Measurement of the $B^0$ Lifetime with Partially Reconstructed $B^0 \to D^{*-} \ell^+ \nu_\ell$ Decays

The technique of partial reconstruction of $D^{*+}$ mesons (charge conjugate states are always implied), in which only the slow pion from the $D^{*-} \rightarrow D^0 \pi^-$ decay is required, has been widely used in the past [1] to select large samples of reconstructed $B$ mesons. This technique provides a way to measure the combination of Cabibbo-Kobayashi-Maskawa angles ($2\beta + \gamma$) with $B^0 \rightarrow D^{*-} \pi^+$ decays [2]. A sample of about 92,000 events selected from the semileptonic decay $B^0 \rightarrow D^{*-} \ell^- \nu_\ell$ with this partial reconstruction method is used to measure the $B^0$ lifetime, $\tau_{B^0}$. Besides providing a validation of the technique, this analysis results in a precise measurement of $\tau_{B^0}$, whose knowledge is important to test the present understanding of the dynamics of heavy meson decays [3] and to reduce the systematic error in the extraction of fundamental parameters, such as $V_{ub}$ and $V_{cb}$ [4].
The data used in this analysis, recorded by the BABAR detector at the PEP-II storage ring during 1999–2000, correspond to an integrated luminosity of 20.7 fb$^{-1}$ collected on the $Y(4S)$ resonance (on-peak events) and 2.6 fb$^{-1}$ collected 40 MeV below the resonance (off-peak) for background studies. Samples of simulated $b\bar{b}$ events, equivalent in luminosity to the on-peak data, were analyzed through the same analysis chain as the real data. A detailed description of the BABAR detector and the algorithms used for track reconstruction, particle identification, and selection of $b\bar{b}$ events is provided elsewhere [5]; a brief summary is given here. Particles with momenta $p \geq 170$ MeV/c are reconstructed by matching hits in the silicon vertex tracker (SVT) with track elements in the drift chamber (DCH). Lower momentum tracks do not penetrate the DCH and are reconstructed in the SVT alone. Electrons are identified with the ratio of the track momentum to the associated energy deposited in the calorimeter (EMC), the transverse profile of the shower, the energy loss in the drift chamber, and the information from the Cherenkov detector (DIRC). The efficiency for electron identification in the acceptance of the electromagnetic calorimeter is about 90%, with a hadron misidentification probability equal to 0.15%. Muons are required to have a path length and hit distribution in the instrumented flux return and energy deposition in the EMC consistent with that expected for a minimum-ionizing particle. The Cherenkov light emission in the DIRC is then employed to further reject kaons misidentified as muons, by requiring muon candidates to have a kaon hypothesis probability less than 5%. These criteria yield 74% muon efficiency with 2.6% hadron misidentification probability.

Semileptonic $B^0$ decays are then selected by searching for the high momentum charged lepton ($\ell = e, \mu$) from the $B^0$ decay and the slow pion ($\pi_s$) from the $D^{*+} \rightarrow D^0 \pi^+_s$ decay. To reject leptons from semileptonic charm decay and misidentified hadrons, the momentum of the lepton candidate in the $Y(4S)$ rest frame ($p_\ell^*$) is required to be in the range $1.4 < p_\ell^* < 2.3$ GeV/c; that of the $\pi_s$ ($p_{\pi_s}^*$) has to be less than 0.19 GeV/c. The kinematics of the decay are exploited for further background suppression as follows. As a consequence of the limited phase space available in the decay $D^{*-} \rightarrow D^0 \pi^-$, the $\pi_s$ is emitted within a 1-rad-wide cone centered about the $D^{*-}$ direction in the $Y(4S)$ rest frame. The $D^{*-}$ four-momentum can therefore be computed by approximating its direction as that of the $\pi_s$, and parametrizing its momentum as a linear function of the $\pi_s$ momentum, with parameters obtained from the simulation. The neutrino invariant mass can be computed from the four-momenta of the $B^0$, $D^{*-}$, and $\ell$ with the relation

$$M_\nu^2 = (P_{B^0} - P_{D^{*-}} - P_\ell)^2.$$ 

The momentum of the $B^0$ in the $Y(4S)$ rest frame, on average 0.34 GeV/c, is neglected. $M_\nu^2$ peaks approximately at zero for signal events, whereas background events are spread over a wide range.

