Sensitivity of Triple Product Asymmetries in $D^+ \to K_S^0 K^+ \pi^- \pi^+$ Decay at LHCb

Shaowei Wu
Supervisor: Dr.Maurizio Martinelli and Dr.Tara Nanut

Contents

1 Introduction

CP-violation of charm decays has never been observed yet. This time we try to apply triple-product asymmetries as a complementary way for searching it. The aim of this project is to calculate sensitivity of certain observables from the likelihood and extropolate it to full LHC Run2.

2 Observables and datasets

1. Observables
   There are four particles in the final state of the decay $D^+ \to K_S^0 K^+ \pi^- \pi^+$. The momentum of K^+, π^- and π^+ are independent and can be used to build the T-odd observable
   $$C_T = \overrightarrow{p_{k^+}} \cdot \overrightarrow{p_{\pi^+}} \times \overrightarrow{p_{\pi^-}}.$$

Asymmetries and CP-violating observable are defined as

$$A_T = \frac{N(D^+, C_T > 0) - N(D^+, C_T < 0)}{N(D^+, C_T > 0) + N(D^+, C_T < 0)},$$

$$\overline{A}_T = \frac{N(D^-, -\overline{C}_T > 0) - N(D^-, -\overline{C}_T < 0)}{N(D^-, -\overline{C}_T > 0) + N(D^-, -\overline{C}_T < 0)},$$

$$a_{CP} = \frac{1}{2}(A_T - \overline{A}_T).$$

2. Dataset
   The datasets include separate samples for K_S^0 decaying in the Vertex detector(LL) or after(DD). Both DD and LL include events of two decay channel samples:

   (a) $D^+ \to K_S^0 \pi^+ \pi^- \pi^+$
   Control sample. No CP-violation is expected in this channel. We use it as an input source for TMVA to search the best BDT cut to suppress background on the signal sample.

   (b) $D^+ \to K_S^0 K^+ \pi^- \pi^+$
   Signal sample. There is possibly CP-violation in this channel.

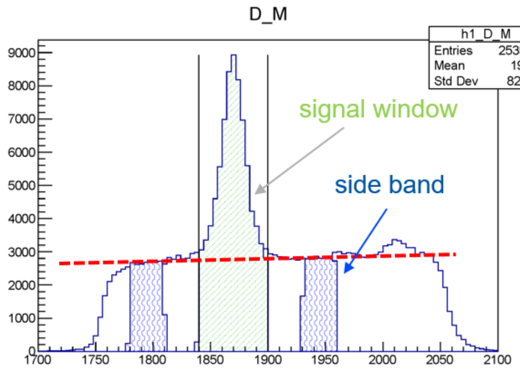
3 MVA

1. Choice of Input Variables
   If distributions of a variable are different between signal and background, then the variable is potential to be used to distinguish between signal and background. A simple and robust way to choose BDT training variables is background subtraction.

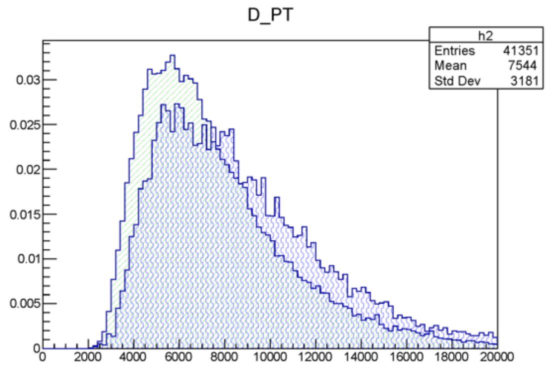
   We define the signal window to be the $\pm$30MeV range symmetric around the expected D^+mass and a lower sideband and an upper sideband, each 30 MeV wide, symmetrically spaced with respect to the center of the nominal signal region.

   After the background subtraction, we choose D_PT, KS_PT, P1_ProbNNpi, P2_ProbNNpi, D_ReFit_chi2, D_ReFit_decayLength/D_ReFit_decayLengthErr, log(D_IPCHI2_OWNPV) as the variables and expressions used in TMVA.

