Measurement of the $B \to X_s \ell^+\ell^-$ Branching Fraction with a Sum over Exclusive Modes

We measure the branching fraction for the flavor-changing neutral-current process $B \rightarrow X_s \ell^+ \ell^-$ with a sample of $89 \times 10^6 \ U(4S) \rightarrow B \overline{B}$ events recorded with the BABAR detector at the PEP-II $e^+e^-$ storage ring. The full reconstruction method exploits the strong kinematic discrimination provided by $
abla m_{ES} = \sqrt{E_{beam}^2 - p_B^2}$ and $\Delta E = E_B - E_{beam}$, where $E_{beam}$ is the beam energy and $E_B(\vec{p_B})$ is the reconstructed $B$-meson energy (three-momentum). These quantities are evaluated in the $e^+e^-$ center-of-mass (CM) frame.

The rare decay $B \rightarrow X_s \ell^+ \ell^-$, which proceeds through the $b \rightarrow s \ell^+ \ell^-$ transition, is interesting because of its rate and charge asymmetry. With loose selection criteria, the average number of candidates per event is five in the signal simulation. $B$ candidates are reconstructed by first selecting the $e^+e^-$ or $\mu^+\mu^-$ pair with the largest value of $|p_{T\ell}$| + $|p_{T\ell}|$. Then, $B \rightarrow X_s \ell^+ \ell^-$ candidates are formed by adding any of the following hadronic topologies: $K^+, K^+\pi^0, K^+\pi^-, K^+\pi^-\pi^0$, $K^+\pi^-\pi^+\pi^0, K^{*0}_0\pi^0, K^{*0}_0\pi^+\pi^-, K^{*0}_0\pi^+\pi^-\pi^0, K^{*0}_0\pi^+\pi^-\pi^0$, and $K^{*0}_0\pi^+\pi^-\pi^0$. We name one member of a pair of charge-conjugate states to refer to both, unless we specify otherwise. A limit is imposed on the number of pions because the expected signal-to-background ratio drops significantly with increasing multiplicity.

With loose selection criteria, the average number of $B$ candidates per event is five in the signal simulation. A likelihood function for the signal is constructed based on the simulated distributions of $\Delta E$, $\log(F_{Bvert})$, and $\cos \theta_B$, where $F_{Bvert}$ is the fit probability for the $B$ vertex constructed from charged daughter particles and $\theta_B$ is the angle between $\vec{p_B}$ and the $e^{-}$ beam axis. We select the candidate with the largest signal likelihood before further cuts are applied. This approach entails a loss of 8% in signal efficiency but reduces the overall background.

The rare decay $B \rightarrow X_s \ell^+ \ell^-$ is motivated by smaller theoretical uncertainties. Exclusive rates to be reconstructed. If the fraction of modes physics in these loops would contribute at the same or lower order as the SM [1, 2]. Recent SM calculations of the inclusive branching fractions predict $B(B \rightarrow X_s e^+ e^-) = (6.9 \pm 1.0) \times 10^{-6}$ for $m_{e^+e^-} > 0.2$ GeV/$c^2$ and $B(B \rightarrow X_s \mu^+ \mu^-) = (4.2 \pm 0.7) \times 10^{-6}$ [1, 3]. Although the branching fraction measurement for inclusive decays is more challenging than for exclusive decays, it is motivated by smaller theoretical uncertainties. Exclusive $B \rightarrow K^{*0} \ell^+ \ell^-$ ($\ell = e, \mu$) decays have been observed by Belle and BABAR [4, 5]. Belle has also reported a measurement of the inclusive branching fraction [6].

The data sample used in this analysis was collected with the BABAR detector [7] at the PEP-II asymmetric-energy $e^+e^-$ storage ring at the Stanford Linear Accelerator Center. The sample consists of $89 \times 10^6 \ B \overline{B}$ events recorded at the $\Upsilon(4S)$, corresponding to an integrated luminosity of 81.9 fb$^{-1}$.

We study the $B \rightarrow X_s \ell^+ \ell^-$ process by fully reconstructing a subset of all possible final states. The reconstructed hadronic systems $X_s$ consist of one $K^{\pm}$ or $K^{*0}$ and up to two pions, with at most one $\pi^0$. This approach allows approximately half of the total inclusive rate to be reconstructed. If the fraction of modes containing a $K^{*0}_{0}$ is assumed equal to that containing a $K^{*0}_{s}$, the missing states represent $\sim30\%$ of the total rate. To compute the inclusive branching fraction, we account for missing modes and selection efficiencies using a $B \rightarrow X_s \ell^+ \ell^-$ decay model constructed as follows. For $m_{X_s} < 1.1$ GeV/$c^2$, exclusive $B \rightarrow K^{(*)} \ell^+ \ell^-$ decays are generated with a $b \rightarrow s \ell^+ \ell^-$ decay model according to [1, 8]. The remaining decays, for $m_{X_s} > 1.1$ GeV/$c^2$, are generated according to a quark-level calculation [1, 9] and the $b$-quark Fermi motion model of [10]. JETSET [11] is then used to hadronize the system consisting of a strange quark and a spectator quark.

