QCD with light Wilson quarks on fine lattices (II)

CERN-PH-TH/2006-261

QCD with light Wilson quarks on fine lattices (II): DD-HMC simulations and data analysis

L. Del Debbio[1], L. Giusti[2,*], M. Lüscher[2], R. Petronzio[3], N. Tantalo[3,4]

[1] *SUPA, School of Physics, University of Edinburgh, Edinburgh EH9 3JZ, UK*

[2] *CERN, Physics Department, TH Division, CH-1211 Geneva 23, Switzerland*

[3] *Università di Roma "Tor Vergata" and INFN sezione "Tor Vergata", Via della Ricerca Scientifica 1, I-00133 Rome, Italy*

[4] *Centro Enrico Fermi, Via Panisperna 89 A, I-00184 Rome, Italy*

Abstract

In this second report on our recent numerical simulations of two-flavour QCD, we provide further technical details on the simulations and describe the methods we used to extract the meson masses and decay constants from the generated ensembles of gauge fields. Among the topics covered are the choice of the DD-HMC parameters, the issue of stability, autocorrelations and the statistical error analysis. Extensive data tables are included as well as a short discussion of the quark-mass dependence in partially quenched QCD, supplementing the physics analysis that was presented in the first paper in this series.

1. Introduction

Lattice QCD with Wilson quarks [1] has seen important algorithmic developments in the last few years [2–8]. As a consequence, a large range of lattice spacings, lattice volumes and quark masses can now be explored, using numerical simulations, thus providing new physics opportunities and a greater lever arm for the extrapolations to the continuum and the chiral limit. Our recent work [9] was the first to fully profit from the technical breakthrough and several other projects, simulating QCD with

* On leave from Centre de Physique Théorique, CNRS Luminy, F-13288 Marseille, France

two [10,11] and three [12–14] flavours of light Wilson quarks, or with two flavours and a twisted mass term [15], are currently underway, all heavily depending on the new generation of algorithms.

The present paper is the second in a series of two papers devoted to the study of two-flavour QCD at small quark masses and lattice spacings. In the first paper [9], the focus was on the physics results, while here we give a fairly detailed technical account of the simulations that we have performed.

Perhaps the most important items that we discuss are the stability of the simulations (sect. 3) and the pattern of autocorrelation times observed in our runs (sect. 4). We also describe, in sect. 5, the methods that we used to extract the meson masses and decay constants from the generated ensembles of gauge fields (extensive data tables are included in appendix C). The paper ends with an addendum to the first paper, where we briefly discuss the quark-mass dependence of various quantities in partially quenched QCD with $2+1$ flavours of quarks.

2. Simulation parameters

We consider the Wilson formulation of lattice QCD, optionally O(a)-improved, with a doublet of mass-degenerate sea quarks. The notation and normalization conventions adopted in this paper coincide with those already used in our previous paper [9]. In particular, the parameters of the lattice theory are the inverse bare coupling β, the sea-quark hopping parameter $\kappa_{\rm sea}$ and the coefficient $c_{\rm sw}$ of the Sheikoleslami-Wohlert improvement term [16,17].

All simulations reported here were performed using the DD-HMC simulation algorithm [7]. As suggested by the name, the algorithm combines domain-decomposition ideas with the HMC algorithm [18]. More precisely, by dividing the lattice into non-overlapping rectangular blocks, a natural separation of the high-frequency from the low-frequency modes of the fields is achieved. Following Sexton and Weingarten [19], the different modes are then evolved using different molecular-dynamics step sizes, which results in a significant acceleration of the simulation.

On a given lattice and at fixed coupling, the simulations progressed from the larger to the smaller quark masses, normally skipping 1500 molecular-dynamics trajectories for thermalization. The number $N_{\rm trj}$ of trajectories generated after thermalization, the separation $N_{\rm sep}$ (in numbers of trajectories) between successive saved field configurations and the number $N_{\rm cfg}$ of saved fields are given in table 1. Different runs at

Table 1. Lattice parameters and simulation statistics

Run	Lattice	β	$c_{\rm sw}$	$\kappa_{\rm sea}$	$N_{\rm trj}$	$N_{\rm sep}$	$N_{\rm cfg}$
A_{1a}	32×24^3	5.6	0	0.15750	6300	100	64
A_{1b}				0.15750	5070	30	169
A_2				0.15800	10800	100	109
A_{3a}				0.15825	6100	100	62
A_{3b}				0.15825	3800	100	38
A_4				0.15835	4950	50	100
B_1	64×32^3	5.8	0	0.15410	5050	50	100
B_2				0.15440	5200	50	101
B_3				0.15455	5150	50	104
B_4				0.15462	5050	50	102
C_1	64×24^3	5.6	0	0.15800	3450	30	116
D_1	48×24^3	5.3	1.90952	0.13550	5150	50	104
D_2				0.13590	5130	30	171
D_3				0.13610	5040	30	168
D_4				0.13620	5010	30	168
D_5				0.13625	5040	30	169
E_1	64×32^3	5.3	1.90952	0.13550	5344	32	168
E_2				0.13590	5024	32	158
E_3				0.13605	5024	32	158

the same lattice parameters (such as A_{3a} and A_{3b}) are distinguished by a lower-case latin index. In our previous paper [9], only the runs A_{1a}, A_2, A_{3a}, A_{3b}, B_1–B_4 and D_1–D_5 were included in the physics analysis. The other runs listed in table 1 merely serve, in sections 3 and 4, to clarify some technical issues.

The DD-HMC simulation algorithm was implemented following the lines of ref. [7]. In particular, for the solution of the Dirac equation on the full lattice, the Schwarz-preconditioned GCR solver described in ref. [6] was used. The so-called replay trick, however, was switched off in the more recent simulations A_{3b}–E_3, because trajectory replays would have been rare and hardly worth the extra effort (see subsect. 3.3).

No attempt was made to tune the DD-HMC algorithm and most of its parameters were actually set to some fixed values, the same as the ones already chosen in ref. [7].

Table 2. DD-HMC parameters, acceptance rate and average solver iteration numbers

Run	Block size	N_2	$P_{\rm acc}$	$\langle N_{\rm GCR}\rangle$	$\langle N_{\rm CG}\rangle$
A_{1a}	$8\times 6^2\times 12$	5	0.81*	23	73
A_{1b}	$8\times 6\times 12^2$	5	0.82	22	89
A_2	$8\times 6^2\times 12$	6	0.79*	39	74
A_{3a}		10	0.89*	54	75
A_{3b}		10	0.86	54	75
A_4		16	0.91	73	75
B_1	$8^3\times 16$	8	0.84	32	85
B_2		10	0.89	52	87
B_3		12	0.87	74	87
B_4		14	0.92	90	88
C_1	$8\times 6\times 12^2$	7	0.81	41	92
D_1	$6^2\times 12^2$	7	0.81	25	120
D_2		8	0.80	41	123
D_3		12	0.87	58	124
D_4		14	0.87	73	125
D_5		18	0.89	87	125
E_1	8^4	9	0.80	25	121
E_2		11	0.84	41	124
E_3		13	0.83	53	125

* Transition probability includes trajectory replays

Among these were the trajectory length $\tau = 0.5$, the integration step numbers $N_0 = 4$ and $N_1 = 5$ associated to the gauge and block fermion forces as well as the admitted tolerances $(r_1, r_2, \tilde{r}_1, \tilde{r}_2) = (10^{-8}, 10^{-7}, 10^{-11}, 10^{-10})$ for the numerical solution of the Dirac equation on the blocks and the full lattice †. The parameters of the Schwarz-preconditioned GCR solver were fixed to the values quoted in ref. [6],

† The trajectory length τ and thus the integration step sizes τ/N_2, etc., refer to a particular normalization of the kinetic term in the molecular-dynamics Hamiltonian. Here the normalizations are the same as in ref. [7], i.e. the term is assumed to be equal to $\frac{1}{2}(\Pi,\Pi) = \sum_{x,\mu} {\rm tr}\{\Pi(x,\mu)^\dagger \Pi(x,\mu)\}$, where $\Pi(x,\mu)$ denotes the canonical momentum of the link variable $U(x,\mu)$.

except for the number $n_{\rm kv}$ of Krylov vectors generated before the GCR recursion is restarted, which was set to 32 in run D_5 and to 24 in all other runs.

What remains to be specified are then the size of the blocks on which the algorithm operates and the integration step number N_2 associated to the block interaction term in the molecular-dynamics Hamiltonian (see table 2). In practice the latter must be increased as one moves to lighter quark masses in order to preserve a high acceptance rate $P_{\rm acc}$. The average number $N_{\rm GCR}$ of GCR solver iterations needed along the trajectories also depends on N_2 (it decreases when N_2 goes up), while the average number $N_{\rm CG}$ of conjugate-gradient iterations required for the computation of the block terms in the molecular-dynamics equations is largely determined by the block size.

With the chosen parameters, the reversibility of the molecular-dynamics trajectories is guaranteed to high precision. In the tests that we have performed, the average absolute deviation of the components of the link variables after a return trajectory was at most 3×10^{-9}, while in the case of the Hamiltonian the observed differences were less than 4×10^{-6}. Deviations larger than 10 times the average occurred in less than 1% of the cases and never went beyond 100 times the average.

3. Spectral gap and stability issues

The Wilson–Dirac operator preserves chiral symmetry only up to lattice effects and is therefore not rigorously protected from having eigenvalues much smaller than the quark mass. Exceptionally small eigenvalues do not invalidate the theory but may lead to instabilities in numerical simulations, depending, to some extent, on which simulation algorithm is used.

