Measurement of $\sin 2\beta$ from $B \to J/\psi K_S^0$ with the CDF detector

T. Affolder,[21] H. Akimoto,[42] A. Akopian,[35] M. G. Albrow,[10] P. Amaral,[7] S. R. Amendolia,[31] D. Amidei,[24] J. Antos,[1] G. Apollinari,[35] T. Arisawa,[42] T. Asakawa,[40] W. Ashmanskas,[7] M. Atac,[10] P. Azzi-Bacchetta,[29] N. Bacchetta,[29] M. W. Bailey,[26] S. Bailey,[14] P. de Barbaro,[34] A. Barbaro-Galtieri,[21] V. E. Barnes,[33] B. A. Barnett,[17] M. Barone,[12] G. Bauer,[22] F. Bedeschi,[31] S. Belforte,[39] G. Bellettini,[31] J. Bellinger,[43] D. Benjamin,[9] J. Bensinger,[4] A. Beretvas,[10] J. P. Berge,[10] J. Berryhill,[7] S. Bertolucci,[12] B. Bevensee,[30] A. Bhatti,[35] C. Bigongiari,[31] M. Binkley,[10] D. Bisello,[29] R. E. Blair,[2] C. Blocker,[4] K. Bloom,[24] B. Blumenfeld,[17] B. S. Blusk,[34] A. Bocci,[31] A. Bodek,[34] W. Bokhari,[30] G. Bolla,[33] Y. Bonushkin,[5] D. Bortoletto,[33] J. Boudreau,[32] A. Brandl,[26] S. van den Brink,[17] C. Bromberg,[25] N. Bruner,[26] E. Buckley-Geer,[10] J. Budagov,[8] H. S. Budd,[34] K. Burkett,[14] G. Busetto,[29] A. Byon-Wagner,[10] K. L. Byrum,[2] M. Campbell,[24] A. Caner,[31] W. Carithers,[21] J. Carlson,[24] D. Carlsmith,[43] J. Cassada,[34] A. Castro,[29] D. Cauz,[39] A. Cerri,[31] P. S. Chang,[1] P. T. Chang,[1] J. Chapman,[24] C. Chen,[30] Y. C. Chen,[1] M. -T. Cheng,[1] M. Chertok,[37] G. Chiarelli,[31] I. Chirikov-Zorin,[8] G. Chlachidze,[8] F. Chlebana,[10] L. Christofek,[16] M. L. Chu,[1] S. Cihangir,[10] C. I. Ciobanu,[27] A. G. Clark,[13] M. Cobal,[31] E. Cocca,[31] A. Connolly,[21] J. Conway,[36] J. Cooper,[10] M. Cordelli,[12] J. Guimaraes da Costa,[24] D. Costanzo,[31] J. Cranshaw,[38] D. Cronin-Hennessy,[9] R. Cropp,[23] R. Culbertson,[7] D. Dagenhart,[41] F. DeJongh,[10] S. Dell'Agnello,[12] M. Dell'Orso,[31] R. Demina,[10] L. Demortier,[35] M. Deninno,[3] P. F. Derwent,[10] T. Devlin,[36] J. R. Dittmann,[10] S. Donati,[31] J. Done,[37] T. Dorigo,[14] N. Eddy,[16] K. Einsweiler,[21] J. E. Elias,[10] E. Engels, Jr.,[32] W. Erdmann,[10] D. Errede,[16] S. Errede,[16] Q. Fan,[34] R. G. Feild,[44] C. Ferretti,[31] I. Fiori,[3] B. Flaugher,[10] G. W. Foster,[10] M. Franklin,[14] J. Freeman,[10] J. Friedman,[22] Y. Fukui,[20] S. Gadomski,[23] S. Galeotti,[31] M. Gallinaro,[35] T. Gao,[30] M. Garcia-Sciveres,[21] A. F. Garfinkel,[33] P. Gatti,[29] C. Gay,[44] S. Geer,[10] D. W. Gerdes,[24] P. Giannetti,[31] P. Giromini,[12] V. Glagolev,[8] M. Gold,[26] J. Goldstein,[10] A. Gordon,[14] A. T. Goshaw,[9] Y. Gotra,[32] K. Goulianos,[35] H. Grassmann,[39] C. Green,[33] L. Groer,[36] C. Grosso-Pilcher,[7] M. Guenther,[33] G. Guillian,[24] R. S. Guo,[1] C. Haber,[21] E. Hafen,[22] S. R. Hahn,[10] C. Hall,[14] T. Handa,[15] R. Handler,[43] W. Hao,[38] F. Happacher,[12] K. Hara,[40] A. D. Hardman,[33] R. M. Harris,[10] F. Hartmann,[18] K. Hatakeyama,[35] J. Hauser,[5] J. Heinrich,[30] A. Heiss,[18] B. Hinrichsen,[23] K. D. Hoffman,[33] C. Holck,[30] R. Hollebeek,[30] L. Holloway,[16] R. Hughes,[27] J. Huston,[25] J. Huth,[14] H. Ikeda,[40] M. Incagli,[31] J. Incandela,[10] G. Introzzi,[31] J. Iwai,[42] Y. Iwata,[15] E. James,[24] H. Jensen,[10] M. Jones,[30] U. Joshi,[10] H. Kambara,[13] T. Kamon,[37] T. Kaneko,[40] K. Karr,[41] H. Kasha,[44] Y. Kato,[28] T. A. Keaffaber,[33] K. Kelley,[22] M. Kelly,[24] R. D. Kennedy,[10] R. Kephart,[10] D. Khazins,[9] T. Kikuchi,[40] M. Kirk,[4] B. J. Kim,[19] H. S. Kim,[23] S. H. Kim,[40] Y. K. Kim,[21] L. Kirsch,[4] S. Klimenko,[11] D. Knoblauch,[18] P. Koehn,[27] A. Köngeter,[18] K. Kondo,[42] J. Konigsberg,[11] K. Kordas,[23] A. Korytov,[11] E. Kovacs,[2] J. Kroll,[30] M. Kruse,[34] S. E. Kuhlmann,[2] K. Kurino,[15] T. Kuwabara,[40] A. T. Laasanen,[33] N. Lai,[7] S. Lami,[35] S. Lammel,[10] J. I. Lamoureux,[4] M. Lancaster,[21] G. Latino,[31] T. LeCompte,[2] A. M. Lee IV,[9] S. Leone,[31] J. D. Lewis,[10] M. Lindgren,[5] T. M. Liss,[16] J. B. Liu,[34] Y. C. Liu,[1] N. Lockyer,[30] O. Long,[30] M. Loreti,[29] D. Lucchesi,[29] P. Lukens,[10] S. Lusin,[43] J. Lys,[21] R. Madrak,[14] K. Maeshima,[10] P. Maksimovic,[14] L. Malferrari,[3] M. Mangano,[31] M. Mariotti,[29] G. Martignon,[29] A. Martin,[44] J. A. J. Matthews,[26] P. Mazzanti,[3] K. S. McFarland,[34] P. McIntyre,[37] E. McKigney,[30] M. Menguzzato,[29] A. Menzione,[31] E. Meschi,[31] C. Mesropian,[35] C. Miao,[24] T. Miao,[10] R. Miller,[25] J. S. Miller,[24] H. Minato,[40] S. Miscetti,[12] M. Mishina,[20] N. Moggi,[31] E. Moore,[26] R. Moore,[24] Y. Morita,[20] A. Mukherjee,[10] T. Muller,[18] A. Munar,[31] P. Murat,[31] S. Murgia,[25] M. Musy,[39] J. Nachtman,[5] S. Nahn,[44] H. Nakada,[40] T. Nakaya,[7] I. Nakano,[15] C. Nelson,[10] D. Neuberger,[18] C. Newman-Holmes,[10] C. -Y. P. Ngan,[22] P. Nicolaidi,[39] H. Niu,[4] L. Nodulman,[2] A. Nomerotski,[11] S. H. Oh,[9] T. Ohmoto,[15] T. Ohsugi,[15] R. Oishi,[40] T. Okusawa,[28] J. Olsen,[43] C. Pagliarone,[31] F. Palmonari,[31] R. Paoletti,[31] V. Papadimitriou,[38] S. P. Pappas,[44] A. Parri,[12] D. Partos,[4] J. Patrick,[10] G. Pauletta,[39] M. Paulini,[21] A. Perazzo,[31] L. Pescara,[29] M. D. Peters,[21] T. J. Phillips,[9] G. Piacentino,[31] K. T. Pitts,[16] R. Plunkett,[10] A. Pompos,[33] L. Pondrom,[43] G. Pope,[32] F. Prokoshin,[8] J. Proudfoot,[2] F. Ptohos,[12] G. Punzi,[31] K. Ragan,[23] D. Reher,[21] A. Ribon,[29] F. Rimondi,[3] L. Ristori,[31] W. J. Robertson,[9] A. Robinson,[23] T. Rodrigo,[6] S. Rolli,[41] L. Rosenson,[22] R. Roser,[10] R. Rossin,[29] W. K. Sakumoto,[34] D. Saltzberg,[5] A. Sansoni,[12] L. Santi,[39] H. Sato,[40] P. Savard,[23] P. Schlabach,[10] E. E. Schmidt,[10] M. P. Schmidt,[44] M. Schmitt,[14] L. Scodellaro,[29] A. Scott,[5] A. Scribano,[31] S. Segler,[10] S. Seidel,[26] Y. Seiya,[40] A. Semenov,[8] F. Semeria,[3] T. Shah,[22] M. D. Shapiro,[21] P. F. Shepard,[32] T. Shibayama,[40] M. Shimojima,[40] M. Shochet,[7] J. Siegrist,[21] G. Signorelli,[31] A. Sill,[38] P. Sinervo,[23] P. Singh,[16] A. J. Slaughter,[44] K. Sliwa,[41] C. Smith,[17] F. D. Snider,[10] A. Solodsky,[35] J. Spalding,[10] T. Speer,[13] P. Sphicas,[22] F. Spinella,[31] M. Spiropulu,[14] L. Spiegel,[10] L. Stanco,[29] J. Steele,[43] A. Stefanini,[31] J. Strologas,[16] F. Strumia,[13] D. Stuart,[10] K. Sumorok,[22] T. Suzuki,[40] R. Takashima,[15] K. Takikawa,[40] M. Tanaka,[40] T. Takano,[28] B. Tannenbaum,[5] W. Taylor,[23] M. Tecchio,[24] P. K. Teng,[1] K. Terashi,[40] S. Tether,[22] D. Theriot,[10] R. Thurman-Keup,[2] P. Tipton,[34] S. Tkaczyk,[10] K. Tollefson,[34] A. Tollestrup,[10] H. Toyoda,[28] W. Trischuk,[23] J. F. de Troconiz,[14] S. Truitt,[24] J. Tseng,[22] N. Turini,[31] F. Ukegawa,[40] J. Valls,[36] S. Vejcik III,[10] G. Velev,[31] R. Vidal,[10] R. Vilar,[6] I. Vologouev,[21] D. Vucinic,[22] R. G. Wagner,[2] R. L. Wagner,[10] J. Wahl,[7] N. B. Wallace,[36] A. M. Walsh,[36] C. Wang,[9] C. H. Wang,[1] M. J. Wang,[1] T. Watanabe,[40] T. Watts,[36] R. Webb,[37] H. Wenzel,[18] W. C. Wester III,[10] A. B. Wicklund,[2] E. Wicklund,[10] H. H. Williams,[30] P. Wilson,[10] B. L. Winer,[27] D. Winn,[24] S. Wolbers,[10] D. Wolinski,[24] J. Wolinski,[25] S. Worm,[26] X. Wu,[13] J. Wyss,[31] A. Yagil,[10] W. Yao,[21] G. P. Yeh,[10] P. Yeh,[1] J. Yoh,[10] C. Yosef,[25] T. Yoshida,[28] I. Yu,[19] S. Yu,[30] A. Zanetti,[39] F. Zetti,[21] and S. Zucchelli[3]

(CDF Collaboration)

