PROCEEDINGS OF SCIENCE

Current progress on the semileptonic form factors for $\bar{B} \to D^* \ell \bar{\nu}$ decay using the Oktay-Kronfeld action

Tanmoy Bhattacharya,[a,1] Benjamin J. Choi,[b,1,*] Rajan Gupta,[a,1] Yong-Chull Jang,[c,1] Seungyeob Jwa,[b,1] Sunghee Kim,[b,1] Sunkyu Lee,[b,1] Weonjong Lee,[b,1] Jaehoon Leem,[d,1] Jeonghwan Pak[b,1] and Sungwoo Park[e,1]

[a] *Theoretical Division T-2, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*

[b] *Lattice Gauge Theory Research Center, CTP and FPRD, Department of Physics and Astronomy, Seoul National University, Seoul 08826, South Korea*

[c] *Physics Department, Brookhaven National Lab, Upton, NY 11973, USA*

[d] *Computational Science and Engineering Team, Innovation Center, Samsung Electronics, Hwaseong, Gyeonggi-do 18448, South Korea*

[e] *Lawrence Livermore National Lab, 7000 East Ave, Livermore, CA 94550, USA*

E-mail: benjaminchoi@snu.ac.kr, wlee@snu.ac.kr

We present recent progress in calculating the semileptonic form factors $h_{A_1}(w)$ for the $\bar{B} \to D^* \ell \bar{\nu}$ decays. We use the Oktay-Kronfeld (OK) action for the charm and bottom valence quarks and the HISQ action for light quarks. We adopt the Newton method combined with the scanning method to find a good initial guess for the χ^2 minimizer in the fitting of the 2pt correlation functions. The main advantage is that the Newton method lets us to consume all the time slices allowed by the physical positivity. We report the first, reliable, but preliminary results for $h_{A_1}(w)/\rho_{A_1}$ at zero recoil ($w = 1$). Here we use a MILC HISQ ensemble ($a = 0.12$ fm, $M_\pi = 220$ MeV, and $N_f = 2 + 1 + 1$ flavors).

The 40th International Symposium on Lattice Field Theory (Lattice 2023)
July 31st - August 4th, 2023
Fermi National Accelerator Laboratory

[1] The LANL-SWME Collaboration
[*] Speaker

https://pos.sissa.it/

PoS(LATTICE2023)245

1. Introduction

We present update of data analysis on the 2pt and 3pt correlation functions to obtain the semileptonic form factors for the $\bar{B} \to D^* \ell \bar{\nu}$ decays. We adopt the Newton method combined with the scanning method [1] to find a good initial guess for the χ^2 minimizer in the fitting of the 2pt correlation functions. We find that the multiple time slice combinations help to distinguish the global minimum of χ^2 and its local minima reliably. The Newton method leads to a self-consistent fit which consumes all the time slices allowed by the physical positivity [2, 3]. The results of data analysis on the 2pt correlation functions are used as inputs to the fitting of the 3pt correlation functions. As a result, we report the first, reliable, but preliminary results for $h_{A_1}(w)/\rho_{A_1}$ at zero recoil ($w = 1$), obtained using the MILC HISQ ensemble in Table 1.

a (fm)	N_f	$N_s^3 \times N_t$	M_π (MeV)	am_l	am_s	am_c
0.1184(10)	$2+1+1$	$32^3 \times 64$	216.9(2)	0.00507	0.0507	0.628

Table 1: Details on the MILC HISQ ensemble used for the numerical study [4].

2. Flow chart of the data analysis

The 2-point (2pt) correlation function is defined as [5],

$$C(t) = \sum_{\alpha=1}^{4} \sum_{\mathbf{x}} \left\langle O_\alpha^\dagger(t,\mathbf{x}) O_\alpha(0) \right\rangle , \tag{1}$$

where an interpolating operator for heavy-light mesons $O_\alpha(t,\mathbf{x})$ is

$$O_\alpha(t,\mathbf{x}) = \left[\bar{\psi}(t,\mathbf{x})\,\Gamma\,\Omega(t,\mathbf{x})\right]_\alpha \chi(t,\mathbf{x}) , \qquad \Omega(t,\mathbf{x}) \equiv \gamma_1^{x_1}\gamma_2^{x_2}\gamma_3^{x_3}\gamma_4^{t} . \tag{2}$$

Here, $\Gamma = \gamma_5$ ($\Gamma = \gamma_j$) for the pseudoscalar (vector) meson. Here, ψ is a heavy quark field in the OK action [6], χ is a light quark field in the HISQ staggered action [7], and the subscript α represents taste degrees of freedom for staggered quarks.

