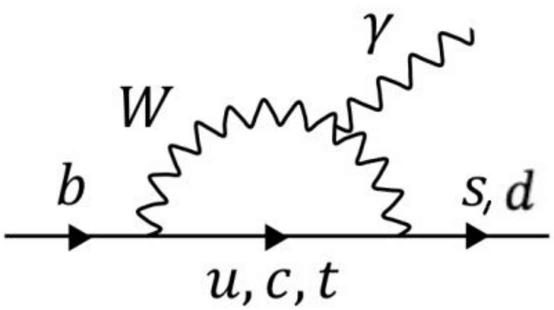
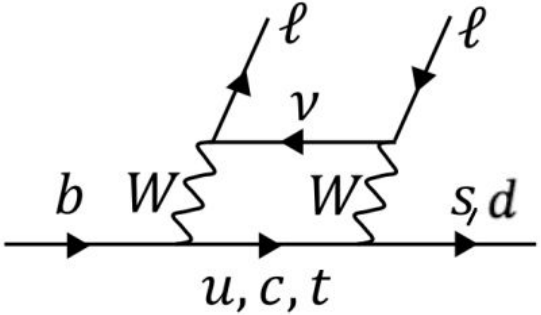
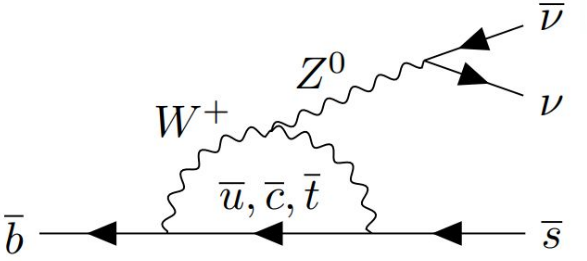
Studies of b⟶s transitions with Belle II

Discrete-2024, 2 Dec 2024
Sasha Glazov, on behalf of Belle II

Study of b⟶s transitions: motivation







- Flavor-changing neutral-current *b ⟶ s(d)* transitions are suppressed in SM, while many BSM theories predict significant contributions
- Several observables have clean theoretical predictions and show tensions between data and SM
- Channels involving 3rd generation interesting in particular due to e.g. connections to anomalies in semi-taunic decays ("$R(D^{(*)})$")

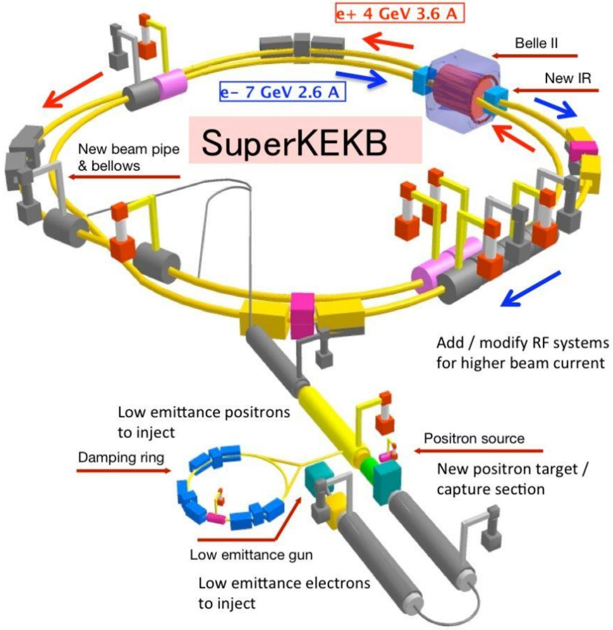
Belle II and SuperKEKB

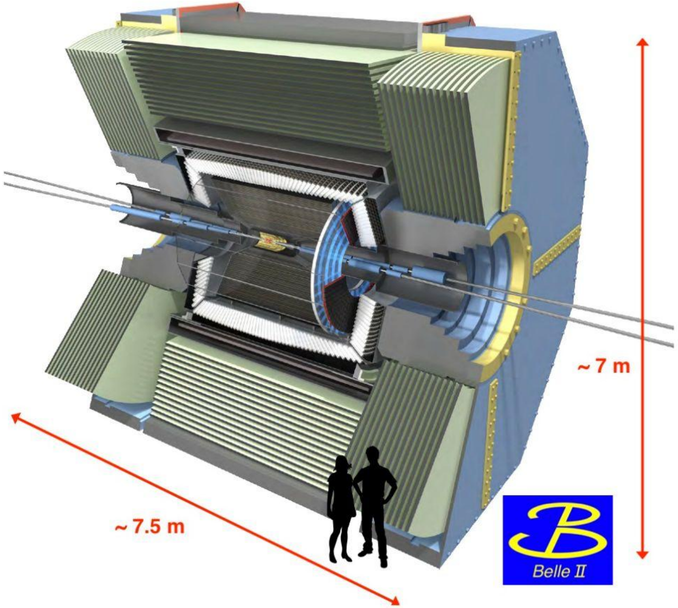
Collected in total:	551 fb^{-1}

Today's talk: Run I Belle II sample: 0.4×10^9 BB plus complete Belle sample.

SuperKEKB:

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



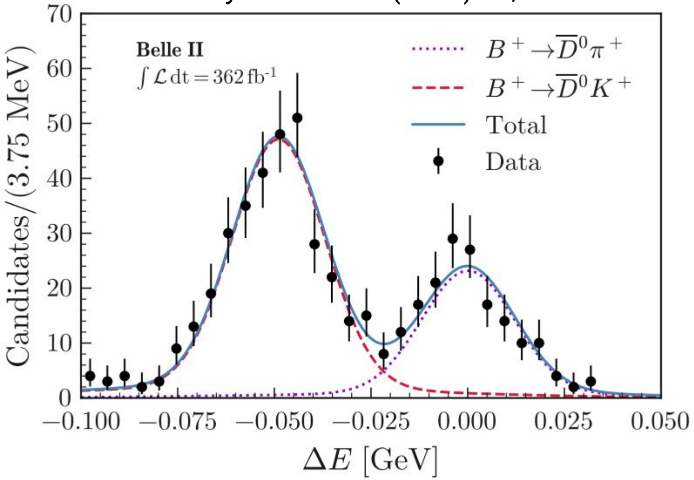


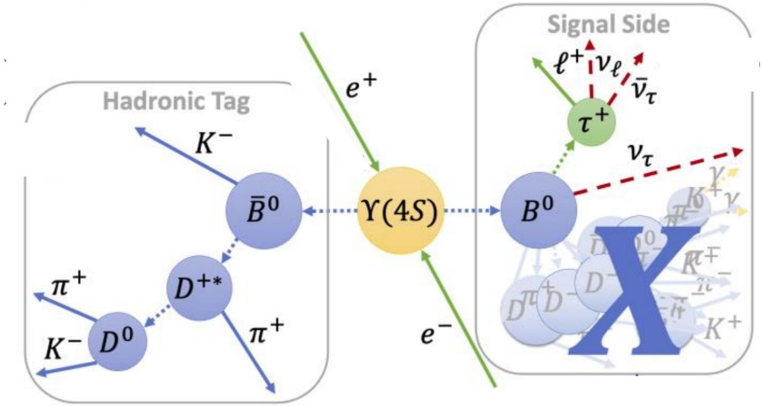
Belle II:

- 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

Experimental techniques at B-factories

From Phys.Rev.D 109 (2024) 11, 112006





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

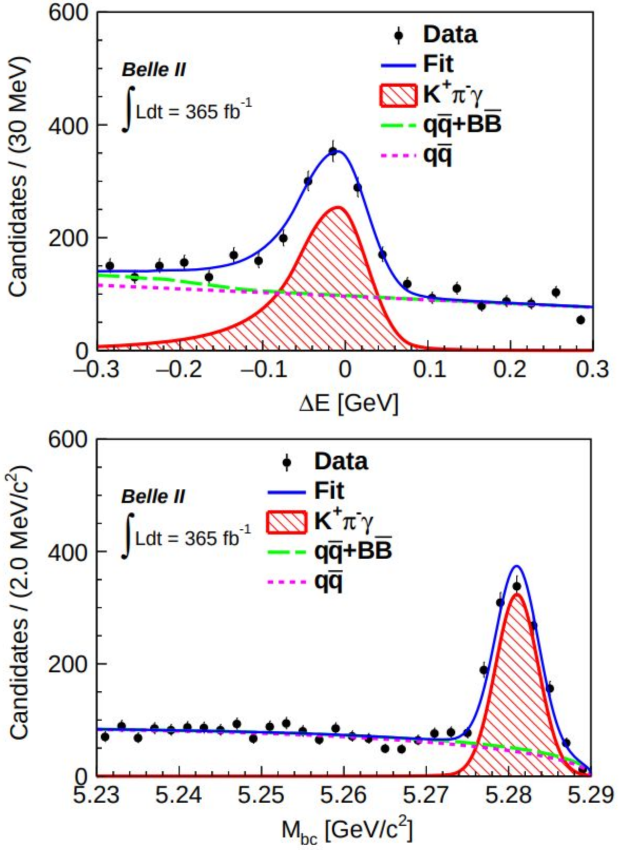
$$M_{\rm bc} = \sqrt{s/(4c^4) - {p_B^*}^2/c^2}$$

- Kinematic constraints using know center of mass energy
- *B*-meson tagging using hadronic or semileptonic *B*-meson decays
- ML-based continuum, π^0 background suppression
- Dedicated off-resonance data taking periods for continuum background studies

Measurement of B →K*γ

arXiv:2411.10127

- Measurement of $B^{+,0} \rightarrow K^{*+,0}\gamma$ using Run I Belle II data
- $K^{*0} \rightarrow K^+\pi^-$, $K^{*0} \rightarrow K_S\pi^0$, $K^{*+} \rightarrow K^+\pi^0$, $K^{*+} \rightarrow K_S\pi^+$ modes considered.
- Dominant background from continuum with $\pi^0(\eta) \rightarrow \gamma\gamma$ faking the prompt photon
- Dedicated MVA to suppress $\pi^0(\eta)$ background and continuum
- 2D unbinned fit in M_{bc}, ΔE
- Large signal, with moderate background → precision measurement.



