PHYSICAL REVIEW D **88,** 054506 (2013)

Nonperturbative renormalization for improved staggered bilinears

Andrew T. Lytle[1,*] and Stephen R. Sharpe[2,†]

[1]*Department of Theoretical Physics, Tata Institute of Fundamental Research, Mumbai 400005, India*
[2]*Physics Department, University of Washington, Seattle, Washington 98195-1560, USA*

(Received 8 July 2013; published 16 September 2013)

We apply nonperturbative renormalization to bilinears composed of improved staggered fermions. We explain how to generalize the method to staggered fermions in a way that is consistent with the lattice symmetries and introduce a new type of lattice bilinear that transforms covariantly and avoids mixing. We derive the consequences of lattice symmetries for the propagator and vertices. We implement the method numerically for hypercubic-smeared (HYP-smeared) and asqtad valence fermion actions, using lattices with asqtad sea quarks generated by the MILC Collaboration. We compare the nonperturbative results so obtained to those from perturbation theory, using both scale-independent ratios of bilinears (of which we calculate 26) and the scale-dependent bilinears themselves. Overall, we find that one-loop perturbation theory provides a successful description of the results for HYP-smeared fermions if we allow for a truncation error of roughly the size of the square of the one-loop term (for ratios) or of size $\mathcal{O}(1) \times \alpha^2$ (for the bilinears themselves). Perturbation theory is, however, less successful at describing the nonperturbative asqtad results.

DOI: 10.1103/PhysRevD.88.054506 PACS numbers: 12.38.Gc

I. INTRODUCTION

Precise knowledge of matching factors (Z factors) between lattice operators and their continuum counterparts is necessary for many phenomenological applications of lattice QCD. Nonperturbative renormalization (NPR) [1] is a widely used method for determining these matching factors, and has been applied successfully to many types of lattice fermion.[1] Compared to perturbative matching, which is necessarily carried out at fixed order, NPR has the great advantage of avoiding truncation errors. While the size of such errors can be estimated, the estimates are necessarily approximate.

In this paper we apply NPR to improved staggered fermions, focusing on matching factors for quark bilinears. There have been relatively few applications of NPR to staggered fermions. Most relevant for our work is a quenched calculation of Z_m, the renormalization factor for the quark mass, using unimproved staggered fermions [3]. This calculation found large discretization errors, which is typical for unimproved staggered fermions. Such errors should be significantly reduced by using improved actions as we do here.

Generalizing NPR to staggered fermions is relatively straightforward, although there are a number of technical details that do not arise with Wilson-like fermions and have not been discussed in previous work. We apply the method to the quark propagator and to quark bilinears having arbitrary spin and taste (but no derivatives). We use two types of improved staggered quarks: "asqtad" [4–6] and hypercubic-smeared (HYP-smeared) improved staggered quarks [7]. A complication arising with staggered quarks is the presence of the taste degree of freedom, which has the consequence that each bilinear comes in 16 possible tastes. In this study we turn the presence of multiple tastes into an advantage. Ratios of matching factors having the same spin but different tastes become unity in the continuum limit, but differ at finite lattice spacing.[2] The differences are proportional to both $\mathcal{O}(\alpha_s)$ (with the coupling evaluated at a scale $\sim 1/a$) and a^2p^2 (where p is the scale at which NPR is implemented). Such ratios are akin to Z_A/Z_V with Wilson fermions. Comparing them to the results from one-loop perturbation theory (PT), and studying their p dependence gives information on the accuracy of truncated PT and may allow discretization effects and perturbative contributions to be disentangled. The multiple tastes of staggered fermions allow us to form many such ratios. We also make the comparison with PT for the Z factors themselves.[3]

Our work was initially motivated by the need for matching factors in two ongoing calculations—that of quark masses by the MILC Collaboration and of electroweak matrix elements by the SWME Collaboration. The former work has determined the light quark masses using first the asqtad action [4–6] (with results reviewed in Ref. [8]) and more recently the HISQ action [9] (with results exemplified

*atlytle@theory.tifr.res.in
†srsharpe@uw.edu

[1]For a review see Ref. [2].

[2]That all ratios (including those containing the taste singlet) become unity in the continuum limit holds only because we consider flavor nonsinglet bilinears, for which there are no quark-disconnected Wick contractions.

[3]We do not present results for the (spin) pseudoscalar in this paper, since for these quantities the chiral limit is complicated by the presence of pion poles. This is discussed further in Sec. II C.

by those of Ref. [10]). These determinations use two-loop matching factors, and the concomitant truncation error is the largest source of error. A nonperturbative determination of the matching factor could firm up and reduce this error. It would also allow check of the consistency of different lattice approaches by comparing with the more precise results for light quark masses obtained in Ref. [11] using a combination of results for m_s/m_c and m_c.

First results using NPR for the matching factor Z_m with asqtad quarks were presented by one of us in Ref. [12] and extended in Ref. [13]. Using the MILC coarse ($a \approx 0.12$ fm) and fine ($a \approx 0.09$ fm) lattices to take the continuum limit, the result obtained for the strange quark mass was $m_s(\overline{\text{MS}}, 2\text{ GeV}) = 103 \pm 3$ MeV, where the error is only statistical. This is somewhat higher than the results one obtains from these two lattices using one-loop [14] ($m_s = 76 \pm 8$ MeV) and two-loop [15] matching ($m_s = 87 \pm 6$ MeV).[4] What is needed, however, is a full error budget for the NPR calculation. One aim of the present work is to study some of the systematic errors that enter into this budget.

The second ongoing calculation which motivates the present work is that of B_K (and related matrix elements) using HYP-smeared staggered fermions on the MILC asqtad configurations [16,17]. This calculation uses one-loop matching for the relevant four-fermion operators, and the truncation error again dominates that from other sources. This error can be significantly reduced using NPR. The present calculation is a step on the way, as the four-fermion operators are essentially composed of products of the bilinears studied here.

For completeness, we recall the main disadvantages of NPR. These are the need for a "window" where nonperturbative and discretization errors are small, the presence of statistical errors, and the possibility of "Gribov noise." Methods exist, however, to systematically reduce the first two errors. The window can be enlarged by combining the step-scaling technique with NPR [18] (a technique we do not use here), and statistical errors can be substantially reduced using momentum sources (which we do use). Gribov noise[5] is the uncertainty caused by the presence of multiple solutions to the gauge-fixing criterion [21–26].

This paper is organized as follows. The following section describes the application of NPR to staggered fermions, beginning with the quark propagator and then discussing bilinears. We introduce and use "covariant bilinears," which transform in irreducible representations (irreps) of the lattice symmetry group, and differ somewhat from the "hypercube bilinears" commonly used in simulations. In Sec. III we briefly describe the numerical methods we use and their implementation. We present our results in Sec. IV, providing a detailed comparison with perturbation theory. We conclude in Sec. V.

Technical results are collected in four appendices. Appendix B sketches the classification of covariant bilinears into irreps of the lattice symmetry group. In Appendix C we explain how lattice symmetries constrain the form of the quark propagator and bilinear amplitudes. In Appendix D we describe how the perturbative calculation of one-loop matching factors changes when moving from hypercube to covariant bilinears. In Appendix E we collect continuum results needed for the renormalization scale evolution of the matching factors.

Preliminary results from this study were presented in Refs. [27,28].

II. NPR FOR STAGGERED FERMIONS

For valence staggered fermions we use either the unimproved action, the HYP-smeared improved action or the asqtad action. The unimproved action is

$$S_{\text{un}} = \sum_n \bar{\chi}(n)\Big[\sum_\mu \eta_\mu(n)\nabla_\mu + m\Big]\chi(n),$$
$$\nabla_\mu\chi(n) = \frac{1}{2}[U_\mu(n)\chi(n+\hat{\mu}) - U^\dagger_\mu(n-\hat{\mu})\chi(n-\hat{\mu})] \tag{1}$$

where $\chi(n)$ is the usual single-component staggered lattice field, $n = (n_1, n_2, n_3, n_4)$ labels lattice sites, $\eta_\mu(n) = (-1)^{n_1+\cdots+n_{\mu-1}}$ is the remnant of the Dirac matrices, and $U_\mu(n)$ are the SU(3) gauge links. All quantities are dimensionless, so that, for example, the bare quark mass is related to the physical mass by $Z_m m = m_{\text{phys}} a$.

The HYP-smeared action is obtained simply by replacing the links with HYP-smeared links, $V_\mu(n)$, obtained as explained in Ref. [7]. We use the HYP-smearing parameters labeled "HYP(1)" in Ref. [29]: $\alpha_1 = 0.75$, $\alpha_2 = 0.6$ and $\alpha_3 = 0.3$.

The asqtad action [4–6] is described in Appendix A. This action is fully tree-level $O(a^2)$ improved, unlike the HYP-smeared action where only taste-breaking terms are improved.

We use configurations from the MILC Collaboration, which are generated with the asqtad action for sea quarks (using the rooting prescription to remove unwanted tastes) and the one-loop improved Symanzik action for gluons [8]. All lattices have an even number of points, L_μ, in each direction, and we use periodic boundary conditions on the propagators in all directions.

Before calculating propagators and vertices, gauge fields are fixed to Landau gauge. On the lattice, this is achieved by maximizing

[4]The result based on four lattice spacings and two-loop matching is $m_s = 88 \pm 5$ MeV [8]. We also note that the most precise determination, (obtained using the ratio m_s/m_c) is $m_s = 92.4 \pm 1.5$ MeV [11].

[5]Gribov noise can be avoided using methods based on the Schrödinger functional [19,20], but these are more complicated in practice than NPR.

$$F_L = \sum_{n,\mu} \mathrm{Tr}(U_\mu(n) + U_\mu(n)^\dagger), \tag{2}$$

for which we use an over-relaxation algorithm. This finds a local maximum, of which there are many, leading to the ambiguity of Gribov copies. We simply assume, following standard practice [21–26], that the differences in the results on different copies are small enough to ignore.

A. Quark propagator

NPR takes place in momentum space, so we must choose the appropriate momentum-space quark fields. The choice is nontrivial for staggered fermions, because the lattice Brillouin zone contains both momentum and taste information [30]. This is the momentum-space analog of the fact that the four-taste Dirac field is built up from staggered fields χ living on a 2^4 hypercube [31]. Motivated by this split into hypercubes, Ref. [3] used the momentum-space field

$$\phi'_A(p') = \sum_y e^{-ip'\cdot y}\chi(y + A), \tag{3}$$

where y is a vector labeling 2^4 hypercubes ($y_\mu = 2n^y_\mu$, with n^y_μ integers), and A is a hypercube vector labeling points within the hypercubes ($A_\mu \in \{0, 1\}$). Thus $y + A$ picks out a particular lattice point. The physical momentum[6] p' lies in a reduced Brillouin zone,

$$-\pi/2 \le p'_\mu < \pi/2, \tag{4}$$

and the label A contains the Dirac and taste indices. The key feature of the choice (3) is that the momentum phase factor does not vary within each hypercube.

This choice is, however, problematic, because $\phi'_A(p')$ does not transform irreducibly under lattice translations. This is clear from the fact that the division of the lattice into 2^4 hypercubes is not invariant under single-site translations. The lack of irreducibility implies that the propagator does not have a simple, continuumlike form.

It is straightforward to avoid this problem by using the definition introduced by Ref. [30]. One uses the standard Fourier transform, without reference to hypercubes, leading to a momentum lying in the usual Brillouin zone, $-\pi \le p_\mu < \pi$. One then breaks this up into 2^4 subzones, each characterized by a hypercube vector B, such that a general momentum is written

$$p_\mu = p'_\mu + \pi B_\mu, \tag{5}$$

with p'_μ constrained as in (4) above. p' is the physical momentum and B contains the spin and taste information. The momentum-space field of Ref. [30] is then

$$\phi_B(p') = \sum_n e^{-ip\cdot n}\chi(n) \tag{6}$$

$$= \sum_{y,A} e^{-ip'\cdot y - ip'\cdot A}(-)^{B\cdot A}\chi(y + A) \tag{7}$$

$$= \sum_A e^{-ip'\cdot A}(-)^{B\cdot A}\phi'_A(p'), \tag{8}$$

(with an identical definition for $\bar\phi$ in terms of $\bar\chi$). The second line shows that this new choice differs from ϕ'_B of Eq. (3) by the presence of a phase factor $\exp(-ip'\cdot A)$ *within* the hypercube. The last line gives the explicit relation between ϕ_B and ϕ'_B. In the continuum limit, when one can set $p' \to 0$, the two fields are simply related by a unitary transformation, and are thus physically equivalent. Away from the continuum, however, they differ in an essential way.

The merits of the choice (6) can be seen by considering the free quark propagator. First we define the propagator (with or without interactions) by

$$\langle \phi_A(p')\bar\phi_B(-q')\rangle = (2\pi)^4\bar\delta(p' - q')S(p')_{AB}, \tag{9}$$

where $\bar\delta$ is the periodic delta function (with period 2π).[7] This form follows from the invariance of the action under two-site translations without the need for phases on the quark fields. The propagator $S(p')$ has implicit color indices and explicit spin-taste indices. Altogether it is a 48×48 matrix. Invariance under global gauge transformations implies, however, that it is proportional to the identity matrix in color space, a property that holds also for its inverse. Thus we keep color indices implicit in the following discussion. For free quarks, the inverse of S is [30]

$$S^{-1}_{\mathrm{free}}(p')_{AB} = m\overline{\overline{(1\otimes I)}}_{AB} + i\sum_\mu \sin(p'_\mu)\overline{\overline{(\gamma_\mu\otimes I)}}_{AB}, \tag{10}$$

where m is the valence quark mass. Here we use the notation of Refs. [32,33] (also briefly explained in Appendix A).

The result (10) has a continuumlike form with a taste-singlet mass term and a taste-singlet derivative term; the only effect of discretization is the replacement of p'_μ with $\sin(p'_\mu)$. In particular, there are no taste-violating terms. This simplicity is guaranteed by the lattice symmetries [30], and does not hold if one uses the field (3).

In fact, one can show that the absence of taste-violating terms holds in the presence of interactions. This was shown in Ref. [30] close to the continuum limit, and is demonstrated for arbitrary p' in Appendix C. The key result is that the propagator satisfies, for each μ,

$$S(p) = \overline{\overline{(I\otimes\xi_\mu)}}S(p)\overline{\overline{(I\otimes\xi_\mu)}} \Leftrightarrow [\overline{\overline{(I\otimes\xi_\mu)}}, S(p)] = 0. \tag{11}$$

[6]This momentum is physical in the sense that the part corresponding to taste degrees of freedom has been removed. It is, however, dimensionless, containing an implicit factor of a.

[7]On a finite lattice one replaces $(2\pi)^4\bar\delta(0)$ with the number of sites, N_{site}.

This implies that $S(p)$ is a taste singlet, i.e. consists only of terms whose matrix structure is $\overline{(\gamma_S \otimes I)}$. We stress that this result holds to all orders in perturbation theory, and, indeed, nonperturbatively.

Constraints on the form of the propagator also arise from lattice rotations and spatial inversions, as discussed in Appendix C. Given that only taste-singlet terms appear, however, these constraints are identical to those that apply to other types of fermions, e.g. Wilson or overlap fermions. The final constraints arise from the $U(1)_\epsilon$ axial symmetry of the staggered action. The net effect is that the form of the inverse propagator is[8]

$$\begin{aligned} S^{-1}(p') = {} & c_S m \overline{\overline{(I \otimes I)}} + c_V p'_\mu \overline{\overline{\gamma_\mu \otimes I}} \\ & + c_T m \sum_{\mu\nu} p'_\mu (p'_\nu)^3 \overline{\overline{\gamma_{\mu\nu} \otimes I}} \\ & + c_A \sum_{\mu\nu\rho} p'_\mu (p'_\nu)^3 (p'_\rho)^5 \overline{\overline{\gamma_{\mu\nu\rho} \otimes I}} \\ & + c_P m \sum_{\mu\nu\rho\sigma} p'_\mu (p'_\nu)^3 (p'_\rho)^5 (p'_\sigma)^7 \overline{\overline{\gamma_{\mu\nu\rho\sigma} \otimes I}}, \end{aligned} \tag{12}$$

where the c_j are constants. Here we are using a somewhat schematic notation in which, for each Dirac structure, we display only the term having the lowest power of p' and m. Thus, for example, in the c_T term, there are terms not shown in which the momentum dependence is $p'_\mu (p'_\nu)^5$, etc. Such terms are suppressed in the continuum limit relative to those shown by powers of a^2. We are also using the shorthand $\gamma_{\mu\nu} = \gamma_\mu \gamma_\nu$, etc. The factors of m arise due to the $U(1)_\epsilon$ symmetry (and are thus absent in the corresponding result for Wilson fermions). As one approaches the continuum limit (i.e. as p', $m \to O(a)$) only the c_S and c_V terms survive, and one is thus guaranteed to obtain the same form as the free propagator, up to mass and wave function renormalization.

With this background we can now return to the application of NPR to staggered fermions. Since the staggered propagator has the same general form as with other fermions, supplemented only by the taste degrees of freedom, one can carry over the formalism of Ref. [1] essentially verbatim. We first calculate $S(p')$ from Eq. (9), and then, for each p', invert the resulting 48×48 matrix to obtain $S^{-1}(p')$. In the RI′ scheme, wave function renormalization is then given by

$$Z'_q(p') = -i \frac{1}{48} \sum_\mu \frac{\tilde{p}'_\mu}{\tilde{p}'^2} \mathrm{Tr}[\overline{\overline{(\gamma_\mu \otimes I)}} S^{-1}(p')]. \tag{13}$$

Here $\tilde{p}' = \sin(p')$ and $\sin(p') + \sin^3(p')/6$, respectively, for HYP and asqtad fermions. These choices are made so that, for both cases, $Z'_q = 1$ in the free theory.

[8]The same constraint applies to the propagator, but for the NPR procedure it is more convenient to focus on S^{-1}.

The shorthand $\tilde{p}'^2$ means $\sum_\mu (\tilde{p}'_\mu)^2$, and the trace is over spin, taste and color indices. As always, with NPR, one aims to work in the window

$$\Lambda^2_{\mathrm{QCD}} \ll p'^2 \ll \left(\frac{\pi}{a}\right)^2, \tag{14}$$

so as to avoid nonperturbative effects and discretization errors. We discuss these constraints further when we present results.

The quark propagator allows one, in principle, to determine the mass renormalization factor Z_m, using

$$\frac{1}{48} \mathrm{Tr}[\overline{\overline{(I \otimes I)}} S^{-1}(p')] = Z'_q(p') \left[Z_m(p') m + C_1 \frac{\langle \bar{\chi} \chi \rangle}{p'^2} \right]. \tag{15}$$

Here we display the leading nonperturbative correction, obtained in Refs. [34,35] using the operator product expansion. In practice, as is well known, this method of determining Z_m has larger nonperturbative corrections than that (to be described in Sec. II C) using vertex functions.

B. Covariant quark bilinears

Before discussing vertex functions we introduce the bilinear operators used in our numerical calculations. The conventional choice for bilinears relies on a partitioning of the lattice into 2^4 hypercubes. For operators at zero momentum, which is all we consider here, these take the form

$$\mathcal{O}_{S \otimes F} = \frac{1}{N_y} \sum_y \sum_{A,B} \bar{\chi}_A(y) \overline{(\gamma_S \otimes \xi_F)}_{AB} \mathcal{U}_{y+A,y+B} \chi_B(y). \tag{16}$$

Here y labels hypercubes as above, with N_y being the total number in the lattice. The hypercube fields are defined by [32]

$$\bar{\chi}_A(y) = \frac{1}{4} \chi(y + A) \quad \text{and} \quad \chi_B(y) = \frac{1}{4} \chi(y + B). \tag{17}$$

The normalization is such that, in the continuum limit, the matrix element of $\mathcal{O}_{S \otimes F}$ is the same as that of $a^3 \int d^4x \bar{Q}(\gamma_S \otimes \xi_F) Q/V$, with V the four-volume [36].

