Search for Rare b to Open-Charm Two-Body Decays of Baryons at LHCb

Promotionsverteidigung

Nis Meinert

Rostock, October 9th 2020

INSTITUT FÜR PHYSIK

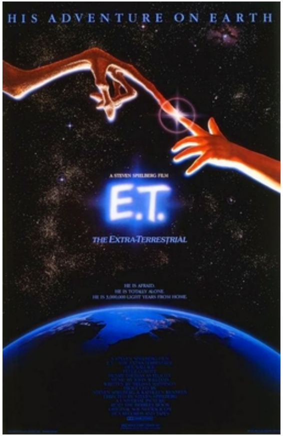
Universität Rostock

Bundesministerium für Bildung und Forschung

Introduction

What if? Consider communication with an alien…

- → Before flying by we want to know: is alien world made out of **anti-matter***?
- → Can we propose an experiment to find out?



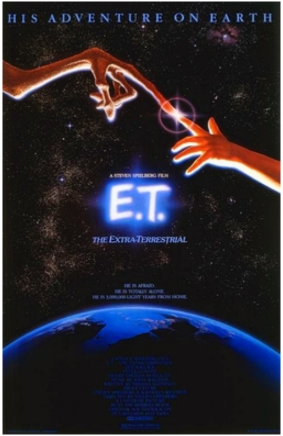
(Image: John Alvin, (c) 1982 Universal Studios)

**cf.* fermions: if ψ obeys $(i\not{\partial} - m)\psi = 0$, then $(\mathrm{CP}\ \psi)$ is also a solution (**C**harge and **P**arity inverted)

Introduction

What if? Consider communication with an alien…

- → Before flying by we want to know: is alien world made out of **anti-matter***?
- → Can we propose an experiment to find out?
- → Where not to look for: QED, QCD (neither **C**harge (C), nor **P**arity (P), nor **T**ime reversal (T) symmetry is broken here
- → What about weak interaction?
  - → P violation measured with Wu-Experiment!
  - → …but asymmetry cancels exactly when using anti-matter
  - → **Is definition of anti-matter ambiguous?**



(Image: John Alvin, (c) 1982 Universal Studios)

* *cf.* fermions: if ψ obeys $(\mathrm{i}\partial\!\!\!/ - m)\psi = 0$, then $(\mathrm{CP}\,\psi)$ is also a solution (**C**harge and **P**arity inverted)

On the other hand…

- Baryon asymmetry of the Universe:

$$\mathcal{O}(\eta_{\text{cosm}}) = \mathcal{O}\left(\frac{n_B - n_{\overline{B}}}{n_\gamma}\right) \sim 10^{-10} \neq 0,$$

  with $n_{B,\overline{B},\gamma}$ = number of baryons, anti-baryons and photons

- Universe is filled with matter (rather than anti-matter*)
- Standard Model of Cosmology: Big Bang produces $n_B = n_{\overline{B}}$ (**CP invariant**)
- Evolution: QED and QCD are **CP invariant**

* *i.e.*, most likely alien world is made out of matter as well

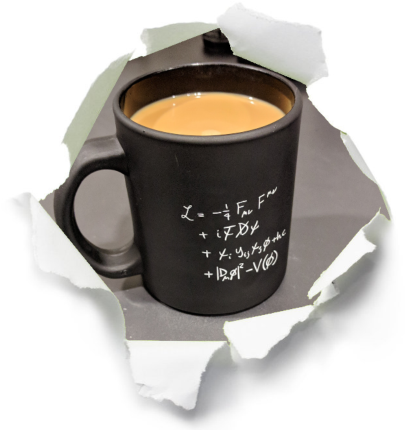
Introduction

Can we explain $\mathcal{O}(\eta_{\text{cosm}})$?

- → **No!** We fail to describe one of the most obvious property of our Universe!

Can we (at least) describe some CP violating processes?

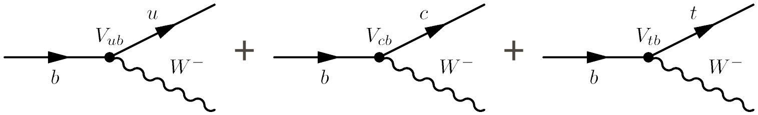
- → Yes! The Standard Model of Particle Physics contains a CP violating sector (weak interaction)
- → However, it does not predict the right order of magnitude of η_{cosm} !



Introduction

- Diagonalizing Yukawa matrices* leads to **quark mass eigenstates** q, but skews **quark flavor eigenstates** q'

$$\mathcal{J}^{\mu}_{cc} \supset (\overline{u'},\, \overline{c'},\, \overline{t'})\, \gamma^{\mu} \frac{1-\gamma^5}{2} \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} \quad \longrightarrow \quad (\overline{u},\, \overline{c},\, \overline{t})\, \gamma^{\mu} \frac{1-\gamma^5}{2} V_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$



- This introduces the non-trivial transformation matrix V_{CKM} for quarks (CKM matrix)
- V_{CKM} is a **unitary** 3×3 matrix (CP violation *in* complex phases)

* *cf.* Higgs mechanism

State-of-the-art

What we know

- → We understand CP violation (CPV) on quark level in the framework of the Standard Model (SM)
- → CPV is predicted and confirmed with high confidence for **mesons*** at particle accelerators

What we don't know

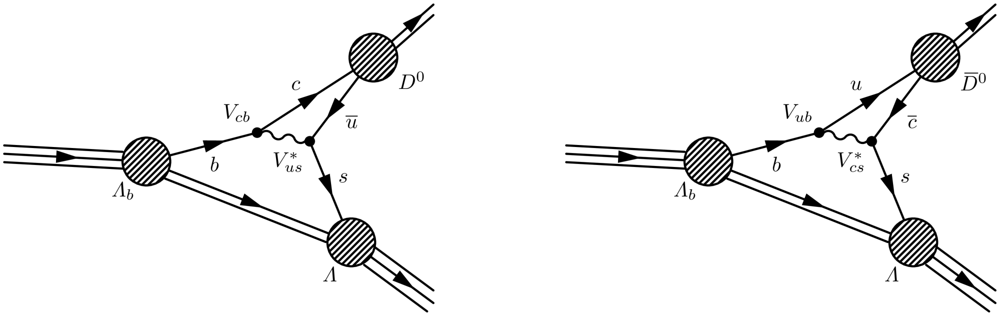
- → Absence of anti-**baryons*** is a large scale CPV which we cannot explain
- → (strong CPV, neutrino CPV, Sakharov conditions, ...)

