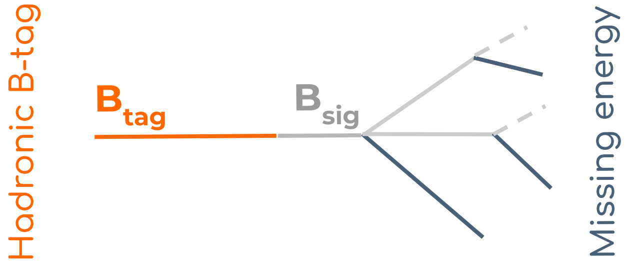
Search for $B^+ \to K^+ \tau^+ \tau^-$ with B-tagging at Belle and Belle II experiments

Journées de Rencontre Jeunes Chercheurs 2022



Vidya Sagar Vobbilisetti

24 Oct 2022

IJCLab
Irène Joliot-Curie
Laboratoire de Physique des 2 Infinis

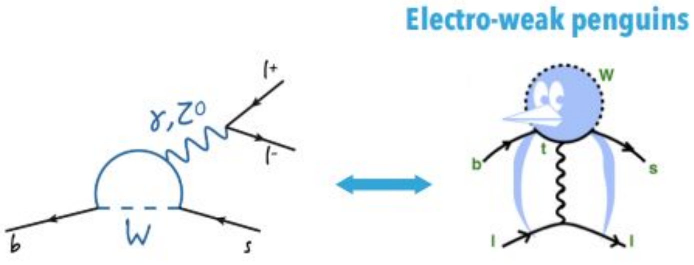
Flavor physics and EW penguins

By precisely measuring the parameters of the Standard Model (SM), we might find signatures of New Physics (NP) beyond the SM.
Or even discover processes that are forbidden in SM.

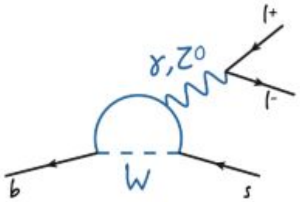
Flavour Changing Neutral Currents (FCNC) **b → s(d)** are one such precision measurements in flavor physics.

The FCNC processes proceed via one-loop diagrams in the SM at lowest order.

Since NP particles may enter the loop diagrams or even mediate FCNCs at tree level, the **b → s(d)** are sensitive to physics beyond the SM.



Current anomalies in B-physics



Semi-leptonic B decays are showing tensions with the SM predictions ⇒ a possible violation of the Lepton Flavor Universality (LFU).

Different behavior for different lepton generations:

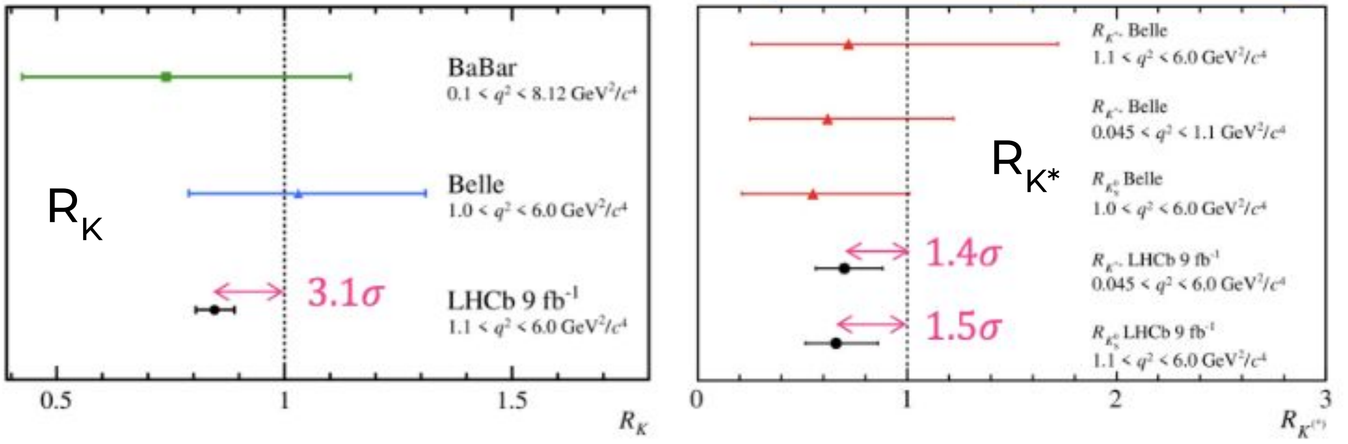
SM

$$R(K^{(*)}) = \frac{\mathcal{B}(B \to K^{(*)}\mu^+\mu^-)}{\mathcal{B}(B \to K^{(*)}e^+e^-)} = 1 \pm \mathcal{O}(10^{-2})$$

$q^2 \in [1(1.1), 6]$ for $R(K^{(*)})$

JHEP 2018, 93 (2018)

arXiv:2103.11769, arXiv:2110.09501



$$R_{K^{(*)}}^{exp} < R_{K^{(*)}}^{SM}$$

Current anomalies in B-physics (cont.)

$$R(K^{(*)}) = \frac{\mathcal{B}(B \to K^{(*)}\mu^+\mu^-)}{\mathcal{B}(B \to K^{(*)}e^+e^-)}$$

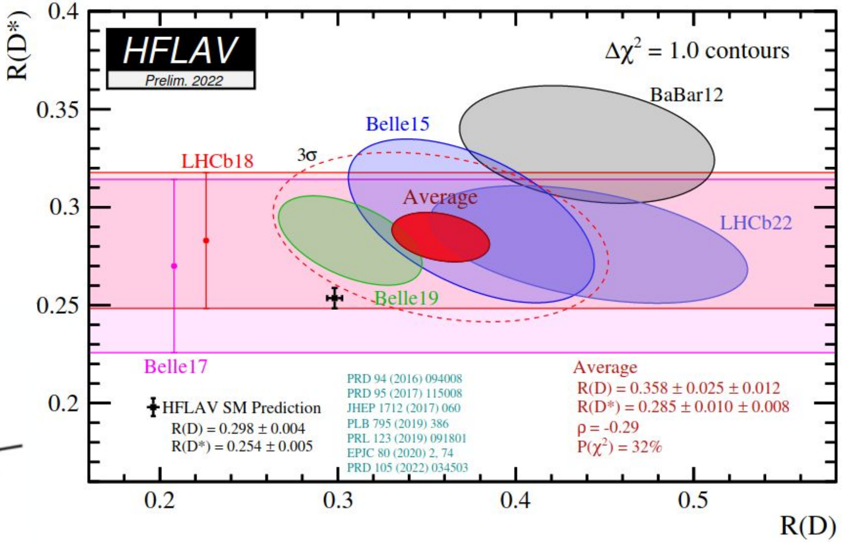
Semi-leptonic B decays are showing tensions with the SM predictions ⇒ a possible violation of the Lepton Flavor Universality (LFU).

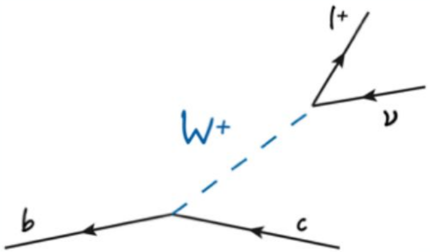
$$R_{K(*)}^{exp} < R_{K(*)}^{SM}$$

SM

$$R(D^{(*)}) = \frac{\mathcal{B}(B \to D^{(*)}\tau\nu_\tau)}{\mathcal{B}(B \to D^{(*)}\ell\nu_\ell)} = 0.300(0.252) \pm \mathcal{O}(10^{-3})$$

$$\ell = e, \mu$$





$$R(D^{(*)})^{exp} > R(D^{(*)})^{SM}$$

NP coupling:
3rd gen > 2nd gen > 1st gen

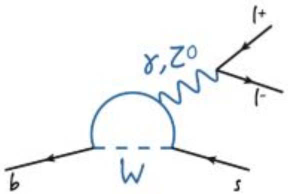
Discrepancy w.r.t. combined average (BaBar, Belle, LHCb):

- R(D): 1.4σ
- R(D*): 2.9σ
- Combined: 3.2σ

IJCLab

Impact on $B^+ \to K^+ \tau^+ \tau^-$ decays

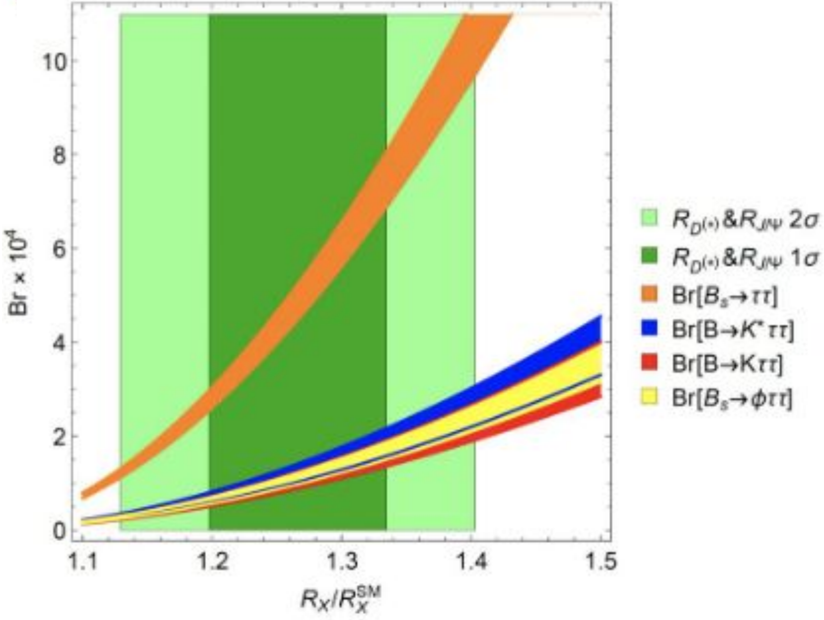
$B^+ \to K^+ \tau^+ \tau^-$ is a FCNC process
⇒ highly suppressed in SM,
⇒ happens through penguin loops
predicted BF: **O(10^{-7})**



τ is 3rd generation and higher mass
⇒ stronger coupling to NP,
like U(1) leptoquark predicts BF: **O(10^{-5} - 10^{-4})**
[2103.16558, 1712.01919]

