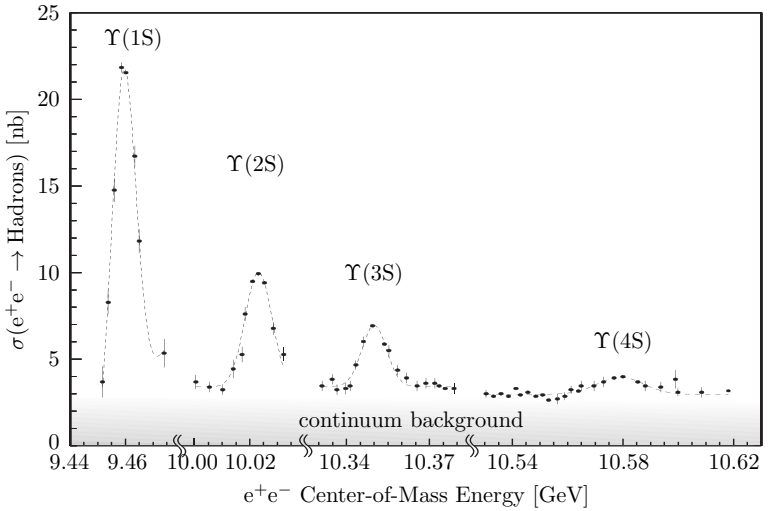
The flavor tagging module for the Belle II Experiment

Fernando Abudinén

May 11, 2015

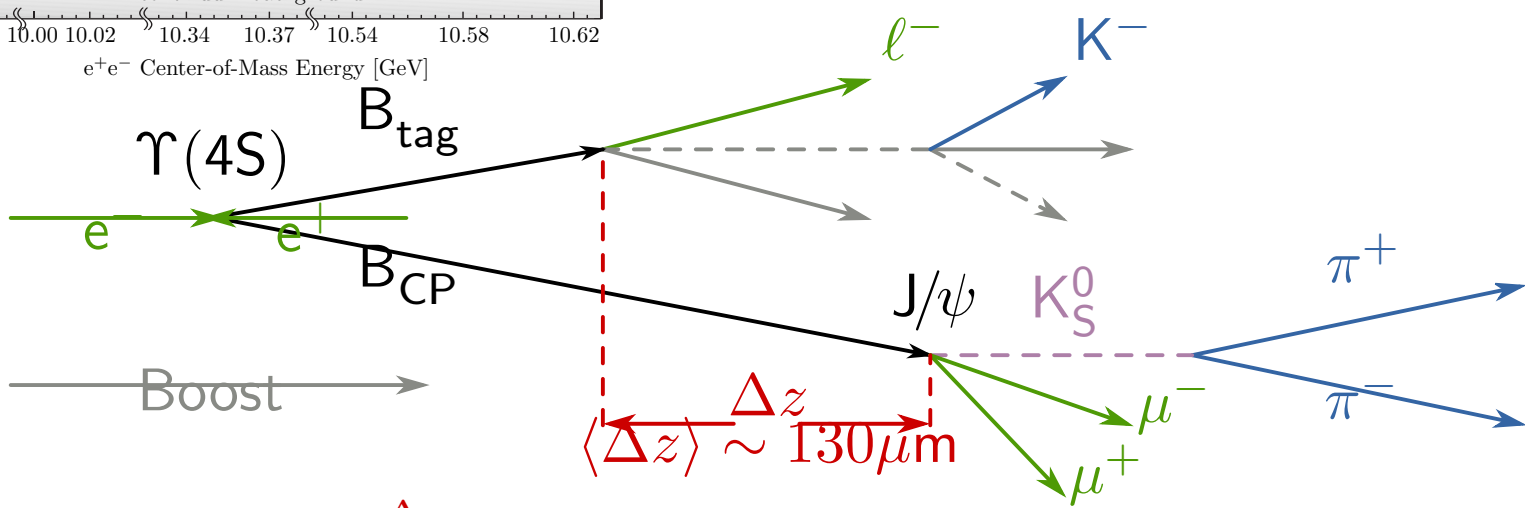
Why flavor tagging?



- $\Upsilon(4S)$ above $B\bar{B}$ prod. threshold

$\Rightarrow$ $B\bar{B}$ at rest in the $\Upsilon(4S)$ frame

- $\sim 48\ \%$ are $B^0\overline{B}^0$ $\quad q_{B^0,\overline{B}^0} = 1, -1$



- With $\Delta t = \dfrac{\Delta z}{\langle\beta\gamma\rangle c}$ $\Rightarrow$

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

Why flavor tagging?

- Only a fraction ε of events is able to be tagged (decays without signatures), and
- a fraction w of them is wrongly classified $\Rightarrow$

$$\mathcal{P}^{\mathsf{Obs}}(\Delta t, q, \varepsilon, w) = \varepsilon \cdot \Big[(1-w)\cdot \mathcal{P}^{\mathsf{Sig}}(\Delta t, q) + w \cdot \mathcal{P}^{\mathsf{Sig}}(\Delta t, -q)\Big]$$

$$= \frac{e^{-|\Delta t|\tau_{\mathsf{B}^0}}}{4\tau_{\mathsf{B}^0}}\varepsilon\left[1 + q\,(1-2w)\,(\mathcal{A}_{CP}\cos(\Delta m \Delta t) + \mathcal{S}_{CP}\sin(\Delta m \Delta t))\right]$$

