FERMILAB-PUB-11-646-E

Measurement of the CP-Violating Phase $\beta_s^{J/\psi\phi}$ in $B_s^0 \to J/\psi\phi$ Decays with the CDF II Detector

T. Aaltonen,[21] B. Álvarez González[z],[9] S. Amerio,[40] D. Amidei,[32] A. Anastassov[x],[15] A. Annovi,[17] J. Antos,[12] G. Apollinari,[15] J.A. Appel,[15] T. Arisawa,[54] A. Artikov,[13] J. Asaadi,[49] W. Ashmanskas,[15] B. Auerbach,[57] A. Aurisano,[49] F. Azfar,[39] W. Badgett,[15] T. Bae,[25] A. Barbaro-Galtieri,[26] V.E. Barnes,[44] B.A. Barnett,[23] P. Barria[hh],[42] P. Bartos,[12] M. Bauce[ff],[40] F. Bedeschi,[42] S. Behari,[23] G. Bellettini[gg],[42] J. Bellinger,[56] D. Benjamin,[14] A. Beretvas,[15] A. Bhatti,[46] D. Bisello[ff],[40] I. Bizjak,[28] K.R. Bland,[5] B. Blumenfeld,[23] A. Bocci,[14] A. Bodek,[45] D. Bortoletto,[44] J. Boudreau,[43] N. Bousson[mm],[43] A. Boveia,[11] L. Brigliadori[ee],[6] C. Bromberg,[33] E. Brucken,[21] J. Budagov,[13] H.S. Budd,[45] K. Burkett,[15] G. Busetto[ff],[40] P. Bussey,[19] A. Buzatu,[31] A. Calamba,[10] C. Calancha,[29] S. Camarda,[4] M. Campanelli,[28] M. Campbell,[32] F. Canelli[11],[15] B. Carls,[22] D. Carlsmith,[56] R. Carosi,[42] S. Carrillo[m],[16] S. Carron,[15] B. Casal[k],[9] M. Casarsa,[50] A. Castro[ee],[6] P. Catastini,[20] D. Cauz,[50] V. Cavaliere,[22] M. Cavalli-Sforza,[4] A. Cerri[f],[26] L. Cerrito[s],[28] Y.C. Chen,[1] M. Chertok,[7] G. Chiarelli,[42] G. Chlachidze,[15] F. Chlebana,[15] K. Cho,[25] D. Chokheli,[13] W.H. Chung,[56] Y.S. Chung,[45] M.A. Ciocci[hh],[42] A. Clark,[18] C. Clarke,[55] G. Compostella[ff],[40] M.E. Convery,[15] J. Conway,[7] M.Corbo,[15] M. Cordelli,[17] C.A. Cox,[7] D.J. Cox,[7] F. Crescioli[gg],[42] J. Cuevas[z],[9] R. Culbertson,[15] D. Dagenhart,[15] N. d'Ascenzo[w],[15] M. Datta,[15] P. de Barbaro,[45] M. Dell'Orso[gg],[42] L. Demortier,[46] M. Deninno,[6] F. Devoto,[21] M. d'Errico[ff],[40] A. Di Canto[gg],[42] B. Di Ruzza,[15] J.R. Dittmann,[5] M. D'Onofrio,[27] S. Donati[gg],[42] P. Dong,[15] M. Dorigo,[50] T. Dorigo,[40] K. Ebina,[54] A. Elagin,[49] A. Eppig,[32] R. Erbacher,[7] S. Errede,[22] N. Ershaidat[dd],[15] R. Eusebi,[49] S. Farrington,[39] M. Feindt,[24] J.P. Fernandez,[29] R. Field,[16] G. Flanagan[u],[15] R. Forrest,[7] M.J. Frank,[5] M. Franklin,[20] J.C. Freeman,[15] Y. Funakoshi,[54] I. Furic,[16] M. Gallinaro,[46] J.E. Garcia,[18] A.F. Garfinkel,[44] P. Garosi[hh],[42] H. Gerberich,[22] E. Gerchtein,[15] S. Giagu,[47] V. Giakoumopoulou,[3] P. Giannetti,[42] K. Gibson,[43] C.M. Ginsburg,[15] N. Giokaris,[3] P. Giromini,[17] G. Giurgiu,[23] V. Glagolev,[13] D. Glenzinski,[15] M. Gold,[35] D. Goldin,[49] N. Goldschmidt,[16] A. Golossanov,[15] G. Gomez,[9] G. Gomez-Ceballos,[30] M. Goncharov,[30] O. González,[29] I. Gorelov,[35] A.T. Goshaw,[14] K. Goulianos,[46] S. Grinstein,[4] C. Grosso-Pilcher,[11] R.C. Group[53],[15] J. Guimaraes da Costa,[20] S.R. Hahn,[15] E. Halkiadakis,[48] A. Hamaguchi,[38] J.Y. Han,[45] F. Happacher,[17] K. Hara,[51] D. Hare,[48] M. Hare,[52] R.F. Harr,[55] K. Hatakeyama,[5] C. Hays,[39] M. Heck,[24] J. Heinrich,[41] M. Herndon,[56] S. Hewamanage,[5] A. Hocker,[15] W. Hopkins[g],[15] D. Horn,[24] S. Hou,[1] R.E. Hughes,[36] M. Hurwitz,[11] U. Husemann,[57] N. Hussain,[31] M. Hussein,[33] J. Huston,[33] G. Introzzi,[42] M. Iori[jj],[47] A. Ivanov[p],[7] E. James,[15] D. Jang,[10] B. Jayatilaka,[14] E.J. Jeon,[25] S. Jindariani,[15] M. Jones,[44] K.K. Joo,[25] S.Y. Jun,[10] T.R. Junk,[15] T. Kamon[25],[49] P.E. Karchin,[55] A. Kasmi,[5] Y. Kato[o],[38] W. Ketchum,[11] J. Keung,[41] V. Khotilovich,[49] B. Kilminster,[15] D.H. Kim,[25] H.S. Kim,[25] J.E. Kim,[25] M.J. Kim,[17] S.B. Kim,[25] S.H. Kim,[51] Y.K. Kim,[11] Y.J. Kim,[25] N. Kimura,[54] M. Kirby,[15] S. Klimenko,[16] K. Knoepfel,[15] K. Kondo,[54,*] D.J. Kong,[25] J. Konigsberg,[16] A.V. Kotwal,[14] M. Kreps[ll],[24] J. Kroll,[41] D. Krop,[11] M. Kruse,[14] V. Krutelyov[c],[49] T. Kuhr,[24] M. Kurata,[51] S. Kwang,[11] A.T. Laasanen,[44] S. Lami,[42] S. Lammel,[15] M. Lancaster,[28] R.L. Lander,[7] K. Lannon[y],[36] A. Lath,[48] G. Latino[hh],[42] T. LeCompte,[2] E. Lee,[49] H.S. Lee[q],[11] J.S. Lee,[25] S.W. Lee[bb],[49] S. Leo[gg],[42] S. Leone,[42] J.D. Lewis,[15] A. Limosani[t],[14] C.-J. Lin,[26] M. Lindgren,[15] E. Lipeles,[41] A. Lister,[18] D.O. Litvintsev,[15] C. Liu,[43] H. Liu,[53] Q. Liu,[44] T. Liu,[15] S. Lockwitz,[57] A. Loginov,[57] D. Lucchesi[ff],[40] J. Lueck,[24] P. Lujan,[26] P. Lukens,[15] G. Lungu,[46] J. Lys,[26] R. Lysak[e],[12] R. Madrak,[15] K. Maeshima,[15] P. Maestro[hh],[42] S. Malik,[46] G. Manca[a],[27] A. Manousakis-Katsikakis,[3] F. Margaroli,[47] C. Marino,[24] M. Martínez,[4] P. Mastrandrea,[47] K. Matera,[22] M.E. Mattson,[55] A. Mazzacane,[15] P. Mazzanti,[6] K.S. McFarland,[45] P. McIntyre,[49] R. McNulty[j],[27] A. Mehta,[27] P. Mehtala,[21] C. Mesropian,[46] T. Miao,[15] D. Mietlicki,[32] A. Mitra,[1] H. Miyake,[51] S. Moed,[15] N. Moggi,[6] M.N. Mondragon[m],[15] C.S. Moon,[25] R. Moore,[15] M.J. Morello[ii],[42] J. Morlock,[24] P. Movilla Fernandez,[15] A. Mukherjee,[15] Th. Muller,[24] P. Murat,[15] M. Mussini[ee],[6] J. Nachtman[n],[15] Y. Nagai,[51] J. Naganoma,[54] I. Nakano,[37] A. Napier,[52] J. Nett,[49] C. Neu,[53] M.S. Neubauer,[22] J. Nielsen[d],[26] L. Nodulman,[2] S.Y. Noh,[25] O. Norniella,[22] L. Oakes,[39] S.H. Oh,[14] Y.D. Oh,[25] I. Oksuzian,[53] T. Okusawa,[38] R. Orava,[21] L. Ortolan,[4] S. Pagan Griso[ff],[40] C. Pagliarone,[50] E. Palencia[f],[9] V. Papadimitriou,[15] A.A. Paramonov,[2] J. Patrick,[15] G. Pauletta[kk],[50] M. Paulini,[10] C. Paus,[30] D.E. Pellett,[7] A. Penzo,[50] T.J. Phillips,[14] G. Piacentino,[42] E. Pianori,[41] J. Pilot,[36] K. Pitts,[22] C. Plager,[8] L. Pondrom,[56] S. Poprocki[g],[15] K. Potamianos,[44] F. Prokoshin[cc],[13] A. Pranko,[26] F. Ptohos[h],[17] E. Pueschel,[10] G. Punzi[gg],[42] A. Rahaman,[43] V. Ramakrishnan,[56] N. Ranjan,[44] I. Redondo,[29] P. Renton,[39] M. Rescigno,[47] T. Riddick,[28] F. Rimondi[ee],[6] L. Ristori[42],[15] A. Robson,[19] T. Rodrigo,[9] T. Rodriguez,[41] E. Rogers,[22] S. Rolli[i],[52] R. Roser,[15] F. Ruffini[hh],[42] A. Ruiz,[9] J. Russ,[10] V. Rusu,[15] A. Safonov,[49] W.K. Sakumoto,[45] Y. Sakurai,[54] L. Santi[kk],[50] K. Sato,[51] V. Saveliev[w],[15] A. Savoy-Navarro[aa],[15] P. Schlabach,[15] A. Schmidt,[24] E.E. Schmidt,[15] T. Schwarz,[15] L. Scodellaro,[9] A. Scribano[hh],[42] F. Scuri,[42] S. Seidel,[35] Y. Seiya,[38] A. Semenov,[13] F. Sforza[hh],[42] S.Z. Shalhout,[7]

Operated by Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy.

T. Shears,[27] P.F. Shepard,[43] M. Shimojima[v],[51] M. Shochet,[11] I. Shreyber-Tecker,[34] A. Simonenko,[13] P. Sinervo,[31] K. Sliwa,[52] J.R. Smith,[7] F.D. Snider,[15] A. Soha,[15] V. Sorin,[4] H. Song,[43] P. Squillacioti[hh],[42] M. Stancari,[15] R. St. Denis,[19] B. Stelzer,[31] O. Stelzer-Chilton,[31] D. Stentz[x],[15] J. Strologas,[35] G.L. Strycker,[32] Y. Sudo,[51] A. Sukhanov,[15] I. Suslov,[13] K. Takemasa,[51] Y. Takeuchi,[51] J. Tang,[11] M. Tecchio,[32] P.K. Teng,[1] J. Thom[g],[15] J. Thome,[10] G.A. Thompson,[22] E. Thomson,[41] D. Toback,[49] S. Tokar,[12] K. Tollefson,[33] T. Tomura,[51] D. Tonelli,[15] S. Torre,[17] D. Torretta,[15] P. Totaro,[40] M. Trovato[ii],[42] F. Ukegawa,[51] S. Uozumi,[25] A. Varganov,[32] F. Vázquez[m],[16] G. Velev,[15] C. Vellidis,[15] M. Vidal,[44] I. Vila,[9] R. Vilar,[9] J. Vizán,[9] M. Vogel,[35] G. Volpi,[17] P. Wagner,[41] R.L. Wagner,[15] T. Wakisaka,[38] R. Wallny,[8] S.M. Wang,[1] A. Warburton,[31] D. Waters,[28] W.C. Wester III,[15] D. Whiteson[b],[41] A.B. Wicklund,[2] E. Wicklund,[15] S. Wilbur,[11] F. Wick,[24] H.H. Williams,[41] J.S. Wilson,[36] P. Wilson,[15] B.L. Winer,[36] P. Wittich[g],[15] S. Wolbers,[15] H. Wolfe,[36] T. Wright,[32] X. Wu,[18] Z. Wu,[5] K. Yamamoto,[38] D. Yamato,[38] T. Yang,[15] U.K. Yang[r],[11] Y.C. Yang,[25] W.-M. Yao,[26] G.P. Yeh,[15] K. Yi[n],[15] J. Yoh,[15] K. Yorita,[54] T. Yoshida[l],[38] G.B. Yu,[14] I. Yu,[25] S.S. Yu,[15] J.C. Yun,[15] A. Zanetti,[50] Y. Zeng,[14] C. Zhou,[14] and S. Zucchelli[ee]6

(CDF Collaboration)[†]

[1]*Institute of Physics, Academia Sinica, Taipei, Taiwan 11529, Republic of China*
[2]*Argonne National Laboratory, Argonne, Illinois 60439, USA*
[3]*University of Athens, 157 71 Athens, Greece*
[4]*Institut de Fisica d'Altes Energies, ICREA, Universitat Autonoma de Barcelona, E-08193, Bellaterra (Barcelona), Spain*
[5]*Baylor University, Waco, Texas 76798, USA*
[6]*Istituto Nazionale di Fisica Nucleare Bologna, [ee]University of Bologna, I-40127 Bologna, Italy*
[7]*University of California, Davis, Davis, California 95616, USA*
[8]*University of California, Los Angeles, Los Angeles, California 90024, USA*
[9]*Instituto de Fisica de Cantabria, CSIC-University of Cantabria, 39005 Santander, Spain*
[10]*Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA*
[11]*Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637, USA*
[12]*Comenius University, 842 48 Bratislava, Slovakia; Institute of Experimental Physics, 040 01 Kosice, Slovakia*
[13]*Joint Institute for Nuclear Research, RU-141980 Dubna, Russia*
[14]*Duke University, Durham, North Carolina 27708, USA*
[15]*Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA*
[16]*University of Florida, Gainesville, Florida 32611, USA*
[17]*Laboratori Nazionali di Frascati, Istituto Nazionale di Fisica Nucleare, I-00044 Frascati, Italy*
[18]*University of Geneva, CH-1211 Geneva 4, Switzerland*
[19]*Glasgow University, Glasgow G12 8QQ, United Kingdom*
[20]*Harvard University, Cambridge, Massachusetts 02138, USA*
[21]*Division of High Energy Physics, Department of Physics, University of Helsinki and Helsinki Institute of Physics, FIN-00014, Helsinki, Finland*
[22]*University of Illinois, Urbana, Illinois 61801, USA*
[23]*The Johns Hopkins University, Baltimore, Maryland 21218, USA*
[24]*Institut für Experimentelle Kernphysik, Karlsruhe Institute of Technology, D-76131 Karlsruhe, Germany*
[25]*Center for High Energy Physics: Kyungpook National University, Daegu 702-701, Korea; Seoul National University, Seoul 151-742, Korea; Sungkyunkwan University, Suwon 440-746, Korea; Korea Institute of Science and Technology Information, Daejeon 305-806, Korea; Chonnam National University, Gwangju 500-757, Korea; Chonbuk National University, Jeonju 561-756, Korea*
[26]*Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[27]*University of Liverpool, Liverpool L69 7ZE, United Kingdom*
[28]*University College London, London WC1E 6BT, United Kingdom*
[29]*Centro de Investigaciones Energeticas Medioambientales y Tecnologicas, E-28040 Madrid, Spain*
[30]*Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[31]*Institute of Particle Physics: McGill University, Montréal, Québec, Canada H3A 2T8; Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6; University of Toronto, Toronto, Ontario, Canada M5S 1A7; and TRIUMF, Vancouver, British Columbia, Canada V6T 2A3*
[32]*University of Michigan, Ann Arbor, Michigan 48109, USA*
[33]*Michigan State University, East Lansing, Michigan 48824, USA*
[34]*Institution for Theoretical and Experimental Physics, ITEP, Moscow 117259, Russia*
[35]*University of New Mexico, Albuquerque, New Mexico 87131, USA*
[36]*The Ohio State University, Columbus, Ohio 43210, USA*
[37]*Okayama University, Okayama 700-8530, Japan*
[38]*Osaka City University, Osaka 588, Japan*
[39]*University of Oxford, Oxford OX1 3RH, United Kingdom*

[40] Istituto Nazionale di Fisica Nucleare, Sezione di Padova-Trento, [ff] University of Padova, I-35131 Padova, Italy
[41] University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA
[42] Istituto Nazionale di Fisica Nucleare Pisa, [gg] University of Pisa,
[hh] University of Siena and [ii] Scuola Normale Superiore, I-56127 Pisa, Italy
[43] University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA
[44] Purdue University, West Lafayette, Indiana 47907, USA
[45] University of Rochester, Rochester, New York 14627, USA
[46] The Rockefeller University, New York, New York 10065, USA
[47] Istituto Nazionale di Fisica Nucleare, Sezione di Roma 1,
[jj] Sapienza Università di Roma, I-00185 Roma, Italy
[48] Rutgers University, Piscataway, New Jersey 08855, USA
[49] Texas A&M University, College Station, Texas 77843, USA
[50] Istituto Nazionale di Fisica Nucleare Trieste/Udine,
I-34100 Trieste, [kk] University of Udine, I-33100 Udine, Italy
[51] University of Tsukuba, Tsukuba, Ibaraki 305, Japan
[52] Tufts University, Medford, Massachusetts 02155, USA
[53] University of Virginia, Charlottesville, Virginia 22906, USA
[54] Waseda University, Tokyo 169, Japan
[55] Wayne State University, Detroit, Michigan 48201, USA
[56] University of Wisconsin, Madison, Wisconsin 53706, USA
[57] Yale University, New Haven, Connecticut 06520, USA

(Dated: December 7, 2011)

We present a measurement of the CP-violating parameter $\beta_s^{J/\psi\phi}$ using approximately 6500 $B_s^0 \to J/\psi\phi$ decays reconstructed with the CDF II detector in a sample of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV corresponding to 5.2 fb^{-1} integrated luminosity produced by the Tevatron Collider at Fermilab. We find the CP-violating phase to be within the range $\beta_s^{J/\psi\phi} \in [0.02, 0.52] \cup [1.08, 1.55]$ at 68% confidence level where the coverage property of the quoted interval is guaranteed using a frequentist statistical analysis. This result is in agreement with the standard model expectation at the level of about one Gaussian standard deviation. We consider the inclusion of a potential S-wave contribution to the $B_s^0 \to J/\psi K^+K^-$ final state which is found to be negligible over the mass interval $1.009 < m(K^+K^-) < 1.028$ GeV/c^2. Assuming the standard model prediction for the CP-violating phase $\beta_s^{J/\psi\phi}$, we find the B_s^0 decay width difference to be $\Delta\Gamma_s = 0.075 \pm 0.035\,(\text{stat}) \pm 0.006\,(\text{syst})$ ps^{-1}. We also present the most precise measurements of the B_s^0 mean lifetime $\tau(B_s^0) = 1.529 \pm 0.025\,(\text{stat}) \pm 0.012\,(\text{syst})$ ps, the polarization fractions $|A_0(0)|^2 = 0.524 \pm 0.013\,(\text{stat}) \pm 0.015\,(\text{syst})$ and $|A_\parallel(0)|^2 = 0.231 \pm 0.014\,(\text{stat}) \pm 0.015\,(\text{syst})$, as well as the strong phase $\delta_\perp = 2.95 \pm 0.64\,(\text{stat}) \pm 0.07\,(\text{syst})$ rad. In addition, we report an alternative Bayesian analysis that gives results consistent with the frequentist approach.

PACS numbers: 13.25.Hw, 14.40.Nd, 14.65.Fy

I. INTRODUCTION

Since the discovery of the simultaneous violation of charge and parity quantum numbers (CP violation) in 1964 in the neutral kaon system [1], CP violation has played a crucial role in the development of the standard model (SM) of particle physics and in searches for "new" physics (NP) beyond the SM. In 1973, be-

* Deceased
† With visitors from [a] Istituto Nazionale di Fisica Nucleare, Sezione di Cagliari, 09042 Monserrato (Cagliari), Italy, [b] University of CA Irvine, Irvine, CA 92697, USA, [c] University of CA Santa Barbara, Santa Barbara, CA 93106, USA, [d] University of CA Santa Cruz, Santa Cruz, CA 95064, USA, [e] Institute of Physics, Academy of Sciences of the Czech Republic, Czech Republic, [f] CERN, CH-1211 Geneva, Switzerland, [g] Cornell University, Ithaca, NY 14853, USA, [h] University of Cyprus, Nicosia CY-1678, Cyprus, [i] Office of Science, U.S. Department of Energy, Washington, DC 20585, USA, [j] University College Dublin, Dublin 4, Ireland, [k] ETH, 8092 Zurich, Switzerland, [l] University of Fukui, Fukui City, Fukui Prefecture, Japan 910-0017, [m] Universidad Iberoamericana, Mexico D.F., Mexico, [n] University of Iowa, Iowa City, IA 52242, USA, [o] Kinki University, Higashi-Osaka City, Japan 577-8502, [p] Kansas State University, Manhattan, KS 66506, USA, [q] Korea University, Seoul, 136-713, Korea, [r] University of Manchester, Manchester M13 9PL, United Kingdom, [s] Queen Mary, University of London, London, E1 4NS, United Kingdom, [t] University of Melbourne, Victoria 3010, Australia, [u] Muons, Inc., Batavia, IL 60510, USA, [v] Nagasaki Institute of Applied Science, Nagasaki, Japan, [w] National Research Nuclear University, Moscow, Russia, [x] Northwestern University, Evanston, IL 60208, USA, [y] University of Notre Dame, Notre Dame, IN 46556, USA, [z] Universidad de Oviedo, E-33007 Oviedo, Spain, [aa] CNRS-IN2P3, Paris, F-75205 France, [bb] Texas Tech University, Lubbock, TX 79609, USA, [cc] Universidad Tecnica Federico Santa Maria, 110v Valparaiso, Chile, [dd] Yarmouk University, Irbid 211-63, Jordan, [ll] University of Warwick, Coventry CV4 7AL, United Kingdom, [mm] CPPM, Aix-Marseille Universite, CNRS/IN2P3, Marseille, France.

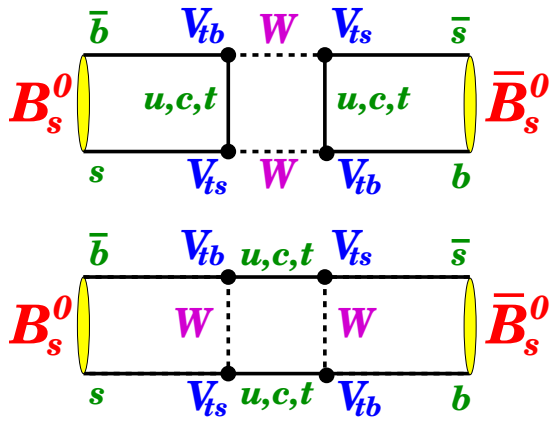

FIG. 1. Lowest order Feynman diagrams that induce B_s^0-$\bar{B}_s^0$ oscillations.

fore the discovery of the fourth (charm) quark [2] and the third generation (bottom [3] and top [4] quarks), Kobayashi and Maskawa proposed an extension to a six-quark model [5] in which CP violation was explained through the quark mixing parametrized by the Cabibbo-Kobayashi-Maskawa (CKM) matrix. A single, irreducible complex phase in the CKM matrix is responsible for all CP-violating effects in the standard model.

