First measurement of the forward-backward asymmetry in bottom-quark pair production at high mass

Aaltonen, T A; Amerio, S; Amidei, D E; et al.,; Canelli, F; Kilminster, B; CDF Collaboration,

Abstract: We measure the particle-level forward-backward production asymmetry in $b\bar{b}$ pairs with masses $m(b\bar{b})$ larger than 150 GeV/$c^2$, using events with hadronic jets and employing jet charge to distinguish $b$ from $\bar{b}$. The measurement uses $9.5 fb^{-1}$ of $p\bar{p}$ collisions at a center of mass energy of 1.96 TeV recorded by the CDF II detector. The asymmetry as a function of $m_{b\bar{b}}$ is consistent with zero, as well as with the predictions of the standard model. The measurement disfavors a simple model including an axigluon with a mass of 200 GeV/$c^2$ whereas a model containing a heavier 345 GeV/$c^2$ axigluon is not excluded.

DOI: [https://doi.org/10.1103/PhysRevD.92.032006](https://doi.org/10.1103/PhysRevD.92.032006)

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: [https://doi.org/10.5167/uzh-121420](https://doi.org/10.5167/uzh-121420)
Journal Article
Accepted Version

Originally published at:
Aaltonen, T A; Amerio, S; Amidei, D E; et al.; Canelli, F; Kilminster, B; CDF Collaboration, (2015). First measurement of the forward-backward asymmetry in bottom-quark pair production at high mass. Physical Review D (Particles, Fields, Gravitation and Cosmology), 92:032006.
DOI: [https://doi.org/10.1103/PhysRevD.92.032006](https://doi.org/10.1103/PhysRevD.92.032006)
First measurement of the forward-backward asymmetry in bottom-quark pair production at high mass

