Measurement of the shape of the $B_{s}^{0} \rightarrow D_{s}^{*-} \mu^{+}\nu_{\mu}$ differential decay rate

LHCb collaboration†

Abstract

The shape of the $B_{s}^{0} \rightarrow D_{s}^{*-} \mu^{+}\nu_{\mu}$ differential decay rate is obtained as a function of the hadron recoil using proton-proton collision data at a centre-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 1.7 fb$^{-1}$ collected by the LHCb detector. The $B_{s}^{0} \rightarrow D_{s}^{*-} \mu^{+}\nu_{\mu}$ decay is reconstructed through the decays $D_{s}^{*-} \rightarrow D_{s}^{-}\gamma$ and $D_{s}^{-} \rightarrow K^{-}K^{+}\pi^{-}$. The differential decay rate is fitted with the Caprini-Lellouch-Neubert (CLN) and Boyd-Grinstein-Lebed (BGL) parametrisations of the form factors, and the relevant quantities for both are extracted.

Submitted to JHEP

© 2020 CERN for the benefit of the LHCb collaboration. [CC BY 4.0 licence]

†Authors are listed at the end of this paper.
1 Introduction

Semileptonic decays of heavy hadrons are commonly used to measure the parameters of the Cabibbo-Kobayashi-Maskawa (CKM) matrix \( |V_{cb}| \), as they involve only one hadronic current that can be parametrised in terms of scalar functions known as form factors. The number of form factors needed to describe a particular decay depends upon the spin of the initial- and final-state hadrons \([3-5]\). For the decay of a pseudoscalar \( B \) meson to a vector \( D^* \) meson, four form factors are required. The determination of the CKM matrix element \( |V_{cb}| \) using \( B \to D^{(*)}\ell\nu_\ell \) decays or via the inclusive sum of all hadronic \( B \to X_c\ell\nu_\ell \) decay channels has been giving inconsistent results during the last thirty years \([6]\). The exclusive determination relies heavily on the parametrisation of the form factors, as it requires an extrapolation of the differential decay rate to the zero recoil point, where the momentum transfer to the lepton system is maximum.

Recently, the LHCb collaboration has measured \( |V_{cb}| \) using \( B^0 \to D_s^{(*)}\mu^+\nu_\mu \) decays\(^1\), with two form-factor parametrisations, giving consistent results \([7]\). The determination of the form factors in \( B^0_s \to D_s^-\ell^+\nu_\ell \) decays obtained using different parametrisations can help to clarify the \( |V_{cb}| \) inconsistency between the exclusive and inclusive approaches. It can also be used to improve the Standard Model (SM) predictions of the \( B^0 \to D_s^-\tau^+\nu_\tau \) branching fraction and the ratio \( R(D_s^+) = B(B^0 \to D_s^-\tau^+\nu_\tau)/B(B^0 \to D_s^-\mu^+\nu_\mu) \). A measurement and precise prediction of the latter could increase the understanding of the current tension between experimental and theoretical values of the equivalent ratio \( R(D^{(*)}) = B(B \to D^{(*)}\tau^+\nu_\tau)/B(B \to D^{(*)}\mu^+\nu_\mu) \) \([6]\). Theoretical predictions on \( B^0_s \) semileptonic decays are expected to be more precise than those on \( B^0 \) or \( B^+ \) decays. For example, the Lattice QCD calculations of the form factors are computationally easier due to the larger mass of the spectator \( s \) quark compared to that of \( u \) or \( d \) quarks \([8,9]\). Despite these advantages, the study of semileptonic \( B^0_s \) decays has received less theoretical attention than the equivalent \( B^0 \) and \( B^+ \) decays due to the lack of experimental results.

This paper reports the first measurement of the shape of the differential decay rate of the \( B^0_s \to D_s^-\mu^+\nu_\mu \) decay as a function of the hadron recoil variable \( w = v_{B^0_s} \cdot v_{D_s^-} \), where \( v_{B^0_s} \) and \( v_{D_s^-} \) are the four-vector velocities of the \( B^0_s \) and \( D_s^- \) mesons, respectively. The spectrum of \( w \) is unfolded for the detector resolution on \( w \) and corrected for the reconstruction and selection efficiency.

The \( B^0_s \to D_s^-\mu^+\nu_\mu \) decay can be described by four form factors, but in the limit of zero recoil only one form factor becomes relevant. This leading form factor is fit using the two most commonly used parametrisations by Caprini-Lellouch-Neubert (CLN) \([10]\) and by Boyd-Grinstein-Lebed (BGL) \([11,13]\). The parameters of the leading form factor for each parametrisation are reported, assuming input from \( B \) decays for the parameters of the sub-leading form factors. The decay is reconstructed in the \( D_s^- \to D^-\gamma \) decay, where the \( D^- \) subsequently decays via \( D^- \to \phi (\rightarrow K^+K^-)\pi^- \) or \( D^- \to K^0 (\rightarrow \pi^- K^+)K^- \). The data used correspond to an integrated luminosity of 1.7 fb\(^{-1}\) collected by the LHCb experiment in 2016 at a centre-of-mass energy of 13 TeV.

\(^1\)The inclusion of charge-conjugate processes is implied throughout this paper.
2 Formalism of the $B^0_s \to D^{*-}_{s} \mu^+ \nu_{\mu}$ decay

The $B^0_s \to D^{*-}_{s} \mu^+ \nu_{\mu}$ decay, with the subsequent $D^{*-}_{s} \to D^{-}_{s} \gamma$ decay, can be described by three angular variables and the squared momentum transfer to the lepton system, defined as $q^2 = (p_{B^0_s} - p_{D^{*-}_{s}})^2$, where $p_{B^0_s}$ and $p_{D^{*-}_{s}}$ are the four-momenta of the $B^0_s$ and $D^{*-}_{s}$ mesons, respectively. The three angular variables, indicated in Fig. 1, are two helicity angles $\theta_{\mu}$ and $\theta_{D_s}$, and the angle $\chi$. The angle between the muon direction and the direction opposite to that of the $B^0_s$ meson in the virtual $W$ rest frame is called $\theta_{\mu}$, while the angle between $D^{-}_{s}$ direction and the direction of the $B^0_s$ meson in the $D^{*-}_{s}$ rest frame is called $\theta_{D_s}$. Finally, $\chi$ is the angle between the two planes formed by the virtual $W$ and $D^{*-}_{s}$ decay products in the $B^0_s$ rest frame [14]. The angles in $B^0_s$ decays are defined such that they are the same for $B^0_s$ and $\bar{B}^0_s$ mesons in the absence of CP violation.

