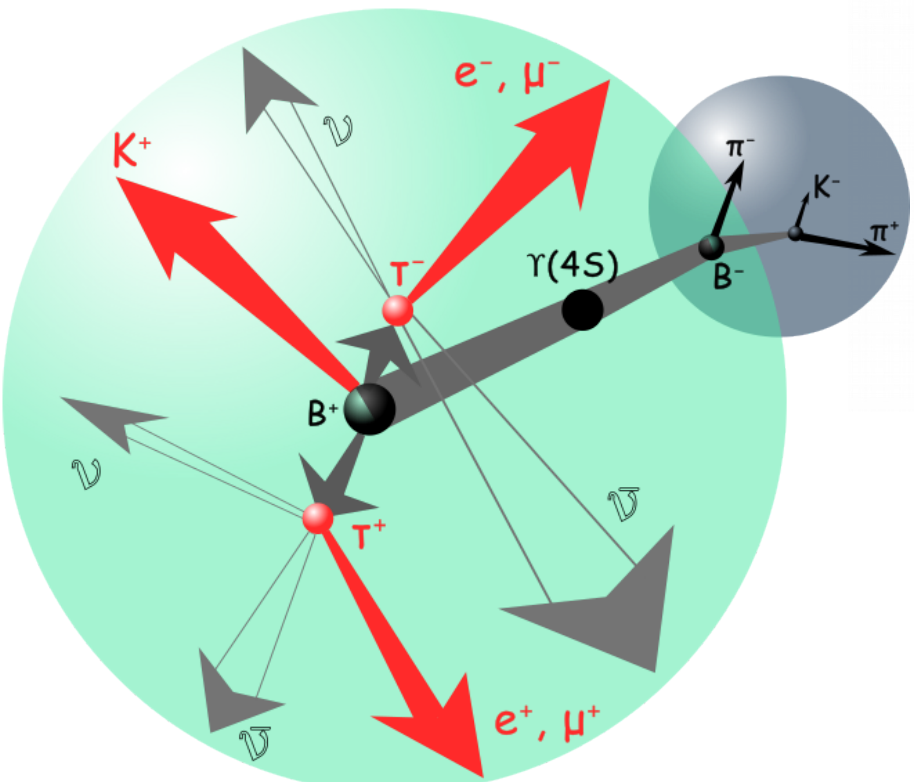
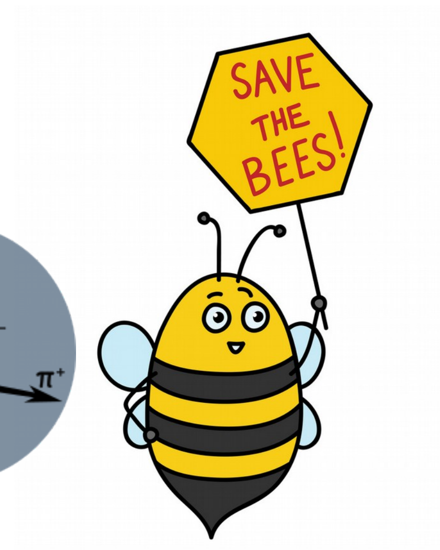
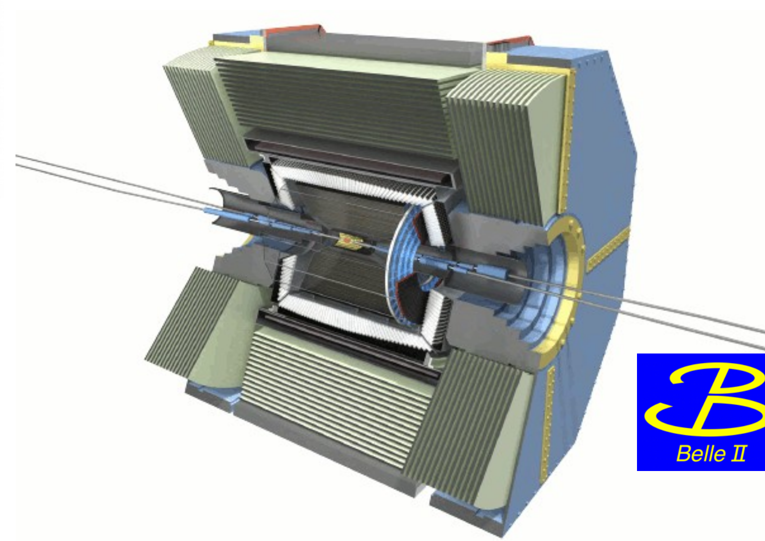
Search for new physics (and else) via modes with missing energy at Belle II

K. Trabelsi
karim.trabelsi@in2p3.fr
(TYL/CPPM)







GDR at Carry-Le-Rouet
November 13, 2025

Belle II, a flavour-factory, a rich physics program...

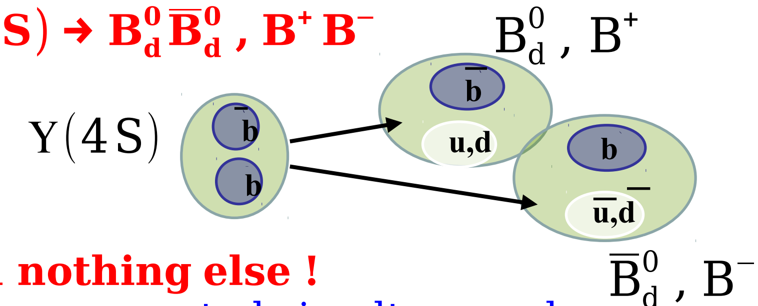
(Belle $\simeq 1\ ab^{-1}$)

◦ We plan to collect $\sim 50\ ab^{-1}$ of e^+e^- collisions at (or close to) the $\Upsilon(4S)$ resonance, where we have:

– a (Super) B-factory ($\sim 1.1 \times 10^9\ B\bar{B}$ pairs per ab^{-1})

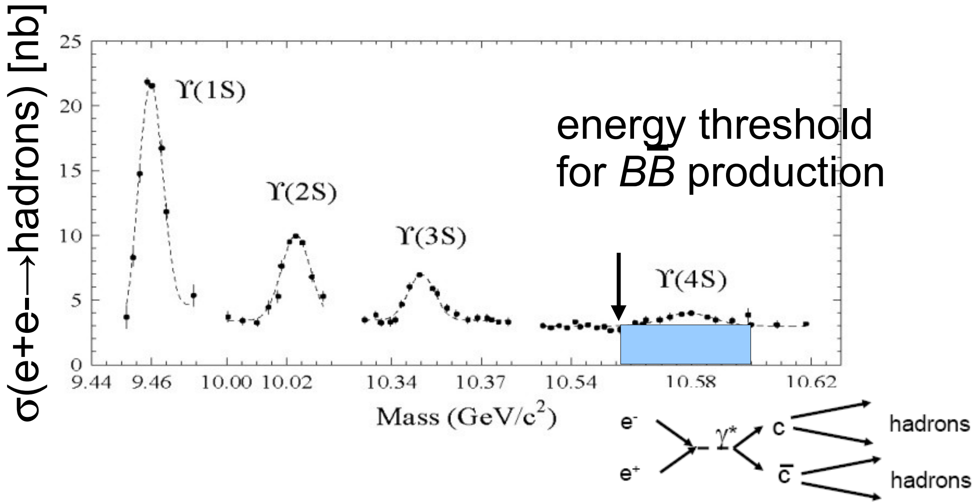
''on resonance'' production

$e^+e^- \to \Upsilon(4S) \to B_d^0 \bar{B}_d^0,\ B^+B^-$



◦ **2 B's and nothing else !**

◦ 2 B mesons are created simultaneously in a L=1 coherent state



– a (Super) charm factory ($\sim 1.3 \times 10^9\ c\bar{c}$ pairs per ab^{-1})

(but also charmonium, X, Y, Z, pentaquarks, tetraquarks, bottomonium...)

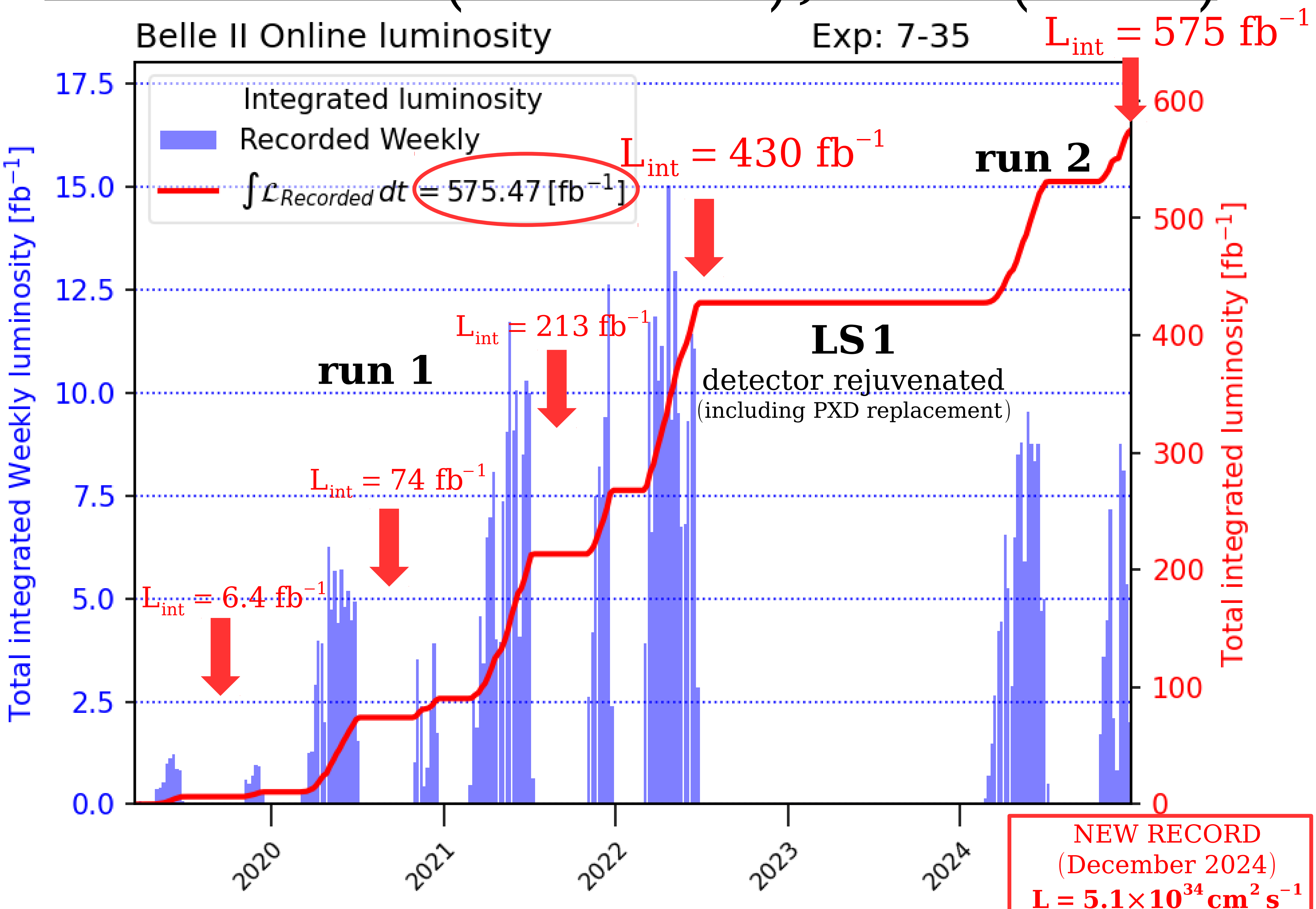
– a (Super) τ factory ($\sim 0.9 \times 10^9\ \tau^+\tau^-$ pairs per ab^{-1})

– exploit the clean e^+e^- environment to probe the existence of exotic hadrons, dark photons/Higgs, light Dark Matter particles, ALPs, LLPs ...

$\Rightarrow$ to reach few $\times 10^{35}\ cm^{-2}s^{-1}$

$\Rightarrow$ cumulate few 10 ab^{-1}

Belle II run 1 (2019 - 2022), run 2 (2024 -)



run 1: data sample used for all results presented here

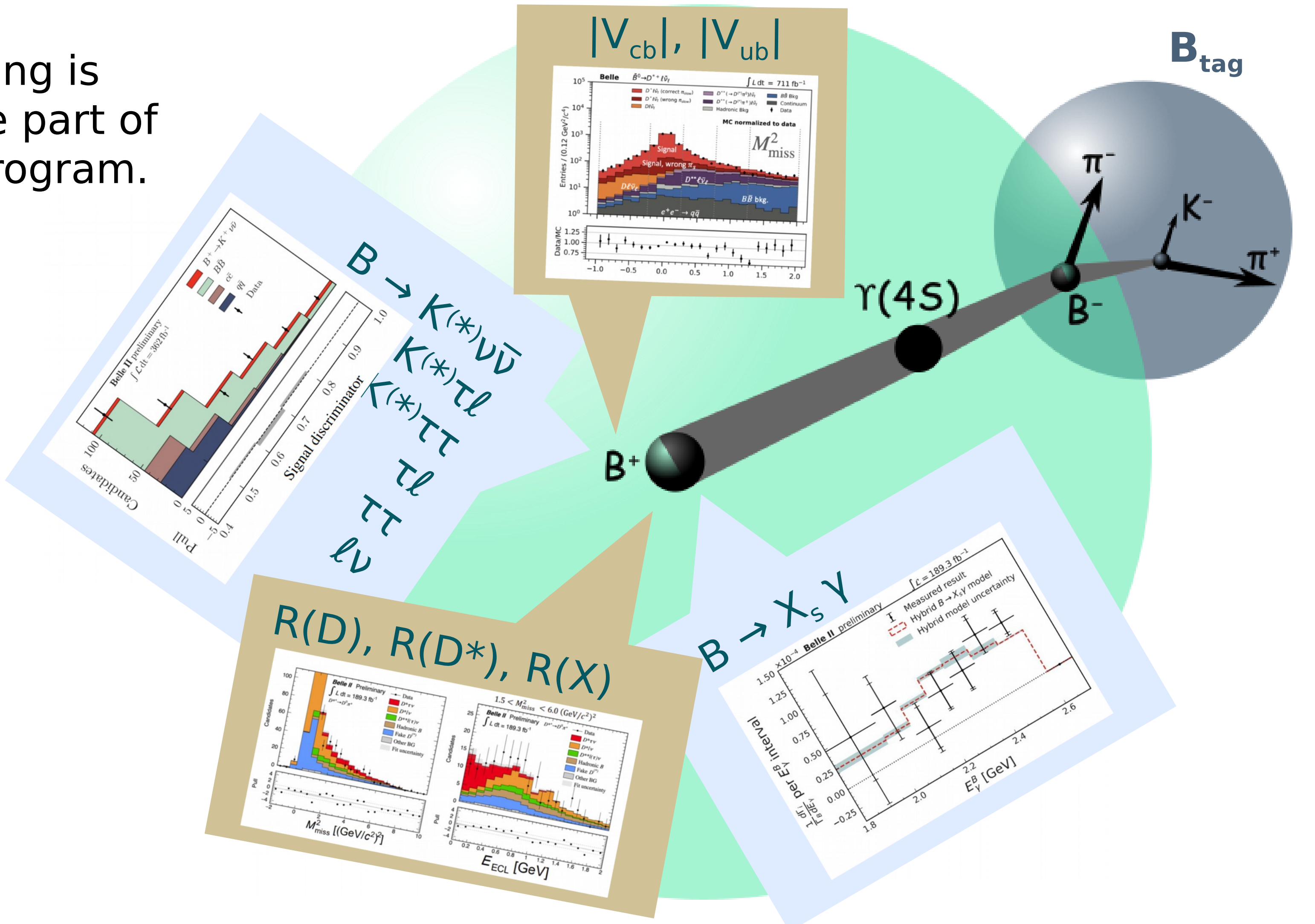
and when possible/relevant add the Belle data sample (+ 1ab^{-1})

Hadronic B-tagging

is widely used in Belle II

It allows neutrino reconstruction at Belle II

Hadronic B-tagging is essential for a large part of Belle II's physics program.



How do we search for B→Kττ ?

The neutrinos escape → using the other B (tag-side)

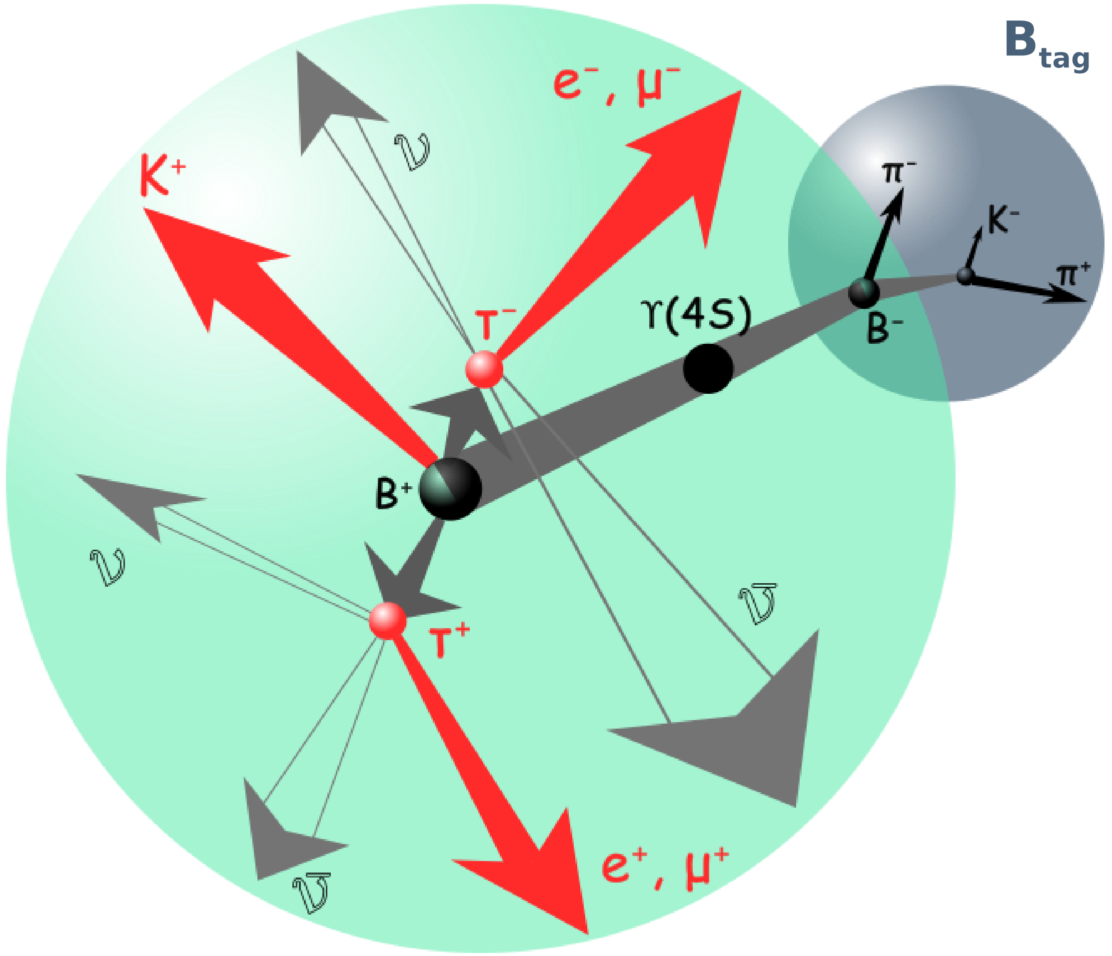
Up to 4 neutrinos in $B^+ \to K^+\tau\tau$
⇒ Cannot reconstruct invariant mass or energy of the B

But, two B-mesons and nothing else in the event !

If the B_{tag} is reconstructed using hadronic decays: Hadronic B-tagging

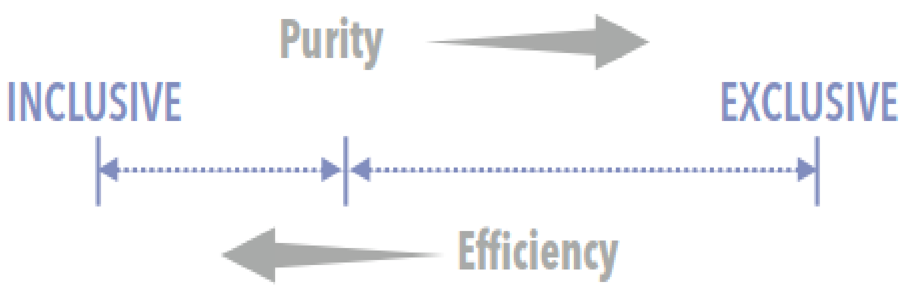
After reconstructing the 3 charged tracks on signal-side and the other B in the event, there will be no additional energy in the calorimeter (E_{ECL}).

⇒ In the rest of the event (ROE), sum of the energies of the clusters should peak at 0.



Missing energy modes and B-tagging

Many interesting B-physics studies involve missing energy: $R(D^{(*)})$, V_{cb}, $K^{(*)}\tau\ell$, $K^{(*)}\tau\tau$, $K^{(*)}\nu\nu$, $\pi l\nu$, $\tau\ell$, $\tau\nu$, $\mu\nu$… which require B-tagging.



Hadronic B-tagging can provide the direction of the B.
Essential in some analysis and unique to B factories!

The 3 important metrics of B-tagging are:

- Efficiency
- Purity
- Data-MC agreement (Calibration factor)

FEI does exclusive B-tagging: Hadronic and Semileptonic

Table 1 Summary of the maximum tag-side efficiency of the Full Event Interpretation and for the previously used exclusive tagging algorithms

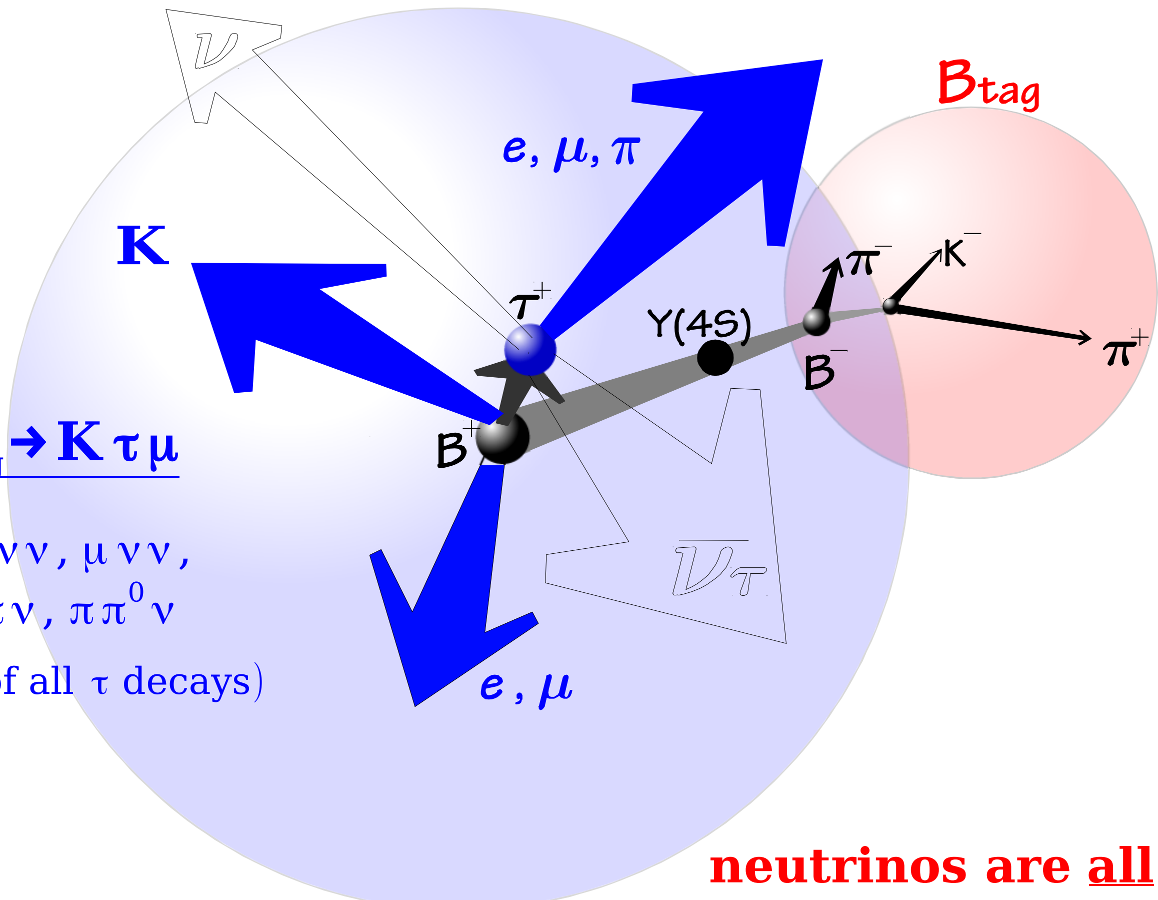
Old measurement in MC	$B^{\pm}$ (%)	B^0 (%)
Hadronic		
FEI with FR channels	0.53	0.33
FEI	0.76	0.46
FR	0.28	0.18
SER	0.4	0.2
Semileptonic		
FEI	1.80	2.04
FR	0.31	0.34
SER	0.3	0.6

[T.Keck et. al, Comput Softw Big Sci (2019) 3: 6]

Exclusive B-tagging:

- **Advantages**: purity, direction of B_{tag}, but also…
  …official training, validation, skims, calibration, systematic (shared knowledge)
- **Disadvantages**: low efficiency…

Event reconstruction in B→Kτμ at B factories



B_{sig}→Kτμ

τ→eνν, μνν,
τ→πν, ππ^{0}ν
(70 % of all τ decays)

B_{tag}

hadronic tag

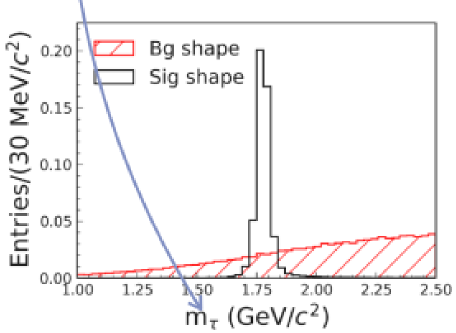
B→D$^{(*)}$π, D$^{(*)}$ρ...
ϵ ~ 0.5%

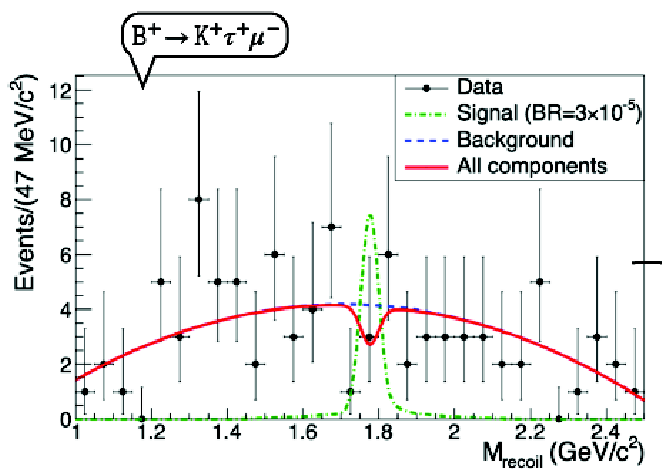
semileptonic tag

B→D$^{(*)}$lν X

neutrinos are all coming from the tau here !

