First observation of the decays $\bar{B}_s^0 \rightarrow D^+_s K^- \pi^+ \pi^-$ and $\bar{B}_s^0 \rightarrow D_{s1}(2536)^+ \pi^-$

LHCb Collaboration ; Anderson, J ; Bernet, R ; Büchler-Germann, A ; Bursche, A ; Chiapolini, N ; De Cian, M ; Elsasser, C ; Müller, K ; Salzmann, C ; Serra, N ; Steinkamp, O ; Straumann, U ; Tobin, M ; Vollhardt, A ; et al

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-75372
Journal Article

Originally published at:
LHCb Collaboration; Anderson, J; Bernet, R; Büchler-Germann, A; Bursche, A; Chiapolini, N; De Cian, M; Elsasser, C; Müller, K; Salzmann, C; Serra, N; Steinkamp, O; Straumann, U; Tobin, M; Vollhardt, A; et al (2012). First observation of the decays $\bar{B}_s^0 \rightarrow D^+_s K^- \pi^+ \pi^-$ and $\bar{B}_s^0 \rightarrow D_{s1}(2536)^+ \pi^-$. Physical Review D (Particles, Fields, Gravitation and Cosmology), 86:112005.
First observation of the decays $\bar{B}_s^0 \to D_s^+ K^- \pi^+ \pi^-$ and $\bar{B}_s^0 \to D_s(2536)^+ \pi^-$

R. Aaij,38,4 C. Abellán Beteta,33,p A. Adamez,11 B. Adeva,34 M. Adinolfi,43 C. Adrover,6 A. Affolder,49 Z. Ajaltouni,5 J. Albrecht,35 F. Alessio,35 M. Alexander,48 S. Ali,38 G. Alkhazov,27 P. Alvarez Cartelle,34 A. A. Alves, Jr.,22 S. Amato,2 Y. Amhis,36 L. Anderlini,17,g J. Anderson,37 R. B. Appleby,51 O. Aquines Gutierrez,10 F. Archilli,18,35 A. Artamonov,32 M. Artuso,5 E. Aslanyan,6 G. Auriemma,22,n S. Bachmann,11 J. J. Back,45 C. Baesso,4 W. Baldini,16 R. J. Barlow,51 C. Barschel,35 S. Barsuk,7 W. Barter,44 A. Bates,48 Th. Bauer,38 A. Bay,36 J. Beddow,35 I. Bediaga,1 S. Belogurov,28 K. Belous,32 I. Belyaev,25 E. Ben-Haim,9 M. Benayoun,8 G. Bencivenni,18 S. Benson,47 J. Benton,43 A. Berezhnuy,29 R. Bernet,37 M.-O. Bettler,44 M. van Beuzekom,38 A. Bien,11 S. Bifani,12 T. Bird,51 A. Bizzeti,173 P. M. Björnstad,51 T. Blake,35 F. Blanc,36 C. Blanks,50 J. Blouw,11 S. Blusk,33 A. Bobrov,31 V. Bocci,22 A. Bondar,31 N. Bondar,27 W. Bonivento,15 S. Borghi,48,51 A. Borgia,53 J. T. V. Bowcock,49 C. Bozzi,16 T. Brambach,9 J. van den Brand,39 J. Bressieux,36 D. Brett,51 M. Britsch,10 T. Britton,53 N. H. Brook,43 H. Brown,49 A. Büchler-Germann,37 I. Burducea,26 A. Bursche,37 J. Buytaert,35 S. Cadeddu,15 O. Callot,7 M. Calvi,20,35 M. Calvo Gomez,33,o A. Camboni,33 P. Campana,18,35 A. Carbone,14,4 G. Carbone,21,1 R. Cardinale,19,a A. Cardini,15 H. Carranza-Mejía,47 L. Carson,50 K. Carvalho Akiba,2 G. Casse,39 M. Cattaneo,35 Ch. Cauet,9 M. Charles,52 Ph. Charpentier,35 P. Chen,36 N. Chiapolini,37 M. Chrzaszcz,23 K. Ciba,35 X. Cid Vidal,34 G. Ciezarek,50 P. E. L. Clarke,47 M. Clemencic,35 H. V. Cliff,44 J. Closier,35 C. Coka,26 V. Coco,38 J. Cogan,6 E. Cogneras,5 P. Collins,35 A. Comerma-Montells,33 A. Contu,52,15 A. Cook,43 M. Coombes,43 S. C. Haines,44 S. Hall,50 T. Hampson,43 S. Hansmann-Menzemer,11 N. Harnew,52 S. T. Harnew,43 A. Bursche,37 J. Buytaert,35 S. Cadeddu,15 O. Callot,7 M. Calvi,20,35 M. Calvo Gomez,33,o A. Camboni,33 P. Campana,18,35 A. Carbone,14,4 G. Carbone,21,1 R. Cardinale,19,a A. Cardini,15 H. Carranza-Mejía,47 L. Carson,50 K. Carvalho Akiba,2 G. Casse,39 M. Cattaneo,35 Ch. Cauet,9 M. Charles,52 Ph. Charpentier,35 P. Chen,36 N. Chiapolini,37 M. Chrzaszcz,23 K. Ciba,35 X. Cid Vidal,34 G. Ciezarek,50 P. E. L. Clarke,47 M. Clemencic,35 H. V. Cliff,44 J. Closier,35 C. Coka,26 V. Coco,38 J. Cogan,6 E. Cogneras,5 P. Collins,35 A. Comerma-Montells,33 A. Contu,52,15 A. Cook,43 M. Coombes,43 S. C. Haines,44 S. Hall,50 T. Hampson,43 S. Hansmann-Menzemer,11 N. Harnew,52 S. T. Harnew,43
The first observation of the decays $B^0 \to D^+_s K^- \pi^+ \pi^-$ and $\bar{B}^0 \to D^+_s K^- \pi^+ \pi^-$ are reported using an integrated luminosity of 1.0 fb$^{-1}$ recorded by the LHCb experiment. The branching fractions, normalized with respect to $B^0 \to D^+_s K^- \pi^+ \pi^-$ and $\bar{B}^0 \to D^+_s K^- \pi^+ \pi^-$, respectively, are measured to be $\frac{\mathcal{B}(B^0 \to D^+_s K^- \pi^+ \pi^-)}{\mathcal{B}(\bar{B}^0 \to D^+_s K^- \pi^+ \pi^-)} = (5.2 \pm 0.5 \pm 0.3) \times 10^{-2}$ and $\frac{\mathcal{B}(\bar{B}^0 \to D^+_s K^- \pi^+ \pi^-)}{\mathcal{B}(B^0 \to D^+_s K^- \pi^+ \pi^-)} = 0.54 \pm 0.07 \pm 0.07$, where the first

