Search for $B \to K^{(*)}\nu\bar{\nu}$ decays at Belle II using semileptonic tagging

Merna Abumusabh, Corentin Santos

IPHC,
CNRS,
University of Strasbourg,

13 November 2025

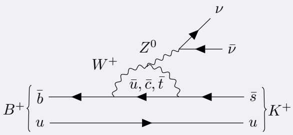
Overview

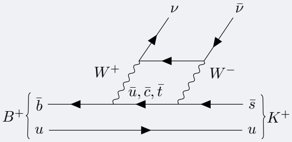
The $b \to s\nu\bar{\nu}$ transitions are flavor-changing neutral currents (FCNC) that are forbidden at tree level in the Standard Model (SM) and proceed through higher-order loop diagrams.

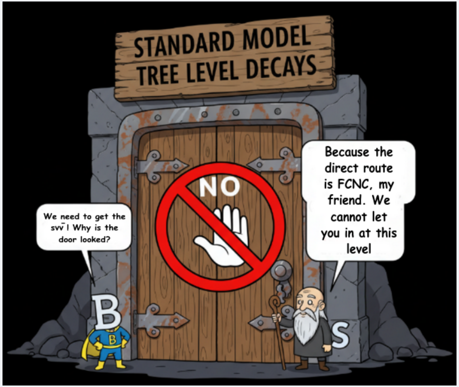
Why $B \to K^{(*)}\nu\bar{\nu}$?

- Precisely predicted in the Standard Model (SM).

- Since neutrinos are not measured, also sensitive to light new particles.



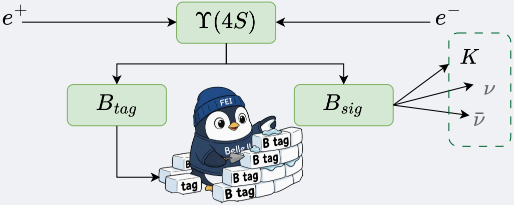




"Based on a true story"

The experimental challange

The challenge is that we have two neutrinos, so we cannot directly reconstruct the final state. To overcome this, we use the fact that at the $\Upsilon(4S)$, B mesons are produced in pairs.



High efficieny

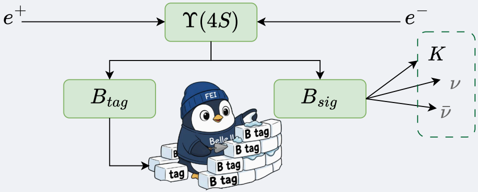
Full Event Interpretation

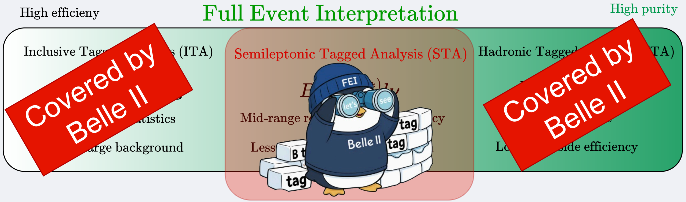
High purity

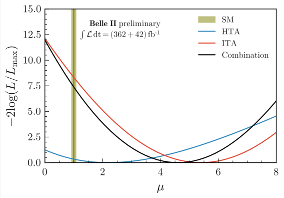
Inclusive Tagged Analysis (ITA)	Semileptonic Tagged Analysis (STA)	Hadronic Tagged Analysis (HTA)
$B \to$ Anything	$B \to D^{(*)} l\nu$	$B \to$ Hadrons
Very large statistics	Mid-range reconstruction efficiency	Cleaner sample
Very large background	Less constrained reconstruction	Lower tag-side efficiency

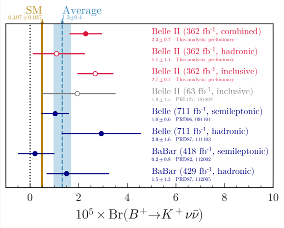
The experimental challenge

The challenge is that we have two neutrinos, so we cannot directly reconstruct the final state. To overcome this, we use the fact that at the $\Upsilon(4S)$, B mesons are produced in pairs.









Current status and targeted channels

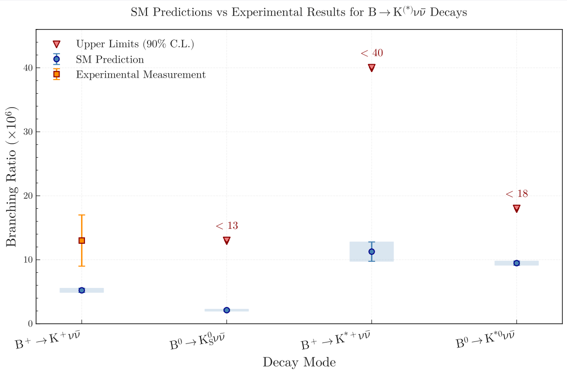
Targeted channels:

$B^+ \longrightarrow K^+\nu\bar{\nu}$

$B^0 \longrightarrow K_S^0\nu\bar{\nu}$

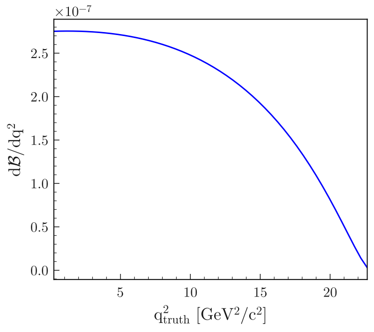
$B^+ \longrightarrow K^{*+}\nu\bar{\nu}$ with $K^{*+} \longrightarrow K^+\pi^0$ and $K^{*+} \longrightarrow K_S^0\pi^+$

