Search for the Flavor-Changing Neutral Current Decay $B_s^0 \to \mu^+\mu^-$ in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV with the D0 Detector

V. M. Abazov,[33] B. Abbott,[70] M. Abolins,[61] B. S. Acharya,[27] M. Adams,[48] T. Adams,[46] M. Agelou,[17] J.-L. Agram,[18] S. H. Ahn,[29] M. Ahsan,[55] G. D. Alexeev,[33] G. Alkhazov,[37] A. Alton,[60] G. Alverson,[59] G. A. Alves,[2] M. Anastasoaie,[32] S. Anderson,[42] B. Andrieu,[16] Y. Arnoud,[13] A. Askew,[74] B. Åsman,[38] O. Atramentov,[53] C. Autermann,[20] C. Avila,[7] F. Badaud,[12] A. Baden,[57] B. Baldin,[47] P. W. Balm,[31] S. Banerjee,[27] E. Barberis,[59] P. Bargassa,[74] P. Baringer,[54] C. Barnes,[40] J. Barreto,[2] J. F. Bartlett,[47] U. Bassler,[16] D. Bauer,[51] A. Bean,[54] S. Beauceron,[16] M. Begel,[66] A. Bellavance,[63] S. B. Beri,[26] G. Bernardi,[16] R. Bernhard,[47,*] I. Bertram,[39] M. Besançon,[17] R. Beuselinck,[40] V. A. Bezzubov,[36] P. C. Bhat,[47] V. Bhatnagar,[26] M. Binder,[24] K. M. Black,[58] I. Blackler,[40] G. Blazey,[49] F. Blekman,[31] S. Blessing,[46] D. Bloch,[18] U. Blumenschein,[22] A. Boehnlein,[47] O. Boeriu,[52] T. A. Bolton,[55] F. Borcherding,[47] G. Borissov,[39] K. Bos,[31] T. Bose,[65] A. Brandt,[72] R. Brock,[61] G. Brooijmans,[65] A. Bross,[47] N. J. Buchanan,[46] D. Buchholz,[50] M. Buehler,[48] V. Buescher,[22] S. Burdin,[47] T. H. Burnett,[76] E. Busato,[16] J. M. Butler,[58] J. Bystricky,[17] W. Carvalho,[3] B. C. K. Casey,[71] N. M. Cason,[52] H. Castilla-Valdez,[30] S. Chakrabarti,[27] D. Chakraborty,[49] K. M. Chan,[66] A. Chandra,[27] D. Chapin,[71] F. Charles,[18] E. Cheu,[42] L. Chevalier,[17] D. K. Cho,[66] S. Choi,[45] T. Christiansen,[24] L. Christofek,[54] D. Claes,[63] B. Clément,[18] C. Clément,[38] Y. Coadou,[5] M. Cooke,[74] W. E. Cooper,[47] D. Coppage,[54] M. Corcoran,[74] J. Coss,[19] A. Cothenet,[14] M.-C. Cousinou,[14] S. Crépé-Renaudin,[13] M. Cristetiu,[45] M. A. C. Cummings,[49] D. Cutts,[71] H. da Motta,[2] B. Davies,[39] G. Davies,[40] G. A. Davis,[50] K. De,[72] P. de Jong,[31] S. J. de Jong,[32] E. De La Cruz-Burelo,[30] C. De Oliveira Martins,[3] S. Dean,[41] F. Déliot,[17] P. A. Delsart,[19] M. Demarteau,[47] R. Demina,[66] P. Demine,[17] D. Denisov,[47] S. P. Denisov,[36] S. Desai,[67] H. T. Diehl,[47] M. Diesburg,[47] M. Doidge,[39] H. Dong,[67] S. Doulas,[59] L. Duflot,[15] S. R. Dugad,[27] A. Duperrin,[14] J. Dyer,[61] A. Dyshkant,[49] M. Eads,[49] D. Edmunds,[61] T. Edwards,[41] J. Ellison,[45] J. Elmsheuser,[24] J. T. Eltzroth,[72] V. D. Elvira,[47] S. Eno,[57] P. Ermolov,[35] O. V. Eroshin,[36] J. Estrada,[47] D. Evans,[40] H. Evans,[65] A. Evdokimov,[34] V. N. Evdokimov,[36] J. Fast,[47] S. N. Fatakia,[58] L. Feligioni,[58] T. Ferbel,[66] F. Fiedler,[24] F. Filthaut,[32] W. Fisher,[64] H. E. Fisk,[47] M. Fortner,[49] H. Fox,[22] W. Freeman,[47] S. Fu,[47] S. Fuess,[47] T. Gadfort,[76] C. F. Galea,[32] E. Gallas,[47] E. Galyaev,[52] C. Garcia,[66] A. Garcia-Bellido,[76] J. Gardner,[54] V. Gavrilov,[34] P. Gay,[12] D. Gelé,[18] R. Gelhaus,[45] K. Genser,[47] C. E. Gerber,[48] Y. Gershtein,[71] G. Ginther,[66] T. Golling,[21] B. Gómez,[70] K. Gounder,[47] A. Goussiou,[52] P. D. Grannis,[67] S. Greder,[18] H. Greenlee,[47] Z. D. Greenwood,[56] E. M. Gregores,[40] Ph. Gris,[12] J.-F. Grivaz,[15] L. Groer,[65] S. Grünendahl,[47] M. W. Grünewald,[28] S. N. Gurzhiev,[36] G. Gutierrez,[47] P. Gutierrez,[70] A. Haas,[65] N. J. Hadley,[57] S. Hagopian,[46] I. Hall,[70] R. E. Hall,[44] C. Han,[60] L. Han,[41] K. Hanagaki,[47] K. Harder,[55] R. Harrington,[59] J. M. Hauptman,[53] R. Hauser,[61] J. Hays,[50] T. Hebbeker,[20] D. Hedin,[49] J. M. Heinmiller,[48] A. P. Heinson,[45] U. Heintz,[58] C. Hensel,[54] G. Hesketh,[59] M. D. Hildreth,[52] R. Hirosky,[75] J. D. Hobbs,[67] B. Hoeneisen,[11] M. Hohlfeld,[23] S. J. Hong,[29] R. Hooper,[71] P. Houben,[31] Y. Hu,[67] J. Huang,[51] I. Iashvili,[45] R. Illingworth,[47] A. S. Ito,[47] S. Jabeen,[54] M. Jaffré,[15] S. Jain,[70] V. Jain,[68] K. Jakobs,[22] A. Jenkins,[40] R. Jesik,[40] K. Johns,[42] M. Johnson,[47] A. Jonckheere,[47] P. Jonsson,[40] H. Jöstlein,[47] A. Juste,[47] M. M. Kado,[43] D. Käfer,[20] W. Kahl,[55] S. Kahn,[68] E. Kajfasz,[14] A. M. Kalinin,[33] J. Kalk,[61] D. Karmanov,[35] J. Kasper,[58] D. Kau,[46] R. Kehoe,[73] S. Kermiche,[14] S. Kesisoglou,[71] A. Khanov,[66] A. Kharchilava,[52] Y. M. Kharzheev,[33] K. H. Kim,[29] B. Klima,[47] M. Klute,[21] J. M. Kohli,[26] M. Kopal,[70] V. M. Korablev,[36] J. Kotcher,[68] B. Kothari,[65] A. Koubarovsky,[35] A. V. Kozelov,[36] J. Kozminski,[61] S. Krzywdzinski,[47] S. Kuleshov,[34] Y. Kulik,[47] S. Kunori,[57] A. Kupco,[17] T. Kurča,[19] S. Lager,[38] N. Lahrichi,[17] G. Landsberg,[71] J. Lazoflores,[46] A.-C. Le Bihan,[18] P. Lebrun,[19] S. W. Lee,[29] W. M. Lee,[46] A. Leflat,[35] F. Lehner,[47,*] C. Leonidopoulos,[65] P. Lewis,[40] J. Li,[72] Q. Z. Li,[47] J. G. R. Lima,[49] D. Lincoln,[47] S. L. Linn,[46] J. Linnemann,[61] V. V. Lipaev,[36] R. Lipton,[47] L. Lobo,[40] A. Lobodenko,[37] M. Lokajicek,[10] A. Lounis,[18] H. J. Lubatti,[76] L. Lueking,[47] M. Lynker,[52] A. L. Lyon,[47] A. K. A. Maciel,[49] R. J. Madaras,[43] P. Mättig,[25] A. Magerkurth,[60] A.-M. Magnan,[13] N. Makovec,[15] P. K. Mal,[27] S. Malik,[56] V. L. Malyshev,[33] H. S. Mao,[6] Y. Maravin,[47] M. Martens,[47] S. E. K. Mattingly,[71] A. A. Mayorov,[36] R. McCarthy,[67] R. McCroskey,[42] D. Meder,[23] H. L. Melanson,[47] A. Melnitchouk,[62] M. Merkin,[35] K. W. Merritt,[47] A. Meyer,[20] H. Miettinen,[74] D. Mihalcea,[49] J. Mitrevski,[65] N. Mokhov,[47] J. Molina,[3] N. K. Mondal,[27] H. E. Montgomery,[47] R. W. Moore,[5] G. S. Muanza,[19] M. Mulders,[47] Y. D. Mutaf,[67] E. Nagy,[14] M. Narain,[58] N. A. Naumann,[32] H. A. Neal,[60] J. P. Negret,[7] S. Nelson,[46] P. Neustroev,[37] C. Noeding,[22] A. Nomerotski,[47] S. F. Novaes,[4] T. Nunnemann,[24] E. Nurse,[41] V. O'Dell,[47] D. C. O'Neil,[5] V. Oguri,[3] N. Oliveira,[3] N. Oshima,[47] G. J. Otero y Garzón,[48] P. Padley,[74] N. Parashar,[56] J. Park,[29] S. K. Park,[29] J. Parsons,[65] R. Partridge,[71] N. Parua,[67] A. Patwa,[68] P. M. Perea,[45] E. Perez,[17] O. Peters,[31] P. Pétroff,[15] M. Petteni,[40] L. Phaf,[31] R. Piegaia,[1] P. L. M. Podesta-Lerma,[30] V. M. Podstavkov,[47] Y. Pogorelov,[52] B. G. Pope,[61] W. L. Prado da Silva,[3] H. B. Prosper,[46] S. Protopopescu,[68] M. B. Przybycien,[50,†] J. Qian,[60]

