Measurement of $B^+ \to \tau^+ \nu_\tau$ branching fraction with a hadronic tagging method at Belle II

INFN Gruppo 1 – 2025, January 16 – Gaudino Giovanni

SSM Scuola Superiore Meridionale

INFN NAPOLI Istituto Nazionale di Fisica Nucleare Sezione di Napoli

Belle II

NAGOYA UNIVERSITY

Leptonic B decays

In SM decays through a $b - u$ quark annihilation mediated by W bosons.

$\longrightarrow$ Decays with helicity suppression

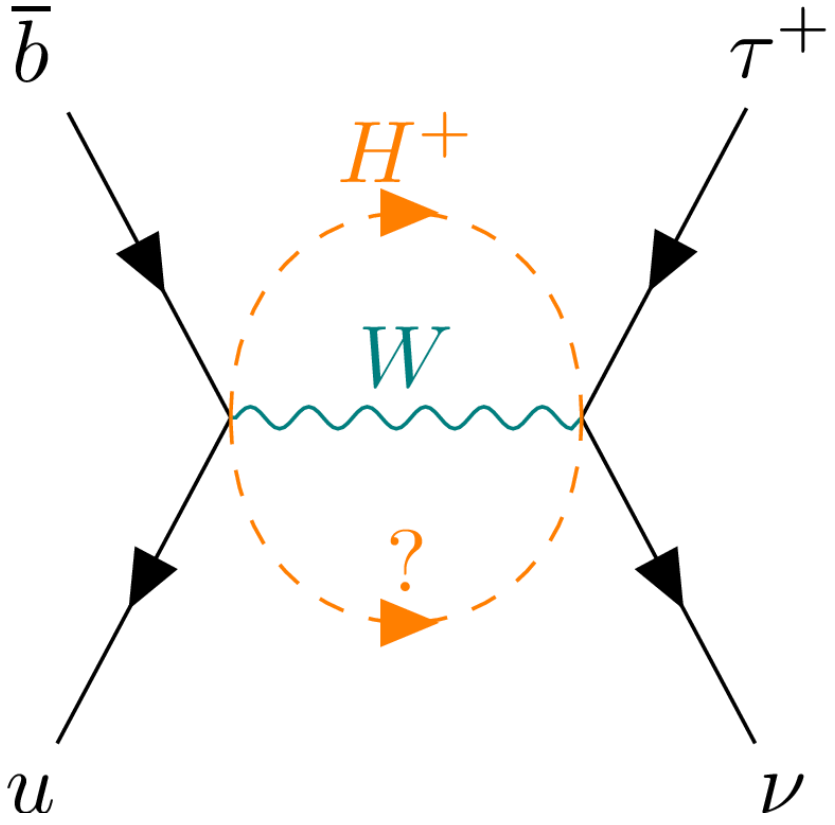
$(l = e, \mu, \tau)$

$$\mathcal{B}(B^+ \to l^+ \nu_l) = \frac{G_F^2 m_B}{8\pi} m_l^2 \left(1 - \frac{m_l^2}{m_B^2}\right)^2 f_B^2 |V_{ub}|^2 \tau_B$$

- Electrons and muons channel strongly suppressed.
- Neither Belle nor BaBar observed at "5σ" $B \to \tau\nu$
- $|V_{ub}|$ measurement with negligible theoretical uncertainty

Experiment	Tag	$\mathcal{B}(10^{-4})$
Belle	Hadronic	$0.72^{+0.27}_{-0.25} \pm 0.11$
BABAR	Hadronic	$1.83^{+0.53}_{-0.49} \pm 0.24$
Belle	Semileptonic	$1.25 \pm 0.28 \pm 0.27$
BABAR	Semileptonic	$1.8 \pm 0.8 \pm 0.2$
PDG		1.09 ± 0.24

711 /fb (Belle)
426 /fb (BABAR)

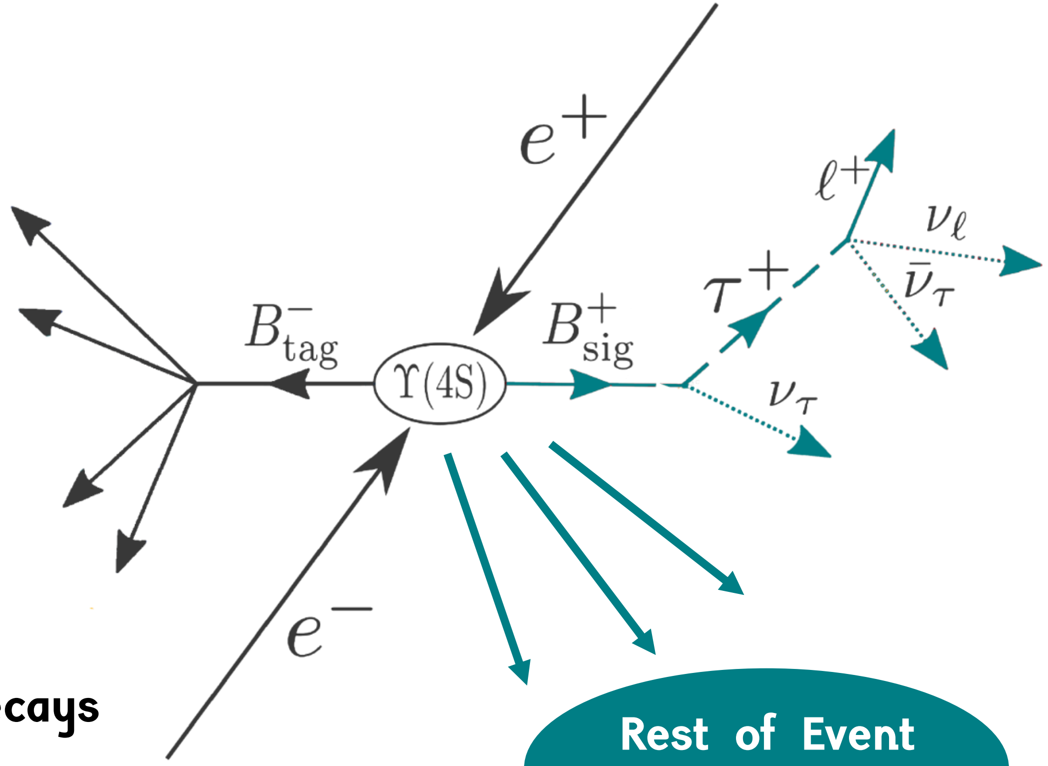


We measured the $\mathcal{B}$ with the Run 1 dataset: 365 /fb

Event Selection and dataset

One B meson is fully reconstructed using a multivariate algorithm, Full Event Interpretation (FEI) with Hadronic Tagging.

1. $\mathcal{O}_{FEI} > 10^{-2}$
2. $-0.15 < \boldsymbol{\Delta E} = \boldsymbol{E_B^*} - \sqrt{s}/2 < 0.1\ GeV$
3. $M_{bc}c^2 = \sqrt{s/4 - (p_B^* c)^2} > 5.27\ GeV$



Backgrounds
$e^+e^- \to q\bar{q}$
$e^+e^- \to \tau^+\tau^-$
$e^+e^- \to B^+B^-$
$e^+e^- \to B^0\overline{B^0}$

Signal is searched through $\boldsymbol{\tau}$ **decays**

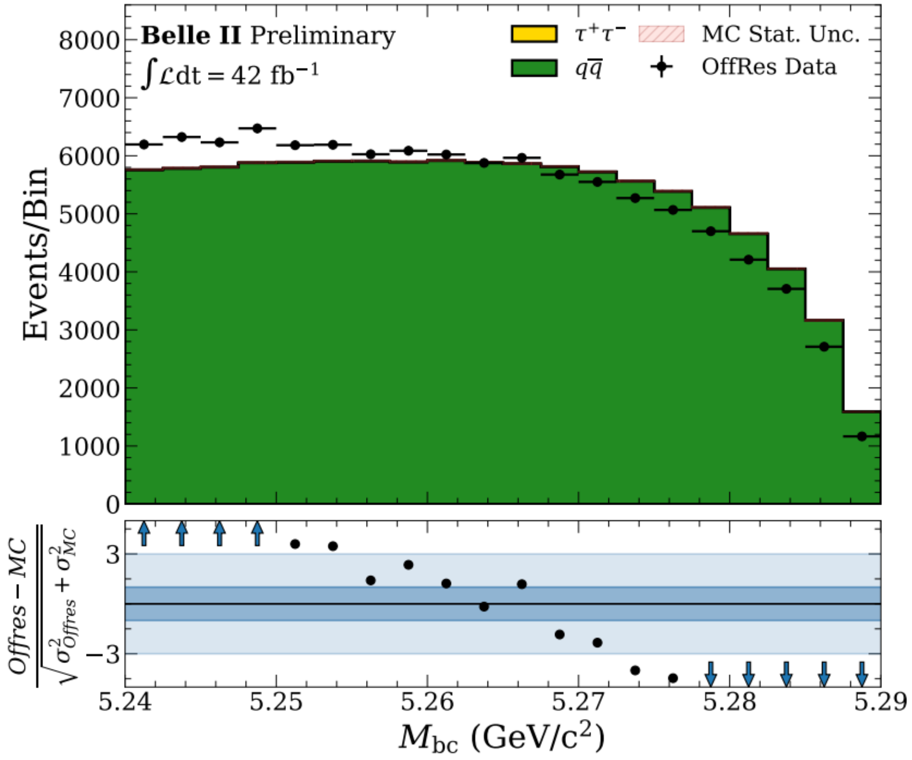
1. $\tau^+ \to \boldsymbol{e}^+\nu_e\bar{\nu}_\tau$
2. $\tau^+ \to \boldsymbol{\mu}^+\nu_\mu\bar{\nu}_\tau$
3. $\tau^+ \to \boldsymbol{\pi}^+\bar{\nu}_\tau$
4. $\tau^+ \to \rho^+\bar{\nu}_\tau$ with $\rho^+ \to \boldsymbol{\pi^+\pi^0}$