The measurement is performed by integrating the full decay rate over the decay angles. Thus, the expression of the $B^0_s \to D^{*-}_{s} \mu^+ \nu_{\mu}$ decay rate is given by

$$\frac{d\Gamma(B^0_s \to D^{*-}_{s} \mu^+ \nu_{\mu})}{dq^2} = \frac{G_F^2 |V_{cb}|^2 |\eta_{EW}|^2 |\vec{p}| q^2}{96 \pi^3 m_{B^0_s}^2} \left(1 - \frac{m_\mu^2}{q^2}\right)^2 \times \left[ (|H_+|^2 + |H_-|^2 + |H_0|^2) \left(1 + \frac{m_\mu^2}{2q^2}\right) + \frac{3}{2} \frac{m_\mu^2}{q^2} |H_t|^2 \right]. \quad (1)$$

In this equation, $G_F$ is the Fermi constant, $V_{cb}$ is the CKM matrix element describing the $b \to c$ transition, $\eta_{EW} = 1.0066$ is the electroweak correction to $V_{cb}$ [15], $m_\mu$ is the muon mass [16], and $H_0$, $H_+$, $H_-$, $H_t$ are the helicity amplitudes of the lepton pair. The magnitude of the $D^{*-}_{s}$ momentum in the $B^0_s$ rest frame is given by $|\vec{p}|$. The dependence of the helicity amplitudes on $w$ can be expressed in different ways, most commonly parametrised in either the CLN or BGL expansion, as discussed further in Sec. 2.1 and Sec. 2.2.

The hadron recoil is related to the squared momentum transfer to the lepton pair, $q^2$,
by

\[ w = \frac{p_{B_0}}{m_{B_0}}, \quad \frac{p_{D^*_s}}{m_{D^*_s}} = \frac{m_{B_0}^2 + m_{D^*_s}^2 - q^2}{2 m_{B_0} m_{D^*_s}}, \quad (2) \]

where \( m_{B_0} \) and \( m_{D^*_s} \) are the masses of the \( B_0 \) and \( D^*_s \) mesons, respectively. The minimal value, \( w = 1 \), corresponds to the situation in which the \( D^*_s \) meson has zero recoil in the \( B_0 \) rest frame. It is also the value for which \( q^2 \) is maximal.

This measurement is only sensitive to a single form-factor contribution while the other form factors are fixed to existing measurements from \( B^+ \) and \( B^0 \) semileptonic decays \([6,17]\). This choice is supported by the good agreement of the form factors at zero recoil between \( B^0 \) and \( B^0 \) decays obtained in Ref. \([18]\).

### 2.1 CLN form-factor parametrisation

For the CLN parametrisation, it is useful to write the helicity amplitudes \( H_0, H_+, H_- \) and \( H_t \) in terms of the form factors \( A_1(w), V(w), A_2(w) \) and \( A_0(w) \) as

\[ H_\pm(w) = m_{B_0}(1 + r) A_1(w) \mp \frac{2}{1 + r} |\vec{p}| V(w), \]

\[ H_0(w) = \frac{m_{B_0} m_{D^*_s} (w - r)(1 + r)^2 A_1(w)}{m_{D^*_s} (1 + r) \sqrt{1 + r^2 - 2wr}}, \quad (3) \]

\[ H_t(w) = \frac{2 |\vec{p}|}{\sqrt{1 + r^2 - 2wr}} A_0(w), \]

where \( r = m_{D^*_s}/m_{B_0} \). The form factors are rewritten in terms of a single leading form factor

\[ h_{A_1}(w) = A_1(w) \frac{1}{R_{D^*_s} - w + 1}, \quad (4) \]

and three ratios of form factors

\[ R_0(w) = \frac{A_0(w)}{h_{A_1}(w)} R_{D^*_s}, \quad R_1(w) = \frac{V(w)}{h_{A_1}(w)} R_{D^*_s}, \quad R_2(w) = \frac{A_2(w)}{h_{A_1}(w)} R_{D^*_s}, \quad (5) \]

where

\[ R_{D^*_s} = \frac{2 \sqrt{r}}{1 + r}. \quad (6) \]

In the CLN parametrisation, the leading form factor and the three ratios are parameterised in terms of \( w \) as

\[ h_{A_1}(w) = h_{A_1}(1)[1 - 8 \rho^2 z(w) + (53 \rho^2 - 15) z^2(w) - (231 \rho^2 - 91) z^3(w)], \]

\[ R_0(w) = R_0(1) - 0.11(w - 1) + 0.01(w - 1)^2, \]

\[ R_1(w) = R_1(1) - 0.12(w - 1) + 0.05(w - 1)^2, \]

\[ R_2(w) = R_2(1) + 0.11(w - 1) - 0.06(w - 1)^2, \quad (7) \]

where the coefficients, originally calculated for \( B \) decays, are assumed to be the same for \( B^0 \) decays. The function \( z(w) \) is defined as

\[ z(w) = \frac{\sqrt{w + 1} - \sqrt{2}}{\sqrt{w + 1} + \sqrt{2}}. \quad (8) \]
Since this measurement is only sensitive to the shape of the form-factor parametrisation the term $h_{A_1}(1)$ is absorbed in the normalisation. The values of $R_1(1)$ and $R_2(1)$ are taken from the HFLAV average of the corresponding parameters, obtained from $B \to D^*\ell\nu_\ell$ decays [6]. The $R_0(1)$ parameter is suppressed by $m_\ell^2/q^2$ in the helicity amplitude and its contribution to the total rate is negligible. It has not been measured, but its value is predicted by the exact heavy quark limit of $R_0(1) = 1$ is therefore used in this measurement [19]. The slope, $\rho^2$, of $h_{A_1}(w)$ is the only parameter fitted in this parametrisation.

2.2 BGL form-factor parametrisation

In the BGL parametrisation [20], the helicity amplitudes are parametrised as

$$H_0(w) = \frac{\mathcal{F}_1(w)}{m_{B_s} \sqrt{1 + r^2 + 2wr}} ,$$

$$H_\pm(w) = f(w) \pm m_{B_s} m_{D_s^*} \sqrt{w^2 - 1} g(w) ,$$

$$H_z(w) = m_{B_s} \sqrt{r(1 + r)} \sqrt{w^2 - 1} \frac{1}{\sqrt{1 + r^2 - 2wr}} \mathcal{F}_2(w) ,$$

where the form factors are defined, following Ref. [21], as

$$f(z) = \frac{1}{P_1^+(z) \phi_f(z)} \sum_{n=0}^{\infty} a_n^f z^n , \quad \mathcal{F}_1(z) = \frac{1}{P_1^+(z) \phi_{\mathcal{F}_1}(z)} \sum_{n=0}^{\infty} a_n^{\mathcal{F}_1} z^n ,$$

$$g(z) = \frac{1}{P_1^-(z) \phi_g(z)} \sum_{n=0}^{\infty} a_n^g z^n , \quad \mathcal{F}_2(z) = \frac{\sqrt{r}}{(1 + r) P_0^-(z) \phi_{\mathcal{F}_2}(z)} \sum_{n=0}^{\infty} a_n^{\mathcal{F}_2} z^n .$$

The $\phi_i$ functions are the so-called outer functions and are defined in Ref. [22], the $P_{1,0}^\pm$-factors are Blaschke factors [7], and the coefficients $a_i^n$, where $i = \{f, g, \mathcal{F}_1, \mathcal{F}_2\}$, are parameters that need to be fitted from data.

