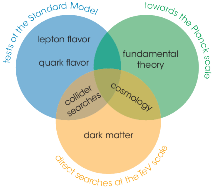
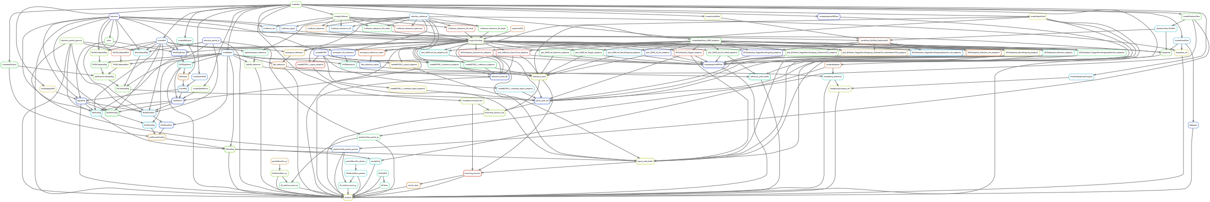
RTG students lecture

III - A Study of the Decay $\Lambda_b^0 \to \Lambda_c^+ \overline{D}^0 K^-$

Marian Stahl

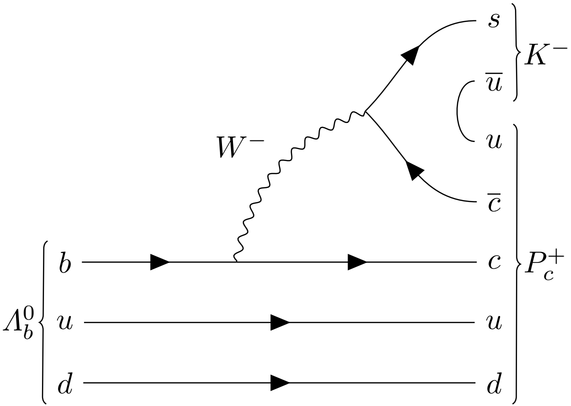
December 12th, 2016





Motivation for $\Lambda_b^0 \to \Lambda_c^+ \overline{D}^0 K^-$

- Can we confirm P_c^+ states seen in $\Lambda_b^0 \to J/\psi\, pK^-$ and $\Lambda_b^0 \to J/\psi\, p\pi^-$ in $\Lambda_b^0 \to \Lambda_c^+ \overline{D}^0 K^-$?



Experimental challenges

- 6 hadrons in final state

  ⇝ larger combinatorial/misID backgrounds, feeddown

- Fewer reconstructed events than $\Lambda_b^0 \to J/\psi\, pK^-$

  Subsequent $\overline{D}^0 \to K^+\pi^-$ and $\Lambda_c^+ \to pK^-\pi^+$ decays, reconstruction efficiency

- Helicity of Λ_c^+ requires $\Lambda_c^+ \to pK^-\pi^+$ amplitudes

  Latest measurement from E791 in 2000 with 950 events [PLB 471, 449]

Phenomenology

- Narrow $P_c^+ \to \Lambda_c^+ \bar{D}^0$ not expected in dynamical diquark-triquark model [PLB 749 454]
- $\Lambda_c^+ \bar{D}^0$ can discriminate between different hypotheses in molecular models [EPJA51 11, 152]
- Using partial reconstruction techniques, it is possible to infer the $\bar{D}^{*0}$ momentum and analyse $\Lambda_c^+ \bar{D}^{*0}$! $\bar{D}^{*0} \to \bar{D}^0 \pi^0/\gamma$

Table 3. Predictions for allowed (✓) and suppressed (×) decays for the different scenarios. The absence of an entry implies that a given channel is not kinematically accessible. The predictions enclosed in brackets are less reliable and can be badly violated if pion-exchange dominates: see the text.

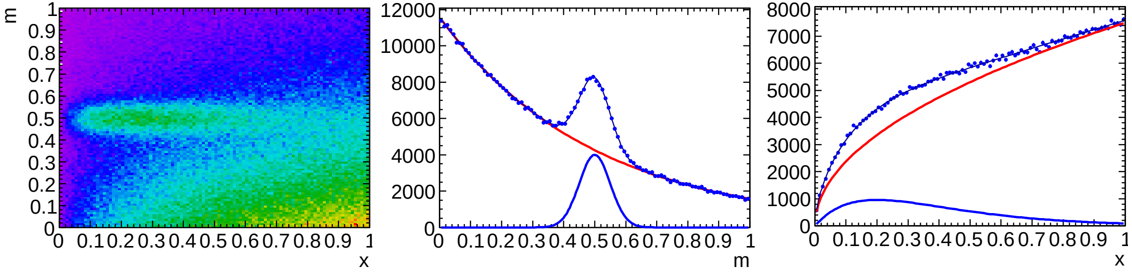
	P_c^*				P_c	
	$\chi_{c1}p$	$\Sigma_c\bar{D}^*$	$\Lambda_c^*\bar{D}$	$J/\psi N^*$	$\Sigma_c^*\bar{D}$	$J/\psi N^*$
$J/\psi N$	✓	✓	✓	✓	✓	✓
$\eta_c N$	×	×	✓	×	×	×
$J/\psi\Delta$	×	✓	×	×	✓	×
$\eta_c\Delta$	×	✓	×	×	✓	×
$\Lambda_c\bar{D}$	✓	[×]	[✓]	×	[×]	×
$\Lambda_c\bar{D}^*$	✓	✓	[✓]	✓	✓	✓
$\Sigma_c\bar{D}$	✓	[×]	✓	×	[×]	×
$\Sigma_c^*\bar{D}$	✓	✓	[×]	✓		
$J/\psi N\pi$	×	✓	×	✓	✓	✓
$\Lambda_c\bar{D}\pi$	×	×	×	×	✓	×
$\Lambda_c\bar{D}^*\pi$	×	✓	×	×		
$\Sigma_c^+\bar{D}^0\pi^0$	×	✓	✓	×		

- Report observation and branching ratio measurement before proceeding with amplitude analysis

$$\frac{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ \overline{D}^{(*)0} K^-)}{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ D_s^-)} = \frac{N(\Lambda_b^0 \to \Lambda_c^+ \overline{D}^{(*)0} K^-)}{\varepsilon(\Lambda_b^0 \to \Lambda_c^+ \overline{D}^{(*)0} K^-)} \cdot \frac{\varepsilon(\Lambda_b^0 \to \Lambda_c^+ D_s^-)}{N(\Lambda_b^0 \to \Lambda_c^+ D_s^-)} \cdot \frac{\mathcal{B}(D_s^- \to K^+ K^- \pi^-)}{\mathcal{B}(\overline{D}^{(*)0} \to K^+ \pi^- (\pi^0/\gamma))}$$

- Select Λ_c^+ and $\overline{D}^0$ signals with dedicated "open charm BDTs" [LHCb-INT-2012-002] [LHCb-ANA-2013-078]. To be updated in the course of this analysis.
  - BDTs trained on $X_b \to X_c \pi$ using kinematic variables and variables which are direct input to conventionally used PID classifiers
  - Their efficiencies are measured on data
- Rely on simulation for trigger-, reconstruction- and pre-selection efficiencies
- Use data driven methods for efficiency of final selection

- Separate (unfold) signal from background distribution in "control" variable x by extracting weights in "discriminating" variable m



- Separate (unfold) signal from background distribution in "control" variable x by extracting weights in "discriminating" variable m
  signal/background only 2 of many possible classes; x,m are vectors in general
- Parametrisation $f(x, m) = N_s \mathcal{P}_s(x, m) + N_b \mathcal{P}_b(x, m)$
  $\mathcal{P}$ denotes a PDF, N a normalisation
- Construct weight function $w(m)$ which projects out signal density

$$N_s \mathcal{P}_s(x) = \int dm\, w(m) f(x, m)$$

- $w(m)$ has to be independent of x $\rightsquigarrow$ $\mathcal{P}_s(x, m)$ and $\mathcal{P}_b(x, m)$ factorise as function of m and x