$$m_\tau^2 = (p_{e^+e^-} - p_K - p_\ell - p_{B_{tag}})^2$$



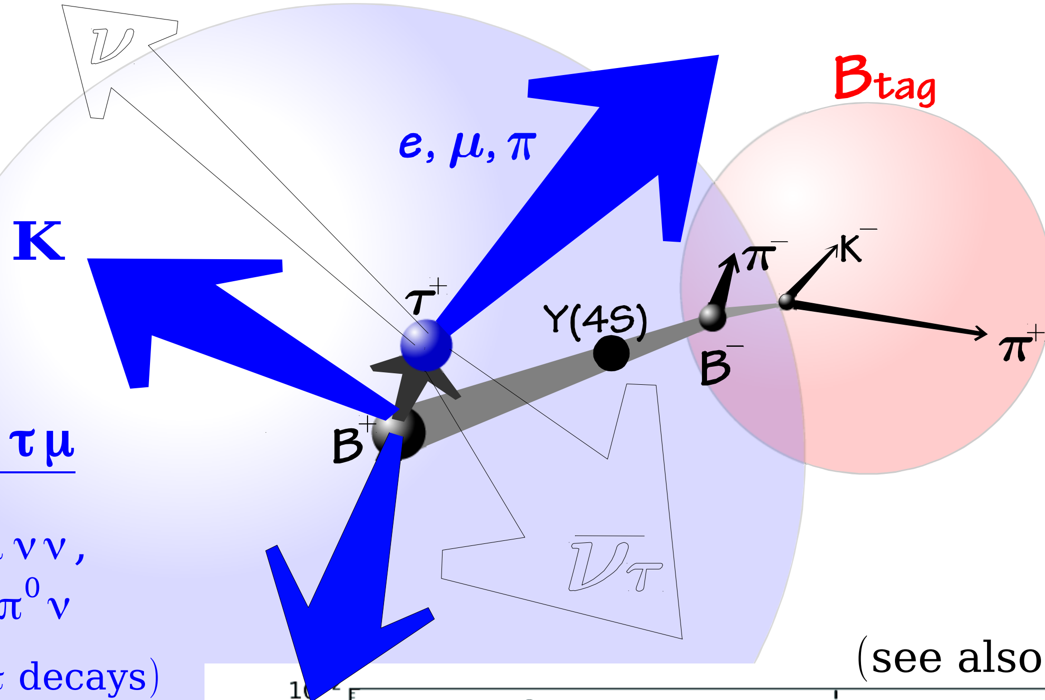


[Belle, PRL 130, 261802 (2023)]

Mode	N_{sig}	ε (%)	$\mathcal{B}^{UL}$ (10^{-5})	$\mathcal{B}^{UL}_{NP}$ (10^{-5})
$B^+ \to K^+\tau^+\mu^-$	-2.1 ± 2.9	0.064	0.59	0.65
$B^+ \to K^+\tau^+e^-$	1.5 ± 5.5	0.084	1.51	1.71
$B^+ \to K^+\tau^-\mu^+$	2.3 ± 4.1	0.046	2.45	2.97
$B^+ \to K^+\tau^-e^+$	-1.1 ± 7.4	0.079	1.53	2.08

PHSP

Event reconstruction in B→Kτμ at B factories



B_{tag}

hadronic tag

$B \to D^{(*)}\pi,\ D^{(*)}\rho \ldots$

$\epsilon \sim 0.5\,\%$

semileptonic tag

$B \to D^{(*)} l \nu\ X$

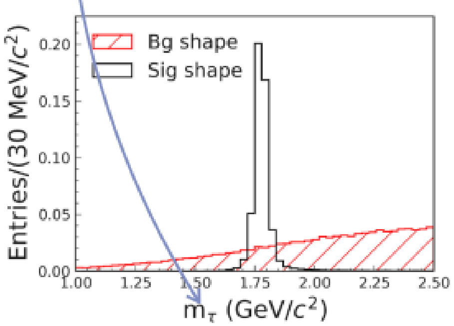
$B_{sig} \to K\tau\mu$

$\tau \to e\nu\nu,\ \mu\nu\nu,$

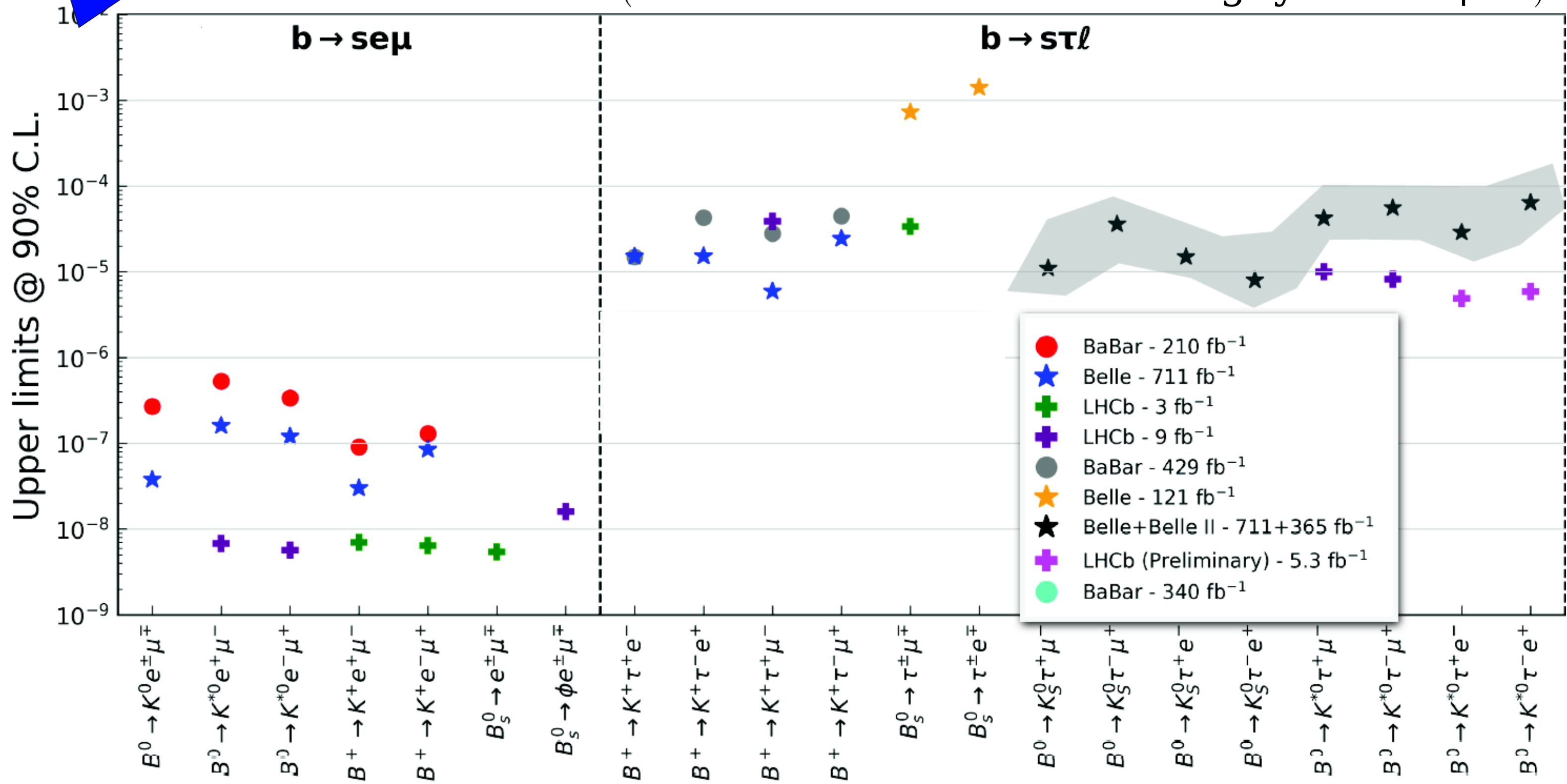
$\tau \to \pi\nu,\ \pi\pi^0\nu$

(70 % of all τ decays)

$$m_\tau^2 = (p_{e^+e^-} - p_K - p_\ell - p_{B_{tag}})^2$$



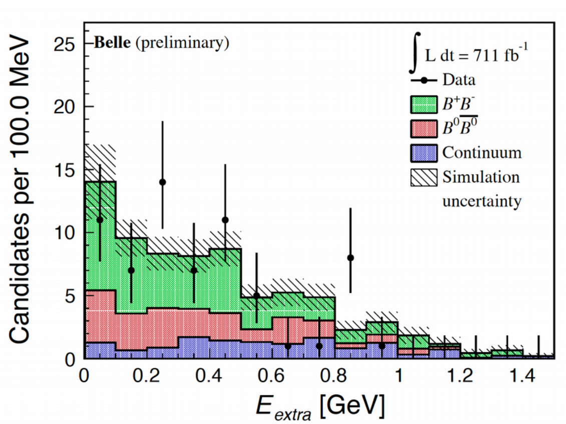
(see also talk of Isaac Consigny for B→ρτl)

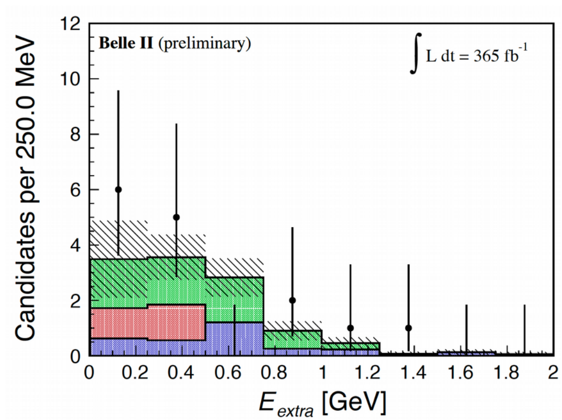


When/why do we use exclusive B-tagging ?

- signal side is reconstructed **exclusively** … examples of 2025…

Search for $B^+ \to K^+ \tau\tau$ [PRELIMINARY]





shown at CKM 2025, Sep 2025

$B_{SM}(B^+ \to K^+\tau\tau) = (1.5 \pm 0.1) \times 10^{-7}$

$\mathcal{B}(B^+ \to K^+\tau^+\tau^-) = 3.13^{+3.70}_{-3.30} \times 10^{-4}$

$\mathcal{B}^{UL}(B^+ \to K^+\tau^+\tau^-) < 8.7 \times 10^{-4}$ at 90% CL

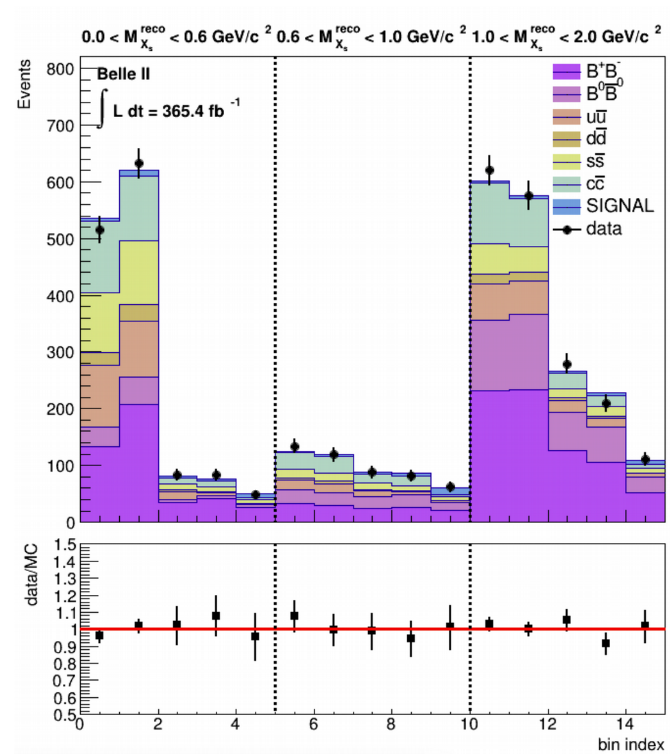
2.6 times better than current world best
Most stringent limit in $B^+ \to K^+\tau\tau$

(see also talk of Mattia Marfoli for $B \to \tau\tau$)

Search for $B \to X_s \nu\bar{\nu}$ [PRELIMINARY]

- $B_{SM} = (2.9 \pm 0.3) \times 10^{-5}$ [JHEP 02 (2015) 184]
- $B < 6.4 \times 10^{-4}$ at 90 % C.L. [ALEPH, EPJC 19 (2001) 213]
- Sum-of-exclusive from 30 decay modes (~90 % of inclusive)

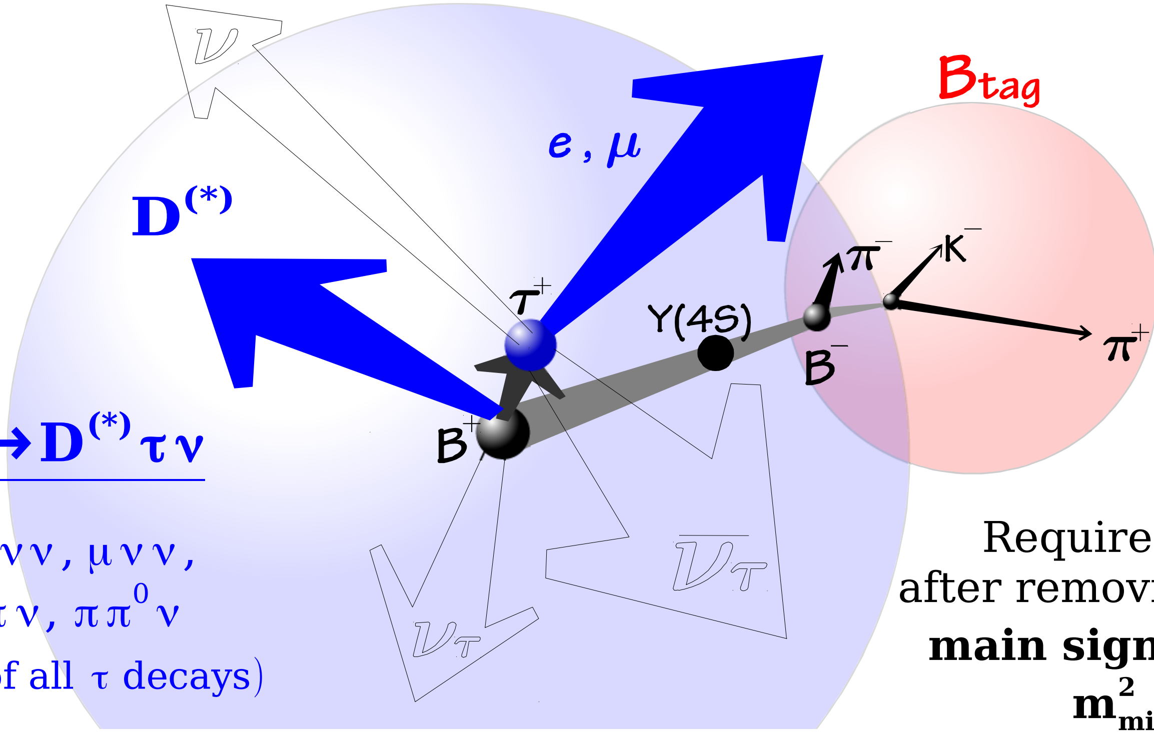
	$B^0\bar{B}^0$			$B^\pm$		
K	K_S^0			$K^\pm$		
$K\pi$	$K^\pm\pi^\mp$	$K_S^0\pi^0$		$K^\pm\pi^0$	$K_S^0\pi^\pm$	
$K2\pi$	$K^\pm\pi^\mp\pi^0$	$K_S^0\pi^\pm\pi^\mp$	$K_S^0\pi^0\pi^0$	$K^\pm\pi^\mp\pi^\pm$	$K_S^0\pi^\pm\pi^0$	$K^\pm\pi^0\pi^0$
$K3\pi$	$K^\pm\pi^\mp\pi^\pm\pi^\mp$	$K_S^0\pi^\pm\pi^\mp\pi^0$	$K^\pm\pi^\mp\pi^0\pi^0$	$K^\pm\pi^\mp\pi^\pm\pi^0$	$K_S^0\pi^\pm\pi^\mp\pi^\pm$	$K_S^0\pi^\pm\pi^0\pi^0$
$K4\pi$	$K^\pm\pi^\mp\pi^\pm\pi^\mp\pi^0$	$K_S^0\pi^\pm\pi^\mp\pi^\pm\pi^\mp$	$K_S^0\pi^\pm\pi^\mp\pi^0\pi^0$	$K^\pm\pi^\mp\pi^\pm\pi^\mp\pi^\pm$	$K_S^0\pi^\pm\pi^\mp\pi^\pm\pi^0$	$K^\pm\pi^\mp\pi^\pm\pi^0\pi^0$
$3K$	$K^\pm K^\mp K_S^0$			$K^\pm K^\mp K^\pm$		
$3K\pi$	$K^\pm K^\mp K^\pm\pi^\mp$	$K^\pm K^\mp K_S^0\pi^0$		$K^\pm K^\mp K^\pm\pi^0$	$K_S^0 K^\pm K^\mp\pi^\pm$	



$B(B \to X_s\nu\bar{\nu}) < 3.6 \times 10^{-4}$ at 90 % C.L.

⇒ The most stringent upper limit on $B \to X_s\nu\bar{\nu}$ decay

Event reconstruction in $B \to D^{(*)} \tau \nu$ at B factories



$B_{sig} \to D^{(*)} \tau \nu$

$\tau \to e\nu\nu,\ \mu\nu\nu,$
$\tau \to \pi\nu,\ \pi\pi^0\nu$
(70 % of all τ decays)

B_{tag}

hadronic tag

$B \to D^{(*)}\pi,\ D^{(*)}\rho \ldots$

$\epsilon \sim 0.5\,\%$

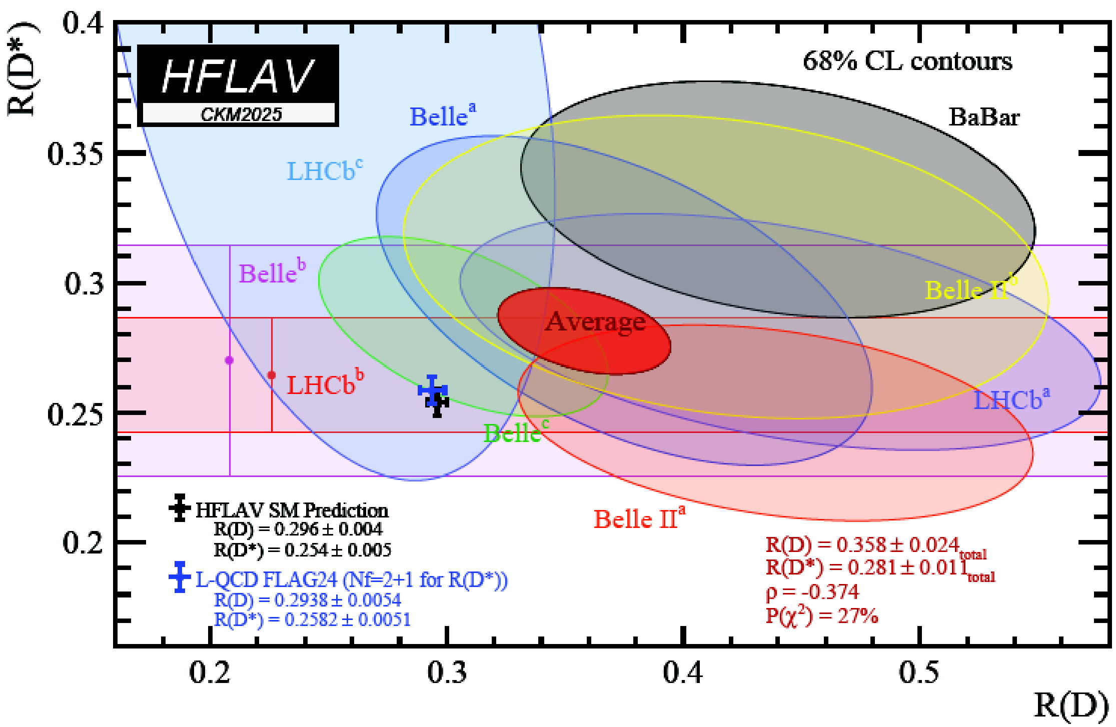
semileptonic tag

$B \to D^{(*)} l \nu X$

Require no particle and no energy left after removing B_{tag} and visible particles of B_{sig}

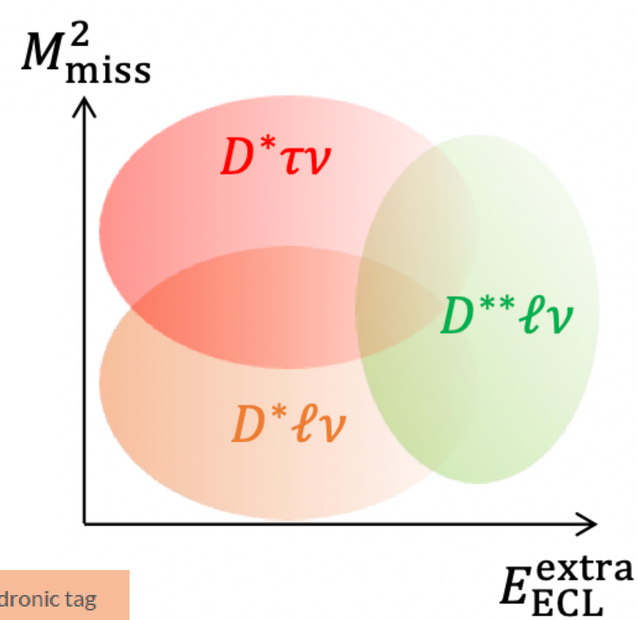
main signal-background discriminator

$$m^2_{miss} = (p_{ee} - p_{tag} - p_{D^{(*)}} - p_l)^2$$



Belle II[a] hadronic tag

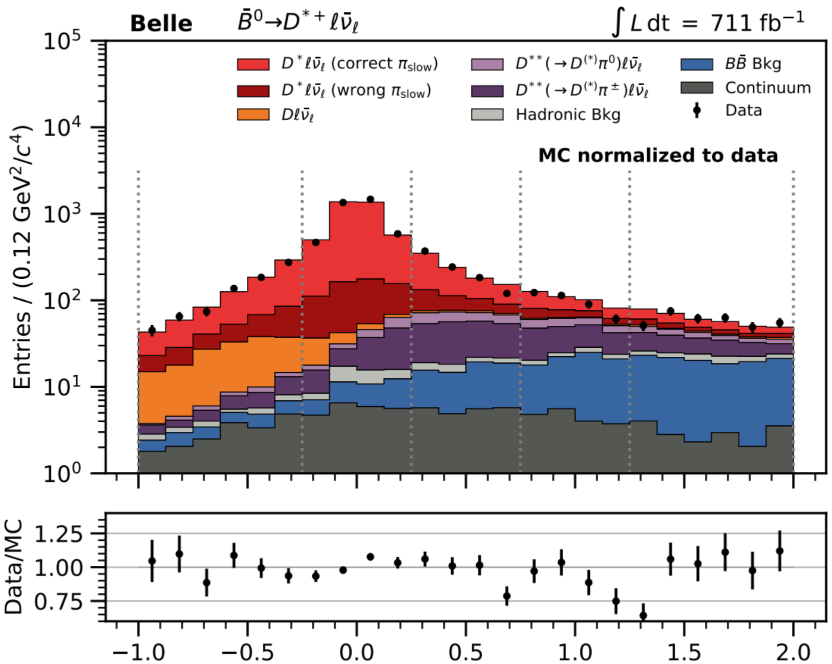
Belle II[b] semileptonic tag (first result with SL B-tag of Belle II)



When/why do we use exclusive B-tagging ?

- not only for search of rare/forbidden decays, or to have high purity…

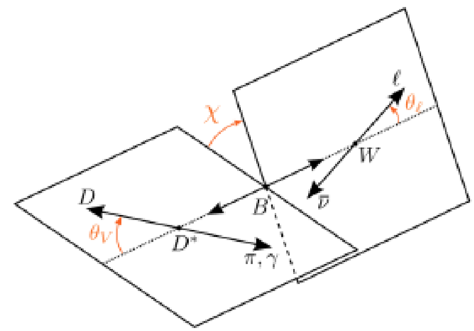
Measurement of angular coefficients with $D^*l\nu$ [Belle, PRD 108(2023)1, 012002/PRL 133 (2024) 131801]

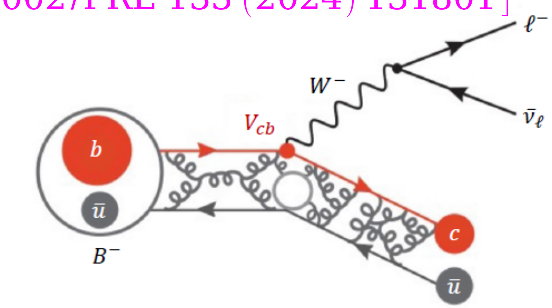


Extraction Method: Missing Mass Squared

$$0 = m_\nu^2 = \mathrm{M}^2_{\mathrm{miss}} = (p_{e^+e^-} - \mathrm{p_B} - \mathrm{p_{D^*}} - \mathrm{p}_\ell)^2$$

$$w = v_B \cdot v_{D^{(*)}} = \frac{m_B^2 + m_{D^{(*)}}^2 - q^2}{2 m_B m_{D^{(*)}}}$$





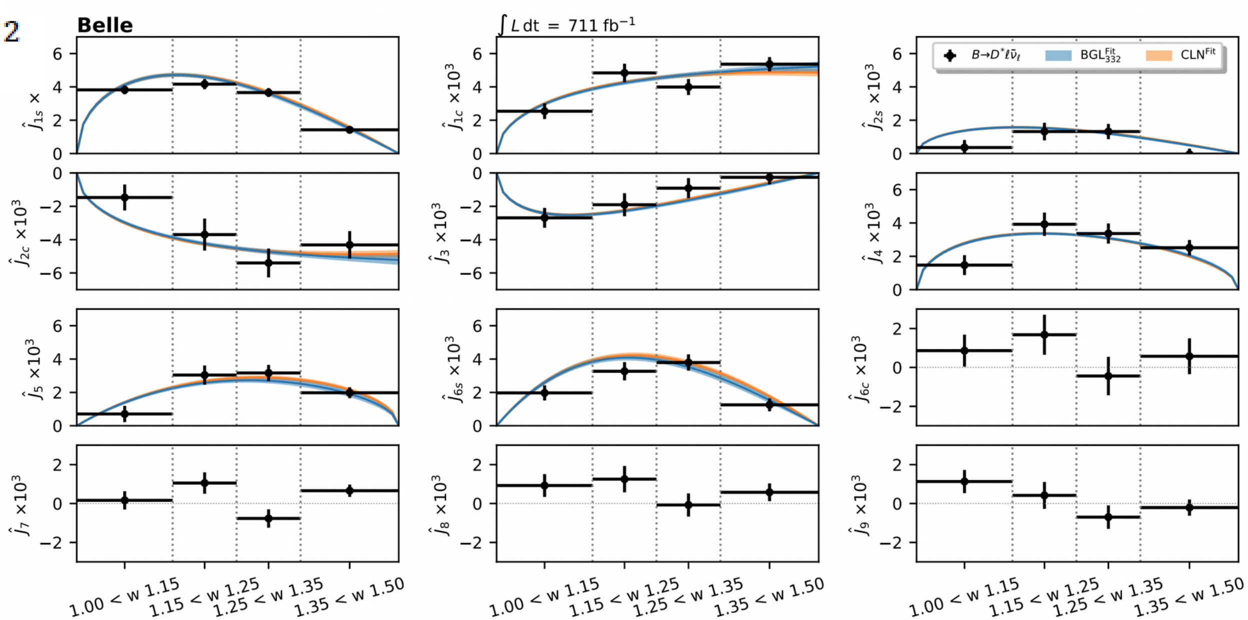
$$\frac{\mathrm{d}\Gamma(B \to D^* \ell \nu_\ell)}{\mathrm{d}w\,\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_V\,\mathrm{d}\chi} = \frac{2G_F^2 \eta_{EW}^2 |V_{cb}|^2 m_B^4 m_{D^*}}{2\pi^4} \times \Big(J_{1s}\sin^2\theta_V + J_{1c}\cos^2\theta_V$$
$$+ (J_{2s}\sin^2\theta_V + J_{2c}\cos^2\theta_V)\cos 2\theta_\ell + J_3 \sin^2\theta_V \sin^2\theta_\ell \cos 2\chi$$
$$+ J_4 \sin 2\theta_V \sin 2\theta_\ell \cos\chi + J_5 \sin 2\theta_V \sin\theta_\ell \cos\chi + (J_{6s}\sin^2\theta_V + J_{6c}\cos^2\theta_V)\cos\theta_\ell$$
$$+ J_7 \sin 2\theta_V \sin\theta_\ell \sin\chi + J_8 \sin 2\theta_V \sin 2\theta_\ell \sin\chi + J_9 \sin^2\theta_V \sin^2\theta_\ell \sin 2\chi \Big) .$$

$$A_{FB} = \frac{3}{2}\frac{(J_{6c} + 2J_{6s})}{3J_{1c} - J_{2c} + 2(3J_{1s} - J_{2s})} ,$$

$$F_L(D^*) = \frac{(3J_{1c} - J_{2c})}{3J_{1c} - J_{2c} + 2(3J_{1s} - J_{2s})} ,$$

