On the computation of hadron-to-hadron transition matrix elements in lattice QCD

ALPHA
Collaboration

John Bulava, Michael Donnellan and Rainer Sommer

NIC, DESY,
Platanenallee 6, 15738 Zeuthen, Germany

E-mail: john.bulava@cern.ch, michael.donnellan@desy.de,
rainer.sommer@desy.de

ABSTRACT: We discuss the accurate determination of matrix elements $\langle f|\hat{h}_{\rm w}|i\rangle$ where neither $|i\rangle$ nor $|f\rangle$ is the vacuum state and $\hat{h}_{\rm w}$ is some operator. Using solutions of the Generalized Eigenvalue Problem (GEVP) we construct estimators for matrix elements which converge rapidly as a function of the Euclidean time separations involved. $|i\rangle$ and $|f\rangle$ may be either the ground state in a given hadron channel or an excited state. Apart from a model calculation, the estimators are demonstrated to work well for the computation of the $B^*B\pi$-coupling in the quenched approximation. They are also compared to a standard ratio as well as to the "summed ratio method" of [1–3]. In the model, we also illustrate the ordinary use of the GEVP for energy levels.

KEYWORDS: Lattice QCD, B-Physics, QCD

Contents

1 Introduction

In lattice QCD, masses, energies and vacuum-to-hadron matrix elements are extracted from the large time asymptotics of Euclidean two-point correlation functions. Convergence to ground state matrix elements and energies proceeds with a rate of order $\exp(-\Delta^{(A)} t)$ where $\Delta^{(A)}$ is the energy difference of the first excited state and the ground state in the hadron channel characterized by a set of quantum numbers A.

Hadron-to-hadron matrix elements of the type $\langle A|\hat{h}_{\mathrm{w}}|B\rangle$ require us to consider in addition three-point functions, which contain two time separations,

$$C^{(3)}(t_2,t_1) = \langle \mathcal{O}^{(A)}(t_2 + t_1)\, h_{\mathrm{w}}(t_1)\, [\mathcal{O}^{(B)}]^*(0)\rangle \tag{1.1}$$

and the corrections to asymptotic behavior are $\mathrm{O}(\exp(-\Delta^{(A)}\,t_2),\ \exp(-\Delta^{(B)}\,t_1))$. Hence one wants both t_1 and t_2 to be large. On the other hand, at large times the statistical errors of the Monte Carlo estimates typically increase: the noise-to-signal ratio grows with a rate $\exp(\delta^{(A)}t_2 + \delta^{(B)}t_1)$, where $\delta^{(i)}$ is a positive energy difference. It is therefore necessary to compromise between the two sources of error. This compromise represents an important limitation to the achievable overall precision (statistical *and* systematic).

As an example, consider the nucleon axial coupling g_A. For this application, $\Delta^{(A)} = \Delta^{(B)}$ due to isospin symmetry and $\delta^{(A)} = \delta^{(B)} \approx m_{\mathrm{nucleon}} - \frac{3}{2}m_\pi$ [4, 5].[1] Due to the symmetry, $t_1 = t_2$ is optimal in this simple but relevant example and

- $t_1 = t_2 \gg 1/\Delta^{(A)} \approx 0.5\,\mathrm{fm}$ is required to keep *systematic corrections* due to excited states small

- but *statistical errors* become too large beyond a time $t = \mathrm{O}(1/\delta^{(A)}) = \mathrm{O}(1\,\mathrm{fm})$.[2]

Numerical results have been shown for this particular example in the reviews [8, 9] and recently in refs. [3, 10].

As a remedy one may try to reduce either the statistical uncertainties or the contamination by excited states. A general idea for reducing statistical fluctuations is to integrate over part of the configuration space analytically or by a multilevel algorithm [11–13]. In the pure gauge theory, a reduction of the growth of statistical errors as a function of time has successfully been achieved by multilevel algorithms [11, 13] as well as by symmetry constrained Monte Carlo [14, 15], but it appears difficult to make further progress in this direction for QCD with dynamical fermions. More radically, in lower-dimensional models a complete rewriting of the path integral led to simulation methods where errors can be kept constant at large time in specific channels [16]. Returning to more moderate gains, in the Heavy Quark Effective Theory (HQET) one is in a special situation because $\delta^{(A)}$ is power divergent and depends strongly on the discretization. An optimization of the discretization of HQET yielded a much reduced $\delta^{(A)}$ [7]. Despite these advances, we do not have a true solution of the signal-to-noise problem either in QCD or in HQET. It is therefore important to efficiently exploit the information present in an available set of gauge fields. In particular, due to translational invariance it is possible to construct volume-averaged estimators which should have reduced variance compared to those in which one of the fields has a fixed position. This requires the stochastic estimation of the all-to-all quark propagator [17–20] rather than the traditional calculation of a point-to-all propagator, and given the aforementioned exponential growth of signal-to-noise in Euclidean time, it is essential to "dilute the noise sources" (notation of [19]) such that each has support only on a single time-slice.

It is then natural to try to reduce the systematic corrections due to excited states. As a first step, one improves the interpolating fields $\mathcal{O}^{(A)}$, usually by smearing (see section 2). In this way one may reduce the prefactor of $\exp(-\Delta^{(A)}\,t_2)$.

---

[1]Another example of interest is the $B^*B\pi$ coupling where $\Delta^{(A)} = \Delta^{(B)}$ due to heavy quark spin symmetry and the energy difference $\delta^{(A)} = \delta^{(B)}$ is discussed in [6, 7].

[2]For our numbers we consider pion masses above the physical one. Close to the physical point the situation is somewhat worse.

But in ref. [21] it has been pointed out that by considering N interpolating fields and the Generalized Eigenvalue Problem (GEVP), one can construct a time-dependent effective GEVP-optimized interpolating field where the gap $\Delta^{(A)}$ is enhanced to

$$\Delta^{(A)} = E_{N+1}^{(A)} - E_n^{(A)} . \tag{1.2}$$

Besides raising the relevant gap for the ground state $n = 1$ by a considerable amount, excited states $n > 1$ can then also be reached in each channel! This was demonstrated to work very well for a decay constant in HQET, i.e. a matrix element of the type $\langle f|\hat{h}_{\mathrm{w}}|0\rangle$. In this work we show that it is also very advantageous for non-vacuum matrix elements.

Moreover, we present a new formula, which involves the 3-point matrix correlation function and the GEVP eigenpairs. There is a summation over the intermediate time t_1 with $t = t_1 + t_2$ held fixed. We consider now the "symmetric case" when initial and final states are related by a symmetry transformation (e.g. for g_A), since the general case without the symmetry is more complicated as we will explain in the following sections. When $\Delta^{(A)}t \gg 1$, the corrections to the matrix element are reduced

$$\begin{aligned} \text{from} \quad & C \exp(-\Delta^{(A)} t/2) && \text{(fixed } t_1 = t_2 = t/2 \text{ GEVP)} \\ \text{to} \quad & C'\Delta^{(A)} t \exp(-\Delta^{(A)} t) && (\text{"summed" GEVP}), \end{aligned}$$

with some coefficients C, C' given by matrix elements which are usually unknown. The $N = 1$ case of the general formula reduces to the summed 3-point function that has previously been used in early investigations of the nucleon sigma term [1] and g_A [22]. The improvement of the convergence rate to the ground state has recently been emphasized in [2, 3]. Let us leave aside the prefactors C, C', about which little can be said on general grounds. The remaining time-dependent factors satisfy $\Delta^{(A)} t \exp(-\Delta^{(A)} t) < \exp(-\Delta^{(A)} t/2)$ for all t. Furthermore, when one is in the asymptotic regime $\Delta^{(A)} t \gg 1$ the gain becomes significant: the "summed" GEVP method requires approximately half the total time separation for the same size systematic corrections.

The derivation of the formula for the matrix element, as well as the associated correction terms, proceeds roughly as follows. We start from the GEVP expression for the energy levels, in a theory with degeneracy $E^{(A)} = E_n^{(B)}$, and augment the theory by a "source" term $\epsilon \hat{h}_{\mathrm{w}}$ in the Hamiltonian. The matrix element $\langle A, n|\hat{h}_{\mathrm{w}}|B, n\rangle$ is then obtained as a derivative with respect to ϵ of the effective (time dependent) energy level at $\epsilon = 0$. This idea is worked out in section 3. In section 4, we report tests of the method for a toy model and also in a quenched QCD/HQET calculation. There we revisit the GEVP for energy levels, using the overlaps computed in the quenched case to fix the parameters of the toy model and examining the convergence of energies in the model. As we will discuss in the conclusions, we expect our method to be advantageous in a number of applications. First we set up the notation and describe the standard GEVP method of ref. [21].

2 Matrix elements from Euclidean correlators

In this section we define the problem more precisely and describe the "standard" solution as well as the one using the GEVP.

We want to compute a matrix element of a local operator $\hat{h}_{\mathrm{w}}(\mathbf{x})$,

$$\mathcal{M}_{mn} = \langle A, m | \hat{h}_{\mathrm{w}}(0) | B, n \rangle \,, \tag{2.1}$$

where $m, n \geq 1$ label the excitations in each channel. The quantum numbers A and B associated with exact symmetries of the (lattice) Hamiltonian including e.g. momentum or flavors charges remain implicit in $\mathcal{M}$. For the lattice Hamiltonian derived from the transfer matrix, we have $\hat{H} | A, m \rangle = E_m^{(A)} | A, m \rangle$. We take the finite (space-) volume normalization of states $\langle A, m | A, m \rangle = 1$, which is easily related to the relativistic one.

