5-5-2005

Measurements of bottom-antibottom azimuthal production correlations in proton-antiproton collisions at √s =1.8 TeV

Darin Acosta
University of Florida-Gainesville, acosta@phys.ufl.edu

Kenneth A. Bloom
University of Nebraska-Lincoln, kenbloom@unl.edu

Collider Detector at Fermilab Collaboration

Measurements of bottom-antibottom azimuthal production correlations in proton-antiproton collisions at $\sqrt{s} = 1.8$ TeV

D. Acosta,[1] T. Affolder,[2] M. G. Albrow,[3] D. Ambrose,[4] D. Amidei,[5] K. Anikeev,[6] J. Antos,[7] G. Apollinari,[3] T. Arisawa,[8] A. Artikov,[9] W. Ashmanskas,[10] F. Azfar,[11] P. Azzi-Bacchetta,[12] N. Bacchetta,[12] H. Bachacou,[13] W. Badgett,[3] A. Barbaro-Galtieri,[13] V. E. Barnes,[14] B. A. Barnett,[15] S. Baroiant,[16] M. Barone,[17] G. Bauer,[6] F. Bedeschi,[18] S. Behari,[15] S. Belforte,[19] W. H. Bell,[20] G. Bellettini,[18] J. Bellinger,[21] D. Benjamin,[22] A. Beretvas,[3] A. Bhatti,[23] M. Binkley,[3] D. Bisello,[12] M. Bishai,[3] R. E. Blair,[10] C. Blocker,[24] K. Bloom,[5] B. Blumenfeld,[15] A. Bocci,[23] A. Bodek,[25] G. Bolla,[14] A. Bolshov,[6] D. Bortoletto,[14] J. Boudreau,[26] C. Bromberg,[27] E. Brubaker,[13] J. Budagov,[9] H. S. Budd,[25] K. Burkett,[3] G. Busetto,[12] K. L. Byrum,[10] S. Cabrera,[22] M. Campbell,[5] W. Carithers,[13] D. Carlsmith,[21] A. Castro,[28] D. Cauz,[19] A. Cerri,[13] L. Cerrito,[29] J. Chapman,[5] C. Chen,[4] Y. C. Chen,[7] M. Chertok,[16] G. Chiarelli,[18] G. Chlachidze,[3] F. Chlebana,[3] M. L. Chu,[7] J. Y. Chung,[30] W.-H. Chung,[21] Y. S. Chung,[25] C. I. Ciobanu,[29] A. G. Clark,[31] M. Coca,[25] A. Connolly,[13] M. Convery,[23] J. Conway,[32] M. Cordelli,[17] J. Cranshaw,[33] R. Culbertson,[3] D. Dagenhart,[24] S. D'Auria,[20] P. de Barbaro,[25] S. De Cecco,[34] S. Dell'Agnello,[17] M. Dell'Orso,[18] S. Demers,[25] L. Demortier,[23] M. Deninno,[28] D. De Pedis,[34] P. F. Derwent,[3] C. Dionisi,[34] J. R. Dittmann,[3] A. Dominguez,[13] S. Donati,[18] M. D'Onofrio,[31] T. Dorigo,[12] N. Eddy,[29] R. Erbacher,[3] D. Errede,[29] S. Errede,[29] R. Eusebi,[25] S. Farrington,[20] R. G. Feild,[35] J. P. Fernandez,[14] C. Ferretti,[5] R. D. Field,[1] I. Fiori,[18] B. Flaugher,[3] L. R. Flores-Castillo,[26] G. W. Foster,[3] M. Franklin,[36] J. Friedman,[6] I. Furic,[6] M. Gallinaro,[23] M. Garcia-Sciveres,[13] A. F. Garfinkel,[14] C. Gay,[35] D. W. Gerdes,[5] E. Gerstein,[37] S. Giagu,[34] P. Giannetti,[18] K. Giolo,[14] M. Giordani,[19] P. Giromini,[17] V. Glagolev,[9] D. Glenzinski,[3] M. Gold,[38] N. Goldschmidt,[5] J. Goldstein,[11] G. Gomez,[39] M. Goncharov,[40] I. Gorelov,[38] A. T. Goshaw,[22] Y. Gotra,[26] K. Goulianos,[23] A. Gresele,[28] C. Grosso-Pilcher,[41] M. Guenther,[14] J. Guimaraes da Costa,[36] C. Haber,[13] S. R. Hahn,[3] E. Halkiadakis,[25] R. Handler,[21] F. Happacher,[17] K. Hara,[42] R. M. Harris,[3] F. Hartmann,[43] K. Hatakeyama,[23] J. Hauser,[44] J. Heinrich,[4] M. Hennecke,[43] M. Herndon,[15] C. Hill,[2] A. Hocker,[25] K. D. Hoffman,[41] S. Hou,[7] B. T. Huffman,[11] R. Hughes,[30] J. Huston,[27] C. Issever,[2] J. Incandela,[2] G. Introzzi,[18] M. Iori,[34] A. Ivanov,[25] Y. Iwata,[45] B. Iyutin,[6] E. James,[3] M. Jones,[14] T. Kamon,[40] J. Kang,[5] M. Karagoz Unel,[46] S. Kartal,[3] H. Kasha,[35] Y. Kato,[47] R. D. Kennedy,[3] R. Kephart,[3] B. Kilminster,[25] D. H. Kim,[48] H. S. Kim,[29] M. J. Kim,[37] S. B. Kim,[48] S. H. Kim,[42] T. H. Kim,[6] Y. K. Kim,[41] M. Kirby,[22] L. Kirsch,[24] S. Klimenko,[1] P. Koehn,[30] K. Kondo,[8] J. Konigsberg,[1] A. Korn,[6] A. Korytov,[1] J. Kroll,[4] M. Kruse,[22] V. Krutelyov,[40] S. E. Kuhlmann,[10] N. Kuznetsova,[3] A. T. Laasanen,[14] S. Lami,[23] S. Lammel,[3] J. Lancaster,[22] K. Lannon,[30] M. Lancaster,[49] R. Lander,[16] A. Lath,[32] G. Latino,[38] T. LeCompte,[10] Y. Le,[15] J. Lee,[25] S. W. Lee,[40] N. Leonardo,[6] S. Leone,[18] J. D. Lewis,[3] K. Li,[35] C. S. Lin,[3] M. Lindgren,[44] T. M. Liss,[29] T. Liu,[3] D. O. Litvintsev,[3] N. S. Lockyer,[4] A. Loginov,[50] M. Loreti,[12] D. Lucchesi,[12] P. Lukens,[3] L. Lyons,[11] J. Lys,[13] R. Madrak,[36] K. Maeshima,[3] P. Maksimovic,[15] L. Malferrari,[28] M. Mangano,[18] G. Manca,[11] M. Mariotti,[12] M. Martin,[15] A. Martin,[35] V. Martin,[46] M. Martínez,[3] P. Mazzanti,[28] K. S. McFarland,[25] P. McIntyre,[40] M. Menguzzato,[12] A. Menzione,[18] P. Merkel,[3] C. Mesropian,[23] A. Meyer,[3] T. Miao,[3] R. Miller,[27] J. S. Miller,[5] S. Miscetti,[17] G. Mitselmakher,[1] N. Moggi,[28] R. Moore,[3] T. Moulik,[14] M. Mulhearn,[6] A. Mukherjee,[3] T. Muller,[43] A. Munar,[4] P. Murat,[3] J. Nachtman,[3] S. Nahn,[35] I. Nakano,[45] R. Napora,[15] F. Niell,[5] C. Nelson,[3] T. Nelson,[3] C. Neu,[30] M. S. Neubauer,[6] C. Newman-Holmes,[3] T. Nigmanov,[26] L. Nodulman,[10] S. H. Oh,[22] Y. D. Oh,[48] T. Ohsugi,[45] T. Okusawa,[47] W. Orejudos,[13] C. Pagliarone,[18] F. Palmonari,[18] R. Paoletti,[18] V. Papadimitriou,[33] J. Patrick,[3] G. Pauletta,[19] M. Paulini,[37] T. Pauly,[11] C. Paus,[6] D. Pellett,[16] A. Penzo,[19] T. J. Phillips,[22] G. Piacentino,[18] J. Piedra,[39] K. T. Pitts,[29] A. Pompoš,[14] L. Pondrom,[21] G. Pope,[26] T. Pratt,[11] F. Prokoshin,[9] J. Proudfoot,[10] F. Ptohos,[17] O. Poukhov,[9] G. Punzi,[18] J. Rademacker,[11] A. Rakitine,[6] F. Ratnikov,[32] H. Ray,[5] A. Reichold,[11] P. Renton,[11] M. Rescigno,[34] F. Rimondi,[28] L. Ristori,[18] W. J. Robertson,[22] T. Rodrigo,[39] S. Rolli,[51] L. Rosenson,[6] R. Roser,[3] R. Rossin,[12] C. Rott,[14] A. Roy,[14] A. Ruiz,[39] D. Ryan,[51] A. Safonov,[16] R. St. Denis,[20] W. K. Sakumoto,[25] D. Saltzberg,[44] C. Sanchez,[30] A. Sansoni,[17] L. Santi,[19] S. Sarkar,[34] P. Savard,[52] A. Savoy-Navarro,[3] P. Schlabach,[3] E. E. Schmidt,[3] M. P. Schmidt,[35] M. Schmitt,[46] L. Scodellaro,[12] A. Scribano,[18] A. Sedov,[14] S. Seidel,[38] Y. Seiya,[42] A. Semenov,[9] F. Semeria,[28] M. D. Shapiro,[13] P. F. Shepard,[26] T. Shibayama,[42] M. Shimojima,[42] M. Shochet,[41] A. Sidoti,[12] A. Sill,[33] P. Sinervo,[52] A. J. Slaughter,[35] K. Sliwa,[51] F. D. Snider,[3] R. Snihur,[49] M. Spezziga,[33] F. Spinella,[18] M. Spiropulu,[2] L. Spiegel,[3] A. Stefanini,[18] J. Strologas,[38] D. Stuart,[2] A. Sukhanov,[1] K. Sumorok,[6] T. Suzuki,[42] R. Takashima,[45] K. Takikawa,[42] M. Tanaka,[10] M. Tecchio,[5] R. J. Tesarek,[3] P. K. Teng,[7] K. Terashi,[23] S. Tether,[6] J. Thom,[3] A. S. Thompson,[20] E. Thomson,[30] P. Tipton,[25] S. Tkaczyk,[3] D. Toback,[40] K. Tollefson,[27] D. Tonelli,[18] M. Tönnesmann,[27] H. Toyoda,[47] W. Trischuk,[52] J. Tseng,[6] D. Tsybychev,[1] N. Turini,[18] F. Ukegawa,[42] T. Unverhau,[20] T. Vaiciulis,[25] A. Varganov,[5] E. Vataga,[18] S. Vejcik III,[3] G. Velev,[3] G. Veramendi,[13] R. Vidal,[3] I. Vila,[39] R. Vilar,[39] I. Volobouev,[13]

M. von der Mey,[44] R. G. Wagner,[10] R. L. Wagner,[3] W. Wagner,[43] Z. Wan,[32] C. Wang,[22] M. J. Wang,[7] S. M. Wang,[1] B. Ward,[20] S. Waschke,[20] D. Waters,[49] T. Watts,[32] M. Weber,[13] W. C. Wester III,[3] B. Whitehouse,[51] A. B. Wicklund,[10] E. Wicklund,[3] H. H. Williams,[4] P. Wilson,[3] B. L. Winer,[30] S. Wolbers,[3] M. Wolter,[51] S. Worm,[32] X. Wu,[31] F. Würthwein,[6] U. K. Yang,[41] W. Yao,[13] G. P. Yeh,[3] K. Yi,[15] J. Yoh,[3] T. Yoshida,[47] I. Yu,[48] S. Yu,[4] J. C. Yun,[3] L. Zanello,[34] A. Zanetti,[19] F. Zetti,[13] and S. Zucchelli[28]

(CDF Collaboration)

[1]*University of Florida, Gainesville, Florida 32611, USA*
[2]*University of California at Santa Barbara, Santa Barbara, California 93106, USA*
[3]*Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA*
[4]*University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA*
[5]*University of Michigan, Ann Arbor, Michigan 48109, USA*
[6]*Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[7]*Institute of Physics, Academia Sinica, Taipei, Taiwan 11529, Republic of China*
[8]*Waseda University, Tokyo 169, Japan*
[9]*Joint Institute for Nuclear Research, RU-141980 Dubna, Russia*
[10]*Argonne National Laboratory, Argonne, Illinois 60439, USA*
[11]*University of Oxford, Oxford OX1 3RH, United Kingdom*
[12]*Universita di Padova, Istituto Nazionale di Fisica Nucleare, Sezione di Padova, I-35131 Padova, Italy*
[13]*Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[14]*Purdue University, West Lafayette, Indiana 47907, USA*
[15]*The Johns Hopkins University, Baltimore, Maryland 21218, USA*
[16]*University of California at Davis, Davis, California 95616, USA*
[17]*Laboratori Nazionali di Frascati, Istituto Nazionale di Fisica Nucleare, I-00044 Frascati, Italy*
[18]*Istituto Nazionale di Fisica Nucleare, University and Scuola Normale Superiore of Pisa, I-56100 Pisa, Italy*
[19]*Istituto Nazionale di Fisica Nucleare, University of Trieste, Italy*
[20]*Glasgow University, Glasgow G12 8QQ, United Kingdom*
[21]*University of Wisconsin, Madison, Wisconsin 53706, USA*
[22]*Duke University, Durham, North Carolina 27708*
[23]*Rockefeller University, New York, New York 10021, USA*
[24]*Brandeis University, Waltham, Massachusetts 02254, USA*
[25]*University of Rochester, Rochester, New York 14627, USA*
[26]*University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA*
[27]*Michigan State University, East Lansing, Michigan 48824, USA*
[28]*Istituto Nazionale di Fisica Nucleare, University of Bologna, I-40127 Bologna, Italy*
[29]*University of Illinois, Urbana, Illinois 61801, USA*
[30]*The Ohio State University, Columbus, Ohio 43210, USA*
[31]*University of Geneva, CH-1211 Geneva 4, Switzerland*
[32]*Rutgers University, Piscataway, New Jersey 08855, USA*
[33]*Texas Tech University, Lubbock, Texas 79409, USA*
[34]*Instituto Nazionale de Fisica Nucleare, Sezione di Roma, University di Roma I, "La Sapienza," I-00185 Roma, Italy*
[35]*Yale University, New Haven, Connecticut 06520, USA*
[36]*Harvard University, Cambridge, Massachusetts 02138, USA*
[37]*Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA*
[38]*University of New Mexico, Albuquerque, New Mexico 87131, USA*
[39]*Instituto de Fisica de Cantabria, CSIC-University of Cantabria, 39005 Santander, Spain*
[40]*Texas A&M University, College Station, Texas 77843, USA*
[41]*Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637, USA*
[42]*University of Tsukuba, Tsukuba, Ibaraki 305, Japan*
[43]*Institut für Experimentelle Kernphysik, Universität Karlsruhe, 76128 Karlsruhe, Germany*
[44]*University of California at Los Angeles, Los Angeles, California 90024, USA*
[45]*Hiroshima University, Higashi-Hiroshima 724, Japan*
[46]*Northwestern University, Evanston, Illinois 60208, USA*
[47]*Osaka City University, Osaka 588, Japan*
[48]*Center for High Energy Physics, Kyungpook National University, Taegu 702-701;*
Seoul National University, Seoul 151-742;
and SungKyunKwan University, Suwon 440-746; Korea
[49]*University College London, London WC1E 6BT, United Kingdom*

[50]Institution for Theoretical and Experimental Physics, ITEP, Moscow 117259, Russia
[51]Tufts University, Medford, Massachusetts 02155, USA
[52]Institute of Particle Physics, University of Toronto, Toronto M5S 1A7, Canada
(Received 24 November 2004; published 5 May 2005)

We have measured the azimuthal angular correlation of $b\overline{b}$ production, using 86.5 pb^{-1} of data collected by Collider Detector at Fermilab (CDF) in $p\overline{p}$ collisions at $\sqrt{s} = 1.8$ TeV during 1994–1995. In high-energy $p\overline{p}$ collisions, such as at the Tevatron, $b\overline{b}$ production can be schematically categorized into three mechanisms. The leading-order (LO) process is ''flavor creation,'' where both b and $\overline{b}$ quarks substantially participate in the hard scattering and result in a distinct back-to-back signal in final state. The ''flavor excitation'' and the ''gluon splitting'' processes, which appear at next-leading-order (NLO), are known to make a comparable contribution to total $b\overline{b}$ cross section, while providing very different opening angle distributions from the LO process. An azimuthal opening angle between bottom and antibottom, $\Delta\phi$, has been used for the correlation measurement to probe the interaction creating $b\overline{b}$ pairs. The $\Delta\phi$ distribution has been obtained from two different methods. One method measures the $\Delta\phi$ between bottom hadrons using events with two reconstructed secondary vertex tags. The other method uses $b\overline{b} \rightarrow (J/\psi X)(\ell X')$ events, where the charged lepton (ℓ) is an electron (e) or a muon (μ), to measure $\Delta\phi$ between bottom quarks. The $b\overline{b}$ purity is determined as a function of $\Delta\phi$ by fitting the decay length of the J/ψ and the impact parameter of the ℓ. Both methods quantify the contribution from higher-order production mechanisms by the fraction of the $b\overline{b}$ pairs produced in the same azimuthal hemisphere, f_{toward}. The measured f_{toward} values are consistent with both parton shower Monte Carlo and NLO QCD predictions.

DOI: 10.1103/PhysRevD.71.092001 PACS numbers: 13.85.−t, 12.38.Qk, 14.65.Fy

I. INTRODUCTION

The dominant b quark production mechanism at the Tevatron is believed to be pair production through the strong interaction. However, predictions from next-to-leading order (NLO) perturbative QCD [1] have so far failed to describe the observed b single quark production cross section [2–8]. Differential cross section measurements have also been systematically higher than theoretical predictions [9–12]. Possible explanations for the disagreement between the measured and predicted cross sections involve improved b fragmentation models [13], and non-perturbative $b\overline{b}$ production mechanisms [14], and supersymmetric production mechanisms [15].

Studying $b\overline{b}$ correlations gives additional insight into the effective contributions from higher-order QCD processes to b quark production at the Tevatron. For example, the lowest-order QCD $b\overline{b}$ production diagrams contain only the b and $\overline{b}$ quarks in the final state. Momentum conservation requires that these quarks be produced back-to-back in azimuthal opening angle, $\Delta\phi$, and with balanced momentum transverse to the beam direction, p_T. However, when higher-order QCD processes are considered, the presence of additional light quarks and gluons in the final state allows the $\Delta\phi$ distribution to become more spread out and the b transverse momenta to become more asymmetric. Previous measurements of azimuthal correlation distributions have yielded varying levels of agreement with NLO predictions [9–12,16]. Additional measurements related to $b\overline{b}$ production are needed to determine whether experimental measurements are consistent with the standard model picture of $b\overline{b}$ production.

The NLO QCD calculation of $b\overline{b}$ production includes diagrams from each production mechanism up to $\mathcal{O}(\alpha_S^3)$. The NLO calculation is the lowest-order approach that returns sensible results because certain classes of diagrams which first appear at $\mathcal{O}(\alpha_S^3)$–often referred to as flavor excitation and gluon splitting diagrams (see below)–provide contributions of approximately the same magnitude as the lowest-order diagrams, which are $\mathcal{O}(\alpha_S^2)$. This contribution can be understood by considering the cross section for $gg \rightarrow gg$ which is approximately 2 orders of magnitude larger than the cross section for $gg \rightarrow b\overline{b}$. Higher-order $b\overline{b}$ diagrams can be formed from the leading-order diagram $gg \rightarrow gg$ by adding a $g \rightarrow b\overline{b}$ vertex to either in the initial or final state, but even with the $\mathcal{O}(\alpha_S)$ suppression, these higher-order diagrams still provide contributions that are numerically comparable to the leading-order terms [1,9]. Therefore, higher-order corrections to $b\overline{b}$ production cannot be ignored, and a recent measurement indicates that the higher-order diagrams contribute a factor of 4 above the leading-order term [16].

An alternative approach to estimating the effects of higher-order corrections is the parton shower model implemented by the PYTHIA [17,18] and HERWIG [19] Monte Carlo programs[1]. The parton shower approach is not exact to any order in α_S but rather tries to approximate corrections to all orders by using leading-order matrix elements

[1]HERWIG and PYTHIA use the exact matrix elements for all parton-parton two-to-two scatterings. However, all two-to-N ($N > 2$) processes are estimated using the ''leading-log'' approximation, which becomes exact in the limit of ''soft'' or ''collinear'' emissions. As a result, such Monte Carlo programs are often said to use the ''leading-log approximation.''

for the hard two-to-two QCD scatter and adding addition initial- and final-state radiation using a probabilistic approach. In this approximation, the diagrams for $b\bar{b}$ production can be divided into three categories:

Flavor creation refers to the lowest-order, two-to-two QCD $b\bar{b}$ production diagrams. This process includes $b\bar{b}$ production through $q\bar{q}$ annihilation and gluon fusion, plus higher-order corrections to these processes. Because this production is dominated by two-body final states, it tends to yield $b\bar{b}$ pairs that are back-to-back in $\Delta\phi$ and balanced in p_T.

Flavor excitation refers to diagrams in which a $b\bar{b}$ pair from the quark sea of the proton or antiproton is excited into the final state because one of the quarks from the $b\bar{b}$ pair undergoes a hard QCD interaction with a parton from the other beam particle. Because only one of the quarks in the $b\bar{b}$ pair undergoes the hard scatter, this production mechanism tends to produce b quarks with asymmetric p_T. Often, one of the b quarks will be produced with high rapidity and not be detected in the central region of the detector.

Gluon splitting refers to diagrams where the $b\bar{b}$ pair arises from a $g \rightarrow b\bar{b}$ splitting in the initial or final state. Neither of the quarks from the $b\bar{b}$ pair participate in the hard QCD scatter. Depending on the experimental range of b quark p_T sensitivity, gluon splitting production can yield a $b\bar{b}$ distribution with a peak at small $\Delta\phi$.

Figure 1 illustrates some lowest-order examples of each type of diagram. The general trend is that flavor creation diagrams, being dominated by two-body $b\bar{b}$ final states, tend to produce back-to-back $b\bar{b}$ pairs balanced in p_T, while flavor excitation and gluon splitting, which necessarily involve multiparticle final states including a $b\bar{b}$ pair and light quarks or gluons, produce $b\bar{b}$ pairs that are more smeared out in $\Delta\phi$ and p_T. Categorizing $b\bar{b}$ diagrams in this scheme becomes ambiguous at higher order in perturbation theory. In the parton shower approximation, flavor creation, flavor excitation, and gluon splitting processes can be separated exactly based on how many b quarks participate in the hard two-to-two scatter. Interference terms among the three production mechanisms, as well

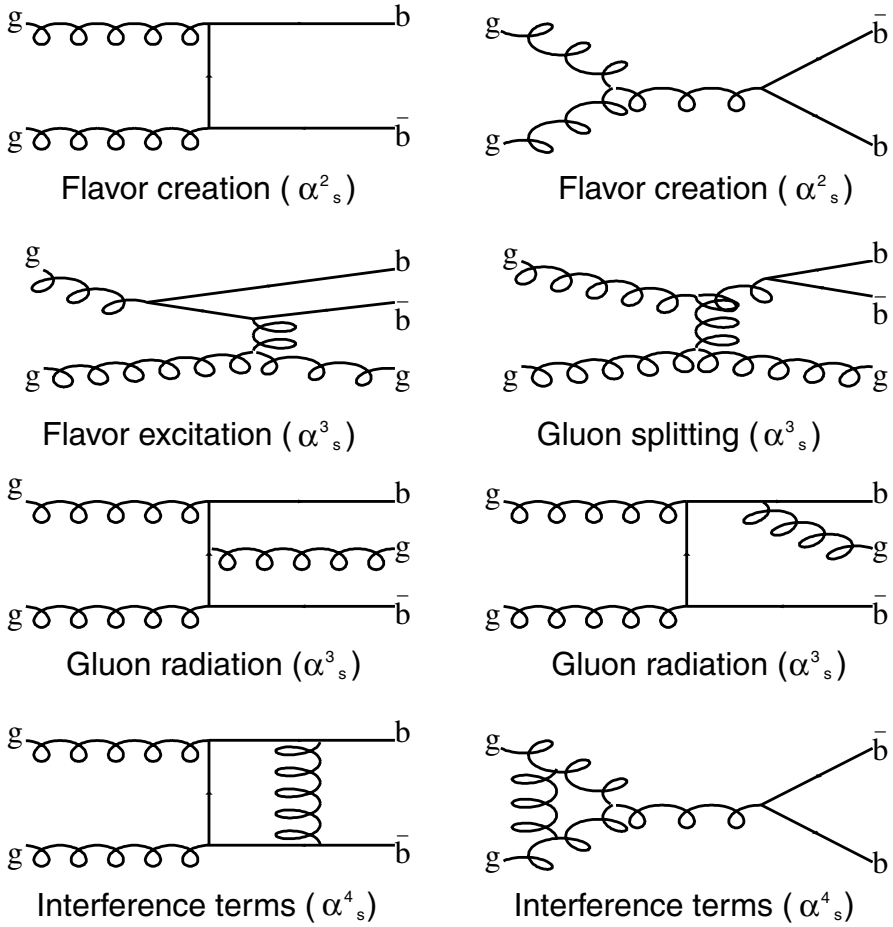

FIG. 1. Example Feynman diagrams that contribute bottom production. The bottom two virtual exchange diagrams enter into the NLO calculation through interferences with leading-order terms. Interferences between the flavor creation, flavor excitation, and gluon splitting diagrams, as well as the virtual exchange diagrams, are ignored in the parton shower approximation.

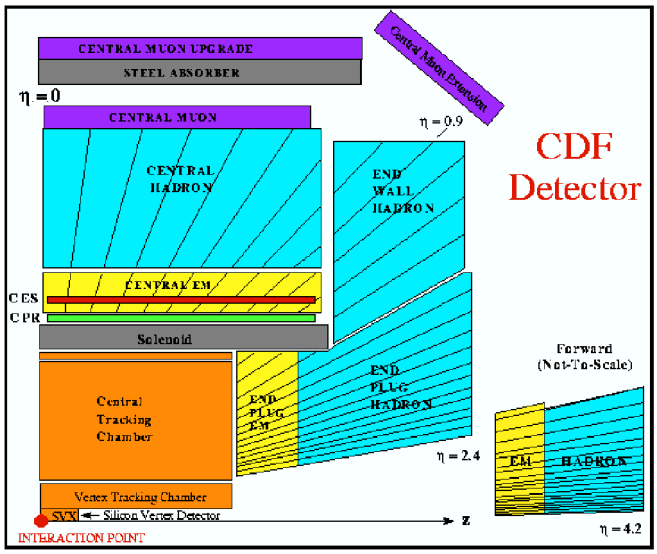

FIG. 2 (color online). Schematic of a quarter cross section of the CDF Run 1b detector.

as virtual exchange diagrams, are neglected as higher-order effects in this approximation.