We measure meson propagators (*i.e.* 2pt correlators in Eq. (1)) on the lattice. In the lattice QCD, the lattice Hilbert space consists of states of quarks and gluons, but the physical Hilbert space consists of states of hadrons. Hence, in order to extract physical information on hadronic states from the 2pt correlator, we use the spectral decomposition in the physical Hilbert space to obtain the fitting functional form $f(t)$. The fitting function of the $m+n$ fit is

$$\begin{aligned} f(t) &= g(t) + g(T-t) , \\ g(t) &= A_0\, e^{-E_0 t} \Big[1 + R_2\, e^{-\Delta E_2 t} \Big(1 + R_4\, e^{-\Delta E_4 t} \Big(\cdots \Big(1 + R_{2m-2}\, e^{-\Delta E_{2m-2} t} \Big) \cdots \Big)\Big) \\ &\qquad -(-1)^t R_1\, e^{-\Delta E_1 t} \Big(1 + R_3\, e^{-\Delta E_3 t} \Big(\cdots \Big(1 + R_{2n-1}\, e^{-\Delta E_{2n-1} t} \Big) \cdots \Big)\Big)\Big] \end{aligned} \tag{3}$$

where $\Delta E_i \equiv E_i - E_{i-2}$, $E_{-1} \equiv E_0$, $R_i \equiv \dfrac{A_i}{A_{i-2}}$ and $A_{-1} \equiv A_0$. The subscript $_0$ represents the ground state in the physical Hilbert space. Hence, A_0 and E_0 represents the amplitude and energy

Symbol	Description	Example
$\bar{\lambda} \pm \sigma(\lambda)$	average and error of a fit parameter λ	$\bar{A}_0 \pm \sigma(A_0), \ldots$
$\lambda_p \pm \sigma_p(\lambda)$	prior information on a fit parameter λ	$[A_0]_p \pm \sigma_p(A_0), \ldots$
$\sigma_p^{\text{mf}}(\lambda)$	$\sigma_p(\lambda)$ obtained by the maximal fluctuation of data	$\sigma_p^{\text{mf}}(A_0), \ldots$
$\sigma_p^{\text{sc}}(\lambda)$	$\sigma_p(\lambda)$ obtained by the signal cut ($\sigma_p^{\text{sc}}(\lambda) = \lvert\bar{\lambda}_p\rvert$)	$\sigma_p^{\text{sc}}(A_0), \ldots$
$\sigma_p^{\max}(\lambda)$	$\min\big(\sigma_p^{\text{mf}}(\lambda),\ \sigma_p^{\text{sc}}(\lambda)\big)$	$\sigma_p^{\max}(A_0), \ldots$
$\sigma_p^{\text{opt}}(\lambda)$	optimal prior width of λ in Eq. (5)	$\sigma_p^{\text{opt}}(A_0), \ldots$
$\sigma_\sigma(\lambda)$	error of error for λ, that is, error of $\sigma(\lambda)$	$\sigma(A_0) \pm \sigma_\sigma(A_0), \ldots$

(a) Notation and convention

Symbol	Description
$\sigma(A_0; \max)$	$\sigma(A_0)$ when we set $\sigma_p(A_0) = \sigma_p^{\max}(A_0)$ and $\sigma_p(E_0) = \sigma_p^{\max}(E_0)\}$
$\sigma(E_0; \max)$	$\sigma(E_0)$ when we set $\sigma_p(A_0) = \sigma_p^{\max}(A_0)$ and $\sigma_p(E_0) = \sigma_p^{\max}(E_0)\}$

(b) Notation used for stability tests

Table 2: Notation and convention used in this article.

of the ground state. In the $m + n$ fit, m (n) is the number of even (odd) time-parity states, which are kept in the fitting, while higher excited states in each time-parity channel are truncated.

Our notation and convention is described in Table 2.