3.1 Spectral gap of the Dirac operator

In a previous dedicated study [20], we computed the distribution of the spectral gap of the hermitian lattice Dirac operator on the lattices $A_1 - A_4$, B_1, B_2, C_1 and D_1. The distributions turned out to be well separated from the origin, thus showing, a posteriori, that the simulations were safe of exceptionally small eigenvalues and the associated instabilities. Moreover, based on the observed scaling properties of the distributions on the A, B and C lattices, we argued that this will always be so in the large-volume regime of the unimproved Wilson theory.

The gap distributions have now also been computed on the lattices B_3, B_4, $D_2 - D_5$, E_2 and E_3. In the following, however, we focus on the improved theory, because

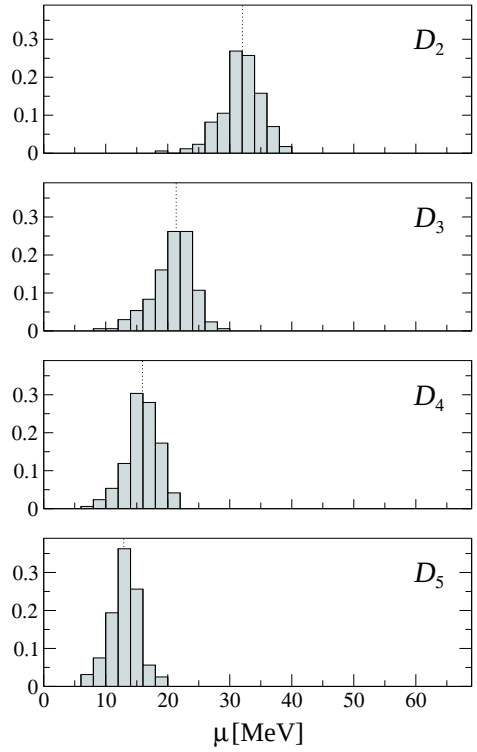

Fig. 1. Normalized histograms of the (unrenormalized) spectral gap μ of the hermitian lattice Dirac operator, as obtained in the runs $D_2 - D_5$. The bin size is 2 MeV and the dotted vertical lines indicate the position of the median $\bar{\mu}$ of the distributions. The data were converted to physical units using $a = 0.0784$ fm [9].

the results obtained on the B lattices are fully in line with the behaviour expected from our previous paper [20].

At first sight, the gap distributions in the improved theory look similar to those in the unimproved theory (see fig. 1). In particular, they are well separated from the origin, on all lattices that we have simulated, and the median of the distributions again turns out to be a practically linear function of the sea-quark mass (fig. 2).

However, the dependence of the width σ of the distributions on the quark mass and the lattice size is different (see table 3) †. In the case of the D-series of lattices, for example, the width decreases by as much as a factor of 1.5 from the largest to the smallest quark mass, while no obvious mass-dependence was seen on the A and

† Following ref. [20], we define the width of the distributions through $\sigma = \frac{1}{2}(v-u)$, where $[u,v]$ is the smallest range in μ, which contains more than 68.3% of the data.

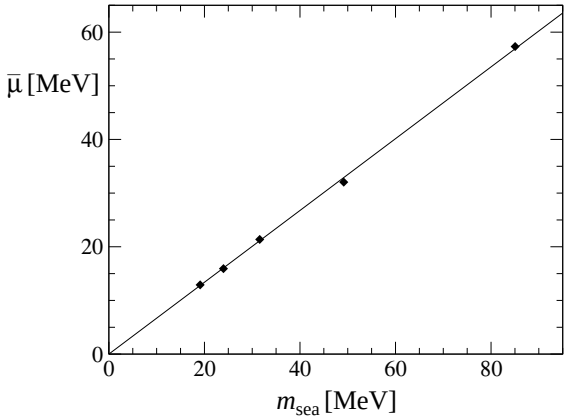

Fig. 2. Median $\bar{\mu}$ of the gap distributions obtained in runs $D_1 - D_5$ (data points), plotted as a function of the bare sea-quark mass $m_{\rm sea}$ (see subsect. 5.2 for the precise definition of the latter). The line is a linear fit of the data without constant term.

B lattices. Moreover, σ does not appear to scale proportionally to the inverse square root of the (four-dimensional) volume V of the lattice (see fig. 3). The widths on the lattices D_2 and E_2, for example, turned out to be nearly the same, contrary to what was expected on the basis of the experience made in the unimproved theory.

Another perhaps not unrelated observation is that the median of the distribution on the D and E lattices is always smaller than the threshold of the spectral density in infinite volume, which we expect to be at $Z_{\rm A} m_{\rm sea}$ [20], $Z_{\rm A}$ being the axial-current renormalization constant ($Z_{\rm A} = 0.75(1)$ on these lattices [21]). The spectral density in finite volume thus has a tail that extends a few MeV below the threshold. On the other hand, the values quoted in table 3 of the average splitting $\langle \Delta \rangle$ of the lowest four eigenvalues suggest that the tail scales to zero in the infinite-volume limit, as it has to be if the density in infinite volume does not extend all the way to zero [20].

At present, however, there is still no theoretical understanding of the dependence of the gap distribution on the quark mass and the lattice size. In particular, the fact that the improved and the unimproved theory behave differently in this respect remains unexplained. Partially quenched (Wilson) chiral perturbation theory may be a framework in which these questions can be addressed [22] and further insight may perhaps also be gained by studying the localization properties of the eigenfunctions and the convergence of the spectral density to the infinite-volume limit. It would be interesting to know, for example, whether the spectral gap coincides with the mobility edge [23] and whether the tail of the spectral density below $Z_{\rm A} m_{\rm sea}$ does in fact disappear in the infinite-volume limit.

Table 3. Median and width of the gap distributions in the improved theory

Run	$\bar{\mu}$	σ	$\bar{\mu} - Z_{\rm A} m_{\rm sea}$	$\langle\Delta\rangle$
D_1	57.3(6)	3.3(4)	−6.5(10)	2.48(12)
D_2	32.0(3)	2.79(24)	−4.8(6)	2.39(7)
D_3	21.4(3)	2.84(23)	−2.3(4)	2.29(7)
D_4	15.9(3)	2.33(18)	−2.1(3)	2.23(6)
D_5	12.9(4)	1.99(15)	−1.4(4)	2.28(5)
E_2	30.3(3)	2.58(19)	−6.6(6)	1.69(8)
E_3	21.3(3)	2.31(19)	−5.8(5)	1.52(7)

All entries are given in MeV

3.2 Accessible range of pion masses on the D and E lattices

When the sea-quark mass decreases, the gap distribution becomes sharper and moves closer to the origin. Eventually the probability for exceptionally small eigenvalues is not completely negligible anymore and one may run into algorithmic instabilities. We have not reached this point yet and consequently cannot say in which way the DD-HMC simulations will be affected. However, in order to be on the safe side, one may prefer to stay in the range of parameters where the gap distribution is well separated from the origin, i.e. where, say, the inequality $\bar{\mu} \geq 3\sigma$ holds [20].

On a given lattice, this bound sets a lower limit on the accessible sea-quark masses and thus on the masses M_π of the pions (the lightest pseudo-scalar mesons made of the sea quarks). Furthermore, if large finite-volume effects are to be avoided, the bound $M_\pi L \geq 3$ (where L denotes the spatial lattice size) should better be respected as well.

In the case of the D and E lattices, the range of pion masses where both conditions are fulfilled can be determined explicitly, using our simulation results. An extrapolation in the sea-quark mass is however still required, but it seems reasonable to extrapolate $\bar{\mu}$ and M_π^2 linearly [9] and to assume that σ drops to values below 2 MeV at small quark masses. For the accessible range of pion masses, we then obtain

$$M_\pi \geq \begin{cases} 314 \text{ MeV} & (D \text{ lattices}), \\ 270 \text{ MeV} & (E \text{ lattices}), \end{cases} \tag{3.1}$$

where the limit is set by the constraint $M_\pi L \geq 3$ on the D lattices. This is not so on

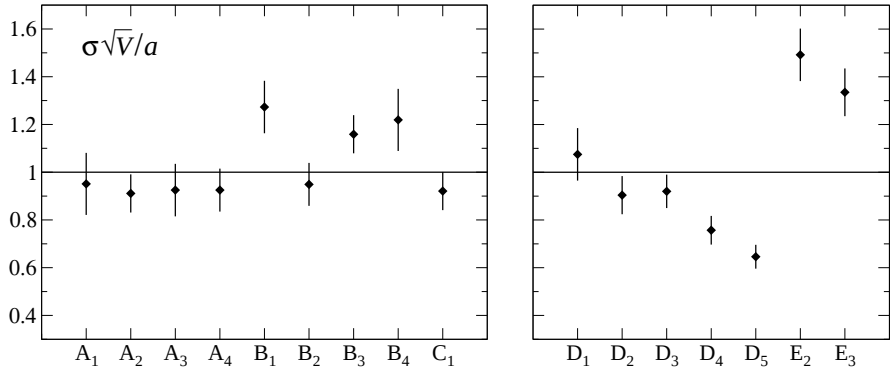

Fig. 3. Width σ of the gap distributions, scaled by the factor $\sqrt{V}/a$, as obtained in the unimproved (left) and the improved theory (right). The statistical errors were determined using the bootstrap method.

the E lattices, but values of $M_\pi L$ as low as 3.4 can still be safely reached, i.e. also in this case, the stability bound is not too restrictive.

3.3 Molecular-dynamics instabilities

Similar to the standard HMC algorithm, the DD-HMC algorithm obtains the next field configuration by integrating the associated molecular-dynamics equations. The numerical integration of these equations is well known to be potentially unstable. If an instability occurs, the energy deficit ΔH at the end of the integration can be large and the new field configuration is then normally rejected. The efficiency of the simulation may thus be affected, but we wish to emphasize that large energy deficits do not invalidate the algorithm unless the reversibility of the molecular-dynamics integration is compromised.