A. Quadt,[21] B. Quinn,[62] K. J. Rani,[27] P. A. Rapidis,[47] P. N. Ratoff,[39] N. W. Reay,[55] S. Reucroft,[59] M. Rijssenbeek,[67] I. Ripp-Baudot,[18] F. Rizatdinova,[55] C. Royon,[17] P. Rubinov,[47] R. Ruchti,[52] G. Sajot,[13] A. Sánchez-Hernández,[30] M. P. Sanders,[41] A. Santoro,[3] G. Savage,[47] L. Sawyer,[56] T. Scanlon,[40] R. D. Schamberger,[67] H. Schellman,[50] P. Schieferdecker,[24] C. Schmitt,[25] A. A. Schukin,[36] A. Schwartzman,[64] R. Schwienhorst,[61] S. Sengupta,[46] H. Severini,[70] E. Shabalina,[48] M. Shamim,[55] V. Shary,[17] W. D. Shephard,[52] D. Shpakov,[59] R. A. Sidwell,[55] V. Simak,[9] V. Sirotenko,[47] P. Skubic,[70] P. Slattery,[66] R. P. Smith,[47] K. Smolek,[9] G. R. Snow,[63] J. Snow,[69] S. Snyder,[68] S. Söldner-Rembold,[41] X. Song,[49] Y. Song,[72] L. Sonnenschein,[58] A. Sopczak,[39] M. Sosebee,[72] K. Soustruznik,[8] M. Souza,[2] B. Spurlock,[72] N. R. Stanton,[55] J. Stark,[13] J. Steele,[56] G. Steinbrück,[65] K. Stevenson,[51] V. Stolin,[34] A. Stone,[48] D. A. Stoyanova,[36] J. Strandberg,[38] M. A. Strang,[72] M. Strauss,[70] R. Ströhmer,[24] M. Strovink,[43] L. Stutte,[47] S. Sumowidagdo,[46] A. Sznajder,[3] M. Talby,[14] P. Tamburello,[42] W. Taylor,[5] P. Telford,[41] J. Temple,[42] S. Tentindo-Repond,[46] E. Thomas,[14] B. Thooris,[17] M. Tomoto,[47] T. Toole,[57] J. Torborg,[52] S. Towers,[67] T. Trefzger,[23] S. Trincaz-Duvoid,[16] B. Tuchming,[17] C. Tully,[64] A. S. Turcot,[68] P. M. Tuts,[65] L. Uvarov,[37] S. Uvarov,[37] S. Uzunyan,[49] B. Vachon,[5] R. Van Kooten,[51] W. M. van Leeuwen,[31] N. Varelas,[48] E. W. Varnes,[42] I. A. Vasilyev,[36] M. Vaupel,[25] P. Verdier,[15] L. S. Vertogradov,[33] M. Verzocchi,[57] F. Villeneuve-Seguier,[40] J.-R. Vlimant,[16] E. Von Toerne,[55] M. Vreeswijk,[31] T. Vu Anh,[15] H. D. Wahl,[46] R. Walker,[40] L. Wang,[57] Z.-M. Wang,[67] J. Warchol,[52] M. Warsinsky,[21] G. Watts,[76] M. Wayne,[52] M. Weber,[47] H. Weerts,[61] M. Wegner,[20] N. Wermes,[21] A. White,[72] V. White,[47] D. Whiteson,[43] D. Wicke,[47] D. A. Wijngaarden,[32] G. W. Wilson,[54] S. J. Wimpenny,[45] J. Wittlin,[58] M. Wobisch,[47] J. Womersley,[47] D. R. Wood,[59] T. R. Wyatt,[41] Q. Xu,[60] N. Xuan,[52] R. Yamada,[47] M. Yan,[57] T. Yasuda,[47] Y. A. Yatsunenko,[33] Y. Yen,[25] K. Yip,[68] S. W. Youn,[50] J. Yu,[72] A. Yurkewicz,[61] A. Zabi,[15] A. Zatserklyaniy,[49] M. Zdrazil,[67] C. Zeitnitz,[23] D. Zhang,[47] X. Zhang,[70] T. Zhao,[76] Z. Zhao,[60] B. Zhou,[60] J. Zhu,[57] M. Zielinski,[66] D. Zieminska,[51] A. Zieminski,[51] R. Zitoun,[67] V. Zutshi,[49] E. G. Zverev,[35] and A. Zylberstejn[17]

