Status of the $B^+ \to K^{*+}\rho^0$ analysis in Belle

Ilya Komarov, Chiara La Licata, Diego Tonelli

Trieste

November 20, 2017

In brief

— Full Belle set and **analysed with Belle 2 software (release 1)**

— Expect O(100) $B^{+} \to K^{*+}\rho^{0}$ decays reconstructed in their $\boldsymbol{K^{+}\pi^{0}\pi^{+}\pi^{-}}$ and $\boldsymbol{K^{0}\pi^{+}\pi^{+}\pi^{-}}$ final states.

— Current best BF and f_L measurements are form Babar, no LHCb measurement so far.
Expect world best results for our analysis.

— Cut and machine-learning-based selection to suppress main backgrounds from continuum and rare B decays

— 6D fit to identify signal fraction of longitudinally polarised decays (f_L)

— Fit and efficiencies validated using control modes, possibly selected as the signal mode, like $B^{+} \to J/\psi(\to \mu\mu)\, K^{*+}$, $B^{+} \to D^{0}\pi^{+}$.

Experimental challenges

- Reconstruct a very rare signal swamped by 10^3-fold larger continuum and several poorly known irreducible peaking backgrounds

- Presence of $\boldsymbol{\pi}^0$ and a wide resonance (ρ) complicates the background discrimination

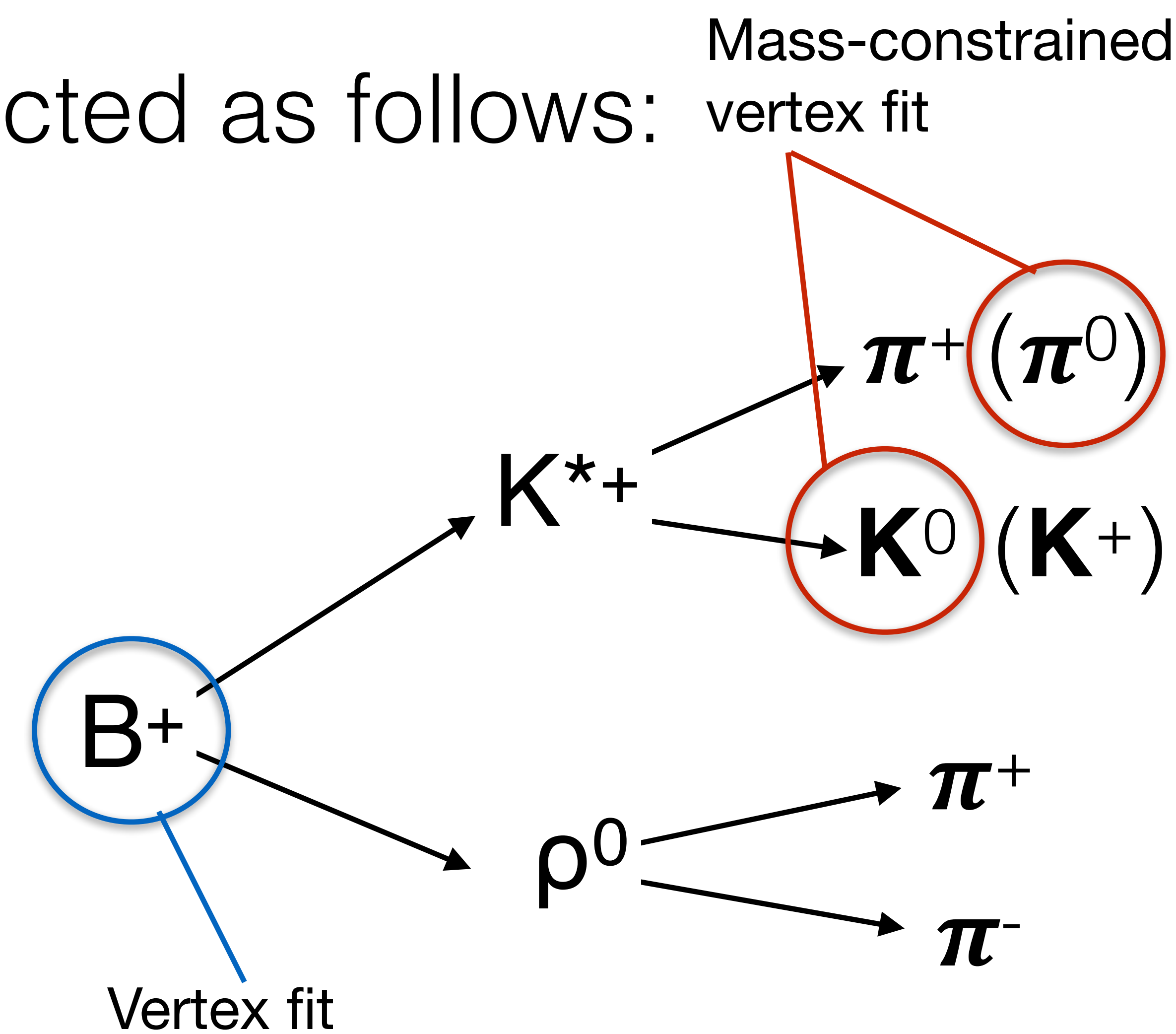
- Need for an angular analysis suggests to use selection requirements that keep to a minimum the correlations with angular variables

Reconstruction

The B-candidates are reconstructed as follows:

For each candidate, we store information about:

— Kinematics
— Track PIDs
— B-Vertex fit quality
— Gen-level info
— Flavour tag
— Continuum suppression variables



Comment on B2BII usability: the software is developing, some hiccups are unavoidable. But this is fully compensated by strong experts support (softare-b2bii@belle2.org)

Signal selection: offenders

- **Fake π^0**: suffer high fake π^0 rate, where π^0 is reconstructed using non-signal γ.

- **Self cross feed**: misreconstructed signal candidates (which remain in signal MC after removing truth-matched candidates)

- **Peaking backgrounds**: from other B decays such as ($B \to K^*K^*$ or $B \to D^0\pi$)

- **Continuum**: usual offender. Candidates built in non-BB events

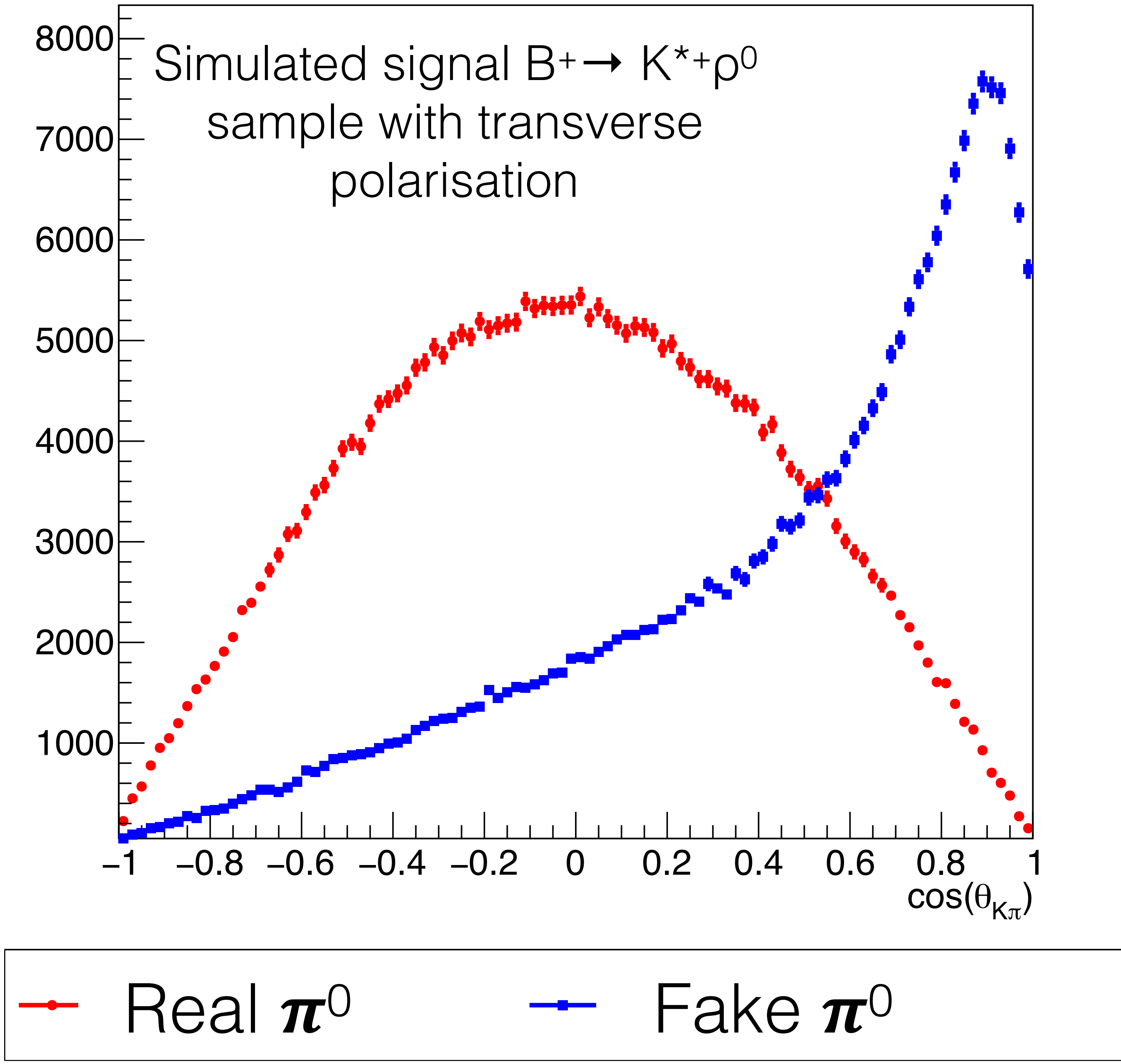
Each source of background needs its own discrimination strategy

π^0 selection

Soft and fake photons (non-photon ECL clusters), cause high rate of fake π^0.

