ProbNN vs. DLL

A PID Comparison

Alex Pearce (University of Manchester)

PPTS Meeting - 8th July 2013

Contents

PID at LHCb

- Most PID power from 2 RICHs
  - Leptons also include CALO, MUON
- PID crucial for heavy flavour physics (RS/WS, rare decays, spectroscopy...)
- Knowledge of PID systems and variables important for almost all analyses

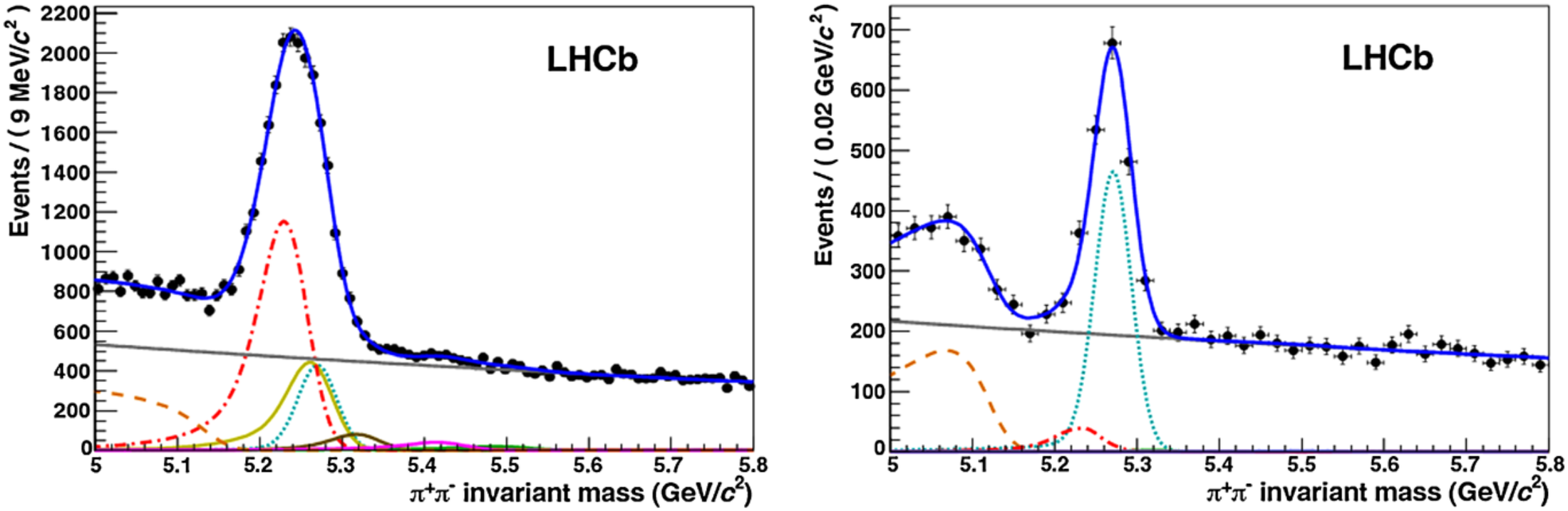


Fig. 2 Invariant mass distribution for $B \to h^+h^-$ decays [6] in the LHCb data before the use of the RICH information (*left*), and after applying RICH particle identification (*right*). The signal under study is the decay $B^0 \to \pi^+\pi^-$, represented by the turquoise *dotted line*. The contributions from different *b*-hadron decay modes ($B^0 \to K\pi$ *red dashed-dotted line*, $B^0 \to$ 3-body *orange dashed-dashed line*, $B_s \to KK$ *yellow line*, $B_s \to K\pi$ *brown line*, $\Lambda_b \to pK$ *purple line*, $\Lambda_b \to p\pi$ *green line*), are eliminated by positive identification of pions, kaons and protons and only the signal and two background contributions remain visible in the plot on the right. The *grey solid line* is the combinatorial background (Color figure online)

