

## Extraction of bare form factors for $B_s \rightarrow K\ell\nu$ decays in nonperturbative HQET

Felix Bahr,<sup>\*</sup> Debasish Banerjee,<sup>\*,†,§</sup> Fabio Bernardoni,<sup>‡</sup> Mateusz Koren,<sup>\*</sup>  
 Hubert Simma<sup>\*</sup> and Rainer Sommer<sup>\*,†</sup>  
 (ALPHA Collaboration)

<sup>\*</sup>*John von Neumann Institute for Computing (NIC), DESY,  
 Platanenallee 6, D-15738 Zeuthen, Germany*

<sup>†</sup>*Institut für Physik, Humboldt-Universität zu Berlin, Newtonstr. 15, 12489 Berlin, Germany*

<sup>‡</sup>*Medizinische Fakultät, Carl Gustav Carus, TU Dresden,  
 Fetscherstraße 74, D-01307 Dresden, Germany*

<sup>§</sup>*debasish@physik.hu-berlin.de*

Received 15 July 2019

Revised 25 September 2019

Accepted 26 September 2019

Published 8 October 2019

We discuss the extraction of the ground state  $\langle K(\mathbf{p})|V_\mu(0)|B_s(\mathbf{0})\rangle$  matrix elements from Euclidean lattice correlation functions. The emphasis is on the elimination of excited state contributions. Two typical gauge-field ensembles with lattice spacings 0.075, 0.05 fm and pion masses 330, 270 MeV are used from the O(*a*)-improved CLS 2-flavor simulations and the final state momentum is  $|\mathbf{p}| = 0.5$  GeV. The b-quark is treated in HQET including the  $1/m_b$  corrections. Fits to two-point and three-point correlation functions and suitable ratios including summed ratios are used, yielding consistent results with precision of around 2% which is *not* limited by the  $1/m_b$  corrections but by the dominating static form factors. Excited state contributions are under reasonable control but are the bottleneck towards precision. We do not yet include a specific investigation of multi-hadron contaminations, a gap in the literature which ought to be filled soon.

**Keywords:** Lattice QCD; heavy quark effective theory; bottom mesons; spectroscopy.

PACS numbers: 12.38.Gc, 12.39.Hg, 14.40.Nd

### 1. Introduction

Decays of b-quarks in the form  $B \rightarrow \pi\ell\nu$  and  $B_s \rightarrow K\ell\nu$  are very relevant in constraining the Standard Model of particle physics and hence also in the search

<sup>§</sup>Corresponding author.

All figures and pictures by the author(s) in this article are distributed under the terms of the Creative Commons Attribution 4.0 (CC BY) License (<https://creativecommons.org/licenses/by/4.0/>).

for deviations from it. Lattice QCD is the method of choice for predicting the necessary form factors.<sup>1–9</sup> While computations are progressing significantly, there are still a few groups carrying them out. Crosschecks by independent (lattice) methods are largely absent. In Ref. 10 we have started an investigation which takes significantly different avenues in almost every choice that can be made. We use (improved) Wilson fermions, treat the b-quark in nonperturbative HQET and compute three-point functions for all time separations rather than fixing its total time-span (called  $\tau = t_K + t_{B_s}$  later) to one or few values. We opt for nonperturbative HQET because a complete and practical method is known to nonperturbatively renormalize the theory and match it to QCD.<sup>11–15</sup> When completed, this will provide an independent crosscheck.

Once these choices are made, one needs to carry out the basic steps as written in Ref. 10:

- (a) obtain the *ground state* matrix elements  $\langle K | V^\mu(0) | B_s \rangle$  that mediate the transition,
- (b) renormalize the currents (and thus matrix elements) and, if an effective theory is used, relate them to QCD (“matching”),
- (c) take the continuum limit of the matrix elements,
- (d) extrapolate to the quark masses realized in Nature,
- (e) map out the  $q^2$  dependence.

In the previous work we saw that the continuum limit is smooth in the static approximation, which is expected to yield the by far dominating part of the full result. In this paper we discuss the step (a) in detail, including the  $1/m$  terms. It is not at all an easy enterprise. As discussed at length below, the issue is that ground state matrix elements are obtained at large Euclidean time separations. At time separations of around  $\frac{1}{2}$  fm — a typical QCD length scale — the low-lying states start dominating the two- and three-point functions. Unfortunately the statistical quality of the Monte Carlo estimates is typically deteriorating fast once 2 fm in total (or 1 fm for K and  $B_s$  each) are reached. It is thus a matter of the details of methods and QCD dynamics whether there is a window to determine the desired form factors at the percent level.

In other words, control over the extraction of the bare form factors is crucial in a proper determination of the decay rates. In this publication we thus explain our methods and their limitations in detail.

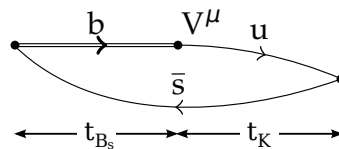


Fig. 1. The correlation function used to extract the  $B_s \rightarrow K$  decay matrix element.

## 2. Setup for the Nonperturbative Evaluation

### 2.1. Form factors from Euclidean correlation functions

To leading order in the weak interactions, the  $B_s \rightarrow K\ell\nu$  decay rate depends on two form factors,  $h_{\parallel}(E_K)$  and  $h_{\perp}(E_K)$ , which are related to the commonly used

$$f_+(q^2) = \frac{1}{\sqrt{2m_{B_s}}} [h_{\parallel} + (m_{B_s} - E_K)h_{\perp}], \quad (2.1)$$

$$f_0(q^2) = \frac{\sqrt{2m_{B_s}}}{m_{B_s}^2 - m_K^2} [(m_{B_s} - E_K)h_{\parallel} + (E_K^2 - m_K^2)h_{\perp}]. \quad (2.2)$$

The kinematics is explained in Ref. 10, and in the rest-frame of the  $B_s$ , we have

$$(2m_{B_s})^{-1/2} \langle K(\mathbf{p}_K) | \hat{V}_M^0(0) | B_s(\mathbf{0}) \rangle = h_{\parallel}(E_K), \quad (2.3)$$

$$(2m_{B_s})^{-1/2} \langle K(\mathbf{p}_K) | \hat{V}_M^k(0) | B_s(\mathbf{0}) \rangle = p_K^k h_{\perp}(E_K), \quad (2.4)$$

where  $(E_K, \mathbf{p}_K)$  is the four-momentum of the final state and  $|B_s(\mathbf{0})\rangle$  is the initial state with the  $B_s$  at rest,  $\mathbf{p}_{B_s} = \mathbf{0}$ . The above hadron-to-hadron matrix elements are in Minkowski space, as indicated by the subscript M. They can be obtained from the Euclidean three-point function (Fig. 1), where the spatial volume is a  $L^3$  torus,

$$\mathcal{C}_{\mu}^{B_s \rightarrow K}(t_K, t_{B_s}; \mathbf{p}_K) = \frac{a^9}{L^3} \sum_{\mathbf{x}_f, \mathbf{x}_v, \mathbf{x}_i} e^{-i\mathbf{p}_K(\mathbf{x}_f - \mathbf{x}_v)} \times \langle \bar{\mathcal{O}}_{us}(x_f) V_{\mu}(x_v) \mathcal{O}_{bs}(x_i) \rangle, \quad (2.5)$$

$$t_K = (x_f - x_v)_0 > 0, \quad t_{B_s} = (x_v - x_i)_0 > 0,$$

with suitable interpolating fields  $\mathcal{O}_{qq'}$ ,  $\bar{\mathcal{O}}_{qq'}$  (see (2.14), (2.16)) and

$$V_{\mu}(x) = \bar{\psi}_u(x) \gamma_{\mu} \psi_b(x). \quad (2.6)$$

Removing the overlaps  $\langle B_s(\mathbf{0}) | \mathcal{O}_{bs} | 0 \rangle$ ,  $\langle 0 | \bar{\mathcal{O}}_{us} | K(\mathbf{p}_K) \rangle$  (amputation), requires also the two-point functions

$$\mathcal{C}^K(t; \mathbf{p}_K) = \frac{a^6}{L^3} \sum_{\mathbf{x}_f, \mathbf{x}_i} e^{-i\mathbf{p}_K(\mathbf{x}_f - \mathbf{x}_i)} \langle \bar{\mathcal{O}}_{us}(x_f) \mathcal{O}_{us}(x_i) \rangle, \quad t = (x_f - x_i)_0, \quad (2.7)$$

$$\mathcal{C}^{B_s}(t) = \frac{a^6}{L^3} \sum_{\mathbf{x}_f, \mathbf{x}_i} \langle \bar{\mathcal{O}}_{bs}(x_f) \mathcal{O}_{bs}(x_i) \rangle, \quad t = (x_f - x_i)_0. \quad (2.8)$$

The explicit relation between the Euclidean correlation functions and the desired matrix elements is (see e.g. Ref. 16)

$$\begin{aligned} \langle K(\mathbf{p}) | \hat{V}_M^0(0) | B_s(\mathbf{0}) \rangle &= \lim_{t \rightarrow \infty} \frac{\mathcal{C}_0^{B_s \rightarrow K}(t, t; \mathbf{p})}{[\mathcal{C}^K(2t; \mathbf{p}) \mathcal{C}^{B_s}(2t)]^{1/2}}, \\ \langle K(\mathbf{p}) | \hat{V}_M^k(0) | B_s(\mathbf{0}) \rangle &= i \lim_{t \rightarrow \infty} \frac{\mathcal{C}_k^{B_s \rightarrow K}(t, t; \mathbf{p})}{[\mathcal{C}^K(2t; \mathbf{p}) \mathcal{C}^{B_s}(2t)]^{1/2}}. \end{aligned} \quad (2.9)$$

A few comments are in order

- We chose to define the correlation functions on a lattice. Replacing  $a^3 \sum_{\mathbf{x}} \rightarrow \int d^3\mathbf{x}$  yields the continuum expressions.
- The derivation of Eq. (2.9) (on a lattice) is based just on the existence of a transfer matrix with standard properties and on interpolating fields which are local in time; nonlocality (smearing) in space is allowed.
- Of course, on a finite lattice the limit  $t \rightarrow \infty$  has to be replaced by a proper procedure. Controlling it is the main topic of this paper. As a first, straightforward, procedure we note that it would be sufficient to have  $t$  (as well as the spatial box length  $L$ ) large compared to the typical QCD scales  $\frac{1}{2}$  fm,  $1/m_\pi$ .
- In order to obtain the physical form factors, the current  $V_\mu$  has to be properly renormalized and the continuum limit has to be taken.

So far we have spoken of standard relativistic (lattice) QCD. The nonperturbative HQET expressions including  $\mathcal{O}(1/m)$  terms are obtained by the following generic expansion of correlation functions<sup>11,17,18</sup>

$$\begin{aligned} \log(\mathcal{C}_\mu^{\text{B}_s \rightarrow \text{K}}) &= \log Z_{V_\mu}^{\text{HQET}} + \log([\mathcal{C}_\mu^{\text{B}_s \rightarrow \text{K}}]_{\text{stat}}) \\ &+ \sum_k \omega_k \rho_\mu^k + \delta_{\mathcal{O}} + \mathcal{O}\left(\frac{1}{m_b^2}\right), \end{aligned} \quad (2.10)$$

where  $[\mathcal{C}_\mu^{\text{B}_s \rightarrow \text{K}}]_{\text{stat}}$  is the correlation function in the static approximation and

$$\rho_\mu^k = \frac{[\mathcal{C}_\mu^{\text{B}_s \rightarrow \text{K}}]_k}{[\mathcal{C}_\mu^{\text{B}_s \rightarrow \text{K}}]_{\text{stat}}}. \quad (2.11)$$

The “stat” terms refer to lowest order (static) HQET, while  $k \in \{\text{kin, spin}\}$  are the terms due to the  $1/m$  corrections to the action, and the remaining terms in the sum over  $k$  are  $1/m$  corrections to the currents, see Subsec. 6.2. In  $\delta_{\mathcal{O}}$ , we collect all terms which arise from the renormalization and  $1/m$  corrections to the interpolating fields. Their form is irrelevant since they cancel in the ratios of interest in Eq. (2.9). Apart from  $\delta_{\mathcal{O}}$ , all divergences strictly cancel within Eq. (2.10) iff the parameters  $\log Z_{V_\mu}^{\text{HQET}}$  and  $\omega_k$  are determined nonperturbatively. The expansion of  $\log(\mathcal{C}_\mu^{\text{B}_s})$  reads the same apart from the current-terms which are absent. Finally, functions of correlation functions such as the ratio of Eq. (2.9) are trivially obtained from the expansion of the correlation functions.

For completeness we remind the reader of the rule of the game: the expansion is “derived” by considering the parameters  $\omega_k$  as  $\mathcal{O}(1/m)$  and expanding in  $1/m$ , despite the fact that individual pieces may be highly divergent. The logarithm is taken in Eq. (2.10) since then all expressions are automatically linear in the parameters and higher order terms in the parameters (and thus  $1/m$ ) are directly avoided. Those would otherwise have to be dropped to preserve renormalizability.

Table 1. Overview of the subset of  $N_f = 2$  CLS ensembles used. They have  $T = 2 \times L$ . Lattice spacings,  $a$ , are taken from Refs. 19 and 20. The pion mass is denoted by  $m_\pi$ , and we have  $m_\pi L \geq 4$  such that finite-volume effects are sufficiently (exponentially) suppressed at the level of accuracy we are aiming for.

Ensemble	$\beta$	$L/a$	$a$ [fm]	$m_\pi$ [MeV]	$m_\pi L$
A5	5.2	32	0.0749(8)	330	4.0
O7	5.5	64	0.0483(4)	270	4.2

## 2.2. Lattice setup

For the lattice computation of the correlation functions we use gauge field configurations generated by the Coordinated Lattice Simulations (CLS) effort<sup>19</sup> with two degenerate flavors of  $O(a)$ -improved Wilson fermions<sup>21</sup> and Wilson gauge action.<sup>22</sup> In the following, we present only results from the two representative ensembles, labeled A5 and O7, see Table 1. The lattice spacings were determined from the Kaon decay constant  $f_K$  at physical pion mass in Ref. 19 and updated in Ref. 23. They do not depend on the quark mass, i.e. we use a mass-independent renormalization scheme.

The natural choice for the mass of the valence up quark is equal to the mass of the degenerate up and down sea quarks,  $m_1$ , and thus fixed by the gauge field ensembles. For the mass of the spectator (valence) strange quark we are free, however, to choose any smooth function  $m_s(m_1)$  passing through the physical point. We follow strategy 1 of Ref. 19 and define  $m_s(m_1)$  by fixing the squared ratio between the Kaon mass and the Kaon decay constant to its physical value  $R_K^{\text{phys}} \equiv (m_K^{\text{phys}}/f_K^{\text{phys}})^2$  with  $m_K^{\text{phys}} = 494.2$  MeV and  $f_K^{\text{phys}} = 155$  MeV. We expect that this will lead to a flat extrapolation to the physical value of  $m_1$ , the “physical point.” At the simulated (unphysical) values of  $m_1$ , however, the resulting Kaon mass is  $m_K(m_1) = f_K(m_1) \cdot (R_K^{\text{phys}})^{1/2}$  and only approximately equal to its physical value (see e.g. Fig. 3).

In the static correlation functions on the right-hand side of Eq. (2.10), the heavy quark field  $\psi_b$  is replaced by a static quark  $\psi_h = \gamma_0 \psi_h$  with the HYP2 discretizations.<sup>24,25</sup> At NLO also the  $k \in \{\text{kin}, \text{spin}\}$  correlation functions are needed, see App. A.1. They are static correlation functions with an extra insertion of  $\mathcal{O}^{\text{kin}}$  or  $\mathcal{O}^{\text{spin}}$  from the HQET action.<sup>11</sup> Finally there are the extra  $1/m$  terms in the current  $V_\mu$ , see Subsec. 6.2 and App. A.1. We implement  $O(a)$  improvement perturbatively by writing the vector current as

$$V_0^{\text{stat}} = \bar{\psi}_u \gamma_0 \psi_h + ac_{V_0}(g_0) \bar{\psi}_1 \sum_l \overleftrightarrow{\nabla}_l^S \gamma_l \psi_h, \quad (2.12a)$$

$$V_k^{\text{stat}} = \bar{\psi}_u \gamma_k \psi_h - ac_{V_k}(g_0) \bar{\psi}_1 \sum_l \overleftrightarrow{\nabla}_l^S \gamma_l \gamma_k \psi_h \quad (2.12b)$$

with<sup>26</sup>

$$c_{V_k} = 0.0518g_0^2, \quad c_{V_0} = 0.0380g_0^2 \quad (2.13)$$

when we evaluate the static matrix elements, but we use  $c_{V_k} = c_{V_0} = 0$  when  $V_\mu^{\text{stat}}$  enters ratios at order  $1/m$  such as Eq. (2.11).