The full reconstruction method exploits the strong kinematic discrimination provided by $m_{ES} = \sqrt{E_{beam}^2 - p_B^2}$ and $\Delta E = E_B - E_{beam}$, where $E_{beam}$ is the beam energy and $E_B(\vec{p_B})$ is the reconstructed $B$-meson energy (three-momentum). These quantities are evaluated in the $e^+e^-$ center-of-mass (CM) frame.

Clean identification of the $B$ decay products is important for minimizing the backgrounds. Electron candidates are required to have a laboratory-frame momentum $p_{\ell} > 0.5$ GeV/$c$ and are identified with measurements from the tracking systems and the electromagnetic calorimeter. Bremsstrahlung photons are recovered by combining an electron with up to three photons within a small angular region around the electron direction [12]. Muon candidates are required to have $p_{\ell} > 1.0$ GeV/$c$ and are identified as penetrating charged particles in the instrumented flux return. Charged kaon identification relies on measurements performed with the Cherenkov ring-imaging detector and the tracking systems. The $K^{*0}_{s}$ candidates are required to have $m_{e^+e^-} - m_{K^{*0}_{s}} > 11.2$ MeV/$c^2$, a decay length greater than 2 mm, and $\cos \delta > 0.99$, where $\delta$ is the angle between the $K^{*0}_{s}$ momentum vector and a line that connects the primary vertex with the $K^{*0}_{s}$ vertex. Charged tracks that do not satisfy tight $e^\pm$ or $K^{\pm}$ identification are considered to be pions. Neutral pions are required to have a laboratory-frame energy greater than 400 MeV, a photon daughter energy greater than 50 MeV, and $m_{\pi^+\pi^-} < 10$ MeV/$c^2$.

The $B$ candidates are reconstructed by first selecting the $e^+e^-$ or $\mu^+\mu^-$ pair with the largest value of $|p_{T\ell}| + |p_{T\ell}|$. Then, $B \rightarrow X_s \ell^+ \ell^-$ candidates are formed by adding any of the following hadronic topologies: $K^+, K^+\pi^0, K^+\pi^-, K^+\pi^-\pi^0, K^+\pi^-\pi^+\pi^0, K^{*0}_0\pi^0, K^{*0}_0\pi^+, K^{*0}_0\pi^+\pi^0, K^{*0}_0\pi^+\pi^-\pi^0$. (We name one member of a pair of charge-conjugate states to refer to both, unless we specify otherwise.) A limit is imposed on the number of pions because the expected signal-to-background ratio drops significantly with increasing multiplicity.
by 34% in the final sample.

Combinatorial backgrounds from nonsignal $B \overline{B}$ and continuum $e^+e^- \rightarrow q \bar{q}$ (where $q = u, d, s, c$) events are reduced by requiring $m_{X} < 1.8 \text{ GeV}/c^2$, $5.20 < m_{ES} < 5.29 \text{ GeV}/c^2$, and $-0.2 < \Delta E < 0.1 \text{ GeV}$. We impose a limit on $m_{X}$ because the signal-to-background ratio drops as $m_{X}$ increases.

The dominant background producing a peak in $m_{ES}$ at the $B$ mass originates from $B \rightarrow J/\psi X$, $\psi(2S)X$ decays with $J/\psi \rightarrow \ell^+\ell^-$ and $\psi(2S) \rightarrow \ell^+\ell^-$. This charmonium background is suppressed by vetoing $B$ candidates with dilepton mass in the ranges $2.70 < m_{e^+e^-} < 3.25$, $2.80 < m_{\mu^+\mu^-} < 3.20$, $3.45 < m_{e^+e^-} < 3.80$, and $3.55 < m_{\mu^+\mu^-} < 3.80 \text{ GeV}/c^2$. In the electron channel, the veto is applied before and after bremsstrahlung recovery to allow for imperfect recovery. The potential peaking background from $B \rightarrow X_s \gamma$ decays with conversion of the photon into an $e^+e^-$ pair in the detector material is removed by requiring $m_{e^+e^-} > 0.2 \text{ GeV}/c^2$. This cut also serves to reduce contributions from $b \rightarrow s e^+e^-$ transitions in which the $e^+e^-$ pair originates from a photon.