2. $_S Plot$
   Before the train we should first extract s-weight to use in TMVA for $D^+ \to K_S^0 \pi^+ \pi^- \pi^+$. The fit result is shown below. The P.D.F. contains two gaussian distributions (blue and green dashed curves) and a cubic polynomial (red dashed curve).



(a) D^+ mass of $D^+ \to K_S^0\pi^+\pi^-\pi^+$(DD)

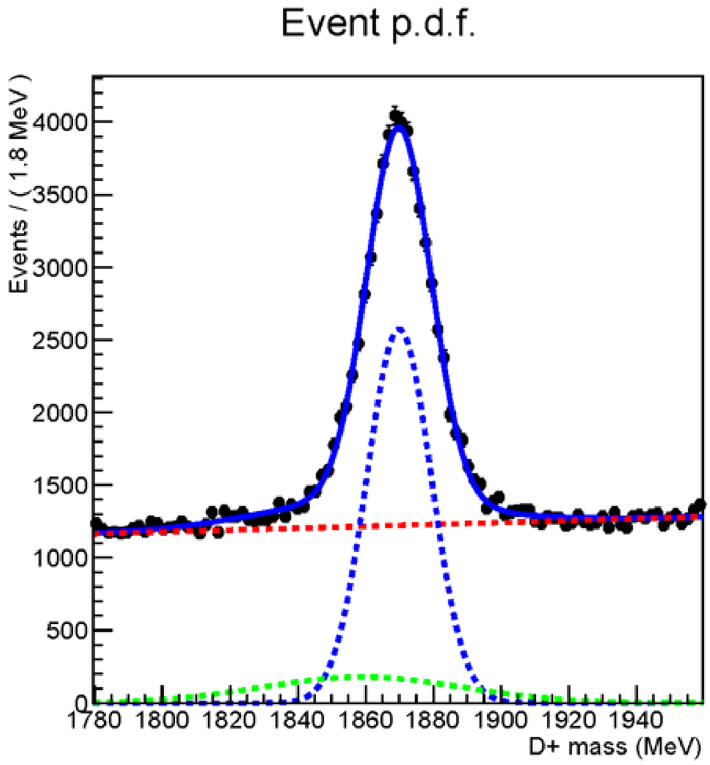


(b) D_PT distribution of estimated signal(blue area) and background(green area)

3. Train with selected variables and expressions
   Considering the overtrain, we should adjust the parameters of BDT method to reduce overtrain as much as possible. The K-S test outputs from the TMVA are usually interpreted as being a signal of good non-overtrained performance if they are close to one, and on the other hand being a signal of overtraining they are very close to zero. Below is the parameter values and K-S test results of DD and LL for $D^+ \to K_S^0\pi^+\pi^-\pi^+$.

Sample category	NTrees	MaxDepth	nCuts
DD	150	2	20
LL	50	2	10

4. Find the best BDT cut and suppress the background
   The calculation of BDT response for $D^+ \to K_S^0\pi^+\pi^-\pi^+$ is based on the result



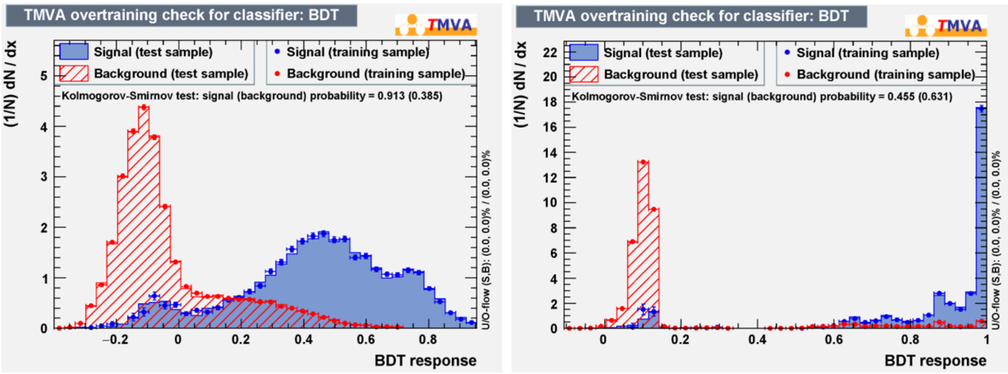
of TMVA training. Then we normalize the signal yield and background yield to $D^+ \to K_S^0 K^+ \pi^- \pi^+$ to calculate the correponding significance after BDT cut. The significance is defined as

$$S = \frac{N_s}{\sqrt{N_s + N_b}}$$

When BDT response > 0.43, the significance reaches the maximum value for DD. While for LL, it is 0.90. (See Figure 3 and 4)

