Improved measurement of the electroweak penguin process $B \to X_s \ell^+ \ell^-$

M. Iwasaki,[43] K. Itoh,[43] H. Aihara,[43] K. Abe,[8] K. Abe,[41] I. Adachi,[8] Y. Asano,[47] T. Aushev,[12] S. Bahinipati,[4] A. M. Bakich,[38] S. Banerjee,[39] I. Bedny,[1] U. Bitenc,[13] I. Bizjak,[13] S. Blyth,[25] A. Bondar,[1] A. Bozek,[26] M. Bračko,[8,19,13] J. Brodzicka,[26] T. E. Browder,[7] M.-C. Chang,[25] P. Chang,[25] Y. Chao,[25] A. Chen,[23] K.-F. Chen,[25] W. T. Chen,[23] B. G. Cheon,[3] R. Chistov,[12] S.-K. Choi,[6] Y. Choi,[37] A. Chuvikov,[33] J. Dalseno,[20] M. Danilov,[12] M. Dash,[48] A. Drutskoy,[4] S. Eidelman,[1] Y. Enari,[21] S. Fratina,[13] N. Gabyshev,[1] T. Gershon,[8] G. Gokhroo,[39] B. Golob,[18,13] A. Gorišek,[13] J. Haba,[8] T. Hara,[30] N. C. Hastings,[8] K. Hayasaka,[21] H. Hayashii,[22] M. Hazumi,[8] T. Hokuue,[21] Y. Hoshi,[41] S. Hou,[23] W.-S. Hou,[25] Y. B. Hsiung,[25] T. Iijima,[21] A. Imoto,[22] K. Inami,[21] A. Ishikawa,[8] H. Ishino,[44] R. Itoh,[8] J. H. Kang,[49] J. S. Kang,[15] N. Katayama,[8] H. Kawai,[2] T. Kawasaki,[28] H. R. Khan,[44] H. Kichimi,[8] H. J. Kim,[16] S. K. Kim,[36] S. M. Kim,[37] K. Kinoshita,[4] P. Križan,[18,13] P. Krokovny,[1] R. Kulasiri,[4] S. Kumar,[31] C. C. Kuo,[23] A. Kuzmin,[1] Y.-J. Kwon,[49] J. S. Lange,[5] G. Leder,[11] S. E. Lee,[36] T. Lesiak,[26] J. Li,[35] S.-W. Lin,[25] D. Liventsev,[12] J. MacNaughton,[11] G. Majumder,[39] F. Mandl,[11] T. Matsumoto,[45] A. Matyja,[26] Y. Mikami,[42] W. Mitaroff,[11] K. Miyabayashi,[22] H. Miyake,[30] H. Miyata,[28] R. Mizuk,[12] D. Mohapatra,[48] G. R. Moloney,[20] T. Nagamine,[42] Y. Nagasaka,[9] I. Nakamura,[8] E. Nakano,[29] M. Nakao,[8] H. Nakazawa,[8] Z. Natkaniec,[26] S. Nishida,[8] O. Nitoh,[46] S. Ogawa,[40] T. Ohshima,[21] T. Okabe,[21] S. Okuno,[14] S. L. Olsen,[7] W. Ostrowicz,[26] H. Ozaki,[8] H. Palka,[26] C. W. Park,[37] H. Park,[16] N. Parslow,[38] L. S. Peak,[38] R. Pestotnik,[13] L. E. Piilonen,[48] N. Root,[1] M. Rozanska,[26] H. Sagawa,[8] Y. Sakai,[8] N. Sato,[21] T. Schietinger,[17] O. Schneider,[17] P. Schönmeier,[42] J. Schümann,[25] C. Schwanda,[11] A. J. Schwartz,[4] M. E. Sevior,[20] H. Shibuya,[40] B. Shwartz,[1] V. Sidorov,[1] J. B. Singh,[31] A. Somov,[4] N. Soni,[31] R. Stamen,[8] S. Stanič,[47,*] M. Starič,[13] K. Sumisawa,[30] T. Sumiyoshi,[45] S. Suzuki,[34] S. Y. Suzuki,[8] O. Tajima,[8] F. Takasaki,[8] K. Tamai,[8] N. Tamura,[28] M. Tanaka,[8] G. N. Taylor,[20] Y. Teramoto,[29] X. C. Tian,[32] T. Tsukamoto,[8] S. Uehara,[8] T. Uglov,[12] K. Ueno,[25] S. Uno,[8] P. Urquijo,[20] Y. Ushiroda,[8] G. Varner,[7] K. E. Varvell,[38] S. Villa,[17] C. C. Wang,[25] C. H. Wang,[24] M.-Z. Wang,[25] Y. Watanabe,[44] Q. L. Xie,[10] B. D. Yabsley,[48] A. Yamaguchi,[42] Y. Yamashita,[27] M. Yamauchi,[8] Heyoung Yang,[36] J. Ying,[32] L. M. Zhang,[35] Z. P. Zhang,[35] V. Zhilich,[1] D. Žontar,[18,13] and D. Zürcher[17]

(Belle Collaboration)

[1]*Budker Institute of Nuclear Physics, Novosibirsk*
[2]*Chiba University, Chiba*
[3]*Chonnam National University, Kwangju*
[4]*University of Cincinnati, Cincinnati, Ohio 45221, USA*
[5]*University of Frankfurt, Frankfurt*
[6]*Gyeongsang National University, Chinju*
[7]*University of Hawaii, Honolulu, Hawaii 96822, USA*
[8]*High Energy Accelerator Research Organization (KEK), Tsukuba*
[9]*Hiroshima Institute of Technology, Hiroshima*
[10]*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing*
[11]*Institute of High Energy Physics, Vienna*
[12]*Institute for Theoretical and Experimental Physics, Moscow*
[13]*J. Stefan Institute, Ljubljana*
[14]*Kanagawa University, Yokohama*
[15]*Korea University, Seoul*
[16]*Kyungpook National University, Taegu*
[17]*Swiss Federal Institute of Technology of Lausanne, EPFL, Lausanne*
[18]*University of Ljubljana, Ljubljana*
[19]*University of Maribor, Maribor*
[20]*University of Melbourne, Victoria*
[21]*Nagoya University, Nagoya*
[22]*Nara Women's University, Nara*
[23]*National Central University, Chung-li*
[24]*National United University, Miao Li*
[25]*Department of Physics, National Taiwan University, Taipei*
[26]*H. Niewodniczanski Institute of Nuclear Physics, Krakow*
[27]*Nihon Dental College, Niigata*
[28]*Niigata University, Niigata*
[29]*Osaka City University, Osaka*
[30]*Osaka University, Osaka*

[31]*Panjab University, Chandigarh*
[32]*Peking University, Beijing*
[33]*Princeton University, Princeton, New Jersey 08544, USA*
[34]*Saga University, Saga*
[35]*University of Science and Technology of China, Hefei*
[36]*Seoul National University, Seoul*
[37]*Sungkyunkwan University, Suwon*
[38]*University of Sydney, Sydney NSW*
[39]*Tata Institute of Fundamental Research, Bombay*
[40]*Toho University, Funabashi*
[41]*Tohoku Gakuin University, Tagajo*
[42]*Tohoku University, Sendai*
[43]*Department of Physics, University of Tokyo, Tokyo*
[44]*Tokyo Institute of Technology, Tokyo*
[45]*Tokyo Metropolitan University, Tokyo*
[46]*Tokyo University of Agriculture and Technology, Tokyo*
[47]*University of Tsukuba, Tsukuba*
[48]*Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, USA*
[49]*Yonsei University, Seoul*

(Received 24 March 2005; published 30 November 2005)

We present an improved measurement of the branching fraction for the electroweak penguin process $B \to X_s \ell^+ \ell^-$, where ℓ is an electron or a muon and X_s is a hadronic system containing an s-quark. The measurement is based on a sample of 152×10^6 $\Upsilon(4S) \to B\bar{B}$ events collected with the Belle detector at the KEKB energy asymmetric e^+e^- collider. The X_s hadronic system is reconstructed from one $K^\pm$ or K^0_S and up to four pions, where at most one pion can be neutral. Averaging over both lepton flavors, the inclusive branching fraction is measured to be $\mathcal{B}(B \to X_s \ell^+ \ell^-) = (4.11 \pm 0.83(\text{stat})^{+0.85}_{-0.81}(\text{syst})) \times 10^{-6}$ for $M_{\ell^+\ell^-} > 0.2\ \text{GeV}/c^2$.