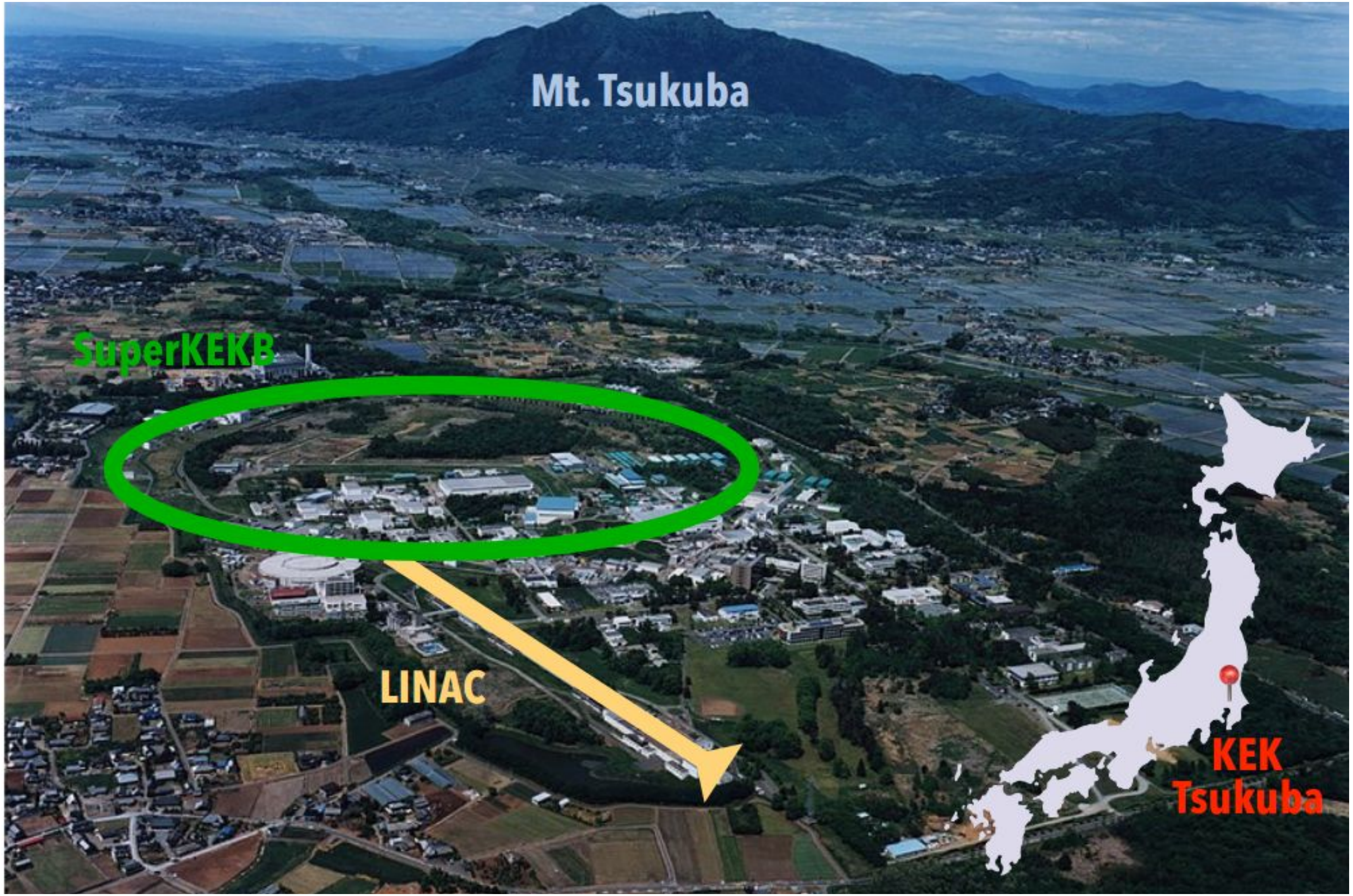
Current (and only) limit:
BF < 2.25 × 10^{-3} @ 90% CL
[BaBar, 1605.09637]

NP coupling: 3^{rd} gen > 2^{nd} gen > 1^{st} gen



B. Capdevila, A. Crivellin, S. Descotes-Genon, L. Hofer, et J. Matias, *arXiv:1712.01919, PRL 120, 181802*

SuperKEKB and Belle II detector

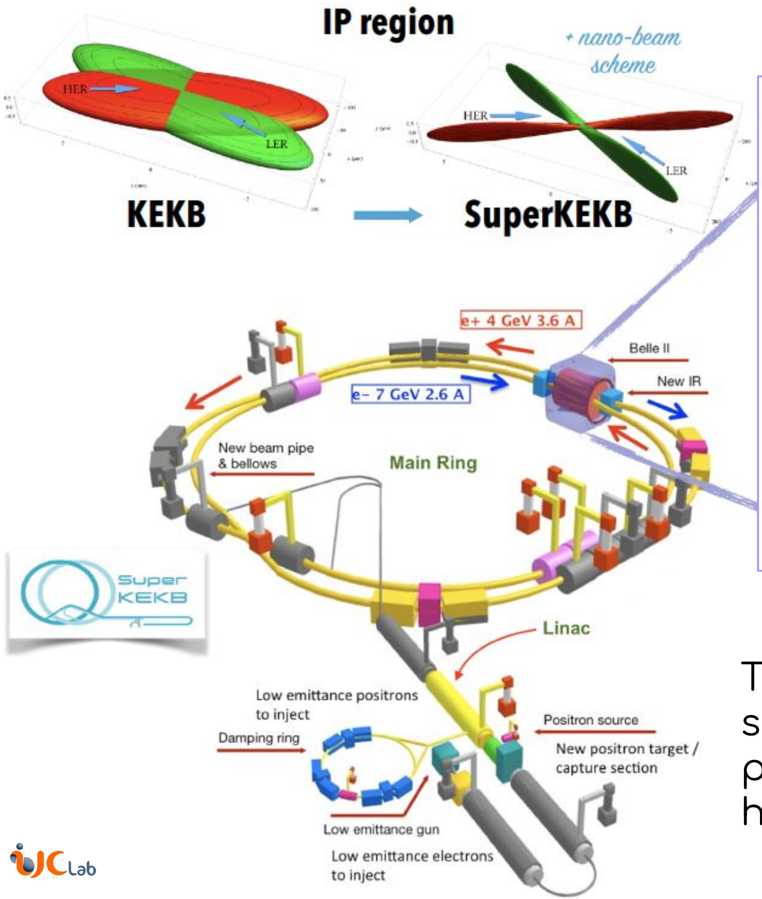


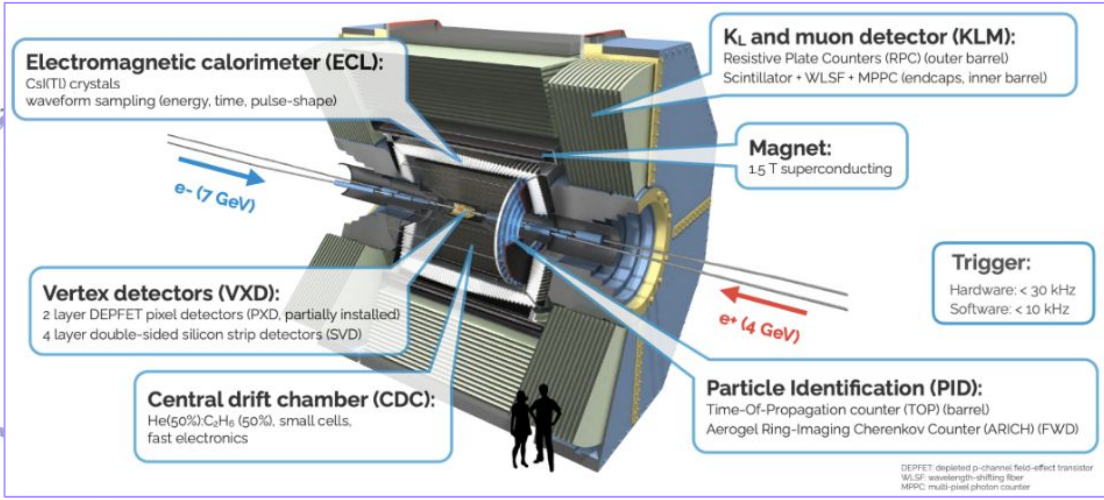
KEKB+Belle collected ~1ab^{-1} in ~10 years (1999-2010)

SuperKEKB + Belle II plans to collect ~50ab^{-1} in ~10 years.

Need to increase instantaneous luminosity substantially!

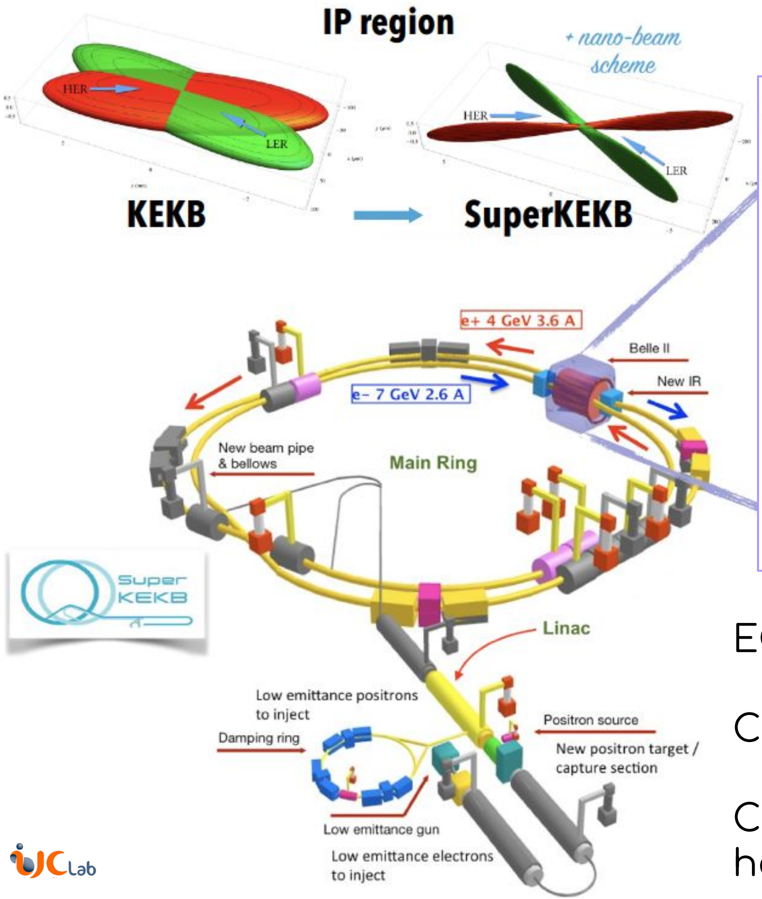
SuperKEKB and Belle II detector

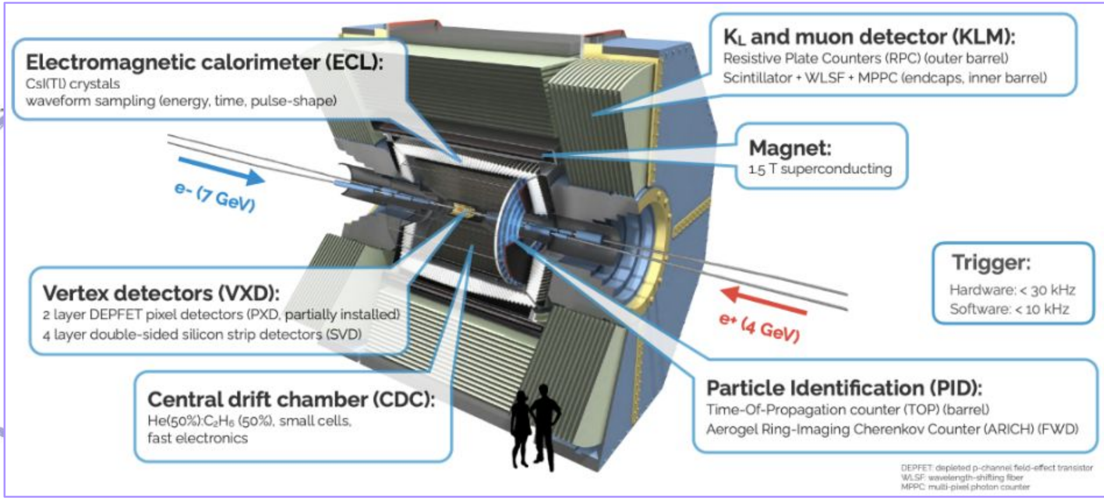




The upgrade KEKB → SuperKEKB has required a substantial redesign of Belle II detector, whose performance is challenged by radiation damage and higher background (design luminosity is x40 higher).

