Measurement of the time-integrated mixing probability $\chi_d$ with a semileptonic double-tagging strategy and 34.6 fb$^{-1}$ of Belle II collision data

(The Belle II Collaboration)

Abstract

We present a first measurement of the time-integrated mixing probability $\chi_d$ using Belle II collision data corresponding to the integrated luminosity of 34.6 fb$^{-1}$. We reconstruct pairs of B mesons both decaying in semileptonic final states. Using a novel extraction methodology, we measure $0.192 \pm 0.010$ (stat.) $\pm 0.016$ (syst.), which is compatible with existing indirect and direct determinations.
1. INTRODUCTION

The properties guiding the time evolution of the neutral $B^0 - \bar{B}^0$ meson system has been studied by a great number of experiments. $B^0 - \bar{B}^0$ mixing was first reported by the ARGUS experiment in Ref. [1] and the observed level of mixing provided first indications for the mass scale of the top quark. When produced in a $\Upsilon(4S)$ decay, the two $B^0$ mesons evolve in a coherent $P$-wave state. For this reason the neutral $B$ mesons’ flavour in the coherent quantum state always strictly maintain opposite flavors. As such, the $B$ mesons’ flavour can only be determined relative to each other as soon as one of the mesons decays.

The mixing properties of the $B^0 - \bar{B}^0$ system can be described by four parameters ($x_d$, $y_d$, $q$, and $p$) and its time evolution is described by the Schrödinger equation, which depends on the relative time difference between the two neutral $B$ meson decays. The heavy ($H$) and light ($L$) mass eigenstates of the system are related to the $B^0$ and $\bar{B}^0$ flavor eigenstates by the equation:

$$|B_{L/H}\rangle = p|B^0_d\rangle \pm q|\bar{B}^0_d\rangle. \quad (1)$$

In the absence of $CP$ violation in mixing unitarity is conserved, $|q/p| = 1$, and we can express $x_p$ and $y_p$ as a function of the mass difference, $\Delta m_d = m_H - m_L$, and the lifetime difference, $\Delta \Gamma_d = \Gamma_L - \Gamma_H$ of the two mass eigenstates:

$$x_d = \Delta m_d/\Gamma_d, \quad y_d = \Delta \Gamma_d/\Gamma_d, \quad (2)$$

with $\Gamma_d = (\Gamma_L + \Gamma_H)/2$ the average decay width. Experimentally, both parameters can be constrained by measuring the time-integrated mixing probability,

$$\chi_d = \frac{\Gamma(B^0 \to \bar{B}^0)}{\Gamma(B^0 \to B^0) + \Gamma(B^0 \to \bar{B}^0)} = \frac{x_d^2 + y_d^2}{2(x_d^2 + 1)}, \quad (3)$$

and the value of $y_d$ can be determined in combination with direct measurements of $\Delta m_d$ and the $B$ meson lifetime. The current most precise value of $\chi_d$ is obtained by combining the information from time-independent and time-dependent measurements: $\chi_d^{WA} = 0.186 \pm 0.001\ [2]$. The world average only using time-independent measurements has a much larger uncertainty and results in $\chi_d^{WA\ t-independent} = 0.182 \pm 0.015$.

In this conference note, we provide a first direct determination of $\chi_d$ using a time-independent approach and semileptonic $B \to X e^+ \nu_e$ decays, recorded by the Belle II data taking campaigns in 2019 and 2020. In total, we analyze an integrated luminosity of 34.6 fb$^{-1}$ of recorded collision data, corresponding to $(37.7 \pm 0.6) \times 10^6$ of $B$ meson pairs. We identify events in which both $B$ mesons decayed via a semileptonic $B \to X e^+ \nu_e$ transition. The value of $\chi_d$ is obtained by determining the number of $e^\pm e^\pm$ same-sign ($N_{SS}$) and $e^\pm e^\mp$ opposite-sign ($N_{OS}$) signal candidates, as the lepton charge of the semileptonic transition encodes directly the flavor of the $B$ meson. Contributions from charged $B$ mesons are subtracted using a correction factor $r_B$ and the number of opposite- and same-sign events are corrected for selection and acceptance effects to determine