Real π^0: both γ are truth-matched

Fake π^0: one γ is not truth-matched



Decouple the purification of photons from the main selection.

π^0 selection

"Default" Belle selection

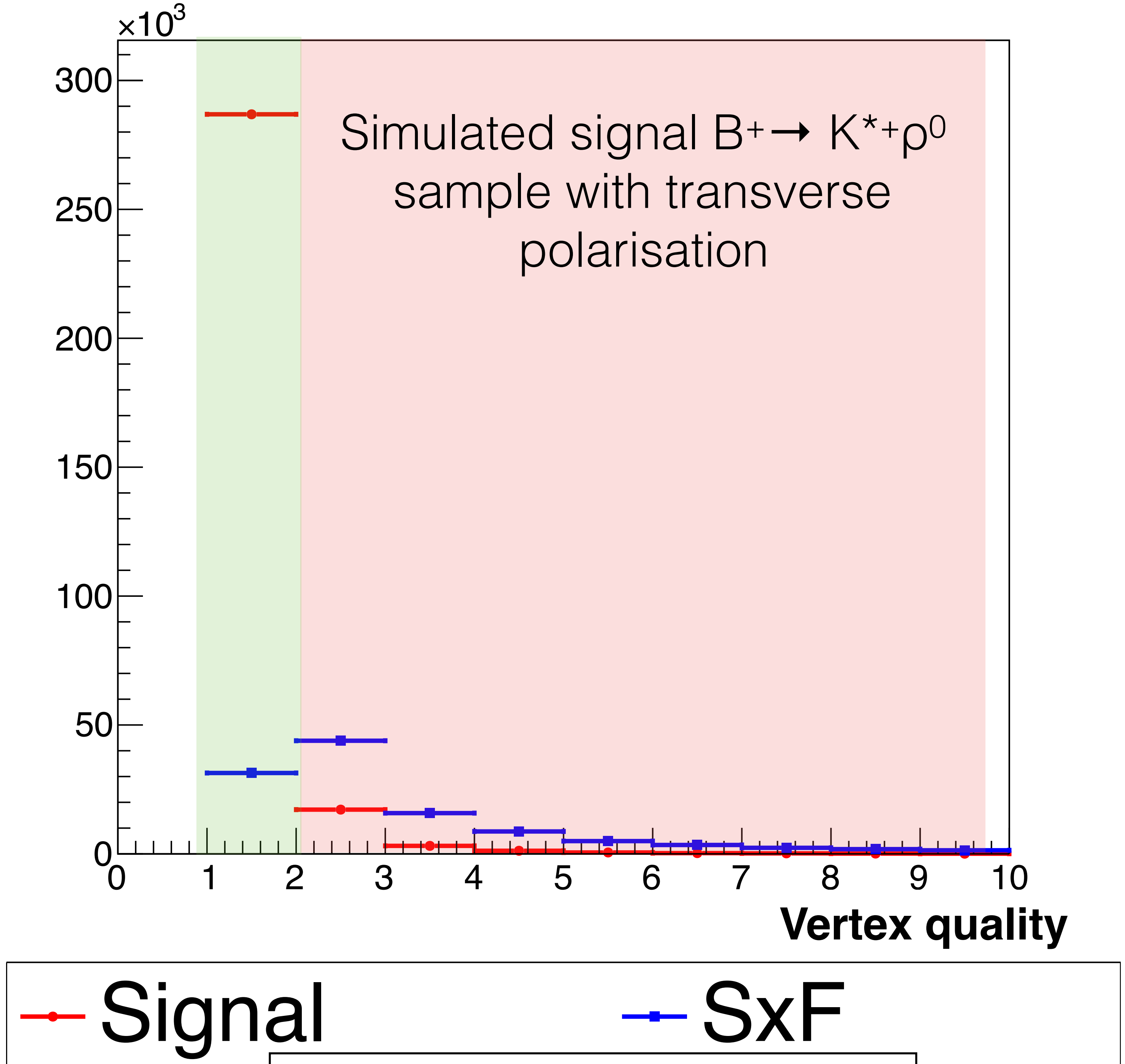
Variable	Cut
Barrel $E(\gamma)$	> 100 MeV
Endcap $E(\gamma)$	> 50 MeV
γ_{E9E25}	> 0.8
$InvM(\pi)$	$\in$[1.2, 1.5] GeV/c^2
$p^{CM}(\pi)$	> 300 MeV/c

γ_{E9E25} - ratio of energy deposited in 3x3 ECL clusters to that in 5x5.

Self cross-feed

Expect a 15-30% fraction of events with multiple candidates

Suppress using candidates with the best B-vertex fit.



Rank based on B-vertex fit quality. First bin is enriched with signal candidates.

Peaking backgrounds

Veto candidates when final-state particles combine to yield invariant masses compatible with known decays

Combination	Veto (GeV/c²)
fake K^* ($K^+\pi^-$)	$\notin[0.842; 0.942]$
fake D^0 ($\pi^+\pi^-\pi^0$)	$\notin[1.6; 2.1]$
fake D^0 ($K^+\pi^-\pi^0$)	$\notin[1.6; 2.1]$

Combination	Veto (GeV/c²)
fake K^* ($K^0_S\pi^+$)	$\notin[0.842; 0.942]$
fake D^0 ($\pi^+\pi^-K^0_S$)	$\notin[1.8; 2]$
fake D^- ($K^0_S\pi^-$)	$\notin[1.8; 2]$

Remainder pollution from rare decays is studied with Rare MC samples (Mixed and Charged)

CB suppression

Baseline signal-to-background ratio is **1:1000**

Goal: Suppress continuum background.

Requirements:

1. Sufficient independence from fit variables (M_{bc}, ΔE, M_K, M_ρ, θ_K, θ_ρ)
2. Should suit both signal and control channels ($B^+ \to K^{*+}\rho^0$ and $B^+ \to J/\psi K^{*+}$) so that we can validate the full analysis on the control mode

