

# Measurement of the CP Asymmetry in $B^+ \rightarrow D_s^+ \bar{D}^0$ and $B^+ \rightarrow D^{(*)+} \bar{D}^0$ Decays

Patrick Odagiu\*

Supervisor: Dr. Alison Tully

École Polytechnique Fédérale de Lausanne - LPHE

JUNE 8, 2020

The  $CP$  asymmetry of  $B^+ \rightarrow D^{(*)+} \bar{D}^0$  and  $B^+ \rightarrow D_s^+ \bar{D}^0$  decays was measured in two datasets. The first corresponds to an integrated luminosity of  $3.0 \text{ fb}^{-1}$  collected during LHC Run I with the LHCb detector, while the second contains  $6.0 \text{ fb}^{-1}$  of Run II data. The datasets were combined to obtain:  $\mathcal{A}^{\text{CP}}(B^+ \rightarrow D^+ \bar{D}^0) = (0.7 \pm 1.0 \pm 0.4)\%$ ,  $\mathcal{A}^{\text{CP}}(B^+ \rightarrow D_s^+ \bar{D}^0) = (-2.6 \pm 0.2 \pm 0.5)\%$ , and  $\mathcal{A}^{\text{CP}}(B^+ \rightarrow D^{*+} \bar{D}^0) = (1.6 \pm 1.5 \pm 0.4)\%$ , where the uncertainties are statistical and systematic. It was observed that the  $D_s^+$  mode shows hints of  $CP$  violation within the Standard Model limits.

## INTRODUCTION

The only known source of both parity and charge conjugation ( $CP$ ) violation in the Standard Model (SM) of particle physics is described by the unitary  $3 \times 3$  Cabibbo-Kobayashi-Maskawa (CKM) matrix [1, 2]. The study of  $CP$  violation is important as it is necessary for the generation of matter-antimatter asymmetry in the early Universe. The weak decays of hadrons are governed by transition amplitudes  $A_f$ , proportional to elements of the CKM matrix. A different decay rate of a hadron, such as the  $B^+$  meson, compared to its antiparticle is possible if there exists interference of at least two quark-level transition processes.

Consider the  $B \rightarrow f$  transition amplitude  $A_f$ , together with its  $CP$  conjugate decay process with amplitude  $\bar{A}_{\bar{f}}$ . Two types of phases can appear in these decay amplitudes.  $CP$  violating phases do change sign under  $CP$  conjugation. In the SM, these phases are known to appear only in interactions including the  $W^\pm$  bosons, hence given the name weak phases. Moreover, a second type of phase can also appear, called strong phase; they are generated by  $CP$ -invariant interactions and are the same in both  $A_f$  and  $\bar{A}_{\bar{f}}$ . For example, consider a two particle decay

$$A_f = |a_1|e^{i(\delta_1+\phi_1)} + |a_2|e^{i(\delta_2+\phi_2)},$$

$$\bar{A}_{\bar{f}} = |a_1|e^{i(\delta_1-\phi_1)} + |a_2|e^{i(\delta_2-\phi_2)},$$

where  $\delta$  is the strong phase,  $\phi$  is the weak phase, and  $a$  is the magnitude of the contribution to the amplitude  $A_f$ . Furthermore, the  $CP$  asymmetry in the decay of a  $B^+$  meson<sup>1</sup> to two charmed mesons is defined as

$$\mathcal{A}^{\text{CP}} \equiv \frac{\Gamma(B^+ \rightarrow D_{(s)}^{(*)+} \bar{D}^0) - \Gamma(B^- \rightarrow D_{(s)}^{(*)-} D^0)}{\Gamma(B^+ \rightarrow D_{(s)}^{(*)+} \bar{D}^0) + \Gamma(B^- \rightarrow D_{(s)}^{(*)-} D^0)}$$

where  $\Gamma$  is the decay rate of the respective channel [3]. Therefore, by taking  $|A_f|^2 = \Gamma$ , the  $CP$  asymmetry in  $B^\pm$  decays is written in terms of strong and weak phases as

$$\mathcal{A}^{\text{CP}} = -\frac{2|a_1 a_2| \sin(\delta_2 - \delta_1) \sin(\phi_2 - \phi_1)}{|a_1|^2 + |a_2|^2 + 2|a_1 a_2| \cos(\delta_2 - \delta_1) \cos(\phi_2 - \phi_1)}.$$

Nonzero  $CP$  asymmetries are expected in the SM due to interference between the tree-level, loop-level, and annihilation type Feynman diagrams, as illustrated in Fig. 1. However, the SM predicts a  $CP$  violation of  $\mathcal{O}(1\%)$  [4–6]. New physics contributions, e.g., supersymmetric ones [7], can enhance the  $CP$  asymmetry in the measured  $B^+$  decays by up to 10% [6–8].

The  $CP$  violation phenomenon is well established experimentally, as it was first observed in neutral kaon decays as early as 1964 [9]. In  $B$  meson systems,  $CP$  violation was first observed in the interference between a decay with and without mixing [10, 11], and directly in  $B^0$  decays [10, 12]. Later, it was also observed in the  $B^\pm$  meson decays [13]. Assuming that the asymmetries are small and higher order terms corresponding to products of asymmetries can be neglected, the total  $CP$  asymmetry is experimentally determined using  $\mathcal{A}^{\text{CP}} \approx A_{\text{raw}} - A_{\text{P}} - A_{\text{D}}$ , where  $A_{\text{raw}}$  is the raw asymmetry,  $A_{\text{P}}$  is the production asymmetry, and  $A_{\text{D}}$  is the asymmetry related to the detection efficiencies. In this paper, only the raw asymmetry is determined using

$$A_{\text{raw}} \equiv \frac{N(B^+ \rightarrow D_{(s)}^{(*)+} \bar{D}^0) - N(B^- \rightarrow D_{(s)}^{(*)-} D^0)}{N(B^+ \rightarrow D_{(s)}^{(*)+} \bar{D}^0) + N(B^- \rightarrow D_{(s)}^{(*)-} D^0)}, \quad (1)$$

where  $N$  is the observed yield in the respective decay mode. The other two asymmetries are obtained from literature.

The main contributions to the studied decays are from tree-level and loop-level amplitudes [14], shown in Fig. 1. This paper presents the measurement of the  $CP$  asymmetry in the  $B^+ \rightarrow D^{(*)+} \bar{D}^0$  and  $B^+ \rightarrow D_s^+ \bar{D}^0$  decay schemes using LHCb data [15] from both LHC runs. The Run I data corresponds to a luminosity of  $3 \text{ fb}^{-1}$ , taken from 2011 to 2012 at  $\sqrt{s} = 7$  or 8 TeV. The Run II data corresponds to  $6.0 \text{ fb}^{-1}$ , taken in the years 2015 to 2018 at  $\sqrt{s} = 13$  TeV. The studied  $D$  mesons are reconstructed from the following decay chains:  $\bar{D}^0 \rightarrow K^+ \pi^-$ ,  $\bar{D}^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ ,  $D^+ \rightarrow K^- \pi^+ \pi^+$ , and  $D_s^+ \rightarrow K^+ K^- \pi^+$ . The  $D^{*+}$  de-excites by emitting a soft pion  $\pi_s^+$  in  $D^{*+} \rightarrow D^0 \pi_s^+$ .