$$\chi_d = \frac{N_{SS}}{N_{SS} + N_{OS} \cdot (\epsilon_{OS}/\epsilon_{SS})^{-1} \cdot (1 + r_B)} \cdot (1 + r_B) \cdot (1 + r_B). \quad (4)$$
Here, $\epsilon_{OS}$ and $\epsilon_{SS}$ denote the efficiency for opposite-sign and same-sign signal, respectively, with $(\epsilon_{OS}/\epsilon_{SS}) = 0.92 \pm 0.01$ as derived from studies on simulated samples. Further, $r_B = f_{+0} \cdot \tau_{+0}^2 = 1.2 \pm 0.1$ with $\tau_{+0} = 1.078 \pm 0.004$ denoting the charged and neutral $B$ meson lifetime ratio and $f_{+0} = B(\Upsilon(4S) \to B^+ B^0)/B(\Upsilon(4S) \to B^0 B^0) = 1.058 \pm 0.024$ [2].

2. **THE BELLE II DETECTOR AND DATA SAMPLE**

The Belle II detector [3, 4] operates at the SuperKEKB asymmetric-energy electron-positron collider [5], located at the KEK laboratory in Tsukuba, Japan. The detector consists of several nested detector subsystems arranged around the beam pipe in a cylindrical geometry. The innermost subsystem is the vertex detector, which includes two layers of silicon pixel detectors and four outer layers of silicon strip detectors. Currently, the second pixel layer is installed in only a small part of the solid angle, while the remaining vertex detector layers are fully installed. Most of the tracking volume consists of a small-cell drift chamber filled with a helium and ethane mixture gas. Outside the drift chamber, a Cherenkov-light imaging and time-of-propagation detector provides charged-particle identification in the barrel region. In the forward endcap, this function is provided by a proximity-focusing, ring-imaging Cherenkov detector with an aerogel radiator. Further out is an electromagnetic calorimeter, consisting of a barrel and two endcap sections made of CsI(Tl) crystals. A uniform 1.5 T magnetic field is provided by a superconducting solenoid situated outside the calorimeter. Multiple layers of scintillators and resistive plate chambers, located between the magnetic flux-return iron plates, constitute the $K_L$ and muon identification system.

The collision data used in this analysis were collected at a center-of-mass (CM) energy of 10.58 GeV, corresponding to the mass of the $\Upsilon(4S)$ resonance. The energies of the electron and positron beams are 7 GeV and 4 GeV, respectively, resulting in a boost of $\beta \gamma = 0.28$ of the CM frame relative to the laboratory frame. In addition, 3.2 fb$^{-1}$ of off-resonance collision data, collected at 60 MeV below the $\Upsilon(4S)$ resonance, is used to model background from $e^+e^-$ continuum processes.

Simulated Monte Carlo (MC) samples of $B \to X e^+ \nu_e$ signal and background processes are used to obtain the reconstruction efficiencies and study the key kinematic distributions. These events are generated with EvtGen [6] and the used branching fractions are summarized in Table 1. Inclusive semileptonic $B \to X e^+ \nu_e$ decays are dominated by $B \to D e^+ \nu_e$ and $B \to D^* e^+ \nu_e$ transitions. We model the $B \to D e^+ \nu_e$ decays using the BGL parametrization [7] with form factor central values and uncertainties taken from the fit in Ref. [8]. For $B \to D^* \ell^+ \nu_\ell$, we use the BGL implementation proposed by Refs. [9, 10] with form factor central values and uncertainties from the fit to the measurement of Ref. [11]. Both backgrounds are normalized to the average branching fraction of Ref. [12] assuming isospin symmetry. Semileptonic $B \to D^{**} \ell^+ \nu_\ell$ decays with $D^{**} = \{D_0^0, D_1^+, D_1, D_2\}$ denoting the four orbitally excited charmed mesons are modeled using the heavy-quark-symmetry-based form factors proposed in Refs. [13, 14]. Non-resonant $B \to D^* \pi (\bar{\nu}_\ell)$ are modeled using the framework of Ref. [15]. To account for the remaining ‘gap’ between the sum of all considered exclusive modes and the inclusive $B \to X_e e^+ \nu_e$ branching fraction is filled with $B \to D^{(*)} \pi \ell^+ \nu_\ell$ and $B \to D^{(*)} \eta \ell^+ \nu_\ell$ decays using a model based on the equidistribution of all final-state particles in phase-space. Semileptonic $B \to X_\mu e^+ \nu_e$ decays are modeled as a mixture of specific exclusive modes and non-resonant contributions. The efficiencies
TABLE I. Branching fractions for \( B \to X_c \ell^+ \nu_\ell \) signal and \( B \to X_u \ell^+ \nu_\ell \) signal processes that were used to generate simulated samples are listed.