Default approach

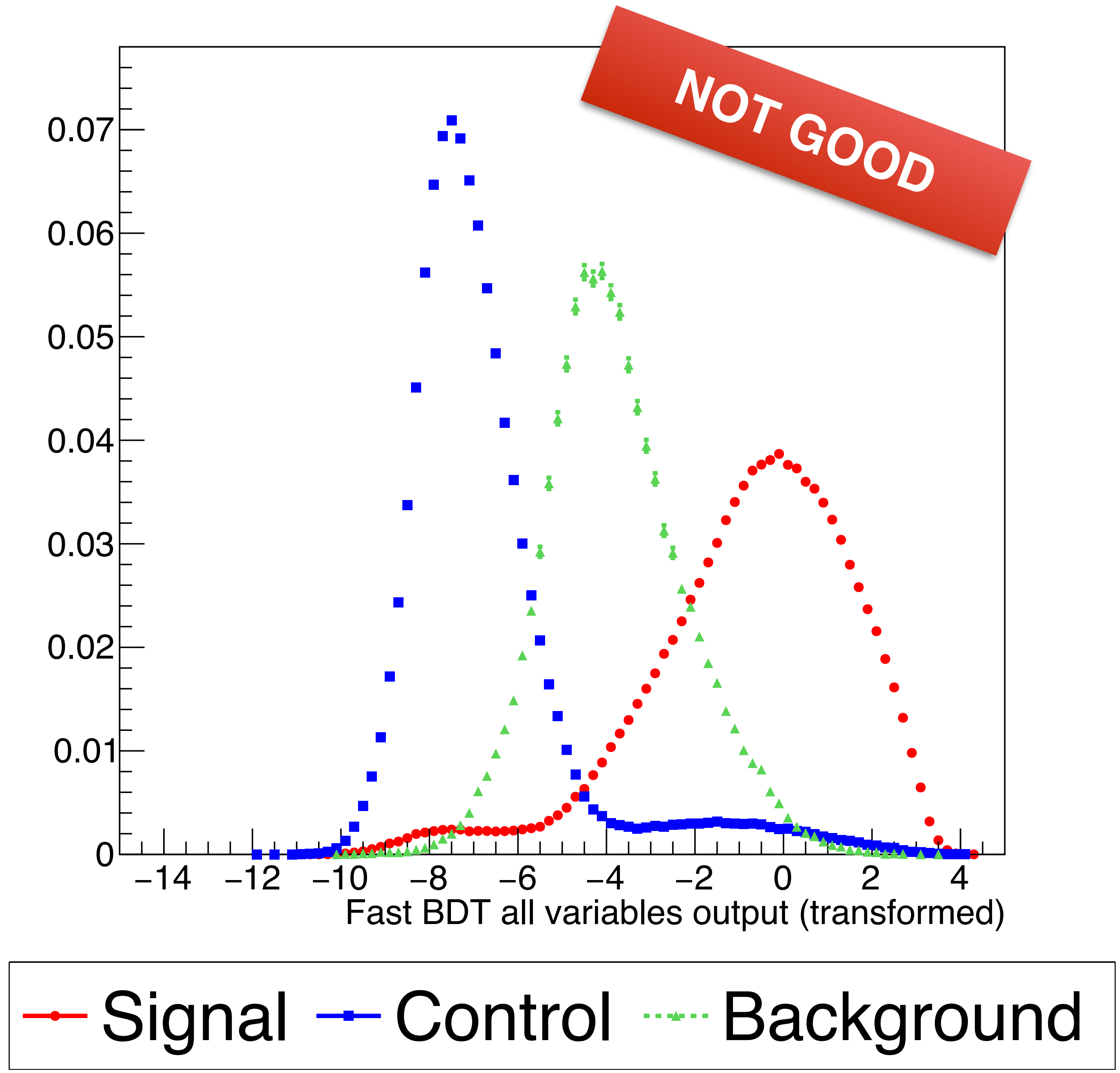
Use maximum discriminating power from event topology:

- Cleo Cones,
- KSFW variables,
- Thrust-related variables,
- dz,
- Flavour tag.

Default approach

The use of discriminating variables that use B thrust and CLEO cones picks up on significant kinematic differences between $K^{*+}\rho^0$ signal and $J/\psi K^{*+}$ control modes yielding different classifier outputs.

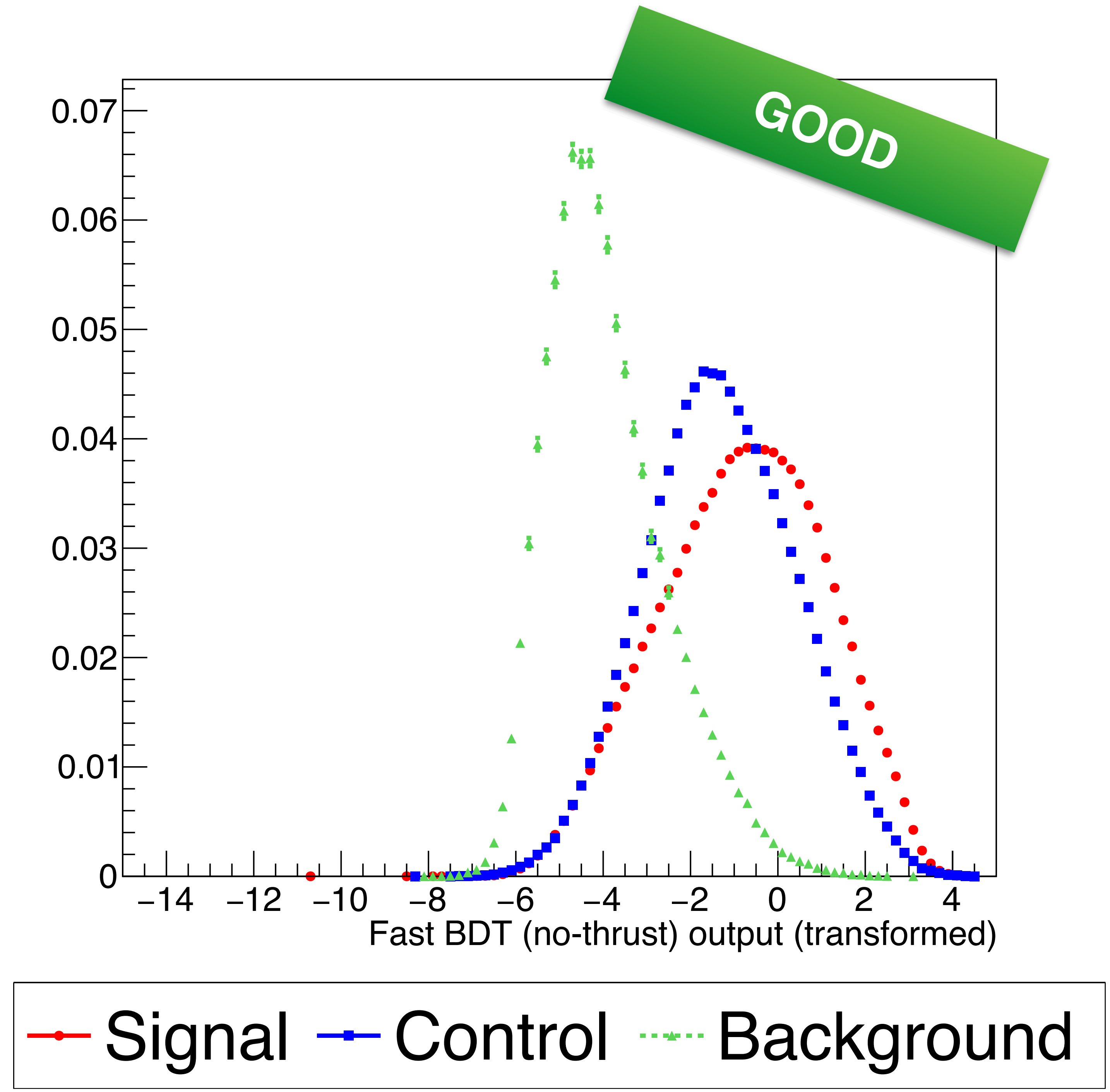
This vanifies our strategy of using the same selection for signal and control.



Current approach

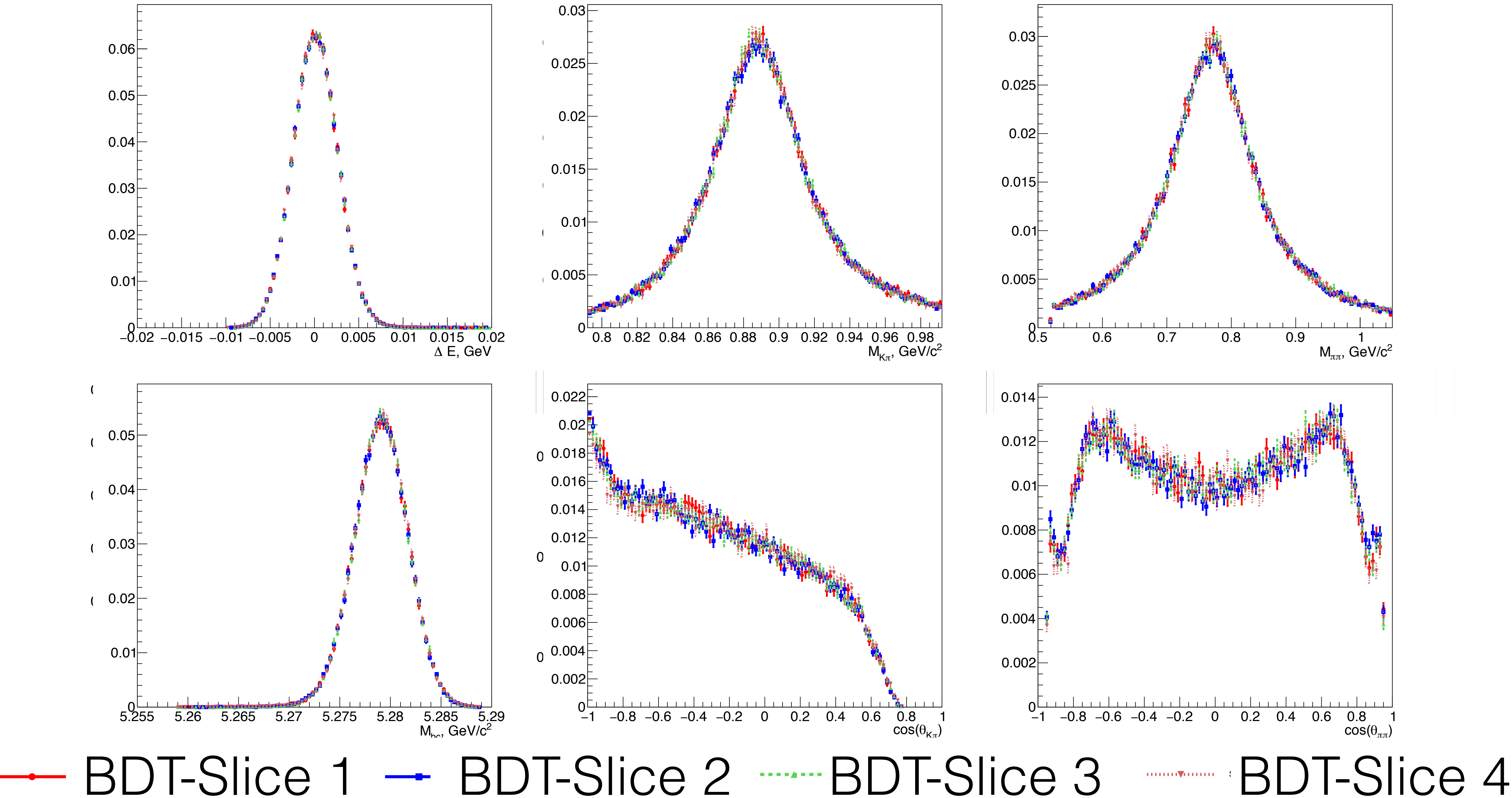
We excluded magnitude of B thrust and Cleo cones. It helped.

