Karlsruhe Institute of Technology

An improved Full Reconstruction tool utilizing NeuroBayes

ACAT 2011

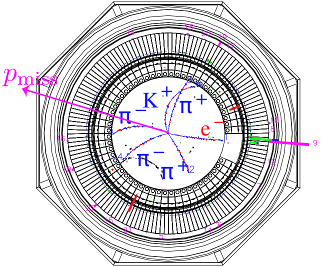
M. Feindt, T. Kuhr, S. Neubauer, D. Zander, A. Zupanc | September 6th, 2011

INSTITUT FÜR EXPERIMENTELLE KERNPHYSIK, KARLSRUHE INSTITUTE OF TECHNOLOGY (KIT)

Contents

The Full Reconstruction method

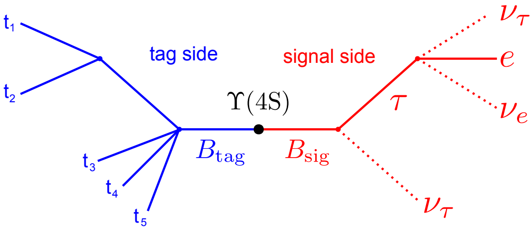
- **B factories:** e^+e^- colliders with center of mass energy of $\sqrt{s} = 10.58$ GeV at the $\Upsilon(4S)$ resonance
- $\mathcal{B}(\Upsilon(4S) \to B\bar{B}) > 96\%$
  - $\Rightarrow\ p_{CMS}(B) + p_{CMS}(\bar{B}) + p_{CMS}(\mathrm{Beam}) = 0$
  - $\Rightarrow$ Clean event topologies (typically ~ 10 tracks)



$B^+_{tag} \to \bar{D}^0\pi^+\pi^+\pi^-$
$\hookrightarrow K^+\pi^-$

$B^-_{sig} \to \tau^-\bar{\nu}_\tau$
$\hookrightarrow e^-\bar{\nu}_e\nu_\tau$

Typical B factory event



Tag and signal side of the full reconstruction

Typical use-cases

The full reconstruction allows to measure missing momentum and additional energy.

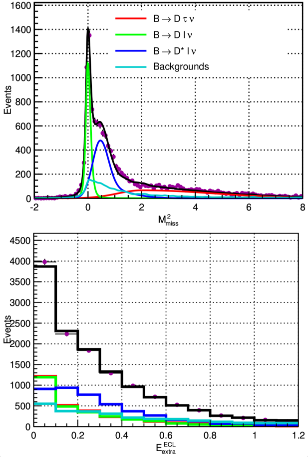
(Semi-)leptonic decays

- $B \to \tau\nu$
- $B \to \ell\nu\gamma$
- $B \to D^{(*)}\ell\nu$
- $B \to D^{(*)}\tau\nu$
- $B \to h\nu\nu$
- $B \to \nu\nu$

Inclusive searches

- $B \to K + X_{c\bar{c}}$

Monte Carlo Simulation:

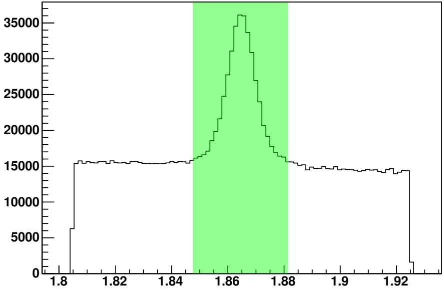


Exemplary reconstruction of the tag side

$$B^+ \to D^0\pi^+\pi^+\pi^-$$

First step

- Reconstruction of D^0 meson (exemplary $\bar{D}^0 \to K^+\pi^-$)
- Specific cuts on Particle ID
- Cut on D^0 mass