The bilinears are made gauge invariant by the inclusion of $\mathcal{U}_{y+A,y+B}$, which is the average over products of gauge links along minimal-length paths connecting the $\bar{\chi}$ and χ fields. We have investigated various choices of links:

(1) For unimproved or asqtad valence quarks, a possible choice is the original gauge links, tadpole-improved: U_μ/u_0. We find that this leads in general to Z factors differing substantially from unity, and poor convergence of perturbative predictions. We do not present results for this choice.

(2) For asqtad valence quarks one can also use the Fat-7 + Lepage smeared links W_μ. The resulting links are closer to unity, and couple less strongly to gluons with momenta of $\mathcal{O}(1/a)$. This is the

choice for which we present results with asqtad quarks.

(3) For HYP valence quarks, we use HYP-smeared links.

The operators (16) do not, in general, transform irreducibly under translations, because they rely on a particular partitioning of the lattice into hypercubes. As discussed in Ref. [37], they can be written as linear combinations of operators with definite, and in general different, transformation properties. These operators are distinguished by having different numbers of derivatives and thus varying dimensions. The operators of lowest dimension are those with no derivatives and thus $d = 3$: these are the "translationally covariant" ("covariant" for short) hypercube operators.

Although noncovariant four-fermion operators are being used in the calculations of B_K with staggered fermions, we have chosen to use covariant bilinears in the present study. This is because these operators are simpler to code, and have simpler renormalization properties. Indeed, if one were calculating matrix elements of staggered bilinears, such as those needed for $K \to \pi$ semileptonic form factors, then covariant bilinears would be a natural choice.

The explicit form of these operators was not determined in Ref. [37], so we construct them here. A simple approach is to adapt the methodology developed in Ref. [38] for the construction of irreducible baryon operators. The key point is that, when separating the quark and antiquark fields in the bilinear, one obtains objects which transform irreducibly under translations if one uses "symmetric shifts." These are shifts in which one averages over forward and backward directions (including, of course, the gauge links necessary for gauge invariance). The operator in Eq. (16) is not of this form. For example, for a vector current with $S = (1000)$ and $F = (0000)$, if $A = (0000)$ then $B = (1000)$ and one only has the link pointing in one direction. A symmetric shift would include terms with $A = (0000)$, $B = (-1000)$ as well as $A = (2000)$, $B = (1000)$ (each weighted by a factor of $1/4$) in addition to the original term (with a weight of $1/2$ since it appears both when shifting the χ field and the $\bar{\chi}$).[9]

This example illustrates the general prescription for converting the hypercube operators (16) into covariant operators. For given values of A and B [and recalling that, for fixed S and F, only one value of B contributes for each A, namely $B =_2 A + S - F$ (with the subscript indicating mod-2 arithmetic)], one replaces $\bar{\chi}_A \mathcal{U}_{y+A,y+B} \chi_B$ with

$$\frac{1}{2N_\Delta} \sum_\Delta (\bar{\chi}_A \mathcal{U}_{y+A,y+A+\Delta} \chi_{A+\Delta} + \bar{\chi}_{B-\Delta} \mathcal{U}_{y+B-\Delta,y+B} \chi_B), \tag{18}$$

where the set of N_Δ allowed vectors Δ are those obtained from $B - A$ by independently changing the signs of the nonzero components, including no changes. For example, if $B - A = (1100)$, then

$$\Delta = (1100), (-1100), (1\ -100), (-1\ -100), \tag{19}$$

and so $N_\Delta = 4$.

After some algebraic manipulations, the resulting operator can be written

$$\mathcal{O}^{\rm cov}_{S\otimes F} = \frac{1}{N_{\rm site}} \sum_n \frac{1}{16} \sum_{A,B} \bar{\chi}(n) \overline{(\gamma_S \otimes \xi_F)}_{n,n+B-A} \times \mathcal{U}_{n,n+B-A} \chi(n + B - A). \tag{20}$$

The factor of $1/16$ is required in order to retain the same normalization as in (16), because of the definition $\chi(y)_A = (1/4)\chi(y + A)$. The double sum over A and B in (20), which is really a single sum since $\overline{(\gamma_S \otimes \xi_F)}$ enforces $B =_2 A + S + F$, corresponds to the sum over Δ in the symmetric shift. This can be made explicit by writing the operator as

$$\mathcal{O}^{\rm cov}_{S\otimes F} = \frac{1}{N_{\rm site}} \sum_n \frac{1}{N_\Delta} \sum_{|\Delta|=|S-F|} \bar{\chi}(n) \overline{(\gamma_S \otimes \xi_F)}_{n,n+S-F} \times \mathcal{U}_{n,n+\Delta} \chi(n + \Delta), \tag{21}$$

where the second sum is over the N_Δ allowed values of Δ. This result makes the presence of symmetric shifts manifest. Note that the sign arising from $\overline{(\gamma_S \otimes \xi_F)}$ is independent of Δ, and that the form (21) removes some redundancy in the sums of (20).

The forms (20) and (21) show explicitly that the covariant bilinears do not require a partitioning of the lattice into hypercubes. This simplifies their numerical implementation, since one can freely sum over n. The computation of the link factors is the most costly part of the calculation, with the cost growing rapidly with $|\Delta|$.

C. Vertex renormalization

To determine matching factors of general bilinears we must calculate the vertex functions. We consider here only the case of exceptional kinematics in which the operator inserts no momentum,

$$\Lambda^{S\otimes F}_{AB}(p') = \frac{1}{N_{\rm site}} \langle \phi^a_A(p') \mathcal{O}^{{\rm cov};(ab)}_{S\otimes F} \bar{\phi}^b_B(-p') \rangle. \tag{22}$$

Like the propagator, the vertex is 48×48 matrix, with the color part being trivial. The new indices a and b in the superscripts are flavor indices. We always choose $a \neq b$ so that the operator is a flavor nonsinglet, which implies that there is only a single quark contraction between the external fields and the operator. The fields in the vertex are valence quarks and antiquarks, as for the propagator.

[9] Here it is convenient to allow the vectors A and B to range outside the hypercube.

One now follows the perturbative renormalization procedure, amputating the vertex with the previously calculated inverse propagators,

$$\Gamma^{S\otimes F}(p') = S^{-1}(p')\Lambda^{S\otimes F}(p')S^{-1}(p'). \tag{23}$$

Matching factors are determined by enforcing the tree-level form of Γ when fields and operators are renormalized,

$$\frac{Z'_q(p')}{Z_{S\otimes F}(p')} = \frac{1}{48}\frac{\mathrm{Tr}[\overline{(\gamma_S\otimes\xi_F)}^{\dagger}\Gamma^{S\otimes F}(p')]}{V_{S\otimes F}(p')}. \tag{24}$$

Here we assume no mixing, which is the case for the covariant bilinears. This is shown nonperturbatively in Appendix C. We have also divided the projected vertex by its tree-level expression, $V_{S\otimes F}$. This has the form $1 + O(a^2)$, and is given explicitly in Eq. (D4). Dividing by V removes some of the discretization errors, and this approach is common practice in NPR.

One can use the following lattice Ward identities to relate matching factors:

$$\frac{1}{m}\mathrm{Tr}[\overline{(I\otimes I)}S^{-1}(p)] = \mathrm{Tr}[\overline{(\gamma_5\otimes\xi_5)}\Gamma^{5\otimes 5}(p)], \tag{25}$$

$$\frac{\partial}{\partial m}\mathrm{Tr}[\overline{(I\otimes I)}S^{-1}(p)] = \mathrm{Tr}[\overline{(I\otimes I)}\Gamma^{I\otimes I}(p)]. \tag{26}$$

These follow by standard manipulations, and hold as written only when m is the valence quark mass [so that the derivative in (26) does not act on sea quark masses] and the operators in the vertices are flavor nonsinglets [so that there are no "quark-disconnected" contractions]. Using the definition (24) for the right-hand sides and inserting the result (15) into the left-hand sides, we find, at sufficiently large p'^2, that

$$\begin{aligned} Z_q(p')Z_m(p') &= \frac{Z_q(p')}{Z_P(p')} = \frac{Z_q(p')}{Z_S(p')} \\ \Rightarrow Z_m(p') &= \frac{1}{Z_P(p')} = \frac{1}{Z_S(p')}. \end{aligned} \tag{27}$$

These are the familiar relations from continuum perturbation theory, which here hold nonperturbatively.

We can now see why it is better to use the vertex rather than the propagator [Eq. (15)] to determine $Z_S = 1/Z_m$. This is because the condensate term in (15), which gives a significant correction at typical values of p', is absent in the scalar vertex. This can be seen by inserting (15) in the left-hand side of (26). The condensate appearing in the operator product expansion is evaluated in the chiral limit, so the $\partial/\partial m$ removes this $1/p'^2$ contribution. By contrast, a similar analysis for the pseudoscalar vertex shows that there is a nonperturbative correction proportional to $\langle\bar{q}q\rangle/(mp'^2)$. This is the well-known pion pole contribution [1], which makes the direct determination of Z_P difficult.

One can also use axial Ward identities to show that $Z_{S\otimes F} = Z_{S5\otimes F5}$, where the subscript $S5\otimes F5$ indicates the bilinear with spin taste $\overline{(\gamma_S\gamma_5\otimes\xi_F\xi_5)}$. We do not reproduce the derivation as this result is already known to hold to all orders in perturbation theory [37].

D. Irreducible representations and perturbative matching

The 16^2 covariant bilinears $\mathcal{O}^{\mathrm{cov}}_{S\otimes F}$ fall into 35 irreps under the lattice symmetry group. These are collected in Table I, organized according to the number of links, i.e. the separation between quark and antiquark fields. To our knowledge, this decomposition into irreps for covariant bilinears has not been demonstrated previously in the literature. Thus we provide a brief demonstration in Appendix B.

As already noted above, matching factors for operators with spin taste $(\gamma_S\otimes\xi_F)$ and $(\gamma_S\gamma_5\otimes\xi_F\xi_5)$ are the same. This reduces the number of independent matching factors from 35 to 19, as described in the caption to the table.

Since our aim is to compare to perturbation theory, we need the one-loop matching factors for the covariant operators. It turns out, for reasons discussed in Appendix D, that they can be obtained from those for hypercube bilinears in a trivial way: one simply has to drop the mixing terms, with the diagonal matching factors being unchanged. The lack of mixing is a direct result of using covariant operators, since different spin tastes lie in different irreps of the lattice symmetry group.

TABLE I. Spin-taste assignments of covariant bilinears forming irreps of the lattice symmetry group. Indices μ, ν and ρ are summed from 1 to 4, except that all are different. If two indices appear in either the spin or the taste, there may be some redundancy, e.g. in $(1\otimes\xi_\mu\xi_\nu)$ one can enforce $\mu<\nu$ so that the dimension of the irrep is 6. Pseudoscalar and axial bilinears are not listed: they can be obtained from scalar and vector, respectively, by multiplication by $\gamma_5\otimes\xi_5$. Bilinears related in this way have the same matching factors. This operation also implies the equality of the matching factors for the three pairs of tensor bilinears within square brackets.

Number of links	S	V	T
4	$(I\otimes\xi_5)$	$(\gamma_\mu\otimes\xi_\mu\xi_5)$	$(\gamma_\mu\gamma_\nu\otimes\xi_\mu\xi_\nu\xi_5)$
3	$(I\otimes\xi_\mu\xi_5)$	$(\gamma_\mu\otimes\xi_5)$ $(\gamma_\mu\otimes\xi_\nu\xi_\rho)$	$[(\gamma_\mu\gamma_\nu\otimes\xi_\mu\xi_5)(\gamma_\mu\gamma_\nu\otimes\xi_\rho)]$
2	$(I\otimes\xi_\mu\xi_\nu)$	$(\gamma_\mu\otimes\xi_\nu)$ $(\gamma_\mu\otimes\xi_\nu\xi_5)$	$[(\gamma_\mu\gamma_\nu\otimes I)(\gamma_\mu\gamma_\nu\otimes\xi_5)]$ $(\gamma_\mu\gamma_\nu\otimes\xi_\nu\xi_\rho)$
1	$(I\otimes\xi_\mu)$	$(\gamma_\mu\otimes I)$ $(\gamma_\mu\otimes\xi_\mu\xi_\nu)$	$[(\gamma_\mu\gamma_\nu\otimes\xi_\nu)(\gamma_\mu\gamma_\nu\otimes\xi_\rho\xi_5)]$
0	$(I\otimes I)$	$(\gamma_\mu\otimes\xi_\mu)$	$(\gamma_\mu\gamma_\nu\otimes\xi_\mu\xi_\nu)$

Expressions for the required diagonal matching factors are given in Ref. [39] in terms of a single lattice loop integral. Numerical values are, however, not given for the HYP-smearing coefficients that we use, nor for mean-field improved asqtad bilinears. We have calculated these values and collect them in Appendix D.

III. NUMERICAL IMPLEMENTATION

We use the Chroma [40] software library for Landau-gauge fixing, HYP smearing, and asqtad inversions. We have added code to implement momentum sources, to invert the unimproved staggered fermion matrix (needed for HYP-smeared fermions), and to construct the bilinears including the gauge links. Stopping criteria for gauge fixing and propagator inversions were set so that the errors are smaller than those from other sources, and in particular from statistics [13].

Our gauge configurations are taken from the MILC coarse ($a \approx 0.12$ fm) and fine ($a \approx 0.09$ fm) ensembles [8], which are generated using asqtad fermions and Symanzik-improved gauge action. Relevant details are given in Tables II and III. We include results for the $\tilde{u}_0$ factors needed for mean-field improvement; these are defined in Appendix D.

The momenta p' that we use are listed in Table IV. These are chosen so that the components are comparable in all four directions (after inclusion of $2\pi/L_{s,t}$ factors), ensuring that no single component becomes too large. This is known to reduce discretization errors. These choices cover the expected NPR window, as will be seen below.

For every gauge-fixed configuration in our ensemble $\{U^i\}$ and each physical momentum p' under consideration, we invert the Dirac operator D on 16 momentum sources, solving $DS = e^{ip\cdot n}$ for $p = p' + \pi B$ to obtain

$$S^i(n, p' + \pi B) = \langle \chi(n)\bar{\phi}_B(-p')\rangle_{U^i}, \tag{28}$$

where color indices are suppressed. We next Fourier transform the free space index with the 16 different momenta $p = -p' + \pi A$, leading to the 16×16 momentum-space propagator matrix $S^i(p')_{AB}$ of Eq. (9). This is then averaged over configurations to obtain the propagator $S(p')_{AB}$ of (9). Lattice symmetries predict that the inverse propagator contains only $1 \otimes 1$ and $\gamma_\mu \otimes 1$ contributions up to terms suppressed by a^4 [see Eq. (12)]. We have checked that noncontinuum terms are in fact consistent with zero within our statistical errors.

Vertex functions are constructed from $S^i(n, p' + \pi B)$ and

TABLE II. Parameters of coarse ensembles. The lattices are of size $20^3 \times 64$. The quoted masses are for the light (average of up and down) quarks, there is in addition a strange sea quark of fixed bare mass $am_{\text{sea,strange}} = 0.05$. Lattice spacings are obtained using $r_1 = 0.3108$ fm and taken from Ref. [8]. Extrapolations to the chiral limit are done with a linear fit. The quoted number of momenta are for valence asqtad/HYP-smeared fermions.

$am_{\text{sea}}/am_{\text{val}}$	Number of configs.	Number of momenta	a^{-1} [GeV]	u_0	$\tilde{u}_0^{\text{HYP}}$	$\tilde{u}_0^{\text{ASQ}}$
0.03/0.03	16	7/10	1.682	0.8696	0.9845	1.0521
0.02/0.02	16	7/10	1.679	0.8688	0.9843	1.0525
0.01/0.01	16	7/10	1.662	0.8677	0.9841	1.0528
Chiral	...	7/10	1.654	...	0.9839	1.0532

TABLE III. Parameters of asqtad fine ensembles. Lattices are of size $28^3 \times 96$, and the strange sea quark mass is $am_{\text{sea,strange}} = 0.031$.

$am_{\text{sea}}/am_{\text{val}}$	Number of configs.	Number of momenta	a^{-1} [GeV]	u_0	$\tilde{u}_0^{\text{HYP}}$	$\tilde{u}_0^{\text{ASQ}}$
0.0124/0.0124	16	8	2.357	0.8788	0.9869	1.0507
0.0093/0.0093	16	8	2.352	0.8785	0.9868	1.0508
0.0062/0.0062	16	8	2.349	0.8782	0.9868	1.0509
Chiral	...	8	2.340	...	0.9867	1.0511

TABLE IV. Physical momenta used in our calculations. The four vectors are in units of $(2\pi/L_s, 2\pi/L_s, 2\pi/L_s, 2\pi/L_t)$, where L_s (L_t) is the number of sites in the spatial (temporal) directions.

Lattice	Fermion	momenta
Coarse	asqtad	(1, 2, 2, 4), (2, 1, 2, 6), (2, 2, 2, 7), (2, 2, 2, 8), (2, 2, 2, 9), (2, 3, 2, 7), (3, 3, 3, 9)
Fine	asqtad	(1, 2, 2, 5), (2, 2, 2, 6), (2, 2, 2, 7), (2, 2, 2, 8), (2, 2, 3, 8), (2, 3, 3, 9), (3, 3, 3, 10), (3, 3, 3, 12)
Coarse	HYP	(1, 1, 1, 4), (1, 1, 1, 6), (1, 2, 1, 5), (1, 2, 2, 4), (2, 1, 2, 6), (2, 2, 2, 7), (2, 2, 2, 8), (2, 2, 2, 9), (2, 3, 2, 7), (3, 3, 3, 9)
Fine	HYP	(1, 2, 2, 5), (2, 2, 2, 6), (2, 2, 2, 7), (2, 2, 2, 8), (2, 2, 3, 8), (2, 3, 3, 9), (3, 3, 3, 10), (3, 3, 3, 12)

$$S^i(p' + \pi A, n) = \langle \phi_A(p') \bar{\chi}(n) \rangle_{U^i}. \quad (29)$$

As usual, the latter propagator may be obtained from the former using the staggered analogue of γ_5-hermiticity of the Dirac operator, $\epsilon D^\dagger \epsilon = D$, where $\epsilon = (-1)^{n_1+n_2+n_3+n_4} \equiv (-1)^n$ is the alternating phase factor. We find

$$S^i(p' + \pi A, n) = (-1)^n S^i(n, p' + \pi \tilde{A})^\dagger, \quad (30)$$

where $\tilde{A} = A +_2 (1, 1, 1, 1)$, and the Hermitian conjugation acts on color indices. The two propagators are then tied together with the bilinear. For example, the pseudoscalar (unamputated) vertex is

$$\begin{aligned}\Lambda^{\gamma_5 \otimes \xi_5}_{AB}(p') &= \langle \phi_A(p') \mathcal{O}^{\mathrm{cov}}_{\gamma_5 \otimes \xi_5} \bar{\phi}_B(-p') \rangle \\ &= \frac{1}{N_{\mathrm{conf}} N_{\mathrm{site}}} \sum_{i,n} S^i(p' + \pi A, n)(-1)^n \\ &\quad \times S^i(n, p' + \pi B) \\ &= \frac{1}{N_{\mathrm{conf}} N_{\mathrm{site}}} \sum_{i,n} S^i(n, p' + \pi \tilde{A})^\dagger S^i(n, p' + \pi B).\end{aligned} \quad (31)$$

In the general the two propagators end at positions differing by a hypercube vector, and are connected by an average over products of links over minimal-length paths [cf. Eq. (21)].