Measurement of B →K*γ

- Comparable statistical and systematic uncertainties for the measured branching fractions
- Dominant systematics from π^0 reconstruction efficiency (3.9%)
- For CP and isospin asymmetry, statistical uncertainties dominate
- Isospin asymmetry is consistent with the SM expectations and previous measurements from Belle and BaBar.

$$\mathcal{B} = \frac{N_S/\epsilon_S + N_{\overline{S}}/\epsilon_{\overline{S}}}{2 \times N_{B\overline{B}} \times f_{+-}(f_{00})},$$

$$\mathcal{A}_{CP} = \frac{N_S/\epsilon_S - N_{\overline{S}}/\epsilon_{\overline{S}}}{N_S/\epsilon_S + N_{\overline{S}}/\epsilon_{\overline{S}}},$$

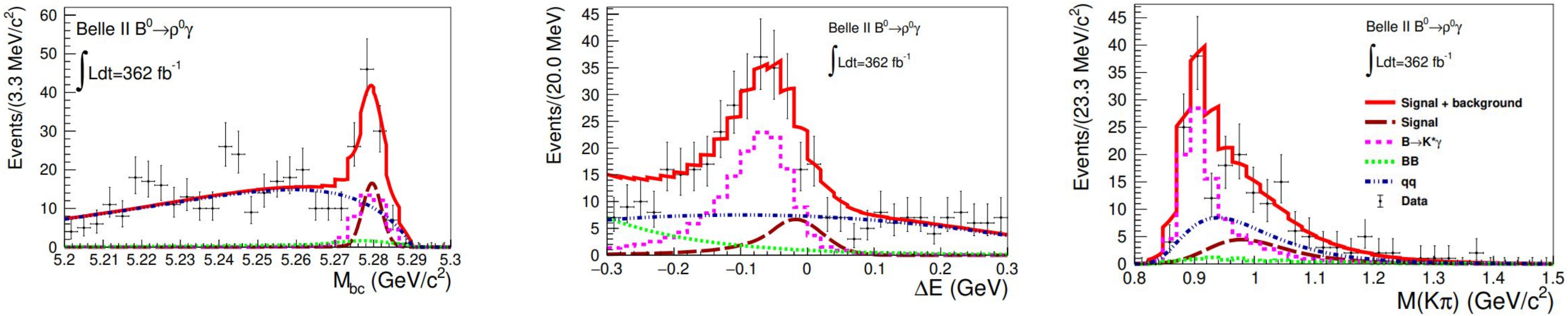
$$\Delta_{0+} = \frac{(\tau_+/\tau_0) \times \mathcal{B}(B^0 \to K^{*0}\gamma) - \mathcal{B}(B^+ \to K^{*+}\gamma)}{(\tau_+/\tau_0) \times \mathcal{B}(B^0 \to K^{*0}\gamma) + \mathcal{B}(B^+ \to K^{*+}\gamma)},$$

$$\Delta\mathcal{A}_{CP} = \mathcal{A}_{CP}(B^+ \to K^{*+}\gamma) - \mathcal{A}_{CP}(B^0 \to K^{*0}\gamma),$$

Channel	$\mathcal{B}$ (10^{-5})	$\mathcal{A}_{CP}$ (%)
$B^0 \to K^{*0}[K^+\pi^-]\gamma$	$4.14 \pm 0.10 \pm 0.11$	$-3.3 \pm 2.3 \pm 0.4$
$B^0 \to K^{*0}[K^0_S\pi^0]\gamma$	$4.07 \pm 0.33 \pm 0.23$	$-$
$B^0 \to K^{*0}\gamma$	$4.14 \pm 0.10 \pm 0.10$	$-3.3 \pm 2.3 \pm 0.4$
$B^+ \to K^{*+}[K^+\pi^0]\gamma$	$3.97 \pm 0.17 \pm 0.20$	$+1.7 \pm 4.0 \pm 0.9$
$B^+ \to K^{*+}[K^0_S\pi^+]\gamma$	$4.06 \pm 0.18 \pm 0.13$	$-3.5 \pm 4.3 \pm 0.7$
$B^+ \to K^{*+}\gamma$	$4.02 \pm 0.13 \pm 0.13$	$-0.7 \pm 2.9 \pm 0.6$
	Δ_{0+} (%)	$\Delta\mathcal{A}_{CP}$ (%)
$B \to K^*\gamma$	$+5.0 \pm 2.0 \pm 1.0 \pm 1.1$	$+2.6 \pm 3.8 \pm 0.7$

Measurement of B →ργ

arXiv:2407.08984



- Measurement using combined Belle+Belle II (run I) data sample
- Dominant backgrounds from B →K*γ and continuum
- 3D unbinned fit in M_{bc}, ΔE and $M(K\pi)$
- Isospin and CP asymmetries are consistent with zero.

$$\mathcal{B}\left(B^+ \to \rho^+\gamma\right) = \left(13.1^{+2.0+1.3}_{-1.9-1.2}\right) \times 10^{-7}$$

$$\mathcal{B}\left(B^0 \to \rho^0\gamma\right) = \left(7.5 \pm 1.3^{+1.0}_{-0.8}\right) \times 10^{-7}$$

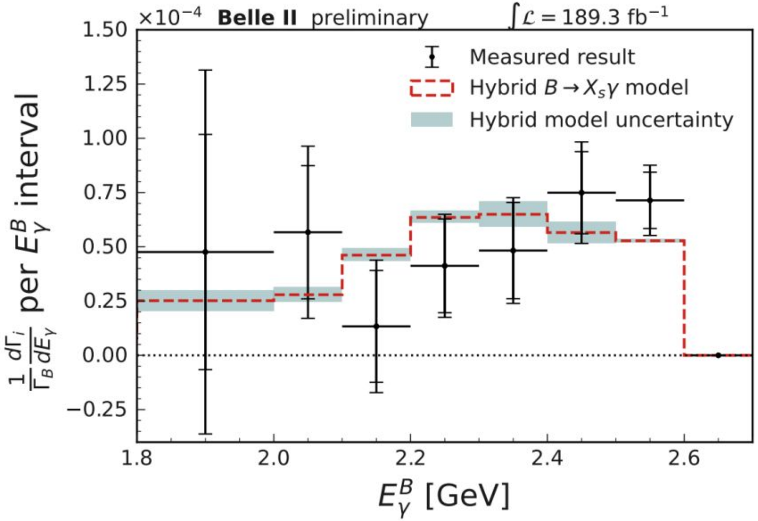
$$A_{CP}\left(B^+ \to \rho^+\gamma\right) = \left(-8.2 \pm 15.2^{+1.6}_{-1.2}\right)\%$$

$$A_{\mathrm{I}}\left(B \to \rho\gamma\right) = \left(10.9^{+11.2+6.8+3.8}_{-11.7-6.2-3.9}\right)\%,$$

Status of inclusive $b \to s\gamma$ analyses

ICHEP 2022

- Inclusive $b \to s\gamma$ measurement is theoretically clean, especially if performed with low E_γ threshold (large continuum background)
- First analysis performed sometime ago, using hadronic tagging, with promising sensitivity
- Interesting results expected with increased luminosity, a parity between statistical and systematic uncertainty is expected to be reached at about 5 ab^{-1}

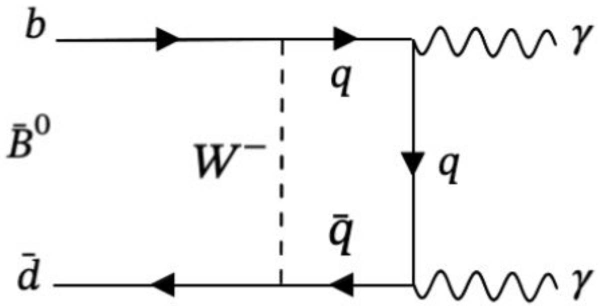


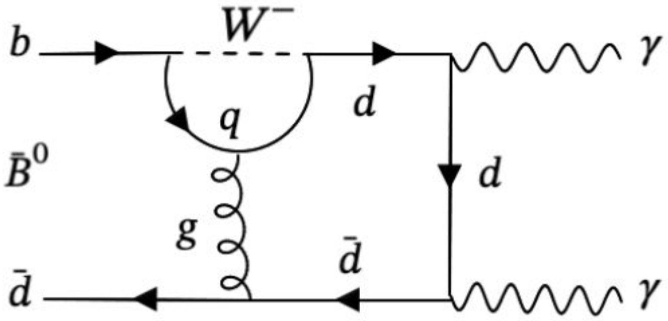
Snowmass (arXiv:2207.06307) projections:

Lower E_γ^B threshold	Statistical uncertainty				Baseline (improved)
	1 ab^{-1}	5 ab^{-1}	10 ab^{-1}	50 ab^{-1}	syst. uncertainty
1.4 GeV	10.7%	6.4%	4.7%	2.2%	10.3% (5.2%)
1.6 GeV	9.9%	6.1%	4.5%	2.1%	8.5% (4.2%)
1.8 GeV	9.3%	5.7%	4.2%	2.0%	6.5% (3.2%)
2.0 GeV	8.3%	5.1%	3.8%	1.7%	3.7% (1.8%)