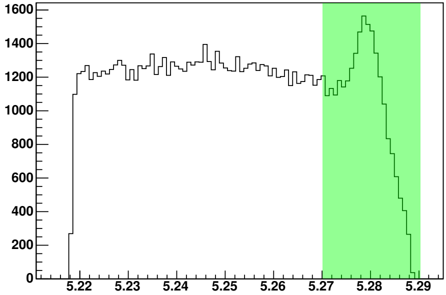
Mass of the D^0 meson

B reconstruction: Important variables

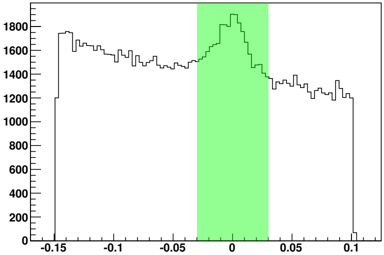
$$M_{bc} \equiv \sqrt{E_{beam}^2 - p_B^2}$$

$$\Delta E \equiv E_{B\text{-meson}} - E_{\text{Beam}}$$

Exemplary reconstruction of the tag side



M_{bc} of the B^+ meson

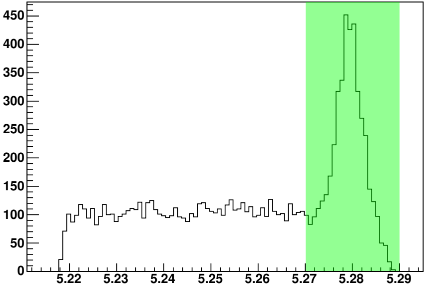


ΔE of the B^+ meson

Often necessary: Improvement of significance

- Usage of multivariate techniques

Exemplary reconstruction of the tag side



M_{bc} of the B^{+} meson

Typical variables:

- Vertex fit information
- Kinematic Variables
- Particle ID information
- M_{bc} or ΔE

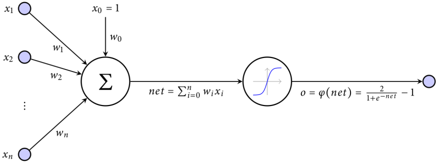
Full Reconstruction

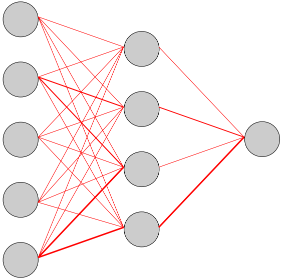
This procedure is performed for hundreds of different channels

Contents

NeuroBayes

- Multivariate analysis software combining a **Neural** Network with sophisticated pre-processing

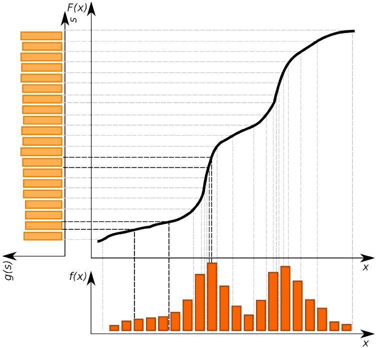




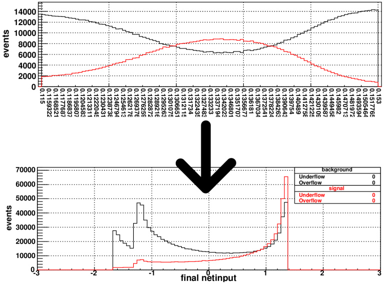
- The output of the Network can be interpreted as **Bayesian** probability

NeuroBayes - Pre-processing

- Input variables are flattened

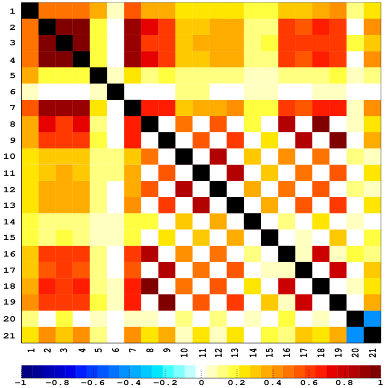


- Purity is taken and transformed to have mean 0 and width 1.



NeuroBayes - Pre-processing

- Input variables are decorrelated



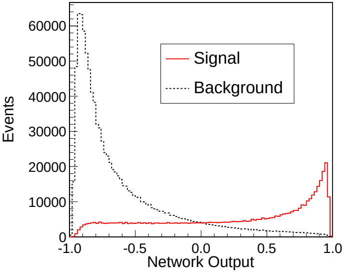
Pre-processing . . .

