We present a new measurement of the CKM matrix element $|V_{cb}|$ from $B^0 \to D^* \ell \nu$ decays, reconstructed with full Belle data set (711 fb$^{-1}$). Two form factor parameterisations, based on work by the CLN and BGL groups, are used to extract the product $F(1)\eta_{EW}|V_{cb}|$ and the decay form factors, where $F(1)$ is the factor normalisation and $\eta_{EW}$ is a small electroweak correction. In the CLN parameterisation we find $F(1)|V_{cb}| = (35.06 \pm 0.15 \pm 0.54) \times 10^{-3}$, $r^2 = 1.106 \pm 0.031 \pm 0.007$, $R_1(1) = 1.229 \pm 0.028 \pm 0.009$, $R_2(1) = 0.852 \pm 0.021 \pm 0.006$. In the BGL parameterisation we find $F(1)|V_{cb}| = 38.73 \pm 0.25 \pm 0.60$, which is higher but consistent with the determination from inclusive semileptonic $B$ decays when correcting for $F(1)\eta_{EW}$. This is the most precise measurement of $F(1)|V_{cb}|$ and form factors that has ever been carried out, and the first direct study of the BGL form factor parameterisation in an experimental measurement.

I. INTRODUCTION

The decay $B^0 \to D^* \ell \nu$ is used to calculate the Cabibbo-Kobayashi-Maskawa (CKM) matrix element $|V_{cb}|$, the magnitude of the coupling between $b$ and $c$ quarks in weak interactions, and is a fundamental parameter of Standard Model (SM). The $B^0 \to D^* \ell \nu$ decay is studied in the context of Heavy Quark Effective Theory (HQET) in which the hadronic matrix elements are parameterised by the form factors that describe this decay. The decay amplitudes of $B^0 \to D^* \ell \nu$ are described by three helicity amplitudes which are extracted from the three polarisation states of the $D^*$ meson: two transverse polarisation terms, $H_{\pm}$, and one longitudinal polarisation term, $H_0$.

There is a long standing tension in the measurement of $|V_{cb}|$ using the inclusive approach, based on the decay mode $B \to X \ell \nu$ and the exclusive approach with $B \to D^* \ell \nu$. Currently, the world averages for $|V_{cb}|$ for inclusive and exclusive decay modes are [1]:

$$|V_{cb}| = (42.2 \pm 0.8) \times 10^{-3} \quad \text{(Inclusive)},$$
$$|V_{cb}| = (39.1 \pm 0.4) \times 10^{-3} \quad \text{(CLN-Exclusive)},$$

where the error is the experimental and theoretical combined. It is thought that the previous theoretical approaches using the CLN form factor parameterisation [2] were model dependent and introduced bias, and therefore model independent form factor approaches based on BGL [3] should be used. In this paper we will perform fits with both approaches for the first time in an experimental paper. In this paper, the decay is reconstructed in the following channel: $B^0 \to D^{*-} \ell^+ \nu_\ell$ where $D^{*-} \to D^0 \pi^-$ and $D^0 \to K^- \pi^+$. This channel offers the best purity for the measurement, which is critical as it is ultimately systematic uncertainty limited. This is experimentally the most precise determination of $|V_{cb}|$ performed with exclusive semileptonic $B$ decays. This result supersedes the previous results on $B \to D^* \ell^+ \nu_\ell$ with an untagged approach from Belle [4]. A major experimental improvement to the efficiency of the track reconstruction software was implemented in 2011, leading to substantially higher slow pion tracking efficiencies and hence much larger signal yields than the previous result.

II. EXPERIMENTAL APPARATUS AND DATA SAMPLES

We use the full $\Upsilon(4S)$ data sample containing $772 \times 10^6 BB$ pairs recorded with the Belle detector [5] at the asymmetric-beam-energy $e^+e^-$ collider KEKB [6]. An additional 88 fb$^{-1}$ of data is collected 60 MeV below the $\Upsilon(4S)$ for the estimation of $q\bar{q}$ ($q = u, d, s, c$) continuum background.

The Belle detector is a large-solid-angle magnetic spectrometer that consists of a silicon vertex detector (SVD), a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC), a barrel-like arrangement of time-of-flight scintillation counters (TOF), and an electromagnetic calorimeter (ECL) comprised of CsI(Tl) crystals located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux-return located outside of the coil is instrumented to detect $K_L^0$ mesons and to identify muons (KLM). The detector is described in detail elsewhere [5]. Two inner detector configurations were used. A 2.0 cm radius beampipe and a 3-layer silicon vertex detector was used for the first sample of $152 \times 10^6 BB$ pairs (referred to as SVD1), while a 1.5 cm radius beampipe, a 4-layer silicon detector and a small-cell inner drift chamber were used to record the remaining $620 \times 10^6 BB$ pairs [7] (referred to as SVD2). We refer to these subsamples later in the paper.

A. Monte Carlo Simulation

Monte Carlo simulated events are used to determined the analysis selection criteria, study the background and estimate the signal reconstruction efficiency. Events with a $BB$ pair are generated using EvtGen [8], and the $B$ meson decays are reproduced based on branching fractions reported in Ref. [9]. The hadronisation process
of $B$ meson decays that do not have experimentally-measured branching fractions is inclusively reproduced by PYTHIA [10]. For the continuum $e^+e^- \rightarrow q\bar{q}$ events, the initial quark pair is hadronised by PYTHIA, and hadron decays are modelled by EvtGen. The final-state radiation from charged particles is added using PHOTOS [11]. Detector responses are simulated with GEANT3 [12].

B. Event reconstruction and selection criteria

Charged particle tracks are required to originate from the interaction point, and to have good track fit quality. The criteria for the track impact parameters in the $r - \phi$ and $z$ directions are: $dr < 2$ cm and $|dz| < 4$ cm, respectively. In addition we require that each track has at least one associated hit in any layer of the SVD detector. For pion and kaon candidates, we use particle identification likelihoods determined using Cherenkov light yield in the ACC, the time-of-flight information from the TOF, and $dE/dx$ from the CDC.