---

$^a$Full author list given at end of the article.
$^b$P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia.
$^{12}$Universitá di Bari, Bari, Italy.
$^{13}$Universitá di Bologna, Bologna, Italy.
$^{14}$Universitá di Cagliari, Cagliari, Italy.
$^{15}$Universitá di Ferrara, Ferrara, Italy.
$^{16}$Universitá di Firenze, Firenze, Italy.
$^{17}$Universitá di Urbino, Urbino, Italy.
$^{18}$Universitá di Modena e Reggio Emilia, Modena, Italy.
$^{19}$Universitá di Genova, Genova, Italy.
$^{20}$Universitá di Milano Bicocca, Milano, Italy.
$^{21}$Universitá di Roma Tor Vergata, Roma, Italy.
$^{22}$Universitá di Roma La Sapienza, Roma, Italy.
$^{23}$Universitá della Basilicata, Potenza, Italy.
$^{24}$LIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain.
$^{25}$Hanoi University of Science, Hanoi, Viet Nam.
$^{26}$Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.
uncertainty is statistical and the second is systematic. The $\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$ decay is of particular interest as it can be used to measure the weak phase $\gamma$. First observation of the $\bar{B}^0 \rightarrow D_s(2536)^+ \pi^-$, $D_s^+ \rightarrow D_s^+ K^- \pi^+ \pi^-$ decay is also presented, and its branching fraction relative to $\bar{B}^0 \rightarrow D_s^+ \pi^- \pi^+ \pi^-$ is found to be $\frac{B(\bar{B}^0 \rightarrow D_s(2536)^+ \pi^-)B(D_s^+ \rightarrow D_s^+ K^- \pi^+ \pi^-)}{B(\bar{B}^0 \rightarrow D_s^+ \pi^- \pi^+ \pi^-)} = (4.0 \pm 1.0 \pm 0.4) \times 10^{-3}$.

DOI: 10.1103/PhysRevD.86.112005 PACS numbers: 13.25.Hw, 13.20.He

I. INTRODUCTION

In the Standard Model (SM), the amplitudes associated with flavor-changing processes depend on four Cabibbo-Kobayashi-Maskawa (CKM) [1,2] matrix parameters. Contributions from physics beyond the Standard Model (BSM) add coherently to these amplitudes, leading to potential deviations in rates and $CP$-violating asymmetries when compared to the SM contributions alone. Since the SM does not predict the CKM parameters, it is important to make precise measurements of their values in processes that are expected to be insensitive to BSM contributions. Their values then provide a benchmark to which BSM-sensitive measurements can be compared.

The least well-determined of the CKM parameters is the weak phase $\gamma \equiv \arg(-V_{ub}^* V_{cd})$, which, through direct measurements, is known to a precision of $\sim 10^0$–$12^0$ [3,4]. It may be probed using time-independent rates of decays such as $B^- \rightarrow DK^- \pi^+ \pi^-$ [5–7] or by analyzing the time-dependent decay rates of processes such as $\bar{B}^0 \rightarrow D_s^+ K^-$ [8–11]. Sensitivity to the weak phase $\gamma$ results from the interference between $b \rightarrow c$ and $b \rightarrow u$ transitions, as indicated in Figs. 1(a)–1(c). Such measurements may be extended to multibody decay modes, such as $B^- \rightarrow DK^- \pi^+ \pi^-$ [12] for a time-independent measurement, or $\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$ in the case of a time-dependent analysis.

The $\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$ decay, while having the same final state as $\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$, receives contributions not only from the $W$-exchange process [Fig. 1(d)], but also from $b \rightarrow c$ transitions in association with the production of an extra $s \bar{s}$ pair [Figs. 1(e) and 1(f)]. The decay may also proceed through mixing followed by a $b \rightarrow u$, $W$-exchange process (not shown). However, this amplitude is Cabibbo-, helicity- and color-suppressed and is therefore negligible compared to the $b \rightarrow c$ amplitude.

This paper reports the first observation of $\bar{B}_s^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$ and $\bar{B}_s^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$ and measurements of their branching fractions relative to $\bar{B}^0 \rightarrow D_s^+ \pi^- \pi^+ \pi^-$ and $\bar{B}_s^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$, respectively. The data sample is based on an integrated luminosity of 1.0 fb$^{-1}$ of $p p$ collisions at $\sqrt{s} = 7$ TeV, collected by the LHCb experiment.

The same data sample is also used to observe the $\bar{B}_s^0 \rightarrow D_s(2536)^+ \pi^-$, $D_s^+ \rightarrow D_s^+ \pi^- \pi^+ \pi^-$ decay for the first time and measure its branching fraction relative to $\bar{B}^0 \rightarrow D_s^+ \pi^- \pi^+ \pi^-$. The inclusion of charge-conjugated modes is implied throughout this paper.

II. DETECTOR AND SIMULATION

The LHCb detector [13] is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed for the study of particles containing $b$ or $c$ quarks. The detector includes a high precision tracking system consisting of a silicon-strip vertex detector surrounding the interaction region, a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about 4 Tm, and three stations of silicon-strip detectors and straw drift-tubes placed downstream. The combined tracking system has a momentum resolution $(\Delta p/p)$ that varies from 0.4% at 5 GeV/$c$ to 0.6% at 100 GeV/$c$, and an impact parameter (IP) resolution of 20 $\mu$m for tracks with high transverse momentum $(p_T)$. Charged hadrons are identified using two ring-imaging Cherenkov detectors. Photon, electron and hadron candidates are identified by a calorimeter system consisting of scintillating-pad and pre-shower detectors, an electromagnetic calorimeter and a hadronic calorimeter. Muons are identified by a system

![Diagram](image.png)

FIG. 1 (color online). Diagrams contributing to the $\bar{B}_s^0$, $\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$ (a–c) and $\bar{B}_s^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$ (d–f) decays, as described in the text. In (a)–(d), the additional $(\pi^+ \pi^-)$ indicates that the $K^- \pi^+ \pi^-$ may be produced either through an excited strange kaon resonance decay, or through fragmentation.
composed of alternating layers of iron and multiwire proportional chambers.

