background suppression

- Selection criteria for particles to ensure high and well-measured efficiency:
  - charged particle momenta and neutral particle energies greater than 100 MeV
  - only in central region
  - charged particles consistent with being from interaction point
- **Signal candidate:**
  - an identified charged kaon that gives the minimal mass of the neutrino pair q^2_{rec} (computed as K^+ recoil)

Event (pre-selection):
- $4 \leq N_{track} \leq 10$
- $E_{total} > 4$ GeV
- $17^o < \vartheta_{miss} < 160^o$

→

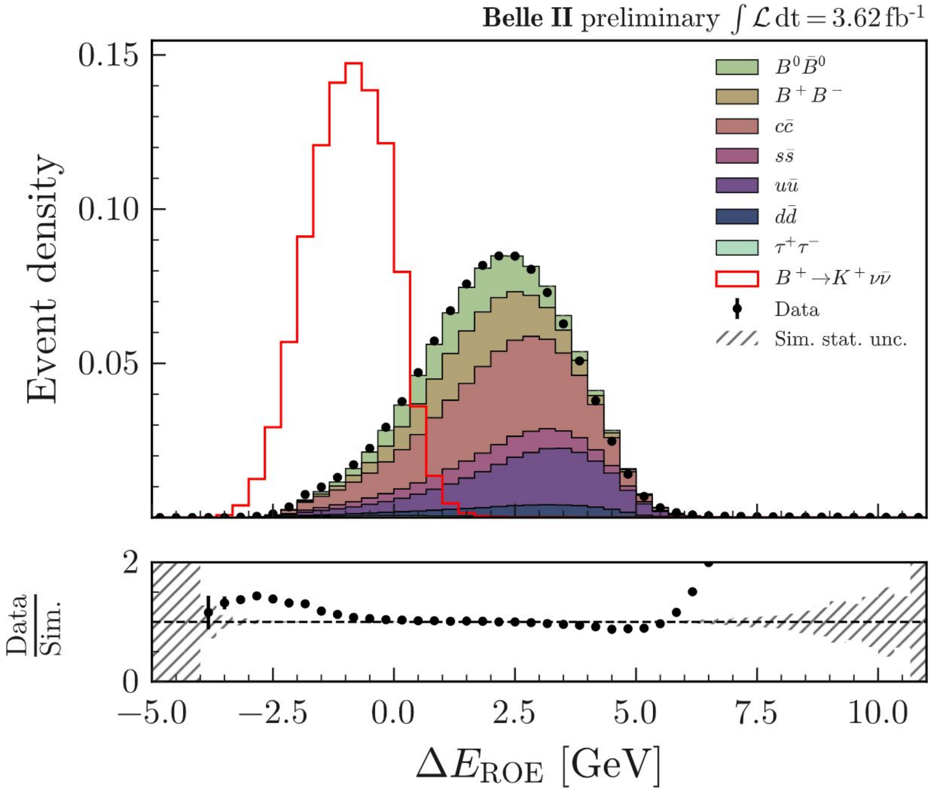
BDT$_1$ (first filter):
- 12 event-shape based kinematic variables

→

BDT$_2$ (final selection):
- *35 input variables*: using signal, event, and their correlations

- **Three-step filter:** basic event cuts, BDT-based filter (BDT$_1$) and final selection (BDT$_2$). BDT$_2$ improves performance in terms of s/√s+b by almost factor 3

Example of a discriminating variable

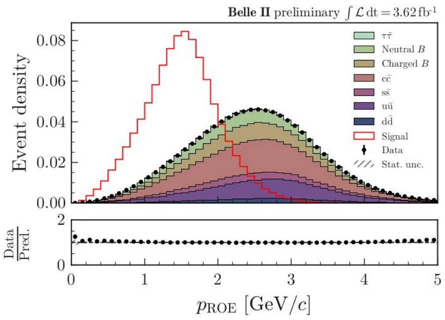


ΔE$_{\mathrm{ROE}}$ is a difference between the second (tag) B-meson energy and the center-of-mass energy/2. Expected to be at zero for fully reconstructed *B* meson decays

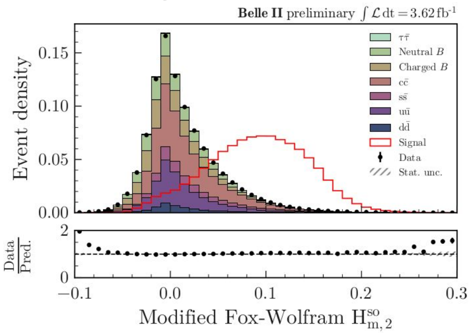
- Above zero for background: extra particles attached
- Below zero for signal: tag B may have missing particles (e.g. for semileptonic decays)

Examples of input variables for BDT_1 and BDT_2

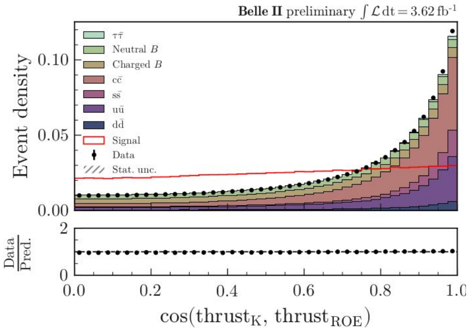
Recoil momentum



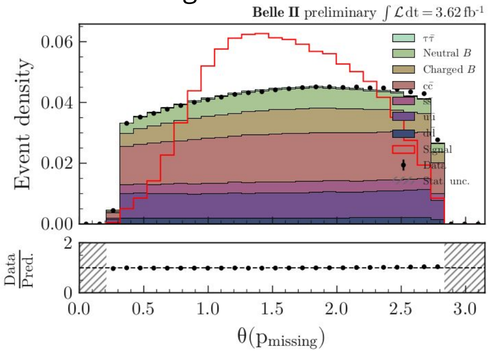
"Missing momentum"



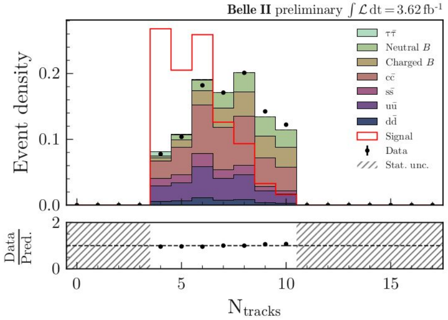
Spherical Jet-like



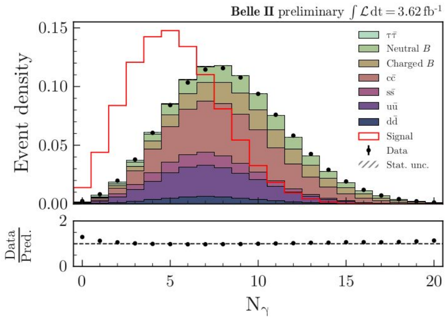
Missing momentum direction



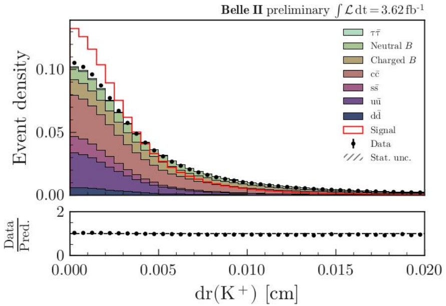
Number of tracks



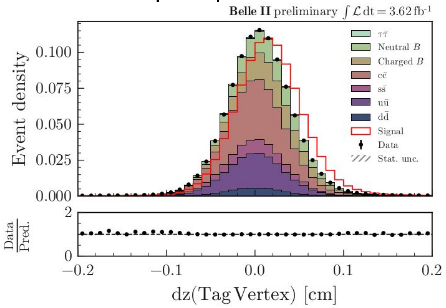
Number of neutrals



K^+ impact parameter

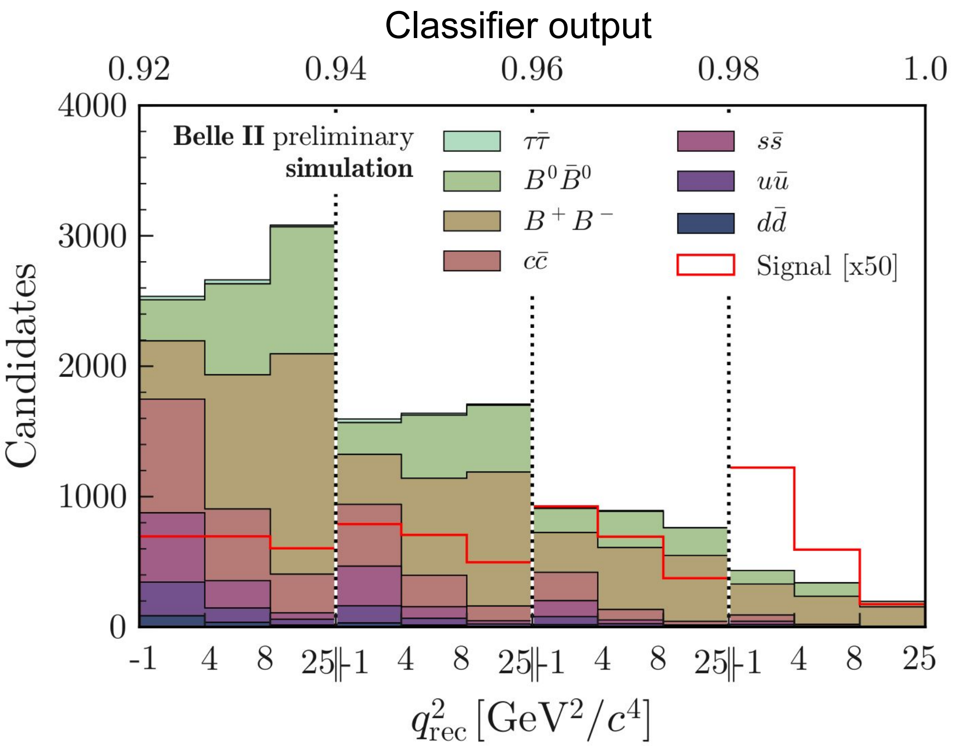


Recoil impact parameter



- Example of input distributions at pre-selection level, 1% of data, with detector-level corrections applied but no physics modeling corrections
- Each variable is examined to have reasonable description by simulation and significant separation power

