Search for the rare $B^+ \to K^+ \nu\nu$ decay at Belle II

Jacopo Cerasoli and Lucas Martel

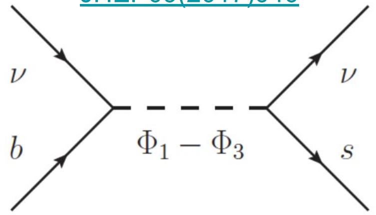
On behalf of the analysis team:
Filippo Dattola, Yulan Fan, Eldar Ganiev, Sasha Glazov, Yubo Han (DESY)
Cyrille Praz (KEK)
Slavomira Stefkova (KIT)
Claudia Cecchi, Elisa Manoni, Stefano Moneta, Roberta Volpe (INFN)
Jacopo Cerasoli, Giulio Dujany, Lucas Martel, Isabelle Ripp-Baudot (IPHC)

GDR-InF annual meeting
Strasbourg
November 6, 2023

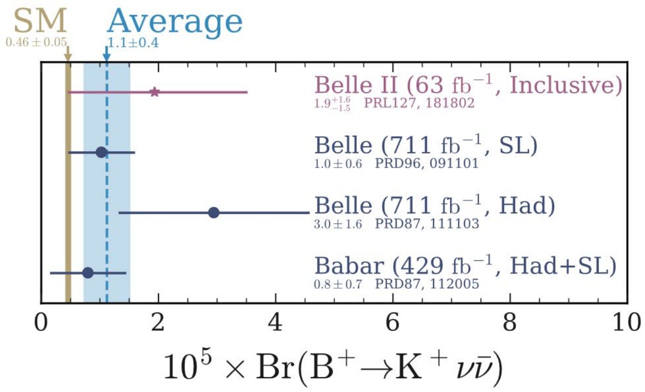
Why $B^+ \to K^+ \nu\nu$?

- $\mathcal{B}_{SM}(B^+ \to K^+\nu\nu) = (5.58 \pm 0.37) \times 10^{-6}$ Phys. Rev. D 107, 119903
- Non-SM particles (e.g. leptoquarks) could significantly modify the BR
- **Indirect way to probe multi-TeV scale**

PhysRevD.98.055003
JHEP09(2017)040



- No evidence for a signal observed (before this analysis)
- **Best upper limit from BaBar: 1.6 x 10^{-5} @ 90% CL**
- **Previous measurement by Belle II on 63 fb^{-1}: 4.1 x 10^{-5} @ 90% CL**

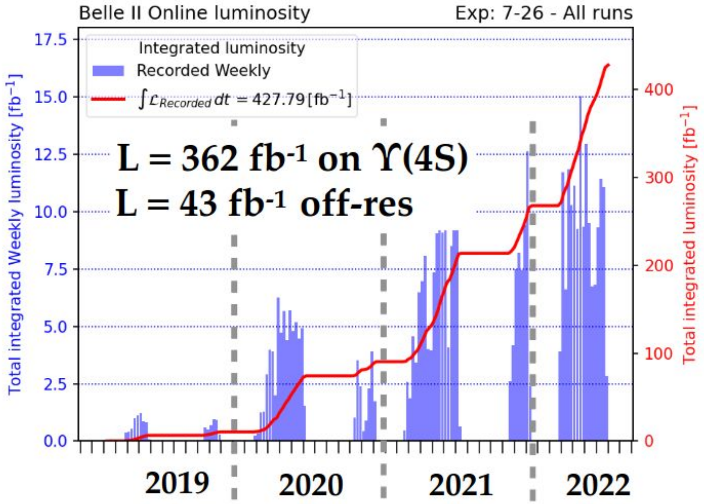


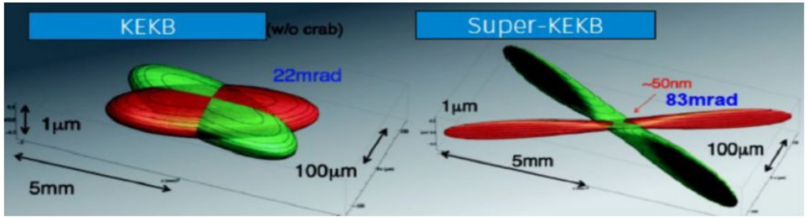
- NB: in this analysis we define **signal strength $\mu = \mathcal{B}_{measured}/\mathcal{B}_{SM}$**

The Belle II experiment at SuperKEKB

KEK Report 2010-1

- **Asymmetric e^+e^- collider at E_{CM} ~ 10.58 GeV**



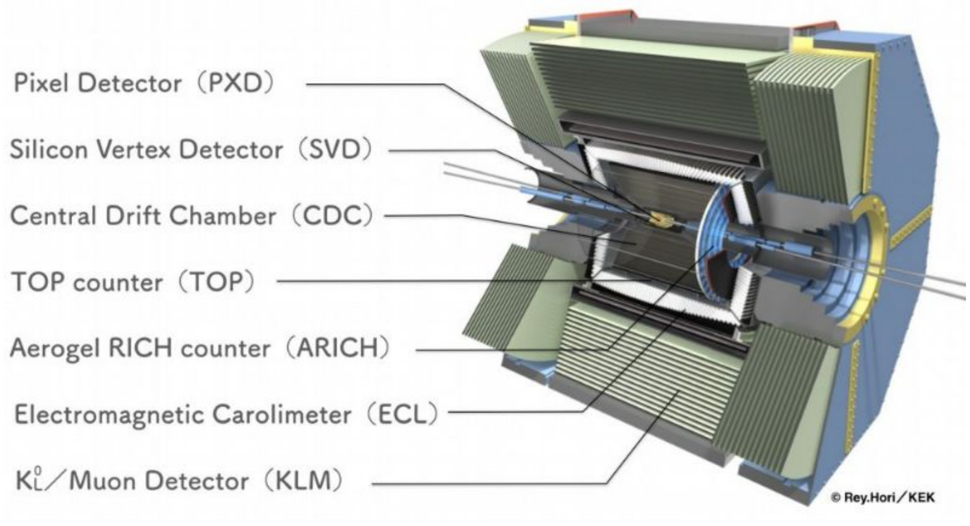


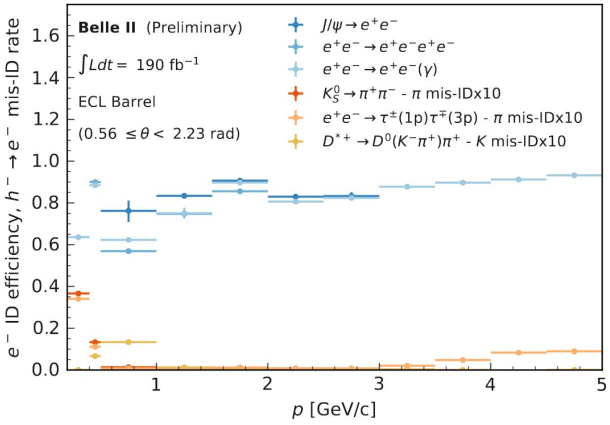
Instantaneous luminosity world record: 4.7×10^{34} cm^{-2} s^{-1}

The Belle II detector

KEK Report 2010-1

- Suited (also) for measurements with **neutrals, missing energy and inclusive decays**
- ~ 4π coverage + known initial 4-momentum → missing energy reconstruction Comput.Soft.Big.Sci 3,6(2019)
- High photon detection efficiency and good energy resolution (π^0 mass resolution ~ 5 MeV)
- Good and similar electrons and muons identification efficiency

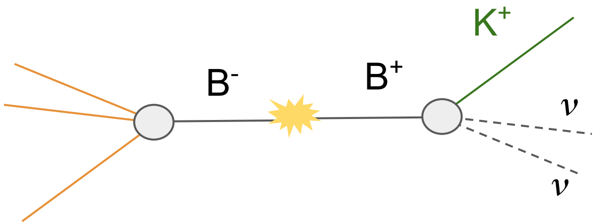




Two independent methods

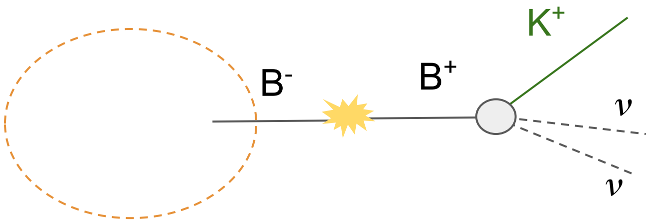
Hadronic tagging analysis (HTA)

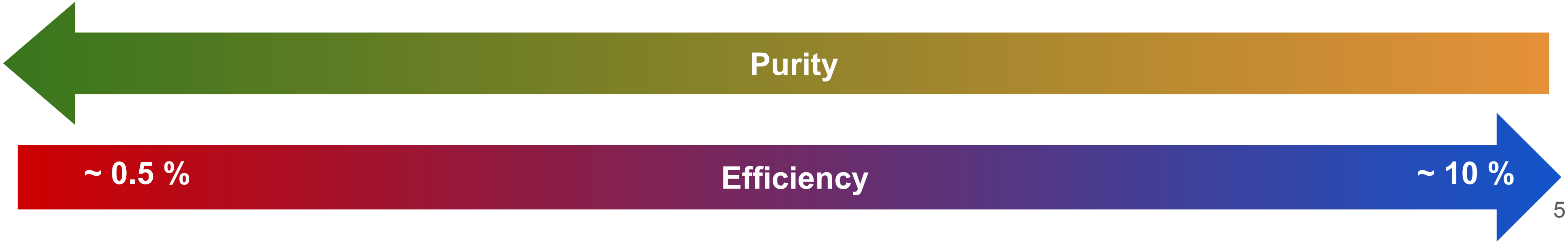
Reconstruct B_{tag} hadronically, pair to it a signal kaon candidate



Inclusive tagging analysis (ITA)

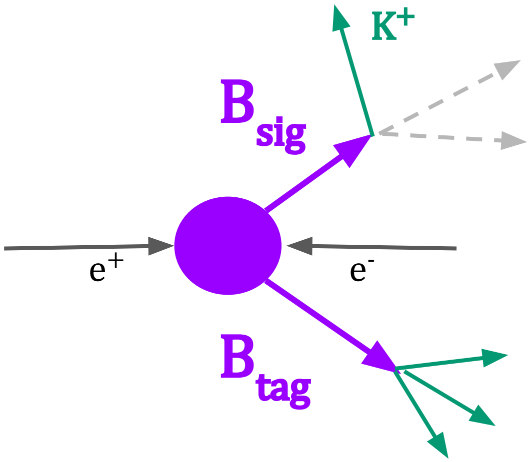
Identify signal kaon candidate and assign everything else to the B_{tag}

