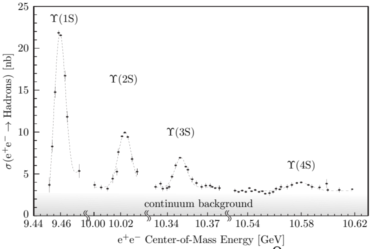
The Belle II flavor tagger

Fernando Abudinén

July 11, 2018

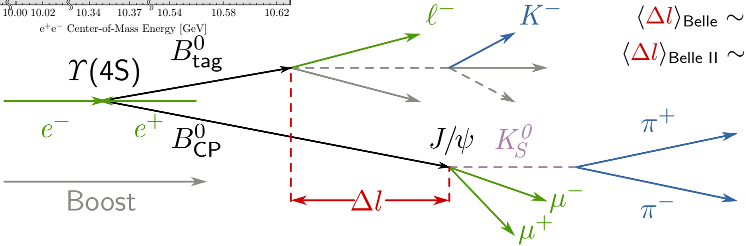


- $\Upsilon(4S)$ above $B\overline{B}$ prod. threshold
- $\Upsilon(4S) \to B\overline{B} > 96\ \%$
- $\dfrac{\Gamma(B^+B^-)}{\Gamma(B^0\overline{B}^0)} \sim 1.06$

$\Rightarrow$ B-Factory

$\langle \Delta l \rangle_{\text{Belle}} \sim 200\ \mu\text{m}$

$\langle \Delta l \rangle_{\text{Belle II}} \sim 130\ \mu\text{m}$



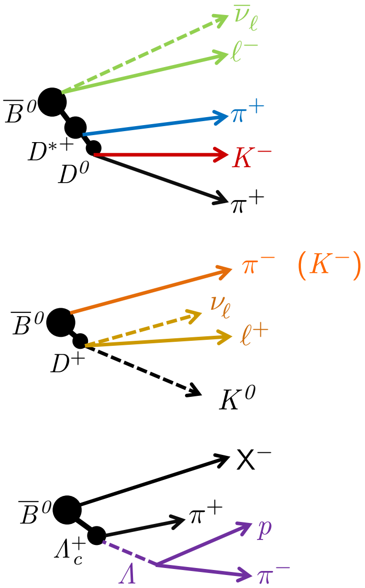
$\Rightarrow$ $\Delta t = \dfrac{\Delta z}{\langle \beta\gamma \rangle c}$ since $B^0\overline{B}^0$ at rest in $\Upsilon(4S)$ frame

$q_{B^0,\overline{B}^0} = 1, -1$

$$\mathcal{P}^{\text{Sig}}(\Delta t, q) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}} \left[1 + q\left(\mathcal{A}_{CP}\cos(\Delta m \Delta t) + \mathcal{S}_{CP}\sin(\Delta m \Delta t)\right)\right]$$

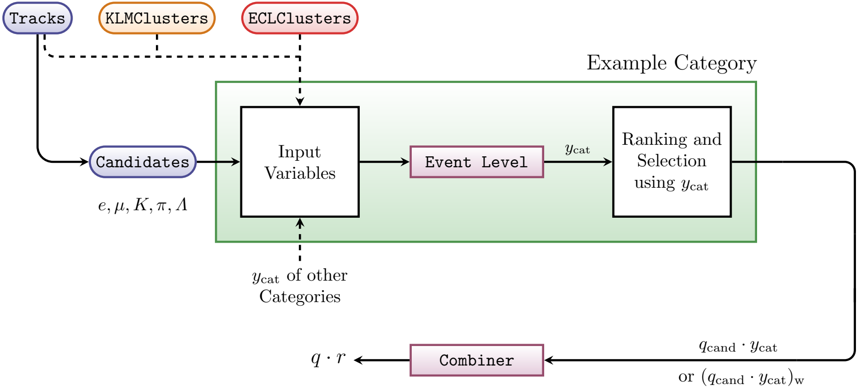
Category-based flavor tagger

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	



Category-based procedure

Starting Info: Objects in the **tag side** (rest of event).

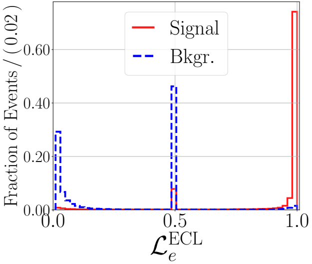


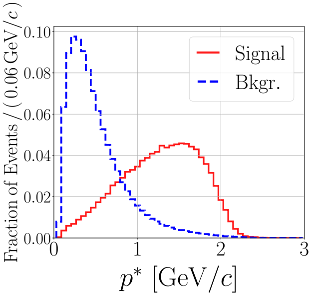
Magenta Boxes: Multi-variate Methods.