In release 1, it's possible to calculate Cleo Cones for the ROE - only. We are studying the effect now.



Dependencies on fit variables

New BDT nicely independent from fit variables:



Preliminary MVA performance

$K^+\pi^0\pi^+\pi^-$ final state

FOM	N^{sig}	$(\epsilon^{\mathrm{sig}})$	N^{bgr}	$(\epsilon^{\mathrm{bgr}})$
50% Efficiency	121.0	(49.0%)	1676.0	(1.69%)
95% Efficiency	231.0	(93.0%)	33528.0	(34.0%)
99% Efficiency	231.0	(93.0%)	33528.0	(34.0%)
99% Purity	8.6	(3.47%)	1.0	(0.00101%)
Custom	206.0	(83.0%)	15657.0	(15.8%)
S/sqrt(B)	8.6	(3.47%)	1.0	(0.00101%)
S/sqrt(S+B)	40.0	(16.2%)	75.0	(0.076%)

$K^0_S\pi^+\pi^+\pi^-$ final state

FOM	N^{sig}	$(\epsilon^{\mathrm{sig}})$	N^{bgr}	$(\epsilon^{\mathrm{bgr}})$
50% Efficiency	103.0	(49.0%)	484.0	(2.32%)
95% Efficiency	196.0	(94.0%)	8808.0	(42.0%)
99% Efficiency	205.0	(99.0%)	15266.0	(73.0%)
99% Purity	5.7	(2.73%)	0	(0%)
Custom	196.0	(94.0%)	8808.0	(42.0%)
S/sqrt(B)	14.6	(7.0%)	3.0	(0.0144%)
S/sqrt(S+B)	70.0	(34.0%)	144.0	(0.69%)

Not bad, but we think we can do better.
Improved tuples with improved selection in progress.

Selection summary

Skim:

$M(B^+)$	$\in[4.8, 5.5]\mathrm{GeV}/c^2$
$M(\rho^0)$	$\in[0.5, 1.2]\mathrm{GeV}/c^2$
$M(K^{*+})$	$\in[0.692, 1.092]\mathrm{GeV}/c^2$
π PID	>0.3
$M(J/\psi)$	$\in[2.95, 3.25]\mathrm{GeV}/c^2$
μ PID	>0.3

π^0

Default selection

K^0_S

good K^0_S

Fit ranges:

$M(B^+)_{bc}$	$\in[5.255, 5.289]\mathrm{GeV}/c^2$
$M(\rho^0)$	$\in[0.52, 1.05]\mathrm{GeV}/c^2$
$M(K^{*+})$	$\in[0.792, 0.992]\mathrm{GeV}/c^2$
ΔE	$\in[-0.02, 0.02]\mathrm{GeV}$
$\cos(\theta_{K\pi})$	$\in[-1, 0.92]$
$\cos(\theta_{\pi\pi})$	$\in[-0.95, 0.95]$

Extra:

K, π PID	>0.6
FBDT	to be defined
B-Vertex	Best candidate
MisRec.	Set of vetoes

Selection summary

Skim:

$M(B^+)$	$\in[4.8, 5.5]\mathrm{GeV}/c^2$
$M(\rho^0)$	$\in[0.5, 1.2]\mathrm{GeV}/c^2$
$M(K^{*+})$	$\in[0.692,1.092]\mathrm{GeV}/c^2$
π P	
M(J	
μ P	

Fit ranges:

$M(B^+)_{bc}$	$\in[5.255, 5.289]\mathrm{GeV}/c^2$
$M(\rho^0)$	$\in[0.52, 1.05]\mathrm{GeV}/c^2$
$M(K^{*+})$	$\in[0.792,0.992]\mathrm{GeV}/c^2$
	2]GeV
	92]
	.95]

- Despite the vetoes, some irreducible peaking backgrounds remain in the sample.
- Characterize them using the official Belle rare sample

π^0

Default selection	good K^0_s

FBDT	to be defined
B-Vertex	Best candidate
MisRec.	Set of vetoes

Working with Rare MC

- Use MCHierarchyTool to access decay string for each candidate

- Identify the processes that contribute backgrounds that survive our selection

- Isolate those that contribute a yield comparable with the uncertainty on the signal yield. Add the inclusive shape of the remainder candidates in the fit.

- Model each of the major contributors exclusively, include them to the fit using up-to-date measurements of BF to constrain the yield.

Preliminary

RareMC breakdown

$K^0\pi^+\pi^+\pi^-$ final state

Decay	#candidates [%] of expected signal yield
$B^+ \rightarrow K^{*+}\pi^+\pi^-$	152
$B^+ \rightarrow K^{*+}\rho^0$	100
$B^+ \rightarrow K^{*+}f_0$	89
$B^+ \rightarrow K^{*+}K^{*0}$	32
$B^+ \rightarrow K^{*+}f_2(1430)$	30
$B^+ \rightarrow K^*(1410)^0\pi^+$	29
$B^+ \rightarrow K^*_0(1430)^+\rho^0$	26
$B^+ \rightarrow a_1(1260)^+K^0$	19
$B^+ \rightarrow \rho^0 K^0_S\pi^+$	10
$B^+ \rightarrow K^{*+}\eta'$	6
Others	20

$K^+\pi^0\pi^+\pi^-$ final state

Decay	#candidates [%] of expected signal yield
$B^+ \rightarrow K^{*+}\pi^+\pi^-$	183
$B^+ \rightarrow K^{*+}\rho^0$	100
$B^+ \rightarrow K^{*+}f_0$	99
$B^+ \rightarrow K^{*+}f_2(1430)$	39
$B^+ \rightarrow K^{*+}K^{*0}$	34
$B^+ \rightarrow K^*(1410)^0\pi^+$	26
$B^+ \rightarrow K^*_0(1430)^+\rho^0$	25
$B^+ \rightarrow \rho^0 K^+\pi^0$	7
$B^+ \rightarrow \rho^+\rho^0$	7
Others	28

PID correction

Belle MC does not describe PID variables correctly. This introduce bias during estimations of selection efficiency and composition of the background sample.
Belle had recipes to account for the differences.

These recipes are not compatible with Belle 2 software.

We developed code to weight tracks of selected samples according to Belle recipes: *Check here*. Working on inclusion of the code to official basf2 release.