T. Aaltonen, 21 S. Amerio,39 D. Amidei,31 A. Anastassov,15 A. Anновi,17 J. Antos,12 G. Appollinari,15 J.A. Appel,15 T. Arisawa,52 A. Artikov,13 J. Asaadi,47 W. Ashmanskas,15 B. Auerbach,2 A. Aurisano,47 F. Azfar,38 W. Badgett,13 T. Bae,25 A. Barbaro-Galtieri,26 V.E. Barnes,43 B.A. Barnett,23 P. Barria,41 P. Bartos,12 M. Bauce,39 F. Bedeschi,41 S. Behr,15 G. Bellettiin,m,41 J. Bellinger,54 D. Benjamin,14 A. Beretvas,15 A. Bhatti,45 K.R. Bland,5 B. Blumenfeld,23 A. Bocci,14 A. Bodek,44 D. Bortoletto,43 J. Boudreau,42 A. Boveia,11 L. Brigliadori,6 C. Bromberg,32 E. Brucken,21 J. Budagov,13 H.S. Budd,44 K. Burkett,15 G. Busetto,39 P. Bussey,19 P. Buttimm,41 A. Buzatu,19 A. Calamba,10 S. Camarda,4 M. Campanelli,28 F. Canelli,11 B. Carls,22 D. Carlsmith,54 R. Carosi,41 S. Carrillo,16 B. Casal,9 M. Casarsa,48 A. Castro,46 P. Catastini,20 D. Cauz,48 V. Cavaliere,22 A. Cerri,26 L. Cerrito,28 Y.C. Chen,1 M. Chertok,7 G. Chiarelli,41 G. Chlachidze,15 K. Cho,25 D. Chokheli,13 A. Clark,18 C. Clarke,53 M.E. Convery,15 J. Conway,7 M. Corbo,15 M. Cordelli,17 C.A. Cox,7 D.J. Cox,7 M. Cremonesi,41 D. Cruz,47 J. Cuevas,9 R. Culbertson,15 N. d’Ascanzo,15 M. Datta,15 K. Hatakeyama,10 S. Donatii,41 M. D’Onofrio,27 M. Dorigo,44 A. Driutti,54 K. Ebina,52 R. Edgar,31 A. Elagin,47 R. Erbacher,7 S. Errede,22 B. Esham,22 S. Farrington,38 J.P. Fernández Ramos,29 R. Field,16 G. Flanagan,15 R. Forrest,7 M. Franklin,20 J.C. Freeman,15 H. Frisch,11 Y. Funakoshi,52 C. Galloni,41 A.F. Garfinkel,43 P. Garosi,41 H. Gerberich,22 E. Gerchtein,15 S. Giagu,46 V. Giakoumopoulos,3 K. Gibson,42 C.M. Ginsburg,15 N. Giokaris,3 P. Giromini,17 V. Glagolev,13 D. Glenzinski,15 M. Gold,34 D. Goldin,47 A. Golossanov,15 G. Gomez,9 G. Gomez-Ceballos,30 M. Goncharov,30 O. González López,29 I. Gorelov,34 A.T. Goshaw,14 K. Goulianos,45 E. Gramellini,6 C. Grosso-Pilcher,11 R.C. Group,51,15 J. Guimaraes da Costa,20 S.R. Halm,15 J.Y. Han,44 F. Happacher,17 K. Hara,49 M. Hare,50 R.F. Harr,53 T. Harrington-Taber,15 K. Hatakeyama,35 C. Hays,38 J. Heinrich,40 S. Henry,47 M. Herndon,54 A. Hocker,15 Z. Hong,47 W. Hopkins,15 S. Hou,1 R.E. Hughes,35 U. Husemann,55 M. Huschle,32 J. Huston,32 G. Intorzia,54 M. Iori,41 R. Ivanov,7 E. James,15 D. Jang,10 B. Jayatilaka,15 E.J. Jeon,25 S. Jindariani,15 M. Jones,43 K.K. Joo,25 S.Y. Jun,10 T.R. Junk,15 M. Kambit,24 T. Kamon,25 P.E. Karchin,15 A. Kasni,5 Y. Kato,9,37 W. Ketchum,11 J. Keung,40 B. Kilminster,15 D.H. Kim,25 H.S. Kim,26 J.E. Kim,25 M.J. Kim,17 S.H. Kim,49 S.B. Kim,25 Y.J. Kim,25 Y.K. Kim,11 N. Kimura,52 M. Kirby,15 K. Knoepfel,15 K. Kondo,52,49 D.J. Kong,29 J. Konigsberg,16 A.V. Kotwal,14 M. Kreps,24 J. Kroll,40 M. Kruse,14 T. Kuhn,24 M. Murata,49 A.T. Laasanen,43 S. Lammel,15 M. Lancaster,28 K. Lannon,35 G. Latino,41 H.S. Lee,25 J.S. Lee,25 S. Leo,22 S. Leone,41 J.D. Lewis,15 A. Limosani,14 E. Lipeles,40 A. Lister,18 H. Liu,51 Q. Liu,43 T. Liu,15 S. Lockwitz,55 A. Loginov,55 D. Luchesis,17 A. Lucà,17 J. Lueck,24 P. Lukian,26 P. Lukens,15 G. Lungu,15 J. Lys,26 R. Lysak,12 R. Madrak,15 P. Maestro,41 S. Malik,45 G. Manca,17 L. Manousakis-Katsikakis,3 L. Marchese,39 F. Margaroli,46 P. Marineo,41 K. Matera,22 M.E. Mattson,53 A. Mazzacane,15 P. Mazzanti,8 R. McNulty,23,27 A. Mehta,27 P. Mehtala,23 C. Mesropian,45 T. Miao,15 D. Mietlicki,11 A. Mitra,1 H. Miyake,49 S. Moed,15 N. Moggi,6 C.S. Moon,15 R. Moore,39 J.M. Morello,41 A. Mukherjee,15 Th. Muller,24 P. Murat,18 M. Mussini,6 J. Nachtmann,57 Y. Nagai,49 J. Nagano,52 I. Nakano,46 A. Napier,50 J. Nett,47 C. Neu,51 T. Nigman,42 L. Nodulman,25 S.H. Oh,25 Y.D. Oh,25 I. Oksuzian,51 T. Okusawa,37 R. Orava,21 L. Oortolani,4 C. Pagliarone,48 E. Palencia,19 P. Palni,34 V. Papadimitriou,15 W. Parker,54 G. Pauletta,54 M. Paulini,10 C. Paus,30 T.J. Phillips,14 G. Picentino,15 E. Pianori,13 J. Pilot,7 K. Pitts,22 C. Plager,8 L. Pondrom,54 S. Poprocki,15 K. Potamianos,26 A. Pranko,26 F. Prokoshin,16 J. Ptollos,8,17 G. Punzi,41 I. Redondo Fernández,28 P. Renton,38 M. Rescigno,46 G. Risse,5,18 L. Ristori,41 A. Robson,19 T. Rodriguez,40 S. Rolli,30 M. Ronzani,41 R. Roser,15 J.L. Rosner,11 F. Ruffini,41 A. Ruiz,9 J. Russ,10 V. Rusu,15 W.K. Sakamoto,44 Y. Sakurai,52 L. Santi,54,48 K. Sato,49 V. Saveliev,41 A. Savoy-Navarro,15 P. Schlabach,15 E.E. Schmidt,15 T. Schwarz,15 L. Scodellaro,4 F. Scuru,41 S. Seidel,34 Y. Seiya,13 A. Semenov,13 F. Sforza,41 S.Z. Shalhout,7 T. Shears,27 P.F. Shepard,12 M. Shimojima,49 M. Shochet,11 I. Shreyber-Tecker,33 A. Simonenko,15 K. Sliwa,10 J.R. Smith,7 F.D. Snider,15 H. Song,42 V. Sorin,4 R. St. Denis,19 M. Stancari,15 D. Steutz,15 J. Strologas,34 Y. Sudo,49 A. Sukhanov,15 I. Suslov,13 K. Takemasa,49 Y. Takeuchi,49 J. Tang,13 M. Tecchio,31 P.K. Teng,17 J. Thom,17 E. Thomson,40 V. Thukral,47 D. Tokar,32 K. Tollesfon,32 T. Tomura,49 D. Tonelli,15 S. Torre,17 D. Torretta,15 P. Totaro,39 M. Trovato,41 F. Ugekawa,49 S. Uozumi,25 F. Vázquez,16 G. Veley,15 C. Vellidis,15 C. Vernieri,41 M. Vidal,43 R. Vilar,9 J. Vizán,55 M. Vogel,34 G. Volpi,17 P. Wagner,40 R. Wallny,15 S.M. Wang,1 D. Waters,28 W.C. Wester III,15 D. Whiteson,40 A.B. Wicklund,2 S. Wilbur,7 H.H. Williams,40
We measure the particle-level forward-backward production asymmetry in $b\bar{b}$ pairs with masses ($m_{b\bar{b}}$) larger than 150 GeV/$c^2$, using events with hadronic jets and employing jet charge to distinguish $b$ from $\bar{b}$. The measurement uses 9.5 fb$^{-1}$ of $p\bar{p}$ collisions at a center of mass energy of 1.96 TeV recorded by the CDF II detector. The asymmetry as a function of $m_{b\bar{b}}$ is consistent with zero, as well as with the predictions of the standard model. The measurement disfavors a simple model including an axigluon with a mass of 200 GeV/$c^2$ whereas a model containing a heavier 345 GeV/$c^2$ axigluon is not excluded.