As the form-factor parametrisation are given by analytic functions, this ensures that the coefficients of the $z$ expansion satisfy the unitarity condition

$$\sum_{n=0}^{\infty} (a_n^g)^2 \leq 1 , \quad \sum_{n=0}^{\infty} (a_n^f)^2 + \sum_{n=0}^{\infty} (a_n^{\mathcal{F}_1})^2 \leq 1 , \quad \sum_{n=0}^{\infty} (a_n^{\mathcal{F}_2})^2 \leq 1 .$$

This analysis is only sensitive to the form factor $f(z)$, and its series is truncated at $N = 2$, following Refs. [17, 20]. The shapes for $\mathcal{F}_i(z)$ and $g(z)$ are constrained by using the results in Ref. [17], where the $a_i^n$ coefficients are fitted using recent Belle measurements with $B^0 \to D^{*-} \ell^+\nu_\ell$ decays [23, 24]. The value of $a_0^f$ in Ref. [17] is determined from the combination of lattice calculations in Ref. [25]. The parameters $a_n^{\mathcal{F}_2}$ for $\mathcal{F}_2(z)$ are fixed from predictions in Ref. [20], where they are called $P_1$. An overview of the fit inputs is given in Tab. 6 in App. B.

3 Detector and simulation

The LHCb detector [26, 27] 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 pp interaction region \[28\], 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 \[29\] placed downstream of the magnet. The tracking system provides a measurement of the momentum, \(p\), of charged particles with a relative uncertainty that varies from 0.5% at low momentum to 1.0% at 200 GeV/c. The minimum distance of a track to a primary vertex (PV), the impact parameter (IP), is measured with a resolution of \((15 + 29/p_T) \mu m\), where \(p_T\) is the component of the momentum transverse to the beam, in GeV/c. Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov detectors \[30\]. Photons, electrons and hadrons are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic and a hadronic calorimeter. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers \[31\]. The online event selection is performed by a trigger \[32\], which 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 hardware muon trigger selects events containing a high-\(p_T\) muon candidate. The software trigger requires three tracks with a significant displacement from any primary pp interaction vertex.

Simulation is required to model the effects of the detector acceptance and the imposed selection requirements. In the simulation, pp collisions are generated using PYTHIA \[33\] with a specific LHCb configuration \[34\]. Decays of unstable particles are described by EVTGEN \[35\], in which final-state radiation is generated using PHOTOS \[36\]. The interaction of the generated particles with the detector, and its response, are implemented using the GEANT4 toolkit \[37\] as described in Ref. \[38\].

The simulation is corrected for mismodeling of the kinematic properties of the generated \(B_s^0\) mesons and of the photons from the \(D_s^{*-}\) decays, as well as for data-simulation differences in the muon trigger efficiency and tracking efficiencies of the final-state particles. Corrections to the \(B_s^0\) and \(\gamma\) kinematic distributions are determined by comparing data and simulated samples of \(B^+ \rightarrow J/\psi K^+\) and \(B_s^0 \rightarrow D_s^{*-}\pi^+\) decays, respectively. Kinematic differences between \(B_s^0\) and \(B^+\) mesons due to their production mechanisms are small and considered to be negligible \[39, 40\]. Corrections to the trigger and tracking efficiencies are evaluated using data and simulated samples of \(B^+ \rightarrow J/\psi K^+\) decays \[41\]. In the simulated signal sample, the form factors are described following the CLN parametrisation with numerical values \(\rho^2 = 1.205\), \(R_1(1) = 1.404\) and \(R_2(1) = 0.854\).

4 Data selection

Candidate \(B_s^0 \rightarrow D_s^{*-}\mu^+\nu_\mu\) decays are selected by pairing \(D_s^{*-}\) and \(\mu^+\) candidates, where the \(D_s^{*-}\) is reconstructed through the \(D_s^-\gamma\) decay. The \(D_s^-\) mesons are reconstructed requiring two opposite-sign kaons and a pion inconsistent with coming from a PV, and forming a common vertex that is displaced from every PV. The final-state hadrons and muon must satisfy strict particle identification (PID) criteria, consistent with the assigned particle hypothesis.

To suppress the combinatorial background in the \(D_s^-\) mass spectrum, only the regions of the \(D_s^- \rightarrow K^+K^-\pi^-\) Dalitz plane compatible with originating from \(\phi\pi^-\) and \(K^*0K^-\)
Figure 2: Distribution of the reconstructed $D_s^-\gamma$ mass, $m(D_s^-\gamma)$, with the fit overlaid. The fit is performed constraining the $D_s^-$ mass to the world-average value [16]. The signal and background components are shown separately with dashed red and dotted green lines, respectively.

are retained by requiring the $K^+K^-$ mass to be within 20 MeV/c$^2$ of the known $\phi$ mass, or the reconstructed $K^+\pi^-$ mass to be within 90 MeV/c$^2$ of the average $K^+(892)^0$ mass [16]. Possible backgrounds arising from the misidentification of one of the $D_s^-$ decay products are removed with explicit vetoes which apply more stringent PID requirements in a small window of invariant mass of the corresponding particle combination. The main contributions that are removed come from $\Lambda_{c}\to K^+p\pi^-$, $D_s^\to K^+\pi^\pi^-$, $D_s^\to K^-\pi^+\pi^-$, and misidentified or partially reconstructed multibody $D$ decays, all originating from semileptonic $b$-hadron decays.

Due to the small mass difference between the $D_s^{*-}$ and $D_s^{-}$ mesons, the photon must be emitted close to the $D_s^{-}$ flight direction. Photons are selected inside a narrow cone surrounding the $D_s^{-}$ candidate, defined in pseudorapidity and azimuthal angle. Only the highest $p_T$ photon inside the cone is combined with the $D_s^-$ candidate. Potential contamination from neutral pions reconstructed as a single merged cluster in the electromagnetic calorimeter is suppressed by employing a neural network classifier trained to separate $\pi^0$ mesons from photons [42].

A fit to the $D_s^-\gamma$ invariant-mass distribution, with the reconstructed $D_s^-$ mass constrained to the known value [16], is performed as shown in Fig. 2. The signal is described by a Gaussian function with a power-law tail on the right hand side of the distribution and the background by an exponential distribution. The power-law tail accounts for cases where additional activity in the calorimeter is mistakenly included in the photon cluster. The sPlot technique [13] is employed to subtract the combinatorial background from random photons. Weighted signal is used to create the templates described in Sec. 5. The correlation between the weights and $w$ is below 4%.

The muon candidate is required to have $p_T$ in excess of 1.2 GeV/c. Background arising from $b$-hadrons decaying into final states containing two charmed hadrons, $H_b\to D_s^{*-}H_c$,
Table 1: Binning scheme used for this measurement.

<table>
<thead>
<tr>
<th>bin</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w$</td>
<td>1.00–1.11</td>
<td>1.11–1.17</td>
<td>1.17–1.22</td>
<td>1.22–1.27</td>
<td>1.27–1.32</td>
<td>1.32–1.38</td>
<td>1.38–1.47</td>
</tr>
</tbody>
</table>

followed by a semileptonic decay of the charmed hadron $H_c \rightarrow \mu^+ \nu_{\mu} X$, where $X$ is one or more hadrons, are suppressed by using a multivariate algorithm based on the isolation of the muon [44]. Finally, the $B^0_s$ meson candidates are formed by combining $\mu^+$ and $D_s^{*-}$ candidates which are consistent with coming from a common vertex.

