Determination of the Ratio of b-Quark Fragmentation Fractions $f(s)/f(d)$ in pp Collisions at root $s=7$ TeV with the ATLAS Detector

Aad, G.; Abbott, B.; Abdallah, J.; Abdinov, O.; Aben, R.; Abolins, M.; AbouZeid, O.S.; Abramowicz, H.; Abreu, H.; Abreu, R.; Dam, Mogens; Hansen, Jørn Dines; Hansen, Jørgen Beck; Xella, Stefania; Hansen, Peter Henrik; Petersen, Troels Christian; Thomsen, Lotte Ansgaard; Mehlhase, Sascha; Jørgensen, Morten Dam; Pingel, Almut Maria; Løvschall-Jensen, Ask Emil; Alonso Diaz, Aléjandro; Monk, James William; Pedersen, Lars Egholm; Wiglesworth, Graig; Galster, Gorm Aske Gram Krohn

Published in:
Physical Review Letters

DOI:
10.1103/PhysRevLett.115.262001

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):

Download date: 29. aug., 2017
Determination of the Ratio of b-Quark Fragmentation Fractions $f_s/f_d$

in $pp$ Collisions at $\sqrt{s} = 7$ TeV with the ATLAS Detector

G. Aad et al.*

(ATLAS Collaboration)

(Received 3 August 2015; published 30 December 2015)

With an integrated luminosity of 2.47 fb$^{-1}$ recorded by the ATLAS experiment at the LHC, the exclusive decays $B^0_s \rightarrow J/\psi K^0$ and $B^0_d \rightarrow J/\psi K^0$ of $B$ mesons produced in $pp$ collisions at $\sqrt{s} = 7$ TeV are used to determine the ratio of fragmentation fractions $f_s/f_d$. From the observed $B^0_s \rightarrow J/\psi \phi$ and $B^0_s \rightarrow J/\psi K^0$ yields, the quantity $(f_s/f_d)[B(B^0_s \rightarrow J/\psi \phi)/B(B^0_s \rightarrow J/\psi K^0)]$ is measured to be $0.199 \pm 0.004$ (stat) $\pm 0.008$ (syst). Using a recent theory prediction for $[B(B^0_s \rightarrow J/\psi \phi)/B(B^0_s \rightarrow J/\psi K^0)]$ yields $(f_s/f_d) = 0.240 \pm 0.004$ (stat) $\pm 0.010$ (syst) $\pm 0.017$ (th). This result is based on a new approach that provides a significant improvement of the world average.


The production rate of $B^0_s$ ($B^0_d$) mesons is a product of the $b\bar{b}$ cross section, the instantaneous luminosity and the probability that the $b$ quark is bound to an $s$ ($d$) quark. The latter, denoted by the fragmentation fraction $f_s$ ($f_d$), depends on the probability that in pQCD-inspired calculations [1,2], a soft gluon splits into $s\bar{s}$ ($d\bar{d}$) and that the overlap of the $b$ and $s$ ($d$) wave functions is sufficiently large to produce a $B^0_s$ ($B^0_d$) bound state. In a similar fashion, $B^+$ mesons, $B^-$ mesons, and $b$ baryons are produced at the LHC with respective fragmentation fractions $f_u$, $f_c$, and $f_{b\text{baryon}}$. The fragmentation fractions are about 40% each for $u$ and $d$ quarks, 10% for $s$ quarks, at the percent level for $c$ quarks, and ~8% for baryon production satisfying the constraint $f_u + f_d + f_s + f_c + f_{b\text{baryon}} = 1$. Precise knowledge of the fragmentation fractions is essential for measuring $b$-hadron cross sections and branching fractions at the LHC. In particular, for rare decays, such as the branching fraction measurement of $B^0_s \rightarrow \mu^+\mu^-$ [3–5], a precise knowledge of $f_s/f_d$ is important since it improves the sensitivity of searches for new physics processes beyond the standard model (SM). The fragmentation ratio $f_s/f_d$ is a universal quantity that was measured by LEP experiments [6], CDF [7], and LHCB [8,9]. This Letter presents a measurement of $f_s/f_d$ using $B^0_s \rightarrow J/\psi \phi$ and $B^0_d \rightarrow J/\psi K^0$ decays.