$B^0 \longrightarrow K^{*0}\nu\bar{\nu}$ with $K^{*0} \longrightarrow K^+\pi^-$



Form factors corrections - In progress

One of the main sources of theoretical uncertainties in the SM differential decay rate of both the $B \to K\nu\bar{\nu}$ and $B \to K^*\nu\bar{\nu}$ decays are coming from the local form factors.



$$\frac{d\mathcal{B}}{dq^2}(B \to K\nu\bar{\nu}) = \tau_B \frac{G_F^2 \alpha_{\mathrm{em}}^2}{256\pi^5} \frac{\lambda_K^{3/2} q^2}{m_B^3} |C_L^{SM}|^2 |\lambda_t|^2 \boxed{[f_+(q^2)]^2}$$

$$f_+(q^2) = \frac{1}{P_+(q^2)} \sum_{n=0}^{N-1} a_n^+ [z^n - (-1)^{n-N} \frac{n}{N} z^N]$$

Form factors corrections - In progress

$$\frac{d\mathcal{B}}{dq^2}(B \to K^*\nu\bar{\nu}) = \tau_B \, \frac{G_F^2 \alpha_{\text{em}}^2}{128\pi^5} \, \frac{\lambda_{K^*}^{1/2} \, q^2}{m_B^3} \, (m_B + m_{K^*})^2 \boxed{\left|C_L^{\text{SM}}\right|^2} \boxed{|\lambda_t|^2} \boxed{F(q^2)}$$

$$F(q^2) = [A_1(q^2)]^2 + \frac{32 \, m_{K^*}^2 m_B^2}{q^2 (m_B + m_{K^*})^2} [A_{12}(q^2)]^2 + \frac{\lambda_{K^*}}{(m_B + m_{K^*})^4} [V(q^2)]^2$$

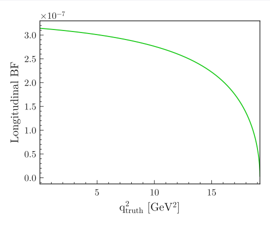
$$F_i(q^2) = P_i(q^2) \sum_k \alpha_k^i \left[z(q^2) - z(0)\right]^k$$

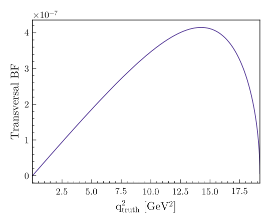
Combined fit
- Low q2 (high recoil): Light Cone Sum Rules (LCSR)
- High q2 (low recoil): Lattice QCD (LQCD)

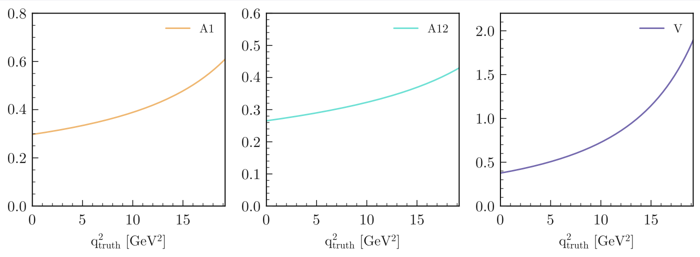
$$z(q^2) = \frac{\sqrt{t_+ - q^2} - \sqrt{t_+ - t_0}}{\sqrt{t_+ - q^2} + \sqrt{t_+ - t_0}}$$

The Form factors Corrections - In progress

For the K^*, there are three form factors $V(q^2)$, $A_1(q^2)$ and $A_{12}(q^2)$ with nine real parameters.



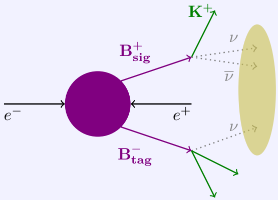




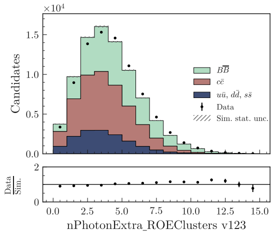
$$F_i(q^2) = P_i(q^2)\, \sum_k \alpha_k^i\, [z(q^2) - z(0)]^k$$

Analysis flow overview

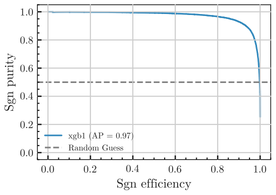
1. Reconstruction and preselection:



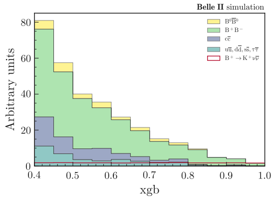
2. Data/simulation corrections:



3. Signal region selection:



4. Signal extraction:



1. Introduction
2. Analysis flow
   - Reconstruction and preselection
   - Data/simulation corrections
   - Signal region selection
   - Systematics
   - Signal extraction
3. Conclusion and outlook

Reconstruction and preselection

- Using the *semileptonic* FEI
- Tagging 4 modes:
  - $De\nu$
  - $D\mu\nu$
  - $D^*e\nu$
  - $D^*\mu\nu$

→ Normally 8 modes, but we do not consider $D^{(*)}\ell\nu\pi$ modes

- Apply a **preselection**
  - → **Best candidate selection** on the FEI output
  - → **No extra** "good" tracks, K_S, Λ or π^0 in the event

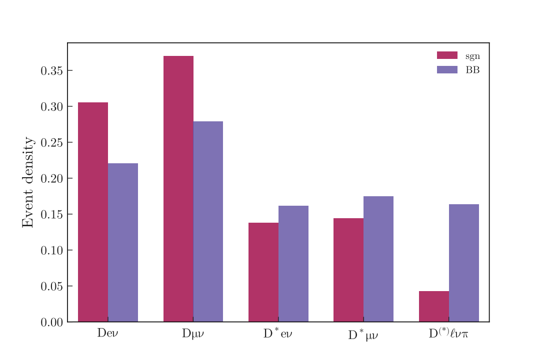


Figure: Comparison of the event density in the various FEI decay modes between signal and $B\bar{B}$ background

