2+1 flavor lattice QCD toward the physical point

著者	Aoki S., Ishikawa K.-I., Ishizuka N., Izubuchi T., Kadoh D., Kanaya K., Kuramashi Y, Namekawa Y., Okawa M., Taniguchi Y., Ukawa A., Ukita N., Yoshie´ T.
journal or publication title	Physical review D
volume	79
number	03
page range	034503
year	2009-02
権利	(C) 2009 The American Physical Society
URL	http://hdl.handle.net/2241/102508

doi: 10.1103/PhysRevD.79.034503

PHYSICAL REVIEW D **79,** 034503 (2009)

2 + 1 flavor lattice QCD toward the physical point

S. Aoki,[1,2] K.-I. Ishikawa,[4] N. Ishizuka,[1,3] T. Izubuchi,[2,5] D. Kadoh,[3] K. Kanaya,[1] Y. Kuramashi,[1,3] Y. Namekawa,[3] M. Okawa,[4] Y. Taniguchi,[1,3] A. Ukawa,[1,3] N. Ukita,[3] and T. Yoshié[1,3]

(PACS-CS Collaboration)

[1]*Graduate School of Pure and Applied Sciences, University of Tsukuba, Tsukuba, Ibaraki 305-8571, Japan*
[2]*Riken BNL Research Center, Brookhaven National Laboratory, Upton, New York 11973, USA*
[3]*Center for Computational Sciences, University of Tsukuba, Tsukuba, Ibaraki 305-8577, Japan*
[4]*Graduate School of Science, Hiroshima University, Higashi-Hiroshima, Hiroshima 739-8526, Japan*
[5]*Institute for Theoretical Physics, Kanazawa University, Kanazawa, Ishikawa 920-1192, Japan*

(Received 5 August 2008; published 5 February 2009)

We present the first results of the PACS-CS project which aims to simulate 2 + 1 flavor lattice QCD on the physical point with the nonperturbatively $O(a)$-improved Wilson quark action and the Iwasaki gauge action. Numerical simulations are carried out at $\beta = 1.9$, corresponding to the lattice spacing of $a = 0.0907(13)$ fm, on a $32^3 \times 64$ lattice with the use of the domain-decomposed HMC algorithm to reduce the up-down quark mass. Further algorithmic improvements make possible the simulation whose up-down quark mass is as light as the physical value. The resulting pseudoscalar meson masses range from 702 MeV down to 156 MeV, which clearly exhibit the presence of chiral logarithms. An analysis of the pseudoscalar meson sector with SU(3) chiral perturbation theory reveals that the next-to-leading order corrections are large at the physical strange quark mass. In order to estimate the physical up-down quark mass, we employ the SU(2) chiral analysis expanding the strange quark contributions analytically around the physical strange quark mass. The SU(2) low energy constants $\bar{l}_3$ and $\bar{l}_4$ are comparable with the recent estimates by other lattice QCD calculations. We determine the physical point together with the lattice spacing employing m_π, m_K and m_Ω as input. The hadron spectrum extrapolated to the physical point shows an agreement with the experimental values at a few % level of statistical errors, albeit there remain possible cutoff effects. We also find that our results of f_π, f_K and their ratio, where renormalization is carries out perturbatively at one loop, are compatible with the experimental values. For the physical quark masses we obtain $m_{ud}^{\overline{\mathrm{MS}}}$ and $m_s^{\overline{\mathrm{MS}}}$ extracted from the axial-vector Ward-Takahashi identity with the perturbative renormalization factors. We also briefly discuss the results for the static quark potential.

DOI: 10.1103/PhysRevD.79.034503 PACS numbers: 11.15.Ha, 12.38.−t, 12.38.Gc

I. INTRODUCTION

Lattice QCD is expected to be an ideal tool to understand the nonperturbative dynamics of strong interactions from first principles. In order to fulfill this promise, the first step should be to establish QCD as the fundamental theory of the strong interaction by reproducing basic physical quantities, *e.g.,* the hadron spectrum, with the systematic errors under control. This is about to be attained thanks to the recent progress of simulation algorithms and the availability of increasingly more powerful computational resources.

Among various systematic errors, the two most troublesome are quenching effects and chiral extrapolation uncertainties. After the systematic studies on the hadron spectrum in quenched and two-flavor QCD [1–3], the CP-PACS and JLQCD collaborations performed a 2 + 1 flavor full QCD simulation employing the nonperturbatively $O(a)$-improved Wilson quark action [4] and the Iwasaki-gauge action [5] on a $(2\ \mathrm{fm})^3$ lattice at three lattice spacings [6,7]. While the quenching effects were successfully removed, we were left with a long chiral extrapolation: the lightest up-down quark mass reached with the plain hybrid Monte Carlo (HMC) algorithm was about 67 MeV corresponding to $m_\pi/m_\rho \approx 0.6$.

The PACS-CS project, which is based on the PACS-CS (Parallel Array Computer System for Computational Sciences) computer with a peak speed of 14.3 Tflops developed at University of Tsukuba [8–10], aims at calculations on the physical point to remove the ambiguity of chiral extrapolations. It employs the same quark and gauge actions as the previous CP-PACS/JLQCD work, but uses a different simulation algorithm: the up-down quark mass is reduced by using the domain-decomposed HMC (DDHMC) algorithm with the replay trick [11,12]. At the lightest up-down quark mass, which is about 3 MeV, several algorithmic improvements are incorporated, including the mass-preconditioning [13,14], the chronological inverter [15], and the deflation technique [16]. For the strange quark part we improve the PHMC algorithm [17–19] with the UV-filtering procedure [20,21].

So far our simulation points cover from 702 MeV to 156 MeV for the pion mass. While we still have to reduce the pion mass by 21 MeV to reach the real physical point, we consider that the findings so far already merits a detailed report. In this paper we focus on the following points: (i) several algorithmic improvements make possible a simulation with the up-down quark mass as light as the physical value. (ii) The range of pion mass we have simulated is sufficiently light to deserve chiral analyses with the chiral perturbation theory (ChPT), which reveals that the strange quark mass is not small enough to be treated by the SU(3) ChPT up to the next-to-leading order (NLO). (iii) The SU(2) chiral analysis on the pion sector and the linear chiral extrapolation for other hadron masses yield the hadron spectrum at the physical point which is compatible with the experimental values at a few % level of statistical errors.

This paper is organized as follows. In Sec. II we present the simulation details. Measurements of hadron masses, pseudoscalar meson decay constants and quark masses are described in Sec. III. In Sec. IV we make chiral analyses on the pseudoscalar meson sector using the SU(3) and SU(2) ChPTs. We present the values of low energy constants and discuss convergences of the SU(3) and SU(2) chiral expansions. The results of hadron spectrum at the physical point are given in Sec. V together with the pseudoscalar meson decay constants and the quark masses. In Sec. VI we show the results for the static quark potential. Our conclusions are summarized in Sec. VII. Appendices are devoted to describe the algorithmic details. Preliminary results have been reported in Refs. [22–24].

II. SIMULATION DETAILS

A. Actions

We employ the Iwasaki-gauge action [5] and the nonperturbatively $O(a)$-improved Wilson quark action as in the previous CP-PACS/JLQCD work. The former is composed of a plaquette and a 1×2 rectangle loop:

$$S_g = \frac{1}{g^2}\left\{c_0 \sum_{\text{plaquette}} \text{tr}U_{\text{pl}} + c_1 \sum_{\text{rectangle}} \text{tr}U_{rtg}\right\} \tag{1}$$

with $c_1 = -0.331$ and $c_0 = 1\text{–}8c_1 = 3.648$. The latter is expressed as

$$\begin{aligned} S_{\text{quark}} = \sum_{q=u,d,s}\Bigg[& \sum_n \bar{q}_n q_n - \kappa_q c_{\text{SW}} \sum_n \sum_{\mu,\nu} \frac{i}{2} \bar{q}_n \sigma_{\mu\nu} F_{\mu\nu}(n) q_n \\ & - \kappa_q \sum_n \sum_\mu \{\bar{q}_n (1-\gamma_\mu) U_{n,\mu} q_{n+\hat{\mu}} \\ & + \bar{q}_n (1+\gamma_\mu) U^\dagger_{n-\hat{\mu},\mu} q_{n-\hat{\mu}}\}\Bigg], \end{aligned} \tag{2}$$

where we consider the case of a degenerate up and down quark mass $\kappa_u = \kappa_d$. The Euclidean gamma matrices are defined in terms of the Minkowski matrices in the Bjorken-Drell convention: $\gamma_j = -i\gamma^j_{\text{BD}}$ $(j = 1, 2, 3)$, $\gamma_4 = \gamma^0_{\text{BD}}$, $\gamma_5 = \gamma^5_{\text{BD}}$, and $\sigma_{\mu\nu} = \frac{1}{2}[\gamma_\mu, \gamma_\nu]$. The field strength $F_{\mu\nu}$ in the clover term is given by

$$F_{\mu\nu}(n) = \frac{1}{4}\sum_{i=1}^{4} \frac{1}{2i}(U_i(n) - U_i^\dagger(n)), \tag{3}$$

$$U_1(n) = U_{n,\mu} U_{n+\hat{\mu},\nu} U^\dagger_{n+\hat{\nu},\mu} U^\dagger_{n,\nu}, \tag{4}$$

$$U_2(n) = U_{n,\nu} U^\dagger_{n-\hat{\mu}+\hat{\nu},\mu} U^\dagger_{n-\hat{\mu},\nu} U_{n-\hat{\mu},\mu}, \tag{5}$$

$$U_3(n) = U^\dagger_{n-\hat{\mu},\mu} U^\dagger_{n-\hat{\mu}-\hat{\nu},\nu} U_{n-\hat{\mu}-\hat{\nu},\mu} U_{n-\hat{\nu},\nu}, \tag{6}$$

$$U_4(n) = U^\dagger_{n-\hat{\nu},\nu} U_{n-\hat{\nu},\mu} U_{n+\hat{\mu}-\hat{\nu},\nu} U^\dagger_{n,\mu}. \tag{7}$$

The improvement coefficient c_{SW} for $O(a)$ improvement was determined nonperturbatively in Ref. [4].

B. Simulation parameters

Our simulations are carried out at $\beta = 1.90$ on a $32^3 \times 64$ lattice for which we use $c_{\text{SW}} = 1.715$ [4]. This β value is one of the three in the previous CP-PACS/JLQCD work, whereas the lattice size is enlarged from $20^3 \times 40$ to investigate the baryon masses. The lattice spacing is found to be 0.0907(14) fm whose determination is explained later. Table I lists the run parameters of our simulations. The six combinations of the hopping parameters (κ_{ud}, κ_s) are chosen based on the previous CP-PACS/JLQCD results. The heaviest combination $(\kappa_{ud}, \kappa_s) = (0.13700, 0.13640)$ in this work corresponds to the lightest one in the previous CP-PACS/JLQCD simulations, which enable us to make a direct comparison of the two results with different lattice sizes. The physical point of the strange quark at $\beta = 1.90$ was estimated as $\kappa_s = 0.136412(50)$ in the CP-PACS/JLQCD work [6,7]. This is the reason why all our simulations are carried out with $\kappa_s = 0.13640$, the one exception being the run at $(\kappa_{ud}, \kappa_s) = (0.13754, 0.13660)$ to investigate the strange quark mass dependence. After more than 1000 MD time for thermalization we calculate hadronic observables solving quark propagators at every 10 trajectories for $\kappa_{ud} \geq 0.13770$ and 20 trajectories for $\kappa_{ud} = 0.13781$, while we measure the plaquette expectation value at every trajectory.

C. Algorithm

Our base algorithm for penetrating into the small mass region for a degenerate pair of up and down quarks is the DDHMC algorithm [11]. The effectiveness of this algorithm for reducing the quark mass was already shown in the $N_f = 2$ case [11,25,26]. We found that it works down to $\kappa_{ud} = 0.13770$ (or $m_\pi \approx 300$ MeV) on our $32^3 \times 64$ lattice. Moving closer to the physical point, however, we found it necessary to add further enhancements including

TABLE I. Simulation parameters. MD time is the number of trajectories multiplied by the trajectory length τ. $\tau_{\rm int}[P]$ denotes the integrated autocorrelation time for the plaquette. CPU time for unit τ using 256 nodes of PACS-CS is also listed.

$\kappa_{\rm ud}$	0.13700	0.13727	0.13754	0.13754	0.13770	0.13781
κ_s	0.13640	0.13640	0.13640	0.13660	0.13640	0.13640
#run	1	1	2	4	2	5
τ	0.5	0.5	0.5	0.5	0.25	0.25
(N_0, N_1, N_2, N_3)	(4,4,10)	(4,4,14)	(4,4,20)	(4,4,28)	(4,4,16)	(4,4,4,6) (4,4,6,6)
ρ	-	-	-	-	-	0.9995
$N_{\rm poly}$	180	180	180	220	180	200
Replay	on	on	on	on	on	off
MD time	2000	2000	2250	2000	2000	990
$\langle P \rangle$	0.569105(18)	0.569727(14)	0.570284(16)	0.570554(17)	0.570573(20)	0.570868(9)
$\langle e^{-dH} \rangle$	0.9922(85)	1.0016(50)	1.0013(56)	0.9993(36)	0.9944(53)	0.970(12)
$P_{\rm acc}$ (HMC)	0.8020(63)	0.8672(47)	0.8573(52)	0.9140(44)	0.8397(41)	0.8033(63)
$P_{\rm acc}$ (GMP)	0.9529(37)	0.9439(34)	0.9331(40)	0.9330(41)	0.9537(26)	0.9670(32)
$\tau_{\rm int}[P]$	8.6(3.1)	20.9(10.2)	9.8(2.8)	6.3(1.4)	25.2(15.2)	2.9(1.9)
CPU hour/unit τ	0.29	0.44	1.3	1.1	2.7	7.1

mass preconditioning, which we call mass-preconditioned DDHMC (MPDDHMC). This is the algorithm we applied at our lightest point at $\kappa_{ud} = 0.13781$.

The characteristic feature of the DDHMC algorithm is a geometric separation of the up-down quark determinant into the UV and the IR parts, which is implemented by domain-decomposing the full lattice into small blocks. We choose 8^4 for the block size, being less than $(1\ {\rm fm})^4$ in physical units and small enough to reside within a computing node of the PACS-CS computer. The latter feature is computationally advantageous since the calculation of the UV part requires no communication between blocks so that the internode communications are sizably reduced.

The UV/IR separation enables the application of multiple time scale integration schemes [27], which reduces the simulation cost substantially. In our simulation points we find that the relative magnitudes of the force terms are

$$\|F_g\|:\|F_{\rm UV}\|:F_{\rm IR}\| \approx 16{:}4{:}1, \tag{8}$$

where we adopt the convention $\|M\|^2 = -2\,{\rm tr}(M^2)$ for the norm of an element M of the SU(3) Lie algebra, and F_g denotes the gauge part and $F_{\rm UV,IR}$ are for the UV and the IR parts of the up-down quarks. The associated step sizes for the forces are controlled by three integers $N_{0,1,2}$ introduced by $\delta\tau_g = \tau/N_0N_1N_2$, $\delta\tau_{\rm UV} = \tau/N_1N_2$, $\delta\tau_{\rm IR} = \tau/N_2$ with τ the trajectory length. The integers $N_{0,1,2}$ should be chosen such that

$$\delta\tau_g\|F_g\| \approx \delta\tau_{\rm UV}\|F_{\rm UV}\| \approx \delta\tau_{\rm IR}\|F_{\rm IR}\|. \tag{9}$$

The relative magnitudes between the forces in Eq. (8) tell us that $\delta\tau_{\rm IR}$ may be chosen roughly 16 times as large as $\delta\tau_g$ and 4 times that of $\delta\tau_{\rm UV}$, which means that we need to calculate $F_{\rm IR}$ an order of magnitude less frequently in the molecular dynamics trajectories. Since the calculation of $F_{\rm IR}$ contains the quark matrix inversion on the full lattice, which is the most computer time consuming part, this integration scheme saves the simulation cost remarkably.

The values for $N_{0,1,2}$ are listed in Table I, where N_0 and N_1 are fixed at 4 for all the hopping parameters, while the value of N_2 is adjusted taking account of acceptance rate and simulation stability. The threshold for the replay trick [11,12] for dealing with instabilities of molecular dynamics trajectories leading to large values of dH is set to be $\Delta H > 2$.

For the strange quark, we employ the UV-filtered PHMC (UVPHMC) algorithm [21]. The UVPHMC action for the strange quark is obtained through the UV-filtering [20] applied after the even-odd site preconditioning for the quark matrix. The domain-decomposition is not used. The polynomial approximation is corrected by the global Metropolis test [28]. Since we find $\|F_s\| \approx \|F_{\rm IR}\|$, the step size is chosen as $\delta\tau_s = \delta\tau_{\rm IR}$. The polynomial order for UVPHMC, which is denoted by $N_{\rm poly}$ in Table I, is adjusted to yield high acceptance rate for the global Metropolis test at the end of each trajectory.

The inversion of the Wilson-Dirac operator D on the full lattice is carried out by the SAP (Schwarz alternating procedure) preconditioned GCR solver. The preconditioning is accelerated with the single-precision arithmetic [29]. We employ the stopping condition $|Dx - b|/|b| < 10^{-9}$ for the force calculation and 10^{-14} for the Hamiltonian, which guarantees the reversibility of the molecular dynamics trajectories to a high precision: $|\Delta U| < 10^{-12}$ for the link variables and $|\Delta H| < 10^{-8}$ for the Hamiltonian at $(\kappa_{ud}, \kappa_s) = (0.13781, 0.13640)$. We describe the details of the DDHMC algorithm and the solver implementation used for $\kappa_{ud} \leq 0.13770$ in Appendix A.

As we reduce the up-down quark mass, we observe a tendency that the fluctuation of $\|F_{\rm IR}\|$ during the molecular

dynamics trajectory increases, which results in a higher replay rate due to the appearance of trajectories with large ΔH. Since ΔH is controlled by the product of $\delta\tau_{\rm IR}$ and $\|F_{\rm IR}\|$, a possible solution to suppress the replay rate is to reduce $\delta\tau_{\rm IR}$. In this case, however, we find the acceptance becoming unnecessarily close to unity. Another solution would be to tame the fluctuation of $\|F_{\rm IR}\|$, and we employ for this purpose the mass preconditioner [13,14] to the IR part of the pseudofermion action. The quark mass in the preconditioner is controlled by an additional hopping parameter $\kappa'_{ud} = \rho\kappa_{ud}$, where ρ should be less than unity so that calculating with the preconditioner is less costly than with the original IR part. The IR force $F_{\rm IR}$ is split into $F'_{\rm IR}$ and $\tilde{F}_{\rm IR}$. The former is derived from the preconditioner and the latter from the preconditioned action.

We employ the mass-preconditioned DDHMC (MPDDHMC) algorithm for the run at the lightest up-down quark mass of $\kappa_{ud} = 0.13781$. With our choice of $\rho = 0.9995$ the relative magnitudes of the force terms become

$$\|F_g\|:\|F_{\rm UV}\|:\|F'_{\rm IR}\|:\|\tilde{F}_{\rm IR}\| \approx 16{:}4{:}1{:}1/7. \qquad (10)$$

According to this result we choose $(N_0, N_1, N_2, N_3) = (4, 4, 4, 6)$ for the associated step sizes. Here the choice of $N_2 = 4$ does not follow the criterion $\delta\tau'_{\rm IR}\|F'_{\rm IR}\| \approx \delta\tilde{\tau}_{\rm IR}\|\tilde{F}_{\rm IR}\|$. This is because we take account of the fluctuations of $\|\tilde{F}_{\rm IR}\|$. The replay trick is not implemented in the runs at $\kappa_{ud} = 0.13781$. For the step size for the strange quark in the UVPHMC algorithm we choose $\delta\tau_s = \delta\tau'_{\rm IR}$ as we observe $\|F_s\| \approx \|F'_{\rm IR}\|$.

The inversion of D during the molecular dynamics steps is also improved at $\kappa_{ud} = 0.13781$ in three ways. (i) We employ the chronological guess using the last 16 solutions to construct the initial solution vector of D^{-1} on the full lattice [15]. In order to assure the reversibility we apply a stringent stopping condition $|Dx - b|/|b| < 10^{-14}$ to the force calculation. (ii) The inversion algorithm is replaced by a nested BiCGStab solver, which consists of an inner solver accelerated with single precision arithmetic and with an automatic tolerance control ranging from 10^{-3} to 10^{-6}, and an outer solver with a stringent tolerance of 10^{-14} operated with the double precision. The approximate solution obtained by the inner solver works as a preconditioner for the outer solver. (iii) We implement the deflation technique to make the solver robust against possible small eigenvalues allowed in the Wilson-type quark action. Once the inner BiCGStab solver becomes stagnant during the inversion of D, it is automatically replaced by the GCRO-DR (generalized conjugate Residual with implicit inner orthogonalization and deflated restarting) algorithm [16]. In our experience the GCRO-DR algorithm is important for calculating D^{-1} but does not save the simulation time at $\kappa_{ud} = 0.13781$. More details of the MPDDHMC algorithm and the improvements are given in Appendix B.

D. Implementation on the PACS-CS computer

All of the simulations reported in this article have been carried out on the PACS-CS parallel computer [10]. PACS-CS consists of 2560 nodes, each node equipped with a 2.8 GHz Intel Xeon single-core processor (*i.e.,* 5.6 Gflops of peak speed) with 2 GBytes of main memory. The nodes are arranged into a $16 \times 16 \times 10$ array and connected by a 3-dimensional hypercrossbar network made of a dual Gigabit Ethernet in each direction. The network bandwidth is 750 MBytes/sec for each node.

The programming language is mainly Fortran 90 with Intel Fortran compiler. To further enhance the performance we used Intel C + + compiler for the single precision hopping matrix multiplication routines which are the most time consuming parts. The Intel compiler enables us to use the Intel streaming SIMD extensions 2 and 3 intrinsics directly without writing assembler language.

We employ a 256 node partition of PACS-CS to execute our $32^3 \times 64$ runs. The sustained performance including communication overhead with our DDHMC code turns out to be 18%. The computer time needed for one MD unit is listed in Table I.