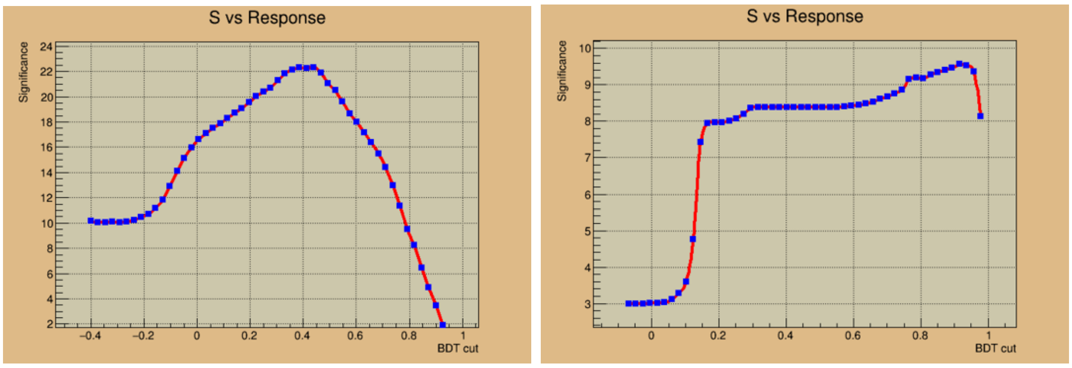
5. Fit mass spectrum and get parameter values
The next step is applying the BDT cut to the $D^+ \to K_S^0 K^+ \pi^- \pi^+$ sample (DD and LL). Before that, we shoud fit D^+/D^-'s mass spectrum for DD/LL to get 4 models ($D^+_{(DD)}$, $D^-_{(DD)}$, $D^+_{(LL)}$ and $D^-_{(LL)}$) with parameters.(See Figure 5)

Figure 2: TMVA overtraining check for classification



(a) Overtraining Check for DD (b) Overtraining Check for LL

Figure 3: Significance vs BDT cut



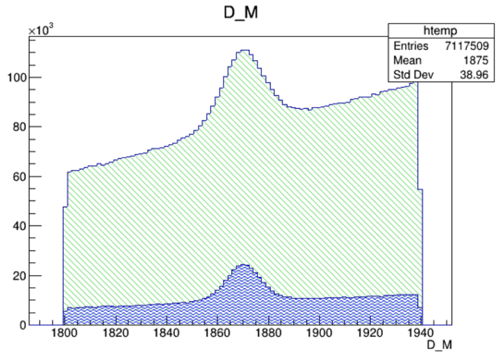
(a) DD (b) LL

4 Simultaneous fit

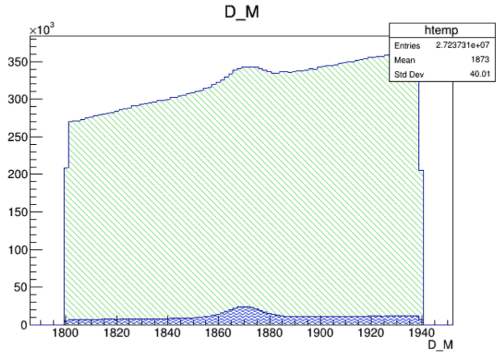
1. Fit models to subsamples

   There are 4 models and 8 subsamples(2 subsamples ($C_T > 0$ and $C_T < 0$) for each model). $A_T^{sig}, \overline{A}_T^{sig}$ are in common among DD and LL. $A_T^{bkg}, \overline{A}_T^{bkg}$ have separate values for DD and LL. We fit the 8 P.D.F.s simultaneously to 8 subsamples to get the result(see Figure 6). $A_T^{sig}, \overline{A}_T^{sig}$ are set to be "blind" using blind analysis using class *RooUnblindPrecision*. Uncertainty of a_{CP} are derived from propagation of uncertainties of $A_T^{sig} and \overline{A}_T^{sig}$.

