1. Introduction

In the Standard Model (SM) charge-parity (CP) violation in the charm sector is expected to be small. Quantitative predictions of CP asymmetries are difficult, since the computation of strong-interaction effects in the non-perturbative regime is involved. In spite of this, it was commonly assumed that the observation of asymmetries of the order of 1% in charm decays would be an indication of new sources of CP violation (CPV). Recent studies, however, suggest that CP asymmetries of this magnitude could still be accommodated within the SM [1–4].

Experimentally, the sensitivity for CPV searches has substantially increased over the past few years. Especially with the advent of the large LHCb data set, CP asymmetries at the O(10⁻²) level are disfavoured [5–9]. With uncertainties approaching O(10⁻³), the current CPV searches start to probe the regime of the SM expectations.

The most simple and direct technique for CPV searches is the computation of an asymmetry between the particle and antiparticle time-integrated decay rates. A single number, however, may not be sufficient for a comprehension of the nature of the CP violating asymmetry. In this context, three- and four-body decays benefit from rich resonance structures with interfering amplitudes modulated by strong-phase variations across the phase space. Searches for localised asymmetries can bring complementary information on the nature of the CPV.

In this Letter, a search for CP violation in the Cabibbo-suppressed decay \( D^+ \rightarrow \pi^- \pi^+ \pi^+ \) is reported. The investigation is performed across the Dalitz plot using two model-independent techniques, a binned search as employed in previous LHCb analyses [10,11] and an unbinned search based on the nearest-neighbour method [12,13]. Possible localised charge asymmetries arising from production or detector effects are investigated using the decay \( D_s^+ \rightarrow \pi^- \pi^+ \pi^+ \), which has the same final state particles as the signal mode, as a control channel. Since it is a Cabibbo-favoured decay, with negligible loop (penguin) contributions, CP violation is not expected at any significant level.

2. LHCb detector and data set

The LHCb detector [14] is a single-arm forward spectrometer covering the pseudorapidity range 2 < \( \eta \) < 5, designed for the study of particles containing b or c quarks. The detector includes a high-precision tracking system consisting of a silicon-strip vertex detector surrounding the pp interaction region, a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about 4 Tm, and three stations of silicon-strip detectors and straw drift tubes placed downstream. The combined tracking system provides a momentum measurement with relative uncertainty that varies from 0.4% at 5 GeV/c to 0.6% at 100 GeV/c, and impact parameter (IP) resolution of 20 \( \mu \)m for tracks with high transverse momentum, \( p_T \). Charged hadrons are identified using two ring-imaging Cherenkov (RICH) detectors [15]. Photon, electron and hadron candidates are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic calorimeter and a hadronic calorimeter. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers [16]. The trigger [17] consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which applies full event reconstruction. At the hardware trigger stage, events are required to have muons with high transverse momentum or hadrons, photons or electrons with high transverse energy deposit in the calorimeters. For hadrons, the transverse energy threshold is 3.5 GeV/c².

The software trigger requires at least one good quality track from the signal decay with high \( p_T \) and high \( \chi^2_0 \), defined as the difference in \( \chi^2 \) of the primary vertex (PV) reconstructed with and without this particle. A secondary vertex is formed by three tracks...
contribution, which are required to satisfy particle identification (PID) criteria.

Figs. 1 and 2. Invariant-mass distributions for (a) $D^+$ and (b) $D_s^+$ candidates in the momentum range $50 < p_T < 100$ GeV/c for magnet up data. The solid (blue) line is the fit function, the (green) dashed line is the signal component and the (magenta) dotted line is the background.

with good quality, each not pointing to any PV, and with requirements on $p_T$, momentum $p$, scalar sum of $p_T$ of the tracks, and a significant displacement from any PV.

