Search for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ Decays with CDF II

T. Aaltonen,[22] B. Álvarez González,[10,x] S. Amerio,[42] D. Amidei,[33] A. Anastassov,[37] A. Annovi,[18] J. Antos,[13] G. Apollinari,[16] J. A. Appel,[16] A. Apresyan,[51] T. Arisawa,[63] A. Artikov,[14] J. Asaadi,[57] W. Ashmanskas,[16] B. Auerbach,[66] A. Aurisano,[57] F. Azfar,[41] W. Badgett,[16] A. Barbaro-Galtieri,[27] V. E. Barnes,[51] B. A. Barnett,[24] P. Barria,[48,46] P. Bartos,[13] M. Bauce,[43,42] G. Bauer,[31] F. Bedeschi,[46] D. Beecher,[29] S. Behari,[24] G. Bellettini,[47,46] J. Bellinger,[65] D. Benjamin,[15] A. Beretvas,[16] A. Bhatti,[53] M. Binkley,[16,a] D. Bisello,[43,42] I. Bizjak,[29,bb] K. R. Bland,[5] B. Blumenfeld,[24] A. Bocci,[15] A. Bodek,[52] D. Bortoletto,[51] J. Boudreau,[50] A. Boveia,[12] L. Brigliadori,[7,6] A. Brisuda,[13] C. Bromberg,[34] E. Brucken,[22] M. Bucciantonio,[47,46] J. Budagov,[14] H. S. Budd,[52] S. Budd,[23] K. Burkett,[16] G. Busetto,[43,42] P. Bussey,[20] A. Buzatu,[32] C. Calancha,[30] S. Camarda,[4] M. Campanelli,[29] M. Campbell,[33] F. Canelli,[12,16] B. Carls,[23] D. Carlsmith,[65] R. Carosi,[46] S. Carrillo,[17,l] S. Carron,[16] B. Casal,[10] M. Casarsa,[16] A. Castro,[7,6] P. Catastini,[21] D. Cauz,[58] V. Cavaliere,[23] M. Cavalli-Sforza,[4] A. Cerri,[27,f] L. Cerrito,[29,r] Y. C. Chen,[1] M. Chertok,[8] G. Chiarelli,[46] G. Chlachidze,[16] F. Chlebana,[16] K. Cho,[26] D. Chokheli,[14] J. P. Chou,[21] W. H. Chung,[65] Y. S. Chung,[52] C. I. Ciobanu,[44] M. A. Ciocci,[48,46] A. Clark,[19] C. Clarke,[64] G. Compostella,[43,42] M. E. Convery,[16] J. Conway,[8] M. Corbo,[44] M. Cordelli,[18] C. A. Cox,[8] D. J. Cox,[8] F. Crescioli,[47,46] C. Cuenca Almenar,[66] J. Cuevas,[10,x] R. Culbertson,[16] D. Dagenhart,[16] N. d'Ascenzo,[44,v] M. Datta,[16] P. de Barbaro,[52] S. De Cecco,[54] G. De Lorenzo,[4] M. Dell'Orso,[47,46] C. Deluca,[4] L. Demortier,[53] J. Deng,[15,c] M. Deninno,[6] F. Devoto,[22] M. d'Errico,[43,42] A. Di Canto,[47,46] B. Di Ruzza,[46] J. R. Dittmann,[5] M. D'Onofrio,[28] S. Donati,[47,46] P. Dong,[16] M. Dorigo,[58] T. Dorigo,[42] K. Ebina,[63] A. Elagin,[57] A. Eppig,[33] R. Erbacher,[8] D. Errede,[23] S. Errede,[23] N. Ershaidat,[44,aa] R. Eusebi,[57] H. C. Fang,[27] S. Farrington,[41] M. Feindt,[25] J. P. Fernandez,[30] C. Ferrazza,[49,46] R. Field,[17] G. Flanagan,[51,t] R. Forrest,[8] M. J. Frank,[5] M. Franklin,[21] J. C. Freeman,[16] Y. Funakoshi,[63] I. Furic,[17] M. Gallinaro,[53] J. Galyardt,[11] J. E. Garcia,[19] A. F. Garfinkel,[51] P. Garosi,[48,46] H. Gerberich,[23] E. Gerchtein,[16] S. Giagu,[55,54] V. Giakoumopoulou,[3] P. Giannetti,[46] K. Gibson,[50] C. M. Ginsburg,[16] N. Giokaris,[3] P. Giromini,[18] M. Giunta,[46] G. Giurgiu,[24] V. Glagolev,[14] D. Glenzinski,[16] M. Gold,[36] D. Goldin,[57] N. Goldschmidt,[17] A. Golossanov,[16] G. Gomez,[10] G. Gomez-Ceballos,[31] M. Goncharov,[31] O. González,[30] I. Gorelov,[36] A. T. Goshaw,[15] K. Goulianos,[53] S. Grinstein,[4] C. Grosso-Pilcher,[12] R. C. Group,[62,16] J. Guimaraes da Costa,[21] Z. Gunay-Unalan,[34] C. Haber,[27] S. R. Hahn,[16] E. Halkiadakis,[56] A. Hamaguchi,[40] J. Y. Han,[52] F. Happacher,[18] K. Hara,[60] D. Hare,[56] M. Hare,[61] R. F. Harr,[64] K. Hatakeyama,[5] C. Hays,[41] M. Heck,[25] J. Heinrich,[45] M. Herndon,[65] S. Hewamanage,[5] D. Hidas,[56] A. Hocker,[16] W. Hopkins,[16,g] D. Horn,[25] S. Hou,[1] R. E. Hughes,[38] M. Hurwitz,[12] U. Husemann,[66] N. Hussain,[32] M. Hussein,[34] J. Huston,[34] G. Introzzi,[46] M. Iori,[55,54] A. Ivanov,[8,p] E. James,[16] D. Jang,[11] B. Jayatilaka,[15] E. J. Jeon,[26] M. K. Jha,[6] S. Jindariani,[16] W. Johnson,[8] M. Jones,[51] K. K. Joo,[26] S. Y. Jun,[11] T. R. Junk,[16] T. Kamon,[57,26] P. E. Karchin,[64] A. Kasmi,[5] Y. Kato,[40,o] W. Ketchum,[12] J. Keung,[45] V. Khotilovich,[57] B. Kilminster,[16] D. H. Kim,[26] H. S. Kim,[26] H. W. Kim,[26] J. E. Kim,[26] M. J. Kim,[18] S. B. Kim,[26] S. H. Kim,[60] Y. K. Kim,[12] N. Kimura,[63] M. Kirby,[16] S. Klimenko,[17] K. Kondo,[63,a] D. J. Kong,[26] J. Konigsberg,[17] A. V. Kotwal,[15] M. Kreps,[25] J. Kroll,[45] D. Krop,[12] N. Krumnack,[5,m] M. Kruse,[15] V. Krutelyov,[57,d] T. Kuhr,[25] M. Kurata,[60] S. Kwang,[12] A. T. Laasanen,[51] S. Lami,[46] S. Lammel,[16] M. Lancaster,[29] R. L. Lander,[8] K. Lannon,[38,w] A. Lath,[56] G. Latino,[47,46] T. LeCompte,[2] E. Lee,[57] H. S. Lee,[12] J. S. Lee,[26] S. W. Lee,[57,y] S. Leo,[47,46] S. Leone,[46] J. D. Lewis,[16] A. Limosani,[15,s] C.-J. Lin,[27] J. Linacre,[41] M. Lindgren,[16] E. Lipeles,[45] A. Lister,[19] D. O. Litvintsev,[16] C. Liu,[50] Q. Liu,[51] T. Liu,[16] S. Lockwitz,[66] A. Loginov,[66] D. Lucchesi,[43,42] J. Lueck,[25] P. Lujan,[27] P. Lukens,[16] G. Lungu,[53] J. Lys,[27] R. Lysak,[13] R. Madrak,[16] K. Maeshima,[16] K. Makhoul,[31] S. Malik,[53] G. Manca,[28,b] A. Manousakis-Katsikakis,[3] F. Margaroli,[51] C. Marino,[25] M. Martínez,[4] R. Martínez-Ballarín,[30] P. Mastrandrea,[54] M. E. Mattson,[64] P. Mazzanti,[6] K. S. McFarland,[52] P. McIntyre,[57] R. McNulty,[28,j] A. Mehta,[28] P. Mehtala,[22] A. Menzione,[46] C. Mesropian,[53] T. Miao,[16] D. Mietlicki,[33] A. Mitra,[1] H. Miyake,[60] S. Moed,[21] N. Moggi,[6] M. N. Mondragon,[16,l] C. S. Moon,[26] R. Moore,[16] M. J. Morello,[16] J. Morlock,[25] P. Movilla Fernandez,[16] A. Mukherjee,[16] Th. Muller,[25] P. Murat,[16] M. Mussini,[7,6] J. Nachtman,[16,n] Y. Nagai,[60] J. Naganoma,[63] I. Nakano,[39] A. Napier,[61] J. Nett,[57] C. Neu,[62] M. S. Neubauer,[23] J. Nielsen,[27,e] L. Nodulman,[2] O. Norniella,[23] E. Nurse,[29] L. Oakes,[41] S. H. Oh,[15] Y. D. Oh,[26] I. Oksuzian,[62] T. Okusawa,[40] R. Orava,[22] L. Ortolan,[4] S. Pagan Griso,[43,42] C. Pagliarone,[58] E. Palencia,[10,f] V. Papadimitriou,[16] A. A. Paramonov,[2] J. Patrick,[16] G. Pauletta,[59,58] M. Paulini,[11] C. Paus,[31] D. E. Pellett,[8] A. Penzo,[58] T. J. Phillips,[15] G. Piacentino,[46] E. Pianori,[45] J. Pilot,[38] K. Pitts,[23] C. Plager,[9] L. Pondrom,[65] K. Potamianos,[51] O. Poukhov,[14,a] F. Prokoshin,[14,z] A. Pronko,[16] F. Ptohos,[18,h] E. Pueschel,[11] G. Punzi,[47,46] J. Pursley,[65] A. Rahaman,[50] V. Ramakrishnan,[65] N. Ranjan,[51] I. Redondo,[30] P. Renton,[41] M. Rescigno,[54] T. Riddick,[29] F. Rimondi,[7,6] L. Ristori,[46,16] A. Robson,[20] T. Rodrigo,[10] T. Rodriguez,[45] E. Rogers,[23] S. Rolli,[61,i] R. Roser,[16] M. Rossi,[58] F. Rubbo,[16] F. Ruffini,[48,46] A. Ruiz,[10] J. Russ,[11] V. Rusu,[16] A. Safonov,[57]

