Calibrating the hadronic FEI

From tagging to calibration factors

Maj Leandro Klančnik Kunstelj

Jožef Stefan Institute,
Ljubljana

September 30, 2026



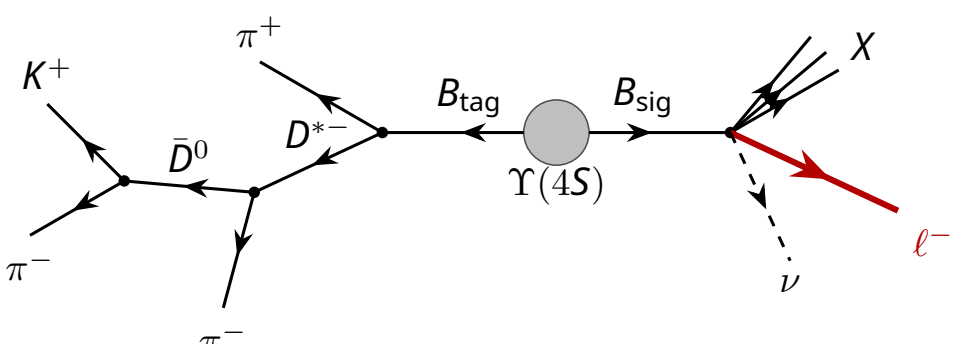


Part I

Tagging and the FEI

Events at Belle II

- Underlying physics processes: $e^+e^- \to \Upsilon(4S) \to B^0\bar{B}^0, B^+B^-(> 96\%)$
- 'Physics' events have *exactly two B* mesons, initial state fully known from the beam energies
- Once one B is fully reconstructed, the momentum of the other one is known: *complete* interpretation of events!
- Crucial for decays with invisible particles (ν), e.g. $B \to X\ell\nu$, $B \to K\nu\bar{\nu}$, $B \to \tau\nu$



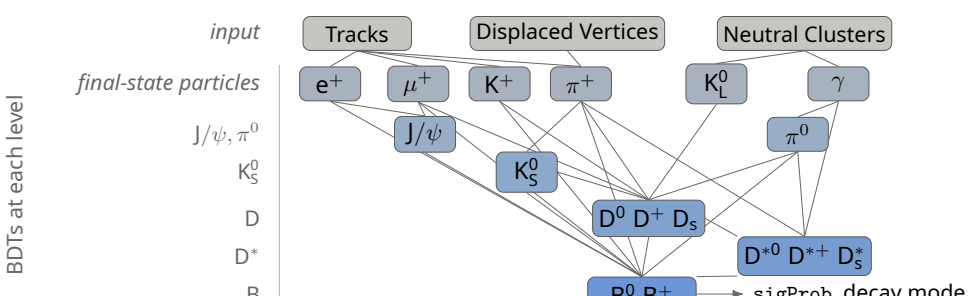
Tagging

- **Tagging**: reconstruct one B (the B_{tag}) from its decay products, and study the other one (the B_{sig}) in the rest of event (ROE)
- Two flavours:
  - ***Hadronic**: B_{tag} in fully hadronic final states – full kinematics, high purity, low efficiency ($\mathcal{O}(0.1\text{–}1\%)$)
  - **Semileptonic**: $B_{\text{tag}} \to D^{(*)}\ell\nu$ – higher efficiency, but loss of ν on the tag side
- Measuring of the hadronically decaying B_{tag} gives us B_{sig}'s momentum and charge
- (Hadronic) tag quality is judged with the beam-constrained mass and energy difference:

$$M_{bc} = \sqrt{E^{*2}_{\text{beam}} - |\vec{p}^{\,*}_B|^2}, \qquad \Delta E = E^*_B - E^*_{\text{beam}}$$

Full Event Interpretation

- Multi-stage machine-learning-based tagging algorithm
- One BDT per particle/channel; inputs additionally include the *daughters' signal probabilities* (output of the BDTs in the previous layer)
- $\mathcal{O}(100)$ explicit channels, $\mathcal{O}(10^4)$ possible chains
- Each B_{tag} gets assigned one of the pre-defined decay channels (modes)