PACS numbers: 14.65.Ha

I. INTRODUCTION

In recent years, the values of the forward-backward asymmetry ($A_{FB}$) of top-quark-pair production measured at the Tevatron proton-antiproton collider at Fermilab have been consistently larger [1] than those predicted by the standard model (SM) [2–7]. Further study of this phenomenon has led to a number of proposed extensions of the SM [8]. One specific class of such models is the low-mass axigluon [9]. These models include a new axial-vector boson with a mass below the $t\bar{t}$ threshold and a natural width broad enough to evade detection in light-quark resonance searches, and predict a non-zero forward-backward asymmetry due to the interference between amplitudes mediated by the gluon and the axigluon. In the simplest models, the axigluon has equal couplings to all the quarks. Hence, a stringent test of such models is the measurement of the forward-backward asymmetry of pair production of other quark flavors, such as bottom. The forward-backward asymmetry for fermion-antifermion production is defined as

$$A_{FB} = \frac{n_F - n_{\bar{F}}}{n_F + n_{\bar{F}}}$$

where $n_F$ is the number of events where the fermion is forward of the antifermion in rapidity [10] ($\Delta y = y_F - y_{\bar{F}} >$...
0) and \( n_y \) the number where it is backward \( (\Delta y < 0) \). This definition is invariant under boosts along the beam axis.

At hadron colliders, \( b\bar{b} \) pairs are almost exclusively produced by the strong interaction of quarks and gluons (QCD). The vast majority of \( b\bar{b} \) pairs are produced via gluon-gluon fusion, which yields no asymmetry due to the symmetric initial state. In the \( q\bar{q} \rightarrow b\bar{b} \) process, a positive \( A_{FB}(b\bar{b}) \) arises from higher-order QCD corrections involving either real or virtual gluons. Since the valence-quark parton-density functions dominate over the gluon parton-density functions at large Bjorken \( x \), we enhance the contribution of the quark-antiquark initial state over the symmetric gluon-fusion background by requiring large values of \( b\bar{b} \) mass, \( m_{b\bar{b}} > 150 \text{ GeV}/c^2 \).

A number of theoretical predictions of the SM value for \( A_{FB}(b\bar{b}) \) at high mass have been reported, using various techniques and kinematic requirements and yielding a range of predictions [4, 7, 11]. Kühn and Rodrigo [4] computed that the asymmetry for \( b\bar{b} \) pairs with \( \sqrt{s} \geq 300 \text{ GeV} \) and production angle \( |\cos \theta^*| < 0.9 \) falls in the range 4.3% to 5.1%. Manohar and Trotta [7] computed \( A_{FB}(b\bar{b}) \) in several bins of \( m_{b\bar{b}} \). They found \( A_{FB}(b\bar{b}) = 0.4 \% \) inclusively and \( A_{FB}(b\bar{b}) = 7.8 \% \) for 350 < \( m_{b\bar{b}} \) < 650 GeV/c\(^2\), depending on the choice of the factorization and renormalization scales. Grinstein and Murphy [11] compute predictions using a variety of kinematic requirements, including an estimate using our kinematic requirements. The results of this calculation [11] are summarized in Table I.

Models containing a low-mass axigluon that has the same couplings to all quark flavours also predict a forward-backward asymmetry in bottom-quark-pair production. This asymmetry arises from the interference between the gluon and axigluon in the same fashion as the \( Z/\gamma^* \) interference of the Drell-Yan process [12]. One key feature of this interference is that the \( A_{FB} \) changes sign at the mass pole of the heavy axial resonance [9]. Observation of such a sign flip would provide significant indication of the underlying dynamics. The predictions of two different representative axigluon models are given in Table I [9]. A relatively light 200 GeV/c\(^2\) axigluon has been previously studied in the context of the top-quark asymmetry [13], and a heavier 345 GeV/c\(^2\) axigluon is of interest because its mass is just below the \( tt \) threshold. Both of these axigluons are assumed to have a width equal to 25% of their mass.

The predictions of both the SM and the models of non-SM physics under consideration are estimated at the parton level, and do not explicitly include the effects of hadronization.

Although the LHCb experiment has measured a related quantity [14], \( A_{FB} \) can only be measured in proton-antiproton collisions [15]. At the Tevatron, the D0 experiment has measured the forward-backward asymmetry of \( B^\pm \) meson production [16]. The momenta of the \( B^\pm \) mesons in that measurement constrain the masses of the \( b\bar{b} \) pairs to be smaller than the range probed in this measurement.

**II. EXPERIMENTAL APPARATUS AND EVENT SELECTION**

The Fermilab Tevatron is a proton-antiproton collider with a center-of-mass energy \( \sqrt{s} = 1.96 \text{ TeV} \). The CDF II detector [17] is an azimuthally symmetric magnetic spectrometer with a large tracking volume inside a solenoid. Outside the solenoid are sampling calorimeters. The calorimeters are further surrounded by the steel flux return of the solenoid and muon detectors [17].