The trigger consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which applies a full event reconstruction. The software trigger requires a two-, three- or four-track secondary vertex with a high \( p_T \) sum of the tracks and a significant displacement from the primary \( pp \) interaction vertices (PVs). At least one track should have \( p_T > 1.7 \text{ GeV}/c \), an IP \( \chi^2 \) greater than 16 with respect to all PVs, and a track fit \( \chi^2/\text{ndf} < 2 \), where ndf is the number of degrees of freedom. The IP \( \chi^2 \) is defined as the difference between the \( \chi^2 \) of the PV reconstructed with and without the considered particle. A multivariate algorithm is used for the identification of secondary vertices [14].

For the simulation, \( pp \) collisions are generated using PYTHIA 6.4 [15] with a specific LHCb configuration [16]. Decays of hadronic particles are described by EVTGEN [17] in which final state radiation is generated using PHOTOS. At least one track should have and a significant displacement from the primary \( \text{PV} \). Tracks used to form the \( IP \) are required to be identified as either a pion or a kaon using information from the ring-imaging Cherenkov detectors, have \( p_T \) in excess of 100 MeV/c and be significantly detached from any reconstructed PV in the event.

Signal \( B^0 \) decay candidates are formed by pairing a \( D^+_s \to K^-\pi^-\pi^- \) candidate with either a \( \pi^-\pi^-\pi^- \) (hereafter referred to as \( X_d \)) or a \( K^-\pi^-\pi^- \) combination (hereafter referred to as \( X_s \)). Tracks used to form the \( D^+_s \) and \( X_{ds} \) are required to be identified as either a pion or a kaon.

Signal \( D^+_s \) candidates are required to have good vertex fit quality, be significantly displaced from the nearest PV and have invariant mass, \( M(K^-\pi^-\pi^-) \), within 20 MeV/c\(^2\) of the \( D^+_s \) mass [21]. To suppress combinatorial and charmless backgrounds, only those \( D^+_s \) candidates that are consistent with decaying through either the \( \phi \) (\( M(K^-\pi^-) < 1040 \text{ MeV}/c^2 \)) or \( \bar{K}^{*0} \) (\( |M(K^-\pi^-) - m_{K^{*0}}| < 75 \text{ MeV}/c^2 \)) resonances are used (here, \( m_{K^{*0}} \) is the \( K^{*0} \) mass [21]). The remaining charmless background yields are determined using the \( D^+_s \) mass sidebands. For about 20% of candidates, when the \( K^- \) is assumed to be a \( \pi^- \), the corresponding \( K^-\pi^-\pi^- \) invariant mass is consistent with the \( D^0 \) mass. To suppress cross feed from \( B^0 \to D^+X \) decays, a tighter particle identification (PID) requirement is applied to the \( K^- \) in the \( D^+_s \to K^-\pi^-\pi^- \) candidates when \( |M(K^-\pi^-\pi^-) - m_{D^0}| < 20 \text{ MeV}/c^2 \) (\( m_{D^0} \) is the \( D^0 \) mass [21]). Similarly, if the invariant mass of the particles forming the \( D^+_s \) candidate, after replacing the \( K^- \) mass with the proton mass, falls within 15 MeV/c\(^2\) of the \( \Lambda_c^+ \) mass, tighter PID selection is applied. The sizes of these mass windows are about 2.5 times the invariant mass resolution and are sufficient to render these cross-feed backgrounds negligible.

Candidates \( X_d \) and \( X_s \) are formed from \( \pi^-\pi^-\pi^- \) or \( K^-\pi^-\pi^- \) combinations, where all invariant mass values up to 3 GeV/c\(^2\) are accepted. To reduce the level of combinatorial background, we demand that the \( X_{ds} \) vertex is displaced from the nearest PV by more than 100 \( \mu \text{m} \) in the direction transverse to the beam axis and that at least two of the daughter tracks have \( p_T > 300 \text{ MeV}/c \).

Backgrounds to the \( X_{ds} \) are suppressed by applying more stringent PID requirements to the \( K^- \) and \( \pi^- \) in \( X_s \). The PID requirements have an efficiency of about 65% for selecting \( X_s \), while rejecting about 97% of the favored three-pion background. To suppress peaking backgrounds from \( B^0 \to D^+_s D^- \) decays, where \( D^+_s \to \pi^+\pi^-\pi^- \) or \( K^-\pi^-\pi^- \), it is required that \( M(X_{ds}) \) is more than 20 MeV/c\(^2\) away from the \( D^+_s \) mass.

Signal \( B \) meson candidates are then formed by combining a \( D^+_s \) with either an \( X_d \) or \( X_s \). The reconstructed \( B \) candidate is required to be well separated from the nearest PV with a decay time larger than 0.2 ps and to have a good quality vertex fit. To suppress remaining charmless backgrounds, which appear primarily in \( B^0 \to D^+_s K^-\pi^-\pi^- \), the vertex separation \( \chi^2 \) between the \( D^+_s \) and \( B \) decay vertices is required to be greater than 9. Candidates passing all selection requirements are refit with both \( D^+_s \) mass and vertex constraints to improve the mass resolution [22].