The data sample used in this analysis corresponds to an integrated luminosity of 1.0 fb$^{-1}$ of pp collisions at a centre-of-mass energy of 7 TeV collected by the LHCb experiment in 2011. The magnetic field polarity is reversed regularly during the data taking in order to minimise effects of charged particle and antiparticle detection asymmetries. Approximately half of the data are collected from the secondary to any primary vertex in the event, and the vertex fit quality, the minimum significance of the displacement from any PV, the minimum significance of the displacement from any PV, and the vertex fit quality, the minimum significance of the displacement from any PV, and with good quality, each not pointing to any PV, and with requirements on $p_T$, momentum $p$, scalar sum of $p_T$ of the tracks, and a significant displacement from any PV.

The data sample is separated according to magnet polarity and background, while for the control mode there are 2938 candidates in the momentum range 50–100 GeV/c, and $p_{D_s^+} > 100$ GeV/c, to take into account the dependence of the mass resolution on the momentum. The parameters are determined by simultaneous fits to these $D_s^+$ and $D_s^0$ subsamples.

The $D^+$ and $D_s^+$ invariant mass distributions and fit results for the momentum range 50 < $p_{D_s^+}$ < 100 GeV/c are shown in Fig. 1 for magnet up data. The total yields after summing over all fits are (2678 ± 7) × 10$^3$ $D^+ → π^−π^+π^+$ and (2704 ± 8) × 10$^3$ $D_s^+ → π^−π^+π^+$ decays. The final samples used for the CPV search consist of all candidates with $M(π^−π^+π^+)$ within ±2σ around $m_{D_s}$, where σ and $m_{D_s}$ are the weighted average of the two fitted Gaussian widths and mean values. The values of σ range from 8 to 12 MeV/c$^2$, depending on the momentum region. For the signal sample there are 3114 × 10$^3$ candidates, including background, while for the control mode there are 2938 × 10$^3$ candidates with purities of 82% and 87%, respectively. The purity is defined as the fraction of signal decays in this mass range.

Fits to the invariant mass distribution $M(π^−π^+π^+)$ are performed for the $D^+$ and $D_s^+$ candidates satisfying the above selection criteria and within the range $1810 < M(π^−π^+π^+) < 1930$ MeV/c$^2$ and $1910 < M(π^−π^+π^+) < 2030$ MeV/c$^2$, respectively. The signal is described by a sum of two Gaussian functions and the background is represented by a third-order polynomial. The data sample is separated according to magnet polarity and candidate momentum ($p_{D_s^+} < 50$ GeV/c, 50 < $p_{D_s^+}$ < 100 GeV/c, and $p_{D_s^+} > 100$ GeV/c), to take into account the dependence of the mass resolution on the momentum. The parameters are determined by simultaneous fits to these $D_s^+$ and $D_s^0$ subsamples.

3. Event selection

To reduce the combinatorial background, requirements on the quality of the reconstructed tracks, their $X^2_{IP}$, $p_T$, and scalar $p_T$ sum are applied. Additional requirements are made on the secondary vertex fit quality, the minimum significance of the displacement from the secondary to any primary vertex in the event, and the $X^2_{IP}$ of the $D_s^+$ candidate. This also reduces the contribution of secondary $D$ mesons from b-hadron decays to 1–2%, avoiding the introduction of new sources of asymmetries. The final state particles are required to satisfy particle identification (PID) criteria based on the RICH detectors.

After these requirements, there is still a significant background contribution, which could introduce charge asymmetries across the Dalitz plot. This includes semileptonic decays like $D^+ → K^−π^+μ^−ν$ and $D^+ → π^−π^+μ^−ν$; three-body decays, such as $D^+ → K^−π^+π^+$; prompt two-body $D^0$ decays forming a threeprong vertex with a random pion; and $D^0$ decays from the $D^+$ chain, such as $D^+ → D^0(K^−π^+, π^−π^+, K^−π^+π^0)π^+$. The contribution from $D^+ → K^−π^+π^+$ and prompt $D^0$ decays that involve the misidentification of the kaon as a pion is reduced to a negligible level with a more stringent PID requirement on the $π^−$ candidate. The remaining background from semileptonic decays is controlled by applying a muon veto to all three tracks, using information from the muon system [18]. The contribution from the $D^+ → K^−π^+π^+$ decay chain is reduced to a negligible level with a requirement on $X^2_{IP}$ of the $π^+$ candidate with lowest $p_T$.

