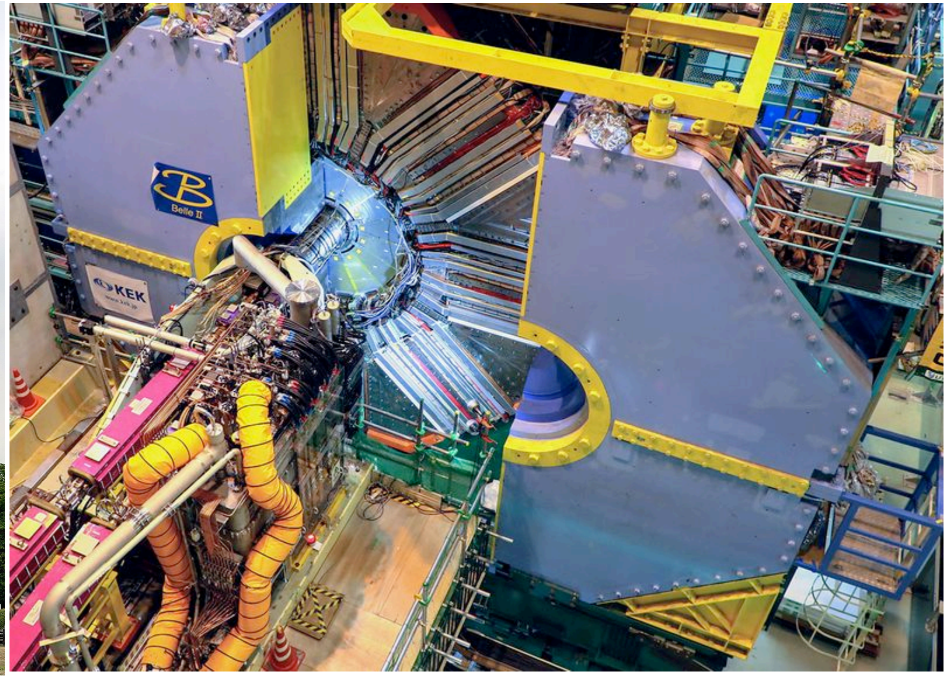


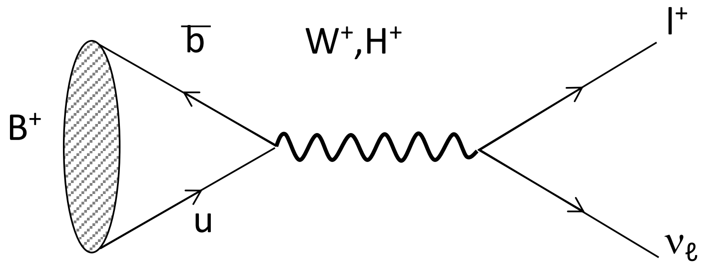
Prospects for $B \rightarrow \tau\nu$ Branching ratio at Belle II

Michele Aversano[1], **Guglielmo De Nardo[2]**, Toru Iijima [134], Giovanni Gaudino[15], Mario Merola[2] and Belle II Collaboration

Nagoya University [1], INFN Napoli – University of Napoli Federico II [2], Nagoya KMI [3], KEK IPNS [4], Scuola Superiore Meridionale [5]

B1 Heavy Flavor and Dark Matter Joint Unit Symposium, 29-30 March 2023, Nagoya (Japan)

B meson purely leptonic decays $B \to l\,\nu$



Very clean theoretically…

…very hard experimentally

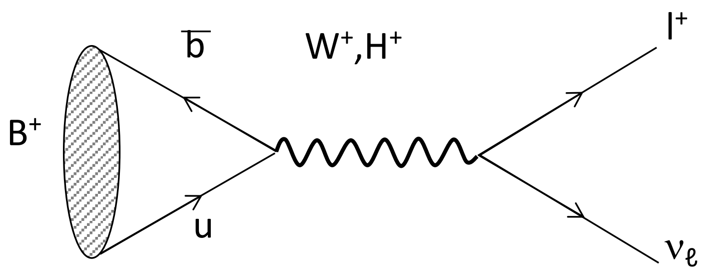
in the SM: $$\mathcal{B}(B \to l\nu) = \frac{G_F^2 m_B}{8\pi} m_l^2 \left(1 - \frac{m_l^2}{m_B^2}\right)^2 f_B^2 |V_{ub}|^2 \tau_B$$

- Helicity suppression by a factor of m_l^2
- being a $b \to u$ transition sensitive to (and suppressed by) $|V_{ub}|$
- Hadronic uncertainty in the decay costant f_B (calculated with lattice QCD)

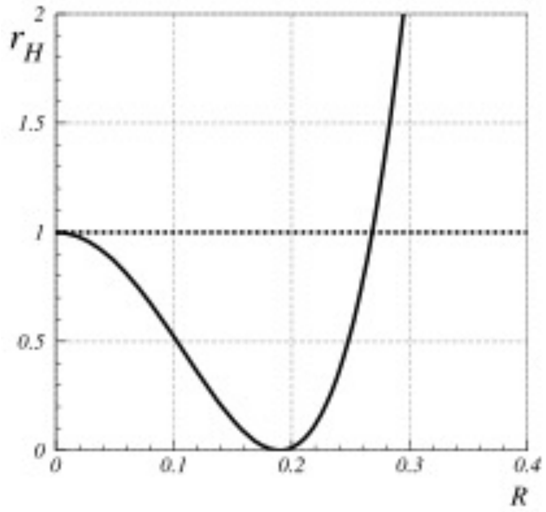
Standard Model Predictions

Mode	$\mathcal{B}(B^+ \to \ell^+\nu_\ell)$	
$\tau\nu_\tau$	$(1.01 \pm 0.29) \times 10^{-4}$	*Accessible with current data sets*
$\mu\nu_\mu$	$\sim 0.45 \times 10^{-6}$	*Need Belle II statistics*
$e\nu_e$	$\sim 0.8 \times 10^{-11}$	*Beyond the reach of experiments*

B meson purely leptonic decays $B \to l\,\nu$



Sensitive to NP contribution (for example: a charged Higgs)



in 2HDM type II

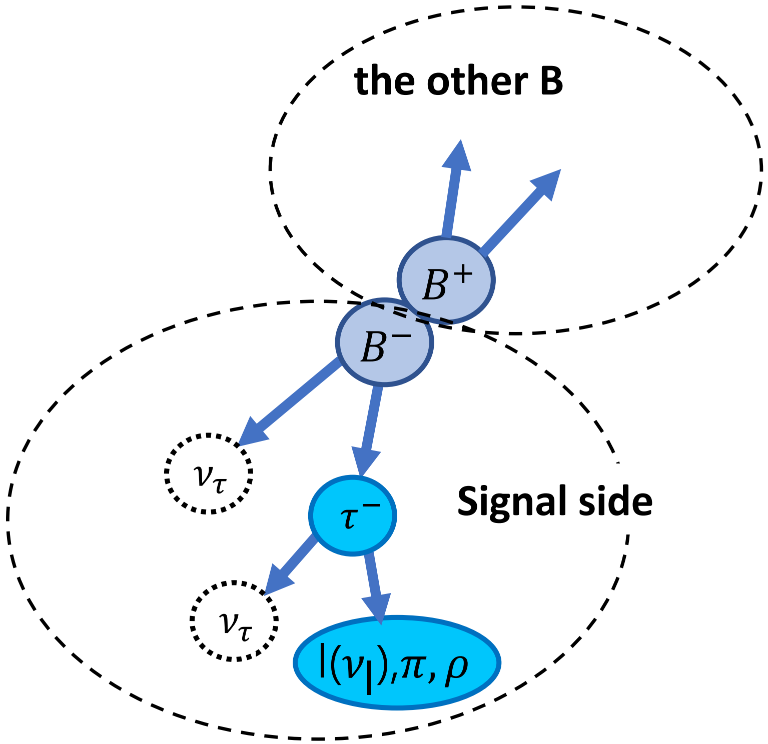
$$\mathcal{B}(B \to l\nu) = \mathcal{B}(B \to l\nu)_{SM} \times r_H$$

$$r_H = (1 - \tan^2\beta\,\frac{m_B^2}{m_H^2})^2$$

Possible test of Lepton Flavor Universality with:

$$R^{\tau\mu} = \frac{\Gamma(B \to \mu\nu)}{\Gamma(B \to \tau\nu)} \qquad R^{\tau e} = \frac{\Gamma(B \to e\nu)}{\Gamma(B \to \tau\nu)}$$

How to search for leptonic decays $B \to \tau\, \nu$ with e$^+$ e$^-$ B factories



$(l = e, \mu)$

Signal is searched through $\boldsymbol{\tau}$ decays (1-prong):

- $\boldsymbol{\tau \to e \nu_e \nu_\tau}$ ~71% of the τ
- $\boldsymbol{\tau \to \mu \nu_\mu \nu_\tau}$ Brancing Fraction
- $\boldsymbol{\tau \to \pi \nu_\tau}$
- $\boldsymbol{\tau \to \rho \nu_\tau}$ with $\boldsymbol{\rho \to \pi^\pm \pi^0}$

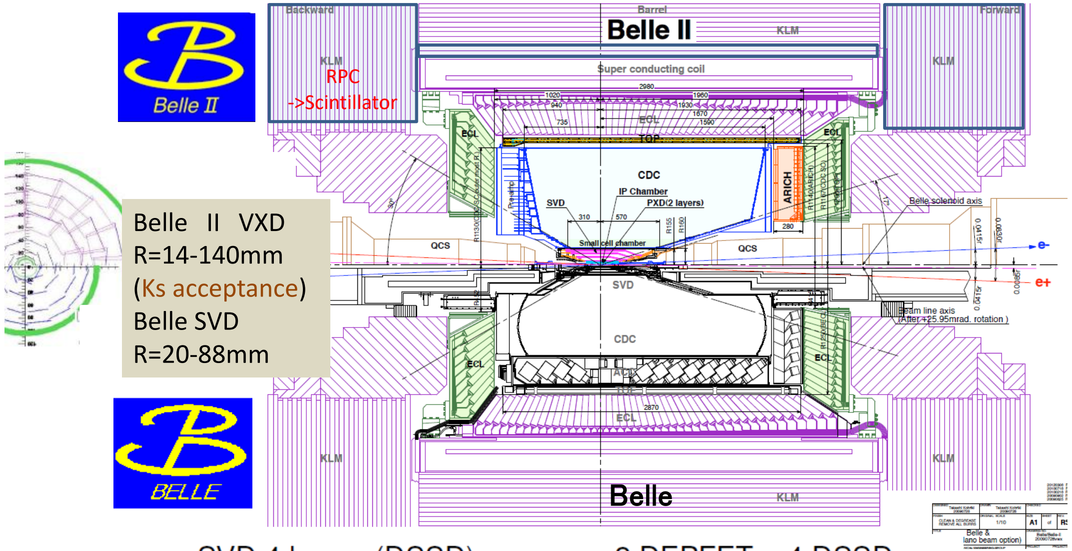
Weak experimental signature: a single charged particle on the signal