To further suppress combinatorial background, a boosted decision tree (BDT) selection [23] with the AdaBoost algorithm [24] is employed. The BDT is trained using simulated \( B^0 \to D^+_s K^-\pi^-\pi^- \) decays for the signal distributions, and the high \( B \) mass sideband in data are used to model the backgrounds. The following 13 variables are used:

(i) \( B \) candidate: IP \( \chi^2 \), vertex separation \( \chi^2 \), vertex fit \( \chi^2 \), and \( p_T \);
(ii) \( D^+_s \) candidate: Flight distance significance from \( B \) vertex;
(iii) \( X_{ds} \) candidate: IP \( \chi^2 \), maximum of the distances of closest approach between any pair of tracks in the decay;
(iv) \( X_{ds} \) daughters: min(IP \( \chi^2 \)), max(IP \( \chi^2 \)), \( \min(p_T) \); and
(v) \( D^+_s \) daughters: min(IP \( \chi^2 \)), max(IP \( \chi^2 \)), \( \min(p_T) \),

where min and max denote the minimum and maximum of the indicated values amongst the daughter particles. The flight distance significance is the separation between the \( D^+_s \) and \( B \) vertices, normalized by the uncertainty. The training produces a single variable, \( x \), that provides discrimination between signal decays and background contributions. The cut value is chosen by optimizing \( S(x_{cut})/\sqrt{S(x_{cut}) + B(x_{cut})} \), where \( S(x_{cut}) \) and \( B(x_{cut}) \) are the expected signal and background yields, respectively.
after requiring $x > x_{\text{cut}}$. At the optimal point, a signal efficiency of $\sim 90\%$ is expected while rejecting about 85\% of the combinatorial background (after the previously discussed selections are applied). After all selections, about 3\% of events have more than one signal candidate in both data and simulation. All candidates are kept for further analysis.

IV. FITS TO DATA

The $B_s^0 \to D_s^+ \pi^- \pi^+ \pi^-$ and $B_s^0 \to D_s^+ K^- \pi^+ \pi^-$ invariant mass spectra are each modeled by the sum of a signal and several background components. The signal shapes are obtained from simulation and are each described by the sum of a crystal ball (CB) [25] shape and a Gaussian function. The CB shape parameter that describes the tail toward low mass is fixed based on simulated decays. A common, freely varying scale factor multiplies the width parameters in the CB and Gaussian functions to account for slightly larger resolution in data than in simulation. For the $B_s^0 \to D_s^+ K^- \pi^+ \pi^-$ mass fit, the difference between the mean $\bar{B}_s^0$ and $B_s^0$ masses is fixed to 87.35 MeV/$c^2$ [21].

Several nonsignal $b$-hadron decays produce broad peaking structures in the $D_s^+ \pi^- \pi^+ \pi^-$ and $D_s^+ K^- \pi^+ \pi^-$ invariant mass spectra. For $B_s^0 \to D_s^+ \pi^- \pi^+ \pi^-$, the only significant source of peaking background is from $B_s^0 \to D_s^+ \pi^- \pi^- \pi^+$, where the photon or $\pi^0$ from the $D_s^+$ decay is not included in the reconstructed decay. Since the full decay amplitude for $B_s^0 \to D_s^+ \pi^- \pi^- \pi^+$ is not known, the simulation may not adequately model the decay. Simulation is therefore used to provide an estimate for the shape, but the parameters are allowed to vary within one standard deviation about the fitted values.

For $B_s^0 \to D_s^+ K^- \pi^+ \pi^-$, backgrounds from $B_s^0 \to D_s^+ K^- \pi^- \pi^- \pi^-$ and from misidentified $B_s^0 \to D_s^+ \pi^- \pi^- \pi^+$ and $B_s^0 \to D_s^+ \pi^- \pi^- \pi^- \pi^-$ decays are considered. The $B_s^0 \to D_s^+ K^- \pi^+ \pi^-$ shape is fixed to be the same as that obtained for the $B_s^0 \to D_s^+ \pi^- \pi^- \pi^-$ component in the $B_s^0 \to D_s^+ \pi^- \pi^- \pi^- \pi^-$ mass fit. This same shape is assumed for both $B_s^0$ and $\bar{B}_s^0$, where for the former, a shift by the $\bar{B}_s^0 - B_s^0$ mass difference is included. For the $B_s^0 \to D_s^+ \pi^- \pi^+ \pi^-$ and $B_s^0 \to D_s^+ \pi^- \pi^- \pi^- \pi^-$ cross feed, simulated decays and kaon misidentification rates taken from $D^*$ calibration data are used to obtain their expected yields and invariant mass shapes. The cross-feed contribution is about 3\% of the $B_s^0 \to D_s^+ \pi^- \pi^- \pi^-$ and $B_s^0 \to D_s^+ \pi^- \pi^- \pi^- \pi^-$ yields; the corresponding cross-feed yields are fixed in the $B_s^0 \to D_s^+ K^- \pi^+ \pi^-$ fit. The shape is obtained by parametrizing the invariant mass spectrum obtained from the simulation after replacing the appropriate $\pi^-$ mass in $X_d$ with the kaon mass. The combinatorial background is described by an exponential function whose slope is allowed to vary independently for both mass fits.

Figure 2 shows the invariant mass distribution for $B_s^0 \to D_s^+ \pi^- \pi^- \pi^-$ candidates passing all selection criteria. The fitted number of $B_s^0 \to D_s^+ \pi^- \pi^- \pi^-$ signal events is $5683 \pm 83$. While it is expected that most of the low mass background emanates from $B_s^0 \to D_s^+ \pi^- \pi^- \pi^-$, contributions from other sources such as $B_s^0 \to D_s^+ \pi^- \pi^- \pi^- \pi^0$ are also possibly absorbed into this background component. Figure 3 shows the invariant mass distribution for $B_s^0 \to D_s^+ K^- \pi^+ \pi^-$ candidates. The fitted signal yields are $402 \pm 33$ $B_s^0 \to D_s^+ K^- \pi^+ \pi^-$ and $216 \pm 21$ $B_s^0 \to D_s^+ K^- \pi^- \pi^-$ events.

![Figure 2](image-url)  
**FIG. 2.** Invariant mass distribution for $B_s^0 \to D_s^+ \pi^- \pi^- \pi^-$ candidates passing all selection criteria. The fitted number of $B_s^0 \to D_s^+ \pi^- \pi^- \pi^-$ signal events is $5683 \pm 83$. While it is expected that most of the low mass background emanates from $B_s^0 \to D_s^+ \pi^- \pi^- \pi^-$, contributions from other sources such as $B_s^0 \to D_s^+ \pi^- \pi^- \pi^- \pi^0$ are also possibly absorbed into this background component.