Idea of this analysis

- → Establish (yet unknown) **baryon** decay which can be used (in the future) to measure CPV parameter of the SM

*meson / anti-meson: $(q\overline{q})$, baryon: (qqq), anti-baryon: $(\overline{qqq})$

The decay $\Lambda_b \to D^0 \Lambda$



- Perfectly suited for LHCb
- CKM angle: $\gamma + \mathcal{O}(\lambda^4)$ (without CPV in D^0 system)
- CPV in $\Lambda_b \to D^0/\overline{D}^0 \Lambda$ and
  - $D^0/\overline{D}^0 \to K^+\pi^-$ (CPV but Cabibbo suppressed, **future prospect**)
  - $D^0/\overline{D}^0 \to K^+K^- / \pi^+\pi^-$ (CPV but Cabibbo suppressed, **future prospect**)
  - $D^0 \to K^-\pi^+$ (no CPV but Cabibbo favored, **this analysis**)

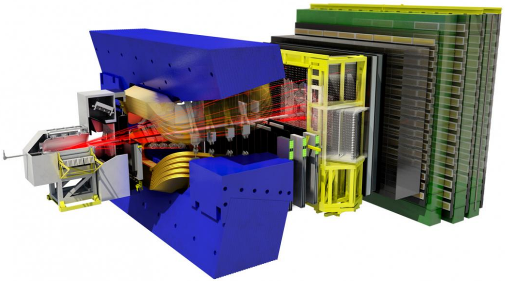
Large Hadron Collider

(Symbolic image, picture taken at CERN site)

- World's largest and highest-energy particle accelerator
- Built by CERN
- Collides proton beams with $\sqrt{s} = 13$ TeV
- b baryon factory (Λ_b is lightest b baryon)

LHCb Experiment

"LHCb is an experiment set up to explore what happened after the Big Bang that allowed matter to survive and build the Universe we inhabit today"



- One of the four major experiments at the LHC
- Dataset: $6\,\text{fb}^{-1}$ at $\sqrt{s} = 13\,\text{TeV}$

Multipurpose detector, *e.g.*:

- Excellent vertex resolution
- Powerful particle identification

Outline

Why / What?

- Pave the way to measure CPV in decays of baryons at colliders
- Prose: search for the decays

$$\Lambda_b/\Xi_b^0 \to [K^-\pi^+]_{D^0}\,[p\pi^-]_{\Lambda}$$

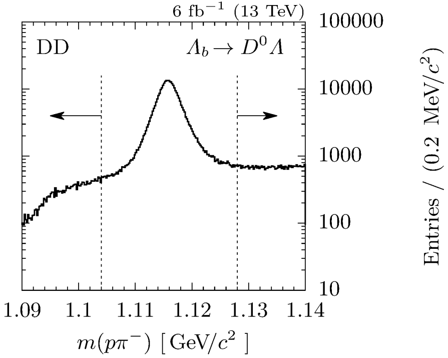
How?

- (Data calibration and pre-processing†)
- Classification of data as *signal* and *background* using MVA*
- (Study various physical backgrounds†)
- Count signal events with fit and estimate stat. significances*
- Estimate branching ratios $\mathcal{B}(\Lambda_b \to D^0\Lambda)/\mathcal{B}(\Lambda_b \to D^0 p\pi^-)^\dagger$ and $\mathcal{B}(\Xi_b^0 \to D^0\Lambda)/\mathcal{B}(\Lambda_b \to D^0\Lambda)^*$

†not discussed here, *partially discussed here

How to find genuine decays? —Example $\Lambda \to p\pi^-$

- Create Λ candidates: $\{\Lambda\} := \{p\} \otimes \{\pi^-\}$
- Use inv. mass $m(\Lambda)^\star \stackrel{?}{=} m(p\pi^-)$ as proxy
- Separate components:
  1. Combinatorial background
     (random track combinations)
  2. physical background
     (*e.g.*, $K_S \to \pi^+\pi^-$)
  3. genuine $\Lambda \to p\pi^-$ decays



(*PDG: $m(\Lambda) \approx 1.117\,\mathrm{GeV}/c^2$)

Technical detail: only (weak) Cabibbo suppressed decays possible for Λ

- *long* lifetime
- some Λ decay outside (inside) of first detector element: refer to as **DD** (**LL**)
- different detector response, separate analysis necessary

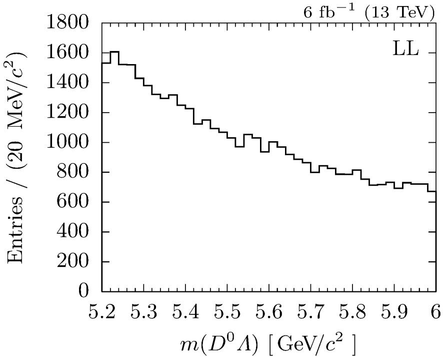
Do the same for $\Lambda_b/\Xi_b^0 \to D^0 \Lambda$?

i.e., $\{\Lambda_b/\Xi_b^0\} := \underbrace{\{K^-\} \otimes \{\pi^+\}}_{\{D^0\}} \otimes \underbrace{\{p\} \otimes \{\pi^-\}}_{\{\Lambda\}}$

Searching for $\Lambda_b/\Xi_b^0 \to D^0\Lambda$

No signal visible!