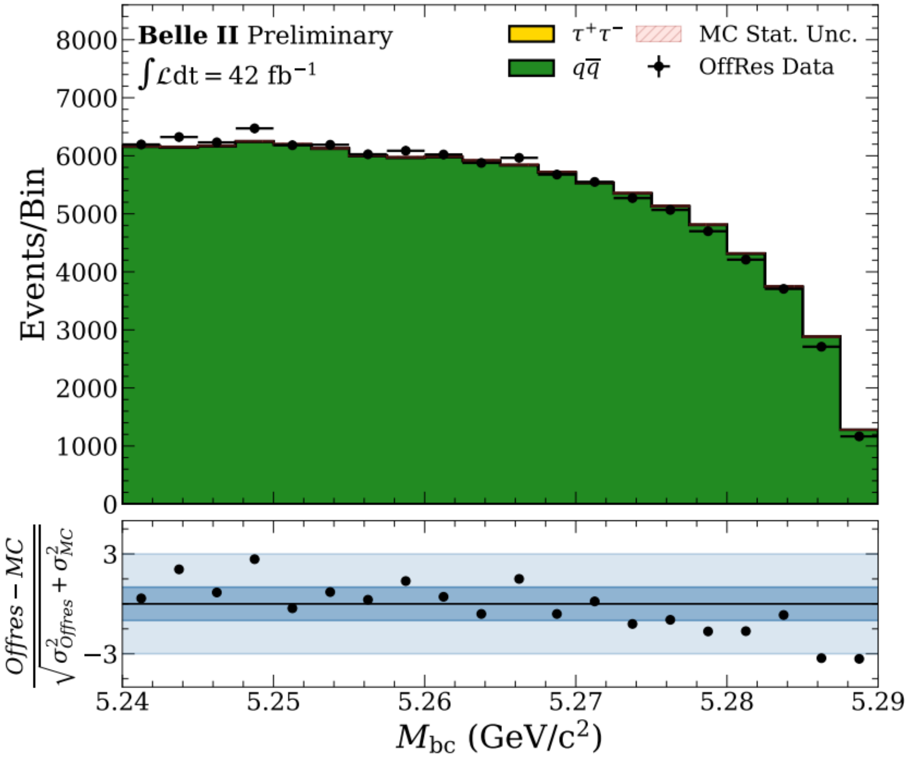
Rest of Event (ROE)

- 0 Extra Tracks
- Extra Clusters clean-up

Continuum Reweighting

We enhance MC simulation accuracy by adjusting events using multivariate analysis (MVA) to identify and correct data-MC differences. We use a Fast Boosted Decision Tree (FBDT) classifier for reweighting. Calibration involves 200/fb of continuum MC events and all off-resonance data (42/fb).





$$\Delta E_{off} = \left(\frac{E_{on}}{E_{off}}\right) E_B^* - \frac{E_{on}}{2}$$

$$M_{bc,off} = \sqrt{\frac{E_{on}^2}{4} - \frac{E_{on}^2}{E_{off}^2} p_B^{*2}}$$

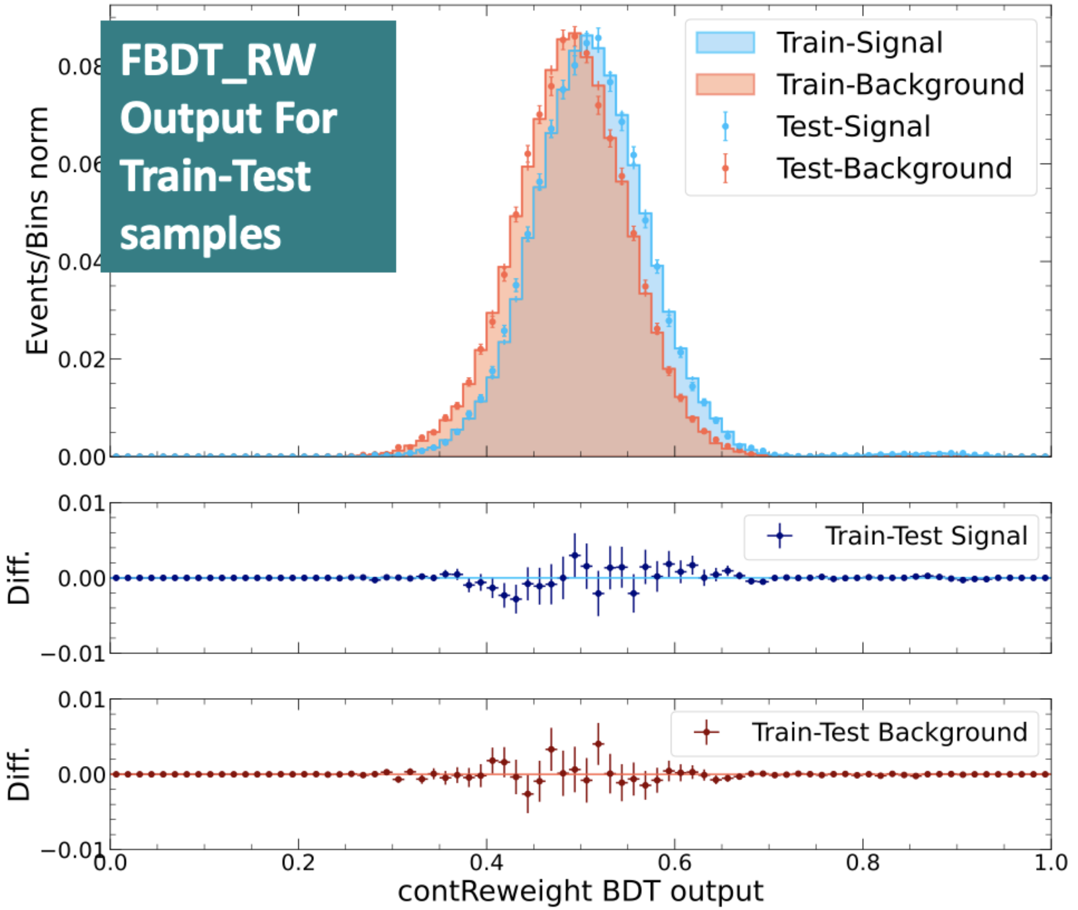
Variables distributions before and after the correction

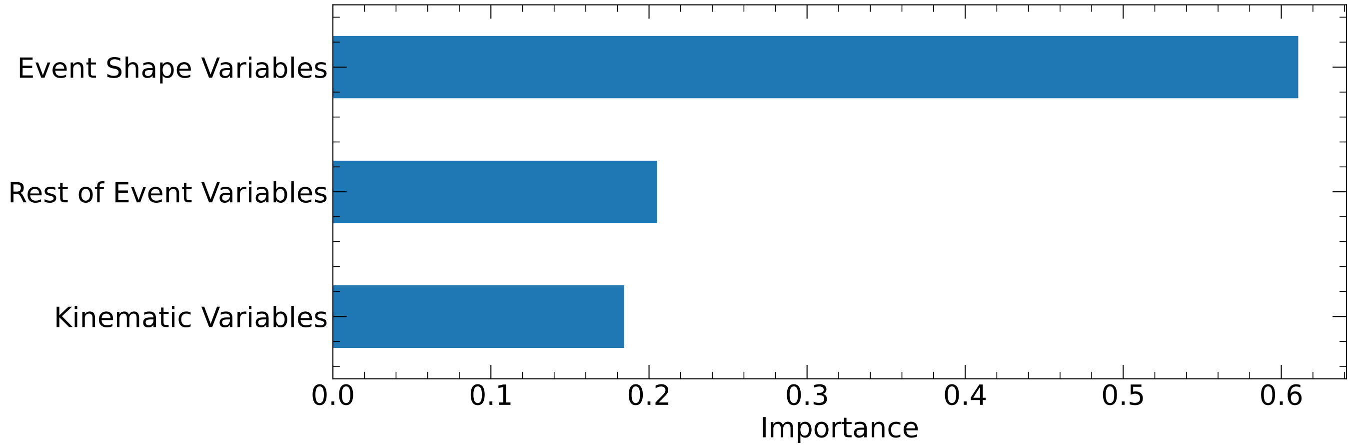
FastBDT: A speed-optimized and cache-friendly implementation of stochastic gradient-boosted decision trees for multivariate classification

Continuum Reweighting

We train a FastBDT using **Off-Res data as "Signal"** and **MC continuum as "Background" to correct the MC shape to Off-Res data.**