For the interpolating field of the  $B_s$  meson we use three different quark bilinears ( $r = 1, 2, 3$ )

$$\mathcal{O}_{bs,r}(x) \equiv \bar{\psi}_h(x) \gamma_5 \psi_{s,r}^{\text{sm}}(x), \quad (2.14a)$$

$$\bar{\mathcal{O}}_{bs,r}(x) \equiv \bar{\psi}_{s,r}^{\text{sm}}(x) \gamma_0 \gamma_5 \gamma_0 \psi_h(x). \quad (2.14b)$$

The smeared light quark fields are constructed by  $I_r^{\text{Bs}}$  iterations of Gaussian smearing<sup>27,28</sup>

$$\psi_{s,r}^{\text{sm}} = (1 + \kappa_G a^2 \Delta)^{I_r^{\text{Bs}}} \psi_s, \quad (2.15)$$

where the gauge links in the covariant Laplacian  $\Delta$  are defined by 3 iterations of (spatial) APE smearing<sup>29</sup> and  $\kappa_G = 0.1$ .

The interpolating field for the Kaon,

$$\mathcal{O}_{su}(x_f) \equiv \bar{\psi}_s^{\text{sm}}(x_f) \gamma_5 \psi_u^{\text{sm}}(x_f), \quad (2.16a)$$

$$\bar{\mathcal{O}}_{su}(x_f) \equiv \bar{\psi}_u^{\text{sm}}(x_f) \gamma_0 \gamma_5 \gamma_0 \psi_s^{\text{sm}}(x_f), \quad (2.16b)$$

is constructed from two smeared quark fields,  $\bar{\psi}_s^{\text{sm}}$  and  $\psi_u^{\text{sm}}$ , each with  $I^K$  iterations of Gaussian smearing.

Integrating out the Grassmann-valued quark fields yields the correlation functions in terms of traces of products of quark propagators and other factors such as smearing operators. In order to make full use of translation invariance, we evaluate the traces by a stochastic estimator which we represent by a single random  $U(1)$  vector with support on time slice  $(x_f)_0$ . All values of  $(x_f)_0$  are averaged over (“full time dilution”). More details can be found in App. A.

The momentum transfer  $q^2$  is computed from the continuum relation

$$q^2 = (m_{B_s}^{\text{phys}})^2 + (m_K^{\text{phys}})^2 - 2m_{B_s}^{\text{phys}} \sqrt{(m_K^{\text{phys}})^2 + \mathbf{p}_K^2} \quad (2.17)$$

with  $\mathbf{p}_K$  in the rest frame of the  $B_s$  and physical values of the masses (using  $m_{B_s}^{\text{phys}} = 5366.77$  MeV). In order to have  $q^2$  fixed to the same value on all ensembles, we need to adjust  $\mathbf{p}_K$  accordingly. This is achieved by using flavor-twisted boundary conditions for the (valence) s and b quarks, i.e. instead of periodic boundary conditions, we impose

$$\psi_s(x_0, \mathbf{x} + \hat{\mathbf{k}}L) = e^{i\theta_k^s} \psi_s(x_0, \mathbf{x}), \quad (2.18a)$$

$$\psi_b(x_0, \mathbf{x} + \hat{\mathbf{k}}L) = e^{i\theta_k^b} \psi_b(x_0, \mathbf{x}), \quad (2.18b)$$

for the unit vectors  $\hat{\mathbf{k}}$  in  $k$ -direction. In this way, the momentum of the  $K$  meson becomes

$$(p_K)_k = \frac{2\pi n_k - \theta_k^s}{L}, \quad (2.19)$$

while the  $B_s$  meson is kept at rest ( $\mathbf{p}_{B_s} \sim \theta^b - \theta^s = 0$ ) by choosing  $\theta^b = \theta^s$ . We fix the value of  $|\mathbf{p}_K| = 0.535$  GeV, and hence  $q^2 = 21.23$  GeV<sup>2</sup>.<sup>a</sup>

### 3. Static Correlation Functions

For  $t_K, t_{B_s} \gg a$ , where any possible violations of positivity are exponentially damped,<sup>b</sup> and assuming infinite  $T$  for the moment, we can decompose the Euclidean correlation functions as (in the following we keep the  $\mathbf{p}$  dependence of  $\mathcal{C}^K$  and  $\mathcal{C}^{B_s \rightarrow K}$  implicit):

$$\mathcal{C}^K(t_K) = \sum_{m=0}^{\infty} (\kappa^{(m)})^2 e^{-E_K^{(m)} t_K} \cong \sum_{m=0}^{N_K-1} (\kappa^{(m)})^2 e^{-E_K^{(m)} t_K}, \quad (3.1)$$

$$\mathcal{C}_{rs}^{B_s}(t_{B_s}) = \sum_{n=0}^{\infty} \beta_r^{(n)} \beta_s^{(n)} e^{-E_{B_s}^{(n)} t_{B_s}} \cong \sum_{n=0}^{N_{B_s}-1} \beta_r^{(n)} \beta_s^{(n)} e^{-E_{B_s}^{(n)} t_{B_s}}, \quad (3.2)$$

$$\begin{aligned} \mathcal{C}_{\mu,r}^{B_s \rightarrow K}(t_K, t_{B_s}) &= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \kappa^{(m)} \varphi_{\mu}^{(m,n)} \beta_r^{(n)} e^{-E_K^{(m)} t_K} e^{-E_{B_s}^{(n)} t_{B_s}} \\ &\cong \sum_{n=0}^{N_K-1} \sum_{m=0}^{N_{B_s}-1} \kappa^{(m)} \varphi_{\mu}^{(m,n)} \beta_r^{(n)} e^{-E_K^{(m)} t_K} e^{-E_{B_s}^{(n)} t_{B_s}}, \end{aligned} \quad (3.3)$$

where the indices  $r, s$  label the smearing levels used for the  $B_s$  meson, while the indices  $m, n$  label the Kaon and  $B_s$  meson energy levels respectively. A rather small number of states leads to a good precision if the time separations  $t_K, t_{B_s}$  are sufficiently large. In our fits we will use  $N_K = 1$  and  $N_{B_s} \leq 3$ .

By comparing the Euclidean representations of the correlation functions with Eq. (2.9) we see that the bare form factors of interest are given by  $\varphi_{\mu}^{(0,0)}$

<sup>a</sup>Note that small changes of the input parameters lead to

$$\delta q^2 \approx 10 \text{ GeV} \cdot \delta m_B + 5 \text{ GeV} \cdot \delta m_K + 2 \text{ GeV}^2 \cdot \delta \vartheta + 5 \text{ GeV}^2 \cdot \frac{\delta a(\beta)}{a(\beta)}.$$

The particular value for  $\mathbf{p}_K$  was chosen on the “N6” lattice which we do not discuss here, but which will be used in the physics analysis. On N6 we have  $L/a = 48$  and  $\beta = 5.5$  such that  $\mathbf{n} = (1, 0, 0)^T$  and  $\theta^s = \mathbf{0}$  lead to  $|\mathbf{p}_K| = 0.535$  GeV.

<sup>b</sup>In the pure Wilson regularization and with static quarks, positivity of the lattice transfer matrix is exact. With the  $O(a)$ -improvement term this may be different, but universality means that nonpositive contributions may appear only at distances close to the lattice spacing. For a concrete example of such contributions we refer to Ref. 30.

from Eq. (3.3), e.g.

$$\langle K(\mathbf{p}) | \hat{V}_M^0(0) | B_s(\mathbf{0}) \rangle = \varphi_0^{(0,0)}. \quad (3.4)$$

In the following we describe two methods to determine them:

- (1) By directly fitting<sup>c</sup> the correlation functions in a window of time extents where one can reliably limit the influence of excited states, statistical noise and other systematic effects (such as the finite- $T$  extent of the available periodic lattices). This will be the focus of Sec. 4.
- (2) By forming suitable ratios of correlation functions (like the one in Eq. (2.9)) where at large enough  $t$  the dependence on all or most of the additional parameters cancels out. This will be the topic of Sec. 5.

As we shall see, for both of these methods it is beneficial to first obtain a good determination of the two-point function parameters, as well as to determine regions in the three-point functions where one is safe from the finite- $T$  effects. These prerequisite steps are described in the following subsections. First we describe our methods to extract  $\kappa^{(0)}$  and  $E_K^{(0)}$  from Eq. (3.1).

### 3.1. Kaon correlation functions

While Eq. (3.1) is true in the limit of infinite  $T$ , in practice we have to take into account that  $T$  is finite. With our periodic boundary conditions in time, the ground-state contribution reads

$$\mathcal{C}^K(t) \cong (\kappa^{(0)})^2 \left( e^{-E_K^{(0)}t} + e^{-E_K^{(0)}(T-t)} \right). \quad (3.5)$$

Since the signal-to-noise problem in the Kaon sector is very mild, we average  $\mathcal{C}^K(t)$  and  $\mathcal{C}^K(T-t)$  and always fit until the middle of the lattice,  $t_{K2}^{\max} = T/2$ . For the start of the fit range, we follow Refs. 19 and 31 and use the (rather conservative) criterion that  $t_{\min}^{K2}$  is the first  $t$  for which the excited-state contribution, estimated by a two-state fit, is less than 25% of the statistical error of the effective mass  $E_{K,\text{eff}}(t) = -\frac{1}{a} \log[\mathcal{C}^K(t+a)/\mathcal{C}^K(t)]$ . In practice, we obtain  $t_{K2}^{\min} \approx 1.3\text{--}1.4$  fm. In this range, we fit the correlator to Eq. (3.5) to get the ground-state energy and amplitude. Note that the gap between the ground state and excited states in the Kaon sector is large, approximately  $1 \text{ GeV} \approx 5 \text{ fm}^{-1}$ .

To check the behavior of the signal-to-noise ratio and the dispersion relation, we calculated the Kaon two-point function at two values of momentum. The effective mass plots for both momenta are presented together with the corresponding fits in Fig. 2. At higher momentum the signal-to-noise problem becomes significantly more severe. The dispersion relation, also including a measurement for  $\mathbf{p} = 0$ , is presented in Fig. 3. The continuum dispersion relation describes the data very well even at the coarsest lattice spacing, showing that the cutoff effects in  $E_K/m_K$  are below our statistical precision.

<sup>c</sup>All the fits in this paper are uncorrelated, i.e. we use a diagonal covariance matrix.



Table 2. Details of the measurements. The hopping parameter of the valence strange quark,  $\kappa_s$ , has been determined as described in the text. The number of configurations on which we computed the correlation functions is denoted by  $N_{\text{cfg}}$ . The number of smearing iterations applied to the light quark(s) in the K and  $B_s$  meson is given by  $I^K$  and  $I_r^{B_s}$ , respectively. The integer vector  $\mathbf{n}$  and the scalar twist angle  $\vartheta$  determine the momentum of the K meson by setting  $\theta^s = \mathbf{n}\vartheta$  in (2.19), i.e.  $\mathbf{p}_K = (2\pi - \vartheta)\mathbf{n}/L$ .

Ensemble	$\kappa_s$	$N_{\text{cfg}}$	$I^K$	$(I_1^{B_s}, I_2^{B_s}, I_3^{B_s})$	$\mathbf{n}$	$\vartheta$
A5	0.13535(2)	1000	35	(30, 60, 155)	$(1, 0, 0)^T$	+0.2141
O7	0.13627(1)	490	75	(68, 135, 338)	$(1, 1, 0)^T$	-0.3593

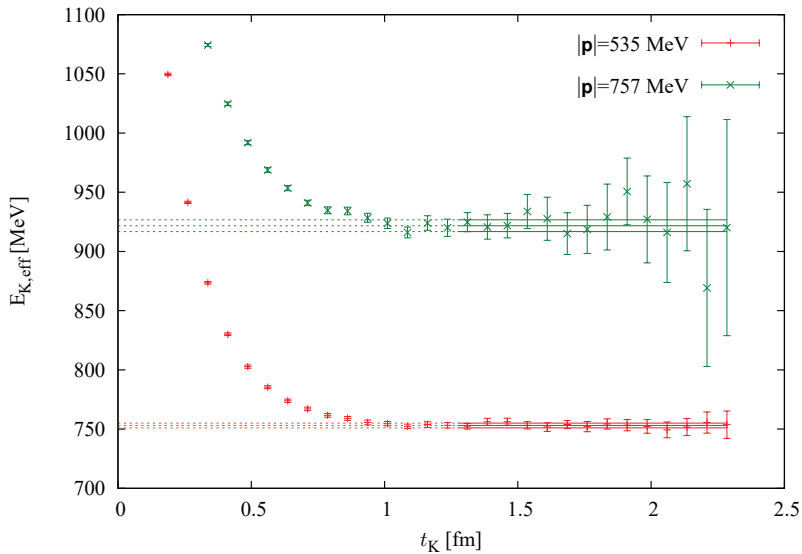


Fig. 2. Effective masses,  $E_{K,\text{eff}}(t_K) = -\partial_t \log \mathcal{C}^K(t_K)$  for ensemble A5 and two values of momentum, together with the fitted energies. In this graph as well as all the following ones we use MeV and fm units just for illustration. We thus convert from lattice units to physical ones *without* taking into account the error in the lattice spacings. The lattice parameters for the Kaon with momentum  $|\mathbf{p}| = 535$  MeV are given in Table 2. For the Kaon at  $|\mathbf{p}| = 757$  MeV, we have used  $\mathbf{n} = (1, 0, 0)$  and  $\vartheta = -0.9263$ .

### 3.2. $B_s$ correlation functions

For the two-point heavy-light ( $B_s$ ) correlator, we have three different smearings (in addition to the unsmeared correlator, which is not used due to its large contamination from highly excited states). With the diagonal and the off-diagonal terms, we have six independent correlators in a symmetric  $\mathcal{C}_{rs}^{B_s}$  matrix. The determination of the parameters of the  $B_s$  correlation functions is divided into several steps, which are described in the following subsections.

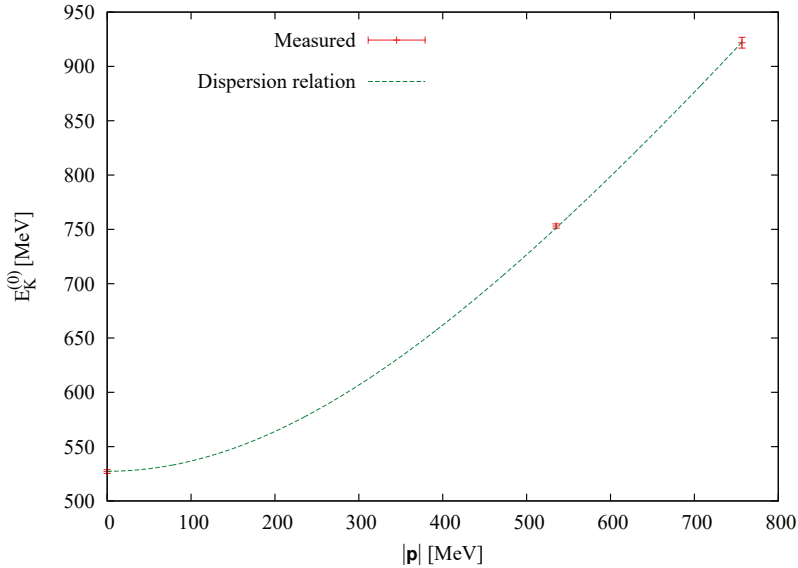


Fig. 3. Kaon energies at three values of momentum for ensemble A5. The lattice parameters are indicated in the caption of Fig. 2. The dashed curve shows the continuum dispersion relation evaluated with the rest mass from Ref. 19 (it is not a fit).

### 3.2.1. Energies from the GEVP

We first determine the energies by solving the generalized eigenvalue problem (GEVP) on  $\mathcal{C}_{rs}^{B_s}$ .<sup>32</sup> We use the full set of three interpolating fields. Thus the GEVP,

$$\mathcal{C}^{B_s}(t)v^{(n)}(t, t_0) = \lambda^{(n)}(t, t_0)\mathcal{C}^{B_s}(t_0)v^{(n)}(t, t_0), \quad n = 1, \dots, N_{\text{GEVP}}, \quad (3.6)$$

yields approximations to the lowest  $N_{\text{GEVP}} = 3$  energy levels via

$$E_{B_s, \text{eff}}^{(n)}(t) = -\tilde{\partial}_t \log(\lambda^{(n)}(t, t_0)) = \frac{1}{2a} \log\left(\frac{\lambda^{(n)}(t - a, t_0)}{\lambda^{(n)}(t + a, t_0)}\right). \quad (3.7)$$

We use  $t_0 = \lceil t/2 \rceil$  since this asymptotically accelerates the convergence (with  $t$ ) to the energies  $E_{B_s}^{(n)}$ .<sup>33</sup> For each  $n$  separately we then find the GEVP estimate of the energy  $E_{B_s}^{(n)}$  by a plateau fit (a weighted average) of  $E_{B_s, \text{eff}}^{(n)}(t)$ . The upper end,  $t_{\text{max}}$ , of the plateau is chosen with a noise criterion: only  $E_{B_s, \text{eff}}^{(n)}(t)$  with a relative error less than 20% (50% for the excited states) are included in the average. The start of the plateau,  $t_{\text{min}}$ , is chosen by requiring  $r(t_{\text{min}}) = \frac{|E_{\text{plat}}^{(n)}(t_{\text{min}}) - E_{\text{plat}}^{(n)}(t_{\text{min}} - \delta t)|}{\sqrt{\sigma^2(t_{\text{min}}) + \sigma^2(t_{\text{min}} - \delta t)}} \leq 3$ , where  $E_{\text{plat}}^{(n)}(t_{\text{min}})$  is the result of the plateau fit starting at  $t_{\text{min}}$  and  $\delta t \approx 2/(E_{B_s}^{(N)} - E_{B_s}^{(0)})$ , which we approximate by choosing  $\delta t = 0.3$  fm. For the ground state this criterion coincides with that of Ref. 34, ensuring that the statistical error dominates over the systematic one. Examples of  $E_{B_s, \text{eff}}^{(0)}(t)$  are presented in Fig. 4 (together with  $E_{K, \text{eff}}(t)$  for comparison). Note that a good estimate of the ground-state energy will be important in Sec. 5.