One of the most promising processes for the search for physics beyond the standard model is through oscillations of B_s^0 and $\bar{B}_s^0$ mesons. The time evolution of the B_s^0 and $\bar{B}_s^0$ mesons can be described by the Schrödinger equation

$$i\frac{d}{dt}\begin{pmatrix} |B_s^0(t)\rangle \\ |\bar{B}_s^0(t)\rangle \end{pmatrix} = \left[\hat{M}_s - \frac{i}{2}\hat{\Gamma}_s\right]\begin{pmatrix} |B_s^0(t)\rangle \\ |\bar{B}_s^0(t)\rangle \end{pmatrix} \quad (1)$$

$$\text{with } \hat{M}_s = \begin{pmatrix} M_0^s & M_{12}^s \\ M_{12}^{s*} & M_0^s \end{pmatrix} \text{ and } \hat{\Gamma}_s = \begin{pmatrix} \Gamma_0^s & \Gamma_{12}^s \\ \Gamma_{12}^{s*} & \Gamma_0^s \end{pmatrix},$$

where $\hat{M}_s$ and $\hat{\Gamma}_s$ are the mass and decay rate 2×2 matrices. The B_s^0 mixing diagrams shown in Fig. 1 give rise to non-zero off-diagonal elements M_{12}^s and Γ_{12}^s. The diagonalization of $\hat{M}_s - i/2\,\hat{\Gamma}_s$ leads to the heavy and light mass eigenstates B_s^H and B_s^L which are admixtures of the flavor eigenstates B_s^0 and $\bar{B}_s^0$:

$$\begin{aligned} |B_s^H\rangle &= p\,|B_s^0\rangle - q\,|\bar{B}_s^0\rangle \\ \text{and}\quad |B_s^L\rangle &= p\,|B_s^0\rangle + q\,|\bar{B}_s^0\rangle, \end{aligned} \quad (2)$$

where p and q are complex quantities which are related to the respective CKM matrix elements (see Fig. 1) through $q/p = (V_{tb}^* V_{ts})/(V_{tb}V_{ts}^*)$ within the SM, and satisfy $|p|^2 + |q|^2 = 1$.

The off-diagonal elements of the mass and decay matrices can be related to the mass difference between the B_s^H and B_s^L mass eigenstates [6]

$$\begin{aligned} \Delta m_s &= m_s^H - m_s^L \\ &= 2|M_{12}^s|\left(1 + \frac{1}{8}\frac{|\Gamma_{12}^s|^2}{|M_{12}^s|^2}\sin^2\phi_s + \ldots\right), \end{aligned} \quad (3)$$

and the corresponding decay width difference [6]

$$\begin{aligned} \Delta\Gamma_s &= \Gamma_s^L - \Gamma_s^H \\ &= 2|\Gamma_{12}^s|\cos\phi_s\left(1 - \frac{1}{8}\frac{|\Gamma_{12}^s|^2}{|M_{12}^s|^2}\sin^2\phi_s + \ldots\right), \end{aligned} \quad (4)$$

where $\phi_s = \arg(-M_{12}^s/\Gamma_{12}^s)$. Typically in the B_s^0 system corrections of order $(\Gamma_{12}^s/M_{12}^s)^2$ can be neglected. The total B_s^0 decay width $\Gamma_s = (\Gamma_s^L + \Gamma_s^H)/2 = \hbar/\tau(B_s^0)$ is related to the mean B_s^0 lifetime $\tau(B_s^0)$, while the mass difference Δm_s is proportional to the frequency of B_s^0-$\bar{B}_s^0$ oscillations, first observed by CDF [7], where the current world average is $\Delta m_s = 17.77 \pm 0.10 \pm 0.07$ ps^{-1} [8]. Assuming no CP violation in the B_s^0 system, which is justified in the SM where the CP-violating phase is expected to be small ($\phi_s^{SM} \approx 0.004$ [9]), the B_s^0 mass eigenstates are also CP eigenstates where Γ_L is the width of the CP-even state corresponding to the short lived state in analogy to the kaon system where the short-lived state (K_S^0) is CP even. Γ_H is the width of the CP-odd state corresponding to the long lived B_s^0 state.

A broad class of theoretically well-founded extensions of the SM predicts new sources of CP-violating phases [10–12]. In the presence of physics beyond the standard model, the quantities describing the B_s^0 system can be modified by a phase ϕ_s^{NP} as follows [9, 12]:

$$\begin{aligned} &\Gamma_{12}^s = {\Gamma_{12}^s}^{SM}, \\ &M_{12}^s = {M_{12}^s}^{SM} \times \Delta_s, \quad \text{where } \Delta_s = |\Delta_s|\, e^{i\,\phi_s^{NP}}. \end{aligned} \quad (5)$$

In this parameterization it is assumed that new physics has a negligible effect on Γ_{12}^s, which is the case for a large class of new physics models and confirmed by experimental data [8], and only M_{12}^s is changed by the factor Δ_s. As the precise determination of the B_s^0 oscillation frequency [8] is well within the standard model expectation, contributions of new physics to the magnitude $|\Delta_s|$ on the level of greater than about 10-20% are unlikely [12]. A currently preferred place to search for new physics is through the phase ϕ_s which is unconstrained by measurements of the B_s^0-$\bar{B}_s^0$ oscillation frequency. Since ϕ_s^{SM} is small, in a new-physics scenario with a large contribution to ϕ_s, the approximation $\phi_s = \phi_s^{SM} + \phi_s^{NP} \approx \phi_s^{NP}$ can be made.

An excellent probe of this new-physics phase [13] is through the decay mode $B_s^0 \to J/\psi\phi(1020)$, with $J/\psi \to \mu^+\mu^-$ and $\phi(1020) \to K^+K^-$. Note that throughout this paper we refer to $\phi(1020)$ just as ϕ for brevity. Figure 2 shows the leading $B_s^0 \to J/\psi\phi$ decay diagram on the left-hand side while the decay topology is indicated on the right. The relative phase between the decay amplitudes with and without mixing is $2\beta_s^{J/\psi\phi}$, which is responsible for CP violation in $B_s^0 \to J/\psi\phi$ decays. Neglecting higher order loop corrections (penguin contributions) and assuming that there is no CP violation present in the decay amplitude, $2\beta_s^{J/\psi\phi}$ can be associated with $e^{i\,2\beta_s^{J/\psi\phi}} = (q/p)\cdot(A_f/\bar{A}_f)$, where A_f and $\bar{A}_f$ are the decay amplitudes in $B_s^0 \to J/\psi\phi$ and $\bar{B}_s^0 \to$

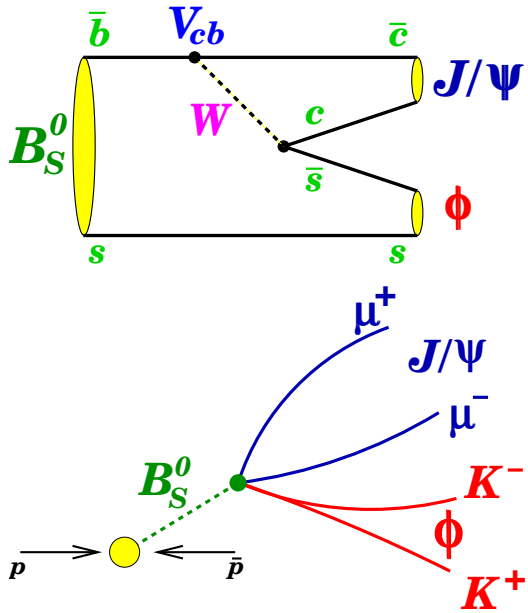

FIG. 2. Leading $B_s^0 \to J/\psi\phi$ decay diagram (top) and decay topology (bottom).

$J/\psi\phi$, respectively. In the standard model this phase is $\beta_s^{SM} = \arg\left[-(V_{ts}V_{tb}^*)/(V_{cs}V_{cb}^*)\right]$, where V_{ij} are again the corresponding elements of the CKM quark mixing matrix. Global fits of experimental data tightly constrain the CP-violating phase to small values in the context of the standard model, $\beta_s^{SM} \approx 0.02$ [9, 14]. The presence of new physics could modify this phase by the same quantity ϕ_s^{NP} that affects the phase ϕ_s allowing $\beta_s^{J/\psi\phi}$ to be expressed as $2\beta_s^{J/\psi\phi} = 2\beta_s^{SM} - \phi_s^{NP}$ [9, 12]. Assuming that new physics effects are much larger than the SM phase, we can again approximate $2\beta_s^{J/\psi\phi} \approx -\phi_s^{NP} \approx -\phi_s$.

The first measurements of the CP-violating phase $\beta_s^{J/\psi\phi}$ by the CDF and D0 experiments [15, 16] each showed a mild inconsistency with the SM prediction where interestingly both results deviated in the same direction. A preliminary combination of the CDF and D0 analyses with samples corresponding to 2.8 fb^{-1} integrated luminosity was inconsistent with the SM expectation at the level of about two standard deviations [17]. In addition, recent dimuon asymmetry results from the D0 collaboration [18] suggest additional indication for effects of physics beyond the standard model in B_s^0 mixing. During the preparation of this manuscript the D0 collaboration released an updated measurement of the CP-violating phase $\beta_s^{J/\psi\phi}$ using a data sample based on 8 fb^{-1} of integrated luminosity [19], while the LHCb collaboration presented a first preliminary measurement of the B_s^0 mixing phase showing confidence regions in agreement with the SM prediction within one standard deviation [20].

This paper presents a measurement of the CP-violating phase $\beta_s^{J/\psi\phi}$ using about four times the integrated luminosity of our previously published analysis [15], as well as additional improvements in flavor tagging and the inclusion of potential S-wave contributions to the $B_s^0 \to J/\psi\phi$ signal. This article is organized as follows. In Sec. II we give an overview of the work flow of the analysis, while we describe the CDF experiment in Sec. III. The data selection is summarized in Sec. IV. The applied flavor tagging is discussed in Sec. V and the likelihood fit function is detailed in Sec. VI. The measurements of the B_s^0 mean lifetime, $\Delta\Gamma_s$, the polarization fraction and the respective systematic uncertainties are described in Sec. VII and Sec. VII A, respectively. The results on $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$ using a frequentist analysis are summarized in Sec. VIII while Sec. IX describes an alternative Bayesian approach. A summary is given in Sec. X.

II. MEASUREMENT OVERVIEW

The measurement of the phase $\beta_s^{J/\psi\phi}$ relies on an analysis of the time-evolution and kinematics of the $B_s^0 \to J/\psi\phi$ decay, which features a pseudoscalar meson decaying to two vector mesons. Consequently, the total spin in the final $J/\psi\phi$ state is either 0, 1 or 2. To conserve the total angular momentum, the orbital angular momentum L between the final state decay products must be either 0, 1 or 2. While the J/ψ and ϕ are CP-even eigenstates, the $J/\psi\phi$ final state has a CP eigenvalue given as $(-1)^L$. Consequently, the states with orbital angular momentum 0 and 2 are CP-even while the state with angular momentum 1 is CP-odd. We use both the decay time of the B_s^0 and the decay angles of the $J/\psi \to \mu^+\mu^-$ and $\phi \to K^+K^-$ mesons to statistically separate the CP-odd and CP-even components of the $J/\psi\phi$ final state.

There are three angles that completely define the directions of the four particles in the final state. We use the angular variables $\vec{\rho} = \{\cos\theta_T, \phi_T, \cos\psi_T\}$ as defined in the transversity basis [21]. In the following relations we use a notation where $\vec{p}(A)_B$ denotes the three-momentum of particle A in the rest frame of particle B. With this notation, the helicity angle ψ_T of the K^+ is defined in the ϕ rest frame as the angle between $\vec{p}(K^+)$ and the negative J/ψ direction:

$$\cos\psi_T = -\frac{\vec{p}(K^+)_\phi \cdot \vec{p}(J/\psi)_\phi}{|\vec{p}(K^+)_\phi| \cdot |\vec{p}(J/\psi)_\phi|}. \tag{6}$$

To calculate the other two angles, we first define a coordinate system through the directions

$$\begin{aligned} \hat{x} &= \frac{\vec{p}(\phi)_{J/\psi}}{\left|\vec{p}(\phi)_{J/\psi}\right|}, \\ \hat{y} &= \frac{\vec{p}(K^+)_{J/\psi} - \left[\vec{p}(K^+)_{J/\psi} \cdot \hat{x}\right]\hat{x}}{\left|\vec{p}(K^+)_{J/\psi} - \left[\vec{p}(K^+)_{J/\psi} \cdot \hat{x}\right]\hat{x}\right|}, \\ \hat{z} &= \hat{x} \times \hat{y}. \end{aligned} \tag{7}$$

With this coordinate system the following angles of the direction of the μ^+ in the J/ψ rest frame are calculated

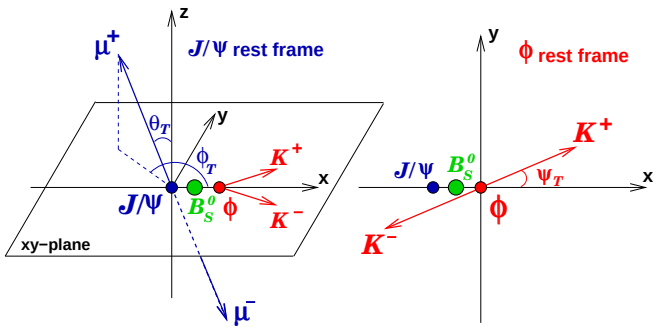

FIG. 3. Illustration of definition of transversity angles θ_T, ϕ_T, and ψ_T.

as

$$\cos\theta_T = \frac{\vec{p}(\mu^+)_{J/\psi}}{|\vec{p}(\mu^+)_{J/\psi}|} \cdot \hat{z}, \tag{8}$$

$$\phi_T = \tan^{-1}\left(\frac{\frac{\vec{p}(\mu^+)_{J/\psi}}{|\vec{p}(\mu^+)_{J/\psi}|} \cdot \hat{y}}{\frac{\vec{p}(\mu^+)_{J/\psi}}{|\vec{p}(\mu^+)_{J/\psi}|} \cdot \hat{x}}\right), \tag{9}$$

where the ambiguity of the angle ϕ_T is resolved using signs of $\vec{p}(\mu^+)_{J/\psi} \cdot \hat{x}$ and $\vec{p}(\mu^+)_{J/\psi} \cdot \hat{y}$. The definitions of the transversity angles are illustrated in Fig. 3.

The decay is further described in terms of the polarization states of the vector mesons, either longitudinal (0), or transverse to their directions of motion, and in the latter case, parallel ($\|$) or perpendicular ($\perp$) to each other. The corresponding amplitudes, which depend on time t, are called A_0, $A_\|$ and $A_\perp$, respectively. The transverse linear polarization amplitudes $A_\|$ and $A_\perp$ correspond to CP-even and CP-odd final states at decay time $t = 0$, respectively. The longitudinal polarization amplitude A_0 corresponds to a CP-even final state. The three states in the transversity basis are easily expressed as linear combinations of states in either the helicity basis ($++$, 00, $--$) or the orbital angular momentum basis (S, P, D). In the helicity basis, $A_\|$ and $A_\perp$ are linear combinations of the states with helicities $++$ and $--$, while the state corresponding to A_0 is the same in both transversity and helicity bases. In terms of the S, P and D-waves, the states described by A_0 and $A_\|$ are linear combinations of S and D waves, while $A_\perp$ corresponds to the P-wave state. Since only differences between the strong phases of these amplitudes are observable, we define the strong phases relative to $A_0(0)$ at time $t = 0$: $\delta_0 = 0$, $\delta_\| = \arg[A_\|(0)A_0^*(0)]$ and $\delta_\perp = \arg[A_\perp(0)A_0^*(0)]$. We note that the strong phases $\delta_\|$ and $\delta_\perp$ are either 0 or π in the absence of final state $J/\psi\phi$ interactions. Deviations of these phases from 0 or π indicate breaking of the factorization hypothesis which assumes no interaction between the J/ψ and ϕ in the final state [9, 13].

If the decay width difference between the B_s^0 mass eigenstates $\Delta\Gamma_s$ is different from zero, a time-dependent angular analysis without flavor tagging is sensitive to $\beta_s^{J/\psi\phi}$ because of the interference between CP-odd and CP-even components [22]. The sensitivity to $\beta_s^{J/\psi\phi}$ can be improved by separating mesons produced as B_s^0 from those produced as $\bar{B}_s^0$ in order to detect CP asymmetries in the fast B_s^0-$\bar{B}_s^0$ flavor oscillations given sufficient decay time resolution. The process of separating B_s^0 mesons from $\bar{B}_s^0$ mesons at production is called flavor tagging.

The angular-dependence and flavor tagged (see Sec. V) time-dependence are combined in an unbinned maximum likelihood fit. The fit is used to extract $\beta_s^{J/\psi\phi}$, the B_s^0 decay width difference $\Delta\Gamma_s$, the average B_s^0 lifetime, the transversity amplitudes and the strong phases. Since a contamination from K^+K^- final states that do not originate from a ϕ decay can contribute to the K^+K^- mass window used to identify ϕ candidates in this analysis, we consider potential contributions from other $B_s^0 \to J/\psi K^+K^-$ decays in our $B_s^0 \to J/\psi\phi$ candidate sample. In such decays the relative angular momentum of the two kaons is assumed to be zero (S-wave) as expected, for example, from $f_0(980) \to K^+K^-$ decays. Continuum $B_s^0 \to J/\psi K^+K^-$ decays with angular momentum higher than zero are expected to be suppressed. In all such cases the K^+K^- system is assumed to be in a partial S-wave whose angular momentum combined with that of the J/ψ leads to a CP-odd final state [23]. The S-wave contribution is included in the time-dependent angular analysis and the S-wave fraction together with its corresponding phase δ_{SW} are determined as parameters in the maximum likelihood fit. The inclusion of the S-wave in the likelihood function constitutes a significant improvement with respect to earlier measurements [15, 16].

Due to the non-Gaussian behavior of the likelihood function with respect to the parameters $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$ [15, 16], we use a frequentist analysis to obtain confidence regions for both parameters. We also determine point estimates for other parameters of interest, like the polarization fractions and the B_s^0 lifetime. In addition, we perform an alternative Bayesian approach, through the use of priors, applied to probability densities determined with Markov chain Monte Carlo.

III. CDF II DETECTOR AND TRIGGER

The CDF II detector employs a cylindrical geometry around the $p\bar{p}$ interaction region with the proton direction defining the positive z-direction. Most of the quantities used for candidate selection are measured in the plane transverse to the z-axis. In the CDF coordinate system, φ is the azimuthal angle, θ is the polar angle measured from the proton direction, and r is the radius perpendicular to the beam axis. The pseudorapidity η is defined as $\eta = -\ln[\tan(\theta/2)]$. The transverse momentum, p_T, is the component of the particle momentum, p, transverse to the z-axis ($p_T = p \cdot \sin\theta$), while $E_T = E \cdot \sin\theta$, with E being the energy measured in the calorimeter.

The CDF II detector features excellent lepton identification and charged particle tracking and is described

in detail elsewhere [24, 25]. The parts of the detector relevant to the reconstruction of the $B_s^0 \to J/\psi\phi$ decays used in this measurement are briefly summarized below. The detector nearest to the $p\bar{p}$ interaction region is a silicon vertex detector (SVX II) [26], which consists of five concentric layers of double-sided sensors located at radii between 2.5 and 10.6 cm plus one additional single-sided layer of silicon (L00) [27] mounted directly onto the beam pipe at radius $r \sim 1.5$ cm. In addition, two forward layers plus one central layer of double-sided silicon located outside the SVX II at radii of 20-29 cm make up the intermediate silicon layers (ISL) [28]. Together with the SVX II, the ISL detector extends the sensitive region of the CDF II tracking detector to $|\eta| \leq 2.0$. The CDF silicon system has a typical hit resolution of ~ 11 μm and provides three-dimensional track reconstruction. It is used to identify displaced vertices associated with bottom hadron decays which are reconstructed with a typical transverse resolution of 10-20 μm. The central outer tracker (COT) [29], an open-cell drift chamber with 30 240 sense wires arranged in 96 layers combined into four axial and four stereo super-layers (SL), provides tracking from a radius of ~ 40 cm out to a radius of 132 cm covering $|z| < 155$ cm, as well as the main measurement of track momentum with a resolution of $\sigma(p_T)/p_T^2 \sim 0.15\%\,[\mathrm{GeV}/c]^{-1}$ for high momentum tracks. The COT also provides specific ionization energy loss, $\mathrm{d}E/\mathrm{d}x$, information for charged particle identification with approximately $1.5\,\sigma$ separation between pions and kaons with momenta greater than 2 GeV/c [30]. The central tracking system is immersed in a superconducting solenoid that provides a 1.4 T axial magnetic field. Right outside the solenoid, the time-of-flight (TOF) detector provides additional particle identification for low-momentum particles.

Central electromagnetic (CEM) [31] and hadronic (CHA) [32] calorimeters ($|\eta| < 1.1$) are located outside the COT and the solenoid, where they are arranged in a projective-tower geometry. The electromagnetic and hadronic calorimeters are lead-scintillator and iron-scintillator sampling devices, respectively. The energy resolution for the CDF central calorimeter is $\sigma(E_T)/E_T = [(13.5\%/\sqrt{E_T})^2 + (1.5\%)^2]^{1/2}$ for electromagnetic showers [31, 33] and $\sigma(E_T)/E_T = [(75\%/\sqrt{E_T})^2 + (3\%)^2]^{1/2}$ for hadrons [25, 32], where E_T is measured in GeV. A layer of proportional chambers (CES), with wire and strip readout, is located six radiation lengths deep in the CEM calorimeters, near the electromagnetic shower maximum. The CES provides a measurement of electromagnetic shower profiles in both the φ and z directions for use in electron identification.

Muon candidates are identified by multi-layer drift chambers and scintillator counters [34]. Four layers of planar drift chambers (CMU) are located outside the central calorimeter at a radius of 347 cm from the beam line. The CMU system covers $|\eta| \leq 0.6$ and can be reached by muons with p_T in excess of ~ 1.4 GeV/c. To reduce the probability of misidentifying penetrating hadrons as muon candidates in the central detector region, four additional layers of drift chambers (CMP) are located behind 0.6 m of steel outside the CMU system. Approximately 84% of the solid angle for $|\eta| \leq 0.6$ is covered by the CMU detector, 63% by the CMP, and 53% by both. To reach these two detectors, particles produced at the primary interaction vertex, with a polar angle of 90°, must traverse material totaling 5.5 and 8.8 pion interaction lengths, respectively. Muons with hits in both the CMU and CMP detectors are called CMUP muons. An additional set of muon chambers (CMX) is located in the pseudorapidity interval $0.6 < |\eta| < 1.0$ to extend the polar acceptance of the muon system to the forward region. Approximately 71% of the solid angle for $0.6 < |\eta| < 1.0$ is covered by the free-standing conical arches of the CMX. The calorimeter, magnet yoke of the detector, and the steel support structure provide shielding of about 6.2 pion interaction lengths.

The data used in this measurement are collected with dimuon triggers [24]. Muons are reconstructed as track stubs in the CMU, CMP and CMX chambers. Muon stubs are matched to tracks reconstructed using COT axial information from the extremely fast trigger (XFT) [35]. The dimuon trigger requires at least one central muon matching the CMU or CMUP chambers, while the second muon can be either central or forward, matching to the CMU or CMX detectors, respectively. The CMU, CMUP, or CMX muons must satisfy $p_T > 1.5$ GeV/c, $p_T > 3.0$ GeV/c and $p_T > 2.0$ GeV/c, respectively. The two trigger muon candidates are required to be oppositely charged, have an opening angle inconsistent with a cosmic ray event, and the invariant mass of the muon pair must satisfy $2.7 < m(\mu^+\mu^-) < 4$ GeV/c^2.