DOI: 10.1103/PhysRevD.72.092005 PACS numbers: 13.20.He, 12.15.Ji, 14.40.Nd, 14.65.Fy

I. INTRODUCTION

In the standard model (SM), the rare decay $B \to X_s \ell^+ \ell^-$ proceeds through a $b \to s \ell^+ \ell^-$ transition, which is forbidden at tree level. Such a flavor-changing neutral current (FCNC) process can occur at higher order via electroweak penguin and W^+W^- box diagrams. The $b \to s\ell^+\ell^-$ transition therefore allows deeper insight into the effective Hamiltonian that describes FCNC processes and is sensitive to the effects of non-SM physics that may enter these loops; see, for example, Refs. [1,2].

Recent SM calculations of the inclusive $B \to X_s \ell^+ \ell^-$ branching fractions for $M_{l^+l^-} > 0.2\ \text{GeV}/c^2$ predict $\mathcal{B}(B \to X_s e^+ e^-) = (4.2 \pm 0.7) \times 10^{-6}$ [1,3], and $\mathcal{B}(B \to X_s \ell^+ \ell^-) = (4.6 \pm 0.8) \times 10^{-6}$ [4]. Both the Belle and *BABAR* collaborations have observed exclusive $B \to K \ell^+ \ell^-$ and $K^* \ell^+ \ell^-$ decays [5–7] and have measured the rate for inclusive $B \to X_s \ell^+ \ell^-$ decay [8,9].

In this analysis, we study the inclusive $B \to X_s \ell^+ \ell^-$ process by semi-inclusively reconstructing the final state from a pair of electrons or muons and a hadronic system consisting of one $K^\pm$ or K^0_S and up to four pions, where at most one pion can be neutral. This semi-inclusive reconstruction approach [10] allows approximately 53% of the full inclusive rate to be reconstructed. If the contribution of the modes containing a K^0_L is taken to be equal to that containing a K^0_S, the missing states that remain unaccounted for represent ~30% of the total rate. We require the hadronic mass for the selected final states to be less than $2.0\ \text{GeV}/c^2$ to reduce combinatorial background. We correct for the missing modes and the effect of the hadronic mass requirement to extract the inclusive $B \to X_s \ell^+ \ell^-$ decay rate for $M_{\ell^+\ell^-} > 0.2\ \text{GeV}/c^2$. This measurement updates and supersedes our previous result [8], which was based on a sample of 65×10^6 $B\bar{B}$ pairs.

II. THE BELLE DETECTOR AND DATA SAMPLE

We use a data sample collected on the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB energy asymmetric e^+e^- collider (3.5 GeV on 8 GeV) [11]. This sample comprises 152×10^6 B meson pairs, corresponding to an integrated luminosity of 140 fb^{-1}. A detailed description of the Belle detector can be found elsewhere [12]. A three-layer silicon vertex detector and a 50-layer central drift chamber (CDC) are used for tracking and identification of charged particles. An array of aerogel threshold Čerenkov counters (ACC) and time-of-flight (TOF) scintillation counters are also used for charged-particle identification. An electromagnetic calorimeter comprised of Tl-doped CsI crystals (ECL) measures the energy of electromagnetic particles and is also used for electron identification. These detectors are located inside a superconducting solenoid

*On leave from Nova Gorica Polytechnic, Nova Gorica.

coil that provides a 1.5 T magnetic field. An iron flux return located outside of the coil is instrumented with resistive plate counters to identify muons (the K_L and muon detection subsystem, KLM).

Particle identification for $e^{\pm}$, $\mu^{\pm}$, $K^{\pm}$, K^0_S, $\pi^{\pm}$, and π^0 is important for this analysis. Electron identification is based on the ratio of the cluster energy to the track momentum (E/p), the specific energy loss (dE/dx) in the CDC, the position and shower shape of the cluster in the ECL, and the response from the ACC. Muon identification is based on the hit positions and the depth of penetration into the ECL and KLM. Electrons and muons are required to have laboratory-frame momenta greater than 0.4 and 0.8 GeV/c, respectively. To select good muon candidates, we apply a kaon veto. We find an electron (muon) selection efficiency of 94.3% (93.2%) with a 0.1% (0.9%) pion to electron (muon) misidentification probability. Bremsstrahlung photons from electrons are recovered by combining each electron with photons within a small angular region around the electron direction. Charged kaon candidates are selected by using information from the ACC, TOF, and CDC. The kaon selection efficiency is 90% with a pion to kaon misidentification probability of 6%. After selecting the electron, muon, and charged kaon candidate tracks, the remaining tracks are assumed to be charged pions. K^0_S candidates are reconstructed from pairs of oppositely charged tracks with $|M_{\pi^+\pi^-} - M_{K^0_S}| < 15\ \mathrm{MeV}/c^2$ ($\sim 6\sigma_{M_{K^0_S}}$). We impose additional K^0_S selection criteria based on the distance and the direction of the K^0_S vertex, and on the impact parameters of the daughter tracks. We require all charged tracks, except those used in the K^0_S reconstruction, to have impact parameters with respect to the nominal interaction point of less than 1.0 cm in the radial direction and 5.0 cm along the beam direction. Neutral pions are required to have a laboratory-frame energy greater than 400 MeV, photon daughter energies greater than 50 MeV, and a $\gamma\gamma$ invariant mass satisfying $|M_{\gamma\gamma} - M_{\pi^0}| < 10\ \mathrm{MeV}/c^2$ ($\sim 3\sigma_{M_{\pi^0}}$).

III. ANALYSIS OVERVIEW

We reconstruct inclusive $B \to X_s \ell^+ \ell^-$ decays with a semi-inclusive reconstruction technique from a pair of electrons or muons and one of 18 different hadronic states. Here the hadronic system consists of one $K^{\pm}$ or K^0_S and up to four pions (at most one pion can be neutral). Compared to a fully inclusive approach, this method has the advantage of having strong kinematical discrimination against background by using the beam-energy constrained mass $M_{\mathrm{bc}} = \sqrt{E^2_{\mathrm{beam}} - p^2_B}$ and the energy difference $\Delta E = E_B - E_{\mathrm{beam}}$, where E_{beam} is the beam energy and E_B (p_B) is the reconstructed B meson energy (3-momentum). All quantities are evaluated in the e^+e^- center-of-mass system (CM).