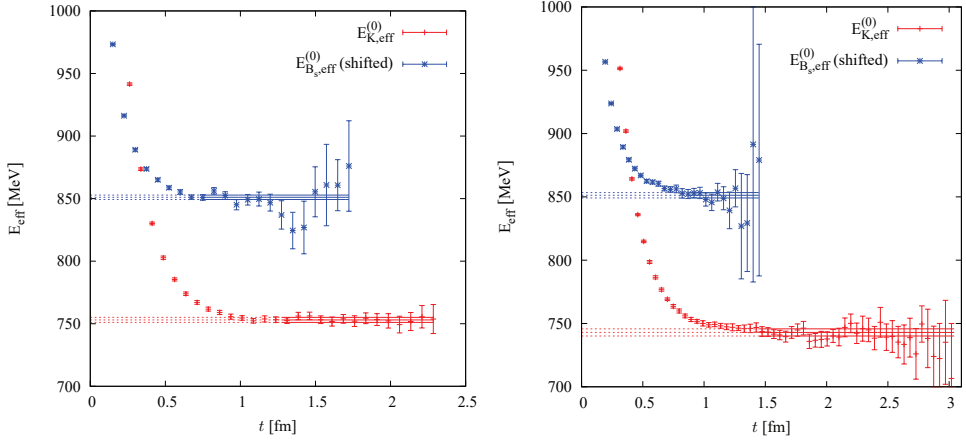


Fig. 4. Dependence of the K and  $B_s$  ground-state energies on  $t$  for ensemble A5 (left) and O7 (right).  $E_{B_s, \text{eff}}^{(0)}$  is shifted vertically for presentation. Dashed (and full) horizontal lines show the plateau fit values (and ranges).

For the excited states one should in principle use a higher value of  $\delta t \approx 2 / (E_{B_s}^{(N)} - E_{B_s}^{(n)})$  but this is not possible with the current precision of the data, as one ends up in regions dominated by noise. However, as opposed to the ground state, for the excited-state energies we only need reasonable rough estimates, as any residual systematic error will be eliminated in the subsequent fitting steps as described in Sec. 4.

### 3.2.2. Amplitudes

Having obtained the energies, we determine the amplitudes  $\beta_r^{(n)}$ . We start with the  $N_{B_s} = 3$  diagonal elements of  $\mathcal{C}^{B_s}$ ,

$$\mathcal{C}_{rr}^{B_s}(t) \cong \sum_{n=0}^{N_{B_s}-1} (\beta_r^{(n)})^2 e^{-E_{B_s}^{(n)} t} \quad (3.8)$$

and estimate the squared amplitudes through a linear fit. These values are used as initial guess for a nonlinear fit for the amplitudes  $\beta_r^{(n)}$  to all the correlators, including the off-diagonal ones. In this way also the relative signs of the amplitudes are obtained.

Then, we try the full fit for both amplitudes and energies whenever possible. In case the full fit becomes unstable with  $N_{B_s} = 3$  we fall back to the fit with fixed energies.

The upper end of the fit range,  $t_{B_2}^{\max}$ , is common for all the six correlators and is determined by the relative noise criterion, i.e. as the last point at which the relative error is smaller than 2.5% for all of the  $\mathcal{C}_{rs}^{B_s}$ . For our ensembles this gives 2.1–2.4 fm. Also the lower end of the fit range,  $t_{B_2}^{\min}$ , is the same for all  $\mathcal{C}_{rs}^{B_s}$ . It is a

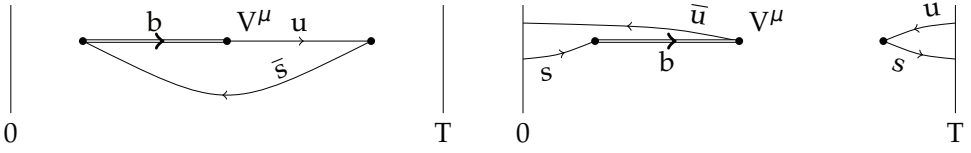


Fig. 5. Contributions to the three-point function corresponding to the  $B_s \rightarrow K$  decay (left), and a wrap-around Kaon (right) (figures taken from Ref. 31).

tunable parameter that is common for the determination of the amplitudes and for the combined fit (cf. Sec. 4).

### 3.3. Three-point static correlation functions

Due to the finite time extent of the lattice, the three-point function also receives contributions from particles propagating (“wrapping”) around the torus in time direction and is the sum of the contributions from the two diagrams shown in Fig. 5. At large enough time separations, the contributions of the corresponding lowest states can be written as

$$\begin{aligned} \mathcal{C}_{\mu,r}^{B_s \rightarrow K}(t_K, t_{B_s}) &\cong \kappa^{(0)} \varphi_{\mu}^{(0,0)} \beta_r^{(0)} e^{-E_{B_s}^{(0)} t_{B_s}} e^{-E_K^{(0)} t_K} \\ &+ \kappa^{(0)} \xi_{\mu,r} e^{-E_{B^*} t_{B_s}} e^{-E_K^{(0)} (T - t_{B_s} - t_K)}, \end{aligned} \quad (3.9)$$

where  $\xi_{\mu,r} = \langle 0 | V_{\mu} | B^* \rangle \langle B^* | P_{hl} | K \rangle$  contains unknown matrix elements of the state  $B^*$ , the lightest heavy–light state contributing to the “wrapper” diagram. At static order the energy  $E_{B^*}$  is equal to  $E_{B_s}^{(0)}$ .

Instead of introducing extra fit parameters  $\xi_{\mu,r}$  (and possibly others, if excited states need to be included), we restrict ourselves to time separations where the wrapper contributions are negligible.

For this purpose, at every fixed  $t_{B_s}$  we fit the three-point function to the form

$$\mathcal{C}_{\mu,r}^{B_s \rightarrow K}(t_K) \cong B_{\mu,r} e^{-Et_K} + C_{\mu,r} e^{+Et_K}, \quad (3.10)$$

with  $B_{\mu,r}$  and  $C_{\mu,r}$  being linear fit parameters (which one can express in terms of the amplitudes and matrix elements from Eq. (3.9)) and  $E$  being a nonlinear fit parameter (although one could in principle set it to  $E_K$  extracted from the two-point function). The fit is done in a region where ground-state dominance in the Kaon sector is expected (starting at approximately  $t_{K3}^{\min} = 0.8$  fm). Figure 6 shows an example of such a fit at fixed  $t_{B_s}$ .

For every  $t_{B_s}$ ,  $\mu$ , and  $r$  separately, we then find  $t_{K3}^{\max,wr}$  as the last  $t_K$  that fulfills the condition:

$$\frac{C_{\mu,r} e^{+Et_K}}{B_{\mu,r} e^{-Et_K} + C_{\mu,r} e^{+Et_K}} < c_{wr} \frac{\delta \mathcal{C}_{\mu,r}^{B_s \rightarrow K}(t_K, t_{B_s})}{\mathcal{C}_{\mu,r}^{B_s \rightarrow K}(t_K, t_{B_s})}, \quad (3.11)$$

with  $c_{wr} = 0.25$ .

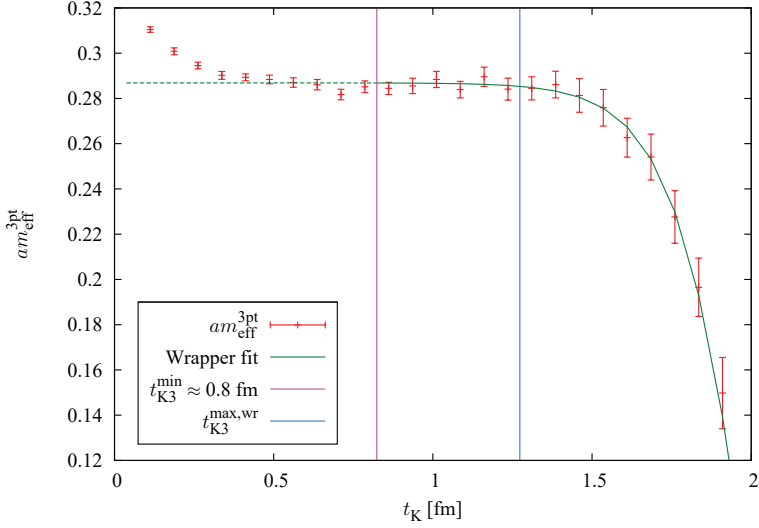


Fig. 6. The effective mass  $am_{\text{eff}}^{3\text{pt}}(t_K) = \log(C_{1,3}^{B_s \rightarrow K}(t_K, t_{B_s})/C_{1,3}^{B_s \rightarrow K}(t_K + a, t_{B_s}))$  with  $t_{B_s} = 8a \approx 0.6$  fm for ensemble A5, together with the fitted function. The start of the finite- $T$  fit and  $t_{K3}^{\text{max},\text{wr}}$  are also shown.

The times  $t_{K3,\mu}^{\text{max},\text{wr}}(t_{B_s})$  are shown by the red lines in Fig. 7 for the two representative ensembles A5 and O7 (together with other constraints for the fits as discussed in Sec. 4). We observe that the wrapper criterion yields considerably different  $t_{K3}^{\text{max},\text{wr}}$  for  $\mu = 0$  and  $\mu = 1$  (with lower  $t_{K3}^{\text{max},\text{wr}}$  for  $\mu = 1$ ). On the other hand, the different smearing levels give very similar limits, but are still all treated separately.

#### 4. Matrix Elements at Static Order from a Combined Fit

From the two-point function fits, we have obtained (estimates of) the energies and amplitudes,  $\kappa^{(m)}$ ,  $\beta_r^{(n)}$ . Taking them as input values, we can now fit Eq. (3.3) for each  $\mu = 0, 1$  with  $\varphi_\mu^{(m,n)}$ , i.e. the form factors, as  $N_K \times N_{B_s}$  linear parameters. These could in principle serve as the final result. However, for better stability, accuracy, and control of the systematic errors, we take all these results only as initial conditions for a combined simultaneous fit to the two-point and the three-point functions. In this way, we can use the information on the energies and amplitudes also from the three-point functions, which contain many data points. To estimate the errors we use<sup>35,36</sup> but with the derivatives with respect to the fit parameters calculated analytically.

The numerical results presented in this section are for the fit with  $N_K = 1$  and  $N_{B_s} = 3$ , corresponding to a total of 20 fit parameters.<sup>d</sup> We find that this

<sup>d</sup> $N_K + N_{B_s}$  energies,  $N_K^2 + N_{B_s}^2$  amplitudes, and  $2 \times N_K \times N_{B_s}$  form factors.

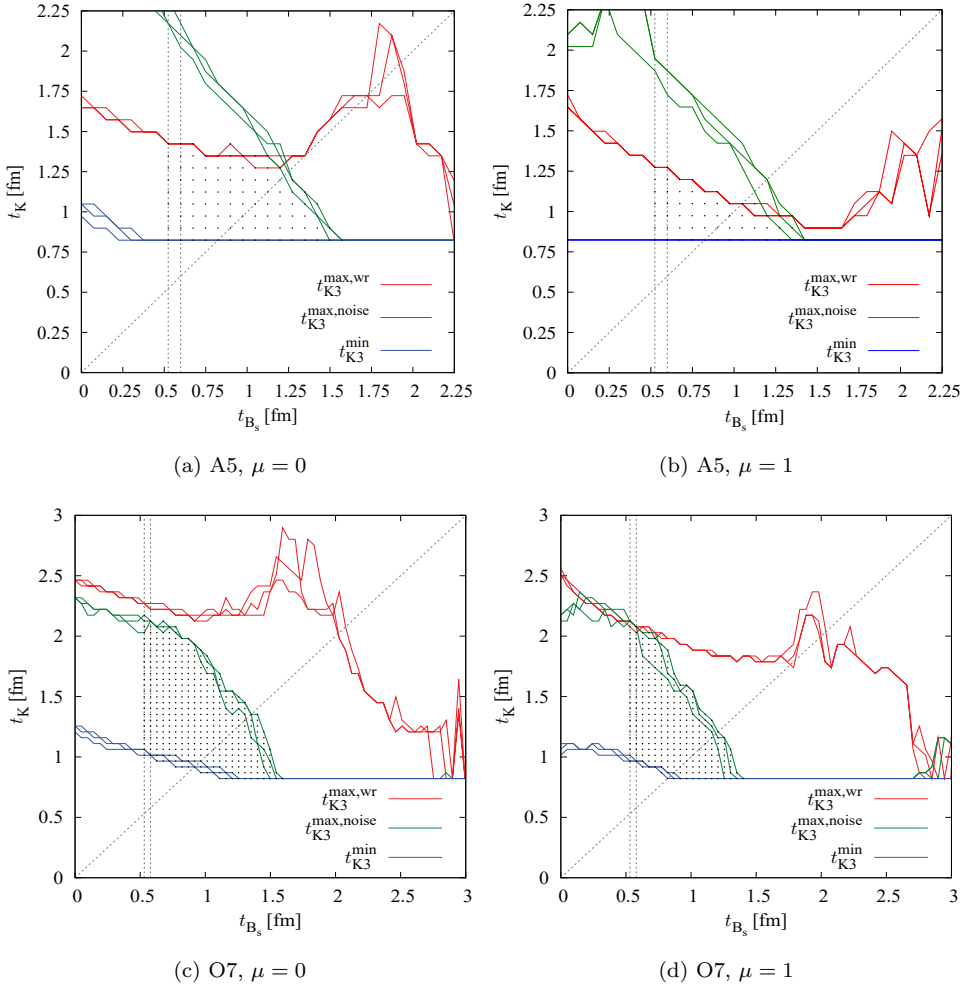


Fig. 7. (Color online) Time ranges  $t_{K3}^{\max,wr}$ ,  $t_{K3}^{\max,noise}$ , and  $t_{K3}^{\min}$  for lattices A5 and O7. The different lines with same color show the values for the three smearings. The vertical dashed lines are the two choices of  $t_{B3}^{\min}$  for the combined fit of Sec. 4 and the dots correspond to the data points included in the fit. To guide the eye, we also show the line  $t_K = t_{B_s}$  which is relevant for the ratios in Sec. 5.

choice gives the best stability with respect to the changes of the fit ranges, and therefore we use it to extract the values of  $\varphi_\mu^{(0)}$ . An alternative fit with  $N_K = 1$ ,  $N_{B_s} = 2$  is discussed in App. B. The fits with both  $N_K > 1$  and  $N_{B_s} > 1$  can easily become very unstable without some form of prior assumptions about the values of the excited-state form factors.

To keep statistical and systematic errors under control, only a small and carefully chosen subset of all the available time separations  $0 \leq t_K + t_{B_s} < T$  of the three-point function in the  $t_{B_s}$ - $t_K$ -plane can be used in the fit.

Moreover, it is important to have a criterion for good determination of the ground-state form factors by the combined fit. However, it is not easy to find a simple strict criterion, like the one from Subsec. 3.1 for  $E_K$ , and we monitor several criteria.

Details of both issues, the choice of the fit ranges and the quality of the fit, are discussed in the following.

#### 4.1. Fit ranges

For the three-point functions we select the data points to be included in the combined fit by the constraints  $t_{B_3}^{\min} \leq t_{B_s}$  and  $t_{K3}^{\min}(t_{B_s}) \leq t_K \leq t_{K3}^{\max}(t_{B_s})$ .

The function  $t_{K3}^{\max}(t_{B_s})$  is chosen, for each  $\mu$  and smearing level separately,<sup>e</sup> as

$$t_{K3}^{\max}(t_{B_s}) = \min\{t_{K3}^{\max,wr}, t_{K3}^{\max,noise}\}, \quad (4.1)$$

where  $t_{K3}^{\max,wr}(t_{B_s})$  is given by the wrapper criterion of (3.11) and  $t_{K3}^{\max,noise}(t_{B_s})$  is determined from a relative noise criterion, i.e. as the largest  $t_K$  which fulfills

$$\frac{\delta\mathcal{C}_{\mu,r}^{B_s \rightarrow K}(t_K, t_{B_s})}{\mathcal{C}_{\mu,r}^{B_s \rightarrow K}(t_K, t_{B_s})} < c_{\text{noise}}, \quad (4.2)$$

with  $c_{\text{noise}} = 2.5\%$ . The corresponding curves are shown in Fig. 7 and illustrate that A5 is an example of a wrapper-limited lattice (due to a small physical time extent). On the other hand, O7 is noise limited and by reducing statistical errors one may hope to improve the determination of the form factors.