W. K. Sakumoto,[52] Y. Sakurai,[63] L. Santi,[59,58] L. Sartori,[46] K. Sato,[60] V. Saveliev,[44,v] A. Savoy-Navarro,[44] P. Schlabach,[16] A. Schmidt,[25] E. E. Schmidt,[16] M. P. Schmidt,[66,a] M. Schmitt,[37] T. Schwarz,[8] L. Scodellaro,[10] A. Scribano,[48,46] F. Scuri,[46] A. Sedov,[51] S. Seidel,[36] Y. Seiya,[40] A. Semenov,[14] F. Sforza,[47,46] A. Sfyrla,[23] S. Z. Shalhout,[8] T. Shears,[28] P. F. Shepard,[50] M. Shimojima,[60,u] S. Shiraishi,[12] M. Shochet,[12] I. Shreyber,[35] A. Simonenko,[14] P. Sinervo,[32] A. Sissakian,[14,a] K. Sliwa,[61] J. R. Smith,[8] F. D. Snider,[16] A. Soha,[16] S. Somalwar,[56] V. Sorin,[4] D. Sperka,[65] P. Squillacioti,[46] M. Stancari,[16] M. Stanitzki,[66] R. St. Denis,[20] B. Stelzer,[32] O. Stelzer-Chilton,[32] D. Stentz,[37] J. Strologas,[36] G. L. Strycker,[33] Y. Sudo,[60] A. Sukhanov,[17] I. Suslov,[14] K. Takemasa,[60] Y. Takeuchi,[60] J. Tang,[12] M. Tecchio,[33] P. K. Teng,[1] J. Thom,[16,g] J. Thome,[11] G. A. Thompson,[23] E. Thomson,[45] P. Ttito-Guzmán,[30] S. Tkaczyk,[16] D. Toback,[57] S. Tokar,[13] K. Tollefson,[34] T. Tomura,[60] D. Tonelli,[16] S. Torre,[18] D. Torretta,[16] P. Totaro,[42] M. Trovato,[49,46] Y. Tu,[45] F. Ukegawa,[60] S. Uozumi,[26] A. Varganov,[33] F. Vázquez,[17,l] G. Velev,[16] C. Vellidis,[3] M. Vidal,[30] I. Vila,[10] R. Vilar,[10] J. Vizán,[10] M. Vogel,[36] G. Volpi,[47,46] P. Wagner,[45] R. L. Wagner,[16] T. Wakisaka,[40] R. Wallny,[9] S. M. Wang,[1] A. Warburton,[32] D. Waters,[29] M. Weinberger,[57] W. C. Wester III,[16] B. Whitehouse,[61] D. Whiteson,[45,c] A. B. Wicklund,[2] E. Wicklund,[16] S. Wilbur,[12] F. Wick,[25] H. H. Williams,[45] J. S. Wilson,[38] P. Wilson,[16] B. L. Winer,[38] P. Wittich,[16,h] S. Wolbers,[16] H. Wolfe,[38] T. Wright,[33] X. Wu,[19] Z. Wu,[5] K. Yamamoto,[40] J. Yamaoka,[15] T. Yang,[16] U. K. Yang,[12,q] Y. C. Yang,[26] W.-M. Yao,[27] G. P. Yeh,[16] K. Yi,[16,n] J. Yoh,[16] K. Yorita,[63] T. Yoshida,[40,k] G. B. Yu,[15] I. Yu,[26] S. S. Yu,[16] J. C. Yun,[16] A. Zanetti,[58] Y. Zeng,[15] and S. Zucchelli[7,7]