Search for $B^0 \to \gamma\gamma$

Phys.Rev.D 110 (2024) 3, L031106



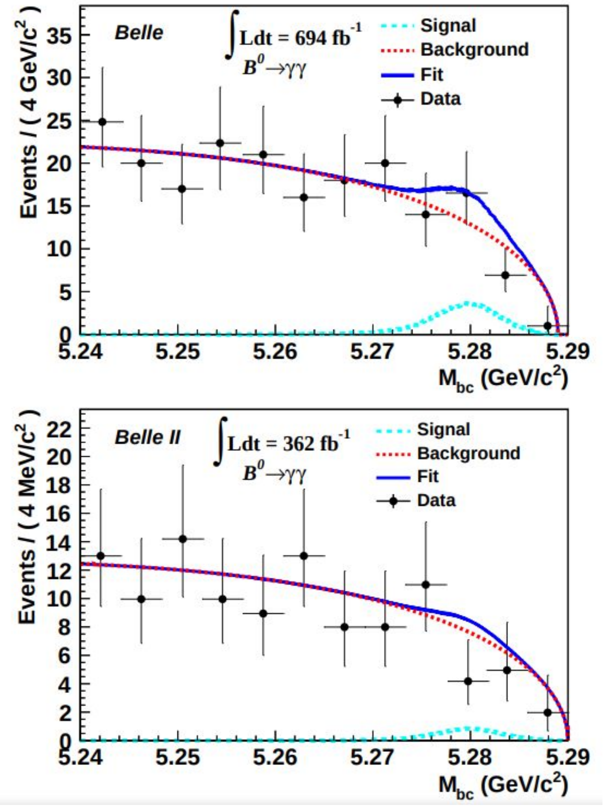


- Analysis based on Belle + Belle II data (694 fb^{-1} + 365 fb^{-1})
- In SM, suppressed vs $B_s \to \gamma\gamma$ as $|V_{td}|/|V_{ts}| = 0.04$.
- Penguin diagram contains sizable long-distance contribution: large SM uncertainties
- $\mathcal{B}_{SM} = (1.4^{+1.4}_{-0.8}) \times 10^{-8}$ (JHEP12(2020)169)

Phys.Rev.D 110 (2024) 3, L031106

Search for $B^0 \to \gamma\gamma$

	$\mathcal{B}(B^0 \to \gamma\gamma)$	UL on $\mathcal{B}(B^0 \to \gamma\gamma)$
Belle	$(5.4^{+3.3}_{-2.6} \pm 0.5) \times 10^{-8}$	$< 9.9 \times 10^{-8}$
Belle II	$(1.7^{+3.7}_{-2.4} \pm 0.3) \times 10^{-8}$	$< 7.4 \times 10^{-8}$
Combined	$(3.7^{+2.2}_{-1.8} \pm 0.5) \times 10^{-8}$	$< 6.4 \times 10^{-8}$



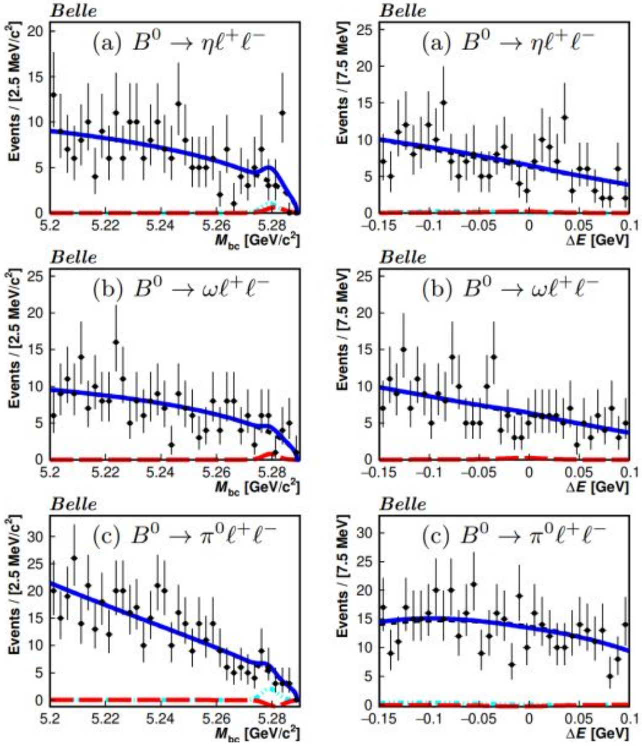
- Dominant background from continuum, suppressed by MVA classifier
- 3D unbinned fit to M_{bc}, ΔE and classifier output
- Comparable sensitivity for Belle and Belle II data
- Consistent with no signal at 2σ level, significantly better vs previous Belle and BaBar results
- Approaching SM sensitivity

Belle searches for b →d ll transitions

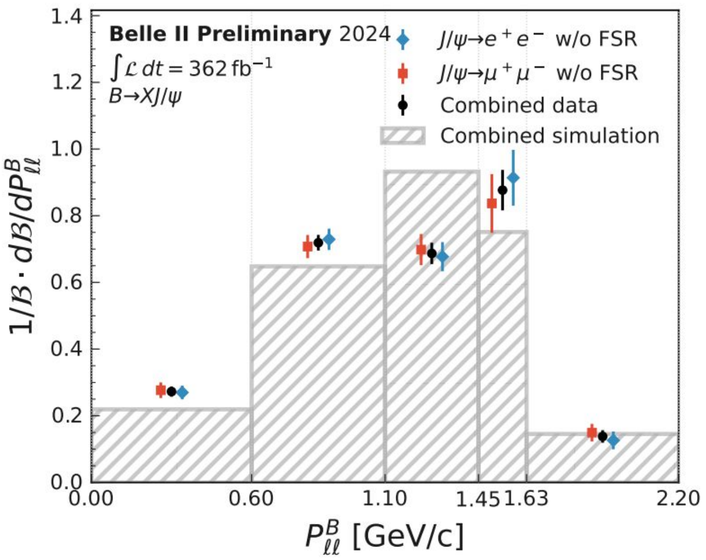
Phys.Rev.Lett. 133 (2024) 10, 101804

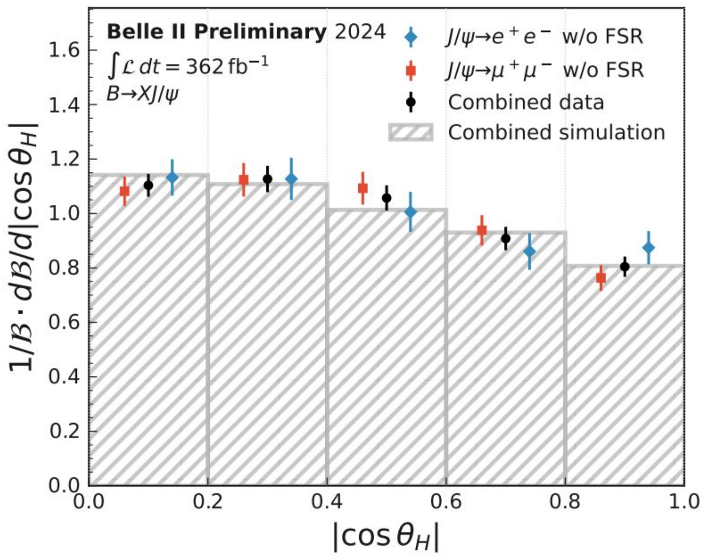
- Analysis using full Belle data sample (711 fb^{-1})
- Focusing on channels complementary to LHCb (with e, π^0 in final state)
- 2D unbinded fit to M_{bc}, ΔE
- Best or first results for a number of channels.

Channel	$\mathcal{B}^{UL}(10^{-8})$
$B^0 \to \eta e^+ e^-$	< 10.5
$B^0 \to \eta \mu^+ \mu^-$	< 9.4
$B^0 \to \eta \ell^+ \ell^-$	< 4.8
$B^0 \to \omega e^+ e^-$	< 30.7
$B^0 \to \omega \mu^+ \mu^-$	< 24.9
$B^0 \to \omega \ell^+ \ell^-$	< 22.0
$B^0 \to \pi^0 e^+ e^-$	< 7.9
$B^0 \to \pi^0 \mu^+ \mu^-$	< 5.9
$B^0 \to \pi^0 \ell^+ \ell^-$	< 3.8
$B^+ \to \pi^+ e^+ e^-$	< 5.4
$B^0 \to \rho^0 e^+ e^-$	45.5
$B^+ \to \rho^+ e^+ e^-$	< 46.7
$B^+ \to \rho^+ \mu^+ \mu^-$	< 38.1
$B^+ \to \rho^+ \ell^+ \ell^-$	< 18.9



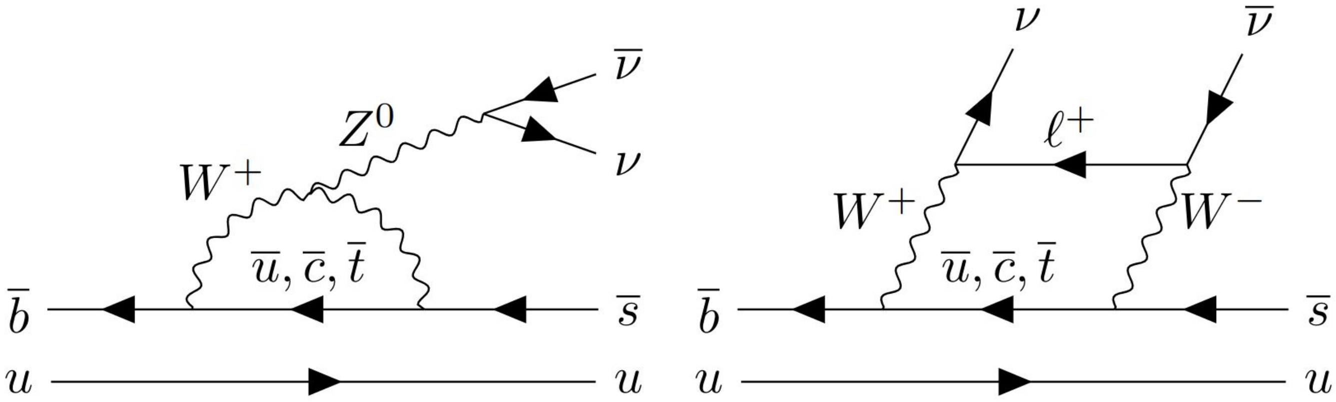
Measurement of inclusive $B \to J/\psi X$





- Belle II analysis of inclusive $B \to J/\psi X$ production employing hadronic tag for companion B
- Differential measurement of the J/ψ momentum and polarization
- Useful as a control channel for (semi) inclusive $B \to X\, ll$ and $B \to X\nu\nu$ measurements.