Refs. [16,20] show that parton shower Monte Carlo programs, which include sizeable contributions from the higher-order b production mechanisms of flavor excitation and gluon splitting, are able to better reproduce the observed b production cross section. Studying $b\bar{b}$ correlations provides a way to tell whether such large contributions from these higher-order processes are supported by the data.

In this paper, we present two new CDF measurements of the $\Delta\phi$ spectrum in $b\bar{b}$ production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV. These measurements are made using approximately 90 pb^{-1} of data collected during the 1994–1995 Tevatron run (known as Run Ib). In addition to providing new information about the entire range of the $b\bar{b}$ $\Delta\phi$ spectrum, these analyses are more sensitive than previous measurements to the low $\Delta\phi$ region, where flavor excitation and gluon splitting make a larger contribution.

One analysis begins with a sample of events containing an 8 GeV electron or muon to enhance the b quark content of the sample by taking advantage of the relatively high semileptonic B branching ratio. These events are then searched for the presence of displaced secondary vertices indicating the decay of a long-lived B hadron, using a vertexing algorithm similar to the SECVTX algorithm used for the top quark analyses [21,22]. This analysis requires that the decay vertices for both B hadrons in the event be reconstructed and extracts the B hadron $\Delta\phi$ distribution from the $\Delta\phi$ distribution measured between the reconstructed secondary vertices. The direction of each B hadron is inferred using the vector sum of the momenta from the secondary vertex tracks and $\Delta\phi$ is defined as the azimuthal angle between the inferred directions of the two B hadrons. This technique yields a high-statistics sample of double-tagged $b\bar{b}$ events and retains sensitivity to $b\bar{b}$ pairs with small opening angles. The second analysis detects the presence of b quark decays in the data entirely through leptonic signatures. The decay of one b is tagged by reconstructing the decay of a $J/\psi \rightarrow \mu^+\mu^-$, which provides the trigger signature that defines this sample. Events are also required to contain an electron or muon consistent with the semileptonic decay of the second b. This approach does not yield as many double-tagged events as the first, but it retains the highest sensitivity for $b\bar{b}$ production at small opening angles and has fewer backgrounds. Both analyses produce consistent results indicating that roughly one fourth of the $b\bar{b}$ pairs produced in the momentum and rapidity range to which these analyses are sensitive have $\Delta\phi < 90°$. In addition, both analyses are at least qualitatively consistent with the contribution from higher-order $b\bar{b}$ predicted by PYTHIA and HERWIG, further supporting the significance of the flavor excitation and gluon splitting production mechanisms at the Tevatron.

II. DETECTOR

The CDF detector has a cylindrical symmetry about the beamline, making it convenient to use a cylindrical coordinate system with the z axis along the proton beam direction. We define r to be the distance from the beamline and ϕ to be the azimuthal angle measured from the direction pointing radially outward in the plane of the Tevatron ring. It is also useful to use the polar angle θ measured with respect to the z axis, and pseudorapidity $\eta = -\ln(\tan(\theta/2))$. In the approximation of massless particles, the pseudorapidity equals the rapidity $y = (1/2)\ln((E + p_z)/(E - p_z))$, which is the invariant boost of the particle along the z axis. The CDF detector is described in detail elsewhere [23]. Figure 2 shows a basic schematic of the CDF detector. In the following, we focus on the elements most relevant to these analyses.

The tracking system, consisting of three different subdetectors, the central tracking chamber (CTC) [24], the vertex detector (VTX), and the silicon vertex detector (SVX′) [25], is immersed in a uniform 1.4 T solenoidal magnetic field in order to measure the charged particle momentum in a plane transverse to the z axis, denoted as $p_T = p\sin\theta$. A charged track reconstruction begins with the measurements made in the CTC, which is a large cylindrical multiwire drift chamber in 3.2 m length along the z axis and centered at the nominal interaction point of the CDF detector. It contains a total of 84 layers of wires positioned between $r = 31$ and 133 cm. The layers are arranged in the alternating groups of 12 with wires parallel to the z axis, known as axial superlayers, and 6 with wires in $\pm 3°$ stereo angles, known as stereo superlayers. The VTX, sitting inside the inner radius of the CTC, is a time projection chamber that provides a precise particle trajectory measurement in the r–z plane and ultimately allows

the determination of the z-location of the primary interaction point. The innermost system is the SVX′ covering from $r = 2.9$ to 8.1 cm. This four-layer detector allows high precision determination of particle trajectories in the r–ϕ plane. Combined, the whole tracking system provides a p_T resolution of $\delta p_T/p_T = [(0.0009 \times p_T)^2 + (0.0066)^2)]^{1/2}$ and an impact parameter resolution of $\delta d_0 = [13 + (40\ \mathrm{GeV}/c)/p_T]\ \mu$m.

The central electro-magnetic calorimeter (CEM) [26], located outside the radius of the CTC and segmented in a projective tower geometry, is designed to be deep enough to contain electro-magnetic showers initiated by electrons or photons. The CEM consists of alternating layers of lead absorber and polystyrene scintillator. A set of wire and strip tracking chambers, known as CES, are embedded in the CEM near the shower maximum or depth of greatest energy deposition, to measure the transverse shower profile. An electron is identified from a track reconstructed in the tracking system that points to an energy deposition in the CEM of appropriate size and matches to a cluster in the CES. For more penetrating particles, the central hadronic calorimeter (CHA) [27] is located behind the CEM. The CHA is constructed from alternating layers of steel absorber and scintillator, and also segmented in a projective tower geometry. The CHA is used in these analyses primarily to reject hadrons that might fake an electron or a muon signature.

Muons are detected by their ability to penetrate the material in the calorimeter. Three sets of chambers are positioned outside the CHA to identify muons. The first set, known as the central muon (CMU) [28] is located at $r = 3.47$ m. A particle traveling perpendicular to the z axis from the primary interaction point must traverse 5.4 pion interaction lengths of material to reach the CMU. An additional set of chambers, the central muon upgrade (CMP) [29] is arranged in a rectangular array around the CMU behind an additional 60 cm of steel shielding to provide further discriminating power between real muons and hadronic punch-through. To penetrate to the CMP, a particle traveling perpendicular to the z axis from the primary interaction point has to pass through 8.4 pion interaction lengths of material. The CMU and the CMP detectors cover $|\eta| < 0.6$. Another set of chambers, the central muon extension (CMX), consisting of four arches of drift chambers located behind 6.2 pion interaction lengths of material, covers $0.6 < |\eta| < 1.0$. In addition, the CMX drift tubes are sandwiched between two layers of scintillator that provide fast timing information to the trigger. Segments reconstructed from hits in the chambers are known as ''stubs.''

The CDF uses a three-level trigger system. The first two levels, named Level-1 and Level-2, are implemented in hardware and reduce the data rate from the full 300 kHz beam crossing rate to a more manageable 20 Hz. The third level, named as Level-3, consists of software algorithms that run a stream-lined version of the full CDF reconstruction software. The triggers used for these analyses rely on lepton identification through matching energy deposition in the CEM (for electron) or muon hits in the CMU, the CMP, and the CMX (for muon) with charged particle tracks reconstructed in the CTC.

III. SECONDARY VERTEX TAG B HADRON CORRELATION ANALYSIS

A. Overview

The $\Delta\phi$ distribution of two reconstructed secondary vertex tags has been obtained from data as a probe to investigate the $b\bar{b}$ production mechanisms and compared to the predictions based on PYTHIA and HERWIG Monte Carlo (MC) programs. We correct our data for detector effects and background contributions using MC information in order to extract the $\Delta\phi$ distribution of B hadrons that can be directly compared to the theoretical predictions. We choose to measure the $\Delta\phi$ distribution of B hadrons rather than b quarks, since our secondary vertex tags are more directly related to B hadrons than b quarks. Converting our measurement from the B hadron level to the b quark level would introduce a dependence on b quark fragmentation models that we wish to avoid.

This analysis uses the largest sample of double-tagged B hadron decays ever collected at a hadron collider, extracted from the data taken by CDF during the 1994–1995 run of the Tevatron (Run Ib). To create a sample enhanced in b quark content, we take advantage of the high purity of CDF lepton triggers as well as the significant impact parameters of B decay daughters. Each candidate event is required to contain a lepton, either an electron or a muon, presumably coming from the semileptonic decay of one B hadron, and the displaced secondary vertices of both B hadrons. After background removal, we obtain a sample of approximately 17 000 events.

B. Secondary vertex tagging

Our secondary vertex tagging algorithm looks for tracks consistent with coming from a secondary vertex, significantly displaced from the primary vertex, using the precise tracking information. This algorithm is based on the BVTX algorithm used for the $B^0 - \overline{B}^0$ mixing analysis [30], which is a modified version of the SECVTX algorithm used for the top quark analysis [21,22]. The main difference between the version of the BVTX used here and the version used for the previous CDF analyses is the ability to locate more than one secondary vertex per jet searched. For extensive details on the BVTX and the modifications made for this analysis, see Refs. [30,31]. Below we summarize the secondary vertex finding approach.

The secondary vertex finding begins by first locating the primary interaction vertex for the event using the precise tracking information. Next the tracks in the event passing quality cuts are grouped into jets using a cone-

based clustering algorithm with a cone size of $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 1.0$. Each jet is then searched for the presence of one or more secondary vertices displaced from the primary. Because the secondary vertex finding is done on a jet-by-jet basis, this algorithm is not able to handle the case where the B decay products are contained in more than one jet. However, the relatively large cone-size used in this analysis was chosen to reduce the number of times the a B decay would span more than one jet. The secondary vertex finding is done in two steps for each jet. The first step finds secondary vertices containing at least three tracks. When the first step fails to find any more secondary vertices in a jet, the second step is attempted in which the individual track cuts are made more stringent and two-track secondary vertices are accepted. Each secondary vertex found is required to be significantly displaced from the primary and not to be consistent with the decay of a K_S^0 or Λ.

C. Sample selection

This analysis starts with the data sample used for the measurement of time dependent $B^0 - \overline{B}^0$ mixing [30], which is a loosely selected sample that requires each event to have at least an electron or a muon with $p_T > 8$ GeV/c identified using the standard CDF lepton identification cuts [31], and at least one reconstructed secondary vertex. This sample is known as the BVTX sample, after the name of the secondary vertex tagging algorithm used to create it. The BVTX sample consists of over 480 000 electron-triggered events and over 430 000 muon-triggered events.

The strategy for extracting candidates from the BVTX sample is as follows: Because the BVTX sample was collected with a number of different lepton triggers, we impose specific trigger requirements to ensure the electron and muon subsamples have comparable kinematic properties. Next, the data sample is reprocessed by the modified version of the BVTX algorithm (see Sec. III B above) and each event is required to contain at least two secondary vertex tags. The separation L_{xy} between each secondary vertex and the primary vertex in the plane perpendicular to the beamline divided by the uncertainty on the measurement ($\sigma_{L_{xy}}$) is required to be $L_{xy}/\sigma_{L_{xy}} \geq 2$. To reduce the chance of tagging the same B decay with two poorly measured tags, the 2-dimensional separation between the secondary vertex tags is also required to be $|\Delta L_{xy}|/\sigma_{\Delta L_{xy}} \geq 2$. ΔL_{xy} is defined to be the distance between the two secondary vertex tags as measured in the plane perpendicular to the beam. Each tag pair is required to have an invariant mass greater than 6 GeV/c^2 to reduce the chance that a tag pair results from a $B \rightarrow D \rightarrow X$ decay chain. For a tag pair failing either the $|\Delta L_{xy}|/\sigma_{\Delta L_{xy}}$ or the invariant mass cuts, only the tag with the longest 2-dimensional separation from the primary vertex is removed. Finally, since the trigger requirements for this sample assume at least one of the B hadrons decay semileptonically, the trigger lepton is required to be within a cone of $\Delta R = 1.0$ of one of the vertices.

D. Sample composition

We have isolated a high purity $b\overline{b}$ sample in Sec. III C with small contamination from other sources. Table I shows the sample composition, including background sources that make a contribution to the sample. We briefly summarize each background contribution below.

A mistag happens when the secondary vertex tagging algorithm tries to fit a vertex from a set of tracks that do not physically originate from a common vertex. Because of errors caused by the tracking performance, it is possible to find a set of prompt tracks that seem to intersect at a vertex displaced from the primary vertex. These vertices distort the correlation spectrum and must be removed. One way to identify mistags is by looking at the distribution of L_{xy}, which is signed based on the inferred direction of the particle, namely, the direction of the secondary vertex, relative to the primary vertex. A particle that seems to be moving out from the primary vertex at the time of decay obtains a positive L_{xy}, while a particle that seems to have been moving towards the primary vertex gets a negative L_{xy}. A particle is deemed to be moving away from the primary vertex if the angle between the tag displacement vector (measured from the primary vertex to the secondary vertex tag) and the tag momentum vector is less than 90°, and towards the primary vertex otherwise. A secondary vertex corresponding to the decay of real long-lived parti-

TABLE I. Different sources of tags and their classification as signal or background.

Scenario	Classification
The tracks in the tag are from the same B decay (including any tracks from a secondary D decay).	Good tag (signal)
The tag contains random prompt tracks not associated with the decay of any long-lived particle.	Mistag (background)
The tracks in the tag are from a B decay (including secondary D decay) that has already been tagged with other tracks.	Sequential double-tag (background)
The tag tracks are from a prompt D decay-in other words, a D not associated with the decay of a B.	Prompt charm (background)

cle is expected to have a positive L_{xy}. However, the finite resolution of the tagging algorithm can yield a negative contribution. As a consequence, mistags make an L_{xy} distribution that is symmetric about zero. We make use of this feature of mistags to subtract them statistically from the data. To understand better how the L_{xy} distribution is used for mistag subtraction, consider the case of an analysis involving only single tags. Half of the total mistag background appears in the negative L_{xy} region. The positive portion of the L_{xy} distribution contains the other half of the mistags, as well as real secondary vertex tags. Therefore, by subtracting twice the number of negative L_{xy} tags from the entire sample of tags, we are left with only the good secondary vertex tags. For analyses such as this one, which considers pairs of secondary vertex tags, the calculation is given by

$$N_{GG} = N_{++} - N_{+-} - N_{-+} + N_{--}, \qquad (1)$$

where N_{GG} is the estimated number of tag pairs in which both tags are legitimate secondary vertex tags, while N_{++} is the number of tag pairs in which both tags have $L_{xy} > 0$, N_{+-} and N_{-+} are the number of tag pairs in which one tag has a positive L_{xy} and the other has a negative L_{xy}, and N_{--} is the number of tag pairs in which both tags have $L_{xy} < 0$. Conceptually, in this equation, we are using the second and third terms to subtract mistags from the tag pairs represented by the first term. However, N_{+-} and N_{-+} each contain a contribution from the case where both tags are mistags and by subtracting them both, this contribution is double-counted. The last term in the equation corrects this. To obtain a mistag-subtracted distribution, Eq. (1) is applied on a bin-by-bin basis.

Another possible source of background involves tagging more than one secondary vertex from a single B decay. These tags, known as sequential tags, are most likely to occur when the B decay involves the production of a D hadron that travels a certain distance from the B decay vertex before itself decaying. The invariant mass cut of 6 GeV$/c^2$ eliminates virtually all contribution from this source. Although this cut does reduce the tagging efficiency at low opening angle, it is necessary to keep the sequential tag background from overwhelming the signal in that region. This efficiency reduction is accounted for in the Monte Carlo modeling of the data. It is also possible that some sequential tag pairs arise from tracking errors that cause tracks originating from a common vertex to be reconstructed as coming from two vertices that are very close together. The cut on the significance of the 2-dimensional separation between the tags ($|\Delta L_{xy}|/\sigma_{\Delta L_{xy}}$) eliminates nearly all these tag pairs. The residual contribution from sequential double tags is estimated in Sec. III G.

Finally, a background source of legitimate secondary vertices is direct $c\bar{c}$ production. In general, most D hadrons have a much smaller lifetime than B hadrons. However, those D hadrons that do live long enough to produce a secondary vertex capable of being tagged by BVTX will not be removed or accounted for by any of the methods mentioned above. In addition, it is possible to have events in which multiple-heavy flavor pairs, such as $b\bar{b} + c\bar{c}$ and $b\bar{b} + b\bar{b}$, are produced. For example, in a flavor creation event, an additional $c\bar{c}$ pair may be produced through gluon splitting. In such events it is possible for the $b\bar{b}$ to contribute one tag and the $c\bar{c}$ to contribute another. Although the rate of multiple-heavy flavor production is much lower than single $b\bar{b}$ production, the opportunity to tag more displaced vertices in a given event can provide an enhancement in tagging acceptance, meaning such processes cannot be discounted outright. Our MC studies indicate that the combined contribution to the tag pair sample from prompt charm and multiple-heavy flavor production is not large, roughly 10%. The subtraction of this contribution and the associated systematic error are described in Sec. III G.

E. Monte Carlo samples

The parton shower MC programs, PYTHIA [17] and HERWIG [19], are used to generate large samples of $b\bar{b}$ events. Because flavor creation, flavor excitation, and gluon splitting mechanisms do not interfere with each other in the parton shower model, each mechanism is generated separately. For PYTHIA, the flavor creation samples are generated as the heavy-flavor production process using massive matrix elements for $q\bar{q} \rightarrow b\bar{b}$ and $gg \rightarrow b\bar{b}$ diagrams. Flavor excitation and gluon splitting samples are produced as the generic QCD 2-to-2 process using massless matrix elements, and then separated from other QCD processes by examining the partons that participate in the 2-to-2 hard scattering. Three PYTHIA samples with different amounts of initial state-radiation (ISR) are generated for each mechanism: The samples are dubbed low, medium, and high ISR, as explained in Appendix A. The choice to investigate different ISR settings in PYTHIA is motivated primarily because the ISR tuning of PYTHIA was changed in the recent past based on studies of heavy-flavor production [32,33], and the new tuning produces a noticeably different $\Delta\phi$ spectrum from the previous version. The low ISR sample corresponds to the most recent ISR settings in PYTHIA while the high ISR sample reflects the previous default settings. In principal, changes in the amount of final state radiation (FSR) would have a similar affect, but such an effect has not been studied here. For all three PYTHIA samples, the underlying event is tuned to match observations in CDF data [34]. On the other hand, because there are fewer parameters to tune, only one HERWIG sample is generated for each mechanism. The HERWIG flavor creation and flavor excitation samples are generated with heavy-flavor production option including massive matrix element treatments of the LO flavor creation and flavor excitation diagrams. As in PYTHIA, the

HERWIG gluon splitting component results from generating all QCD 2-to-2 processes using massless matrix elements and retaining those events classified as gluon splitting based on the partons involved in the hard scattering. In addition, a small sample of $c\bar{c}$ events was generated using PYTHIA, for the purpose of evaluating the possible effects of residual prompt charm as a background for this analysis. Both PYTHIA and HERWIG generation used the CTEQ5L parton distribution functions. See Appendix A for more information about the PYTHIA and HERWIG parameters used for this analysis. For all samples, heavy-flavor decays are handled by the CLEO QQ MC program [35]. Finally, to make the MC data resemble the actual data as closely as possible, the MC events are passed through a CDF detector simulation, a CDF trigger simulation, and the same reconstruction and analysis code used for the actual data. Additional details regarding the generation of MC samples can be found in Ref. [31].

F. Data-Monte Carlo comparison

After the Monte Carlo samples have been passed through the detector simulation as described above, the Monte Carlo predictions for the secondary vertex tag distributions can be compared directly with data. Distributions involving individual tags have similar shapes for flavor creation, flavor excitation, and gluon splitting, and so these distributions can be used to check whether the detector simulation adequately models detector effects. Distributions involving tag pairs, and therefore correlations, give information about how well the Monte Carlo models describe $b\bar{b}$ production.

Figure 3 shows the comparison of the trigger electron p_T and E_T and the trigger muon p_T between the data and each Monte Carlo sample. Figure 4 shows the comparison of the secondary vertex tag properties in the Monte Carlo and data. In each case, the agreement between the measured spectrum from the data and the predicted spectra for each Monte Carlo sample indicates that the effects of trigger and reconstruction thresholds are adequately modeled in the simulation. In addition, examining the Monte Carlo truth information for the B hadrons tagged by the analysis code allows a determination of the effective minimum B p_T sensitivity for this analysis. For the B producing the 8 GeV/c trigger lepton, this measurement is sensitive only to B hadrons with a minimum p_T of 14 GeV/c,

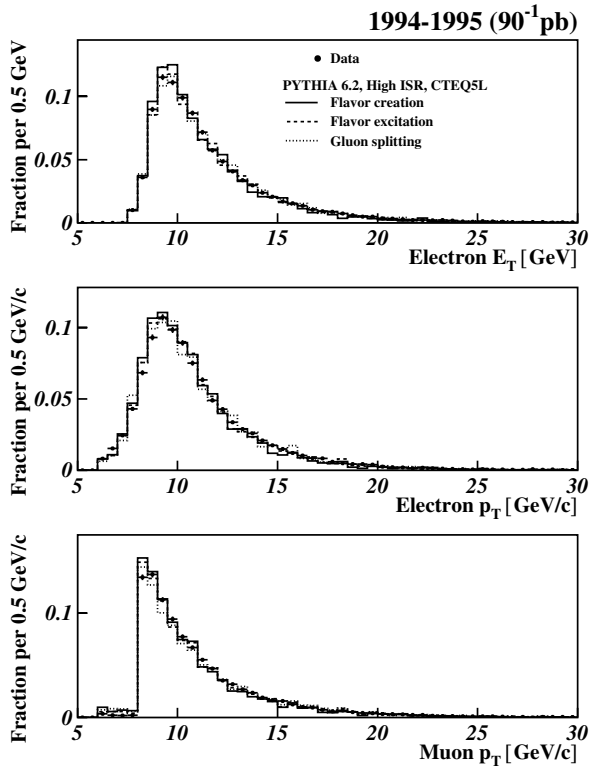

FIG. 3. The trigger electron p_T and E_T and the trigger muon p_T distributions from data compared to the high ISR PYTHIA Monte Carlo sample. The comparisons of data to other Monte Carlo samples is similar and can be found in Ref. [31]. The muon trigger threshold is clearly visible in the lower plot. The small number of muons below the 8 GeV trigger threshold come from events containing a second muon that passes the offline selection.

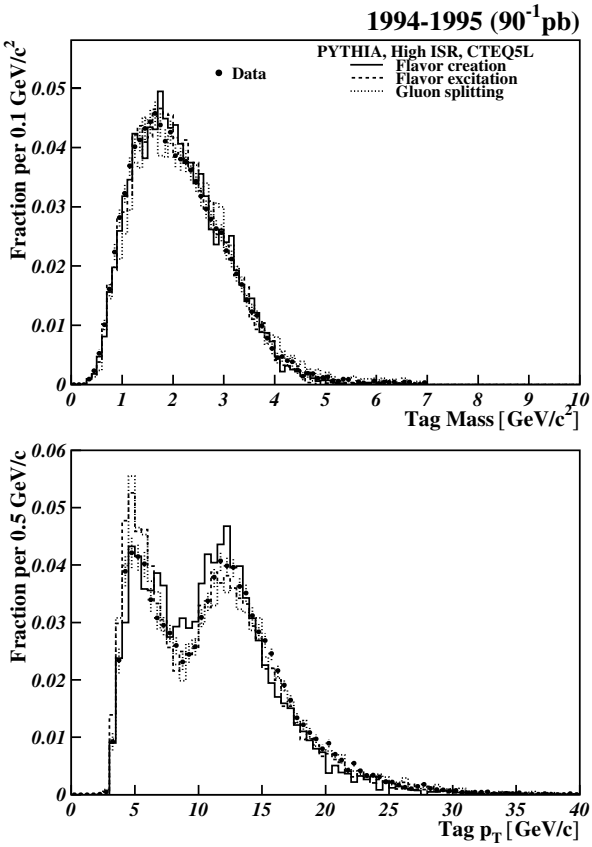

FIG. 4. The secondary vertex tag distributions from electron data compared to the high ISR PYTHIA Monte Carlo sample. Comparisons involving muon data and comparisons to other Monte Carlo samples are similar and can be found in Ref. [31].

while the requirement that the other B be tagged by the BVTX algorithm sets a minimum p_T acceptance of 7.5 GeV/c.

Comparing tag pair correlations between the Monte Carlo samples and the data reveals whether PYTHIA or HERWIG provide an adequate model of the higher-order contributions to $b\bar{b}$ production. This analysis focuses on the transverse opening angle, $\Delta\phi$. For tag pairs, $\Delta\phi$ is defined as the angle between the p_T vectors determined by taking the vector sum of the p_T from all the tracks involved in the tag. The $\Delta\phi$ distribution is interesting to study because it is sensitive to contributions from flavor excitation and gluon splitting. Also, the broadness of the back-to-back peak in $\Delta\phi$ is sensitive to the amount of initial-state radiation present in the Monte Carlo. It should be noted that the shape of the $\Delta\phi$ distribution and the relative contributions from the three production mechanisms depend on the p_T cuts placed on each of the B hadrons.