Amputation and determination of the Z factors is then performed using Eqs. (23) and (24). For a given value of p', these involve manipulations of 16×16 matrices, which can be done in the analysis phase of the calculation. The Z factor corresponding to a given irrep is determined by averaging the traces of the amputated vertex functions in that irrep; e.g., for the vector we compute Z_V from $\frac{1}{4}\,\mathrm{Tr} \sum_\mu (\overline{\gamma_\mu \otimes 1}) \Gamma^{\gamma_\mu \otimes 1}(p')$.

For both asqtad and HYP-smeared fermions, and for both coarse and fine lattices, we use valence quarks with bare masses equal to those of the light (asqtad) sea quarks. Thus our calculations are unquenched for asqtad valence fermions. For HYP-smeared fermions we are, however, using a mixed fermion action (different discretizations of valence and sea quarks) and, in addition, a partially quenched setup (because Z factors for HYP-smeared and asqtad fermions are different, so that the physical masses of sea and valence quarks differ even though the bare masses are equal). For both types of valence fermions we extrapolate our final results to the chiral limit using a linear fit. If the dependence on quark masses is linear and weak, this extrapolation will remove partial quenching effects for the HYP-smeared fermions. Residual effects from using a mixed action should vanish in the continuum limit, and thus appear as additional discretization errors for $a > 0$. Examples of the chiral fits for $Z_{\gamma_\mu \otimes 1}$ are shown in Figs. 1 and 2. These are typical in terms of the quality of fits,

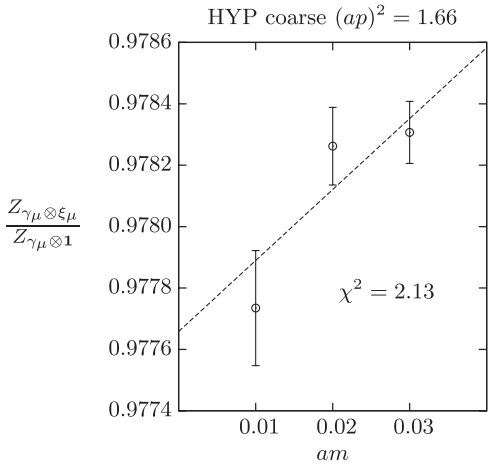

FIG. 1. Example of chiral extrapolation for the ratio $Z_{\gamma_\mu \otimes \xi_\mu}/Z_{\gamma_\mu \otimes 1}$ on the HYP coarse ensemble.

although the extent of the chiral extrapolation is greater for scalar bilinears. We also use linear chiral fits to determine values for $1/a$, $\tilde{u}_0^{\mathrm{HYP}}$ and $\tilde{u}_0^{\mathrm{ASQ}}$ which we use in subsequent analysis. These are shown in Tables II and III. We stress that these are very mild extrapolations, so that none of our conclusions would be changed were we to take the values of these quantities from, say, the lattice spacing with the smallest values of the valence quark masses.

The only exception to the above discussion of chiral extrapolations are the matching factors for pseudoscalar bilinears. As discussed after Eq. (27), these are singular in the $m \to 0$ limit [1]. It is possible to remove the singular part in various ways, but in this work we have chosen to exclude the pseudoscalars from our analysis.

Although we can extrapolate to the chiral limit for the two light quarks, our calculations have the strange sea-quark mass fixed at approximately its physical value. Strictly speaking, this means that our NPR results are not in

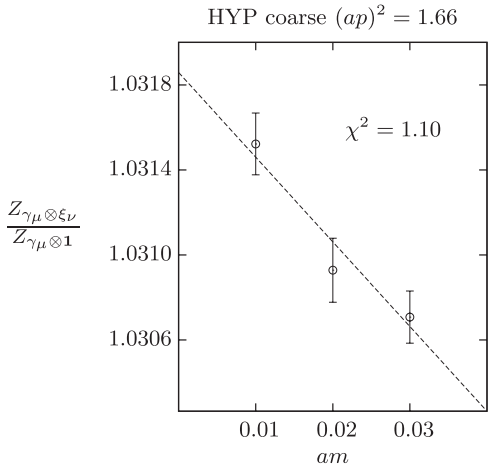

FIG. 2. Example of chiral extrapolation for the ratio $Z_{\gamma_\mu \otimes \xi_\nu}/Z_{\gamma_\mu \otimes 1}$ on the HYP coarse ensemble.

the desired mass-independent renormalization scheme. However, given the mild dependence on quark mass that we observe, we expect that this shortcoming will have little impact on the final results. In particular, we assume the error that this introduces to be smaller than the truncation errors in the one-loop PT expressions to which we compare.

We compute diagonal Z factors for all 256 choices of spin and taste. We have checked in some cases that off-diagonal contributions to Z factors are consistent with zero, as expected given that the covariant bilinears do not mix. We use 16 decorrelated configurations, which we find to be sufficient when using momentum sources. We then combine the 256 choices into the irreps listed in Table I, which further reduces the errors. All errors are obtained using single-elimination jackknife.

IV. RESULTS

We divide our discussion of the results into three parts. In the first two we consider ratios of Z factors in which the numerator and denominator have the same spins but different tastes. Specifically, we consider ratios in which the denominators are taste singlets[10]:

$$\frac{Z_{S\otimes F}(p,a)}{Z_{S\otimes 1}(p,a)} = 1 + \frac{\alpha(\mu_0)}{4\pi}[C^{\mathrm{LAT}}_{S\otimes F} - C^{\mathrm{LAT}}_{S\otimes 1}] + \mathcal{O}([ap]^2) + \ldots. \tag{32}$$

As discussed in Appendix D, PT predicts these ratios to be independent of the NPR momentum p since they are dominated by contributions from loop momenta near the cutoff scale. In particular, α is to be evaluated at a scale $\mu_0 \sim 1/a$ which is not related to $|p|$. This is illustrated by the right-hand side of Eq. (32), which shows the one-loop expression for the simplified case of no mean-field improvement.[11] These ratios are thus good quantities to use to test the accuracy of PT since one does not have to worry about anomalous dimensions. They are analogous to Z_A/Z_V with Wilson-like fermions, with the analogue of the lack of p dependence being the fact that Z_A/Z_V calculated in different ways should agree up to discretization errors.

The lack of dependence of the ratios on p does not carry over to discretization effects [represented in Eq. (32) by the $(ap)^2$ term] or to nonperturbative effects, which behave as inverse powers of p and are important only for small p. One can hope to disentangle these effects by studying the p and a dependence of the ratios, as we discuss below.

[10]In this section we will denote the NPR momentum scale by p, which has physical units. Thus ap here corresponds to the p' used in previous sections. We also make explicit that the Z factors depend separately on p and a in general.

[11]The C^{LAT} are finite lattice constants. For the general expression including mean-field improvement see Eqs. (D6)–(D10).

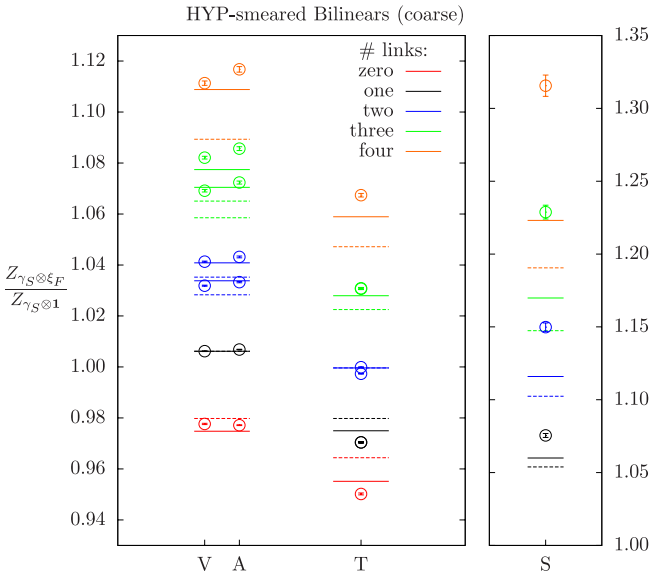

FIG. 3 (color online). Comparison of Z-factor ratios obtained using NPR to one-loop perturbation theory for HYP fermions on coarse lattices. V, A, T and S refer to bilinears with vector, axial, tensor and scalar spins, respectively. Horizontal lines show perturbative predictions, with solid/dotted lines showing results with/without mean-field improvement. Results are in the chiral limit for the momentum described in the text.

In the final part of this section we present results for the denominators of the ratios. These do have anomalous dimensions, so we can see how well the p dependence agrees with the perturbative predictions. These predictions can be made using continuum perturbation theory, for which results are known to three- or four-loop order (as described in Appendix E).

A. Ratios for HYP-smeared bilinears

We begin by discussing the results with HYP-smeared fermions. In Figs. 3 and 4 we display results for all ratios at a fixed NPR momentum. We choose $ap = (2, 2, 2, 7)$ in units of $(2\pi/L_s, 2\pi/L_s, 2\pi/L_s, 2\pi/L_t)$, so that $(ap)^2 \approx 1.66$ and 0.81, respectively, on coarse and fine lattices. This turns out to correspond to nearly the same physical value, $|p| \approx 2.1$ GeV, for both lattice spacings. We expect that this choice satisfies the window condition (14) for both lattice spacings.

These figures show the comparison of the 26 ratios involving bilinears with vector, axial, tensor and scalar spins to one-loop PT. We show perturbative predictions both without [Eq. (32)] and with [Eq. (D9)] mean-field improvement [41].[12] For these predictions we use $\mu_0 = 1.8/a$ leading to $\alpha(\mu_0) \approx 0.24$ and 0.21 on the coarse and fine lattices, respectively. For the mean-field improved

[12]We stress that we are using mean-field improvement to obtain an improved perturbative prediction for the *same* operators.

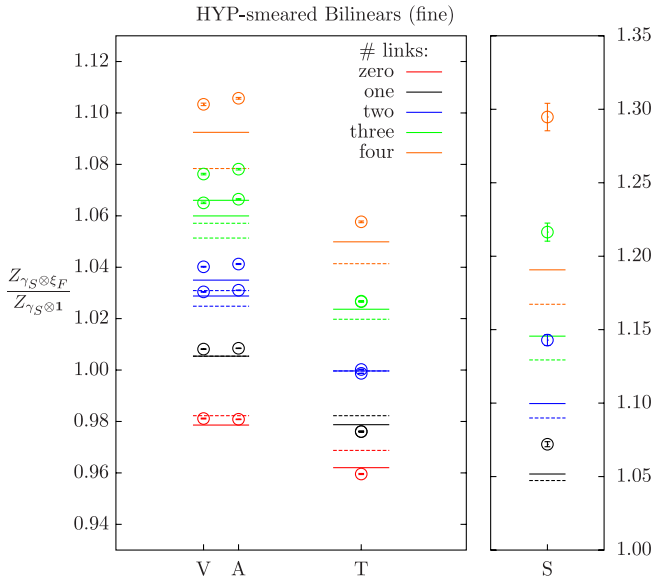

FIG. 4 (color online). As for Fig. 3 but on the fine lattices.

prediction, we also need values for $\tilde{u}_0^{\mathrm{HYP}}$, which are given in Table II. The color coding in the plots indicates the number of links in the operators in the numerators of the ratios. The denominators have 1-link operators for spins V and A, 2-link operators for T, and 0-link operators for S.

Overall the one-loop prediction works well. We highlight certain features. First, the statistical errors in the NPR results are very small, particularly for spins V, A and T. Second, PT correctly captures the ordering with link number, and placement relative to unity. This ordering is the dominant feature of the results, and indicates that the fluctuations in individual smeared links (which reduce their [gauge-fixed] average values below unity) are the largest contributor to Z factors differing from unity. Bilinears with more links thus have smaller matrix elements and require larger Z factors to attain the canonical normalization. This argument would imply that ratios involving 1-link V and A ratios and 2-link T ratios should lie close to unity, since the numerators and denominators have the same number of links. This is indeed what is observed.

Third, PT correctly predicts the "fine structure" within a given link number. For example, for spins V and A, there are two ratios involving three-link numerators, and two involving two-links (see Table I). The predicted orderings and splittings match well with PT. There is a similar fine structure for the tensors, though this is hard to discern from the figure. For one-link numerators, there are two ratios, which are predicted to be equal to all orders in PT. The NPR results for these two ratios are indistinguishable. The same is true for three-link numerators. For two-link numerators there are also two ratios, but in this case they are predicted to be equal at one-loop order but not to all orders. Here the NPR results for the two ratios do differ, but the difference is very small (and consistent with a two-loop or higher-order perturbative effect).

Fourth, we recall that matching factors for spins V and A are predicted to be equal to all-orders in PT. We observe very small (subpercent level) differences. Differences can arise due to long-distance nonperturbative effects, and so these effects are small in these channels.

Fifth, we note that the NPR results on the fine lattices are all slightly closer to unity than those on the coarse lattices. This is qualitatively what one would expect if the dominant contribution to the difference from unity was perturbative, since $\alpha(\mu_0)$ decreases with a if $\mu_0 \sim 1/a$. However, a complete interpretation of this result requires understanding the contributions of nonperturbative and discretization errors, which we discuss below.

We now discuss the level of quantitative agreement between the NPR results and one-loop PT. With the couplings we have chosen, the agreement is at the subpercent level for spins V, A and T, and at the 5–10% level for scalars. We cannot, of course, expect perfect agreement because we have truncated PT. One way of estimating the uncertainty in the one-loop prediction is to vary the scale at which α is evaluated over a reasonable range. Were we to use $\mu_0 = 1/a$ rather than $1.8/a$, the couplings would become roughly 30% larger ($\alpha \approx 0.32$ and 0.27 on the coarse and fine lattices, respectively). This would lead to a much improved quantitative agreement with the scalar ratios, while that with spins V, A and T would be less good.[13] But the most important point is that, within this perturbative uncertainty the PT and NPR results agree.

As can be seen from the numerical values, the coefficients of α in the one-loop predictions have magnitudes smaller than unity for all except the three and four-link scalar ratios. Our way of estimating the uncertainties in perturbative predictions assumes that a small one-loop coefficient implies that higher orders are also small. An alternate, and more conservative, approach is to say that, for all ratios, two-loop effects are of size $\mathcal{O}(1) \times \alpha^2 \approx$ 0.05–0.09. This gives a larger estimate than that obtained above except for the three- and four-link scalar ratios, for which the two estimates agree.

Our final comment on these two figures is that we find the impact of mean-field improvement to be fairly minor, below the level of the uncertainty due to the choice of α. The effects are small because $\tilde{u}_0^{\mathrm{HYP}}$ is very close to unity, and lies close to its perturbative prediction [Eq. (D11)]. Evaluating α at the scale $1/a$ the prediction for the coarse and fine lattices are 0.985 and 0.988, respectively, to be compared to 0.984 and 0.987 (from Tables II and III).

We next display the NPR renormalization-scale dependence of the ratios. Figures 5–7 show this respectively for

[13]It would be interesting to use an approximate scale-setting method to better predict the appropriate value of α to use for each quantity. Our data suggests that a lower scale would be found for scalars than for the other bilinears.

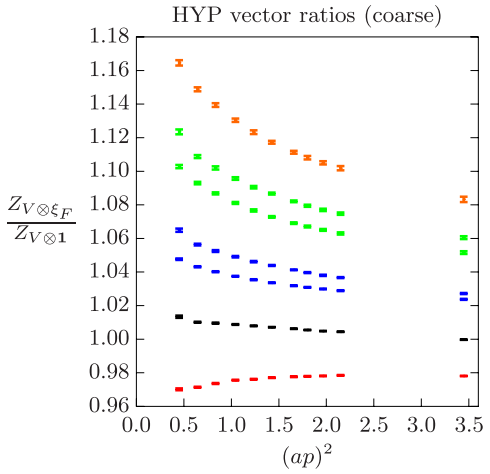

FIG. 5 (color online). Scale dependence of the ratios of vector HYP-smeared bilinear Z factors on coarse MILC lattices. The color coding indicates the link number of the numerator and corresponds to that in Fig. 3 (where the results for $(ap)^2 = 1.66$ are shown).

the S, T and V ratios on the coarse lattices. We omit the axial ratios since they are very similar to the vectors. We recall that PT predicts to all orders that the ratios should be independent of p, up to discretization effects at large $(ap)^2$ and nonperturbative effects at small p^2. We might hope that the window in which such effects would be small runs (for the coarse lattices) from $|p| \approx 1$ GeV($\Rightarrow (ap)^2 \approx 0.4$) up to $(ap)^2 \sim 2$–3, i.e. the entire width of our data set. In fact we find moderate scale dependence for vector ratios (a doubling of the separation from unity from the high end to the low end of the range), a significantly smaller dependence for the tensor ratios, but a very strong dependence for the scalars.

A possible interpretation of these results is as follows. The curvature at small $(ap)^2$ suggests nonperturbative effects proportional to powers of $1/p^2$. These are largest for the scalar ratios, and for these the lower edge of the NPR window should be moved up to $(ap)^2 \approx 1.5$ (corresponding to $|p| \approx 2$ GeV). The data above this value can be reasonably well fit by a straight line, consistent with discretization errors. Extrapolating to $(ap)^2 = 0$ removes these discretization errors. For the vector and tensor ratios the lower edge of the window can be placed at $(ap)^2 \approx 1$ (corresponding to $|p| \approx 1.6$ GeV). This is how many NPR results have traditionally been analyzed (see, e.g., Ref. [3]).

We do not carry out these extrapolations quantitatively, because there is clearly an uncertainty introduced by the choice of fitting window, and we are in this work not aiming to quote results with a full error analysis. Nevertheless, what is clear from the figures is that, after extrapolation, the overall features found at $(ap)^2 = 1.66$ and shown in Fig. 3 would still hold. The only change would be that the ratios would be pushed further away from unity: by 15–20% for V, A and T ratios and by $\approx 50\%$ for the scalar ratios. Thus for quantitative agreement at $ap = 0$ for V, A and T ratios one needs to use $\alpha \approx 0.28$, corresponding to $\mu_0 \approx 1.3/a \approx 2.2$ GeV, while for scalars one needs $\alpha \approx 0.48$, corresponding to $\mu_0 \approx a/2 \approx 0.8$ GeV. In the former case the scale is reasonable and the value of α small enough for reasonable convergence, but for the scalars the convergence of PT is suspect.