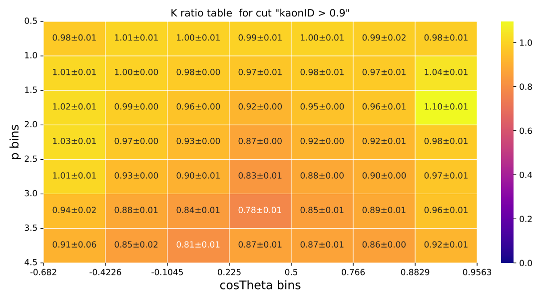
Corrections overview

Corrections	**Status**
PID	
FEI	
Signal Form Factors	
Continuum	
Extra Energy	
Backgrounds	

Table: Table of the corrections and their status

PID and FEI corrections

- PID and FEI corrections given by Belle II collaboration



modes	cal factor	error
$De\nu$	1.18	0.09
$D\mu\nu$	1.03	0.08
$D^*e\nu$	0.97	0.07
$D^*\mu\nu$	0.91	0.06

Table: Calibration factors and errors for FEI on B^+, on previous MC

Continuum correction (BDTc)

Control sample: *off-resonance sample*

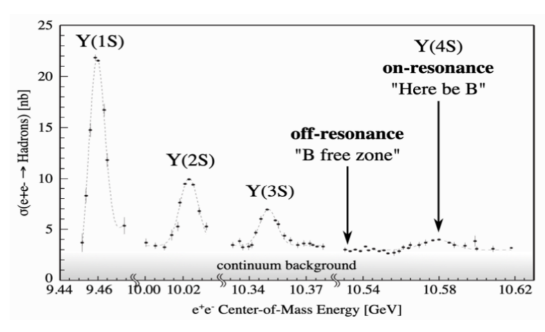


Figure: Energy spectrum showing the on-resonance and off-resonance peaks

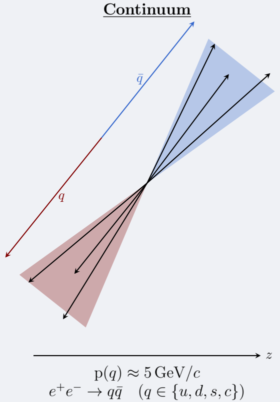


Figure: Continuum event geometry

Continuum correction (BDTc)

Correction: Reweighting the MC to the data based on the BDTc output in the *off-resonance sample*, $weight = \frac{BDTc}{1-BDTc}$

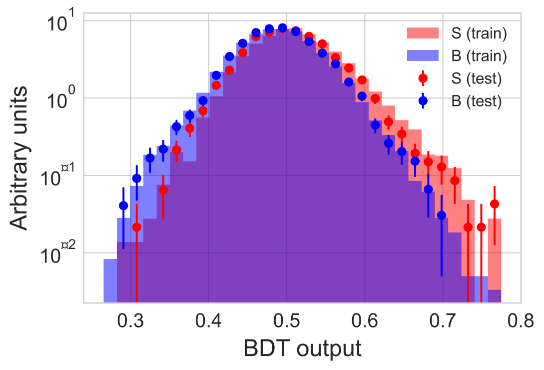


Figure: BDT output for offres data (in red) and offres MC (in blue)

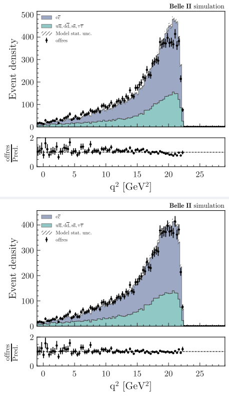


Figure: Plots of q^2 before and after BDTc correction in off-resonance sample

Extra energy correction (EExtra)

Mismodelling: Numerous unmatched clusters were observed in the ECL, mainly due to beam background
$\longrightarrow$ Bias the **extra energy** in the ECL

Control sample: *extra tracks* in the rest of event
Definition: events with no extra "very good" tracks but "good" extra tracks;
$\longrightarrow$ **Extra tracks** in the sideband have: $p_t \in]0.1, 0.2[$ Gev/c, $dr \in]0.5, 2[$ cm, $dz \in]2, 4[$ cm

Extra energy correction (EExtra)

Correction: compute weights from *extra tracks* sideband
Check: compute weights from *extra tracks* + *wrong-charge* double sideband and apply to *wrong-charge sideband* for check

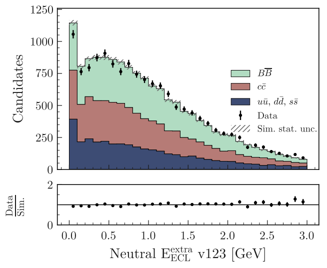


Figure: ECL extra energy in the *wrong-charge sideband* before the corrections

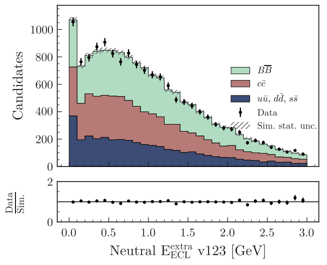


Figure: ECL extra energy in the *wrong-charge sideband* after the corrections

Background corrections

Correction: Reweight the simulation branching fractions to their *PDG* value
The reweighting is done only if the considered decay is coming from the B_{sig}

Specific correction: Correct q^2 distribution based on models for *specific backgrounds*:

- $B^+ \longrightarrow K^+ K^0_L K^0_L$
- $B^+ \longrightarrow K^+ n\bar{n}$

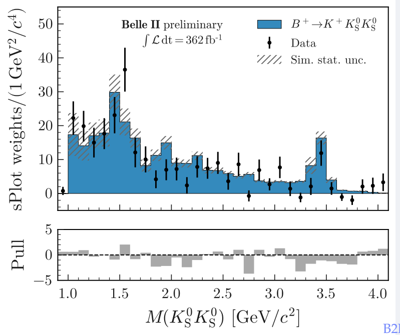


Figure: Distribution of invariant $K^0_S K^0_S$ mass in background-subtracted data and signal simulation for $B^+ \longrightarrow K^+ K^0_S K^0_S$ candidates, in previous Belle II analysis [2]

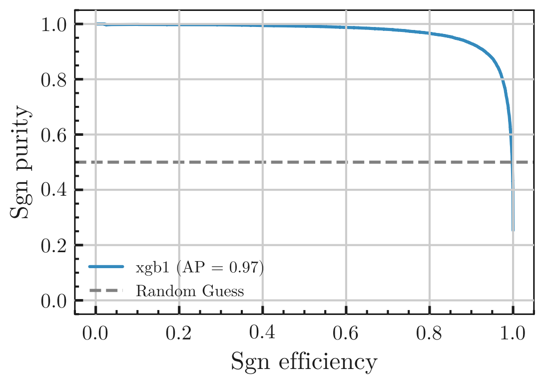
Corrections overview

Corrections	**Status**
PID	✓
FEI	*waiting for correction factors*
Signal Form Factors	✓
Continuum	✓
Extra Energy	✓
Backgrounds	*WiP*

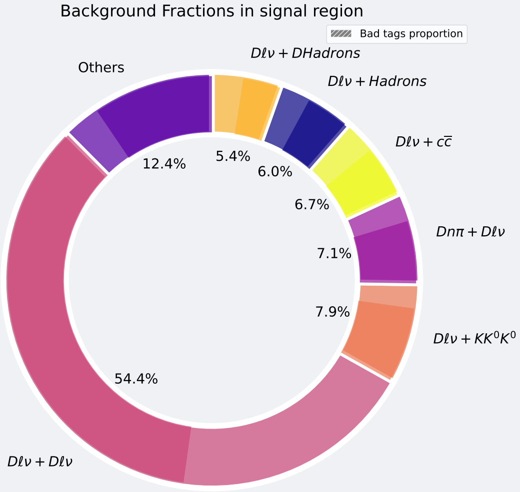
Table: Table of the corrections and their status

BDT

- Using XGBClassifier
- Using Shapley values for feature importance
- Using 27 variables (*c.f.* Slide 39 for more details)
- F_1-score for signal on validation data: 97%
- Efficiency after cut on BDT output > 0.4: 78.5×10^{-4}
- Purity after cut on BDT output > 0.4: 7%



Backgrounds



Proportion of well-tagged events: 40%

Study of $D\ell\nu + D\ell\nu$ backgrounds

- **Verification:** 199/247 signal ℓ without track found

$\longrightarrow$ Signal ℓ lost outside acceptance

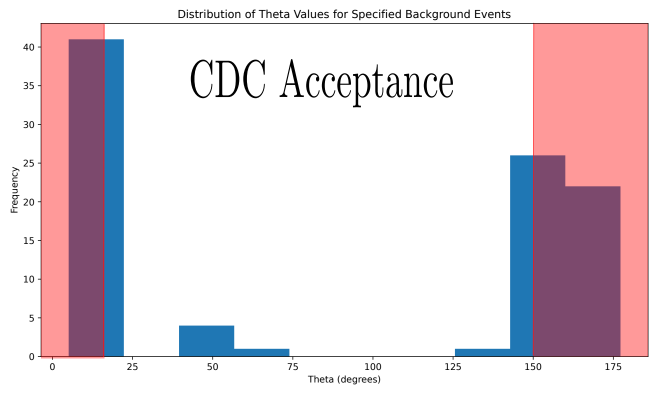


Figure: Distribution of the ℓ angle with respect to the beam axis for well tagged events

Validation

Validation of the signal efficiency using $B^+ \to J/\psi(\mu^+\mu^-)K^+$ **signal embedding**:

1. Reconstruct $B^+ \to J/\psi(\mu^+\mu^-)K^+$ **both in data and MC**
   $\longrightarrow$ contains full kinematic information of the B decay
2. Simulate desired signal decay (here $B^+ \to K^+\nu\bar{\nu}$)
3. The J/ψ decay products are **removed from the events**, leaving only *rest of event*
4. The simulated signal decay products are **added**
5. Use this new events to **compute the efficiency** of the analysis selection **in data and MC**

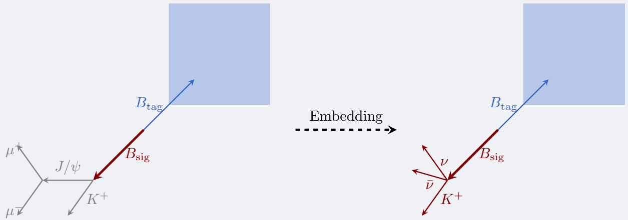


Figure: Scheme of the signal embedding procedure

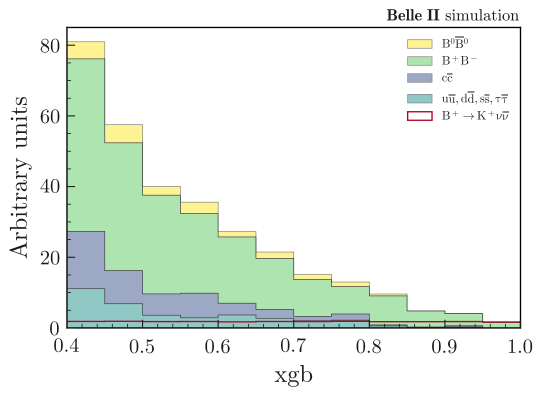
Systematics overview

Systematics	**Status**
$B\bar{B}$ background normalization	*WiP*
$q\bar{q}$ background normalization	*WiP*
PID	✓
Signal Form Factors	✓
Continuum	✓
Extra energy	✓
Backgrounds	*WiP*
Signal efficiency	*WiP*
Simulated sample size	✓