(D0 Collaboration)

[1]*Universidad de Buenos Aires, Buenos Aires, Argentina*
[2]*LAFEX, Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil*
[3]*Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil*
[4]*Instituto de Física Teórica, Universidade Estadual Paulista, São Paulo, Brazil*
[5]*Simon Fraser University, Burnaby, Canada, University of Alberta, Edmonton, Canada, McGill University, Montreal, Canada, and York University, Toronto, Canada*
[6]*Institute of High Energy Physics, Beijing, People's Republic of China*
[7]*Universidad de los Andes, Bogotá, Colombia*
[8]*Charles University, Center for Particle Physics, Prague, Czech Republic*
[9]*Czech Technical University, Prague, Czech Republic*
[10]*Institute of Physics, Academy of Sciences, Center for Particle Physics, Prague, Czech Republic*
[11]*Universidad San Francisco de Quito, Quito, Ecuador*
[12]*Laboratoire de Physique Corpusculaire, IN2P3-CNRS, Université Blaise Pascal, Clermont-Ferrand, France*
[13]*Laboratoire de Physique Subatomique et de Cosmologie, IN2P3-CNRS, Universite de Grenoble 1, Grenoble, France*
[14]*CPPM, IN2P3-CNRS, Université de la Méditerranée, Marseille, France*
[15]*Laboratoire de l'Accélérateur Linéaire, IN2P3-CNRS, Orsay, France*
[16]*LPNHE, Universités Paris VI and VII, IN2P3-CNRS, Paris, France*
[17]*DAPNIA/Service de Physique des Particules, CEA, Saclay, France*
[18]*IReS, IN2P3-CNRS, Université Louis Pasteur, Strasbourg, France, and Université de Haute Alsace, Mulhouse, France*
[19]*Institut de Physique Nucléaire de Lyon, IN2P3-CNRS, Université Claude Bernard, Villeurbanne, France*
[20]*RWTH Aachen, III Physikalisches Institut A, Aachen, Germany*
[21]*Universität Bonn, Physikalisches Institut, Bonn, Germany*
[22]*Universität Freiburg, Physikalisches Institut, Freiburg, Germany*
[23]*Universität Mainz, Institut für Physik, Mainz, Germany*
[24]*Ludwig-Maximilians-Universität München, München, Germany*
[25]*Fachbereich Physik, University of Wuppertal, Wuppertal, Germany*
[26]*Panjab University, Chandigarh, India*
[27]*Tata Institute of Fundamental Research, Mumbai, India*
[28]*University College Dublin, Dublin, Ireland*
[29]*Korea Detector Laboratory, Korea University, Seoul, Korea*
[30]*CINVESTAV, Mexico City, Mexico*
[31]*FOM-Institute NIKHEF and University of Amsterdam/NIKHEF, Amsterdam, The Netherlands*
[32]*University of Nijmegen/NIKHEF, Nijmegen, The Netherlands*