The data used for this analysis were collected by the CDF II detector using three online event selections (triggers), which require at least one jet with transverse energy \( E_T > 50, 70, \text{ or } 100 \text{ GeV} \), respectively [10]. To control the trigger rate, the two lower-threshold triggers only accept one event out of every 100 or 8 events that satisfy their requirements, respectively, while the highest-threshold trigger accepts all events meeting its requirements. After data-quality requirements and trigger acceptance rates, the integrated luminosities of the samples are \( 95 \text{ pb}^{-1}, 1.2 \text{ fb}^{-1} \), and \( 9.5 \text{ fb}^{-1} \), respectively.

The offline event selection requires at least two jets [18, 19] with \( E_T \geq 20 \text{ GeV} \) and pseudorapidity \( |\eta| < 1.1 \). Of the jets that pass these requirements, exactly two must contain a secondary vertex (separated transversely from the primary pp interaction vertex) consistent with the decay of a \( b \) quark (\( b \)-tagged jets), identified using a secondary-vertex identification algorithm [20]. The observed invariant mass \( m_{b\bar{b}} \) of the two \( b \)-tagged jets is required to be at least 150 GeV/c\(^2\).

In order to maximize the statistical significance of the analysis, we separate the data into several subsamples depending on \( m_{b\bar{b}} \) (three subsamples, see Sec. III), on the estimated charges of the partons that produced the jets (four subsamples, see Sec. III A), and on the quality of the \( b \) tags (three subsamples, see Sec. III C), yielding a total of 36 subsamples.

**III. METHODOLOGY**

We discriminate jets originated from \( b \) quarks (\( b \) jets) from jets originated from \( b \) quarks (\( b \) jets) using the momentum-weighted average of the charges of the particles associated with each jet (jet charge),

\[
Q_{jet} = \frac{\sum_i q_i (\hat{p}_i \cdot \vec{p}_{jet})^{0.5}}{\sum_i (\hat{p}_i \cdot \vec{p}_{jet})^{0.5}},
\]

where the sum \( i \) is over all tracks in the jet, \( q_i \) is the charge of the corresponding particle, and \( \hat{p}_i \) and \( \vec{p}_{jet} \) are the momentum vectors of the particle and of the jet, respectively. The exponent 0.5 was chosen to maximize the power of the jet charge in separating \( b \) jets from \( b \) jets [21].
The asymmetry depends on the mass of the $b\bar{b}$ pair, both in the SM and in models with an axigluon. To study the behavior of $A_{FB}$ as a function of $m_{b\bar{b}}$, we divide the sample into several ranges (bins) of $m_{b\bar{b}}$. The choice of $m_{b\bar{b}}$ bins is motivated by the trigger. Each jet-energy-trigger threshold efficiently selects events only over a limited range of $m_{b\bar{b}}$. Events with $150 < m_{b\bar{b}} < 225$ GeV/$c^2$ are selected with a jet transverse-energy threshold of 50 GeV, and the thresholds of 70 and 100 GeV are used to select events with $m_{b\bar{b}}$ in the ranges $225 < m_{b\bar{b}} < 325$ GeV/$c^2$ and 325 GeV/$c^2 < m_{b\bar{b}}$, respectively.

The measurement of $A_{FB}(b\bar{b})$ must account for non-$b\bar{b}$ backgrounds, detector effects, and the dilution of the asymmetry due to misidentification of $b$ jets as $\bar{b}$ jets and vice versa. We therefore define a detector-level $A_{FB}$,

$$A_{FB}^{det} = \frac{1}{2P - 1} \left( \frac{N_F - N_{F}^{bkgd}}{N_F - N_{F}^{bkgd} + N_B - N_{B}^{bkgd}} - \frac{N_B - N_{B}^{bkgd}}{N_F - N_{F}^{bkgd} + N_B - N_{B}^{bkgd}} \right),$$

(3)

where $N_F$ and $N_B$ are the observed numbers of events, $N_{F}^{bkgd}$ and $N_{B}^{bkgd}$ are the estimated numbers of background events, and $P$ is the probability to make the charge assignment correctly. To account for additional effects, such as the finite resolution and acceptance of the CDF II detector, we employ a Bayesian technique to measure the $A_{FB}$ at particle level. “Particle level” refers to quantities reconstructed from final-state, color-confined particles with lifetimes greater than 10 ps [22].

The background levels are determined with a data-driven technique based on the $b$-tagged vertex mass as described in Sec. III C. The calibration of the charge misidentification and the calculation of $P$ is presented in Sec. III A, and the background asymmetries are discussed in Sec. III D. Finally, we use a Markov-chain Monte Carlo calculation to derive particle-level results by identifying the maxima of the marginalized posterior probability densities and by constructing associated credible intervals. The correction to the particle level, including the effects of mismeasurement and acceptance, are described in Sec. III G.

### A. Identification of the $b$ jet using jet charge

The forward or backward assignment is performed using the momentum-weighted track charge, or “jet charge” (see Eq. (2)), for each of the two $b$-tagged jets. Distributions of the jet charge ($Q_{jet}$) for $b$ jets and $\bar{b}$ jets are shown in Fig. 1. We use the difference of the two jet charges $\Delta Q = Q_1 - Q_2$ to make the assignment: if $\Delta Q$ is negative, the jet with charge $Q_1$ is considered to be the $b$ jet (because the $b$ quark is negatively charged), and if $\Delta Q$ is positive, the jet with charge $Q_2$ is considered to be the $\bar{b}$ jet.