<table>
<thead>
<tr>
<th>( B )</th>
<th>( B^+ )</th>
<th>( B^0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( B \to X_c \ell^+ \nu_\ell )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( B \to D \ell^+ \nu_\ell )</td>
<td>( (2.4 \pm 0.1) \times 10^{-2} )</td>
<td>( (2.3 \pm 0.1) \times 10^{-2} )</td>
</tr>
<tr>
<td>( B \to D^* \ell^+ \nu_\ell )</td>
<td>( (5.6 \pm 0.2) \times 10^{-2} )</td>
<td>( (5.1 \pm 0.1) \times 10^{-2} )</td>
</tr>
<tr>
<td>( B \to D_0^* \ell^+ \nu_\ell )</td>
<td>( (0.4 \pm 0.1) \times 10^{-2} )</td>
<td>( (0.5 \pm 0.2) \times 10^{-2} )</td>
</tr>
<tr>
<td>( B \to D^1_0 \ell^+ \nu_\ell )</td>
<td>( (0.4 \pm 0.1) \times 10^{-2} )</td>
<td>( (0.5 \pm 0.1) \times 10^{-2} )</td>
</tr>
<tr>
<td>( B \to D_1 \ell^+ \nu_\ell )</td>
<td>( (0.8 \pm 0.5) \times 10^{-2} )</td>
<td>( (0.7 \pm 0.1) \times 10^{-2} )</td>
</tr>
<tr>
<td>( B \to D_2^* \ell^+ \nu_\ell )</td>
<td>( (0.4 \pm 0.1) \times 10^{-2} )</td>
<td>( (0.3 \pm 0.1) \times 10^{-2} )</td>
</tr>
<tr>
<td>( B \to D \pi \ell^+ \nu_\ell )</td>
<td>( (0.1 \pm 0.1) \times 10^{-2} )</td>
<td>( (0.1 \pm 0.1) \times 10^{-2} )</td>
</tr>
<tr>
<td>( B \to D^* \pi \ell^+ \nu_\ell )</td>
<td>( (0.1 \pm 0.1) \times 10^{-2} )</td>
<td>( (0.1 \pm 0.1) \times 10^{-2} )</td>
</tr>
<tr>
<td>( B \to D \pi \pi \ell^+ \nu_\ell )</td>
<td>( (0.1 \pm 0.1) \times 10^{-2} )</td>
<td>( (0.1 \pm 0.1) \times 10^{-2} )</td>
</tr>
<tr>
<td>( B \to D^* \pi \pi \ell^+ \nu_\ell )</td>
<td>( (0.2 \pm 0.2) \times 10^{-2} )</td>
<td>( (0.2 \pm 0.2) \times 10^{-2} )</td>
</tr>
<tr>
<td>( B \to D \eta \ell^+ \nu_\ell )</td>
<td>( (0.1 \pm 0.1) \times 10^{-2} )</td>
<td>( (0.1 \pm 0.1) \times 10^{-2} )</td>
</tr>
<tr>
<td>( B \to D^* \eta \ell^+ \nu_\ell )</td>
<td>( (0.1 \pm 0.1) \times 10^{-2} )</td>
<td>( (0.1 \pm 0.1) \times 10^{-2} )</td>
</tr>
<tr>
<td>( B \to X_u \ell^+ \nu_\ell )</td>
<td>( (2.2 \pm 0.3) \times 10^{-3} )</td>
<td>( (2.1 \pm 0.3) \times 10^{-3} )</td>
</tr>
<tr>
<td>( B \to X \ell^+ \nu_\ell )</td>
<td>( (10.99 \pm 0.28) \times 10^{-2} )</td>
<td>( (10.33 \pm 0.28) \times 10^{-2} )</td>
</tr>
</tbody>
</table>