Also the function  $t_{K3}^{\min}(t_{B_s})$  is selected by an automatic criterion<sup>f</sup> analogous to the one described in Subsec. 3.1: for each  $t_{B_s}$ ,  $\mu$ , and smearing, we do a two-exponential fit to  $t_K \in [0.4 \text{ fm}, t_{K3}^{\max}(t_{B_s})]$  (with the energies fixed to the ones extracted from the two-point function). We then find the minimum time  $t_{K3*}^{\min}(t_{B_s})$  at which the excited-state contribution is smaller than 25% of the statistical error. The final value is set to

$$t_{K3}^{\min}(t_{B_s}) = \max\{0.8 \text{ fm}, t_{K3*}^{\min}(t_{B_s})\} \quad (4.3)$$

to avoid values getting too small in the regions with large noise, where the two-state fit does not work well. The resulting  $t_{K3}^{\min}$  is shown by the blue curves in Fig. 7.

The remaining tunable parameters of the fit are  $t_{B_2}^{\min}$  and  $t_{B_3}^{\min}$ . In Ref. 10 we used  $t_{B_2}^{\min} \approx 0.45 \text{ fm}$ , which seems sufficient when  $N_{B_s} = 3$ . With better smearings and improved analysis methods, we can use slightly lower values, with two choices:  $t_{B_2}^{\min} \approx 0.32 \text{ fm}$  (“aggressive”) and  $0.38 \text{ fm}$  (“conservative”).

Concerning  $t_{B_3}^{\min}$  it was observed in Ref. 10 that the fit performs better if one takes a positive  $t_{B\Delta}^{\min} \equiv t_{B_3}^{\min} - t_{B_2}^{\min}$ . This can be understood by noting that the

<sup>e</sup>In Refs. 10 and 31 we used one and the same value of  $t_{K3,\mu}^{\max}$  for all  $t_{B_s}$ , corresponding roughly to the one that produced a rectangle of largest area in the plane of the allowed points (for both values of  $\mu$  separately).

<sup>f</sup>In Ref. 10  $t_{K3}^{\min}$  was a manually tunable parameter of the fit procedure, chosen in common for all smearings and  $\mu$ .

amplitudes in the two-point function are proportional to  $|\beta_i^{(n)}|^2$ , while in the three-point functions they are only proportional to  $\beta_i^{(n)}$ . Thus, the suppression of excited states is expected to be stronger in the two-point functions if  $\beta_i^{(0)} > \beta_i^{(1)}$ . We find that this is in fact the case, except for the lowest smearing.<sup>§</sup>

With increasing time separation  $t_{B_s}$  the suppression of the first excited state is proportional to  $e^{-(E_{B_s}^{(1)} - E_{B_s}^{(0)})t}$ , which approximately amounts to 0.67 at  $t = 0.2$  fm, 0.4 at  $t = 0.4$  fm, and 0.25 at  $t = 0.5$  fm. Although it is not immediately obvious which value one should take, this justifies a nonzero value of  $t_{B\Delta}^{\min}$ . We keep  $t_{B\Delta}^{\min} \approx 0.2$  fm, which gives a good compromise in suppressing the higher excited-state contributions and keeping the number of data points large enough for good fit stability.

#### 4.2. Fit quality and results

The combined fit is “uncorrelated,” i.e. we use as weights which define  $\chi^2$ , the inverse diagonal elements of the covariance matrix,  $C_{\text{dat}}$ . For such a big matrix the full  $C_{\text{dat}}$  is very badly determined and conditioned (in fact, if the number of data points is bigger than the number of measurements, it is not possible to invert<sup>37,38</sup>).

As a way to monitor the quality of the uncorrelated fit, we use  $\chi_{\text{exp}}^2$ , the value expected when the fit describes the data, given normally distributed data with the measured  $C_{\text{dat}}$ .<sup>39</sup> It can be expressed as  $\chi_{\text{exp}}^2 = \text{tr } C_{\text{dat}} K_{\text{fit}}$ , where  $K_{\text{fit}}$  is a matrix given in terms of the fit functions and the weights defining  $\chi^2$ . Also including autocorrelations  $\chi_{\text{exp}}^2$  can be estimated efficiently,<sup>40</sup> but its uncertainty can be large as is common for treating autocorrelations. Thus it is not easy to exactly quantify what is a good fit. As a rule of thumb, we consider fits with  $\chi^2 \lesssim 2\chi_{\text{exp}}^2$  as acceptable. We often find values of  $\chi_{\text{exp}}^2$  which are a factor of two to five smaller than the number of degrees of freedom.

Another way of monitoring the fit quality is to look at the contributions to the  $\chi^2$  coming from different correlation functions. In particular, an unusually large contribution from  $C^{B_s}$  should be monitored as it can indicate that  $t_{B2}^{\min}$  was chosen too small (even though the overall  $\chi^2$  looks acceptable because of the many points in the three-point functions).

Apart from monitoring the  $\chi^2$ , we analyze the stability of the extracted fit values and errors with respect to changes of  $t_{B2}^{\min}$  and  $t_{B3}^{\min}$ .

The stability plots are always organized in the following way: the data is divided in groups with different  $t_{B2}^{\min}$ , which is also plotted on the  $x$ -axis in the middle of each group. Inside every group  $t_{B\Delta}^{\min}$  varies from 0 to approximately 0.5 fm. Additionally, we highlight in different colors the values for selected  $t_{B2/3}^{\min}$ , which are described in the following.

<sup>§</sup>For the ensembles we analyzed, typical values are:  $\beta_3^{(1)}/\beta_3^{(0)} \approx 0.25 - 0.4$ ,  $\beta_2^{(1)}/\beta_2^{(0)} \approx 0.7$ ,  $\beta_1^{(1)}/\beta_1^{(0)} \approx 1$ .



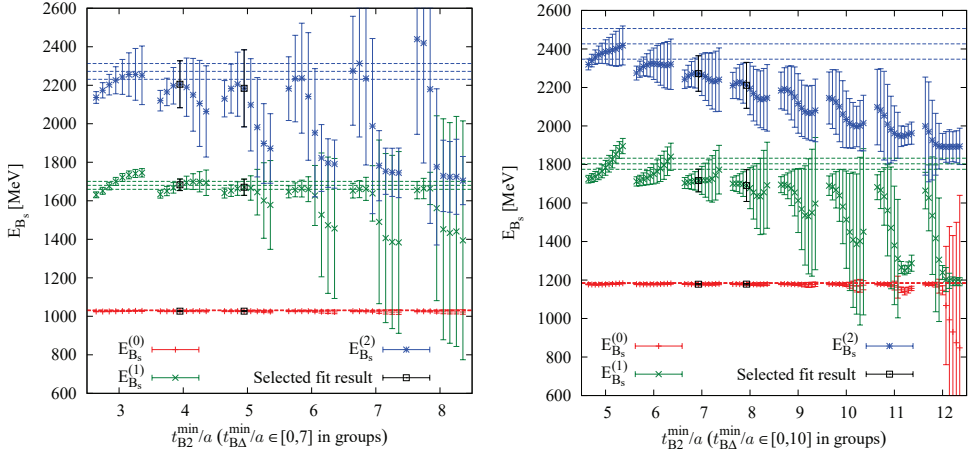


Fig. 8. (Color online) Stability plots for all three  $B_s$  energies on ensemble A5 (left) and O7 (right). The selected  $t_{B2/3}^{\min}$  values are highlighted with a black color, while the dashed lines are values from the GEVP, cf. Subsec. 3.2.1. The behavior at large  $t_{B2}^{\min}$  and  $t_{B3}^{\min}$  can be traced to the fact that the fit can no longer resolve three separate states. The selected fit values are discussed in Subsec. 4.1. A zoom into the behavior of  $E_{B_s}^{(0)}$  is shown in Fig. 9.

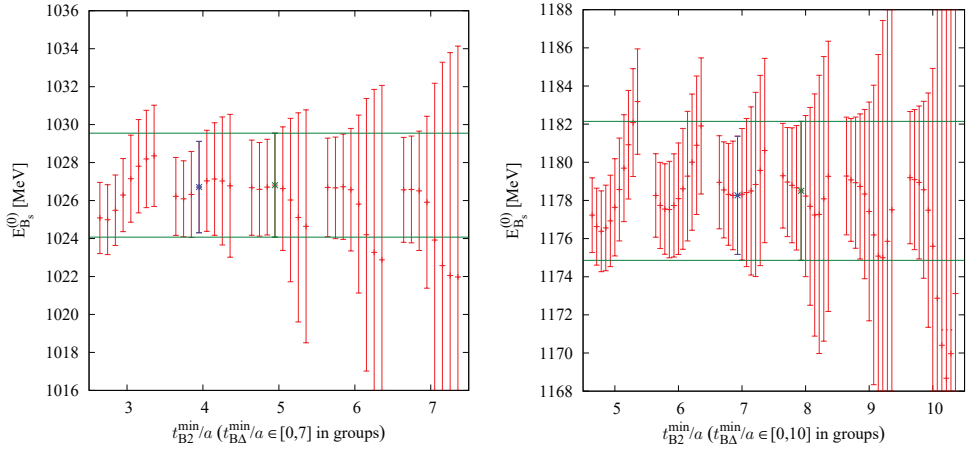


Fig. 9. Stability plots for  $E_{B_s}^{(0)}$  for ensemble A5 (left) and O7 (right).

In choosing  $t_{B2}^{\min}$  and  $t_{B3}^{\min}$  one must find a good window between too small values, where the fit is plagued by contamination from the higher excited states (this can be seen by high  $\chi^2$  as well as by lack of stability of the extracted fit parameters with respect to small changes in  $t_{B2/3}^{\min}$ ) and too high values, where it can no longer resolve the three  $B_s$  states. The latter is demonstrated in Fig. 8. At large  $t_{B2/3}^{\min}$  we see that for both ensembles at least one of the energies is no longer resolved. Therefore, we refrain from showing the largest values of  $t_{B2}^{\min}$  in Figs. 9–11.

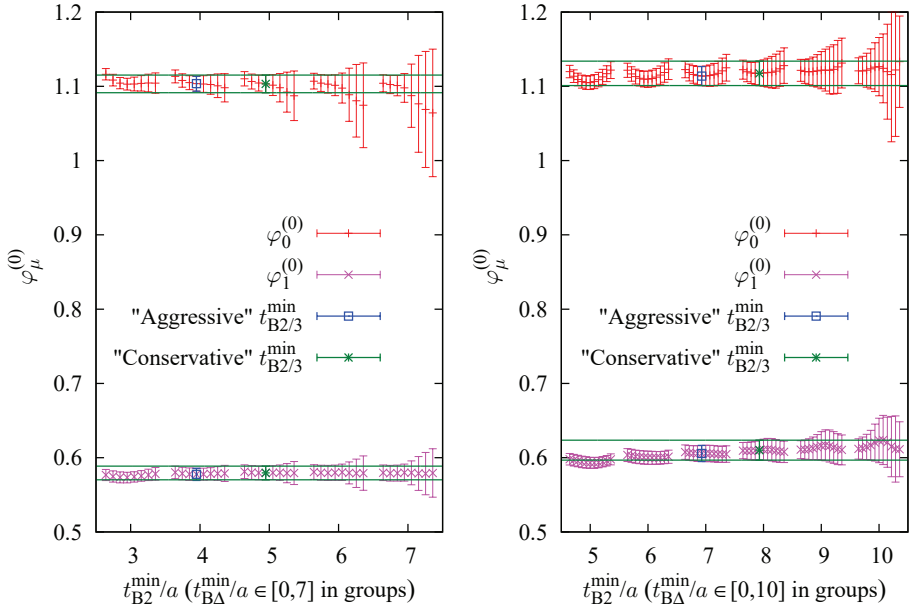


Fig. 10. Stability plots of the ground-state form factors for ensemble A5 (left) and O7 (right). The results are rather stable with respect to changes of  $t_{B2/3}^{\min}$ , although errors grow quickly.

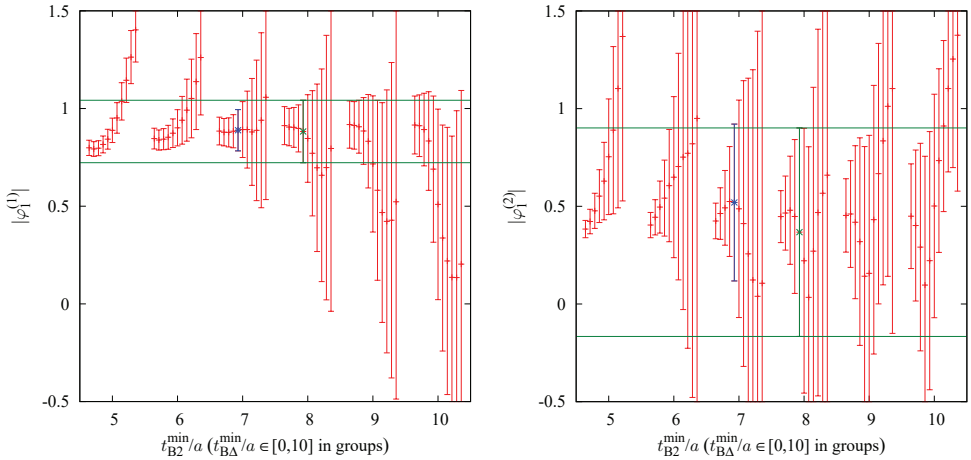


Fig. 11. Excited-state form factors at  $\mu = 1$  for ensemble O7.

In general, the ground-state energies (cf. Fig. 9) and the ground-state form factors (cf. Fig. 10) are reasonably stable with respect to changes of  $t_{B2/3}^{\min}$  (within quickly growing errors). We also show representative examples for the excited-state form factors in Fig. 11. Their precision, especially of  $\varphi_\mu^{(2)}$ , is much worse than of their ground-state counterparts.

Table 3. Results for the ground-state form factors  $\varphi_\mu^{(0,0)}$ . The second error in the result for  $\mathcal{R}_{\mu=1}^{\text{II}}$  for ensemble A5 is the systematic error due to the wrapper effect. The last column shows the results of the linear fit to  $\mathcal{S}_{\mu,3}^{\text{I}}(\tau)$ , see the shaded band in Fig. 14.

id	$\mu$	Fit		Ratio $\mathcal{R}^{\text{II}}$	Summed ratio $\mathcal{M}^{\text{I}}$
		$t_{\text{B}2}^{\text{min}} \approx 0.32 \text{ fm}$	$t_{\text{B}2}^{\text{min}} \approx 0.38 \text{ fm}$		
A5	0	1.103(12)	1.104(10)	1.129(18)	1.097(17)
A5	1	0.579(9)	0.578(8)	0.579(13)(05)	0.587(11)
O7	0	1.118(16)	1.114(13)	1.113(19)	1.109(13)
O7	1	0.610(13)	0.606(11)	0.602(15)	0.601(10)

The values for the ground-state form factors extracted with both choices of  $t_{\text{B}2}^{\text{min}}$  are collected in Table 3, together with the results from ratio methods described in the next section. Of course, the errors of all fit parameters are computed including correlations and autocorrelations encoded in  $C_{\text{dat}}$ .

## 5. Matrix Elements at Static Order from Ratios

### 5.1. Ordinary ratios

An alternative to extracting the ground-state matrix elements from fitting the correlation functions directly is to construct an appropriate ratio of the correlation functions such that the dependence on all or most of the other parameters cancels in the limit of large time.

One example of such a ratio was actually used to *define* the form factor, cf. Eq. (2.9). In the following it will be beneficial to generalize this definition to the case where  $t_K \neq t_{\text{B}_s}$  in the three-point function ( $\tau \equiv t_K + t_{\text{B}_s}$ ):

$$\mathcal{R}_{\mu,r}^{\text{I}}(t_K, t_{\text{B}_s}) = \frac{\mathcal{C}_{\mu,r}^{\text{B}_s \rightarrow \text{K}}(t_K, t_{\text{B}_s})}{[\mathcal{C}^{\text{K}}(\tau)\mathcal{C}_{rr}^{\text{B}_s}(\tau)]^{1/2}} \exp\left\{\left(\tilde{E}_{\text{B}_s}^{(0)}(\tau) - \tilde{E}_{\text{K}}^{(0)}(\tau)\right) \frac{t_{\text{B}_s} - t_K}{2}\right\}. \quad (5.1)$$

In the limit  $t_K, t_{\text{B}_s} \rightarrow \infty$ , this ratio converges to the desired bare form factor  $\varphi_\mu^{(0,0)}$  if additional parameters  $\tilde{E}_{\text{K}}(\tau), \tilde{E}_{\text{B}_s}^{(0)}(\tau)$  satisfy  $\tilde{E}(\tau) = E + \mathcal{O}(\exp(-\Delta E\tau))$ . Possible choices will be discussed below.