Table: Table of the systematics and their status

Signal extraction

- **Create histograms** in signal region (BDT > 0.4) with 12 bins
- Run a **maximum likelihood fit** with pyhf on the data
- Systematics are incorporated as **nuisance parameters** to constrain the fit and estimate them
- Compute the **profile likelihood ratio** with the signal strength μ as the parameter of interest
- Get the **expected error** on branching ratio and expected upper limits



$\longrightarrow$ Finalising the fit and the limit computation

Next steps

Corrections and systematics are implemented (see summary table) but still need some work
BDT work properly but can still be optimized (reduce number of features)
After some more checks, we then plan to:

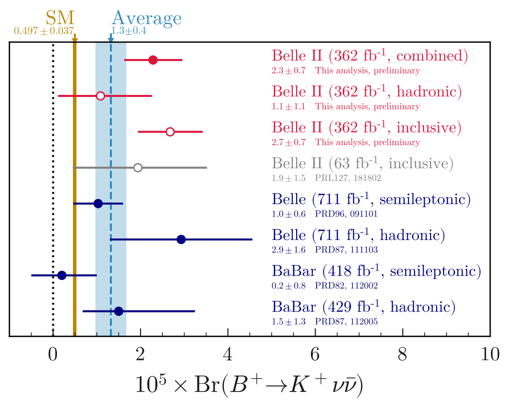
- Update FEI corrections when available
- (redo) Detailed study of the backgrounds
- Finish work on systematics

ETA: for $B^+ \rightarrow K^+\nu\bar{\nu}$, summer 2026 (end of my PhD in September 2026)
Will run this analysis flow for $B^0 \rightarrow K^{*0}\nu\bar{\nu}$ afterwards (Merna's work).

4 Supplementary Material

- Selection
- PID corrections
- Shapley values

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



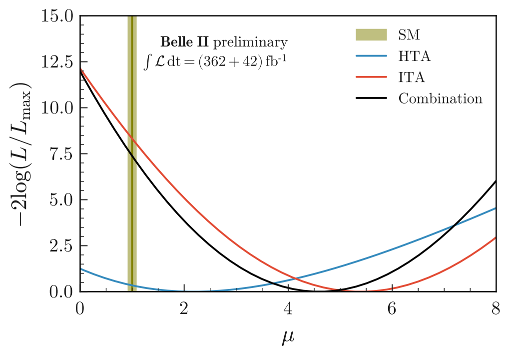
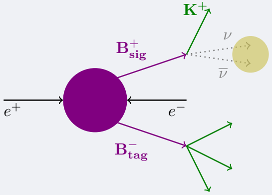


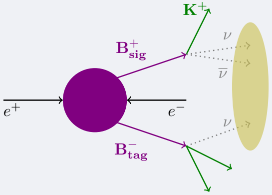
Figure: from arXiv:2301.06990

Tagged analyses

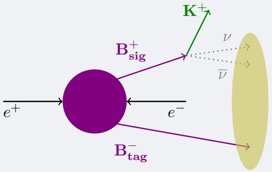
Hadronic Tagged Analysis (HTA)



Semileptonic Tagged Analysis (STA)



Inclusive Tagged Analysis (ITA)



Tagging efficiency

$\mathcal{O}\,(0.1\%)$ $\mathcal{O}\,(10\%)$

Tagging purity

$80\% - 20\%$ $\mathcal{O}\,(1\%)$

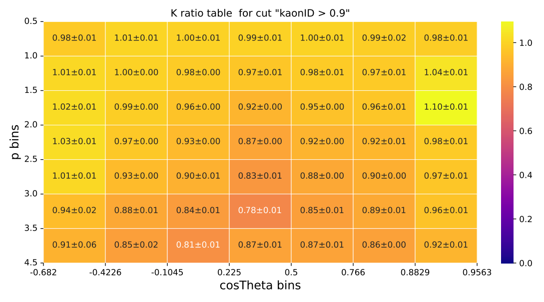
Why semileptonic tagging?

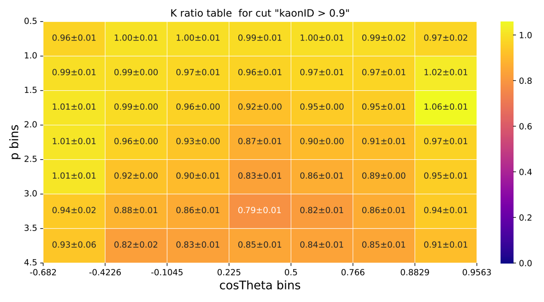
- Higher efficiency than HTA
- Higher purity than ITA
- Complementary to the other two Belle II analyses
- Complementary to the Belle and BaBar analyses

Analysis	Uncertainty on the BF naively scaled to 362 fb^{-1}
Belle HTA	2.29×10^{-5}
Belle STA	0.80×10^{-5}
BaBar HTA	1.41×10^{-5}
BaBar STA	0.81×10^{-5}
Belle II HTA	1.20×10^{-5}
Belle II ITA	0.71×10^{-5}

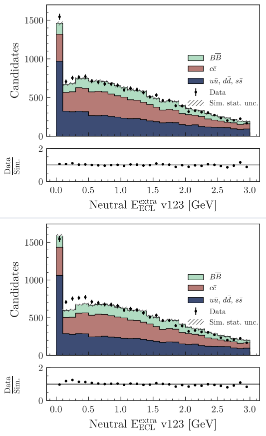
Table: Comparison of the uncertainties on the branching fraction of the various analyses, scaled to 362 fb^{-1}

