PHYSICAL REVIEW D **71,** 114511 (2005)

A lattice determination of moments of unpolarized nucleon structure functions using improved Wilson fermions

M. Göckeler,[1,2] R. Horsley,[3] D. Pleiter,[4] P. E. L. Rakow,[5] and G. Schierholz[4,6]

(QCDSF Collaboration)

[1]*Institut für Theoretische Physik, Universität Leipzig, D-04109 Leipzig, Germany*
[2]*Institut für Theoretische Physik, Universität Regensburg, D-93040 Regensburg, Germany*
[3]*School of Physics, University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom*
[4]*John von Neumann-Institut für Computing NIC, Deutsches Elektronen-Synchrotron DESY, D-15738 Zeuthen, Germany*
[5]*Theoretical Physics Division, Department of Mathematical Sciences, University of Liverpool, Liverpool L69 3BX, United Kingdom*
[6]*Deutsches Elektronen-Synchrotron DESY, D-22603 Hamburg, Germany*

(Received 14 October 2004; published 23 June 2005)

Within the framework of quenched lattice QCD and using $O(a)$ improved Wilson fermions and nonperturbative renormalization, a high statistics computation of low moments of the unpolarized nucleon structure functions is given. Particular attention is paid to the chiral and continuum extrapolations.

DOI: 10.1103/PhysRevD.71.114511 PACS numbers: 12.38.Gc, 11.15.Ha

I. INTRODUCTION

The results of a lattice simulation of quantum chromodynamics (QCD) give in principle a direct probe of certain low energy aspects of the theory, such as hadronic masses and matrix elements. This is at present the only way of getting these quantities from QCD, without additional model-dependent assumptions. A useful theoretical tool in conjunction with QCD and deep inelastic scattering (DIS) experiments is the operator product expansion (OPE). At leading twist the OPE relates moments of an experimentally measured structure function, generically denoted by F, to certain matrix elements v_n where

$$\int_0^1 dx x^{n-2} F^{\mathrm{NS}}(x, Q^2) = f E^S_{F;v_n}\left(\frac{M^2}{Q^2}, g^S(M)\right) v_n^S(g^S(M)). \tag{1}$$

F is a function of two variables—Q^2, the spacelike momentum transfer to the nucleon and x, the Bjorken variable (f is a normalization factor). v_n are the nucleon matrix elements of certain operators and E are the associated Wilson coefficients. These are perturbatively known at high energies where the coupling constant g becomes small and are found in a specified scheme S at scale M. Usually, of course, we take $S \equiv \overline{\mathrm{MS}}$ at scale $M \equiv \mu \sim$ few GeV. We also assume that Q^2 is large enough, so that higher twist terms i.e. $O(1/Q^2)$ terms are negligible.

As will be discussed later (Sec. IV C), lattice computations are presently restricted to determining nonsinglet (NS) nucleon structure functions

$$F^{\mathrm{NS}}(x, Q^2) \equiv F^p(x, Q^2) - F^n(x, Q^2), \tag{2}$$

i.e. the difference between proton, p, and neutron, n, results. Note, in particular, that this means that nucleon matrix elements of gluonic operators have canceled.

In this article we shall only be concerned with unpolarized structure functions. The same matrix elements v_n contribute to the scattering of charged leptons and of neutrinos, but the weights f are different in the two cases. Thus for charged lepton-nucleon DIS, $lN \to lX$, which is mediated by a photon, we have $F^{\mathrm{NS}} \to F_2^{\gamma;\mathrm{NS}}$ with $n = 2, 4, \ldots$ and $f^\gamma = 2(e_u^2 - e_d^2)/2 = 1/3$. For neutrino-nucleon charged weak current interactions for example $\nu N \to lX$, ($l^+ N \to \overline{\nu} X$) or $lN \to \nu X$, ($\overline{\nu} N \to l^+ X$) which are mediated by W^+, W^- respectively, then we have $F^{\mathrm{NS}} \to F_2^{W^\pm;\mathrm{NS}}$ and $f^{W^+} = 2(-1)$, $f^{W^-} = 2(+1)$ upon neglecting the Cabibbo-Kobayashi-Maskawa mixing matrix with $n = 3, 5, \ldots$. Alternatively setting $F^{\mathrm{NS}} \to 2xF_1^{\mathrm{NS}}$ in all cases one has the same matrix elements and fs as for F_2^{NS} in Eq. (1), but different Wilson coefficients. The additional F_3 structure functions, occurring because of parity nonconservation also obey Eq. (1) with $F^{\mathrm{NS}} \to xF_3^{W^\pm;\mathrm{NS}}$ and again $f^{W^\pm} = \mp 2$ with $n = 2, 4, \ldots$. Similar expressions hold for the neutral currents, but with more complicated expressions for the fs.

In all cases the relevant matrix elements are given by first defining the sequence of quark bilinear forms

$$\mathcal{O}_q^{\mu_1 \cdots \mu_n} = i^{n-1} \overline{q} \gamma^{\mu_1} \overleftrightarrow{D}^{\mu_2} \cdots \overleftrightarrow{D}^{\mu_n} q, \qquad q = u, d, \tag{3}$$

where $\overleftrightarrow{D} = \frac{1}{2}(\vec{D} - \overleftarrow{D})$. Symmetrizing the indices and removing traces, gives the Lorentz decomposition of the proton (i.e. nucleon, N) matrix element of [1]

$$\langle N(\vec{p}) | [\mathcal{O}_q^{\{\mu_1 \cdots \mu_n\}} - \mathrm{Tr}] | N(\vec{p}) \rangle^S \equiv 2 v_n^{(q)S} [p^{\mu_1} \cdots p^{\mu_n} - \mathrm{Tr}]. \tag{4}$$

For example we have for $n = 2$,

$$\langle N(\vec{p})|i\overline{q}\left[\frac{1}{2}[\gamma^{\mu_1}\vec{D}^{\mu_2}+\gamma^{\mu_2}\vec{D}^{\mu_1}]-\frac{1}{4}\vec{\not{D}}\eta^{\mu_1\mu_2}\right]q|N(\vec{p})\rangle^S$$
$$=2v_2^{(q)S}\left[p^{\mu_1}p^{\mu_2}-\frac{1}{4}m_N^2\eta^{\mu_1\mu_2}\right], \quad (5)$$

where $\eta = \mathrm{diag}(-1,1,1,1)$. More complicated expressions hold for higher moments. Finally, the nonsinglet, NS, matrix element is defined as

$$v_n^S \equiv v_n^{(u)S} - v_n^{(d)S}. \quad (6)$$

In this article, we shall compute v_2, v_3 and v_4 in the quenched approximation ($n_f = 0$), by finding the appropriate matrix elements in Eq. (4). As will be seen most effort will be spent on v_2, as this is technically less complicated than the higher moments, and also numerically the lattice results are more precise. The lattice approach discretizes Euclidean space-time, with lattice spacing a, in the path integral and simulates the resulting high-dimensional integral for the partition function using Monte Carlo techniques. Matrix elements can then be obtained from suitable ratios of correlation functions [2,3]. Note that the lattice program is rather like an experiment: careful account must be taken of error estimations and extrapolations. There are three limits to consider:

(i) The spatial box size L_S must be large enough so that finite size effects are small. Currently sizes of $\gtrsim 2$ fm seem large enough [4,5] (the nucleon diameter is about 1.5 fm) [6]. This situation is probably more favorable for quenched QCD where one drops the fermion determinant in the path integral, see Sec. IV, as due to the suppression of the pion cloud we would expect the radius of the nucleon to be somewhat smaller. This indeed seems to be the case, see e.g. [7].

(ii) The continuum limit, $a \to 0$. We use $O(a)$ improved Wilson fermions, where the discretization effects of the action and matrix elements have been arranged to be $O(a^2)$. For unimproved Wilson fermions, or where one has not succeeded in entirely $O(a)$ improving the matrix element we should extrapolate in a rather than a^2.

(iii) The chiral limit, when the quark mass approaches zero. There has recently been much activity on deriving formulas for this limit [8–13]. However while most of these results are valid around the physical pion mass on the lattice, it is difficult to calculate quark propagators at quark masses much below the strange quark mass. Thus the use of these formulas is not straightforward.

In addition the lattice matrix element must also be renormalized.

Previous lattice studies gave discrepancies to the phenomenological results, especially for v_2. In this work we want to try to narrow down the sources for this difference. In particular we shall present here nonperturbative (NP) results for the renormalization constants, and as many previous studies used results based on perturbation theory compare with these other results. We also consider $O(a)$ improvement and operator mixing to enable a reliable continuum extrapolation to be performed.

Compared to our previous work [3] we have improved our techniques in several respects:

(i) We employ nonperturbatively improved Wilson fermions instead of unimproved Wilson fermions. This should reduce cutoff effects.

(ii) Modified operators are used for v_3, v_4, which improves the numerical signal.

(iii) In [3] we had simulations for a single lattice spacing only. Here we shall present results for three different values of the lattice spacing so that we can monitor lattice artifacts.

(iv) As mentioned before, the 1-loop perturbative renormalization factors of [3] have been replaced by nonperturbatively calculated renormalization constants. In addition we shall pay close attention to possible mixing problems of the operators involved.

(v) We have increased the number of quark masses at each value of the lattice spacing in order to improve the chiral extrapolation.

The organization of this paper is as follows. In Sec. II various continuum results for the β and γ functions and for the Wilson coefficients in the $\overline{\mathrm{MS}}$ scheme are collated and renormalization group invariant quantities are introduced, while in Sec. III some NS phenomenological results are discussed to compare later with the lattice results. Section IV describes our lattice techniques: choice of operators, operator mixing problems, $O(a)$ improvement and gives the bare, i.e. unrenormalized results. Section V discusses and compares various renormalization results: one-loop perturbation theory and a tadpole improvement, together with the RI′-MOM nonperturbative method. In Sec. VI we discuss the results and give continuum and chiral extrapolations. Finally in Sec. VII we present our conclusions.

II. CONTINUUM QCD RESULTS

In this section we shall consider the right-hand side (rhs) of Eq. (1). Much of the functional form is already known: the lattice input is reduced, for each moment, to the computation of a single number.

The running of the coupling constant as the scale M is changed is controlled by the β function. This is defined as

$$\beta^S(g^S(M)) \equiv \frac{\partial g^S(M)}{\partial \log M}\Big|_{\text{bare}}$$
$$= -b_0(g^S)^3 - b_1(g^S)^5 - b_2^S(g^S)^7 - b_3^S(g^S)^9 - \ldots, \quad (7)$$

with the bare parameters being held constant. This function is given perturbatively as a power series expansion in the

coupling constant. The first two coefficients in the expansion are universal, i.e. scheme independent. In the $\overline{\rm MS}$ scheme where $(S, M) = (\overline{\rm MS}, \mu)$, the expansion is now known to four loops [14,15]. The three-loop result for quenched QCD is

$$b_0 = \frac{11}{(4\pi)^2}, \qquad b_1 = \frac{102}{(4\pi)^4}, \qquad b_2^{\overline{\rm MS}} = \frac{1}{(4\pi)^6}\left[\frac{2857}{2}\right]. \tag{8}$$

We may immediately integrate Eq. (7) to obtain the solution,

$$M = \Lambda^S \exp\left[\frac{1}{2b_0 g^S(M)^2}\right][b_0 g^S(M)^2]^{b_1/2b_0^2} \times \exp\left\{\int_0^{g^S(M)} d\xi\left[\frac{1}{\beta^S(\xi)} + \frac{1}{b_0\xi^3} - \frac{b_1}{b_0^2\xi}\right]\right\}, \tag{9}$$

where Λ^S is an integration constant.

While Eq. (1) is the conventional definition of the moment of a structure function, for us it is more convenient to rewrite it using renormalization group invariant (RGI) functions. The bare operator, or matrix element, must first be renormalized

$$\mathcal{O}^S(M) = Z^S_{\mathcal{O};\rm bare}(M)\mathcal{O}_{\rm bare}, \tag{10}$$

giving γ, the anomalous dimension of the operator,

$$\gamma^S_{\mathcal{O}}(g^S(M)) \equiv \frac{\partial \log Z^S_{\mathcal{O}}(M)}{\partial \log M}\bigg|_{\rm bare} = d_{\mathcal{O};0}(g^S)^2 + d^S_{\mathcal{O};1}(g^S)^4 + d^S_{\mathcal{O};2}(g^S)^6 + \ldots. \tag{11}$$

The first coefficient is again scheme independent. One may also change the scale and/or scheme $(S, M) \to (S', M')$ for the operator by

$$\mathcal{O}^{S'}(M') = Z^{S'}_{\mathcal{O};S}(M', M)\mathcal{O}^S(M). \tag{12}$$

This leads to two anomalous dimension equations, obtained by either differentiating with respect to M' or M. Integrating these equations gives

$$Z^{S'}_{\mathcal{O};S}(M', M) = \frac{Z^{S'}_{\mathcal{O};\rm bare}(M')}{Z^S_{\mathcal{O};\rm bare}(M)} \equiv \frac{\Delta Z^S_{\mathcal{O}}(M)}{\Delta Z^{S'}_{\mathcal{O}}(M')}, \tag{13}$$

where we have defined

$$[\Delta Z^S_{\mathcal{O}}(M)]^{-1} = [2b_0 g^S(M)^2]^{-d_{\mathcal{O};0}/2b_0} \times \exp\left\{\int_0^{g^S(M)} d\xi\left[\frac{\gamma^S_{\mathcal{O}}(\xi)}{\beta^S(\xi)} + \frac{d_{\mathcal{O};0}}{b_0\xi}\right]\right\}. \tag{14}$$

From Eqs. (12) and (13), we see that we can define a RGI operator by

$$\mathcal{O}^{\rm RGI} = \Delta Z^S_{\mathcal{O}}(M)\mathcal{O}^S(M) \equiv Z^{\rm RGI}_{\mathcal{O}}\mathcal{O}_{\rm bare}. \tag{15}$$

Then obviously $\mathcal{O}^{\rm RGI}$ is independent of the scale and scheme. The γ function thus controls how the matrix element changes as the scale M is varied. Note also that the normalization of $\mathcal{O}^{\rm RGI}$ depends on the convention chosen for $\Delta Z^S_{\mathcal{O}}(M)$, here given in Eq. (14).

As the left-hand side of Eq. (1) is a physical quantity, the RGI form for the Wilson coefficient is given by

$$E^{\rm RGI}_{F;v_n}(Q^2) = \Delta Z^S_{v_n}(M)^{-1}E^S_{F;v_n}\left(\frac{M^2}{Q^2}, g^S(M)\right). \tag{16}$$

It is convenient to choose $M^2 \equiv Q^2$, as then

$$E^{\rm RGI}_{F;v_n}(Q^2) = \Delta Z^S_{v_n}(Q)^{-1}E^S_{F;v_n}(1, g^S(Q)), \tag{17}$$

and $E^S_{F;v_n}(1, g^S(Q))$ has no large numbers in it, so that a perturbative power series in g^S becomes tenable. In two schemes S and S' from Eq. (17) we have

$$\frac{E^S_{F;v_n}(1, g^S(Q))}{E^{S'}_{F;v_n}(1, g^{S'}(Q))} = \frac{\Delta Z^S_{v_n}(Q)}{\Delta Z^{S'}_{v_n}(Q)} \to 1 \quad \text{as } Q^2 \to \infty, \tag{18}$$

because in this limit $g^{S'} = g^S + \ldots \to 0$. Hence $E^S_{F;n}(1, 0)$ is independent of the scheme. With our convention for f this is 1, so that

$$E^S_{F;v_n}(1, g^S) = 1 + e^S_{F;n;0}(g^S)^2 + e^S_{F;n;1}(g^S)^4 + \ldots. \tag{19}$$

Practically we shall here only consider the $n = 2$, 3 and 4 moments. For these moments we have, for quenched QCD (i.e. $n_f = 0$) [16–19]

$$e^{\overline{\rm MS}}_{F_2;n;0} = \frac{1}{(4\pi)^2}\left\{\frac{4}{9}; \frac{29}{9}; \frac{91}{15}\right\} \qquad n = \{2; 3; 4\}, \tag{20}$$

$$e^{\overline{\rm MS}}_{F_2;n;1} = \frac{1}{(4\pi)^4}\left\{\frac{363\,604}{3645} - \frac{1024}{15}\zeta(3);\ 62.74;\ \frac{1\,112\,778\,271}{3\,645\,000} - \frac{1220}{9}\zeta(3)\right\},$$

for $n = 2$, 3 and 4 respectively, where $\zeta(3) = 1.20206\ldots$. The operator has anomalous dimensions given by [20–22]

$$d_{v_n;0} = -\frac{1}{(4\pi)^2}\left\{\frac{64}{9}; \frac{100}{9}; \frac{628}{45}\right\} \qquad n = \{2; 3; 4\},$$
$$d^{\overline{\rm MS}}_{v_n;1} = -\frac{1}{(4\pi)^4}\left\{\frac{23\,488}{243}; \frac{34\,450}{243}; \frac{5\,241\,914}{30\,375}\right\}, \tag{21}$$

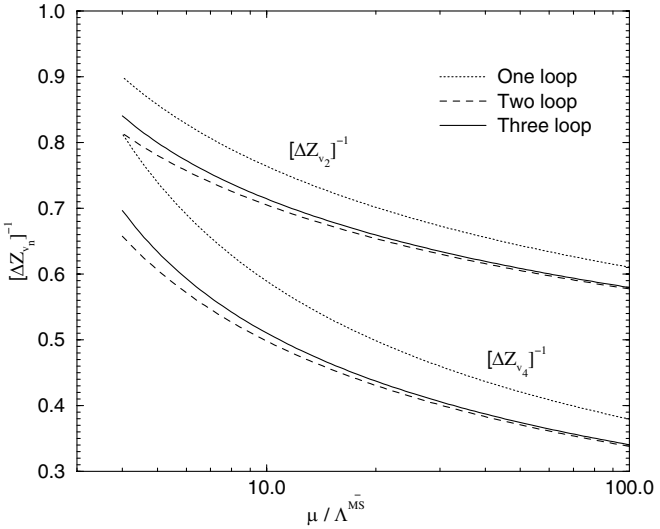

FIG. 1. The one-, two- and three-loop results for $[\Delta Z^{\overline{\rm MS}}_{v_2}(\mu)]^{-1}$ and $[\Delta Z^{\overline{\rm MS}}_{v_4}(\mu)]^{-1}$ for quenched QCD versus $\mu/\Lambda^{\overline{\rm MS}}$.

$$d^{\overline{\rm MS}}_{v_n;2} = -\frac{1}{(4\pi)^6}\left\{\frac{11\,028\,416}{6561} + \frac{2560}{81}\zeta(3);\ \frac{64\,486\,199}{26\,244} + \frac{2200}{81}\zeta(3); \frac{245\,787\,905\,651}{82\,012\,500} + \frac{11\,512}{405}\zeta(3)\right\},$$

again for $n = 2$, 3 and 4, respectively. Solving first Eq. (9) for $g^{\overline{\rm MS}}$ and then using Eq. (14) gives the results for $[\Delta Z^{\overline{\rm MS}}_{v_n}(\mu)]^{-1}$ shown in Fig. 1. Note that by loop expansion, we mean using the β and γ function result to the appropriate order; we do not expand Eqs. (9) and (14) any further, but solve them numerically.

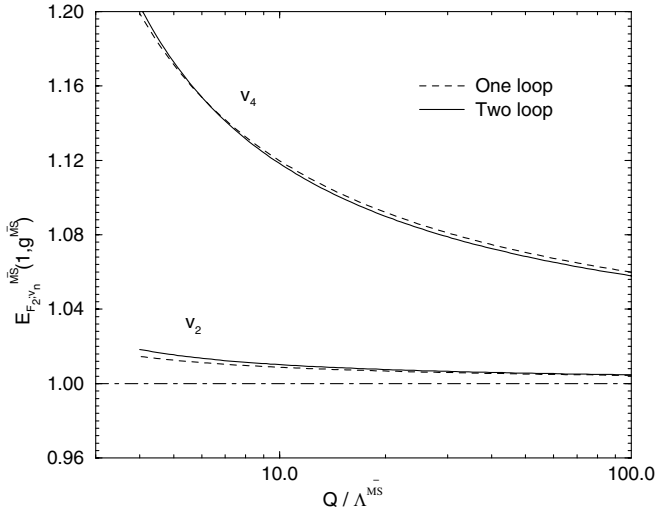

FIG. 2. The one- and two-loop results for $E^{\overline{\rm MS}}_{F_2;v_n}$, $n = 2, 4$ for quenched QCD versus $Q/\Lambda^{\overline{\rm MS}}$. The two-loop results at 2 GeV are 1.011(1), 1.130(3) for $n = 2, 4$ respectively where the error is a reflection of the error in $\Lambda^{\overline{\rm MS}} r_0$.

To determine a result in GeV, we shall use the r_0 scale here [23]. From [24,25] we take $\Lambda^{\overline{\rm MS}} r_0 = 0.602(48)$ for quenched QCD and together with the scale choice $r_0 = 0.5$ fm $\equiv 1/(394.6$ MeV) this gives for an energy of $Q \equiv \mu = 2$ GeV for example, $\mu/\Lambda^{\overline{\rm MS}} \sim 8.4$. The Wilson coefficient, $E^{\overline{\rm MS}}_{F_2;v_n}$ can also be found and is shown in Fig. 2 for $n = 2, 4$. To obtain $E^{\rm RGI}_{F_2;v_n}$ from Eq. (17) we must simply multiply the results from Fig. 1 with those of Fig. 2. From this latter figure we see that the change from the tree level result for the $n = 2$ moment in the Wilson coefficient is at most $\sim$1% for $Q \sim 2$ GeV and is practically negligible. This is not so for the higher moments, when the Wilson coefficient deviates significantly from one. Useful values for $\Delta Z^{\overline{\rm MS}}_{v_n}$, relevant for the forthcoming lattice results, are given in Table VI in Appendix A.

III. PHENOMENOLOGY AND EXPERIMENTAL DATA

Ideally we would like to make a direct comparison between the theoretical and experimental result, by rewriting Eq. (1) as,

$$\int_0^1 dx x^{n-2} F^{\rm NS}(x, Q^2) = f E^{\rm RGI}_{F;v_n}(Q^2) v^{\rm RGI}_n. \qquad (22)$$

The rhs of this equation has a clean separation between a number $v^{\rm RGI}_n$, which can only be obtained using a nonperturbative method (e.g. the lattice approach) and a function, $E^{\rm RGI}_{F;v_n}(Q^2)$, which describes all the momentum behavior of the moment.

More conventional and practical however is to use parton densities. Usually phenomenological fits using parton densities are obtained from global fits (such as MRST [26] and CTEQ [27]) to the data. In this section we shall compare whether taking moments of the structure function gives the same answer as taking moments of the parton density. This could also help in estimating the error in the phenomenological fit. Parton densities q, $\overline{q}$ are implicitly defined by

$$\int_0^1 dx x^{n-1}[q^S(x, Q) + (-1)^n \overline{q}^S(x, Q)] = v^{(q)S}_n(Q). \quad (23)$$

We may relate the structure function to the parton density via a convolution. Defining similar but separate Mellin transformations for even and odd n by

$$h_n = \int_0^1 dx x^{n-1} h(x) \qquad n = \text{even/odd},$$
$$h(x) = \int_{c-i\infty}^{c+i\infty} \frac{dn}{2\pi i} x^{-n} h_n, \qquad (24)$$

where in the inverse transformation n in h_n is analytically continued from even/odd integer n values to complex numbers and c is chosen so that all singularities lie to the left of the line $n = c$, then gives

$$F^{\rm NS}(x,Q^2) = fx\int_0^1 \frac{dy}{y} E^S(x/y,Q)[(u^S(y,Q)-d^S(y,Q)) + \eta(\overline{u}^S(y,Q)-\overline{d}^S(y,Q))], \quad (25)$$

where $\eta = +1$ for even n (i.e. for $F_2^{\gamma;\rm NS}$) and $\eta = -1$ for odd n (i.e. for $F_2^{W^\pm;\rm NS}$). To lowest order in the coupling constant we get from Eq. (19), $E^S(z,Q) = \delta(1-z) + O(g^S)^2$ and so [28]

$$F^{\rm NS}(x,Q^2) = fx[(u^S(x,Q)-d^S(x,Q)) + \eta(\overline{u}^S(x,Q) - \overline{d}^S(x,Q))] + \ldots. \quad (26)$$

The parton densities are usually determined from global fits to the data, with an assumed functional form, typically for MRST results like

$$xq^{\overline{\rm MS}}(x,Q_0) = A_q x^{\lambda_q}(1-x)^{\eta_q}(1+\epsilon_q\sqrt{x}+\gamma_q x), \quad (27)$$

with parameters A_q, λ_q, η_q, ϵ_q and γ_q at some given reference scale Q_0. For the MRST results given here we use the fit MRST2001E [30] together with the error analysis of [26] at a scale of $Q_0^2 = 4$ GeV^2. As a comparison we also consider the CTEQ fit CTEQ61M, with errors calculated from [27]. Practically, in both cases, we use the parton distribution calculator [31] to compute the moments in Eq. (23).

Let us now briefly consider some lepton-nucleon DIS experimental results. While $F_2^{\gamma;p}(x,Q^2)$ is well known, experiments with deuterium to find $F_2^{\gamma;n}(x,Q^2)$ are much more difficult, due to target nuclear effects. We shall use here the results from [32] which employ both proton and neutron (deuterium) targets in the same experiment, which thus minimizes systematic errors. In [33] this has been combined with the world data [34] and is shown in Fig. 3 in the form of a series of bins at different x values. (Naively, if there were no QCD interactions, the parton model would give a delta-function distribution at $x = 1/3$. This distribution has been considerably washed out here though.) There is a paucity of data for larger x. However $F_2^{\gamma;\rm NS}$ is dropping rapidly to zero, so any error here will not affect the low moments. As shown in the figure, we have simply made a linear extrapolation to $x = 1$, to estimate this region. To find the moments for Eq. (1), we simply need to find the area under the bins weighted with the appropriate x moment, i.e.

$$\frac{1}{f}\int_0^1 dx x^{n-2}F_2^{\gamma;\rm NS}(x,Q_0^2) \approx \frac{1}{f}\sum_{{\rm bins},b}\frac{1}{n-1}(x_{b+1}^{n-1}-x_b^{n-1})F_2^{\gamma;\rm NS}(x_b,Q_0^2) \approx \begin{cases} 0.164(4)(1) & n=2 \\ 0.0289(10)(10) & n=4, \end{cases} \quad (28)$$

where the first error is from the data and the second error is the effect of dropping the estimated last bin. As expected, we see that higher moments are more sensitive to this bin.

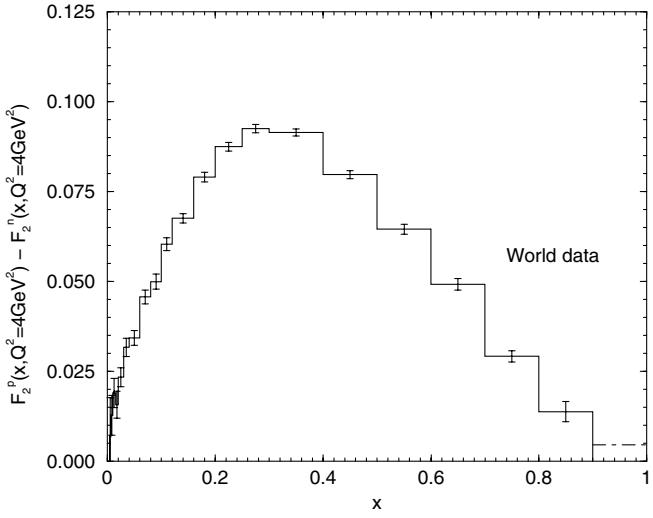

FIG. 3. World experimental data for $F_2^{\gamma;\rm NS}(x,Q_0^2)$ [33] at $Q_0^2 = 4$ GeV^2 in the form of bins, plotted against x using a linear scale. Errors in the bins are also shown. The dot-dashed line is a rough estimate, obtained by a linear extrapolation of the last bin to nought (at $x = 1$).