in the MC are corrected using data-driven methods to account for differences in particle identification and reconstruction efficiencies.

3. ANALYSIS STRATEGY

Our general analysis strategy is as follows: we first identify samples of same-sign and opposite-sign di-electron candidates. We then apply a selection based on tracks and global event properties to enrich the samples with double-semileptonic \( B^0 - \bar{B}^0 \) (“signal”) decays. For events passing the selection, we build a variable that can distinguish between signal and misreconstructed events, which is fitted to extract \( N_{SS} \) and \( N_{OS} \). We describe these steps in more detail below.

3.1. Di-electron candidate selection

We build electron pair candidates from single electron candidates that satisfy the following criteria:

- High center-of-mass momentum: \( |p_\ell^*| > 1 \text{ GeV} \).
- Impact parameter consistent with the interaction point: the track must pass within 4 cm along the beam axis in the laboratory frame and within 2 cm transverse to the beam axis of the interaction point.
High electron ID likelihood: ID > 0.9.

Each electron candidate is corrected for Bremsstrahlung radiation using a dedicated search algorithm that matches electromagnetic clusters in the calorimeter to electron candidates. We then separate the electron pair candidates according to the reconstructed charges into SS and OS categories.

### 3.2. Signal enrichment

Using the correspondence between the generated particles and the reconstructed tracks, we classify the pair as “Signal” if both electrons are daughters of different $B$ mesons, and label the rest “Other.” We use the off-resonance dataset to describe the continuum background distributions.

We use the following selections, in this order, to suppress the major backgrounds:

- **Photon and $J/\psi$ veto**: we check whether either of the signal electron candidates can match with any opposite-sign track in the event under the electron mass hypothesis to give an invariant mass near zero ($m_{ee} < 0.2$ GeV) to reject electrons from photon conversions or near the $J/\psi$ resonance (2.92 GeV $< m_{ee} < 3.14$ GeV). We discard di-electron candidates if such a match is found ($\epsilon_{SS} = 0.958$, $\epsilon_{OS} = 0.896$).

- **Best candidate selection**: if multiple di-electron pair candidates exist in an event, we randomly choose one as the sole candidate in the event ($\epsilon_{SS} = 0.999$, $\epsilon_{OS} = 0.990$).

- **Event track number selection**: we discard events for which there are less than five total tracks that pass the same impact parameter selections as the signal leptons ($\epsilon_{SS} = 0.923$, $\epsilon_{OS} = 0.917$).

- **Continuum suppression**: we retain events with a normalized Fox-Wolfram moment [16] value $R_2 < 0.3$ ($\epsilon_{SS} = 0.889$, $\epsilon_{OS} = 0.895$).

- **Electron ID coverage selection**: we discard events for which electron ID corrections are not available for both electron candidates ($\epsilon_{SS} = 0.890$, $\epsilon_{OS} = 0.894$).

After these selections, the remaining backgrounds come largely from mis-identified electrons, same-$B$ secondary decays $B \rightarrow (X_c \rightarrow X_s e^- \nu_e) e^+ \nu_e$ (OS only), and continuum events.