$\Rightarrow$ $N_s \mathcal{P}_s(x) = \int dm\, w(m) \left[N_s \mathcal{P}_s(x) \mathcal{P}_s(m) + N_b \mathcal{P}_b(x) \mathcal{P}_b(m) \right]$, implying that

$$\int dm\, w(m) \mathcal{P}_s(m) = 1 \quad \text{and} \quad \int dm\, w(m) \mathcal{P}_b(m) = 0$$

- Any $w(m)$ orthogonal to $\mathcal{P}_b(m)$ but not to $\mathcal{P}_s(m)$ possible

- Choose $w(m)$ to give most sensitivity on $\mathcal{P}_s(x)$

$$\sum_{\text{events}} w^2(m) = \text{min} \quad \Rightarrow \quad \int dx\, dm\, w^2(m) f(x,m) = \text{min}$$

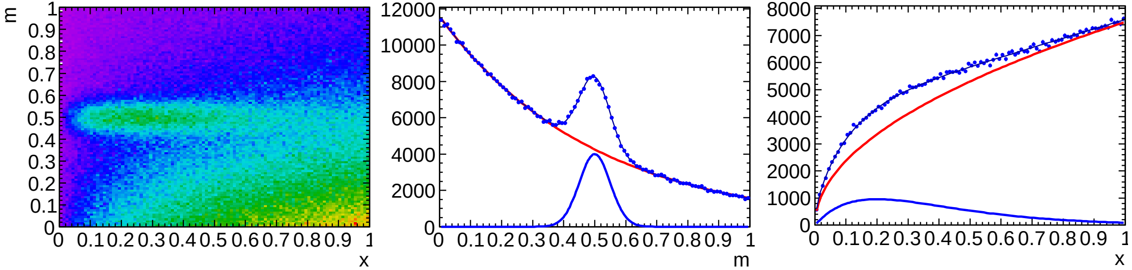
- Lagrange multiplier problem, solved by

$$w(m) = \frac{\langle V_{ss}\rangle \mathcal{P}_s(m) + \langle V_{sb}\rangle \mathcal{P}_b(m)}{N_s \mathcal{P}_s(m) + N_b \mathcal{P}_b(m)} \quad \text{with} \quad \langle V^{-1}_{\text{n}j}\rangle = \int dm \frac{\mathcal{P}_{\text{n}}(m)\mathcal{P}_j(m)}{N_s \mathcal{P}_s(m) + N_b \mathcal{P}_b(m)}$$

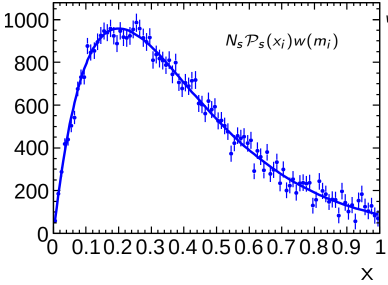
- Note that $V^{-1}_{\text{n}j} = \dfrac{\partial^2(-\mathcal{L})}{\partial N_{\text{n}} \partial N_j}$, where $\mathcal{L}$ is the likelihood

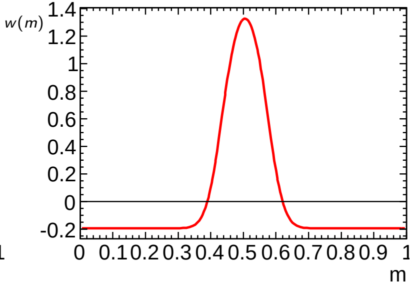
$$\ln \mathcal{L} = \sum_{\text{events}} \ln \left\{ N_s \mathcal{P}_s(m) + N_b \mathcal{P}_b(m) \right\} - N_s - N_b$$

which is minimised in a fit where only N_s and N_b are free parameters



sWeighted events can become **negative!!!**





- Efficiencies are evaluated on control samples (data or MC) in bins of kinematic variables ω ω denotes binning scheme in kinematic phase space
  - Because the efficiency depends on these variables
  - Because the distribution of these variables is expected to change from control to signal sample

- Efficiencies are evaluated on control samples (data or MC) in bins of kinematic variables ω ω denotes binning scheme in kinematic phase space
  - Because the efficiency depends on these variables
  - Because the distribution of these variables is expected to change from control to signal sample
- Efficiencies are "applied" event-by-event

$$\frac{N_s}{\varepsilon_s} \cdot \frac{\varepsilon_r}{N_r} = \frac{N_s}{N_r} \prod_k \sum_{\text{events } i} \frac{\mathcal{P}_s(\omega_{k,i}) w(m_i)}{\varepsilon_k(\omega_{k,i})} \cdot \prod_\ell \sum_{\text{events } j} \frac{\varepsilon_\ell(\omega_{\ell,j})}{\mathcal{P}_r(\omega_{\ell,j}) w(m_j)}$$

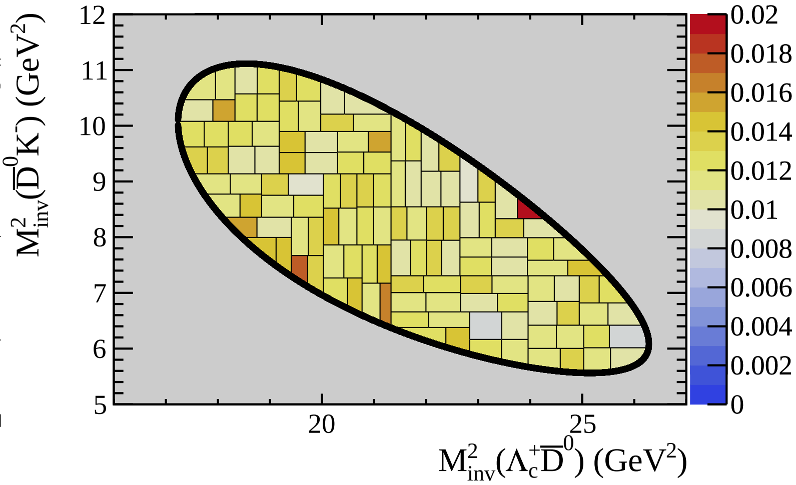
where k, ℓ denotes an efficiency class, s is the $\Lambda_b^0 \to \Lambda_c^+ \overline{D}^{(*)0} K^-$ signal and r the $\Lambda_b^0 \to \Lambda_c^+ D_s^-$ reference-signal N, $\mathcal{P}$, w as in slides before

- N.B.: Product of classes means that factorisation is assumed

Efficiency class	control sample	Phase space signal	reference
Generator level detector acceptance Trigger Reconstruction Stripping pre-selection stream Offline pre-selection	signal MC	Dalitz plot $(M^2_{\text{inv}}(\Lambda_c^+\bar{D}^0), M^2_{\text{inv}}(\bar{D}^0K^-))$	p_T, η
BDT Λ_c^+	$\Lambda_b^0 \to \Lambda_c^+\pi^-$ data	p_T, Flight D istance χ^2	p_T, FDχ^2
BDT $\bar{D}^0$	$B^+ \to \bar{D}^0\pi^+$ data	p_T, FDχ^2	-
BDT D_s^-	$B_s^0 \to D_s^-\pi^+$ data	-	p_T, FDχ^2
PID K^-	$D^{*-} \to [K^-\pi^+]_{D^0}\,\pi^-$ data	p_T, η, `nTracks`	-

- Uncertainties in signal and reference cancel to large extent
- Signal MC needs to be reweighted in p_T, η
- Using adaptive binning to increase sensitivity
- Uncertainties on efficiencies from Wilson confidence interval best average coverage