Default: Belle II's Fast-boosted decision tree. ▸ FBDT

Cross check: 3-layer Perceptron (Only Combiner) ▸ FANN Library.

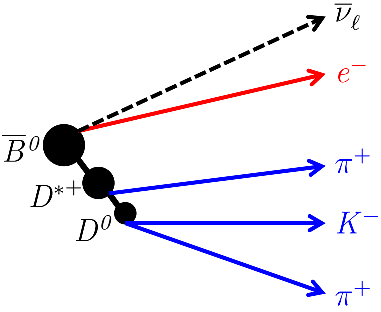
Tagging variables





Two types:

- Particle identification (PID)
- Kinematic



Tagging variables

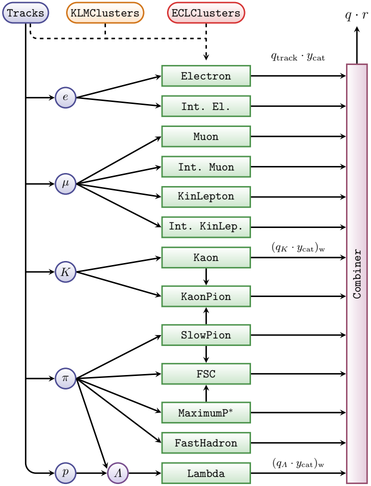
PID: Likelihoods combining info from different subdetectors.

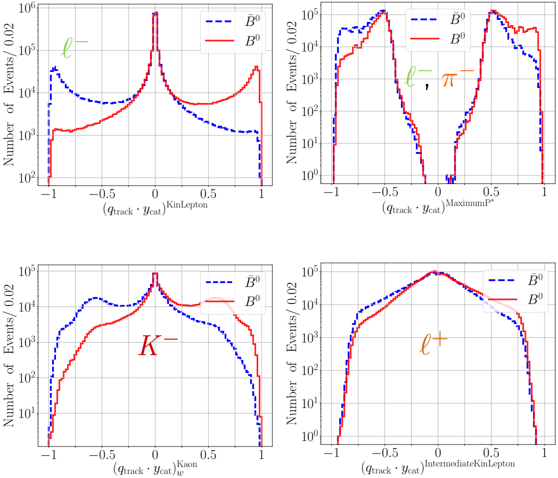
Kin.: Simple: Momentum, transverse momentum, impact params., polar angle.
Elaborated: recoil mass, energy in W boson direction, miss. momentum., cosines to thrust axis, and others.

- Total: 220 Variables. Unique variables: 108.

$\Rightarrow$ Optimized for CPU: Each Variable is calculated only once for each particle list! (108 instead of 220 calculations).

Category-based procedure

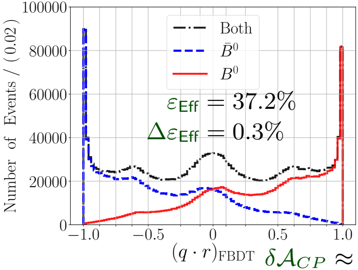




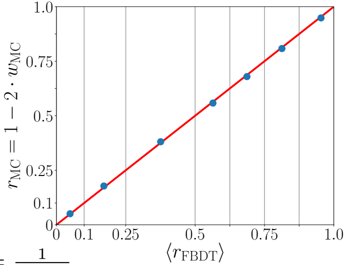
Training: Signal MC $\Upsilon(4S) \to B^0{}_1(\to J/\Psi K^0_S)\ B^0{}_2(\to \text{generic})$

Category-based tagger with Belle II MC

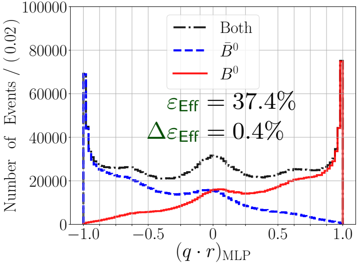
Output

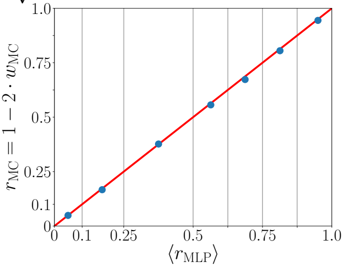


Calibration



$$\delta\mathcal{A}_{CP} \approx \frac{1}{\sqrt{\varepsilon_{\text{Eff}} \cdot N}} = \frac{1}{\sqrt{N_{\text{Eff}}}}$$

