Measurement of time-dependent CP asymmetries in $B^0 \to D^{(*)\pm}\pi^\mp$ and $B^0 \to D^{\pm}\rho^\mp$ decays

We present updated results on time-dependent CP asymmetries in fully reconstructed $B^0 \rightarrow D^{(*)}\pi \pi$ and $B^0 \rightarrow D^{(*)}\rho \rho$ decays in approximately $232 \times 10^6$ $Y(4S) \rightarrow BB$ events collected with the BABAR detector at the PEP-II asymmetric-energy $B$ factory at SLAC. From a time-dependent maximum-likelihood fit we obtain for the parameters related to the CP violation angle $2\beta + \gamma$: $a^{D_{\rho \rho}}_{\pi \pi} = -0.010 \pm 0.023 \pm 0.007$, $c^{D_{\rho \rho}}_{\pi \pi} = -0.033 \pm 0.042 \pm 0.012$, $a^{D_{\rho \rho}}_{\pi \pi} = -0.040 \pm 0.023 \pm 0.010$, $c^{D_{\rho \rho}}_{\pi \pi} = 0.049 \pm 0.042 \pm 0.015$, $c^{D_{\rho \rho}}_{\pi \pi} = -0.024 \pm 0.031 \pm 0.009$, $t^{D_{\rho \rho}}_{\pi \pi} = -0.098 \pm 0.055 \pm 0.018$, where the first error is statistical and the second is systematic. Using other measurements and theoretical assumptions, we interpret the results in terms of the angles of the Cabibbo-Kobayashi-Maskawa unitarity triangle and find $|\sin(2\beta + \gamma)| > 0.64 (0.40)$ at 68% (90%) confidence level.

DOI: 10.1103/PhysRevD.73.111101

PACS numbers: 12.15.Hh, 11.30.Er, 13.25.Hw

In the standard model, CP violation in the weak interactions between quarks manifests itself as a nonzero area of the Cabibbo-Kobayashi-Maskawa (CKM) unitarity triangle [1]. While the measurement of $|\sin2\beta|$ is now quite precise [2,3], the constraints on the other two angles of the unitarity triangle, $\alpha$ and $\gamma$, are still limited by statistical and theoretical uncertainties.

This paper presents updates for the measurements of CP asymmetries in $B^0 \rightarrow D^{(*)}\pi \pi$ decays [4], as reported in Ref. [5], with a larger data sample ($\times 2.6$), and in addition includes the measurement of the CP asymmetry in the decay mode $B^0 \rightarrow D^{(*)}\rho \rho$. We denote these decays as $B^0 \rightarrow D^{(*)}\pi \pi$, where $\pi \pi$ is a charged pion or $\rho$ meson.

The time evolution of $B^0 \rightarrow D^{(*)}\pi \pi$ decays is sensitive to $\gamma$ because the CKM-favored decay amplitude $B^0 \rightarrow D^{(*)}\pi \pi$, which is proportional to the CKM matrix elements $V_{cb}V_{ud}^\ast$, and the doubly CKM-suppressed decay amplitude $B^0 \rightarrow D^{(*)}\rho \rho$, which is proportional to $V_{cd}V_{ub}^\ast$, interfere due to $B^0 - \bar{B}^0$ mixing. The relative weak phase between these two amplitudes is $\gamma$. With $B^0 - \bar{B}^0$ mixing, the total weak phase difference between the interfering amplitudes is $2\beta + \gamma$.

Neglecting the very small decay width difference between the two $B^0$ mass eigenstates [6], the proper-time distribution of the $B^0 \rightarrow D^{(*)}\pi \pi$ decays is given by

$$f^\pm(\eta, \Delta t) = \frac{e^{-|\Delta t|/\tau}}{4\tau} \left[ 1 \pm \frac{\eta}{2} \sin(\Delta m_d \Delta t) \right] + \eta C \cos(\Delta m_d \Delta t),$$