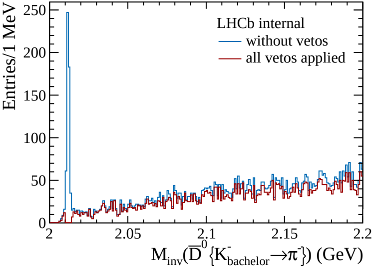
Efficiency class
Generator level detector acc
Trigger
Reconstruction
Stripping pre-selection stream
Offline pre-selection
BDT Λ_c^+
BDT $\bar{D}^0$
BDT D_s^-
PID K^-

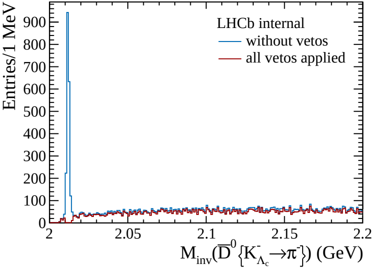


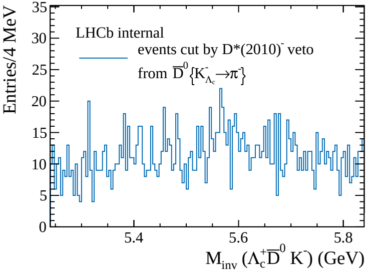
- Uncertainties in sign
- Signal MC needs to
- Using adaptive binning to increase sensitivity
- Uncertainties on efficiencies from Wilson confidence interval best average coverage

- Six hadrons in final state $\rightsquigarrow$ large combinatorics!
- Particle identification is not entirely accurate $\rightsquigarrow$ misidentification
- Decay topology reduces misidentification backgrounds

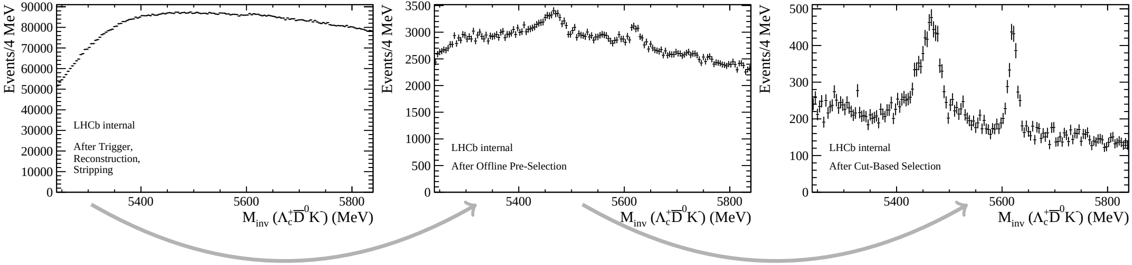
- Six hadrons in final state $\rightsquigarrow$ large combinatorics!
- Particle identification is not entirely accurate $\rightsquigarrow$ misidentification
- Decay topology reduces misidentification backgrounds







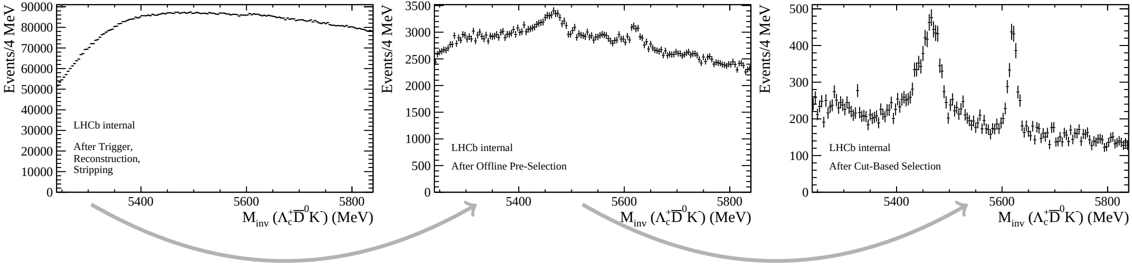
- Two possibilities: include in fit or apply veto
- Combinatorial events in veto region can be signal $\rightsquigarrow$ check signal mass projection
- In this analysis $\phi \rightarrow K^-_{\Lambda_c^+}\{p \rightarrow K^+\}$, $D^0 \rightarrow K^-_{\Lambda_c^+}\{p \rightarrow \pi^+\}$, $D^+/D_s^+ \rightarrow K^-_{\Lambda_c^+}\pi^+_{\Lambda_c^+}\{p \rightarrow \pi^+/K^+\}$, $D^{*-} \rightarrow \bar{D}^0\{K^-_{\Lambda_c^+} \rightarrow \pi^-\}$, $D^{*-} \rightarrow \bar{D}^0\{K^-_{\text{bachelor}} \rightarrow \pi^-\}$, $\phi \rightarrow K^-_{\text{bachelor}}\{p \rightarrow K^+\}$ are vetoed



sanity cuts, loose mass cuts on charm daugters ($\varepsilon \approx 100$ %), $\overline{D}^0/\Lambda_c^+$ BDT cuts ($\varepsilon \approx 97/98$ %), K^- PID cut ($\varepsilon \approx 98$ %)

veto cuts, mass cuts on charm daugters ($\varepsilon \approx 99$ % each), $\overline{D}^0/\Lambda_c^+$ BDT cuts and K^- PID cut (combined $\varepsilon \approx 80$ %)

- Cut-based selection in principle good enough
  - Signal and background clearly separable
  - ⇒ Amplitude analysis can use sWeighted distributions
  - Can we get better sensitivity ($\equiv$ increase statistical power)?



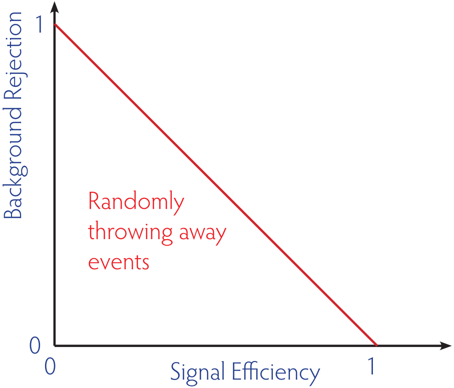
sanity cuts, loose mass cuts on charm daugters ($\varepsilon \approx 100$ %), $\overline{D}^0/\Lambda_c^+$ BDT cuts ($\varepsilon \approx 97/98$ %), K^- PID cut ($\varepsilon \approx 98$ %)

veto cuts, mass cuts on charm daugters ($\varepsilon \approx 99$ % each), $\overline{D}^0/\Lambda_c^+$ BDT cuts and K^- PID cut (combined $\varepsilon \approx 80$ %)

- Cut-based selection in principle good enough
  - Signal and background clearly separable
  - $\Rightarrow$ Amplitude analysis can use sWeighted distributions
  - Can we get better sensitivity ($\equiv$ increase statistical power)?
- Use multivariate Classification! $\leftarrow$ quality quantifyable!
- But optimal working point will depend on many factors $\rightsquigarrow$ systematic studies

ROC Curves

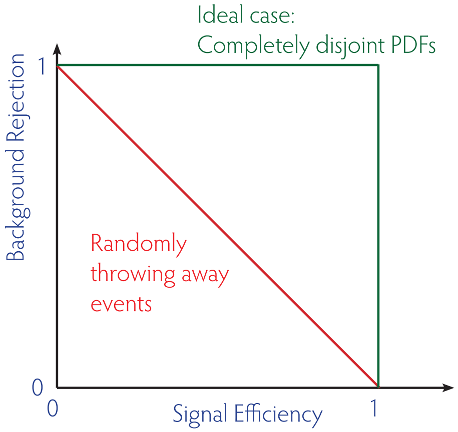
- Receiver Operating Characteristics - originally from signal transmission in electrical engineering