DLL

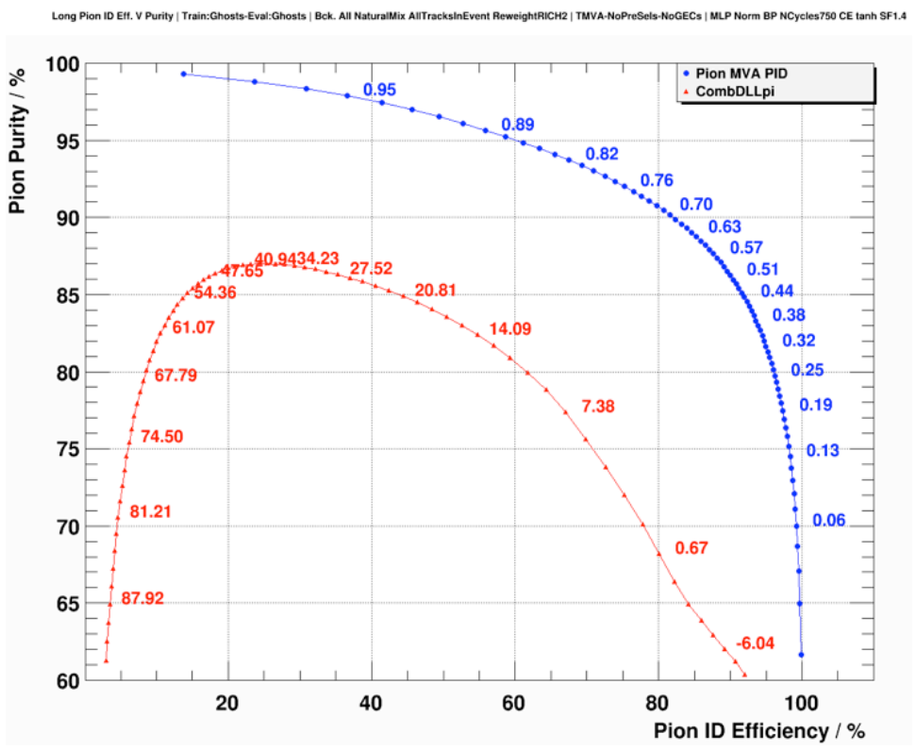
- Tracking propagates one cone of photons per particle ID hypothesis, compare predicted pixel hits with those seen by HPDs
- Tracks given a "log likelihood" for each hypothesis, based on how well hits match
- Likelihood is relative to pion hypothesis, *a priori* zero, hence "delta"

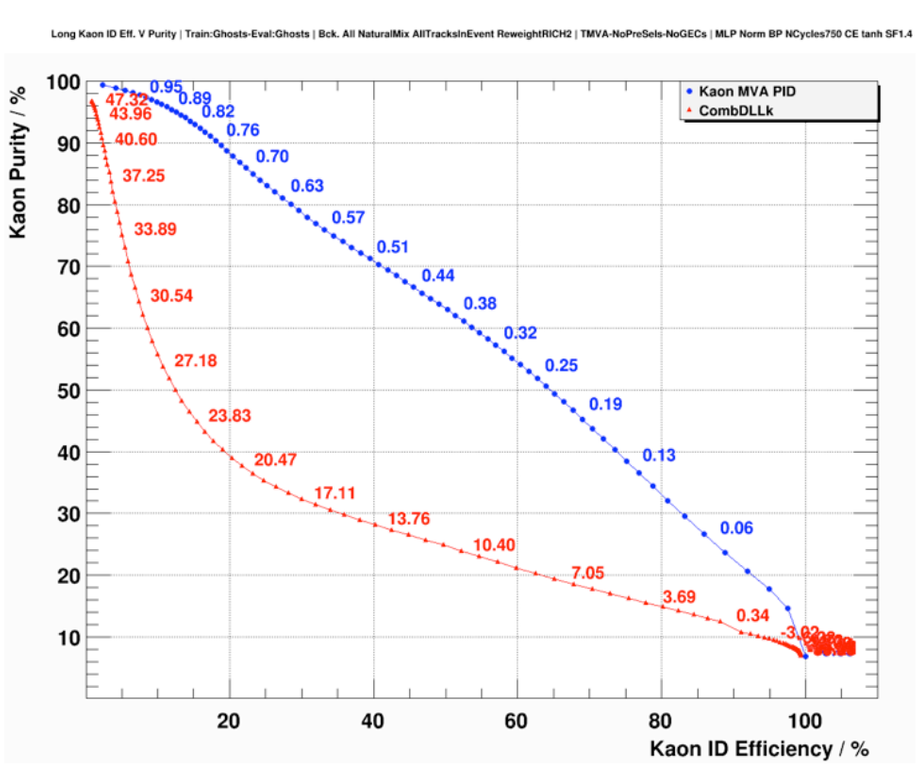
✓ Perform well
✓ Very widely used
✓ Easy to understand

× Not on stable mathematical footing
× Requires long list of cuts per track to exclude non-pion hypothesis
  - Complex and time-consuming PID calibration
  - Tighter cuts non-obvious

ProbNN

- There's more to PID than Cherenkov radiation! Correlations to PID include
  - Tracking performance
  - Track kinematics
- Combine information with an MVA, more information should mean *better PID*
- One network per hypothesis, each response is exclusive: no need to combine them
- Range [0, 1]; cut on **x** amounts to (100*__x__)% purity (in MC)
- Performs better than DLL across the board (in MC)