where $\tau$ is the $B^0$ lifetime, $\Delta m_d$ is the $B^0 - \bar{B}^0$ mixing frequency, and $\Delta t = t_{\text{rec}} - t_{\text{tag}}$ is the time difference between the $B^0 \rightarrow D^{(*)}\pi \pi$ decay ($t_{\text{rec}}$) and the decay of the other $B$ ($B_{\text{tag}}$) from the $Y(4S) \rightarrow B^0\bar{B}^0$ decay. In this equation the upper (lower) sign refers to the flavor of $B_{\text{tag}}$ as $B^0$ ($\bar{B}^0$), while $\eta = +1$ ($-1$) and $\zeta = +$ ($-$) refer to the final state $D^{(*)}\pi \pi$ ($D^{(*)}\rho \rho$). The sine term is due to interference between direct decay and decay after $B^0 - \bar{B}^0$ mixing. The cosine term arises from interference between decay amplitudes with different weak and strong phases (direct CP violation) or from CP violation in mixing. The $S$ and $C$ asymmetry parameters can be expressed as

$$S_\pm = - \frac{2\text{Im}(\lambda_\pm)}{1 + |\lambda_\pm|^2} \quad \text{and} \quad C = 1 - r^2 / 1 + r^2,$$

where $r = |\lambda_+| = 1/|\lambda_-|$ and

$$\lambda_\pm = \frac{g A(B^0 \rightarrow D^{(*)}\pi \pi)}{p A(B^0 \rightarrow D^{(*)}\rho \rho)} = r^{\pm 1} e^{-i(2\beta + \gamma \mp \delta)}.$$  

Here $g/p$ is a function of the elements of the mixing Hamiltonian [6], and $\delta$ is the relative strong phase between the two contributing amplitudes. In the standard model, CP violation in mixing is negligible and thus $|g/p| = 1$. In these equations, the parameters $r$ and $\delta$ depend on the choice of the final state. They will be indicated as $r_{D\pi}$, $\delta_{D\pi}$ for the $B^0 \rightarrow D^\pm \pi^\mp$ mode, $r_{D\rho}$, $\delta_{D\rho}$ for $B^0 \rightarrow D^\pm \rho^\mp$, and $r_{D\pi}$, $\delta_{D\pi}$ for $B^0 \rightarrow D^{\ast} \pi^\mp$ [7,8].

Interpreting the $S$ parameters in terms of the angles of the unitarity triangle requires knowledge of the corresponding $r$ parameters. The values of $r$ are expected to be small ($\sim 0.02$) and therefore cannot be extracted from the measurement of $C$. They can be estimated, assuming SU(3) symmetry and neglecting contributions from $W$-exchange diagrams, from the ratios of branching fractions $B(B^0 \rightarrow D^{(*)}\pi \pi)/B(B^0 \rightarrow D^{(*)}\pi^\mp)$ and $B(B^0 \rightarrow D^{\ast}\rho^\mp)/B(B^0 \rightarrow D^\pm \rho^\mp)$ [5,9,10].

This measurement is based on $232 \times 10^6$ $Y(4S) \rightarrow BB$ decays, collected with the BABAR detector [11] at the PEP-II asymmetric-energy $B$ factory at SLAC. We use a Monte Carlo simulation of the BABAR detector based on GEANT4 [12] to validate the analysis procedure and to estimate some of the backgrounds.

The event selection criteria are unchanged from our previous publication [5], except for the application of a kaon veto on the pion candidate in the decay modes
$D^{(*)-} \pi^+$ to suppress $B^0 \to D^{(*)-} K^+$ background events, and for the addition of the decay mode $B^0 \to D^- \rho^+$. The $D^{*-}$ is reconstructed through its decay to $\bar{D}^0 \pi^-$, where the $\bar{D}^0$ decays into $K^- \pi^+$, $K^- \pi^- \pi^0$, $K^- \pi^- \pi^+ \pi^-$, or $K^0_S \pi^+ \pi^-$. The $D^-$ is reconstructed through its decay into $K^- \pi^- \pi^-$ or $K^0_S \pi^+ \pi^-$. The $\rho^+$ decay is reconstructed in the final state $\pi^+ \pi^0$. For the $CP$ analysis we require the $\pi^+ \pi^0$ invariant mass ($m_{\pi^+ \pi^0}$) to be in the window $620 < m_{\pi^+ \pi^0} < 920 \text{ MeV}/c^2$. Exploiting the polarization of the $\rho$ meson from the decay $B^0 \to D^- \rho^+$, we require the cosine of the $\rho^+$ helicity angle $\theta_{hel}$, defined as the angle between the charged pion and the $D^-$ momentum in the $\rho^+$ rest frame, to satisfy $|\cos \theta_{hel}| > 0.4$.