(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, I-40127 Bologna, Italy*
[7]*University of Bologna, I-40127 Bologna, Italy*
[8]*University of California, Davis, Davis, California 95616, USA*
[9]*University of California, Los Angeles, Los Angeles, California 90024, USA*
[10]*Instituto de Fisica de Cantabria, CSIC-University of Cantabria, 39005 Santander, Spain*
[11]*Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA*
[12]*Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637, USA*
[13]*Comenius University, 842 48 Bratislava, Slovakia; Institute of Experimental Physics, 040 01 Kosice, Slovakia*
[14]*Joint Institute for Nuclear Research, RU-141980 Dubna, Russia*
[15]*Duke University, Durham, North Carolina 27708, USA*
[16]*Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA*
[17]*University of Florida, Gainesville, Florida 32611, USA*
[18]*Laboratori Nazionali di Frascati, Istituto Nazionale di Fisica Nucleare, I-00044 Frascati, Italy*
[19]*University of Geneva, CH-1211 Geneva 4, Switzerland*
[20]*Glasgow University, Glasgow G12 8QQ, United Kingdom*
[21]*Harvard University, Cambridge, Massachusetts 02138, USA*
[22]*Division of High Energy Physics, Department of Physics, University of Helsinki and Helsinki Institute of Physics, FIN-00014, Helsinki, Finland*
[23]*University of Illinois, Urbana, Illinois 61801, USA*
[24]*The Johns Hopkins University, Baltimore, Maryland 21218, USA*
[25]*Institut für Experimentelle Kernphysik, Karlsruhe Institute of Technology, D-76131 Karlsruhe, Germany*
[26]*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*
[27]*Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[28]*University of Liverpool, Liverpool L69 7ZE, United Kingdom*
[29]*University College London, London WC1E 6BT, United Kingdom*
[30]*Centro de Investigaciones Energeticas Medioambientales y Tecnologicas, E-28040 Madrid, Spain*
[31]*Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[32]*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*
[33]*University of Michigan, Ann Arbor, Michigan 48109, USA*
[34]*Michigan State University, East Lansing, Michigan 48824, USA*