- Speeds up the training process
- Facilitates the weight finding
- Increases the robustness of the algorithm

Probability

1) Signal to background ratio in training is the same as on data

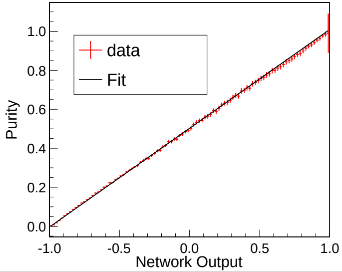
(Output of NeuroBayes+1)/2 is the signal probability by construction.



2) Signal to background ratio in training is **not** the same as on data

- It is often necessary to artificially enhance the signal component for the training.
- The output can be corrected:

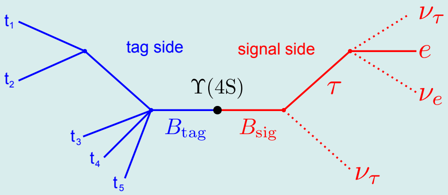
$$o_p = \frac{1}{1 + \left(\frac{1}{o_t} - 1\right)\frac{P_p(B)}{P_p(S)}\frac{P_t(S)}{P_t(B)}}$$



Contents

Efficiency and Purity

Two premises



1. The more **correct** tag side *B* mesons, the more signal side *B* mesons are available for analysis.
   $\Rightarrow$ **Need for good efficiency**
2. The more **incorrect** tag side *B* mesons, the more background pollutes the signal side.
   $\Rightarrow$ **Need for good purity**

Channels

Channel			$\mathcal{BR}$	Channel			$\mathcal{BR}$
B^+	$\rightarrow$	$\bar{D}^0\pi^+$	0.484%	B^0	$\rightarrow$	$D^-\pi^+$	0.268%
B^+	$\rightarrow$	$\bar{D}^0\pi^+\pi^0$	1.340%	B^0	$\rightarrow$	$D^-\pi^+\pi^0$	0.760%
B^+	$\rightarrow$	$\bar{D}^0\pi^+\pi^+\pi^-$	1.100%	B^0	$\rightarrow$	$D^-\pi^+\pi^+\pi^-$	0.800%
B^+	$\rightarrow$	$D_S^+\bar{D}^0$	1.000%	B^0	$\rightarrow$	$\bar{D}^0\pi^0$	0.026%
B^+	$\rightarrow$	$\bar{D}^{0*}\pi^+$	0.519%	B^0	$\rightarrow$	$D_S^+D^-$	0.720%
B^+	$\rightarrow$	$\bar{D}^{0*}\pi^+\pi^0$	0.980%	B^0	$\rightarrow$	$D^{*-}\pi^+$	0.276%
B^+	$\rightarrow$	$\bar{D}^{0*}\pi^+\pi^+\pi^-$	1.030%	B^0	$\rightarrow$	$D^{*-}\pi^+\pi^0$	1.500%
B^+	$\rightarrow$	$\bar{D}^{0*}\pi^+\pi^+\pi^-\pi^0$	1.800%	B^0	$\rightarrow$	$D^{*-}\pi^+\pi^+\pi^-$	0.700%
B^+	$\rightarrow$	$D_S^{+*}\bar{D}^0$	0.760%	B^0	$\rightarrow$	$D^{*-}\pi^+\pi^+\pi^-\pi^0$	1.760%
B^+	$\rightarrow$	$D_S^+\bar{D}^{0*}$	0.820%	B^0	$\rightarrow$	$D_S^{+*}D^-$	0.740%
B^+	$\rightarrow$	$D_S^{+*}\bar{D}^{0*}$	1.710%	B^0	$\rightarrow$	$D_S^+D^{*-}$	0.800%
B^+	$\rightarrow$	$\bar{D}^0K^+$	0.037%	B^0	$\rightarrow$	$D_S^{+*}D^{*-}$	1.770%
B^+	$\rightarrow$	$D^-\pi^+\pi^+$	0.107%	B^0	$\rightarrow$	$J/\psi K_S^0$	0.087%
B^+	$\rightarrow$	$J/\psi K^+$	0.101%	B^0	$\rightarrow$	$J/\psi K^+\pi^-$	0.120%
B^+	$\rightarrow$	$J/\psi K^+\pi^+\pi^-$	0.107%	B^0	$\rightarrow$	$J/\psi K_S^0\pi^+\pi^-$	0.100%
B^+	$\rightarrow$	$J/\psi K^+\pi^0$	0.047%				
B^+	$\rightarrow$	$J/\psi K_S^0\pi^+$	0.094%				