PID corrections - Kaon

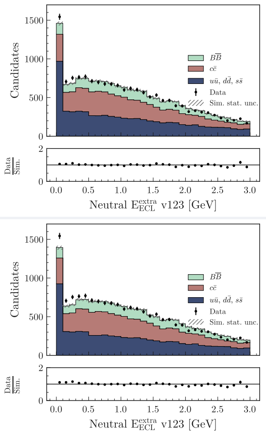




Extra tracks sideband (continued)



Extra Energy in kid sideband before and after wc+kid correction



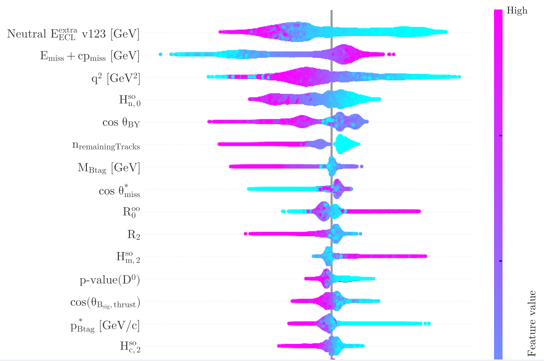
Extra Energy in kid sideband before and after extra track+kid correction

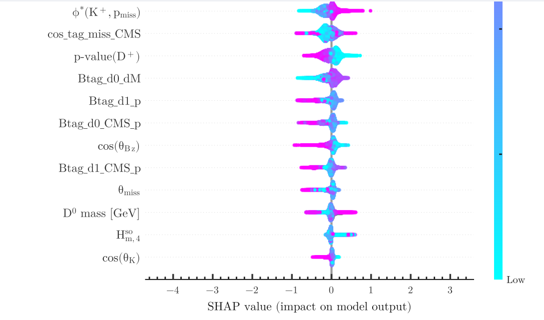
4 Supplementary Material

- Selection
- PID corrections
- Shapley values

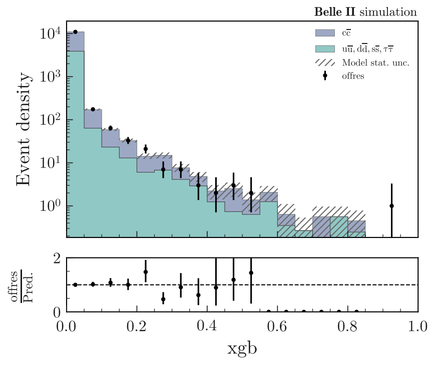
BDT input variables

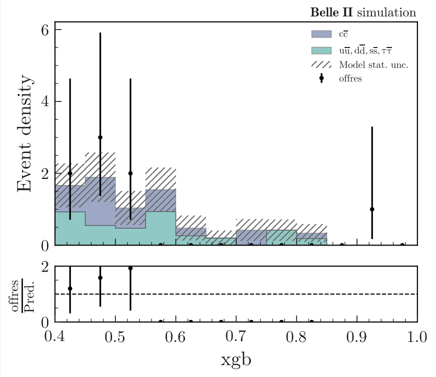
- The color gradient indicates the value of the feature
- The x-axis is the Shapley value (see Slide 54); the more on the right(left) it is, the more it contributes to the signal(background) prediction



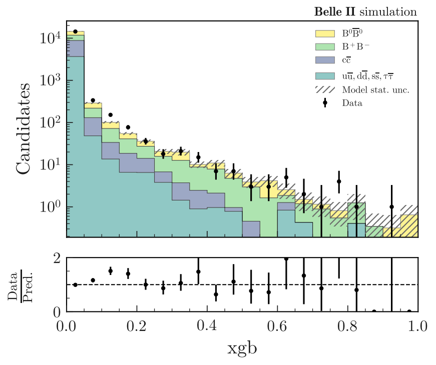


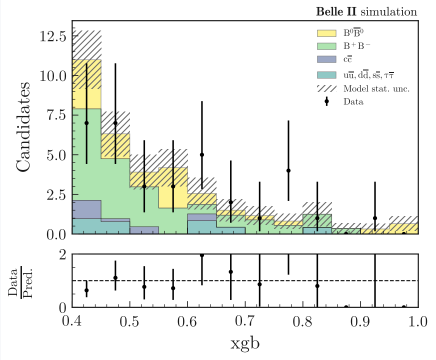
Selection on sidebands (offres)



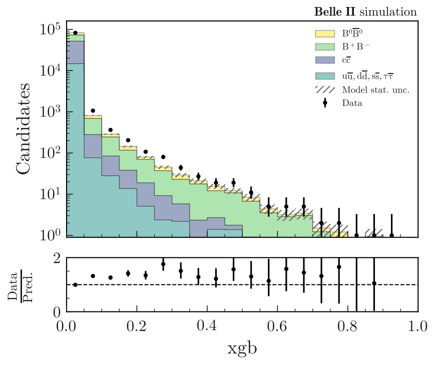


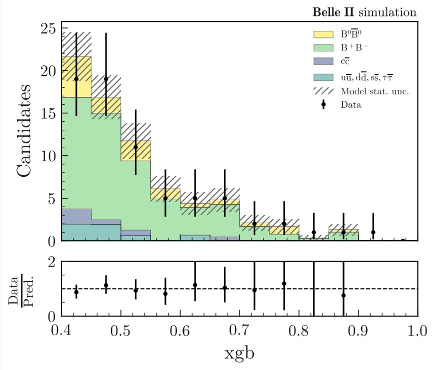
Selection on sidebands (wc)



