Search for CP violation in the charm sector at LHCb

Maxime Schubiger, Maurizio Martinelli, Olivier Schneider

SPS annual meeting
25. August 2016

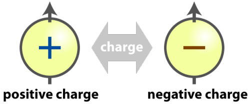
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

CP violation

- CP violation has been observed in weak interactions
- CP violation exists in the Standard Model (SM) but it is too small to explain a matter-dominated universe
- New Physics (NP) processes could fill this gap of the SM

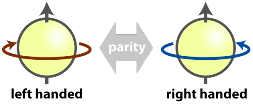
Charge conjugation symmetry C

$$q \rightarrow -q$$



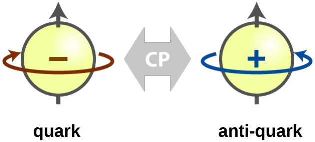
Parity symmetry P

$$\vec{x} \rightarrow -\vec{x}$$



CP symmetry

$$(q, \vec{x}) \rightarrow (-q, -\vec{x})$$



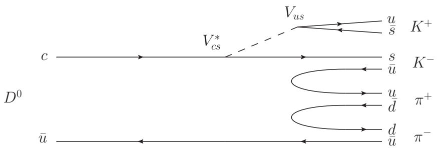
CP violation in charm decays

- CP violation in charm decays is a very good probe for NP beyond the SM
  - CP violation in charm decays is expected to be very small in SM $\mathcal{O}(10^{-4} - 10^{-3})$
    - see PRD 51 (1995) 3478
  - NP could enhance this CP violation up to $\mathcal{O}(10^{-2})$
    - see PRD 75 (2007) 036008
  - Effects from both direct (decay) and indirect (mixing) CP violation can be searched for

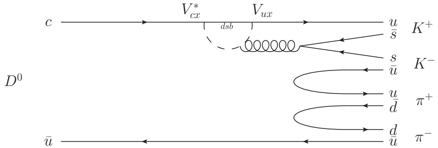
CP violation in $D^0 \to K^+K^-\pi^+\pi^-$ decays

- $D^0 \to K^+K^-\pi^+\pi^-$ is a good probe for the search of CP violation in charm decays
  - It is singly Cabibbo suppressed with two competing diagrams

Tree

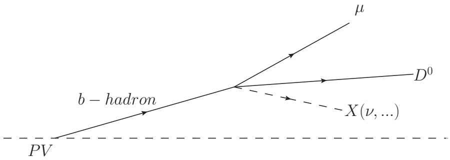


Loop

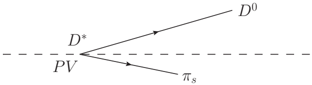


  - Many competing amplitudes could enhance local CP violation effects
  - It has two tagging channels with large statistics

Semileptonic (180k)



Prompt (250k)



Previous results on $D^0 \to K^+K^-\pi^+\pi^-$

FOCUS (PLB 610 (2005) 225) and CLEO (PRD 85 (2012) 122002) did an amplitude analysis with ∼**1300** and ∼**3000** signal events respectively

- The sensitivity achieved is **5%** − **30%** per amplitude and no CP violation has been found

BaBar (PRD 81 (2010) 111103) and LHCb (JHEP 10 (2014) 005) did an integrated measurement of the CP violation

→ Best result has been achieved by LHCb with : $A_{CP} = (1.8 \pm 2.9(\text{stat}) \pm 0.4(\text{syst})) \times 10^{-3}$, ie. compatible with no CP violation.

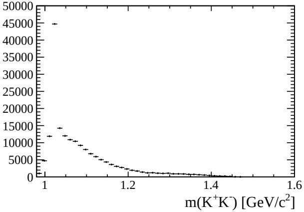
Goals of the analysis :

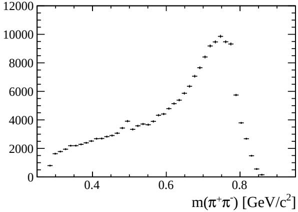
- Do a complete amplitude analysis
- Use Run 1 data (∼**200000**) with possible addition of Run 2
- Achieve sensitivity < **1%** on all amplitudes

