Per-event flavor tagger

M. Bertemes, S. Raiz

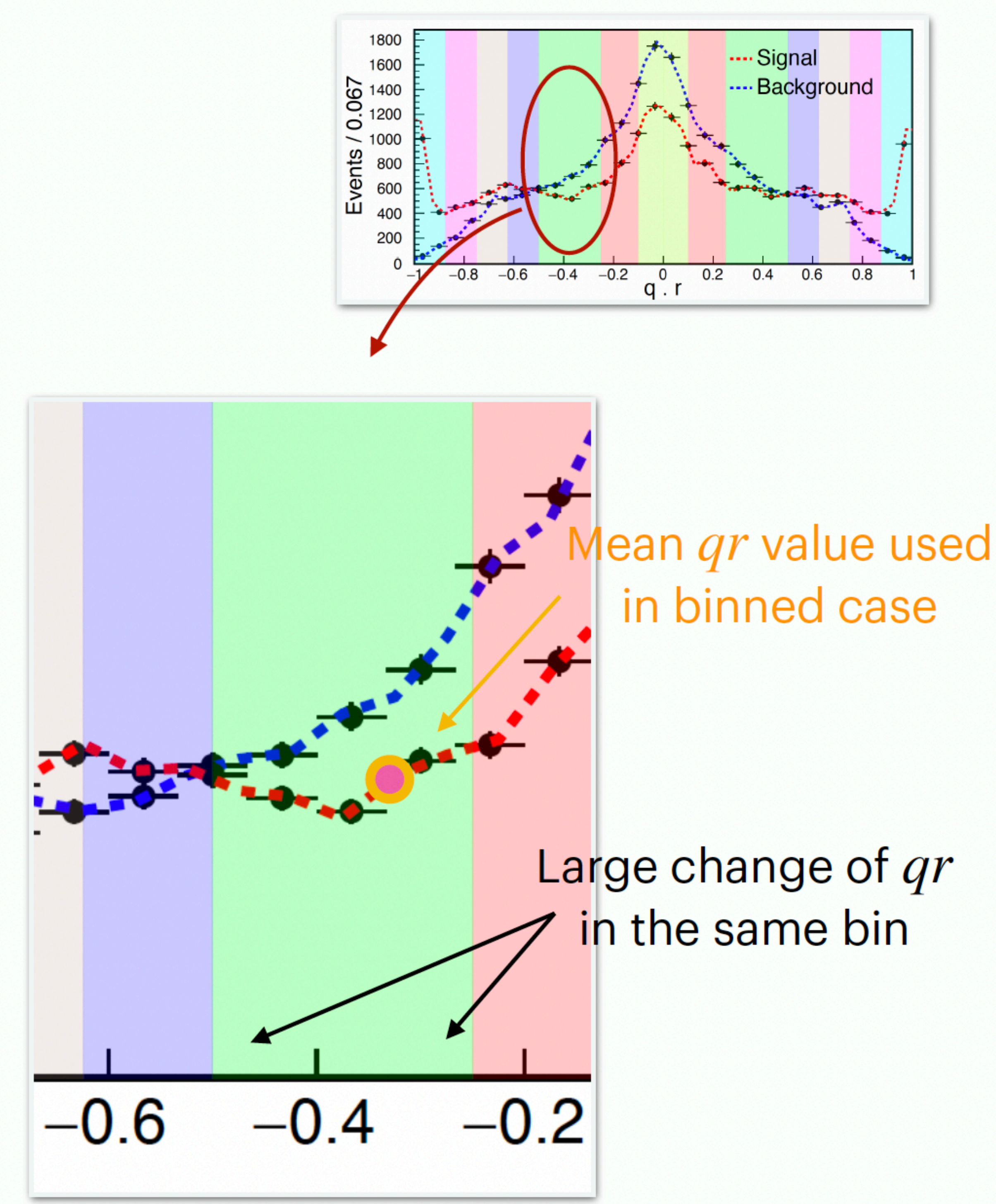
Flavor tagger in Belle II

qr in Belle II analyses was always binned.

qr can change a lot inside the same bin: taking mean value can be inaccurate.