E. Efficiency of DDHMC algorithms

The efficiency of the DDHMC algorithm may be clarified in comparison with that of the HMC algorithm. For $N_f = 2$ QCD simulations with the Wilson-clover quark action, an empirical cost formula suggested for the HMC algorithm based on the CP-PACS and JLQCD $N_f = 2$ runs was as follows [30]:

$$\text{cost[Tflops} \cdot \text{years]} = C\left[\frac{\#\text{conf}}{1000}\right] \cdot \left[\frac{0.6}{m_\pi/m_\rho}\right]^6 \cdot \left[\frac{L}{3\ \text{fm}}\right]^5 \cdot \left[\frac{0.1\ \text{fm}}{a}\right]^7 \qquad (11)$$

with $C \approx 2.8$. A strong quark mass dependence in the above formula $1/(m_\pi/m_\rho)^6 \sim 1/m_{ud}^3$ stems from three factors: (i) the number of iterations for the quark matrix inversion increases as the condition number which is proportional to $1/m_{ud}$, (ii) to keep the acceptance rate constant we should take $\delta\tau \propto m_{ud}$ for the step size in the molecular dynamics trajectories, and (iii) the autocorrelation time of the HMC evolution was consistent with an $1/m_{ud}$ dependence in the CP-PACS runs [2].

To estimate the computational cost for $N_f = 2 + 1$ QCD simulations with the HMC algorithm, we assume that the strange quark contribution is given by half of Eq. (11) at $m_\pi/m_\rho = 0.67$ which is a phenomenologically estimated ratio of the strange pseudoscalar meson "$m_{\eta_{ss}}$" and m_ϕ:

$$\frac{m_{\eta_{ss}}}{m_\phi} = \frac{\sqrt{2m_K^2 - m_\pi^2}}{m_\phi} \approx 0.67.$$

Since the strange quark is relatively heavy, its computa-

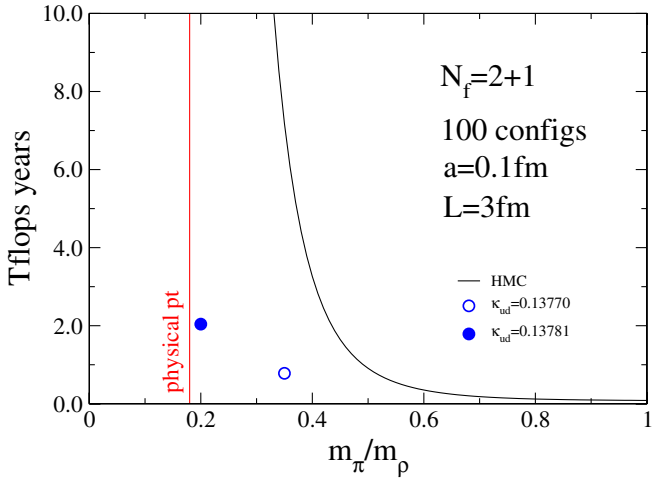


FIG. 1 (color online). Simulation cost at $(\kappa_{ud}, \kappa_s) = (0.13770, 0.13640)$ by DDHMC (blue open circle) and $(\kappa_{ud}, \kappa_s) = (0.13781, 0.13640)$ by MPDDHMC (blue closed circle) for 10000 trajectories. Solid line indicates the cost estimate of $N_f = 2 + 1$ QCD simulations with the HMC algorithm at $a = 0.1$ fm with $L = 3$ fm for 100 independent configurations. Vertical line denotes the physical point.

tional cost occupies only a small fraction as the up-down quark mass decreases. In Fig. 1 we draw the cost formula for the $N_f = 2 + 1$ case as a function of m_π/m_ρ, where we take #conf $= 100$, $a = 0.1$ fm and $L = 3$ fm in Eq. (11) as a representative case. We observe a steep increase of the computational cost below $m_\pi/m_\rho \simeq 0.5$. At the physical point the cost expected from Eq. (11) would be $O(100)$ Tflops · years.

Let us now see the situation with the DDHMC algorithm. The blue open symbol in Fig. 1 denotes the measured cost at $(\kappa_{ud}, \kappa_s) = (0.13770, 0.13640)$, which is the lightest point implemented with the DDHMC algorithm. Here we assume that we need 100 MD time separation between independent configurations. We observe a remarkable reduction in the cost by a factor 20–30 in magnitude. The majority of this reduction arises from the multiple time scale integration scheme and the GCR solver accelerated by the SAP preconditioning with the single-precision arithmetic. Roughly speaking, the improvement factor is $O(10)$ for the former and 3–4 for the latter. The cost of the MPDDHMC algorithm at $(\kappa_{ud}, \kappa_s) = (0.13781, 0.13640)$ is plotted by the blue closed symbol in Fig. 1. In this case, the reduction is mainly owing to the multiple time scale integration scheme armored with the mass-preconditioning and the chronological inverter for F'_{IR} and $\tilde{F}_{\mathrm{IR}}$. As we already noted, the GCRO-DR solver does not accelerate the inversion albeit it renders the solver robust against the small eigenvalues of the Wilson-Dirac operator.

Since we find in Table I that $\tau_{\mathrm{int}}[P]$ is roughly independent of the up-down quark mass employed in the DDHMC algorithm, the cost is expected to be proportional to $1/m_{ud}^2$. Assuming this quark mass dependence for the MPDDHMC algorithm, we find that simulations at the physical point is feasible, at least for $L \approx 3$ fm lattices, with $O(10)$ Tflops computers, which are already available at present.

F. Autocorrelations and statistical error analysis

The autocorrelation function $\Gamma(\tau)$ of a time series of an observable $\mathcal{O}$ in the course of a numerical simulation is given by

$$\Gamma(\tau) = \langle \mathcal{O}(\tau_0)\mathcal{O}(\tau_0 + \tau)\rangle - \langle \mathcal{O}(\tau_0)\rangle^2. \qquad (12)$$

In Fig. 2 we show the plaquette history and the normalized autocorrelation function $\rho(\tau) = \Gamma(\tau)/\Gamma(0)$ at $\kappa_{ud} = 0.13727$ as an example. The integrated autocorrelation time is estimated as $\tau_{\mathrm{int}}[P] = 20.9(10.2)$ following the definition in Ref. [11,31]

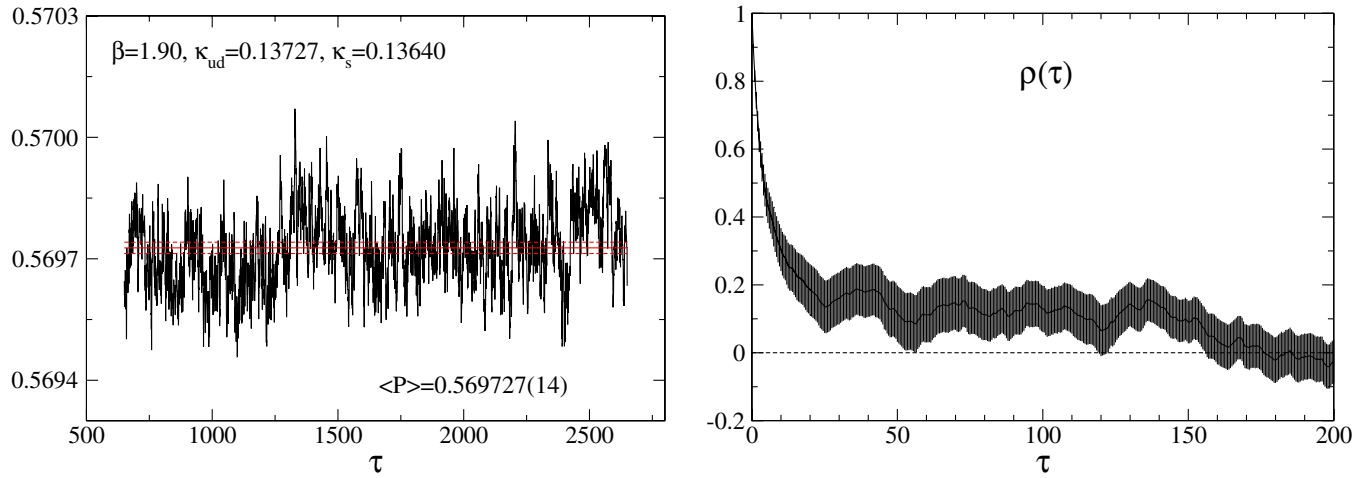


FIG. 2 (color online). Plaquette history (left) and normalized autocorrelation function (right) for $(\kappa_{ud}, \kappa_s) = (0.13727, 0.13640)$. Horizontal lines in the left denote the average value of the plaquette with one standard deviation error band.

$$\tau_{\rm int}(\tau) = \frac{1}{2} + \sum_{0<\tau\le W} \rho(\tau), \tag{13}$$

where the summation window W is set to the first time lag τ such that $\rho(\tau)$ becomes consistent with zero within the error bar. In this case we find $W = 119.5$. The choice of W is not critical for estimate of $\tau_{\rm int}$ in spite of the long tail observed in Fig. 2. Extending the summation window, we find that $\tau_{\rm int}[P]$ saturates at $\tau_{\rm int}[P] \approx 25$ beyond $W = 200$, which is within the error bar of the original estimate.

Our simulations at $\kappa_{ud} = 0.13700$ and 0.13727 are fast enough to be executed by a single long run of 2000 MD units. The simulations at $\kappa_{ud} \ge 0.13754$ become increasingly CPU time consuming so that we had to execute multiple runs in parallel. The data obtained from different runs are combined into a single extended series, for which we define the above autocorrelation function $\Gamma(\tau)$ as if it were a single run. The results for $\tau_{\rm int}[P]$ are listed in Table I. Although we hardly observe any systematic quark mass dependence for the integrated autocorrelation time, the statistics of the individual run may not be sufficiently large to derive a definite conclusion.

Since the plaquette is a local quantity, it is inappropriate to make use of its integrated autocorrelation time for estimate of the autocorrelation time of nonlocal physical quantities such as hadron masses. In the physics analysis we estimate the statistical errors with the jackknife method in order to take account of the autocorrelation. For the simulations at $\kappa_{ud} \ge 0.13754$ we apply the jackknife analysis after combining the different runs into a single series. The bin size dependence of the statistical error is investigated for each physical observable. For a cross-check we also carry out the bootstrap error estimation with 1000 samples. In all cases we find the two estimates agree for the magnitude of errors within 10%. We follow the procedure given in Appendix B of Ref. [2] in estimating the errors for the chiral fit parameters.

III. MEASUREMENTS OF HADRONIC OBSERVABLES

A. Hadron masses, quark masses and decay constants

We measure the meson and baryon correlators at the unitary points where the valence quark masses are equal to the sea quark masses. For the meson operators we employ

$$M_\Gamma^{fg}(x) = \bar{q}_f(x)\Gamma q_g(x), \tag{14}$$

where f and g denote quark flavors and Γ are 16 Dirac matrices $\Gamma = {\rm I}, \gamma_5, \gamma_\mu, i\gamma_\mu\gamma_5$ and $i[\gamma_\mu, \gamma_\nu]/2$ ($\mu, \nu = 1, 2, 3, 4$). The octet baryon operators are given by

$$\mathcal{O}_\alpha^{fgh}(x) = \epsilon^{abc}((q_f^a(x))^T C\gamma_5 q_g^b(x))q_{h\alpha}^c(x), \tag{15}$$

where a, b, c are color indices, $C = \gamma_4\gamma_2$ is the charge conjugation matrix and $\alpha = 1, 2$ labels the z component of the spin 1/2. The Σ- and Λ-like octet baryons are distinguished by the flavor structures:

$$\Sigma - \text{like:} \ -\frac{\mathcal{O}^{[fh]g} + \mathcal{O}^{[gh]f}}{\sqrt{2}}, \tag{16}$$

$$\Lambda - \text{like:} \ \frac{\mathcal{O}^{[fh]g} - \mathcal{O}^{[gh]f} - 2\mathcal{O}^{[fg]h}}{\sqrt{6}}, \tag{17}$$

where $\mathcal{O}^{[fg]h} = \mathcal{O}^{fgh} - \mathcal{O}^{gfh}$. We define the decuplet baryon operators for the four z components of the spin 3/2 as

$$D_{3/2}^{fgh}(x) = \epsilon^{abc}((q_f^a(x))^T C\Gamma_+ q_g^b(x))q_{h1}^c(x), \tag{18}$$

$$D_{1/2}^{fgh}(x) = \epsilon^{abc}[((q_f^a(x))^T C\Gamma_0 q_g^b(x))q_{h1}^c(x) - ((q_f^a(x))^T C\Gamma_+ q_g^b(x))q_{h2}^c(x)]/3, \tag{19}$$

$$D_{-1/2}^{fgh}(x) = \epsilon^{abc}[((q_f^a(x))^T C\Gamma_0 q_g^b(x))q_{h2}^c(x) - ((q_f^a(x))^T C\Gamma_- q_g^b(x))q_{h1}^c(x)]/3, \tag{20}$$

$$D_{-3/2}^{fgh}(x) = \epsilon^{abc}((q_f^a(x))^T C\Gamma_- q_g^b(x))q_{h2}^c(x), \tag{21}$$

where $\Gamma_\pm = (\gamma_1 \mp \gamma_2)/2$, $\Gamma_0 = \gamma_3$ and the flavor structures should be symmetrized.

We calculate the meson and the baryon correlators with point and smeared sources and a local sink. For the smeared source we employ an exponential smearing function $\Psi(|\vec{x}|) = A_q \exp(-B_q|\vec{x}|)$ ($q = ud, s$) with $\Psi(0) = 1$ for the ud and s quark propagators. The parameters A_q and B_q are adjusted from a couple of configurations after the beginning of the production run such that the pseudoscalar meson effective masses reach a plateau as soon as possible. Their values are given in Table II. The point and smeared sources allow the hadron propagators with nonzero spatial momentum, and we calculate them for $\vec{p} = (0, 0, 0)$, $(\pi/16, 0, 0)$, $(0, \pi/16, 0)$, $(0, 0, \pi/16)$.

In order to increase the statistics we calculate the hadron correlators with four source points at $(x_0, y_0, z_0, t_0) = (17, 17, 17, 1)$, $(1, 1, 1, 9)$, $(25, 25, 25, 17)$, and $(9, 9, 9, 25)$ for $\kappa_{ud} \ge 0.13754$. They are averaged on each configuration before the jackknife analysis. This procedure reduces the statistical errors by typically 20–40% for the

TABLE II. Smearing parameters A and B for the ud and s quark propagators.

κ_{ud}	κ_s	#source	A_{ud}	B_{ud}	A_s	B_s
0.13700	0.13640	1	1.2	0.21	1.2	0.28
0.13727	0.13640	1	1.2	0.19	1.2	0.25
0.13754	0.13640	4	1.2	0.17	1.2	0.25
0.13754	0.13660	4	1.2	0.17	1.2	0.25
0.13770	0.13640	4	1.2	0.09	1.2	0.21
0.13781	0.13640	4	1.2	0.07	1.2	0.20

vector meson and the baryon masses and less than 20% for the pseudoscalar meson masses compared to a single source point. For further enhancement of the signal we average zero momentum hadron propagators over possible spin states on each configuration: three polarization states for the vector meson and two (four) spin states for the octet (decuplet) baryons.

We extract the meson and the baryon masses from the hadron propagators with the point sink and the smeared source, where all the valence quark propagators in the mesons and the baryons have the smeared sources. Figures 3–6 show effective mass plots for the meson and the baryon propagators with the smeared source for $\kappa_{ud} \geq 0.13754$. While the excited state contributions are effectively suppressed and good plateaus start at small values of t, the statistical errors in the vector meson and the baryon channels grow rapidly for large times. In Fig. 7 we plot the noise-to-signal ratio of the π, ρ and nucleon propagators normalized by the value at $t = 10$. Their t-dependences are argued in Ref. [32], which predicts constant behavior for π, $\exp(-(m_\rho - m_\pi)t)$ for ρ and $\exp(-(m_N - 3/2m_\pi)t)$ for nucleon. The blue (red) line shows the expected t-dependence of the noise-to-signal ratio at the heaviest (lightest) case. The pion channel shows very mild t-dependence. Although we observe steeper t-dependences for the ρ and the nucleon channels, their growth rates seem hardly consistent with the theoretical expectations.

The hadron masses are extracted by uncorrelated χ^2 fits to the propagators without taking account of correlations between different time slices, since we encounter instabilities for correlated fits using covariance matrix. We assume

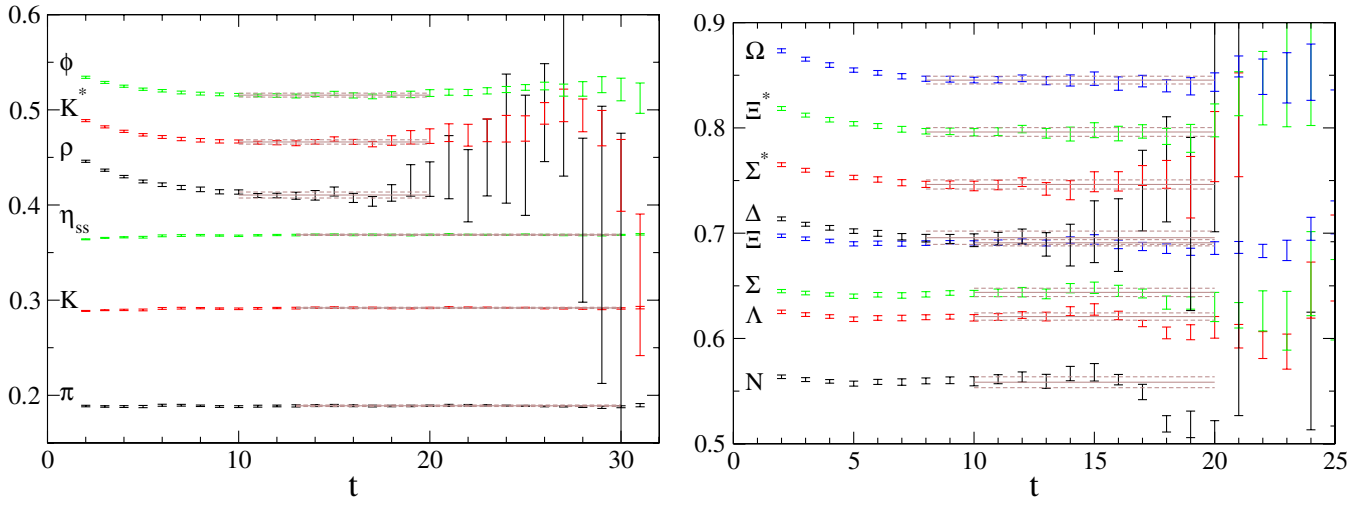


FIG. 3 (color online). Effective masses for the mesons (left) and the baryons (right) at $(\kappa_{ud}, \kappa_s) = (0.13754, 0.13640)$. Horizontal lines represent the fitting results with 1 standard deviation error band.

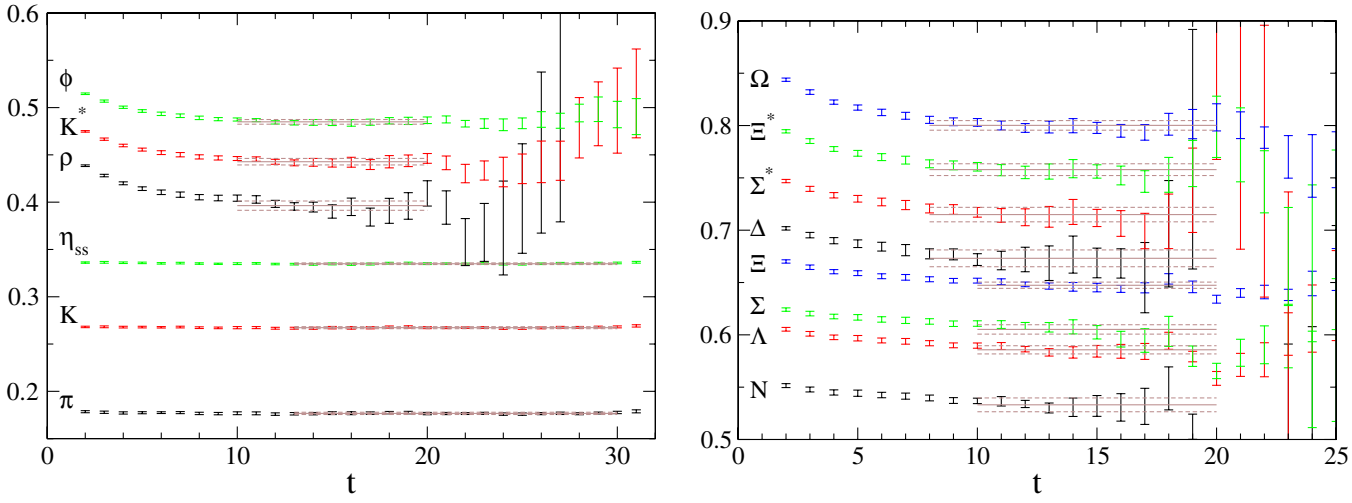


FIG. 4 (color online). Same as Fig. 3 for $(\kappa_{ud}, \kappa_s) = (0.13754, 0.13660)$.