The $B^0$ decay point is determined from a vertex fit of the $\pi_s$ and $\ell$ tracks, constrained to the beam spot position in the plane perpendicular to the beam axis (the $x$-$y$ plane). The beam spot is determined on a run-by-run basis using two-prong events [5]. Its size in the horizontal direction is 120 $\mu$m. Although the beam spot size in the vertical ($y$) direction is only 5.6 $\mu$m, a beam spot constraint of 50 $\mu$m is applied to account for the flight of the $B^0$ in the $x$-$y$ plane. Only events with vertex fit probability, $P_V$, greater than 0.1% are retained.

A selection is applied on the combined likelihood for $p_\ell^*$, $p_{\pi_s}^*$, and $P_V$, which results in a signal-to-background ratio of about one-to-one in the signal region, $M_\nu^2 > -2$ GeV$^2$/c$^4$. Figure 1 shows the $M_\nu^2$ distribution of data events used to measure $\tau_{B^0}$ when the $\ell$ and the $\pi_s$ have opposite-sign charges. Same-sign events are used as a background control sample. The individual distributions shown in Fig. 1 are obtained by fitting to the data the contributions from continuum events, obtained from the off-peak data, and from $b\bar{b}$ combinatorial background, $B^0$ signal, and $B^+$ resonant background, as predicted by the simulation. The $B^+$ resonant background is due to intermediate production of higher mass charm resonances (denoted as $D^{**}$). The fit determines the composition of the selected sample, which is reported in Table I for the events in the signal region.

The PEP-II collider produces $b\bar{b}$ pairs moving along the beam axis ($z$ direction) with an average Lorentz boost of $\langle \beta \gamma \rangle = 0.55$. Hence, the two $B$ decay vertices are separated on average by $|\Delta z| = 255$ $\mu$m. The position of the $B^0 \rightarrow D^{*-} \ell^+ \nu_\ell$ ("decay") vertex is reconstructed as described above. The decay point of the other $B$ is determined from a selection of the remaining tracks in the...
event using the following criteria. In events that have another lepton with momentum $p^* > 1.1 \text{ GeV}/c$, the other $B$ vertex is computed with only this lepton track constrained to the beam spot in the $x$-$y$ plane. Otherwise, all the tracks with a center-of-mass angle greater than $90^\circ$ with respect to the $\pi_\ell$ direction are considered. This requirement is used to remove most of the tracks from the decay of the $D^{0}$ daughter of the $D^{*-}$, which would otherwise bias the reconstruction of the other $B$ vertex position. Simulation shows that in about 75% of signal events the other vertex has no tracks from the $D^{0}$ decay. The selected tracks are then constrained to the beam spot in the $x$-$y$ plane. The track with the largest contribution to the vertex $\chi^2$, if greater than 6, is removed and the fit iterated until no track fails this requirement. Vertices composed of just one track that is not a high momentum lepton are rejected in order to reduce the number of poorly measured vertices. The lifetime is determined by measuring the quantity $\Delta z = z_{\text{decay}} - z_{\text{other}}$, where $z_{\text{decay}}(z_{\text{other}})$ is the position along the beam line of the decay (other) vertex. The proper time difference is then computed with the relation $\Delta t = \Delta z/(c \beta \gamma)$. A fit with a double Gaussian to the $\Delta t$ residuals in the Monte Carlo simulation shows that one-half of the events are contained in the narrower Gaussian, which has a width of 0.7 ps. The width of the wider Gaussian is 2.3 ps. To remove badly reconstructed vertices, all events for which either $|\Delta z| > 3 \text{ mm}$ or $\sigma_{\Delta z} > 500 \mu\text{m}$ are rejected, where $\sigma_{\Delta z}$ is the uncertainty on $\Delta z$ computed for each event. The last cut removes 3.9% of the events.

$\tau_{b^0}$ is obtained from a binned maximum likelihood fit to the two-dimensional $\Delta t, \sigma_{\Delta t}$ distribution. The $\Delta t$ distribution of signal events, $f(\Delta t, \sigma_{\Delta t}, \tau_{b^0})$, is described by the convolution of the decay probability distribution with the experimental resolution function, which is parametrized by the sum of three Gaussian distributions. The two narrow Gaussians, which account for more than 99% of the events, have the form

$$G[\delta(\Delta t), \sigma_{\Delta t}] = \frac{1}{\sqrt{2\pi} \sigma_{\Delta t}} \exp\left(-\frac{[\delta(\Delta t) - b]^2}{2\sigma_{\Delta t}^2}\right),$$

where $\delta(\Delta t) = \Delta t - \Delta t_{\text{true}}$ is the difference between the measured and the true value of $\Delta t$, $b$ is a bias due to the charm tracks in the other vertex and resolution effects, and the scale factor $S$ accounts for possible misestimation of the calculated error $\sigma_{\Delta t}$ on the proper time difference. The third Gaussian of fixed bias ($-2 \text{ ps}$) and width (8 ps) accounts for badly measured events ("outliers").