The ratio of fragmentation fractions $f_s/f_d$ is extracted from the measured $B^0_s \rightarrow J/\psi \phi$ and $B^0_d \rightarrow J/\psi K^0$ signal yields, $N_{B^0_s}$ and $N_{B^0_d}$. These are converted into $B^0_s$ and $B^0_d$ meson yields after dividing by the branching fractions of the relevant decays and correcting for the relative efficiency $\mathcal{R}_{\text{eff}}$ that is expressed as a product of acceptance and selection efficiency ratios for the two modes and is determined from Monte Carlo (MC) simulations:

$$\frac{f_s}{f_d} = \frac{N_{B^0_s}}{N_{B^0_d}} \frac{B(B^0_s \rightarrow J/\psi K^0) B(K^0 \rightarrow K^+\pi^-)}{B(B^0_d \rightarrow J/\psi \phi) B(\phi \rightarrow K^+K^-)} \mathcal{R}_{\text{eff}}.$$

where the $J/\psi$, $\phi$, and $K^0$ are reconstructed in their $J/\psi \rightarrow \mu^+\mu^-$, $\phi \rightarrow K^+K^-$, and $K^0 \rightarrow K^+\pi^-$ final states [10], respectively. The data sample consists of $pp$ collisions collected with the ATLAS detector at $\sqrt{s} = 7$ TeV corresponding to an integrated luminosity of 2.47 ± 0.04 fb$^{-1}$. The ATLAS multipurpose detector is described in detail in Ref. [11].

The PYTHIA 6 and 8 [12,13] MC generators with parameters tuned to reproduce ATLAS data [14] are used to simulate background and signal events, respectively. For the signal channels, the angular distributions are produced with the measured polarization parameters [15]. The detector response for the generated events is simulated with GEANT4 [16,17].

The $B^0_s \rightarrow J/\psi \phi$ and $B^0_d \rightarrow J/\psi K^0$ signal candidates consist of two muons and two hadrons originating from a common secondary vertex. The $J/\psi$ candidates are selected from the dimuon trigger sample requiring two oppositely charged muon candidates, each having a transverse momentum of $p_T > 4$ GeV. Reconstructed muon candidates are categorized either as combined or segment-tagged muons. A combined muon consists of an inner detector (ID) track combined with a muon spectrometer (MS) track using tight matching criteria, while a segment-tagged muon requires an ID track and track segments in the MS that are not reconstructed as an MS track [11]. The two muons, of which at least one must be a combined muon, are fitted to originate from the same two-track vertex. The

*Full author list given at the end of the article.

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.
vertex fit chi-square per degree of freedom (dof) is required to be $\chi^2$/dof < 10. To improve the sample purity, each muon track must have at least one hit in the pixel detector, more than five hits in the silicon strip detector and at least one hit in the transition radiation tracker that reduces the pseudorapidity coverage to $|\eta| < 2.0$ [18].

Since the dimuon mass resolution is different for muons reconstructed in the end caps ($1.05 < |\eta| < 2.5$) and for muons reconstructed in the barrel ($|\eta| < 1.05$), all accepted $J/\psi$ candidates are divided into three classes: two barrel muons (BB), one end-cap and one barrel muon (EB), and two end-cap muons (EE). The parameters describing the dimuon mass distribution in the $J/\psi$ signal region for the three pseudorapidity classes in data and in $B_d^0 \to J/\psi\phi$ and $B_d^0 \to J/\psi K^{*0}$ MC signal samples are extracted from maximum-likelihood fits. Signal events are selected requiring mass windows of ±3σ around the $J/\psi$ peak in data and simulations. For data, the selected signal regions are 2.991–3.197 GeV for BB, 2.955–3.235 GeV for EB, and 2.914–3.275 GeV for EE classes, while in simulations they are slightly smaller.

The $B_d^0$ candidates are reconstructed from a $J/\psi$ candidate plus two oppositely charged hadrons with a kaon mass hypothesis assigned. The dimuon mass is constrained to the $J/\psi$ mass [15], and the $J/\psi$ and two kaons have to originate from the same vertex. All combinations are accepted if $p_T(B_d^0) > 8$ GeV, $\chi^2$/dof < 3 for the vertex fit and the $K^+K^-$ invariant mass lies in the range determined by ±2 natural widths ($\Gamma_\phi$) around the $\phi$ mass peak, 1011 < $m_{K^+K^-} < 1028$ MeV. The $m_{K^+K^-}$ distribution is modeled with a Breit-Wigner line shape convolved with a Crystal Ball function [19]. The selected mass window retains 85% of signal events.