There are two possible approaches to normalizing the relative contributions in Monte Carlo from flavor creation, flavor excitation, and gluon splitting. PYTHIA and HERWIG each provide predictions for the cross section of each production mechanism, and these cross sections can be used to normalize their contributions relative to one another. Alternatively, one could take the position that PYTHIA and HERWIG may not correctly model the amount of each contribution, and the relative contributions should be determined to provide the best match to data. In this analysis, both approaches are examined. As described in the sections below, the data is compared to the Monte Carlo predictions in two ways. First, the Monte Carlo prediction for the cross section of each production mechanism is used

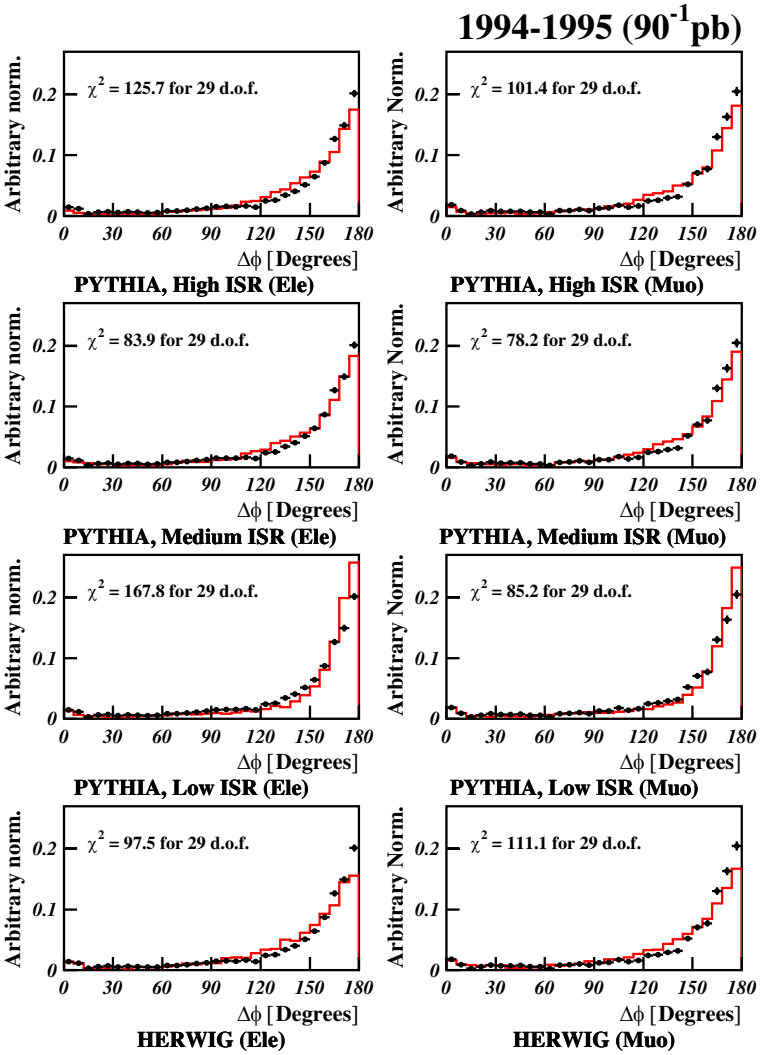

FIG. 5 (color online). Comparisons between the shape of the tag $\Delta\phi$ distribution for data (points) and the tag $\Delta\phi$ distribution predicted by the various Monte Carlo samples (line). The contributions from flavor creation, flavor excitation and gluon splitting are added together according to the individual Monte Carlo cross section predictions for these contributions and only a single common normalization is varied to get the best match to data. The χ^2 values shown in the plots account only for the data and Monte Carlo statistical errors.

TABLE II. Comparison of the effective contributions from flavor creation (FC), flavor excitation (FE), and gluon splitting (GS) to fits of the Monte Carlo $\Delta\phi$ to the data. The fit χ^2 takes into account Monte Carlo statistics in addition to errors on the data. The "χ^2 probability" entry refers to the probability of getting a worse fit, according to the χ^2 distribution.

	Electrons				Muons			
	Fixed normalization		Floating normalization		Fixed normalization		Floating normalization	
PYTHIA	FC:	43.7%	FC:	66.1%	FC:	41.4%	FC:	64.5%
High ISR	FE:	40.7%	FE:	1.7%	FE:	41.5%	FE:	8.1%
	GS:	15.6%	GS:	32.2%	GS:	17.1%	GS:	27.4%
χ^2/d.o.f.	125.7/29		34.1/27		101.4/29		39.3/27	
χ^2 probability	5.26×10^{-14}		0.136		5.83×10^{-10}		0.0595	
PYTHIA	FC:	47.7%	FC:	65.3%	FC:	46.5%	FC:	63.2%
Medium ISR	FE:	35.8%	FE:	0.2%	FE:	35.3%	FE:	8.6%
	GS:	16.5%	GS:	34.5%	GS:	18.2%	GS:	28.2%
χ^2/d.o.f.	83.9/29		38.6/27		78.2/29		34.0/27	
χ^2 probability	3.07×10^{-7}		0.0688		4.11×10^{-6}		0.166	
PYTHIA	FC:	68.3%	FC:	37.6%	FC:	63.9%	FC:	48.5%
Low ISR	FE:	12.0%	FE:	51.4%	FE:	13.9%	FE:	34.6%
	GS:	19.7%	GS:	11.0%	GS:	22.2%	GS:	16.9%
χ^2/d.o.f.	167.8/29		33.9/27		85.2/29		46.0/27	
χ^2 probability	1.76×10^{-21}		0.169		1.96×10^{-7}		0.0127	
HERWIG	FC:	57.6%	FC:	70.9%	FC:	55.7%	FC:	74.8%
	FE:	24.0%	FE:	0.0%	FE:	23.1%	FE:	0.0%
	GS:	18.4%	GS:	29.1%	GS:	21.2%	GS:	25.2%
χ^2/d.o.f.	97.5/29		65.6/27		111.1/29		70.2/27	
χ^2 probability	2.45×10^{-9}		4.65×10^{-5}		1.52×10^{-11}		1.05×10^{-5}	

to normalize the flavor excitation and gluon splitting components relative to the flavor creation contribution. In this "fixed normalization" scheme, the data is compared to the Monte Carlo using one arbitrary global normalization parameter. The arbitrary global normalization is included because this analysis attempts only a shape comparison, not an absolute cross section measurement. In addition, the Monte Carlo and data are compared using a "floating normalization" scheme. In this comparison, each production mechanism is given an independent arbitrary normalization constant and the three normalizations are varied to yield the best match to data.

Figure 5 shows the comparison of the BVTX tag $\Delta\phi$ distribution from data to the distribution predicted by each Monte Carlo sample when the relative normalization of each production mechanism is based on the Monte Carlo prediction for the cross sections of the different production mechanisms. From these $\Delta\phi$ comparisons it can be seen that each Monte Carlo model matches the qualitative features of the data, although there are definite differences in shape, as reflected by the poor χ^2 values. For the PYTHIA sample with low initial-state radiation (ISR), the peak in the back-to-back region is too narrow, while for the medium and high ISR samples, the back-to-back peak is too broad. Similarly, the HERWIG Monte Carlo sample also has a peak that is too broad at high $\Delta\phi$, perhaps even more so than in PYTHIA. However, aside from these discrepancies at

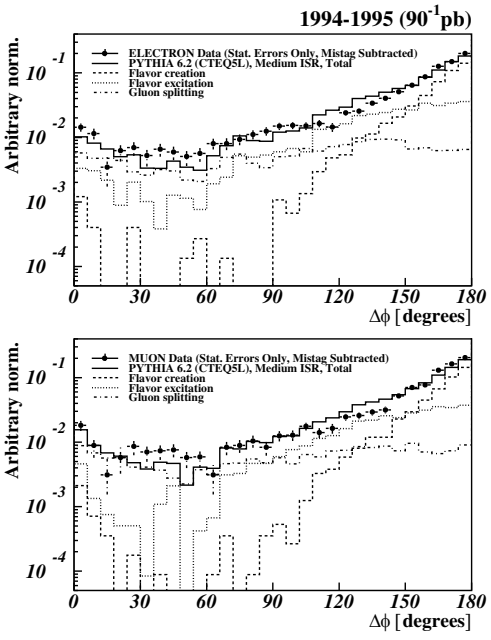

FIG. 6. A detailed comparison between the $\Delta\phi$ distribution from data (points, statistical errors only) and the $\Delta\phi$ distribution from PYTHIA with medium ISR (solid line). In addition, the contributions from flavor creation (dashes), flavor excitation (dots), and gluon splitting (dash-dots) are shown. The contributions are normalized according to PYTHIA's cross section predictions and an arbitrary global normalization is used to give the best shape fit between data and Monte Carlo.

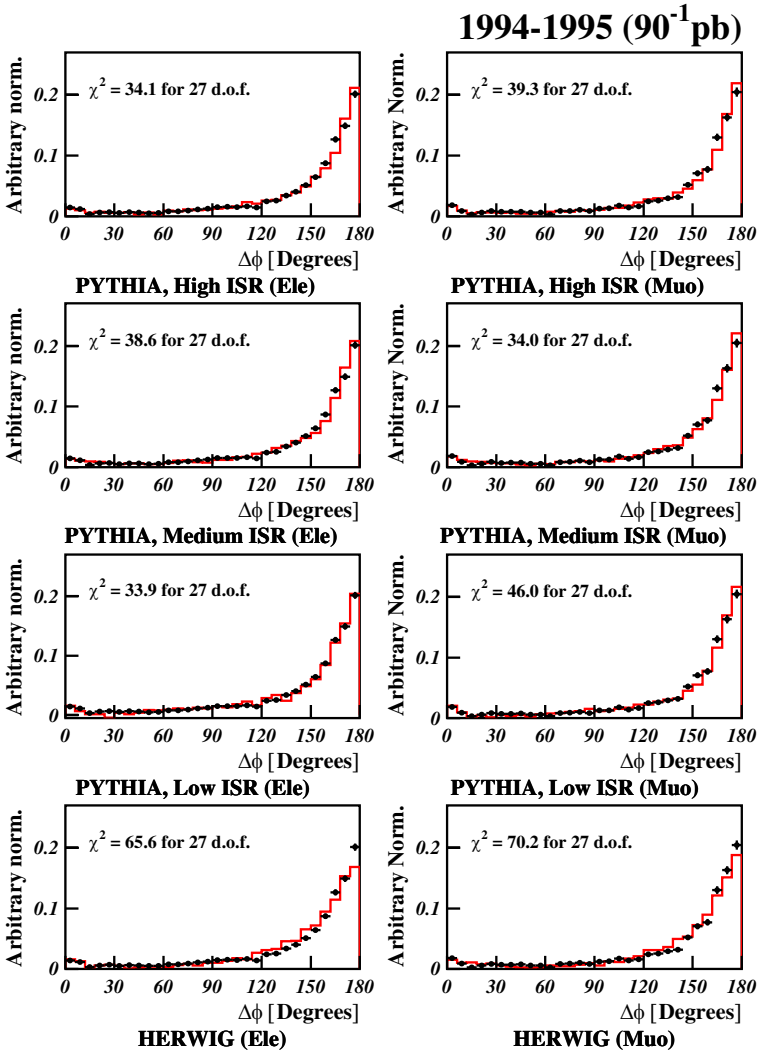

FIG. 7 (color online). Comparisons between the shape of the tag $\Delta\phi$ distribution for data (points) and the tag $\Delta\phi$ distribution predicted by the various Monte Carlo samples (line). In these comparisons, the normalizations of each production mechanism were allowed to vary independently and were chosen to give the best fit between the Monte Carlo and the data. Again, the fit χ^2 takes into account Monte Carlo statistics in addition to errors on the data.

high $\Delta\phi$, the rest of the $\Delta\phi$ distribution matches reasonably well between Monte Carlo and data using the normalizations predicted by the Monte Carlo generators for the different production mechanisms. The χ^2 values between the $\Delta\phi$ curves from Monte Carlo and data are listed in Table II. On the basis of these χ^2 values, it appears that PYTHIA with the medium ISR value provides the best match to data when using the Monte Carlos default normalization for the three production mechanisms. Figure 6 shows the break down of the contributions from the individual production mechanisms to the overall $\Delta\phi$ shape for this PYTHIA sample.

However, since, in the parton shower approximation, the contributions from flavor creation, flavor excitation, and gluon splitting may be generated separately, each component can have a separate, arbitrary normalization and the three components can be fit for the combination of normalizations that gives the best match to the shape of the $\Delta\phi$ spectrum from data. These fits are shown in Fig. 7. When the normalizations of the individual components are allowed to float with respect to one another, one can obtain rather good agreement in shape between data and both the low ISR and high ISR PYTHIA samples. The fit of the low ISR PYTHIA Monte Carlo to the data increases the broader contribution from flavor excitation to compensate for the narrowness of the back-to-back peak from flavor creation. For the high ISR PYTHIA samples, the peak at high $\Delta\phi$ is made narrower to match the data by all but eliminating the contribution from flavor excitation. A comparison of the relative fractions of each production mechanism in the two PYTHIA fits is shown in Fig. 8. The fit of the HERWIG sample to the data also tries to compensate for the excessive broadness of the HERWIG flavor creation peak at high $\Delta\phi$, but even after completely eliminating the flavor exci-

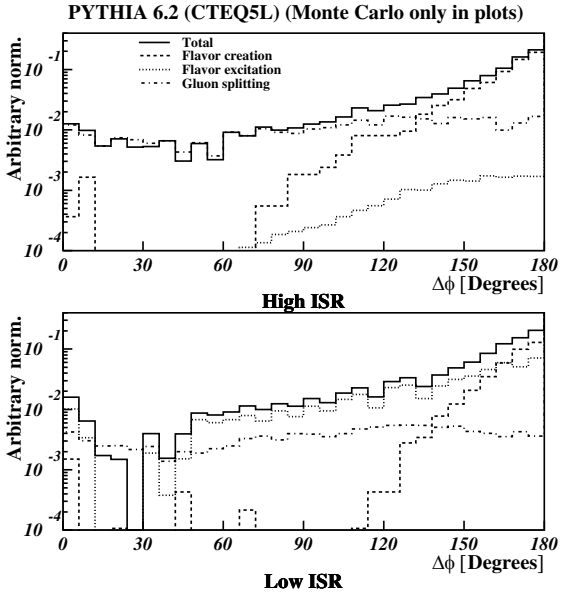

FIG. 8. A comparison of the contributions from flavor creation (dashes), flavor excitation (dots), and gluon splitting (dash-dots) to the total $\Delta\phi$ shapes (solid) for PYTHIA with high ISR (top) and low ISR (bottom). Electron Monte Carlo is shown in the plots. The muon plots can be found in Ref. [31]. The normalization of each component is set by the best fit of the three components to the spectrum from data. Note that mistag subtraction applied to the individual PYTHIA contributions can result in negative values for bins with few entries. Consequently, the total PYTHIA distribution can be less than one of the components in some bins.

tation contribution, the remaining contribution from flavor creation at high $\Delta\phi$ is too broad to model the data. Table II compares the fit quality and effective contribution from flavor creation, flavor excitation, and gluon splitting in the fits of the various Monte Carlo samples to the data. Both low ISR and high ISR PYTHIA samples can be made to fit the data with approximately the same fit quality, which is unexpected, especially since the low ISR sample accomplishes this fit with a high flavor excitation content while the high ISR sample fits with almost no flavor excitation contribution. In the end, there seems to be an ambiguity in PYTHIA that allows a trade-off between initial-state-radiation and the amount of flavor excitation.

In general, any of the Monte Carlo samples compared to the data shows reasonable qualitative agreement. The Monte Carlo sample that best matches the data is PYTHIA with medium or high ISR settings, when the individual normalizations of the flavor creation, flavor excitation, and gluon splitting are allowed to float separately to best fit the data. Although the fit using PYTHIA with low ISR is not so poor as to rule this model out completely, studies indicate that PYTHIA with high initial state radiation does a better job of matching both the underlying event and minimum bias data at CDF [34]. Therefore, we select the PYTHIA sample, with high ISR and the relative normalizations of flavor creation, flavor excitation, and gluon splitting fixed by our fit to the $\Delta\phi$ distribution of the data, as the best Monte Carlo model of the data. Comparisons indicate that the differences between PYTHIA with medium or high ISR settings are minor. Figure 9 shows a comparison of other correlations between the data and PYTHIA with high ISR. Although these plots show good agreement between the data and PYTHIA for the overall shapes of the distributions, the shapes of the individual contributions from flavor creation, flavor excitation, and gluon splitting are not distinct enough to allow a separation of the components as was done for the $\Delta\phi$ distribution.

It is interesting to note that before allowing the normalizations of each production mechanism to float in the fits, the agreement between HERWIG and the data is no worse than the agreement between the low ISR PYTHIA sample and the data. However, because the disparity between the data and the low ISR PYTHIA sample comes from the narrowness in the flavor creation peak at high $\Delta\phi$, when the normalizations are allowed to float, the fit can alleviate the disagreement by increasing the peak width through a higher contribution from flavor excitation. In contrast, for HERWIG, once the contribution from flavor excitation has been reduced to zero, the fit has no way to make the width of the back-to-back flavor creation peak smaller, short of the unphysical situation of setting the flavor excitation normalization negative. If there were some other parameter for HERWIG, like PYTHIAS initial state-radiation parameter, PARP(67), that could be used to tune the width of the back-to-back flavor creation peak, it may be possible to achieve good agreement between HERWIG and the data as well.

The results presented here can be compared to another analysis of lepton tags in heavy-flavor events presented in Ref. [16]. That analysis compares HERWIG to double-tagged events using higher E_T jet samples. In addition to using a sample of double-tagged events at higher momentum, Ref. [16] also differs from this analysis in that it uses calorimeter based jets as opposed to the tracking jets utilized here, and the quantity measured is the azimuthal opening angle between the tagged jets rather than the angle between the tags themselves. In agreement with this analysis, that one clearly shows the importance of the higher-order contributions in heavy-flavor production, and also shows an agreement with HERWIG that is better than the agreement seen in this analysis. Perhaps this suggests that the disagreement shown here between HERWIG and the data is related to HERWIGS ability to model low p_T b production or b fragmentation.

G. Corrections and systematics

The correlations examined so far in the data involve pairs of BVTX tags, rather than pairs of B hadrons. There are detector effects, such as the tagging efficiency for pairs of B hadrons as a function of $\Delta\phi$, that distort the shape of the measured tag pair correlations from the true B hadron distribution. In addition, residual contributions

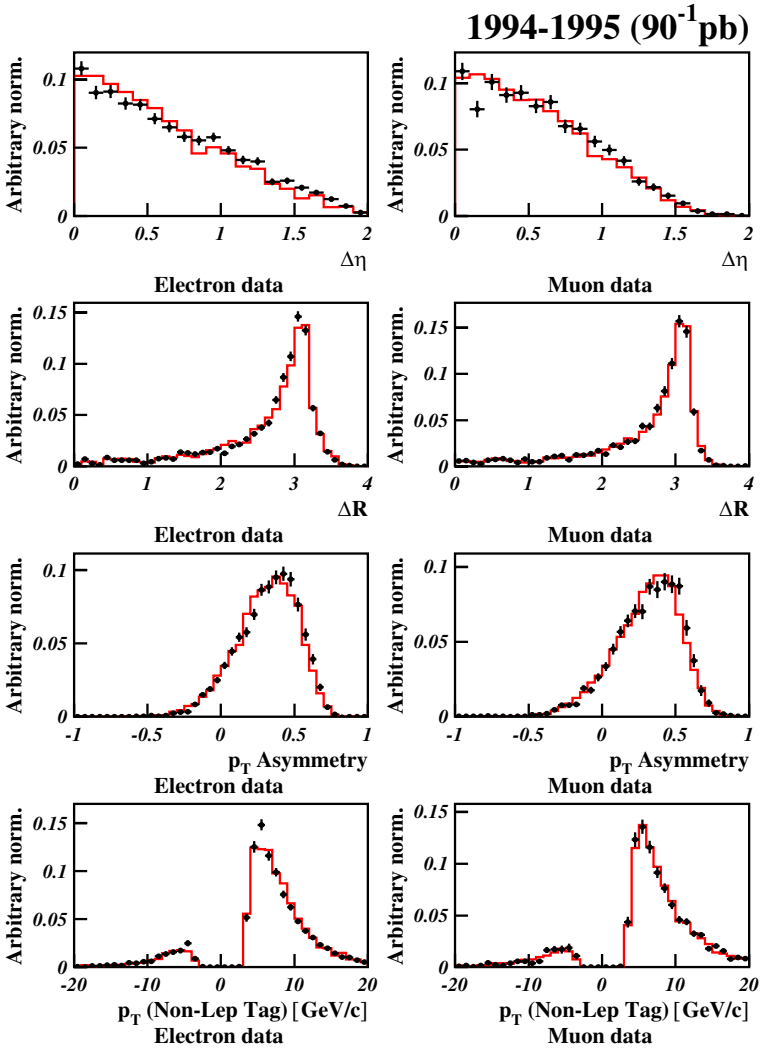

FIG. 9 (color online). A comparison of PYTHIA with high ISR to the data for several different correlation quantities. The normalizations for the three production mechanisms in PYTHIA have been determined by the fit of the PYTHIA $\Delta\phi$ distributions to data. The p_T asymmetry is given by $A_{p_T} = (p_T(\text{lep}) - p_T(\text{non} - \text{lep})/(p_T(\text{lep}) + p_T(\text{non} - \text{lep})$. In the $p_T(\text{Non} - \text{LepTag})$ plot, the sign of the p_T is determined by the opening angle between the lepton-tag and the non-lepton tag: negative for tag pairs with $\Delta\phi < 90°$, positive otherwise. The data are shown as points with statistical error bars only. The solid line is PYTHIA with high ISR.

from background can affect the shape of the tag pair distribution. For the comparison between MC and data, the detector effects are accounted for by using a detector and trigger simulation to adjust the MC to match the conditions in the data, while the backgrounds are assumed to be negligible. However, since the MC models examined in Sec. III F match the data reasonably well, MC events can be used to determine the relationship between the measured tag pair distribution and the actual B hadron distribution. In the sections below, two kinds of corrections to the tag pair $\Delta\phi$ distribution are considered: a correction for the relative tagging efficiency, which is a detector effect, and a correction for the contributions from mistags, prompt charm, and sequential decays that remain in the data after the steps taken in Sec. III D to remove backgrounds. In addition, the MC is used to estimate the systematic uncertainties associated with correcting for the relative tagging efficiency and removing background events. These corrections and systematic errors are evaluated using several different MC samples to account for uncertainties involved in the MC model itself.

The BVTX tagging algorithm is not equally effective for all topologies of $b\bar{b}$ production. In particular, it becomes more difficult for the BVTX algorithm to reconstruct both displaced secondary vertices as the opening angle between the two B hadrons decreases. This effect becomes especially severe when the two B hadrons are both contained within the cone of a single jet for track clustering purposes. Furthermore, correlations between opening angle and $p_T(B)$ for the various production mechanisms can lead to differences in the relative efficiency for reconstructing tag pairs at different opening angles. These effects distort the

shape of the $\Delta\phi$ distribution measured for tags from the true $B\overline{B}$ $\Delta\phi$ distribution.

We correct for these relative efficiency effects using the MC that best matches the data, as determined in Sec. III F. Because we are only examining the shape of the $\Delta\phi$ distribution, our goal in making this correction is only to account for differences in the relative efficiency of the tagging algorithm, as a function of $\Delta\phi$. We do not attempt to correct for effects that impact all parts of the $\Delta\phi$ spectrum equally. For example, an overall shift in the muon trigger efficiency would not affect this correction. To determine the correction for each bin we take the ratio of the number of tag pairs reconstructed in the MC to the number of pairs that could have been reconstructed if the tagging algorithm had perfect efficiency. The number of tag pairs that would have been reconstructed assuming perfect efficiency is determined by looking at the generator-level B hadron $\Delta\phi$ distribution. For electron MC, to simulate the electron trigger, we require one B hadron in the event to have a $p_T > 14.0$ GeV/c and $|\eta| < 1.0$. For the muon MC, we demand one B hadron with $p_T > 14.0$ GeV/c and $|\eta| < 0.6$. For both cases, we require a second B hadron with $p_T > 7.5$ GeV/c and $|\eta| < 1.0$. The cuts placed on the generator-level MC were determined by examining the p_T and η distributions for B hadrons from MC events in which two BVTX tags were reconstructed. The p_T and η values were chosen by determining the cuts for which 90% of the B hadrons in the double-tagged MC events would pass. We take the $\Delta\phi$ distribution resulting from the event selection above and convolute it with a Gaussian resolution function with a width of 0.1086 radians, characteristic of the $\Delta\phi$ resolution of the BVTX tagging algorithm as measured in MC.

In order to minimize the effect of statistical fluctuations in the tagging efficiency determined from MC, we fit the tagging efficiency to an empirical function of the following form:

$$\epsilon(\Delta\phi) = P_1 \exp\left\{-\frac{1}{2}\left(\frac{\Delta\phi^2}{P_2^2}\right)\right\} + P_3 \exp\left\{-\frac{1}{2}\left[\frac{(\Delta\phi - \pi)^2}{P_4^2}\right]\right\} + P_5 \mathrm{freq}\left(\frac{\Delta\phi - P_6}{P_7}\right) + P_8(\Delta\phi) + P_9, \quad (2)$$

where freq is the error function. The relative efficiency curve resulting from this fit is shown in Fig. 10. The sharp step around $\phi = 60°$, which is modeled by the error function term, comes from the transition from the case of finding secondary vertex tags in two separate jets to finding secondary vertex tags in the same jet. Since we are only interested in the effect of the efficiency on the shape of the $\Delta\phi$ distribution, and not on its absolute normalization, we have rescaled the curve in Fig. 10 so that the relative efficiency in the last $\Delta\phi$ bin is defined to be unity. Thus this curve shows the effect of the BVTX tagging efficiency for a given bin relative to the last $\Delta\phi$ bin.

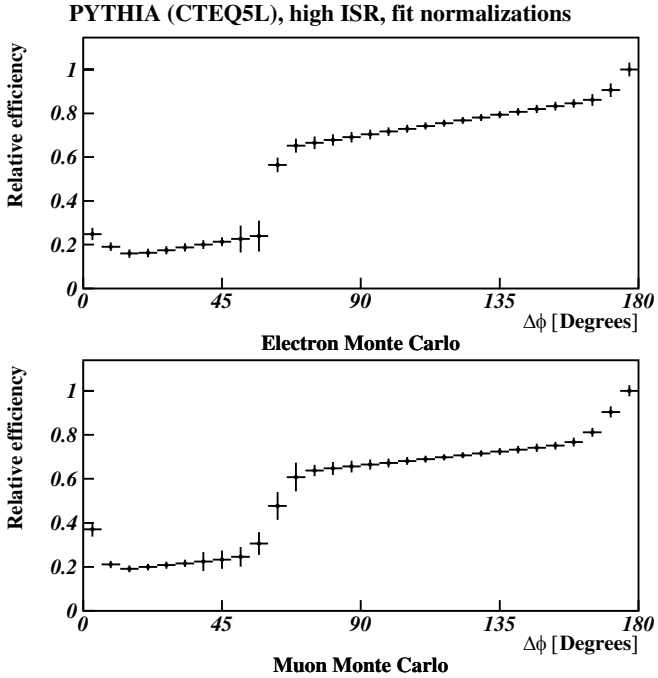

FIG. 10. The bin-by-bin values for the relative efficiency returned by the fit. The curves have been normalized so that the last $\Delta\phi$ bin has a value of one by definition. The error bars on these curves indicate the statistical error on the bin values returned from the fit. The statistical errors for the fit are correlated from bin to bin.

There are two main contributions to the systematic uncertainty associated with the relative tagging efficiency correction. First, the statistical errors on the fit value for the relative efficiency correction factor should be propagated into systematic uncertainties on the corrected $\Delta\phi$ distribution. There is an additional systematic uncertainty that comes from the model used to calculate the relative efficiency correction. The PYTHIA MC sample, with high ISR and with the normalization of the different production mechanisms taken from the fit to the $\Delta\phi$ distribution in the data, is used as our baseline for the relative efficiency correction. However, other models, like the lower ISR PYTHIA sample or HERWIG also match the data to varying degrees and so could also have been used. To account for this ambiguity, we compare the relative efficiency corrections from other MC models to our baseline model. In the worst case, the difference for the bin-by-bin relative efficiency correction factor is approximately equal in magnitude to the statistical error from the fit. Therefore, to account for modeling uncertainties in the relative efficiency correction, we increase the systematic error associated with the correction by a factor of $\sqrt{2}$.