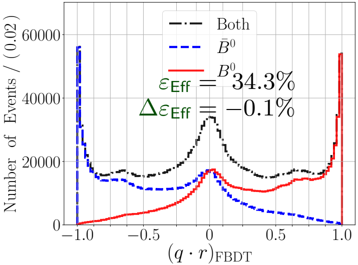


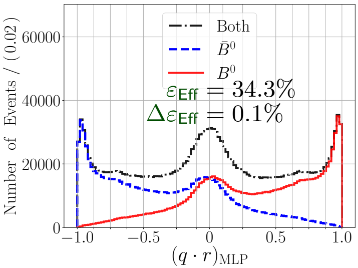


Category-based tagger with Belle MC

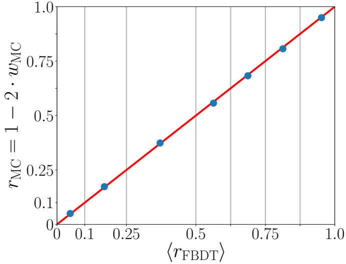
LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

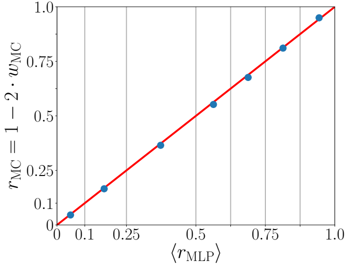
Output





Calibration



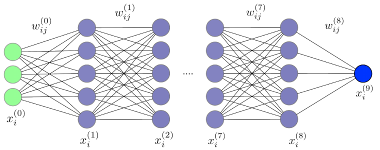


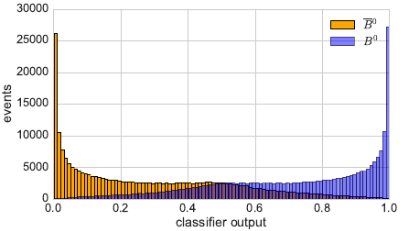
Deep-neural flavor tagger

- 10 tracks at maximum
- sorted by momentum and grouped by charge.

Input: PID variables, momentum, azim. and polar angles, impact params., hit counts in trackers. Total = 140

Training: `Pylearn2` Library ⇒ Theano⇒ GPUs.



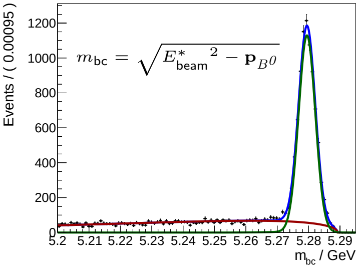


Effective efficiencies:

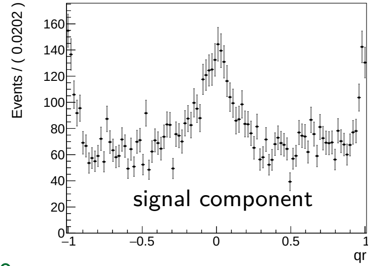
ε_{eff} [%]	MC	
	Belle II*	Belle
Category-based	37.2	34.3
Deep neural	40.7	34.4

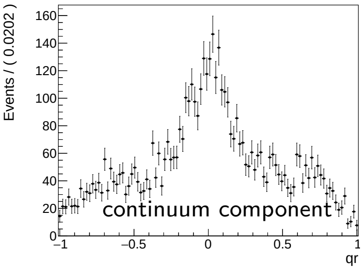
* No beam bkg. and IP always at zero.

Category-based tagger with Belle data

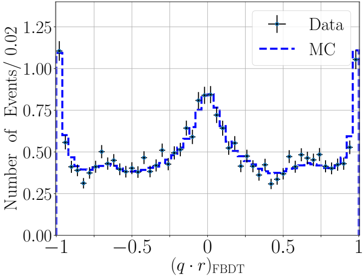


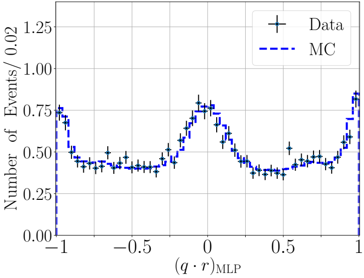
- Splot performed with converted Belle data using m_{bc} as discriminating variable.
- Full Belle 0.8 ab^{-1}
  $B^0 \to J/\psi K_S^0$





- Belle Data distribution weighted with splot output variable (signal component). B2TIP (Belle II physics book)





- Nice overlap of output distribution for Belle MC and Belle data ☺.

Belle II: $\varepsilon_{\text{Eff}} = 33.6 \pm 0.5\%$ on Belle data (Assuming MC calibration).

- $\varepsilon_{\text{Eff}}(\text{Belle}) = 30.1 \pm 0.4\%$, $\varepsilon_{\text{Eff}}(\text{BaBar}) = 33.1 \pm 0.3\%$.