Signal extraction

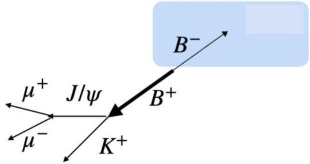


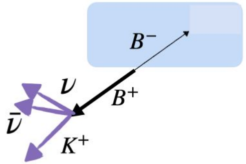
(3 bins in q^2_{rec}) x (4 bins in classifier output)

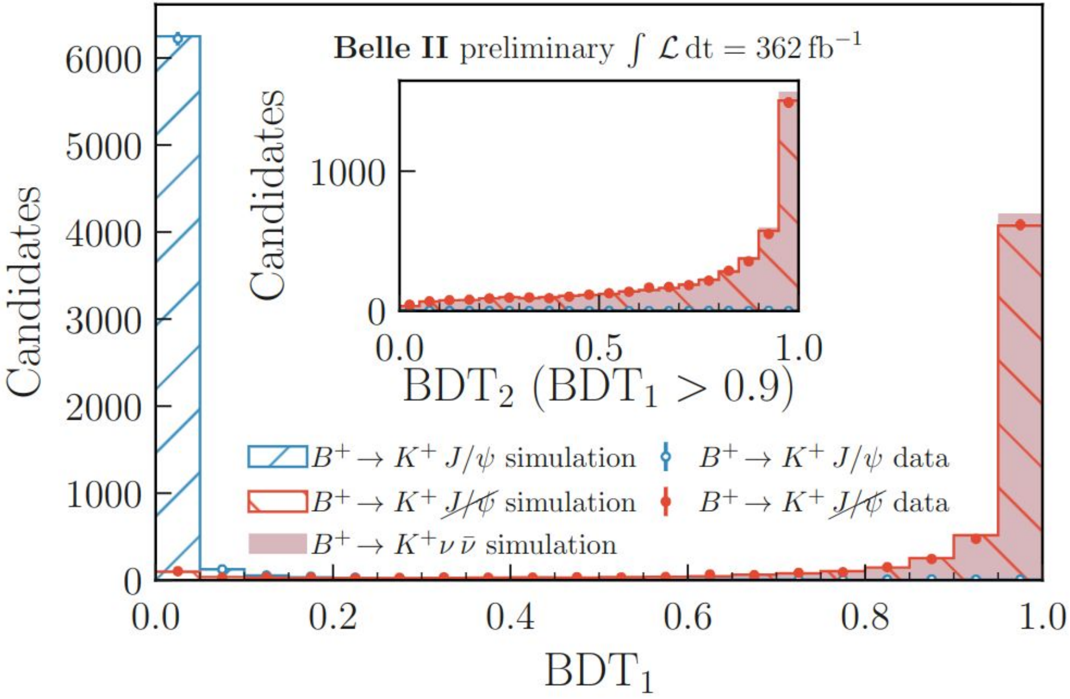
- Define the signal region at the plateau of the classifier sensitivity which corresponds to signal efficiency of **8%**
- Further subdivide it in 4 bins of classifier output and 3 bins in q^2_{rec}
- Binned profile maximum likelihood fit to data using **signal** and **7** background templates
- Systematic uncertainties varied in the fit

Main backgrounds are from charged and to lesser extent neutral *B* decays; continuum sources are checked/constraint using data taken below *Y(4S)* resonance.

Signal efficiency validation

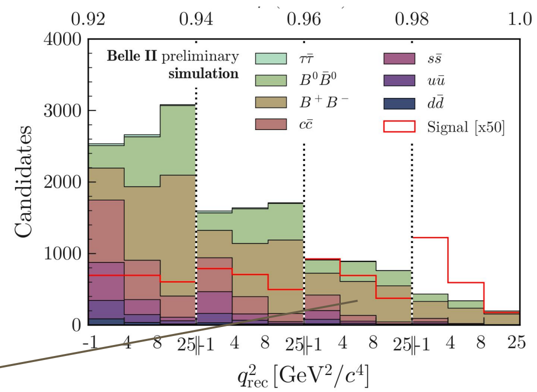


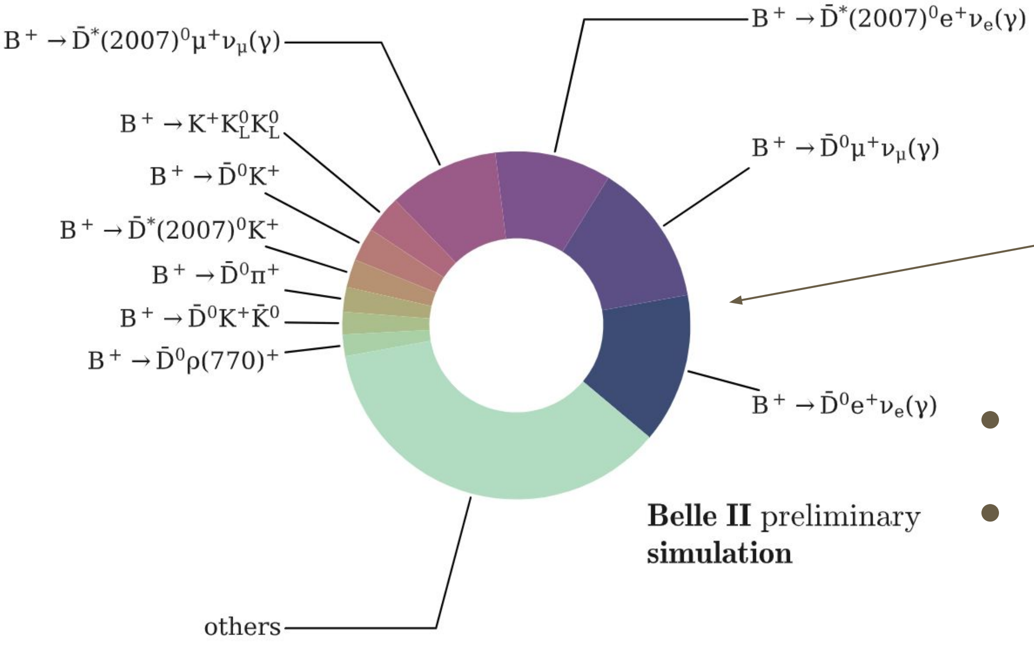




- Use cleanly reconstructed $B^+ \to K^+ J/\psi(\to \mu^+\mu^-)$ decays with $\mu^+\mu^-$ pair removed and K^+ kinematics adjusted (“signal embedding”) to validate the signal efficiency in simulation. The ratio of data/simulation efficiency in the signal region is **1.00±0.03**