→ could lead to first observation of CP violation in charm

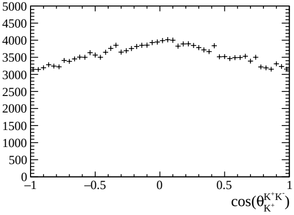
$D^0 \to K^+K^-\pi^+\pi^-$

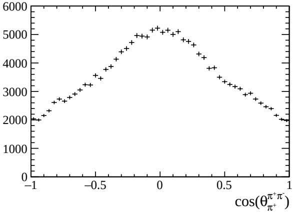
- $D^0 \to K^+K^-\pi^+\pi^-$ is a four-body decay
- It is described by a 5D phase space
- The following 5 variables (Cabibbo-Maksimowicz variables) describe the phase space :

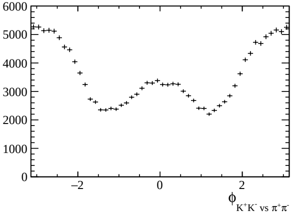




Full Run 1 data signal distribution





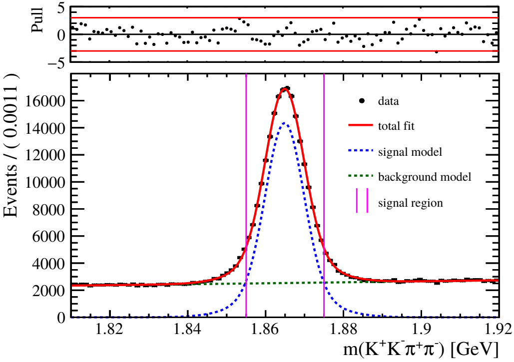


Selection

- Create a Boosted Decision Tree (BDT)
- Selected 15 most discriminating variables
  - Only topological variables, no particle identification (PID)
- Signal is 2012 MC sample, MU + MD, 180k events
- Backgroud is 2012 data sidebands, MU + MD, 180k events
- Signal and background samples are split in two for Training and Testing
- Significance is optimised on data for the BDT and the PID variables

Mass distribution

Resulting mass distribution with optimised cuts (full Run I)



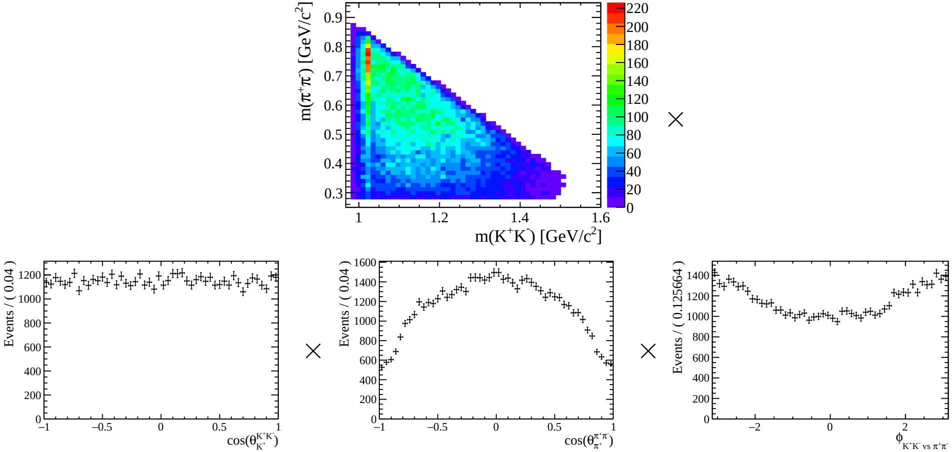
- Parameters of the distribution inside $\pm 2\sigma$ signal region :
  - $N_{\text{sig}} = 161103 \pm 551$
  - $N_{\text{bkg}} = 46277 \pm 122$
  - purity $= 78\%$
  - significance $= 354$

Study of the background

- The purity of the selected data is $\sim 80\%$
- The remaining $\sim 20\%$ need to be handled carefully
- Different strategies have been studied
  - background subtraction with the sPlot technique
  - background description with empirical 5D pdf
  - background parametrisation
- All have their advantages and drawbacks

