Development and automation of the Flavour Tagger's calibration at Belle II

GFlaT - GdR InF 2024

Giulio Dujany, Isabelle Ripp-Baudot, Corentin Santos, Petros Stavroulakis

Institut Pluridisciplinaire Hubert Curien
University of Strasbourg

08 November 2024

Table of Contents

CP-asymmetry

- Want to measure the *CP*-asymmetry

$$a^{CP} = \frac{N_{B^0} - N_{\bar{B}^0}}{N_{B^0} + N_{\bar{B}^0}}$$

⟶ Need to know the flavour of the B meson at the time of its decay

- **Self-tagged** decays: one of the final state particles gives away the flavour of the B meson
- *What if the signal is not self-tagged ?*

CP-asymmetry

- Want to measure the *CP*-asymmetry

$$a^{CP} = \frac{N_{B^0} - N_{\bar{B}^0}}{N_{B^0} + N_{\bar{B}^0}}$$

$\longrightarrow$ Need to know the flavour of the B meson at the time of its decay

- **Self-tagged** decays: one of the final state particles gives away the flavour of the B meson
- *What if the signal is not self-tagged ?*

CP-asymmetry

- Want to measure the CP-asymmetry

$$a^{CP} = \frac{N_{B^0} - N_{\bar{B}^0}}{N_{B^0} + N_{\bar{B}^0}}$$

$\longrightarrow$ Need to know the flavour of the B meson at the time of its decay

- **Self-tagged** decays: one of the final state particles gives away the flavour of the B meson
- *What if the signal is not self-tagged ?*

CP-asymmetry

- Want to measure the *CP*-asymmetry

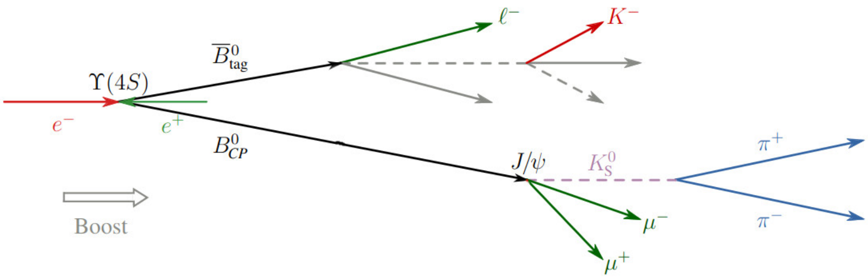
$$a^{CP} = \frac{N_{B^0} - N_{\bar{B}^0}}{N_{B^0} + N_{\bar{B}^0}}$$

$\longrightarrow$ Need to know the flavour of the B meson at the time of its decay

- **Self-tagged** decays: one of the final state particles gives away the flavour of the B meson
- *What if the signal is not self-tagged ?*

$B\bar{B}$ decays at Belle II

- Each B meson comes from $\Upsilon(4S) \longrightarrow$ Can infer the properties of the signal side from the tag side

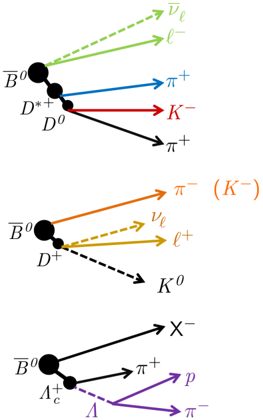


- Can rely on the B_{tag} being self-tagged

Categories...

- 13 signature categories were define:

Categories	Targets
Electron	e^-
Intermediate Electron	e^+
Muon	μ^-
Intermediate Muon	μ^+
KinLepton	e^-
Intermediate KinLepton	ℓ^+
Kaon	K^-
KaonPion	K^-, π^+
SlowPion	π^+
FastHadron	π^-, K^-
MaximumP	ℓ^-, π^-
FSC	ℓ^-, π^+
Lambda	Λ
Total= 13	



- Classification $\longrightarrow$ Classifier $\longrightarrow$ Machine Learning

$\Longrightarrow$ Introducing the FLAVOUR TAGGER

Parameters

- FLAVOUR TAGGER characterized by its **effective tagging efficiency**, ε:

$$\varepsilon = \frac{N_{B^0}^{\text{tag}} + N_{\bar{B}^0}^{\text{tag}}}{N_{B^0} + N_{\bar{B}^0}}$$

- Mistagging quantified by the **wrong tag fraction**, w:

$$N_{B^0}^{\text{tag}} = \varepsilon\left(\left(1-w\right)N_{B^0} + wN_{\bar{B}^0}\right)$$
$$N_{\bar{B}^0}^{\text{tag}} = \varepsilon\left(\left(1-w\right)N_{\bar{B}^0} + wN_{B^0}\right)$$

Observed vs True CP-asymmetry

- Observed CP-asymmetry written as

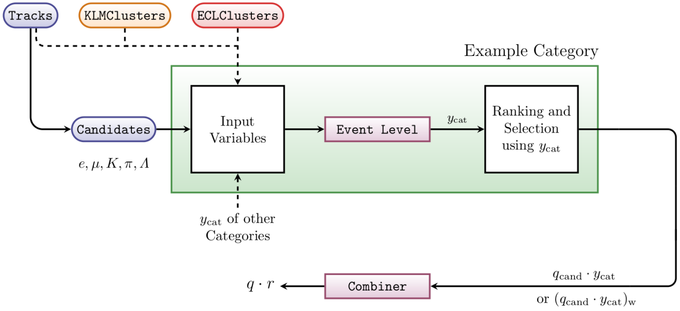
$$
\begin{aligned}
a_{\mathrm{obs}}^{CP} &= \frac{N_{B^0}^{\mathrm{tag}} - N_{\bar{B}^0}^{\mathrm{tag}}}{N_{B^0}^{\mathrm{tag}} + N_{\bar{B}^0}^{\mathrm{tag}}} \\
&= (1-2w)\,\frac{N_{B^0} - N_{\bar{B}^0}}{N_{B^0} + N_{\bar{B}^0}} \\
&= (1-2w)\,a^{CP}
\end{aligned}
$$