The matrix elements are computed from correlation functions

$$C_{ij}^{(3)}(t_2, t_1) = \langle \mathcal{O}_i^{(A)}(t_2 + t_1) \, h_{\mathrm{w}}(t_1) \, [\mathcal{O}_j^{(B)}(0)]^* \rangle \tag{2.2}$$

$$C_{ij}^{(A)}(t) = \langle \mathcal{O}_i^{(A)}(t) \, [\mathcal{O}_j^{(A)}(0)]^* \rangle \,, \quad C_{ij}^{(B)}(t) = \langle \mathcal{O}_i^{(B)}(t) \, [\mathcal{O}_j^{(B)}(0)]^* \rangle \tag{2.3}$$

where $\mathcal{O}_j^{(B)}(t)$ are interpolating fields localized on a time-slice t with j enumerating different fields. They carry the quantum numbers B in the usual way. We now turn to different ways of analyzing the correlation functions.

2.1 Standard ratios

We consider $m = n = 1$ for describing the "standard" method, since it is largely restricted to ground states. One defines a ratio

$$\mathcal{R}(t_2, t_1) = \frac{C_{ij}^{(3)}(t_2, t_1)}{[C_{ii}^{(A)}(t) C_{jj}^{(B)}(t)]^{1/2}} \exp\left((E_B^{\mathrm{eff}}(t) - E_A^{\mathrm{eff}}(t))(t_1 - t_2)/2 \right) \tag{2.4}$$

for fixed i, j with

$$t = t_2 + t_1 \,, \quad E_A^{\mathrm{eff}}(t) = -\partial_t \log(C_{ii}^{(A)}(t)) \,, \quad E_B^{\mathrm{eff}}(t) = -\partial_t \log(C_{jj}^{(B)}(t)) \,.$$

Our lattice derivative is defined as $\partial_t f(t) = \frac{1}{a}[f(t + a) - f(t)]$. When the sectors (A) and (B) are related by a symmetry of the theory, the exponential factor in eq. (2.4) is unity, as $E_A^{\mathrm{eff}}(t) = E_B^{\mathrm{eff}}(t)$.

Many variations of the ratio are possible, e.g. replacing $E_A^{\mathrm{eff}}(t) \to E_A^{\mathrm{eff}}(t_1)$. The ratio has a quantum mechanical representation (based on the transfer matrix of the lattice theory)[3]

$$\mathcal{R}(t_2, t_1) = \mathcal{M}_{11} + c_1 \exp(-(E_2^{(A)} - E_1^{(A)})t_2) + c_2 \exp(-(E_2^{(B)} - E_1^{(B)})t_1) + \dots \,. \tag{2.5}$$

These correction terms have already been mentioned in the introduction. Note that replacing $\mathcal{O}_j^{(A)} \to \sum_k \alpha_k^{(A)} \mathcal{O}_k^{(A)}$ and $\mathcal{O}_i^{(B)} \to \sum_k \alpha_k^{(B)} \mathcal{O}_k^{(B)}$ with a specific choice of fixed coefficients α does not change anything in this formula except for modifying the prefactors c_1, c_2. Instead, in the following section we turn to the use of the GEVP in order to change the exponential rates of the correction terms.

---

[3]For simplicity, we everywhere neglect terms which decay exponentially with the time extent of the lattice.

2.2 Summed ratios

An improved asymptotic convergence is provided by the effective matrix element

$$\mathcal{M}_{11}^{\text{summed}}(t) = -\partial_t a \sum_{t_1} \mathcal{R}(t - t_1, t_1) = \mathcal{M}_{11} + \mathrm{O}(t\Delta\,\mathrm{e}^{-t\Delta}), \tag{2.6}$$

$$\Delta = \mathrm{Min}(E_2^{(A)} - E_1^{(A)}, E_2^{(B)} - E_1^{(B)}). \tag{2.7}$$

Eq. (2.7) can be seen by explicit summation over t_1 of the transfer matrix representation of eq. (2.4) and it is the $N = 1$ case of eq. (2.15) (taking the limit $t_0 \to t$). For the degenerate case $E_n^{(A)} = E_n^{(B)}$ it has been used long ago [1, 22] and its improved convergence rate has recently been emphasized in refs. [2, 3]. In ref. [3] the generalization to non-degenerate spectra was introduced.

2.3 GEVP improvement

We here summarize ref. [21] and apply it to the present case. We assume that we have N linearly independent fields $\mathcal{O}_j$, with couplings to the low lying states. The labels A, B are dropped where statements independent of the channel are made. The GEVP [23] ($[C\,v_n]_i = \sum_{j=1}^{N} C_{ij}\,[v_n]_j$),

$$C(t)\,v_n(t, t_0) = \lambda_n(t, t_0)C(t_0)v_n(t, t_0), \tag{2.8}$$

constructed from the matrices $C^{(A)}, C^{(B)}$ at times $t > t_0$ yields effective energies [24]

$$E_n^{\text{eff}}(t, t_0) = -\partial_t \log(\lambda_n(t, t_0)) \tag{2.9}$$

which converge as [21]

$$E_n^{\text{eff}}(t, t_0) = E_n + \mathrm{O}(\exp(-\Delta_{N+1,n}\,t)), \ \Delta_{N+1,n} = E_{N+1} - E_n \tag{2.10}$$

provided one takes $t_0 \geq t/2$, which we use here.[4]

The starting point for computing matrix elements is an operator (in each channel) which satisfies [21]

$$\hat{\mathcal{A}}_n^{\text{eff}}(t)|0\rangle = |n\rangle + \mathrm{O}(\exp(-\Delta_{N+1,n}\,t)). \tag{2.11}$$

With the definitions

$$v_n(t) \equiv v_n(t + a, t), \quad (u, w) = \sum_{i=1}^{N} u_i^* w_i \tag{2.12}$$

$$R_n(t) = (v_n(t),\, C(t)\,v_n(t))^{-1/2} \exp(E_n^{\text{eff}}(t + a, t)\,t/2), \tag{2.13}$$

the explicit construction of $\hat{\mathcal{A}}_n^{\text{eff}}(t)$ is given by

$$[\hat{\mathcal{A}}_n^{\text{eff}}(t)]^\dagger = \mathrm{e}^{-\hat{H}t}\,R_n(t)\,(v_n(t),\,\widehat{\mathcal{O}}^\dagger). \tag{2.14}$$

---

[4] For fixed t_0 one has $E_n^{\text{eff}}(t, t_0) = E_n + \mathrm{O}(\exp(\min(\Delta_{n+1,n}, \Delta_{n,n-1})t))$ instead [24].

With respect to [21] we have here made a specific choice for the relation of t_0 and t, denoting the resulting v_n as $v_n(t)$ with a single argument.

We can then obtain the desired matrix element

$$\mathcal{M}_{mn} = \mathcal{M}_{mn}^{\mathrm{eff}}(t_2, t_1) + \mathrm{O}(\exp(-\Delta_{N_A+1,m}^{(A)} t_1), \exp(-\Delta_{N_B+1,n}^{(B)} t_2)) \qquad (2.15)$$

from

$$\begin{aligned} \mathcal{M}_{mn}^{\mathrm{eff}}(t_2, t_1) &= \langle 0|[\hat{\mathcal{A}}_n^{\mathrm{eff}}(t_2)]^{(A)} \, \hat{h}_{\mathrm{w}} \, [[\hat{\mathcal{A}}_n^{\mathrm{eff}}(t_1)]^{(B)}]^\dagger|0\rangle \\ &= (v_m^{(A)}(t_2), C^{(3)}(t_2, t_1) v_n^{(B)}(t_1)) \, R_m^{(A)}(t_2) \, R_n^{(B)}(t_1) \,. \end{aligned} \qquad (2.16)$$

Here we have reintroduced the labels A, B. Eq. (2.16) reduces to eq. (2.4) for $N_A = 1 = N_B$, but taking N_A, N_B larger improves the convergence and enables access to excited states.

As before, one can formulate a simpler effective matrix element when (A) and (B) are related by a symmetry and only the $m = n$ matrix elements are required. The symmetry means

$$E_n^{\mathrm{eff},B}(t, t_0) = E_n^{\mathrm{eff},A}(t, t_0), \quad v_n^{(A)}(t, t_0) = v_n^{(B)}(t, t_0) \qquad (2.17)$$

for all t, t_0 and n. The ratio (remember that we use the shorthand $v_n^{(A)}(t) = v_n^{(A)}(t + a, t)$)

$$\mathcal{M}_{nn}^{\mathrm{eff}'}(t_2, t_1) = \frac{(v_n^{(A)}(t_2), C^{(3)}(t_2, t_1) v_n^{(A)}(t_1))}{(v_n^{(A)}(t_2), C^{(A)}(t_2 + t_1) v_n^{(A)}(t_1))} \qquad (2.18)$$

satisfies eq. (2.15) as well but may have reduced statistical errors. The leading error is minimized by the choice $t_2 = t_1$.

3 Improved method: sGEVP

Here we combine the improvement by summation of section 2.2 with the GEVP of section 2.3.