It is interesting to ask whether the $(ap)^2$ corrections are of the expected size. If the ratios are described approximately by $R(ap = 0)[1 + x(ap)^2]$, then, if we take the relevant scale for cutoff effects to be π/a, and assume that the (approximate) improvement of the actions leads to a reduction by $\sim\alpha$, then we would expect $|x| \approx \alpha/\pi^2 \sim 0.03$. In fact, we find, for example, that $x \approx -0.015$ for four-link vector ratios, $x \approx -0.007$ for four-link tensors, and $x \approx -0.06$ for three-link scalars. These are of the expected size or somewhat smaller. For ratios involving smaller numbers of links, which lie closer to unity, we see that the slopes, x, have yet smaller magnitudes. For example, the slope of the two-link tensor ratios are almost

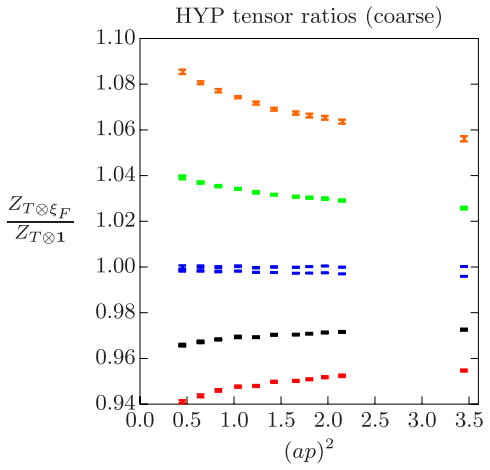

FIG. 6 (color online). As for Fig. 5 but for the tensor ratios.

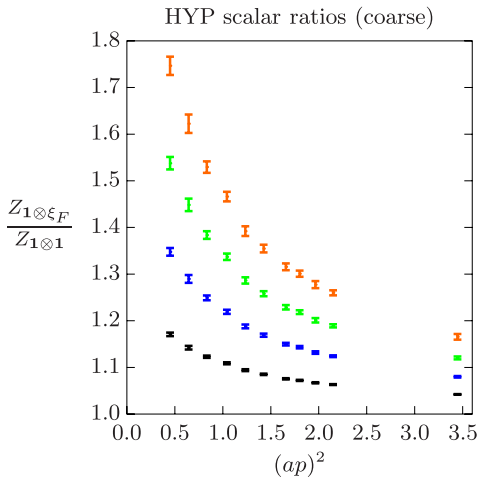

FIG. 7 (color online). As for Fig. 5 but for the scalar ratios.

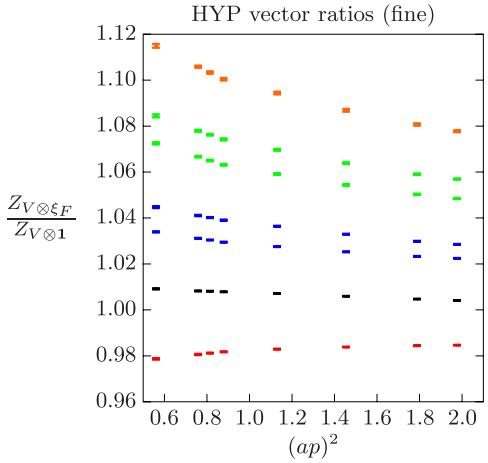

FIG. 8 (color online). Vector ratios vs. $(ap)^2$ for HYP-smeared bilinears on the fine lattices. Note that the results at $(ap)^2 = 0.81$ are the same as those in Fig. 4. Notation as in Fig. 5.

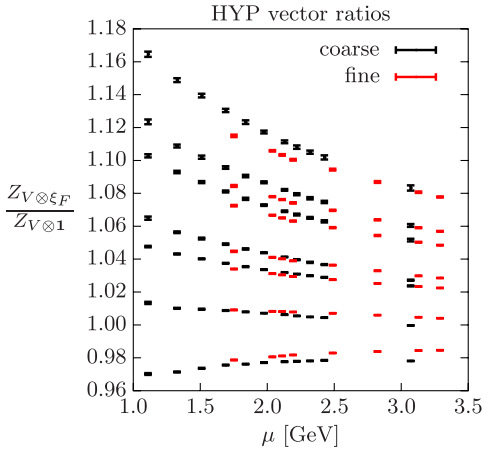

FIG. 9 (color online). Comparison of vector ratios for HYP-smeared bilinears from coarse (black) and fine (red) lattices, plotted against $\mu = |p|$. The coarse results are the same as those presented in Fig. 5. The link numbers for the coarse results can be determined by referring to the latter plot; those for the fine results are the same as for the nearest coarse points.

zero. This suggests that there is an additional suppression arising from a cancellation of discretization effects which follows approximately that for the perturbative corrections.

The corresponding plots for the fine lattices are qualitatively similar and, for the sake of brevity, we display only the results for the vector bilinears [Fig. 8]. Note that the range of $(ap)^2$ that is covered is smaller than on the coarse lattices. Since the ratio of squared lattice spacings is $(a_{\text{coarse}}/a_{\text{fine}})^2 \approx 2$, the lower edge of the NPR window should be halved compared to the coarse lattices (since it is set by a physical momentum). Thus the lower edge for V, A and T ratios should move from $(ap)^2 \approx 1$ to $(ap)^2 \approx 0.5$, as a result of which the entire momentum range shown in Fig. 8 should lie in the window. This is consistent with our results, which are approximately linear across the figure. The same is true for the tensor ratios, while for the scalars the lower edge of the window must be moved up.

Comparing Figs. 5 and 8, we see that, aside from the one-link (black) points, all the ratios move towards unity as one goes from the coarse to the fine lattices at a fixed value of $(ap)^2$. The one-link points start very close to unity and remain there. This same "collapse towards unity" occurs for the tensor and scalar ratios. This is qualitatively what we expect, because discretization errors should be similar for both lattice spacings at fixed $(ap)^2$, while nonperturbative $1/p^n$ effects should be small as we are in the NPR window, and so the change in the ratios should (if one-loop PT is reasonably accurate) fall-like $\alpha(\mu_0)$ with $\mu_0 \sim 1/a$. In fact, it may be that the discretization errors scale approximately in this fashion too.

Pursuing this a little more quantitatively, we find that the values of the slopes x are approximately the same for corresponding quantities at the two lattice spacings. This holds for all the ratios. The uncertainties in our estimates are large enough to accommodate a possible factor of $\alpha(1/a)$ reduction in slope for the fine lattices, but we do not claim to have found such a reduction. The approximate equality of slopes implies that the values after extrapolation to $ap = 0$ on the fine lattices remain closer to unity than the corresponding values on the coarse lattices. This is what one expects from the perturbative prediction.

Finally, we show, in Figs. 9–11, a direct comparison of the results for the ratios at the two lattice spacings. To make the plots readable, we plot versus $\mu \equiv |p|$ rather than $(ap)^2$. This prevents the points from overlapping and distributes them more evenly in the horizontal direction. The coarse results are identical to those in Figs. 5–7, respectively, except that the color coding is no longer used. A disadvantage of this presentation is that the discretization errors are, at fixed μ, roughly half as large for the fine lattice points as for the coarse points. An advantage is that we expect nonperturbative effects to be similar. Thus one cannot, from these plots alone, easily disentangle

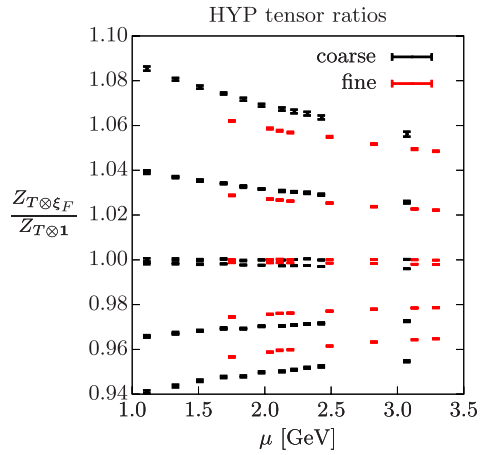

FIG. 10 (color online). As for Fig. 9 except for tensor ratios.

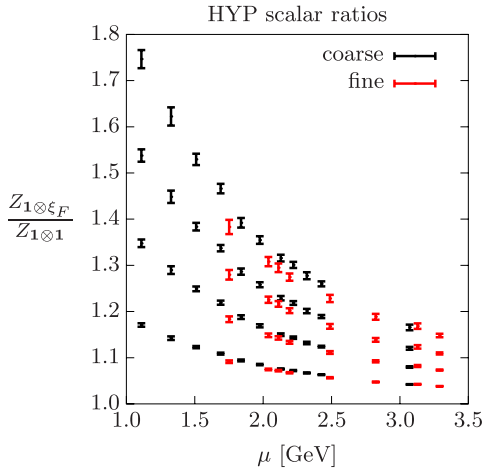

FIG. 11 (color online). As for Fig. 9 except for scalar ratios.

the perturbative, discretization and nonperturbative contributions. Nevertheless, one does see the general trend noted above that the ratios move towards unity on the fine lattices.

B. Ratios for asqtad bilinears

We now turn to the asqtad bilinears, for which the results turn out to be less well represented by PT, and harder to understand. We begin with plots of all ratios at our canonical momentum on the coarse and fine lattices, Figs. 12 and 13. Mean-field improvement is necessary to obtain even reasonably accurate predictions, so we show only the corresponding results. The greater importance of mean-field improvement for asqtad fermions is

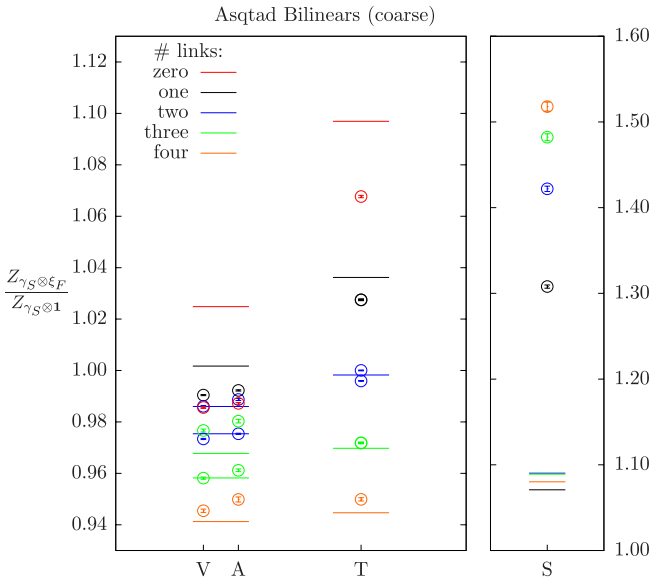

FIG. 12 (color online). As for Fig. 3 but for asqtad fermions, and with only mean-field improved perturbative predictions shown.

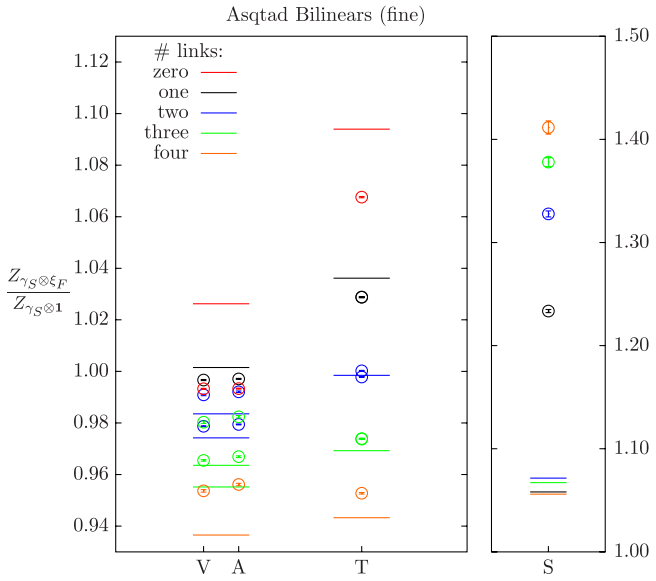

FIG. 13 (color online). As for Fig. 12, but for the fine lattices.

related to the result that the corresponding fat links ("Fat-7 + Lepage") have traces that are significantly further from unity than the HYP-smeared links indicating larger fluctuations. For example, on the coarse lattices, $\tilde{u}_0^{\mathrm{ASQ}} = 1.053$ to be compared to $\tilde{u}_0^{\mathrm{HYP}} = 0.984$.

The asqtad results differ in several noteworthy ways from those with HYP-smeared fermions. First, the ordering of the tensor bilinears by link number is reversed. This is also true for the two-, three- and four-link V and A spins. This can be qualitatively understood as follows. The average of the smeared link in the asqtad operators is larger than unity.[14] Bilinear matrix elements are thus expected to grow with the number of links, leading to Z factors that must decrease to compensate. This is the same argument used above for the HYP-smeared bilinears, except in that case it leads to the opposite ordering because $\tilde{u}_0^{\mathrm{HYP}} < 1$. Here, however, the argument fails for the scalars, which have the same ordering as for HYP fermions (although they are, in relative terms, more bunched together). Of course this argument is naive, as there are correlations between fluctuations in the links, something that is approximately accounted for by PT. Indeed, mean-field improved PT does predict the observed ordering for tensor ratios and the two- to four-link V and A spins. Nevertheless, the gross structure is reproduced in its entirety only for the tensor ratios, with the predictions for the scalars simply being poor. This situation is not improved by changes to the value of α.

On the positive side we note that the all-orders predictions of degeneracy are borne out at about the same level as for HYP-smeared operators. In addition, the quantitative disagreements with PT for the V, A and T ratios are at the

[14]This is possible because the Fat-7 + Lepage links are linear combinations of different paths.

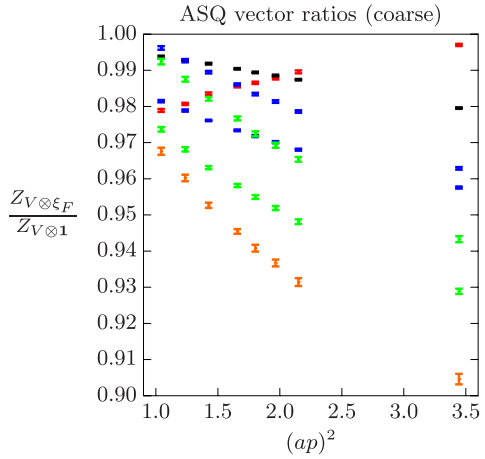

FIG. 14 (color online). Scale dependence of asqtad vector ratios on coarse MILC lattices. Notation as in Fig. 5.

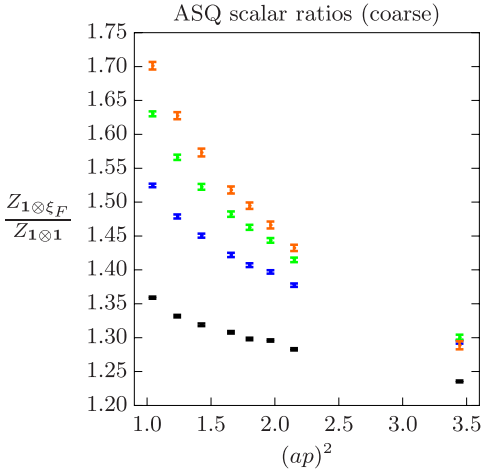

FIG. 16 (color online). As for Fig. 14 but for scalar ratios.

few percent level, which could be understood as generic $\mathcal{O}(\alpha^2)$ effects.

Finally, we comment on the changes as one goes from coarse to fine lattices, which exhibit a more complicated pattern than for the HYP-smeared operators. The NPR results for the scalar, vector and axial ratios do move towards unity as a decreases (as in the HYP case), but the tensor ratios are almost unchanged. The perturbative predictions for scalar ratios move toward unity, while those for vector and axial ratios are almost unchanged, and the predictions for tensors move slightly *away* from unity. This complicated pattern of perturbative predictions is due to the use of mean-field improvement and the fact that the nonperturbative value of $\tilde{u}_0^{\mathrm{ASQ}}$ drops by a smaller factor from coarse to fine lattices than is predicted by perturbation theory. For example, evaluating α at the scale $1/a$, Eq. (D12) predicts 1.12 and 1.10 for coarse and fine lattices, respectively, to be compared to the measured values 1.053 and 1.051.

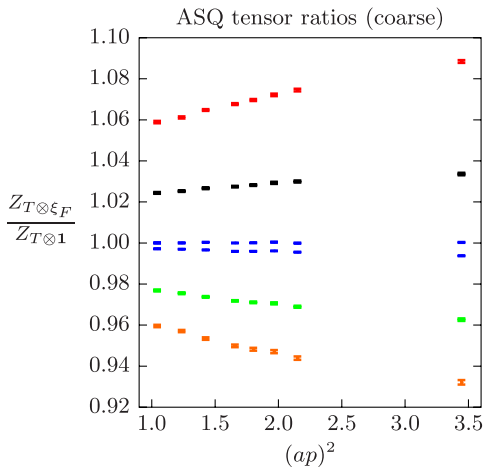

FIG. 15 (color online). As for Fig. 14 but for tensor ratios.

The momentum dependence of the asqtad ratios on the coarse lattices are shown in Figs. 14–16, with one example (the tensors) of the corresponding behavior on the fine lattices shown in Fig. 17. Note that on the coarse lattices the range of $(ap)^2$ is smaller than in the corresponding HYP plots, so any curvature due to nonperturbative effects will be harder to see. This explains why the curves for V and T spins appear more linear. For these cases the NPR window appears to cover the entire range of our data (consistent with the results from HYP-smeared bilinears), while for the scalars the lower cutoff again needs to be moved up to $(ap)^2 \sim 1.5$ on the coarse lattices.

We first discuss the vector and tensor ratios. Although the plots look superficially different from those with HYP-smeared fermions, we note that all the slopes have the same signs in the two cases (comparing data with the same number of links), and indeed have the same ordering of

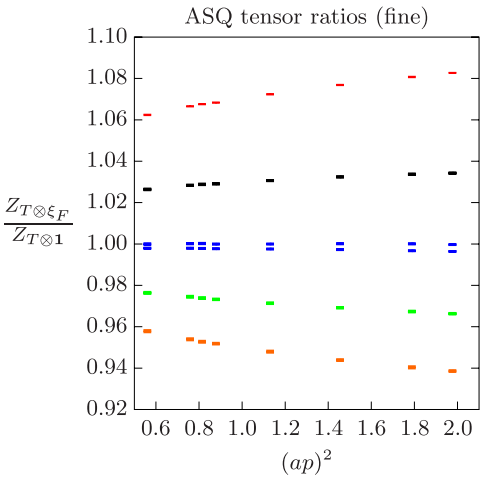

FIG. 17 (color online). Tensor ratios vs. $(ap)^2$ for HYP-smeared bilinears on the fine lattices. Note that the results at $(ap)^2 = 0.81$ are the same as those in Fig. 13. Notation as in Fig. 15.

magnitudes. The only change in the discretization effects between HYP and asqtad cases is that the slope coefficients x are about twice as large for asqtad bilinears. This is consistent with the general experience that HYP smearing leads to smaller discretization effects. We also find that, as for HYP-smeared bilinears, the slopes at the two lattice spacings are similar.

For tensor ratios, the ordering seen in Fig. 12 remains valid over our entire momentum range, and also after extrapolation to $ap = 0$. Thus by a small rescaling of α one can retain quantitative agreement with one-loop PT at $ap = 0$. On the other hand, extrapolating the fine lattice results to $ap = 0$ leads to values which lie a little further from unity (cf. Figs. 15 and 17), which is not consistent with PT.

For the vector ratios extrapolation to $ap = 0$ reshuffles the ordering, with the zero-link ratio now having the smallest Z factor. Thus the perturbative predictions of Fig. 12 become worse for the vectors after extrapolation, even after possible rescalings of α.

The ap dependence of the scalar ratios, by contrast, is similar to that for the HYP-smeared bilinears. The ordering is maintained by extrapolations to $ap = 0$, with slope coefficients that are similar (not differing by a factor of two). However, the already very poor perturbative predictions become even worse after the extrapolations.