- 1.3M events, Train/Test sample 80%/20%



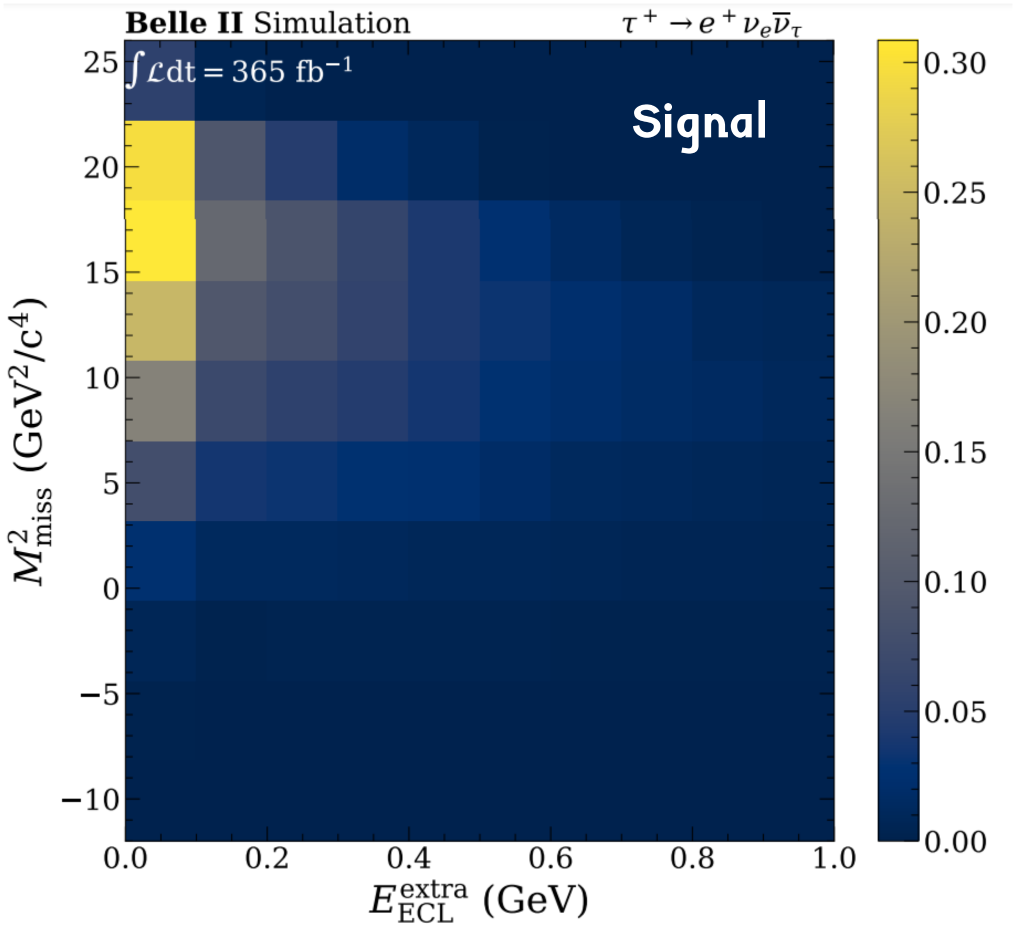


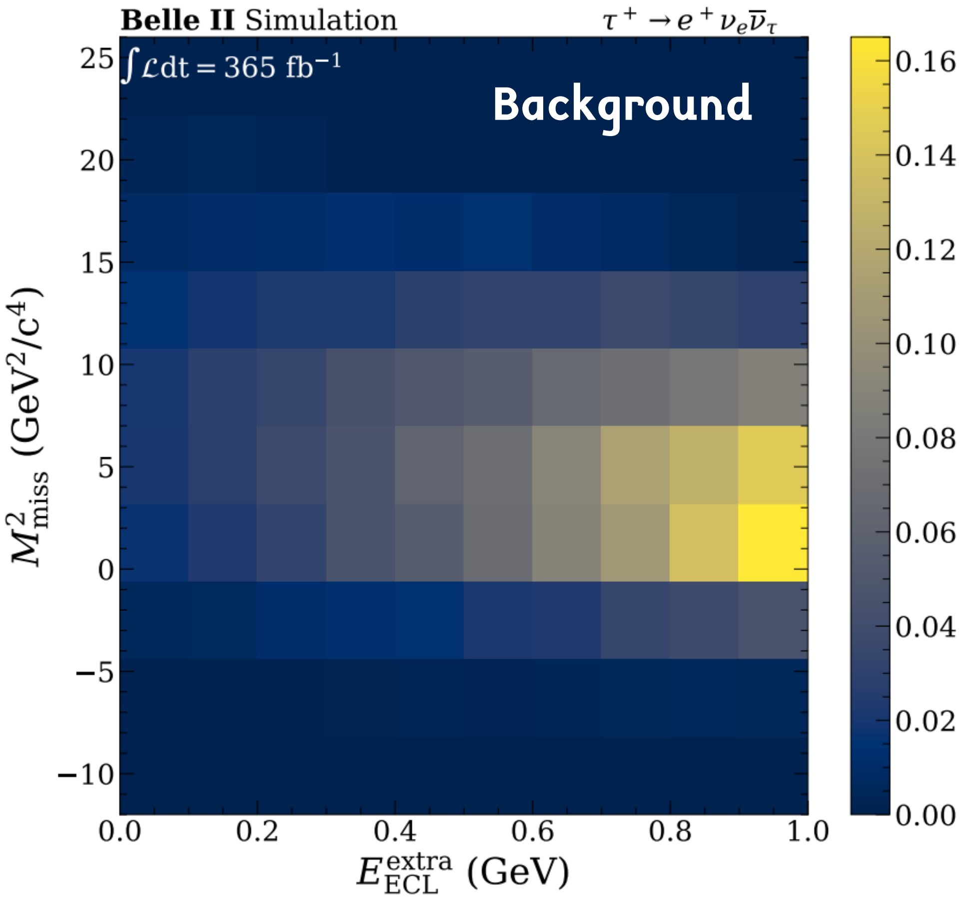
The discriminator output is transformed in an event-by-event weight to correct MC shape:

$$w_i = \frac{\mathcal{O}_{CR,i}}{1 - \mathcal{O}_{CR,i}}$$

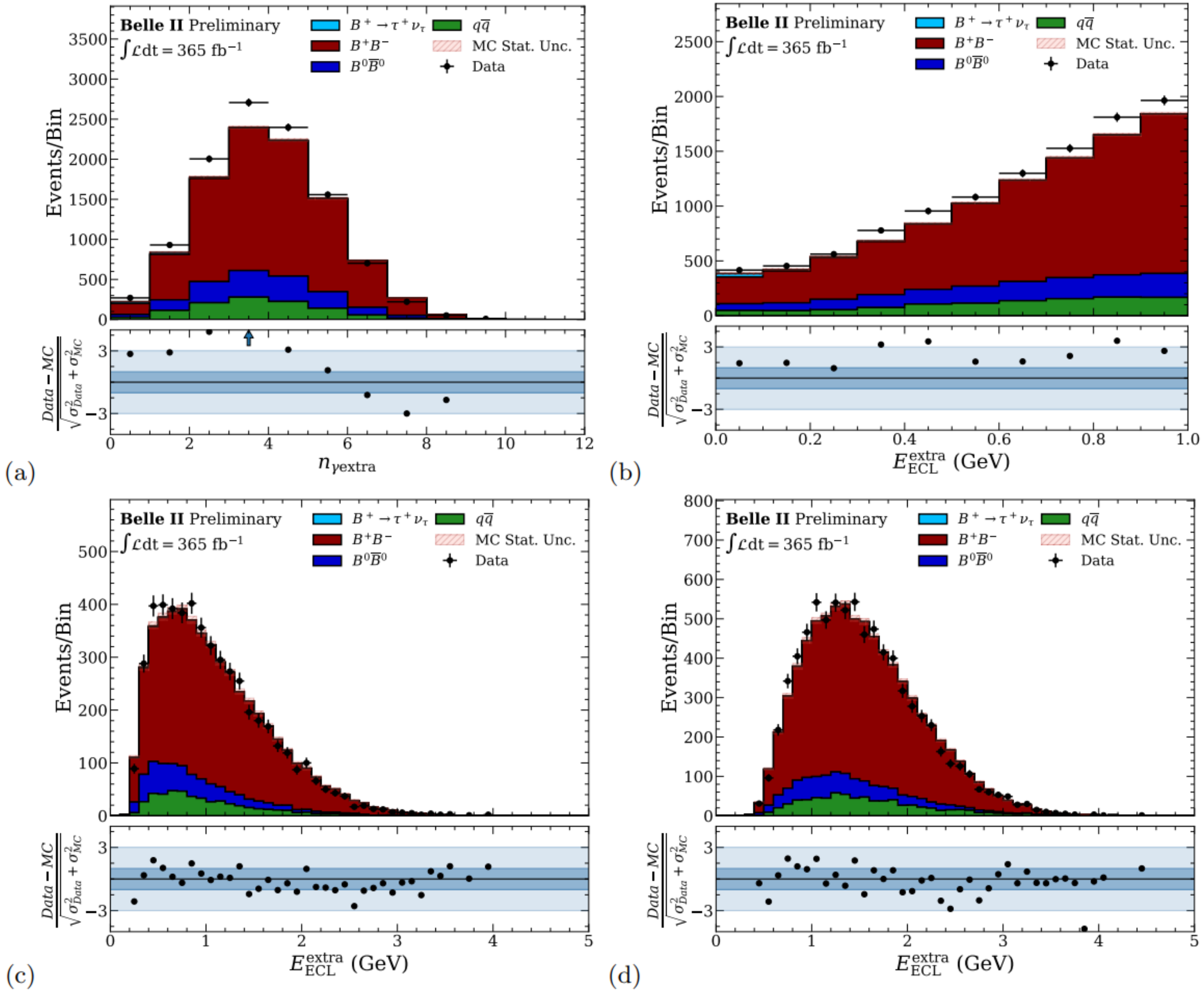
Selection Optimization

The optimization is done extracting the signal yield with a 2D Fit E_{ECL}^{extra} vs M_{miss}^2 in the signal region $[0{,}1]GeV \times [-10,\ 26]GeV^2$





Correction from the extra clusters multiplicity



$n_{\gamma extra}$ (a) and E_{extra}^{ECL} (b) in data and simulation for E extra ECL < 1 GeV.

E_{extra}^{ECL} with $n_{\gamma extra} = 3$ (c) and $n_{\gamma extra} = 5$ (d).

The number of events in simulation is scaled to the one in data to compare just the shapes.

Control sample study

The residual Data/MC disagreement related to the wrong simulation of the number of neutral clusters (γ) in ECL $\rightarrow$ to be reweighted using Control Samples.

Main Sample	Continuum	$B\bar{B}$	Sig. $\tau \to \ell\nu\nu$	Sig. $\tau \to h\nu$
Control Sample	Off-resonance Data	Extra Tracks	$B \to D^*\ell\nu$	Double Tag

- In Extra Tracks control sample, we require other extra charged tracks in addition to the signal one.
- In Double Tag control sample, we use the Hadronic Tagging multivariate FEI algorithm to reconstruct also the signal B.
- $B \to D^*\ell\nu$ control sample is reconstructed using the hadronic FEI

$\mathcal{B}(B \to \tau\nu)$ Extraction

The Branching Fraction is extracted by Simultaneous Binned Maximum Likelihood 2D Fit on E_{ECL}^{extra} and M_{miss}^2.

- **5 Free Parameters: 4 Background Yield + 1 common BF**
- **PDFs from the MC.**