The paper is structured as follows. In Section I, the LHCb detector is described, together with the simulations that were employed in this analysis. Section II details the use of multivariate methods and of fitting to compute the signal yield. The final results are presented in Section III.

\* Accompanying webiste: <https://bb511.github.io/B2DD/>.

<sup>1</sup> Charge conjugation is implied throughout, unless otherwise stated.

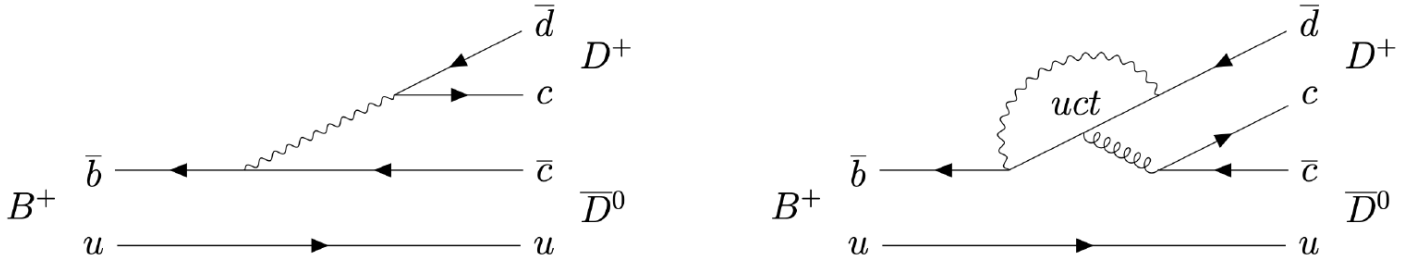


FIG. 1. The (left) tree diagram and (right) loop diagram contributions to the decay  $B^+ \rightarrow D^{(*)+} \bar{D}^0$ . Similar diagrams, with the  $\bar{d}$  quark replaced by an  $\bar{s}$ , apply to  $B^+ \rightarrow D_s^{(*)+} \bar{D}^0$  [16]. The  $D^{*+}$  modes are the same, but first de-excite by emitting a soft pion.

## I. LHCb DATA AND SIMULATION

The LHCb detector [15, 17] is a single-arm forward detector that covers the pseudo-rapidities between  $2 < \eta < 5$ . Originally, it was designed for the study of particles that contain  $b$  or  $c$  quarks [18]. The detector schematic is presented in Fig. 2a). Following this schematic, the various parts of the detector are described below.

The LHCb detector includes a tracking system made from a silicon-strip vertex detector surrounding the interaction point (VELO) [19], a large silicon detector (TT), a dipole magnet with a bending power of  $\sim 4$  Tm [20], and three stations of silicon-strip detectors and straw drift tubes representing the outer tracker (T1, T2, T3) [21]. Trackers enable physicists to measure the trajectory of each particle passing through the detector. The VELO, which is placed close to the beam pipe, uses silicon microstrip detectors to record the paths of particles going through it and is especially important in distinguishing  $B$  mesons from the  $pp$  interaction point, as the former fly distances of  $\sim 1$  cm. Charged particles collide with silicon atoms and thus free electrons, which in turn create an electric current that traces the path of the original particle. The outer tracker is situated further from the beam pipe and is made up of thousands of gas-filled straw tubes. When a charged particle passes through, it ionizes the gas, producing electrons. The position of the track is determined by measuring the time taken for the electrons to reach the central anode wire of their respective tube. The asymmetries in the detection efficiency due to nonuniformities in the detector response are reduced by periodically reversing the polarity of the magnet while taking data. Particles normally travel in straight lines, but the presence of a magnetic field curves the paths of charged particles, with positive and negative particle trajectories being bent in opposite directions. The curvature of the path can be analyzed to obtain the momentum of the particle, thus working towards establishing its identity. The tracking system measures the momentum  $p$  of charged particles with very high precision, giving a relative uncertainty that varies from 0.5% at low momentum to 1.0% at high  $p$ . The impact parameter (IP), which is the minimum distance of a track to its vertex of origin, is measured with a high resolution.

The charged hadrons are identified using the two ring-imaging Cherenkov (RICH1 and RICH2) detectors [22]. RICH detectors work by measuring Cherenkov radiation. This type of radiation is produced when a charged particle passes through a certain medium, in this case a dense gas, faster than light does. In its path, the particle emits a cone of light which is reflected by mirrors built into the RICH detector onto an array of sensors. The shape of the cone of light depends on the particle's velocity and thus its speed is determined. This additional bit of information is then combined with the particle trajectory, recorded using the tracking system, to determine its mass, its charge, and finally, its identity. The RICH detectors can identify the various charged particles resulting from the decay of  $B$  mesons, such as pions, kaons, and protons.

Furthermore, final state particles such as photons, electrons, and hadrons are distinguished by a calorimeter system which consists of the scintillating-pad and preshower detectors (SPD), an electromagnetic calorimeter (ECAL) and a hadronic calorimeter (HCAL). The calorimeter system is designed to stop particles as they pass through and thus measure the amount of energy lost as a particle comes to a halt. The ECAL is responsible for determining the energy of lighter particles, such as electrons and photons, while the HCAL measures the energy of protons, neutrons, and other particles that contain quarks. Both calorimeters consist of alternating layers of metal and plastic plates. Relatively heavy particles produce showers of secondary particles as they hit these metal plates. In turn, the showers excite molecules in the plastic plates, which then emit ultraviolet (UV) light. The amount of UV produced is proportional to the energy of the particles that entered the calorimeter system. Calorimeters are crucial in identifying particles that possess no electrical charge, such as photons and neutrons.

Muons are present in the final stages of many  $B$  meson decays and thus their detection is important for LHCb. Muons are measured by a system of alternating layers of iron and multiwire proportional chambers (M) [23]. Located at the far end of the detector, the muon system includes five rectangular stations, which increase in size. Each station has chambers filled with a three gas mixture; muons ionize the gases and electrodes record the products. This works on the same principle as the outer tracker.