- Background (at this stage) mainly combinatorial
- Need to increase
  - purity of $\{D^0\}$
  - purity of $\{\Lambda\}$
  - filter $\{D^0\} \otimes \{\Lambda\}$ for $\Lambda_b/\Xi_b^0 \to D^0\Lambda$ decays by constraining physical properties
- (Carefully analyse physical backgrounds†)

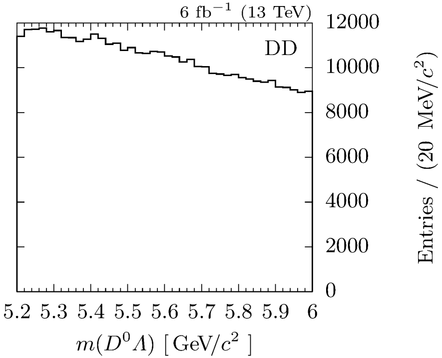


(PDG: $m(\Lambda_b) \approx 5.620\,\mathrm{GeV}/c^2$,
PDG: $m(\Xi_b^0) \approx 5.792\,\mathrm{GeV}/c^2$)

Searching for $\Lambda_b/\Xi_b^0 \to D^0\Lambda$

No signal visible!

- Background (at this stage) mainly combinatorial
- Need to increase
  - purity of $\{D^0\}$
  - purity of $\{\Lambda\}$
  - filter $\{D^0\} \otimes \{\Lambda\}$ for $\Lambda_b/\Xi_b^0 \to D^0\Lambda$ decays by constraining physical properties
- (Carefully analyse physical backgrounds†)



(PDG: $m(\Lambda_b) \approx 5.620\,\text{GeV}/c^2$,
PDG: $m(\Xi_b^0) \approx 5.792\,\text{GeV}/c^2$)

Searching for $\Lambda_b/\Xi_b^0 \to D^0\Lambda$

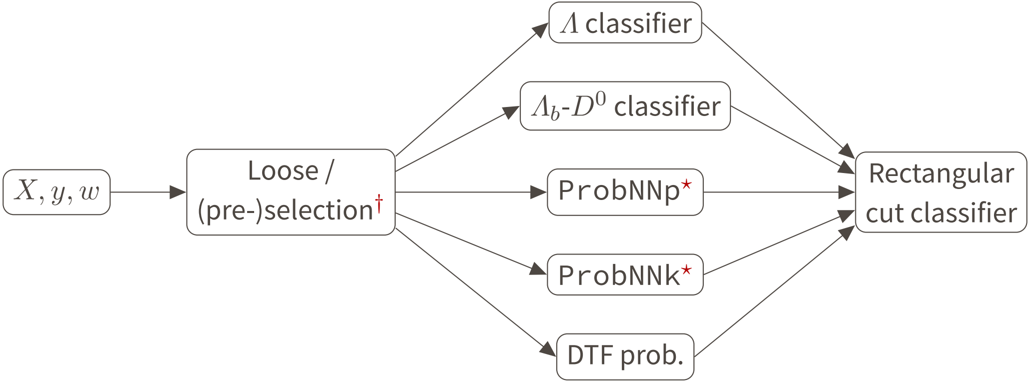
Use ML to rescue! Classification problem

- Separate ***signal*** (genuine $\Lambda_b/\Xi_b^0 \to D^0\Lambda$)
- ...from (combinatorial) ***background***
- Labels?
  - use MC simulated decays for class *signal*†
  - use recorded data from sidebands for class *background*

(Image: R. Diepenheim on thenounproject.com)

† calibration needed (not discussed here)

MVA methodoloy

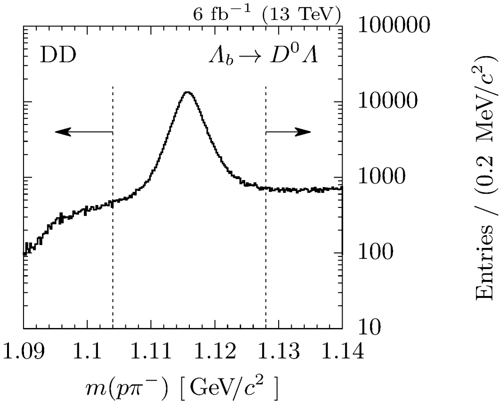


Training: classify m data $X \in \mathbb{R}^{m \times n}$ with n features using labels $y \in \{\text{sig., bkg.}\}^m$ and calibration factors† $w \in \mathbb{R}^m$ with $\mathcal{O}(m) = 10k, n = 18$

⋆ pre-trained shallow NN

Example: Λ classifier

- 5 features of $\{\Lambda\}$:
  1. transverse momentum
  2. angle between momentum and connecting line between pp-IA point (PV) and Λ decay vertex
  3. χ^2 improvement of PV with Λ track
  4. fit prob. of decay vertex
  5. significance of flight distance
- Separate classifiers for LL and DD
- Optimize hyper-parameters of 5 classifiers using grid search w.r.t. ROC-AUC* value

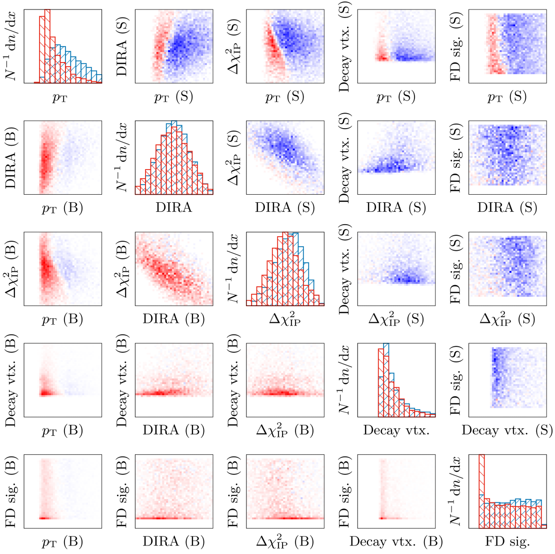


* Prob. that clf. will rank randomly chosen *sig.* instance higher than randomly chosen *bkg.* one