Earlier studies of the phenomenon suggest that the instabilities are caused by exceptionally small eigenvalues of the lattice Dirac operator [24–26]. Even if the gap distribution is safely separated from zero, it is possible that the Dirac operator develops such eigenvalues somewhere along the molecular-dynamics trajectories. The probability for this depends on how accurately the molecular-dynamics equations are solved, i.e. on the integration step sizes and the solver residues.

In our simulations, the probability for $|\Delta H|$ to be larger than 2 was always fairly small and often equal to zero (runs $B_1 - B_4$, for example). The worst cases in the unimproved and the improved theory were the runs A_4 and D_5 respectively, where the threshold of 2 was passed by 1.4% and 0.7% of the trajectories. Energy deficits $|\Delta H|$ larger than 10^3 were never seen, but values above 100 did occur, although very rarely so.

4. Autocorrelation times

The dynamical properties of the simulation algorithms used in lattice QCD are still largely unknown. It is not clear, for example, whether there are several relevant time scales and how they depend on the lattice parameters and the chosen algorithm. We shall not attempt to answer these difficult questions here, however, and merely give an account of our empirical studies of the autocorrelations in the runs listed in table 1.

4.1 Determination of autocorrelation times

Following the standard conventions, we define the integrated autocorrelation time $\tau_{\rm int}$ of an infinite series $a_1, a_2, a_3, \ldots$ of measured values of an observable A through

$$\tau_{\rm int} = \frac{1}{2} + \sum_{t=1}^{\infty} \frac{\Gamma(t)}{\Gamma(0)}, \tag{4.1}$$

where $\Gamma(t)$ denotes the autocorrelation function of the series. In practice only a finite number N of measurements can be made and the estimation of the autocorrelation time from the available data then requires some ad hoc choices to be made.

For the autocorrelation function we use the approximation

$$\Gamma(t) \simeq \frac{1}{N-t} \sum_{i=1}^{N-t} (a_i - \bar{a}_-)(a_{i+t} - \bar{a}_+), \qquad 0 \le t < N, \tag{4.2}$$

in which $\bar{a}_-$ and $\bar{a}_+$ are, respectively, the averages of the first $N-t$ and the last $N-t$ elements of the series $a_1, \ldots, a_N$. The sum in eq. (4.1) is then truncated at some value $W \ll N$ of the time lag t, referred to as the summation window, which should ideally be such that the remainder of the sum can be safely neglected.

If the autocorrelation function is well behaved, as in the case shown in the upper plot of fig. 4, the choice of the summation window is not critical and any reasonable prescription will do. The rule adopted here is to stop the summation in eq. (4.1) at the first value of t where the normalized autocorrelation function is equal to zero within two times its statistical error, the latter being estimated using the Madras-Sokal approximation (see appendix E of ref. [7]).

In practice the calculated autocorrelation functions may have long tails and they may also vary significantly with the selected range of the data series. An example illustrating this behaviour is shown in the lower plot in fig. 4. In all these cases, we

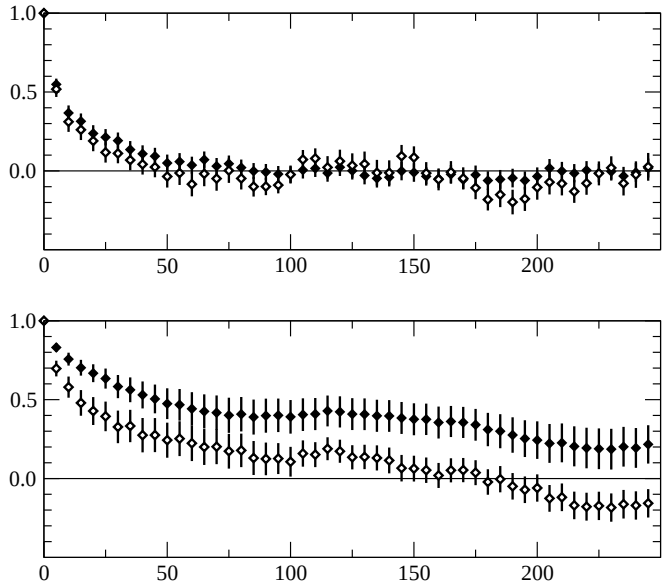

Fig. 4. Normalized autocorrelation functions $\Gamma(t)/\Gamma(0)$, plotted versus the time lag t given in numbers of trajectories, of the plaquette P (upper plot) and the solver iteration number N_{GCR} (lower plot). The data shown were calculated using the last 4000 trajectories of run B_2 (full points) or only the first 2000 of these (open points).

divide the data series into large bins, calculate the bin averages and estimate the statistical variance σ^2 of the total average assuming these are statistically independent. The integrated autocorrelation time is then given by

$$\tau_{\mathrm{int}} = \frac{\sigma^2}{2\sigma_0^2}, \tag{4.3}$$

where σ_0 denotes the naive statistical error. Evidently, the results obtained in this way are rough estimates that could easily be wrong by factor 2 or so.

4.2 Reference autocorrelation times

The integrated autocorrelation times of the Wilson plaquette P and the GCR solver iteration number N_{GCR} are listed in table 4. These two quantities are unphysical, but they are readily accessible and are useful reference cases that probe the dynamics of the simulation at both short and long distances.

In order to facilitate the comparison of the figures quoted in the table, the autocorrelation times were determined using data series of a fixed length equal to 4000 trajectories. The autocorrelation times are given in numbers of trajectories and er-

Table 4. Autocorrelation times of the plaquette P and the solver iteration number N_{GCR}

Run	$\tau_{\mathrm{int}}[P]$	$\tau_{\mathrm{int}}[N_{\mathrm{GCR}}]$	Run	$\tau_{\mathrm{int}}[P]$	$\tau_{\mathrm{int}}[N_{\mathrm{GCR}}]$
A_{1a}	25(5)	43*	C_1	17(3)	35(7)
A_{1b}	29(6)	38*	D_1	11(1)	10(2)
A_2	23(4)	46*	D_2	17(3)	21(4)
A_{3a}	14(2)	53(10)	D_3	16(2)	19(3)
A_{3b}	28*	53*	D_4	16(2)	15(2)
A_4	19(4)	45*	D_5	32(6)	24(5)
B_1	14(2)	50*	E_1	33*	14(3)
B_2	12(2)	39*	E_2	19(3)	11(2)
B_3	9(1)	45*	E_3	27(5)	25(5)
B_4	14(2)	51*			

* Estimate based on data binning

ror estimates are quoted only in those cases where the autocorrelation function was well behaved. In these regular situations, the binning method always gave consistent results.

In all simulations of the improved theory, except for run E_1 perhaps, the autocorrelation times were safely determined and turned out to be reasonably small. This was not so in the simulations of the unimproved theory, where the autocorrelation function of the GCR iteration number typically had a tail similar to the one shown in the lower plot in fig. 4. O(a) improvement thus appears to have the side-effect of reducing the autocorrelation times.

The regularity of run C_1 then remains unexplained, however, and the differences in the autocorrelation times could actually also very well be related to the fact that the physical volumes of the $C_1 - E_3$ lattices are larger, by a factor of two or more, than the volumes of the other lattices. Presumably the size of the blocks, on which the DD-HMC algorithm operates, matters as well, although the comparison of the runs A_{1a} and A_{1b} does not suggest this to be so.

4.3 Autocorrelations of physical quantities

The meson masses and all other physical quantities were calculated after finishing the simulations, using the generated ensembles of saved gauge-field configurations (see table 1). A fairly large number of trajectories was skipped between successive

saved configurations so that the statistical correlations in these sets of fields can be expected to be small.

In order to find out whether the residual correlations are relevant for the determination of the statistical errors, the basic two-point correlation functions were averaged over small bins of successive configurations. The physical quantities were then extracted from the binned data and their statistical errors were estimated using the jackknife method (appendix A). If there were significant statistical correlations in the data, the errors would increase with the bin size, but this was not the case and we therefore concluded that it was safe to proceed without data binning.

5. Computation of meson masses and decay constants

The masses and matrix elements tabulated in appendix C were calculated using a combination of methods, most of which being entirely standard by now. We consider two valence quarks, labelled r and s, and study the vector and pseudo-scalar mesons in the $\bar{r}s$-channel. The masses of the valence quarks may be set to the sea-quark mass, but we are also interested in the partially quenched situation where one of the quark masses is different from the sea-quark mass.

5.1 Two-point correlation functions

The pseudo-scalar density, the axial current and the vector current in the $\bar{r}s$-channel are given by

$$P^{rs} = \bar{r}\gamma_5 s, \qquad A_\mu^{rs} = \bar{r}\gamma_\mu\gamma_5 s, \qquad V_\mu^{rs} = \bar{r}\gamma_\mu s. \tag{5.1}$$

All masses and decay constants we are interested in were extracted from the two-point functions

$$f_{\rm PP}(x_0) = a^3 \sum_{x_1,x_2,x_3} \langle P^{rs}(x) P^{sr}(0)\rangle \,, \tag{5.2}$$

$$f_{\rm AP}(x_0) = a^3 \sum_{x_1,x_2,x_3} \langle A_0^{rs}(x) P^{sr}(0)\rangle \,, \tag{5.3}$$

$$f_{\rm VV}(x_0) = a^3 \sum_{x_1,x_2,x_3} \sum_{k=1}^{3} \langle W_k^{rs}(x) W_k^{sr}(0)\rangle \,, \tag{5.4}$$

where W_μ^{rs} is a linear combination of the vector current V_μ^{rs} and a Jacobi smeared form of it [27], slightly tuned so as to suppress the high-energy intermediate states in the two-point function.

