The following full text is a publisher's version.

For additional information about this publication click this link. http://hdl.handle.net/2066/128944

Please be advised that this information was generated on 2017-08-27 and may be subject to change.
Measurement of $CP$-Violating Asymmetries in $B^0$ Decays to $CP$ Eigenstates

CP-violating asymmetries in the time distributions of decays of $B^0$ and $\bar{B}^0$ mesons provide a direct test of the standard model of electroweak interactions [1]. For the neutral $B$ decay modes reported here, corrections to CP-violating effects from strong interactions are absent, in contrast to the $K^0_L$ modes in which CP violation was discovered [2].

Using a data sample of $23 \times 10^6$ $B\bar{B}$ pairs recorded at the Y(4S) resonance by the Babar detector at the PEP-II asymmetric-energy $e^+e^-$ collider at the Stanford Linear Accelerator Center, we have fully reconstructed a sample $B_{CP}$ of neutral $B$ mesons decaying to the CP eigenstates $J/\psi K^0_S$, $\psi(2S)K^0_S$, and $J/\psi K^0_L$. We examine each of the events in this sample for evidence that the other neutral $B$ meson decayed as either a $B^0$ or a $\bar{B}^0$, designated as a $B^0$ or $\bar{B}^0$ flavor tag. The final $B_{CP}$ sample contains about 360 signal events.

When the Y(4S) decays, the $P$-wave $B\bar{B}$ state evolves coherently until one of the mesons decays. In one of four time-order and flavor configurations, if the tagging meson $B_{tag}$ decays first, and as a $B^0$, the other meson must be a $\bar{B}^0$ at that same time $t_{tag}$. It then evolves independently and can decay into a CP eigenstate $B_{CP}$ at a later time $t_{CP}$. The time between the two decays $\Delta t = t_{CP} - t_{tag}$ is a signed quantity made measurable by producing the Y(4S) with a boost $\beta = 0.56$ along the collision ($z$) axis, with nominal energies of 9.0 and 3.1 GeV for the electron and positron beams. The measured distance $\Delta z = \beta \gamma c \Delta t$ between the two decay vertices provides a good estimate of the corresponding time interval $\Delta t$; the average value of $|\Delta z|$ is $\beta \gamma c t_{CP} = 250$ $\mu$m.

The decay-time distribution for events with a $B^0$ or a $\bar{B}^0$ tag can be expressed in terms of a complex parameter $\lambda$ that depends on both $B^0\bar{B}^0$ mixing and on the amplitudes describing $\bar{B}^0$ and $B^0$ decay to a common final state $f$ [3]. The distribution $f_{\pm}(f_-)$ of the decay rate when the tagging meson is a $B^0(\bar{B}^0)$ is given by

\[ f_{\pm}(\Delta t) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{2\tau_{B^0}(1 + |\lambda|^2)} \times \left[ \frac{1 + |\lambda|^2}{2} \right] \pm \frac{1 - |\lambda|^2}{2} \cos(\Delta m_{B^0} \Delta t) , \]

where $\tau_{B^0}$ is the $B^0$ lifetime and $\Delta m_{B^0}$ is the mass difference determined from $B^0\bar{B}^0$ mixing [4], and where the lifetime difference between neutral $B$ mass eigenstates is assumed to be negligible. The first oscillatory term in Eq. (1) is due to interference between direct decay and decay after mixing. A difference between the $B^0$ and $\bar{B}^0$ distributions or a $\Delta t$ asymmetry for either tag is evidence for CP violation.

If all amplitudes contributing to $B^0 \rightarrow f$ have the same weak phase, a condition satisfied in the standard model for charmonium-containing $b \rightarrow c\bar{c}d$ decays, then $|\lambda| = 1$. For these CP eigenstates the standard model predicts $\lambda = \eta_f e^{-2i\beta}$, where $\eta_f$ is the CP eigenvalue of the state $f$ and $\beta = \arg [-V_{cd}V_{cb}^*/V_{ub}V_{ub}^*]$ is an angle of the unitarity triangle of the three-generation Cabibbo-Kobayashi-Maskawa (CKM) matrix [5]. Thus, the time-dependent CP-violating asymmetry is $A_{CP}(\Delta t) = \frac{f_+(\Delta t) - f_-(\Delta t)}{f_+(\Delta t) + f_-(\Delta t)}$ where $f_+ = -\eta_f \sin 2\beta \sin(\Delta m_{B^0} \Delta t)$ and $f_- = +1$ for $J/\psi K^0_S$ and $\psi(2S)K^0_L$.