The performance of the jet-charge algorithm is calibrated with data, by dividing the distribution of $Q_{jet}$ into bins. The bin edges are $-0.25, 0,$ and $0.25$, for a total of four bins (the lowest- and highest-charge bins are open). We arbitrarily assign $Q_{jet}$ values of $-0.5, -0.25, 0.25,$ and $0.5$ to the jets falling into each respective bin. This translates into five bins of $|\Delta Q|$ with values $0, 0.25, 0.5, 0.75,$ and $1.0$. The bin with $|\Delta Q| = 0$ is not informative because it gives no indication of how to assign the $b$ jet, so only four bins of $|\Delta Q|$ are used.

### B. Charge-identification probability

The probability $P$ of correctly assigning the $b$ jet and $\bar{b}$ jet, introduced in Eq. (3), is estimated from the data. We evaluate this in each bin of $|\Delta Q|$, so that

$$P_{0.25} = \frac{p_{0.5}(1 - p_{0.25})}{p_{0.5}(1 - p_{0.25}) + (1 - p_{0.5})p_{0.25}},$$

$$P_{0.5} = \frac{p_{0.5}^2 + (1 - p_{0.5})^2}{p_{0.25}p_{0.5} + (1 - p_{0.25})(1 - p_{0.5})},$$

and

$$P_{0.75} = \frac{p_{0.5}^2}{p_{0.5}^2 + (1 - p_{0.5})^2},$$

(4)

where the subscript of $P$ indicates $|\Delta Q|$, and the various $p_{|Q_{jet}|}$ are the probabilities that a jet with a binned charge of $|Q_{jet}|$ is correctly identified as $b$ or $\bar{b}$ by the sign of $Q_{jet}$. These expressions are derived by exhaustively considering all the possibilities. For example, $|\Delta Q| = 1.0$ implies that one jet has $Q_{jet} = +0.5$ and the other jet has $Q_{jet} = -0.5$. If the $b$ jet has charge $-0.5$, then the use of $\Delta Q$ allows for a correct assignment of the $b$ and $\bar{b}$ jets. This case occurs with probability $p_{0.5}^2$, because both jets (with $|Q_{jet}| = 0.5$) have a jet charge whose sign matches the sign of the charge of the originating quark. The other possibility is that we misidentify both jets, which occurs with probability $(1 - p_{0.5})^2$. The denominator in Eq. (4) is the sum of these two probabilities, and the numerator only contains the correct case.

In order to measure $p_{|Q_{jet}|}$, we measure the number of opposite-charge events, $N_{OC}$, and number of same-charge events, $N_{SC}$, in the data in each $m_{b\bar{b}}$ subsample. These
numbers are corrected for the presence of background using the calibrated \( b \)-fractions (see Sec. III C) and the background model described in Sec. III D. We assume that the remaining sample is composed of \( \bar{b}b \) with no contamination from \( bb \) or \( \bar{b}b \), and compute the opposite-charge fraction \( F_{OC} = N_{OC}/(N_{OC} + N_{SC}) \).

In events in which both jets have \( |Q_{jet}| = 0.25 \), the opposite-sign fraction is expressed as

\[
F_{OC}^{0.25−0.25} = p_{0.25}^2 + (1 − p_{0.25})^2,
\]

where the term \( p_{0.25}^2 \) arises from events in which both \( b \) jets have the correct sign, and the term \( (1 − p_{0.25})^2 \) arises from events in which both \( b \) jets have the wrong sign. Similarly, we have

\[
F_{OC}^{0.25−0.5} = p_{0.25}p_{0.5} + (1 − p_{0.25})(1 − p_{0.5}), \quad \text{and (6)}
\]

\[
F_{OC}^{0.5−0.5} = p_{0.5}^2 + (1 − p_{0.5})^2, \quad \text{(7)}
\]

for events containing one jet with \( |Q_{jet}| = 0.25 \) and one jet with \( |Q_{jet}| = 0.5 \), and for events in which both jets have \( |Q_{jet}| = 0.5 \), respectively. We measure each \( F_{OC} \) value in the data and solve for each \( p_{|Q_{jet}|} \), and then for each of the four \( F_{\Delta Q} \).

### C. Sample purity

Equation (3) also requires knowledge of the rate of background events, \( N_f^{bkg} \) and \( N_{bkg} \). We obtain the background yields by estimating the number of \( \bar{b}b \) events \( N_{\bar{b}b} \) and subtracting it from the number of events observed.

We divide the data into subsamples with varying \( b \)-tag quality, which in turn provides subsamples of varying \( \bar{b}b \) purity and improves the statistical power of the measurement. Since we apply \( b \)-tags to jets based on the presence of a secondary vertex, the quality of the \( b \)-tag is based on the confidence level of the identification of the secondary vertex. Specifically, the significance is based on the distance \( L_{2D} \), in the plane perpendicular to the beam, between the primary and secondary vertices, projected onto the jet momentum [20]. Jets with a significance \( \langle |L_{2D}|/\sigma(L_{2D}) \rangle \) greater than 20 are high significance or “H” tags, and jets with a lower significance are “L” tags. This results in the following independent subsamples ranked in order of increasing \( \bar{b}b \) purity: LL (both jets are L tagged), LH (one L and one H tag), and HH (both H tags).