[1]*Institute of Physics, Academia Sinica, Taipei, Taiwan 11529, Republic of China*
[2]*Argonne National Laboratory, Argonne, Illinois 60439*
[3]*Istituto Nazionale di Fisica Nucleare, University of Bologna, I-40127 Bologna, Italy*
[4]*Brandeis University, Waltham, Massachusetts 02254*
[5]*University of California at Los Angeles, Los Angeles, California 90024*
[6]*Instituto de Fisica de Cantabria, University of Cantabria, 39005 Santander, Spain*
[7]*Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637*
[8]*Joint Institute for Nuclear Research, RU-141980 Dubna, Russia*
[9]*Duke University, Durham, North Carolina 27708*
[10]*Fermi National Accelerator Laboratory, Batavia, Illinois 60510*
[11]*University of Florida, Gainesville, Florida 32611*
[12]*Laboratori Nazionali di Frascati, Istituto Nazionale di Fisica Nucleare, I-00044 Frascati, Italy*
[13]*University of Geneva, CH-1211 Geneva 4, Switzerland*
[14]*Harvard University, Cambridge, Massachusetts 02138*
[15]*Hiroshima University, Higashi-Hiroshima 724, Japan*
[16]*University of Illinois, Urbana, Illinois 61801*
[17]*The Johns Hopkins University, Baltimore, Maryland 21218*
[18]*Institut für Experimentelle Kernphysik, Universität Karlsruhe, 76128 Karlsruhe, Germany*
[19]*Korean Hadron Collider Laboratory: Kyungpook National University, Taegu 702-701; Seoul National University, Seoul 151-742; and SungKyunKwan University, Suwon 440-746; Korea*
[20]*High Energy Accelerator Research Organization (KEK), Tsukuba, Ibaraki 305, Japan*
[21]*Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, California 94720*
[22]*Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*
[23]*Institute of Particle Physics: McGill University, Montreal H3A 2T8; and University of Toronto, Toronto M5S 1A7; Canada*
[24]*University of Michigan, Ann Arbor, Michigan 48109*
[25]*Michigan State University, East Lansing, Michigan 48824*
[26]*University of New Mexico, Albuquerque, New Mexico 87131*
[27]*The Ohio State University, Columbus, Ohio 43210*
[28]*Osaka City University, Osaka 588, Japan*
[29]*Universita di Padova, Istituto Nazionale di Fisica Nucleare, Sezione di Padova, I-35131 Padova, Italy*
[30]*University of Pennsylvania, Philadelphia, Pennsylvania 19104*
[31]*Istituto Nazionale di Fisica Nucleare, University and Scuola Normale Superiore of Pisa, I-56100 Pisa, Italy*
[32]*University of Pittsburgh, Pittsburgh, Pennsylvania 15260*
[33]*Purdue University, West Lafayette, Indiana 47907*
[34]*University of Rochester, Rochester, New York 14627*
[35]*Rockefeller University, New York, New York 10021*
[36]*Rutgers University, Piscataway, New Jersey 08855*
[37]*Texas A&M University, College Station, Texas 77843*
[38]*Texas Tech University, Lubbock, Texas 79409*
[39]*Istituto Nazionale di Fisica Nucleare, University of Trieste/Udine, Italy*
[40]*University of Tsukuba, Tsukuba, Ibaraki 305, Japan*
[41]*Tufts University, Medford, Massachusetts 02155*
[42]*Waseda University, Tokyo 169, Japan*
[43]*University of Wisconsin, Madison, Wisconsin 53706*
[44]*Yale University, New Haven, Connecticut 06520*

(Received 8 September 1999; published 9 March 2000)

This paper reports an updated measurement of the standard model CP violation parameter $\sin 2\beta$ using the CDF Detector at Fermilab. The entire run I data sample of 110 pb^{-1} of proton-antiproton collisions at $\sqrt{s}=1.8$ TeV is used to identify a signal sample of ~ 400 $B \to J/\psi K_S^0$ events, where $J/\psi \to \mu^+\mu^-$ and $K_S^0 \to \pi^+\pi^-$. The flavor of the neutral B meson is identified at the time of production by combining information from three tagging algorithms: a same-side tag, a jet-charge tag, and a soft-lepton tag. A maximum likelihood fitting method is used to determine $\sin 2\beta = 0.79^{+0.41}_{-0.44}(\mathrm{stat}+\mathrm{syst})$. This value of $\sin 2\beta$ is consistent with the standard model prediction, based upon existing measurements, of a large positive CP-violating asymmetry in this decay mode.

PACS number(s): 12.15.Hh, 13.20.He, 14.40.Nd

I. INTRODUCTION

The first observation of a violation of charge-conjugation parity (CP) invariance was in the neutral kaon system in 1964 [1]. To date, violation of CP symmetry has not been directly observed in any other system. The study of CP violation in the B system is an ideal place to test the predictions of the standard model [2–4]. The decays of neutral B mesons into CP eigenstates are of great interest, in particular the CP-odd state, $B\to J/\psi K_S^0$ [5,6]. The decay $B\to J/\psi K_S^0$ is a popular mode in which to observe a CP-violating asymmetry because it has a distinct experimental signature and is known theoretically to be free of large hadronic uncertainties [7]. Furthermore, the contribution to the asymmetry due to penguin diagrams, which is difficult to calculate, is negligible because the penguin contribution is small and the tree level and penguin diagrams contribute with the same weak phase [8]. Previous work searching for a CP-violating asymmetry in the decay $B\to J/\psi K_S^0$ has been presented by the OPAL Collaboration [9]. An initial study on the measurement of $\sin 2\beta$ by the CDF Collaboration is given in Ref. [10]. The result reported here incorporates and supersedes Ref. [10]. This paper reports a measurement of $\sin 2\beta$ that is the best direct indication of a CP-violating asymmetry in the neutral B meson system.

Within the framework of the standard model, CP nonconservation arises through a non-trivial phase in the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix [11]. The CKM matrix V is the unitary matrix that transforms the mass eigenstates into the weak eigenstates:

$$V=\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}$$

$$\simeq\begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho-i\eta) \\ -\lambda & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix}+O(\lambda^4).$$

The second matrix is a useful phenomenological parametrization of the quark mixing matrix suggested by Wolfenstein [12], in which λ is the sine of the Cabibbo angle. The condition of unitarity, $V^\dagger V=1$, yields several relations, the most important of which is a relation between the first and third columns of the matrix, given by

$$V_{ub}^* V_{ud}+V_{cb}^* V_{cd}+V_{tb}^* V_{td}=0.$$

This relation, after division by $V_{cb}^* V_{cd}$, is displayed graphically in Fig. 1 as a triangle in the complex (ρ-η) plane, and is known as the unitarity triangle [13]. CP violation in the standard model manifests itself as a nonzero value of η, the height of the triangle.

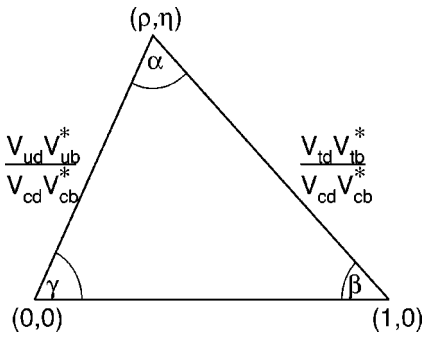

FIG. 1. The unitarity triangle indicating the relationship between the CKM elements.

CP nonconservation is expected to manifest itself in the B_d^0 system [2] as an asymmetry in particle decay rate versus antiparticle decay rate to a particular final state:

$$A_{CP}=\frac{N(\overline{B}^0\to J/\psi K_S^0)-N(B^0\to J/\psi K_S^0)}{N(\overline{B}^0\to J/\psi K_S^0)+N(B^0\to J/\psi K_S^0)},$$

where $N(\overline{B}^0\to J/\psi K_S^0)$ is the number of mesons decaying to $J/\psi K_S^0$ that were produced as $\overline{B}^0$ and $N(B^0\to J/\psi K_S^0)$ is the number of mesons decaying to $J/\psi K_S^0$ that were produced as B^0 [3]. It should be noted that the definition of A_{CP} is the negative of that in Refs. [8] and [9].

In the standard model, the CP asymmetry in this decay mode is proportional to $\sin 2\beta$: $A_{CP}(t)=\sin 2\beta \sin(\Delta m_d t)$, where β is the angle of the unitarity triangle shown in Fig. 1, t is the proper decay time of the B^0 meson and Δm_d is the mass difference between the heavy and light B^0 mass eigenstates. In a hadron collider, $B\overline{B}$ pairs are produced as two incoherent meson states. Consequently, the asymmetry can be measured as either a time-dependent or time-integrated quantity. The time-dependent analysis is however statistically more powerful. In this paper, we take advantage of this fact and employ a sample of events that have a broad range of time resolutions.

It is possible to combine information from several measurements to indirectly constrain the allowed range of $\sin 2\beta$. Based on global fits to these measurements, it is found that the standard model prefers a large positive value of $\sin 2\beta$ and that the fits are in good agreement with each other [14–17]. One recent global fit finds $\sin 2\beta=0.75\pm 0.09$ [17]. However, the sign of the expected asymmetry depends on the sign of the product of B_B and B_K, which are the ratios between the short distance contributions to $B\overline{B}$ and $K\overline{K}$ mixing respectively and their values in the vacuum insertion approximation [18].

To measure this asymmetry, the flavor of the B meson (whether it is a B^0 or a $\overline{B}^0$) must be identified (tagged) at the time of production. The effectiveness of a tagging algorithm depends on both the efficiency for assigning a flavor tag and the probability that the flavor tag is correct. The true asymmetry is ''diluted'' by misidentifying a B^0 meson as a $\overline{B}^0$ meson or *vice versa*. We define the tagging dilution as D

$=(N_R-N_W)/(N_R+N_W)$, where $N_R(N_W)$ is the number of right (wrong) tags. The observed asymmetry, given by $A_{CP}^{\rm obs} = DA_{CP}$, is reduced in magnitude by this dilution parameter. As can be seen from the relation above, maximal sensitivity to the asymmetry is achieved when the dilution factor is large. The statistical uncertainty on $\sin 2\beta$ is inversely proportional to $\sqrt{\epsilon D^2}$, where the efficiency ϵ is the fraction of events that are tagged. This analysis combines three tagging algorithms in order to minimize the statistical uncertainty of the measurement.

A. The CDF detector

The collider detector at Fermilab (CDF) detector is described in detail elsewhere [19,20]. The CDF detector systems that are relevant for this analysis are (i) a silicon vertex detector (SVX) [21], (ii) a time projection chamber (VTX), (iii) a central tracking chamber (CTC), (iv) electromagnetic and hadronic calorimeters, (v) a preshower detector (CPR, central preradiator), (vi) a shower maximum detector (CES, central electron strip chamber), and (vii) a muon system. The CDF coordinate system has the *z*-axis pointing along the proton momentum, with the *x*-axis located in the horizontal plane of the Tevatron storage ring, pointing radially outward, so that the *y*-axis points up.

The SVX consists of four layers of silicon axial-strip detectors located between radii of 2.9 and 7.9 cm and extending ± 25 cm in z from the center of the detector. The geometrical acceptance of the SVX is $\sim 60\%$ because the $p\bar{p}$ interactions are distributed with a Gaussian profile along the beam axis with a standard deviation of ~ 30 cm, which is large relative to the length of the detector. The SVX is surrounded by the VTX, which is used to determine the z coordinate of the $p\bar{p}$ interaction (the primary vertex). Momenta of charged particles are measured in three dimensions using the CTC, an 84-layer drift chamber that covers the pseudorapidity interval $|\eta|<1.1$, where $\eta=-\ln[\tan(\theta/2)]$, and the angle θ is measured from the z-axis. The SVX, VTX, and CTC are immersed in a 1.4 T solenoidal magnetic field. The momentum transverse to the beamline (P_T) of a charged particle is determined using the SVX and CTC detectors. The combined CTC/SVX P_T resolution is $\delta P_T/P_T = [(0.001\ c/\mathrm{GeV}\cdot P_T)^2+(0.0066)^2]^{1/2}$. The typical uncertainty on the B meson decay distance is about $60\mu m$. The CTC also provides measurements of the energy loss per unit distance, dE/dx, of a charged particle.