A measurement of $A_{CP}$ requires determination of the experimental $\Delta t$ resolution and the fraction of events in which the tag assignment is incorrect. A mistag fraction $w$ reduces the observed asymmetry by a factor $(1 - 2w)$.

Several samples of fully reconstructed $B^0$ mesons are used in this measurement. The $B_{CP}$ sample contains candidates reconstructed in the CP eigenstates $J/\psi K^0_S, \psi(2S)K^0_S$, and $J/\psi K^0_L$. The $J/\psi$ and $\psi(2S)$ mesons are reconstructed through their decays to $e^+e^-$ and $\mu^+\mu^-$; the $\psi(2S)$ is
also reconstructed through its decay to \( J/\psi \pi^+ \pi^- \). A sample of \( B \) decays \( B_{\text{flav}} \) [6] used in the determination of the mistag fractions and \( \Delta t \) resolution functions consists of the channels \( D^{(*)-}h^+(h^+ = \pi^+, \rho^+, \omega) \) and \( J/\psi K^{*0}(K^{*0} \rightarrow K^+ \pi^-) \). A control sample of charged \( B \) mesons decaying to the final states \( J/\psi K^{(*)+}, \psi(2S)K^+ \), and \( D^{(*)0} \pi^+ \) is used for validation studies.

A description of the BaBar detector can be found in Ref. [7]. Charged particles are detected and their momenta measured by a combination of a silicon vertex tracker (SVT) consisting of five double-sided layers and a central drift chamber (DCH), in a 1.5-T solenoidal field. The average vertex resolution in the \( z \) direction is 70 \( \mu \)m for a fully reconstructed \( B \) meson. We identify leptons and hadrons with measurements from all detector systems, including the energy loss \( (dE/dx) \) in the DCH and SVT. Electrons and photons are identified by a CsI electromagnetic calorimeter (EMC). Muons are identified in the instrumented flux return (IFR). A Cherenkov ring imaging detector (DIRC) covering the central region, together with the \( dE/dx \) information, provides \( K-\pi \) separation of at least 3 standard deviations for \( B \) decay products with momentum greater than 250 MeV/c in the laboratory.

We select events with a minimum of three reconstructed charged tracks, each having a laboratory polar angle between 0.41 and 2.54 rad and an impact parameter in the plane transverse to the beam less than 1.5 cm from the beam line. The event must have a total measured energy in the laboratory greater than 4.5 GeV within the fiducial regions for charged tracks and neutral clusters. To help reject continuum background, the second Fox-Wolfram moment [8] must be less than 0.5.

An electron candidate must have a ratio of calorimeter energy to track momentum, an EMC cluster shape, a DCH \( dE/dx \), and a DIRC Cherenkov angle (if available) consistent with an electron.

A muon candidate must satisfy requirements on the measured and expected number of interaction lengths penetrated, the position match between the extrapolated DCH track and IFR hits, and the average and spread of the number of IFR hits per layer.

A track is identified as a kaon candidate by means of a neural network that uses \( dE/dx \) measurements in the DCH and SVT, and comparison of the observed pattern of detected photons in the DIRC with that expected for kaon and pion hypotheses.

Candidates for \( J/\psi \rightarrow \ell^+ \ell^- \) must have at least one decay product identified as a lepton (electron or muon) candidate or, if outside the calorimeter acceptance, must have DCH \( dE/dx \) information consistent with the electron hypothesis. Tracks in which the electron has radiated are combined with bremsstrahlung photons, reconstructed as clusters with more than 30 MeV lying within 35 mrad in polar angle and 50 mrad in azimuth of the projected photon position on the EMC. The second track of a \( \mu^+ \mu^- \) pair, if within the acceptance of the calorimeter, must be consistent with being a minimum ionizing particle. Two identified electron or muon candidates are required for \( J/\psi \) or \( \psi(2S) \rightarrow \ell^+ \ell^- \) reconstruction in the higher-background \( \psi(2S)K^0 \) and \( J/\psi K^0 \) channels.