Now, $B^0 \to \pi^0\pi^0$ and $D^0 \to \pi^0\pi^0$ analyses starting to explore use of event-per-event qr: not only maximally exploiting flavor tagger information (slightly better tagging-efficiency), but also further signal/background separation.

qr (or w) is directly taken from data and is an additional fit variable.



Binned → Unbinned

Investigated methods still not working: value is biased or no fit convergence. Technical or conceptual issue in Likelihood?

$$\mathcal{L} = \prod_{b=1}^{\text{Bins}}\left[\prod_{i=1}^{\text{Cand}}\left[BF\cdot\varepsilon_{sig}\cdot\varepsilon^{b}_{sig}\cdot n_{B\bar{B}}\cdot[1+(1-2\chi_d)[q(1-2w_b)]A_{CP}]\cdot\mathscr{P}_{sig}\right]\right]$$

Fraction of events in the bin b

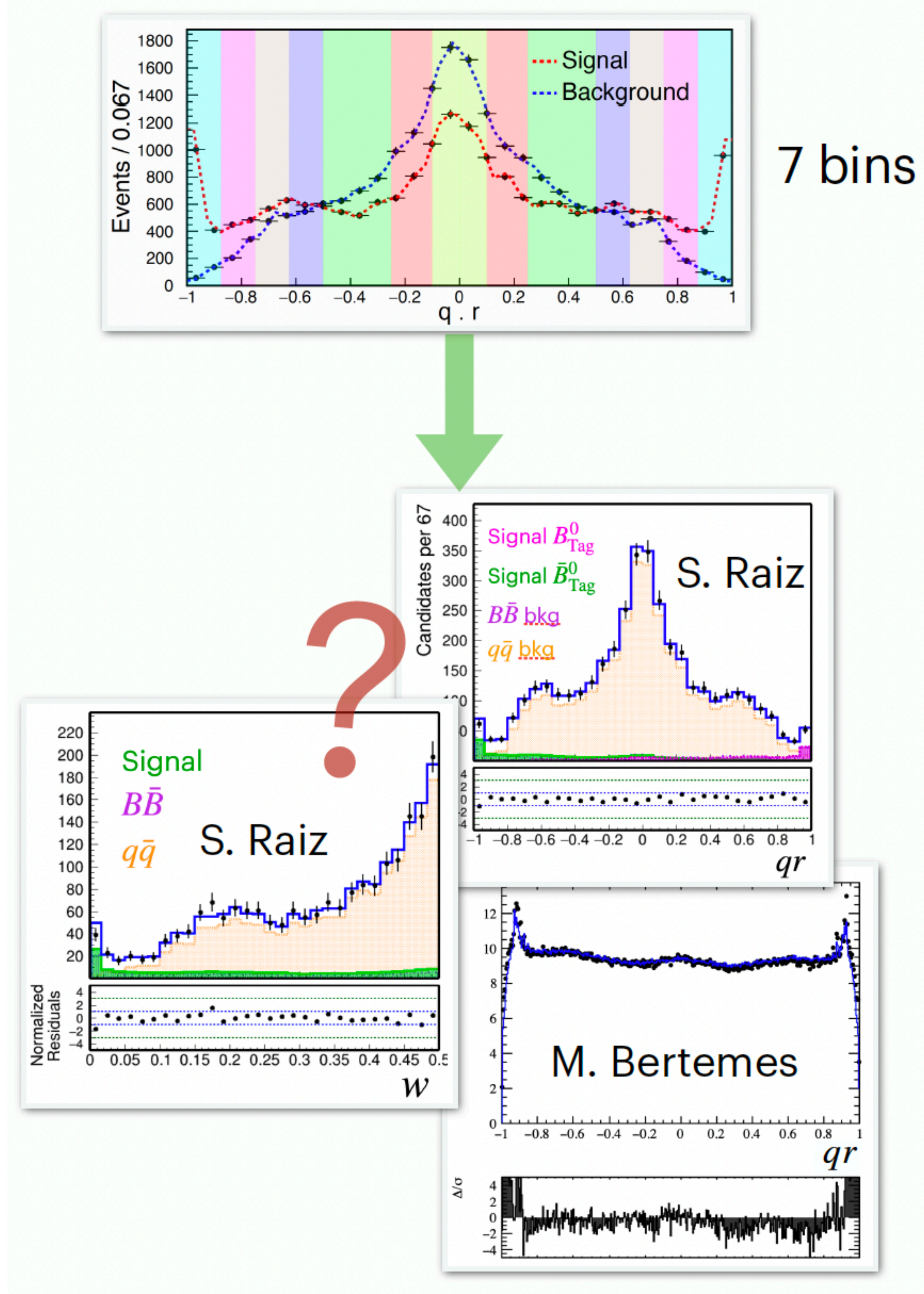
Mean wrong-tag for the bin b

$$\mathcal{L} = \prod_{i=1}^{\text{Cand}}\left[BF\cdot\varepsilon_{sig}\cdot n_{B\bar{B}}\cdot[1+(1-2\chi_d)[q(1-2w)]A_{CP}]\cdot\mathscr{P}_{sig}\cdot\mathscr{P}_{sig}(w)\right]$$

Mean wrong-tag for the event

Pdf of w

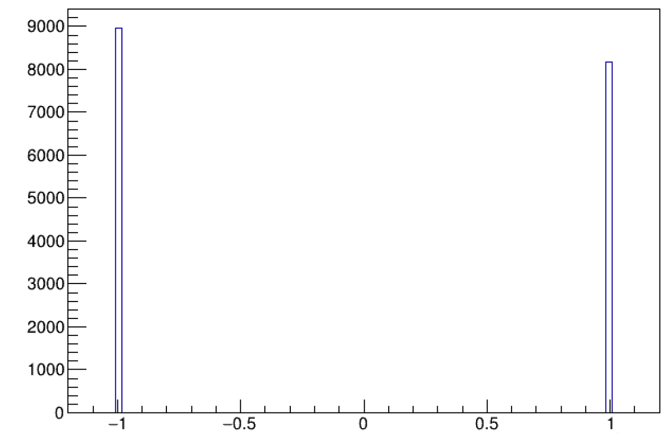
Simplified Likelihood with only signal, $\Delta w = 0, \mu = 0$



Start from the basics

Let's consider the B^+B^- case:

$$N^{\pm} = \frac{N_{\text{tot}}}{2}(1 \mp A_{CP})$$

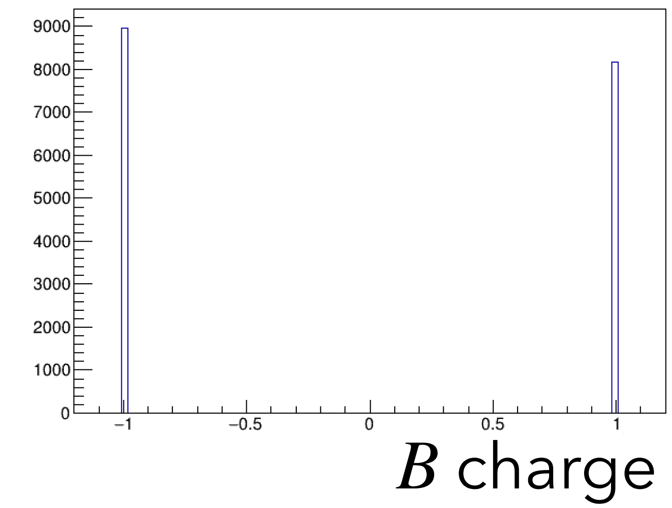


B charge

Start from the basics

Let's consider the B^+B^- case:

$$N^{\pm} = \frac{N_{\text{tot}}}{2}(1 \mp A_{CP})$$



Pass to the $B^0\overline{B}^0$ case, considering a perfect flavor tagger:

$$N^{B^0,\overline{B}^0} = \frac{N_{\text{tot}}}{2}(1 + q \cdot A_{CP})$$

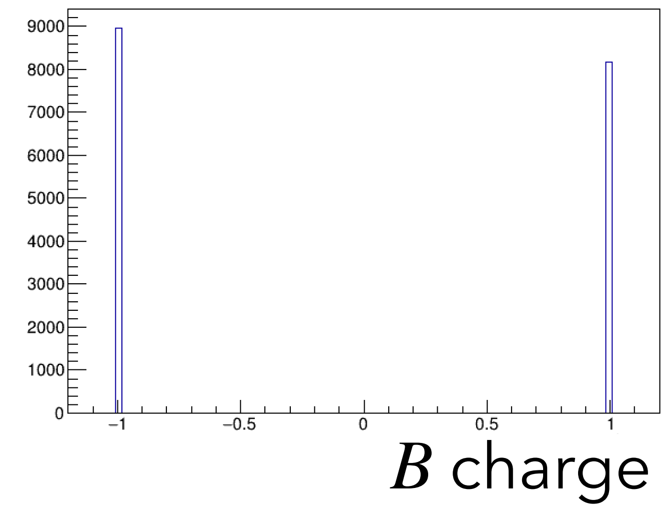
q = charge of B_{tag}

Start from the basics

Let's consider the B^+B^- case:

1

$$N^{\pm} = \frac{N_{\text{tot}}}{2}(1 \mp A_{CP})$$



Pass to the $B^0\overline{B}^0$ case, considering a perfect flavor tagger:

2

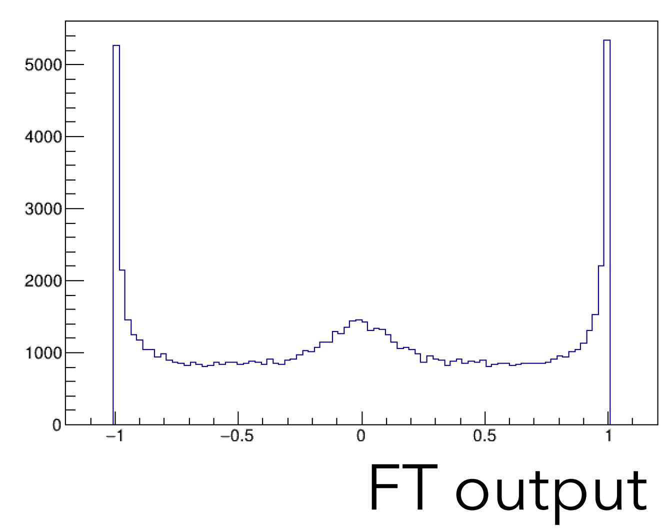
$$N^{B^0,\overline{B}^0} = \frac{N_{\text{tot}}}{2}(1 + q \cdot A_{CP})$$

q = charge of B_{tag}

In reality, there is some dilution factor r
(let's not consider Δw, μ, …):

3

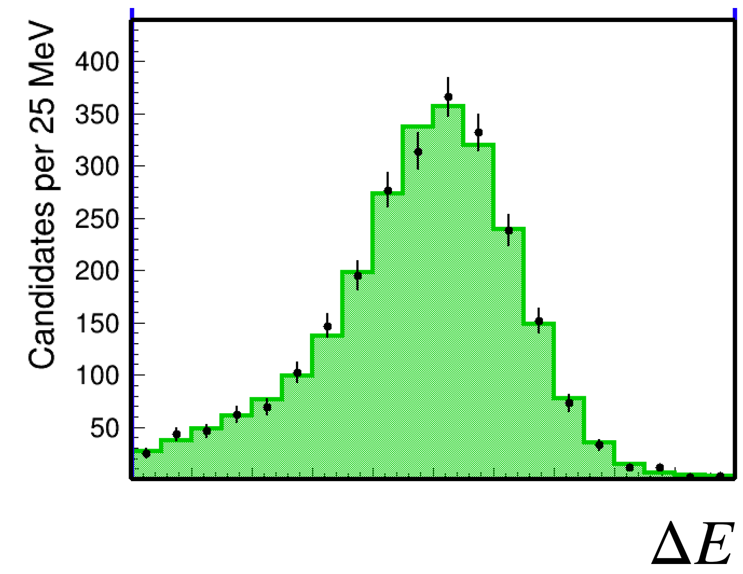
$$N^{B^0,\overline{B}^0} = \frac{N_{\text{tot}}}{2}(1 + qr \cdot A_{CP})$$



Toy fitter

Consider a very simple fitter. Fit signalMC using ΔE as only variable.