Niklaus Berger – SMIPP – WS 2013 – Slide 6

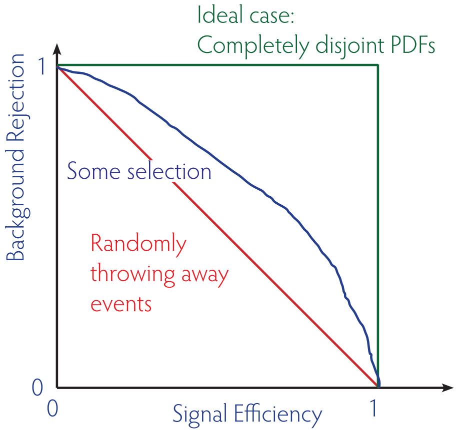
ROC Curves

- Receiver Operating Characteristics - originally from signal transmission in electrical engineering



ROC Curves

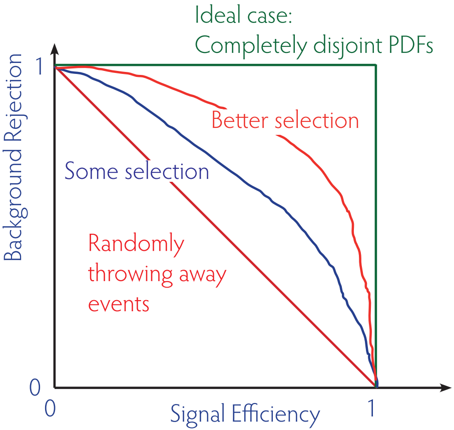
- Receiver Operating Characteristics - originally from signal transmission in electrical engineering



Niklaus Berger – SMIPP – WS 2013 – Slide 8

ROC Curves

- Receiver Operating Characteristics - originally from signal transmission in electrical engineering



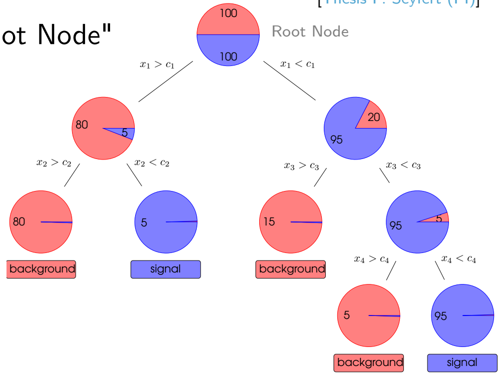
- How far you can go to the upper right is limited by Neyman-Pearson

- ROC-AreaUnderCurve is commonly used measure of quality
- Maximise ROC-AUC by studying different MachineLearning methods and their *hyperparameters*

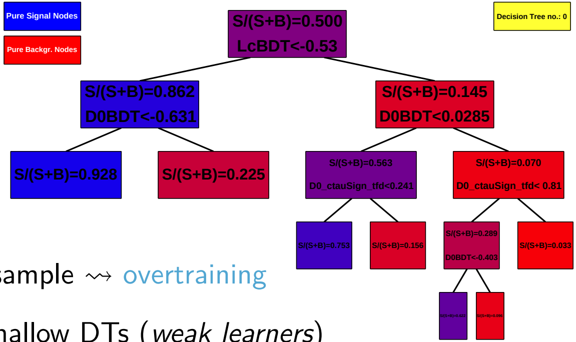
Niklaus Berger – SMIPP – WS 2013 – Slide 10

[Thesis P. Seyfert (PI)]

- Start with training (S,B *known*) sample at "Root Node"
- Split sample using cut c_1 that gives best separation gain usually Gini Index $= p(1-p)$ with purity p
- Continue splitting until reaching
  - Minimal number of events per node
  - Maximum number of nodes
  - Maximum depth
  - Insufficient separation gain
- DTs will be 100 % correct on training sample $\rightsquigarrow$ overtraining

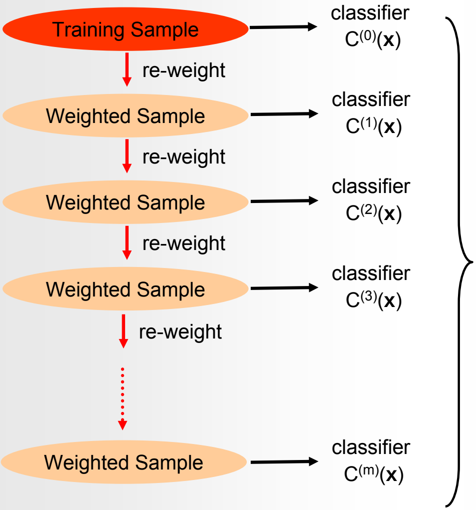


- Start with training (S,B *known*) sample at "Root Node"
- Split sample using cut c_1 that gives best separation gain usually Gini Index $= p(1-p)$ with purity p
- Continue splitting until reaching
  - Minimal number of events per node
  - Maximum number of nodes
  - Maximum depth
  - Insufficient separation gain



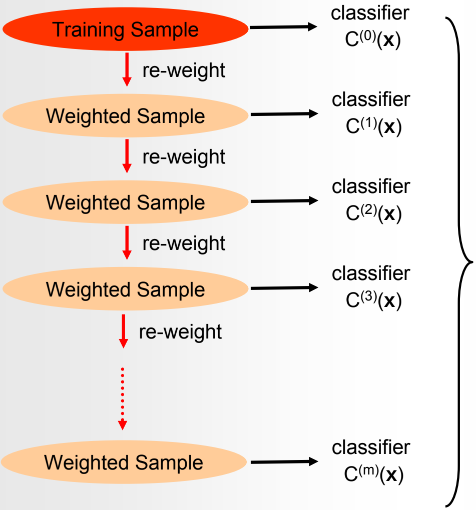
- DTs will be 100 % correct on training sample $\rightsquigarrow$ overtraining
- Can be avoided using combination of shallow DTs (*weak learners*)
- There are algorithms to systematically combine these weak learners: boosting

Boosting



$$y(x) = \sum_{i}^{N_{Classifier}} w_i C^{(i)}(x)$$

Adaptive Boosting (AdaBoost)



- AdaBoost re-weights events misclassified by previous classifier by:

$$\frac{1-f_{err}}{f_{err}} \quad \text{with:}$$

$$f_{err} = \frac{\text{misclassified events}}{\text{all events}}$$

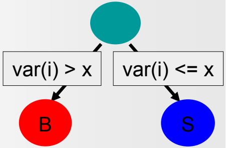
- AdaBoost weights the classifiers also using the error rate of the individual classifier according to:

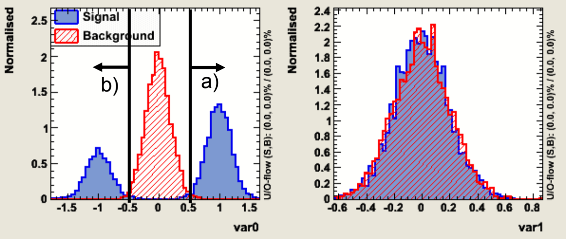
$$y(x) = \sum_{i}^{N_{Classifier}} \log\left(\frac{1-f_{err}^{(i)}}{f_{err}^{(i)}}\right) C^{(i)}(x)$$

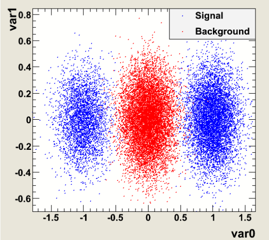
AdaBoost: A simple demonstration

The example: (somewhat artificial…but nice for demonstration) :

- Data file with three “bumps”
- Weak classifier (i.e. one single simple “cut” ↔ decision tree stumps)







Two reasonable cuts: a) Var0 > 0.5 → ε_{signal}=66% ε_{bkg} ≈ 0% misclassified events in total 16.5%

or

b) Var0 < -0.5 → ε_{signal}=33% ε_{bkg} ≈ 0% misclassified events in total 33%

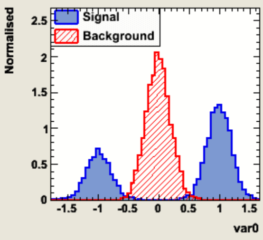
the training of a single decision tree stump will find “cut a)”