The mistag subtraction scheme used for this analysis relies on the assumption that 100% of legitimate tags and

50% of mistags have positive L_{xy}. The true fraction may be somewhat different. For example, if most of the events contain at least one B hadron, then the L_{xy} distribution of mistags may be biased towards positive values by the presence of actual displaced tracks in the events. Furthermore, the bias in L_{xy} may depend on the topology of the event. To investigate any possible bias in the L_{xy} distribution of mistags, we examined MC events containing mistags identified by matching tracking information to MC truth information. From MC sample to MC sample, the fraction of legitimate secondary vertex tags that have positive L_{xy} varies from 0.97 to 1.0. For mistags, the positive L_{xy} fraction varies from 0.45 to 0.55. To estimate the possible effect of using the wrong fractions when performing mistag subtraction, we redo the mistag subtraction in the data using different assumptions about the positive L_{xy} fraction for good tags and mistags. The mistag subtraction formula [Eq. (1)], generalized for an arbitrary fraction p of good tags with positive L_{xy} and an arbitrary fraction q of mistags with positive L_{xy}, is given by

$$N_{GG} = \frac{(q-1)^2}{(q-p)^2}N_{++} + \frac{q(q-1)}{(q-p)^2}N_{+-}, + \frac{q^2}{(q-p)^2}N_{--} + \frac{q(q-1)}{(q-p)^2}N_{-+}. \quad (3)$$

Changing the positive L_{xy} fractions from mistag subtraction affects both the normalization and the shape of the $\Delta\phi$ distribution. However, we are only concerned about the shape for this analysis. Therefore, before we compare the shape of the $\Delta\phi$ distribution using the standard mistag subtraction scheme to the shape obtained using alternative values for the positive L_{xy} fractions, we normalize the distributions to unit area. To estimate the systematic error from mistag subtraction, we take the $\Delta\phi$ distributions calculated varying the p and q values in Eq. (3) within their allowed ranges and fit them to the functional form for the $\Delta\phi$ distribution, given below, in order to minimize the effect of statistical fluctuations:

$$f(\Delta\phi) = P_1 \exp\left[-\frac{1}{2}\left(\frac{\Delta\phi}{P_2}\right)^2\right] + P_3 \exp[P_4(\Delta\phi - \pi)^2 + P_5(\Delta\phi)] + P_6\Delta\phi + P_7. \quad (4)$$

We then calculate the maximum deviation between the result from the default mistag subtraction scheme and the results obtained from varying the positive L_{xy} fractions. This maximum deviation is assigned as the systematic error on the $\Delta\phi$ shape from mistag subtraction.

The bin-by-bin contribution to the double-tag $\Delta\phi$ distribution from prompt charm is estimated primarily using MC. The overall amounts of prompt charm and $b\bar{b} + c\bar{c}$ double tags are estimated by comparing the relative rate of obtaining a double-tagged $b\bar{b}$ MC event to the rates for double-tagging $c\bar{c}$ and $b\bar{b} + c\bar{c}$ MC events. This approach estimates that 2.9% (6.0%) of the tag pairs in this sample come from $c\bar{c}$ production for electron (muon) data, and 1.8% of the tag pairs in both the electron and the muon samples comes from $b\bar{b} + c\bar{c}$ production. The $\Delta\phi$ shape for the $c\bar{c}$ and $b\bar{b} + c\bar{c}$ contributions is estimated by applying the measured relative tagging efficiency as a function of $\Delta\phi$ to the generator level $c\bar{c}$ and $b\bar{b} + c\bar{c}$ $\Delta\phi$ distributions. The resulting estimated contamination from prompt charm to the double-tag $\Delta\phi$ distribution is shown in Fig. 11. The systematic error on this correction is estimated by performing several checks on the data. One check involves comparing the $\Delta\phi$ spectrum for double-tagged events in which the invariant mass of the tracks for each tag is greater than 2 GeV/c^2 to the spectrum when both tags have an invariant mass less than 2 GeV/c^2. The former sample is enhanced in $b\bar{b}$ content relative to prompt charm, while the latter sample has a greater contribution from prompt charm. Both subsamples have far fewer statistics than the main sample. The $\Delta\phi$ shapes of these two subsamples agree within the statistics of the samples, suggesting a negligible contribution from prompt charm. An alternative estimation of the prompt charm contribution can be obtained by fitting the tag mass distribution to template shapes derived from $b\bar{b}$ events (including tags of secondary charmed mesons) and $c\bar{c}$ events. The results of these fits suggest a prompt charm contamination roughly

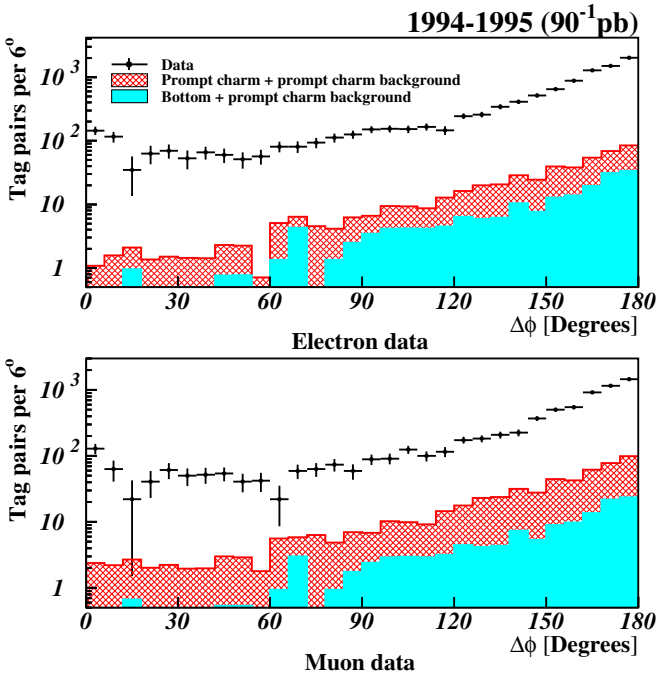

FIG. 11 (color online). The estimated shape of the background from direct $c\bar{c}$ production (the hatched area) and multiple-heavy flavor ($b\bar{b} + b\bar{b}$ and $b\bar{b} + c\bar{c}$) production (the solid area). The points with error bars show the data.

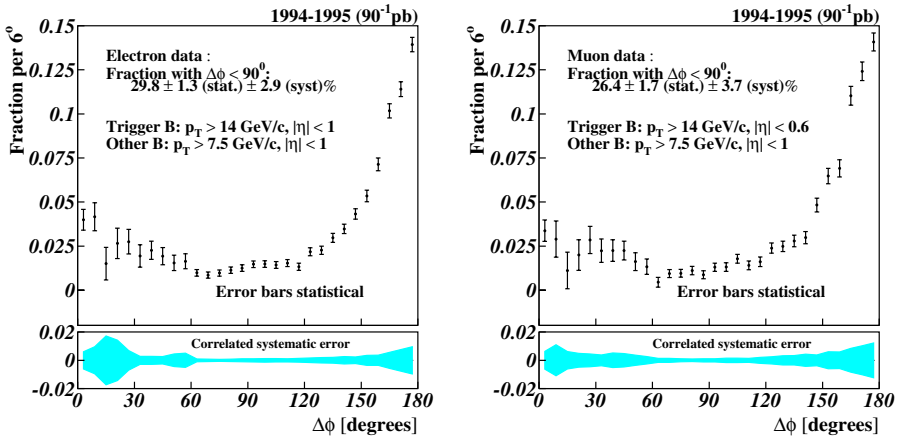

FIG. 12 (color online). The final, corrected $\Delta\phi$ distribution for the double-tagged electron (left) and muon (right) data. The corrections made to the data include mistag subtraction, sequential removal, prompt charm subtraction, and the relative tagging efficiency correction. The error bars display the statistical error on the points. The filled region at the bottom indicates the systematic errors. The systematic errors are correlated from bin to bin. Mistag subtraction provides the dominant contribution to the systematic errors.

a factor of 2 larger than the MC estimates, although still a relatively small contribution at 7.1% for the electron data and 13.3% for the muon data. As a result of the differences between these two alternate estimates of the prompt charm contribution and the MC method used to set the normalization of our prompt charm correction, we set the systematic error on the prompt charm correction equal to the size of the correction in each bin.

The MC is also used to determine the residual contribution from sequential double tags. Based on examining MC events in which two tags are identified to come from the same B decay, we determine that after mistag subtraction, 25.9% of the tags removed by the 6 GeV/c^2 mass cut were from sequential tag pairs. Furthermore, using Monte Carlo it was also determined that for every 100 mistags removed by the 6 GeV/c^2 mass cut, roughly 2.41 events remained in this sample, yielding an efficiency for this cut of 97.6%. In the data, after mistag subtraction, the 6 GeV/c^2 invariant mass cut removes 471 tags from the electron sample and 598 tags from the muon sample. Using the numbers derived from the MC above, this means that of the tags removed by the invariant mass cut, 122.1 electron and 155.0 muon tags come from sequential double-tag pairs, and an estimated 2.9 electron sequential tag pairs and 3.7 muon sequential tag pairs remain in the data after this cut. The $\Delta\phi$ distribution of the sequential tag pairs is also determined using MC to be well described by the positive half of a Gaussian distribution with a mean of zero and a width of 0.122 radians. To correct for the sequential double-tag contribution in the data, we take the estimated number of sequential double tags, with a half-Gaussian distribution as described above, and subtract them from the $\Delta\phi$ bins in the data. The systematic error on this correction is set equal to the size of the correction.

H. Final distribution and comments

Figure 12 shows the final, corrected tag $\Delta\phi$ distribution, including systematic errors. To obtain this distribution, the contributions from residual sequentials and prompt charm

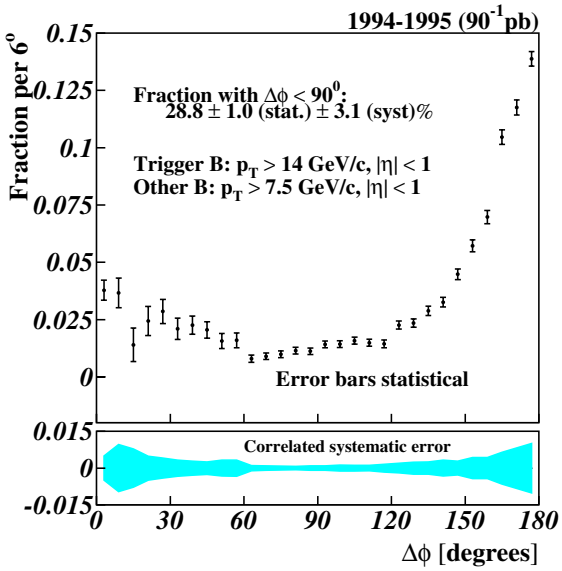

FIG. 13 (color online). The combined, corrected electron and muon $\Delta\phi$ distribution from the double-tagged analysis. In making this plot, we ignored the difference in η acceptance between the electron and muon triggers. The corrections made to the data include mistag subtraction, sequential removal, prompt charm subtraction, and the relative tagging efficiency correction. The error bars display the statistical error on the points. The filled region at the bottom indicates the systematic errors. The systematic errors are correlated from bin to bin. Mistag subtraction provides the dominant contribution to the systematic errors.

TABLE III. The corrected fraction of combined electron and muon data in each bin as well as a break down of the components of the systematic errors on each bin. The total systematic error is the sum in quadrature of the individual components.

				Systematic error components			
Bin	Fraction	Statistical error	Systematic error	Sequential	Prompt charm	Mistag subtraction	Relative efficiency
$0° - 6°$	0.03901	0.00462	0.00593	0.00060	0.00051	0.00421	0.00411
$6° - 12°$	0.03765	0.00684	0.01042	0.00044	0.00082	0.00982	0.00336
$12° - 18°$	0.01347	0.00774	0.00833	0.00013	0.00125	0.00810	0.00149
$18° - 24°$	0.02498	0.00674	0.00544	0	0.00084	0.00472	0.00257
$24° - 30°$	0.02942	0.00561	0.00472	0	0.00087	0.00370	0.00279
$30° - 36°$	0.02152	0.00493	0.00372	0	0.00074	0.00309	0.00194
$36° - 42°$	0.02323	0.00420	0.00336	0	0.00069	0.00256	0.00206
$42° - 48°$	0.02077	0.00379	0.00298	0	0.00101	0.00211	0.00185
$48° - 54°$	0.01568	0.00349	0.00472	0	0.00093	0.00171	0.00380
$54° - 60°$	0.01651	0.00344	0.00461	0	0.00043	0.00137	0.00438
$60° - 66°$	0.00751	0.00167	0.00114	0	0.00087	0.00054	0.00049
$66° - 72°$	0.00869	0.00151	0.00102	0	0.00084	0.00037	0.00044
$72° - 78°$	0.00973	0.00153	0.00090	0	0.00073	0.00030	0.00045
$78° - 84°$	0.01156	0.00156	0.00079	0	0.00059	0.00025	0.00047
$84° - 90°$	0.01100	0.00155	0.00097	0	0.00085	0.00022	0.00040
$90° - 96°$	0.01423	0.00157	0.00130	0	0.00084	0.00023	0.00046
$96° - 102°$	0.01395	0.00160	0.00128	0	0.00121	0.00026	0.00040
$102° - 108°$	0.01559	0.00162	0.00120	0	0.00117	0.00033	0.00040
$108° - 114°$	0.01474	0.00163	0.00172	0	0.00106	0.00044	0.00034
$114° - 120°$	0.01370	0.00177	0.00212	0	0.00159	0.00057	0.00029
$120° - 126°$	0.02203	0.00187	0.00259	0	0.00195	0.00071	0.00045
$126° - 132°$	0.02244	0.00193	0.00268	0	0.00242	0.00082	0.00045
$132° - 138°$	0.02813	0.00213	0.00344	0	0.00246	0.00088	0.00059
$138° - 144°$	0.03128	0.00223	0.00303	0	0.00328	0.00080	0.00069
$144° - 150°$	0.04471	0.00249	0.00465	0	0.00279	0.00051	0.00106
$150° - 156°$	0.05622	0.00275	0.00466	0	0.00444	0.00018	0.00137
$156° - 162°$	0.06983	0.00306	0.00716	0	0.00419	0.00113	0.00169
$162° - 168°$	0.10516	0.00341	0.00914	0	0.00583	0.00267	0.00319
$168° - 174°$	0.11783	0.00346	0.01105	0	0.00688	0.00457	0.00391
$174° - 180°$	0.13944	0.00336	0.00419	0	0.00779	0.00662	0.00419

are removed from the mistag-subtracted distributions. Then the relative efficiency corrections derived in Sec. III G are applied. Systematic errors from the various corrections are combined in quadrature to give the total systematic error. Mistag subtraction gives the largest contribution to the systematic error. The final corrected tag $\Delta\phi$ distribution provides a measurement of the $B - \overline{B}$ $\Delta\phi$ distribution where the B providing the trigger electron (muon) has $p_T > 14.0$ GeV/c and $|\eta| < 1.0(0.6)$, and the other B has $p_T > 7.5$ GeV/c and $|\eta| < 1.0$, with a $\Delta\phi$ resolution of 6.22°. This distribution can be compared to generator-level $\Delta\phi$ distributions from Monte Carlo that have been convoluted with a Gaussian resolution function to account for our $\Delta\phi$ resolution. Finally, ignoring the

TABLE IV. The number of events in the towards and away regions before and after applying corrections to the data.

	Electrons		Muons	
	Towards	Away	Towards	Away
Mistag-subtracted data	1210	8887	832	6260
Charm contamination	42.1	442.6	52.9	500.3
Sequential contamination	2.9	0.0	3.7	0.0
Relative efficiency Correction factor	0.326	1.0	0.376	1.0
Corrected data	3573.6	8444.4	2062.2	5759.7

TABLE V. The break down of the systematic errors by contribution. The total systematic error is the quadrature sum of the individual components.

	Electrons	Muons
Towards fraction	29.8%	26.4%
Statistical error	±1.3%	±1.7%
Mistag subtraction systematic error	±2.0%	±2.8%
Sequential removal systematic error	±0.05%	±0.09%
Charm subtraction systematic error	±1.3%	±2.2%
Relative efficiency correction systematic error	±1.6%	±1.1%
Total systematic error	±2.9%	±3.7%

small difference in η acceptance between the electron and muon samples, these two distributions can be combined to give the overall B hadron $\Delta\phi$ distribution, shown in Fig. 13. Table III specifies the corrected fraction in each $\Delta\phi$ bin as well as the break down of the systematic errors for each bin.

From the corrected data, we can also calculate the fraction of tag pairs in the ''towards'' region, defined by $\Delta\phi < 90°$. This fraction is of interest because $\Delta\phi$ production in the towards region is dominated by the higher-order production diagrams. The towards fraction provides a single figure of merit to indicate the relative sizes of the contributions from flavor excitation and gluon splitting. To account for correlated systematic errors, we calculate the towards fraction for our data by essentially repeating the analysis with two $\Delta\phi$ bins instead of 30, and then taking the ratio of the towards bin over the total. For the electron data, we obtain a towards fraction of $29.8 \pm 1.3(\text{stat.}) \pm 2.9(\text{syst.})\%$. For muon data, we obtain a towards fraction of $26.4 \pm 1.7(\text{stat.}) \pm 3.7(\text{syst.})\%$. The electron and muon samples are combined to give a towards fraction of $28.8 \pm 1.0(\text{stat.}) \pm 3.1(\text{syst.})\%$. Table IV shows the uncorrected number of tag pairs in the towards and ''away'' bins in the data and gives the corrections applied to obtain the final number. Table V breaks down the contributions to the systematic uncertainty on the towards fraction.

IV. J/ψ–LEPTON b QUARK CORRELATION MEASUREMENT

A. Overview

This measurement is optimized to measure the region in phase space least understood in experimental measurements and theoretical predictions: small $\Delta\phi$ where both bottom quarks point in the same azimuthal direction. As stated previously, earlier bottom quark angular production measurements had little sensitivity to this region. A study of opposite side flavor tags using soft leptons for the CDF $\sin 2\beta$ measurement [36] showed a significant number of tags at small opening angles between fully reconstructed bottom decays and the soft leptons. Figure 14 shows the

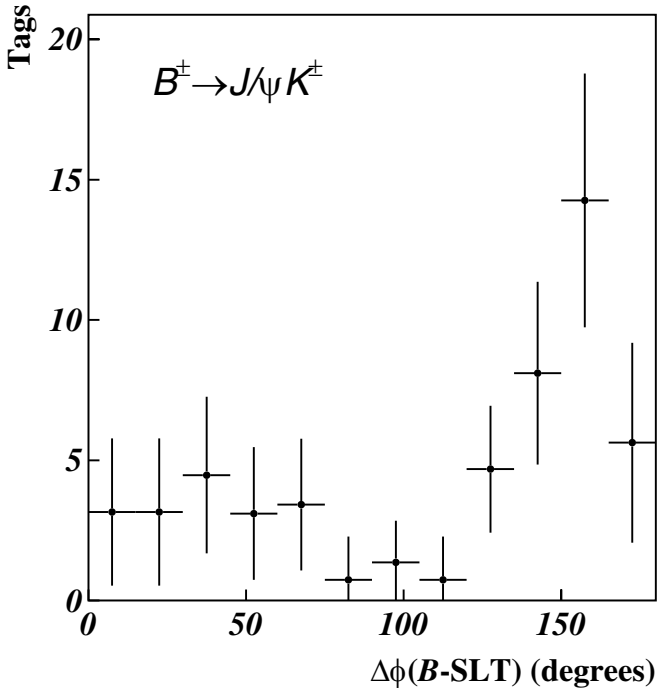

FIG. 14. Sideband-subtracted $\Delta\phi$ distribution between fully reconstructed $B^+ \to J/\psi K^+$ and soft leptons.

sideband-subtracted $\Delta\phi$ distribution between $B^+ \to J/\psi K^+$ candidates and the soft leptons. About 30% of the soft leptons are in the same azimuthal hemisphere, a fraction much larger than expected from parton shower flavor creation Monte Carlo ($\approx 5\%$ for PYTHIA flavor creation).

This analysis uses the bottom pair decay signature of $b \to J/\psi X$, $\overline{b} \to \ell^+ X$. The impact parameter of the additional lepton and the pseudo-$c\tau$ of the J/ψ are fit simultaneously in order to determine the $b\overline{b}$ fraction of the two $\Delta\phi$ regions. Angular requirements that were necessary in previous di-lepton measurements because of double sequential semileptonic decay backgrounds ($b \to c\ell^- X; c \to \ell^+ X'$) are avoided by the chosen signal. B_c is the only particle that decays directly into J/ψ and an addition lepton. The only other source of candidates where the additional lepton and J/ψ candidates originate from the same displaced decay are hadrons that fake leptons or decay-in-flight of kaons and pions. The number of events from B_c or from ''fake'' leptons can be estimated well by using techniques from Ref. [37]. Thus, no angle requirement between the two candidate bottom decay products are necessary, yielding uniform efficiency over the entire $\Delta\phi^{b\overline{b}}$ range. Because of the limited size of the data sample, only f_{toward}, the fraction of $b\overline{b}$ pairs in the same azimuthal hemisphere, can be measured.

The selection criteria used in this analysis have similar bottom momenta and rapidity acceptances to CDF's Run II displaced track (SVT) [38] and J/ψ triggers, and the addi-

tional leptons have momenta very similar to the opposite side taggers planned for Run II (opposite kaon, opposite lepton and jet charge flavor taggers). Therefore, this measurement aids in the development and understanding of flavor taggers for such Run II measurements as the B_s mass difference.

B. Sample selection

The signal searched for in this analysis is $b \to J/\psi X$, $\overline{b} \to \ell^+ X'$ where ℓ can be an electron or muon. In this section, the Run Ib J/ψ data set is described. The offline selection criteria for both the J/ψ and the additional lepton are also described.

1. $J/\psi \to \mu^+\mu^-$ *selection*

This analysis uses the CDF Run Ib J/ψ data set obtained between January 1994 and July 1995. The CDF J/ψ triggers and offline J/ψ selection criteria utilized are the same as the B_c discovery analysis at CDF [37]. In order to understand the trigger efficiencies, we confirm that the J/ψ candidate's muons are the two muon candidates which triggered the event.

After confirming the trigger, the position of the extrapolated track at the muon chamber is compared to the position of the muon stub using a χ^2 matching test, taking into account the effects of multiple scattering and energy loss in material. The positions are required to match within 3 standard deviations in the r$-\phi$ projection and within 3.5 standard deviations in the r$-$z projection.

Next, we require a high quality track for both muon candidates. The pseudolifetime ($c\tau$) of the J/ψ candidates is used in this analysis to determine the bottom purity. Therefore, SVX$'$ information is required in order to improve the precision of the $c\tau$ measurement.

In order to reject J/ψ candidates with muons originating from different primary interactions, the z position difference between the two tracks is required to be less than 5 cm at the beamline. A vertex constrained fit of the two muon candidates is performed [39]. The χ^2 probability of the vertex fit is required to be better than 1%. The vertex constrained mass of the J/ψ candidate is required to be $2.9\ \text{GeV} < \text{M}_{J/\psi} < 3.3\ \text{GeV}$.

A total of 177 650 events pass the above selection cuts. Figure 15 shows the J/ψ mass distribution for these events. In order to estimate the number of $J/\psi \to \mu^+\mu-$ candidates, the mass has been fit with two Gaussians (used to model the J/ψ signal) and a linear background term. A linear background has been assumed in many previous CDF J/ψ analyses [37,40,41]. The background under the mass peak is caused by irreducible decay-in-flight and punch-though backgrounds, Drell-Yan muons and double sequential semileptonic decays where $b \to c\mu^- X$, $c \to s\mu^+ X'$. From the fit, 137780 ± 440 J/ψ candidates are in the sample. For this measurement, the J/ψ mass signal region is defined to be within ± 50 MeV of the Particle

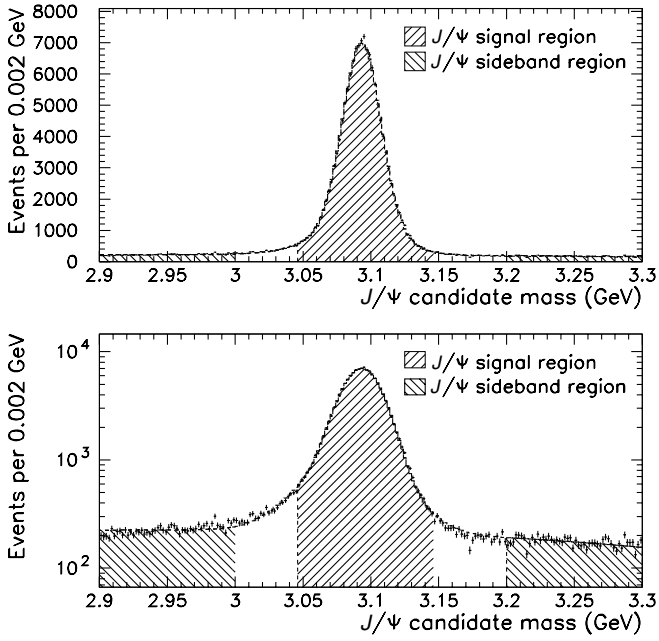

FIG. 15. Di-muon invariant mass distribution from events passing selection criteria. Top: linear scale. Bottom: logarithmic scale.

Data Group [42] world average value (3096.87 MeV). The sideband regions are chosen to be $2.900\ \text{GeV} \leq \text{M}_{J/\psi} \leq 3.000\ \text{GeV}$ and $3.200\ \text{GeV} \leq \text{M}_{J/\psi} \leq 3.300\ \text{GeV}$. The sideband regions contain 20 180 events. The events in these regions are used later in the analysis to describe the $c\tau$ shape of J/ψ background in the mass signal region.

The ratio between the number of background J/ψ events in the mass signal region to the background in the mass sideband region (R_{side}) was determined to be $R_{\text{side}} = 0.501 \pm 0.000043$ by the mass fit. To estimate the systematic uncertainty of this ratio, a 2nd order polynomial is used to describe the background term. The resulting fit value is $R_{\text{side}} = 0.545 \pm 0.008$. The difference between the two fits is taken to be the systematic uncertainty yielding $R_{\text{side}} = 0.501 \pm 0.044$.

2. CMUP μ *selection requirements*

The additional (non-J/ψ) muon is required to have muon stubs in both the CMU and CMP (a CMUP muon). Requiring both CMU and CMP muon stubs maximizes the amount of material traversed by the candidate, reducing the background due to hadronic punch-though of the calorimeter. The χ^2 matching requirements are the same as for the J/ψ muons. The muon candidates are required to have a $p_T > 3$ GeV; muons with lower p_T will typically range out prior to the CMP due to energy loss in the calorimeter and the CMP steel.