(see also talk of Claire Chevallier)

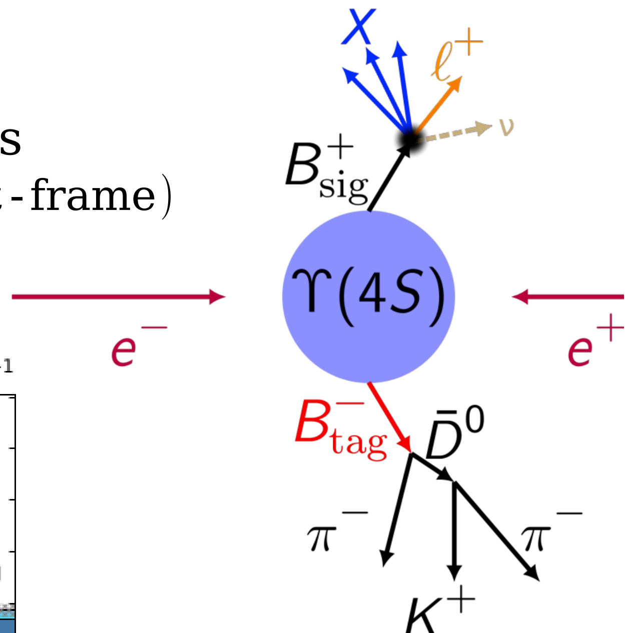
w bin	A_{FB}	$F_L(D^*)$
$1.00 < w < 1.15$	0.23 ± 0.03	0.29 ± 0.04
$1.15 < w < 1.25$	0.30 ± 0.03	0.45 ± 0.04
$1.25 < w < 1.35$	0.29 ± 0.03	0.47 ± 0.03
$1.35 < w < 1.50$	0.16 ± 0.03	0.70 ± 0.03

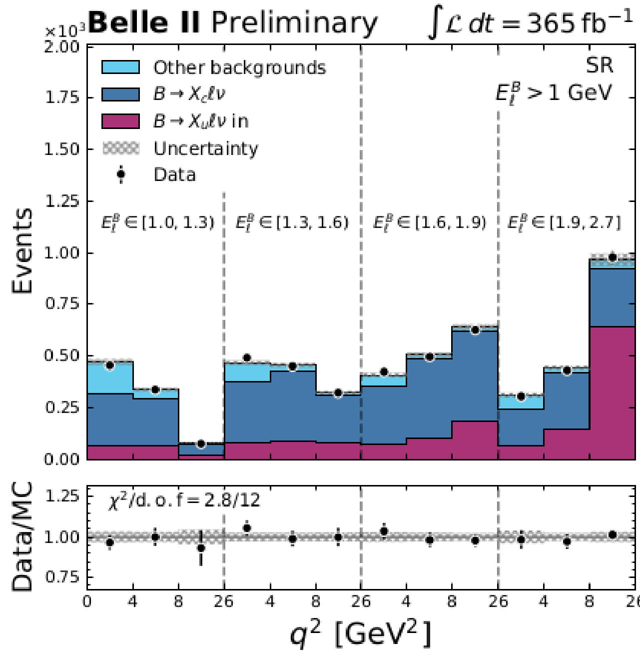


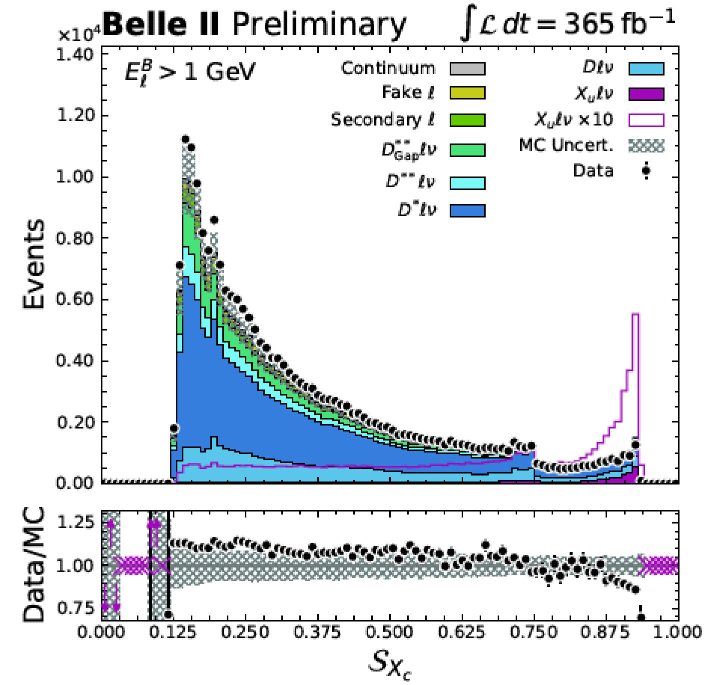
$|V_{ub}|$ from inclusive $B \to X_u l \nu$ decays (had tag)

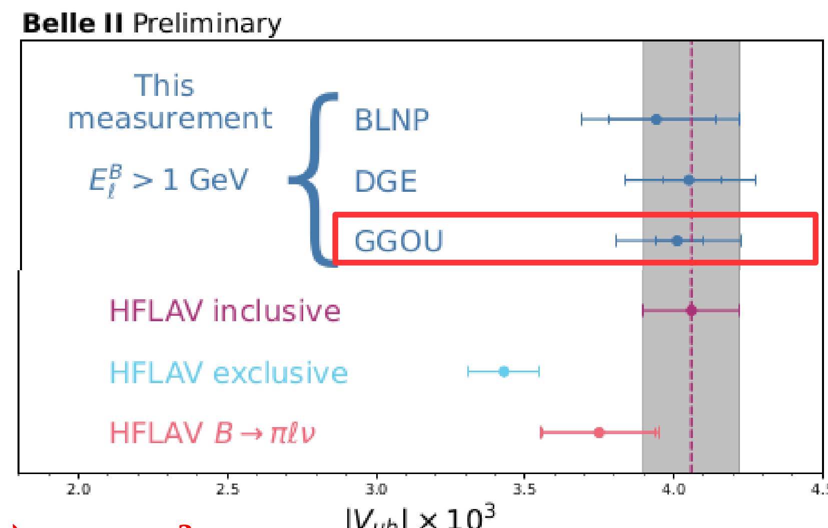
[PRELIMINARY]

- First Belle II measurement
- Hadronic B-tagging
- 3 main kinematical variables
  - $E_l^{(B)}$: lepton energy (in B_{sig} rest-frame)
  - M_X: mass of hadronic system
  - q^2: momentum transfer









Extract $|V_{ub}|$ from partial BR using the predicted partial decay rate over a given phase-space region

$$|V_{ub}| = \sqrt{\frac{\Delta\mathcal{B}(B \to X_u \ell \nu)}{\tau_B \Delta\Gamma(B \to X_u \ell \nu)}}$$

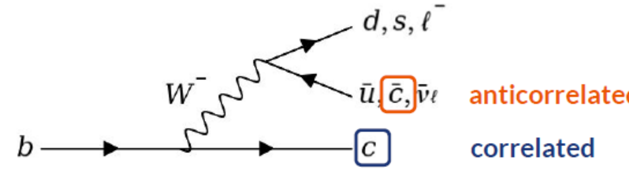
$$|V_{ub}|_{GGOU} = (4.01 \pm 0.11(\text{stat}) \pm 0.16(\text{syst})\ ^{+0.09}_{-0.07}(\text{theo})) \times 10^{-3}$$

$$|V_{ub}|^{HFLAV}_{incl} = (4.06 \pm 0.16) \times 10^{-3}$$

When/why do we use exclusive B-tagging ?

- signal side is **partially** reconstructed…
  - measurement of inclusive differential BFs:

    $B^0 \to \Lambda_c^- X$, $B^0 \to \Lambda_c^+ X$, $B^+ \to \Lambda_c^- X$, $B^+ \to \Lambda_c^+ X$



Experimental results on inclusive (only BaBar hep-ex/0606026)

$B(B^+ \to D^0 X) = (8.6 \pm 0.7)\%$ | $B(B^0 \to D^0 X) = (8.1 \pm 1.5)\%$

$B(B^+ \to \bar{D}^0 X) = (79 \pm 4)\%$ | $B(B^0 \to \bar{D}^0 X) = (47.4 \pm 2.8)\%$

$B(B^+ \to D^+ X) = (2.5 \pm 0.5)\%$ | $B(B^0 \to D^+ X) < 3.9\%$

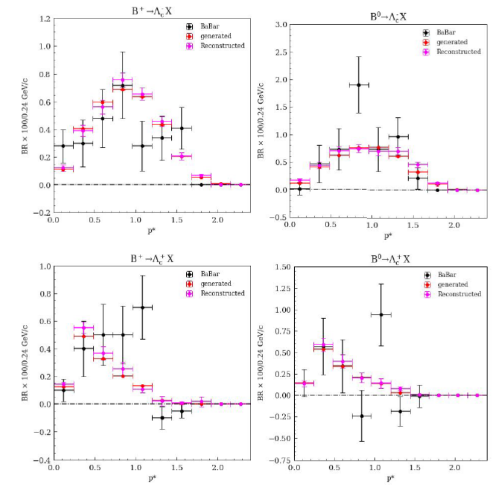
$B(B^+ \to D^- X) = (9.9 \pm 1.2)\%$ | $B(B^0 \to D^- X) = (36.9 \pm 3.3)\%$

$B(B^+ \to D_s^+ X) = (7.9^{+1.4}_{-1.3})\%$ | $B(B^0 \to D_s^+ X) = (10.3^{+2.1}_{-1.8})\%$

$B(B^+ \to D_s^- X) = (1.10^{+0.40}_{-0.32})\%$ | $B(B^0 \to D_s^- X) < 2.6\%$

$B(B^+ \to \Lambda_c^+ X) = (2.1^{+0.9}_{-0.6})\%$ | $B(B^0 \to \Lambda_c^+ X) < 3.1\%$

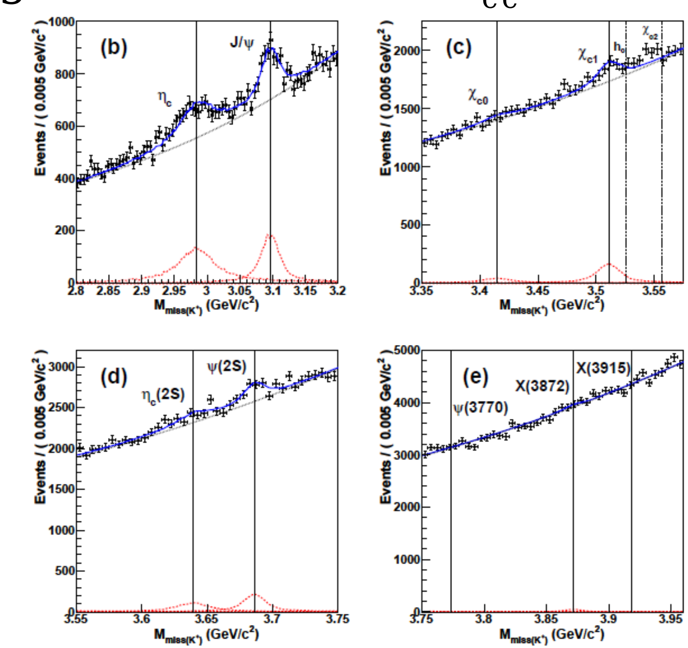
$B(B^+ \to \Lambda_c^- X) = (2.8^{+1.1}_{-0.9})\%$ | $B(B^0 \to \Lambda_c^- X) = (5.0^{+2.1}_{-1.5})\%$



- $B^{+/0} \to X_s \gamma$, $J/\psi X$…
- Measurements of the absolute branching fractions of $B^+ \to X_{c\bar{c}} K^+$

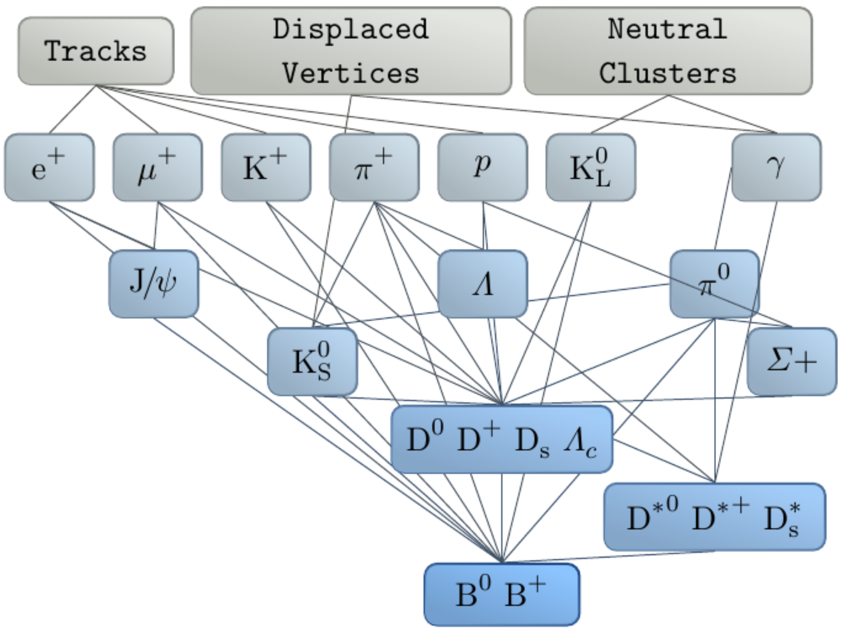
arXiv:1709.06108, Phys. Rev. D 97, 012005 (2018)

Mode	Yield	Significance (σ)	$\epsilon(10^{-3})$	$\mathcal{B}$ (10^{-4})	World average for $\mathcal{B}$ (10^{-4}) [10]
η_c	2590 ± 180	14.2	2.73 ± 0.02	$12.0 \pm 0.8 \pm 0.7$	9.6 ± 1.1
J/ψ	1860 ± 140	13.7	2.65 ± 0.02	$8.9 \pm 0.6 \pm 0.5$	10.26 ± 0.031
χ_{c0}	430 ± 190	2.2	2.67 ± 0.02	$2.0 \pm 0.9 \pm 0.1$ (< 3.3)	$1.50^{+0.15}_{-0.14}$
χ_{c1}	1230 ± 180	6.8	2.68 ± 0.02	$5.8 \pm 0.9 \pm 0.5$	4.79 ± 0.23
$\eta_c(2S)$	1050 ± 240	4.1	2.77 ± 0.02	$4.8 \pm 1.1 \pm 0.3$	3.4 ± 1.8
$\psi(2S)$	1410 ± 210	6.6	2.79 ± 0.02	$6.4 \pm 1.0 \pm 0.4$	6.26 ± 0.24
$\psi(3770)$	-40 ± 310	-	2.76 ± 0.02	$-0.2 \pm 1.4 \pm 0.0$ (< 2.3)	4.9 ± 1.3
$X(3872)$	260 ± 230	1.1	2.79 ± 0.01	$1.2 \pm 1.1 \pm 0.1$ (< 2.6)	(< 3.2)
$X(3915)$	80 ± 350	0.3	2.79 ± 0.01	$0.4 \pm 1.6 \pm 0.0$ (< 2.8)	-



Hadronic B-tagging tool at Belle/Belle II

called Full Event Interpretation (FEI)



Designed for Belle II software, now used with Belle data also.

For each decay, BDTs trained on MC.

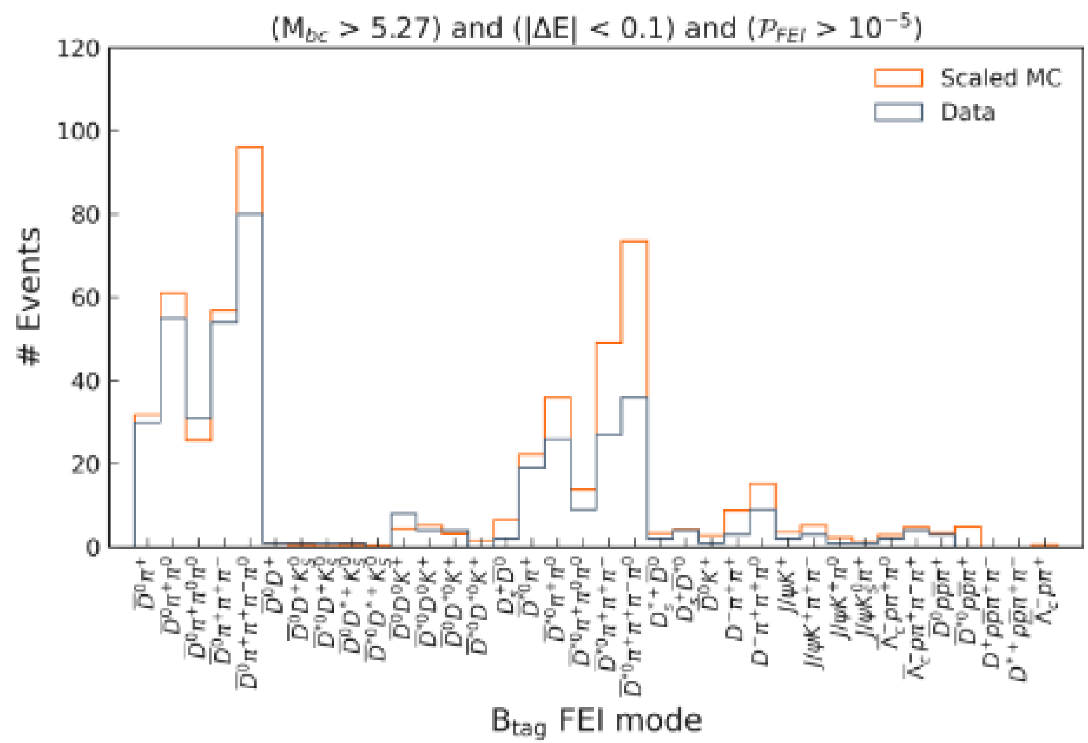
B^+-tagging uses 36 decays.
But only 12 of them, essentially $B \rightarrow D^{(*)} m\pi^{\pm} n\pi^0$, gives ~90% of the efficiency.

Total efficiency < 1%.

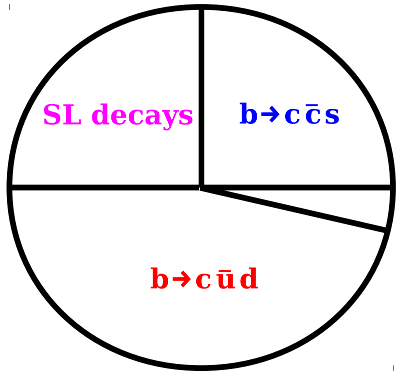
Hierarchical reconstruction...
o(10^4) B total decay chains

Uses machine learning: over 200 BDTs **trained on simulated BB data**

Outputs:
- List of tagged B candidates (each in a specific B decay cascade)
- A "signal probability" for each...



Why is B-decay modeling so hard ?



Inclusive decays for b→c transition

A.Lenz et al, arXiv:1305.5390, 1404.6197

$B(b\to c\bar{u}d) = 0.446\pm0.014$

$B(b\to c\bar{c}s) = 0.232\pm0.007$

$B(b\to c e \nu_e) = 0.116\pm0.008$

$B(b\to c \mu \nu_\mu) = 0.116\pm0.008$

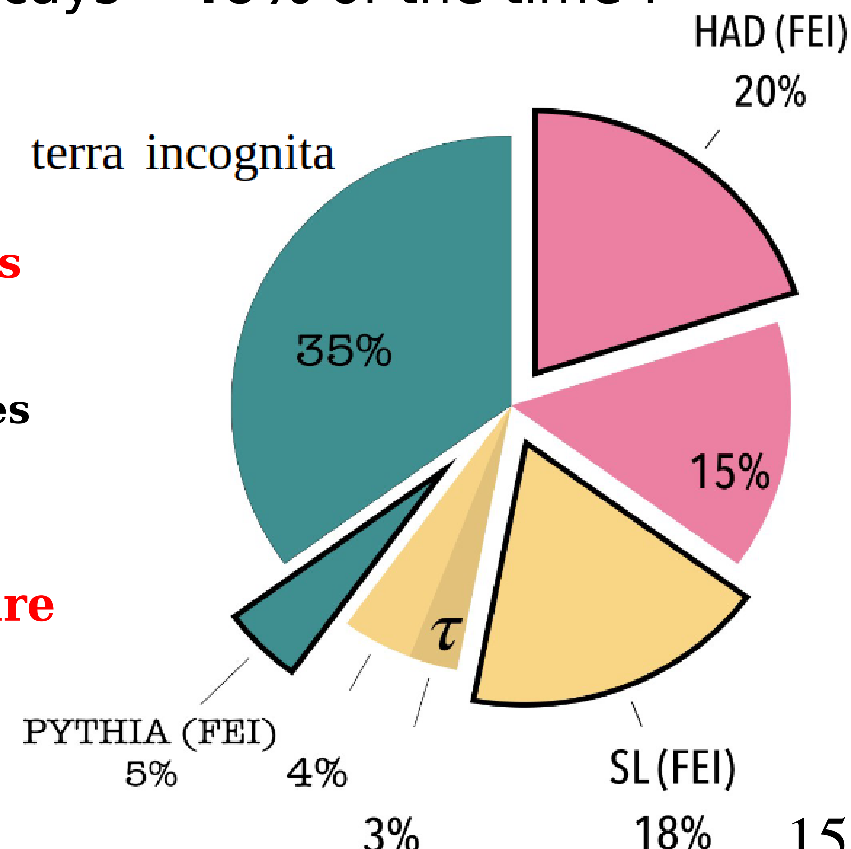
$B(b\to c \tau \nu_\tau) = 0.027\pm0.001$

$B(b\to c\bar{u}s) = 0.024\pm0.001$

$B(b\to c\bar{c}d) = 0.0126\pm0.0005$

We will see that we (and PDG) use a 30-year-old measurement with ~75% uncertainty for one of the largest hadronic B-decays...

But on top of that, we don't know how B decays **~40%** of the time !
We ask **PYTHIA** to (poorly) generate them.



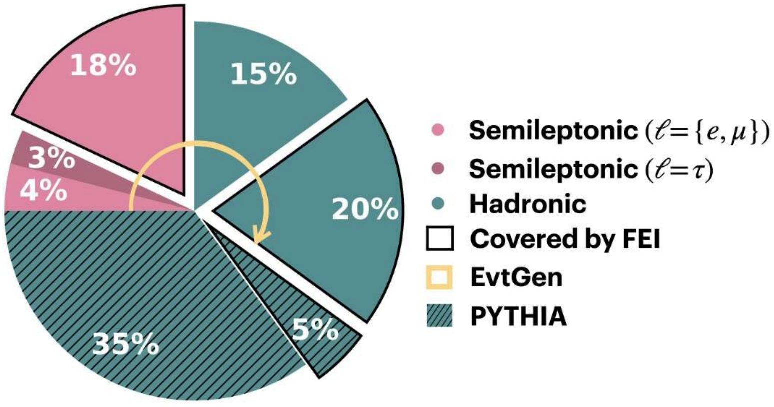
B-tagging is key tool for missing energy analyses

- low efficiency (efficiency for hadronic B-tagging < 1%)
- **and ML can't (always) save you...**
  B-tagging algorithms are trained using MC samples
- 40% of hadronic B decays generated by PYTHIA ...
- and even among the EvtGen part...

most BFs measured are old measurements from ARGUS, CLEO...

lot of hadronic B decays to understand/measure
⇒ new contributions to B-tagging ??

How are B decays generated ?



EvtGen

Hadronic B-decays: $\sim 75\%$ of the total branching fraction

```
Decay B+

0.054900000   anti-D*0  e+  nu_e       BGL 0.02596 -0.06049 0.01311 0.01713 0.00753 -0.09346,
0.023100000   anti-D0  e+  nu_e        BGL 0.0126 -0.094 0.34 -0.1 0.0115 -0.057 0.12 0.4;
0.007570000   anti-D_10  e+  nu_e       LLSW 0.71 -1.6 -0.5 2.9;
0.003890000   anti-D_0*0  e+  nu_e      LLSW 0.68 -0.2 0.3;
0.004310000   anti-D'_10  e+  nu_e      LLSW 0.68 -0.2 0.3;
0.003730000   anti-D_2*0  e+  nu_e      LLSW 0.71 -1.6 -0.5 2.9;
              .
              .
              .
              .
0.000383590   D+  anti-D0               PHSP;
0.000392390   D*+  anti-D0              SVS;
0.000630000   anti-D*0  D+              SVS;
0.000810000   anti-D*0  D*+             SVV_HELAMP 0.56 0.0 0.96 0.0 0.47 0.0;
              .
              .
```

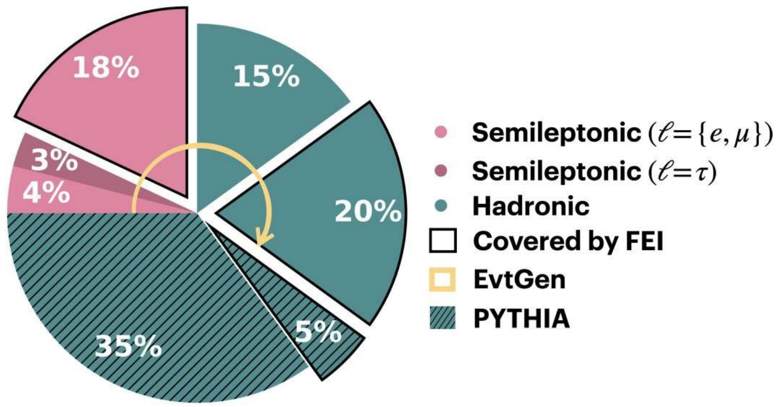
The largest decays are at 10^{-2}, 10^{-3}
so talking about $o(10^4)$ decay channels
we only list $o(10^3)$ explicitly

This is from PDG and some guestimates…
but what about the rest ?

How are B decays generated ?

EvtGen + PYTHIA

Hadronic B-decays: ~ 75% of the total branching fraction but only about half of it is measured
PYTHIA is used to generate the other half in MC



Quark transition	modeID in PYTHIA v8	$\mathcal{B}^{\text{Belle}}(\%)$	$\mathcal{B}^{\text{Belle II}}(\%)$
u anti-d anti-c u	23	31.23	20.26
u anti-d anti-c u	43	-	3.87
u anti-s anti-c u	43	2.23	2.02
c anti-s anti-c u	43	-	6.66
c anti-d anti-c u	43	-	0.36
u anti-d anti-u u	23	-	0.27
c anti-s anti-u u	23	-	0.36
u anti-u anti-d u	23	-	0.18
d anti-d anti-d u	23	-	<0.01
s anti-s anti-d u	23	-	0.01
u anti-u anti-s u	23	-	0.20
d anti-d anti-s u	23	-	0.16
s anti-s anti-s u	23	-	0.13
anti-s u	91	-	0.45
anti-cd_1 uu_1	63	3.40	2.97
anti-cd_1 uu_1	64	1.27	-
anti-cs_0 cu_0	63	0.85	-
anti-cs_1 uu_1	63	0.18	0.81
anti-cs_1 uu_1	64	0.04	-
anti-cd_0 cu_0	63	0.04	-
Total PYTHIA contribution		39.24	38.71

- PYTHIA is called for quark fragmentation according to relative rates determined by the parameters of the StringFlav class
- We use the default values for most parameters, with the production of some excited mesons turned off, like $a_1^{\pm}$, a_1^0, D^{**}...