A perplexing feature of the results for momentum dependence is that, at the highest values of $(ap)^2$, both vector and scalar ratios become much closer to the perturbative predictions, particularly in terms of the ordering and relative splittings. We do not understand why this should be.

For the sake of brevity, we do not show the direct comparisons of coarse and fine asqtad ratios versus $|p|$. These plots are both messy and hard to interpret, adding little to the preceding discussion.

In summary, one-loop PT fails to provide even a qualitative description of many of the features observed for the asqtad bilinears, with the exception of the tensor ratios. One should keep in mind, however, that the disagreements with the vectors are well within the expected size of generic two-loop contributions.

C. Results for denominators

Finally we turn to a discussion of the denominators in the ratios, namely the matching factors $Z_{1\otimes 1}$, $Z_{\gamma_\mu\otimes 1}$ and $Z_{\gamma_\mu\gamma_\nu\otimes 1}$ which we label simply Z_S, Z_V and Z_T, respectively. Unlike the ratios, these quantities have anomalous dimensions (even Z_V in the RI′ scheme), so that they do depend on $\mu = |p|$ even in the absence of nonperturbative effects and discretization errors. This dependence is described in perturbation theory by the result Eq. (D6), the ingredients for which are collected in appendices D and E. In brief, one runs in the continuum (in the RI′ scheme) from μ to $\mu_0 \sim 1/a$, and then matches to the lattice

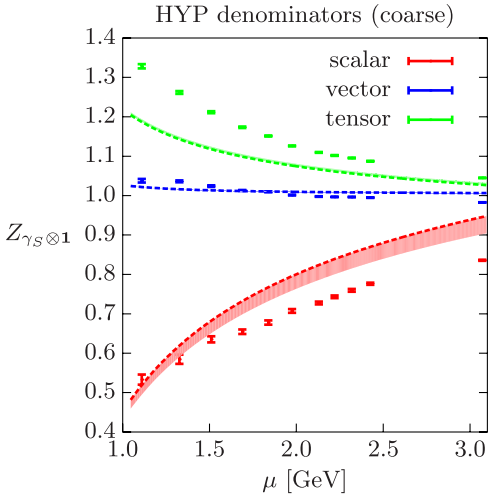

FIG. 18 (color online). Comparison of scale dependence of the HYP-smeared taste-singlet scalar, vector, and tensor Z factors computed nonperturbatively on the coarse lattices to the perturbative prediction described in the text. The colored bands give the variation in the perturbative prediction arising from varying the intermediate matching scale between $1/a$ (dotted line) and $2/a$.

scheme at that scale. This matching is done, for technical reasons, using the $\overline{\text{MS}}$ scheme as an intermediate step.

As above, we consider first the HYP-smeared bilinears. Results from the coarse and fine lattices are shown in Figs. 18 and 19, respectively. Note again that the range of μ differs in the two cases. For the perturbative results, we use the non-mean-field improved result (which lies very close to the mean-field improved result), and display a band to give an indication of the uncertainty due to truncation errors. This is obtained by varying μ_0 between $1/a$ and $2/a$, a range for which $\Delta\alpha \approx \alpha^2$. We stress that the weakest link in the perturbative result is the one-loop matching between the lattice and $\overline{\text{MS}}$ schemes; all other running or matching is done at three- or four-loop order. We also note that, as for the Z-factor ratios, this estimate of truncation errors is not the most conservative when Z is close to unity, because there can be generic $\mathcal{O}(\alpha^2)$ terms of size 5% to 9%.

The figures show good qualitative agreement between the NPR and PT results in all three channels. The ordering is correct and the μ dependence is reasonably well predicted. Quantitatively the perturbative prediction undershoots the separation from unity for Z_T and Z_V, even allowing for the predicted uncertainty band. This mismatch is small enough, however, that it could be due to generic two-loop contributions. The level of quantitative agreement is somewhat worse than that found above for the ratios: for these, PT could reproduce all the vector and tensor ratios with choices of μ_0 lying in the range $1/a - 2/a$.

Unlike the ratios, the Z factors do not themselves have a good continuum limit, due to the nonvanishing anomalous

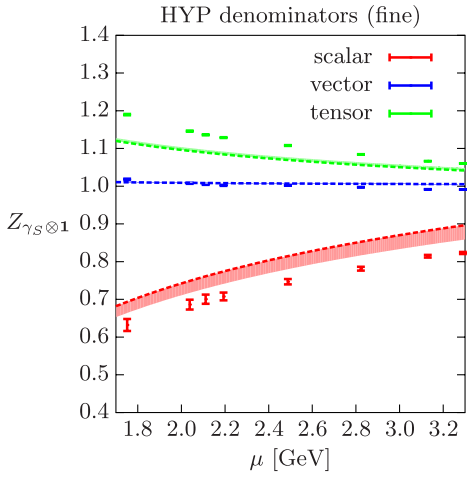

FIG. 19 (color online). As for Fig. 18 but on the fine lattices.

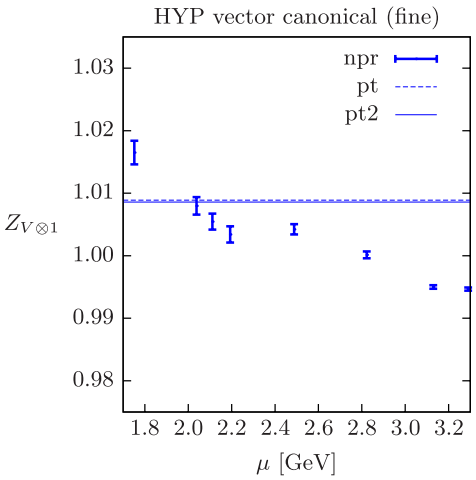

FIG. 21 (color online). As for Fig. 20 but for Z_V.

dimensions. To take a continuum limit one must multiply them by hadronic matrix elements of the corresponding bilinears, which we do not have available here. Because of this, there is no general expectation that results from the fine lattices should lie closer to unity than those from the coarse lattices, even ignoring discretization errors. What one might expect, however, is that the perturbative prediction should become more accurate, since the intermediate matching scale μ_0 is higher. We do in fact see a small improvement between Fig. 18 and 19.

In order to disentangle the predicted running with μ from discretization effects, we can run our results in the RI′ scheme from μ to a canonical scale which we choose to be 2 GeV. This running is done at three- or four-loop order using continuum anomalous dimensions (see Appendix E). The hope is that the data will significantly "flatten," leaving a residual $(a\mu)^2$ dependence. In Figs. 20–22, we show results after this running for the fine lattices. Results on the coarse lattices are similar and are not shown. The perturbative predictions are obtained as above, but with p replaced by 2 GeV. We recall that the NPR window covers the range of results shown in these plots.

For Z_S, we see that the flattening is successful, although the perturbative prediction for the absolute value misses the data. Nevertheless, a generic two-loop term would be sufficient to make up the gap. Z_V varies by $\approx$ 1–2% over the momentum range shown, and, if extrapolated to $a\mu = 0$, will lie quite close to the perturbative prediction. Z_T varies more rapidly, and, if extrapolated linearly in $(ap)^2$ to $a = 0$, will become ≈ 1.16. This is $\sim$5% above the perturbative prediction, a difference which could be bridged by two-loop perturbative contributions. The slope coefficient is $x \approx -0.023$ and is comparable to that for ratios.

We now turn to the asqtad denominators, for which we show the running with NPR scale on the fine lattices in Fig. 23. Results are similar on the coarse lattices.

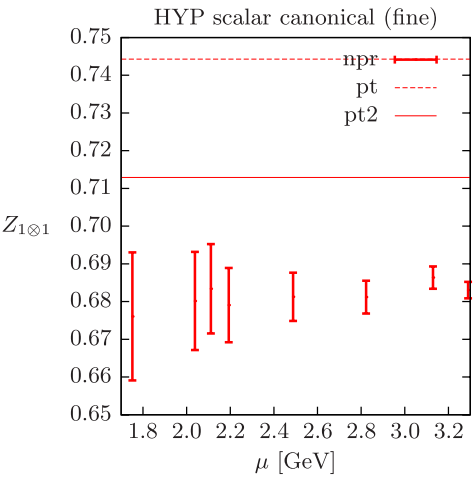

FIG. 20 (color online). Predictions for $Z_S^{\mathrm{RI}'}$ (2 GeV) with HYP-smeared fermions on fine MILC lattices. The Z factor is run from NPR scale μ to 2 GeV using continuum perturbation theory. The perturbative predictions use an intermediate conversion scale of $1/a$ ("pt") or $2/a$ ("pt2").

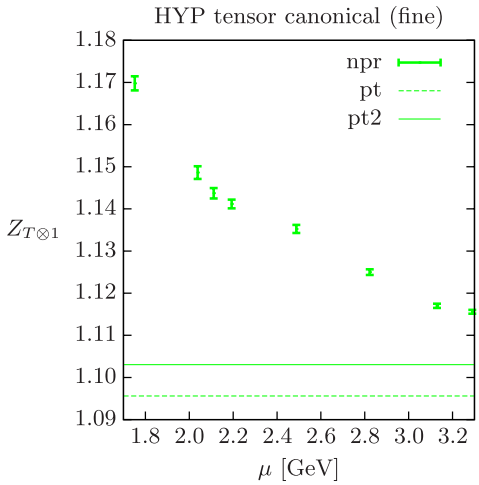

FIG. 22 (color online). As for Fig. 20 but for Z_T.

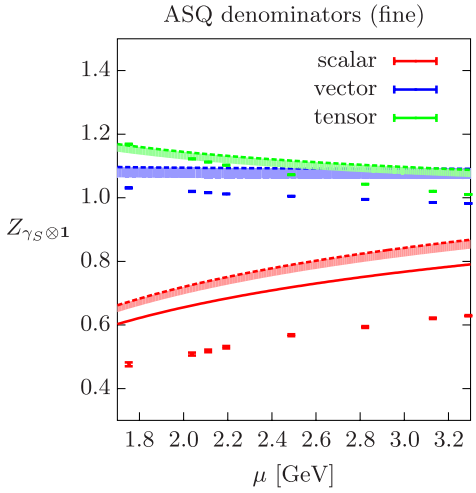

FIG. 23 (color online). Comparison of the asqtad Z_S, Z_V and Z_T on fine lattices with mean-field improved PT. Details as in Fig. 19, except that the two-loop prediction for Z_S (red solid line) is also shown.

We compare here to mean-field improved perturbation theory, since without mean-field improvement the asqtad ratios are poorly represented, as discussed above. We note that mean-field improvement impacts the predictions for Z_V and Z_T, but not that for Z_S. For Z_S, we also show the perturbative result including the two-loop lattice to $\overline{\text{MS}}$ matching factor from Ref. [15]. This has a much weaker dependence on the intermediate matching scale than that using one-loop matching, and we show the result only for intermediate scale $1/a$. We stress that one cannot directly

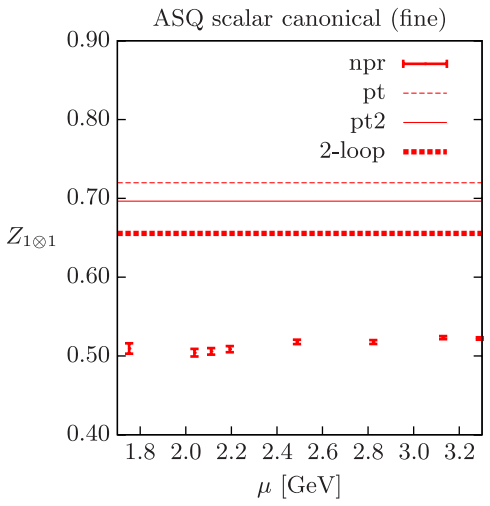

FIG. 24 (color online). Predictions for $Z_S^{\text{RI}'}$ (2 GeV) with asqtad fermions on fine MILC lattices. The Z factor is run from NPR scale μ to 2 GeV using continuum perturbation theory. The one-loop perturbative predictions use an intermediate conversion scale of $1/a$ ("pt") or $2/a$ ("pt2"). Also shown is the two-loop perturbative prediction (with intermediate scale $1/a$).

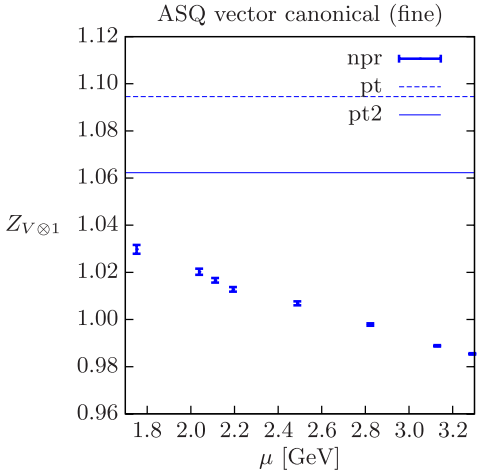

FIG. 25 (color online). As for Fig. 24 except for Z_V.

gauge the rate of convergence of the perturbative series from a comparison of "one-loop" and "two-loop" results, since both are composed of several components, some of which are being evaluated at three- or four-loop order [see Eq. (D6)]. What one can see, however, is that shift between "one-loop" and "two-loop" results is of the ~5% size expected of a generic two-loop term on the fine lattices. Compared to the corresponding HYP-smeared results (Fig. 19), we observe that the NPR result for Z_S is much further from unity, and also further from the perturbative predictions.

We learn more from the results after flattening, shown in Figs. 24–26. For Z_S, it is striking that (as for the HYP bilinears), the results show little indication of $(a\mu)^2$ effects, indicating that the four-loop anomalous dimension is giving a good representation of the μ dependence. On the other hand, the value itself lies ~0.2 below the "one-loop" and ~0.15 below the "two-loop" predictions, indicating a failure of convergence since this gap is too large to be bridged by a generic $\mathcal{O}(\alpha^3)$ term. We note that

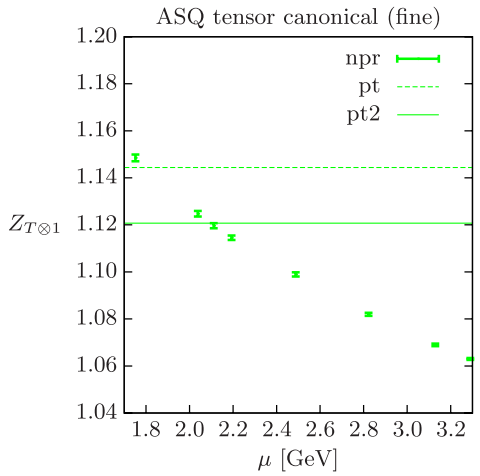

FIG. 26 (color online). As for Fig. 24 except for Z_T.

this gap is the reason why, as described in the Introduction, the value of m_s obtained from our NPR results lies significantly above that obtained using two-loop matching. Specifically, on the fine lattices, NPR yields $m_s = 105$ MeV [13] while the two-loop result is 86 MeV [15].

The situation is much better for Z_V and Z_T. For Z_V there is a mild μ dependence, which brings the result at $a\mu = 0$ close to the one-loop prediction. For Z_T, the μ dependence is somewhat stronger and leads to a value at $a\mu = 0$ of $Z_T(a\mu = 0) \approx 1.18$, within 5% of the perturbative prediction. In both cases, the gap can be bridged by a generic two-loop contribution.

V. CONCLUSIONS

We have implemented nonperturbative renormalization for general staggered-fermion bilinear operators, using a method that is consistent with the symmetries of the staggered action. We have shown how those symmetries constrain the propagator and vertex functions to have the expected continuum forms at leading nontrivial order in an expansion in the lattice spacing. We have also introduced "covariant bilinears," which transform irreducibly under the lattice symmetries and thus do not mix, unlike the traditional "hypercube bilinears."

We have calculated Z factors for 30 different operators having spins V, A, T and S. It is well known that, for unimproved staggered fermions, many Z factors, particularly those for scalars, lie very far from unity and have perturbative expansions which are not convergent [3,36]. We have rechecked this result ourselves. It is also well known that these problems can be substantially improved using smeared lattice links and other forms of action improvement. Here we have used HYP-smeared and asqtad fermions. By studying many operators we are able to give a general judgement on the utility of perturbation theory for these two types of fermion. A useful tool in this regard is the use of ratios for which the overall running due to anomalous dimensions cancels, allowing a study of the approach to the continuum limit.

Overall, we find that the HYP-smeared Z factors lie relatively close to unity and can be predicted by one-loop PT as long as one includes a generic uncertainty of relative size $\mathcal{O}(1) \times \alpha(1/a)^2$. This holds both for ratios and for the Z factors themselves. In fact, PT works more accurately than this for the vector and tensor ratios, with an uncertainty given by the square of the one-loop term sufficing. The detailed ordering of these ratios is predicted very well. We also find that discretization errors proportional to $(ap)^2$ are of the expected size or smaller.

For the asqtad bilinears, one-loop PT is less successful. Only for the tensor ratios does it approach the efficacy observed in the HYP-smeared case, while for the scalars there appears to be a breakdown in convergence.

These results have implications for extracting physical predictions from staggered simulations. The recent calculation of B_K using HYP-smeared fermions used one-loop perturbative results for the needed Z factors [16,17]. The anomalous dimension of the operator which appears is roughly comparable to that for the tensor bilinear, and thus we can use the latter as a guide to how well one-loop PT reproduces the Z factor obtained using NPR. We find in this case (see, e.g., Figs. 18 and 19) that one-loop PT gives a good estimate as long as one uses an error estimate of $\mathcal{O}(1) \times \alpha^2$. This is, in fact, the estimate used in Refs. [16,17].

As for quark masses obtained using one- or two-loop perturbative matching, the results of Figs. 23 and 24 show that there is a substantial gap between the perturbative and nonperturbative results for $Z_S = 1/Z_m$ with asqtad fermions. This gap is larger than a straightforward estimate of the truncation error. This suggests that the systematic error in the quark masses obtained in Refs. [8,14,15] may be larger than previously estimated. To study this point further, it will be important to use NPR with nonexceptional momenta [28].

Finally, we note that present large-scale simulations with staggered fermions now use HISQ rather than asqtad quarks. HISQ quarks combine the advantages of HYP smearing with the full $\mathcal{O}(a^2)$ improvement of asqtad quarks (and in addition reduce discretization errors for heavier quarks) [9]. Thus we expect the success of PT in describing Z factors for HYP-smeared operators to carry over to operators composed of HISQ quarks.

ACKNOWLEDGMENTS

Computations for this work were carried out on USQCD Collaboration clusters at Fermilab. The USQCD Collaboration is funded by the Office of Science of the U.S. Department of Energy. The work of A. L. was supported in part by STFC Grants No. ST/G000557/1 and No. ST/J000396/1. The work of S. S. is supported in part by the U.S. Department of Energy Grant No. DE-FG02-96ER40956.