The $B_d^0$ candidates are reconstructed in a similar way. Here, one track of the $K^{*0}$ decay is assigned a kaon mass hypothesis and the other track a pion mass hypothesis. Since ATLAS has limited kaon-pion separation capability in the momentum range relevant for this analysis, both $K\pi$ mass assignment combinations are tested. That with mass closest to the nominal $K^{*0}$ mass is chosen, yielding the correct $K\pi$ selection for 86% of all $K^{*0}$ candidates. The probability density function (PDF) for the invariant mass of correctly selected $K\pi$ candidates is modeled with a relativistic Breit-Wigner line shape convolved with a Crystal Ball function, while that where the $K$ and $\pi$ are swapped is modeled with a Gaussian function. The decay $B_d^0 \to J/\psi\phi$ produces a peaking background in $B_d^0 \to J/\psi K^{*0}$ that appears in the low $K^{*0}$ mass region. To remove this contribution, the selected $K^{*0}$ region is constrained to one $K^{*0}$ decay width around the $K^{*0}$ mass peak, corresponding to 847 < $m_{K\pi} < 942$ MeV for data. Since the $K^{*0}$ line shape is narrower in the MC simulations than in data, the $K\pi$ mass selection needs to be adjusted in simulations to produce identical efficiencies in data and simulations. For the $K^+K^-$ mass selection, a similar procedure is used.

The signal-to-background ratios for $B_d^0 \to J/\psi\phi$ and $B_d^0 \to J/\psi K^{*0}$ decays are optimized using three variables with high background suppression power: the $\chi^2$/dof of the $B$ vertex fit, the transverse decay length $L_{xy}$ defined as the length of the vector from the primary vertex (PV) [20] to the $B$ decay vertex in the transverse plane, and the pointing angle $\alpha$ defined as the angle between the $B$ meson transverse momentum and $L_{xy}$. If more than one PV candidate exists, the one is selected for which the sum of squared transverse momenta of all tracks originating from the vertex, $\sum p_T^2$, yields the highest value. The $\chi^2$/dof, $L_{xy}$ and $\alpha$ selection criteria are optimized using simulated $B_d^0 \to J/\psi\phi$ and $B_d^0 \to J/\psi K^{*0}$ events for signal and data sidebands for background.

To produce similar $p_T$ and $\eta$ distributions in data and MC, data-driven weights are obtained by the following procedure. Sideband-subtracted $B_d^0 \to J/\psi\phi$ ($B_d^0 \to J/\psi K^{*0}$) $p_T$ and $\eta$ distributions from data are compared with corresponding distributions in simulation in the signal region. 5.32 < $m_{J/\psi\phi}$ < 5.42 (5.21 < $m_{J/\psi K^{*0}} < 5.35$) GeV. The upper and lower sidebands 5.20 < $m_{J/\psi\phi}$ < 5.25 (5.09 < $m_{J/\psi K^{*0}} < 5.16$) GeV and 5.48 < $m_{J/\psi\phi}$ < 5.53 (5.40 < $m_{J/\psi K^{*0}} < 5.47$) GeV are selected such that their summed yields represent the expected backgrounds in the signal region for the data. The weights are obtained by dividing the yield in each $p_T$ and $\eta$ bin in data by the corresponding yield of the MC sample using only events with odd event numbers. Thus, for each bin $(i)$ and $(j)$ of the $p_T$ and $\eta$ distributions, a weight is determined as a product of a $p_T$–dependent and $\eta$–dependent weights:

$$W_{ij}(p_T, \eta) = \frac{n_{ij}^{\text{data/MC}(p_T)}}{n_{ij}^{\text{MC}(p_T)}} \frac{n_{ij}^{\text{data}(\eta)}}{n_{ij}^{\text{MC}(\eta)}},$$

where $n_{ij}^{\text{data/MC}(p_T)}$ is the normalized number of entries in the $p_T$ bin $i$ and $n_{ij}^{\text{data/MC}(\eta)}$ is that in the $\eta$ bin $j$. To obtain good agreement between data and simulation, the procedure is repeated twice. The two sets of weights are multiplied and are used to correct the $p_T$ and $\eta$ distributions of the MC sample with even event numbers. From the corrected MC samples, distributions for $\chi^2$/dof, $L_{xy}$, and $\alpha$ are determined, which are in good agreement with those measured in the data. The correlation between $p_T$ and $\eta$ is small and is accounted for in the systematic error.