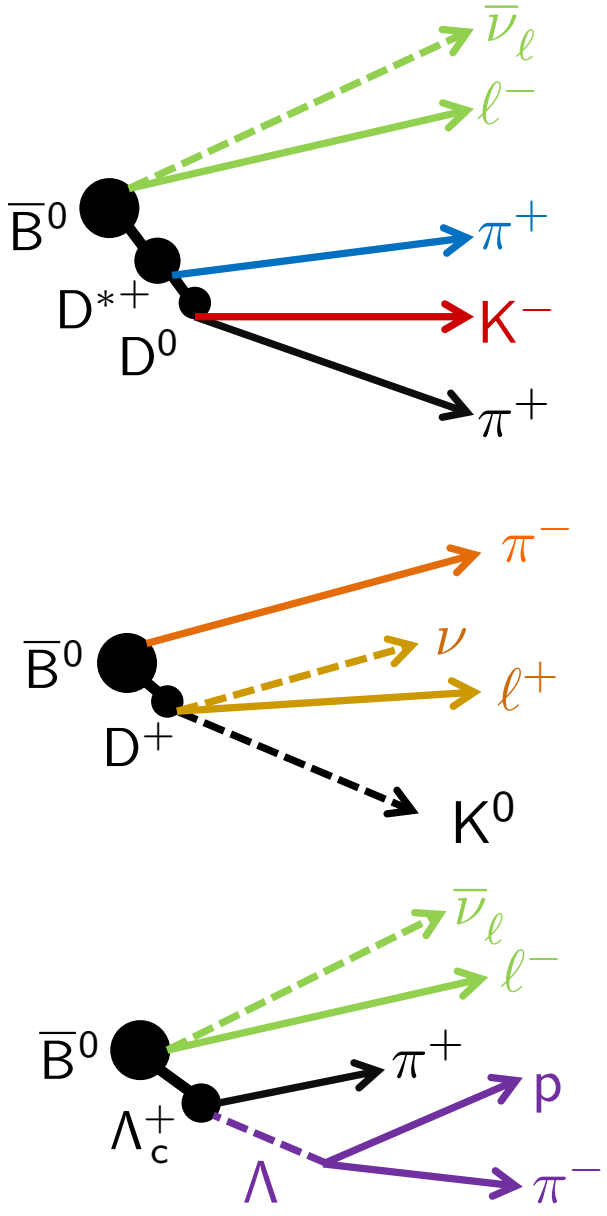
$$= \frac{e^{-|\Delta t|\tau_{\mathsf{B}^0}}}{4\tau_{\mathsf{B}^0}}\varepsilon\left[1 + q\cdot r\,(\mathcal{A}_{CP}\cos(\Delta m \Delta t) + \mathcal{S}_{CP}\sin(\Delta m \Delta t))\right]$$

- Statistical significance $\sim r\sqrt{\varepsilon}$

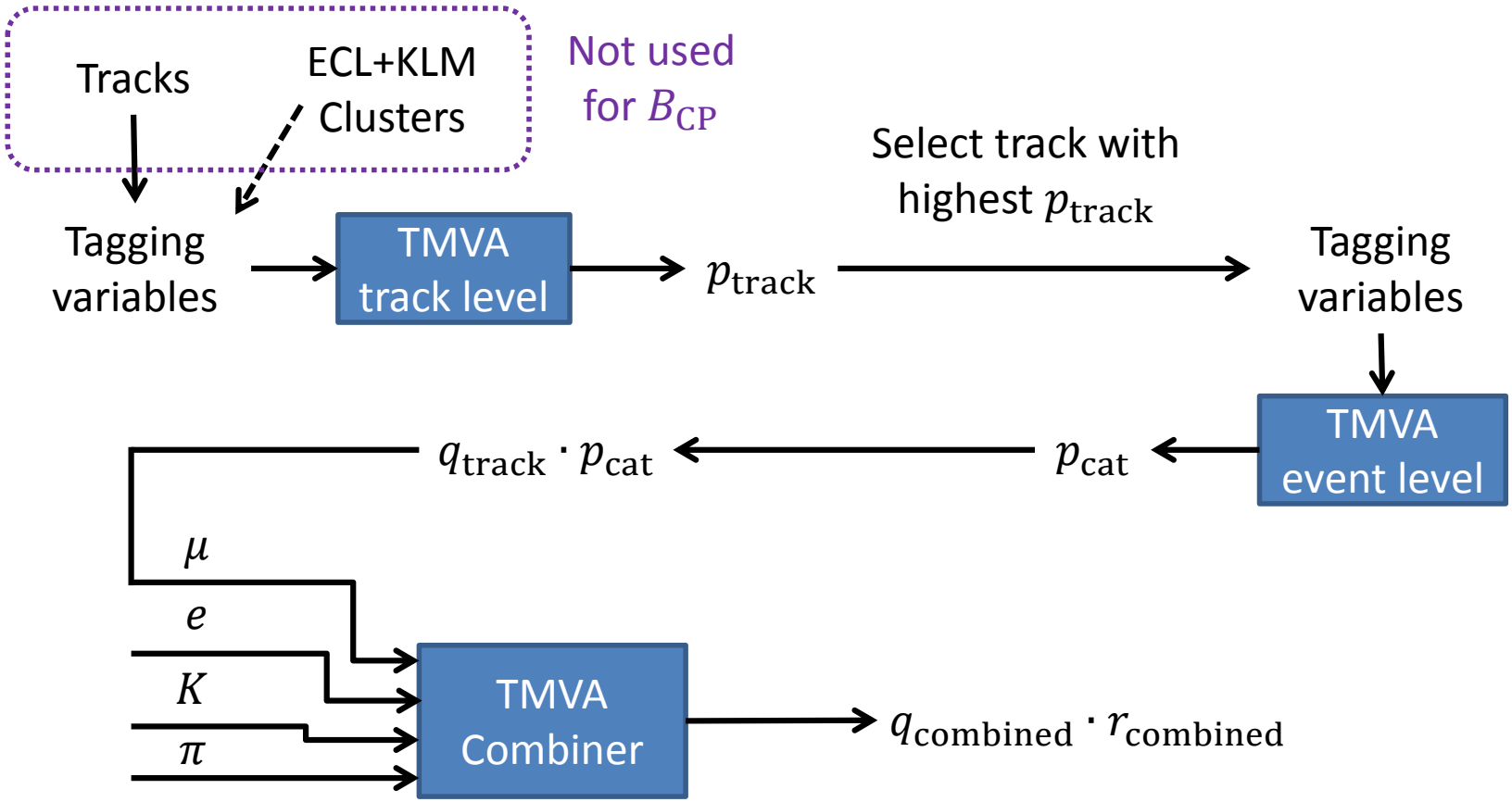
$\Rightarrow$ Efficiency for several events: $\varepsilon_{\mathsf{Eff}} = r^2\varepsilon$