3.1 Symmetric case

We consider the symmetric case eq. (2.17) and drop the labels A and B. As derived in appendix A, the effective matrix element

$$\mathcal{M}_{nn}^{\mathrm{eff,s}}(t, t_0) = -\partial_t \left\{ \frac{|(u_n, [K(t)[\lambda_n(t, t_0)]^{-1} - K(t_0)]u_n)|}{(u_n, C^{(A)}(t_0)u_n)} \right\}, \qquad (3.1)$$

$$K_{ij}(t) \equiv a \sum_{t_1} C_{ij}^{(3)}(t - t_1, t_1), \quad u_n \equiv v_n(t, t_0) \qquad (3.2)$$

converges to the exact matrix element as

$$\mathcal{M}_{nn}^{\mathrm{eff,s}}(t, t_0) = \mathcal{M}_{nn} + \mathrm{O}(\Delta_{N+1,n} \, t \exp(-\Delta_{N+1,n} \, t)) \,. \qquad (3.3)$$

The formula assumes $t_0 \geq t/2$ and the exact size of the corrections does in general depend on how we choose t_0, e.g. $t_0 = t - a$ vs. $t_0/t =$ fixed. We shall demonstrate in section 4 that the corrections in eq. (3.1) are very small generically. The label "s" stands for summed, since $K_{ij}(t)$ is a 3-point function summed over one argument.

3.2 Asymmetric case

In the situation when eq. (2.17) is not satisfied or if we want a matrix element $\mathcal{M}_{mn}$ with $n \neq m$, we first define the estimator for the difference $E_n^{(B)} - E_m^{(A)}$,

$$\Sigma(t,t_0) = E_n^{\text{eff},B}(t,t_0) - E_m^{\text{eff},A}(t,t_0) \overset{t\to\infty}{\sim} E_n^{(B)} - E_m^{(A)} \tag{3.4}$$

as well as the energy shifted correlation function

$$D_{ij}(t,t_0) = e^{-t\Sigma(t,t_0)} C_{ij}^{(A)}(t) . \tag{3.5}$$

The summed three-point function is then defined by

$$K_{ij}(t,t_0) = a \sum_{t_1} e^{-(t-t_1)\Sigma(t,t_0)} C_{ij}^{(3)}(t-t_1,t_1) . \tag{3.6}$$

Everywhere we take $t_0 \geq t/2$. An approximation to the matrix element is

$$\mathcal{M}_{mn}^{\text{eff},s}(t,t_0) = -\partial_t \left\{ \frac{|(u_m(t,t_0), [K(t,t_0)[\lambda_n^{(B)}(t,t_0)]^{-1} - K(t_0,t_0)]w_n(t,t_0))|}{[(u_m(t,t_0), D^{(A)}(t_0)u_m(t,t_0))(w_n(t,t_0), C^{(B)}(t_0)w_n(t,t_0))]^{1/2}} \right\} , \tag{3.7}$$

with

$$D(t)\, u_m(t,t_0) = \tilde{\lambda}_m(t,t_0) D(t_0)u_m(t,t_0) , \tag{3.8}$$

$$C^{(B)}(t)\, w_n(t,t_0) = \lambda_n^{(B)}(t,t_0) C^{(B)}(t_0)w_n(t,t_0) . \tag{3.9}$$

We have observed numerically that in the case of $\Sigma(t,t_0) \neq 0$, it converges as

$$\mathcal{M}_{mn}^{\text{eff},s}(t,t_0) = \mathcal{M}_{mn} + \mathrm{O}(\Delta\, t \exp(-\Delta\, t_0)) , \tag{3.10}$$

see section 4. The gap Δ is given by the minimum one in the two channels,

$$\Delta = \mathrm{Min}\left(E_{N_A+1}^{(A)} - E_m^{(A)} ,\ E_{N_B+1}^{(B)} - E_n^{(B)} \right) . \tag{3.11}$$

Since the exponential convergence is now governed by t_0, there is no obvious advantage compared to eq. (2.16) unless one takes $t_0 \approx t$. If statistical precision is good enough to allow for such large t_0, 3-point functions and 2-point functions with a maximal time extent of t are sufficient to obtain a convergence rate of $\mathrm{O}(\Delta\, t \exp(-\Delta\, t))$ as in the symmetric case. Eq. (3.10) has not been proved formally, but our numerical investigation of toy models leaves little doubt that it is correct.

4 Demonstrations

We carry out two sets of demonstrations of how the various estimators for matrix elements work. First we consider toy models, prescribing spectra and matrix elements and do not take statistical errors into account. We construct "difficult" (large corrections due to excited states) and "easy" toy models. The second set of experiments is a quenched computation of the $B^*B\pi$-coupling, where realistic statistical errors are present.

4.1 Models

4.1.1 Definition of the models

We first specify the spectra in dimensionless form. Two different ones are used below,

$$r_0 E_n^{(l)} = n \,, \quad r_0 E_n^{(h)} = 1.1 \times n \,. \tag{4.1}$$

The length factor r_0 is in principle arbitrary, setting the overall scale of the theory, but we think of it as $r_0 \approx 0.5 \mathrm{fm}$. Level splittings of around $1/r_0$ are realistic in QCD, as the particle data book and lattice computations show.

Next the overlaps

$$\psi_{\mathrm{in}} = \langle 0|\mathcal{O}_i|n\rangle \tag{4.2}$$

need to be fixed. In our HQET applications (see section 4.2), we use spatially smeared quark fields to construct the fields $\mathcal{O}_i$. We computed their overlaps ψ_{in} for $n = 1, \ldots, 5$ and $i = 1, \ldots 7$ using the GEVP "creation operator" $[\hat{\mathcal{A}}_n^{\mathrm{eff}}(t)]^\dagger$. For details we refer to the following section. Here we just take the approximate matrix

$$\psi^{\mathrm{S}} = \begin{pmatrix} 0.92 & 0.03 & -0.10 & -0.01 & -0.02 \\ 0.84 & 0.40 & 0.03 & -0.06 & 0.00 \\ 0.56 & 0.56 & 0.47 & 0.26 & 0.04 \end{pmatrix} \tag{4.3}$$

corresponding to smearing levels $1, 4, 7$ which is typically done in practice [25]. We observed a strong decay of the overlaps $\psi_{\mathrm{in}}^{\mathrm{S}}$ with increasing n which suggests that a truncation with $\psi_{\mathrm{in}} = 0$ for $n > 5$ is realistic at reasonable time separations of the correlation functions, say $t > r_0/2$. In any case, what we discuss here remains a model, but we expect it to be quite realistic.

The matrix ψ^{S} represents a relatively comfortable situation which we may not always have. For that reason we also construct a more challenging case

$$\psi_{\mathrm{in}}^{\mathrm{C}}|_{n \leq 3} = \begin{pmatrix} 0.9 & 0.1 & -0.1 \\ 0.8 & 0.4 & 0.2 \\ 0.6 & 0.6 & 0.5 \end{pmatrix} \qquad \psi_{\mathrm{in}}^{\mathrm{C}}|_{4 \leq n \leq 20} = \begin{pmatrix} -1/(3n^2) \\ 2\,n^{-2} - (2n)^{-3/2} \\ 1/(n-1) \end{pmatrix} \tag{4.4}$$

with a slow decay in n. We set $\psi_{\mathrm{in}} = 0$ for $n > 20$.

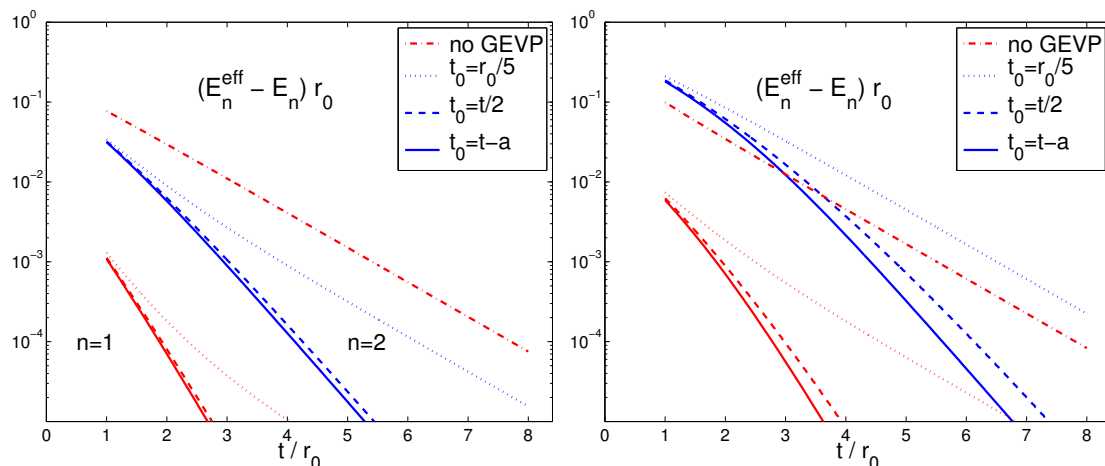With the model matrix elements (again we note that these are not completely unrealistic)

$$\mathcal{M}_{nn} = 0.7 \frac{6}{n+5} \,, \qquad \mathcal{M}_{n,n+m} = \frac{M_{nn}}{3m} \text{ for } m > 0 \,, \tag{4.5}$$

and assuming the sectors A, B to be related by a symmetry,

$$C_{ij}^{(A)}(t) = \sum_n \psi_{\mathrm{in}}^{\mathrm{S}} (\psi^{\mathrm{S}})_{jn}^* \mathrm{e}^{-E_n^{(l)} t} \ = \ C_{ij}^{(B)}(t) \,, \tag{4.6}$$

the model is completely defined. In particular we have

$$C_{ij}^{(3)}(t_2, t_1) = \sum_{n,m} \psi_{\mathrm{in}}^{\mathrm{S}} \mathrm{e}^{-E_n^{(l)} t_2} \mathcal{M}_{nm} \mathrm{e}^{-E_m^{(l)} t_1} (\psi^{\mathrm{S}})_{jm}^* \,. \tag{4.7}$$

Figure 1. Corrections of eq. (2.9). The ground state $n = 1$ is plotted in red and has the small corrections. The first excitation (blue) is above. Dotted lines are for $t_0 = r_0/5$, while dashed lines are $t_0 = t/2$ and full lines $t_0 = t - a$. Shown on the left is the model Sl and on the right Cl. The dashed-dotted lines show the corrections of the standard effective mass of the correlator C_{22} which approaches the ground state energy.