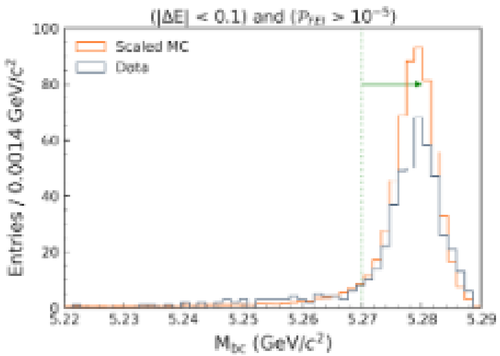
The StringFlav parameters as well as relative fractions assigned to different quark transitions need to be tuned

- Fragmentation compares the final state with the explicitly listed decays, and if found, performed again to produce an alternative final state
- Therefore, to exclude that a particular decay is generated by PYTHIA, it can be explicitly listed in DECAY.DEC with a branching fraction of 0%

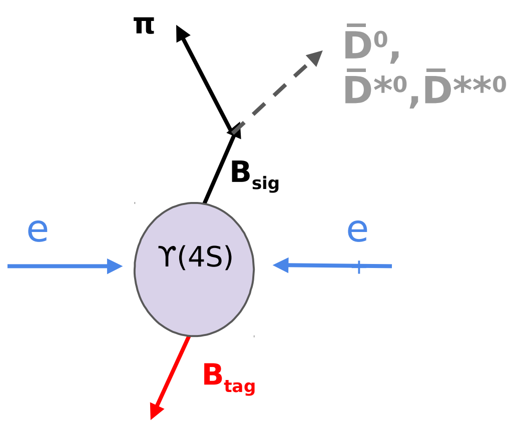
Need to know what not to generate as well

Ideal control sample to study B-tagging

First idea, use $B \to J/\psi K$:

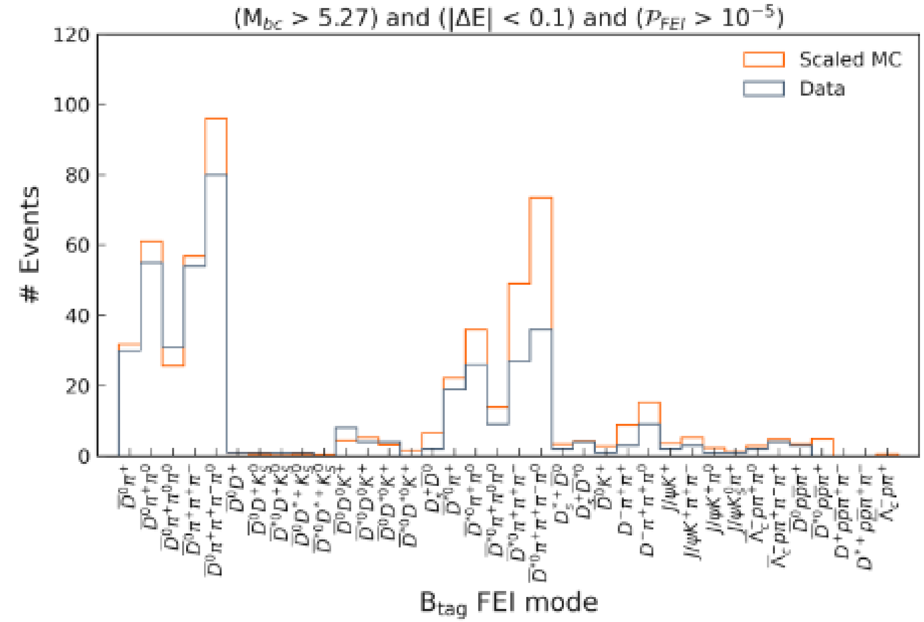


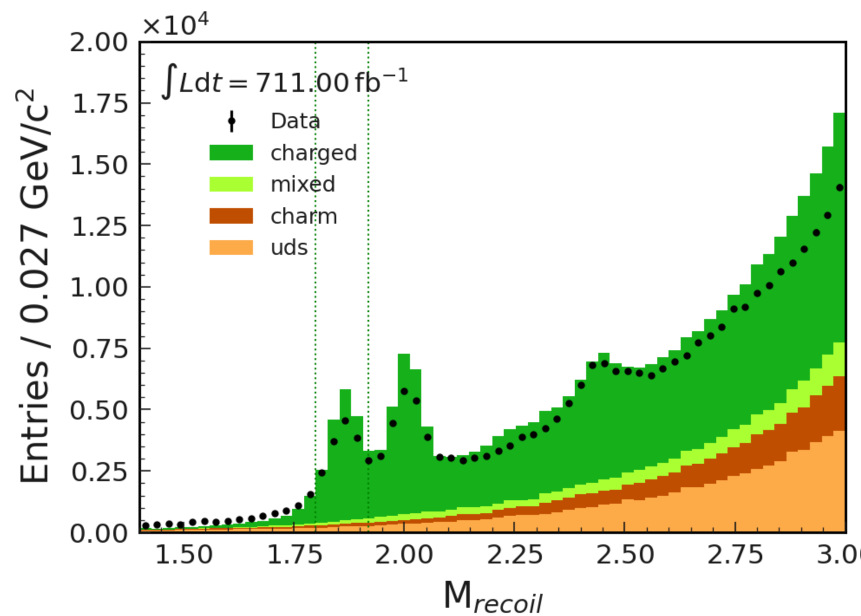
clean, allow first estimation (large MC/data differences)
⇒ but too limited stat (~ 400 evts after B-tagging)



We can look for D^0, D^{*0} and even D^{**0} in the recoil mass of a fully reconstructed B and a $\pi^\pm$

Within a narrow region around the peak, we know that one B decays to $D^0\pi^+$ and we can study the other B (decaying hadronically)





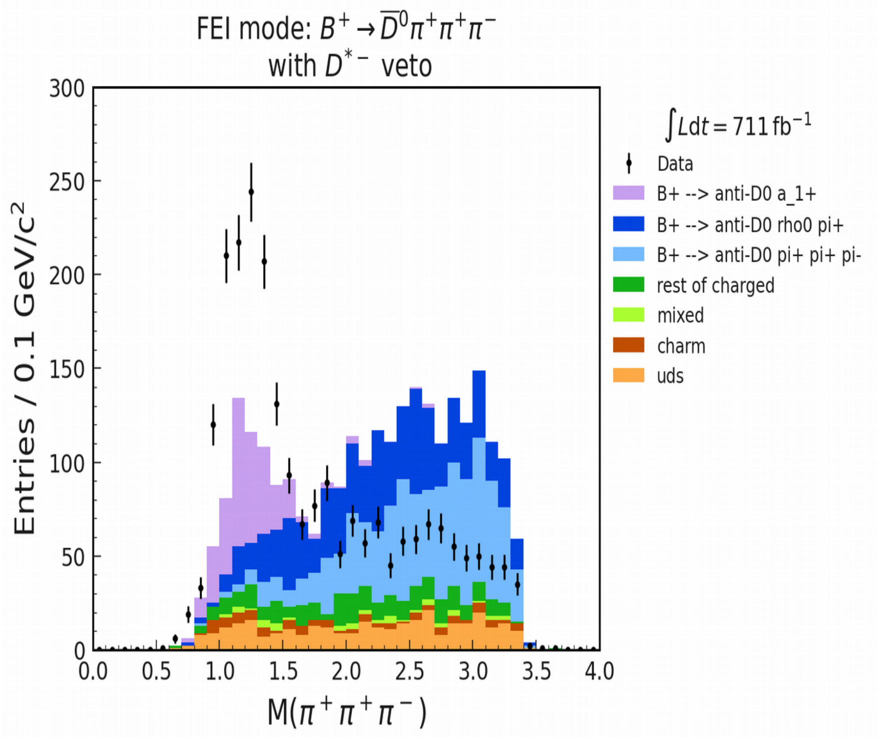
~16k events in a 3σ window around each peak in data.

Need to calibrate the algorithm, but more importantly, need to improve MC for training.

Improving MC model: an example

Let's take one final state for example: $B^+ \to \bar{D}^0 \pi^+ \pi^+ \pi^-$.
It can be produced through many intermediate states:

Decay	Belle	Belle II
$B^+ \to \overline{D}^0 \pi^- \pi^+ \pi^+$	0.46	0.51
$B^+ \to \overline{D}^0 \rho(770)^0 \pi^+; \rho(770)^0 \to \pi^+ \pi^-$	0.39	0.42
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to \rho(770)^0 \pi^+; \rho(770)^0 \to \pi^+ \pi^-$	0.13	0.14
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to f_0(600) \pi^+; f_0(600) \to \pi^+ \pi^-$	0.05	-
$B^+ \to \overline{D}_1(2420)^0 \pi^+; \overline{D}_1(2420)^0 \to D^*(2010)^- \pi^+; D^*(2010)^- \to \overline{D}^0 \pi^-$	0.04	0.02
$B^+ \to \overline{D}_1(2430)^0 \pi^+; \overline{D}_1(2430)^0 \to D^*(2010)^- \pi^+; D^*(2010)^- \to \overline{D}^0 \pi^-$	0.03	0.02
$B^+ \to \overline{D}_2^*(2460)^0 \pi^+; \overline{D}_2^*(2460)^0 \to D^*(2010)^- \pi^+; D^*(2010)^- \to \overline{D}^0 \pi^-$	0.01	0.01
$B^+ \to D^*(2010)^- \pi^+ \pi^+; D^*(2010)^- \to \overline{D}^0 \pi^-$	-	0.09
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to \pi^+ \pi^+ \pi^-$	-	0.07
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to f_0(500) \pi^+; f_0(500) \to \pi^+ \pi^-$	-	0.05
$B^+ \to \overline{D}_1(2420)^0 \pi^+; \overline{D}_1(2420)^0 \to \overline{D}^0 \pi^- \pi^+$	-	0.02
$B^+ \to \overline{D}^0 K^*(892)^+; K^*(892)^+ \to K^0 \pi^+; K^0 \to K_S^0; K_S^0 \to \pi^+ \pi^-$	-	0.01
Rest of Exclusive	0.03	0.03
Sum of Exclusive	1.12	1.38
Sum of Pythia	0	0
Total Sum	1.12	1.38



Comparing with data clearly shows that a_1^+ component is underestimated, and the $\rho^0\pi^+$ and direct $\pi^+ \pi^+ \pi^-$ components are overestimated.

Similarly , for other final states

$B^+ \to \bar{D}^0\ \pi^+\ \pi^+\ \pi^-\ \pi^0$

Decay	Belle	Belle II
$B^+ \to \overline{D}^{*0}\pi^-\pi^+\pi^+\pi^0$	1.80	1.80
$B^+ \to \overline{D}^{*0}\omega(782)\pi^+$; $\omega(782) \to \pi^-\pi^+\pi^0$	0.40	0.41
Rest of Exclusive	0.02	0.05
Sum of Exclusive	2.22	2.25
$B^+ \to \overline{D}^{*0}\rho(770)^0\rho(770)^+$; $\rho(770)^0 \to \pi^+\pi^-$; $\rho(770)^+ \to \pi^+\pi^0$	0.49	0.20
$B^+ \to \overline{D}^{*0}\rho(770)^+\pi^+\pi^-$; $\rho(770)^+ \to \pi^+\pi^0$	0.40	0.20
$B^+ \to \overline{D}^{*0}\rho(770)^0\pi^+\pi^0$; $\rho(770)^0 \to \pi^+\pi^-$	0.40	0.20
$B^+ \to \overline{D}^{*0}\rho(770)^-\pi^+\pi^+$; $\rho(770)^- \to \pi^-\pi^0$	0.20	0.10
$B^+ \to \overline{D}^{*0}\eta\pi^+$; $\eta \to \pi^-\pi^+\pi^0$	0.14	0.07
$B^+ \to \overline{D}_1(2430)^0\rho(770)^0\pi^+$; $\overline{D}_1(2430)^0 \to \overline{D}^{*0}\pi^0$; $\rho(770)^0 \to \pi^+\pi^-$	0.03	-
Rest of PYTHIA	0.02	0.01
Sum of PYTHIA	1.68	0.77
Total Sum	3.90	3.03

blue means generated by PYTHIA

$\bar{D}^{*0}\ \pi^+\ \pi^+\ \pi^-$

TABLE VI: Contents of the DECAY file concerning the $B^+ \to \overline{D}^{*0}\pi^+\pi^+\pi^-$ final state and corresponding measurements in PDG [in %]. The rows in blue correspond to decays produced by Pythia.

Decay	Belle	Belle II	Marker	Ref
$B^+ \to \overline{D}^*(2007)^0\pi^-\pi^+\pi^+$	1.03	-	■	[2], [7]
$B^+ \to \overline{D}^*(2007)^0a_1(1260)^+$; $a_1(1260)^+ \to \rho(770)^0\pi^+$; $\rho(770)^0 \to \pi^+\pi^-$	0.66	0.58	★	
$B^+ \to \overline{D}^*(2007)^0a_1(1260)^+$; $a_1(1260)^+ \to f_0(600)\pi^+$; $f_0(600) \to \pi^+\pi^-$	0.25	-	★	
$B^+ \to \overline{D}^*(2007)^0a_1(1260)^+$; $a_1(1260)^+ \to \pi^+\pi^+\pi^-$	-	0.28	★	
$B^+ \to \overline{D}^*(2007)^0a_1(1260)^+$; $a_1(1260)^+ \to f_0(500)\pi^+$; $f_0(500) \to \pi^+\pi^-$	-	0.20	★	
$B^+ \to \overline{D}^*(2007)^0\rho(770)^0\pi^+$; $\rho(770)^0 \to \pi^+\pi^-$	-	0.04	★	
Rest of Exclusive	0.02	0.05		
Sum of Exclusive	1.96	1.15		
$B^+ \to \overline{D}^*(2007)^0f_0(980)\pi^+$; $f_0(980) \to \pi^+\pi^-$	0.05	-	★	
$B^+ \to \overline{D}^*(2007)^0\pi^+\pi^+\pi^-$	-	0.20	■	
Rest of Pythia	0.00	0.00		
Sum of Pythia	0.05	0.20		
Total Sum	2.01	1.35		

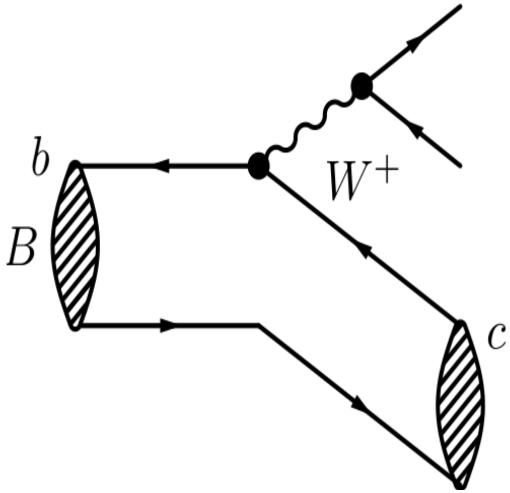
$B^+ \to \bar{D}^0\ \pi^+\ \pi^+\ \pi^-\ \pi^0$

Marker convention:
★ : Old/No measurement
■ : Double counting

TABLE IX: Contents of the DECAY file concerning the $B^+ \to \overline{D}^0\pi^+\pi^+\pi^-\pi^0$ final state and corresponding measurements in PDG [in %]. The rows in blue correspond to decays produced by Pythia.

Decay	Belle	Belle II	Markers	Ref
$B^+ \to D^*(2010)^-\pi^0\pi^+\pi^+$; $D^*(2010)^- \to \overline{D}^0\pi^-$	1.02	1.03	★	[8]
$B^+ \to \overline{D}^*(2007)^0\pi^-\pi^+\pi^+$; $\overline{D}^*(2007)^0 \to \overline{D}^0\pi^0$	0.64	-	■	
$B^+ \to \overline{D}^*(2007)^0a_1(1260)^+$; $\overline{D}^*(2007)^0 \to \overline{D}^0\pi^0$; $a_1(1260)^+ \to \rho(770)^0\pi^+$; $\rho(770)^0 \to \pi^+\pi^-$	0.41	0.38	★	
$B^+ \to \overline{D}^0\omega(782)\pi^+$; $\omega(782) \to \pi^-\pi^+\pi^0$	0.37	0.37	★	[9]
$B^+ \to \overline{D}^*(2007)^0a_1(1260)^+$; $\overline{D}^*(2007)^0 \to \overline{D}^0\pi^0$; $a_1(1260)^+ \to f_0(600)\pi^+$; $f_0(600) \to \pi^+\pi^-$	0.16	-	★	
$B^+ \to D^*(2010)^-\rho(770)^+\pi^+$; $D^*(2010)^- \to \overline{D}^0\pi^-$; $\rho(770)^+ \to \pi^+\pi^0$	0.14	0.14	★	
$B^+ \to \overline{D}^*(2007)^0a_1(1260)^+$; $\overline{D}^*(2007)^0 \to \overline{D}^0\pi^0$; $a_1(1260)^+ \to \pi^+\pi^+\pi^-$	-	0.18	★	
$B^+ \to \overline{D}^*(2007)^0a_1(1260)^+$; $\overline{D}^*(2007)^0 \to \overline{D}^0\pi^0$; $a_1(1260)^+ \to f_0(500)\pi^+$; $f_0(500) \to \pi^+\pi^-$	-	0.13	★	
Rest of Exclusive	0.03	0.10		
Sum of Exclusive	2.75	2.32		
$B^+ \to \overline{D}^0\rho(770)^+\pi^+\pi^-$; $\rho(770)^+ \to \pi^+\pi^0$	0.20	0.30		
$B^+ \to \overline{D}^0\rho(770)^0\rho(770)^+$; $\rho(770)^0 \to \pi^+\pi^-$; $\rho(770)^+ \to \pi^+\pi^0$	0.20	0.20		
$B^+ \to \overline{D}^0\rho(770)^-\pi^+\pi^+$; $\rho(770)^- \to \pi^-\pi^0$	0.10	0.10		
$B^+ \to \overline{D}^0\rho(770)^0\pi^+\pi^0$; $\rho(770)^0 \to \pi^+\pi^-$	0.10	0.20		
$B^+ \to \overline{D}^0\eta\pi^+$; $\eta \to \pi^-\pi^+\pi^0$	0.05	0.07	★	
$B^+ \to \overline{D}_1(2430)^0\pi^+\pi^0$; $\overline{D}_1(2430)^0 \to D^*(2010)^-\pi^+$; $D^*(2010)^- \to \overline{D}^0\pi^-$	0.05	-		
$B^+ \to \overline{D}_0^*(2300)^0\rho(770)^0\pi^+$; $\overline{D}_0^*(2300)^0 \to \overline{D}^0\pi^0$; $\rho(770)^0 \to \pi^+\pi^-$	0.03	-		
$B^+ \to \overline{D}^*(2007)^0f_0(980)\pi^+$; $\overline{D}^*(2007)^0 \to \overline{D}^0\pi^0$; $f_0(980) \to \pi^+\pi^-$	0.03	-	■	
$B^+ \to \overline{D}_2^*(2460)^0\rho(770)^0\pi^+$; $\overline{D}_2^*(2460)^0 \to \overline{D}^0\pi^0$; $\rho(770)^0 \to \pi^+\pi^-$	0.02	-		
$B^+ \to \overline{D}_2^*(2460)^0\pi^+\pi^0$; $\overline{D}_2^*(2460)^0 \to D^*(2010)^-\pi^+$; $D^*(2010)^- \to \overline{D}^0\pi^-$	0.01	-		
$B^+ \to \overline{D}^*(2007)^0\pi^+\pi^+\pi^-$; $\overline{D}^*(2007)^0 \to \overline{D}^0\pi^0$	-	0.13	■	
$B^+ \to \overline{D}^0\pi^+\pi^+\pi^-\pi^0$	-	0.10		
Rest of Pythia	0.01	0.01		
Sum of Pythia	0.79	1.10		
Total Sum	3.54	3.42	★	

Model for B → $D^{(*,\,**)}$ nπ mπ^0 decays



X = π, ρ, a_1, ωπ, ρππ, ηπ

Y = D, D*, D**

Happens through 2 channels,
one with spectator quarks (call Y) and
one from the W (call X).

We modify the DECAY table to latest PDG/paper interpretations and this model to see the impact.

Essentially validation, we do not want to fine-tune (except set 0 there is no signal).

2 primary rules:

- D^0 X: D^{*0} X : D^{**0} X ~= 1 : 1 : 1
  (based on observation from D π^- : D* π^- : D** π^- and D ρ^- : D* ρ^-)
- Y π^- : Y ρ^- : Y a_1^- ~= 1 : 2.5 : 2.5
  (based on predictions and confirmed with $\tau \rightarrow h\,\nu$ decays)

Additional information:

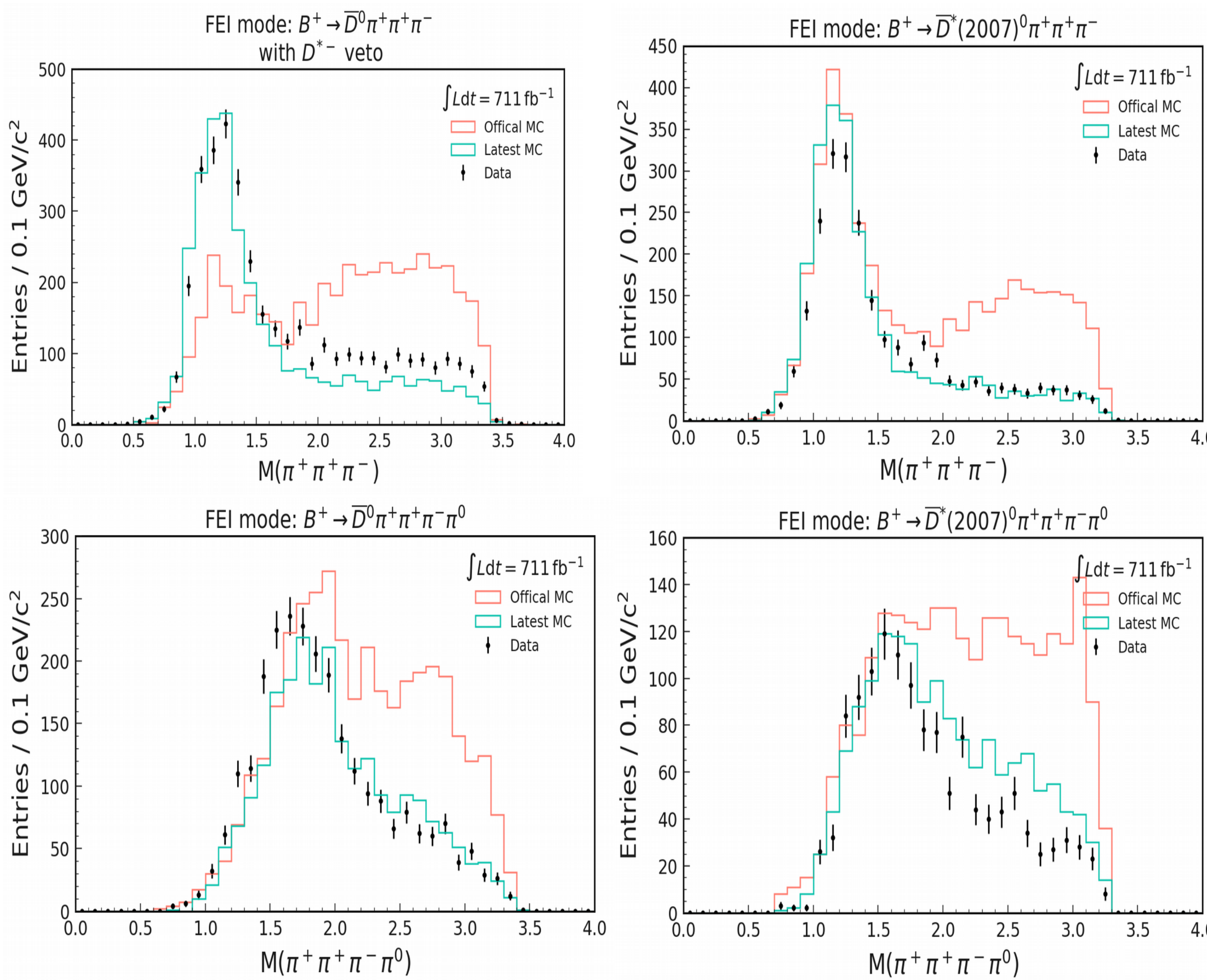
- 3π π^0 is hard to model without some sort of ρ' resonance
  - For **ωπ** we fix from measurements.
  - For **ρππ** and **ηπ**, we let PYTHIA generate it**.**
- Decays of D** particles is synchronized with Belle II
- The fraction of 4 different D** is fixed based on observations.

Decay description is improved !