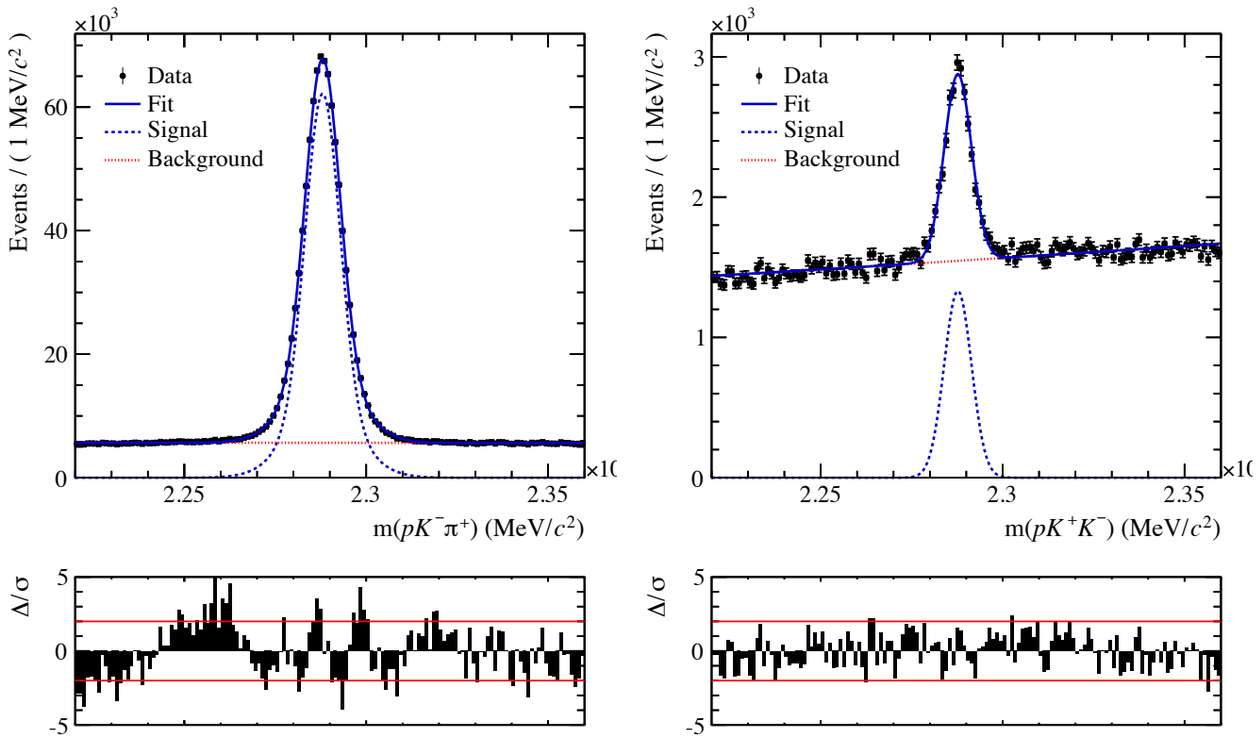


- Already in your ntuples with `TupleToolPid`, same as DLL

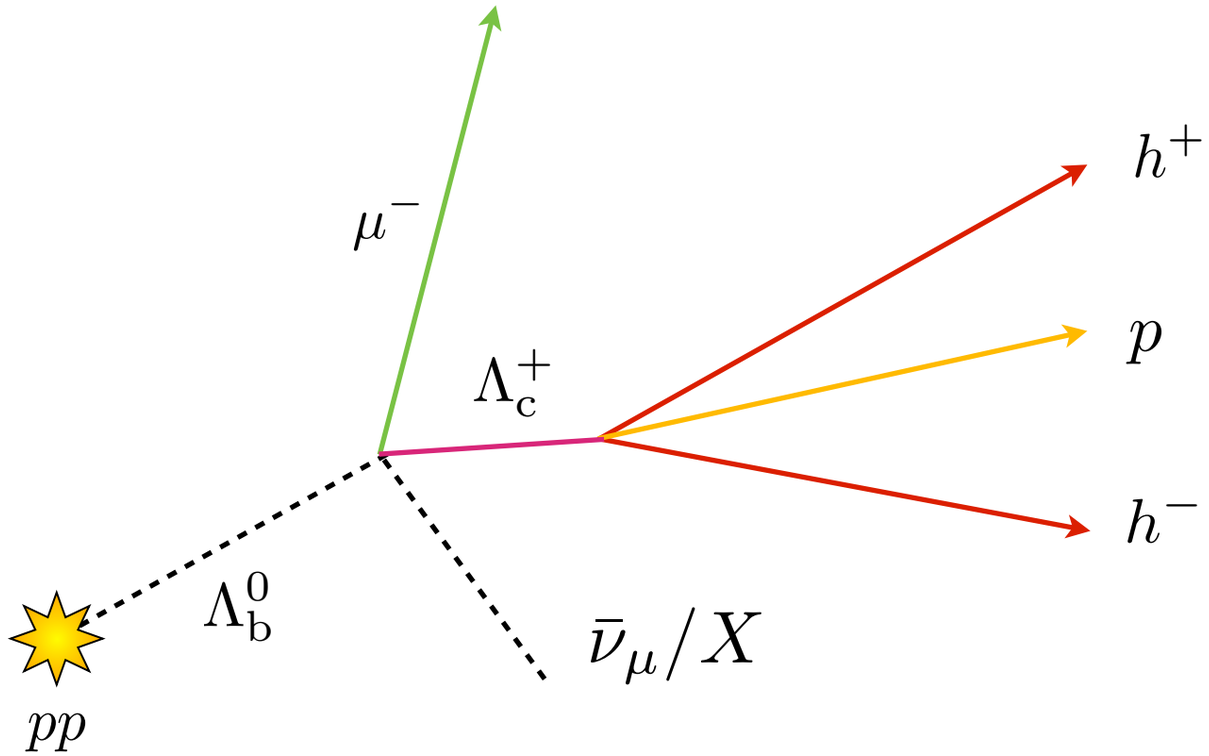
Data

- 2011 shown today, analysis will combine with 2012 (3 fb^{-1})
- Reco 14, Stripping 20r1, DaVinci v33r4 (V2 tunings $\geq$ v33r3p3)

- Branching fraction and ΔA_{CP} analysis on three Λ_c modes from semileptonic Λ_b
  - $pK^-\pi^+$
  - pK^+K^-
  - $p\pi^+\pi^-$