Example: Λ classifier

Λ classifier

- 5 features of $\{\Lambda\}$
- Separate classifiers for LL and **DD**
- Optimize hyper-parameters of 5 classifiers using grid search w.r.t. ROC-AUC* value

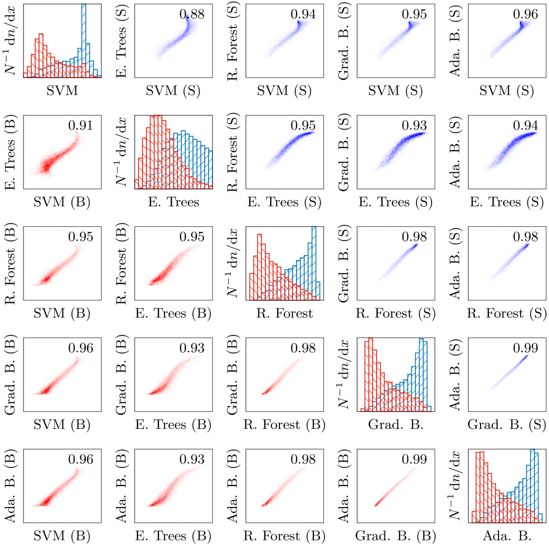


(track type **DD**)

Example: Λ classifier

Λ classifier

- 5 features of $\{\Lambda\}$
- Separate classifiers for LL and **DD**
- Optimize hyper-parameters of 5 classifiers using grid search w.r.t. ROC-AUC* value

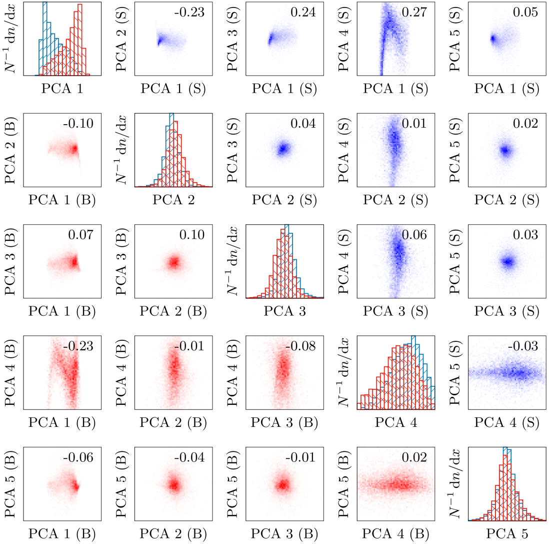


(track type **DD**)

Example: Λ classifier

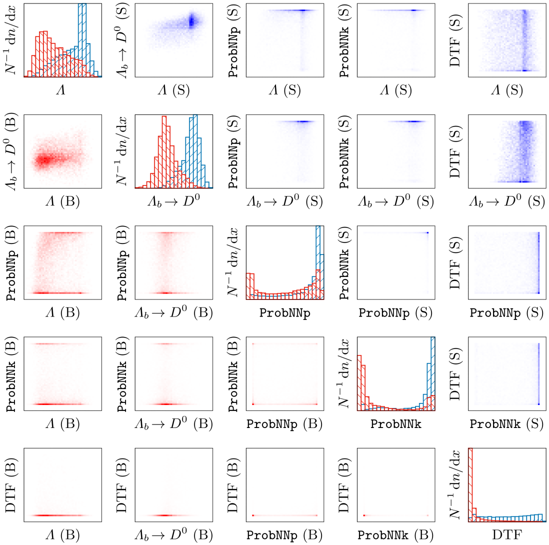
Λ classifier

- 5 features of $\{\Lambda\}$
- Separate classifiers for LL and **DD**
- Optimize hyper-parameters of 5 classifiers using grid search w.r.t. ROC-AUC* value



(track type **DD**)

Fusion of high level classifiers



(track type DD)

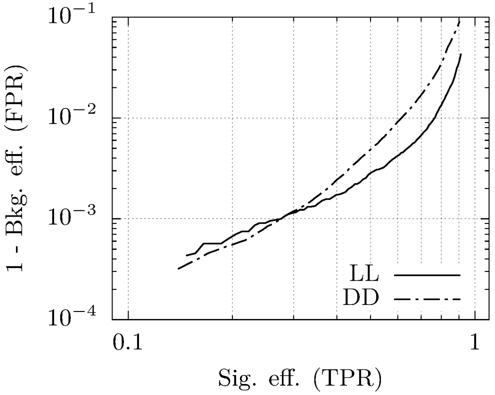
Fusion of high level classifiers

- → Optimize rectangular cut classifier (to avoid MC fidelity issues†)
- → …using classifier outputs as features
  1. Λ classifier
  2. Λ_b-D^0 classifier
  3. `ProbNNp`
  4. `ProbNNk`
  5. Decay tree fit
- → (Compare rect. cut clf. eff. against decision tree†)

Classifier fusion

Optimize rect. cut clf.