For both modes, the dominant background originates from a $J/\psi$ produced at the PV plus two oppositely charged hadrons (direct $J/\psi$) [21]. Since the hadrons are not associated with any $B_d^0 (B_d^0)$ decay, the $J/\psi K^+K^- (J/\psi K^+\pi^-)$ invariant-mass spectrum does not peak but decreases with mass. Another large background consists of two random low-momentum, oppositely
charged muons combined with two random charged hadrons. Here, the dimuon mass distribution does not peak at the $J/\psi$ nor does the four-particle mass show any peaking structure. Inclusive decays $B \to J/\psi X$, where $X$ is a single hadron or a collection of hadrons, provide a source of background that is very similar to the signal. If $X$ consists of exactly two charged-particle tracks (without any $\pi^0$), the mode is topologically indistinguishable from the signal mode. Self-cross-feed, in which one or both hadrons from the $\phi(K^0)$ decay are replaced with random hadrons, is negligible. In addition, peaking backgrounds from $B_d^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi K^+\pi^-$ contribute to $B_s^0 \to J/\psi\phi$ while $B_d^0 \to J/\psi K^+\pi^-$ also contributes to $B_d^0 \to J/\psi K^{*0}$.

To reduce these backgrounds, the $\chi^2$/dof, $L_{xy}$, and $\alpha$ selections are optimized for each mode separately by determining the maximum value of $S/\sqrt{S+B}$ as a function of selected values for the observable to be optimized, where $S$ represents the signal yield obtained from simulation and $B$ is the background extracted from data sidebands. For the $B_s^0$ ($B_d^0$) mode, the optimization yields $\chi^2$/dof < 2.4 (2.6), $L_{xy} > 0.26 (0.30)$ mm, and $\alpha < 0.14 (0.12)$ rad. In combination with the $J/\psi$ mass requirement, the $\chi^2$/dof selection reduces the combinatorial background significantly, while the $L_{xy}$ and $\alpha$ selections remove most of the direct $J/\psi$ background.

In the final sample, the signal yields $N_{B_s^0}$ and $N_{B_d^0}$ are extracted from unbinned extended maximum-likelihood fits to the $J/\psi K^+K^-$ and $J/\psi K^+\pi^-$ invariant-mass spectra, respectively. The $B_s^0$ signal PDF is modeled with three Gaussian functions with common mean that is determined from the fit, while widths and fractions are fixed to the values obtained from MC simulations. To account for possible width differences in the two narrowest Gaussian functions between data and simulation, an additional scale factor is introduced, which is left free in the fit. The peaking background PDF is modeled with a Crystal Ball function with parameters fixed to the values obtained in simulations.

The peaking background yield of $652 \pm 93$ events is calculated from the $B_d^0$ signal yield. The selection efficiencies of both peaking background modes are determined from simulation and are fixed in the fit to data. The remaining residual backgrounds are modeled with an exponential function leaving fraction and exponent free in the fit to data.

The $B_d^0$ signal PDF is parametrized with three Gaussian functions that describe both the correctly reconstructed and swapped $K^+\pi^-$ events. The PDF of the peaking background is modeled with a sum of Crystal Ball and Gaussian functions for which the relative $B_d^0 \to J/\psi K^+\pi^-$ yield with respect to that of the $B_d^0 \to J/\psi K^{*0}$ signal is determined from the corresponding branching fractions and selection efficiencies, yielding $(4.7 \pm 2.4)\%$. Most of the residual background is modeled with an exponential function, while partially reconstructed $B \to J/\psi X$ decays require parametrization with a complementary error function. All parameters of the residual background PDFs are left free in the fit.

Figure 1 shows the measured $J/\psi\phi$ and $J/\psi K^{*0}$ invariant-mass spectra with fits overlaid. The fits yield $N_{B_s^0} = 6640 \pm 100$ $B_s^0 \to J/\psi\phi$ and $N_{B_d^0} = 36290 \pm 320$ $B_d^0 \to J/\psi K^{*0}$ signal events. The $\chi^2$/dof values of the fits are 0.959 for $B_s^0$ and 0.945 for $B_d^0$, indicating that both fits describe the data well.