In Table I we give estimates of $v_n^{\overline{\rm MS}}(Q_0)$, $v_n^{\rm RGI}(Q_0)$, using estimates of the Wilson coefficients and $[\Delta Z_{v_n}^{\overline{\rm MS}}(2\text{ GeV})]^{-1}$ given in the figure caption. These numbers are similar to the quenched results, as can be seen from the caption of Fig. 2 and Table VI. We find that there is good agreement between the two methods, with the lowest moment from MRST or CTEQ being slightly smaller than the experi-

TABLE I. Values of $v_n^{\overline{\rm MS}}(Q_0)$ at $Q_0 = 2$ GeV. The Wilson coefficients (needed for "World") have been computed from Eq. (19), using $\Lambda^{\overline{\rm MS}}|_{n_f=4} = 250(50)$ MeV, giving $E_{F_2;v_2} = 1.018(3)$, $E_{F_2;v_4} = 1.200(30)$, where the error is a reflection of the error in $\Lambda^{\overline{\rm MS}} r_0$. Similarly, to convert to the RGI form, $[\Delta Z_{v_n}^{\overline{\rm MS}}(2\text{ GeV})]^{-1}|_{n_f=4}$ from Eq. (15) again uses $\Lambda_{n_f=4}^{\overline{\rm MS}}$ giving 0.713(40), 0.581(50) and 0.502(54) for $n = 2$, 3 and 4 respectively.

	World	MRST	CTEQ
$v_2^{\overline{\rm MS}}(Q_0)$	0.161(4)	0.157(9)	0.155(17)
$v_2^{\rm RGI}$	0.226(14)	0.220(18)	0.217(27)
$v_3^{\overline{\rm MS}}(Q_0)$		0.0565(26)	0.0551(51)
$v_3^{\rm RGI}$		0.0972(95)	0.095(12)
$v_4^{\overline{\rm MS}}(Q_0)$	0.0241(13)	0.0231(11)	0.0230(23)
$v_4^{\rm RGI}$	0.0480(58)	0.0460(54)	0.0458(67)

mental result. Thus these global fits describe the low moment data well [36]. In the future for definiteness we use the MRST results.

IV. THE LATTICE APPROACH

Euclidean lattice operators [38] are defined by

$$\mathcal{O}^{\Gamma}_{q;\mu_1\cdots\mu_n} = \overline{q}\Gamma_{\mu_1\cdots\mu_i}\overleftrightarrow{D}_{\mu_{i+1}}\cdots\overleftrightarrow{D}_{\mu_n}q, \tag{29}$$

where q is taken to be either a u or d quark and Γ is an arbitrary product of Dirac gamma matrices. In the following the q index will be usually suppressed. We have used the lattice definitions

$$\begin{aligned}\vec{D}_\mu q(x) &= \frac{1}{2a}[U_\mu(x)q(x+a\hat{\mu}) - U^\dagger_\mu(x-a\hat{\mu})q(x-a\hat{\mu})],\\ \overline{q}(x)\tilde{D}_\mu &= \frac{1}{2a}[\overline{q}(x+a\hat{\mu})U^\dagger_\mu(x) - \overline{q}(x-a\hat{\mu})U_\mu(x-a\hat{\mu})],\end{aligned} \tag{30}$$

and $\overleftrightarrow{D} = \frac{1}{2}(\vec{D} - \tilde{D})$. For the operators corresponding to Eq. (3), we shall only need $\Gamma = \gamma$. However for the discussion on mixing under renormalization, we shall also use $\Gamma = \gamma\gamma_5$ and $\sigma\gamma_5$.

A. Choice of lattice operators

We now take the simplifying choices of two momenta $\vec{p} = \vec{0}$ or $\vec{p} = (p_1, 0, 0) \equiv \vec{p}_1$ with p_1 being the lowest nonzero momentum possible on a periodic lattice i.e. $p_1 = 2\pi/L_S$ where the number of sites in each spatial direction is $N_S \equiv L_S/a$. We take our lattice operators as [40]

$$\begin{aligned}\mathcal{O}_{v_{2a}} &= \mathcal{O}^\gamma_{\{14\}},\\ \mathcal{O}_{v_{2b}} &= \mathcal{O}^\gamma_{\{44\}} - \frac{1}{3}(\mathcal{O}^\gamma_{\{11\}} + \mathcal{O}^\gamma_{\{22\}} + \mathcal{O}^\gamma_{\{33\}}),\\ \mathcal{O}_{v_3} &= \mathcal{O}^\gamma_{\{441\}} - \frac{1}{2}(\mathcal{O}^\gamma_{\{221\}} + \mathcal{O}^\gamma_{\{331\}}),\\ \mathcal{O}_{v_4} &= \mathcal{O}^\gamma_{\{1144\}} + \mathcal{O}^\gamma_{\{2233\}} - \frac{1}{2}(\mathcal{O}^\gamma_{\{1133\}} + \mathcal{O}^\gamma_{\{1122\}}\\ &\quad + \mathcal{O}^\gamma_{\{2244\}} + \mathcal{O}^\gamma_{\{3344\}}).\end{aligned} \tag{31}$$

Of course, there are other possibilities. However these will all require nonzero momenta in two directions or suffer from more severe mixing problems. As we shall see, even a nonzero momentum in one direction leads to a strong degradation of the signal and with two nonzero momenta very little signal is observed [41].

The transformation properties under the hypercubic group $H(4)$ are given in Table II [42]. Note, in particular, that the "off-diagonal" $\mathcal{O}_{v_{2a}}$ and "diagonal" operators $\mathcal{O}_{v_{2b}}$ for v_2 belong to different representations, in distinction to the continuum operator. Thus we expect that although these operators should have different lattice artifacts and renormalization factors, in the continuum limit both should lead to the same result—potentially a useful check.

TABLE II. Transformation of the various operators under the hypercubic group $H(4)$ [42], where l represents the dimension of the representation $\tau^{(l)}_k$, k labels different representations of the same dimension and C is the charge conjugation parity of the operator.

Operator	$(\tau^{(l)}_k, C)$
$\mathcal{O}_{v_{2a}}$	$(\tau^{(6)}_3, +)$
$\mathcal{O}_{v_{2b}}$	$(\tau^{(3)}_1, +)$
$\mathcal{O}_{v_3}$	$(\tau^{(8)}_1, -)$
$\mathcal{O}_{v_4}$	$(\tau^{(2)}_1, +)$

B. Mixing of lattice operators

A given operator of engineering dimension $d_\mathcal{O}$ can mix with operators (with the same quantum numbers) of lower dimension, the same dimension or higher dimension. $O(a)$ improvement involves mixing with one dimension higher operators i.e. irrelevant operators where the choice of improvement coefficients is conventionally treated separately to mixing with operators of dimension $\leq d_\mathcal{O}$ i.e. relevant operators. We shall follow this practice here.

1. Operator mixing with additional relevant operators

While for $\mathcal{O}_{v_2}$ there is no mixing with relevant operators, unfortunately for $\mathcal{O}_{v_3}$ and $\mathcal{O}_{v_4}$ there are other relevant operators transforming identically under $H(4)$ which can thus mix with the original operator [42]. Specifically we have [43]

(i) Operators mixing with $\mathcal{O}_{v_3}$:

$$\begin{aligned}\mathcal{O}^{m_1}_{v_3} &= \mathcal{O}^\gamma_{\langle\langle 144\rangle\rangle} - \frac{1}{2}(\mathcal{O}^\gamma_{\langle\langle 133\rangle\rangle} + \mathcal{O}^\gamma_{\langle\langle 122\rangle\rangle}),\\ \mathcal{O}^{m_2}_{v_3} &= \mathcal{O}^{\gamma\gamma_5}_{||432||} + 3\mathcal{O}^{\gamma\gamma_5}_{|432|} = 2\mathcal{O}^{\gamma\gamma_5}_{||342||},\end{aligned} \tag{32}$$

where using the notation of [42] we have defined:

$$\begin{aligned}\mathcal{O}^\Gamma_{|\nu_1\nu_2\nu_3|} &= \mathcal{O}^\Gamma_{\nu_1\nu_2\nu_3} - \mathcal{O}^\Gamma_{\nu_1\nu_3\nu_2} - \mathcal{O}^\Gamma_{\nu_3\nu_1\nu_2}\\ &\quad + \mathcal{O}^\Gamma_{\nu_3\nu_2\nu_1},\\ \mathcal{O}^\Gamma_{||\nu_1\nu_2\nu_3||} &= \mathcal{O}^\Gamma_{\nu_1\nu_2\nu_3} - \mathcal{O}^\Gamma_{\nu_1\nu_3\nu_2} + \mathcal{O}^\Gamma_{\nu_3\nu_1\nu_2} - \mathcal{O}^\Gamma_{\nu_3\nu_2\nu_1}\\ &\quad - 2\mathcal{O}^\Gamma_{\nu_2\nu_3\nu_1} + 2\mathcal{O}^\Gamma_{\nu_2\nu_1\nu_3},\\ \mathcal{O}^\Gamma_{\langle\langle\nu_1\nu_2\nu_3\rangle\rangle} &= \mathcal{O}^\Gamma_{\nu_1\nu_2\nu_3} + \mathcal{O}^\Gamma_{\nu_1\nu_3\nu_2} - \mathcal{O}^\Gamma_{\nu_3\nu_1\nu_2}\\ &\quad - \mathcal{O}^\Gamma_{\nu_3\nu_2\nu_1}.\end{aligned} \tag{33}$$

(ii) Operators mixing with $\mathcal{O}_{v_4}$:

$$\begin{aligned}
\mathcal{O}^{m_1}_{v_4} =& -\mathcal{O}^{\gamma}_{1144} - \mathcal{O}^{\gamma}_{4114} - \mathcal{O}^{\gamma}_{1441} - \mathcal{O}^{\gamma}_{4411} + 2\mathcal{O}^{\gamma}_{1414} \\
&+ 2\mathcal{O}^{\gamma}_{4141} - \mathcal{O}^{\gamma}_{2233} - \mathcal{O}^{\gamma}_{3223} - \mathcal{O}^{\gamma}_{2332} - \mathcal{O}^{\gamma}_{3322} \\
&+ 2\mathcal{O}^{\gamma}_{2323} + 2\mathcal{O}^{\gamma}_{3232} + \frac{1}{2}(+\mathcal{O}^{\gamma}_{1133} + \mathcal{O}^{\gamma}_{3113} \\
&+ \mathcal{O}^{\gamma}_{1331} + \mathcal{O}^{\gamma}_{3311} - 2\mathcal{O}^{\gamma}_{1313} - 2\mathcal{O}^{\gamma}_{3131} \\
&+ \mathcal{O}^{\gamma}_{1122} + \mathcal{O}^{\gamma}_{2112} + \mathcal{O}^{\gamma}_{1221} + \mathcal{O}^{\gamma}_{2211} - 2\mathcal{O}^{\gamma}_{1212} \\
&- 2\mathcal{O}^{\gamma}_{2121} + \mathcal{O}^{\gamma}_{2244} + \mathcal{O}^{\gamma}_{4224} + \mathcal{O}^{\gamma}_{2442} + \mathcal{O}^{\gamma}_{4422} \\
&- 2\mathcal{O}^{\gamma}_{2424} - 2\mathcal{O}^{\gamma}_{4242} + \mathcal{O}^{\gamma}_{3344} + \mathcal{O}^{\gamma}_{4334} \\
&+ \mathcal{O}^{\gamma}_{3443} + \mathcal{O}^{\gamma}_{4433} - 2\mathcal{O}^{\gamma}_{3434} - 2\mathcal{O}^{\gamma}_{4343}), \\
\mathcal{O}^{m_2}_{v_4} =& \frac{1}{2}(+\mathcal{O}^{\gamma\gamma_5}_{1234} - \mathcal{O}^{\gamma\gamma_5}_{3214} - \mathcal{O}^{\gamma\gamma_5}_{1432} + \mathcal{O}^{\gamma\gamma_5}_{3412} \\
&+ \mathcal{O}^{\gamma\gamma_5}_{2143} - \mathcal{O}^{\gamma\gamma_5}_{4123} - \mathcal{O}^{\gamma\gamma_5}_{2341} + \mathcal{O}^{\gamma\gamma_5}_{4321} - \mathcal{O}^{\gamma\gamma_5}_{1324} \\
&+ \mathcal{O}^{\gamma\gamma_5}_{2314} + \mathcal{O}^{\gamma\gamma_5}_{1423} - \mathcal{O}^{\gamma\gamma_5}_{2413} - \mathcal{O}^{\gamma\gamma_5}_{3142} + \mathcal{O}^{\gamma\gamma_5}_{4132} \\
&+ \mathcal{O}^{\gamma\gamma_5}_{3241} - \mathcal{O}^{\gamma\gamma_5}_{4231}) + \mathcal{O}^{\gamma\gamma_5}_{1243} - \mathcal{O}^{\gamma\gamma_5}_{4213} - \mathcal{O}^{\gamma\gamma_5}_{1342} \\
&+ \mathcal{O}^{\gamma\gamma_5}_{4312} + \mathcal{O}^{\gamma\gamma_5}_{2134} - \mathcal{O}^{\gamma\gamma_5}_{3124} - \mathcal{O}^{\gamma\gamma_5}_{2431} + \mathcal{O}^{\gamma\gamma_5}_{3421} \\
\mathcal{O}^{m_3}_{v_4} =& \frac{1}{2} i(+\mathcal{O}^{\sigma\gamma_5}_{2413} - \mathcal{O}^{\sigma\gamma_5}_{2431} + \mathcal{O}^{\sigma\gamma_5}_{1324} - \mathcal{O}^{\sigma\gamma_5}_{1342} \\
&- \mathcal{O}^{\sigma\gamma_5}_{3412} + \mathcal{O}^{\sigma\gamma_5}_{3421} - \mathcal{O}^{\sigma\gamma_5}_{1234} + \mathcal{O}^{\sigma\gamma_5}_{1243} + 2\mathcal{O}^{\sigma\gamma_5}_{2314} \\
&- 2\mathcal{O}^{\sigma\gamma_5}_{2341} + 2\mathcal{O}^{\sigma\gamma_5}_{1423} - 2\mathcal{O}^{\sigma\gamma_5}_{1432}). \qquad (34)
\end{aligned}$$

$\mathcal{O}^{m_3}_{v_4}$ is an operator of one lower dimension and opposite chiral properties than the other $\mathcal{O}_{v_4}$ operators. It is also possible that four-fermion operators can mix: we shall not consider this here though.

In Appendix B we illustrate, by an example for $\mathcal{O}_{v_4}$, how $\mathcal{O}^{m_1}_{v_4}$ and $\mathcal{O}^{m_2}_{v_4}$ can be derived. The other mixing operators follow by similar considerations.

While this list contains all operators allowed by group theoretical arguments, not all the operators contribute, see Sec. V.

2. Operator improvement

As we are using $O(a)$ improved fermions, to achieve the elimination of $O(a)$ terms in matrix elements, the corresponding operators must also include additional irrelevant terms, with coefficients appropriately chosen. Presently only for the lowest moment (i.e. v_2) are these extra operators known [44]. In this case we should replace the operators $\mathcal{O}^{\gamma}_{\mu\nu}$ in v_{2a} or v_{2b} by [45]

$$\begin{aligned}
\mathcal{O}^{\gamma}_{\mu\nu} \to (1 + am_q c_0)\mathcal{O}^{\gamma}_{\mu\nu} + \sum_{i=1}^{3} ac_i \mathcal{O}^{(i)}_{\mu\nu} \equiv (1 + am_q c_0)\overline{q}\gamma_\mu \overleftrightarrow{D}_\nu q + ac_1\left[i\sum_\lambda \overline{q}\,\sigma_{\mu\lambda}\overleftrightarrow{D}_{[\nu}\overleftrightarrow{D}_{\lambda]}q \right] \\
+ ac_2[-\overline{q}\overleftrightarrow{D}_{\{\mu}\overleftrightarrow{D}_{\nu\}}q] + ac_3\left[i\sum_\lambda \partial_\lambda(\overline{q}\sigma_{\mu\lambda}\overleftrightarrow{D}_\nu q) \right], \qquad (35)
\end{aligned}$$

where m_q is the bare lattice quark mass, related to the hopping parameter κ by

$$am_q = \frac{1}{2}\left(\frac{1}{\kappa} - \frac{1}{\kappa_c}\right). \qquad (36)$$

So we see that there are potentially four additional improvement operators, defined in Eq. (35) together with four unknown improvement coefficients, $c_i(g_0)$, for $i = 0, \ldots, 3$, which are functions of the gauge coupling constant g_0. The $i = 3$ operator only contributes for nonforward matrix elements, which are not considered here and so this term may be dropped. Also ultimately c_0 will not concern us as we are only interested in the result in the chiral limit. For on-shell matrix elements the equation of motion may be used to eliminate one of the improvement terms, for example, we can choose c_0 and c_1 as linear functions of c_2. First order perturbation theory gives the relation between the improvement coefficients for v_{2b} of

$$\begin{aligned}
c_0 &= 1 - c_2 + \frac{g_0^2 C_F}{16\pi^2}(17.203\,77 - 8.690\,45c_2) + O(g_0^4), \\
c_1 &= c_2 + O(g_0^2), \qquad (37)
\end{aligned}$$

where $C_F = 4/3$. A similar expression holds for v_{2a} [44]. With these values of the improvement coefficients, $O(a)$ corrections to the v_{2b} matrix element have been eliminated, at least to lowest order perturbatively. We see that we cannot determine the $O(g_0^2)$ term of c_1 because it is the coefficient of an operator that vanishes at tree level. Nonperturbatively the improvement coefficients have not yet been determined. However one might suspect, that choosing $c_2 = 0$ also gives in this case a small c_1 coefficient.

For higher moments, the bases for the improved operators become increasingly cumbersome, as not only would we expect more possible irrelevant operators built from the original operator together with an additional covariant derivative, but also four-fermion operators may play a

role. This does not necessarily detract from the original operator though, because we can always attempt to make a continuum extrapolation in a rather than a^2 if we cannot motivate why the irrelevant matrix elements are small.

C. Determining the matrix element

Matrix elements are determined from the ratio R of three-point to two-point correlation functions,

$$R(t,\tau;\vec{p};\mathcal{O}) = \frac{C_{\frac{1}{2}(1+\gamma_4)}(t,\tau;\vec{p};\mathcal{O})}{C_{\frac{1}{2}(1+\gamma_4)}(t;\vec{p})}, \tag{38}$$

where $C_{\frac{1}{2}(1+\gamma_4)}(t;\vec{p})$ is the unpolarized nucleon two-point function with a source at time 0 and sink at time t, while the also unpolarized three-point function $C_{\frac{1}{2}(1+\gamma_4)}(t,\tau;\vec{p};\mathcal{O})$ has an operator insertion at time τ. If we consider a region $0 \ll \tau \ll t \lesssim \frac{1}{2}L_T$ (i.e. well inside the nucleon branch of the propagator) where L_T is the temporal extension of the lattice, then the transfer matrix formalism upon projecting out the ground state nucleon leads to

$$\begin{aligned} R(t,\tau;\vec{p};\mathcal{O}_{v_{2a}}) &= ip_1 v_{2a}, \\ R(t,\tau;\vec{p};\mathcal{O}_{v_{2b}}) &= -\frac{E_{\vec{p}}^2 + \frac{1}{3}\vec{p}^2}{E_{\vec{p}}} v_{2b}, \\ R(t,\tau;\vec{p}_1;\mathcal{O}_{v_3}) &= -ip_1 E_{\vec{p}_1} v_3, \\ R(t,\tau;\vec{p}_1;\mathcal{O}_{v_4}) &= p_1^2 E_{\vec{p}_1} v_4, \end{aligned} \tag{39}$$

for the bare matrix elements v_n. For some more details see e.g. [3,39]. In general we would expect that the best signals with the smallest noise are seen for zero momentum. We take, and have numerically checked [3,46], that the standard dispersion relation $E_{\vec{p}}^2 = m_N^2 + \vec{p}^2$ holds for $\vec{p} = \vec{p}_1$.

The nucleon three-point correlation function is depicted in Fig. 4. While the two-point correlation function consists of one diagram, for the three-point correlation function we have two diagrams—a ''quark line connected'' contribution, and a ''quark line disconnected'' contribution, left and right diagrams in Fig. 4 respectively. This is not the usual field theoretic splitting of the Green's function into connected and disconnected diagrams. As quarks can travel backwards in time as well as forwards, we would expect that the quark line connected term would also give a contribution to the antiquark parton density defined in Eq. (23). As quark line disconnected diagrams can, by definition, only interact with the hadron via the exchange of gluons then the numerical results suffer from large short distance (i.e. ultraviolet) fluctuations. So a very large number of configurations is required, which is very expensive in computer time. We have not computed this term here. To cancel any effects of these disconnected terms, if flavor $SU_F(2)$ symmetry is a good symmetry, it is sufficient to consider NS matrix elements, i.e. the u quark matrix element minus the d quark matrix element. In Appendix C , for completeness, we give explicit expressions for the relevant three-point correlation functions.

A further class of disconnected terms are given by gluon matrix elements. Again on the lattice these are difficult to compute, see [47], but again they cancel upon considering nonsinglet matrix elements.

All these gluonic or sea-quark effects are concentrated at small x, and thus for higher moments, i.e. $n = 3$ or 4, are naturally suppressed anyway. Thus disconnected contributions may be less significant, so that the computation of singlet matrix elements is then more reliable and we can consider just a u or d operator matrix element. Although the quenched approximation does not handle the sea-quarks correctly, we might also expect for these higher moments that quenching has less effect. But these statements are hard to quantify, and as this is all less likely to be the case for the lowest moment anyway, we shall consider mainly the nonsinglet results here.

D. Raw results for lattice matrix elements

We now discuss our raw numerical results for the lattice operators and the numerical significance of the additional improvement operators and/or additional relevant operators to the nucleon matrix element. Since our original publication [3] which employed unimproved Wilson fer-

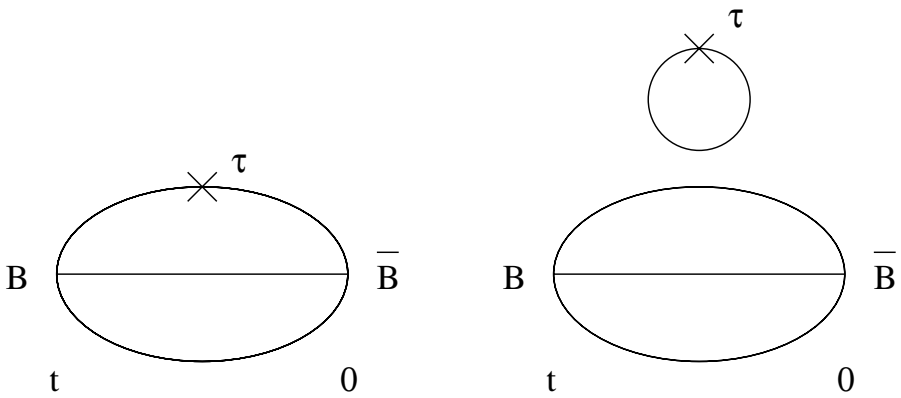

FIG. 4. The 3-point quark correlation function for a baryon.

mions at $\beta = 6.0$, we have used quenched configurations with $O(a)$ improved fermions: thus for the action we take the standard Wilson gluon action, while for the fermion propagator the standard Wilson fermion is used together with a "clover" improvement term. The nonperturbative coefficient c_{sw} of the improvement term was taken from [48]. This means that on-shell quantities, such as masses, have only $O(a^2)$ discretization effects. Simulations were performed at three $\beta \equiv 6/g_0^2$ values, $\beta = 6.0, 6.2, 6.4$. At each β value, four or more quark masses (degenerate in u and d) were used, at each mass a statistic of several hundred configurations was generated. Antiperiodic fermion boundary conditions were taken in the time direction and periodic in the remaining spatial directions. Further details of our runs are given in Appendix D in Table VII.

To improve the overlap with the nucleon we employed Jacobi smearing, and used a nonrelativistic (NR) projection of the nucleon. Jacobi smearing is described, for example, in [39], where a hopping parameter (κ_s) expansion of order n_s of a Wilson fermion operator restricted to a time plane smears out the original point quark source. For $\beta = 6.0$, 6.2 and 6.4 we use $(\kappa_s, n_s) = (0.21, 50)$, $(0.21, 100)$ and $(0.21, 150)$ respectively giving a root-mean-square radius of about 0.4 fm, a reasonable fraction of the nucleon radius $\sim$0.8 fm. The NR projection, where in our Dirac gamma matrix representation only the upper two components of the spinors are used, is briefly described in [46] and Appendix C.

The nucleon sink positions were chosen as $t = 13a$, $17a$ and $23a$ for $\beta = 6.0$, 6.2 and 6.4 respectively and the fit ranges in τ were taken as $[4a, 9a]$, $[6a, 11a]$ and $[7a, 16a]$. These fit range values were not so critical, but allowed a splitting of the range $[0, t]$ into roughly three equal parts each piece being roughly the same physical size.

In Appendix D, the data are presented in Tables VII, VIII, IX, X, XI, XII, and XIII giving separately the u and d contributions. As discussed previously though, as the disconnected diagrams have not been computed, the most physically significant result is for the nonsinglet matrix elements. Thus we have also repeated the data analysis directly for these matrix elements. This allows, in particular, a better estimation of the error. These numbers are also given in the tables in Appendix D. As an example of the typical ratios obtained, in Fig. 5 we show $R(17a, \tau; \vec{0}; \mathcal{O}_{\upsilon_{2b}})$, $R(17a, \tau; \vec{p}_1; \mathcal{O}_{\upsilon_{2b}})$, $R(17a, \tau; \vec{p}_1; \mathcal{O}_{\upsilon_{2a}})$, left to right pictures, respectively, for $\beta = 6.2$ and $\kappa = 0.1344$. We seek a plateau in the region $0 \ll \tau/a \ll 17$. The region chosen is denoted by vertically dashed lines. Clearly the operator corresponding to υ_{2b} delivers a better signal for the ratio than the operator for υ_{2a}, although even for the υ_{2b} operator we see that it is better to choose zero momentum rather than nonzero momentum.