Experimental features to exploit:

- Large missing momentum and energy from many (2 or 3) neutrinos
- Particle Identification of the charged particle decay product
- Kinematics constraints from two body decays in sequence for hadronic channels, mass contraint for r channel

…evidence of the companion B meson and nothing else…

Belle II detector and its unique features



Belle II VXD
R=14-140mm
(Ks acceptance)
Belle SVD
R=20-88mm

SVD 4 layers (DSSD)	→	2 DEPFET + 4 DSSD
CDC:		small cell, long lever arm
ACC+TOF	→	TOP+ARICH (Better K/p separation)
ECL:		waveform sampling
KLM: RPC	→	Scintillator+SiPM

Exactly 2 (quantum correlated) B meson produced at Y(4S) and trigger efficiency close to 100%

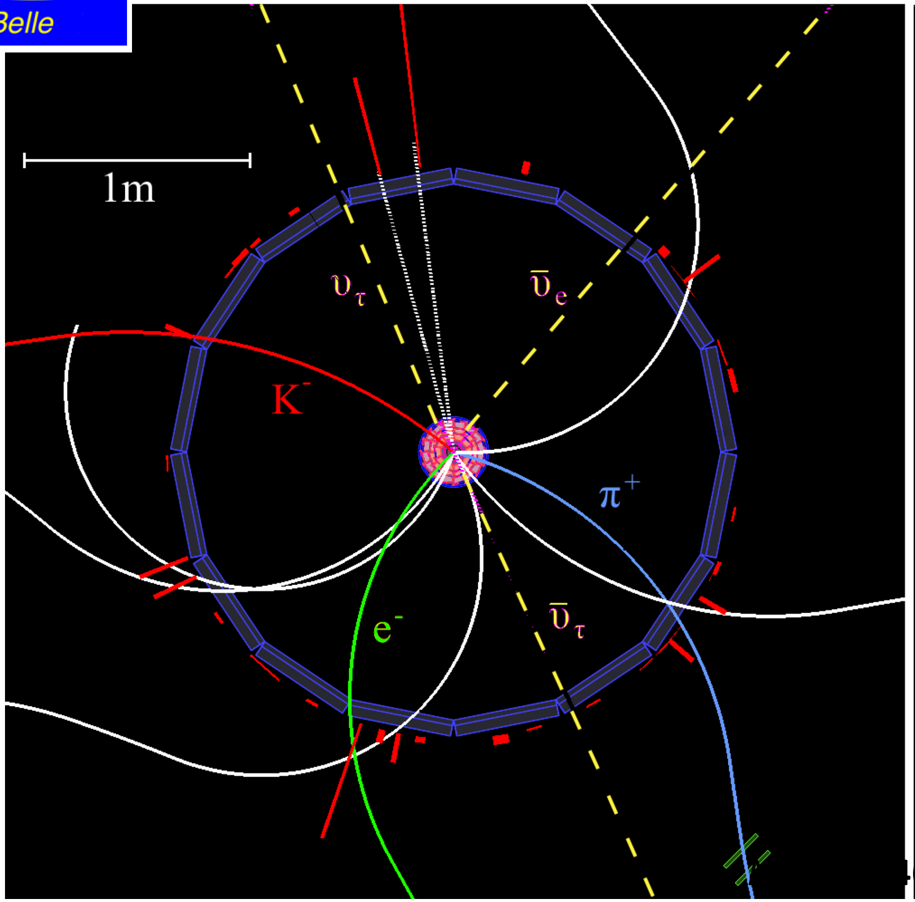
Belle II will accumulate by 2035 (5×10^{10} B pairs)

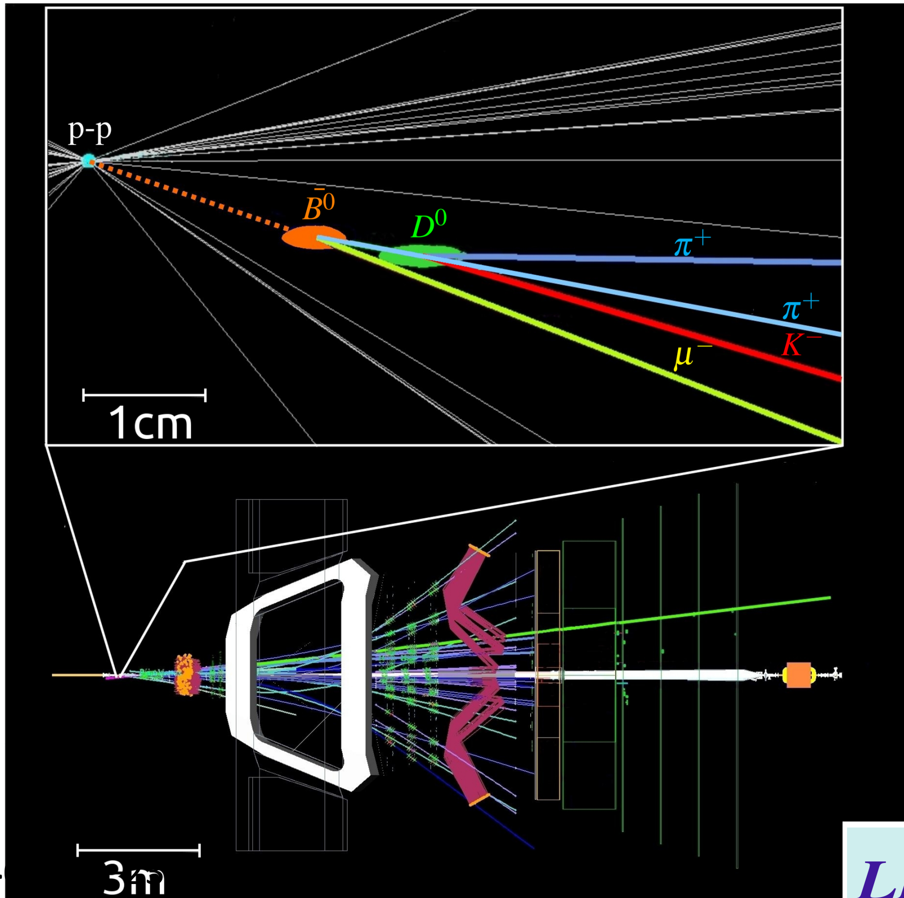
Excellent efficiency and resolution in tracking as well as in detecting photons, π^0, K_L

Electrons and muon performances both excellent

e+ e- environment is "clean" enough→ see next slide

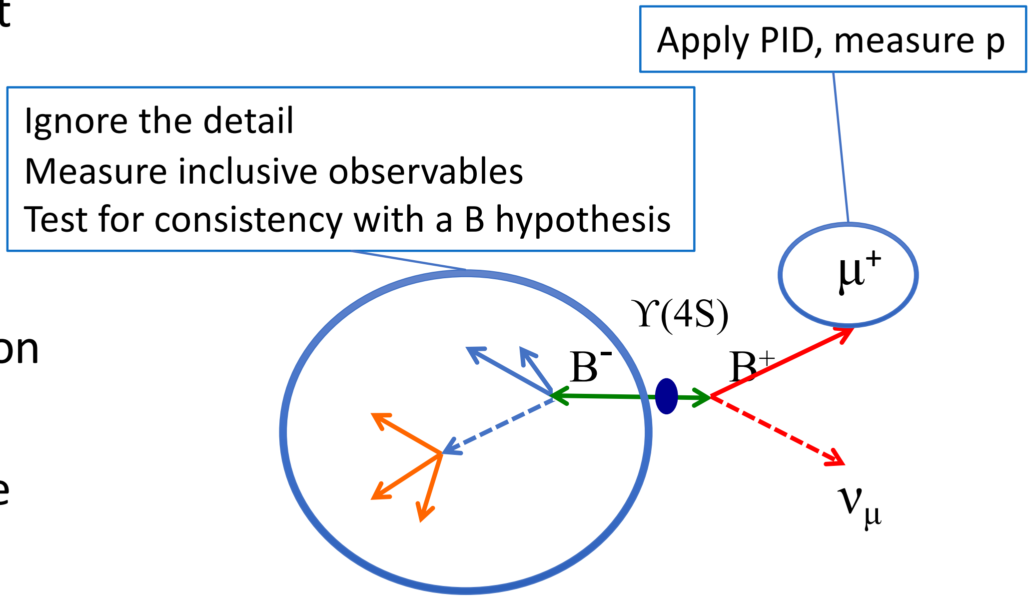
$B \to \tau\nu$ can be measured only in a clean environment as the one in Belle II


1m
υ_τ
$\bar{\upsilon}_e$
K^-
π^+
e^-
$\bar{\upsilon}_\tau$


p-p
$\bar{B}^0$
D^0
π^+
π^+
K^-
μ^-
1cm
3m

Untagged analyses still doable ($B \rightarrow \mu\nu$ and $B \rightarrow e\nu$)

- Inclusive on the rest of the event when the signal signature strong enough
- B → π l ν
  - Loose neutrino reconstruction
- B → μ ν
  - Monochromatic muon in the final state in B rest frame
    - Smeared in the CM frame