The additive systematic uncertainties result from the $B_s^0 \to J/\psi\phi$ and $B_d^0 \to J/\psi K^{*0}$ signal and background parametrizations. The contribution from the signal shape parametrization is calculated by varying the five fixed parameters within $\pm 1\sigma$ in a multivariate Gaussian function that takes into account all correlations. For nonpeaking backgrounds, the exponential function is replaced with a second-order polynomial for the $B_s^0$ and with a second-order polynomial plus an error function for the $B_d^0$. The difference in signal yield with respect to the nominal fit is taken as a systematic error. For peaking backgrounds, the

![FIG. 1](color online) The invariant-mass spectra of $B_s^0 \to J/\psi\phi$ (left panel) and $B_d^0 \to J/\psi K^{*0}$ decays (right panel) for data (points with error bars), total fit (solid line), signal (dashed line), residual background (yellow shaded), partially reconstructed events (magenta shaded) and peaking background (blue shaded).
TABLE I. Measured $B_s^0$ and $B_d^0$ signal yields, the efficiency ratio $\mathcal{R}_{\text{eff}}$ extracted from simulations, world averages for $\phi$ and $K^{*0}$ decay branching fractions, as well as corresponding systematic uncertainties $\sigma$ on $(f_s/f_d)\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)/\mathcal{B}(B_d^0 \rightarrow J/\psi K^{*0})$.

<table>
<thead>
<tr>
<th>Observable</th>
<th>Value</th>
<th>$\sigma$</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{B_s^0}$</td>
<td>6640 ± 100 ± 220</td>
<td>3.3%</td>
<td></td>
</tr>
<tr>
<td>$N_{B_d^0}$</td>
<td>36290 ± 320 ± 650</td>
<td>1.8%</td>
<td></td>
</tr>
<tr>
<td>$\mathcal{R}_{\text{eff}}$</td>
<td>0.799 ± 0.001 ± 0.010</td>
<td>1.3%</td>
<td></td>
</tr>
<tr>
<td>$B(\phi \rightarrow K^+ K^-)$</td>
<td>0.489 ± 0.005</td>
<td>1.0%</td>
<td>[15]</td>
</tr>
<tr>
<td>$B(K^{*0} \rightarrow K^+ \pi^-)$</td>
<td>0.66503 ± 0.00014</td>
<td>0.02%</td>
<td>[15]</td>
</tr>
<tr>
<td>Total</td>
<td>4.1%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

fixed parameters are varied by ±1σ, and the difference with respect to the nominal yield is taken as a systematic error. In addition, since S-wave contributions from $B_s^0 \rightarrow J/\psi K^+ K^-$ and $B_d^0 \rightarrow J/\psi f_0 (980)$ decays to $B_s^0 \rightarrow J/\psi \phi$ and $B_d^0 \rightarrow J/\psi K^{*0}$ are neglected in the fits, an uncertainty is derived using the ATLAS measured contribution of 2.4% [22] for $B_s^0 \rightarrow J/\psi \phi$, and the contribution of 1% for $B_d^0 \rightarrow J/\psi K^{*0}$ derived from the MC simulation. All additive systematic errors are added in quadrature, yielding total additive uncertainties of 220 $N_{B_s^0}$ and 650 $N_{B_d^0}$ events.

The multiplicative systematic uncertainty includes contributions from the relative efficiency and the branching fractions of the $\phi$ and $K^{*0}$ decays. The uncertainty on the relative efficiency is dominated by the uncertainty on the $\phi/K^{*0}$ selection (1.2%), which is obtained by varying the fixed fit parameters in the $\phi$ and $K^{*0}$ fits by ±1σ and adding all contributions in quadrature. Other uncertainties from the $J/\psi$ selection (0.2%), reweighting (0.4%), $B_s^0$ and $B_d^0$ lifetimes (0.002%), and the contribution due to uncertainties in the polarization parameters (0.01%) are negligible. Varying the selection criteria of $\chi^2$/dof, $L_{xy}$ and $\alpha$ gives negligible contributions. Table I summarizes the contributions of the additive and multiplicative systematic errors.