Background description

- Empirical description of the data sidebands with histograms
- Used 4 histograms to build the pdf
  - 2D × 1D × 1D × 1D
  - $(m_{KK}, m_{\pi\pi}) \times cos(\theta_K) \times cos(\theta_\pi) \times \phi$



Fitting the signal

A fitter has been developed in order to interact directly with the well-known minimizer *Minuit*. The main component is the construction of the likelihood that will be minimized :

$$\mathcal{L} = -2 \cdot ln \left(\prod_{data} \mathcal{L}(\boldsymbol{x}_i) \right)$$

$$\mathcal{L}(\boldsymbol{x}_i) = \frac{|\sum_{amp} c_j \cdot A_j(\boldsymbol{x}_i)|^2}{\int_{MC} |\sum_{amp} c_j \cdot A_j(\boldsymbol{p})|^2 d\boldsymbol{p}}$$

where :

- c_j : are the complex fit parameters
- A_j : are the amplitudes described by the multiplication of a Blatt-Weisskopf factor, a Breit-Wigner lineshape and a spin factor
- $\boldsymbol{x}_i$: are the data events
- $\boldsymbol{p}$: are the MC events

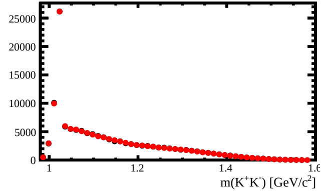
Amplitude model

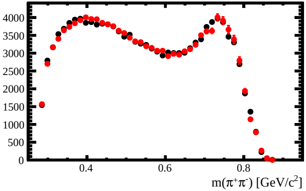
- A new amplitude model will be developed
- Starting point is CLEO model
  - 11 amplitudes
  - PRD 85 (2012) 122002

$$D^0 \to \phi(1020)^0(\to K^+, K^-), \rho(770)^0(\to \pi^+, \pi^-)$$
$$D^0 \to \phi(1020)^0(\to K^+, K^-), \rho(770)^0(\to \pi^+, \pi^-) \text{ (D wave)}$$
$$D^0 \to \phi(1020)^0(\to K^+, K^-), \text{NonRes}(\pi^+, \pi^-)_S$$
$$D^0 \to K^*(892)^0(\to K^+, \pi^-), \overline{K^*(892)^0}(\to K^-, \pi^+)$$
$$D^0 \to K_1(1270)^+(\to K^*(892)^0(\to K^+, \pi^-), \pi^+), K^-$$
$$D^0 \to K_1(1270)^-(\to \overline{K^*(892)^0}(\to K^-, \pi^+), \pi^-), K^+$$
$$D^0 \to K_1(1270)^+(\to \rho(770)^0(\to \pi^+, \pi^-), K^+), K^-$$
$$D^0 \to K_1(1270)^-(\to \rho(770)^0(\to \pi^-, \pi^+), K^-), K^+$$
$$D^0 \to K^*(1410)^+(\to K^*(892)^0(\to K^+, \pi^-), \pi^+), K^-$$
$$D^0 \to K^*(1410)^-(\to \overline{K^*(892)^0}(\to K^-, \pi^+), \pi^-), K^+$$
$$D^0 \to \text{NonRes}(K^-, \pi^+)_P, \text{NonRes}(K^+, \pi^-)_S$$

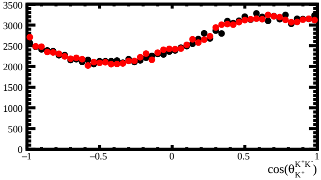
Pseudo-experiment

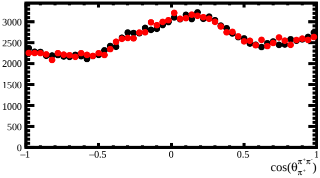
- Fitting procedure has been tested on pseudo-experiments
- Good agreement between generated and fitted sample
- Test done for signal model