We require a \( J/\psi \) candidate to have \( 2.95 \leq m_{\ell^+ \ell^-} \leq 3.14 \) GeV/c\(^2\) or \( 3.06 \leq m_{\mu^+ \mu^-} \leq 3.14 \) GeV/c\(^2\), and a \( \psi(2S) \rightarrow \ell^+ \ell^- \) candidate to have \( 3.44 \leq m_{\ell^+ \ell^-} \leq 3.74 \) GeV/c\(^2\) or \( 3.64 \leq m_{\mu^+ \mu^-} \leq 3.74 \) GeV/c\(^2\). Requirements are made on the lepton helicity angle in order to provide further discrimination against background. For the \( \psi(2S) \rightarrow J/\psi \pi^+ \pi^- \) mode, mass-constrained \( J/\psi \) candidates are combined with pairs of oppositely charged tracks considered as pions; the resulting mass must be within 15 MeV/c\(^2\) of the \( \psi(2S) \) mass [4].

A \( K^0 \) candidate must satisfy \( 489 < m_{\pi^+ \pi^-} < 507 \) MeV/c\(^2\). The distance between the \( J/\psi \) or \( \psi(2S) \) and \( K^0 \) vertices is required to be at least 1 mm.

Pairs of \( \pi^0 \) candidates with total energy above 800 MeV are considered as \( K^0 \) candidates for the \( J/\psi \) or \( \psi(2S) \) mode. We determine the most probable \( K^0 \) decay point along the path defined by the initial \( K^0 \) momentum vector and the \( J/\psi \) or \( \psi(2S) \) vertex by maximizing the product of probabilities for the daughter \( \pi^0 \) mass-constrained fits. Allowing for vertex resolution, we require the displacement from the \( J/\psi \) or \( \psi(2S) \) decay point to be between \(-10 \) and \(+40 \) cm and the \( \pi^0 \pi^0 \) mass evaluated at this point to be between 470 and 550 MeV/c\(^2\).

A \( K^0 \) candidate is formed from a cluster not matched to a reconstructed track. For the EMC the cluster must have energy above 200 MeV, while for the IFR the cluster must have at least two layers. We determine the \( K^0 \) energy by combining its direction with the reconstructed \( J/\psi \) momentum, assuming the decay \( B^0 \rightarrow J/\psi K^0 \). To reduce photon backgrounds, EMC clusters consistent with a \( \pi^0 \rightarrow \gamma \gamma \) decay are rejected and the transverse missing momentum of the event projected on the \( K^0 \) candidate direction must be consistent with the \( K^0 \) momentum. In addition, the center-of-mass \( J/\psi \) momentum is required to be greater than 1.4 GeV/c.

\( B_{\text{CP}} \) candidates used in the analysis are selected by requiring that the difference \( \Delta E \) between the energy of the \( B_{\text{CP}} \) candidate and the beam energy in the center-of-mass frame be less than 3 standard deviations from zero and that, for \( K^0 \) modes, the beam-energy-substituted mass \( m_{\text{ES}} = \sqrt{(E_{\text{beam}})^2 - (p_B)^2} \) must be greater than 5.2 GeV/c\(^2\). The resolution for \( \Delta E \) is about 10 MeV, except for \( J/\psi K^0 \) (3 MeV) and the \( K^0 \rightarrow \pi^0 \pi^0 \) mode (33 MeV). For the purpose of determining numbers of events, purities, and efficiencies, a signal region \( m_{\text{ES}} > 5.27 \) GeV/c\(^2\) is used for all modes except \( J/\psi K^0 \).

Figure 1 shows the resulting \( \Delta E \) and \( m_{\text{ES}} \) distributions for \( B_{\text{CP}} \) candidates containing a \( K^0 \) and \( \Delta E \) for the candidates containing a \( K^0 \) or \( K^0 \) or \( K^0 \). The \( B_{\text{CP}} \) sample is composed of 890 events in the signal region, with an estimated background of 260 events, predominantly in the \( J/\psi K^0 \) channel. For that channel, the composition, effective \( \eta_{\text{CP}} \), and \( \Delta E \) distributions of the individual background sources are
taken either from a Monte Carlo simulation (for $B$ decays to $J/\psi$) or from the $m_{\ell^+\ell^-}$ sidebands in data.

For flavor tagging, we exploit information from the incompletely reconstructed other $B$ decay in the event. The charge of energetic electrons and muons from semileptonic $B$ decays, kaons, soft pions from $D^*$ decays, and high momentum charged particles is correlated with the flavor of the decaying $b$ quark: e.g., a positive lepton yields a $B^0$ tag. Each event is assigned to one of four hierarchical, mutually exclusive tagging categories or is excluded from further analysis. The mistag fractions and efficiencies of all categories are determined from data.