Figure 4: $D^{\pm}$mass spectrum. Blue area →after BDT cut. Green area →before BDT cut



(a) DD



(b) LL

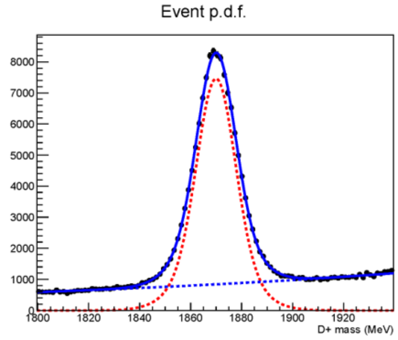
Variable	Value	Uncertainty
signal yield (DD)	226757	± 546
signal yield (LL)	166245	±578
background yield (DD)	168905	±494
background yield (LL)	796116	±1035
A_T	2.3×10^{-3}	2.5×10^{-3}
$\overline{A}_T$	-1.4×10^{-3}	2.5×10^{-3}
a_{CP}	1.8×10^{-3}	2.0×10^{-3}

2. Extrapolation

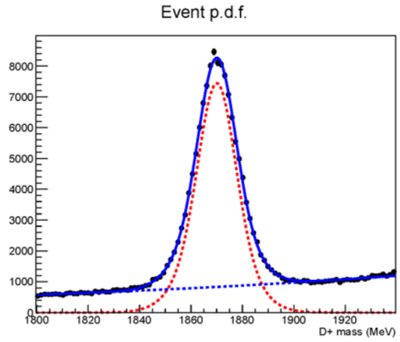
The integrated luminosity of our sample is L=1.33/fb^{-1}. For LHC Run2, L^f= 6/fb^{-1}

Let N_S be the number of signal events in sample. Suppose $N_S \propto L$. For a large enough sample, the uncertainty is inversely proportional to the square root of sample size N.

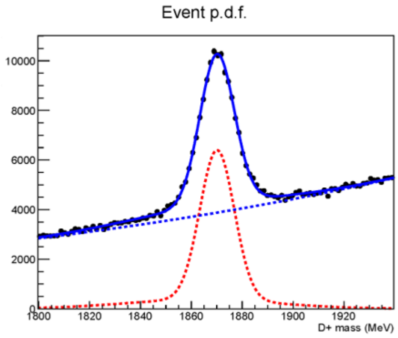
$$\left.\begin{array}{r} \dfrac{N_S}{N_S^f} = \dfrac{L}{L^f} \\[2em] \dfrac{\sigma_{a_{CP}}(N_S)}{\sigma_{a_{CP}^f}(N_S^f)} = \sqrt{\dfrac{N_S^f}{N_S}} \end{array}\right\} \Rightarrow \sigma_{a_{CP}^f}(N_S^f) = \sigma_{a_{CP}}(N_S)\sqrt{\frac{L}{L^f}} = 2.0\times10^{-3}\times\sqrt{\frac{1.33}{6}} = 9.6\times10^{-4}$$



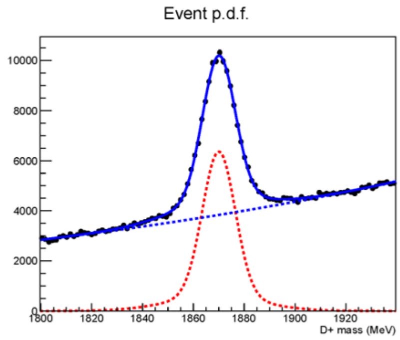
(a) Fit result of D^+(DD)



(b) Fit result of D^-(DD)



(c) Fit result of D^+(LL)



(d) Fit result of D^-(LL)