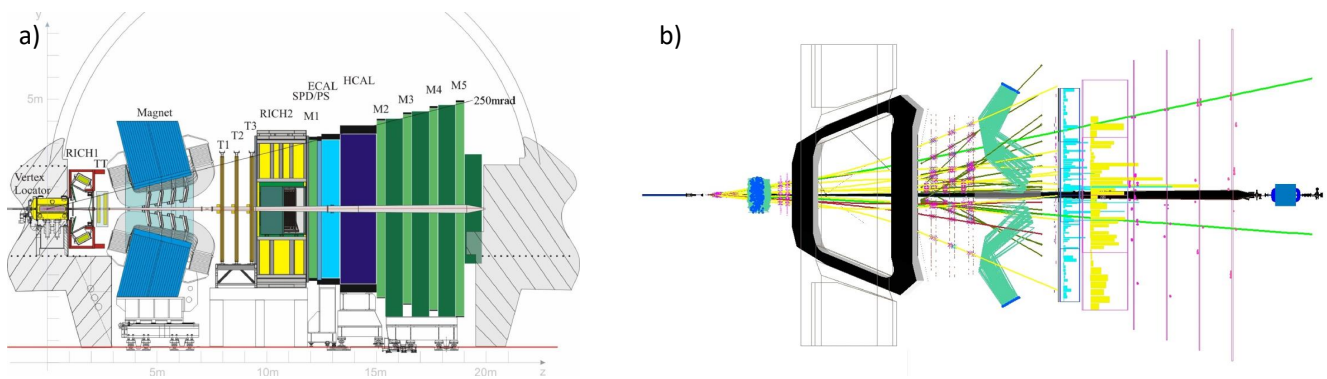


FIG. 2. Schematics of a) the LHCb detector [15] and b) a  $B_s^0$  meson decaying leptonically into a muon pair at the LHCb detector. Under the LHCb coordinate system, the  $z$  axis is along the beam pipe with the  $pp$  interaction point on the left-hand side. The  $y$  axis points from the floor of the cavern up and the  $x$  axis is perpendicular to the plane of the figure. The muons in b) are represented by the straight green lines. Of course, only specific hadronic decays are of interest in this work; however, the leptonic decay shown in this figure is instructive of how particles are reconstructed at LHCb.

Once a signal is recorded in the physical detector, a selection is applied by a trigger system while still actively taking data [24]. The LHCb trigger is essential in selecting the collisions of interest for  $B$  meson studies from all the recorded events. The trigger is based on a two level system and exploits the fact that hadrons containing the  $b$  quark are relatively heavy and long lived. The selection consists of a hardware stage, based on the calorimeter and muon system responses, and a software stage, which facilitates the complete event reconstruction. A full event reconstruction culminates as exemplified in Fig. 2b), and the decay data is saved for further analysis. To pass the hardware stage, events require a high transverse momentum ( $p_T$ ) muon to be detected in the muon chambers or a high energy hadron, photon, or electron to be detected in the HCAL or in the ECAL. The software trigger requires the existence of an up to four-tracks secondary decay vertex with a high  $p_T$  sum and a significant displacement from the primary IP. Thus, at least one track should possess a  $p_T$  of about 1 GeV and a relatively high  $\chi_{IP}^2$ , where  $\chi_{IP}^2$  is the difference in the  $\chi^2$  of a primary vertex fit that has original particle reconstruction being applied and a fit that was performed without using the reconstruction information.

As mentioned in the introduction, the processes of interest are  $B^+ \rightarrow D^{(*)+} \bar{D}^0$  and  $B^+ \rightarrow D_s^+ \bar{D}^0$ , with the subsequent decays:  $\bar{D}^0 \rightarrow K^+ \pi^-$ ,  $\bar{D}^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$ ,  $D^+ \rightarrow K^- \pi^+ \pi^+$ ,  $D_s^+ \rightarrow K^+ K^- \pi^+$ , and  $D^{*+} \rightarrow D^0 \pi_s^+$ . These subdecays are reconstructed in the LHCb detector as outlined in the previous paragraph; a more in-depth explanation of the trigger system in the context of the studied decays is found in Ref. [16] and is beyond the scope here. After passing the trigger selection, the event is stored on disk and is in principle ready for offline analysis. However, other selection cuts are often made before an analysis is performed on the data. Thus, a broad selection is applied on the raw trigger information, to construct a proper data set on which, say, a  $CP$  asymmetry measurement can be performed.

The data samples for this analysis are the refined Run I and Run II data sets collected by the LHCb experiment. This refinement was performed in two steps. First, a centralised preselection was used, also known as stripping. The stripped [25] data sets were obtained by applying a number of loose requirements to select the particular decay channels of interest. Then, a momentum scale calibration was applied [26] to account for the possible momentum measurement bias due to imperfect knowledge of the magnetic field in the tracker. Secondly, a further preselection, detailed in Ch. 4.6 of Ref. [16], is performed to reduce the background levels with little loss of signal.

Simulated particle decay samples, known as Monte Carlo data sets, were quintessential for this measurement, since they were used for the training of the multivariate selection algorithm, and for constraining the signal shapes of the invariant mass fits. These simulations model  $pp$  collisions in which the studied decays were then generated. All known issues regarding the simulation, including the mass scale, momentum resolution, and the RICH response, were corrected through data-driven methods [27]. The MC data was compared to the actual data and validated [16].

## II. EVENT SELECTION

The event selection consists of three parts, as follows. First, a list of potentially useful variables is constructed, and the shapes of the MC and data distributions are overlaid for comparison. Thus, the variables for which the signal shape can be easily discriminated from the background are picked. Then, a Boosted Decision Tree (BDT) is trained on the chosen variable combinations, removing some variables from the training set in an iterative process. Finally, a set of definite training variables is established, and a cut which maximizes the signal to background yield is applied on the corresponding BDT response.

Variable	Formula/Shorthand	Brief Description
Invariant Mass	$m = \sqrt{\left(\sum_i^N E_i\right)^2 - \left(\sum_i^N \mathbf{p}_i\right)^2}$	Peaks at the decaying particle mass for system of decay particles. Combinatorial backgrounds will have a smooth distribution.
Transverse Momentum	$p_T = \sqrt{p_x^2 + p_y^2}$	The component of a particle's momentum, $\perp$ to beam direction. The high $B$ mass usually results in high $p_T$ decay products.
Decay Time	$\tau = L/\gamma v$	The time between the production of a particle and its decay. $L$ is the travel distance, $\gamma$ the lorentz factor and $v$ the velocity.
Decay Time Significance	$\tau/\Delta\tau$	The decay time divided by its uncertainty. Useful in distinguishing particles with large decay time uncertainties.
Impact Parameter	IP	The distance of closest approach between the primary vertex (PV) and the extrapolated particle trajectory.
Impact Parameter $\chi^2$	$\text{IP}\chi^2$	The difference between the fit quality of the PV reconstructed with and without the particle under consideration.
Vertex $\chi^2$	$V\chi^2$	The $\chi^2$ of the vertex fit.
Log likelihood difference	$\text{DLL}_{K\pi}$	The difference in likelihood of the $\pi$ and the $K$ mass hypotheses. A large $\text{DLL}_{K\pi}$ value means that the particle is likely a kaon.
Dalitz Mass	$m^2(\pi K)$	The invariant mass of a $\pi$ and $K$ in a three body decay of a $D$ meson.

TABLE I. The main particle properties used throughout the analysis. These properties apply to the  $B^\pm$  mesons,  $D$  mesons, or the final decay products represented by  $\pi$ s and  $K$ s. Naturally, not every particle in the studied decay chains pertains to all the variables described here, e.g., the decay time of the  $\pi$  meson is not recorded since it is a final state particle.