5 Signal yield

The signal yield is determined using a template fit to the distribution of the corrected mass [45],

$$m_{\text{corr}} = \sqrt{m_{D_s^{*-}\mu^+}^2 + |p_{\perp}|^2 + |p_{\perp}|},$$  
(12)

where $m_{D_s^{*-}\mu^+}$ is the measured mass of the $D_s^{*-}\mu^+$ candidate, and $p_{\perp}$ is the momentum of the candidate transverse to the $B^0_s$ flight direction. When only one massless final-state particle is missing from the decay, $m_{\text{corr}}$ peaks at the $B^0_s$ mass. Only candidates in the window $3500 < m_{\text{corr}} < 5367$ MeV/$c^2$ are considered.

Extended binned maximum-likelihood fits to the $m_{\text{corr}}$ distribution are performed independently in seven bins of the reconstructed hadronic recoil, $w$, to obtain the raw yields $N_{\text{meas}}$ per bin. The binning scheme, detailed in Tab. 1, is chosen such that each $w$ bin has roughly the same signal yield, based on simulation. Obtaining the value of $w$ requires the knowledge of the momentum of the $B^0_s$ meson, which in the decays under study can be solved up to a quadratic ambiguity. By imposing momentum balance against the visible system with respect to the flight direction, and assuming the mass of the $B^0_s$ meson, the momentum of the $B^0_s$ meson can be estimated. To resolve the ambiguity in the solutions, a multivariate regression algorithm based on the flight direction is used [46], yielding a purity on the solutions of around 70%. The $m_{\text{corr}}$ distribution is fitted using shapes (templates) of signal and of background distributions mostly obtained from simulation. These simulated decays are selected as described in Sec. 4, and are corrected for the simulation mismodeling as described in Sec. 3.

The largest contribution to the background is due to $B^0_s \rightarrow D_s^{*-}\tau^+\nu_{\tau}$ decays, with $\tau^- \rightarrow \mu^- \nu_{\mu}\nu_{\tau}$. A small source of background is formed by excited $D_s^{*-}$ mesons decaying into a $D_s^{*-}$ resonance. The only excited state decaying into $D_s^{*-}$ is the $D_{s1}(2460)^-$ meson, and hence templates for $B^0_s \rightarrow D_{s1}(2460)^-\mu^+\nu_{\mu}$ and $B^0_s \rightarrow D_{s1}(2460)^-\tau^+\nu_{\tau}$ decays are included in the fit. The background arising from $b$ hadrons decaying into final states containing two charmed hadrons, $H_b \rightarrow D_s^{*-}H_c$, is also addressed. The template for this process is generated using simulated events of $B^0_s$, $B^0$, $B^+$ and $\Lambda_b^0$ decays, with an appropriate admixture of final states, based on their production rates, branching ratios and relative reconstruction efficiencies taken from simulation. The last background considered in the fit is the combinatorial background, arising from random combinations of tracks. This template is obtained from a data sample where the $D_s^{*-}$ meson and the muon have the same charge.
The free parameters in the fit are the signal yield, the relative abundances of $B^0_s \rightarrow D_s^- \tau^+ \nu_\tau$ and $B^0_s \rightarrow D_{s1}(2460)^- \mu^+ \nu_\mu$ candidates with respect to that of the signal, and the fraction of combinatorial background. The total fraction of backgrounds from $H_c \rightarrow \mu^+ \nu_\mu X$ decays is fixed to the expected value using the measured branching fractions and selection efficiencies obtained from simulation. A 40% uncertainty is assigned to this component to account for the uncertainties on the branching fractions [16]. The $B^0_s \rightarrow D_{s1}(2460)^- \tau^+ \nu_\tau$ contribution is also fixed assuming a value of its ratio with respect to the muonic mode equal to the SM prediction for $B(B^+ \rightarrow D^{*-0} \tau^+ \nu_\tau)/B(B^+ \rightarrow D^{*0} \mu^+ \nu_\mu)$ [19] under the assumption that this ratio is identical for $B^0_s$ meson decays. The contribution of this decay to the fit is negligible. The Barlow-Beeston “lite” technique [47,48] is applied to account for the limited size of the simulation samples. The distributions of $m_{\text{corr}}$ with the fit overlaid are shown in Fig. 3.

Using the fractions obtained from the fit, data and simulated distributions of the angular variables $\cos(\theta_\mu)$, $\cos(\theta_{D_s})$, and $\chi$, as defined in Sec. 2, are shown in Fig. 4. All distributions show good agreement between data and simulation, justifying performing the measurement of the differential decay rate only as a function of $w$.

6 Efficiency correction

This analysis requires a precise measurement of all contributions to the efficiency as a function of the true value of the hadronic recoil $w_{\text{true}}$ extracted from simulation. However, the overall normalisation of the efficiency is not determined as only its dependency with $w_{\text{true}}$ is relevant.

The total efficiency is the product of the geometrical acceptance of the detector, the efficiency of reconstructing all tracks, the trigger requirements, and the full set of kinematic, PID and background rejection requirements. Most of the contributions to the total efficiency are obtained using simulation. Only the particle identification and the $D^-$ selection efficiencies are derived from data using control samples. The muon and hadron PID efficiencies are taken from large data samples of $J/\psi \rightarrow \mu^+ \mu^-$ and $D^{*-} \rightarrow D^{*0} \pi^+$ decays, respectively [49]. These samples are then used to determine the PID efficiencies in bins of $p$, $p_T$ and number of tracks in the event. The $D^-_s$ selection efficiency accounts for selecting the regions in the Dalitz plane, as well as the vetoes described in Sec. 4. This efficiency is determined from a sample of fully reconstructed $B^0_s \rightarrow D^-_s \pi^+$ decays as a function of the $D^-_s$ meson $p_T$. The efficiencies extracted from data are convolved with the simulation to obtain their dependency on $w_{\text{true}}$.

The efficiencies derived from simulation are extracted by comparing the generator-level simulation, based on PYTHIA [33] and EvtGen [35], to the final reconstructed and selected simulation sample used for the template fit, omitting the particle identification and the $D^-_s$ selection criteria.

7 Form factor fits

The measured $B^0_s \rightarrow D_s^- \mu^+ \nu_\mu$ spectrum from Sec. 5 must be unfolded to account for the resolution which is 0.07 in the $w$ variable. The unfolding procedure uses a migration matrix determined from simulation, defined as the probability that a candidate generated in bin $j$ of the $w_{\text{true}}$ distribution appears in bin $i$ of the $w$ distribution. The unfolded spectrum is
Figure 3: Distribution of the corrected mass, $m_{\text{corr}}$, for the seven bins of $w$, overlaid with the fit results. The $B_s^0 \to D_{s1}(2460)^- \tau^+ \nu_{\tau}$ and the $B_s^0 \to D_{s1}(2460)^- \mu^+ \nu_{\mu}$ components are combined in $B_s^0 \to D_{s1}(2460)^- \ell^+ \nu_{\ell}$. Below each plot, differences between the data and fit are shown, normalised by the uncertainty in the data.
then corrected bin-by-bin using the efficiency described in Sec. 6. The combination of the migration matrix and the total efficiency, called the response matrix, is shown in App. D.

The unfolding procedure adopted is based on the singular value decomposition (SVD) method [50] using the RooUnfold package [51]. The SVD method includes a regularisation procedure that depends upon a parameter $k$, ranging between unity and the number of degrees of freedom, seven in this case. Using simulation, the optimal value for $k$ is found to be $k = 5$, which minimises the difference between the yield from the unfolding procedure and the expected yield in each bin. The final yields are normalised to unity.