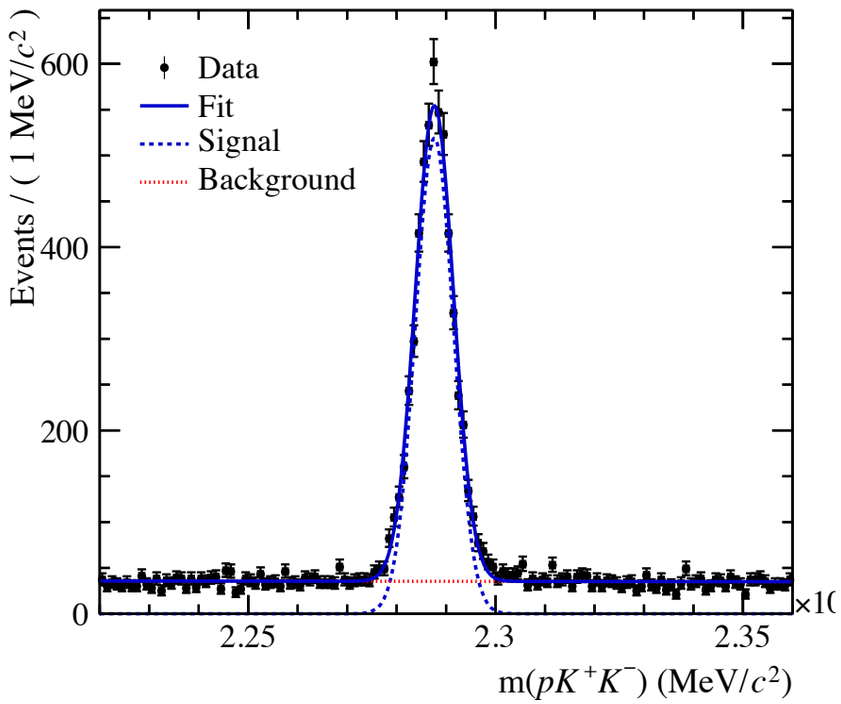


Stripped & Triggered

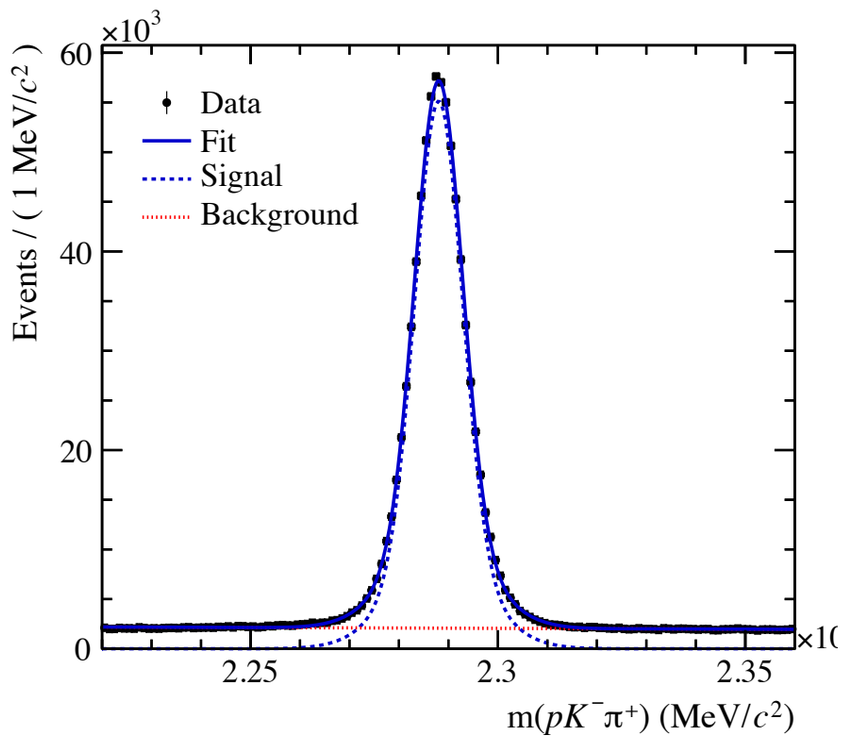


Comparison - DLL

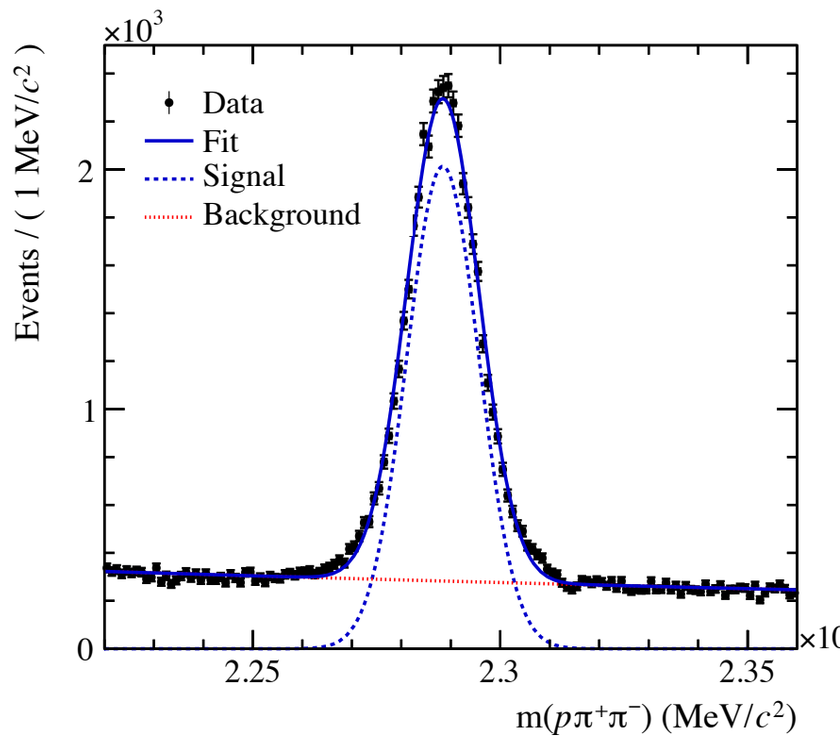
- Initial selection found by optimising significance ($\pm\ 3\sigma$ mass window), $S/\sqrt{S+B}$
- Tighten selection for higher purity, if required