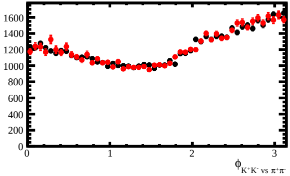




+ fit
+ simulated events







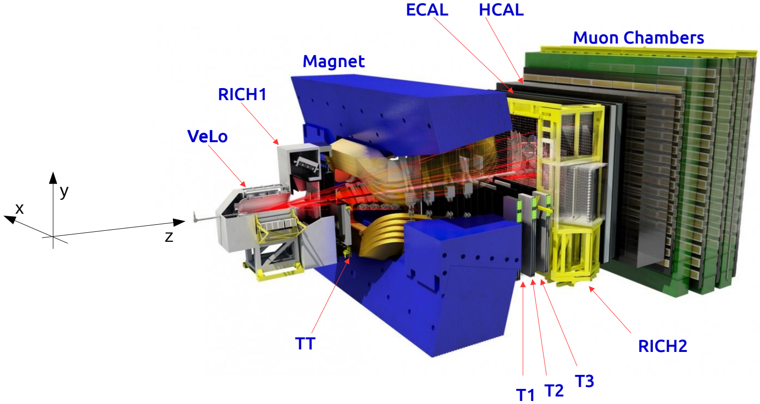
Conclusion

- The first part of the analysis is to fit the data with an amplitude model
- The background will be fitted by an empirical pdf
- Various signal models will be tested
- Systematics uncertainties will be studied
- Once the best model and all uncertainties are determined, the data will be split between the D^0 and $\bar{D}^0$ in order to search for CP violation
- Any hint of CP violation of $\mathcal{O}(10^{-2})$ or above would be a clear sign for NP

Backup Slides

The LHCb detector

- One of the four main experiments of the LHC
- Forward arm spectrometer
- Main goals of LHCb
    - Search for New Physics (NP) beyond the Standard Model (SM)
    - Study *b*- and *c*-hadrons
    - Study for CP violation, i.e. understand the differences between matter and anti-matter



Selection

- Stripping requirements
  - Stripping 20, semileptonic.dst
  - Stripping line : b2D0MuX2K2PiB2DMuNuXLine
- Trigger requirements (L0 && HLT1 && HLT2)
  - L0 : D0_L0HadronDecision_TOS || Mu_L0MuonDecision_TOS
  - HLT1 : B_Hlt1TrackAllL0Decision_TOS
    || Mu_Hlt1TrackAllMuonDecision_TOS
  - HLT2 : Mu_Hlt2SingleMuonDecision_TOS
    || B_Hlt2TopoMu2BodyBBDTDecision_TOS
    || B_Hlt2TopoMu3BodyBBDTDecision_TOS
    || B_Hlt2TopoMu4BodyBBDTDecision_TOS

Stripping selection requirements

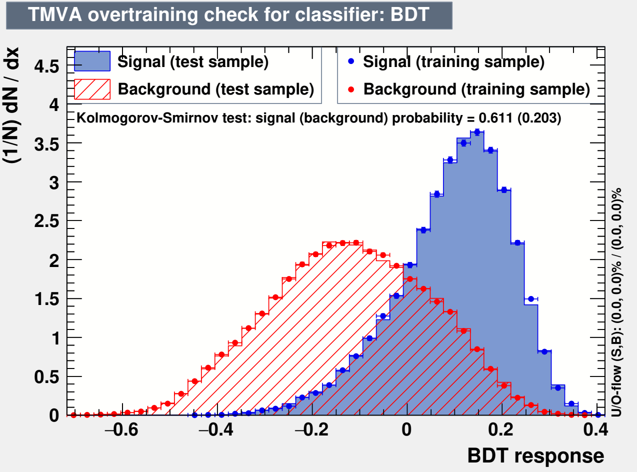
Table 44: Stripping selection requirements.