The final suppression of the combinatorial background is achieved with a likelihood based on nine variables: (i) $\Delta E$, (ii) $\Delta E_{ROE}$, (iii) $m_{ES}^{ROE}$, where $ROE$ refers to the rest of the event (all charged tracks and photon candidates not included in the $B$ candidate), (iv) the separation $\Delta z$ between the two leptons along the beam direction measured at their points of closest approach to the beam axis, (v) $\log(P_{B_{max}})$, (vi) $\cos \theta_{miss}$, where $\theta_{miss}$ is the angle between the missing momentum vector for the whole event and the $z$ axis in the CM frame, (vii) $\cos \theta_B$, (viii) $|\cos \theta_T|$, where $\theta_T$ is the angle between the thrust axes of the $B$ candidate and the ROE in the CM frame, and (ix) the ratio $R_2$ of the second and zeroth-order Fox-Wolfram moments [13]. The variables $\Delta E$, $\Delta E_{ROE}$, and $m_{ES}^{ROE}$ are most effective at rejecting $B \overline{B}$ background, especially for events with two semileptonic decays that have large missing energy. The event-shape variables $|\cos \theta_T|$ and $R_2$ are most effective at suppressing continuum events.

A likelihood value is computed as the product of nine $R$-nels in the regions $m_{X} < 0.6$, $0.6 < m_{X} < 1.1$, and $1.1 < m_{X} < 1.8 \text{ GeV}/c^2$, and results in progressively harder $L_R$ cuts for increasing $m_{X}$.

After applying all selection criteria, we obtain a sample of 349 (222) events in the electron (muon) channel. According to the simulation, the remaining background consists mostly of $B \overline{B}$ events.

We perform an extended unbinned maximum-likelihood fit to the $m_{ES}$ distribution to extract the signal yield $N_{sig}$ as well as the combinatorial background shape and yield. The likelihood function consists of four components: signal, charmonium peaking background, hadronic peaking background (see below), and combinatorial background.

The shapes of the signal and charmonium peaking background are described by the same Gaussian PDF, with a width of 2.80, 2.61, and 2.74 MeV/c$^2$ in the electron, muon, and electron+muon channels, respectively. The Gaussian mean and width are determined from signal-like $B \rightarrow J/\psi X$, $\psi(2S)X$ decay candidates satisfying all selection criteria but failing the charmonium veto (charmonium-veto sample). The charmonium peaking background is estimated from the simulation to contribute $0.4 \pm 0.2$ (1.4 $\pm$ 0.5) events in the electron (muon) channel. The size and shape of the hadronic peaking background component arising from $B \rightarrow D^{(*)+}n\pi (n > 0)$ decays with misidentification of two charged pions as leptons, are derived directly from data by performing the analysis without the lepton identification requirements. Taking the $\pi \rightarrow e$ misidentification rates into account, we estimate the hadronic peaking background to be $< 0.1$ and $2.4 \pm 0.8$ events in the electron and muon channels, respectively. The error on the misidentification rate derived from data control samples dominates the uncertainty. The PDF for the combinatorial background arising from continuum and $B \overline{B}$ events is an ARGUS function [14] with an endpoint equal to $E_{beam}$.

The fit results are presented in Fig. 1(a) and Table I. The statistical significance is $S = \sqrt{2 \ln(L_{max}/L_0)}$, where $L_{max}$ is the maximum likelihood for the fit and $L_0$ is that for a fit with signal yield fixed to zero. The $B \rightarrow X_s \ell^+\ell^-$ signal yield is obtained from a fit to the combined electron and muon data. Figure 1(b) shows the $m_{ES}$ distribution of $B \rightarrow X_s e^+\mu^-$ candidates selected using the nominal criteria but requiring leptons of different flavor. An ARGUS function fits this distribution well, consistent with that sample being pure combinatorial background.
TABLE I: Signal yield, significance, efficiency, and branching fraction in the electron and muon channels, as well as for the electron+muon average (first, second, and third rows, respectively). For the signal yield, the first error is statistical and the second error is systematic. For the signal efficiency, the first error corresponds to the experimental systematic uncertainty arising from detector modeling, hadronization, and Monte Carlo statistics, whereas the second error corresponds to the uncertainties in the signal model. For the branching fraction, the errors correspond to statistical, experimental systematic, and signal model systematic uncertainties.