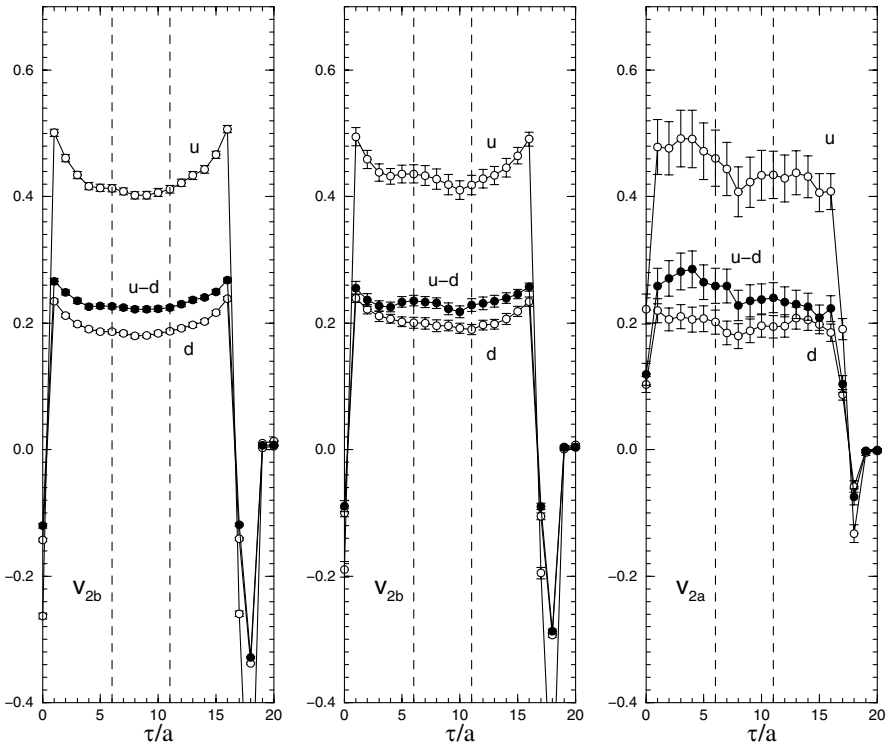

FIG. 5. υ_{2b} versus τ/a from the ratios $R(17a, \tau; \vec{0}; \mathcal{O}_{\upsilon_{2b}})$, $R(17a, \tau; \vec{p}_1; \mathcal{O}_{\upsilon_{2b}})$, Eq. (39) left and middle pictures, respectively, and υ_{2a} from $R(17a, \tau; \vec{p}_1; \mathcal{O}_{\upsilon_{2a}})$, right picture for $\mathcal{O}^{(u)}$, $\mathcal{O}^{(d)}$ (open circles) and NS (i.e. $\mathcal{O}^{(u)} - \mathcal{O}^{(d)}$) (solid circles) for $\beta = 6.2$ at $\kappa = 0.1344$. The chosen fit intervals are denoted by vertical dotted lines.

In this section, as mentioned before, we wish to merely estimate the numerical significance of extra operators, as described in Sec. IV B. To this end, as mixing coefficients and improvement coefficients are not so well known, we make a series of plots either comparing the ratio of the additional improvement operators [i.e. the υ_n constructed using the $\mathcal{O}^{(i)}$ in Eqs. (35) and (39)] to the original operator,

$$r^i_{\upsilon_n} \equiv \frac{\upsilon^i_n}{\upsilon_n}, \tag{40}$$

or compare directly the $\upsilon_n^{m_i}$ derived from Eqs. (32) and (34) to the original operator, υ_n. It will be seen in Sec. V that the ratios $r^i_{\upsilon_n}$ have a physical significance.

These are all plotted against the square of the pseudoscalar meson mass. As $(am_{ps})^2$ is proportional to the quark mass for small quark mass, using this quantity avoided the necessity of first determining the critical hopping parameter κ_c in the quark mass definition, Eq. (36). We have thus used the dimensionless extrapolation parameter $(r_0/a)^2 \times (am_{ps})^2 \equiv (r_0 m_{ps})^2$, with r_0/a being given by the Padé formula in [35]. Using $r_0 = 0.5$ fm this enables us to get an idea of how close our simulation points are to the chiral limit in physical units.

1. υ_2

In Fig. 6 we show ar^1_{2a}, ar^2_{2a} plotted against $(r_0 m_{ps})^2$, together with a linear chiral extrapolation. Also plotted is the approximate value of the pseudoscalar mass corresponding to degenerate u/d or s quark masses. So we see that our simulation runs over the range from about 2 to 3 times the strange quark mass to a little under the strange quark mass. It is also noticeable that we are a long way from simulating with a light u/d quark mass—in fact within our errors, there is no difference between linearly extrapolating to the chiral limit or to the pion mass. Nevertheless we see that the effect of any extra improvement operators is likely to be small, of the order of a few percent.

The above result is for the off-diagonal operator, which needs a nonzero momentum in its evaluation. This figure is to be compared with the result using the diagonal operator, which advantageously may use zero momentum. In Fig. 7 we show this result. In comparison with the previous picture, the errors are considerably reduced, as $\vec{p} = \vec{0}$

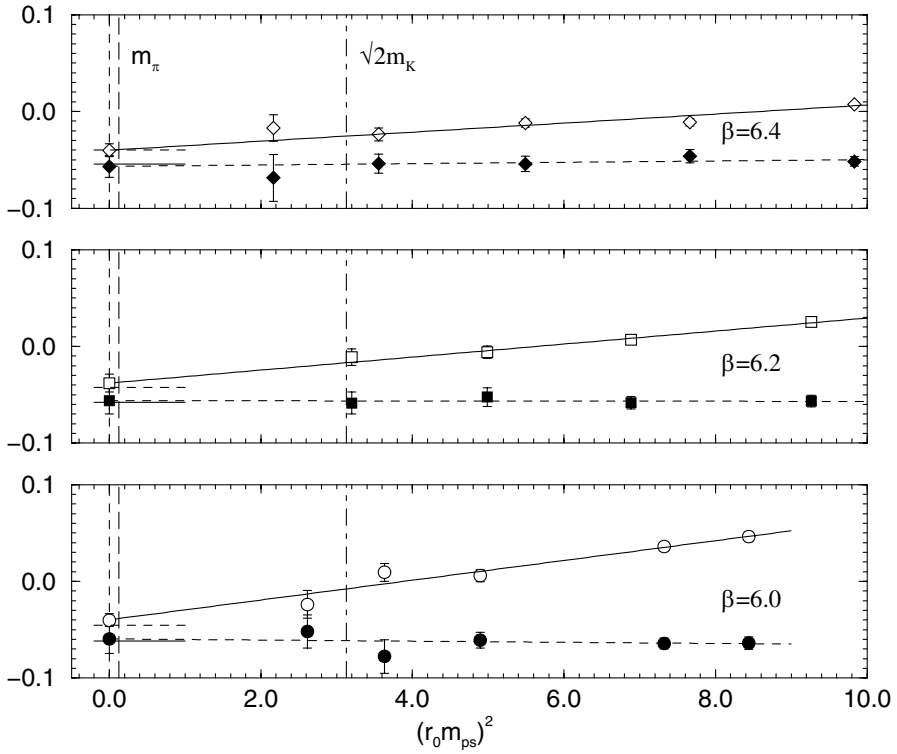

FIG. 6. The ratios $ar^1_{\upsilon_{2a}}$, $ar^2_{\upsilon_{2a}}$ plotted against $(r_0 m_{ps})^2$. The solid symbols at nonzero $(r_0 m_{ps})^2$ are for $ar^1_{\upsilon_{2a}}$, while the open symbols denote $ar^2_{\upsilon_{2a}}$. Also shown is a linear chiral extrapolation (dashed line and solid line for $ar^1_{\upsilon_{2a}}$, $ar^2_{\upsilon_{2a}}$ respectively). The result in the chiral limit $(r_0 m_{ps})^2 \equiv 0$ is also indicated, again using solid, open symbols for $ar^1_{\upsilon_{2a}}$, $ar^2_{\upsilon_{2a}}$ respectively. The small horizontal lines around the chiral limit represent the perturbative estimate, see Sec. V, with again a dashed line corresponding to $ar^1_{\upsilon_{2a}}$ and a solid line to $ar^2_{\upsilon_{2a}}$ respectively. The rough values of the (hypothetical) pseudoscalar mass composed from u/d or s quark masses, evaluated from m_π and m_K respectively are shown as dashed vertical lines.

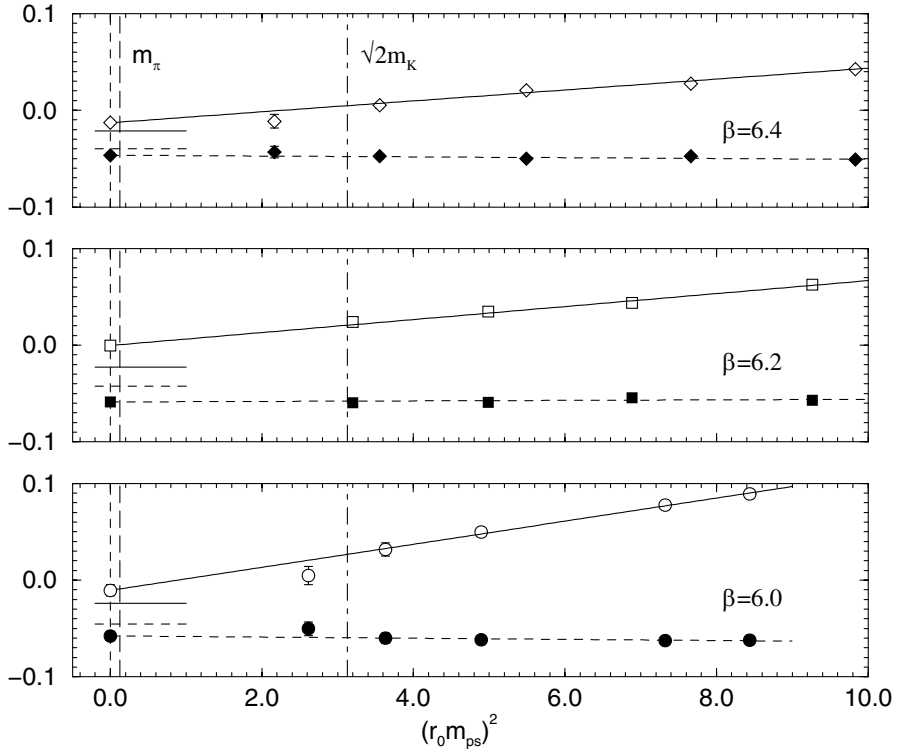

FIG. 7. $ar^1_{v_{2b}}$, $ar^2_{v_{2b}}$ versus $(r_0 m_{ps})^2$ for $\beta = 6.0$, 6.2 and 6.4 with $\vec{p} = \vec{0}$. The notation is the same as for Fig. 6.

is used (see also Fig. 5). Indeed using $\vec{p} = \vec{p}_1$ (see Appendix D for the numbers) we see that the errors grow again, although they are never as large as for the off-diagonal operator. Again, the effect of any extra improvement operators is small.

Thus all the numerical results and linearly chirally extrapolated results for $ar^i_{v_2}$ look small giving $av^i_2 \lesssim 5\%$ of v_2 and some indeed are consistent with zero. From Eq. (37), we see that if we choose $c_2 = 0$ then $c_1 \equiv O(g_0^2)$, which is likely to be small. So these results lead numerically to a small additional improvement term and we conclude that the effect of these terms in v_2 can be assumed to be negligible.

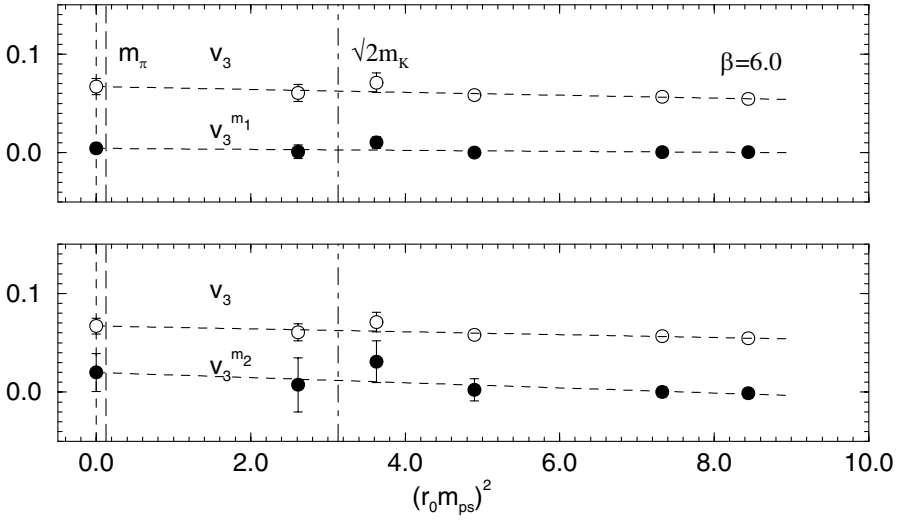

FIG. 8. Mixing terms, $v_3^{m_1}$, $v_3^{m_2}$ for $\beta = 6.0$ (solid circles). Also shown is v_3 (open circles) and linear chiral extrapolations. The same notation as for Fig. 6.

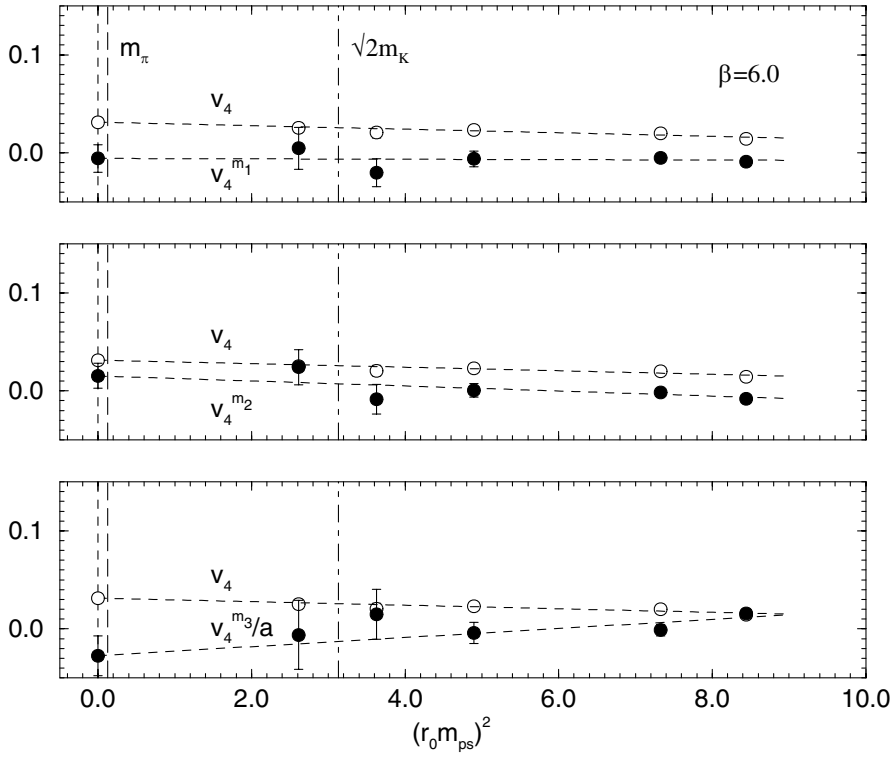

FIG. 9. v_4 for $\beta = 6.0$ together with $v_4^{m_1}$, $v_4^{m_2}$, $v_4^{m_3}/a$. The notation is the same as for Fig. 6.

2. v_3 and v_4

We now present our results for the higher moments. In Fig. 8 we show $v_3^{m_1}$, $v_3^{m_2}$ for $\beta = 6.0$, together with a linear chiral extrapolation. Also shown, for comparison, is the operator v_3. As expected while the magnitude of the noise has increased in comparison with v_2 and scatters more, an acceptable signal is still seen. We find a clear separation between v_3 and the mixing operators (indeed they are consistent with ≈ 0). For higher β values, the data fluctuates more and it becomes more difficult to disentangle the results.

In Fig. 9 we plot $v_4^{m_1}$, $v_4^{m_2}$, $v_4^{m_3}/a$, together with v_4. While there is a reasonable signal for v_4, the improvement terms fluctuate a lot (again becoming worse the higher the value of β is). Indeed even finding a plateau for the ratio $R(t, \tau; \vec{p}; \mathcal{O})$ in Eq. (38) becomes problematical. It would appear that while, numerically, $v_4^{m_1}$, $v_4^{m_2}$ are small in comparison with v_4 the situation is less clear-cut for $v_4^{m_3}/a$, although it is only including the heaviest quark mass point here that leads to a nonzero result in the chiral limit. Nevertheless as most of the results for the mixing terms are much smaller than v_4 we shall ignore any effects from them.

V. OPERATOR RENORMALIZATION

A lattice operator (or matrix element) must, in general, be renormalized. Again, we shall discuss mixing and $O(a)$ operator improvement separately.

A. Operator mixing and renormalization

For operator mixing we can generally write

$$\mathcal{O}_i^S(M) = \sum_j Z^S_{\mathcal{O}_i \mathcal{O}_j;\mathrm{LAT}}(M, a)\mathcal{O}_j^{\mathrm{LAT}}(a), \qquad (41)$$

working in a scheme S at scale M. So we must determine a matrix of renormalization constants. While for those mixing operators of the same dimension as the original operator, low order perturbation results can be calculated at least in principle, for lower-dimensional operators this is not reliable: the renormalization constant is proportional to a positive power of $1/a$ and nonperturbative terms may contribute. For the higher moments, both cases occur. For v_3 both the m_1 and m_2 operators have the same dimension as the original operator, and for v_4, m_1 and m_2 have the same dimension, while m_3 is dimension one lower, see Eqs. (32) and (34).

1. Renormalization and relevant operator mixing

In this section we want to make a few comments on the operator mixings seen for the operators we use in this paper, and contrast them with the mixing problems for the operators which we rejected.

In the continuum, symmetry under $O(4)$ in the Euclidean case or the Lorentz group in the Minkowski case imposes strong constraints on which operators can have unpolarized forward nucleon matrix elements. One way of stating the rule is to say that the only operators with nonzero unpo-

larized forward matrix elements are those which have $J^P = 0^+$ in the nucleon's rest frame. The quantum numbers $J^P = 0^+$ only occur in $O(4)$ representations of the form $U_+^{(\mu,\mu)}$ in the notation of [49]. Considering the $O(4)$ classification of the operators is relevant, because although bare lattice matrix elements only respect the symmetries of the hypercubic group $H(4)$, renormalized quantities should respect the full continuum symmetry group.

If we look at the mixing operators for v_3, as listed in Eq. (32), we see that under the hypercubic group they transform exactly the same way as v_3, but under $O(4)$ they both transform according to the representation $U^{(1/2,3/2)}$, which means that by the continuum symmetries their renormalized forward matrix elements must be zero,

$$v_3^{m_i;S} = 0. \tag{42}$$

Similarly the mixing operators for v_4, listed in Eq. (34), all transform according to the $U^{(0,2)}$ representation of $O(4)$, so

$$v_4^{m_i;S} = 0. \tag{43}$$

If the renormalized mixing operators all have the value zero, we can express all the bare lattice matrix elements in terms of a single renormalized matrix element. Again taking the v_3 system as our exemplar,

$$v_3 = (Z^{-1})_{v_3;v_3} v_3^S, \qquad v_3^{m_i} = (Z^{-1})_{v_3^{m_i};v_3} v_3^S, \tag{44}$$

temporarily dropping, for clarity, other arguments of the renormalization constants. Since all the lattice matrix elements are multiples of the physically interesting matrix element v_3^S we have a choice in how we calculate the renormalized matrix element. We could add up all the terms in Eq. (41), assuming that we know the mixing Zs well enough. Or we could equally well just invert Eq. (44) and calculate v_3^S from v_3 alone,

$$v_3^S = \frac{v_3}{(Z^{-1})_{v_3;v_3}}. \tag{45}$$

If we calculate the renormalized matrix element from v_3 alone, we should really renormalize by dividing by $(Z^{-1})_{v_3;v_3}$ instead of multiplying by $Z_{v_3;v_3}$. In practice the difference is minor,

$$\frac{1}{(Z^{-1})_{v_3;v_3}} \approx Z_{v_3;v_3} - \sum_i \frac{Z_{v_3;v_3^{m_i}} Z_{v_3^{m_i};v_3}}{Z_{v_3^{m_i};v_3^{m_i}}}. \tag{46}$$

The difference involves the product of two off-diagonal Zs, and so is $\mathcal{O}(g_0^4)$ in perturbation theory, see Sec. V B 1, and is still likely to be tiny nonperturbatively. Note that Eq. (46) tells us that mixing with lower dimension operators is no more dangerous than mixing with operators of the same dimension, in a case like this where the lower-dimensional operator has zero renormalized matrix element. This is because the product $Z_{v_3;v_3^{m_i}} Z_{v_3^{m_i};v_3}$ is dimensionless, even when $v_3^{m_i}$ is a lower-dimensional operator.

We conclude that for the operators we use in this paper, mixing is relatively benign, because continuum symmetry says that the renormalized mixing matrix elements are zero.

Finally we want to contrast this with an example where this is not so, to show the importance of choosing the operators carefully. We could have tried to measure v_3 with the operator

$$O^\gamma_{\{444\}} - O^\gamma_{\{411\}} - O^\gamma_{\{422\}} - O^\gamma_{\{433\}}. \tag{47}$$

This however would have much worse mixing problems because it mixes with operators which are allowed to have nucleon matrix elements in the continuum, for example

$$O^\gamma_{\{444\}} + O^\gamma_{\{411\}} + O^\gamma_{\{422\}} + O^\gamma_{\{433\}}, \qquad \overline{q}\gamma_4 q \quad \text{or} \quad \overline{q}\overleftrightarrow{D}_4 q. \tag{48}$$

Now we would not have the option of using Eq. (45), we would have to use Eq. (41) and to get a reliable answer we would need to know the off-diagonal elements of Z accurately, especially those for the operators of lower dimension. This is why we rejected the operator Eq. (47), and why we think that the mixing operators of Sec. IV B 1 are not a problem.

2. Renormalization and operator improvement

For $O(a)$ improved operators, we know that the operator depends linearly on the improvement coefficients c_i as shown in Eq. (35). Our experience from extensive perturbative calculations in [44] leads us to expect a result of the form

$$\Lambda^{\mathrm{LAT}}_{\mathcal{O}^i} = \frac{\zeta_i}{a} \Lambda^{\mathrm{LAT}}_{\mathcal{O}} + O(a^0), \tag{49}$$

where Λ^{LAT} is the lattice amputated three-point Green's function, $\mathcal{O}$ is the relevant operator, and $\mathcal{O}^i$ are the improvement or irrelevant operators in Eq. (35). Note that there is nothing forbidding the improvement terms mixing with $\mathcal{O}/a$. The only case that does not mix is the mass improvement term $am_q\mathcal{O}$,

$$am_q \Lambda^{\mathrm{LAT}}_{\mathcal{O}} = O(a). \tag{50}$$

We can calculate the c_i dependence of the renormalization constant by requiring that the renormalized $\Lambda^S_{\mathcal{O}}$ should be independent of the c_i at leading order in a and thus we have, temporarily suppressing additional M, a arguments

$$\begin{aligned}
\Lambda^S_{\mathcal{O}} &= \frac{Z^S_{\mathcal{O};\mathrm{LAT}}(\{c_i\})}{Z^S_{q;\mathrm{LAT}}} \left[\Lambda^{\mathrm{LAT}}_{\mathcal{O}} + \sum_{i\neq 0} c_i \zeta_i \Lambda^{\mathrm{LAT}}_{\mathcal{O}} + O(a) \right] \\
&= \frac{Z^S_{\mathcal{O};\mathrm{LAT}}(\{c_i\})}{Z^S_{q;\mathrm{LAT}}} \left(1 + \sum_{i\neq 0} c_i \zeta_i \right) \Lambda^{\mathrm{LAT}}_{\mathcal{O}} + O(a) \\
&\equiv \frac{Z^S_{\mathcal{O};\mathrm{LAT}}(\{0\})}{Z^S_{q;\mathrm{LAT}}} \Lambda^{\mathrm{LAT}}_{\mathcal{O}} + O(a),
\end{aligned} \tag{51}$$

and so we can write

$$\mathcal{O}^S(M) = Z^S_{\mathcal{O};\mathrm{LAT}}(M, a, \{c_i\})\mathcal{O}(a, \{c_i\}, c_0) \equiv \frac{Z^S_{\mathcal{O};\mathrm{LAT}}(M, a, \{0\})}{1 + \sum_{i\neq 0} c_i\zeta_i}\mathcal{O}(a; \{c_i\}, c_0), \quad (52)$$

where the coefficients $\zeta_i \equiv \zeta_i(g_0) = O(g_0^2)$ have to be determined. Again in [44] we have given these coefficients for v_{2a} and v_{2b} in lowest order perturbation theory using Eq. (49). [Requiring $O(a)$ improvement also determines the c_i as given, for example, in Eq. (37).] Numerically these coefficients turn out to be quite small. In addition to the perturbative ζ_i's we can also find nonperturbative ζ_i's by looking at our measured nucleon matrix elements and requiring that the improved result and the unimproved agree to leading order in a, i.e. from Eq. (52) we have

$$\frac{v_n + a\sum_{i\neq 0} c_i v_n^i}{1 + \sum_{i\neq 0} c_i\zeta_i} = v_n + O(a), \quad (53)$$

giving, cf. Eq. (49),

$$v_n^i = \frac{1}{a}\zeta_i v_n + O(a^0), \quad (54)$$

or

$$\zeta_i = a r^i_{v_n}|_{am_q=0} + O(a), \quad (55)$$

see Eq. (40). We would expect ζ_i calculated in this way to agree up to ambiguities of $O(a)$. In Figs. 6 and 7 we compare the values for ζ_i obtained this way with the 1-loop perturbative results from [44], shown in the figures as short horizontal lines about the chiral limit. Although, not unexpectedly, there are differences to the perturbative result, the ratios remain small. From Eq. (37) we may choose $c_2 = 0$ and so $c_1 = O(g_0^2)$ is also small. Hence the change in the denominator of Eq. (52) from 1 is small and so does not change the renormalization constant perceptibly. Thus we shall ignore any small effects here, just as in Sec. IV D where our conclusion was that we could numerically drop the $O(a)$ improvement terms.

B. Determining the renormalization constants

To define the renormalization constants a renormalization procedure must be prescribed. Often the renormalization constants are defined first in a momentum of MOM scheme by computing the (Landau) gauge fixed two-quark Green's function with one operator insertion and setting

$$[\Lambda^{\mathrm{MOM}}_{\mathcal{O}}]_{p^2=\mu_p^2} = \Lambda^{\mathrm{MOM}}_{\mathcal{O}}|_{\mathrm{BORN}}, \quad (56)$$

where $\Lambda^{\mathrm{MOM}}_{\mathcal{O}}$ is the renormalized amputated three-point Green's function for the operator $\mathcal{O}$. The rhs of this equation is the tree level value (or Born approximation) of the amputated function. This definition may be used for perturbative computations, see e.g. [50], where $[\Lambda]$ in Eq. (56) means that, in a given basis, we drop terms not proportional to the Born term. This may be modified, using a trace condition for the definition of the renormalization constants, to give the alternative RI′-MOM scheme [51] which is also suitable for nonperturbative calculations. A discussion and some results for this nonperturbative method will be given in Sec. V B 3.

The resulting $Z^{\mathrm{MOM}}_{\mathcal{O}_i;\mathrm{LAT}}$ may, if wished, be converted to another scheme, i.e. to the $\overline{\mathrm{MS}}$ scheme using Eq. (12) where $Z^{\overline{\mathrm{MS}}}_{\mathcal{O}_i;\mathcal{O}_j;\mathrm{MOM}}(\mu, \mu_p)$ is perturbatively calculable. The $\overline{\mathrm{MS}}$ scheme is particularly convenient, as the renormalization constants are independent of the gauge fixing condition chosen.