From the ratio $N_{B_s^0}/N_{B_d^0}$ after efficiency correction and division by $\phi$ and $K^{*0}$ decay branching fractions, ATLAS measures

$$\frac{f_s}{f_d} \mathcal{B}(B_s^0 \rightarrow J/\psi \phi)/\mathcal{B}(B_d^0 \rightarrow J/\psi K^{*0}) = 0.199 \pm 0.004(\text{stat}) \pm 0.008(\text{syst}).$$

A perturbative QCD prediction [23] yields

$$\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)/\mathcal{B}(B_d^0 \rightarrow J/\psi K^{*0}) = 0.83^{+0.03 -0.02} (\omega_B)^{+0.01 -0.00} (f_M)^{+0.01 -0.00} (a_1)^{+0.01 -0.02} (m_c),$$

where the uncertainties result from the shape parameter $\omega_B$ of the $B$ meson wave function, meson decay constants $f_M$, Gegenbauer moments $a_1$ in the wave functions of the light vector mesons and the $c$-quark mass. Adding all contributions linearly yields a 7.1% theory error. Using this prediction, the ratio of fragmentation fractions is measured to be

$$\frac{f_s}{f_d} = 0.240 \pm 0.004(\text{stat}) \pm 0.010(\text{syst}) \pm 0.017(\text{th}).$$

Figure 2 (right panel) shows the ATLAS $f_s/f_d$ measurement in comparison with results from LEP [6], CDF [6,7], and LHCb [8,9]. The ratio $f_s/f_d$ may depend on $p_T$ and $\eta$ of the $B$ meson; e.g., LHCb observes a $p_T$ but no $\eta$ dependence of $f_s/f_d$ [8]. Figure 2 (left panel) shows the $p_T$ dependence of $f_s/f_d$ for ATLAS and that of other
experiments. To investigate the $p_T$ and $\eta$ dependence of $f_s/f_d$, the data sample is divided into six $p_T$ bins in the range $8 \text{ GeV} < p_T < 50 \text{ GeV}$ and into four $\eta$ bins for $|\eta| < 2.5$ such that the number of events in each bin is approximately equal. The $f_s/f_d$ distributions as a function of $p_T$ and $\eta$ have been fitted with a uniform (first-order polynomial) distribution yielding fit $p$ values 0.54 (0.66) and 0.66 (0.49), respectively. No significant $f_s/f_d$ dependence on $p_T$ and $|\eta|$ is seen at the present level of accuracy.

In summary, this Letter reports on the first ATLAS measurement of the ratio of $B_s^0 \rightarrow J/\psi f$ and $B_d^0 \rightarrow J/\psi K^{*0}$ branching fractions multiplied by the ratio of fragmentation fractions $f_s/f_d$ from which $f_s/f_d$ is determined. The data were produced at the LHC in $pp$ collisions at $\sqrt{s} = 7 \text{ TeV}$ and correspond to an integrated luminosity of $2.47 \text{ fb}^{-1}$. This $f_s/f_d$ measurement, obtained with a new approach, agrees with the LHCb [8,9] results, improving the world average considerably. A comparison with the CDF [6,7] measurement and the LEP [6] average confirms the universality of $f_s/f_d$. The ATLAS data show no dependence on $p_T$ nor on $|\eta|$ within the kinematic range tested.

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC, and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST, and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR, and VSC CR, Czech Republic; DLR, DMSF, and HGF, Germany; INFN, Italy; INFN, Italy; KIT/GridKA, Germany; INRN, India; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; FOM and NWO, Netherlands; BRF and RCN, Norway; MNiSW and NCN, Poland; GRICES and FCT, Portugal; FOM and NWO, Netherlands; BRF and RCN, Norway; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SER, SNSF, and Cantons of Bern and Geneva, Switzerland; NSC, Taiwan; TAEK, Turkey; STFC, the Royal Society and Leverhulme Trust, United Kingdom; DOE and NSF, USA. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular, from CERN and the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA), and in the Tier-2 facilities worldwide.

[5] CMS Collaboration, Measurement of the $B_s^0 \rightarrow \mu^+\mu^-$ Branching Fraction and Search for $B^0 \rightarrow \mu^+\mu^-$ Decays with the CMS Experiment, Phys. Rev. Lett. 111, 101804 (2013).
[7] T. Aaltonen et al. (CDF Collaboration), $BR(B^0_s \rightarrow J/\psi\phi)$ measurement and extraction of the fragmentation fractions, Public CDF Note No. 10795, 2012.
[10] Charge conjugation is implied unless stated otherwise.
[18] ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector and the $z$ axis along the beam pipe. The $x$ axis points from the IP to the center of the LHC ring, and the $y$ axis points upward. Cylindrical coordinates $(r, \phi)$ are used in the transverse plane, $\phi$ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln(\tan(\theta/2))$.
[20] The PV is the parton interaction vertex. The one of interest is that where the $B$ meson is produced.