<table>
<thead>
<tr>
<th></th>
<th>$N_{s_{\text{sig}}}$</th>
<th>$S$</th>
<th>$\epsilon$ (%)</th>
<th>$B$ ($10^{-3}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>29.2</td>
<td>8.3 ± 1.3</td>
<td>4.0</td>
<td>2.74 ± 0.27 ± 0.49</td>
<td>6.0 ± 1.7 ± 0.7 ± 1.1</td>
</tr>
<tr>
<td>11.2</td>
<td>6.2 ± 0.9</td>
<td>2.0</td>
<td>1.26 ± 0.12 ± 0.25</td>
<td>5.0 ± 2.8 ± 0.6 ± 1.0</td>
</tr>
<tr>
<td>40.1</td>
<td>10.4 ± 1.7</td>
<td>4.3</td>
<td>2.00 ± 0.19 ± 0.37</td>
<td>5.6 ± 1.5 ± 0.6 ± 1.1</td>
</tr>
</tbody>
</table>

FIG. 2: Differential branching fraction as a function of (a) hadronic mass and (b) dilepton mass, averaged over electron and muon channels for data (points) and signal Monte Carlo (histogram). The outer (inner) error bars correspond to the total (statistical) uncertainties.

The branching fraction is $B = N_{s_{\text{sig}}} / (2N_{B\overline{B}} \epsilon)$, where $N_{B\overline{B}} = (88.9 \pm 1.0) \times 10^6$ is the number of $B\overline{B}$ pairs and $\epsilon$ is the signal efficiency. The yield $N_{s_{\text{sig}}}$ includes a contribution from events containing a true $B \rightarrow X_s \ell^+\ell^-$ decay but having a selected $B$ candidate with incomplete or wrong decay products (cross-feed events). We estimate the cross-feed contribution to the signal yield to be $1.5 \pm 1.5 (0.6 \pm 0.6)$ events in the electron (muon) channel by including in the fit a component whose shape and normalization are taken from simulation. The corresponding uncertainties are estimated with an ensemble of simulated data-sized experiments. The signal efficiency is adjusted to reflect this contribution. Figure 2 shows the $B \rightarrow X_s \ell^+\ell^-$ differential branching fraction as a function of $m_{X_s}$ and $m_{\ell^+\ell^-}$, obtained by applying the nominal likelihood fit procedure to the data in bins of $m_{X_s}$ and $m_{\ell^+\ell^-}$. We use the nominal Gaussian shape for all bins, as we found no significant shape dependence on $m_{X_s}$ or $m_{\ell^+\ell^-}$.

We evaluate the systematic uncertainties in the signal yield by varying the signal Gaussian parameters (mean and width) and the signal shape (using asymmetric signal shapes) within the constraints allowed by the charmonium-veto sample. The amount of charmonium and hadronic peaking background is varied, as well as the shape of the latter.

Uncertainties in the signal efficiency originate from the detector modeling and from the simulation of signal decays. From control data samples we find uncertainties of 1.3% per track (0.8% per $\pi^+$) in the tracking efficiency; 0.85% per electron, 1.55% per muon, 1.0% per $K^+$, and 1.4% per $\pi^+$ in the charged-particle identification efficiency; 1.5% per $K^0_S$ and 5.6% per $\pi^0$ in the reconstruction efficiency. We check the efficiency of the cut in $L_R$ with the charmonium-veto sample and take the discrepancy with the simulation (3.7%) as the uncertainty. The fraction of signal cross feed included in the signal Gaussian is varied by ±100%. Altogether, the uncertainty from detector modeling is 8.3%.

The dominant source of uncertainty (18%) arises from the decay model. The fractions of $B \rightarrow K\ell^+\ell^-$ and $B \rightarrow K^*\ell^+\ell^-$ decays are both increased (or decreased) together by the theoretical uncertainties [1], resulting in the largest contribution (16%). Parameters of the Fermi motion model are varied within limits allowed by measurements of hadronic moments in semileptonic B decays [15] and the photon spectrum in inclusive $B \rightarrow X_s \gamma$ decays [16]. The transition point (in $m_{X_s}$) between the $B \rightarrow K^*\ell^+\ell^-$ and $b \rightarrow s \ell^+\ell^-$ decay models is varied by ±0.1 GeV/c².

Hadronization uncertainties affecting the region $m_{X_s} > 1.1$ GeV/c² total 4.9% and are evaluated as follows. The ratio between the generator yield for decay modes containing a $K^0_S$ and that for modes containing a charged kaon is varied within the range 0.50 ± 0.05, to allow for isospin violation. Similarly, the ratio between modes with one and no $\pi^0$ is varied within the range 1.0 ± 0.5, and the ratio between two-body and three-body hadronic systems is varied within the range 0.5 ±0.3. Uncertainties in the last two ratios are set by the level of discrepancy between data and simulation as measured in [17]. The uncertainty in the fraction of modes with pion or kaon multiplicities different from those used in this analysis, or with photons that do not originate from $\pi^0$ decays but rather from $\eta, \eta', \omega$, etc, is estimated by varying these different fractions by ±50%.