![Figure 3](image-url)  
**FIG. 3.** Invariant mass distribution for $B_s^0 \to D_s^+ K^- \pi^+ \pi^-$ candidates. The fitted signal (dashed lines) and background shapes (shaded/hatched regions) are shown, as described in the text.
TABLE I. Summary of event yields from data in the $D_s^+$ signal and sidebands regions and the background corrected yield. The signal and sideband regions require $D_s^+$ candidates to have invariant mass $|M(K^-K^+\pi^+) - m_{D_s^+}| < 20$ MeV/$c^2$ and $35 < |M(K^-K^+\pi^+) - m_{D_s^+}| < 55$ MeV/$c^2$, respectively, where $m_{D_s^+}$ is the $D_s^+$ mass [21].

| Decay | Signal Region | Sideband Region | Corrected Yield |
|-------|---------------|-----------------|-----------------|
| $B_0^0 \to D_s^+ \pi^-\pi^+\pi^-$ | 5683 ± 83 | 61 ± 16 | 5622 ± 85 |
| $B_0^0 \to D_s^+ K^-\pi^+\pi^-$ | 216 ± 21 | 0$^{+5}_{-0}$ | 216 ± 22 |
| $B^0 \to D^+_s K^-\pi^+\pi^-$ | 402 ± 33 | 9 ± 5 | 393 ± 33 |

The $D_s^+$ mass sidebands, defined to be from 35 to 55 MeV/$c^2$ on either side of the nominal $D_s^+$ mass, are used to estimate the residual charmless background that may contribute to the observed signals. The numbers of $B^0$ decays in the $D_s^+$ sidebands are $61 \pm 16$, $0^{+5}_{-0}$, and $9 \pm 5$ for the $B_0^0 \to D_s^+\pi^-\pi^+\pi^-$, $B_0^0 \to D_s^+K^-\pi^+\pi^-$, and $B^0 \to D_s^+K^-\pi^+\pi^-$ decays, respectively; they are subtracted from the observed signal yields to obtain the corrected number of signal decays. The yields in the signal and sideband regions are summarized in Table I.

V. MASS DISTRIBUTIONS OF $X_{d,s}$ AND TWO-BODY MASSES

In order to investigate the properties of these $B^0$($K^0$) decays, $s$Weights [26] obtained from the mass fits are used to determine the underlying $X_{d,s}$ invariant mass spectra as well as the two-body invariant masses amongst the three daughter particles. Figure 4 shows (a) the $\pi^-\pi^+\pi^-$ mass, (b) the smaller $\pi^-\pi^-$ mass and (c) the larger $\pi^-\pi^-$ mass in $B_0^0 \to D_s^+ \pi^-\pi^+\pi^-$ data and simulated decays. A prominent peak, consistent with the $a_1(1260)^- \to \pi^-\pi^+\pi^-$, is observed, along with structures consistent with the $\rho^0$ in the two-body masses. There appears to be an offset in the peak position of the $a_1(1260)^-$ between data and simulation. Since the mean and width of the $a_1(1260)^-$ resonance are not well known, and their values may even be process dependent, this level of agreement is reasonable. A number of other spectra have been compared between data and simulation, such as the $p_T$ spectra of the

FIG. 4 (color online). Invariant mass distributions for (a) $X_d$, (b) smaller $\pi^-\pi^-$ mass in $X_d$ and (c) the larger $\pi^-\pi^-$ mass in $X_d$, from $B^0 \to D_s^+ \pi^-\pi^+\pi^-$ decays using $s$Weights. The points are the data and the solid line is the simulation. The simulated distribution is normalized to have the same yield as the data.

FIG. 5 (color online). Invariant mass distributions for (a) $X_s$, (b) $\pi^+\pi^-$ in $X_s$ and (c) the $K^-\pi^+$ in $X_s$, from $B^0 \to D_s^+ K^-\pi^+\pi^-$ data using $s$Weights. The points are data and the solid line is the simulation. The simulated distribution is normalized to have the same yield as the data.
$D_s^+$, $X_d$ and the daughter particles, and excellent agreement is found.

Figure 5 shows the corresponding distributions for the $B_s^0 \rightarrow D_{sJ}^- \ K^- \pi^+ \pi^-$ decay. A peaked structure at low $K^- \pi^+ \pi^-$ mass, consistent with contributions from the lower-lying excited strange mesons, such as the $K_1(1270)^-$ and $K_1(1400)^-$, is observed. As many of these states decay through $\bar{K}^0$ and $\rho^0$ mesons, significant contributions from these resonances are observed in the $K^- \pi^+$ and $\pi^+ \pi^+$ invariant mass spectra, respectively. The simulation provides a reasonable description of the distributions in the data.

Figure 6 shows the same distributions for $\bar{B}_s^0 \rightarrow D_{sJ}^+ \ K^+ \pi^- \pi^-$. The $K^- \pi^+ \pi^-$ invariant mass is quite broad, with little indication of any narrow structures. There are indications of $K^{*0}$ and $\rho^0$ contributions in the $K^- \pi^+$ and $\pi^- \pi^+$ invariant mass spectra, respectively, but the contribution from resonances such as the $K_1(1270)^-$ or $K_1(1400)^-$ appear to be small or absent. In the $K^- \pi^+$ invariant mass spectrum, there may be an indication of a $K^{*0}(1430)^0$ contribution. The simulation, which models the $K^- \pi^+ \pi^-$ final state as 10% $K_1(1270)^-$, 10% $K_1(1400)^-$, 40% $\bar{K}^{*0}$ and 40% $K^- \rho^0$, provides a reasonable description of the data, which suggests that processes such as those in Figs. 1(e) and 1(f) constitute a large portion of the total width for this decay.