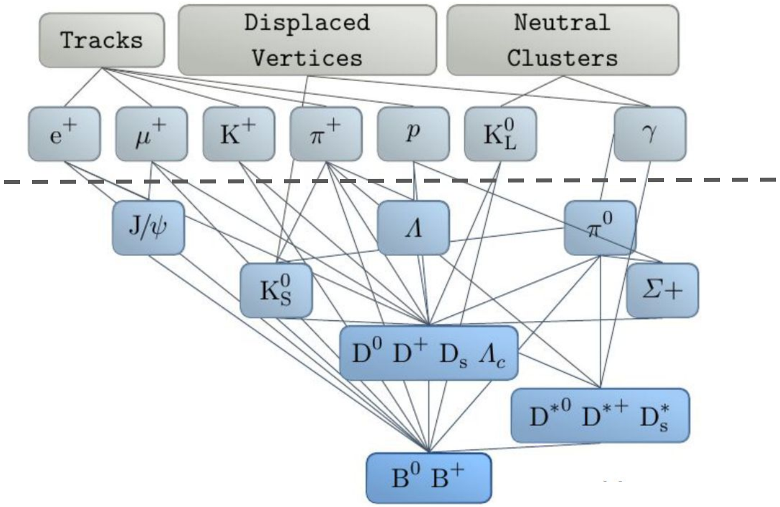




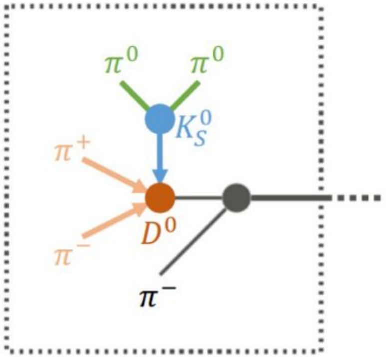
Hadronic tagging analysis (HTA)

Hadronic tagging



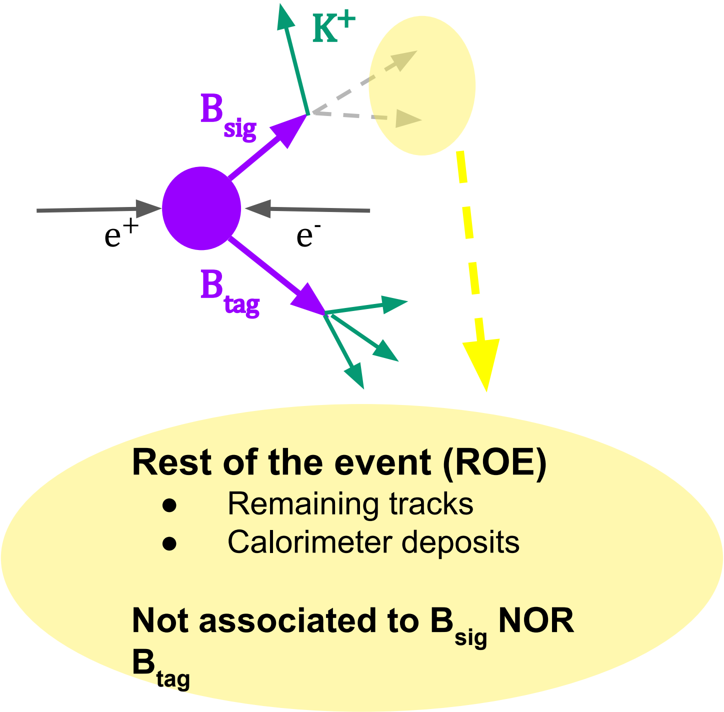


- Two B-mesons produced ➡ Signal event split in two sides: “**signal**” and “**tag**”
- Full Event Interpretation (**FEI**) algorithm
- Use final state particles to hierarchically reconstruct the **most probable** Btag
- Reconstruction done within a list of **O(10^4) fully hadronic** decay chains



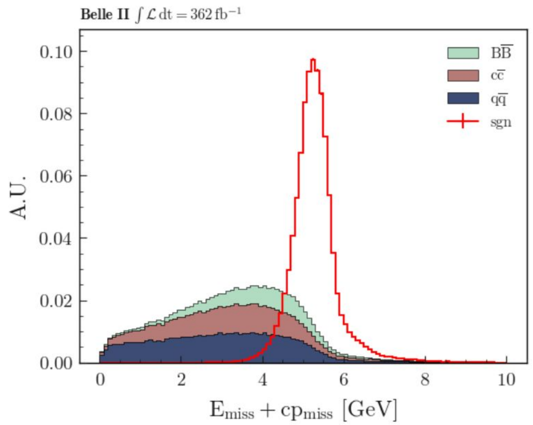
Event reconstruction

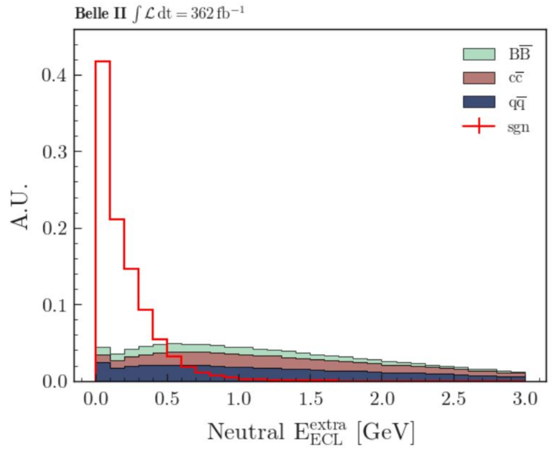
1. K^+ candidate: track with at least 1 pixel hit and requirement on kaon PID
   - ~68 % kaon efficiency with ~1.2% $K \to \pi$ mis ID
2. Identify rest-of-event (ROE):
   - Charged particles, photons, K^0_S
3. Event requirements:
   - B_{tag} and K^+ of opposite charge
   - $N_{tracks} < 12$
   - No clean tracks in ROE
   - No K^0_S, π^0 or Λ^0 in ROE



Background suppression

- Build a BDT based on XGBoost to distinguish between signal and background
- 12 features used in the training:
  - Extra calorimeter energy
  - Event topology
  - Signal K^+ kinematics
  - D meson suppression variables
  - Missing quantities (E,p) in the event

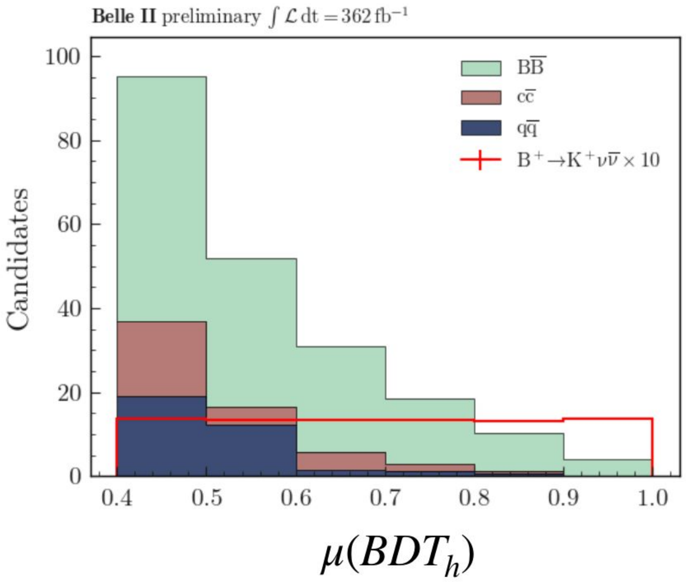




Background suppression

- Build a BDT based on XGBoost to distinguish between signal and background
- 12 features used in the training:
  - Extra calorimeter energy
  - Event topology
  - Signal K^+ kinematics
  - D meson suppression variables
  - Missing quantities (E,p) in the event
- Define signal search region from BDT output (0.4% signal eff.)

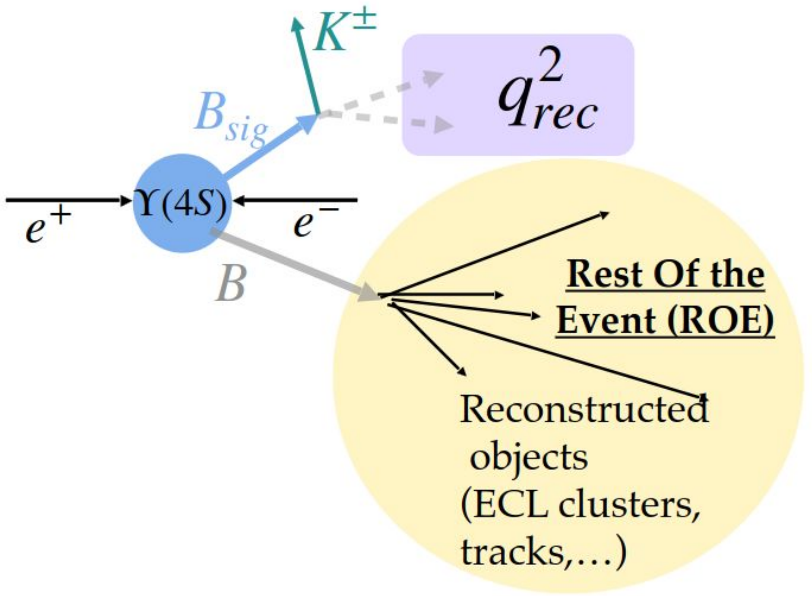
➡ Low efficiency but high sample purity



Inclusive tagging analysis (ITA)