Neutral $D^0$ meson candidates are reconstructed only in the clean $D^0 \rightarrow K^-\pi^+$ decay channel. The daughter tracks are fit to a common vertex using a Kalman fit algorithm, with a $\chi^2$-probability requirement of greater than $10^{-3}$ to reject background. The reconstructed $D^0$ mass is required to be in a window of $\pm 13.75$ MeV/$c^2$ from the nominal $D^0$ mass of $1.865$ GeV/$c^2$, corresponding to a width of $2.5 \sigma$, determined from data.

The $D^0$ candidates are combined with an additional pion that has a charge opposite that of the kaon, to form $D^{*+}$ candidates. Pions produced in this transition are close to kinematic threshold, with a mean momentum of approximately $100$ MeV/$c$, hence are denoted slow pions, $\pi^s$. There are no SVD hit requirements for slow pions. Another vertex fit is performed between the $D^*$ and the $\pi^s$ and a $\chi^2$-probability requirement of greater than $10^{-3}$ is again imposed. The invariant mass difference between the $D^*$ and the $D^0$ candidates, $\Delta m = m_{D^*} - m_{D^0}$, is first required to be less than $165$ MeV/$c^2$ for the background fit, and further tightened for the signal yield determination.

Although the contribution from $e^+e^- \rightarrow q\bar{q}$ continuum is relatively small in this analysis, we further suppress prompt charm by imposing an upper threshold on the $D^*$ momentum of $2.45$ GeV/$c$ in the CM frame (Fig. 1).

Candidate $B$ mesons are reconstructed by combining $D^*$ candidates with an oppositely charged electron or muon. Electron candidates are identified using the ratio of the energy detected in the ECL to the momentum of the track, the ECL shower shape (E9/E25), the distance between the track at the ECL surface and the ECL cluster centre, the energy loss in the CDC ($dE/dx$) and the response of the ACC. For electron candidates we search for nearby bremsstrahlung photons in a cone of 3 degrees around the electron track, and sum the momenta with that of the electron. Muons are identified by their penetration range and transverse scattering in the KLM detector. In the momentum region relevant to this analysis, charged leptons are identified with an efficiency of about $90\%$, while the probabilities to misidentify a pion as an electron or muon is $0.25\%$ and $1.5\%$ respectively. We impose lower thresholds on the momentum of the leptons, such that they reach the respective particle identification detectors for good hadron fake rejection. Here we impose lab frame momentum thresholds $0.3$ GeV/$c$ for electrons and $0.6$ GeV/$c$ for muons. We furthermore require an upper threshold of $2.4$ GeV/$c$ in the CM frame to reject continuum events.

III. DECRY KINEMATICS

The tree level transition of the $B^0 \rightarrow D^{*--}e^+\nu_e$ decay is shown in Fig. 2. Three angular angular variables and the hadronic recoil are used to describe this decay. The latter is defined as follows. Iwhere $q^2$ is the momentum transfer between the $B$ and the $D^*$ meson, and $m_B, m_{D^*}$ are the the masses of $B$ and $D^*$ mesons respectively. The range of $w$ is restricted by the value of $q^2$ such that the minimum value of $q^2 = 0$ corresponds to the maximum value of $w$,

$$w_{\text{max}} = \frac{m_B^2 + m_{D^*}^2}{2m_Bm_{D^*}}.$$  

FIG. 1. The $D^*$ momenta in the CM frame, for on-resonance and scaled off-resonance data.

FIG. 2. Tree level Feynman diagram for $B^0 \rightarrow D^{*--}e^+\nu_e$. 

The tree level transition of the $B^0 \rightarrow D^{*--}e^+\nu_e$ decay is shown in Fig. 2. Three angular angular variables and the hadronic recoil are used to describe this decay. The latter is defined as follows. Iwhere $q^2$ is the momentum transfer between the $B$ and the $D^*$ meson, and $m_B, m_{D^*}$ are the the masses of $B$ and $D^*$ mesons respectively. The range of $w$ is restricted by the value of $q^2$ such that the minimum value of $q^2 = 0$ corresponds to the maximum value of $w$,
The three angular variables are depicted in Fig. 3 and are defined as follows:

- $\theta_l$: the angle between the $D^*$ and the lepton, defined in the rest frame of $W$ boson.
- $\theta_v$: the angle between the $D^0$ and the $D^*$, defined in the rest frame of $D^*$ meson.
- $\chi$: the angle between the two planes formed by the decays of the $W$ and the $D^*$ meson, defined in the rest frame of the $B^0$ meson.

**FIG. 3.** Definition of the angles $\theta_l$, $\theta_v$ and $\chi$ for the decay $B^0 \to D^*\ell^+\nu_\ell$.

### IV. SEMILEPTONIC DECAYS

In the massless lepton limit, the differential decay rate of $B \to D^*\ell\nu$ decays is given by [2]

$$\frac{d\Gamma(B \to D^*\ell\nu)}{dw d\theta_l d\cos\theta_V d\chi} = \frac{\eta_{EW}^2 m_B m_{D^*}^2 G_F^2 |V_{cb}|^2 \sqrt{w^2 - 1}(1 - 2wr + r^2)}{4(4\pi)^2} \left\{ (1 - \cos\theta_l)^2 \sin^2 \theta_V^2 H_+^2 + (1 + \cos\theta_l)^2 \sin^2 \theta_V^2 H_-^2 + 4 \sin \theta_V^2 \cos \theta_V^2 H_0^2 - 2 \sin \theta_V^2 \cos \theta_V^2 \sin 2\chi H_+ H_- - 4 \sin \theta_V \cos \theta_V \sin \theta_R \cos \theta_V \cos \chi H_0 H_0 - 4 \sin \theta_V \cos \theta_V \cos \theta_V \cos \chi H_0 H_0 \right\},$$

where $r = m_{D^*}/m_B$, $G_F = (1.6637 \pm 0.00001) \times 10^{-5} h^2 GeV^{-2}$ and $\eta_{EW}$ is a small electroweak correction (equal to 1.006 in Ref. [13]).

#### A. The CLN Parameterisation

The helicity amplitudes $H_{\pm,0}$ in Eq. [4] are given in terms of three form factors. In the Caprini-Lellouch-Neubert (CLN) parameterisation [2] one writes these expressions in terms of the form factor $h_{A_1}(w)$ and the form factor ratios $R_{1,2}(w)$. They are defined as follows.