Hadronic FEI

- Hadronic FEI: 36 B^+ and 32 B^0 decay modes – for B^+, the leading 11 carry $\sim$80% of the tagging efficiency
- Dominated by $B \to \bar{D}^{(*)} + n\pi(\pi^0)$; calibration is done per mode, others grouped into "rest"

B^+ modes		B^0 modes	
$\bar{D}^0\pi^+$	$\bar{D}^{*0}\pi^+$	$D^-\pi^+$	$D^{*-}\pi^+$
$\bar{D}^0\pi^+\pi^0$	$\bar{D}^{*0}\pi^+\pi^0$	$D^-\pi^+\pi^0$	$D^{*-}\pi^+\pi^0$
$\bar{D}^0\pi^+\pi^+\pi^-$	$\bar{D}^{*0}\pi^+\pi^+\pi^-$	$D^-\pi^+\pi^+\pi^-$	$D^{*-}\pi^+\pi^+\pi^-$
$\bar{D}^0\pi^+\pi^+\pi^-\pi^0$	$\bar{D}^{*0}\pi^+\pi^+\pi^-\pi^0$	$D^-\pi^+\pi^+\pi^-\pi^0$	$D^{*-}\pi^+\pi^+\pi^-\pi^0$
$D^-\pi^+\pi^+$	$D^-\pi^+\pi^+\pi^0$	$\bar{D}^0\pi^+\pi^-$	$\bar{\Lambda}_c^- p\pi^+\pi^-$
$\bar{\Lambda}_c^- p\pi^+\pi^-\pi^+$	*rest*	*rest*	

- Modes shown are the ones calibrated individually

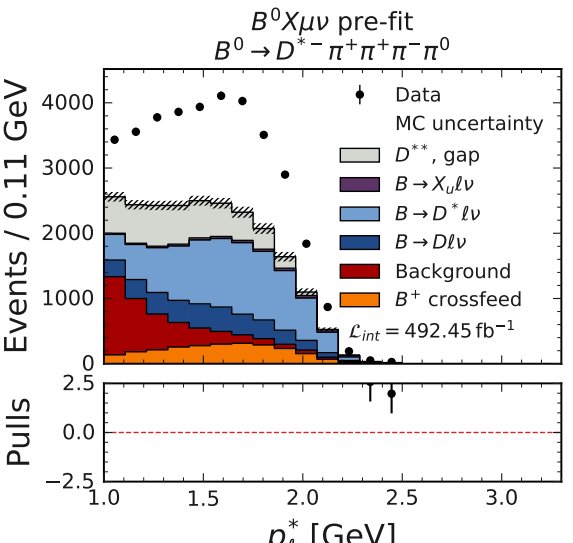
Part II

Calibration

Why calibrate?

- The FEI is trained on MC – *not* a perfect description of data!
  - Many hadronic BFs are poorly known or unmeasured
  - Imperfect decay models, resonant substructure
  - Detector response: PID, tracking, γ/π^0 reconstruction
- $\Rightarrow$ tagging efficiency in data $\neq$ in MC, varying across modes
- How to correct for this? A **calibration factor** for each mode (and `sigProb` benchmark):

$$CF = \frac{\varepsilon_{\text{tag}}^{\text{data}}}{\varepsilon_{\text{tag}}^{\text{MC}}} = \frac{N^{\text{data}}}{N^{\text{MC}}}$$



$B^0 X\mu\nu$ pre-fit, $B^0 \to D^{*-}\pi^+\pi^+\pi^-\pi^0$ tag ($CF \approx 1.85$)

Calibration strategies

- Measure N^{data} on a **calibration mode**: a B_{sig} decay with a well-known BF, reconstructed in the ROE of the B_{tag}
- Main channels: $B \to D^{(*)}\pi$ (hadronic, peaks in recoil mass) and $B \to X\ell\nu$ (inclusive semileptonic, fit in p_ℓ^*)
- Differences between data and MC not limited to mere normalization; various sources of systematic uncertainty must be considered