IV. DATA RECONSTRUCTION AND SELECTION

We use a data sample corresponding to an integrated luminosity of 5.2 fb^{-1} collected with all CDF II detector subsystems functioning. In addition, all analyzed data passed the dimuon trigger requirements given above. We begin our offline reconstruction of the $B_s^0 \to J/\psi(\to \mu^+\mu^-)\phi(\to K^+K^-)$ decay mode by requiring two muon candidate tracks that extrapolate to a track segment in the muon detectors reapplying the appropriate transverse momentum requirements for the respective trigger muons using offline-reconstructed quantities. To reconstruct the J/ψ candidate, a kinematic fit constraining the two oppositely charged muon candidate tracks to a common interaction point (vertex) is applied. All other charged particles in the event are assumed to be kaon candidates and combined as opposite-charge pairs to reconstruct ϕ meson candidates. Finally, all four candidate tracks are combined in a kinematic fit that constrains the muon candidates to the J/ψ world average mass [8] and requires the four tracks to originate from a common three-dimensional vertex point.

For each event a primary interaction point is constructed from all reconstructed tracks in an event, excluding the J/ψ and ϕ candidate tracks. This interaction point is used in the calculation of the B_s^0 proper decay time, $ct = m(B_s^0)L_{xy}(B_s^0)/p_T(B_s^0)$, where $L_{xy}(B_s^0)$ is the distance from the primary vertex point to the $B_s^0 \to J/\psi\phi$ decay vertex projected onto the momentum of the B_s^0 in the plane transverse to the proton beam direction, $m(B_s^0)$ is the nominal mass of the B_s^0 meson [8], and $p_T(B_s^0)$ is its measured transverse momentum.

A. Basic selection criteria

For the final selection of $B_s^0 \to J/\psi\phi$ candidates, we use the Neurobayes [36] artificial neural network (ANN) to distinguish signal events from background. Prior to the training of the ANN, we apply basic selection requirements in order to ensure track and vertex quality, as well as the reconstructed particle candidates to have kinematic properties appropriate for B_s^0, J/ψ, and ϕ mesons. These basic selection criteria are listed in Table I summarizing the standard quality requirements that were imposed on track candidates. We require kaon candidates to have transverse momentum greater than 400 MeV/c and all kinematic fits are required to have $\chi^2_{r\phi} < 50$, where $\chi^2_{r\phi}$ is the χ^2 of the two-dimensional $r\phi$-vertex fit for four degrees of freedom (dof). To select $B_s^0 \to J/\psi\phi$ candidates, we require that the invariant mass of the muon pair lies within the mass region $3.04 < m(\mu^+\mu^-) < 3.14$ GeV/c^2 corresponding to an interval around the world average J/ψ mass [8] of about $\pm 2.5\sigma$ where σ is the J/ψ mass resolution. The invariant mass of the kaon pair is required to be within $1.009 < m(K^+K^-) < 1.028$ GeV/c^2 corresponding to a $\pm 2.5\sigma$ interval around the nominal ϕ mass [8] where σ corresponds to the ϕ mass resolution. The mass of the reconstructed $J/\psi\phi$ candidate has to be in the mass window $5.1 < m(J/\psi K^+K^-) < 5.6$ GeV/c^2 corresponding to a ± 250 MeV/c^2 interval around the nominal B_s^0 mass [8]. Additionally we require that the transverse momentum of the ϕ candidate is greater than 1.0 GeV/c and the B_s^0 candidate has a transverse momentum of more than 4.0 GeV/c.

B. Monte Carlo simulation

We use simulated $B_s^0 \to J/\psi\phi$ Monte Carlo (MC) event samples to describe the signal in the training of the artificial neural network. These MC samples are also employed in the determination of the transversity angle efficiencies due to the non-hermeticity of the CDF II detector (see Sec. VI). We simulate the generation and fragmentation of b quarks using the BGENERATOR program [37]. It is based on next-to-leading-order QCD calculations and the Peterson fragmentation function [38] tuned to the b-quark momentum spectrum measured at CDF [24]. The decay of the B_s^0 meson is simulated with the EVTGEN decay package [39]. The interaction of the generated particles with the CDF II detector is simulated with the full GEANT [40] based CDF II detector simulation package [41]. We subject the simulated events to the same trigger requirements and reconstruction process as our data events. The B_s^0 decays are simulated according to the phase space available to the decay averaging over the spin states of the decay daughters. This procedure ensures that all transversity angles are generated flat for B_s^0 decays.

TABLE I. Basic selection requirements applied to the $B_s^0 \to J/\psi\phi$ four-track system as used in training the artificial neural network.

Quantity	Selection requirement
COT hits	≥ 2 stereo and ≥ 2 axial super-layers with ≥ 5 hits each
r-ϕ silicon hits	≥ 3
Kaon track p_T	> 0.4 GeV/c
Vertex $\chi^2_{r\phi}$ (4 dof)	< 50
J/ψ mass region	$3.04 < m(\mu^+\mu^-) < 3.14$ GeV/c^2
ϕ mass region	$1.009 < m(K^+K^-) < 1.028$ GeV/c^2
B_s^0 mass region	$5.1 < m(J/\psi K^+K^-) < 5.6$ GeV/c^2
$p_T(\phi)$	> 1.0 GeV/c
$p_T(B_s^0)$	> 4.0 GeV/c

C. Selection using artificial neural network

The information from several kinematic variables is combined into a single discriminant by the artificial neural network. Based on the discriminant, an event is classified as background-like or signal-like on a scale of -1 to $+1$. Correlations between variables are taken into account by the ANN, and the weight of each variable in the overall discriminant depends on its correlation with other variables. Our artificial neural network is trained on a signal sample based on 350 000 $B_s^0 \to J/\psi\phi$ Monte Carlo events. The background sample used in training the ANN consists of $\sim 300\,000$ data events taken from the B_s^0 invariant mass sideband regions $(5.2, 5.3) \cup (5.45, 5.55)$ GeV/c^2.

We use the following variables as input to the ANN listed in order of discriminating power and relevance to the final discriminant: the transverse momentum p_T of the ϕ meson, the kaon likelihood [22] based on TOF and $\mathrm{d}E/\mathrm{d}x$ information, the muon likelihood [42] for the J/ψ muon daughters, $\chi^2_{r\phi}$ for the B_s^0 decay vertex reconstruction, the transverse momentum p_T of the B_s^0 meson, and the probabilities to reconstruct vertices of the B_s^0, ϕ, and J/ψ candidates. These vertex probabilities are χ^2 probabilities for the three-dimensional vertex fit, while $\chi^2_{r\phi}$ is the goodness of fit for the two-dimensional vertex fit.

The muon and kaon likelihoods are quantities used for particle identification. The algorithm determining the muon likelihood is described in Ref. [42]. The kaon likelihood [22] is a combined discriminant constructed from the kaon track specific energy loss, $\mathrm{d}E/\mathrm{d}x$, and its time-of-flight information. Both likelihood variables have been calibrated on large control samples.

D. Optimization of selection

Once the artificial neural network is trained, we choose a discriminant cut value that provides the best expected average resolution (sensitivity) to the CP-violating phase $\beta_s^{J/\psi\phi}$. This optimization differs from our previous analysis [15] where the significance of B_s^0 signal yield was maximized using $S/\sqrt{S+B}$ as figure of merit, in which the signal S was obtained from Monte Carlo simulation of B_s^0 events and the background B taken from the B_s^0 sideband regions. In the current optimization we study the sensitivity to $\beta_s^{J/\psi\phi}$ as a function of ANN discriminant C_{NN} using pseudoexperiments that are generated to mimic our data with a specific signal-to-background ratio for a given cut value on C_{NN}. Using the likelihood fit function described in Sec. VI, we repeat the entire analysis procedure for each pseudoexperiment and evaluate the distributions of the estimated variance on $\beta_s^{J/\psi\phi}$ to find the ANN discriminant cut value that gives the best expected average resolution for $\beta_s^{J/\psi\phi}$ [43].

In detail, the pseudoexperiments are created by randomly sampling the probability density functions (PDF) for variables used in the fit to describe the data as outlined in Sec. VI. We simulate the effect of varying the cut on the ANN output variable by generating pseudoexperiments at different values of S/N_{TOT}, where N_{TOT} is the total number of events in the $J/\psi\phi$ invariant mass window, and S is the number of B_s^0 signal events. N_{TOT} and S are determined from mass fits to the data for different cut values of C_{NN}. The input values of all other parameters in the PDF are kept the same for all pseudoexperiments corresponding to the same ANN cut; only N_{TOT} and S/N_{TOT} are varied. We take the parameter input values from the results of the unbinned maximum likelihood fit of $B_s^0 \to J/\psi\phi$ decays using data corresponding to an integrated luminosity of 2.8 fb^{-1} [17]. The only exceptions are the parameters describing the tagging power, which correspond to the total tagging effectiveness of both the opposite and same side tagging algorithms (see Sec. V) valid for the full data set of 5.2 fb^{-1} integrated luminosity. To verify in our optimization that the expected average resolution to $\beta_s^{J/\psi\phi}$ is independent of the true values of our parameters of interest, $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$, we generate pseudoexperiments at a few points in ($\beta_s^{J/\psi\phi}$, $\Delta\Gamma_s$) parameter space: ($\beta_s^{J/\psi\phi} = 0.5$, $\Delta\Gamma_s = 0.12$ ps^{-1}), ($\beta_s^{J/\psi\phi} = 0.02$, $\Delta\Gamma_s = 0.1$ ps^{-1}), and ($\beta_s^{J/\psi\phi} = 0.3$, $\Delta\Gamma_s = 0.09$ ps^{-1}). We also verify that the widths of the uncertainty distributions do not vary as a function of artificial neural network discriminant [43].

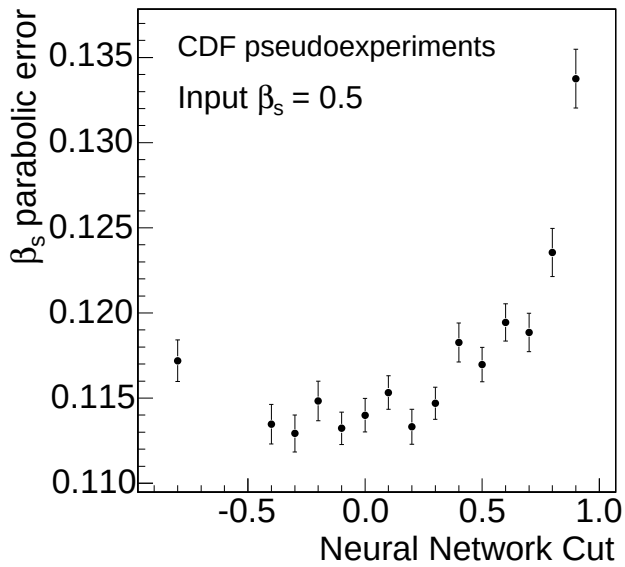

FIG. 4. Magnitude of expected uncertainty on $\beta_s^{J/\psi\phi}$ as a function of cut value on the artificial neural network discriminant for pseudoexperiments generated with input values $\beta_s^{J/\psi\phi} = 0.5$ and $\Delta\Gamma_s = 0.12$ ps^{-1}.

The most probable values of the $\beta_s^{J/\psi\phi}$ uncertainty for pseudoexperiments generated at ($\beta_s^{J/\psi\phi} = 0.5$, $\Delta\Gamma_s = 0.12$ ps^{-1}) are shown as a function of the cut value on the ANN output variable in Fig. 4. It is apparent that tight cuts on C_{NN} correspond to larger $\beta_s^{J/\psi\phi}$ expected uncertainties. It should be emphasized that this technique is not intended to guarantee a particular uncertainty on $\beta_s^{J/\psi\phi}$, but is merely used to identify the trend in $\beta_s^{J/\psi\phi}$ uncertainty size as a function of cut value on the ANN output variable.

The trend in size of the $\beta_s^{J/\psi\phi}$ uncertainty distributions is similar for the other sets of ($\beta_s^{J/\psi\phi}$, $\Delta\Gamma_s$) at which pseudoexperiments are generated. The expected statistical uncertainty on $\beta_s^{J/\psi\phi}$ shows a shallow minimum around $C_{NN} \sim 0$. We adopt a cut on the ANN output discriminant at > 0.2, where the uncertainty on $\beta_s^{J/\psi\phi}$ is small, allowing us to avoid adding unnecessary amounts of background which we would include by going to a lower cut value on C_{NN}. For comparison, an optimization using $S/\sqrt{S+B}$ as figure of merit yields a cut value of the ANN discriminant of about 0.9 resulting in an almost 20% larger expected uncertainty on $\beta_s^{J/\psi\phi}$.

A cut on the ANN output discriminant of 0.2 yields 6504 ± 85 $B_s^0 \to J/\psi\phi$ signal events, as extracted by a fit to the invariant mass with a single Gaussian with flat background as shown in Fig. 5 on the left-hand side. The signal and sideband regions used to describe sig-

nal and background events in the PDF (see Sec. VI) are indicated by the areas with vertical and horizontal lines, respectively. They explicitly correspond to $5.340 < m(J/\psi K^+K^-) < 5.393$ GeV/c^2 for the signal region and $(5.287, 5.314) \cup (5.419, 5.446)$ GeV/c^2 for the sidebands. The right-hand side of Fig. 5 shows the $J/\psi K^+K^-$ invariant mass distribution with an additional lifetime requirement $ct > 60$ μm on the B_s^0 candidate indicating the combinatorial background to be mainly prompt.

E. Initial study of S-wave contribution

Since this analysis considers a potential S-wave contribution, which is not coming from $\phi \to K^+K^-$ decays, to the $B_s^0 \to J/\psi K^+K^-$ signal in the description of the likelihood function (see Sec. VI), we perform an initial study to obtain an estimate of the size of the S-wave contribution using a data sample corresponding to integrated luminosity of 3.8 fb^{-1}. We select ϕ candidates in a larger K^+K^- mass window $0.98 < m(K^+K^-) < 1.08$ GeV/c^2 and form B_s^0 candidates. Taking B_s^0 candidates from the signal region defined above, we obtain the K^+K^- invariant mass distribution shown in Fig. 6. Using a binned likelihood method, we fit the K^+K^- invariant mass distribution with a ϕ signal component modeled by a template from $B_s^0 \to J/\psi\phi$ MC simulation allowing for a mass-dependent width consistent with the parameterization in Eq. (17) (see Sec. VI). The combinatoric background is modeled by a histogram taken from the B_s^0 sidebands. An additional component, which takes into account B^0 reflections where the pion from a B^0 decay is misidentified as a kaon, is obtained from an inclusive $B^0 \to J/\psi X$ MC simulation. The fractions of combinatorial background and B^0 reflection are fixed from a fit to the B_s^0 invariant mass, which prevents these components from absorbing a non-resonant component in the K^+K^- mass distribution if present. Together with the ϕ signal and the fixed components, an additional contribution is included in the fit allowing for a possible non-resonant S-wave contribution from $B_s^0 \to J/\psi K^+K^-$ or $B_s^0 \to J/\psi f_0(980)$. This component is modeled either flat in K^+K^- mass or following a mass parameterization suggested in Ref. [44]. In either case the S-wave fraction is found to be compatible with zero as indicated in Fig. 6. From this study we do not expect a significant S-wave contribution across the ϕ mass region.

V. FLAVOR TAGGING

To maximize the sensitivity to the CP-violating phase $\beta_s^{J/\psi\phi}$, we employ flavor tagging algorithms to determine whether the reconstructed $J/\psi\phi$ candidate was a B_s^0 meson or its anti-particle $\bar{B}_s^0$ at the time of production. Flavor tagging algorithms assign each B_s^0 meson candidate a tagging decision and a tagging dilution. The tagging decision can be $\xi = +1$ or -1, corresponding to a B_s^0 meson at production or a $\bar{B}_s^0$ meson, respectively. A value of $\xi = 0$ means the tagging algorithm failed and no tag is assigned. The tagging dilution $\mathcal{D}$ is related to the probability that the tagging decision is correct, $\mathcal{D} = 1 - 2p_W$, where p_W is the probability of an incorrect tag or mis-tag. In general, the dilution is obtained by counting the number of correctly and incorrectly assigned tags $\mathcal{D} = (N_R - N_W)/(N_R + N_W)$, where N_R is the number of correct tags and N_W is the number of incorrect tags. The flavor tagging performance is quantified by the product between the tagging efficiency and the squared dilution $\varepsilon\mathcal{D}^2$. The tagging efficiency ε is defined as the number of events that receive a tag, divided by the total number of events considered. We use two types of flavor tagging algorithms: a same side (SST) and an opposite side (OST) tagger. Due to the use of information from different event hemispheres, there is no overlap between both taggers by construction. In particular, the SST algorithm considers only tracks within a cone of $\sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.7$ around the B_s^0 candidate while the OST tagger only includes tracks outside that cone, allowing us to treat SST and OST as uncorrelated tagging methods.

A. Same side tagging

In this analysis we employ a same side kaon tagging (SSKT) algorithm which uses the charge of the kaon produced in association with the b ($\bar{b}$) quark in the fragmentation process forming the $\bar{B}_s^0$ (B_s^0) meson as illustrated in Fig. 7. To determine the b ($\bar{b}$) production flavor we attempt to find kaon tracks produced in the hadronization of the B_s^0 meson. The strangeness of the B_s^0 meson preferentially produces associated kaons in the fragmentation process. The charges of these nearby kaons are correlated to the b quark content of the B_s^0 meson and provide an opportunity to identify the initial flavor of the B_s^0 meson. However, the B_s^0 meson can also be accompanied by a neutral kaon which cannot be used to tag the B_s^0 flavor and therefore lowers the tagging power. Misidentification of the associated charged kaon leads to a further decrease of the tagging dilution. The SSKT algorithm was developed on a simulated high statistics Monte Carlo B_s^0 sample, using the $B_s^0 \to J/\psi\phi$ and $B_s^0 \to D_s^-\pi^+$ decay modes. We use particle identification ($\mathrm{d}E/\mathrm{d}x$ and time-of-flight) to help identify the associated track as a kaon [7, 45, 46].

We calibrate the SSKT algorithm [47] by measuring the B_s^0 mixing frequency on a data set corresponding to 5.2 fb^{-1} integrated luminosity. Using CDF's Silicon Vertex Trigger [48], we select events that contain $B_s^0 \to D_s^-\pi^+$ candidates. The trigger configuration used to collect this heavy flavor data sample is described in Ref. [49]. $B_s^0 \to D_s^-\pi^+$ events are fully reconstructed in three D_s^- decay modes: $D_s^- \to \phi\pi^-$ with $\phi \to K^+K^-$ (5600 events), $D_s^- \to K^{*0}K^-$ with $K^{*0} \to K^+\pi^-$ (2760 events), and $D_s^- \to \pi^-\pi^-\pi^+$ (2650 events). We

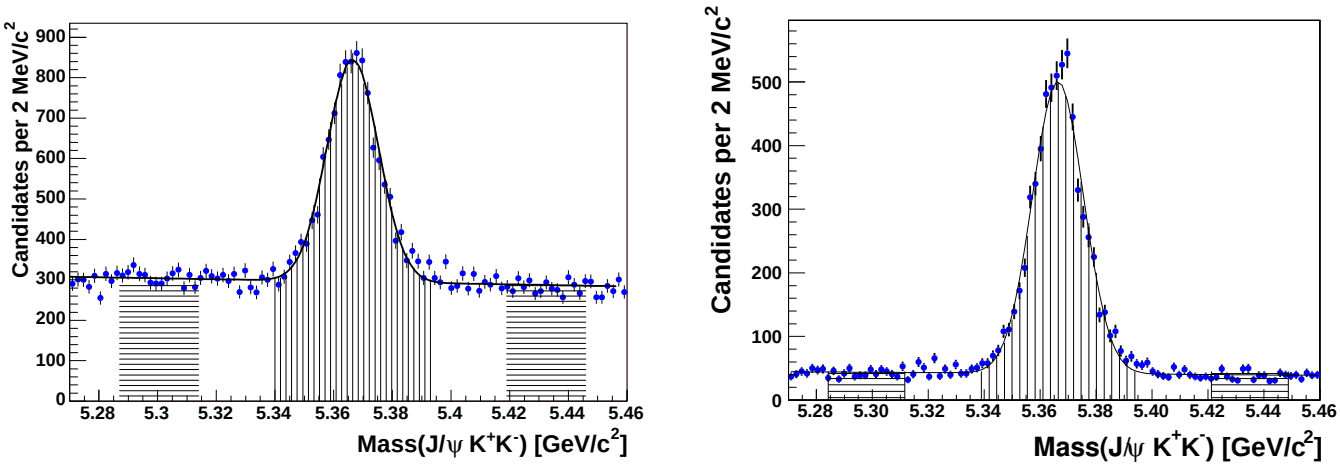

FIG. 5. $J/\psi K^+K^-$ invariant mass distribution with a cut of 0.2 on the ANN discriminant (left) and in addition with a lifetime requirement $ct > 60$ μm on the B_s^0 candidate (right). The areas with vertical (horizontal) lines indicate the signal (sideband) regions used in Sec. VI.

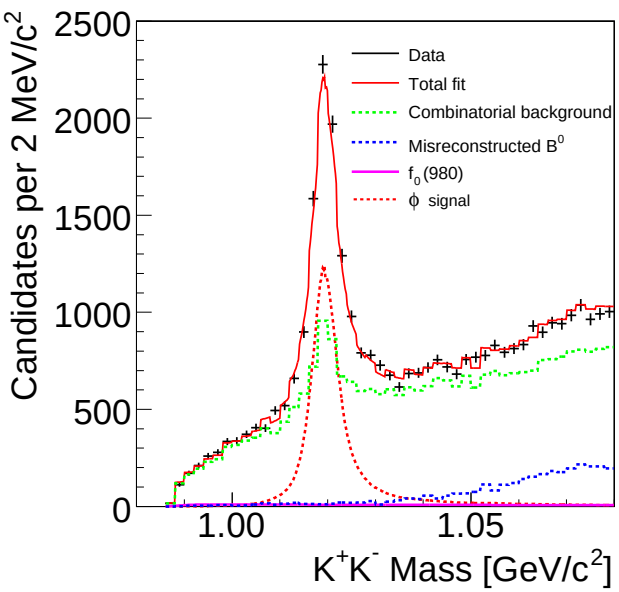

FIG. 6. (color online). K^+K^- invariant mass distribution with combinatorial background, B^0 reflection, and potential S-wave contribution.

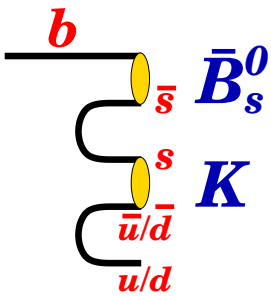

FIG. 7. Illustration of b quark fragmentation into $\bar{B}_s^0$ meson.

also include the decay mode $B_s^0 \to D_s^- \pi^+\pi^+\pi^-$, with $D_s^- \to \phi\pi^-$ and $\phi \to K^+K^-$ (1850 events). To illustrate this sample of B_s^0 candidates, the left-hand side of Fig. 8 shows the invariant mass of $B_s^0 \to D_s^- \pi^+$ candidates with $D_s^- \to \phi\pi^-$ including background contributions.

The calibration of the SSKT is achieved via an amplitude scan of the mixing frequency Δm_s. The probability for observing a B_s^0 meson in a B_s^0 or $\bar{B}_s^0$ flavor eigenstate as a function of time is

$$P(t)_{B_s^0,\bar{B}_s^0} \propto 1 \pm \mathcal{A}\,\mathcal{D}_p \cos \Delta m_s t, \qquad (10)$$

where $\mathcal{D}_p$ is the event by event predicted dilution and $\mathcal{A}$ is a Fourier-like coefficient called "amplitude". The amplitude scan consists of a series of steps in which the mixing frequency Δm_s is fixed at values between zero and 30 ps^{-1}. At each step, the likelihood function based on the above probability density function, is maximized and the best fit value of the amplitude parameter is determined. Whenever the mixing frequency is fixed to values far from the true mixing frequency, the best fit value of the amplitude parameter is consistent with zero. On the contrary, when values of Δm_s close to the true B_s^0 mixing frequency are probed, the best fit value of the amplitude parameter is inconsistent with zero. If the dilution $\mathcal{D}_p$, which is predicted on an event-by-event basis by the tagging algorithms, is correct, the amplitude $\mathcal{A}$ will be close to unity at the true value of Δm_s. Deviations from unity indicate that the predicted dilution has to be re-scaled by the actual value of the amplitude parameter at the amplitude maximum. This value of $\mathcal{A}$ is also called the dilution scale factor $\mathcal{S}_\mathcal{D}$. If the dilution scale factor is larger (smaller) than unity, the tagging algorithm under (over) estimates the predicted dilution. Multiplying the predicted SSKT dilution by $\mathcal{S}_\mathcal{D}$ will then provide on average the correct event-by-event dilution.