The improvement is not limited to calibration factors, but more importantly in the invariant masses (of intermediate particles), which are used as training variables in FEI

3π± case:

3π± π⁰ case:



improving description of hadronic B decays ⇒ improve B-tagging efficiency

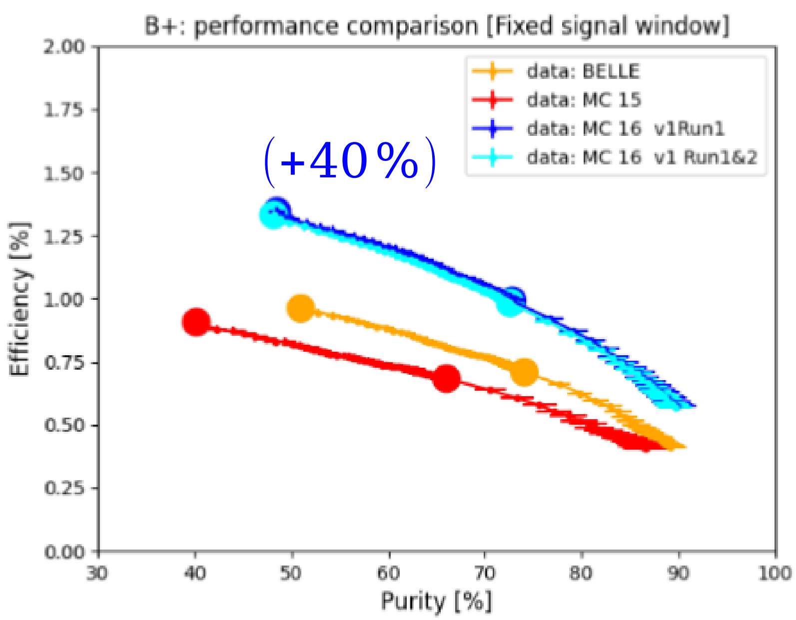
Hadronic B-tagging performance
(in recently reprocessed data, run1+2)

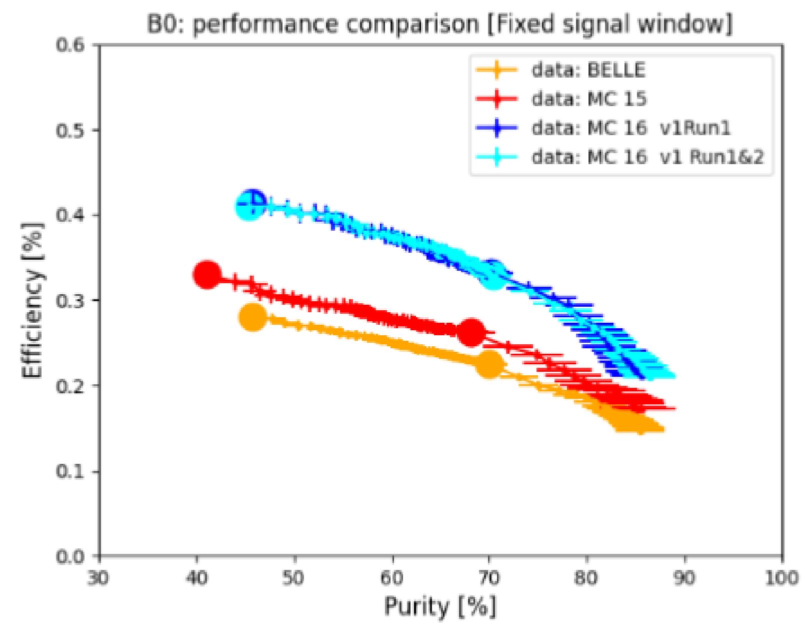
B^0:

- Purity: from 33% to 37%   Efficiency: from 0.38% to 0.47%

B^+:

- Purity: from 34% to 39%   Efficiency: from 1.06% to 1.54%

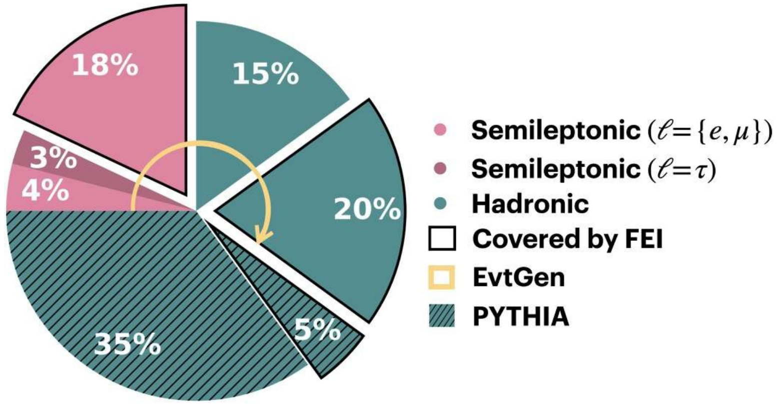




– improvement is clear … and already available for coming results
– now finally better than Belle
– run 2 seems to be of comparable quality

Summary

- knowledge of hadronic B-decays is essential for B-tagging
- a large part (50%) of the hadronic B decays not measured…
  …and PYTHIA is generating something…



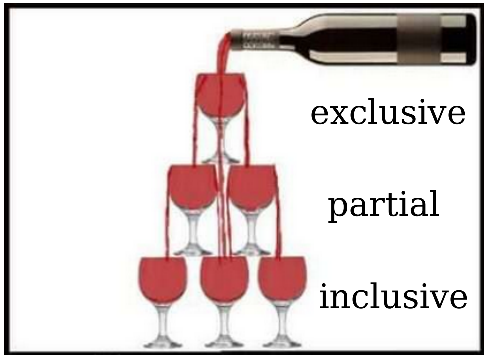
- **clear improvement of hadronic B-tagging in newly reprocessed data thanks to long term efforts on hadronic B decays modeling**

Nice perspectives for using run 1 + run 2 + Belle with hadronic B-tagging for missing energy modes searches

Soon more results with SL B-tagging (especially for rare B decays)
(see talk of Corentin Santos/Merna Abumusabh for $B \to K \nu \bar{\nu}$)

… and inclusive B-tagging

… and ultimately combination of B-taggings



SuperKEKB, the first new collider in particle physics since the LHC in 2008 (electron-positron (e^+e^-) rather than proton-proton (p-p))

Phase 1

Background, Optics commissioning

Feb - June **2016**

Brand new 3 km positron ring

Phase 2: Pilot run

Superconducting Final Focus

add positron damping ring

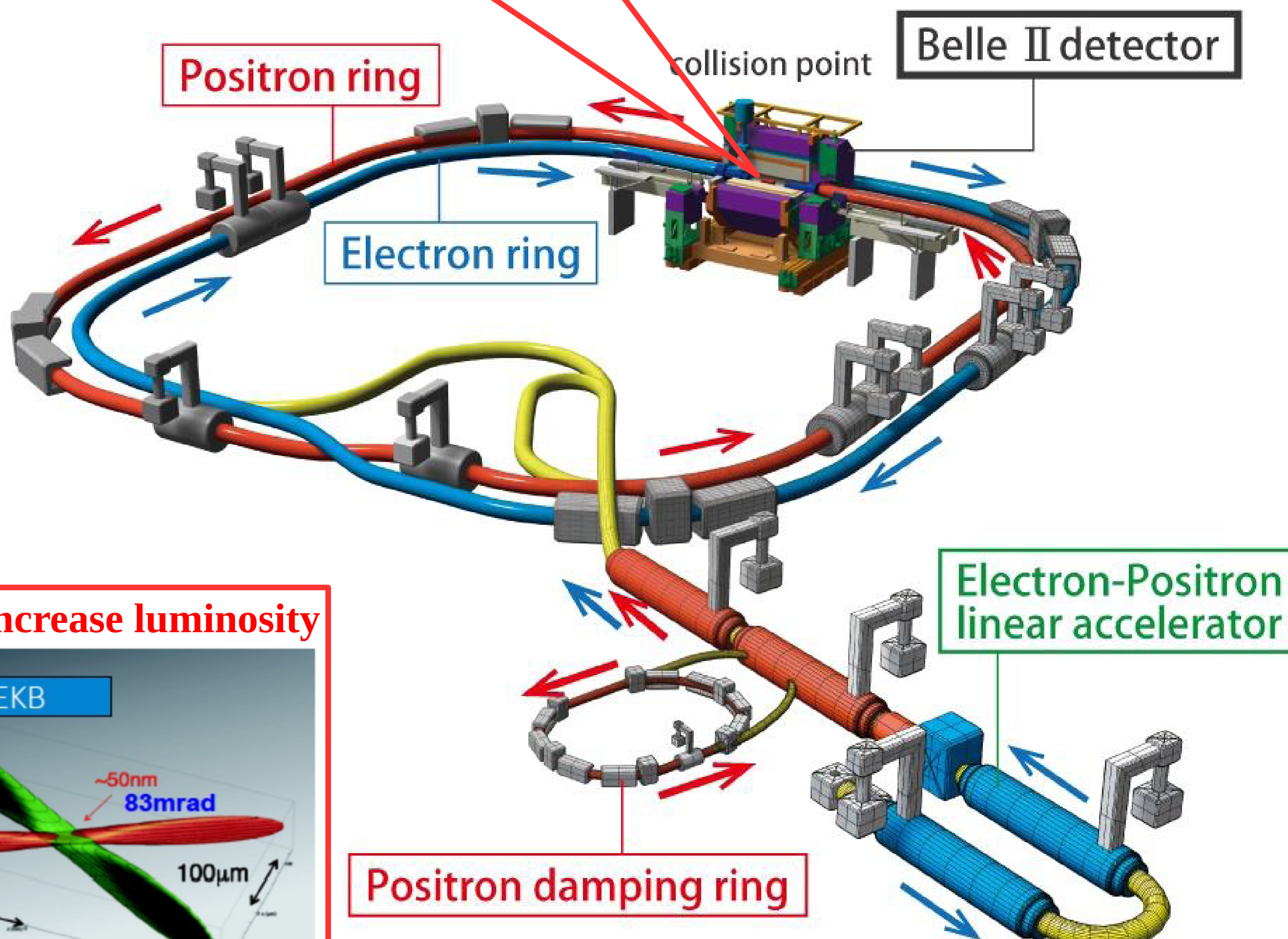
First Collisions ($0.5\ fb^{-1}$)

April 27 - July 17, **2018**

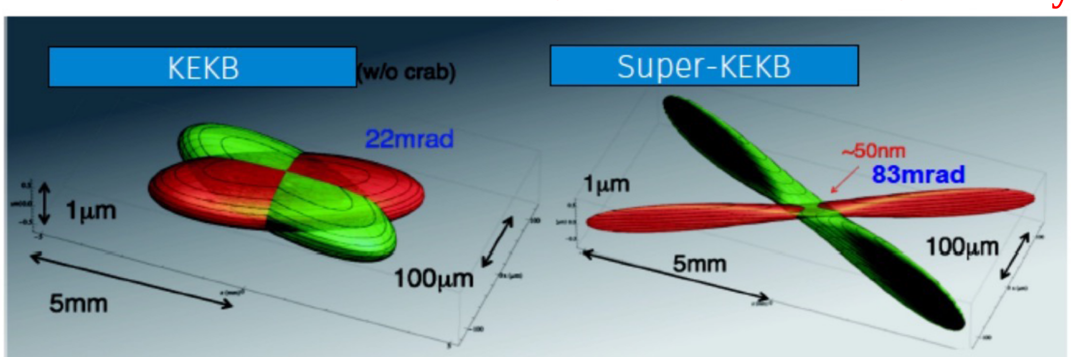
Phase 3: Physics run

Since April, **2019**

Squeeze strongly at IP



Nano-beams and more beam current to increase luminosity



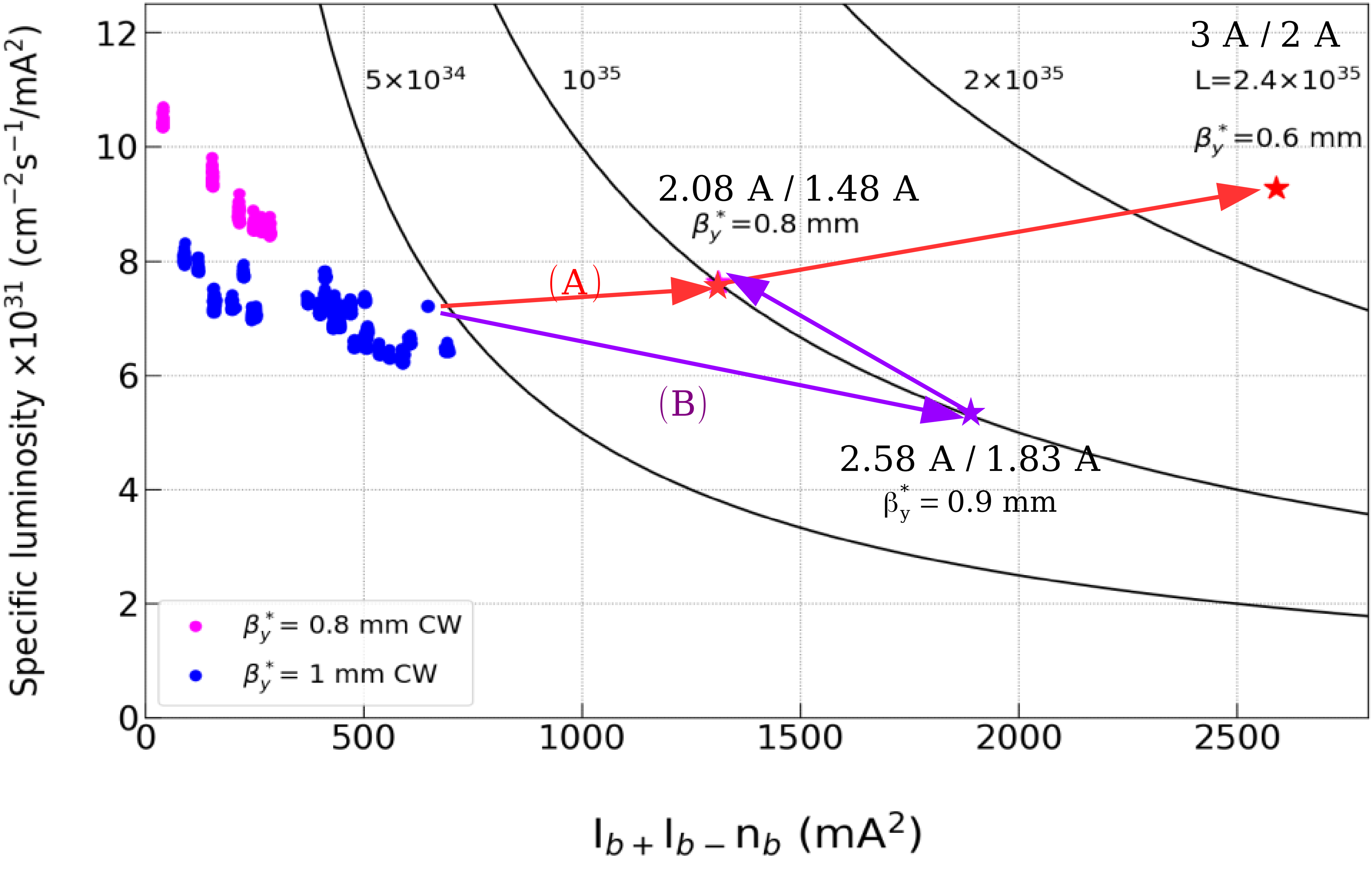
	E (GeV) LER/HER	β^*_y (mm) LER/HER	β^*_x (cm) LER/HER	φ (mrad)	I (A) LER/HER	L ($cm^{-2}s^{-1}$)
KEKB	3.5/8.0	5.9/5.9	120/120	11	1.6/1.2	2.1 x 10^{34}
SuperKEKB	4.0/7.0	0.27/0.30	3.2/2.5	41.5	3.6/2.6	80 x 10^{34}

factor 20 factor 2-3

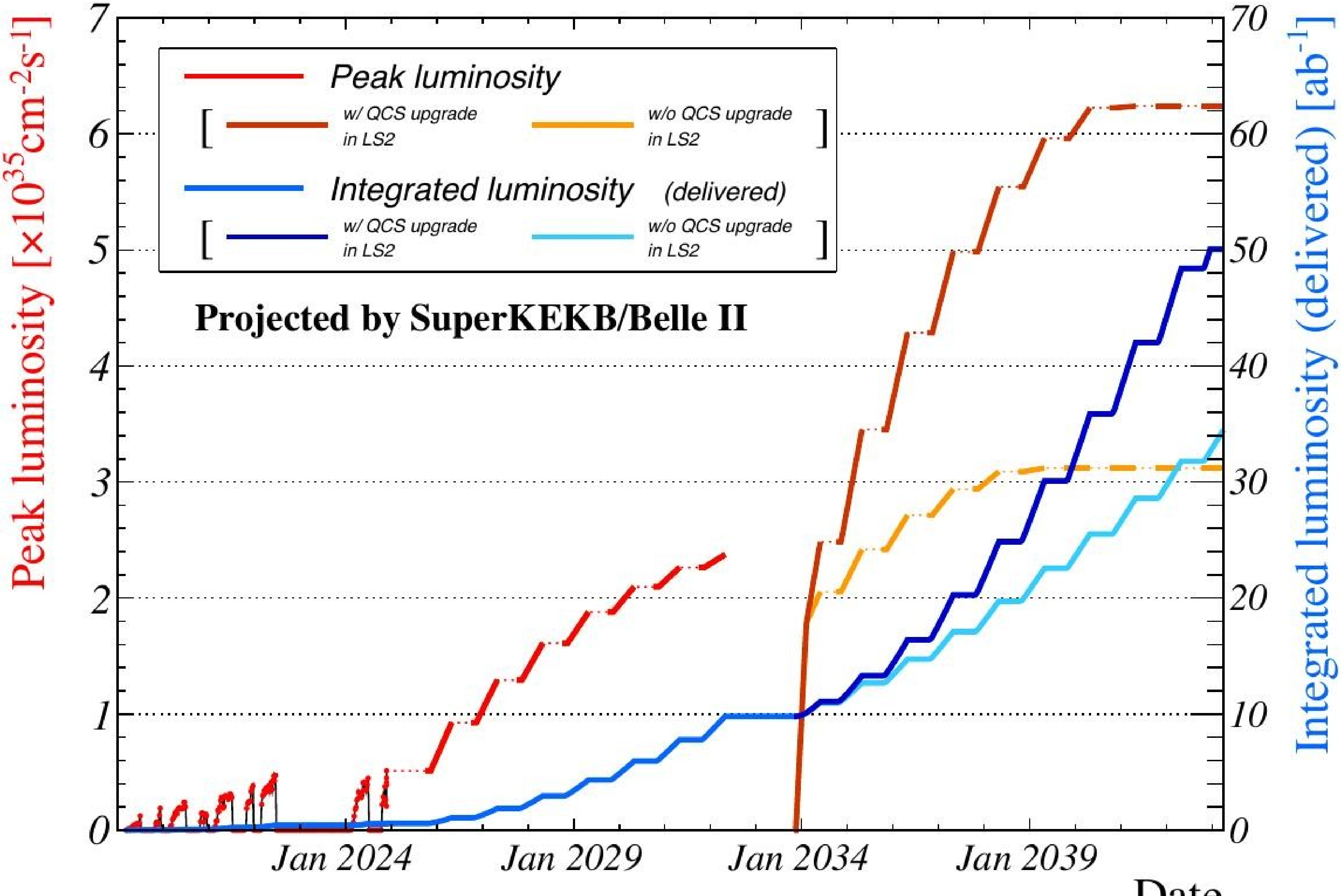
⇒ to reach ~ $6\times10^{35}\ cm^{-2}s^{-1}$

⇒ cumulate 50 ab^{-1} by ~ 2040

Strategy toward > 10^{35} $cm^{-2}s^{-1}$

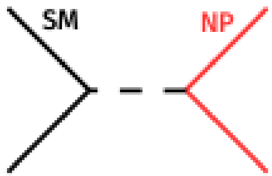


Luminosity projection plot
(plan for the coming years)



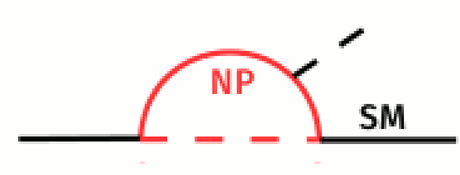
Rare/Forbidden B decays

energy frontier



limited by beam energy

intensity frontier

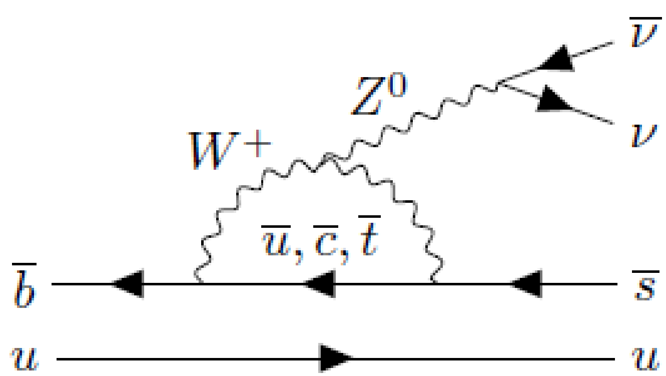


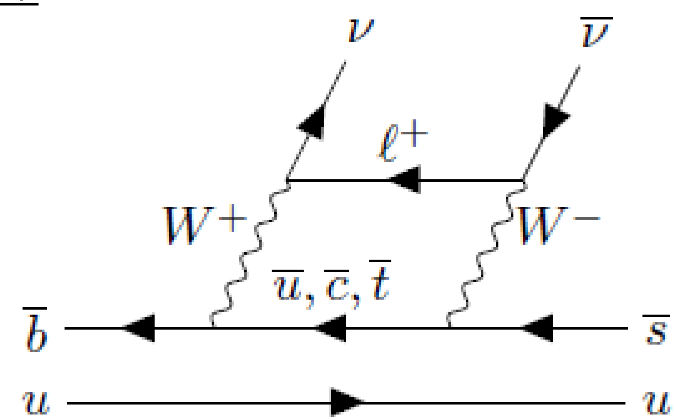
limited by statistics

➔ NP beyond the direct reach of the LHC

New particles can for example contribute to loop or tree level diagrams
by enhancing/suppressing decay rates, introducing new sources of CP violation or modifying the angular distribution of the final-state particles

$B \to K \nu \bar{\nu}$



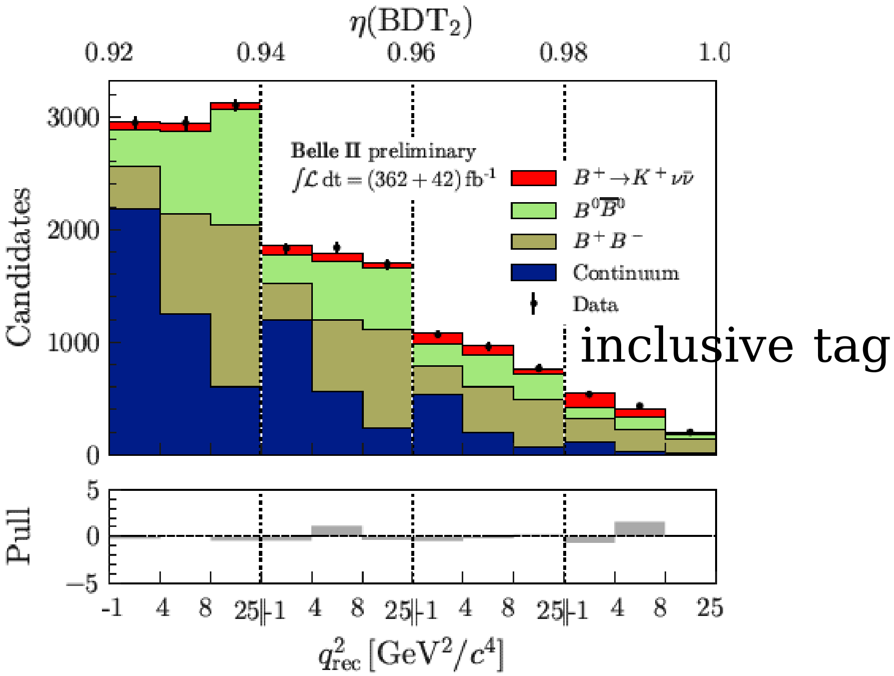


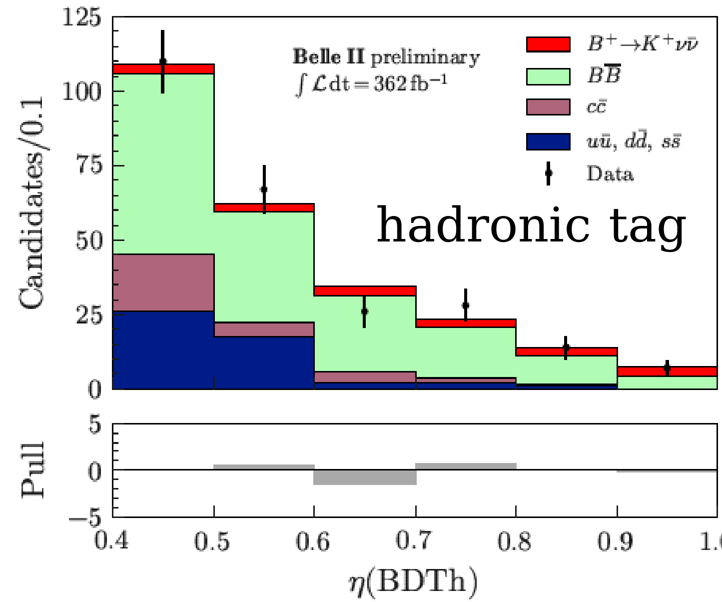
- $B \to K \nu \nu$ is known with high accuracy
  $B(B \to K \nu \nu) = (5.6 \pm 0.4) \times 10^{-6}$ [arXiv:2207.13371]
- Extensions beyond SM may lead to significant rate increase
- Very challenging experimentally, not yet observed
  - Low branching fraction, high background contributions
  - 3-body kinematics, no good kinematics
- Unique for Belle II

Evidence of $B \to K \nu \bar{\nu}$

[arXiv:2311.14647]
PRD109, 112006 (2024)

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





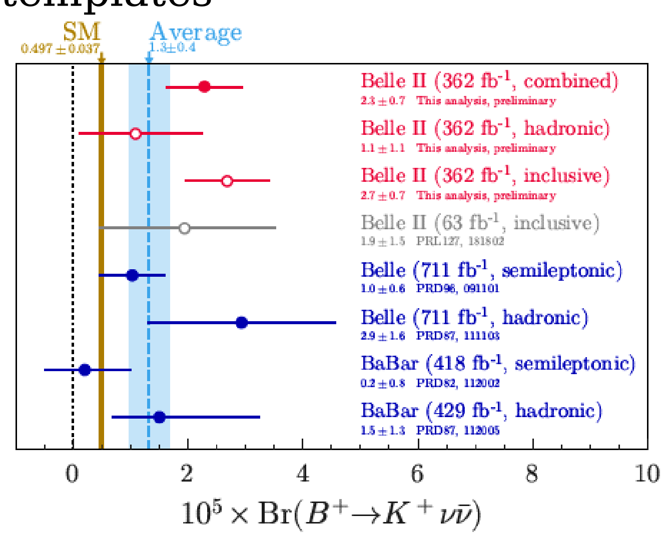
- Maximum likelihood fit to data using signal and background templates

$$\mathbf{B_{incl} = (2.7 \pm 0.5\,(stat) \pm 0.5\,(syst)) \times 10^{-5}}$$

$$\mathbf{B_{had} = (1.1\,^{+0.9}_{-0.8}\,(stat)\,^{+0.8}_{-0.5}\,(syst)) \times 10^{-5}}$$

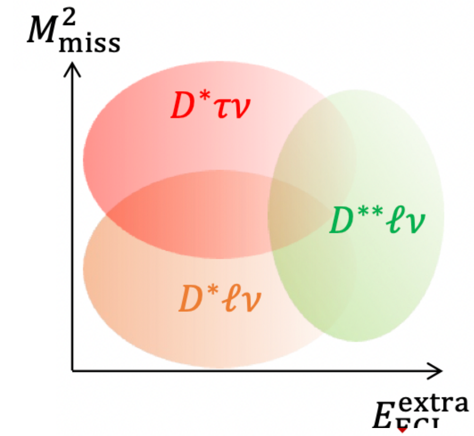
- For inclusive analysis, evidence for $B \to K \nu \bar{\nu}$ at 3.5σ branching fraction within $\mathbf{3\sigma}$ of SM
- For hadronic tag, the result is consistent with null hypothesis and SM at 1.1σ and 0.6σ

$\Rightarrow$ Combination of two analyses provides first evidence of the decay at 2.7σ from SM



First Belle II result on $B(B\to D^* \tau\nu)/B(B\to D^* l\nu)$

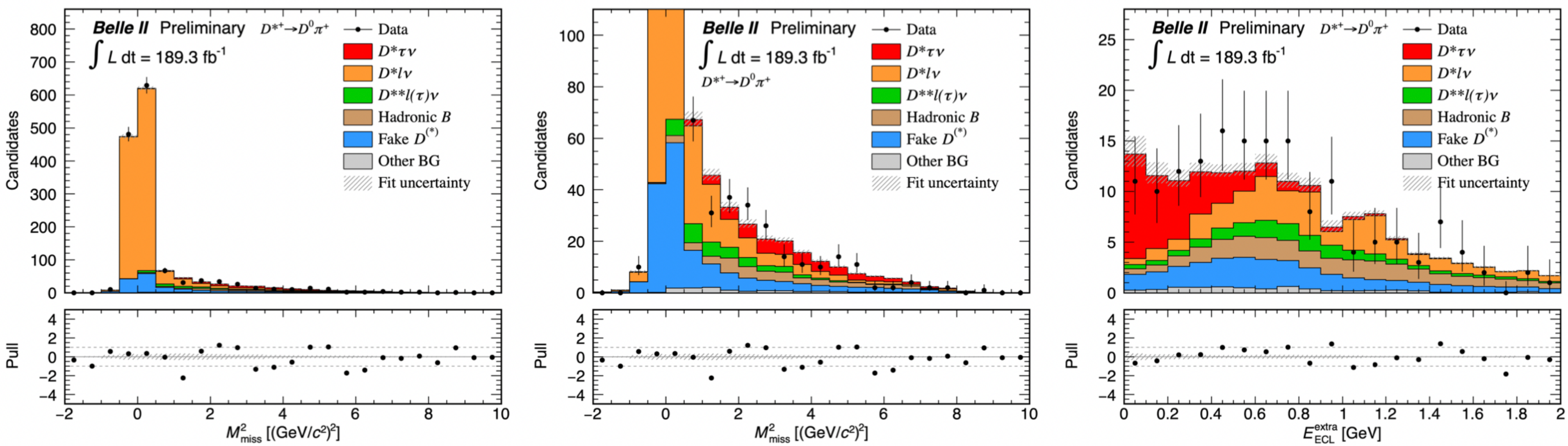
- Half of available sample (200 million $B\bar{B}$ pairs)
- Fully reconstruct the partner B in the event to suppress bckg
- Reconstruct numerator and denominator with ~ same selections
- Two-dimensional fit of missing mass and total residual energy in calorimeter determines signal yields
- Data sidebands validate understanding of sample composition



Post-fit distributions for $D^{*+}\to D^0\pi^+$

$$M^2_{miss} = (p_B - p_{X_c} - p_\mu)^2$$

$1.5 < M^2_{miss} < 6.0\ (\mathrm{GeV}/c^2)^2$



$$R(D^*) = 0.267\ ^{+0.041}_{-0.039}(\mathrm{stat})\ ^{+0.028}_{-0.033}(\mathrm{syst})$$

Not leading, 40% improvement in statistical precision over Belle at the same sample size
Consistent with WA

$R_{\tau/\ell}(D^*)$-$R_{\tau/\ell}(D)$ H-tag leptonic τ from Belle II using 365 fb^{-1} (Results)

NEW!

