$N = 5, 6, 7, 8$: Nested hypothesis tests and truncation dependence of $|V_{cb}|$

Florian U. Bernlochner,[1] Zoltan Ligeti,[2] and Dean J. Robinson[2,3]

[1]*Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany*
[2]*Ernest Orlando Lawrence Berkeley National Laboratory, University of California, Berkeley, California 94720, USA*
[3]*Santa Cruz Institute for Particle Physics and Department of Physics, University of California Santa Cruz, Santa Cruz, California 95064, USA*

(Received 1 March 2019; published 23 July 2019)

The determination of $|V_{cb}|$ from exclusive semileptonic $B \to D^* \ell \nu$ decays is sensitive to the choice of form factor parametrization. Larger $|V_{cb}|$ values are obtained by fitting the Boyd-Grinstein-Lebed (BGL) versus the Caprini-Lellouch-Neubert (CLN) parametrization to recent Belle measurements. For the BGL parametrization, published fits use different numbers of parameters. We propose a method based on nested hypothesis tests to determine the optimal number of BGL parameters to fit the data, and find that six parameters are optimal to fit the Belle tagged and unfolded measurement [1]. We further explore the differences between fits that use different numbers of parameters. The fits which yield $|V_{cb}|$ values in better agreement with determinations from inclusive semileptonic decays tend to exhibit tensions with heavy quark symmetry expectations. These have to be resolved before the determinations of $|V_{cb}|$ from exclusive and inclusive decays can be considered understood.

DOI: 10.1103/PhysRevD.100.013005

I. INTRODUCTION

In 2017, the Belle Collaboration presented, for the first time, unfolded measurements of the differential decay distributions for $\bar{B} \to D^* \ell \bar{\nu}$ decays [1], and another measurement appeared more recently [2]. The unfolded measurement [1] permitted outside groups to perform their own fits to the data, using different parametrizations of the $\bar{B} \to D^* \ell \bar{\nu}$ form factors to extract $|V_{cb}|$. The choice of form factor parametrizations can have a sizable impact on the extracted value of $|V_{cb}|$. This is because heavy quark symmetry gives the strongest constraints on the differential rate at zero recoil (maximal dilepton invariant mass, q^2) [3–10], resulting in both continuum methods and lattice QCD giving the most precise information on the normalization of the rate at zero recoil. However, phase space vanishes near maximal q^2 as $\sqrt{q^2_{\max} - q^2}$, so the measured q^2 spectrum has to be fitted over some range to extract $|V_{cb}|$. This results in sensitivity to the functional form of the fitted parametrization.

Fitting Belle's unfolded measurement [1] to the BGL parametrization [11,12] yielded higher values of $|V_{cb}|$ [13,14] than fitting the CLN [15] parametrization to the same dataset. (To our knowledge, during 1997–2017, all *BABAR* and Belle measurements of $|V_{cb}|$ from $\bar{B} \to D^* \ell \bar{\nu}$ used the CLN parametrization.) The BGL results are in better agreement with $|V_{cb}|$ extracted from inclusive $B \to X_c \ell \bar{\nu}$ decays [16]:

$$|V_{cb}|_{\mathrm{CLN}} = (38.2 \pm 1.5) \times 10^{-3}, \quad [1], \tag{1a}$$

$$|V_{cb}|_{\mathrm{BGL}_{332}} = (41.7^{+2.0}_{-2.1}) \times 10^{-3}, \quad [13], \tag{1b}$$

$$|V_{cb}|_{\mathrm{BGL}_{222}} = (41.9^{+2.0}_{-1.9}) \times 10^{-3}, \quad [14]. \tag{1c}$$

Here the BGL_{ijk} notation highlights that these fits have different numbers of parameters (the notation is defined below in Sec. II), in particular eight and six parameters, respectively. In Ref. [2], the Belle Collaboration published an "untagged" measurement of $\bar{B} \to D^* \ell \bar{\nu}$, without fully reconstructing the second B meson in the collision using hadronic decay modes. In that analysis, fits to the CLN and a five-parameter version of the BGL parametrization were performed [2], and the results are in agreement:

$$|V_{cb}|_{\mathrm{CLN}} = (38.4 \pm 0.9) \times 10^{-3}, \tag{2a}$$

$$|V_{cb}|_{\mathrm{BGL}_{122}} = (38.3 \pm 1.0) \times 10^{-3}. \tag{2b}$$

The BGL method implements constraints on the shapes of the $B \to D^*$ form factors based on analyticity and unitarity [17–19]. Three conveniently chosen linear combinations of form factors are expressed in terms of power series in a small conformal parameter, $0 < z \ll 1$. As indicated in Eqs. (1) and (2), there are varying choices for the total number of coefficients, N, in the three power series, ranging from $N = 5$ [2] to $N = 6$ [14,20] and $N = 8$ [13,21,22]. The CLN [15] prescription uses similar analyticity and unitarity constraints on the $B \to D$ form factor, heavy quark effective theory (HQET) [7,8] relations between the $B \to D$ and $B \to D^*$ form factors, and QCD sum rule calculations [23–25] of the order $\Lambda_{\rm QCD}/m_{c,b}$ subleading Isgur-Wise functions [9,10]. It has four fit parameters. [This version of the CLN parametrization, as used to extract $|V_{cb}|$, is not self consistent at $\mathcal{O}(\Lambda_{\rm QCD}/m_{c,b})$ [26].]