Signal	5021 ± 78
Background	821 ± 13
Significance	66 ± 1



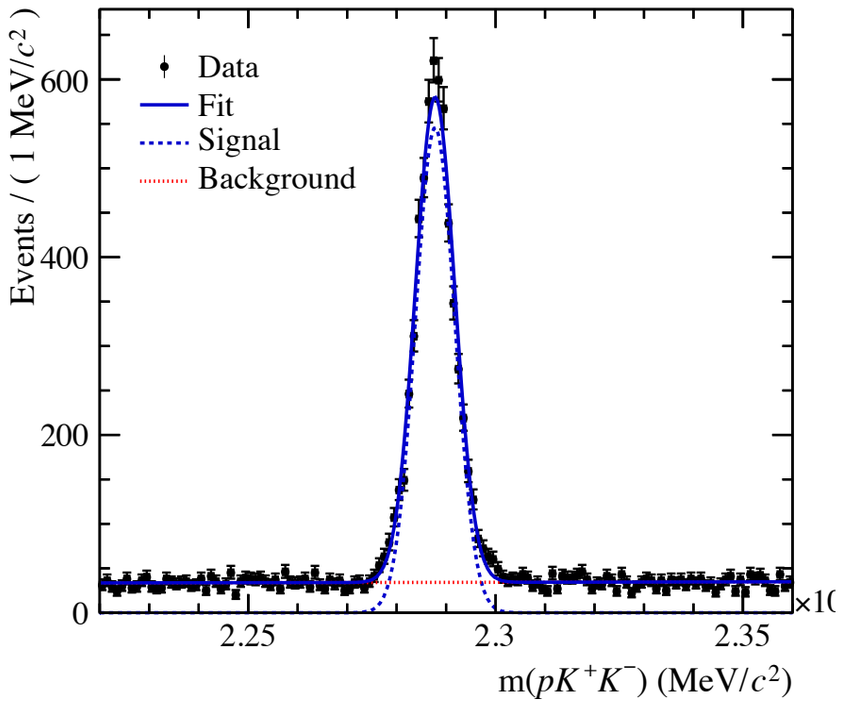
Signal	727,817 ± 989
Background	59,195 ± 157
Significance	820 ± 1



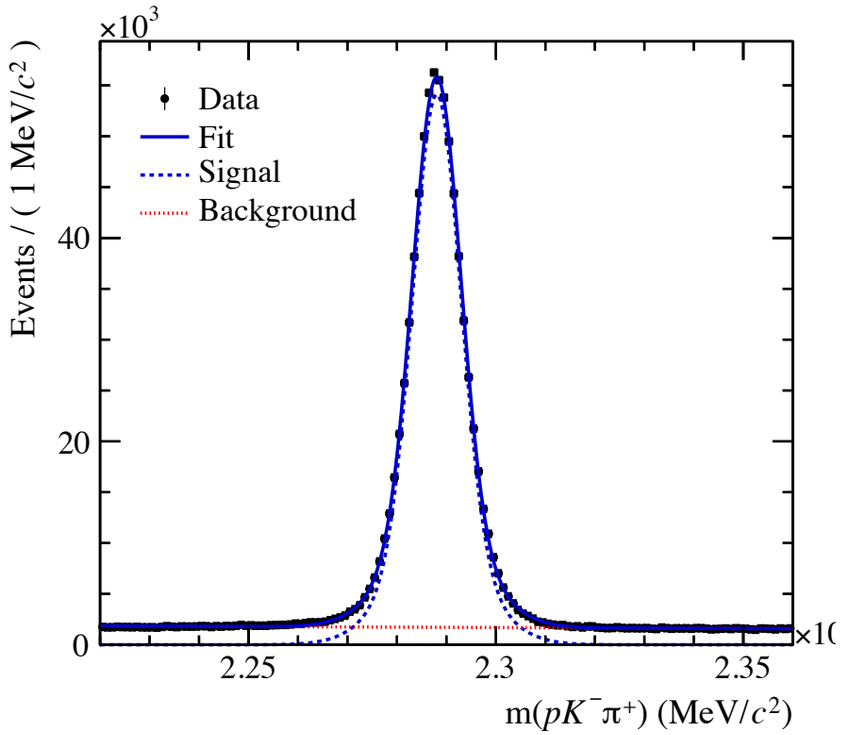
Signal	36,575 ± 233
Background	12,328 ± 75