To determine fit parameters, A_0, E_0, $\{R_j, \Delta E_j\}$, we use sequential Bayesian method. We obtain fit parameters which minimizes the χ^2. Using the fit function given in Eq. (3), we adopt the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [8–11] for the χ^2 minimizer, which belongs to the quasi-Newton method. The quasi-Newton method needs an initial guess for the fit parameters. Here, we denote the initial guess as A_0^g, E_0^g, $\{R_j^g, \Delta E_j^g\}$. The superscript ^{g} represents the "initial guess". A good initial guess reduces the number of iterations in the BFGS minimizer, which saves the computing cost dramatically [1]. In order to find a good initial guess directly from the data, we use the multi-dimensional Newton method [12, 13] combined with the scanning method [1]. We present the flow chart for the sequential Bayesian method in the following to describe the logistics.

Step-1 Do the 1st fit. [*e.g.*][1] Do 1+0 fit with 2 parameters: $\{A_0, E_0\}$

Step-2 Feed the previous fit results as prior information for the next fit: $\lambda_p = \bar{\lambda}$, $\sigma_p(\lambda) = \sigma_p^{\max}(\lambda)$. We do not impose any constraint on new fit parameters which are not included in the previous fit. [*e.g.*] Results for the 1+0 fit are used as prior information on the 1+1 fit such that $[A_0]_p = \bar{A}_0$, $\sigma_p(A_0) = \sigma_p^{\max}(A_0) = \sigma_p^{\text{sc}}(A_0)$ and $[E_0]_p = \bar{E}_0$, $\sigma_p(E_0) = \sigma_p^{\max}(E_0) = \sigma_p^{\text{mf}}(E_0)$ with no constraint on R_1 and ΔE_1.

In order to give a general picture of our methodology, let us consider the $m+n$ fit, in which we want to determine $N = 2(m + n)$ fit parameters. Hence, we should select N time slices such as $\{t_1, t_2, \ldots, t_N\}$, to determine the initial guess A_0^g, E_0^g, $\{R_j^g, \Delta E_j^g\}$ by solving the following

[1] It stands for *exempli gratia* in Latin, which means *for example*.

equations using the Newton method.

$$q(t_j) \equiv \frac{f(t_j) - C(t_j)}{C(t_j)} = 0\,, \quad \text{with} \quad j = 1, \cdots, N \tag{4}$$

where $C(t)$ is data for the 2pt correlation functions in Eq. (1) and $f(t)$ is the fit function in Eq. (3).

Step-2A We find all the possible combinations of N time slices which satisfy the following conditions.

- We choose N time slices within the fit range: $t_{\text{min}} \le t \le t_{\text{max}}$.
- t_{min} should be included.
- The number of even time slices should be equal to that of odd time slices in order to avoid a bias.

Step-2B Use the Newton method for each time slice combination to obtain a good initial guess for the χ^2 minimizer:

Step-2B1 Take i-th time slice combination ($1 \le i \le N_c$, N_c is the total number of the time slice combinations).

Step-2B2 Recycle fit results from the previous fit ($[m-1]+n$ fit or $m+[n-1]$ fit) to set part of the initial guess for the Newton method.

Step-2B3 Use the scanning method [1] to set the remaining part of the initial guess for the Newton method.

Step-2B4 Run the Newton method.

Step-2B5 If the Newton method finds roots, then save them. If it fails, discard the i-th time slice combination. [*e.g.*] The failure rate is about $\approx 21\%$ for the 1+1 fit.

Step-2B6 Take the next $\big($*i.e.* $(i+1)$-th$\big)$ time slice combination, and repeat the loop (**Step-2B1**–**Step-2B6**) until we consume all the time slice combinations.

Step-2C Perform the least χ^2 fitting with the initial guess obtained by the Newton method.

Step-2D Sort results for χ^2, check the χ^2 distribution, and find out whether the χ^2 minimizer converges to the global minimum or a local minimum.

Step-3 Perform stability tests to obtain optimal prior widths. [*e.g.*] Determine $\{\sigma_p^{\text{opt}}(A_0), \sigma_p^{\text{opt}}(E_0)\}$.

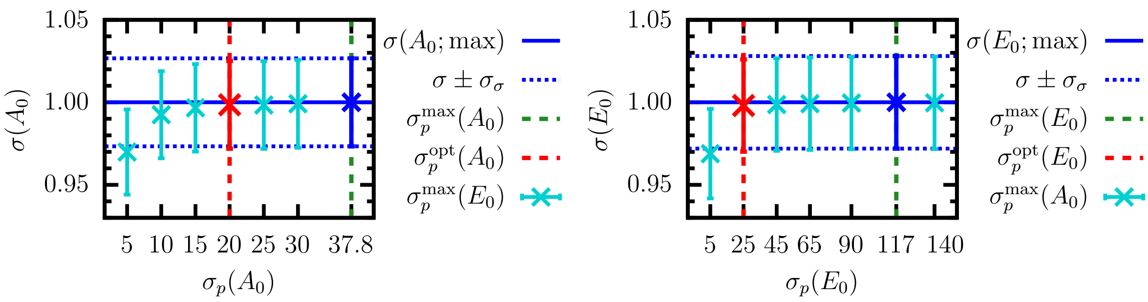
Step-4 Save the current fitting results ([*e.g.*] 1+1 fit) into the 1st fitting.