The central and endwall calorimeters are arranged in projective towers and cover the central region $|\eta|<1.05$. In the central electromagnetic calorimeter, proportional chambers (CES), are embedded near shower maximum for position measurements. The CPR is located on the inner face of the central calorimeter and consists of proportional chambers. The muon system consists of three different subsystems each containing four layers of drift chambers. The central muon chambers, located behind ~ 5 absorption lengths of calorimeter, cover 85% of the azimuthal angle ϕ in the range $|\eta|<0.6$. Gaps in ϕ are filled in part by the central muon upgrade chambers with total coverage in ϕ of 80% and $|\eta|<0.6$. These chambers are located behind a total of ~ 8 absorption lengths. Finally, the central extension muon chambers provide 67% coverage in ϕ for the region $0.6<|\eta|<1.0$ behind a total of ~ 6 absorption lengths.

Muons, used to reconstruct the J/ψ meson and by the soft lepton tagging algorithm (SLT), are identified by combining a muon track segment with a CTC track. SVX information is used when available. Electrons, which are used by the SLT, are identified by combining a CTC track with information from the central calorimeters, the central strip chambers, dE/dx, and the CPR detectors.

Dimuon events are collected using a three-level trigger. The first-level trigger system requires two charged track segments in the muon chambers. The second level trigger requires a CTC track, with P_T greater than ~ 2 GeV/c, to match a muon chamber track segment. The third level, implemented with online track reconstruction software, requires two oppositely charged CTC tracks to match muon track segments and a dimuon invariant mass between 2.8 and 3.4 GeV/c^2. Approximately two thirds of all $J/\psi\to\mu^+\mu^-$ events recorded enter on a dedicated J/ψ trigger, where the two reconstructed muons are from the J/ψ. This fraction is consistent with expectations. The majority of the remaining events, referred to as ''volunteers,'' enter the sample through a single inclusive muon trigger caused by one of the two muons from the J/ψ decay, or, through a dimuon trigger where one of the two trigger muons was from the J/ψ and the second ''trigger muon'' is a fake muon, primarily due to punch-through.

B. Overview of the analysis

This analysis builds on the work of several previous analyses using the various B enriched data sets recorded by the CDF detector. The $B\to J/\psi K_S^0$ decay mode is reconstructed in a manner similar to the CDF measurements of the branching ratio [22,23] and the B lifetime [24]. The three tagging algorithms are then applied to the $B\to J/\psi K_S^0$ sample and the observed asymmetry, given by $A_{CP}^{\rm obs}=DA_{CP}$, is then determined. In order to extract a value of $\sin 2\beta$ from the observed asymmetry, tagging dilution parameters are required for the three tagging algorithms. These dilution parameters are determined from an analysis of the calibration samples. In particular, the same-side tagging (SST) dilutions are determined from a combination of results from Ref. [10] and measurements on a sample of ~ 1000 $B^\pm\to J/\psi K^\pm$ decays. The jet-charge tag algorithm (JETQ) and soft-lepton tag algorithm (SLT) dilutions are determined from the $B^\pm\to J/\psi K^\pm$ sample and $\sim 40\,000$ inclusive $B\to J/\psi X$ events. The dilutions and efficiencies are then combined for each event and a maximum likelihood fitting procedure is used to extract the result for $\sin 2\beta$. The fit includes the possibility that the tagging dilutions and efficiencies have inherent asymmetries. In addition, the backgrounds, divided into prompt and long-lived categories, are also allowed to have an asymmetry. In the end, these possible asymmetries are found not to be significant.

Each flavor tagging method, SST, SLT, and JETQ, has been previously verified in a B^0-$\bar{B}^0$ mixing analysis. Our

previously published measurement of $\sin 2\beta$ used the B^0-$\overline{B^0}$ mixing analysis of Ref. [25] to establish the viability of the SST method [26]. Here we report work that uses the same algorithm for events where the two muons are contained within the SVX detector acceptance and uses a modified version of the algorithm for events with less precise flight path information, i.e. events not fully contained within the SVX detector acceptance.

The two additional tagging algorithms used are based on the B^0-$\overline{B^0}$ mixing analysis of Ref. [27]. These mixing analyses use decays of B mesons with higher P_T ($\sim$ a factor of 2 higher) than the B mesons in this analysis. This is due to the lower trigger threshold for $J/\psi \to \mu^+\mu^-$ than for the inclusive lepton triggers used to select the mixing analyses samples. The SLT algorithm is similar to that in Ref. [27], except the lepton P_T threshold has been lowered to increase the efficiency of tagging lower P_T B mesons. The JETQ algorithm is also similar to the algorithm used in the mixing analysis [27] except the acceptance cone defining the jet has been enlarged and impact parameter weighting of tracks has been added to reduce the fraction of incorrectly tagged events.

II. SAMPLE SELECTION

Four event samples, $B \to J/\psi K_S^0$, $B^\pm \to J/\psi K^\pm$, inclusive $B \to J/\psi X$ decays, and an inclusive lepton sample [25] are used in the determination of $\sin 2\beta$. The B mesons are reconstructed using the decay modes $J/\psi \to \mu^+\mu^-$ and $K_S^0 \to \pi^+\pi^-$. The $B \to J/\psi K_S^0$ candidates form the signal sample, the $B^\pm \to J/\psi K^\pm$ sample is used to determine the tagging dilutions, and the inclusive J/ψ decays are used to constrain ratios of efficiencies. The inclusive lepton sample was used in Refs. [10,25] in the determination of the SST dilution.

The selection criteria are largely the same as in Ref. [10]. The criteria for the $B \to J/\psi K_S^0$ sample provide an optimal value of the ratio $S^2/(S+N_{\rm bck})$, where S is the number of signal events and $N_{\rm bck}$ is the number of background events within three standard deviations of the B mass. The square root of this ratio enters into the uncertainty on the measurement of $\sin 2\beta$. The J/ψ is identified by selecting two oppositely charged muon candidates, each with $P_T > 1.4$ GeV/c. Additional selection criteria are applied to ensure good matching between the CTC track and the muon chamber track segment. A J/ψ candidate is defined as a $\mu^+\mu^-$ pair within $\pm 5\sigma$ of the world average mass of 3.097 GeV/c^2 [8], where σ is the mass uncertainty calculated for each event.

The K_S^0 candidates are found by matching pairs of oppositely charged tracks, assumed to be pions. The K_S^0 candidates are required to travel a significant distance $L_{xy} > 5\sigma_L$, and to have $P_T > 700$ MeV/c in order to improve the signal-to-background ratio. The quantity $L_{xy} = \mathbf{X} \cdot \hat{P}_T$ is the two-dimensional (2D) flight distance, where $\mathbf{X}$ is the vector pointing from the production vertex to the decay vertex, and σ_L is the measurement uncertainty on L_{xy}. This flight distance is used to calculate the proper decay time $t = L_{xy} M_0 / P_T$, where M_0 is the world average B^0 mass of 5.2792 GeV/c^2 [8]. In about 15% of the K_S^0 decays, SVX information is available for one or both tracks. When the decay vertex location in the radial direction is found to lie beyond the second layer of the SVX detector, the SVX information is not used. The J/ψ and K_S^0 candidates are combined into a four particle fit to the hypothesis $B \to J/\psi K_S^0$ and the $\mu^+\mu^-$ and $\pi^+\pi^-$ are constrained to the appropriate masses and separate decay vertices. The K_S^0 and B are constrained to point back to their points of origin. In order to further improve the signal-to-background ratio, B candidates are accepted for $P_T(B) > 4.5$ GeV/c and fit quality criteria are applied to the J/ψ and B candidates.

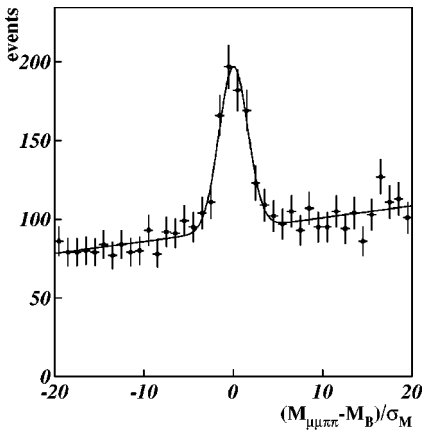

FIG. 2. The normalized mass distribution of the $J/\psi K_S^0$ candidates. The curve is a Gaussian signal plus linear background from a maximum likelihood fit.

The data are divided into two samples, one called the SVX sample, the other the non-SVX sample. The SVX sample requires both muon candidates to have at least three out of four possible hits that are well measured by the silicon vertex detector. This is the sample of B candidates with precise decay length information and is similar to the sample that was used in the previously published CDF $\sin 2\beta$ analysis. The non-SVX sample is the subset of events in which one or both muon candidates are not measured in the silicon vertex detector. About 30% of the events in this sample have one muon candidate track with high quality SVX information. Events of this type lie mostly at the boundaries of the SVX detector.

We define a normalized mass $M_N = (m_{\mu\mu\pi\pi} - M_0)/\sigma_{\rm fit}$, where $m_{\mu\mu\pi\pi}$ is the four-track mass coming from the vertex and mass-constrained fit of the B candidate. The uncertainty, $\sigma_{\rm fit}$, is from the fit, typically ~ 10 MeV/c^2. The normalized mass distribution is shown in Fig. 2 and contains 4156 entries, from which we observe 395 ± 31 signal events with a signal-to-noise ratio of 0.7. The SVX sample contains 202 ± 18 events (signal-to-noise ratio of 0.9) and the non-SVX sample contains 193 ± 26 events (signal-to-noise ratio of 0.5) as shown in Fig. 3. The event yields reported here come from the full unbinned likelihood fit which will be described in detail later.

The criteria used to select the $B^\pm \to J/\psi K^\pm$ decays are the same as described for $B \to J/\psi K_S^0$ decays except for the $K^\pm$

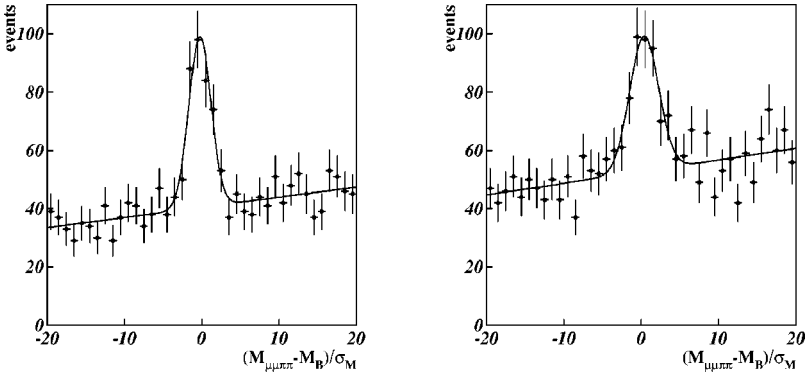

FIG. 3. Left: Normalized mass distribution of the $J/\psi K_S^0$ candidates where both muons have good SVX information providing a high precision decay length measurement. Right: Normalized mass distribution of the $J/\psi K_S^0$ candidates in the non-SVX sample. Either one or both muons are missing good SVX information, leading to a low resolution decay length. For both plots, the curves are Gaussian signals plus linear background.

selection. Since the CDF detector has limited particle identification separation power at high P_T using the dE/dx system, candidate kaons are defined as any track with $P_T > 2$ GeV/c. The $\mu^+\mu^- K^\pm$ mass distribution is shown in Fig. 4 and the number of $J/\psi K^\pm$ candidates is 998 ± 51.

The inclusive $J/\psi \to \mu^+\mu^-$ sample is a superset from which the $B \to J/\psi K_S^0$ and $B^\pm \to J/\psi K^\pm$ samples are derived. The inclusive sample is $\sim 80\%$ prompt J/ψ from direct $c\bar{c}$ production. In order to enrich the sample in $B \to J/\psi X$ decays, both muons are required to have good SVX information and the J/ψ 2D travel distance must be > 200 μm from the beamline. This results in a sample of about 40 000 $B \to J/\psi X$ decays.