AdaBoost: A simple demonstration

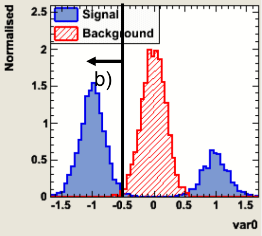
The first "tree", choosing cut a) will give an error fraction: err = 0.165

➔ before building the next "tree": weight wrong classified training events by (1-err/err)) ≈ 5

➔ the next "tree" sees essentially the following data sample:

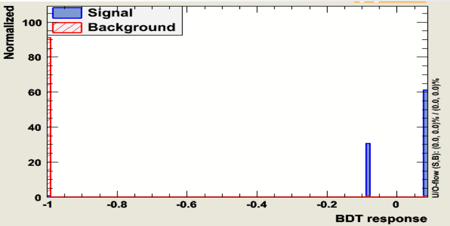


re-weight



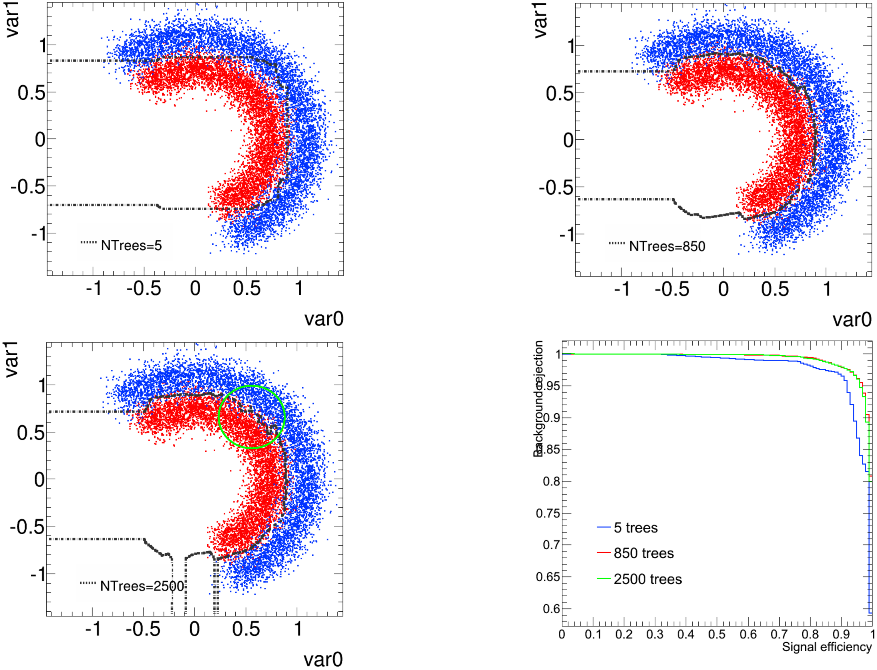
.. and hence will chose: "cut b)": Var0 < -0.5

The combined classifier: Tree1 + Tree2 the (weighted) average of the response to a test event from both trees is able to separate signal from background as good as one would expect from the most powerful classifier



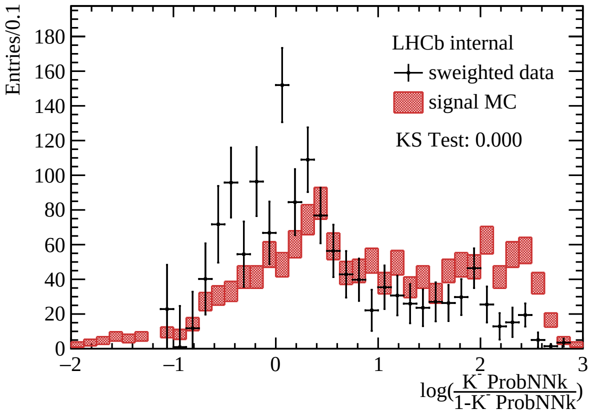
Toy Example (TMVA)

[plots: C. Böser]

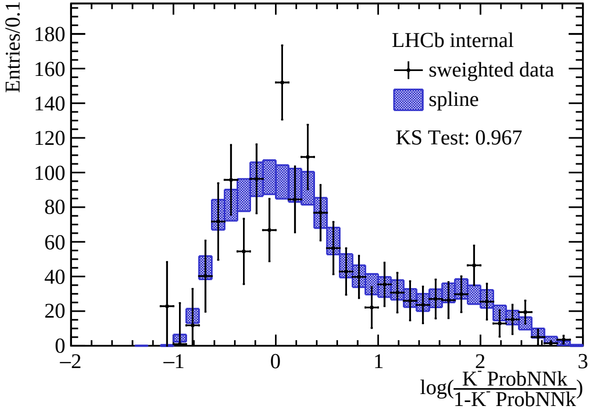


- 25 variables with good separation power have been identified
- How to define signal training sample? background from sidebands in data ✓
  - Train directly on data? Overtraining, ε bias $\rightsquigarrow$ k-fold cross validation
  - Use signal simulation? Need well simulated variables

- 25 variables with good separation power have been identified
- How to define signal training sample? background from sidebands in data ✓
  - ✗ Train directly on data? Overtraining, ε bias, statistics insufficient!
  - ✗ Use signal simulation? Need well simulated variables



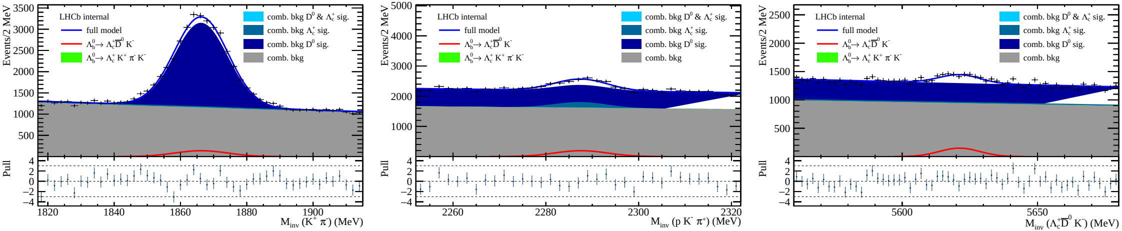
- 25 variables with good separation power have been identified
- How to define signal training sample? background from sidebands in data ✓
  - ✗ Train directly on data? Overtraining, ε bias, statistics insufficient!
  - ✗ Use signal simulation? Need well simulated variables
  - Use splines or KernelDensityEstimators to get smoothed signal PDF. Sample from it.
    Need to do this 25 dimensional to capture correlations ⇝ GAN[arXiv:1406.2661]



- 25 variables with good separation power have been identified
- How to define signal training sample? background from sidebands in data ✓
  - ✗ Train directly on data? Overtraining, ε bias, statistics insufficient!
  - ✗ Use signal simulation? Need well simulated variables
  - Use splines or KernelDensityEstimators to get smoothed signal PDF. Sample from it. Need to do this 25 dimensional to capture correlations $\rightsquigarrow$ GAN[arXiv:1406.2661]
  - Reweight signal MC to splines? tradeoff agreement $\leftrightarrow$ effective MC statistics
  - Study correlations and use multiple stages of training?
  - Study ongoing. Best solution is probably a mix
  - Currently: 25 $\times$ 1D splines

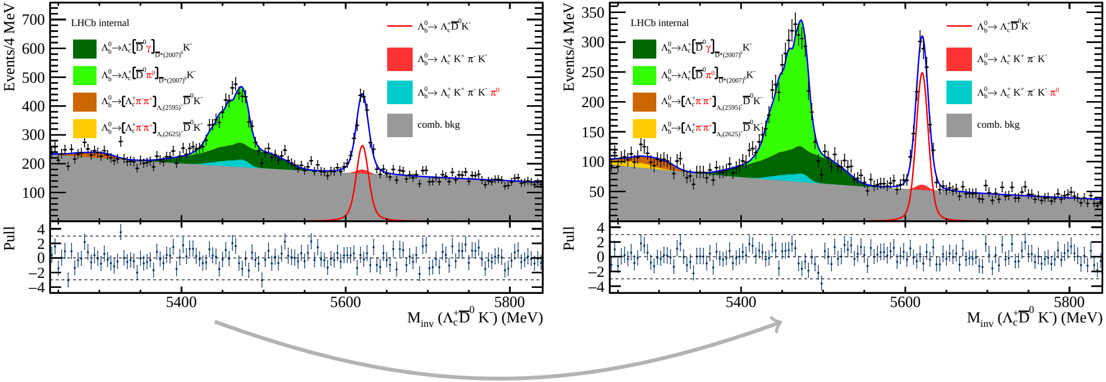
- Wait... where did the sWeights come from?