Fits to the invariant mass distribution $M(π^−π^+π^+)$ are performed for the $D^+$ and $D_s^+$ candidates satisfying the above selection criteria and within the range $1810 < M(π^−π^+π^+) < 1930$ MeV/c$^2$ and $1910 < M(π^−π^+π^+) < 2030$ MeV/c$^2$, respectively. The signal is described by a sum of two Gaussian functions and the background is represented by a third-order polynomial. The data sample is separated according to magnet polarity and candidate momentum ($p_{D_s^+} < 50$ GeV/c, 50 < $p_{D_s^+}$ < 100 GeV/c, and $p_{D_s^+} > 100$ GeV/c), to take into account the dependence of the mass resolution on the momentum. The parameters are determined by simultaneous fits to these $D_s^+$ and $D_s^0$ subsamples.

The $D^+$ and $D_s^+$ invariant mass distributions and fit results for the momentum range 50 < $p_{D_s^+}$ < 100 GeV/c are shown in Fig. 1 for magnet up data. The total yields after summing over all fits are (2678 ± 7) × 10$^3$ $D^+ → π^−π^+π^+$ and (2704 ± 8) × 10$^3$ $D_s^+ → π^−π^+π^+$ decays. The final samples used for the CPV search consist of all candidates with $M(π^−π^+π^+)$ within ±2σ around $m_{D_s}$, where σ and $m_{D_s}$ are the weighted average of the two fitted Gaussian widths and mean values. The values of σ range from 8 to 12 MeV/c$^2$, depending on the momentum region. For the signal sample there are 3114 × 10$^3$ candidates, including background, while for the control mode there are 2938 × 10$^3$ candidates with purities of 82% and 87%, respectively. The purity is defined as the fraction of signal decays in this mass range.

The $D^+ → π^−π^+π^+$ and $D_s^+ → π^−π^+π^+$ Dalitz plots are shown in Fig. 2, with $s_{low}$ and $s_{high}$ being the lowest and highest invariant mass squared combination, $M^2(π^−π^+)$, respectively.
4. Binned analysis

4.1. Method

The binned method used to search for localised asymmetries in the $D^+ \to \pi^+ \pi^+ \pi^+$ decay phase space is based on a bin-by-bin comparison between the $D^+$ and $D^-$ Dalitz plots [19, 20]. For each bin of the Dalitz plot, the significance of the difference between the number of $D^+$ and $D^-$ candidates, $S_{CP}^i$, is computed as

$$S_{CP}^i = \frac{N^+_i - \alpha N^-_i}{\sqrt{\alpha(N^+_i + N^-_i)}}, \quad \alpha = \frac{N^+}{N^-},$$

where $N^+_i$ ($N^-_i$) is the number of $D^+$ ($D^-$) candidates in the $i$th bin and $N^+ + N^-$ is the sum of $N^+_i$ ($N^-_i$) over all bins. The parameter $\alpha$ removes the contribution of global asymmetries which may arise due to production [21, 22] and detection asymmetries, as well as from CPV. Two binning schemes are used, a uniform grid with bins of equal size and an adaptive binning where the bins have the same population.

In the absence of localised asymmetries, the $S_{CP}^i$ values follow a standard normal Gaussian distribution. Therefore, CPV can be detected as a deviation from this behaviour. The numerical comparison between the $D^+$ and $D^-$ Dalitz plots is made by a $\chi^2$ test, with $\chi^2 = \sum_i (S_{CP}^i)^2$. A p-value for the hypothesis of no CPV is obtained considering that the number of degrees of freedom (ndf) is equal to the total number of bins minus one, due to the constraint on the overall $D^+/D^-$ normalisation.