Unfortunately, the definition given in Eq. (56) has its limitations: $\mathcal{O}^{m_2}_{v_3}$, $\mathcal{O}^{m_1}_{v_4}$ $\mathcal{O}^{m_2}_{v_4}$ and $\mathcal{O}^{m_3}_{v_4}$ all have vanishing Born matrix elements between quark states as they involve commutators of covariant derivatives. So in order to be able to compute the renormalization constants we would have to consider more general Green's functions (e.g. quark-gluon). At present we must simply ignore this problem though.

After these more general remarks, we shall now give various results for the renormalization constants for one-loop perturbation theory, tadpole-improved renormalization-group-improved boosted perturbation theory (TRB-PT) and finally a nonperturbative determination of the relevant constants.

1. Perturbation theory

One-loop perturbation theory [52] yields [54]

$$Z^S_{\mathcal{O}_i;\mathcal{O}_j;\mathrm{LAT}}(M, a) = \delta_{\mathcal{O}_i;\mathcal{O}_j} + g_0^2[d_{\mathcal{O}_i;\mathcal{O}_j;0}\ln(aM) - B^S_{\mathcal{O}_i;\mathcal{O}_j}(c_{sw})] + O(g_0^4). \quad (57)$$

In the $\overline{\mathrm{MS}}$ scheme ($M \equiv \mu$) we have [44]

$$B^{\overline{\mathrm{MS}}}_{v_{2a}}(c_{sw}) = \frac{C_F}{(4\pi)^2}(1.279\,59 - 3.872\,97c_{sw} - 0.678\,26c_{sw}^2),$$
$$B^{\overline{\mathrm{MS}}}_{v_{2b}}(c_{sw}) = \frac{C_F}{(4\pi)^2}(2.561\,84 - 3.969\,80c_{sw} - 1.039\,73c_{sw}^2), \quad (58)$$

with $C_F = 4/3$. The calculations have been extended by S. Capitani [55] to now include $B^{\overline{\mathrm{MS}}}_{v_3}$, $B^{\overline{\mathrm{MS}}}_{v_4}$. For $\mathcal{O}_{v_3}$, off-diagonal elements in Eq. (57) have also been computed

$$
\begin{aligned}
B^{\overline{\rm MS}}_{v_3}(c_{sw}) &= \frac{C_F}{(4\pi)^2}(-12.127\,40 - 2.921\,69c_{sw} \\
&\quad - 0.981\,66c_{sw}^2), \\
B^{\overline{\rm MS}}_{v_3;v_3^{m_1}}(c_{sw}) &= \frac{C_F}{(4\pi)^2}(-0.368\,48 - 0.032\,760c_{sw} \\
&\quad + 0.029\,137c_{sw}^2), \\
B^{\overline{\rm MS}}_{v_3^{m_1}}(c_{sw}) &= \frac{C_F}{(4\pi)^2}(-14.851\,57 - 2.152\,28c_{sw} \\
&\quad - 1.707\,41c_{sw}^2), \\
B^{\overline{\rm MS}}_{v_3^{m_1};v_3}(c_{sw}) &= \frac{C_F}{(4\pi)^2}(-3.306\,05 + 0.333\,35c_{sw} \\
&\quad - 0.370\,50c_{sw}^2).
\end{aligned} \tag{59}
$$

For $B^{\overline{\rm MS}}_{v_4}$ we have [56]

$$
\begin{aligned}
B^{\overline{\rm MS}}_{v_4}(c_{sw}) &= \frac{C_F}{(4\pi)^2}(-25.503\,03 - 2.417\,88c_{sw} \\
&\quad - 1.128\,26c_{sw}^2).
\end{aligned} \tag{60}
$$

The lowest order anomalous dimension coefficient, being universal, is given for the three moments by Eq. (21) and in our basis the coefficients $d_{v_3;v_3^{m_1};0}$ and $d_{v_3^{m_1};v_3;0}$ vanish, while $d_{v_3^{m_1};0} = -28/(9(4\pi)^2)$.

For consistency at this order in perturbation theory, we take $c_{sw} = 1$ (the tree level value). Setting $\mu = 1/a$ to avoid large logarithmic factors gives, for example, at $\beta = 6.0$, the results

$$
\begin{aligned}
&Z^{\overline{\rm MS}}_{v_{2a};\rm LAT} = 1.0276, \qquad Z^{\overline{\rm MS}}_{v_{2b};\rm LAT} = 1.0207, \\
&Z^{\overline{\rm MS}}_{v_3;\rm LAT} = 1.1354, \qquad Z^{\overline{\rm MS}}_{v_3;v_3^{m_1};\rm LAT} = 0.003\,142, \\
&\qquad\qquad Z^{\overline{\rm MS}}_{v_4;\rm LAT} = 1.2453.
\end{aligned} \tag{61}
$$

While we see that for v_2 first order perturbation theory changes the tree level result ($\equiv 1$) very little, there are perceptible differences for the higher moments. Note also that the mixing renormalization constant for v_3 is very small in comparison to the diagonal renormalization constant, $Z^{\overline{\rm MS}}_{v_3;\rm LAT}$. In addition, although the mixing operator signal is rather noisy, $v_3^{m_1} \ll v_3$ as we have seen in Sec. IV D 2. Thus assuming that for a nonperturbative evaluation $Z^{\overline{\rm MS}}_{v_3;\rm LAT} \gg Z^{\overline{\rm MS}}_{v_3;v_3^{m_1};\rm LAT}$, as is the case for the perturbative result, we can ignore the effects of the mixing term in the future.

We also use the three-loop result from Table VI for $\Delta Z^{\overline{\rm MS}}_{v_n}(\mu)$ to find the RGI factor $Z^{\rm RGI}_{v_n}$, Eq. (15).

2. TRB perturbation theory

To improve the perturbative renormalization results of the last section, we shall apply TRB-PT [58] which we shall now describe. The renormalized operator is given by

$$
\mathcal{O}^{\overline{\rm MS}}(\mu) = Z^{\overline{\rm MS}}_{\mathcal{O};\rm LAT}(\mu, a)\mathcal{O}(a). \tag{62}
$$

As in Eq. (11), we may define a γ function either in the $\overline{\rm MS}$ scheme or what we shall formally call here the *LAT* scheme. Additionally as expansions in the bare coupling constant seem to be badly convergent, we choose to expand in the boosted coupling constant and thus we have

$$
\begin{aligned}
\frac{\partial}{\partial \log a}\log Z^{\overline{\rm MS}}_{\mathcal{O};\rm LAT}(\mu, a)\,\Big|_{\mu} &= \gamma^{\rm LAT}_{\mathcal{O}}(g_{\Box}) \\
&= d_{\mathcal{O};0}g^2_{\Box} + d^{\rm LAT}_{\mathcal{O};1}g^4_{\Box} + \ldots,
\end{aligned} \tag{63}
$$

where

$$
g^2_{\Box} = \frac{g_0^2}{u_0^4}, \qquad u_0^4 = \left\langle \frac{1}{3}{\rm Tr}U^{\Box} \right\rangle, \tag{64}
$$

and $U^{\Box}$ is the product of links around an elementary plaquette. Expanding [59] u_0 we have $u_0 = 1 - \frac{1}{4}g_0^2 p_1 + O(g_0^4)$ where $p_1 = 1/3$.

From Eq. (13), i.e. integrating Eq. (11) for $(S, M) \equiv (\overline{\rm MS}, \mu)$ and (LAT, a^{-1}), Eq. (63), gives

$$
Z^{\overline{\rm MS}}_{\mathcal{O};\rm LAT}(\mu, a) = \frac{\Delta Z^{\rm LAT}_{\mathcal{O}}(a)}{\Delta Z^{\overline{\rm MS}}_{\mathcal{O}}(\mu)}, \tag{65}
$$

and thus from Eq. (62), the RGI quantity may be written as

$$
\mathcal{O}^{\rm RGI} = \Delta Z^{\overline{\rm MS}}_{\mathcal{O}}(\mu)\mathcal{O}^{\overline{\rm MS}}(\mu) = \Delta Z^{\rm LAT}_{\mathcal{O}}(a)\mathcal{O}(a). \tag{66}
$$

Expanding Eq. (65) and comparing with Eq. (57) enables an expression to be found for $d^{\rm LAT}_{\mathcal{O};1}$ of

$$
d^{\rm LAT}_{\mathcal{O};1} = d^{\overline{\rm MS}}_{\mathcal{O};1} + d_{\mathcal{O};0}(t_1 - p_1) - 2b_0 B^{\overline{\rm MS}}_{\mathcal{O}}(1), \tag{67}
$$

for $O(a)$ improved Wilson fermions and where

$$
g^{\overline{\rm MS}} = g_0\left(1 + \frac{1}{2}t_1 g_0^2 + \ldots\right), \tag{68}
$$

at the scale $\mu = 1/a$. t_1 is known and is given by $t_1 = 0.468\,201$ [60]. Hence $d^{\rm LAT}_{\mathcal{O};1}$ may be computed. Values are given in Table III.

TABLE III. Values of $d^{\rm LAT}_{\mathcal{O};1}$ for $O(a)$ improved fermions from Eq. (67).

$\mathcal{O}$	$d^{\rm LAT}_{\mathcal{O};1}$
v_{2a}	$-152.14/(4\pi)^4$
v_{2b}	$-176.31/(4\pi)^4$
v_3	$91.828/(4\pi)^4$
v_4	$382.32/(4\pi)^4$

For two loops a simple exact analytic expression is possible for $\Delta Z^S_{\mathcal{O}}(M)$ of

$$\Delta Z^{\mathrm{LAT}}_{\mathcal{O}}(a) = [2b_0 g^2_{\Box}]^{d_{\mathcal{O};0}/2b_0}\left[1 + \frac{b_1}{b_0} g^2_{\Box}\right]^{((b_0 d^{\mathrm{LAT}}_{\mathcal{O};1} - b_1 d_{\mathcal{O};0})/2b_0 b_1)}. \tag{69}$$

The expression in Eq. (69) is the result of renormalization-group-improved boosted perturbation theory. We can finally tadpole improve it to obtain to this order

$$\Delta Z^{\mathrm{LAT}}_{\mathcal{O}}(a) = u_0^{1-n_D}[2b_0 g^2_{\Box}]^{d_{\mathcal{O};0}/2b_0}\left[1 + \frac{b_1}{b_0} g^2_{\Box}\right]^{((b_0 d^{\mathrm{LAT}}_{\mathcal{O};1} - b_1 d_{\mathcal{O};0})/2b_0 b_1) + (p_1/4)(b_0/b_1)(1-n_D)}, \tag{70}$$

where n_D is the number of derivatives in the operator. Note that for one derivative operators, TI has no effect, i.e. Eq. (69) is the same as Eq. (70). Again we have the factor, Eq. (15),

$$Z^{\mathrm{RGI}}_{\mathcal{O}}(a) \equiv \Delta Z^{\mathrm{LAT}}_{\mathcal{O}}(a). \tag{71}$$

Thus $\Delta Z^{\mathrm{LAT}}_{\mathcal{O}}(a)$ is the function that takes you directly from the lattice result to the RGI result.

Finally if we additionally wish to TI the improvement coefficients [44], then we replace g_0^2 by $g^2_{\Box}$ in Eq. (37). Numerical results from this procedure for c_0 in υ_2 with $c_2 = 0$ are given in Table IV. The associated κ_c, necessary for the computation of am_q, Eq. (36) are given in [61].

3. Nonperturbative Z determinations

We now look at the RI′-MOM nonperturbative determination of renormalization constants, using the method proposed by Martinelli *et al.* [51] which mimics, up to a point, the approach of the perturbative lattice procedure, by defining

$$Z^{\mathrm{RI'-MOM}}_{\mathcal{O};\mathrm{LAT}}(\mu_p, a) = \frac{Z^{\mathrm{RI'-MOM}}_{q;\mathrm{LAT}}(p)}{\frac{1}{12}\mathrm{Tr}[\Lambda_{\mathcal{O}}(p)\Lambda^{-1}_{\mathcal{O},\mathrm{BORN}}(p)]}\Bigg|_{p^2=\mu_p^2}, \tag{72}$$

where the wave function renormalization constant $Z^{\mathrm{RI'-MOM}}_{q;\mathrm{LAT}}(p)$ can be fixed from the conserved vector current or from the Fourier transformed quark propagator $S_q(p)$ by

$$Z^{\mathrm{RI'-MOM}}_{q;\mathrm{LAT}}(p) = \frac{\mathrm{Tr}(-i\sum_\lambda \gamma_\lambda \sin(ap_\lambda) a S_q^{-1}(p))}{12\sum_\lambda \sin^2(ap_\lambda)}. \tag{73}$$

There are still various possibilities for $Z_{q;\mathrm{LAT}}$, see e.g. [62] for different definitions. Again $\Lambda_{\mathcal{O}}$ is the amputated three-point Green's function for the operator $\mathcal{O}$. For our implementation using a "momentum source" see [63]. For the higher derivative operators considered here, this is a non-covariant renormalization condition, depending on the momentum direction [63,64]. Numerically this is a small effect for the momenta considered here.

The nonperturbative results for $Z^{\mathrm{RI'-MOM}}_{\mathcal{O};\mathrm{LAT}}$ should now be brought to an RGI form, which can only be done perturbatively. In order to avoid problems caused by the noncovariance of the renormalization condition Eq. (72) we first transform $Z^{\mathrm{RI'-MOM}}_{\mathcal{O};\mathrm{LAT}}$ perturbatively to a covariant scheme S like $\overline{\mathrm{MS}}$ or MOM employing a conversion factor of the form

$$1 + c_1^S(g^{\overline{\mathrm{MS}}})^2 + c_2^S(g^{\overline{\mathrm{MS}}})^4 + \ldots. \tag{74}$$

For $S = \overline{\mathrm{MS}}$ the general expression to one-loop order can be found in [63], while an explicit formula for υ_{2b} is given in [64]. The one-loop expressions for $S = \mathrm{MOM}$ are also known. In the case of υ_{2a} and υ_{2b}, three-loop expressions can be derived from [65] and will be used in the following.

In a second step, multiplying the resulting numbers by $\Delta Z^S_{\mathcal{O}}$ we obtain $Z^{\mathrm{RGI}}_{\mathcal{O}}$, see Eq. (15). Thus $\Delta Z^S_{\mathcal{O}}$ has to be found, which as in Sec. V B 2 is again the computation of the perturbative coefficients of the anomalous dimension

TABLE IV. Z^{RGI} results (and some c_0 results) at $\beta = 6.0$, 6.2 and 6.4, for υ_n, $n = 2$, 3 and 4. PT denotes the perturbative results from Sec. V B 1 and TRB-PT from Sec. V B 2. Note that to obtain Z^{RGI} for PT the three-loop results from Table VI have been used. Results from both NP variations, I and II are shown.

β		6.0	6.2	6.4
$Z^{\mathrm{RGI}}_{\upsilon_{2a}}$	PT	1.416	1.475	1.522
	TRB-PT	1.536	1.571	1.604
	NP-I	1.46	1.51	1.56
	NP-II	1.45	1.50	1.55
$c^{\mathrm{TI}}_{\upsilon_{2a};0}$	TI	1.232	1.218	1.205
$Z^{\mathrm{RGI}}_{\upsilon_{2b}}$	PT	1.407	1.465	1.513
	TRB-PT	1.519	1.555	1.589
	NP-I	1.46	1.50	1.55
	NP-II	1.45	1.49	1.54
$c^{\mathrm{TI}}_{\upsilon_{2b};0}$	TI	1.245	1.229	1.216
$Z^{\mathrm{RGI}}_{\upsilon_3}$	PT	1.928	2.038	2.129
	TRB-PT	2.268	2.330	2.389
	NP-I	2.2	2.3	2.4
	NP-II	2.1	2.2	2.3
$Z^{\mathrm{RGI}}_{\upsilon_4}$	PT	2.367	2.548	2.700
	TRB-PT	3.156	3.242	3.325
	NP-I	3.1	3.3	3.5
	NP-II	2.9	3.1	3.2

γ^S. For $S = \overline{\rm MS}$ the anomalous dimension is known to three loops for v_2, v_3 and v_4. For $S =$ MOM only the first two loops (or coefficients) are available in the case of v_3 and v_4, while in the case of v_2 we can make use of the three-loop calculation in [65]. For example, expanding in $g^{\overline{\rm MS}}$, $\gamma^S_{\mathcal{O}} = d_{\mathcal{O};0}(g^{\overline{\rm MS}})^2 + d^S_{\mathcal{O};1}(g^{\overline{\rm MS}})^4 + \ldots$, we have to two loops, similarly to Eq. (70),

$$\Delta Z^S_{\mathcal{O}} = [2b_0(g^{\overline{\rm MS}})^2]^{d_{\mathcal{O};0}/2b_0} \times \left[1 + \frac{b_1}{b_0}(g^{\overline{\rm MS}})^2\right]^{((b_0 d^S_{\mathcal{O};1} - b_1 d_{\mathcal{O};0})/2b_0 b_1)}. \quad (75)$$

Depending on the choice of S and the coupling in which γ^S is expanded, there are several possibilities, of course equivalent if one knew the whole power series. We shall briefly describe two methods here (I and II), whose difference we shall use to estimate the potential error due to unknown higher terms in the perturbative expansion. In method I we choose $S = \overline{\rm MS}$ and expand in $g^{\overline{\rm MS}}$. In method II we work in the MOM scheme (see [63]), and therefore it may be more natural to expand in other coupling constants defined using momentum renormalization conditions. In [66] several possibilities are given. We shall use here the coupling defined by using the three gluon vertex, the $\widetilde{\rm MOM}gg$ scheme, in the notation of [66].

If we plot $Z^{\rm RGI}_{v_n}$ against $(r_0\mu_p)^2$ and the perturbative expressions are sufficiently well known, we would expect to see a plateau where $Z^{\rm RGI}_{v_n}$ is independent of $(r_0\mu_p)^2$. This region occurs when μ_p is not too small $\sim\Lambda$ otherwise nonperturbative effects play a role, nor when it is too large as lattice artifacts then occur. Unfortunately these $O(r_0\mu_p)^2$ effects may become large, but depend very much on the operator considered. In perturbation theory in the chiral limit $O(a)$ terms of the Green's function have opposite chirality to the leading term, they disappear when the trace in Eq. (72) is taken. (For explicit lowest order results see [44].) Condensates may spoil this at low μ_p^2, but here we are looking at higher momentum scales. Thus we shall take the plateaus in the chiral limit as our renormalization constants.

We have made determinations of $Z^{\rm RI'-MOM}_{\mathcal{O};\rm LAT}$ using Eq. (72) for $\beta = 6.0$, 6.2 and 6.4 on $24^3 \times 48$, $24^3 \times 48$, $32^3 \times 40$ lattices, respectively. For each β three or more quark masses were used and a linear extrapolation in am_q was performed to the chiral limit. More details will be given in [67].

C. Comparison of $Z^{\rm RGI}$ results

We shall now use these results to find $Z^{\rm RGI}_{v_2}$, $Z^{\rm RGI}_{v_3}$ and $Z^{\rm RGI}_{v_4}$. In Figs. 10–12 we plot $Z^{\rm RGI}_{v_2}$, $Z^{\rm RGI}_{v_3}$ and $Z^{\rm RGI}_{v_4}$ for $\beta =$ 6.2 as computed from these techniques. In the case of the nonperturbative Zs a reasonable plateau is seen for large

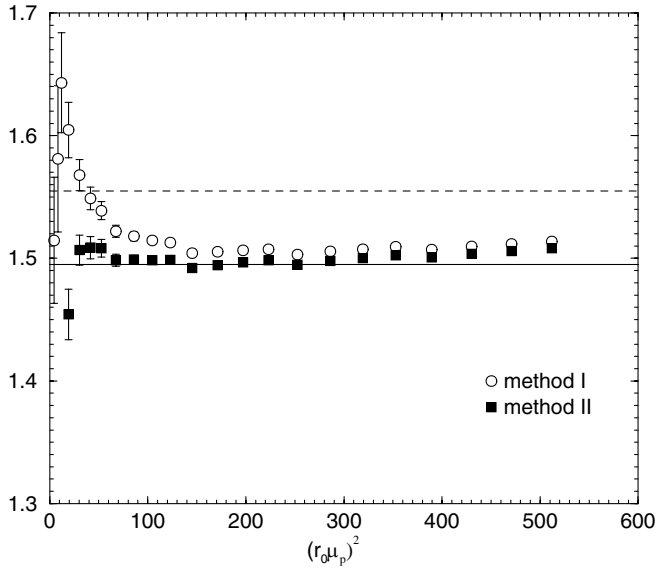

FIG. 10. $Z^{\rm RGI}_{v_{2b}}$ at $\beta = 6.2$ versus $(r_0\mu_p)^2$. NP method I is denoted by circles, while method II is given by solid squares. The dashed line is the TRB-PT result, while the solid line is the NP estimate from method II.

$(r_0\mu_p)^2$ enabling a value for the renormalization constant to be found. Method II seems to reach a plateau faster than method I, so we shall use these results (taken around $\mu_p =$ 5 GeV), with the difference to method I giving a rough estimate of the error. Also shown in the plots are the TRB-PT results. The NP and TRB-PT results lie close to each other, with a maximum discrepancy of about 8%. In the NP determination of the Zs, the plateaus become better as β increases. This is shown in Fig. 13.

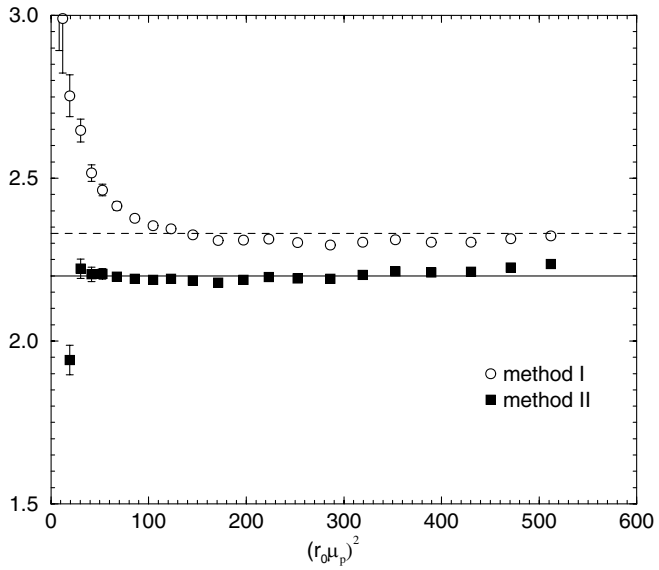

FIG. 11. $Z^{\rm RGI}_{v_3}$ at $\beta = 6.2$ versus $(r_0\mu_p)^2$. The same notation as for Fig. 10.

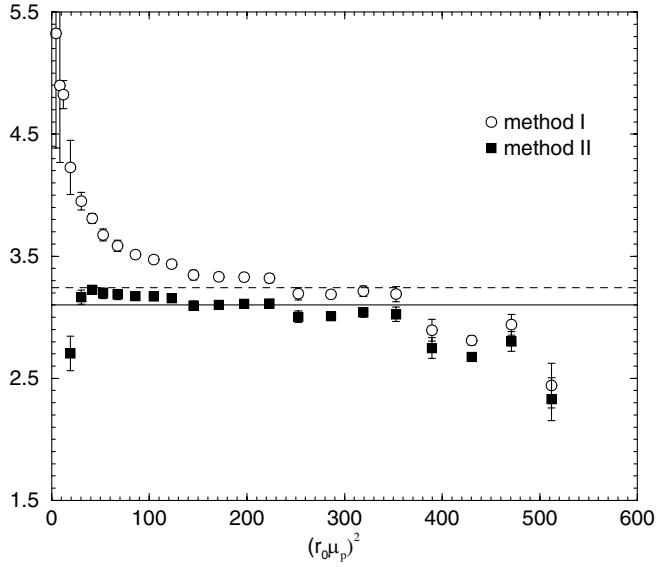

FIG. 12. $Z^{\mathrm{RGI}}_{v_4}$ at $\beta = 6.2$ versus $(r_0\mu_p)^2$. The same notation as for Fig. 10.

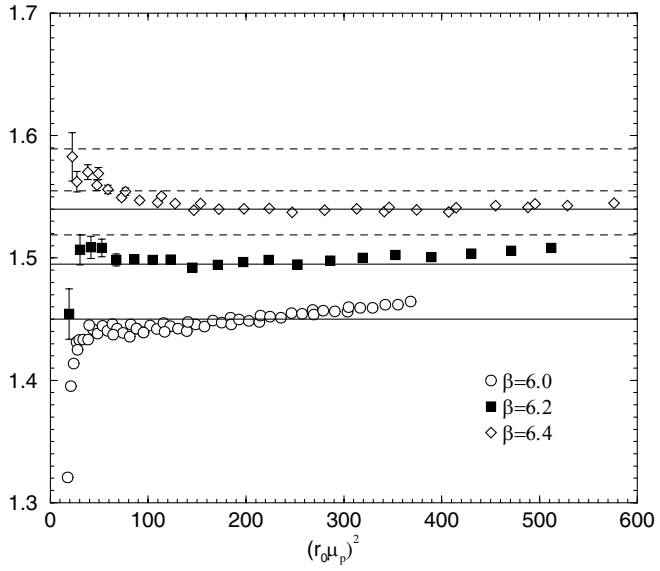

FIG. 13. $Z^{\mathrm{RGI}}_{v_{2b}}$ versus $(r_0\mu_p)^2$ for $\beta = 6.0$ (circles), 6.2 (solid squares), 6.4 (diamonds) using method II. The corresponding dashed lines are the TRB-PT results, while the solid lines are the NP estimates.

In Table IV we give the results from PT, TRB-PT and NP method (both variants) for various renormalization constants [68] at $\beta = 6.0$, 6.2 and 6.4.

VI. RESULTS: CHIRAL AND CONTINUUM EXTRAPOLATIONS

A. The phenomenological approach

From the bare results given in Appendix D and the Zs in Table IV we can now construct our renormalized matrix elements. As well as the continuum extrapolation $a \to 0$, as the quark masses presently used are rather heavy we also need to extrapolate these renormalized results to the chiral limit. In this section we shall consider both these limits. In Fig. 14 we show the results for v_{2b} for $\beta = 6.0$, 6.2 and 6.4 with both $c_0 = c_0^{\mathrm{TI}}$, as given in Table IV, and $c_0 = 0$. We would expect that for c_0 set to the TI value the results are practically $O(a)$ improved for nonzero quark mass.

Also shown in the figure is a linear extrapolation $v_n \equiv F^{(n)}_\chi(r_0 m_{ps})$ where

$$F^{(n)}_\chi(x) = a_n x^2 + b_n, \tag{76}$$

with $n = 2$. This fit describes the data well and not surprisingly using either c_0 value gives the same result in the chiral limit. Indeed the v_{2b} $O(a)$ improved results seem to be independent of the quark mass.

A similar situation holds for v_{2b}, v_{2a}, v_3 and v_4 all evaluated with nonzero momentum but with larger (and increasing) error bars. In Fig. 15 we show the results for v_{2a} for $\beta = 6.0$, 6.2 and 6.4 together with a linear chiral extrapolation (using the TI value for c_0). Immediately noticeable when comparing with Fig. 14 is the large increase in the error bars and less consistent ordering of gradients with increasing β.

In Fig. 16 we show the results for v_3 and in Fig. 17 the analogous results for v_4. As expected the results become noisier for increasing β. Perhaps surprisingly the results for v_4 seem to be as consistent over our β range as those of v_3. We have no real explanation for this.

