Measurement of the $B^0\bar{B}^0$ Oscillation Frequency with Inclusive Dilepton Events

The $B^0 - \bar{B}^0$ oscillation frequency $\Delta m_d$ is directly related to the Cabibbo-Kobayashi-Maskawa (CKM) matrix element $|V_{td}|$ [1,2]. Thus, its precise measurement is of fundamental importance; in particular, when combined with the knowledge of the $B_s - \bar{B}_s$ oscillation frequency, it allows a stringent constraint on the Unitarity Triangle of the CKM matrix.

In this Letter, we present a measurement of the time dependence of $B^0 - \bar{B}^0$ mixing using data collected with the BABAR detector at the PEP-II asymmetric energy $e^+ e^-$ collider operated at or near the $Y(4S)$ resonance. The data sample corresponds to an integrated luminosity of 20.7 fb$^{-1}$ on the $Y(4S)$ resonance (on resonance), and 2.6 fb$^{-1}$ collected at 40 MeV lower energies (off resonance). $B\bar{B}$ pairs from the $Y(4S)$ decay move along the high-energy beam direction ($z$) with a nominal Lorentz boost $(\beta \gamma) = 0.55$. Therefore, the two $B$ decay vertices are separated by about 260 $\mu$m on average.
The measurement technique is based on the identification of events containing two leptons from semileptonic decays of $B$ mesons. The flavor of the $B$ mesons at the time of their decay is determined or "tagged" by the charge of the leptons. Thus, for $Y(4S)$ resonance decays into $B^0 \bar{B}^0$ pairs, neglecting backgrounds, opposite-sign (+) and same-sign (−) lepton pairs correspond to unmixed and mixed events, respectively. Because the $B^0 \bar{B}^0$ pair is in a coherent $P$-wave state, the time evolution of the $B$ mesons is a function of the proper time difference $\Delta t$ between the two $B$ decays:

$$S_\pm (\Delta t; \Delta m_d) = \frac{e^{-|\Delta t|/\tau}}{4\tau} (1 \pm \cos \Delta m_d \Delta t),$$

where $\tau$ is the $B^0$ lifetime, and the lifetime difference between the two mass eigenstates is neglected. The corresponding time-dependent asymmetry is $[S_+ (\Delta t) - S_- (\Delta t)]/[S_+ (\Delta t) + S_- (\Delta t)] = \cos \Delta m_d \Delta t$.

This simple picture is modified by the effects of detector resolution and the presence of backgrounds. The most important background, about 55% of the selected sample, is due to $B^+ B^-$ events, which are not removed by the event selection criteria. The fraction of $B^+ B^-$ events is determined from the data itself in order to reduce systematic uncertainties. Other non-negligible backgrounds are leptons from the $b \to c \to \ell$ decay chain (cascade decays), which are also the main source of wrong tags, and hadrons that are misidentified as leptons. Signal and background probability density functions (PDF) for opposite- and same-sign events are included in the full PDF. The corresponding likelihood function, combining opposite- and same-sign dilepton events, is maximized to determine $\Delta m_d$.

The $B\bar{B}$ system is described in detail elsewhere [3]. Charged particle tracking is provided by a five-layer, double-sided silicon vertex tracker (SVT) and a 40-layer drift chamber (DCH), both operating inside a 1.5-T superconducting solenoidal magnet. The CsI(Tl) electromagnetic calorimeter (EMC) detects photons and electrons. Particle identification is provided by a ring-imaging Cherenkov detector (DIRC) and specific ionization measurements $dE/dx$ in the DCH. Muons are identified with the instrumented flux return (IFR), segmented to contain resistive plate chambers.

Events are selected with more than five reconstructed charged tracks, at least three of which must originate from the interaction region and be reconstructed in the DCH, a normalized second Fox-Wolfram moment [4] less than 0.4 and the event aplanarity greater than 0.01, and an invariant mass squared of the event greater than 20 GeV$^2$/c$^4$.

Electrons are selected by requirements on the ratio of the energy deposited in the EMC to the momentum measured in the DCH, the lateral shape of the energy deposition in the EMC, and $dE/dx$ in the DCH. Muons are identified on the basis of the energy in the EMC, as well as the strip multiplicity, track continuity, and penetration depth in the IFR. Lepton candidates consistent with the kaon hypothesis as measured in the DIRC are rejected. Electron (muon) selection efficiencies and misidentification rates at high momentum are about 92% (75%) and 0.15% (3%), respectively.