We refer to this model as SlSl. Replacing ψ^S by ψ^C defines the model ClCl and finally with $\psi^S, E_n^{(l)}$ for channel A and $\psi^C, E_n^{(h)}$ for channel B we define the model SlCh. In other words we have the following table.

model	$\psi^{(A)}$	$\psi^{(B)}$	$E_n^{(A)}$	$E_n^{(B)}$
SlSl	ψ^S	ψ^S	$E_n^{(l)}$	$E_n^{(l)}$
ClCl	ψ^C	ψ^C	$E_n^{(l)}$	$E_n^{(l)}$
SlCh	ψ^S	ψ^C	$E_n^{(l)}$	$E_n^{(h)}$

4.1.2 Energies from the GEVP

The corrections of the effective energies extracted from the GEVP, eq. (2.9), compared to the exact energies is shown in Fig. 1. We see how $t_0 \geq t/2$ accelerates the convergence. As expected Cl is a more challenging situation with larger corrections. One also sees that at short time ($t/r_0 \leq 2$) the dependence on t_0 is typically not very dramatic. This feature has been observed in a number of practical applications [25, 26]. Still, it appears dangerous to rely on this in general. In the left hand plot, we also observe the difference of the GEVP and a standard effective mass (dashed-dotted line). Here C_{22} is shown (the corrections for C_{11} are quite a bit smaller).

4.1.3 Matrix elements

Let us start with the easiest situation, the extraction of ground state matrix elements $m = n = 1$ in the symmetric case. These are shown for two models in Fig. 2. The labeling
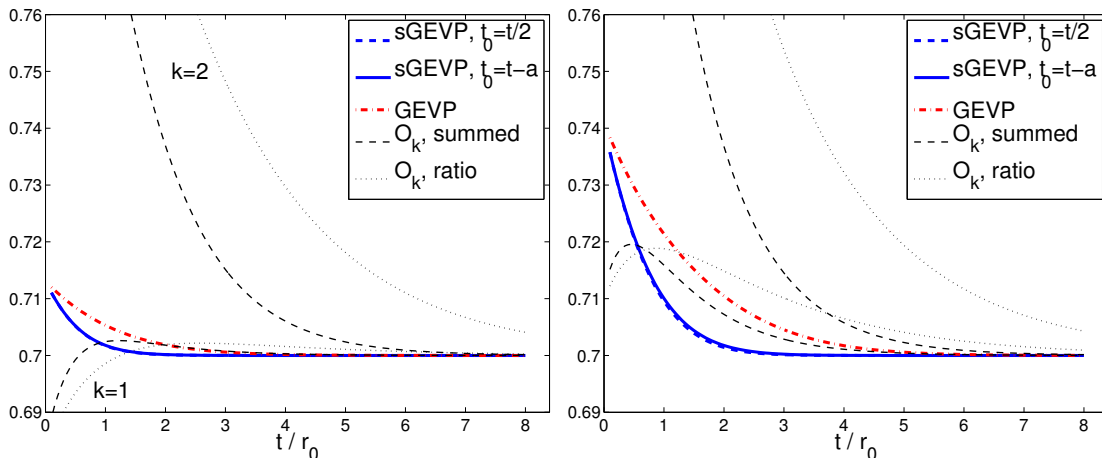
Figure 2. Effective ground state matrix elements, model SlSl on the left and model ClCl on the right, both shown as a function of the total time-separation of the 3-point function. On the left side, the sGEVP estimates for $t_0 = t/2$ and $t_0 = t - a$ can't be distinguished in the figure. For the non-GEVP cases we show two different interpolating fields, O_k, $k = 1, 2$.

of the different estimates is as follows:

"ratio"	dotted, black	eq. (2.4) with $t_1 = t_2 = t/2$
"summed"	dashed, black	eq. (2.6)
"GEVP"	dashed-dotted, red	eq. (2.18) with $t_1 = t_2 = t/2$
"sGEVP"	blue	eq. (3.1)

The scale of the y-axis covers a variation of 10%. On the x-axis in this and the following figures we have for each method considered the total time extent of the 3-point functions since in a MC computation this generically governs the statistical accuracy. The graphs illustrate that the improved *asymptotics* of the sGEVP estimate (compared to the GEVP and the single operators) ($N = 1$) go hand in hand with smaller corrections at *moderate* time separations, $t \approx r_0 \ldots 2r_0$.[5] Among the estimates which do not use a GEVP, the summed method is generically better, at least when t is not too small.

Diagonal ($m = n$) matrix elements for the degenerate case are shown in Fig. 3 on the left. For $n > 1$ only GEVP and sGEVP can be used for a systematic computation. Even though the scale of the y-axis is enlarged, we observe that sGEVP also works rather well for determining excited states. Note that with a GEVP with $N = 3$ states (as is used here), the convergence of the $m = n = 3$ matrix elements is rather slow, but we show them anyway for illustration. It is strongly recommended to use a larger N in a real computation of $\mathcal{M}_{33}$ if statistical errors allow.

On the right of Fig. 3, we show the matrix elements $\mathcal{M}_{12}$. Here the sGEVP means eq. (3.7) with the energy shifts. The improvement compared to the standard application of the GEVP, eq. (2.18), is present but is not as impressive as on the left side, where no

---

[5]Recall that the gaps of the models are $\Delta_{n+1,n} = 1/r_0$.
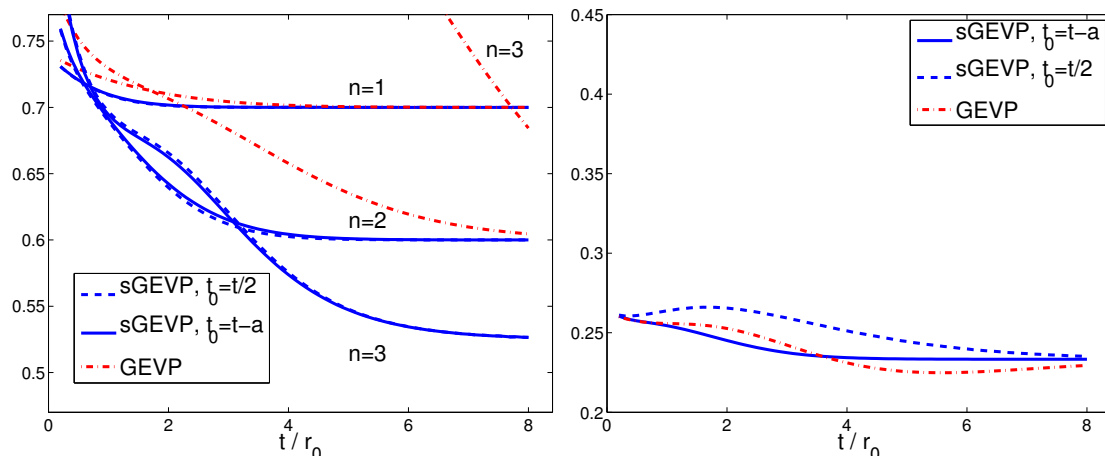
Figure 3. Effective matrix elements in the model ClCl. Left: $\mathcal{M}_{11}, \mathcal{M}_{22}, \mathcal{M}_{33}$, from top to bottom; right: $\mathcal{M}_{12}$.
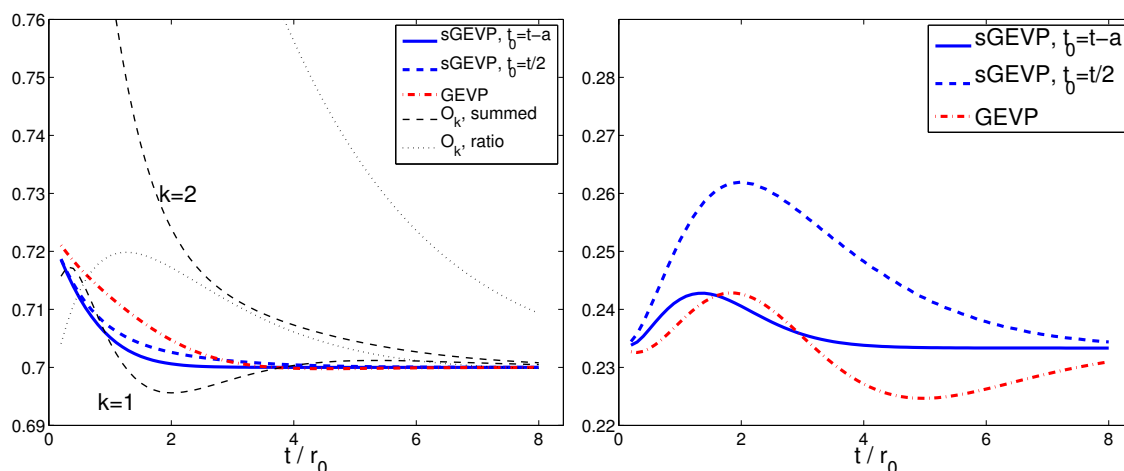
Figure 4. Effective ground state matrix element $\mathcal{M}_{11}$ (left) and $\mathcal{M}_{12}$ in model SlCh.

energy shifts are needed. We do not show levels above $n = 2$ since there a larger GEVP would be recommended as we discussed for the diagonal case.