The result of the Δm_s amplitude scan is shown in Fig. 8. Maximizing the likelihood as a function of Δm_s

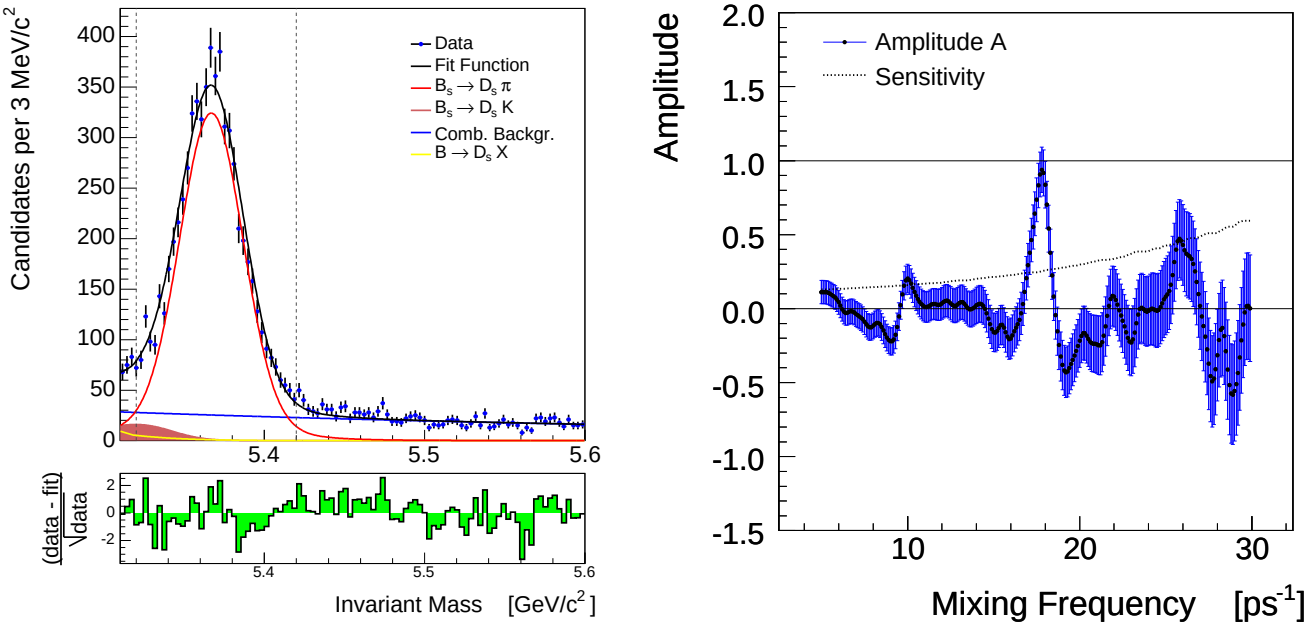

FIG. 8. (color online). Left: Invariant mass of $B_s^0 \to D_s^- \pi^+$ candidates with $D_s^- \to \phi\pi^-$ and $\phi \to K^+K^-$ including background contributions. The pull distribution at the bottom shows the difference between data and fit value normalized to the data uncertainty. Right: The measured amplitude $\mathcal{A}$ and uncertainties versus the B_s^0 oscillation frequency Δm_s for the combined set of $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^- \pi^+\pi^+\pi^-$ candidates. The sensitivity curve represents $1.645\,\sigma_{\mathcal{A}}$, where $\sigma_{\mathcal{A}} \propto e^{(\Delta m_s \sigma_t)^2/2}$ is the expected uncertainty on $\mathcal{A}$ for a given value of Δm_s. Note that the uncertainties shown are correlated between data points.

measures the frequency of B_s^0 oscillations at $\Delta m_s = 17.79 \pm 0.07\,(\text{stat})\ \text{ps}^{-1}$ in agreement with the world average value of Δm_s [8]. At this point the maximum amplitude is consistent with one and the measured dilution scale factor for the SSKT algorithm is thus consistent with unity, indicating that the initial calibration based on simulated events was accurate. We find $\mathcal{S}_{\mathcal{D}} = 0.94 \pm 0.15\,(\text{stat}) \pm 0.13\,(\text{syst})$. This is the first time the SSKT dilution was calibrated using data only. We measure a tagging efficiency of $\varepsilon = (52.2 \pm 0.7)\%$ and an average predicted dilution of $\sqrt{<\mathcal{D}_p^2>} = (27.5 \pm 0.3)\%$. The total tagging power is found to be $\varepsilon {\mathcal{S}_{\mathcal{D}}}^2 <\mathcal{D}_p^2> = (3.5 \pm 1.4)\%$.

B. Opposite side tagging

The opposite side tagger capitalizes on the fact that most b quarks produced in $p\bar{p}$ collisions originate from $b\bar{b}$ pairs. The b or $\bar{b}$ quark on the opposite side of the B_s^0 or $\bar{B}_s^0$ candidate hadronizes into a B hadron, whose flavor can be inferred using its decay products.

The OST is a combination of several algorithms: the soft muon tagger (SMT) [42], the soft electron tagger (SET) [50], and the jet charge tagger (JQT) [51]. The JQT combines all tracks from the fragmentation of the opposite-side b quark into a single jet charge measurement. The charge of the jet is determined by the momentum-weighted sum over the momenta p_T^i of all tracks in the jet $Q_{jet} = \sum_i Q^i p_T^i (1 + P_{trk}^i) / \sum_i p_T^i (1 + P_{trk}^i)$, where $Q_i = \pm 1$ is the electric charge of track and P_{trk}^i is the probability of the track being part of the b jet. Jets are reconstructed by a cone-clustering algorithm and separated into three classes, based on their probability of containing a b quark. A class 1 jet has a vertex displaced with respect to the primary $p\bar{p}$ interaction vertex, while a class 2 jet contains at least one track displaced with respect to the primary vertex. If there are no class 1 and 2 jets found, the jet with the highest transverse momentum in the event is used. These jets constitute class 3 jets which can be identified for nearly 100% of the events.

The lepton taggers, SMT and SET, utilize the charge of a muon or electron from a $b \to c\,\ell^-\bar{\nu}_\ell$ transition to determine the production flavor of the parent B meson. Several variables used to identify electrons and muons are combined with a multivariate technique into a global likelihood to select the lepton candidates used in the tagging algorithms [42, 50].

The outcomes of the three OST algorithms are combined to give a single tag decision and predicted dilution [7]. We optimize the total dilution by combining the taggers using an artificial neural network trained on data from semileptonic $B \to \ell\nu X$ decays. The neural network handles correlations between the jet charge tagger and the lepton taggers and improves the tagging per-

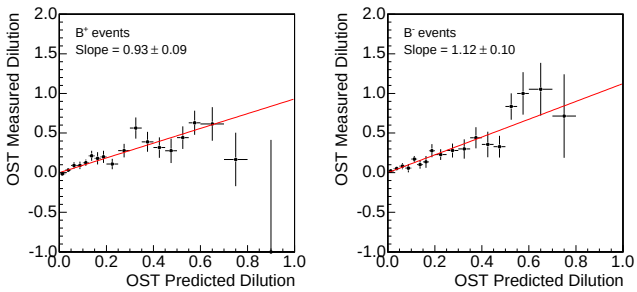

FIG. 9. Measured versus predicted dilution for B^+ (left) and B^- candidates (right) from $B^\pm \to J/\psi K^\pm$ events.

formance by 15% relative to other combinations that do not account for correlations.

We calibrate the OST performance on $B^+ \to J/\psi K^+$ decays. Since charged B mesons do not oscillate, the production flavor of the B^+ is identified by the charge of the kaon daughter track. Knowing the true flavor of the B^+ meson as well as the flavor predicted by the OST algorithm, we can easily identify the dilution measured from the number of correct and false tags as a function of the predicted dilution. This dependence is shown for B^+ and B^- mesons in Fig. 9 and fitted with a linear function whose slope is expected to be one for a perfectly functioning OST algorithm. The actual measured slope of the linear fitting function provides the OST scale factor $\mathcal{S}_\mathcal{D}$ which is determined to be 0.93 ± 0.09 (1.12 ± 0.10) for B^+ (B^-) mesons.

Although the dilution scale factors determined separately from B^+ and B^- decays are both within uncertainties consistent with unity and with each other, we use two scale factors for the opposite side tagger, one for B^+ mesons and one for B^- mesons, in order to allow for any potential asymmetry in the tagging algorithms. As a cross-check we determine the scale factors in different data taking periods and find that the scale factors are stable throughout all parts of the data. We measure a tagging efficiency of $\varepsilon = (94.2 \pm 0.4)\%$ and an average predicted dilution from the $B^\pm \to J/\psi K^\pm$ signal events of $\sqrt{<\mathcal{D}_p^2>} = (11.04 \pm 0.18)\%$. The average OST dilution scale factor is $\mathcal{S}_\mathcal{D} = 1.03 \pm 0.06$. The total OST tagging power is $\varepsilon {\mathcal{S}_\mathcal{D}}^2 <\mathcal{D}_p^2> = (1.2 \pm 0.2)\%$.

VI. THE LIKELIHOOD FUNCTION

A simultaneous unbinned maximum likelihood fit to our data including information on the $J/\psi K^+K^-$ invariant mass, B_s^0 candidate decay time, and transversity angular variable $\vec{\rho}$ is performed to extract the main parameters of interest, $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$, plus additional physics parameters, which include the B_s^0 meson mass, the mean B_s^0 width $\Gamma_s = \hbar/\tau(B_s^0)$, the polarization amplitudes in the transversity basis $|A_0(0)|^2$, $|A_\parallel(0)|^2$, and $|A_\perp(0)|^2 \equiv 1-|A_0(0)|^2-|A_\parallel(0)|^2$, the corresponding CP-conserving strong phases $\delta_\parallel$ and $\delta_\perp$, the fraction of the S-wave component and its corresponding phase δ_{SW}. The likelihood function also includes other technical parameters of less interest referred to as "nuisance parameters" such as the B_s^0 signal fraction f_s, parameters describing the $J/\psi\phi$ mass distribution, the B_s^0 decay time plus angular distributions of background events, parameters used to describe the estimated decay time uncertainty distributions for signal and background events, scale factors between the estimated decay time and mass uncertainties and their true uncertainties, as well as tagging dilution scale factors and efficiency and asymmetry parameters. There are a total of 35 fit parameters in the likelihood function, 11 of which we consider parameters of physical interest with $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$ being the main physics parameters. All the physics parameters are listed in Table II.

TABLE II. Physics parameters of interest in the likelihood fit with $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$ as main physics parameters.

Parameter	Description
$\beta_s^{J/\psi\phi}$	CP-violating phase [rad]
$\Delta\Gamma_s$	$\Gamma_s^L - \Gamma_s^H$ [ps^{-1}]
Δm_s	B_s^0-$\bar{B}_s^0$ oscillation frequency [ps^{-1}]
$\tau(B_s^0)$	B_s^0 mean lifetime [ps]
$m(B_s^0)$	B_s^0 mass [MeV/c^2]
$\|A_0(0)\|^2$	Longitudinally-polarized transition probability at $t = 0$
$\|A_\parallel(0)\|^2$	Parallel component of transversely-polarized transition probability at $t = 0$
$\delta_\perp$	$\arg[A_\perp(0)A_0^*(0)]$
$\delta_\parallel$	$\arg(A_\parallel(0)A_0^*(0)]$
f_{SW}	Fraction of S-wave in $J/\psi K^+K^-$ sample
δ_{SW}	Relative phase of S-wave contribution

The distributions of the transversity angular variables $\vec{\rho} = (\cos\theta_T, \phi_T, \cos\psi_T)$ observed with the CDF II detector are different from the true distributions because the efficiency of detecting the final state muons and kaons from J/ψ and ϕ decays is non-uniform. The angular efficiency function is parametrized in three dimensions using a set of real spherical harmonics and Legendre polynomials as basis functions with ranges $0 < \psi_T < \pi$, $0 < \theta_T < \pi$ and $0 < \phi_T < 2\pi$ [52]:

$$\varepsilon(\psi_T, \theta_T, \phi_T) = \sum_{lmk} a_{lm}^k P_k(\cos\psi_T) Y_{lm}(\theta_T, \phi_T). \quad (11)$$

The parameters a_{lm}^k are obtained from simulated events where all transversity angles are generated flat for B_s^0 decays. These MC events, which have been passed through the full CDF II detector simulation, enable us to examine how the initially flat distributions are sculpted by the detector acceptance, thus allowing us to determine the angle-dependent efficiencies of the reconstructed particle

candidates. There are no predictions for the distribution of the background transversity angles, but we find that they can be represented as the product of three independent functions of $\cos\theta_T$, ϕ_T, and $\cos\psi_T$ that are constant in time:

$$f(\cos\theta_T) = \frac{a_0 - a_1\cos^2(\theta_T)}{2a_0 - 2a_1/3},$$
$$f(\phi_T) = \frac{1 + b_1\cos(2\phi_T + b_0)}{2\pi},$$
$$f(\cos\psi_T) = \frac{c_0 + c_1\cos^2(\psi_T)}{2c_0 + 2c_1/3}. \quad (12)$$

The parameters $a_{0,1}$, $b_{0,1}$ and $c_{0,1}$ are determined as best fit estimates from the maximum likelihood optimization. The above functions follow closely the shapes of the angular efficiencies, which suggests that the underlying transversity angle distributions of the background events are flat.

To set up the full unbinned maximum likelihood fit, we define a set of probability density functions (PDFs), $P(\vec{x}\,|\,\vec{\mu})$, which give the probability density of observing the measured variables $\vec{x}_i$ for an event i, given a set of unknown parameters $\vec{\mu}$. In our likelihood function the measured variables $\vec{x}_i$ for each event i are the $J/\psi K^+K^-$ invariant mass value m and its uncertainty σ_m, the B_s^0 candidate proper decay length ct and uncertainty σ_{ct}, the angular distributions of $\vec{\rho} = (\cos\theta_T, \phi_T, \cos\psi_T)$ in the transversity basis, and the predicted dilution $\mathcal{D}_p$ and tag decision ξ for the SSKT and OST method as described in Sec. V. Among the unknown fit parameters $\vec{\mu} = (\vec{\theta}, \vec{\nu})$ are the physics parameters $\vec{\theta}$ described in Table II as well as the nuisance parameters $\vec{\nu}$ discussed above.

The likelihood function for our dataset of N events is given as

$$\mathcal{L}(\vec{x}\,|\,\vec{\theta},\vec{\nu}) = \prod_{i=1}^{N} P(\vec{x}_i\,|\,\vec{\theta},\vec{\nu}). \quad (13)$$

We minimize

$$-\log\mathcal{L}(\vec{x}\,|\,\vec{\theta},\vec{\nu}) = -\sum_{i=1}^{N}\log P(\vec{x}_i\,|\,\vec{\theta},\vec{\nu}) \quad (14)$$

using the MINUIT program package [53]. The likelihood function is composed of separate probability density functions for signal events, P_s, and for background events, P_b. Both the signal and background components contain PDFs describing the measured variables $\vec{x}_i$ of the B_s^0 candidate described above.

The full likelihood function, including flavor tagging, can be expressed for signal and background events as

$$\mathcal{L} = \prod_{i=1}^{N}\left[f_s\cdot P_s(m|\sigma_m)\cdot P_s(\xi)\cdot P_s(\theta_T,\phi_T,\psi_T,ct|\sigma_{ct},\xi,\mathcal{D}_p)\cdot P_s(\sigma_{ct})\cdot P_s(\mathcal{D}_p)\right.$$
$$\left. +\,(1-f_s)\cdot P_b(m)\cdot P_b(\xi)\cdot P_b(ct|\sigma_{ct})\cdot P_b(\theta_T)\cdot P_b(\phi_T)\cdot P_b(\psi_T)\cdot P_b(\sigma_{ct})\cdot P_b(\mathcal{D}_p)\right], \quad (15)$$

where the product runs over all N events in the data sample and f_s and $(1-f_s)$ are the fraction of signal and background events, respectively. For the case of the fit without flavor tagging, the likelihood function reduces to

$$\mathcal{L}^{\text{notag}} = \prod_{i=1}^{N}[f_s\cdot P_s(m|\sigma_m)\cdot P_s(\theta_T,\phi_T,\psi_T,ct|\sigma_{ct})\cdot P_s(\sigma_{ct})$$
$$+\,(1-f_s)\cdot P_b(m)\cdot P_b(ct|\sigma_{ct})\cdot P_b(\theta_T)\cdot P_b(\phi_T)\cdot P_b(\psi_T)\cdot P_b(\sigma_{ct})] \quad (16)$$

which simply corresponds to the flavor tagged case with no tag decision ($\xi = 0$) or a tagging dilution of zero ($\mathcal{D}_p = 0$).

In the following, we describe the individual elements of the full likelihood function in more detail starting with the signal mass PDF $P_s(m|\sigma_m)$. We model the signal mass distribution with a Gaussian distribution of variable width. To form the probability density function, $P_s(m|\sigma_m)$, we use the candidate-by-candidate observed mass uncertainty σ_m multiplied by a scale factor s_m, which is a fit parameter and accounts for a collective misestimation of the mass uncertainties. The PDF is normalized over the range $5.2 < m(J/\psi K^+K^-) < 5.6$ GeV/c^2. The background mass PDF, $P_b(m)$, is parametrized as a first order polynomial. Since the distributions of the decay time uncertainty σ_{ct} and the event-specific dilution $\mathcal{D}_p$ are observed to be different in signal and background, we include their PDFs explicitly in the likelihood. The signal PDFs $P_s(\sigma_{ct})$ and $P_s(\mathcal{D}_p)$ are determined from sideband-subtracted data distributions, while the background PDFs $P_b(\sigma_{ct})$ and $P_b(\mathcal{D}_p)$ are determined from the $J/\psi K^+K^-$ invariant mass sidebands. The PDFs of

the decay time uncertainties, $P_s(\sigma_t)$ and $P_b(\sigma_t)$, are described with a sum of normalized Gamma function distributions as described below, while the dilution PDFs $P_s(\mathcal{D})$ and $P_b(\mathcal{D})$ are included in the likelihood as histograms that have been extracted from data.

For the time and angular dependence of the signal PDF $P_s(\theta_T, \phi_T, \psi_T, ct|\sigma_{ct}, \xi, \mathcal{D}_p)$, we follow the method derived in Ref. [52]. This PDF includes the additional contribution from S-wave $B_s^0 \to J/\psi K^+K^-$ decays, with fraction f_{SW} and relative phase δ_{SW} between the S-wave and P-wave amplitude. The major difference between our treatment and that of Ref. [52] is a refinement of the model used to describe the line shape of the K^+K^- invariant mass $\mu \equiv m(K^+K^-)$ [54]. As in Ref. [52] we use a flat model for the S-wave component, whereas for the P-wave $B_s^0 \to J/\psi\phi$ component we instead use an asymmetric relativistic Breit-Wigner distribution with mass-dependent width

$$|BW(\mu)|^2 = \frac{\mu}{m_\phi} \cdot \Gamma_1 \cdot \frac{E(K^+K^-)}{E(\phi)} \cdot \frac{1}{(m_\phi^2 - \mu^2)^2 + m_\phi^2 \cdot \Gamma_{\rm tot}^2}, \tag{17}$$

where $E(\phi)$ (K^+K^-) is the energy of the ϕ (K^+K^-) in the decay of $B_s^0 \to J/\psi\phi$ $(B_s^0 \to J/\psi K^+K^-)$. This treatment assumes a two-body decay, where the other daughter particle is the J/ψ, and the total decay width $\Gamma_{\rm tot} = \Gamma_1 + \Gamma_2 + \Gamma_3$, where $\Gamma_{1,2,3}$ are the partial decay widths for the decays $\phi \to K^+K^-$ $(48.8 \pm 0.5\%)$, $\phi \to K_L^0 K_S^0$ $(34.2 \pm 0.4\%)$, and $\phi \to \rho\pi$ plus $\phi \to \pi^+\pi^-\pi^0$ $(15.32 \pm 0.32\%)$, respectively [8]. Following Ref. [55] we describe the B_s^0 decay rate as a function of the transversity angles, decay time and, in addition, the K^+K^- invariant mass. When both a ϕ component with kaons in a relative P-wave and a S-wave component are present, the amplitudes must be summed and then squared. The P-wave amplitude has a resonant structure due to the ϕ propagator, while the S-wave amplitude is flat, but can have an arbitrary phase δ_{SW} with respect to the P-wave

$$\begin{aligned}
\rho_B(\theta_T, \phi_T, \psi_T, t, \mu) &= \frac{9}{16\pi} \left| \left[\sqrt{1 - f_{SW}} BW(\mu) \mathbf{A}(t) + e^{i\delta_{SW}} \sqrt{f_{SW}} \frac{h(\mu)}{\sqrt{3}} \mathbf{B}(t) \right] \times \hat{n} \right|^2 , \\
\rho_{\bar{B}}(\theta_T, \phi_T, \psi_T, t, \mu) &= \frac{9}{16\pi} \left| \left[\sqrt{1 - f_{SW}} BW(\mu) \bar{\mathbf{A}}(t) + e^{i\delta_{SW}} \sqrt{f_{SW}} \frac{h(\mu)}{\sqrt{3}} \bar{\mathbf{B}}(\mathbf{t}) \right] \times \hat{n} \right|^2 .
\end{aligned} \tag{18}$$

In our analysis we accept events for which the reconstructed K^+K^- mass μ lies within a window $\mu_{\rm lo} = 1.009 < \mu < \mu_{\rm hi} = 1.028$ GeV/c^2. The ϕ mass distribution is described by the Breit-Wigner function given in Eq. (17). The S-wave mass distribution is given by the flat function $h(\mu) = \frac{1}{\sqrt{\Delta\mu}}$ between $\mu_{\rm lo}$ and $\mu_{\rm hi}$, where $\Delta\mu = \mu_{\rm hi} - \mu_{\rm lo}$. The likelihood function used in the maximum likelihood fit is obtained by numerically integrating Eq. (18) over the K^+K^- invariant mass μ. $\mathbf{A}(t)$ and $\bar{\mathbf{A}}(\mathbf{t})$ are time-dependent complex vector functions describing the P-wave component in the transversity basis. They are defined as

$$\begin{aligned}
\mathbf{A}(t) &= \left(\mathcal{A}_0(t) \cos\psi_T, -\frac{\mathcal{A}_\parallel(t) \sin\psi_T}{\sqrt{2}}, i\frac{\mathcal{A}_\perp(t) \sin\psi_T}{\sqrt{2}} \right), \\
\bar{\mathbf{A}}(t) &= \left(\bar{\mathcal{A}}_0(t) \cos\psi_T, -\frac{\bar{\mathcal{A}}_\parallel(t) \sin\psi_T}{\sqrt{2}}, i\frac{\bar{\mathcal{A}}_\perp(t) \sin\psi_T}{\sqrt{2}} \right),
\end{aligned}$$

with

$$\begin{aligned}
\mathcal{A}_i(t) &= \frac{e^{-\Gamma_s t/2}}{\sqrt{\tau_H + \tau_L \pm \cos 2\beta_s (\tau_L - \tau_H)}} \\
&\quad \times \left[E_+(t) \pm e^{2i\beta_s} E_-(t) \right] a_i , \\
\bar{\mathcal{A}}_i(t) &= \frac{e^{-\Gamma_s t/2}}{\sqrt{\tau_H + \tau_L \pm \cos 2\beta_s (\tau_L - \tau_H)}} \\
&\quad \times \left[\pm E_+(t) + e^{-2i\beta_s} E_-(t) \right] a_i ,
\end{aligned} \tag{19}$$

where $i \in \{0, \parallel, \perp\}$ and the upper sign applies to the CP-even final states (0 and $\parallel$), while the lower sign applies to the CP-odd final state ($\perp$). Furthermore,

$$E_\pm(t) \equiv \frac{1}{2} \left[e^{+\left(\frac{-\Delta\Gamma_s}{4} + i\frac{\Delta m_s}{2}\right)t} \pm e^{-\left(\frac{-\Delta\Gamma_s}{4} + i\frac{\Delta m_s}{2}\right)t} \right], \tag{20}$$

and the a_i are complex amplitude parameters satisfying

$$\sum_i |a_i|^2 = 1 . \tag{21}$$

The S-wave component is described in the transversity basis as

$$\begin{aligned}
\mathbf{B}(t) &= (\mathcal{B}(t), 0, 0) , \\
\bar{\mathbf{B}}(t) &= \left(\bar{\mathcal{B}}(t), 0, 0\right) ,
\end{aligned} \tag{22}$$

where the time-dependent amplitudes are

$$\begin{aligned}\mathcal{B}(t) &= \frac{e^{-\Gamma_s t/2}}{\sqrt{\tau_H + \tau_L - \cos 2\beta_s\,(\tau_L - \tau_H)}} \\ &\quad \times \left[E_+(t) - e^{2i\beta_s} E_-(t)\right], \\ \bar{\mathcal{B}}(t) &= \frac{e^{-\Gamma_s t/2}}{\sqrt{\tau_H + \tau_L - \cos 2\beta_s\,(\tau_L - \tau_H)}} \\ &\quad \times \left[-E_+(t) + e^{-2i\beta_s} E_-(t)\right]. \end{aligned} \tag{23}$$

In Eq. (18) the unit vector $\hat{n}$ is defined as $\hat{n} = (\sin\theta_T \cos\phi_T, \sin\theta_T \sin\phi_T, \cos\theta_T)$ in the transversity basis and the strong phases $\delta_\parallel$ and $\delta_\perp$ appear from terms of the form $A_i A_0^* = |A_i||A_0|e^{\arg(A_i A_0^*)}$, where $i = '\parallel'$ or $i = '\perp'$. The decay width difference $\Delta\Gamma_s$ and the B_s^0 oscillation frequency Δm_s are encoded in Eq. (20).