Further background suppression to reduce the large combinatorial backgrounds is necessary. The main contribution to the combinatorial background comes from pairs of semileptonic $b \to c$ decays in $B\bar{B}$ events. In these events, $B \to X_s \ell^+ \ell^-$ candidates are reconstructed with the decay products from both $B\bar{B}$ mesons. This background has a significant amount of missing energy due to the neutrinos from the semileptonic decays. Another contribution to the combinatorial background comes from continuum events, which are efficiently suppressed with event-shape variables.

There are two background sources that can peak in M_{bc} and ΔE. The first comes from the charmonium decays $B \to J/\psi X$ and $B \to \psi(2S)X$ with $J/\psi(\psi(2S)) \to \ell^+\ell^-$. This charmonium background is efficiently removed with requirements on the dilepton mass $M_{\ell^+\ell^-}$. The removed events provide a large control sample of decays with a signature identical to that of the signal. The second comes from $B \to K^{\pm}(K^0_S) n\pi$ ($n > 1$) decays with misidentification of two charged pions as leptons. We estimate these peaking background contaminations, then subtract them from the signal yield.

For the $B \to X_s \ell^+ \ell^-$ event simulation, we use EVTGEN [13] as the event generator, JETSET [14] to hadronize the system consisting of a strange quark and a spectator quark, and GEANT [15] for the detector simulation. In the event generation, $B \to X_s \ell^+ \ell^-$ events are produced with a combination of exclusive and inclusive models. In the hadronic mass region $M_{X_s} < 1.1\ \mathrm{GeV}/c^2$, exclusive $B \to K^{(*)}\ell^+\ell^-$ decays are generated according to Refs. [1,16], where the relevant form factors are computed using light-cone QCD sum rules. In the region $M_{X_s} > 1.1\ \mathrm{GeV}/c^2$, event generation is based on a nonresonant model following Refs. [1,17] and the Fermi motion model of Ref. [18]. We use the measured branching fraction $(4.8^{+1.0}_{-0.9} \pm 0.3 \pm 0,1) \times 10^{-7}$ for $B \to K\ell^+\ell^-$ [6], and theoretical calculations $(1.19 \pm 0.39) \times 10^{-6}$ for $B \to K^*\ell^+\ell^-$ ($M_{l^+l^-} > 0.2\ \mathrm{GeV}/c^2$) [1] and $(4.2 \pm 0.7) \times 10^{-6}$ for $B \to X_s\ell^+\ell^-$ ($M_{l^+l^-} > 0.2\ \mathrm{GeV}/c^2$ and entire M_{X_s} region before the M_{X_s} requirement) [1] for the event generation.

IV. EVENT SELECTION

Events are required to have a well determined primary vertex, be tagged as multihadron events, and contain two electrons (or muons) having laboratory-frame momenta greater than 0.4 GeV/c (or 0.8 GeV/c for muons). Dilepton candidates consist of e^+e^- or $\mu^+\mu^-$ pairs. The two leptons are required to originate from a common vertex and satisfy the requirement $|\Delta z| < 0.015$ cm. Here, Δz is the distance between the two leptons along the beam direction; the z-coordinate of each lepton is determined at the point of closest approach to the beam axis.

Charmonium backgrounds are reduced by removing B candidates with a dilepton mass in the ranges $-0.40\ \mathrm{GeV}/c^2 < M_{ee(\gamma)} - M_{J/\psi} < 0.15\ \mathrm{GeV}/c^2$,

$-0.25\ \mathrm{GeV}/c^2 < M_{\mu\mu} - M_{J/\psi} < 0.10\ \mathrm{GeV}/c^2$, $-0.25\ \mathrm{GeV}/c^2 < M_{ee(\gamma)} - M_{\psi(2S)} < 0.10\ \mathrm{GeV}/c^2$, and $-0.15\ \mathrm{GeV}/c^2 < M_{\mu\mu} - M_{\psi(2S)} < 0.10\ \mathrm{GeV}/c^2$. If one of the electrons from a J/ψ or $\psi(2S)$ decay erroneously picks up a random photon in the bremsstrahlung-recovery process, the dilepton mass can increase sufficiently to evade the above requirements. Therefore the charmonium veto is applied to the dilepton mass before and after bremsstrahlung recovery. Using the simulation, we estimate the remaining peaking charmonium background to be 1.20 ± 0.28 events and 1.33 ± 0.21 events for the e^+e^- and $\mu^+\mu^-$ channels, respectively.

The requirement $M_{l^+l^-} > 0.2\ \mathrm{GeV}/c^2$ is applied. For the e^+e^- channel, this removes potential peaking background from the conversion of photons from radiative $B \to X_s\gamma$ decays and from π^0 Dalitz decays.

Starting from a $\ell^+\ell^-$ pair, $B \to X_s\ell^+\ell^-$ candidates are formed by adding either a $K^\pm$ or a K^0_S and up to four pions, but no more than one π^0. In this manner, eighteen different hadronic topologies are reconstructed: $K^\pm$, $K^\pm\pi^0$, $K^\pm\pi^\mp$, $K^\pm\pi^\mp\pi^0$, $K^\pm\pi^\mp\pi^\pm$, $K^\pm\pi^\mp\pi^\pm\pi^0$, $K^\pm\pi^\mp\pi^\pm\pi^\mp$, $K^\pm\pi^\mp\pi^\pm\pi^\mp\pi^0$, $K^\pm\pi^\mp\pi^\pm\pi^\mp\pi^\pm$, K^0_S, $K^0_S\pi^0$, $K^0_S\pi^\pm$, $K^0_S\pi^\pm\pi^0$, $K^0_S\pi^\pm\pi^\mp$, $K^0_S\pi^\pm\pi^\mp\pi^0$, $K^0_S\pi^\pm\pi^\mp\pi^\mp$, $K^0_S\pi^\pm\pi^\mp\pi^\mp\pi^0$, and $K^0_S\pi^\pm\pi^\mp\pi^\mp\pi^\pm$.

After forming the $B \to X_s\ell^+\ell^-$ candidates, we apply additional requirements to suppress background. The largest background sources are random combinations from continuum $q\bar{q}$ ($q = u, d, s, c$) production or from semileptonic B decays. We reject the $q\bar{q}$ background by a requirement on a Fisher discriminant [19] ($\mathcal{F}_{\mathrm{FW}}$) based on a modified set of Fox-Wolfram moments [20] that characterize the event topology. In the semileptonic B decay background, both B mesons decay into leptons or two leptons are produced from the $b \to c \to s, d$ decay chain. This B decay background is rejected by the requirement on another Fisher discriminant variable ($\mathcal{F}_{\mathrm{miss}}$) constructed from the total visible energy $E_{\mathrm{vis}} = \sum_i E_i$ and the missing mass $M_{\mathrm{miss}} = \sqrt{(2E_{\mathrm{beam}} - \sum E_i)^2 - |\sum \vec{p}_i|^2}$, where $(\vec{p}_i, E_i)$ are the reconstructed four momenta, in the CM frame, of all tracks (assumed to be pions) and all photons in the event.

We further reduce the background by requiring $-0.10\ \mathrm{GeV} < \Delta E < 0.05\ \mathrm{GeV}$ for the dielectron channel ($-0.05\ \mathrm{GeV} < \Delta E < 0.05\ \mathrm{GeV}$ for the dimuon channel), and $\chi^2_{\mathrm{vtx}}/NDF < 10$. Here, χ^2_{vtx} is the χ^2 of the B vertex constructed from the charged daughter particles, excluding the K^0_S daughters. We also reject candidates with an X_s invariant mass greater than $2\ \mathrm{GeV}/c^2$. This condition removes a large fraction of the combinatorial background while retaining 99% of the signal.