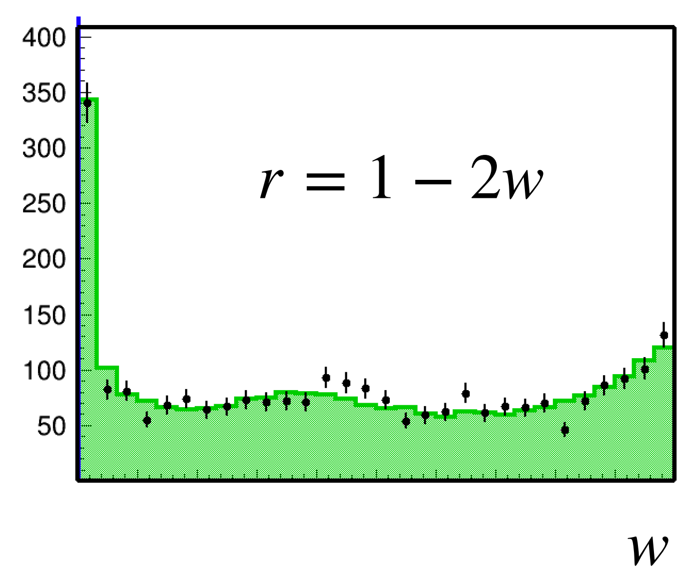
$$\mathscr{L} = \prod_{\mathrm{i=1}}^{\mathrm{Cand}} \left[N_{\mathrm{sig}} \cdot \mathscr{P}_{sig}(\Delta E) \right]$$



Add an asymmetry A:

2

$$\mathscr{L} = \prod_{\mathrm{i=1}}^{\mathrm{Cand}} \left[N_{\mathrm{sig}} \cdot \left[1 + q \cdot A \right] \cdot \mathscr{P}_{sig}(\Delta E) \right]$$



and a diluition factor r:

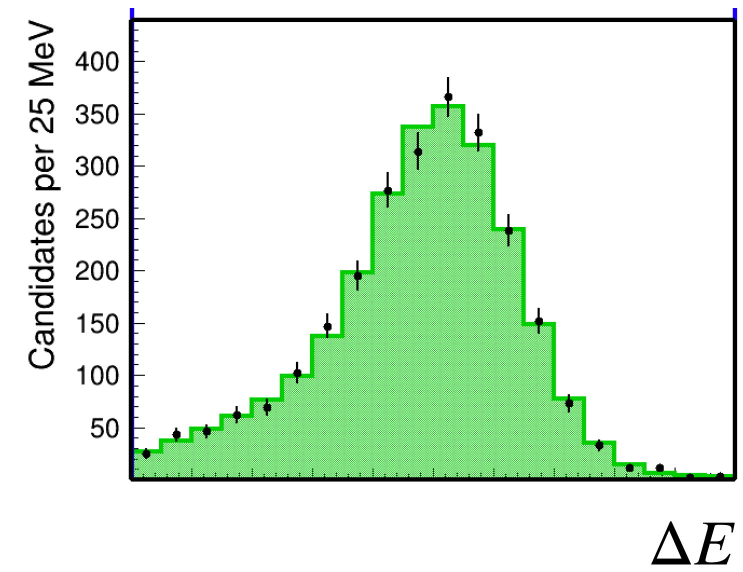
3

$$\mathscr{L} = \prod_{\mathrm{i=1}}^{\mathrm{Cand}} \left[N_{\mathrm{sig}} \cdot \left[1 + qr \cdot A \right] \cdot \mathscr{P}_{sig}(\Delta E) \cdot \mathscr{P}_{sig}(qr) \right]$$

Toy fitter

Consider a very simple fitter. Fit signalMC using ΔE as only variable.