- Wait... where did the sWeights come from?
- Use 3D fit to data after offline pre-selection stage



- Separate pure combinatorial and single charm backgrounds
  - Will study MVA with multiple background sources in the future
- Separate double charm from single and no-charm signal
  - Only $\Lambda_b^0 \to \Lambda_c^+ K^+ \pi^- K^-$ contributes at tree-level
- 3D Fits repeated after final selection to extract $\Lambda_b^0 \to \Lambda_c^+ K^+ \pi^- K^- (\pi^0)$ yield. This yield will be fixed in the final 1D fit

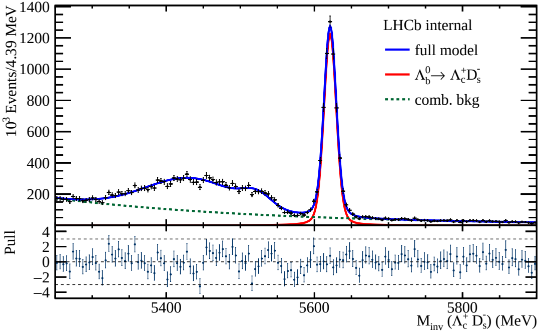
1D mass fits

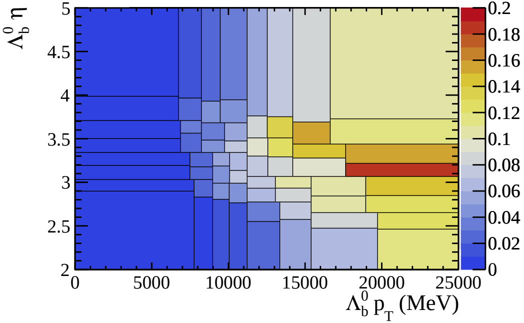


cut-based $\rightarrow$ MVA selection $\varepsilon_{\text{relative}} \approx 90\%$

- Shapes of partially reconstructed Λ_b^0 decays from KDE PDFs of simulated data
- Further partially reconstructed decays absorbed in background or negligible

$$\frac{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ \bar{D}^{(*)0} K^-)}{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ D_s^-)} = \frac{N(\Lambda_b^0 \to \Lambda_c^+ \bar{D}^{(*)0} K^-)}{\varepsilon(\Lambda_b^0 \to \Lambda_c^+ \bar{D}^{(*)0} K^-)} \cdot \frac{\varepsilon(\Lambda_b^0 \to \Lambda_c^+ D_s^-)}{N(\Lambda_b^0 \to \Lambda_c^+ D_s^-)} \cdot \frac{\mathcal{B}(D_s^- \to K^+ K^- \pi^-)}{\mathcal{B}(\bar{D}^{(*)0} \to K^+ \pi^- (\pi^0/\gamma))}$$





- Reference channel workflow similar to signal channel
- Use reference channel to validate open charm BDTs
- $\dfrac{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ \bar{D}^{(*)0} K^-)}{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ D_s^-)} \approx 0.13$

- Exotica are excellent laboratory to study the poorly understood dynamics and binding mechanisms of QCD
- Absence of exotica in the light quark sector
  - Inconclusive searches for decades
  - Two waves of hints for exotic KN resonances
- Large number of Tetraquark candidates observed with $c\overline{c}$ or $b\overline{b}$ content
- First $uudc\overline{c}$ pentaquark candidates observed in 2015 at LHCb
- Lots of theoretical predictions waiting to be tested
- One of them: search for $uudc\overline{c}$ pentaquarks in $\Lambda_b^0 \to \Lambda_c^+ \overline{D}^{(*)0} K^-$
  - Here, $\Lambda_b^0 \to \Lambda_c^+ \overline{D}^{(*)0} K^-$ has been observed for the first time
  - Challenges: Efficiencies as function of kinematics, optimisation of signal selection
  - Amplitude analysis proven to be feasible, but helicity of Λ_c^+s needed

Backup slides start here

EventType	Decay	feeds into Λ_b^0	expected $f_{\rm fd}$
15196200	$\Lambda_b^0 \to \Lambda_c^+ \left[\bar{D}^0 \gamma\right]_{\bar{D}^*(2007)^0} K^-$	✗	1.4
15196400	$\Lambda_b^0 \to \Lambda_c^+ \left[\bar{D}^0 \pi^0\right]_{\bar{D}^*(2007)^0} K^-$	✗	2.3
15196201	$\Lambda_b^0 \to \Lambda_c^+ \left[\left[\bar{D}^0 \gamma\right]_{\bar{D}^*(2007)^0} K^-\right]_{D_{s1}(2536)^-}$	✗	0.06
15196406	$\Lambda_b^0 \to \Lambda_c^+ \left[\left[\bar{D}^0 \pi^0\right]_{\bar{D}^*(2007)^0} K^-\right]_{D_{s1}(2536)^-}$	✗	0.10
15196401	$\Lambda_b^0 \to \Lambda_c^+ \bar{D}^0 \left[K^- \pi^0\right]_{K^*(892)^-}$	✗	0.25?
15196402	$\Lambda_b^0 \to \left[\Lambda_c^+ \pi^0\right]_{\Sigma_c(2455)^+} \bar{D}^0 K^-$	✗	0.05
15196403	$\Lambda_b^0 \to \left[\Lambda_c^+ \pi^0\right]_{\Sigma_c(2520)^+} \bar{D}^0 K^-$	✗	$\lesssim 0.02$
15196404	$\Lambda_b^0 \to \left[\Lambda_c^+ K^- \pi^0\right]_{\Xi_c(2980)^0} \bar{D}^0$	✗	negl.
15196405	$\Lambda_b^0 \to \left[\Lambda_c^+ K^- \pi^0\right]_{\Xi_c(3080)^0} \bar{D}^0$	✗	negl.
15198002	$\Lambda_b^0 \to \left[\Lambda_c^+ \pi^- \pi^+\right]_{\Lambda_c(2595)^+} \bar{D}^0 K^-$	✗	negl.
15198003	$\Lambda_b^0 \to \left[\Lambda_c^+ \pi^- \pi^+\right]_{\Lambda_c(2625)^+} \bar{D}^0 K^-$	✗	negl.
16196440	$\Xi_b^0 \to \left[\Lambda_c^+ \pi^0\right]_{\Sigma_c(2455)^+} \bar{D}^0 K^-$	✓	negl.
16196441	$\Xi_b^0 \to \left[\Lambda_c^+ \pi^0\right]_{\Sigma_c(2520)^+} \bar{D}^0 K^-$	[✓]	negl.
16197030	$\Xi_b^- \to \left[\Lambda_c^+ \pi^-\right]_{\Sigma_c(2455)^0} \bar{D}^0 K^-$	✓	negl.
16197031	$\Xi_b^- \to \left[\Lambda_c^+ \pi^-\right]_{\Sigma_c(2520)^0} \bar{D}^0 K^-$	[✓]	negl.
16196442	$\Xi_b^0 \to \left[\left[\pi^0 p\right]_{\Sigma^+} K^- \pi^+ K^-\right]_{\Omega_c^0} \bar{D}^0$	[✓]	negl.
16196443	$\Xi_b^0 \to \Lambda_c^+ \bar{D}^0 \left[K^- \pi^0\right]_{K^*(892)^-}$	[✓]	negl.
16196444	$\Xi_b^0 \to \left[p K^- \pi^+ \pi^0\right]_{\Xi_c^+} \bar{D}^0 K^-$	[✓]	negl.