Table I summarizes the results of the analysis. For the combined $B \rightarrow X_s \ell^+\ell^-$ branching fraction we assume equal branching fractions in the electron and muon channels, and average over both channels. Table II shows the partial branching fractions in several dilepton and hadronic mass ranges.

We search for CP violation in the $B \rightarrow X_s \ell^+\ell^-$ decay by performing separate fits to $B$ and $\overline{B}$ final states, where the final state flavor is determined by the kaon and pion charges (modes with $X_s = K^0_S, K^0_S\pi^0$, or $K^0_S\pi^+\pi^-$ are not
TABLE II: Partial branching fractions in bins of dilepton and hadronic mass averaged over electrons and muons. The first error is statistical and the second error is systematic.

<table>
<thead>
<tr>
<th>$m_{\ell^+\ell^-}$ (GeV/$c^2$)</th>
<th>$B$ ($10^{-6}$)</th>
<th>$m_{X_s}$ (GeV/$c^2$)</th>
<th>$B$ ($10^{-6}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2 – 1.0</td>
<td>0.08 ± 0.56 &amp; 0.06</td>
<td>0.4 – 0.6</td>
<td>0.53 ± 0.17 &amp; 0.04</td>
</tr>
<tr>
<td>1.0 – 2.0</td>
<td>1.6 ± 0.6 &amp; 0.5</td>
<td>0.6 – 0.8</td>
<td>0.32 ± 0.24 &amp; 0.04</td>
</tr>
<tr>
<td>2.0 – $m_{J/\psi}$</td>
<td>1.8 ± 0.8</td>
<td>0.8 – 1.0</td>
<td>0.64 ± 0.40 &amp; 0.08</td>
</tr>
<tr>
<td>$m_{J/\psi} - m_{\ell^+\ell^-}$</td>
<td>1.0 ± 0.8 &amp; 0.2</td>
<td>1.0 – 1.8</td>
<td>3.5 ± 2.1 &amp; 0.7</td>
</tr>
<tr>
<td>$m_{\ell^+\ell^-} - 5.0$</td>
<td>0.64 ± 0.32 &amp; 0.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0 – 2.45</td>
<td>1.8 ± 0.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.8 – 5.0</td>
<td>0.50 ± 0.25 &amp; 0.07</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

used). We find $N_{\ell\ell}^P = 14.7 ± 6.5$ (stat) and $N_{\ell\ell}^B = 22.9 ± 7.4$ (stat), corresponding to an asymmetry $A_{CP} \equiv (N_{\ell\ell}^P - N_{\ell\ell}^B)/(N_{\ell\ell}^P + N_{\ell\ell}^B) = -0.22 ± 0.26$ (stat). The systematic uncertainty is taken to be the statistical uncertainty in the asymmetry measured in the charmonium-veto sample $A_{CP}^{QCD} = -0.005 ± 0.016$ (stat). In summary, we observe a signal of 40 ± 10 (stat) ± 2 (syst) $B \rightarrow X_s \ell^+\ell^-$ events with a statistical significance of 4.3 $\sigma$. The corresponding branching fraction $B(B \rightarrow X_s \ell^+\ell^-) = (5.6 ± 1.5$ (stat) ± 0.6 (exp syst) ± 1.1 (model syst)) $\times 10^{-6}$ for $m_{\ell^+\ell^-} > 0.2$ GeV/$c^2$ agrees well with predictions [1] and a previous measurement [6]. In restricted dilepton mass ranges below and above the charmonium region (see bottom of Table II), the partial branching fractions also agree with predictions [18]. At low hadronic mass, the partial branching fractions are consistent with existing $B \rightarrow K^{(*)}\ell^+\ell^-$ measurements [4, 5]. We determine the direct CP asymmetry to be $A_{CP} = -0.22 ± 0.26$ (stat) ± 0.02 (syst), in agreement with a predicted asymmetry of $(0.19 ± 0.15) \times 10^{-2}$ in the SM [19]. Overall, we find no evidence for contributions from physics beyond the SM.

We thank Gudrun Hiller and Tobias Hurth for helpful discussions. 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 the A. P. Sloan Foundation, Research Corporation, and Alexander von Humboldt Foundation.