The relation between the above fits is nontrivial, and has not been studied systematically. The goal of this paper is to explore their differences, and to devise a quantitative method to identify the optimal number of parameters in the BGL framework. Using a prescription based on a nested hypothesis test, we find that at least six parameters are required to describe the data from Ref. [1]. The $N = 5$ and 6 fits we study in detail yield $|V_{cb}|$ values in better agreement with determinations from inclusive semileptonic decays, but they exhibit tensions with expectations from heavy quark symmetry.

II. FORMALISM AND NOTATIONS

The vector and axial-vector $\bar{B} \to D^*$ form factors are defined as

$$\begin{aligned} \langle D^*|\bar{c}\gamma^\mu b|\bar{B}\rangle &= i\sqrt{m_B m_{D^*}}\, h_V \varepsilon^{\mu\nu\alpha\beta}\epsilon^*_\nu v'_\alpha v_\beta, \\ \langle D^*|\bar{c}\gamma^\mu\gamma^5 b|\bar{B}\rangle &= \sqrt{m_B m_{D^*}}[h_{A_1}(w+1)\epsilon^{*\mu} \\ &\quad - h_{A_2}(\epsilon^*\cdot v)v^\mu - h_{A_3}(\epsilon^*\cdot v)v'^\mu], \end{aligned} \tag{3}$$

where v (v') is the four-velocity of the B (D^*). The form factors $h_{V,A_{1,2,3}}$ depend on $w = v\cdot v' = (m_B^2 + m_{D^*}^2 - q^2)/(2m_B m_{D^*})$. In the heavy quark limit, $h_{A_1} = h_{A_3} = h_V = \xi$ and $h_{A_2} = 0$, where ξ is the Isgur-Wise function [3,4]. Each of these form factors can be expanded in powers of $\Lambda_{\rm QCD}/m_{c,b}$ and α_s.

In the massless lepton limit (i.e., $\ell = e$ or μ), the differential $B \to D^*\ell\bar{\nu}$ rate is given by

$$\begin{aligned} \frac{{\rm d}\Gamma}{{\rm d}w} &= \frac{G_F^2|V_{cb}|^2\eta_{\rm ew}^2 m_B^5}{48\pi^3}\sqrt{w^2-1}(w+1)^2 r^3(1-r)^2 \\ &\quad\times\left[1+\frac{4w}{w+1}\frac{1-2wr+r^2}{(1-r)^2}\right][\mathcal{F}(w)]^2, \end{aligned} \tag{4}$$

where $r = m_{D^*}/m_B$, and $\mathcal{F}(w)$ can be written in terms of $h_{A_1}(w)$ and the two form factor ratios (see, e.g., Ref. [27]):

$$R_1(w) = \frac{h_V}{h_{A_1}}, \qquad R_2(w) = \frac{h_{A_3} + r h_{A_2}}{h_{A_1}}. \tag{5}$$

All measurable information is then contained in the three functions $\mathcal{F}(w)$ and $R_{1,2}(w)$. Throughout this paper, $\mathcal{F}(1) = 0.906$ [28] and $\eta_{\rm ew} = 1.0066$ [29] are used to convert fit results for $|V_{cb}|\mathcal{F}(1)\eta_{\rm ew}$ to values of $|V_{cb}|$. In the heavy quark limit, $R_{1,2}(w) = 1 + \mathcal{O}(\Lambda_{\rm QCD}/m_{c,b}, \alpha_s)$ and $\mathcal{F}(w) = \xi(w)$. Thus, $R_{1,2}(w) - 1$ parametrize deviations from the heavy quark limit.

The BGL framework is defined by expanding three form factors g, f, and $\mathcal{F}_1$, which are linear combinations of those defined in Eq. (3), in power series of the form $1/[P_i(z)\phi_i(z)] \times \sum a_n^i z^n$, where $i = g, f, \mathcal{F}_1$ (see, e.g., Ref. [12], and note that $\mathcal{F}_1 \neq \mathcal{F}$). Here $z = z(w)$ is a conformal parameter that maps the physical region $1 < w < 1.5$ onto $0 < z < 0.056$, and $P_i(z)$ and $\phi_i(z)$ are known functions [14]. There are two notations in the literature for the coefficients of these power series, which map onto each other via

$$\{a_n, b_n, c_n\} \quad [14] \longleftrightarrow \{a_n^g, a_n^f, a_n^{\mathcal{F}_1}\} \quad [13]. \tag{6}$$

In the remainder of this paper, we adopt the former notation, so that a_n, b_n, and c_n are the coefficients of g, f, and $\mathcal{F}_1$, respectively. (The convention for the sign of g, and thus the a_n, in Ref. [14] is opposite to that used in Refs. [13,22].) Note that c_0 is fixed by b_0 [12,14], and the fits are performed for the rescaled parameters

$$\{\tilde{a}_n, \tilde{b}_n, \tilde{c}_n\} = \eta_{\rm ew}|V_{cb}|\{a_n, b_n, c_n\}, \tag{7}$$

and $|V_{cb}|$ is determined by $|\tilde{b}_0|$.