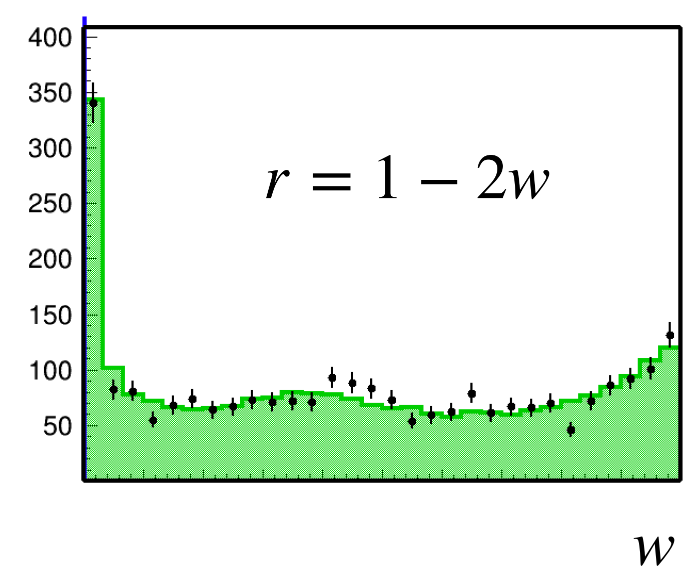
$$\mathscr{L} = \prod_{\text{i=1}}^{\text{Cand}} \left[N_{\text{sig}} \cdot \mathscr{P}_{sig}(\Delta E) \right]$$



Add an asymmetry A:

2 $\mathscr{L} = \prod_{\text{i=1}}^{\text{Cand}} [\ldots]$

A results are biased wrt true value



and a diluition factor r:

3 $\mathscr{L} = \prod_{\text{i=1}}^{\text{Cand}} [\ldots \mathscr{P}_{sig}(qr)]$

A results are biased wrt true value

Differences btw (1) and (2)

First two cases seem identical, but fitter implementation is different:

(1) B^+B^- case

$$\mathscr{L} = \prod_{\text{i=1}}^{\text{Cand}^{+1}} \left[N_{\text{sig}} \cdot \left[\frac{1-A}{2} \right] \cdot \mathscr{P}_{sig}(\Delta E) \right] \times \prod_{\text{i=1}}^{\text{Cand}^{-1}} \left[N_{\text{sig}} \cdot \left[\frac{1+A}{2} \right] \cdot \mathscr{P}_{sig}(\Delta E) \right]$$

Simultaneous fit in two bins (charge=+1 and charge=-1).

(2) $B^0\overline{B}^0$ case

$$\mathscr{L} = \prod_{\text{i=1}}^{\text{Cand}} \left[N_{\text{sig}} \cdot \left[1 + q \cdot A \right] \cdot \mathscr{P}_{sig}(\Delta E) \right]$$

No simultaneous fit. q is directly inserted in $\mathscr{L}$.

Logarithms of Likelihoods mathematically equivalent, but case (2) gives biased A. Maybe some bug in the code. Will check using configuration (1) to fit $B^0\overline{B}^0$ sample.

Challenge failed?