[33]*Joint Institute for Nuclear Research, Dubna, Russia*
[34]*Institute for Theoretical and Experimental Physics, Moscow, Russia*
[35]*Moscow State University, Moscow, Russia*
[36]*Institute for High Energy Physics, Protvino, Russia*
[37]*Petersburg Nuclear Physics Institute, St. Petersburg, Russia*
[38]*Lund University, Lund, Sweden, Royal Institute of Technology and Stockholm University, Stockholm, Sweden, and Uppsala University, Uppsala, Sweden*
[39]*Lancaster University, Lancaster, United Kingdom*
[40]*Imperial College, London, United Kingdom*
[41]*University of Manchester, Manchester, United Kingdom*
[42]*University of Arizona, Tucson, Arizona 85721, USA*
[43]*Lawrence Berkeley National Laboratory and University of California, Berkeley, California 94720, USA*
[44]*California State University, Fresno, California 93740, USA*
[45]*University of California, Riverside, California 92521, USA*
[46]*Florida State University, Tallahassee, Florida 32306, USA*
[47]*Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA*
[48]*University of Illinois at Chicago, Chicago, Illinois 60607, USA*
[49]*Northern Illinois University, DeKalb, Illinois 60115, USA*
[50]*Northwestern University, Evanston, Illinois 60208, USA*
[51]*Indiana University, Bloomington, Indiana 47405, USA*
[52]*University of Notre Dame, Notre Dame, Indiana 46556, USA*
[53]*Iowa State University, Ames, Iowa 50011, USA*
[54]*University of Kansas, Lawrence, Kansas 66045, USA*
[55]*Kansas State University, Manhattan, Kansas 66506, USA*
[56]*Louisiana Tech University, Ruston, Louisiana 71272, USA*
[57]*University of Maryland, College Park, Maryland 20742, USA*
[58]*Boston University, Boston, Massachusetts 02215, USA*
[59]*Northeastern University, Boston, Massachusetts 02115, USA*
[60]*University of Michigan, Ann Arbor, Michigan 48109, USA*
[61]*Michigan State University, East Lansing, Michigan 48824, USA*
[62]*University of Mississippi, University, Mississippi 38677, USA*
[63]*University of Nebraska, Lincoln, Nebraska 68588, USA*
[64]*Princeton University, Princeton, New Jersey 08544, USA*
[65]*Columbia University, New York, New York 10027, USA*
[66]*University of Rochester, Rochester, New York 14627, USA*
[67]*State University of New York, Stony Brook, New York 11794, USA*
[68]*Brookhaven National Laboratory, Upton, New York 11973, USA*
[69]*Langston University, Langston, Oklahoma 73050, USA*
[70]*University of Oklahoma, Norman, Oklahoma 73019, USA*
[71]*Brown University, Providence, Rhode Island 02912, USA*
[72]*University of Texas, Arlington, Texas 76019, USA*
[73]*Southern Methodist University, Dallas, Texas 75275, USA*
[74]*Rice University, Houston, Texas 77005, USA*
[75]*University of Virginia, Charlottesville, Virginia 22901, USA*
[76]*University of Washington, Seattle, Washington 98195, USA*

(Received 15 October 2004; published 22 February 2005)

We present the results of a search for the flavor-changing neutral current decay $B_s^0 \to \mu^+\mu^-$ using a data set with integrated luminosity of 240 pb^{-1} of $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV collected with the D0 detector in run II of the Fermilab Tevatron collider. We find the upper limit on the branching fraction to be $\mathcal{B}(B_s^0 \to \mu^+\mu^-) \leq 5.0 \times 10^{-7}$ at the 95% C.L. assuming no contributions from the decay $B_d^0 \to \mu^+\mu^-$ in the signal region. This limit is the most stringent upper bound on the branching fraction $B_s^0 \to \mu^+\mu^-$ to date.

DOI: 10.1103/PhysRevLett.94.071802 PACS numbers: 13.20.He, 12.15.Mm, 14.40.Nd

The purely leptonic decays $B_{d,s}^0 \to \mu^+\mu^-$ [1] are flavor-changing neutral current (FCNC) processes. In the standard model (SM), these decays are forbidden at the tree level and proceed at a very low rate through higher-order diagrams. The SM leptonic branching fractions ($\mathcal{B}$) were calculated including QCD corrections in Ref. [2]. The latest SM prediction [3] is $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.42 \pm 0.54) \times 10^{-9}$, where the error is dominated by

nonperturbative uncertainties. The leptonic branching fraction of the B_d^0 decay is suppressed by Cabibbo-Kobayashi-Maskawa (CKM) matrix elements $|V_{td}/V_{ts}|^2$ leading to a predicted SM branching fraction of $(1.00 \pm 0.14) \times 10^{-10}$. The best existing experimental bound for the branching fraction of $B_s^0(B_d^0)$ is presently $\mathcal{B}(B_s^0(B_d^0) \to \mu^+\mu^-) < 7.5 \times 10^{-7}(1.9 \times 10^{-7})$ at the 95% C.L. [4].