At this stage, there is an average of 1.6 B candidates per event in the signal simulation. In order to select the most signal-like B candidate, we construct a likelihood ratio based on the following six discriminant variables: ΔE, $\Delta E^{\mathrm{ROE}} = E_{\mathrm{ROE}} - E_{\mathrm{beam}}$, χ^2_{vtx}, $\cos\theta_B$, $\mathcal{F}_{\mathrm{FW}}$, and $\mathcal{F}_{\mathrm{miss}}$. The energy of the rest of the event, E_{ROE}, is calculated by summing the energies of all charged tracks and neutral calorimeter clusters not included in the B candidate, and $\cos\theta_B$ is the cosine of the B flight direction with respect to the e^- beam direction in the CM frame. The variables ΔE, ΔE^{ROE}, and $\mathcal{F}_{\mathrm{miss}}$ are effective at rejecting $B\bar{B}$ background, especially for events with two semileptonic decays, which have large missing energy. For continuum suppression, the event-shape variables $\mathcal{F}_{\mathrm{FW}}$ and $\cos\theta_B$ are useful. The variable χ^2_{vtx} is effective for rejecting the random combinatorial background in the high multiplicity modes.

The probability density functions (PDFs) of these six variables are determined from Monte Carlo (MC) events, separately for signal and background, with the exception of signal PDFs for ΔE and χ^2_{vtx}. We determine the latter two from the charmonium data sample, a large control sample which is topologically identical to the signal, satisfying all selection criteria except for the charmonium veto and falling in the charmonium window (charmonium-veto sample), because we observe some discrepancy between MC and data in the charmonium-veto samples. To determine the PDF from data, we subtract the background using the shapes obtained from the M_{bc} sidebands. The number of background events in the M_{bc} signal region is normalized by fitting a Gaussian (signal) and an ARGUS function [21] (background) to the M_{bc} distribution. For ΔE, we use separate PDFs for the dielectron and dimuon channels.

We then calculate the likelihoods $\mathcal{L}_{\mathrm{sig}}$ and $\mathcal{L}_{\mathrm{BG}}$, which are products of the PDFs of the above six discriminants for the signal and the background, respectively. In each event, only the B candidate with the largest value of the likelihood ratio $\mathcal{R} = \mathcal{L}_{\mathrm{sig}}/(\mathcal{L}_{\mathrm{sig}} + \mathcal{L}_{\mathrm{BG}})$ is retained. According to the MC simulation, this choice implies that 84% of the selected $B \to X_s\ell^+\ell^-$ candidates in $B \to X_s\ell^+\ell^-$ events have all daughter particles correctly assigned.

To check the distributions of the likelihood ratio $\mathcal{R}$, we compare data and MC, separately for signal [Fig. 1(a)] and background [Fig. 1(b)] events. For the signal, we use charmonium-veto candidates from data and $B \to X_s\ell^+\ell^-$ signal MC simulation. In the $\mathcal{R}$ distribution for charmonium-veto events, we subtract the background using shapes obtained from the M_{bc} sidebands and normalization from the number of background events in the M_{bc} signal region, in the same way as for the signal PDF determination. For the background we use $B \to X_s e^\pm\mu^\mp$ candidates that are reconstructed using the nominal selection criteria but requiring that the two leptons have different flavor, in data and in a background MC sample that contains $b \to c$ decays and $q\bar{q}$ events. We observe good agreement in both cases.

To suppress the combinatorial background effectively, we also use the likelihood ratio $\mathcal{R}$. We apply a requirement on $\mathcal{R}$, which is optimized to maximize the statistical

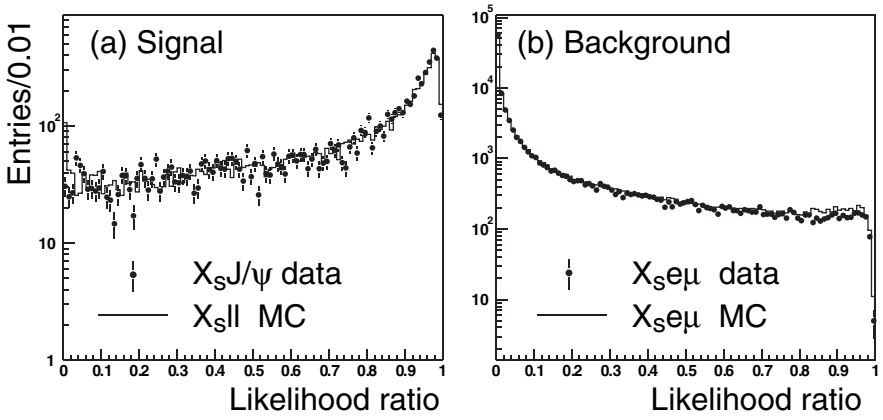

FIG. 1. Likelihood ratio for (a) signal and (b) background events. Points represent data and solid histograms represent MC. Samples used are described in the text.

significance of the signal $S/\sqrt{S+B}$. Here S and B denote the numbers of the signal and background events, respectively. This MC optimization is performed separately in the regions $M_{X_s} < 1.1\ \mathrm{GeV}/c^2$ and $1.1\ \mathrm{GeV}/c^2 < M_{X_s} < 2.0\ \mathrm{GeV}/c^2$, resulting in the requirements $\mathcal{R} > 0.3$ and $\mathcal{R} > 0.9$, respectively.

After applying all selection criteria, 155 $B \to X_s e^+ e^-$ and 112 $B \to X_s \mu^+ \mu^-$ candidates remain in the M_{bc} signal region, defined as $5.27\ \mathrm{GeV}/c^2 < M_{\mathrm{bc}} < 5.29\ \mathrm{GeV}/c^2$. According to the simulation, the background remaining at this stage of the analysis consists mostly of $B\bar{B}$ events (80% and 72% of the total background in the electron and muon channels, respectively). In MC signal events, the probability that all daughter particles from a selected $B \to X_s \ell^+ \ell^-$ candidate are correctly assigned is 91%.

V. MAXIMUM LIKELIHOOD FIT

We perform an extended unbinned maximum likelihood fit to the M_{bc} distribution in the region $M_{\mathrm{bc}} > 5.2\ \mathrm{GeV}/c^2$ to extract the signal yield as well as the shape and yield of the combinatorial background. The fit is performed separately for the dielectron and dimuon channels. The likelihood function $\mathcal{L}$ is expressed as

$$\mathcal{L} = \frac{e^{-(N_{\mathrm{sig}}+N_{\mathrm{peak}}+N_{\mathrm{pc}}+N_{\mathrm{cf}}+N_{\mathrm{comb}})}}{N!} \prod_{i=1}^{N} [(N_{\mathrm{sig}} + N_{\mathrm{peak}})\mathcal{P}_i^{\mathrm{sig}} + N_{\mathrm{pc}}\mathcal{P}_i^{\mathrm{pc}} + N_{\mathrm{cf}}\mathcal{P}_i^{\mathrm{cf}} + N_{\mathrm{comb}}\mathcal{P}_i^{\mathrm{comb}}],$$

where N and i denote the total number and index of candidate events, respectively. N_{sig}, N_{peak}, N_{pc}, N_{cf}, and N_{comb} represent the yields of the signal, peaking background, combinatorial background from peaking background, cross-feed background (i.e. events containing a true $B \to X_s \ell^+ \ell^-$ decay, but a selected B candidate with incomplete or wrong decay products), and combinatorial background, respectively. As explained below, we use the same PDF, $\mathcal{P}_i^{\mathrm{sig}}$, for the signal and peaking background components.