To study and distinguish expansions truncated at different orders in z, we denote by ${\rm BGL}_{n_a n_b n_c}$ a BGL fit with the parameters

$$\{a_{0,\ldots,n_a-1}, b_{0,\ldots,n_b-1}, c_{1,\ldots,n_c}\}. \tag{8}$$

The total number of fit parameters is $N = n_a + n_b + n_c$. The BGL parametrization used in Refs. [14,20] is ${\rm BGL}_{222}$, while that used in Refs. [13,22] is ${\rm BGL}_{332}$.

III. NESTED HYPOTHESIS TESTS: FIXING THE OPTIMAL NUMBER OF COEFFICIENTS

Our aim is to construct a prescription to determine the optimal number of parameters to fit a given dataset. This can be achieved by use of a nested hypothesis test: a test of an N-parameter fit hypothesis versus a fit using one additional parameter (the alternative hypothesis).

Such a hypothesis test requires an appropriate statistical measure or test statistic. A suitable choice is the difference in χ^2,

$$\Delta\chi^2 = \chi^2_N - \chi^2_{N+1}. \tag{9}$$

n_c \ n_a	1	2	3	1	2	3	1	2	3
1	33.2 38.6 ± 1.0	31.6 38.6 ± 1.0	31.2 38.6 ± 1.0	33.0 39.0 ± 1.5	29.1 40.7 ± 1.6	28.9 40.7 ± 1.6	30.4 40.7 ± 1.7	29.1 40.6 ± 1.8	28.9 40.6 ± 1.8
2	32.9 38.8 ± 1.1	31.3 38.7 ± 1.1	31.1 38.8 ± 1.0	32.7 39.5 ± 1.7	**27.7** $\mathbf{41.7 \pm 1.8}$	27.7 41.6 ± 1.8	29.2 41.8 ± 2.0	27.7 41.8 ± 2.0	27.7 41.7 ± 2.0
3	31.7 39.0 ± 1.1	31.3 38.6 ± 1.2	31.0 38.6 ± 1.1	29.1 41.9 ± 2.0	27.7 41.8 ± 2.0	27.6 41.7 ± 2.0	29.2 41.8 ± 2.0	27.6 41.7 ± 1.9	23.2 41.4 ± 2.0
	$n_b = 1$			$n_b = 2$			$n_b = 3$		

FIG. 1. The χ^2 (upper entry) and $|V_{cb}| \times 10^3$ (lower entry) values for the $\mathrm{BGL}_{n_a n_b n_c}$ fits used for the nested hypothesis test. The number of free parameters in a given fit is $N = n_a + n_b + n_c$ and the bold entry is the selected BGL_{222} hypothesis $\{a_0, a_1, b_0, b_1, c_1, c_2\}$. Cells corresponding to $N = 5, 6, 7, 8$ are highlighted blue, green, orange, and red, respectively.

The fit with one additional parameter—the $(N+1)$-parameter fit—has one fewer degree of freedom (d.o.f.) (number of bins minus the number of parameters). In the large number of d.o.f. limit, $\Delta\chi^2$ is distributed as a χ^2 with a single d.o.f. [30]. One may reject or accept the alternative hypothesis by choosing a decision boundary. If, for instance, we choose $\Delta\chi^2 = 1$ as the decision boundary, we would reject the $(N+1)$-parameter hypothesis in favor of the N-parameter fit 68% of the time, if the N-parameter hypothesis is true.

We seek a prescription to incrementally apply this nested hypothesis test, starting from a suitably small initial number of parameters (to avoid possible overfitting), until we reach the simplest (smallest-N) fit containing the initial parameters, that is preferred over all hypotheses that nest it or are nested by it. For a set of BGL fits, we thus propose the following prescription starting from a suitable low-N fit $\mathrm{BGL}_{n_a n_b n_c}$:

(i) Carry out fits with one parameter added (a "descendant" fit) or, when permitted, removed (a "parent" fit); i.e., for $\mathrm{BGL}_{(n_a \pm 1) n_b n_c}$, $\mathrm{BGL}_{n_a (n_b \pm 1) n_c}$, $\mathrm{BGL}_{n_a n_b (n_c \pm 1)}$.

(ii) For each descendant (parent) hypothesis, accept it over $\mathrm{BGL}_{n_a n_b n_c}$ if $\Delta\chi^2$ is above (below) the decision boundary value.

(iii) Repeat (i) and (ii) recursively, until a "stationary" fit is reached, that is preferred over its parents and descendants.

(iv) If there are multiple stationary fits, choose the one with the smallest N, then the smallest χ^2.

The optimal truncation order obtained this way depends on the precision of the available experimental data. Our prescription attempts to minimize the residual model dependence (caused by this truncation) with respect to the experimental uncertainty.

Figure 1 shows the fitted χ^2 values for the set of 27 different $\mathrm{BGL}_{n_a n_b n_c}$ fits with $n_i = 1, 2, 3$. A suitable choice for a starting fit is BGL_{111} or one of the three possible fits with $N = 4$. Using the decision boundary of $\Delta\chi^2 > 1$, one then obtains a single stationary solution, BGL_{222}, shown in bold. For example, one path to BGL_{222} is $111 \to 211 \to 221 \to 222$, while another is $121 \to 131 \to 231 \to 232 \to 222$.

