ANALYSIS OF MULTIBODY BEAUTY TO OPEN-CHARM DECAYS AT LHCb[*]

Wojciech Krupa

on behalf of the LHCb Collaboration

AGH University of Science and Technology
Faculty of Physics and Applied Computer Science
al. Mickiewicza 30, 30-059 Kraków, Poland

The LHCb detector is designed for precise measurements of beauty and charm quark decays and searching for new physics beyond the Standard Model. Its precise tracking, excellent particle identification and high trigger efficiency allow to measure the CKM matrix parameters with unprecedented accuracy. Selected results regarding decays of beauty mesons to final states containing open charm particles as well as the description of data analysis techniques applied for such complicated multibody decays are presented.

1. Introduction

The quark sector consists of 6 particles grouped in 3 families (quark u and d, s and c, b and t). The Cabibbo–Kobayashi–Maskawa (CKM) matrix gives information on the strength of flavour-changing charged weak decays of the quarks. The CKM matrix is a unitarity matrix, *i.e.* fulfills the requirement: $\sum_i V_{ij}V_{ik}^* = 0$ ($j \neq k$, V_{ij}, V_{ik} *etc.* — elements of the CKM matrix), thus providing 9 unitarity conditions, among them 6 can be presented as triangles involving different physics processes. The most interesting ones are: $V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$ and $V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0$ because they may be measured in experiment within the current precision.

Uncertainty of the CKM angle $\gamma = \arg(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*})$ is still large. Therefore, further measurements are performed. The CKM angle γ in LHCb is measured both in decays with loops and in tree-level processes. The advantage of the former is a large number of expected signal events, but the theoretical

---

uncertainties make these measurements difficult to interpret. The latter are rare (the typical probability of the tree-level B meson decay is of the order of 10^{-3} and below). However, these decays are theoretically very clean $(\frac{\delta\gamma}{\gamma} < 10^{-7})$.

A discrepancy between the CKM angle γ measurements in tree-level decays and loops processes may indicate a large contribution of new physics beyond of Standard Model[1]. The CKM angle γ is determined as $65.3^{+1.0}_{-2.5}{}^{\circ}$ in all decays and $76.8^{+5.1}_{-5.7}{}^{\circ}$ in tree-level days only. The summary of the LHCb results is given in [1].

2. The LHCb detector

The LHCb spectrometer is especially designed to study heavy-flavour physics. It covers the pseudorapidity range $(2 < \eta < 5)$, which is sufficient to detect most of the final states of the B-mesons decays.
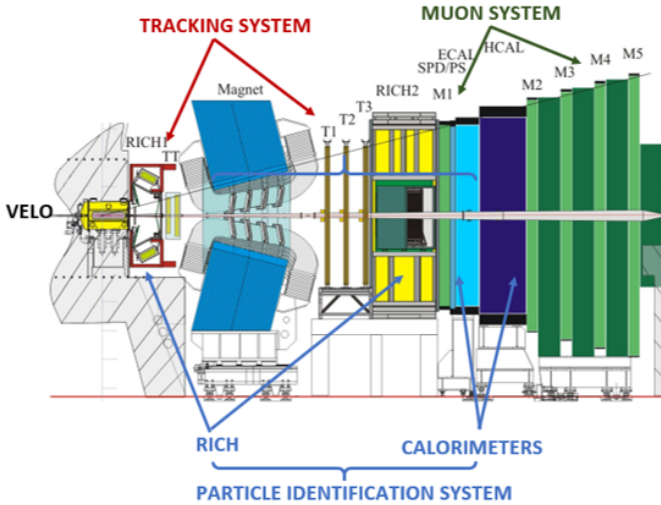

Fig. 1. The LHCb detector scheme [2].

The main detector properties that enable precision measurements are: excellent vertex resolution: ~ 13 μm, Impact Parameter (IP) resolution: $\sigma_{\mathrm{IP}} \sim 20$ μm, momentum resolution: $\frac{\sigma p}{p} \sim 0.5$ %, particle identification of hadrons: $\eta(h-h) \sim 90\%$, identification of muons: $\eta \sim 97\%$ and good decay time resolution ~ 45 fm. The scheme of the spectrometer is presented in Fig. 1. The details of the LHCb spectrometer can be found in [2].

---

[1] New physics beyond the Standard Model effects should contribute to loops decays. This contribution in tree-level decays is negligible.

3. Beauty to charm decays

The CKM angle γ in the tree-level processes can be determined by measurement of the interference between favourite $b \to c$ (V_{cb}) and suppressed $b \to u$ (V_{ub}) transition amplitudes[2].