Background composition

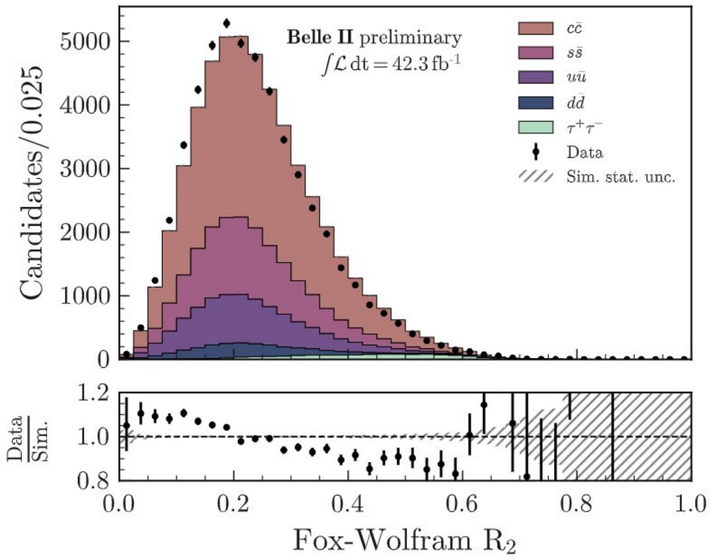


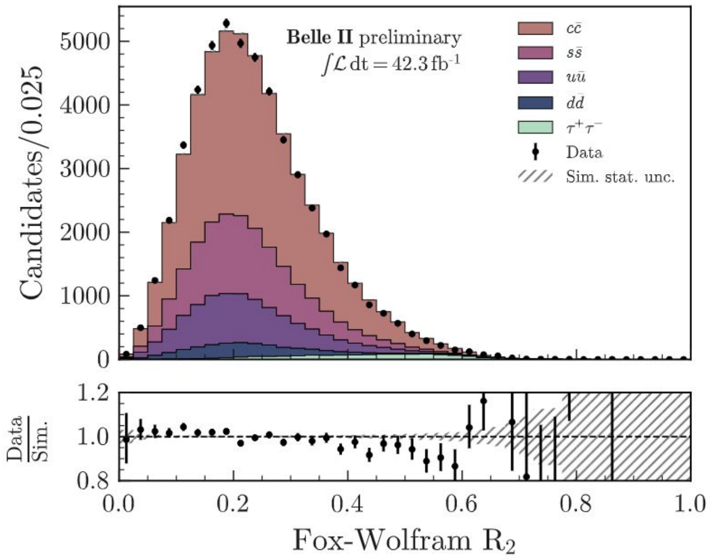


Charged (*B*+*B*-) background composition

- Main background from charged B^+B^- decays
- Main subcomponents are
  - Semileptonic (K^+ from D decays), **47%**
  - Hadronic involving charm ($B \to D^{(*)}K^+$), **38%**
  - Light hadronic, involving K_L , **14%**
  - $B \to \tau(K^+\nu)\nu$, $B \to K^*\nu\nu$, **1%**

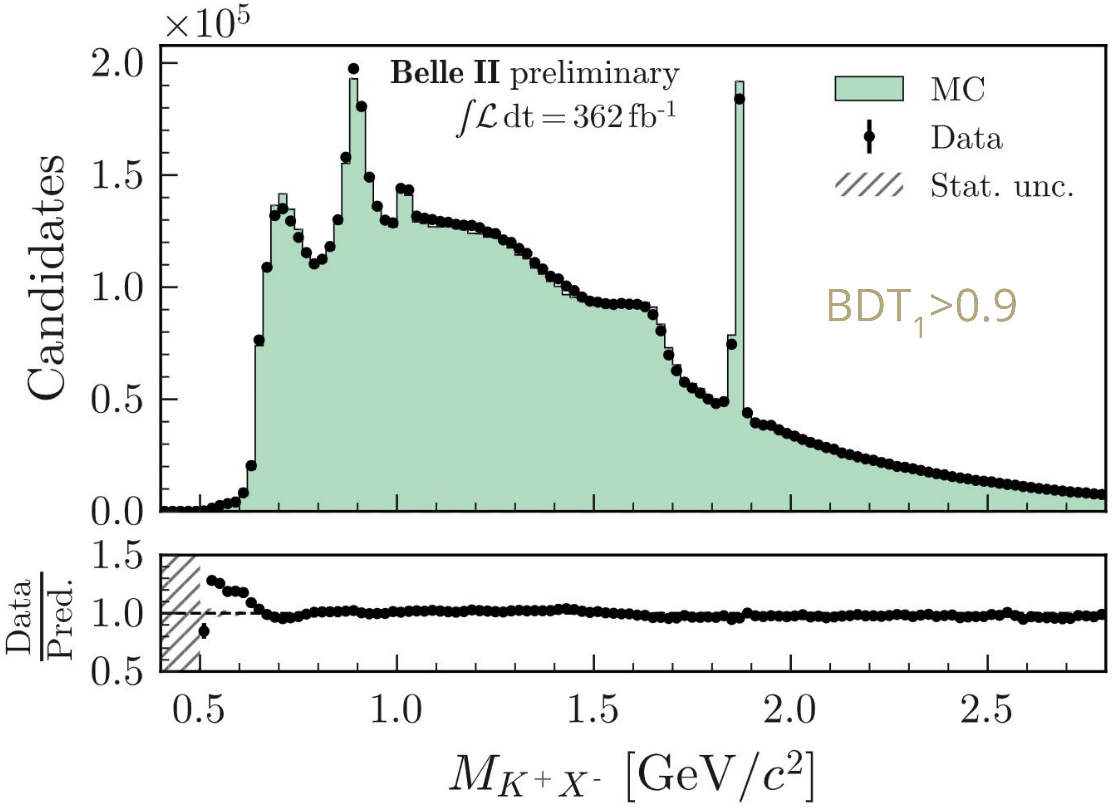
Continuum background





- Continuum *qq* background is checked using data taken below *Y*(4*S*) resonance
- Modeling is improved using a dedicated BDT$_c$ trained to distinguish between data and simulation; simulated events with large separation are suppressed by a weight = BDT$_c$/(1-BDT$_c$)
- While shapes are improved, there is a global 40% normalisation difference in the signal region between data and simulation (→ systematic uncertainty)

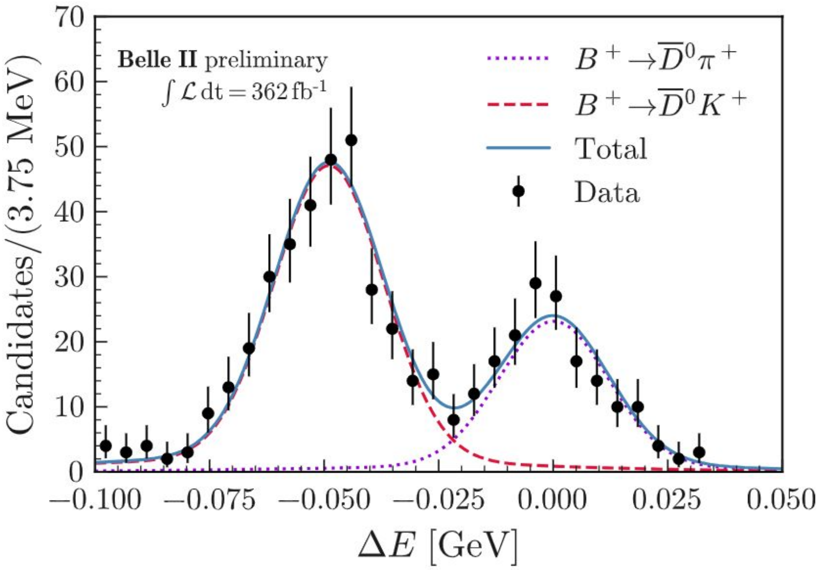
Background from $B \to D(\to K^+ X) l \nu$ (~50%)



X can be a pion, kaon or proton, depending on particle identification likelihoods

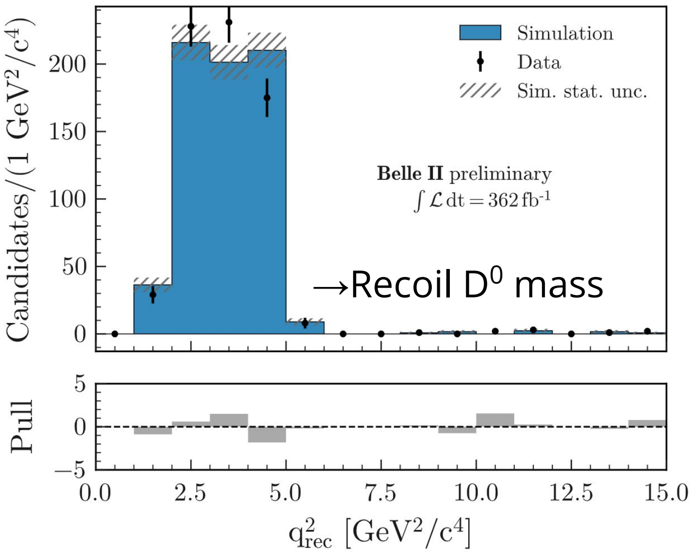
- Semileptonic decays suppressed by several MVA input variables, which match signal candidate K^+ with ROE particles, checked at each selection step
- Systematic uncertainties on decay branching fractions, enlarged for $B \to D^{**} l \nu$ decays
- In general, higher multiplicity decays contribute less to the background