A lepton tag requires an electron or muon candidate with a center-of-mass momentum $p_{cm} > 1.0$ or 1.1 GeV/c, respectively. This efficiently selects primary leptons and reduces contamination due to oppositely charged leptons from semileptonic charm decays. Events meeting these criteria are assigned to the lepton category unless the lepton charge and the net charge of all kaon candidates indicate opposite tags. Events without a lepton tag but with a nonzero net kaon charge are assigned to the kaon category.

All remaining events are passed to a neural network algorithm whose main inputs are the momentum and charge of the track with the highest center-of-mass momentum, and the outputs of secondary networks, trained with Monte Carlo samples to identify primary leptons, kaons, and soft pions. Based on the output of the neural network algorithm, events are tagged as $B^0$ or $\bar{B}^0$ and assigned to the NT1 (more certain tags) or NT2 (less certain tags) category, or not tagged at all. The tagging power of the NT1 and NT2 categories arises primarily from soft pions and from recovering unidentified isolated primary electrons and muons.

Table I shows the number of tagged events and the signal purity, determined from fits to the $m_{ES}$ ($K^0_L$ modes) or $\Delta E$ ($K^0_S$ mode) distributions. The measured efficiencies for the four tagging categories are summarized in Table II.

The uncertainty in the $\Delta t$ measurement is dominated by the measurement of the position $z_{tag}$ of the tagging vertex. The tagging vertex is determined by fitting the tracks not belonging to the $B_{CP}$ (or $B_{flav}$) candidate to a common vertex. Reconstructed $K^0_L$ and $\Lambda$ candidates are used as input to the fit in place of their daughters. Tracks from $\gamma$ conversions are excluded from the fit. To reduce contributions from charm decay, which bias the vertex estimation, the track with the largest vertex $\chi^2$ contribution greater than 6 is removed and the fit is redone until no track fails the $\chi^2$ requirement or fewer than two tracks remain. The average resolution for $\Delta z = z_{CP} - z_{tag}$ is 190 $\mu$m. The time interval $\Delta t$ between the two $B$ decays is then determined from the $\Delta z$ measurement, including an event-by-event correction for the direction of the $B$ with respect to the $z$ direction in the $\Upsilon(4S)$ frame. An accepted candidate must have a converged fit for the $B_{CP}$ and $B_{tag}$ vertices, an error of less than 400 $\mu$m on $\Delta z$, and a measured $|\Delta z| < 3$ mm; 86% of the $B_{CP}$ events satisfy this requirement.

The sin2$\beta$ measurement is made with an unbinned maximum likelihood fit to the $\Delta t$ distribution of the combined $B_{CP}$ and $B_{flav}$ tagged samples. The $\Delta t$ distribution of the former is given by Eq. (1), with $|\alpha| = 1$. The latter evolves according to the known rate for flavor oscillations in neutral $B$ mesons. The amplitudes for $B_{CP}$ asymmetries and for $B_{flav}$ flavor oscillations are reduced by the same factor.

![Diagram](image)

FIG. 1. (a) Distribution of $m_{ES}$ and $\Delta E$ for $B_{CP}$ candidates having a $K^0_L$ in the final state; (b) distribution of $\Delta E$ for $J/\psi K^0_L$ candidates.

<table>
<thead>
<tr>
<th>Sample</th>
<th>$N_{tag}$</th>
<th>Purity (%)</th>
<th>sin2$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J/\psi K^0_L, \psi(2S)\Lambda^0$</td>
<td>273</td>
<td>96 ± 1</td>
<td>0.25 ± 0.22</td>
</tr>
<tr>
<td>$J/\psi K^0_L$</td>
<td>256</td>
<td>39 ± 6</td>
<td>0.87 ± 0.51</td>
</tr>
<tr>
<td>Full CP sample</td>
<td>529</td>
<td>69 ± 2</td>
<td>0.34 ± 0.20</td>
</tr>
</tbody>
</table>

Table I. Number of tagged events, signal purity, and result of fitting for CP asymmetries in the full CP sample and in various subsamples, as well as in the $B_{flav}$ and charged $B$ control samples. Purity is the fitted number of signal events divided by the total number of events in the $\Delta E$ and $m_{ES}$ signal region defined in the text. Errors are statistical only.