A CPV signal is established if a p-value lower than $3 \times 10^{-7}$ is found, in which case it can be converted to a significance for the exclusion of CP symmetry in this channel. If no evidence of CPV is found, this technique provides no model-independent way to set an upper limit.

4.2. Control mode and background

The search for local asymmetries across the $D^+_s \to \pi^- \pi^+ \pi^+$ Dalitz plot is performed using both the uniform and the adaptive (“$D^+_s$ adaptive”) binning schemes mentioned previously. A third scheme is also used: a “scaled $D^-$” scheme, obtained from the $D^+$ adaptive binning by scaling the bin edges by the ratios of the maximum values of $s_{\text{high}}(D^+_s)/s_{\text{high}}(D^-)$ and $s_{\text{low}}(D^+_s)/s_{\text{low}}(D^-)$. This scheme provides a one-to-one mapping of the corresponding Dalitz plots and allows to probe regions in the signal and control channel phase spaces where the momentum distributions of the three final state particles are similar.

The study is performed using $\alpha = 0.992 \pm 0.001$, as measured for the $D^+_s$ sample, and different granularities: 20, 30, 40, 49 and 100 adaptive bins for both the $D^+_s$ adaptive and scaled $D^+$ schemes, and $5 \times 5, 6 \times 7, 8 \times 9$ and $12 \times 12$ bins for the uniform grid scheme. Only bins with a minimum occupancy of 20 entries are considered. The p-values obtained are distributed in the range 4–87%, consistent with the hypothesis of absence of localised asymmetries. As an example, Fig. 3 shows the distributions of $S_{CP}^i$ for the $D^+_s$ adaptive binning scheme with 49 bins.

As a further cross-check, the $D^+_s$ sample is divided according to magnet polarity and hardware trigger configurations. Typically, the p-values are above 1%, although one low value of 0.07% is found for a particular trigger subset of magnet up data with 40 adaptive bins. When combined with magnet down data, the p-value increases to 11%.

The possibility of local asymmetries induced by the background under the $D^+$ signal peak is studied by considering the candidates with mass $M(\pi^+ \pi^- \pi^+ \pi^+)$ in the ranges 1810–1835 MeV/c² and 1905–1935 MeV/c², for which $\alpha = 1.000 \pm 0.002$. Using a uniform grid with four different granularities, the p-values are computed for each of the two sidebands. The data are also divided according to the magnet polarity. The p-values are found to be within 0.4–95.5%, consistent with differences in the number of $D^+$ and $D^-$ candidates arising from statistical fluctuations. Since the selection criteria suppress charm background decays to a negligible level, it is assumed that the background contribution to the signal is similar to the sidebands. Therefore, asymmetries eventually observed in the signal mode cannot be attributed to the background.

4.3. Sensitivity studies

To study the CPV sensitivity of the method for the current data set, a number of simulated pseudo-experiments are performed with sample size and purity similar to that observed in data. The $D^+ \to \pi^- \pi^+ \pi^+$ decays are generated according to an amplitude model inspired by E791 results [23], where the most important contributions originate from $\rho(770) \pi^+$, $\sigma(500) \pi^+$ and $f_2(1270) \pi^+$ resonant modes. Background events are generated evenly in the Dalitz plot. Since no theoretical predictions on the presence or size of CPV are available for this channel, various scenarios are studied by introducing phase and magnitude differences between the main resonant modes for $D^+$ and $D^-$. The sensitivity for different binning strategies is also evaluated.