The values of the form-factor parameters are derived from a $\chi^2$ fit with

$$\chi^2 = \sum_{i,j} \left( N_{\text{corr},i}^{\text{unf}} - N_{\text{exp},i} \right) C_{ij}^{-1} \left( N_{\text{corr},j}^{\text{unf}} - N_{\text{exp},j} \right).$$

In this expression, $N_{\text{corr},i}^{\text{unf}}$ is the normalised, unfolded and efficiency-corrected yield in bin $i(j)$, $N_{\text{exp},i}$ is the expected yield in bin $i(j)$ obtained from integrating $d\Gamma_{i(j)}/dw$ from the CLN or BGL parametrisation over the bin, and $C_{ij}$ is the covariance matrix describing the statistical uncertainties from the yields and efficiency corrections. This $\chi^2$ function is minimised for the CLN and BGL parametrisations separately. The unitarity constraint for the BGL parametrisation is considered in the minimisation by adding a penalty.
function to the $\chi^2$ defined in Eq. 13. This function resembles a step function by raising the unitarity constraint to the power of 20: $\sum_{n=0}^{\infty} (a_n^f)^2 + \sum_{n=0}^{\infty} (a_n^F)^2 \cdot 20$. For the CLN parametrisation, the fitted value is $\rho^2 = 1.16 \pm 0.05$, while for the BGL parametrisation, the fitted values are $a_1^f = -0.002 \pm 0.034$, and $a_2^f = 0.93^{+0.05}_{-0.20}$, where the uncertainties are only statistical in nature.

8 Systematic uncertainties

Systematic uncertainties on the form-factor parameters and $N_{\text{corr}}^{\text{unf}}$ originate from the fitted $D_s^- \to \rho K$ and $B_s^0$ yields, the efficiency corrections and the form-factor fit. They are determined on the normalised and efficiency-corrected yields, as well as on the parameters $\rho^2$, $a_1^f$ and $a_2^f$. Their impact on the form-factor fits is assessed by repeating the fit with different conditions and comparing the obtained values to the nominal values. A summary of all systematic uncertainties for both the CLN and BGL parametrisations is shown in Tab. 2.

The size of the simulated samples, which are very CPU intensive to generate, is the dominating systematic uncertainty on the form-factor parameters. The simulated sample size is accounted for in the fit by applying the Barlow-Beeston “lite” technique [47,48] when determining the signal yield. Its relative contribution to the systematic uncertainty is assessed by not applying this technique and comparing the obtained uncertainties. The uncertainties due to the size of the control samples used to determine the efficiencies and corrections are obtained by varying each of the efficiency and correction inputs within their uncertainty, repeating this 1000 times, and taking the spread as the uncertainty on the form-factor parameters or $N_{\text{corr}}^{\text{unf}}$.

The uncertainty on the SVD unfolding procedure is determined by repeating the regularisation procedure with a different regularisation parameter, $k$. The nominal value used is $k = 5$, which is changed to $k = 4$ and $k = 6$, and the difference with the nominal value is assigned as the systematic uncertainty.

Two systematic uncertainties are determined to account for assumptions in the simulation. Radiative corrections simulated by the PHOTOS package are known to be incomplete [36,52]. The difference in the form factor measured from simulated samples with and without PHOTOS is evaluated and a third of the difference is assigned following Ref. [53]. The efficiency due to the detector acceptance, and thus the shape of the efficiency correction, may be affected by the form factors in the HQET model used to generate the simulation, which are based on the 2016 HFLAV averages [54]. This is studied by weighting both the generator level and fully reconstructed simulated samples to the 2019 HFLAV averages [54]: $\rho^2 = 1.122 \pm 0.024$, $R_1(1) = 1.270 \pm 0.026$, and $R_2(1) = 0.852 \pm 0.018$, with correlations $\text{corr}[\rho^2, R_1(1)] = -0.824$, $\text{corr}[\rho^2, R_2(1)] = 0.566$, and $\text{corr}[R_1(1), R_2(1)] = -0.715$. The values of each pair are varied within one standard deviation of their mean, taking into account their correlation. The value of $R_0(1)$ is varied by a 20% uncertainty accounting for finite $b$- and $c$-quark masses [19]. These variations result in small changes of the total efficiency and the average difference is taken as the uncertainty.

The trigger corrections applied to the simulated samples depend on the kinematics of the candidates. To estimate the effect of the choice of the binning scheme used to make these corrections a different binning scheme is used and the corrections re-evaluated. Moreover, the impact of the detector occupancy is assessed by adding the number of
Table 2: Summary of the systematic and statistical uncertainties on the parameters $\rho^2$, $a_f^1$ and $a_f^2$ from the unfolded CLN and BGL fits. The total systematic uncertainty is obtained by adding the individual components in quadrature.

<table>
<thead>
<tr>
<th>Source</th>
<th>$\sigma(\rho^2)$</th>
<th>$\sigma(a_f^1)$</th>
<th>$\sigma(a_f^2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation sample size</td>
<td>0.053</td>
<td>0.036</td>
<td>+0.04 -0.35</td>
</tr>
<tr>
<td>Sample sizes for efficiencies and corrections</td>
<td>0.020</td>
<td>0.016</td>
<td>+0.02 -0.16</td>
</tr>
<tr>
<td>SVD unfolding regularisation</td>
<td>0.008</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>Radiative corrections</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulation FF parametrisation</td>
<td>0.007</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>Kinematic weights</td>
<td>0.024</td>
<td>0.013</td>
<td></td>
</tr>
<tr>
<td>Hardware-trigger efficiency</td>
<td>0.001</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>Software-trigger efficiency</td>
<td>0.004</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>$D_s^-$ selection efficiency</td>
<td></td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>$D_s^{*-}$ weights</td>
<td>0.002</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>External parameters in fit</td>
<td>0.024</td>
<td>0.002</td>
<td>0.04</td>
</tr>
<tr>
<td>Total systematic uncertainty</td>
<td>0.068</td>
<td>0.046</td>
<td>+0.06 -0.38</td>
</tr>
<tr>
<td>Statistical uncertainty</td>
<td>0.052</td>
<td>0.034</td>
<td>+0.05 -0.20</td>
</tr>
</tbody>
</table>

tracks reconstructed in each event as a binning variable. The systematic uncertainty due to the selection of muons is estimated by changing the PID requirements of the control sample. The effect of the $B_s^0$ and $\gamma$ kinematic corrections is also assessed by changing the weighting schemes to include more bins in $p$ and $p_T$. The possible systematic uncertainty due to the kinematic dependence of the $D_s^-$ selection efficiency is assessed by extracting the efficiency as a function of $p$ instead of $p_T$ from the $B_s^0 \rightarrow D_s^- \pi^+$ control sample.

The systematic uncertainty due to the photon background subtraction, performed through the sPlot method with fits to the $D_s^{*-}$ invariant mass, is assessed by implementing the fit with a third-order Chebyshev polynomial for the background description, and repeating the background subtraction process.