Channels

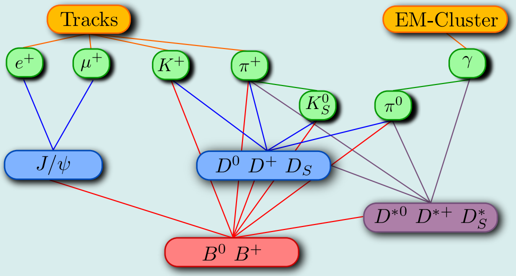
Channel			$\mathcal{BR}$	Channel			$\mathcal{BR}$
D^0	$\rightarrow$	$K^-\pi^+$	3.89%	D^+	$\rightarrow$	$K^-\pi^+\pi^+$	9.40%
D^0	$\rightarrow$	$K^-\pi^+\pi^+\pi^-$	8.09%	D^+	$\rightarrow$	$K_S^0\pi^+$	1.49%
D^0	$\rightarrow$	$K^-\pi^+\pi^0$	6.90%	D^+	$\rightarrow$	$K_S^0\pi^+\pi^0$	6.90%
D^0	$\rightarrow$	$\pi^+\pi^-$	0.14%	D^+	$\rightarrow$	$K^-\pi^+\pi^+\pi^0$	6.08%
D^0	$\rightarrow$	$\pi^+\pi^-\pi^0$	1.44%	D^+	$\rightarrow$	$K_S^0\pi^+\pi^+\pi^-$	3.10%
D^0	$\rightarrow$	$K_S^0\pi^0$	1.22%	D^+	$\rightarrow$	$K^+K^-\pi^+$	0.98%
D^0	$\rightarrow$	$K_S^0\pi^+\pi^-$	2.94%	D^+	$\rightarrow$	$K^+K^-\pi^+\pi^0$	1.50%
D^0	$\rightarrow$	$K_S^0\pi^+\pi^-\pi^0$	5.40%	D^{+*}	$\rightarrow$	$D^0\pi^+$	67.70%
D^0	$\rightarrow$	K^+K^-	0.39%	D^{+*}	$\rightarrow$	$D^+\pi^0$	30.70%
D^0	$\rightarrow$	$K^+K^-K_S^0$	0.47%				
D^{0*}	$\rightarrow$	$D^0\pi^0$	61.9%	D^{0*}	$\rightarrow$	$D^0\gamma$	38.10%
D_S^+	$\rightarrow$	$K^+K_S^0$	1.49%	D_S^+	$\rightarrow$	$K^+K^-\pi^+\pi^+\pi^-$	0.88%
D_S^+	$\rightarrow$	$K^+\pi^+\pi^-$	0.69%	D_S^+	$\rightarrow$	$\pi^+\pi^+\pi^-$	1.10%
D_S^+	$\rightarrow$	$K^+K^-\pi^+$	5.50%	D_S^{+*}	$\rightarrow$	$D_S^+\gamma$	94.20%
D_S^+	$\rightarrow$	$K^+K^-\pi^+\pi^0$	5.60%	J/ψ	$\rightarrow$	e^-e^+	5.94%
D_S^+	$\rightarrow$	$K^+K_S^0\pi^+\pi^-$	0.96%	J/ψ	$\rightarrow$	$\mu^-\mu^+$	5.93%
D_S^+	$\rightarrow$	$K^-K_S^0\pi^+\pi^+$	1.64%				

Channels

Channels		$\mathcal{BR}$	Channels		$\mathcal{BR}$	Channels		$\mathcal{BR}$
D^+	7	29.4%	D^{*+}	2	98.4%	B^+	17	12.0%
D^0	10	37.9%	D^{*0}	2	100.0%	B^0	15	10.4%
D_S^+	8	17.9%	D_S^{*+}	1	94.2%			
J/Ψ	2	11.9%						

Exclusive reconstruction of **1104 decay channels** and code maintenance would be a futile task.