Electrons from photon conversions are rejected by requirements on the invariant mass and distance of closest approach in combination with all other oppositely charged loosely selected electron candidates. Events with $J/\psi$ or $\psi(2S)$ decays to lepton pairs are rejected by veto of corresponding mass windows, again with looser identification requirements on the second lepton.

Events with at least two leptons are retained and the two highest momentum leptons in the $Y(4S)$ rest frame are used in the following.

The two lepton tracks and a beam spot constraint are used in a vertex fit to find the primary vertex of the event in the transverse plane. The positions of closest approach of the two tracks to this vertex in the transverse plane are computed and their $z$ coordinates are denoted as $z_1$ and $z_2$, where the subscripts 1 and 2 refer to the highest and second highest momentum leptons in the $Y(4S)$ rest frame. The time difference $\Delta t = \Delta z / (\beta \gamma) c$ is obtained from the measured $\Delta z = z_1 - z_2$ and average boost $(\beta \gamma)$. Since the boost is known to good precision, the $\Delta z$ measurement dominates the $\Delta t$ resolution.

To improve the $\Delta z$ and $\Delta t$ resolution, reduce the fraction of incorrectly measured tracks, and minimize related systematic uncertainties, charged tracks are required to satisfy the following criteria. Lepton candidates must have a distance of closest approach to the nominal beam position of less than 1 cm (6 cm) transverse to (along) the beam direction, at least 12 hits in the DCH and four $z$-coordinate hits in the SVT, a momentum between 0.7 and 2.5 GeV/c in the $Y(4S)$ rest frame and between 0.5 and 5.0 mV/c in the laboratory frame, and a polar angle between 0.5 and 2.6 rad in the laboratory frame. The total error on $\Delta z$, computed on an event-by-event basis, is required to be less than 175 $\mu$m. The vertex fit constrains the lepton tracks to originate from the same point in the transverse plane, thereby neglecting the nonzero flight length for $B$ mesons. As a consequence, the $\Delta z$ resolution function is $\Delta z$ dependent, becoming worse at higher $|\Delta z|$. Neglecting this dependence introduces a small bias, discussed below.

The separation between direct leptons and background from cascade decays is achieved with a neural network that combines five discriminating variables for each event and provides two outputs, one for each lepton, chosen to vary between 0 for cascade leptons and 1 for direct leptons. The discriminating variables are the momenta of the two leptons, their opening angle, the total visible energy, and the missing momentum of the event, all computed in the $Y(4S)$ rest frame. The first two variables are very powerful in discriminating between direct and cascade leptons. The third efficiently removes direct-cascade lepton pairs from the same $B$ decay and further reduces contributions from photon conversions. Minimization of the total error on
leads to the requirement that both neural network outputs be greater than 0.8.

The numbers of selected on-resonance and off-resonance events are 99,010 and 428, respectively. The combined requirements give a direct dilepton purity and efficiency of about 83% and 9%, respectively, based on Monte Carlo simulation. Semileptonic $B$ decays in the simulation have been modeled separately for each charm meson involved. Form factors from Heavy-Quark Effective Theory are used for $B 	o D^* \ell \nu$ [5], while current models are used for $B 	o D^{(*)} \pi \ell \nu$ [6], and $B 	o D \ell \nu$ and $B 	o D^{(*)} \ell \nu$ [7]. The measured branching fractions for decays to $D^{(*)}$ and $D^{(*)}\pi$ states are fixed to their world averages [8], and unmeasured processes have inferred rates from isospin arguments. Events from $B\bar{B}$ decays are grouped in three topologies, each of which is assigned its own PDF with different $\Delta t$ dependence and tagging properties.