Multidimensional Fitter Framework

In brief: Custom C++ wrap around RooFit providing simple access to configuration of multidimensional multicomponent fit. Logger and plotter included. Fully defined from config files:

1. Define observables

```
B__Mbc_corr 5.26 5.29 MB
```

2. Define fit variables

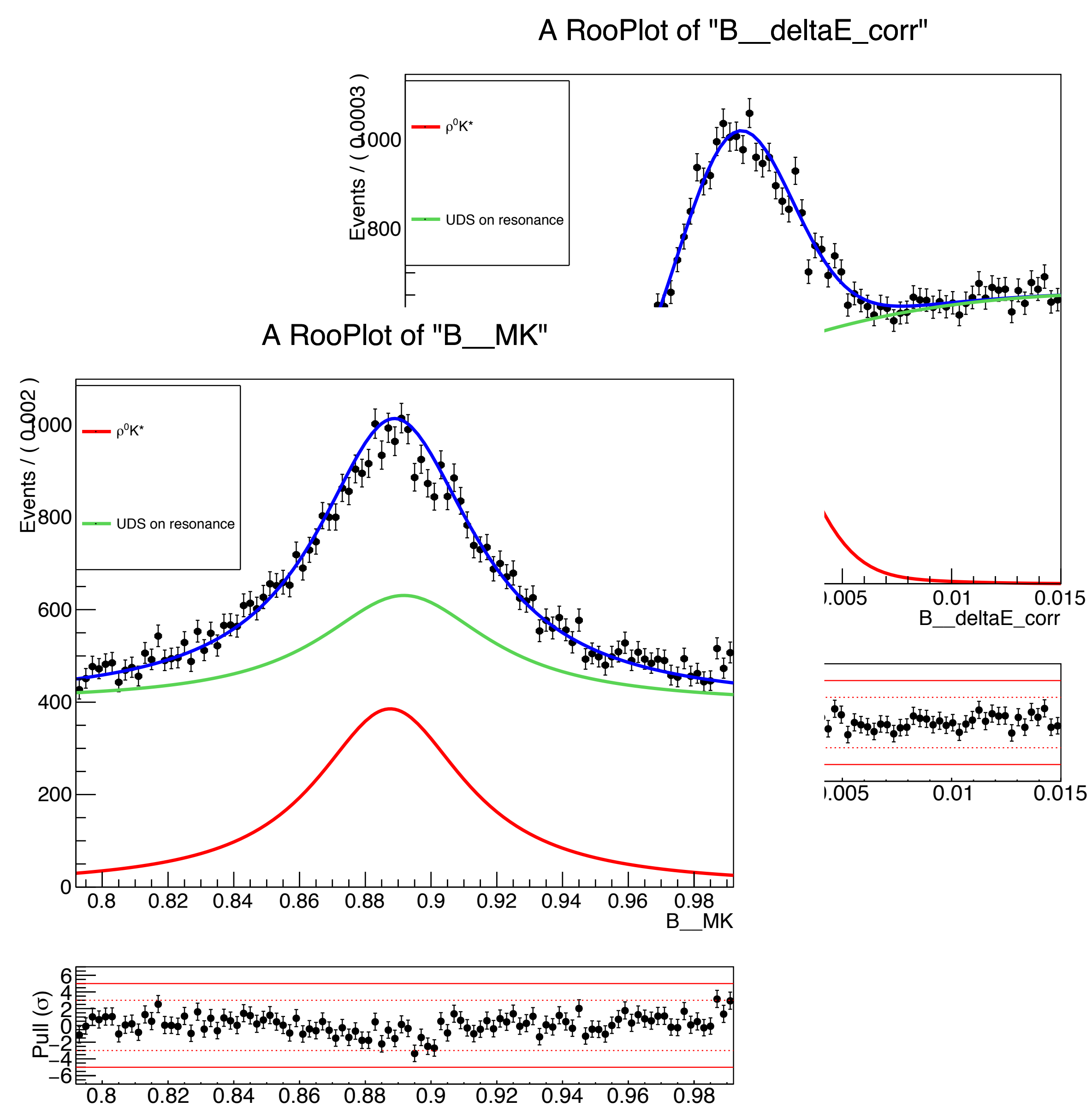
```
Sign_MB_CB_alpha 1 4 2 Unfixed
```

3. Define contributions

```
Sign  : CB(MB) Pol5(HR)
```

4. Define input data

```
MC : 1 - workspace.root
```



Multidimensional Fitter Framework

— Framework can be used for any fits. No restrictions on dimensionality or number of components. Just add what you need in config file.

— Easy to use for toy studies: classes in framework create RooAbsPdf from descriptions in config files and built-in logger will keep track of results of all fits.

— Smart plotter keeps track of all drawings - legends and colours are defined in config and are consistent across all plots.

—Package contains detailed instructions, examples, documentation

Framework is publicly accessible: Stash repository

Selection optimisation

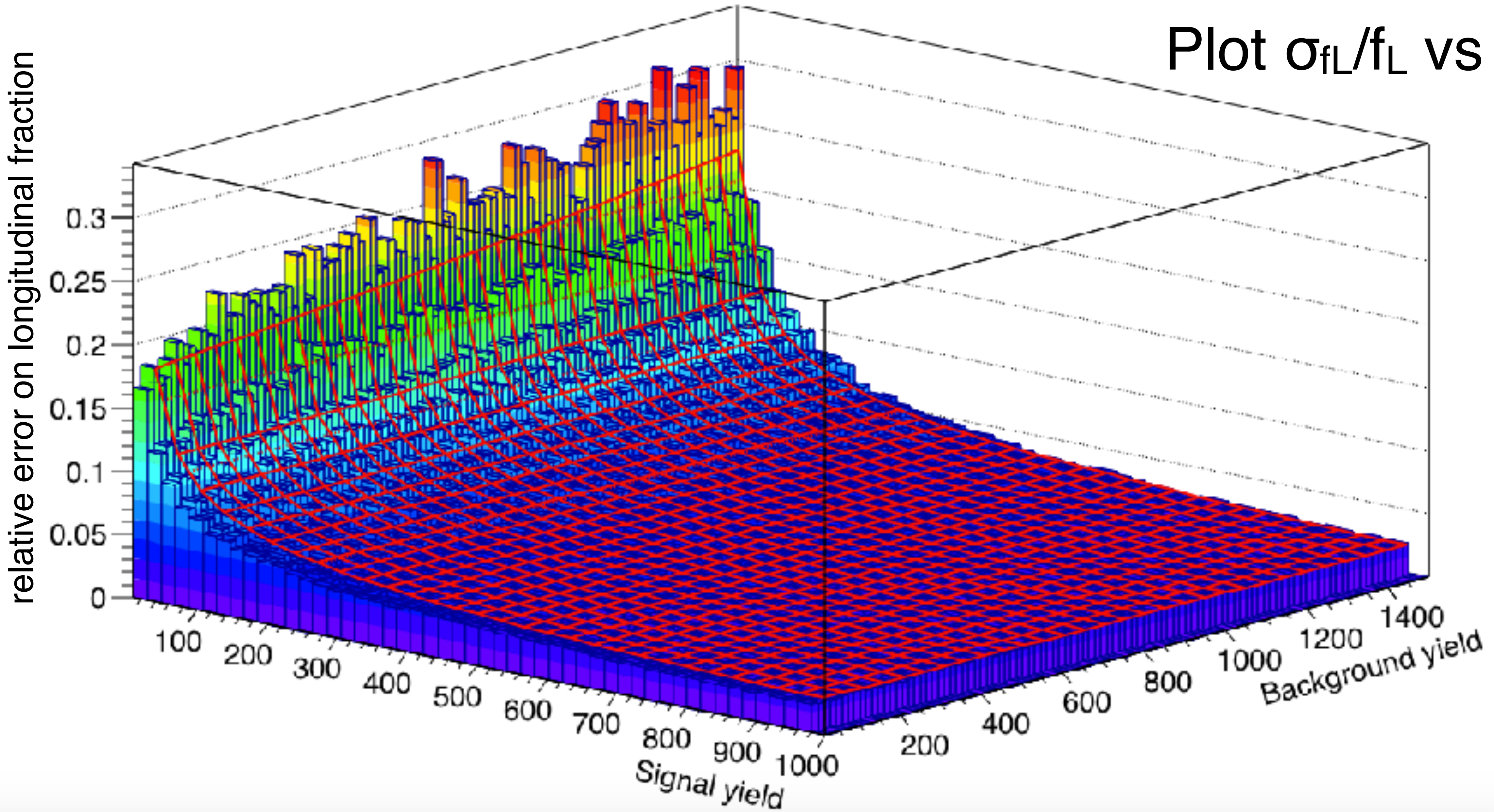
Low signal, high background ⇒ selection optimisation.

To get an idea, BaBar sees 85 signal events with a continuum background of ~2500. **[Phys. Rev. D83:051101,2011]**

- define a figure of merit: the average expected uncertainty on the polarization fraction σ_{fL}/f_L
- find empirically the dependence of σ_{fL}/f_L on the signal and background yields: $\sigma_{fL}/f_L \approx f(S,B)$
- minimize f(S,B) over the space of the cuts in the discriminating variables identified

Figure of merit

Plot σ_{f_L}/f_L vs S and B



$$f(S,B) = \frac{S}{(S^{p0} + p1 * B^{p2})}$$

Consistency check

In order to check our approximated procedure compare out findings with the real world uncertainty obtained by BaBar

Using 85 signal events overlapping about 2680 continuum background events in the $K^0{}_S\pi^+$ final state BaBar obtains a relative uncertainty on the longitudinal fraction of 17% (only statistical uncertainty considered)

In our toy MC, this point in (S,B) space corresponds to a relative uncertainty of 14%.