- Minimize FPR* at given TPR**
- Numerically noisy: use simulated annealing for finding minimum
- Set thresholds at TPR $\approx 50\,\%$ (LL) and TPR $\approx 30\,\%$ (DD)
- (Refine *optimal* TPR with in a second, tight max. step only using Λ/Λ_b-D^0 classifiers†)

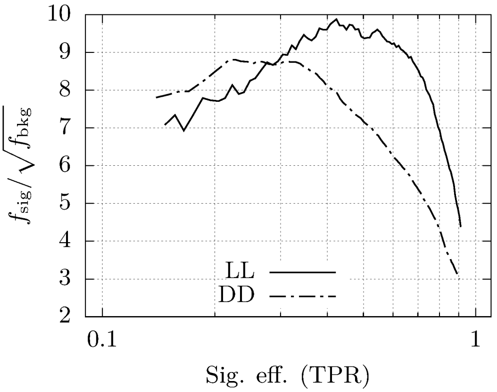


*FPR: false positive rate, *i.e.*, genuine bkg. predicted as sig.
**TPR: true positive rate, *i.e.*, genuine sig. predicted as sig.

Classifier fusion

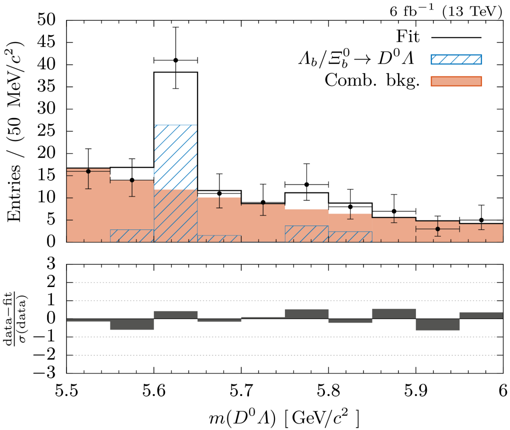
Optimize rect. cut clf.

- Minimize FPR* at given TPR**
- Numerically noisy: use simulated annealing for finding minimum
- Set thresholds at TPR $\approx 50\,\%$ (LL) and TPR $\approx 30\,\%$ (DD)
- (Refine *optimal* TPR with in a second, tight max. step only using Λ/Λ_b-D^0 classifiers†)



*FPR: false positive rate, *i.e.*, genuine bkg. predicted as sig.
**TPR: true positive rate, *i.e.*, genuine sig. predicted as sig.

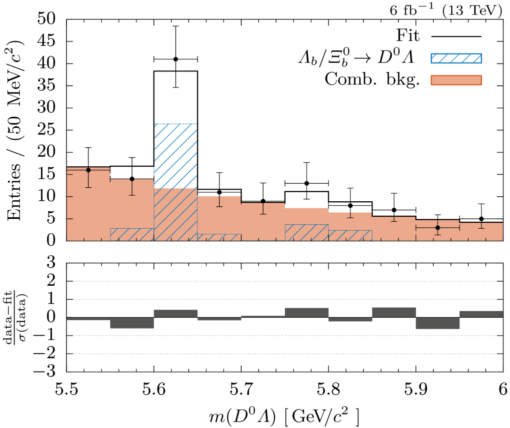
Fit invariant mass



Result: $n(\Lambda_b) = 31(7)$ and $n(\Xi_b^0) = 6(4)$
(only stat. uncertainty given)

- Fit remaining 260 signal candidates (rec. data)
- Simultaneously fit rec. and MC sim. data of both track types (6 samples)
  - 23 parameters
  - most of them shared between at least two samples
- Vary fit model to check for systematic error
  - change parametrization
  - change inv. mass range
  - … (8 variations in total†)

Fit invariant mass



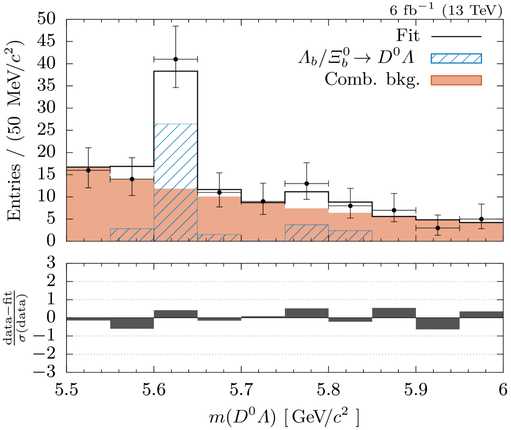
Result: $n(\Lambda_b) = 31(7)$ and $n(\Xi_b^0) = 6(4)$
(only stat. uncertainty given)

Sanity checks

→ (Bias of estimated yields†)

→ (Validity of (stat.) error estimates†)

Fit invariant mass



Result: $n(\Lambda_b) = 31(7)$ and $n(\Xi_b^0) = 6(4)$
(only stat. uncertainty given)

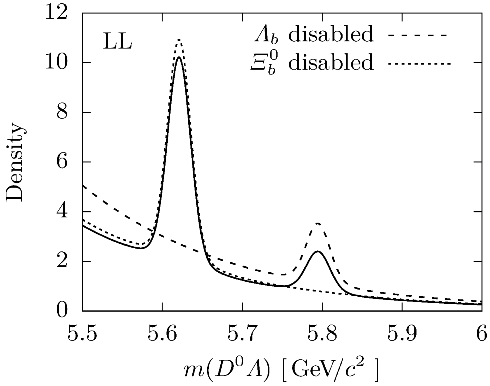
Stat. significance of yields

- Fix all parameters from MC sim.
- Run fit twice
  - $\mathcal{L}$ for signal comp. **enabled**
  - $\mathcal{L}$ for signal comp. **disabled**
- Stat. sign. given by $\sqrt{-2\Delta\ln\mathcal{L}}$