The signal PDF $\mathcal{P}_i^{\mathrm{sig}}$ is described by a Gaussian for the $\mu^+\mu^-$ channel as well as for the e^+e^- channel, since the bremsstrahlung recovery and selection procedure for the e^+e^- channel lead to a negligible radiative tail in the M_{bc} distribution. The Gaussian shape parameters are determined from fits of the sum of a Gaussian and an ARGUS function to the charmonium-veto data sample. The fits result in a mean M_{bc} of $m_{\mathrm{sig}} = 5279.31 \pm 0.05\ \mathrm{MeV}/c^2$ and a M_{bc} resolution of $\sigma_{\mathrm{sig}} = 2.62 \pm 0.04\ \mathrm{MeV}/c^2$ for the e^+e^- channel, and $m_{\mathrm{sig}} = 5279.03 \pm 0.04\ \mathrm{MeV}/c^2$ and $\sigma_{\mathrm{sig}} = 2.53 \pm 0.04\ \mathrm{MeV}/c^2$ for the $\mu^+\mu^-$ channel. In the simulation, the Gaussian fit results for the M_{bc} distributions of correctly reconstructed signal events are in agreement with the shape parameters extracted from the fits to the charmonium-veto sample. The signal yield N_{sig} is a free parameter in the likelihood fit.

We consider two separate contributions to the peaking background: the charmonium peaking background, already described, and the hadronic peaking background. The latter arises from the $B \to D^{(*)} n\pi\ (n > 0) \to X_s \pi^+ \pi^-$ decays where two hadrons are misidentified as leptons. The normalization and shape of this background are determined directly from the data by repeating the selection without the lepton identification requirements, and fitting the M_{bc} distribution with the sum of a Gaussian and an ARGUS function. We obtain $m_{\mathrm{peak}} = 5279.16 \pm 0.04\ \mathrm{MeV}/c^2$ and $\sigma_{\mathrm{peak}} = 2.60 \pm 0.02\ \mathrm{MeV}/c^2$, consistent with the signal shape of the charmonium peaking background. Hence we use a common PDF for these two peaking backgrounds, which is the same as the one determined for the signal. Taking into account the momentum and angular dependence of the $\pi \to \ell$ misidentification probabilities (on average 0.08% for electrons and 0.92% for muons), the hadronic peaking background is estimated to be $N_{\text{h-peak}} = 0.030 \pm 0.001$ events for the dielectron sample and

$N_{\text{h-peak}} = 1.78 \pm 0.05$ events for the dimuon sample. The charmonium peaking background is estimated from the simulation to be $N_{\text{c-peak}} = 1.20 \pm 0.28$ and $N_{\text{c-peak}} = 1.33 \pm 0.21$, respectively. In the likelihood fits, the peaking background normalization is fixed to $N_{\text{peak}} = N_{\text{c-peak}} + N_{\text{h-peak}}$. Figures 2(a) and 2(b) show the M_{bc} distributions for the MC charmonium events for the dielectron and dimuon channels, and 2(c) and 2(d) show the M_{bc} distributions for $B \to X_s h^+ h^-$ candidates in data for the dielectron and dimuon channels, respectively.

The other background PDFs, $\mathcal{P}_i^{\text{pc}}$, $\mathcal{P}_i^{\text{cf}}$, and $\mathcal{P}_i^{\text{comb}}$ are given by an ARGUS shape. They describe the combinatorial contribution from peaking background events, from cross-feed events, and from continuum and $B\bar{B}$ events, respectively. The ARGUS cutoff is determined by the beam energy in the $\Upsilon(4S)$ rest frame, $E_{\text{beam}} = 5.290$ GeV. The values of the ARGUS shape parameter for each background component are determined from the peaking background M_{bc} distribution shown in Fig. 2 ($\mathcal{P}_i^{\text{pc}}$), from incorrectly reconstructed signal MC events ($\mathcal{P}_i^{\text{cf}}$), and from the fit to the $B \to X_s e^{\pm} \mu^{\mp}$ data selected using the nominal selection criteria but requiring that the two leptons have different flavor ($\mathcal{P}_i^{\text{comb}}$). We fix these three ARGUS shape parameters. We also fix the number of peaking background events N_{pc} and the number of cross-feed events N_{cf}, from the peaking background M_{bc} distribution shown in Fig. 2 and from incorrectly reconstructed signal MC events, respectively. Here the signal MC generations are based on the branching fractions in the Ref. [1]. The yield N_{comb} is taken as a free parameter in the likelihood fit.

VI. RESULTS

Using the fit parametrization described above, we fit the M_{bc} distributions in data separately for the $B \to X_s e^+ e^-$ and $B \to X_s \mu^+ \mu^-$. The fit results are displayed in Fig. 3 and summarized in Table I. The significance is $S = \sqrt{-2\ln(\mathcal{L}_{\text{max}}^0/\mathcal{L}_{\text{max}})}$, where $\mathcal{L}_{\text{max}}$ represents the maximum likelihood for the fit and $\mathcal{L}_{\text{max}}^0$ denotes the maximum likelihood when the signal yield is constrained to be 0. To include the systematic uncertainty of the signal yield estimation into the significance, the fixed parameters in the fit are varied individually by $\pm 1\sigma$. The minimum significance is quoted. The $B \to X_s \ell^+ \ell^-$ signal yield presented in Table I is obtained by a fit to the combined electron and muon samples. Figure 3(d) shows the M_{bc} distribution for $B \to X_s e^{\pm} \mu^{\mp}$ candidates. Applying the ARGUS fit to the M_{bc} distribution, there is no evidence for a peaking background as expected.

Figures 4(a) and 4(b) show the distributions of the hadronic mass M_{X_s} and $q^2 \equiv M^2_{\ell^+\ell^-}$ for the electron and