Direct dilepton events are described by the convolution of an oscillatory term for neutral $B$ decays, or an exponential function for charged $B$ decays, with the resolution function $R$:

$$S^{n(c)}_{\pm} = e^{-|\Delta t|/\tau_{n(c)}} \frac{1 \pm D_{\text{sig}}^{n(c)} \xi_{n(c)}}{4\tau_{n(c)}} \otimes R,$$

for neutral ($n$) and charged ($c$) events, where $\tau_{n(c)}$ is the $B$ meson lifetime, $\xi_n = \cos \Delta m_d \Delta t$, and $\xi_c = 1$. The resolution function $R$ is taken as the sum of three Gaussian distributions, with three widths and two fractions as free parameters. We find that this functional form is a good description of the vertex resolution for leptonic $J/\psi$ decays, both in data and Monte Carlo simulation. The correction factors $D_{\text{sig}}^{n(c)} = 0.95$ account for the (small) fraction of wrongly tagged direct dilepton events. These events are due to hadrons from the $B$ vertex that are misidentified as leptons or leptons from the decay of resonances (e.g., events where only one lepton comes from a $J/\psi$ produced at the $B$ vertex). Both of these sources give almost random tagging and, in the absence of such events, $D_{\text{sig}}^{n(c)}$ would be exactly 1. A small fraction of events of the type $b \to \tau^- \to \ell^-$, which have the correct charge correlation, are also included in the signal topology. Neglecting the $\tau$ lepton lifetime introduces a negligible bias on the $\Delta m_d$ measurement.

Opposite $B$ cascade (OBC) events, 9% of the selected sample, contain one lepton from a $b \to \ell$ decay and one from a $b \to c \to \ell$ decay of the companion $B$ meson. These events are the main source of wrong tags. Their PDFs are modeled by the convolution of $\Delta t$-dependent terms of a form similar to the signal with a resolution model that takes into account the effect of the charmed meson lifetimes by convoluting three Gaussians with a single-sided exponential decay distribution. Since both short-lived $D^0$ and $D_s$, and long-lived $D^+$ mesons, are involved in cascade decays, the global OBC PDFs are

$$C_{\text{OBC}}^{n(c)} = \frac{e^{-|\Delta t|/\tau_{n(c)}}}{4\tau_{n(c)}} \sum_i f_i^{n(c)} (1 \pm D^{n(c)}_{\text{OBC}} \xi_{n(c)}) \otimes R_{\text{OBC}},$$

where the index $i$ runs over the short- and long-lived charm meson components. The two resolution functions $R_{\text{OBC}}$ allow for different effective charm lifetimes and parameters of the three Gaussians, since the resolution function depends on the $B$ and $D$ flight lengths. Because of the different decay processes involved, the relative fractions $f_i^{n(c)}$ of short- and long-lived charm mesons are also different in neutral and charged $B$ events. With the chosen sign convention for $\Delta z$, the sign for the single-sided exponential of the resolution function must be flipped 25% of the time to account for events where the most energetic lepton originates from a cascade decay. If particle identification were perfect and cascade leptons originated only from the $b \to c \to \ell^+$ process, then flavor tagging would always be wrong and the factors $D^{n(c)}_{\text{OBC}}$ would be exactly $-1$. Hadron misidentification (PID) and resonance decays, as well as leptons originating from the $b \to e\tau(\to \ell^+\nu)$ chain, give a fraction of right tags (15%) even in the OBC topology. These two processes have been factorized by writing $D^{n(c)}_{\text{OBC}} = D^{n(c)}_{\text{PID}} \cdot D^{n(c)}_{b\to c\tau}$ and assuming no correlation between the two terms.

Same $B$ cascade (SBC) events, 4% of the selected sample, contain two leptons from a single $B$ meson, obtained via the decay chain $b \to c\ell^-\nu$, with $c \to x\ell^+\nu$. SBC events are insensitive to mixing and, in the case of perfect particle identification and in the absence of resonances, would always give opposite-sign leptons. The PDFs are

$$C_{\text{SBC}}^{n(c)} = \frac{e^{-|\Delta t|/\tau^{n(c)}_{\text{SBC}}}}{4\tau^{n(c)}_{\text{SBC}}} \sum_i f_i^{n(c)} (1 \pm D^{n(c)}_{\text{SBC}}) \otimes R_{\text{SBC}},$$

where $\tau^{n(c)}_{\text{SBC}}$ are effective lifetimes and $D^{n(c)}_{\text{SBC}}$ are corrections for wrong tags in the SBC topology. The resolution $R$ is taken to be the same as for signal events, with no significant bias on the final result.