As the impact parameter is used to estimate the bottom purity of the muons, the same track quality is required as for the J/ψ muons. Additionally, the muon candidate's track projection is required to be in the fiducial volume

of the CMU and CMP. The z positions of the J/ψ candidate and the CMUP muon are required to be within 5 cm of each other at the beamline.

In total, 247 CMUP candidate muons are found in the J/ψ sample, out of which 51(142) CMUP candidates are in events where the J/ψ candidate is in the mass sideband (signal) region.

Of the 142 CMUP candidates with J/ψ candidates in the J/ψ mass signal region, 64 events have the CMUP muon and the J/ψ candidate in the same hemisphere in the azimuthal angle (which will be known as toward); the other 78 events have the CMUP muon and J/ψ candidate in the opposite hemisphere in the azimuthal angle (which is denoted away).

3. Electron selection criteria

A method of the finding soft (relatively low momenta) electrons was developed for bottom flavor tagging in CDF's B_d mixing and $\sin(2\beta)$ measurements [30,36]. These electrons have a relatively high purity, a transverse momenta greater than 2 GeV, and an understood efficiency. The rate of hadrons faking an electron was studied extensively in Ref. [37], making it possible to estimate the background due to hadrons faking electrons. The selection criteria is identical to Ref. [37] in order to use the same fake rate estimates.

The selection criteria requires a high quality track which is consistent with an electron in the various detector systems. Information on the energy (charge) deposited, the cluster location, and track matching χ^2 variables are all used in order to reduce the rate of a hadron in the detectors' fiducial volume faking an electron to $(6.4 \pm 0.6) \times 10^{-4}$.

One source of electron background is photon conversions, where a photon interacts with detector material and converts into a e^+e^- pair. In addition, 1.2% of all neutral pions decay into γe^+e^- directly (Dalitz decay). To reduce this background, conversions are searched for and vetoed by looking for a conversion partner track.

The conversion requirements are the same as Ref. [30]. Unfortunately, the conversion removal is not totally efficient. Therefore, some of the soft electron candidates are residual conversion electrons, where either the conversion pair track is not found due to tracking inefficiencies at low p_T or the conversion electron selection is not fully efficient. The rate of residual conversions is studied more in Sec. IV C 4.

In total, 514 candidate electrons are found after conversion removal; 92 events have the J/ψ candidate in the mass sidebands and 312 events have the J/ψ candidate in the mass signal region. In the J/ψ mass signal region, 107(205) of the events are in the toward(away) regions in $\Delta\phi$. In the J/ψ mass signal region, 6(9) events were vetoed as conversions in the toward (away) $\Delta\phi$ bin. In the J/ψ mass sideband region, 5(4) events were vetoed as conversions.

C. Signal and background description

The signal and backgrounds for both the $J/\psi + \mu$ and $J/\psi + e$ samples are very similar. The basic technique to determine the amounts of the various signal and background components is with a simultaneous fit of the pseudo-$c\tau$ (defined in Sec. IV C 1) of the J/ψ and the signed impact parameter of the non-J/ψ lepton. The impact parameter is signed to distinguish between residual electron conversions and electrons from bottom decay, as described in Sec. IV C 5. The impact parameter is signed positive if the primary vertex lies outside the r$-\phi$ projection of the particle's helix fit.

As the J/ψ and additional lepton originate from separate bottom hadron decays, the impact parameter of the additional lepton and the $c\tau$ of the J/ψ are not strongly correlated for the signal. The backgrounds in this analysis have two categories: one in which the impact parameter and $c\tau$ are uncorrelated, and other where the impact parameter and $c\tau$ are strongly correlated. The impact parameter and the $c\tau$ become strongly correlated when both the J/ψ and the additional lepton candidate originate from the same displaced vertex.

In uncorrelated sources, the impact parameter and $c\tau$ shapes describing the background are determined independently. J/ψ candidates are assumed to originate from three sources: direct J/ψ production (including feed-down from χ_{c1}, χ_{c2}, and $\psi(2s)$) where the J/ψ decays at the primary vertex, J/ψ from bottom decay (including the feed-down from higher $c\bar{c}$ resonances), and the non-J/ψ background described by the events in the J/ψ mass sidebands. Lepton candidates are assumed to originate from the following sources: directly produced fake or real leptons from the primary vertex, leptons from bottom decay (including $b \rightarrow cX \rightarrow \ell X'$), lepton candidates with the fake J/ψ candidate, and residual conversion electrons.

In addition, two correlated sources of backgrounds exist. The first source is $B_c \rightarrow J/\psi \ell^+ X$, which is a small but irreducible background. The impact parameter of the additional lepton and the $c\tau$ of the J/ψ is described by Monte Carlo techniques and the overall size of the background is also estimated (see Sec. IV C 7).

The other correlated source of background occurs when a bottom hadron decays into a J/ψ and a hadron which is misidentified as a lepton. For electrons, this background is due to hadrons (mostly $\pi^\pm$ and $K^\pm$) showering early in the calorimeter and passing the electron identification selection. For muons, there are two sources of this background. The largest source of correlated background is due to decay-in-flight of charged pions and kaons, which result in a real muon. These real muons are denoted as ''fakes'' in this analysis. The other, smaller correlated fake muon background is caused by hadrons punching through the calorimeter and muon steel shielding. These background sources are more fully described in Sec. IV C 8.

The following sections provide a description of the techniques used to determine the impact parameter and $c\tau$ shapes of the various sources and to estimate of the number of residual conversions, $B_c \to J/\psi \ell X$ events, and $b \to J/\psi \ell_{\text{fake}} X$ events in the sample.

1. J/ψ $c\tau$ signal and background distributions

The direct J/ψ and bottom decay $J/\psi c\tau$ shapes are determined from a fit to the data, using a technique previously establish in Ref. [40,41]. The relatively long average lifetime of the bottom hadron (1.564 ± 0.014 ps [42]) allows one to distinguish between these two sources. First, the signed transverse decay length L_{xy} is determined in the r$-\phi$ plane.

$$L_{xy} \equiv \frac{(\vec{X}_{\text{SV}} - \vec{X}_{\text{PV}}) \cdot \vec{p}_T^{J/\psi}}{p_T^{J/\psi}}, \tag{5}$$

where $\vec{X}_{\text{SV}}$ and $\vec{X}_{\text{PV}}$ are the locations of the J/ψ vertex and the primary vertex in the transverse plane, and $\vec{p}_T^{J/\psi}$ is the vector transverse momentum of the J/ψ. Directly produced J/ψ have a symmetric L_{xy} distribution around zero and bottom J/ψ events will predominately have a positive sign. If the bottom decay was fully reconstructed, one could determine the proper decay length exactly ($c\tau_{\text{proper}}$) from the measured L_{xy} and p_T. Because the bottom hadron is not fully reconstructed, a "pseudoproper decay length" ($c\tau$) is constructed using the kinematics of the J/ψ only:

$$c\tau \equiv \frac{L_{xy} \cdot m_{J/\psi}}{p_T^{J/\psi} \cdot F_{\text{corr}}(p_T^{J/\psi})}. \tag{6}$$

$F_{\text{corr}}(p_T^{J/\psi})$, determined in Ref. [40], is the average correction factor for the partial reconstruction of the bottom hadron.

The events in the J/ψ mass sidebands are used to model the fake J/ψ background under the J/ψ mass signal peak. Two components are fit using an unbinned log-likelihood technique: events from the primary vertex (direct) and events with lifetime from heavy flavor (predominantly from $b \to c\mu^- X \to \mu^+ X'$). The direct events are described by a symmetric resolution function chosen to be a Gaussian plus two symmetric exponentials. The events with lifetime are fit with a positive only exponential.

Once the $c\tau$ shape of the mass sideband events is found, the $c\tau$ shapes of directly produced and bottom decay J/ψ can be determined. The shape of the directly produced J/ψ ($F_{\text{direct}}^{c\tau}(x)$) is parameterized by a Gaussian with two symmetric exponential tails; this shape is the assumed resolution function of the $c\tau$ measurement. The shape of J/ψ events from bottom decay ($F_b^{c\tau}(x)$) is therefore described as a positive exponential convoluted with the $c\tau$ resolution function. The background shape (g_{back}) is fixed to the value obtained in the fit of the sideband region and the background fraction (f_{back}) is fixed to the value predicted by the J/ψ candidate mass fit. Figure 16 shows the fit result of the signal region. The fit average bottom proper decay length of 442 ± 5 μm is consistent with previous measurements at the Tevatron [40,41]. The fit yields a bottom fraction of $16.6\% \pm 0.2\%$ or equivalently 22150 ± 270 J/ψ from bottom decay.

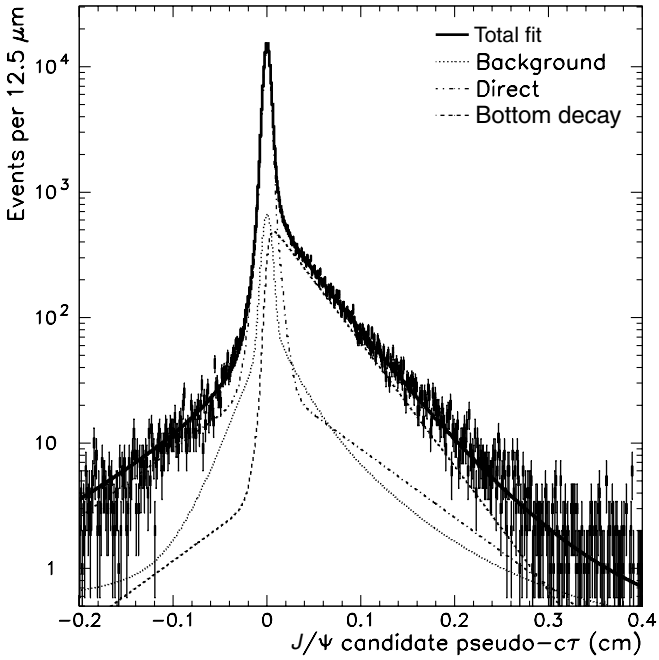

FIG. 16. Fit of J/ψ signal region.

2. Lepton impact parameter signal distributions

The impact parameter shape of bottom decay leptons is determined by Monte Carlo simulation, using the prescription from Ref. [20] for parton shower Monte Carlo programs. In this prescription, separate samples of flavor creation, flavor excitation, and gluon splitting events are generated and then combined with the relative rates predicted by the Monte Carlo programs. In Appendix B, the generation of the simulated events is described.

Figure 17 shows the unsigned impact parameter for direct bottom electrons ($b \to ce^-$) and sequential charm electrons ($b \to c \to se^+$) in the flavor creation sample. The impact parameter distributions are very similar and cannot be fit for separately. The uncertainty on the relative rate of these two sources is one of the systematic uncertainties treated in Sec. IV D 1. The bottom decay impact parameter shape is fit with two symmetric exponentials and a Gaussian. The fit to the simulated bottom impact parameter distribution for electrons is shown in Fig. 18.

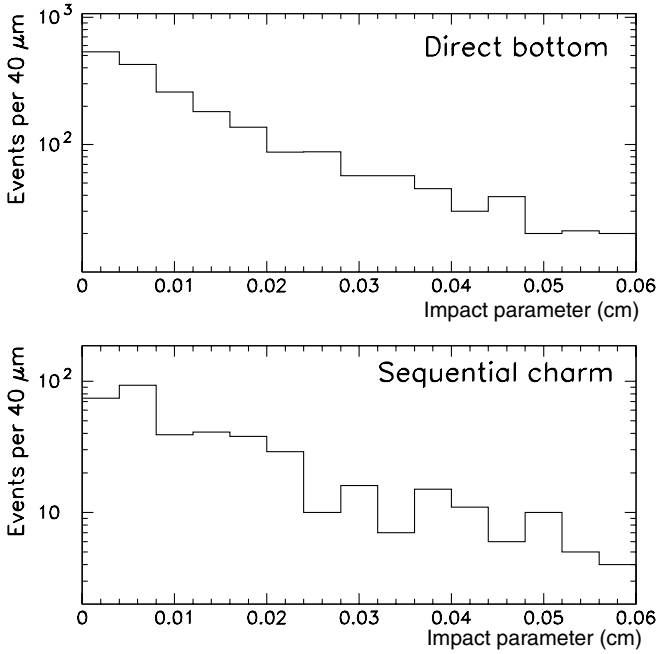

FIG. 17. The impact parameter of events with electrons passing requirement in the flavor creation Monte Carlo sample. Top: Direct bottom. Bottom: sequential charm.

3. Lepton impact parameter background distributions

In previous analyses [11,12], the impact parameter distributions for particles originating at the primary vertex were determined using jet data. Unfortunately, any data sample will have low level contamination of heavy flavor (charm or bottom) at the $\approx 0.1 - 1\%$ level, which is larger than the non-Gaussian effects in the impact parameter resolution. In this analysis, the impact parameter shapes of directly produced particles are determined with a Monte Carlo technique.

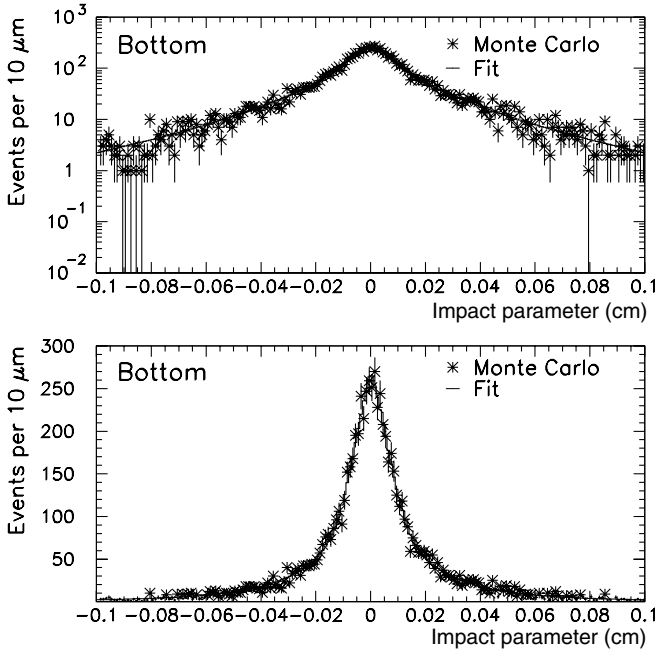

FIG. 18. The signed impact parameter distribution for the combined electron bottom Monte Carlo. Top: Logarithmic scale. Bottom: linear scale.

PYTHIA is used to generate the light quark and gluon subprocesses, which are then passed through a detector simulation. To be included in the electron sample, the candidates must be a quality track with a $p_T > 2$ GeV and extrapolate into the electron fiducial region. For the muon direct sample, the candidates must have a quality track with a $p_T > 3$ GeV and extrapolate into the CMUP muon fiducial region. The simulation events which pass the selection criteria are fit with two symmetric exponentials and a Gaussian. The fit shapes are then used as input to the likelihood fit for particles originating from the primary vertex (direct tracks).

4. Size of residual conversion background

One obvious source of electron background is residual conversions left in the sample due to the inefficiency of finding the conversion pair. In order to estimate the number of residual conversions, a technique similar to Ref. [43] is used. It is assumed that there are two independent causes for the lack of removal of a conversion electron: the track pair is lost due to tracking inefficiencies at low momenta or the selection requirements are not fully efficient. By measuring these two efficiencies and the rate of conversion removal with the chosen conversion selection requirements, one determines the residual electrons ($N_{\rm resid}$). The conversion electron that passes the electron identification criteria is denoted as the conversion candidate, and other electron that did not pass the electron identification criteria is denoted as the pair candidate.

The number of residual electrons is equal to

$$N_{\rm resid} = N_{\rm tag} \cdot \left(\frac{1}{\epsilon_{\rm cnv}({\rm cut})} \cdot \frac{1}{\epsilon_{\rm cnv}(p_T)} - 1 \right), \qquad (7)$$

where $N_{\rm tag}$ is the number of the conversions removed, $\epsilon_{\rm cnv}({\rm cut})$ is the conversion finding efficiency, and $\epsilon_{\rm cnv}(p_T)$ is the tracking efficiency of the conversion pair.

The efficiency $\epsilon_{\rm cnv}({\rm cut})$ is measured using different sets of conversion requirements, the tight (standard) and a loose set of cuts. Assuming that the loose cuts are fully efficient, the ratio of conversion pairs fit with tight and loose cuts yields $\epsilon_{\rm cnv}({\rm cut}) = 72.3 \pm 6.5\%$. In order to test this assumption, 2 additional wider sets of cuts are used which yield no extra conversion candidates.

The tracking efficiency of the pair candidates ($\epsilon_{\rm cnv}(p_T)$) is estimated with a Monte Carlo technique similar to Ref. [37]. The generation of the simulated conversions is detailed in Appendix B. $\epsilon_{\rm cnv}(p_T)$ is estimated by comparing the p_T distribution of the conversion partner in the simulation sample to the p_T distribution of the conversion

partners in the data sample. The simulation is normalized to the data in the p_T range where the tracking is fully efficient ($p_T > 0.5$ GeV). The ratio of the number of events seen in data versus the number of normalized conversion candidates in simulation gives $\epsilon_{\rm cnv}(p_T) = 69 \pm 5(\text{stat}) \pm 9(\text{syst})\%$. The systematic error includes the uncertainty in the conversion's p_T spectra.

The ratio of the number of residual to found conversions was found to be $R_{\rm conv} = 1.00 \pm 0.38$. The conversion veto removes 6(9) electron candidates in the toward (away) $\Delta\phi$ region. Thus, approximately 6.0 (9.0) residual conversion are in toward (away) $\Delta\phi$ region. About 5% of the electron candidates are residual conversions and have to be included in the $c\tau$–impact parameter fit. A total of 9 conversions (4 toward, 5 away) are found in events with the J/ψ candidate in the mass sideband region.

5. Impact parameter shape of residual conversion background

For conversion electrons from a primary photon, the primary vertex alway lies outside of the helix projection with perfect tracking. To distinguish between conversions and bottom decay electrons, the impact parameter is signed such that the impact parameter is positive if the primary vertex is outside the r–ϕ projection of the track's helix and is negative otherwise. Conversion electrons are positively signed, and Dalitz decay electrons and bottom decay electrons are equally negatively and positively signed.

The vast majority of the conversion candidates are signed positive as predicted (shown in Fig. 19). Unfortunately, there is a large positive tail which is not consistent with the number of silicon hits assigned to the track. Since at least 3 SVX′ hits are required, one would expect the conversion candidates either to originate from the first two silicon layers or the beam pipe, or to be a π^0 Dalitz decay from the primary vertex. These sources would produce conversions with a impact parameter less than 0.04 cm. Therefore, a fraction of the conversion candidates must have misassigned silicon hits and originate outside of the SVX′. The measured conversion radius for 25 of the 62 conversion candidates is greater than 6 cm, which is outside of the second silicon layer.

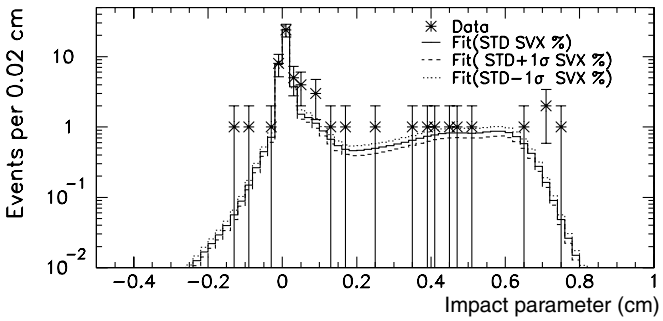

FIG. 19. The signed impact parameter distribution for conversions found in data. Solid line: Monte Carlo fit using the central value of $f_{\rm conv}^{\rm Good\ SVX}$. Dashed line: Monte Carlo fit increasing $f_{\rm conv}^{\rm Good\ SVX}$ by one sigma. Dotted line: Monte Carlo fit decreasing $f_{\rm conv}^{\rm Good\ SVX}$ by one sigma.

To construct an impact parameter shape for residual conversions, two Monte Carlo samples are generated: a sample in which the silicons hits are correctly assigned (low conversion radius, $R_{\rm conv} < 6$ cm) and a sample in which silicon hits are falsely added (mostly at high conversion radius, $R_{\rm conv} > 6$ cm). The relative fractions of the two components are determined by data. The fraction of conversions with $R_{\rm conv} < 6$ cm ($f_{\rm conv}^{\rm Good\ SVX}$) in data and simulation are matched, with the uncertainty in the fraction in data included as a systematic uncertainty in the shape.

Figure 19 shows the impact parameter of the candidates found in data and the combined conversion impact parameter shape normalized to data. The combined impact parameter shape describes the data adequately, including both the negative tail and large positive tail.

6. Additional lepton impact parameter distributions in events with fake J/ψ

Events in which the J/ψ candidate is in the mass sideband regions are used to describe the impact parameter shape of leptons in events with a fake J/ψ. The composition of events in the J/ψ mass sidebands are unknown; therefore, the shapes have to be fit in a similar manner as Sec. IV C 1.

For the muon sample, there is no knowledge of the contributions to the additional muon's impact parameter distribution. Therefore, the shape is parameterized with a Gaussian and symmetric exponential.

The impact parameter shape in the electron sample for events in the J/ψ mass sideband has an additional complication; the sample contains residual conversions. The number of found conversions in the signal region is used as a constraint on the number of residual conversion events in the signal region. Thus, the number of residual conversion events fit in the sideband region has to be known. In the sidebands, the residual conversion fraction is fit for $f_{\rm conv} = \frac{r_{\rm conv} \cdot n_{\rm convside}}{N_{\rm sideband}}$, where $r_{\rm conv}$ and $n_{\rm convside}$ are the fit ratio of residual to found conversions and fit number of "found" conversions, respectively. These quantities are constrained by the estimate of $R_{\rm conv}$ and the number of found conversions in the sidebands, $N_{\rm convside}$. Since $r_{\rm conv}$ is a component of the signal region fit, the signal and sideband regions have to be fit simultaneously.

7. $B_c \to J/\psi \ell^+ X$ background

B_c decay is the only known process that yields a lepton and a J/ψ from the same displaced vertex. CDF's measurement [37] of the $\frac{\sigma(B_c)\cdot BR(B_c \to J/\psi \ell \nu)}{\sigma(B^+)\cdot BR(B^+ \to J/\psi K)}$ and the B_c lifetime allow one to estimate both the number and impact parameter–$c\tau$ shape of this background. The estimated number of

$B_c \to J/\psi \ell X$ events in the samples is used as a constraint in the fit.

Taking into account correlated uncertainties, the number of $B_c \to J/\psi \ell X$ events estimated in the sample are $N^e_{B_c} = 10.0^{+3.5}_{-3.3}$ and $N^\mu_{B_c} = 7.2^{+2.6}_{-2.4}$. According to Monte Carlo, over 99% of the B_c passing the selection requirements have $\Delta\phi < \frac{\pi}{2}$ between the lepton and the J/ψ. Thus, this background is included in the fit only for the toward $\Delta\phi$ region.

Using a $B_c \to J/\psi \ell \nu$ Monte Carlo sample described in Appendix B, the impact parameter–$c\tau$ shape is determined. The shape takes into account the correlation in the impact parameter of the addition lepton and the $c\tau$ of the J/ψ. As the $c\tau$ of the J/ψ increases, the impact parameter of the additional lepton can be larger.

8. $b \to J/\psi \ell_{\mathrm{fake}}$ background

The other source of background where the impact parameter and $c\tau$ are correlated is bottom hadrons decaying to a real J/ψ with a hadron from the same decay faking an lepton. The sources and rates for faking leptons were studied extensively in Ref. [37] and are used in this analysis. In Monte Carlo, more than 99% of the events have the J/ψ and the fake lepton candidate in the same azimuthal hemisphere, and therefore this background is included only in the toward $\Delta\phi$ region.

The estimates of the amount and shapes of these backgrounds are made using a Monte Carlo sample of $b \to J/\psi X$ events, described in Appendix B. Hadrons can fake an electron by showering early in the calorimeter. ''Fake muons'' can be caused by decay-in-flight of charged pions and kaons, and by hadrons not being completely absorbed in the calorimeter and leaving hits in the muon chamber.

The rate that a hadron will fake an electron was studied in Ref. [37]. The number of $b \to J/\psi e_{\mathrm{fake}} X$ events is determined by using these fake rates and the Monte Carlo sample. To be included in the fake rate calculation, an event must pass the J/ψ selection and have a charged hadron in the electron identification fiducial volume with a track with SVX′ hits and a $p_T > 2$ GeV. Ideally, one would then apply the appropriate fake rate for the particle's p_T and isolation for each track passing the selection, yielding the total fake rate. The isolation is defined to be the scalar sum of the momenta of particles within a cone $\Delta R < 0.2$, divided by the momentum of the particle in consideration. Unfortunately, the Monte Carlo used does not include particles from the underlying event, fragmentation, gluon radiation or the other bottom hadron in the event. Thus, the isolation in the simulation does not represent the data. A large fraction ($\sim 70\%$) of the events have a small isolation ($I < 0.2$); this value of the isolation is used as a central value of the estimate. The fake rates using the other isolation value estimates are used to quantify the systematic uncertainty of the background estimate.

Normalizing the Monte Carlo to the estimated number of $b \to J/\psi X$ events in data yields a estimate of $2.85 \pm 0.03(\mathrm{stat.}) \pm 0.75(\mathrm{syst.})$ events, where the second error is the systematic uncertainty due to modeling (or lack thereof) of the track isolation.

The impact parameter-$c\tau$ shape of this background is determined by a fit to the Monte Carlo. The Monte Carlo sample used in the fit consists of events which can included in the electron fake rate calculation.

Decay-in-flight (DIF) of charged pions and kaons to muons is also a source of correlated background, as long as the track is reconstructible. The probability of a decay-in-flight to have a reconstructible track is greatly reduced by the SVX′ requirements, as shown in Ref. [37].

The number of correlated background events from decay-in-flight is determined in a manner similar to the fake electron estimate. The Monte Carlo is normalized in the same manner as the fake electron calculation. The J/ψ candidate is required to pass the selection criteria in Sec. IV B 1, and the decay-in-flight candidates are required to have a SVX′ track with $p_T > 3$ GeV and project into the CMU and CMP fiducial volumes. The probability of decaying-in-flight is determined for the given p_T and particle species. In Monte Carlo, $64.2 \pm 0.3\%$ of the particles passing the requirements are kaons. The decay-in-flight background is estimated to average 9.9 ± 2.1 events. The error includes the 12% Monte Carlo calculation systematic uncertainty and a 17% reconstruction efficiency systematic uncertainty quoted in Ref. [37].

In Ref. [37], the decay-in-flight estimate was done using data. In that analysis, the kaon fraction was measured to be $(44 \pm 4.4)\%$. The difference between the kaon fraction in Ref. [37] and the simulation could lead to a large systematic difference, because of the difference in the kaon and pion decay-in-flight probabilities. To estimate this uncertainty, the Monte Carlo events are reweighted in order to match the kaon fraction measured by [37]. With the reweighting, the estimated number of correlated decay-in-flight background is 8.7 ± 2.0. The difference between the two estimates is conservatively used as the systematic error, yielding a final decay-in-flight estimate of $N^{\mu,\mathrm{DIF}}_{B_{\mathrm{fake}}} = 9.9 \pm 2.4$.