Significance	165 ± 1

Comparison - ProbNN

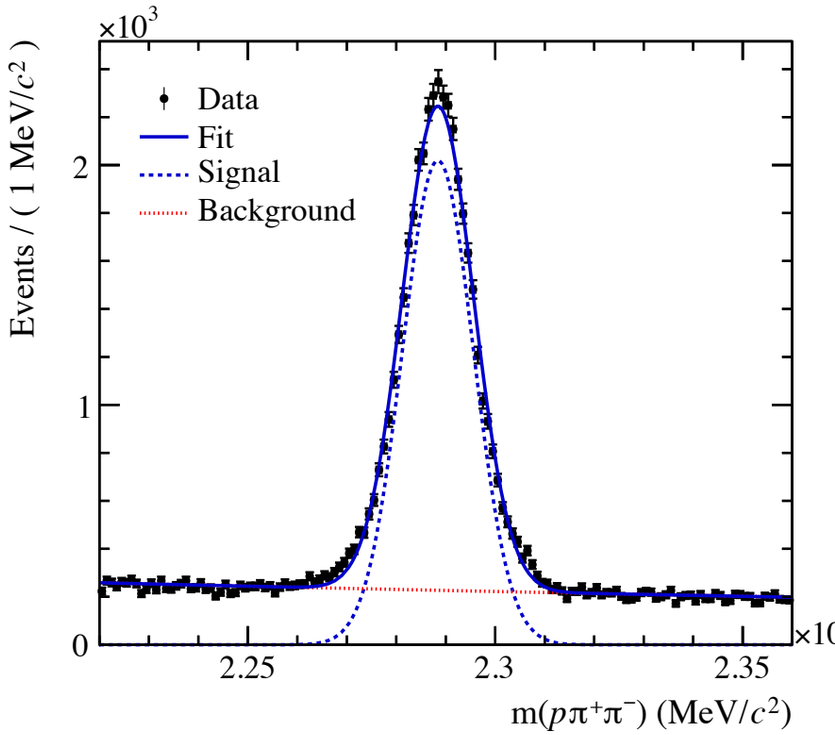
- Initial selection found by optimising significance ($\pm\ 3\sigma$ mass window), $S/\sqrt{S+B}$
- Tighten selection for higher purity, if required