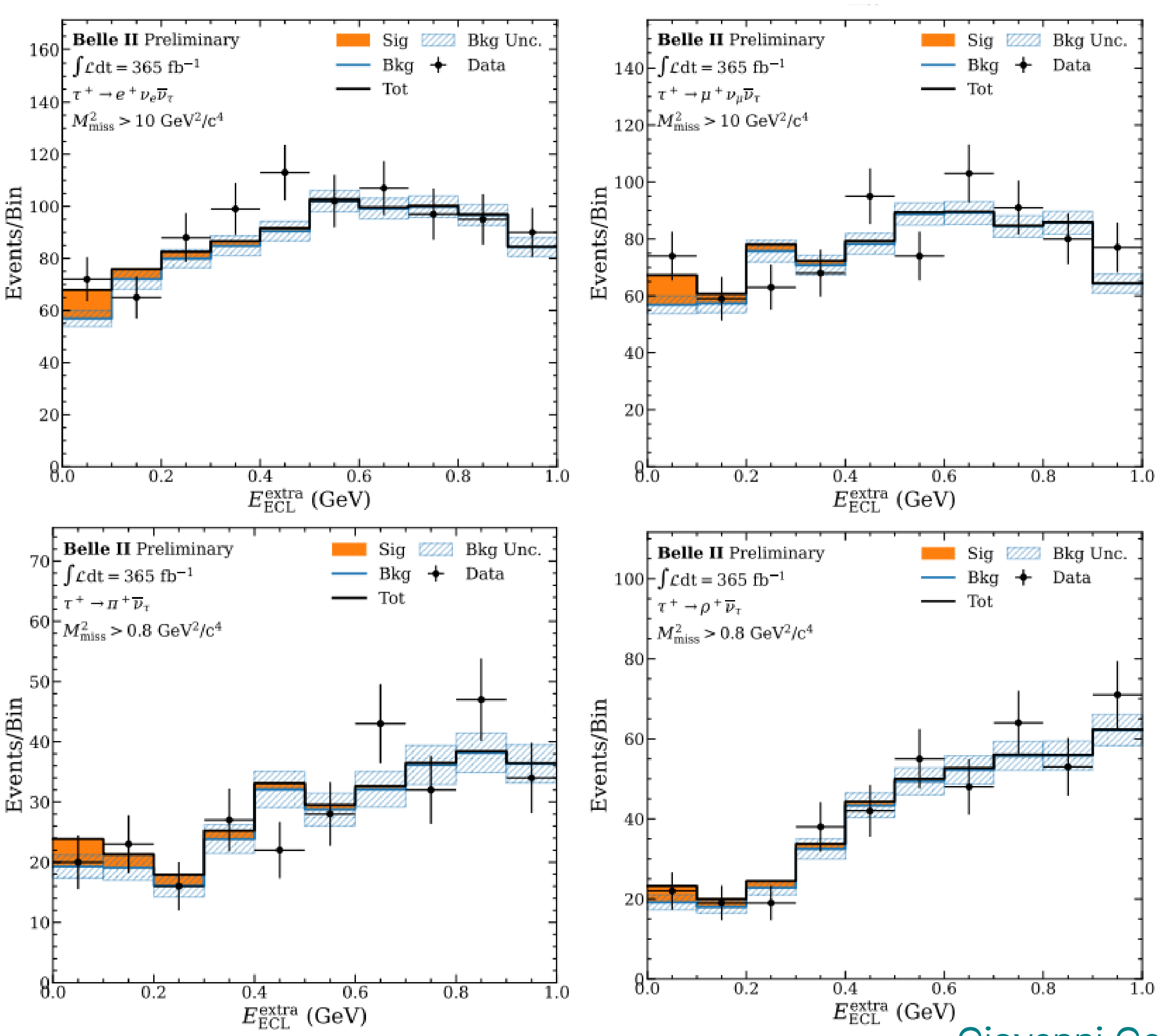
$$\boldsymbol{n_{s,k} = 2N_{\Upsilon(4S)} \cdot f_{+-} \cdot \frac{N_{reco}^{k}}{N_{gen}} \cdot B(B^+ \to \tau^+ \nu_\tau)}$$

We use f_{+-} value from HFLAV latest review: arXiv:2411.18639

Post Fit Distributions

Post–fit distributions of E_{ECL}^{extra} for the **leptonic** channels in the signal enriched region $M_{miss}^2 > 10\ GeV^2/c^4$

Post–fit distributions of E_{ECL}^{extra} for the **hadronic** channels in the signal enriched region $M_{miss}^2 > 0.8\ GeV^2/c^4$



Systematics Uncertainty

We compute the systematic uncertainties on data varying the MC shape according to the considered source.

Source	Sys.Unc.
Simulation statistics	13.3%
Fit variables PDF corrections	5.5%
Decays branching fractions in MC	4.1%
Tag B^- reconstruction efficiency	2.2%
Continuum reweighting	1.9%
π^0 reconstruction efficiency	0.9%
Continuum normalization	0.7%
Particle identification	0.6%
Number of produced $B\overline{B}$ pairs	1.5%
Fraction of B^+B^- pairs	1.4%
Tracking efficiency	0.2%
Total	15.4%

Not considered in the significance (Number of produced $B\overline{B}$ pairs, Fraction of B^+B^- pairs, Tracking efficiency)

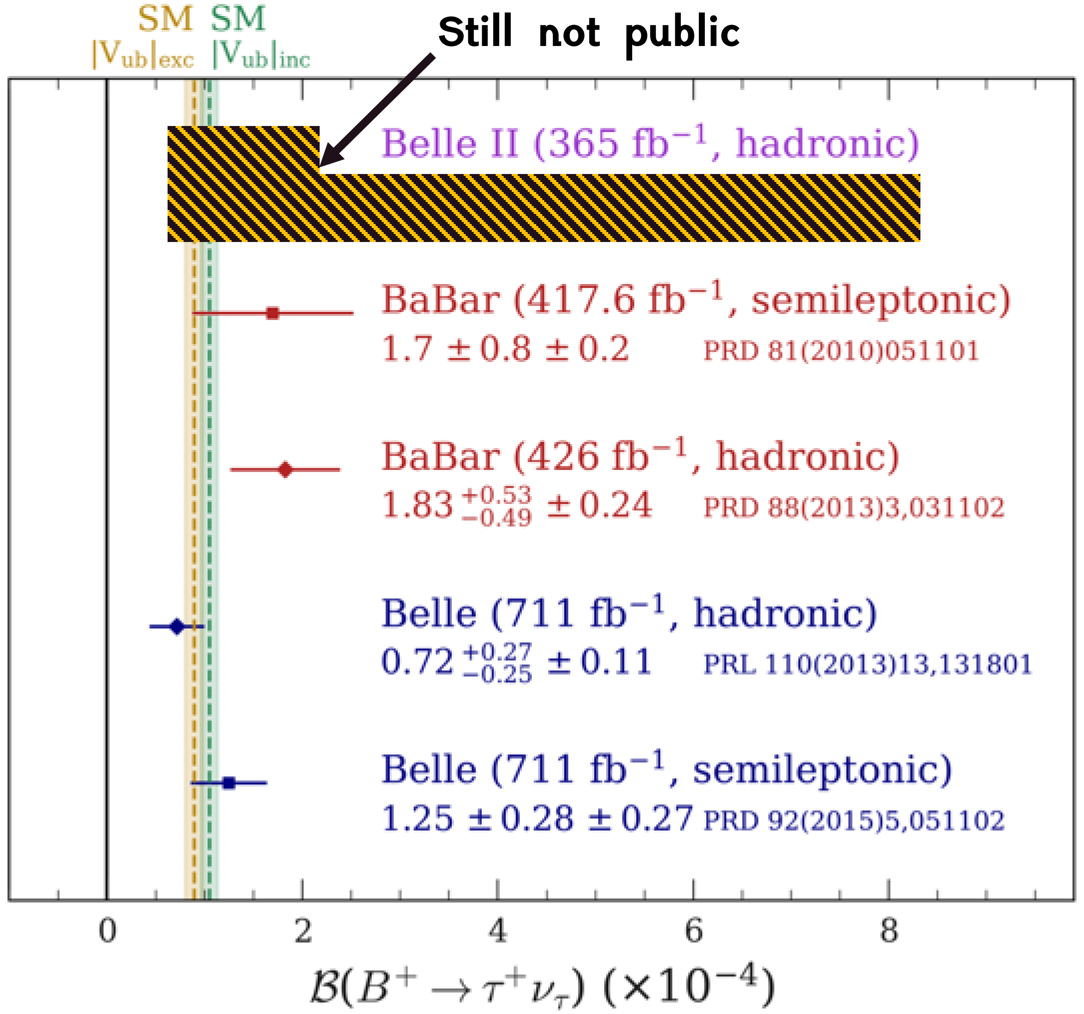
Conclusions

This analysis will be the main argument of my Ph.D. Thesis in next October.

Since the end of November, we are passing all the collaboration step in order to publish as soon as possible.

We aim to have a public result in these weeks and present it at Moriond or maybe earlier.

In this work we cooperated with Nagoya University, in particular with professor Iijima and the Ph.D. student Michele



- $|V_{ub}|_{inc}$ from arXiv:2411.18639 Inclusive from GGOU
  - $|V_{ub}|_{exc}$ from arXiv:2411.18639 from $B \to \pi \ell \nu$

That’s all!

Thanks for the attention

Motivations of (Semi)Leptonic B decays

Lepton-Flavor Universality tests | SM Precision Measurements

- In SM, the W boson couples equally to τ, μ, e → Lepton-Flavor Universality (LFU)
- (Semi)Leptonic B decays are sensitive to new physics beyond SM

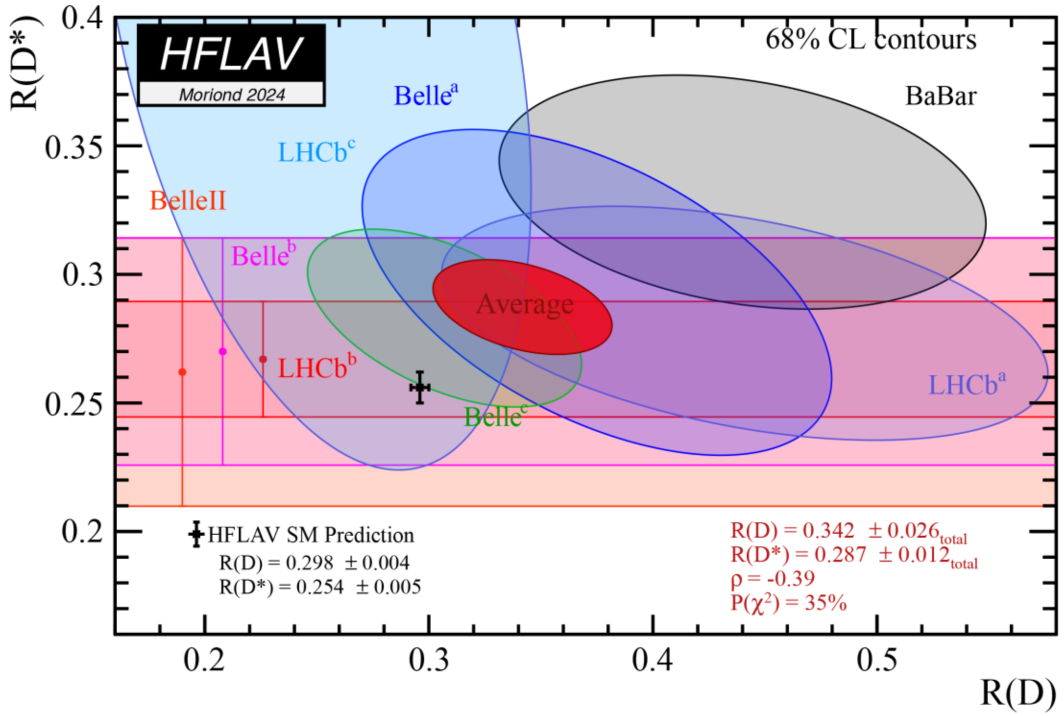
$$R\left(H_{\tau/\ell}\right) = \frac{B(B \to H\tau\nu)}{B(B \to H\ell\nu)}$$