SuperKEKB and Belle II detector



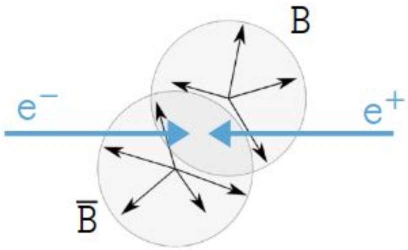


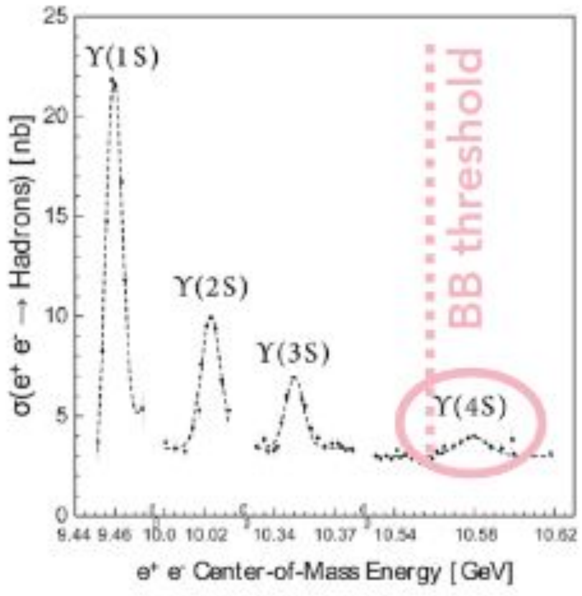
ECL provides information about the energy deposited.

CDC provides information on tracks.

CDC along with other dedicated detectors (ARICH & TOP) help identify different particles (PID): e vs μ vs π vs K

Principle of B-factories





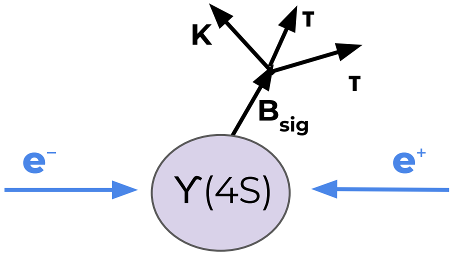
e^+e^- collisions at Υ(4S) @ 10.58 GeV
(above the threshold to produce BB pairs)

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

happens along with:

$$e^+e^- \rightarrow \begin{cases} e^+e^-(\gamma) \\ \mu^+\mu^-(\gamma) \\ e^+e^-e^+e^- \\ \tau^+\tau^-(\gamma) \\ e^+e^-\mu^+\mu^- \\ \gamma\gamma(\gamma) \\ u\bar{u}(\gamma) \\ d\bar{d}(\gamma), s\bar{s}(\gamma) \\ c\bar{c}(g) \end{cases}$$

Signal events at B-factories



e^+e^- collisions at Υ(4S) @ 10.58 GeV
(above the threshold to produce BB pairs)

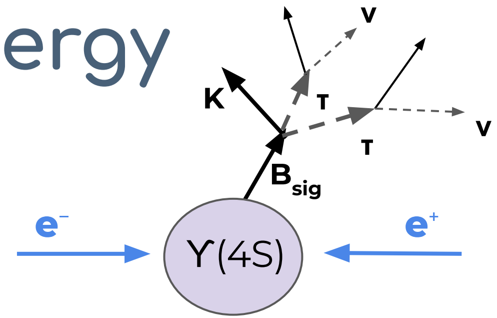
$e^+ e^- \to \Upsilon(4S) \to B\bar{B}$

happens along with:

$$e^+e^- \to \begin{cases} e^+e^-(\gamma) \\ \mu^+\mu^-(\gamma) \\ e^+e^-e^+e^- \\ \tau^+\tau^-(\gamma) \\ e^+e^-\mu^+\mu^- \\ \gamma\gamma(\gamma) \\ u\bar{u}(\gamma) \\ d\bar{d}(\gamma), s\bar{s}(\gamma) \\ c\bar{c}(g) \end{cases}$$

These are the events we are searching for among all the possibilities.

Signal events have missing energy



e^+e^- collisions at Υ(4S) @ 10.58 GeV
(above the threshold to produce BB pairs)

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

happens along with:

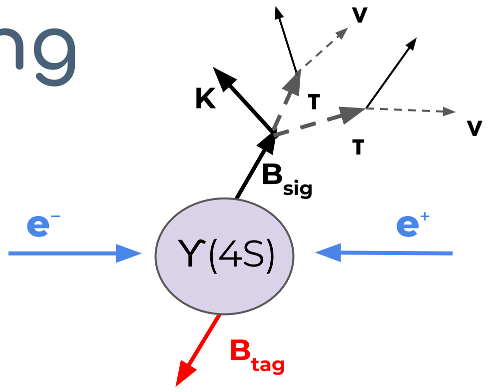
$$e^+e^- \rightarrow \begin{cases} e^+e^-(\gamma) \\ \mu^+\mu^-(\gamma) \\ e^+e^-e^+e^- \\ \tau^+\tau^-(\gamma) \\ e^+e^-\mu^+\mu^- \\ \gamma\gamma(\gamma) \\ u\bar{u}(\gamma) \\ d\bar{d}(\gamma), s\bar{s}(\gamma) \\ c\bar{c}(g) \end{cases}$$

But the τ particles decay into neutrinos, which can't be detected.

⇒ Missing energy

⇒ The B_{sig} can't be fully-reconstructed.

Missing energy needs B-tagging



e^+e^- collisions at ϒ(4S) @ 10.58 GeV
(above the threshold to produce BB pairs)

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

happens along with:

$$e^+e^- \rightarrow \begin{cases} e^+e^-(\gamma) \\ \mu^+\mu^-(\gamma) \\ e^+e^-e^+e^- \\ \tau^+\tau^-(\gamma) \\ e^+e^-\mu^+\mu^- \\ \gamma\gamma(\gamma) \\ u\bar{u}(\gamma) \\ d\bar{d}(\gamma), s\bar{s}(\gamma) \\ c\bar{c}(g) \end{cases}$$

But the τ particles decay into neutrinos, which can't be detected.

⇒ Missing energy

⇒ The B_{sig} can't be fully-reconstructed.

So we fully-reconstruct the other B (**B_{tag}**) in the event to

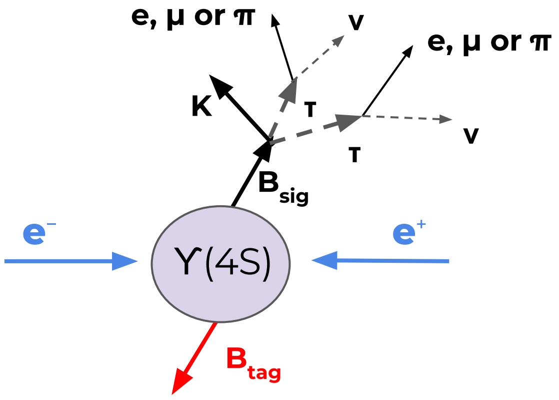
- to distinguish $B\bar{B}$ event from others
- constrain the kinematics.

Analysis procedure

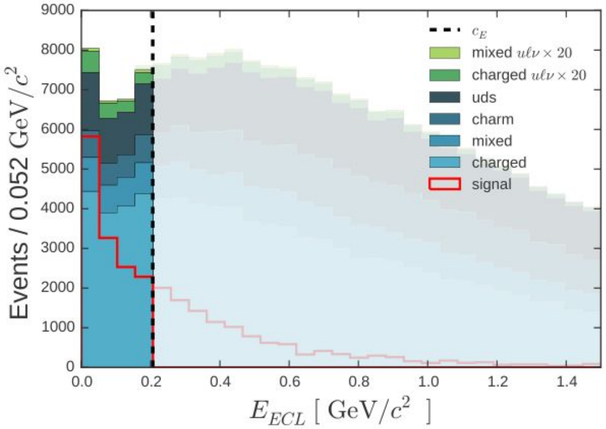
- We start by reconstructing one B (**B_{tag}**) completely.
- And look for a K and a combination pair of e, μ or π in the rest of the event.

- i.e., we reconstruct everything in the event except for the 2-4 neutrinos in the final state.