A small residual background remains (0.3% of the total sample) where both leptons are from an unrecognized $J/\psi$ decay. These are described by a term $\Psi = \delta(\Delta t) \otimes R$, whose normalization is obtained from simulation. Events where one lepton originates from a cascade decay and the other from a $B$ decay to $\tau$ or to a resonance, and events where both leptons come from cascade decays (0.3% of the total sample), are assigned the OBC event topology with no significant bias on $\Delta m_d$.

The fraction $f_{\text{cont}} = 3.4\%$ and $\Delta t$ dependence of the continuum background are determined from off-resonance data. The $\Delta t$ dependence is parametrized for opposite- and same-sign leptons as $Q_{\pm} = \tau_{\text{cont}}^{-1} e^{-\tau_{\text{cont}}|\Delta t|/\tau_{\pm}}$, with $f_+ + f_- = 1$.

The full likelihood function is the product of likelihoods for opposite- and same-sign events, which can be schematically written as
\[ L = (1 - f_{\text{cont}})(1 - f_{J/\psi})(1 - f_c)(f_{\text{sig}}^n S^n + f_{\text{OBC}}^n C_{\text{OBC}}^n + f_{\text{SBC}}^n C_{\text{SBC}}^n) + f_c(f_{\text{sig}}^c S^c + f_{\text{OBC}}^c C_{\text{OBC}}^c + f_{\text{SBC}}^c C_{\text{SBC}}^c) + (1 - f_{\text{cont}})f_{J/\psi} \Psi + f_{\text{cont}} Q. \]

where the $J/\psi$ term and its relative abundance $f_{J/\psi}$ are present for opposite-sign events only, and $f_{\text{sig}}^{n(c)} = (1 - f_{\text{OBC}}^{n(c)} - f_{\text{SBC}}^{n(c)})$. The fraction $f_c$ of charged $B$ events in the selected sample and the OBC fraction $f_{\text{OBC}}^n$ in neutral $B$ events are extracted from the fit. The OBC fraction $f_{\text{OBC}}^n$ in charged $B$ events is scaled with $f_{\text{OBC}}^n$ according to the value of the ratio $f_{\text{OBC}}^c/f_{\text{OBC}}^n$ determined with the Monte Carlo simulation. This parametrization of OBC events significantly reduces the related systematic uncertainty. The SBC fractions are computed for simulated events and fixed in the fit. The various parameters for the OBC resolution functions are taken from a fit to Monte Carlo events. The factor $D_{\text{PID}}^{1,c}$ is fitted and all the other corrections for wrong tags scale with $D_{\text{PID}}^{1,c}$ according to ratios determined with simulated events.

In summary, the values for $\Delta m_d$, $f_c$, $f_{\text{OBC}}^n$, $D_{\text{PID}}^{1,c}$, $f_{J/\psi}^n$, and the widths and relative fractions of the Gaussian components for the signal resolution are determined in the likelihood fit. The $B$ meson lifetimes are fixed to the values quoted in [8].

The result of a binned maximum likelihood fit to the data sample with the requirement $|\Delta t| < 12$ ps yields $\Delta m_d = 0.488 \pm 0.012$ ps$^{-1}$ and $f_c = 0.554 \pm 0.014$. Figures 1a and 1b show the $\Delta t$ distributions for opposite- and same-sign dilepton events, respectively, along with the result of the fit. Figure 1c shows the resulting asymmetry as a function of $\Delta t$. The widths of the three Gaussians for the signal resolution function are $0.55 \pm 0.09$, $1.06 \pm 0.23$, and $4.8 \pm 0.7$ ps, and the corresponding fractions of events are 76%, 22%, and 2%. The probability to obtain a worse fit is 65%, evaluated with an ensemble of data-sized experiments that are generated with a parametrized simulation based on the observed total PDF. The global fit is also performed on a sample from full Monte Carlo simulation, where the fitted results for parameters are consistent with generated values.

The fit result is found to be stable and consistent under a variety of choices for free parameters, where fixed values obtained from Monte Carlo simulation are substituted.

A summary of the systematic uncertainties is given in Table I, where the total is estimated to be 0.0087 ps$^{-1}$. The most important contributions are due to the $B$ meson lifetimes, the $\Delta t$ resolution function, and the modeling of OBC events. Varying the neutral and charged $B$ meson lifetimes independently within their known errors [8] contributes uncertainties of 0.005 and 0.004 ps$^{-1}$, respectively, on $\Delta m_d$.