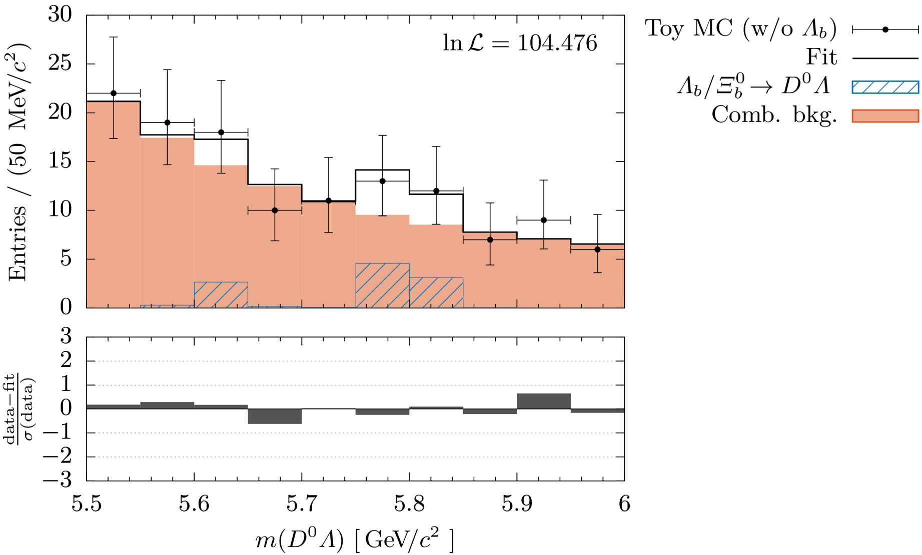
Statistical significance of yields

Wilks theorem

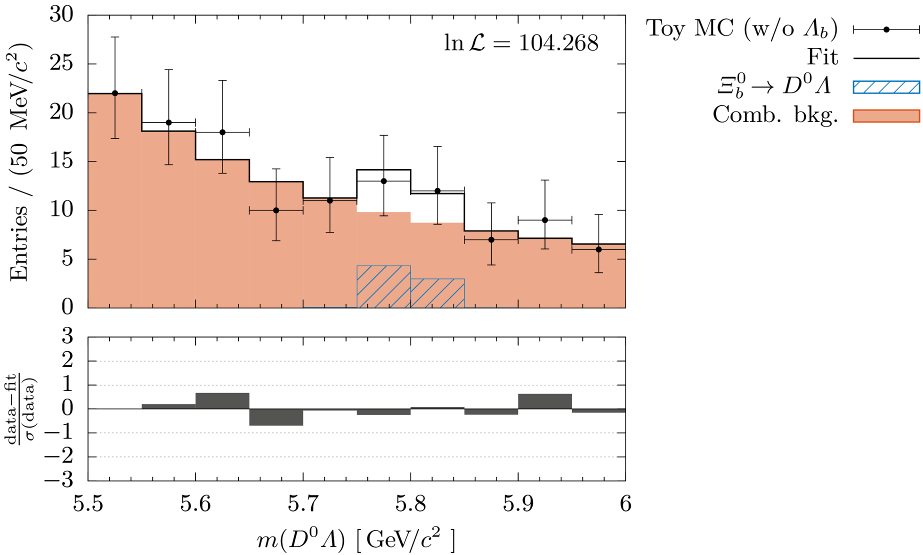
- Here: if sample size tends to infinity $2\Delta\ln\mathcal{L}$ tends to a χ^2-distribution with one DoF
- Hence: stat. significance of rej. null hypothesis *no signal* is $\sqrt{-2\Delta\ln\mathcal{L}}$
- Test assumption with pseudo experiment (*Toy MC*)
  - sample Toy MC from fitted PDF w/o signal component
  - run fit twice (w/ and w/o signal component)



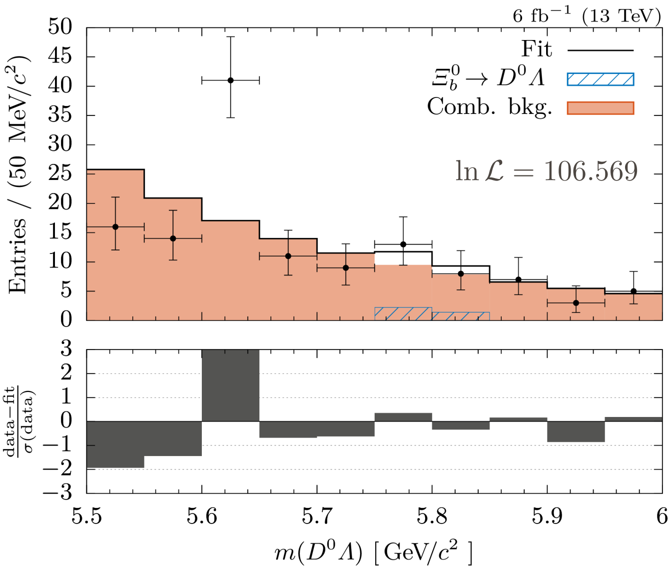
Statistical significance of yields



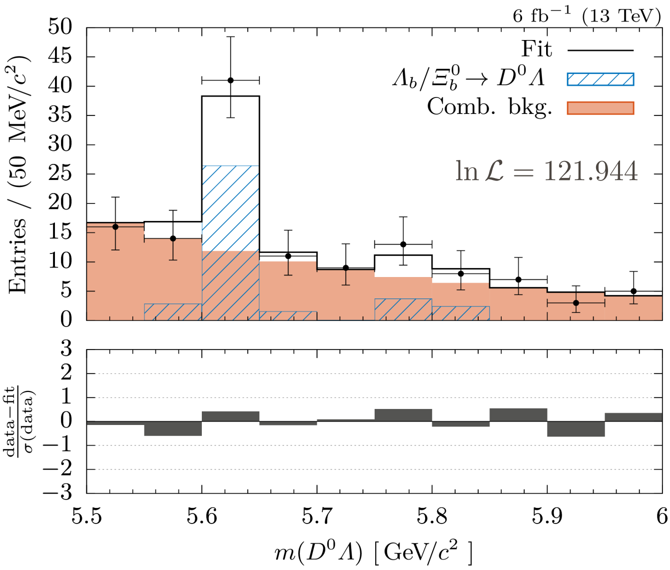
Statistical significance of yields



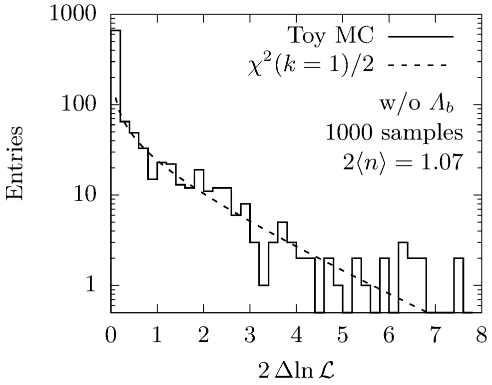
Statistical significance of yields



Statistical significance of yields



Statistical significance of yields



Wilks theorem

- → Distribution of $\Delta\ln\mathcal{L}$ indeed follows a χ^2-distribution with 1 DoF
- → Wilks theorem is applicable!