Kaon identification check



Important background from $B^+ \to \pi^+ D^{(*)}$ with misidentified pions

$B^+ \to h^+ D^0(\to K^+ \pi^-)$ with $K^+\pi^-$ removed

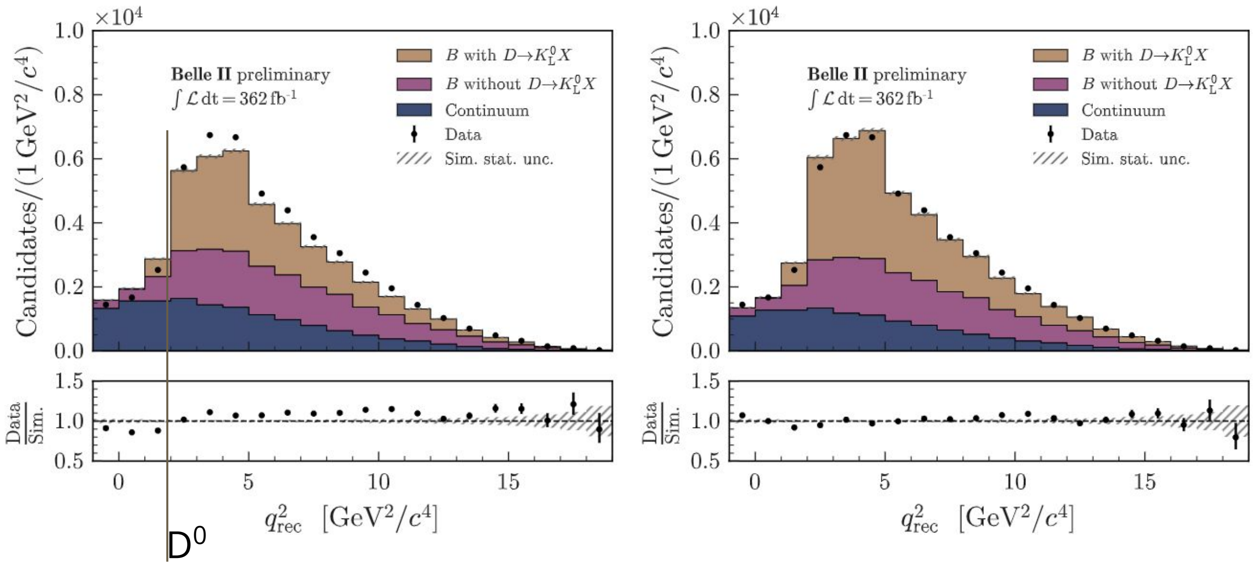


- Kaon identification efficiency and pion → kaon "fake" rate are determined using high-statistics low-background samples of $D^{*+} \to \pi^+ D^0(\to K^- \pi^+)$ decays
- Checked using $B^+ \to h^+ D^0(\to K^+ \pi^-)$ decays. Use $\Delta E = E^*_B - E_{beam}$ distribution to separate $h^+ = \pi^+$ vs K^+. Remove D^0 decay products to mimic the signal, select signal region in the analysis.

Pion sideband checks

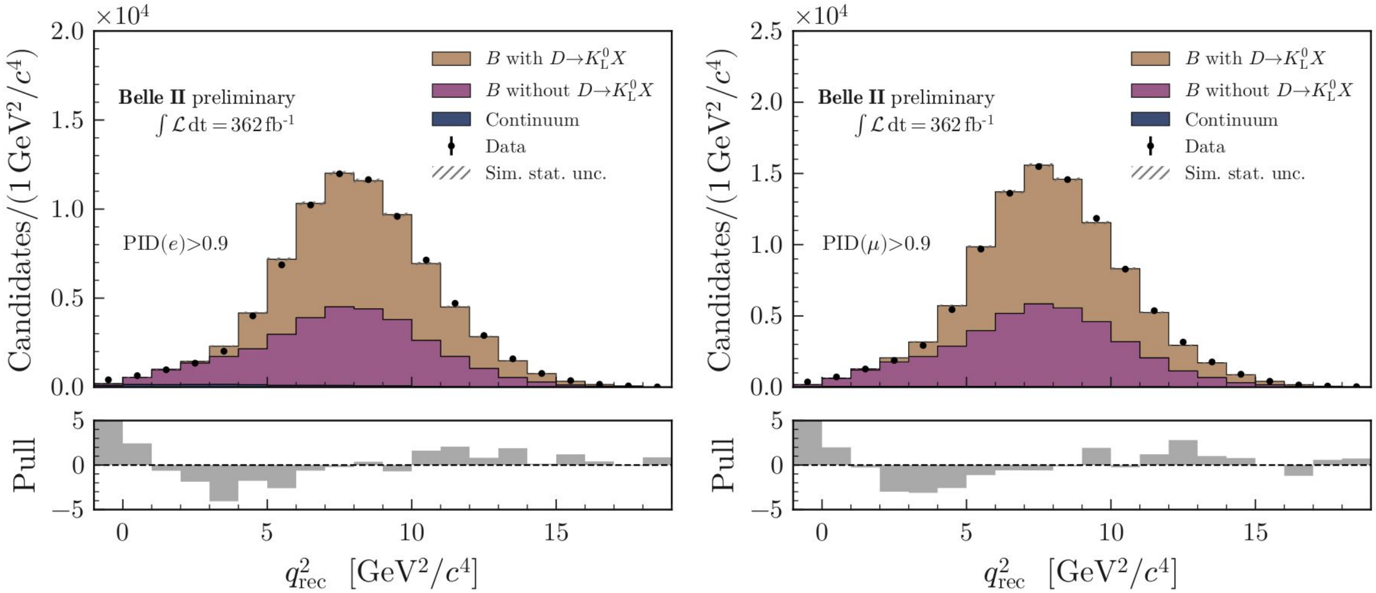
Check of $B \to K^+ D^{(*)}$ background

Fit



- Use reconstructed $B^+ \to \pi^+ X$ events (“pion sideband”) to check $B \to K^+ X$ background.
- Continuum background normalized using off-resonance data
- Excess observed at charm mass threshold, attributed to charm meson decays involving K_L
- Scaling factor of +(30+-2)% determined using 3-component fit to q^2_{rec} distribution
- Extra uncertainty for on/off-resonance data relative normalization of 5%

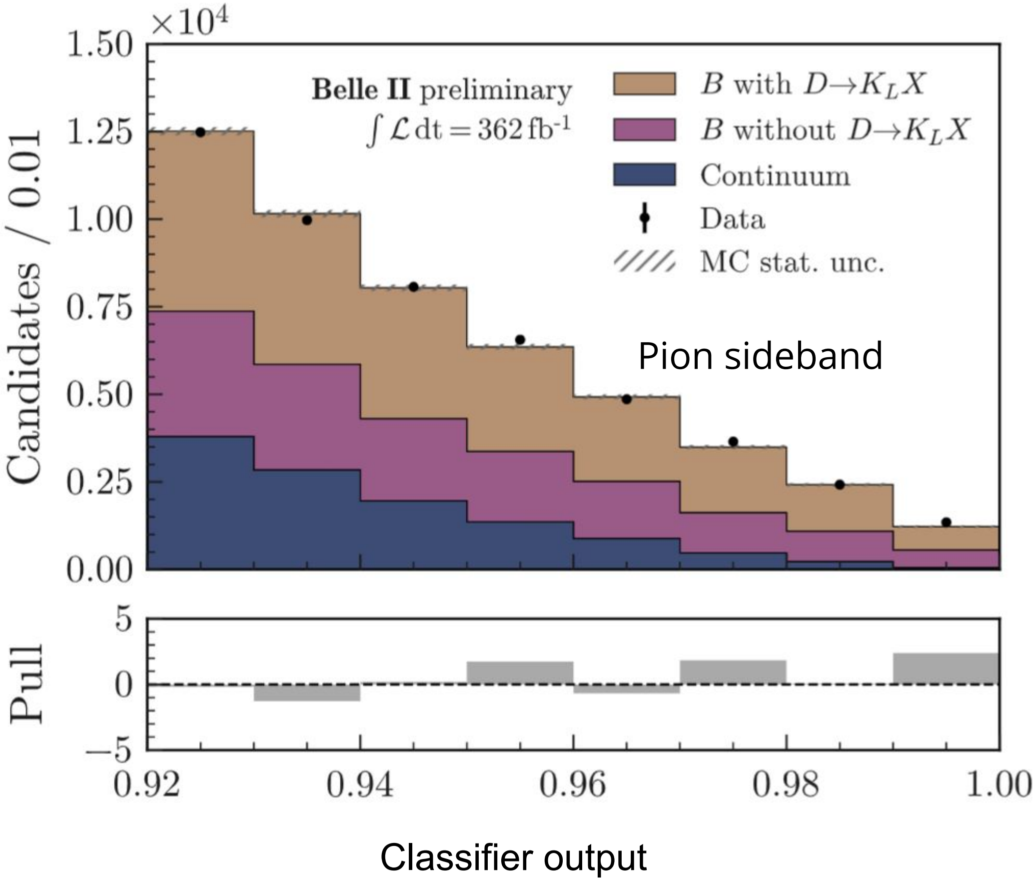
Lepton sideband checks



← Post-fit distributions

- Similar analysis performed using electron and muon instead of kaon identification. Scaling factors of (35+-1)% and (38+-1)% required.
- →correction of (30 +- 10)% applied for the nominal sample.