$B^+$ background events that peak in the $M_{\ell^2}$ signal region are described by an identical function, with the same resolution parameters as for the $B^0$ signal events, and an effective $B^+$ lifetime of 1.57 ps. This value, obtained by fitting simulated $B^+$ events, is smaller than the value of 1.655 ps generated in the simulation due to tracks from the decay of the $D^{0}$ or the $D^{*+}$ produced in the decay vertex being included in the other vertex.

The $\Delta t$ distribution of continuum background events is modeled as the sum of a nonzero lifetime and a zero-lifetime component, convolved with the same single Gaussian resolution function. The parameters of the resolution function, as well as the lifetime and the fraction of events with nonzero lifetime, are determined with the selected off-peak events.

The $\Delta t$ distribution of the combinatorial $B\bar{B}$ background is modeled as the sum of a nonzero and a zero-lifetime component, with a resolution function that is the sum of three Gaussians. All parameters are determined from the data by fitting the measured $\Delta t$ distribution of the events in the sideband region, $-10 < M_{\ell^2} < -4 \text{ GeV}/c^2$. The Monte Carlo simulation shows, however, that there are small differences in the lifetime and in the fraction of events with nonzero lifetime between the signal region and the sideband. These differences are also observed in the data by separately fitting the signal region and sideband events in the same sign $\ell\pi_\ell$ background control sample. The results from the like-sign fits are used to scale the two background parameters from the sideband to the signal region.

The function used to fit the data is the weighted sum of the four contributions:

$$f(\Delta t, \sigma_{\Delta t}) = \frac{1}{2\tau_{b^0}} \exp\left(-|\Delta t|/\tau_{b^0}\right),$$

where

$$\mathcal{F}(\Delta t, \sigma_{\Delta t}, \tau_{b^0}) = [1 - f_{B^+}(M_{\ell^2}) - f_c(M_{\ell^2}) - f_{B\bar{B}}(M_{\ell^2})] \mathcal{F}_{B^+}(\Delta t, \sigma_{\Delta t}, \tau_{b^0}) + f_{B^+}(M_{\ell^2}) \mathcal{F}_{B^+}(\Delta t, \sigma_{\Delta t}) + f_{B\bar{B}}(M_{\ell^2}) \mathcal{F}_{B\bar{B}}(\Delta t, \sigma_{\Delta t}),$$

with the experimental resolution function, which is parametrized by the sum of three Gaussian distributions. The two narrow Gaussians, which account for more than 99% of the events, have the form

$$G[\delta(\Delta t), \sigma_{\Delta t}] = \frac{1}{\sqrt{2\pi} \sigma_{\Delta t}} \exp\left(-\frac{[\delta(\Delta t) - b]^2}{2\sigma_{\Delta t}^2}\right),$$

where $\delta(\Delta t) = \Delta t - \Delta t_{\text{true}}$ is the difference between the measured and the true value of $\Delta t$, $b$ is a bias due to the charm tracks in the other vertex and resolution effects, and the scale factor $S$ accounts for possible misestimation of the calculated error $\sigma_{\Delta t}$ on the proper time difference. The third Gaussian of fixed bias ($-2 \text{ ps}$) and width (8 ps) accounts for badly measured events ("outliers").

$B^+$ background events that peak in the $M_{\ell^2}$ signal region are described by an identical function, with the same resolution parameters as for the $B^0$ signal events, and an effective $B^+$ lifetime of 1.57 ps. This value, obtained by fitting simulated $B^+$ events, is smaller than the value of 1.655 ps generated in the simulation due to tracks from the decay of the $D^{0}$ or the $D^{*+}$ produced in the decay vertex being included in the other vertex.

The $\Delta t$ distribution of continuum background events is modeled as the sum of a nonzero lifetime and a zero-lifetime component, convolved with the same single Gaussian resolution function. The parameters of the resolution function, as well as the lifetime and the fraction of events with nonzero lifetime, are determined with the selected off-peak events.