The correlation functions were evaluated in the standard manner by first expressing them as an expectation value of a product of two quark propagators. These were calculated by solving the lattice Dirac equation, using the Schwarz-preconditioned GCR solver [6] and requiring the normalized residue of the solution to be less than 10^{-10}. In order to reduce the statistical fluctuations, the results were averaged over time-reflections and 5 distant source points in the case of the A and B runs and over 3 source points in the case of the D runs.

5.2 Masses and matrix elements

On a lattice of infinite time-like extent, and at large times x_0, the correlation function $f_{\mathrm{PP}}(x_0)$ is saturated by the one-particle pseudo-scalar meson state in the $\bar{r}s$-channel. If we denote the mass of the meson by M_{PS} and the associated vacuum-to-meson matrix element by G_{PS}, the asymptotic form of the correlation function is

$$f_{\mathrm{PP}}(x_0) = -\frac{G_{\mathrm{PS}}^2}{M_{\mathrm{PS}}} \mathrm{e}^{-M_{\mathrm{PS}} x_0} + \ldots , \tag{5.5}$$

where the ellipsis stands for a series of more rapidly decaying terms. The mass M_{V} of the $\bar{r}s$ vector meson may be defined similarly through the asymptotic behaviour of the vector correlation function $f_{\mathrm{VV}}(x_0)$, but the definition requires further explanation if the meson is unstable in infinite volume (see subsect. 5.6).

Next we note that the ratio

$$m_{\mathrm{eff}}(x_0) = \left\{ \tfrac{1}{2} \left(\partial_0 + \partial_0^*\right) f_{\mathrm{AP}}(x_0) + c_{\mathrm{A}} a \partial_0^* \partial_0 f_{\mathrm{PP}}(x_0) \right\} \,/\, f_{\mathrm{PP}}(x_0) \tag{5.6}$$

converges to a constant m_{rs} at large times x_0, for any fixed value of the parameter c_{A}, because both $f_{\mathrm{AP}}(x_0)$ and $f_{\mathrm{PP}}(x_0)$ are proportional to $\mathrm{e}^{-M_{\mathrm{PS}} x_0}$ in this limit. Moreover, in the continuum limit, $m_{\mathrm{eff}}(x_0)$ is expected to converge to the sum of the bare current-quark masses of the r and the s quark, at all times x_0, with a rate proportional to a in the unimproved theory (where we set c_{A} to zero) or a^2 if the improvement coefficients c_{sw} and c_{A} are properly tuned [17,28,29] †.

All our numerical data for $m_{\mathrm{eff}}(x_0)$ in fact turned out to be statistically consistent with a constant value, over a large range of x_0, and the quark mass sum m_{rs}

† The effects of the $1 + \mathrm{O}(am)$ renormalization factors (C.2) are expected to be small in practice and are neglected here for simplicity.

was therefore always unambiguously and accurately determined. In particular, the current-quark mass $m_{\rm sea} = \frac{1}{2}m_{rr}$ of the sea quarks is obtained by setting the hopping parameters of the valence quarks to $\kappa_{\rm sea}$. Whether in general m_{rs} coincides with $\frac{1}{2}(m_{rr} + m_{ss})$, as one expects to be the case if the lattice effects are small, is a question to which we shall return in sect. 6.

The bare pseudo-scalar decay constant $F_{\rm PS}$ in the $\bar{r}s$-channel is normally extracted from the asymptotic behaviour of the two-point functions $f_{\rm AP}(x_0)$ and $f_{\rm PP}(x_0)$. In this paper, however, we first computed m_{rs}, $M_{\rm PS}$ and $G_{\rm PS}$ and then used the formula

$$F_{\rm PS} = \frac{m_{rs}}{M_{\rm PS}^2} G_{\rm PS} \tag{5.7}$$

for the decay constant. Starting from eq. (5.6), it is straightforward to show that equivalent results are obtained in this way, up to small corrections of $\mathrm{O}(a^2)$. Note that $F_{\rm PS}$ is automatically $\mathrm{O}(a)$-improved if m_{rs} is.

5.3 Spectral decomposition in finite volume

On a finite lattice with time-like extent T, the calculation of the pseudo-scalar and vector meson masses requires some care and must address the issue of higher-states contributions. This is, incidentally, not so in the case of the quark mass sum m_{rs}, which is expected to be independent of the lattice size up to lattice-spacing effects.

For $0 < x_0 < T$, the correlation function $f_{\rm PP}(x_0)$ (and similarly $f_{\rm VV}(x_0)$) can be expanded in a rapidly convergent series of the form

$$f_{\rm PP}(x_0) = -\sum_{i=0}^{\infty}\sum_{j=i}^{\infty} c_{ij} h(x_0; E_i, E_j), \tag{5.8}$$

$$h(t; E, E') = \exp\{-Et - E'(T-t)\} + \exp\{-E't - E(T-t)\}, \tag{5.9}$$

where $0 = E_0 < E_1 < E_2 < \ldots$ are the intermediate-state energies and $c_{ij} \geq 0$ the associated spectral weights ‡. In the channel considered here, the lowest intermediate state is the $\bar{r}s$ pseudo-scalar meson state at zero spatial momentum. Then come the multi-meson scattering states and more and more complicated states as one moves up the energy scale.

‡ Equation (5.8) assumes the existence of a positive hermitian transfer matrix which may not be guaranteed in the improved theory. It seems likely to us, however, that a transfer matrix can still be defined, as is the case in $\mathrm{O}(a^2)$-improved gauge theories [30], although complex energy values and negative weights may occur at energies on the order of the cutoff scale $1/a$.

At large x_0 and T, the dominant term in the series (5.8) is thus the one where $E_i = 0$ and $E_j = M_{\rm PS}$. Moreover, using the product inequality (B.3), the contributions of all higher-energy states can be shown to be exponentially suppressed with respect to this term. In practice their effects are seen in the simulation data only when either x_0 or $T - x_0$ is not too large. The leading terms in this range are then

$$f_{\rm PP}(x_0) = c_0 h(x_0; 0, M_0) + c_1 h(x_0; 0, M_1) + \ldots, \qquad M_0 = M_{\rm PS}, \tag{5.10}$$

where M_1 denotes the energy of the next-to-lowest state in the $\bar{r}s$-channel (if the spatial volume of the lattice is large enough, this will be a three-meson state with all particles at rest).

Note that each term in the spectral series (5.8) decreases exponentially in the range $0 \le x_0 \ll \frac{1}{2}T$, with an exponent equal to $E_j - E_i$ that can be as small as the pseudo-scalar meson mass, for example, even if both E_i and E_j are not small. The presence of such contributions complicates the analysis of the correlation functions considerably unless the time-like extent T of the lattice is sufficiently large to strongly suppress them. This condition was barely satisfied in the case of the run A_4, which is why we decided to discard it from the physics analysis (as already mentioned in the first paper in this series).

5.4 Effective masses and matrix elements

Slightly departing from what is usually done, we define the effective pseudo-scalar meson mass $M_{\rm eff}(x_0)$ in the $\bar{r}s$-channel to be the value of $M \ge 0$ where

$$\frac{h(x_0 - a; 0, M)}{h(x_0; 0, M)} = \frac{f_{\rm PP}(x_0 - a)}{f_{\rm PP}(x_0)}. \tag{5.11}$$

Using the results obtained in appendix B, it is not difficult to prove that this equation has one and only one solution. Moreover, with this definition of the effective mass it is guaranteed that $M_{\rm eff}(x_0) = M_{\rm PS}$ at large x_0, up to exponentially small terms. We then also introduce the effective matrix element

$$G_{\rm eff}(x_0) = \left\{ -M_{\rm eff}(x_0) \frac{f_{\rm PP}(x_0)}{h(x_0; 0, M_{\rm eff}(x_0))} \right\}^{1/2}, \tag{5.12}$$

which converges to $G_{\rm PS}$ in the large-time limit.

The asymptotic behaviour of the effective mass at large x_0 and T can be worked

out explicitly, starting from the spectral representation (5.10). Setting

$$\epsilon(x_0) = \frac{c_1 h(x_0; 0, M_1)}{c_0 h(x_0; 0, M_0)}, \qquad \delta(x_0) = \left\{ M \frac{\partial}{\partial M} \ln h(x_0; 0, M) \right\}_{M=M_0}, \tag{5.13}$$

and going through a few lines of algebra, it is straightforward to derive the expansion

$$M_{\mathrm{eff}}(x_0) = M_{\mathrm{PS}} \left\{ 1 + \frac{\epsilon(x_0) - \epsilon(x_0 - a)}{\delta(x_0) - \delta(x_0 - a)} + \ldots \right\}, \tag{5.14}$$

where the ellipsis stands for terms that are exponentially small with respect to the next-to-leading term. A similar formula,

$$G_{\mathrm{eff}}(x_0) = G_{\mathrm{PS}} \left\{ 1 + \tfrac{1}{2}\epsilon(x_0) + \tfrac{1}{2}\left(1 - \delta(x_0)\right) \frac{\epsilon(x_0) - \epsilon(x_0 - a)}{\delta(x_0) - \delta(x_0 - a)} + \ldots \right\}, \tag{5.15}$$

is obtained in the case of the effective matrix element.

5.5 Fit procedures

From the point of view of the statistical error analysis, the correlation functions f_{PP}, f_{AP} and f_{VV} are the primary quantities, while the effective quark mass sums, meson masses and matrix elements are functions of these. The statistical errors of all these quantities tend to be strongly correlated. We took the correlations fully into account, from the primary quantities to the final results, by propagating the errors using the jackknife method (appendix A). In particular, fitted and interpolated values were always considered to be functions of the input data, which allows their errors to be calculated in the standard manner.