Search for $B^+ \to K^+\nu\nu$ motivation



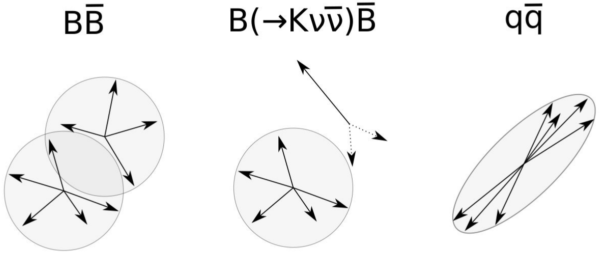
- The $B \to K^+ \nu\nu$ process is known with high accuracy in the SM:

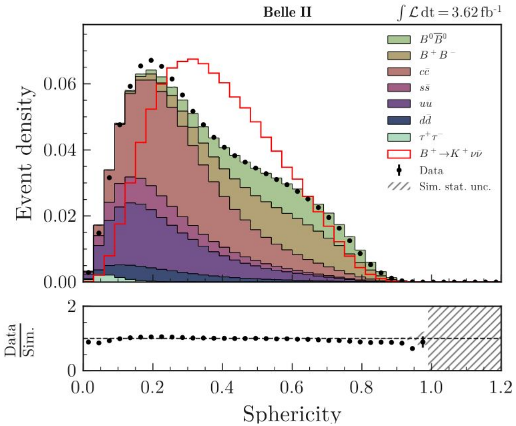
  $\mathcal{B}(B \to K^+ \nu\nu) = (5.6 \pm 0.4) \times 10^{-6}$ (arXiv:2207.13371)

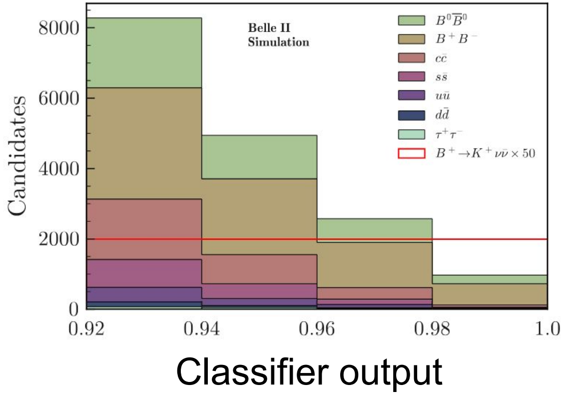
- Extensions beyond SM may lead to significant rate increase
- Very challenging experimentally, not yet observed
  - Low branching fraction, high background contributions
  - 3-body kinematics, no good kinematic variable to fit
- Unique for Belle II

Analysis strategy

Phys.Rev.D 109 (2024) 11, 112006



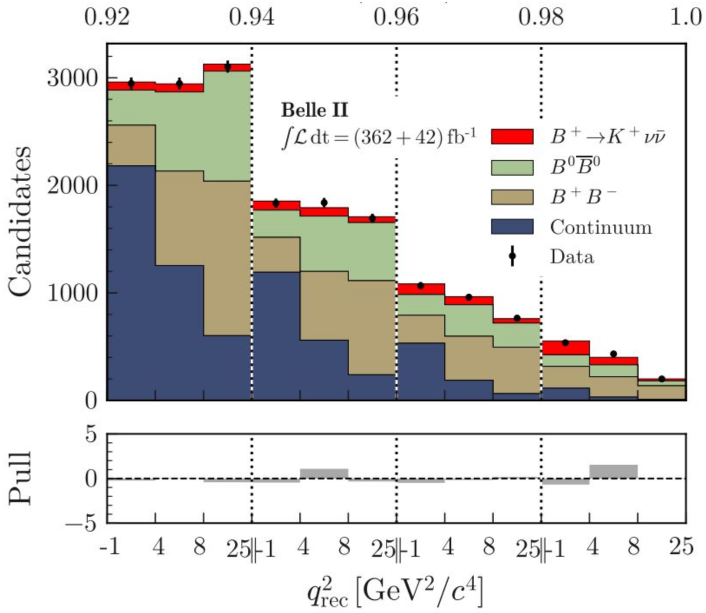


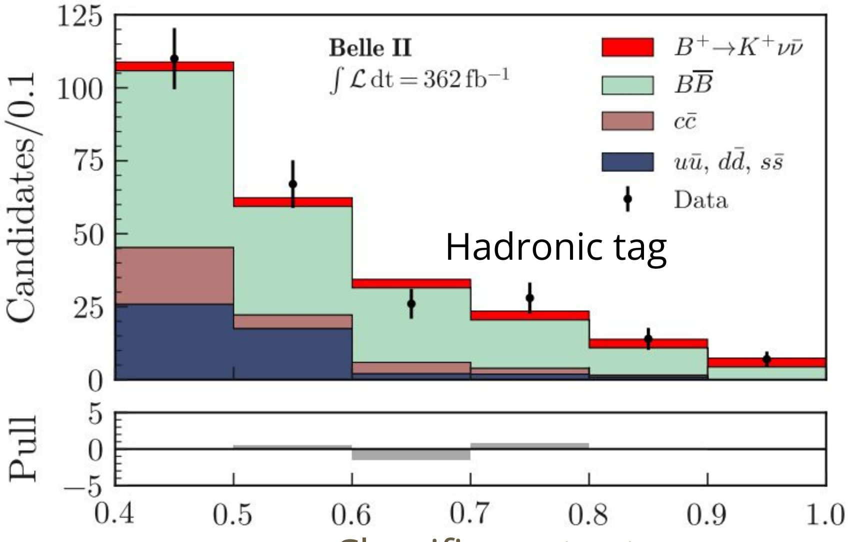


- Two analyses: more sensitive **inclusive** (total efficiency: 8%) and conventional **hadronic** tagging (total efficiency: 0.4%)
- Use event properties to suppress background with multiple variables combined
- Use classifier output as (one of) the fit variable(s), use **simulation** for signal and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

Results

Classifier output

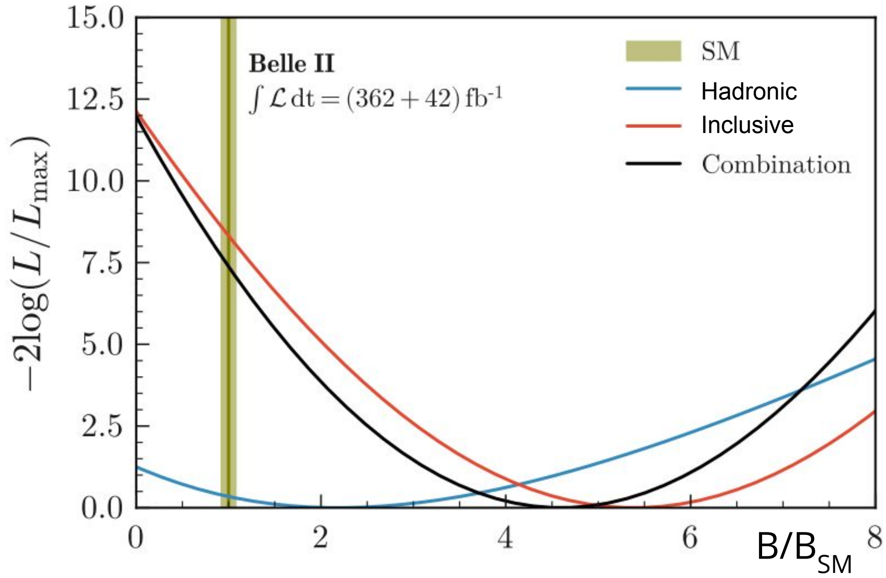


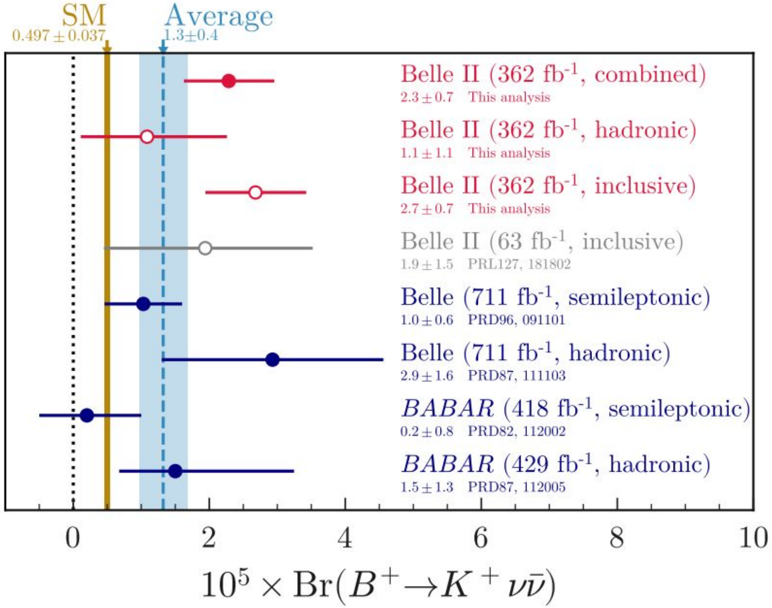


Classifier output

- Maximum likelihood fit to data using signal and background templates
- Branching fractions: $B_{\text{incl.}}$=(2.7±0.5(stat)±0.5(stat)) x 10^{-5} , $B_{\text{had.}}$=($1.1^{+0.9}_{-0.8}$(stat)$^{+0.8}_{-0.5}$(syst)) x 10^{-5}
- For inclusive analysis, evidence for *B→Kνν* 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