Step-5 Take the next fitting ([*e.g.*] 2+1 fit) as the current fitting.

Step-6 Go back to **Step-2** and repeat the loop until we consume all the time slices allowed by the physical positivity [2, 3]. [*e.g.*] 1+0 fit $\longrightarrow$ 1+1 fit $\longrightarrow$ 2+1 fit $\longrightarrow$ 2+2 fit $\longrightarrow \cdots$.

The stability test condition for the sequential Bayesian method is

(i) We have freedom to adjust λ_p so that $|\bar{\lambda} - \lambda_p| \ll \sigma_\sigma(\lambda)$. One way to achieve this criterion is $|\bar{\lambda} - \lambda_p| < 10^{-4}\sigma(\lambda) \ll \sigma_\sigma(\lambda)$. Here the scaling factor 10^{-4} is obtained empirically for our statistical sample of ≈ 1000 gauge configurations.



(a) $\sigma_p^{\mathrm{opt}}(A_0)$ in the unit of $\sigma(A_0;\max)$.

(b) $\sigma_p^{\mathrm{opt}}(E_0)$ in the unit of $\sigma(E_0;\max)$.

Figure 1: Results of the stability tests to find optimal prior widths for the 2+1 fit.

(ii) $\sigma_p(\lambda)$ should not disturb $\sigma(\lambda)$ such that $|\sigma(\lambda,\sigma_p(\lambda)) - \sigma(\lambda,\max)| \ll \sigma_\sigma(\lambda)$.

(iii) Based on condition (i) and (ii), we find the optimal prior width σ_p^{opt}. Here, σ_p^{opt} is the minimum value of the prior width which does not disturb $\sigma(\lambda)$ of the fitting results such that

$$\sigma_p^{\mathrm{opt}} = \min(\sigma_p) \text{ for } \forall\sigma_p \in \{\sigma_p | \bar{\lambda}(\sigma_p) \doteq \lim_{\sigma_p^t\to\infty} \bar{\lambda}(\sigma_p^t),\ \sigma(\lambda,\sigma_p) \doteq \lim_{\sigma_p^t\to\infty} \sigma(\lambda,\sigma_p^t)\} \tag{5}$$

Here the new equal symbol ($\doteq$) means that they are equal within the statistical uncertainty of $\sigma_\sigma(\lambda)$.

Let us explain the stability test (Step-3), using the $2+1$ fit as an example. In Fig. 1, we describe how to set the optimal prior widths $\sigma_p^{\mathrm{opt}}(A_0)$ and $\sigma_p^{\mathrm{opt}}(E_0)$ in the $2+1$ fit. In the case of $\sigma(A_0)$ in Fig. 1 (a), we plot $\sigma(A_0)$ as a function of $\sigma_p(A_0)$ in the unit of $\sigma(A_0;\max)$ while we fix $\sigma_p(E_0)$ to its maximum value: $\sigma_p(E_0) = \sigma_p^{\max}(E_0) = \sigma_p^{\mathrm{mf}}(E_0)$. Here we find that $\sigma_p^{\mathrm{opt}}(A_0) = 20 \times [\sigma(A_0;\max)]$, which corresponds to the red (dashed) lines and red cross symbol in Fig. 1 (a). Here the blue cross symbol and green (dashed) lines represents $\sigma(A_0;\max)$ and the blue dotted lines represents the error of error $\sigma_\sigma(A_0)$ of $\sigma(A_0;\max)$.

In Fig. 1 (b), we present the same kind of a plot for $\sigma(E_0)$ with the same color convention as Fig. 1 (a). Here we find that $\sigma_p^{\mathrm{opt}}(E_0) = 25 \times [\sigma(E_0;\max)]$.