Not exactly spot on, but the difference is sufficiently small for confirming the soundness of the procedure in addition the difference might comfortably be explained by the assumptions made in our work

Summary

The measurement of $B^+ \to K^{*+}\rho^0$ BF and polarization fractions in Belle data is in an advanced state:

- signal selection (the crux of this analysis) nearly finalized
- good handles on fake pi0 and multiple candidates
- good discrimination of continuum
- now studying remaining peaking backgrounds
- same selection used for the J/Psi control mode. Considering to also add an hadronic control mode.
- the fitter framework is up and running. Need to fill up the details and test it.

Backup

π^0 selection

"Default" Belle selection

Variable	Cut
Barrel E(γ)	> 100 MeV
Endcap E(γ)	> 50 MeV
γ_{E9E25}	> 0.8
InvM(π)	$\in$[1.2, 1.5] GeV/c^2
$p^{CM}(\pi)$	> 300 MeV/c

Current π^0 selection.

A FastBDT trained for each of the two photons.

Inputs:

- π^0 mass
- π^0 opening angle
- π^0 χ^2 prob.
- γ min C2HDist
- γ cluster E9E25
- γ Energy
- γ cos(θ)

Min. cluster-to-hist distance

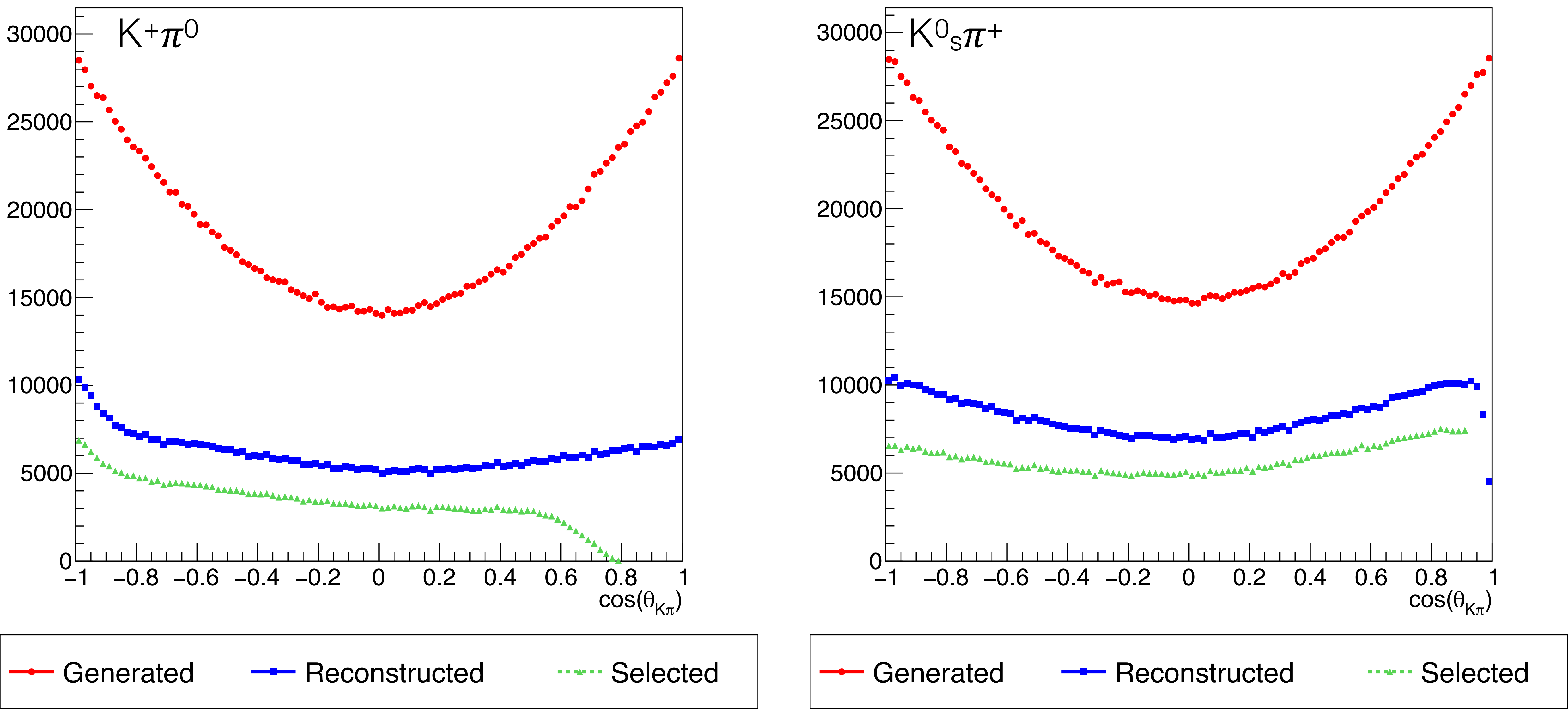
Ratio of energy deposited in 3x3 ECL clusters to that in 5x5.

However, classifier might correlate with fit variables.

Keep it as "plan B"

Selection bias

Selection scalps variable distributions.
This bias will be taken into account in the fit.



Correlations

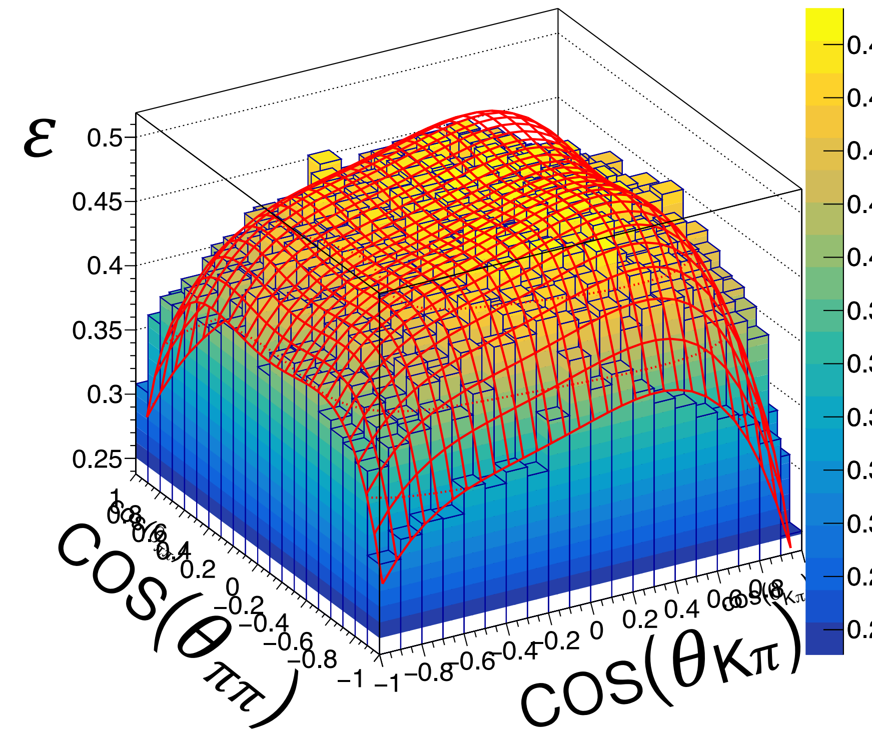
Physics correlations between fit variables are known, but nontrivial acceptance/efficiency correlations are introduced by the selection.

They will affect the ML estimator through the normalization. We've done some exploratory work to investigate this but will dig deeper once the selection will be finalized.

Correlations in fit:

$$\mathrm{PDF} = \mathrm{PDF}(M_{K\pi}, M_{\pi\pi}, M_{bc}, \Delta E) \times \times \mathrm{PDF}(\theta_{K\pi}, \theta_{\pi\pi})$$

$$\mathrm{PDF}(\theta_{K\pi}, \theta_{\pi\pi}) = \mathrm{PDF}^{\mathrm{GEN}}(\theta_{K\pi}, \theta_{\pi\pi}) \times \times \varepsilon(\theta_{K\pi}, \theta_{\pi\pi})$$



$\varepsilon(\theta_{K\pi}, \theta_{\pi\pi})$ from simulation

Fitter status

$B^+ \rightarrow \rho^0 K^{*+}$ yield and longitudinal fraction given by an unbinned maximum likelihood

6 discriminating variables:

- B mass
- ρ^0 mass
- ρ^0 helicity angle
- delta E
- K^* mass
- K^* helicity angle

Checked to be independent (apart from helicity angles)
Effectively a 6 x 1D fit.