APPENDIX A: NOTATION AND CONVENTIONS

1. Staggered matrix conventions

We use the notation of Refs. [32,33], which introduce two sets of matrices unitarily equivalent to the general spin-taste matrices. A basis for the latter is $(\gamma_S \otimes \xi_F)$, where a general spin matrix is labeled by the hypercube vector S,

$$\gamma_S = \gamma_1^{S_1}\gamma_2^{S_2}\gamma_3^{S_3}\gamma_4^{S_4}, \tag{A1}$$

while the general taste matrix is labeled by another such vector F,

$$\xi_F = \xi_1^{F_1}\xi_2^{F_2}\xi_3^{F_3}\xi_4^{F_4}, \tag{A2}$$

with $\xi_\mu = \gamma_\mu^*$. The two unitarily equivalent sets of matrices are then

$$\overline{(\gamma_S \otimes \xi_F)}_{AB} \equiv \frac{1}{4}\,\mathrm{Tr}[\gamma_A^\dagger \gamma_S \gamma_B \gamma_F^\dagger], \tag{A3}$$

$$\overline{\overline{(\gamma_S \otimes \xi_F)}}_{AB} \equiv \sum_{CD} \frac{(-)^{A\cdot C}}{4} \overline{(\gamma_S \otimes \xi_F)}_{CD} \frac{(-)^{D\cdot B}}{4}. \tag{A4}$$

Using these relations one can trace the connection between the 16^2 choices for the indices AB in the propagator (10) to the spin and taste indices in the more familiar form $(\gamma_S \otimes \xi_F)$.

2. Definition of the asqtad action

The asqtad action is [4–6]

$$S_{\mathrm{asqtad}} = \sum_n \Big[\bar\chi(n) \sum_\mu \eta_\mu(n) \Big(\nabla_\mu^{\mathrm{F7L}} \chi(n) + \frac{1}{8}[\nabla_\mu^{\mathrm{T1}} - \nabla_\mu^{\mathrm{T3}}]\chi(n) \Big) + (m/u_0)\bar\chi(n)\chi(n) \Big], \tag{A5}$$

$$\nabla_\mu^{\mathrm{F7L}} \chi(n) = \frac{1}{2}[W_\mu(n)\chi(n+\hat\mu) - W_\mu^\dagger(n-\hat\mu)\chi(n-\hat\mu)], \tag{A6}$$

$$\nabla_\mu^{\mathrm{T1}} \chi(n) = \frac{1}{2u_0}[U_\mu(n)\chi(n+\hat\mu) - U_\mu^\dagger(n-\hat\mu)\chi(n-\hat\mu)], \tag{A7}$$

$$\nabla_\mu^{\mathrm{T3}} \chi(n) = \frac{1}{6u_0^3}[U(n, n+3\hat\mu)\chi(n+3\hat\mu) - U(n, n-3\hat\mu)\chi(n-3\hat\mu)], \tag{A8}$$

where $W_\mu(n)$ is a smeared link constructed using the Fat7 blocking transformation [4,5] combined with Lepage's prescription [6] and tadpole improvement [41], and $U(n, n \pm 3\hat\mu)$ are products of three thin links in the μ direction starting at position n. Finally, u_0 is the tadpole improvement factor [41], which we take to be the fourth root of the average plaquette.

APPENDIX B: IRREDUCIBLE REPRESENTATIONS FOR COVARIANT BILINEARS

In this appendix we sketch the demonstration that the covariant bilinears $\mathcal{O}^{\mathrm{cov}}_{S\otimes F}$ of Eq. (21) fall into the irreps listed in Table I under the lattice symmetry group. Although this result is likely known to workers in the field, we have not found a demonstration in the literature. In particular, in their seminal work on staggered fermions, Golterman and Smit described the full lattice group [30], but focused on constructing operators transforming as irreps of the smaller time slice group, which classifies eigenstates of the transfer matrix [38,42]. Verstegen subsequently classified the irreps of bilinears living on a single 2^4 hypercube [43]. The symmetry group in this case is smaller than that for the zero momentum covariant bilinears, since translations are excluded. Thus, although Verstegen's work will be useful in the following, the irreps he finds are in general smaller than those for covariant bilinears.

Perturbative calculations of matching factors also give information on the irreps, since operators living in different irreps have different matching factors. This information is, however, incomplete since results are available only at finite order (usually one-loop), and differences could show up at higher order.

In the subsequent discussion we use the presentation of the lattice group and method of analysis (as well as the notation) of Ref. [44]. We refer to this reference for most of the technical details. An alternative approach is that of Ref. [45].

For operators having zero physical momentum, the group of transformations is

$$\mathcal{G}_0 = \Gamma_{4,1} \rtimes W_4, \qquad \Gamma_{4,1} = \{\Xi_\mu, C_0\}, \qquad W_4 = \{R_{\mu\nu}, I_s\}. \tag{B1}$$

Here W_4 is the hypercubic group generated by rotations $R_{\mu\nu}$ and spatial inversion I_s, while $\Gamma_{4,1}$ is the Clifford group in 5-dimensional Minkowski space generated by lattice charge conjugation, C_0, and single-site translations Ξ_μ.[15] The symbol "$\rtimes$" indicates a semidirect product. In the analysis of Verstegen, the translations Ξ_μ are absent, leaving only the group W_4 combined with C_0. Thus the constraints he finds are weaker than those obtained from $\mathcal{G}_0$.

Under translations the covariant bilinears pick up a sign $(-)^{\tilde F_\mu}$, where $\tilde F_\mu = \sum_{\nu\neq\mu} F_\nu$. This is shown in the following appendix. Similarly the bilinears have a definite parity under C_0 (which is straightforward to calculate but not needed in the following). Thus the bilinears reside in 1-dimensional irreps of $\Gamma_{4,1}$ characterized by five parities. In Ref. [44] these are called

$$\Delta^{(4,1)}(\pm, \pm, \pm, \pm, \xi_C), \tag{B2}$$

where the first four arguments are the translation signs under Ξ_1, Ξ_2, Ξ_3 and Ξ_4, while the last is the parity under C_0. Since $\mathcal{G}_0$ is a semidirect product, one must, for each irrep $\Delta^{(4,1)}$, find the subgroup of W_4 which leaves the irrep invariant. The bilinears are then classified into irreps of this "little group." These induce representations of the full group that are known to be irreducible.

[15] In general Ξ_μ are single-site translations with the momentum factor $e^{ip'_\mu}$ removed, but this removal is not needed as $p' = 0$.

The action of the rotations and spatial inversion which form W_4 is discussed in the next appendix. All we need to know here is that both transformations act simultaneously on spin and taste indices. Thus $O^{\rm cov}_{S\otimes F}$ is transformed, up to a sign, into $O^{\rm cov}_{S_R\otimes F_R}$, where S_R and F_R are the hypercube vectors obtained from S by F by the transformation under consideration.

We now begin the classification of bilinears into irreps. For taste-singlet bilinears, the $\Gamma_{4,1}$ irrep is $\Delta^{(4,1)}(+,+,+,+,\xi_C)$ and the little group is the full W_4 [44]. The same little group holds for taste ξ_5 [$F=(1111)$] for which the irrep is $\Delta^{(4,1)}(-,-,-,-,\xi_C)$. In both cases we can use the analysis of Verstegen, who shows (see his Table III for irreps of the rotation subgroup, together with the discussion in his Sec. V of how inversion combines irreps) that each of the five types of spin lives in a single irrep.[16] Explicitly, the irreps for taste singlets have spin tastes

$$(I\otimes I),\qquad (\gamma_\mu\otimes I),\qquad (\gamma_{\mu\nu}\otimes I),\qquad (\gamma_{\mu5}\otimes I)\quad\&\quad(\gamma_5\otimes I). \tag{B3}$$

Here μ and ν run from 1 to 4 except that $\mu<\nu$. These are the five taste-singlet irreps appearing in Table I. The same five spins apply also to taste ξ_5.

Next we consider bilinears with taste ξ_μ and $\xi_{\mu5}$. Choosing $\mu=4$ for definiteness, the $\Gamma_{4,1}$ irreps are

$$\Delta^{(4,1)}(-,-,-,+,\xi_C)\quad\text{and}\quad\Delta^{(4,1)}(+,+,+,-,\xi_C), \tag{B4}$$

respectively. In both cases the little group is $W_3\times Z_2$, with W_3 the cubic group $\{R_{ij},I_s\}$ while Z_2 is generated by the axis inversion symmetry in the fourth direction [44]. Determining the transformations under rotations, one finds that the following spins live in 3-dimensional irreps of W_3 (either the 1 or the $\bar{1}$ in the notation of Ref. [46]):

$$\gamma_j,\qquad \gamma_{j5},\qquad \gamma_{j4},\quad\text{and}\quad \epsilon_{jkl}\gamma_{kl}\qquad (j,k,l=1\text{–}3). \tag{B5}$$

The remaining spins (I, γ_5, γ_4 and γ_{45}) live in one of the two 1-dimensional irreps. Extending these irreps of the little group to the full group by acting with the "missing" generators, i.e. R_{4k}, their size is multiplied by a factor of 4, the dimension of the orbit of the $\Gamma_{4,1}$ irrep under W_4. Thus (choosing taste vector for definiteness) one ends up with four 12-dimensional irreps and four 4-dimensional irreps:

$$(\gamma_\mu\otimes\xi_\nu),\quad(\gamma_{\mu5}\otimes\xi_\nu),\quad(\gamma_{\mu\nu}\otimes\xi_\nu),\quad(\gamma_{\mu\rho}\otimes\xi_\nu),\quad(I\otimes\xi_\nu),\quad(\gamma_5\otimes\xi_\nu),\quad(\gamma_\nu\otimes\xi_\nu),\quad(\gamma_{\nu5}\otimes\xi_\nu), \tag{B6}$$

where $\mu\neq\nu$, $\rho\neq\nu$ and $\mu<\rho$. These are the eight taste-vector irreps appearing in Table I. The same set of spins appear for the axial taste bilinears (with $\xi_\mu\to\xi_{\mu5}$).

In this case, the results differ from those obtained for single-hypercube bilinears. For example, Verstegen finds that the spin-scalar, taste-vector bilinears split into two irreps, a 1-dimensional irrep $\sum_\mu(I\otimes\xi_\mu)$ and a 3-dimensional irrep consisting of the differences $(I\otimes\xi_\mu)-(I\otimes\xi_\nu)$. For covariant bilinears, by contrast, one has a single 4-dimensional irrep, $(I\otimes\xi_\mu)$.

Finally, we consider the taste tensors. If the taste is ξ_{12}, the $\Gamma_{4,1}$ irrep is $\Delta^{(4,1)}(-,-,+,+,\xi_C)$. The little group is $D_4\otimes D_4$, where the first dihedral group D_4 is generated by R_{12} and I_1 (the axis inversion operator in the first direction), while the second D_4 is generated by R_{34} and I_3. D_4 has four 1-dimensional and one 2-dimensional irreps. The bilinears decompose into a single 4-dimensional irrep of $D_4\times D_4$ (spin $\gamma_j\gamma_k$, with $j=1,2$ and $k=3,4$), four 2-dimensional irreps (spins γ_j, γ_{j5}, γ_k and γ_{k5}), and four 1-dimensional irreps (spins I, γ_5, γ_{12} and γ_{34}). The orbit in this case is six dimensional, so the induced irreps of $\mathcal{G}_0$ are the 24-dimensional

$$(\gamma_{\mu\rho}\otimes\xi_{\mu\nu}), \tag{B7}$$

with $\mu<\nu$, $\rho\neq\mu$ and $\rho\neq\nu$, the four 12-dimensional irreps

$$(\gamma_\mu\otimes\xi_{\mu\nu})\quad(\gamma_{\mu5}\otimes\xi_{\mu\nu})\quad(\gamma_\rho\otimes\xi_{\mu\nu})\quad\&\quad(\gamma_{\rho5}\otimes\xi_{mu\nu}), \tag{B8}$$

and the four 6-dimensional irreps,

$$(I\otimes\xi_{\mu\nu})\quad(\gamma_5\otimes\xi_{\mu\nu})\quad(\gamma_{\mu\nu}\otimes\xi_{\mu\nu})\quad\&\quad(\gamma_{\rho\sigma}\otimes\xi_{\mu\nu}), \tag{B9}$$

with indices constrained as above together with σ differing from μ, ν and ρ. Altogether, these are the nine taste-tensor irreps appearing in Table I.

APPENDIX C: SYMMETRY CONSTRAINTS ON PROPAGATOR AND VERTICES

In this appendix we describe how lattice translation symmetry constrains the form of the quark propagator and the vertices of covariant bilinears.

The fermion fields transform under translations as [30]

$$\chi(n)\to\zeta_\mu(n)\chi(n+\hat{\mu})\quad\text{and}\quad\bar{\chi}(n)\to\zeta_\mu(n)\bar{\chi}(n+\hat{\mu}). \tag{C1}$$

The translation phases can be chosen to be

[16] Verstegen's rotations and inversions are about the center of the hypercube, rather than the standard choice of being about a lattice point. These choices differ, however, by translations, which, for the taste-singlet and ξ_5 operators, are simply signs and do not lead to changes in the dimensionality of the resulting irreps.

$$\zeta_\mu(n) = (-)^{\sum_{\nu>\mu} n_\nu} = (-)^{n_\zeta \cdot \hat{\mu}} = (-)^{n \cdot \hat{\mu}_\eta}, \tag{C2}$$

where

$$n_\zeta = (n_2 + n_3 + n_4, n_3 + n_4, n_4, 0) \tag{C3}$$

and

$$n_\eta = (0, n_1, n_1 + n_2, n_1 + n_2 + n_3), \tag{C4}$$

and we have used the identity

$$n_\zeta \cdot m = n \cdot m_\eta. \tag{C5}$$

Thus the momentum-space field (6) transforms as

$$\chi_A(p') \to \sum_n e^{-ip' \cdot n} (-)^{A \cdot n} (-)^{n \cdot \hat{\mu}_\eta} \chi(n + \hat{\mu}) \tag{C6}$$

$$= e^{ip'_\mu} (-)^{A_\mu} \sum_m e^{-ip' \cdot m} (-)^{m \cdot (A + \hat{\mu}_\eta)} \chi(m) \tag{C7}$$

$$= e^{ip'_\mu} (-)^{A_\mu} \delta_{A +_2 \hat{\mu}_\eta, C} \phi_C(p') \tag{C8}$$

$$= e^{ip'_\mu} \overline{\overline{(I \otimes \xi_\mu)}}_{AC} \phi_C(p'), \tag{C9}$$

where $+_2$ indicates addition mod 2. In the last step we have used

$$\overline{\overline{(I \otimes \xi_\mu)}}_{AC} = (-)^{A_\mu} \delta_{A +_2 \hat{\mu}_\eta, C}. \tag{C10}$$

One can similarly show that

$$\bar{\phi}_B(-q') \to e^{-iq'_\mu} \bar{\phi}_D(-q') \overline{\overline{(I \otimes \xi_\mu)}}_{DB}. \tag{C11}$$

These results show explicitly how translations by a single site correspond (once momentum factors are removed) to taste rotations [30].

Using the translation invariance of the action, one thus learns that the momentum space propagator (9) satisfies

$$S(p')_{AB} N_{\text{site}} \equiv \langle \phi_A(p') \bar{\phi}_B(-p') \rangle = \overline{\overline{(I \otimes \xi_\mu)}}_{AC} \langle \phi_C(p') \bar{\phi}_D(-p') \rangle \overline{\overline{(I \otimes \xi_\mu)}}_{DB} \tag{C12}$$

from which the result (11) follows. As explained in the main text, it follows that the propagator is taste singlet.

We now turn to the implications of translation invariance for the (unamputated) vertex, Eq. (22). As for the propagator, translating the external fields lead to multiplications by $\overline{\overline{(I \otimes \xi_\mu)}}$, as well as to phase factors which cancel in our kinematics. To determine the effect of translations on the bilinear operator (21) we first note that the $\bar{\chi}$ and χ fields together lead to the sign $(-)^{(S-F)_\zeta \cdot \hat{\mu}}$. Combining this with the sign resulting from translating the phases in the operator,

$$\overline{(\gamma_S \otimes \xi_F)}_{x+\hat{\mu}, x+\hat{\mu}+S-F} = (-)^{\tilde{F}_\mu} (-)^{(S-F)_\zeta \cdot \hat{\mu}} \overline{(\gamma_S \otimes \xi_F)}_{x, x+S-F}, \tag{C13}$$

we find (dropping flavor indices for clarity)

$$\mathcal{O}^{\text{cov}}_{S \otimes F} \to (-)^{\tilde{F}_\mu} \mathcal{O}^{\text{cov}}_{S \otimes F}. \tag{C14}$$

Combining these results we see that the vertex functions satisfy, for each μ,

$$\Lambda^{S \otimes F}(p') = (-)^{\tilde{F}_\mu} \overline{\overline{(I \otimes \xi_\mu)}} \Lambda^{S \otimes F}(p') \overline{\overline{(I \otimes \xi_\mu)}}. \tag{C15}$$

This implies that $\Lambda(p)$ must have taste F, because

$$\overline{\overline{(I \otimes \xi_\mu)}}\, \overline{\overline{(\gamma_S \otimes \xi_F)}}\, \overline{\overline{(I \otimes \xi_\mu)}} = (-)^{\tilde{F}_\mu} \overline{\overline{(\gamma_S \otimes \xi_F)}}. \tag{C16}$$

When we amputate the vertex using inverse propagators, which we know, from above, are taste singlets, the resulting amputated vertex will also have taste F. This shows that, if one uses covariant bilinears, there can be no mixing with other tastes.

We next discuss the constraints due to spatial inversion symmetry, I_s. This acts on the fields as

$$\chi(n) \to \eta_4(n) \chi(n_S), \qquad \bar{\chi}(n) \to \eta_4(n) \bar{\chi}(n_S), \tag{C17}$$

where $n_S = I_s^{-1} n$. By manipulations analogous to those given above, one can rewrite these transformations in terms of the momentum-space fields, finding

$$\phi(p') \to \overline{\overline{(\gamma_4 \otimes \xi_4)}} \phi(p'_S), \qquad \bar{\phi}(-p') \to \phi(-p'_S) \overline{\overline{(\gamma_4 \otimes \xi_4)}}. \tag{C18}$$

From the invariance of the action under I_s it follows that

$$S(p') = \overline{\overline{(\gamma_4 \otimes \xi_4)}} S(p'_S) \overline{\overline{(\gamma_4 \otimes \xi_4)}} \Rightarrow S^{-1}(p') = \overline{\overline{(\gamma_4 \otimes \xi_4)}} S^{-1}(p'_S) \overline{\overline{(\gamma_4 \otimes \xi_4)}}. \tag{C19}$$

The conjugation by $\overline{\overline{(\gamma_4 \otimes \xi_4)}}$ flips the sign of each spatial component of all spin and taste matrices. Since we know, however, that S^{-1} is a taste singlet, the effect of the conjugation is to replace each γ_μ with its spatial inverse. The relation (C19) thus has exactly the same implication as the corresponding result for fermions without the taste degree of freedom. As noted in the main text, combined with rotations, one finds that S^{-1} has the form given in Eq. (12). The appearance of odd powers of p' in this result is due to the spatial inversion symmetry.