[35]Institution for Theoretical and Experimental Physics, ITEP, Moscow 117259, Russia
[36]University of New Mexico, Albuquerque, New Mexico 87131, USA
[37]Northwestern University, Evanston, Illinois 60208, USA
[38]The Ohio State University, Columbus, Ohio 43210, USA
[39]Okayama University, Okayama 700-8530, Japan
[40]Osaka City University, Osaka 588, Japan
[41]University of Oxford, Oxford OX1 3RH, United Kingdom
[42]Istituto Nazionale di Fisica Nucleare, Sezione di Padova-Trento, I-35131 Padova, Italy
[43]University of Padova, I-35131 Padova, Italy
[44]LPNHE, Universite Pierre et Marie Curie/IN2P3-CNRS, UMR7585, Paris, F-75252 France
[45]University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA
[46]Istituto Nazionale di Fisica Nucleare Pisa, I-56127 Pisa, Italy
[47]University of Pisa, I-56127 Pisa, Italy
[48]University of Siena I-56127 Pisa, Italy
[49]Scuola Normale Superiore, I-56127 Pisa, Italy
[50]University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA
[51]Purdue University, West Lafayette, Indiana 47907, USA
[52]University of Rochester, Rochester, New York 14627, USA
[53]The Rockefeller University, New York, New York 10065, USA
[54]Istituto Nazionale di Fisica Nucleare, Sezione di Roma 1, I-00185 Roma, Italy
[55]Sapienza Università di Roma, I-00185 Roma, Italy
[56]Rutgers University, Piscataway, New Jersey 08855, USA
[57]Texas A&M University, College Station, Texas 77843, USA
[58]Istituto Nazionale di Fisica Nucleare Trieste/Udine, I-34100 Trieste, Italy
[59]University of Udine, I-33100 Udine, Italy
[60]University of Tsukuba, Tsukuba, Ibaraki 305, Japan
[61]Tufts University, Medford, Massachusetts 02155, USA
[62]University of Virginia, Charlottesville, Virginia 22906, USA
[63]Waseda University, Tokyo 169, Japan
[64]Wayne State University, Detroit, Michigan 48201, USA
[65]University of Wisconsin, Madison, Wisconsin 53706, USA
[66]Yale University, New Haven, Connecticut 06520, USA
(Received 10 July 2011; published 1 November 2011; corrected 17 November 2011)

A search has been performed for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays using 7 fb^{-1} of integrated luminosity collected by the CDF II detector at the Fermilab Tevatron collider. The observed number of B^0 candidates is consistent with background-only expectations and yields an upper limit on the branching fraction of $\mathcal{B}(B^0 \to \mu^+\mu^-) < 6.0 \times 10^{-9}$ at 95% confidence level. We observe an excess of B_s^0 candidates. The probability that the background processes alone could produce such an excess or larger is 0.27%. The probability that the combination of background and the expected standard model rate of $B_s^0 \to \mu^+\mu^-$ could produce such an excess or larger is 1.9%. These data are used to determine $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (1.8^{+1.1}_{-0.9}) \times 10^{-8}$ and provide an upper limit of $\mathcal{B}(B_s^0 \to \mu^+\mu^-) < 4.0 \times 10^{-8}$ at 95% confidence level.

DOI: 10.1103/PhysRevLett.107.191801 PACS numbers: 13.20.He, 12.15.Mm, 12.60.Jv

Studies of flavor-changing neutral current (FCNC) decays have played an important role in formulating the theoretical description of particle physics known as the standard model (SM). In the SM all neutral currents conserve flavor so that FCNC decays do not occur at lowest order. The decays of B_s^0 mesons (with a quark content of $\bar{b}s$) and B^0 mesons ($\bar{b}d$) into a dimuon pair ($\mu^+\mu^-$) [1] are examples of FCNC processes that can occur in the SM through higher order loop diagrams. Their branching fractions are predicted in the SM to be $(3.2 \pm 0.2) \times 10^{-9}$ and $(1.0 \pm 0.1) \times 10^{-10}$, respectively [2]. A wide variety of beyond-SM theories predict significant increases over the SM branching fraction [3], making the study of $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays one of the most sensitive indirect searches for new physics. Published upper limits [4–6] contribute significantly to our knowledge of the available new physics parameter space [7–11].

We report a search for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays using $p\bar{p}$ data corresponding to an integrated luminosity of 7 fb^{-1} collected with the Collider Detector at Fermilab (CDF II). The sensitivity of this analysis is significantly improved with respect to the previous analysis [4] due to the higher integrated luminosity of the event sample, a 20% increase in the signal acceptance, and the use of an improved neural-network (NN) discriminant that provides approximately twice the background rejection for the same signal efficiency.

A detailed description of the CDF II detector can be found in Ref. [12]. A charged particle tracking system provides precise vertex determination and momentum measurements in a pseudorapidity range $|\eta| < 1.0$. Additionally, the system measures the ionization per unit path length dE/dx for particle identification. Beyond the tracking detectors are electromagnetic and hadronic calorimeters, which are surrounded by drift chambers used to detect muons in the central region (C) $|\eta| < 0.6$ and the forward region (F) $0.6 < |\eta| < 1.0$.