### 3.3. The extraction variable

Our signal consists of two electrons from semileptonic $B$ decays, each of which has a mean energy above that of the primary background sources. Therefore, the sum of the magnitude of the momentum in the center-of-mass frame of the two electrons

$$p_{ee} = |p^*_{e1}| + |p^*_{e2}| \ ,$$

provides discrimination between signal and background. This variable has not been exploited thus far in time-integrated $\chi_d$ measurements. In Fig. [1] we show this spectrum in OS and SS for MC and data.
FIG. 1. The $p_{ee}$ spectrum for opposite-sign (left) and same-sign (right) di-electron samples after all selections and before fitting. The shaded, stacked histograms show the expected spectra from the sum of Signal MC (green), Other $B \bar{B}$ MC (orange), and scaled off-resonance data (purple). The black points indicate the spectrum as measured in data. The shaded area shows the size of the systematic uncertainty from lepton identification efficiencies, signal and background shapes, and the statistical uncertainty of the off-resonance data. The lower distributions show the normalized residuals between data and MC calculated as $(N_{\text{data}} - N_{\text{MC}})/\sqrt{\sigma_{\text{data}}^2 + \sigma_{\text{MC}}^2}$.

4. FITTING PROCEDURE

The number of same-sign and opposite-sign $B \to X e^+ \nu_e$ candidates is determined by a simultaneous binned likelihood fit to the $p_{ee}$ distribution of both samples and in 11 bins ranging from 2 - 6 GeV. For each of the 11 $p_{ee}$ bins, the free parameters of the fit are:

a) The number of same-sign and opposite-sign $B \to X e^+ \nu_e$ candidates: $N_{SS}, N_{OS}$

b) The number of background events in each sample from $B$ meson decays: $N_{BSS}, N_{BOS}$

c) The number of background events in each sample from continuum processes: $N_{CSS}, N_{COS}$

and the correspond to the yields of the three event categories considered. The total likelihood function is

$$ L = \prod_{i}^{\text{bins}} P(n_i; \nu_i) \times \prod_{k} \mathcal{G}_k, $$

(6)

with $n_i$ denoting the number of observed data events and $\nu_i$ the total number of expected events in a given bin $i$. Here, $\mathcal{G}_k$ are nuisance-parameter (NP) constraints, whose role is to incorporate systematic uncertainties and the number of expected continuum events, as
determined from the off-resonance data, directly into the fit. The number of expected events in a given bin, $\nu_i$, is estimated using MC and off-resonance data and is given by

$$\nu_i = N_{SS} \cdot f_{i,SS} + N_{BSS} \cdot f_{i,BSS} + N_{CSS} \cdot f_{i,CSS},$$

(7)
or

$$\nu_i = N_{OS} \cdot f_{i,OS} + N_{BOS} \cdot f_{i,BOS} + N_{COS} \cdot f_{i,COS},$$

(8)

for same-sign or opposite-sign events, respectively. Here, the $f_i$ correspond to the fraction of events of each category to being reconstructed in bin $i$ as determined by the MC simulation or the off-resonance data. The NP constraints are constructed such that they take into account uncertainties due to the electron identification efficiency, signal and background template compositions, and the statistical uncertainty of the template in question. They are incorporated using multivariate Gaussian distributions, $G_k = G_k(0; \theta_k, \Sigma_k)$. Here the $\Sigma_k$ denotes the systematic covariance matrix for a given template $k$ and $\theta_k$ is a vector of NPs. The covariance $\Sigma_k$ is the sum over all possible uncertainty sources for a given template. The fractions in Eqs. (7) and (8) are allowed to change within these systematic uncertainties according to

$$f_i = \frac{N_{i,MC} (1 + \theta_i)}{\sum_j N_{j,MC} (1 + \theta_j)},$$

(9)

with $N_{i,MC}$ denoting the number of expected events of a given category in bin $i$ as estimated by MC, and $\theta_j$ the $j^{th}$ element of the NP vector $\theta_k$. The likelihood function is maximized numerically to determine all components and NP constraints. Confidence intervals for the $N_{SS}$, $N_{OS}$ components are constructed using the profile likelihood ratio method.