Potential Λ_b^0 and Ξ_b decays which cross feed into the $\Lambda_c^+ \bar{D}^0 K^-$ invariant mass distribution. The third column indicates if the decay feeds into the Λ_b^0 signal (✓), or not (✗). A [✓] indicates that the tails of the distribution feed into the Λ_b^0 signal. Particles labelled in red are not reconstructed, whereas blue labelled particles are required to be within [2270,2305] MeV of their invariant mass to mimic a Λ_c^+. The last column gives the expected feeddown fraction w.r.t. the signal yield.

Cross feed into $\Lambda_c^+ \bar{D}^0 K^-$

EventType	Decay	feeds into Λ_b^0	expected f_{fd}
15196200	$\Lambda_b^0 \to \Lambda_c^+ \left[\bar{D}^0\gamma\right]_{\bar{D}^*(2007)^0} K^-$	✗	1.4
15196400	$\Lambda_b^0 \to \Lambda_c^+ \left[\bar{D}^0\pi^0\right]_{\bar{D}^*(2007)^0} K^-$	✗	2.3
15196201	$\Lambda_b^0 \to \Lambda_c^+ \left[\left[\bar{D}^0\gamma\right]_{\bar{D}^*(2007)^0} K^-\right]_{D_{s1}(2536)^-}$	✗	0.06
15196406	$\Lambda_b^0 \to \Lambda_c^+ \left[\left[\bar{D}^0\pi^0\right]_{\bar{D}^*(2007)^0} K^-\right]_{D_{s1}(2536)^-}$	✗	0.10
15196401	$\Lambda_b^0 \to \Lambda_c^+ \bar{D}^0 \left[K^-\pi^0\right]_{K^*(892)^-}$	✗	0.25?
15196402	$\Lambda_b^0 \to [\Lambda_c^+\pi^0]_{\Sigma_c(2455)^+} \bar{D}^0 K^-$	✗	0.05
15196403	$\Lambda_b^0 \to [\Lambda_c^+\pi^0]_{\Sigma_c(2520)^+} \bar{D}^0 K^-$	✗	$\lesssim 0.02$
15196404	$\Lambda_b^0 \to [\Lambda_c^+ K^-\pi^0]_{\Xi_c(2980)^0} \bar{D}^0$	✗	negl.
15196405	$\Lambda_b^0 \to [\Lambda_c^+ K^-\pi^0]_{\Xi_c(3080)^0} \bar{D}^0$	✗	negl.
15198002	$\Lambda_b^0 \to [\Lambda_c^+\pi^-\pi^+]_{\Lambda_c(2595)^+} \bar{D}^0 K^-$	✗	negl.
15198003	$\Lambda_b^0 \to [\Lambda_c^+\pi^-\pi^+]_{\Lambda_c(2625)^+} \bar{D}^0 K^-$	✗	negl.
16196440	$\Xi_b^0 \to [\Lambda_c^+\pi^0]_{\Sigma_c(2455)^+} \bar{D}^0 K^-$	✓	negl.
16196441	$\Xi_b^0 \to [\Lambda_c^+\pi^0]_{\Sigma_c(2520)^+} \bar{D}^0 K^-$	[✓]	negl.
16197030	$\Xi_b^- \to [\Lambda_c^+\pi^-]_{\Sigma_c(2455)^0} \bar{D}^0 K^-$	✓	negl.
16197031	$\Xi_b^- \to [\Lambda_c^+\pi^-]_{\Sigma_c(2520)^0} \bar{D}^0 K^-$	[✓]	negl.
16196442	$\Xi_b^0 \to [[\pi^0 p]_{\Sigma^+} K^-\pi^+K^-]_{\Omega_c^0} \bar{D}^0$	[✓]	negl.
16196443	$\Xi_b^0 \to \Lambda_c^+ \bar{D}^0 \left[K^-\pi^0\right]_{K^*(892)^-}$	[✓]	negl.
16196444	$\Xi_b^0 \to [pK^-\pi^+\pi^0]_{\Xi_c^+} \bar{D}^0 K^-$	[✓]	negl.

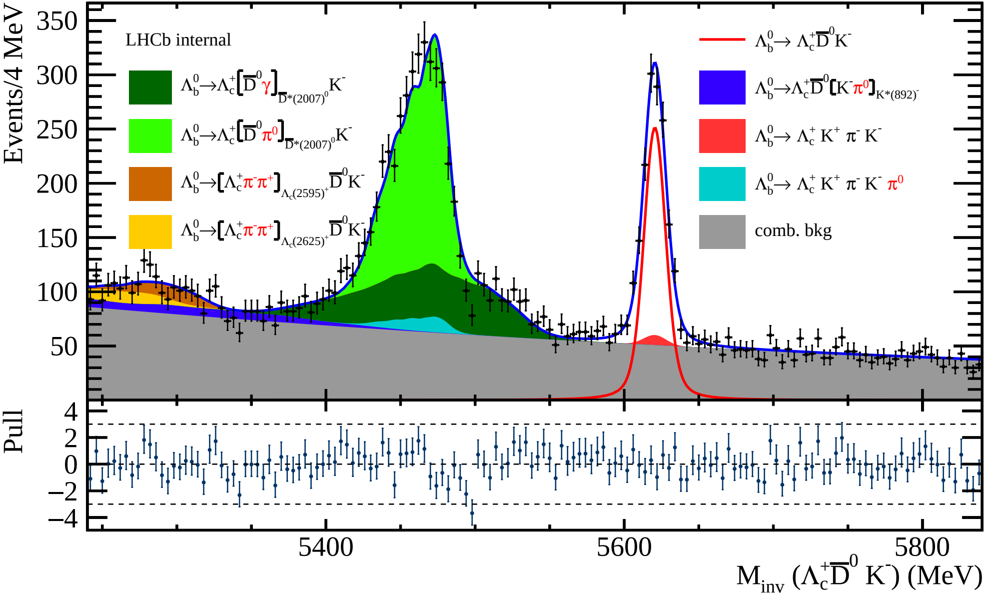
Potential Λ_b^0 and Ξ_b decays which cross feed into the $\Lambda_c^+ \bar{D}^0 K^-$ invariant mass distribution. The third column indicates if the decay feeds into the Λ_b^0 signal (✓), or not (✗). A [✓] indicates that the tails of the distribution feed into the Λ_b^0 signal. Particles labelled in red are not reconstructed, whereas blue labelled particles are required to be within [2270,2305] MeV of their invariant mass to mimic a Λ_c^+. The last column gives the expected feeddown fraction w.r.t. the signal yield.