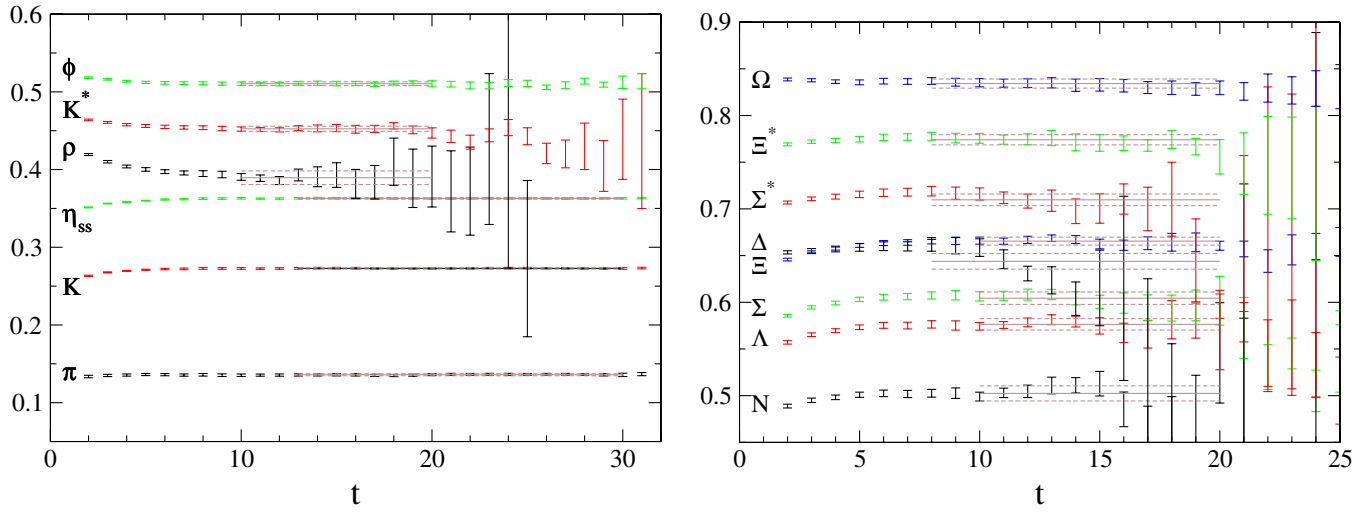


FIG. 5 (color online). Same as Fig. 3 for $(\kappa_{ud}, \kappa_s) = (0.13770, 0.13640)$.

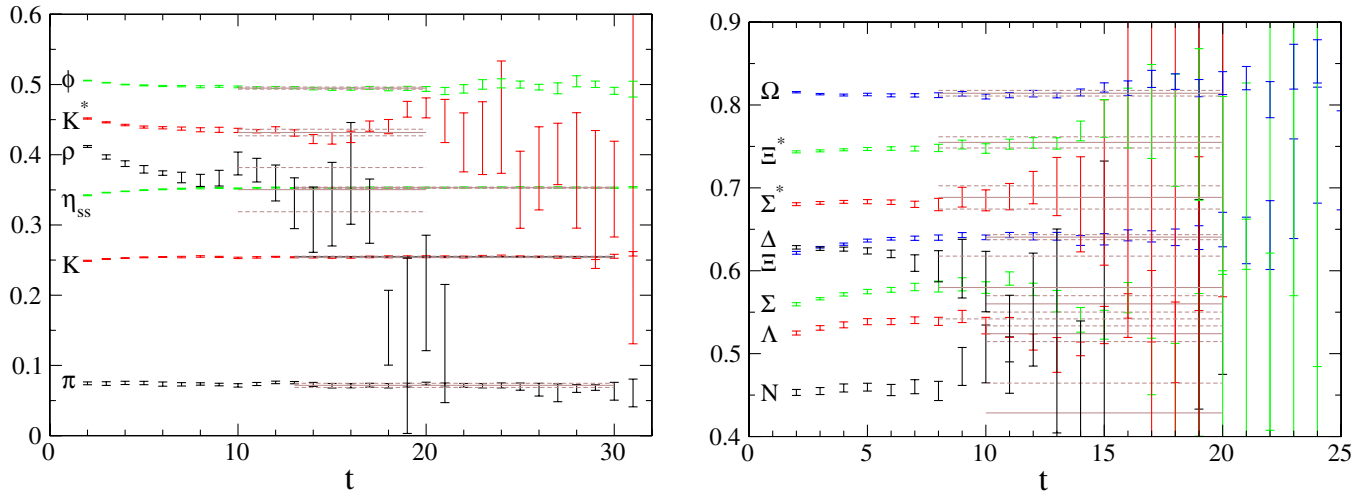


FIG. 6 (color online). Same as Fig. 3 for $(\kappa_{ud}, \kappa_s) = (0.13781, 0.13640)$.

a single hyperbolic cosine function for the mesons and a single exponential form for the baryons. The lower end of the fit range $t_{\rm min}$ is determined by investigating stability of the fitted mass. On the other hand, the choice of $t_{\rm max}$ gives little influence on the fit results as far as the effective mass exhibits a plateau and the signal is not lost in the noise. We employ the same fit range $[t_{\rm min}, t_{\rm max}]$ for the same particle type: [13, 30] for pseudoscalar mesons, [10, 20] for vector mesons, [10, 20] for octet baryons and [8, 20] for decuplet baryons. These fit ranges are independent of the quark masses. Resulting hadron masses are summarized in Table III.

Statistical errors are estimated with the jackknife procedure. In Fig. 8 we show the bin size dependence of the error for m_π and $m_{\eta_{\rm ss}}$. We observe that the magnitude of error reaches a plateau after 100–200 MD time depending on the quark mass. Since similar binsize dependences are found for other particle types, we employ a binsize of 250 MD time for the jackknife analysis at $0.13770 \geq \kappa_{ud} \geq 0.13754$. At our lightest point $\kappa_{ud} = 0.13781$ with the statistics of 990 MD units, we had to reduce the bin size to 110 MD units.

We define the bare quark mass based on the axial vector Ward-Takahashi identity (AWI) by the ratio of matrix elements of the pseudoscalar density P and the fourth component of the axial vector current A_4:

$$\bar{m}_f^{\rm AWI} + \bar{m}_g^{\rm AWI} = \frac{\langle 0|\nabla_4 A_4^{\rm imp}|{\rm PS}\rangle}{\langle 0|P|{\rm PS}\rangle}, \tag{22}$$

where $|{\rm PS}\rangle$ denotes the pseudoscalar meson state at rest and f and g ($f, g = ud, s$) label the flavors of the valence quarks. We employ the nonperturbatively $O(a)$-improved axial vector current $A_4^{\rm imp} = A_4 + c_A \bar{\nabla}_4 P$ with $\bar{\nabla}_4$ the sym-

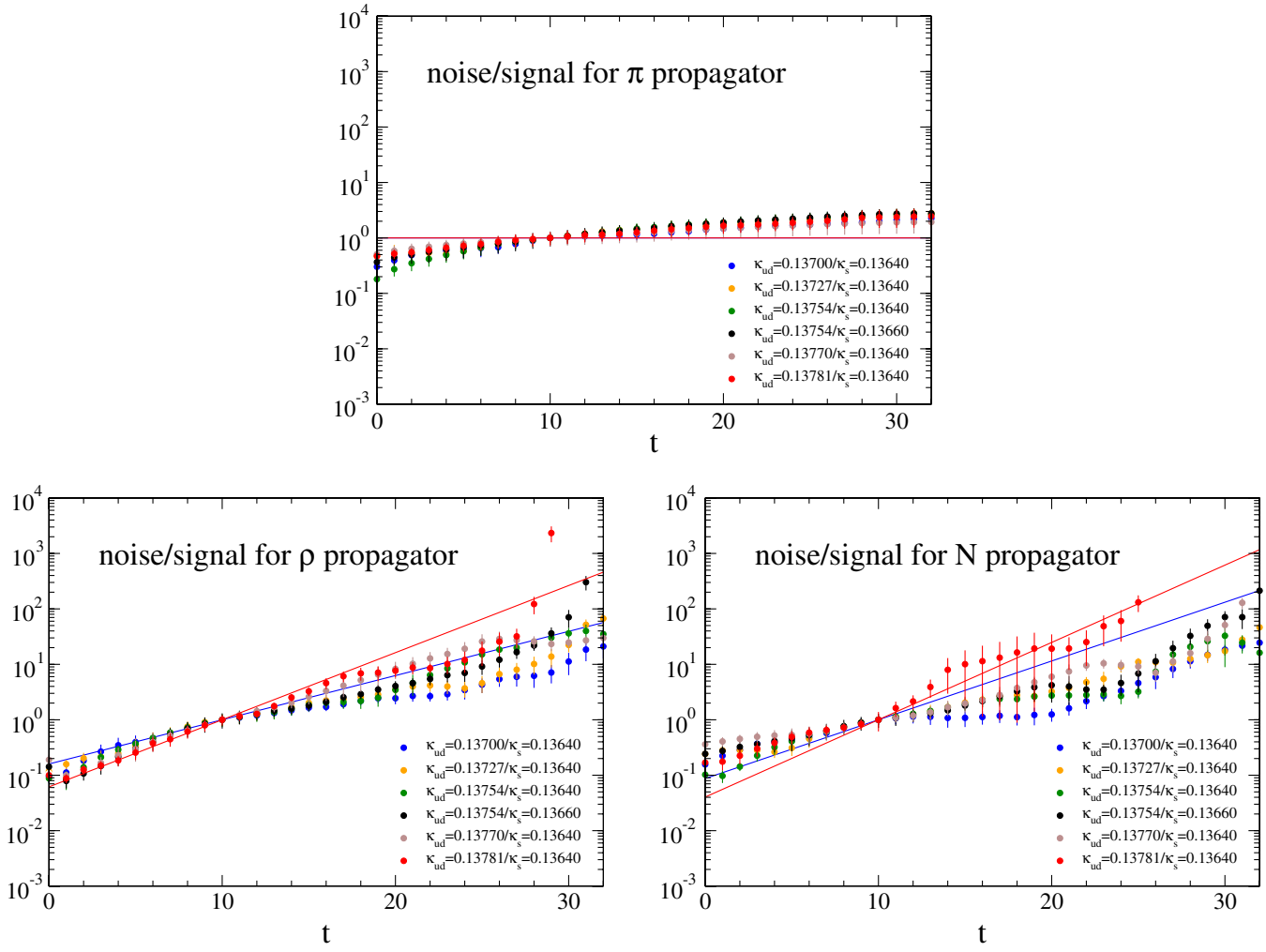


FIG. 7 (color online). Noise-to-signal ratio for the π, ρ and nucleon propagators. The data are normalized by the value at $t = 10$. See text for red and bule lines.

metric lattice derivative, and $c_A = -0.03876106$ as determined in Ref. [33]. In practice the AWI quark mass is determined by

$$\bar{m}_f^{\mathrm{AWI}} + \bar{m}_g^{\mathrm{AWI}} = m_{\mathrm{PS}} \left| \frac{C_A^s}{C_P^s} \right|, \tag{23}$$

where m_{PS}, C_A^s and C_P^s are obtained by applying a simultaneous χ^2 fit to

$$\langle A_4^{\mathrm{imp}}(t) P^s(0) \rangle = 2C_A^s \frac{\sinh(-m_{\mathrm{PS}}(t - T/2))}{\exp(m_{\mathrm{PS}} T/2)} \tag{24}$$

with a smeared source and

$$\langle P(t) P^s(0) \rangle = 2C_P^s \frac{\cosh(-m_{\mathrm{PS}}(t - T/2))}{\exp(m_{\mathrm{PS}} T/2)} \tag{25}$$

with a smeared source, where T denotes the temporal extent of the lattice. We employ the fit range of $[t_{\mathrm{min}}, t_{\mathrm{max}}] = [13, 25]$ for the former and [13, 30] for the latter at all the hopping parameters. The renormalized quark mass in the continuum $\overline{\mathrm{MS}}$ scheme is defined as

$$m_f^{\overline{\mathrm{MS}}} = \frac{Z_A}{Z_P} m_f^{\mathrm{AWI}}, \tag{26}$$

with

$$m_f^{\mathrm{AWI}} = \frac{(1 + b_A \frac{m_f^{\mathrm{VWI}}}{u_0})}{(1 + b_P \frac{m_f^{\mathrm{VWI}}}{u_0})} \bar{m}_f^{\mathrm{AWI}}. \tag{27}$$

The renormalization factors $Z_{A,P}$ and the improvement coefficients $b_{A,P}$ are perturbatively evaluated up to one-loop level [34–36] with the tadpole improvement. The VWI quark masses in the ma corrections are perturbatively obtained from the AWI quark masses:

$$\frac{m_f^{\mathrm{VWI}}}{u_0} = \frac{Z_A}{Z_P Z_m} \bar{m}_f^{\mathrm{AWI}}. \tag{28}$$

In Table IV we list the values of $m_{ud}^{\overline{\mathrm{MS}}}$ and $m_s^{\overline{\mathrm{MS}}}$ renormal-

TABLE III. Meson and baryon masses in lattice units at each combination of κ_{ud} and κ_s. χ^2/dof for the fit is also presented in the second row of each channel. The fit range is [13,30] for pseudoscalar mesons, [10,20] for vector mesons, [10,20] for octet baryons, and [8,20] for decuplet baryons.

κ_{ud}	0.13700	0.13727	0.13754	0.13754	0.13770	0.13781
κ_s	0.13640	0.13640	0.13640	0.13660	0.13640	0.13640
π	0.32242(65)	0.26191(73)	0.18903(79)	0.17671(129)	0.13593(140)	0.07162(299)
	0.015	0.008	0.002	0.001	0.001	0.004
K	0.36269(61)	0.32785(74)	0.29190(67)	0.26729(110)	0.27282(103)	0.25454(97)
	0.016	0.015	0.002	0.001	0.001	0.025
η_{ss}	0.39947(58)	0.38380(74)	0.36870(71)	0.33490(93)	0.36289(103)	0.35306(82)
	0.017	0.015	0.000	0.002	0.001	0.016
ρ	0.5060(30)	0.4566(36)	0.4108(31)	0.3963(53)	0.3895(94)	0.3503(315)
	0.043	0.229	0.017	0.090	0.005	0.418
K^*	0.5314(23)	0.4954(32)	0.4665(23)	0.4428(37)	0.4525(35)	0.4316(47)
	0.088	0.068	0.007	0.014	0.003	0.092
ϕ	0.5560(17)	0.5325(28)	0.5156(21)	0.4849(26)	0.5105(26)	0.4949(15)
	0.124	0.015	0.002	0.007	0.001	0.026
N	0.7277(22)	0.6487(56)	0.5584(53)	0.5331(71)	0.5025(87)	0.4285(360)
	0.077	0.027	0.358	0.014	0.171	1.138
Λ	0.7557(23)	0.6913(45)	0.6208(36)	0.5857(42)	0.5764(65)	0.5240(95)
	0.115	0.029	0.089	0.015	0.018	0.154
Σ	0.7606(20)	0.7039(51)	0.6437(39)	0.6052(48)	0.6044(71)	0.5601(99)
	0.072	0.030	0.041	0.091	0.043	1.377
Ξ	0.7859(25)	0.7399(43)	0.6910(30)	0.6474(32)	0.6655(46)	0.6405(31)
	0.139	0.035	0.028	0.020	0.010	0.008
Δ	0.8290(42)	0.7694(84)	0.6956(66)	0.6731(86)	0.6438(90)	0.5798(378)
	0.046	0.022	0.102	0.038	0.860	0.421
Σ^*	0.8537(35)	0.8039(74)	0.7464(43)	0.7149(74)	0.7097(67)	0.6885(140)
	0.037	0.010	0.022	0.036	0.179	0.031
Ξ^*	0.8788(30)	0.8395(67)	0.7964(41)	0.7579(60)	0.7740(58)	0.7549(67)
	0.036	0.005	0.005	0.035	0.022	0.255
Ω	0.9038(29)	0.8754(61)	0.8456(37)	0.8001(49)	0.8342(52)	0.8142(34)
	0.050	0.003	0.009	0.032	0.015	0.206

ized at the scale of $1/a$, whose statistical errors are provided by the jackknife analysis with the bin size chosen as in the hadron mass measurements.

The bare pseudoscalar meson decay constant on the lattice is defined by

$$|\langle 0|A_4^{\mathrm{imp}}|\mathrm{PS}\rangle| = f_{\mathrm{PS}}^{\mathrm{bare}} m_{\mathrm{PS}}. \tag{29}$$

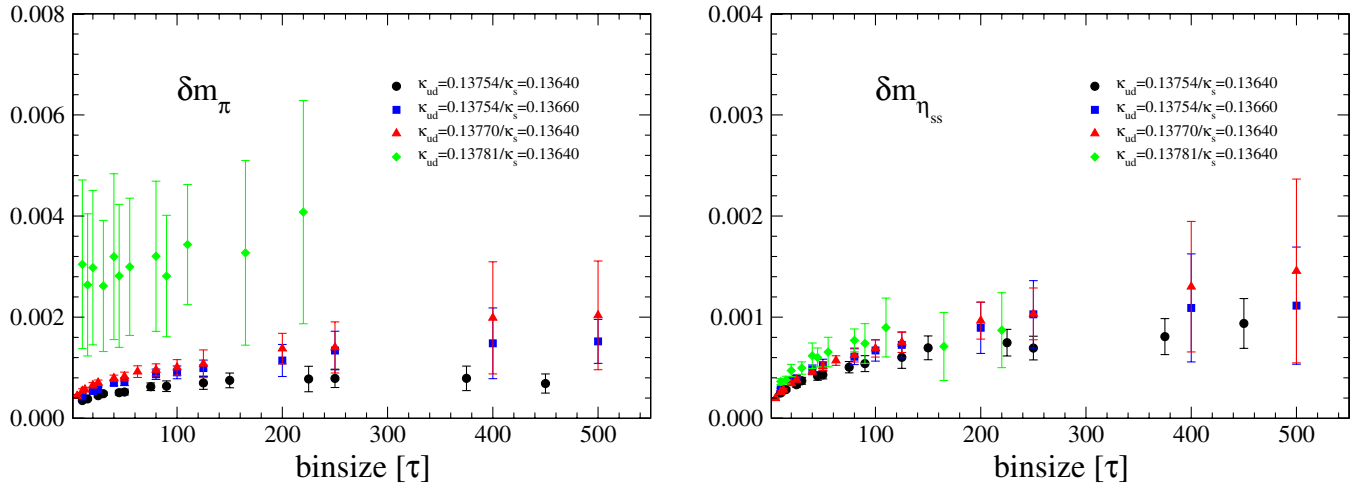


FIG. 8 (color online). Binsize dependence of the magnitude of error for m_π (left) and $m_{\eta_{ss}}$ (right) at $\kappa_{ud} \geq 0.13754$.

TABLE IV. Quark masses in the $\overline{\rm MS}$ scheme at the scale of $1/a$ and pseudoscalar decay constants at each combination of κ_{ud} and κ_s. Both are renormalized at one-loop level. The values for $m_\pi^2/m_{ud}^{\rm AWI}$ and f_K/f_π are also listed.

κ_{ud}	0.13700	0.13727	0.13754	0.13754	0.13770	0.13781
κ_s	0.13640	0.13640	0.13640	0.13660	0.13640	0.13640
$am_{ud}^{\overline{\rm MS}}$	0.030753(110)	0.020834(66)	0.011028(80)	0.009666(105)	0.005644(120)	0.001609(118)
$am_s^{\overline{\rm MS}}$	0.047142(110)	0.044674(72)	0.042355(79)	0.035571(98)	0.041285(94)	0.039913(62)
m_s/m_{ud}	1.5329(20)	2.1443(39)	3.841(21)	3.680(30)	7.32(14)	24.8(1.8)
af_π	0.0898(12)	0.0853(18)	0.07481(51)	0.07262(60)	0.06973(78)	0.0656(35)
af_K	0.0942(13)	0.0916(15)	0.08432(56)	0.08058(40)	0.08089(57)	0.0777(13)
$am_\pi^2/m_{ud}^{\rm AWI}$	3.708(13)	3.610(14)	3.558(16)	3.542(25)	3.585(29)	3.732(73)
f_K/f_π	1.0485(13)	1.0739(57)	1.1271(16)	1.1095(64)	1.1601(73)	1.186(48)

with $|{\rm PS}\rangle$ the pseudoscalar meson state at rest consisting of f and g valence quarks. We evaluate $f_{\rm PS}^{\rm bare}$ from the formula

$$f_{\rm PS}^{\rm bare} = \left|\frac{C_A^s}{C_P^s}\right|\sqrt{\frac{2|C_P^l|}{m_{\rm PS}}}, \tag{30}$$

where we extract $m_{\rm PS}$, C_A^s, C_P^s and C_P^l from a simultaneous fit of Eqs. (24) and (25) and

$$\langle P(t)P^l(0)\rangle = 2C_P^l\frac{\cosh(-m_{\rm PS}(t-T/2))}{\exp(m_{\rm PS}T/2)} \tag{31}$$

with a local source. The fit ranges are [13, 25], [13, 30] and [15, 25], respectively, at all the hopping parameters. The bare decay constant $f_{\rm PS}^{\rm bare}$ is renormalized perturbatively with

$$f_{\rm PS} = 2\kappa u_0 Z_A\left(1 + b_A\frac{m_f^{\rm VWI} + m_g^{\rm VWI}}{2u_0}\right)f_{\rm PS}^{\rm bare}, \tag{32}$$

where $m_f^{\rm VWI}$ is estimated by Eq. (28). Table IV summarizes the results for $f_{\rm PS}$ with the statistical errors evaluated by the jackknife analysis with the bin size chosen as in the hadron mass measurements.

TABLE V. PACS-CS and CP-PACS/JLQCD results for the hadron masses at $(\kappa_{ud}, \kappa_s) = (0.13700, 0.13640)$. $[t_{\rm min}, t_{\rm max}]$ denotes the fitting range.

	lattice size	m_π	m_ρ	$m_{\rm N}$
PACS-CS	$32^3 \times 64$	0.32242(65)	0.5060(30)	0.7277(22)
$[t_{\rm min}, t_{\rm max}]$		[13,30]	[10,20]	[10,20]
CP-PACS/JLQCD	$20^3 \times 40$	0.32247(74)	0.5157(21)	0.7337(28)
$[t_{\rm min}, t_{\rm max}]$		[9,17]	[9,15]	[9,15]

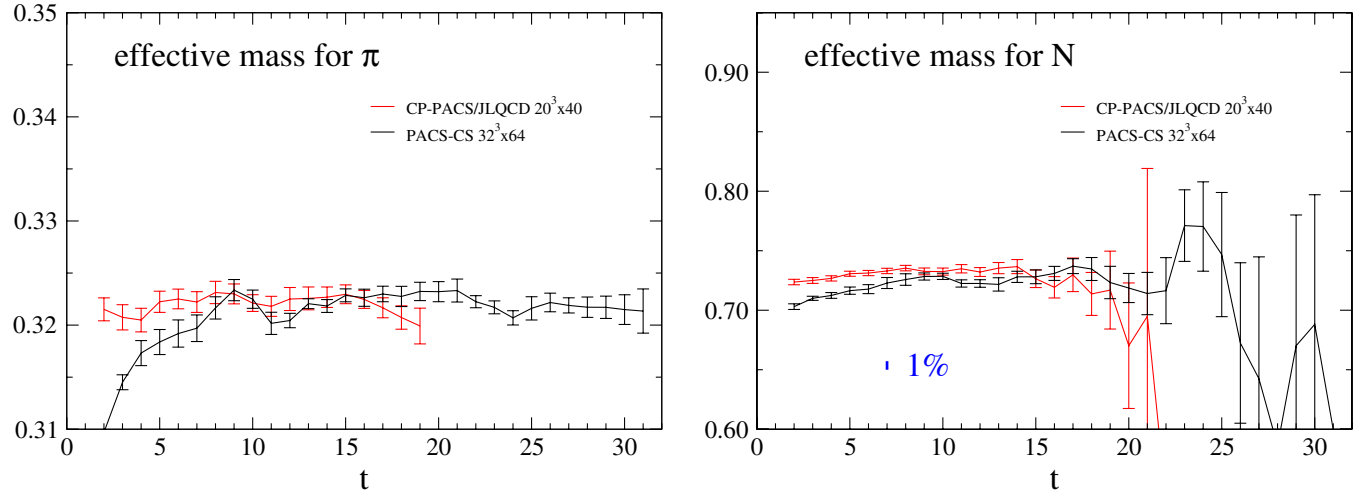


FIG. 9 (color online). Effective masses for the π (left) and the nucleon (right) at $(\kappa_{ud}, \kappa_s) = (0.13700, 0.13640)$. Black and red symbols denote the PACS-CS and the CP-PACS/JLQCD results, respectively.