The time-dependent functions in the PDFs above are convolved with a resolution function composed of two independent Gaussians with candidate-by-candidate expected width σ_{ct}. The widths of each Gaussian function are multiplied by independent scale factors s_{ct1} and s_{ct2} which are freely floated in the maximum likelihood fit to account for an overall mis-estimation of the decay-time resolution. The PDF describing the decay time distributions for signal events as part of $P_s(\theta_T, \phi_T, \psi_T, ct|\sigma_{ct}, \xi, \mathcal{D}_p)$ in Eq. (15) is of the form

$$\begin{aligned} P_s(ct|\sigma_{ct}) &= P_s(ct, \sigma_{ct}|c\tau, s_{ct1,2}) \\ &= F(ct, c\tau) \otimes G(ct, \sigma_{ct}|f_{s_{ct1}}, s_{ct1}, s_{ct2}), \end{aligned} \tag{24}$$

where $\tau = \tau(B_s^0)$ and $F(ct, c\tau)$ represents the time dependence of the signal events which, e.g., for an exponential decay is given as $e^{-\frac{ct}{c\tau}}/(c\tau)$. The symbol "$\otimes$" denotes a convolution which is with respect to the decay time resolution function defined as

$$G(ct, \sigma_{ct}|f_{s_{ct1}}, s_{ct1}, s_{ct2}) = f_{s_{ct1}} \frac{1}{\sqrt{2\pi} s_{ct1}\sigma_{ct}} e^{-\frac{c^2t^2}{2s_{ct1}^2\sigma_{ct}^2}} + (1 - f_{s_{ct1}}) \frac{1}{\sqrt{2\pi} s_{ct2}\sigma_{ct}} e^{-\frac{c^2t^2}{2s_{ct2}^2\sigma_{ct}^2}}, \tag{25}$$

where $f_{s_{ct1}}$ is the fraction of the first resolution Gaussian. From the distribution of decay time uncertainties, we find the average of the decay time resolution function at $\sigma_{ct} \sim 30$ μm, with a root-mean-square deviation of about 12 μm [15].

As we are using candidate-by-candidate expected decay time uncertainties, which are not distributed identically for the signal and background events, it is necessary to include a PDF for the separate uncertainty distributions [56]. The PDF describing the decay time uncertainty $P_s(\sigma_{ct})$ is constructed from normalized Gamma distributions

$$\Gamma_{\rm norm}(x) \equiv \frac{x^a e^{-x/b}}{b^{a+1}\Gamma(a+1)}, \tag{26}$$

where a and b define the mean and width of the distribution. Each function has different values of a and b. We find these values from a separate lifetime-only fit before running the full angular analysis and fix these parameters within uncertainties in the full likelihood fit. We handle the background lifetime resolution PDF $P_b(\sigma_{ct})$ in the same way as the signal distribution, using a sum of three normalized Gamma distributions $\Gamma_{\rm norm}$.

The portion of the likelihood function $P_s(\theta_T, \phi_T, \psi_T, ct|\sigma_{ct}, \xi, \mathcal{D}_p)$ related to tagging can be written as follows:

$$\begin{aligned} P_s(ct, \psi_T, \theta_T, \phi_T|\sigma_{ct}, \mathcal{D}_{SS}, \mathcal{D}_{OS}, \xi_{SS}, \xi_{OS}) &= \frac{1 + \xi_{SS} s_{SS} \mathcal{D}_{SS}}{1 + |\xi_{SS}|} \frac{1 + \xi_{OS} s_{OS} \mathcal{D}_{OS}}{1 + |\xi_{OS}|} \rho_B(ct, \psi_T, \theta_T, \phi_T) \otimes G(ct, \sigma_{ct}) \\ &+ \frac{1 - \xi_{SS} s_{SS} \mathcal{D}_{SS}}{1 + |\xi_{SS}|} \frac{1 - \xi_{OS} s_{OS} \mathcal{D}_{OS}}{1 + |\xi_{OS}|} \rho_{\bar{B}}(ct, \psi_T, \theta_T, \phi_T) \otimes G(ct, \sigma_{ct}), \end{aligned} \tag{27}$$

where s_{SS} and s_{OS} are the SSKT and OST dilution scale factors. Since the two flavor taggers search for tagging information (tracks, jets) in complementary regions in space, we treat them as independent tags. We have verified that the tagging decisions and predicted dilutions are indeed independent.

The probability of a particular combined tag decision is dependent on the efficiency of each of the taggers

$$P(\xi) = \begin{cases} (1-\varepsilon_{SS})(1-\varepsilon_{OS}) & (\xi_{SS} = 0,\ \xi_{OS} = 0) \\ \varepsilon_{SS}(1-\varepsilon_{OS}) & (\xi_{SS} = \pm 1,\ \xi_{OS} = 0) \\ (1-\varepsilon_{SS})\varepsilon_{OS} & (\xi_{SS} = 0,\ \xi_{OS} = \pm 1) \\ \varepsilon_{SS}\varepsilon_{OS} & (\xi_{SS} = \pm 1,\ \xi_{OS} = \pm 1). \end{cases} \tag{28}$$

The predicted dilution distributions are different for signal and background events. Therefore the likelihood

function in Eq. (15) contains the corresponding dilution PDFs separately for signal, $P_s(\mathcal{D}_p)$, and for background events, $P_b(\mathcal{D}_p)$. These dilution distributions are measured with data. In the case of the signal, they are the sideband-subtracted distributions taken from the B_s^0 invariant mass signal region and stored as histograms. The histograms are normalized to represent probability densities for the dilution.

Knowledge of the fractions of positively and negatively charged events is sufficient to describe any tagging asymmetry present in the background. Therefore, the probability density $P_b(\xi)$ is equal to the fraction of events with positive tags for $\xi = +1$, and equal to the fraction of events with negative tags for $\xi = -1$. The background dilution distributions are handled analogously to the signal dilution. The dilution distributions are taken from the B_s^0 invariant mass sideband region, normalized to form a probability density, and again stored as histograms.

The background proper decay time PDF, $P_b(ct|\sigma_{ct})$ is parametrized as a prompt peak modeled by a δ-function plus two positive exponentials and one negative exponential. This function is convolved with the same resolution function as the signal decay time dependence. In our parameterization, the prompt peak models the majority of the combinatorial background events, which are expected to have no significant lifetime, the negative exponentials defined for $t > 0$ account for a small fraction of longer lived background such as other B hadron decays, and the positive exponential defined for $t < 0$ takes into account events with a mis-reconstructed vertex. The background decay time PDF reads

$$P_b\left(ct|\sigma_{ct}\right) = \left\{ f_g\,\delta(ct) + (1-f_g)\left[f_{++}\frac{e^{-\frac{ct}{\lambda_{++}}}}{\lambda_{++}} + (1-f_{++})\left(f_{-}\frac{e^{\frac{ct}{\lambda_{-}}}}{\lambda_{-}} + (1-f_{-})\frac{e^{-\frac{ct}{\lambda_{+}}}}{\lambda_{+}}\right)\right]\right\} \otimes G(ct,\sigma_{ct}|f_{s_{ct1}},s_{ct1},s_{ct2}), \tag{29}$$

where f_g is the fraction of the prompt background, λ_{++}, λ_+ and λ_- are the effective lifetimes of the background events distributed according to the long and short lived positive exponential as well as the negative exponential, respectively, while f_{++} and f_- are their corresponding fractions.

VII. B_S^0 MEAN LIFETIME, DECAY WIDTH DIFFERENCE AND POLARIZATION FRACTIONS

We use the likelihood function presented in the previous section to extract measurements of the physics parameters of interest. Before applying the unbinned maximum likelihood fit on data, we perform an extensive set of tests of the fitting procedure using simulated pseudoexperiments. We observe that, with the current statistics, the maximized likelihood function returns biased results for the parameters of interest. Moreover, the likelihood function shows non-Gaussian behavior with respect to the $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$ parameters. For these reasons we employ frequentist techniques to determine confidence level (C.L.) regions in the $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ plane as described in Sec. VIII.

However, we find that the likelihood function constructed according to the standard model expectation, in which the CP-violation parameter $\beta_s^{J/\psi\phi}$ is fixed to a value very close to zero ($\beta_s^{J/\psi\phi} = 0.02$), returns unbiased results for the mean B_s^0 lifetime $\tau(B_s^0)$, the decay width difference $\Delta\Gamma_s$, the polarization fractions $|A_\parallel(0)|^2$ and $|A_0(0)|^2$ as well as the strong phase $\delta_\perp$. We also observe that the likelihood function is Gaussian with respect to all these parameters. Under these favorable circumstances we provide point estimates with Gaussian uncertainties for the following quantities:

$$\begin{aligned} c\tau(B_s^0) &= 458.6 \pm 7.6\,(\text{stat}) \pm 3.6\,(\text{syst})\ \mu\text{m}, \\ \Delta\Gamma_s &= 0.075 \pm 0.035\,(\text{stat}) \pm 0.006\,(\text{syst})\ \text{ps}^{-1}, \\ |A_\parallel(0)|^2 &= 0.231 \pm 0.014\,(\text{stat}) \pm 0.015\,(\text{syst}), \\ |A_0(0)|^2 &= 0.524 \pm 0.013\,(\text{stat}) \pm 0.015\,(\text{syst}), \\ \delta_\perp &= 2.95 \pm 0.64\,(\text{stat}) \pm 0.07\,(\text{syst})\ \text{rad}. \end{aligned} \tag{30}$$

The systematic uncertainties on the measured quantities are discussed in detail in Sec. VII A and given above for completeness. We are unable to also quote a result for $\delta_\parallel$ since the fit prefers a value of $\delta_\parallel$ at the boundary of π resulting in a non-Gaussian likelihood shape around the minimum.

The results in Eq. 30 show good agreement with previous measurements [15, 16]. The B_s^0 mean lifetime can be calculated as $\tau(B_s^0) = 1.529 \pm 0.025\,(\text{stat}) \pm 0.012\,(\text{syst})$ ps, which is the most precise single measurement of this quantity. It can be compared to the most recent measurement from the D0 collaboration using a data sample based on 8 fb^{-1} of integrated luminosity [19] quoting $\tau(B_s^0) = 1.443^{+0.038}_{-0.035}$ ps and to the Particle Data Group (PDG) average of $\tau(B_s^0) = 1.472^{+0.024}_{-0.026}$ ps [8]. The $\Delta\Gamma_s$ value is of comparable precision to the current world average of $\Delta\Gamma_s = 0.062^{+0.034}_{-0.037}$ ps^{-1} [8]. Our central value is somewhat smaller than the most recent measurement of $\Delta\Gamma_s = 0.163^{+0.065}_{-0.064}$ ps^{-1} from the D0 collaboration [19] but compares well to the PDG average as well as the SM prediction $\Delta\Gamma_s = 0.090 \pm 0.024$ ps^{-1} [9].

In addition to comparing the fit results with predictions and other measurements, three cross-checks are performed using alternative versions of the fit. First we use the likelihood without flavor tagging to check for any bias which could be introduced by the tagging. The fit without flavor tagging does not have sensitivity to $\delta_\perp$. This quantity is thus omitted from Table III for fits without flavor tagging. The second and third checks are done without the S-wave component included in the likelihood fit, once with flavor tagging included and once without. These checks are made to determine whether the S-wave component has a significant effect on the fit results and to provide a direct comparison with previous CDF results [15] which did not account for the S-wave component. Table III shows the results of these cross-checks, that demonstrate good agreement between different versions of the fit and our main results which include both flavor tagging and the S-wave component.

Since the unbinned maximum likelihood method does not readily provide a goodness of fit estimator, we present fit projections onto the data to support the quality of the fit. The likelihood function, in which all parameters are fixed to their best-fit values, is overlaid on top of data distributions. Such projections are performed for both signal and background events and separately in the subspaces of the B_s^0 decay time, decay time expected uncertainty and transversity angles. The fit projections for the proper decay time and proper decay time uncertainties are shown in Figs. 10 and 11. Fit projections for the transversity angles $\cos\theta_T$, ϕ_T, and $\cos\psi_T$ from the sideband-subtracted signal region and the background (sideband) region are shown in Figs. 12 and 13, respectively. The good agreement between the data and fit projections validates our parameterization of both the signal and background distributions plus their uncertainties.

A. Systematic uncertainties

To assess systematic uncertainties on quantities of interest other than $\beta_s^{J/\psi\phi}$, namely $\Delta\Gamma_s$, the B_s^0 mean lifetime, the polarization fractions, and the strong phase $\delta_\perp$, we set the CP-violating phase to its standard model expectation $\beta_s^{J/\psi\phi} = 0.02$. This choice in addition improves the statistical behavior of the likelihood function by eliminating biases on these parameters.

Systematic uncertainties are assigned by considering several effects that are not accounted for in the likelihood fit [57]. Such effects include potential misparameterization in the fit model, impact of particular assumptions in the fit model, and physical effects which are not well known or fully incorporated into the model. To estimate the size of the systematic uncertainties, we generate two sets of pseudoexperiments by extracting random numbers distributed according to our PDF: one set with each of the considered systematic variations and another set of default pseudoexperiments. Each pair of modified and not modified pseudoexperiments are generated with the same random seeds. The unbinned likelihood function is maximized over the modified pseudoexperiments as well as over the corresponding default ones. For each systematic effect, the associated uncertainty is the difference between the mean of the best fit value for the pseudoexperiments with the systematic alteration included, and the equivalent mean value for the reference set of pseudoexperiments generated with the default model. The individual systematic uncertainties are summed in quadrature and presented in Table IV to give the total contribution to the uncertainties for each parameter which are due to sources of systematic uncertainty.

One source of systematic uncertainty is the modeling of the angular efficiency of the detector described in Sec. VI. We model the detector efficiency with a linear combination of Legendre polynomials and spherical harmonics as described in detail in Ref. [52]. The expansion coefficients of these functions are obtained by fitting a three-dimensional efficiency distribution obtained using simulated events. This simulated sample is re-weighted to match the p_T distributions observed in data. If the modeling is inaccurate, or the p_T re-weighting incorrect, a systematic uncertainty could be introduced. We test these effects separately. The former effect is investigated by using the default fit model on pseudoexperiments, generated with angular efficiencies taken directly from the background angular distributions rather than the default parameterization. The latter effect is investigated by generating pseudoexperiments with non-reweighted MC events as input for the angular efficiencies. The second test is a rather extreme case, but shows only a small systematic effect.

The next systematic uncertainty that we consider is the modeling of the signal B_s^0 mass distribution, which is fitted by default with a single Gaussian distribution. If a double Gaussian model is used, the fit quality would be comparable and we evaluate a systematic uncertainty from the comparison of the two fit models. To test the size of a potential systematic effect, we generate pseudoexperiments with a double-Gaussian signal mass model, extracted from data, and fit with the usual single-Gaussian parameterization.

Similarly, the model used for the mass distribution of combinatorial background events can be changed to a second order polynomial with comparable fit quality. To study this effect, we generate pseudoexperiments with a second-order polynomial background model instead of the default first order polynomial and fit with the default straight line. The results of both tests and the corresponding systematic uncertainties are listed in Table IV.

A particularly important effect to consider for the lifetime measurement is the lifetime resolution model. In our standard fit we model the detector resolution by convolving each lifetime component with a two-Gaussian resolution function. To test the effect of a misparameterization, we generate pseudoexperiments with

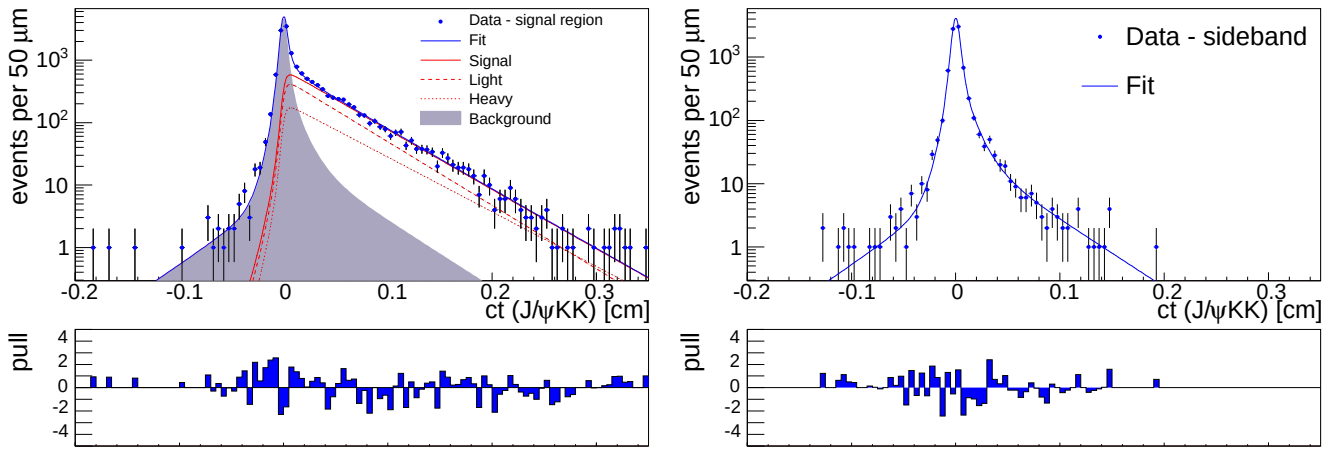

FIG. 10. (color online). Proper decay time fit projections for the B_s^0 signal (left) and background (right) regions. The dashed distributions labeled as "Light" and "Heavy" indicate the contribution of the B_s^L and B_s^H, respectively. The pull distributions at the bottom show the difference between data and fit value normalized to the data uncertainty.

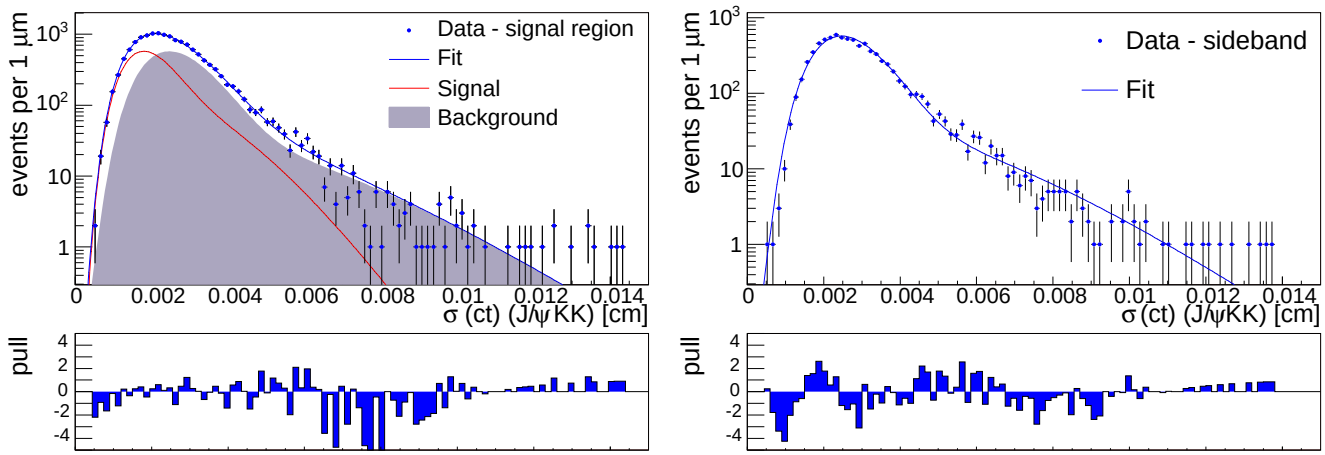

FIG. 11. (color online). Proper decay time uncertainty fit projections for the B_s^0 signal (left) and background (right) regions. The pull distributions at the bottom show the difference between data and fit value normalized to the data uncertainty.

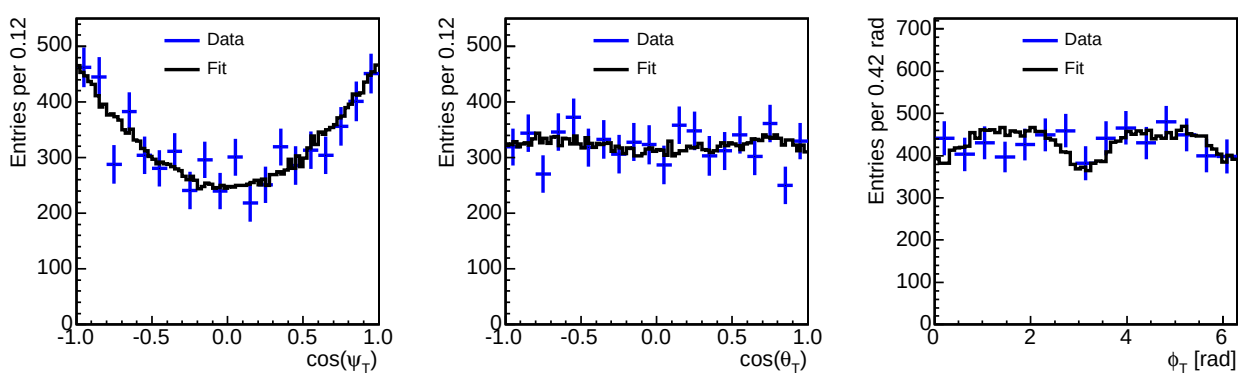

FIG. 12. Fit projections for transversity angles for sideband-subtracted signal.

TABLE III. Results of alternative fits to cross-check the main SM results. All uncertainties quoted are statistical only. As shown in Ref. [22] the untagged analysis is not sensitive to the strong phase $\delta_\perp$.