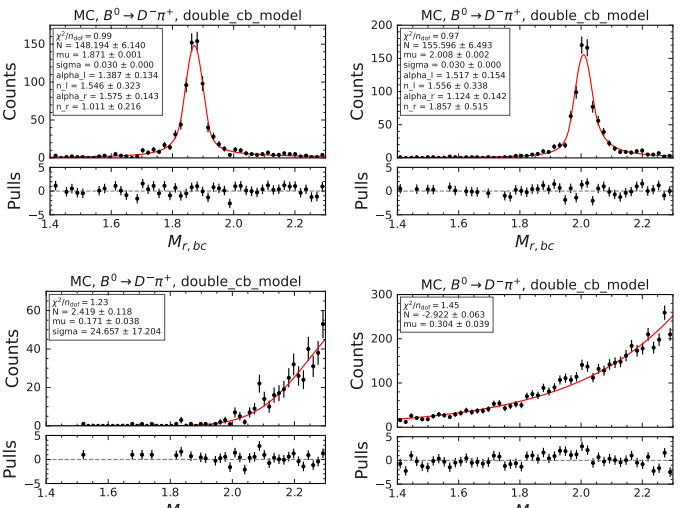
Part III

Example A: $D^{(*)}\pi$ calibration

$D\pi$ calibration

Method

- Reconstruct B_{tag} with the FEI, add a π from the ROE
- Recoil mass $M_r \equiv \sqrt{(p_{e^+e^-} - p_{B_{\text{tag}}} - p_\pi)^2}$ peaks at m_D and m_{D^*} (1.86 and 2.01 GeV)
- Fit procedure:
  - Signal shapes (double-sided CB) from truth-matched MC $D\pi$, $D^*\pi$
  - $D^{(*)}\rho$ (Gaussian) and combinatorial background (exponential)
- Then: combined fit to MC, then to data
- $CF = N^{\text{data}}_{D\pi} / N^{\text{MC}}_{D\pi}$, per FEI mode

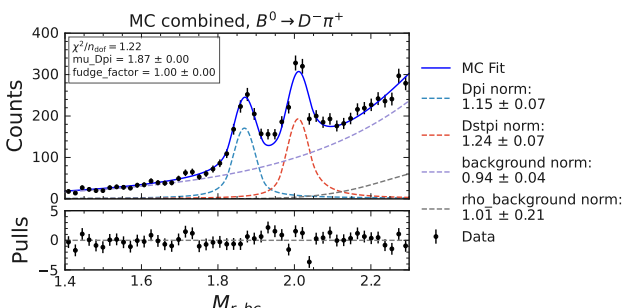


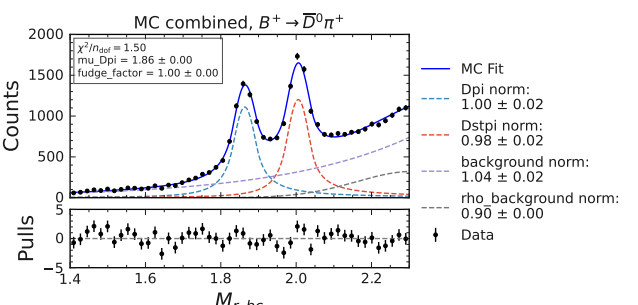
MC component fits, $B^0 \to D^-\pi^+$ tag: $D\pi$, $D^*\pi$ (top), $D^{(*)}\rho$, combinatorial (bottom)

$D\pi$ calibration

Combined MC fit

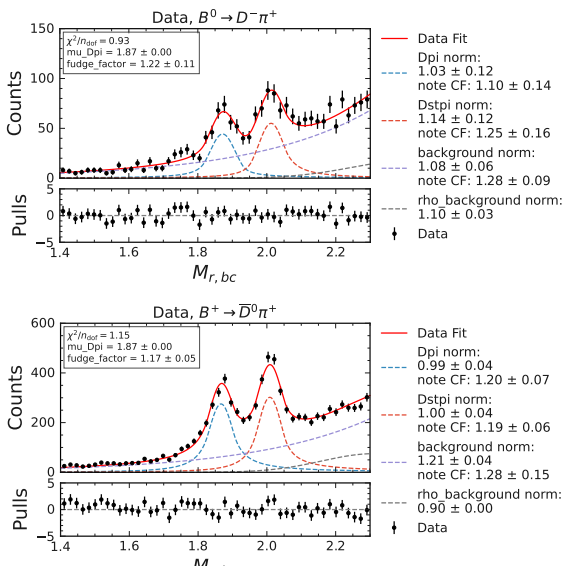
- All components summed and fit together to the full MC sample
- Shapes fixed from the component fits; floating: component normalisations, the $D\pi$ mean and a resolution fudge factor
- Constraints: $D^*\pi$ mean $= D\pi$ mean $+ \Delta M$ (PDG), and $N_{D^{(*)}\rho}/N_{D^*\pi}$ fixed to its MC ratio
- The same model, with fixed constraints, then fit to data