The last limit to be taken is the continuum limit, $a \to 0$. As discussed in Sec. IV we believe that for v_2 the improvement terms are numerically small and so can be neglected. Thus we can make an extrapolation in a^2 (rather than a). While we cannot be so confident in this for the higher moments, based on the experience with the lowest moment, we shall also assume this for these higher moments. In Table V we give first the RGI values at $\beta = 6.0, 6.2, 6.4$, for v^{RGI}_{2a}, v^{RGI}_{2b} both for nonzero and zero momentum, v^{RGI}_3 and v^{RGI}_4. These have all been obtained using the NP method II results, as given in Table IV.

We now use these results to perform a continuum extrapolation. In Fig. 18 we plot the continuum extrapolations for the various v^{RGI}_2. A very consistent picture is obtained firstly between the different representations (''a'' and ''b'') and secondly between the different momenta used in the ''b'' representation. As expected though using a nonzero momentum gives a much noisier signal: in the extreme case between the off-diagonal and diagonal representations the error is about $\sim O(2)$ larger. We shall present our final result using v^{RGI}_{2b} for $\vec{p} = \vec{0}$ only. In Fig. 19 we show the results for v^{RGI}_3 and v^{RGI}_4. Using the modified operators of Eq. (31) enables a relatively smooth extrapolation to the continuum limit, giving results with about a 20%—30% error.

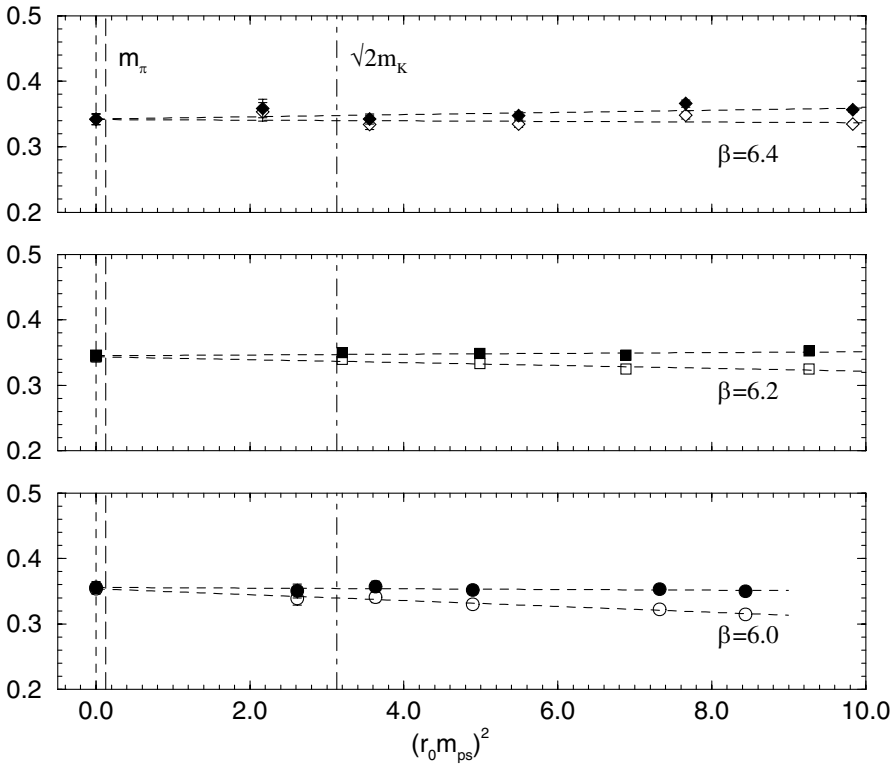

FIG. 14. v_{2b}^{RGI} versus $(r_0 m_{ps})^2$ for $\beta = 6.0$ (circles), 6.2 (squares) and 6.4 (diamonds). The solid symbols are obtained when using the TI value for c_0 while the open symbols have $c_0 = 0$. Also shown are linear extrapolations to the chiral limit (dashed lines). Other notation as in Fig. 6.

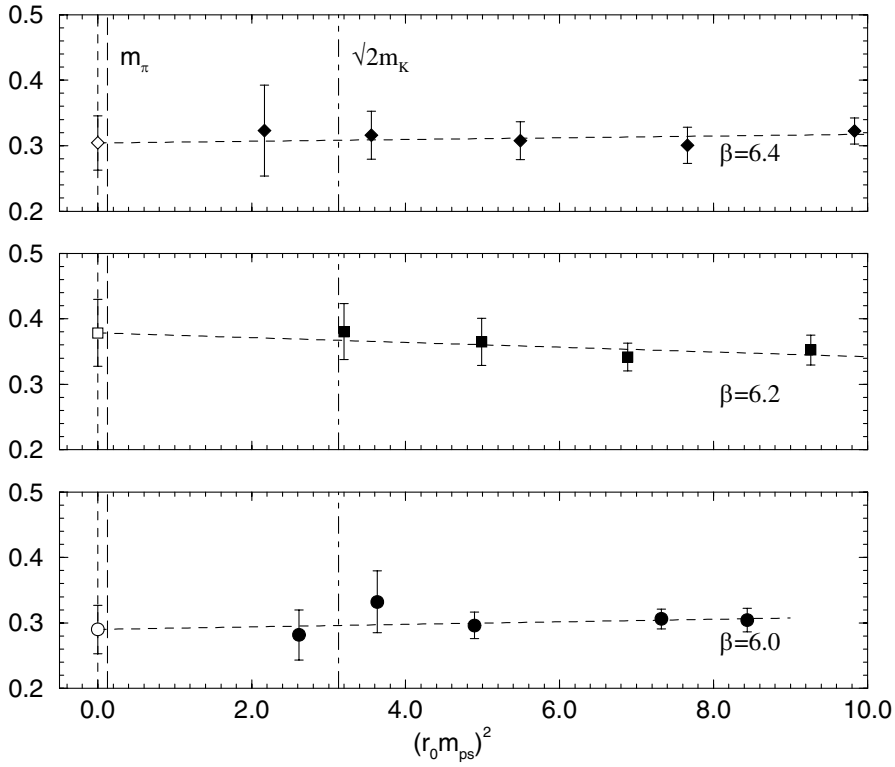

FIG. 15. v_{2a}^{RGI} versus $(r_0 m_{ps})^2$ for $\beta = 6.0$ (circles), 6.2 (squares) and 6.4 (diamonds). Also shown are linear extrapolations to the chiral limit (dashed lines). Other notation as in Fig. 6.

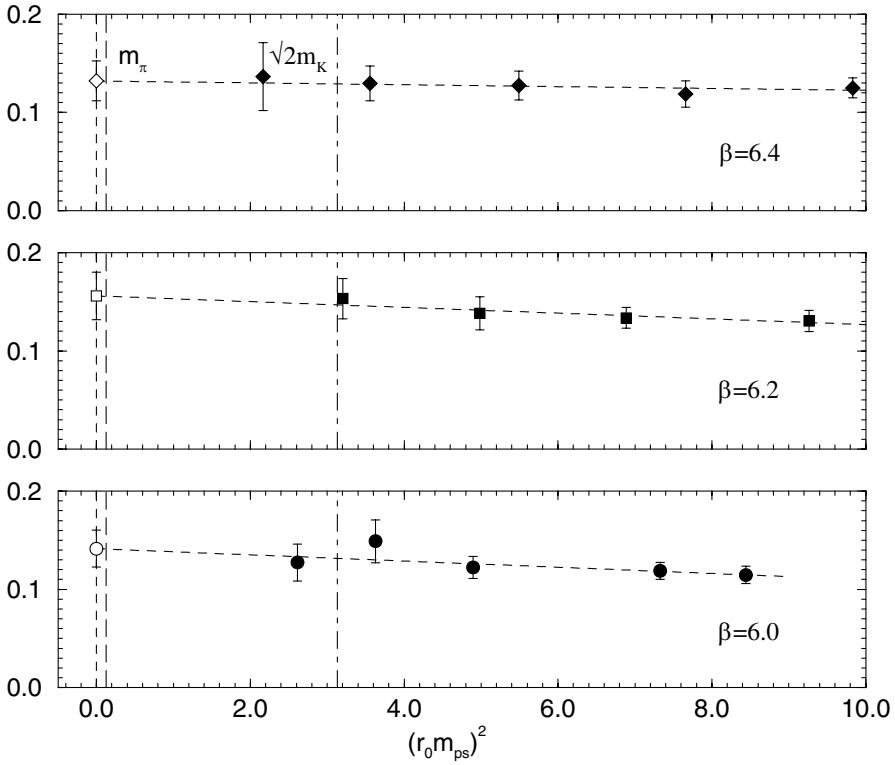

FIG. 16. v_3^{RGI} versus $(r_0 m_{ps})^2$ for $\beta = 6.0$ (circles), 6.2 (squares) and 6.4 (diamonds). Also shown are linear extrapolations to the chiral limit (dashed lines). Other notation as in Fig. 6.

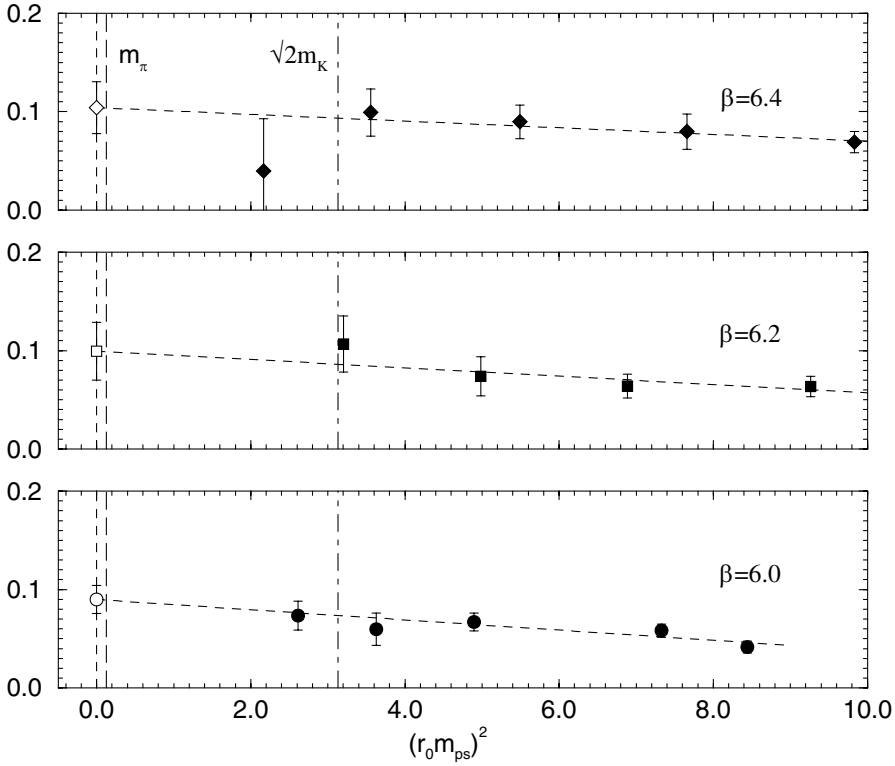

FIG. 17. v_4^{RGI} versus $(r_0 m_{ps})^2$ for $\beta = 6.0$ (circles), 6.2 (squares) and 6.4 (diamonds). Also shown are linear extrapolations to the chiral limit (dashed lines). Other notation as in Fig. 6.

TABLE V. v_n^{RGI} results for $O(a)$ improved fermions at $\beta = 6.0$, 6.2 and 6.4, for v_n^{RGI}, $n = 2$, 3 and 4. "∞" denotes the continuum extrapolation ($a = 0$) limit.

β	$\vec{p}$	6.0	6.2	6.4	∞
v_{2a}^{RGI}	$\vec{p}_1$	0.290(37)	0.379(51)	0.304(41)	0.343(56)
v_{2b}^{RGI}	$\vec{p}_1$	0.338(16)	0.328(16)	0.340(16)	0.335(22)
v_{2b}^{RGI}	$\vec{0}$	0.354(8)	0.344(8)	0.342(8)	0.335(11)
v_3^{RGI}	$\vec{p}_1$	0.141(19)	0.156(24)	0.132(20)	0.137(28)
v_4^{RGI}	$\vec{p}_1$	0.090(14)	0.099(29)	0.104(26)	0.110(33)

Finally, we convert our results to the $\overline{\mathrm{MS}}$ scheme at a scale of $\mu = 2\,\mathrm{GeV}$, using the three-loop result for $[\Delta Z_{v_n}^{\overline{\mathrm{MS}}}(\mu)]^{-1}$ given in Table VI. We find

$$\begin{aligned} v_2^{\overline{\mathrm{MS}}}(2\ \mathrm{GeV}) &= 0.245(9), \\ v_3^{\overline{\mathrm{MS}}}(2\ \mathrm{GeV}) &= 0.083(17), \\ v_4^{\overline{\mathrm{MS}}}(2\ \mathrm{GeV}) &= 0.059(18). \end{aligned} \tag{77}$$

The total error is the combined errors from the three-point functions and chiral/continuum fits together with the error for the renormalization constant. For v_{2b} this indicates that the dominant error is now possibly coming from the renormalization constant; the opposite is so for the higher moments.

How reasonable are the results in comparison with experiment or the MRST phenomenological fits? The MRST numbers from Sec. III in Table I are also plotted in Figs. 18 and 19. We see that for v_{2b}^{RGI}, the discrepancy between the experimental result and the lattice result stubbornly remains—and has persisted ever since the first pioneering works in the field [2]. It is also notable that v_4^{RGI} in particular is too large in comparison with the phenomenological result. For both v_3^{RGI} and v_4^{RGI} the chiral and continuum extrapolations are more problematic than for v_{2b}. This can probably only be resolved by increasing the statistic of the ensembles and by additional simulations at other β values.

Finally for completeness we give the results for $v_n^{(q)\overline{\mathrm{MS}}}$ for $q = u$, d separately. We find

$$\begin{aligned} v_2^{(u)\overline{\mathrm{MS}}}(2\ \mathrm{GeV}) &= 0.436(18), \\ v_3^{(u)\overline{\mathrm{MS}}}(2\ \mathrm{GeV}) &= 0.136(25), \\ v_4^{(u)\overline{\mathrm{MS}}}(2\ \mathrm{GeV}) &= 0.096(25), \end{aligned} \tag{78}$$

for the u quark and

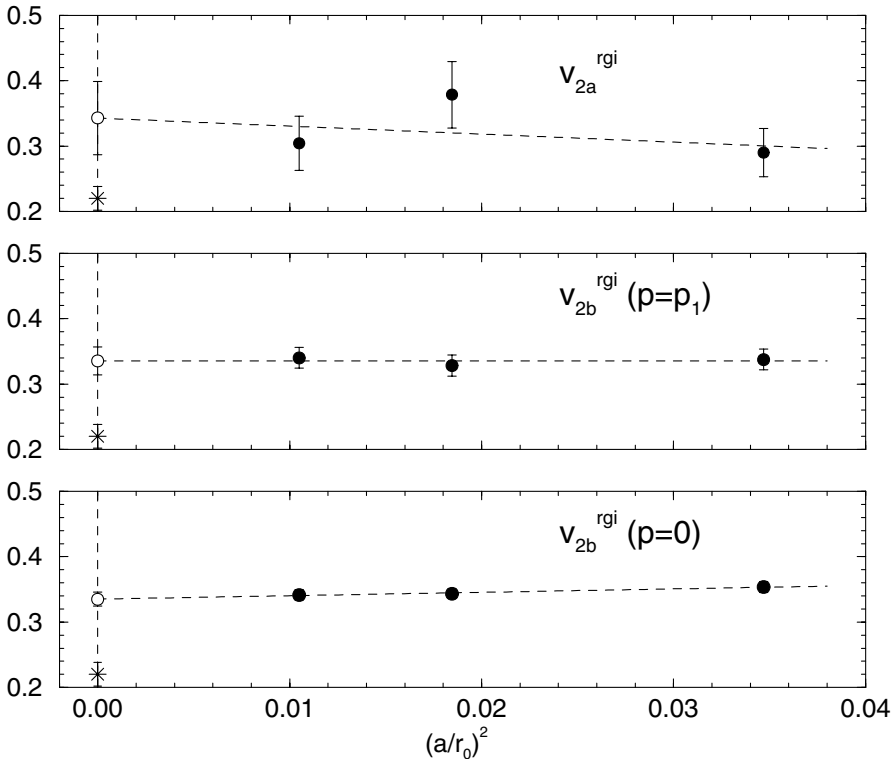

FIG. 18. v_{2a}^{RGI} and v_{2b}^{RGI} (both for $\vec{p} = \vec{p}_1$ and $\vec{0}$) versus $(a/r_0)^2$, using the results at $\beta = 6.0$, 6.2, 6.4 (solid circles). A linear continuum extrapolation in a^2 is also given (dashed line and open circle). The star is the MRST value given in Table I.

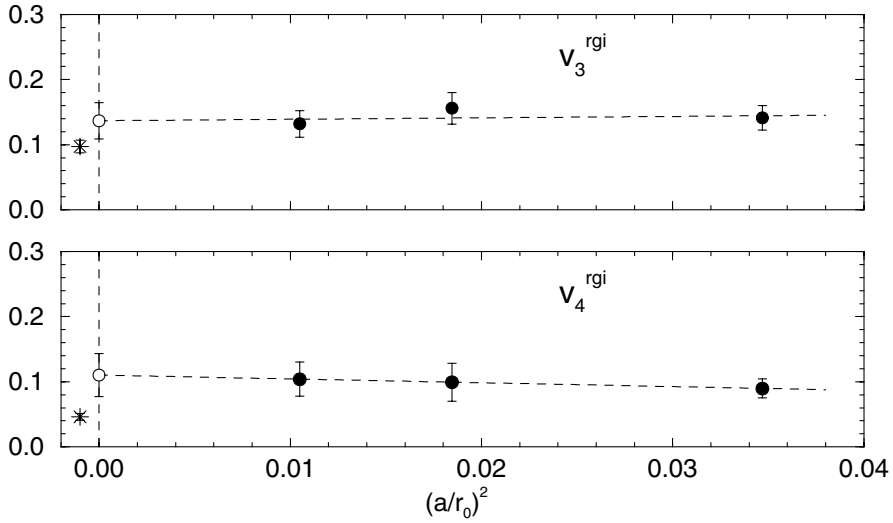

FIG. 19. v_3^{RGI} and v_4^{RGI} versus $(a/r_0)^2$, using the results at $\beta = 6.0, 6.2, 6.4$. A continuum extrapolation is also given. The same notation as for Fig. 18.

$$
\begin{aligned}
v_2^{(d)\overline{\mathrm{MS}}}(2\ \mathrm{GeV}) &= 0.191(7), \\
v_3^{(d)\overline{\mathrm{MS}}}(2\ \mathrm{GeV}) &= 0.052(11), \\
v_4^{(d)\overline{\mathrm{MS}}}(2\ \mathrm{GeV}) &= 0.027(15),
\end{aligned}
\tag{79}
$$

for the d quark. As discussed previously in Sec. IV C, the quark line disconnected, $qldis$, contribution to the matrix element is not computed. Thus on the rhs of Eqs. (78) and (79) there should be an extra term $v_n^{\overline{\mathrm{MS}}}|_{qldis}$, which is the same for u and d quarks, and of course, cancels for the NS results of Eq. (77).

B. Chiral perturbation theory

Although linear extrapolations in m_{ps}^2 seem to describe the results presented earlier quite well, it is not clear that for the quark masses used here and in other simulations higher order effects and/or chiral logarithms can be neglected. There has recently been a flurry of interest in this direction. In [9], based on chiral perturbation theory proposed in [8], a fit function model is used which tries to take into account the ''pion cloud'' around the nucleon, giving with $v_n^{\mathrm{RGI}} \equiv F_\chi^{(n)}(r_0 m_{ps})$,

$$
F_\chi^{(n)}(x) = a_n x^2 + b_n\left(1 - c x^2 \ln\frac{x^2}{(x^2 + (r_0\Lambda_\chi)^2)}\right), \tag{80}
$$

where Λ_χ is a parameter, the chiral scale, usually taken to be of $O(4\pi f_\pi) \sim O(1\ \mathrm{GeV})$ (where $f_\pi \approx 93$ MeV). For $\Lambda_\chi = 0$ or large pseudoscalar mass the equation reduces to the linear case, Eq. (76). Unfortunately as most of the masses used in the present simulation are larger than the strange quark mass, Eq. (80) may need higher order terms in chiral perturbation theory to be valid at these large masses. These first results have been confirmed by further chiral perturbation computations [10–13,70]. In particular in [13], quenched QCD was considered, with the result that at least for the nucleon there are no additional quenched chiral logarithms present, so-called ''hairpin diagrams,'' so the structure of the result in Eq. (80) remains unchanged. Furthermore in (unquenched) QCD we have

$$
c \equiv \frac{3g_A^2 + 1}{(4\pi r_0 f_\pi)^2} \sim 0.66, \tag{81}
$$

while for quenched QCD, assuming that $\alpha^{(2)} \sim \alpha^{(1)}$, $\beta^{(2)} \sim \beta^{(1)}$ in [13], then $c \sim 0.28$. Of course, in principle these formulas, Eq. (80), are valid in the continuum so we should first take the continuum limit and then apply chiral perturbation theory. Thus we should interpolate the values for v_n in Figs. 14–17 to a set of constant $(r_0 m_{ps})^2$ for each value of β. For each of these values of $(r_0 m_{ps})^2$ a continuum extrapolation should be performed. A chiral extrapolation of the data can then be attempted. Unfortunately our ''grid'' of data points is not fine enough and also for the higher moments is too noisy for this procedure. Thus we shall try a ''halfway house'' approach and attempt a simultaneous continuum and chiral extrapolation of the data,

$$
v_n^{\mathrm{RGI}} = F_\chi^{(n)}(r_0 m_{ps}) + c_n\left(\frac{a}{r_0}\right)^2 + d_n a r_0 m_{ps}^2, \tag{82}
$$

where the first term represents chiral physics, given by Eq. (76) or Eq. (80), the second term discretization effects and the last term accounts for residual $O(a)$ effects $\propto a m_q$, see Eq. (35). With this type of fit the number of free parameters is slightly reduced in comparison with the previous fit procedure given in Sec. VI A and so tends to produce smaller error bars. We shall restrict our results

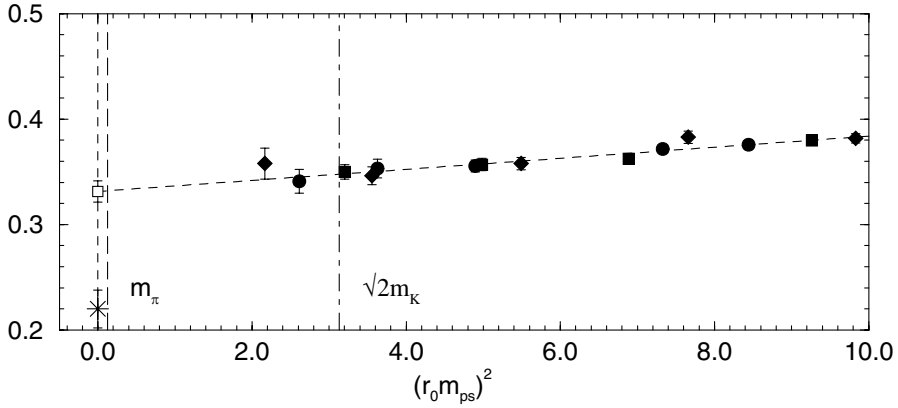

FIG. 20. $v_{2b}^{\rm RGI} - c_2(a/r_0)^2 - d_2 a r_0 m_{ps}^2 \equiv F_\chi^{(2)} = a_2(r_0 m_{ps})^2 + b_2$ [i.e. using the chiral function of Eq. (76)] versus $(r_0 m_{ps})^2$, dashed line. Filled circles, squares and diamonds represent $\beta = 6.0$, 6.2 and 6.4 respectively. The open square represents the chirally extrapolated value. Again the MRST phenomenological value of $v_2^{\rm RGI}$ is represented by a star.

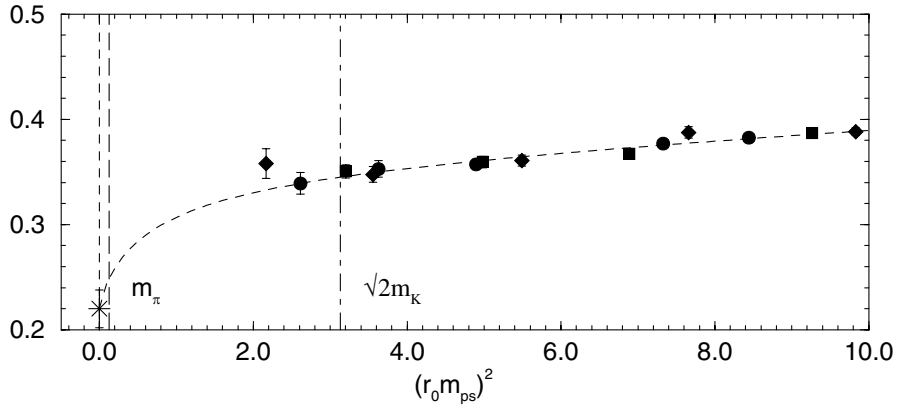

FIG. 21. $v_{2b}^{\rm RGI} - c_2(a/r_0)^2 - d_2 a r_0 m_{ps}^2 \equiv F_\chi^{(2)}$ [using the chiral function of Eq. (80)] versus $(r_0 m_{ps})^2$, dashed line. The chiral limit was fixed to the MRST phenomenological value. Other notation is the same as in Fig. 20.

here to our best data set—$v_{2b}^{\rm RGI}$ and first check that using Eq. (76) for $F_\chi^{(2)}$ reproduces our previous results. In Fig. 20 we first fit $v_{2b}^{\rm RGI}$ to Eq. (82) with $F_\chi^{(2)}$ given by the linear function of Eq. (76), $F_\chi^{(2)} = a_2(r_0 m_{ps})^2 + b_2$, and then plot $v_{2b}^{\rm RGI} - c_2(a/r_0)^2 - d_2 a r_0 m_{ps}^2$ for our three β values. There is no perceptible difference between using $c_0 = c_0^{\rm TI}$ or $c_0 = 0$, for definiteness we show the result for $c_0 = 0$. The points lie reasonably well on a straight line, with extrapolated result 0.331(10) consistent with our previously obtained value in Table V. The alternative possibility of $v_2^{\rm RGI} - (F_\chi^{(2)}(r_0 m_{ps}) - d_2 a r_0 m_{ps}^2)$ versus $(a/r_0)^2$ would display the $O(a^2)$ lattice discretization errors. However from Fig. 18 we see that the $O(a^2)$ effects are small and this alternative plot just reproduces them again.

Bolstered by this result, we now try to use $F_\chi^{(2)}$ from Eq. (80) as shown in Fig. 21. However it is difficult to detect any nonlinearities in the data and a 6-parameter fit $(a_2, b_2, c, \Lambda_\chi, c_2, d_2)$ fails. We were forced to see if such a 6-parameter fit could be plausible, by fixing the chiral limit, b_2, to be the MRST phenomenological value given in Table I. Note that there is no reason that the quenched QCD value must be the same as the QCD value, however for many hadronic quantities there appears to be little difference between the quenched and unquenched QCD values. As expected, while the fit (dashed) line and the numerical results are in reasonable agreement, all of the curvature of the fit takes place at small quark mass values where there is no data. Also the fitted parameter result for the chiral scale, $\Lambda_\chi \sim 505(48)$ MeV, is small in comparison with the expected value discussed earlier. For c we found 0.39(6) which lies between the unquenched and quenched values.