3. Application of the Newton method to the 2+2 fit

3.1 Example for Step-2A

As explained in Ref. [1], we use the multi-dimensional Newton method [12, 13] to obtain a good initial guess for the χ^2 minimizer in the $m+n$ correlator fit. For the $2+2$ fit, we should determine 8 fit parameters: A_0, E_0, $\{R_j, \Delta E_j\}$ with $j = 1,2,3$. Hence, we need an initial guess: A_0^g, E_0^g, $\{R_j^g, \Delta E_j^g\}$ $(j = 1,2,3)$.

In order to find an initial guess, we need 8 time slices (T_8) so that we can use the Newton method to find roots:

$$q(t_k) \equiv \frac{f(t_k) - C(t_k)}{C(t_k)} = 0\,, \quad k = 1,2,\ldots,8 \tag{6}$$

where $t_k \in T_8$, $f(t)$ is the fitting function, and $C(t)$ is the 2pt correlator data. The 8 time slices in T_8 should be chosen within the fit range $t_{\min} \le t \le t_{\max}$ with $t_{\min} = 3$ and $t_{\max} = 30$. It is required

ID	N_r	χ^2/d.o.f.	note
`2+2/G`	1000	0.4091(82)	**global minimum**
`2+2/L1`	167	0.6093(88)	$\Delta E_3 < 0$
`2+2/L2`	54	3.766(26)	$R_3 < O(10^{-9}) \approx 0$ or $R_3 < 0$

Table 3: Patterns of the χ^2 distribution. Here, N_r means the number of roots obtained by the Newton method. Here `G` represents the global minimum and `L` represents the local minima.

to set t_1 to $t_1 = t_{\rm min}$. The number of even (odd) time slices is 14 (13) except for $t_1 = t_{\rm min}$. Hence, the total number of the possible combinations for T_8 is 286,286.

$$ {}_{14}C_4 \times {}_{13}C_3 = 286{,}286 \tag{7} $$

3.2 Description of Step-2B

In **Step-2B**, we run the Newton method. For example, in 2+2 fit, we select a time slice combination out of the 286,286 combinations [**Step-2B1**]. We need another initial guess as input to run the Newton method: A_0^{gn}, E_0^{gn}, $\{R_j^{gn}, \Delta E_j^{gn}\}$ where the superscript ^{gn} indicates the initial guess for the Newton method. We use fit results from the 2+1 fit to set up A_0^{gn}, E_0^{gn}, $\{R_j^{gn}, \Delta E_j^{gn}\}$ $(j = 1, 2)$ [**Step-2B2**]. We use the scanning method in Ref. [1] to determine R_3^{gn} and ΔE_3^{gn} [**Step-2B3**].

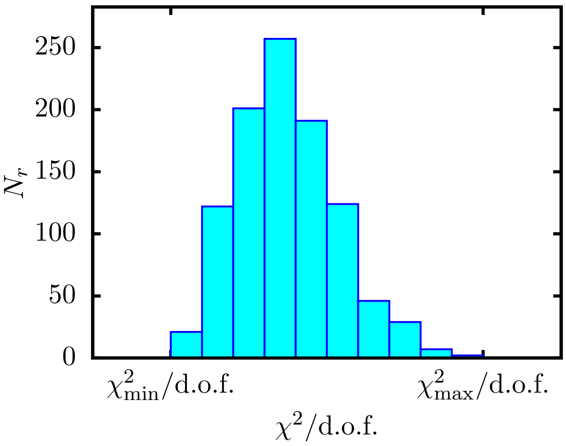
Now, we run the Newton method to find roots: A_0^g, E_0^g, $\{R_j^g, \Delta E_j^g\}$ [**Step-2B4**]. If the Newton method finds a root, save them. Otherwise, discard it [**Step-2B5**]. It turns out that the Newton method can find 16,574 roots out of the 286,286 combinations, while the rest fails. We select 1,242 roots randomly out of the whole 16,574 roots in order to monitor statistics for the χ^2 distribution.