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

$$R_1(w) = R_1(1) - 0.12(w - 1) + 0.05(w - 1)^2,$$

$$R_2(w) = R_2(1) - 0.11(w - 1) - 0.06(w - 1)^2,$$

where $z = (\sqrt{w} + 1 - \sqrt{2})/(\sqrt{w} + 1 + \sqrt{2})$, and there are four independent parameters in total. After integrating over the angles, the $w$ distribution is proportional to

$$\mathcal{F}(w) = h_{A_1}^2(w) \left( 1 + 4 \frac{w}{w + 1} \frac{1 - 2wr + r^2}{(1 - r^2)} \right)^{-1} \left[ 2 - 2wr + r^2 \left( 1 + R_1^2 \frac{w - 1}{w + 1} \right) + \left( 1 + (1 - R_2(w)) \frac{w - 1}{1 - r} \right)^2 \right].$$

#### B. The BGL Parameterisation

A more general parameterisation comes from Boyd, Grinstein and Lebed (BGL) [3], recently used in Refs. [14, 15]. In their approach, the helicity amplitudes $H_i$ are given by

$$H_0(w) = \mathcal{F}_1(w) \sqrt{w^2 - 1},$$

$$H_{\pm}(w) = f(w) \mp m_B m_{D^*} \sqrt{w^2 - 1} g(w).$$

The relation between the form factors in the BGL and CLN notations are

$$f = \sqrt{m_B m_{D^*}} (1 + w) h_{A_1},$$

$$g = h_{V}/\sqrt{m_B m_{D^*}},$$

$$\mathcal{F}_1 = (1 + w)(m_B - m_{D^*}) \sqrt{m_B m_{D^*}} A_5,$$

and

$$R_1(w) = (w + 1) m_B m_{D^*} \frac{g(w)}{f(w)},$$

$$R_2(w) = \frac{w - r}{w - 1} - \frac{\mathcal{F}_1}{m_B (w - 1) f_1(w)}. $$

The three BGL form factors can be written as a series in $z$,

$$f(z) = \frac{1}{P_1(z) \phi_f(z)} \sum_{n=0}^{\infty} a_{f_n} z^n,$$

$$\mathcal{F}_1(z) = \frac{1}{P_1(z) \phi_{\mathcal{F}_1}(z)} \sum_{n=0}^{\infty} a_{\mathcal{F}_1} z^n,$$

$$g(z) = \frac{1}{P_1(z) \phi_g(z)} \sum_{n=0}^{\infty} a_{g_n} z^n.$$
where $z_p$ is defined as

$$z_p = \frac{\sqrt{t_+ - m_p^2} - \sqrt{t_- - t^*}}{\sqrt{t_+ - m_p^2} + \sqrt{t_- - t^*}}.$$  \hspace{1cm} (12)$$

and $t^\pm = (m_B \pm m_{D^\pm})^2$. The product is extended to include all the $B_c$ resonances below the $B - D^*$ threshold of 7.29 GeV with the appropriate quantum numbers ($1^+$ for $f$ and $F_1$, and $1^-$ for $g$). We use the the $B_c$ resonances listed in Table I. The $B_c$ resonances also enter the $1^-$ 

| Type | Mass (GeV/c^2) |
|------|----------------|
| $1^-$ | 6.337          |
| $1^-$ | 6.899          |
| $1^-$ | 7.012          |
| $1^-$ | 7.280          |
| $1^-$ | 6.730          |
| $1^+$ | 6.736          |
| $1^+$ | 7.135          |
| $1^+$ | 7.142          |

unitarity bounds as single particle contributions. The outer functions $\phi_i$ for $i = g, f, F_1$ are as follows:

$$\phi_g(z) = \sqrt{\frac{n_I}{3\pi \chi^T_{-1}(0)}} \times \frac{2^2 r^2 (1+z)^2 (1-z)^{-1/2}}{(1+r)(1-z) + 2\sqrt{1+z}}.$$  \hspace{1cm} (13)

$$\phi_f(z) = \frac{4r}{m_B^2} \sqrt{\frac{n_I}{3\pi \chi^T_{1+}(0)}} \times \frac{(1+z)(1-z)^{3/2}}{[(1+r)(1-z) + 2\sqrt{1+z}]^4}.$$  \hspace{1cm} (14)

$$\phi_{F_1}(z) = \frac{4r}{m_B^2} \sqrt{\frac{n_I}{6\pi \chi^T_{1+}(0)}} \times \frac{(1+z)(1-z)^{5/2}}{[(1+r)(1-z) + 2\sqrt{1+z}]^5}.$$  \hspace{1cm} (15)

where $\chi^T_{1+}(0)$ and $\chi^T_{-1}(0)$ are constants given in Table II and $n_I = 2.6$ represents the number of spectator quarks (three), decreased by a large and conservative SU(3) breaking factor. At zero recoil ($w = 1$ or $z = 0$) there is a relation between two of the form factors,

$$F_1(0) = (m_B - m_{D^*})f(0).$$  \hspace{1cm} (16)

The coefficients of the expansions in Eq. 10 are subject to unitarity bounds based on analyticity and the operator product expansion applied to correlators of two hadronic

| Input | Value |
|-------|-------|
| $m_{B^0}$ | 5.279 GeV/c^2 |
| $m_{D^{*+}}$ | 2.010 GeV/c^2 |
| $\eta_{EW}$ | 1.0066 |
| $\chi^T_{-1}(0)$ | $5.28 \times 10^{-4}$ GeV^{-2} |
| $\chi^T_{1+}(0)$ | $3.07 \times 10^{-4}$ GeV^{-2} |

**V. BACKGROUND ESTIMATION**

The most powerful discriminator against background is the cosine of the angle between the $B$ and the $D^*\ell$ momentum vectors in the CM frame under the assumption that the $B$ decays to $D^*\ell\nu$. In the CM frame, the $B$ direction lies on a cone around the $D^*\ell$ axis with an opening angle $2 \cos \theta_{B,D^*\ell}$, defined as:

$$\cos \theta_{B,D^*\ell} = \frac{2E^*_B|E_{D^*\ell}| - m^2_B - m^2_{D^*\ell}}{2|p^*_B||p^*_{D^*\ell}|},$$  \hspace{1cm} (17)

where $E^*_B$ is half of the CM energy and $|p^*_B|$ is $\sqrt{E^*_B^2 - m^2_B}$. The quantities $E^*_{D^*\ell}$, $p^*_{D^*\ell}$, and $m_{D^*\ell}$ are determined from the reconstructed $D^*\ell$ system.