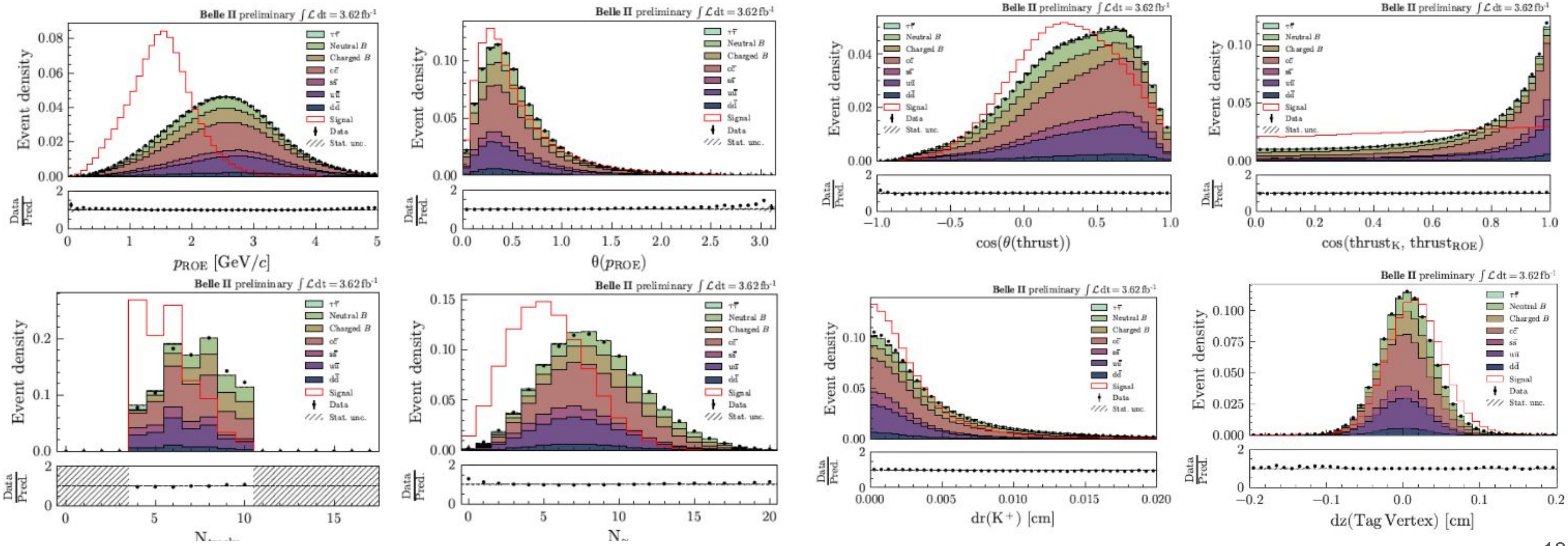
Event reconstruction

1. K^+ candidate: track with at least 1 pixel hit and requirement on kaon PID
   - ~68 % kaon efficiency with ~1.2% $K \to \pi$ mis ID
2. Identify rest-of-event (ROE):
   - Charged particles, photons, K^0_S
3. Compute q^2 of neutrino pair: $q^2_{rec} = \frac{s}{4} + M_K^2 - \sqrt{s} E_K^*$
   - Keep the candidate with lowest q^2 in the event
4. Apply event-cleaning requirements:
   - $3 < N_{tracks} < 11$
   - $17° < \Theta_{miss}^* < 160°$
   - $E_{total} > 4$ GeV



Discriminating variables

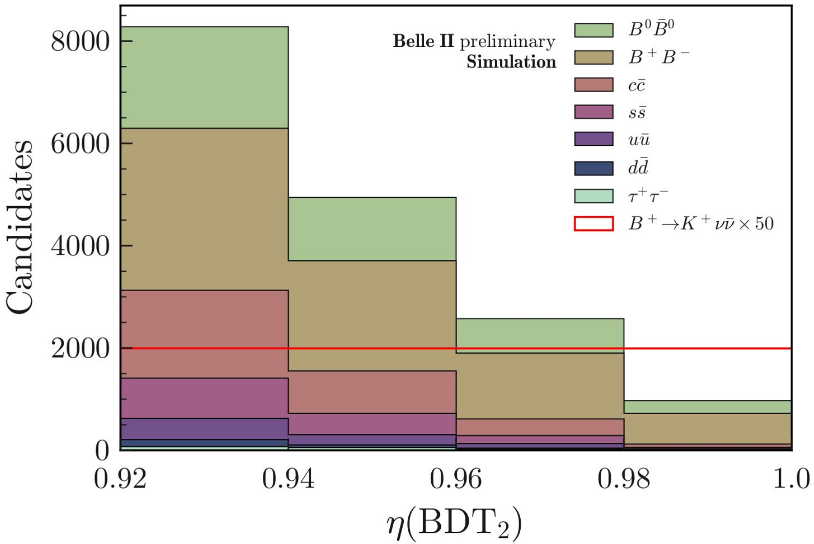
- Examples of **discriminating variables** after event reconstruction (~1% of data is left)
- Variables with good signal-background discrimination are checked for reasonable data-MC agreement



Background suppression

- Background further suppressed with **two BDTs** in sequence:
  - BDT1 uses 12 input variables, BDT2 uses 35 variables
  - BDT2 trained after cut on BDT1 and output is flattened on signal sample and called η(BDT2)

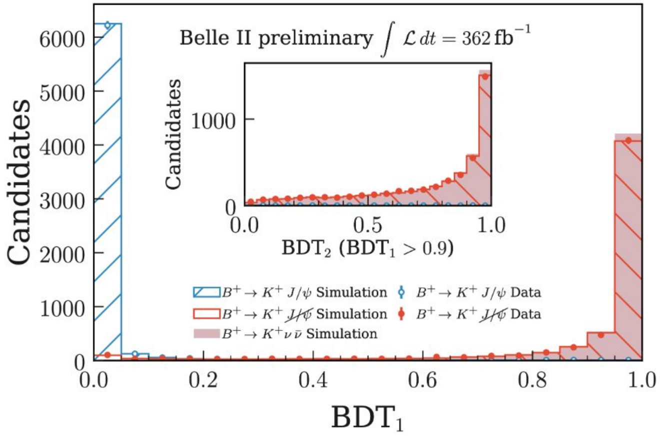
- **Signal region** defined as:
  - BDT1 > 0.9
  - η(BDT2) > 0.92



Consistency checks

Signal efficiency validation

- Analysis strategy thoroughly validated on several control samples
- Signal efficiency of BDT selection validated using **embedding procedure**:
    - Select $B^+ \rightarrow K^+ J/\Psi (\rightarrow \mu^+ \mu^-)$ candidates in data and MC
    - Remove muons and replace K^+ with the ones from signal MC (charged track, neutral clusters and PID values associated to K^+)
    - Adjust K^+ kinematics in order to match original B^+ momentum and decay vertex

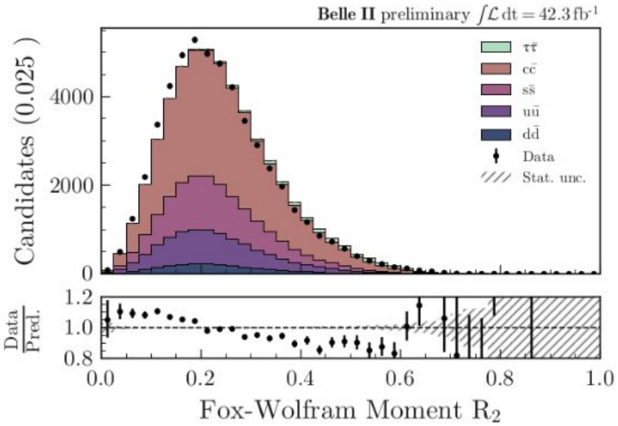


data/MC efficiency ratio under control
Uncertainty assigned as systematic

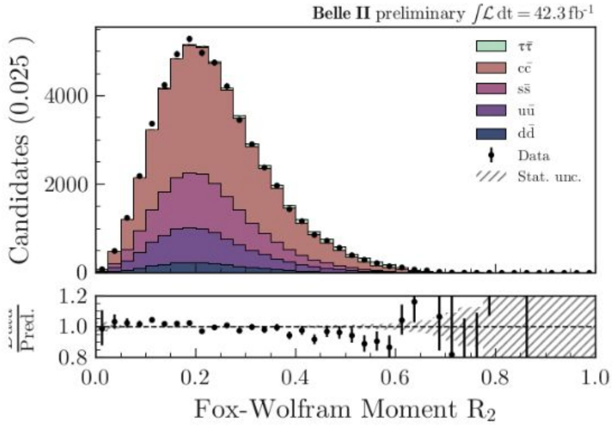
Background validation: continuum

- **Continuum background** ($e^+e^- \rightarrow$ *uu, dd, cc, ss*) represents ~30-40% of total background in signal region
- Correction derived from off-resonance data
- **Overall normalization correction factor applied**
- **Shape corrected by applying event-by-event weight**:
  - BDTc trained to separate off-resonance in simulation and data
  - Weight defined as BDTc / (1 - BDTc)

Before corrections

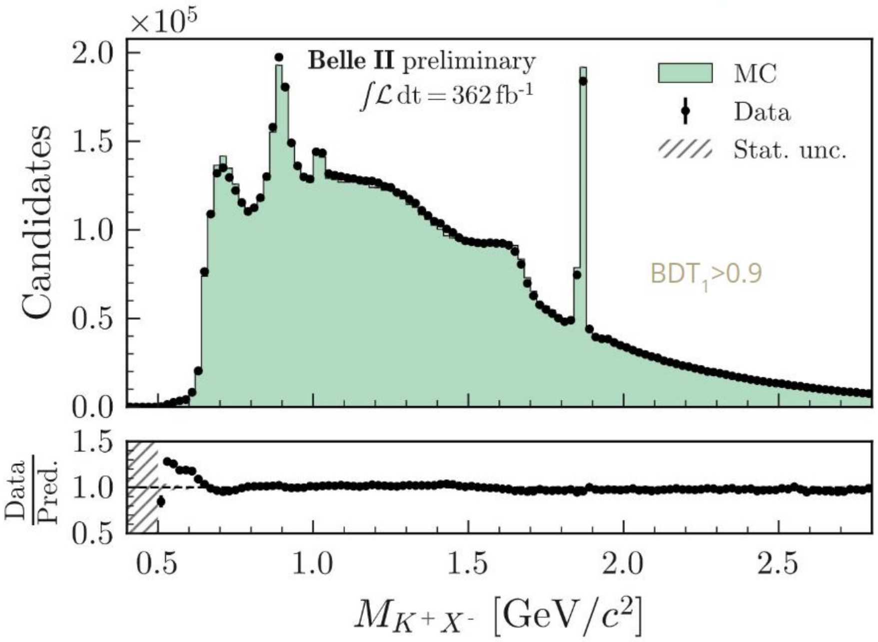


After corrections



Background validation: $B \to D\,(\to K^+ X)\,l\,\nu$

- **Semileptonic *B* decays with K^+ coming from *D* mesons** represent ~50-60% of *B* background
- Distributions checked throughout the analysis → well modeled by the simulation