The potential of LHCb experiment to determine the CKM angle γ lies in various possible measurements through processes like $B^{\pm} \to D^0 X$ (Fig. 2), $B^0 \to D^{\pm} X$ and $B_s^0 \to D_s^{\pm} X$ (where X stands for one or more pions and kaons). The LHCb measurements cover methods which have different sensitivity to the CKM angle γ depending on the type of the D meson final states and state X. Some of these methods are: GLW (Gronau–London–Wyler, final states of D meson decay: $KK/\pi\pi$) [3], ADS (Atwood–Dunietz–Soni, final states: $K\pi, K\pi\pi\pi$) [4], GGSZ (Giri–Grossman–Soffer–Zupan, three-body final states) [5] and finally Dalitz and time-dependent analysis. For a full review of these methods, see [1].
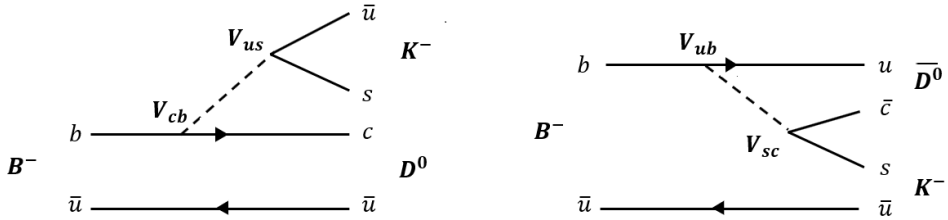

Fig. 2. Feynman diagram of $B \to D^0/\bar{D}^0 \; K^-$ decay.

4. Multibody decays analysis methods

Many of $B \to DK$ decays include multibody final states. A typical visible Branching Ratio is well below 10^{-5}, which leads to small signal-to-background ratio. A linear selection technique, where respective cuts are applied one at a time, is usually not effective enough to remove the background. Instead, a number of multivariate selection methods based on the machine learning algorithms are proved to be much more adequate.

4.1. Multivariate analysis methods

In multivariate approach, a model (classifier) is trained to classify events into classes in multidimensional space. A simple example of the effectiveness of this approach has been shown in Fig. 3.

---

[2] Suppressed and favoured decays depend on different quarks transitions that occur in the decay and corresponding coupling constants of the CKM mixing matrix (*e.g.* V_{ub}, V_{cd}).
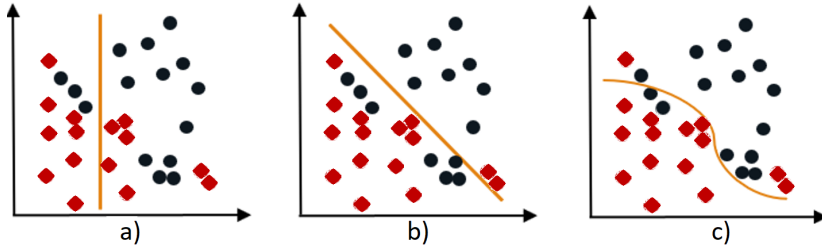
Fig. 3. Results of a division of events belong to two different classes (circle, rectangles) by (a) one-dimensional cut, (b) linear function, (c) non-linear function.

Let us assume that our events can be completely described using two variables and, thus, can be represented as points on a 2-dimensional plane. It is clear that, in the presented case, any simple linear cut will lead to a selection of a sample featuring either poor efficiency or purity. In the case where more sophisticated multivariate methods are used, the selection algorithm can be made very effective. Depending on a particular event distribution, models of different complexity can be used to resolve the boundary between different classes. The classifier training procedure can be regarded as an optimization problem and, in turn, its result can be represented as a hyperplane in a space of variables used to describe events.

In $B \to DK$ analysis, events are classified as signal or background. The training of a classifier is carried on Monte Carlo events (which modeled signal) and selected data events as background. Boosted Decision Tree is an example of the model effective in classification.

In the Boosted Decision Tree method (BDT), a model has a tree-like structure (Fig. 4). On the top, there are all events which will be classified. Every decision (D_1, D_2, D_3) splits them into sets (depending on the values of parameters: x_1, x_2, x_3). At the bottom, there are classes (K_1, K_2 ...).
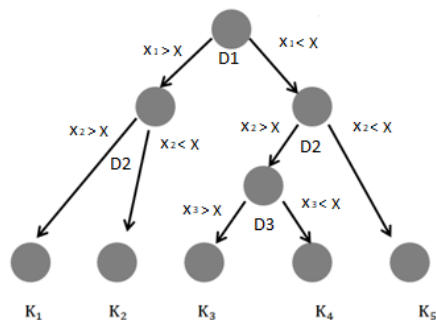

Fig. 4. Scheme of the classifier of the Boosted Decision Tree methods.

4.2. Verification

Several methods may be used in verification of a effectiveness of a trained classifier. Confusion Matrix (CM) is one among them. True Positives (TP) and True Negatives (TN) are these events, which have been classified properly (genuine signal events as the signal and background events as background) and misclassified — False Positives (FP) or False Negatives (FN). The CM consists of these four quantities. Metrices describe some efficiency of the classifier. Accuracy is a ratio of all properly classified events (TP and TN) to all events. True Positive Rate (TPR) is signal classification efficiency. There is also Misclassification Rate (ratio of a number of misclassified events to all events), False Positive Rate (FPR, the efficiency of misclassification of the signal) and more.

The CM depends on the *discrimination threshold*. Classifier estimates *a classifier response value* for every event. Afterward, the objects with the classifier response value above or below discrimination threshold are classified as signal or background. Information from many CM may be summarized. Very useful is Receiver Operating Characteristic (ROC) curve. It is a representation of the dependency of TPR to FPR for different discrimination threshold.