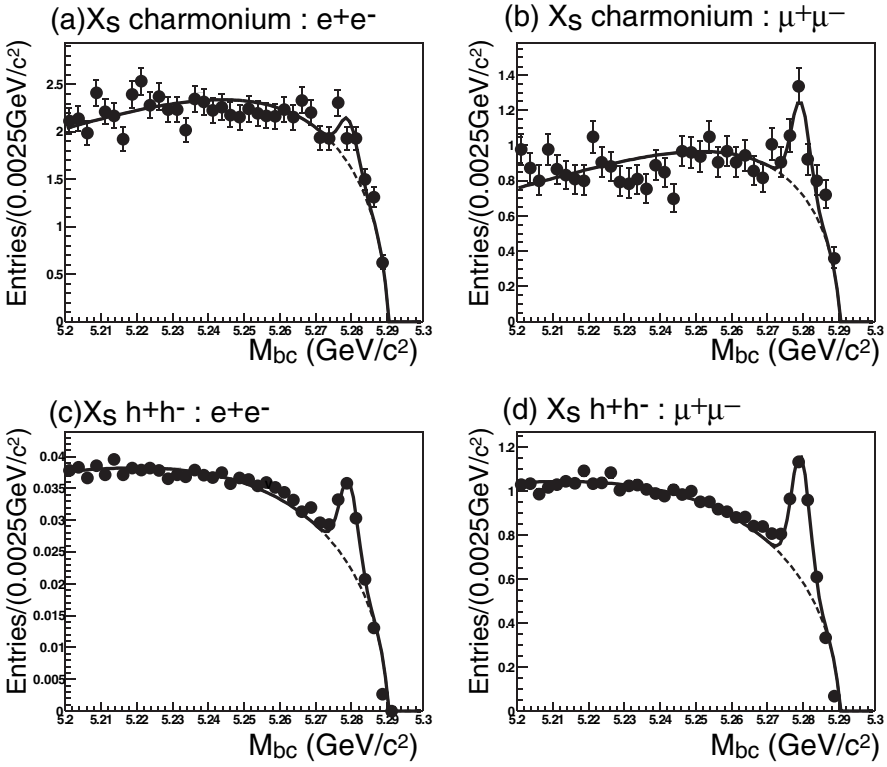

FIG. 2. M_{bc} distributions for MC charmonium events for the (a) dielectron and (b) dimuon channels, and for $B \to X_s h^+ h^-$ candidates in data for the (c) dielectron and (d) dimuon channels. The normalization corresponds to the expected charmonium and hadronic peaking background in 140 fb^{-1} of data.

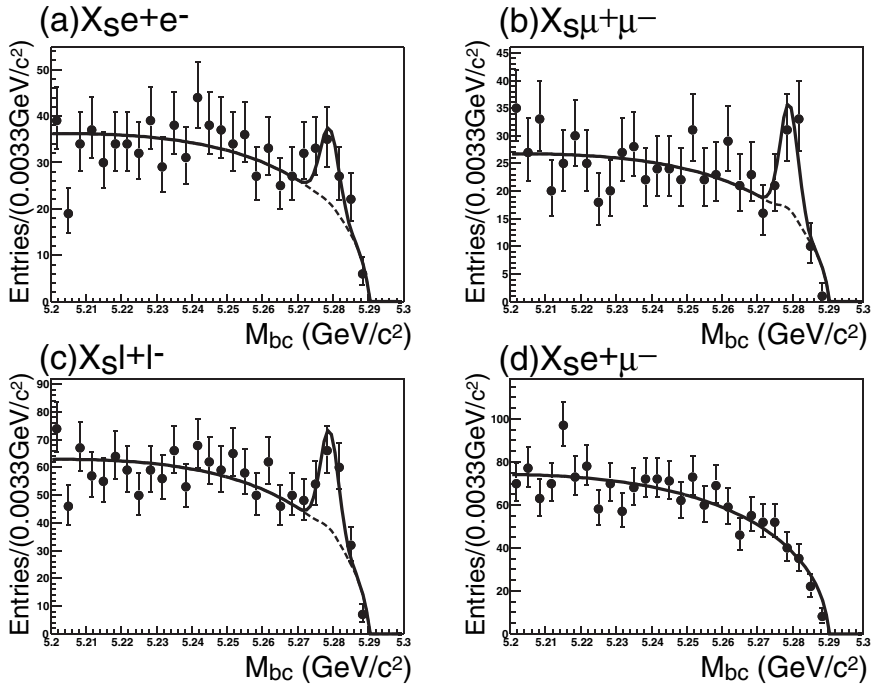

FIG. 3. $M_{\rm bc}$ distributions of (a) $B \to X_s e^+ e^-$, (b) $B \to X_s \mu^+ \mu^-$, (c) $B \to X_s \ell^+ \ell^-$ ($\ell = e, \mu$), and (d) $B \to X_s e^\pm \mu^\mp$ candidates in data. The solid lines represent the result of the fits, and dashed lines represent the sum of all background components under the signal peaks, respectively.

TABLE I. Results of the fit to the data: number of signal candidates in the $M_{\rm bc}$ signal box, fitted signal yield $N_{\rm sig}$, estimated peaking backgrounds $N_{\rm peak}$ (fixed in the fit), and significance including systematics.

Mode	Candidates	$N_{\rm sig}$	$N_{\rm peak}$	Significance
$X_s e^+ e^-$	155	31.8 ± 10.2	1.24 ± 0.28	3.2
$X_s \mu^+ \mu^-$	112	36.3 ± 9.3	3.11 ± 0.22	4.4
Combined	267	68.4 ± 13.8	4.35 ± 0.36	5.4

muon channels combined, obtained by repeating the likelihood fit in bins of M_{X_s} and q^2. Figure 4(a) indicates that the observed signal includes contributions from final states across a range of hadronic masses, including hadronic systems with a mass above that of the $K^*(892)$.

The branching fraction $\mathcal{B}$ for the signal is calculated as

$$\mathcal{B} = \frac{N_{\rm sig}}{2 N_{B\bar{B}} \epsilon}, \tag{1}$$

where $N_{B\bar{B}} = (152.0 \pm 0.7) \times 10^6$ is the number of $B\bar{B}$ pairs produced in 140 fb^{-1} and ϵ is the signal efficiency.

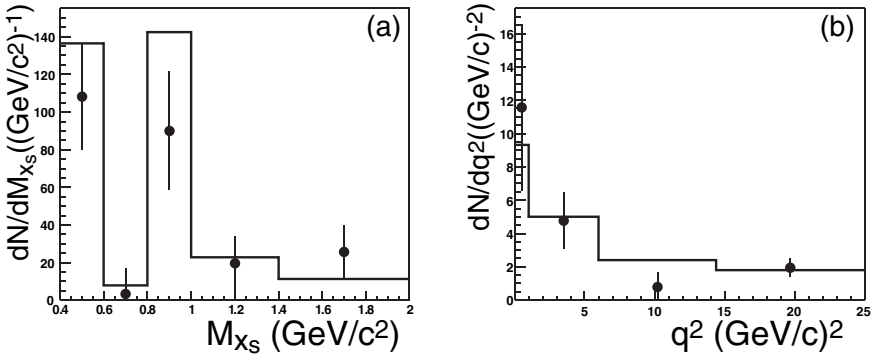

FIG. 4. Distributions for the signal yield as a function of (a) hadronic mass M_{X_s} and (b) $q^2 \equiv M^2_{\ell^+\ell^-}$ for the $B \to X_s \ell^+ \ell^-$ signal. The points with error bars are the data (electron and muon channels combined) while the histogram is the MC signal normalized to the integrated luminosity of 140 fb^{-1}.

VII. SYSTEMATIC UNCERTAINTIES

Systematic uncertainties are of two different types: those that affect the extraction of the number of signal events and those that affect the calculation of the branching fraction. The systematic uncertainties are summarized in Table II.

Uncertainties affecting the extraction of the signal yield are evaluated by varying the signal Gaussian parameters (mean and width) and the background shape parameter within $\pm 1\sigma$ of the measured values from the charmonium-veto data sample (signal) and the $B \to X_s e^{\pm} \mu^{\mp}$ data sample (background). We estimate the uncertainties in the shape of the combinatorial contribution in the peaking background events and in the cross-feed events by comparing the signal yields obtained with and without the corresponding PDFs in the unbinned maximum likelihood fit to $M_{\rm bc}$.

Uncertainties affecting the signal efficiency originate from the detector modeling, from the simulation of signal decays, and from the estimate of the number of B mesons in the sample. By far the largest component is that due to the simulation of signal decays, discussed in detail below.

TABLE II. Relative systematic uncertainties (in %) on the $B \to X_s e^+ e^-$ and $B \to X_s \mu^+ \mu^-$ branching fraction measurements. The different contributions are summed in quadrature. PID in the table denotes particle identification.