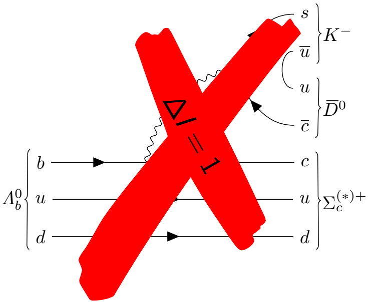
Cross feed into $\Lambda_c^+ \bar{D}^0 K^-$

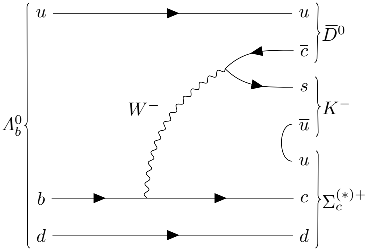
EventType	Decay	feeds into Λ_b^0	expected f_{fd}
15196200	$\Lambda_b^0 \to \Lambda_c^+ \left[\bar{D}^0 \gamma\right]_{\bar{D}^*(2007)^0} K^-$	✗	1.4
15196400	$\Lambda_b^0 \to \Lambda_c^+ \left[\bar{D}^0 \pi^0\right]_{\bar{D}^*(2007)^0} K^-$	✗	2.3
15196201	$\Lambda_b^0 \to \Lambda_c^+ \left[\left[\bar{D}^0 \gamma\right]_{\bar{D}^*(2007)^0} K^-\right]_{D_{s1}(2536)^-}$	✗	0.06
15196406	$\Lambda_b^0 \to \Lambda_c^+ \left[\left[\bar{D}^0 \pi^0\right]_{\bar{D}^*(2007)^0} K^-\right]_{D_{s1}(2536)^-}$	✗	0.10
15196401	$\Lambda_b^0 \to \Lambda_c^+ \bar{D}^0 \left[K^- \pi^0\right]_{K^*(892)^-}$	✗	0.25?
15196402	$\Lambda_b^0 \to [\Lambda_c^+ \pi^0]_{\Sigma_c(2455)^+} \bar{D}^0 K^-$	✗	0.05
15196403	$\Lambda_b^0 \to [\Lambda_c^+ \pi^0]_{\Sigma_c(2520)^+} \bar{D}^0 K^-$	✗	$\lesssim 0.02$
15196404	$\Lambda_b^0 \to [\Lambda_c^+ K^- \pi^0]_{\Xi_c(2980)^0} \bar{D}^0$	✗	negl.
15196405	$\Lambda_b^0 \to [\Lambda_c^+ K^- \pi^0]_{\Xi_c(3080)^0} \bar{D}^0$	✗	negl.
15198002	$\Lambda_b^0 \to [\Lambda_c^+ \pi^- \pi^+]_{\Lambda_c(2595)^+} \bar{D}^0 K^-$	✗	negl.
15198003	$\Lambda_b^0 \to [\Lambda_c^+ \pi^- \pi^+]_{\Lambda_c(2625)^+} \bar{D}^0 K^-$	✗	negl.
16196440	$\Xi_b^0 \to [\Lambda_c^+ \pi^0]_{\Sigma_c(2455)^+} \bar{D}^0 K^-$	✓	negl.
16196441	$\Xi_b^0 \to [\Lambda_c^+ \pi^0]_{\Sigma_c(2520)^+} \bar{D}^0 K^-$	[✓]	negl.
16197030	$\Xi_b^- \to [\Lambda_c^+ \pi^-]_{\Sigma_c(2455)^0} \bar{D}^0 K^-$	✓	negl.
16197031	$\Xi_b^- \to [\Lambda_c^+ \pi^-]_{\Sigma_c(2520)^0} \bar{D}^0 K^-$	[✓]	negl.
16196442	$\Xi_b^0 \to [[\pi^0 p]_{\Sigma^+} K^- \pi^+ K^-]_{\Omega_c^0} \bar{D}^0$	[✓]	negl.
16196443	$\Xi_b^0 \to \Lambda_c^+ \bar{D}^0 \left[K^- \pi^0\right]_{K^*(892)^-}$	[✓]	negl.
16196444	$\Xi_b^0 \to [p K^- \pi^+ \pi^0]_{\Xi_c^+} \bar{D}^0 K^-$	[✓]	negl.

Potential Λ_b^0 and Ξ_b decays which cross feed into the $\Lambda_c^+ \bar{D}^0 K^-$ invariant mass distribution. The third column indicates if the decay feeds into the Λ_b^0 signal (✓), or not (✗). A [✓] indicates that the tails of the distribution feed into the Λ_b^0 signal. Particles labelled in red are not reconstructed, whereas blue labelled particles are required to be within [2270,2305] MeV of their invariant mass to mimic a Λ_c^+. The last column gives the expected feeddown fraction w.r.t. the signal yield.

Adding $\Lambda_b^0 \to \Lambda_c^+ \overline{D}^0 [K^- \pi^0]_{K^*(892)^-}$







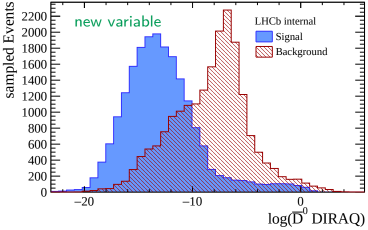
- Estimate feeddown fraction from CDF PRD 79 032001

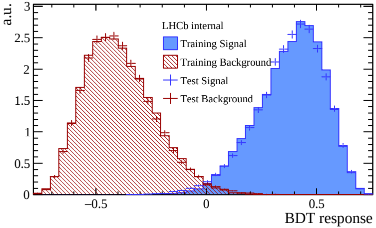
$$\frac{\mathcal{B}(\Lambda_b^0 \to \Sigma_c(2455)^+ \pi^0 \mu^- \overline{\nu}_\mu)}{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ \mu^- \overline{\nu}_\mu)} = 0.054$$

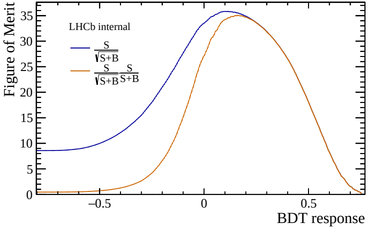
- Correct for additional π^0 assuming additional pions are Poissonian and using

$$\frac{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ \mu^- \overline{\nu}_\mu)}{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ \pi^+ \pi^- \mu^- \overline{\nu}_\mu)} = 1.1 = \frac{\mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ \mu^- \overline{\nu}_\mu)}{0.6 \cdot \mathcal{B}(\Lambda_b^0 \to \Lambda_c^+ 2\pi \mu^- \overline{\nu}_\mu)}$$

- Additional $d\overline{d}$ pair in semileptonic diagram $\rightsquigarrow$ factor 0.5
- Estimate $< 2\%$ of $\Lambda_b^0 \to \Lambda_c^+ \left[\overline{D}^0 \pi^0\right]_{\bar{D}^*(2007)^0} K^-$







DIRAQ $= (1 - \text{DIRA}) \cdot \Delta\text{DIRA}$ requires vertex-momentum covariances

$\overline{D}^0$ BDT	Λ_c^+ BDT	$\arctan(\overline{D}^0\ c\tau_{\text{DTF}}$ sign.)	$\log(\frac{K^-\ \texttt{ProbNNk}}{1-K^-\ \texttt{ProbNNk}})$	$\arctan(\Lambda_c^+\ c\tau_{\text{DTF}}$ sign.)
$\log(K^-\ \Delta p_{\text{TDTF}})$	$\log(\overline{D}^0$ DIRAQ)	$\log(K^-\ p_{\text{TDTF}})$	$\log(\Lambda_b^0$ BPV IP$\chi^2_{\text{DTF}})$	$\overline{D}^0\ \alpha_{\texttt{AP}}$
$\log(\Lambda_c^+$ DIRAQ)	$\log(\Lambda_c^+$ BPV IP$\chi^2_{\text{DTF}})$	$\log(K^-$ BPV IP$\chi^2_{\text{DTF}})$	$\log(1 - \Lambda_b^0$ DIRA)	$\log(\overline{D}^0\ \Delta p_{\text{TDTF}})$
$\log(\frac{K^-\ \texttt{ProbNNghost}}{1-K^-\ \texttt{ProbNNghost}})$	$\log(\overline{D}^0\ p_{\text{TDTF}})$	$\log(\Lambda_b^0\ \Delta p_{\text{TDTF}})$	$\log(\Lambda_b^0\ \texttt{BPV}\tau_{\text{DTF}})$	$\log(\overline{D}^0$ BPV IP$\chi^2_{\text{DTF}})$
$\log(\Lambda_b^0\ \Delta M_{\text{DTF}})$	$\log(\Lambda_c^+\ p_{\text{TDTF}})$	$\log(\Lambda_b^0\ \texttt{BPVPDS}_{\text{DTF}})$	$\log(\Lambda_b^0\ p_{\text{TDTF}})$	$\log(\Lambda_c^+\ \Delta p_{\text{TDTF}})$