B. Comparison with the previous CP-PACS/JLQCD results

A comparison between the present PACS-CS results and those with the previous CP-PACS/JLQCD work [6,7] obtained with the same gauge and quark actions is possible at $(\kappa_{ud}, \kappa_s) = (0.13700, 0.13640)$, except that the lattice sizes are different: $32^3 \times 64$ for the former and $20^3 \times 40$ for the latter. In Table V we list the PACS-CS and the CP-PACS/JLQCD results for the hadron masses at $(\kappa_{ud}, \kappa_s) = (0.13700, 0.13640)$. While the results for m_π are consistent within the errors, we find a 1–2% deviation for m_ρ and m_N. This is pictorially confirmed in Fig. 9 which shows the effective masses for the π meson and the nucleon. The pion effective masses are almost degenerate from $t = 8$ to 17, while a slight discrepancy is observed for the nucleon results. The ρ and nucleon masses may be suffering from finite-size effects.

IV. CHIRAL ANALYSIS ON PSEUDOSCALAR MESON MASSES AND DECAY CONSTANTS

The analysis of chiral behavior of pseudoscalar meson masses and decay constants occupy an important place in lattice QCD. Theoretically the main points to examine are the presence of chiral logarithms as predicted by ChPT and the convergence of the ChPT series itself. The viability of ChPT is relevant also for studies of finite-size effects. The low energy constants are important from phenomenological points of view. And finally, the chiral analysis is required to pin down the physical point in the parameter space of the simulations. We begin with a discussion of a subtle point in the chiral analysis when Wilson-clover quark action with implicit chiral symmetry breaking is employed.

A. Chiral perturbation theory for $O(a)$-improved Wilson-type quark action

Our simulations are carried out with a non-perturbatively $O(a)$-improved Wilson quark action. At present we correct $O(ma)$ terms in the AWI quark masses and decay constants by one-loop perturbation theory. These corrections are expected to be very small in magnitude, and hence leading scaling violations in meson masses and decay constants from our simulations can be taken as $O(a^2)$. In this case, the NLO formula of Wilson chiral perturbation theory for the SU(3) flavor case [37], which incorporates the leading contributions of the implicit chiral symmetry breaking effects of the Wilson-type quarks, are given by

$$\frac{m_\pi^2}{2m_{ud}} = B_0\Big\{1 + \mu_\pi - \frac{1}{3}\mu_\eta + \frac{2B_0}{f_0^2}(16m_{ud}(2L_8 - L_5) + 16(2m_{ud} + m_s)(2L_6 - L_4)) - \frac{2H''}{f_0^2}\Big\}, \tag{33}$$

$$\frac{m_K^2}{(m_{ud} + m_s)} = B_0\Big\{1 + \frac{2}{3}\mu_\eta + \frac{2B_0}{f_0^2}(8(m_{ud} + m_s)(2L_8 - L_5) + 16(2m_{ud} + m_s)(2L_6 - L_4)) - \frac{2H''}{f_0^2}\Big\}, \tag{34}$$

$$f_\pi = f_0\Big\{1 - 2\mu_\pi - \mu_K + \frac{2B_0}{f_0^2}(8m_{ud}L_5 + 8(2m_{ud} + m_s)L_4) - \frac{2H'}{f_0^2}\Big\}, \tag{35}$$

$$f_K = f_0\Big\{1 - \frac{3}{4}\mu_\pi - \frac{3}{2}\mu_K - \frac{3}{4}\mu_\eta + \frac{2B_0}{f_0^2}(4(m_{ud} + m_s)L_5 + 8(2m_{ud} + m_s)L_4) - \frac{2H'}{f_0^2}\Big\}, \tag{36}$$

where the quark masses are defined by the axial-vector Ward-Takahashi identities: $m_{ud} = m_{ud}^{\mathrm{AWI}}$ and $m_s = m_s^{\mathrm{AWI}}$. $L_{4,5,6,8}$ are the low energy constants and μ_{PS} is the chiral logarithm defined by

$$\mu_{\mathrm{PS}} = \frac{1}{16\pi^2}\frac{\tilde{m}_{\mathrm{PS}}^2}{f_0^2}\ln\Big(\frac{\tilde{m}_{\mathrm{PS}}^2}{\mu^2}\Big), \tag{37}$$

where

$$\tilde{m}_\pi^2 = 2m_{ud}B_0, \tag{38}$$

$$\tilde{m}_K^2 = (m_{ud} + m_s)B_0, \tag{39}$$

$$\tilde{m}_\eta^2 = \frac{2}{3}(m_{ud} + 2m_s)B_0 \tag{40}$$

with μ the renormalization scale.

The two additional parameters H'' and H' are associated with the $O(a^2)$ contributions distinguishing the Wilson ChPT from that in the continuum. Since these parameters are independent of the quark masses, their contributions can be absorbed into B_0 and f_0 by the following redefinitions:

$$B_0' = B_0\Big(1 - \frac{2H''}{f_0^2}\Big), \tag{41}$$

$$f_0' = f_0\Big(1 - \frac{2H'}{f_0^2}\Big), \tag{42}$$

Indeed reexpansion of the terms in the curly brackets of (33) to (36) gives rise only to terms of form $O(m_q \cdot a^2)$ and $O(m_q \ln m_q \cdot a^2)$, which are NNLO in the order counting of WChPT analysis and hence can be ignored. Thus, up to

NLO, WChPT formula are equivalent to the continuum form. Note that the expressions in terms of the VWI quark masses take different forms and cannot be reduced to those of the continuum ChPT. Hereafter we concentrate on the continuum ChPT.

B. SU(3) chiral perturbation theory

The SU(3) ChPT formula in the continuum up to NLO [38] is given by

$$\frac{m_\pi^2}{2m_{ud}} = B_0\Big\{1 + \mu_\pi - \frac{1}{3}\mu_\eta + \frac{2B_0}{f_0^2}(16m_{ud}(2L_8 - L_5) + 16(2m_{ud} + m_s)(2L_6 - L_4))\Big\}, \quad (43)$$

$$\frac{m_K^2}{(m_{ud} + m_s)} = B_0\Big\{1 + \frac{2}{3}\mu_\eta + \frac{2B_0}{f_0^2}(8(m_{ud} + m_s)(2L_8 - L_5) + 16(2m_{ud} + m_s)(2L_6 - L_4))\Big\}, \quad (44)$$

$$f_\pi = f_0\Big\{1–2\mu_\pi - \mu_K + \frac{2B_0}{f_0^2}(8m_{ud}L_5 + 8(2m_{ud} + m_s)L_4)\Big\}, \quad (45)$$

$$f_K = f_0\Big\{1 - \frac{3}{4}\mu_\pi - \frac{3}{2}\mu_K - \frac{3}{4}\mu_\eta + \frac{2B_0}{f_0^2}(4(m_{ud} + m_s)L_5 + 8(2m_{ud} + m_s)L_4)\Big\}, \quad (46)$$

There are six unknown low energy constants B_0, f_0, $L_{4,5,6,8}$ in the expressions above. $L_{4,5,6,8}$ are scale-dependent so as to cancel that of the chiral logarithm given by (37). We can determine these parameters by applying a simultaneous fit to $m_\pi^2/(2m_{ud})$, $m_K^2/(m_{ud} + m_s)$, f_π and f_K.

In order to provide an overview of our data we plot in Fig. 10, a comparison of the PACS-CS (red symbols) and the CP-PACS/JLQCD results (black symbols) for $m_\pi^2/m_{ud}^{\rm AWI}$ and f_K/f_π as a function of $m_{ud}^{\rm AWI}$. The two data sets show a smooth connection at $\kappa_{ud} = 0.13700$ ($m_{ud}^{\rm AWI} = 0.028$). More important is the fact that an almost linear quark mass dependence of the CP-PACS/JLQCD results in heavier quark mass region changes into a convex behavior, both for $m_\pi^2/m_{ud}^{\rm AWI}$ and f_K/f_π, as $m_{ud}^{\rm AWI}$ is diminished in the PACS-CS results. This is a characteristic feature expected from the ChPT prediction in the small quark mass region due to the chiral logarithm. This curvature drives up the ratio f_K/f_π toward the experimental value as the physical point is approached.

Having confirmed signals for the presence of the chiral logarithm, we apply the SU(3) ChPT formulas (43)–(46) to our results. We choose the four simulation points at $\kappa_{ud} \geq 0.13754$. In Fig. 10 these four points lie to the left and around the turning point of the curvature. They also correspond to the region where the ρ meson mass satisfies the condition $m_\rho > 2m_\pi$, and hence lie to the left of the threshold singularity in the complex energy plane for the ρ meson. The heaviest pion mass at $(\kappa_{ud}, \kappa_s) = (0.13754, 0.13640)$ is about 410 MeV with the use of $a^{-1} = 2.176(31)$ GeV determined below. The measured bare AWI quark masses, but corrected for the $O(ma)$ corrections at one-loop perturbation theory, are used for m_{ud} and m_s in Eqs. (43)–(46).

We present the fit results for the low energy constants in Table VI. The results are quoted both without (w/o FSE) and with (w/ FSE) finite-size corrections in the ChPT

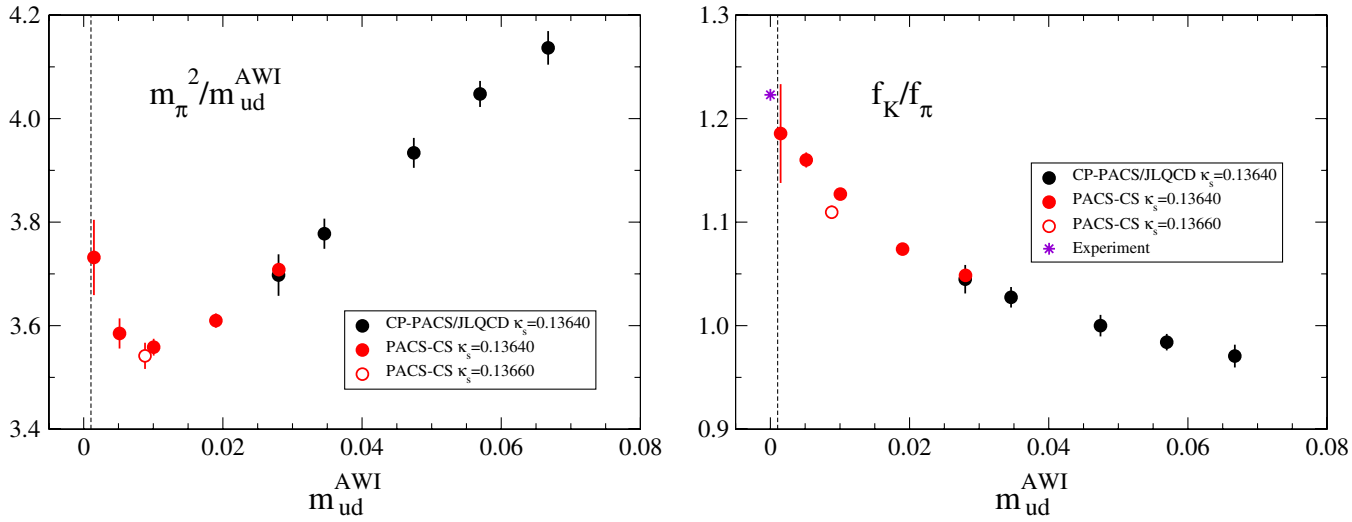


FIG. 10 (color online). Comparison of the PACS-CS (red) and the CP-PACS/JLQCD (black) results for $m_\pi^2/m_{ud}^{\rm AWI}$ (left) and f_K/f_π (right) as a function of $m_{ud}^{\rm AWI}$. Vertical line denotes the physical point and star symbol represents the experimental value.

TABLE VI. Results for the low energy constants in the SU(3) ChPT together with the phenomenological estimates and the RBC/UKQCD and MILC results. f_0 is perturbatively renormalized at one-loop level. $L_{4,5,6,8}$ are in units of 10^{-3} at the scale of 770 MeV. $\langle \bar{u}u \rangle$ and $\langle \bar{u}u \rangle_0$ are renormalized in the $\overline{\text{MS}}$ scheme at 2 GeV.

	PACS-CS		phenomenology [39,40]	RBC/UKQCD [41]	MILC [42]
	w/o FSE	w/ FSE			
aB_0	1.789(34)	1.778(34)	-	2.35(16)	-
af_0	0.0534(38)	0.0546(39)	-	0.0541(40)	-
f_0 [GeV]	0.1160(88)	0.1185(90)	0.115	0.0935(73)	-
f_π/f_0	1.159(57)	1.145(56)	1.139	1.33(7)	$1.21(5)(^{+13}_{-3})$
$m_{ud}^{\text{ph}}B_0$ [GeV^2]	0.00859(10)	0.00859(11)	0.0181	-	-
$m_s^{\text{ph}}B_0$ [GeV^2]	0.2550(36)	0.2534(36)	0.434	-	-
$a^3\langle \bar{u}u \rangle_0$	$-(0.132(6))^3$	$-(0.133(6))^3$	-	-	-
$\langle \bar{u}u \rangle_0$ [GeV^3]	$-(0.286(15))^3$	$-(0.290(15))^3$	-	-	$-(0.242(9)(^{+5}_{-17})(4))^3$
L_4	−0.04(10)	−0.06(10)	0.00(80)	0.139(80)	$0.1(3)(^{+3}_{-1})$
L_5	1.43(7)	1.45(7)	1.46(10)	0.872(99)	$1.4(2)(^{+2}_{-1})$
$2L_6 - L_4$	0.10(2)	0.10(2)	0.0(1.0)	−0.001(42)	$0.3(1)(^{+2}_{-3})$
$2L_8 - L_5$	−0.21(3)	−0.21(3)	0.54(43)	0.243(45)	0.3(1)(1)
χ^2/dof	4.2(2.7)	4.4(2.8)	-	0.7	-

formulas (see Sec. IV D). We also list the phenomenological estimates with the experimental inputs [39,40], and the results obtained by recent 2 + 1 flavor lattice QCD calculations [41,42]. The renormalization scale is set to be 770 MeV. The MILC results for the low energy constants quoted at the scale of m_η are converted according to Ref. [38]

$$L_4(\mu) = L_4(m_\eta) - \frac{1}{256\pi^2}\ln\left(\frac{\mu^2}{m_\eta^2}\right), \tag{47}$$

$$L_5(\mu) = L_5(m_\eta) - \frac{3}{256\pi^2}\ln\left(\frac{\mu^2}{m_\eta^2}\right), \tag{48}$$

$$(2L_6 - L_4)(\mu) = (2L_6 - L_4)(m_\eta) - \left(\frac{2}{9}\right)\frac{1}{256\pi^2}\ln\left(\frac{\mu^2}{m_\eta^2}\right), \tag{49}$$

$$(2L_8 - L_5)(\mu) = (2L_8 - L_5)(m_\eta) + \left(\frac{4}{3}\right)\frac{1}{256\pi^2}\ln\left(\frac{\mu^2}{m_\eta^2}\right) \tag{50}$$

with μ the renormalization scale. For L_4 and L_5 governing the behavior of f_π and f_K, we find that all the results are compatible. On the other hand, some discrepancies are observed for the results of $2L_6 - L_4$ and $2L_8 - L_5$ contained in the ChPT formulas for m_π^2 and m_K^2.

For later convenience we convert the SU(3) low energy constants B_0, f_0, $L_{4,5,6,8}$ to the SU(2) low energy constants B, f, $l_{3,4}$ defined by

$$\frac{m_\pi^2}{2m_{ud}} = B\left\{1 + \mu_\pi(B_0 \to B, f_0 \to f) + 4\frac{\bar{m}_\pi^2}{f^2}l_3\right\}, \tag{51}$$

$$f_\pi = f\left\{1 - 2\mu_\pi(B_0 \to B, f_0 \to f) + 2\frac{\bar{m}_\pi^2}{f^2}l_4\right\} \tag{52}$$

with $\bar{m}_\pi^2 = m_{ud}B$. The NLO relations are given by [38]

$$B = B_0\left(1 - \frac{1}{3}\bar{\mu}_\eta + \frac{32\bar{m}_K^2}{f_0^2}(2L_6 - L_4)\right), \tag{53}$$

$$f = f_0\left(1 - \bar{\mu}_K + \frac{16\bar{m}_K^2}{f_0^2}L_4\right), \tag{54}$$

$$l_3 = -8L_4 - 4L_5 + 16L_6 + 8L_8 - \frac{1}{18}\bar{\nu}_\eta, \tag{55}$$

$$l_4 = 8L_4 + 4L_5 - \frac{1}{2}\bar{\nu}_K, \tag{56}$$

where $\bar{\mu}_{K,\eta}$ and $\bar{\nu}_{K,\eta}$ are defined by

$$\bar{\mu}_{K,\eta} = \frac{\bar{m}_{K,\eta}^2}{16\pi^2 f_0^2}\ln\left(\frac{\bar{m}_{K,\eta}^2}{\mu^2}\right), \tag{57}$$

$$\bar{\nu}_{K,\eta} = \frac{1}{32\pi^2}\left(\ln\left(\frac{\bar{m}_{K,\eta}^2}{\mu^2}\right) + 1\right) \tag{58}$$

TABLE VII. Results for the low energy constants in the SU(2) ChPT obtained by the conversion from those in the SU(3) ChPT. The RBC/UKQCD and MILC results are also given for comparison. f and f_0 are perturbatively renormalized at one-loop level. $\langle \bar{u}u \rangle$ and $\langle \bar{u}u \rangle_0$ are renormalized in the $\overline{\text{MS}}$ scheme at 2 GeV.

	PACS-CS		RBC/UKQCD [41]	MILC [42]
	w/o FSE	w/ FSE		
aB	1.950(31)	1.935(30)	2.457(78)	-
af	0.0582(19)	0.0588(19)	0.0661(18)	-
f [GeV]	0.1263(51)	0.1277(51)	0.1143(36)	-
f_π/f	1.065(8)	1.062(8)	-	$1.052(2)(^{+6}_{-3})$
$m_{ud}^{\text{ph}}B$ [GeV^2]	0.009364(36)	0.009352(34)	-	-
$m_s^{\text{ph}}B$ [GeV^2]	0.2780(52)	0.2758(49)	-	-
$a^3\langle \bar{u}u \rangle$	$-(0.143(3))^3$	$-(0.144(3))^3$	-	-
$\langle \bar{u}u \rangle$ [GeV^3]	$-(0.310(9))^3$	$-(0.312(10))^3$	-	$-(0.278(1)(^{+2}_{-3})(5))^3$
$\bar{l}_3$	3.50(11)	3.47(11)	2.87(28)	$1.2(6)(^{+1.0}_{-1.5})$
$\bar{l}_4$	4.22(10)	4.21(11)	4.10(5)	$4.4(4)(^{+4}_{-1})$
B/B_0	1.090(15)	1.089(15)	-	-
f/f_0	1.089(45)	1.078(44)	-	$1.15(5)(^{+13}_{-3})$
$\langle \bar{u}u \rangle/\langle \bar{u}u \rangle_0$	1.268(10)	1.245(10)	-	$1.52(17)(^{+38}_{-15})$

with

$$\bar{m}_K^2 = m_s B_0, \tag{59}$$

$$\bar{m}_\eta^2 = \frac{4}{3} m_s B_0, \tag{60}$$

and $\bar{l}_i$ ($i = 3, 4$) are defined at the renormalization scale $\mu = m_\pi = 139.6$ MeV [43]:

$$l_i = \frac{\gamma_i}{32\pi^2}\left(\bar{l}_i + \ln\frac{m_\pi^2}{\mu^2}\right) \tag{61}$$

with

$$\gamma_3 = -\frac{1}{2}, \tag{62}$$

$$\gamma_4 = 2. \tag{63}$$

In Table VII we summarize the results for the SU(2) low energy constants obtained by the conversion from the

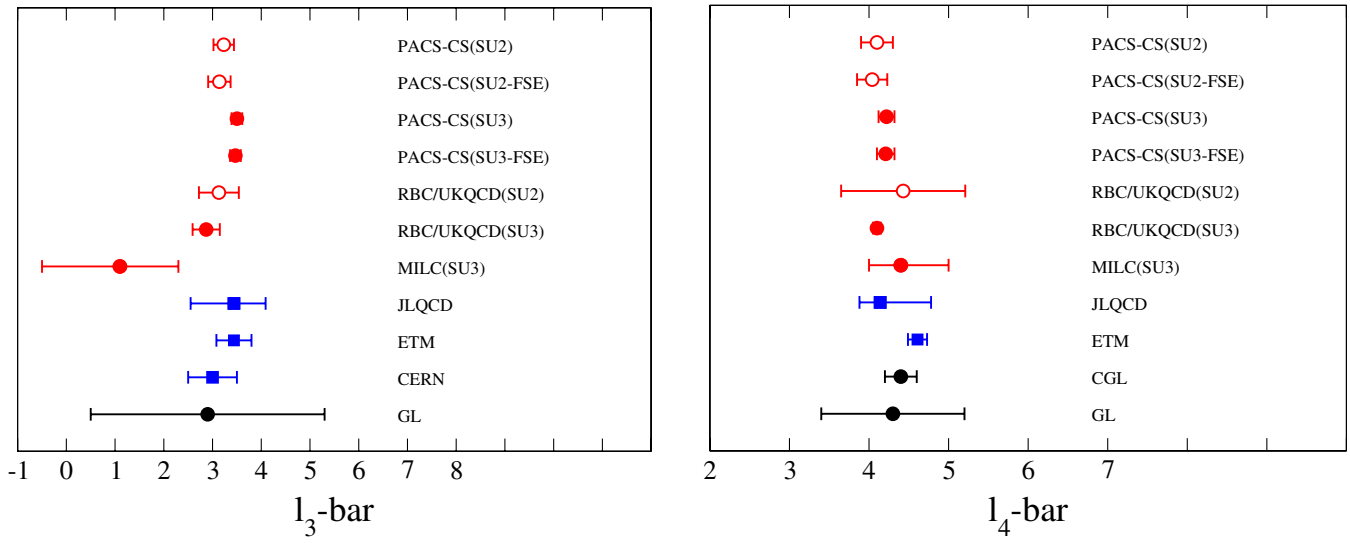


FIG. 11 (color online). Comparison of the results for $\bar{l}_3$ and $\bar{l}_4$. Black symbols denote the phenomenological estimates. Blue ones are for 2 flavor lattice results. Red closed (open) symbols represent the results for the SU(3) (SU(2)) ChPT analyses in the 2 + 1 flavor dynamical simulations. See text and Table VIII for details.