We estimate the number of events in which both \( b \)-tagged jets are genuine \( b \) jets by counting events in which one or both of the \( b \)-tagged jets have a negative \( L_{2D} \). These “negative” tags are predominantly false tags from light-flavor jets and are a consequence of the finite position resolution of the tracking system. We expect the rate of false tags from this source to be equal for positive and negative tags. There are additional false positive tag contributions from hyperon decays and from interactions between high-momentum particles and the detector material. This results in an excess of positive over negative false tags. We exploit this relationship between the number of positively- and negatively-tagged light-flavor jets and compute the number of true \( \bar{b}b \) events using

\[
N_{\bar{b}b} = \frac{1}{\xi}(N_{++} − \lambda N_{−−} + \lambda^2 N_{−−}), \quad (8)
\]

where \( N_{++} \) is the number of observed positive double-tag events, \( N_{−−} \) is the number of events with one of the tags negative, \( N_{−−} \) is the number of events with both tags negative, \( \lambda \) is the ratio of positive to negative false tags, and \( \xi = 1 − 2\lambda r + \lambda^2 r^2 \) is a factor that accounts for the presence of negatively-tagged \( b \) jets. Here, \( r = \epsilon_{\bar{b}}/\epsilon_{b} \) is the ratio of the negative to positive tag efficiencies for \( b \) jets, estimated using Monte Carlo simulation and data. Finally, \( N_{\bar{b}b} \) is corrected by a scale factor derived from MC to account for a bias resulting from the presence of charm jets [23].

![Graph](image-url)

**FIG. 2.** Distribution of net \( m_{VTX} \), for H (a) and L (b) tagged data, with fit overlaid. The negative tag rates are subtracted from the positive tag rates. The histograms are stacked on top of one another.

We measure the ratio of positive to negative light-flavor tag rates \( \lambda \) following Ref. [23]. Since the MC indicates no strong dependence of \( \lambda \) on jet \( E_T \), we use the same value for all three dijet mass bins. We study the distribution of the invariant mass of all charged particles associated with the secondary vertex, \( m_{VTX} \) [20], for both negative and positive tags. We subtract the negative tag distribution...
from the positive tag distribution, yielding the “net” distribution. The relative rates of MC-derived templates for $\delta$, $c$, and light-flavor jets, left as free parameters, are fit in the net $m_{\mathrm{VTX}}$ distribution observed in data (see Fig. 2).

These net flavor fractions allow for the normalization of the complete positive- and negative-tagged distributions of $m_{\mathrm{VTX}}$ for all jet flavors. We find that the negative-tagged component of the MC templates must be scaled by 1.5 (1.3) to match the total number of low (high) significance negative tags observed in data. This scaling is propagated into the positive-tagged region of the templates as this population of “fake” tags is considered to be positive-negative symmetric, leaving the original “net” distributions unchanged [20]. From these corrected templates, we then obtain $\lambda$ for low- and high-significance $b$-tagged jets. We also correct the values of $r$ used in the three dijet mass bins for the higher negative tag rates observed in the data.

D. Background model

The non-$b\bar{b}$ background includes a wide variety of physics processes, such as $b+\text{mistag}$, $u\bar{u}$, $d\bar{d}$, gluon jets, etc. Of primary concern is the scattering of the valence quarks, which proceed through the $t$-channel exchange of a gluon and exhibits a forward-backward asymmetry due to the forward Rutherford peak. Rather than attempting to tune the simulation to reproduce all of these processes in the proper amounts, we use data in a control region that is expected to be enhanced in background and depleted of $b\bar{b}$ signal. The control region is defined by loosening the requirements of the $b$-tagging algorithm (looser than the low-significance “L” tag), and requiring that at least one of the tagged jets is negatively tagged. The sample of events satisfying these criteria is non-overlapping with the signal sample, and is composed almost entirely of background but is kinematically very similar to the signal sample. The forward-backward asymmetry of this sample is assumed to approximate the forward-backward asymmetry of the background in the signal sample (see Fig. 3).

The asymmetry of the background is negative, consistent with $u-\bar{u}$ scattering as the most important component. The $u$ quark tends to follow the incoming proton direction in $t$-channel scattering which would seem to yield a positive $A_{FB}$, but the opposite charge of the $u$ and $b$ quarks, combined with the definition of the asymmetry in terms of the jet charge, reverses the sign of the asymmetry.

Together with the background yields, the estimated background asymmetries are used to estimate the number of forward and backward background events, $N_F^{\text{bkgd}}$ and $N_B^{\text{bkgd}}$, in Eq. (3).

E. Jet energy mismeasurement

Although we correct the energy of each jet for known effects [19], the energy of a measured jet does not exactly match the energy of the corresponding particle-level jet [22]. Mismeasurement of jet energies principally affects the measurement of the dijet mass. Since we use wide mass bins, this effect is small. We estimate the effect using MC samples of dijet events produced with the PYTHIA event generator at leading order in the strong coupling constant, using the CTEQ5L parton distribution functions. We select events that have a pair of bottom quarks, whether produced directly or in the parton shower, to produce the signal model.

This signal model allows the determination of a matrix relating the measured dijet mass to particle-level dijet mass. The matrix is evaluated from MC, and depends on $|\Delta Q|$, but it is shown summed over bins of $|\Delta Q|$. 

![FIG. 3. Estimated asymmetry of the backgrounds in each $m_{b\bar{b}}$ bin with tag type LL, as a function of $|\Delta Q|$.