In Subsec. 5.2 we will see the advantage of this definition, while its obvious disadvantage is that in the denominator one needs the two-point correlation functions at time  $\tau = t_K + t_{\text{B}_s}$ . This means that, e.g. for a simple estimate of a plateau in the three-point function at  $t_K \approx t_{\text{B}_s} \approx 1 \text{ fm}$  one needs the two-point correlation functions at a large time separation  $\tau \approx 2 \text{ fm}$ , which is particularly problematic in the case of  $\mathcal{C}^{\text{B}_s}$  where the signal-to-noise problem is much more severe than for  $\mathcal{C}^{\text{K}}$ .

For lattices with short time extent, like A5,  $\tau$  may come close to the middle of the lattice and we need to take into account the influence of the second term in Eq. (3.5). We do that by multiplying the ratio by an additional factor  $(1 + e^{E_{\text{K}}(T-4t)})^{1/2}$  which

cancels the unwanted contribution up to excited states which are negligible in this region.

In addition, one can consider another ratio

$$\mathcal{R}_{\mu,r}^{\text{II}}(t_K, t_{B_s}) = \frac{\mathcal{C}_{\mu,r}^{B_s \rightarrow K}(t_K, t_{B_s})}{[\mathcal{C}^K(t_K) \mathcal{C}_{rr}^{B_s}(t_{B_s})]^{1/2}} \exp \left\{ \tilde{E}_{B_s}^{(0)}(\tau) \frac{t_{B_s}}{2} + \tilde{E}_K^{(0)}(\tau) \frac{t_K}{2} \right\} \quad (5.2)$$

which has a more favorable signal-to-noise behavior of the denominator but requires energy estimates even for the choice  $t_K = t_{B_s}$ .

A similar ratio is

$$\mathcal{R}_{\mu,i}^{\text{III}}(t_K, t_{B_s}) = \frac{\mathcal{C}_{\mu,r}^{B_s \rightarrow K}(t_K, t_{B_s})}{\mathcal{N}^K \mathcal{C}^K(t_K) \mathcal{N}_r^{B_s} \mathcal{C}_{rr}^{B_s}(t_{B_s})}, \quad (5.3)$$

where the additional normalization factors  $\mathcal{N}^K$ ,  $\mathcal{N}_r^{B_s}$  can be expressed in terms of the correlation function parameters  $\kappa^{(0)}$  and  $\beta_r^{(0)}$  respectively. In practice, our data for  $\mathcal{R}^{\text{III}}$  shows very similar or slightly inferior behavior compared to  $\mathcal{R}^{\text{II}}$ , therefore we do not present the numerical results for this ratio.

Let us now discuss the choice of the ground-state energy estimates. One can set them to the effective masses of the corresponding two-point functions, avoiding any fit procedure. This, however, results in large statistical fluctuations in the large time region. For our data, it is beneficial to instead use the time-independent estimates  $E_K^{(0)}$ ,  $E_{B_s}^{(0)}$  extracted in Subsecs. 3.1 and 3.2 respectively. The plateaux in both cases were chosen in a conservative manner, and as a crosscheck we calculated the ratios for different choices of the plateau ranges of both energies. From that exercise we find that the systematic error associated with using the fitted ground-state energies is negligible with respect to our statistical uncertainties.

The resulting ratios  $\mathcal{R}^{\text{I}}$  and  $\mathcal{R}^{\text{II}}$  are presented in Figs. 12 and 13, respectively. In addition to the data we also plot the fit results from Sec. 4, including both the fitted values of  $\varphi_\mu^{(0,0)}$  and the respective ratios of correlation functions reconstructed by inserting the fit parameters ( $\varphi_\mu^{(0,n)}$ ,  $\kappa^{(0)}$ ,  $\beta_r^{(n)}$ ,  $E_K^{(0)}$  and  $E_{B_s}^{(n)}$ ) into Eqs. (3.1)–(3.3). There is a good agreement of the actual ratios with their fitted counterparts at large values of  $t$ , where excited states neglected in the fits are irrelevant.

On ensemble A5, due to the short time extent of the lattice, the wrapper criterion Eq. (3.11) becomes relevant. For  $\mu = 1$  it restricts the available times to approximately 1.1 fm, cf. Fig. 7. At larger times we observe that the ratio  $\mathcal{R}_1^{\text{II}}$  starts to grow rapidly, while for  $\mathcal{R}_1^{\text{I}}$  this is most likely masked by the large uncertainties. Furthermore, for  $\mu = 0$  the data is close to violating the wrapper criterion (see Fig. 7) at the maximal time separations used.

On both lattices, A5 and O7, we see that at large times,  $t \gtrsim 1$  fm,  $\mathcal{R}^{\text{II}}$  is clearly superior to  $\mathcal{R}^{\text{I}}$  in terms of the signal-to-noise and has a comparable precision to the results of the combined fit. To use  $\mathcal{R}^{\text{II}}$  for the extraction of  $\varphi_\mu^{(0,0)}$ , we need to select a suitable plateau: looking at the bands from the fit which give us an estimate of the excited-state contamination from the  $B_s$  sector, we start the plateau at  $t^{\text{min}} = 1.1$  fm

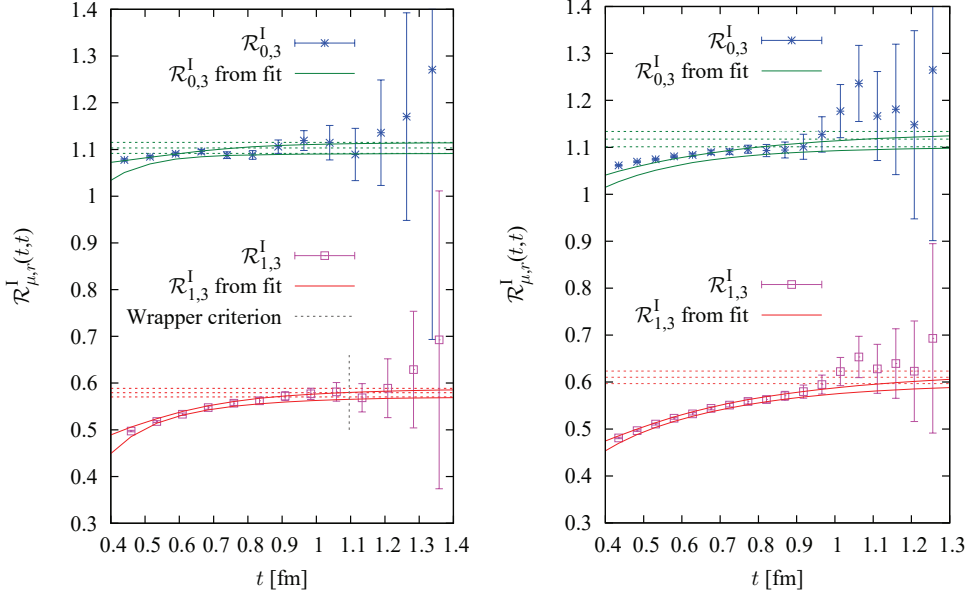


Fig. 12. Overview of the ratios  $\mathcal{R}^I$  with  $t_K = t_{B_s}$  for ensembles A5 (left) and O7 (right). Only the highest smearing,  $r = 3$ , is presented. For comparison we also show the results from the combined fit in Subsec. 4: the dashed horizontal lines are the fitted values of  $\varphi_{\mu}^{(0,0)}$  and the curves ( $1\sigma$  bands) are the respective ratios of correlation functions reconstructed from the fitted parameters.

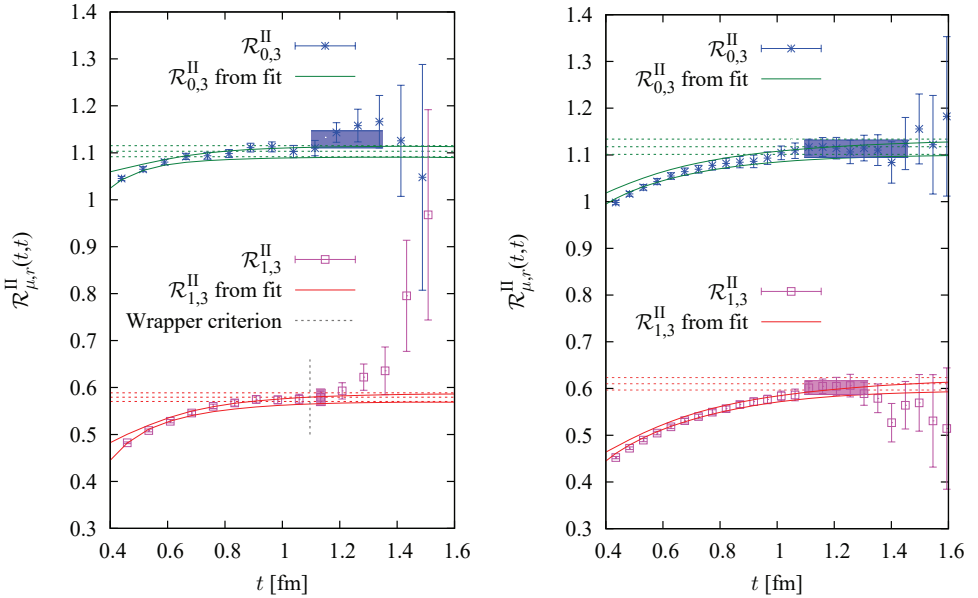


Fig. 13. Overview of the ratios  $\mathcal{R}^{II}$  with  $t_K = t_{B_s}$  for ensemble A5 (left) and O7 (right), with the fit curves analogous as in Fig. 12. Selected plateaux are also shown. For  $\mu = 1$  on ensemble A5 the wrapper contribution limits the available time range to  $t \leq t_{K3}^{\max,wr}$ , cf. Fig. 7.

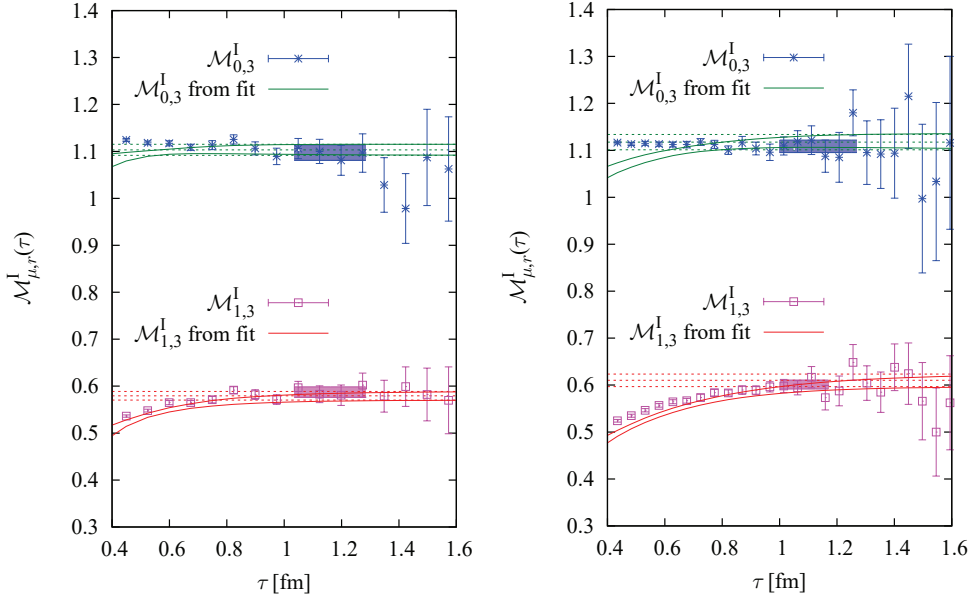


Fig. 14. Overview of the  $\mathcal{M}^I$  for ensemble A5 (left) and O7 (right), together with the fit curves analogous as in Figs. 12 and 13. The shaded bands represent  $c_2$  of fits  $\mathcal{S}_{\mu,r}^I(\tau) = c_1 + c_2\tau$  to the sum in the indicated range of  $\tau$ , chosen such that the wrapper effects are negligible.

and fit until the loss of precision of the signal below 5% or until we hit the wrapper criterion. On A5 for  $\mu = 1$  there is no valid plateau, because the wrapper is hit before  $t^{\min}$ . In this case, we quote the first data point above 1.1 fm as the final result and add to it a systematic error associated with the wrapper contribution estimated by the fit.

## 5.2. Summed ratios

A way to get improved convergence to the ground state is to sum<sup>h</sup> the ratio<sup>41–43</sup>  $\mathcal{R}^I$  and determine the matrix element from

$$\mathcal{M}_{\mu,r}^I(\tau) = \partial_\tau \mathcal{S}_{\mu,r}^I(\tau) = \partial_\tau a \sum_{t_{B_s}=0}^{\tau} \mathcal{R}_{\mu,r}^I(\tau - t_{B_s}, t_{B_s}). \quad (5.4)$$

The asymptotic excited-state contaminations are then  $\mathcal{O}(\tau \Delta e^{-\tau \Delta})$ , where  $\Delta = \min(E_K^{(1)} - E_K^{(0)}, E_{B_s}^{(1)} - E_{B_s}^{(0)})$ , as opposed to  $\mathcal{O}(e^{-\tau \Delta/2})$  in ordinary ratios.<sup>42,43,i</sup> The accelerated convergence of the summed ratios can be a decisive advantage in

<sup>h</sup>We write this in the form used in our analysis, but note that often it is advantageous and natural to sum the position of the operator that causes the transition over all space-time instead.<sup>41,42</sup>

<sup>i</sup>One could in principle attempt to sum the other ratios as well. However, the improvement in convergence is only proven to work (at least asymptotically) for  $\mathcal{R}^I$ . Empirically we see that it does not work for  $\mathcal{R}^{II}$  and  $\mathcal{R}^{III}$ .

the case of ensembles with a limited extent in time such as A5: contaminations by wrappers are less relevant. Example results are shown in Fig. 14. We observe that the convergence is indeed improved.

In fact, especially for  $\mu = 0$  the plateaux seem to start *very* early. This early onset of the plateaux is discussed in App. C where we show that it is caused by an accidental cancellation of the excited-state contributions in the  $B_s$  and  $K$  sector and therefore should be treated with caution.

Note, however, that the ratios reconstructed from the fits of Sec. 4 only include the influence of the excited states in the  $B_s$  sector. As in the previous subsection, these fit bands can be used to estimate the start of the plateaux. We choose  $\tau^{\min} = 1$  fm and  $\tau^{\max}$  by the 5% relative noise criterion.<sup>j</sup> In this range we extract the form factor from the slope of the linear fit directly to  $\mathcal{S}^I(\tau)$ . Note that on both lattices we are not affected by the wrapper criterion as opposed to ordinary ratios.

The results from the different methods are gathered in Table 3 and are in good agreement. One slight difference is seen in  $\mathcal{R}_0^{\text{II}}$  on ensemble A5. However, in this case if we start the plateau one data point earlier (which, judging from Fig. 13, seems legitimate) the result goes down to 1.117(15) which removes most of the difference. In general the results from the fits and the summed ratio have similar precision, in which case we prefer to quote the result from the summed ratio as the final results, as the latter method is simpler.

### 5.3. GEVP method

We have also implemented the GEVP method of Ref. 42, both the ratio one, Eq. (2.16) of Ref. 42, and the summed one, Eq. (3.7) of Ref. 42. In both cases we specialized to using the GEVP only in the  $B_s$  channel. The results are very similar to the ones of the previous two sections (with the largest wave function), but with errors which are a little bit larger. The fact that we do not see a significant improvement might be due to our interpolating fields, which likely do not distinguish between single hadron and (excited) multi-hadron states. The GEVP is then not able to significantly reduce the multi-hadron contributions. We comment further on this issue in the conclusions.

## 6. Matrix Elements at $1/m$ Order

The HQET expansion of B-meson observables becomes a precision tool only when  $1/m$  terms are included. We now discuss the determination of these crucial terms for the matrix elements. As in the static approximation we have the option to perform fits to the (two-point and three-point) correlation functions, or to consider ratios or summed ratios. Fits have been discussed and applied in Ref. 18 for the somewhat simpler case of the B-to-vacuum matrix element  $f_{B_s}$ . As there, in our present case,

<sup>j</sup>One could also choose 10%, the plateaux would be a bit longer but that does not help in reducing the statistical error.

all parameters  $\beta_s^{(n)}$ ,  $E^{(n)}$  get triplicated with static and  $1/m$  pieces corresponding to kin and spin insertions. Furthermore pure exponentials turn into exponentials plus terms of the form  $E_n^{\text{kin}} t \exp(-E_n^{\text{stat}} t)$ . This proliferation of terms in addition to the  $1/m$  corrections to the matrix elements themselves makes an analysis in terms of fits cumbersome and difficult. We therefore concentrate on the analysis of ratios, which we already have seen to be just as good as fits in the static approximation. We will also see that there are some simplifications in the  $1/m$  expansion of ratios which make them rather accessible.