Phase differences in the range 0.5–4.0° and magnitude differences in the range 0.5–4.0% are tested for $\rho(770) \pi^+$, $\sigma(500) \pi^+$ and $f_2(1270) \pi^+$ modes. The study shows a sensitivity (p-values below $10^{-7}$) around 1° to 2° in phase differences and 2% in amplitude in these channels. The sensitivity decreases when the number of bins is larger than 100, so a few tens of bins approaches the optimal choice. A slightly better sensitivity for the adaptive binning strategy is found in most of the studies.
Since the presence of background tends to dilute a potential sign of CPV, additional pseudo-experiment studies are made for different scenarios based on signal yields and purities attainable on data. Results show that better sensitivities are found for higher yields, despite the lower purity.

5. Unbinned analysis

5.1. k-Nearest neighbour analysis technique

The unbinned model-independent method of searching for CPV in many-body decays uses the concept of nearest neighbor events in a combined $D^+$ and $D^-$ samples to test whether they share the same parent distribution function $f(N_{+}, N_{-})$. To find the $n_{k}$ nearest neighbour events of each $D^+$ and $D^-$ event, the Euclidian distance between points in the Dalitz plot of three-body $D^+$ and $D^-$ decays is used. For the whole event sample a test statistic $T$ for the null hypothesis is calculated,

$$ T = \frac{1}{n_{k}(N_{+} + N_{-})} \sum_{i=1}^{N_{+} + N_{-}} \sum_{k=1}^{n_{k}} I(i,k), $$

where $I(i,k) = 1$ if the $i$th event and its $k$th nearest neighbour have the same charge and $I(i,k) = 0$ otherwise and $N_{+} (N_{-})$ is the number of events in the $D^+ (D^-)$ sample. The test statistic $T$ is the mean fraction of like-charged neighbour pairs in the combined $D^+$ and $D^-$ decays sample. The advantage of the $k$-nearest neighbour method (kNN), in comparison with other proposed methods for unbinned analyses [24], is that the calculation of $T$ is simple and fast and the expected distribution of $T$ is well known: for the null hypothesis it follows a Gaussian distribution with mean $\mu_T$ and variance $\sigma_T$ calculated from known parameters of the distributions,

$$ \mu_T = \frac{N_{+}(N_{+} - 1) + N_{-}(N_{-} - 1)}{N(N-1)}, $$

$$ \lim_{N,n_{k},D \rightarrow \infty} \sigma_T^2 = \frac{1}{N n_{k}} \left( \frac{N_{+} N_{-}}{N^2} + \frac{N_{+}^2 N_{-}^2}{N^2} \right), $$

where $N = N_{+} + N_{-}$ and $D$ is a space dimension. For $N_{+} = N_{-}$ a reference value

$$ \mu_{TR} = \frac{1}{2} \left( \frac{N - 2}{N - 1} \right) $$

is obtained and for a very large number of events $N$, $\mu_T$ approaches 0.5. However, since the observed deviations of $\mu_T$ from $\mu_{TR}$ are sometimes tiny, it is necessary to calculate $\mu_T - \mu_{TR}$. The convergence in Eq. (4) is fast and $\sigma_T$ can be obtained with a good approximation even for space dimension $D = 2$ for the current values of $N_{+}, N_{-}$ and $n_{k}$ [13,24].

The kNN method is applied to search for CPV in a given region of the Dalitz plot in two ways: by looking at a “normalization” asymmetry $(N_{+} \neq N_{-}$ in a given region) using a pull $(\mu_T - \mu_{TR})/\Delta(\mu_T - \mu_{TR})$ variable, where the uncertainty on $\mu_T$ is $\Delta\mu_T$ and the uncertainty on $\mu_{TR}$ is $\Delta\mu_{TR}$, and looking for a “shape” or particle density function (pdf) asymmetry using another pull $(T - \mu_T)/\sigma_T$ variable.

As in the binned method, this technique provides no model-independent way to set an upper limit if no CPV is found.

5.2. Control mode and background

The Cabibbo-favoured $D^+_s$ decays serve as a control sample to estimate the size of production and detection asymmetries and systematic effects. The sensitivity for local CPV in the Dalitz plot of the kNN method can be increased by taking into account only events from the region where CPV is expected to be enhanced by the known intermediate resonances in the decays. Since these regions are characterised by enhanced variations of strong phases, the conditions for observation of CPV are more favourable. Events from other regions are expected to only dilute the signal of CPV.