Combination and comparison with other measurements

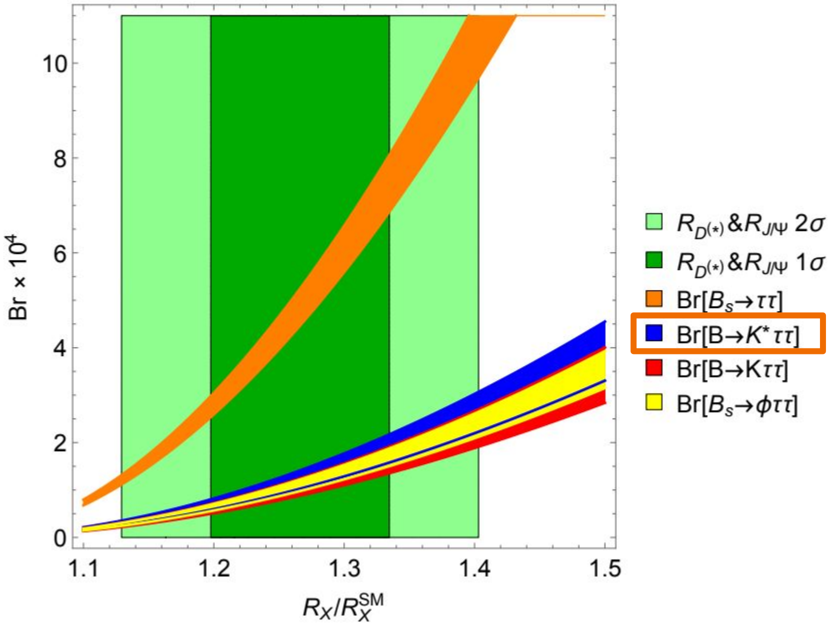




- Inclusive and hadronic measurements are combined, taking into account common correlated uncertainties. The resulting branching fraction is
  $B_{comb}(B^+ \to K^+ \nu\nu) = (2.3 \pm 0.7) \times 10^{-5} = [2.4 \pm 0.5(stat)^{+0.5}_{-0.4}(syst)] \times 10^{-5}$
  significance of observation is **3.5σ** the result is within **2.7σ** vs standard model
- 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

Search for $B^0 \to K^{*0}\tau\tau$

- Very small branching fraction in SM: 1×10^{-7}
- However, NP models that describe $b \to c\tau l$ anomalies ("R_X") may generate $\times10^4$ increase in the branching fraction
- Experimentally very challenging
  - Low efficiency
  - A lot of missing energy
  - Large backgrounds
  - Low K*0 momentum
- Last limit from Belle based on 711 fb^{-1} : $< 3.1\times10^{-3}$ @90% CL (PRD 108 011102 (2023))

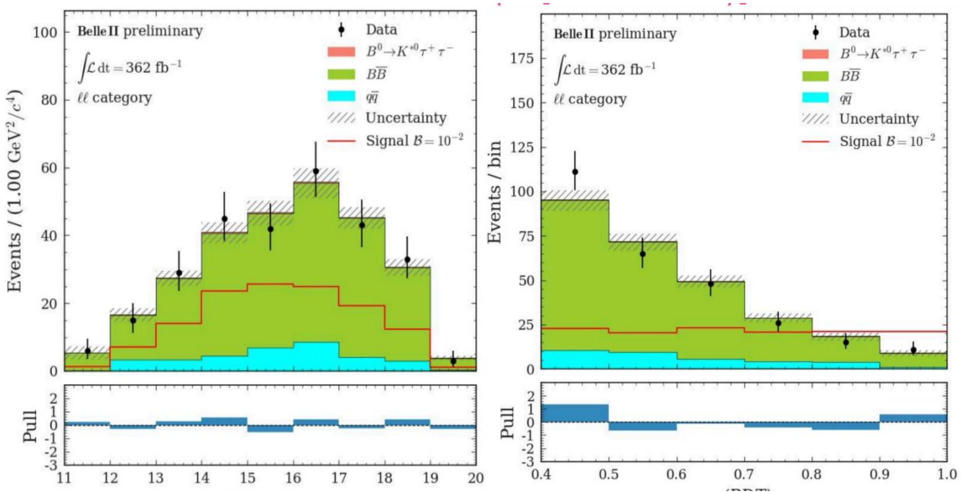


Search for $B^0 \to K^{*0}\tau\tau$

ICHEP 24

- Analyses uses hadronic tagging for companion B, based on 365 fb^{-1} Belle II data.
- Several τ decays considered: *ll*, $l\pi$, $\pi\pi$, $\rho\pi$. The best sensitivity for *ll* channel
- Binned likelihood fit to MVA classifier output (BDT), that is trained using missing energy, extra energy in the calorimeter, etc.
- Multiple validation channels
- Main backgrounds from B decays

$\to$Twice better limit vs Belle due to better tagging efficiency, more τ decay channels, MVA.

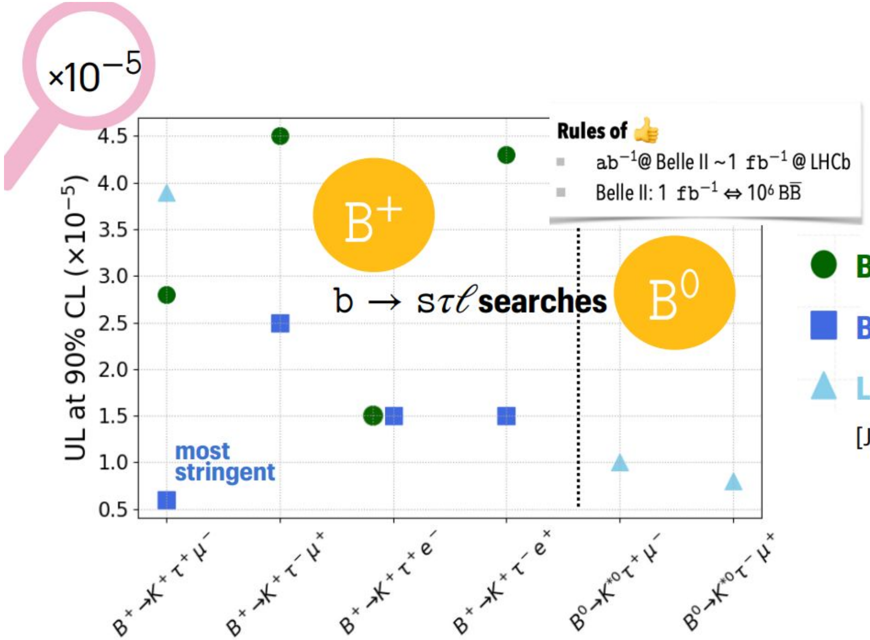


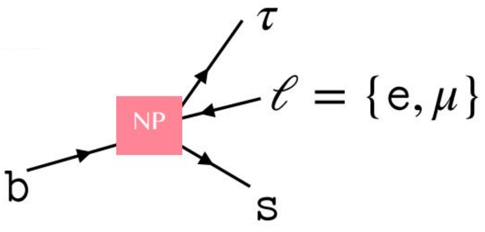
$M^2(\tau\tau)$ [GeV^2/c^4]   Classifier output

$B(B\to K^{*0}\tau\tau) < 1.8 \times 10^{-3}$ *at 90% CL*

Search for forbidden $B^0 \to K_S \tau l$

- *R(X)* anomalies and *$B(B^+ \to K^+ \nu\nu)$* excess can be explained by a new heavy particles coupled differently to 3rd generation leptons
- BSM Models can generate LFV decays with branching fractions ~ 10^{-5}
- Recent experimental limits approach their level