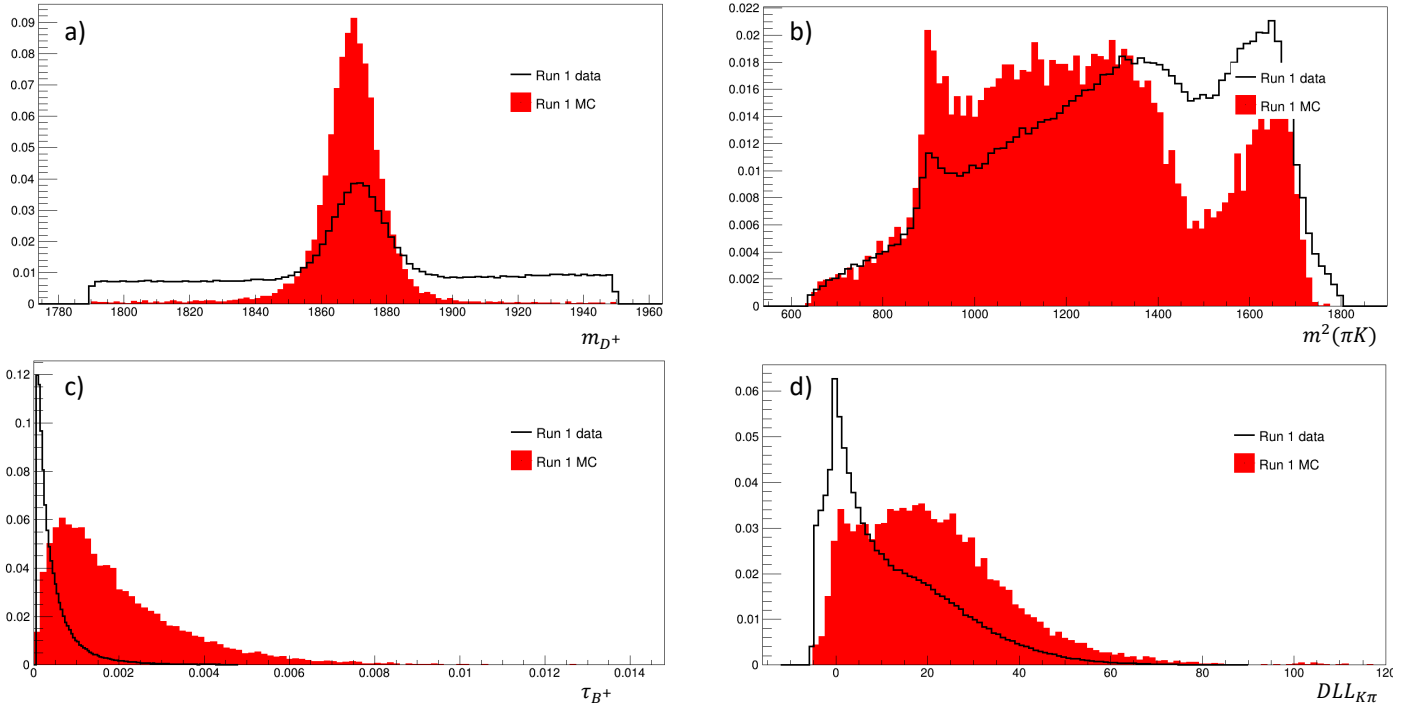


FIG. 3. Four preliminary histograms, plotted for the decay  $B^+ \rightarrow D^+ \bar{D}^0$ , with final states  $D^+ \rightarrow \pi^+ \pi^+ K^-$  and  $\bar{D}^0 \rightarrow K^+ \pi^-$ . Short descriptions of the illustrated  $x$ -axis variables are available above in Tab. 1; the  $y$ -axis is the normalised number of events. The subfigures show a) The invariant mass of the  $D^+$  meson, sharply peaked in the MC and broader in the real data, as expected, b) The Dalitz mass between the  $\pi^+$  and  $K^-$  in the  $D^+$  decay, c) The lifetime of the  $B^+$ , sharply peaked in the data at lower values due to short-lived background particles, and d) The  $\text{DLL}_{K\pi}$  of the  $K$  meson, where the MC is again peaked at higher values, since a real kaon corresponds to a higher  $\text{DLL}_{K\pi}$ . All of the variable distributions shown above passed the preliminary variable selection. However, these four figures represent only a small subset of the initial MVA training variable set. All the preliminary histograms, together with the initial training set for each decay chain, can be found on the [website](#) complementing this paper.



### A. Preliminary Discriminating Variable Selection

A large array of potentially useful variables are selected from the refined analysis data samples described in [Sec. I](#). This initial selection is conducted based on general physical considerations of what would constitute the best training set for a multivariate analysis algorithm (MVA). The main particle properties considered in the selection are in [Tab. 1](#). Most, but not all of them are applicable for every particle in the studied decay chains.

Thus, if a particle contains one of the properties listed in [Tab. 1](#), the distribution of that variable is plotted for both real data and the simulation. The two resulting histograms are then normalized with respect to their total number of events and overlaid. The obtained plots are used to determine if the variable under consideration is indeed a good candidate for MVA training. If the shape of the signal MC sample distribution is quite distinct from the real data sample shape, then the variable is deemed as good; if they match, the variable is removed and the BDT will not be trained on it. This is to ensure that the MVA learns how to discriminate between signal and background with high efficiency. Some example plots corresponding to the  $B^+ \rightarrow D^+ \bar{D}^0$  decay channel are shown in [Fig. 3](#).

### B. Multivariate Analysis

To reduce the combinatorial background, while maximizing the ratio of signal efficiency to background rejection, a multivariate selection based on a BDT [\[28\]](#) is applied. Optimizing this BDT represents the main body of work for this analysis. The optimization process is described in this section. However, an introduction to MVAs is given first.

Generally, MVA algorithms take as input a signal sample, a background sample, and a set of discerning variables. The MVA classifies the variable space as signal or background, depending on the properties of the respective data. MVAs can exploit correlations between input variables and thus perform better than the traditional method of applying rectangular cuts [\[16\]](#). The output of the MVA is a single variable, denoted  $BDT\_G$  in the plots, that represents how signal-like a certain event is. A decision tree has a structure that is very similar to a flowchart [\[29\]](#). The first decision node receives the full training sample and applies a requirement on one of the considered discriminating variables; the applied cut gives the best signal-background separation. Subsequently, two sub-samples are created, which are used as inputs for the further nodes. The process continues until either the maximum number of requirements are applied, as set by the user, or there are no events left. The resulting nodes are thus classified as signal/background, depending on the main contribution. In case an event is misclassified by a decision tree, it is assigned a higher priority to ensure that the next decision tree is more likely to perform the classification properly. This process was named boosting and thus the term Boosted Decision Trees was coined to denote this type of MVA.

The training and reweighting of the training sample is repeated until a decision forest is obtained, which is subsequently combined into a single classifier by taking the weighted average of the forest. Due to being based on simple 1D optimisation, BDTs are insensitive to the inclusion of poorly discriminating input variables. As a caveat, this may limit their relative performance when compared to more sophisticated MVA methods [\[16\]](#). A total of 16 BDTs were trained and used for the studied decay modes.