3.3 Description of Step-2C and Step-2D

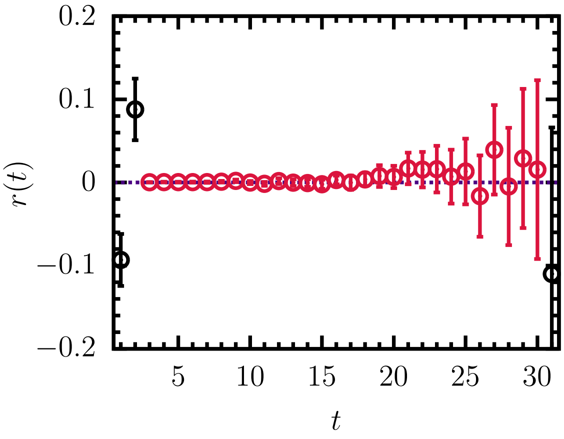
For each root that the Newton method can find successfully, we use it as an input to perform the least χ^2 fitting [**Step-2C**]. For each root, we determine the statistics for A_0, E_0, $\{R_j, \Delta E_j\}$ $(j = 1, 2, 3)$ and χ^2/d.o.f, using the jackknife resampling. Using the 1,242 roots, we check whether the χ^2 minimizer reaches the global minimum or local minima.

We summarize patterns for the χ^2 distribution in the Table 3. Among 1,242 roots, 1,000 roots converges to the global minimum (pattern ID = `2+2/G`). We find two local minima of χ^2: the pattern ID = `2+2/L1` (167 roots) and the pattern ID = `2+2/L2` (54 roots). The `2+2/L1` pattern gives consistent results of $\Delta E_3 = -0.409(55) < 0$, which is definitely unphysical and wrong. The `2+2/L2` pattern gives consistent results of $R_3 < O(10^{-9}) \approx 0$ or $R_3 < 0$, which are unphysical and wrong. There are 21 values of the χ^2/d.o.f. between the `2+2/L1` and `2+2/L2` patterns, which also gives wrong results for R_3 or ΔE_3. Table 3 shows that we can find the global minimum of the χ^2 distribution with the Newton method reliably. In addition, we find that the local minima of the χ^2 distribution always come up with unphysical (= wrong) results for R_3 or ΔE_3.

In Fig. 2 (a), we show the histogram of the χ^2/d.o.f. distribution of the `2+2/G` pattern. This indicates that the χ^2 minimizer find the global minimum with about $\approx 80\%$ probability which indicates that once out of five times the χ^2 minimizer converges to local minima. Hence, this is a clear advantage of using the Newton method to find a good initial guess for the χ^2 minimizer. Out of the set of multiple roots of the Newton method, a subset find the global minimum for the



(a) Histogram of the χ_d^2 distribution. We use the same notation as in Table 3. Here $\chi_d^2 \equiv \chi^2/\text{d.o.f.}$, $\chi^2_{\min} \equiv \min(\chi^2)$, and $\chi^2_{\max} \equiv \max(\chi^2)$.



(b) Residual $r(t) = \dfrac{f(t) - C(t)}{|C(t)|}$. Here the red color represents the fit range ($3 \le t \le 30$). The errors are purely statistical.

Figure 2: χ^2 distribution and residual coming from the data analysis on the B-meson propagators.

χ^2 distribution, and another subset reach the local minima, which we can discard without loss of generality.

In Fig. 2 (a), $\Delta\chi^2/\text{d.o.f.} = \chi^2_{\max}/\text{d.o.f.} - \chi^2_{\min}/\text{d.o.f.} \cong 10^{-8}$ is significantly less than the statistical error ($\cong 0.0082$) of the χ^2/d.o.f.. This indicates that the χ^2 distribution has sharp peak.

The residual plot of B meson 2pt correlator for the 2 + 2 fit is given in Fig. 2 (b). Note that we consumed up all the possible time slices allowed by the physical positivity [2, 3].

4. Preliminary result on $h_{A_1}(w = 1)/\rho_{A_1}$: $B \to D^*\ell\nu$ form factor at zero recoil

The semileptonic form factor $h_{A_1}(w)$ at zero recoil (*i.e.* $w = 1$) can be obtained with the Hashimoto ratio [14]:

$$h_{A_1}(w=1) = \rho_{A_1}\sqrt{\frac{\langle D_0^*|A_j^{cb}|B_0\rangle_\text{L}\,\langle B_0|A_j^{bc}|D_0^*\rangle_\text{L}}{\langle D_0^*|V_4^{cc}|D_0^*\rangle_\text{L}\,\langle B_0|V_4^{bb}|B_0\rangle_\text{L}}}\,, \qquad \rho_{A_1} = \sqrt{\frac{Z_A^{cb}\,Z_A^{bc}}{Z_V^{cc}\,Z_V^{bb}}}\,. \tag{8}$$