The impact parameter–$c\tau$ shape of the decay-in-flight background is determined by the same Monte Carlo sample. In Ref. [12,37], it is shown that the impact parameter distribution of reconstructible decay-in-flight particles with SVX′ information have the same impact parameter distribution as the parent particle. Similar to the fake electron shape, the Monte Carlo events which could be used in the DIF rate calculation are fit in order to determine the DIF impact parameter-$c\tau$ shape.

Hadrons can also mimic muons by not being completely absorbed by the calorimeter and leaving hits in the muon chambers. The probability of a track punching-through the calorimeter was determined in Ref. [37]. The selection

criteria of the punch-through estimate is the same as the decay-in-flight estimate. The punch-though probability of the tracks passing the requirements is calculated from its particle type and momentum, yielding the final estimate. An average 1.76 ± 0.70 punch-though events are expected in the data, including a 40% systematic error used in [37].

As the punch-though rate is much larger for K^+ than K^- or $\pi^\pm$, the large difference in kaon fraction between [37] and Monte Carlo (shown in the decay-in-flight estimate) is a significant systematic shift in the punch-through estimate. To be conservative, we reweight the data with the kaon fraction measured in Ref. [37]; 1.23 ± 0.46 events are expected. The difference between the two predictions is used as the estimate of the systematic uncertainty in the prediction, yielding a final estimate of the average number of correlated backgrounds from punch-through of $N^{\mu,\mathrm{PT}}_{B_{\mathrm{fake}}} = 1.76 \pm 0.88$ events. In Ref. [44], the decay-in-flight and punch-through backgrounds are shown to have the same lifetime shape in the B_c lifetime fit. The decay-in-flight and punch-through backgrounds are assumed to have the same impact parameter-$c\tau$ shape.

D. Unbinned likelihood fit results

An unbinned log-likelihood fit is used to determine the estimated number of $b\bar{b}$ pairs and backgrounds in the two $\Delta\phi$ regions. Inputs to the fit are the J/ψ's $c\tau$, the additional lepton's impact parameter, and the shapes and the background estimates described in the previous sections. The shapes are used to determine the sample compositions, with the background estimates used as constants. In Appendix C, the complete details of the log-likelihood function is given.

The log-likelihood ($-2\ln\mathcal{L}$) is minimized for both data sets using MINUIT [45]. The fit parameter errors are defined by $\pm 1\sigma(\Delta\mathcal{L} = 1)$ contours of the likelihood function using the MINOS option. The results of the fit are shown in Table VI. In order to display the fit result, the log-likelihood function has been integrated in regions of impact parameter–$c\tau$ space. As examples of the fit quality, Figs. 20 and 21 show the fit results projected onto the impact parameter and $c\tau$ axis for the electron sample in the away $\Delta\phi$ region and muon sample in the toward $\Delta\phi$ regions, respectively.

TABLE VI. Fit results and constraints for the electron and muon samples. Appendix C describes all the fit parameters in detail. Variables n_{signal} and n_{conv} are not fit parameters but are functions of fit parameters.

Fit parameter	Electron	Electron constraint	Muon	Muon constraint
$n^t_{b\bar{b}}$	$29.6^{+11.7}_{-10.4}$		$23.0^{+7.6}_{-6.9}$	
n^t_{bd}	$1.5^{+8.5}_{-8.1}$		$1.6^{+4.6}_{-2.9}$	
$n^t_{b\mathrm{conv}}$	0.6 (fixed)		N/A	
n^t_{db}	0 (fixed)		0 (fixed)	
n^t_{dd}	$37.0^{+8.0}_{-7.3}$		$11.3^{+5.1}_{-4.5}$	
$n^t_{d\mathrm{conv}}$	$2.8^{+2.1}_{-1.7}$		N/A	
n^t_{side}	$45.4^{+6.9}_{-6.2}$	45	$32.9^{+5.7}_{-5.1}$	34
$n^t_{B_{\mathrm{fake}}}$	$2.8^{+0.7}_{-0.7}$	2.85 ± 0.75	$10.7^{+2.5}_{-2.5}$	11.7 ± 2.6
$n^t_{B_c}$	$10.0^{+3.2}_{-3.3}$	$10.0^{+3.5}_{-3.3}$	$5.1^{+2.5}_{-2.5}$	$7.2^{+2.6}_{-2.4}$
n^t_{signal}	107.1	107	68.2	64
n^t_{conv}	5.6	6	N/A	
$n^a_{b\bar{b}}$	$124.7^{+17.9}_{-16.7}$		$43.6^{+10.2}_{-9.0}$	
n^a_{bd}	$-1.4^{+12.5}_{-12.2}$		$8.1^{+8.0}_{-7.5}$	
$n^a_{b\mathrm{conv}}$	1.2 (fixed)		N/A	
n^a_{db}	0 (fixed)		0 (fixed)	
n^a_{dd}	$49.5^{+9.2}_{-8.5}$		$16.0^{+5.5}_{-5.2}$	
$n^a_{d\mathrm{conv}}$	$6.0^{+2.6}_{-2.2}$		N/A	
n^a_{side}	$47.6^{+7.1}_{-6.5}$	47	$18.2^{+4.5}_{-3.9}$	17
n^a_{signal}	204.9	205	76.8	78
n^a_{conv}	9.5	9	N/A	
r_{side}	$0.505^{+0.043}_{-0.043}$	0.501 ± 0.044	$0.501^{+0.043}_{-0.043}$	0.501 ± 0.044
r_{conv}	$0.99^{+0.31}_{-0.28}$	1.00 ± 0.37	N/A	

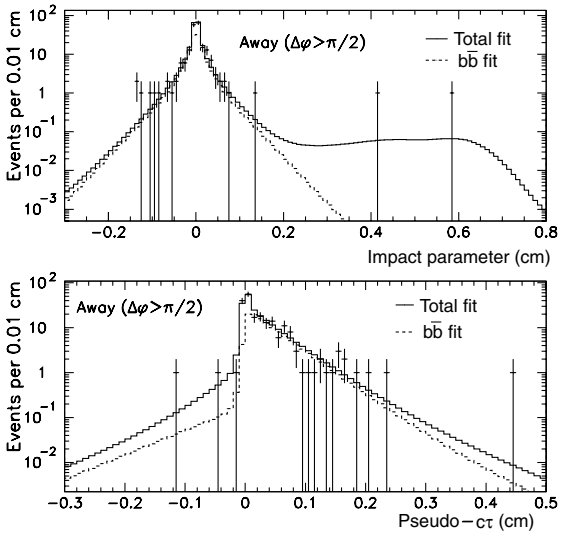

FIG. 20. Result of the $c\tau$−impact parameter fit for the electron sample in the away bin. Top: Projection onto impact parameter axis. Bottom: projection onto $c\tau$ axis.

The toward fraction measured in the two samples are

$$f^{\mu}_{\text{toward}} = 34.5^{+9.2}_{-8.2}\%, \tag{8}$$

$$f^{e}_{\text{toward}} = 19.2^{+6.5}_{-5.9}\%. \tag{9}$$

The measurement error includes both the statistical error as well as systematic uncertainties due to the constraints.

As a test of the fitting technique, a set of 1000 toy Monte Carlo "experiments" have been generated. The results of the study, given in Appendix D, show that the fit results are unbiased and have proper errors.

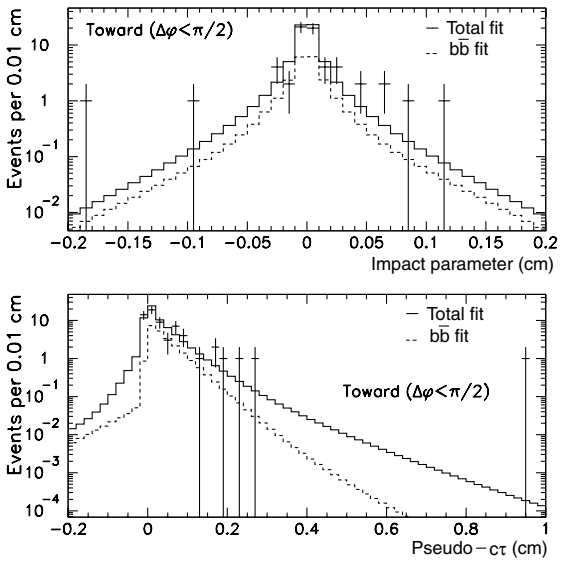

FIG. 21. Result of the $c\tau$−impact parameter fit for the muon sample in the toward bin. Top: Projection onto impact parameter axis. Bottom: projection onto $c\tau$ axis.

D. Fit systematics

By measuring the fraction of bottom quark pairs produced in the same hemisphere f_{toward}, the systematic uncertainties are minimized. The selection in both the $\Delta\phi < \pi/2$ and $\Delta\phi > \pi/2$ regions are the same; therefore, the uncertainties in the lepton selection efficiency, tracking efficiency, luminosity, etc. mostly cancel in the fraction measurement. In this section, the systematic uncertainty in the log-likelihood not already included in the fit is estimated. The estimated size of these systematic uncertainties are collected in Table VII.

The sequential charm fraction (f_{seq}) that is used in the bottom impact parameter shape ($F_b^{d_0}$) is derived from the simulation. The uncertainty in the sequential charm fraction leads to a systematic uncertainty in the determination of $F_b^{d_0}$, as sequential charm leptons have a larger impact parameter than direct bottom leptons. In Ref. [12], the relative systematic uncertainty in f_{seq} was studied. The relative uncertainty in f_{seq} was $\pm 19\%$, which is used in this analysis. The value of f_{seq} is varied by $\pm 1\sigma$, $F_b^{d_0}$ is refit, and then the new $F_b^{d_0}$ shapes are used to refit f_{toward}. The maximum differences of $\pm 0.1\%$ and $\pm 0.3\%$ are assigned as the systematic uncertainty for the electron and muon samples, respectively.

The bottom hadrons' lifetimes (B^+, B^0, B_s, and Λ_b) and their decay products' impact parameters are strongly correlated. In order to estimate the uncertainty caused by bottom lifetime uncertainty, two additional Monte Carlo samples were generated using BGENERATOR, a fast $b\bar{b}$ Monte Carlo that approximates the NLO prediction by Ref. [1]. All the bottom hadron lifetimes are shifted by $\pm 1\sigma$ from their PDG values [42]. The $F_b^{c\tau}$ shapes determined by these samples are then used in a refit of f_{toward}. The estimate of the systematic uncertainty due to the bottom lifetime is chosen to be the greatest differences from the standard fit. The uncertainties estimated in the electron and muon samples are $\pm 0.3\%$ and $\pm 2.2\%$, respectively.

TABLE VII. Summary of the estimated values of the systematic uncertainty for f_{toward}.

Source	Electron	Muon
Sequential rate	$\pm 0.1\%$	$\pm 0.3\%$
B Lifetime	$\pm 0.3\%$	$\pm 2.2\%$
Fragmentation fractions	$\pm 0.1\%$	$\pm 0.2\%$
$n_{b\text{conv}}/n_{d\text{conv}}$ ratio	$\pm 0.1\%$	
Residual conversion shape	$\pm 0.2\%$	
Direct impact parameter shape	$+0.3 - 0.4\%$	$+7.4 - 1.0\%$
f_{back} (for $F^{c\tau}_{\text{Direct}}$ and $F_b^{c\tau}$)	$\pm 0.02\%$	$\pm 0.01\%$
N_{db}	$\pm 0.1\%$	$\pm 1.9\%$
Total	$+0.5 - 0.6\%$	$+8.0 - 3.1\%$

B_s, B^+, and B^- have proper decay lengths of $\sim 470\mu$m, whereas Λ_B has a proper decay length of 387μm. Thus, the uncertainty in the fraction of bottom quarks fragmenting to Λ_b leads to the largest uncertainty of the $F_b^{c\tau}$ shape. Using BGENERATOR, samples are generated with the Λ_b fragmentation fraction varied by $\pm 1\sigma$ from the PDG values [46]. The new $F_b^{c\tau}$ shapes are used to refit f_{toward}, with the maximum difference from the standard fit used as the estimate of the systematic uncertainty, yielding a systematic uncertainty of $\pm 0.1\%$ and $\pm 0.2\%$ for the electron and muon samples.

Because of the limited number of residual conversions in the sample, the number of conversions pairing with J/ψ for bottom decay ($n_{b\text{conv}}$) and with direct J/ψ ($n_{d\text{conv}}$) cannot be fit independently. Thus, the ratio between $n_{b\text{conv}}$ and $n_{d\text{conv}}$ is fixed to the fit ratio between J/ψ from bottom decay and directly produced J/ψ. In order to estimate the systematic uncertainty due to this assumption, the fit of the data is redone with either $n_{b\text{conv}}$ or $n_{d\text{conv}}$ fixed to zero; the difference between fits are used as an estimate of the systematic uncertainty, yielding a systematic uncertainty of $\pm 0.1\%$ for the electron sample.

The residual conversion shape (F_{conv}) is determined using data and simulation. In data, the conversion radii of the found conversions indicate that a large fraction of the conversion candidates have at least 1 SVX′ hit misassigned to the track. The shape F_{conv} is the sum of two shapes: $F_{\text{conv}}^{\text{Good SVX}}$, which describes the shape of residual conversion where SVX′ hits are assumed to be correctly assigned, and $F_{\text{conv}}^{\text{Bad SVX}}$, which describes the shape of residual conversion where at least 1 SVX′ hit is assumed to be incorrectly assigned. $F_{\text{conv}}^{\text{Good SVX}}$ and $F_{\text{conv}}^{\text{Bad SVX}}$ are determined using Monte Carlo described in Sec. IV C 5. The value of $f_{\text{conv}}^{\text{Good SVX}}$ is changed by $\pm 1\sigma$ in F_{conv} in order to estimate the systematic uncertainty due to the F_{conv} shape used. The maximum difference of $\pm 0.2\%$ is assigned as a conservative estimate of the systematic uncertainty due to the residual conversion impact parameter shape used.

The direct impact parameter shapes ($F_{\text{direct}}^{d_0}$) are determined by a fit to Monte Carlo samples in Sec. IV C 3. The finite size of the Monte Carlo samples lead to an uncertainty in the fit parameters of the shapes. In order to estimate the uncertainty in f_{toward} due the $F_{\text{direct}}^{d_0}$ shape uncertainty, each parameter is fixed to a value $\pm 1\sigma$ from the best fit value and the $F_{\text{direct}}^{d_0}$ shape is refit. The new shape is then used in the impact parameter–$c\tau$ fit. The largest negative and positive differences from the standard fit is conservatively assigned as the systematic error, $+0.3 - 0.4\%$ for the electron sample and $+7.4 - 1.0\%$ for the muon sample.

The direct and bottom $c\tau$ shapes are determined by a fit to the data. In the fit, the fraction of fake J/ψ events (f^{back})

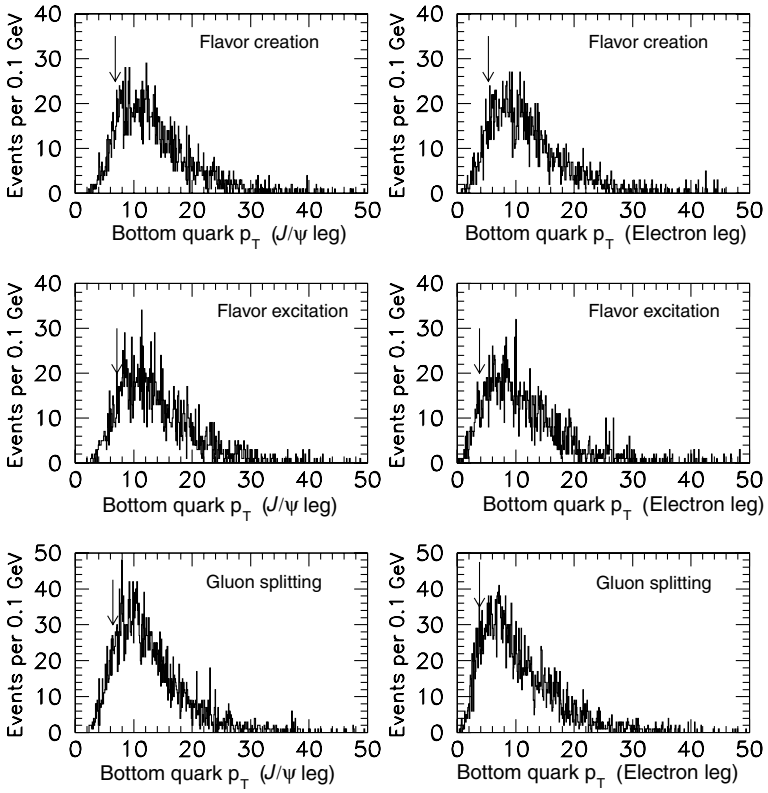

FIG. 22. p_T of the bottom quarks in events that pass selection in the electron PYTHIA samples. The arrows indicate the 90% acceptance value.

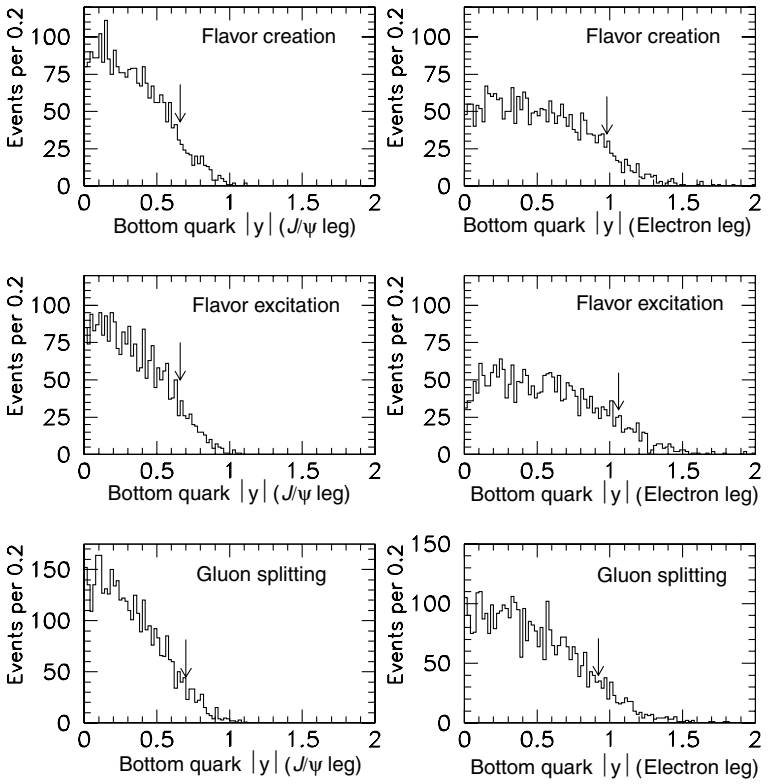

FIG. 23. $|y|$ of the bottom quarks in events that pass selection in the electron PYTHIA samples. The arrows indicate the 90% acceptance value.

is fixed at the predicted fraction. In order to estimate the effect on f_{toward}, the value of f^{back} is change by $\pm 1\sigma$ and the J/ψ mass signal region $c\tau$ fit is redone. The resulting $F^{c\tau}_{\text{direct}}$ and $F^{c\tau}_{b}$ shapes are used in a refit of f_{toward}. The greatest difference from the standard fit is chosen to be a conservative estimate of systematic uncertainty, yielding uncertainty estimates of $\pm 0.015\%$ and $\pm 0.01\%$ for the electron and muon channels.

In this analysis, the number of events with a directly produced J/ψ with a lepton from bottom decay (n_{db}) is assumed to be zero. In order to measure the effects of this assumption, a fit of f_{toward} is performed where n_{db} is a free parameter. The difference in this fit from the standard fit is assigned as the systematic uncertainty due to n_{db}. We assign a $\pm 1.9\%$ uncertainty to f^{μ}_{toward} and $\pm 0.1\%$ to f^{e}_{toward}.

The individual systematic uncertainties are added in quadrature in order to determine the combined systematic uncertainty. The systematic uncertainties for the electron and muon samples are $+0.5 - 0.6\%$ and $+8.0 - 3.1\%$, respectively.

2. Correction to b quark level

At this time, no fragmenting NLO QCD calculation of bottom production at the Tevatron exists. In order to compare to next-to-leading-order calculations, one must "correct" the experimental measurement to the bottom quark level ($f^{\text{corr}}_{\text{toward}}$), using similar technique as in Ref. [11,12,47].

The correction is

$$C_{B\to b} = \frac{f^{b\bar{b}^{90\%}}_{\text{toward},mc}}{f^{b\to J/\psi X;\bar{b}\to \ell Y}_{\text{toward},mc}} \tag{10}$$

TABLE VIII. 90% acceptance requirements on the bottom quarks decaying to a J/ψ or a lepton predicting by PYTHIA Monte Carlo and a detector simulation. Top: electron. Bottom: muon.

Sample	$p_T^{J/\psi}$	$y^{J/\psi}$	p_T^e	y^e
FC	6.8 GeV	0.66	5.3 GeV	0.98
FE	7.1 GeV	0.66	3.8 GeV	1.06
GS	6.4 GeV	0.70	3.8 GeV	0.92
Ave	6.8 GeV	0.67	4.3 GeV	0.99

Sample	$p_T^{J/\psi}$	$y^{J/\psi}$	p_T^μ	y^μ
FC	7.3 GeV	0.66	6.6 GeV	0.60
FE	7.0 GeV	0.66	5.8 GeV	0.66
GS	6.6 GeV	0.68	5.7 GeV	0.58
Ave	7.0 GeV	0.67	6.0 GeV	0.61

TABLE IX. Correction factor between the experimental measurement and the bottom quarks. The errors quoted are statistical only. Top: electron sample. Bottom: muon sample.

	FC	FE	GS	Combined
$f^{b\bar{b}^{90\%}}_{\text{toward},mc}$	$5.1 \pm 0.1\%$	$21.4 \pm 0.5\%$	$46.4 \pm 0.5\%$	$26.4 \pm 0.2\%$
$f^{b\to J/\psi X;\bar{b}\to \ell Y}_{\text{toward},mc}$	$5.8 \pm 0.4\%$	$23.4 \pm 0.8\%$	$47.8 \pm 0.8\%$	$27.3 \pm 0.5\%$
$C_{B\to b}$	0.879 ± 0.063	0.915 ± 0.038	0.971 ± 0.020	0.967 ± 0.019

	FC	FE	GS	Combined
$f^{b\bar{b}^{90\%}}_{\text{toward},mc}$	$3.5 \pm 0.2\%$	$19.5 \pm 0.8\%$	$47.2 \pm 0.8\%$	$25.5 \pm 0.4\%$
$f^{b\to J/\psi X;\bar{b}\to \ell Y}_{\text{toward},mc}$	$3.4 \pm 0.5\%$	$20.4 \pm 1.3\%$	$49.3 \pm 1.2\%$	$26.3 \pm 0.7\%$
$C_{B\to b}$	1.029 ± 0.164	0.956 ± 0.072	0.957 ± 0.028	0.968 ± 0.026

$f^{b\to J/\psi X;\bar{b}\to \ell Y}_{\text{toward},mc}$ is the f_{toward} prediction by PYTHIA, where the $\Delta\phi$ is calculated between the J/ψ and the additional lepton, which both meet the selection criteria. The quantity $f^{b\bar{b}^{90\%}}_{\text{toward},mc}$ is the fraction of bottom quarks produced by PYTHIA in the same hemisphere which pass the following criteria:

(a) $p_T^{b_1} > p_T^{J/\psi}$ and $|y^{b_1}| < y^{J/\psi}$

(b) $p_T^{b_2} > p_T^{\ell}$ and $|y^{b_2}| < y^{\ell}$

$b_{1,2}$ can be either the bottom quark or antiquark. No requirements are made on the decay products of the bottom quarks.

The distributions of the p_T and rapidity (y) of bottom quarks in Monte Carlo events that pass J/ψ and lepton selection (shown in Fig. 22 and 23) are used to determine the p_T and rapidity regions for the calculation of the correction factor. $p_T^{J/\psi}$ is defined to be the value of the bottom quark transverse momentum in which 90% of the b quarks which decay to a J/ψ have a higher momenta. $y^{J/\psi}$ is defined to be the value of bottom quark's $|y|$ in which 90% of the b quarks which decay to a J/ψ have a lower $|y|$. p_T^{ℓ} and y^{ℓ} are defined in a similar manner for the bottom quark that decayed into the additional lepton. Table VIII shows the value determined in both the electron and muon samples for three different production mechanisms. The rapidities of all three mechanisms are very similar and are determined by the detector geometry. The p_T values are different in the three mechanisms. Flavor creation produces two bottom quarks with similar momenta, while gluon splitting and flavor excitation produce quarks with dissimilar p_T. The values of $y^{J/\psi}$, y^{ℓ}, $p_T^{J/\psi}$, and p_T^{ℓ} used in the correction factor calculation is the average of the three production mechanisms. To estimate the size of the systematic uncertainty, $C_{B\to b}$ is calculated for each production mechanism separately. The systematic uncertainty is estimated as the largest difference between the individual and combined production mechanisms.

Table IX shows the $f^{b\bar{b}^{90\%}}_{\text{toward},mc}$, $f^{b\to J/\psi X;\bar{b}\to \ell Y}_{\text{toward},mc}$, and $C_{B\to b}$ for the complete sample and the three

TABLE X. Compilation of the corrected data results, the PYTHIA predictions, and the NLO predictions of f_{toward} Top: electron. Bottom: muon.