<table>
<thead>
<tr>
<th>$B_{flav}$ sample</th>
<th>$B^0$ tags</th>
<th>$\bar{B}^0$ tags</th>
</tr>
</thead>
<tbody>
<tr>
<td>4637</td>
<td>141</td>
<td>96 ± 2</td>
</tr>
<tr>
<td>3965</td>
<td>132</td>
<td>97 ± 2</td>
</tr>
</tbody>
</table>

Charged $B$ sample | 5165 | 90 ± 1 | 0.02 ± 0.05 |
factor $(1 - 2w)$ due to mistags. The distributions are both
convoluted with a common $\Delta t$ resolution function and
corrected for backgrounds, incorporated with different
assumptions about their $\Delta t$ evolution and convoluted with
a separate resolution function. Events are assigned signal
and background probabilities based on fits to $m_{ES}$ (all
modes except $J/\psi K_0^{(*)}$) or $\Delta E (J/\psi K_0^{(*)})$ distributions.

The $\Delta t$ resolution function for signal candidates is rep-
resented by a sum of three Gaussian distributions with dif-
ferent means and widths. For the core and tail Gaussians,
the widths are scaled by the event-by-event measurement
error derived from the vertex fits; the combined rms error is
1.1 ps. A separate offset for the core distribution is al-
lowed for each tagging category to account for small shifts
caused by inclusion of residual charm decay products in
the tag vertex; a common offset is used for the tail compo-
nent. The third Gaussian (of fixed 8 ps width) accounts for
the fewer than 1% of events with incorrectly reconstructed
vertices. Identical resolution function parameters are used
for all modes, since the $B_{tag}$ vertex precision dominates the
$\Delta t$ resolution.

A total of 35 parameters are varied in the final fit, in-
cluding the values of $\sin2\beta$ (1), the average mistag frac-
tion $w$ and the difference $\Delta w$ between $B^0$ and $\bar{B}^0$ mistags
for each tagging category (8), parameters for the signal $\Delta t$
resolution (9), and parameters for background time depen-
dence (6), $\Delta t$ resolution (3) and mistag fractions (8). The
determination of the mistag fractions and signal $\Delta t$ res-
olution function is dominated by the high-statistics $B_{flav}$
sample, while background parameters are governed by events
with $m_{ES} < 5.27$ GeV/$c^2$ (except $J/\psi K_0^{(*)}$). We fix
$\tau_{B^0} = 1.548$ ps and $\Delta m_{B^0} = 0.472 \pm 0.016$ ps$^{-1}$ [4]. The largest correlation between $\sin2\beta$ and any linear combination of
the other free parameters is 0.076.

The measurement of $\sin2\beta$ was performed as a blind
analysis by hiding the value of $\sin2\beta$ obtained from the
fit, as well as the $CP$ asymmetry in the $\Delta t$ distribution,
until the analysis was complete. This allowed us to study
statistical and systematic errors without knowing the nu-
merical value of $\sin2\beta$.

The measured mistag rates obtained from the like-
lihood fit for the four tagging categories are summarized
in Table II. As a check, the mistag rates were evaluated
with a sample of about 16 000 $D^{*-} \ell^+ \nu_\ell$ events and found
not to be consistent with the results from the hadronic decay
sample.

The combined fit to the $CP$ decay modes and the flavor
decay modes yields

$$\sin2\beta = 0.34 \pm 0.20 \text{ (stat)} \pm 0.05 \text{ (syst)}.$$ 

The decay asymmetry $A_{CP}$ as a function of $\Delta t$ and the
log likelihood as a function of $\sin2\beta$ are shown in Fig. 2. If $|\lambda|$ is allowed to float in the fit, the value obtained is
consistent with 1 and there is no significant difference in
the value of $\eta_{\ell}$, $\lambda$ (identified with $\sin2\beta$ in the
standard model) and our quoted result. Repeating the fit
with all parameters fixed to their determined values except
$\sin2\beta$, we find that a total contribution of $\pm 0.02$ to the
error on $\sin2\beta$ is due to the combined statistical uncertainties
in mistag rates, $\Delta t$ resolution, and background parameters.