[Sec. IIa](#) describes the process through which initial sets of variables are selected to train each BDT. However, listing all these variables here would over-saturate the text. A complete inventory of the initial training variables is available on this paper's [website](#), under [ML training](#). The [Sec. IIa](#) variable sets are refined through an iterative process, guided by two factors. First, the final training set should contain only variables that are minimally correlated, as to mitigate over-training [\[29\]](#); the over-training is shown through figures such as [Fig. 4c](#)). The correlation between different quantities is illustrated through correlation matrices, illustrated in [Fig. 4b](#)). Secondly, the ratio between signal efficiency and background rejection should be high, since this will lead to a higher precision in the  $CP$  asymmetry measurement. This efficiency is determined by the so called ROC curves, as exemplified in [Fig. 4a](#)).

The two contributions counter-balance one another. If a lot of variables are eliminated, as to guarantee minimum overtraining, the signal efficiency over the background rejection decreases; however, a high ratio is obtained through the inclusion of many variables, some of which are bound to have high correlation coefficients and thus produce an overtrained BDT. Therefore, each of the 16 BDTs were trained using 15 distinct variable mixtures. All the resulting correlation matrices, ROC curves, and overtraining figures ( $\sim 100$  in total) are on the [website](#). By performing an overall comparison of the plots, obtained for each of the 16 BDTs, it was determined that the best training variables are the ones illustrated in [Tab. 2](#). Furthermore, the input variables undergo an additional procedure to reduce correlations, in the form of training data transformations. These are described in [Fig. 4](#).

Finally, the  $k$ -folding technique is used to mitigate the eventual performance issues and eliminate biases that small data sets might experience in training the BDT by increasing the overall statistics. Withholding a subset of the signal and background samples is quintessential for post-training performance validation. Ideally, the size of the training and testing samples is the same. However, as stated at the start of the paragraph, this might hinder the performance of small data sets. To avoid this possible complication,  $k = 5$  BDTs are trained for each variable combination in each decay chain, while withholding a different 20% of the sample for testing every time. Thus, the performance validation is done by evaluating the BDT on its excluded data. This procedure allows the majority of the signal and background samples to be used. Finally, the five different BDT responses are incorporated into the same analysis data set.

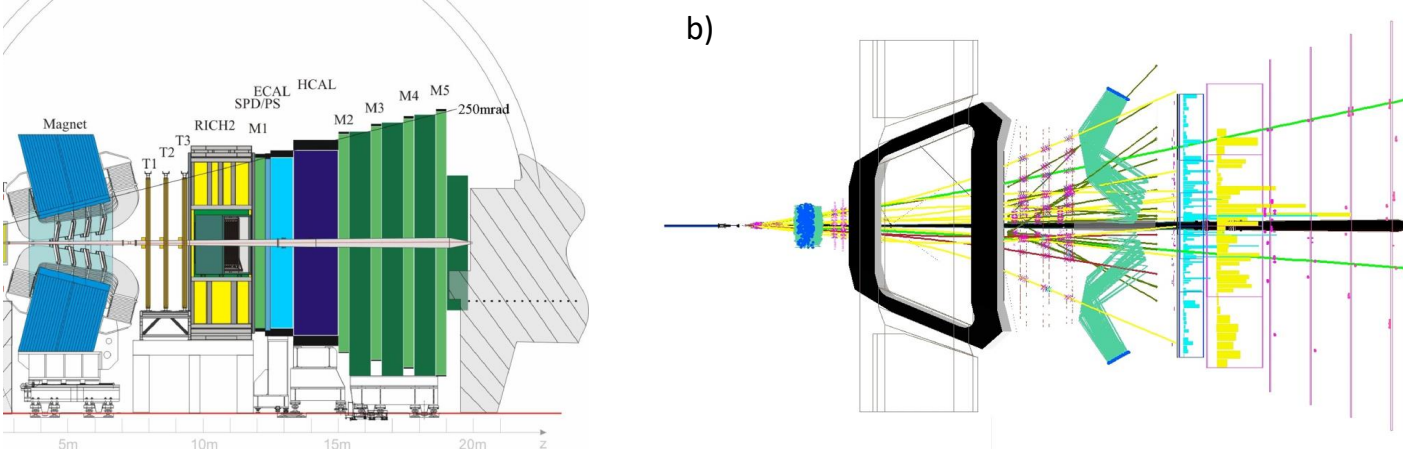


FIG. 4. The selection process of the optimal set of discriminating variables, involving mainly three figures for each variable set: a) The signal efficiency of the BDT, b) The correlation matrix, and c) The overtraining of the BDT, with a Gaussian transformation applied to the data samples. Notice that for each set of variables, 5 different BDTs are trained on distinct transformations, in a). These are denoted: no transformation (I), uniformisation (U), gaussian transformation (G), variable decorrelation (D), and principal component decomposition (P). Applying two of these transformations, e.g., gaussian and decorrelation, is represented with GD. The complete description of each of these transformations is found in Ref. [29]. The presented plots are of the 1st decay in Tab. 2. The arrows indicate the process of selecting a particular variable training set. The full set of plots can be found on the [website](#).

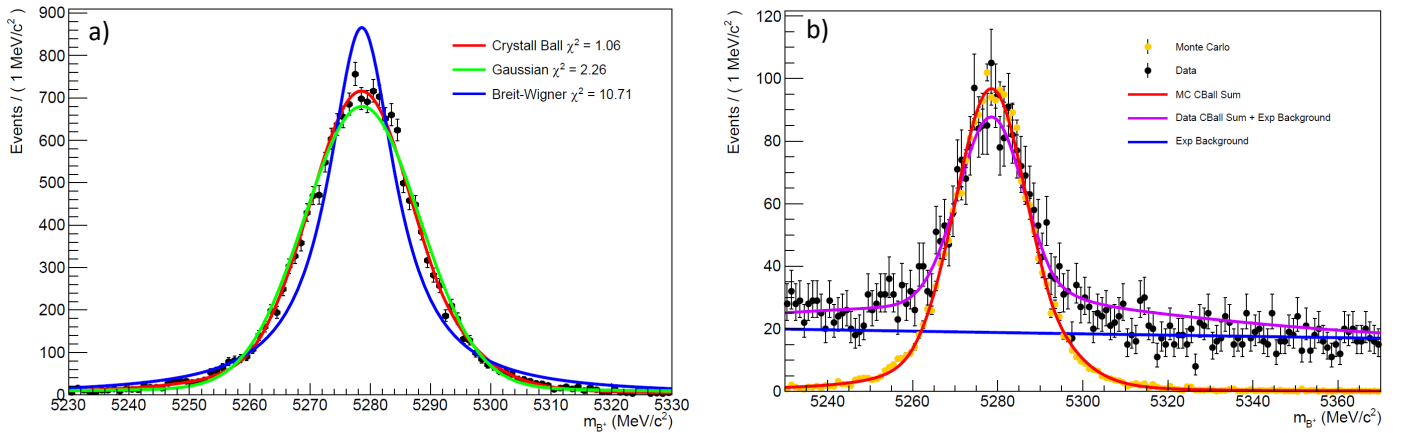


FIG. 5. The fitting of the data and the MC for the 1st decay listed in Tab. 2: a) The best fit model determination in MC and b) The full fit with no BDT cuts, for the real Run I data and overlaid MC. Notice that in a), the Crystall Ball gives the best goodness of fit parameter ( $\chi^2$ ), while the Breit-Wigner [30] performs the worst. In b), the MC is scaled to facilitate the overlaying. The full set of fitting plots can be found on the accompanying [website](#).