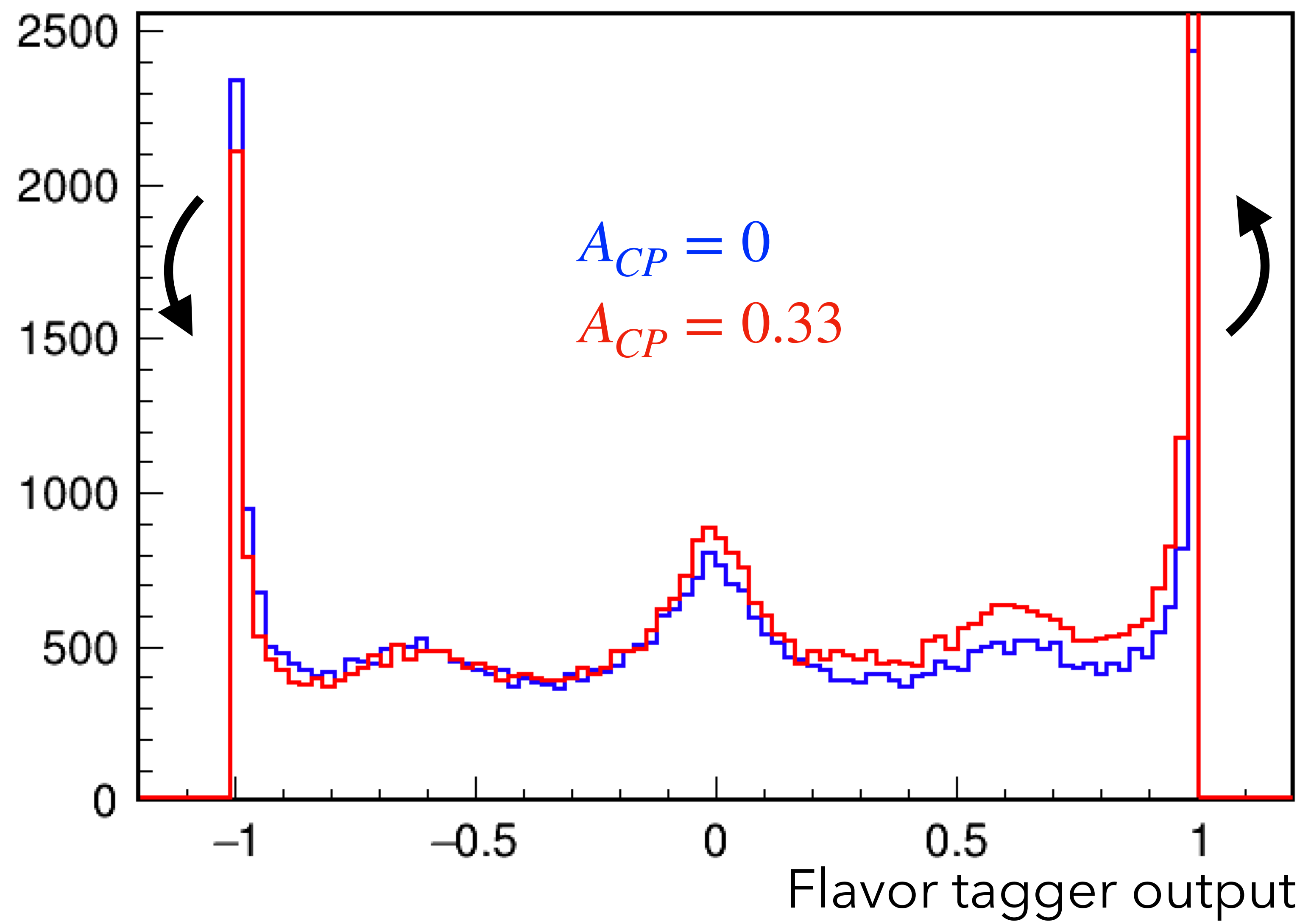
Yes, but also new ideas to investigate:

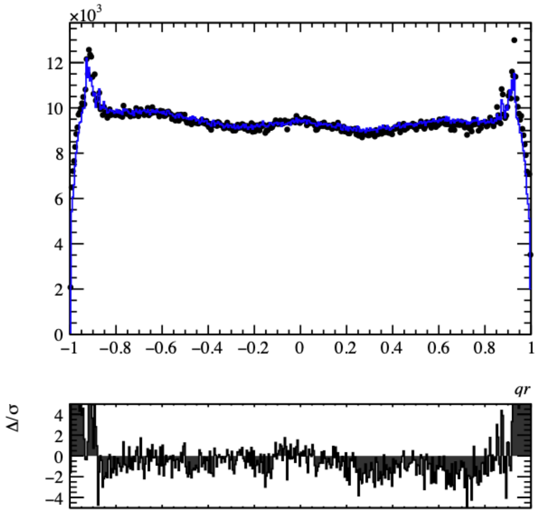
- Logarithm of Likelihoods (1) and (2) mathematically equivalent. Maybe some bug in the code. Will check using configuration (1) to fit $B^0\overline{B}^0$ sample.
- Extended ML fit does not converge, while non-extended does (with biased A). Need to understand why,

Michel's case

Use D^0 mass and qr as fit variables. Fit converges, but A_{CP} is biased because qr template has a fixed A_{CP}=0 (https://indico.belle2.org/event/9872/contributions/68321/attachments/24934/36867/b2gm_pi0pi0.pdf).

Example with $B^0 \to \pi^0\pi^0$ signalMC





Flavor tagger output

Reweight of the template inside the minimisation (based on scanned A_{CP}) could be the solution.

Backup

Calibration

Usual method employed by LHCb and charm flavor tagger:

$$qr_{\mathrm{event}} = p_1 + p_2 \cdot qr_{\mathrm{FBDT}}$$

Diluition

Calibration factors from control channels

Output of the flavor tagger