The online (trigger) requirements used to collect the data sample and the initial set of baseline requirements used in the analysis are the same as those described in Ref. [13]. The events are collected using a set of dimuon triggers [12] and must satisfy either of two sets of requirements corresponding to different topologies: CC events have both muon candidates detected in the central region, while CF events have one central muon and another muon detected in the forward region. Since the expected signal-to-background ratios are different, the two topologies are treated separately. The acceptance of the analysis is improved by 20% by using additional forward muon candidates and by using muon candidates that traverse detector regions previously excluded due to their rapidly changing trigger efficiency. The larger data sample has allowed us to obtain a detailed understanding of the trigger performance in these regions so that we can confidently include these muon candidates in the current analysis. The baseline selection requires high quality muon candidates with transverse momentum relative to the beam direction of $p_T > 2.0(2.2)$ GeV$/c$ in the central (forward) region. The muon pairs are required to have an invariant mass in the range $4.669 < m_{\mu\mu} < 5.969$ GeV$/c^2$ and are constrained to originate from a common well measured 3D vertex. A likelihood method [14] together with a dE/dx based selection [15] are used to further suppress contributions from hadrons misidentified as muons. The baseline requirements also demand that the measured proper decay length of the B candidate λ with its uncertainty σ_λ satisfy $\lambda/\sigma_\lambda > 2$, the 3D opening angle between the momentum of the dimuon pair and the displacement vector between the primary $p\bar{p}$ collision vertex and the dimuon vertex $\Delta\Omega < 0.7$ rad, and the B-candidate track isolation [16] $I > 0.50$. There are 48 279 CC and 52 179 CF muon pairs that fulfill the trigger and baseline selection requirements.

A sample of $B^+ \to J/\psi K^+$ events serves as a normalization mode. The $B^+ \to J/\psi K^+$ sample is collected using the same dimuon triggers and selection requirements so that common systematic uncertainties are suppressed. An additional requirement on the kaon candidate $p_T > 1$ GeV$/c$ is made to limit the p_T range to a region where the tracking efficiency is well understood.

For the final selection, we define search regions around the known B_s^0 and B^0 masses [17]. These regions correspond to approximately $\pm 2.5\sigma_m$, where $\sigma_m \approx 24$ MeV$/c^2$ is the estimated two-track mass resolution. The sideband regions $5.0 < m_{\mu\mu} < 5.169$ GeV$/c^2$ and $5.469 < m_{\mu\mu} < 5.969$ GeV$/c^2$ are used to estimate combinatorial backgrounds. Backgrounds from $B \to h^+ h'^-$ decays (where $h, h' = \pi^\pm$ or $K^\pm$), which peak in the signal mass region, are estimated separately.

Fourteen variables are used to construct a NN discriminant ν_N that ranges from 0 to 1 and enhances the signal-to-background ratio [18]. The variables include dimuon vertex related information (e.g., λ/σ_λ), the impact parameters with respect to the primary vertex and transverse momenta of the muons, the isolation of the B candidate, and the opening angle $\Delta\Omega$. The NN is trained with background events sampled from the sideband regions and signal events generated with a simulation described below. Only a fraction of the total number of background and simulated signal events are used to train the NN. The remainder are used to test for NN overtraining and to determine the signal and background efficiencies. Several tests are done to ensure ν_N is independent of $m_{\mu\mu}$.

All selection criteria were finalized before revealing the content of the signal regions. The optimization of the criteria used the expected upper limit on the $B_s^0 \to \mu^+\mu^-$ branching fraction as a figure of merit. To exploit the difference in the $m_{\mu\mu}$ distributions between signal and background and the improved suppression of combinatorial background at large ν_N, the data are divided into subsamples in the $(\nu_N, m_{\mu\mu})$ plane. The CC and CF samples are each divided into 40 subsamples. There are eight bins in ν_N with bin boundaries 0.70, 0.76, 0.85, 0.90, 0.94, 0.97, 0.987, 0.995, and 1. Within each ν_N bin we employ five $m_{\mu\mu}$ bins, each 24 MeV$/c^2$ wide, centered on the world average $B_s^0(B^0)$ mass. The expected backgrounds and efficiencies are calculated in each bin separately.

For measuring efficiencies, estimating backgrounds, and optimizing the analysis, samples of $B_s^0(B^0) \to \mu^+\mu^-$, $B^+ \to J/\psi K^+$, and $B \to h^+ h'^-$ are generated with the PYTHIA program [19] and a CDF II detector simulation. The p_T spectrum and the I distribution of the B mesons are weighted to match distributions measured in samples of $B^+ \to J/\psi K^+$ and $B_s^0 \to J/\psi \phi$ events [12].

We use a relative normalization to determine the $B_s^0 \to \mu^+\mu^-$ branching fraction:

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = \frac{N_s}{N_+} \frac{\alpha_+}{\alpha_s} \frac{\epsilon_+}{\epsilon_s} \frac{1}{\epsilon_N} \frac{f_+}{f_s} \mathcal{B}(B^+),$$