So it would seem that any possible nonlinearities can only show up at rather small quark mass outside the present range of data, where we would be able to match onto chiral perturbation theory. Teraflop simulations will be necessary to reach more physical pion masses. At present we shall stick to the simplest extrapolation possible.

VII. CONCLUSIONS

In this article we have computed v_n, $n = 2, 3, 4$. There is still a difference between the lattice results and the phenomenological results, particularly apparent for v_2 where there is a $\sim$40% discrepancy.

We have tried here to narrow down the range of possibilities for the disagreement, on the experimental side by comparing the global MRST/CTEQ fits with the experimental data: there is good agreement.

On the lattice side, we have $O(a)$ improved the lowest moment, investigated possible mixing operators for the higher moments and discussed renormalization, both perturbatively and nonperturbatively. At least for v_2 there do not seem to be large $O(a^2)$ corrections and so a continuum extrapolation can be reliably performed. For the higher moments, we have introduced modified operators for v_3, v_4, which reduces the numerical noise and improves the signal. It is difficult to see if there are any $O(a)$ corrections. However at present the simplest assumption that these corrections are small also fits with the data.

Although this is only a partial study here of mixing operators for the higher operators, due in part to the present difficulty of even defining renormalization constants for several of these operators, presently we find little sign of problems. Indeed even for $v_4^{m_3}$ the mixing with a lower-dimensional operator appears harmless. In other situations this is not the case and there may be significant changes, see e.g. [71].

We have discussed and compared various renormalization procedures ranging from simple perturbation theory to TRB perturbation theory to a nonperturbative method. Using NP results eliminates uncertainty concerning the renormalization constant. It is seen though that between TRB perturbation theory and nonperturbative results there can be up to an 8% difference—far less than the difference between the lattice and phenomenological result for v_2^{RGI}.

Finally although at present we see little numerical evidence of chiral logarithms, this is perhaps telling us that we must go to significantly smaller quark masses before the chiral extrapolation ''bends'' over. Thus we have an uncontrolled systematic error in the chiral extrapolation, which can only be removed if a match onto chiral perturbation theory can be performed. However preliminary results at lighter quark mass for unimproved Wilson fermions also show a linear behavior and so do not seem to improve the situation [5,72].

Whether quenching effects are significant remains unclear, but recent unquenched results [5,73], do not seem to reveal any significant differences between quenching and unquenching, at least in the quark mass range considered. Finally there are hints of a different situation for overlap fermions [74], which might suggest again that one has to simulate at light quark masses close to the chiral limit—a challenge for the lattice.

ACKNOWLEDGMENTS

The numerical calculations were performed on the APE1000 and Quadrics at DESY (Zeuthen) and the Cray T3E at ZIB (Berlin). We thank the operating staff for their support. This work has been supported in part by the EU Integrated Infrastructure Initiative Hadron Physics (I3HP) under Contract No. RII3-CT-2004-506078 and by the DFG (Forschergruppe Gitter-Hadronen-Phänomenologie). We also wish to thank L. Mankiewicz for much help with obtaining the experimental data used in Sec. III and S. Capitani for one-loop perturbative results for v_3 and v_4.

APPENDIX A: SOME VALUES OF $[\Delta Z_{v_n}^{\overline{\mathrm{MS}}}(\mu)]^{-1}$

Some useful values of $[\Delta Z_{v_n}^{\overline{\mathrm{MS}}}(\mu)]^{-1}$ are given in Table VI.

APPENDIX B: GROUP PROPERTIES OF MIXING OPERATORS—AN EXAMPLE

When QCD is put on the lattice, the necessary analytic continuation from Minkowski to Euclidean space replaces the Lorentz group by the orthogonal group $O(4)$, which by the discretization of space-time is further reduced to the hypercubic group $H(4) \subset O(4)$. Since $H(4)$ is only a finite group, the restrictions imposed by symmetry are less stringent than in the continuum and the possibilities for mixing

TABLE VI. Useful values of $[\Delta Z_{v_n}^{\overline{\mathrm{MS}}}(\mu)]^{-1}$ $(n = 2, 3, 4)$ and $\alpha_s^{\overline{\mathrm{MS}}}(\mu) \equiv (g^{\overline{\mathrm{MS}}}(\mu))^2/4\pi$. The errors are a reflection of the error in $\Lambda^{\overline{\mathrm{MS}}} r_0$. The lattice simulations performed here, as described in Sec. IV, give the above $1/a$ values (found from using r_0/a in [35], namely $r_0/a = 5.37, 7.36, 9.76$ at $\beta = 6.0, 6.2, 6.4$ respectively, together with the scale $r_0 = 0.5$ fm).

μ	One-loop	Two-loop	Three-loop
		$[\Delta Z_{v_2}^{\overline{\mathrm{MS}}}(\mu)]^{-1}$	
2.00 GeV	0.783(10)	0.721(8)	0.732(9)
2.12 GeV ($1/a$ at $\beta = 6.0$)	0.776(10)	0.715(8)	0.726(8)
2.90 GeV ($1/a$ at $\beta = 6.2$)	0.743(8)	0.688(7)	0.696(7)
3.85 GeV ($1/a$ at $\beta = 6.4$)	0.718(7)	0.668(6)	0.674(6)
		$[\Delta Z_{v_3}^{\overline{\mathrm{MS}}}(\mu)]^{-1}$	
2.00 GeV	0.682(14)	0.596(10)	0.609(11)
2.12 GeV ($1/a$ at $\beta = 6.0$)	0.673(13)	0.589(10)	0.602(10)
2.90 GeV ($1/a$ at $\beta = 6.2$)	0.629(11)	0.555(8)	0.564(8)
3.85 GeV ($1/a$ at $\beta = 6.4$)	0.596(9)	0.529(7)	0.537(7)
		$[\Delta Z_{v_4}^{\overline{\mathrm{MS}}}(\mu)]^{-1}$	
2.00 GeV	0.619(15)	0.520(11)	0.534(13)
2.12 GeV ($1/a$ at $\beta = 6.0$)	0.609(15)	0.512(10)	0.526(12)
2.90 GeV ($1/a$ at $\beta = 6.2$)	0.559(12)	0.476(9)	0.486(10)
3.85 GeV ($1/a$ at $\beta = 6.4$)	0.522(10)	0.448(8)	0.456(8)
		$\alpha_s^{\overline{\mathrm{MS}}}(\mu)$	
2.00 GeV	0.268(10)	0.195(6)	0.201(6)
2.12 GeV ($1/a$ at $\beta = 6.0$)	0.261(9)	0.191(5)	0.196(6)
2.90 GeV ($1/a$ at $\beta = 6.2$)	0.228(7)	0.170(5)	0.174(5)
3.85 GeV ($1/a$ at $\beta = 6.4$)	0.205(6)	0.156(3)	0.159(4)

increase, sometimes in a way which is not easily anticipated by our (continuum) intuition. In this appendix we illustrate the problem by an example. For a more complete and systematic treatment we refer to [42].

As a further symmetry we have charge conjugation. It operates on the fermion fields $q(x)$, $\overline{q}(x)$ and on the lattice gauge field $U_\mu(x)$ according to

$$\begin{aligned} q(x) &\xrightarrow{C} C\overline{q}(x)^T, \\ \overline{q}(x) &\xrightarrow{C} -q(x)^T C^{-1}, \\ U_\mu(x) &\xrightarrow{C} U_\mu(x)^* \end{aligned} \tag{B1}$$

with the charge conjugation matrix C satisfying

$$C\gamma_\mu^T C^{-1} = -\gamma_\mu. \tag{B2}$$

So we get, e.g.,

$$\begin{aligned} \mathcal{O}^{\gamma}_{\mu_1\mu_2\ldots\mu_n} &\xrightarrow{C} (-1)^n \mathcal{O}^{\gamma}_{\mu_1\mu_n\mu_{n-1}\ldots\mu_2}, \\ \mathcal{O}^{\gamma\gamma_5}_{\mu_1\mu_2\ldots\mu_n} &\xrightarrow{C} (-1)^{n-1} \mathcal{O}^{\gamma\gamma_5}_{\mu_1\mu_n\mu_{n-1}\ldots\mu_2}. \end{aligned} \tag{B3}$$

Identifying elements R of $H(4)$ with the corresponding 4×4 matrices in the defining representation we find that R acts on $\overline{q}\gamma_\mu q$ as follows:

$$\overline{q}\gamma_\mu q \xrightarrow{R} \sum_\nu R_{\nu\mu}\bar{q}\gamma_\nu q, \tag{B4}$$

i.e., the four operators $\overline{q}\gamma_\mu q$ form a basis for the defining representation of $H(4)$. More generally, we get for the action of R:

$$\mathcal{O}^{\gamma}_{\mu_1\mu_2\ldots\mu_n} \xrightarrow{R} \sum_{\nu_1,\nu_2,\ldots,\nu_n} R_{\nu_1\mu_1}R_{\nu_2\mu_2}\cdots R_{\nu_n\mu_n}\mathcal{O}^{\gamma}_{\nu_1\nu_2\ldots\nu_n} \tag{B5}$$

and

$$\mathcal{O}^{\gamma\gamma_5}_{\mu_1\mu_2\ldots\mu_n} \xrightarrow{R} \det R \sum_{\nu_1,\nu_2,\ldots,\nu_n} R_{\nu_1\mu_1}R_{\nu_2\mu_2}\cdots R_{\nu_n\mu_n}\mathcal{O}^{\gamma\gamma_5}_{\nu_1\nu_2\ldots\nu_n}. \tag{B6}$$

Thus the 4^n operators $\mathcal{O}^{\gamma}_{\mu_1\mu_2\ldots\mu_n}$ as well as the operators $\mathcal{O}^{\gamma\gamma_5}_{\mu_1\mu_2\ldots\mu_n}$ form a basis for a representation of $H(4)$, which for $n>1$ is reducible. It is helpful to consider these operators as forming an orthonormal basis of the representation space. (Orthonormal) bases of irreducible subspaces have been given in [42].

Operators transforming according to the same irreducible representation of $H(4)$ (and having the same C parity) will in general mix with each other so that one has to consider appropriate linear combinations. Writing down these linear combinations one has to choose the bases in the different (equivalent) representation spaces such that they transform *identically* under $H(4)$ and not just equivalently.

Consider two bases which are known to transform according to the same irreducible representation. In order to check whether they even transform identically it is sufficient to compare their transformation behavior under a set of generators. For $H(4)$ there is a set of three generators $\{\alpha, \beta, \gamma\}$ given by

$$\alpha = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \beta = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad \gamma = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{B7}$$

According to (B5), α interchanges 1 and 2, β performs a cyclic permutation of the values of the indices, and γ produces a factor of -1 for every index 1, e.g., for the operator $\mathcal{O}^{\gamma}_{3341}$ one finds

$$\mathcal{O}^{\gamma}_{3341} \xrightarrow{\alpha} \mathcal{O}^{\gamma}_{3342}, \qquad \mathcal{O}^{\gamma}_{3341} \xrightarrow{\beta} \mathcal{O}^{\gamma}_{4412}, \qquad \mathcal{O}^{\gamma}_{3341} \xrightarrow{\gamma} -\mathcal{O}^{\gamma}_{3341}. \tag{B8}$$

As an example for mixing operators, let us consider $\mathcal{O}_{v_4}$. It belongs to a doublet of operators transforming according to the $H(4)$ representation $\tau_1^{(2)}$ (in the notation of [75] and Table II) and has positive C parity. Indeed, the two operators

$$u_1 := \frac{\sqrt{6}}{2}(\mathcal{O}^{\gamma}_{\{1122\}} + \mathcal{O}^{\gamma}_{\{3344\}} - \mathcal{O}^{\gamma}_{\{1133\}} - \mathcal{O}^{\gamma}_{\{2244\}}), \tag{B9}$$

and

$$\begin{aligned} u_2 &:= \frac{1}{\sqrt{2}}(2\mathcal{O}^{\gamma}_{\{1144\}} + 2\mathcal{O}^{\gamma}_{\{2233\}} - \mathcal{O}^{\gamma}_{\{1122\}} - \mathcal{O}^{\gamma}_{\{3344\}} \\ &\quad - \mathcal{O}^{\gamma}_{\{1133\}} - \mathcal{O}^{\gamma}_{\{2244\}}) \\ &= \sqrt{2}\mathcal{O}_{v_4} \end{aligned} \tag{B10}$$

form an orthonormal basis for this representation. It is obvious that both operators remain unchanged under the action of γ. The generator α acts on them as follows:

$$u_1 \xrightarrow{\alpha} \frac{1}{2}u_1 - \frac{1}{2}\sqrt{3}u_2, \qquad u_2 \xrightarrow{\alpha} -\frac{1}{2}\sqrt{3}u_1 - \frac{1}{2}u_2, \tag{B11}$$

while β gives

$$u_1 \xrightarrow{\beta} \frac{1}{2}u_1 + \frac{1}{2}\sqrt{3}u_2, \qquad u_2 \xrightarrow{\beta} \frac{1}{2}\sqrt{3}u_1 - \frac{1}{2}u_2. \tag{B12}$$

It is straightforward to check that the operators

$$w_1 := \frac{1}{4\sqrt{3}}(-\mathcal{O}^{\gamma}_{1122} - \mathcal{O}^{\gamma}_{2112} - \mathcal{O}^{\gamma}_{1221} - \mathcal{O}^{\gamma}_{2211} + 2\mathcal{O}^{\gamma}_{1212} + 2\mathcal{O}^{\gamma}_{2121} - \mathcal{O}^{\gamma}_{3344} - \mathcal{O}^{\gamma}_{4334} - \mathcal{O}^{\gamma}_{3443} - \mathcal{O}^{\gamma}_{4433} + 2\mathcal{O}^{\gamma}_{3434} + 2\mathcal{O}^{\gamma}_{4343} + \mathcal{O}^{\gamma}_{1133} + \mathcal{O}^{\gamma}_{3113} + \mathcal{O}^{\gamma}_{1331} + \mathcal{O}^{\gamma}_{3311} - 2\mathcal{O}^{\gamma}_{1313} - 2\mathcal{O}^{\gamma}_{3131} + \mathcal{O}^{\gamma}_{2244} + \mathcal{O}^{\gamma}_{4224} + \mathcal{O}^{\gamma}_{2442} + \mathcal{O}^{\gamma}_{4422} - 2\mathcal{O}^{\gamma}_{2424} - 2\mathcal{O}^{\gamma}_{4242}), \tag{B13}$$

and

$$\begin{aligned} w_2 &:= \frac{1}{12}(+\mathcal{O}^{\gamma}_{1122} + \mathcal{O}^{\gamma}_{2112} + \mathcal{O}^{\gamma}_{1221} + \mathcal{O}^{\gamma}_{2211} - 2\mathcal{O}^{\gamma}_{1212} - 2\mathcal{O}^{\gamma}_{2121} + \mathcal{O}^{\gamma}_{3344} + \mathcal{O}^{\gamma}_{4334} + \mathcal{O}^{\gamma}_{3443} + \mathcal{O}^{\gamma}_{4433} - 2\mathcal{O}^{\gamma}_{3434} - 2\mathcal{O}^{\gamma}_{4343} \\ &\quad + \mathcal{O}^{\gamma}_{1133} + \mathcal{O}^{\gamma}_{3113} + \mathcal{O}^{\gamma}_{1331} + \mathcal{O}^{\gamma}_{3311} - 2\mathcal{O}^{\gamma}_{1313} - 2\mathcal{O}^{\gamma}_{3131} + \mathcal{O}^{\gamma}_{2244} + \mathcal{O}^{\gamma}_{4224} + \mathcal{O}^{\gamma}_{2442} + \mathcal{O}^{\gamma}_{4422} - 2\mathcal{O}^{\gamma}_{2424} - 2\mathcal{O}^{\gamma}_{4242} \\ &\quad - 2\mathcal{O}^{\gamma}_{1144} - 2\mathcal{O}^{\gamma}_{4114} - 2\mathcal{O}^{\gamma}_{1441} - 2\mathcal{O}^{\gamma}_{4411} + 4\mathcal{O}^{\gamma}_{1414} + 4\mathcal{O}^{\gamma}_{4141} - 2\mathcal{O}^{\gamma}_{2233} - 2\mathcal{O}^{\gamma}_{3223} - 2\mathcal{O}^{\gamma}_{2332} - 2\mathcal{O}^{\gamma}_{3322} + 4\mathcal{O}^{\gamma}_{2323} \\ &\quad + 4\mathcal{O}^{\gamma}_{3232}), \\ &= \frac{1}{6}\mathcal{O}^{m_1}_{v_4}, \end{aligned} \tag{B14}$$

transform identically to u_1, u_2:

$$w_1 \xrightarrow{\alpha} \frac{1}{2}w_1 - \frac{1}{2}\sqrt{3}w_2, \qquad w_2 \xrightarrow{\alpha} -\frac{1}{2}\sqrt{3}w_1 - \frac{1}{2}w_2, \tag{B15}$$

$$w_1 \xrightarrow{\beta} \frac{1}{2}w_1 + \frac{1}{2}\sqrt{3}w_2, \qquad w_2 \xrightarrow{\beta} \frac{1}{2}\sqrt{3}w_1 - \frac{1}{2}w_2, \tag{B16}$$

and they are invariant under the action of γ. Hence any linear combination $r \cdot u_i + s \cdot w_i$ $(i = 1, 2)$ has the same transformation properties under $H(4)$ as u_i. In particular, $\mathcal{O}_{v_4}$ may mix with $\mathcal{O}^{m_1}_{v_4}$, and the renormalization of $\mathcal{O}_{v_4}$ will in general involve also $\mathcal{O}^{m_1}_{v_4}$.

Also the operators

$$w_1' := \frac{1}{4}(+\mathcal{O}^{\gamma\gamma_5}_{1324} - \mathcal{O}^{\gamma\gamma_5}_{2314} - \mathcal{O}^{\gamma\gamma_5}_{1423} + \mathcal{O}^{\gamma\gamma_5}_{2413} + \mathcal{O}^{\gamma\gamma_5}_{3142} - \mathcal{O}^{\gamma\gamma_5}_{4132} - \mathcal{O}^{\gamma\gamma_5}_{3241} + \mathcal{O}^{\gamma\gamma_5}_{4231} + \mathcal{O}^{\gamma\gamma_5}_{1234} - \mathcal{O}^{\gamma\gamma_5}_{3214} - \mathcal{O}^{\gamma\gamma_5}_{1432} + \mathcal{O}^{\gamma\gamma_5}_{3412} + \mathcal{O}^{\gamma\gamma_5}_{2143} - \mathcal{O}^{\gamma\gamma_5}_{4123} - \mathcal{O}^{\gamma\gamma_5}_{2341} + \mathcal{O}^{\gamma\gamma_5}_{4321}), \tag{B17}$$

and

$$\begin{aligned} w_2' &:= \frac{1}{4\sqrt{3}}(+\mathcal{O}^{\gamma\gamma_5}_{1234} - \mathcal{O}^{\gamma\gamma_5}_{3214} - \mathcal{O}^{\gamma\gamma_5}_{1432} + \mathcal{O}^{\gamma\gamma_5}_{3412} + \mathcal{O}^{\gamma\gamma_5}_{2143} - \mathcal{O}^{\gamma\gamma_5}_{4123} - \mathcal{O}^{\gamma\gamma_5}_{2341} + \mathcal{O}^{\gamma\gamma_5}_{4321} - \mathcal{O}^{\gamma\gamma_5}_{1324} + \mathcal{O}^{\gamma\gamma_5}_{2314} + \mathcal{O}^{\gamma\gamma_5}_{1423} - \mathcal{O}^{\gamma\gamma_5}_{2413} \\ &\quad - \mathcal{O}^{\gamma\gamma_5}_{3142} + \mathcal{O}^{\gamma\gamma_5}_{4132} + \mathcal{O}^{\gamma\gamma_5}_{3241} - \mathcal{O}^{\gamma\gamma_5}_{4231} + 2\mathcal{O}^{\gamma\gamma_5}_{1243} - 2\mathcal{O}^{\gamma\gamma_5}_{4213} - 2\mathcal{O}^{\gamma\gamma_5}_{1342} + 2\mathcal{O}^{\gamma\gamma_5}_{4312} + 2\mathcal{O}^{\gamma\gamma_5}_{2134} - 2\mathcal{O}^{\gamma\gamma_5}_{3124} - 2\mathcal{O}^{\gamma\gamma_5}_{2431} \\ &\quad + 2\mathcal{O}^{\gamma\gamma_5}_{3421}), \\ &= \frac{1}{2\sqrt{3}}\mathcal{O}^{m_2}_{v_4}, \end{aligned} \tag{B18}$$

transform in exactly the same way under α, β, and γ, i.e. under $H(4)$, as follows from (B6). Therefore $\mathcal{O}^{m_2}_{v_4}$ may mix with $\mathcal{O}_{v_4}$, too.

APPENDIX C: TWO- AND THREE-POINT CORRELATION FUNCTIONS

1. General formulas

In this appendix we shall give explicit expressions for the correlation functions in terms of quark propagators. We start with the two-point correlation function. A suitable proton operator is

$$B_\alpha(t; \vec{p}) = \sum_{\vec{x}} e^{-i\vec{p}\cdot\vec{x}} \epsilon^{abc} u^a_\alpha(\vec{x}, t)[u^b(\vec{x}, t)^{T_D} C\gamma_5 d^c(\vec{x}, t)], \tag{C1}$$

($a\ldots$ denote color indices, $\alpha\ldots$ Dirac indices, T_D is the transpose in Dirac space and $\gamma_5 = \gamma_1\gamma_2\gamma_3\gamma_4$). C, the charge conjugation matrix, has the defining property given in Eq. (B2) and is antisymmetric. As we take our gamma matrices to be Hermitian then C may be taken as unitary. Thus $-C = C^{T_D}$ and $C^{-1} = C^\dagger$. One possible choice, used in our computer program, is $C = \gamma_4\gamma_2$ (so that $-C = C^{-1}$) together with the Dirac basis

$$\gamma_i = \begin{pmatrix} 0 & -i\sigma_i \\ i\sigma_i & 0 \end{pmatrix}, \qquad \gamma_4 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \tag{C2}$$

but the results given here should not depend on this particular choice. The last two-quark fields in Eq. (C1) form a di-quark structure, while the first quark carries the spin index. With the help of

$$\overline{B}_\alpha(t;\vec{p}) = \sum_{\vec{x}} e^{i\vec{p}\cdot\vec{x}} \epsilon^{abc} \overline{u}^a_\alpha(\vec{x},t)[\overline{d}^b(\vec{x},t)\gamma_5 C \overline{u}^c(\vec{x},t)^{T_D}], \tag{C3}$$

the two-point correlation function is formed in the usual way

$$C_\Gamma(t;\vec{p}) = \mathrm{Tr}_D \Gamma \langle B(t,\vec{p})\overline{B}(0,\vec{p})\rangle, \tag{C4}$$

where we have introduced an arbitrary Dirac matrix, Γ, which for unpolarized nucleons, the case considered here, should be taken as $\frac{1}{2}(1+\gamma_4)$. This equation may be rewritten using quark propagators propagating from a (point) source to a sink. As we are averaging over the gauge fields we may shift all sources to $(\vec{0},0)$. Some algebra then yields

$$\begin{aligned} C_\Gamma(t;\vec{p}) = -N_S^3 \sum_{\vec{x}} e^{-i\vec{p}\cdot\vec{x}} \epsilon^{abc}\epsilon^{a'b'c'} \langle \mathrm{Tr}_D[\Gamma G^{(u)aa'}(\vec{x},t;\vec{0},0)]\mathrm{Tr}_D[\tilde{G}^{(d)bb'}(\vec{x},t;\vec{0},0)G^{(u)cc'}(\vec{x},t;\vec{0},0)] \\ + \mathrm{Tr}_D[\Gamma G^{(u)aa'}(\vec{x},t;\vec{0},0)\tilde{G}^{(d)bb'}(\vec{x},t;\vec{0},0)G^{(u)cc'}(\vec{x},t;\vec{0},0)]\rangle_{\{U\}}, \end{aligned} \tag{C5}$$

where N_S^3 is the number of spatial lattice sites and we have defined a tilde in Dirac space by

$$\tilde{X} = (C\gamma_5 X \gamma_5 C)^{T_D}. \tag{C6}$$

The problem is thus reduced to finding the propagator or Green's function for quark q from a source $(\vec{0},0)$ to $(\vec{x},t)$. In general, the quark propagator from y to x is defined by

$$G^{(q)ab}_{\alpha\beta}(x;y) = \langle q^a_\alpha(x)\overline{q}^b_\beta(y)\rangle_q, \tag{C7}$$

and can be computed from

$$\sum_w M^{(q)}(x;w)G^{(q)}(w;y) = \delta_{xy}, \tag{C8}$$

where $M^{ab}_{\alpha\beta}$ is the Wilson (clover) matrix, given in Appendix D.

For the three-point correlation function,

$$C_\Gamma(t,\tau;\vec{p};\mathcal{O}_q) = \mathrm{Tr}_D\Gamma\langle B(t,\vec{p})\mathcal{O}_q(\tau)\overline{B}(0,\vec{p})\rangle, \tag{C9}$$

we shall only consider the quark line connected term, i.e. the left diagram of Fig. 4. First we rewrite the operator insertion generally as

$$\mathcal{O}_q(\tau) = \sum_{\vec{y}} \mathcal{O}_q(\vec{y},\tau) = \sum_{\vec{y},v,w} \overline{q}^a_\alpha(v) O^{ab}_{\alpha\beta}(v,w;\vec{y},\tau) q^b_\beta(w) \tag{C10}$$

i.e. sum over spatial planes, where $(\vec{y},\tau)$ is the "center of mass of the operator." For operators without derivatives and with exactly one derivative we have $\overline{q}\gamma q$ and $\frac{1}{2}\overline{q}\gamma(\vec{D}-\overleftarrow{D})q$ for $\mathcal{O}_q$ respectively [$\vec{D}$ and $\overleftarrow{D}$ are defined in Eq. (30)] while for two and three derivative operators, to minimize the extension on the lattice we "integrate by parts" and choose $-\overline{q}\gamma\overleftarrow{D}\vec{D}q$ and $\frac{1}{2}\overline{q}\gamma(\overleftarrow{D}\overleftarrow{D}\vec{D}-\overleftarrow{D}\vec{D}\vec{D})q$ respectively. This also allows the higher derivative operators to be built from the previously constructed lower derivative operators.