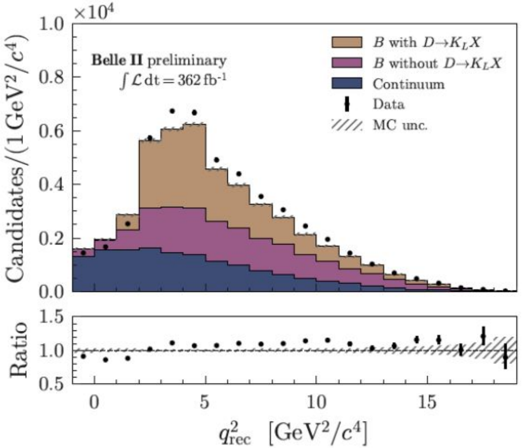


- Example: invariant mass of K^+ and a charged particle from ROE after BDT1>0.9

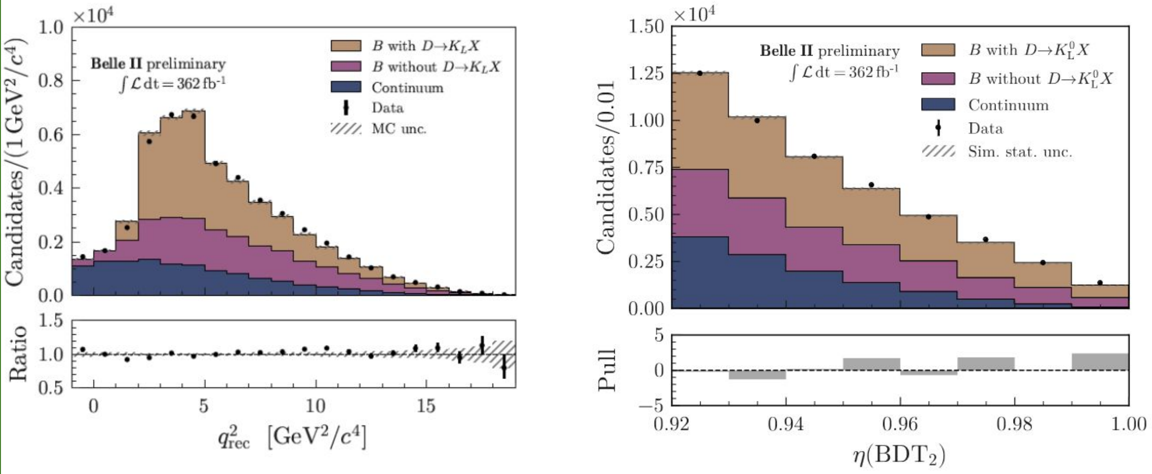
Background validation: $B \to D \to K^0_L\ X$

- **Hadronic *B* decays with D → K^0_L** represent ~20-40% of *B* background
- Sizable and poorly-known branching fractions
- Pion-enriched sample used to determine corrections: decays with B → D → K^0_LX scaled by 1.30

Before scaling

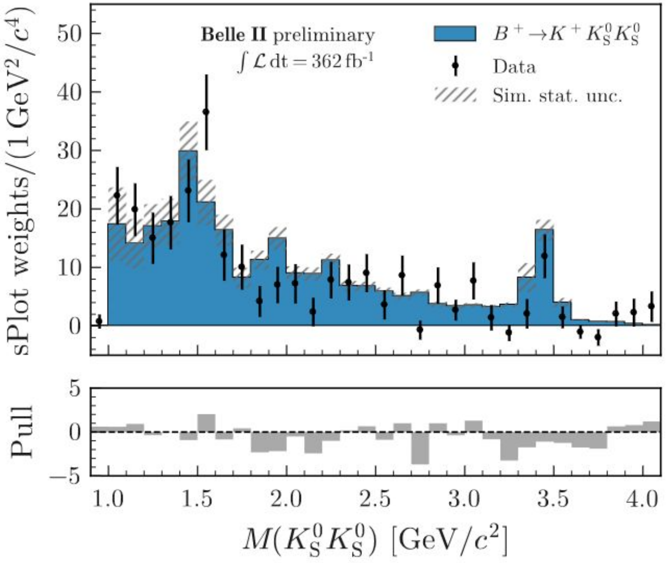


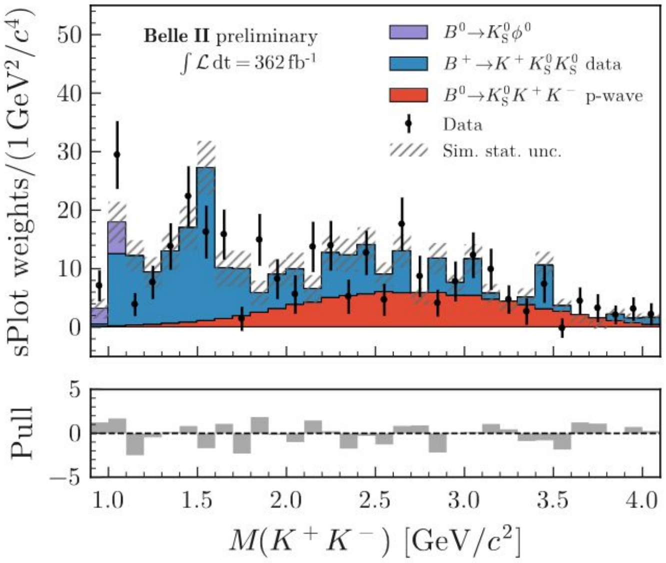
After scaling



Background validation: $B \to K^+ X^0 X^0$

- Remaining background from $B^+ \to K^+ K^0 K^0$ and $B^+ \to K^+ nn$
- $B^+ \to K^+ K^0 K^0$ validated using $B^+ \to K^+ K^0_S K^0_S$ and $B^0 \to K^0_S K^+ K^-$





- sWeighted distribution of $K^0_S K^0_S$ and $K^+ K^-$ invariant mass shows good data-MC agreement
- $B^+ \to K^+ nn$ modeled with threshold enhancement using measurements of $B^+ \to K^+ pp$

Results

Systematic uncertainties

ITA

- Main systematic uncertainty from **BB normalization**
- Also significant: **simulated** sample **size**, **branching fractions** of $B^+ \to K^+ K^0 K^0$ and $B \to D^{(**)} K^+$ decays
- Total syst. unc. on μ = +1.0 / - 0.9

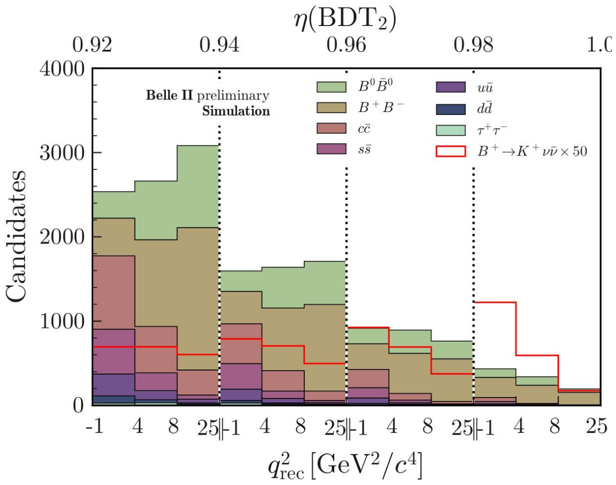
HTA

- Main systematic uncertainty from **BB normalization**
- Also significant: **photon energy** correction, **simulated** sample **size** and **continuum normalization**
- Main systematic uncertainties actually linked to **sample size**
- Total syst. unc. on μ = +1.6 / - 0.7

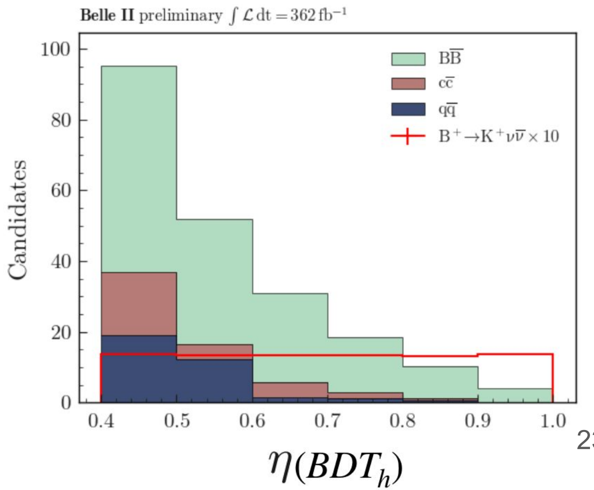
Signal extraction setup

- **Binned maximum likelihood fit**, ITA:
  - Signal region split into q^2 bins: [-1, 4, 8, 25] GeV^2
  - Off-resonance data used to constraint background
  - Total of 24 bins:
    - η(BDT2) x q^2 x on/off resonance

      4 x 3 x 2



- **Similar strategy for HTA**, simpler setup: 1-D in 6 bins of η(BDT)
- Poisson uncertainties for data counts, systematic uncertainties included as gaussian constraints



HTA measurement in data

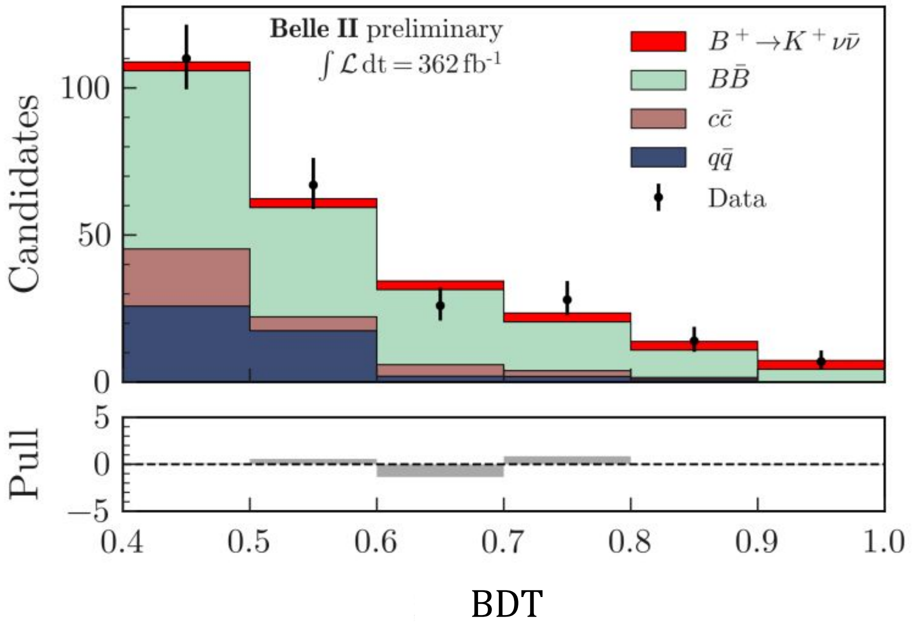
- Finally, performing the HTA fit in data we get :

$$\mu = 2.2 \pm 2.3(\text{stat})^{+1.6}_{-0.7}(\text{syst})$$

- Giving:

$$BR(B^+ \to K^+ \nu\bar{\nu}) = [1.1^{+0.9}_{-0.8}(\text{stat})^{+0.8}_{-0.5}(\text{sys})] \times 10^{-5}$$

- Significance with respect to background only hypothesis (μ = 0): **1.1σ**
- W.r.t. SM signal: **0.6σ**



This **improves** on previous **hadronic tag** results:

- **30% improvement** in uncertainty w.r.t **Belle** hadronic tag measurement with a **2x smaller** dataset
- **15% improvement** in uncertainty w.r.t **BaBar** hadronic tag measurement with a **20% smaller** dataset