TABLE VIII. Comparison of $\bar{l}_{3,4}$.

group	#flavor	quark action	ChPT	$\bar{l}_3$	$\bar{l}_4$
this work	2 + 1	NP clover	SU(2) w/o FSE	3.23(21)	4.10(20)
			SU(2) w/ FSE	3.14(23)	4.04(19)
			SU(3) w/o FSE	3.50(11)	4.22(10)
			SU(3) w/ FSE	3.47(11)	4.21(11)
RBC/UKQCD [41]	2 + 1	DWF	SU(3)	2.87(28)	4.10(5)
			SU(2)	3.13(33)(24)	4.43(14)(77)
MILC [42]	2 + 1	KS	SU(3)	$1.1(6)(^{+1.0}_{-1.5})$	$4.4(4)(^{+4}_{-1})$
JLQCD [44]	2	Overlap	SU(2)	$3.44(57)(^{+0}_{-68})(^{+32}_{-0})$	$4.14(26)(^{+49}_{-0})(^{+32}_{-0})$
ETM [45]	2	TM	SU(2)	3.44(8)(35)	4.61(4)(11)
CERN [26]	2	Wilson + NP clover	SU(2)	3.0(5)(1)	-
CGL [46]	-	–	SU(2)	-	4.4(2)
GL [43]	-	–	SU(2)	2.9(2.4)	4.3(9)

SU(3) low energy constants. The vacuum condensations are defined by

$$\langle \bar{u}u\rangle_0 \equiv \langle \bar{u}u\rangle|_{m_{ud}=m_s=0} = -\frac{1}{2}f_0^2 B_0, \tag{64}$$

$$\langle \bar{u}u\rangle \equiv \langle \bar{u}u\rangle|_{m_{ud}=0,m_s=m_s^{\rm physical}} = -\frac{1}{2}f^2 B. \tag{65}$$

These quantities are perturbatively renormalized at the scale of 2 GeV.

In Fig. 11 we compare our results for $\bar{l}_{3,4}$ with those obtained by other groups whose numerical values are listed in Table VIII. Black symbols denote the phenomenological estimates, blue symbols represent the results obtained by the SU(2) ChPT fit on 2 flavor dynamical configurations and red closed (open) symbols are for those obtained by the SU(3) (SU(2)) ChPT fit on 2 + 1 flavor dynamical configurations. For $\bar{l}_3$ all the results reside between 3.0 and 3.5, except for the MILC result which is sizably smaller and marginally consistent with others within a large error. On the other hand, we find a good consistency among the results for $\bar{l}_4$.

We have found that the SU(3) ChPT fit gives reasonable values for the low energy constants. However, we are concerned with a rather large value of $\chi^2/\mathrm{dof} = 4.2(2.9)$ (see Table VI). Figures. 12 and 13 show how well the data for m_π^2/m_{ud}, $2m_K^2/(m_{ud}+m_s)$, f_π and f_K are described by the SU(3) ChPT up to NLO. The filled and open circles are our data, and the fit results are plotted by blue triangles. We note in passing that, for the Wilson-clover quark action, $m_s^{\rm AWI}$ varies at $O(a^2)$ as m_{ud} varies even if κ_s is held fixed. Thus we are not able to draw a line with a fixed value for $m_s^{\rm AWI}$. The blue star symbols represent the extrapolated values at the physical point whose determination will be explained below in Sec. V.

The points around $m_{ud}^{\rm AWI} \approx 0.01$ corresponds to $(\kappa_{ud}, \kappa_s) = (0.13754, 0.13640)$ and $(0.13754, 0.13660)$. Marked deviations between circles and triangles show

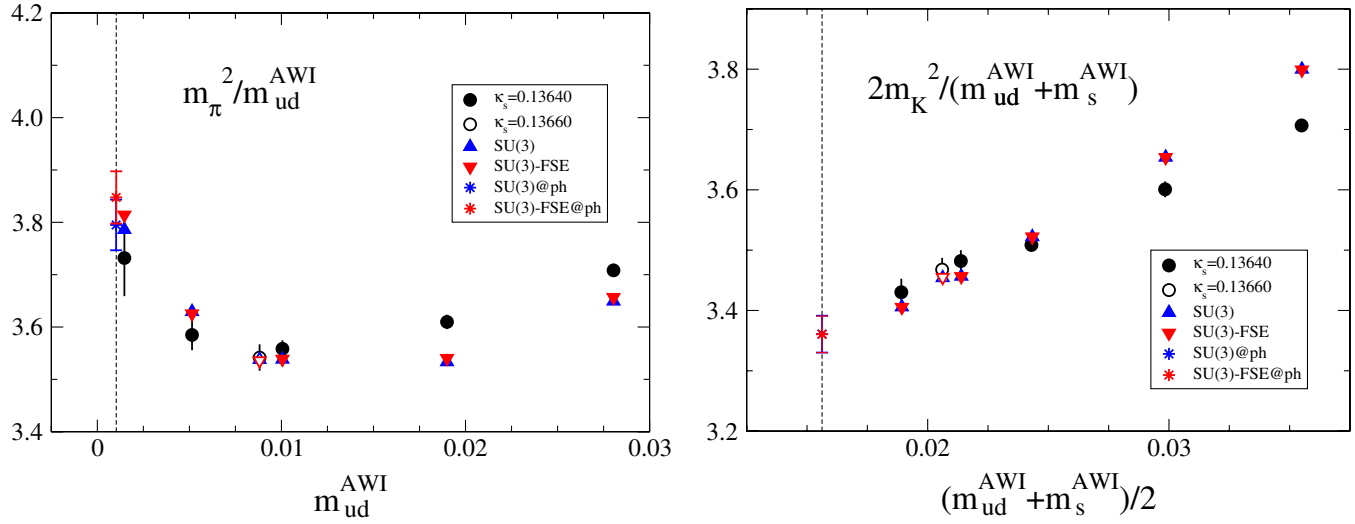


FIG. 12 (color online). SU(3) ChPT fit for $m_\pi^2/m_{ud}^{\rm AWI}$ (left) and $2m_K^2/(m_{ud}^{\rm AWI}+m_s^{\rm AWI})$ (right).

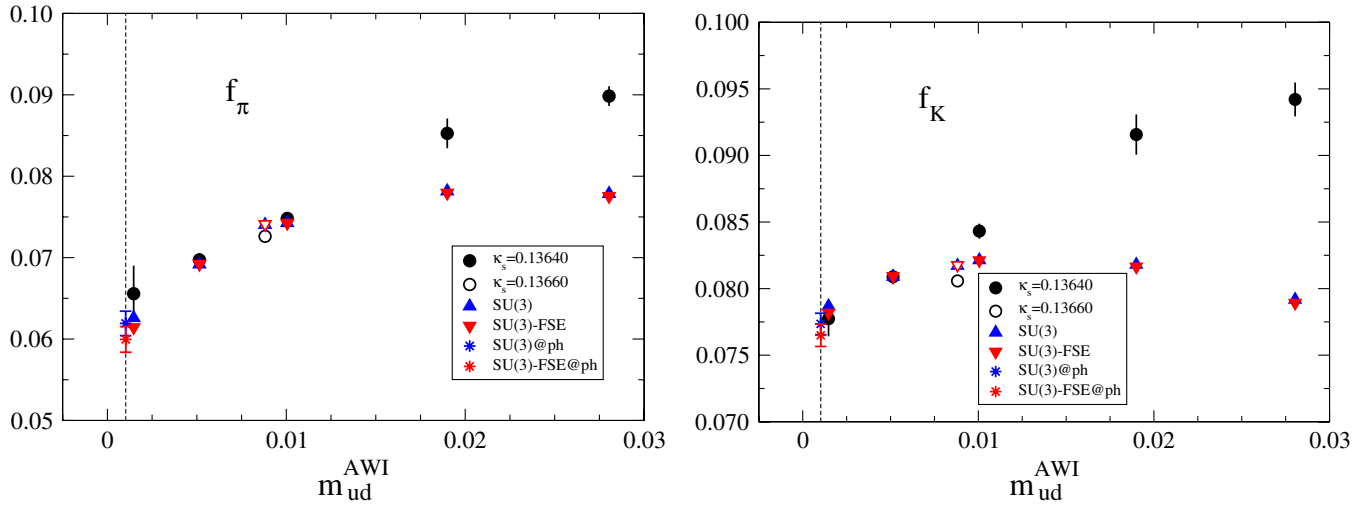


FIG. 13 (color online). SU(3) ChPT fit for f_π (left) and f_K (right).

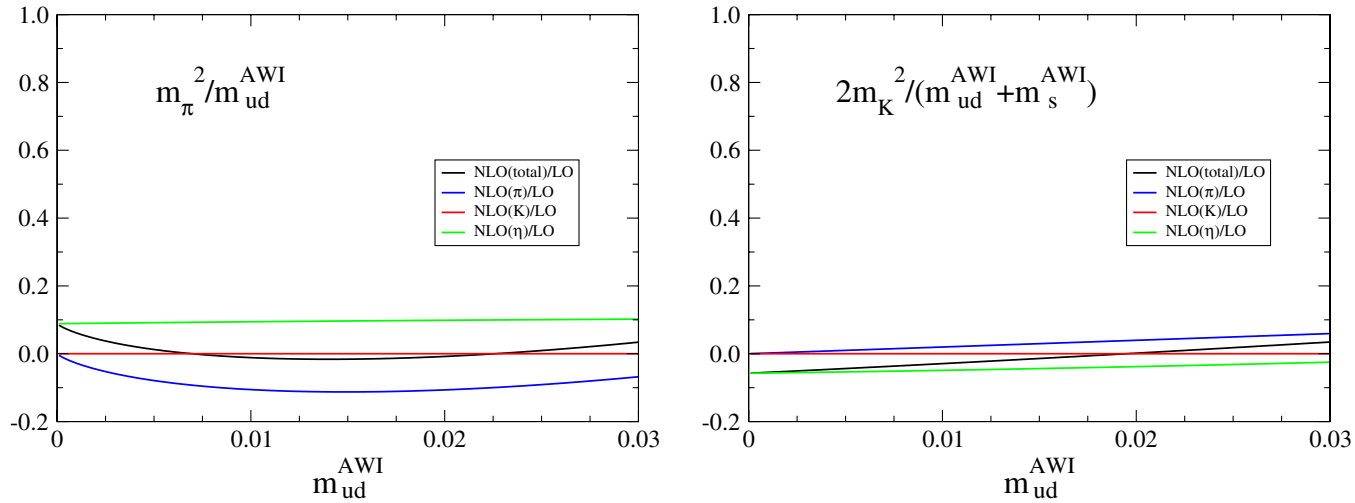


FIG. 14 (color online). Ratio of the NLO contribution to the LO one in the SU(3) ChPT fit for $m_\pi^2/m_{ud}^{\mathrm{AWI}}$ (left) and $2m_K^2/(m_{ud}^{\mathrm{AWI}} + m_s)$ (right).

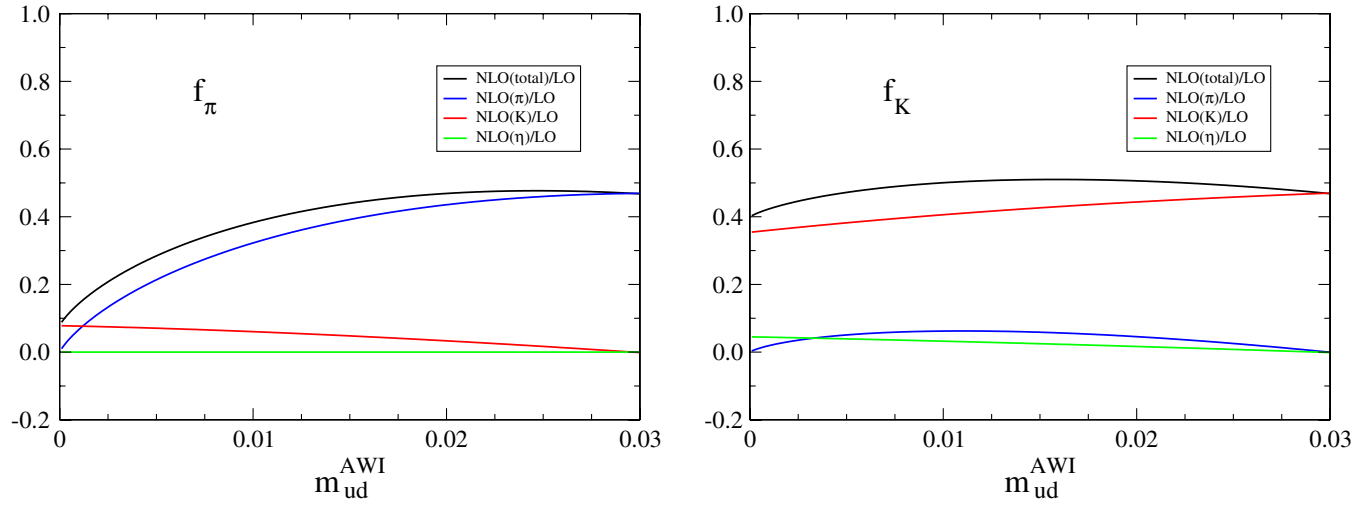


FIG. 15 (color online). Ratio of the NLO contribution to the LO one in the SU(3) ChPT fit for f_π (left) and f_K (right).

that the SU(3) ChPT poorly accounts for the strange quark mass dependence of f_π and f_K. This flaw is mainly responsible for the large value of χ^2/dof.

In order to investigate the origin of discrepancy between the data and the fit more closely, we draw the relative magnitude of the NLO contribution to the LO one for m_π^2/m_{ud}, $2m_K^2/(m_{ud}+m_s)$, f_π and f_K as a function of $m_{ud}^{\rm AWI}$ in Figs. 14 and 15. The strange quark mass is fixed at the physical value, and the contributions from π, K and η loops are separately drawn. The relative magnitudes are at most 10% for m_π^2/m_{ud} and $2m_K/(m_{ud}+m_s)$. We find, however, significant NLO contributions for the decay constants. For f_π the relative magnitude rapidly increases from 10% at $m_{ud}=0$ to 40% at around $m_{ud}=0.01$. The situation is worse for f_K for which the NLO contribution is about 40% of the LO one even at $m_{ud}=0$, most of which arises from the K loop.

C. SU(2) chiral perturbation theory

The bad convergences of the chiral expansions for f_π and f_K tell us that the strange quark mass is not light enough to be appropriately treated by the NLO SU(3) ChPT. There are two alternative choices for further chiral analysis. One is to extend SU(3) ChPT to NNLO, and the other is to use SU(2) ChPT with the aid of an analytic expansion for the strange quark contribution around the physical strange quark mass.

The former method, which has been employed by the MILC collaboration in an incomplete fashion [42], is very demanding: we cannot determine the additional low energy constants at NNLO without significantly increasing the data points. There is in addition no guarantee that the expansion is controlled at NNLO. We therefore consider that the latter route is more natural. This alternative was employed by the RBC/UKQCD collaboration [41]. Since they had data only at a single strange quark mass, they could not study the strange quark mass dependence. This we shall do with our data thanks to the second choice of the strange quark mass at $\kappa_{ud}=0.13754$.

For m_π and f_π the SU(2) ChPT formulas of (51) and (52) are employed. The low energy constants B and f are functions of the strange quark mass. Assuming that we run simulations close enough around the physical point for the strange quark mass so that a linear expansion in m_s is sufficient, we write $B=B_s^{(0)}+m_sB_s^{(1)}$ and $f=f_s^{(0)}+m_sf_s^{(1)}$, where it should be noted that $B_s^{(0)}\neq B_0$ and $f_s^{(0)}\neq f_0$.

For the kaon sector we treat the K mesons as matter fields in the isospin 1/2 linear representation, and couple pions in SU(2) invariant ways (see, *e.g.*, Ref. [47]). For m_K and f_K this leads to the following fit formulas:

$$m_K^2=\alpha_m+\beta_m m_{ud}+\gamma_m m_s, \tag{66}$$

$$f_K=\bar{f}\left\{1+\beta_f m_{ud}-\frac{3}{4}\frac{\tilde{m}_\pi^2}{16\pi^2 f^2}\ln\left(\frac{\tilde{m}_\pi^2}{\mu^2}\right)\right\} \tag{67}$$

with $\bar{f}=\bar{f}_s^{(0)}+m_s\bar{f}_s^{(1)}$. In these formulas, the linear expansion in m_s should be regarded as that around the physical strange quark mass.

We apply a simultaneous fit to m_π, f_π and f_K employing the formulas of Eqs. (51), (52), and (67). The kaon mass m_K^2 is independently fitted according to Eq. (66). Calling the four data points corresponding to $\kappa_{ud}\geq 0.13754$ as Range I, the fit results for B, f, $\bar{l}_3$, $\bar{l}_4$ at the physical strange quark mass are presented in Table IX and Fig. 11 both

TABLE IX. Results for the low energy constants in the SU(2) ChPT fit together with the phenomenological estimates and the RBC/UKQCD results. $B=B_s^{(0)}+m_sB_s^{(1)}$ and $f=f_s^{(0)}+m_sf_s^{(1)}$ are given at the physical strange quark mass. f and f_0 are perturbatively renormalized at one-loop level. Range I, II, III denote the selection of data sets corresponding to $\kappa_{ud}\geq 0.13754$, $\kappa_{ud}\geq 0.13727$, $0.13770\geq\kappa_{ud}\geq 0.13727$, respectively. $\langle\bar{u}u\rangle$ and $\langle\bar{u}u\rangle_0$ are renormalized in the $\overline{\rm MS}$ scheme at 2 GeV.

	PACS-CS						phenomenology	RBC/UKQCD [41]
	Range I		Range II		Range III			
	w/o FSE	w/ FSE	w/o FSE	w/ FSE	w/o FSE	w/ FSE		
aB	1.907(36)	1.891(35)	1.941(20)	1.931(21)	1.947(20)	1.942(20)	-	2.414(61)(115)
af	0.0573(23)	0.0581(21)	0.0547(13)	0.0553(14)	0.0541(13)	0.0544(13)	-	0.0665(21)(47)
f [GeV]	0.1248(51)	0.1264(47)	0.1181(30)	0.1194(31)	0.1158(28)	0.1165(28)	0.1219(7) [48]	0.1148(41)(81)
f_π/f	1.063(8)	1.060(7)	1.074(5)	1.072(5)	1.078(5)	1.077(5)	1.072(7) [48]	1.080(8)
$m_{ud}^{\rm ph}B$ [GeV^2]	0.009345(27)	0.009332(26)	0.009381(16)	0.009372(17)	0.009391(17)	0.009387(16)	-	0.00937(57)(64)
$m_s^{\rm ph}B$ [GeV^2]	0.2709(43)	0.2686(43)	0.2782(25)	0.2768(26)	0.2794(26)	0.2787(26)	-	0.270(16)(18)
$a^3\langle\bar{u}u\rangle$	$-(0.141(3))^3$	$-(0.142(3))^3$	$-(0.138(2))^3$	$-(0.138(2))^3$	$-(0.137(2))^3$	$-(0.137(2))^3$	-	-
$\langle\bar{u}u\rangle$ [GeV^3]	$-(0.307(8))^3$	$-(0.309(7))^3$	$-(0.297(5))^3$	$-(0.299(5))^3$	$-(0.293(5))^3$	$-(0.294(4))^3$	-	$-(0.255(8)(8)(13))^3$
$\bar{l}_3$	3.23(21)	3.14(23)	3.32(10)	3.28(11)	3.31(10)	3.30(10)	2.9(2.4) [43]	3.13(33)(24)
$\bar{l}_4$	4.10(20)	4.04(19)	4.32(9)	4.28(10)	4.36(9)	4.34(9)	4.4(2) [46]	4.43(14)(77)
B/B_0	1.066(15)	1.064(15)	-	-	-	-	-	1.03(5)
f/f_0	1.073(55)	1.065(58)	-	-	-	-	-	1.229(59)
$\langle\bar{u}u\rangle/\langle\bar{u}u\rangle_0$	1.228(13)	1.205(14)	-	-	-	-	-	1.55(21)
χ^2/dof	0.33(68)	0.43(77)	2.0(1.0)	2.3(1.1)	2.8(1.8)	3.0(1.8)	-	0.3

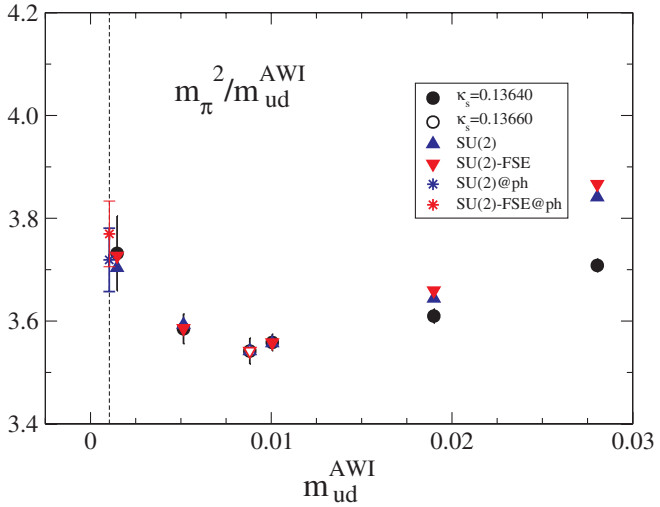


FIG. 16 (color online). SU(2) ChPT fit for $m_\pi^2/m_{ud}^{\mathrm{AWI}}$.

without and with finite-size corrections. We find that they are consistent with those obtained by the NLO conversion from the SU(3) low energy constants given in Table VII. Although our result for $\langle \bar{u}u \rangle$ is about 50% smaller than that of RBC/UKQCD, the difference comes from estimates of the renormalization factor: we use one-loop perturbation while they employ the nonperturbative RI-MOM scheme. This is verified by the observation that the value of f and the renormalization-free quantities $m_{ud}B$ and m_sB show consistency between our results and those of RBC/UKQCD.