NLO MRST99 (Electron)	
$\langle k_T\rangle = 0.0$ GeV	$16.9\% \pm 0.2\%(\text{stat.})^{+3.6\%}_{-3.8\%}(\text{sys.})$
$\langle k_T\rangle = 1.0$ GeV	$19.4\% \pm 0.3\%(\text{stat.})^{+4.8\%}_{-4.1\%}(\text{sys.})$
$\langle k_T\rangle = 2.0$ GeV	$23.2\% \pm 0.4\%(\text{stat.})^{+5.8\%}_{-5.0\%}(\text{sys.})$
$\langle k_T\rangle = 3.0$ GeV	$31.9\% \pm 0.6\%(\text{stat.})^{+5.3\%}_{-5.7\%}(\text{sys.})$
$\langle k_T\rangle = 4.0$ GeV	$44.9\% \pm 0.7\%(\text{stat.})^{+5.4\%}_{-4.8\%}(\text{sys.})$
NLO CTEQ5M (electron)	
$\langle k_T\rangle = 0.0$ GeV	$16.5\% \pm 0.2\%(\text{stat.})^{+3.7\%}_{-3.3\%}(\text{sys.})$
$\langle k_T\rangle = 1.0$ GeV	$19.1\% \pm 0.3\%(\text{stat.})^{+4.9\%}_{-3.9\%}(\text{sys.})$
$\langle k_T\rangle = 2.0$ GeV	$23.1\% \pm 0.4\%(\text{stat.})^{+5.2\%}_{-5.0\%}(\text{sys.})$
$\langle k_T\rangle = 3.0$ GeV	$31.7\% \pm 0.6\%(\text{stat.})^{+6.0\%}_{-5.4\%}(\text{sys.})$
$\langle k_T\rangle = 4.0$ GeV	$45.1\% \pm 0.7\%(\text{stat.})^{+5.5\%}_{-4.6\%}(\text{sys.})$
PYTHIA (electron)	$26.4\% \pm 0.2\%(\text{stat.})$
Data (electron)	$18.6^{+6.3+0.5}_{-5.7-0.6} \pm 1.7\%$
NLO MRST99 (Muon)	
$\langle k_T\rangle = 0.0$ GeV	$16.7\% \pm 0.3\%(\text{stat.})^{+5.2\%}_{-3.6\%}(\text{sys.})$
$\langle k_T\rangle = 1.0$ GeV	$22.7\% \pm 0.6\%(\text{stat.})^{+6.1\%}_{-7.0\%}(\text{sys.})$
$\langle k_T\rangle = 2.0$ GeV	$24.7\% \pm 0.6\%(\text{stat.})^{+6.7\%}_{-6.5\%}(\text{sys.})$
$\langle k_T\rangle = 3.0$ GeV	$32.1\% \pm 0.6\%(\text{stat.})^{+6.4\%}_{-6.9\%}(\text{sys.})$
$\langle k_T\rangle = 4.0$ GeV	$44.9\% \pm 0.9\%(\text{stat.})^{+5.1\%}_{-6.6\%}(\text{sys.})$
NLO CTEQ5M (Muon)	
$\langle k_T\rangle = 0.0$ GeV	$16.8\% \pm 0.3\%(\text{stat.})^{+4.1\%}_{-3.8\%}(\text{sys.})$
$\langle k_T\rangle = 1.0$ GeV	$22.1\% \pm 0.5\%(\text{stat.})^{+5.6\%}_{-5.9\%}(\text{sys.})$
$\langle k_T\rangle = 2.0$ GeV	$23.8\% \pm 0.9\%(\text{stat.})^{+7.4\%}_{-5.0\%}(\text{sys.})$
$\langle k_T\rangle = 3.0$ GeV	$31.9\% \pm 0.9\%(\text{stat.})^{+5.7\%}_{-6.5\%}(\text{sys.})$
$\langle k_T\rangle = 4.0$ GeV	$44.4\% \pm 1.2\%(\text{stat.})^{+7.1\%}_{-5.0\%}(\text{sys.})$
PYTHIA(Muon)	$25.5\% \pm 0.4\%(\text{stat.})$
Data(Muon)	$33.4^{+8.9+7.7}_{-7.9-3.0} \pm 2.3\%$

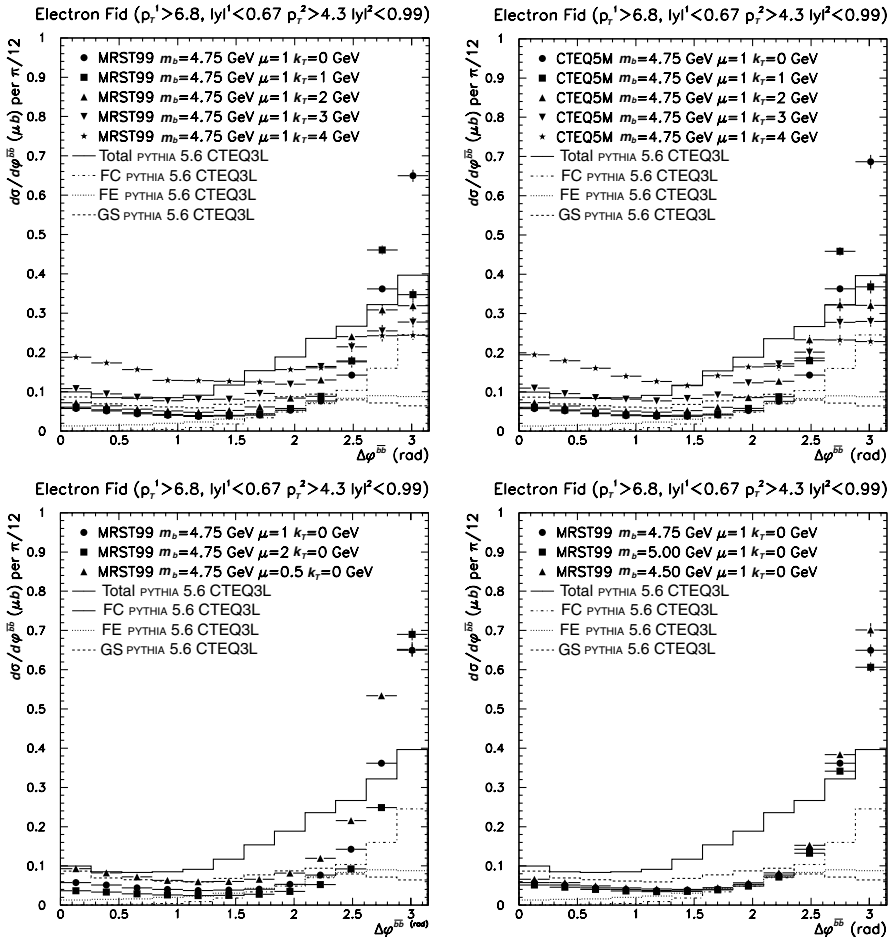

FIG. 24. LO prediction [1] of bottom spectra in the acceptance region. The PYTHIA prediction is shown as a reference. Top Left: MRST99 PDF varying the additional k_T smearing. Top right: CTEQ5M PDF varying the additional k_T smearing. Bottom left: MRST99 PDF varying the renormalization scale μ. Bottom right: MRST99 PDF varying the bottom quark mass m_B.

separate production mechanisms. The values estimated are

$$C^e_{B\to b} = 0.967 \pm 0.019(\text{stat.}) \pm 0.088(\text{syst.})\mathrm{C}^{\mu}_{\mathrm{B\to b}} = 0.968 \pm 0.026(\text{stat.}) \pm 0.061(\text{syst.}) \quad (11)$$

The measured toward fraction for the bottom quarks ($f^{\text{corr}}_{\text{toward}}$) extracted using the correction factor ($C_{B\to b}$) is

$$f^{\text{corr},e}_{\text{toward}} = 18.6^{+6.3+0.5}_{-5.7-0.6} \pm 1.7\%, \quad (12)$$

$$f^{\text{corr},\mu}_{\text{toward}} = 33.4^{+8.9+7.7}_{-7.9-3.0} \pm 2.3\%, \quad (13)$$

where the first error is the fit error, the second error is the additional shape systematic uncertainties on f_{toward}, and the third error is the uncertainty due to the correction to the bottom quark kinematics.

E. Data-theory comparisons

The measured toward fraction corrected to the quark level is compared to the NLO QCD predictions [1], using the same requirements as for the correction of the experimental measurements. The NLO prediction ($f^{\text{NLO}}_{\text{toward}}$) is made using $m_b = 4.75$ GeV, a renormalization/factorization scale $\mu = \sqrt{m_b^2 + (p_T^b + p_T^{\bar{b}})/2}$ and CTEQ5M [48] and MRST99 [49] parton distribution functions (PDFs). To estimate the systematic uncertainty in the NLO calculation, m_b is varied from 4.5–5.0 GeV, and μ is varied from 0.5–2.0. To study the effects of large initial-state parton transverse momenta (k_T), the NLO prediction is also made with $\langle k_T\rangle$ values of 0–4 GeV. The k_T effects are implemented in the same manner as in Ref. [50], where a $\langle k_T\rangle$ of 3–4 GeV per parton is predicted at the Tevatron. $f^{\text{NLO}}_{\text{toward}}$ is predicted using MRST99 and CTEQ5M PDFs, respectively, for the different input values of $\langle k_T\rangle$, m_b, and μ. Table X shows the summary of the predictions. The NLO predictions do not depend strongly on the PDF selected. The measured $f^{\text{corr}}_{\text{forward}}$ for both the electron and muon sample are consistent with the NLO prediction for values of $\langle k_T\rangle$ between zero and 3 GeV. The renormalization/factorization scale uncertainties in the NLO predictions and the statistical uncertainty in the measurement of

$f_{\text{toward}}^{\text{corr}}$ prohibit a more precise determination of $\langle k_T \rangle$ from this analysis.

Figure 24 illustrates the effects of varying the PDF, $\langle k_T \rangle$, m_b, and renormalization/factorization scale on the NLO predictions in the electron acceptance region. Varying m_b mass does not affect the predicted shape, but instead only affects the total cross section predicted. The two different PDFs studied yield very similar shape and total cross section predictions. Only scale and $\langle k_T \rangle$ variations yield appreciably different shape and total cross section predictions. Varying the renormalization/factorization scale changes the total cross section as expected; lowering the scale increases the total cross section. In addition, varying the scale changes the predicted rate at large $\Delta\phi^{b\bar{b}}$ (> 2.9 radians) relative to the rest of the distribution, while the shape of <2.9 radian region varies little. Varying the scale changes the relative rates of $p\bar{p} \rightarrow b\bar{b}$ to $p\bar{p} \rightarrow b\bar{b}g$ in the NLO prediction. Varying the $\langle k_T \rangle$ on the other hand, changes the predicted $\Delta\phi^{b\bar{b}}$ in a more continuous manner. With the increased number of $J/\psi + \ell$ expected in Run II, a differential azimuthal cross section measurement with 6–12 bins in $\Delta\phi$ should be able to separate scale uncertainty and k_T smearing effects.

V. CONCLUSIONS

We have presented two new measurements of the $\Delta\phi$ distribution for bottom antibottom pairs produced through QCD interactions at the Tevatron. These measurements are specifically targeted to measure the $\Delta\phi$ distribution down to arbitrarily small opening angles for $b\bar{b}$ pairs produced with low transverse momentum, where previous measurements have lacked sensitivity. The small $b\bar{b}$ opening angle region is of interest because in this region, the higher-order $b\bar{b}$ production mechanisms dominate over flavor creation. The data presented here are consistent with other measurements and cannot be described solely by flavor creation. Both measurements indicate that a significant fraction of the $b\bar{b}$ pairs (roughly 25%) are produced with $\Delta\phi < 90°$, in agreement with the conclusion from previous analyses [11,20] that flavor excitation and gluon splitting play a significant role in $b\bar{b}$ production at the Tevatron. The results of these measurements are consistent with the parton shower Monte Carlo models of PYTHIA and HERWIG as implemented in Ref. [20] and with NLO QCD predictions. Neither nonperturbitive [14] nor supersymmetric [15] production mechanisms are needed in order to describe the measured $\Delta\phi$ spectra.

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 fuer Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Particle Physics and Astronomy Research Council and the Royal Society, U.K.; the Russian Foundation for Basic Research; the Comision Interministerial de Ciencia y Tecnologia, Spain; work supported in part by the European Community's Human Potential Programme under contract HPRN-CT-20002, Probe for New Physics; and this work was supported by Research Fund of Istanbul University Project No. 1755/21122001.

APPENDIX A: MONTE CARLO PARAMETERS FOR THE SECONDARY VERTEX TAG CORRELATION ANALYSIS

The specific Monte Carlo generator settings used for the secondary vertex tag correlation analysis are specified below. For a more explicit discussion of the Monte Carlo, consult Ref. [31].

1. PYTHIA

Version 6.203 of PYTHIA [17,18] was used for this analysis. Flavor creation events are generated with the process MSEL = 5, while flavor excitation and gluon splitting use MSEL = 1. The parameter PARP(67), which holds the value that is multiplied by the Q^2 of the hard scatter to determine the maximum virtuality of the initial-state shower, was used to control the amount of initial state radiation in the PYTHIA samples. Three different values of PARP(67) were used and for each setting, other PYTHIA parameters were manipulated to give the best match to the CDF data. The naming convention for the three ISR samples is given in Table XI. This tuning was done with the CTEQ5L [51] parton distribution functions and different PYTHIA parameters may be required to achieve the same tuning with a different set of parton distribution functions. Table XII gives the values of the parameters used for this tuning. Parameters not mentioned in Table XII were left at the default values for this version of PYTHIA.

TABLE XI. The naming convention for the three ISR samples for PYTHIA.

Sample name	ISR setting
High	PARP(67) = 4.0
Medium	PARP(67) = 3.0
Low	PARP(67) = 1.0

TABLE XII. The table above shows the PYTHIA setting used to tune the underlying event to data for the CTEQ5L parton distribution set and three different initial-state-radiation settings.

Parameter	Meaning	PARP(67)=3.0,4.0	PARP(67)=1.0
MSTP(81)	Multiple-parton interaction switch	1 (Multiple-parton interactions ON)	
MSTP(82)	Model of multiple-parton interactions	3 (Varying impact parameter assuming A single Gaussian matter distribution)	
PARP(82)	p_T turn-off when using single Gaussian model of multiple interactions	1.7	1.6
PARP(85)	Probability that a multiple-parton interaction produces two gluons with color connections to the ''nearest neighbors''	1.0	
PARP(86)	Probability that an MPI produces two gluons either as described above or as a closed gluon loop. The rest of the MPIs produce quark-antiquark pairs	1.0	
PARP(89)	Determines the reference energy E0	1800	

2. HERWIG

Version 6.400 of HERWIG [19] was used for this analysis. The flavor creation and flavor excitation samples were generated with IPROC = 1705, and the gluon splitting sample was made using IPROC = 1500. Only one HERWIG sample was generated for each production mechanism, again using the CTEQ5L parton distribution functions. $\Lambda_{\rm QCD}$ was set to 192 MeV and the CLPOW parameter was set to 1.25 to match the observed frequency of B baryons at CDF. All other HERWIG parameters were left at their default values for this version.

APPENDIX B: J/ψ–LEPTON CORRELATION MONTE CARLO EVENT SAMPLES

Various simulated data sets have been necessary for the measurement of $f_{\rm toward}$ in the J/ψ–lepton data sample. The follow section will detail the method of generating the simulated events used in the measurement.

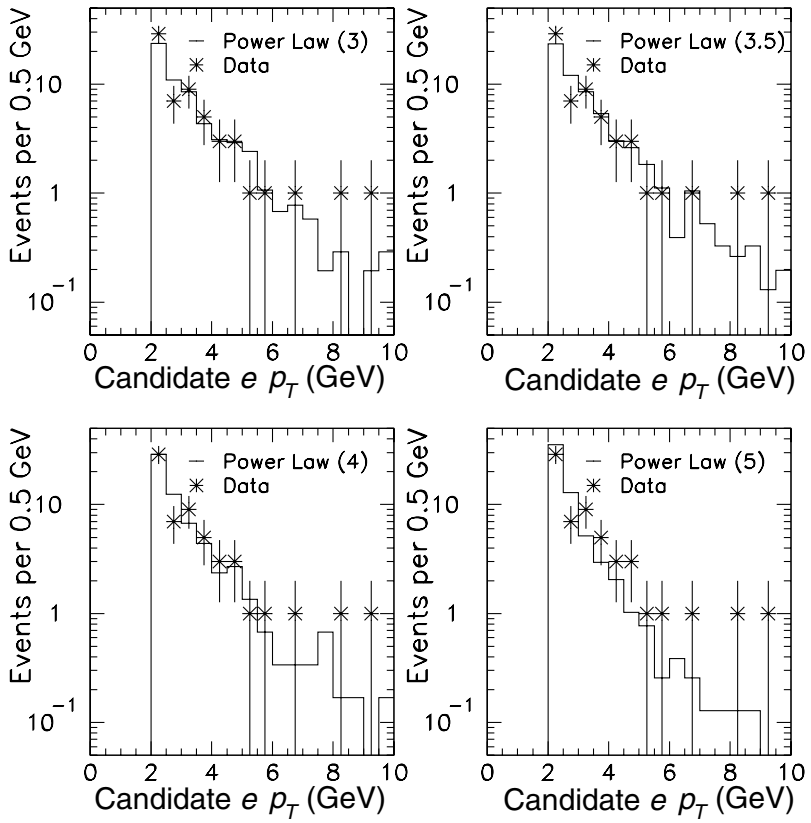

FIG. 25. The p_T spectra of the SLT conversion candidates. Top left: 3rd order power law. Top right: 3.5 order power law. Bottom Left: 4th order power law. Bottom right: 5th order power law.

For the bottom impact parameter description, PYTHIA [17,18] with the CTEQ3L [48] parton distribution functions is used. The bottom quarks are hadronized using the Bowler fragmentation function [52] and using the LUND string fragmentation model. The resulting bottom hadrons are decayed using the CLEO decay model [35]. The events are then passed through a detector simulation [53] and the trigger simulation. The same selection criteria is applied to the J/ψ candidates in Monte Carlo as in data.

For the bottom decay impact parameter shape for muons, the muons are required to be fiducial in both the CMU and the CMP muon subsystems with a SVX′ track and a $p_T > 3$ GeV.

For the bottom decay shapes for electrons, the electrons are required to be in the CEM fiducial region with a SVX′ quality track with a $p_T > 2$ GeV. The efficiency of the electron identification criteria is simulated in the same manner as Ref. [37]. The CPR, the CES, and the CTC dE/dx selection criteria do not depend on the isolation of the electron, due to the fine segmentation of the CPR, the CES and the CTC. Therefore, the efficiencies as a function of p_T of the CPR and the CES selection derived by Ref. [54] using conversions can be used. The CTC dE/dx efficiency as a function of p is defined by the selection criteria. The rate of signal events being removed as conversions and the $E_{\rm had}/E_{\rm EM}$ and E/p efficiencies depend on the isolation of the track. Therefore, the values simulated in the Monte Carlo have to be used to determine the efficiency.

For residual conversion studies and impact parameter shape determination, the necessary simulated sample was generated in the following manner. A sample of π^0 is simulated in the detector. A soft electron conversion candidate is required to have a found track with SVX′ information in the electron fiducial region with a $p_T > 2$ GeV. The efficiency of electron identification requirements is simulated in the same manner as for the bottom impact parameter shape sample.

The simulated π^0 are generated with a power law spectra for p_T and a flat η distribution. The order of the power law is varied in order to match the p_T spectra of the found SLT conversion candidates in data. The shape of the found pair candidates' p_T is used as a cross check of the power law description of the conversions. Figure 25 shows the Monte Carlo spectra normalized to the data for a power law of 3, 3.5, 4, and 5. The 3.5 order power law describes the data well and is used for the calculation of the efficiency. The 3rd and 4th order power law is used as a estimate of the systematic uncertainty.

The p_T of the pair candidates is shown in Fig. 26. The 3.5 order power law describes the shape of events with $p_T > 0.5$ GeV (where the tracking is assumed to be fully

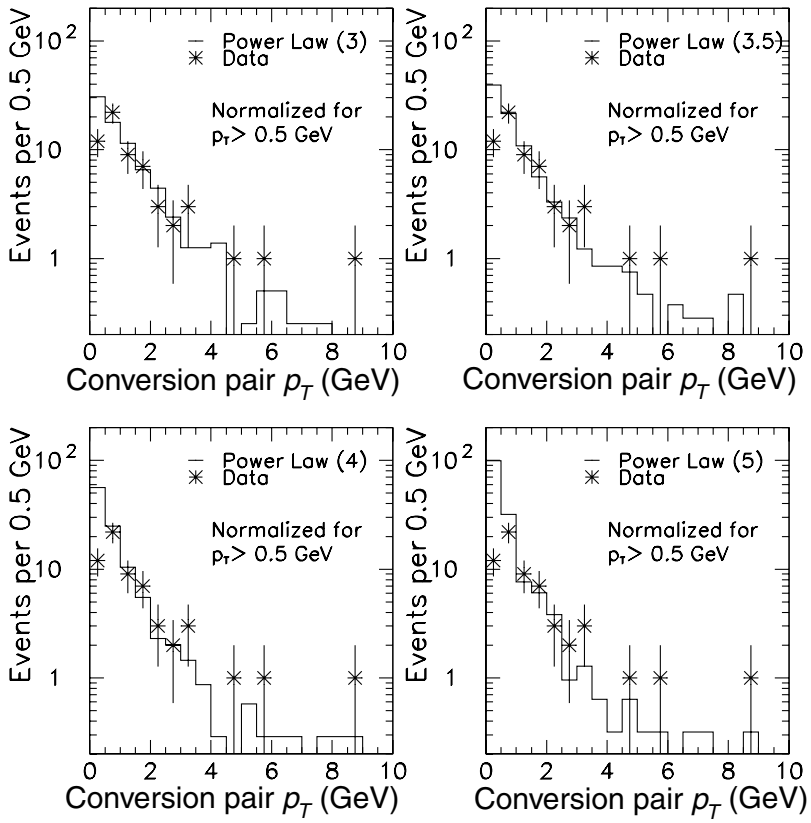

FIG. 26. The p_T spectra of the pair candidates. The Monte Carlo is normalized to the data with $p_T > 0.5$ GeV. Top left: 3rd order power law. Top right: 3.5 order power law. Bottom left: 4th order power law. Bottom right: 5th order power law.

efficient). Half the difference between the 3rd and 4th order power law spectra is used as the systematic uncertainty of the estimate.

The impact parameter-$c\tau$ shape of the B_c background was determined using a Monte Carlo sample. The B_c mesons are generated according to the NLO fragmentation model from Ref. [55] with a flat rapidity spectra. The particles are decayed using the semileptonic decay model of Ref. [56] and passed through a detector and trigger simulation. The selection criteria used is identical to Sec. IV C 2.

Single bottom quarks are generated according to the next-to-leading order QCD predictions by Ref. [57] and fragmented using the Peterson fragmentation model [58]. The resulting bottom hadrons are decayed using the CLEO decay model [35], requiring a $J/\psi \to \mu^+\mu^-$ decay. The events are then passed through a detector simulation [53] and the trigger simulation. For both the calculation of electron and muon fake rates, the J/ψ is required to pass the same selection criteria as data. The sample is normalized to the number of J/ψ events from bottom decay fit in data in sSec. IV C 1.

APPENDIX C: J/ψ–LEPTON CORRELATION LOG-LIKELIHOOD FUNCTION

The fit parameters will be in lower case, the constraints will be upper case, and the errors on the constraints (if applicable) are denoted Δ(Constraint). The superscripts indicate the additional lepton used (e for electrons, μ for muons) and the $\Delta\phi$ region (t for toward, a for away). For example, $N_{B_c}^{e,t}$ is the number of $B_c \to J/\psi eX$ background events estimated in the electron sample in the toward $\Delta\phi$ region.

1. Data

The inputs to the fit on an event-by-event basis are the J/ψ candidate's $c\tau$ and the additional lepton candidate's impact parameter. In the following sections, x will denote the impact parameter, and y will denote the $c\tau$. The numbers of candidates in the J/ψ mass sideband and signal regions in both $\Delta\phi$ regions are used as a constraint in the likelihood. In the electron sample, the numbers of found conversions in J/ψ mass sideband and signal regions in both $\Delta\phi$ regions also used as a constraint. Conversion constraints are discussed later in their respective sections.

$$P(n_{\text{signal}}, N_{\text{signal}}) = \frac{(n_{\text{signal}})^{N_{\text{signal}}}}{N_{\text{signal}}!}\, e^{-n_{\text{signal}}} \qquad \text{(C1)}$$

with the appropriate n_{signal} and N_{signal} for the given sample and $\Delta\phi$ region. n_{signal} is not a fit parameter, but is a function of the other fit parameters, shown in Appendix C 8.

2. $b \to J/\psi X; \overline{b} \to \ell X'$ signal

The shapes which are used for the $b \to J/\psi X; \overline{b} \to \ell X'$ signal are described by the fit functions in Secs. IV C 1 and IV C 2. The impact parameter and $c\tau$ are assumed to be uncorrelated. Therefore, the shape which describes the signal is the product of the impact parameter shape ($F_b^{d_0}(x)$) and the $c\tau$ shape ($F_b^{c\tau}(y)$) for bottom decay. The parameters that are used in $F_b^{d_0}(x)$ are different for the electron and muon fits.

The number of $b\overline{b}$ events fit is $n_{b\overline{b}}$ with the superscripts given by sample and $\Delta\phi$ region. For example, the number of $b\overline{b}$ events fit in the toward $\Delta\phi$ region in the electron sample is $n_{b\overline{b}}^{e,t}$. The $b\overline{b}$ contribution of the shape component of the likelihood is given by $\frac{n_{b\overline{b}}}{n_{\text{signal}}} \cdot F_b^{d_0}(x) \cdot F_b^{c\tau}(y)$ with the appropriate superscripts for the additional lepton type and $\Delta\phi$ region.

3. Unconstrained, uncorrelated backgrounds

The impact parameter-$c\tau$ shapes of the three sources of uncorrelated backgrounds without constraints, considered in this analysis, are constructed using the functions derived in Secs. IV C 1 and IV C 2. The fit parameters for these three backgrounds are:

(a) n_{dd}: the number of events with the J/ψ candidate and with the additional lepton candidate both directly produced,
(b) n_{bd}: the number of events with the J/ψ candidate from bottom decay and with the additional lepton candidate produced directly,
(c) n_{db}: the number of events with a directly produced J/ψ candidate and with an additional lepton candidate from bottom decay,

where the superscripts indicate of the sample and $\Delta\phi$ region. The number of events with a directly produced J/ψ candidate and an additional lepton candidate from bottom decay is assumed to be small and n_{db} is fixed to zero. This parameter is released and fit for as an estimate of systematic uncertainty due to this assumption.

The shape component of the likelihood for these three backgrounds is assembled in the same manner as the $b\overline{b}$ signal.

4. Residual conversion background

The total number of predicted residual conversions is $R_{\text{conv}} \cdot N_{\text{conv}}$, where R_{conv} is the ratio between the number of residual versus found conversions and N_{conv} is the number of found conversions in the sample. The number of found conversions removed from the two $\Delta\phi$ regions with the J/ψ candidate in the mass signal region is $N_{\text{conv}}^t = 6$ and $N_{\text{conv}}^t = 9$, respectively. In Sec. IV C 4, R_{conv} is estimated to be 1.00 ± 0.37, using data and Monte Carlo techniques. Residual conversions are assumed to pair with all three sources of uncorrelated J/ψ candidates: fake J/ψ

, directly produced J/ψ, and bottom decay J/ψ. The same value of the fit parameter $r_{\rm conv}$ is used for all sources of J/ψ candidates that pair with the residual conversions. The value of $r_{\rm conv}$ is constrained as a Gaussian probability in the likelihood.

$$G(r_{\rm conv} - R_{\rm conv}, \Delta R_{\rm conv}) = \frac{1}{\sqrt{2}\Delta R_{\rm conv}} e^{-1/2(\frac{r_{\rm conv}-R_{\rm conv}}{\Delta R_{\rm conv}})^2}. \quad \text{(C2)}$$

The fit parameters that set the scale for the number of residual conversions events with the J/ψ candidate from bottom decay and direct production are $n_{b\rm conv}$ and $n_{d\rm conv}$. The parameters represent the number of found conversions with the J/ψ candidate from the given source. The number of residual conversions fit from these two sources are $r_{\rm conv} \cdot n_{b\rm conv}$ and $r_{\rm conv} \cdot n_{d\rm conv}$. The number of residual conversions already included in the sideband shape component is $r_{\rm side} n^e_{\rm side} f^{d_0}_{\rm conv}$, where $f^{d_0}_{\rm conv} = \frac{r_{\rm conv} \cdot n_{\rm conv\ side}}{N_{\rm sideband}}$ is the fit fraction of J/ψ mass sideband events where the electron is a residual conversion.