Parameter	Un-tagged (with S-wave)	Tagged (without S-wave)	Un-tagged (without S-wave)
$c\tau(B_s^0)$ [μm]	456.9 ± 8.0	459.1 ± 7.7	457.2 ± 7.9
$\Delta\Gamma_s$ [ps^{-1}]	0.069 ± 0.030	0.073 ± 0.030	0.070 ± 0.040
$\lvert A_\parallel(0)\rvert^2$	0.232 ± 0.032	0.232 ± 0.014	0.233 ± 0.016
$\lvert A_0(0)\rvert^2$	0.521 ± 0.013	0.523 ± 0.012	0.520 ± 0.013
$\delta_\perp$ [rad]	–	2.8 ± 0.6	–

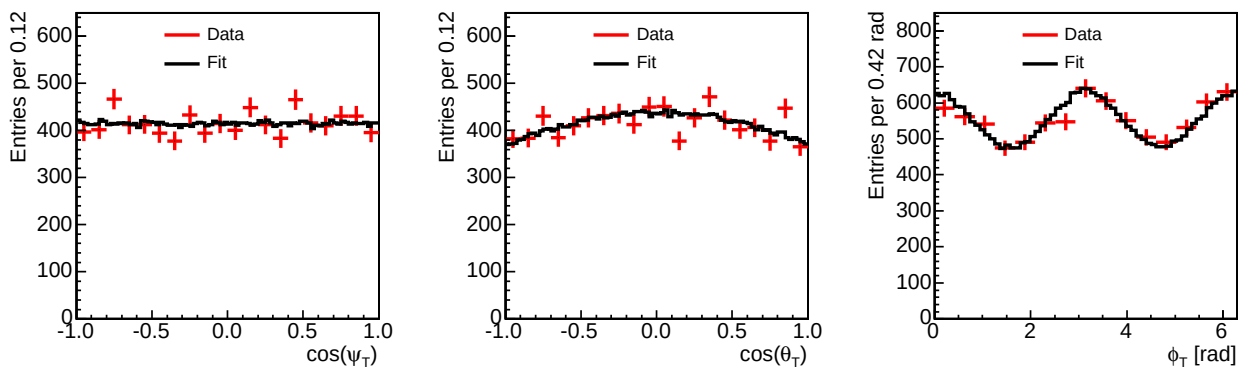

FIG. 13. Fit projections for transversity angles in background region.

TABLE IV. Summary of systematic uncertainties assigned to the five physics quantities discussed in Sec. VII.

Source of systematic effect	$\Delta\Gamma_s$ [ps^{-1}]	$c\tau(B_s^0)$ [μm]	$\lvert A_\parallel(0)\rvert^2$	$\lvert A_0(0)\rvert^2$	$\delta_\perp$ [rad]
Signal efficiency					
Parameterization	0.0024	0.96	0.0076	0.008	0.016
MC re-weighting	0.0008	0.94	0.0129	0.0129	0.022
Signal mass model	0.0013	0.26	0.0009	0.0011	0.009
Background mass model	0.0009	1.4	0.0004	0.0005	0.004
Resolution model	0.0004	0.69	0.0002	0.0003	0.022
Background lifetime model	0.0036	2.0	0.0007	0.0011	0.058
Background angular distribution					
Parameterization	0.0002	0.02	0.0001	0.0001	0.001
σ_{ct} correlation	0.0002	0.14	0.0007	0.0007	0.006
Non-factorization	0.0001	0.06	0.0004	0.0004	0.003
$B^0 \to J/\psi K^*$ cross-feed	0.0014	0.24	0.0007	0.0010	0.006
SVX alignment	0.0006	2.0	0.0001	0.0001	0.020
Mass resolution	0.0001	0.58	0.0004	0.0004	0.002
σ_{ct} modeling	0.0012	0.17	0.0005	0.0007	0.013
Pull bias	0.0028		0.0013	0.0021	
Totals	0.006	3.6	0.015	0.015	0.07

a three-Gaussian resolution model, extracted from data, and fit with the default two-Gaussian model.

As well as the lifetime resolution, the modeling of the various components of the background lifetime can systematically affect the measured B_s^0 lifetime. To check the effect of any inaccuracy in our background lifetime model as described in Section VI, we generate pseudoexperiments with the decay time of the background events taken from histograms of the B_s^0 mass sidebands and fit with the default model.

We consider three possible sources of systematic uncertainty related to the transversity angles of the background events: mis-modeling of the parameterization described in Section VI, ignoring the observed small correlations between the three angles, and correlations between the angles and the expected proper decay time uncertainty, σ_{ct}. The effect of these sources of uncertainty is checked using the actual data distributions from the mass sidebands to generate pseudoexperiments and test the difference between our model and the true distributions. For the parameterization check, we simply use the data background angular distributions in the generation of the pseudoexperiments before fitting with the default model. To check the effect of neglecting the small correlations between the angles, we generate pseudoexperiments where two of the background angles are sampled randomly from the data distributions, and the third one from a two-dimensional histogram according to the sampled value of the second angle. To check the effect of ignoring correlations between the transversity angles and σ_{ct}, we sample the ϕ_T angle distribution, found to have the largest correlation with σ_{ct}, using a two-dimensional histogram of ϕ_T versus σ_{ct} in order to generate the pseudoexperiments. The effect of ignoring these very small correlations results in an almost negligible systematic uncertainty on the measurements (see Table IV).

In the default fit, we do not account for contamination from $B^0 \to J/\psi K^{*0}$ events mis-reconstructed as $B_s^0 \to J/\psi\phi$ decays (B^0 cross-feed). A small fraction of these events lies in the B_s^0 mass signal region. The first step in identifying the size of the systematic effect is to estimate the size of this contribution by using measured production fractions of the B_s^0 and B^0 mesons, their relative decay rates to $J/\psi\phi$ and $J/\psi K^{*0}$, respectively, and the probability for each type of event to pass our final selection criteria when reconstructed under the $B_s^0 \to J/\psi\phi$ hypothesis. Both the production fractions and the branching fractions are taken from Ref. [8]. We estimate the efficiencies using simulation, with both $B_s^0 \to J/\psi\phi$ and $B^0 \to J/\psi K^{*0}$ modes reconstructed as $B_s^0 \to J/\psi\phi$ decay. The fraction f of B^0 cross-feed events in the B_s^0 sample is calculated as

$$f(B^0 \text{ in } B_s^0 \text{ sample}) = \\ = \frac{f(\bar{b} \to B^0)\,\mathcal{B}(B^0 \to J/\psi K^{*0})\,\epsilon(B^0)}{f(\bar{b} \to B_s^0)\,\mathcal{B}(B_s^0 \to J/\psi\phi)\,\epsilon(B_s^0)}. \quad (31)$$

Using Eq. (31), we find that the fraction of B^0 cross-feed into the signal sample of this analysis is $(1.6 \pm 0.6)\%$. To make a conservative estimate of the systematic uncertainty that this effect will add to the measurement of the parameters of interest, we generate pseudoexperiments with a fraction of 2.2% B^0 cross-feed, and fit with the default model which does not account for this component. The cross-feed component is generated using values of the B^0 lifetime, decay width and transversity amplitudes from the CDF angular analysis of $B^0 \to J/\psi K^{*0}$ [58].

A systematic uncertainty can be introduced by the assumption that the silicon detector is perfectly aligned, when it could actually be mis-aligned by bowing of the detector layers of up to 50 μm. A study on the effect of the limited knowledge of the CDF II silicon detector alignment concluded that a conservative estimation of the systematic uncertainty on the decay length $c\tau$ in CDF lifetime measurements is given by a 2 μm systematic uncertainty on $c\tau$ [59]. This study was done by fully reconstructing both data and simulation under different silicon alignment assumptions, including shifts of ± 50 μm in all silicon detector components. The lifetime was fitted in several $B \to J/\psi X$ channels, and the worst shift was taken as the systematic uncertainty on the lifetime due to the assumption of perfect silicon alignment.

We use the value of 2 μm systematic uncertainty on $c\tau(B_s^0)$ to also assess secondary effects on the other parameters of interest. Due to correlations between the B_s^0 lifetime and the other physics parameters, it is expected that an additional uncertainty on the lifetime measurement will also cause uncertainties in the measurement of the other parameters. To quantify the effect on the other parameters, we generate pseudoexperiments in which the decay time in each event is randomly shifted ± 2 μm and fit in the usual manner to allow for comparisons between the input and fitted values of the parameters of interest.

In the fit model, we treat the mass resolution identically for signal and background events. The effect of any inaccuracy in this assumption can be tested by generating pseudoexperiments with mass uncertainty distributions modeled by histograms of B_s^0 sideband data for background events and sideband-subtracted signal region data for signal events separately, and then fitted with the default model.

Finally, we consider the effect of mis-parameterization of the σ_{ct} distributions. To account for a possible effect, we fit with the default model to pseudoexperiments generated with the uncertainty distributions taken from data histograms rather than the model described in Section VI. This systematic check also accounts for any effect caused by small observed correlations between σ_{ct} and the invariant mass by sampling the background uncertainties from separate upper and lower sideband histograms according to the generated B_s^0 mass.

The total systematic uncertainty assigned to $\Delta\Gamma_s$ is 0.058 ps^{-1}, while 3.6 μm is assigned to the measurement of $c\tau(B_s^0)$. Both $|A_\parallel(0)|^2$ and $|A_0(0)|^2$ are assigned an uncertainty of 0.015. Finally, $\delta_\perp$ is assigned a 0.07 rad uncertainty.

VIII. FREQUENTIST ANALYSIS OF $\beta_S^{J/\psi\phi}$ AND $\Delta\Gamma_S$

As anticipated in Sec. VII, due to the pathological behavior of the likelihood function with respect to $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$, we perform a frequentist analysis to determine confidence regions for these parameters. In addition, we perform a cross-check in which we determine credible intervals for $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$ using Bayesian techniques described in Sec. IX. In the main analysis we use profile-likelihood ratio ordering [60] to determine the confidence level region in the $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ space. The coverage of the confidence region against deviations of the nuisance parameters from their measured values is confirmed by explicitly checking the effect of variations of the nuisance parameters on the profile-likelihood shape. Following our previous publication [15], the confidence regions are corrected to guarantee a coverage of the true value with at least the nominal confidence level. We thus choose this method to quote the main results of this paper.

The likelihood function has several symmetries that are discussed in detail in Ref. [52]. In absence of the S-wave component in $B_s^0 \to J/\psi K^+K^-$ decays, the likelihood function is symmetric under the simultaneous transformations: $\beta_s^{J/\psi\phi} \to \frac{\pi}{2} - \beta_s^{J/\psi\phi}$, $\Delta\Gamma_s \to -\Delta\Gamma_s$, $\delta_\perp \to \pi - \delta_\perp$, and $\delta_\parallel \to -\delta_\parallel$. An approximate symmetry is also present under the above simultaneous transformations of only $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$. The approximate symmetry produces a local minimum in the $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ space in addition to the global minimum. We account for this effect by performing the likelihood scan with $\delta_\parallel$ being started in the fit separately in the range $[0, \pi]$ and then in the range $[\pi, 2\pi]$. At each point in the $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ plane, we choose the deeper of the two $-2\log\mathcal{L}$ likelihood values (absolute minimum) as evaluated for the different $\delta_\parallel$ ranges. This procedure guarantees that we use the global, not the local minimum, at each point.

Once we have minimized the likelihood function on data with respect to the $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$ parameters, we proceed with determining the 68% and 95% confidence regions. Constructing correct and informative confidence regions from highly multi-dimensional likelihoods is challenging and, as in our case, evaluating the full 35-dimensional confidence space is computationally prohibitive. To construct a proper coverage adjustment, which ensures that the quoted 68% (95%) confidence levels do indeed contain the true values of $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$ at least 68% (95%) of the time, the choice of the ordering algorithm is important. We choose the profile-likelihood ratio ordering method [60] described below. The obtained profile-likelihood ratio is then used as a χ^2 variable to derive confidence regions in the two-dimensional space of $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$. However, simulations show that the observed profile-likelihood ratio deviates from a true χ^2 distribution. In particular, the resulting confidence regions contain true values of the parameters of interest with lower probability than the nominal confidence level and, in addition, the profile-likelihood ratio appears to depend on the true values of the nuisance parameters, which are unknown. We therefore use a large number of pseudo-experiments to derive the actual profile-likelihood ratio distribution relevant for our data. The effect of systematic uncertainties is accounted for by randomly sampling a limited number of points in the space of all nuisance parameters and using the most conservative of the resulting profile-likelihood ratio distributions to calculate the final confidence region. In the following, the coverage adjustment procedure is described in detail.

A. Coverage adjustment

To construct coverage adjusted confidence level regions for each point in the $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ plane, we start with calculating a p-value, which, given a certain hypothesis, describes the probability to observe data as discrepant or more discrepant than the data observed in our experiment. The set of all points with a p-value larger than $1-x$ forms the $x\,\%$ C.L. region. In particular, the set of points with p-value larger than $1-0.95=0.05$ outlines the 95% confidence region.

Since a main goal of our analysis is to determine the compatibility of our data with the standard model expectation for $\beta_s^{J/\psi\phi}$, we start by calculating the SM p-value. We generate pseudoexperiments at the standard model expected point in the $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ plane ($\beta_s^{J/\psi\phi} = 0.02$, $\Delta\Gamma_s = 0.090$ ps^{-1}). When generating the pseudoexperiments, we use the best fit values for all nuisance parameters as observed in our data while $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$ are fixed to the SM expected values. The likelihood function corresponding to each pseudoexperiment is first maximized with all parameters floating, and then maximized a second time with $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$ fixed to their SM values while the remaining fit parameters (nuisance parameters) are independently floating. We then form twice the negative difference between the logarithms of the likelihood values obtained in each of the two steps to obtain a profile-likelihood ratio value $-2\Delta\log\mathcal{L}$. The profile-likelihood ratio distribution from 1000 pseudoexperiments is used to obtain the standard model p-value, which is the fraction of pseudoexperiments with $-2\Delta\log\mathcal{L}$ larger than the corresponding quantity observed in data.

We construct the cumulative distribution of $-2\Delta\log\mathcal{L}$ to obtain a mapping between the p-value $= 1-$ C.L. and $-2\Delta\log\mathcal{L}$, as shown in Fig. 14 by the solid black histogram which has been interpolated. In an ideal situation, when the likelihood function is Gaussian with respect to $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$, this dependence should be a χ^2 distribution with two degrees of freedom as indicated by the green line. It is evident from Fig. 14 that, at least with our current data sample size, we are not in an asymptotic, Gaussian regime. To test the dependence of

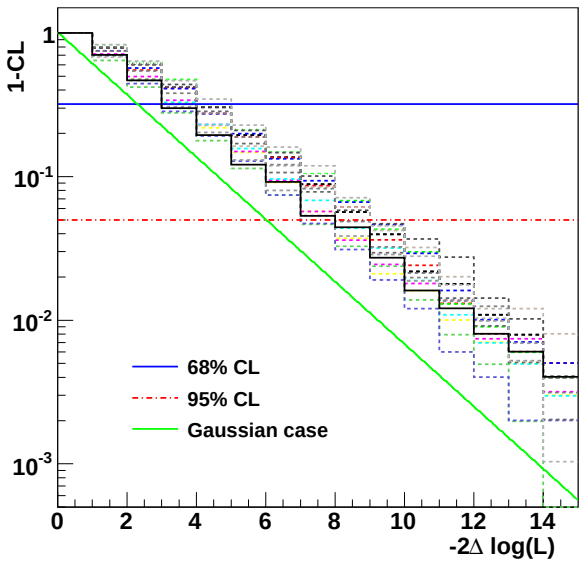

FIG. 14. (color online). Mapping of p-value (1 − C.L.) as a function of twice the negative difference of log-likelihoods ($-2\Delta \log \mathcal{L}$) as evaluated in pseudoexperiments. The ideal dependence is a χ^2 distribution with two degrees of freedom as shown by the solid (green) line. The actual observed mapping for our data is shown by the black histogram, while the corresponding distributions for the alternative ensembles are displayed by the colored, dashed histograms.

the obtained mapping on the chosen SM point for $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$, we construct similar maps between the confidence level and $-2\Delta \log \mathcal{L}$ for other random points in the $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ plane and find very similar dependencies. Consequently, we consider the mapping determined at the SM point to apply for all points in the $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ plane.

To obtain confidence regions in $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$, we determine profile-likelihood ratios for a grid on the $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ plane. In a Gaussian regime, the points with p-value = 0.05, corresponding to a confidence level of 95%, are identified by the intersection of the two-dimensional profile-likelihood function and a horizontal plane which is 5.99 units above the global minimum. The value 5.99 is the point on the $-2\Delta \log \mathcal{L}$ axis where the χ^2 distribution with two degrees of freedom (green line) intersects the $1 - 0.95 = 0.05$ level (red dashed line) in Fig. 14. The 68% C.L. is correspondingly obtained by the top horizontal (blue) line. The intersection between the 0.05 level and the actual mapping (black histogram) is at $-2\Delta \log \mathcal{L} = 7.34$ which means that the 95% confidence region is obtained by taking the intersection of the two-dimensional profile-likelihood function and a horizontal plane which is 7.34 units above the global minimum. In this case we find the standard model p-value for $\beta_s^{J/\psi\phi}$ to be 0.27. Clearly, this procedure leads to confidence regions larger than in the ideal, Gaussian case.

In order to guarantee additional coverage over a conservative range of possible values of nuisance parameters, sixteen alternative ensembles are generated. As we do not know the true values for these nuisance parameters, we compute the coverage over a wide range of possible values but always within their physically allowed range [61]. In particular, each alternative ensemble is produced by generating pseudoexperiments with nuisance parameters randomized uniformly within $\pm 5\,\sigma$ of their best fit values as obtained from maximizing the likelihood function on data. In these pseudoexperiments, the parameters $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$ are again fixed to their standard model expectation. We choose a random variation of $\pm 5\,\sigma$ over the nuisance parameters because we aim to cover the space of nuisance parameters with a C.L. much larger than the anticipated C.L. for our final result. Exceptions to this approach are the strong phases which are generated only within the range from zero to 2π and the dilution scale factors which are generated so that the dilution is always between zero and one. The other exception to applying a $\pm 5\,\sigma$ range is the phase δ_{SW}, which is generated flat between 0 and 2π. Since the S-wave fraction f_{SW} is consistent with zero as discussed in Sec. VIII B, we lack sensitivity to the associated phase and choose to vary it over its full range possible.

To determine this additional coverage adjustment, we again generate 1000 pseudoexperiments for each of the sixteen alternative ensembles. The same profile-likelihood ratio procedure is performed on each ensemble, and the broadest and thus most conservative p-value is taken to form the final confidence regions. The colored, dashed lines in Fig. 14 show the resulting mappings between 1 − C.L. and $-2\Delta \ln \mathcal{L}$ for each of the sixteen alternative ensembles. We use the p-value of the most conservative ensemble to determine the corresponding 68% and 95% confidence regions for our data. The intersection between the 0.05 (0.32) confidence level (1 − C.L.) and the most conservative mapping is at $-2\Delta \log \mathcal{L} = 8.79$ (4.19). This means that the 95% (68%) confidence region with guaranteed coverage is obtained by taking the intersection of the two-dimensional profile-likelihood ratio and a horizontal plane which is 8.79 (4.19) units above the global minimum. Note that this procedure of randomly sampling a limited number of points in the space of all nuisance parameters and using the most conservative of the resulting profile-likelihood ratio distributions automatically accounts for the effect of systematic uncertainties.

A similar coverage adjustment procedure is carried out to determine individual confidence intervals for $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$ separately. When determining the $\beta_s^{J/\psi\phi}$ confidence interval, $\Delta\Gamma_s$ is randomized in the pseudoexperiment generation and treated analogously with the other nuisance parameters. We again generate 1000 pseudoexperiments per alternative ensemble for the final coverage adjustment in the one-dimensional case. Using a similar approach as in the two-dimensional case, we obtain again

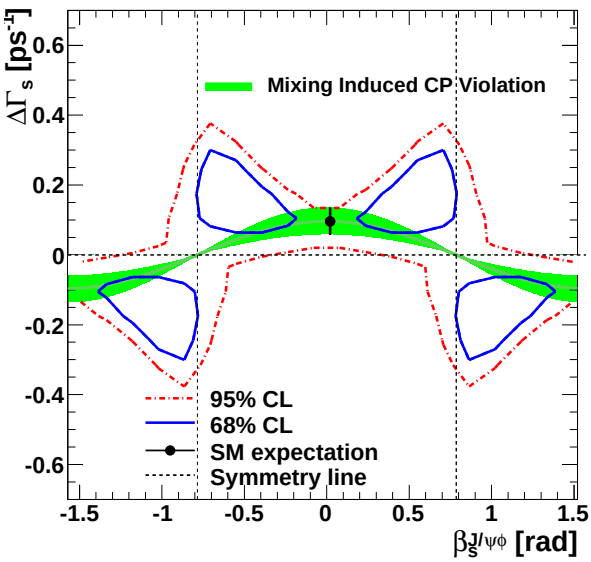

FIG. 15. (color online). Confidence regions in $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ plane for the fit without application of flavor tagging. The solid (blue) and dot-dashed (red) contours show the 68% and 95% confidence regions, respectively. The dotted lines are the symmetry axes corresponding to the profiled likelihood invariance under $\beta_s^{J/\psi\phi} \to \frac{\pi}{2} - \beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s \to -\Delta\Gamma_s$. In addition, the likelihood is invariant under $\beta_s^{J/\psi\phi} \to -\beta_s^{J/\psi\phi}$. The shaded (green) band is the theoretical prediction of mixing-induced CP violation.

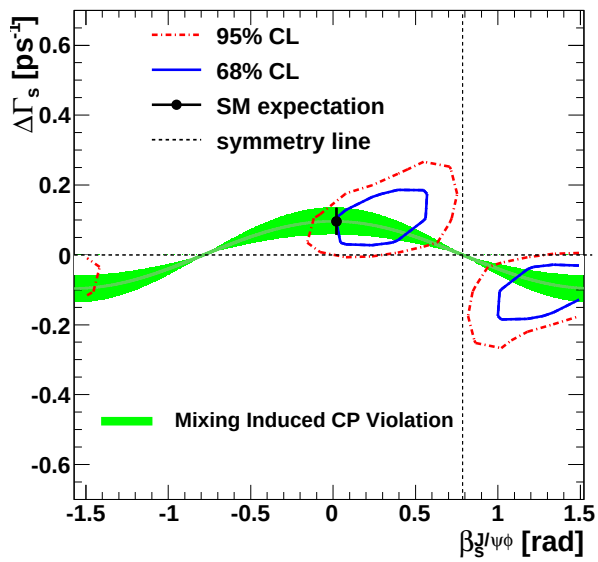

FIG. 16. (color online). Confidence regions in $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ plane for the fit including flavor tagging information. The solid (blue) and dot-dashed (red) contours show the 68% and 95% confidence regions, respectively. The dotted lines are the symmetry axes corresponding to the profiled likelihood invariance under $\beta_s^{J/\psi\phi} \to \frac{\pi}{2} - \beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s \to -\Delta\Gamma_s$. The shaded (green) band is the theoretical prediction of mixing-induced CP violation.

mappings of $1-$C.L. versus $-2\Delta\log\mathcal{L}$ for alternative ensembles with randomized nuisance parameters.

B. Results using frequentist approach

We present frequentist $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ confidence regions and p-values obtained according to the procedure described in Sec. VIII above. The $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ confidence regions without the application of flavor tagging are shown in Fig. 15. The SM prediction is indicated by the black marker, and the 68% and 95% C.L. regions are shown as solid (blue) and dot-dashed (red) contours, respectively. We find the p-value for $\beta_s^{J/\psi\phi}$ to agree with the standard model prediction to be 0.10. The shaded (green) band is the theoretical prediction of mixing-induced CP violation. As discussed above, in the absence of an S-wave component, the likelihood function is symmetric under the simultaneous transformations $\beta_s^{J/\psi\phi} \to \frac{\pi}{2} - \beta_s^{J/\psi\phi}$, $\Delta\Gamma_s \to -\Delta\Gamma_s$, $\delta_\perp \to \pi - \delta_\perp$ and $\delta_\parallel \to -\delta_\parallel$. In addition, if no flavor tagging information is used, an additional symmetry is present in the likelihood $\beta_s^{J/\psi\phi} \to -\beta_s^{J/\psi\phi}$. As a consequence of these symmetries, the likelihood function has four global maxima as can be seen in Fig. 15.