The quark mass sum, the pseudo-scalar meson masses and matrix elements, and the masses of the vector mesons were all obtained by fitting the corresponding effective quantity $P_{\mathrm{eff}}(x_0)$ in a range $t_0 \leq x_0 \leq t_1$ of time with the chosen fit function $\Phi(x_0)$. We performed correlated least-squares fits, where the values of the fit parameters were determined by minimizing

$$\chi^2 = \sum_{x_0, y_0 = t_0}^{t_1} \left[P_{\mathrm{eff}}(x_0) - \Phi(x_0) \right] (C^{-1})_{x_0 y_0} \left[P_{\mathrm{eff}}(y_0) - \Phi(y_0) \right], \tag{5.16}$$

the matrix C being the statistical error covariance of $P_{\mathrm{eff}}(t_0), \ldots, P_{\mathrm{eff}}(t_1)$. The quark mass sum m_{rs}, for example, was computed by fitting $m_{\mathrm{eff}}(x_0)$ to a constant as shown in fig. 5a.

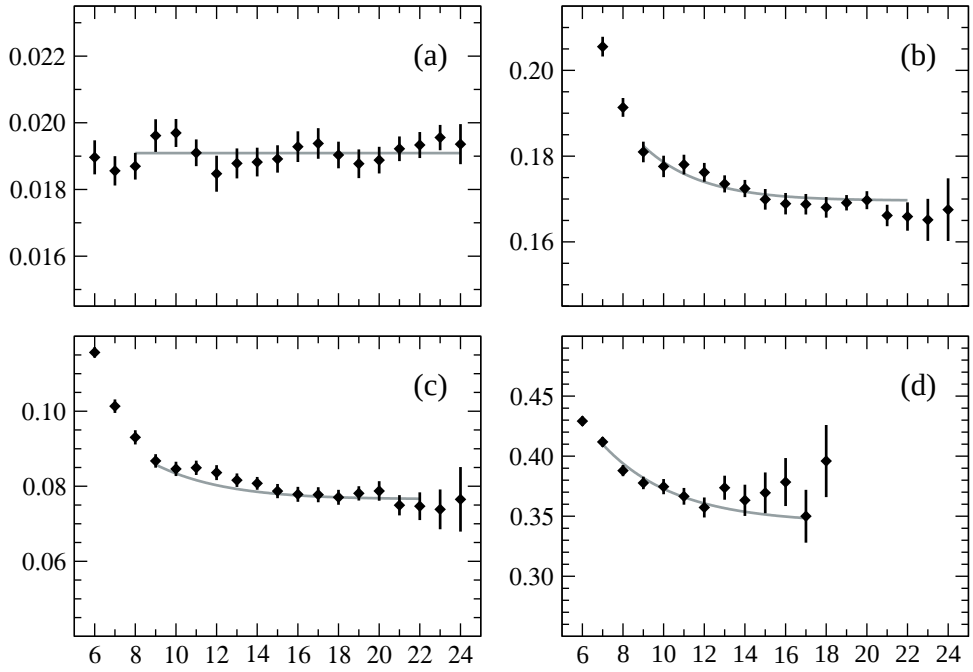

Fig. 5. Sample plots illustrating the dependence on x_0/a of the effective quark mass sum (figure a), the pseudo-scalar mass and matrix element (figures b and c) and the vector meson mass (figure d), all given in lattice units. The data points shown are from run D_4 and the valence quark masses were both set to the sea-quark mass in this example. The solid lines are the fits discussed in the text.

In the case of the pseudo-scalar meson masses, we fitted the data with the asymptotic expression (5.14). We first calculated the mass M_π of the pions, i.e. the mesons made of the sea-quarks, by substituting $M_1 = 3M_\pi$ for the energy of the next-higher state (thus assuming the latter is a three-pion state with small interaction energy) and adjusting M_π and c_1/c_0 so as to minimize χ^2. While the fit curves obtained in this way represent the data very well, it should be noted that the fitted value of M_π is largely determined by the data at large times x_0, where a fit to a constant would give nearly the same results (see fig. 5b).

Once M_π was determined, the mesons made of a sea quark and a valence quark with a mass different from the sea quark were considered. Here we set $M_1 = M_{\rm PS} + 2M_\pi$ and otherwise proceeded as in the degenerate case. Next the matrix elements $G_{\rm PS}$ were computed by fitting the data with the asymptotic expression (5.15), using the same values of M_1 as in the fits of the effective meson masses (fig. 5c). We did not set c_1/c_0 to the previously computed values, but it turned out that the two fits

gave consistent results for this parameter.

5.6 Energy spectrum in the vector channel

At small sea-quark masses, the vector mesons become resonances that decay into two (or more) pseudo-scalar mesons. As was shown long ago [31], resonances give rise to a characteristic volume-dependent pattern of the energy spectrum which allows their masses and decay widths to be determined, in principle, from simulation data.

As before, we considered the channels where one or both of the r and s quarks is a sea quark. Starting from the correlation functions $f_{\mathrm{VV}}(x_0)$, the lowest energy M_{V} in this channel was calculated by fitting the effective mass with the asymptotic formula (5.14) (with M_{PS} replaced by M_{V}). For the lowest excited-state energy we substituted

$$M_1 = (M_{\mathrm{PS}}^2 + k^2)^{1/2} + (M_\pi^2 + k^2)^{1/2}, \qquad k = 2\pi/L, \tag{5.17}$$

in this case, L being the spatial size of the lattice. Excellent fits were obtained with this ansatz and M_{V} was determined quite accurately on all lattices.

We refer to the energy values M_{V} as the vector meson masses in this paper, even in those cases where the meson is likely to become a resonance in the infinite volume limit (we estimate this to be so at the lightest quark masses in each series of lattices and perhaps at some of the second-to-lightest as well). This use of language is only slightly incorrect, however, because in all our simulations M_{V} turned out to be at most 20% larger than $M_{\mathrm{PS}} + M_\pi$ and significantly smaller than M_1, in which case the true resonance energy is expected to be close to M_{V} [31].

We finally note that the statistical errors in the vector channel tend to be larger than those in the pseudo-scalar channel. The effect could be related to the resonance character of the vector mesons and it is conceivable that a coupled channel analysis, such as the one recently presented by Aoki et al. [32], will not only allow the vector meson decays to be studied but may also help to reduce the statistical errors.

6. Quark-mass dependence in partially quenched QCD

The most important physical results of our simulations were already presented in our first paper in this series [9]. We now discuss the dependence of the quantities tabulated in appendix C on the quark masses in some further detail, focusing on the empirical facts rather than on their possible theoretical interpretation.

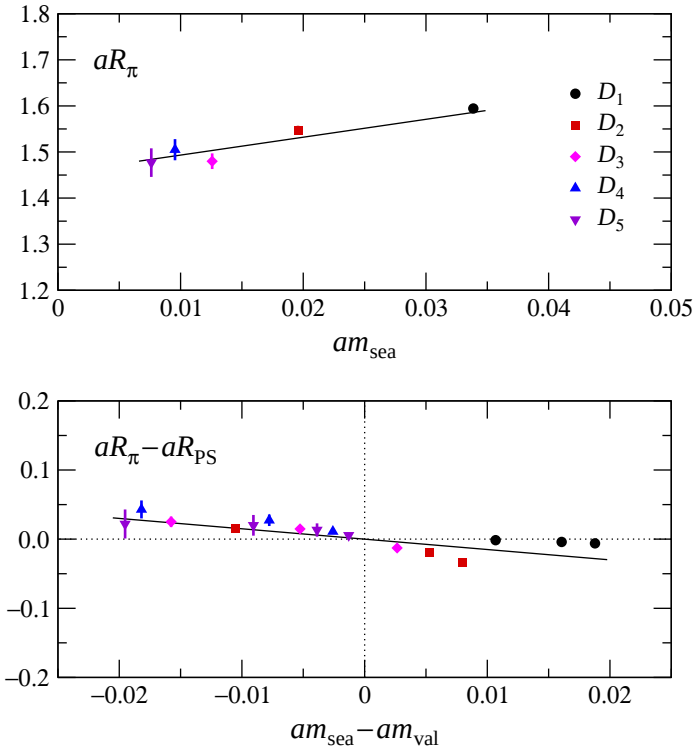

Fig. 6. Results for the ratio R_π and the difference $R_\pi - R_{\rm PS}$ obtained on the D-series of lattices. The solid lines represent the global linear fit (6.2). Note that the points in the lower plot do not have to line up within errors, since $R_\pi - R_{\rm PS}$ is a function of two independent variables rather than of $m_{\rm sea} - m_{\rm val}$ alone.

As before we set $m_{\rm sea} = \frac{1}{2}m_{rr}$ if the r quark is a sea quark and we now also set $m_{\rm val} = \frac{1}{2}m_{ss}$ if the s quark is a valence quark. The figures in the tables are all for the mixed case, where one quark is a sea quark and the other a valence quark. We are thus considering partially quenched QCD with $2 + 1$ flavours of quarks.

6.1 Quark and pseudo-scalar meson masses

We first remark that the quark mass sum m_{rs} turns out to be equal to $m_{\rm sea} + m_{\rm val}$ within statistical errors, on all lattices and for all quark-mass combinations. The ratio $m_{rs}/(m_{\rm sea} + m_{\rm val})$ is obtained with better statistical precision than the quark masses, but the largest deviation seen in this case is only 0.6%. The additivity of the current quark masses (which is an exact property of the theory in the continuum limit) is thus accurately guaranteed on the lattices that we have simulated.