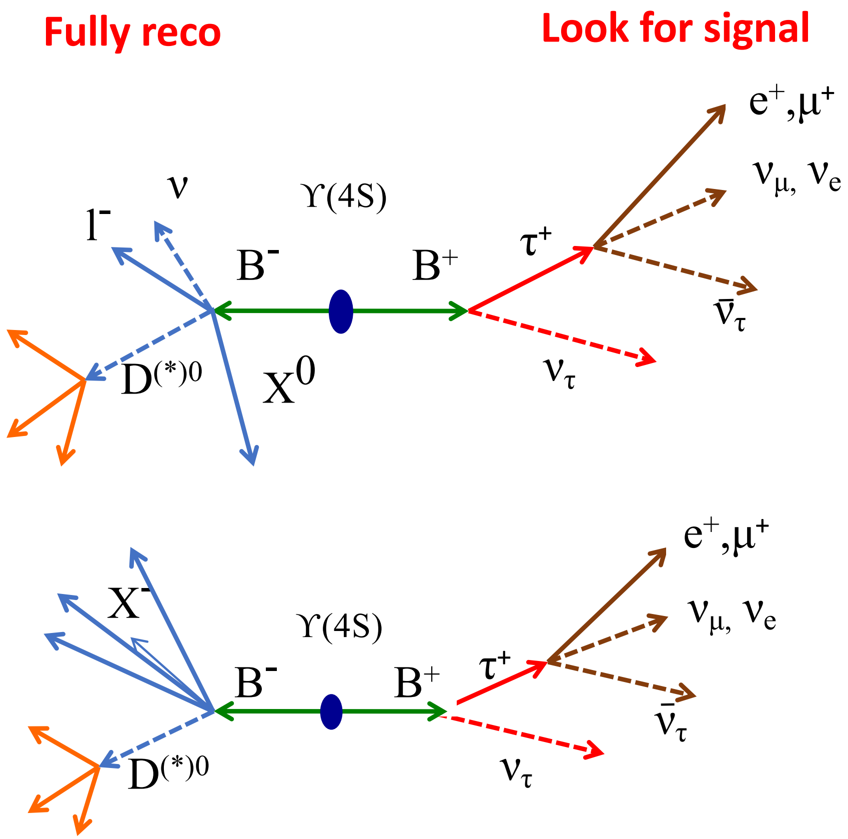


High efficiency and large backgrounds, too

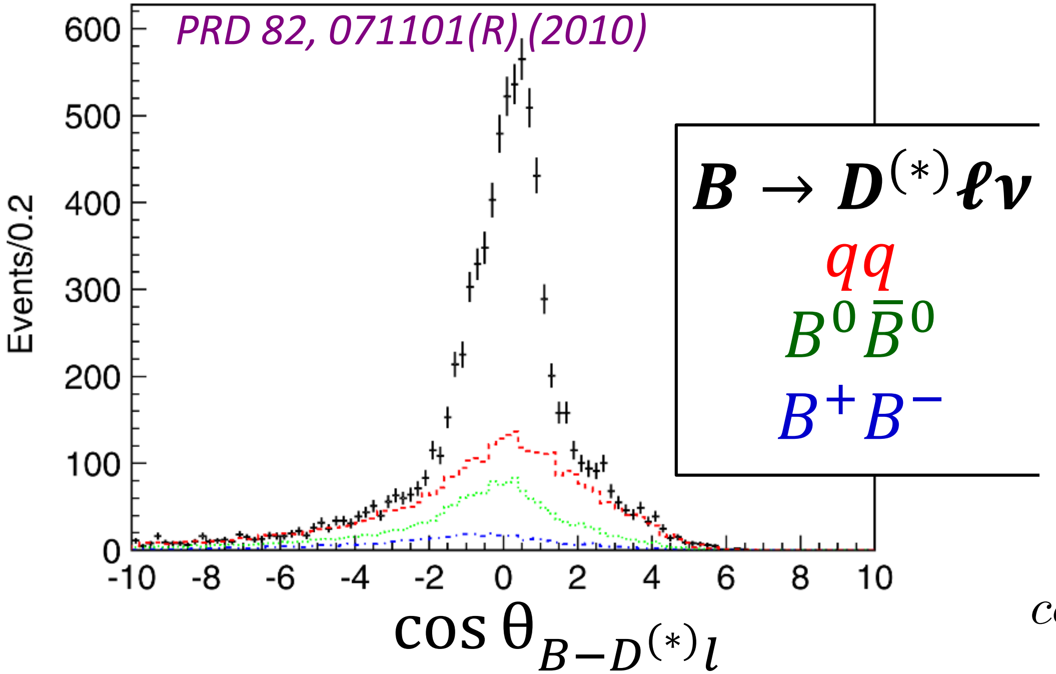
Full event reconstruction (tagged analyses)

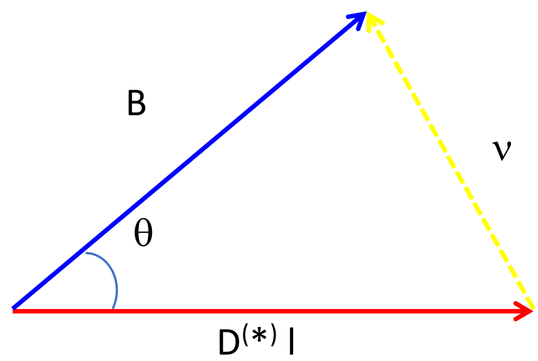
- For signal with weak signature like
  - Decay with missing momentum (many neutrinos in the final state)
  - Inclusive analyses
- background rejection improved fully reconstructing the companion B (tag)

- Tag with semileptonic decays
  - PRO: Higher efficiency $\varepsilon_{tag} \sim 1.5\%$
    CON: more backgrounds, B momentum not measured
- Tag with hadronic decays
  - PRO: much cleaner events,
    B momentum reconstructed
    CON: smaller efficiency $\varepsilon_{tag} \sim 0.2\text{-}0.5\%$



Tag with B semileptonic decays



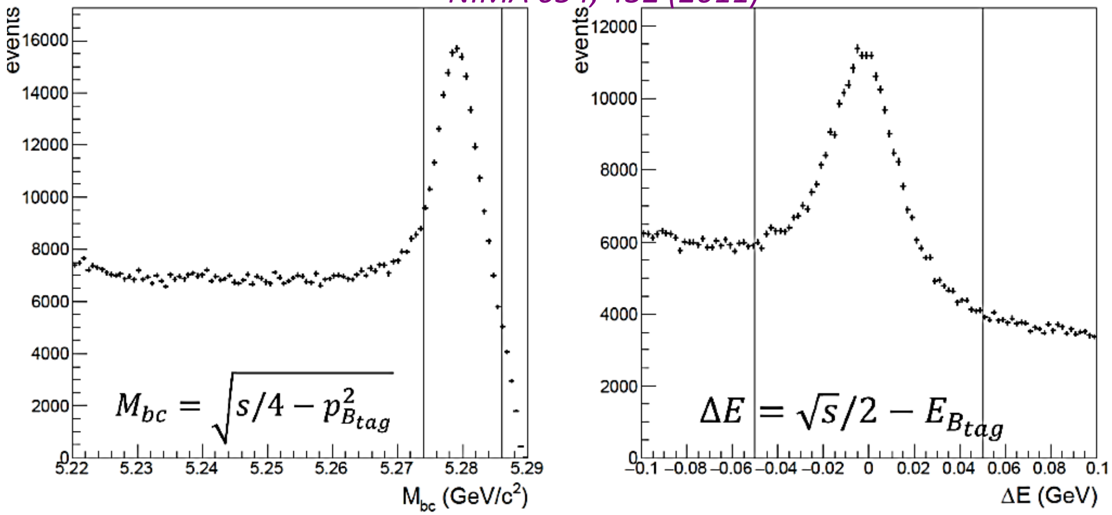


$$\cos\theta_{B-D^{(*)}l} = \frac{2E_{beam}E_{D^{(*)}l} - m_B^2 - m_{D^{(*)}l}^2}{2\, p_B\, p_{D^{(*)}l}}$$

Tag with B hadronic decays

$$\boldsymbol{M_{bc}} = \sqrt{E_{\text{beam}}^2 - \vec{p}_{B_{\text{tag}}}^{\,2}} \qquad \boldsymbol{\Delta E} = E_{\text{beam}} - E_{B_{\text{tag}}}$$

NIMA 654, 432 (2011)



Babar and Belle pioneered a tag reconstruction technique

- reconstructing D/D* mesons in as many as possible decay trees
- Combining the seed D/D* with an hadronic system of charged and neutral pions to make fully reconstructed B candidates.

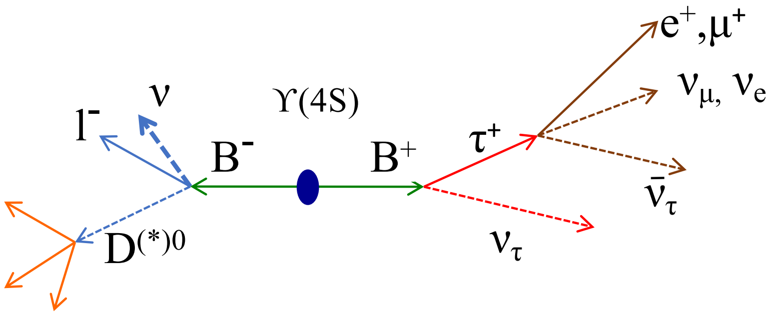
Many combinations per event!

BaBar determined the purity on experimental data to rank the B decay modes

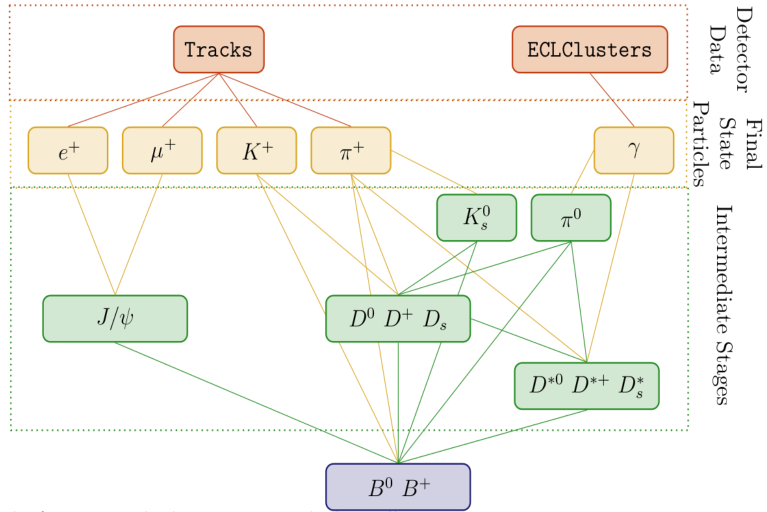
Belle used a NN tool to determine the quality of the tag (output of the classifer)

This has been refined in Belle II

Belle II Full Event Interpretation (FEI)



T. Keck et al., "The Full Event Interpretation", Comput. Softw. Big Sci. 3, 6(2019)



Keck, T., Abudinén, F., Bernlochner, F.U. *et al.* The Full Event Interpretation. *Comput Softw Big Sci* **3**, 6 (2019).

-> multivariate method to separate the two B mesons

-> hierarchical reconstruction of the B and D decay chains (e.g. $B \to D\,n\pi,\ B \to D^*\,n\pi, B \to J/\psi K, \dots$)

- Input variables used to train the multivariate classifiers:
  - PID, tracks momenta, impact parameters (charged FS particles);
  - cluster info, energy and direction (photons);
  - invariant mass, angle between photons, energy and direction (π^0);
  - released energy, invariant mass, daughter momenta and vertex quality ($D^{(*)}_{(s)}$, J/ψ);
  - the same as previous step plus vertex position, ΔE (B);
  - additionally, for each particle the classifier output of the daughters are also used as discriminating variables.