Our first goal is to have a running fit machinery. Hence assume various simplifications

* only uds continuum background is used.

* Simple model for each pdf — do not care about accurate modelling for the moment

* Assume cross-feed-free sample

* Testing the fit without any selection to decouple the issues associated with the fit machinery to those associated with the possible non-independence of the pdf on each other

First tests

2 fit components:

- Signal $B^+ \rightarrow \rho^0 K^{*+}$
- continuum uds background *

final pdf -> $P = f_{sig} * P_{sig} + (1-f_{sig}) * P_{bkg}$

$P_{sig} = f_{ln} * P_{ln} + (1-f_{ln}) * P_{tr}$

P_{tr} = P_ΔE * P_Bmass * P_Kmass * P_ρmass * P_Khel_tr* P_ρhel_tr

P_{ln} = P_ΔE * P_Bmass * P_Kmass * P_ρmass * P_Khel_ln* P_ρhel_ln

a) toy MC studies (draw events from the pdf and fit them under various configurations)

b) fit simulated signal and uds background MC without any selection applied

- 6D fit -> 2 floating parameters: the fraction of signal: f_{sig}, f_{ln}

fitter validation - toy MC studies

1- 500 sets of 10^5 events each generated with different configurations:

(a) signal fraction (f_{sig}) = 0.1

(b) signal fraction (f_{sig}) = 0.4

(c) signal fraction (f_{sig}) = 0.9

(i) longitudinal signal fraction (f_{ln}) = 0.2

(ii) longitudinal signal fraction (f_{ln}) = 0.465

(iii) longitudinal signal fraction (f_{ln}) = 0.8

for each signal fraction 3 samples with different longitudinal fractions are generated

2- 800 sets of 10^3 events each generated with:

(a) signal fraction (f_{sig}) = 0.1 longitudinal signal fraction (f_{ln}) = 0.78

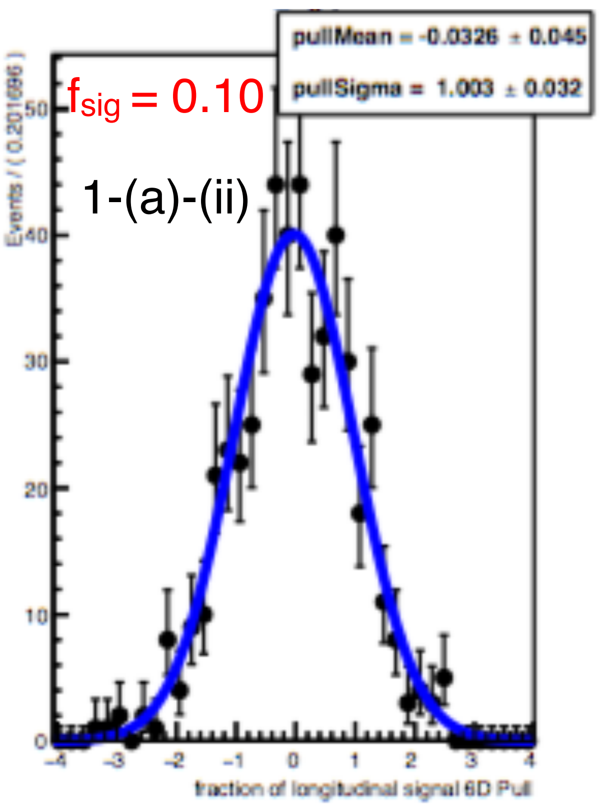
~100 signal events are produced to mimic what we expect

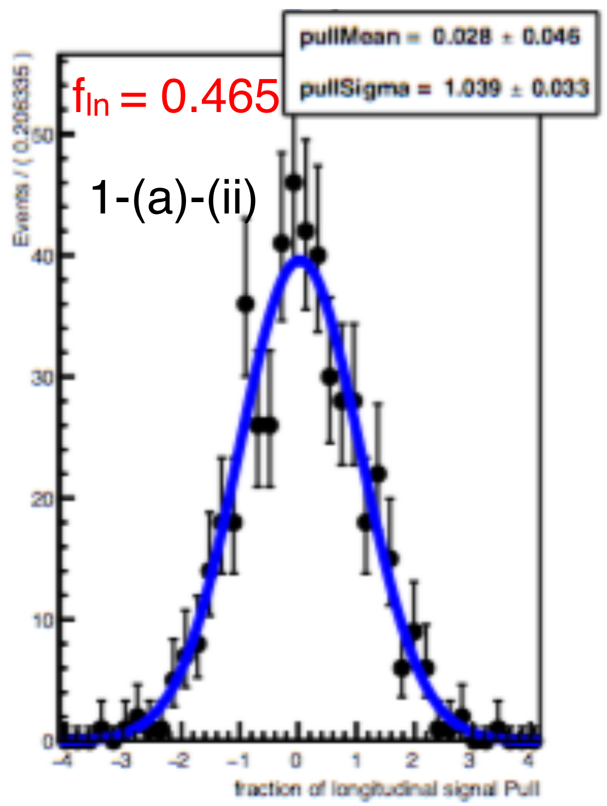
$\longrightarrow$ pull distributions of f_{sig} and f_{ln} produced and fitted with a gaussian function

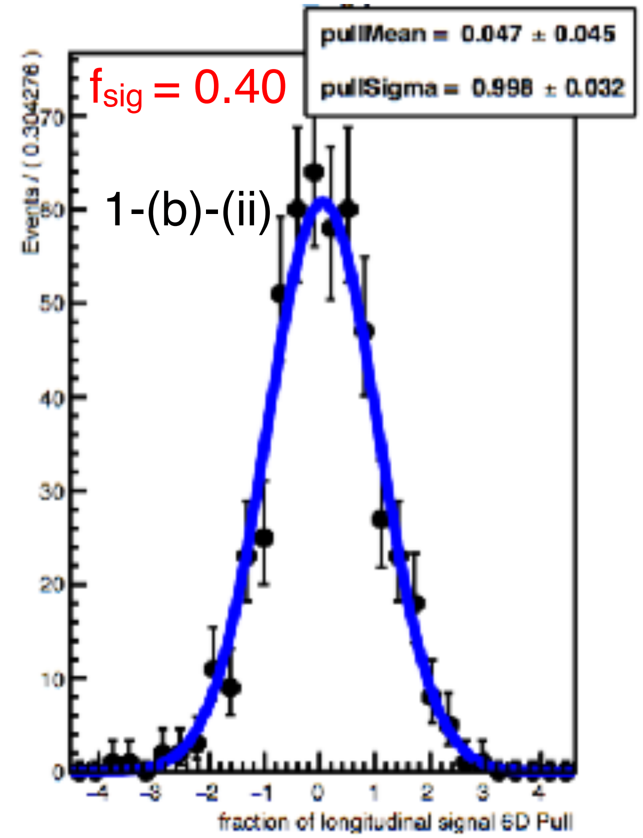
fit estimates are unbiased and the uncertainties are gaussian in all the configurations

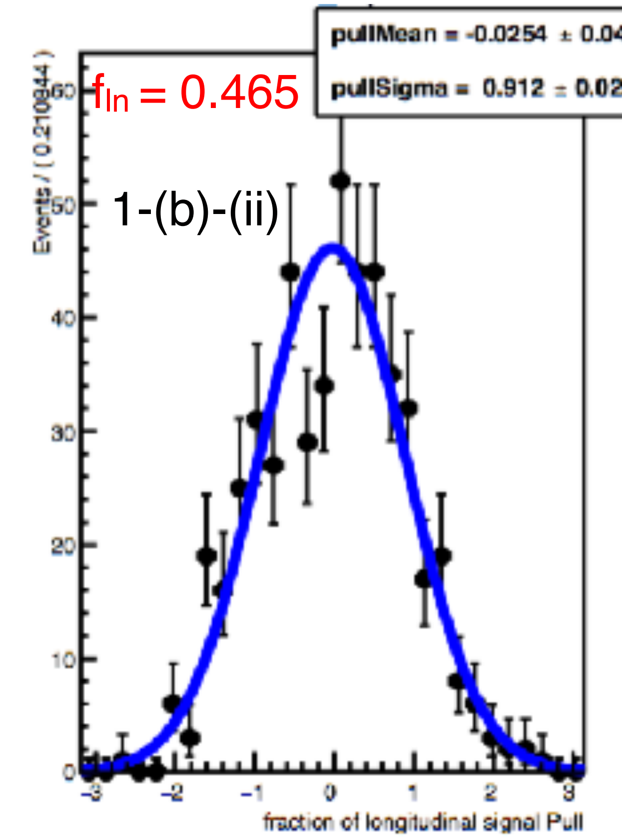
Example: toy pulls

Generated 500 sets of 10^5 events with fixed fraction of signal (fsig = 0.1, 0.4, 0.9) and fixed fln (=0.465)