Also shown in Fig. 1 are the $|V_{cb}|$ values for all 27 fits. These results are consistent with the statement made in Ref. [13] that the extracted values of $|V_{cb}|$ remain stable when one adds more fit parameters to the BGL_{332} fit. This stability can be seen directly by comparing the preferred BGL_{222} fit with its descendants. One may notice that the χ^2 of the BGL_{333} fit is substantially smaller than those of its parents. However, our procedure starting from $N = 3$ or 4 fits always terminates before reaching so many parameters. Plotting the fitted BGL_{333} distributions, one sees that its small χ^2 is due to fitting fluctuations in the data, and should be seen as an overfit.

The unitarity constraints, $\sum_{n=0}^{\infty} |a_n|^2 \le 1$ and $\sum_{n=0}^{\infty} (|b_n|^2 + |c_n|^2) \le 1$, can be imposed on the fits. The stationary fit in our approach, BGL_{222}, is far from saturating these bounds [14]. While the form factors must obey the unitarity constraints, statistical fluctuations in their binned measurements may cause the central values to appear to violate unitarity[1] (at a modest confidence level). This can occur because such fits may yield large coefficients for higher-order terms to accommodate "wiggles" in the data. In this paper, we do not impose unitarity as a constraint; fits whose central values violate unitarity (at a modest confidence level) may suggest an overfit. This is the case for the BGL_{333} fit, providing another reason to limit the number of fit coefficients, as proposed in our method.

IV. COMPARING $N = 5$ FITS WITH BGL_{222}

To explore the differences between the various five-parameter fits and the BGL_{222} fit, we perform such fits to Belle's unfolded data [1]. (The untagged Belle measurement [2] is not unfolded, and cannot be analyzed at this point outside the Belle framework. With limited statistics,

[1] We thank Paolo Gambino for raising this question.

TABLE I. Summary of the BGL_{222}, BGL_{122}, BGL_{212}, and BGL_{221} fits to the tagged and unfolded Belle measurement [1].

	BGL_{222}	BGL_{122}	BGL_{212}	BGL_{221}
$\chi^2/\mathrm{n.d.f.}$	$27.7/34$	$32.7/35$	$31.3/35$	$29.1/35$
$\lvert V_{cb}\rvert \times 10^3$	41.7 ± 1.8	39.5 ± 1.7	38.7 ± 1.1	40.7 ± 1.6
$R_1(1)$	0.45 ± 0.31	1.30 ± 0.09	0.86 ± 0.37	0.48 ± 0.34
$R_1'(1)$	4.23 ± 1.28	0.26 ± 0.27	2.34 ± 1.60	4.02 ± 1.44
$R_2(1)$	1.00 ± 0.19	1.03 ± 0.20	1.05 ± 0.20	0.82 ± 0.10
$R_2'(1)$	-0.53 ± 0.43	-0.29 ± 0.51	-0.25 ± 0.52	-0.02 ± 0.05

the differences between the fits we perform on the unfolded data contain fluctuations, which are different from those of the folded measurement.) There are six possible fits with five parameters, as shown in Fig. 1. Here we focus on comparing BGL_{122}, BGL_{212}, and BGL_{221}, which set a_1, b_1, or c_2, respectively, to zero. (We do not study further the BGL_{311}, BGL_{131}, and BGL_{113} fits, as each removes two and adds one parameter to the BGL_{222} fit.)

The results of the BGL_{222} fit and the three five-parameter fits for the physical observables $\lvert V_{cb}\rvert$, $R_{1,2}(1)$, and $R'_{1,2}(1)$ are shown in Table I. (Our BGL_{222} fit results vary slightly from those in Ref. [20], due to using $m_B = 5.280$ GeV versus 5.279 GeV.) The best-fit parameters [rescaled as in Eq. (7)] and correlations for these four fits are shown in Fig. 2.

The results for the BGL_{222} fit in Fig. 2 suggest that, if one wants to reduce the number of fit parameters from six to five, the BGL_{122} fit might be the least optimal choice, as the significance of a nonzero value for $\lvert a_1\rvert$ is greater than for $\lvert b_1\rvert$, which is turn greater than for $\lvert c_2\rvert$. This is in line with the observation that, compared to the BGL_{222} fit, the value of χ^2 increases the most for BGL_{122}, followed by

	Param	Value $\times 10^2$	Correlation $\tilde a_0$	$\tilde a_1$	$\tilde b_0$	$\tilde b_1$	$\tilde c_1$	$\tilde c_2$
BGL_{222}	$\tilde a_0$	0.0379 ± 0.0249	1.000	-0.952	-0.249	0.417	0.137	-0.054
	$\tilde a_1$	2.6954 ± 0.9320		1.000	0.383	-0.543	-0.268	0.165
	$\tilde b_0$	0.0550 ± 0.0023			1.000	-0.793	-0.648	0.461
	$\tilde b_1$	-0.2040 ± 0.1064				1.000	0.542	-0.333
	$\tilde c_1$	-0.0433 ± 0.0264					1.000	-0.953
	$\tilde c_2$	0.5350 ± 0.4606						1.000

	Param	Value $\times 10^2$	Correlation $\tilde a_0$	$\tilde b_0$	$\tilde b_1$	$\tilde c_1$	$\tilde c_2$
BGL_{122}	$\tilde a_0$	0.1066 ± 0.0070	1.000	0.271	-0.163	-0.316	0.297
	$\tilde b_0$	0.0521 ± 0.0022		1.000	-0.767	-0.612	0.432
	$\tilde b_1$	-0.0446 ± 0.0839			1.000	0.489	-0.287
	$\tilde c_1$	-0.0193 ± 0.0252				1.000	-0.956
	$\tilde c_2$	0.2654 ± 0.4492					1.000