Tag side reconstruction

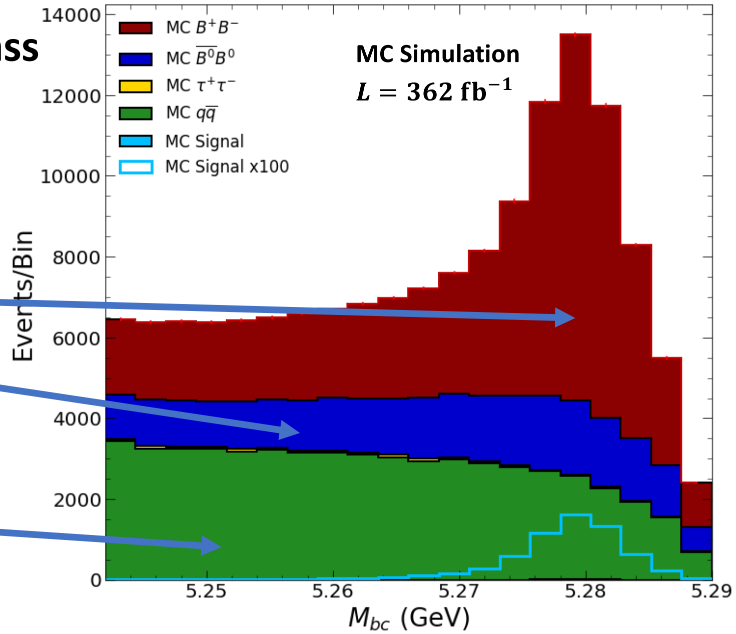
- **Beam energy-constrained mass**

$$M_{bc} = \sqrt{E_{beam}^2 - p_B^2}$$

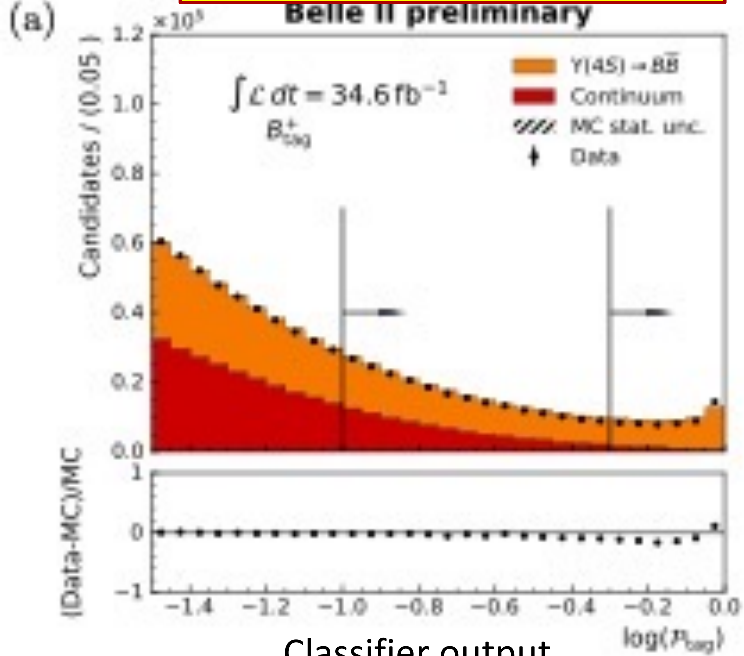
Peaking and Combinatorial B+

Combinatorial B^0

Continuum Background



Belle II Coll., arXiv:2008.06096



Classifier output

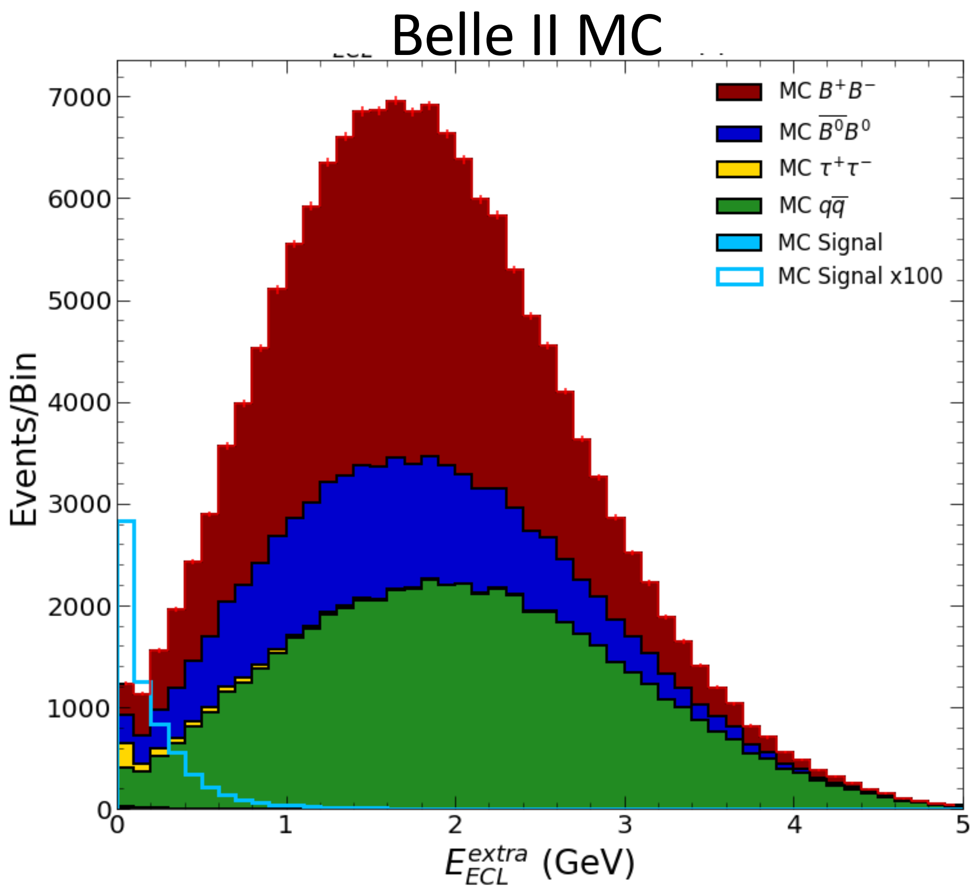
	B^+	B^0
ϵ_{tag}	0.30%	0.23%

Currently we are using this in Belle II for $B \to \tau\nu$

Tag MVA output > 0.01
(tight requirement to select a sample enriched in good tags)

$M_{bc} > 5.27$ GeV

Most discriminating signal (Babar, Belle and Belle II analyses)



No activity expected in the electromagnetic calorimeter is expected*

***after removing the signal τ decays product and and the tag B decay product**

Most discriminating variable for signal:
→ E_{ECL}^{extra}, the extra energy not associated with the B_{tag} and B_{sig} (Rest of Event or ROE).

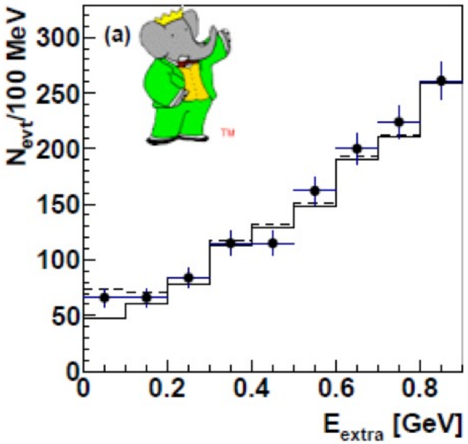
Signal would peak at low E_{ECL}^{extra}, background smooth increasing function of E_{ECL}^{extra}.

potential peaking background at low energy must be suppressed / correctly estimated.

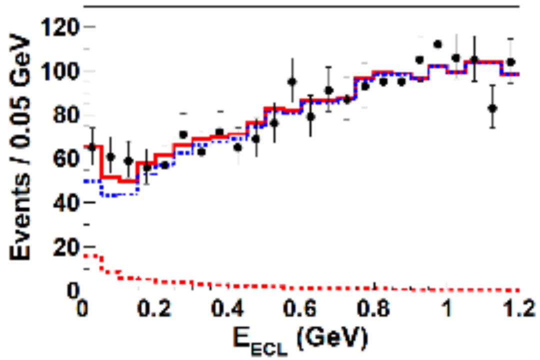
E_{ECL}^{extra} between 0 and 1 GeV is used to extract the BR($B \to \tau\nu$) in our current Belle II analysis

BaBar and Belle past measurements

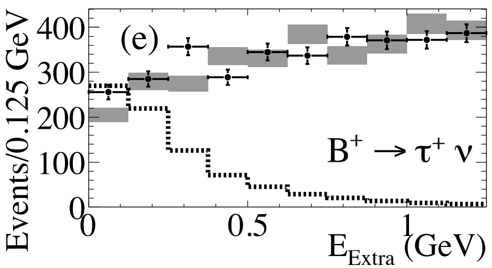
BaBar Hadronic tag
PRD 88 031102



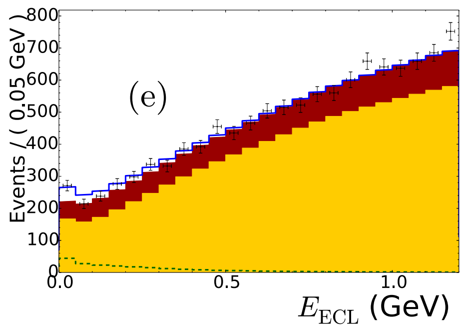
Belle Hadronic tag
PRL 110 131801



BaBar SL tag
PRD 81 051101



Belle SL tag
PRD 92 051102



BaBar and Belle measurements used very same approach for both hadronic and semileptonic tags

Analysis	Measurement (10^{-4})
BABAR Had	1.8 ± 0.5 ± 0.2
BABAR SL	1.7 ± 0.8 ± 0.2
Belle Had	0.7 ± 0.3 ± 0.1
Belle SL	1.25 ± 0.28 ± 0.27

Belle results more SM-like,
Babar a bit in excess.

Actually both consistent with each other because of the large uncertaintites

Signal Events Selection

- Particle identification criteria and π^0 reconstruction define four different signal cathegories ($e, \mu, \pi, \rho = \pi\pi^0$)
- Exploit signal kinematics with requirements on missing momentum, charged particle momentum, missing mass
- Continuum suppression exploiting event topology

All the cuts have been optimized:

minimize a figure of merit (FOM) obtained through $\boldsymbol{Extended\ Maximum\ Likelihood\ fits}$ on the variable $\boldsymbol{E_{extra}}$ through a study with 10,000 pseudo-experiments (ToyMC study).

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

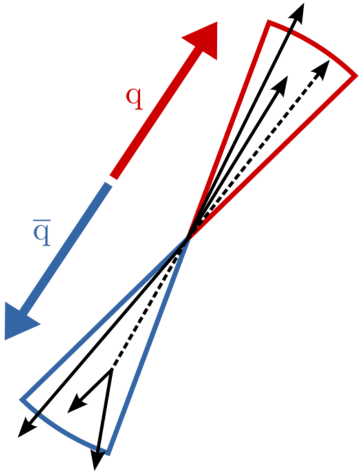
where $\bar{N}_S$ and $\bar{\sigma}_S$ are the mean signal yield and error of the ToyMC.

Continuum Suppression

Exploit the different topology between **event shapes of continuum and $B\overline{B}$**, i.e. the momentum-weighted distribution of all particles in the detector

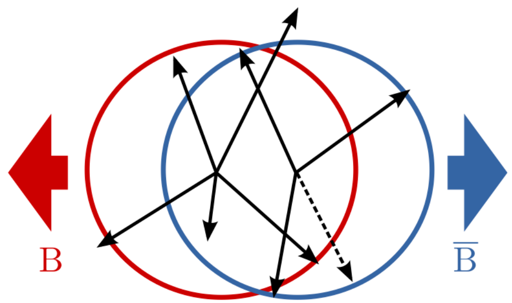
cosTBTO: cosine of angle between thrust axis of B_{sig} and thrust axis of ROE.