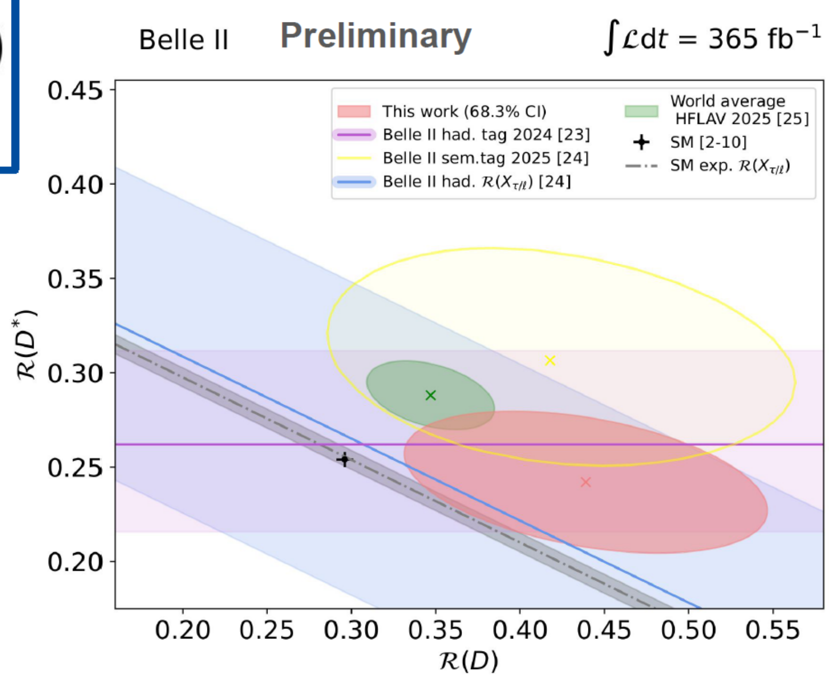
Preliminary

$$R(D^*) = 0.242 \pm 0.019\,(\text{stat.}) \pm 0.016\,(\text{syst.})$$
$$R(D) = 0.439 \pm 0.055\,(\text{stat.}) \pm 0.046\,(\text{syst.})$$
$$\rho = -0.40\,(\text{stat.}) - 0.20\,(\text{syst.})$$

Source	$R(D^*)$	$R(D)$	ρ
Simulation sample size	4.8%	8.4%	−0.44
gap-mode branching fraction	2.7%	3.1%	0.00
$B \to D^{**}\tau^-/(\ell^-)\nu_\ell$ branching fractions	0.3%	1.3%	0.25
Hadronic B decay branching fractions	1.6%	1.5%	−0.26
Form factors	0.5%	0.9%	−0.70
Fraction of misreconstructed $D^{(*)}$	0.5%	1.2%	0.00
Continuum background	2.4%	2.1%	0.93
Fit biases	0.9%	1.2%	0.00
Low-momentum π^0, γ efficiency	2.2%	2.4%	0.99
Other efficiency corrections	0.7%	1.4%	0.92
B-tagging efficiency of data	0.9%	1.8%	−1.00
B-tagging efficiency of $B \to D\tau\nu$	0.1%	1.8%	1.00
M^2_{miss} resolution	0.5%	0.8%	0.48
Total systematic uncertainty	6.7%	10.3%	−0.20
Statistical uncertainty	8.3%	16.3%	−0.40

Systematics determined with toy pseudoexperiments varying pdf with normal /uniform sampling

Most precise $R(D^*)$-$R(D)$ measurement using H-tag

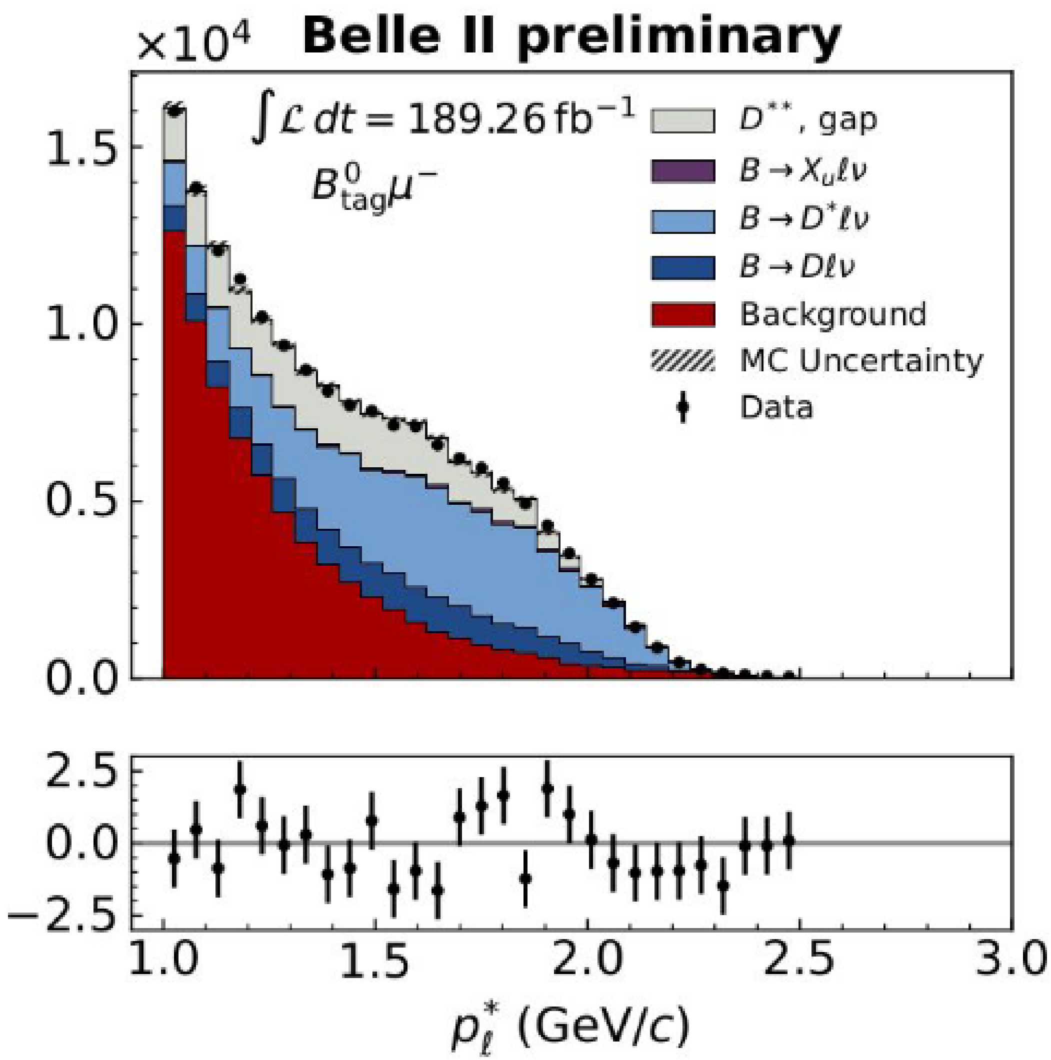


In agreement with SM within 1.5σ
0.5σ for $R(D^*)$ and 2σ for $R(D)$
In agreement with the world average within 1.3σ

19

B^+-tagging: standard calibration sample

BDTs are trained on **MC**
⇒ The performance has to be calibrated with data.



Traditionally, this calibration is done with semi-leptonic B on the signal side because it has large branching fraction.

Fit the lepton momentum in B rest frame.
No clear peak
⇒ Complex template fitting strategy
⇒ Low signal-side purity

Systematically limited

- Highly dependent on the SL decay model including D** and SL gap components
- Significant cross-feed from B^0

But, if MC is not optimal, the BDT selection will not be optimal.

This cannot be easily studied with semi-leptonic B because there are no peaking structures.

An orthogonal sample is needed not only to provide calibration factors but to study the sources of discrepancy.

True lies and hard truths

(summarized by Peter Lewis)

In analysis

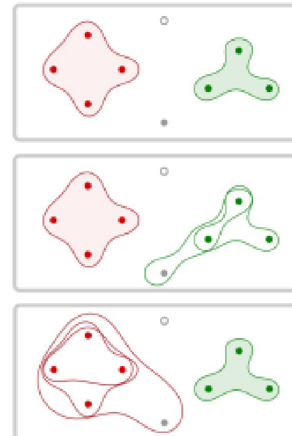
So you have your tags… *what do you do now?*

1. Build your *signal-side B* candidate
2. Combine *tag* and *signal Bs* to make $\Upsilon(4S)$ candidates

FEI efficiency ϵ_{FEI} enters in one of two ways:

Lies!

- BF(signal): $\epsilon_{total} = \epsilon_{FEI} * \epsilon_{sig}$
- BF(signal)/BF(normalization): FEI efficiency *cancels*



But FEI is trained on MC: ϵ_{FEI} needs a *calibration…*

20

We now know that it is not possible to disentangle sig/tag efficiency, so a calibration may only be valid for the mode it is calibrated on (!)

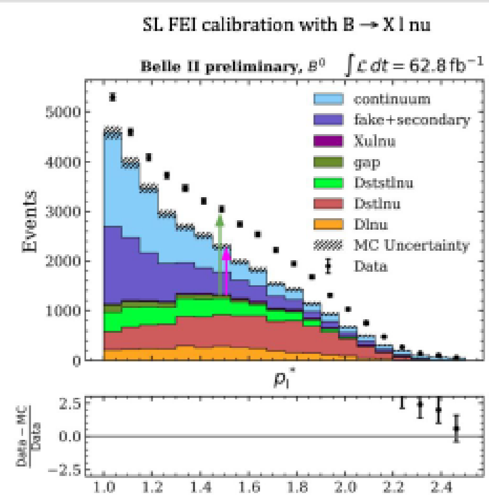
Calibration

Measure *known decay* in FEI-tagged events

Lie → $N_{\text{expected}} = N_{\Upsilon(4S)} \epsilon_{\text{FEI}} \cdot \mathcal{B}\epsilon$

Calibration factors are **data**/**MC** ratios of measured yield

Oversimplification → To use: correct simulated FEI efficiency with the calibration factor



◦ having several calibration procedure (learn a lot about signal-side dependencies)

◦ the closer the calibration factors are from 1, the better is our MC (so is the cross-feed simulation, the signal-side dependencies…)

FEI metrics in data

Calculated directly on data:

- Calibration factor

$$= \frac{\text{Signal yield in data}}{\text{Signal yield in MC}}$$

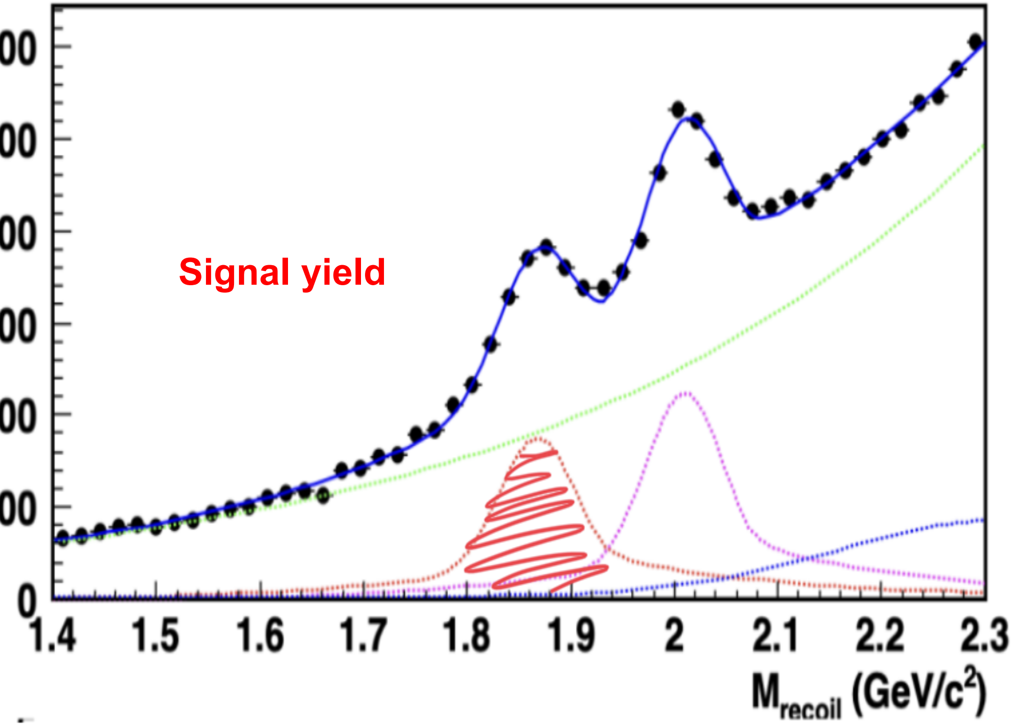
- Purity

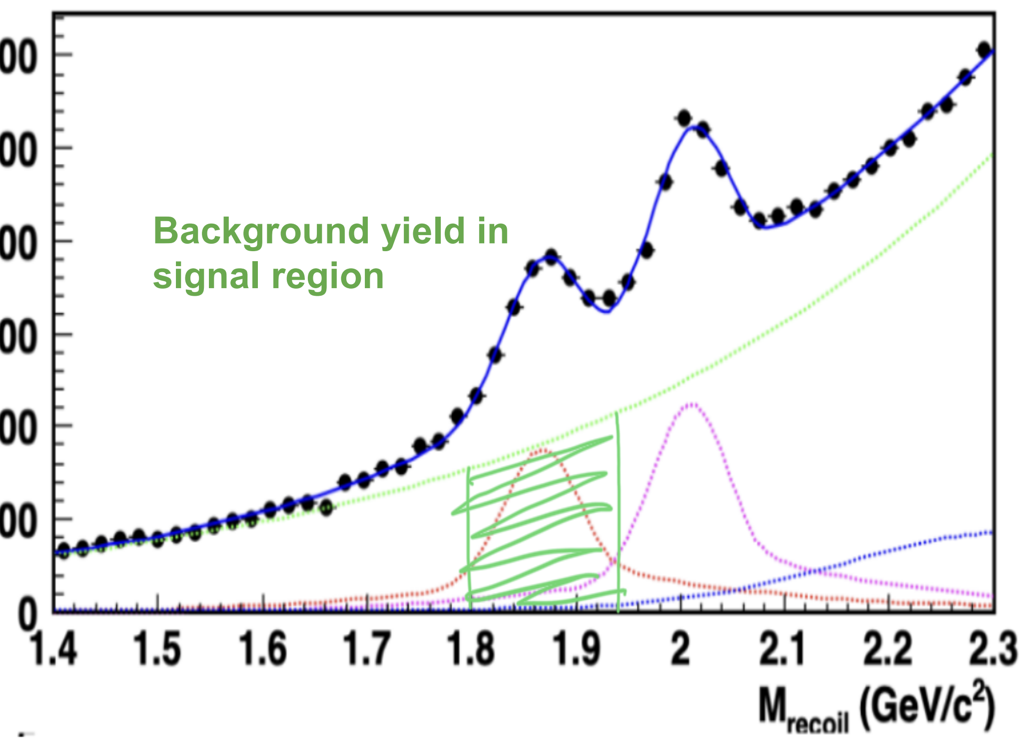
$$= \frac{\text{Signal yield}}{\text{Signal yield + Background yield in signal region}}$$

- Efficiency

$$= \frac{\text{Signal yield}}{n_{BB} \quad BF_{B+ \to D\pi} \quad \epsilon_{\pi}}$$

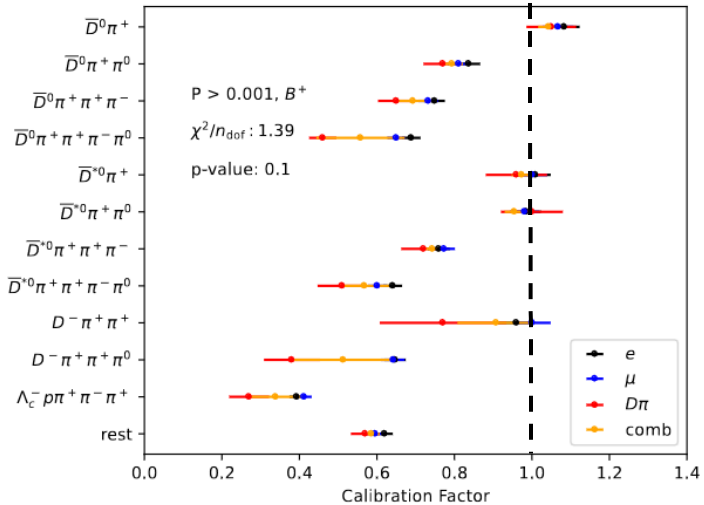
n_{BB} = 392.5 x 10^6
$BF_{B+ \to D\pi}$ = 0.467 x10^{-2}
ϵ_{π} = 90%





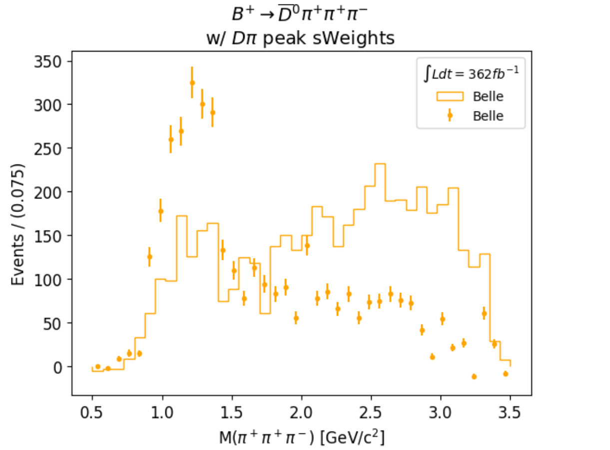
But why calibration factors are still far from 1 ?

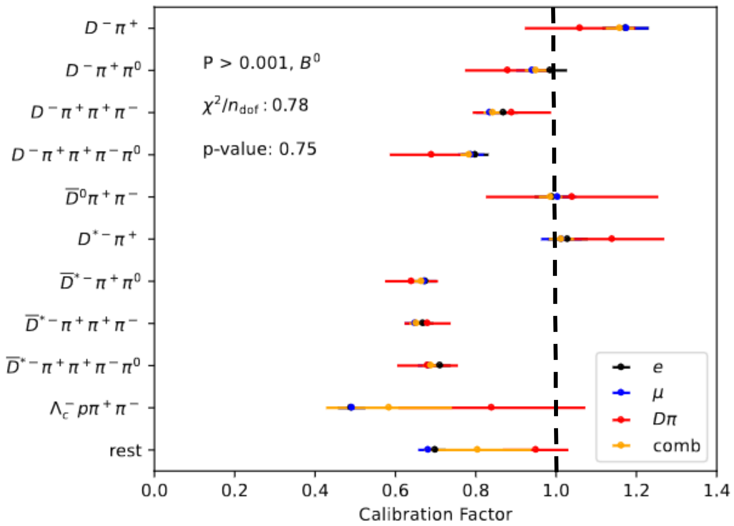
- The fit allows to obtain calibration factors but also thanks to splot, obtain the distributions for B_{tag} decays: invariant mass of intermediate states, sigprob…

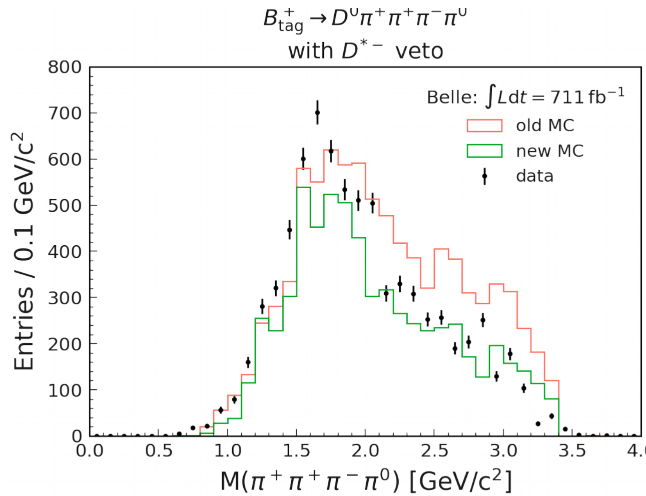


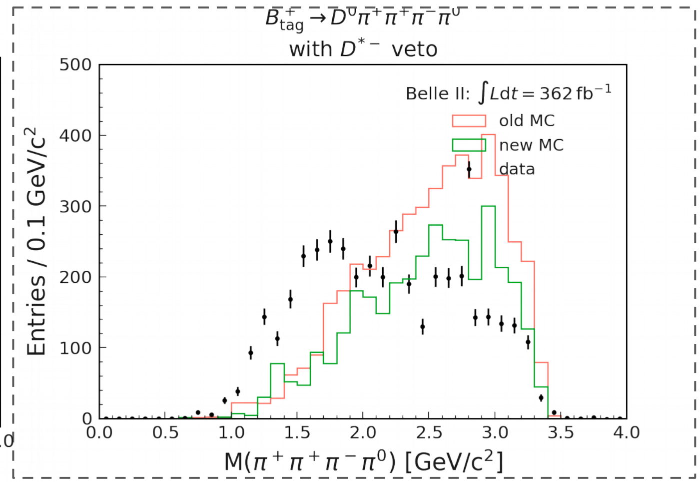
Understanding of $B \to D^{(*)} n\pi m\pi^0$ decays is essential for B-tagging…

$\Rightarrow$ significant differences between data and MC (e.g. $n+m \geq 3$)









Belle and Belle II have different PYTHIA.
But the distribution is different in the data itself
$\Rightarrow$ bias introduced by training on MC !!

$\Rightarrow$ Need to understand and improve the MC modeling of B decays

Improving MC model: an example

Let's take one final state for example: $B^+ \to \bar{D}^0 \pi^+ \pi^+ \pi^-$.
It can be produced through many intermediate states:

Decay	Belle	Belle II
$B^+ \to \overline{D}^0 \pi^-\pi^+\pi^+$	0.46	0.51
$B^+ \to \overline{D}^0 \rho(770)^0\pi^+; \rho(770)^0 \to \pi^+\pi^-$	0.39	0.42
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to \rho(770)^0\pi^+; \rho(770)^0 \to \pi^+\pi^-$	0.13	0.14
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to f_0(600)\pi^+; f_0(600) \to \pi^+\pi^-$	0.05	-
$B^+ \to \overline{D}_1(2420)^0\pi^+; \overline{D}_1(2420)^0 \to D^*(2010)^-\pi^+; D^*(2010)^- \to \overline{D}^0\pi^-$	0.04	0.02
$B^+ \to \overline{D}_1(2430)^0\pi^+; \overline{D}_1(2430)^0 \to D^*(2010)^-\pi^+; D^*(2010)^- \to \overline{D}^0\pi^-$	0.03	0.02
$B^+ \to \overline{D}_2^*(2460)^0\pi^+; \overline{D}_2^*(2460)^0 \to D^*(2010)^-\pi^+; D^*(2010)^- \to \overline{D}^0\pi^-$	0.01	0.01
$B^+ \to D^*(2010)^-\pi^+\pi^+; D^*(2010)^- \to \overline{D}^0\pi^-$	-	0.09
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to \pi^+\pi^+\pi^-$	-	0.07
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to f_0(500)\pi^+; f_0(500) \to \pi^+\pi^-$	-	0.05
$B^+ \to \overline{D}_1(2420)^0\pi^+; \overline{D}_1(2420)^0 \to \overline{D}^0\pi^-\pi^+$	-	0.02
$B^+ \to \overline{D}^0 K^*(892)^+; K^*(892)^+ \to K^0\pi^+; K^0 \to K_S^0; K_S^0 \to \pi^+\pi^-$	-	0.01
Rest of Exclusive	0.03	0.03
Sum of Exclusive	1.12	1.38
Sum of Pythia	0	0
Total Sum	1.12	1.38

The $\pi^+ \pi^+ \pi^-$ could be directly generated, could come through $\rho^0\pi^+$ or through an intermediate a_1^+ resonance.

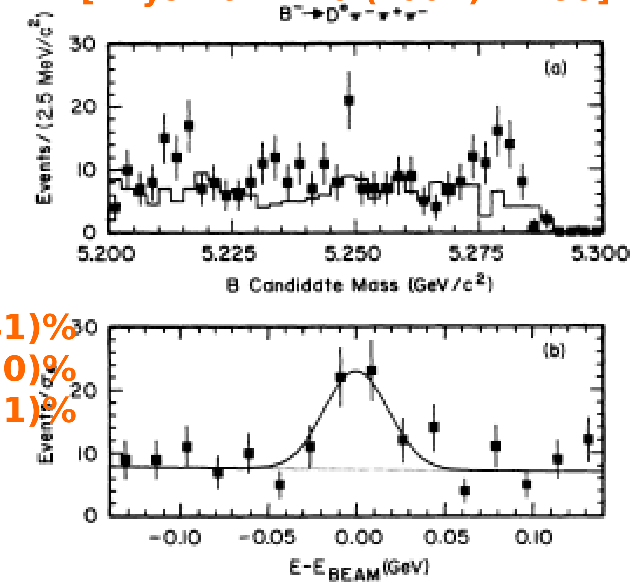
Improving MC model: an example

Let's take one final state for example: $B^+ \to \overline{D}^0 \pi^+ \pi^+ \pi^-$.
It can be produced through many intermediate states:

Decay	Belle	Belle II
$B^+ \to \overline{D}^0 \pi^- \pi^+ \pi^+$	0.46	0.51
$B^+ \to \overline{D}^0 \rho(770)^0 \pi^+; \rho(770)^0 \to \pi^+ \pi^-$	0.39	0.42
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to \rho(770)^0 \pi^+; \rho(770)^0 \to \pi^+ \pi^-$	0.13	0.14
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to f_0(600) \pi^+; f_0(600) \to \pi^+ \pi^-$	0.05	-
$B^+ \to \overline{D}_1(2420)^0 \pi^+; \overline{D}_1(2420)^0 \to D^*(2010)^- \pi^+; D^*(2010)^- \to \overline{D}^0 \pi^-$	0.04	0.02
$B^+ \to \overline{D}_1(2430)^0 \pi^+; \overline{D}_1(2430)^0 \to D^*(2010)^- \pi^+; D^*(2010)^- \to \overline{D}^0 \pi^-$	0.03	0.02
$B^+ \to \overline{D}_2^*(2460)^0 \pi^+; \overline{D}_2^*(2460)^0 \to D^*(2010)^- \pi^+; D^*(2010)^- \to \overline{D}^0 \pi^-$	0.01	0.01
$B^+ \to D^*(2010)^- \pi^+ \pi^+; D^*(2010)^- \to \overline{D}^0 \pi^-$	-	0.09
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to \pi^+ \pi^+ \pi^-$	-	0.07
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to f_0(500) \pi^+; f_0(500) \to \pi^+ \pi^-$	-	0.05
$B^+ \to \overline{D}_1(2420)^0 \pi^+; \overline{D}_1(2420)^0 \to \overline{D}^0 \pi^- \pi^+$	-	0.02
$B^+ \to \overline{D}^0 K^*(892)^+; K^*(892)^+ \to K^0 \pi^+; K^0 \to K_S^0; K_S^0 \to \pi^+ \pi^-$	-	0.01
Rest of Exclusive	0.03	0.03
Sum of Exclusive	1.12	1.38
Sum of Pythia	0	0
Total Sum	1.12	1.38

(0.51 ± 0.41)%
(0.42 ± 0.30)%
(0.14 ± 0.11)%

[Phys.Rev.D 45 (1992) 21-35]



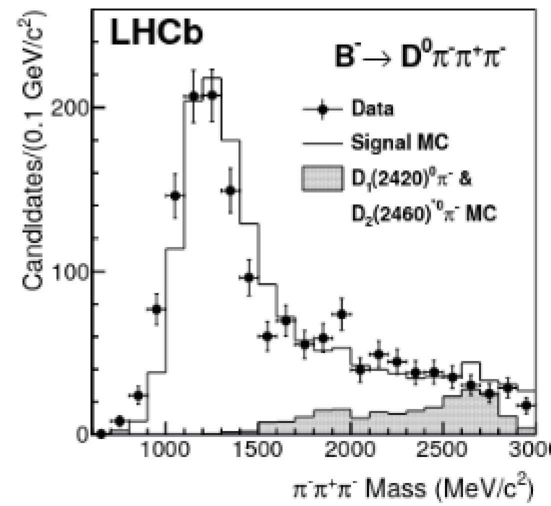
In 1992, CLEO experiment measured these 3 values but with ~75% uncertainty!