Figures. 16 and 17 show that the quark mass dependences of $m_\pi^2/m_{ud}^{\mathrm{AWI}}$, f_π and f_K are reasonably described by the SU(2) ChPT formulas of (51), (52), and (67). The resulting χ^2/dof is 0.33(72), which is an order of magnitude smaller than in the SU(3) case. In Fig. 18 we illustrate the relative magnitude of the NLO contribution against the LO value for m_π^2/m_{ud}, f_π and f_K as a function of m_{ud}^{AWI} fixing the strange quark mass at the physical value. The convergences for f_π and f_K are clearly better than the SU(3) case.

In order to investigate the stability of the fit, we try two additional choices of the data sets for the SU(2) ChPT fit: Range II $(\kappa_{ud}, \kappa_s) = (0.13781, 0.13640)$, (0.13770, 0.13640), (0.13754, 0.13640), (0.13754, 0.13660), and (0.13727, 0.13640) includes one more data at a heavier pion mass added to Range I, and Range III $(\kappa_{ud}, \kappa_s) = (0.13770, 0.13640)$, (0.13754, 0.13640), (0.13754, 0.13660), and (0.13727, 0.13640) removes the point with the lightest pion mass from Range II. The results for B, f, $\bar{l}_3$, $\bar{l}_4$, and corresponding χ^2/dof are given in Table IX. While inclusion of the data at $\kappa_{ud} = 0.13727$ increases the value of χ^2/dof, the results for B, f, $\bar{l}_3$, $\bar{l}_4$ are consistent among the three cases within the error bars.

D. Finite size effects based on chiral perturbation theory

We evaluate finite-size effects based on the NLO formulas of ChPT. In the case of SU(3) ChPT the finite size effects defined by $R_X = (X(L) - X(\infty))/X(\infty)$ for $X = m_\pi, m_K, f_\pi, f_K$ are given by [39]:

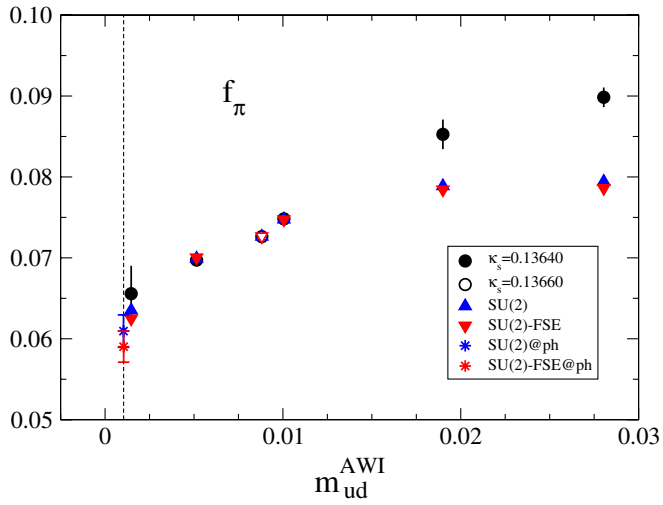
$$R_{m_\pi} = \frac{1}{4}\xi_\pi \tilde{g}_1(\lambda_\pi) - \frac{1}{12}\xi_\eta \tilde{g}_1(\lambda_\eta), \tag{68}$$

$$R_{m_K} = \frac{1}{6}\xi_\eta \tilde{g}_1(\lambda_\eta), \tag{69}$$

$$R_{f_\pi} = -\xi_\pi \tilde{g}_1(\lambda_\pi) - \frac{1}{2}\xi_K \tilde{g}_1(\lambda_K), \tag{70}$$

$$R_{f_K} = -\frac{3}{8}\xi_\pi \tilde{g}_1(\lambda_\pi) - \frac{3}{4}\xi_K \tilde{g}_1(\lambda_K) - \frac{3}{8}\xi_\eta \tilde{g}_1(\lambda_\eta) \tag{71}$$

with



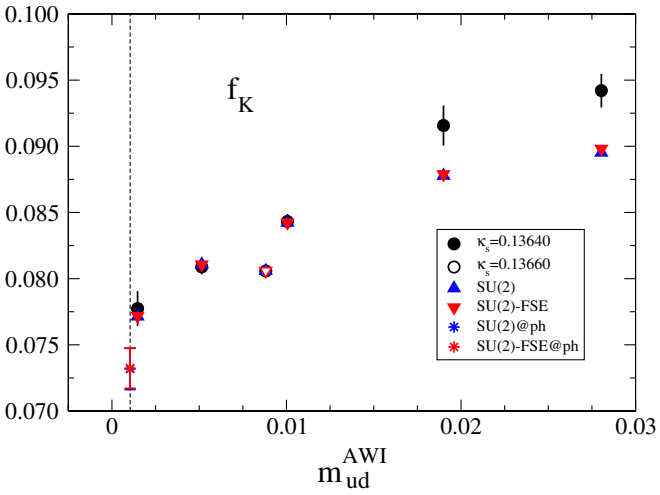


FIG. 17 (color online). SU(2) ChPT fit for f_π (left) and f_K (right).

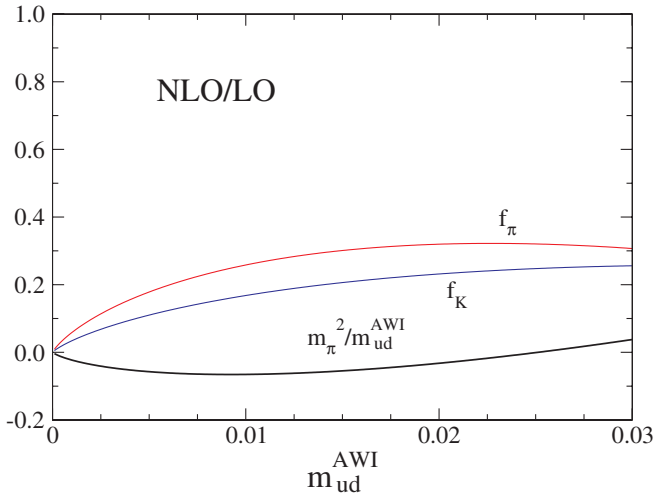


FIG. 18 (color online). Ratio of the NLO contribution to the LO one in the SU(2) ChPT fit for $m_\pi^2/m_{ud}^{\mathrm{AWI}}$, f_π and f_K.

$$\xi_{\mathrm{PS}} \equiv \frac{2m_{\mathrm{PS}}^2}{(4\pi f_\pi)^2}, \tag{72}$$

$$\lambda_{\mathrm{PS}} \equiv m_{\mathrm{PS}} L, \tag{73}$$

$$\tilde{g}_1(x) = \sum_{n=1}^{\infty} \frac{4m(n)}{\sqrt{n}x} K_1(\sqrt{n}x), \tag{74}$$

where K_1 is the Bessel function of the second kind and $m(n)$ denotes the multiplicity of the partition $n = n_x^2 + n_y^2 + n_z^2$. The authors in Ref. [39] expect that the above formulas are valid for $m_\pi L > 2$, in which our simulation points reside. In Figs. 12 and 13 we also plotted the ChPT fit results including finite size effects. The results are almost degenerate with the fit results without finite size effects except at the lightest simulation point at $\kappa_{ud} = 0.13781$ and the extrapolated values at the physical point. This feature is understood by looking at Fig. 19 where we plot the magnitude of R_X for $X = m_\pi, m_K, f_\pi$, and f_K with $L = 2.9$ fm as a function of m_π keeping the strange quark mass fixed at the physical value. The expected finite size effects are less than 2% for m_{PS} and f_{PS} at our simulation points. For m_{PS} this is true even at the physical point, while the value of f_π is decreased by 4% due to the finite size effects.

We can repeat the above study for the SU(2) case. The NLO formulas for m_π and f_π are given by [39]

$$R'_{m_\pi} = \frac{1}{4}\xi_\pi \tilde{g}_1(\lambda_\pi), \tag{75}$$

$$R'_{f_\pi} = -\xi_\pi \tilde{g}_1(\lambda_\pi). \tag{76}$$

In Figs. 16 and 17 we hardly detect finite size effects for $m_{ud}^{\mathrm{AWI}} > 0.001$. Figure 20 shows R'_X for $X = m_\pi, f_\pi$ with $L = 2.9$ fm as a function of m_π. The situation is similar to the SU(3) case: although finite size effects increase as m_π decreases, their magnitudes are at most 2% for m_π and 4% for f_π even at the physical point, which is easily expected by comparing the expressions of R and R'.

In Tables VI, VII, VIII, IX, X, and XI we compare the results with and without the finite size corrections in the one-loop ChPT analyses. Although we do not find any discrepancy beyond the statistical errors for all the cases, we take the results obtained with the finite size corrections as our best estimate.

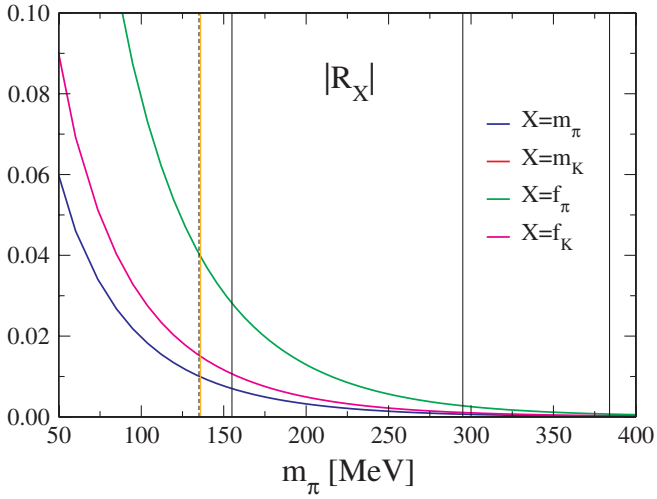


FIG. 19 (color online). $|R_X|$ ($R_{m_{\mathrm{PS}}} > 0$ and $R_{f_{\mathrm{PS}}} < 0$) for $X = m_\pi, m_K, f_\pi$, and f_K with $L = 2.9$ fm as a function of m_π at the physical strange quark mass based on the NLO SU(3) ChPT. Dotted vertical line denotes the physical point and the solid ones are for our simulation points. Orange vertical line represents the pion mass with $m_\pi L = 2$.

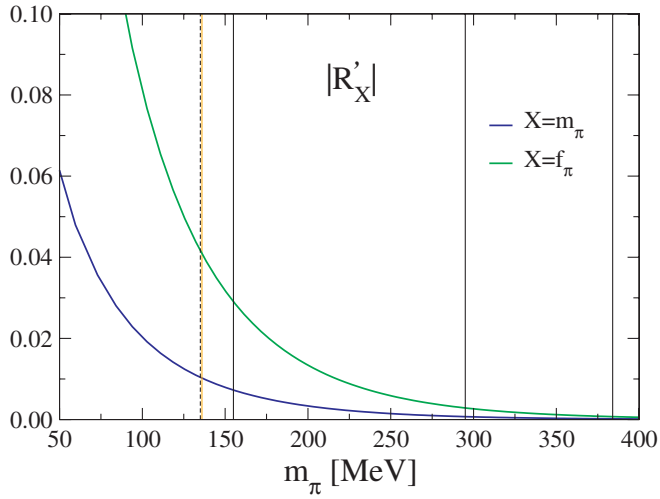


FIG. 20 (color online). $|R'_X|$ ($R'_{m_{\mathrm{PS}}} > 0$ and $R'_{f_{\mathrm{PS}}} < 0$) for $X = m_\pi, f_\pi$ with $L = 2.9$ fm as a function of m_π based on the NLO SU(2) ChPT. Dotted vertical line denotes the physical point and the solid ones are for our simulation points. Orange vertical line represents the pion mass with $m_\pi L = 2$.

TABLE X. Cutoff, renormalized quark masses, pseudoscalar meson decay constants determined with m_π, m_K, m_Ω inputs. Quark masses are renormalized at 2 GeV.

	physical point		experiment [49]	RBC/UKQCD [41]	MILC [42]
	w/o FSE	w/ FSE			
a^{-1} [GeV]	2.176(31)	2.176(31)	-	1.729(28)	continuum
$m_{ud}^{\overline{\rm MS}}$ [MeV]	2.509(46)	2.527(47)	-	3.72(16)(33)(18)	3.2(0)(1)(2)(0)
$m_s^{\overline{\rm MS}}$ [MeV]	72.74(78)	72.72(78)	-	107.3(4.4)(9.7)(4.9)	88(0)(3)(4)(0)
m_s/m_{ud}	29.0(4)	28.8(4)	-	28.8(0.4)(1.6)	27.2(1)(3)(0)(0)
f_π [MeV]	132.6(4.5)	134.0(4.2)	$130.7 \pm 0.1 \pm 0.36$	124.1(3.6)(6.9)	input
f_K [MeV]	159.2(3.2)	159.4(3.1)	$159.8 \pm 1.4 \pm 0.44$	149.6(3.6)(6.3)	$156.5(0.4)(^{+1.0}_{-2.7})$
f_K/f_π	1.201(22)	1.189(20)	1.223(12)	1.205(18)(62)	$1.197(3)(^{+6}_{-13})$

TABLE XI. Meson and baryon masses at the physical point in physical units. m_π, m_K, m_Ω are inputs.

channel	experiment [GeV] [49]	physical point [GeV]
π	0.1350	-
K	0.4976	-
ρ	0.7755	0.776(34)
K^*	0.8960	0.896(9)
ϕ	1.0195	1.0084(40)
N	0.9396	0.953(41)
Λ	1.1157	1.092(20)
Σ	1.1926	1.156(17)
Ξ	1.3148	1.304(10)
Δ	1.232	1.275(39)
Σ^*	1.3837	1.430(23)
Ξ^*	1.5318	1.562(9)
Ω	1.6725	-

Let us add a cautionary note that the finite-size formulas analyzed here lose viability when $m_\pi L$ becomes too small. Precisely at what values of $m_\pi L$ this takes place is not well controlled theoretically, however. Direct simulations on a larger lattice is required to pin down the actual magnitude of finite-size effects at the physical point. The need for such calculations are even more for baryons whose sizes are larger than mesons.

V. RESULTS AT THE PHYSICAL POINT

We need three physical inputs to determine the up-down and the strange quark masses and the lattice cutoff. We choose m_π, m_K and m_Ω. The choice of m_Ω has both theoretical and practical advantages: the Ω baryon is stable in the strong interactions and its mass, being composed of three strange quarks, is determined with good precision with small finite size effects.

For the pseudoscalar meson sector, we employ SU(2) chiral expansion as explained in the previous Section. For the vector mesons and the baryons we use a simple linear formula $m_{\rm had} = \alpha_h + \beta_h m_{ud}^{\rm AWI} + \gamma_h m_s^{\rm AWI}$, employing the data set in the same range $\kappa_{ud} \geq 0.13754$ as for the pseudoscalar meson sector. We do not rely on heavy meson effective theory (HMET) [50] or heavy baryon ChPT (HBChPT) [51] since they show very poor convergences even at the physical point [52]. The finite size effects are not taken into account for the vector mesons and the baryons. In Figs. 21–23, we show linear chiral extrapolations of the vector meson, the octet and the decuplet baryon masses, respectively. Red symbols represent the fit results at the measured values of $m_{ud}^{\rm AWI}$. The extrapolated values at the physical point are also denoted by blue star symbols, which should be compared with the experimental values plotted at $m_{ud}^{\rm AWI} = 0$.

Since the linear fit is applied to the data set at $\kappa_{ud} \geq 0.13754$, blue symbols at $\kappa_{ud} < 0.13754$ express the predictions from the fit results. We observe that the quark mass dependence of m_Ω is remarkably well described by the linear function, which assures that m_Ω is a good quantity for the physical input in the sense that its chiral behavior is easily controlled.

The results for the physical quark masses and the lattice cutoff are presented in Table X. The errors are statistical. They are provided with and without the finite size corrections based on the NLO SU(2) ChPT analyses. Both are almost degenerate. We also list the RBC/UKQCD and the MILC results in Table X for comparison. We find that our quark masses are sizably smaller than the RBC/UKQCD results. The deviation may be attributed to the renormalization factors: We employ the perturbative renormalization factors at one-loop level, while the RBC/UKQCD results are obtained with the nonperturbative ones. Actually, the ratio of the quark masses shows a good consistency. It is also suggestive that the MILC values for the quark masses, which are obtained with two-loop renormalization factors, reside between ours and the RBC/UKQCD ones. A nonperturbative calculation of the renormalization factor is in progress using the Schrödinger functional scheme.

Table X also lists the results for the pseudoscalar meson decay constants at the physical point using the physical quark masses and the cutoff determined above, which should be compared with the experimental values $f_\pi = 130.7$ MeV, $f_K = 159.8$ MeV, and $f_K/f_\pi = 1.223$ [49]. We observe a good consistency within the error

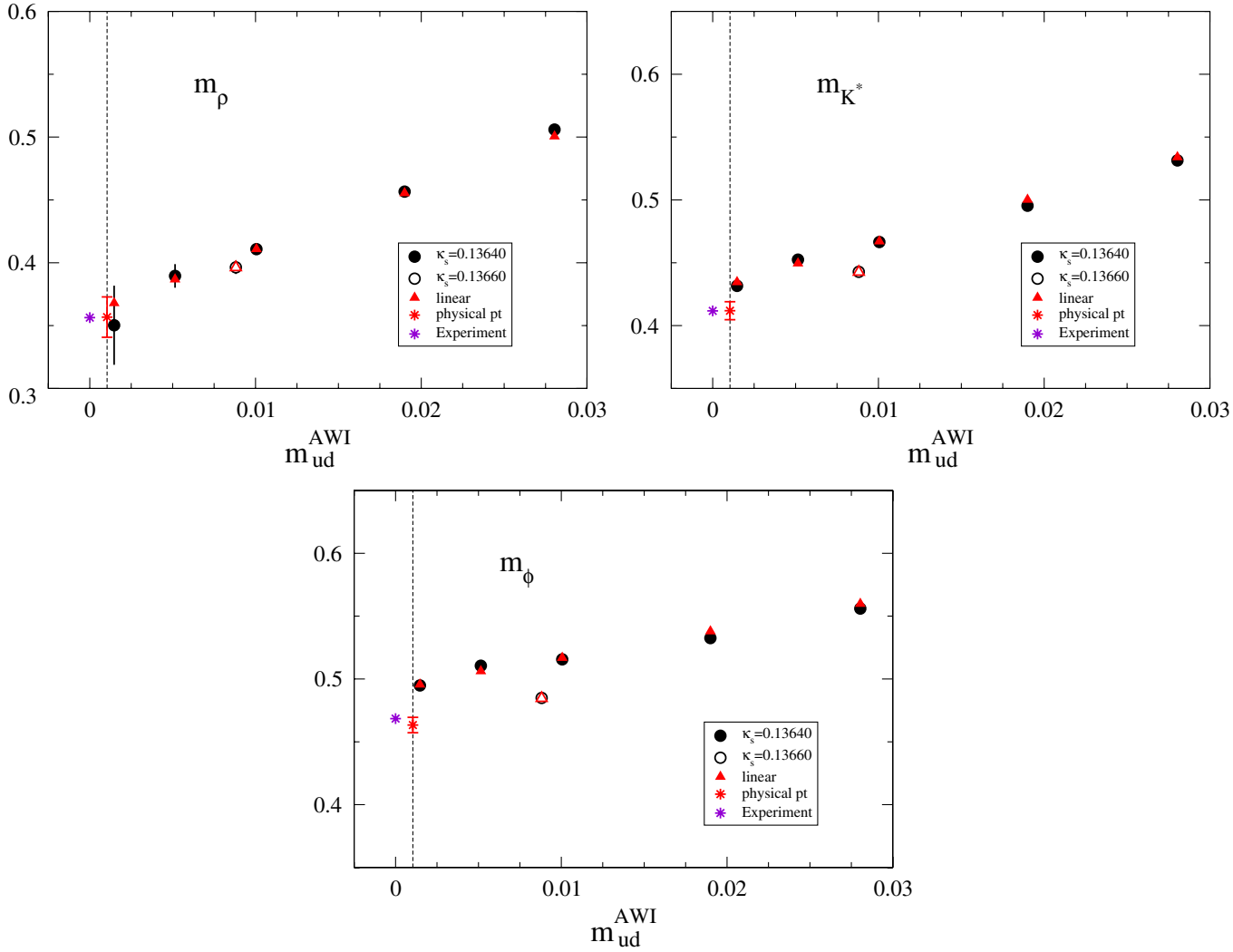


FIG. 21 (color online). Linear chiral extrapolation for the vector meson masses in lattice units. Red triangles represent the fit results at the measured values of m_{ud}^{AWI}. Red star symbol denotes the extrapolated value at the physical point. Experimental value in lattice units is also plotted at $m_{ud}^{\mathrm{AWI}} = 0$ for comparison.

of 2–3%. The ratio is 3% smaller than the experimental value in the case of the SU(2) ChPT fit with the finite size corrections. The RBC/UKQCD results for f_π and f_K are rather smaller than ours and the experimental values. Again, this may be explained by the difference of the renormalization methods. The MILC results are free from finite renormalization. A nonperturbative calculation of Z_A is in progress.