Some algebra yields the result for Eq. (C9) of

$$C_\Gamma(t,\tau;\vec{p};\mathcal{O}_q) = -N_S^3 \sum_{\vec{y},v,w} \langle \mathrm{Tr}_{\mathrm{CD}}[\Sigma^{(q)}_\Gamma(\vec{0},0;v;\vec{p},t)O(v,w;\vec{y},\tau)G^{(q)}(w;0,0)]\rangle_{\{U\}}, \tag{C11}$$

where $\Sigma^{(q)}_\Gamma(\vec{0},0;v;\vec{p},t)$ is given by

$$\Sigma^{(q)}_\Gamma(\vec{0},0;v;\vec{p},t) = \sum_{\vec{x}} S^{(q)}_\Gamma(\vec{x},t;\vec{0},0;\vec{p})G^{(q)}(\vec{x},t;v), \tag{C12}$$

in terms of

$$\begin{aligned} S^{(u)a'a}_\Gamma(\vec{x},t;\vec{0},0;\vec{p}) = e^{-i\vec{p}\cdot\vec{x}}\epsilon^{abc}\epsilon^{a'b'c'}[\tilde{G}^{(d)bb'}(\vec{x},t;\vec{0},0)G^{(u)cc'}(\vec{x},t;\vec{0},0)\Gamma + \mathrm{Tr}_D[\tilde{G}^{(d)bb'}(\vec{x},t;\vec{0},0)G^{(u)cc'}(\vec{x},t;\vec{0},0)]\Gamma \\ + \Gamma G^{(u)bb'}(\vec{x},t;\vec{0},0)\tilde{G}^{(d)cc'}(\vec{x},t;\vec{0},0) + \mathrm{Tr}_D[\Gamma G^{(u)bb'}(\vec{x},t;\vec{0},0)]\tilde{G}^{(d)cc'}(\vec{x},t;\vec{0},0)]], \end{aligned} \tag{C13}$$

when $q = u$ and a slightly simpler expression for $S^{(d)}_\Gamma$ namely

$$S_\Gamma^{(d)a'a}(\vec{x},t;\vec{0},0;\vec{p}) = e^{-i\vec{p}\cdot\vec{x}}\epsilon^{abc}\epsilon^{a'b'c'}[\tilde{G}^{(u)bb'}(\vec{x},t;\vec{0},0)\tilde{\Gamma}\tilde{G}^{(u)cc'}(\vec{x},t;\vec{0},0) + \mathrm{Tr}_D[\Gamma G^{(u)bb'}(\vec{x},t;\vec{0},0)]\tilde{G}^{(u)cc'}(\vec{x},t;\vec{0},0)]. \tag{C14}$$

Practically we must find $\Sigma_\Gamma^{(q)}$ from a second Green's function using these rather ugly looking $S_\Gamma^{(q)}$ expressions as sources. By considering

$$\sum_v \Sigma_\Gamma^{(q)}(\vec{0},0;v;\vec{p},t)M^{(q)}(v;v') = S_\Gamma^{(q)}(\vec{v}',t;\vec{0},0;\vec{p})\delta_{v_4',t}, \tag{C15}$$

we see that this is the wrong way around for the inversion in Eq. (C8) but taking color/spin components and using $M^{ab}_{\alpha\beta}(x;y) = (\gamma_5 M^*(y;x)^{ba}\gamma_5)_{\beta\alpha}$ the equation for Σ can be rewritten as

$$\sum_v M^{(q)}(v';v)\gamma_5\Sigma_\Gamma^{(q)\dagger_{CD}}(\vec{0},0;v;\vec{p},t) = \gamma_5 S_\Gamma^{(q)\dagger_{CD}}(\vec{v}',t;\vec{0},0;\vec{p})\delta_{v_4',t}, \tag{C16}$$

where $\dagger_{CD}$ is the Hermitian conjugate in color and spin space. We see that in this form Σ is rather like a Green's function from the source given on the rhs of the equation.

Thus finding the three-point correlation functions is a two step process: first the usual Green's function from $(\vec{0},0)$ to any point x is found, stage I. Then a second inversion (stage II) is made with the source given from Eq. (C13) if the inserted operator consists of u quarks or using Eq. (C14) for d quarks.

The advantage of this procedure is that by tying together the two Green's functions appropriately any operator can be inserted with no additional computational cost. There is no restriction on the derivative structure and Dirac matrix γ. The disadvantage is that for each nucleon state (i.e. Γ determining whether the nucleon is unpolarized or polarized), momentum $\vec{p}$ and nucleon sink position (t) a separate inversion is required. Thus results for a range of t values are expensive and practically we must restrict ourselves to one value.

2. The nonrelativistic projection

To improve the overlap with the nucleon we have used Jacobi smearing [39], and nonrelativistic, NR, projection of the nucleon wave function. For completeness we now briefly describe our implementation of this projection. Rather than using the nucleon operator of Eq. (C1) we shall replace it by

$$B_\alpha^{\mathrm{NR}}(t;\vec{p}) = \sum_{\vec{x}} e^{-i\vec{p}\cdot\vec{x}}\epsilon^{abc}u_\alpha^a(\vec{x},t)[u_\beta^b(\vec{x},t) \times (C\gamma_5\frac{1}{2}(1+\gamma_4))_{\beta\gamma}d_\gamma^c(\vec{x},t)], \tag{C17}$$

i.e. we replace the matrix $C\gamma_5 \to C\gamma_5\frac{1}{2}(1+\gamma_4)$. Both operators, Eqs. (C1) and (C17), behave the same way under rotations and reflections in the spatial directions, and both have the same quantum numbers (color neutral, spin $\frac{1}{2}$, isospin $\frac{1}{2}$ and parity $+$), and both will therefore overlap with the proton.

As we shall see, not only is B^{NR} computationally cheaper by a factor of 2, but it has a better overlap with the proton. This can be easiest shown if we use the Dirac basis Eq. (C2) first giving

$$C\gamma_5 = \gamma_4\gamma_2\gamma_5 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad C\gamma_5\frac{1}{2}(1+\gamma_4) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{C18}$$

If we now write out Eq. (C17) for a spin-up (i.e. $\alpha = 1$) proton we have

$$B_1^{\mathrm{NR}} = \epsilon^{abc}(u_1^a u_1^b d_2^c - u_1^a u_2^b d_1^c), \tag{C19}$$

setting $\vec{p} = \vec{0}$ and suppressing the coordinate index for simplicity. This is not quite the final form. When we sum over all fermion line diagrams, only the part of the operator which is antisymmetric under the interchange of the two u quarks makes any contribution to the measured Green's function. Since the color wave function ϵ^{abc} is completely antisymmetric, this means that the part of Eq. (C19) which survives is

$$B_1^{\mathrm{NR}} = \epsilon^{abc}\left(u_1^a u_1^b d_2^c - \frac{1}{2}u_1^a u_2^b d_1^c - \frac{1}{2}u_2^a u_1^b d_1^c\right). \tag{C20}$$

Since the mid-1960s it has been known that the lowest-lying octet and decuplet baryons are well described as a **56** of $SU(6)$. According to $SU(6)$ the flavor/spin wave function of the spin-up proton is, e.g. [76]

$$p_\uparrow = \sqrt{\frac{1}{6}}(2u_\uparrow u_\uparrow d_\downarrow - u_\uparrow u_\downarrow d_\uparrow - u_\downarrow u_\uparrow d_\uparrow). \tag{C21}$$

Comparing this with Eq. (C20) we see that this is exactly the wave function we have been using. The wave function of Eq. (C21) makes some very successful predictions, for example, that the ratio of proton to neutron magnetic moments should be $-\frac{3}{2}$. The experimental value is -1.460, in good agreement with the $SU(6)$ prediction. This success suggests that Eq. (C20) is close to the true proton wave function. We would expect this wave function to work even better with heavier quarks, so it is an appropriate wave function to use on the lattice. If we carry out the same

exercise with Eq. (C21) we obtain

$$B_1 = \epsilon^{abc}\Big(u_1^a u_1^b d_2^c - \frac{1}{2}u_1^a u_2^b d_1^c - \frac{1}{2}u_2^a u_1^b d_1^c + \frac{1}{2}u_1^a u_3^b d_4^c + \frac{1}{2}u_3^a u_1^b d_4^c - \frac{1}{2}u_1^a u_4^b d_3^c - \frac{1}{2}u_4^a u_1^b d_3^c\Big), \quad \text{(C22)}$$

as the explicit component expression. This has many terms involving the 3rd and 4th Dirac components with amplitudes just as large as the terms with upper components only. In any sort of constituent quark model we would expect these terms to be small in the ground state. Adding them into the wave function not only increases the cost of the computation, it also degrades the signal by adding terms which are likely to have more overlap with excited baryon states.

The correlation functions are constructed also using

$$\overline{B}_\alpha^{\mathrm{NR}}(t;\vec{p}) = \sum_{\vec{x}} e^{i\vec{p}.\vec{x}}\epsilon^{abc}\overline{u}_\alpha^a(\vec{x},t) \times \Big[\overline{d}^b(\vec{x},t)\frac{1}{2}(1+\gamma_4)\gamma_5 C\overline{u}^c(\vec{x},t)^{T_D}\Big], \quad \text{(C23)}$$

i.e. replacing $\gamma_5 C \to \frac{1}{2}(1+\gamma_4)\gamma_5 C$. Thus the NR projection can be obtained by projecting out the positive eigenvalues of γ_4, i.e. by replacing each quark field, q, by

$$q \to \frac{1}{2}(1+\gamma_4)q, \qquad \overline{q} \to \overline{q}\frac{1}{2}(1+\gamma_4), \quad \text{(C24)}$$

everywhere and considering polarization matrices Γ which satisfy

$$\frac{1}{2}(1+\gamma_4)\Gamma = \Gamma = \Gamma\frac{1}{2}(1+\gamma_4). \quad \text{(C25)}$$

This gives the NR nucleon two-point function, Eq. (C4). So for the Dirac basis, Eq. (C2), only the components $\alpha = 1, 2$ are nonzero, as discussed above. In Fig. 22 we show a comparison for the nucleon two-point correlation function using the relativistic and nonrelativistic operators when applied to Eq. (C4). The gradient of the left branch measures the nucleon mass. It can be seen that this branch has been extended by about 5 units of t/a when using the NR operator as opposed to the relativistic operator.

For the three-point functions the new tilde, replacing Eq. (C6), is given by

$$\tilde{X} = \Big(C\gamma_5\frac{1}{2}(1+\gamma_4)X\frac{1}{2}(1+\gamma_4)\gamma_5 C\Big)^{T_D}, \quad \text{(C26)}$$

and obeys

$$\frac{1}{2}(1+\gamma_4)\tilde{X} = \tilde{X} = \tilde{X}\frac{1}{2}(1+\gamma_4). \quad \text{(C27)}$$

Considering Eqs. (C13) and (C14) then Eqs. (C25) and (C27) imply that

$$\frac{1}{2}(1+\gamma_4)S^{(q)} = S^{(q)} = S^{(q)}\frac{1}{2}(1+\gamma_4). \quad \text{(C28)}$$

This identity reduces the number of independent components from 4 to 2 for the source for the stage II inversion, Eq. (C16). Again, for the Dirac representation Eq. (C2) only the first two components are needed.

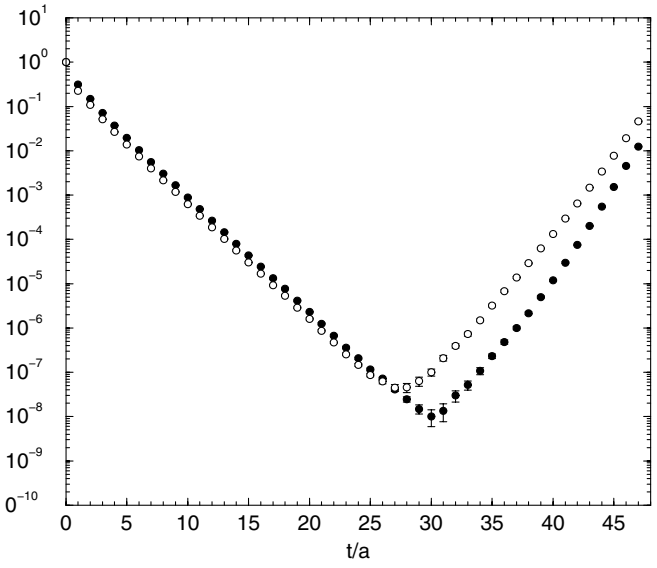

FIG. 22. Relativistic (open circles) and nonrelativistic (solid circles) two-point operators, Eq. (C4) with $\vec{p} = \vec{0}$, for $\beta = 6.2$ at $\kappa = 0.1344$.

TABLE VII. Parameter values used in the simulations, together with the measured pseudoscalar and nucleon masses. Note that the statistic refers to the number of (independent) configurations used for the 3-point functions; the masses have sometimes been determined with a larger statistic.

β	c_{sw}	κ	Volume	# configs.	am_{ps}	am_N
6.0	1.769	0.1320	$16^3 \times 32$	$O(445)$	0.5412(9)	0.9735(40)
6.0	1.769	0.1324	$16^3 \times 32$	$O(560)$	0.5042(7)	0.9353(25)
6.0	1.769	0.1333	$16^3 \times 32$	$O(560)$	0.4122(9)	0.8241(34)
6.0	1.769	0.1338	$16^3 \times 32$	$O(520)$	0.3549(12)	0.7400(85)
6.0	1.769	0.1342	$16^3 \times 32$	$O(735)$	0.3012(10)	0.7096(48)
6.2	1.614	0.1333	$24^3 \times 48$	$O(300)$	0.4136(6)	0.7374(21)
6.2	1.614	0.1339	$24^3 \times 48$	$O(300)$	0.3565(7)	0.6655(28)
6.2	1.614	0.1344	$24^3 \times 48$	$O(300)$	0.3034(6)	0.5963(29)
6.2	1.614	0.1349	$24^3 \times 48$	$O(470)$	0.2431(6)	0.5241(39)
6.4	1.526	0.1338	$32^3 \times 48$	$O(220)$	0.3213(8)	0.5718(28)
6.4	1.526	0.1342	$32^3 \times 48$	$O(120)$	0.2836(9)	0.5266(31)
6.4	1.526	0.1346	$32^3 \times 48$	$O(220)$	0.2402(8)	0.4680(37)
6.4	1.526	0.1350	$32^3 \times 48$	$O(320)$	0.1923(9)	0.4156(34)
6.4	1.526	0.1353	$32^3 \times 64$	$O(260)$	0.1507(8)	0.3580(47)

TABLE VIII. The bare results for v_2 from Eq. (39) for $\beta = 6.0$, $c_{sw} = 1.769$.

κ	0.1320	0.1324	0.1333	0.1338	0.1342
			$\vec{p} = \vec{p}_1$		
$v_{2a}^{(u)}$	0.359(19)	0.348(16)	0.335(21)	0.382(45)	0.312(35)
$v_{2a}^{(d)}$	0.1696(89)	0.1557(82)	0.144(11)	0.163(19)	0.123(17)
v_{2a}	0.189(11)	0.1925(95)	0.192(13)	0.219(31)	0.188(25)
$av_{2a}^{(u;1)}$	−0.0254(17)	−0.0261(15)	−0.0288(40)	−0.0288(40)	−0.0226(36)
$av_{2a}^{(d;1)}$	−0.013 21(93)	−0.013 72(85)	−0.0136(12)	−0.0115(21)	−0.0130(25)
$av_{2a}^{(1)}$	−0.012 11(98)	−0.012 38(91)	−0.0117(14)	−0.0170(29)	−0.009 74(292)
$av_{2a}^{(u;2)}$	0.0190(14)	0.0162(13)	0.006 76(165)	0.001 88(206)	−0.003 61(302)
$av_{2a}^{(d;2)}$	0.010 25(82)	0.009 33(75)	0.005 92(104)	−0.000 17(142)	0.001 36(227)
$av_{2a}^{(2)}$	0.008 72(79)	0.006 93(82)	0.001 11(117)	0.002 04(200)	−0.004 48(260)
			$\vec{p} = \vec{0}$		
$v_{2b}^{(u)}$	0.4066(32)	0.4108(27)	0.4130(48)	0.4196(78)	0.414(11)
$v_{2b}^{(d)}$	0.1894(17)	0.1886(16)	0.1851(26)	0.1845(41)	0.1798(55)
v_{2b}	0.2171(18)	0.2221(16)	0.2278(29)	0.2350(52)	0.2336(69)
$av_{2b}^{(u;1)}$	−0.028 44(39)	−0.028 95(40)	−0.030 02(81)	−0.0328(13)	−0.0311(25)
$av_{2b}^{(d;1)}$	−0.014 90(26)	−0.015 03(27)	−0.015 85(51)	−0.0185(11)	−0.0189(16)
$av_{2b}^{(1)}$	−0.013 54(21)	−0.013 90(25)	−0.014 06(54)	−0.0141(11)	−0.0117(16)
$av_{2b}^{(u;2)}$	0.039 54(52)	0.035 52(52)	0.0266(11)	0.0235(19)	0.0173(43)
$av_{2b}^{(d;2)}$	0.020 15(31)	0.018 22(35)	0.014 94(72)	0.0173(43)	0.0152(30)
$av_{2b}^{(2)}$	0.019 36(29)	0.017 22(32)	0.011 35(69)	0.007 41(162)	0.001 08(219)
			$\vec{p} = \vec{p}_1$		
$v_{2b}^{(u)}$	0.3918(54)	0.3959(49)	0.3908(88)	0.417(24)	0.370(19)
$v_{2b}^{(d)}$	0.1842(30)	0.1807(28)	0.1727(49)	0.186(13)	0.155(11)
v_{2b}	0.2076(31)	0.2152(32)	0.2180(59)	0.231(13)	0.215(14)
$av_{2b}^{(u;1)}$	−0.026 42(60)	−0.014 72(41)	−0.0288(12)	−0.0266(32)	−0.0268(41)
$av_{2b}^{(d;1)}$	−0.014 20(37)	−0.014 88(42)	−0.015 63(85)	−0.0171(20)	−0.0185(47)
$av_{2b}^{(1)}$	−0.012 22(42)	−0.013 20(39)	−0.012 97(90)	−0.0102(24)	−0.008 98(402)
$av_{2b}^{(u;2)}$	0.037 14(80)	0.033 57(78)	0.0244(17)	0.0168(45)	0.0114(63)
$av_{2b}^{(d;2)}$	0.019 44(48)	0.019 44(48)	0.014 33(11)	0.0149(26)	0.0168(58)
$av_{2b}^{(2)}$	0.017 77(52)	0.016 04(50)	0.009 88(109)	0.003 53(335)	−0.003 15(519)

APPENDIX D: TABLES

The action used here, in the quenched limit, is

$$S = \frac{1}{3}\beta\sum_{\square}\mathrm{Tr}_C\mathrm{Re}[1 - U_{\mu\nu}^{\square}] + a^4\sum_{xy;q=u,d}\overline{q}(x)M^{(q)}(x;y)q(y), \tag{D1}$$

where $U_{\mu\nu}^{\square}$ is the product of links around an elementary plaquette in the $\mu - \nu$ plane and the Wilson (clover) fermion matrix is given by

$$\sum_{xy}\overline{q}(x)M^{(q)}(x;y)q(y) = \sum_x\left\{\frac{1}{a}\overline{q}(x)q(x) - \frac{\kappa}{a}\sum_\mu\overline{q}(x)U_\mu^\dagger(x - a\hat{\mu})[1 + \gamma_\mu]q(x - a\hat{\mu})\right.$$
$$\left. - \frac{\kappa}{a}\sum_\mu\overline{q}(x)U_\mu(x)[1 - \gamma_\mu]q(x + a\hat{\mu}) - 2\kappa a\, c_{sw}\, g_0\sum_{\mu\nu}\frac{1}{4}\overline{q}(x)\sigma_{\mu\nu}F_{\mu\nu}^{\mathrm{clover}}(x)q(x)\right\}, \tag{D2}$$

TABLE IX. The bare results for v_2 from Eq. (39) for $\beta = 6.2$, $c_{sw} = 1.614$.

κ	0.1333	0.1339	0.1344	0.1349
		$\vec{p} = \vec{p}_1$		
$v_{2a}^{(u)}$	0.403(24)	0.393(23)	0.413(38)	0.422(42)
$v_{2a}^{(d)}$	0.187(11)	0.179(11)	0.182(17)	0.175(18)
v_{2a}	0.217(14)	0.214(13)	0.232(23)	0.246(28)
$av_{2a}^{(u;1)}$	−0.0252(18)	−0.0247(18)	−0.0251(28)	−0.0269(32)
$av_{2a}^{(d;1)}$	−0.012 94(96)	−0.012 24(97)	−0.0130(16)	−0.0123(18)
$av_{2a}^{(1)}$	−0.0122(10)	−0.0125(11)	−0.0122(19)	−0.0144(22)
$av_{2a}^{(u;2)}$	0.0117(12)	0.004 53(116)	−0.001 10(179)	−0.004 93(254)
$av_{2a}^{(d;2)}$	0.006 22(71)	0.003 08(68)	−0.0067(110)	−0.002 24(144)
$av_{2a}^{(2)}$	0.005 50(71)	0.001 43(81)	−0.001 40(146)	−0.002 73(209)
		$\vec{p} = \vec{0}$		
$v_{2b}^{(u)}$	0.4100(28)	0.4020(39)	0.4071(52)	0.4052(58)
$v_{2b}^{(d)}$	0.1920(14)	0.1839(21)	0.1834(28)	0.1775(35)
v_{2b}	0.2179(17)	0.2181(22)	0.2237(34)	0.2278(39)
$av_{2b}^{(u;1)}$	−0.025 54(35)	−0.024 94(48)	−0.027 75(65)	−0.028 83(96)
$av_{2b}^{(d;1)}$	−0.013 16(21)	−0.013 02(31)	−0.014 44(42)	−0.015 23(68)
$av_{2b}^{(1)}$	−0.012 39(20)	−0.011 91(28)	−0.013 23(42)	−0.013 57(68)
$av_{2b}^{(u;2)}$	0.028 01(41)	0.020 91(58)	0.018 74(79)	0.0152(12)
$av_{2b}^{(d;2)}$	0.014 43(26)	0.011 41(38)	0.010 71(56)	0.009 59(87)
$av_{2b}^{(2)}$	0.013 60(24)	0.009 49(35)	0.007 77(58)	0.005 48(97)
		$\vec{p} = \vec{p}_1$		
$v_{2b}^{(u)}$	0.4153(56)	0.3987(58)	0.419(14)	0.406(19)
$v_{2b}^{(d)}$	0.1966(29)	0.1840(32)	0.1938(77)	0.186(11)
v_{2b}	0.2186(33)	0.2146(40)	0.2255(81)	0.220(10)
$av_{2b}^{(u;1)}$	−0.025 36(49)	−0.023 77(76)	−0.0269(11)	−0.0288(19)
$av_{2b}^{(d;1)}$	−0.013 16(30)	−0.012 51(59)	−0.014 38(63)	−0.0148(14)
$av_{2b}^{(1)}$	−0.012 19(30)	−0.011 44(42)	−0.012 33(82)	−0.0136(15)
$av_{2b}^{(u;2)}$	0.027 61(67)	0.01973(92)	0.0174(15)	0.0151(25)
$av_{2b}^{(d;2)}$	0.014 37(42)	0.010 88(74)	0.0105(10)	0.009 70(179)
$av_{2b}^{(2)}$	0.013 26(44)	0.009 01(56)	0.006 64(125)	0.004 54(215)

where the hopping parameter, κ, is related to the (bare) quark mass via Eq. (36), and we are taking mass degenerate u and d quarks. In Eq. (D2) the quark fields are normalized according to the lattice conventions i.e. they correspond to the continuum fields by rescaling $q \to 1/\sqrt{2\kappa}q$. This introduces a further factor $1/(2\kappa)$ on the rhs of Eq. (39) when using the raw output for the two- and three- point functions. The last term in Eq. (D2), sufficient for on-shell $O(a)$ improvement [with a to be determined function $c_{sw}(g_0)$] has a clover field strength tensor given by

$$F_{\mu\nu}^{\text{clover}}(x) = \frac{1}{8ig_0a^2} \sum_{\pm\mu,\pm\nu} [U_{\mu\nu}^{\square}(x) - U_{\mu\nu}^{\square}(x)^{\dagger}], \tag{D3}$$

where we have extended the definition of the plaquette, so that the μ, ν directions can be negative.

Thus we can rewrite Eq. (D2) as

TABLE X. The bare results for v_2 from Eq. (39) for $\beta = 6.4$, $c_{sw} = 1.526$.

κ	0.1338	0.1342	0.1346	0.1350	0.1353
			$\vec{p} = \vec{p}_1$		
$v_{2a}^{(u)}$	0.368(21)	0.348(30)	0.340(28)	0.339(34)	0.329(63)
$v_{2a}^{(d)}$	0.173(11)	0.161(15)	0.150(12)	0.140(16)	0.122(24)
v_{2a}	0.196(12)	0.185(17)	0.191(18)	0.199(23)	0.205(44)
$av_{2a}^{(u;1)}$	−0.0210(13)	−0.0181(17)	−0.0208(18)	−0.0214(24)	−0.0242(55)
$av_{2a}^{(d;1)}$	−0.010 92(75)	−0.0092(10)	−0.0104(11)	−0.0103(14)	−0.009 81(291)
$av_{2a}^{(1)}$	−0.010 10(73)	−0.008 53(99)	−0.0104(11)	−0.0107(16)	−0.0141(39)
$av_{2a}^{(u;2)}$	0.004 70(68)	−0.002 03(101)	−0.001 81(99)	−0.005 25(152)	−0.0045 6(307)
$av_{2a}^{(d;2)}$	0.003 23(45)	0.000 043(690)	0.000 384(646)	−0.000 831(942)	−0.000 49(205)
$av_{2a}^{(2)}$	0.001 47(44)	−0.002 06(64)	−0.002 29(75)	−0.004 72(121)	−0.003 52(269)
			$\vec{p} = \vec{0}$		
$v_{2b}^{(u)}$	0.4088(27)	0.4140(47)	0.4017(47)	0.3951(70)	0.397(13)
$v_{2b}^{(d)}$	0.1914(14)	0.1881(23)	0.1842(24)	0.1781(39)	0.1674(71)
v_{2b}	0.2174(17)	0.2259(32)	0.2174(31)	0.2169(47)	0.2294(90)
$av_{2b}^{(u;1)}$	−0.022 67(23)	−0.021 91(58)	−0.022 90(38)	−0.022 95(75)	−0.0213(22)
$av_{2b}^{(d;1)}$	−0.011 61(15)	−0.011 18(31)	−0.011 90(28)	−0.012 65(50)	−0.0120(16)
$av_{2b}^{(1)}$	−0.011 05(13)	−0.010 76(34)	−0.010 93(26)	−0.010 28(50)	−0.009 88(133)
$av_{2b}^{(u;2)}$	0.019 43(28)	0.014 06(68)	0.011 22(48)	0.007 28(90)	0.002 06(265)
$av_{2b}^{(d;2)}$	0.010 15(19)	0.007 97(33)	0.006 72(37)	0.006 11(66)	0.004 85(201)
$av_{2b}^{(2)}$	0.009 26(15)	0.006 24(44)	0.004 45(32)	0.001 15(64)	−0.002 61(160)
			$\vec{p} = \vec{p}_1$		
$v_{2b}^{(u)}$	0.4097(42)	0.4059(97)	0.4012(91)	0.385(16)	0.390(34)
$v_{2b}^{(d)}$	0.1906(23)	0.1828(46)	0.1813(46)	0.1697(72)	0.162(15)
v_{2b}	0.2191(27)	0.2231(61)	0.2200(59)	0.215(12)	0.228(25)
$av_{2b}^{(u;1)}$	−0.022 38(38)	−0.0212(11)	−0.022 19(70)	−0.0209(15)	−0.0252(40)
$av_{2b}^{(d;1)}$	−0.011 56(24)	−0.010 44(57)	−0.011 89(48)	−0.012 51(91)	−0.0113(24)
$av_{2b}^{(1)}$	−0.010 79(21)	−0.010 88(67)	−0.010 22(46)	−0.008 44(101)	−0.0142(29)
$av_{2b}^{(u;2)}$	0.018 98(44)	0.0142(13)	0.010 42(78)	0.005 66(163)	0.007 48(445)
$av_{2b}^{(d;2)}$	0.010 03(28)	0.007 31(63)	0.006 75(56)	0.006 39(111)	0.003 58(326)
$av_{2b}^{(2)}$	0.008 95(23)	0.007 06(79)	0.003 66(55)	−0.000 46(124)	0.004 85(456)

$$\sum_{xy} \overline{q}(x) M^{(q)}(x;y) q(y) \to \sum_{x} \{ \overline{q}(x)(\vec{\not{\partial}} + m_q) q(x) + m_{0c} \overline{q}(x) q(x) - \frac{1}{2} a \sum_{\mu} \overline{q}(x) \vec{\Delta}_{\mu}^{-} \vec{\Delta}_{\mu}^{+} q(x) - a\, c_{sw}\, g_0 \sum_{\mu\nu} \frac{1}{4} \overline{q}(x) \sigma_{\mu\nu} F_{\mu\nu}^{\text{clover}}(x) q(x) \}, \tag{D4}$$

where

$$\vec{\Delta}_{\mu}^{+} q(x) = \frac{1}{a}[U_{\mu}(x) q(x + a\hat{\mu}) - q(x)], \qquad \vec{\Delta}_{\mu}^{-} q(x) = \frac{1}{a}[q(x) - U_{\mu}^{\dagger}(x - a\hat{\mu}) q(x - a\hat{\mu})], \tag{D5}$$

so that, see Eq. (30)

TABLE XI. The bare results for v_3 and v_4 from Eq. (39) for $\beta = 6.0$, $c_{sw} = 1.769$.