- The extra energy in calorimeter (E_{ECL}) should peak at 0.*



$BF(\tau^+ \to e^+ \bar{\nu} \nu) = 17\%$
$BF(\tau^+ \to \mu^+ \bar{\nu} \nu) = 17\%$
$BF(\tau^+ \to \pi^+ \bar{\nu}) = 10\%$

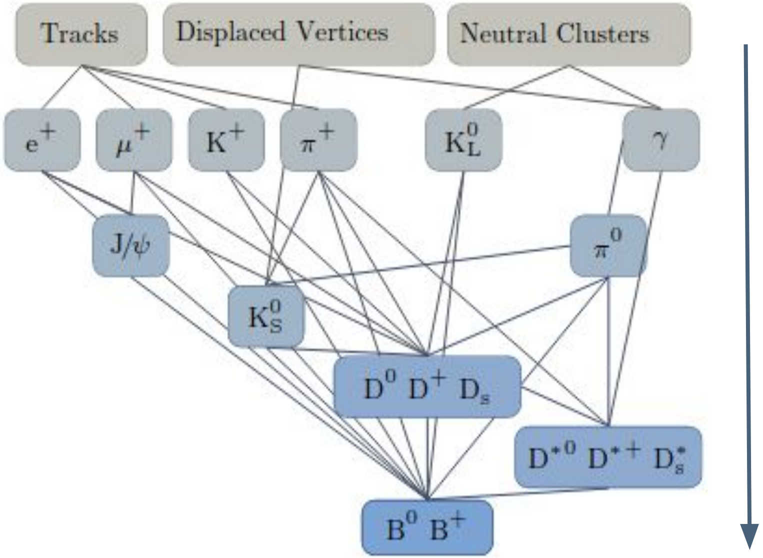


Just for illustration

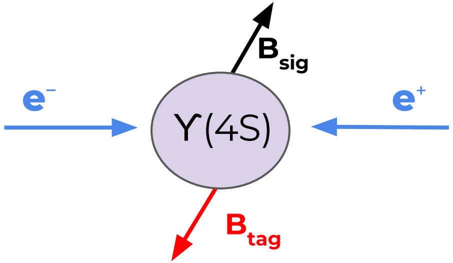
*If we do π^0 veto

B⁺ tagging

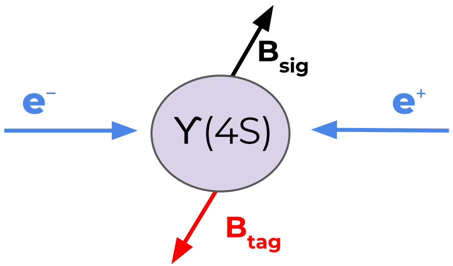
[T.Keck et. al, Comput Softw Big Sci (2019) 3: 6]



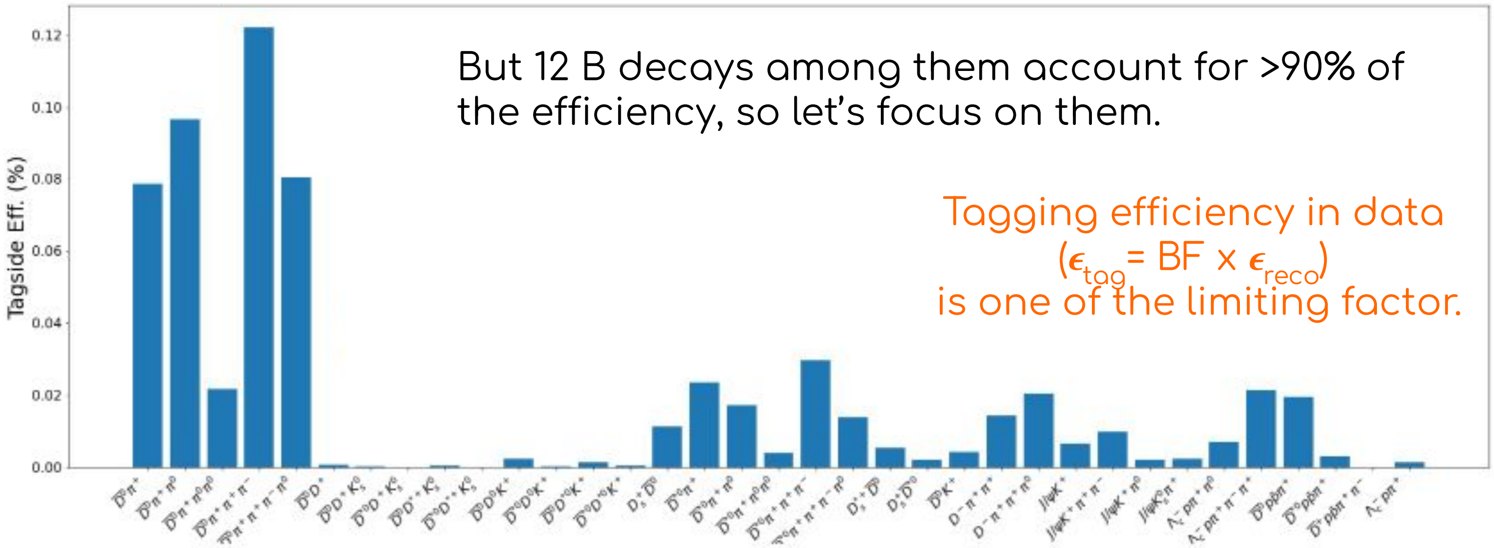
In FEI, Belle II's B-tagging algorithm:
BDTs are trained on MC for some final states in a hierarchical structure starting from tracks and clusters.



B⁺ tagging: 36 final states!



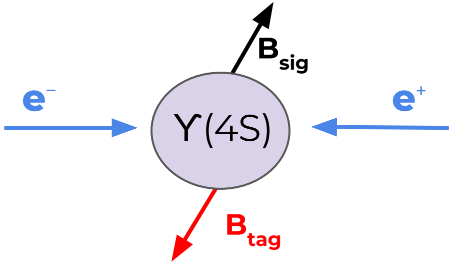
The hadronic FEI algorithm reconstructs B in 36 different B decays.



But 12 B decays among them account for >90% of the efficiency, so let's focus on them.

Tagging efficiency in data
(ϵ_{tag} = BF x ϵ_{reco})
is one of the limiting factor.

B^+ tagging: Effectively 12 final states!



In Hadronic tagging, we essentially reconstruct (12 decays)
$B \to D^{(*)}\ (n\pi^+)\ (m\pi^0)$ final states:

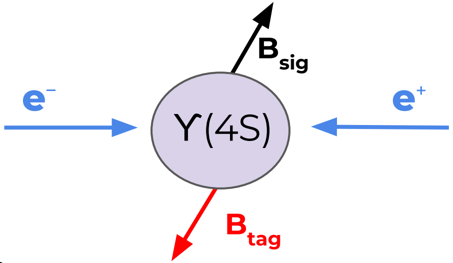
- $\overline{D}^0\pi^+$
- $\overline{D}^{*0}\pi^+$
- $\overline{D}^0\pi^+\pi^0$
- $\overline{D}^{*0}\pi^+\pi^0$
- $\overline{D}^0\pi^+\pi^+\pi^-$
- $\overline{D}^{*0}\pi^+\pi^+\pi^-$
- $\overline{D}^0\pi^+\pi^0\pi^0$
- $\overline{D}^{*0}\pi^+\pi^0\pi^0$
- $\overline{D}^0\pi^+\pi^+\pi^-\pi^0$
- $\overline{D}^{*0}\pi^+\pi^+\pi^-\pi^0$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

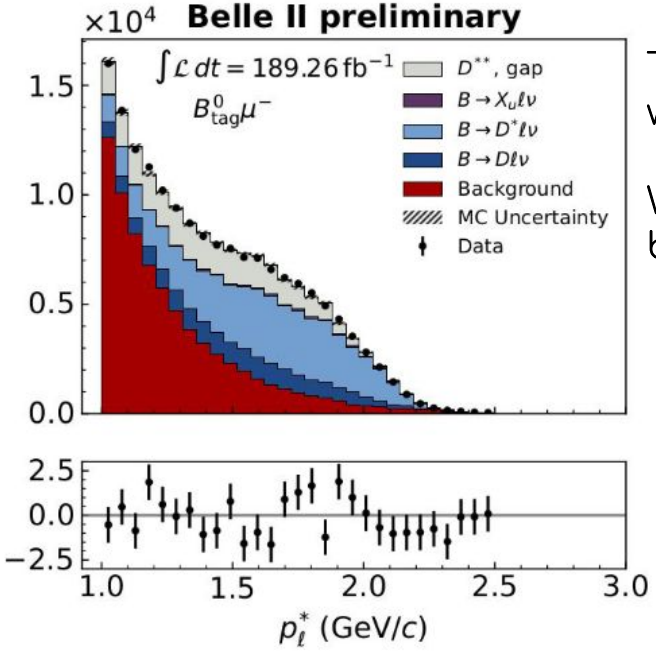
More $\pi \Rightarrow$ More complex,
but higher Branching Fraction

Tagging efficiency in data
(ϵ_{tag} = BF x ϵ_{reco})
is one of the limiting factor.

B^+ tagging: Traditional calibration sample

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





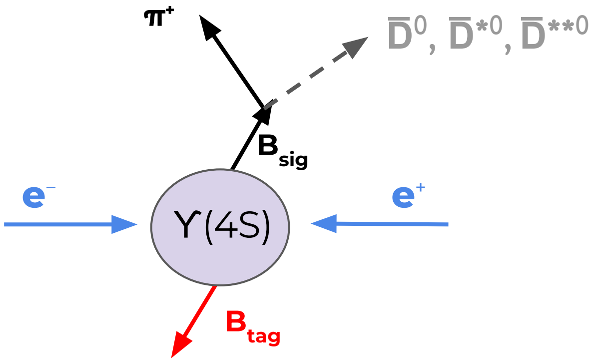
Traditionally, this calibration is done with semi-leptonic B on the signal side.

Which works well because it has large branching fraction.

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

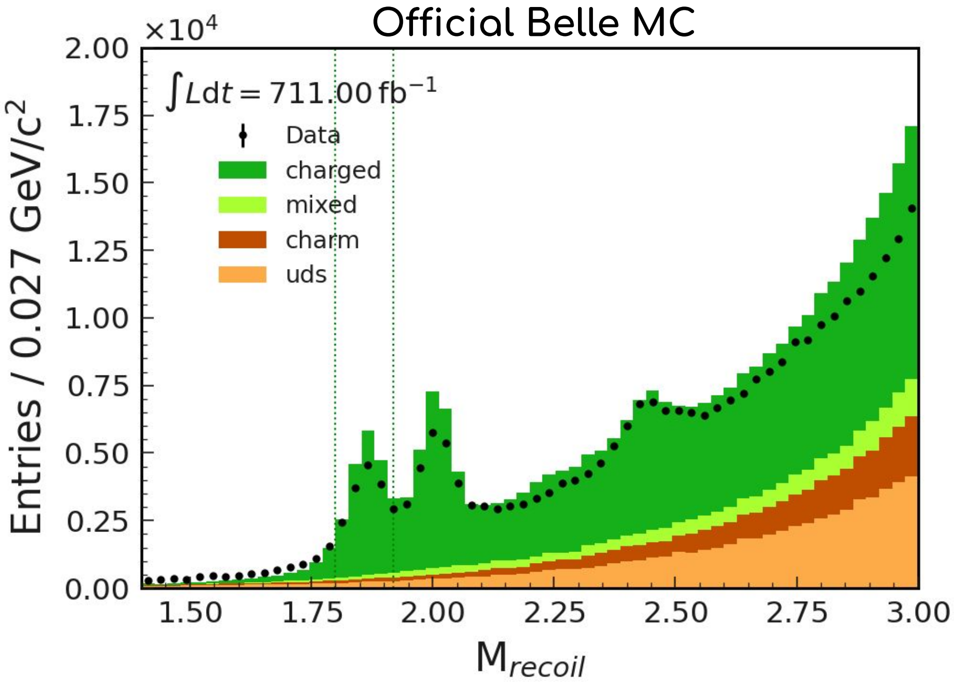
This cannot be studied with semi-leptonic B because there are no peaking structures.

Ideal control sample to study 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 π±

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.