Actual Categories and Variables

Categories	Targets
Electron	e^-
(Intermediate Electron)	e^+
Muon	μ^-
(Intermediate Muon)	μ^+
KinLepton	e^-
Kaon	K^-
KaonPion	K^-, π^+
SlowPion	π^+
FastPion	π^-
MaximumP	ℓ^-, π^-
FSC	ℓ^-, π^+
Lambda	Λ
Total= 10 (12)	

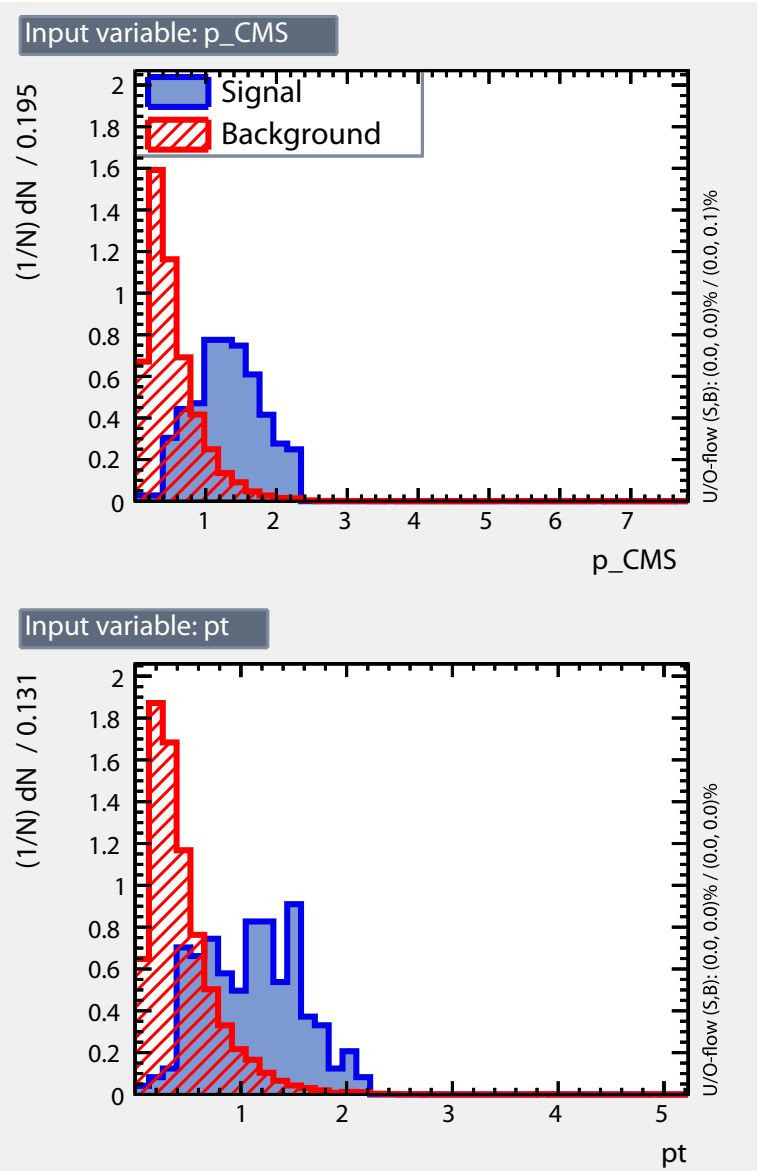


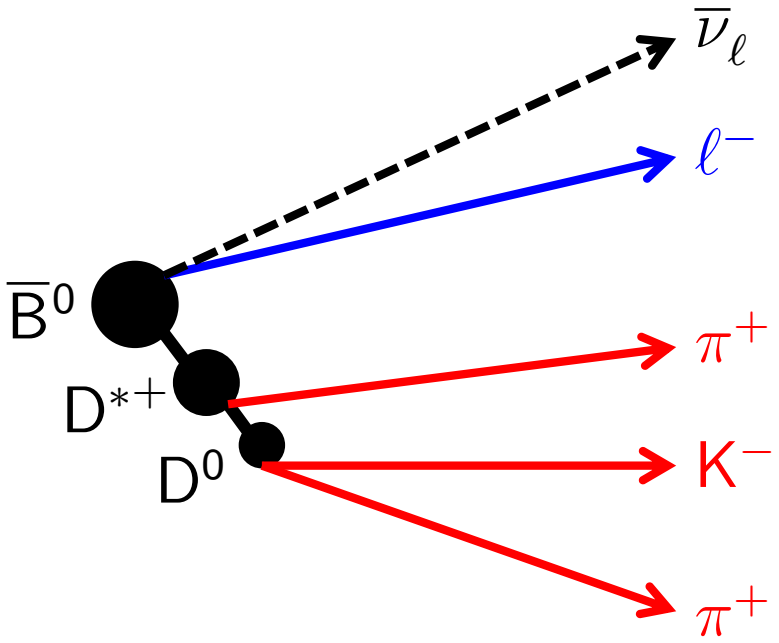
Flavor tagging



- TMVA Multivariate Method ‘FastBDT’ $p \mathrel{\hat{=}}$ probability

Tagging Variables





- E.g. $\cos(\theta)$, d_0, z_0, E_W^{90}

Combiner

Combines all event level outputs:

$\Rightarrow$ 10 Inputs: $q_{\text{track}} \cdot p_{\text{Cat}} \in [-1, 1]$ of each Category

$\Rightarrow$ Output: $y_{\text{Out}} = (q \cdot r)_{\text{Output}} = (q \cdot r)_{\text{Combined}} \in [-1, 1]$

Methods Explored:

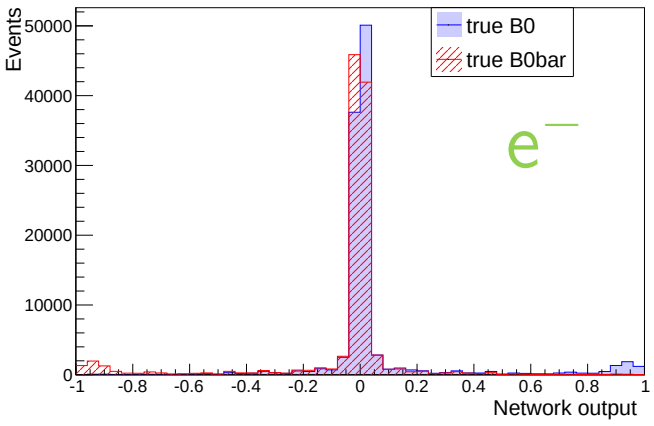
- TMVA '`FastBDT`'
- FANN Multiple Layer Perceptron (MLP)

Training sample (500 MC kEvents):

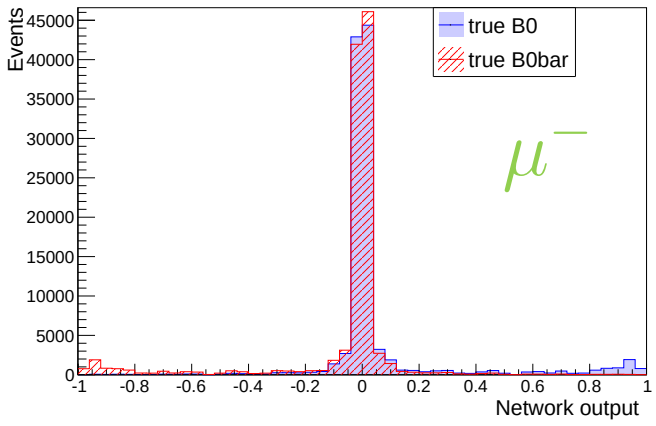
- $\Upsilon(4S) \to \mathrm{B^0}_1 \mathrm{B^0}_2 \to$
  $B_1 \to J/\Psi K^0_S$ $\quad$ $B_2 \to$ generic