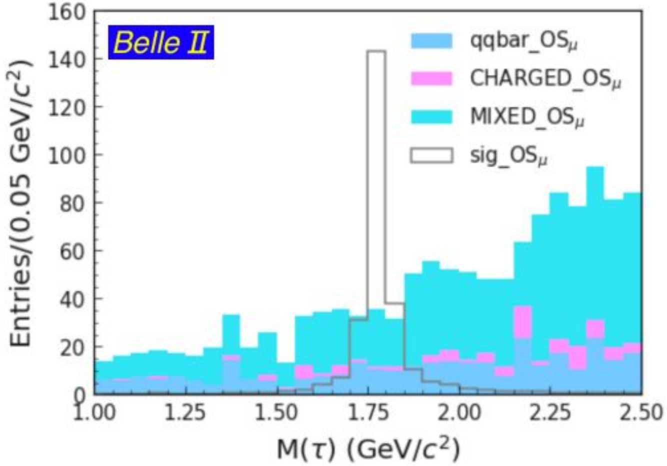
- **BaBar** (428 fb^{-1}) $B^+ \to K^+\tau^\pm\ell^\mp$ [PRD86, 012004, 2012]
- **Belle** (711 fb^{-1}) $B^+ \to K^+\tau^\pm\ell^\mp$ [PRL130, 261802, 2023]
- **LHCb** (9 fb^{-1}) $B^+ \to K^+\tau^+\mu^-$, $B^0 \to K^{*0}\tau^\pm\mu^\mp$ [JHEP06,129,2020] [JHEP06,143,2023]

No search yet in $B^0 \to K^0_S\tau^\pm\ell^\mp$

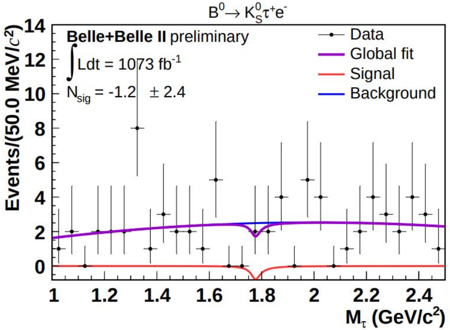
Search for forbidden $B^0 \to K_S \tau l$

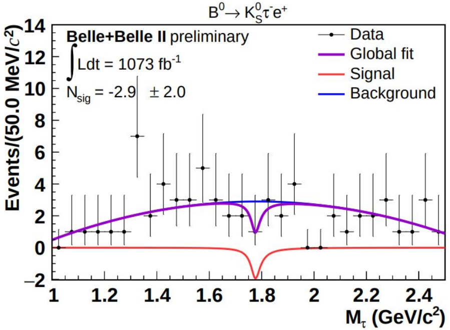
- Analysis based on Belle and Belle II data sample (711 and 365 fb^{-1})
- Hadronic tagging employed for companion B^0, its kinematics plus signal K_S and l are used to reconstruct τ mass, used as the fit variable
- Four channels considered, with e and μ leptons separated in different charges. τ one-prong decays into μ, e and π are used.
- Dedicated veto for semileptonic decays plus BDT for other backgrounds.

$$M_\tau = \left[m^2_{B_{PDG}} + m^2_{K\ell} - 2\left(E^*_B E^*_{K\ell} + \left|\vec{p}^{\,*}_B\right| \left|\vec{p}^{\,*}_{K\ell}\right| \cos\theta\right)\right]^{\frac{1}{2}}$$



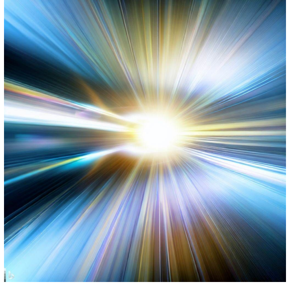
$$\mathcal{B}(B^0 \to K^0_S \tau^+ \mu^-) < 1.1 \times 10^{-5}$$
$$\mathcal{B}(B^0 \to K^0_S \tau^- \mu^+) < 3.6 \times 10^{-5}$$
$$\mathcal{B}(B^0 \to K^0_S \tau^+ e^-) < 1.5 \times 10^{-5}$$
$$\mathcal{B}(B^0 \to K^0_S \tau^- e^+) < 0.8 \times 10^{-5}$$





Comparable to the best existing limits

Summary



- SuperKEKB and Belle II provide unique opportunities for studies of $b \to s(d)$ transitions
- Many new results in the recent years, focused on the strengths of the detector: inclusive measurements, final states including photons, electrons and missing energy,
- Evidence for $B^+ \to K^+ \nu\nu$ decay with a branching fraction 2.7 standard deviations above the standard model
- Best limit for $B \to \gamma\gamma$, approaching SM sensitivity
- Precision measurements of radiative $B \to K^* \gamma$ decays
- First results on $b \to dll$ transitions from Belle and combined analysis of $B \to \rho\gamma$
- Best limits for $B^0 \to K^{*0} \tau\tau$
- First search for $B^0 \to K_S \tau l$, sensitivity similar to other similar channels with LFV

Backup

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

→

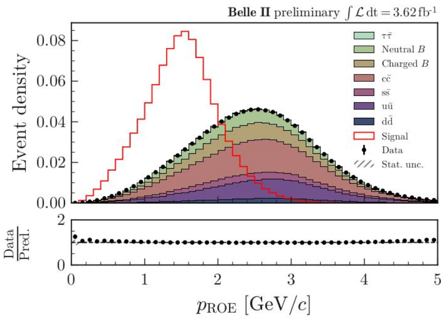
BDT$_2$ (final selection):

- *35 input variables*: using signal, event, and their correlations

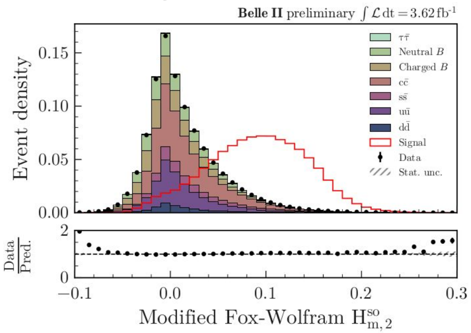
- **Three-step filter:** basic event cuts, BDT-based filter (BDT$_1$) and final selection (BDT$_2$). BDT$_2$ improves performance in terms of s/√s+b by almost factor 3

Examples of input variables for BDT_1 and BDT_2

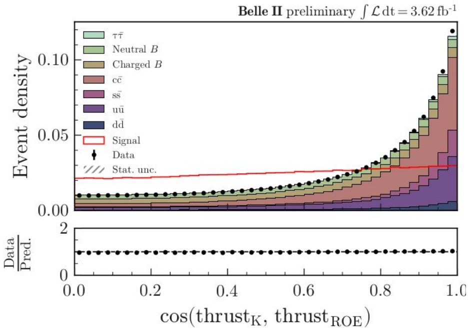
Recoil momentum



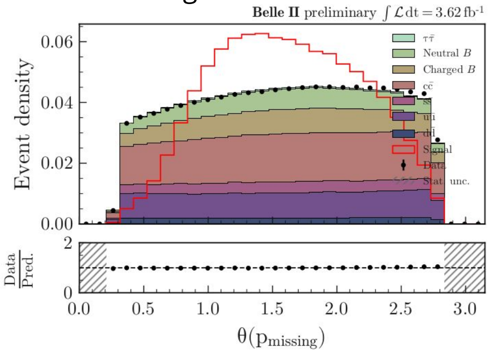
"Missing momentum"