κ	0.1320	0.1324	0.1333	0.1338	0.1342
			$\vec{p} = \vec{p}_1$		
$v_3^{(u)}$	0.0981(55)	0.0973(52)	0.0961(72)	0.114(15)	0.0955(125)
$v_3^{(d)}$	0.0435(26)	0.0408(25)	0.0378(36)	0.0407(61)	0.0342(62)
v_3	0.0546(34)	0.0566(31)	0.0582(46)	0.0710(98)	0.0606(85)
$v_3^{(u;m_1)}$	0.000 51(215)	0.001 65(221)	0.001 93(373)	0.006 94(638)	−0.000 15(818)
$v_3^{(d;m_1)}$	0.000 07(129)	−0.001 40(145)	−0.002 08(232)	−0.001 58(379)	−0.001 60(557)
$v_3^{(m_1)}$	0.000 45(150)	0.000 56(155)	−0.000 09(298)	0.0103(57)	0.001 13(687)
$v_3^{(u;m_2)}$	−0.004 90(717)	−0.0102(73)	−0.0070(132)	0.0350(269)	−0.0205(311)
$v_3^{(d;m_2)}$	−0.003 42(459)	−0.0108(54)	−0.0118(100)	0.0053(167)	0.0146(227)
$v_3^{m_2}$	−0.001 34(487)	0.000 02(522)	0.0022(112)	0.0308(212)	0.0073(273)
			$\vec{p} = \vec{p}_1$		
$v_4^{(u)}$	0.0272(28)	0.0330(27)	0.0342(39)	0.0339(74)	0.0331(64)
$v_4^{(d)}$	0.0128(15)	0.0128(16)	0.0112(24)	0.0141(42)	0.007 80(485)
v_4	0.0143(19)	0.0201(18)	0.0232(27)	0.0206(55)	0.0253(47)
$v_4^{(u;m_1)}$	−0.005 88(626)	−0.006 25(648)	−0.0048(104)	−0.0109(217)	0.0196(243)
$v_4^{(d;m_1)}$	0.003 55(394)	−0.001 13(450)	0.001 44(708)	0.0086(129)	0.0152(153)
$v_4^{(m_1)}$	−0.009 05(399)	−0.005 12(392)	−0.006 12(799)	−0.0202(144)	0.0047(214)
$v_4^{(u;m_2)}$	−0.006 80(598)	−0.002 73(503)	−0.000 53(841)	−0.0052(198)	0.0163(201)
$v_4^{(d;m_2)}$	0.001 80(298)	−0.001 88(359)	−0.001 42(591)	0.0066(100)	−0.0068(150)
$v_4^{(m_2)}$	−0.008 03(382)	−0.001 69(327)	0.000 38(692)	−0.0085(150)	0.0243(180)
$v_4^{(u;m_3)}/a$	0.0178(75)	0.005 21(918)	0.0082(157)	0.0028(330)	0.0378(447)
$v_4^{(d;m_3)}/a$	0.002 44(508)	0.005 23(630)	0.0093(108)	−0.0111(188)	0.0434(496)
$v_4^{(m_3)}/a$	0.0156(48)	−0.001 27(587)	−0.0041(107)	0.0150(255)	−0.0062(351)

TABLE XII. The bare results for v_3 and v_4 from Eq. (39) for $\beta = 6.2$, $c_{sw} = 1.614$.

κ	0.1333	0.1339	0.1344	0.1349
		$\vec{p} = \vec{p}_1$		
$v_3^{(u)}$	0.1087(72)	0.1075(64)	0.110(11)	0.115(12)
$v_3^{(d)}$	0.0493(36)	0.0467(30)	0.0478(53)	0.0455(54)
v_3	0.0594(41)	0.0608(40)	0.0628(71)	0.0697(88)
$v_3^{(u;m_1)}$	−0.001 41(292)	0.002 78(316)	0.000 02(551)	−0.004 74(849)
$v_3^{(d;m_1)}$	−0.002 33(195)	−0.000 23(187)	−0.003 64(328)	−0.007 16(453)
$v_3^{(m_1)}$	0.000 62(175)	0.002 94(225)	0.003 64(415)	0.003 19(699)
$v_3^{(u;m_2)}$	−0.0011(101)	−0.0038(120)	0.0201(228)	0.0449(378)
$v_3^{(d;m_2)}$	−0.000 88(704)	−0.005 97(810)	0.0136(159)	0.0262(224)
$v_3^{(m_2)}$	0.001 91(505)	0.000 98(765)	0.0135(143)	0.0223(288)
		$\vec{p} = \vec{p}_1$		
$v_4^{(u)}$	0.0333(41)	0.0336(55)	0.0331(69)	0.0442(85)
$v_4^{(d)}$	0.0127(21)	0.0129(30)	0.007 57(388)	0.008 03(561)
v_4	0.0205(31)	0.0206(37)	0.0238(62)	0.0344(89)
$v_4^{(u;m_1)}$	0.0212(116)	−0.0113(147)	0.0339(256)	0.0375(332)
$v_4^{(d;m_1)}$	0.0102(69)	−0.0117(90)	0.0176(136)	0.0088(189)
$v_4^{(m_1)}$	0.0118(73)	0.0083(906)	0.0172(180)	0.0270(267)
$v_4^{(u;m_2)}$	0.0074(101)	−0.0007(130)	0.0157(213)	0.0357(310)
$v_4^{(d;m_2)}$	0.000 98(587)	−0.003 84(687)	0.0082(121)	0.0092(168)
$v_4^{(m_2)}$	0.007 56(606)	0.002 18(904)	0.0100(150)	0.0257(249)
$v_4^{(u;m_3)}/a$	0.0225(166)	−0.0201(192)	0.0324(364)	−0.0271(578)
$v_4^{(d;m_3)}/a$	0.0090(110)	−0.000 02(1285)	0.0198(235)	0.0291(390)
$v_4^{(m_3)}/a$	0.0136(104)	−0.0134(129)	0.0097(275)	−0.0520(497)

TABLE XIII. The bare results for v_3 and v_4 from Eq. (39) for $\beta = 6.4$, $c_{sw} = 1.526$.

κ	0.1338	0.1342	0.1346	0.1350	0.1353
			$\vec{p} = \vec{p}_1$		
$v_3^{(u)}$	0.0996(62)	0.0940(88)	0.0963(88)	0.0946(104)	0.0898(179)
$v_3^{(d)}$	0.0453(31)	0.0413(43)	0.0410(40)	0.0381(54)	0.0313(71)
v_3	0.0543(37)	0.0516(53)	0.0554(59)	0.0564(73)	0.0594(148)
$v_3^{(u;m_1)}$	−0.000 88(245)	−0.002 50(353)	−0.000 78(418)	−0.006 59(559)	0.0180(117)
$v_3^{(d;m_1)}$	−0.001 66(157)	−0.003 01(268)	−0.002 70(278)	−0.004 00(400)	−0.003 94(813)
$v_3^{(m_1)}$	−0.000 56(139)	0.000 18(258)	0.001 87(271)	−0.003 41(430)	0.0213(109)
$v_3^{(u;m_2)}$	−0.0194(96)	0.0135(151)	−0.0240(180)	−0.0067(282)	0.0557(718)
$v_3^{(d;m_2)}$	−0.0121(58)	−0.0016(110)	−0.0152(110)	−0.0039(183)	0.0212(400)
$v_3^{(m_2)}$	−0.007 52(523)	0.0131(103)	−0.0057(109)	0.0014(191)	0.0383(582)
			$\vec{p} = \vec{p}_1$		
$v_4^{(u)}$	0.0355(49)	0.0382(78)	0.0440(75)	0.0496(104)	0.0403(227)
$v_4^{(d)}$	0.0136(32)	0.0136(37)	0.0161(52)	0.0169(72)	0.0303(137)
v_4	0.0216(27)	0.0249(51)	0.0280(46)	0.0310(69)	0.0123(166)
$v_4^{(u;m_1)}$	−0.0067(130)	0.0220(222)	−0.0004(217)	0.0194(296)	0.0878(792)
$v_4^{(d;m_1)}$	−0.006 81(859)	0.0162(132)	−0.0117(151)	0.0046(199)	0.0106(445)
$v_4^{(m_1)}$	−0.000 76(782)	−0.0006(131)	0.0085(154)	0.0107(230)	0.0727(662)
$v_4^{(u;m_2)}$	−0.0055(112)	0.0189(190)	0.0064(187)	0.0356(250)	0.0102(674)
$v_4^{(d;m_2)}$	−0.000 45(699)	0.0163(141)	0.0050(117)	0.0264(164)	0.0029(425)
$v_4^{(m_2)}$	−0.004 55(654)	0.001 74(936)	−0.0001(127)	0.0075(208)	0.0038(508)
$v_4^{(u;m_3)}/a$	0.0306(223)	−0.0381(328)	0.0570(402)	0.142(57)	−0.069(127)
$v_4^{(d;m_3)}/a$	0.0182(124)	−0.0155(202)	0.0325(221)	0.0961(364)	−0.100(84)
$v_4^{(m_3)}/a$	0.0072(124)	−0.0211(255)	0.0100(259)	0.0488(449)	0.028(108)

$$\vec{D}_\mu = \frac{1}{2}(\vec{\Delta}_\mu^+ + \vec{\Delta}_\mu^-). \tag{D6}$$

am_q is defined in Eq. (36) and

$$am_{0c} = \frac{1}{2}\left(\frac{1}{\kappa_c} - 8\right). \tag{D7}$$

In this latter form Eq. (D4) shows the additional $O(a)$ operators most clearly.

In Table VII we give our parameter values used in the quenched fermion simulations together with the pseudoscalar, am_{ps}, and nucleon mass, am_N.

We now give a series of tables tabulating the bare matrix elements $v_n^{(q)}$ for $q = u$, d, the improvement operator matrix elements $av_2^{(q;i)}$ for $i = 1, 2$ and the mixing operators $v_3^{(q;m_i)}$ for $i = 1, 2$ and $v_4^{(q;m_i)}$ for $i = 1, 2$ and $v_4^{(q;m_3)}/a$.

[1] The nucleon states are normalized with the convention $\langle N(\vec{p})|N(\vec{p}')\rangle = (2\pi)^3 2E_{\vec{p}}\delta(\vec{p} - \vec{p}')$.

[2] G. Martinelli and C. T. Sachrajda, Nucl. Phys. **B316**, 355 (1989).

[3] M. Göckeler, R. Horsley, E.-M. Ilgenfritz, H. Perlt, P. Rakow, G. Schierholz, and A. Schiller, Phys. Rev. D **53**, 2317 (1996).

[4] M. Göckeler, R. Horsley, E.-M. Ilgenfritz, H. Perlt, P. Rakow, G. Schierholz, and A. Schiller, Nucl. Phys. Proc. Suppl. **B53**, 81 (1997).

[5] T. Bakeyev, D. Galletly, M. Göckeler, M. Gürtler, R. Horsley, B. Joó, A. D. Kennedy, B. Pendleton, H. Perlt, D. Pleiter, P. E. L. Rakow, G. Schierholz, A. Schiller, T. Streuer, and H. Stüben, Nucl. Phys. Proc. Suppl. **B128**,

82 (2004).
[6] Additionally all our current lattices have $m_{ps}L_S \gtrsim 4$ where m_{ps} is the pseudoscalar mass.
[7] M. Göckeler, R. Horsley, D. Pleiter, P. E. L. Rakow, and G. Schierholz, hep-ph/0108105.
[8] A. W. Thomas, W. Melnitchouk, and F. M. Steffens, Phys. Rev. Lett. **85**, 2892 (2000).
[9] W. Detmold, W. Melnitchouk, J. W. Negele, D. B. Renner, and A. W. Thomas, Phys. Rev. Lett. **87**, 172001 (2001).
[10] J.-W. Chen and X. Ji, Phys. Lett. B **523**, 107 (2001).
[11] J.-W. Chen and X. Ji, Phys. Rev. Lett. **87**, 152002 (2001).
[12] D. Arndt and M. J. Savage, Nucl. Phys. **A697**, 429 (2002).
[13] J.-W. Chen and M. J. Savage, Nucl. Phys. **A707**, 452 (2002).
[14] T. van Ritbergen, J. A. M. Vermaseren, and S. A. Larin, Phys. Lett. B **400**, 379 (1997).
[15] J. A. M. Vermaseren, S. A. Larin, and T. van Ritbergen, Phys. Lett. B **405**, 327 (1997).
[16] W. A. Bardeen, A. J. Buras, D. W. Duke, and T. Muta, Phys. Rev. D **18**, 3998 (1978).
[17] W. L. van Neerven and E. B. Zijlstra, Phys. Lett. B **272**, 127 (1991).
[18] S. A. Larin and J. A. M. Vermaseren, Z. Phys. C **57**, 93 (1993).
[19] W. L. van Neerven and A. Vogt, Nucl. Phys. **B568**, 263 (2000).
[20] A. González-Arroyo, C. López, and F. J. Ynduráin, Nucl. Phys. **B153**, 161 (1979).
[21] S. A. Larin, T. van Ritbergen, and J. A. M. Vermaseren, Nucl. Phys. **B427**, 41 (1994).
[22] A. Rétey and J. A. M. Vermaseren, Nucl. Phys. **B604**, 281 (2001).
[23] R. Sommer, Nucl. Phys. **B411**, 839 (1994).
[24] S. Capitani, M. Lüscher, R. Sommer, and H. Wittig, Nucl. Phys. **B544**, 669 (1999).
[25] S. Booth, M. Göckeler, R. Horsley, A. C. Irving, B. Joó, S. Pickles, D. Pleiter, P. E. L. Rakow, G. Schierholz, Z. Sroczynski, and H. Stüben, Phys. Lett. B **519**, 229 (2001).
[26] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Eur. Phys. J. C **28**, 455 (2003).
[27] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. Nadolsky, and W. K. Tung, J. High Energy Phys. 07 (2002) 12.
[28] There are many references to the relationship between structure functions and parton densities. See for example [29].
[29] J. T. Londergan and A. W. Thomas, Prog. Part. Nucl. Phys. **41**, 49 (1998).
[30] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Eur. Phys. J. C **23**, 73 (2002).
[31] See http://durpdg.dur.ac.uk/hepdata/pdf3.html.
[32] A. Arneodo *et al.*, Phys. Rev. D **50**, R1 (1994).
[33] A. Bruell, Habilitation thesis, Ruprecht-Karls-Universität Heidelberg, 1995.
[34] P. Amaudruz *et al.*, Nucl. Phys. **B371**, 3 (1992).
[35] M. Guagnelli, R. Sommer, and H. Wittig, Nucl. Phys. **B535**, 389 (1998).
[36] Note that a recent analysis [37] gives similar results of $v_2^{\overline{\rm MS}}(Q_0) \approx 0.159$ and $v_4^{\overline{\rm MS}}(Q_0) \approx 0.0245$.
[37] J. Blümlein, H. Böttcher, and A. Guffanti, Nucl. Phys. Proc. Suppl. **B135**, 152 (2004).
[38] Our Euclidean conventions are described in [39].
[39] C. Best, M. Göckeler, R. Horsley, E.-M. Ilgenfritz, H. Perlt, P. Rakow, A. Schäfer, G. Schierholz, A. Schiller, and S. Schramm, Phys. Rev. D **56**, 2743 (1997).
[40] $\mathcal{O}_{v_{2a}}$, $\mathcal{O}_{v_{2b}}$ are the same operators as we used previously in [3], while there are some modifications to $\mathcal{O}_{v_3}$ and $\mathcal{O}_{v_4}$. For $\mathcal{O}_{v_3}$ we have effectively made the proper rotation (i.e. preserving γ_5) of $4 \leftrightarrow 1$ (and $2 \leftrightarrow 3$) of the operator in [3]. This means that the measured ratio in Eq. (39) is $-ip_1 E_{\vec{p}_1} v_3$ rather than $-p_1^2 v_3$ and hence the signal is better by a factor $O(E_{\vec{p}_1}/p_1)$. For $\mathcal{O}_{v_4}$ we have symmetrized over the 2, 3 indices of the operator given in [3], which makes the measurement of the ratio for the new operator numerically a little less noisy.
[41] R. C. Brower, S. Huang, J. W. Negele, A. Pochinsky, and B. Schreiber, Nucl. Phys. Proc. Suppl. **B53**, 318 (1997).
[42] M. Göckeler, R. Horsley, E.-M. Ilgenfritz, H. Perlt, P. Rakow, G. Schierholz, and A. Schiller, Phys. Rev. D **54**, 5705 (1996).
[43] Using the operators given in [3] we would have $\mathcal{O}_{v_3}^{m_1} = \mathcal{O}^\gamma_{\langle\langle 411\rangle\rangle} - \frac{1}{2}(\mathcal{O}^\gamma_{\langle\langle 422\rangle\rangle} + \mathcal{O}^\gamma_{\langle\langle 433\rangle\rangle})$, $\mathcal{O}_{v_3}^{m_2} = 2\mathcal{O}^{\gamma\gamma_5}_{||213||}$ and $\mathcal{O}_{v_4}^{m_1} = -\mathcal{O}^\gamma_{1144} - \mathcal{O}^\gamma_{4114} - \mathcal{O}^\gamma_{1441} - \mathcal{O}^\gamma_{4411} + 2\mathcal{O}^\gamma_{1414} + 2\mathcal{O}^\gamma_{4141} - \mathcal{O}^\gamma_{2233} - \mathcal{O}^\gamma_{3223} - \mathcal{O}^\gamma_{2332} - \mathcal{O}^\gamma_{3322} + 2\mathcal{O}^\gamma_{2323} + 2\mathcal{O}^\gamma_{3232} + \mathcal{O}^\gamma_{1133} + \mathcal{O}^\gamma_{3113} + \mathcal{O}^\gamma_{1331} + \mathcal{O}^\gamma_{3311} - 2\mathcal{O}^\gamma_{1313} - 2\mathcal{O}^\gamma_{3131} + \mathcal{O}^\gamma_{2244} + \mathcal{O}^\gamma_{4224} + \mathcal{O}^\gamma_{2442} + \mathcal{O}^\gamma_{4422} - 2\mathcal{O}^\gamma_{2424} - 2\mathcal{O}^\gamma_{4242}$, $\mathcal{O}_{v_4}^{m_2} = +\mathcal{O}^{\gamma\gamma_5}_{1234} - \mathcal{O}^{\gamma\gamma_5}_{3214} - \mathcal{O}^{\gamma\gamma_5}_{1432} + \mathcal{O}^{\gamma\gamma_5}_{3412} + \mathcal{O}^{\gamma\gamma_5}_{2143} - \mathcal{O}^{\gamma\gamma_5}_{4123} - \mathcal{O}^{\gamma\gamma_5}_{2341} + \mathcal{O}^{\gamma\gamma_5}_{4321} + \mathcal{O}^{\gamma\gamma_5}_{1243} - \mathcal{O}^{\gamma\gamma_5}_{4213} - \mathcal{O}^{\gamma\gamma_5}_{1342} + \mathcal{O}^{\gamma\gamma_5}_{4312} + \mathcal{O}^{\gamma\gamma_5}_{2134} - \mathcal{O}^{\gamma\gamma_5}_{3124} - \mathcal{O}^{\gamma\gamma_5}_{2431} + \mathcal{O}^{\gamma\gamma_5}_{3421}$, $\mathcal{O}_{v_4}^{m_3} = i(\mathcal{O}^{\sigma\gamma_5}_{2314} - \mathcal{O}^{\sigma\gamma_5}_{2341} + \mathcal{O}^{\sigma\gamma_5}_{1423} - \mathcal{O}^{\sigma\gamma_5}_{1432} + \mathcal{O}^{\sigma\gamma_5}_{2413} - \mathcal{O}^{\sigma\gamma_5}_{2431} + \mathcal{O}^{\sigma\gamma_5}_{1324} - \mathcal{O}^{\sigma\gamma_5}_{1342})$.
[44] S. Capitani, M. Göckeler, R. Horsley, H. Perlt, P. E. L. Rakow, G. Schierholz, and A. Schiller, Nucl. Phys. **B593**, 183 (2001).
[45] Often, alternatively, the improved operator is rewritten as $(1 + am_q b_\mathcal{O})(\mathcal{O} + \sum_{i=1}^3 ac_i' \mathcal{O}^{(i)})$. In the remainder of this article it should hopefully be clear from context whether we are referring to the operator or its improved version.
[46] M. Göckeler, R. Horsley, E.-M. Ilgenfritz, H. Perlt, P. Rakow, G. Schierholz, and A. Schiller, Nucl. Phys. Proc. Suppl. **B42**, 337 (1995).
[47] M. Göckeler, R. Horsley, E.-M. Ilgenfritz, H. Oelrich, H. Perlt, P. E. L. Rakow, G. Schierholz, A. Schiller, and P. Stephenson, Nucl. Phys. Proc. Suppl. **B53**, 324 (1997).
[48] M. Lüscher, S. Sint, R. Sommer, P. Weisz, and U. Wolff, Nucl. Phys. **B491**, 323 (1997).
[49] M. Baake, B. Gemünden, and R. Oedingen, J. Math. Phys. (N.Y.) **24**, 1021 (1983).
[50] M. Göckeler, R. Horsley, E.-M. Ilgenfritz, H. Perlt, P. Rakow, G. Schierholz, and A. Schiller, Nucl. Phys. **B472**, 309 (1996).
[51] G. Martinelli, C. Pittori, C. T. Sachrajda, M. Testa, and A. Vladikas, Nucl. Phys. **B445**, 81 (1995).
[52] For a general review of lattice perturbation theory, see [53].
[53] S. Capitani, Phys. Rep. **382**, 113 (2003).
[54] For diagonal elements we write $B^S_{\mathcal{O}_i;\mathcal{O}_i} \equiv B^S_{\mathcal{O}_i}$ and $d_{\mathcal{O}_i;\mathcal{O}_i;0} \equiv d_{\mathcal{O}_i;0}$.
[55] S. Capitani (private communication).

[56] The number for $c_{sw} = 0$ was incorrectly found in [3,50]. We shall use the result of [57] in the following.
[57] S. Capitani, Nucl. Phys. **B597**, 313 (2001).
[58] S. Capitani, M. Göckeler, R. Horsley, D. Pleiter, P. Rakow, H. Stüben, and G. Schierholz, Nucl. Phys. Proc. Suppl. **B106**, 299 (2002).
[59] Appropriate numerical values of $u_0^4 = \langle \frac{1}{3} \mathrm{Tr} U_\Box \rangle$ are given in [25]. For other g_0^2 numbers a simple interpolation between these values can be performed, or alternatively a Padé fit, including the known first three coefficients of the plaquette expansion, can be made.
[60] M. Lüscher and P. Weisz, Phys. Lett. B **349**, 165 (1995).
[61] T. Bakeyev, M. Göckeler, R. Horsley, D. Pleiter, P. E. L. Rakow, G. Schierholz, and H. Stüben, Phys. Lett. B **580**, 197 (2004).
[62] K. G. Chetyrkin and A. Rétey, Nucl. Phys. **B583**, 3 (2000).
[63] M. Göckeler, R. Horsley, H. Oelrich, H. Perlt, D. Petters, P. E. L. Rakow, A. Schäfer, G. Schierholz, and A. Schiller, Nucl. Phys. **B544**, 699 (1999).
[64] M. Göckeler, R. Horsley, H. Oelrich, H. Perlt, D. Petters, P. E. L. Rakow, A. Schäfer, G. Schierholz, and A. Schiller, Nucl. Phys. Proc. Suppl. **B73**, 291 (1999).
[65] J. A. Gracey, Nucl. Phys. **B667**, 242 (2003).
[66] K. G. Chetyrkin and A. Rétey, hep-ph/0007088.
[67] M. Göckeler *et al.* (unpublished).
[68] For an independent nonperturbative calculation of $Z^{\mathrm{RGI}}_{\nu_{2b}}$ see [69] giving 1.21(5), 1.32(6) and 1.37(6) for $\beta = 6.0$, 6.2, 6.4 respectively. These results are about 15%–10% smaller than our numbers. The numbers do not have to agree though, as they can have different $O(a^2)$ errors. Indeed the ratio of the two numbers against $(a/r_0)^2$ slopes towards 1 as expected, indicating that the results are compatible with each other. Appropriate continuum extrapolations of the renormalized matrix elements confirm this giving answers agreeing to about 5%.
[69] M. Guagnelli, K. Jansen, F. Palombi, R. Petronzio, A. Shindler, and I. Wetzorke, Eur. Phys. J. C **40**, 69 (2005).
[70] These works find the leading chiral logarithm behavior, $F_\chi^{(n)}(x) \sim b_n(1 - cx^2 \ln(x^2/(r_0\Lambda_\chi)^2))$, which is built into the model of Eq. (80).
[71] M. Göckeler, R. Horsley, W. Kürzinger, H. Oelrich, D. Pleiter, P. E. L. Rakow, A. Schäfer, and G. Schierholz, Phys. Rev. D **63**, 074506 (2001).
[72] M. Göckeler, R. Horsley, D. Pleiter, P. E. L. Rakow, A. Schäfer, and G. Schierholz, Nucl. Phys. Proc. Suppl. **B119**, 398 (2003).
[73] D. Dolgov, R. Brower, S. Capitani, P. Dreher, J. W. Negele, A. Pochinsky, D. B. Renner, N. Eicker, Th. Lippert, K. Schilling, R. G. Edwards, and U. M. Heller, Phys. Rev. D **66**, 034506 (2002).
[74] D. Galletly, M. Gürtler, R. Horsley, B. Joó, A. D. Kennedy, H. Perlt, B. J. Pendleton, P. E. L. Rakow, G. Schierholz, A. Schiller, and T. Streuer, Nucl. Phys. Proc. Suppl. **B129 & B130**, 456 (2004).
[75] M. Baake, B. Gemünden, and R. Oedingen, J. Math. Phys. (N.Y.) **23**, 944 (1982).
[76] F. E. Close, *An Introduction to Quarks and Partons* (Academic Press, London, 1981).