where N_s is the number of $B_s^0 \to \mu^+\mu^-$ candidate events. The observed number of $B^+ \to J/\psi K^+$ candidates is $N_+ = 22\,388 \pm 196$ and 9943 ± 138 in the CC and CF channels, respectively. The contribution of $B^+ \to J/\psi \pi^+$ events is negligible. We use $\mathcal{B}(B^+) = \mathcal{B}(B^+ \to J/\psi K^+ \to \mu^+\mu^- K^+) = (6.01 \pm 0.21) \times 10^{-5}$ [17] and the ratio of B-meson production fractions $f_+/f_s = 3.55 \pm 0.47$ [17]. The parameter α_s (α_+) is the acceptance of the trigger and ϵ_s (ϵ_+) is the efficiency of the reconstruction requirements for the signal (normalization) mode. The reconstruction efficiency includes trigger, track, muon, and baseline requirement efficiencies. The NN efficiency ϵ_N only applies to the signal mode since it is not used to select the $B^+ \to J/\psi K^+$ sample. The expression for $\mathcal{B}(B^0 \to \mu^+\mu^-)$ is derived by replacing B_s^0 with B^0 and f_+/f_s with $f_+/f_d = 1$. The ratios of acceptances α_+/α_s are 0.307 ± 0.018 and 0.197 ± 0.014 for the CC and CF topologies, respectively. These ratios are measured using simulated events. The uncertainties include contributions from systematic variations of the modeling of the B-meson p_T distributions and the longitudinal beam profile. The ratio of reconstruction efficiencies is $\epsilon_+/\epsilon_s = 0.81 \pm 0.03$ as determined from studies using samples of $J/\psi \to \mu^+\mu^-$ and $B^+ \to J/\psi K^+$ events collected with the same triggers. The uncertainty in ϵ_+/ϵ_s is dominated by kinematic differences between $J/\psi \to \mu^+\mu^-$ and $B_s^0(B^0) \to \mu^+\mu^-$ decays. The ϵ_N is estimated from the simulation. We assign a relative systematic uncertainty on ϵ_N of 4%–7%, depending on ν_N bin, using comparisons of the NN performance in simulated and observed $B^+ \to J/\psi K^+$ event samples, and the statistical uncertainty on studies of the p_T and I distributions from observed $B_s^0 \to J/\psi \phi$ event samples. The $B^0 \to \mu^+\mu^-$ decay is determined to have the same acceptances and efficiencies. Treating CC and CF together, about 90% of simulated $B_s^0 \to \mu^+\mu^-$ events surviving the initial requirements have $\nu_N > 0.70$, with about 45% having $\nu_N > 0.995$. The expected SM yield of $B_s^0 \to \mu^+\mu^-$ events ranges from 0.05 in the lowest ν_N bin to 1.0 events in the highest ν_N bin summing the CC and CF contributions. The expected SM yield of $B^0 \to \mu^+\mu^-$ events is about 30 times smaller.

The expected background is obtained by summing contributions from the combinatorial background and from $B \to h^+h'^-$ decays. To estimate the combinatorial background, we fit the $m_{\mu\mu}$ distribution of sideband events with $\nu_N > 0.70$ to a linear function. We only use events with $m_{\mu\mu} > 5$ GeV$/c^2$ in order to suppress contributions from $b \to \mu^+\mu^- X$ decays. The slopes are then fixed, and the normalization is determined for each ν_N bin separately using the relevant sideband events. In addition to the statistical uncertainties of the slope and normalization parameters, systematic uncertainties are assigned by comparing results derived using alternative fit functions and ranges. The systematic uncertainties vary from about 7% for the lower ν_N bins to about 45% for the highest ν_N bins. The $B \to h^+h'^-$ contributions are estimated using efficiencies determined from the simulation, probabilities of misidentifying hadrons as muons measured in data, and normalizations derived from their branching fractions [15,17]. The hadron misidentification probabilities are parametrized as a function of hadron p_T and instantaneous luminosity using a $D^0 \to K^-\pi^+$ data sample obtained from $D^{*+} \to D^0\pi^+$ decays. In addition to the statistical uncertainties from the D^0 sample, systematic uncertainties are assigned to account for residual variations of the misidentification probability due to variations in detector performance (primarily arising from occupancy and calibration effects) and for branching fraction uncertainties. For the B_s^0 modes there is an additional uncertainty from f_+/f_s. The estimated $B \to h^+h'^-$ background is approximately one quarter of the total background in the $B^0 \to \mu^+\mu^-$ search while in the $B_s^0 \to \mu^+\mu^-$ search it is a factor of 10 smaller than both the combinatorial background and the SM signal. The expected background is shown in Fig. 1 for the $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ searches. The background estimates are cross-checked using three sets of independent control samples: $\mu^+\mu^-$ events with $\lambda < 0$ and $\mu^\pm\mu^\pm$ events, both of which are dominated by combinatorial backgrounds, and a misidentified-muon enhanced $\mu^+\mu^-$ sample with at least one muon candidate failing the muon quality requirements.

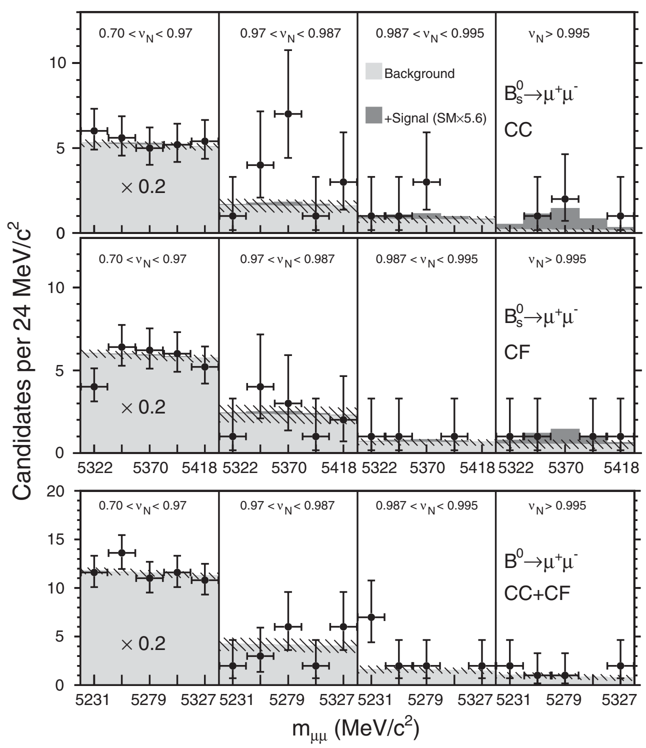

FIG. 1. For the B_s^0 and B^0 signal regions, the observed number of events (points) is compared to the total expected background (light gray) and its uncertainty (hatched) using the $(\nu_N, m_{\mu\mu})$ bins from the optimization. The background uncertainty is the quadrature sum of the relevant systematic uncertainties. The top and middle rows show the results in the B_s^0 mass signal region for the CC and CF channels, respectively. The bottom row shows the results in the B^0 mass signal region for the CC and CF channels combined. The results for the first five ν_N bins are combined (and scaled by 0.2) while the results for the last three bins are each shown separately. Also shown is the expected contribution from $B_s^0 \to \mu^+\mu^-$ events (dark gray) using a branching fraction that corresponds to the central value from the fit to the data, which is 5.6 times the expected SM value.