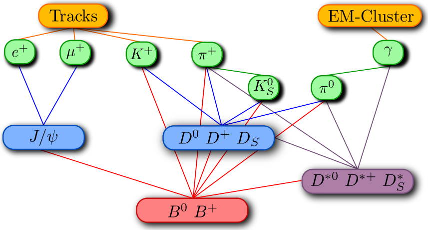
Hierarchical Reconstruction



Intermediate Cuts

How to make the smartest cuts possible?

- How do you compare $D^0 \to K^- \pi^+$ to $D^0 \to K^0_S \pi^+ \pi^- \pi^0$?
- Even worse: The cut depends on the next level: The D^0 meson in $B^+ \to \bar{D}^0 \pi^+ \pi^+ \pi^-$ should get a different cut than in $B^+ \to \bar{D}^0 \pi^+$.
- **Solution:** Multiply the signal probability (given by the NeuroBayes training) of all children and use that to cut.



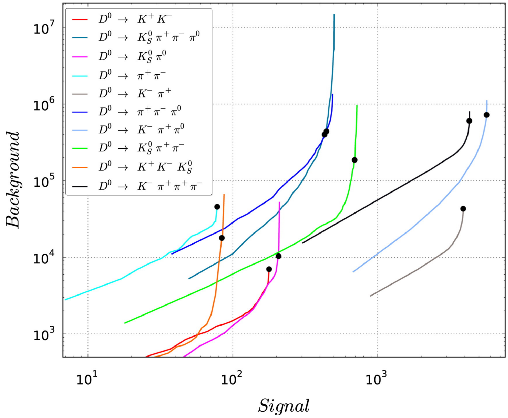
Cut decisions are postponed to a later level

Not only the reconstruction is **hierarchical**, but also the **information flow**.

Channels

Pre-cuts

- Cut on the product of the **signal probabilities (= NeuroBayes outputs)** of the children.



- Choose a cut to have roughly the same slope for all curves.
- This slope corresponds to the number of candidates.
- Very soft cuts, usage of probability product on next level.
- Trade-off: Efficiency ↔ Purity and CPU time.

Contents

Usage

- Run the Full Reconstruction on your data set
- The “fullrecon table” is added to certain events of your data set (Belle specific)
- Choose your candidate

```
-----------------------------------------------------------------------------------------------
 ID TAG_ID DECAYDDEC(1)DDEC(2)DDEC(3)DDEC(4) NBRANK NBOUTCONT_NBCONT_NB MCINFO DELTAE MBC NFS BREC
-----------------------------------------------------------------------------------------------
 1 521 521743 421421 0 0 0 2 0.09 0 0.00 -2 -0.09 5.29 7 1
 2 -521-521743 421421 0 0 0 3 0.01 0 0.00 -2 0.06 5.26 7 13
 3 -521-521743 421421 0 0 0 1 0.25 0 0.00 1 -0.09 5.28 7 25
-----------------------------------------------------------------------------------------------
```

Results

What is the result of the Full Reconstruction?

- A collection of B_{tag} candidates
- Visualisation of results: *beam constrained mass* M_{bc}

$$M_{bc} \equiv \sqrt{E_{beam}^2 - p_B^2}$$

Any basis for comparison?

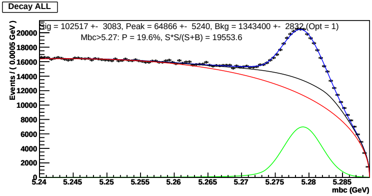
- Yes! The cut-based predecessor to this Full Reconstruction.