III. TAGGING ALGORITHMS

Three tagging algorithms are used, two opposite-side tag algorithms and one same-side tag (SST) algorithm. The idea behind the SST algorithm [26] exploits the local correlation between the B meson and the charge of a nearby track to tag the flavor of the B meson. We employ the SST algorithm described in detail in Refs. [10,25]. We consider all charged tracks that pass through all stereo layers of the CTC and within a cone of radius $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.7$ centered along the B meson direction. Candidate tracks must be consistent with originating from the primary vertex and have a $P_T > 400$ MeV/c. If more than one candidate is found, the track with the smallest $P_T^{\rm rel}$ is chosen, where $P_T^{\rm rel}$ is the track momentum transverse to the momentum sum of the track and the B meson. A tagging track with negative charge indicates a $\bar{B}^0$ meson, while a positive track indicates a B^0 meson.

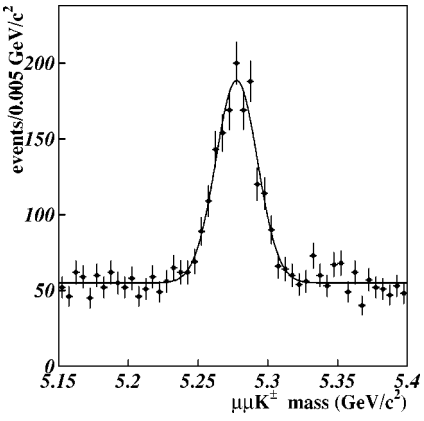

FIG. 4. The mass distribution of the $J/\psi K^\pm$ candidates both with and without SVX information. The curve is a Gaussian signal plus linear background from the likelihood fit.

The performance of the SST algorithm could depend on the availability of precise vertex information. When using the SVX sample, the SST algorithm of Ref. [10] and tagging dilution parameter $D = (16.6 \pm 2.2)\%$ is used. This dilution result is obtained by extrapolating the value obtained in the mixing analysis in Ref. [25] to the lower P_T of the $B \to J/\psi K_S^0$ sample. When using the non-SVX sample, the SST algorithm is modified slightly by dropping the SVX information for all candidate tagging tracks and adjusting the track selection criteria in order to increase the geometrical acceptance. A dilution scale factor f_D, defined by $D_{\text{non-SVX}} = f_D D_{\text{SVX}}$, is derived from the $B^\pm \to J/\psi K^\pm$ sample. This relates the SVX sample SST algorithm performance to that of the non-SVX sample SST algorithm. To measure this quantity, we compare the tagging track using SVX information to the track we obtain when all SVX information is ignored. This provides a measure of the effectiveness of the SVX information. We find a value of $f_D = (1.05 \pm 0.17)$, apply it to the measured SST dilution for SVX tracks, and obtain $D = (17.4 \pm 3.6)\%$.

Opposite-side tagging refers to the identification of the flavor of the ''opposite'' B in the event at the time of production. As mentioned earlier, two algorithms are employed: soft-lepton tag (SLT) and jet-charge tag (JETQ) algorithms.

The SLT algorithm is described in detail in Ref. [27]. The SLT algorithm associates the charge of the lepton (electron or muon) with the flavor of the parent B-meson, which in turn is anticorrelated with the produced flavor of the B-meson that decays to $J/\psi K_S^0$. These leptons are considered ''soft'' because their momenta are on average considerably lower than the high momentum leptons from W boson, Z boson, and top quark decays. A soft muon tag is defined as a charged track reconstructed in the CTC (CTC track) with $P_T > 2$ GeV/c that has been matched to a track segment in a muon system. A soft electron tag is defined as a CTC track with $P_T > 1$ GeV/c that has been successfully extrapolated into the calorimeters, CPR and CES detectors and passed selection criteria. In particular, the CPR and CES position information is required to match with the CTC track and the shower profiles must be consistent with an electron. In addition, the electron candidate CTC track must have a dE/dx

TABLE I. Summary of tagging algorithms performance. All numbers listed are in percent. The efficiencies are obtained from the $B \to J/\psi K_S^0$ sample. The dilution information is derived from the $B^\pm \to J/\psi K^\pm$ sample.

Tag side	Tag type	Class	Efficiency	Dilution
Same-side	SST	μ_1, μ_2 in SVX	35.5 ± 3.7	16.6 ± 2.2
	SST	μ_1 or μ_2 non-SVX	38.1 ± 3.9	17.4 ± 3.6
Opposite side	SLT	all events	5.6 ± 1.8	62.5 ± 14.6
	JETQ	all events	40.2 ± 3.9	23.5 ± 6.9

deposition consistent with an electron. Photon conversions are explicitly rejected. A dilution of $D=(62.5\pm 14.6)\%$ is obtained by applying the SLT algorithm to the $B^\pm \to J/\psi K^\pm$ sample.

If a soft lepton is not found, we try to identify a jet produced by the opposite B. We calculate a quantity called the jet charge Q_{jet} of this jet:

$$Q_{\rm jet}=\frac{\sum_i q_i P_{Ti}[2-(T_p)_i]}{\sum_i P_{Ti}[2-(T_p)_i]},$$

where q_i and P_{Ti} are the charge and transverse momentum of the ith track in the jet with $P_T > 750$ MeV/c. The quantity T_p is the probability that track i originated from the $p\bar{p}$ interaction point. The quantity $(2-T_p)$ is constructed such that a displaced (prompt) track has the value $T_p \sim 0(1)$, and the quantity $(2-T_p)$ is $\sim 2(1)$. Tracks that arise from B decays are displaced from the primary vertex and give a probability distribution T_p peaked near zero, lending larger weight to the sum. For tracks that emanate from the primary vertex, T_p is a flat distribution between 0 and 1, giving less weight to the jet charge quantity. For b-quark jets, the sign of the jet charge is on average the same as the sign of the b-quark that produced the jet, so the sign of the jet charge may be used to identify the flavor at production of the B hadron which decayed to $J/\psi K_S^0$. This algorithm is conceptually similar to that used in Ref. [27] except that jet clustering and weighting factors are optimized for this sample. This optimization was performed by maximizing ϵD^2 on a sample of $B^\pm \to J/\psi K^\pm$ events generated by a Monte Carlo program.

Jets are found with charged particles instead of the more commonly used calorimeter clusters. The algorithm is optimized using Monte Carlo generated data. All tracks in an event with $P_T > 1.75$ GeV/c are identified as seed tracks. For pairs of seed tracks, the quantity $Y_{ij}=2E_iE_j(1-\cos\theta_{ij})$ is calculated, where E_i, E_j are the energies and θ_{ij} is the angle between the ith and jth seed tracks. Seed tracks are combined in pairs as long as Y_{ij}, the JADE distance measure, is less than 24 GeV^2. After mergings, each set of seed tracks defines a jet. The remaining tracks ($P_T < 1.75$ GeV/c) are combined with the jet that minimizes the distance measure provided that $Y_{ij} < 24$ GeV^2. Any tracks unassociated with a track-group are discarded. This is a modified version of the JADE clustering algorithm [28].

Tracks within a cone of $\Delta R < 0.7$ with respect to the $B \to J/\psi K_S^0$ direction are excluded from clustering to avoid overlap with the SST candidate tracks. The B meson decay products (μ^+, μ^-, π^+, and π^-) are also explicitly excluded from the track-group. A jet can consist of a single track with $P_T > 1.75$ GeV/c. If multiple jets are found, we choose the one that is most likely a B jet, based on an algorithm that uses the track impact parameter information first, if available, and then the jet P_T. The momentum and impact parameter weighted charge, $Q_{\rm jet}$, is calculated for the jet and normalized such that $|Q_{\rm jet}| \leqslant 1$. Only tracks with $P_T > 0.750$ GeV/c are used to weight the charge. The parameter $Q_{\rm jet} > 0.2$ selects the $\bar{b}$ quark decays and $Q_{\rm jet} < -0.2$ selects the b quark decays. The value $|Q_{\rm jet}| \leqslant 0.2$ is considered untagged. A dilution of $D=(23.5\pm 6.9)\%$ is found by applying the JETQ algorithm to the $B^\pm \to J/\psi K^\pm$ sample.

We use a sample of 998 ± 51 $B^\pm \to J/\psi K^\pm$ decays to determine the tagging dilutions for the opposite-side algorithms. Using both real data and simulated data, we have verified that $D(B^\pm)$ is consistent with $D(B^0)$ for the opposite-side flavor tagging algorithms. At the Tevatron, the strong interaction creates $b\bar{b}$ pairs at a production energy sufficiently high that the fragmentation processes that create the B mesons are largely uncorrelated. For example, the b quark could hadronize as a B^- meson, while independently, the $\bar{b}$ quark could hadronize as a B^+, B^0 or B_s^0 meson. These opposite side dilution numbers are valid for both the SVX and non-SVX samples. The tagging dilutions and efficiencies are presented in Table I.

Each event has the opportunity to be tagged by two tag algorithms: one same-side and one opposite-side. We followed the prescription outlined in Ref. [27] in which the SLT tag is used if both the SLT and JETQ tags are available. This is done to avoid correlations between the two opposite side tagging algorithms. The result of the SLT algorithm is used because the dilution of the SLT algorithm is much larger than that of the JETQ algorithm. Given the low efficiency for lepton tags (6%) the potential overlap is small. As mentioned earlier, tracks eligible for the SST algorithm are excluded from the JETQ track list, thus ensuring these two algorithms are orthogonal. There is however an overlap between the SST and the SLT algorithms in which the lepton is used as the SST track. In order to use the dilution measured in Ref. [10], we use the identical SST algorithm on the SVX sample, and therefore permit this overlap. We allow leptons in the cone to account for $b\bar{b}$ production from the higher-

TABLE II. Definition of tags. For the case of the SST algorithm, the tag depends upon the charge of a track (t^+,t^-) near the B; for the SLT algorithm, the tag depends upon the charge of a lepton in the event (l^+,l^-); for the JETQ algorithm, the tag depends upon the average weighted charge of tracks in a jet ($Q_{\rm jet}$).

Tag	Positive (+) tag $B^0\to J/\psi K_S^0$	Negative (−) tag $\overline{B}^0\to J/\psi K_S^0$	No tag
SST	Single track t^+	Single track t^-	No track
SLT	Single lepton l^-	Single lepton l^+	No lepton
JETQ	$Q_{\rm jet}<-0.20$	$Q_{\rm jet}>0.20$	$\lvert Q_{\rm jet}\rvert\leq 0.20$

order gluon splitting process where the $b\to lX$ decay is located nearby the fully reconstructed $B\to J/\psi K_S^0$. This overlap occurs in three events in the signal region and the final result changes negligibly if these events are removed from the sample.

Based upon the tagging efficiency of each individual tagging algorithm, we can calculate the expected fraction of events which will be tagged by two, one or zero algorithms. We find the expected efficiency of each combination of tags (e.g. events tagged by both SST and SLT, events tagged by JETQ only, etc.) is consistent with estimates derived from a study of tagging efficiencies as applied to the $B^\pm\to J/\psi K^\pm$ sample. Tag efficiencies are higher, typically by $\sim$10%, in the trigger volunteer sample, except for the JETQ tagging algorithm, in which the efficiency increases by about 17%. These higher efficiencies are due to the increased average charged-track multiplicity of the trigger volunteer sample. Thus trigger samples that do not include volunteers, as planned for run II, will have lower tagging efficiencies. It is found that $\sim$80% of the events in the entire $B\to J/\psi K_S^0$ sample are tagged by at least one tagging algorithm.

Tag sign definition

An event is tagged if it satisfies the criteria of any of the three tag algorithms. For all tag algorithms, the flavor tag refers to whether the candidate $B\to J/\psi K_S^0$ was produced as a B^0 or $\overline{B}^0$. The sign of all tag algorithms follow the convention established by the same-side tag algorithm discussed in Ref. [10]: The positive tag (+ tag) is defined as the identification of a $\overline{b}$-quark and therefore a B^0 meson. The negative tag (− tag) is defined as the identification of a b-quark and therefore a $\overline{B}^0$ meson. A null tag (or tag 0) means the criteria of the tag algorithms were not satisfied, and the flavor of the B is not identified. A summary is provided in Table II.