Main Decay	Charged Meson	Neutral Meson	Discriminating Variables
$B^+ \rightarrow D^+ \bar{D}^0$	$D_{(s)}^+ \rightarrow K^- \pi^+ \pi^+$	$\bar{D}^0 \rightarrow K^+ \pi^-$	$\tau/\Delta\tau, V\chi^2, m_{D^+}, m_{\bar{D}^0}, \max(m^2(\pi K)), p_T, DLL_{K\pi}$
		$\bar{D}^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$	
$B^+ \rightarrow D_s^+ \bar{D}^0$	$D_{(s)}^+ \rightarrow K^- \pi^+ \pi^+$	$\bar{D}^0 \rightarrow K^+ \pi^-$	$\tau/\Delta\tau, V\chi_{B^+}^2, m_{D_s^+}, m_{\bar{D}^0}, m^2(\pi K), p_T, DLL_{K\pi}$
		$\bar{D}^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$	
$B^+ \rightarrow D^{*+} \bar{D}^0$	$D^{*+} \rightarrow D^0 \pi_s^+$	$\bar{D}^0 \rightarrow K^+ \pi^-$	$\tau/\Delta\tau, V\chi^2, \Delta m_{D^{*+}}, m_{\bar{D}^0}, m_{D^0}, p_T, DLL_{K\pi}$
	$D^0 \rightarrow \pi^+ K^-$	$\bar{D}^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$	
	$D^{*+} \rightarrow D^0 \pi_s^+$	$\bar{D}^0 \rightarrow K^+ \pi^-$	$\tau/\Delta\tau, V\chi^2, \Delta m_{D^{*+}}, m_{\bar{D}^0}, m_{D^0}, m^2(\pi K), p_T, DLL_{K\pi}$
	$D^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$	$\bar{D}^0 \rightarrow K^+ \pi^- \pi^+ \pi^-$	

TABLE II. The selected BDT discriminating variables for all the studied decay chains. These sets of variables are obtained after the refinement procedure described in Sec. IIb. The meaning of the presented variables is outlined in Tab. 1. The  $p_T$  and  $DLL_{K\pi}$  variables are included for all the final state particles, in every decay chain. The  $V\chi^2$  variable is selected for all the mesons in a decay chain, unless stated otherwise. Throughout the last four decay chains, all variables related to the  $D^{*+}$  meson are ignored, except for  $\Delta m_{D^{*+}}$ . The  $\Delta m_{D^{*+}}$  variable is the reconstructed mass difference of the  $\pi_s$  and  $D^0$  coming from the  $D^{*+}$  decay.

### C. Fitting the Spectra

A cut must be applied at a certain value of the BDT response variable to obtain maximum signal yield while minimizing the background. This is established by doing a series of fits on the MC and real data  $m_{B^+}$  spectra.

The functional shape of the signal is determined by fitting three different models to the signal MC: a sum of Crystal Ball (CB) distributions [31], a sum of Gaussians, and a Breit-Wigner distribution [30]. The motivation for fitting a Breit-Wigner comes from theory, as this particular distribution models the resonance curve produced by a particle, in this case the  $B$  meson. However, it is found that it performs the worst due to detector resolution effects that appear in the MC simulation. Conversely, the CB model gives the best  $\chi^2$  when compared to the other two, and thus it is used in fitting the data spectrum. The Gaussian sum lies in between the other two fits in terms of  $\chi^2$  value. An example of all three fits is shown in Fig. 5a); the complete set of these fits, performed in this analysis, can be found on the paper's website. For some of the studied decay chains, presented in Tab. 2, the Crystal Ball  $\chi^2$  and the Gaussian fit  $\chi^2$  were quite close. Thus, for these particular modes, it was considered using the Gaussian sum since it contains fewer parameters. However, the Crystal Ball sum was ultimately used for all the decays because the added complexity of using an inhomogeneous fitting paradigm would lead to more complications than benefits.

Prior to any data fitting, events are selected to contain at least one charged  $B$  meson decaying as in Tab. 2. Thus, the amount of data to be fitted is reduced considerably and the  $B$  meson mass resonance is visible even without applying cuts. The data sample is then fitted using a composite model of the previously determined Crystal Ball sum and an exponential. The Crystal Ball parameters are mostly fixed to be the ones obtained from fitting the corresponding Monte Carlo sample. The sole exception is the width which is allowed to vary slightly due to the MC being narrower. The exponential was chosen to fit the combinatorial background, which looks like fake signal.

The result of this fit is illustrated in Fig. 5b); again, the complete set of such fits is available on the paper's website. Overlaid on top of the data signal is the fitted Monte Carlo simulation, scaled for comparison purposes. Notice that the CB sum fits the data spectrum quite well, which means that the fitting procedure was successful. Additionally, the exponential is plotted separately and shows the amount of combinatorial background present in the data sample.

Both the exponential background and the signal CB sum are integrated in the range of 5230 – 5305 MeV/ $c^2$ . These integration limits on the mass were chosen based on the Monte Carlo plot, as they span the full signal range. The exponential integral gives the background yield in the data, denoted by  $B$ . Alternatively, the integral of the data Crystal Ball sum gives the signal yield, denoted by  $S$ . The significance  $f$  of the signal is defined as

$$f = \frac{S}{\sqrt{S+B}} \quad (2)$$

and is calculated using the previously obtained values.

The value of  $f$  is used as a metric to determine the particular cut that must be applied on the BDT response to obtain a maximum amount of signal and a minimum amount of background. The BDT response, obtained in Sec. IIb, lies in the range  $-0.5 < BDT\_G < 0.5$ . In Fig. 5b), no BDT cuts were applied, which is equivalent with applying a  $> -0.5$  cut. The optimal BDT cut is obtained iteratively. A total of 300 equally spaced values are taken from the BDT response interval. For each of these values, the data sample is cut such that no event has a lower BDT response than the considered one. The previously described fitting process is performed on the obtained data sample, but all the parameters, except the yields, are completely fixed on the values that were obtained from the initial  $> -0.5$  BDT cut fit. The significance  $f$  is calculated for each of the cut samples and plotted. The results are shown in Fig. 6. The maximum  $f$  and the corresponding BDT cut is determined for each decay chain, marked by a red dot in the mentioned figure. Each obtained BDT cut is ultimately applied on the data sample prior to computing the corresponding raw  $CP$  asymmetry.