Continuum

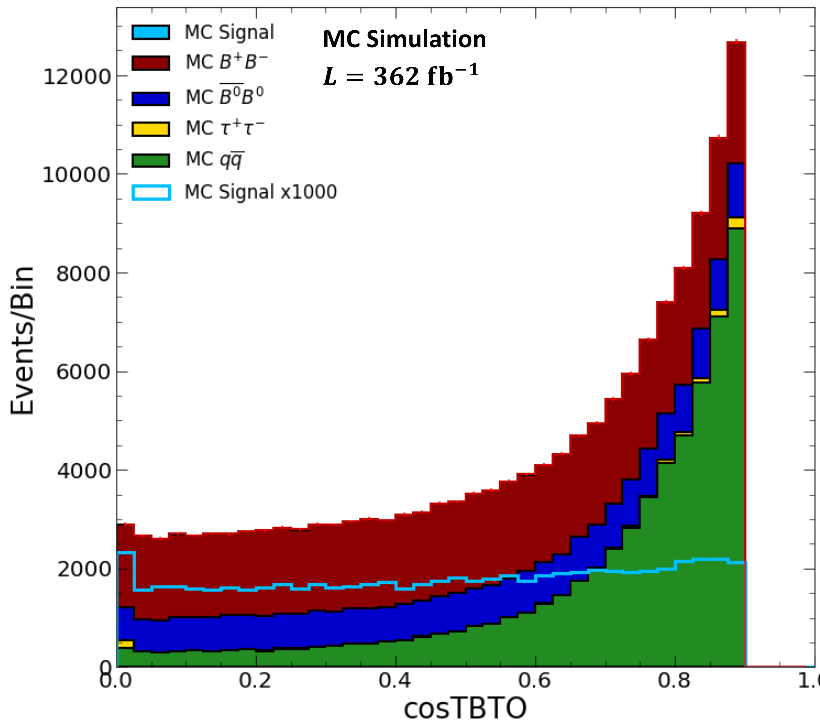


$e^+e^- \rightarrow q\overline{q} \quad (q \in \{u, d, s, c\})$

$B\overline{B}$



$e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\overline{B}$



(pre-cut at 0.9 applied)

Several observables describing the event shape are combined in a BDT

Many observables available

- Input Variables: R2, Cosθ_{th} , Cleo Cones and Kakuno Super Fox-Wolfram (KSFW) moments: 30 variables
  - R2: $R_2 = H_2/H_0$ $\quad H_l = \sum_{i,j} \frac{|\vec{p}_i||\vec{p}_j|}{W^2} P_l(\cos\vartheta_{ij})$
    are the Fox-Wolfram m
  - Cosθ_{th} : $\left|\cos(\vartheta_{thrust})\right| = \frac{|\vec{p}_B \cdot \hat{T}|}{|\vec{p}_B|}$ where *T* is the thrust axis of the rest of the event
  - Cleo Cones: momentum flow around the B thrust axis in 9 angular bins
  - KSFW: $KSFW = \sum_{l=0}^{4} R_l^{so} + \sum_{l=0}^{4} R_l^{oo} + \gamma \sum_{n=1}^{N_t} |(P_t)_n|$

so: particles from b-tag and ROE are considered
oo: particles from ROE only are considered

scalar sum of the transverse momentum of each particle

c: charged,
n: neutral,
m: missing

$$R_l^{so} = \frac{\alpha_{cl} H_{cl}^{so} + \alpha_{nl} H_{nl}^{so} + \alpha_{ml} H_{ml}^{so}}{E_{\text{beam}}^* - \Delta E}$$

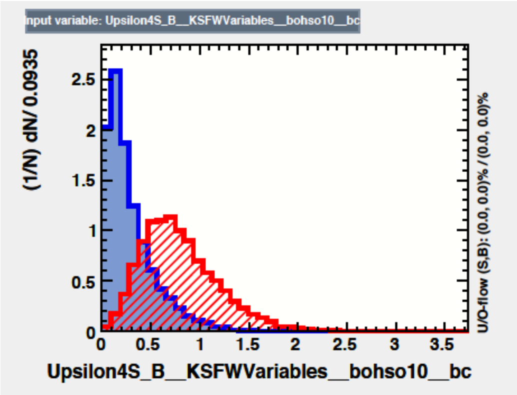
l odd $\quad H_{cl}^{so} = \sum_i \sum_{jx} Q_i Q_{jx} |p_{jx}| P_l(\cos\theta_{i,jx})$

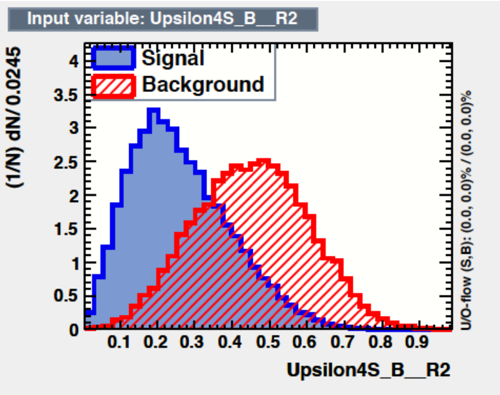
l even $\quad H_{xl}^{so} = \sum_i \sum_{jx} |p_{jx}| P_l(\cos\theta_{i,jx})$

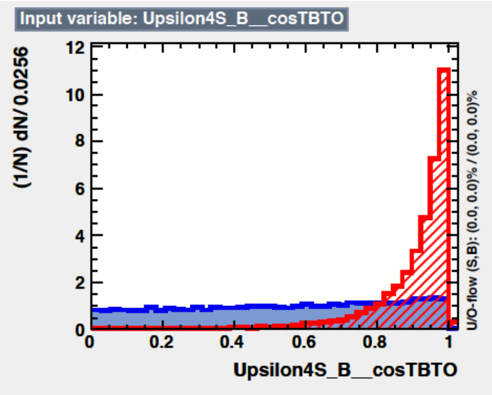
l odd $\quad R_l^{oo} = \sum_j \sum_k \beta_l Q_j Q_k |p_j||p_k| P_l(\cos\theta_{j,k})$

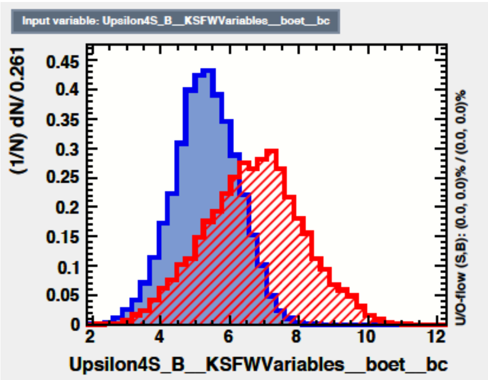
l even $\quad R_l^{oo} = \sum_j \sum_k \beta_l |p_j||p_k| P_l(\cos\theta_{j,k})$

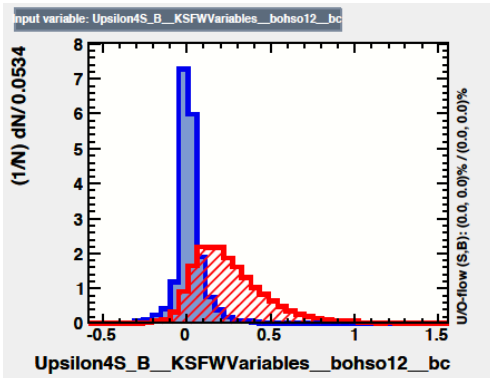
Few examples showing separation









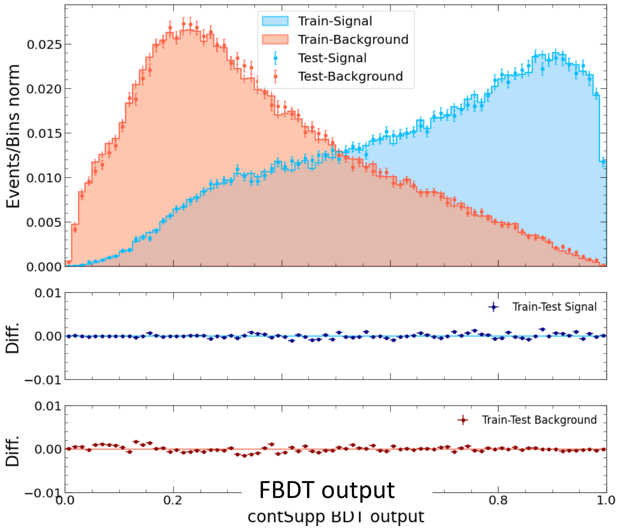


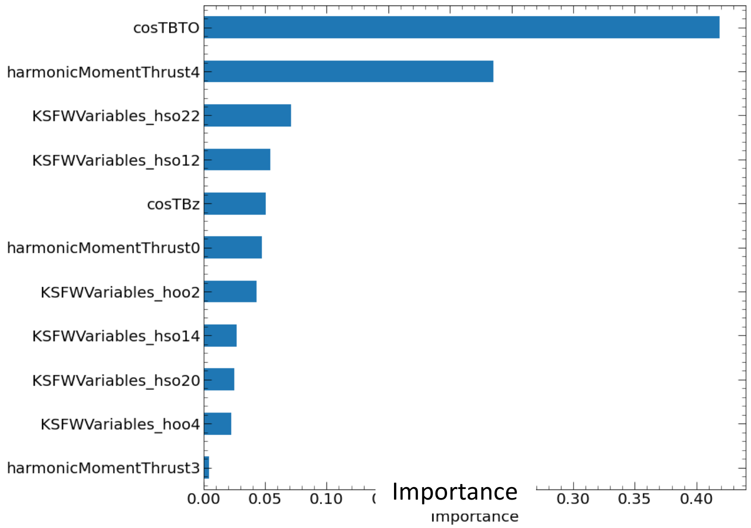
Signal: $B \to \tau\nu$

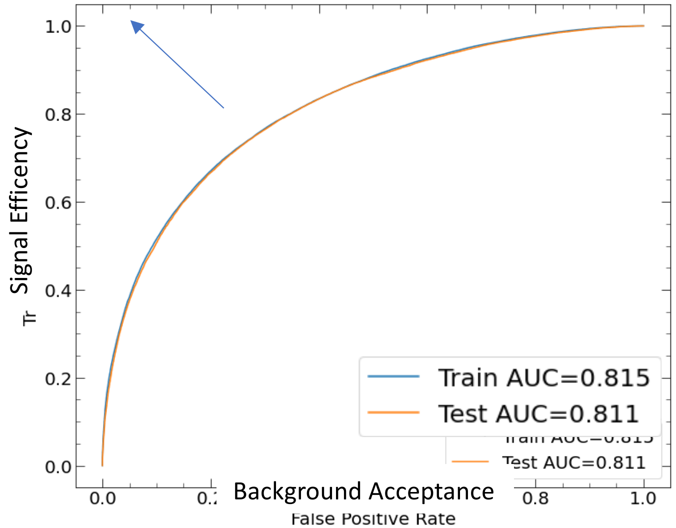
Background: $q\bar{q}$

Continuum Suppression FBDT

We train a multivariate classifier, a Fast Boosted Decision Tree (FBDT), with ones with the highest separation power