However still statistically limited

ITA measurement in data

- Performing the **ITA** fit in data we get:

$$\mu = 5.6 \pm 1.1(\text{stat})^{+1.0}_{-0.9}(\text{syst})$$

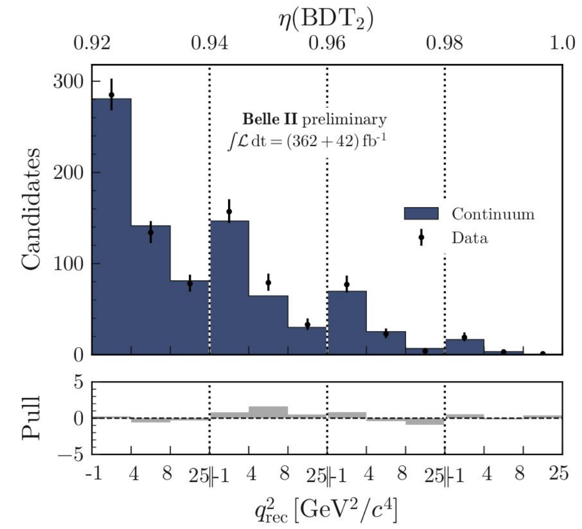
- Which gives:

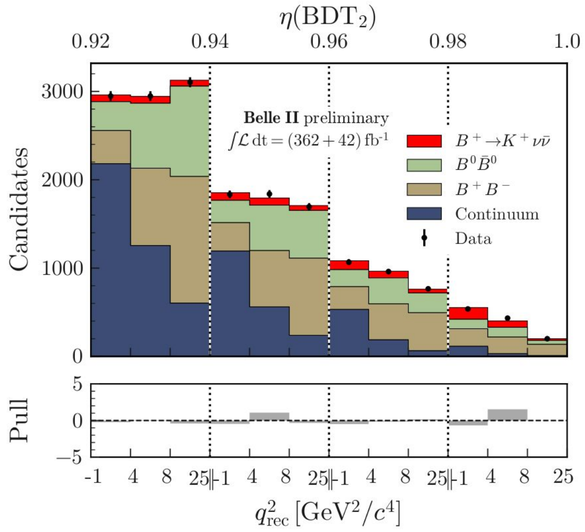
$$BR(B^+ \to K^+\nu\bar{\nu}) = [2.8 \pm 0.5(\text{stat}) \pm 0.5(\text{sys})] \times 10^{-5}$$

- Significance with respect to background only hypothesis (μ = 0): **3.6σ**
- W.r.t. SM signal: **3.0σ**

Competitive result despite small sample size:

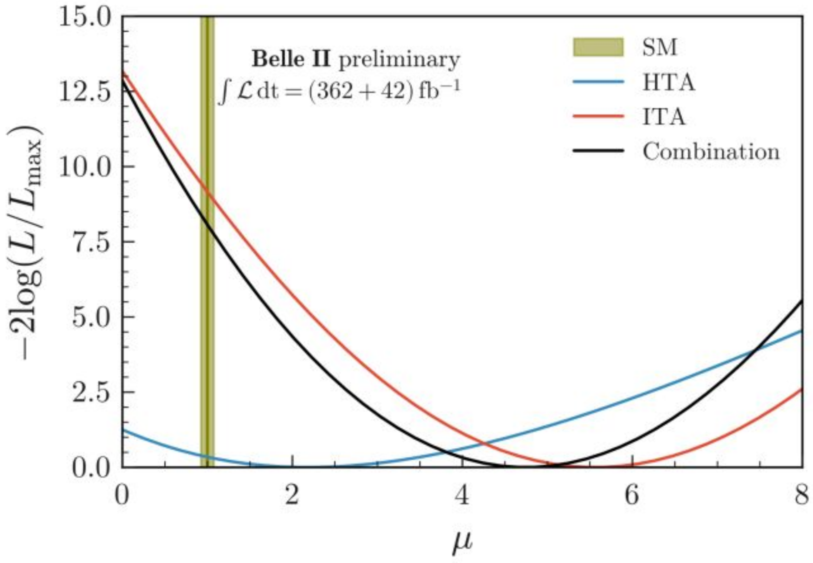
- First **evidence** of the $B^+ \to K^+ \nu\nu$ decay
- **Tension** seen with SM expectations





Combination of the measurements

- A combination of both ITA and HTA results is performed
- Correlations among common systematic uncertainties are taken into account
- Overlap between the two samples = 2% of ITA sample ⇔ 50% of HTA sample
- The combination **improves the precision** of the I-only measurement by **10%**

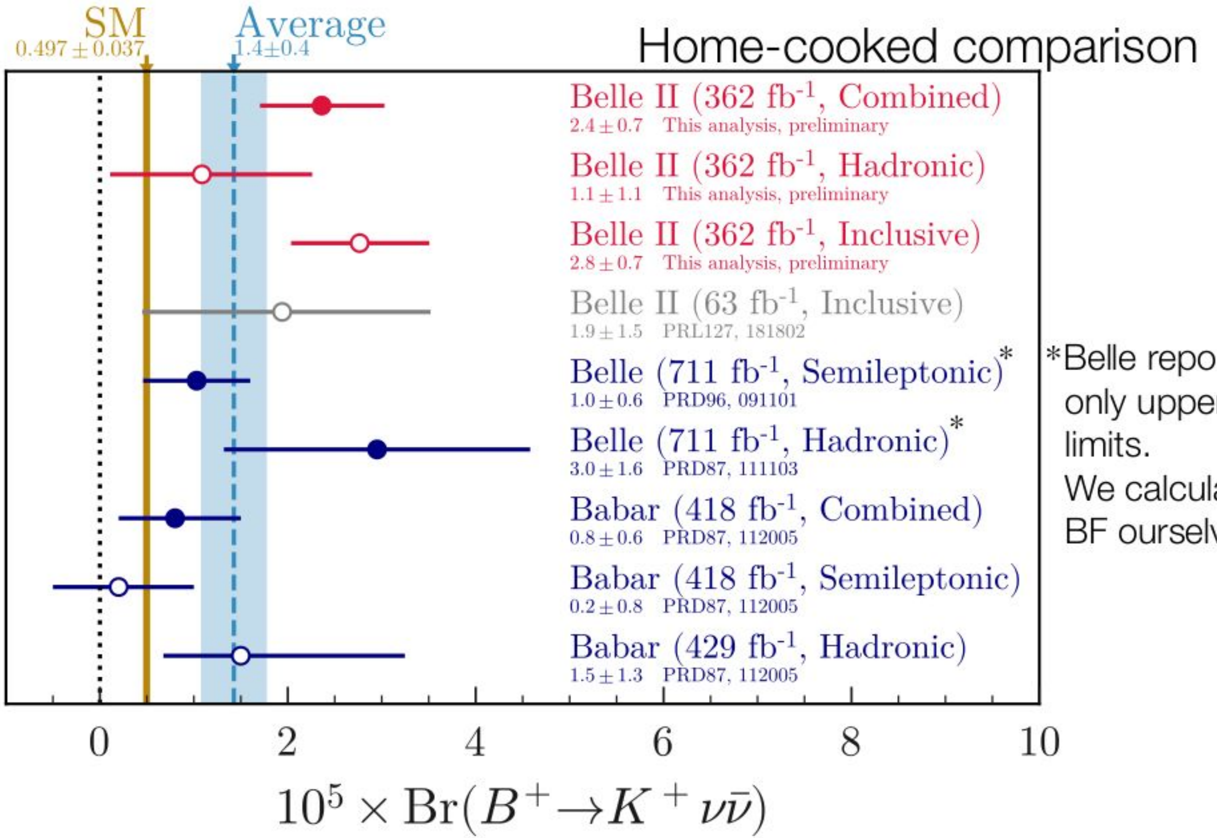


$\mu = 4.7 \pm 1.0(stat) \pm 0.9(syst)$

$BR(B^+ \to K^+ \nu\bar{\nu}) = [2.4 \pm 0.5(stat)^{+0.5}_{-0.4}(syst)] \times 10^{-5}$

3.6σ significance w.r.t background-only hyp. (2.8σ w.r.t SM) → **First evidence of the $B^+ \to K^+$ + inv. decay !** (and $b \to s$ + inv.)

New experimental state of the art



*Belle reports only upper limits. We calculate BF ourselves

ITA result has some tension with previous semi-leptonic tag measurements
a 2.4σ tension with BaBar
a 1.9σ tension with Belle

HTA result in agreement with all the previous measurements

Overall compatibility is good: $\chi^2/ndf = 4.3/4$

Summary and outlook

- Two complementary **Belle II analyses** targeting the observation of the $\mathbf{B^+ \to K^+ + \nu\nu}$ decay

- **Hadronically** tagged analysis shows competitive results w.r.t. previous similar measurements

- **Inclusively** tagged analysis allows to make the most of the early Belle II dataset → significant gain in sensitivity compared to previous measurements

- Combination of the result allows for a first evidence of the decay (**3.6σ** away from null hypothesis) and shows a **2.8σ** tension with SM expectations

- Really exciting result ! Additional work needed to get a clearer picture:

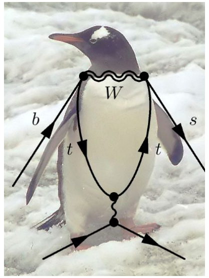
  - Complementary semileptonic tag analysis
  - other b→s + inv. Modes
  - Opportunity to bring out Belle dataset

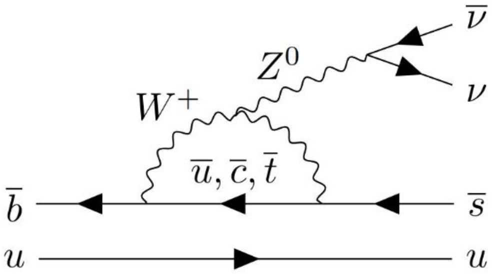
Thank you !

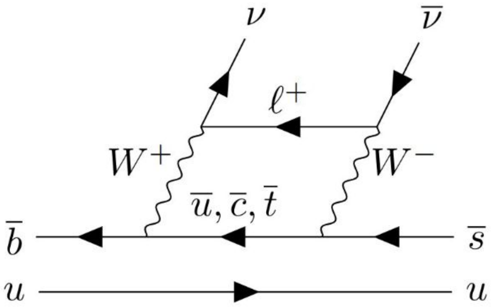
Backup

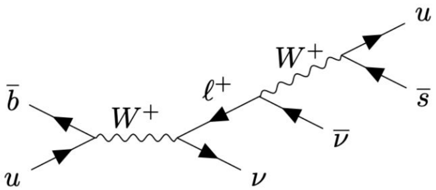
Why $B^+ \to K^+ \nu\nu$?