![FIG. 4. Smearing matrix relating measured dijet mass to particle-level dijet mass. The matrix is evaluated from MC, and depends on $|\Delta Q|$, but it is shown summed over bins of $|\Delta Q|$.

| $M_{bb}$ | $|\Delta Q|$ | $bkgd$ | $F$ |
|---------|-------------|-------|------|
| [0,150] | 0.00 | 0.07 | 0.00 |
| [150,225] | 0.81 | 0.06 | 0.00 |
| [225,325] | 0.12 | 0.80 | 0.05 |
| [325,\infty] | 0.00 | 0.14 | 0.95 |

The matrix is evaluated from MC, and depends on $|\Delta Q|$, but it is shown summed over bins of $|\Delta Q|$. 

| $M_{bb}$ | $|\Delta Q|$ | $bkgd$ | $F$ |
|---------|-------------|-------|------|
| [0,150] | 0.00 | 0.07 | 0.00 |
| [150,225] | 0.81 | 0.06 | 0.00 |
| [225,325] | 0.12 | 0.80 | 0.05 |
| [325,\infty] | 0.00 | 0.14 | 0.95 |
FIG. 5. Detector-level dijet mass spectra in each of the three $m_{bb}$ subsamples, showing the spectrum resulting from the non-$b\bar{b}$ component only as well as from the sum of the $b\bar{b}$ and non-$b\bar{b}$ components (“Model $\pm 1\sigma$”). The agreement between data and simulation is good, suggesting that the particle-level dijet mass in MC simulation and the smearing matrix are also correct within the uncertainties. The non-$b\bar{b}$ component is not included in the smearing matrix.

### TABLE II. Numbers of events observed in the various subsamples.

| $m_{bb}$ range | $\Delta Q$ | tag type | $N(\Delta y > 0)$ | $N(\Delta y < 0)$ |
|----------------|-----------|----------|------------------|------------------|
| [150, 225]     | 0.25      | LL       | 483              | 472              |
|                |           | HH       | 465              | 469              |
|                | 0.5       | LL       | 734              | 699              |
|                |           | HH       | 1223             | 1240             |
|                | 0.75      | LL       | 460              | 474              |
|                |           | HH       | 854              | 882              |
|                | 1.0       | LL       | 97               | 81               |
|                |           | HH       | 130              | 141              |
|                |           |          | 101              | 106              |
| [225, 325]     | 0.25      | LL       | 1014             | 984              |
|                |           | HH       | 1419             | 1520             |
|                | 0.5       | LL       | 1465             | 1499             |
|                |           | HH       | 2197             | 2230             |
|                | 0.75      | LL       | 915              | 979              |
|                |           | HH       | 1455             | 1558             |
|                | 1.0       | LL       | 712              | 735              |
|                |           | HH       | 140              | 160              |
|                |           |          | 283              | 256              |
|                |           |          | 149              | 140              |
| [325, 1960]    | 0.25      | LL       | 1565             | 1636             |
|                |           | HH       | 2214             | 2401             |
|                | 0.5       | LL       | 965              | 937              |
|                |           | HH       | 3196             | 3254             |
|                | 0.75      | LL       | 1298             | 1318             |
|                |           | HH       | 2282             | 2512             |
|                | 1.0       | LL       | 958              | 963              |
|                |           | HH       | 189              | 191              |

**F. Selection effects**

We also estimate the effect of the potentially asymmetric acceptance of the detector and analysis selection. We apply the analysis event selection to the simulated events to estimate the fractions of forward ($\epsilon_F$) and backward ($\epsilon_B$) events that pass the analysis selection. We perform this estimation in each bin of particle-level dijet mass. Only the ratio of the forward to the backward acceptance affects the analysis, so we compute the ratio $R = \epsilon_F / \epsilon_B$. We also estimate the uncertainty on this ratio by varying the renormalization and factorization scales used in the signal model by a factor of two and the jet-energy scale within its uncertainty.
G. Extraction of the particle-level asymmetry

To infer the asymmetry at the particle level, we construct a Bayesian model to describe the data. This model combines all effects discussed in the preceding sections, propagates the uncertainties, and allows the data to constrain the uncertainties when possible. The parameters of the model and their assumed prior probability distributions are as follows:

1. $f_{MQT}$ is the $b\bar{b}$ fraction in each bin of detector-level dijet mass $M$, charge difference $Q$, and tag-quality $T$. The prior probability distribution for $f$ is a normal distribution centered at the calibrated value, with a width equal to the residual uncertainty from the calibration.

2. $P_{MQ}$ is the correct charge probability described in Eq. (4). The priors for $p_{0.5}$ and $p_{0.25}$ are normal distributions with mean and uncertainty taken from the calibration.

3. $F_{MQT}^{bkgd}$ and $B_{MQT}^{bkgd}$ are the rates of forward and backward events in the background-dominated sideband. From these, we calculate the background asymmetry, which is assumed to be consistent with the asymmetry of the background in the signal region. The prior for each of these is the gamma distribution.

4. $J$ is the shift in the jet energy scale. We coherently shift all of the jet energies in every MC event by $J$ times the jet-energy uncertainty [19]. The prior is a normal distribution with a mean of zero and standard deviation of one.

5. $S_{M' MQ}(J)$ is a matrix describing the contribution of events with particle-level dijet mass $M'$ to the various bins of measured dijet mass and charge difference. The matrix is a function of $J$. The prior is taken from the rate and uncertainty in simulation, and the matrix is normalized so that $\sum_{M'} S_{M' MQ}(J) = 1$.

6. $\sigma_{MQT}$ is the rate of events in each bin of detector-level $b\bar{b}$ mass, charge difference, and tag quality. This parameter has a uniform prior over the non-negative range. This parameter is necessary because the simulation does not accurately predict the overall event rate as a function of mass, charge difference, and tag quality.

7. $R_{M'}$ is the ratio of the forward to the backward acceptance. The prior is a normal distribution with mean and width taken from the calibration described in Sec. III F.