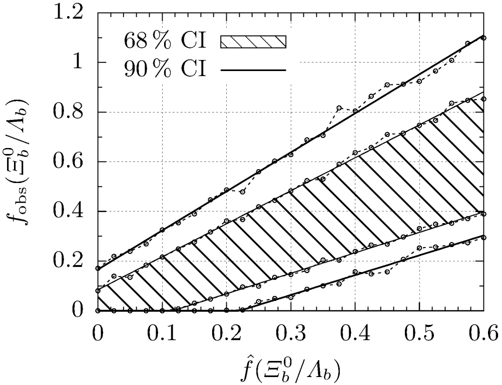
Estimated **yield significances**: **5.5** for $\Lambda_b \to D^0\Lambda$ and **1.8** for $\Xi_b^0 \to D^0\Lambda$

Bayesian CIs

- (Different likelihood scans assuming flat prior[†])

Frequentist CIs

- Toy MC: draw random events from fitted PDF and vary Ξ_b^0/Λ_b ratio (400 fits for each $\hat{f}(\Xi_b^0/\Lambda_b)$)
- CI from interval in $\hat{f}(\Xi_b^0/\Lambda_b)$ at $f_{\text{obs}}(\Xi_b^0/\Lambda_b) = 0.20(15)$
- Construct CIs in $f_{\text{obs}}(\Xi_b^0/\Lambda_b)$ by
  - **central method**
  - shortest method



(central in f_{obs})

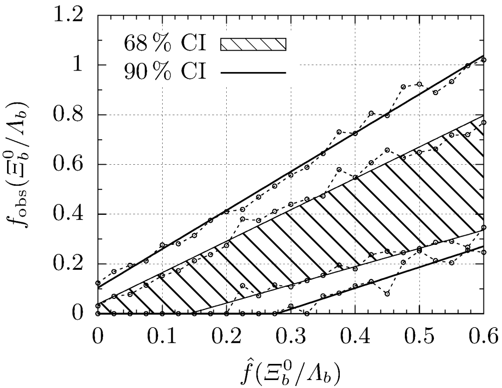
Branching Ratios —Example $f(\Xi_b^0/\Lambda_b)$

Bayesian CIs

- (Different likelihood scans assuming flat prior[†])

Frequentist CIs

- Toy MC: draw random events from fitted PDF and vary Ξ_b^0/Λ_b ratio (400 fits for each $\hat{f}(\Xi_b^0/\Lambda_b)$)
- CI from interval in $\hat{f}(\Xi_b^0/\Lambda_b)$ at $f_{\text{obs}}(\Xi_b^0/\Lambda_b) = 0.20(15)$
- Construct CIs in $f_{\text{obs}}(\Xi_b^0/\Lambda_b)$ by
  - central method
  - **shortest method**



(shortest in f_{obs})

Results / Outlook

Search for new decays

- $\Lambda_b \to D^0 \Lambda$ with a stat. sign. of 5.5 σ (discovery)
- $\Xi_b^0 \to D^0 \Lambda$ with a stat. sign. of 1.8 σ

Measurement of branching fractions / ratios

- $\mathcal{B}(\Lambda_b \to D^0 \Lambda) = (9.9 \pm 2.3_{\text{stat}} \pm 1.6_{\text{sys}} \pm 1.1_{\text{ext}}) \times 10^{-6}$
- $\frac{f_{\Xi_b^0}}{f_{\Lambda_b}} \times \frac{\mathcal{B}(\Xi_b^0 \to D^0 \Lambda)}{\mathcal{B}(\Lambda_b \to D^0 \Lambda)} < 0.5 \quad (\text{CL} = 95\,\%)$

Outlook

- Collect $\sim$ 20x more data (in reach within next LHC runs)
- Benchmark SM by measuring CPV parameter γ using suppressed D^0 decays

Results / Outlook

Search for new decays

- $\Lambda_b \to D^0 \Lambda$ with a stat. sign. of 5.5 σ (discovery)
- $\Xi_b^0 \to D^0 \Lambda$ with a stat. sign. of 1.8 σ

Measurement of branching fractions / ratios

- $\mathcal{B}(\Lambda_b \to D^0 \Lambda) = (9.9 \pm 2.3_{\text{stat}} \pm 1.6_{\text{sys}} \pm 1.1_{\text{ext}}) \times 10^{-6}$
- $\frac{f_{\Xi_b^0}}{f_{\Lambda_b}} \times \frac{\mathcal{B}(\Xi_b^0 \to D^0 \Lambda)}{\mathcal{B}(\Lambda_b \to D^0 \Lambda)} < 0.5 \quad (\text{CL} = 95\,\%)$

Outlook

- Collect $\sim$ 20x more data (in reach within next LHC runs)
- Benchmark SM by measuring CPV parameter γ using suppressed D^0 decays

INSTITUT FÜR PHYSIK

Universität Rostock

Bundesministerium für Bildung und Forschung

Thank you for your attention

BACKUP

The CKM matrix

$$V_{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - \mathrm{i}\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - \mathrm{i}\eta) & -A\lambda^2 & 1 \end{pmatrix}}_{\text{cf. Wolfenstein: Phys. Rev. Lett. 51 (1945), 21}} + \mathcal{O}(\lambda^3)$$

$$\lambda \approx .23, A \approx .81, \rho \approx .14, \eta \approx .35$$

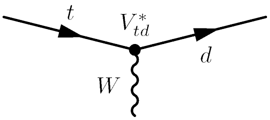
- One non-trivial complex phase, encoded in matrix elements
  - V_{ub} and V_{td} (up to $\mathcal{O}(\lambda^2)$)
  - V_{cd}, V_{cs} and V_{ts} (up to $\mathcal{O}(\lambda^6)$)
- CP violation if and only if $\eta \neq 0$

CP violation (?)