- *Dilution factor*: $r \equiv |1-2w| \in [0,1]$

FLAVOUR TAGGER workflow

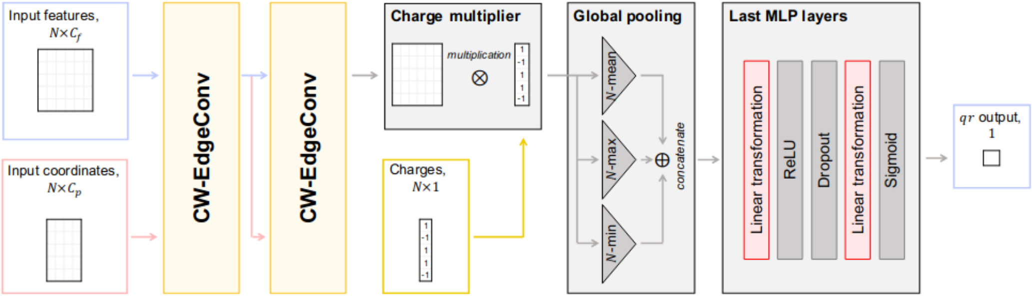
Two Boosted Decision Trees (BDTs):

1. Reconstructs and classifies the tracks among the 13 categories with a probability
2. Selects the particle with the highest probability and computes qr, where q is the flavour of the particle



A new Flavour Tagger: GFlaT

- Based on a Graph Neural Network [arXiv:2402.17260]
- Same first step as the previous Flavour Tagger, but takes into consideration all probabilities from the 13 categories to compute r



Wrong tag fraction

- Need to discriminate between:
  - Same or Opposite Flavours (SF or OF)
  - Flavour of the B_{tag} (+ or -)
- Define 7 bins of qr
- Compute the wrong tag fractions $w_i^{\pm}$ for each bin $i \in [\![0, 6]\!]$

$$w_i^{\pm} = \frac{f_i^{\pm} - R}{(1 - R)\left(1 + f_i^{\pm}\right)}$$

with

$$f_i^{\pm} = \frac{n_{\text{OF},i}^{\pm}}{n_{\text{OF},i}^{\pm} + n_{\text{SF},i}^{\pm}}, \quad R = \frac{\chi_d}{1 - \chi_d}$$

and deduce

$$\boxed{w_i = \frac{1}{2}\left(w_i^{+} + w_i^{-}\right)}$$

Effective tagging efficiency

- Define a tagging efficiency for each bin $i \in [\![0, 6]\!]$

$$\varepsilon_i^{\pm} = \frac{n_{\mathrm{OF},i}^{\pm} + n_{\mathrm{SF},i}^{\pm}}{N_{B^0}^{\mathrm{tag}} + N_{\bar{B}^0}^{\mathrm{tag}}}$$

and compute the effective tagging efficiency

$$\boxed{\varepsilon = \sum_{i=0}^{6} \varepsilon_i \left(1 - 2w_i\right)^2}$$

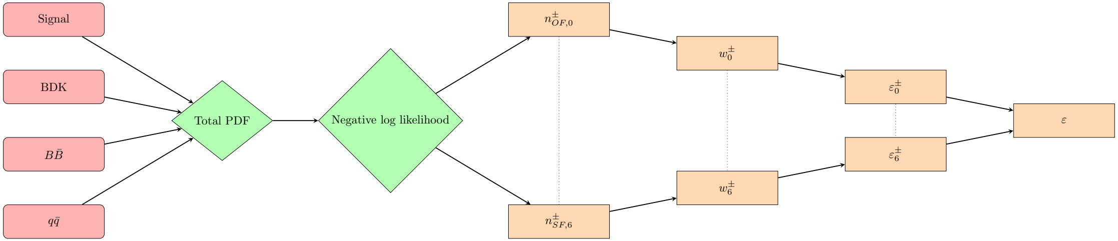
Our signal

- Need self-tagged decays:

$$B^0 \to D^-\pi^+$$
$$B^0 \to D^{*-}(\to \bar{D}^0(K^+\pi^-)\pi^-)\pi^+$$
$$B^0 \to D^{*-}(\to \bar{D}^0(K^+\pi^-\pi^0)\pi^-)\pi^+$$
$$B^0 \to D^{*-}(\to \bar{D}^0(K^+\pi^-\pi^-\pi^+)\pi^-)\pi^+$$

- Also consider $B \to D^{(*)-}K^+$ equivalents, called "BDK"s

Calibration workflow



Technical details

- Studied the correlations for a simultaneous fit
- Reduced $q\bar{q}$ background using a new cut on a geometric variable
- Changed the systematics computation strategy for bootstrapping
- Automated the workflow from reconstruction to calibration using b2luigi

Final results

Comparing results between strategies:

$\varepsilon[\%]$	Previous strategy	New strategy
Monte Carlo	40.1	41.5
Data	37.4	37.5

Conclusion

- Calibration of the Flavour Tagger is now fully automated
- Issue with the reconstruction taking longer than what b2luigi can manage
- While this work was carried out exclusively for the time-independent case, it will be of use for time-dependent *CP*-violation study too (Hawaii Belle II team)