The beam-energy substituted mass, $m_{ES} \equiv \sqrt{s/4 - p_B^e}$, and the difference between the $B$ candidate’s measured energy and the beam energy, $\Delta E \equiv E^e_B - \langle \sqrt{s} \rangle/2$, are used to identify the final sample, where $E^e_B$ ($p_B$) is the energy (momentum) of the $B$ candidate in the nominal $e^+e^-$ center-of-mass frame, and $\langle \sqrt{s} \rangle$ is the total center-of-mass energy. The $\Delta E$ signal region is defined as $|\Delta E| < 3\sigma$, where the resolution $\sigma$ is mode dependent and approximately 20 MeV, as determined from data. Figure 1 shows the $m_{ES}$ distribution for candidates with $m_{ES} > 5.2 \text{ GeV}/c^2$ in the $\Delta E$ signal region. These candidates satisfy the tagging and vertexing requirements, which are described later. Each distribution is fit to the sum of an Argus function [13], which accounts for the background from random combinations of tracks (combinatorial background), and a Gaussian distribution with a fitted width of about 2.5 MeV$/c^2$, which describes the signal and the backgrounds that peak in the $m_{ES}$ signal region (peaking background). Signal yields and sample purities are determined in the $m_{ES}$ signal region, with $m_{ES} > 5.27 \text{ GeV}/c^2$, and are summarized in Table I. Backgrounds from $B^0$ and $B^+$ decays that peak in the $m_{ES}$ signal region are estimated using Monte Carlo events and are mostly due to charged final states. They are also reported in Table I.

For the $B^0 \to D^- \rho^+ \pi^0$ mode we consider additional sources of background with the same final state $D^\pm \pi^+ \pi^0$, where the $\pi^+ \pi^0$ system is not produced through the $\rho^+$ resonance. Interfering sources of background can introduce a dependence of the $\lambda^\alpha_{\rho\pi}$ parameters of Eq. (3) on $m_{\pi^+ \pi^0}$. The dependency has been studied using the distribution of $m_{\pi^+ \pi^0}$.

The possible background contributions have been evaluated with a sample of 130,273 $B^0 \to D^- \pi^+ \pi^0$ candidates, on which the requirements on the $\rho$ helicity and on $m_{\pi^+ \pi^0}$ have been removed. Three interfering components are considered: $B^0 \to D^- \rho^+$ (the signal), $B^0 \to D^- \rho^0$ (1450) with a pole mass of $(1465 \pm 25) \text{ MeV}/c^2$ and a width of $(400 \pm 60) \text{ MeV}/c^2$ [6] for the $\rho^0$; both described with $P$-wave relativistic Breit-Wigner functions [14,15], and a nonresonant component, $B^0 \to D^- (\pi^+ \pi^0)_{int}$. Contributions from the decay modes $B^0 \to D^+ \pi^0 (D^{*-} \to D^- \pi^0)$ and $B^0 \to D^{*+} \pi^0 (D^{*-} \to D^- \pi^0)$ are negligible due to the kinematic constraints imposed on the $\rho$ daughter particles. We perform a fit to the binned $m_{\pi^+ \pi^0}$ distribution to extract the amplitudes of the three components, where for each bin the combinatorial background has been subtracted, as estimated from the corresponding $m_{ES}$ distribution, and the number of peaking background events has been estimated using fully simulated Monte Carlo events. The result of the fit is shown in Fig. 2. The fraction of $B^0 \to D^- \rho^+ (1450)$ and $B^0 \to D^- (\pi^+ \pi^0)_{int}$ events in the mass window $620 < m_{\pi^+ \pi^0} < 920 \text{ MeV}/c^2$ is found to be smaller than 0.02 at 90% confidence level (C.L.)

Table I. Signal yields, sample purities $P$, and fractions of peaking backgrounds, $f_{peak}$, for the selected samples for events that satisfy the tagging and vertexing requirements described in the text.