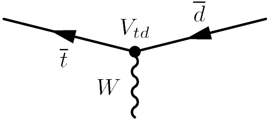
- If $\eta \neq 0$ some V_{ij} carry a complex phase (*weak* phase)
  - amplitude: $\mathcal{A}(t \to dW^{+}) \sim V_{td}^{*}$
  - amplitude: $\bar{\mathcal{A}}(\bar{t} \to \bar{d}W^{-}) \sim V_{td}$
  - CPV since $\mathcal{A} \neq \bar{\mathcal{A}}$ (?)

… not quite

- Amplitude $\mathcal{A}$ is not observable …
- … but branching fraction $\mathcal{B} \sim |\mathcal{A}|^2$ is
  - CPV needs at least two interfering decay modes with …
  - … one **CP odd** and one **CP even** phase



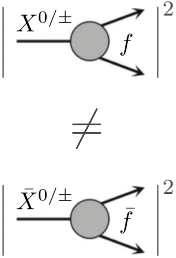
$t \to dW^{+}$



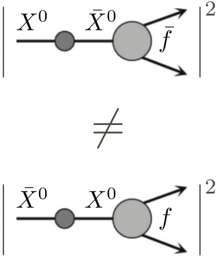
$\bar{t} \to dW^{-}$

Types of CPV

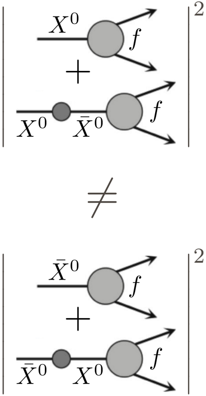
(1) Direct CPV



(2) CPV in mixing



(3) CPV in interference of mixing and decay



(Images: *CP Violation*, I. I. Bigi and A. I. Sanda)

- CP odd: from CKM matrix
- CP even:
  - (1): strong phase difference between both amplitudes (*e.g.*, *tree* and *penguin*)
  - (2), (3): $\pi/2$ (**constant!**)

The CKM Matrix

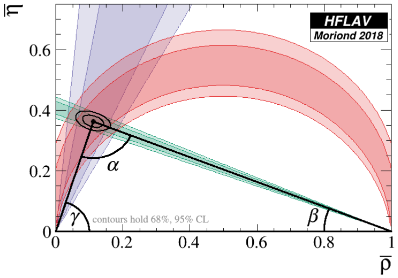
Another parametrization – Prog. Part. Nucl. Phys. 47 (2001)

$$V_{\text{CKM}} = \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}|e^{-i\tilde{\gamma}} \\ -|V_{cd}|e^{+i\phi_4} & |V_{cs}|e^{-i\phi_6} & |V_{cb}| \\ |V_{td}|e^{-i\tilde{\beta}} & -|V_{ts}|e^{+i\phi_2} & |V_{tb}| \end{pmatrix}$$

$$\begin{aligned} \gamma &\equiv \tilde{\gamma} - \phi_4, & \phi_2 &\approx \eta\lambda^2 \\ \beta &\equiv \tilde{\beta} + \phi_4, & \phi_4 &\approx \eta A^2\lambda^4 \\ \alpha &\equiv \pi - \beta - \gamma, & \phi_6 &\approx \eta A^4\lambda^6 \\ \beta_s &\equiv \phi_s \equiv \phi_2 + \phi_6 \end{aligned}$$

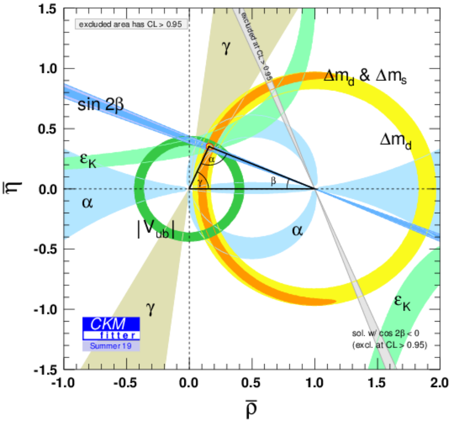
- From unitarity: 6 triangles
  - $V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0$
  - $V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$
  - … 4 more
- Angles $\tilde{\alpha}$, $\tilde{\beta}$, $\tilde{\gamma}$ and $\phi_{2,4,6}$ **depend on phase convention** (*i.e.*, not observable)
- Phases of products $V_{ij}V_{kl}V_{il}^*V_{kj}^*$ are **invariant** and **observable** (*e.g.*, α, β, γ, …)

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$$



Probing the SM by Overconstraining

- **Motivation:**
  - We utterly fail explaining CPV on cosmological scale!
  - Is the CKM matrix the only source of CPV?
  - Is the SM incomplete / are there more particles?
- **Strategy:**
  - Overconstrain CKM triangles by measuring
    - sides – *e.g.*, $|V_{td}V_{tb}^*|$
    - angles – *e.g.*, $\gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right)$
  - Is $\beta_s = \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right)$ tiny? ($\mathcal{O}(\lambda^2)$)



Any deviation would point towards **new physics**, *e.g.*, 4th quark family?

Branching ratios