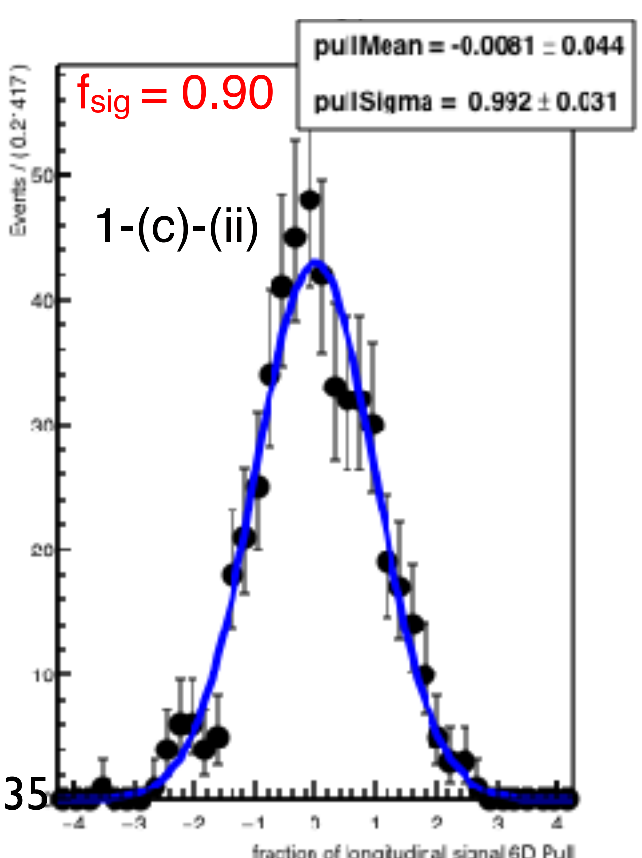


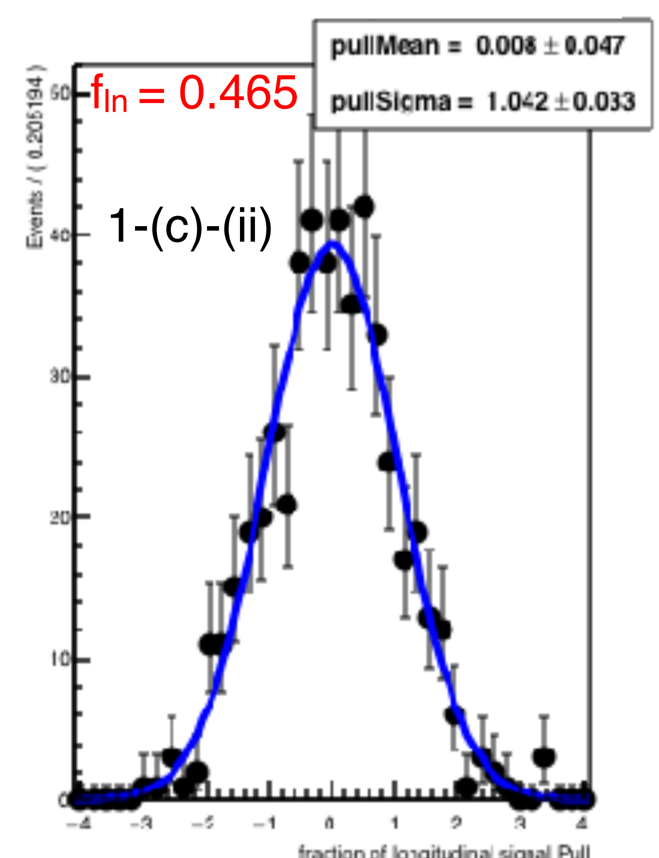




mean and sigma of pull distribution of fsig

	pull mean	pull sigma
f_{sig}= 0.1	-0.03 +-0.05	1.003 +-0.032
f_{sig}= 0.4	-0.047 +-0.045	0.998 +-0.032
f_{sig}= 0.9	-0.0081 +-0.044	0.992 +-0.031



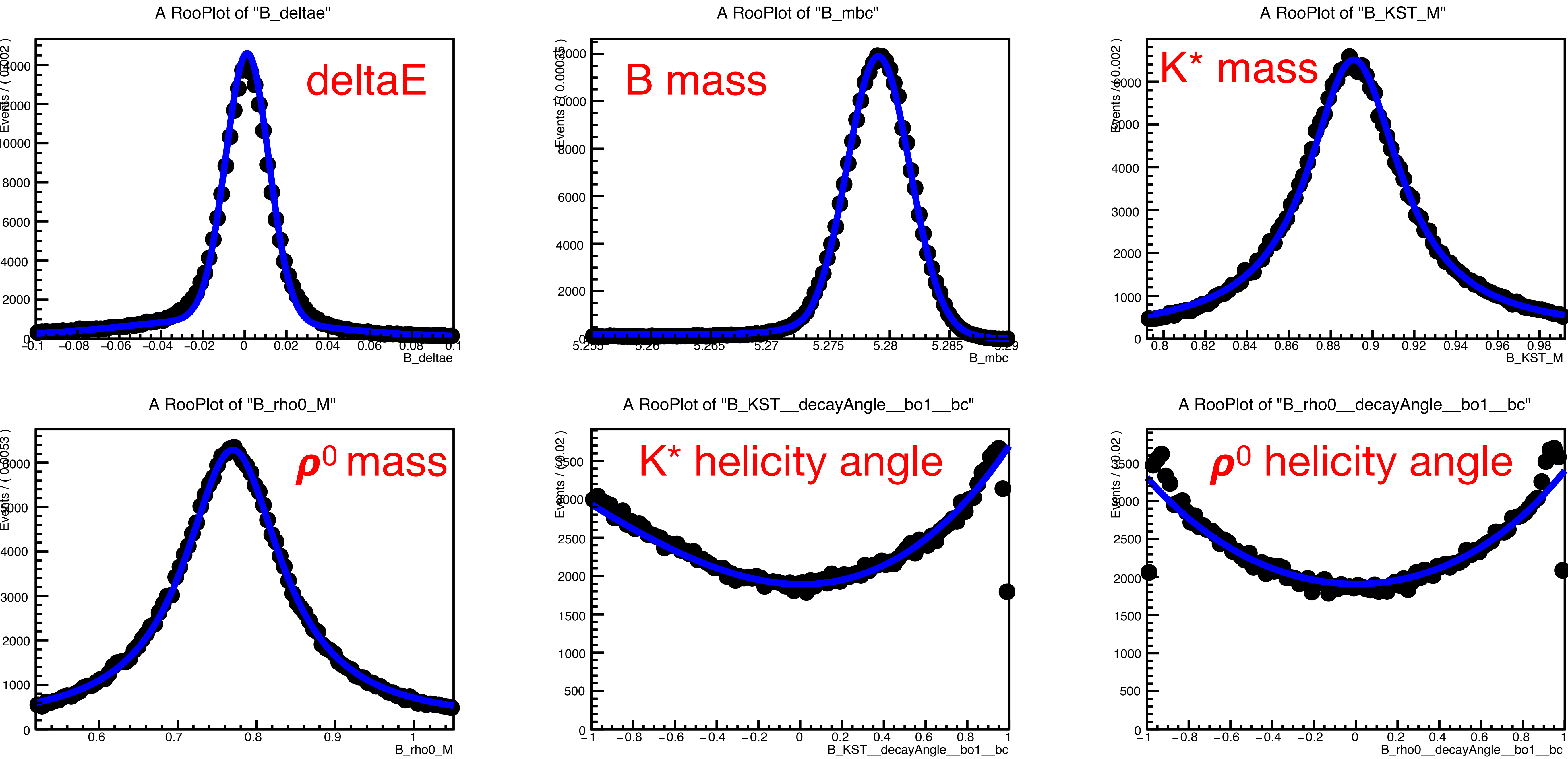


mean and sigma of pull distribution of fln

	pull mean	pull sigma
f_{sig}= 0.1	0.028 +-0.05	1.039 +-0.033
f_{sig}= 0.4	-0.0254 +-0.041	0.912 +-0.029
f_{sig}= 0.9	0.008 +-0.047	1.042 +-0.033

Fit on simulated data: signal + uds background

projections of the 6 observables:



Floating Parameter	InitialValue	FinalValue +/- Error
f_ln	5.0000e-01	4.5642e-01 +/- 1.34e-03
f_sig_6D	5.0000e-01	9.4904e-01 +/- 5.88e-04

$f_{\mathrm{ln\ expected}} = 0.476$

$f_{\mathrm{sig\ expected}} = 0.973$

Not worried right now for biases — plenty of known mismodelings