Phys. Rev. D 107, 119903









Long-distance double-charged contribution: ~ 0.6×10^{-6}

- FCNC suppressed by GIM mechanism
- **Precisely known in the SM**, no photon contribution and cc loops:

$$\mathcal{B}(B^+ \to K^+ \nu \overline{\nu}) = (5.58 \pm 0.37) \times 10^{-6}$$

Systematic uncertainties - HTA

Source	Uncertainty size	Impact on σ_μ
Normalization $B\bar{B}$ background	30%	0.91
Normalization continuum background	50%	0.58
Leading B-decays branching fractions	$O(1\%)$	0.10
Branching fraction for $B^+ \to K^+ K^0_L K^0_L$	20%	0.20
Branching fraction for $B \to D^{(**)}$	50%	< 0.01
Branching fraction for $B^+ \to K^+ n\bar{n}$	100%	0.05
Branching fraction for $D \to K_L X$	10%	0.03
Continuum background modeling, $\mathrm{BDT_c}$	100% of correction	0.29
Number of $B\bar{B}$	1.5%	0.07
Track finding efficiency	0.3%	0.01
Signal kaon PID	$O(1\%)$	< 0.01
Extra photon multiplicity	$O(20\%)$	0.61
K^0_L efficiency	17%	0.31
Signal SM form factors	$O(1\%)$	0.06
Signal efficiency	16%	0.42
Simulated sample size	$O(1\%)$	0.60

statistical uncertainty on $\mu = 2.3$

Systematic uncertainties

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

Total statistical uncertainty on μ: ± 1.0

From data-simulation difference in off-resonance

20% to cover possible K^0_L-K^0_S BR differences

50% since BRs are poorly known

BDT input variables (ITA)

Variables related to the kaon candidate

- Radial distance between the POCA of the K^+ candidate track and the IP (BDT_2)
- Cosine of the angle between the momentum line of the signal kaon candidate and the z axis (BDT_2)

Variables related to the tracks and energy deposits of the rest of the event (ROE)

- Two variables corresponding to the x, z components of the vector from the average interaction point to the ROE vertex (BDT_2)
- p-value of the ROE vertex fit (BDT_2)
- Variance of the transverse momentum of the ROE tracks (BDT_2)
- Polar angle of the ROE momentum (BDT_1, BDT_2)
- Magnitude of the ROE momentum (BDT_1, BDT_2)
- ROE-ROE (oo) modified Fox-Wolfram moment calculated in the c.m. (BDT_1, BDT_2)
- Difference between the ROE energy in the c.m. and the energy of one beam of c.m. ($\sqrt{s}/2$) (BDT_1, BDT_2)

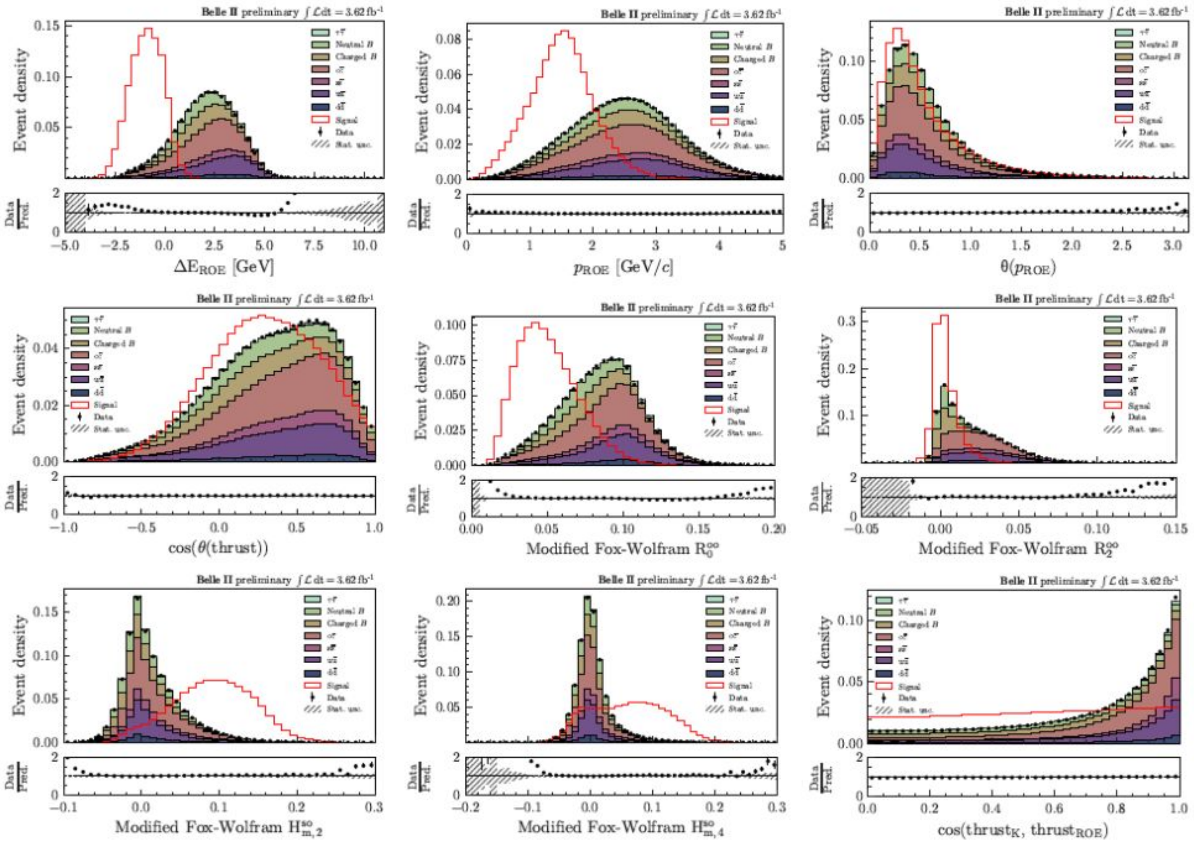
Variables related to the entire event

- Number of charged lepton candidates ($e^\pm$ or $\mu^\pm$) (BDT_2)
- Number of photon candidates, number of charged particle candidates (BDT_2)
- Square of the total charge of tracks in the event (BDT_2)
- Cosine of the polar angle of the thrust axis in the c.m. (BDT_1, BDT_2)
- Harmonic moments with respect to the thrust axis in the c.m. [41] (BDT_1, BDT_2)
- Modified Fox-Wolfram moments calculated in the c.m. [42] (BDT_1, BDT_2)
- Polar angle of the missing three-momentum in the c.m. (BDT_2)
- Square of the missing invariant mass (BDT_2)
- Event sphericity in the c.m. [40] (BDT_2)
- Normalized Fox-Wolfram moments in the c.m. [41] (BDT_1, BDT_2)
- Cosine of the angle between the momentum line of the signal kaon track and the ROE thrust axis in the c.m. (BDT_1, BDT_2)
- Radial and longitudinal distance between the POCA of the K^+ candidate track and the tag vertex (BDT_2)

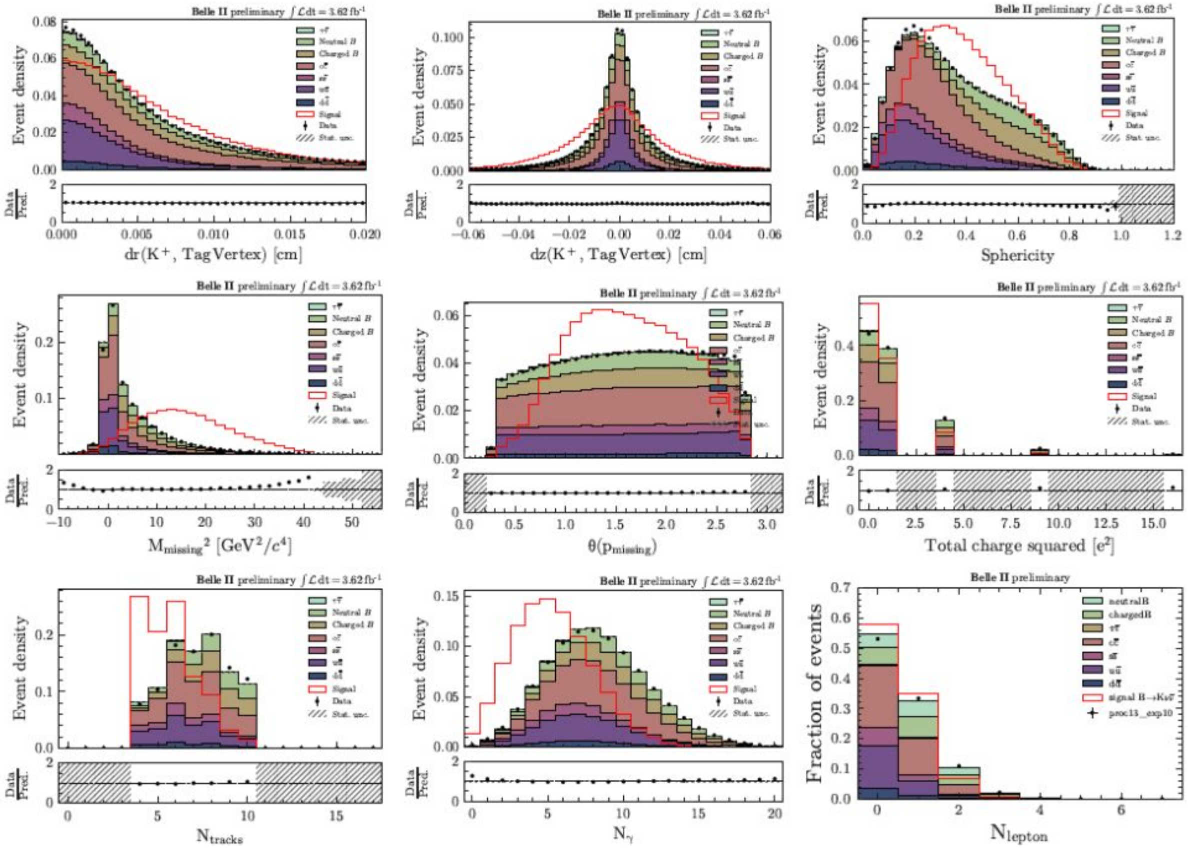
Variables related to the D^0/D^+ suppression

- Radial distance between the best D^+ candidate vertex and the IP (BDT_2)
- χ^2 of the best D^0 candidate vertex fit and the best D^+ candidate vertex fit (BDT_2)
- Mass of the best D^0 candidate (BDT_2)
- Median p-value of the vertex fits of the D^0 candidates (BDT_2)