Signal	5422 ± 80
Background	814 ± 13
Significance	69 ± 1



Signal	717,675 ± 942
Background	48,940 ± 136
Significance	819 ± 1

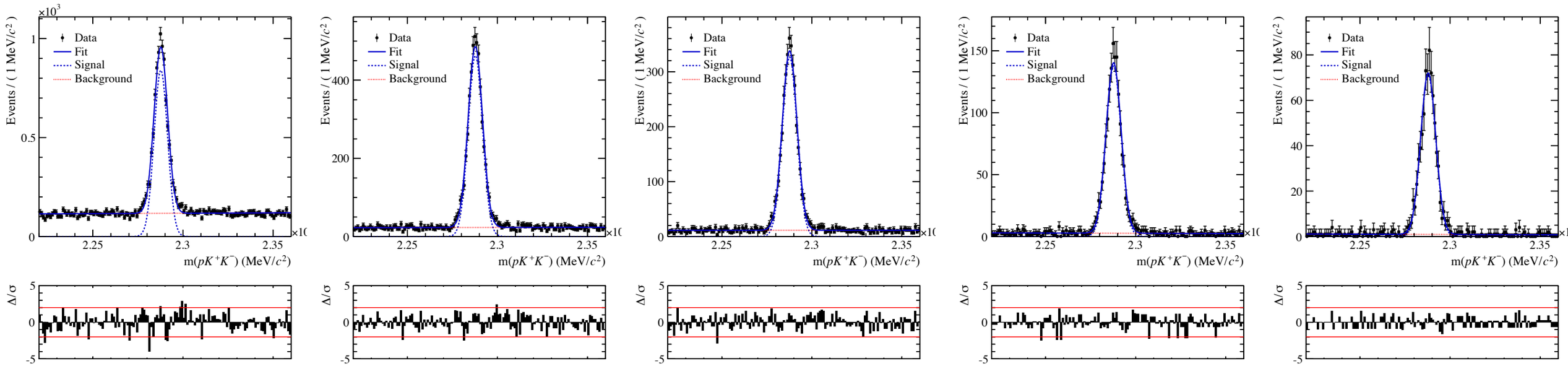


Signal	36,343 ± 225
Background	9813 ± 67
Significance	169 ± 1

Comparison - Purification

- So, SCS modes improved, CF about the same
- For this analysis, purity is preferable over signal significance
- Purifying signal trivial, just increase each track's ProbNN threshold
- Not so obvious how to proceed with DLL

Increasing ProbNN cut on all tracks, $pK^+K^- \rightarrow$



0.3 0.5 0.6 0.8 0.9

- Not perfect, closer inspection needed by analysts
- In our case, proton cut is more discriminatory than pion and kaon

Conclusions

- ProbNN generally gives purer, more significant samples with ‘optimal’ cuts
- Higher purity simple to achieve
- Less PID calibration fuss
- A+ variables, would recommend