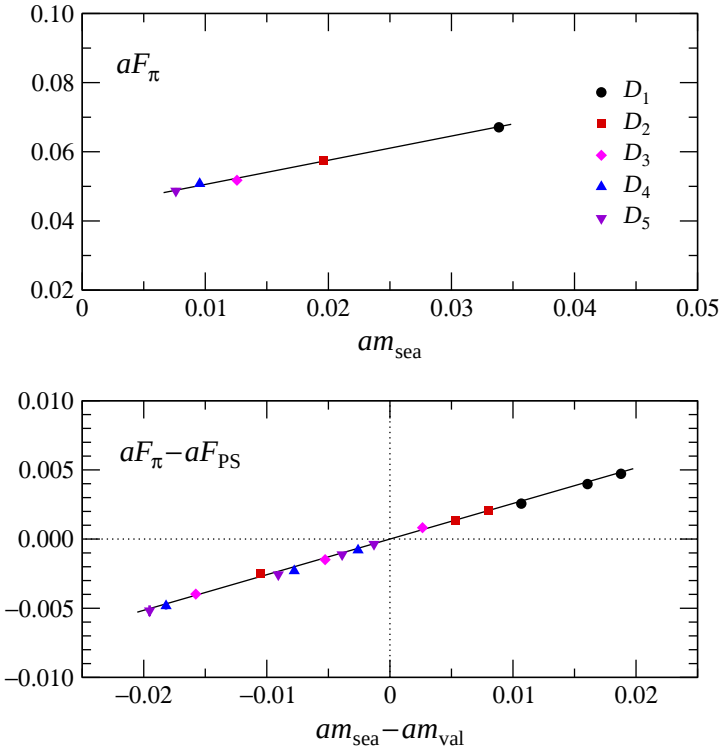

Fig. 7. Dependence of the bare pion decay constant $F_\pi = F_{\rm PS}|_{m_{\rm val}=m_{\rm sea}}$ and of the difference $F_\pi - F_{\rm PS}$ on the quark masses, as determined on the D-series of lattices. The solid lines represent the global linear fit (6.3).

Next we consider the ratios

$$R_{\rm PS} = \frac{M_{\rm PS}^2}{m_{\rm sea} + m_{\rm val}}, \qquad R_\pi = R_{\rm PS}|_{m_{\rm val}=m_{\rm sea}} = \frac{M_\pi^2}{2m_{\rm sea}}, \tag{6.1}$$

which are independent of the quark masses to lowest order of chiral perturbation theory. However, this is not so at next-to-leading order and the numerically calculated ratios are in fact weakly mass-dependent (see fig. 6). An empirical fit

$$R_{\rm PS} = a_0 + a_1(m_{\rm sea} + m_{\rm val}) + a_2 m_{\rm sea} \tag{6.2}$$

represents the data quite well in the given range of masses except perhaps for the points where $m_{\rm val} \ll m_{\rm sea}$. In the case of the D-series of lattices, for example, the data for $R_{\rm PS}$ deviate from the fit by no more than 2% and most points are within a margin of 1%.

6.2 Pseudo-scalar decay constant and vector meson mass

As can be seen from the tables in appendix C, the calculated values of $F_{\rm PS}/M_{\rm V}$ are nearly independent of the quark masses. This comes a bit as a surprise, and could merely be an accidental agreement in a limited range of masses, since there does not appear to be any obvious physical connection between the pseudo-scalar decay constant and the vector meson mass.

The mass dependence of these two quantities is thus practically the same and it suffices to consider one of them. Focusing on the decay constant, a simple linear expression,

$$F_{\rm PS} = b_0 + b_1(m_{\rm sea} + m_{\rm val}) + b_2 m_{\rm sea}, \tag{6.3}$$

turns out to fit the available data for $F_{\rm PS}$ very well. On the D-series of lattices, for example, the fit matches the data within statistical errors and the maximal relative deviation in the given range of masses is only 1.6% (see fig. 7).

It is tempting to use these fits to extrapolate the decay constant to the chiral limit, but as already emphasized in our previous paper [9], such extrapolations are difficult to justify and asymptotically inconsistent with one-loop chiral perturbation theory. On the other hand, the observed linearity of the pseudo-scalar decay constant in the range of masses covered by the simulations is striking and calls for a theoretical explanation.

7. Concluding remarks

Numerical lattice QCD is currently in an interesting transition phase. The valence approximation is now practically overcome, but important physical effects of the light sea quarks, such as the decay of the rho meson or the anomaly-driven mass splitting between the eta and the pions, still have not or only barely been studied directly. Simulations at smaller quark masses and on larger lattices than reported here will probably be required for this. Our experience however suggests that the prospects for such simulations, using O(a)-improved Wilson quarks, are now quite good.

So far the DD-HMC algorithm performed well and we did not run into any instabilities or other technical difficulties. As one moves to smaller quark masses and smaller lattice spacings, there may be some room for further algorithmic improvements, but

the development of variance-reduction methods is likely to be more rewarding at this point, particularly so if disconnected quark-line diagrams and multi-particle amplitudes are to be computed.

The numerical simulations were performed on PC clusters at CERN, the Centro Enrico Fermi, the Institut für Theoretische Physik der Universität Bern (with a contribution from the Schweizerischer Nationalfonds) and on a CRAY XT3 at the Swiss National Supercomputing Centre (CSCS). We are grateful to all these institutions for the continuous support given to this project.

Appendix A. Statistical error analysis

In the physics analysis of the runs $A_1 - A_3$, $B_1 - B_4$ and $D_1 - D_5$, we kept track of the statistical errors using the jackknife method. In particular, any correlations among the errors of different observables were always properly taken into account. Here we summarize our conventions and briefly explain the basic procedures that we used.

A.1 Jackknife samples

Let A_r, $r = 1, \ldots, R$, be a set of primary stochastic observables and $a_{r,1}, \ldots, a_{r,N}$ a sequence of N measured values of these. In lattice QCD the most common primary observables are the Wilson loops and sums of products of quark propagators. The jackknife method assumes that the measured values are unbiased and statistically independent. We shall thus take it for granted that the residual autocorrelations are negligible in the cases of interest (see sect. 4).

The averages $\bar{a}_r$ of the observables A_r and the associated statistical error covariance C_{rs} are given by

$$\bar{a}_r = \frac{1}{N} \sum_{i=1}^{N} a_{r,i}, \tag{A.1}$$

$$C_{rs} = \frac{1}{N(N-1)} \sum_{i=1}^{N} (a_{r,i} - \bar{a}_r)\,(a_{s,i} - \bar{a}_s)\,. \tag{A.2}$$

If we introduce the jackknife samples

$$a^J_{r,i} = \bar{a}_r + c_N\,(\bar{a}_r - a_{r,i})\,, \qquad c_N = (N(N-1))^{-1/2}\,, \tag{A.3}$$

an equivalent expression for the error matrix is

$$C_{rs} = \sum_{i=1}^{N} \left(a_{r,i}^J - \bar{a}_r\right)\left(a_{s,i}^J - \bar{a}_s\right). \tag{A.4}$$

Note that our definition of the jackknife samples slightly departs from the standard conventions, where $c_N = 1/(N-1)$. The modification is numerically insignificant in practice, but leads to some simplifications when data from different simulations are to be combined (see subsect. A.3).

A.2 Error propagation

Apart from estimating the primary observables, one may be interested in evaluating various functions $f(A_1, \ldots, A_R)$ of them, which may involve fit procedures and other complicated operations. The standard stochastic estimate of such an observable is

$$\bar{f} = f(\bar{a}_1, \ldots, \bar{a}_R) \tag{A.5}$$

and the associated series of jackknife estimates is defined by

$$f_i^J = f(\bar{a}_{1,i}^J, \ldots, \bar{a}_{R,i}^J), \qquad i = 1, \ldots, N. \tag{A.6}$$

A little algebra then shows that the expression

$$\sigma^2 = \sum_{i=1}^{N} \left(f_i^J - \bar{f}\right)^2 \tag{A.7}$$

provides an estimate of the statistical variance of $\bar{f}$, which coincides with the usual error propagation formula (the one that involves the gradient of f) up to terms of order $1/N$. Similarly the error covariance of f and any other function g is obtained by summing $(f_i^J - \bar{f})(g_i^J - \bar{g})$ over the jackknife samples.

In practice the error formula (A.7) proves to be very convenient. If an observable is a function of previously calculated observables, for example, one can take advantage of the fact that the composition of functions is associative, i.e. the jackknife series f_i^J is simply obtained by inserting the jackknife series of the arguments, independently of whether these are primary or not. The data analysis can thus proceed in steps, starting from the primary observables and progressing to more and more complicated observables.

A.3 Combining data from different runs

Simulations of lattice QCD at different sea-quark masses, lattice spacings, etc., can be assumed to be statistically independent. The statistical variance of any observable that depends on data from several simulations is therefore the sum of the associated partial variances. This rule can easily be accommodated in the jackknife analysis by embedding the jackknife series of the observables in extended series that include all simulations on which the observable depends.

The method is best explained by considering two simulations, where N_1 measurements of some observables A_r are made in the first and N_2 measurements of some other observables B_s in the second. The associated jackknife series $a^J_{r,1}, \ldots, a^J_{r,N_1}$ and $b^J_{s,1}, \ldots, b^J_{s,N_2}$ are then computed as before, starting from the primary observables in each simulation. Next they are embedded in extended series

$$a^J_{r,1}, \ldots, a^J_{r,N_1}, \underbrace{\bar{a}_r, \ldots, \bar{a}_r}_{N_2 \text{ elements}} \qquad \text{and} \qquad \underbrace{\bar{b}_s, \ldots, \bar{b}_s}_{N_1 \text{ elements}}\, b^J_{s,1}, \ldots, b^J_{s,N_2} \tag{A.8}$$

of length $N_1 + N_2$ such that the first N_1 elements are occupied by the jackknife series from the first simulation and the last N_2 elements by those from the second simulation.