VI. FIRST OBSERVATION
OF $\bar{B}_s^0 \rightarrow D_{sJ}(2536)^+ \pi^-$

A search for excited $D_s^+$ states, such as $D_{sJ}^- \rightarrow D_{sJ}^- \ K^- \pi^+ \pi^+$, contributing to the $\bar{B}_s^0 \rightarrow D_{sJ}^+ \ K^+ \pi^- \pi^-$ final state is performed. Signal candidates within $\pm 40 \text{MeV}/c^2$ of the nominal $B_s^0$ mass are selected, and from them the invariant mass difference, $\Delta M = M(D_{sJ}^+ \pi^- \pi^+) - M(D_s^+)$ is formed, where both $\pi^- \pi^+$ combinations are included. The $\Delta M$ distribution for candidates in the $\bar{B}_s^0$ signal window is shown in Fig. 7. A peak corresponding to the $D_{sJ}(2536)^+$ is observed, whereas no significant structures are observed in the upper $B_s^0$ mass sideband ($5450$–$5590 \text{MeV}/c^2$). The distribution is fitted to the sum of a signal Breit-Wigner shape convolved with a Gaussian resolution function, and a second order polynomial to describe the background contribution. The Breit-Wigner width is set to $0.92 \text{MeV}/c^2$ [21], and the Gaussian resolution is fixed to $3.8 \text{MeV}/c^2$ based on simulation. A signal yield of $20.0 \pm 5.1$ signal events is observed at a mass difference of $565.1 \pm 1.0 \text{MeV}/c^2$, which is consistent with the known $D_{sJ}(2536)^+ - D_s^+$ mass difference of $566.63 \pm 0.35 \text{MeV}/c^2$ [21]. The significance of the signal is 5.9, obtained by fitting the invariant mass distribution with the mean mass difference fixed to $566.63 \text{MeV}/c^2$ [21], and computing $\sqrt{-2\ln(L_0/L_{\text{max}})}$. Here, $L_{\text{max}}$ and $L_0$ are the fit likelihoods with the signal yields left free and fixed to zero, respectively. Several
variations in the background shape were investigated, and in all cases the signal significance exceeded 5.5. This decay is therefore observed for the first time. To obtain the yield in the normalization mode \( (\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-) \), the signal function is integrated from 40 MeV/c\(^2\) below to 40 MeV/c\(^2\) above the nominal \( B^0 \) mass. A yield of 5505 ± 85 events is found in this restricted mass interval.

VII. SELECTION EFFICIENCIES

The ratios of branching fractions can be written as

\[
\frac{\mathcal{B}(\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-)}{\mathcal{B}(\bar{B}^0 \rightarrow D_s^+ \pi^- \pi^- \pi^-)} = \frac{Y(\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-)}{Y(\bar{B}^0 \rightarrow D_s^+ \pi^- \pi^- \pi^-)} \times \epsilon_{\text{rel}}
\]

(1)

and

\[
\frac{\mathcal{B}(\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-)}{\mathcal{B}(\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-)} = \frac{Y(\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-)}{Y(\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-)} \times \epsilon_{\text{rel}} \times f_s/f_d,
\]

(2)

where \( Y \) are the measured yields, \( \epsilon_{\text{rel}} = \epsilon(\bar{B}^0 \rightarrow D_s^+ \pi^- \pi^- \pi^-)/\epsilon(\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-) \) and \( \epsilon_{\text{rel}} = \epsilon(\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-)/\epsilon(\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-) \) are the relative selection efficiencies (including trigger), and \( f_s/f_d = 0.267 ± 0.021 \) [27] is the \( B^0 \) fragmentation fraction relative to \( B^0 \). The ratios of selection efficiencies are obtained from simulation, except for the PID requirements, which are obtained from a dedicated \( D^{*+} \) calibration sample, weighted to match the momentum spectrum of the particles that form \( X_d \) and \( X_s \). The selection efficiencies for each decay are given in Table II. The efficiency of the \( \bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^- \pi^- \) decay is about 35% larger than the values obtained in either the \( \bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^- \) or \( \bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^- \) decay; the efficiencies of the latter two are consistent with each other. The lower efficiency is due almost entirely to the tighter PID requirements on the \( K^- \) and \( \pi^+ \) in \( X_s \). Two additional multiplicative correction factors, also shown in Table II, are applied to the measured ratio of branching fractions in Eqs. (1) and (2). The first is a correction for the \( D_s^+ \) mass veto on \( M(X_{d,s}) \), and the second is due to the requirement that \( M(X_{d,s}) < 3 \text{ GeV/c}^2 \). The former, which represents a small correction, is estimated from the \( s\text{Weight-ed} \) distributions of \( M(X_{d,s}) \) shown previously. For the latter, the fraction of events with \( M(X_{d,s}) > 3 \text{ GeV/c}^2 \) is obtained from simulation and scaled by the ratio of yields in data relative to simulation for the mass region \( 2.6 < M(X_{d,s}) < 3.0 \text{ GeV/c}^2 \). A 50% uncertainty is assigned to the estimated correction. Based on the qualitative agreement between data and simulation in the \( M(X_{d,s}) \) distributions (see Sec. V) and the fact that the phase space approaches zero as \( M(X_{d,s}) \rightarrow 3.5 \text{ GeV/c}^2 \), this uncertainty is conservative. The relative efficiency between \( \bar{B}^0 \rightarrow D_s^+(2536)^+ \pi^- \), \( D_s^+ \rightarrow D_s^+ K^- \pi^+ \pi^- \) and \( \bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^- \) is estimated from simulation and is found to be \( 0.90 ± 0.05 \).

VIII. SYSTEMATIC UNCERTAINTIES

Several uncertainties contribute to the ratio of branching fractions. The sources and their values are listed in Table III. The largest uncertainty, which applies only to the ratio \( \mathcal{B}(\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-)/\mathcal{B}(\bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^-) \), is from the \( b \) hadronization fraction, \( f_s/f_d = 0.267 ± 0.021 \) [27], which is 7.9%. Another large uncertainty results from the required correction factor to account for the signal with \( M(X_{d,s}) > 3 \text{ GeV/c}^2 \). Those corrections are described in Sec. VII.

The selection efficiency depends slightly on the modeling of the \( X_{d,s} \) decay. The momentum spectra of the \( \bar{B}, D_s^+ \), \( X_d \), and the \( X_{d,s} \) daughters have been compared to simulation, and excellent agreement is found. The selection efficiency is consistent with being flat as a function of \( M(X_{d,s}) \) at the level of two standard deviations or less. To assess a potential systematic uncertainty due to a possible \( M(X_{d,s}) \)-dependent efficiency, the relative differences between the nominal selection efficiencies and the ones obtained by reweighting the measured efficiencies by the \( X_{d,s} \) mass spectra in data are computed. The relative deviations of 0.5%, 1.1%, and 1.2% for \( \bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^- \), \( \bar{B}^0 \rightarrow D_s^+ \pi^- \pi^- \pi^- \) and \( \bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^- \), respectively, are the assigned uncertainties. The systematic uncertainty on the BDT efficiency is determined by fitting the \( \bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^- \) mass distribution in data with and without the BDT requirement. The efficiency is found to agree with simulation to better than the 1% uncertainty assigned to this source. In total, the simulated efficiencies have uncertainties of 1.6 and 1.9% in the two ratios of branching fractions. The PID efficiency uncertainty is dominated by the usage of the \( D^{*+} \) calibration sample to determine