Source	$X_s e^+ e^-$	$X_s \mu^+ \mu^-$
Signal shape	± 1.4	± 0.5
Background shape	± 7.8	± 4.7
Peaking background statistics	± 0.9	± 0.6
Peaking background PID error	< 0.1	± 0.5
Peaking background shape	± 4.3	± 2.1
Cross-feed events	± 4.1	± 2.2
Signal yield total	± 9.9	± 5.7
Tracking efficiency	± 3.5	± 3.5
Lepton identification efficiency	± 4.1	± 5.9
Kaon identification efficiency	± 0.8	± 0.8
$\pi^{\pm}$ identification efficiency	± 0.6	± 0.5
K^0_S efficiency	± 0.7	± 0.8
π^0 efficiency	± 0.3	± 0.3
$\mathcal{R}$ requirement efficiency	± 5.4	± 4.5
Fermi motion model	$^{+6.5}_{-2.4}$	$^{+6.1}_{-2.3}$
$\mathcal{B}(B \to K \ell^+ \ell^-)$	± 9.9	± 10.5
$\mathcal{B}(B \to K^* \ell^+ \ell^-)$	± 7.0	± 7.8
K^*–X_s transition	± 4.5	± 4.7
Hadronization	± 8.5	± 8.2
Missing modes	± 4.5	± 4.4
Monte Carlo statistics	± 1.6	± 1.5
Efficiency total	$^{+19.0}_{-18.1}$	$^{+19.7}_{-18.9}$
$B\bar{B}$ counting	± 0.5	± 0.5
Total	$^{+21.5}_{-20.6}$	$^{+20.5}_{-19.7}$

The detector modeling uncertainty is composed of the following uncertainties determined from the data: the uncertainty in the tracking efficiency of 1.0% per track; the uncertainty in the charged-particle identification efficiency of 2.05% per electron, 2.95% per muon, 1.0% per kaon and 0.8% per pion; and the uncertainty in the reconstruction efficiency of 4.5% per K^0_S and 3.3% per π^0. The efficiency of the likelihood ratio requirement, which suppresses combinatorial background, is checked with the charmonium-veto sample and the level of discrepancy with the simulation is taken as the corresponding uncertainty.

The dominant source of uncertainty arises from modeling the signal decays. Parameters of the Fermi motion model are varied in accordance with measurements of hadronic moments in semileptonic B decays [22] and the photon spectrum in inclusive $B \to X_s \gamma$ decays [23]. The fractions of exclusive $B \to K \ell^+ \ell^-$ and $B \to K^* \ell^+ \ell^-$ decays are varied according to theoretical uncertainties [1]. The transition point in M_{X_s} between pure $K^* \ell^+ \ell^-$ and nonresonant $X_s \ell^+ \ell^-$ final states is varied by ± 0.1 GeV$/c^2$.

The nonresonant Monte Carlo event generator relies on JETSET to fragment and hadronize the system consisting of a final state s quark and a spectator quark from the B meson. Since the signal efficiencies depend strongly on the particle content of the final state, uncertainties in the number of charged and neutral pions and in the number of charged and neutral kaons translate into a significant uncertainty in the signal efficiency (for $M_{X_s} > 1.1$ GeV$/c^2$).

The ratio of the generator yield for decay modes containing a K^0_S to that for modes containing a charged kaon is varied according to 0.50 ± 0.11, to allow for isospin violation in the decay chain. The ratio of the generator yield for decay modes containing one π^0 meson to that for modes containing none is varied according to 1.00 ± 0.19. Uncertainties in these two ratios are set by the level of discrepancy between $B \to J/\psi X$ data and MC $B \to X_s \ell^+ \ell^-$ samples, by checking the level of agreement of the multiplicities of charged and neutral pions, and K^0_S and $K^{\pm}$ candidates.

The 18 modes selected in this analysis only capture about 53% of the full inclusive rate. If the contribution of the modes containing a K^0_L is taken to be equal to that containing a K^0_S, the missing states that remain unaccounted for represent $\sim$30% of the total rate. To account for the uncertainty in such remaining missing states, we estimate the uncertainty in the fraction of modes with too many pions or kaons (two extra kaons may be produced via $s\bar{s}$ popping), as well as the contributions of modes with photons that do not originate from π^0 decays but rather from η, η', etc. For final states with $M_{X_s} > 1.1$ GeV$/c^2$, we vary these fractions by $\pm 5\%$ per π^0, $\pm 20\%$ for η, $\pm 30\%$ for $N_\pi > 5$, and $\pm 50\%$ for η' and others.

Including systematic uncertainties, the measured branching fractions for $M_{\ell^+\ell^-} > 0.2$ GeV$/c^2$ are

TABLE III. Summary of results: signal yield ($N_{\rm sig}$), statistical significance, total signal efficiency ϵ (including the fraction of X_s states considered in this analysis), and branching fraction ($\mathcal{B}$). The first and second errors quoted on $N_{\rm sig}$ and $\mathcal{B}$ are statistical and systematic, respectively. The first error on ϵ corresponds to uncertainties in detector modeling, $B\bar{B}$ counting, and Monte Carlo statistics, and the second error on ϵ to the uncertainties in the signal model.

Mode	$N_{\rm sig}$	Significance	ϵ (%)	$\mathcal{B}$ ($\times 10^{-6}$)
$X_s e^+ e^-$	$31.8 \pm 10.2 \pm 3.1$	3.2	$2.59 \pm 0.20^{+0.45}_{-0.42}$	$4.04 \pm 1.30^{+0.87}_{-0.83}$
$X_s \mu^+ \mu^-$	$36.3 \pm 9.3 \pm 2.1$	4.4	$2.89 \pm 0.24^{+0.52}_{-0.49}$	$4.13 \pm 1.05^{+0.85}_{-0.81}$
$X_s \ell^+ \ell^-$	$68.4 \pm 13.8 \pm 5.0$	5.4	$2.74 \pm 0.22^{+0.48}_{-0.45}$	$4.11 \pm 0.83^{+0.85}_{-0.81}$

TABLE IV. $B \to X_s \ell^+ \ell^-$ branching fractions ($\mathcal{B}$) for different bins in M_{X_s} and $q^2 = M^2_{\ell^+\ell^-}$. The first and second errors are statistical and systematic, respectively.

M_{X_s} in GeV/c^2	$\mathcal{B}$ ($\times 10^{-7}$)	q^2 in (GeV/c)^2	$\mathcal{B}$ ($\times 10^{-7}$)
[0.4, 0.6]	$3.8 \pm 1.0^{+0.3}_{-0.3}$	[0.04, 1.0]	$11.3 \pm 4.8^{+4.6}_{-2.7}$
[0.6, 0.8]	$0.4 \pm 0.9^{+0.1}_{-0.1}$	[1.0, 6.0]	$14.9 \pm 5.0^{+4.1}_{-3.2}$
[0.8, 1.0]	$6.6 \pm 2.3^{+0.7}_{-0.7}$	[6.0, 14.4]	$7.3 \pm 6.1^{+1.8}_{-1.9}$
[1.0, 1.4]	$10.5 \pm 6.9^{+2.0}_{-2.1}$	[14.4, 25.0]	$4.2 \pm 1.2^{+0.6}_{-0.7}$
[1.4, 2.0]	$46.6 \pm 23.4^{+22.5}_{-10.9}$		

$$\mathcal{B}(B \to X_s e^+ e^-) = (4.04 \pm 1.30^{+0.87}_{-0.83}) \times 10^{-6}, \quad (2)$$

$$\mathcal{B}(B \to X_s \mu^+ \mu^-) = (4.13 \pm 1.05^{+0.85}_{-0.81}) \times 10^{-6}, \quad (3)$$

$$\mathcal{B}(B \to X_s \ell^+ \ell^-) = (4.11 \pm 0.83^{+0.85}_{-0.81}) \times 10^{-6}, \quad (4)$$

where the first error is statistical and the second error is systematic. The combined $B \to X_s \ell^+ \ell^-$ branching fraction is the weighted average of the branching fractions for the electron and muon channels, assuming the individual branching fractions to be equal for $M_{\ell^+\ell^-} > 0.2$ GeV/c^2. Table III summarizes the results of the analysis and lists both the statistical and systematic errors on the signal yields, the signal efficiencies and the branching fractions.