	Param	Value $\times 10^2$	Correlation $\tilde a_0$	$\tilde a_1$	$\tilde b_0$	$\tilde c_1$	$\tilde c_2$
BGL_{212}	$\tilde a_0$	0.0672 ± 0.0288	1.000	-0.972	0.128	-0.061	0.053
	$\tilde a_1$	1.4254 ± 1.0155		1.000	-0.074	-0.005	0.010
	$\tilde b_0$	0.0511 ± 0.0014			1.000	-0.420	0.342
	$\tilde c_1$	-0.0140 ± 0.0223				1.000	-0.976
	$\tilde c_2$	0.2187 ± 0.4367					1.000

	Param	Value $\times 10^2$	Correlation $\tilde a_0$	$\tilde a_1$	$\tilde b_1$	$\tilde b_1$	$\tilde c_1$
BGL_{221}	$\tilde a_0$	0.0399 ± 0.0270	1.000	-0.965	-0.294	0.472	0.330
	$\tilde a_1$	2.5020 ± 0.9984		1.000	0.380	-0.555	-0.408
	$\tilde b_0$	0.0537 ± 0.0021			1.000	-0.774	-0.787
	$\tilde b_1$	-0.1618 ± 0.1020				1.000	0.799
	$\tilde c_1$	-0.0141 ± 0.0082					1.000

FIG. 2. Fit coefficients and correlation matrices for the six-parameter BGL_{222} fit and 3 five-parameter BGL fits to the tagged and unfolded Belle measurement [1].

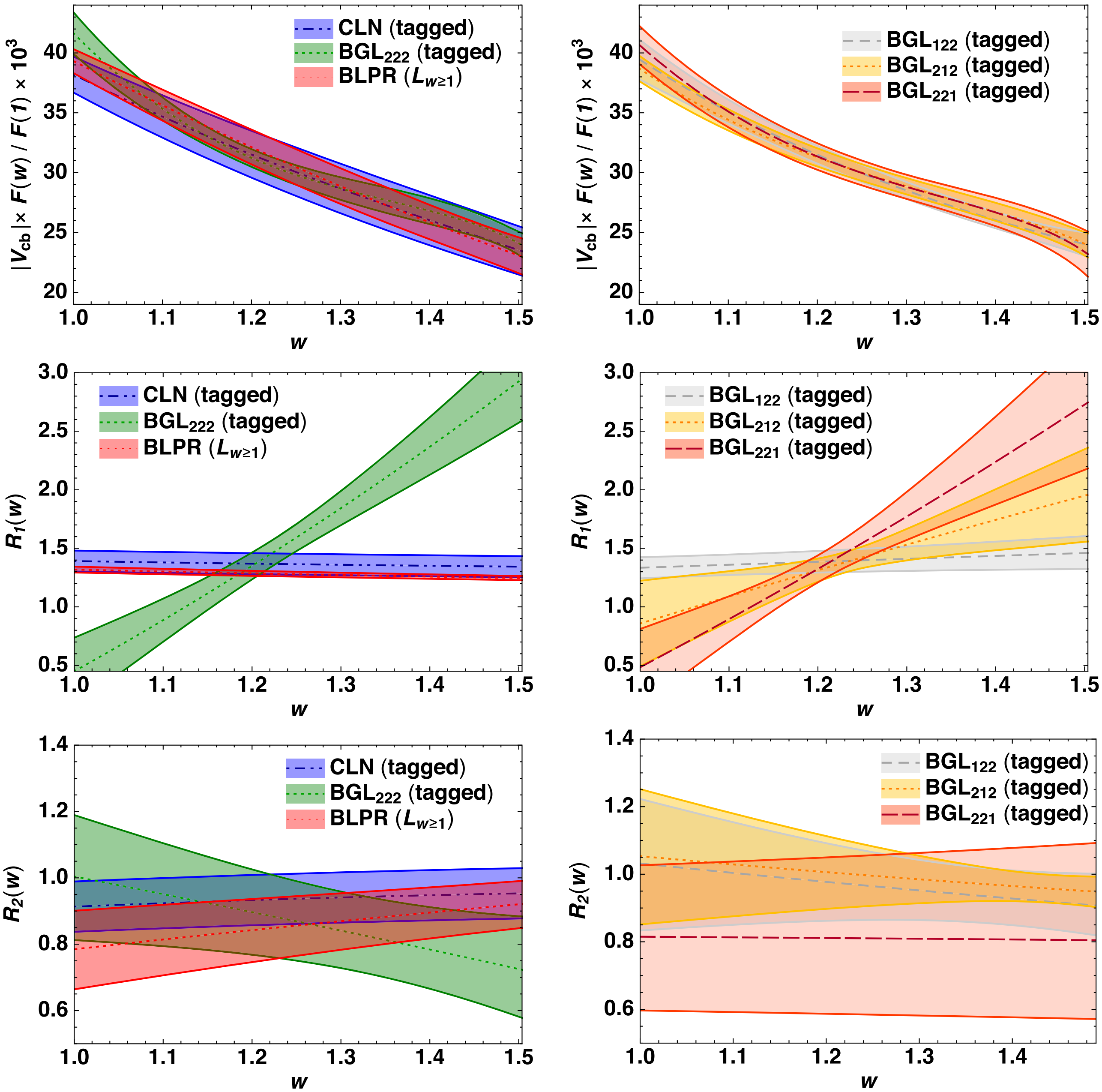

FIG. 3. The form factor $\mathcal{F}(w)$ (top), $R_1(w)$ (middle), and $R_2(w)$ (bottom) for the six fits described in the text.