Samples

- Sig $\rightarrow$ MC $q\bar{q} + \tau^{+}\tau^{-}$(Continuum)
- Bkg $\rightarrow$ MC $B\bar{B}$

- Train/Test $\rightarrow$ 80%/20%
- Sig/Bkg events ratio = 1

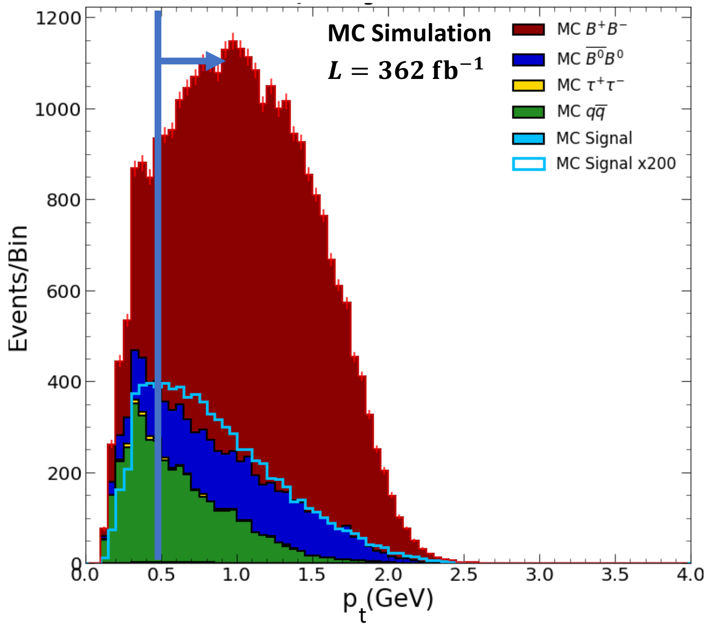
No overtaining and good performance.

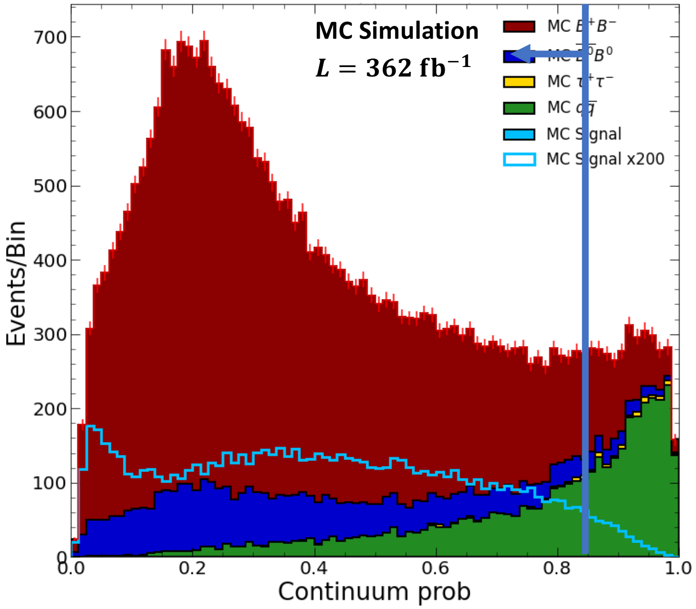
Signal Events Selection optimization by grid search

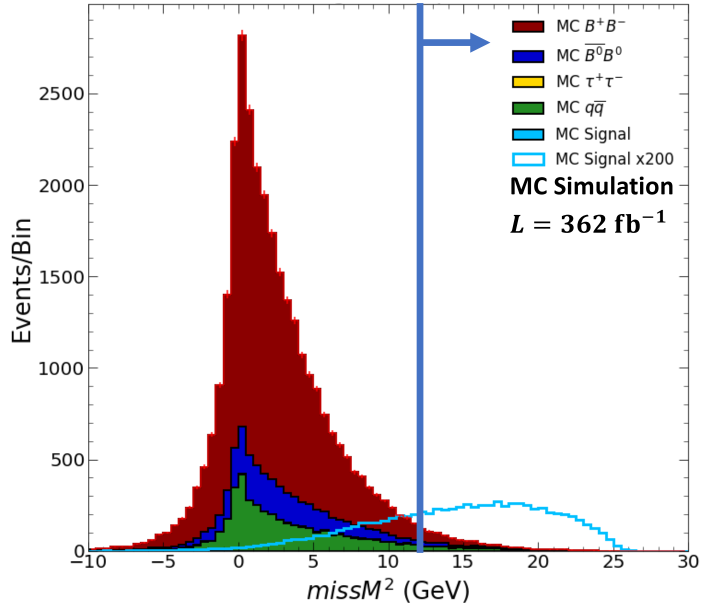
Best Cuts:

	e ID	μ ID	sigProb(FEI)	M_{bc} (GeV)	pt candidate (GeV)	ContSupp	MissM^2 (GeV²)	FOM
e	>0.9		>0.01	>5.27	>0.5	<0.85	>12	0.59
μ		>0.9	>0.01	>5.27	>0.5	<0.7	>11	0.74

Electron (similar for muon)







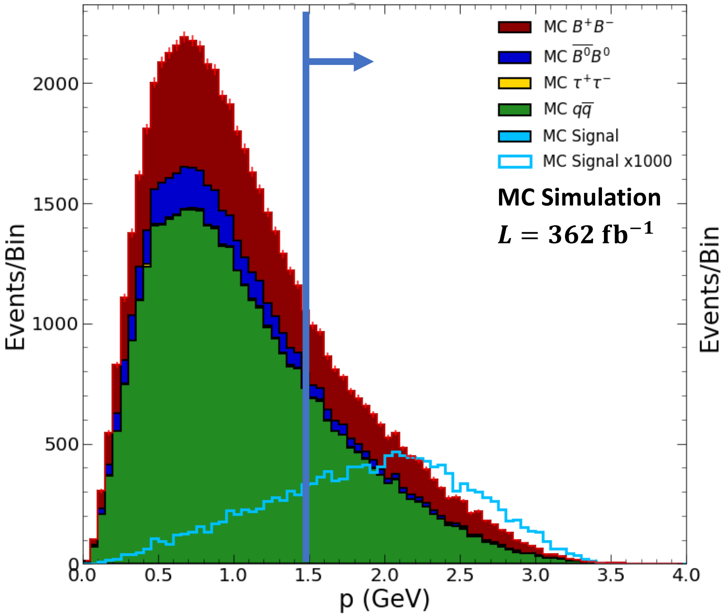
$missM^2$: missing Mass² associated to ROE.

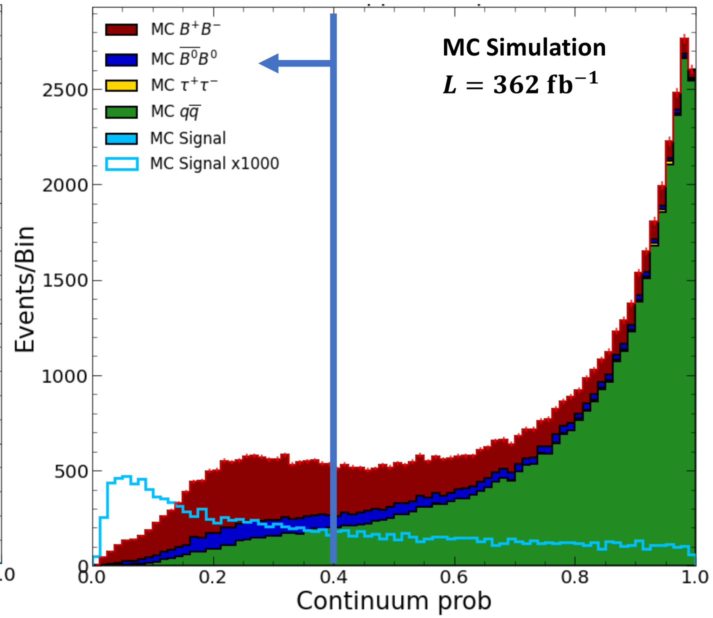
Signal Events Selection

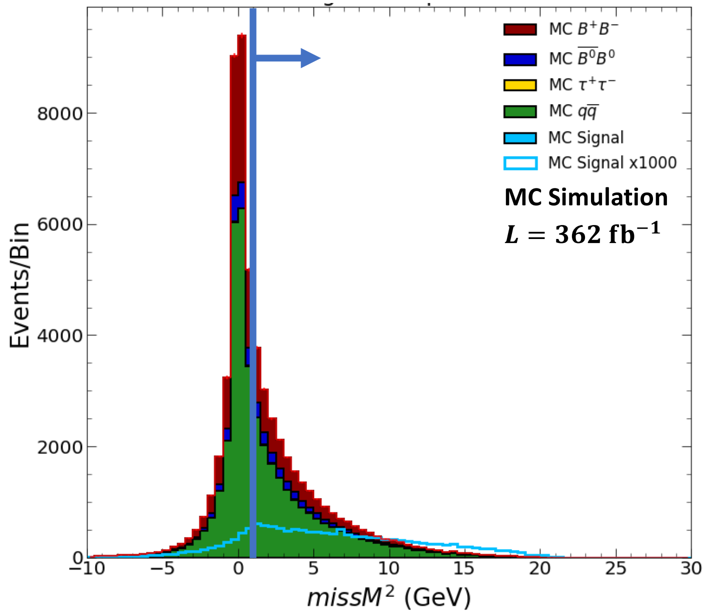
Best Cuts:

	e ID	μ ID	sigProb(FEI)	M_{bc} (GeV)	p candidate (GeV)	ContSupp	MissM^2 (GeV²)	FOM
π	<0.9	<0.9	>0.01	>5.27	>1.5	<0.4	>1	1.11
ρ	<0.9	<0.9	>0.01	>5.27	>1.5	<0.4	>1	1.45

Rho (similar for pion)







$BR(B \to \tau\nu)$ Extraction

The Branching ratio BR is estimated by means of a maximum likelihood fit on E_{extra} simultaneously on the four τ decay modes (the BR being a common parameter)

The Likelihood for each k-mode: (k = e, μ, π, ρ)

$$\boldsymbol{L_k = \frac{e^{-(n_{s,k}+n_{b,k})}}{(n_{s,k}+n_{b,k})!} \prod_{i=1}^{n_{s,k}+n_{b,k}} \left\{ n_{s,k} \cdot P_k^s\left(E_{extra}^{i,k}\right) + n_{b,k} \cdot P_k^b\left(E_{extra}^{i,k}\right) \right\}}$$

$n_{s,k}$ e $n_{b,k}$ sig and bkg yields.