IV. DILUTIONS, EFFICIENCIES AND TAGGING ASYMMETRIES

The dilutions and efficiencies described earlier need to be generalized in order to accommodate possible detector asymmetries in the analysis. For example, the CTC has a small ($\sim$1%) bias toward reconstructing more tracks of positive charge at low transverse momentum. This small bias is due to the tilted drift cell that is necessary to compensate for the Lorentz angle of the drift electrons, and a known asymmetry in background tracks from beam pipe interactions. The formalism for measuring and correcting for these possible tagging asymmetries in this multitag analysis is provided below.

For B mesons decaying to a CP eigenstate, the decay rate as a function of proper time t can be written as

$$h_\pm(t)=\frac{e^{-t/\tau}}{2\tau}[1\pm\Lambda_{CP}\sin(\Delta m_d t)],$$

where $h_+(t)$ is the decay rate for B's produced as type ''+,'' $h_-(t)$ is the decay rate for B's produced as type ''−,'' and $\Lambda_{CP}=-\sin 2\beta$ is the asymmetry due to CP violation. Particle type ''+'' refers to a $B\to J/\psi K_S^0$ decay and particle type ''−'' refers to a $\overline{B}\to J/\psi K_S^0$ decay.

To allow for an imperfect and (possibly) asymmetric tagging algorithm, the following definitions are used. For those B mesons of (produced) type +, a fraction ϵ_R^+ will be actually tagged +, fraction ϵ_W^+ will be tagged as −, and fraction ϵ_0^+ will not be tagged, i.e. tag 0. Similarly, for those B mesons of (produced) type −, ϵ_R^- will be tagged −, fraction ϵ_W^- will be tagged as +, and fraction ϵ_0^- will be tagged as 0. Because, by definition, $\epsilon_R^+ + \epsilon_W^+ + \epsilon_0^+ = 1$ and $\epsilon_R^- + \epsilon_W^- + \epsilon_0^- = 1$, there are four independent numbers that characterize a general asymmetric tagging algorithm.

We define the efficiencies and dilutions for the general asymmetric tagging algorithm as $\epsilon_+=(\epsilon_R^+ + \epsilon_W^-)/2$, $\epsilon_- = (\epsilon_R^- + \epsilon_W^+)/2$, $\epsilon_0=(\epsilon_0^+ + \epsilon_0^-)/2$ and

$$D_+=\frac{\epsilon_R^+ - \epsilon_W^-}{\epsilon_R^+ + \epsilon_W^-},\quad D_-=\frac{\epsilon_R^- - \epsilon_W^+}{\epsilon_R^- + \epsilon_W^+},\quad D_0=\frac{\epsilon_0^+ - \epsilon_0^-}{\epsilon_0^+ + \epsilon_0^-}.$$

The observed decay rate as a function of time for events tagged as +, − or 0 is given by

$$h_+(t)=\frac{e^{-t/\tau}}{\tau}\epsilon_+[1+\Lambda_{CP}D_+\sin(\Delta m_d t)],$$

$$h_-(t)=\frac{e^{-t/\tau}}{\tau}\epsilon_-[1-\Lambda_{CP}D_-\sin(\Delta m_d t)],$$

and

$$h_0(t)=\frac{e^{-t/\tau}}{\tau}\epsilon_0[1+\Lambda_{CP}D_0\sin(\Delta m_d t)].$$

Note that $\epsilon_+ + \epsilon_- + \epsilon_0 = 1$ and $\epsilon_+ D_+ - \epsilon_- D_- + \epsilon_0 D_0 = 0$, so there are four independent parameters remaining. For example,

$$D_0 = \frac{\epsilon_- D_- - \epsilon_+ D_+}{1 - \epsilon_+ - \epsilon_-}.$$

Combining tags in an event

Tagging information for each event is combined to reduce the uncertainty on the CP asymmetry. The tags are weighted for each event by the dilution of the individual tag algorithms. This procedure must also combine the efficiencies in a similar manner. The algorithm used to combine multiply-tagged events is as follows. We define the tags for two tagging algorithms as q_1 and q_2 (each taking the values -1, 0, and 1), the individual dilutions as D_1 and D_2, and the individual efficiencies as ϵ_{q_1} and ϵ_{q_2}. We then define the dilution-weighted tags $\mathcal{D}_i = q_i D_i$, the product of the tag and the dilution. We calculate the combined dilutions and efficiencies as

$$\mathcal{D}_{q_1 q_2} = \frac{\mathcal{D}_1 + \mathcal{D}_2}{1 + \mathcal{D}_1 \mathcal{D}_2}, \quad \epsilon_{q_1 q_2} = \epsilon_{q_1} \epsilon_{q_2} (1 + \mathcal{D}_1 \mathcal{D}_2),$$

where $\mathcal{D}_{q_1 q_2}$ is the combined dilution-weighted tag, and $\epsilon_{q_1 q_2}$ is the combined efficiency. In this manner, tags in agreement as well as tags in conflict are handled properly: in the cases where the charge of the two tags agree, the effective dilution is increased; in the cases where the two tags disagree, the effective dilution is decreased.

To help understand the expression for combined dilution $\mathcal{D}$, we examine several limiting cases. In the case of a perfect first tagging algorithm, $|\mathcal{D}_1| = 1$, the combined tag always equals the value of the perfect algorithm ($\mathcal{D}_{q_1 q_2} = \mathcal{D}_1$), independently of the second tagging algorithm. For the case where the first tagging algorithm is random, $|\mathcal{D}_1| = 0$, the combined tag always equals the value of second algorithm ($\mathcal{D}_{q_1 q_2} = \mathcal{D}_2$). In the case where the result of first tagging algorithm is equal and opposite to the result of the second tagging algorithm ($\mathcal{D}_1 = -\mathcal{D}_2$), the $\mathcal{D}_{q_1 q_2} = 0$. This is expected when the two tagging algorithms have equal power but give the opposite answer.

To understand the combined efficiency $\epsilon_{q_1 q_2}$, we consider an example. There are nine possible efficiencies for the combined tagging algorithms, $\epsilon_{q_1 q_2}$. The individual efficiencies for perfectly efficient symmetric tagging algorithms have the values $\epsilon_+ = \epsilon_- = 0.5$ and $\epsilon_0 = 0$ ($\epsilon_+ + \epsilon_- + \epsilon_0 = 1$). In this case, five of the nine combined efficiencies are trivially zero. For the case of two perfect tagging algorithms giving the opposite result ($\mathcal{D}_1 = -\mathcal{D}_2$ and $|\mathcal{D}_1| = 1$), then the combined efficiency must be $\epsilon_{q_1 q_2} = 0$, independent of the magnitude of ϵ_{q_1} and ϵ_{q_2}. This is expected because, by definition, perfect tagging algorithms cannot disagree. There are only two remaining nonzero cases to examine for the perfectly efficient tagging algorithm. For the case in which they agree, the combined efficiencies are $\epsilon_{+1,+1} = 0.5$ and $\epsilon_{-1,-1} = 0.5$.

V. THE LIKELIHOOD FUNCTION

An extended log-likelihood method is used to determine the best value for $\sin 2\beta$, a free parameter in the fit. It is helpful to refer to the parameters collectively as a vector $\vec{p}$ with 65 components. The remaining 64 parameters describe other features of the data (signal and background) which need to be determined simultaneously, but have only technical importance.

The main ingredient of the likelihood function is the product $\Pi_i \mathcal{P}_i$ where i runs over all the selected events and $\mathcal{P}_i$ is the probability distribution in the measured quantities: the normalized mass, the flight-time, and the tags (q_1, q_2, q_3). The tags, although discrete variables, are conceptually thought of as analogous to continuous variables, such as the measured mass. The parameters $\vec{p}$ control the shape of the $\mathcal{P}_i$. There is a separate set of parameters for the SVX sample and the non-SVX sample to control the shape of the components of $\mathcal{P}_i$. This is especially important for the parts of the function that specify the distribution of the measured flight-time and mass, but also the distribution of SST tags.

The form for $\mathcal{P}_i$ assumes that all events are of three types: signal, prompt background, and long-lived background. Each possibility is included in $\mathcal{P}_i$. Because the distributions in mass, flight-time, and tag are different for the three types, $\mathcal{P}_i$ contains separate components $\mathcal{P}_S$, $\mathcal{P}_P$, and $\mathcal{P}_L$, which are the overall distributions for signal, prompt background, and long-lived background respectively. Additional parameters—a separate set of parameters for SVX and non-SVX—specify the relative quantities of each event-type. Each of the components $\mathcal{P}_S$, $\mathcal{P}_P$, and $\mathcal{P}_L$ is expressed as the product of a time-function (T_S, T_P, T_L), a mass-function (M_S, M_P, M_L), and a tagging-efficiency-function ($\mathcal{E}_S, \mathcal{E}_P, \mathcal{E}_L$).

The time-function T_S is the probability distribution for the observed-time given the observed tags, and therefore has a dependence on the measured time and its uncertainty, the measured tags and dilutions, and $\sin 2\beta$. The B^0 lifetime τ and mixing parameter Δm_d are constrained at the world averages: $\tau = (1.54 \pm 0.04)$ ps and $\Delta m_d = (0.464 \pm 0.018)\hbar\,\mathrm{ps}^{-1}$ [8]. The T_P function is a simple Gaussian representing the prompt J/ψ background, and depends on the measured time and uncertainty. There are two time-uncertainty scale factors in $\vec{p}$, one for SVX events and one for the non-SVX events, to allow for the possibility that the measured time-uncertainties are different from the true uncertainties by a constant factor. The T_L function has positive and negative exponentials in time to represent positive and negative long-lived background. The positive long-lived background arises primarily from real B decays, while the negative long-lived background is used to describe non-Gaussian tails in the lifetime resolution.

The mass-function M_S is a Gaussian representing the normalized mass, and also includes a mass-uncertainty scale parameter. The mass-functions M_P and M_L are linear in mass and normalized over the $\pm 20\sigma$ mass window.

The tagging-efficiency-function $\mathcal{E}_S$ gives the probability of obtaining the observed combination of tags for a signal

TABLE III. The dilutions determined from the $B^{\pm}\to J/\psi K^{\pm}$ sample and the efficiency ratios determined from the inclusive J/ψ sample are shown. $D_{\rm ave}$ is the average dilution. The SST dilutions utilize additional information as described in the text.

Tag	ϵ_+/ϵ_-	$D_+(\%)$	$D_-(\%)$	$D_{\rm ave}(\%)$
SST_{SVX}	1.031 ± 0.011	16.1 ± 5.1	17.1 ± 5.2	16.6 ± 2.2
$\mathrm{SST}_{\text{non-SVX}}$	1.037 ± 0.010	17.0 ± 5.7	17.8 ± 5.8	17.4 ± 3.6
SLT	0.978 ± 0.047	76.9 ± 19.6	46.4 ± 21.8	62.5 ± 14.6
JETQ	0.977 ± 0.015	20.7 ± 9.3	26.5 ± 8.3	23.5 ± 6.9

event. In addition to the observed tags for the event, it also depends on the individual tagging efficiencies and dilutions. The prompt and long-lived background tagging-efficiency-functions, $\mathcal{E}_{\rm P}$ and $\mathcal{E}_{\rm L}$, give the probability of obtaining the observed combination of tags for prompt and long-lived background events; they depend on individual background tagging efficiencies, but no dilutions are involved because there is no right or wrong sign in the tagging background. For each individual tagging algorithm, the efficiencies and the dilutions (each a component of $\vec{p}$) float and are allowed to be different for $+$ and $-$ tags and the corresponding efficiencies and the dilutions for the tag-0 cases follow by normalization. However, for the signal, there are constraints on the individual tagging efficiencies and dilutions based on the available measurements and their uncertainties.