BGL_{212}, and then BGL_{221}. This suggests that among the five-parameter fits, setting $c_2 = 0$ (the BGL_{221} fit) may instead be the preferred option—though inferior, according to our method, to the BGL_{222} fit for the Belle tagged and unfolded dataset [1].

The top row in Fig. 3 shows $\mathcal{F}(w)$ normalized to the lattice QCD value of $\mathcal{F}(1)$, as $|V_{cb}|\mathcal{F}(w)/\mathcal{F}(1)$ for six fits. The left-side plots show three previously published fits: the BGL_{222} and CLN fit results, based on the 2017 Belle tagged measurement, and the "BLPR" result of Ref. [26], which performed a HQET-based fit to both $B \to D^* l\bar{\nu}$ and $B \to D l\bar{\nu}$ data to determine the subleading $\mathcal{O}(\Lambda_{\mathrm{QCD}}/m_{c,b})$ Isgur-Wise functions, using also lattice QCD information. The right-side plots in Fig. 3 show the BGL_{122}, BGL_{212}, and BGL_{221} fits, based on the 2017 Belle tagged measurement [1]. The shaded bands indicate the uncertainties.

TABLE II. Fit coefficients used to construct the ensembles of toy experiments. The third-order terms $\{\tilde{a}_2, \tilde{b}_2, \tilde{c}_3\}$ are taken either as 1 or 10 times the second-order terms $\{\tilde{a}_1, \tilde{b}_1, \tilde{c}_2\}$ in the BGL_{222} fit shown in Fig. 2.

Parameter	$1\times$ scenario	$10\times$ scenario
$\tilde{a}_2 \times 10^2$	2.6954	26.954
$\tilde{b}_2 \times 10^2$	−0.2040	−2.040
$\tilde{c}_2 \times 10^2$	0.5350	5.350

TABLE III. The frequency of the selected hypotheses for ensembles created with the two scenarios for the higher-order terms, as estimated with an ensemble size of 250 pseudodata sets.

	BGL_{122}	BGL_{212}	BGL_{221}	BGL_{222}	BGL_{223}	BGL_{232}	BGL_{322}	BGL_{233}	BGL_{323}	BGL_{332}	BGL_{333}
$1\times$ scenario	6%	0%	37%	27%	6%	6%	11%	0%	2%	4%	0.4%
$10\times$ scenario	0%	0%	8%	38%	14%	8%	16%	3%	4%	8%	1%

The BGL_{222} and BGL_{221} fits have the largest differential rates near zero recoil ($w = 1$), corresponding to the largest extracted values of $|V_{cb}|$.

The value of $|V_{cb}|$ extracted from the BGL_{122} fit to the 2017 Belle unfolded measurement [1] is more than 1σ smaller than in the six-parameter BGL_{222} fit to the same data. This raises several questions: Would a BGL_{222} fit to the 2018 Belle measurement [2] find a larger value of $|V_{cb}|$ than that in Eq. (2b), closer to its inclusive determination? The consistency of the fitted BGL_{122} coefficients from the 2017 and 2018 Belle measurements is only at about the 2σ level for $\tilde{a}_0$.

Also shown in Fig. 3 are the fit results for the form factor ratios $R_{1,2}(w)$. The BGL_{222} fit to the tagged Belle measurement [1] indicated a substantial deviation from heavy quark symmetry, in particular for the R_1 form factor ratio [20]. The central values, for fixed quark mass parameters, at order $\mathcal{O}(\Lambda_{\mathrm{QCD}}/m_{c,b}, \alpha_s)$, are [20]

$$R_1(1) = 1.34 - 0.12\eta(1) + \cdots,$$
$$R_1'(1) = -0.15 + 0.06\eta(1) - 0.12\eta'(1) + \cdots, \qquad (10)$$

where $\eta(w)$ is a ratio of a subleading and the leading Isgur-Wise function. With $\eta(1)$ and $\eta'(1)$ of order unity, $R_1(1)$ cannot be much below 1, and $|R_1'(1)|$ cannot be large, without a breakdown of heavy quark symmetry. Preliminary lattice QCD calculations [31,32] also do not indicate $\mathcal{O}(1)$ violations of heavy quark symmetry. Figure 3 shows that the BGL_{122} fit exhibits better agreement with heavy quark symmetry expectations for $R_1(w)$. However, this likely arises because $R_1(w) \propto (w+1)g/f$, so setting $a_1 = 0$ constrains the shape of the numerator. By contrast, the BGL_{212}, BGL_{221}, and BGL_{222} fits prefer $a_1 \neq 0$, and yield $R_1(w)$ in some tension with heavy quark symmetry and lattice QCD.