The $\Delta t$ distribution of the combinatorial $B\bar{B}$ background is modeled as the sum of a nonzero and a zero-lifetime component, with a resolution function that is the sum of three Gaussians. All parameters are determined from the data by fitting the measured $\Delta t$ distribution of the events in the sideband region, $-10 < M_{\ell^2} < -4 \text{ GeV}/c^2$. The Monte Carlo simulation shows, however, that there are small differences in the lifetime and in the fraction of events with nonzero lifetime between the signal region and the sideband. These differences are also observed in the data by separately fitting the signal region and sideband events in the same sign $\ell\pi_\ell$ background control sample. The results from the like-sign fits are used to scale the two background parameters from the sideband to the signal region.

The function used to fit the data is the weighted sum of the four contributions:

$$f(\Delta t, \sigma_{\Delta t}) = \frac{1}{2\tau_{b^0}} \exp\left(-|\Delta t|/\tau_{b^0}\right),$$

where

$$\mathcal{F}(\Delta t, \sigma_{\Delta t}, \tau_{b^0}) = [1 - f_{B^+}(M_{\ell^2}) - f_c(M_{\ell^2}) - f_{B\bar{B}}(M_{\ell^2})] \mathcal{F}_{B^+}(\Delta t, \sigma_{\Delta t}, \tau_{b^0}) + f_{B^+}(M_{\ell^2}) \mathcal{F}_{B^+}(\Delta t, \sigma_{\Delta t}) + f_{B\bar{B}}(M_{\ell^2}) \mathcal{F}_{B\bar{B}}(\Delta t, \sigma_{\Delta t}),$$

where the functions $\mathcal{F}_{B^+}, \mathcal{F}_{B\bar{B}}, \mathcal{F}_c$, and $\mathcal{F}_{B\bar{B}}$ describe the measured decay time difference distributions for the signal, peaking $B^+$, continuum, and $B\bar{B}$ combinatorial background, respectively. $f_{B^+}, f_c,$ and $f_{B\bar{B}}$ are the probabilities that the event is from the $B^+$, continuum, or $B\bar{B}$ background, computed for each event on the basis of the measured value of $M_{\ell^2}$. Along with $\tau_{b^0}$, the scale factor of the first Gaussian, $S_1 = 1.02 \pm 0.02$, the scale factor of the second Gaussian,
$S_2 = 2.4 \pm 0.1$, the bias of the first Gaussian, $b_1 = -0.120 \pm 0.009$ ps, and the fraction of outliers, $f_0 = (0.2 \pm 0.1)$% are obtained from the fit. The fraction of events contained in the second Gaussian, $f_2$, and its bias $b_2$ are fixed to 7% and $-0.85$ ps, respectively.

The result of the fit is $\tau_{B^0}^{\text{raw}} = 1.482 \pm 0.012$ ps, where the error is statistical only. Figure 2 shows the comparison between the measured $\Delta t$ distribution and the fit result. The probability of obtaining a lower likelihood if the assumed probability distributions correctly describe the data is evaluated to be 18% with a Monte Carlo technique. This raw lifetime must be corrected for the bias induced by the tracks from the $D^0$ that are not rejected by the $\pi_s$ cone cut. A multiplicative correction factor of $R_{\tau_{B^0}} = 1.032 \pm 0.007(\text{stat}) \pm 0.007(\text{syst})$ is computed from the simulation. The statistical error arises from the number of simulated events. The dominant systematic uncertainty corresponds to the full variation in $R_{\tau_{D^0}} (0.66\%)$ obtained by smearing the $\Delta t$ resolution in the simulation to match that in the data. A second systematic uncertainty is computed by comparing in data and simulation the fraction of charged tracks from $D^0$ decays outside the $\pi_s$ cone for a subset of events in which the $D^0$ is fully reconstructed in the $K^- \pi^-, K^+ \pi^- \pi^0$, and $K^+ \pi^- \pi^+ \pi^-$ final states. The maximum discrepancy between data and simulation corresponds to a variation of $\pm 0.24\%$ in the value of $R_{\tau_{D^0}}$. The corrected value of the $B^0$ lifetime is then

$$\tau_{B^0} = \tau_{B^0}^{\text{raw}} R_{\tau_{D^0}} = 1.529 \pm 0.012 \text{ ps}.$$

The systematic error on $\tau_{B^0}$ is computed by adding in quadrature the contributions from several sources, described below and summarized in Table II.