The remaining background in the sample is split into the following categories.

- $B \to D^{**}\ell\nu$, both resonant where $D^{**}$ decays to a $D^*$, and non-resonant $B \to D^*\pi\ell\nu$ decays.
- Correlated cascade decays where the $D^*$ and $\ell$ originate from the same $B$, e.g. $B \to D^*\tau\nu$ ($\tau \to \ell\nu\nu$), and $B \to D^*D, D \to \ell X$.
- Uncorrelated decays, where the $D^*$ and $\ell$ originate from different $B$ mesons in the event.
- Mis-identified leptons (fake leptons): the probability for a hadron being identified as a lepton is small but not negligible in the low momentum region, and is higher for muons.
- Fake $D^*$ candidates, where the $D^*$ is incorrectly reconstructed.
- $q\bar{q}$ continuum, typically $e^+e^- \to c\bar{c}$. 

\[\sum_{i=0}^{\infty} (a_i)^2 < 1,\]
\[\sum_{i=0}^{\infty} [(a_i)^2 + (a_i^*)^2] < 1.\]  \hspace{1cm} (18)
To model the $B \to D^{∗+}\ell\nu$ component, which is comprised of four $P$-wave resonant modes ($D_1, D_0^*, D_1^*, D_2^*$) for both neutral and charged $B$ decays, we correct the branching ratios and form factors. The $P$-wave charm mesons are categorised according to the angular momentum of the light constituent, $j_t$, namely the $j_t^P = 1/2^−$ doublet of $D_0^*$ and $D_1^*$ and the $j_t^P = 3/2^−$ doublet $D_1$ and $D_2^*$. The shapes of the $B \to D^{∗+}\ell\nu$ $q^2$ distributions are corrected to matched the predictions of the LLSW model [10]. An additional contribution from non-resonant modes is considered, although the rate appears to be consistent with zero in recent measurements.

To estimate the background yields we perform a binned maximum log likelihood fit of the $D^*\ell$ candidates in three variables, $\Delta m, \cos \theta_{B,D^*\ell}$, and $p_\ell^t$. The bin ranges are as follows:

- $\Delta m$: 5 equidistant bins in the range $[0.141, 0.156]$ GeV/c$^2$.
- $\cos \theta_{B,D^*\ell}$: 15 equidistant bins in the range $[-10, 5]$.
- $p_\ell^t$: 2 bins in the ranges $[0.6, 0.85], [3.0]$ GeV/c for muons and $[0.3, 0.80, 3.0]$ GeV/c for electrons.

Prior to the fit, the residual continuum background is estimated from off-resonance data and scaled by the off-resonance integrated luminosities and the $1/s$ dependence of the $e^+e^-\to q\bar{q}$ cross section. The kinematics of the off and on-resonant continuum background is expected to be slightly different and therefore binned correction weights are determined using MC and applied to the scaled off-resonance data. The remaining background components are modelled with MC simulation after correcting for the most recent decay modelling parameters, and for differences in reconstruction efficiencies between data and MC. Corrections are applied to the lepton identification efficiencies, hadron misidentification rates, and slow pion tracking efficiencies. The data/MC ratios for high momentum tracking efficiencies are consistent with unity and are only considered in the systematic uncertainty estimates. The results from the background fits are given in Table [III] and Fig. [4].

VI. MEASUREMENT OF DIFFERENTIAL DISTRIBUTIONS

Measurement of the decay kinematics requires good knowledge of the signal $B$ direction to constrain the neutrino momentum 4-vector. To determine the $B$ direction we estimate the CM frame momentum vector of the non-signal $B$ meson by summing the momenta of the remaining particles in the event ($p^\text{incl}_i$) and choose the direction on the cone that minimises the difference to $-p^\text{incl}_i$. To determine $p^\text{incl}_i$, we exclude tracks that do not pass near the interaction point. The impact parameter requirements depend on the transverse momentum of the track, $p_T$, and are set to:

- $p_T < 250$ MeV/c: $dr < 20$ cm, $dz < 100$ cm,
- $p_T < 500$ MeV/c: $dr < 15$ cm, $dz < 50$ cm,
- $p_T \geq 500$ MeV/c: $dr < 10$ cm, $dz < 20$ cm.

Some track candidates may be counted multiple times, due to low momentum particles spiralling in the central drift chamber, or due to fake tracks fit to a similar set of detector hits as the primary track. This can be removed by looking for pairs of tracks with similar kinematics, travelling in the same direction with the same electric charge, or in the opposite direction with the opposite electric charge. Isolated clusters that are not matched to the signal particles (i.e. from photons or $\pi^0$ decays) are required to have lower energy thresholds to mitigate beam induced background, and are 50, 100 and 150 MeV in the barrel, forward end-cap and backward end-cap regions respectively. We compute $\vec{p}_{\text{incl}}$ by summing the 3-momenta of the selected particles:

$$\vec{p}_{\text{incl}} = \sum_i \vec{p}_i,$$

where the index $i$ denotes all isolated clusters and tracks that pass the above criteria. This vector is then translated into the CM frame. There is no mass assumption used for the charged particles. The energy component, $E^\text{incl}_i$, is set to the experiment dependent beam energies through $E^\text{beam} = \sqrt{5}/2$.

We find that the resolutions of the kinematic variables are 0.020 for $w$, 0.038 for $\cos \theta_v$, 0.044 for $\cos \theta_{\ell\nu}$ and 0.210 for $\chi$. Based on these resolutions, and the available data sample, we split each distribution into 10 equidistant bins for the $|V_{cb}|$ and form factor fits.

A. Fit to the CLN Parameterisation

We perform a binned $\chi^2$ fit to determine the following quantities in the CLN parameterisation: the product $F_1|V_{cb}|$, and the three parameters $\rho^2$, $R_1(1)$ and $R_2(1)$ that parameterise the form factors. We use a set of one-dimensional projections of $w, \cos \theta_v, \cos \theta_{\ell\nu}$ and $\chi$. This reduces complications in the description of the six background components and their correlations across four dimensions. This approach introduces finite bin to bin correlations that must be accounted for in the $\chi^2$ calculation.