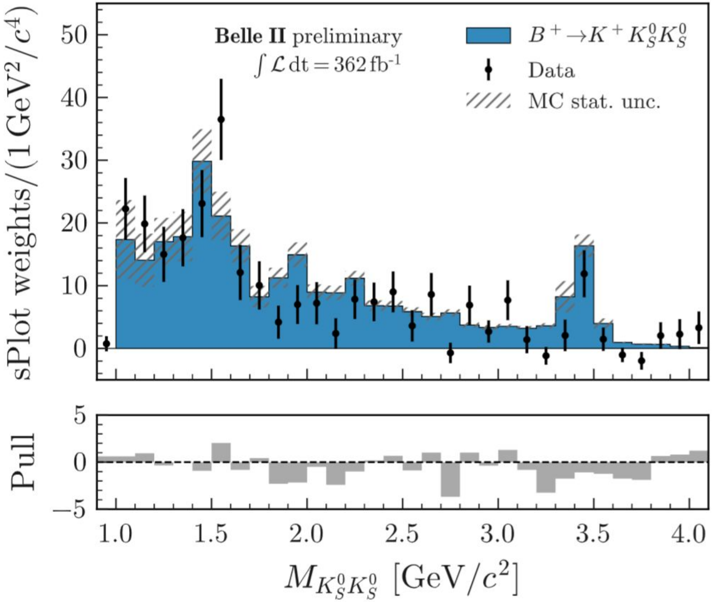
Classifier output for pion sideband



- Good description of the BDT_2 classifier output distribution, post-fit to $q2_{rec}$ distribution.

Background from $B^+ \to K^+ K^0 K^0$

Most signal-like backgrounds



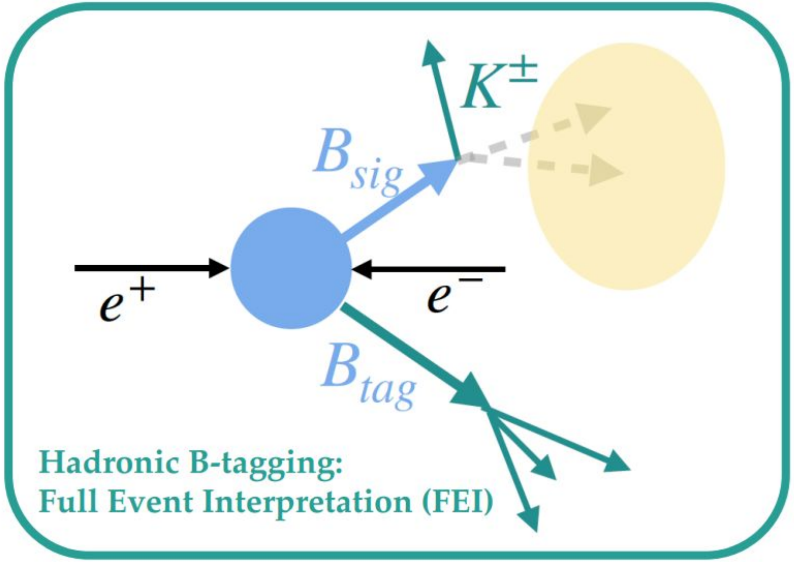
$\leftarrow B^+ \to K^+ K_S K_S$ decays

- Backgrounds from $B^+ \to K^+ nn$ and $B^+ \to K^+ K^0 K^0$ have branching fractions of few x 10^{-5}, however K_L and neutrons can escape EM calorimeter
- $B^+ \to K^+ K^0 K^0$ modeled based on BaBar analysis (arXiv:1201.5897)
- Dedicated checks of K_L's performance in calorimeter using radiative φ production
- Dedicated checks using $B^+ \to K^+ K_S K_S$ and $B^0 \to K_S K^+ K^-$ control channels

Systematic uncertainties of the inclusive analysis

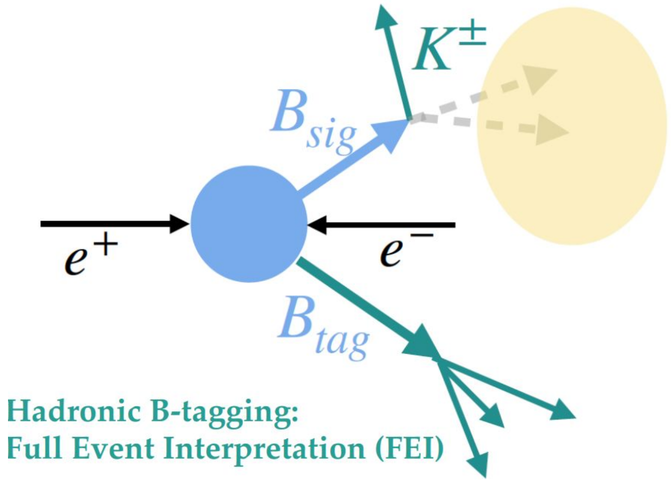
Source	Correction	Uncertainty type, parameters	Uncertainty size	Impact on σ_μ
Normalization of $B\bar{B}$ background	—	Global, 2	50%	0.90
Normalization of continuum background	—	Global, 5	50%	0.10
Leading B-decay branching fractions	—	Shape, 5	$O(1\%)$	0.22
Branching fraction for $B^+ \to K^+ K^0_{\mathrm{L}} K^0_{\mathrm{L}}$	q^2 dependent $O(100\%)$	Shape, 1	20%	0.49
p-wave component for $B^+ \to K^+ K^0_{\mathrm{S}} K^0_{\mathrm{L}}$	q^2 dependent $O(100\%)$	Shape, 1	30%	0.02
Branching fraction for $B \to D^{**}$	—	Shape, 1	50%	0.42
Branching fraction for $B^+ \to K^+ n\bar{n}$	q^2 dependent $O(100\%)$	Shape, 1	100%	0.20
Branching fraction for $D \to K^0_{\mathrm{L}} X$	+30%	Shape, 1	10%	0.14
Continuum-background modeling, $\mathrm{BDT_c}$	Multivariate $O(10\%)$	Shape, 1	100% of correction	0.01
Integrated luminosity	—	Global, 1	1%	< 0.01
Number of $B\bar{B}$	—	Global, 1	1.5%	0.02
Off-resonance sample normalization	—	Global, 1	5%	0.05
Track-finding efficiency	—	Shape, 1	0.3%	0.20
Signal-kaon PID	p, θ dependent $O(10-100\%)$	Shape, 7	$O(1\%)$	0.07
Photon energy	—	Shape, 1	0.5%	0.08
Hadronic energy	−10%	Shape, 1	10%	0.37
K^0_{L} efficiency in ECL	−17%	Shape, 1	8%	0.22
Signal SM form-factors	q^2 dependent $O(1\%)$	Shape, 3	$O(1\%)$	0.02
Global signal efficiency	—	Global, 1	3%	0.03
Simulated-sample size	—	Shape, 156	$O(1\%)$	0.52

Hadronic tag



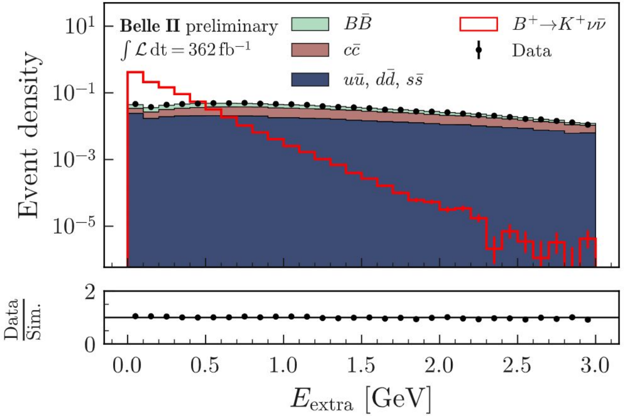
Hadronic tag selection

- Reconstruct Btag in one of 35 hadronic decay modes using FEI algorithm (arxiv:2008.06096)
- Require good B_{tag} reconstruction with a good p-value
- Same requirements for the signal K^{+} as inclusive analysis
- Similar event requirements; additional cut on opposite charge of B_{tag} and K^{+}
- BDT to further suppress background