The Dalitz plot for the control channel $D^+_s \rightarrow \pi^- \pi^+ \pi^+$ is partitioned into three (P1–P3) or seven (R1–R7) regions shown in Fig. 4. The division R1–R7 is such that regions enriched in resonances are separated from regions dominated by smoother distributions of events. Region R3 is further divided into two regions of $s_{\text{high}}$ at masses smaller (R3l) and larger (R3r) than the $\rho^0(770)$ resonance, in order to study possible asymmetries due to a sign change of the strong phase when crossing the resonance pole. The three regions P1–P3 correspond to a more complicated structure of resonances in the signal decay $D^+_s \rightarrow \pi^- \pi^+ \pi^+$ (see Fig. 11).

The value of the test statistic $T$ measured using the kNN method with $n_k = 20$ for the full Dalitz plot (called R0) of $D^+_s \rightarrow \pi^- \pi^+ \pi^+$ candidates is compared to the expected Gaussian $T$ distribution with $\mu_T$ and $\sigma_T$ calculated from data. The calculated $p$-value is 44% for the hypothesis of no CPV asymmetry. The $p$-values are obtained by integrating the Gaussian $T$ distribution from a given value up to its maximum value of 1. The results are shown in Fig. 5 separately for each region. They do not show any asymmetry between $D^+_s$ and $D^+_s$ samples. Since no CPV is expected in the control channel, the local detection asymmetries are smaller than the present sensitivity of the kNN method. The production asymmetry is accounted for in the kNN method as a deviation of the measured value of $\mu_T$ from the
Fig. 5. (a) Pull values of $T$ and (b) the corresponding p-values for $D_s^{+} \rightarrow \pi^- \pi^+ \pi^+$ control sample candidates restricted to each region, obtained using the kNN method with $n_k = 20$. The horizontal blue lines in (a) represent $-3$ and $+3$ pull values. The region R0 corresponds to the full Dalitz plot. Note that the points for the overlapping regions are correlated.

Fig. 6. (a) Raw asymmetry $A = (N_- - N_+) / (N_- + N_+)$ and (b) the pull values of $\mu_T$ for $D_s^{+} \rightarrow \pi^- \pi^+ \pi^+$ control sample candidates restricted to each region. The horizontal lines in (b) represent $+3$ and $+5$ pull values. The region R0 corresponds to the full Dalitz plot. Note that the points for the overlapping regions are correlated.

Fig. 7. (a) Pull values of $T$ and (b) the corresponding p-values for the background candidates restricted to each region obtained using the kNN method with $n_k = 20$. The horizontal blue lines in (a) represent $-3$ and $+3$ pull values. The region R0 corresponds to the full Dalitz plot. Note that the points for the overlapping regions are correlated.

reference value $\mu_{TR}$. In the present sample, the obtained value $\mu_T - 0.5 = (84 \pm 15) \times 10^{-7}$, with $(\mu_T - \mu_{TR}) / \Delta (\mu_T - \mu_{TR}) = 5.8 \sigma$, in the full Dalitz plot is a consequence of the observed global asymmetry of about 0.4%. This value is consistent with the previous measurement from LHCb [22]. The comparison of the raw asymmetry $A = (N_- - N_+) / (N_- + N_+)$ and the pull values of $\mu_T$ in all regions are presented in Fig. 6. The measured raw asymmetry is similar in all regions as expected for an effect due to the production asymmetry. It is interesting to note the relation $\mu_T - \mu_{TR} \approx A^2 / 2$ at order $1 / N$ between the raw asymmetry and the parameters of the kNN method.

A region-by-region comparison of $D_s^{+}$ candidates for magnet down and magnet up data gives insight into left-right detection asymmetries. No further asymmetries, except for the global production asymmetry discussed above, are found.