$D\pi$ calibration

Post-fit distributions and systematics



Data fits: $B^0 \to D^-\pi^+$ (top), $B^+ \to \bar{D}^0\pi^+$ (bottom)

Systematics: vary one fit ingredient, refit, take the shift in CF:

Source ($\times 10^{-2}$)	B^+	B^0
Fixed signal	1.0	0.2
Gaussian tail	1.1	0.6
Zero $D^{(*)}\rho$	0.9	0.6
Crossfeed	<0.1	<0.1
Random cand. sel.	1.6	1.8
Bad tag	2.7	0.3
Lumi scaling	0.9	1.0
Branching fraction	2.4	4.2
Total	4.4	4.7

From reference analysis

Part IV

Example B: $X\ell\nu$ calibration

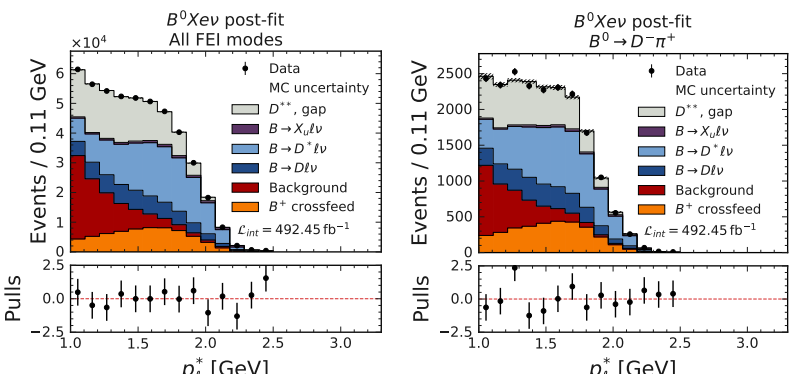
$X\ell\nu$ calibration

Method

- B_{tag} from the FEI, one lepton (e or μ) in the ROE; fit variable: lepton momentum p^*_ℓ
- Templates: $D\ell\nu$, $D^*\ell\nu$, D^{**}/gap, $X_u\ell\nu$, background (continuum + generic), crossfeed
- Binned maximum-likelihood fit (HistFactory model, `pyHF`) on the full MC; only the $X\ell\nu$ and background normalisations float
- Systematics as constrained nuisance parameters:
  - Lepton PID (correction tables) and $D^{(*)}\ell\nu$ form factors (BGL eigen-variations)
  - $X\ell\nu$ composition fractions (using HFLAV BFs) and crossfeed normalisation
  - MC statistics (per-bin uncorrelated)
- $CF = N^{\text{fit}}_{X\ell\nu}/N^{\text{MC}}_{X\ell\nu}$, uncertainty propagated from the nuisance-parameter covariance

$X\ell\nu$ calibration

Post-fit distributions and uncertainties



$B^0Xe\nu$ post-fits: all FEI modes (left), $B^0 \to D^-\pi^+$ tag (right)

Nuisance parameters absorb shape differences

Source (%)	all	$D^-\pi^+$
MC statistics	0.10	0.48
$X\ell\nu$ BFs	2.70	2.70
Tracking	0.35	0.35
$D\ell\nu$ FF	0.01	0.01
$D^*\ell\nu$ FF	0.60	0.99
Lepton PID	0.62	0.62
Fit (N^{data})	4.1	7.0
Total on CF	5.0	7.7

Relative uncertainty breakdown, $B^0e\nu$, $\texttt{sigProb} > 0.001$

Results: combined calibration factors

- All of the above results in per-mode CFs from $Xe\nu$, $X\mu\nu$ and $D\pi$, finally combine into a single CF per FEI mode (red)
- This is the final result from examples A and B, ready for common use as corrections
- Mostly $CF \sim 1.1$–1.4 with some outliers