As a preparation for expanding the ratios, we start with the underlying two- and three-point functions. First note that we choose the arbitrary interpolating fields  $\mathcal{O}_{\text{bs},r}$ , Eq. (2.14), not to contain a  $1/m$  piece. Therefore, the  $1/m$  expansion of the two-point function reads

$$\log(\mathcal{C}^{\text{B}_s}(t)) = -m_{\text{bare}} t + \log(\mathcal{C}^{\text{B}_s, \text{stat}}(t)) + \sum_{k \in \{\text{kin}, \text{spin}\}} \omega_k \left( \frac{\mathcal{C}^{\text{B}_s, k}(t)}{\mathcal{C}^{\text{B}_s, \text{stat}}(t)} \right). \quad (6.1)$$

Only the (dimensionful) HQET parameters of the action,  $m_{\text{bare}}$  as well as  $\omega_k \sim 1/m$ , appear. Here and everywhere below all  $\mathcal{O}(1/m^2)$  terms are dropped without notice. Note also that  $m_{\text{bare}}$  drops out in the expressions for matrix elements.

The energy  $E^{\text{B}_s} = \lim_{t \rightarrow \infty} -\partial_t \log(\mathcal{C}^{\text{B}_s}(t))$  is  $1/m$ -expanded as  $E^{\text{B}_s} = m_{\text{bare}} + E^{\text{stat}} + \omega_{\text{kin}} E^{\text{kin}} + \omega_{\text{spin}} E^{\text{spin}}$  with  $E^x$  appearing in the large time behavior of<sup>44</sup>

$$-\partial_t \log \mathcal{C}_{rr}^{\text{B}_s, \text{stat}}(t) = E^{\text{stat}} + \mathcal{O}\left(e^{-\Delta E^{\text{stat}} t}\right), \quad (6.2)$$

$$-\partial_t \frac{\mathcal{C}_{rr}^{\text{B}_s, k}(t)}{\mathcal{C}_{rr}^{\text{B}_s, \text{stat}}(t)} = E^k + \mathcal{O}\left(te^{-\Delta E^{\text{stat}} t}\right), \quad k \in \{\text{kin}, \text{spin}\}. \quad (6.3)$$

All  $E^x$  refer to the ground state; the first excited state contribution leads to the term with  $\Delta E^{\text{stat}} = E_2^{\text{stat}} - E_1^{\text{stat}}$ . Because one always first expands in  $1/m$  and then takes a limit of large time, terms such as  $\exp(-\Delta E^{\text{kin}} t)$  do not appear.

In the numerical applications we will take  $\partial_t$  to be the forward derivative,  $a\partial_t f(t) = f(t+a) - f(t)$ . On integrating Eq. (6.3), we get:

$$\frac{\mathcal{C}_{rr}^{\text{B}_s, k}(t)}{\mathcal{C}_{rr}^{\text{B}_s, \text{stat}}(t)} = A_r^{\text{B}_s, k} - E^k t + \mathcal{O}\left(te^{-\Delta E^{\text{stat}} t}\right). \quad (6.4)$$

The integration constants  $A_r^{\text{B}_s, k}$  do depend on the kin and spin insertions as well as the smearing level used.

In complete analogy we have for the three-point functions ( $t = t_{\text{B}_s} = t_K$ ):

$$\frac{\mathcal{C}_{\mu, r}^{\text{B}_s \rightarrow \text{K}, k}(t, t)}{\mathcal{C}_{\mu, r}^{\text{B}_s \rightarrow \text{K}, \text{stat}}(t, t)} = A_{\mu, r}^{\text{B}_s \rightarrow \text{K}, k} - E^k t + \mathcal{O}\left(te^{-\Delta E^{\text{stat}} t}\right). \quad (6.5)$$

Now we turn to the ratios and insert the  $1/m$  expansion of the quantities which enter their definition in (5.1) and (5.2). To understand the structure,



consider the expansion

$$\mathcal{R}_{\mu,r}^I(t) = \mathcal{R}_{\mu,r}^{I,\text{stat}}(t) \left[ 1 + \sum_k \omega_k \rho_{\mu,r}^{I,k}(t) + \sum_j \omega_{\mu j} \rho_{\mu,r}^{I,j}(t) \right]$$

with

$$\rho_{\mu,r}^{I,k}(t, t) = \frac{\mathcal{C}_{\mu,r}^{B_s \rightarrow K, k}(t, t)}{\mathcal{C}_{\mu,r}^{B_s \rightarrow K, \text{stat}}(t, t)} - \frac{1}{2} \frac{\mathcal{C}_r^{B_s, k}(2t)}{\mathcal{C}_r^{B_s, \text{stat}}(2t)}, \quad k \in \{\text{kin}, \text{spin}\}, \quad (6.6)$$

$$\rho_{\mu,r}^{I,j}(t, t) = \frac{\mathcal{C}_{\mu,r}^{B_s \rightarrow K, j}(t, t)}{\mathcal{C}_{\mu,r}^{B_s \rightarrow K, \text{stat}}(t, t)}. \quad (6.7)$$

The sums over  $k$  run as indicated and for  $j$  the range is seen in Table 5. It follows that the desired  $1/m$  corrections to the ground-state matrix elements are given by the large time limits,

$$\rho_{\mu}^k = \lim_{t \rightarrow \infty} \rho_{\mu,r}^{I,k}(t) = A_{\mu,r}^{B_s \rightarrow K, k} - \frac{1}{2} A_r^{B_s, k}, \quad (6.8)$$

$$\rho_{\mu}^j = \lim_{t \rightarrow \infty} \rho_{\mu,r}^{I,j}(t). \quad (6.9)$$

These equalities are what we alluded to before as simplification in the  $1/m$  expansion. In fact, since the ground-state matrix elements in the last expression are just given by the  $A$  terms, it should not come as a surprise that the same final formulae for  $\rho_{\mu}^j$  and  $\rho_{\mu}^k$  are obtained if one starts from ratios  $\mathcal{R}^{\text{II}}$  or  $\mathcal{R}^{\text{III}}$ .

The strategy to obtain the  $1/m$  terms of the matrix elements is then to extract  $A_r^{B_s, k}$  and  $A_{\mu,r}^{B_s \rightarrow K, k}$  from fits to Eqs. (6.4) and (6.5) and  $\rho_{\mu}^j$  from Eq. (6.9).

### 6.1. Numerical results for the kin and spin insertions at $1/m$ order

As a first analysis we performed combined fits with parameters  $A_r^{B_s, k}$ ,  $A_{\mu,r}^{B_s \rightarrow K, k}$  and  $E^k$  to Eqs. (6.4) and (6.5) in the time region where a linear behavior in  $t$  is observed. This was done for fixed insertion  $k \in \{\text{kin}, \text{spin}\}$  and fixed smearing level, typically the largest smearing. The resulting errors on the matrix elements were rather large because the data did not constrain the fitted energies  $E^k$  so well. An improvement could be achieved by determining  $E^k$  from the GEVP at order  $1/m$ , exactly as described in Ref. 45 and then using that as a constraint in Eqs. (6.4) and (6.5). The GEVP takes into account information from all smearing levels of the two-point functions. In a little more detail, we expand the ground-state GEVP eigenvalues in  $1/m$  (see Ref. 45 for explicit formulae) as  $\lambda^{(0)}(t, t_0) = \lambda^{\text{stat}}(t, t_0) + \omega_{\text{kin}} \lambda^{\text{kin}}(t, t_0) + \omega_{\text{spin}} \lambda^{\text{spin}}(t, t_0)$  and then form

$$E_{\text{eff}}^k(t, t_0) = -\partial_t \log(\lambda^k(t, t_0)). \quad (6.10)$$

We consider just  $t_0 \geq t/2$  as the *asymptotic* convergence is proven to be much better<sup>45</sup> under that condition. For the present case, it turns out that there is rather

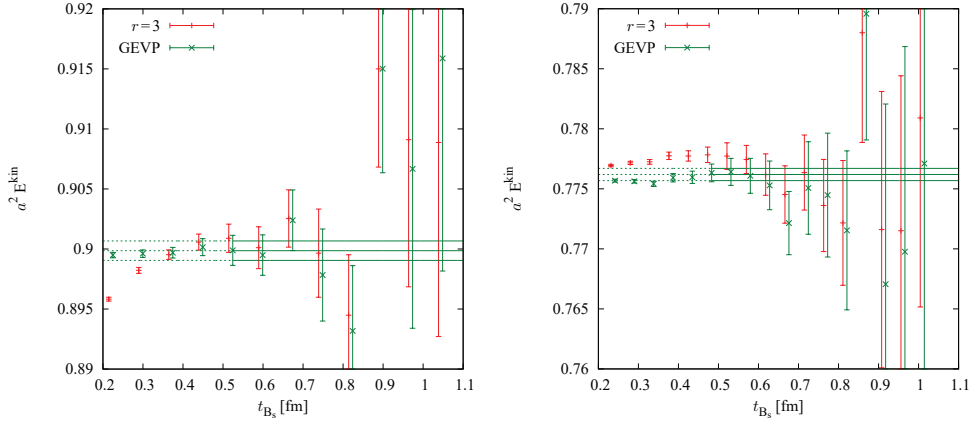


Fig. 15. Estimates for  $a^2 E_{\text{kin}}$  for ensembles A5 (left) and O7 (right). The plus symbols show Eq. (6.3) for the highest smearing and the crosses show the GEVP results of Eq. (6.10). The band represents the GEVP (weighted) plateaux average, taken over the range 0.5–1.0 fm.

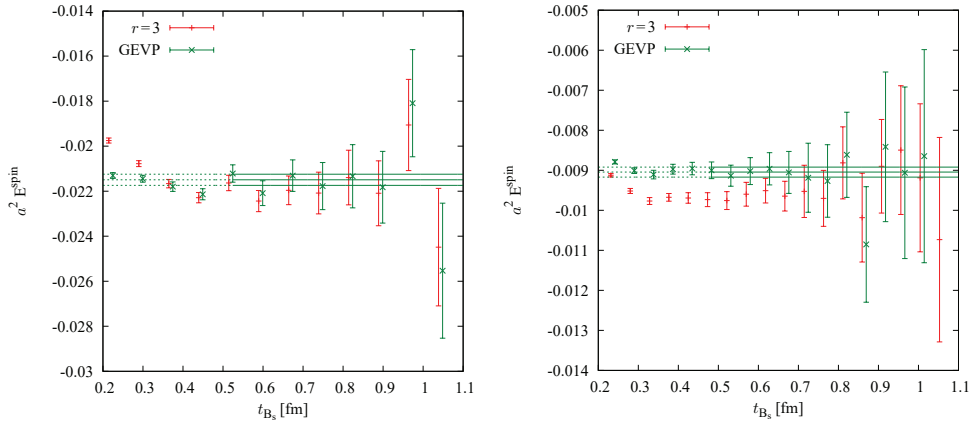


Fig. 16. Results for  $a^2 E_{\text{spin}}$  analogous to Fig. 15.

little dependence on  $t_0$  in practice but errors of course grow as it is increased. We then form a weighted average of the first (up to) three values with  $t_0$  at and above  $t/2$ . These averages are shown as  $E_{\text{eff}}^k(t)$  in Figs. 15 and 16 together with the effective masses of the best smearing level. The GEVP estimates behave significantly better than just the best smearing which we show for comparison. Our final numbers come from plateau fits of the GEVP estimates starting at  $t = 0.5$  fm, and are shown as bands in Figs. 15 and 16. The (uncorrelated) fits correspond to weighted averages with errors of the average computed by standard error propagation including the correlation of the points. Including points beyond about  $t = 1$  fm is not reasonable given their large uncertainties.

Table 4. Matrix elements at  $\mathcal{O}(1/m)$ : kin and spin contributions.

$\mu$	$k$	$\rho_\mu^k$ A5	$\rho_\mu^k$ O7
0	kin	-0.554(13)	-0.493(12)
1	kin	-0.316(24)	-0.335(21)
0	spin	0.363(3)	0.348(2)
1	spin	-0.223(5)	-0.177(3)

With these energies we determine estimates of  $A_r^{B_s,k}(t)$  by computing the weighted (plateau) average of  $\mathcal{C}_{rr}^{B_s,k}(t)/\mathcal{C}_{rr}^{B_s,\text{stat}}(t) + E^k t$  starting from  $t_{B_2}^{\min}$ , and similar for  $A_{\mu,r}^{B_s \rightarrow K,k}$ . These estimates are shown in Figs. 17 and 18 as a function of  $t_{B_2}^{\min}$  and  $t_{B_3}^{\min}$ . Precise, early and long plateaux are present for the two-point function  $A_r^{B_s,k}$  and we choose  $t_{B_2}^{\min} \approx 0.6$  fm for the final plateau averages. In contrast, the results for the three-point function  $A_{\mu,r}^{B_s \rightarrow K,k}$  show a stronger dependence on the start of the plateau averaging as long as  $t_{B_3}^{\min} < 0.8$  fm and in fact we would like the  $t$ -independence to be more convincing in one or two cases. Nevertheless, taking  $t_{B_3}^{\min} = 0.8$  fm is reasonable and we collect the results combined into  $\rho_\mu^k$  in Table 4. The bands in Figs. 17 and 18 show the chosen fit values.

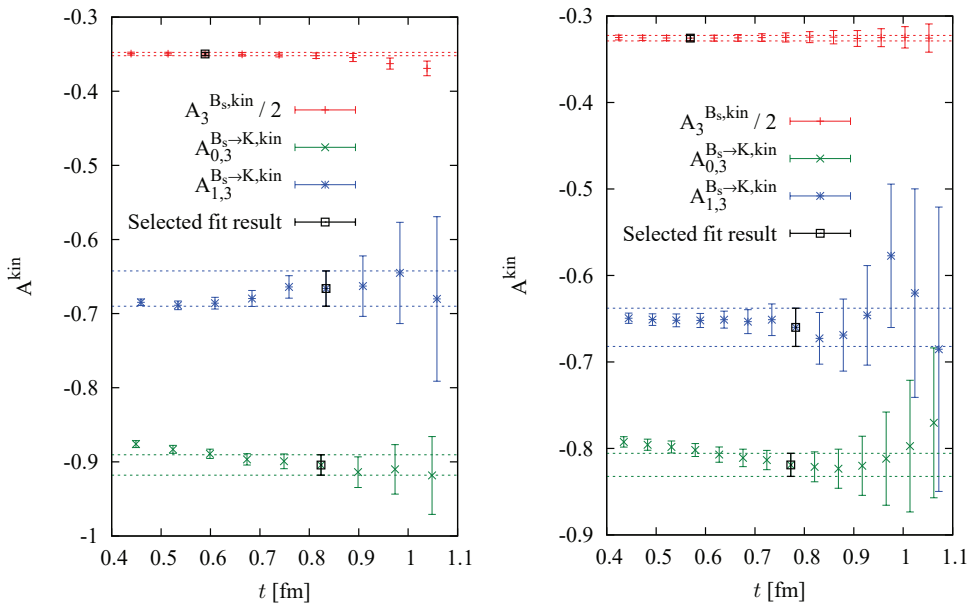


Fig. 17. Results for  $A_{\mu,3}^{B_s \rightarrow K,\text{kin}}$  and  $A_3^{B_s,\text{kin}}$  for ensembles A5 (left) and O7 (right). The horizontal axis refers to start  $t$ -values,  $t_{B_3}^{\min}$  and  $t_{B_2}^{\min}$ , of the plateaux averages. To guide the eye we add the band passing through the selected plateau average.

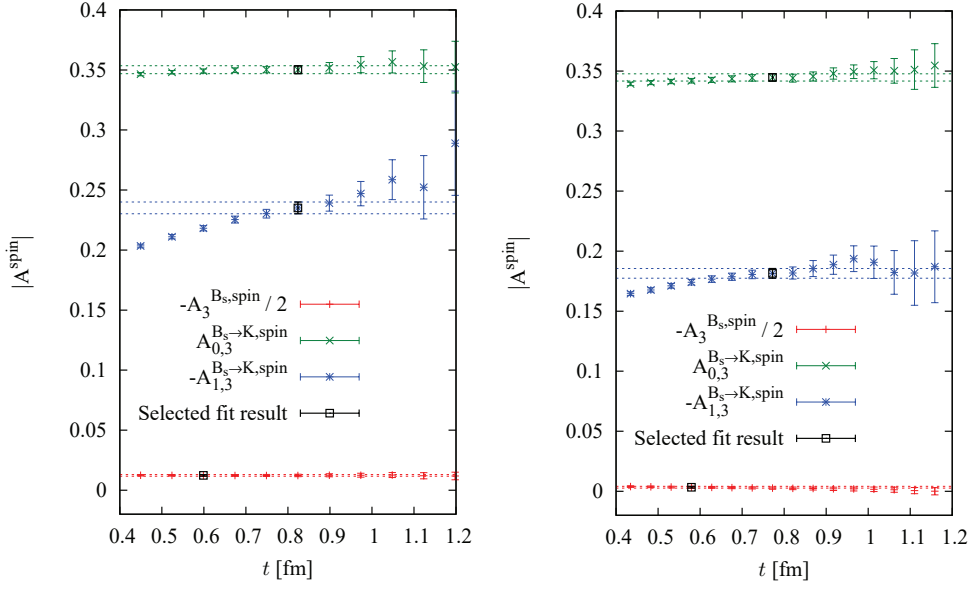


Fig. 18. Same as Fig. 17 for the spin amplitudes  $A_{\mu,3}^{B_s \to K,spin}$  and  $A_3^{B_s,spin}$ .