Once the flavor tagging information is added to the analysis, the $\beta_s^{J/\psi\phi} \to -\beta_s^{J/\psi\phi}$ symmetry is removed and the likelihood function has only two global maxima corresponding to the likelihood invariance under $\beta_s^{J/\psi\phi} \to \frac{\pi}{2} - \beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s \to -\Delta\Gamma_s$. The $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ confidence regions for the flavor tagged analysis, after coverage adjustment, are shown in Fig. 16. Our sensitivity to $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$ has substantially improved compared to our previously published measurement [15], as evidenced by the decrease in size of the confidence region. The result is also more consistent with the standard model prediction.

To illustrate the effect of the coverage adjustment, the left-hand side of Fig. 17 compares the 68% and 95% C.L. contours after coverage adjustment with the corresponding contours before the coverage adjustment procedure. A small increase in the size of the contours can be seen. As a further cross check, we also performed the same fit setting the S-wave fraction to zero as shown on the right-hand side of Fig. 17. The contour regions corresponding to a profile-likelihood ratio variation of $-2\Delta\log\mathcal{L} = 2.30$ (blue) and $-2\Delta\log\mathcal{L} = 5.99$ (red) are compared when including (solid) and not including (dashed) the S-wave fraction in the likelihood fit. The contours are almost identical.

The one-dimensional likelihood scan in the quantity $\beta_s^{J/\psi\phi}$ after coverage adjustment is shown in Fig. 18 on the left-hand side. In a Gaussian scenario the 68% (95%)

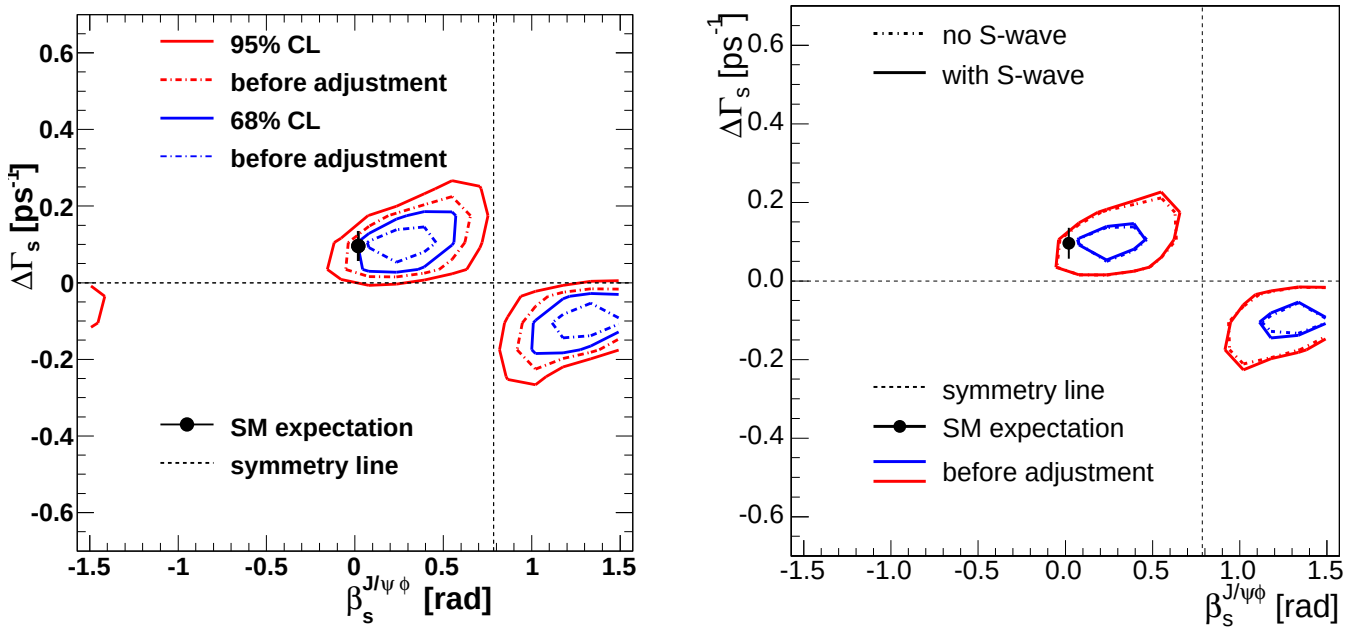

FIG. 17. (color online). Left: Confidence regions in $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ plane for the fit including flavor tagging information before (dashed) and after (solid) performing the coverage adjustment. Right: Comparison of including (solid) and not including (dashed) the S-wave contribution in the likelihood fit.

C.L. range is between the points of intersection of the profile-likelihood scan curve and a horizontal line which is one unit (four units) above the global minimum. In our case after coverage adjustment the solid (blue) and dot-dashed (red) horizontal lines which indicate the 68% and 95% C.L. ranges are at 2.74 and 7.11 units above the global minimum, respectively. We obtain

$$\begin{aligned}\beta_s^{J/\psi\phi} &\in [0.02, 0.52] \cup [1.08, 1.55] \text{ at 68\% confidence level,}\\ &\in [-\pi/2, -1.46] \cup [-0.11, 0.65] \cup [0.91, \pi/2] \text{ at 95\% confidence level.}\end{aligned}$$

We find the standard model p-value for $\beta_s^{J/\psi\phi}$ to be 0.30 corresponding to about one Gaussian standard deviation from the SM expectation as is also evidenced in Fig. 16. In comparison with the recent measurement of $\beta_s^{J/\psi\phi}$ from the D0 collaboration using a data sample based on 8 fb^{-1} of integrated luminosity [19], we find a similar region to constrain $\beta_s^{J/\psi\phi}$ at the 68% C.L. and obtain a similar p-value for comparison with the SM expectation. However, our result constrains $\beta_s^{J/\psi\phi}$ to a narrower region at the 95% confidence level.

In addition, we quote a confidence interval for the S-wave fraction after performing a likelihood scan for f_{SW} as shown in Fig. 19. We also show a quadratic fit overlaid indicating the parabolic shape of the likelihood around the minimum which we integrate to calculate upper limits on the S-wave fraction. The upper limit on the S-wave fraction over the mass interval $1.009 < m(K^+K^-) < 1.028$ GeV/c^2 corresponding to the selected K^+K^- signal region is 4% of the total signal at the 68% confidence level, and $f_{SW} < 6\%$ at 95% C.L. Since the analysis is limited to events in a narrow K^+K^- mass range around the ϕ signal, the observed S-wave fraction is small and its effect on the observables quoted in this analysis is minor. We verified with pseudoexperiments that a sizeable amount of S-wave would affect the measured value of $\beta_s^{J/\psi\phi}$. In contrast to our result, the recent D0 publication [19] quotes a sizeable fraction of $17.3 \pm 3.6\%$ for the S-wave fraction over almost the same K^+K^- mass range. We also perform a likelihood scan to determine the associated S-wave phase, but, as expected from simulated experiments, we find that we are not sensitive to δ_{SW} with the current data sample size.

Finally, we perform a flavor tagged analysis with $\Delta\Gamma_s$ Gaussian constrained to the theoretical prediction of $2\,|\Gamma_{12}^s| = (0.090 \pm 0.024)$ ps^{-1} [9]. Under this constraint, $\beta_s^{J/\psi\phi}$ is found in the range $[0.05, 0.40] \cup [1.17, 1.49]$ at the 68% confidence level, and within $[-\pi/2, -1.51] \cup [-0.07, 0.54] \cup [1.03, \pi/2]$ at 95% C.L. as shown in Fig. 18 on the right-hand side. The p-value for the SM expected

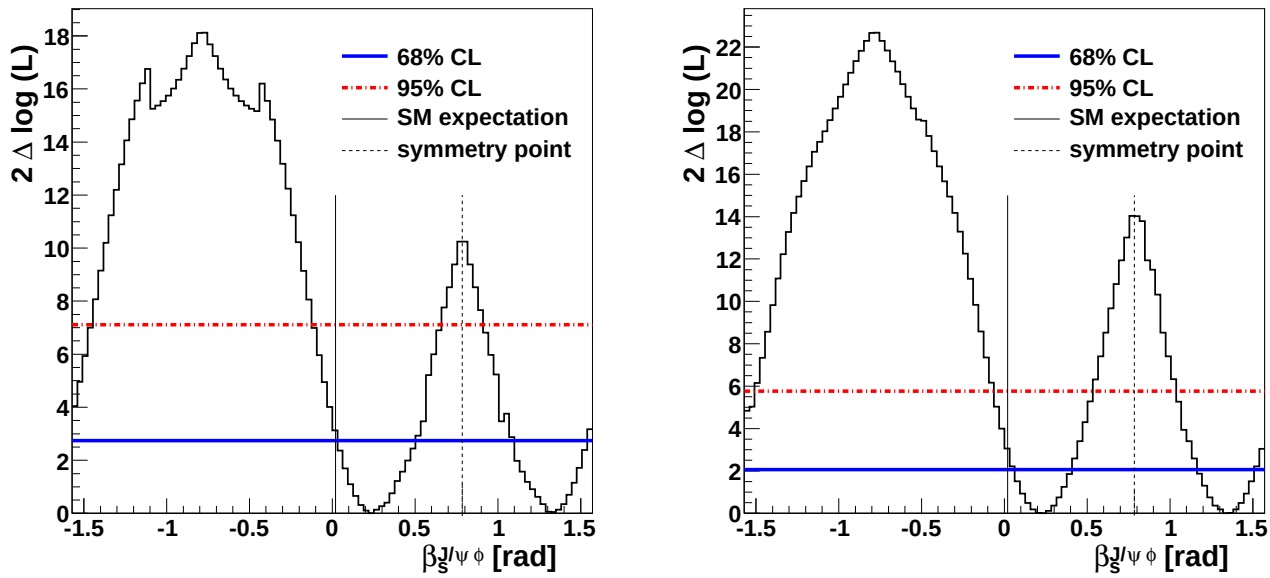

FIG. 18. Likelihood scan in $\beta_s^{J/\psi\phi}$ with no constraint (left) and with $\Delta\Gamma_s$ constrained to the SM prediction (right). After coverage adjustment the solid (blue) and dot-dashed (red) horizontal lines indicate the 68% (95%) C.L. range above the global minimum.

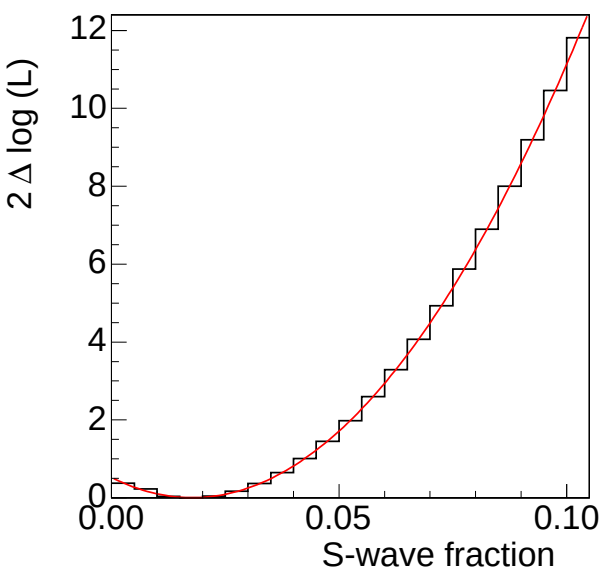

FIG. 19. Likelihood scan of the S-wave fraction with a quadratic fit overlaid indicating the parabolic shape of the likelihood around the minimum.

value of $\beta_s^{J/\psi\phi}$ from this constrained fit is 0.21, corresponding to a deviation from the SM expectation of $1.3\,\sigma$ significance. We note that the likelihood scans in Fig. 18 exhibit small deviations from the symmetry in $\beta_s^{J/\psi\phi}$ that is expected according to our discussion above. The reason is given by the small S-wave fraction that our likelihood fit finds as well as the choice of binning and numerical precision in determining the displayed $-2\Delta\log\mathcal{L}$ values.

IX. RESULTS ON $\beta_S^{J/\psi\phi}$ AND $\Delta\Gamma_S$ IN A BAYESIAN APPROACH

In addition to the frequentist results shown in the previous section, we use a Bayesian analysis to provide cross-checks on the determination of the physics parameters. We use Bayesian inference via integration of the posterior density obtained from the likelihood function described in Sec. VI over the nuisance parameters and over those physics parameters in which we are not presently interested.

The starting point for this Bayesian approach is the likelihood function, $\mathcal{L}(\vec{x}\,|\,\vec{\theta},\vec{\nu}\,)$, where $\vec{x}$ are the experimental measurements including the B_s^0 candidate decay time and invariant mass, the transversity angles and tagging information, while $\vec{\mu}=(\vec{\theta},\vec{\nu}\,)$ is a vector distinguishing the physics parameters $\vec{\theta}$ described in Table II from the remaining nuisance parameters $\vec{\nu}$ in the fit describing features such as background shapes, tagging performance, and detector resolution (see Sec. VI). In our analysis the dimensionality of $\vec{\theta}$ and $\vec{\nu}$ is 11 and 24, respectively.

Within the Bayesian approach to statistical inference, Bayes' theorem defines the posterior probability density

given the observed data set $\vec{x}$

$$p(\vec{\theta}\,|\,\vec{x}) = \frac{p(\vec{x}\,|\,\vec{\theta}\,)\,p(\vec{\theta}\,)}{\int p(\vec{x}\,|\,\vec{\theta}\,)\,p(\vec{\theta}\,)\;d^N\theta}\,, \tag{32}$$

where $p(\vec{x}\,|\,\vec{\theta}\,)$ is the likelihood function $\mathcal{L}(\vec{x}\,|\,\vec{\theta}\,)$ and $p(\vec{\theta}\,)$ is the prior probability density for $\vec{\theta}$, which describes the knowledge about parameters $\vec{\theta}$ that we assume prior to our measurement. The projection of the N-dimensional posterior density onto M parameters of physical interest corresponds to an integration over all the other $N-M$ parameters, where the limits of integration cover all possible values for the $N-M$ parameters. This projection gives a new posterior density $p(\vec{\alpha}\,|\,\vec{x})$ for parameters $\alpha_i \in \vec{\theta}$. For a single variable α and a uniform prior density this expression reduces to

$$p(\alpha\,|\,\vec{x}) = \int p(\vec{\theta}\,|\,\vec{x})\,(\theta)\;d^{N-1}\theta\,. \tag{33}$$

We compute a representation of the 35-dimensional likelihood function including nuisance parameters, which we store in a computer-readable data format (Ntuple). These Ntuples contain Markov chains, which have been generated by a Markov chain Monte Carlo (MCMC) technique [62]. Projections of the Markov chains onto subspaces of physical parameters of interest, in particular $\beta_s^{J/\psi\phi}$, $\Delta\Gamma_s$, etc., are then performed. In addition, we compute credible intervals for certain parameters and credible contours for pairs of parameters derived from these projections, as discussed in more detail below.

From a Markov chain projection one may easily draw conclusions about specific values of parameters such as $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$ with a view toward propagating that information into global fits of, e.g., CKM parameters and incorporating certain prior information, about, e.g., mixing-induced *CP* violation and the values of the strong phases that appear in $B_s^0 \to J/\psi\phi$ decays. By projecting a Markov chain onto a subspace of parameters of dimension M, i.e., making a histogram or a scatter-plot from the Ntuples, one is in fact performing a numerical integration of the posterior density, over the other $N-M$ parameters. The normalization factor, i.e., the denominator in Eq. (32), is easily identified as the total number of points of the Markov Chain.

We can define a credible interval $[\alpha_L, \alpha^U]$ for the parameter α with probability content β through

$$\beta = \int_{\alpha_L}^{\alpha^U} p(\alpha\,|\,\vec{x})\;d\alpha = P(\alpha_L < \alpha < \alpha^U). \tag{34}$$

The credible interval $[\alpha_L, \alpha^U]$ contains a fraction β of the posterior density about α but it is not unique. However, we can build a "shortest interval" using the straightforward algorithm of maximum probability ordering by accepting into the interval the largest values of the PDF $p(\alpha\,|\,\vec{x})$. Using the same algorithm as in the one-dimensional case, we build credible contours in the $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ plane. A credible interval (or contour) does not necessarily cover the true value of a parameter (or parameters) with any given frequentist probability. On the other hand, regarding the technical aspect it allows for the combination of experimental results and theoretical inputs in a straightforward manner. In addition, the Bayesian technique can be trivially modified to incorporate other conditions on, for example, mixing-induced *CP* violation or constraints on the strong phase angles through non-uniform prior probability densities (see Sec. IX B and Sec. IX B).

To verify convergence of the Markov chain, we generate sixteen independent chains. A burn-in phase of approximately 10 000 steps is identified. We discard the first 250 000 states in each chain and keep the following one million states. This means the probability densities shown below are based on sixteen million states.

A. Results using Bayesian approach

The projections of the sixteen chains onto the variables $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$ are displayed in Fig. 20 together with the sum of all sixteen chains. The close agreement between the sixteen independent chains on the left-hand side of Fig. 20 also demonstrates the convergence of the computation. Using this Bayesian analysis of the data, we obtain

$$\begin{aligned}
\beta_s^{J/\psi\phi} &\in [0.11, 0.41] \cup [1.16, 1.47] \;\text{ as 68\% credible interval,}\\
&\in [-0.04, 0.59] \cup [0.98, 1.62] \;\text{ as 95\% credible interval,}\\
\Delta\Gamma_s &\in \left[-0.14\ \text{ps}^{-1}, -0.06\ \text{ps}^{-1}\right] \cup \left[0.06\ \text{ps}^{-1}, 0.14\ \text{ps}^{-1}\right] \;\text{ as 68\% credible interval,}\\
&\in \left[-0.18\ \text{ps}^{-1}, -0.02\ \text{ps}^{-1}\right] \cup \left[0.02\ \text{ps}^{-1}, 0.18\ \text{ps}^{-1}\right] \;\text{ as 95\% credible interval.}
\end{aligned}$$

The joint posterior probability densities and credible contours in the $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ plane are displayed in Fig. 21. The narrow band shown in the figure is the theoretical prediction of mixing-induced *CP* violation using $2|\Gamma_{12}^s| =$

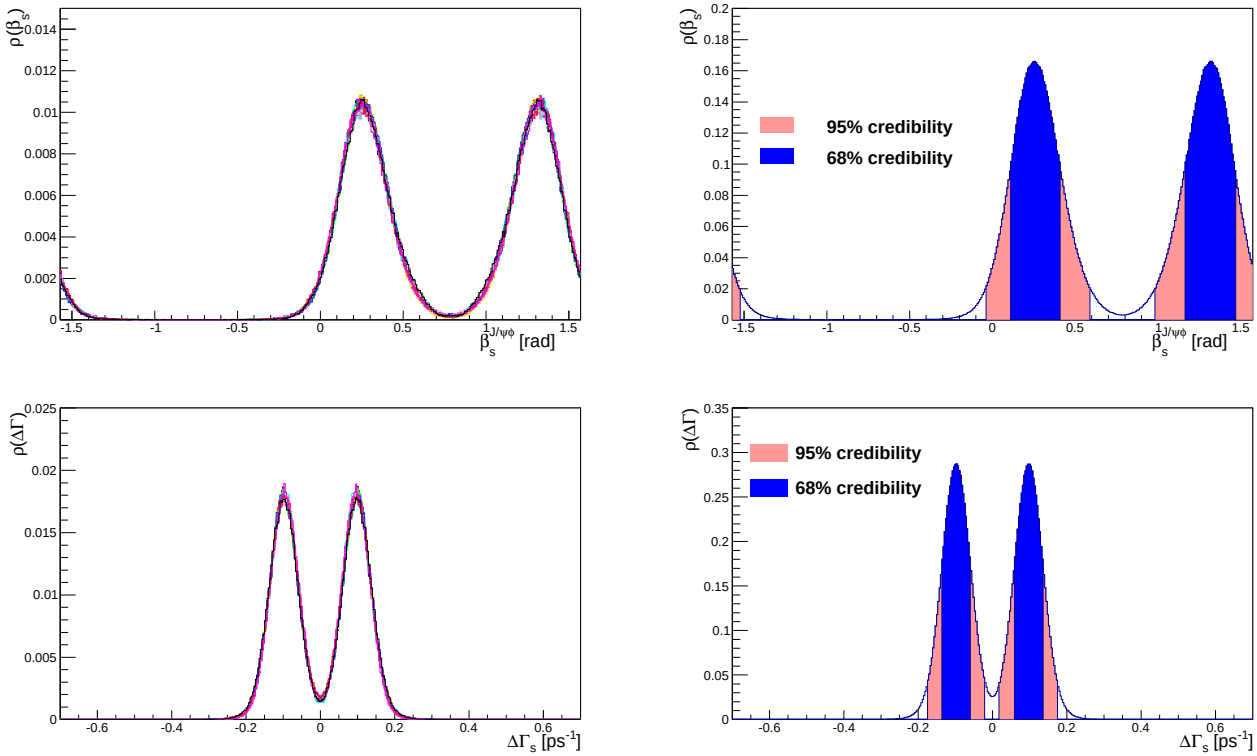

FIG. 20. (color online). Bayesian posterior densities for the variables $\beta_s^{J/\psi\phi}$ (top) and $\Delta\Gamma_s$ (bottom). The left plots show projections of sixteen independent Markov Chains, while the right two plots show the posterior densities with 68% and 95% credible intervals in dark-solid (blue) and light-solid (red) areas, respectively.

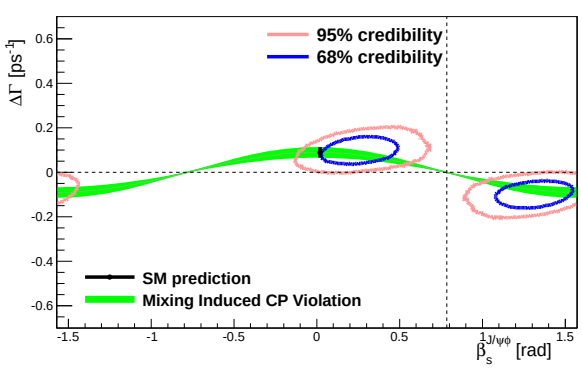

FIG. 21. (color online). Joint posterior probability density in the $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ plane for the combined analysis. The dark-solid (blue) and light-solid (red) contours show the 68% and 95% credible regions, respectively. The narrow band (green) is the theoretical prediction of mixing-induced CP violation.

(0.090 ± 0.024) ps^{-1} [9, 10] overlapping with our result.

Furthermore, we calculate

$$\int_{\mathcal{A}_{\beta_s}} p(\beta_s^{J/\psi\phi})\, d\beta_s^{J/\psi\phi} = 0.119, \qquad (35)$$

where $\mathcal{A}_{\beta_s} = \{\beta_s^{J/\psi\phi}; p(\beta_s^{J/\psi\phi}) < p(\beta_s^{SM})\}$, indicating that, as the contour is expanded to enclose larger credibility, the standard model prediction β_s^{SM} is first included within the enclosed region at a credibility of 88.1%.

We also examine the posterior density in the variables $\delta_\parallel$ versus $\delta_\perp$ as shown in Fig. 22. It is predicted that the phases in $B_s^0 \to J/\psi\phi$ match those in the equivalent decay $B^0 \to J/\psi K^{*0}$ to within 10° [63]. The measured values of these phases in the $B^0 \to J/\psi K^{*0}$ decay ($\delta_\parallel = -2.93 \pm 0.08\,(\text{stat}) \pm 0.04\,(\text{syst})$ rad and $\delta_\perp = 2.91 \pm 0.05\,(\text{stat}) \pm 0.03\,(\text{syst})$ rad [64]) are overlaid in form of a green ellipse in Fig. 22. The width of this ellipse includes the 10° theoretical uncertainty added in quadrature with the experimental uncertainties on $\delta_\parallel$ and $\delta_\perp$. For one mode of the probability density good agreement between the B_s^0 and B^0 system is observed as predicted in Ref. [63].