5. RESULTS

Figure 2 shows the post-fit $p_{ee}$ distribution for same-sign and opposite-sign candidates. The goodness-of-fit cannot be judged based on the residuals alone, but also depends on the NP pulls, which we show in Appendix A. From this fit, we measure:

$$\chi^2_d = 0.192 \pm 0.010 \text{ (stat.)} \pm 0.016 \text{ (syst.)}$$

(10)

where the first uncertainty is the statistical uncertainty and the second uncertainty is the systematic uncertainty. The uncertainty is dominated by the statistical uncertainty of the same-sign sample, the off-resonance data sample, and the systematic uncertainty of the electron identification. The size of these systematic uncertainties is expected to decrease with larger control samples. The obtained value of $\chi^2_d$ is compatible with the world average from time-independent and time-dependent determinations

$$\chi^2_{d,WA}^{t-in} = 0.182 \pm 0.015, \quad \chi^2_{d,WA} = 0.186 \pm 0.001.$$  

(11)

and the reported measurement already has a similar precision to the time-independent world average.
FIG. 2. The $p_{ee}$ spectrum for opposite-sign (left) and same-sign (right) di-electron samples after fitting. The shaded, stacked histograms show the fitted spectra from the sum of Signal MC (green), Other $B\bar{B}$ MC (orange), and scaled off-resonance data (purple). The black points indicate the spectrum as measured in data. The shaded area around the MC expectation correspond to the post-fit uncertainty. The lower distributions show the normalized residuals between data and MC calculated as $(N_{\text{data}} - N_{\text{MC}}) / \sqrt{\sigma_{\text{data}}^2 + \sigma_{\text{MC}}^2}$.

6. ACKNOWLEDGEMENTS

We thank the SuperKEKB group for the excellent operation of the accelerator; the KEK cryogenics group for the efficient operation of the solenoid; and the KEK computer group for on-site computing support. This work was supported by the following funding sources: Science Committee of the Republic of Armenia Grant No. 18T-1C180; Australian Research Council and research grant Nos. DP180102629, DP170102389, DP170102204, DP150103061, FT130100303, and FT130100018; Austrian Federal Ministry of Education, Science and Research, and Austrian Science Fund No. P 31361-N36; Natural Sciences and Engineering Research Council of Canada, Compute Canada and CANARIE; Chinese Academy of Sciences and research grant No. QYZDJ-SSW-SLH011, National Natural Science Foundation of China and research grant Nos. 11521505, 11575017, 11675166, 11761141009, 11705209, and 11975076, LiaoNing Revitalization Talents Program under contract No. XLYC1807135, Shanghai Municipal Science and Technology Committee under contract No. 19ZR1403000, Shanghai Pujiang Program under Grant No. 18PJ1401000, and the CAS Center for Excellence in Particle Physics (CCEPP); the Ministry of Education, Youth and Sports of the Czech Republic under Contract No. LTT17020 and Charles University grants SVV 260448 and GAUK 404316; European Research Council, 7th Framework PIEF-GA-2013-622527, Horizon 2020 Marie Skłodowska-Curie grant agreement No. 700525 ‘NIOBE,’ and Horizon 2020 Marie Skłodowska-Curie RISE project JENNIFER2 grant agreement No. 822070 (European grants); L’Institut National de Physique Nucléaire et de Physique des Particules (IN2P3) du CNRS (France); BMBF, DFG, HGF, MPG, AvH Foundation, and Deutsche Forschungsgemeinschaft (DFG) under Germany’s Excellence Strategy – EXC2121 “Quantum Universe”.)


Appendix A: Nuisance parameter pulls

We include distributions of the NP pulls from the fit shown in Fig. 2 for the signal (Fig. 3), other $B\bar{B}$ (Fig. 4), and off-resonance (Fig. 5) templates.

FIG. 3. Pulls on the nuisance parameters for the OS (left) and SS (right) signal templates.

FIG. 4. Pulls on the nuisance parameters for the OS (left) and SS (right) $B\bar{B}$ background templates.
FIG. 5. Pulls on the nuisance parameters for the OS (left) and SS (right) off-resonance templates.