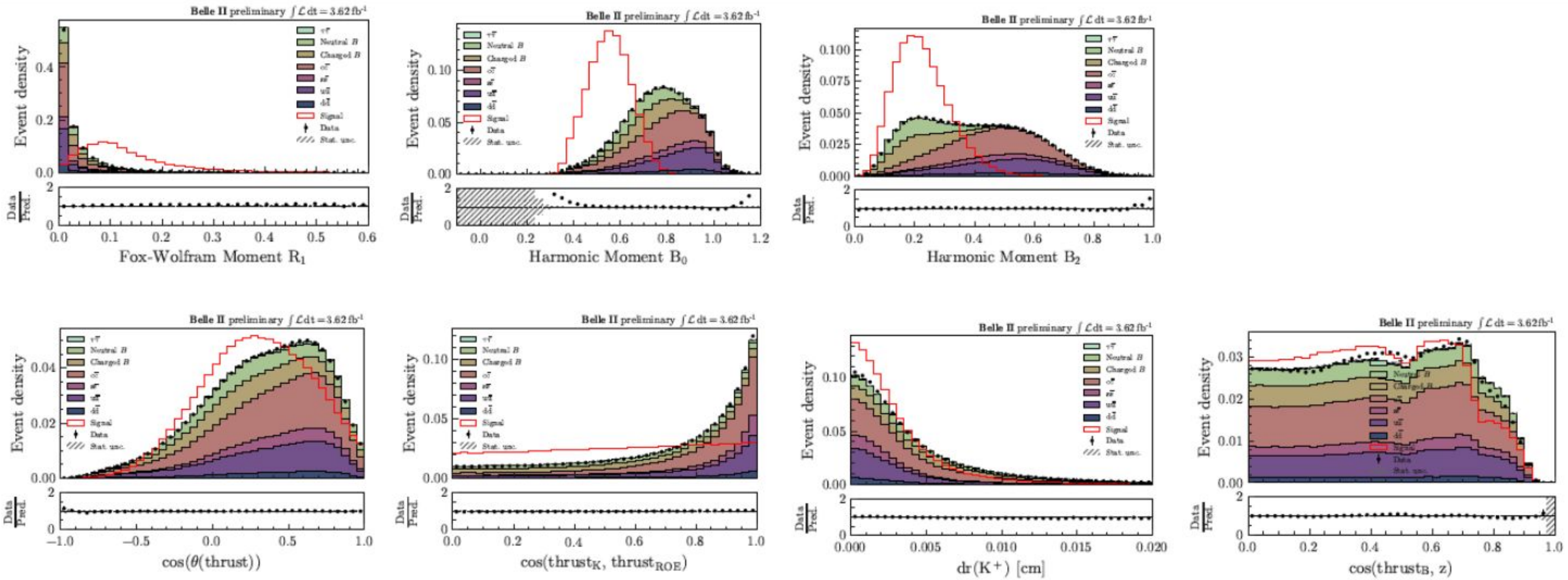
BDT input variables (ITA)



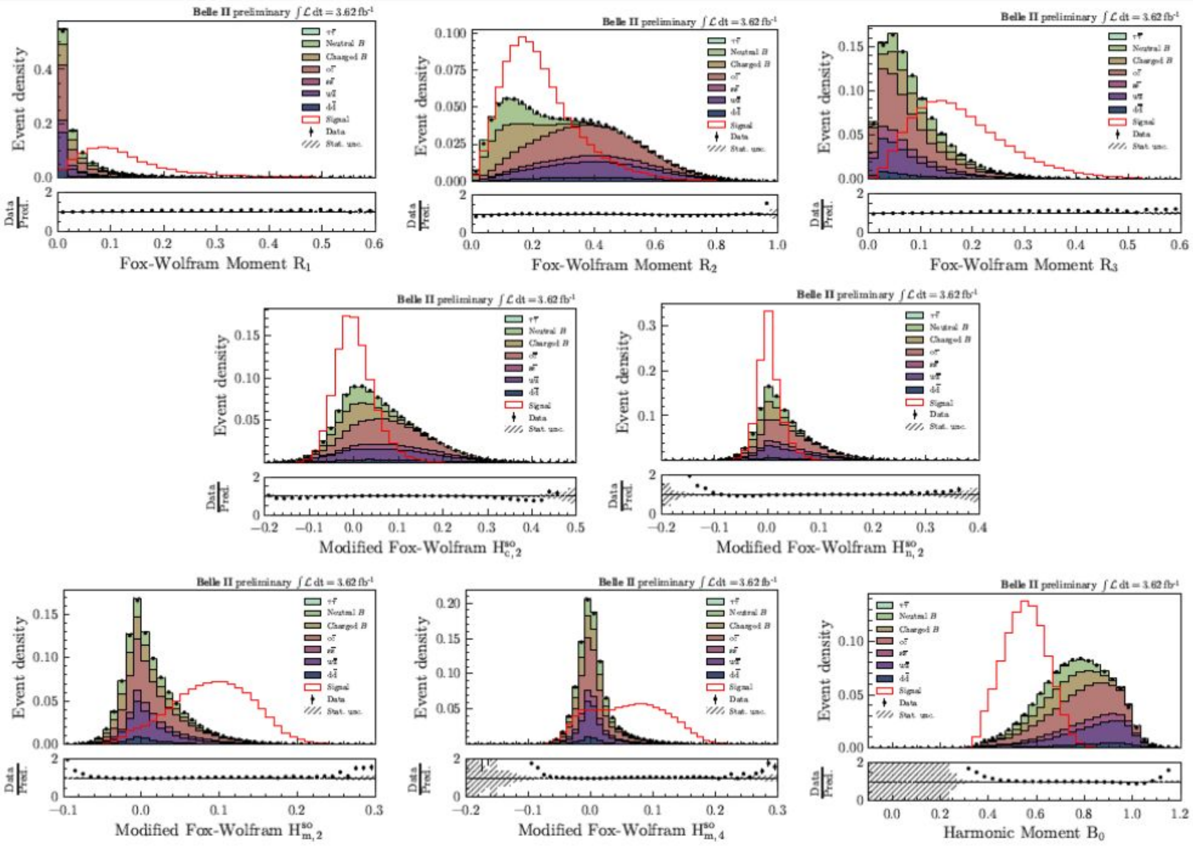
BDT input variables (ITA)



BDT input variables (ITA)



BDT input variables (ITA)

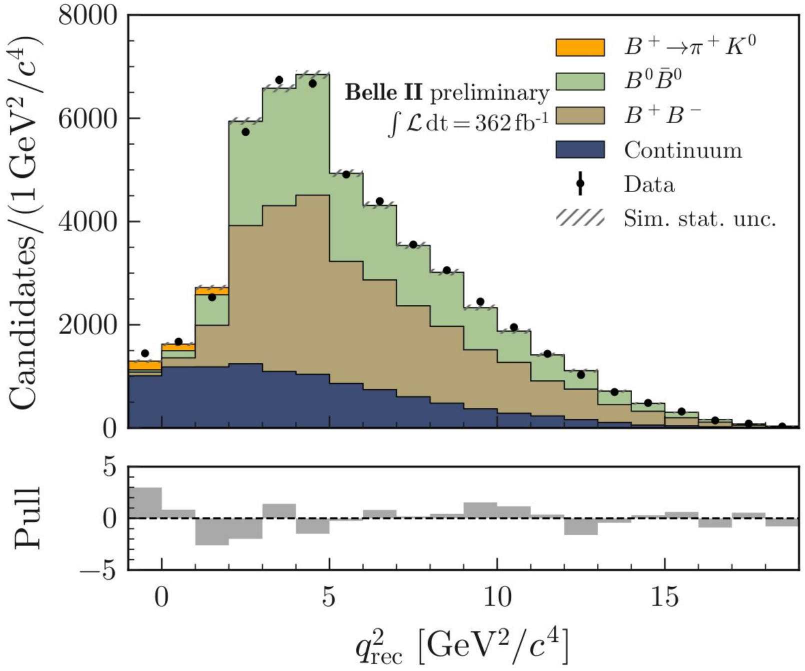


Measurement of $B^+ \rightarrow \pi^+ K^0$ (ITA only)

- **Measure the known and rare $B^+ \rightarrow \pi^+ K^0$ decay** to validate ITA analysis strategy
- Full nominal analysis chain except:
  - Pion ID instead of kaon ID
  - Different q^2 boundaries
  - Only on-resonance data
  - Only normalization systematics

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

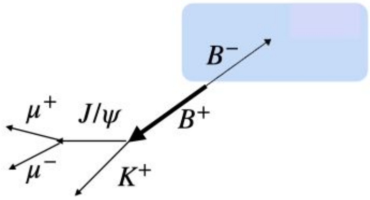
Consistent with PDG value:
$\mathcal{B}_{PDG}(B^+ \rightarrow \pi^+ K^0) = (2.3 \pm 0.08) \times 10^{-5}$

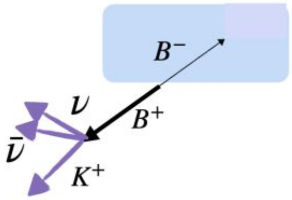


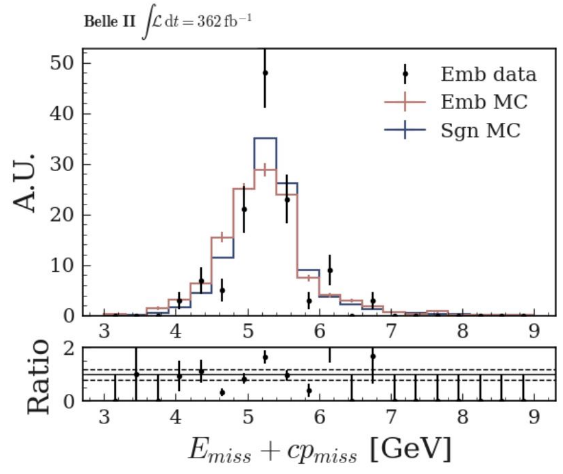
Validation of signal efficiency - HTA

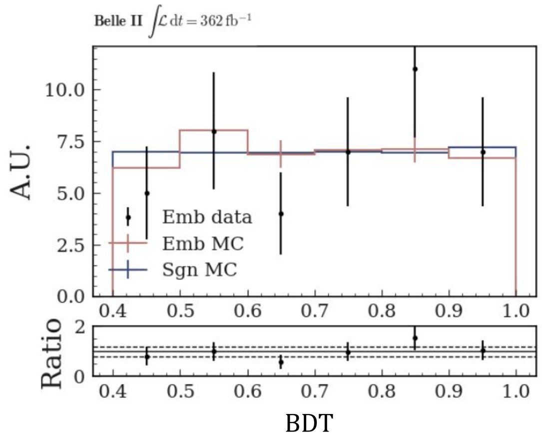
- Validate signal behavior by **embedding** signal MC into data events
- Use $B^+ \to K^+ J/\psi(\mu\mu)$ events, replace B decays by simulated signal and match kinematics
- Done for both data and simulation
- Data/MC efficiency ratio = 0.67 ± 0.06