The systematic error due to the uncertain knowledge of the resolution function for OBC events is estimated by varying the parameters within their errors from the fit to simulated events, including the effect of correlations. A possible scale uncertainty between data and simulation is estimated by allowing a conservative increase of 20% in the OBC resolution width. The overall uncertainty due to the OBC resolution function is 0.0026 ps$^{-1}$. The assumed form for the signal and OBC resolution functions does not incorporate the $\Delta t$ dependence brought about by neglecting the $B$ flight length in determining the vertex separation. The systematic effect introduced by this simplification, as well as the boost approximation, has been studied with large parametrized Monte Carlo samples. For this purpose, the predicted $\Delta t$ dependence is taken from full simulation. Neglecting the $\Delta t$ dependence results in a bias for $\Delta m_d$ of $-0.0045$ ps$^{-1}$. The fit result has been

\begin{table}[h]
\centering
\begin{tabular}{|l|c|}
\hline
Source & $\sigma(\Delta m_d)$ ps$^{-1}$ \\
\hline
$B$ lifetimes & 0.0064 \\
OBC resolution/lifetimes & 0.0026 \\
$\Delta t$ dependence of resolution & 0.0043 \\
z scale and SVT alignment & 0.0020 \\
OBC fractions/wrong tags & 0.0020 \\
Hadron misidentification & 0.0010 \\
$J/\psi$ fraction & 0.0003 \\
Continuum parametrization & 0.0009 \\
Binned fit bias & 0.0006 \\
Beam energy uncertainty & 0.0005 \\
Total & 0.0087 \\
\hline
\end{tabular}
\end{table}
corrected to account for this bias and a corresponding systematic error of 0.0043 ps$^{-1}$ is assigned. Knowledge of the absolute $z$ scale of the detector and the residual uncertainties in the SVT local alignment give a combined error of 0.0020 ps$^{-1}$.

Systematic effects due to the limited knowledge of the parameters of the OBC PDF, which are taken from simulated events, are greatly reduced by fitting the fractions of OBC and the short-lived charm component in neutral $B$ events. The remaining systematic uncertainty (0.0020 ps$^{-1}$) is estimated by varying the otherwise fixed charm-related parameters (the amount of $D_s$, $f_1$ and the various fractions of cascades) by 10%, both coherently and independently. This is a conservative range, given our present knowledge of the physics processes involved.

The ratios between the various wrong-tag factors due to PID are conservatively varied by 30% in the fit. The maximum effect is obtained when the signal and cascade PID wrong-tag corrections are varied in opposite directions. In this case, the total systematic error is 0.0010 ps$^{-1}$.

The uncertainty on the fraction of $J/\psi$ is 30%, which contributes an error on $\Delta m_d$ of 0.0003 ps$^{-1}$. The effective lifetime, the fraction of same-sign events, and the fraction of continuum events are varied independently, giving a combined systematic error of 0.0009 ps$^{-1}$. The dependence of the fit result on the number of bins has been estimated with a parametrized Monte Carlo simulation. A shift of $-0.0006$ ps$^{-1}$ in $\Delta m_d$ is observed and a corresponding correction applied with a systematic error of 0.0006 ps$^{-1}$. The uncertainty (0.1%) on the absolute scale of the beam energies gives an error of 0.0005 ps$^{-1}$.

In conclusion, the neutral $B$ meson oscillation frequency has been measured with an inclusive dilepton sample to be

$$\Delta m_d = 0.493 \pm 0.012 \text{(stat)} \pm 0.009 \text{(syst)} \text{ ps}^{-1}.$$

This result is the single most precise measurement to date and is consistent with a recent BABAR measurement with a fully reconstructed $B^0$ sample [9], as well as the world average of previous measurements [8].

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 thank SLAC for its support and kind hospitality. This work is supported by DOE and NSF (U.S.A.), NSERC (Canada), IHEP (China), CEA and CNRS-IN2P3 (France), BMBF (Germany), INFN (Italy), NFR (Norway), MIST (Russia), and PPARC (United Kingdom). Individuals have received support from the A.P. Sloan Foundation, Research Corporation, and Alexander von Humboldt Foundation.

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