The latter sample has a significant contribution from $B \to h^+h'^-$ backgrounds. We compare the predicted and observed number of events in each of these control samples for all 80 subsamples and observe no significant discrepancies.

Two fits are performed on the data, a background-only fit (b) and a signal-plus-background fit ($s + b$) for which the branching fraction of the signal is left floating. A log-likelihood ratio is formed, $-2\ln Q$, where $Q = \mathcal{L}(s + b|\text{data})/\mathcal{L}(b|\text{data})$ and $\mathcal{L}(h|x)$ is the likelihood of hypothesis h given observation x; this likelihood is obtained by multiplying Poisson probabilities over all 80 subsamples and is minimized with respect to the nuisance parameters that model our systematic uncertainties. To evaluate the consistency of the data in the signal region with our background model, we compare the observed value of $-2\ln Q$ with the distribution of $-2\ln Q$ obtained from an ensemble of background-only simulated experiments. The effects of systematic uncertainties are included in the simulated experiments by randomly choosing the nuisance parameters from Gaussian distributions. The fraction of simulated experiments with a value of $-2\ln Q$ less than that observed in the data is used to determine the p value for the background-only hypothesis.

The data in the signal regions are shown in Fig. 1 using the $(\nu_N, m_{\mu\mu})$ binning from the optimization. In the B^0 search region the data are consistent with the background prediction and have a p value of 23%. In the B_s^0 search region the data exceed the background prediction and have a p value of 0.27%. The excess is concentrated in bins with $\nu_N > 0.97$. If we restrict ourselves to only the two highest ν_N bins ($\nu_N > 0.987$), which together account for 85% of the signal acceptance, we find a p value of 0.66%. For the $B_s^0 \to \mu^+\mu^-$ analysis we also produce an ensemble of simulated experiments that includes a $B_s^0 \to \mu^+\mu^-$ contribution at the expected SM branching fraction [2] and yields a p value of 1.9%. The corresponding p value for the two highest ν_N bins alone is 4.3%.

We use a modified frequentist approach [20,21] that includes the effects of systematic uncertainties to calculate expected and observed limits. We calculate expected limits of $\mathcal{B}(B^0 \to \mu^+\mu^-) < 4.6 \times 10^{-9}$ and $\mathcal{B}(B_s^0 \to \mu^+\mu^-) < 1.5 \times 10^{-8}$ at the 95% confidence level (C.L.), a factor of 3.3 improvement relative to our previous analysis [4]. We calculate observed limits of $\mathcal{B}(B^0 \to \mu^+\mu^-) < 6.0(5.0) \times 10^{-9}$ and $\mathcal{B}(B_s^0 \to \mu^+\mu^-) < 4.0(3.5) \times 10^{-8}$ at 95% (90%) C.L. If we assume the observed excess in the B_s^0 region is due to signal, we determine $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (1.8^{+1.1}_{-0.9}) \times 10^{-8}$ using the data $-2\ln Q$ distribution and taking the central value from the minimum and the associated uncertainty as the interval corresponding to a change of one unit. By examining the interval corresponding to a change of 2.71 units we set bounds of $4.6 \times 10^{-9} < \mathcal{B}(B_s^0 \to \mu^+\mu^-) < 3.9 \times 10^{-8}$ at the 90% C.L. As a cross-check we use a Bayesian technique to make a point estimate and to derive bounds at 90% C.L. and obtain results very similar to those reported here. Using the central value for the fitted $B_s^0 \to \mu^+\mu^-$ branching fraction, we produce an ensemble of simulated experiments and find a p value of 50%.

The source of the data excess in the $0.970 < \nu_N < 0.987$ bin of the B_s^0 signal region is investigated. The same events, same fits, and same methodologies are used for both the B_s^0 and B^0 searches. Because the data in the B^0 search region show no excess, problems with the background estimates are ruled out. In particular, the only peaking background in this mass region is from $B \to h^+h'^-$ decays, whose contribution to the B^0 search region is 10 times larger than to the B_s^0 search region. Problems with the NN are ruled out by the many studies performed. These NN studies find no evidence of a ν_N-$m_{\mu\mu}$ correlation, no evidence of overtraining, and no evidence of a significant mismodeling of the ν_N shape, even in the region $0.995 < \nu_N$. In short, there is no evidence that the excess in this bin is caused by a mistake or systematic error in our background estimates or our modeling of the ν_N performance and distribution. The most plausible remaining explanation is that this is a statistical fluctuation. For our central result we use the full set of bins that had been established *a priori* since this represents an unbiased choice. As discussed above, if we remove the $0.970 < \nu_N < 0.987$ bin the results are not significantly affected.