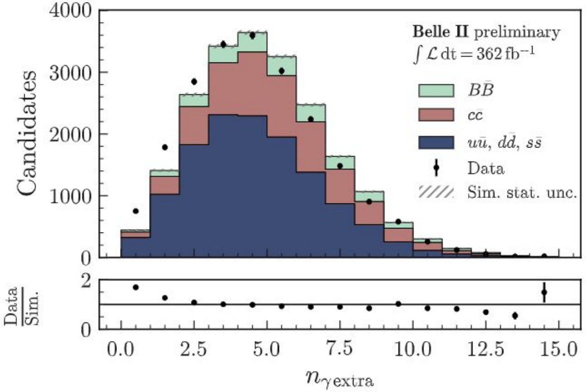


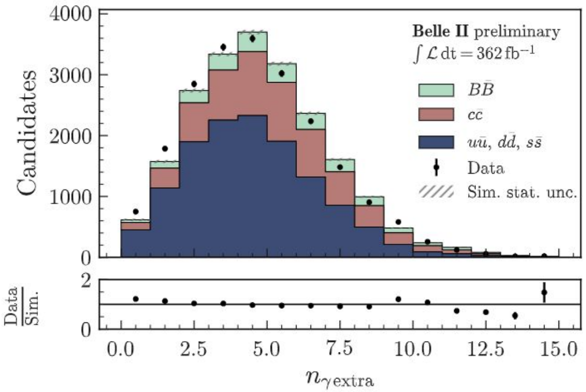
Hadronic tag selection

Before BDT



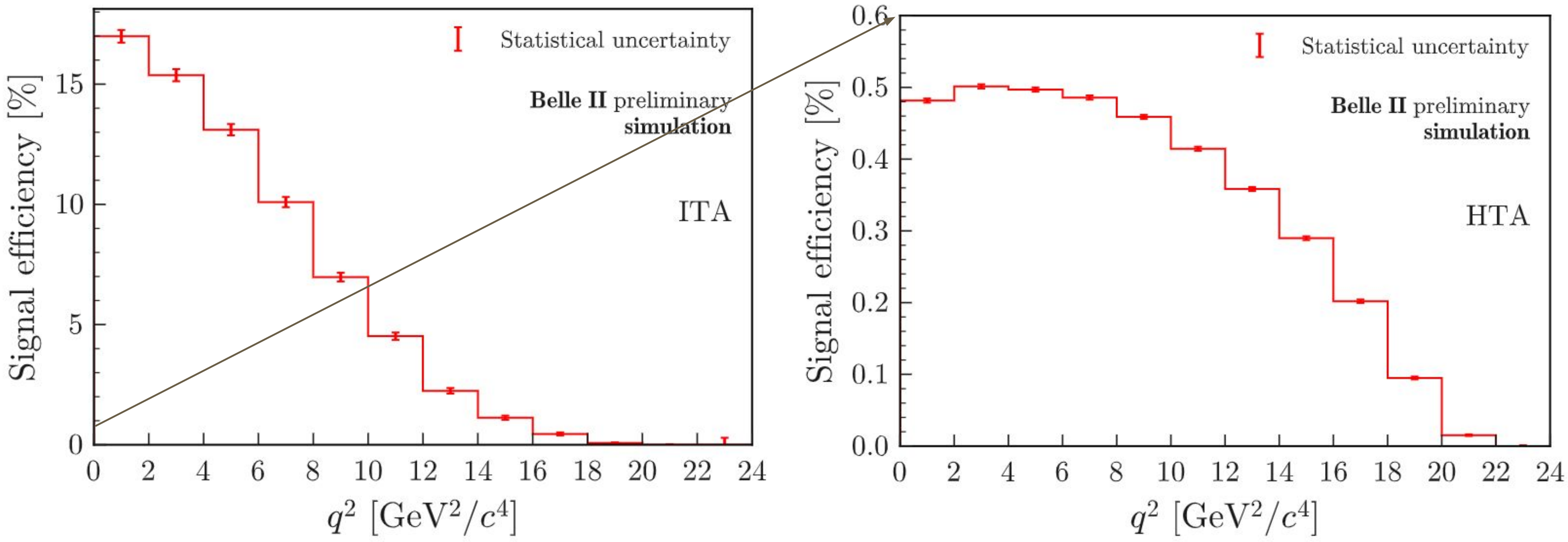
Pion sideband, before and after correction





- Most discriminating variable is E_{extra}: extra energy in the calorimeter vs the signal K^+ and B^-_{tag}
- Dedicated correction using photon multiplicity distribution in data sidebands
- Residual difference is included as systematics

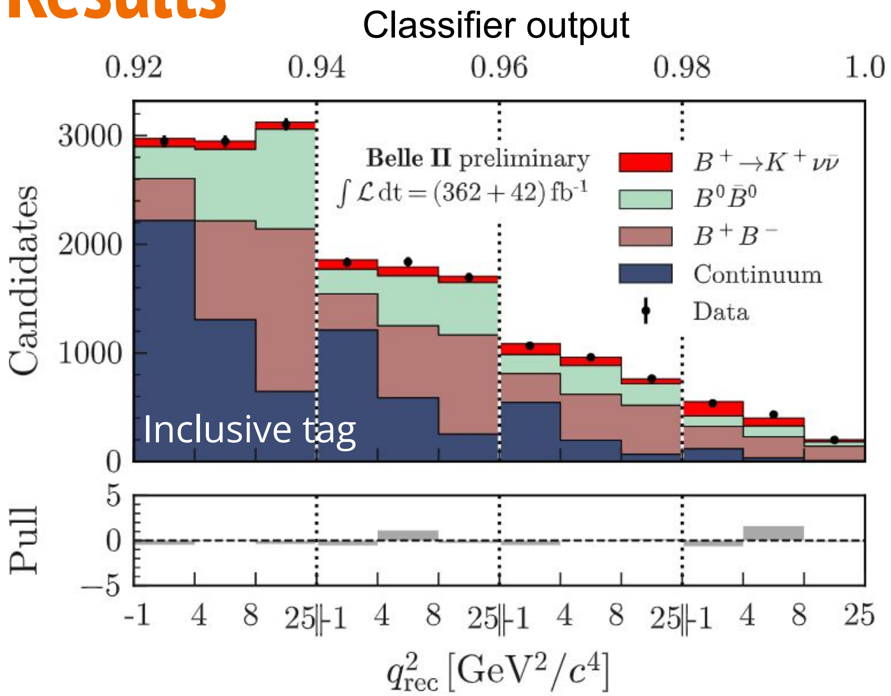
Signal efficiency for the inclusive and hadronic analyses

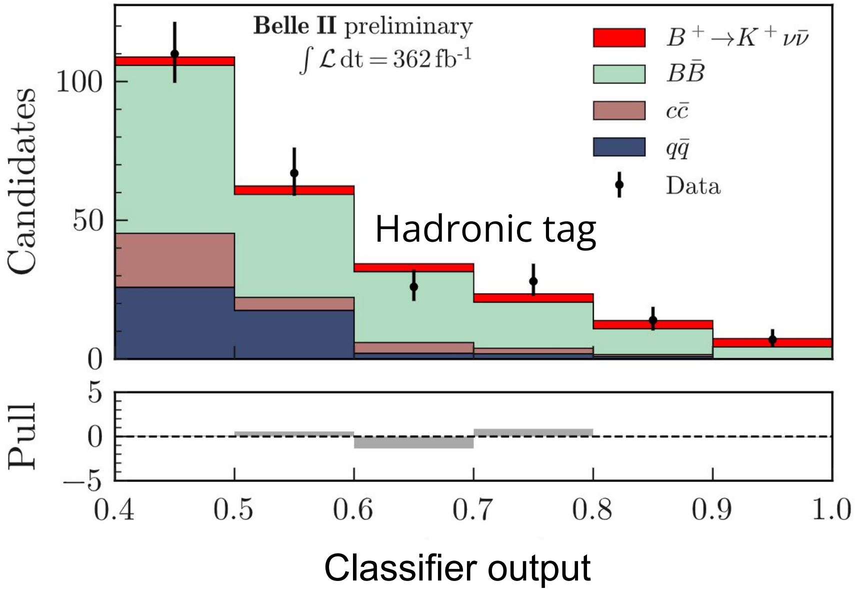


- Much lower efficiency for the hadronic tag analysis, however flatter q^2 distribution

Results

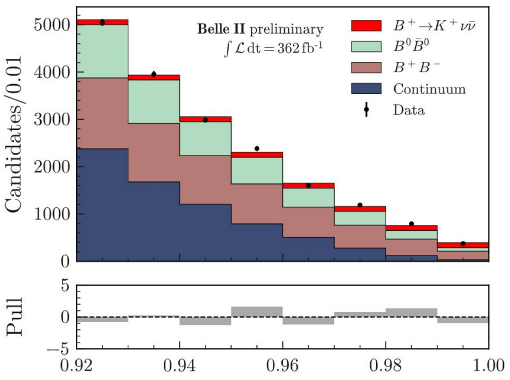
Results

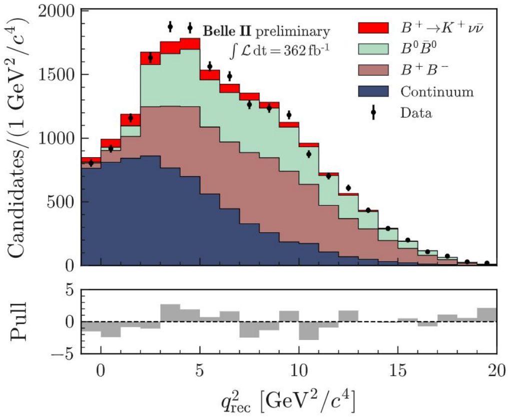




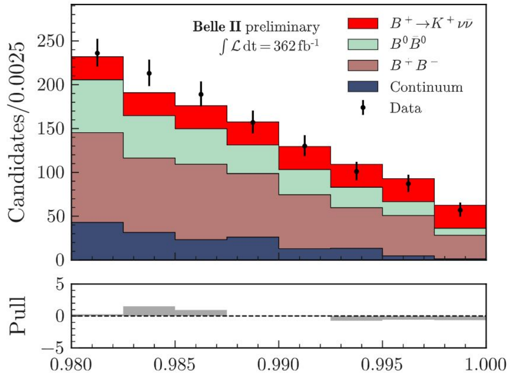
- Maximum likelihood fit to data using signal and background templates
- Branching fractions: $B_{\text{incl.}}=(2.7\pm0.5(\text{stat})\pm0.5(\text{stat}))\times10^{-5}$, $B_{\text{had.}}=(1.1^{+0.9}_{-0.8}(\text{stat})^{+0.8}_{-0.5}(\text{syst}))\times10^{-5}$
- For inclusive analysis, **evidence for $B\to K\nu\bar{\nu}$** at 3.5σ, branching fraction within 2.9σ of standard model (both considering total uncertainty)
- For hadronic tag, the result is consistent with null hypothesis and SM at 1.1σ and 0.6σ