We choose equidistant binning in each kinematic observable, as described above, and set the ranges accordingly to their kinematically allowed limits. The exception is $w$: while the kinematically allowed range is between 1 and 1.504, we restrict this to between 1 and 1.50 such that we can ignore the finite mass of the lepton in the interaction.

The number of expected events in a given bin $i$, $N^\text{theory}_i$, is given by

$$N^\text{theory}_i = N_{B^0}B(D^{∗+} \to D_0^0\pi^+) \times B(D^0 \to K^-\pi^+)\mathcal{B}_{\text{int}}\Gamma_i,$$
where $N_{B^0}$ is the number of $B^0$ mesons in the data sample, $B(D^{*+} \rightarrow D^0\pi^+)$ and $B(D^0 \rightarrow K^-\pi^+)$ are the $D^*$ and $D$ branching ratios into the final state studied in this analysis, $\tau_{B^0}$ is the $B^0$ lifetime, and $\Gamma_j$ is the width obtained by integrating the CLN theoretical expectation within the corresponding bin boundaries. The expected number of events, $N_i^{\text{exp}}$, must take into account finite detector resolution and efficiency,

$$N_i^{\text{exp}} = \sum_{j=1}^{40} (R_{ij} \epsilon_j N_j^{\text{theory}}) + N_i^{\text{bkg}},$$

where $\epsilon_j$ is probability that an event generated in bin $j$ is reconstructed and passes the analysis selection criteria, and $R_{ij}$ is the detector response matrix (the probability that an event generated in bin $j$ is observed in bin $i$). $N_i^{\text{bkg}}$ is the number of expected background events as constrained from the total background yield fit.

In the nominal fit we use the following $\chi^2$ function based on a forward folding approach:

$$\chi^2 = \sum_{i,j} (N_{i\text{obs}} - N_i^{\text{exp}}) C_{ij}^{-1} (N_{j\text{obs}} - N_j^{\text{exp}}),$$

where $N_{i\text{obs}}$ are the number of events observed in bin $i$ of our data sample, and $C_{ij}^{-1}$ is the inverse of the covariance matrix. The covariance matrix is the variance-covariance matrix whose diagonal elements are the variances and the off-diagonal elements are the covariance of the elements from the $i^{\text{th}}$ and $j^{\text{th}}$ positions. The covariance is calcu-
TABLE III. Signal and background fractions (%) for events selected in the signal region of (|cos θ_{B_0,D+ℓ}| < 1, 0.144 < Δm < 0.147, p_ℓ < 0.80, p_ν > 0.85).

|                | SVD1(e) | SVD1(µ) | SVD2 (e) | SVD2 (µ) |
|----------------|---------|---------|----------|----------|
| Signal Events  | 83.31 ± 0.60 | 83.84 ± 0.54 | 84.80 ± 0.19 | 84.20 ± 0.21 |
| Fake ℓ         | 0.10 ± 0.17  | 1.16 ± 0.73  | 0.10 ± 0.82  | 1.21 ± 0.40  |
| Fake D*        | 3.17 ± 0.10  | 3.02 ± 0.06  | 3.08 ± 0.01  | 2.96 ± 0.02  |
| D**            | 6.07 ± 0.42  | 4.19 ± 0.25  | 5.32 ± 0.14  | 3.82 ± 0.09  |
| Signal Corr.   | 1.29 ± 0.35  | 2.08 ± 0.40  | 1.49 ± 0.07  | 2.52 ± 0.15  |
| Uncorrelated   | 6.05 ± 0.52  | 5.23 ± 0.61  | 5.19 ± 0.15  | 5.27 ± 0.26  |
| Continuum      | 4.29 ± 0.67  | 4.64 ± 0.77  | 4.68 ± 0.40  | 5.43 ± 0.47  |

lated for each pair of bins in either w, cos θ_ℓ, cos θ_V and χ. The off-diagonal elements are calculated as,

\[ C_{ij} = N_{ij} - N_{pi}p_{pj}, \quad \forall i \neq j \tag{21} \]

where \( p_{ij} \) is the relative probability of the two-dimensional histograms between observable pairs, \( p_i \) and \( p_j \) are the relative probabilities of the one dimensional histograms of each observable, and \( N \) is the total size of the sample. The diagonal elements are the variances of \( N_{isp} \) and are calculated as,

\[ \sigma_i^2 = \sum_{j=1}^{40} \left[ R_{ij}^2 \left( \epsilon_j (N_j^i) + c_j^i \right)^2 + R_{ij} \frac{1}{N_{data}} (N_j^i)^2 \right] + \frac{1}{N_{MC}} \]  

\[ + \left[ R_{ij} \frac{1}{N_{data}} (N_j^i)^2 \right] + \frac{\sigma^2(N_i^{bkg})}{N_{MC}} \]  

which uses the Poisson uncertainty associated with the number of events in the MC and data in each bin, and the final term is the total error associated with the background arising from the background fit procedure. We have tested this fit procedure using MC simulated data samples and all results are consistent with expectations, showing no signs of bias. The results from the fit are summarised in Table IV and the fit correlation coefficients are given in Table V. We find reasonable \( p \)-values and results consistent with the world averages. The comparison between data and the form factor fit are shown in Fig. 5.