Combiner Input

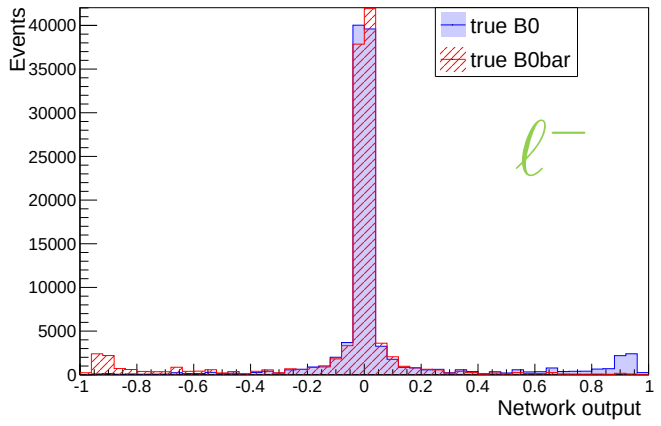
Electron category input



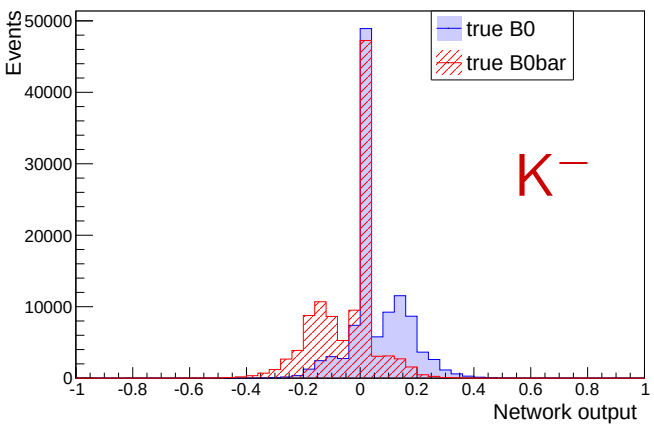
Muon category input



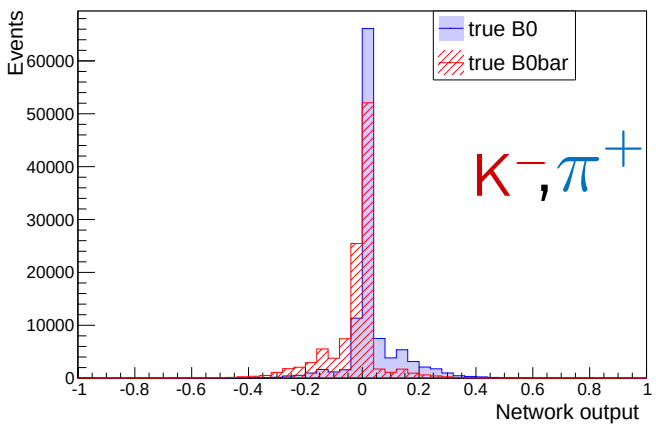
KinLepton category input



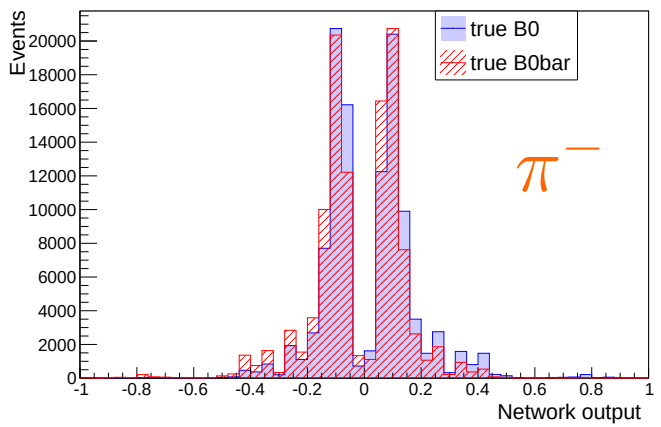
Kaon category input



KaonPion category input

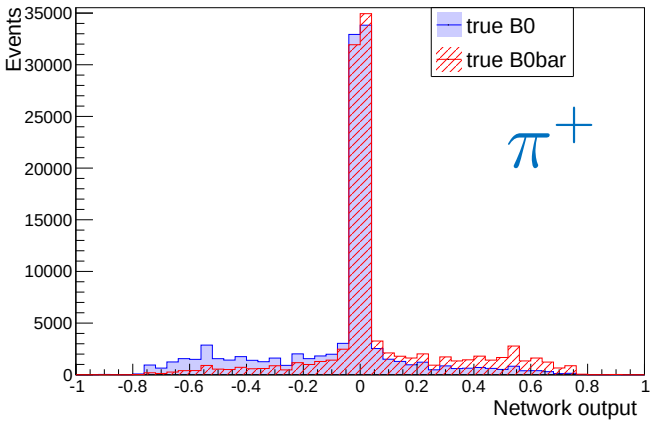


FastPion category input

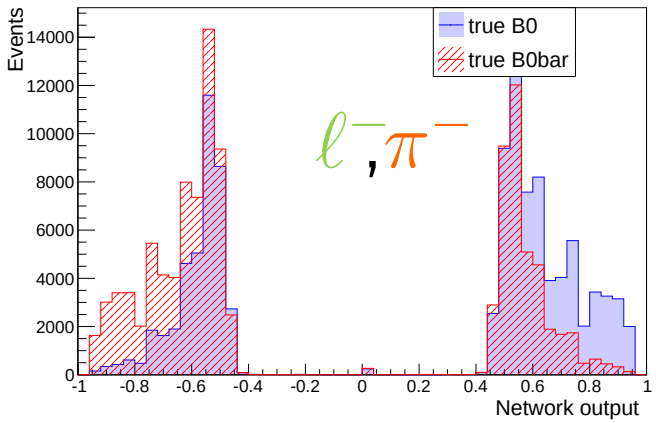


Combiner Input

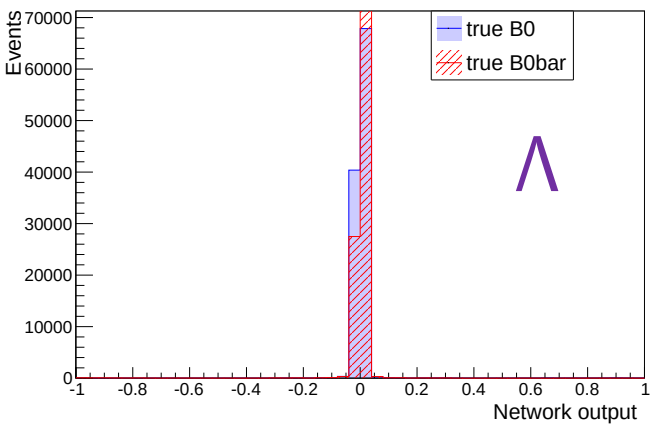
SlowPion category input



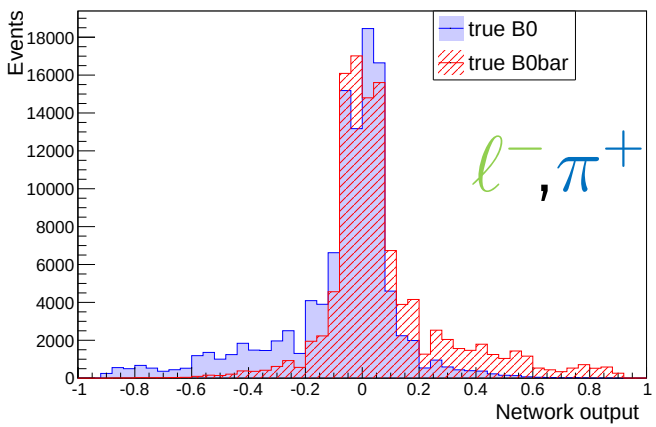
MaximumP category input



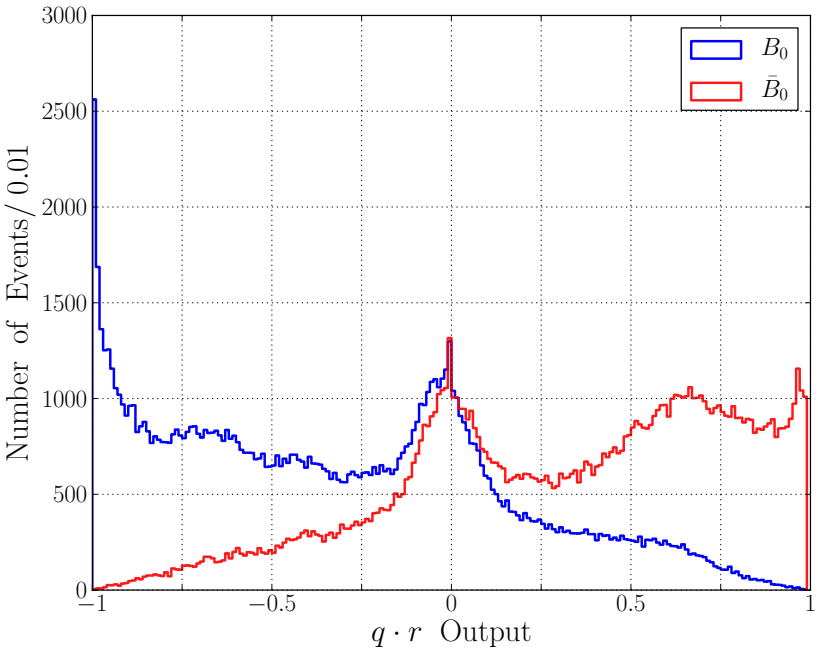
Lambda category input