Area Under Curve (AUC) defines the quality of a classifier. For random classification, True Positive Rate is equal to False Positive Rate, which leads to AUC value equal 0.5 (Fig. 5 (c)). For the better classifier, AUC value will be closer to 1 (Fig. 5 (a)–(b)).
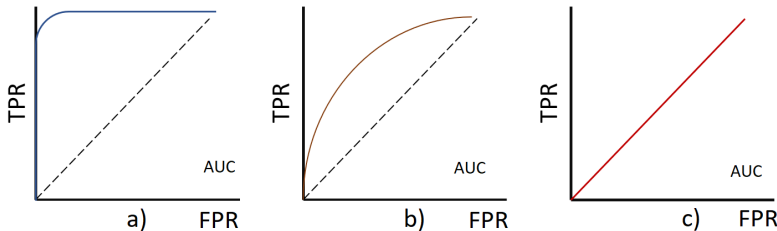

Fig. 5. The ROC curves for three different classifiers: (a) perfect classifier, (b) good classifier, (c) random guessing.

The multidimensional approach in multibody decays analysis is used in the background subtraction. An example of the usage of BDT methods has been presented in Fig. 6. The plots show the difference between $D_s^{\pm*}$ and $D_s^{\pm}$ mesons mass for selected events collected in the LHCb spectrometer during Run 2. The signal peak, which should be around 140 MeV, is hidden in a large background (Fig. 6 (a)). Various classifier response value (BTD) cuts are used for each plot (Fig. 6 (b)–(d)). The expected signal peak becomes visible for higher BDT cut (Fig. 6 (d)).
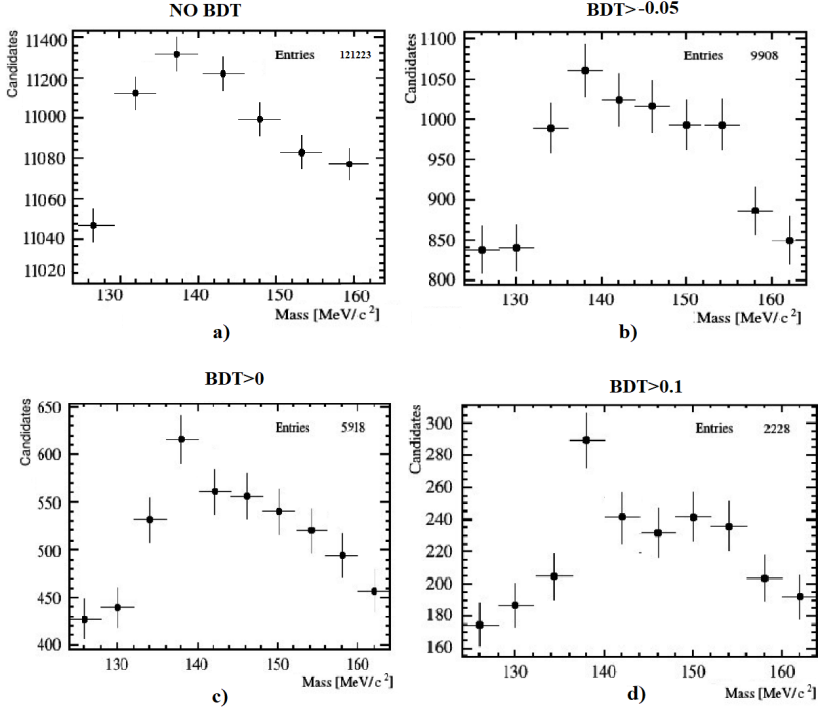
Fig. 6. Distributions of the mass difference between $D_s^{\pm *}$ and $D_s^{\pm}$ candidates for selected events collected in the LHCb spectrometer during Run 2 for: (a) only rectangular cut, (b) BDT > -0.05 cut, (c) BDT > 0 cut, (d) BDT > 0.1 cut.

5. Conclusion

Beauty to open charm decays ($B \to DK$) are crucial in the determination of the CKM angle γ and searches of new physics beyond the Standard Model. Precise measurements of this angle require performing analysis of many decay modes. These analyses are demanding because of the rarity of $B \to DK$ decays and became possible due to a usage of a multidimensional approach. Application of method such as Boosted Decision Tree requires taking into account some aspects of proper training or verification of the utility of a classifier. The multidimensional approach is used not only in data analysis but also in many other fields of physics such as particle track reconstruction, identification of particles, data quality monitoring, fake track rejection and more.

REFERENCES

[1] R. Aaij *et al.* [LHCb Collaboration], LHCb-CONF-2017-004.

[2] R. Aaij *et al.* [LHCb Collaboration], *Int. J. Mod. Phys. A* **30**, 1530022 (2015).

[3] M. Gronau, D. London, *Phys. Lett. B* **253**, 483 (1991).

[4] D. Atwood, I. Dunietz, A. Soni, *Phys. Rev. Lett.* **78**, 3257 (1997).

[5] A. Giri, Y. Grossman, A. Soffer, J. Zupan, *Phys. Rev. D* **68**, 054018 (2003).