TABLE IV. Statistical correlation matrix of the fit to the full sample in the CLN parameterisation.

|          | \( \rho^2 \) | \( R_1(1) \) | \( R_2(1) \) | \( F(1)[V_{cb}] \) |
|----------|--------------|--------------|--------------|------------------|
| \( \rho^2 \) | 1.000        | 0.533        | -0.873       | 0.722            |
| \( R_1(1) \) | 1.000        | -0.629       | -0.014       | -0.333           |
| \( R_2(1) \) | 1.000        | -0.333       | -0.014       | -0.333           |
| \( F(1)[V_{cb}] \) | 1.000        |              |              |                  |

B. Fit to the BGL Parameterisation

To perform the fit to the BGL parameterisation we follow the approach described in Ref. [14]. We similarly truncate the series in the expansion for \( a_f \) and \( a_g \) terms at \( O(z^2) \) and order \( O(z^3) \) for \( F_1 \). This results in five free parameters (one more than in the CLN fit), defined as \( a_{f(i)} = [V_{cb}]_{\eta}^{F(i)} \) where \( i = 0, 1 \), \( a_{f(i)}^{s} = [V_{cb}]_{\eta}^{s} \) where \( i = 1 \), and \( a_{f(i)}^{s} = [V_{cb}]_{\eta}^{s} \) where \( i = 1, 2 \) This number of free parameters can describe the data well, while higher order terms will not be well constrained unless additional information from lattice is introduced. We perform a \( \chi^2 \) fit to the data with same procedure as for the CLN fit described above. The resulting value for \( |V_{cb}| \) is larger than that from the CLN parameterisation, and consistent with the inclusive approach. The fit results are given in Table VI and Fig. 6. The linear correlation coefficients are listed in Table VII. Correlations can be high in this fit approach. Only the SVD1+SVD2 combined samples are fitted as the fit does not converge well with the smaller SVD1 data sets in this parameterisation.

VII. SYSTEMATIC UNCERTAINTIES

To estimate the systematic uncertainties on the partial branching fractions, CLN form factor parameters, and \( |V_{cb}| \), we consider the following sources: background component normalisations, MC tracking efficiency, charm branching fractions, \( B \to D^{**}ℓν \) branching fractions and decay differentials, the \( B^0 \) lifetime, and the number of \( B^0 \) mesons in the data sample. The systematic uncertainties on the branching fraction, \( F(1)[V_{cb}] \) and CLN form factor parameters from the CLN fit are summarised in Table VIII. These systematic errors are compatible with BGL approach as well.

We estimate systematic uncertainties by varying each possible uncertainty source such as the PDF shape and the signal reconstruction efficiency with the assumption of a Gaussian error, unless otherwise stated. This is done via sets of pseudo-experiments in which each independent systematic uncertainty parameter is randomly varied us-
FIG. 5. Results of the fit with the CLN form factor parameterisation. The results from the SVD1 and SVD2 samples are added together. The electron modes are on the left and muon modes on the right. The points with error bars are the on-resonance data. Where not shown, the uncertainties are smaller than the black markers. The histograms are, top to bottom, the signal component, $B \to D^{**}$ background, signal correlated background, uncorrelated background, fake $\ell$ component, fake $D^*$ component and continuum.

The parameters varied are split into two categories, those that affect only the normalisation, and those that affect the differentials (shapes). We first list the latter contributions.

- The tracking efficiency corrections for low momen-
TABLE V. Fit Results for the four sub-samples in the CLN parameterisation where the following parameters are floated: $\rho^2$, $R_1(1)$, $R_2(1)$ along with $\mathcal{F}(1)|V_{cb}|$. The $p$-value corresponds to the $\chi^2$/ndf using the statistical errors only.

| Parameter | SVD1 $e$ | SVD1 $\mu$ | SVD2 $e$ | SVD2 $\mu$ |
|-----------|----------|----------|----------|----------|
| $\rho^2$  | 1.165 $\pm$ 0.099 | 1.165 $\pm$ 0.102 | 1.087 $\pm$ 0.046 | 1.095 $\pm$ 0.051 |
| $R_1(1)$  | 1.326 $\pm$ 0.106 | 1.336 $\pm$ 0.102 | 1.117 $\pm$ 0.040 | 1.289 $\pm$ 0.048 |
| $R_2(1)$  | 0.767 $\pm$ 0.073 | 0.777 $\pm$ 0.074 | 0.861 $\pm$ 0.030 | 0.882 $\pm$ 0.034 |
| $\mathcal{F}(1)|V_{cb}|\times 10^3$ | 34.66 $\pm$ 0.48 | 35.01 $\pm$ 0.50 | 35.25 $\pm$ 0.23 | 34.98 $\pm$ 0.24 |
| $\chi^2$/ndf | 35/36 | 36/36 | 44/36 | 43/36 |
| $p$-value | 0.52 | 0.47 | 0.17 | 0.20 |
| B.F [%] | 4.84 $\pm$ 0.06 | 4.91 $\pm$ 0.06 | 4.88 $\pm$ 0.03 | 4.82 $\pm$ 0.03 |

TABLE VI. Fit results for the electron and muon sub-samples in the BGL parameterisation where the following parameters are floated: $\tilde{a}_0^l$, $\tilde{a}_1^l$, $\tilde{a}_2^l$, $\tilde{a}_3^l$, $\tilde{a}_4^l \times 10^2$ along with $\mathcal{F}(1)|V_{cb}|\times 10^3$ (derived from $\tilde{a}_0^l$). The $p$-value corresponds to the $\chi^2$/ndf using the statistical errors only.

| Parameter | $e$ | $\mu$ |
|-----------|-----|-----|
| $\tilde{a}_0^l \times 10^2$ | 0.0566 $\pm$ 0.0005 | 0.0561 $\pm$ 0.0006 |
| $\tilde{a}_1^l \times 10^2$ | 0.0758 $\pm$ 0.0243 | 0.0644 $\pm$ 0.0281 |
| $\tilde{a}_2^l \times 10^2$ | 0.0264 $\pm$ 0.0010 | 0.0288 $\pm$ 0.0106 |
| $\tilde{a}_3^l \times 10^2$ | 0.2771 $\pm$ 0.1866 | 0.3712 $\pm$ 0.2078 |
| $\tilde{a}_4^l \times 10^2$ | 0.0963 $\pm$ 0.0027 | 0.1110 $\pm$ 0.0031 |
| $\mathcal{F}(1)|V_{cb}|\times 10^3$ | 38.89 $\pm$ 0.34 | 38.56 $\pm$ 0.38 |
| $\chi^2$/ndf | 48/35 | 40/35 |
| $p$-value | 0.08 | 0.26 |

TABLE VII. Statistical correlation matrix of the fit to the full sample in the BGL parameterisation.