8. $A_{M'}$ is the $b\bar{b}$ asymmetry in bins of particle-level mass. This is the parameter we wish to measure. We use a uniform prior from $[-1, 1]$.

9. $A_{M'}^{acc}$ is the $b\bar{b}$ asymmetry after acceptance and selection effects. It is a function of $A_{M'}$ and $R_{M'}$,

$$A_{M'}^{acc} = \frac{R_{M'}(1 + A_{M'}) - (1 - A_{M'})}{R_{M'}(1 + A_{M'}) + (1 - A_{M'})}.$$

We compute the rate $\theta$ of forward and backward events expected in data, and compare this rate to that observed in data (see Table II) via a Poisson likelihood, given an observation of $k$ events, $L(\theta|k) = \theta^k e^{-\theta}/k!$. The rates are

$$\theta_{MQT}^{Forward} = \left[ f_{MQT} \sum_{M'} 1 + \frac{A_{M'}^{acc}(2P_{MQ} - 1)}{2} S_{M' MQ}(J) \right] + (1 - f_{MQT}) B_{MQT}^{bkgd} \times \sigma_{MQT}, \quad (9)$$

and

$$\theta_{MQT}^{Backward} = \left[ f_{MQT} \sum_{M'} 1 - \frac{A_{M'}^{acc}(2P_{MQ} - 1)}{2} S_{M' MQ}(J) \right] + (1 - f_{MQT}) B_{MQT}^{bkgd} \times \sigma_{MQT}. \quad (10)$$

The prior probability densities described above, together with this likelihood, fully specify the posterior probability density for the parameters. To estimate this posterior density, we employ Markov-chain Monte Carlo sampling [24]. This technique provides us with samples from the posterior probability distribution over the parameter space. We marginalize the nuisance parameters and obtain the posterior density for $A_{M'}$, the asymmetry in each bin of particle-level $b\bar{b}$ mass. The marginal distributions are shown in Fig. 6, and the marginalized samples are provided in the supplemental material [25].

IV. RESULTS AND CONCLUSIONS

To characterize the posterior and describe the measurement, we find the highest probability-density credible intervals at 68% and 95% credibility for $A_{M'}$ in each particle-level mass bin. The posterior densities, along with the intervals describing them, are shown in Fig. 6. The red vertical bands (with solid lines) represent the theoretical predictions from the SM [11], while the blue and magenta bands (with dashed and dotted lines, respectively) represent the predictions from two axigluon models (see Table I). Although the measurement is performed at the particle level, the predictions are at the parton level and do not include the effects of hadronization. Because hadronization effects are not expected to be large, we do not hesitate to interpret the results.

The measured asymmetries, summarized in Fig. 7, are $-0.6^{+4.2}_{-5.2}$, $-7.4^{+8.9}_{-7.4}$, and $6.1^{+13.4}_{-11.3}$% in the low, middle,
FIG. 6. Marginal posterior probability distribution of asymmetry in each bin of particle-level $b\bar{b}$ mass. The inner and outer bands represent the 68% and 95% credible intervals, respectively.

FIG. 7. Values of maximum a posteriori signal asymmetry as a function of $b\bar{b}$ mass. The error bars represent the 68% credible intervals.

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean World Class University Program, the National Research Foundation of Korea; the Science and Technology Facilities Council and the Royal Society, United Kingdom; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; the Academy of Finland; the Australian Research Council (ARC); and the EU community Marie Curie Fellowship Contract No. 302103.
polar and azimuthal angles, respectively, and pseudorapidity and rapidity defined by:
\[ \eta = -\ln \tan(\theta/2) \]
and
\[ y = (1/2) \ln(E + p_z/E - p_z), \]
respectively. The transverse momentum \( p_T \) (transverse energy \( E_T \)) is defined to be
\[ p \sin \theta \ (E \sin \theta). \]

[11] B. Grinstein and C. W. Murphy, Phys. Rev. Lett. 111, 062003 (2013); C. W. Murphy, (2015), arXiv:1504.02493 [hep-ph].

[12] T. A. Aaltonen et al. (CDF Collaboration), Phys. Rev. D 89, 072005 (2014); S. Drell and T.-M. Yan, Phys. Rev. Lett. 25, 316 (1970).

[13] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. D 88, 072003 (2013).

[14] R. Aaij et al. (LHCb collaboration), Phys. Rev. Lett. 113, 082003 (2014).

[15] The forward-backward asymmetry must vanish in a symmetric proton-proton initial state.

[16] V. M. Abazov et al. (D0 Collaboration), Phys. Rev. Lett. 114, 051803 (2015).

[17] D. E. Acosta et al. (CDF Collaboration), Phys. Rev. D 71, 032001 (2005).

[18] We use a fixed-size cone algorithm to identify jets with a cone radius \( \Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} = 0.7 \).

[19] A. Bhatti et al., Nucl. Instrum. Methods A 566, 375 (2006).

[20] D. Acosta et al. (CDF Collaboration), Phys. Rev. D 71, 052003 (2005).

[21] T. Aaltonen et al., Phys. Rev. D 88, 032003 (2013).

[22] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. D 78, 052006 (2008).

[23] T. Aaltonen et al. (CDF Collaboration), Phys. Rev. D 85, 032005 (2012).

[24] A. Patil, D. Huard, and C. J. Fonnesbeck, Journal of Statistical Software 35, 1 (2010).

[25] See supplemental material at [URL will be provided by publisher].

[26] Ž. Šidák, Journal of the American Statistical Association 62, 626 (1967).