Table XI lists the numerical values for the hadron masses extrapolated to the physical point which is estimated by the SU(2) ChPT analyses on the pseudoscalar meson sector with the finite size corrections. Figure 24 depicts the comparison of the light hadron spectrum with the experimental values for the case of the physical point estimated with the finite size corrections. The largest discrepancy between our results and the experimental values is at most 3%, albeit errors are still not small for the ρ meson, the nucleon and the Δ baryon. The results are clearly encouraging, but further work is needed to remove the cutoff errors of $O((a\Lambda_{\mathrm{QCD}})^2)$. Another important future work is proper treatment of the ρ meson and the Δ baryon as resonances on the lattice. In the present work what we have measured are the lowest energy levels in the ρ meson and the Δ baryon channels, which are different from the true resonance masses. If we assume the physical value of 149 MeV for the ρ meson decay width, the Lüscher's finite size formula [53] tells us that the resonance mass is expected to be 5% higher than our measured ρ meson mass at $(\kappa_{ud}, \kappa_s) = (0.13781, 0.13640)$. This correction is within the statistical error. For other hopping parameters the deviations are less than 1%. We are planing to calculate the rho and the Delta resonance masses and their decay widths following the method in Ref. [54], where the rho resonance mass and its decay width were successfully extracted in two-flavor QCD based on the Lüscher's formula [53].

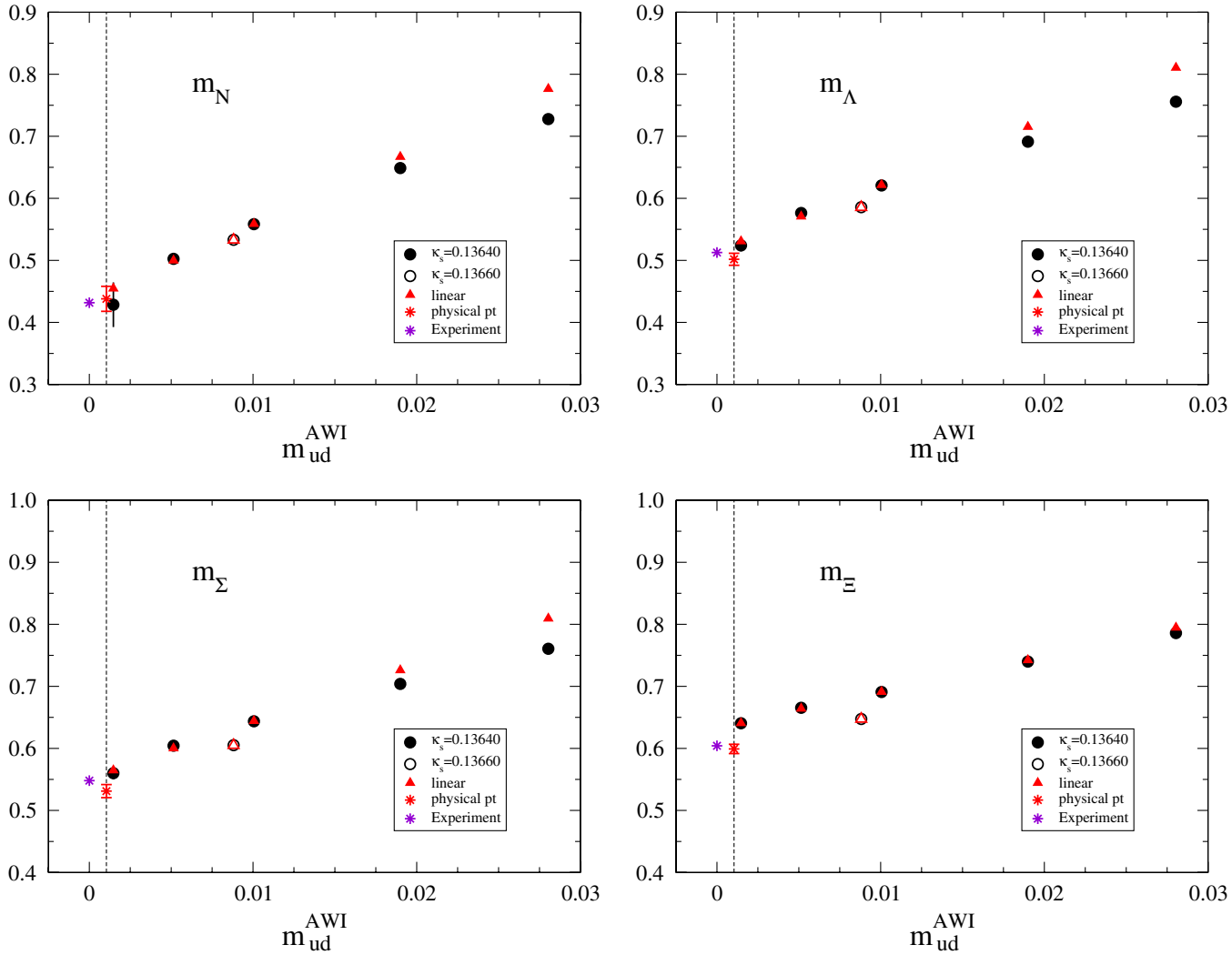


FIG. 22 (color online). Same as Fig. 21 for the octet baryons.

VI. STATIC QUARK POTENTIAL

In addition to the hadronic observables presented so far, we also calculate the Sommer scale which is a popular gluonic observable. In order to calculate the static quark potential we measure the temporal and the spatial Wilson loops with the use of the smearing procedure of Ref. [55]. The number of smearing steps is determined to be 20 after examining the sufficient overlap of the Wilson loops onto the ground state. The potential $V(r)$ is extracted from the Wilson loops applying a correlated fit of the form

$$W(r, t) = C(r)\exp(-V(r)t), \tag{77}$$

where the same fitting range $[t_{\mathrm{min}}, t_{\mathrm{max}}] = [5, 8]$ is chosen for all the simulations after investigating the effective potential

$$V_{\mathrm{eff}}(r, t) = \ln\left[\frac{W(r, t)}{W(r, t+1)}\right]. \tag{78}$$

Figure 25 shows a typical case of $V_{\mathrm{eff}}(r, t)$ with $r = 4, 8, 12$ at $\kappa_{ud} = 0.13770$. We find that plateau starts at $t = 4$ and signals are lost beyond $t = 7$. A result of $V(r)$ at $\kappa_{ud} = 0.13770$ is plotted in Fig. 26 as a representative case. Since good rotational symmetry and no sign of the string breaking are observed, we employ the following fitting form for the potential:

$$V(r) = V_0 - \frac{\alpha}{r} + \sigma r, \tag{79}$$

where V_0, α, σ are unknown parameters. The fitting range is $[r_{\mathrm{min}}, r_{\mathrm{max}}] = [3, 16]$.

The Sommer scale r_0 is a phenomenological quantity defined by

$$r_0^2 \left.\frac{dV(r)}{dr}\right|_{r=r_0} = 1.65. \tag{80}$$

Given Eq. (79) we obtain

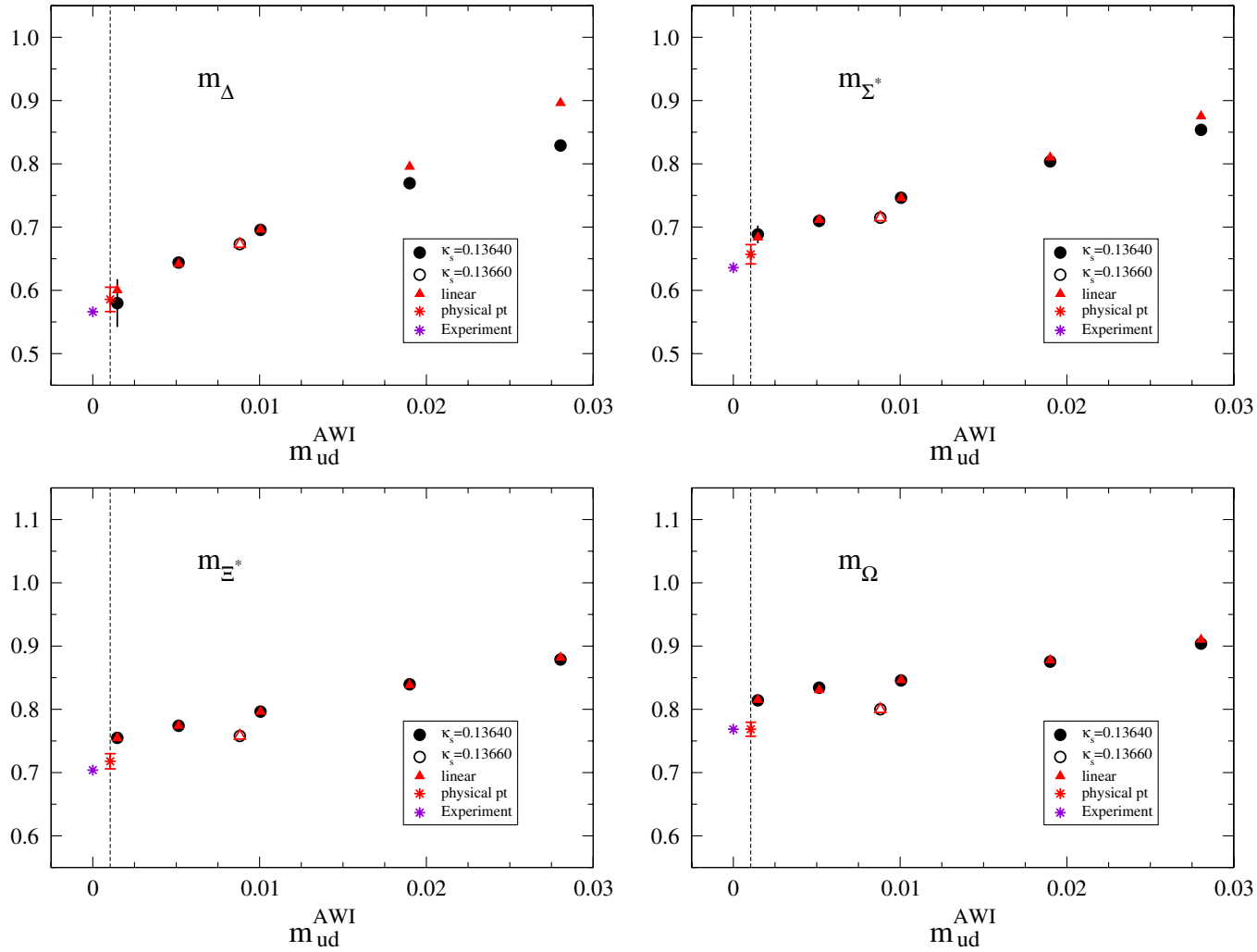


FIG. 23 (color online). Same as Fig. 21 for the decuplet baryons.

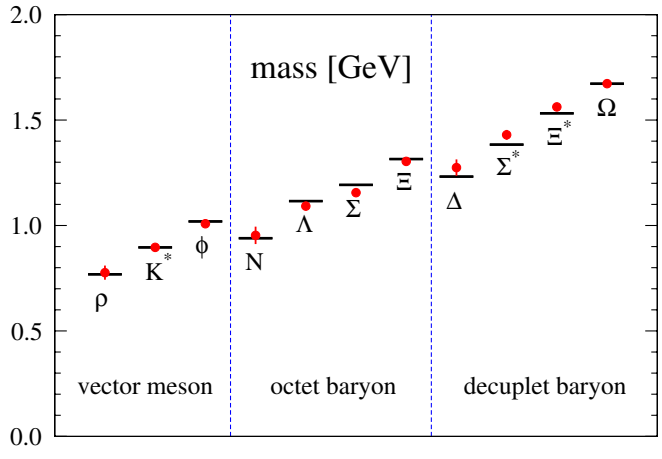


FIG. 24 (color online). Light hadron spectrum extrapolated to the physical point using m_π, m_K and m_Ω as input. Horizontal bars denote the experimental values.

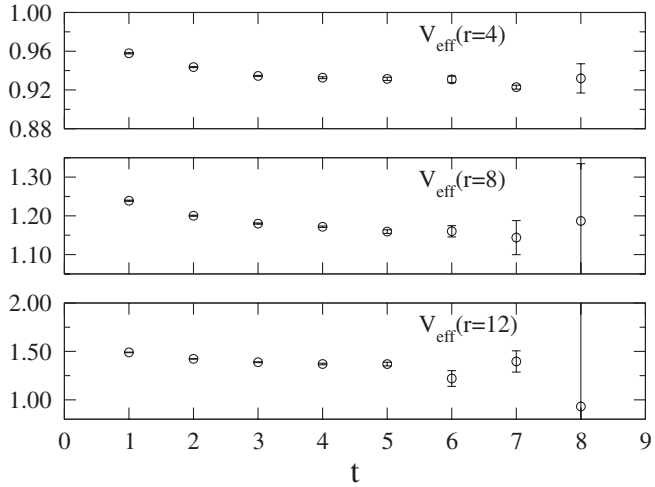


FIG. 25. Effective potential $V_{\rm eff}(r,t)$ with $r = 4,\ 8,\ 12$ at $\kappa_{ud} = 0.13770$ as a representative case.

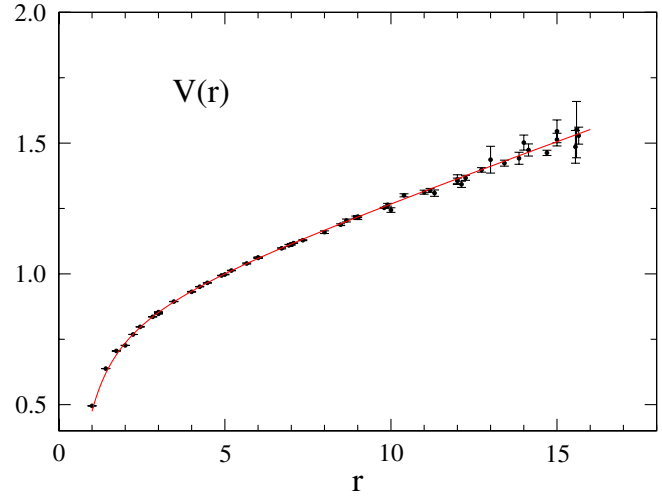


FIG. 26 (color online). Static quark potential $V(r)$ at $\kappa_{ud} = 0.13770$ as a representative case. Solid line denote the fit result with Eq. (79).

$$r_0 = \sqrt{\frac{1.65 - \alpha}{\sigma}}. \tag{81}$$

In Table XII we list the results for r_0 including the systematic errors due to the choices of $t_{\rm min}$ and $r_{\rm min}$.

At $(\kappa_{ud}, \kappa_s) = (0.13700, 0.13640)$, our result is compared with those of CP-PACS/JLQCD [56] in Table XIII. The two results are in reasonable agreement given the sizable magnitude of systematic errors caused by the shortness of plateau of effective masses for potentials.

In order to extrapolate r_0 to the physical point we employ a linear form $1/r_0 = \alpha_r + \beta_r \cdot m_{ud}^{\rm AWI} + \gamma_r \cdot m_s^{\rm AWI}$ for the data set at $\kappa_{ud} \geq 0.13754$. We illustrate the chiral extrapolation in Fig. 27, where the fit results are plotted by red triangles at the measured values of $m_{ud}^{\rm AWI}$. The extrapolated result of r_0 at the physical point is $5.427(51)(+81)(-2)$, which is $0.4921(64)(+74)(-2)$ fm in physical units with the aid of $a^{-1} = 2.176(31)$ GeV. The first error is statistical and the second and the third ones are the systematic uncertainties originating from the choice of $t_{\rm min}$ and $r_{\rm min}$, respectively.

TABLE XII. r_0 at each hopping parameter and the physical point. The first error at the physical point is statistical and the second and the third ones are the systematic uncertainties due to the choice of $t_{\rm min}$ and $r_{\rm min}$, respectively.

κ_{ud}	κ_s	r_0
0.13700	0.13640	4.813(30)(+40)(+13)
0.13727	0.13640	4.879(38)(+35)(+74)
0.13754	0.13640	5.121(21)(+82)(+9)
0.13754	0.13660	5.276(28)(+85)(+8)
0.13770	0.13640	5.176(23)(+54)(+8)
0.13781	0.13640	5.276(33)(+112)(−3)
physical point		5.427(51)(+81)(−2)

TABLE XIII. PACS-CS and CP-PACS/JLQCD results for r_0 in lattice units at $(\kappa_{ud}, \kappa_s) = (0.13700, 0.13640)$. Meaning of errors are the same as in Table XI.

	lattice size	r_0
PACS-CS	$32^3 \times 64$	4.813(30)(+40)(+13)
CP-PACS/JLQCD	$20^3 \times 40$	4.741(33)(+323)(+30)

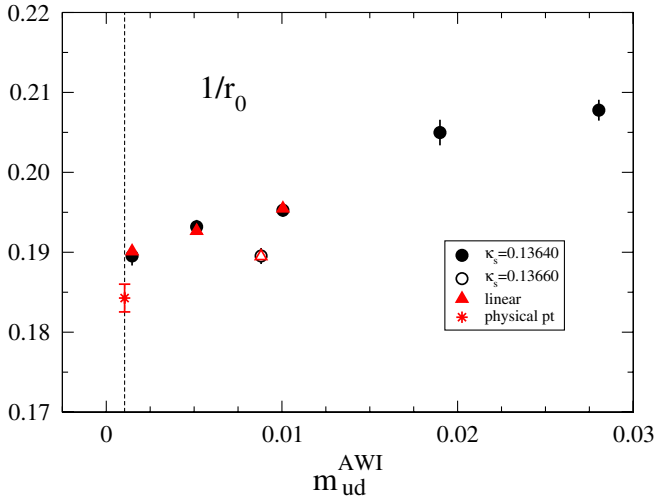


FIG. 27 (color online). Linear chiral extrapolation for $1/r_0$ at the physical point. Red triangles denote the fit results at the measured values of $m_{ud}^{\rm AWI}$.

VII. CONCLUSION

We have presented the first results of the PACS-CS project which aims at a 2 + 1 flavor lattice QCD simulation at the physical point using the $O(a)$-improved Wilson quark action. The DDHMC algorithm, coupled with several algorithmic improvements, have enabled us to reach $m_\pi = 156$ MeV, which corresponds to $m_{ud}^{\overline{\rm MS}}(\mu = 2\ {\rm GeV}) = 3.6$ MeV. We are almost on the physical point, except that the strange quark mass is about 20% larger than the physical value.

We clearly observe the characteristic features of the chiral logarithm in the ratios $m_\pi^2/m_{ud}^{\rm AWI}$ and f_K/f_π. We find that our data are not well described by the NLO SU(3) ChPT, due to bad convergence of the strange quark contributions. We instead employ the NLO SU(2) ChPT for m_π and f_π, and an analytic expansion around the physical strange quark mass for m_K and f_K in order to estimate the physical point. The low energy constants obtained in this way are compatible with phenomenological estimates and other recent lattice calculations.

Thanks to the enlarged physical volume compared to the previous CP-PACS/JLQCD work, we obtain good signals not only for the meson masses but also for the baryon masses. After linear chiral extrapolations of the vector and baryon masses the hadron spectrum at the physical

point shows a good agreement with the experimental values, albeit some of the hadrons have rather large errors and scaling violations remain to be examined. We find smaller values for the physical quark masses compared to the recent estimates in the literature. This may be due to the one-loop estimate of the renormalization factor.

At present the simulation at the physical point is under way, and the statistics of the run at $\kappa_{ud} = 0.13781$ is being accumulated. We are evaluating the nonperturbative renormalization factors for the quark masses and the pseudoscalar meson decay constants in order to remove perturbative uncertainties.

Once these calculations are accomplished, the next step is to investigate the finite size effects at the physical point, and then to reduce the discretization errors by repeating the calculations at finer lattice spacings.

ACKNOWLEDGMENTS

Numerical calculations for the present work have been carried out on the PACS-CS computer under the ''Interdisciplinary Computational Science Program'' of Center for Computational Sciences, University of Tsukuba. We thank T. Sakurai and H. Tadano for a series of informative discussions on single-precision acceleration of the solver. One of the authors (Y. K.) thank A. Kennedy for valuable discussions on the algorithmic improvements. A part of the code development has been carried out on Hitachi SR11000 at Information Media Center of Hiroshima University. This work is supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (Nos. 16740147, 17340066, 18104005, 18540250, 18740130, 19740134, 20340047, 20540248, 20740123, 20740139).

APPENDIX A: DDHMC ALGORITHM

In Appendix A, we describe our implementation details of the Lüscher's DDHMC algorithm [11] employed for our $\kappa_{ud} \leq 0.13770$ runs.

1. Domain decomposed HMC effective action

In this work we employ the $O(a)$-improved Wilson fermions. Before applying the domain-decomposition preconditioning for the quark determinant, we first apply Jacobi preconditioning to split the local clover term. The $O(a)$-improved Wilson-Dirac operator D is expressed as

$$D = 1 + T + M, \tag{A1}$$

where T is the local clover term, M is the hopping term. Jacobi preconditioning transforms the up-down quark determinant $|\det[D]|^2$ to

$$|\det[D]|^2 = |\det[1+T]|^2 |\det[\tilde{D}]|^2, \tag{A2}$$

where $\tilde{D} \equiv 1 + (1+T)^{-1}M = 1 + \tilde{M}$. By splitting lattice sites into even and odd domains, $\tilde{D}$ has the following 2×2 blocked matrix form,

$$\tilde{D} = \begin{pmatrix} \tilde{D}_{EE} & \tilde{D}_{EO} \\ \tilde{D}_{OE} & \tilde{D}_{OO} \end{pmatrix}, \tag{A3}$$

where the suffix $E(O)$ means the even (odd) domain. Applying the domain decomposition preconditioning for this form, we obtain

$$|\det[D]|^2 = |\det[1+T]|^2 |\det[\tilde{D}_{EE}]|^2 |\det[\tilde{D}_{OO}]|^2 | \times \det[\hat{D}_{EE}]|^2, \tag{A4}$$

where $\hat{D}_{EE}$ is the Schur complement of $\tilde{D}$ and expressed as

$$\hat{D}_{EE} = 1 - (\tilde{D}_{EE})^{-1}\tilde{D}_{EO}(\tilde{D}_{OO})^{-1}\tilde{D}_{OE}. \tag{A5}$$

Our domain decomposition is based on the four dimensional checkerboard coloring.