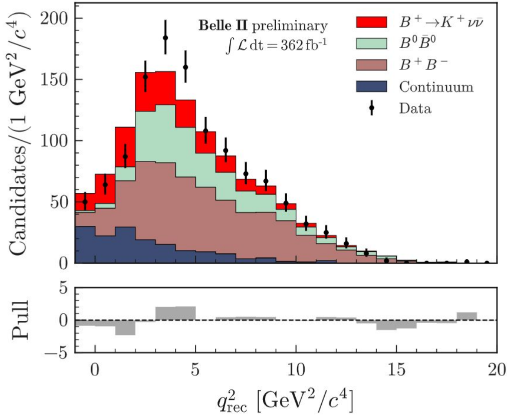
Post-fit distributions for inclusive analysis





Classifier output

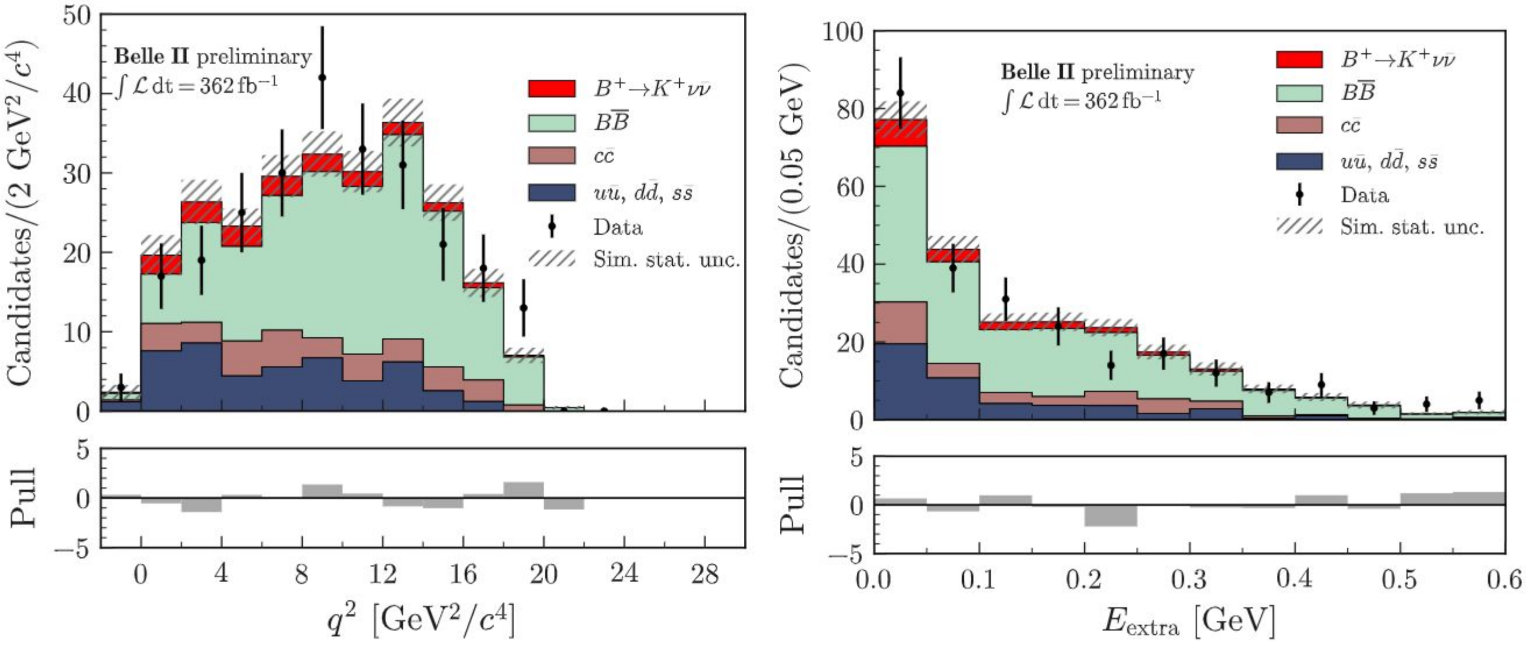




Classifier output

- Post-fit distributions for the inclusive analysis shown for the signal region and separately for the region with maximal sensitivity, classifier output>0.98
- Possible shape differences for q^2_{rec}, however given that fit is performed using coarse binning, not conclusive

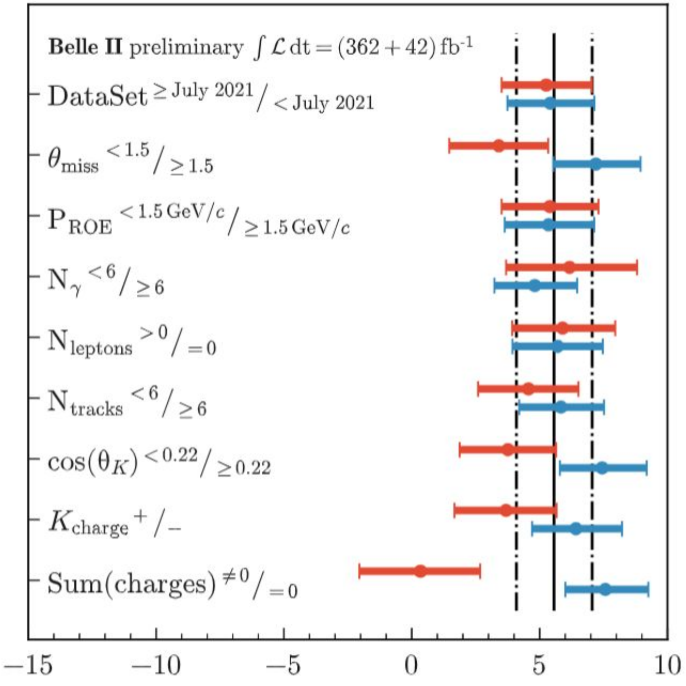
Post-fit distributions for hadronic tag analysis



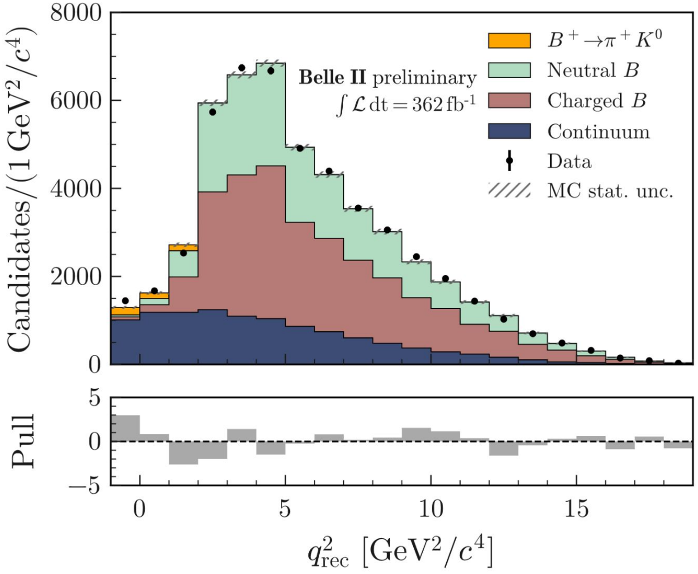
- Good description of q^2 (calculating using tagged-B) and E_{extra} distributions

Cross checks

Inclusive



μ =B/B$_{SM}$

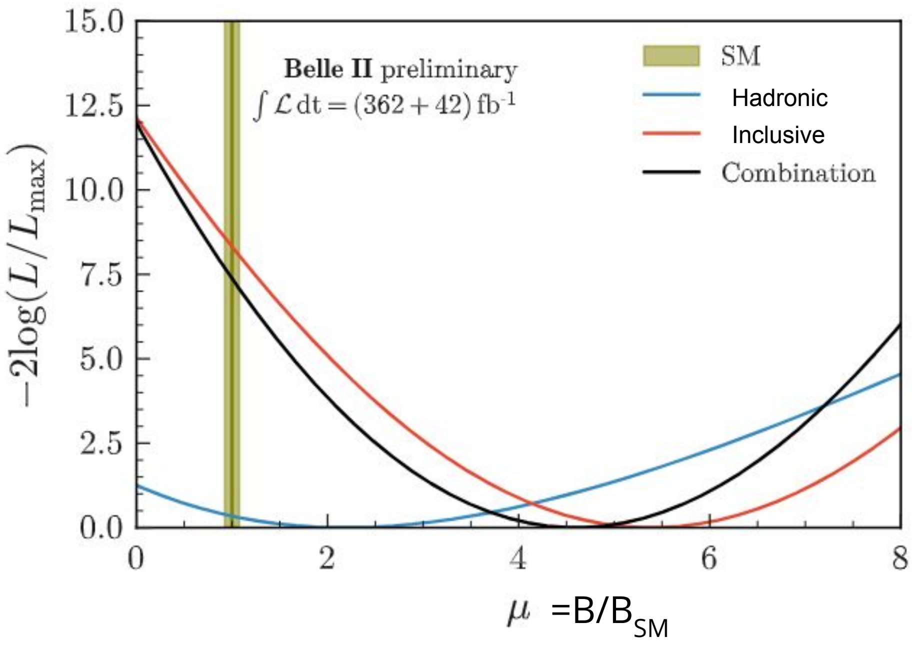


$\mathcal{B}(B^+ \to \pi^+ K^0) = (2.5 \pm 0.5) \times 10^{-5}$

PDG: $(2.38 \pm 0.08) \times 10^{-5}$

- Multiple checks of the analyses stability, including tests dividing data into approximately equal sub-samples. Reported here as measured branching fraction divided by SM expectation, μ=B/B$_{SM}$.
- Control measurement of $B^+ \to \pi^+ K^0$ decay