For the unamputated vertex, one finds that inversion symmetry leads to

$$\Lambda_{S \otimes F}(p') = \eta_4(\Delta) \overline{\overline{(\gamma_4 \otimes \xi_4)}} \Lambda_{S \otimes F}(p'_S) \overline{\overline{(\gamma_4 \otimes \xi_4)}}, \tag{C20}$$

where $\Delta = S - F$. Multiplying from left and right with $S^{-1}(p')$ and using the relation (C19), one can convert this into a result of the same form for the amputated vertex,

$$\Gamma_{S\otimes F}(p') = \eta_4(\Delta)\overline{(\gamma_4\otimes\xi_4)}\Gamma_{S\otimes F}(p'_S)\overline{(\gamma_4\otimes\xi_4)}. \quad \text{(C21)}$$

We consider only the consequences of this result for the momentum independent part of the vertex, i.e. that which survives in the continuum limit (when multiplied by an appropriate matching factor). Then $\Gamma_{S\otimes F}$ is simply a 16×16 matrix, having taste F (from translation invariance) but as yet undetermined spin,

$$\Gamma_{S\otimes F} = \sum_{S'} c^F_{SS'}\overline{(\gamma_{S'}\otimes\xi_F)}. \quad \text{(C22)}$$

The constraint (C21) implies that the only nonvanishing constants, $c^F_{SS'}$ are those for which S' satisfies $\eta_4(S'-S)=1$. This is because

$$\overline{(\gamma_4\otimes\xi_4)}\,\overline{(\gamma_{S'}\otimes\xi_F)}\,\overline{(\gamma_4\otimes\xi_4)} = \eta_4(S'-F)\overline{(\gamma_{S'}\otimes\xi_F)} \quad \text{(C23)}$$

and

$$\eta_4(S'-F) = \eta_4(S'-S)\eta_4(\Delta). \quad \text{(C24)}$$

Thus, inversions alone allow several choices of S', those satisfying $S'_1+S'_2+S'_3 =_2 S_1+S_2+S_3$.

To further constrain the propagator and vertices we turn to the final discrete symmetry, namely rotations. Here the analysis is more involved, since rotations mix bilinears. Consider the $(\mu\nu)$ rotation generator defined such that

$$p'_R = R^{-1}p', \qquad (p'_R)_\mu = -p'_\nu,$$
$$(p'_R)_\nu = p'_\nu, \qquad (p'_R)_\rho = p'_\rho, \quad \text{(C25)}$$

where μ, ν and ρ are all different. We find that the inverse propagator satisfies

$$S^{-1}(p') = \mathcal{R}S^{-1}(p'_R)\mathcal{R}^{-1}, \quad \text{(C26)}$$

where

$$\mathcal{R} = \frac{1}{2}\overline{([I+\gamma_{\mu\nu}]\otimes[I+\xi_{\mu\nu}])}. \quad \text{(C27)}$$

The key property of $\mathcal{R}$ is that it rotates the spin and taste matrices, e.g.,

$$\mathcal{R}\overline{(\gamma_\mu\otimes I)}\mathcal{R}^{-1} = -\overline{(\gamma_\nu\otimes I)}, \quad \text{(C28)}$$

$$\mathcal{R}\overline{(\gamma_\nu\otimes I)}\mathcal{R}^{-1} = \overline{(\gamma_\mu\otimes I)}, \quad \text{(C29)}$$

$$\mathcal{R}\overline{(I\otimes\xi_\mu)}\mathcal{R}^{-1} = -\overline{(I\otimes\xi_\nu)}. \quad \text{(C30)}$$

The result (C26) is the final input which leads to the general form of the propagator, Eq. (12). Given that S^{-1} is a taste singlet, (C26) enforces that each γ_μ must be multiplied by a power of p_μ.

The implication of rotation invariance for amputated vertices is

$$\Gamma_{S\otimes F}(p') = \psi(S,F)\mathcal{R}\Gamma_{S_R\otimes F_R}(p'_R)\mathcal{R}^{-1} \quad \text{(C31)}$$

$$\psi(S,F) = \frac{1}{16}\,\mathrm{tr}[\mathcal{R}\overline{(\gamma_{S_R}\otimes\xi_{F_R})}\mathcal{R}^{-1}\overline{(\gamma_S\otimes\xi_F)}^\dagger]. \quad \text{(C32)}$$

Note that, unlike for translations and spatial inversion, the vertices on the two sides of (C31) involve different operators. This is as expected since the operators fall into nontrivial irreps under the full lattice group. Note that the signs $\psi(S,F)$ are such that (C31) is satisfied if $\Gamma_{S\otimes F}(p') = \overline{(\gamma_S\otimes\xi_F)}$. We will show that, up to a constant, this is the only momentum-independent solution to (C31) that is also consistent with the relations from translations and spatial inversion.

Indeed, from translations and spatial inversion we know the form of the momentum-independent part of the amputated vertices to be that of Eq. (C22). Applying (C31) we learn that the constants satisfy

$$c^F_{SS'} = c^{F_R}_{S_RS'_R}\frac{\psi(S,F)}{\psi(S',F)}. \quad \text{(C33)}$$

At first sight, this appears to simply relate the constants appearing in the expansions of $\Gamma_{S\otimes F}$ and $\Gamma_{S_R\otimes F_R}$. However, if we apply (C33) twice, we obtain

$$c^F_{SS'} = c^F_{SS'}\frac{\psi(S,F)\psi(S_R,F_R)}{\psi(S',F)\psi(S'_R,F_R)}. \quad \text{(C34)}$$

Here we have used the result that $(S_R)_R = S$ for hypercube vectors, since their elements are binary numbers. It is straightforward to show that

$$\psi(S,F)\psi(S_R,F_R) = (-)^{S_\mu+S_\nu}(-)^{F_\mu+F_\nu}, \quad \text{(C35)}$$

so that (C34) becomes

$$c^F_{SS'} = c^F_{SS'}(-)^{S_\mu+S_\nu+S'_\mu+S'_\nu}. \quad \text{(C36)}$$

Thus we learn that the only nonvanishing constants are those for which $S'_\mu+S'_\nu =_2 S_\mu+S_\nu$ for all pairs (μ,ν). The only solutions are $S'=S$ and $S' =_2 S+(1111)$. This ambiguity is expected, since rotations alone allow mixing, e.g. between $\overline{(\gamma_\mu\otimes I)}$ and $\overline{(\gamma_{\mu5}\otimes I)}$. However, if we also enforce spatial inversion invariance, which, as seen above, implies $\eta_4(S')=\eta_4(S)$, then we find that only $S'=S$ is allowed. Thus we finally attain the desired result that

$$\Gamma_{S\otimes F}(p') \propto \overline{(\gamma_S\otimes\xi_F)} + O(a), \quad \text{(C37)}$$

where the $O(a)$ indicates momentum and mass dependent terms.

For completeness we note that one can obtain covariant operators containing derivatives by adding appropriate signs in the sum over Δ in Eq. (21). In particular, if the derivative is in the μth direction, instead of adding the two terms in (18), one takes the difference when $\Delta_\mu = 1$. This leads to the correspondence,

$$\frac{1}{V}a^3\int d^4x\bar{Q}\partial_\mu(\gamma_S\otimes\xi_F)Q \simeq \frac{1}{N_{\rm site}}\sum_n\frac{1}{N_\Delta}\sum_{|\Delta|=|S-F|}\Delta_\mu\bar{\chi}(n)\overline{(\gamma_S\otimes\xi_F)}_{n,n+S-F} \times \mathcal{U}_{n,n+\Delta}\chi(n+\Delta). \quad \text{(C38)}$$

The only difference from (21) is the factor of Δ_μ. This construction only works if $\Delta_\mu \neq 0$, i.e. if the spin taste of the operator is such that the $\bar{\chi}$ and χ fields are already separated in the μth direction. If they are not, one must use a two-step difference to get an operator containing a derivative [37].

APPENDIX D: PERTURBATIVE MATCHING FOR COVARIANT BILINEARS

In this appendix we describe briefly how the use of covariant bilinears changes the one-loop matching factors compared to those for hypercube bilinears. The latter have been calculated for our choices of fermion and gauge action in Ref. [39], following the earlier work of Refs. [30,33,36,47]. We also present numerical results for our choices of action, since these are not given in Ref. [39].

It is instructive to compare the tree-level matrix elements of the hypercubic and covariant bilinears between external quark "states" with physical momenta $p'+\pi C$ (outgoing from $\bar{\chi}$) and $p'+\pi D$ (incoming to χ). As explained in Ref. [36], the matrix element of a hypercubic operator is

$$M(S\otimes F;\text{hyp})^{(0)}_{CD} = \sum_{MN}E_M(p')E_N(-p')\overline{\overline{(\gamma_{MSN}\otimes\xi_{MFN})}}_{CD}, \quad \text{(D1)}$$

where, like S and F, M and N are hypercube vectors. (Note that "hyp" indicates hypercubic operator and should not be confused with "HYP" for HYP-smearing.) The functions that enter are

$$E_M(k) = \prod_\mu\frac{1}{2}(e^{-ik_\mu/2}+(-)^{\tilde{M}_\mu}e^{ik_\mu/2}), \quad \text{(D2)}$$

which are thus products of cosines or sines for the different components. We see from (D1) that, even in this tree-level matrix element, all combinations of spin and taste appear which satisfy $S'-F' = {}_2S-F$, i.e. which have the same number of links. This mixing is, however, suppressed by powers of a, since if $M\neq 0$ then $\tilde{M}\neq 0$, and there is at least one factor of $\sin(p'_\mu/2)\propto ap'^{\rm phys}_\mu$ in E_M. These factors of a correspond to the fact that the hypercube operators, when written in terms of irreps of the translation group, break up into the desired dimension-3 bilinear plus additional dimension-4 and higher operators containing derivatives.

If one projects out the part of this vertex with the same spin and taste as the initial bilinear (as one does in NPR), one finds

$$\frac{1}{48}\text{Tr}[\overline{\overline{(\gamma_S\otimes\xi_F)}}^\dagger M(S\otimes F;\text{hyp})^{(0)}] = \sum_M E_M(p')E_M(-p')(-)^{(S-F)\cdot\tilde{M}} \quad \text{(D3)}$$

$$= \prod_\mu\cos[p'_\mu(S-F)_\mu] \equiv V_{S\otimes F}(p'). \quad \text{(D4)}$$

To obtain the last line we have used the sum rule given in Eq. (A8) of Ref. [47]. Thus the tree-level kinematic factor associated with the hypercube bilinear is the vertex factor $V_{S\otimes F}$. This factor necessarily tends to unity in the continuum limit, but the $O(a^2)$ corrections contained in the cosines can be significant in practice.

At one-loop level, some of the hypercube operators mix. This mixing arises from the so-called X-diagrams (see, e.g., Ref. [39] for a figure explaining this terminology), in which the momentum flowing through the bilinear is not of $O(a)$ but rather of $O(1)$ (since it is inside a loop integral). This mixing is not suppressed by powers of a.

Now consider the covariant bilinears of Eq. (21). A key point is that the sign arising from the matrix $\overline{(\gamma_S\otimes\xi_F)}$ is independent of Δ. Thus the sum over Δ can be done, and leads exactly to the vertex factor $V_{S\otimes F}(p')$. The tree-level vertex is simply

$$M(S\otimes F;\text{cov})^{(0)}_{CD} = V_{S\otimes F}(p')\overline{\overline{(\gamma_S\otimes\xi_F)}}_{CD}, \quad \text{(D5)}$$

with no mixing with other spins and tastes. Because of the lack of mixing, one can read off the kinematical factor associated with this vertex without the need for projection. The result is that the kinematical factor is the same as that for hypercube operators.

Since there is no mixing in the vertex (D5), irrespective of the value of p', we expect that there will be no mixing between covariant bilinears in the one-loop calculation. This is indeed what we find by explicit calculation. Furthermore, it turns out that the diagonal (nonmixing) parts of the matching factors are identical to those for hypercube bilinears. For the X-diagrams, this is because the same factor $V_{S\otimes F}$ occurs in both tree-level vertices. For the "Y-diagrams" (those involving a gluon coupling to the vertex—see, e.g., Ref. [39]), the reason for the equality is similar. The remaining diagrams (self-energy and tadpoles) are the same for both operators.

Thus, we arrive at a very simple result. At one-loop order, the diagonal matching of covariant operators is identical to that for hypercube operators, while the off-diagonal matching coefficients vanish. We stress that the equality of diagonal matching factors should not hold at higher orders in perturbation theory. One way to see this is that the hypercube operators with different spin tastes that arise due to one-loop mixing can mix back with the

original operators at two-loop order. Such contributions are not present for the covariant operators.

The rest of this appendix is devoted to providing numerical results for bilinear one-loop matching factors for the Symanzik gauge action and our choices of valence fermions and links in the bilinears. Analytic formulas are given in Ref. [39], but that work quotes numerical values for several choices of fermion actions and operators differing from those we use. In particular, when HYP-smearing we use the HYP(1) choice of smearing parameters.

We are ultimately interested in perturbative predictions for the matching factors $Z_{S\otimes F}$ relating operators in the "lattice scheme" (i.e. the bare operators we place on the lattice and use in simulations) to those in the RI′ scheme. These are the matching factors we obtain nonperturbatively in our simulations using Eq. (24). However, perturbative calculations typically give results for matching from the lattice scheme to an intermediate continuum scheme, usually $\overline{\text{MS}}$. Thus to obtain the full matching factors, one must determine the matching between $\overline{\text{MS}}$ and RI′ schemes. This latter matching can be done in the continuum.

These considerations lead to the "master formula,"[17]

$$Z_{\mathcal{O}}^{\text{RI}',\text{LAT}}(\mu,a) = \exp\left[-\int_{\lambda(\mu_0)}^{\lambda(\mu)} d\lambda \frac{\gamma_{\mathcal{O}}(\lambda)}{\beta(\lambda)}\right] \times Z_{\mathcal{O}}^{\text{RI}',\overline{\text{MS}}}(\mu_0) \times Z_{\mathcal{O}}^{\overline{\text{MS}},\text{LAT}}(\mu_0,a). \quad \text{(D6)}$$

Here $\lambda = \alpha/(4\pi)$, with α always evaluated in the $\overline{\text{MS}}$ scheme. $\gamma_{\mathcal{O}}$ is the anomalous dimension of the operator ($\mathcal{O}$ is shorthand for $S\otimes F$), and $\beta(\lambda)$ the β function. In words, this equation says that one way of matching from the RI′ scheme at scale[18] μ to the lattice scheme with spacing a is to first run in the RI′ scheme to an intermediate scale $\mu_0 \approx 1/a$, then convert to the $\overline{\text{MS}}$ scheme at that scale, and finally convert to the lattice scheme at scale $1/a$. This formula allows one to have large values of the ratio μ/μ_0, with the first factor on the right-hand side summing the appropriate logarithms.

The one-loop results for matching from the lattice to the $\overline{\text{MS}}$ scheme have the form

$$Z_{\mathcal{O}}^{\overline{\text{MS}},\text{LAT}}(\mu_0,a) = \tilde{u}_0^{N_u}\left\{1 + \frac{\alpha(\mu_0)}{4\pi}[-2\gamma_{\mathcal{O}}^{(0)}\log(\mu_0 a) + C_{\mathcal{O}}^{\overline{\text{MS}}} - C_{\mathcal{O}}^{\text{LAT}}]\right\}, \quad \text{(D7)}$$

where $\gamma_{\mathcal{O}}^{(0)}$ is the one-loop anomalous dimension of the bilinear (defined precisely in the following appendix) and the C are finite constants. The continuum constants are $C_I^{\overline{\text{MS}}} = C_P^{\overline{\text{MS}}} = 10/3$, $C_V^{\overline{\text{MS}}} = C_A^{\overline{\text{MS}}} = 0$, and $C_T^{\overline{\text{MS}}} = 2/3$, and they do not depend on the taste. The factor of $\tilde{u}_0^{N_u}$ arises from possible mean-field improvement. This will be discussed below, including the appropriate values of N_u. Without such improvement, $\tilde{u}_0 = 1$. We stress again that one should choose $\mu_0 \approx 1/a$ when using this result; extending to other values of μ_0 requires resumming the leading logarithms using Eq. (D6).

A very important feature of the results (D6) and (D7) is that the anomalous dimensions depend only on the spin S but not on the taste F. The same is true for $Z_{\mathcal{O}}^{\text{RI}',\overline{\text{MS}}}$, and, as seen above, the $C_{\mathcal{O}}^{\overline{\text{MS}}}$. This implies that if one takes ratios of matching factors having different tastes but the same spin, then most of the terms in Eq. (D6) will cancel, yielding

$$\frac{Z_{S\otimes F}(\mu,a)}{Z_{S\otimes I}(\mu,a)} = \frac{Z_{S\otimes F}^{\overline{\text{MS}},\text{LAT}}(\mu_0,a)}{Z_{S\otimes I}^{\overline{\text{MS}},\text{LAT}}(\mu_0,a)} \quad \text{(D8)}$$

$$= \tilde{u}_0^{|S|-|S-F|}\left[1 + \frac{\alpha(\mu_0)}{4\pi}\delta_{S\otimes I}^{S\otimes F}\right], \quad \text{(D9)}$$

$$\delta_{S\otimes I}^{S\otimes F} = C_{S\otimes F}^{\text{LAT}} - C_{S\otimes 1}^{\text{LAT}}. \quad \text{(D10)}$$

Here we have taken the denominators (arbitrarily) to be taste singlets. The first line shows the cancellation of all except the lattice to $\overline{\text{MS}}$ matching factors, and has the important consequence that the ratios are predicted to be independent of μ. This holds to all orders in PT, and arises simply because it is only for momenta near the lattice cutoff $1/a$ that taste dependence enters. The lack of dependence on μ need not hold, however, for discretization errors, so the ratios can depend on powers of $(a\mu)^2$.

The second line of Eq. (D9) gives the one-loop result for the ratios, which, as shown in the third line, depends only on the (difference of the) finite lattice constants C^{LAT}. The values of these constants depend on whether mean-field improvement (along the lines of Ref. [41]) has been implemented. In ratios, mean-field improvement amounts to dividing the links in the bilinears by the fourth root of the plaquette built from those links, $\tilde{u}_0$. It is expected (and found) that bilinears with such rescaled links will have better behaved perturbative expansions [41]. Since we have not implemented this rescaling in our nonperturbative simulations, we must multiply by the rescaled bilinear by $\tilde{u}_0$ raised to the power of the number of links in the bilinear. These powers involve the length of the hypercube vectors S and $S - F$, where, e.g., $|S| = \sum_\mu |S_\mu|$. Although it might appear that multiplying and dividing by the same factors of $\tilde{u}_0$ would lead to no change, this is not the case because for the external factors we use the nonperturbatively determined value, while the impact of mean-field improvement

[17]For the sake of clarity, we have made the dependence on a explicit on the left-hand side, although this is left implicit in the main text.

[18]In the main text this scale is denoted p', but this symbol is used for a dimensionless lattice momentum earlier in this appendix, so we use μ here to denote a dimensionful energy scale.

TABLE V. Results for $C^{\rm LAT}_{S\otimes I}$ for the five choices of fermion action and operators explained in the text: (a) Naive with mean-field improvement, (b) HYP-smeared, (c) HYP-smeared with mean-field improvement, (d) asqtad with smeared links in operators, (e) asqtad with smeared and mean-field improved links in operators. The indices μ and ν are different.

Spin (S)	(a)	(b)	(c)	(d)	(e)
I	34.12	3.29	2.71	4.83	4.83
γ_μ	0	0	0	−1.91	−6.57
$\gamma_{\mu\nu}$	−1.54	−1.53	−0.96	0.23	−9.08

TABLE VI. Results for $\delta^{I\otimes F}_{I\otimes I}$, i.e. the finite coefficients for ratios involving scalar bilinears. The column "links" gives the number of links in the operator with the given taste. Choices of action and operators are as in Table V.