Because of the relatively small number of residual conversions, fitting all three pairing of J/ψ candidate types with conversions is not possible. In order to constrain this component of the fit farther, the ratio of between $n_{b\rm conv}$ and $n_{d\rm conv}$ is assumed to be the same as the ratio between J/ψ mesons from bottom decay and J/ψ mesons produced directly (at the primary vertex). The fraction of J/ψ mesons from bottom decay is fit to be $16.6\% \pm 0.2\%$ in Sec. IV C 1, yielding the relationship $n^{t/a}_{b\rm conv} = 0.2 \cdot n^{t/a}_{d\rm conv}$. As an estimate of the systematic uncertainty, $n^{t/a}_{b\rm conv}$ and $n^{t/a}_{d\rm conv}$ are fixed to zero in separate fits in order to probe the full range of ratio $n^{t/a}_{b\rm conv}$: $n^{t/a}_{d\rm conv}$.

The number of found conversions in the two $\Delta\phi$ regions is used as a constraint on the fit of the residual conversions. The number of found conversions fit is the number of residual conversions fit divided by the ratio of residual versus found conversions:

$$n_{\rm conv} \equiv n_{b\rm conv} + n_{d\rm conv} + \frac{r_{\rm side} n^e_{\rm side} n_{\rm convside}}{N_{\rm sideband}}. \quad \text{(C3)}$$

The constraint using the number of found conversions is the Poisson probability of finding $N_{\rm conv}$ conversion candidates with a mean value of number of found conversion fit.

$$P(n_{\rm conv}, N_{\rm conv}) = \frac{(n_{\rm conv})^{N_{\rm conv}}}{N_{\rm conv}!} e^{-(n_{\rm conv})}. \quad \text{(C4)}$$

5. Fake J/ψ backgrounds

The fake J/ψ impact parameter-$c\tau$ background shape ($F^{\mu}_{\rm sideband}$) is determined in Sec. IV C 6 from a fit to the data for the muon sample. The predicted number of events for this background is the ratio between the number of fake J/ψ events in the J/ψ mass signal and sideband region ($R_{\rm side}$) times the number of events seen in data with the J/ψ candidate in the mass sideband regions ($N_{\rm side}$) for the given sample and $\Delta\phi$ region.

In Sec. IV B, the ratio is determined to be $R_{\rm side} = 0.501 \pm 0.044$ from a fit of the total J/ψ data sample. The same value for the fit parameter $r_{\rm side}$ is used in both $\Delta\phi$ regions in the sample, but can be different in the electron and muon samples. The fit value of $r_{\rm side}$ is constrained using a Gaussian factor in the likelihood function.

$$G(r_{\rm side} - R_{\rm side}, \Delta R_{\rm side}) = \frac{1}{\sqrt{2}\Delta R_{\rm side}} e^{-1/2(\frac{r_{\rm side}-R_{\rm side}}{\Delta R_{\rm side}})^2}. \quad \text{(C5)}$$

The corresponding fit parameter $n_{\rm side}$, for the given sample and $\Delta\phi$ region, is constrained using the Poisson probability of measuring $N_{\rm side}$ events for a sample with an average of $n_{\rm side}$ events.

$$P(n_{\rm side}, N_{\rm side}) = \frac{(n_{\rm side})^{N_{\rm side}}}{N_{\rm side}!} e^{-n_{\rm side}}. \quad \text{(C6)}$$

The contribution of the shape component of the likelihood is $\frac{r_{\rm side} n_{\rm side}}{n_{\rm signal}} \cdot F_{\rm side}(x, y)$ for the given sample and $\Delta\phi$ region.

The fake J/ψ background component in the electron sample is treated differently due to the presence of residual conversions in the background. The fake J/ψ shape is fit at the same time as the J/ψ signal region. The $f_{\rm conv}$ component of the fake J/ψ impact parameter–$c\tau$ shape is a composite of two variables which are constrained. $f_{\rm conv} = \frac{r_{\rm conv} \cdot n_{\rm convside}}{N_{\rm sideband}}$ where $r_{\rm conv}$, $n_{\rm convside}$, and $N_{\rm sideband}$ are the fit ratio of residual to found conversion, fit number of found conversions, and $N_{\rm sideband} = N^{e,t}_{\rm side} + N^{\mu,a}_{\rm side}$. The parameter $n_{\rm convside}$ is constrained by the number of conversions found in the sideband using the Poisson probability:

$$P(n_{\rm convside}, N_{\rm convside}) = \frac{(n_{\rm convside})^{N_{\rm convside}}}{N_{\rm convside}!} e^{-n_{\rm convside}}. \quad \text{(C7)}$$

6. $b_c \to J/\psi \ell X$ backgrounds

$B_c \to J/\psi \ell X$ background is predicted to only populate the toward region in $\Delta\phi$. The expected number of B_c events is constrained in the likelihood as a Gaussian probability factor.

$$G(n_{B_c} - N_{B_c}, \Delta N_{B_c}) = \frac{1}{\sqrt{2}\Delta N_{B_c}} e^{-1/2(\frac{n_{B_c}-N_{B_c}}{\Delta N_{B_c}})^2}, \quad \text{(C8)}$$

where ΔN_{B_c} is the positive-sided error of N_{B_c} if $(n_{B_c} - N_{B_c}) \geq 0.0$, and the negative-sided error otherwise.

7. $b \to J/\psi \ell_{\rm fake} X$ backgrounds

As in the B_c background, $b \to J/\psi \ell_{\rm fake} X$ background events are only expected to populate the toward region in $\Delta\phi$, and therefore the background is only fit for in the toward region in the two samples. The expected number is

$b \rightarrow J/\psi \ell_{\text{fake}} X$ events is used as a constraint to the likelihood in the same way as B_c.

$$G(n_{B_{\text{fake}}} - N_{B_{\text{fake}}}, \Delta N_{B_{\text{fake}}}) = \frac{1}{\sqrt{2}\Delta N_{B_{\text{fake}}}} e^{-1/2(\frac{n_{B_{\text{fake}}} - N_{B_{\text{fake}}}}{\Delta N_{B_{\text{fake}}}})^2}.$$

8. n_{signal} sums

The number of events fit in the J/ψ mass signal region is a function of fit parameters described previously in this section. Listed below are the functions for number of events fit for the two samples and $\Delta\phi$ regions.

$$n^{\mu,t}_{\text{signal}} = n^{\mu,t}_{b\bar{b}} + n^{\mu,t}_{bd} + n^{\mu,t}_{db} + n^{\mu,t}_{dd} + r^{\mu}_{\text{side}} \cdot n^{\mu,t}_{\text{side}} + n^{\mu,t}_{B_c} + n^{\mu,t}_{B_{\text{fake}}}, \tag{C9}$$

$$n^{\mu,a}_{\text{signal}} = n^{\mu,a}_{b\bar{b}} + n^{\mu,a}_{bd} + n^{\mu,a}_{db} + n^{\mu,a}_{dd} + r^{\mu}_{\text{side}} \cdot n^{\mu,a}_{\text{side}}, \tag{C10}$$

$$n^{e,t}_{\text{signal}} = n^{e,t}_{b\bar{b}} + n^{e,t}_{bd} + n^{e,t}_{db} + n^{e,t}_{dd} + r^{e}_{\text{side}} \cdot n^{e,t}_{\text{side}} + n^{e,t}_{B_c} + n^{e,t}_{B_{\text{fake}}} + r_{\text{conv}} \cdot n^{t}_{b\text{conv}} + r_{\text{conv}} \cdot n^{t}_{d\text{conv}}, \tag{C11}$$

$$n^{e,a}_{\text{signal}} = n^{e,a}_{b\bar{b}} + n^{e,a}_{bd} + n^{e,a}_{db} + n^{e,a}_{dd} + r^{e}_{\text{side}} \cdot n^{e,a}_{\text{side}} + r_{\text{conv}} \cdot n^{t}_{b\text{conv}} + r_{\text{conv}} \cdot n^{t}_{d\text{conv}}. \tag{C12}$$

9. Impact parameter–$c\tau$ shape component

The complete functions for the shape components of the fit are listed below for the two samples and $\Delta\phi$ regions. As a reminder, x is the additional lepton candidate's impact parameter and y is the J/ψ candidate's $c\tau$.

$$F^{\mu,t}_{\text{shape}}(x,y) = \frac{1}{n^{\mu,t}_{\text{signal}}}[n^{\mu,t}_{b\bar{b}} F^{c\tau,\mu}_{b}(y) F^{d_0,\mu}_{b}(x) + n^{\mu,t}_{bd} F^{c\tau,\mu}_{b}(y) F^{d_0,\mu}_{\text{direct}}(x) + n^{\mu,t}_{db} F^{c\tau,\mu}_{\text{direct}}(y) F^{d_0,\mu}_{b}(x) \tag{C13}$$

$$+ n^{\mu,t}_{dd} F^{c\tau,\mu}_{\text{direct}}(y) F^{d_0,\mu}_{\text{direct}}(x) + n^{\mu,t}_{B_{\text{fake}}} F^{\mu}_{B_{\text{fake}}}(x,y) + n^{\mu,t}_{B_c} F^{\mu}_{B_c}(x,y) + r^{\mu}_{\text{side}} n^{\mu,t}_{\text{side}} F^{\mu}_{\text{side}}(x,y)] \tag{C14}$$

$$F^{\mu,a}_{\text{shape}}(x,y) = \frac{1}{n^{\mu,a}_{\text{signal}}}[n^{\mu,a}_{b\bar{b}} F^{c\tau,\mu}_{b}(y) F^{d_0,\mu}_{b}(x) + n^{\mu,a}_{bd} F^{c\tau,\mu}_{b}(y) F^{d_0,\mu}_{\text{direct}}(x) + n^{\mu,a}_{db} F^{c\tau,\mu}_{\text{direct}}(y) F^{d_0,\mu}_{b}(x) \tag{C15}$$

$$+ n^{\mu,a}_{dd} F^{c\tau,\mu}_{\text{direct}}(y) F^{d_0,\mu}_{\text{direct}}(x) + r^{\mu}_{\text{side}} n^{\mu,a}_{\text{side}} F^{\mu}_{\text{side}}(x,y)] \tag{C16}$$

$$F^{e,t}_{\text{shape}}(x,y) = \frac{1}{n^{e,t}_{\text{signal}}}[n^{e,t}_{b\bar{b}} F^{c\tau,e}_{b}(y) F^{d_0,e}_{b}(x) + n^{e,t}_{bd} F^{c\tau,e}_{b}(y) F^{d_0,e}_{\text{direct}}(x) + n^{e,t}_{db} F^{c\tau,e}_{\text{direct}}(y) F^{d_0,e}_{b}(x) \tag{C17}$$

$$+ n^{e,t}_{dd} F^{c\tau,e}_{\text{direct}}(y) F^{d_0,e}_{\text{direct}}(x) + n^{e,t}_{B_{\text{fake}}} F^{e}_{B_{\text{fake}}}(x,y) + n^{e,t}_{B_c} F^{e}_{B_c}(x,y) \tag{C18}$$

$$+ r^{e}_{\text{side}} n^{e,t}_{\text{side}} F^{e}_{\text{side}}(x,y) + r_{\text{conv}} n^{t}_{b\text{conv}} F_{\text{conv}}(x) F^{c\tau,e}_{b}(y) + r_{\text{conv}} n^{t}_{d\text{conv}} F_{\text{conv}}(x) F^{c\tau,e}_{\text{direct}}(y)] \tag{C19}$$

$$F^{e,a}_{\text{shape}}(x,y) = \frac{1}{n^{e,a}_{\text{signal}}}[n^{e,a}_{b\bar{b}} F^{c\tau,e}_{b}(y) F^{d_0,e}_{b}(x) + n^{e,a}_{bd} F^{c\tau,e}_{b}(y) F^{d_0,e}_{\text{direct}}(x) + n^{e,a}_{db} F^{c\tau,e}_{\text{direct}}(y) F^{d_0,e}_{b}(x) \tag{C20}$$

$$+ n^{e,a}_{dd} F^{c\tau,e}_{\text{direct}}(y) F^{d_0,e}_{\text{direct}}(x) + r^{e}_{\text{side}} n^{e,a}_{\text{side}} F^{e}_{\text{side}}(x,y) + r_{\text{conv}} n^{a}_{b\text{conv}} F_{\text{conv}}(x) F^{c\tau,e}_{b}(y) \tag{C21}$$

$$+ r_{\text{conv}} n^{a}_{d\text{conv}} F_{\text{conv}}(x) F^{c\tau,e}_{\text{direct}}(y)] \tag{C22}$$

10. Bin constraints component

The constraints which are specific to a given region in $\Delta\phi$ and sample are listed below:

$$C^{\mu,t}_{\text{bin}} = P(n^{\mu,t}_{\text{signal}}, N^{\mu,t}_{\text{signal}}) \cdot P(n^{\mu,t}_{\text{side}}, N^{\mu,t}_{\text{side}}) \cdot G(n^{\mu,t}_{B_c} - N^{\mu,t}_{B_c}, \Delta N^{\mu,t}_{B_c}) \cdot G(n^{\mu,t}_{B_{\text{fake}}} - N^{\mu,t}_{B_{\text{fake}}}, \Delta N^{\mu,t}_{B_{\text{fake}}}) \tag{C23}$$

$$C^{\mu,a}_{\text{bin}} = P(n^{\mu,a}_{\text{signal}}, N^{\mu,a}_{\text{signal}}) \cdot P(n^{\mu,a}_{\text{side}}, N^{\mu,a}_{\text{side}}) \tag{C24}$$

$$C^{e,t}_{\text{bin}} = P(n^{e,t}_{\text{signal}}, N^{e,t}_{\text{signal}}) \cdot P(n^{e,t}_{\text{side}}, N^{e,t}_{\text{side}}) \cdot G(n^{e,t}_{B_c} - N^{e,t}_{B_c}, \Delta N^{e,t}_{B_c}) \tag{C25}$$

$$\times G(n^{e,t}_{B_{\text{fake}}} - N^{e,t}_{B_{\text{fake}}}, \Delta N^{e,t}_{B_{\text{fake}}}) \cdot P(n^{t}_{\text{conv}}, N^{t}_{\text{conv}}) \tag{C26}$$

$$C^{e,a}_{\text{bin}} = P(n^{e,a}_{\text{signal}}, N^{e,a}_{\text{signal}}) \cdot P(n^{e,a}_{\text{side}}, N^{e,a}_{\text{side}}) \cdot P(n^{a}_{\text{conv}}, N^{a}_{\text{conv}}) \tag{C27}$$

11. Global constraints component

The global constraints are the simplest component of the likelihood. The functions of the global constraints are listed below:

$$C^{\mu}_{\text{global}} = G(r^{\mu}_{\text{side}} - R_{\text{side}}, \Delta R_{\text{side}}) \tag{C28}$$

$$C^{e}_{\text{global}} = G(r^{e}_{\text{side}} - R_{\text{side}}, \Delta R_{\text{side}}) \tag{C29}$$

$$\times G(r_{\text{conv}} - R_{\text{conv}}, \Delta R_{\text{conv}}) \tag{C30}$$

12. Log-likelihood function

Finally, the log-likelihood can be assembled from the functions developed in the previous sections. The likelihood function for the muon sample is

$$\mathcal{L} = C^{\mu}_{\text{global}} \prod_{i}^{t,a} \left[C^{\mu,i}_{\text{bin}} \prod_{j}^{N^{\mu,i}_{\text{signal}}} (F^{\mu,i}_{\text{shape}}(x^{\mu}_{i,j}, y^{\mu}_{i,j})) \right],$$

where $x^{\mu}_{i,j}$ and $y^{\mu}_{i,j}$ are the impact parameter of the additional muon candidate and the $c\tau$ of the J/ψ candidate for j^{th} event in the i^{th} $\Delta\phi$ region in the J/ψ signal region.

The likelihood function for the electron sample is similar to the muon likelihood. The electron likelihood includes conversion terms and the fit of the J/ψ mass sideband region.

$$\mathcal{L} = C^{e}_{\text{global}} \prod_{i}^{t,a} \left[C^{e,i}_{\text{bin}} \prod_{j}^{N^{e,i}_{\text{signal}}} (F^{e,i}_{\text{shape}}(x^{e}_{i,j}, y^{e}_{i,j})) \right] \times P(n_{\text{conv side}}, N_{\text{conv side}}) \cdot \prod_{k}^{N_{\text{sideband}}} (F_{\text{conv}}(x^{e}_{k}, y^{e}_{k}))$$

where $x^{e}_{i,j}$ and $y^{e}_{i,j}$ are the impact parameter of the additional electron candidate and the $c\tau$ of the J/ψ candidate for j^{th} event in the i^{th} $\Delta\phi$ region in the J/ψ signal region, and x^{e}_{k} and y^{e}_{k} are the impact parameter of the additional electron candidate and the $c\tau$ of the J/ψ candidate for k^{th} event in the J/ψ sideband region.

APPENDIX D: TESTS OF J/ψ–LEPTON CORRELATION LOG-LIKELIHOOD FIT

The impact parameter–$c\tau$ likelihood fit is tested using a large set of toy Monte Carlo samples. First, the input means for the various fit components are chosen to be similar to the results in data. The constrained terms are chosen to be consistent with the constraint. These inputs are Poisson fluctuated to determine the composition of each sample. Each event is assigned an impact parameter and $c\tau$ according to the shape function used to describe that type of event.

Next, the fit constraints not yet varied (R_{side}, R_{conv}, $N_{B_{\text{fake}}}$, N_{B_c}, N^{t}_{conv}, N^{a}_{conv}, and $N_{\text{conv side}}$) are fluctuated using the appropriate statistic. The fluctuated constraints are then used in the fit of the toy Monte Carlo sample.

A total of 1000 samples are generated and fit for both the electron and muon samples. The fit values are not forced to be non-negative. The pull is calculated for each fit value relative to the nonfluctuated input quantities. The pull is equal to width of the $\frac{n-\mu}{\sigma_n}$ distribution where n is the fit value, σ_n is the fit error returned, and μ is the average value of the parameter input. The bias, which is the mean of the $\frac{n-\mu}{\sigma_n}$ distribution, is also measured. Finally, the average difference between the fitted value and input parameter is calculated, $\overline{(x - \overline{x})}$.

The pulls, biases, and average differences for all variables are acceptable for both test samples. All pulls are within $\pm 6\%$ of 1 and all biases are within $\pm 0.12\sigma$ of 0. Allowing the likelihood to have negative components yields fit results with meaningful fit values and errors.

Figure 27 shows the minimum log-likelihood distributions of both samples. We find that 19.6% of the muon toy Monte Carlo samples and 49.8% of the electron toy Monte Carlo samples have a higher minimum log-likelihood than the data. The minimum log-likelihood distributions along with the biases, pulls, and average differences give confidence that the likelihood is working properly.

The muon toy Monte Carlo samples have an input mean of $f^{\text{input}}_{\text{towards}} = 34.5\%$ and a fit mean of $f^{\text{fit}}_{\text{towards}} = 34.5 \pm 0.4\%$. The width of the fit f_{toward} distribution is $10.9 \pm 0.3\%$, which is consistent with the error seen in data of $+9.2 - 8.2\%$. The electron toy Monte Carlo

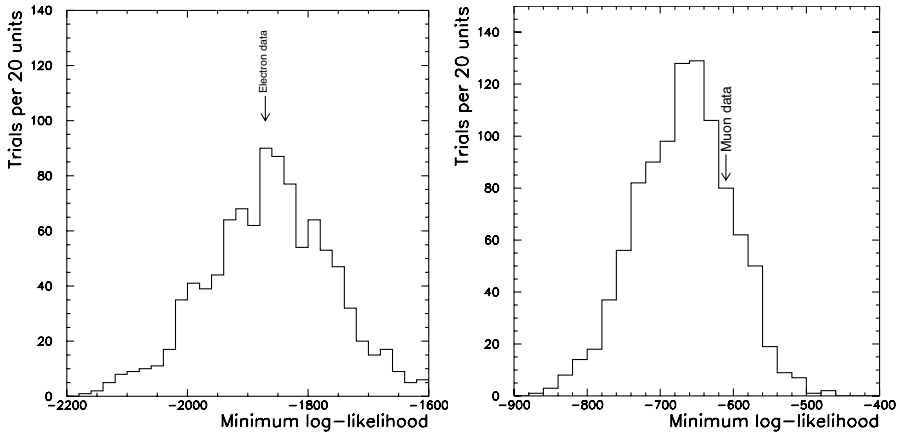

FIG. 27. The minimum log-likelihood distributions of the toy Monte Carlo left: Electrons right: Muons.

samples have an input mean of $f^{\text{input}}_{\text{towards}} = 19.2\%$ and a fit mean of $f^{\text{fit}}_{\text{towards}} = 18.6 \pm 0.2\%$. The width of the fit f_{toward} distribution is 6.0 ± 0.1, which is consistent with the error seen in data of $+6.5 - 5.8\%$.

[1] M. Mangano, P. Nason, and G. Ridolfi, Nucl. Phys. **B373**, 295 (1992).
[2] C. Albajar *et al.*, Phys. Lett. B **256**, 121 (1991).
[3] F. Abe *et al.*, Phys. Rev. Lett. **71**, 500 (1993).
[4] F. Abe *et al.*, Phys. Rev. Lett. **71**, 2396 (1993).
[5] F. Abe *et al.*, Phys. Rev. Lett. **75**, 1451 (1995).
[6] D. Acosta *et al.*, Phys. Rev. D **65**, 052005 (2002).
[7] S. Abachi *et al.*, Phys. Rev. Lett. **74**, 3548 (1995).
[8] B. Abbott *et al.*, Phys. Rev. Lett. **85**, 5068 (2000).
[9] C. Albajar *et al.*, Z. Phys. C **61**, 41 (1994).
[10] B. Abbott *et al.*, Phys. Lett. B **487**, 264 (2000).
[11] F. Abe *et al.*, Phys. Rev. D **53**, 1051 (1996).
[12] F. Abe *et al.*, Phys. Rev. D **55**, 2546 (1997).
[13] M. Cacciari and P. Nason, Phys. Rev. Lett. **89**, 122003 (2002).
[14] F. Halzen, W. Y. Keung, and D. M. Scott, Phys. Rev. D **27**, 1631 (1983).
[15] E. L. Berger *et al.*, Phys. Rev. Lett. **86**, 4231 (2001).
[16] D. Acosta *et al.*, Phys. Rev. D **69**, 072004 (2004).
[17] T. Sjostrand and M. Bengtsson, Comput. Phys. Commun. **43**, 367 (1987).
[18] H.-U. Bengtsson and T. Sjostrand, Comput. Phys. Commun. **46**, 43 (1987).
[19] G. Marchesini *et al.*, Comput. Phys. Commun. **67**, 465 (1992).
[20] R. D. Field, Phys. Rev. D **65**, 094006 (2002).
[21] F. Abe *et al.*, Phys. Rev. D **50**, 2966 (1994).
[22] T. Affolder *et al.*, Phys. Rev. D **64**, 032002 (2001).
[23] F. Abe *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **271**, 387 (1988).
[24] F. Abe *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **271**, 387 (1988).
[25] S. Cihangir *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **360**, 137 (1995).
[26] L. Balka *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **267**, 272 (1988).
[27] S. Bertolucci *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **267**, 301 (1988).
[28] G. Ascoli *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **268**, 33 (1988).
[29] G. Brandenburg *et al.*, CDF Public Note 6362, CDF (2002).
[30] F. Abe *et al.*, Phys. Rev. D **60**, 072003 (1999).
[31] K. Lannon, Ph.D. thesis, University of Illinois at Urbana-Champaign (2003), http://library.fnal.gov/archive/thesis/fermilab-thesis-2003-21.pdf.
[32] E. Norrbin and T. Sjostrand, Phys. Lett. B **442**, 407 (1998).
[33] E. Norrbin and T. Sjostrand, Eur. Phys. J. C **17**, 137 (2000).
[34] R. D. Field (private communication), http://www.phys.ufl.edu/~rfield.
[35] CLEO Collaboration, P. Avery *et al.*, Cornell Internal Note CSN-112 (1985).
[36] T. Affolder *et al.*, Phys. Rev. D **61**, 072005 (2000).
[37] F. Abe *et al.*, Phys. Rev. D **58**, 112004 (1998).
[38] A. Bardi *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **485**, 178 (2002).
[39] J. Marriner, CDF Public Note 1996, CDF (1993).
[40] F. Abe *et al.*, Phys. Rev. D **57**, 5382 (1998).
[41] T. Affolder *et al.*, Phys. Rev. Lett. **85**, 2886 (2000).
[42] D. E. Groom *et al.*, Eur. Phys. J. C **15**, 1 (2000).
[43] O. Long, Ph.D. thesis, University of Pennsylvania, 1998.
[44] J. Suzuki, Ph.D. thesis, University of Tsukuba, 1998.
[45] F. James and M. Roos, Comput. Phys. Commun. **10**, 343 (1975).
[46] K. Hagiwara *et al.*, Phys. Rev. D **66**, 010001 (2002).
[47] F. Abe *et al.*, Phys. Rev. D **61**, 032001 (2000).
[48] H. L. Lai *et al.*, Phys. Rev. D **51**, 4763 (1995).
[49] A. D. Martin *et al.*, Eur. Phys. J. C **4**, 463 (1998).
[50] L. Apanasevich *et al.*, Phys. Rev. D **59**, 074007 (1999).
[51] CTEQ Collaboratrion, H. L. Lai *et al.*, Eur. Phys. J. C **12**, 375 (2000).
[52] M. G. Bowler, Z. Phys. C **11**, 169 (1981).
[53] CDF Collaboration, M. Shapiro *et al.*, CDF Public Note 1810, (1992).
[54] F. Abe *et al.*, Phys. Rev. D **50**, 2966 (1994).
[55] E. Braaten *et al.*, Phys. Rev. D **48**, R5049 (1993).
[56] N. Isgur *et al.*, Phys. Rev. D **39**, 799 (1989).
[57] P. Nason *et al.*, Nucl. Phys. **B327**, 49 (1989).
[58] C. Peterson *et al.*, Phys. Rev. D **27**, 105 (1983).