The decay amplitude of $B_{d,s}^0 \to \mu^+\mu^-$ can be significantly enhanced in some extensions of the SM. For instance, in the type-II two-Higgs-doublet model the branching fraction depends only on the charged Higgs boson mass M_{H^+} and $\tan\beta$, the ratio of the two neutral Higgs field vacuum expectation values, with the branching fraction growing as $(\tan\beta)^4$ [5]. In the minimal supersymmetric standard model (MSSM), however, $\mathcal{B}(B_s^0 \to \mu^+\mu^-) \propto (\tan\beta)^6$, leading to an enhancement of up to 3 orders of magnitude [6] compared to the SM, even if the MSSM with minimal flavor violation (MFV) is considered; i.e., the CKM matrix is the only source of flavor violation. An observation of $B_s^0 \to \mu^+\mu^-$ would then immediately lead to an upper bound on the heaviest mass in the MSSM Higgs sector [7] if MFV applies. In minimal supergravity models, an enhancement of $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ is correlated [8] with a sizeable positive shift in $(g-2)_\mu$ that also requires large $\tan\beta$. A large value of $\tan\beta$ is theoretically well motivated by grand unified theories based on minimal SO(10). These models predict large enhancements of $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ as well [8,9]. Finally, FCNC decays of $B_{d,s}^0$ are also sensitive to supersymmetric models with nonminimal flavor violation structures such as the generic MSSM [10] and R parity violating supersymmetry [11].

In this Letter we report on a search for the decay $B_s^0 \to \mu^+\mu^-$ using a data set of integrated luminosity of 240 pb^{-1} recorded with the D0 detector in the years 2002–2004. Our mass resolution is not sufficient to readily separate B_s^0 from B_d^0 leptonic decays. For the final calculation of the upper limit on $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ we assumed that there is no contribution from $B_d^0 \to \mu^+\mu^-$ decays in our search region due to its suppression by $|V_{td}/V_{ts}|^2$, which holds in all models with MFV.

The D0 detector is described in detail elsewhere [12]. The main elements, relevant for this analysis, are the central tracking and muon detector systems. The central tracking system consists of a silicon microstrip tracker (SMT) and a central fiber tracker (CFT), both located within a 2 T superconducting solenoidal magnet. Located outside the calorimeter, the muon detector consists of a layer of tracking detectors and scintillation trigger counters in front of toroidal magnets (1.8 T), followed by two more similar layers behind the toroids, allowing for efficient muon detection out to η of about ± 2, where $\eta = -\ln[\tan(\theta/2)]$ is the pseudorapidity and θ is the polar angle measured relative to the proton beam direction.

Four versions of dimuon triggers were used in the data selection of this analysis. A trigger simulation was used to estimate the trigger efficiency for the signal and normalization samples. These efficiencies were also checked with data samples collected with single muon triggers. The preselection of events was started by requiring two muons of opposite charge which form a common secondary 3D vertex. Each muon candidate had to have $p_T > 2.5$ GeV/c, $|\eta| < 2.0$, and a sufficient number of hits in the central tracking station. In order to select well-measured secondary vertices, we determined the two-dimensional decay length L_{xy} in the plane transverse to the beam line, and required the uncertainty δL_{xy} to be less than 0.15 mm. L_{xy} was calculated as $L_{xy} = (\vec{l}_{vtx} \cdot \vec{p}_T^{\,B}/p_T^B)$, where p_T^B is the transverse momentum of the candidate B_s^0, and $\vec{l}_{vtx}$ represents the vector pointing from the primary vertex to the secondary vertex. The error on the transverse decay length, δL_{xy}, was calculated by taking into account the uncertainties in both the primary and secondary vertex positions. The primary vertex itself was found for each event using a beam-spot constrained fit as described in Ref. [13]. To ensure a similar p_T dependence of the $\mu^+\mu^-$ system in both the signal and the normalization channel, p_T^B had to be greater than 5 GeV/c. A total of 38 167 events survive these preselection requirements. The effects of these criteria on the number of events are shown in Table I.

For the final event selection, we required the candidate events to pass additional criteria. The long lifetime of the B_s^0 mesons allows us to reject random combinatorics background. We therefore used the decay length significance $L_{xy}/\delta L_{xy}$ as one of the discriminating variables, since it gives better discriminating power than the transverse decay length alone, as large values of L_{xy} may originate due to large uncertainties.

The fragmentation characteristics of the b quark are such that most of its momentum is carried by the b hadron. Thus the number of extra tracks near the B_s^0 candidate tends to be small. The second discriminant was therefore an isolation variable, $\mathcal{I}$, of the muon pair, defined as

$$\mathcal{I} = \frac{|\vec{p}(\mu^+\mu^-)|}{|\vec{p}(\mu^+\mu^-)| + \sum\limits_{\text{track } i \neq B} p_i(\Delta\mathcal{R} < 1)}. \qquad (1)$$

Here, $\sum_{\text{track } i \neq B} p_i$ is the scalar sum over all tracks excluding the muon pair within a cone of $\Delta\mathcal{R} < 1$ around the momentum vector $\vec{p}(\mu^+\mu^-)$ of the muon pair where $\Delta\mathcal{R} = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$.

The final discriminating variable was the pointing angle α, defined as the angle between the momentum vector $\vec{p}(\mu^+\mu^-)$ of the muon pair and the vector $\vec{l}_{vtx}$ between the primary and secondary vertices. This requirement ensured consistency between the direction of the decay vertex and the momentum vector of the B_s^0 candidate.

An optimization based on these discriminating variables was done on signal Monte Carlo (MC) events in the B_s^0

TABLE I. Number of candidate events in data satisfying successive preselection requirements.