| $\tilde{a}_0^l$ | $\tilde{a}_1^l$ | $\tilde{a}_2^l$ | $\tilde{a}_3^l$ | $\tilde{a}_4^l$ |
|-----------|-------------|-------------|-------------|-------------|
| $\tilde{a}_0^l$ | 1.0000 | -0.7794 | -0.6685 | -0.0386 |
| $\tilde{a}_1^l$ | 1.0000 | 0.4724 | -0.4124 | -0.0770 |
| $\tilde{a}_2^l$ | 1.0000 | -0.9814 | 0.0713 | 0.0000 |
| $\tilde{a}_3^l$ | 1.0000 | -0.5096 | 0.0000 |

TABLE VIII. Combined results for the full data set in the BGL scheme where the errors shown are statistical only.

| Parameter | Values |
|-----------|--------|
| $\tilde{a}_0^l \times 10^2$ | 0.0564 $\pm$ 0.0004 |
| $\tilde{a}_1^l \times 10^2$ | 0.0701 $\pm$ 0.0183 |
| $\tilde{a}_2^l \times 10^2$ | 0.0276 $\pm$ 0.0071 |
| $\tilde{a}_3^l \times 10^2$ | 0.3242 $\pm$ 0.1388 |
| $\tilde{a}_4^l \times 10^2$ | 0.1037 $\pm$ 0.0020 |
| $\mathcal{F}(1)|V_{cb}|\times 10^3$ | 38.73 $\pm$ 0.25 |

The lepton identification efficiencies are varied according to their respective uncertainties, which are dominated by contributions that are correlated across all bins in $p_{\text{lab}}$ and $\theta_{\text{lab}}$.

The results from the background normalisation fit are varied within their fitted uncertainties. We take into account finite correlations between the fit results of each component.

The uncertainty of the decays $\bar{B} \rightarrow D^{**}\ell^-\bar{\nu}_\ell$ are twofold: the indeterminate composition of each $D^{**}$ state and the uncertainty in the form-factor parameters used for the MC sample production. The composition uncertainty is estimated based on uncertainties of the branching fractions: $\pm 6\%$ for $B \rightarrow D_1 \rightarrow D^*\pi \ell\bar{\nu}_\ell$, $\pm 12\%$ for $B \rightarrow D_2 \rightarrow D^*\pi \ell\bar{\nu}_\ell$, $\pm 24\%$ for $B \rightarrow D_1 \rightarrow D^{*}\pi \ell\bar{\nu}_\ell$ and $\pm 17\%$ for $B \rightarrow D_2 \rightarrow D^{*}\pi \ell\bar{\nu}_\ell$. If the experimentally-measured branching fractions are not applicable, we vary the branching fractions continuously from 0\% to 200\% in the MC expectation. We estimate an uncertainty arising from the LLSW model parameters by changing the correction fac-
FIG. 6. Results of the fit with the BGL form factor parameterisation. The results from the SVD1 and SVD2 samples are added together. The electron modes are on the left and muon modes on the right. The points with error bars are the on-resonance data. Where not shown, the uncertainties are smaller than the black markers. The histograms are, top to bottom, the signal component, $B \rightarrow D^{*\ast}$ background, signal correlated background, uncorrelated background, fake $\ell$ component, fake $D^*$ component and continuum.

- The relative number of $B^0 \bar{B}^0$ meson pairs compared to $B^+ B^-$ pairs collected by Belle has a small uncertainty and affects only the relative composition of cross-feed signal events from $B^+$ and $B^0$ decays.
- Charged hadron particle identification uncertainties are determined with data using $D^*$ tagged...
charm decays.

The uncertainties that only affect the overall normalisation are: the tracking efficiency for high momentum tracks, the branching ratios $\mathcal{B}(D^{+} \rightarrow D^{0}\pi^{+})$, and $\mathcal{B}(D^{0} \rightarrow K^{-}\pi^{+})$, the total number of $\Upsilon(4S)$ events in the sample, and the $B^{0}$ lifetime.

**VIII. RESULTS**

The full results for the CLN fit are given below, where the first uncertainty is statistical, and the second systematic.

$$
\rho^{2} = 1.106 \pm 0.031 \pm 0.007,
R_{1}(1) = 1.229 \pm 0.028 \pm 0.009,
R_{2}(1) = 0.852 \pm 0.021 \pm 0.006,
\mathcal{B}(B^{0} \rightarrow D^{*-}\ell^{+}\nu_{\ell}) = (4.86 \pm 0.02 \pm 0.15)\%,
\mathcal{F}(1)|V_{cb}|\eta_{\text{EW}} \times 10^{3} = 35.06 \pm 0.15 \pm 0.54.
$$

These results are consistent with, and more precise than those published in Refs. [14, 15]. We find the value of branching fraction is insensitive to the choice of parameterisation. We also present the results for $|V_{cb}|$ from the BGL fit, where the first uncertainty is statistical, and the second systematic.

$$
\mathcal{F}(1)|V_{cb}|\eta_{\text{EW}} \times 10^{3} = 38.73 \pm 0.25 \pm 0.60.
$$

These results are consistent with those based on a preliminary tagged approach by Belle [20], as performed in Refs. [14, 15]. Both sets of fits give acceptable $\chi^{2}$/ndf: therefore the data does not discriminate between the parameterisations. The result with the BGL parameterisation has a larger fit uncertainty.

Taking the value of $\mathcal{F}(1) = 0.906 \pm 0.013$ from Lattice QCD in Ref. [21] and $\eta_{\text{EW}} = 1.0066$ from Ref. [13], we find the following values for $|V_{cb}|$: $(38.4 \pm 0.2 \pm 0.6 \pm 0.6) \times 10^{-3}$ (CLN+LQCD) and $(42.5 \pm 0.3 \pm 0.7 \pm 0.6) \times 10^{-3}$ (BGL+LQCD). The value of $|V_{cb}|$ using CLN parameterisation is consistent with the world average value where as the value we get using BGL parameterisation is higher but consistent with the inclusive $|V_{cb}|$ value shown in Eq 2 and Eq 3 respectively.

We perform a lepton flavour universality (LFU) test by forming a ratio of the branching fractions of modes with electrons and muons. The corresponding value of this ratio is

$$
\frac{\mathcal{B}(B^{0} \rightarrow D^{*-}\ell^{+}\nu_{\ell})}{\mathcal{B}(B^{0} \rightarrow D^{*-}\mu^{+}\nu_{\mu})} = 1.01 \pm 0.01 \pm 0.03,
$$

where the first error is statistical and the second is systematic. The systematic uncertainty is dominated by the electron and muon identification uncertainties, as all others cancel in the ratio. This is the most stringent test of LFU in $B$ decays. This result is consistent with unity.