The number of nearest neighbour events $n_k$ is the only parameter of the kNN method. The results for the control channel show no significant dependence of p-values on $n_k$. Higher values of $n_k$ reduce statistical fluctuations due to the local population density and should be preferred. On the other hand, increasing the number of nearest neighbours with limited number of events in the sample can quickly increase the radius of the local region under investigation.

The kNN method also is applied to the background events, defined in Section 4.2. Contrary to the measurements for the $D_s^{+} \rightarrow \pi^- \pi^+ \pi^+$ candidates, for background no production asymmetry is observed. The calculated $\mu_T - 0.5 = (-5.80 \pm 0.46) \times 10^{-7}$ for the full Dalitz plot is very close to the value $\mu_{TR} - 0.5 = (-5.8239 \pm 0.0063) \times 10^{-7}$ expected for an equal number of events in $D^+$ and $D^-$ samples (Eq. (5)). The measured pull values of $T$ and the corresponding p-values obtained using the kNN method with $n_k = 20$ are presented for the background in Fig. 7, separately for each region. The comparison of normalisation asymmetries and pull values of $\mu_T$ in all regions are presented in Fig. 8. All the kNN method results are consistent with no significant asymmetry.

5.3. Sensitivity studies

The sensitivity of the kNN method is tested with the same pseudo-experiment model described in Section 4.3. If the simulated asymmetries are spread out in the Dalitz plot the events may be moved from one region to another. For these asymmetries it is
observed that the difference in shape of the probability density functions is in large part absorbed in the difference in the normalisation. This indicates that the choice of the regions is important for increasing the sensitivity of the kNN method. In general the method applied in a given region is sensitive to weak phase differences greater than (1–2)$^\circ$ and magnitude differences of (2–4)%.

6. Results

6.1. Binned method

The search for CPV in the Cabibbo-suppressed mode $D^+ \rightarrow \pi^-\pi^+\pi^+$ is pursued following the strategy described in Section 4. For the total sample size of about 3.1 million $D^+$ and $D^-$ candidates, the normalisation factor $\alpha$, defined in Eq. (1), is 0.990 $\pm$ 0.001. Both adaptive and uniform binning schemes in the Dalitz plot are used for different binning sizes.

The $S_{CP}^a$ values across the Dalitz plot and the corresponding histogram for the adaptive binning scheme with 49 and 100 bins are illustrated in Fig. 9. The $p$-values for these and other binning choices are shown in Table 1. All $p$-values show statistical agreement between the $D^+$ and $D^-$ samples.

The same $\chi^2$ test is performed for the uniform binning scheme, using 20, 32, 52 and 98 bins also resulting in $p$-values consistent with the null hypothesis, all above 90%. The $S_{CP}^a$ distribution in the Dalitz plot for 98 bins and the corresponding histogram is shown in Fig. 10.

As consistency checks, the analysis is repeated with independent subsamples obtained by separating the total sample accord-
Fig. 10. (a) Distribution of $S_{i\text{CP}}$ with 98 bins in the uniform binning scheme for the total $D^+ \rightarrow \pi^- \pi^+ \pi^+$ data sample and (b) the corresponding one-dimensional $S_{i\text{CP}}$ distribution (b) with a standard normal Gaussian function superimposed (solid line).

Fig. 11. Dalitz plot for $D^+ \rightarrow \pi^- \pi^+ \pi^+$ candidates divided into (a) seven regions R1–R7 and (b) three regions P1–P3.

Fig. 12. (a) Raw asymmetry and (b) the pull values of $\mu_T$ for $D^+ \rightarrow \pi^- \pi^+ \pi^+$ candidates restricted to each region. The horizontal lines in (b) represent pull values $+3$ and $+5$. The region R0 corresponds to the full Dalitz plot. Note that the points for the overlapping regions are correlated.