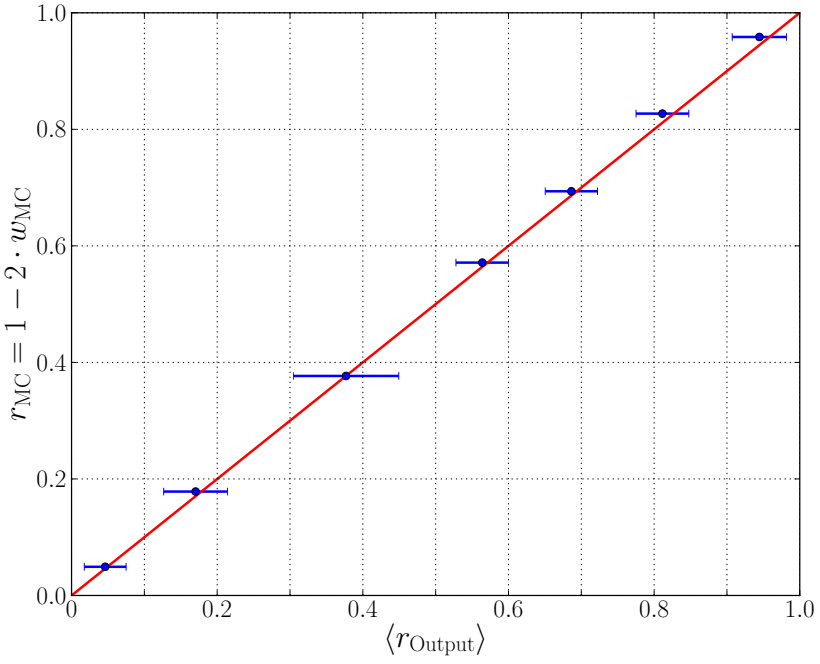


FSC category input



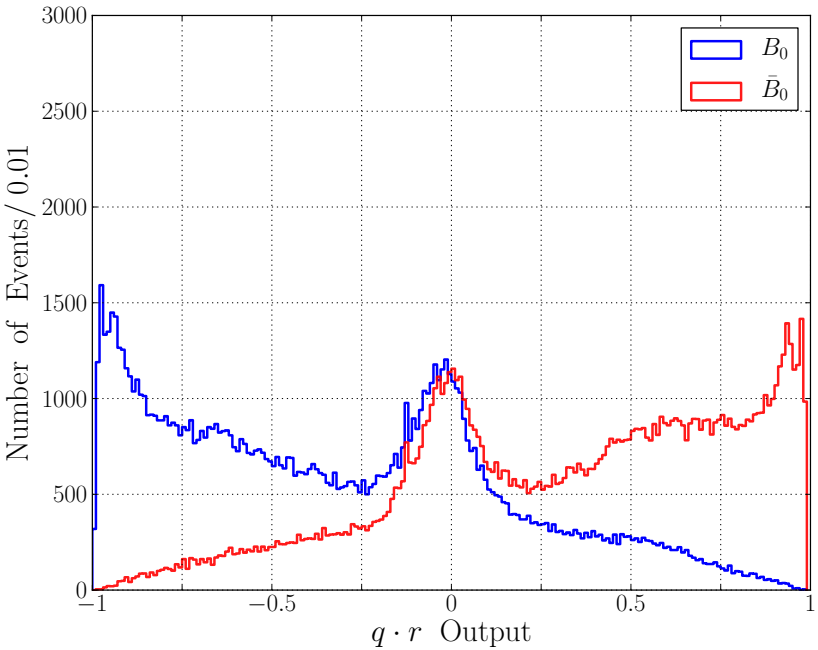
TMVA FastBDT Output

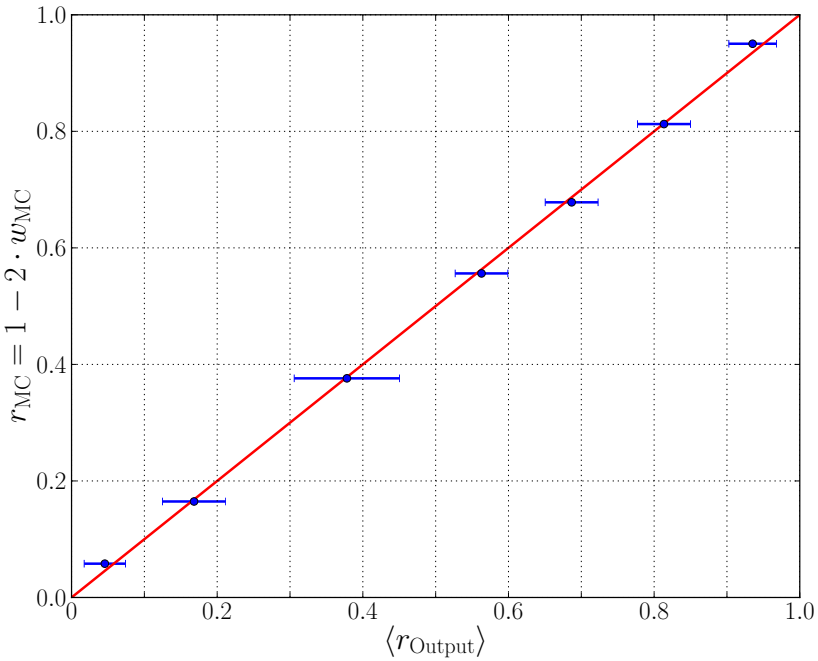




Total Efficiency = 31.3 %

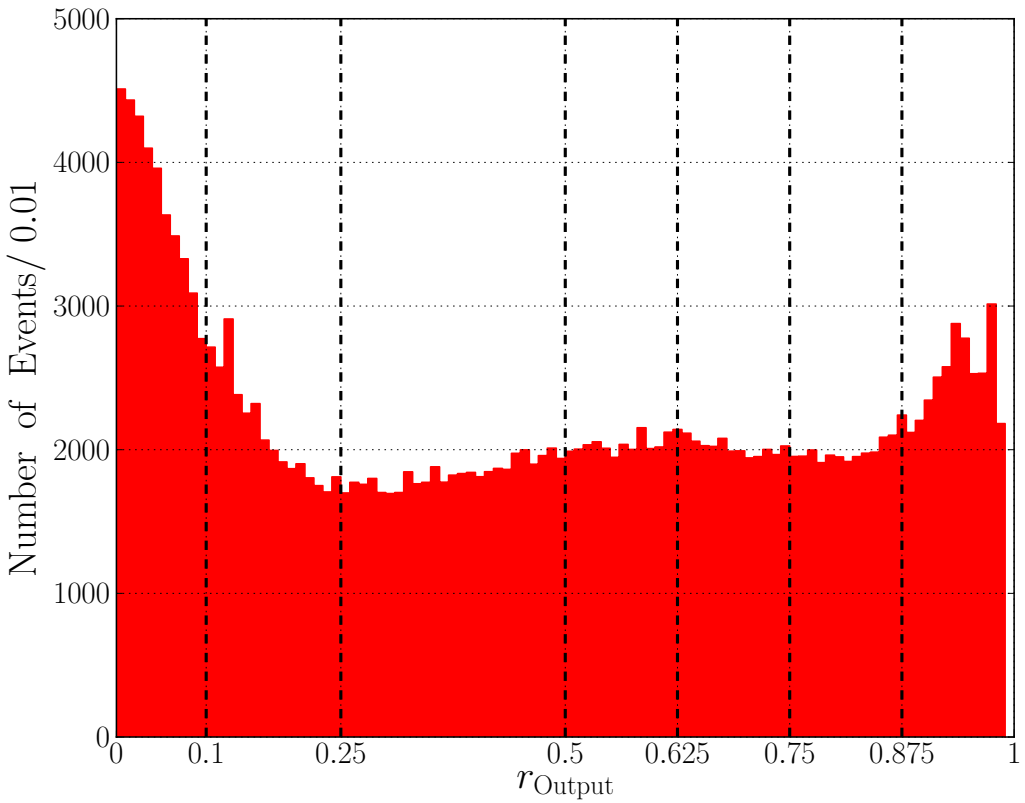
FANN MLP Output





Total Efficiency = 31.6 %

Efficiency Calculation



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

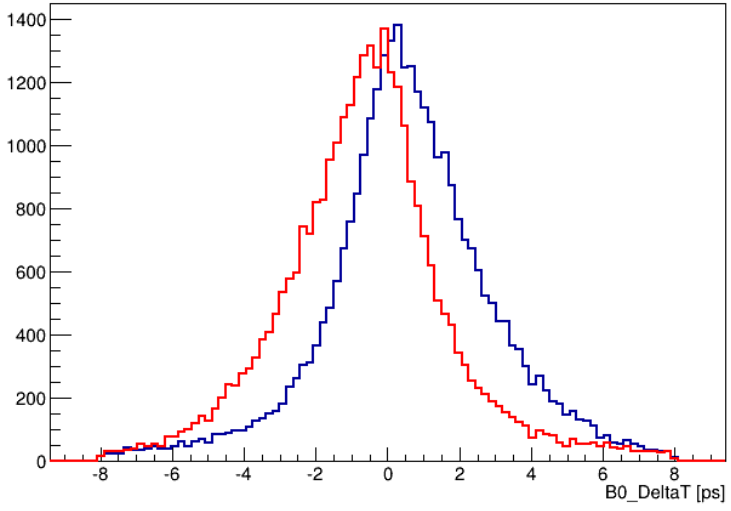
$$\varepsilon_{\text{Eff}} = \sum_i \varepsilon_i \cdot \langle r_i \rangle^2$$

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

Δt Distribution

MC Analysis of $B_{CP} \to J/\psi K^0_S$

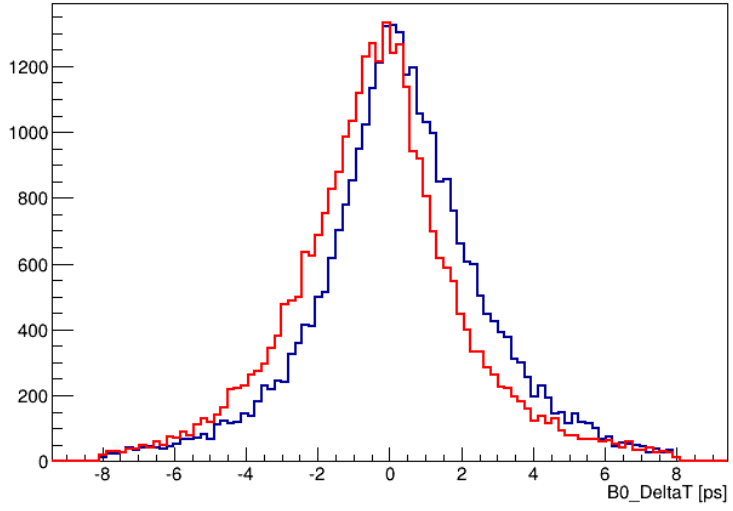
a) q from MC



B^0

$\overline{B}^0$

b) q tagged (Total Effcy. $= 34\ \%$)



Impact on $\mathcal{P}(\Delta t, q, r)$:

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

Outlook

- A good Flavor Tagging is indispensable for the Study of time dependent CP Violation.
- Up to now: Flavor Tagging Interface for Belle II reaches a slightly better efficiency ($\sim 34\%$) than its predecessors. (Belle $\sim 29\%$, BaBar $\sim 33\%$)
- Still several studies to be made: Comparison and optimization of methods, categories, variables etc.
- Challenge: the development and the accuracy of the Flavor tagging is coupled to the whole software development.