A. The likelihood function definition

The negative log-likelihood $l(\vec{p})$ is given by

$$l(\vec{p})=N_{\rm S}^{\rm SVX}+N_{\rm B}^{\rm SVX}+N_{\rm S}^{\text{non-SVX}}+N_{\rm B}^{\text{non-SVX}}-\sum_i \ln(\mathcal{P}_i)+\sum_j \frac{1}{2}\left(\frac{f_j(\vec{p})-\langle f_j\rangle}{\sigma_j}\right)^2.$$

The four free parameters $N_{\rm S}^{\rm SVX}$, $N_{\rm B}^{\rm SVX}$, $N_{\rm S}^{\text{non-SVX}}$, and $N_{\rm B}^{\text{non-SVX}}$ refer to the number of signal and background events in the SVX and non-SVX respectively. The summation over j represents a summation over all of the constraints we place on the parameters. The constraints in general connect some function $f_j(\vec{p})$ of the parameters with the corresponding value $\langle f_j\rangle$ and uncertainty σ_j determined by other measurements.

The summation over i above runs over all data events that satisfy our selection criteria; $\mathcal{P}_i$ is the probability for the ith event, and implicitly depends on $\vec{p}$. The function $\mathcal{P}_i$ is given by

$$\mathcal{P}_i=N_{\rm S}\mathcal{P}_{\rm S}+N_{\rm B}[(1-F_{\rm L})\mathcal{P}_{\rm P}+F_{\rm L}\mathcal{P}_{\rm L}].$$

All events are classified as either type SVX or type non-SVX: the $N_{\rm S}$, $N_{\rm B}$, and $F_{\rm L}$ in the expression above are actually parameters $N_{\rm S}^{\rm SVX}$, $N_{\rm B}^{\rm SVX}$, and $F_{\rm L}^{\rm SVX}$ (the long-lived fraction of SVX background) for SVX-type events and $N_{\rm S}^{\text{non-SVX}}$, $N_{\rm B}^{\text{non-SVX}}$, and $F_{\rm L}^{\text{non-SVX}}$ for non-SVX-type events. Although the lifetime resolution for non-SVX events is poor relative to the SVX events, the information is used in the likelihood function.

The functions $\mathcal{P}_{\rm S}$, $\mathcal{P}_{\rm P}$, and $\mathcal{P}_{\rm L}$ are the probabilities for the signal, prompt background, and long-lived backgrounds. They are given by the products of time, mass, and tagging-efficiency functions:

$$\mathcal{P}_{\rm S}=T_{\rm S}M_{\rm S}\mathcal{E}_{\rm S},\quad \mathcal{P}_{\rm P}=T_{\rm P}M_{\rm P}\mathcal{E}_{\rm P},\quad \mathcal{P}_{\rm L}=T_{\rm L}M_{\rm L}\mathcal{E}_{\rm L}.$$

The signal time function is specified by

$$T_{\rm S}=\frac{1}{2}g^*h(t),\quad \sigma=S_t\sigma_t,$$

$$h(t)=\frac{e^{-t/\tau}}{\tau}\epsilon_{q_1q_2}[1+\Lambda_{CP}\mathcal{D}_{q_1q_2}\sin(\Delta m t)],$$

where $g^*h(t)$ represents the convolution of $h(t)$ with a Gaussian of width σ and depends implicitly on the values of the flight-time-uncertainty σ and $\sin 2\beta$. The S_t above is $S_t^{\rm SVX}$ (the SVX lifetime error scale) for SVX events and $S_t^{\text{non-SVX}}$ for non-SVX events. The σ_t is the uncertainty on the flight-time t of the B-candidate, determined independently for each event. The prompt background allows the determination of $S_t^{\rm SVX}$ and $S_t^{\text{non-SVX}}$ using the global fit. Knowledge of the individual tag dilutions is incorporated through the constraints.

The signal mass function is

$$M_{\rm S}=\frac{1}{\sqrt{2\pi}S_m}e^{-0.5(M_B/S_m)^2},$$

where M_B is the normalized mass of the B-candidate and S_m is the B-mass error scale.

In an analogous fashion to $\mathcal{D}$, the combined signal tagging-efficiency function $\mathcal{E}_{\rm S}$, calculated by combining three tags as in Sec. IV A, depends on the eight tagging dilution components (as in Table III) of $\vec{p}$ and the eight individual $+$ and $-$ tagging-efficiency components. The combined efficiency $\mathcal{E}_{\rm S}$ is the efficiency for obtaining the particular combination of tags observed in the event.

The prompt background time and mass functions are

$$T_{\rm P}=\frac{1}{2\sqrt{2\pi}\sigma}e^{-t^2/(2\sigma^2)},$$

$$\sigma = S_t \sigma_t, \quad S_t = S_t^{\mathrm{SVX}} \text{ or } S_t^{\text{non-SVX}},$$

$$M_{\mathrm{P}} = (1 + \varsigma_m^{\mathrm{P}} M_B)/(2W), \quad W = 20,$$

where W represents the normalized-mass window-size ($\pm 20\sigma$), and ς_m^{P} is the mass-slope of the prompt background.

The combined prompt-background tagging-efficiency function $\mathcal{E}_{\mathrm{P}}$ is given by the product of the individual prompt background tagging-efficiencies: $\mathcal{E}_{\mathrm{P}} = \Pi_k \mathcal{E}_{\mathrm{P}}^k$ where k runs over the tags. The individual prompt background tagging-efficiencies are parametrized as

$$\mathcal{E}_{\mathrm{P}}^k = \begin{cases} \epsilon_{\mathrm{P}}^k (1 - A_{\mathrm{P}}^k)/2, & q^k = -1, \\ 1 - \epsilon_{\mathrm{P}}^k, & q^k = 0, \\ \epsilon_{\mathrm{P}}^k (1 + A_{\mathrm{P}}^k)/2, & q^k = 1, \end{cases}$$

where q^k is the tag-result of the kth tagging algorithm, and ϵ_{P}^k and A_{P}^k are components of $\vec{p}$ (specifically $\epsilon_{\mathrm{P}}^{\mathrm{SST_{SVX}}}$, $A_{\mathrm{P}}^{\mathrm{SST_{SVX}}}$, $\epsilon_{\mathrm{P}}^{\mathrm{SST_{non\text{-}SVX}}}$, $A_{\mathrm{P}}^{\mathrm{SST_{non\text{-}SVX}}}$, $\epsilon_{\mathrm{P}}^{\mathrm{JCH}}$, $A_{\mathrm{P}}^{\mathrm{JCH}}$, $\epsilon_{\mathrm{P}}^{\mathrm{SLT}}$, and $A_{\mathrm{P}}^{\mathrm{SLT}}$). The A_{P}^k parameters are the asymmetries of the kth algorithm in tagging the prompt background. The $\mathrm{SST_{SVX}}$ and $\mathrm{SST_{non\text{-}SVX}}$ are mutually exclusive—k always runs over three tags.

The long-lived time function T_{L} is given by

$$T_{\mathrm{L}} = \begin{cases} F_- \dfrac{1}{2\tau_-} e^{t/\tau_-}, & t < 0, \\ (1 - F_-) \dfrac{1}{2\tau_+} e^{-t/\tau_+}, & t \geqslant 0, \end{cases}$$

where F_- is one of F_-^{SVX} and $F_-^{\text{non-SVX}}$, τ_+ is one of τ_+^{SVX} and $\tau_+^{\text{non-SVX}}$, and τ_- is one of τ_-^{SVX} and $\tau_-^{\text{non-SVX}}$.

The long-lived mass and tagging-efficiency functions are

$$M_{\mathrm{L}} = (1 + \varsigma_m^{\mathrm{L}} M_B)/(2W), \quad \mathcal{E}_{\mathrm{L}} = \prod_k \mathcal{E}_{\mathrm{L}}^k,$$

$$\mathcal{E}_{\mathrm{L}}^k = \begin{cases} \epsilon_{\mathrm{L}}^k (1 - A_{\mathrm{L}}^k)/2, & q^k = -1, \\ 1 - \epsilon_{\mathrm{L}}^k, & q^k = 0, \\ \epsilon_{\mathrm{L}}^k (1 + A_{\mathrm{L}}^k)/2, & q^k = 1, \end{cases}$$

where the notation is exactly analogous to the M_{P} and $\mathcal{E}_{\mathrm{P}}$ defined above.

To further illustrate the role of constraint terms in the negative log-likelihood function we highlight the dilution constraints. There are two dilution parameters, D_+ and D_-, per tagging method, the eight parameters in $l(\vec{p})$ representing the tagging dilutions that float in the fit that locates the minimum of $l(\vec{p})$. The probability $\mathcal{P}_i$ of the ith $J/\psi K_S^0$ candidate depends on these parameters through T_{S} and $\mathcal{E}_{\mathrm{S}}$. Each tagging method also has its own calibration information derived from other decay modes. For example, the dilutions are constrained using results from the $J/\psi K^{\pm}$ calibration sample. In addition, the D_+ and D_- dilutions for the SST SVX sample are constrained to the average dilution ($D_{\mathrm{ave}} = 16.6 \pm 2.2\%$) obtained after extrapolating the mixing analysis dilution to lower P_T [10,25]. The available calibration information for each tagging method is represented in $l(\vec{p})$ by constraint terms. These terms cause the function $l(\vec{p})$ to increase as the dilution parameters wander from the values preferred by the calibration. When locating the minimum of $l(\vec{p})$ we are then simultaneously determining $\sin 2\beta$ and the eight dilution parameters, so that the uncertainty on $\sin 2\beta$ from the fit includes contributions from all of the calibration uncertainties.

There are similar constraint terms for the efficiency ratios for each tagging method (ϵ_+/ϵ_-). The efficiency ratios ϵ_+/ϵ_- for each tag algorithm are constrained using the inclusive $B \to J/\psi X$ sample. We fit the J/ψ mass distributions for the number of $+$ and $-$ tags. The ratio of the number of $+$ tags to the number of $-$ tags constrains ϵ_+/ϵ_-. The $B \to J/\psi X$ sample is assumed to have negligible intrinsic CP asymmetry. In addition, the B^0 lifetime τ_{B^0} and mixing parameter Δm_d are free parameters in the fit, and there are terms to constrain each to its world average [8]. The parameter τ_{B^0} is constrained to 1.56 ± 0.04 ps and the parameter Δm_d is constrained to $0.464 \pm 0.018 \hbar$ ps^{-1}. Although constraining Δm_d to the world average is the most natural procedure, we also have the option of determining Δm_d and $\sin 2\beta$ simultaneously from the $J/\psi K_S^0$ data by removing the constraint on Δm_d.

The calibration measurements are summarized in Table III. The efficiency ratios are consistent with expectations. For SST, the ratios are greater than unity due to a higher efficiency for reconstructing tracks with positive charge in the CTC.

B. Fits to toy Monte Carlo data

As a check of the fitting procedure several sets of ~ 1000 toy Monte Carlo data samples were generated, each set generated with a different value of $\sin 2\beta$. The number of events, SVX/non-SVX ratio, signal-to-background ratios, tagging efficiencies and dilutions, mass uncertainty and its scale factor, background lifetimes, time uncertainties and scale factors, and other kinematic features of the generation procedure were all tuned to be similar to the composition of the data sample.

The left plot in Fig. 5 shows the distribution of the appropriate uncertainty (allowing for asymmetric errors [29]) on $\sin 2\beta$ returned from the Monte Carlo fits with generated $\sin 2\beta = 0.5$. The typical value of the uncertainty on $\sin 2\beta$ returned from these fits is ~ 0.44, though there is a long tail extending out to ~ 0.7. The width of the distribution is determined by Poisson fluctuations in the number of Monte Carlo events that are tagged. The right plot in Fig. 5 shows $[\sin 2\beta(\mathrm{fit}) - 0.5]/\sigma$, where σ is the appropriate $+$ or $-$ uncertainty on $\sin 2\beta$.