**IX. CONCLUSION**

In this conference paper we present a new study by the Belle experiment of the decay $B \rightarrow D^{*-}\ell\nu$. We present the most precise measurement of $|V_{cb}|$ from exclusive decays, and the first direct measurement using the BGL parameterisation. The BGL parameterisation gives a higher value for $|V_{cb}|$, which is closer to that expected from the inclusive approach [1] [22, 24]. We also place stringent bounds on lepton flavour universality violation, which has been observed to be consistent with zero.

**ACKNOWLEDGMENTS**

We thank the KEKB group for the excellent operation of the accelerator, the KEK cryogenics group for the efficient operation of the solenoid, and the KEK computer group and the National Institute of Informatics for valuable computing and SINET3 network support. We acknowledge support from the Ministry of Education, Culture, Sports, Science, and Technology of Japan and the Japan Society for the Promotion of Science; the Australian Research Council and the Australian Department of Education, Science and Training; the National Natural Science Foundation of China under contract No. 10575109 and 10775142; the Department of Science and Technology of India; the BK21 program of the Ministry of Education of Korea, the CHEP SRC program and Basic Research program (grant No. R01- 2005-000-10089-0) of the Korea Science and Engineering Foundation, and the Pure Basic Research Group program of the Korea Research Foundation; the Polish State Committee for Scientific Research; the Ministry of Education and Science of the Russian Federation and the Russian Federal Agency for Atomic Energy; the Slovenian Research Agency; the Swiss National Science Foundation; the National Science Council and the Ministry of Education of Taiwan; and the U.S. Department of Energy.

[1] Y. Amhis et al. (HFLAV Collab.), Eur. Phys. J. C 77, no. 12, 895 (2017).
[2] I. Caprini, L. Lellouch and M. Neubert, Nucl. Phys. B 530 (1998) 153.
[3] C. G. Boyd, B. Grinstein, and R. F. Lebed, Phys. Rev. D 56, 6895 (1997).
[4] W. Dungel et al. (Belle Collab.), Phys. Rev. D 82, 112007 (2010).
| Systematic Uncertainties                  | $\rho^2$ | $R_1(1)$ | $R_2(1)$ | $F(1)|V_{cb}|$ [%] | B.F. [%] |
|------------------------------------------|---------|----------|----------|---------------------|---------|
| Slow pion efficiency                    | 0.005   | 0.002    | 0.001    | 0.65                | 1.29    |
| Lepton ID combined                      | 0.001   | 0.006    | 0.004    | 0.68                | 1.38    |
| $B(B \to D^{**} \ell \nu)$              | 0.002   | 0.001    | 0.002    | 0.26                | 0.52    |
| $B \to D^{**} \ell \nu$ Form factors    | 0.003   | 0.001    | 0.004    | 0.10                | 0.22    |
| $f_+/f_0$                                | 0.001   | 0.002    | 0.002    | 0.52                | 1.06    |
| Fake $e/\mu$                            | 0.004   | 0.006    | 0.001    | 0.11                | 0.21    |
| Norm. continuum                         | 0.002   | 0.002    | 0.001    | 0.01                | 0.06    |
| Fast track efficiency                   | -       | -        | -        | 0.53                | 1.05    |
| $N(T(4S))$                               | -       | -        | -        | 0.68                | 1.37    |
| $B^0$ life time                         | -       | -        | -        | 0.13                | 0.26    |
| $K/\pi$ ID                              | -       | -        | -        | 0.39                | 0.77    |
| $B(D^{**} \to D^0 \pi^+_0)$             | -       | -        | -        | 0.37                | 0.74    |
| $B(D^0 \to K \pi)$                      | -       | -        | -        | 0.51                | 1.02    |
| **Total Systematic Error**              | 0.008   | 0.009    | 0.007    | 1.60                | 3.21    |

[5] A. Abashian et al. (Belle Collab.), Nucl. Instr. and Meth. A 479, 117 (2002).
[6] S. Kurokawa and E. Kikutani, Nucl. Instr. and. Meth. A 499, 1 (2003), and other papers included in this volume.
[7] Z. Natkaniec et al. (Belle SVD2 Group), Nucl. Instr. and Meth. A 560, 1 (2006).
[8] D.J. Lange, Nucl. Instr. and Meth. A 462, 152 (2001).
[9] C. Patrignani et al. (Particle Data Group), Chin. Phys. C 40, no. 10, 100001 (2016).
[10] T. Sjostrand, S. Mrenna and P. Skands, J. High Energy Phys. 2006, 05, 026 (2006).
[11] N. Davidson, T. Przedzinski and Z. Was, Comput. Phys. Commun. 199, 86 (2016).
[12] R. Brun et al., GEANT 3.21, CERN Report DD/EE/84-1, 1984 (unpublished).
[13] A. Sirlin, Nucl. Phys. B 3196, 83 (1982).
[14] B. Grinstein and A. Kobach, Phys. Lett. B 771, 359 (2017).
[15] D. Bigi, P. Gambino and S. Schacht, Phys. Lett. B 769, 441 (2017).
[16] A.K. Leibovich, Z. Ligeti, I.W. Stewart and M.B. Wise, Phys. Rev. D 57, 308 (1998).
[17] B. Aubert et al. (BaBar Collab.), "|V_{cb}|," Phys. Rev. D 77, 032002 (2008).
[18] B. Aubert et al. (BaBar Collab.), Phys. Rev. Lett. 100, 231803 (2008).
[19] B. Aubert et al. (BaBar Collab.), Phys. Rev. D 79, 012002 (2009).
[20] A. Abdesselam et al. (Belle Collab.), [arXiv:1702.01521] [hep-ex].
[21] J. A. Bailey et al. (Fermilab Lattice, MILC), Phys. Rev. D 89, 114504 (2014).
[22] P. Urquijo et al. (Belle Collab., Phys. Rev. D 75, 032001 (2007).
[23] C. Schwanda et al. (Belle Collab.), Phys. Rev. D 75, 023005 (2007).