Where:

$$\boldsymbol{n_{s,k} = N^{MEASURED}(\tau \to k - mode) = N_{BB} \cdot \epsilon_k \cdot BR(B \to \tau\nu)}$$

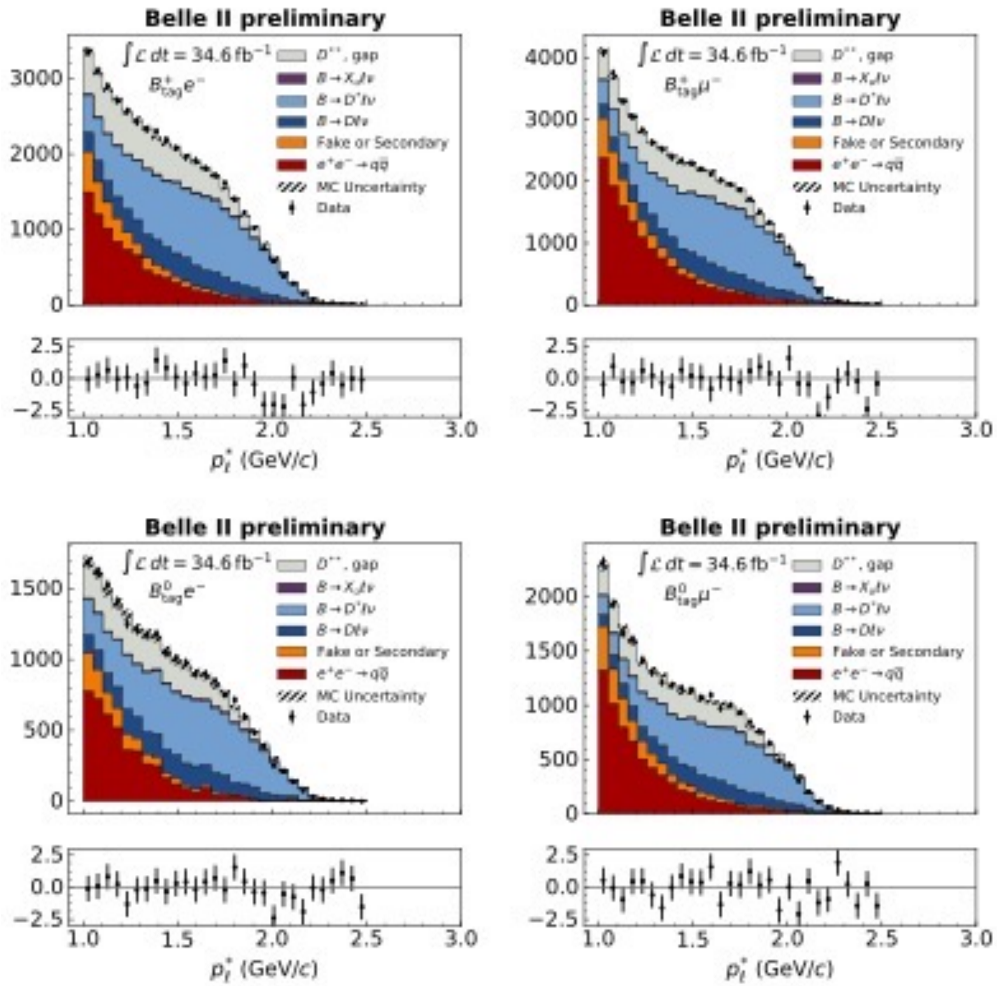
- PDFs of signal and background are taken from the MC simulation.
- Largest source of systematics are MC mismodelling of signal efficiency (including the tag B reconstruction) and PDF shapes.
- Data control sample are used to study this effects and extract correction factor with systematics uncertainties

Tag side MC / data corrections

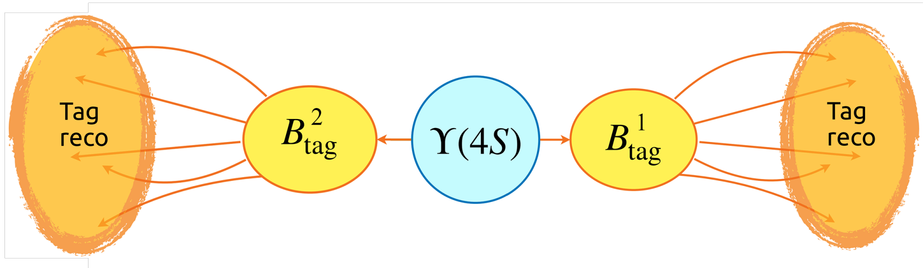
Belle II Coll., arXiv:2008.06096

Control sample of inclusive semileptonic decays

Mostly used for extract tag B reconstruction efficiency from data



Double tags



Two B_{tag} (opposite charge) - Reconstruction of $\Upsilon(4S) \rightarrow B^{+}B^{-}$.

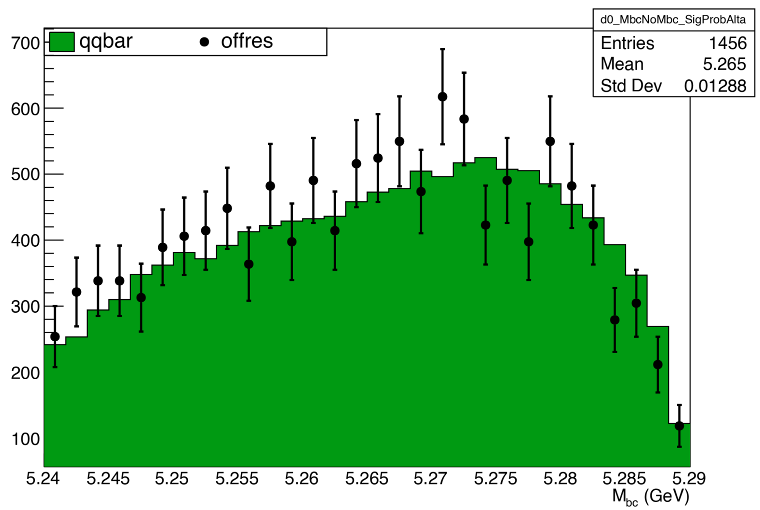
Loose cuts:

- $M_{bc} > 5.24\ GeV$
- $|\Delta E| < 0.3\ GeV$
- TagProb > 0.001
- cos θ_{thrust} < 0.9

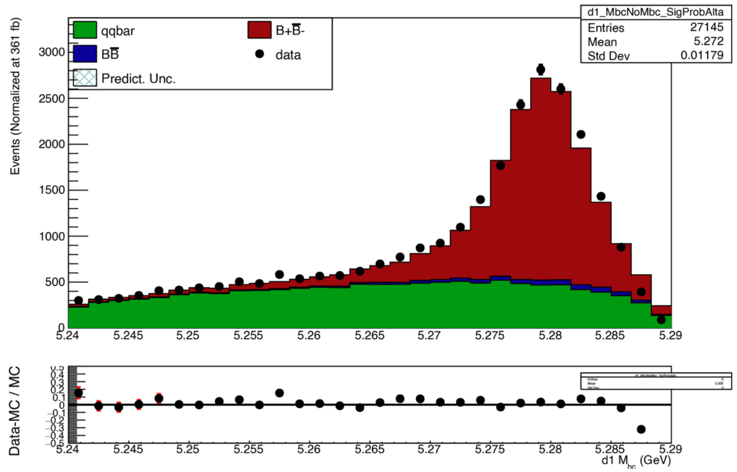
Best Candidate selection with respect to the Tag Probability of the first B_{tag}

0 Extra Tracks in the rest of event.

Off-resonance data – no B expected



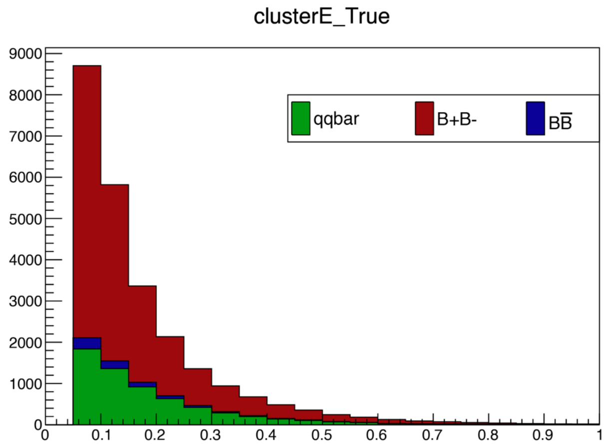
On-resonance data – fully reconstructed $\Upsilon(4S) \rightarrow B^{+}B^{-}$



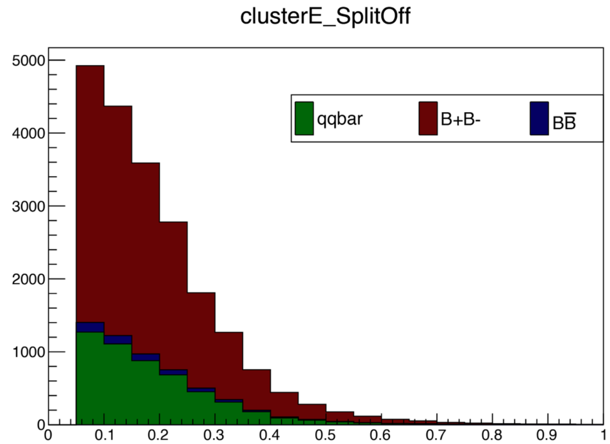
Extra cluster characterization from double tag study

Energy of un-assigned clusters in the electromagnetic calorimeter from MC simulation augmented with beam backgrounds from experimental data

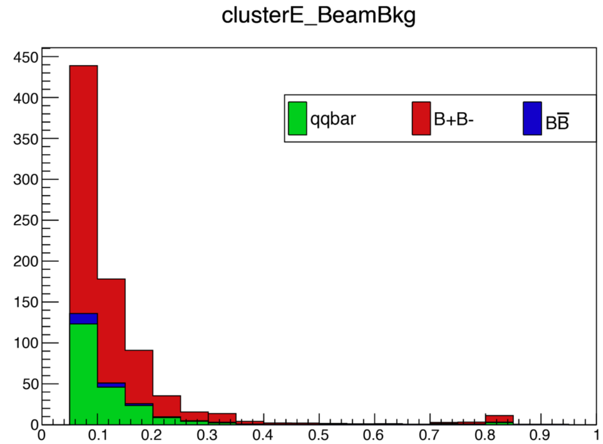
Phyiscs photons



Cluster split off from charged tracks (reconstruction defects)



clusters from beam background



$BR(B \to \tau\nu)$ extraction with ToyMC

The sensitivity is estimated by producing 10,000 pseudo-datasets by a $Simultaneous\ fit$ on E_{extra} between 0 and 1 GeV.