The branching fractions for each M_{X_s} and q^2 bin are also measured, and summarized in Table IV. Figures 5(a) and 5(b) show the distributions of the differential branching fractions as a function of hadronic mass M_{X_s} and $q^2 \equiv M^2_{\ell^+\ell^-}$, respectively, for electron and muon channels combined. The two peaks in the first and third M_{X_s} bins correspond to the $K\ell^+\ell^-$ and $K^*\ell^+\ell^-$ contributions, respectively.

VIII. SUMMARY

Using a sample of 152×10^6 $\Upsilon(4S) \to B\overline{B}$ events, we measure the branching fraction for the rare decay $B \to X_s \ell^+ \ell^-$, where $\ell = e$ or μ and X_s is a hadronic system reconstructed semi-inclusively in 18 different hadronic states (with up to four pions). For $M_{\ell^+\ell^-} > 0.2$ GeV/c^2, we observe a signal of 68.4 ± 13.8(stat) $\pm$ 5.0(syst) events and obtain a branching fraction (averaged over lepton species) of

$$\mathcal{B}(B \to X_s \ell^+ \ell^-) = (4.11 \pm 0.83(\text{stat})^{+0.85}_{-0.81}(\text{syst})) \times 10^{-6},$$

with a statistical significance of 5.4σ.

This result is consistent with the recent prediction by Ali *et al.* [1,3,4], our previous inclusive $B \to X_s \ell^+ \ell^-$ measurement [8], and that of the *BABAR* collaboration [9].

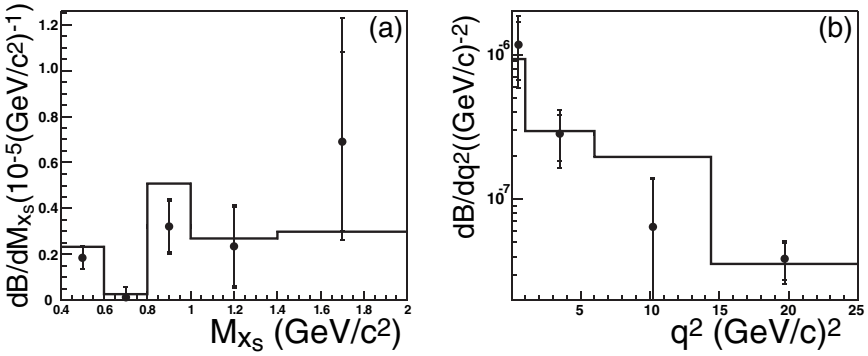

FIG. 5. Differential branching fraction as a function of (a) hadronic mass M_{X_s} and (b) $q^2 \equiv M^2_{\ell^+\ell^-}$ for the $B \to X_s \ell^+ \ell^-$ signal. The points and error bars represent the data (electron and muon channels combined) while the histograms represent the MC signal. The outer (inner) error bars represent the total (statistical) errors.

ACKNOWLEDGMENTS

The authors wish to thank Gudrun Hiller, Tobias Hurth, and Gino Isidori for their helpful suggestions. We have also benefited from discussions with Stephane Willocq regarding Monte Carlo generation with EVTGEN. We thank the KEKB group for the excellent operation of the accelerator, the KEK cryogenics group for the efficient operation of the solenoid, and the KEK computer group and the National Institute of Informatics for valuable computing and SuperSINET network support. We acknowledge support from the Ministry of Education, Culture, Sports, Science, and Technology of Japan and the Japan Society for the Promotion of Science; the Australian Research Council and the Australian Department of Education, Science and Training; the National Science Foundation of China under Contract No. 10175071; the Department of Science and Technology of India; the BK21 program of the Ministry of Education of Korea and the CHEP SRC program of the Korea Science and Engineering Foundation; the Polish State Committee for Scientific Research under Contract No. 2P03B 01324; the Ministry of Science and Technology of the Russian Federation; the Ministry of Education, Science and Sport of the Republic of Slovenia; the Swiss National Science Foundation; the National Science Council and the Ministry of Education of Taiwan; and the U.S. Department of Energy.

[1] A. Ali, E. Lunghi, C. Greub, and G. Hiller, Phys. Rev. D **66**, 034002 (2002).
[2] T. Hurth, Rev. Mod. Phys. **75**, 1159 (2003).
[3] A. Ali, in Proceedings of the 31st International Conference on High Energy Physics (ICHEP 2002), Amsterdam, The Netherlands, 2002, hep-ph/0210183.
[4] A. Ghinculov, T. Hurth, G. Isidori, and Y. P. Yao, Nucl. Phys. **B685**, 351 (2004).
[5] K. Abe *et al.* (Belle Collaboration), Phys. Rev. Lett. **88**, 021801 (2002).
[6] A. Ishikawa *et al.* (Belle Collaboration), Phys. Rev. Lett. **91**, 261601 (2003).
[7] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **91**, 221802 (2003).
[8] J. Kaneko *et al.* (Belle Collaboration), Phys. Rev. Lett. **90**, 021801 (2003).
[9] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **93**, 081802 (2004).
[10] The semi-inclusive technique was introduced in the context of inclusive $B \to X_s\gamma$ decays; see M. S. Alam *et al.* (CLEO Collaboration), Phys. Rev. Lett. **74**, 2885 (1995).
[11] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods Phys. Res., Sect. A **499**, 1 (2003).
[12] A. Abashian *et al.* (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 117 (2002).
[13] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).
[14] T. Sjöstrand, Comput. Phys. Commun. **82**, 74 (1994).
[15] R. Brun *et al.*, GEANT 3 Manual, CERN Program Library Long Writeup W5013, 1994.
[16] A. Ali, P. Ball, L. T. Handoko, and G. Hiller, Phys. Rev. D **61**, 074024 (2000).
[17] F. Krüger and L. M. Sehgal, Phys. Lett. B **380**, 199 (1996).
[18] A. Ali and E. Pietarinen, Nucl. Phys. **B154**, 519 (1979); G. Altarelli *et al.*, Nucl. Phys. **B208**, 365 (1982).
[19] R. A. Fisher, Ann. Eugen. **7**, 179 (1936).
[20] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).
[21] H. Albrecht *et al.* (ARGUS Collaboration), Phys. Lett. B **192**, 245 (1987).
[22] D. Cronin-Hennessy *et al.* (CLEO Collaboration), Phys. Rev. Lett. **87**, 251808 (2001).
[23] S. Chen *et al.* (CLEO Collaboration), Phys. Rev. Lett. **87**, 251807 (2001).