IJCLab

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)%

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

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

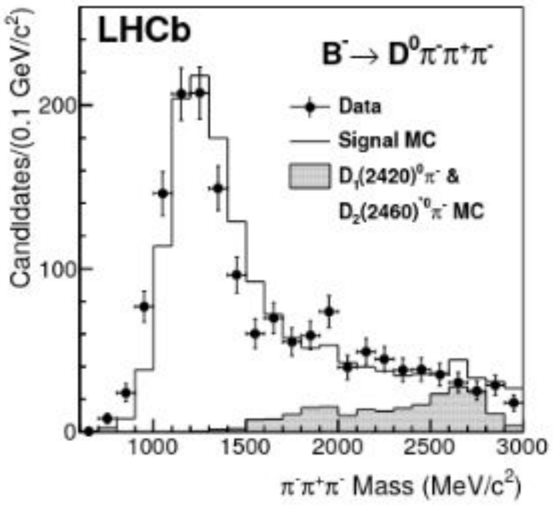
IJCLab

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



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

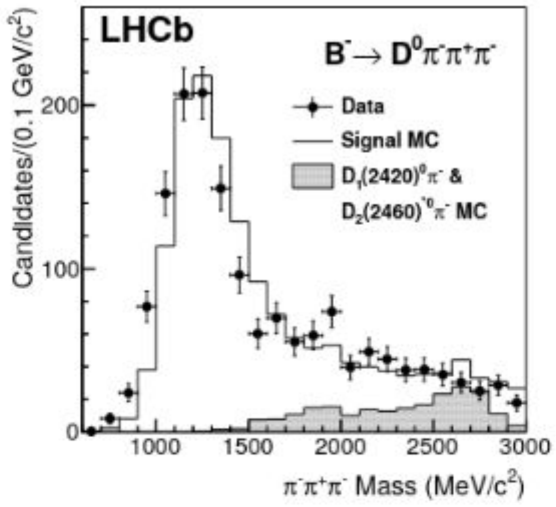
So we are still suck with a 30 year old CLEO measurement in PDG.

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



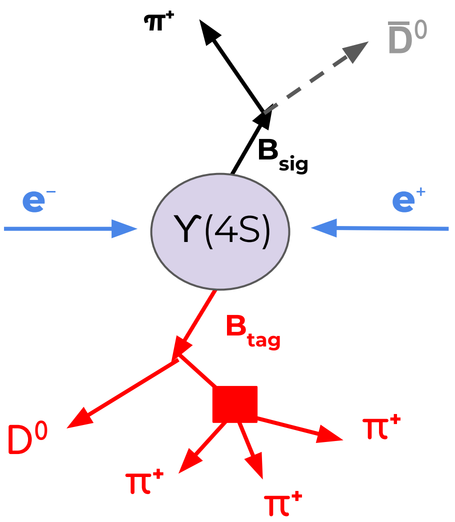
But looking at this plot, it looks like most contribution comes through a_1^+ resonance (mass 1400 MeV/c^2).

Improving MC model: An example

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

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

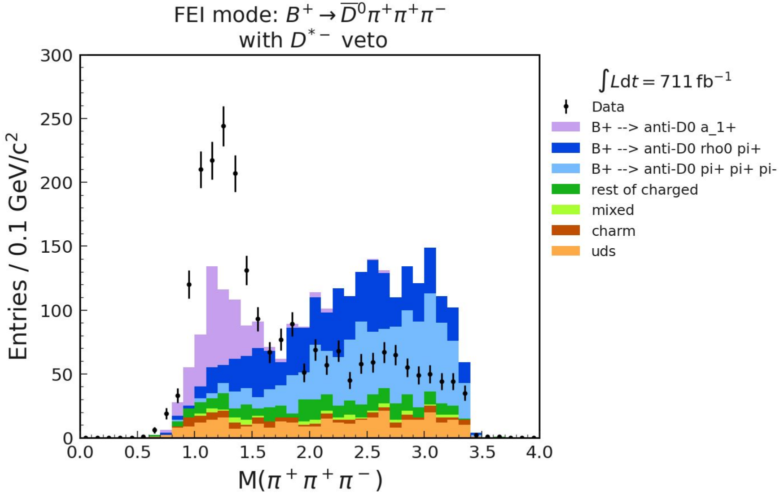
Can be compared with data at Belle, if we reconstruct one B as $B^+ \rightarrow \bar{D}^0 \pi^+$ and other B as $B^- \rightarrow D^0 \pi^+ \pi^+ \pi^-$



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



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

BELLE2-NOTE-PH-2022-002

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

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

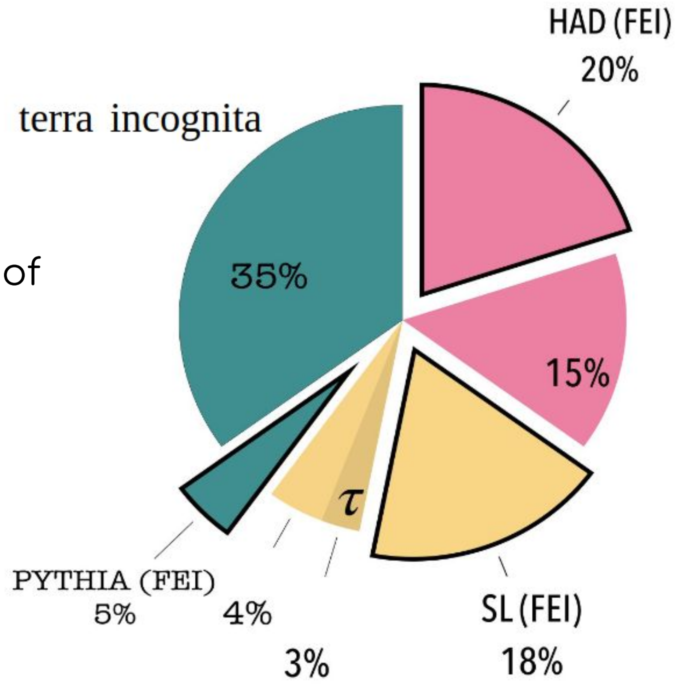
TABLE IX: Contents of the DECAY file concerning the $B^+ \to \overline{D}^0\pi^+\pi^+\pi^-\pi^0$ final state and corresponding measurements in PDG [in %]. The rows in blue correspond to decays produced by Pythia.

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	★	

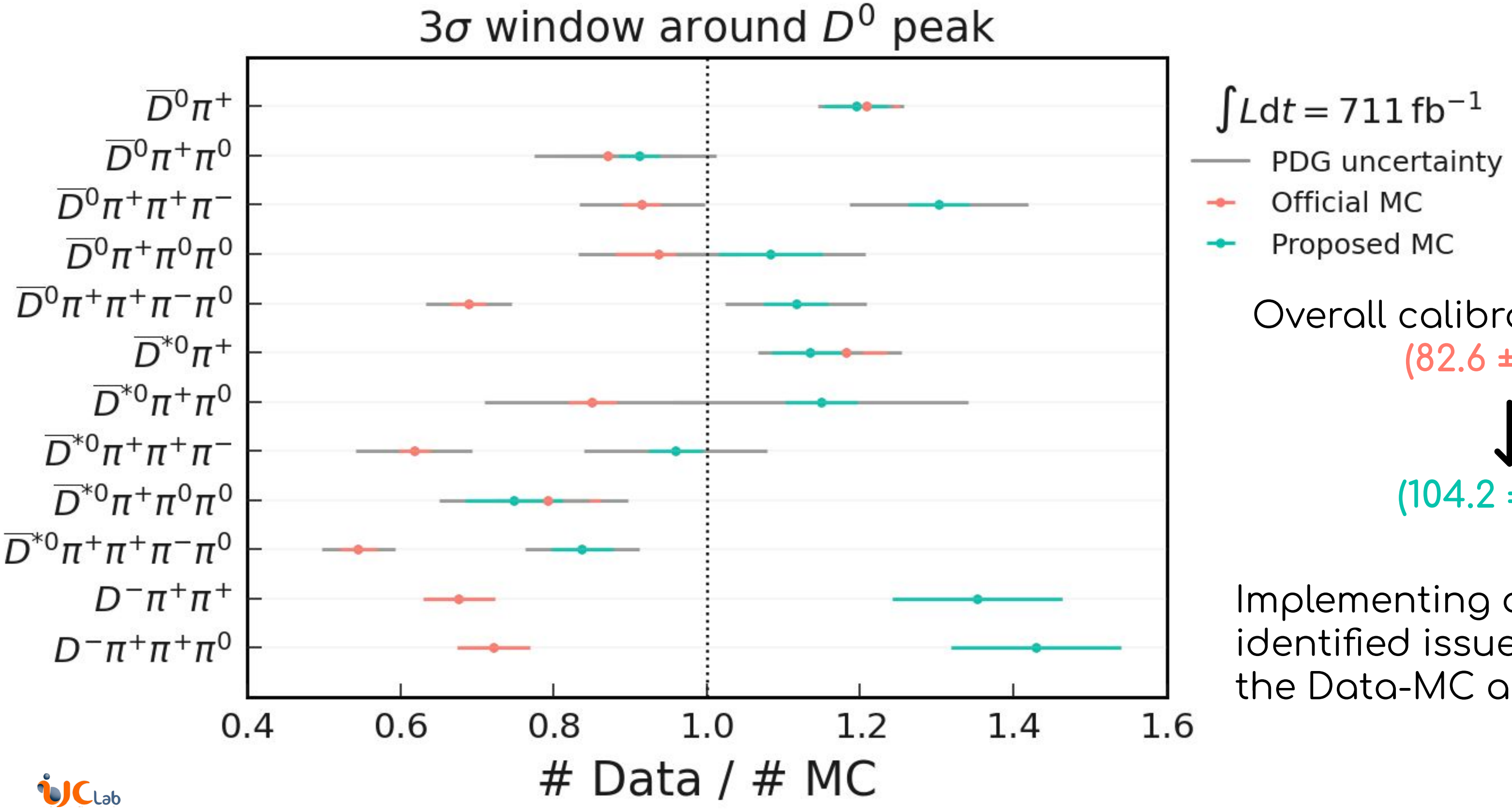
Why is B-decay modeling so hard?

We already saw that we (and PDG) uses a 30-year-old measurement with ~75% uncertainty for one of the largest hadronic B-decay.