The fractions of $B^+$, continuum, and combinatorial $B\bar{B}$ events are varied by the uncertainties obtained from the $M^2_t$ fit (see Table I). The parameters of the continuum and combinatorial $B\bar{B} \Delta t$ distributions are varied by their uncertainties, accounting for their correlations. As described above, the fraction of events with nonzero lifetime and the lifetime of the combinatorial $B\bar{B}$ background computed from the sideband, $(82.0 \pm 0.9)$% and $1.412 \pm 0.013$ ps, respectively, are corrected with the same-sign control sample. This method is validated by a simulation study, and the statistical error of the validation is included in the background systematic error. The effective $B^+$ lifetime is varied by $\pm 3\%$, which is the sum in quadrature of the world average error on the $B^+$ lifetime and the statistical and systematic uncertainties of the $D^0$ bias correction.

The parameters of the signal resolution function that are not determined in the fit to the data are varied within conservative ranges ($f_2$ between 0.03 and 0.13, and $b_2$ between $-1.5$ and 0 ps). Several different analytical expressions are used to represent the small fraction of outliers. The fit is also performed by allowing the scale factor and the bias of the narrow Gaussians to depend linearly on $\sigma_{\Delta t}$ or on the lepton polar angle. The maximum change from the nominal fit is taken as the systematic error due to the parametrization of the resolution function.

The bias due to the event selection is found to be compatible with zero from a study of the true proper

![FIG. 2. $\Delta t$ distribution for selected events in the data (points) in linear (upper) and logarithmic (lower) scale. The lifetime fit result is superimposed on the data. The hatched histograms show the contributions from the background sources described in the text.](image)

<table>
<thead>
<tr>
<th>Source</th>
<th>$\sigma_{\tau_{B^0}} / \tau_{B^0}$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuum fraction and parametrization</td>
<td>0.36</td>
</tr>
<tr>
<td>$B\bar{B}$ fraction and parametrization</td>
<td>0.68</td>
</tr>
<tr>
<td>$B^+$ fraction and parametrization</td>
<td>0.64</td>
</tr>
<tr>
<td>Resolution model</td>
<td>1.14</td>
</tr>
<tr>
<td>Event selection bias</td>
<td>0.30</td>
</tr>
<tr>
<td>$D^0$ bias ($R_{\tau_{D^0}}$)</td>
<td>0.95</td>
</tr>
<tr>
<td>Bias due to charm from the other $B$</td>
<td>0.21</td>
</tr>
<tr>
<td>$z$ scale</td>
<td>0.40</td>
</tr>
<tr>
<td>Total</td>
<td>1.89</td>
</tr>
</tbody>
</table>
The time difference of signal events selected in the simulation. The statistical precision of this test is taken as a systematic error.

The statistical and systematic errors on $R_{D^0}$ are propagated to the final error. A possible bias induced by the presence of tracks from charm decays produced by the other $B$ meson is investigated in the simulation by varying within their uncertainties the relative fractions of charmless, single charm, and double charm events, as well as the relative fractions of $D^+, D^0, D_s$, and $\Lambda_c$ hadrons. The $\Delta t$ length scale is determined with an uncertainty of 0.4% from secondary interactions with a beam pipe section of known length. The result shows no significant dependence on selection criteria (angular width of the $\pi_s$ cone used to reject $D^0$ tracks, soft pion momentum, lepton momentum, polar and azimuthal angle, and alignment conditions). No difference in the result is observed if $\tau_{B^0}$ is determined with an unbinned maximum likelihood fit. A final relative error of $\pm 1.9\%$ is found by adding in quadrature the uncertainties from the above sources, as listed in Table II.

In conclusion, a sample of about 92,000 $B^0 \rightarrow D^{*+} \ell^+ \nu_\ell$ decays is selected by partial reconstruction of the $D^{*-} \rightarrow D^0 \pi^- \pi^+$ decay. It is used for a measurement of the $B^0$ lifetime. The value obtained,

$$\tau_{B^0} = 1.529 \pm 0.012(\text{stat}) \pm 0.029(\text{syst}) \text{ ps},$$

is consistent with a recent $BABAR$ measurement [6] and with the world average [4]. It is currently the most precise single measurement of this quantity.

We are grateful for the excellent luminosity and machine conditions provided by our PEP-II colleagues. 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 (Germany), INFN (Italy), NFR (Norway), MIST (Russia), and PPARC (United Kingdom). Individuals have received support from the Swiss NSF, 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.