Adjusting Cut on NeuroBayes output

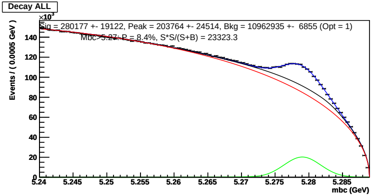
Cut-based predecessor

NeuroBayes Full Reconstruction

B^+

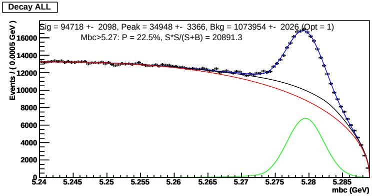


sig: 103,000 **bg: 1,300,000** pur: 19.6 %

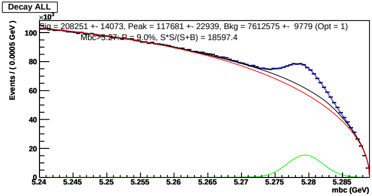


sig: 280,000 **bg: 11,000,000** pur: 8.4 %

B^0



sig: 95,000 **bg: 1,100,000** pur: 22.5 %



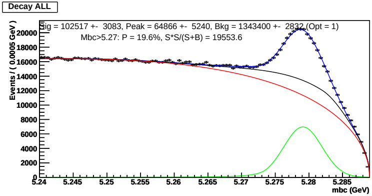
sig: 209,000 **bg: 7,600,000** pur: 9.0 %

Adjusting Cut on NeuroBayes output

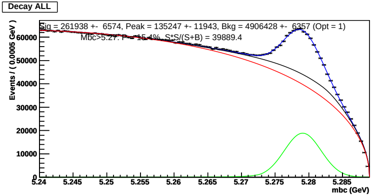
Cut-based predecessor

NeuroBayes Full Reconstruction

B^+

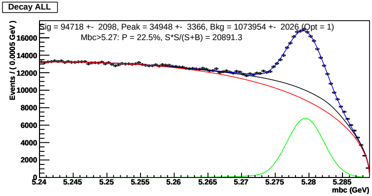


sig: 103,000 **bg: 1,300,000** pur: 19.6 %

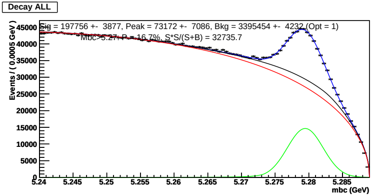


sig:262,000 **bg: 4,900,000** pur: 15.4 %

B^0



sig: 95,000 **bg: 1,100,000** pur: 22.5 %

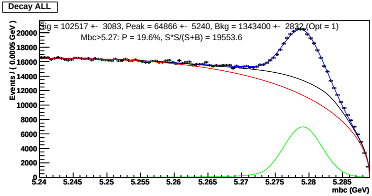


sig:198,000 **bg: 3,400,000** pur: 16.7 %

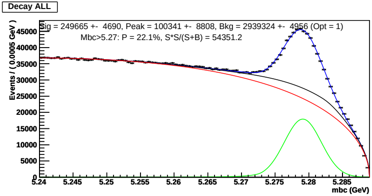
Adjusting Cut on NeuroBayes output

Cut-based predecessor | NeuroBayes Full Reconstruction

B^+

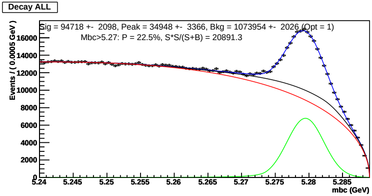


sig: 103,000 **bg: 1,300,000** pur: 19.6 %

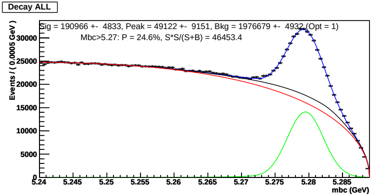


sig: 250,000 **bg: 2,900,000** pur: 22.1 %

B^0



sig: 95,000 **bg: 1,100,000** pur: 22.5 %



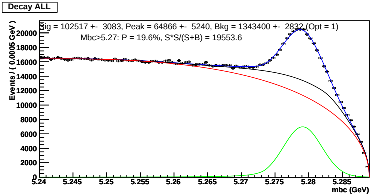
sig: 191,000 **bg: 2,000,000** pur: 24.6 %