→ Use as calibration factor and propagate uncertainty



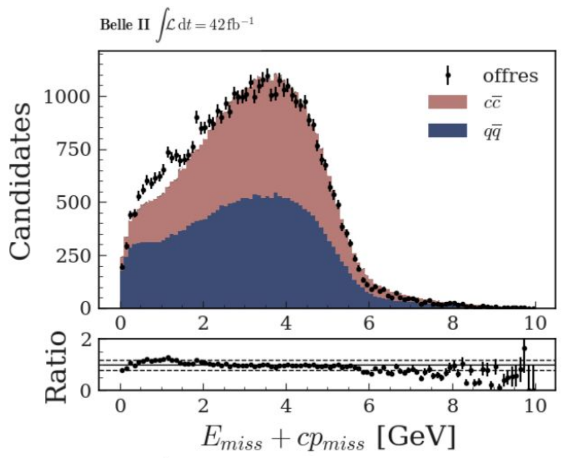


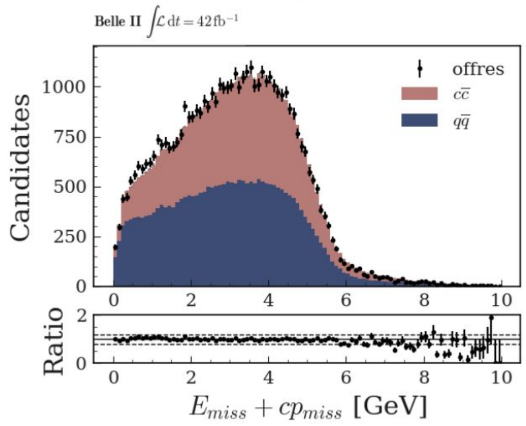




Validation of background using off-resonance data

- $c\bar{c}$ and light-$q\bar{q}$ background simulation studied in off-resonant data (collected 60 MeV below Υ(4S) mass)
- Overall acceptable agreement, but some discrepancies are seen
- In normalization: data/MC ratio = 0.82 ± 0.01 → reweighting of the simulation
- In shape: devise a correction using an additional BDT to correct simulation and derive a systematic uncertainty
- After corrections, data/MC agreement greatly improves



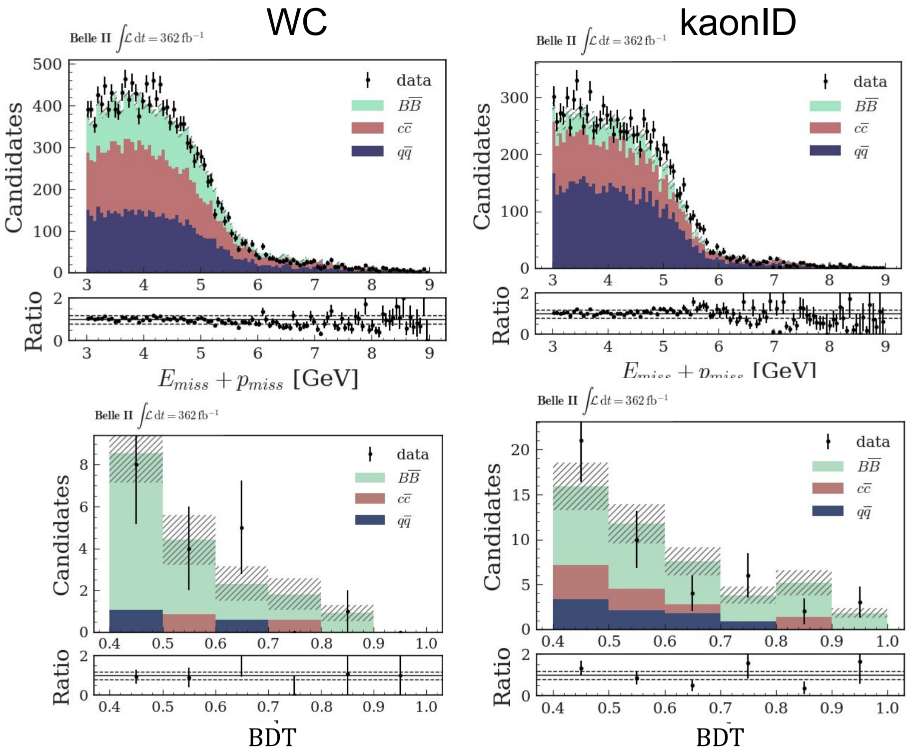


Validation of $B\bar{B}$ background using signal sidebands - HTA

- On-resonance data: need to limit signal contamination
- Same selection as signal
- Some cuts inverted to avoid looking at the SR:
  - "**Wrong charge**": the B_{tag} and B_{sig} are required to be of same electrical charge
  - "**kaonID**" the reconstructed signal kaon is required to be compatible with the pion hypothesis
- Overall acceptable data-MC agreement
- data/MC ratios are computed:

1.6 ± 0.6	**1.24 ± 0.27**
wrong charge	kaonID

- Compatible with 1 but large stat uncertainty
  → treated as systematic uncertainty

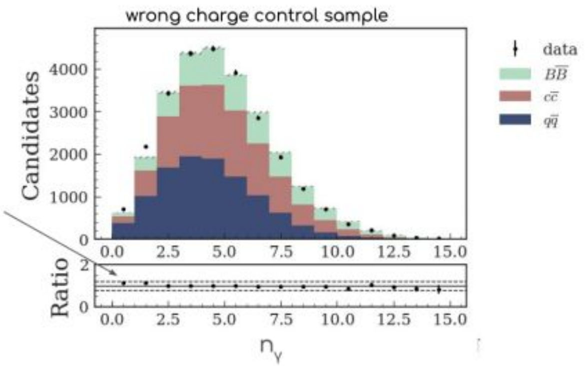


Neutral Extra energy

Correct simulation to account for residual data-simulation discrepancy by using extra photon multiplicity.

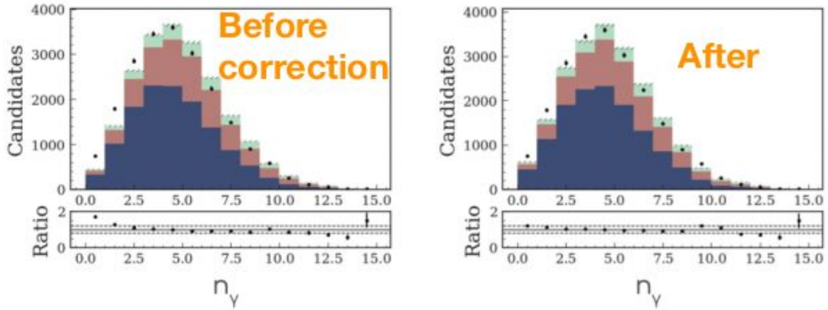
- Use wrong-charge sideband to derive the correction
- Use pion-ID sideband to validate it and estimate systematic uncertainty.

$$w_{n_\gamma} = \frac{n_{data}(n_\gamma)}{n_{MC}(n_\gamma)}$$



Apply the weight $w_{n\gamma}$ in the signal region based on the associated n_γ.

pion-ID sideband



Data-simulation agreement is improved but residual discrepancy persists.

Assign 100% of residual discrepancy as systematic uncertainty

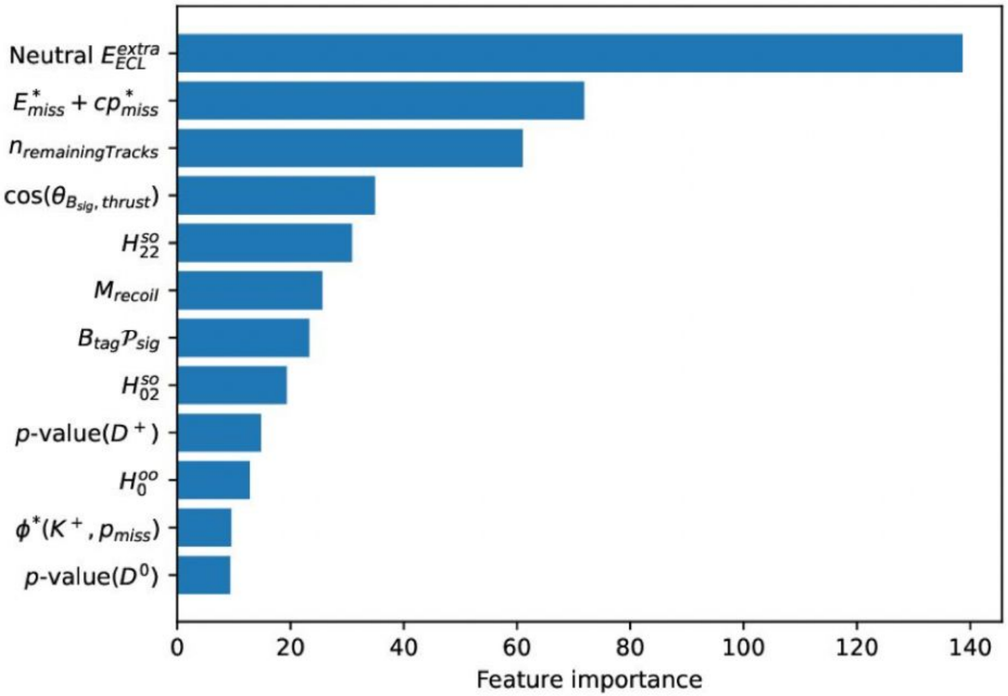
Knunu expectations - HTA

Observables	Belle $0.71\,\mathrm{ab}^{-1}$ $(0.12\,\mathrm{ab}^{-1})$	Belle II $5\,\mathrm{ab}^{-1}$	Belle II $50\,\mathrm{ab}^{-1}$
$\mathrm{Br}(B^+ \to K^+\nu\bar{\nu})$	$< 450\%$	30%	11%
$\mathrm{Br}(B^0 \to K^{*0}\nu\bar{\nu})$	$< 180\%$	26%	9.6%
$\mathrm{Br}(B^+ \to K^{*+}\nu\bar{\nu})$	$< 420\%$	25%	9.3%
$F_L(B^0 \to K^{*0}\nu\bar{\nu})$	–	–	0.079
$F_L(B^+ \to K^{*+}\nu\bar{\nu})$	–	–	0.077
$\mathrm{Br}(B^0 \to \nu\bar{\nu}) \times 10^6$	< 14	< 5.0	< 1.5
$\mathrm{Br}(B_s \to \nu\bar{\nu}) \times 10^5$	< 9.7	< 1.1	–

Expected sensitivities

BDT features - HTA

- Sum of photon energy deposits in ECL in ROEh
- Number of tracks in ROEh
- Sum of the missing energy and absolute missing three-momentum vector
- Azimuthal angle between the signal kaon and the missing momentum vector
- Cosine of the angle between the thrust axis of the signal kaon candidate and the thrust axis of the ROEh
- Kakuno-Super-Fox-Wolfram moments H_{22}^{so}, H_{02}^{so}, H_0^{oo}
- Invariant mass of the tracks and energy deposits in ECL in the recoil of the signal kaon
- p-value of B_{tag}
- p-value of the vertex fit of the signal kaon and one or two tracks in the event to reject fake kaons coming from D^0 or D^+ decays



Prospects - HTA