V. TOY STUDIES

To validate the prescription outlined above, and to demonstrate that it yields an unbiased value of $|V_{cb}|$, we carried out a toy MC study using ensembles of pseudodata sets. These were generated using the BGL_{333} parametrization, i.e., with nine coefficients. The six lower-order coefficients $\{\tilde{a}_{0,1}, \tilde{b}_{0,1}, \tilde{c}_{1,2}\}$ were chosen to be identical to the BGL_{222} fit results of Fig. 2. The third-order terms $\{\tilde{a}_2, \tilde{b}_2, \tilde{c}_3\}$ were chosen according to two different scenarios: Either 1 or 10 times the size of the $\{\tilde{a}_1, \tilde{b}_1, \tilde{c}_2\}$ coefficients in the BGL_{222} fit, as shown in Table II. We call these the "$1\times$" and "$10\times$" scenarios, respectively. Ensembles were constructed as follows: First, predictions for the 40 bins of the tagged measurement [1] were produced. Ensembles of pseudodata sets were then generated using the full experimental covariance, assuming Gaussian errors, and then each pseudodata set was fit according to the nested hypothesis test prescription.

The frequency with which particular BGL_{ijk} parametrizations are selected are shown in Table III, for both the

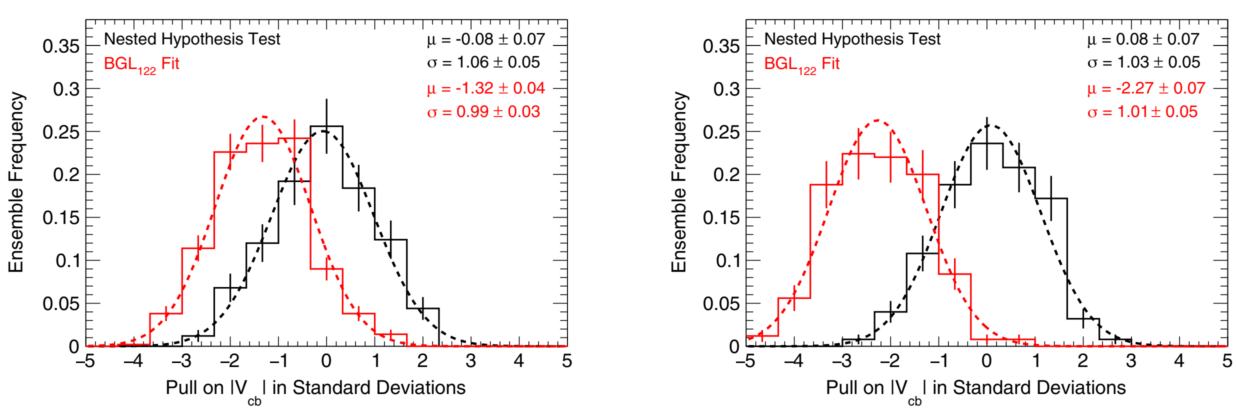

FIG. 4. The pull constructed from a large ensemble of pseudoexperiments using third-order terms of the $1\times$ scenario (left plot) and $10\times$ scenario (right plot) described in the text. The pull of the fits selected by the nested hypothesis prescription (black) show no bias or undercoverage of uncertainties. Also shown in red is the pull from a BGL_{122} fit, showing a large bias on the value of $|V_{cb}|$. Mean (μ) and standard deviation (σ) from normal distributions fitted to the ensembles are also provided.

$1\times$ and $10\times$ scenarios. For each selected fit hypothesis, the recovered value, $|V_{cb}|_{\rm rec}$, and the associated uncertainty, σ, may then be used to construct a pull, i.e., the normalized difference $(|V_{cb}|_{\rm rec} - |V_{cb}|_{\rm true})/\sigma$, where $|V_{cb}|_{\rm true}$ is the "true" value used to construct the ensembles. If a fit or a procedure is unbiased, the corresponding pull distribution should follow a standard normal distribution (mean of zero, standard deviation of unity). In Fig. 4, the pull distributions for both the $1\times$ and $10\times$ scenarios are shown and compared to that of the BGL_{122} parametrization. One sees that the nested hypothesis test proposed in this paper selects fit hypotheses that provide unbiased values for $|V_{cb}|$ in both scenarios. However, the BGL_{122} fit shows significant biases. In the ensemble tests, the BGL_{122} fits have mean χ^2 values of 41.0 and 56.6, respectively (with 35 d.o.f.). For the $1\times$ scenario, this produces an acceptable fit probability on average. Nonetheless, the recovered value of $|V_{cb}|$ is biased by about 1.3σ.

VI. CONCLUSIONS

We studied the differences of the determinations of $|V_{cb}|$ from exclusive semileptonic $B \to D^*\ell\nu$ decays, depending on the truncation order of the BGL parametrization of the form factors used to fit the measured differential decay distributions. Since the 2018 untagged Belle measurement [2] used a five-parameter BGL fit, Refs. [14,20] used a six-parameter fit, and Refs. [13,22] used an eight-parameter one, we explored differences between the five-, six-, seven-, and eight-parameter fits.

We proposed using nested hypothesis tests to determine the optimal number of fit parameters. For the 2017 Belle analysis [1], six parameters are preferred. Including additional fit parameters only improves χ^2 marginally. Comparing the result of the BGL_{122} fit used in the 2018 untagged Belle analysis [2] to the corresponding fit to the 2017 tagged Belle measurement [1], up to 2σ differences occur, including in the values of $|V_{cb}|$. This indicates that more precise measurements are needed to resolve tensions between various $|V_{cb}|$ determinations, and that the truncation order of the BGL expansion of the form factors has to be chosen with care, based on data.