Improving MC model: an example

Let's take one final state for example: $B^+ \to \overline{D}^0 \pi^+ \pi^+ \pi^-$.
It can be produced through many intermediate states:

Decay	Belle	Belle II
$B^+ \to \overline{D}^0 \pi^- \pi^+ \pi^+$	0.46	0.51
$B^+ \to \overline{D}^0 \rho(770)^0 \pi^+; \rho(770)^0 \to \pi^+ \pi^-$	0.39	0.42
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to \rho(770)^0 \pi^+; \rho(770)^0 \to \pi^+ \pi^-$	0.13	0.14
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to f_0(600) \pi^+; f_0(600) \to \pi^+ \pi^-$	0.05	-
$B^+ \to \overline{D}_1(2420)^0 \pi^+; \overline{D}_1(2420)^0 \to D^*(2010)^- \pi^+; D^*(2010)^- \to \overline{D}^0 \pi^-$	0.04	0.02
$B^+ \to \overline{D}_1(2430)^0 \pi^+; \overline{D}_1(2430)^0 \to D^*(2010)^- \pi^+; D^*(2010)^- \to \overline{D}^0 \pi^-$	0.03	0.02
$B^+ \to \overline{D}_2^*(2460)^0 \pi^+; \overline{D}_2^*(2460)^0 \to D^*(2010)^- \pi^+; D^*(2010)^- \to \overline{D}^0 \pi^-$	0.01	0.01
$B^+ \to D^*(2010)^- \pi^+ \pi^+; D^*(2010)^- \to \overline{D}^0 \pi^-$	-	0.09
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to \pi^+ \pi^+ \pi^-$	-	0.07
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to f_0(500) \pi^+; f_0(500) \to \pi^+ \pi^-$	-	0.05
$B^+ \to \overline{D}_1(2420)^0 \pi^+; \overline{D}_1(2420)^0 \to \overline{D}^0 \pi^- \pi^+$	-	0.02
$B^+ \to \overline{D}^0 K^*(892)^+; K^*(892)^+ \to K^0 \pi^+; K^0 \to K_S^0; K_S^0 \to \pi^+ \pi^-$	-	0.01
Rest of Exclusive	0.03	0.03
Sum of Exclusive	1.12	1.38
Sum of Pythia	0	0
Total Sum	1.12	1.38

Phys.Rev.D 84 (2011) 092001



In 2011 (~20 years later), LHCb looked at this final state, but did not provide individual measurements.

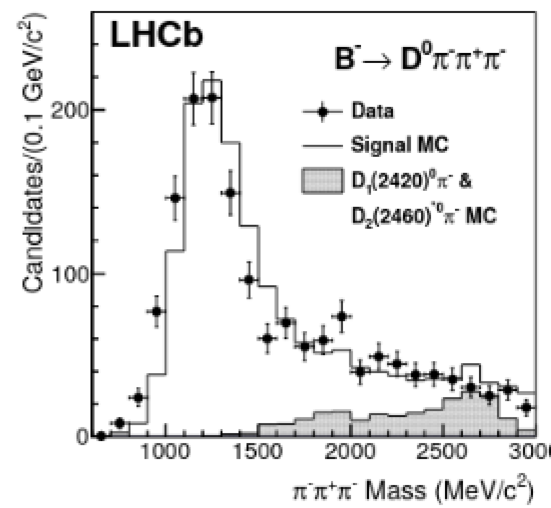
So we are still suck with a 30 year old CLEO measurement in PDG.

Improving MC model: an example

Let's take one final state for example: $B^+ \to \bar{D}^0 \pi^+ \pi^+ \pi^-$.
It can be produced through many intermediate states:

Decay	Belle	Belle II
$B^+ \to \overline{D}^0 \pi^- \pi^+ \pi^+$	0.46	0.51
$B^+ \to \overline{D}^0 \rho(770)^0 \pi^+; \rho(770)^0 \to \pi^+ \pi^-$	0.39	0.42
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to \rho(770)^0 \pi^+; \rho(770)^0 \to \pi^+ \pi^-$	0.13	0.14
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to f_0(600) \pi^+; f_0(600) \to \pi^+ \pi^-$	0.05	-
$B^+ \to \overline{D}_1(2420)^0 \pi^+; \overline{D}_1(2420)^0 \to D^*(2010)^- \pi^+; D^*(2010)^- \to \overline{D}^0 \pi^-$	0.04	0.02
$B^+ \to \overline{D}_1(2430)^0 \pi^+; \overline{D}_1(2430)^0 \to D^*(2010)^- \pi^+; D^*(2010)^- \to \overline{D}^0 \pi^-$	0.03	0.02
$B^+ \to \overline{D}_2^*(2460)^0 \pi^+; \overline{D}_2^*(2460)^0 \to D^*(2010)^- \pi^+; D^*(2010)^- \to \overline{D}^0 \pi^-$	0.01	0.01
$B^+ \to D^*(2010)^- \pi^+ \pi^+; D^*(2010)^- \to \overline{D}^0 \pi^-$	-	0.09
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to \pi^+ \pi^+ \pi^-$	-	0.07
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to f_0(500) \pi^+; f_0(500) \to \pi^+ \pi^-$	-	0.05
$B^+ \to \overline{D}_1(2420)^0 \pi^+; \overline{D}_1(2420)^0 \to \overline{D}^0 \pi^- \pi^+$	-	0.02
$B^+ \to \overline{D}^0 K^*(892)^+; K^*(892)^+ \to K^0 \pi^+; K^0 \to K_S^0; K_S^0 \to \pi^+ \pi^-$	-	0.01
Rest of Exclusive	0.03	0.03
Sum of Exclusive	1.12	1.38
Sum of Pythia	0	0
Total Sum	1.12	1.38

Phys.Rev.D 84 (2011) 092001



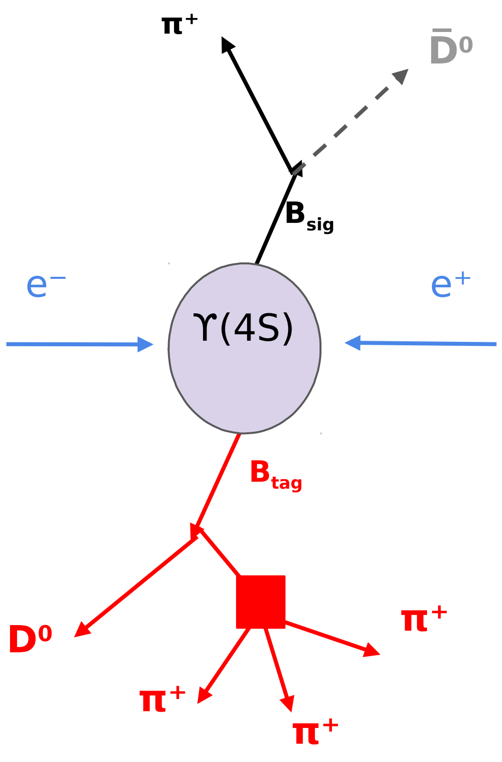
But looking at this plot, it looks like most contribution comes through $a_1{}^+$ resonance (mass 1400 MeV/c^2).

Improving MC model: an example

Let's take one final state for example: $B^+ \to \bar{D}^0 \pi^+ \pi^+ \pi^-$.
It can be produced through many intermediate states:

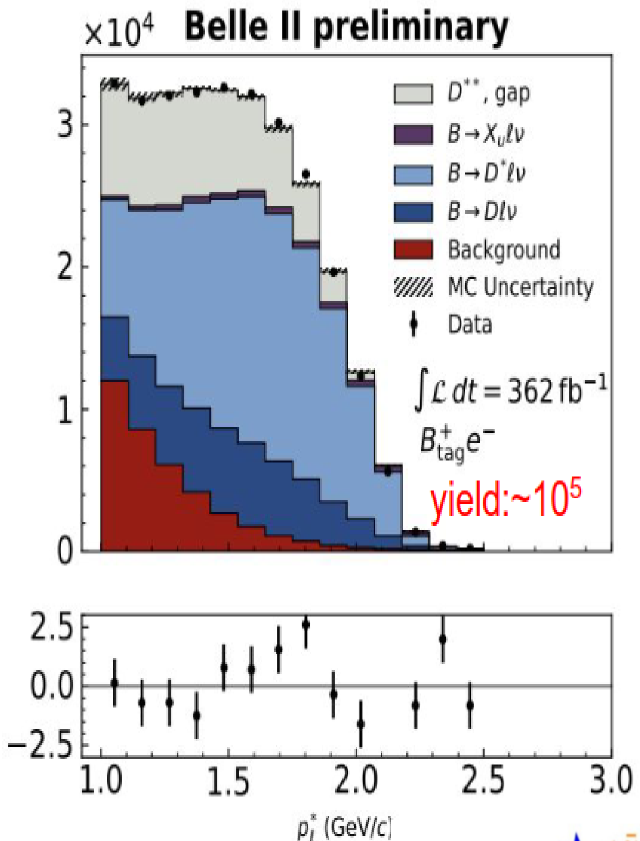
Decay	Belle	Belle II
$B^+ \to \overline{D}^0 \pi^- \pi^+ \pi^+$	0.46	0.51
$B^+ \to \overline{D}^0 \rho(770)^0 \pi^+; \rho(770)^0 \to \pi^+\pi^-$	0.39	0.42
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to \rho(770)^0\pi^+; \rho(770)^0 \to \pi^+\pi^-$	0.13	0.14
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to f_0(600)\pi^+; f_0(600) \to \pi^+\pi^-$	0.05	-
$B^+ \to \overline{D}_1(2420)^0\pi^+; \overline{D}_1(2420)^0 \to D^*(2010)^-\pi^+; D^*(2010)^- \to \overline{D}^0\pi^-$	0.04	0.02
$B^+ \to \overline{D}_1(2430)^0\pi^+; \overline{D}_1(2430)^0 \to D^*(2010)^-\pi^+; D^*(2010)^- \to \overline{D}^0\pi^-$	0.03	0.02
$B^+ \to \overline{D}_2^*(2460)^0\pi^+; \overline{D}_2^*(2460)^0 \to D^*(2010)^-\pi^+; D^*(2010)^- \to \overline{D}^0\pi^-$	0.01	0.01
$B^+ \to D^*(2010)^-\pi^+\pi^+; D^*(2010)^- \to \overline{D}^0\pi^-$	-	0.09
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to \pi^+\pi^+\pi^-$	-	0.07
$B^+ \to \overline{D}^0 a_1(1260)^+; a_1(1260)^+ \to f_0(500)\pi^+; f_0(500) \to \pi^+\pi^-$	-	0.05
$B^+ \to \overline{D}_1(2420)^0\pi^+; \overline{D}_1(2420)^0 \to \overline{D}^0\pi^-\pi^+$	-	0.02
$B^+ \to \overline{D}^0 K^*(892)^+; K^*(892)^+ \to K^0\pi^+; K^0 \to K_S^0; K_S^0 \to \pi^+\pi^-$	-	0.01
Rest of Exclusive	0.03	0.03
Sum of Exclusive	1.12	1.38
Sum of Pythia	0	0
Total Sum	1.12	1.38

Can be compared with data at Belle, if we reconstruct one B as $B^+ \to \bar{D}^0 \pi^+$ and other B as $B^- \to D^0 \pi^+ \pi^+ \pi^-$

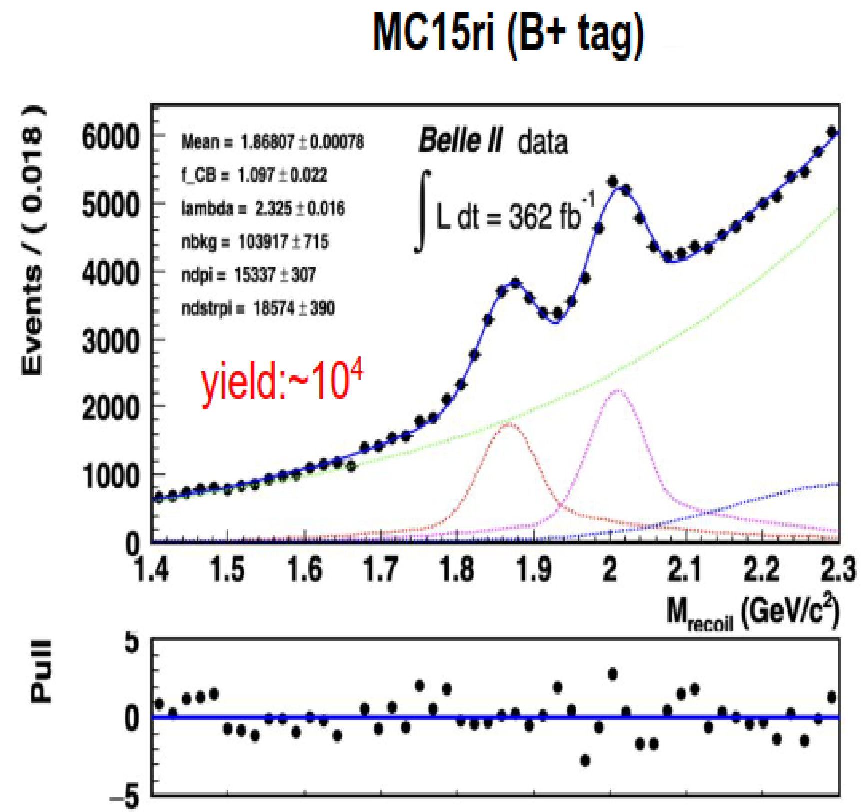


Had FEI calibration with Xlν and Dπ samples

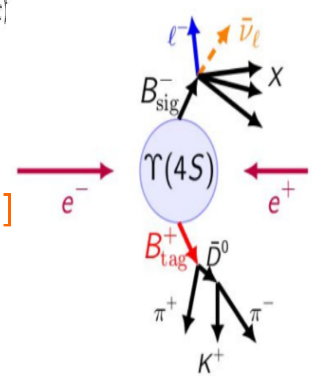
Xℓν sample:
High statistics, low purity



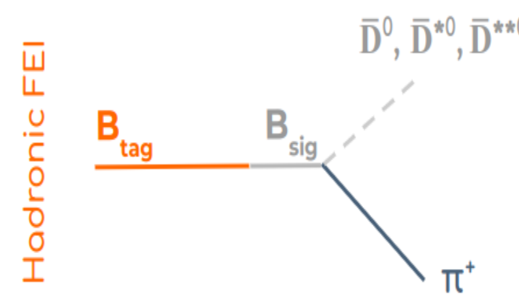
Dπ sample:
Low statistics, high purity



[Florian, William, Daniel Jacobi: BELLE2-NOTE-PH-2023-008]

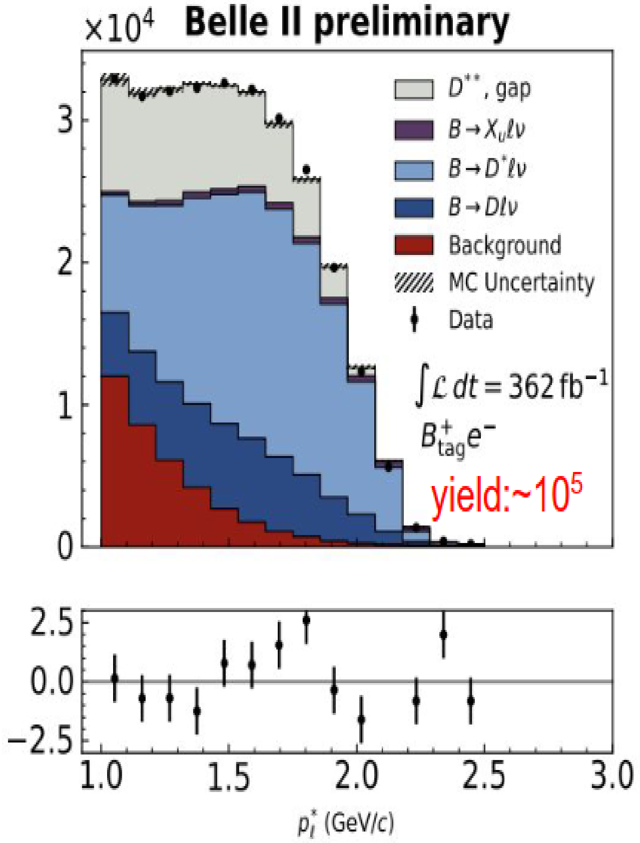


[Karim, Meihong, Niharika, Vidya: BELLE2-NOTE-PH-2023-004]

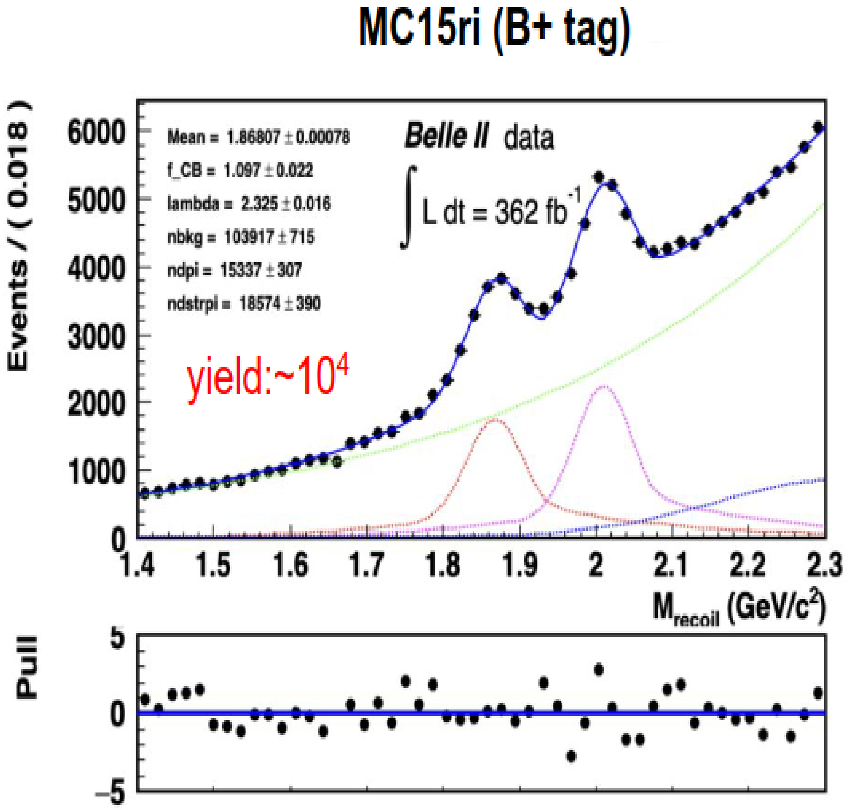


Had FEI calibration with $Xl\nu$ and $D\pi$ samples

$X\ell\nu$ sample:
High statistics, low purity



$D\pi$ sample:
Low statistics, high purity



Calibration factors are calculated from signal yields i.e., correctly-reconstructed B_{tag}.
Hence, applicable on Signal MC.

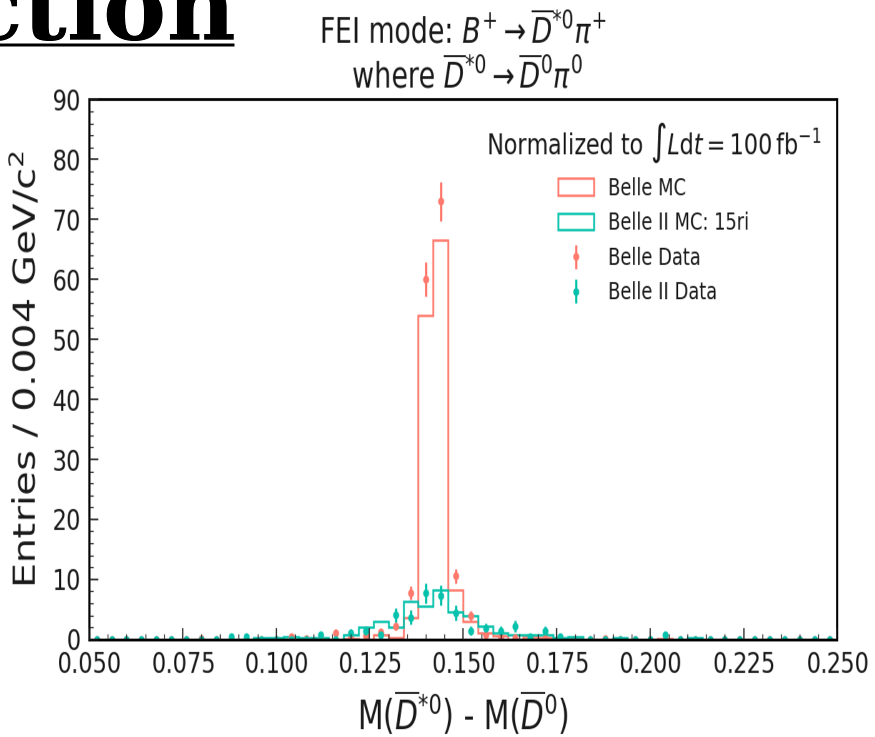
$D^{*0} \rightarrow D^0 \pi^0$ reconstruction

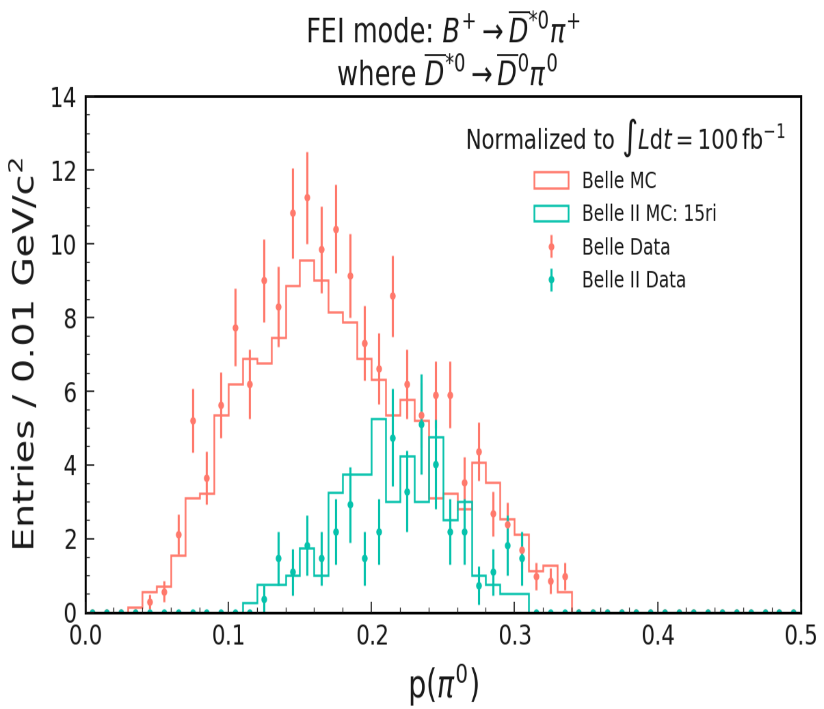
In Belle II, the yield of $D^{*0} \rightarrow D^0 \pi^0$ is much worse than Belle.