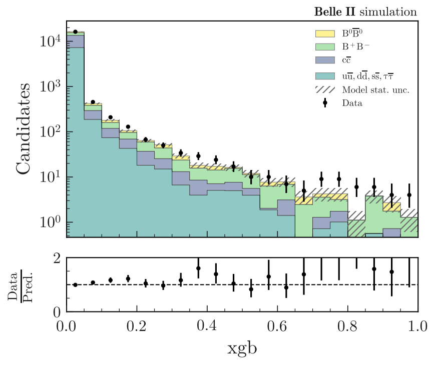


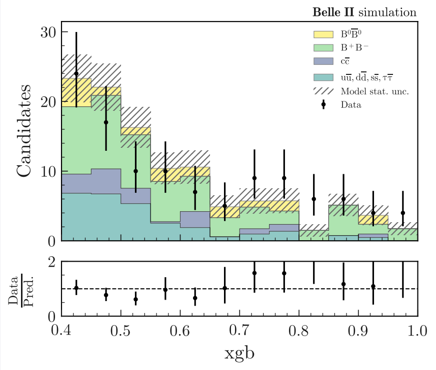
Selection on sidebands (extra)



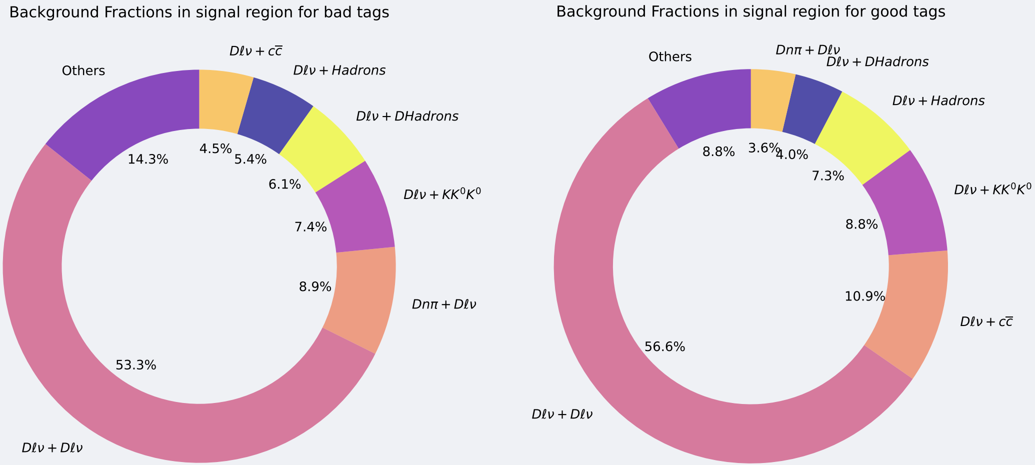


Selection on sidebands (kid)

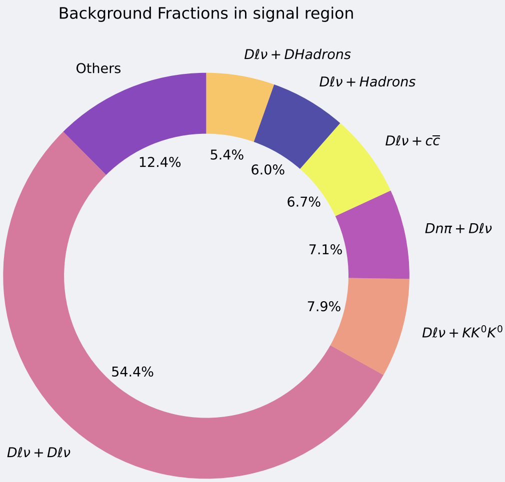




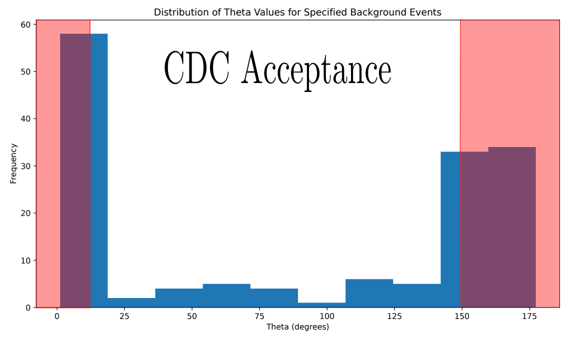
Backgrounds



Backgrounds



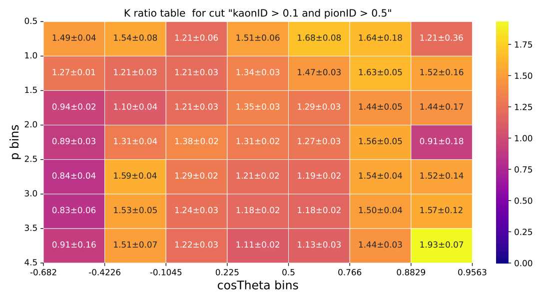
$D\ell\nu + D\ell\nu$

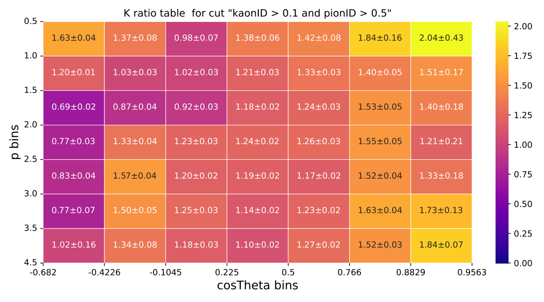


4 Supplementary Material

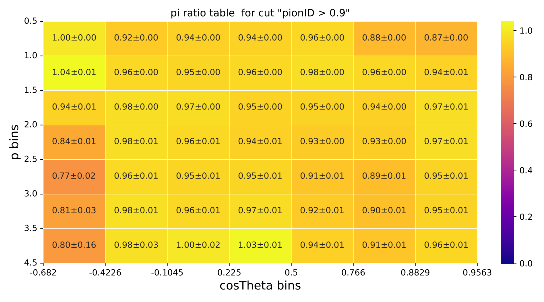
- Selection
- PID corrections
- Shapley values