5 Conclusion

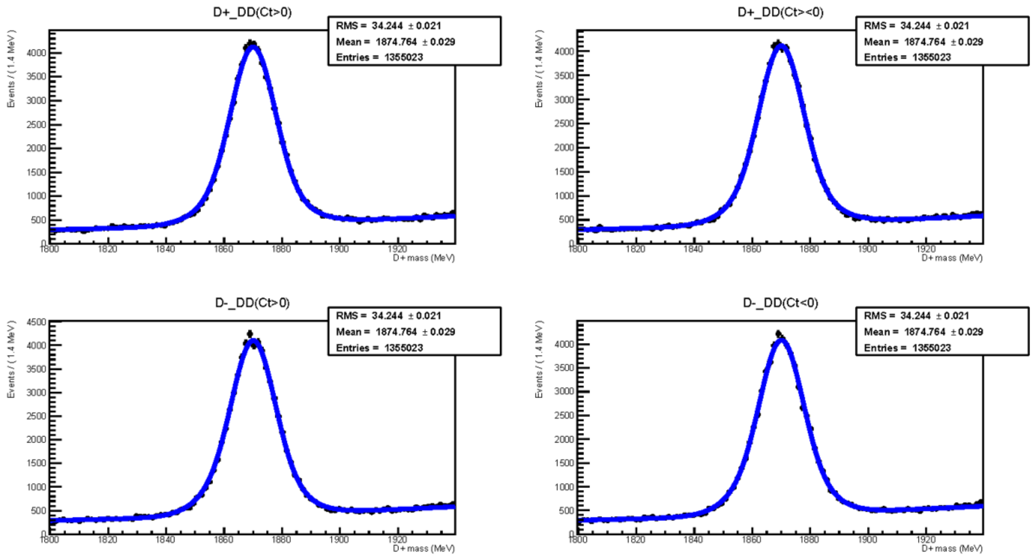
We use background subtraction as a simple and robust way to sort out variables for MVA. Then a small control sample is used to train the classifier. BDT cut proved to be an effective method to suppress background. We fit the samples after cut and extrapolate the uncertainty to the full LHC Run 2. This is the first time that T-odd observables are used in charm baryon decay.

6 Acknowledgement

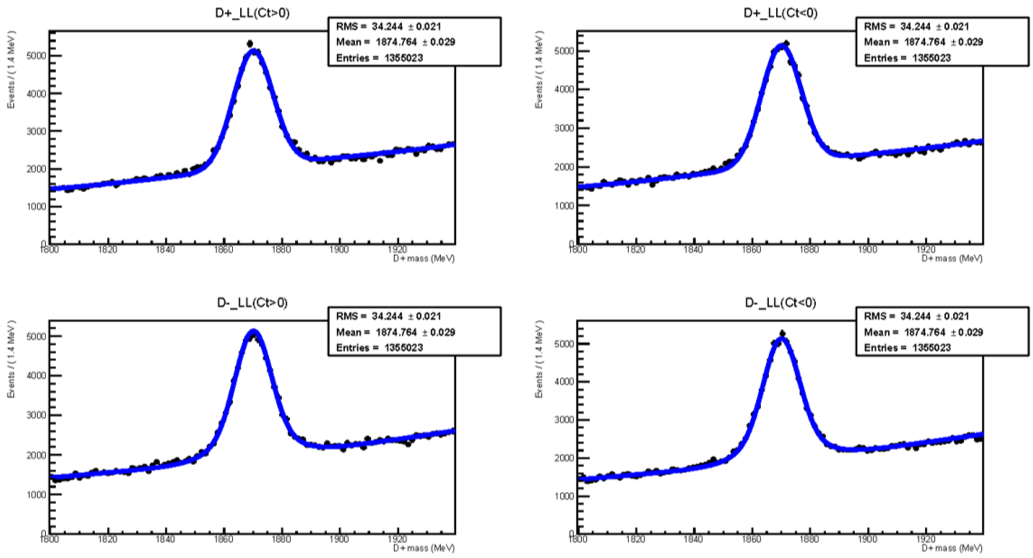
I would like to express my deepest appreciation to all those who provided me the possibility to complete this report. A special gratitude I give to my supervisors, Dr.Maurizio Martinelli and Dr.Tara Nanut, whose contribution in stimulating suggestions and encouragement, helped me to coordinate my project especially in writing this report.

I would like to express my gratitude towards my parents and member of CERN and Sichuan University for their kind co-operation and encouragement which help me in

Figure 6: Fit result for 8 subsamples



(a) DD



(b) LL

completion of this project.

My thanks and appreciations also go to my colleague in developing the project and people who have willingly helped me out with their abilities.

References

[1] Maurizio Martinelli: Measurements of T-odd observables,
https://arxiv.org/abs/1509.04615