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



Improving MC model ⇒ B^+ tagging



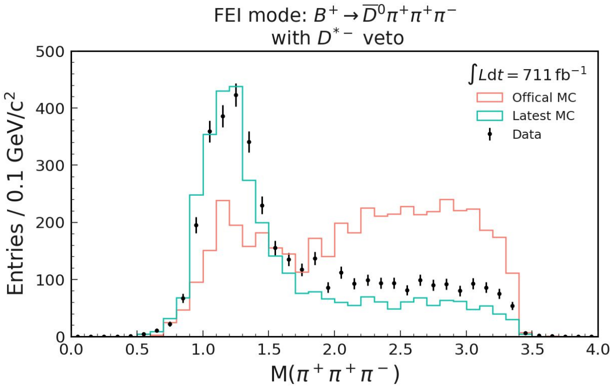
Overall calibration factor:
(82.6 ± 0.9)%
↓
(104.2 ± 1.2)%

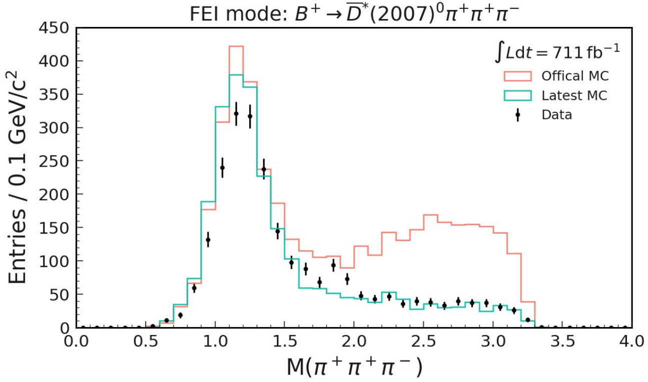
Implementing all the identified issues improves the Data-MC agreement!

Decay description is improved!

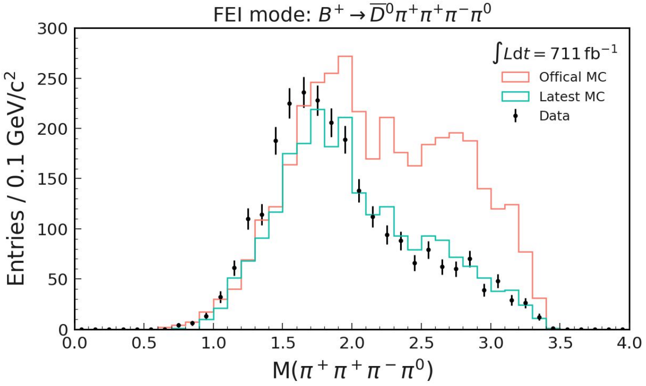
The improvement is not limited to calibration factors, but more importantly in the invariant masses (of intermediate particles), which are used as training variables in FEI

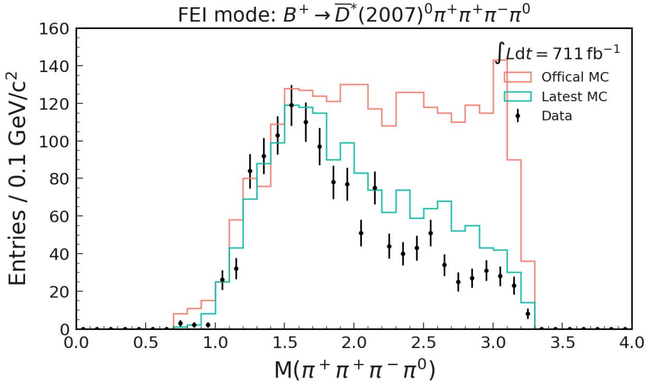
3π± case:





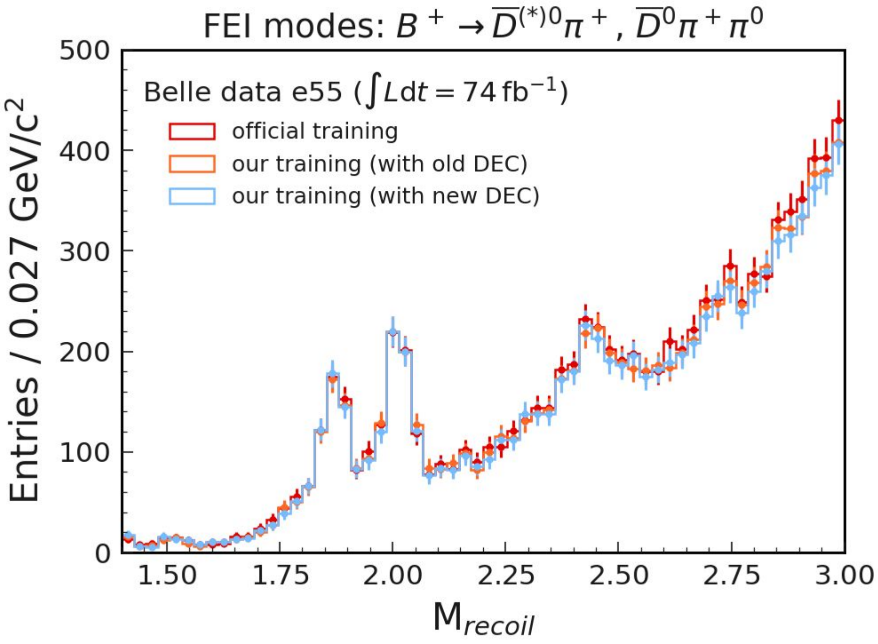
3π± π⁰ case:



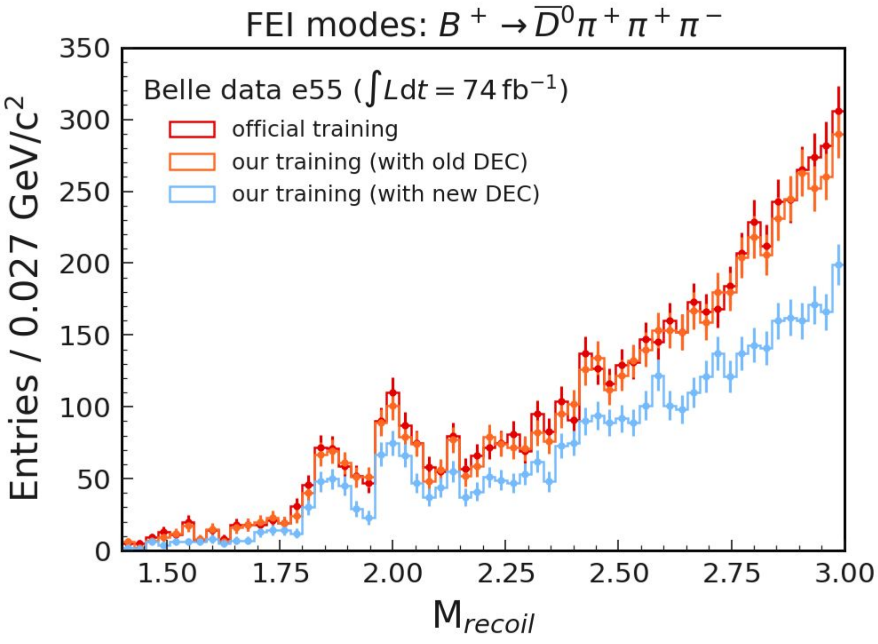


Retraining FEI: Validation

Once we have a new model for how the **B → D$^{(*)}$ (nπ^+) (mπ^0)** decays, we can train BDTs again with it and see performance:

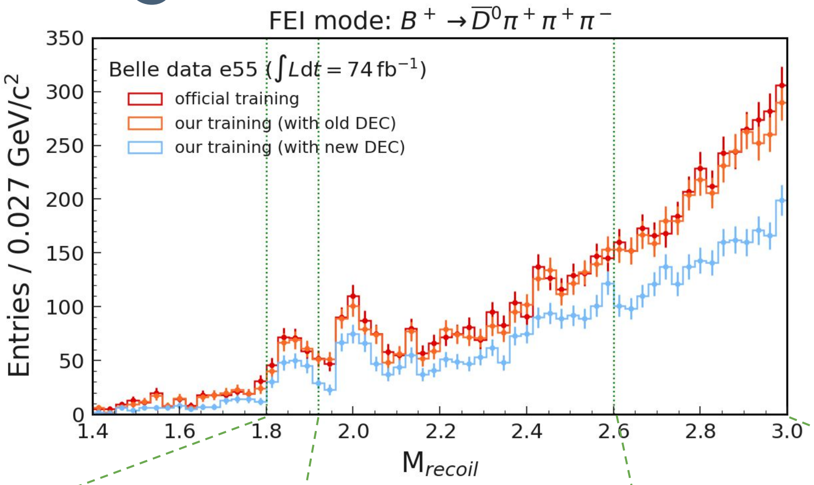


Nothing changes in the FEI modes where we did not change anything.



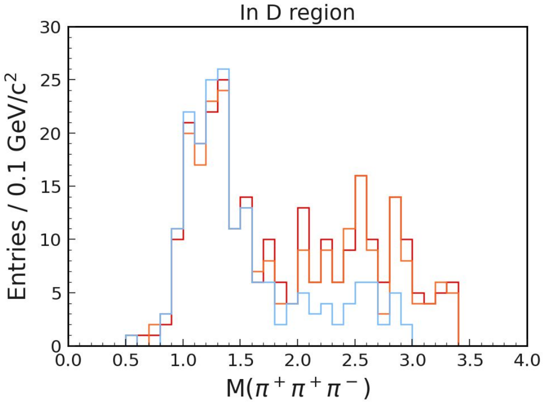
There is a significant background reduction in FEI modes where MC model is improved.

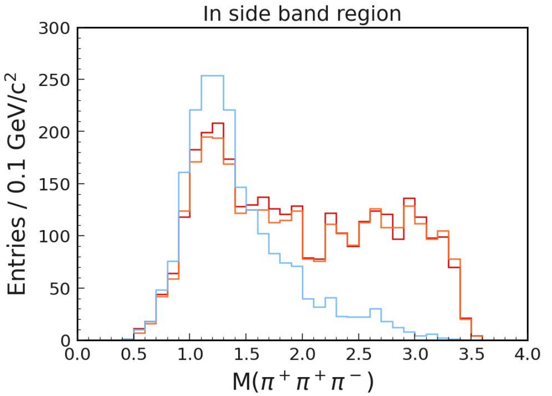
Retraining FEI: Effective cuts



The new training is learning the a_1^+ cut from the MC we give it!

Can we apply this cut manually instead?



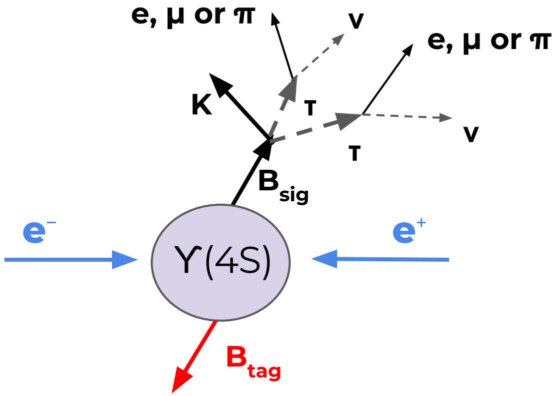


Back to signal-side

Once we have a reliable B_{tag}
We add K + 2<e, μ or π> to it.

For the signal events, the extra energy left in the calorimeter will peak at 0, because neutrinos don't leave energy behind.

We next have to identify what other events when mis-reconstructed can mimic signal.
And train BDTs to suppress such background.