The one-loop matching calculation of ρ_{A_1} is underway [15]. Here we present preliminary results on blind (*i.e.* $\rho_{A_1} = 1$) $h_{A_1}(w = 1)$ in this work. The subscript $_0$ in Eq. (8) represents the ground states at zero momentum ($\vec{p}_X = 0$ with $X = B,\ D^*$). We want to extract the four ground state matrix elements: $\langle D_0^*|A_j^{cb}|B_0\rangle_\text{L}$, $\langle B_0|A_j^{bc}|D_0^*\rangle_\text{L}$, $\langle D_0^*|V_4^{cc}|D_0^*\rangle_\text{L}$, and $\langle B_0|V_4^{bb}|B_0\rangle_\text{L}$ from the 3pt correlation functions. The 3pt correlation functions are calculated on the lattice and so the Hilbert space consists in quark and gluon states. The Hilbert space for physical observables such as the matrix elements consists in hadronic states. For example, when we fit the 2pt correlation functions for B-meson propagators, we obtain results for A_0, E_0 (*i.e.* information on the B_0 ground state), R_1, ΔE_1 (*i.e.* the B_1 excited state with odd time-parity), R_2, ΔE_2 (*i.e.* the B_2 excited state with even time-parity) and so on. We can obtain similar results for D^*-meson propagators. The fitting functional form for the 3pt correlation functions calculated on the lattice in the $B \to D^*$ channel is

$$f_{T_\text{sep}}^{B\to D^*}(t) = \langle D_0^*|\,A_j^{cb}\,|B_0\rangle_\text{L}\;k_0^{D^*}(t)\,k_0^B\,(T_\text{sep}-t) + \langle D_0^*|\,A_j^{cb}\,|B_2\rangle_\text{L}\;k_0^{D^*}(t)\,k_2^B\,(T_\text{sep}-t)$$

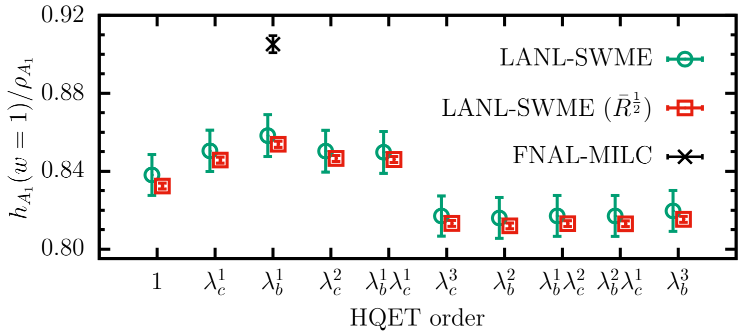


Figure 3: $h_{A_1}(w=1)/\rho_{A_1}$ in the order of the HQET power counting. Here $\lambda_c \cong 1/5$, and $\lambda_b \cong 1/17$.

$$
\begin{aligned}
&+ \langle D_1^*| A_j^{cb} |B_1\rangle_{\mathrm{L}}\, k_1^{D^*}(t)\, k_1^B(T_{\mathrm{sep}}-t) + \langle D_2^*| A_j^{cb} |B_0\rangle_{\mathrm{L}}\, k_2^{D^*}(t)\, k_0^B(T_{\mathrm{sep}}-t) \\
&+ \langle D_1^*| A_j^{cb} |B_3\rangle_{\mathrm{L}}\, k_1^{D^*}(t)\, k_3^B(T_{\mathrm{sep}}-t) + \langle D_2^*| A_j^{cb} |B_2\rangle_{\mathrm{L}}\, k_2^{D^*}(t)\, k_2^B(T_{\mathrm{sep}}-t) \\
&+ \langle D_3^*| A_j^{cb} |B_1\rangle_{\mathrm{L}}\, k_3^{D^*}(t)\, k_1^B(T_{\mathrm{sep}}-t) + \langle D_3^*| A_j^{cb} |B_3\rangle_{\mathrm{L}}\, k_3^{D^*}(t)\, k_3^B(T_{\mathrm{sep}}-t)\ , \qquad (9)
\end{aligned}
$$

$$
\begin{aligned}
k_0^X(t) &= \sqrt{A_0^X}\, e^{-E_0^X t}\ , & k_1^X(t) &= \sqrt{A_0^X R_1^X}\, e^{-(E_0^X+\Delta E_1^X)\,t}\, (-1)^{t+1}\ , \\
k_2^X(t) &= \sqrt{A_0^X R_2^X}\, e^{-(E_0^X+\Delta E_2^X)\,t}\ , & k_3^X(t) &= \sqrt{A_0^X R_1^X R_3^X}\, e^{-(E_0^X+\Delta E_1^X+\Delta E_3^X)\,t}\, (-1)^{t+1}\ . \qquad (10)
\end{aligned}
$$