With this assignment, and if the extended series are treated as ordinary jackknife series, the correct error correlation matrix of the full set $A_1, \ldots, A_R, B_1, \ldots, B_S$ of observables is obtained. Moreover, we may define the jackknife series of any observable $f(A_1, \ldots A_R, B_1, \ldots, B_S)$ in the standard manner and compute its variance using eq. (A.7). The embedding trick thus allows the statistical errors to be propagated as if there were a single simulation.

Appendix B. Properties of the auxiliary function $h(t; E, E')$

The symmetry properties

$$h(t; E, E') = h(T - t; E, E') = h(t; E', E) \tag{B.1}$$

are an immediate consequence of the definition (5.9) of the function $h(t; E, E')$. It is also straightforward to verify that

$$h(t; E, E') = 2\mathrm{e}^{-\frac{1}{2}(E+E')T} \cosh\left(\tfrac{1}{2}(E' - E)(T - 2t)\right), \tag{B.2}$$

and $h(t; E, E')$ is thus a convex function of t which attains its minimum at $t = \frac{1}{2}T$.

B.1 Product inequality

We now show that the inequality

$$h(t; E, E' + E'') \leq h(t; E, E')h(t; 0, E'') \tag{B.3}$$

holds for all values of the arguments t, E, E' and E''. To this end, first note that

$$\cosh(\alpha + \beta) \leq \cosh(\alpha + \beta) + \cosh(\alpha - \beta) = 2\cosh\alpha\cosh\beta. \tag{B.4}$$

Substituting $\alpha = \frac{1}{2}(E' - E)(T - 2t)$ and $\beta = \frac{1}{2}E''(T - 2t)$, this inequality becomes

$$\cosh\left(\tfrac{1}{2}(E' + E'' - E)(T - 2t)\right) \leq$$
$$2\cosh\left(\tfrac{1}{2}(E' - E)(T - 2t)\right)\cosh\left(\tfrac{1}{2}E''(T - 2t)\right), \tag{B.5}$$

which is easily seen to coincide with (B.3) after inserting the representation (B.2).

B.2 Monotonicity property

If t and s are in the range $s < t \leq \frac{1}{2}T$, and if $M > 0$, it follows from eq. (B.2) that the ratio

$$r = \frac{h(s; 0, M)}{h(t; 0, M)} \tag{B.6}$$

is greater than 1. A less obvious statement is that the ratio increases monotonically from $r = 1$ to $r = \infty$ when M goes from zero to infinity.

In order to show this, we insert eq. (B.2) and work out the quotient

$$q = \frac{r - 1}{r + 1} = \tanh\left(\tfrac{1}{2}M(t - s)\right)\tanh\left(\tfrac{1}{2}M(T - t - s)\right). \tag{B.7}$$

In the specified range of t and s, the arguments of the hyperbolic functions in this equation are non-negative and monotonically growing with M. The quotient thus rises monotonically from 0 to 1 when M goes from zero to infinity, which proves our claim, since r and q are monotonically related to each other.

Appendix C. Tables of meson masses and decay constants

The simulation results tabulated in this appendix were obtained following the lines of sect. 5. In all cases, the r quark was taken to be a sea quark, i.e. the associated hopping parameter κ_r was set to $\kappa_{\rm sea}$. The hopping parameter κ_s of the other quark, on the other hand, ranged over 4 or 5 values, one of which being $\kappa_{\rm sea}$.

For each series of runs, we quote the quark mass sums m_{rs}, the pseudo-scalar meson masses $M_{\rm PS}$ and matrix elements $G_{\rm PS}$, and the vector meson masses $M_{\rm V}$, all given in lattice units (tables 5, 7 and 9). Some combinations of these quantities are printed in tables 6, 8 and 10. The errors given in brackets are statistical only.

If so desired, the quoted results can be converted to physical units by substituting the estimates 0.0717(15), 0.0521(7) and 0.0784(10) fm for the spacings of the A, B and D lattices [9]. The quark mass sums m_{rs}, the matrix elements $G_{\rm PS}$ and the decay constants $F_{\rm PS}$ then also need to be renormalized,

$$m_{rs} \to Z_{\rm A} Z_{\rm P}^{-1} m_{rs}, \qquad G_{\rm PS} \to Z_{\rm P} G_{\rm PS}, \qquad F_{\rm PS} \to Z_{\rm A} F_{\rm PS}, \tag{C.1}$$

where $Z_{\rm A}$ and $Z_{\rm P}$ denote the (mass-independent) renormalization constants of the non-singlet axial current and density. Moreover, in order to guarantee the $\mathrm{O}(a)$ improvement of these quantities in the improved theory, the renormalization constants must be modified according to

$$Z_{\rm X} \to Z_{\rm X}\big(1 + \bar{b}_{\rm X} a m_{\rm sea} + \tfrac{1}{2}\tilde{b}_{\rm X} a m_{rs}\big), \tag{C.2}$$

with properly adjusted coefficients $\bar{b}_{\rm X}$ and $\tilde{b}_{\rm X}$ [17,33] (the figures quoted in tables 9 and 10 include the contribution of the operator improvement term proportional to $c_{\rm A}$ but not the $1 + \mathrm{O}(am)$ renormalization factors).

Table 5. Results for m_{rs}, $M_{\rm PS}$, $G_{\rm PS}$ and $M_{\rm V}$ (lattices $A_1 - A_3$)

Run	κ_r	κ_s	am_{rs}	$aM_{\rm PS}$	$a^2G_{\rm PS}$	$aM_{\rm V}$
A_1	0.15750	0.15750	0.0548(5)	0.2726(19)	0.0881(12)	0.389(4)
		0.15800	0.0472(6)	0.2536(19)	0.0859(12)	0.379(5)
		0.15825	0.0434(6)	0.2438(20)	0.0848(13)	0.373(5)
		0.15835	0.0419(6)	0.2398(21)	0.0844(13)	0.371(5)
A_2	0.15800	0.15750	0.0359(3)	0.2137(18)	0.0703(12)	0.344(3)
		0.15800	0.0285(3)	0.1913(19)	0.0682(13)	0.334(4)
		0.15825	0.0249(3)	0.1790(21)	0.0671(14)	0.329(5)
		0.15835	0.0235(3)	0.1738(22)	0.0666(15)	0.328(5)
A_3	0.15825	0.15750	0.0281(4)	0.185(3)	0.0617(19)	0.327(5)
		0.15800	0.0208(4)	0.160(3)	0.0599(22)	0.317(7)
		0.15825	0.0172(4)	0.147(4)	0.0593(23)	0.312(8)
		0.15835	0.0158(4)	0.141(4)	0.0592(24)	0.311(9)

Table 6. Combinations of m_{rs}, $M_{\rm PS}$, $G_{\rm PS}$ and $M_{\rm V}$ (lattices $A_1 - A_3$)

Run	κ_r	κ_s	$aM^2_{\rm PS}/m_{rs}$	$aF_{\rm PS}$	$F_{\rm PS}/M_{\rm V}$
A_1	0.15750	0.15750	1.357(17)	0.0650(7)	0.1669(21)
		0.15800	1.363(19)	0.0630(7)	0.1664(24)
		0.15825	1.369(21)	0.0619(8)	0.166(3)
		0.15835	1.372(22)	0.0615(8)	0.166(3)
A_2	0.15800	0.15750	1.272(20)	0.0553(7)	0.161(3)
		0.15800	1.282(25)	0.0532(7)	0.159(3)
		0.15825	1.29(3)	0.0522(8)	0.159(3)
		0.15835	1.29(3)	0.0518(8)	0.158(4)
A_3	0.15825	0.15750	1.22(4)	0.0505(8)	0.154(3)
		0.15800	1.23(5)	0.0486(10)	0.153(4)
		0.15825	1.25(6)	0.0474(11)	0.152(5)
		0.15835	1.26(7)	0.0469(12)	0.151(6)

Table 7. Results for m_{rs}, $M_{\rm PS}$, $G_{\rm PS}$ and $M_{\rm V}$ (lattices $B_1 - B_4$)

Run	κ_r	κ_s	am_{rs}	$aM_{\rm PS}$	$a^2G_{\rm PS}$	$aM_{\rm V}$
B_1	0.15410	0.15410	0.03889(18)	0.1958(9)	0.0453(5)	0.2896(17)
		0.15425	0.03631(18)	0.1892(9)	0.0447(5)	0.2858(17)
		0.15440	0.03375(18)	0.1824(10)	0.0441(6)	0.2820(18)
		0.15455	0.03120(18)	0.1754(10)	0.0435(6)	0.2782(19)
B_2	0.15440	0.15410	0.02696(13)	0.1619(11)	0.0384(7)	0.2518(21)
		0.15425	0.02440(14)	0.1546(12)	0.0379(7)	0.2475(22)
		0.15440	0.02187(14)	0.1470(12)	0.0374(7)	0.2432(24)
		0.15455	0.01935(14)	0.1391(13)	0.0370(8)	0.239(3)
B_3	0.15455	0.15410	0.02185(12)	0.1416(12)	0.0333(7)	0.2418(24)
		0.15425	0.01927(12)	0.1329(13)	0.0326(7)	0.238(3)
		0.15440	0.01668(12)	0.1235(14)	0.0318(7)	0.233(3)
		0.15455	0.01409(13)	0.1132(15)	0.0310(8)	0.230(3)
B_4	0.15462	0.15410	0.02029(16)	0.1328(10)	0.0317(6)	0.237(3)
		0.15425	0.01774(16)	0.1242(11)	0.0312(6)	0.233(3)
		0.15440	0.01521(17)	0.1151(12)	0.0307(6)	0.229(3)
		0.15455	0.01269(17)	0.1055(14)	0.0302(7)	0.224(4)
		0.15462	0.01151(17)	0.1008(15)	0.0300(8)	0.223(4)