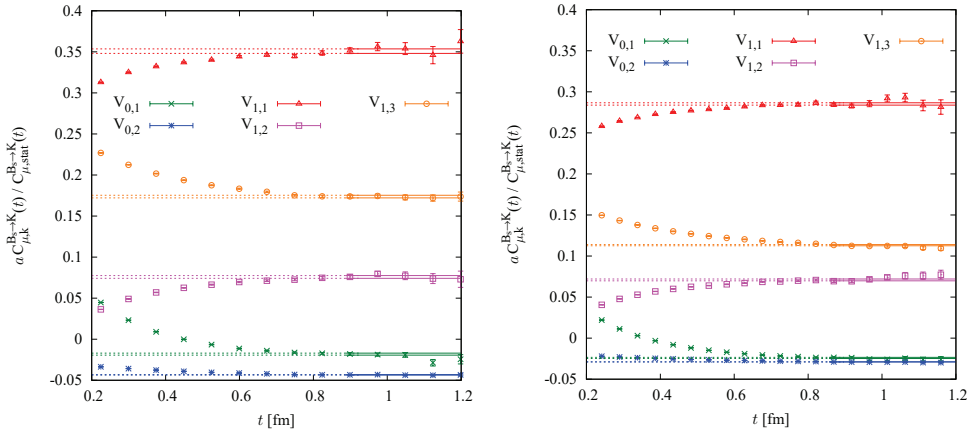


Fig. 19. Overview of the current insertions for ensembles A5 (left) and O7 (right). Fit bands are plateaus averages starting at 0.86 fm. For A5 the  $V_{1,4}$  contribution is exactly equal to  $V_{1,3}$ , while on the O7 they are not exactly identical but still indistinguishable on the plot, therefore  $V_{1,4}$  was not plotted.

## 6.2. Current insertions at $1/m$ order

The  $1/m$  vector current contributions are

$$V_0^{\text{HQET}}(x) = Z_{V_0}^{\text{HQET}} \left( V_0^{\text{stat}}(x) + \sum_{j=1}^2 \omega_{0,j} V_{0,j}(x) \right), \quad (6.11)$$

$$V_i^{\text{HQET}}(x) = Z_{V_i}^{\text{HQET}} \left( V_i^{\text{stat}}(x) + \sum_{j=1}^4 \omega_{i,j} V_{i,j}(x) \right), \quad (6.12)$$

$$V_{\mu,j}(x) = \bar{\psi}_u \Gamma_{\mu,j} \psi_h,$$

with the operators  $\Gamma_{\mu,j}$  detailed in Table 5. Note that with a momentum purely along the  $x$ -axis (A5 ensemble), only  $i = 1$  contributes and  $\rho_1^3 = \rho_1^4$  is exact.

Table 5. Overview of the  $1/m$  vector current insertions, their tree-level matching coefficients, and the results extracted from the highest light-quark smearing,  $r = 3$ . We use symmetric covariant derivatives  $\nabla_i^S$ .

$\mu$	$j$	$\Gamma_{\mu}^j$	$\omega_{\mu,j}^{\text{tree}} \cdot m_h$	$a\rho_{\mu}^j$ A5	$a\rho_{\mu}^j$ O7
0	1	$\sum_l \gamma_l \frac{1}{2} (\nabla_l^S - \overleftarrow{\nabla}_l^S)$	1/2	-0.0182(12)	-0.0241(7)
0	2	$\sum_l \gamma_l \frac{1}{2} (\nabla_l^S + \overleftarrow{\nabla}_l^S)$	1/2	-0.0434(4)	-0.0289(2)
$i$	1	$\sum_l \frac{1}{2} (\nabla_l^S - \overleftarrow{\nabla}_l^S) \gamma_l \gamma_i$	1/2	0.3508(27)	0.2852(15)
$i$	2	$\frac{1}{2} (\nabla_i^S - \overleftarrow{\nabla}_i^S)$	-1	0.0759(17)	0.0709(11)
$i$	3	$\sum_l \frac{1}{2} (\nabla_l^S + \overleftarrow{\nabla}_l^S) \gamma_l \gamma_i$	1/2	0.1737(16)	0.1132(8)
$i$	4	$\frac{1}{2} (\nabla_i^S + \overleftarrow{\nabla}_i^S)$	-1	0.1737(16)	0.1133(8)

The results obtained for Eq. (6.7) are presented in Fig. 19. We see plateaux starting at roughly 0.8–0.9 fm and average from 0.85 fm on. The precision is better than that of the kin and spin terms. A full quantitative error budget for the  $1/m$  contributions to the form factors has to wait until the corresponding nonperturbative matching coefficients  $\omega$  are available.

## 7. Discussion

Flavor changing transitions are important channels for learning about possible limitations of the standard model. Exclusive semi-leptonic decays of B or  $B_s$  mesons are particularly clean theoretically. However, lattice computations of the relevant form factors are needed and they are nontrivial in practice. A major reason is the infamous signal-to-noise problem. At large Euclidean time separations the noise in Monte Carlo evaluation of the correlation functions is too large to determine the form factors (matrix elements) with interesting precision.

While the issue is not new, we have exposed the problem in a few graphs more clearly than often done, see Figs. 4 and 14. For the three-point functions and derived (summed) ratios, this was possible because we have evaluated correlation functions at all time separations. We are only considering the pseudoscalar sector, where the signal-to-noise problem is very mild for a relativistic formulation and at zero momentum. However, we need finite momentum (see Fig. 2 for the momentum

dependence) and for reasons explained in the beginning of the paper we use HQET for the b-quark (see Fig. 4 for the difference of Kaon and static  $B_s$  meson). In HQET, the signal-to-noise problem becomes worse as one decreases the lattice spacing. Nevertheless, we are determining the matrix elements at total time separations  $\tau$  of the three-point functions of around 2 fm and for  $a \approx 0.05$  fm. Only for the summed ratio we use about half that time separation (Fig. 14), in agreement with the predicted better suppression of excited states after summation.

Despite our use of HQET, these separations are larger or similar to the ones typically used (e.g. most recently Ref. 9 used a fixed  $\tau \approx 2.2$  fm).

We have presented good evidence that the chosen plateaux or fit-windows are reasonably safe, but nevertheless it would be better to have larger times accessible. Maybe multilevel strategies<sup>46,47</sup> will help to reach those separations in the future. At present we derive our confidence from the good agreement of

- fits with  $t_{B_s} \gtrsim 0.3\text{--}0.6$  and 3 states in the  $B_s$  sector (and also with  $2B_s$  states and  $t_{B_s} \gtrsim 0.5\text{--}0.7$ , cf. App. B),
- ratios of Eq. (5.2) with  $\tau = t_{B_s} + t_K \gtrsim 2$  fm,
- and summed ratios of Eq. (5.4) at total separation  $\tau \gtrsim 1$  fm.

Note also that the precision of the different estimates is quite comparable. It is therefore preferable to use the technically easier ratio methods.

At the lowest (static) order in  $1/m$  the most relevant form factors for  $\mu = 1$  have an accuracy around 2%, which is good for precision physics. The first-order corrections in  $1/m$  are actually more precise than that. The relative errors induced into the form factors are given by the absolute ones of the quantities  $\rho$  in Tables 4 and 5 multiplied with the appropriate  $\omega$ -coefficients. From the errors of  $\rho^{\text{kin}}$  and  $\rho^{\text{spin}}$ , using the nonperturbative  $\omega_{\text{kin}}, \omega_{\text{spin}}$ ,<sup>13</sup> we get error contributions of 1% and 0.5% to the form factors, respectively. The uncertainties of  $\rho^j$ , translate into below 0.5% errors assuming coefficients  $\omega_j$  which do not exceed the tree-level values by more than a factor of two. Of course, the uncertainties in the coefficients  $\omega$  are to be added separately.

A positive result of our detailed analysis is thus that  $1/m$  corrections can be determined precisely, when the coefficients (HQET parameters) are known with reasonable accuracy. Once they are available it will be possible to provide further crosschecks on the existing analysis, see e.g. Fig. 13 of Ref. 9.

We would finally like to point out a possible danger in current and future semileptonic form factor computations on the lattice. It consists in the contribution of multihadron states, such as  $|B_s^{(*)}, \pi\rangle$  or  $|B_s^{(*)}, K\rangle$ , to our Euclidean correlation functions. Formally, these contributions are just particular ones in the sums over excited states in Eqs. (3.1)–(3.3) and are thus covered by our analysis. However, there are two properties which make such states special. First, when the spatial volume becomes large and the light-quark masses small, there are several low-lying states with small gaps  $E^{(n)} - E^{(0)}$ . All used methods may have difficulties in

separating those. This has also been pointed out recently in Ref. 48. Second, normalized overlaps  $\beta_r^{(n)}/\beta_r^{(0)}$  may be very similar for different smearing levels  $r$  when  $n$  corresponds to a multihadron state.<sup>k</sup> As mentioned in Subsec. 5.3 the GEVP method is of little help in such a situation. For B mesons these contaminations at finite  $t_{B_s}$  have not been investigated at all, while there is considerable discussion in the analogous determination of nucleon matrix elements (see Refs. 50, 51 and references therein). It appears likely that form factors  $B_s \rightarrow K$ ,  $B \rightarrow \pi$  are in somewhat better shape than nucleon matrix elements, since larger Euclidean time separations are reached. Nevertheless systematic studies, especially theoretical ones such as the ones carried out in chiral perturbation theory<sup>52,53</sup> for nucleon matrix elements are urgently needed to make quantitative statements.

## Acknowledgments

We would like to acknowledge useful discussions with Oliver Bär, Stefan Schaefer, Michele Della Morte and Brian Colquhoun.

We gratefully acknowledge the Gauss Centre for Supercomputing (GCS) for providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS share of the supercomputer JUQUEEN at Jülich Supercomputing Centre (JSC). GCS is the alliance of the three national supercomputing centers HLRS (Universität Stuttgart), JSC (Forschungszentrum Jülich), and LRZ (Bayerische Akademie der Wissenschaften), funded by the German Federal Ministry of Education and Research (BMBF) and the German State Ministries for Research of Baden-Württemberg (MWK), Bayern (StMWFK) and Nordrhein-Westfalen (MIWF). We acknowledge PRACE for awarding us access to resource JUQUEEN in Germany at Jülich and to resource SuperMUC in Germany at München. We also thank the LRZ for a CPU time grant on SuperMUC, project pr85ju, and DESY for access to the PAX cluster in Zeuthen.

## Appendix A. Computation of the Correlation Functions

Integration over the fermion fields yields the correlation functions as

$$\mathcal{C}_{\mu,r}^{B_s \rightarrow K}(t_K, t_{B_s}; \mathbf{p}) = \langle \hat{\mathcal{C}}_{\mu,r}^{B_s \rightarrow K} \rangle_U, \quad \mathcal{C}^K(t; \mathbf{p}) = \langle \hat{\mathcal{C}}^K \rangle_U, \quad \mathcal{C}_{rr'}^{B_s}(t) = \langle \hat{\mathcal{C}}_{rr'}^{B_s} \rangle_U \quad (\text{A.1})$$

averaged over the gauge fields,  $U$ , (with effective action including the log of the quark determinant). The dependence of the functions

$$\hat{\mathcal{C}}_{\mu,r}^{B_s \rightarrow K} = \frac{a^9}{L^3} \text{Tr} [\gamma_5 P_f F_- W_K S_u \gamma_\mu P_v F_+ S_h P_l \gamma_5 W_r S_s W_K], \quad (\text{A.2})$$

$$\hat{\mathcal{C}}^K = \frac{a^6}{L^3} \text{Tr} [\gamma_5 P_f F_- W_K S_u W_K \gamma_5 P_l F_+ W_K S_s W_K], \quad (\text{A.3})$$

<sup>k</sup>See the discussion of the field in the chiral effective theory representing the action density at finite flow time,  $E(x, t)$  in Ref. 49.

$$\hat{\mathcal{C}}_{rr'}^{\text{B}_s} = \frac{a^6}{L^3} \text{Tr}[\gamma_5 P_f S_h P_i \gamma_5 W_{r'} S_s W_r] \quad (\text{A.4})$$

on the times  $(x_0)_s$  ( $s = f, v, i$ ) as well as the gauge fields is suppressed. In the above expressions  $S_q$  are the quark propagators,  $W_K = G^{I^K}$ ,  $G = 1 - \kappa_G a^2 \Delta$  and  $W_r = G^{I_r^{\text{B}_s}}$  are the smearing operators,  $P_s$  is the projection on time slice  $(x_0)_s$  and  $F_{\pm} = \exp(\pm i \mathbf{x} \mathbf{p})$  is the multiplication with the Fourier phases on the corresponding time slice. Evaluation of the full traces which include the space coordinates becomes possible by representing the projector

$$P_f = \langle \eta \eta^\dagger \rangle_\eta \quad (\text{A.5})$$

in terms of a random  $U(1)$  field  $\eta$  with support on time slice  $(x_0)_f$ . This yields

$$\hat{\mathcal{C}}_{\mu,r}^{\text{B}_s \rightarrow \text{K}} = \frac{a^9}{L^3} \langle \eta^\dagger F_- \gamma_5 W_K S_u P_v \gamma_\mu F_+ S_h P_i \gamma_5 W_r S_s W_K \eta \rangle_\eta \quad (\text{A.6})$$

$$= \langle \phi_u^\dagger \gamma_5 P_v \gamma_\mu F_+ S_h P_i \gamma_5 W_r \phi_s \rangle_\eta, \quad (\text{A.7})$$

$$\phi_s = S_s W_K \eta, \quad \phi_u = S_u W_K F_-^\dagger \eta \quad (\text{A.8})$$

and similar for the two-point functions. In this form we see that for each vector  $\eta$  two solutions of the Dirac equation are needed in order to compute the fields  $\phi_u, \phi_s$ . The static propagator  $S_h$  is inserted by explicit forward propagation. Translation invariance in time is used by averaging over all source time slices  $(x_0)_f$ , each one with a random  $U(1)$  field  $\eta$ . Since the averages in Eqs. (A.1) and (A.5) are independent of each other, any number of  $\eta$  fields per gauge field is correct; we use a single one per gauge field and time  $(x_0)_f$ .

In practice, we compute propagators of periodic quark fields in a gauge field  $e^{i\theta_\mu a/L} U(x, \mu)$ , which includes the constant  $U(1)$  background field  $\theta_\mu^q a$  and only use the integer part of the momentum in the Fourier factors  $F_{\pm}$ . This is equivalent to Eq. (2.18), apart from a phase in the Gaussian smearing, which we set to zero. Choosing a different phase in the smearing along the lines of Ref. 54 might be a further optimization.

### A.1. Improvement and $1/m$ terms

The  $1/m$  terms are simple generalizations of the above. First,  $[\mathcal{C}_\mu^{\text{B}_s \rightarrow \text{K}}]_k$ ,  $k \in \text{kin}$ , spin and  $[\mathcal{C}_{rr'}^{\text{B}_s}]_k$ , are given by replacing  $S_h \rightarrow S_k$ , where the latter are

$$S_{\text{kin}} = S_h \nabla_i^* \nabla_i S_h, \quad (\text{A.9})$$

$$S_{\text{spin}} = \frac{i}{4} S_h [\gamma_i, \gamma_j] \hat{F}_{ij} S_h \quad (\text{A.10})$$

and the chromo-magnetic field strength tensor  $\hat{F}_{ij}$  is discretized in terms of the clover leaf, see e.g. Ref. 21.

Second, the NLO three-point functions  $[\mathcal{C}_\mu^{\text{B}_s \rightarrow \text{K}}]_j$  are given by the substitution  $\gamma_\mu \rightarrow \Gamma_\mu^j$  in Eqs. (A.2) and (A.7), with  $\Gamma_\mu^j$  listed in Table 5. The  $\mathcal{O}(a)$  improvement corrections to the static three-point functions are just linear combinations thereof.



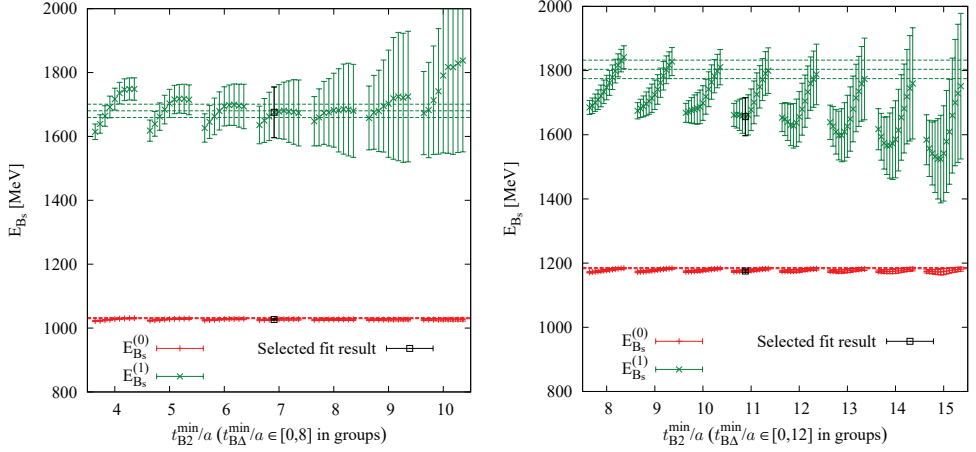


Fig. B.1. (Color online) Stability plots for the  $B_s$  energies on ensemble A5 (left) and O7 (right). The selected  $t_{B2/3}^{\min}$  values are highlighted with a black color, while the dashed lines are values from the GEVP, cf. Subsec. 3.2.1. A zoom into the behavior of  $E_{B_s}^{(0)}$  is shown in Fig. B.2.