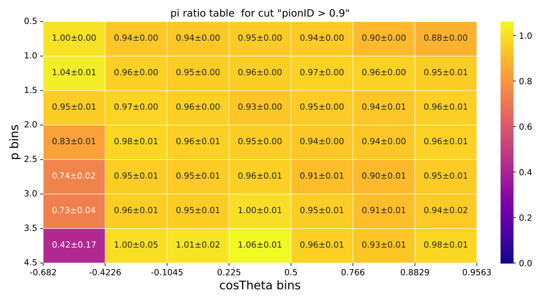
PID corrections - Kaon sideband



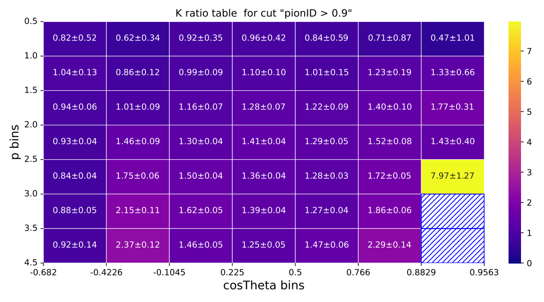


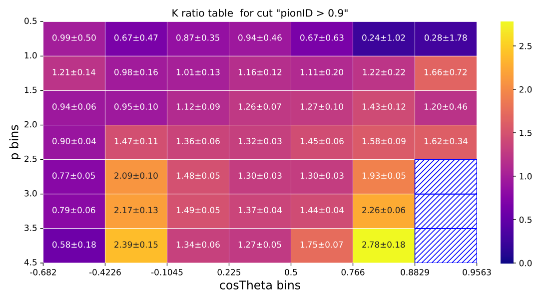
PID corrections - Pion sideband



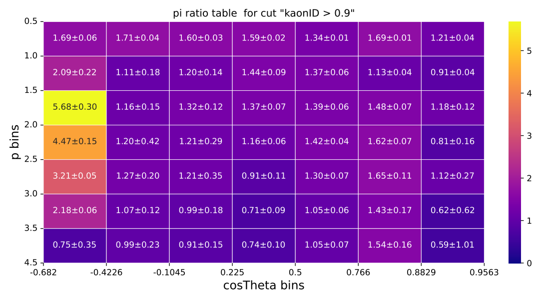


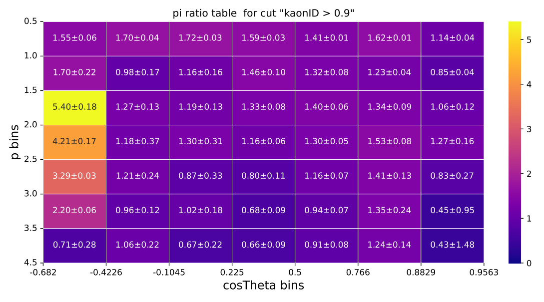
PID corrections - Pion sideband



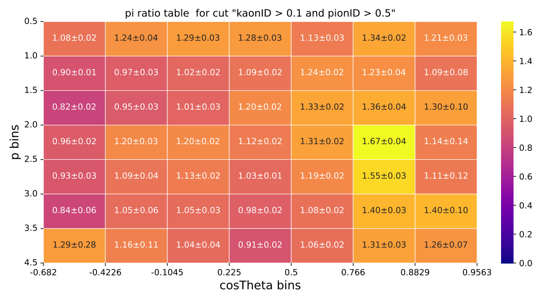


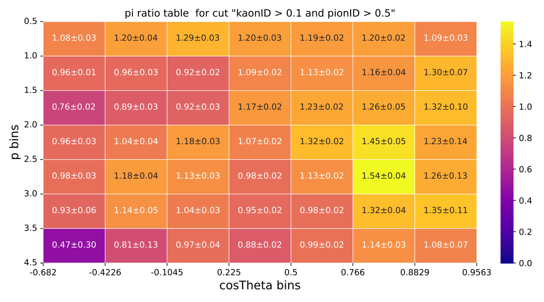
PID corrections - Pion





PID corrections - Pion





Shapley values - Explanation

- Method to attribute the prediction of a model to its features
- Originates from cooperative game theory:
  - Imagine a set N (of n players) and a gain function v that assigns a value to each coalition S
  - Question: How to fairly distribute the total gain among the players?
  - Answer: Shapley values

$$\varphi_i(v) = \frac{1}{n} \sum_{S \subseteq N \setminus \{i\}} \binom{n-1}{|S|}^{-1} \left(v(S \cup \{i\}) - v(S)\right)$$

Shapley values - Example

```python
import xgboost as xgb
import shap as sh

# ------ XGBoost Classifier ------ #
bdt = xgb.XGBClassifier(**param)
bdt.fit(X_train, y_train, sample_weight=weights_train)

# ------ Shapley Values ------ #
explainer_xgb = sh.TreeExplainer(bdt)
explanation = explainer_xgb(X_test)

# ------ SHAP Interpreter Plot ------ #
sh.plots.beeswarm(explanation,max_display=len(branches))
```

Listing: Example Python Code

[1] D. Bečirević, G. Piazza, and O. Sumensari, *Revisiting*

$$b \longrightarrow k^{(*)}\nu\bar{\nu}$$

decays in the standard model and beyond, The European Physical Journal C **83**, 10.1140/epjc/s10052-023-11388-z (2023), http://dx.doi.org/10.1140/epjc/s10052-023-11388-z.

[2] Belle II Collaboration, *Evidence for $B^+ \longrightarrow K^+\nu\bar{\nu}$ decays*, (2024), arXiv:2301.06990.