We look forward to more precise experimental measurements, more complete fit studies inside the experimental analysis frameworks, as well as better understanding of the composition of the inclusive semileptonic rate as a sum of exclusive channels [33,34]. Improved lattice QCD results, including finalizing the form factor calculations in the full w range [31,32], are also expected to be forthcoming. These should all contribute to a better understanding of the determinations of $|V_{cb}|$ from exclusive and inclusive semileptonic decays, which is important for CKM fits, new physics sensitivity, ϵ_K, and rare decays.

ACKNOWLEDGMENTS

We thank Toru Iijima for organizing the KMI workshop "Hints for New Physics in Heavy Flavors," and thank him and Marina Artuso, Ben Grinstein, Shoji Hashimoto, Aneesh Manohar, Sheldon Stone, and Mike Williams for useful questions and conversations. F. B. was supported by DFG Emmy-Noether Grant No. BE 6075/1-1. Z. L. and D. R. were supported in part by the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. D. R. was also supported in part by NSF Grant No. PHY-1720252.

[1] A. Abdesselam *et al.* (Belle Collaboration), arXiv:1702.01521.
[2] A. Abdesselam *et al.* (Belle Collaboration), arXiv:1809.03290.
[3] N. Isgur and M. B. Wise, Phys. Lett. B **232**, 113 (1989).
[4] N. Isgur and M. B. Wise, Phys. Lett. B **237**, 527 (1990).
[5] M. A. Shifman and M. B. Voloshin, Yad. Fiz. **47**, 801 (1988) [Sov. J. Nucl. Phys. **47**, 511 (1988)].
[6] S. Nussinov and W. Wetzel, Phys. Rev. D **36**, 130 (1987).
[7] E. Eichten and B. R. Hill, Phys. Lett. B **234**, 511 (1990).
[8] H. Georgi, Phys. Lett. B **240**, 447 (1990).
[9] M. E. Luke, Phys. Lett. B **252**, 447 (1990).
[10] A. F. Falk, B. Grinstein, and M. E. Luke, Nucl. Phys. **B357**, 185 (1991).
[11] C. G. Boyd, B. Grinstein, and R. F. Lebed, Nucl. Phys. **B461**, 493 (1996).
[12] C. G. Boyd, B. Grinstein, and R. F. Lebed, Phys. Rev. D **56**, 6895 (1997).
[13] D. Bigi, P. Gambino, and S. Schacht, Phys. Lett. B **769**, 441 (2017).
[14] B. Grinstein and A. Kobach, Phys. Lett. B **771**, 359 (2017).
[15] I. Caprini, L. Lellouch, and M. Neubert, Nucl. Phys. **B530**, 153 (1998).
[16] Y. Amhis *et al.* (HFLAV Collaboration), Eur. Phys. J. C **77**, 895 (2017).
[17] C. Bourrely, B. Machet, and E. de Rafael, Nucl. Phys. **B189**, 157 (1981).
[18] C. G. Boyd, B. Grinstein, and R. F. Lebed, Phys. Rev. Lett. **74**, 4603 (1995).
[19] C. G. Boyd, B. Grinstein, and R. F. Lebed, Phys. Lett. B **353**, 306 (1995).
[20] F. U. Bernlochner, Z. Ligeti, M. Papucci, and D. J. Robinson, Phys. Rev. D **96**, 091503 (2017).

[21] D. Bigi, P. Gambino, and S. Schacht, J. High Energy Phys. 11 (2017) 061.
[22] S. Jaiswal, S. Nandi, and S. K. Patra, J. High Energy Phys. 12 (2017) 060.
[23] M. Neubert, Z. Ligeti, and Y. Nir, Phys. Lett. B **301**, 101 (1993).
[24] M. Neubert, Z. Ligeti, and Y. Nir, Phys. Rev. D **47**, 5060 (1993).
[25] Z. Ligeti, Y. Nir, and M. Neubert, Phys. Rev. D **49**, 1302 (1994).
[26] F. U. Bernlochner, Z. Ligeti, M. Papucci, and D. J. Robinson, Phys. Rev. D **95**, 115008 (2017).
[27] A. V. Manohar and M. B. Wise, Cambridge Monogr. Part. Phys., Nucl. Phys., Cosmol. **10**, 1 (2000).
[28] J. A. Bailey *et al.* (Fermilab Lattice and MILC Collaborations), Phys. Rev. D **89**, 114504 (2014).
[29] A. Sirlin, Nucl. Phys. **B196**, 83 (1982).
[30] S. S. Wilks, Ann. Math. Stat. **9**, 60 (1938).
[31] V. Aviles-Casco *et al.*, EPJ Web Conf. **175**, 13003 (2018).
[32] T. Kaneko, Y. Aoki, B. Colquhoun, H. Fukaya, and S. Hashimoto (JLQCD Collaboration), Proc. Sci., LATTICE2018 (**2018**) 311.
[33] F. U. Bernlochner, Z. Ligeti, and S. Turczyk, Phys. Rev. D **85**, 094033 (2012).
[34] F. U. Bernlochner, D. Biedermann, H. Lacker, and T. Luck, Eur. Phys. J. C **74**, 2914 (2014).