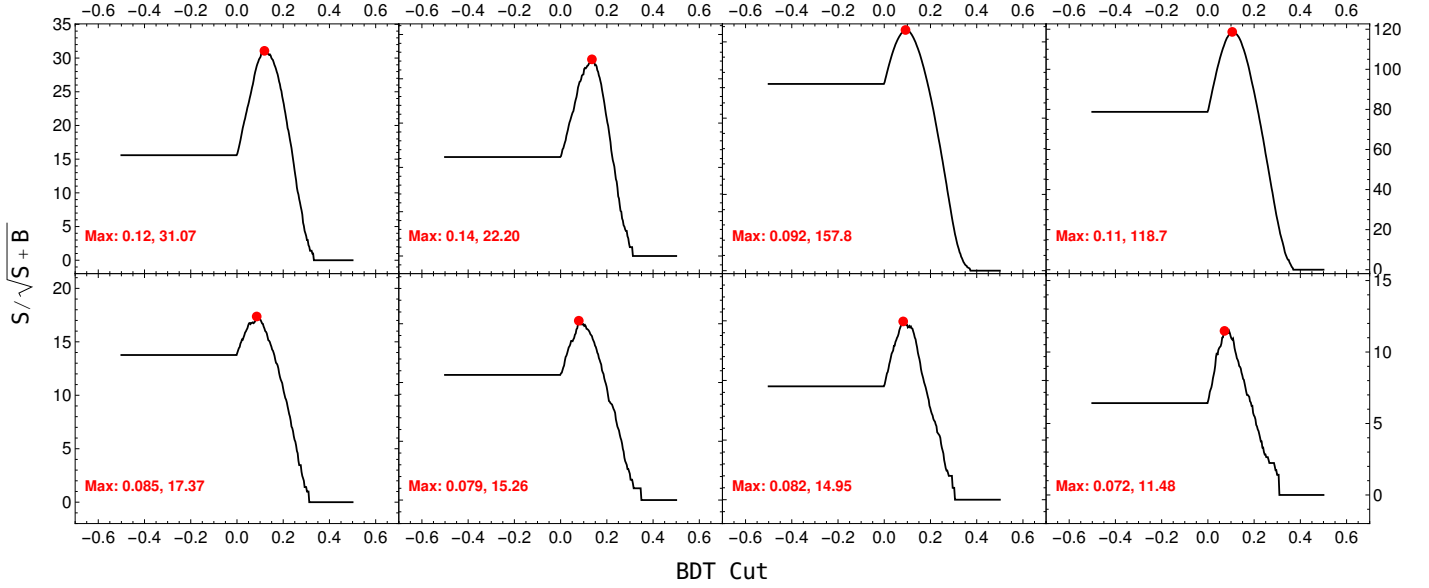


FIG. 6. The figure of merit plotted for 300 equally spaced BDT cuts in the range -0.5 to 0.5, for each of the studied decay chains. The BDT cut corresponding to a maximum signal yield and minimum background yield is denoted by the red dot and is shown on the bottom left of each figure, along with the produced signal significance, in this order. Notice that at high BDT cuts there is no signal left in the sample and the significance is 0. The presented plot is for Run I data. The combined runs plot is on the [website](#). The left hand  $y$ -axis applies to the first two plots on each respective row, while the right hand one corresponds to the last two.

### III. RESULTS

The optimal BDT cut, which maximizes the figure of merit, is determined for each decay as shown in Fig. 6. This cut is then applied on the data sample. The yield of reconstructed  $B^+$  and  $B^-$  mesons is then determined from fitting the data spectrum one more time and the raw asymmetry is calculated using Eq. (1).

As stated in the introduction, the  $CP$  decay asymmetry is defined as  $\mathcal{A}^{CP} = A_{\text{raw}} - A_P - A_D$ , up to first order. In this paper, only the raw asymmetry is determined. The production asymmetry and the detection asymmetry values are taken from literature. Thus, the detection asymmetry used here is the one determined in the previous measurement [27], while the production asymmetry is taken from Ref. [32]. The values of  $A_D + A_P$  for the  $B^+ \rightarrow D_s^+ \bar{D}^0$  mode is  $(-1.4 \pm 0.5)\%$ , while for  $B^+ \rightarrow D^+ \bar{D}^0$  it is  $(-0.3 \pm 0.4)\%$ .

The three asymmetries are combined. The results are

$$\begin{aligned}\mathcal{A}^{CP}(B^+ \rightarrow D^+ \bar{D}^0) &= (0.3 \pm 2.3 \pm 0.4)\%, \\ \mathcal{A}^{CP}(B^+ \rightarrow D_s^+ \bar{D}^0) &= (0.4 \pm 0.5 \pm 0.5)\%, \\ \mathcal{A}^{CP}(B^+ \rightarrow D^{*+} \bar{D}^0) &= (-2.6 \pm 3.2 \pm 0.4)\%\end{aligned}$$

for Run I only and

$$\begin{aligned}\mathcal{A}^{CP}(B^+ \rightarrow D^+ \bar{D}^0) &= (0.7 \pm 1.0 \pm 0.4)\%, \\ \mathcal{A}^{CP}(B^+ \rightarrow D_s^+ \bar{D}^0) &= (-2.6 \pm 0.2 \pm 0.5)\%, \\ \mathcal{A}^{CP}(B^+ \rightarrow D^{*+} \bar{D}^0) &= (1.6 \pm 1.5 \pm 0.4)\%\end{aligned}$$

for both runs combined, where the uncertainties are statistical and systematic, in this order. Only the  $D_s^+$  mode of the combined run shows signs of  $CP$  violation.

The same detection and production asymmetry values were used in the combined Run I and II data for simplicity. They would need to be computed separately for Run II and then combined accordingly with the Run I values to make a more definitive claim about the observed  $CP$  violation. However, this calculation is outside the scope of this work. Furthermore, the same  $A_P + A_D$  was used for the  $D^+$  and  $D^{*+}$  decay modes since they present matching final states.

### CONCLUSION

The  $CP$  asymmetries of the  $B^+ \rightarrow D^{(*)+} \bar{D}^0$  and the  $B^+ \rightarrow D_s^+ \bar{D}^0$  decays were measured using a multivariate analysis method. A set of potential training variables was selected, shown in Tab. 1, together with examples of their distributions displayed in Fig. 3. This initial set was then refined as exemplified in Fig. 4 and a Boosted Decision Tree was trained on the refined variables, for each of the studied decay channels presented in Tab. 2. The BDT response was then analyzed by fitting the spectrum as shown in Fig. 5. The optimal BDT cut was determined as illustrated in Fig. 6 for Run I data. Finally, these BDT cuts were applied on their respective decays and the corresponding  $CP$  asymmetry was calculated.

The obtained values do not show evidence of  $CP$  violation, except for the  $D_s^+$  mode which hints towards it. Moreover, the statistical uncertainties of all the measured decay modes were improved when compared to the previous Run I measurements [27], especially the  $D^{*+}$  mode [14]. By using the combined run data, the statistical uncertainty of the performed measurements was even further improved. The presented values could be refined more by independently determining the  $A_P + A_D$  value for the Run II data.