Finally, consider the situation where the spectra of the A sector and the B sector are different, as in the model SlCh. Example applications are $B \to \pi$ transitions or elastic form factors with momentum transfer. On the left side of Fig. 4 the $\mathcal{M}_{11}$ matrix element is shown. We again observe an impressive advantage of the GEVP methods, in particular of sGEVP over the standard ratio eq. (2.4). On the right side of the figure we study $\mathcal{M}_{12}$, where eq. (2.4) is not applicable. In this particular case, the amplitudes of the corrections of the GEVP effective matrix elements are relatively small and interfere destructively. It therefore happens to be more accurate than sGEVP for a range of t.

In conclusion, the study of the models shows that the asymptotic convergence formulae also provide a very good estimate of the relative advantages of the different methods at intermediate t. In particular, consider first the degenerate case. The comparison of the

asymptotic behavior, $\mathcal{M}_{nn}^{\mathrm{eff,s}}(t,t_0) - \mathcal{M}_{nn} \sim \mathrm{O}(t\Delta \exp(-t\Delta))$ vs. $\mathcal{M}_{nn}^{\mathrm{eff}}(t_0,t_0) - \mathcal{M}_{nn} \sim \mathrm{O}(t\Delta \exp(-t_0\Delta))$ suggests that $t_0 \approx t$ is needed to reach similar accuracy in the two cases and indeed we find this generically to be the case. One then needs the 3-point functions at twice the total time separation in GEVP compared to sGEVP. In the non-degenerate case, the convergence is governed by t_0 in both GEVP and sGEVP. Here a very significant improvement is the change from a standard ratio eq. (2.4) to GEVP or sGEVP, see the left of Fig. 4. The right side of that figure shows that sGEVP yields considerable further improvement over GEVP when a large t_0 is chosen.

4.2 The $B^*B\pi$-coupling in the quenched approximation

In the static approximation for the b-quark, the $B^*B\pi$-coupling is denoted by $\hat{g}$. It is a leading order low energy constant in the heavy meson chiral Lagrangian [27–29] . As such, it is of considerable interest for chiral extrapolations of lattice results, employing a systematic expansion in $1/m_{\mathrm{b}}$ and $m_\pi^2/(8\pi^2 F_\pi^2)$. The bare matrix element is[6]

$$\hat{g} = \tfrac{1}{2}\langle B^0(\mathbf{0})|A_k(0)|B_k^{*+}(\mathbf{0})\rangle\,, \qquad A_\mu(x) = \overline{\psi}_{\mathrm{d}}(x)\gamma_\mu\gamma_5\psi_{\mathrm{u}}(x) \tag{4.8}$$

with $|B_k^*(0)\rangle$ polarized along the k-axis, see also [2, 30]. Note that here we use the normalization of states $\langle B(\mathbf{p})|B(\mathbf{p})\rangle = \langle B_k^*(\mathbf{p})|B_k^*(\mathbf{p})\rangle = 2L^3$, which corresponds to the non-relativistic one in the infinite volume limit. We do not include the renormalization factor of the axial current anywhere.

Our interpolating fields for B and B^* are related by the exact spin symmetry of the static approximation and are generated by gauge-covariant Gaussian wave functions inserted between the static and the light quark field. Such gauge invariant interpolating fields were introduced in ref. [22]. We use exactly the ones of ref. [2] with width $r_{\mathrm{wf}}/r_0 = 0.36, 0.51, 0.62, 0.71, 0.87, 1.01, 1.13$ (eq. (2.5) of ref. [2]). For the present demonstration we work in the quenched approximation and the light quark mass is set to the mass of the strange as in [25]. An ensemble of one hundred gauge configurations is used on a 32×16^3 lattice with spacing $a \approx 0.1$fm and statistical errors are kept small by an all-to-all method [19] in combination with the static action "HYP2" [7] as done previously [2]. Here we use one hundred fully time-diluted noise sources per configuration.

4.2.1 Approximate overlaps

We first pick five fields $\mathcal{O}_i$ from our set with $r_{\mathrm{wf}}/r_0 = 0.36, 0.51, 0.62, 0.71, 1.13$. With the operator eq. (2.14) we can then compute the overlaps
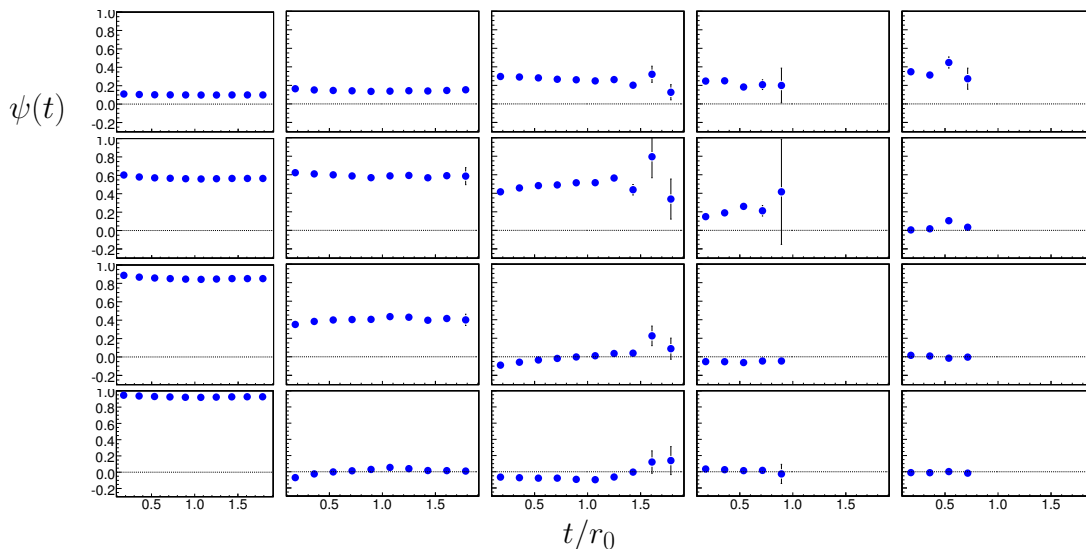
$$\psi_{\mathrm{in}} = \langle 0|\mathcal{O}_i|n\rangle = \psi_{\mathrm{in}}(t) + \mathrm{O}(\exp(-(E_{N+1} - E_n)t)) \tag{4.9}$$

$$\psi_{\mathrm{in}}(t) = \langle 0|\widehat{\mathcal{O}}_i\,\hat{\mathcal{A}}_n^{\mathrm{eff}}(t)|0\rangle = \sum_j C_{ij}(t)[v_n]_j(t)R_n(t) \tag{4.10}$$

where $n = 1,\ldots,5$ labels the excitations. The normalization of the fields $\mathcal{O}_i$ is irrelevant for all applications, but in order to have the interpretation of an overlap, we choose the

---

[6]We thank Fabio Bernardoni for discussions on the effective theory and a check of the normalization of $\hat{g}$.

Figure 5. Overlaps for the static B-meson interpolating fields. Column n refers to state n, row 1 to the local (time component of the) axial current. Rows 2-4 correspond to the three interpolating fields used in our GEVP computation of $\hat{g}$ with radii $r_{\rm wf}/r_0 = 0.36\,, 0.62\,, 1.13$ respectively.
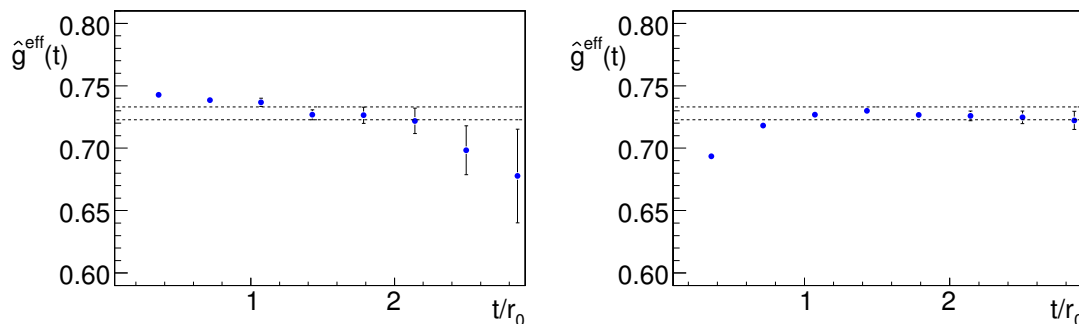

Figure 6. The matrix element $\hat{g}$ as a function of t/r_0. Left: sGEVP estimate eq. (3.1) with $t_0 = t/2$, right: GEVP estimate with $t_0 = t/2$. The error band is our best estimate determined previously with very high statistics [2].

normalization such that $C_{ii}(0) = 1$. In this case a value of one for $\psi_{\rm in}^2$ means that $\widehat{\mathcal{O}}_i|0\rangle = |n\rangle$ without corrections, i.e. 100% overlap. Furthermore, we fix the signs by the convention $\psi_{\rm in} > 0$ for the value i which maximizes $|\psi_{\rm in}|$ at fixed n.