Combination of the inclusive and hadronic tag



Full pyhf model of the data will be released when paper is accepted by the Journal

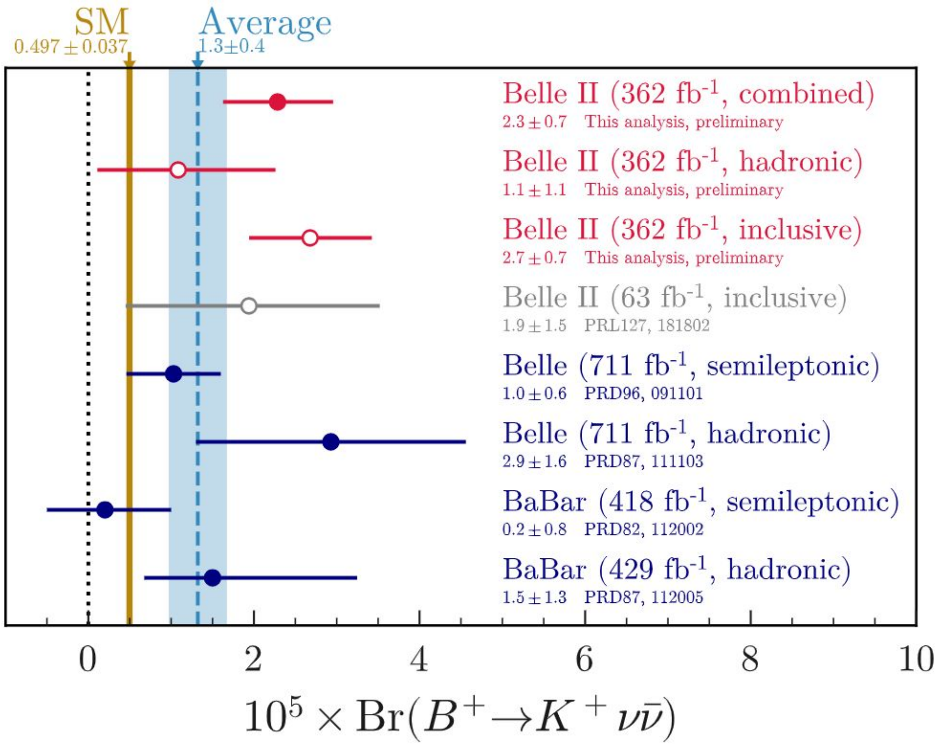
Inclusive and hadronic measurements are combined, taking into account common correlated uncertainties. The resulting branching fraction is

$$B_{comb}(B^+ \to K^+ \nu\nu) = (2.3 \pm 0.7) \times 10^{-5} = [2.3 \pm 0.5(\mathrm{stat})^{+0.5}_{-0.4}(\mathrm{syst})] \times 10^{-5}$$

Significance of observation is **3.5σ** the result is within **2.7σ** vs standard model

Comparison with other measurements

*



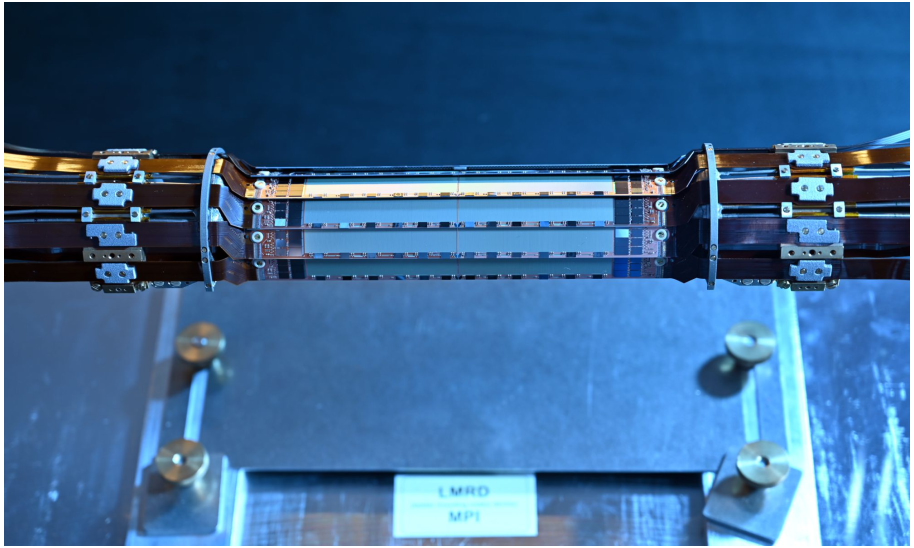
- Belle reports upper limits only: for this comparison branching fractions are computed using published number of events and efficiency
- Inclusive result has comparable accuracy to previous best measurements, hadronic tag is the best among hadronic tag measurements
- Some tensions between inclusive and semileptonic results for Belle and BaBar, however overall compatibility of the results is good with **χ^2/dof = 5.6/5**

Summary and Outlook

- Two independence analyses using inclusive and hadronic tagging methods
- **Evidence** for ***$B^+ \to K^+ \nu\bar{\nu}$*** decay with a branching fraction **2.7σ** above the standard model
- More results expected:
  - Inclusive tag analysis using Belle data sample
  - Several analyses improvements in the pipeline, to be used with more data
  - Other channels: $K^*\nu\nu$, $K_S\nu\nu$

Backup

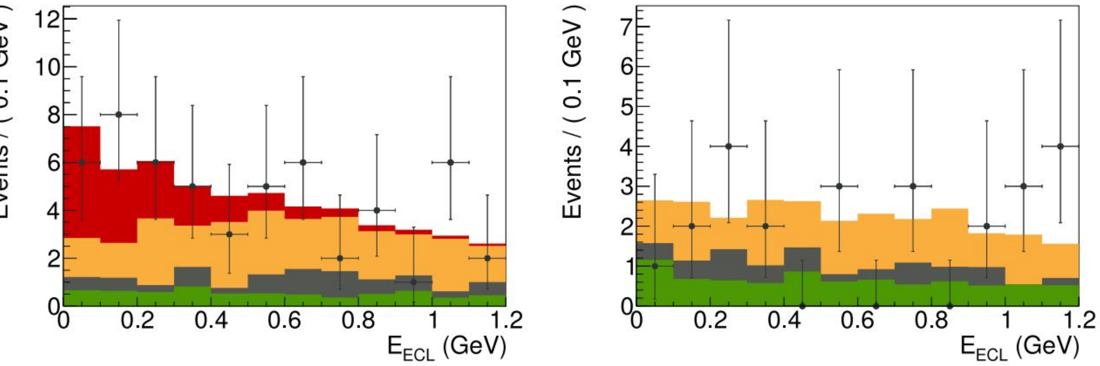
Upgrades during LS1



Since summer 2022 data taking, SuperKEKB and Belle II are in LS1, until fall 2023. Several improvements for accelerator complex, to reduce background and improve luminosity, and detector upgrades, such as installation of the complete two-layer vertex pixel detector. Details on SuperKEKB upgrade can be found [here]

Current limits for B → K(*)*νν*

From arXiv:1702.03224, Belle



(c) $B^+ \to K^{*+}\nu\bar{\nu}$ (d) $B^0 \to K^{*0}\nu\bar{\nu}$

Channel	Efficiency	Expected limit	Observed limit
$K^+\nu\bar{\nu}$	2.16×10^{-3}	0.8×10^{-5}	1.9×10^{-5}
$K_S^0\,\nu\bar{\nu}$	0.91×10^{-3}	1.2×10^{-5}	1.3×10^{-5}
$K^{*+}\nu\bar{\nu}$	0.57×10^{-3}	2.4×10^{-5}	6.1×10^{-5}
$K^{*0}\nu\bar{\nu}$	0.51×10^{-3}	2.4×10^{-5}	1.8×10^{-5}

- The best limit for $B^0 \to K^{*0}\nu\nu$ is obtained by Belle semileptonic tag analysis, and it is used by PDG.
- The same paper reports 2.3σ fluctuation for $B^+ \to K^{*+}\nu\nu$ leading to a significantly worse limit
- The paper also reports isospin averaged limit, $B(B \to K^*\nu\nu) < 2.7 \times 10^{-5}$ (90% cl), which is probably a better guess of the current status