- 
- [1] Nicola Cabibbo, “Leptonic Decays in the Unitary Symmetry Scheme,” in *International Conference on Fundamental Aspects of Weak Interactions* (1964) pp. 299–302.
- [2] Makoto Kobayashi and Toshihide Maskawa, “CP Violation in the Renormalizable Theory of Weak Interaction,” *Prog. Theor. Phys.* **49**, 652–657 (1973).
- [3] Wei-Shu Hou, “Source of CP Violation for the Baryon Asymmetry of the Universe,” *Chin. J. Phys.* **47**, 134 (2009), [arXiv:0803.1234 \[hep-ph\]](#).
- [4] Run-Hui Li *et al.*, “Decays of  $B$  meson to two charmed mesons,” *Phys. Rev. D* **81**, 034006 (2010), [arXiv:0910.1424 \[hep-ph\]](#).
- [5] Hui-Feng Fu *et al.*, “Semi-leptonic and non-leptonic  $B$  meson decays to charmed mesons,” *Chin. Phys. Lett.* **28**, 121301 (2011), [arXiv:1202.1221 \[hep-ph\]](#).
- [6] Lin-Xia Lu *et al.*, “The Double charm decays of  $B$  Mesons in the mSUGRA model,” *Commun. Theor. Phys.* **56**, 125–133 (2011), [arXiv:1008.4987 \[hep-ph\]](#).
- [7] C.S. Kim *et al.*, “Studying Double Charm Decays of  $B(u,d)$  and  $B(s)$  Mesons in the MSSM with R-parity Violation,” *Phys. Rev. D* **79**, 055004 (2009), [arXiv:0812.4136 \[hep-ph\]](#).
- [8] Yuan-Guo Xu and Ru-Min Wang, “Studying the fourth generation quark contributions to the double charm decays  $B_{(s)} \rightarrow D_{(s)}^{(*)} D_s^{(*)}$ ,” *Int. J. Theor. Phys.* **55**, 5290–5306 (2016).
- [9] J.H. Christenson *et al.*, “Evidence for the  $2\pi$  Decay of the  $K_2^0$  Meson,” *Phys. Rev. Lett.* **13**, 138–140 (1964).
- [10] Bernard Aubert *et al.* (BaBar), “Observation of CP violation in the  $B^0$  meson system,” *Phys. Rev. Lett.* **87**, 091801 (2001), [arXiv:hep-ex/0107013](#).
- [11] Kazuo Abe *et al.* (Belle), “Observation of large CP violation in the neutral  $B$  meson system,” *Phys. Rev. Lett.* **87**, 091802 (2001), [arXiv:hep-ex/0107061](#).
- [12] Y. Chao *et al.* (Belle), “Evidence for direct CP violation in  $B^0 \rightarrow K^+ \pi^-$  decays,” *Phys. Rev. Lett.* **93**, 191802 (2004), [arXiv:hep-ex/0408100](#).
- [13] R. Aaij *et al.* (LHCb), “Observation of CP violation in  $B^\pm \rightarrow DK^\pm$  decays,” *Phys. Lett. B* **712**, 203–212 (2012), [arXiv:1203.3662 \[hep-ex\]](#).
- [14] Bernard Aubert *et al.* (BaBar), “Measurement of branching fractions and CP-violating charge asymmetries for  $B$  meson decays to  $D^*$  anti- $D^*$ ,” *Phys. Rev. D* **73**, 112004 (2006), [arXiv:hep-ex/0604037](#).
- [15] Jr. Alves, A. Augusto *et al.* (LHCb), “The LHCb Detector at the LHC,” *JINST* **3**, S08005 (2008).
- [16] Alison Maria Tully, *Doubly charmed  $B$  decays with the LHCb experiment*, Ph.D. thesis, Cambridge U. (2019).
- [17] Roel Aaij *et al.* (LHCb), “LHCb Detector Performance,” *Int. J. Mod. Phys. A* **30**, 1530022 (2015), [arXiv:1412.6352 \[hep-ex\]](#).
- [18] H. Dijkstra *et al.*, “LHCb Letter of Intent, LHCb Collaboration,” (1995).
- [19] P. Sievers (LHCb), “Performance tests of large pitch silicon strip sensors for the LHCb inner tracker,” *Nucl. Instrum. Meth. A* **512**, 245–249 (2003).
- [20] J. André *et al.*, “Status of the LHCb dipole magnet,” *IEEE Trans. Appl. Supercond.* **14**, 509–513 (2004).
- [21] R Arink *et al.* (LHCb Outer Tracker Group), “Performance of the LHCb Outer Tracker,” *JINST* **9**, P01002 (2014), [arXiv:1311.3893 \[physics.ins-det\]](#).
- [22] M. Adinolfi *et al.* (LHCb RICH Group), “Performance of the LHCb RICH detector at the LHC,” *Eur. Phys. J. C* **73**, 2431 (2013), [arXiv:1211.6759 \[physics.ins-det\]](#).
- [23] Jr. Alves, A.A. *et al.*, “Performance of the LHCb muon system,” *JINST* **8**, P02022 (2013), [arXiv:1211.1346 \[physics.ins-det\]](#).
- [24] Barbara Sciascia (LHCb), “LHCb Run 2 trigger performance,” *PoS BEAUTY2016*, 029 (2016).
- [25] F. Stagni *et al.*, “LHCbDirac: Distributed computing in LHCb,” *J. Phys. Conf. Ser.* **396**, 032104 (2012).
- [26] R Aaij *et al.* (LHCb), “Measurement of the  $\Lambda_b^0$ ,  $\Xi_b^-$  and  $\Omega_b^-$  baryon masses,” *Phys. Rev. Lett.* **110**, 182001 (2013), [arXiv:1302.1072 \[hep-ex\]](#).
- [27] R. Aaij *et al.* (LHCb), “Measurement of the  $CP$  asymmetry in  $B^- \rightarrow D_s^- D^0$  and  $B^- \rightarrow D^- D^0$  decays,” *JHEP* **05**, 160 (2018), [arXiv:1803.10990 \[hep-ex\]](#).
- [28] Byron P. Roe *et al.*, “Boosted decision trees, an alternative to artificial neural networks,” *Nucl. Instrum. Meth. A* **543**, 577–584 (2005), [arXiv:physics/0408124](#).
- [29] Andreas Hocker *et al.*, “TMVA - Toolkit for Multivariate Data Analysis,” (2007), [arXiv:physics/0703039](#).
- [30] G. Breit and E. Wigner, “Capture of Slow Neutrons,” *Phys. Rev.* **49**, 519–531 (1936).
- [31] Diego Martínez Santos and Frederic Dupertuis, “Mass distributions marginalized over per-event errors,” *Nucl. Instrum. Meth. A* **764**, 150–155 (2014), [arXiv:1312.5000 \[hep-ex\]](#).
- [32] Roel Aaij *et al.* (LHCb), “Measurement of the  $B^\pm$  production asymmetry and the  $CP$  asymmetry in  $B^\pm \rightarrow J/\psi K^\pm$  decays,” *Phys. Rev. D* **95**, 052005 (2017), [arXiv:1701.05501 \[hep-ex\]](#).