| Quantity | \( \bar{B}^0 \rightarrow D_s^+ \pi^- \pi^- \pi^- \) | \( \bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^- \) | \( \bar{B}^0 \rightarrow D_s^+ K^- \pi^+ \pi^- \) |
|----------|-----------------|-----------------|-----------------|
| Total \( \epsilon \) (10\(^{-4}\)) | 4.97 ± 0.08 | 3.67 ± 0.10 | 3.59 ± 0.10 |
| \( D_s^+ \) veto corrected | 1.013 ± 0.003 | 1.013 ± 0.003 | 1.017 ± 0.005 |
| \( M > 3 \text{ GeV/c}^2 \) corrected | 1.02 ± 0.01 | 1.04 ± 0.02 | 1.14 ± 0.07 |
The efficiencies of a given PID requirement [28]. This uncertainty is assessed by comparing the PID efficiencies obtained directly from simulated signal decays with the values obtained using a simulated $D^{+}$ calibration sample that is re-weighted to match the kinematics of the signal decay particles. Using this technique, an uncertainty of 2% each on the $B^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$ and $B^0 \rightarrow D_s^+ K^- \pi^- \pi^-$ PID efficiencies is obtained, which is 100% correlated, and a 1% uncertainty for $B^0 \rightarrow D_s^+ \pi^- \pi^- \pi^-$. The trigger is fully simulated, and given the identical number of tracks and the well-modeled $p_T$ spectra, the associated uncertainty cancels to first order. Based on previous studies [12], a 2% uncertainty is assigned.

The uncertainties in the signal yield determinations have contributions from both the background and signal modeling. The signal shape uncertainty was estimated by varying all the fixed signal shape parameters one at a time by one standard deviation, and adding the changes in yield in quadrature (0.5%). A double Gaussian signal shape model was also tried, and the difference was negligible. For the combinatorial background, the shape was modified from a single exponential to either the sum of two exponentials, or a linear function. For $B^0 \rightarrow D_s^+ \pi^- \pi^- \pi^-$, the difference in yield was 0.4%. For $B^0 \rightarrow D_s^+ K^- \pi^- \pi^-$, the maximum change was 4%, and for $B^0 \rightarrow D_s^+ K^- \pi^- \pi^-$, the maximum shift was 1%. In the $B^0 \rightarrow D_s^+ K^- \pi^- \pi^-$ mass fit, the $B^0 \rightarrow D_s^+ K^- \pi^- \pi^-$ contribution was modeled using the shape from the $B^0 \rightarrow D_s^+ \pi^- \pi^- \pi^-$ mass fit. To estimate an uncertainty from this assumption, the data were fitted with the shape obtained from $B^0 \rightarrow D_s^+ K^- \pi^- \pi^-$ simulation. A deviation of 5.5% in the fitted $B^0 \rightarrow D_s^+ K^- \pi^- \pi^-$ yield is found, with almost no change in the $B^0 \rightarrow D_s^+ K^- \pi^- \pi^-$ yield. The larger sensitivity on the $B^0$ yield than the $B^0$ yield arises because these background contributions have a rising edge in the vicinity of the $B^0$ mass peak, which is far enough below the $B^0$ mass peak to have negligible impact. These yield uncertainties are added in quadrature to obtain the values shown in Table III. The uncertainties due to the finite simulation sample sizes are 3.0%.

The major source of systematic uncertainty on the branching fraction for $B^0 \rightarrow D_s^+ (2536)^+ \pi^- , D_s^+ \rightarrow D_s^+ \pi^- \pi^- \pi^-$, is from the relative efficiency (5%), and on the fraction of events with $M > 3$ GeV/$c^2$ (10%). This 10% uncertainty is conservatively estimated by assuming a flat distribution in $M(X_d)$ up to 3 GeV/$c^2$ and then a linear decrease to zero at the phase space limit of $\approx 3.5$ GeV/$c^2$. Other systematic uncertainties related to the fit model are negligible. Thus in total, a systematic uncertainty of 11% is assigned to the ratio $B(B^0 \rightarrow D_s^+ (2536)^+ \pi^- , D_s^+ \rightarrow D_s^+ \pi^- \pi^- \pi^-) / B(B^0 \rightarrow D_s^+ \pi^- \pi^- \pi^-)$.

### IX. Results and Summary

This paper reports the first observation of the $B^0 \rightarrow D_s^+ K^- \pi^+ \pi^- \pi^-$, $B^0 \rightarrow D_s^+ K^- \pi^- \pi^- \pi^-$ and $B^0 \rightarrow D_s^+ (2536)^+ \pi^- , D_s^+ \rightarrow D_s^+ \pi^- \pi^- \pi^-$ decays. The ratios of branching fractions are measured to be

$$\frac{B(B^0 \rightarrow D_s^+ K^- \pi^+ \pi^- \pi^-)}{B(B^0 \rightarrow D_s^+ K^- \pi^- \pi^- \pi^+)} = (5.2 \pm 0.5 \pm 0.3) \times 10^{-2},$$

$$\frac{B(B^0 \rightarrow D_s^+ K^- \pi^- \pi^- \pi^+)}{B(B^0 \rightarrow D_s^+ K^- \pi^- \pi^- \pi^-)} = (5.2 \pm 0.5 \pm 0.3) \times 10^{-2},$$

and

$$\frac{B(B^0 \rightarrow D_s^+ (2536)^+ \pi^- , D_s^+ \rightarrow D_s^+ \pi^- \pi^- \pi^-) / B(B^0 \rightarrow D_s^+ \pi^- \pi^- \pi^-)}{(4.0 \pm 1.0 \pm 0.4) \times 10^{-3}},$$

where the uncertainties are statistical and systematic, respectively. The $B^0 \rightarrow D_s^+ K^- \pi^+ \pi^- \pi^-$ branching fraction is consistent with expectations from Cabibbo suppression. This decay is particularly interesting because it can be used in a time-dependent analysis to measure the CKM phase $\gamma$. Additional studies indicate that this decay mode, with selections optimized for only $B^0 \rightarrow D_s^+ K^- \pi^+ \pi^-$, can contribute about an additional 35% more signal events relative to the signal yield in $B^0 \rightarrow D_s^+ K^- \pi^- \pi^- \pi^-$ alone.