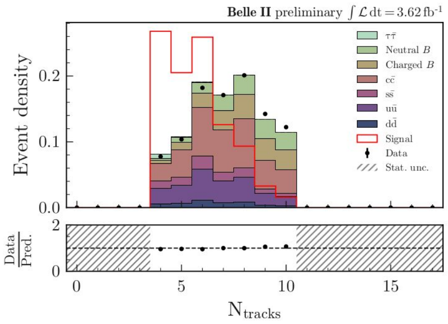
Spherical Jet-like



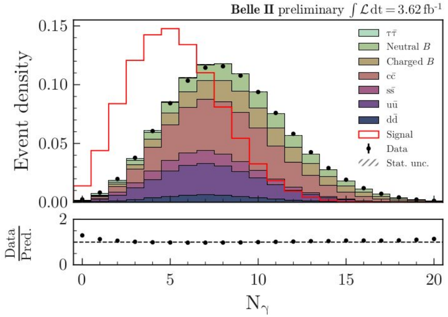
Missing momentum direction



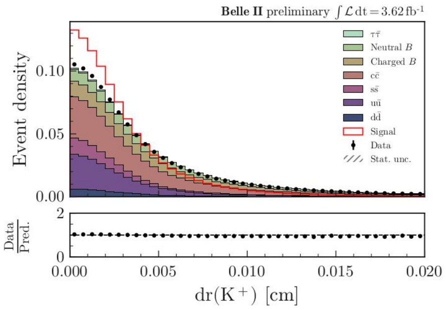
Number of tracks



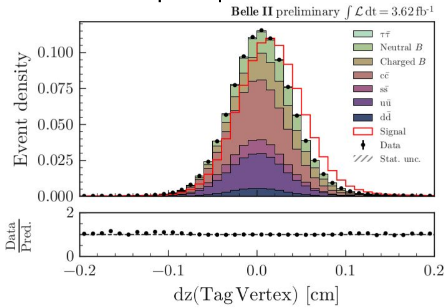
Number of neutrals



K^+ impact parameter

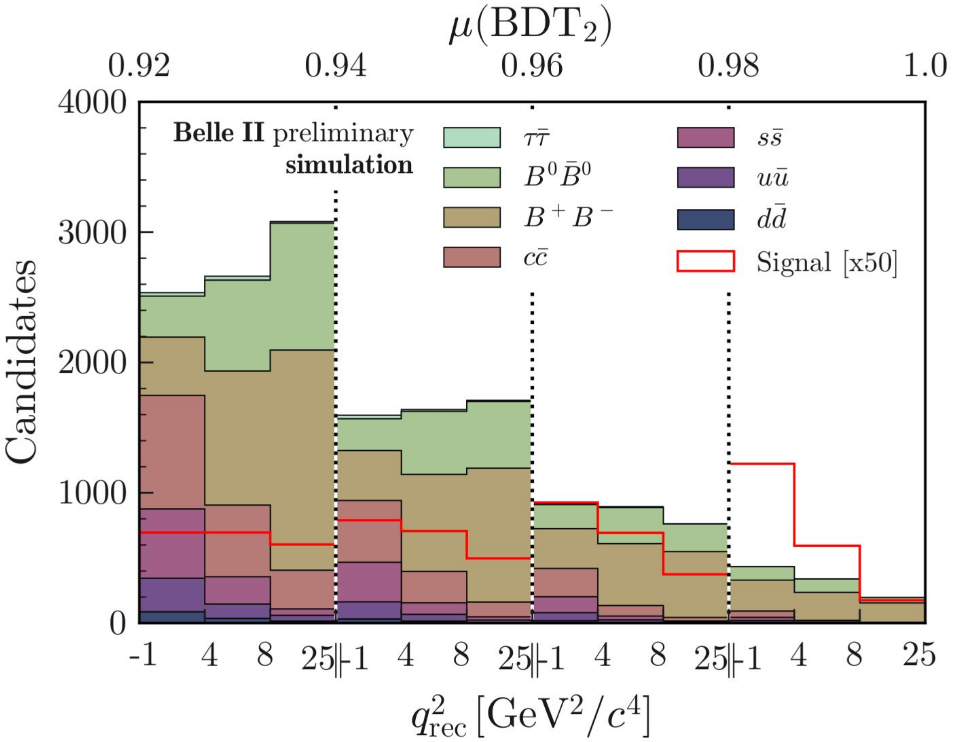


Recoil impact parameter



- 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

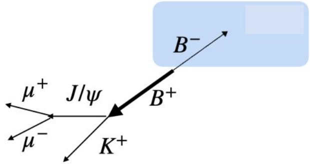


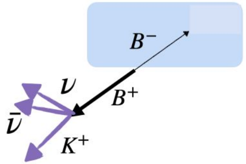
(3 bins in q^2_{rec}) x (4 bins in $\mu(\mathrm{BDT}_2)$)

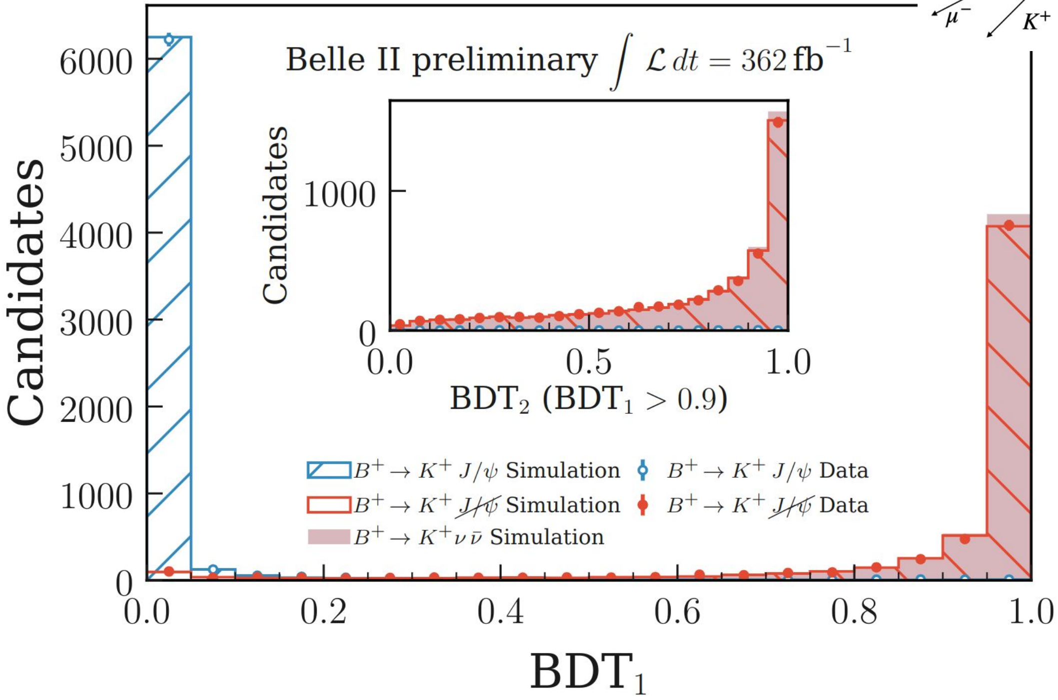
- Define the signal region at the plateau of the classifier sensitivity which corresponds to signal efficiency of **8%**
- Further subdivide it in 4 bins of classifier output $\mu(\mathrm{BDT}_2)$ and 3 bins in q^2_{rec}
- Binned profile maximum likelihood fit to data using **signal** and **7** background templates
- Systematic uncertainties varied in the fit

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

Signal efficiency validation

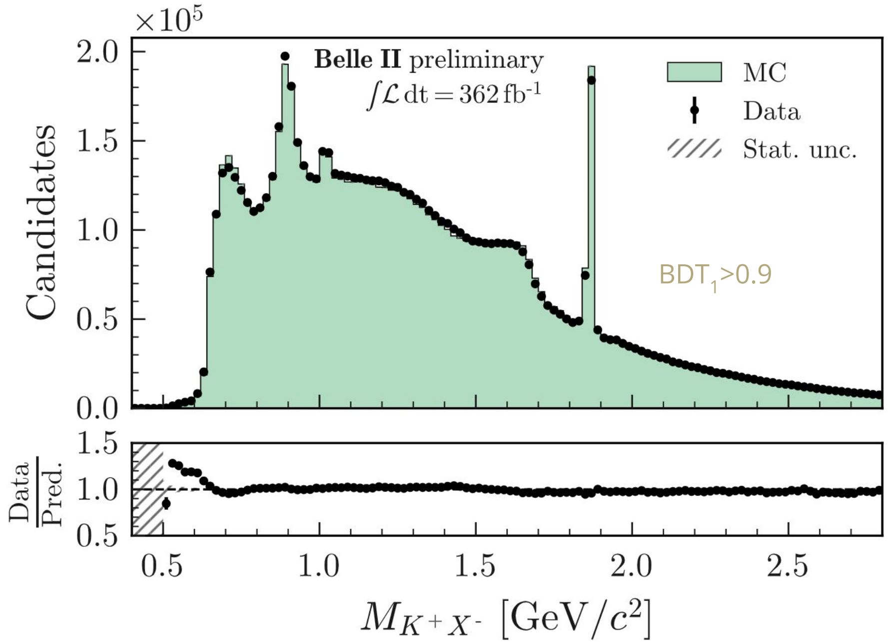


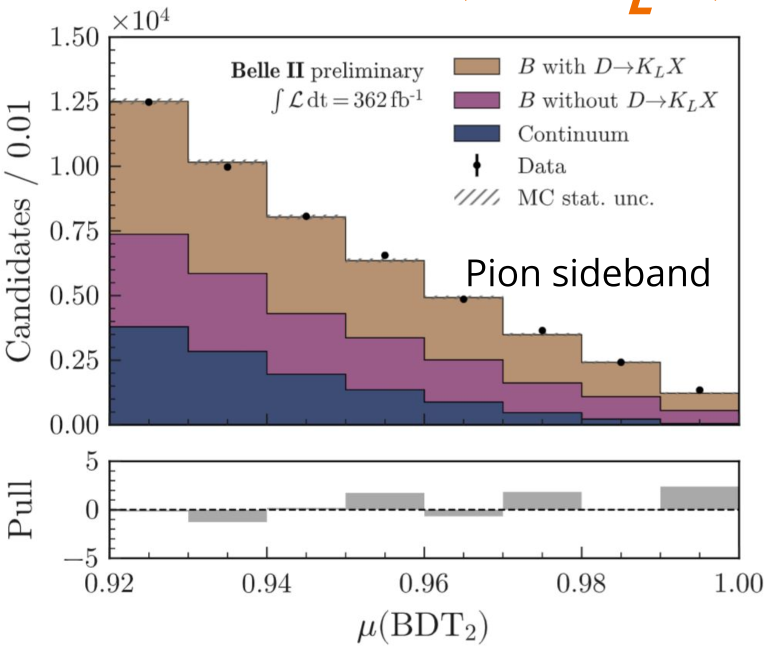




- Use cleanly reconstructed $B^+ \to K^+ J/\psi(\to\mu^+\mu^-)$ decays with $\mu^+\mu^-$ pair removed and K^+ kinematics adjusted to validate the signal efficiency in simulation. The ratio of data/simulation efficiency in the signal region is **1.00±0.03**