$H = D^{(*)}, X, \pi, \ldots$

$\ell = e, \mu$

Tension of $R\left(D^{(*)}_{\tau/\ell}\right)$ with SM ~3σ

$$\mathcal{B}(B^+ \to \ell^+\nu) = \frac{G_F^2 m_B m_\ell^2}{8\pi}\left[1 - \frac{m_\ell^2}{m_B^2}\right] f_B^2 \left|V_{ub}\right|^2 \tau_B$$



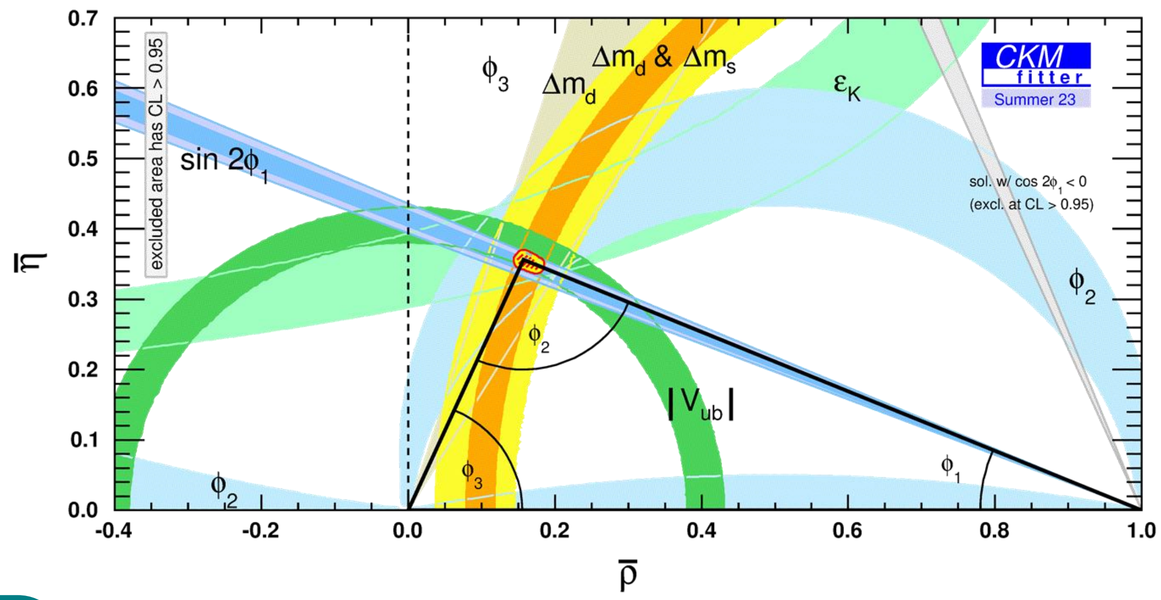
Motivations of (Semi)Leptonic B decays

Lepton–Flavour Universality tests

SM Precision Measurements

- $|V_{ub}|$ and $|V_{cb}|$ important to **constrain** CKM Unitarity
- **Precisely** measured with semileptonic B decays
- Independent measurement using leptonic B decays

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



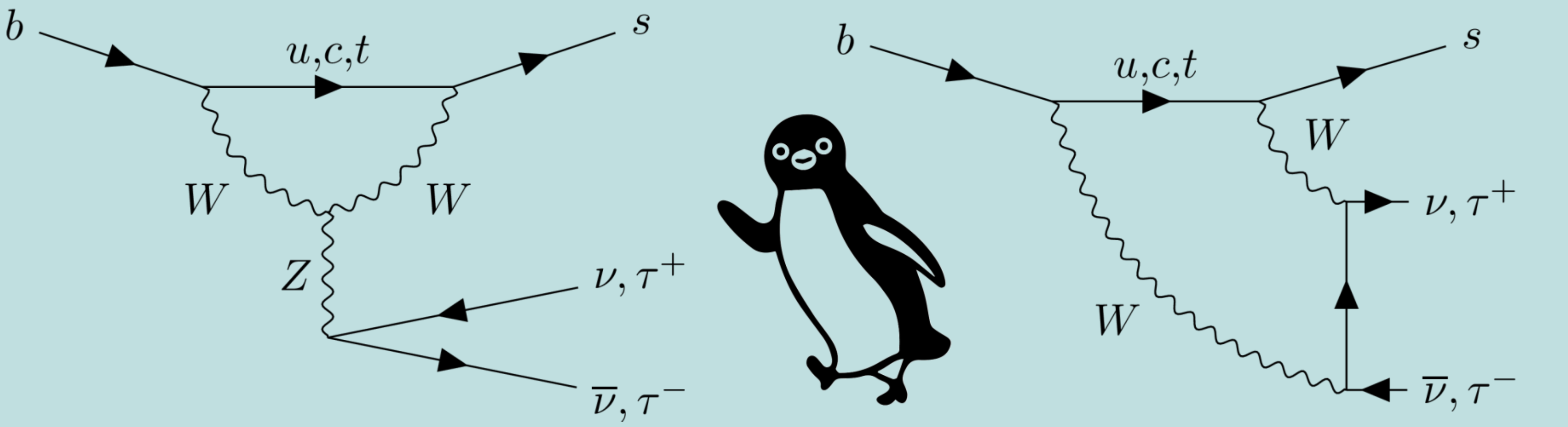
$$\mathcal{B}(B^+ \to \ell^+\nu) = \frac{G_F^2 m_B m_\ell^2}{8\pi}\left[1 - \frac{m_\ell^2}{m_B^2}\right] f_B^2 \, |V_{ub}|^2 \, \tau_B$$

Longstanding **tension** between exclusive and inclusive determinations

Motivations of (Semi)Leptonic B decays

Lepton-Flavour Universality tests | SM Precision Measurements | **Electroweak Penguins**

- Flavor-changing neutral currents are not possible at tree level in the **Standard Model (SM)**
- Branching fractions predicted in the range 10^{-7}–10^{-4} with 5–30% uncertainties (dominated by soft QCD effects).
- Highly sensitive to potential **non-SM contributions.**
- Belle II published last year the first evidence of $\boldsymbol{B \to K\nu\bar{\nu}}$



Standard Model Feynman diagram of Penguins with missing energy

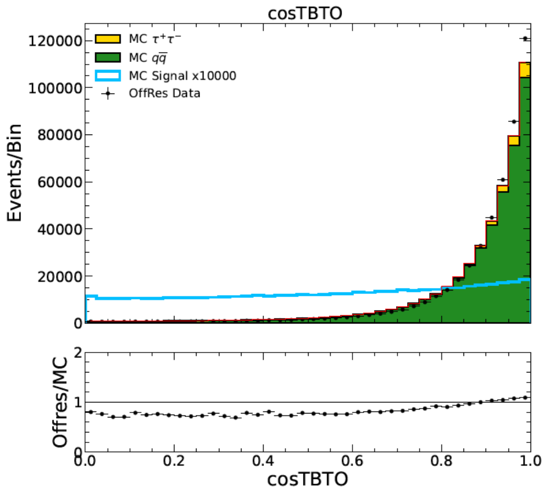
2025, January 16

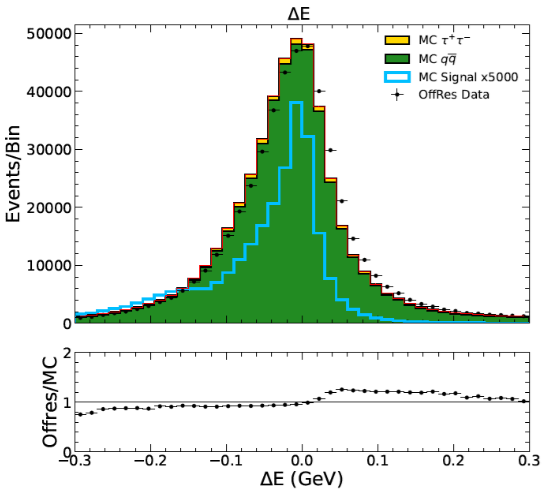
Giovanni Gaudino

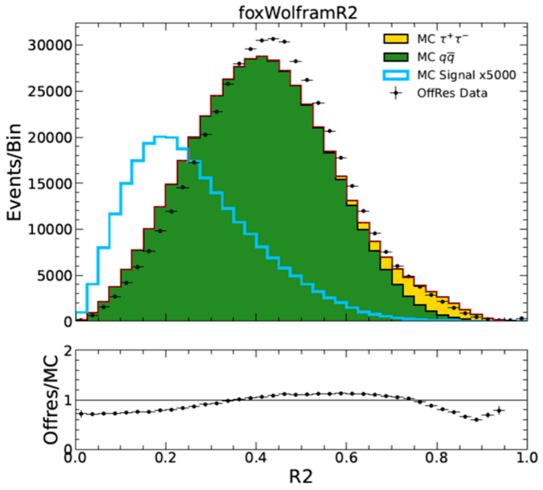
Continuum Preselection

Event shape variables are crucial in discriminating between the continuum and $B\bar{B}$ components. In order to suppress continuum background, we first apply loose cuts (also to use the same FEI Performance cuts)

- $cos\theta_{BTO} < 0.9$;
- $-0.15 < \Delta E < 0.1$;
- $\mathrm{R2} < 0.6$ (99% of taupair component removed)