In the form-factor fit, the parameters $R_1(1)$ and $R_2(1)$ are fixed to the HFLAV averages \cite{6}. The uncertainties on these values are propagated to the CLN fit outcome by changing $R_1(1)$ and $R_2(1)$ within one standard deviation from their average, while accounting for the correlation. Since this uncertainty is related to the fit parametrisation only, it is not included as an uncertainty on the fit yields. For the BGL fit, the values of the external parameters of the $f(z)$, $g(z)$ and $F_1(z)$ functions are varied simultaneously within their uncertainty. When the uncertainties are asymmetric the largest is chosen. This process is repeated 1000 times applying the unitarity constrain and the difference between the average of the variations and the nominal value is assigned as a systematic uncertainty.

Systematic uncertainties induced by the tracking corrections, detector occupancy and PID efficiencies are found to be negligible as they do not affect the corrected mass distribution nor the shape of the efficiency correction.
9 Results and conclusions

A measurement of the leading parameters of the form factor describing the semileptonic transition \( B_0^s \to D_{s}^{*-} \mu^+ \nu_\mu \) has been performed. Using the CLN parametrisation the result obtained is

\[
\rho^2 = 1.16 \pm 0.05 \text{ (stat)} \pm 0.07 \text{ (syst)},
\]

where the mass of the muon has not been neglected. To compare with other published results, the fit is repeated assuming a massless muon, resulting in a small shift of the central value of the \( \rho^2 \) parameter of about 1.5%, as shown in Tab. 3. The world-average value of \( \rho^2 \) for the equivalent \( B^0 \to D^{*-} \mu^- \nu_\mu \) decay is \( \rho^2 = 1.122 \pm 0.015 \text{ (stat)} \pm 0.019 \text{ (syst)} \) \cite{6}. The results agree as expected assuming SU(3) symmetry. The measurement is also in agreement with the value obtained in Ref. \cite{7}, \( \rho^2 = 1.23 \pm 0.17 \text{ (stat)} \pm 0.05 \text{ (syst)} \pm 0.01 \text{ (ext)} \), where the last uncertainty comes from external inputs. That analysis uses \( B^0_s \to D_{s}^{*-} \mu^+ \nu_\mu \) decays from an independent data set, and where the photon from the \( D_{s}^{*-} \) decay is not reconstructed. A comparison with the normalised \( \Delta \Gamma/\Delta w \) spectra inferred from the CLN and BGL parametrisations in Ref. \cite{7} gives consistent results with the measured \( w \) spectrum in this paper, which is shown in App. C.

Using the BGL parametrisation, the results obtained are

\[
\begin{align*}
a^f_1 &= -0.002 \pm 0.034 \text{ (stat)} \pm 0.046 \text{ (syst)}, \\
a^f_2 &= 0.93^{+0.05}_{-0.20} \text{ (stat)} ^{+0.06}_{-0.38} \text{ (syst)}. 
\end{align*}
\]

In Fig. 5 the \( \Delta \chi^2 \) contours for \( a^f_1 \) versus \( a^f_2 \) are shown. The unitarity constraint results in a non-gaussian distribution of the uncertainty on the \( a^f_2 \) parameter. The fits to the differential decay rate using both parametrisations are shown in Fig. 6. The fit probabilities are 8.2% and 1.3% for the CLN and BGL parametrisations, respectively. The low values of the probabilities are caused by the third bin in \( w \), which is higher than expected in both CLN and BGL parametrisations.

The unfolded spectrum as a function of \( w \) with the systematic uncertainty per bin is given in Tab. 4. The correlations between bins are given in Tab. 5 and the covariance matrix is presented in Tab. 7, both in App. D.

The prediction of the decay rate can also be transformed to a prediction of the expected event yields taking into account the efficiency and resolution, which then is fit to the experimental spectrum. Both procedures provide similar results with small differences induced by slightly different bin-by-bin correlations as shown in Tab. 3. The detector response combined with the reconstruction efficiency is presented in App. D.

In conclusion, this paper presents for the first time the unfolded normalised differential decay rate for \( B^0_s \to D_{s}^{*-} \mu^+ \nu_\mu \) decays as a function of the recoil parameter \( w \), which can be compared directly to theoretical predictions. The form-factor parameters using the CLN and BGL parametrisations are also presented. Both parametrisations give consistent results when compared to data.

Acknowledgements

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at the
Figure 5: $\Delta\chi^2$ contours for $a_1^f$ versus $a_2^f$. The black cross marks the best-fit central value. The solid (dashed) contour encloses the $\Delta\chi^2 = 2.3$ (6.17) region. The observed shape is due to the applied unitarity condition, see Eq. (11).

Figure 6: Unfolded normalised differential decay rate with the fit superimposed for the CLN parametrisation (green), and BGL (red). The band in the fit results includes both the statistical and systematic uncertainty on the data yields.
Table 3: Results from different fit configurations, where the first uncertainty is statistical and the second systematic.

<table>
<thead>
<tr>
<th>Fit Configuration</th>
<th>Result</th>
<th>Statistical Uncertainty</th>
<th>Systematic Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLN fit</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unfolded fit</td>
<td>$\rho^2 = 1.16 \pm 0.05 \pm 0.07$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unfolded fit with massless leptons</td>
<td>$\rho^2 = 1.17 \pm 0.05 \pm 0.07$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Folded fit</td>
<td>$\rho^2 = 1.14 \pm 0.04 \pm 0.07$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BGL fit</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unfolded fit</td>
<td>$a_1^f = -0.002 \pm 0.034 \pm 0.046$</td>
<td>$a_2^f = 0.93^{+0.05+0.06}_{-0.20-0.38}$</td>
<td></td>
</tr>
<tr>
<td>Folded fit</td>
<td>$a_1^f = 0.042 \pm 0.029 \pm 0.046$</td>
<td>$a_2^f = 0.93^{+0.05+0.06}_{-0.20-0.38}$</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Fraction of the unfolded yields corrected for the global efficiencies, $N_{\text{corr}}^{\text{unf}}$, for each $w$ bin. Also shown in this table is the breakdown of the systematic and statistical uncertainties on $N_{\text{corr}}^{\text{unf}}$. These are shown as a fraction of the unfolded yield.