## Appendix B. Alternative Static Fit with $N_K = 1$ , $N_{B_s} = 2$

In this section we discuss a simpler version of the fit which includes only one excited state in the  $B_s$  sector. We use only the two highest  $B_s$  smearings, therefore we have a total of 12 fit parameters (as opposed to 20 in the fit with  $N_{B_s} = 3$ ).

Similar to Sec. 4 we show the  $B_s$  energy states in Fig. B.1. Here, the growth of the errors at larger  $t_{B2/3}^{\min}$  is much milder than for the  $N_{B_s} = 3$  fit, but on the other hand the stability of the results with respect to changes of  $t_{B2/3}^{\min}$  is in some cases rather unsatisfying. It can be seen even more clearly for  $E_{B_s}^{(0)}$  (cf. Fig. B.2) which, in particular for O7, shows a characteristic drift upwards with growing  $t_{B3}^{\min}$ .

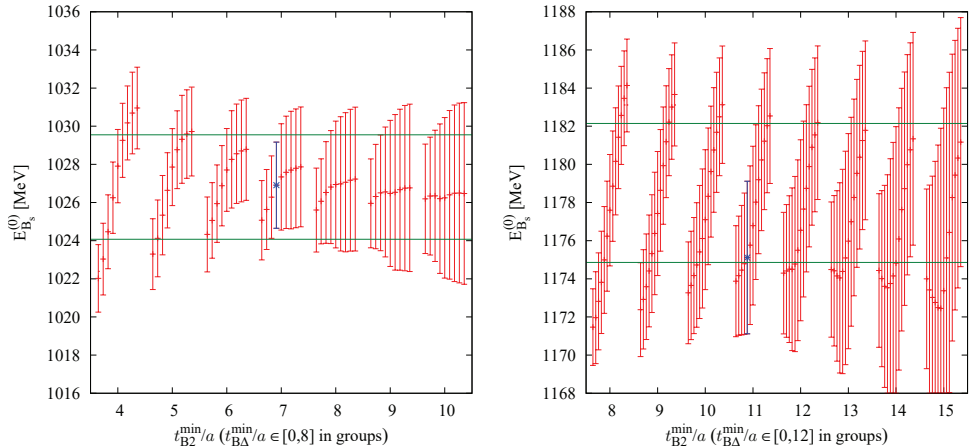


Fig. B.2. Stability plots for  $E_{B_s}^{(0)}$  for ensembles A5 (left) and O7 (right).

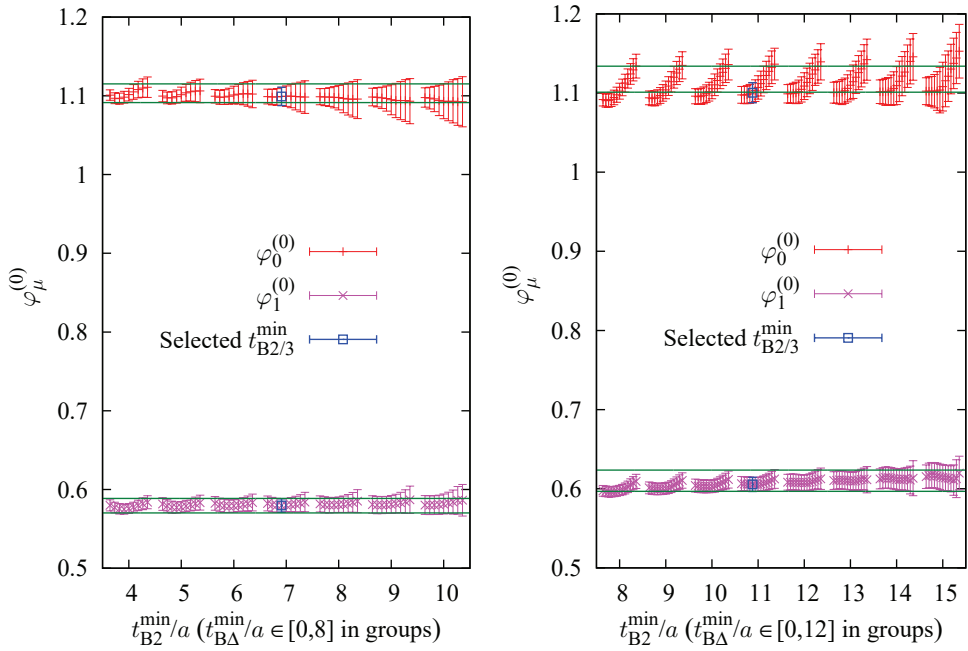


Fig. B.3. Stability plots of the ground-state form factors for ensembles A5 (left) and O7 (right).

On the other hand, considering these drifts in energies (and amplitudes as well), the ground-state form factors show reasonable stability and agreement with the  $N_{B_s} = 3$  fit results, cf. Fig. B.3. The case of  $\varphi_0^{(0)}$  for O7 is by far the worst, but most look more like the stable and consistent  $\varphi_1^{(0)}$ .

When choosing a final form factor value from the fit, one clearly has to keep  $t_{B2/3}^{\min}$  slightly higher than for the  $N_{B_s} = 3$  fit. We use  $t_{B2}^{\min} \approx 0.52$  fm and the same  $t_{B\Delta}^{\min}$  as before. The values summarized in Table B.1 are consistent within errors with other methods and give similar precision.

Table B.1. Results for the ground-state form factors  $\varphi_\mu^{(0)}$  using the fit with  $N_{B_s} = 2$ .

id	$\mu$	Fit value
A5	0	1.099(12)
A5	1	0.580(9)
O7	0	1.100(13)
O7	1	0.606(9)

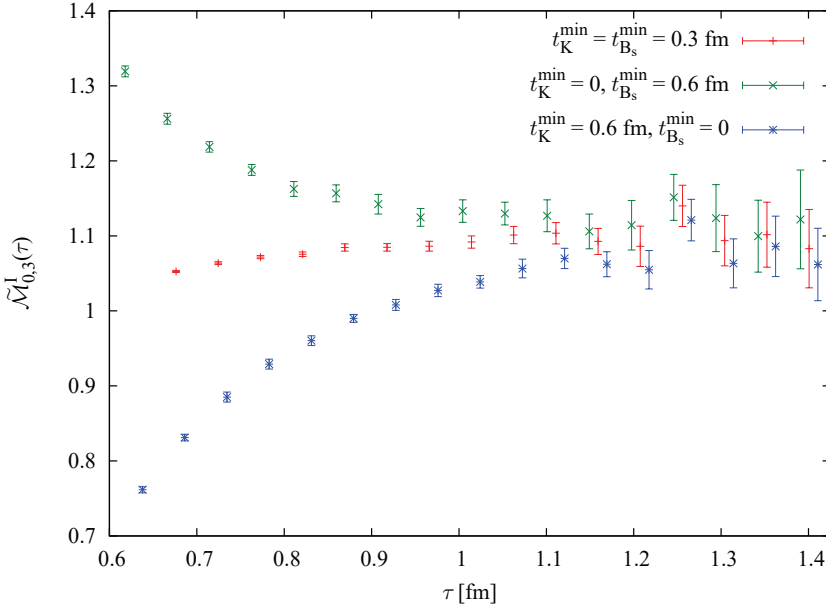


Fig. B.4. The summed ratio for ensemble O7,  $\mu = 0$  with three versions: with symmetric sum limits and with either only  $t_{B_s}^{\min}$  or  $t_K^{\min}$  larger than 0. Note that due to different (worse) convergence properties, these plots cannot be directly compared with the ones in Fig. 14.

### Appendix C. Note on the Convergence of the Static Summed Ratio

In Subsec. 5.2, devoted to static summed ratios, we observe that the plateaux for  $\mu = 0$  in Fig. 14 seem to start very early. Taking the results for O7 at face value, one could start the plateau fit as early as 0.4 fm. To see whether this is a genuine ground state dominance we investigate truncated sums of the form:

$$\tilde{\mathcal{M}}_{\mu,r}^I(\tau) = \partial_\tau a \sum_{t_{B_s}=t_{B_s}^{\min}}^{\tau-t_K^{\min}} \mathcal{R}_{\mu,r}^I(\tau-t_{B_s}, t_{B_s}). \quad (\text{C.1})$$

By putting  $t_{B_s}^{\min}$  or  $t_K^{\min}$  larger than 0 we can suppress the excited states in the  $B_s$  and K sector respectively.<sup>1</sup> The results of this procedure are shown in Fig. B.4. We see that the excited states in the K sector push the result upwards, while the excited states in the  $B_s$  sector push the results downwards. Their cancellation results in a flat “fake” plateau which can start at very early times. A similar but much less pronounced effect is at work for  $\mu = 1$ . One should therefore not rely solely on the flatness of the plots but also devise an independent criterion for the beginning of the plateaux, as was done in Subsec. 5.2.

<sup>1</sup>Note however that this increases the excited state corrections from  $\sim \tau \Delta \exp(-\tau \Delta)$  to  $\sim \tau' \Delta \exp(-\tau' \Delta)$  with  $\tau' = \tau - t_{B_s}^{\min} - t_K^{\min}$ .

## References

1. FERMILAB LATTICE, MILC Collabs. (J. A. Bailey *et al.*), *Phys. Rev. D* **92**, 014024 (2015), arXiv:1503.07839.
2. J. M. Flynn *et al.*, *Phys. Rev. D* **91**, 074510 (2015), arXiv:1501.05373.
3. C. M. Bouchard, G. P. Lepage, C. Monahan, H. Na and J. Shigemitsu, *Phys. Rev. D* **90**, 054506 (2014), arXiv:1406.2279.
4. B. Colquhoun, R. J. Dowdall, J. Koponen, C. T. H. Davies and G. P. Lepage, *Phys. Rev. D* **93**, 034502 (2016), arXiv:1510.07446.
5. B. Colquhoun, S. Hashimoto and T. Kaneko, *EPJ Web Conf.* **175**, 13004 (2018), arXiv:1710.07094.
6. Z. Gelzer *et al.*, *EPJ Web Conf.* **175**, 13024 (2018), arXiv:1710.09442.
7. C. J. Monahan, C. M. Bouchard, G. P. Lepage, H. Na and J. Shigemitsu, *Phys. Rev. D* **98**, 114509 (2018), arXiv:1808.09285.
8. FERMILAB LATTICE, MILC Collabs. (Y. Liu *et al.*), *EPJ Web Conf.* **175**, 13008 (2018), arXiv:1711.08085.
9. FERMILAB LATTICE, MILC Collabs. (A. Bazavov *et al.*),  $B_s \rightarrow K \ell \nu$  decay from lattice QCD, arXiv:1901.02561.
10. F. Bahr *et al.*, *Phys. Lett. B* **757**, 473 (2016).
11. J. Heitger and R. Sommer, *J. High Energy Phys.* **0402**, 022 (2004).
12. B. Blossier, M. Della Morte, N. Garron and R. Sommer, *J. High Energy Phys.* **1006**, 002 (2010).
13. B. Blossier *et al.*, *J. High Energy Phys.* **1209**, 132 (2012).
14. M. Della Morte, S. Dooling, J. Heitger, D. Hesse and H. Simma, *J. High Energy Phys.* **1405**, 060 (2014).
15. D. Hesse and R. Sommer, *J. High Energy Phys.* **1302**, 115 (2013).
16. C. Best *et al.*, *Phys. Rev. D* **56**, 2743 (1997), arXiv:hep-lat/9703014.
17. R. Sommer, Introduction to non-perturbative heavy quark effective theory, in *Modern Perspectives in Lattice QCD: Quantum Field Theory and High Performance Computing. Proceedings*, International School, 93rd Session, Les Houches, France, 3–28 August 2009 (2010), pp. 517–590.
18. B. Blossier *et al.*, *J. High Energy Phys.* **1012**, 039 (2010).
19. P. Fritzsch *et al.*, *Nucl. Phys. B* **865**, 397 (2012), , arXiv:1205.5380.
20. G. P. Engel, L. Giusti, S. Lottini and R. Sommer, *Phys. Rev. Lett.* **114**, 112001 (2015).
21. ALPHA Collab. (K. Jansen and R. Sommer), *Nucl. Phys. B* **530**, 185 (1998), arXiv:hep-lat/9803017.
22. K. G. Wilson, *Phys. Rev. D* **10**, 2445 (1974).
23. ALPHA Collab. (S. Lottini), *Acta Phys. Polon. Suppl.* **7**, 565 (2014), arXiv:1406.2939.
24. ALPHA Collab. (M. Della Morte *et al.*), *Phys. Lett. B* **581**, 93 (2004), arXiv:hep-lat/0307021.
25. M. Della Morte, A. Shindler and R. Sommer, *J. High Energy Phys.* **0508**, 051 (2005), arXiv:hep-lat/0506008.
26. A. Grimbach, D. Guazzini, F. Knechtli and F. Palombi, *J. High Energy Phys.* **03**, 039 (2008), arXiv:0802.0862.
27. C. Alexandrou, F. Jegerlehner, S. Gusken, K. Schilling and R. Sommer, *Phys. Lett. B* **256**, 60 (1991).
28. S. Gusken, U. Löw, K. H. Mütter, R. Sommer, A. Patel and K. Schilling, *Phys. Lett. B* **227**, 266 (1989).
29. APE Collab. (M. Albanese *et al.*), *Phys. Lett. B* **192**, 163 (1987).
30. M. Lüscher and P. Weisz, *Commun. Math. Phys.* **97**, 59 (1985).

31. F. T. Bahr, Form factors for semileptonic  $B_s \rightarrow K\ell\nu$  decays in lattice QCD, Ph.D. thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät (2015), doi:10.18452/17384.
32. M. Lüscher and U. Wolff, *Nucl. Phys. B* **339**, 222 (1990).
33. B. Blossier, M. Della Morte, G. von Hippel, T. Mendes and R. Sommer, *J. High Energy Phys.* **0904**, 094 (2009).
34. F. Bernardoni *et al.*, *Phys. Lett. B* **730**, 171 (2014).
35. ALPHA Collab. (U. Wolff), *Comput. Phys. Commun.* **156**, 143 (2004), arXiv:hep-lat/0306017.
36. ALPHA Collab. (S. Schaefer, R. Sommer and F. Virotta), *Nucl. Phys. B* **845**, 93 (2011), arXiv:1009.5228.
37. C. Michael, *Phys. Rev. D* **49**, 2616 (1994).
38. C. Michael and A. McKerrell, *Phys. Rev. D* **51**, 3745 (1995).
39. B. Bunk, private communications (1985).
40. M. Bruno and R. Sommer, in preparation, see also “Data analysis” at <https://indico.desy.de/indico/event/9420/>.
41. L. Maiani, G. Martinelli, M. L. Paciello and B. Taglienti, *Nucl. Phys. B* **293**, 420 (1987).
42. J. Bulava, M. Donnellan and R. Sommer, *J. High Energy Phys.* **01**, 140 (2012), arXiv:1108.3774.
43. S. Capitani, B. Knippschild, M. Della Morte and H. Wittig, *PoS LATTICE2010*, 147 (2010), arXiv:1011.1358.
44. F. Bernardoni *et al.*, *Phys. Lett. B* **730**, 171 (2014).
45. B. Blossier, M. Della Morte, G. von Hippel, T. Mendes and R. Sommer, *J. High Energy Phys.* **0904**, 094 (2009), arXiv:0902.1265.
46. M. Cè, L. Giusti and S. Schaefer, *Phys. Rev. D* **95**, 034503 (2017), arXiv:1609.02419.
47. M. Cè, L. Giusti and S. Schaefer, *EPJ Web Conf.* **175**, 11005 (2018), arXiv:1711.01592.
48. S. Hashimoto, Hints and challenges in heavy flavor physics, in *36th Int. Symp. on Lattice Field Theory (Lattice 2018)*, East Lansing, MI, USA, 22–28 July 2018, arXiv:1902.09119.
49. O. Bär and M. Golterman, *Phys. Rev. D* **89**, 034505 (2014), arXiv:1312.4999.
50. O. Bär, *EPJ Web Conf.* **175**, 01007 (2018), arXiv:1708.00380.
51. J. Green, Systematics in nucleon matrix element calculations, in *36th Int. Symp. on Lattice Field Theory (Lattice 2018)*, East Lansing, MI, USA, 22–28 July 2018, arXiv:1812.10574.
52. O. Bär, *Phys. Rev. D* **94**, 054505 (2016), arXiv:1606.09385.
53. O. Bär,  $N\pi$ -state contamination in lattice calculations of the nucleon axial form factors, arXiv:1812.09191.
54. G. S. Bali, B. Lang, B. U. Musch and A. Schäfer, *Phys. Rev. D* **93**, 094515 (2016), arXiv:1602.05525.