Figure 5 shows examples of $\psi_{\rm in}(t)$. Even if these are not precision determinations of the overlaps, they show interesting features. The local field shown in the first row has considerable overlap with all states considered. It is a bad interpolating field for ground state physics. However the other fields with reasonable radii display a rather strong decay of the overlaps with growing n, indicating that the smeared fields provide a good basis of interpolating fields which couple little to excited states. Indeed, this figure demonstrates that these wave functions considerably reduce the overlaps to high excited states. Conversely, this also means that high excited states are difficult to access with these fields.
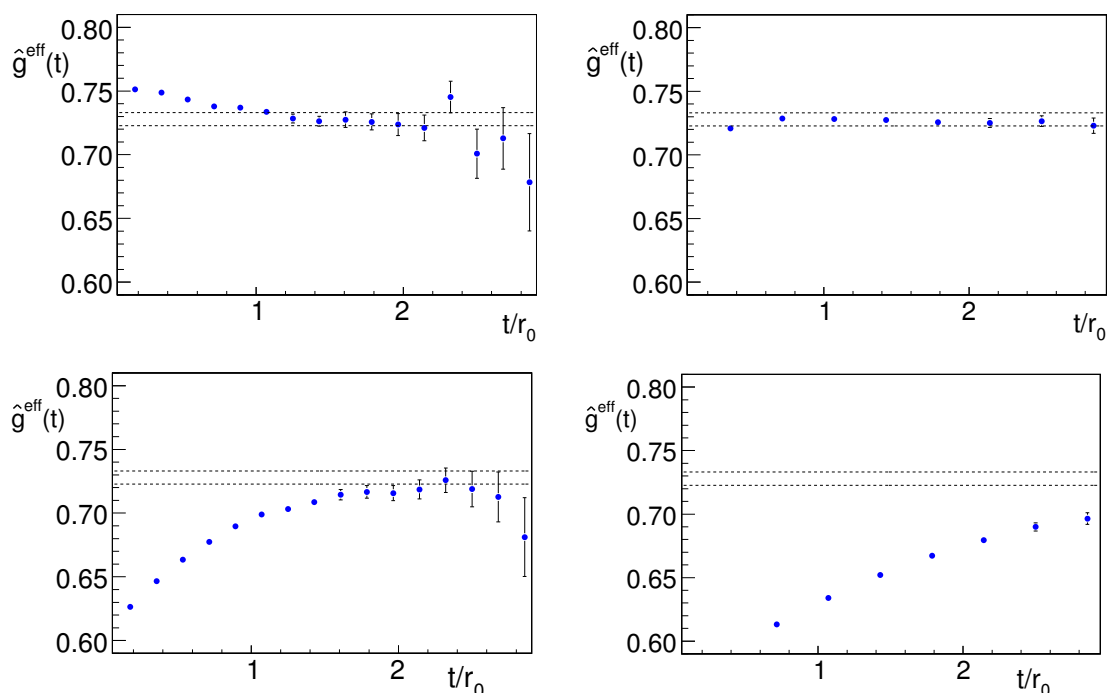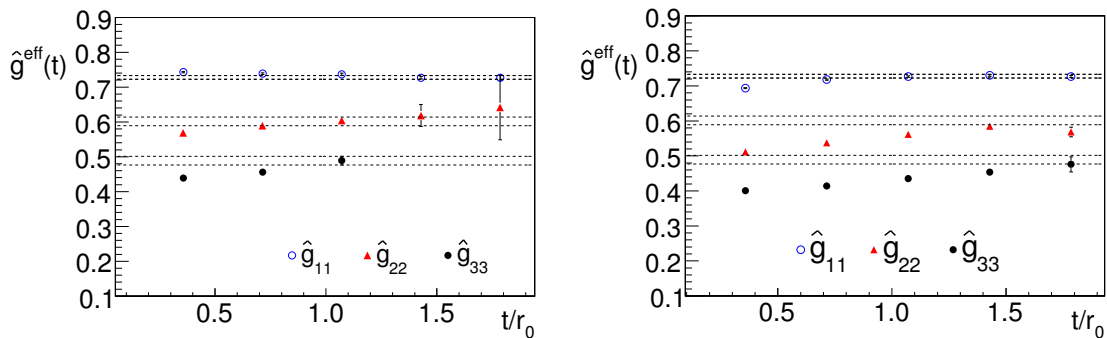
Figure 7. Estimates of $\hat{g}$ as a function of t/r_0. Left: summed ratio estimate eq. (2.6), right: ratio eq. (2.4). On the top the interpolating field with the biggest overlap with the ground state is shown ($r_{\mathrm{wf}} = 1.13\,r_0$). The bottom two figures are for $r_{\mathrm{wf}} = 0.62\,r_0$.
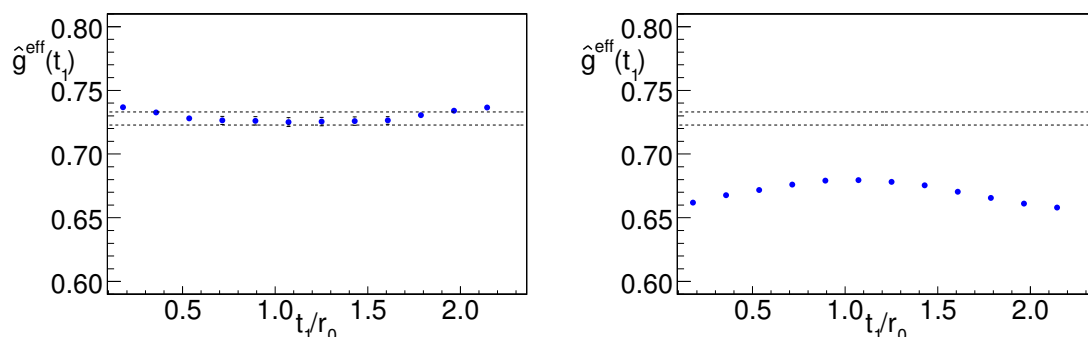
Figure 8. Standard plateau plot for $\hat{g}$. The ratio $\mathcal{R}(t - t_1, t_1)$ is considered as a function of t_1/r_0 with t fixed at $t = 2.14 r_0$. The latter is the value used in the so far most complete determination [31], while earlier $t/r_0 \approx 4$ was used on a lattice with $a = 0.2\,\mathrm{fm}$ [32]. Left: wave-function with $r_{\mathrm{wf}} = 1.13\,r_0$, right: $r_{\mathrm{wf}} = 0.62\,r_0$.

Reading off approximate plateau values, we extract the model ψ^{S} for section 4.1. This model yields a qualitative understanding of the corrections. We believe that computing overlaps as done here may also be very useful for understanding the systematic errors in present extractions of nucleon matrix elements, namely the question of the magnitude of excited state contamination. Given an approximate knowledge of the spectrum, this contamination can be roughly estimated when the overlaps are known. Indeed, let us

Figure 9. The diagonal matrix elements $\hat{g}_{nn}$ as a function of t/r_0. Left: sGEVP estimate eq. (3.1) with $t_0 = t/2$, right: GEVP estimate with $t_0 = t/2$. An $N = 3$ GEVP is used. The matrix elements are seen to be ordered $\hat{g}_{n+1,n+1} < \hat{g}_{nn}$.

apply our approximate knowledge of $\psi_{1n} < 0.1$ for $n > 1$, together with the plausible assumption that matrix elements $\mathcal{M}_{mn}$ are of roughly the same magnitude as $\mathcal{M}_{11}$. Then, at time separations $t = r_0$ and for the best wave function, excited states make rather small corrections of order $0.1\mathrm{e}^{-1}$, i.e. of the order of a few per cent. Therefore the matrix element $\hat{g}$ is a rather easy test case and all methods should be successful.

4.2.2 The matrix element $\hat{g}$

In this section we show numerical results for $\hat{g}$, computed with the various methods introduced above. The GEVP estimates for the ground state, displayed in Fig. 6, exhibit no corrections exceeding 2% once $t = r_0$ has been reached. For smaller t, sGEVP has smaller corrections than GEVP. However, at large times the statistical errors are increasing faster for sGEVP.

For this particular matrix element and for the best interpolating field, the corrections for the summed ratio (Fig. 7, top) are somewhat larger than those of the standard ratio and again the summed method suffers from larger statistical errors at large times. However, in the case of a less optimal interpolating field (bottom of Fig. 7), the summed ratio exhibits its superiority.

For comparison, we also show the frequently used analysis where t is kept fixed (here at a value used previously in determinations of $\hat{g}$ [31]) and one looks for a plateau as a function of t_1. With our precision one can observe the lack of a plateau in Fig. 8 for $r_{\mathrm{wf}} = 0.62\,r_0$, but with errors at a 1% level a false "plateau" would be observed for $t/4 \leq t_1 \leq 3t_1/4$. This demonstrates the danger inherent in this method. The left hand side of the figure shows that for $\hat{g}$ a plateau *with the correct height* is obtained for a larger smearing radius.