<table>
<thead>
<tr>
<th>Decay mode</th>
<th>Yield</th>
<th>$P(%)$</th>
<th>$f_{peak}(%)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^0 \to D^- \pi^+$</td>
<td>105038</td>
<td>87.0±0.3</td>
<td>1.6±0.1</td>
</tr>
<tr>
<td>$B^0 \to D^- \rho^+$</td>
<td>14002</td>
<td>93.2±0.2</td>
<td>1.0±0.1</td>
</tr>
<tr>
<td>$B^0 \to D^- \rho^0$</td>
<td>8736</td>
<td>81.7±0.4</td>
<td>1.3±0.2</td>
</tr>
</tbody>
</table>

FIG. 1 (color online). $m_{ES}$ distributions in the signal region for, from top to bottom, the $B \to D^- \pi^+$, $B \to D^- \pi^0$, and $B \to D^- \rho^+$ sample for the events that satisfy the tagging and vertexing requirements described in the text, fit with the function described in the text. The dashed lines indicate the sum of the combinatorial and peaking background contributions.
The proper-time interval $\Delta t$ between the two $B$ decays is calculated from the measured separation $\Delta z$, between the $B_{\text{rec}}$ and $B_{\text{tag}}$ decay points along the beam direction. We determine the $B_{\text{rec}}$ decay point from its charged tracks. The $B_{\text{tag}}$ decay point is obtained by fitting tracks that do not belong to $B_{\text{rec}}$, imposing constraints from the $B_{\text{rec}}$ momentum and the beam-spot location. We accept events with calculated $\Delta t$ uncertainty of less than 2.5 ps and $|\Delta t| < 20$ ps. The average $\Delta t$ resolution is approximately 1.1 ps. We use multivariate algorithms that identify signatures in the $B_{\text{tag}}$ decay products to determine ("tag") the flavor to be either a $B^0$ or a $\bar{B}^0$ [2]. Primary leptons from semileptonic $B$ decays are selected from identified electrons and muons and from isolated energetic tracks. The charges of identified kaons and soft pions from $D^{\pm}$ decays are also used to extract flavor information. Each event with an estimated mistag probability less than 45% is assigned to the lepton tagging category, while other events are divided into categories. The lepton tagging category contains events with an identified lepton, while other events are divided into categories based on their estimated mistag probability. The effective efficiency of the tagging algorithm, defined as $Q = \sum \epsilon_i (1 - w_i)^2$, where $\epsilon_i$ and $w_i$ are the efficiency and the mistag probability, respectively, for category $i$, is $30.1 \pm 0.5%$.

Since the expected CP asymmetry in the selected $B$ decays is small, this measurement is sensitive to the interference between the $b \to u$ and $b \to c$ amplitudes in the decay of $B_{\text{tag}}$. To account for this "tagside interference," we use a parametrization different from Eq. (2), which is described in Ref. [16] and summarized here. For each tagging category $i$, independent of the decay mode $\mu \in \{ D_{\pi}, D^{* \pi}, D_{\rho} \}$, the tagside interference is parametrized in terms of the effective parameters $r_i^\mu$ and $\delta_i^\mu$. Neglecting terms of order $(r_i^\mu)^2$ and $(r_i^\mu)^2$, the $\Delta t$ distributions are written as

$$f_i^{\pm \mu}(\eta, \Delta t) = \frac{e^{-|\Delta t|/\tau}}{4\tau} \left[ 1 + (a_i^\mu \mp \eta b_i^\mu - \eta c_i^\mu) \times \sin(\Delta m_{\mu} \Delta t) + \eta \cos(\Delta m_{\mu} \Delta t) \right],$$

where, in the standard model,

$$a_i^\mu = 2 r_i^\mu \sin(2\beta + \gamma) \cos \delta_i^\mu,$$

$$b_i^\mu = 2 r_i^\mu \sin(2\beta + \gamma) \cos \delta_i^\mu,$$

$$c_i^\mu = 2 \cos(2\beta + \gamma)(r_i^\mu \sin \delta_i^\mu - r_i^\mu \sin \delta_i^\mu).$$

Semileptonic $B$ decays do not have a doubly CKM-suppressed amplitude contribution, and hence $r_i^\mu = 0$. In the following, we quote results for the six $a_i^\mu$ and $c_i^\mu$ parameters, which are independent of the unknown parameters $r_i^\mu$ and $\delta_i^\mu$. The other $b_i^\mu$ and $c_i^\mu$ parameters depend on $r_i^\mu$ and $\delta_i^\mu$, and do not contribute to the interpretation of the result in terms of $\sin(2\beta + \gamma)$. Note that all tagging categories contribute to the measurement of the $a_i^\mu$ parameters.