Variable	Requirement	Number of candidates
Mass (GeV/c^2)	$4.5 < m_{\mu^+\mu^-} < 7.0$	405 307
Muon quality		234 792
χ^2/d.o.f of vertex	<10	146 982
Muon p_T (GeV/c^2)	>2.5	129 558
Muon $\lvert\eta\rvert$	<2.0	125 679
Tracking hits	CFT ≥ 4, SMT ≥ 3	92 678
δL_{xy} (mm)	<0.15	90 935
B_s^0 candidate p_T^B (GeV/c^2)	>5.0	38 167

mass region $4.53 < M_{\mu^+\mu^-} < 6.15$ GeV/c^2 with $m_{B_s^0} = 5369.6 \pm 2.4$ MeV/c^2 [14] and on data events in regions outside the signal window, i.e., in the sidebands. The mass scale throughout this analysis is shifted downward with respect to the world average B_s^0 mass by 30 MeV/c^2 to compensate for the shift in the momentum scale of the D0 tracking system. The mass shift was found by linear interpolation to the B_s^0 mass of the measured mass shifts between the J/ψ and the Υ resonances relative to their world average values [14]. The mass shift is smaller than the MC predicted mass resolution for two-body decays of $\sigma = 90$ MeV/c^2 at the B_s^0 mass.

In order to avoid biasing the optimization procedure, data candidates in the signal mass region were not examined until completion of the analysis, and events in the sideband regions around the B_s^0 mass were used instead. The start (end) of the upper (lower) sideband was chosen such that they were at least 3σ (270 MeV/c^2) away from the B_s^0 mass. The widths of the sidebands used for background estimation were chosen to be 6σ each. The size of the blind signal region was $\pm 3\sigma$ around the B_s^0 mass. To determine the limit on the branching fraction, we used a smaller mass region of $\pm 2\sigma$.

A random-grid search [15] and an optimization procedure [16] were used to find the optimal values of the discriminating variables, by maximizing the variable $P = \epsilon_{\mu\mu}^{B_s^0}/(a/2 + \sqrt{N_{\text{back}}})$. Here, $\epsilon_{\mu\mu}^{B_s^0}$ is the reconstruction efficiency of the signal events relative to the preselection (estimated using MC simulations), and N_{back} is the expected number of background events interpolated from the sidebands. The constant a is the number of standard deviations corresponding to the confidence level at which the signal hypothesis is tested. This constant a was set to 2.0, corresponding to about the 95% C.L. Figure 1 shows the distribution of the three discriminating variables after the preselection for signal MC events and data in the sideband regions. After optimization, we found the following values for the discriminating variables and MC signal efficiencies relative to the preselected sample: $L_{xy}/\delta L_{xy} > 18.5$ (47.5%), $\mathcal{I} > 0.56$ (97.4%), and $\alpha < 0.2$ rad (83.4%). A linear extrapolation of the sideband population for the whole data sample into the (± 180 MeV/c^2) signal region yields an expected number of 3.7 ± 1.1 background events.

Upon examining the data in all mass regions, four events are observed in the signal region, entirely consistent with the background events as estimated from sidebands. We examined the four observed events in detail by studying various kinematic variables, e.g., p_T of the muons, isolation, etc., and found them to be compatible with background events. Figure 2 shows the remaining events populating the lower and upper sidebands as well as the signal region almost equally.

In the absence of an apparent signal, a limit on the branching fraction $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ can be computed by normalizing the upper limit on the number of events in the B_s^0 signal region to the number of reconstructed $B^\pm \to J/\psi K^\pm$ events:

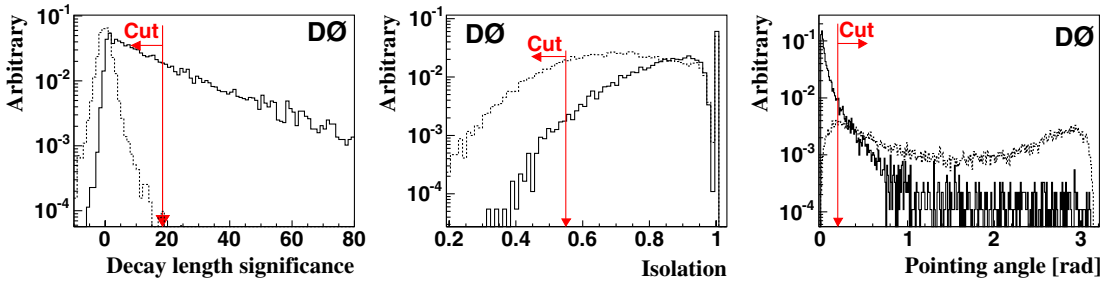

FIG. 1 (color online). Discriminating variables after the preselection for signal MC (solid line) and data events (dashed line) from the sidebands. The arrows indicate the discriminating values that were obtained after optimization. The normalization is done on the number of signal MC and sideband data events after preselection.

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) \leq \frac{N_{\rm ul}}{N_{B^\pm}} \frac{\epsilon^{B^\pm}_{\mu\mu K}}{\epsilon^{B^0_s}_{\mu\mu}} \left[\mathcal{B}(B^\pm \to J/\psi(\mu^+\mu^-)K^\pm) \Big/ \left(\frac{f_{b\to B_s^0}}{f_{b\to B_{u,d}}} + R \frac{\epsilon^{B^0_d}_{\mu\mu}}{\epsilon^{B^0_s}_{\mu\mu}} \right) \right], \quad (2)$$

where $N_{\rm ul}$ is the upper limit on the number of signal decays, estimated from the number of observed events and expected background events and $N_{B^\pm}$ is the observed number $B^\pm \to J/\psi K^\pm$ events. The efficiencies of the signal and normalization channels obtained from MC simulations are $\epsilon^{B^0_s}_{\mu\mu}$ and $\epsilon^{B^\pm}_{\mu\mu K}$, respectively. The factor $\mathcal{B}(B^\pm \to J/\psi(\mu^+\mu^-)K^\pm)$ is the product of the measured branching fractions $\mathcal{B}(B^\pm \to J/\psi K^\pm) = (1.00 \pm 0.04) \times 10^{-3}$ and $\mathcal{B}(J/\psi \to \mu^+\mu^-) = (5.88 \pm 0.10) \times 10^{-2}$ [14]. The ratio $f_{b\to B_s^0}/f_{b\to B_{u,d}}$ of a B_s^0 meson being produced in the fragmentation compared to a $B_{u,d}$ meson is taken to be 0.270 ± 0.034. This ratio has been calculated using the latest world average fragmentation values [14] for B_s^0 and $B_{u,d}$ mesons, where the uncertainty on the ratio is conservatively calculated assuming a full anticorrelation among the individual $B_{u,d}$ and B_s^0 fragmentation uncertainties.