The $B^0 \rightarrow D_s^+ K^- \pi^+ \pi^- \pi^-$ branching fraction is about 50% of that for $B^0 \rightarrow D_s^+ K^- \pi^- \pi^- \pi^-$. Compared to the $B^0 \rightarrow D_s^+ K^- \pi^- \pi^- \pi^-$ decay that proceeds only via a $W$-exchange diagram, where $B(B^0 \rightarrow D_s^+ K^-) / B(B^0 \rightarrow D_s^+ K^-) \sim 0.1$ [21], the ratio $B(B^0 \rightarrow D_s^+ K^- \pi^+ \pi^- \pi^-) / B(B^0 \rightarrow D_s^+ K^- \pi^+ \pi^- \pi^-) \sim 0.1$ is about five times larger. A consistent explanation of this larger $B^0 \rightarrow D_s^+ K^- \pi^- \pi^- \pi^- \pi^-$ branching fraction is that only about 1/5 of the rate is from the $W$-exchange process [Fig. 1(d)] and about 4/5 comes from the diagrams shown in Figs. 1(e) and 1(f). The observed $M(X_d)$, $M(K^+ \pi^+ \pi^-)$ and $M(\pi^- \pi^- \pi^-)$ distributions in Fig. 6 also support this explanation, as evidenced by the qualitative agreement with the simulation.

### Acknowledgments

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of

TABLE III. Summary of systematic uncertainties (in %) on the measurements of the ratios of branching fractions.

| Source                  | $\frac{B(B^0 \rightarrow D_s^+ K^- \pi^- \pi^- \pi^-)}{B(B^0 \rightarrow D_s^+ K^- \pi^- \pi^- \pi^-)}$ | $\frac{B(B^0 \rightarrow D_s^+ K^- \pi^- \pi^- \pi^-)}{B(B^0 \rightarrow D_s^+ K^- \pi^- \pi^- \pi^-)}$ | $\frac{B(B^0 \rightarrow D_s^+ (2536)^+ \pi^- , D_s^+ \rightarrow D_s^+ \pi^- \pi^- \pi^-)}{B(B^0 \rightarrow D_s^+ \pi^- \pi^- \pi^-)}$ |
|------------------------|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|
| $f_s / f_s$            | 7.9                                             | 7.9                                             | 7.9                                             |
| $M(X_{j,2}) > 3$ GeV/$c^2$ | 2.2                                             | 2.2                                             | 2.2                                             |
| Efficiency             | 1.6                                             | 1.6                                             | 1.6                                             |
| PID                    | 2.2                                             | 2.2                                             | 2.2                                             |
| Trigger                | 2.0                                             | 2.0                                             | 2.0                                             |
| Signal yields          | 4.0                                             | 4.0                                             | 4.0                                             |
| Simulated sample size  | 3.0                                             | 3.0                                             | 3.0                                             |
| Total                  | 6.4                                             | 13.4                                            | 13.4                                            |
the LHC. We thank the technical and administrative staff at the LHCb institutes. We acknowledge support from CERN and from the following national agencies: CAPES, CNPq, FAPERJ and FINEP (Brazil); NSFC (China); CNRS/IN2P3 and Region Aquervgne (France); BMBF, DFG, HGF and MPG (Germany); SFI (Ireland); INFN (Italy); FOM and NWO (The Netherlands); SCSR (Poland); ANCS/IFA (Romania); MinES, Rosatom, RFBR and NRC “Kurchatov Institute” (Russia); MinECo, XuntaGal and GENCAT (Spain); SNSF and SER (Switzerland); NAS Ukraine (Ukraine); STFC (United Kingdom); and the NSF (USA). We also acknowledge the support received from the ERC under FP7. The Tier1 computing centers are supported by IN2P3 (France), KIT and BMBF (Germany), INFN (Italy), NWO and SURF (The Netherlands), CIEMAT, IFAE and UAB (Spain), GridPP (United Kingdom). We are thankful for the computing resources put at our disposal by Yandex LLC (Russia), as well as to the communities behind the multiple open source software packages that we depend on.

[1] N. Cabibbo, Phys. Rev. Lett. 10, 531 (1963).
[2] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. 49, 652 (1973).
[3] M. Bona et al. (UTfit Collaboration) (Nucl. Phys. B, Proc. Suppl., to be published). Additional information available at http://www.utfit.org/UTfit/.
[4] S. Descotes-Genon et al. (CKMFitter Collaboration) (Nucl. Phys. B, Proc. Suppl., to be published). Updated results and plots available at http://ckmfitter.in2p3.fr.
[5] I. Dunietz, Phys. Lett. B 270, 75 (1991); Z. Phys. C 56, 129 (1992); D. Atwood, G. Eilam, M. Gronau, and A. Soni, Phys. Lett. B 341, 372 (1995); D. Atwood, I. Dunietz, and A. Soni, Phys. Rev. Lett. 78, 3257 (1997).
[6] M. Gronau and D. London, Phys. Lett. B 253, 483 (1991); M. Gronau and D. Wyler, Phys. Lett. B 265, 172 (1991).
[7] A. Giri, Y. Grossman, A. Soffer, and J. Zupan, Phys. Rev. D 68, 054018 (2003).
[8] I. Dunietz and R. G. Sachs, Phys. Rev. D 37, 3186 (1988); 39, 3515 (1989).
[9] R. Aleksan, I. Dunietz, and B. Kayser, Z. Phys. C 54, 653 (1992).
[10] I. Dunietz, Phys. Rev. D 52, 3048 (1995).
[11] R. Fleischer, Nucl. Phys. B671, 459 (2003).
[12] R. Aaij et al. (LHCb collaboration), Phys. Rev. D 84, 092001 (2011).
[13] A. A. Alves Jr. et al. (LHCb collaboration), JINST 3, S08005 (2008).
[14] V.Y. Gligorov, C. Thomas, and M. Williams, Report No. LHcb-PUB-2011-016.
[15] T. Sjöstrand, S. Mrenna, and P. Skands, J. High Energy Phys. 05 (2006) 026.