In our group, similar efforts have been made for $B^+ \to K^+ \tau^+ l^-$ reconstruction.

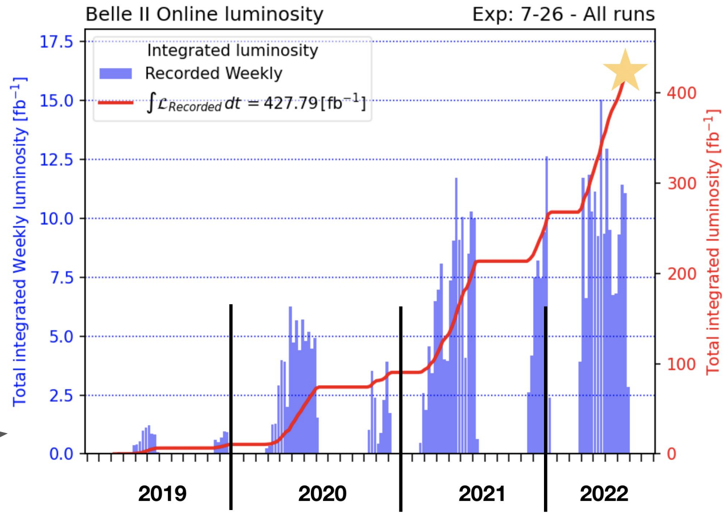
Estimated sensitivity:
BF: O(10^{-4}) at 90% CL with Belle data
BF: O(10^{-5}) at 90% CL with full Belle II data
[1808.10567]

So, in the order of NP predictions!

Summary

- $B^+ \to K^+ \tau^+ \tau^-$ has two 3rd gen. leptons
  ⇒ Good probe for New Physics

- Search status:
  - Only 1 result (from BaBar) so far.
  - Searching in Belle + early Belle II data with hadronic B-tagging
  - Belle II is taking data!



Better MC modeling of hadronic B decays can improve B-tagging performance (calibration factor and background rejection)

Collected : ~420 fb^{-1}
10 year goal : 50 ab^{-1}

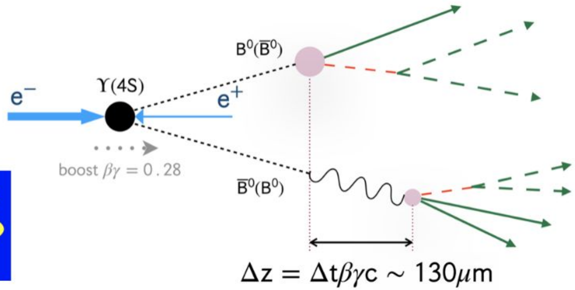
Analysis procedure

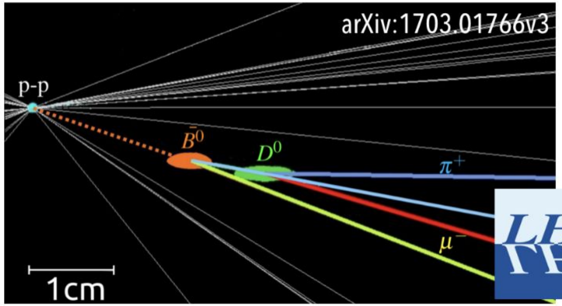
1. There are two B in one event
2. One B is fully reconstructed
3. Many B modes, and as soon as more than two π in $B \to D^{(*)}X$, is complex but high BF.
4. In other B you can probe modes with neutrinos (even 4!)
5. Belle (II) has a large advantage over LHCb for this search. (different situation than $B^+ \to K^+ l^+ l^-$)

Backup

Belle II vs LHCb

Belle II $e^+e^- \to \Upsilon(4S) \to B\bar{B}$		**LHCb** $pp \to B\bar{B}X$	
Two B's and nothing else	Higher tagging efficiency	Large pp background	
Small cross section $\sigma_{bb} \sim 1$ nb but $\sigma_{bb}/\sigma_{tot} \sim 1/4$		Large cross section $\sigma_{bb} \sim 248\ \mu$b but $\sigma_{bb}/\sigma_{tot} \sim 10^{-2}$	
Mostly $B^{+/0}$		Not only $B^{+/0}$: B_s, B_c, Λ_b	Better on heavy hadrons
Efficient, simple trigger		Complex triggers	
Momentum conservation, ~ hermetic detector		p_T conservation, no hermeticity	Higher sensitivity for modes with muons
Similar performance for and e and μ	LFU tests	Better performance for μ than for e	
High neutrals efficiencies		Poor neutrals efficiencies	
B meson decay lengths: hundreds of μm		B meson decay lengths: mm	Good separation between vertices





SuperKEKB vs LHC

superKEKB

Clean: only 1 B-$\bar{B}$ pair,

Constrained kinematics: known $E_{CMS}(B)$

$\sim$ 60 B's per sec, $\sim 1/4$ of total events
high reconstruction efficiency, "no" trigger

Ideal for decays with π^0, γ, ν

$p(B) \sim 1.5$ GeV
flight distance ~ 0.1 mm
$\Rightarrow$ decay-time resolution ~ 0.30 ps

LHC

B hadrons $+\ \mathcal{O}(100)$ charged particles

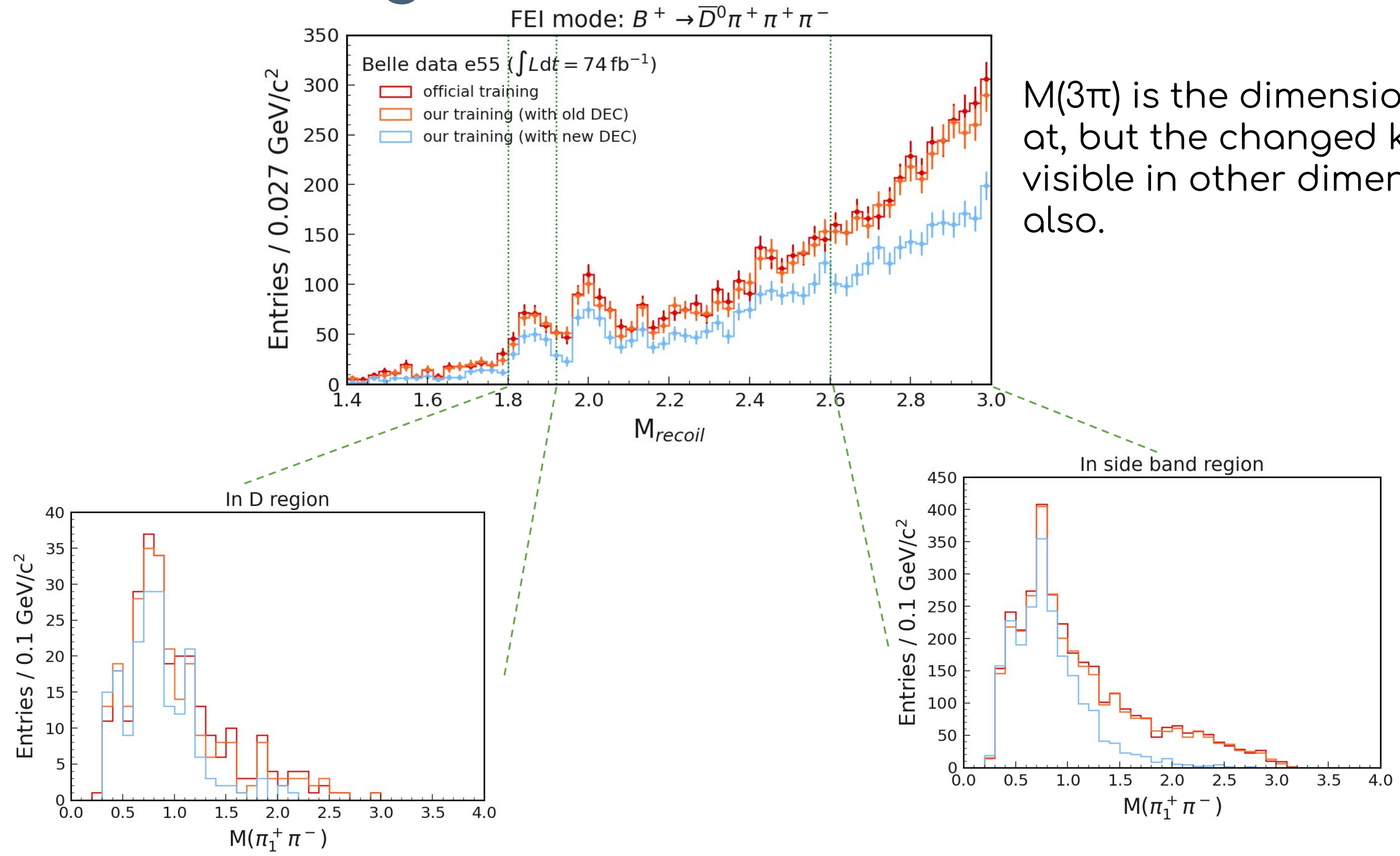
Unconstrained kinematics

$\sim$ 20'000 B's per sec., 1% of total events
low reconstruction efficiency, need trigger

Ideal for very rare decays to charged particles

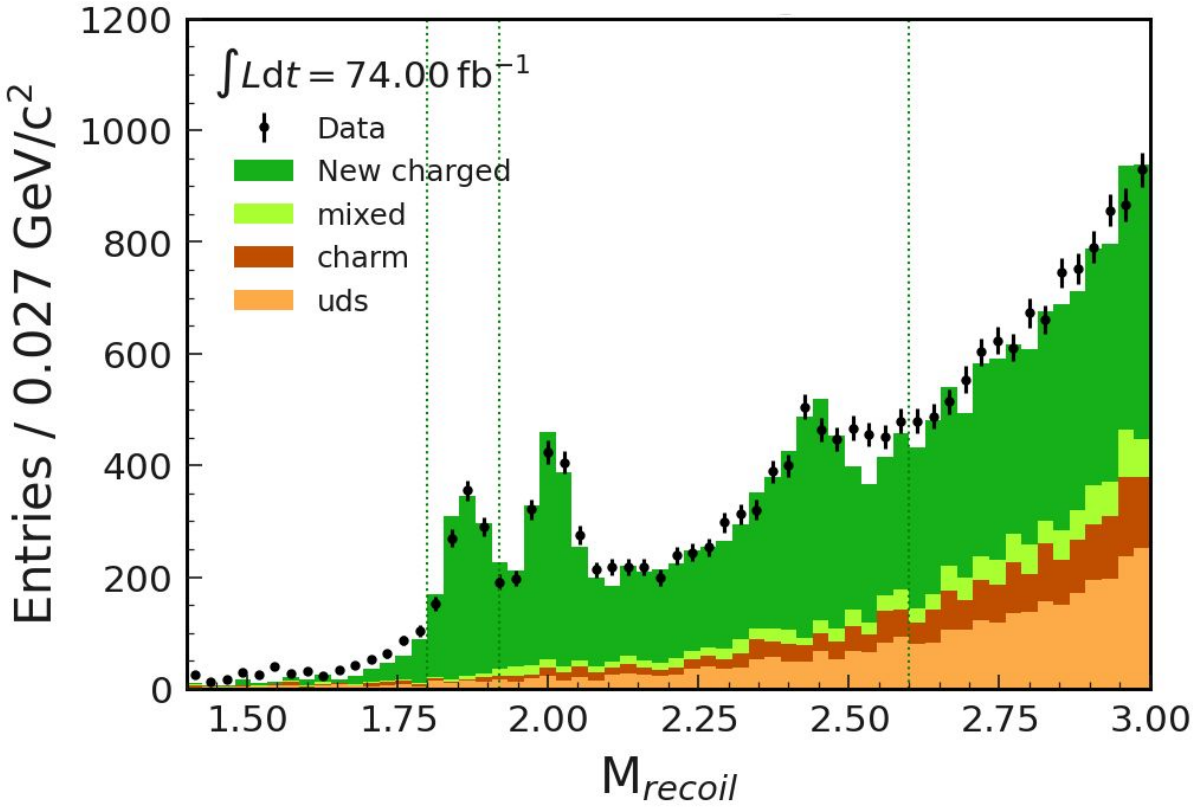
$p(B) \sim 100$ GeV
flight distance ~ 1 cm
$\Rightarrow$ decay-time resolution ~ 0.05 ps

Retraining FEI: Effective cuts



M(3π) is the dimension we usually look at, but the changed kinematics is visible in other dimensions like M(2π) also.