4.2.3 Excited state matrix elements $\hat{g}_{nm}$

For excited state matrix elements, Fig. 9, only the GEVP estimates are applicable. They appear to work quite well for the first excitation and also for the second excitation a reasonable estimate can be obtained. The sGEVP again seems superior, as the deviations
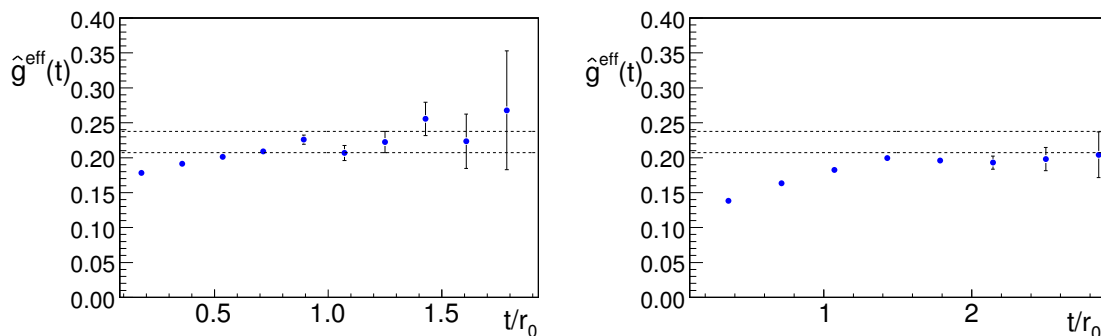
Figure 10. The off-diagonal matrix element $\hat{g}_{12}$ as a function of t/r_0. Left: sGEVP estimate eq. (3.1) with $t_0 = t - a$, right: GEVP estimate with $t_0 = t/2 - a$. An $N = 3$ GEVP is used.

from our estimated asymptotic values are smaller. Figure 10 demonstrates these same features for an off-diagonal matrix element.

5 Conclusions

In this paper we have introduced the GEVP method with summation, denoted sGEVP, and we have examined several alternative methods for computing hadron-to-hadron matrix elements. They have rather different asymptotic corrections due to excited states:

ratio	eq. (2.4)	:	$\exp(-\Delta_{2,1}t/2)$	(just ground state)
summed ratio	eq. (2.6)	:	$t\Delta_{2,1}\exp(-t\Delta_{2,1})$	(just ground state)
GEVP	eq. (2.16)	:	$\exp(-\Delta_{N+1,1}t/2)$	
sGEVP	eq. (3.1)	:	$t\Delta_{N+1,1}\exp(-t\Delta_{N+1,1})$	(equal energy case)
sGEVP	eq. (3.7)	:	$t\Delta\exp(-t_0\Delta)$	(general case)

In the last case, Δ is given by eq. (3.11) and one will typically use $t_0 = t/2$. The form of the leading correction term of sGEVP is derived in the appendix for the equal energy case, while for the general one we deduced it from the numerical investigation of toy models. The GEVP correction term is known from [21] and for "ratio" and "summed ratio" it follows directly from the transfer matrix representation.

We investigated two toy models constructed to be quite representative for heavy-light meson matrix elements. In these models, the *asymptotic* forms of the corrections have been found to be a good guideline for the behavior at *intermediate* values of t, of the order $t = (2-3)r_0$. In particular we found that generically sGEVP has the smallest systematic errors, followed by GEVP. As a rule of thumb, sGEVP requires half the time separation of GEVP for the same systematic accuracy.

A Monte Carlo computation of the $B^*B\pi$ coupling $\hat{g}$ confirms our findings in the models concerning the systematic errors. In addition it allows us to make statements about the statistical errors which have to be balanced with systematic ones due to excited states. Statistical errors grow more quickly as a function of t for sGEVP compared to GEVP, but a comparison at roughly the same amount of excited state contamination corresponds to a

m	n	$\mathcal{M}_{mn}^{\mathrm{eff}}(t,t)$	$\mathcal{M}_{mn}^{\mathrm{eff,s}}(t,t/2)$	$\mathcal{M}_{mn}^{\mathrm{eff,s}}(t,t-a)$
1	1	0.004	0.003	0.003
2	1	0.010	0.013	0.009
2	2	0.032	0.012	0.013
3	3	0.063	0.012	0.012

Table 1. Statistical errors of various estimators for $\hat{g}_{mn} = \mathcal{M}_{mn}$ for $t \approx r_0$.

factor two between the values of t. A comparison at roughly fixed systematic error is shown in table 1. We observe a minor difference for the ground state in advantage for sGEVP and the ratio of errors grows up to a factor five in the error for $\mathcal{M}_{33}$ for the considered matrix element.

In the comparison of the different methods, one also has to consider the numerical effort to compute the effective matrix elements. We assume that one wants to control the corrections by computing the t-dependence of the estimators. In the summed cases, eq. (2.6) and eq. (3.1), this can often be done with a fixed number of quark propagator computations yielding a result for all t, by computing sequential propagators. The computation of $\hat{g}$ is such a case. In fact, since we have used a full all-to-all computation with "time dilution" (in the notation of [19]), also the GEVP estimate is obtained at the same expense. In contrast, if one only uses translation invariance on a time slice ("time-slice-to-all"), and for example varies t, keeping $t_1 = t_2 = t/2$ in eq. (2.4) or eq. (2.18), then the required number of propagator computations is proportional to the number of t-values considered. In this situation the sGEVP method has an additional advantage.

Taking statistical and excited state errors as well as the effort into account, sGEVP seems to be the overall most accurate, safe and efficient method. Given the difficulty in evaluating relevant correlation functions at large time separations and assessing the systematic errors, it still appears advisable to compare the different approaches in most cases.

In our opinion the sGEVP method (and maybe the GEVP method) should be applied to nucleon matrix elements such as g_A or moments of structure functions, where large time separations are difficult to reach [3, 8–10] and it is non-trivial to estimate possible contamination by excited states. In order to appreciate the last point, recall that in the standard ratio method, the systematic error drops like $\exp(-\Delta_{2,1}t/2)$. In order to see such a term, one has to change t to t' such that the error term changes appreciably, say by a factor of three. One then needs $t' - t \approx 2/\Delta_{2,1} \approx 1\,\mathrm{fm}$.[7] The summed ratio reduces this requirement by a factor of about two and the GEVP methods by a larger factor. This gains security in the detection of possible systematic errors.

---

[7]We here again assume a gap of around $400\,\mathrm{MeV}$. Close to the chiral limit lower energy states with a gap of $2m_\pi$ exist, but probably have small overlaps with the typically considered interpolating fields.

Acknowledgments

A   Derivation of the sGEVP method

Here we give a derivation of the formulae of section 3.1.

A.1   Linear perturbation of the original theory

We are here interested in the matrix element $\mathcal{M}_{nn}$ assuming the degeneracy of sectors A and B via eq. (2.17) and $\hat{h}_{\mathrm{w}}(\mathbf{x})^{\dagger} = \hat{h}_{\mathrm{w}}(\mathbf{x})$.[8] To arrive at an expression for the matrix element, we augment the original theory with Hamiltonian $\hat{H}$ (defined through the transfer matrix) by adding a perturbation term with strength ϵ,

$$\hat{H}(\epsilon) = \hat{H} + \epsilon\, \hat{h}_{\mathrm{w}}(0)\,. \tag{A.1}$$

The twofold degenerate levels with energy $E_n^{(A)} = E_n^{(B)} \equiv E_n$ are then split to $E_n^{\pm}(\epsilon)$. From standard degenerate perturbation theory one has

$$E_n^{\pm}(\epsilon) = E_n \pm \epsilon\mathcal{M}_{nn} + \mathrm{O}(\epsilon^2) \tag{A.2}$$

with eigenstates $|\pm, n\rangle = [|B, n\rangle \pm |A, n\rangle]/\sqrt{2}$ and

$$\mathcal{M}_{nn} = E_n'(0) \equiv \left. \frac{d}{d\epsilon}E_n^+(\epsilon)\right|_{\epsilon=0}\,. \tag{A.3}$$

A.2   GEVP in the augmented theory

The desired $E_n'(0)$ is efficiently computed with a GEVP method as follows. We combine the interpolating fields $\mathcal{O}_i^{(A)}, \mathcal{O}_j^{(B)}$ from section 2

$$\mathcal{O}_i(t) = \mathcal{O}_i^{(A)}(t)\,,\ i = 1 \ldots N\,,\quad \mathcal{O}_{i+N}(t) = \mathcal{O}_i^{(B)}(t)\,,\ i = 1 \ldots N\,. \tag{A.4}$$

Since A, B correspond to different sectors (e.g. different flavours) we have $\langle A, m|\mathcal{O}_i^{(B)}|0\rangle = 0 = \langle B, n|\mathcal{O}_i^{(A)}|0\rangle$. Expanding the path integral to first order[9] in ϵ one then sees immediately that the combined $2N \times 2N$ matrix correlation function

$$C_{ij}(t, \epsilon) = \langle \mathcal{O}_i(t)\, \mathcal{O}_j^{\dagger}(0)\rangle \tag{A.5}$$

---

[8]The operator $\hat{h}_{\mathrm{w}}(\mathbf{x})$ typically comes from the expansion of the electroweak hamiltonian density in terms of $1/M_W$, but other applications are possible. For example the field $h_{\mathrm{w}}(x)$ representing $\hat{h}_{\mathrm{w}}(\mathbf{x})$ in the path integral may be $h_{\mathrm{w}}(x) = A_k^+(x) + A_k^-(x)$, with $A_k^{\pm}(x) = \overline{\psi}(x)\tau^{\pm}\gamma_k\gamma_5\psi$, with $\tau^{\pm}$ the raising and lowering Pauli matrices in SU(2) flavor space. In this case, the matrix elements sought are the $B^*B\pi$ coupling $\hat{g}$ or the nucleon axial coupling g_{A}.