B. Constrained results

Figure 21 shows that our measurement in the $\beta_s^{J/\psi\phi}$-$\Delta\Gamma_s$ plane is consistent with the hypothesis of mixing-induced CP violation as well as with the hypothesis that the measured CP violation originates from the standard model. We apply the hypothesis of mixing-induced CP

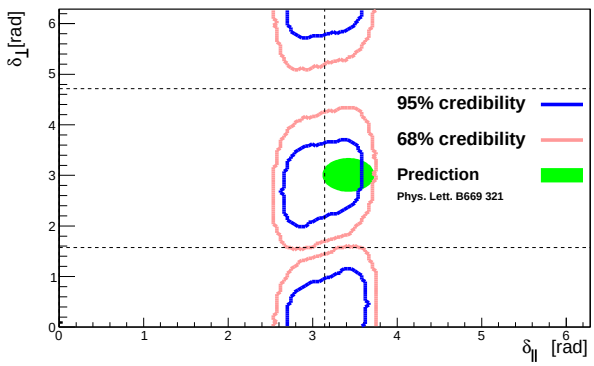

FIG. 22. (color online). Posterior density in the strong phases $\delta_\perp$ versus $\delta_\parallel$ overlaid with the prediction that the phases in $B_s^0 \to J/\psi\phi$ match those in $B^0 \to J/\psi K^{0*}$ decays to within 10° [63]. The dark-solid (blue) and light-solid (red) contours show the 68% and 95% credible regions, respectively. The width of the light-shaded (green) ellipse includes the 10° theoretical uncertainty added in quadrature with the experimental uncertainties on $\delta_\parallel$ and $\delta_\perp$ from $B^0 \to J/\psi K^{0*}$.

violation, together with the theoretical calculation of Γ_{12}^s, to our data in form of a prior density during the computation of the MCMC. We carry out this calculation by simply re-weighting the Markov chains using a new flat prior density derived from the theoretical calculation which gives $2\,|\Gamma_{12}^s| = (0.090 \pm 0.024)$ ps^{-1} [9].

The 68% credible interval on the CP-violating quantity $\beta_s^{J/\psi\phi}$ is $\beta_s^{J/\psi\phi} \in [0.09, 0.32] \cup [1.24, 1.48]$. The posterior density in $\beta_s^{J/\psi\phi}$ alone is shown in Fig. 23 on the left-hand side. Again, we calculate the quantity

$$\int_{\mathcal{A}_{\beta_s}} p(\beta_s^{J/\psi\phi})\; d\beta_s^{J/\psi\phi} = 0.131, \tag{36}$$

where $\mathcal{A}_{\beta_s} = \{\beta_s^{J/\psi\phi}; p(\beta_s^{J/\psi\phi}) < p(\beta_s^{SM})\}$. The SM is first included at a credibility of 86.9%.

As can be seen in Fig. 22, the theoretical predictions of Ref. [63] are consistent with one of the modes of the probability density in the $\delta_\perp$ versus $\delta_\parallel$ plane but not with the other mode. In the following, conditional posterior densities are used to show that the favored mode of probability from Fig. 24 corresponds to the solution $\beta_s^{J/\psi\phi} \in [0.09, 0.32]$.

The sequence of plots in Fig. 24 illustrates this statement. The figures in the top row show the conditional posterior density in the parameters $\delta_\parallel$ and $\delta_\perp$ after imposing the requirement $|\beta_s^{J/\psi\phi}| < \pi/4$. It can be seen that this condition completely eliminates one of the modes of the probability density in the $\delta_\parallel$ versus $\delta_\perp$ parameter space. The plots in the bottom row of Fig. 24 show that, conversely, the condition $\pi/2 < \delta_\perp < 3\pi/2$ eliminates one mode of the probability projected onto $\beta_s^{J/\psi\phi}$, while there is virtually no impact on the other mode with $|\beta_s^{J/\psi\phi}| < \pi/4$.

TABLE V. Summary of the sensitivity study. The 68% credibility interval on $\beta_s^{J/\psi\phi}$ is given for the unconstrained result and when $2\,|\Gamma_{12}^s|$ is constrained to its SM prediction.

Variation	Constrained	Unconstrained
Default	[0.09,0.32]	[0.11,0.41]
Flat $\sin 2\beta_s^{J/\psi\phi}$	[0.08,0.31]	[0.09,0.37]
Flat $\cos\delta_\perp$	[0.09,0.33]	[0.10,0.43]
Flat $\cos\delta_\parallel$	[0.09,0.32]	[0.11,0.41]
Previous three together	[0.07,0.31]	[0.09,0.39]
Flat in amplitudes	[0.09,0.32]	[0.11,0.41]
Gaussian mix.-ind. CP viol.	[0.09,0.34]	

These conditional probabilities allow us to visualize what is happening in the larger space of all four parameters ($\beta_s^{J/\psi\phi}$, $\Delta\Gamma_s$, $\delta_\parallel$ and $\delta_\perp$), and to identify the prediction of Ref. [63] with only one of the solutions for $\beta_s^{J/\psi\phi}$, namely $\beta_s^{J/\psi\phi} \in [0.09, 0.32]$ at the 68% credible interval. This result confirms the early finding in our previous publication [15] which indicated that the solution centered in $0 \le \beta_s^{J/\psi\phi} \le \pi/4$ and $\Delta\Gamma_s > 0$ corresponds to $\cos(\delta_\perp) < 0$ and $\cos(\delta_\perp - \delta_\parallel) > 0$, while the opposite is true for the solution centered in $\pi/4 \le \beta_s^{J/\psi\phi} \le \pi/2$ and $\Delta\Gamma_s < 0$.

C. Sensitivity analysis

It is a well-known fact that Bayesian results depend on the chosen prior densities; a flat prior in a given metric might, in general, not be flat in another metric. To study such effects, we carried out a sensitivity analysis in order to characterize the degree to which the Bayesian results of this section depends upon the chosen input priors. The sensitivity analysis was performed by weighting the probability density with the Jacobian of the transformation from the default parameterization to the desired parameterization. Using this technique we checked the variation of the Bayesian result with respect to the following six variations. First, the prior is taken flat in $\sin 2\beta_s^{J/\psi\phi}$ rather than flat in $\beta_s^{J/\psi\phi}$; second, the prior is taken flat in $\cos\delta_\perp$, and third, the prior is taken flat in $\cos\delta_\parallel$. Afterwards, all three conditions are applied together at the same time. Fifth, the prior is taken flat in the amplitudes $A_\parallel(0)$ and $A_\perp(0)$ rather than in their squares and finally, the mixing-induced CP violation constraint is taken as a Gaussian rather than flat constraint. The effect of changing the priors on the 68% credibility intervals on $\beta_s^{J/\psi\phi}$ is summarized in Table V. Modest changes are observed for the unconstrained result and the result with $|\Gamma_{12}^s|$ constrained to $2\,|\Gamma_{12}^s| = (0.090 \pm 0.024)$ ps^{-1} [9]. Only the effect on the first $\beta_s^{J/\psi\phi}$ credibility interval is shown, since the second interval can be trivially derived from the numbers in Table V.

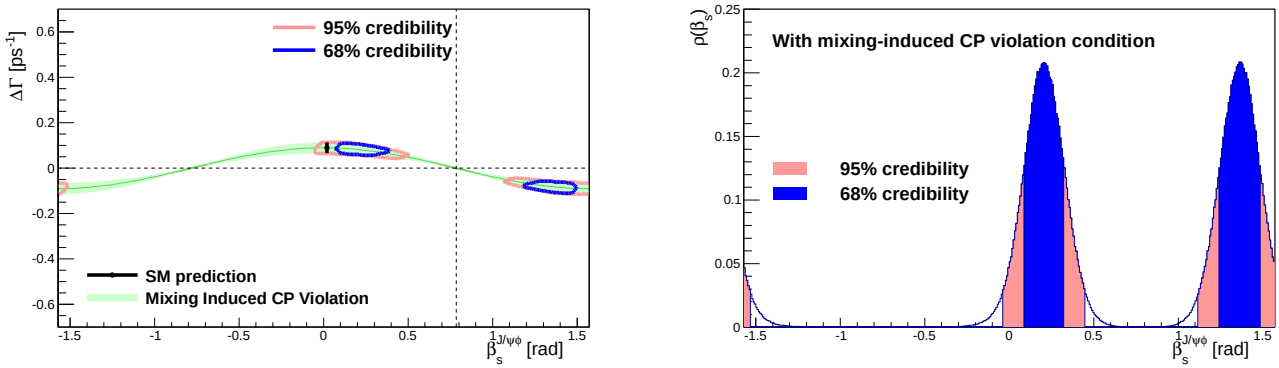

FIG. 23. (color online). Two-dimensional posterior densities for the variables $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma_s$ (left), and one-dimensional posterior density for the variable $\beta_s^{J/\psi\phi}$ (right) including a prior density for mixing-induced CP violation. The dark-solid (blue) and light-solid (red) contours show the 68% and 95% credible regions, respectively.

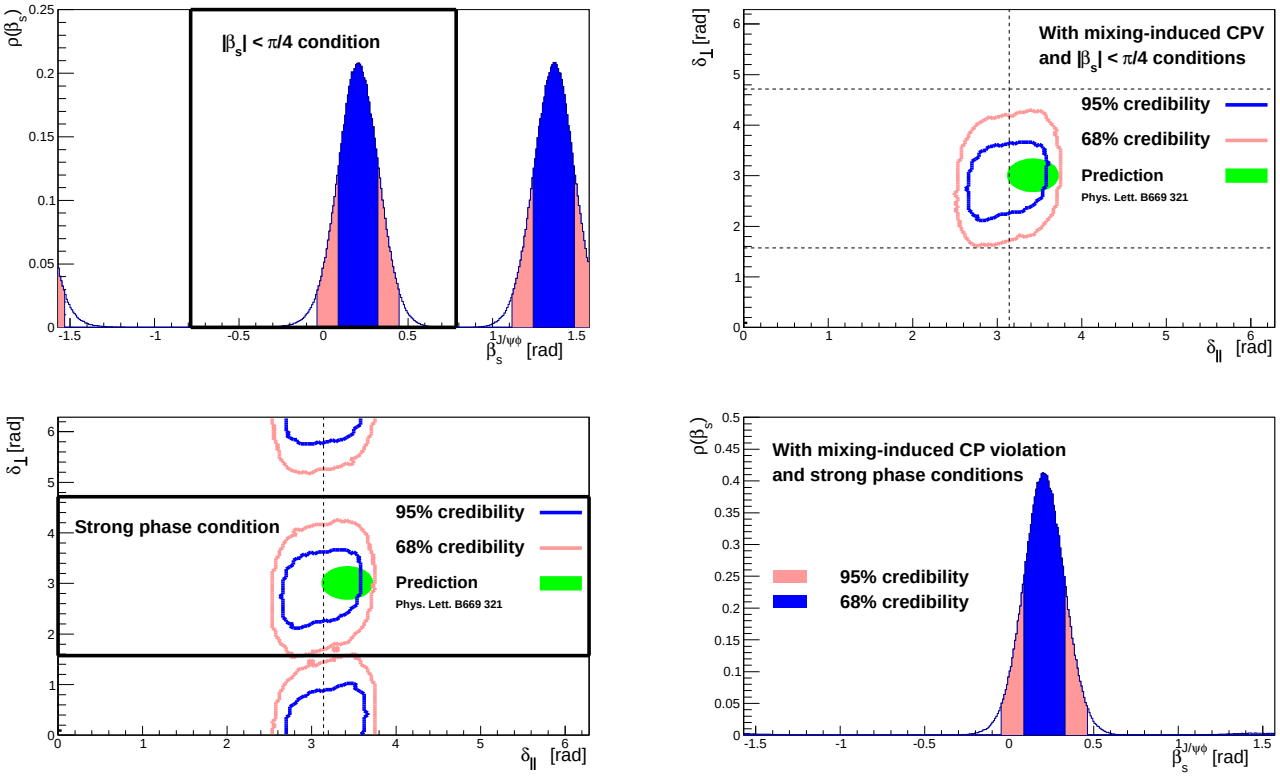

FIG. 24. (color online). Conditional posterior densities for $\delta_\perp$ versus $\delta_\parallel$ and for $\beta_s^{J/\psi\phi}$. The extra conditions applied to $\beta_s^{J/\psi\phi}$ (top row) and $\delta_\perp$ (bottom row) are shown on the left, and the resulting conditional probabilities are displayed on the right. The theoretical prediction of Ref. [63] is indicated as green ellipse in the bottom left plot. All plots in this figure are subject to the constraint of mixing-induced CP violation. The dark-solid (blue) and light-solid (red) contours show the 68% and 95% credible regions, respectively.

X. SUMMARY

In summary, we have presented a measurement of CP violation in $B_s^0 \to J/\psi\phi$ decays using 6500 signal events from a data sample with 5.2 fb^{-1} integrated luminosity collected with the CDF II detector operating at the Tevatron $p\bar{p}$ collider. We find the CP-violating phase $\beta_s^{J/\psi\phi}$ to be within the range $\beta_s^{J/\psi\phi} \in [0.02, 0.52] \cup [1.08, 1.55]$ at 68% confidence level and within the interval $[-\pi/2, -1.46] \cup [-0.11, 0.65] \cup [0.91, \pi/2]$ at 95% C.L.

where the coverage property of the quoted interval is guaranteed using a frequentist statistical analysis. Assuming the standard model expectation for $\beta_s^{J/\psi\phi}$, the probability to observe a fluctuation as large as in our data or larger is given by a p-value of 0.30 corresponding to about one Gaussian standard deviation. This result shows less of a discrepancy with the SM expectation than our previously published result using 1.3 fb^{-1} of integrated luminosity [15]. In comparison with the recent measurement of $\beta_s^{J/\psi\phi}$ from the D0 collaboration using a data sample based on 8 fb^{-1} of integrated luminosity [19], our result agrees within uncertainty but constrains $\beta_s^{J/\psi\phi}$ to a narrower region. With $\Delta\Gamma_s$ constrained to its SM prediction, we find $\beta_s^{J/\psi\phi}$ in the range $[0.05, 0.40] \cup [1.17, 1.49]$ at the 68% C.L. The measurement of the CP-violating phase $\beta_s^{J/\psi\phi}$ is still statistics-limited. It will improve with the final 2011 CDF dataset approximately doubling the current integrated luminosity.

This analysis also incorporates the possibility of contributions to the $B_s^0 \to J/\psi K^+K^-$ final state in the region of the ϕ resonance from $B_s^0 \to J/\psi f_0$ and $B_s^0 \to J/\psi K^+K^-$ (non-resonant) decays. We measure the S-wave contribution over the mass interval $1.009 < m(K^+K^-) < 1.028$ GeV/c^2 corresponding to the selected K^+K^- signal region to be less than 6% (4%) at the 95% (68%) confidence level. We do not confirm a sizeable fraction of $17.3 \pm 3.6\%$ for the S-wave fraction as quoted over almost the same K^+K^- mass range in a recent D0 publication [19].

Assuming the standard model prediction for the CP-violating phase $\beta_s^{J/\psi\phi}$, we measure several other parameters describing the B_s^0 system. These include the B_s^0 mean lifetime $\tau(B_s^0)$, the decay width difference $\Delta\Gamma_s$ between the heavy and light B_s^0 mass eigenstates, the transversity amplitudes $|A_\parallel(0)|^2$ and $|A_0(0)|^2$, as well as the strong phase $\delta_\perp$. The measurements for $\tau(B_s^0) = 1.529 \pm 0.025\,(\text{stat}) \pm 0.012\,(\text{syst})$ ps and $\Delta\Gamma_s = 0.075 \pm 0.035\,(\text{stat}) \pm 0.006\,(\text{syst})$ ps^{-1} are the most precise measurements of these quantities using a single decay mode. They are also in good agreement with the PDG world averages [8]. The measurements of the transversity amplitudes are consistent with previous measurements in the $B_s^0 \to J/\psi\phi$ system. Finally, we report an alternative Bayesian analysis based on Markov chain integration that gives results consistent with the frequentist approach.

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, Culture, Sports, Science and Technology 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; the Bundesministerium für Bildung und Forschung, Germany; the Korean World Class University Program, the National Research Foundation of Korea; the Science and Technology Facilities Council and the Royal Society, UK; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; the Academy of Finland; and the Australian Research Council (ARC).

[1] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, Phys. Rev. Lett. **13**, 138 (1964).
[2] J. J. Aubert *et al.* (E598 Collaboration), Phys. Rev. Lett. **33**, 1404 (1974); J. E. Augustin *et al.* (SLAC-SP-017 Collaboration), *ibid.* **33**, 1406 (1974).
[3] S. W. Herb *et al.*, Phys. Rev. Lett. **39**, 252 (1977).
[4] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 2626 (1995); S. Abachi *et al.* (D0 Collaboration), *ibid.* **74**, 2632 (1995).
[5] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
[6] M. Kreps, A. Lenz, and O. Leroy, arXiv:1103.4962 (2011).
[7] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. Lett. **97**, 242003 (2006).
[8] K. Nakamura *et al.* (Particle Data Group), J. Phys. G **37**, 075021 (2010).
[9] A. Lenz and U. Nierste, J. High Energy Phys. 06, 072 (2007).
[10] B. A. Dobrescu, P. J. Fox, and A. Martin, Phys. Rev. Lett. **105**, 041801 (2010).
[11] S. Nandi and A. Soni, Phys. Rev. D **83**, 114510 (2011); J. Drobnak, S. Fajfer, and J. F. Kamenik, Phys. Lett. B **701**, 234 (2011); R.-M. Wang, Y.-G. Xu, Q. Chang, and Y.-D. Yang, Phys. Rev. D **83**, 095010 (2011); J. Shelton and K. M. Zurek, *ibid.* **83**, 091701 (2011); J. Girrbach *et al.*, J. High Energy Phys. 06, 044 (2011), [Erratum-ibid.07:001,2011]; A. J. Buras, M. Nagai, and P. Paradisi, *ibid.* 05, 005 (2011); Z. J. Ajaltouni and E. Di Salvo, J. Phys. D **37**, 125001 (2010); Y. Bai and A. E. Nelson, Phys. Rev. D **82**, 114027 (2010); L. L. Everett, J. Jiang, P. G. Langacker, and T. Liu, *ibid.* **82**, 094024 (2010); J. P. Saha, B. Misra, and A. Kundu, *ibid.* **81**, 095011 (2010); G. Isidori, Y. Nir, and G. Perez, Ann. Rev. Nucl. Part. Sci. **60**, 355 (2010); F. J. Botella, G. C. Branco, and M. N. Rebelo, Phys. Lett. B **687**, 194 (2010); C.-W. Chiang *et al.*, J. High Energy Phys. 04, 031 (2010).
[12] Z. Ligeti, M. Papucci, and G. Perez, Phys. Rev. Lett. **97**, 101801 (2006).
[13] I. Dunietz, R. Fleischer, and U. Nierste, Phys. Rev. D **63**, 114015 (2001).
[14] J. Charles *et al.* (CKMfitter Group), Eur. Phys. J. C **41**, 1 (2005); M. Bona *et al.* (UT*fit* Collaboration), J. High

Energy Phys. 10, 081 (2006).
[15] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **100**, 161802 (2008).
[16] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **101**, 241801 (2008).
[17] M. Paulini, Int. J. Mod. Phys. A **24**, 4413 (2009).
[18] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **82**, 032001 (2010); **84**, 052007 (2011).
[19] V. M. Abazov *et al.* (D0 Collaboration), arXiv:1109.3166 (2011).
[20] LHCb Collaboration, Report No. LHCb-CONF-2011-049 (2011).
[21] A. S. Dighe, I. Dunietz, and R. Fleischer, Eur. Phys. J. C **6**, 647 (1999).
[22] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **100**, 121803 (2008).
[23] S. Stone and L. Zhang, Phys. Rev. D **79**, 074024 (2009).
[24] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **71**, 032001 (2005).
[25] R. Blair *et al.* (CDF Collaboration), FERMILAB-PUB-96-390-E(1996).
[26] A. Sill (CDF Collaboration), Nucl. Instrum. Methods A **447**, 1 (2000).
[27] C. S. Hill (CDF Collaboration), Nucl. Instrum. Methods A **530**, 1 (2004).
[28] A. A. Affolder *et al.* (CDF Collaboration), Nucl. Instrum. Methods A **453**, 84 (2000).
[29] A. A. Affolder *et al.* (CDF Collaboration), Nucl. Instrum. Methods A **526**, 249 (2004).
[30] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **103**, 031801 (2009); **106**, 181802 (2011).
[31] L. Balka *et al.* (CDF Collaboration), Nucl. Instrum. Methods A **267**, 272 (1988).
[32] S. Bertolucci *et al.* (CDF Collaboration), Nucl. Instrum. Methods A **267**, 301 (1988).
[33] A. Abulencia *et al.* (CDF Collaboration), J. Phys. G **34**, 2457 (2007).
[34] G. Ascoli *et al.* (CDF Collaboration), Nucl. Instrum. Methods A **268**, 33 (1988); T. Dorigo (CDF Collaboration), **461**, 560 (2001).
[35] E. J. Thomson *et al.*, IEEE Trans. Nucl. Sci. **49**, 1063 (2002).
[36] M. Feindt and U. Kerzel, Nucl. Instrum. Methods A **559**, 190 (2006).
[37] V. V. Anisovich and B. C. Metsch, Phys. Rev. D **46**, 3195 (1992).
[38] C. Peterson, D. Schlatter, I. Schmitt, and P. M. Zerwas, Phys. Rev. D **27**, 105 (1983).
[39] D. J. Lange, Nucl. Instrum. Methods A **462**, 152 (2001).
[40] R. Brun, R. Hagelberg, M. Hansroul, and J. C. Lassalle, CERN-DD-78-2-REV(1978); R. Brun and F. Carminati, in 9th International Conference on Computing in High-Energy Physics (CHEP 91), Tsukuba, Japan, 1991(1991).
[41] E. Gerchtein and M. Paulini, ECONF **C0303241**, TUMT005 (2003).
[42] G. Giurgiu, Ph.D. thesis, Carnegie Mellon University, FERMILAB-THESIS-2005-41(2005).
[43] E. Pueschel, Ph.D. thesis, Carnegie Mellon University, FERMILAB-THESIS-2010-70(2010).
[44] S. M. Flatté, Phys. Lett. B **63**, 224 (1976).
[45] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. Lett. **97**, 062003 (2006).
[46] A. Belloni, Ph.D. thesis, MIT, FERMILAB-THESIS-2007-36(2007).
[47] J. Morlock, Ph.D. thesis, University of Karlsruhe, Report No. IEKP-KA/2010-13(2010).
[48] B. Ashmanskas *et al.* (CDF Collaboration), Nucl. Instrum. Methods A **518**, 532 (2004).
[49] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. Lett. **96**, 191801 (2006).
[50] V. Tiwari, Ph.D. thesis, Carnegie Mellon University, FERMILAB-THESIS-2007-09(2007).
[51] C. Lecci, Ph.D. thesis, University of Karlsruhe, FERMILAB-THESIS-2005-89(2005).
[52] F. Azfar *et al.*, J. High Energy Phys. 11, 158 (2010).
[53] F. James and M. Roos, Comput. Phys. Commun. **10**, 343 (1975).
[54] E. P. Wigner, Phys. Rev. **98**, 145 (1955).
[55] A. S. Dighe, I. Dunietz, H. J. Lipkin, and J. L. Rosner, Phys. Lett. B **369**, 144 (1996).
[56] G. Punzi, arXiv:physics/0401045 (2004).
[57] L. B. Oakes, Ph.D. thesis, Oxford University, FERMILAB-THESIS-2010-60(2010).
[58] K. Anikeev, Ph.D. thesis, MIT, FERMILAB-THESIS-2004-12(2004).
[59] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **104**, 102002 (2010); **106**, 121804 (2011).
[60] G. J. Feldman and R. D. Cousins, Phys. Rev. D **57**, 3873 (1998).
[61] B. L. Berger and D. D. Boos, J. Am. Stat. Assoc. **89**, 1012 (1994).
[62] W. Gilkes, S. Richardson, and D. Spiegelhalter, *Markov Chain Monte Carlo in Practice* (Chapman and Hall / CRC Interdisciplinary Statistics, 1996).
[63] M. Gronau and J. L. Rosner, Phys. Lett. B **669**, 321 (2008).
[64] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **76**, 031102 (2007).