The branching fraction ratio $R = \mathcal{B}(B_d^0)/\mathcal{B}(B_s^0)$ of $B^0_{d,s}$ mesons decaying into two muons multiplied by the total detection efficiency ratio [17] is $R\epsilon^{B^0_d}_{\mu\mu}/\epsilon^{B^0_s}_{\mu\mu}$. Any non-negligible contribution due to B_d^0 decays ($R > 0$) would make the limit on the branching fraction $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ as given in Eq. (2) smaller. Our limit presented for $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ is therefore conservative.

Using the $B^\pm \to J/\psi K^\pm$ mode [18] has the advantage that the efficiencies to detect the $\mu^+\mu^-$ system in signal and normalization events are similar, and systematic effects tend to cancel. A pure sample of $B^\pm \to J/\psi K^\pm$ events was obtained by applying the following selection criteria. The mass-constrained vertex fit of the two muons to form a J/ψ was required to have a $\chi^2/\text{d.o.f.} < 10$, similar to the $\mu^+\mu^-$ vertex criterion in the $B_s^0 \to \mu^+\mu^-$ search. The combined vertex fit of the J/ψ and the additional $K^\pm$ [$p_T(K^\pm) > 0.9$ GeV/c] had to have $\chi^2 < 20$ for 3 d.o.f. The requirements on the three discriminating variables were also applied. The mass spectrum of the reconstructed $B^\pm \to J/\psi K^\pm$ for the full data sample after all analysis requirements is shown in Fig. 3. A fit using a Gaussian for the signal and a second order polynomial for the background yields $741 \pm 31(\text{stat}) \pm 22(\text{syst})$ $B^\pm$ candidates, where the systematic uncertainty was estimated by varying the fit range, background, and signal shape hypotheses.

The p_T distribution of the $B^\pm$ in data has a slightly harder spectrum than that from MC simulations. Therefore, MC events of the signal and normalization channels have been reweighted accordingly. In addition, the observed widths of known $\mu^+\mu^-$ resonances [J/ψ and $\Upsilon(1S)$] are $(27 \pm 4)\%$ larger than predicted by MC simulations. The $\pm 2\sigma$ signal mass region using the MC mass resolution therefore corresponds to $\pm 1.58\sigma$ when the data mass resolution is considered, and the efficiency is corrected accordingly. To within errors, the MC calculation correctly reproduces the efficiency of the cuts on the discriminating variables when applied to the normalization channel.

The final corrected value for the efficiency ratio is then given by $\epsilon^{B^\pm}_{\mu\mu K}/\epsilon^{B^0_s}_{\mu\mu} = 0.247 \pm 0.009(\text{stat}) \pm 0.017(\text{syst})$, where the first uncertainty is due to limited MC statistics and the second accounts for the $B^\pm/B_s^0$ lifetime ratio uncertainties and for uncertainties in data-MC differences. All systematic uncertainties entering the calculation of the branching fraction limit are listed in Table II.

We have used a prescription [19] to construct a confidence interval with the Feldman-Cousins ordering scheme. The expected background was modeled as a Gaussian distribution with its mean value equal to the expected number of background events and its standard deviation equal to the background uncertainty. The uncertainty on the number of $B^\pm$ events as well as the uncertainties on the fragmentation ratio and branching fractions for $B^\pm \to J/\psi(\mu^+\mu^-)K^\pm$ were added in quadrature to the efficiency uncertainties and parametrized as a Gaussian distribution. The resulting branching fraction limit [20] including all the statistical and systematic uncertainties at a 95% (90%) C.L. is given by

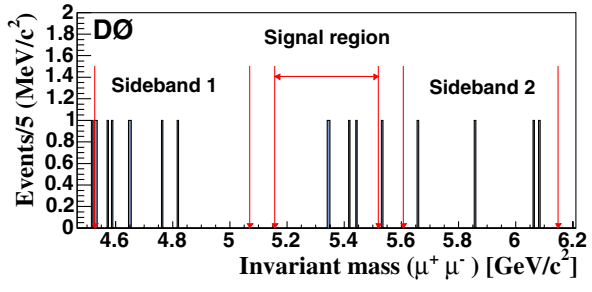

FIG. 2 (color online). Invariant mass of the remaining events of the full data sample after optimized requirements on the discriminating variables.

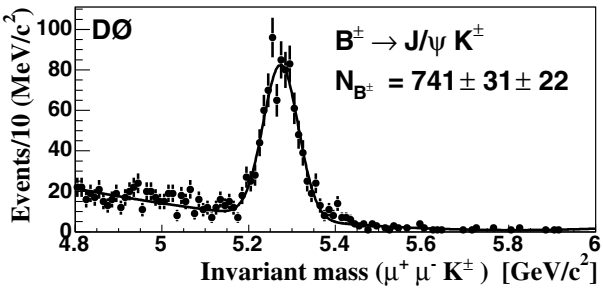

FIG. 3. Invariant mass distribution for candidates in the normalization channel $B^\pm \to J/\psi K^\pm$.