[9]See for example [21], section 3.2.

has a simple block structure,

$$C(t, \epsilon) = \begin{pmatrix} C^{(A)}(t) & \epsilon K(t) \\ \epsilon K(t)^\dagger & C^{(A)}(t) \end{pmatrix} + \mathrm{O}(\epsilon^2) \qquad (A.6)$$

up to first order in ϵ. The entries $C^{(A)} = C^{(B)}$ were defined in eq. (2.3) and K in eq. (3.2).

The generalized eigenvalues λ_n, eq. (2.8), determine effective energies

$$E_n^{\mathrm{eff}}(t, t_0, \epsilon) = -\partial_t \log(\lambda_n(t, t_0, \epsilon)). \qquad (A.7)$$

In the augmented theory, an extra argument ϵ has been added to λ_n for clarity. Eq. (2.10) describes the corrections by which $E_n(t, t_0, \epsilon)$ differ from the exact energy levels. Differentiating that equation with respect to ϵ yields

$$E_n^{\mathrm{eff}\prime}(t, t_0) \equiv \frac{d}{d\epsilon} E_n^{\mathrm{eff}}(t, t_0, \epsilon)\Big|_{\epsilon=0} = \mathcal{M}_{nn} + \mathrm{O}(\Delta_{N+1,n} t \exp(-\Delta_{N+1,n} t)). \qquad (A.8)$$

It remains to give an explicit expression for $E_n^{\mathrm{eff}\prime}(t, t_0)$ in terms of the correlation functions, which is equivalent to a solution of the GEVP to first order in ϵ. The $2N \times 2N$ GEVP equation, $C(t, \epsilon) v_n(t, t_0, \epsilon) = \lambda_n(t, t_0, \epsilon) C(t_0, \epsilon) v_n(t, t_0, \epsilon)$, separates into the two independent ones

$$[C^{(A)}(t) \pm \epsilon K(t)] u_n^\pm(t, t_0, \epsilon) = \lambda_n^\pm(t, t_0, \epsilon) [C^{(A)}(t_0) \pm \epsilon K(t_0)] u_n^\pm(t, t_0, \epsilon) \qquad (A.9)$$

with $v_n^\pm = \frac{1}{\sqrt{2}} \begin{pmatrix} u_n^\pm \\ \pm u_n^\pm \end{pmatrix}$. The expansion of such a GEVP in ϵ was written down in [21] with the intention that ϵ is given by the HQET expansion parameter. We here just use the solution. Its first order term in ϵ yields the desired matrix element in the form eq. (3.1) in terms of the generalized eigenvectors u_n of the lowest order ($\epsilon = 0$) GEVP of size $N \times N$ in a single channel A.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

[1] L. Maiani, G. Martinelli, M. Paciello and B. Taglienti, *Scalar densities and baryon mass differences in lattice QCD with Wilson fermions*, *Nucl. Phys.* **B 293** (1987) 420 [INSPIRE].

[2] ALPHA collaboration, J. Bulava, M. Donnellan and R. Sommer, *The $B^* B\pi$ coupling in the static limit*, PoS(LATTICE 2010)303 [arXiv:1011.4393] [INSPIRE].

[3] S. Capitani, M. Della Morte, B. Knippschild and H. Wittig, *Systematic errors in extracting nucleon properties from lattice QCD*, PoS(LATTICE 2010)147.

[4] G.P. Lepage, *The analysis of algorithms for lattice field theory*, in *From actions to answers*, T. DeGrand and D. Toussaint eds., World Scientic, Singapore (1989).

[5] M. Lüscher, *Computational strategies in lattice QCD*, arXiv:1002.4232 [INSPIRE].

[6] S. Hashimoto, *Computation of the heavy-light decay constant using nonrelativistic lattice QCD*, *Phys. Rev.* **D 50** (1994) 4639 [hep-lat/9403028] [INSPIRE].

[7] M. Della Morte, A. Shindler and R. Sommer, *On lattice actions for static quarks*, *JHEP* **08** (2005) 051 [hep-lat/0506008] [INSPIRE].

[8] D.B. Renner, *Status and prospects for the calculation of hadron structure from lattice QCD*, PoS(LAT2009)018 [arXiv:1002.0925] [INSPIRE].

[9] C. Alexandrou, *Hadron structure and form factors*, PoS(LATTICE 2010)001 [arXiv:1011.3660] [INSPIRE].

[10] S. Dinter et al., *Precision study of excited state effects in nucleon matrix elements*, *Phys. Lett.* **B 704** (2011) 89 [arXiv:1108.1076] [INSPIRE].

[11] G. Parisi, R. Petronzio and F. Rapuano, *A measurement of the string tension near the continuum limit*, *Phys. Lett.* **B 128** (1983) 418 [INSPIRE].

[12] U. Wolff, *Asymptotic freedom and mass generation in the O(3) nonlinear σ-model*, *Nucl. Phys.* **B 334** (1990) 581 [INSPIRE].

[13] M. Lüscher and P. Weisz, *Locality and exponential error reduction in numerical lattice gauge theory*, *JHEP* **09** (2001) 010 [hep-lat/0108014] [INSPIRE].

[14] M. Della Morte and L. Giusti, *Symmetries and exponential error reduction in Yang-Mills theories on the lattice*, *Comput. Phys. Commun.* **180** (2009) 819 [arXiv:0806.2601] [INSPIRE].

[15] M. Della Morte and L. Giusti, *A novel approach for computing glueball masses and matrix elements in Yang-Mills theories on the lattice*, *JHEP* **05** (2011) 056 [arXiv:1012.2562] [INSPIRE].

[16] U. Wolff, *Strong coupling expansion Monte Carlo*, PoS(LATTICE 2010)020 [arXiv:1009.0657] [INSPIRE].

[17] R. Sommer, *Leptonic decays of B and D mesons*, *Nucl. Phys. Proc. Suppl.* **42** (1995) 186 [hep-lat/9411024] [INSPIRE].

[18] UKQCD collaboration, M. Foster and C. Michael, *Quark mass dependence of hadron masses from lattice QCD*, *Phys. Rev.* **D 59** (1999) 074503 [hep-lat/9810021] [INSPIRE].

[19] J. Foley et al., *Practical all-to-all propagators for lattice QCD*, *Comput. Phys. Commun.* **172** (2005) 145 [hep-lat/0505023] [INSPIRE].

[20] C. Morningstar et al., *Improved stochastic estimation of quark propagation with Laplacian Heaviside smearing in lattice QCD*, *Phys. Rev.* **D 83** (2011) 114505 [arXiv:1104.3870] [INSPIRE].

[21] B. Blossier, M. Della Morte, G. von Hippel, T. Mendes and R. Sommer, *On the generalized eigenvalue method for energies and matrix elements in lattice field theory*, *JHEP* **04** (2009) 094 [arXiv:0902.1265] [INSPIRE].

[22] S. Güsken et al., *Nonsinglet axial vector couplings of the baryon octet in lattice QCD*, *Phys. Lett.* **B 227** (1989) 266 [INSPIRE].

[23] C. Michael and I. Teasdale, *Extracting glueball masses from lattice QCD*, *Nucl. Phys.* **B 215** (1983) 433 [INSPIRE].

[24] M. Lüscher and U. Wolff, *How to calculate the elastic scattering matrix in two-dimensional quantum field theories by numerical simulation*, *Nucl. Phys.* **B 339** (1990) 222 [INSPIRE].

[25] Alpha collaboration, B. Blossier et al., *HQET at order 1/m: II. Spectroscopy in the quenched approximation*, *JHEP* **05** (2010) 074 [`arXiv:1004.2661`] [inSPIRE].

[26] J.J. Dudek, R.G. Edwards, N. Mathur and D.G. Richards, *Charmonium excited state spectrum in lattice QCD*, *Phys. Rev.* **D 77** (2008) 034501 [`arXiv:0707.4162`] [inSPIRE].

[27] G. Burdman and J.F. Donoghue, *Union of chiral and heavy quark symmetries*, *Phys. Lett.* **B 280** (1992) 287 [inSPIRE].

[28] M.B. Wise, *Chiral perturbation theory for hadrons containing a heavy quark*, *Phys. Rev.* **D 45** (1992) 2188 [inSPIRE].

[29] T.-M. Yan et al., *Heavy quark symmetry and chiral dynamics*, *Phys. Rev.* **D 46** (1992) 1148 [*Erratum ibid.* **D 55** (1997) 5851] [inSPIRE].

[30] UKQCD collaboration, G. de Divitiis et al., *Towards a lattice determination of the B* B pi coupling*, *JHEP* **10** (1998) 010 [`hep-lat/9807032`] [inSPIRE].

[31] D. Becirevic, B. Blossier, E. Chang and B. Haas, *g(B*Bπ)-coupling in the static heavy quark limit*, *Phys. Lett.* **B 679** (2009) 231 [`arXiv:0905.3355`] [inSPIRE].

[32] H. Ohki, H. Matsufuru and T. Onogi, *Determination of B*Bπ coupling in unquenched QCD*, *Phys. Rev.* **D 77** (2008) 094509 [`arXiv:0802.1563`] [inSPIRE].