The $k_j^X(t)$ ($j=0,1,2,3$, $X=B,D^*$) comes from the fit results for the 2pt correlation functions for the B and D^* mesons. Hence, we determine the lattice matrix elements simply by a linear fit. As a result, we obtain $\langle D_0^*|A_j^{cb}|B_0\rangle_{\mathrm{L}}$. We can apply the same kind of fitting to the $D^* \to B$, $B \to B$, and $D^* \to D^*$ channels. As a results, we obtain the rest of the lattice matrix elements: $\langle B_0|A_j^{bc}|D_0^*\rangle_{\mathrm{L}}$, $\langle D_0^*|V_4^{cc}|D_0^*\rangle_{\mathrm{L}}$, $\langle B_0|V_4^{bb}|B_0\rangle_{\mathrm{L}}$.

In Fig. 3 we present results for $h_{A_1}(w=1)/\rho_{A_1}$. Here the green circles represent our results for $h_{A_1}(w=1)/\rho_{A_1}$ with no contamination from exited states, while the red squares represent those obtained using the $\bar{R}$ ratio [16] which include some contamination from excited states by construction. The black cross represents the FNAL-MILC result for $h_{A_1}(w=1)/\rho_{A_1}$ which is obtained using the $\bar{R}$ ratio with the Fermilab action for bottom and charm quarks, and the asqtad action for light quarks with $N_f = 2+1$ [16].

When we do the linear fit over the 3pt correlation functions, we could not use the full covariance fitting but the diagonal approximation [17] due to unwanted bias by strong correlation between different time slices. We find that off-diagonal elements of the correlation matrix $\rho(t_i, t_j)$ with $t_i \neq t_j$ are close to one. This issue needs further investigation.

Acknowledgments

We would like to thank Andreas Kronfeld, and Carlton Detar for helpful discussion on many issues on theory and fitting. We would like to thank the MILC collaboration and Chulwoo Jung for providing the HISQ lattice ensembles to us. The research of W. Lee is supported by the Mid-Career Research Program (Grant No. NRF-2019R1A2C2085685) of the NRF grant funded by the Korean government (MSIT). W. Lee would like to acknowledge the support from the KISTI supercomputing center through the strategic support program for the supercomputing application research (No. KSC-2018-CHA-0043, KSC-2020-CHA-0001, KSC-2023-CHA-0010). Computations were carried out in part on the DAVID cluster at Seoul National University.

References

[1] T. Bhattacharya *et al.* *PoS* **LATTICE2021** (2021) 136 [2204.05848].

[2] M. Luscher *Commun. Math. Phys.* **54** (1977) 283.

[3] M. Luscher and P. Weisz *Nucl. Phys. B* **240** (1984) 349.

[4] A. Bazavov *et al.* *Phys. Rev.* **D87** (2013) 054505 [1212.4768].

[5] A. Bazavov *et al.* *Phys. Rev.* **D85** (2012) 114506 [1112.3051].

[6] M.B. Oktay and A.S. Kronfeld *Phys. Rev.* **D78** (2008) 014504 [0803.0523].

[7] E. Follana *et al.* *Phys. Rev.* **D75** (2007) 054502 [hep-lat/0610092].

[8] C.G. Broyden *IMA Journal of Applied Mathematics* **6** (1970) 76.

[9] R. Fletcher *The Computer Journal* **13** (1970) 317.

[10] D. Goldfarb *Mathematics of Computation* **24** (1970) 23.

[11] D.F. Shanno *Mathematics of Computation* **24** (1970) 647.

[12] W.H. Press *et al.*, *Numerical Recipes*, Cambridge University Press, 3 ed. (2007), pp 477–483.

[13] C.G. Broyden *Mathematics of Computation* **19** (1965) 577.

[14] S. Hashimoto *et al.* *Phys. Rev. D* **61** (1999) 014502 [hep-ph/9906376].

[15] J.A. Bailey and S. Lee, in preparation.

[16] Fermilab Lattice and MILC collaboration *Phys. Rev. D* **89** (2014) 114504 [1403.0635].

[17] B. Yoon *et al.* *J. Korean Phys. Soc.* **63** (2013) 145 [1101.2248].