<table>
<thead>
<tr>
<th>$w$ bin</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction of $N_{\text{corr}}^{\text{unf}}$</td>
<td>0.183</td>
<td>0.144</td>
<td>0.148</td>
<td>0.128</td>
<td>0.117</td>
<td>0.122</td>
<td>0.158</td>
</tr>
<tr>
<td>Uncertainties (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulation sample size</td>
<td>3.5</td>
<td>3.0</td>
<td>2.8</td>
<td>3.1</td>
<td>3.4</td>
<td>3.0</td>
<td>3.7</td>
</tr>
<tr>
<td>Sample sizes for effs and corrections</td>
<td>3.6</td>
<td>3.2</td>
<td>3.0</td>
<td>2.8</td>
<td>2.8</td>
<td>2.7</td>
<td>2.8</td>
</tr>
<tr>
<td>SVD unfolding regularisation</td>
<td>0.5</td>
<td>0.5</td>
<td>0.1</td>
<td>0.7</td>
<td>1.2</td>
<td>0.0</td>
<td>0.5</td>
</tr>
<tr>
<td>Radiative corrections</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.3</td>
<td>0.4</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Simulation FF parametrisation</td>
<td>0.3</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.4</td>
<td>0.2</td>
</tr>
<tr>
<td>Kinematic weights</td>
<td>2.4</td>
<td>1.0</td>
<td>1.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.9</td>
</tr>
<tr>
<td>Hardware-trigger efficiency</td>
<td>0.3</td>
<td>0.3</td>
<td>0.0</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
<td>0.1</td>
</tr>
<tr>
<td>Software-trigger efficiency</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>$D_s^-$ selection efficiency</td>
<td>0.5</td>
<td>0.2</td>
<td>0.3</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
<td>0.3</td>
</tr>
<tr>
<td>$D_s^{*-}$ weights</td>
<td>0.0</td>
<td>2.3</td>
<td>0.8</td>
<td>2.9</td>
<td>2.0</td>
<td>0.9</td>
<td>0.4</td>
</tr>
<tr>
<td>Total systematic uncertainty</td>
<td>5.6</td>
<td>5.1</td>
<td>4.4</td>
<td>5.2</td>
<td>5.0</td>
<td>4.2</td>
<td>4.8</td>
</tr>
<tr>
<td>Statistical uncertainty</td>
<td>3.4</td>
<td>2.9</td>
<td>2.7</td>
<td>3.1</td>
<td>3.2</td>
<td>2.9</td>
<td>3.4</td>
</tr>
</tbody>
</table>
Table 5: Correlation matrix for the unfolded data set in bins of $w$, including both statistical and systematic uncertainties.

<table>
<thead>
<tr>
<th>$w$ bin</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.44</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.13</td>
<td>0.60</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.19</td>
<td>0.32</td>
<td>0.48</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.30</td>
<td>0.30</td>
<td>0.15</td>
<td>0.60</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.34</td>
<td>0.38</td>
<td>0.33</td>
<td>0.22</td>
<td>0.54</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.27</td>
<td>0.34</td>
<td>0.34</td>
<td>0.27</td>
<td>0.07</td>
<td>0.32</td>
<td>1</td>
</tr>
</tbody>
</table>

and NCN (Poland); MEN/IFA (Romania); MSHE (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); DOE NP and NSF (USA). We acknowledge the computing resources that are provided by CERN, IN2P3 (France), KIT and DESY (Germany), INFN (Italy), SURF (Netherlands), PIC (Spain), GridPP (United Kingdom), RRCKI and Yandex LLC (Russia), CSCS (Switzerland), IFIN-HH (Romania), CBPF (Brazil), PL-GRID (Poland) and OSC (USA). We are indebted to the communities behind the multiple open-source software packages on which we depend. Individual groups or members have received support from AvH Foundation (Germany); EPLANET, Marie Skłodowska-Curie Actions and ERC (European Union); ANR, Labex P2IO and OCEVU, and Région Auvergne-Rhône-Alpes (France); Key Research Program of Frontier Sciences of CAS, CAS PIFI, and the Thousand Talents Program (China); RFBR, RSF and Yandex LLC (Russia); GVA, XuntaGal and GENCAT (Spain); the Royal Society and the Leverhulme Trust (United Kingdom).
Appendices

A  Fitted yields and efficiency

Figure 7 shows the total efficiency that is applied to the unfolded signal yields, as a function of \( w_{\text{true}} \). This is the combination of the reconstruction and selection efficiencies, including the acceptance of the LHCb detector.

![Figure 7](image)

Figure 7: Total efficiency as a function of \( w_{\text{true}} \), including the acceptance of the LHCb detector as well as the reconstruction and selection efficiencies.

B  Inputs for BGL fit

Table 6 gives an overview of the fit inputs for the BGL fit.

C  Comparison with LHCb-PAPER-2019-041

The \( w \) spectrum measured in this analysis can be compared with the results obtained in Ref. [7]. In Ref. [7], the form-factor parameters of the \( B^0 \rightarrow D_s^{-} \mu^+ \nu_{\mu} \) decay are measured using a version of the CLN and BGL parametrisations. From this, the normalised \( \Delta \Gamma / \Delta w \) spectrum can be inferred, which is shown in Fig. 8. The spectrum measured in this paper is consistent with the normalised spectra inferred from both CLN and BGL parametrisations used in Ref. [7].
Table 6: Fit inputs used for the BGL fit, taken from Ref. [17] and Ref. [20].

<table>
<thead>
<tr>
<th>BGL parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a'_0$</td>
<td>$0.01221 \pm 0.00016$</td>
</tr>
<tr>
<td>$a'_1$</td>
<td>$0.0042 \pm 0.0022$</td>
</tr>
<tr>
<td>$a'_2$</td>
<td>$-0.069^{+0.041}_{-0.037}$</td>
</tr>
<tr>
<td>$a_0^g$</td>
<td>$0.024^{+0.021}_{-0.009}$</td>
</tr>
<tr>
<td>$a_1^g$</td>
<td>$0.05^{+0.39}_{-0.72}$</td>
</tr>
<tr>
<td>$a_2^g$</td>
<td>$1.0^{+0.0}_{-2.0}$</td>
</tr>
<tr>
<td>$a_0^{F_2}$</td>
<td>$0.0595 \pm 0.0093$</td>
</tr>
<tr>
<td>$a_1^{F_2}$</td>
<td>$-0.318 \pm 0.170$</td>
</tr>
</tbody>
</table>

Figure 8: Comparison between the $w$ spectrum measured in this paper to the normalised $\Delta \Gamma/\Delta w$ spectra inferred from the CLN and BGL parametrisations in Ref. [7].

D Covariance and response matrices

This section contains the information needed to reproduce a form-factor fit using, for example, different fit parametrisations. To perform the fit using the unfolded, efficiency-corrected and normalised yields given in Tab. 4, the corresponding covariance matrix with the combined statistical uncertainties is given in Tab. [7].
Table 7: Covariance matrix for the unfolded data set in bins of $w$, including both statistical and systematic uncertainties in units of $10^{-5}$.

<table>
<thead>
<tr>
<th>$w$ bin [$10^{-5}$]</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4.73 7.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.21 3.81 5.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.87 2.12 2.81 6.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2.74 1.80 0.78 3.37 5.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2.42 1.82 1.38 0.98 2.17 3.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>3.24 2.69 2.43 2.02 0.44 1.69 8.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8: Response matrix, containing the migration from $w_{\text{true}}$ to $w$ bins together with the total efficiency in units of $10^{-4}$.

<table>
<thead>
<tr>
<th>$[10^{-4}]$</th>
<th>$w_{\text{true}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w$</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>132.0</td>
</tr>
<tr>
<td>2</td>
<td>22.4</td>
</tr>
<tr>
<td>3</td>
<td>6.0</td>
</tr>
<tr>
<td>4</td>
<td>4.6</td>
</tr>
<tr>
<td>5</td>
<td>1.4</td>
</tr>
<tr>
<td>6</td>
<td>0.8</td>
</tr>
<tr>
<td>7</td>
<td>−0.1</td>
</tr>
</tbody>
</table>

To transform theoretical predictions into expected signal yields, the response matrix, given in Tab. 8 is needed. This contains the migration matrix (from the true value of $w$ to the reconstructed one) combined with the reconstruction efficiency. The migration matrix is normalised such that the entries within a given bin of $w$ sum up to unity. The absolute efficiencies have not been measured for this analysis.
References

[14] P. Colangelo and F. De Fazio, Scrutinizing $\bar{B}\to D^*(D\pi)\ell^-\nu_\ell$ and $\bar{B}\to D^*(D\gamma)\ell^-\nu_\ell$ in search of new physics footprints, JHEP 06 (2018) 082, arXiv:1801.10468.
[18] M. Bordone, N. Gubernari, D. van Dyk, and M. Jung, Heavy-Quark expansion for $\bar{B}_s \to D_s^{(*)}$ form factors and unitarity bounds beyond the $SU(3)_F$ limit, arXiv:1912.09335.