Efficiency of Actual Categories

Category	Actual $\epsilon[\%]$	BaBar $\epsilon[\%]$
Electron	5.33	6.09
Muon	5.40	4.27
KinLepton	5.94	2.83
Kaon	19.06	19.74
KaonPion	13.46	10.23
SlowPion	6.11	5.70
FastPion	2.39	–
MaximumP	12.55	11.19
FSC	11.49	6.26
Lambda	1.781	0.29

Background: Train a Multi-Layer Perceptron

a) Inputs $(q \cdot p)_k \hat{=} (q \cdot p)_{\text{Cat}}$

b) Target $\hat{=}$ q_{MC}

- One Hidden Layer (Kolmogorov)

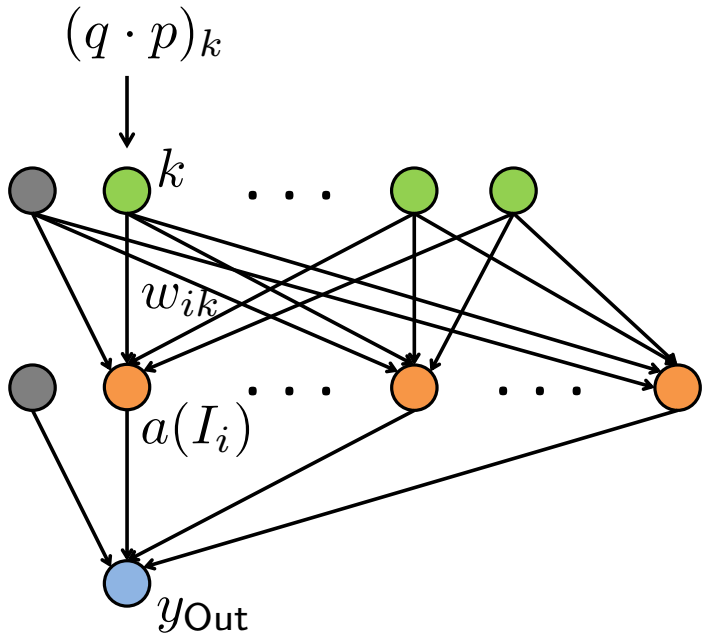
$\Rightarrow$ Hidden layer: $n_{\text{hidden}} = 3 \cdot n_{\text{input}}$

$\Rightarrow$ Activation function
$a_i = \tanh(I_i) \in [-1, 1]$

$\Rightarrow$ Input: $I_i = \sum_{k=0}^{n_{\text{input}}} w_{ik}(q \cdot p)_k$

- Training Sample$= \sim 110$ kEvents

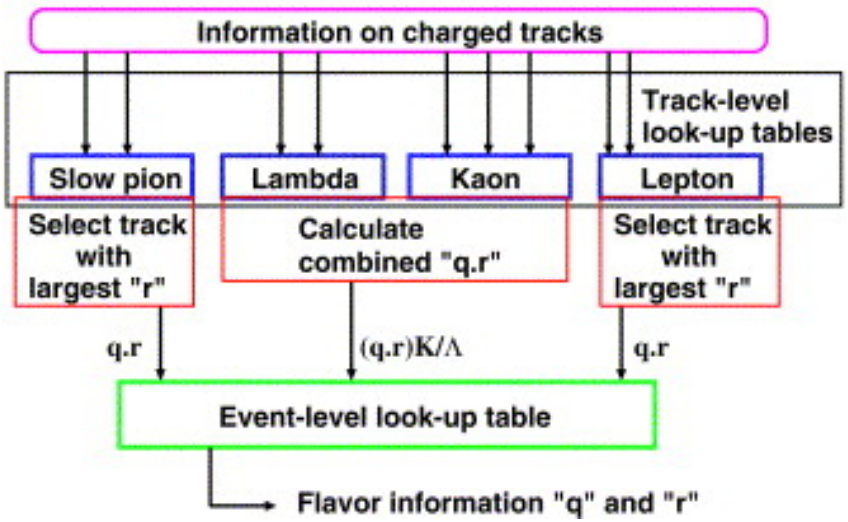
$\Rightarrow$ Output neuron: $y_{\text{Out}} = (q \cdot r)_{\text{Combined}}$

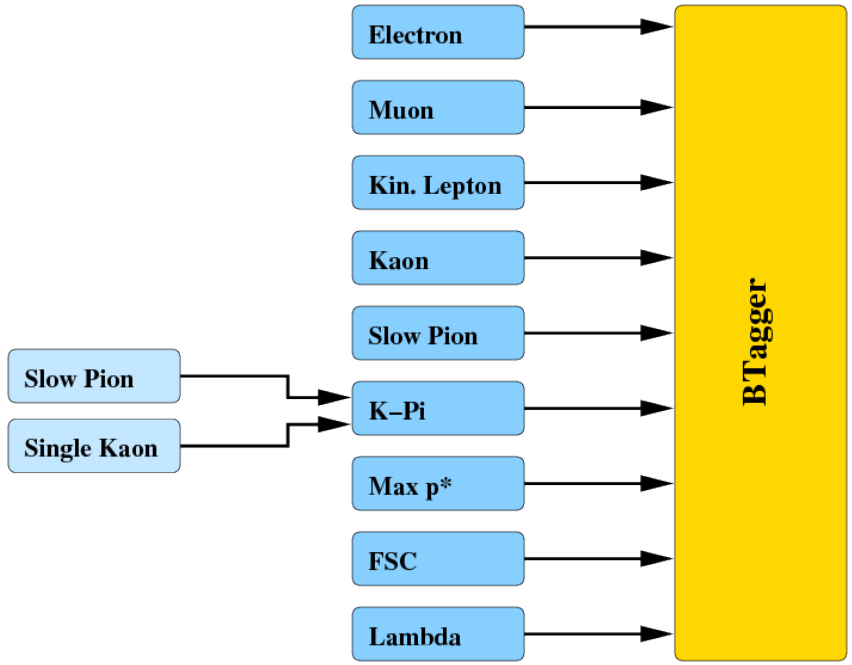


Training:

- Cost function$=$ MSE
- Training Algorithm$=$ BP
- NCycles$=\sim$2000
- Monitoring$=$ True

Belle vs. Babar Scheme





⇒ Each step: Trained neural network (NN) or other TMVA method.