In summary, we have performed a search for $B^0 \to \mu^+\mu^-$ and $B_s^0 \to \mu^+\mu^-$ decays using 7 fb^{-1} of integrated luminosity collected by the CDF II detector at the Fermilab Tevatron. The data in the B^0 search region are consistent with background expectations and the world's most stringent upper limit on $\mathcal{B}(B^0 \to \mu^+\mu^-)$ is established. The data in the B_s^0 search region are in excess of the background predictions with a p value of 0.27%. A fit to the data determines $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (1.8^{+1.1}_{-0.9}) \times 10^{-8}$ including all uncertainties.

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 Institut National de Physique Nucleaire et Physique des Particules/CNRS; 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).

[a]Deceased.
[b]Visitor from Istituto Nazionale di Fisica Nucleare, Sezione di Cagliari, 09042 Monserrato (Cagliari), Italy.
[c]Visitor from University of California, Irvine, Irvine, CA 92697, USA.
[d]Visitor from University of California, Santa Barbara, Santa Barbara, CA 93106, USA.
[e]Visitor from University of California, Santa Cruz, Santa Cruz, CA 95064, USA.
[f]Visitor from CERN, CH-1211 Geneva, Switzerland.
[g]Visitor from Cornell University, Ithaca, NY 14853, USA.
[h]Visitor from University of Cyprus, Nicosia CY-1678, Cyprus.
[i]Visitor from Office of Science, U.S. Department of Energy, Washington, D.C. 20585, USA.
[j]Visitor from University College Dublin, Dublin 4, Ireland.
[k]Visitor from University of Fukui, Fukui City, Fukui Prefecture, Japan 910-0017.
[l]Visitor from Universidad Iberoamericana, Mexico D.F., Mexico.
[m]Visitor from Iowa State University, Ames, IA 50011, USA.
[n]Visitor from University of Iowa, Iowa City, IA 52242, USA.
[o]Visitor from Kinki University, Higashi-Osaka City, Japan 577-8502.
[p]Visitor from Kansas State University, Manhattan, KS 66506, USA.
[q]Visitor from University of Manchester, Manchester M13 9PL, United Kingdom.
[r]Visitor from Queen Mary, University of London, London, E1 4NS, United Kingdom.
[s]Visitor from University of Melbourne, Victoria 3010, Australia.
[t]Visitor from Muons, Inc., Batavia, IL 60510, USA.
[u]Visitor from Nagasaki Institute of Applied Science, Nagasaki, Japan.
[v]Visitor from National Research Nuclear University, Moscow, Russia.
[w]Visitor from University of Notre Dame, Notre Dame, IN 46556, USA.
[x]Visitor from Universidad de Oviedo, E-33007 Oviedo, Spain.
[y]Visitor from Texas Tech University, Lubbock, TX 79609, USA.
[z]Visitor from Universidad Tecnica Federico Santa Maria, 110v Valparaiso, Chile.
[aa]Visitor from Yarmouk University, Irbid 211-63, Jordan.
[bb]On leave from J. Stefan Institute, Ljubljana, Slovenia.

[1] Throughout this Letter inclusion of charge conjugate modes is implied.
[2] E. Gamiz *et al.* (HPQCD Collaboration), Phys. Rev. D **80**, 014503 (2009); A. J. Buras, M. V. Carlucci, S. Gori, and G. Isidori, J. High Energy Phys. 10 (2010) 009.
[3] C. Hamzaoui, M. Pospelov, and M. Toharia, Phys. Rev. D **59**, 095005 (1999); S. R. Choudhury and N. Gaur, Phys. Lett. B **451**, 86 (1999); K. S. Babu and C. Kolda, Phys. Rev. Lett. **84**, 228 (2000).
[4] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **100**, 101802 (2008).
[5] V. Abazov *et al.* (D0 Collaboration), Phys. Lett. B **693**, 539 (2010).
[6] R. Aaij *et al.* (LHCb Collaboration), Phys. Lett. B **699**, 330 (2011).
[7] A. Dedes, H. K. Dreiner, and U. Nierste, Phys. Rev. Lett. **87**, 251804 (2001); R. Arnowitt *et al.*, Phys. Lett. B **538**, 121 (2002).
[8] S. Baek *et al.*, J. High Energy Phys. 06 (2005) 017.
[9] R. Ruiz de Austri, R. Trotta, and L. Roszkowski, J. High Energy Phys. 05 (2006) 002; J. Ellis *et al.*, J. High Energy Phys. 05 (2006) 063.
[10] S. Baek, P. Ko, and W. Y. Song, Phys. Rev. Lett. **89**, 271801 (2002).
[11] A. Buras *et al.*, J. Phys. Conf. Ser. **171**, 012004 (2009).
[12] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **71**, 032001 (2005).
[13] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. Lett. **93**, 032001 (2004).
[14] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. Lett. **97**, 242003 (2006).
[15] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **106**, 181802 (2011).
[16] $I = |\vec{p}_T^{\,\mu\mu}|/(\sum_i p_T^i + |\vec{p}_T^{\,\mu\mu}|)$, where $\vec{p}^{\,\mu\mu}$ is the momentum of the dimuon pair; the sum is over all tracks with $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \leq 1$; $\Delta\phi$ and $\Delta\eta$ are the relative azimuthal angle and pseudorapidity of track i with respect to $\vec{p}^{\,\mu\mu}$.
[17] K. Nakamura *et al.* (Particle Data Group), J. Phys. G **37**, 075021 (2010).
[18] M. Feindt and U. Kerzel, Nucl. Instrum. Methods Phys. Res., Sect. A **559**, 190 (2006).
[19] T. Sjöstrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001).
[20] T. Junk, Nucl. Instrum. Methods Phys. Res., Sect. A **434**, 435 (1999).
[21] A. L. Read, CERN Yellow Report No. 2000-005, 2000.