The kNN method is applied to the Cabibbo-suppressed mode $D^+ \rightarrow \pi^- \pi^+ \pi^+$ with the two region definitions shown in Fig. 11. To account for the different resonance structure in $D^+$ and $D_s^+$ decays, the region R1–R7 definition for the signal mode is different from the definition used in the control mode (compare Figs. 4(a) and 11(a)). The region P1–P3 definitions are the same. The results for the raw asymmetry are shown in Fig. 12. The production asymmetry is clearly visible in all the regions with the same magnitude as in the control channel (see Fig. 6). It is accounted for in the kNN method as a deviation of the measured value of $\mu_T$ from the reference value $\mu_{TR}$ shown in Fig. 12. In the signal sample the values $\mu_T - 0.5 = (98 \pm 15) \times 10^{-7}$ and $(\mu_T - \mu_{TR})/\Delta(\mu_T - \mu_{TR}) = 6.5 \sigma$ in the full Dalitz plot are a consequence of the 0.4% global asymmetry similar to that observed in the control mode and consistent with the previous measurement from LHCb [21].

The pull values of $T$ and the corresponding p-values for the hypothesis of no CPV are shown in Fig. 13 for the same regions. To check for any systematic effects, the test is repeated for samples separated according to magnet polarity. Since the sensitivity of the method increases with $n_k$, the analysis is repeated with $n_k = 500$ for all the regions. All p-values are above 20%, consistent with no CP asymmetry in the signal mode.

7. Conclusion

A search for CPV in the decay $D^+ \rightarrow \pi^- \pi^+ \pi^+$ is performed using $pp$ collision data corresponding to an integrated luminosity...
of 1.0 fb$^{-1}$ collected by the LHCb experiment at a centre-of-mass energy of 7 TeV. Two model-independent methods are applied to a sample of 3.1 million $D^+ \rightarrow \pi^- \pi^+ \pi^-$ decay candidates with 82% signal purity.

The binned method is based on the study of the local significances $S^2_{CP}$ in bins of the Dalitz plot, while the unbinned method uses the concept of nearest neighbour events in the pooled $D^+$ and $D^-$ sample. Both methods are also applied to the Cabibbo-favoured $D^+_s \rightarrow \pi^- \pi^+ \pi^-$ decay and to the mass sidebands to control possible asymmetries not originating from CPV.

No single bin in any of the binning schemes presents an absolute $S^2_{CP}$ value larger than 3. Assuming no CPV, the probabilities of observing local asymmetries across the phase-space of the $D^+$ meson decay as large or larger than those in data are above 50% in all the tested binned schemes. In the unbinned method, the p-values are above 30%. All results are consistent with no CPV.

Acknowledgements

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at the LHCb institutions. We acknowledge support from CERN and from the national agencies: CAPES, CNPq, FAPERJ and FINENP (Brazil); NSFC (China); CNRS/IN2P3 and Region Auvergne (France); BMBF, DFG, HGF and MPG (Germany); SFI (Ireland); INFN (Italy); FOM and NWO (The Netherlands); MEN/IFA (Romania); MinES, Rosatom, RFBR and NRC “Kurchatov Institute” (Russia); MinECo, XuntaGal and GENCAT (Spain); SNSF and SER (Switzerland); NAS Ukraine (Ukraine); STFC (United Kingdom); NSF (USA). We also acknowledge the support received from the ERC under FP7. The Tier1 computing centres are supported by IN2P3 (France), KIT and BMBF (Germany), INFN (Italy), NWO and SURF (The Netherlands), PIC (Spain); GridPP (United Kingdom). We are thankful for the computing resources put at our disposal by Yandex LLC (Russia), as well as to the communities behind the multiple open source software packages that we depend on.

References

[6] LHCb Collaboration, A search for time-integrated CP violation in $D^0 \rightarrow K^- K^+$ and $D^0 \rightarrow \pi^- \pi^+$ decays, LHCb-CONF-2013-003.

LHCb Collaboration