Retraining FEI: Data-MC agreement

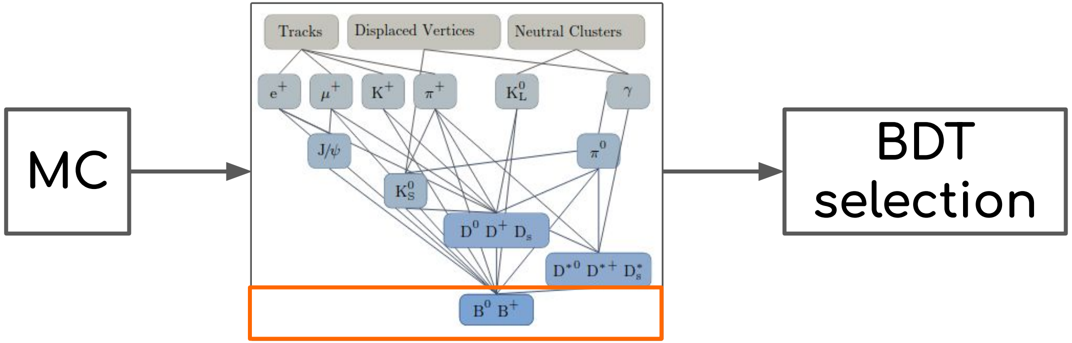


After reconstructing all MC and data with the training based on new DEC, the Data - MC agreement improves too! (even at higher M_{recoil} !)

Improving B^+ tagging

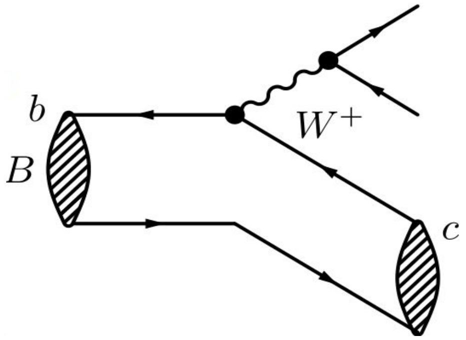
- Training is done on MC. If MC does not resemble data:
  - Biases enter in selection conditions.
  - The efficiency looks different in MC and data.

We are studying the main modes of hadronic tag and improving their MC model to look closer to data.



- Can we replace the last stage (B^+ reconstruction) BDT → cut-based
  to avoid (re)training-time and be more robust?

Model for $B \rightarrow D^{(*,**)}\, n\pi\, m\pi^0$ 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 want to 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:

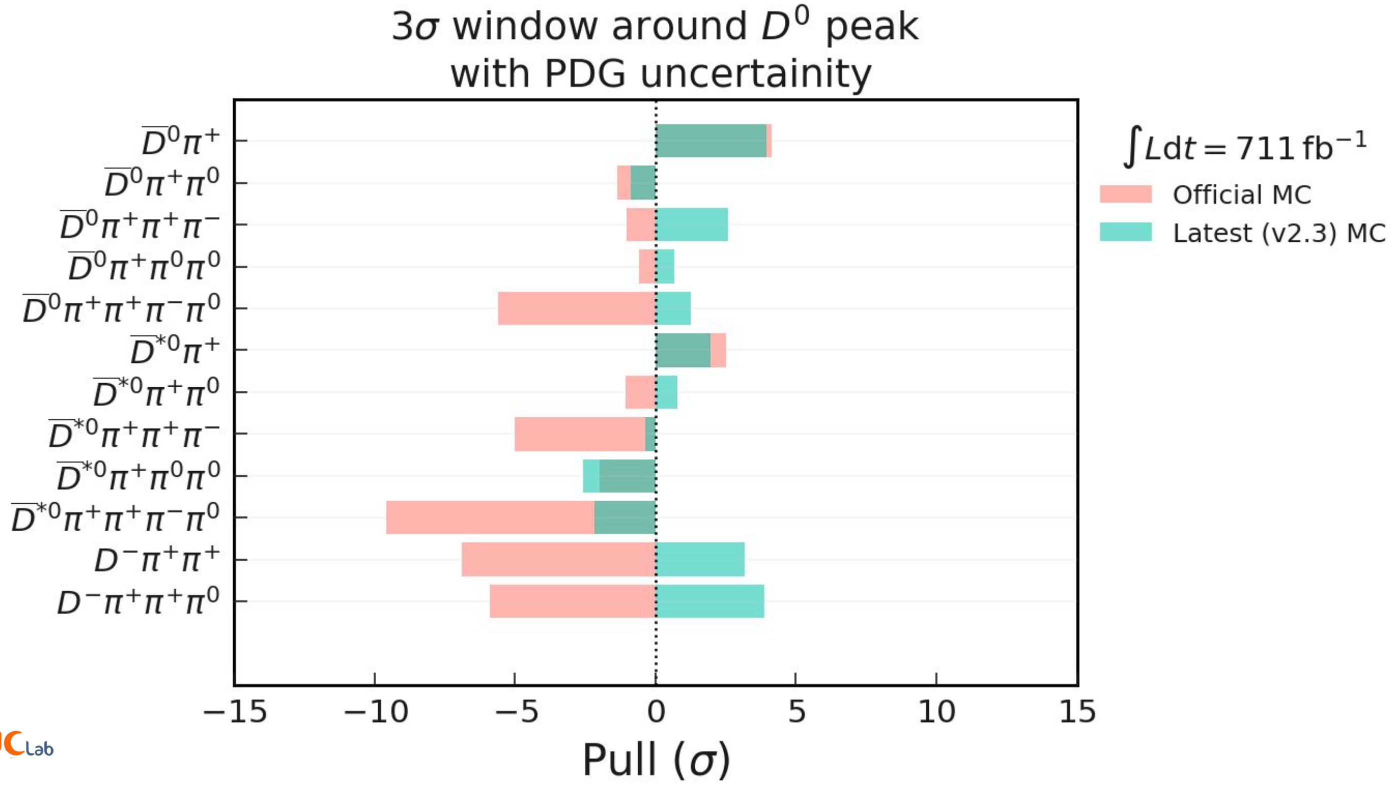
- 3π π^0 is hard to model without some sort of ρ' resonance
  - For **ωπ** we fix from measurements.
  - For **ρππ** and **ηπ**, we let PYTHIA generate it.
- Decays of D** particles is synchronized with Belle II
- The fraction of 4 different D** is fixed based on observations.

*See backup

Pulls of calibration factors

per mode

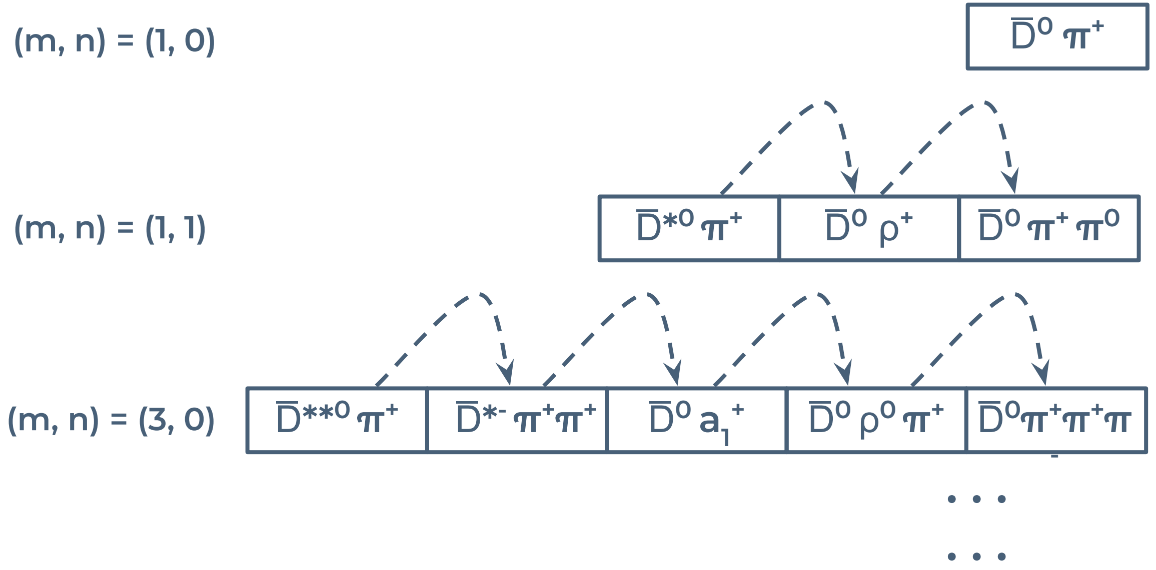
Another way to visualize the improvement in the calibration factors:



Alternative FEI algorithm

Alternatively, using FEI
particle list of $\bar{D}^0$,
we want to reconstruct B^+
particle list manually

in orders of $\bar{D}^0$ **($m\,\pi^+$) ($n\,\pi^0$)**:



Reconstructing in this order,
going to the next step only if it fails, ⇒ Simpler best candidate selection
using the constraints of intermediate resonances when possible ⇒ Higher purity

Let's call this algorithm “ FREE ”

LFU violation in b → s $l^+ l^-$: Projection

- Belle II, enjoys nearly symmetric electron/muon reconstruction performance, and can:
    - provide independent check of $R(K^{(*)})$ anomalies with > 5-10 ab^{-1}

The Belle II Physics Book, PETP 2019, 123C01 (2019)