Adjusting Cut on NeuroBayes output

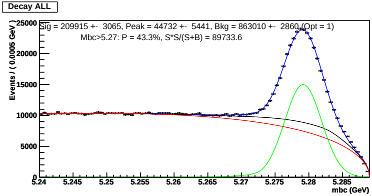
Cut-based predecessor

NeuroBayes Full Reconstruction

B^+

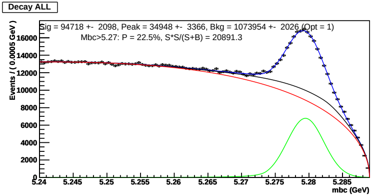


sig: 103,000 **bg: 1,300,000** pur: 19.6 %

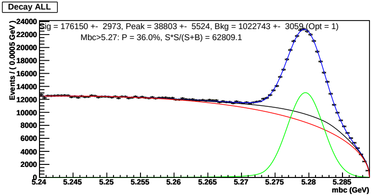


sig: 210,000 **bg: 860,000** pur: 43.3 %

B^0



sig: 95,000 **bg: 1,100,000** pur: 22.5 %

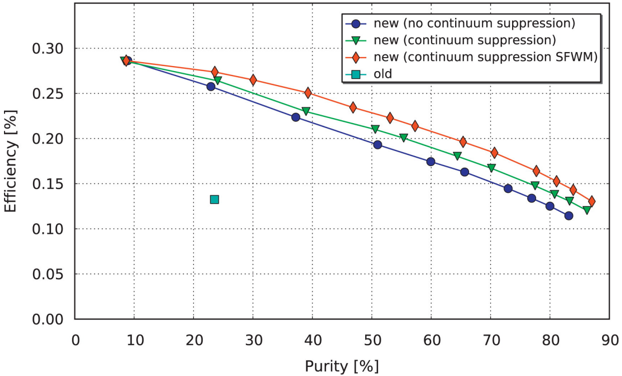


sig: 176,000 **bg: 1,000,000** pur: 36.0 %

Purity-Efficiency of B_{tag} Sample

Purity-Efficiency

- Classification of B^+ mesons, cut on the signal probability

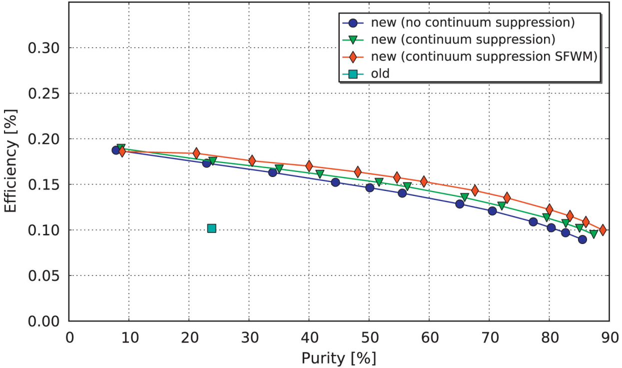


With high efficiency: 2.1 Million B^+_{tag} mesons

Purity-Efficiency of B_{tag} Sample

Purity-Efficiency

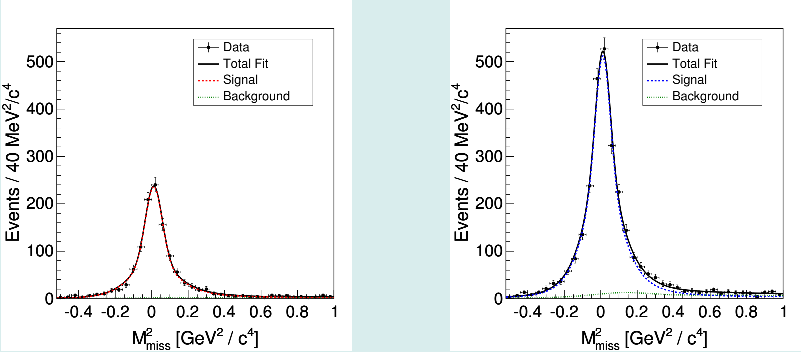
- Classification of B^0 mesons, cut on the signal probability



With high efficiency: 1.4 Million B^0_{tag} mesons

Signal Side

Exemplary signal side reconstruction: $B^0 \rightarrow D^{*-} l^+ \nu_l$



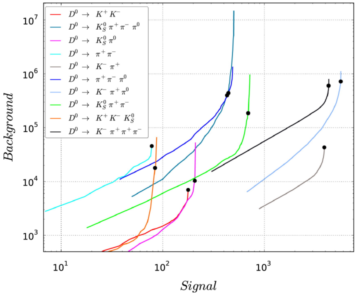
Comparison between the cut based **(left)** and the NeuroBayes **(right)** Full Reconstruction.