E > 0.09 GeV cut for γ is too tight for slow π^0
Should be loosened.

```
if convertedFromBelle:
    gamma_cut = 'goodBelleGamma == 1 and clusterBelleQuality == 0'
else:
    gamma_cut = '[[clusterReg == 1 and E > 0.10] or [clusterReg == 2 and E > 0.09] or [clusterReg == 3 and E > 0.16]]'
if specific:
    gamma_cut += ' and isInRestOfEvent > 0.5'

gamma = Particle('gamma',
                 MVAConfiguration(variables=['clusterReg', 'clusterNHits', 'clusterTiming', 'extraInfo(preCut_rank)',
                                             'clusterE9E25', 'pt', 'E', 'pz'],
                                  target='isPrimarySignal'),
                 PreCutConfiguration(userCut=gamma_cut,
                                     bestCandidateMode='highest',
                                     bestCandidateVariable='E',
                                     bestCandidateCut=40),
                 PostCutConfiguration(bestCandidateCut=20, value=0.01))
gamma.addChannel(['gamma:FSP'])
```

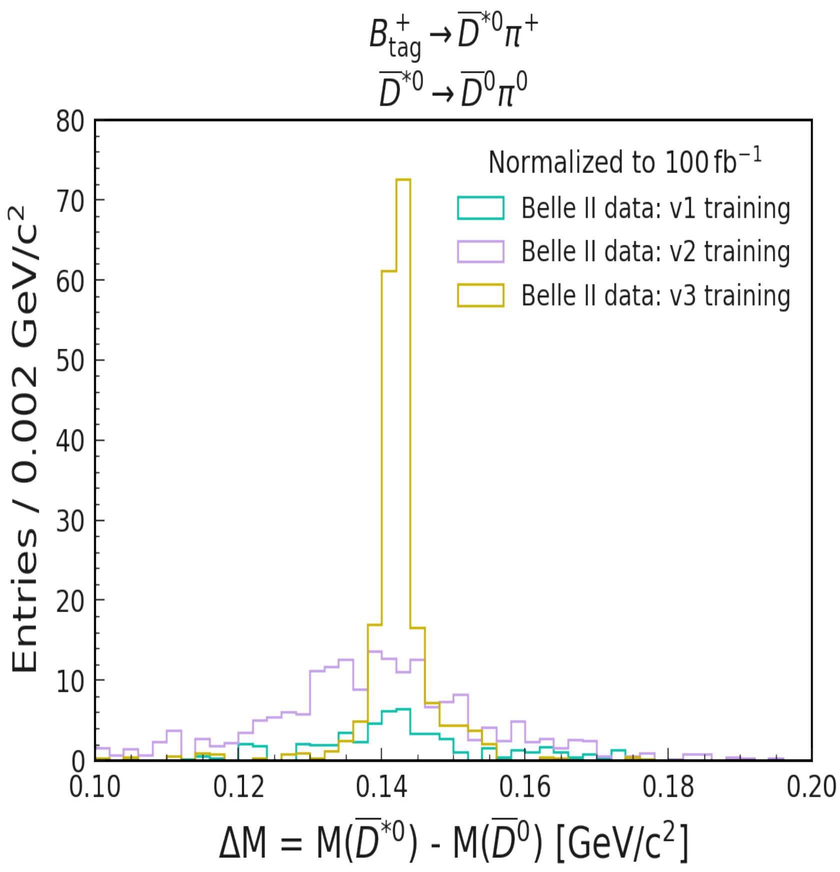


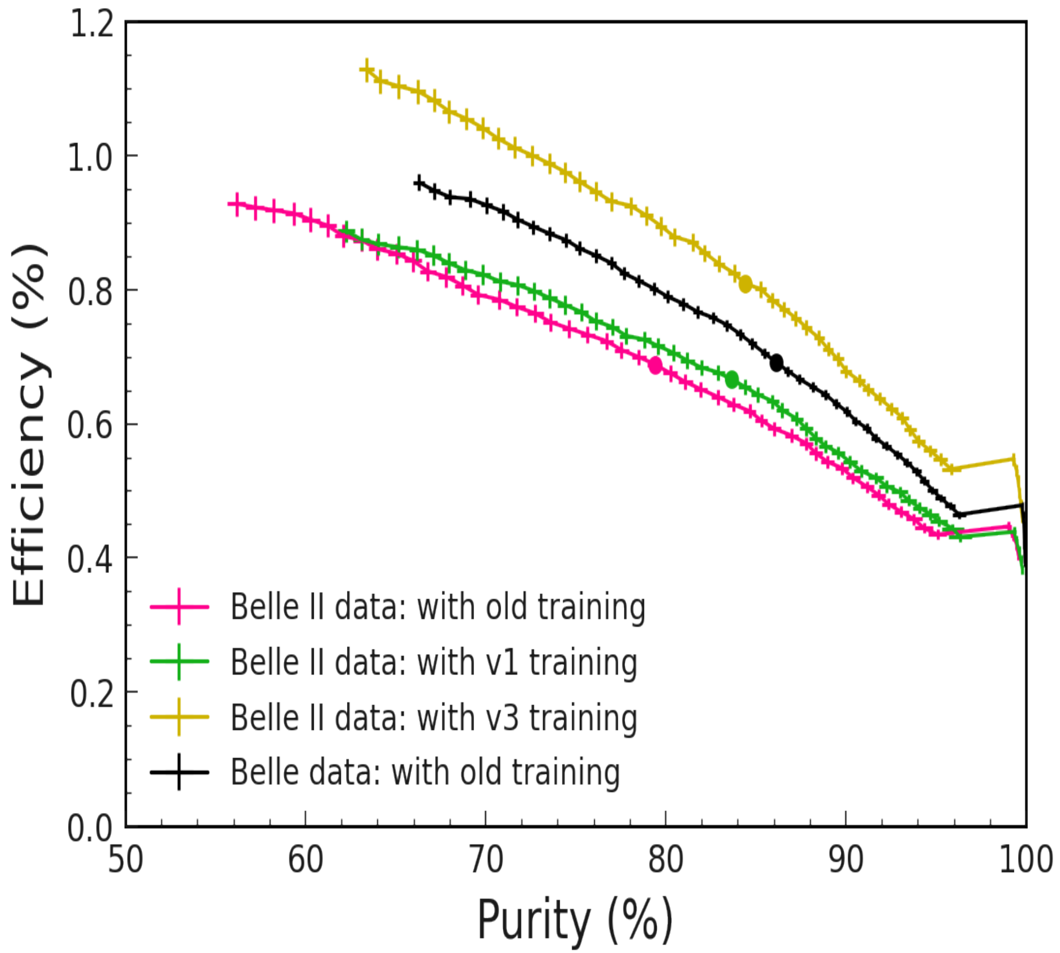


$D^{*0} \rightarrow D^0 \pi^0$ reconstruction

Optimize ΔM for D^{*0} reconstruction

Along with looser preselection for photons, mass-constraint is applied for π^0 candidates in Belle II.
This will improve ΔM distribution which is used in preselection and training for D^{*0}.
Retraining FEI provides expected results:





Improving metrics of FEI

For Hadronic B^+:

- ➢ Updated decay model for the most efficient B decay modes

  Belle 0.75 → 1.04 : 39% ⬆ in Calibration factor

  0.65 → 0.81 : 25% ⬆ in Calibration factor

- ➢ Training with the MCri-up (new DECAY.dec)

  56% → 63% : 12% ⬆ in purity

- ➢ Loosen the γ preselection and mass-constraint π^0

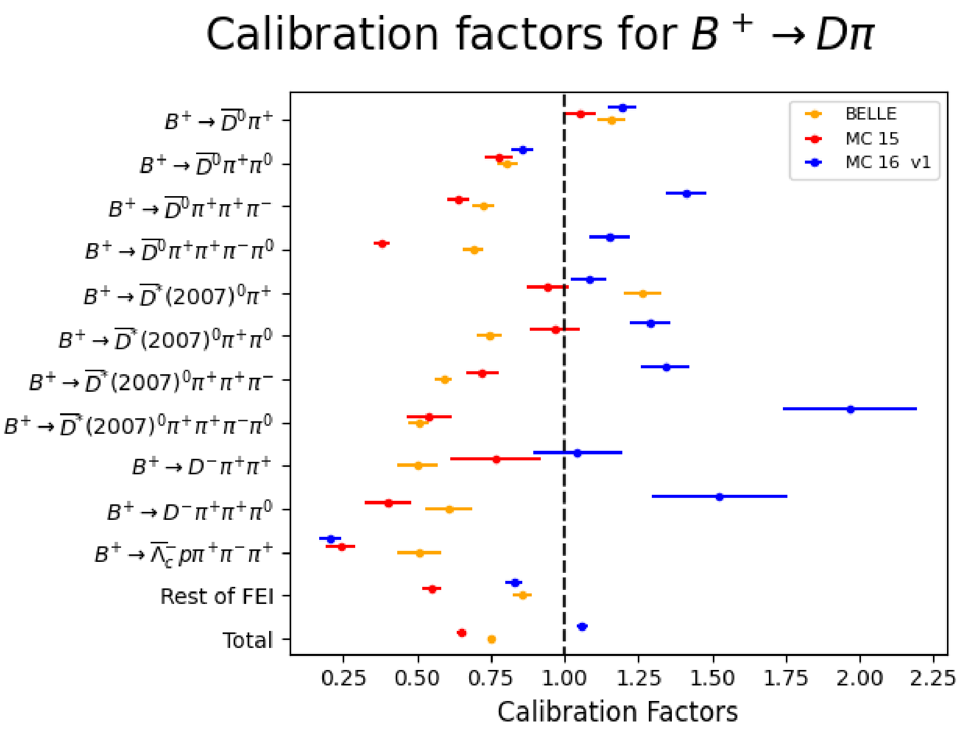
  0.93% → 1.13% : 21% ⬆ in efficiency

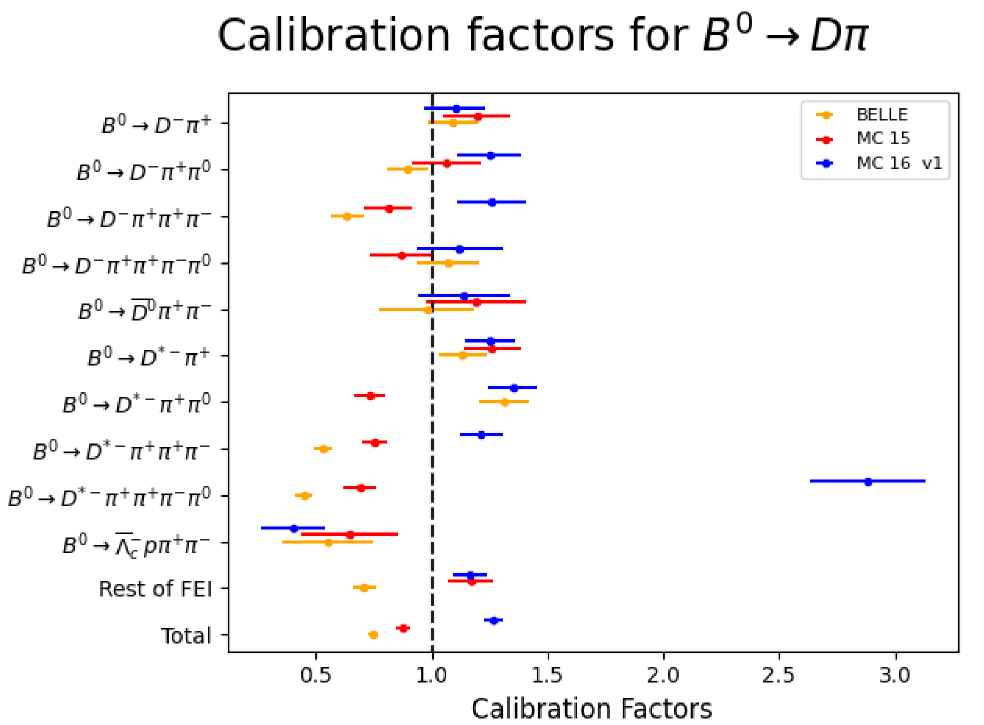
All these improvements are default for MC16/proc16 (shared knowledge)
still studying the impact on SL FEI

Hadronic B-tagging with proc16

see M.Marfoli's talk (+BELLE2-NOTE-PH-2025-033)

Calibration factors at P_{FEI}>0.001 for Belle - MC15 - MC16





General improvements

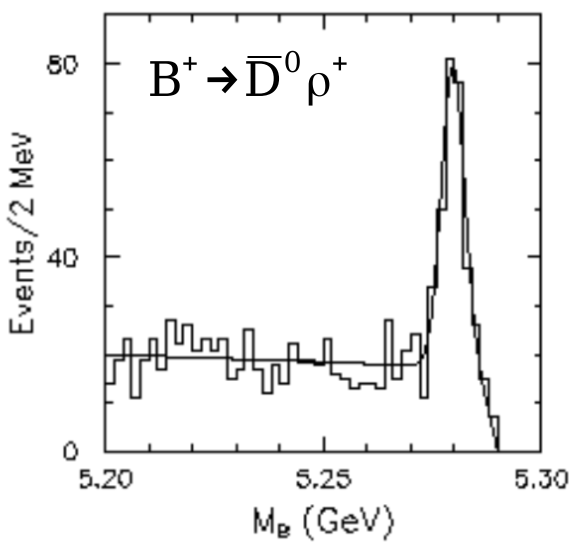
B^+: from 0.65+/-0.01 to 1.04+/-0.02

B^0: from 0.88+/-0.03 to 1.22+/-0.04

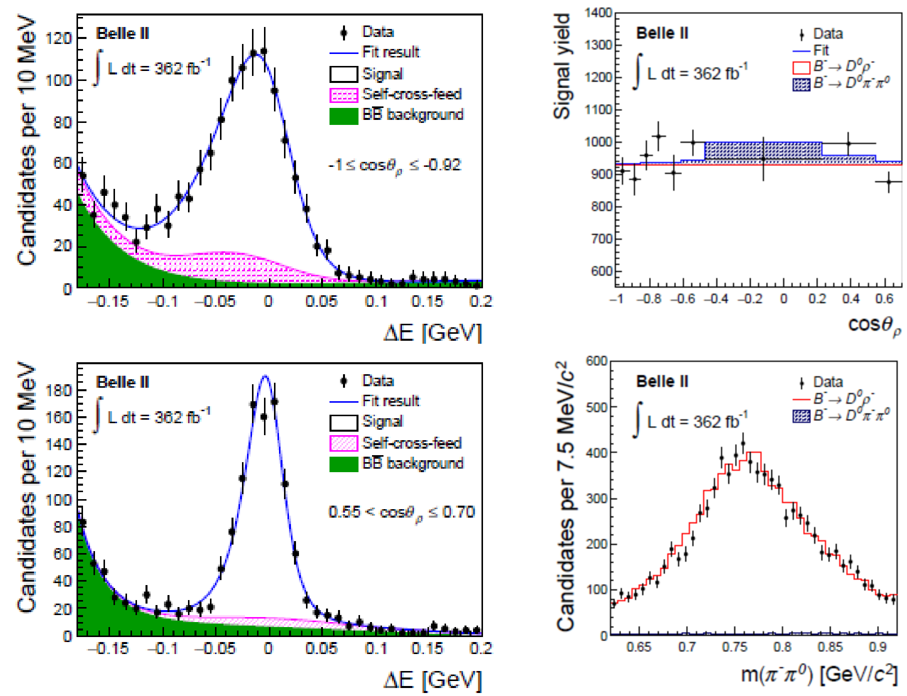
Still some discrepancies, especially in $D^{(*)}3(^0)$ and Λ_c modes

First, understand better the B decays...

CLEO 0.89 fb^{-1}
[PRD 50 (1994) 43]
30 years ago
Uses M_{bc}
$B = (1.34 \pm 018)\%$
13% uncertainty !



[arXiv:2404.10874 , Phys Rev D. 109, L111103]



$B(B^+ \to D^0 \rho^+) = (0.94 \pm 0.02 \pm 0.05)\%$

- World's best result with more than 2× improvement
- Factorisation test: in agreement with prediction
- Systematically limited by uncertainty on π^0 efficiency

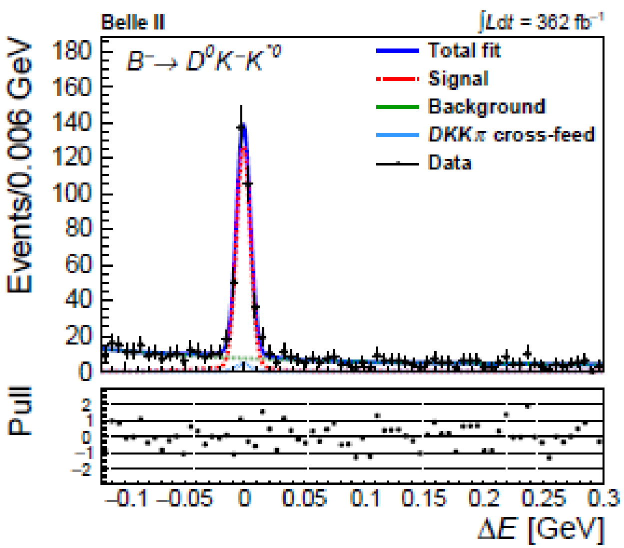
First, understand better the B decays…

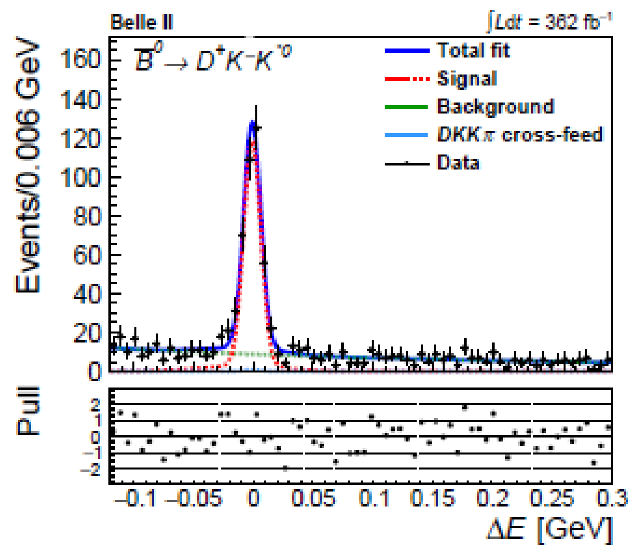
B→DKK: largely unexplored sector

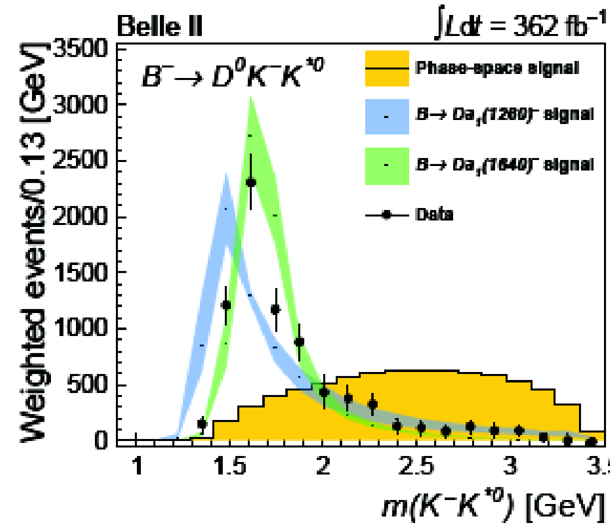
- few % of B branching fraction expected
- Only 0.3% measured so far

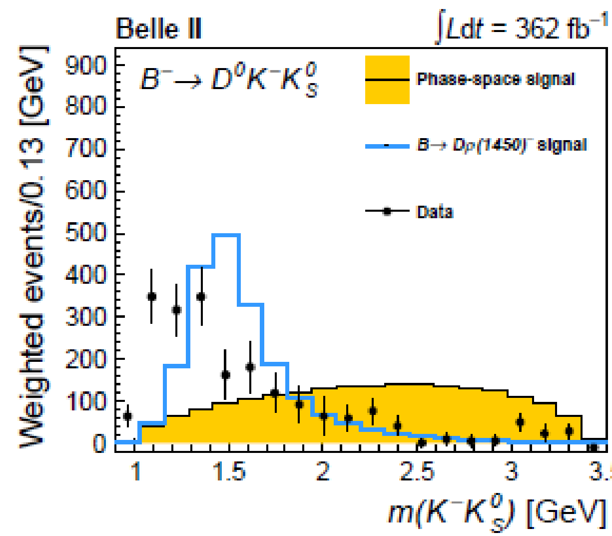
[arXiv:2406.06277, JHEP 08 (2024) 206]

Measurement of the branching fractions of $B\to D^{(*)}KK_S^{(*)}$…





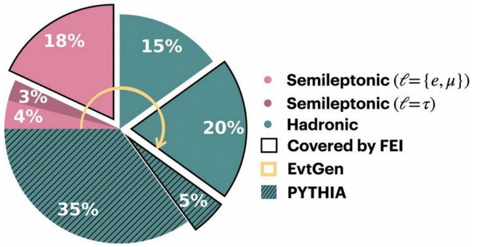




- Efficiency correction applied in the planes $m(D^{(*)}K^-)$ and $m(K^-K_{(S)}^{(*)0})$
- Extraction of bkg-subtracted and efficiency corrected invariant mass and helicity
- Dominant transitions $J^P = 1^{-/+}$
- $B\to D^{(*)}D_s(\to KK^{(*)})$ are used as control modes

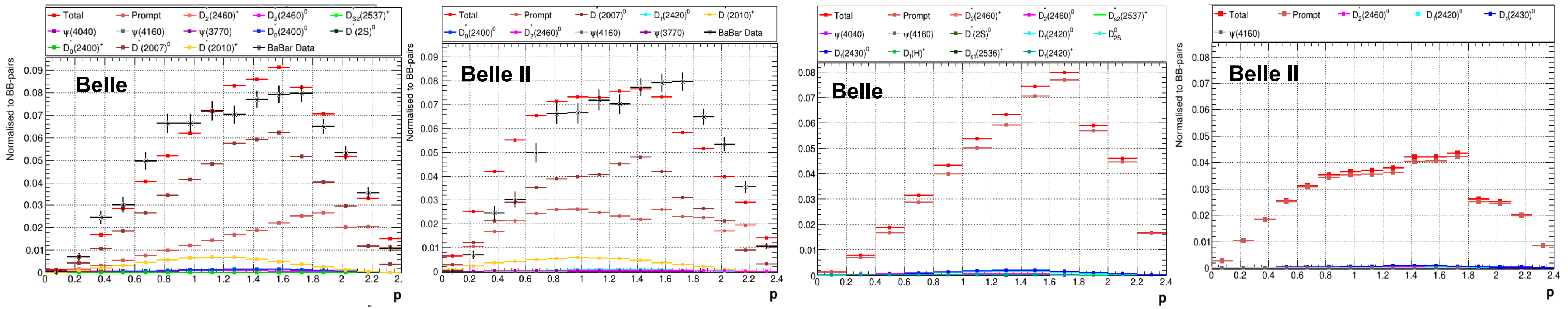
Further improvements → inclusive

- need more measurements to ''constrain'' our MC
- B→DX (**but also B→D*X**), on-going analysis…
- difference between Belle and Belle II MC shows room for improvement:



Correlated

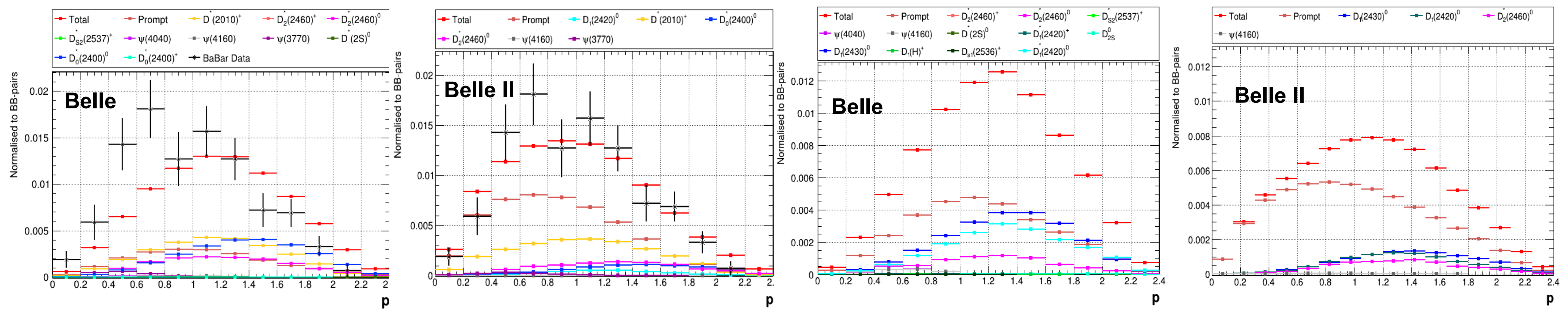
$B^{\pm}\to\bar{D}^0 X$ $B^{\pm}\to\bar{D}^{*0} X$



BR : *Belle* = 0.791 ; *Belle II* = 0.829
BaBar = 0.786±0.016±0.027

BR : *Belle* = 0.496 ; *Belle II* = 0.442

$B^{\pm}\to D X$ $B^{\pm}\to D^{*} X$

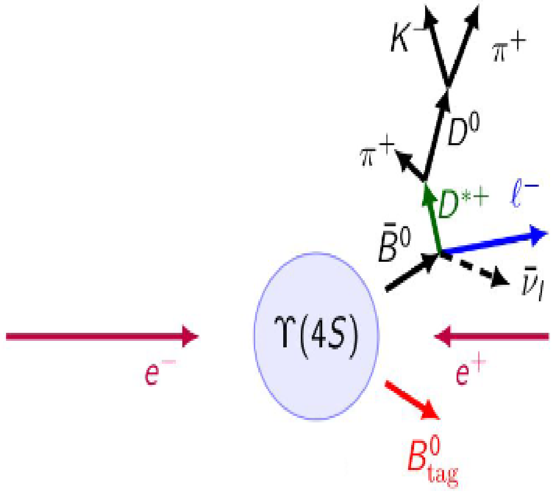


BR : *Belle* = 0.087 ; *Belle II* = 0.095
BaBar = 0.099±0008±0.005

BR : *Belle* = 0.080 ; *Belle II* = 0.074

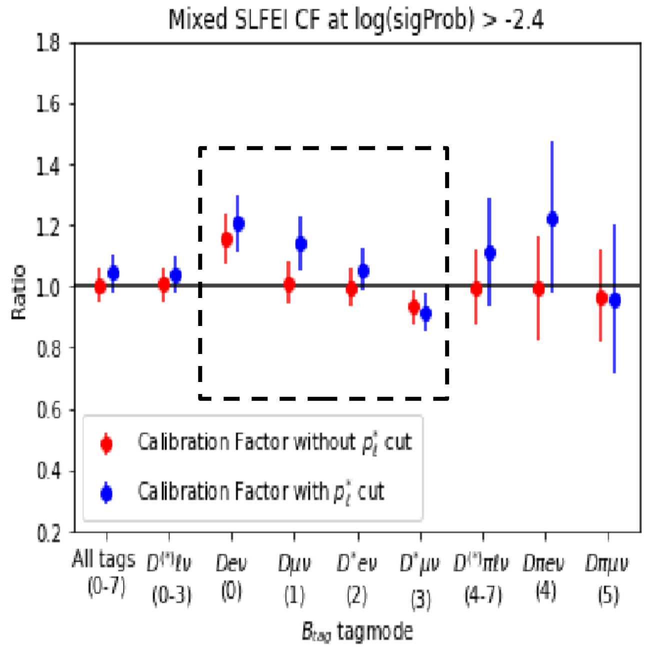
Momentum spectra, for correlated cases in the B to D$^{(*)}$ cases considering the B$^+$ rest frame.

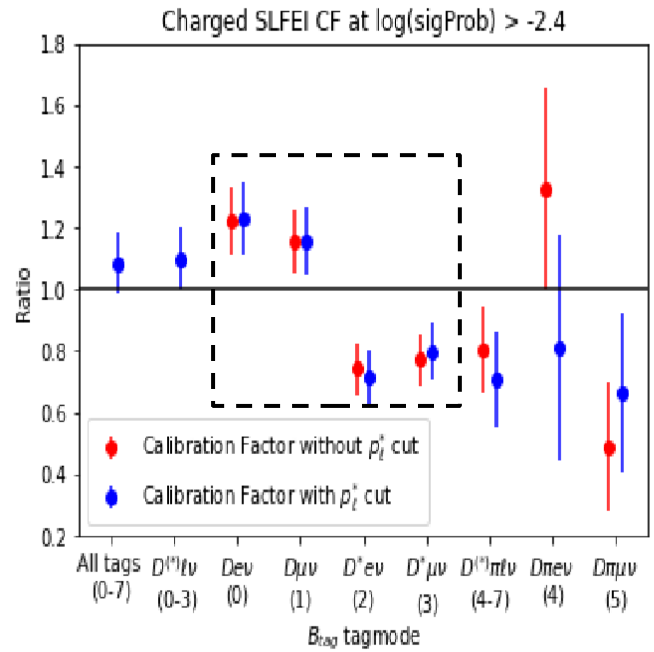
SL FEI calibration with $D^* l \nu$ sample

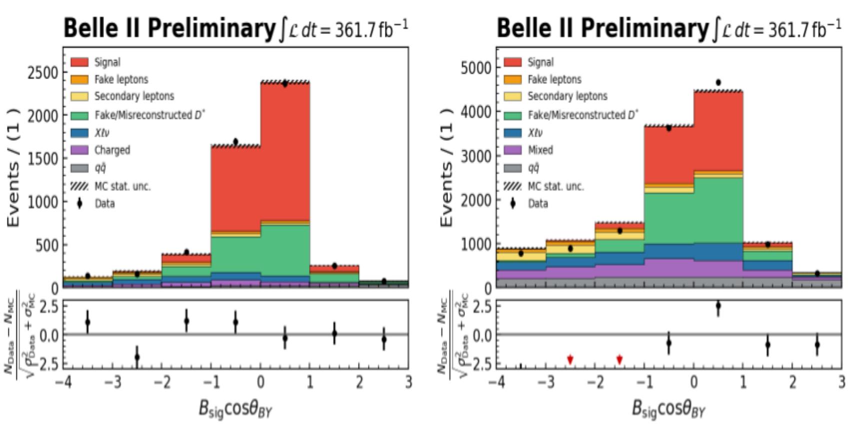


Consistent selection between B_{sig} and B_{tag}

The calibration factors for MC15ri:







(a) $B^0 \to D^{*-}\ell^+\nu_\ell$ (b) $B^+ \to \bar{D}^{*0}\ell^+\nu_\ell$

FIG. 1: Data-MC $\cos\theta_{BY}$ distributions, following reconstruction of an $\Upsilon(4S)$ candidate and the additional selections listed in Table V with all dataset corrections.

Recommendations:

- ➢ Use only the 4 $D^{(*)}\ell\nu$ modes (select after BCS).
- ➢ Apply mode-dependent CF, not the overall.
- ➢ The p_ℓ^* selection could be analysis dependant.

RC in progress to approve the procedure.
Yet to check for MC15rd (Not used for this winter).