Taste (F)	Links	(a)	(b)	(c)	(d)	(e)
ξ_μ	1	21.84	2.80	2.23	1.99	6.64
$\xi_{\mu\nu}$	2	32.02	5.32	4.17	1.58	10.88
$\xi_{\mu 5}$	3	37.41	7.66	5.93	0.21	14.16
ξ_5	4	41.52	9.90	7.59	−1.51	17.10

in the differences δ is evaluated in one-loop perturbation theory. In effect, we are summing certain classes of diagrams to all orders in PT by using the nonperturbative $\tilde{u}_0$.

We present results for the taste singlet constants $C^{\rm LAT}_{S\otimes F1}$ and the differences $\delta^{S\otimes F1}_{S\otimes I}$ in Tables V, VI, VII, and VIII for the following choices of action and operators. In all cases the gauge action is the tree-level improved Symanzik action.[19]

(a) Mean-field improved naive staggered fermions with operators containing mean-field improved thin links. In this case, $\tilde{u}_0$ is determined from the thin link plaquette, and equals the u_0 discussed in the main text. Mean-field improvement of the links replaces U_μ with $U_\mu/\tilde{u}_0$. Mean-field improvement of the action follows the prescription explained, for the present context, in Refs. [36,39,47]. The improvement of the action has no impact on the differences δ, but does change the constants C, because the power of $\tilde{u}_0$ in Eq. (D7) becomes $N_u = 1 - |S - F|$.

(b) HYP-smeared staggered fermions with operators containing HYP-smeared links. No mean-field improvement is used, so that $\tilde{u}_0 = 1$. As noted above, HYP(1) smearing is used.

(c) As in (b), except with mean-field improved HYP-smeared links, with $\tilde{u}_0$ being the fourth root of the average plaquette composed of HYP-smeared links. The action is also mean-field improved, so that $N_u = 1 - |S - F|$ as in case (a).

(d) Asqtad fermions with operators containing smeared ("Fat7 + Lepage") links. No mean-field improvement.

(e) As in (d), but with mean-field improvement of the links in the operators, $\tilde{u}_0$ now being the fourth root of the plaquette composed of the same smeared links as used in the operators. Note that the asqtad action already includes some tadpole improvement, and no further improvement is made to the action. This means that $N_u = -|S - F|$ in Eq. (D7).

[19]The numerical results for cases (a) and (d) are directly obtained from those in Ref. [39], while those for cases (b), (c) and (e) are new. The latter is new because Ref. [39] did not consider mean-field improvement of the asqtad bilinears.

TABLE VII. Results for $\delta^{\hat{\mu}\otimes F}_{\hat{\mu}\otimes I}$, i.e. the finite coefficients for ratios involving vector bilinears. Notation as in Table V. The indices μ, ν and ρ are all different.

Taste (F)	Links	(a)	(b)	(c)	(d)	(e)
ξ_μ	0	−5.32	−1.05	−0.48	3.26	−1.40
$\xi_{\mu\nu}$	1	−3.46	0.32	0.32	0.09	0.09
ξ_ν	2	0.40	1.47	0.89	−3.23	1.42
$\xi_{\nu 5}$	2	0.51	1.83	1.25	−2.65	2.00
$\xi_{\nu\rho}$	3	3.06	3.04	1.88	−6.03	3.27
ξ_5	3	3.44	3.38	2.23	−5.49	3.82
$\xi_{\mu 5}$	4	5.80	4.64	2.91	−8.78	5.18

We present results only for scalar, vector and tensor bilinears, since multiplication of the operators by $\gamma_5 \otimes \xi_5$ leaves the constants unchanged. Thus those for pseudoscalars can be obtained from the results for scalars, and results from axial bilinears from those for vectors. In addition, three pairs of tensor matching factors are equal, as displayed in Table VIII.

For completeness, we also give the expressions for $\tilde{u}_0$ in PT. For HYP(1) smearing, we find

$$\tilde{u}_0^{\rm HYP} = 1 - C_F \frac{\alpha}{4\pi} 0.4331, \tag{D11}$$

where $C_f = 4/3$, while for "Fat-7 + Lepage" smearing,

$$\tilde{u}_0^{\rm ASQ} = 1 + C_F \frac{\alpha}{4\pi} 3.4897. \tag{D12}$$

We comment briefly on the values of the constants. $C^{\rm LAT}_{\gamma_\mu\otimes 1} = 0$ for naive and HYP-smeared quarks (see

TABLE VIII. Results for $\delta^{\hat{\mu}\,\hat{\nu}\otimes F}_{\hat{\mu}\,\hat{\nu}\otimes I}$, i.e. the finite coefficients for ratios involving tensor bilinears. Notation as in Table V. The indices μ, ν, ρ and σ are all different.

Taste (F)	Links	(a)	(b)	(c)	(d)	(e)
$\xi_{\mu\nu}$	0	2.74	−1.85	−0.69	8.73	−0.58
ξ_μ, $\xi_{\rho 5}$	1	−0.91	−1.05	−0.47	3.82	−0.84
$\xi_{\mu\rho}$	2	−0.36	−0.02	−0.02	−0.09	−0.09
ξ_5	2	0	0	0	0	0
ξ_ρ, $\xi_{\mu 5}$	3	1.57	1.17	0.59	−3.54	1.11
$\xi_{\rho\sigma}$	4	3.72	2.45	1.30	−6.82	2.49

Table V) because the taste-singlet vector bilinear is the conserved current. This is not the case for the asqtad action (due to the distance 3 Naik term), and so the constant need not (and does not) vanish. To give an idea of the size of the corrections, we note that on the coarse MILC lattices, the momenta within the window where NPR can be used range roughly from $\mu = 1.0/a \approx 1.7$ GeV to $\mu \approx 3$ GeV, so that $4\pi/\alpha(\mu)$ ranges from 38 to 52. Thus one needs perturbative coefficients C and δ to have magnitudes $\lesssim 10$ to have reasonable convergence. We see from the tables that this is the case except for the scalar (and pseudoscalar) bilinears with the naive staggered action and operators [case (a)]. This is one of the reasons why we do not present numerical results for this case in the main text. The constants are smallest for HYP-smeared operators, suggesting the PT should be better behaved in these cases. We also note that, while mean-field improvement reduces the magnitude of the corrections for the HYP-smeared action and operators, this is not uniformly the case for the asqtad action, where for scalars corrections are increased.

We close this section by describing two ways of rewriting the perturbative results that might have some practical utility. The first involves ratios of the differences of the initial ratios from unity:

$$\frac{\left(\frac{Z_{S\otimes F1}(\mu,a)}{Z_{S\otimes I}(\mu,a)}\tilde{u}_0^{-|S|+|S-F1|}-1\right)}{\left(\frac{Z_{S\otimes F2}(\mu,a)}{Z_{S\otimes I}(\mu,a)}\tilde{u}_0^{-|S|+|S-F2|}-1\right)}=\frac{\delta_{S\otimes I}^{S\otimes F1}}{\delta_{S\otimes I}^{S\otimes F2}}+\mathcal{O}(\alpha^2). \quad \text{(D13)}$$

The utility of this double ratio is that the coupling constant cancels in the one-loop contribution, so one obtains a simple numerical prediction. In practice, however, there are two difficulties: the $(a\mu)^2$ discretization errors need not cancel, and the relative size of the $\mathcal{O}(\alpha^2)$ contributions are typically different for tastes $F1$ and $F2$. Because of these difficulties, we have found it more useful to simply compare the initial single ratios to PT.

An alternative way of presenting PT results for ratios is to define $\alpha_{\rm eff}$ as follows:

$$\frac{Z_{S\otimes F}(\mu)}{Z_{S\otimes I}(\mu)}\tilde{u}_0^{-|S|+|S-F|}-1=\frac{\alpha_{\rm eff}}{4\pi}\delta_{S\otimes I}^{S\otimes F}+\mathcal{O}(\alpha^2). \quad \text{(D14)}$$

If the one-loop results gave a perfect representation of the data, $\alpha_{\rm eff}$ would be the almost the same for all ratios and independent of μ. There would be some variation since $\alpha_{\rm eff}$ is the coupling evaluated at a scale which we know to be of $\mathcal{O}(1/a)$ but whose precise value varies between ratios. It would then be interesting to take the ratio of the values of $\alpha_{\rm eff}$ at our two different lattice spacings, since this should lie in the range $\alpha(1/a_{\rm coarse})/\alpha(1/a_{\rm fine}) = 1.22$ to $\alpha(2/a_{\rm coarse})/\alpha(2/a_{\rm fine}) = 1.15$. In forming this ratio one should work at fixed $a\mu$ (rather than at fixed μ), so as to better cancel lattice artifacts.

Again, in practice we have found that the combination of noncanceling discretization errors and taste-dependent higher-order corrections makes this method difficult to use quantitatively. Thus in the main text we make a more qualitative comparison between the results on the two lattice spacings.

APPENDIX E: CONTINUUM PERTURBATIVE RESULTS

In this appendix we collect the results from the literature that allow us to predict the matching factors using perturbation theory using the master formula Eq. (D6).

There are many ways of writing the matching factor, with or without intermediate schemes, and with the running over the large range of scales taking place in different schemes. We have chosen the specific form (D6) for the following reasons. First, by doing the running from μ to $\mu_0 \approx 1/a$ first, we can, if we wish, move the running to the other side of the equation, and so convert the lattice results into a scale-independent form. Second, we need to use the intermediate $\overline{\rm MS}$ scheme because the matching to the lattice is only available in this scheme (as discussed in the previous appendix).

When evaluating the master expression, we have used the highest order available in the literature for each part. For the beta function in the $\overline{\rm MS}$ scheme, in the convention where

$$\beta(\lambda)=\frac{d\lambda}{d\ln(\mu^2)}=-\beta^{(0)}\lambda^2-\beta^{(1)}\lambda^3-\cdots, \quad \text{(E1)}$$

we have (setting here, and in the following, $N_c = N_f = 3$)

$$\beta^{(0)}=9, \quad \beta^{(1)}=64, \quad \beta^{(2)}=643.83, \quad \beta^{(3)}=12090.4. \quad \text{(E2)}$$

For the anomalous dimensions in the RI′ scheme, whose perturbative expansion we define as

$$\gamma_{\mathcal{O}}(\lambda)=-\frac{d\ln(Z_{\mathcal{O}})}{d\ln(\mu^2)} \quad \text{(E3)}$$

$$=\gamma_{\mathcal{O}}^{(0)}\lambda+\gamma_{\mathcal{O}}^{(1)}\lambda^2+\cdots, \quad \text{(E4)}$$

the coefficients are known to four loops for the scalar [48],

$$\gamma_S^{(0)}=-4, \quad \gamma_S^{(1)}=-108.67, \quad \gamma_S^{(2)}=-3576.95, \quad \gamma_S^{(3)}=-147207, \quad \text{(E5)}$$

and three loops for the tensor [49],

$$\gamma_T^{(0)}=1.33, \quad \gamma_T^{(1)}=34.44, \quad \gamma_T^{(2)}=976.64. \quad \text{(E6)}$$

The vector current also has a nonvanishing anomalous dimension in the RI′ scheme, which we determine below.

The conversion factors from RI′ to $\overline{\rm MS}$ can be obtained for the scalar bilinear from Ref. [48] and for the tensor from Ref. [49]. The results are

$$Z_S^{\mathrm{RI}',\overline{\mathrm{MS}}}(\mu_0) \approx 1 - 5.33\lambda - 121.37\lambda^2 - 3564.54\lambda^3, \quad \text{(E7)}$$

$$Z_T^{\mathrm{RI}',\overline{\mathrm{MS}}}(\mu_0) \approx 1 + 35.07\lambda^2 + 1207.96\lambda^3, \quad \text{(E8)}$$

with the coupling constant evaluated at scale μ_0.

To obtain the result for the vector bilinear, we first note that $Z_V^{\mathrm{RI},\overline{\mathrm{MS}}} = 1$, as shown in Ref. [1]. Here RI refers to the original regularization independent scheme of Ref. [1], in which the condition used to determine Z_q from the quark propagator differs from that in the RI′ scheme. The condition determining Z_q/Z_V, Eq. (24), is, however, the same in both schemes, from which we learn that [50]

$$Z_V^{\mathrm{RI}',\mathrm{RI}} = Z_q^{\mathrm{RI}',\mathrm{RI}}. \quad \text{(E9)}$$

Combining these results, we find the desired matching factor

$$Z_V^{\mathrm{RI}',\overline{\mathrm{MS}}} = Z_V^{\mathrm{RI}',\mathrm{RI}} Z_V^{\mathrm{RI},\overline{\mathrm{MS}}} \quad \text{(E10)}$$

$$= Z_q^{\mathrm{RI}',\mathrm{RI}} \quad \text{(E11)}$$

$$= 1 + c_2\lambda^2 + c_3\lambda^3 + \cdots \quad \text{(E12)}$$

$$\approx 1 + 9.17\lambda^2 + 342.01\lambda^3, \quad \text{(E13)}$$

where the numerical values are from Ref. [50]. This, in turn, can be used to obtain the anomalous dimension,

$$\gamma_V^{\mathrm{RI}'} = -\frac{d\ln Z_V^{\mathrm{RI}',\overline{\mathrm{MS}}}}{d\ln(\mu^2)} \quad \text{(E14)}$$

$$= 2c_2\beta^{(0)}\lambda^3 + (2c_2\beta^{(1)} + 3c_3\beta^{(0)})\lambda^4 + \cdots \quad \text{(E15)}$$

$$\approx 165\lambda^3 + 10407.5\lambda^4. \quad \text{(E16)}$$

[1] G. Martinelli, C. Pittori, C. T. Sachrajda, M. Testa, and A. Vladikas, Nucl. Phys. **B445**, 81 (1995).
[2] Y. Aoki, Proc. Sci., LAT2009 (2009) 012.
[3] S. Aoki *et al.* (JLQCD Collaboration), Phys. Rev. Lett. **82**, 4392 (1999).
[4] K. Orginos and D. Toussaint (MILC Collaboration), Phys. Rev. D **59**, 014501 (1998).
[5] K. Orginos, D. Toussaint, and R. Sugar (MILC Collaboration), Phys. Rev. D **60**, 054503 (1999).
[6] G. Lepage, Phys. Rev. D **59**, 074502 (1999).
[7] A. Hasenfratz and F. Knechtli, Phys. Rev. D **64**, 034504 (2001).
[8] A. Bazavov, D. Toussaint, C. Bernard, J. Laiho, C. DeTar *et al.*, Rev. Mod. Phys. **82**, 1349 (2010).
[9] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. Lepage, J. Shigemitsu, H. Trottier, and K. Wong (HPQCD Collaboration and UKQCD Collaboration), Phys. Rev. D **75**, 054502 (2007).
[10] A. Bazavov *et al.* (MILC Collaboration), Phys. Rev. D **82**, 074501 (2010).
[11] C. Davies, C. McNeile, K. Wong, E. Follana, R. Horgan, K. Hornbostel, G. P. Lepage, J. Shigemitsu, and H. Trottier, Phys. Rev. Lett. **104**, 132003 (2010).
[12] A. T. Lytle, Proc. Sci., LAT2009 (2009) 202.
[13] A. T. Lytle, Ph.D. thesis, University of Washington, 2010.
[14] C. Aubin *et al.* (HPQCD Collaboration, MILC Collaboration, and UKQCD Collaboration), Phys. Rev. D **70**, 031504 (2004).
[15] Q. Mason, H. D. Trottier, R. Horgan, C. T. Davies, and G. Lepage (HPQCD Collaboration), Phys. Rev. D **73**, 114501 (2006).
[16] T. Bae, Y.-C. Jang, C. Jung, H.-J. Kim, J. Kim, J. Kim, K. Kim, W. Lee, S. R. Sharpe, and B. Yoon, Phys. Rev. D **82**, 114509 (2010).
[17] T. Bae *et al.*, Phys. Rev. Lett. **109**, 041601 (2012).
[18] R. Arthur and P. Boyle (RBC Collaboration and UKQCD Collaboration), Phys. Rev. D **83**, 114511 (2011).
[19] M. Luscher, R. Narayanan, P. Weisz, and U. Wolff, Nucl. Phys. **B384**, 168 (1992).
[20] M. Della Morte, R. Hoffmann, F. Knechtli, J. Rolf, R. Sommer, I. Wetzorke, and U. Wolff (ALPHA Collaboration), Nucl. Phys. **B729**, 117 (2005).
[21] M. L. Paciello, S. Petrarca, B. Taglienti, and A. Vladikas, Phys. Lett. B **341**, 187 (1994).
[22] L. Giusti, M. Paciello, C. Parrinello, S. Petrarca, and B. Taglienti, Int. J. Mod. Phys. A **16**, 3487 (2001).
[23] L. Giusti, S. Petrarca, B. Taglienti, and N. Tantalo, Phys. Lett. B **541**, 350 (2002).
[24] C. Gattringer, M. Gockeler, P. Huber, and C. Lang, Nucl. Phys. **B694**, 170 (2004).
[25] A. Lokhov, O. Pene, and C. Roiesnel, arXiv:hep-lat/0511049.
[26] I. Bogolubsky, G. Burgio, M. Muller-Preussker, and V. Mitrjushkin, Phys. Rev. D **74**, 034503 (2006).
[27] A. T. Lytle and S. R. Sharpe, Proc. Sci., LATTICE2011 (2011) 237.
[28] A. T. Lytle and S. R. Sharpe, Proc. Sci., LATTICE2012 (2012) 242.
[29] W.-j. Lee and S. R. Sharpe, Phys. Rev. D **68**, 054510 (2003).
[30] M. F. Golterman and J. Smit, Nucl. Phys. **B245**, 61 (1984).
[31] H. Kluberg-Stern, A. Morel, O. Napoly, and B. Petersson, Nucl. Phys. **B220**, 447 (1983).
[32] D. Daniel and T. Kieu, Phys. Lett. B **175**, 73 (1986).
[33] D. Daniel and S. N. Sheard, Nucl. Phys. **B302**, 471 (1988).
[34] H. Politzer, Nucl. Phys. **B117**, 397 (1976).
[35] P. Pascual and E. de Rafael, Z. Phys. C **12**, 127 (1982).
[36] A. Patel and S. R. Sharpe, Nucl. Phys. **B395**, 701 (1993).

[37] S. R. Sharpe and A. Patel, Nucl. Phys. **B417**, 307 (1994).
[38] M. F. Golterman and J. Smit, Nucl. Phys. **B255**, 328 (1985).
[39] J. Kim, W. Lee, and S. R. Sharpe, Phys. Rev. D **81**, 114503 (2010).
[40] R. G. Edwards and B. Joo (SciDAC Collaboration, LHPC Collaboration, and UKQCD Collaboration), Nucl. Phys. B, Proc. Suppl. **140**, 832 (2005).
[41] G. Lepage and P. B. Mackenzie, Phys. Rev. D **48**, 2250 (1993).
[42] M. F. Golterman, Nucl. Phys. **B273**, 663 (1986).
[43] D. Verstegen, Nucl. Phys. **B249**, 685 (1985).
[44] G. Kilcup and S. R. Sharpe, Nucl. Phys. **B283**, 493 (1987).
[45] M. F. Golterman, Nucl. Phys. **B278**, 417 (1986).
[46] J. E. Mandula, G. Zweig, and J. Govaerts, Nucl. Phys. **B228**, 91 (1983).
[47] W.-j. Lee and S. R. Sharpe, Phys. Rev. D **66**, 114501 (2002).
[48] K. Chetyrkin and A. Retey, Nucl. Phys. **B583**, 3 (2000).
[49] J. Gracey, Nucl. Phys. **B662**, 247 (2003).
[50] P. Huber, J. High Energy Phys. 11 (2010) 107.