[23] Belle collaboration, A. Abdesselam et al., Precise determination of the CKM matrix element $|V_{cb}|$ with $\bar{B}^0 \to D^{*-}\ell^+\nu_\ell$ decays with hadronic tagging at Belle, arXiv:1702.01521.


[53] BaBar collaboration, B. Aubert et al., *Measurements of the semileptonic decays $\bar{B} \to D\ell\bar{\nu}$ and $\bar{B} \to D^*\ell\bar{\nu}$ using a global fit to $DX\ell\bar{\nu}$ final states*, Phys. Rev. D79 (2009) 012002, arXiv:0809.0828.

LHCb collaboration

L. Xu\textsuperscript{3}, M. Xu\textsuperscript{7}, Q. Xu\textsuperscript{5}, Z. Xu\textsuperscript{4}, Z. Yang\textsuperscript{3}, Z. Yang\textsuperscript{65}, Y. Yao\textsuperscript{67}, L.E. Yeomans\textsuperscript{59}, H. Yin\textsuperscript{7}, J. Yu\textsuperscript{1,oa}, X. Yuan\textsuperscript{67}, O. Yushchenko\textsuperscript{43}, K.A. Zarebski\textsuperscript{15,c}, M. Zavertyaev\textsuperscript{15,e}, M. Zdybai\textsuperscript{53}, M. Zeng\textsuperscript{3}, D. Zhang\textsuperscript{3}, S. Zhang\textsuperscript{4}, W.C. Zhang\textsuperscript{3,z}, Y. Zhang\textsuperscript{47}, A. Zhelezov\textsuperscript{16}, Y. Zheng\textsuperscript{5}, X. Zhou\textsuperscript{5}, Y. Zhou\textsuperscript{5}, X. Zhu\textsuperscript{3}, V. Zhukov\textsuperscript{13,39}, J.B. Zonneveld\textsuperscript{57}, S. Zucchelli\textsuperscript{19,e}.

\textsuperscript{1}Centro Brasileiro de Pesquisas Físicas (CBPF), Rio de Janeiro, Brazil
\textsuperscript{2}Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil
\textsuperscript{3}Center for High Energy Physics, Tsinghua University, Beijing, China
\textsuperscript{4}School of Physics State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing, China
\textsuperscript{5}University of Chinese Academy of Sciences, Beijing, China
\textsuperscript{6}Institute Of High Energy Physics (IHEP), Beijing, China
\textsuperscript{7}Institute of Particle Physics, Central China Normal University, Wuhan, Hubei, China
\textsuperscript{8}Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IN2P3-LAPP, Annecy, France
\textsuperscript{9}Université Clermont Auvergne, CNRS/IN2P3, LPC, Clermont-Ferrand, France
\textsuperscript{10}Aix Marseille Univ, CNRS/IN2P3, CPPM, Marseille, France
\textsuperscript{11}Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay, France, Orsay, France
\textsuperscript{12}LPNHE, Sorbonne Université, Paris Diderot Sorbonne Paris Cité, CNRS/IN2P3, Paris, France
\textsuperscript{13}I. Physikalisches Institut, RWTH Aachen University, Aachen, Germany
\textsuperscript{14}Physikalisches Institut, Technische Universität Dortmund, Dortmund, Germany
\textsuperscript{15}Max-Planck-Institut für Kernphysik (MPIK), Heidelberg, Germany
\textsuperscript{16}Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany
\textsuperscript{17}School of Physics, University College Dublin, Dublin, Ireland
\textsuperscript{18}INFN Sezione di Bari, Bari, Italy
\textsuperscript{19}INFN Sezione di Bologna, Bologna, Italy
\textsuperscript{20}INFN Sezione di Ferrara, Ferrara, Italy
\textsuperscript{21}INFN Sezione di Firenze, Firenze, Italy
\textsuperscript{22}INFN Laboratori Nazionali di Frascati, Frascati, Italy
\textsuperscript{23}INFN Sezione di Genova, Genova, Italy
\textsuperscript{24}INFN Sezione di Milano-Bicocca, Milano, Italy
\textsuperscript{25}INFN Sezione di Milano, Milano, Italy
\textsuperscript{26}INFN Sezione di Padova, Padova, Italy
\textsuperscript{27}INFN Sezione di Pisa, Pisa, Italy
\textsuperscript{28}INFN Sezione di Roma Tor Vergata, Roma, Italy
\textsuperscript{29}INFN Sezione di Roma La Sapienza, Roma, Italy
\textsuperscript{30}NIKHEF National Institute for Subatomic Physics, Amsterdam, Netherlands
\textsuperscript{31}NIKHEF National Institute for Subatomic Physics and VU University Amsterdam, Amsterdam, Netherlands
\textsuperscript{32}Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest-Magurele, Romania
\textsuperscript{33}Petersburg Nuclear Physics Institute NRC Kurchatov Institute (PNPI NRC KI), Gatchina, Russia
\textsuperscript{34}Institute of Theoretical and Experimental Physics NRC Kurchatov Institute (ITEP NRC KI), Moscow, Russia
\textsuperscript{35}Budker Institute of Nuclear Physics (SB RAS), Novosibirsk, Russia
\textsuperscript{36}Yandex School of Data Analysis, Moscow, Russia
\textsuperscript{37}Institute of Nuclear Physics, Moscow State University (SINP MSU), Moscow, Russia
\textsuperscript{38}Institute for Nuclear Research of the Russian Academy of Sciences (INR RAS), Moscow, Russia
\textsuperscript{39}ICCUB, Universitat de Barcelona, Barcelona, Spain
\textsuperscript{40}Instituto Galego de Física de Altas Enerxías (IGFAE), Universidade de Santiago de Compostela, Santiago de Compostela, Spain

27
\textsuperscript{t}Scuola Normale Superiore, Pisa, Italy
\textsuperscript{u}Università di Modena e Reggio Emilia, Modena, Italy
\textsuperscript{v}Università di Siena, Siena, Italy
\textsuperscript{w}MSU - Iligan Institute of Technology (MSU-IIT), Iligan, Philippines
\textsuperscript{x}Novosibirsk State University, Novosibirsk, Russia
\textsuperscript{y}INFN Sezione di Trieste, Trieste, Italy
\textsuperscript{z}School of Physics and Information Technology, Shaanxi Normal University (SNNU), Xi’an, China
\textsuperscript{aa}Physics and Micro Electronic College, Hunan University, Changsha City, China
\textsuperscript{ab}Universidad Nacional Autonoma de Honduras, Tegucigalpa, Honduras