- $\boldsymbol{BR}$ set to the PDG value. $\boldsymbol{BR_{PDG}(B \to \tau\nu) = (1.09 \pm 0.24)\times 10^{-4}}$
- PDFs from the MC.

The Likelihood for each k-mode: (k = e, μ, π, ρ)

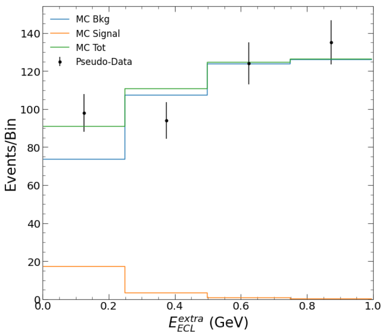
$$\boldsymbol{L_k = \frac{e^{-(n_{s,k}+n_{b,k})}}{(n_{s,k}+n_{b,k})!} \prod_{i=1}^{n_{s,k}+n_{b,k}} \left\{ n_{s,k} \cdot P_k^s\left(E_{extra}^{i,k}\right) + n_{b,k} \cdot P_k^b\left(E_{extra}^{i,k}\right) \right\}}$$

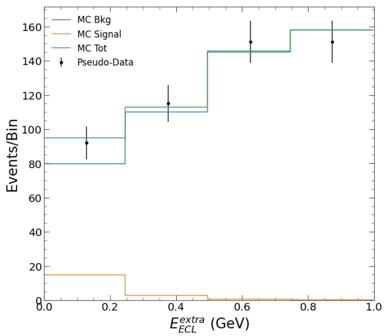
$n_{s,k}$ e $n_{b,k}$ sig and bkg yields.

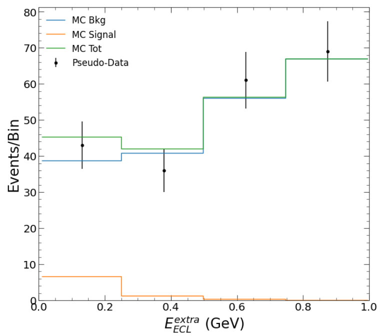
Where:

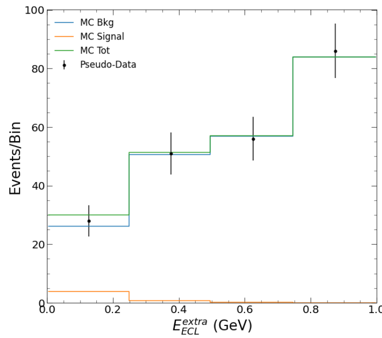
$$\boldsymbol{n_{s,k} = N^{MEASURED}(\tau \to k - mode) = N_{BB} \cdot \epsilon_k \cdot BR(B \to \tau\nu)}$$

$L = 362 fb^{-1}$









$BR(B \to \tau\nu)$ extraction with ToyMC

The sensitivity is estimated by producing 10,000 pseudo-datasets by a $Simultaneous\ fit$ on E_{extra} between 0 and 1 GeV.

- BR set to the PDG value. $BR_{PDG}(B \to \tau\nu) = (1.09 \pm 0.24)\times10^{-4}$

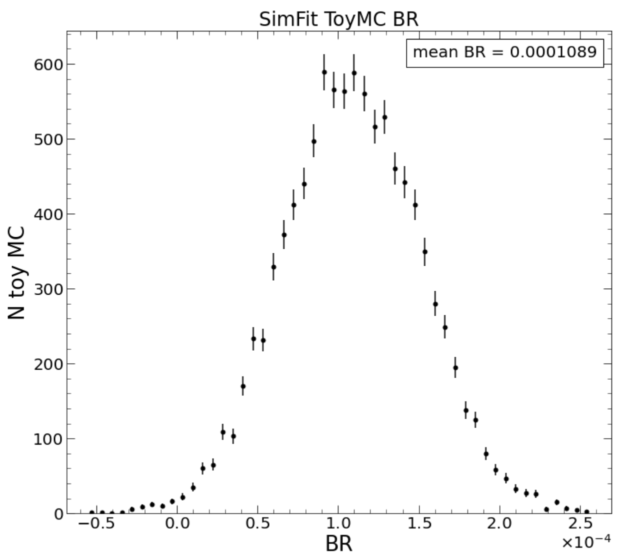
ToyMC result for 362 fb^{-1} :

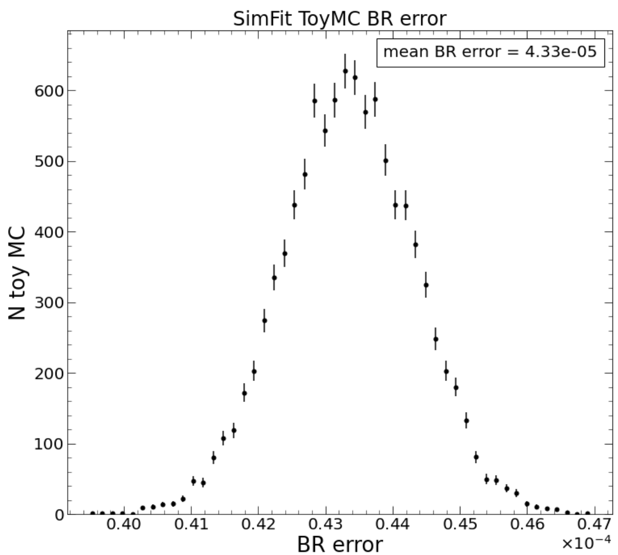
$$\boxed{BR_{simultaneous} = (1.09 \pm 0.43)\times10^{-4}}$$

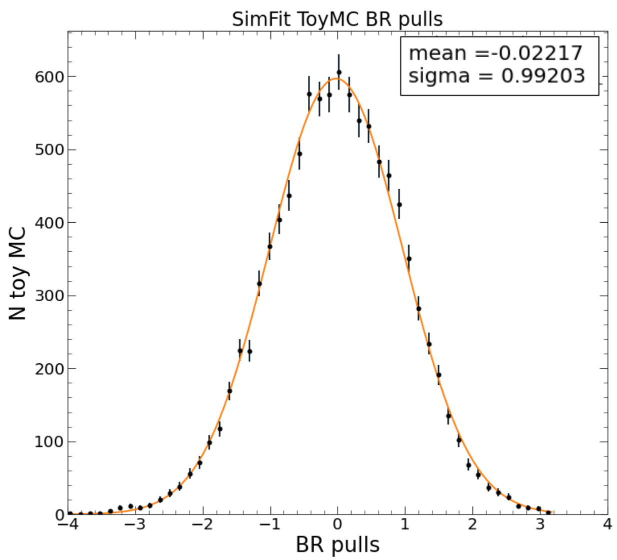
The Bias ($BR_{Input} - BR_{fit}$) is negligible.

Statistical uncertainty ~ 39%.

Belle ~ 38%.
BABAR ~ 29%.







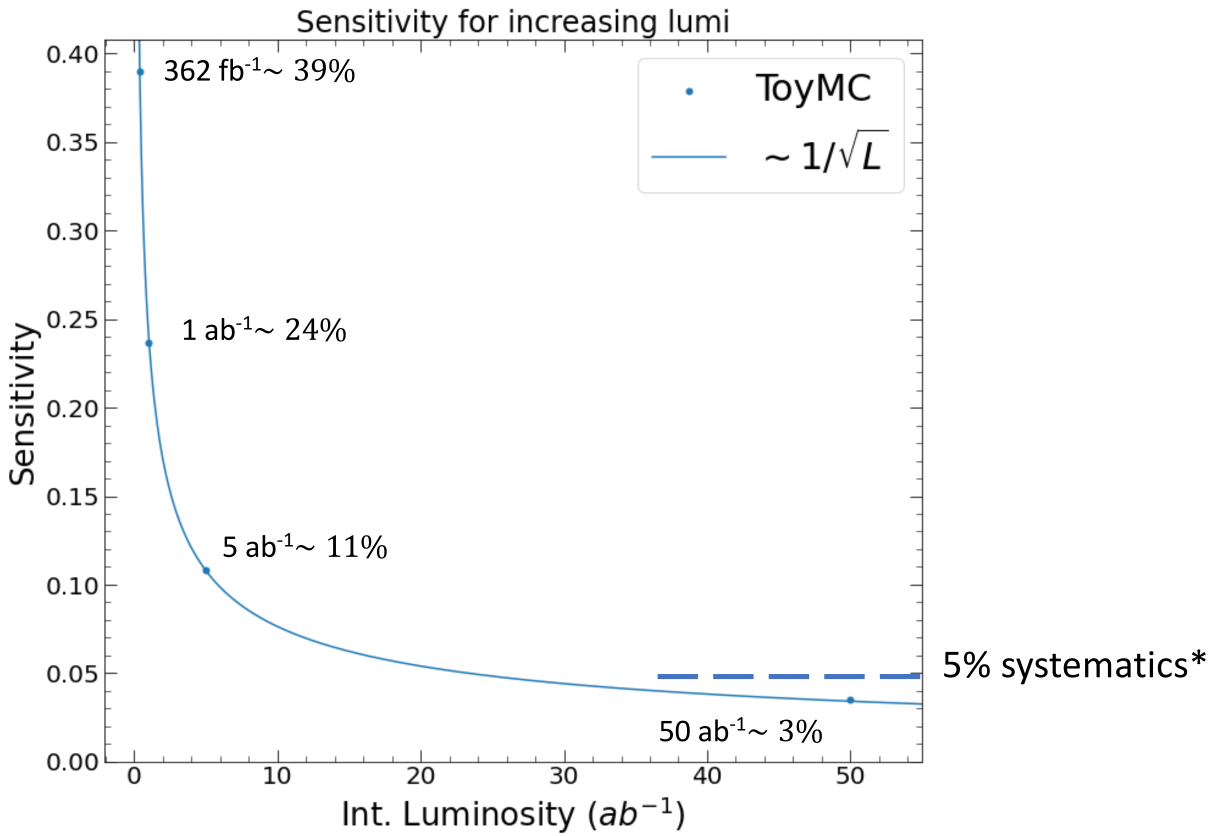
$BR(B \to \tau\nu)$ extrapolation with ToyMC

Extrapolation of statistical uncertainty with toy MC assumimng PDG branching ratio

Toward a 5σ measurment with a single measurement around 1 ab^{-1}

At ~ 50 ab^{-1} the systematics dominate the uncertainty*

systematics uncertainty extrapolation from Belle II Physics Book:
Belle II Coll., Prog. Theor. Exp. Phys. 2019, 123C01
arXiv:1808.10567



Conclusions

- It's important to have a measurement of the purely leptonic decay $\boldsymbol{BR(B \to \tau\nu)}$ from Belle II
  - Complements on-going determinations of semileptonic decays with τ lepton branching ratio from Belle II and LHCb
  - Belle and BaBar measurements a bit in tension and call for an improvement in precision.
  - It appears to be possible only at Belle II

- A Belle II analysis with hadronic B tags is on-going and we are aming at a public result by summer 2023