The dominant sources of systematic error are the as-
sumed parametrization of the $\Delta t$ resolution function (0.04),
due in part to residual uncertainties in the SVT alignment,
and uncertainties in the level, composition, and $CP$ asym-
metry of the background in the selected $CP$ events (0.02).
The systematic errors from uncertainties in $\Delta m_{B^0}$ and $\tau_{B^0}$
and from the parametrization of the background in the se-
lected $B_{flav}$ sample are found to be negligible. An increase
of 0.02 $\Delta t$ ps$^{-1}$ in the assumed value for $\Delta m_{B^0}$ decreases
$\sin2\beta$ by 0.012.

The large sample of reconstructed events allows a num-
ber of consistency checks, including separation of the data
by decay mode, tagging category, and $B_{tag}$ flavor. The re-
results of fits to these subsamples are shown in Table I for the
dependent $B^0$ events. Table I also shows results of
fits with the samples of non-$CP$ decay modes, where no
statistically significant $CP$ asymmetry is found.

Our measurement of $\sin2\beta$ is consistent with, but im-
proves substantially on the precision of, previous determi-
nations [9]. The central value is consistent with the range
implied by measurements and theoretical estimates of the
magnitudes of CKM matrix elements [10]; it is also con-
sistent with no $CP$ asymmetry at the 1.7$\sigma$ level.

We thank our PEP-II colleagues for their extraordi-
nary achievement in reaching design luminosity and high

<table>
<thead>
<tr>
<th>Category</th>
<th>$w$ (%)</th>
<th>$\Delta w$ (%)</th>
<th>$Q$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lepton</td>
<td>10.9 ± 0.4</td>
<td>11.6 ± 2.0</td>
<td>3.1 ± 3.1</td>
</tr>
<tr>
<td>Kaon</td>
<td>36.5 ± 0.7</td>
<td>17.1 ± 1.3</td>
<td>−1.9 ± 1.9</td>
</tr>
<tr>
<td>NT1</td>
<td>7.7 ± 0.4</td>
<td>21.2 ± 2.9</td>
<td>7.8 ± 4.2</td>
</tr>
<tr>
<td>NT2</td>
<td>13.7 ± 0.5</td>
<td>31.7 ± 2.6</td>
<td>−4.7 ± 3.5</td>
</tr>
<tr>
<td>All</td>
<td>68.9 ± 1.0</td>
<td></td>
<td>26.7 ± 1.6</td>
</tr>
</tbody>
</table>
FIG. 2. The raw asymmetry in the number of $B^0$ and $\bar{B}^0$ tags in the signal region, $(N_{B^0} - N_{\bar{B}^0})/(N_{B^0} + N_{\bar{B}^0})$, with asymmetric binomial errors, as a function of $\Delta t$ for (a) the $J/\psi K^0_S$ and $\phi(2S)K^0_S$ modes ($\eta_f = -1$) and (b) the $J/\psi K^0_L$ mode ($\eta_f = +1$). The solid curves represent the time-dependent asymmetries determined for the central values of $\sin 2\beta$ from the fits for these samples. Eight events that lie outside the plotted interval were also used in the fits. The probability of obtaining a lower likelihood, evaluated using a Monte Carlo technique, is 60%. (c) Variation of the log likelihood as a function of $\sin 2\beta$ for the modes containing $K^0_S$ (dashed curve), the $J/\psi K^0_S$ mode (dotted curve), and the entire sample (solid curve). For the latter, solid lines indicate the central value and values of the log likelihood corresponding to 1 statistical standard deviation.

reliability in a remarkably short time. The collaborating institutions thank SLAC for its support and the kind hospitality extended to them. This work has been supported by the U.S. Department of Energy and National Science Foundation, the Natural Sciences and Engineering Research Council (Canada), the Institute of High Energy Physics (China), Commissariat à l’Energie Atomique and Institut National de Physique Nucléaire et de Physique des Particules (France), Bundesministerium für Bildung und Forschung (Germany), Istituto Nazionale di Fisica Nucleare (Italy), the Research Council of Norway, the Ministry of Science and Technology of the Russian Federation, and the Particle Physics and Astronomy Research Council (United Kingdom). Individuals have received support from the Swiss National Foundation, the A. P. Sloan Foundation, the Research Corporation, and the Alexander von Humboldt Foundation.

*Also with Università di Perugia, Perugia, Italy.
†Also with Università della Basilicata, Potenza, Italy.
‡Deceased.

[6] Throughout this paper, flavor-eigenstate decay modes imply also their charge conjugate.