Candidate	Variable	2K2Pi	K3Pi
K	track $\chi^2/\mathrm{n_{dof}}$	< 3	
	p	$> 2\,\mathrm{GeV}/c$	
	p_T	$> 300\,\mathrm{MeV}/c$	
	$\chi^2(\mathrm{IP})$	> 9	
	PIDK	> 4	> 8
	Prob(ghost)	< 0.5	
π	track $\chi^2/\mathrm{n_{dof}}$	< 3	
	p	$> 2\,\mathrm{GeV}/c$	
	p_T	$> 300\,\mathrm{MeV}/c$	
	$\chi^2(\mathrm{IP})$	> 9	
	PIDK	< 10	
	Prob(ghost)	< 0.5	
μ	track $\chi^2/\mathrm{n_{dof}}$	< 3	
	p	$> 2\,\mathrm{GeV}/c$	
	p_T	$> 1.2\,\mathrm{GeV}/c$	
	$\chi^2(\mathrm{IP})$	> 9	
	PIDmu	> 0	
	Prob(ghost)	< 0.5	
D^0	mass	$\epsilon[1.8, 1.925]\,\mathrm{GeV}/c^2$	
	$\sum_{\mathrm{daughters}} p_T^i$	$> 1.8\,\mathrm{GeV}/c$	
	$\chi^2(\mathrm{Vtx})/\mathrm{n_{dof}}$	< 6	
	cos(DIRA)	> 0.99	
B	$M(p_{D^0} + p_\mu)$	$< 6.2\,\mathrm{GeV}/c^2$	
	mass	$\epsilon[2.5, 6.0]\,\mathrm{GeV}/c^2$	
	$\chi^2(\mathrm{Vtx})/\mathrm{n_{dof}}$	< 6	
	cos(DIRA)	> 0.999	

BDT variables

- B p_{T} corrected mass
- B $\log(\chi^2_{\mathrm{IP}})$
- B mass
- B log(min $\Delta\chi^2$ 1 track)
- B χ^2_{vtx}
- D^0 log(min $\Delta\chi^2$ M 1 track)
- D^0 log(min $\Delta\chi^2$ 2 tracks)
- D^0 $p_{\mathrm{T,asym}}$
- D^0 log(DLS)
- D^0 $\log(\tau)$
- D^0 log(min $\Delta\chi^2$ 1 track)
- D^0 p
- D^0 p_{T}
- D^0 $\log(\chi^2_{\mathrm{IP}})$
- μ $\log(\chi^2_{\mathrm{IP}})$

BDT response

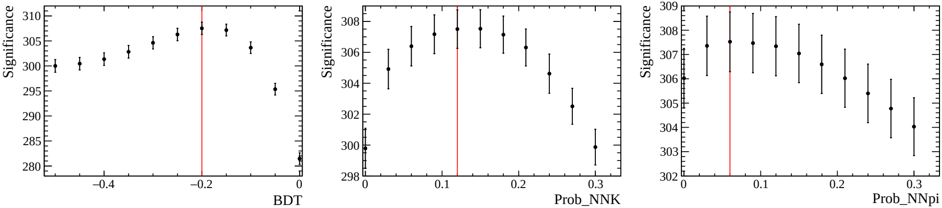


Selection optimisation

- Optimisation is done on data
- Optimisation is done on the BDT, Prob_NNK on kaons and Prob_NNpi on pions
- Significance ($S/\sqrt{S+B}$) in $\pm 2\sigma$ signal region is used as FoM
- The 3 sets of cuts are applied simultaneously
- Explore a volume of BDT $\in [-0.5, 0]$, Prob_NNK and Prob_NNpi $\in [0, 0.3]$ (2000 fits)

$\Rightarrow$ Cuts chosen: BDT>-0.2, Prob_NNK>0.12 and Prob_NNpi>0.06

2012 plots:



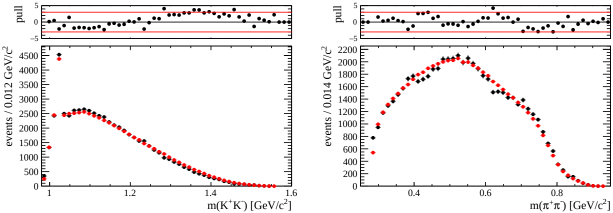
- Similar behaviour is observed for 2011 data sample

Expected systematics

- The method chosen will be insensitive to :
  - $D^0 - \overline{D^0}$ production asymmetry
  - D^0 flavour-tagging asymmetry
  - K^+K^- $(\pi^+\pi^-)$ reconstruction asymmetry (to first order)
- It will have to be studied where the main systematics come from :
  - the model
  - the background description
  - the efficiency determination

Fit of the sidebands

- The 2012 sidebands have been fitted
- Integration MC sample : 1 M phase space events



$\phi\pi\pi$
ρKK
$K^{*}K\pi$
$\overline{K^{*}}K\pi$
$KK\pi\pi$