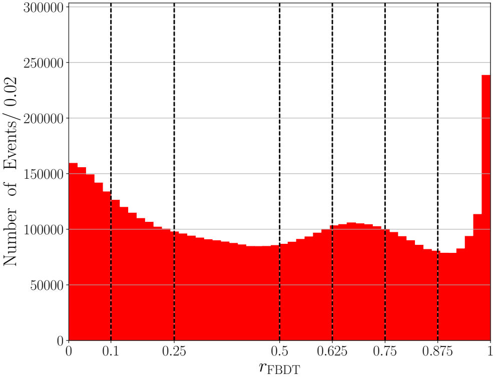
Conclusions and Outlook

- Performance of category-based and of deep-neural taggers are similar with Belle MC.
- On Belle II MC without beam bkg. and beam spread, the deep-neural algorithm performs better.
- The category-based tagger has been validated on Belle data.

$\Rightarrow$ For the deep-neural tagger is ongoing.

- Calibration of both taggers using Belle data (flavor-mixing measurement) is ongoing.
- Benchmark calibration with first $\sim 20\,\text{fb}^{-1}$ Belle II commisioning data possible (Preparation ongoing).
- We want to use both algorithms for better understanding of MC/Data differences.

Efficiency Calculation



- Binning $\Rightarrow$ correction with real data!
- Efficiency:

$$\varepsilon_{\mathsf{Eff}} = \sum_i \varepsilon_i \cdot \langle r_i \rangle^2$$

- $r_{\mathsf{MC}} = 1 - 2 \cdot w_{\mathsf{MC}}$
- Calibration: r_{MC} linear to r_{Output}

Tagging variables

Categories	Discriminating input variables
Electron Int. Electron	$\mathcal{L}_e$, p^*, p^*_t, p, p_t, $\cos\theta$, d_0, $\lvert\boldsymbol{x}\rvert$, M^2_rec, E^W_{90}, p^*_miss, $\cos\theta^*_\mathsf{miss}$, $\cos\theta^*_\mathsf{T}$, p-val.
Muon Int. Muon	$\mathcal{L}_\mu$, p^*, p^*_t, p, p_t, $\cos\theta$, d_0, $\lvert\boldsymbol{x}\rvert$, M^2_rec, E^W_{90}, p^*_miss, $\cos\theta^*_\mathsf{miss}$, $\cos\theta^*_\mathsf{T}$, p-val.
Kin. Lepton Int. Kin. Lep.	$\mathcal{L}_e$, $\mathcal{L}_\mu$, p^*, p^*_t, p, p_t, $\cos\theta$, d_0, $\lvert\boldsymbol{x}\rvert$, M^2_rec, E^W_{90}, p^*_miss, $\cos\theta^*_\mathsf{miss}$, $\cos\theta^*_\mathsf{T}$, p-v.
Kaon	$\mathcal{L}_K$, p^*, p^*_t, p, p_t, $\cos\theta$, d_0, $\lvert\boldsymbol{x}\rvert$, $n_{K^0_S}$, $\sum p^2_\mathsf{t}$, M^2_rec, E^W_{90}, p^*_miss, $\cos\theta^*_\mathsf{miss}$, $\cos\theta^*_\mathsf{T}$, χ^2
Slow Pion Fast Hadron	$\mathcal{L}_\pi$, $\mathcal{L}_\mathsf{e}$, $\mathcal{L}_K$, p^*, p^*_t, p, p_t, $\cos\theta$, d_0, $\lvert\boldsymbol{x}\rvert$, $n_{K^0_S}$, $\sum p^2_\mathsf{t}$, M^2_rec, E^W_{90}, p^*_miss, $\cos\theta^*_\mathsf{miss}$, $\cos\theta^*_\mathsf{T}$, p-val.
Kaon-Pion	$\mathcal{L}_K$, y_Kaon, y_SlowPion, $\cos\theta^*_{K\pi}$, $q_K \cdot q_\pi$
Maximum P*	p^*, p^*_t, p, p_t, d_0, $\lvert\boldsymbol{x}\rvert$, $\cos\theta^*_\mathsf{T}$
FSC	$\mathcal{L}_{K\mathsf{Slow}}$, p^*_Slow, p^*_Fast, $\cos\theta^*_\mathsf{T,\ Slow}$, $\cos\theta^*_\mathsf{T,\ Fast}$, $\cos\theta^*_\mathsf{SlowFast}$, $q_\mathsf{Slow} \cdot q_\mathsf{Fast}$
Lambda	$\mathcal{L}_\mathsf{p}$, $\mathcal{L}_\pi$, p^*_Λ, p_Λ, p^*_p, p_p, p^*_π, p_π, q_Λ, M_Λ, $n_{K^0_S}$, $\cos\theta_{\boldsymbol{x}_\Lambda,\boldsymbol{p}_\Lambda}$, $\lvert\boldsymbol{x}_\Lambda\rvert$, σ^{zz}_Λ, p-v.

Optimized for CPU: Each Variable is calculated only once for each particle list! (108 instead of 220 calculations)

5 Particle lists: e, μ, K, π, Λ