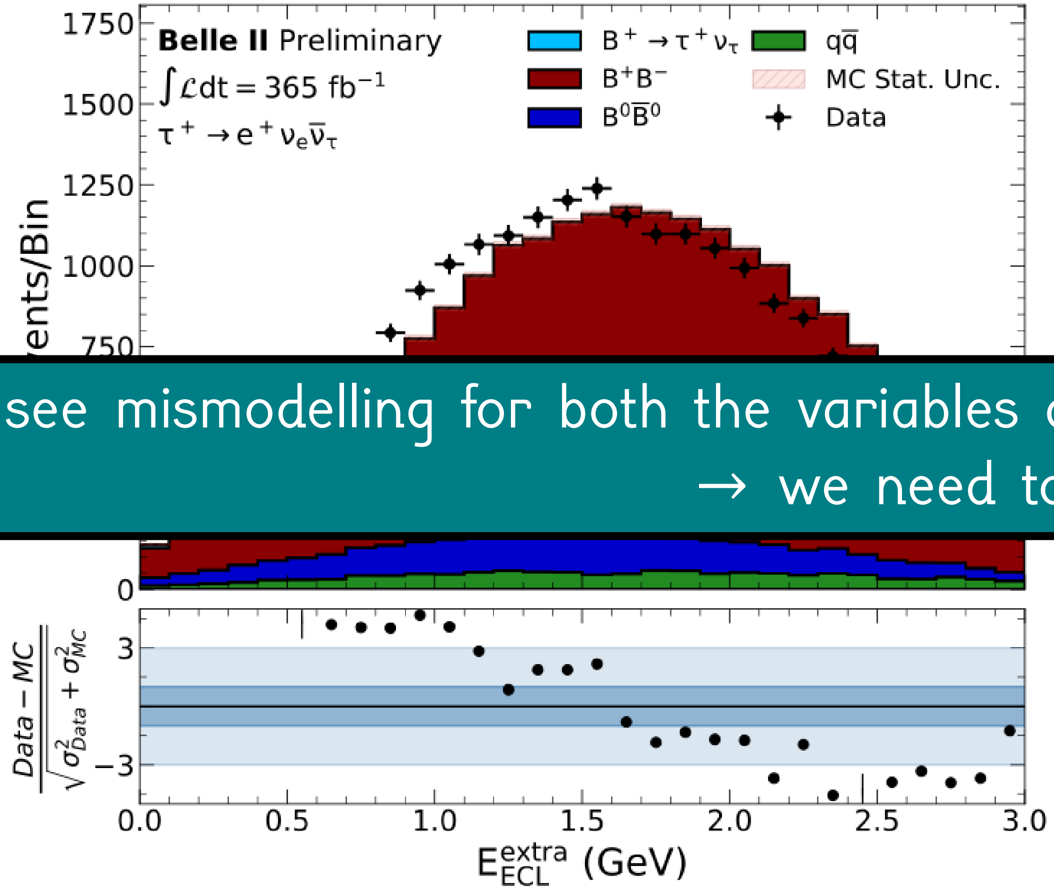


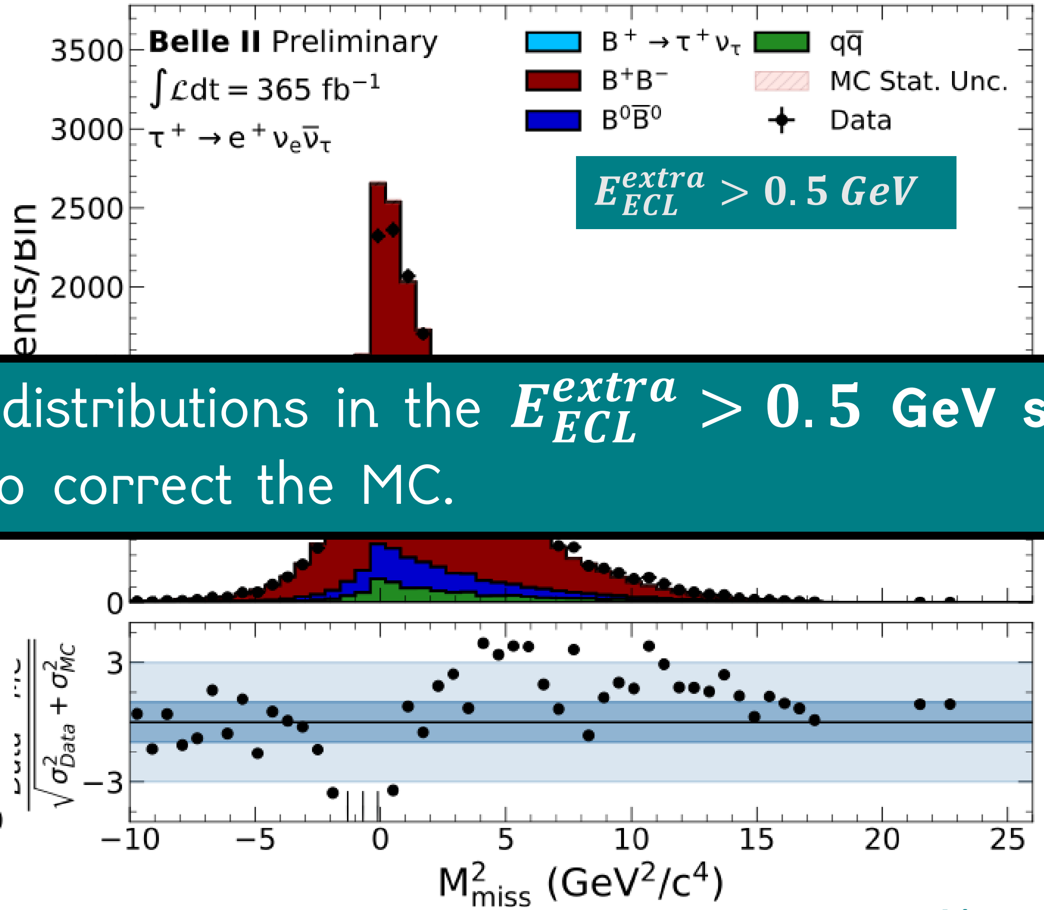
Selections	$B \to \tau\nu$	$B \to \tau\nu(\mathrm{isSig})$	$q\bar{q}$	$\tau\tau$
preselection	100%	51%	100%	100%
$\cos\theta_{\mathrm{TBTO}} < 0.9$	86%	45%	40%	17%
$-0.15 < \Delta E < 0.1\,\mathrm{GeV}$	74%	38%	35%	8%
$\mathrm{R2} < 0.6$	73%	37%	34%	1%

Checks and corrections on Data – MC agreement

Most discriminating variables for signal:

- E_{ECL}^{extra}, the extra energy not associated with the B_{tag} and B_{sig} (Rest of Event).
- $M_{miss}^2 = E_{miss}^2 - p_{miss}^2$, squared magnitude of the four-momentum p_{miss} with the Extra Event definition.





We see mismodelling for both the variables distributions in the $E_{ECL}^{extra} > 0.5$ **GeV sideband** → we need to correct the MC.

Fit Variables Correction

First, **we apply Particle ID, Tag and π^0 reconstruction efficiencies corrections and reweight the Branching Ratio** of simulations to the last PDG averages.

We also use the «**Photon Efficiency Data/MC Ratio» correction** from neutral group study, which gives a weight for each cluster with energy greater than 200 MeV.
The «new» E_{ECL}^{extra} for B^+B^- and $B^0\overline{B^0}$ is computed by summing all the cluster energy contributions, removing randomly some of them after extracting a random number between 0 and 1 (1–Weight = probability to kill a cluster).

Example:

	γ_1	γ_2	γ_3	γ_4
Energy (GeV)	0.25	0.2	0.15	0.4
Weight	0.9	0.9	1	0.85
N_{rand}	0.4	0.95	0.2	0.4
	OK	**Kill**	**OK**	**OK**

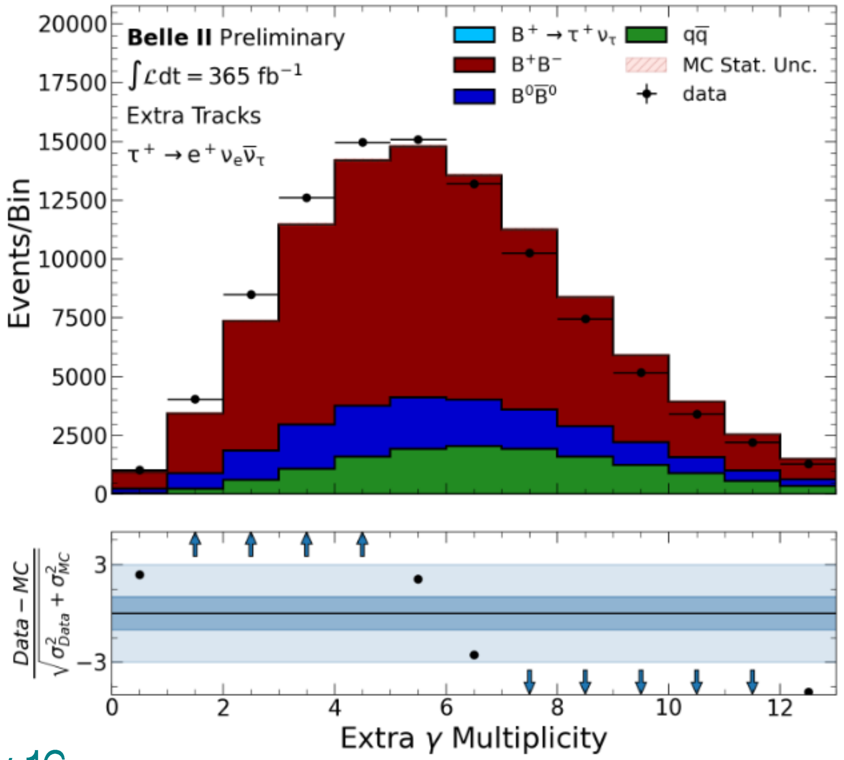
$$E_{ECL}^{extra} = \gamma_1 + \gamma_3 + \gamma_4 = 0.8\ GeV$$
$$n_\gamma^{extra} = 3$$

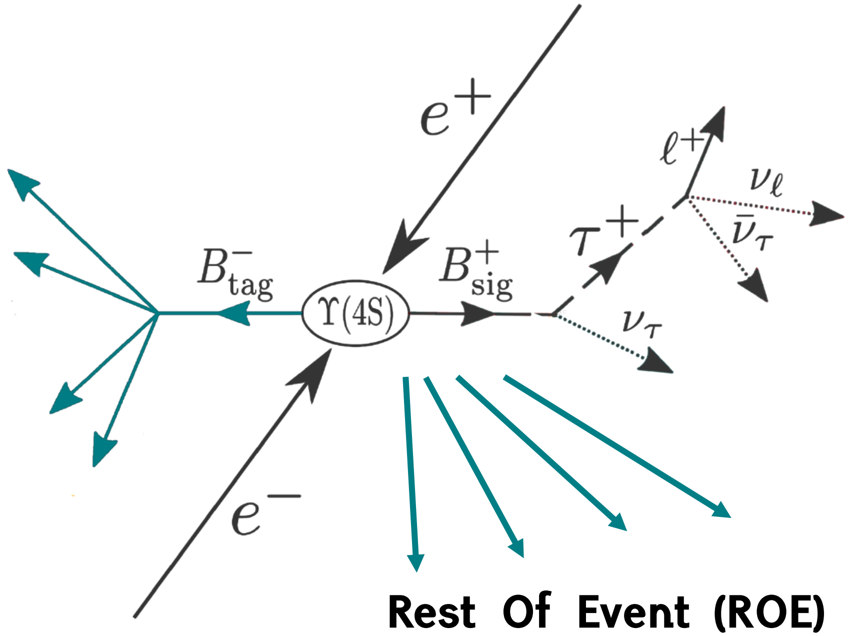
Extra Tracks Control Sample

For the main channel, we require no charged tracks in the ROE.