TABLE II. Relative uncertainties used in the calculation of an upper limit of $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$.

Source	Relative uncertainty [%]
$\epsilon_{\mu\mu K}^{B^\pm}/\epsilon_{\mu\mu}^{B_s^0}$	7.7
Number of $B^\pm \to J/\psi K^\pm$ events	5.1
$\mathcal{B}(B^\pm \to J/\psi K^\pm)$	4.0
$\mathcal{B}(J/\psi \to \mu^+\mu^-)$	1.7
$f_{b\to B_s^0}/f_{b\to B_{u,d}^0}$	12.7
Background uncertainty	29.7

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) \leq 5.0 \times 10^{-7} (4.1 \times 10^{-7}).$$

We also used a Bayesian approach with flat prior and Gaussian (smeared) uncertainties [22] and obtained the limit of $\mathcal{B}(B_s^0 \to \mu^+\mu^-) \leq 5.1 \times 10^{-7} (4.1 \times 10^{-7})$ at the 95% (90%) C.L. This new result is presently the most stringent bound on $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$, improving the previously published value [4] and can be used to constrain models of new physics beyond the SM.

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the Department of Energy and National Science Foundation (USA), Commissariat à l'Energie Atomique and CNRS/Institut National de Physique Nucléaire et de Physique des Particules (France), Ministry of Education and Science, Agency for Atomic Energy and RF President Grants Program (Russia), CAPES, CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil), Departments of Atomic Energy and Science and Technology (India), Colciencias (Colombia), CONACyT (Mexico), KRF (Korea), CONICET and UBACyT (Argentina), The Foundation for Fundamental Research on Matter (The Netherlands), PPARC (United Kingdom), Ministry of Education (Czech Republic), Natural Sciences and Engineering Research Council and WestGrid Project (Canada), BMBF and DFG (Germany), A. P. Sloan Foundation, Research Corporation, Texas Advanced Research Program, and the Alexander von Humboldt Foundation.

*Visiting from University of Zurich, Zurich, Switzerland.
†Visiting from Institute of Nuclear Physics, Krakow, Poland.

[1] Charge conjugated states are included implicitly.
[2] G. Buchalla and A. J. Buras, Nucl. Phys. **B400**, 225 (1993); M. Misiak and J. Urban, Phys. Lett. B **451**, 161 (1999); G. Buchalla and A. J. Buras, Nucl. Phys. **B548**, 309 (1999).
[3] A. J. Buras, Phys. Lett. B **566**, 115 (2003).
[4] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. Lett. **93**, 032001 (2004).
[5] H. E. Logan and U. Nierste, Nucl. Phys. **5B86**, 39 (2000).
[6] K. S. Babu and C. F. Kolda, Phys. Rev. Lett. **84**, 228 (2000); A. Dedes *et al.*, Fermilab Report No. FERMILAB-PUB-02-129-T, 2002; S. R. Choudhury and N. Gaur, Phys. Lett. B **451**, 86 (1999).
[7] A. Dedes and T. Huffman, Phys. Lett. B **600**, 261 (2004).
[8] A. Dedes *et al.*, Phys. Rev. Lett. **87**, 251804 (2001).
[9] T. Blazek *et al.*, Phys. Lett. B **589**, 39 (2004); R. Dermisek *et al.*, J. High Energy Phys. 04 (2003) 37.
[10] C. Bobeth *et al.*, Phys. Rev. D **66**, 074021 (2002).
[11] R. Arnowitt *et al.*, Phys. Lett. B **538**, 121 (2002).
[12] D0 Collaboration, V. Abazov *et al.* (unpublished); T. LeCompte and H. T. Diehl, Annu. Rev. Nucl. Part. Sci. **50**, 71 (2000).
[13] DELPHI Collaboration, J. Abdallah *et al.*, Eur. Phys. J. C **32**, 185 (2004).
[14] S. Eidelman *et al.*, Phys. Lett. B **592**, 1 (2004).
[15] N. Amos *et al.*, in *Proceedings of Computing in High Energy Physics (CHEP'95)*, edited by R. Shellard and T. Nguyen (World Scientific, River Edge, NJ, 1996), p. 215.
[16] G. Punzi, in *Proceedings of the Conference on Statistical Problems in Particle Physics, Astrophysics and Cosmology (Phystat 2003)*, edited by L. Lyons *et al.* (SLAC, Menlo Park, CA, 2003), p. 79.
[17] The ratio $\epsilon_{\mu\mu}^{B_d^0}/\epsilon_{\mu\mu}^{B_s^0}$ has been determined from simulation to be 0.92 ± 0.04, with the uncertainty due to limited MC statistics.
[18] In addition to $B^\pm \to J/\psi K^\pm$, the other possible normalization channel is $B_s^0 \to J/\psi\phi$. We have not used it due to low statistics, a large uncertainty on its branching fraction, and a poorly known mixture of *CP* even and *CP* odd decay modes with lifetime differences.
[19] J. Conrad *et al.*, Phys. Rev. D **67**, 012002 (2003); G. J. Feldman and R. D. Cousins, Phys. Rev. D **57**, 3873 (1998).
[20] This limit is derived with the world average fragmentation value [14] of $f_{b\to B_s^0}/f_{b\to B_{u,d}} = 0.270 \pm 0.034$. A fragmentation ratio based on Tevatron data alone [21], but with larger uncertainty, gives $f_{b\to B_s^0}/f_{b\to B_{u,d}} = 0.42 \pm 0.14$ and results in a limit of 4.1×10^{-7} (3.1×10^{-7}) at a 95% (90%) C.L. using the method of Ref. [19].
[21] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. Lett. **84**, 1663 (2000).
[22] I. Bertram *et al.*, Fermilab Report No. FERMILAB-TM-2104, 2000.