Reduced Channels

- The question has popped up:

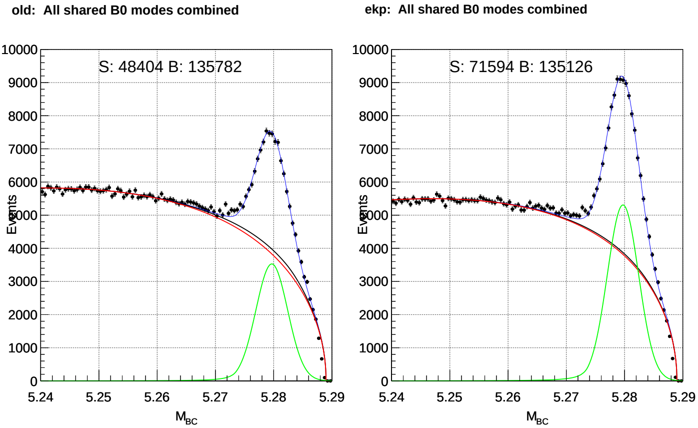
  “How much of the improvement is due to new channels and how much due to NeuroBayes?”

- Hard to answer, as the two aspects are connected.



Approximation: Only use channels that are common to both Full Reconstructions.

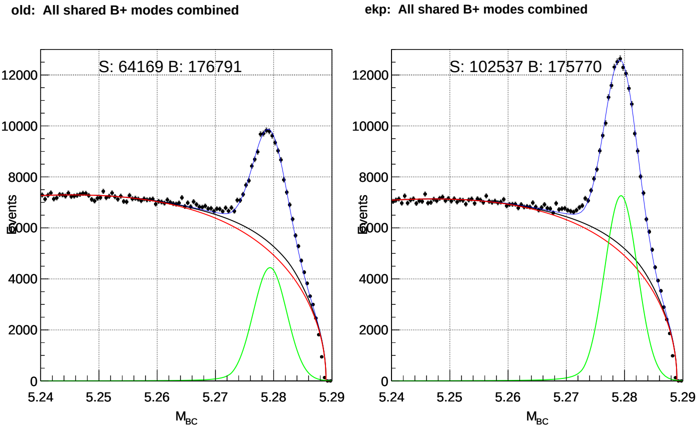
Reduced Channels B^0



Still a factor ~ 1.5 improvement

Tag side for only common channels for the cut based **(left)** and the NeuroBayes **(right)** Full Reconstruction.

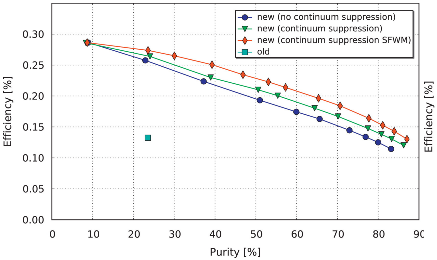
Reduced Channels B^+

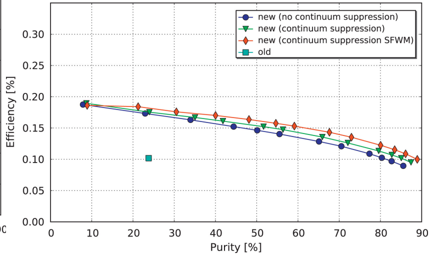


Still a factor ~ 1.6 improvement

Tag side for only common channels for the cut based **(left)** and the NeuroBayes **(right)** Full Reconstruction.

Summary and Conclusion





- The Full Reconstruction allows to **measure missing momentum** and is therefore **crucial for numerous analyses**.
- NeuroBayes and a hierarchical information flow $\Rightarrow$ **improvement of factor $\sim$ 2**.
- Corresponds to 10 years of data taking.

Publication

ARTICLE IN PRESS

Nuclear Instruments and Methods in Physics Research A ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

A hierarchical NeuroBayes-based algorithm for full reconstruction of B mesons at B factories

M. Feindt, F. Keller, M. Kreps [1], T. Kuhr, S. Neubauer *, D. Zander, A. Zupanc

Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, Campus Süd, Postfach 69 80, 76128 Karlsruhe, Germany

http://www.sciencedirect.com/science/article/pii/S0168900211011193

available online 16 June 2011, ISSN 0168-9002, DOI: 10.1016/j.nima.2011.06.008.