Background from $B \to D(\to K^{+} X) l\nu$ and $B \to K^{+} D(\to K_L X)$

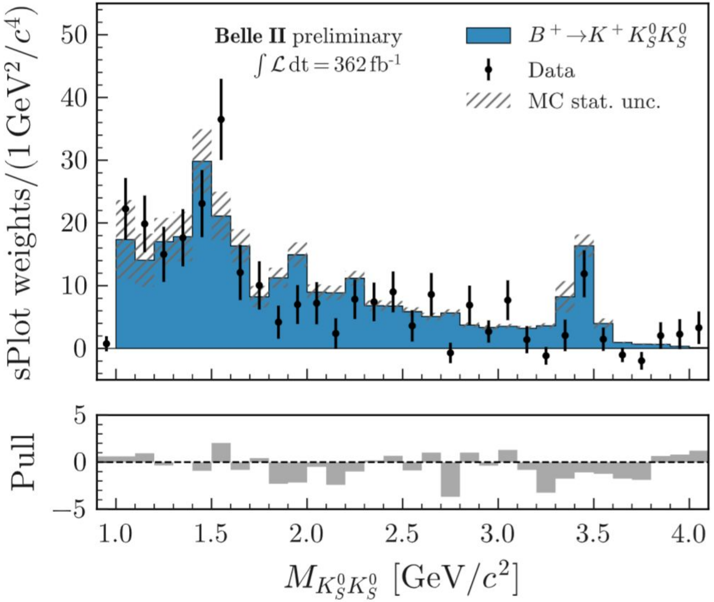




- Main backgrounds: semileptonic $B \to D(\to K^{+}X)l\nu$ decays and prompt $B \to K^{+}X$ production (>90%)
- Semileptonic decays suppressed by several MVA variables, checked at each selection step
- Prompt K^{+} production studied using prompt π^{+} from $B^{+} \to \pi^{+}X$ (and l^{+} from $B^{+} \to l^{+}X$) decays
- Systematic uncertainties on decay branching fractions, enlarged for $D(\to K_L X)$ and $B \to D^{**} l\, \nu$

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

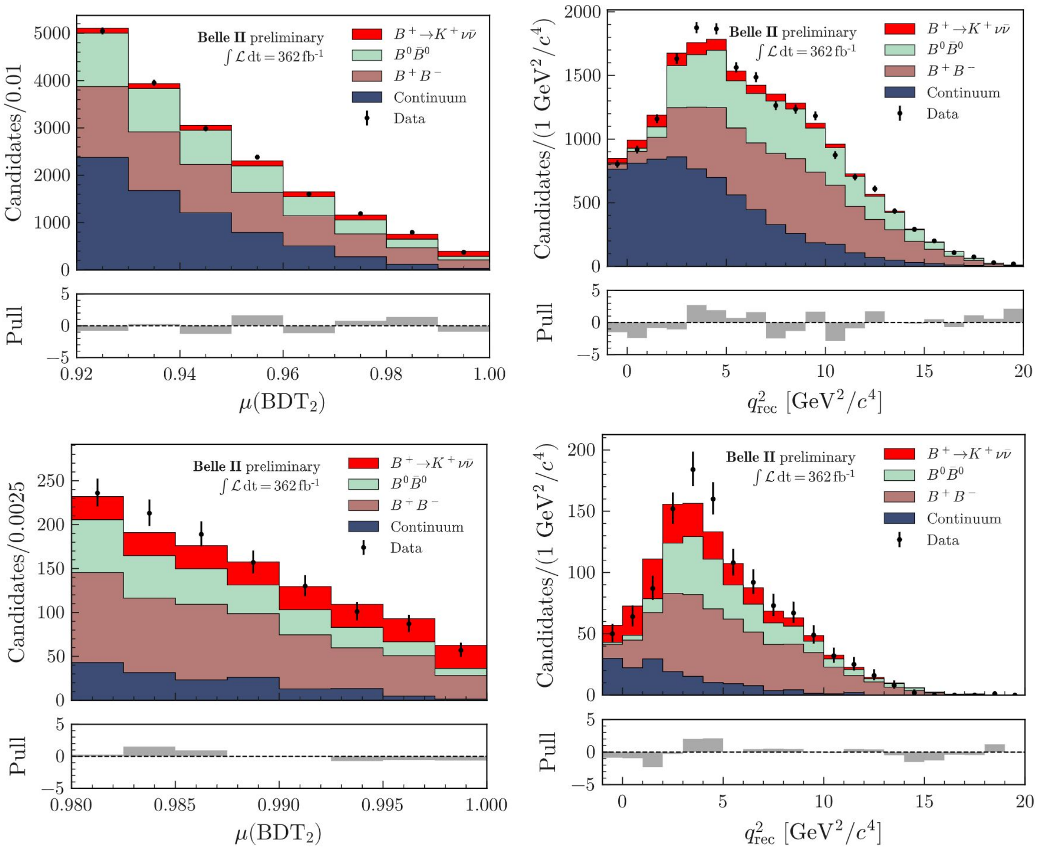
Most signal-like backgrounds



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

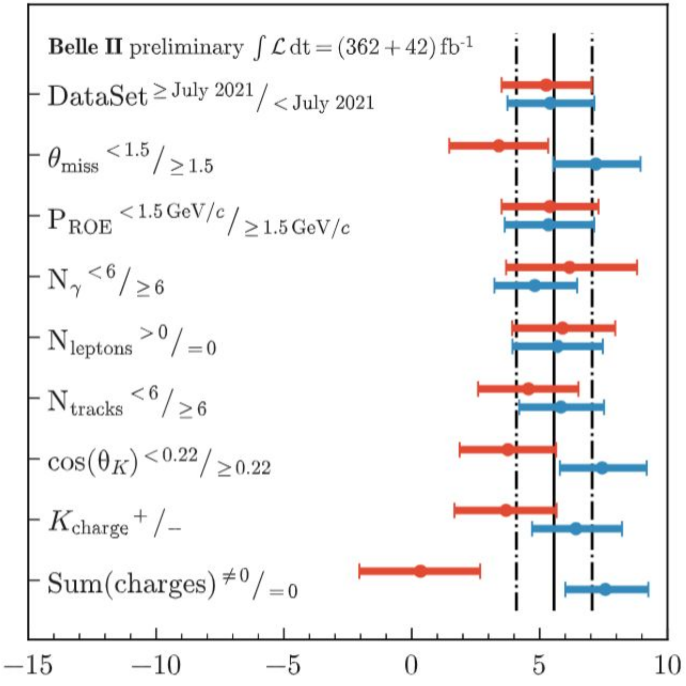
- Backgrounds from $B^+ \to K^+$ *nn* and $B^+ \to K^+K^0K^0$ have branching fractions of few x 10^{-5}, however K_L and neutrons can escape EM calorimeter
- $B^+ \to K^+K^0K^0$ modeled based on BaBar analysis (arXiv:1201.5897)
- Dedicated checks of K_L's performance in calorimeter using radiative φ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

Post-fit distributions for inclusive analysis

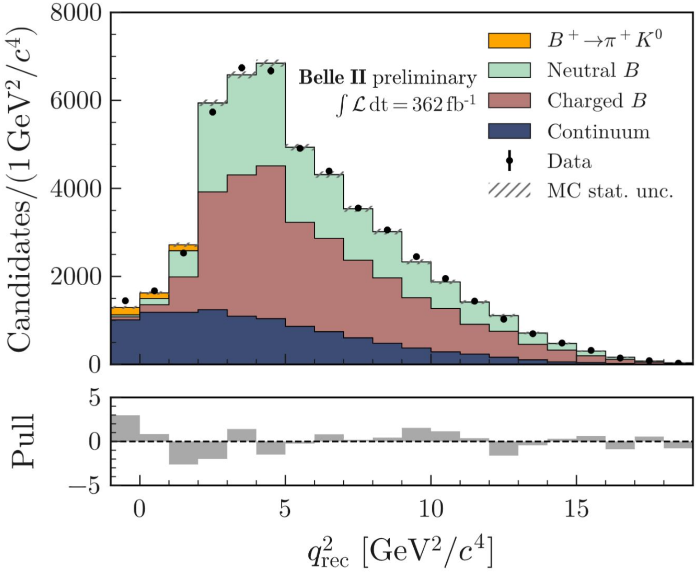


- Post-fit distributions for the inclusive analysis shown for the signal region and separately for the region with maximal sensitivity, $\mu(\mathrm{BDT}_2)>0.98$

Cross checks



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

Systematic uncertainties of the inclusive analysis

Source	Correction	Uncertainty type	Uncertainty size	Impact on σ_μ
Normalization of $B\bar{B}$ background	—	Global, 2 NP	50%	0.88
Normalization of continuum background	—	Global, 5 NP	50%	0.10
Leading B-decays branching fractions	—	Shape, 5 NP	$O(1\%)$	0.22
Branching fraction for $B^+ \to K^+ K^0_L K^0_L$	q^2 dependent $O(100\%)$	Shape, 1 NP	20%	0.49
p-wave component for $B^+ \to K^+ K^0_S K^0_L$	q^2 dependent $O(100\%)$	Shape, 1 NP	30%	0.02
Branching fraction for $B \to D^{(**)}$	—	Shape, 1 NP	50%	0.42
Branching fraction for $B^+ \to n\bar{n}K^+$	q^2 dependent $O(100\%)$	Shape, 1 NP	100%	0.20
Branching fraction for $D \to K_L X$	+30%	Shape, 1 NP	10%	0.14
Continuum background modeling, BDT$_c$	Multivariate $O(10\%)$	Shape, 1 NP	100% of correction	0.01
Integrated luminosity	—	Global, 1 NP	1%	< 0.01
Number of $B\bar{B}$	—	Global, 1 NP	1.5%	0.02
Off-resonance sample normalization	—	Global, 1 NP	5%	0.05
Track finding efficiency	—	Shape, 1 NP	0.3%	0.20
Signal kaon PID	p, θ dependent $O(10 - 100\%)$	Shape, 7 NP	$O(1\%)$	0.07
Photon energy scale	—	Shape, 1 NP	0.5%	0.08
Hadronic energy scale	−10%	Shape, 1 NP	10%	0.36
K^0_L efficiency in ECL	−17%	Shape, 1 NP	8%	0.21
Signal SM form factors	q^2 dependent $O(1\%)$	Shape, 3 NP	$O(1\%)$	0.02
Global signal efficiency	—	Global, 1 NP	3%	0.03
MC statistics	—	Shape, 156 NP	$O(1\%)$	0.52