An unbinned maximum-likelihood fit is applied to the $\Delta t$ distribution of the selected $B$ candidates in the $\Delta E$ signal region. The whole $m_{\text{ES}}$ range is used to determine the signal probability of each event on the basis of the Argus plus Gaussian fit described previously. The effect of finite $\Delta t$ resolution is described by convoluting Eq. (4) with a resolution function composed of three Gaussian distributions. Incorrect tagging dilutes the parameters $a_i^\mu$, $c_i^\mu$, and the coefficient of $\cos(\Delta m_{\mu} \Delta t)$ by a factor $D_\mu = 1 - 2w_i$ [2]. The parameters of the resolution function and those associated with flavor tagging are determined simultaneously from the fit to the data and are consistent with previous BABAR analyses [2]. The $\Delta t$ distribution of the combinatorial background is parametrized using two empirical components: a prompt component with zero lifetime and a component with an effective lifetime. The components are convoluted with the sum of two Gaussians, and the resolution parameters of the two Gaussians, including the effective dilution parameters, the effective lifetime, and the relative fraction of the two components, are determined from the fit to the data. The peaking background coming from $B^\pm$ mesons is modeled by an exponential with the $B^\pm$ lifetime. Its relative fraction is fixed to the value estimated from simulations. The resolution function is the same as the signal resolution, while the dilution parameters are fixed to the values obtained from a $B^+ \to \mu^+ \nu$ control sample. The peaking backgrounds from $B^0$ mesons, whose amounts are also fixed to the value estimated using simulation, are modeled with a likelihood similar to the signal likelihood, but without $CP$ violation (all the $a$, $b$, $c$ parameters set to zero). Possible $CP$ violation in this background is taken into account in the evaluation of the systematic uncertainties. The resolution and the dilution parameters are the same as for the signal.
From the unbinned maximum-likelihood fit we obtain
\[ a^{D\pi} = -0.010 \pm 0.023 \pm 0.007, \]
\[ c_{lep}^{D\pi} = -0.033 \pm 0.042 \pm 0.012, \]
\[ a^{D\rho} = -0.040 \pm 0.023 \pm 0.010, \]
\[ c_{lep}^{D\rho} = 0.049 \pm 0.042 \pm 0.015, \]
\[ a^{D\pi} = -0.024 \pm 0.031 \pm 0.009, \]
\[ c_{lep}^{D\rho} = -0.098 \pm 0.055 \pm 0.018, \]
where the first quoted error is statistical and the second is systematic. The largest correlation with any linear combination of other fit parameters is about 20% and 30% for the \( a^\mu \) and the \( c_{lep}^\rho \) parameters, respectively. Figure 3 shows the fitted \( \Delta t \) distributions for events from the lepton tagging category, which has the lowest level of background and mistag probability. The various contributions to the systematic uncertainties of the \( a \) and \( c_{lep} \) parameters are shown in Table II.

![Graph showing distributions of \( \Delta t \) for different tagging categories.](image)

**TABLE II.** Systematic uncertainties on the \( a \) and \( c_{lep} \) parameters (in units of 10^{-3}).

<table>
<thead>
<tr>
<th>Source</th>
<th>( \sigma_a )</th>
<th>( \sigma_c )</th>
<th>( \sigma_{lep} )</th>
<th>( \sigma_{lep} )</th>
<th>( \sigma_{lep} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertexing (( \sigma_{vt} ))</td>
<td>0.37</td>
<td>0.64</td>
<td>0.80</td>
<td>1.14</td>
<td>0.47</td>
</tr>
<tr>
<td>Fit (( \sigma_{fit} ))</td>
<td>0.51</td>
<td>0.95</td>
<td>0.95</td>
<td>0.99</td>
<td>0.75</td>
</tr>
<tr>
<td>Model (( \sigma_{mod} ))</td>
<td>0.12</td>
<td>0.13</td>
<td>0.12</td>
<td>0.13</td>
<td>0.01</td>
</tr>
<tr>
<td>Tagging (( \sigma_{tag} ))</td>
<td>0.07</td>
<td>0.16</td>
<td>0.11</td>
<td>0.14</td>
<td>0.06</td>
</tr>
<tr>
<td>Background (( \sigma_{bkg} ))</td>
<td>0.13</td>
<td>0.10</td>
<td>0.10</td>
<td>0.09</td>
<td>0.28</td>
</tr>
<tr>
<td>( m_{\pi\pi} ) dependence (( \sigma_{\delta} ))</td>
<td>( \cdots )</td>
<td>( \cdots )</td>
<td>( \cdots )</td>
<td>( \cdots )</td>
<td>( 0.16 )</td>
</tr>
<tr>
<td>Total (( \sigma_{tot} ))</td>
<td>0.66</td>
<td>1.17</td>
<td>0.97</td>
<td>1.53</td>
<td>0.94</td>
</tr>
</tbody>
</table>