Table 8. Combinations of m_{rs}, $M_{\rm PS}$, $G_{\rm PS}$ and $M_{\rm V}$ (lattices $B_1 - B_4$)

Run	κ_r	κ_s	$aM^2_{\rm PS}/m_{rs}$	$aF_{\rm PS}$	$F_{\rm PS}/M_{\rm V}$
B_1	0.15410	0.15410	0.986(7)	0.0460(4)	0.1587(17)
		0.15425	0.986(8)	0.0453(4)	0.1586(18)
		0.15440	0.986(8)	0.0447(4)	0.1585(19)
		0.15455	0.986(9)	0.0441(5)	0.1584(20)
B_2	0.15440	0.15410	0.973(14)	0.0395(4)	0.1567(19)
		0.15425	0.979(15)	0.0387(4)	0.1562(20)
		0.15440	0.988(17)	0.0379(4)	0.1556(21)
		0.15455	1.000(19)	0.0370(4)	0.1549(23)
B_3	0.15455	0.15410	0.918(15)	0.0363(3)	0.1502(21)
		0.15425	0.916(17)	0.0356(4)	0.1497(23)
		0.15440	0.914(19)	0.0348(4)	0.149(3)
		0.15455	0.910(22)	0.0340(4)	0.148(3)
B_4	0.15462	0.15410	0.869(13)	0.0365(4)	0.154(3)
		0.15425	0.869(15)	0.0359(4)	0.154(3)
		0.15440	0.871(18)	0.0352(5)	0.154(3)
		0.15455	0.878(23)	0.0344(5)	0.153(4)
		0.15462	0.88(3)	0.0340(6)	0.153(4)

Table 9. Results for m_{rs}, $M_{\rm PS}$, $G_{\rm PS}$ and $M_{\rm V}$ (lattices $D_1 - D_5$)

Run	κ_r	κ_s	am_{rs}	$aM_{\rm PS}$	$a^2G_{\rm PS}$	$aM_{\rm V}$
D_1	0.13550	0.13550	0.06771(21)	0.3286(10)	0.1069(15)	0.464(3)
		0.13590	0.05704(22)	0.3017(10)	0.1030(15)	0.447(3)
		0.13610	0.05165(23)	0.2873(11)	0.1008(15)	0.438(3)
		0.13620	0.04893(24)	0.2799(12)	0.0998(15)	0.434(4)
D_2	0.13590	0.13550	0.04968(13)	0.2758(8)	0.0920(11)	0.4173(24)
		0.13590	0.03914(14)	0.2461(9)	0.0891(11)	0.401(3)
		0.13610	0.03383(14)	0.2301(9)	0.0880(12)	0.394(4)
		0.13620	0.03112(15)	0.2218(10)	0.0878(13)	0.390(4)
D_3	0.13610	0.13550	0.04092(14)	0.2440(10)	0.0811(12)	0.382(3)
		0.13590	0.03041(14)	0.2110(11)	0.0780(13)	0.363(4)
		0.13610	0.02514(15)	0.1929(12)	0.0766(14)	0.354(5)
		0.13620	0.02249(15)	0.1832(13)	0.0760(15)	0.349(5)
D_4	0.13620	0.13550	0.03728(14)	0.2335(11)	0.0813(11)	0.374(4)
		0.13590	0.02686(15)	0.1993(12)	0.0785(12)	0.356(4)
		0.13610	0.02168(15)	0.1800(13)	0.0771(13)	0.348(5)
		0.13620	0.01909(15)	0.1695(14)	0.0765(13)	0.345(6)
D_5	0.13625	0.13550	0.03474(13)	0.2249(11)	0.0784(13)	0.376(5)
		0.13590	0.02428(13)	0.1881(11)	0.0747(14)	0.359(6)
		0.13610	0.01910(13)	0.1672(13)	0.0729(15)	0.350(7)
		0.13620	0.01651(14)	0.1559(14)	0.0722(16)	0.346(8)
		0.13625	0.01522(14)	0.1499(15)	0.0719(17)	0.344(9)

Table 10. Combinations of m_{rs}, $M_{\rm PS}$, $G_{\rm PS}$ and $M_{\rm V}$ (lattices $D_1 - D_5$)

Run	κ_r	κ_s	$aM_{\rm PS}^2/m_{rs}$	$aF_{\rm PS}$	$F_{\rm PS}/M_{\rm V}$
D_1	0.13550	0.13550	1.594(9)	0.0671(9)	0.1445(20)
		0.13590	1.596(10)	0.0645(9)	0.1444(21)
		0.13610	1.598(11)	0.0631(9)	0.1440(23)
		0.13620	1.601(12)	0.0624(9)	0.1438(23)
D_2	0.13590	0.13550	1.531(9)	0.0601(6)	0.1441(17)
		0.13590	1.547(10)	0.0576(7)	0.1435(20)
		0.13610	1.565(12)	0.0562(7)	0.1428(22)
		0.13620	1.581(14)	0.0556(7)	0.1424(24)
D_3	0.13610	0.13550	1.454(11)	0.0558(6)	0.1461(20)
		0.13590	1.465(14)	0.0533(7)	0.1467(23)
		0.13610	1.480(17)	0.0518(7)	0.146(3)
		0.13620	1.492(19)	0.0510(8)	0.146(3)
D_4	0.13620	0.13550	1.462(13)	0.0556(6)	0.1487(22)
		0.13590	1.478(17)	0.0531(7)	0.149(3)
		0.13610	1.494(20)	0.0516(7)	0.148(3)
		0.13620	1.505(23)	0.0508(7)	0.147(3)
D_5	0.13625	0.13550	1.456(15)	0.0539(8)	0.143(3)
		0.13590	1.457(19)	0.0512(8)	0.143(3)
		0.13610	1.464(24)	0.0498(8)	0.142(3)
		0.13620	1.47(3)	0.0490(9)	0.142(4)
		0.13625	1.48(3)	0.0487(9)	0.142(4)

References

[1] K. G. Wilson, Phys. Rev. D10 (1974) 2445

[2] M. Hasenbusch, Phys. Lett. B519 (2001) 177

[3] M. Hasenbusch, K. Jansen, Nucl. Phys. B659 (2003) 299

[4] M. Della Morte et al. (ALPHA collab.), Comput. Phys. Commun. 156 (2003) 62

[5] M. Lüscher, JHEP 0305 (2003) 052

[6] M. Lüscher, Comput. Phys. Commun. 156 (2004) 209

[7] M. Lüscher, Comput. Phys. Commun. 165 (2005) 199

[8] C. Urbach, K. Jansen, A. Shindler, U. Wenger, Comput. Phys. Commun. 174 (2006) 87

[9] L. Del Debbio, L. Giusti, M. Lüscher, R. Petronzio, N. Tantalo, QCD with light Wilson quarks on fine lattices (I): first experiences and physics results, hep-lat/0610059, to appear in JHEP

[10] M. Göckeler et al. (QCDSF–UKQCD collab.), PoS (LAT2006) 160 and 179

[11] H. Meyer, O. Witzel (ALPHA collab.), PoS (LAT2006) 032

[12] K-I. Ishikawa et al. (PACS-CS collab.), PoS (LAT2006) 027

[13] Y. Kuramashi et al. (PACS-CS collab.), PoS (LAT2006) 029

[14] A. Ukawa et al. (PACS-CS collab.), PoS (LAT2006) 039

[15] K. Jansen, C. Urbach (ETM collab.), PoS (LAT2006) 203

[16] B. Sheikholeslami, R. Wohlert, Nucl. Phys. B259 (1985) 572

[17] M. Lüscher, S. Sint, R. Sommer, P. Weisz, Nucl. Phys. B478 (1996) 365

[18] S. Duane, A. D. Kennedy, B. J. Pendleton, D. Roweth, Phys. Lett. B195 (1987) 216

[19] J. C. Sexton, D. H. Weingarten, Nucl. Phys. B380 (1992) 665

[20] L. Del Debbio, M. Lüscher, L. Giusti, R. Petronzio, N. Tantalo, JHEP 0602 (2006) 011

[21] M. Della Morte et al. (ALPHA collab.), JHEP 0507 (2005) 007

[22] S. R. Sharpe, Phys. Rev. D74 (2006) 014512

[23] M. Golterman, Y. Shamir, B. Svetitsky, Phys. Rev. D71 (2005) 071502; *ibid.* D72 (2005) 034501

[24] C. R. Allton et al. (UKQCD collab.), Phys. Rev. D70 (2004) 014501

[25] Y. Namekawa et al. (CP-PACS collab.), Phys. Rev. D70 (2004) 074503

[26] M. Della Morte, R. Hoffmann, F. Knechtli, U. Wolff, Comput. Phys. Commun. 165 (2005) 49

[27] C. R. Allton et al. (UKQCD collab.), Phys. Rev. D47 (1993) 5128

[28] K. Jansen, R. Sommer (ALPHA collab.), Nucl. Phys. B530 (1998) 185 [E: *ibid.* B643 (2002) 517]
[29] M. Della Morte, R. Hoffmann, R. Sommer, JHEP 0503 (2005) 029
[30] M. Lüscher, P. Weisz, Nucl. Phys. B240 (1984) 349
[31] M. Lüscher, Nucl. Phys. B364 (1991) 237
[32] S. Aoki et al. (CP-PACS collab.), PoS (LAT2006) 110
[33] T. Bhattacharya, R. Gupta, W. Lee, S. R. Sharpe, J. M. S. Wu, Phys. Rev. D73 (2006) 034504