In this control samples, $N^{ROE}_{Tracks} > 1$ → **same background composition but negligible signal events.**

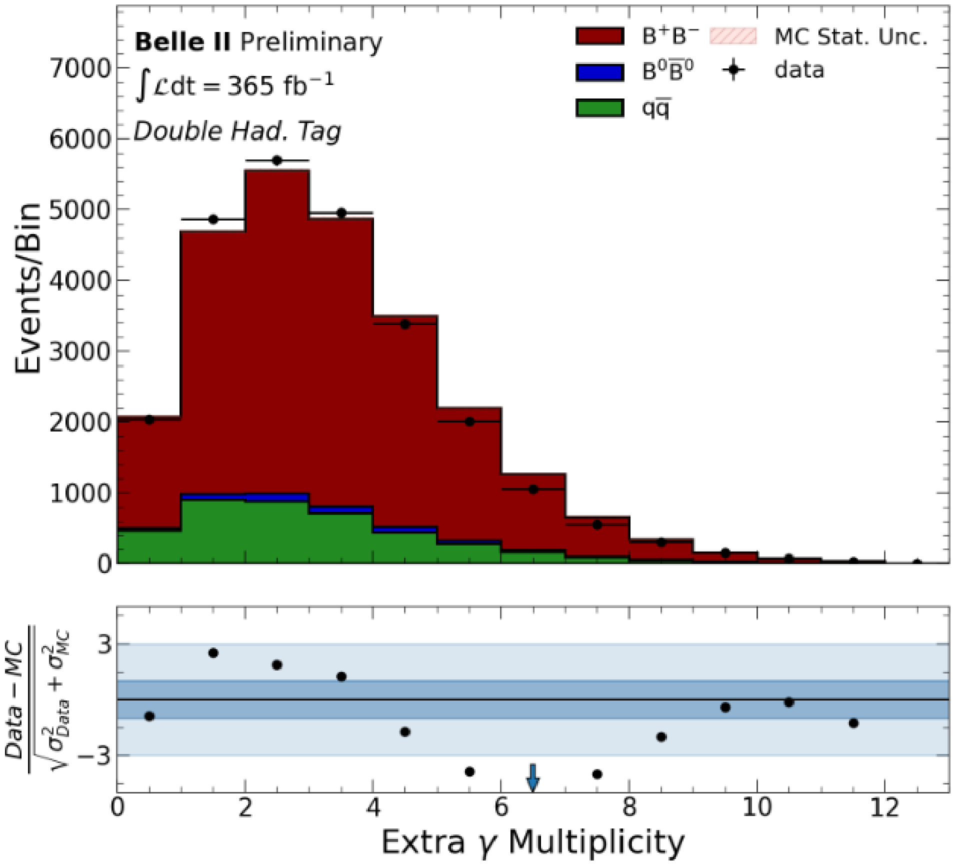
	N^{ROE}_{Tracks}=0	N^{ROE}_{Tracks}=1	N^{ROE}_{Tracks}>1
% Signal	95	4	1

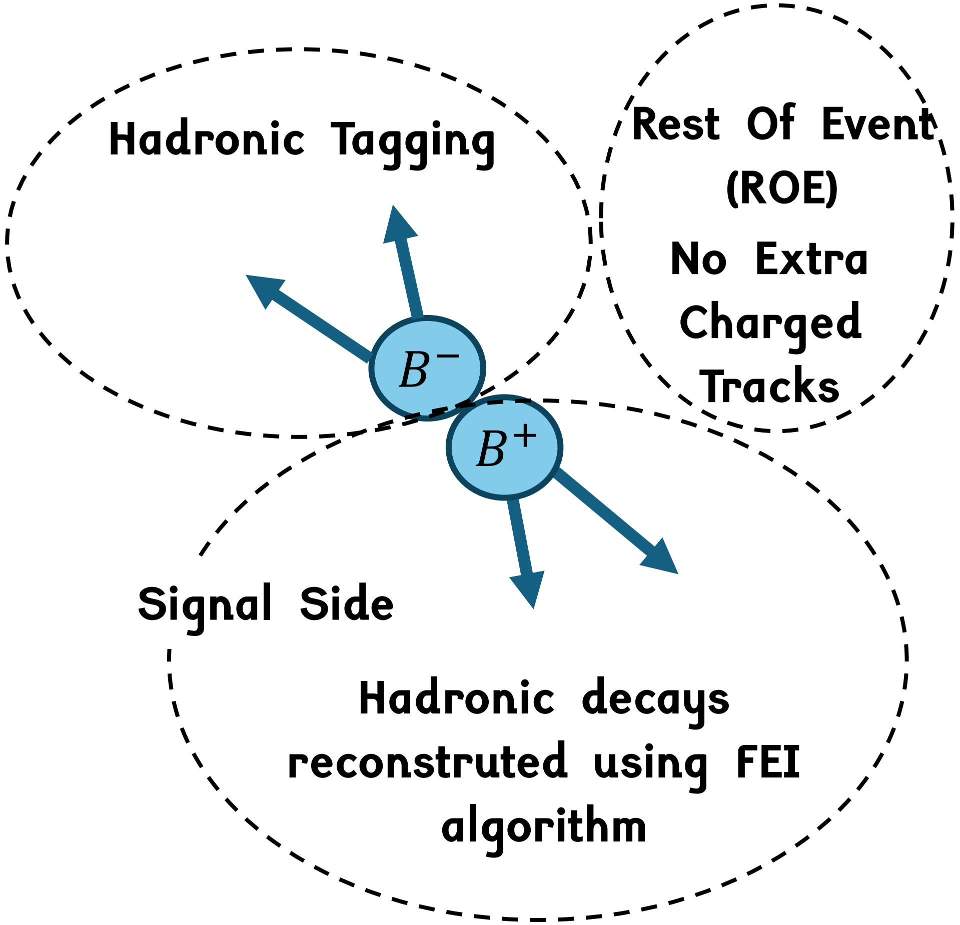




Double Tag Control Sample

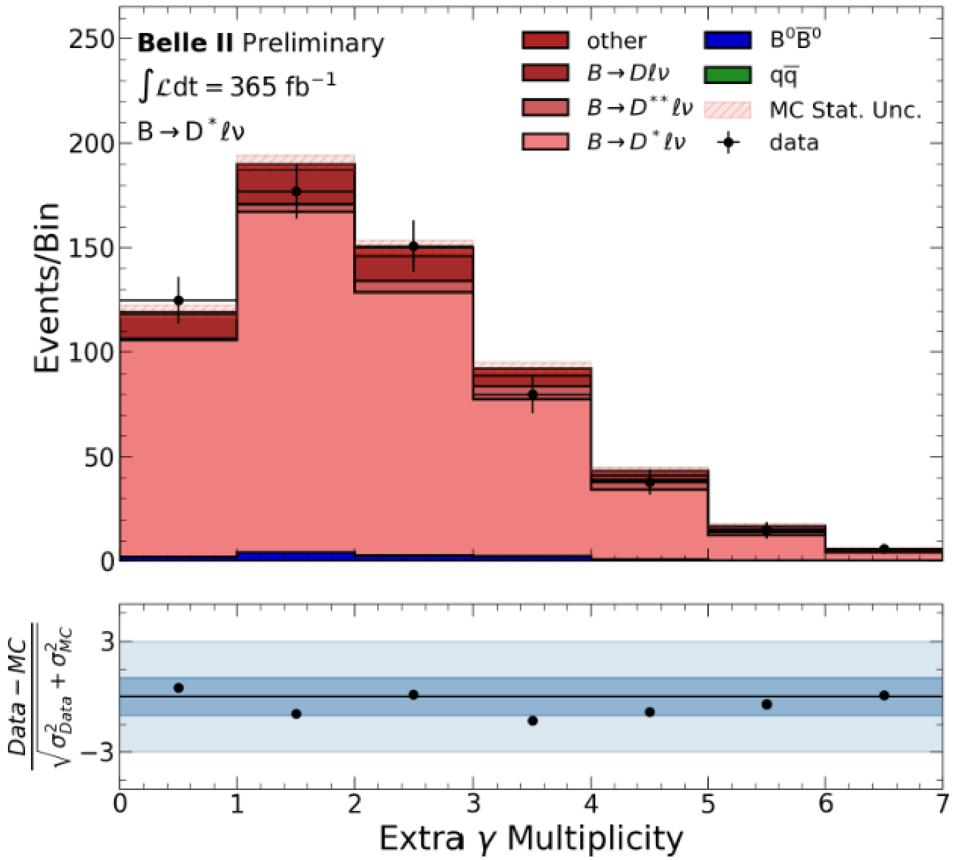
We reconstruct the **two B candidates using the Hadronic Tagging FEI alogritm**.
As for the haronic signal channel, the decay is fully reconstructed.
No signal events.

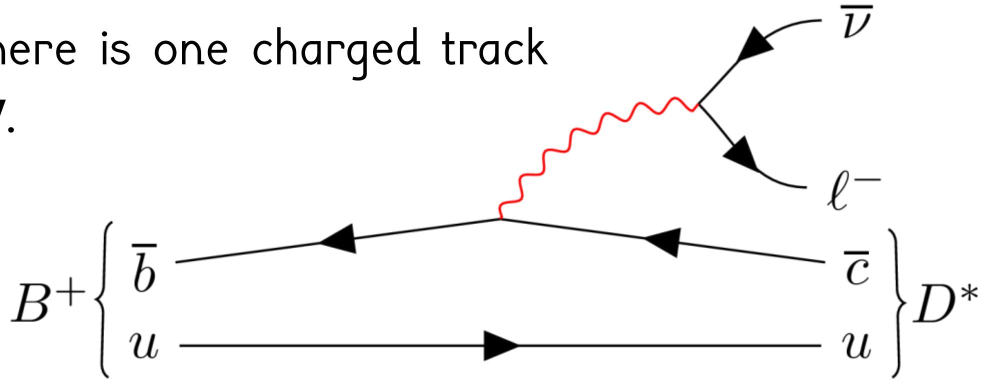




$B \to D^*\ell\nu$ Control Sample

This control sample resembles the leptonic signal since there is one charged track and the D^* is fully reconstructed $\to$ $\boldsymbol{E_{ECL}^{extra}}$ **peaks at 0 GeV**.