Backup

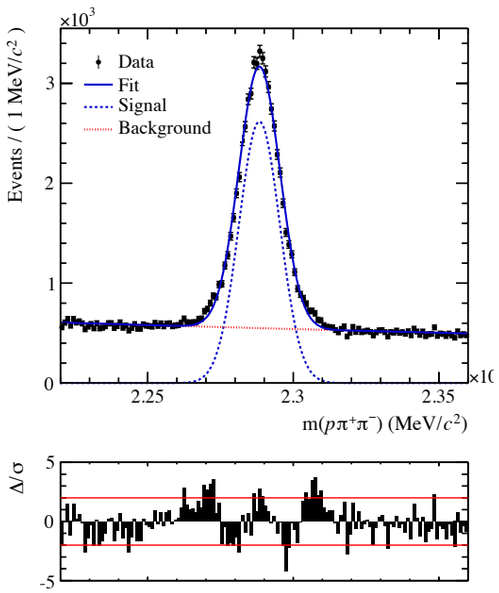
‘Optimal’ Selection

	ProbNN	DLL
$pK^-\pi^+$	proton_ProbNNp>0.2 h1_ProbNNk>0.05 h2_ProbNNpi>0.1	proton_PIDp > 14 proton_PIDp - proton_PIDK > 0 h1_PIDe < 6, h1_PIDK > 0 h1_PIDK - h1_PIDp > -20 h2_PIDK < 10, h2_PIDe < 4
pK^+K^-	proton_ProbNNp>0.45 h{1,2}_ProbNNk>0.45	proton_PIDp > 19 proton_PIDp - proton_PIDK > 7 h{1,2}_PIDK > 15 h{1,2}_PIDK - h1_PIDp > -8 h{1,2}_PIDe < 4
$p\pi^+\pi^-$	proton_ProbNNp > 0.45 h{1,2}_ProbNNpi > 0.6	proton_PIDp > 19 proton_PIDp - proton_PIDK > 9 h{1,2}_PIDe < 3 h{1,2}_PIDK < 10

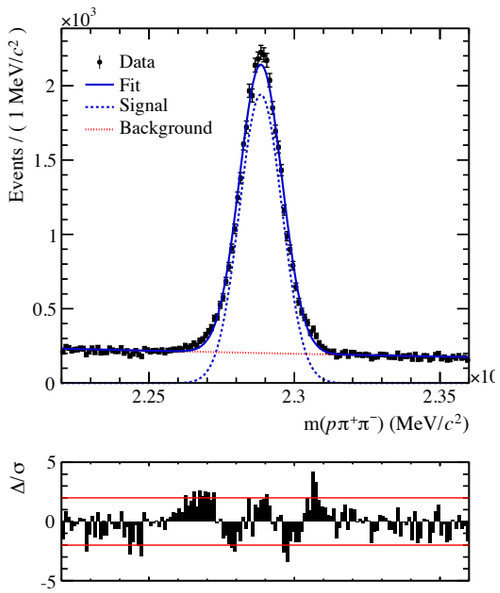
h1 and h2 represent the opposite-sign and same-sign meson, with respect to the proton, respectively

Purity Increase - $p\pi\pi$

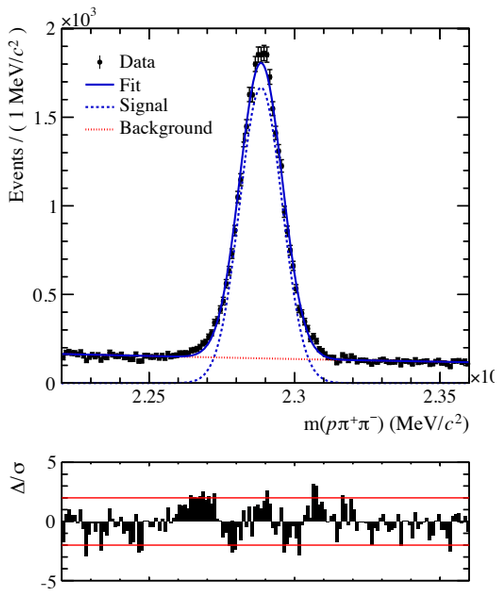
Increasing ProbNN cut on all tracks, $p\pi^+\pi^- \rightarrow$



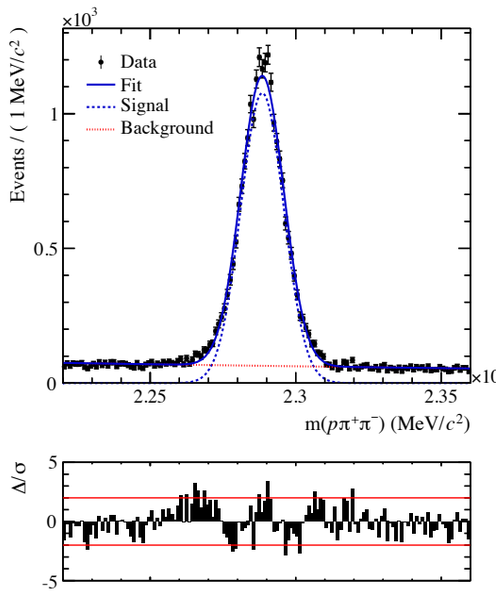
0.5



0.6



0.8



0.9