The impact of a possible systematic mismeasurement of \( \Delta t \) (\( \sigma_{\Delta t} \)) has been estimated by comparing different parametrizations of the resolution function, varying the position of the beam spot and the absolute \( z \) scale within their uncertainties, and loosening and tightening the quality criteria on the reconstructed decay points. We also estimate...
the impact of the uncertainties on the alignment of the silicon vertex tracker (SVT) by repeating the measurement using simulated events, with the SVT intentionally misaligned. For the systematic uncertainty of the fit ($\sigma_{\text{fit}}$), we quote the upper limit on the bias on the $a^D$ and $c^D$ parameters, as estimated from samples of fully simulated events. The model error ($\sigma_{\text{mod}}$) contains the uncertainty on the $B^0$ lifetime and $\Delta m_{d}$, varied by the uncertainties on the world averages [6] and also by allowing them to vary in the fit. When fit for, the values obtained are consistent with the world averages. The tagging error ($\sigma_{\text{tag}}$) is estimated considering possible differences in tagging efficiency between $B^0$ and $\bar{B}^0$, different mistag fractions for the decay modes $D\pi$, $D^{*}\pi$, $D\rho$, and different $\Delta t$ resolutions for correctly and incorrectly tagged events. We also account for uncertainties in the background ($\sigma_{\text{bkg}}$) by varying the effective lifetimes, dilutions, $m_{\text{ES}}$ shape parameters, signal fractions, and background $CP$ asymmetry. The dependence of $a^{D\pi}$ and $c^{D\rho}_{\text{lep}}$ on the background contribution from $B^0 \to D^{-}\rho^{+}(1450)$ and $B^0 \to D^{-}(\pi^{+}\pi^{0})_{\text{tot}} (\sigma_{\chi})$ is estimated from the fit to the $\pi\pi^{0}$ invariant mass spectrum in Fig. 2.

As a cross-check, we perform the same fits on a sample of 6843 $B^{-} \to D^{(*)0}\pi^{-}$ candidates, where, as expected, we find no $CP$ asymmetries. We combine our results with the result obtained on the partially reconstructed $B \to D^{(*)}\pi^{0}$ sample [10] and use a frequentist method described in Ref. [10] to set a constraint on $2\beta + \gamma$. The confidence level as a function of $|\sin(2\beta + \gamma)|$ is shown in Fig. 4. We set the lower limits $|\sin(2\beta + \gamma)| > 0.64 (0.40)$ at 68% (90%) C.L.

In conclusion, we have studied the time-dependent $CP$-violating asymmetries in fully reconstructed $B^0 \to D^{(*)}\pi^{0}$, $B^0 \to D^{(*)}\pi^{-}$, and $B^0 \to D^{-}\rho^{+}$ decays in a sample of $232 \times 10^{6} Y(4S) \to BB$ decays, and have measured the $CP$-violating parameters listed in Eq. (6). We interpret the result in terms of $\sin(2\beta + \gamma)$ and find that $|\sin(2\beta + \gamma)| > 0.64 (0.40)$ at 68% (90%) C.L. These results are consistent with and supersede our previous measurement.

We are grateful for the excellent luminosity and machine conditions provided by our PEP-II colleagues, and for the substantial dedicated effort from the computing organizations that support BABAR. The collaborating institutions wish to thank SLAC for its support and kind hospitality. This work is supported by DOE and NSF (USA), NSERC (Canada), IHEP (China), CEA and CNRS-IN2P3 (France), BMBF and DFG (Germany), INFN (Italy), FOM (The Netherlands), NFR (Norway), MIST (Russia), and PPARC (United Kingdom). Individuals have received support from CONACyT (Mexico), Marie Curie EIF (European Union), the A.P. Sloan Foundation, the Research Corporation, and the Alexander von Humboldt Foundation.

\begin{itemize}
  \item [4] Charge conjugation is implied in this paper, unless otherwise stated. The superscript (*) indicates that the superscripted symbol must be considered both with and without the *.
  \item [8] According to Ref. [7], the strong phase difference for $B^0 \to D^{(*)}\pi^{0}$ is given by $\delta^{D\pi}\pi^{0}$, but this does not affect this measurement.
\end{itemize}