D^* decays

- $D^* \to D\gamma$
- $D^* \to D\pi^0$

D decays

- $D \to K\pi$
- $D \to K\pi\pi\pi$
- $D \to K_S\pi\pi$

$\to$ 0 Extra Tracks (from IP)

We also use it to validate the signal efficiency between data and simulation.
We find a Data/MC ratio after all the selection and calibrations equal to 0.96 ± 0.04.

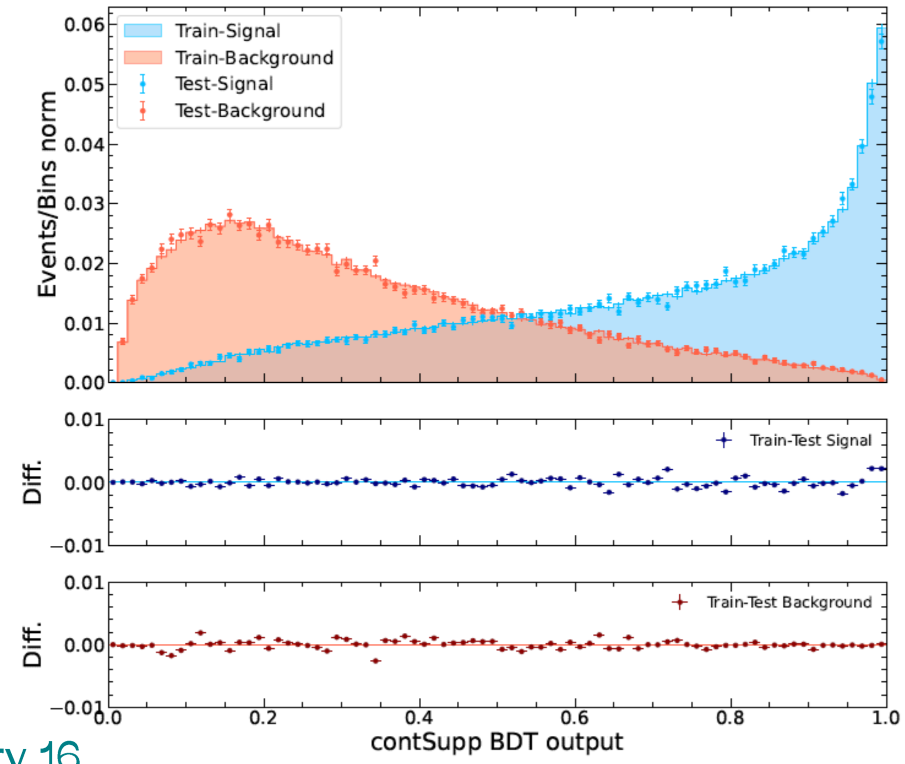
Continuum Suppression

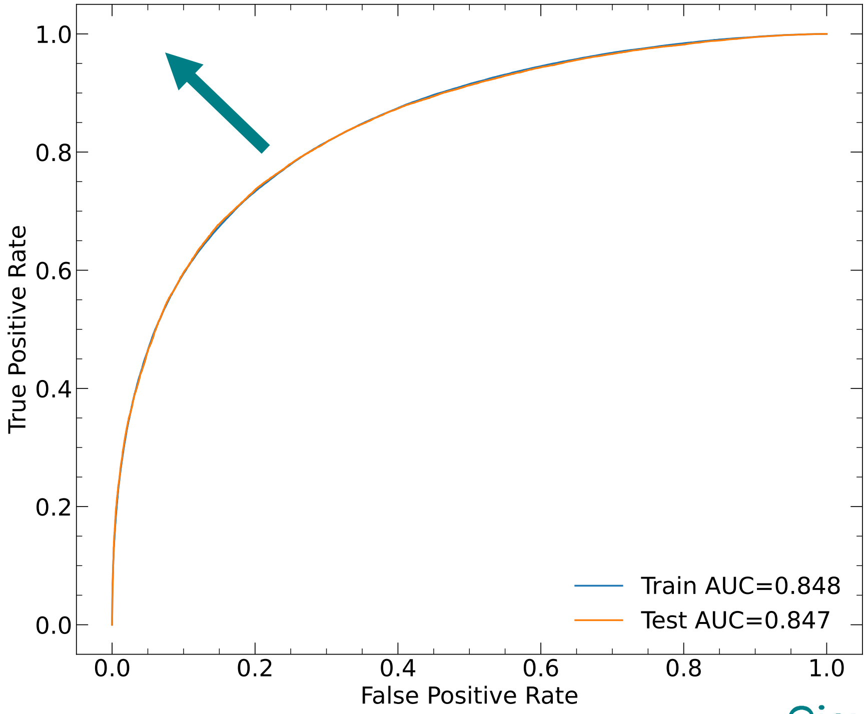
To suppress continuum, we train **2 FastBDT**, one for Leptons and one for Hadrons, using **MC continuum as "Signal"** and **MC $B\overline{B}$ as "Background"**.

- 300K events, Train/Test sample 80%/20%
- Signal/Background events ratio = 1
- Features = only variables with good Data/MC agreement and less correlated with our fit variables.

In the training, the weights from continuum reweighting are used.

Leptons (hadrons in backup)





Selection Optimization

The optimization is done extracting the signal yield with a 2D Fit E_{ECL}^{extra} vs M_{miss}^2 in the signal region $[0,1]GeV \times [-10,\ 26]GeV^2$

The cuts have been optimized:

- minimize a FOM obtained through 5000 ToyMC study on the variables M_{miss}^2 and E_{ECL}^{extra} for each cut combination.

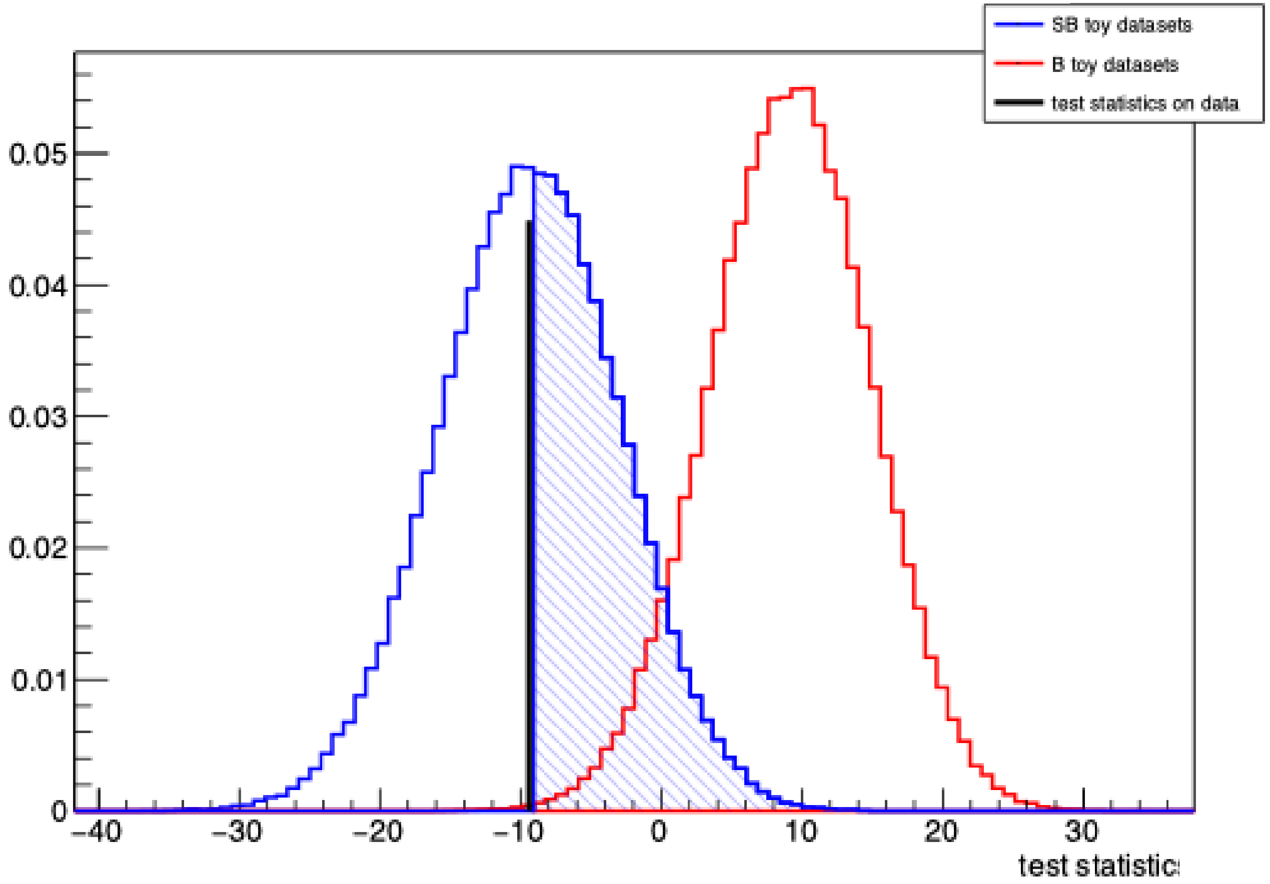
$$\boldsymbol{FOM = \frac{\bar{\sigma}_S}{\bar{N}_S}}$$

($\bar{N}_S$ and $\bar{\sigma}_S$ are the mean signal yield and error of the ToyMC)

$$\left(\epsilon = \frac{n_{sel}}{n_{gen}}\right)$$

	eID	μID	πID	sigProb	M_{bc} (GeV)	p (GeV)	ContSupp	$\epsilon(10^{-4})$
e	>0.9			>0.01	>5.27	>0.5	<0.8	7.3
μ		>0.9		>0.01	>5.27	>0.5	<0.6	7.6
π			>0.6	>0.01	>5.27	>1.4	<0.6	3.4
ρ			>0.6	>0.01	>5.27	>1.65	<0.7	3.1

Significance



Significance from null hypothesis with 1.000.000 toys:

3.15σ only statistical unc.

3.01σ convolving the signal likelihood with a Gaussian whose width is equal to the systematic uncertainty.

Test statistics: $-2log(\mathcal{L}/\mathcal{L}_0)$

Embedding Procedure for Signal efficiency

The other way to check the signal efficiency is the embedding procedure, using the $B^+ \to J/\psi(\to \ell\ell)K^+$ clean sample.
Following the basf2 procedure: software page.
After all the procedure we found an efficiency ratio:

$$\frac{\varepsilon_{data}}{\varepsilon_{MC}} = 1.02 \pm 0.18$$

The uncertanties is still large, but it is a double check for the signal efficiency found in the $B \to D^*\ell\nu$ control sample.