The results from this and other samples generated at different values of $\sin 2\beta$ support that the fitting procedure provides an unbiased estimate of the value of $\sin 2\beta$ of the parent distribution. The distribution of the difference between

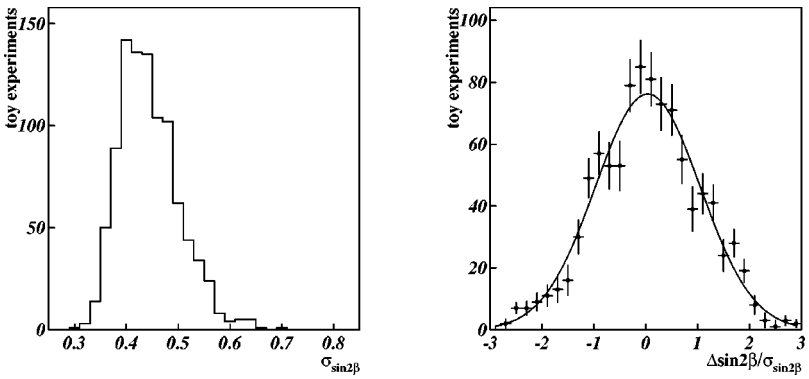

FIG. 5. Left: Distribution of $\sigma_{\sin 2\beta}$ from fits to multiple Monte Carlo datasets generated with $\sin 2\beta = 0.5$. Right: Distribution of normalized $\sin 2\beta$ deviations, i.e. $(\text{fit-}\sin 2\beta - 0.5)/\sigma_{\sin 2\beta}$, and a Gaussian fit to that distribution. The mean of the Gaussian fit is 0.038 ± 0.033 and the width is 1.01 ± 0.03, consistent with expectation.

the fit-$\sin 2\beta$ and the true $\sin 2\beta$ of the parent distribution is well approximated by a Gaussian and the fit-uncertainty on $\sin 2\beta$ provides a good estimate of the σ of that Gaussian.

C. Systematic uncertainties

Systematic uncertainties on the measurement of $\sin 2\beta$ due to flavor tagging, the B lifetime and Δm_d are included as constraints in the fit. We evaluated the systematic uncertainties due to the uncertainty in the B^0 mass, trigger bias and K_L^0 regeneration.

The systematic uncertainty arising from the B mass is studied using 1000 simulated experiments. The data were generated at the nominal B mass and three full likelihood fits were performed on each experiment. One fit was performed using the normalized mass calculated with the nominal B mass and two additional fits were performed using B masses shifted by ± 1 MeV/c^2. The shifts observed in $\sin 2\beta$ from fits to the simulated experiments are consistent with a random distribution centered on zero with an rms of 0.019. The change in the observed rms spread of $\sin 2\beta$ is <0.019 when combined in quadrature. We also fit the data with the B mass shifted by 1 MeV/c^2 and found the value of $\sin 2\beta$ changed by 0.013, which consistent with the simulation results. We conclude the additional uncertainty on $\sin 2\beta$ due to the uncertainty on the B mass is <0.019 and is negligible.

The data are assumed to be a 50:50 mix of $B^0/\overline{B}^0$. A possible charge bias arising from the trigger is considered. Events that are triggered on the two muons from the J/ψ decay do not contribute to the charge bias. The remaining 30% contain some events in which the trigger was from one of the J/ψ muons and the other lepton candidate was from the opposite side B. The magnitude of the charge bias in the trigger has been measured to be $<1\%$ at a threshold of $P_T = 2$ GeV/c and is consistent with zero for $P_T > 3$ GeV/c, rendering this uncertainty negligible.

Possible contamination of our data from K_L^0 regeneration from the material in the inner detector has been considered. Reconstruction of the K_L^0 as a K_S^0 causes the event to be entered with the incorrect sign in the asymmetry. This effect shifts $\sin 2\beta$ by less than 0.003, which is neglected. The results of the systematic studies are shown in Table IV.

We have evaluated the contribution to the sample from $B^0 \to J/\psi K^*$, with $K^* \to K_S^0 \pi^0$ and the π^0 not reconstructed and find it to be a negligible contribution. The same is true with $\Lambda_B \to J/\psi \Lambda$ and $\Lambda \to p\pi^-$ and the Λ reconstructed as $K_S^0 \to \pi^+\pi^-$; $B_s \to J/\psi\phi$, $\phi \to K_S^0 K_L^0$; and $B_s \to J/\psi K_S^0$.

Many checks of the data and analysis have been performed to increase our confidence in the result. In order to check the sensitivity of the result to the dilutions, we imposed alternative JETQ and SLT dilution parameters taken from our various mixing analyses that use the inclusive lepton sample [27]. We observe the expected shift in the value of $\sin 2\beta$ and small changes in the uncertainty. The signal sample selection criteria have been varied, and other than a sensitivity to the SST tag track P_T threshold, as discussed in Ref. [10], we find no unexpected sensitivity in the result.

D. Final result

The maximum likelihood function fitting procedure returns a stable value for $\sin 2\beta$ and the uncertainties are approximately Gaussian. Even though asymmetric dilutions are permitted in the fit, no significant asymmetry is observed. Furthermore, the background asymmetries are consistent with zero.

Using the entire data set and three tagging algorithms, we find

$$\sin 2\beta = 0.79^{+0.41}_{-0.44}.$$

The asymmetry is shown in Fig. 6 for the SVX and non-SVX events separately. The asymmetry for the SVX events is displayed as a function of lifetime, while the asymmetry for the non-SVX events is shown in a single, time-integrated

TABLE IV. Systematic uncertainties in the measurement of $\sin 2\beta$. The items labeled ''in fit'' are parameters that are allowed to float in the fit but are constrained by their measured uncertainties. The uncertainty returned from the likelihood fit includes the contributions from these sources.

Parameter	$\delta \sin 2\beta$	In fit
Dilution and efficiency	0.16	Yes
Δm_d	Negligible	Yes
τ_{B^0}	Negligible	Yes
m_B	Negligible	Yes
Trigger bias	Negligible	No
K_L^0 regeneration	Negligible	No

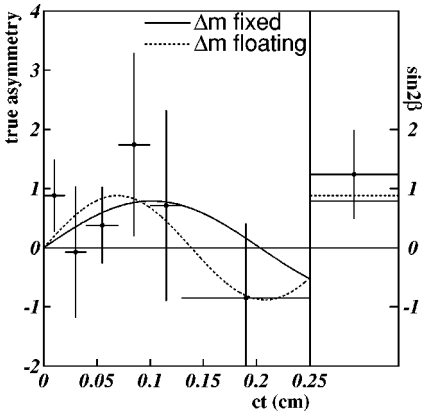

FIG. 6. The true asymmetry ($\sin 2\beta \sin \Delta m_d t$) as a function of lifetime for $B \to J/\psi K_S^0$ events. The data points are sideband-subtracted and have been combined according to the effective dilution for single and double-tags. The non-SVX events are shown on the right.

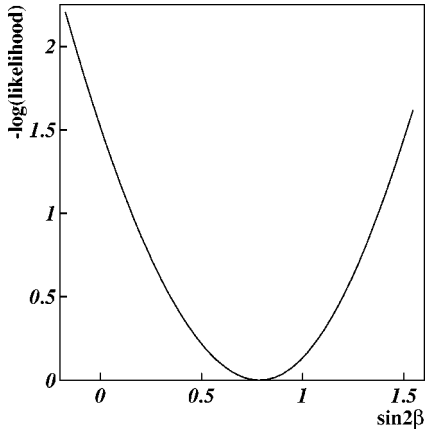

FIG. 7. A scan of the log-likelihood function. The value of $\sin 2\beta$ is scanned, and at each step, the function is minimized.

bin since the decay length information is of low resolution. Although plotted as a time-integrated point, lifetime information for the non-SVX events is utilized in the maximum likelihood function. The positive asymmetry preferred by the fit can be seen. The curves displayed in the plot are the results from the full maximum likelihood fit using all data. In order to display the data, we have combined the effective dilution for single and double-tag events after having subtracted the background. The full maximum likelihood fit uses the SVX and non-SVX samples and treats properly the decay length, dilution and uncertainty for each event.

The uncertainty can be divided into statistical and systematic terms:

$$\sin 2\beta = 0.79 \pm 0.39(\text{stat}) \pm 0.16(\text{syst}).$$

The systematic term predominantly reflects the uncertainty in the result due to the uncertainty in the dilution parameters. Although the dilution parameters are not precisely determined, due to the limited statistics of the $B^{\pm} \to J/\psi K^{\pm}$ calibration sample, this uncertainty term does not dominate the overall uncertainty on $\sin 2\beta$. Furthermore, the uncertainty on $\sin 2\beta$ will not be dominated by the uncertainty on the dilution parameters in future runs because the uncertainty scales inversely with increasing statistics of the calibration samples.

It is of interest to determine the quantitative statistical significance of whether this result supports $\sin 2\beta > 0.0$ and hence provides an indication of CP symmetry violation in the b quark system. A scan through the likelihood function as $\sin 2\beta$ is varied is shown in Fig. 7 and demonstrates that the uncertainties follow Gaussian statistics. Using the Feldman-Cousins frequentist approach [30], we calculate a confidence interval of $0.0 < \sin 2\beta < 1$ at 93%. An alternative approach is the Bayesian method, where a flat prior distribution in $\sin 2\beta$ is assumed and a probability that $\sin 2\beta > 0.0$ of 95% is calculated. Finally, if the true value of $\sin 2\beta$ is zero, and the measurement uncertainty is 0.44 (Gaussian uncertainty), the probability of obtaining $\sin 2\beta > 0.79$ is 3.6%. This value is obtained by simply integrating the Gaussian distribution from 0.79 to ∞. The toy Monte Carlo calculation is in good agreement with the calculated probability.

It is possible to remove the constraint that ties Δm_d to the world average value and to fit for $\sin 2\beta$ and Δm_d simultaneously. In this case the result is $\sin 2\beta = 0.88^{+0.41}_{-0.44}$ and $\Delta m_d = 0.68 \pm 0.17 \hbar$ ps^{-1}. The value of Δm_d from the fit agrees with the world value at the level of $\sim 1.2\sigma$. This agreement increases our confidence in the main result. Figure 8 shows the 1σ ''error ellipse'' contour in $\sin 2\beta$-Δm_d parameter space for the fit when both parameters float freely, and for comparison the nominal $\sin 2\beta$ result with the world average Δm_d and uncertainty. From the roughly circular shape of the contour, the Δm_d and $\sin 2\beta$ parameters are largely uncorrelated in the fit.

A time-integrated measurement to check the final result was performed. This simplified analysis does not use the time dependence of the asymmetry and ignores the small tagging asymmetry corrections applied in the full maximum likelihood fit. Each event falls into one of 12 classifications depending upon the type of flavor tags available for that event. Each event can be associated with only one class of tag combination. The effective tagging efficiency for the entire sample, ϵD^2, is $(6.3 \pm 1.7)\%$. A value of $\sin 2\beta$ for each class is calculated and a weighted average from the 12 classes is determined. Ignoring correlations in the dilution, $\sin 2\beta = 0.71 \pm 0.63$. This value is consistent with the final result and demonstrates the improvement in the uncertainty of $\sin 2\beta$ provided by the full maximum likelihood procedure. This improvement agrees well with improvements observed using the toy Monte Carlo calculation.

Table V summarizes fit results for various tag-dataset combinations. The three tagging algorithms contribute roughly equally to the precision of the $\sin 2\beta$ measurement. Although the SVX and non-SVX sample sizes are approximately equal, the SVX events contribute more significantly to the final result. The main reasons for this are that the precision lifetime information from the SVX allows a better determination of where the decay takes place along the oscillation curve and the better signal-to-background level from eliminating the prompt background.

The row in Table V labeled SVX SST is the result obtained when this analysis restricts the data set to the SVX sample and uses only the SST algorithm. This procedure essentially repeats the published CDF $\sin 2\beta$ analysis that obtained $\sin 2\beta = 1.8 \pm 1.1(\text{stat}) \pm 0.3(\text{syst})$. The small difference is due to sample selection.

VI. MIXING IN THE $B \to J/\psi K^*$ SAMPLE AS A CHECK

A control sample of $B^0 \to J/\psi K^*(892)^0$ decays, where $K^*(892)^0 \to K^\pm \pi^\mp$, can be analyzed for the presence of an oscillation due to mixing (Δm_d is well measured) in order to verify the tag algorithms and likelihood fitting procedure. The three flavor tagging algorithms are used to determine the neutral B flavor at the time of production and the dilution parameters are constrained using the same values as in the $B \to J/\psi K_S^0$ analysis. The charge of the kaon is used to differentiate the B^0 from $\overline{B^0}$ at the time of decay. After correcting for tagging dilutions, the amplitude of the oscillation still differs from unity due to the probability that the $K^+\pi^-$ is reconstructed as $K^-\pi^+$, which occurs about $P_K = 5\%$ of the time due to the wide K^* resonance.

The $J/\psi K_S^0$-$J/\psi K^*(892)^0$ analogy is, however, not perfect. In order to achieve similar signal-to-background ratios, the selection criteria for the $B \to J/\psi K^*(892)^0$ are more severe, which changes the kinematic properties of one sample with respect to the other. The largest backgrounds for both decay modes are at short decay distances and they decrease as the flight path increases. This works to our advantage in the CP analysis but reduces the sensitivity of the mixing analysis. In particular, due to the different oscillation phase in the CP analysis versus this mixing analysis [$\sin(\Delta m_d t) \to \cos(\Delta m_d t)$], the smallest signal-to-background ratio occurs at the peak of the mixing amplitude for $B \to J/\psi K^*(892)^0$ data set, where as a very favorable signal-to-background ratio occurs at the peak of the $B \to J/\psi K_S^0$ oscillation. In both the $J/\psi K_S^0$ and $J/\psi K^*(892)^0$ modes, 75–80 % of the background is prompt, i.e. consistent with having zero lifetime.

The sample is constructed using similar criteria to that used to reconstruct the $B \to J/\psi K$ decay modes in this paper. The J/ψ selection for this decay mode is the same as the $J/\psi K_S^0$ analysis. Pion and kaon tracks are required to have $P_T > 500$ MeV/c. The reconstructed $K^*(892)^0$ candidates are required to have an invariant mass within 80 MeV/c^2 of the world average of 896.10 ± 0.28 MeV/c^2 [8] $K^*(892)^0$ mass. The K^* candidate must have $P_T > 3$ GeV/c. The four-track fit for $J/\psi K^*$ is the same as the fit for $J/\psi K_S^0$, except the four tracks are required to meet at a common point and the K^* mass is not constrained. If a candidate event has two tracks that satisfy two $K^*(892)^0$ combinations ($K^+\pi^-/K^-\pi^+$) then the combination with a $K\pi$ mass closest to the mean $K^*(892)^0$ mass is chosen. Finally, if multiple K^* candidates are found in an event, the $K^*(892)^0$ candidate chosen is the one that gives the best four-track fit. All four charged tracks (μ, μ, K, π) must originate from a common vertex and a $P_T(B) > 4.5$ GeV/c is required. A total signal sample of 226 ± 24 events where both muon candidates have precision lifetime information and 231 ± 28 events where $\leqslant 1$ muon candidate has precision lifetime information are found.

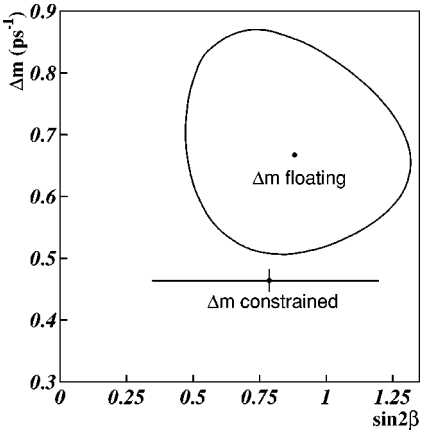

FIG. 8. The 1σ (39%) $\sin 2\beta$-Δm_d contour from a fit with Δm_d constrained only by the $B \to J/\psi K_S^0$ data. Also shown is the nominal fit with $\Delta m_d = (0.464 \pm 0.018)\hbar$ ps^{-1} [8].

The maximum likelihood fit to the $J/\psi K^*(892)^0$ data is implemented in the same way as previously described for $J/\psi K_S^0$ except for the time-function T_S in which $h(t)$ is replaced by

$$h(t) = \frac{e^{-t/\tau}}{\tau} \epsilon_{q_1 q_2} [1 + \mathcal{D}_K \mathcal{D}_{q_1 q_2} \cos(\Delta m_d t)].$$

Here $\mathcal{D}_K = q_K D_K$, where q_K is the charge of the $K^\pm$ from the decay of the $K^*(892)^0$, and D_K is the dilution arising from the inability to correctly distinguish the charged kaon

TABLE V. Fit $\sin 2\beta$ results for the three tagging algorithms. The combined χ^2 for the SST, JETQ, and SLT tagging algorithms is 4.63 for 2 degrees of freedom, giving a probability of ~10%.

Data	Tag(s)	$\sin 2\beta$	+ error	− error
All	all	0.79	0.41	0.44
	SST	2.03	0.84	0.77
	JETQ	−0.31	0.81	0.85
	SLT	0.52	0.61	0.75
SVX	All	0.54	0.52	0.57
	SST	1.77	1.04	1.01
Non-SVX	All	1.24	0.75	0.70

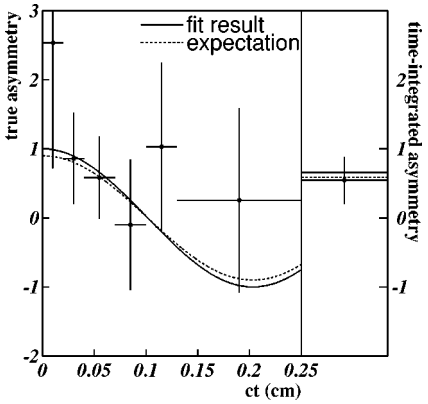

FIG. 9. The true asymmetry ($D_K \cos \Delta m_d t$) as a function of lifetime for $B^0 \to J/\psi K^*(892)^0$ events. The data points are sideband-subtracted and have been combined according to the effective dilution for single and double-tags. The time-integrated asymmetry for non-SVX events is shown on the right. The solid curve represents the maximum likelihood fit in which Δm_d is fixed and the dashed curve is the expectation when we also fix D_K.

from the charged pion in the $K^*(892)^0$ decay. The dilution D_K is the free parameter in this fit and is analogous to $\sin 2\beta$ in the $J/\psi K_S^0$ fit, the parameters in each case representing the amplitude of an oscillation. The amplitude is expected to be $D_K = 1 - 2P_K = 0.9^{+0.1}_{-0.2}$ where D_K is the dilution factor coming from incorrect K-π assignment [25].

When Δm_d is fixed to the world average, we measure $D_K = 1.00 \pm 0.37$, which is consistent with expectation. When Δm_d is allowed to float, we measure: $D_K = 0.96 \pm 0.38$ and $\Delta m_d = 0.40 \pm 0.18 \hbar$ ps^{-1}, which is consistent with the world average $\Delta m_d = (0.464 \pm 0.018)\hbar$ ps^{-1} [8]. The results of the fits are shown in Fig. 9. Although the statistics are not sufficient for a precise measurement of Δm_d, this check on an independent sample of events is entirely consistent with our expectation.

VII. CONCLUSION

We have presented a measurement of $\sin 2\beta$ using ~ 400 $B \to J/\psi K_S^0$ events reconstructed with the CDF detector. We find

$$\sin 2\beta = 0.79^{+0.41}_{-0.44}(\text{stat} + \text{syst})$$

with the uncertainty dominated by the statistical contribution.

We have calculated the statistical significance of whether this result supports $\sin 2\beta > 0.0$ and hence provides indication for CP symmetry violation in the b quark system. Using the Feldman-Cousins [30] method, a 93% confidence interval of $0.0 < \sin 2\beta < 1.00$ is found. Alternative methods yield similar limits. This measurement is the best direct indication that CP invariance is violated in the b quark system and is consistent with the standard model expectation of a large positive value of $\sin 2\beta$ [14–17]. The sign of our result supports the favored positive signs for B_B and B_K. With an anticipated luminosity of 2 fb^{-1} in run II, we expect, based on a simple extrapolation of this measurement, an uncertainty on $\sin 2\beta$ of ~ 0.08. Detector upgrades in progress should further reduce this uncertainty.

ACKNOWLEDGMENTS

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation, the Italian Istituto Nazionale di Fisica Nucleare, the Ministry of Education, Science and Culture of Japan, the Natural Sciences and Engineering Research Council of Canada, the National Science Council of the Republic of China, the Swiss National Science Foundation, the A. P. Sloan Foundation, and the Bundesministerium für Bildung und Forschung, Germany.

[1] J. H. Christenson *et al.*, Phys. Rev. Lett. **13**, 138 (1964).
[2] A. B. Carter and A. I. Sanda, Phys. Rev. Lett. **45**, 952 (1980); Phys. Rev. D **23**, 1567 (1981).
[3] I. I. Bigi and A. I. Sanda, Nucl. Phys. **B193**, 85 (1981); **B281**, 41 (1987).
[4] I. I. Bigi, V. A. Khoze, N. G. Uraltsev, and A. I. Sanda, in *CP Violation*, edited by C. Jarlskog (World Scientific, Singapore, 1989), p. 175.
[5] I. Dunietz and J. L. Rosner, Phys. Rev. D **34**, 1404 (1986).
[6] For a review of CP violation in B decays, Y. Nir and H. R. Quinn, Annu. Rev. Nucl. Part. Sci. **42**, 221 (1992).
[7] N. G. Deshpande, X. G. He, and S. Oh, Z. Phys. C **74**, 359 (1997).
[8] Particle Data Group, C. Caso *et al.*, Eur. Phys. J. C. **3**, 1 (1998).
[9] K. Ackerstaff *et al.*, Eur. Phys. J. C **5**, 379 (1998).
[10] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **81**, 5513 (1998); K. Kelley, Ph.D. thesis, Massachusetts Institute of Technology, 1999.
[11] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963); M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
[12] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).
[13] L.-L. Chau and W.-Y. Keung, Phys. Rev. Lett. **53**, 1802 (1984); C. Jarlskog and R. Stora, Phys. Lett. B **208**, 268 (1988); J. D. Bjorken, Phys. Rev. D **39**, 1396 (1989).
[14] S. Herrlich and U. Nierste, Phys. Rev. D **52**, 6505 (1995).
[15] A. Ali and D. London, Nucl. Phys. B (Proc. Suppl.) **54A**, 297 (1997).
[16] P. Paganini *et al.*, Phys. Scr. **58**, 556 (1998).
[17] S. Mele, Phys. Rev. D **59**, 113011 (1999).
[18] Y. Grossman, B. Kayser, and Y. Nir, Phys. Lett. B **415** 90 (1997); I. I. Bigi and A. I. Sanda, Phys. Rev. D **60**, 033001 (1999).
[19] CDF Collaboration, F. Abe *et al.*, Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988).
[20] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **74**, 2626 (1995); see also Phys. Rev. D **50**, 2966 (1994).

[21] D. Amidei *et al.*, Nucl. Instrum. Methods Phys. Res. A **350**, 73 (1994); P. Azzi *et al.*, *ibid.* **360**, 137 (1995).
[22] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **76**, 2015 (1996).
[23] CDF Collaboration, F. Abe *et al.*, Phys. Rev. D **54**, 6596 (1996).
[24] CDF Collaboration, F. Abe *et al.*, Phys. Rev. D **57**, 5382 (1998).
[25] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **80**, 2057 (1998); Phys. Rev. D **59**, 032001 (1999); P. Maksimović, Ph.D. thesis, Massachusetts Institute of Technology, 1998.
[26] M. Gronau, A. Nippe, and J. L. Rosner, Phys. Rev. D **47**, 1988 (1993); M. Gronau and J. L. Rosner, *ibid.* **49**, 254 (1994).
[27] CDF Collaboration, F. Abe *et al.*, Phys. Rev. D **60**, 072003 (1999); O. Long, Ph.D. thesis, University of Pennsylvania, 1998; M. Peters, Ph.D. thesis, University of California, Berkeley, 1998.
[28] W. Bartel *et al.*, Z. Phys. C **33**, 23 (1986).
[29] F. James, *MINUIT—Function Minimization and Error Analysis*, Version 94.1, CERN Program Library Long Writeup D506, CERN, Geneva, Switzerland, 1994.
[30] G. J. Feldman and R. D. Cousins, Phys. Rev. D **57**, 3873 (1998).