The operator $\hat{D}_{EE}$ can be further preconditioned by the spin and hopping structure because $\tilde{D}_{EO}$ ($\tilde{D}_{EO}$) only connects domain surface sites. Let P_E^{spin} (P_O^{spin}) be the spin and site projection operator to the even (odd) domain sites,

$$P_E^{\text{spin}}\psi_n = \begin{cases} 0 & \text{if } n \text{ is located on the bulk site of the even domain,} \\ \frac{1}{2}(1+\gamma_\mu)\psi_n & \text{if } n \text{ is located in the even domain and } n+\hat{\mu} \text{ is in the odd domain with one value of } \mu \text{ only,} \\ \frac{1}{2}(1-\gamma_\mu)\psi_n & \text{if } n \text{ is located in the even domain and } n-\hat{\mu} \text{ is in the odd domain with one value of } \mu \text{ only,} \\ \psi_n & \text{otherwise.} \end{cases} \tag{A6}$$

This projection operator satisfies the following relations.

$$(P_E^{\text{spin}})^2 = P_E^{\text{spin}}, \tag{A7}$$

$$\tilde{D}_{OE} = \tilde{D}_{OE}P_E^{\text{spin}}, \tag{A8}$$

and the same relation holds for the odd domain case. With these properties, $\hat{D}_{EE}$ satisfies

$$\hat{D}_{EE} = 1 - P_E^{\text{spin}} + \hat{D}_{EE}P_E^{\text{spin}}. \tag{A9}$$

This means that $\hat{D}_{EE}$ is a triangular matrix in view of the projection space. Thus we have

$$\det[\hat{D}_{EE}] = \det[P_E^{\text{spin}}\hat{D}_{EE}P_E^{\text{spin}}], \tag{A10}$$

where the matrix dimension of the operator $P_E^{\rm spin}\hat{D}_{EE}P_E^{\rm spin}$ is effectively reduced. We define

$$\hat{D}_{EE}^{\rm spin} \equiv P_E^{\rm spin}\hat{D}_{EE}P_E^{\rm spin}, \tag{A11}$$

for the reduced operator.

Since the domain block lattice extent we use is 8 and is an even number, the domain restricted operator $\tilde{D}_{EE}$ ($\tilde{D}_{OO}$) can be further preconditioned by the even-odd site preconditioning which is widely used for full lattice case in the literature.

$$\det[\tilde{D}_{EE}] = \det[(\hat{D}_{EE})_{ee}], \tag{A12}$$

$$(\hat{D}_{EE})_{ee} = 1 - (\tilde{M}_{EE})_{eo}(\tilde{M}_{EE})_{oe}, \tag{A12b}$$

where the suffices eo and oe mean hopping from an odd-site to an even-site and *vice versa*. For the odd domain operator $\tilde{D}_{OO}$ the same relation exists. Our even-odd site preconditioning is based on the four dimensional checkerboard coloring again.

After applying all these preconditioning we obtain the following lattice QCD partition function for degenerate up-down quarks.

$$\mathcal{Z} = \int \mathcal{D}P\mathcal{D}U\mathcal{D}\phi_{Ee}^\dagger\mathcal{D}\phi_{Ee}\mathcal{D}\phi_{Oe}^\dagger\mathcal{D}\phi_{Oe}\mathcal{D}\chi_E^\dagger\mathcal{D}\chi_E e^{-H[P,U,\phi_{Ee},\phi_{Oe},\chi_E]}, \tag{A13a}$$

$$H[P, U, \phi_{Ee}, \phi_{Oe}, \chi_E] = \frac{1}{2}\,\mathrm{Tr}[P^2] + S_g[U] + S_{\rm clv}[U] + \sum_{X=E,O} S_{q{\rm UV},X}[U, \phi_{Xe}] + S_{q{\rm IR}}[U, \chi_E], \tag{A13b}$$

where P is the canonical momenta for U, $S_g[U]$ the gauge action, and

$$S_{\rm clv}[U] = -2\log[\det[(1+T)]], \tag{A14a}$$

$$S_{q{\rm UV},X}[U, \phi_{Xe}] = |((\hat{D}_{XX})_{ee})^{-1}\phi_{Xe}|^2, \tag{A14b}$$

$$S_{q{\rm IR}}[U, \chi_E] = |(\hat{D}_{EE}^{\rm spin})^{-1}\chi_E|^2, \tag{A14c}$$

where χ_E is projected so as to satisfy $P_E^{\rm spin}\chi_E = \chi_E$. Our DDHMC algorithm is based on this partition function and the UVPHMC algorithm for strange quark is simply added to this form.

2. Multi time scale molecular dynamics integrator

We employ the Sexton-Weingarten multiple time scale molecular dynamics (MD) integrator [27]. The ordering to evolve link variables and momenta is arbitrary in the simple leap-frog integrator, and it is known that the so-called QPQ-ordering has better performance than that of the PQP-ordering [19,57,58]. While an actual performance comparison is not made systematically, this leads us to implement the QPQ-ordered multi time step integrator expecting better performance.

Suppose that there is a Hamiltonian H expressed as a sum of N potentials:

$$H = T(p) + \sum_{i=0}^{N-1} V_i(q), \tag{A15}$$

where q represents dynamical variables, $T(p)$ the kinetic term $p^2/2$, and p is the canonical momenta. This leads to the following equation of motion:

$$\dot{q} = p = -\{H, q\}_P, \tag{A16a}$$

$$\dot{p} = F = -\{H, p\}_P = \sum_{i=0}^{N-1} F_i, \tag{A16b}$$

$$F_i = -\frac{\partial V_i}{\partial q} = -\{V_i, p\}_P, \tag{A16c}$$

$$\{X, Y\}_P = \frac{\partial X}{\partial q}\frac{\partial Y}{\partial p} - \frac{\partial X}{\partial p}\frac{\partial Y}{\partial q}, \tag{A16d}$$

where $\{X, Y\}_P$ is Poisson bracket, and the dot $\dot{}$ is the abbreviation for the time derivative $d/d\tau$. The formal solution is written as

$$\begin{pmatrix} q \\ p \end{pmatrix}(\tau) = \exp\{\tau\hat{L}_H\}\begin{pmatrix} q \\ p \end{pmatrix}(0), \tag{A17}$$

where $\exp\{\tau\hat{L}_H\}$ is the exponentiation of the Liouvillean $\hat{L}_H X = -\{H, X\}_P$. In our case $\hat{L}_H X = \hat{L}_T X + \sum_{i=0}^{N-1}\hat{L}_{V_i}X$, where $\hat{L}_T X = -\{T, X\}_P$ and $\hat{L}_{V_i}X = -\{V_i, X\}_P$. We assume that the numbering of the potential V_i is ordered so as to satisfy $|F_i| < |F_{i-1}|$. Any molecular dynamics integrator is an approximation/decomposition of the operator exponential $\exp\{\tau\hat{L}_H\}$ using $\exp\{\tau\hat{L}_T\}$ and $\exp\{\tau\hat{L}_{V_i}\}$.

To explain symplectic molecular dynamics integrators we introduce the following mapping:

$$Q(\delta\tau) \equiv \exp\{\tau\hat{L}_T\}:(q, p) \to (q + \delta\tau p, p), \tag{A18a}$$

$$P_i(\delta\tau) \equiv \exp\{\tau\hat{L}_{V_i}\}:(q, p) \to (q, p + \delta\tau F_i). \tag{A18b}$$

Using these operators we can derive the following multi time scale integrators.

The PQP-ordered multi-time step integration operator $S^{\rm PQP}(\tau)$ is defined as

$$S^{\mathrm{PQP}}(\tau,(N_0,N_1,\ldots,N_{N-1})) = S^{\mathrm{PQP}}_{N-1}(\tau,(N_0,N_1,\ldots,N_{N-1})), \tag{A19}$$

where S^{PQP}_{N-1} is recursively defined as

$$\begin{aligned} S^{\mathrm{PQP}}_i(\tau,(N_0,N_1,\ldots,N_i)) &\equiv \left[P_i\left(\frac{\tau}{2N_i}\right)S^{\mathrm{PQP}}_{i-1}\left(\frac{\tau}{N_i},(N_0,N_1,\ldots,N_{i-1})\right)P_i\left(\frac{\tau}{2N_i}\right)\right]^{N_i}, \\ S^{\mathrm{PQP}}_0(\tau,N_0) &\equiv \left[P_0\left(\frac{\tau}{2N_0}\right)Q\left(\frac{\tau}{N_0}\right)P_0\left(\frac{\tau}{2N_0}\right)\right]^{N_0}, \end{aligned} \tag{A20}$$

where N_i is the step number for each time scale. The momentum is updated by $\delta\tau_i F_i$ with $\delta\tau_i = \tau/(\prod_{j=0,i}N_j)$ at depth i.

The QPQ-ordered multi time step integrator is used for our productive runs. The QPQ-ordered integrator S^{QPQ} is defined as

$$S^{\mathrm{QPQ}}(\tau,(N_0,N_1,\ldots,N_{N-1})) = S^{\mathrm{QPQ}}_{N-1}(\tau,(N_0,N_1,\ldots,N_{N-1})), \tag{A21}$$

where S^{QPQ}_{N-1} is recursively defined as

$$\begin{aligned} S^{\mathrm{QPQ}}_i(\tau,(N_0,\ldots,N_i)) &\equiv \left[S^{\mathrm{QPQ}}_{i-1}\left(\frac{\tau}{N_i},(N_0,\ldots,N_{i-1})\right)P_i\left(\frac{\tau}{N_i}\right)S^{\mathrm{QPQ}}_{i-1}\left(\frac{\tau}{N_i},(N_0,\ldots,N_{i-1})\right)\right]^{(N_i/2)(1+\delta_{i,N-1})}, \\ S^{\mathrm{QPQ}}_0(\tau,N_0) &\equiv \left[Q\left(\frac{\tau}{2N_0}\right)P_0\left(\frac{\tau}{N_0}\right)Q\left(\frac{\tau}{2N_0}\right)\right]^{N_0/2}. \end{aligned} \tag{A22}$$

In this case the division numbers, N_i, should be chosen from even numbers except for the outermost division number N_{N-1}.

The integrator described above is based on the nesting of the simple leap-frog integrator. We also note that we have not yet tried the so-called Omelyan integrator [58,59] for the recurrence kernel, albeit it is generally known to be a better scheme and may be used for our case. The multi time step integrator with the Omelyan kernel has been used in Refs. [41,60].

3. UV part solver

The UV part of the HMC algorithm is governed by the action Eq. (A14b). This contains the inversion of $(\hat{D}_{EE})_{ee}$ and $(\hat{D}_{OO})_{ee}$. In our parallel implementation of the algorithm, each block lattice is completely contained in a single node. This means that there is no ghost site exchange for multiplying $\tilde{D}_{EE}$. In this case SSOR preconditioning with natural site ordering is more efficient than the even-odd site preconditioning [61].

We solve the linear equation

$$(\hat{D}_{EE})_{ee}x_e = b_e, \tag{A23}$$

using SSOR preconditioned GCR solver where x_e and b_e carry the even-site data in the even-domain. We implemented the SSOR preconditioner with single precision arithmetic.

To solve Eq. (A23) with an SSOR preconditioner, we transform Eq. (A23) back to the unpreconditioned form,

$$\tilde{D}_{EE}y = c, \tag{A24a}$$

$$c = \begin{pmatrix} b_e \\ 0 \end{pmatrix}, \tag{A24b}$$

$$x_e = y_e, \tag{A24c}$$

where y_e is the even-site components of the full even-domain vector y. The right-hand vector c has zero for the odd-site components and has b_e for the even-site components.

We make use of the block/domain independence among the computational nodes and matrix structure of the domain operator $\tilde{D}_{EE}$ to solve Eq. (A24a). With the natural site-ordering in each block, $\tilde{D}_{EE}$ can be decomposed as

$$\tilde{D}_{EE} = 1 - L - U, \tag{A25}$$

where L is the forward hopping term and U is the backward hopping term. The L and U are strictly triangular for the natural site ordering because of the Dirichlet boundary condition for each block in a domain. Equation (A24a) is solved by

$$d = (1-\omega L)^{-1}c, \tag{A26a}$$

$$(\hat{D}_{EE})_{\mathrm{SSOR}}M_{\mathrm{SSOR}}z = d, \tag{A26b}$$

$$y = (1-\omega U)^{-1}M_{\mathrm{SSOR}}z, \tag{A26c}$$

where ω is an over-relaxation parameter to be tuned, $(\hat{D}_{EE})_{\mathrm{SSOR}}$ and M_{SSOR} are defined by

$$\begin{aligned} (\hat{D}_{EE})_{\mathrm{SSOR}} = \frac{1}{\omega}[&(1-\omega L)^{-1} + (1-\omega U)^{-1} \\ &+ (\omega-2)(1-\omega L)^{-1}(1-\omega U)^{-1}], \end{aligned} \tag{A27}$$

$$M_{\mathrm{SSOR}} = \sum_{j=0}^{N_{\mathrm{SSOR}}} [(1 - (\hat{D}_{EE})_{\mathrm{SSOR}})^j]_{32\,\mathrm{bit}}. \tag{A28}$$

The preconditioner M_{SSOR} is computed in single precision, and Eq. (A26b) is solved using GCR solver to double precision. The inverses $(1 - \omega L)^{-1}$ and $(1 - \omega U)^{-1}$ are easily solved by forward and backward substitutions, and Eq. (A27) is computed using the Eisenstat trick. The parameter ω is tuned to ~ 1.2 and N_{SSOR} to $5 \sim 10$ to achieve optimal performance. The maximal Krylov subspace dimension N_{KV} for GCR solver is chosen to avoid frequent restarting and residual stagnation, and our experience tells that $N_{\mathrm{KV}} \sim O(10)$ is sufficient.

4. IR part solver

The IR part of the DDHMC algorithm contains the linear equation as

$$\hat{D}_{EE}^{\mathrm{spin}} x_E = b_E, \tag{A29}$$

where $\hat{D}_{EE}^{\mathrm{spin}}$ is defined by Eq. (A29). Equation (A29) is solved with the restrictions $(1 - P_E^{\mathrm{spin}})x_E = 0$ and $(1 - P_E^{\mathrm{spin}})b_E = 0$.

As described in Ref. [11], directly solving Eq. (A29) is rather slow because the operator $\hat{D}_{EE}^{\mathrm{spin}}$ contains the domain inversions $(\tilde{D}_{EE})^{-1}$ and $(\tilde{D}_{OO})^{-1}$ with double precision. Instead of solving Eq. (A29), the solution x_E can be expressed using the unpreconditioned operator $\tilde{D}$ as

$$x_E = P_E^{\mathrm{spin}}(b_E - y_E), \tag{A30a}$$

$$\tilde{D}y = w, \tag{A30b}$$

$$w = \begin{pmatrix} 0 \\ \tilde{D}_{OE} b_E \end{pmatrix}, \tag{A30c}$$

where y_E is the even domain component of y, and w is the full lattice vector for which the even domain components are set to zero.

Equation (A30b) is efficiently solved with the GCR-SAP solver [29] via

$$\tilde{D} M_{\mathrm{SAP}} z = w, \tag{A31a}$$

$$y = M_{\mathrm{SAP}} z. \tag{A31b}$$

The SAP preconditioner M_{SAP} is computed in single precision as

$$M_{\mathrm{SAP}} = \left[K \sum_{j=0}^{N_{\mathrm{SAP}}} (1 - \tilde{D}K)^j \right]_{32\,\mathrm{bit}}, \tag{A32}$$

where

$$K = \begin{pmatrix} A_{EE} & 0 \\ -A_{OO}\tilde{D}_{OE}A_{EE} & A_{OO} \end{pmatrix}, \tag{A33a}$$

$$A_{EE} = (1 - \omega U)^{-1} M_{\mathrm{SSOR}} (1 - \omega L)^{-1}, \tag{A33b}$$

and A_{OO} similar to A_{EE}. The operator A_{EE} (A_{OO}) is the approximation for $(\tilde{D}_{EE})^{-1}$ ($(\tilde{D}_{OO})^{-1}$) via the SSOR fixed iteration M_{SSOR} defined in Eq. (A28).

Thus the solver for Eq. (A29) contains several tunable parameters; ω, N_{SSOR}, N_{SAP}, and N_{KV} the maximal Krylov subspace dimension for GCR. We observed that $\omega \sim 1.2$, $N_{\mathrm{SSOR}} = 1$, $N_{\mathrm{SAP}} = 10 \sim 20$ and $N_{\mathrm{KV}} = 40 \sim 100$ show satisfactory performance.

5. Dead/alive link method

Lüscher's DDHMC algorithm was originally proposed for the plaquette Wilson gauge action and the unimproved Wilson fermion [11]. He restricted the link variables evolved by the MD integrator to a subset. The link variables which connect the domain interfaces and are located parallel to the domain surfaces are kept fixed during the MD evolution (dead links), and only the remaining bulk links are evolved (alive links). The choice of the set of dead links are dictated by the condition that the alive links are decoupled. The method has the benefit that if the layout of the domain decomposition is properly matched to the compute node location there is no need to exchange link data during the MD evolution. Thus the algorithm becomes a semilocal update algorithm. To ensure the ergodicity a random parallel translation of the lattice coordinate origin is required after each HMC evolution.

In our case we employed the Iwasaki-gauge action and the $O(a)$-improved Wilson fermion. These actions have a larger lattice extent compared to the unimproved action (the rectangular part of the gauge action and the clover term of the fermion action), and one may worry about the semilocality of the MD evolution. Since the extension is still within two sites, we can conclude that the dead links are still domain connecting ones and those on the thin surface of the domains. Thus we can apply the same dead/alive link method as that for the unimproved case. For more extended gauge or fermion action the number of dead links should be enlarged to decouple the active links.

The efficiency of the dead/alive link method depends on the ratio of the number of active links and all links, which is estimated as

$$(N_B - 1)(N_B - 2)^3 / N_B^4, \tag{A34}$$

where N_B is a domain block size assuming N_B^4 blocking. In this paper we employed $N_B = 8$, which results in $\sim 37\%$ for the ratio. We employed the same algorithm for the random parallel translation as in Ref. [11].

APPENDIX B: MASS PRECONDITIONED DDHMC (MPDDHMC) ALGORITHM

As the up-down quark mass is reduced toward the physical point, we observed strong MD instability with the DDHMC algorithm. The origin of the instability is the appearance of near zero or negative eigenvalues in the $\tilde{D}$ spectra [62,63]. The corresponding eigenmodes yield a strong MD force and large fluctuations for the IR action (A14c) as described in the main text. We could handle the instability by reducing $\delta\tau_{\rm IR}$. However this results in very high values of the HMC acceptance, *e.g.,* $\gtrsim 90\%$, which is unnecessarily large compared with the optimal acceptance ratio, *e.g.,* ~60–70% for 2nd order MD integrator [64].

We introduce Hasenbusch's heavy mass preconditioner [13,14] to stabilize the IR part (A14c), and call the resulting algorithm MPDDHMC algorithm. We also implement several improvements in the algorithm. Our simulation with the lightest up-down quark mass corresponding to $\kappa_{ud} = 0.13781$ is finally carried out by the MPDDHMC algorithm. Here we describe the implementation details of the MPDDHMC algorithm.

1. Hasenbusch's heavy mass preconditioning for DDHMC algorithm

The mass preconditioner is introduced for the IR part action Eq. (A14c). The action is transformed and split into two pieces as

$$|\det[\hat{D}^{\rm spin}_{EE}]|^2 = \left|\det\left[\frac{\hat{D}^{\rm spin}_{EE}}{\hat{D}'^{\rm spin}_{EE}}\right]\right|^2 |\det[\hat{D}'^{\rm spin}_{EE}]|^2 = |\det[R_{EE}]|^2 |\det[\hat{D}'^{\rm spin}_{EE}]|^2, \tag{B1}$$

where the primed operator $\hat{D}'^{\rm spin}_{EE}$ is defined with the modified hopping parameter $\kappa' \equiv \rho\kappa$ keeping the clover term unchanged. Introducing the pseudofermion fields for each determinant we obtain

$$|\det[\hat{D}^{\rm spin}_{EE}]|^2 = \int \mathcal{D}\chi^\dagger_E \mathcal{D}\chi_E \mathcal{D}\zeta^\dagger_E \mathcal{D}\zeta_E \times e^{-S_{q{\rm I}\tilde{\rm R}}[U,\zeta_E] - S_{q{\rm IR}'}[U,\chi_E]}, \tag{B2a}$$

$$S_{q{\rm I}\tilde{\rm R}} = |(R_{EE})^{-1}\zeta_E|^2, \tag{B2b}$$

$$S_{q{\rm IR}'} = |(\hat{D}'^{\rm spin}_{EE})^{-1}\chi_E|^2. \tag{B2c}$$

The action Eq. (A14c) is replaced by Eqs. (B2b) and (B2c). Our MPDDHMC algorithm is based on this action.

The parameter ρ is a tunable parameter and should be chosen close to but less than $\rho = 1$ while keeping $R \sim 1$ so as to achieve optimal performance. For example, since the DDHMC simulation at $\kappa = 0.13770$ ran successfully, we use $\rho = 0.9995$ at $\kappa_{ud} = 0.13781$ since we expect $\kappa' \sim 0.13770$ would lead to a stabilized behavior for Eq. (B2c).

2. Solver improvements

As the quark masses are taken small, we encountered a solver stagnation or failure due to the presence of the near zero modes or the negative (real part) eigenmodes. In this case the GCR-SAP solver sometimes does not converge. Although this difficulty could be cured by changing the solver algorithm or finely tuning the solver parameters, *i.e.* ω, $N_{\rm SSOR}$, *etc.*, applying such remedies causes violation of the reversibility of the MD evolution when a loose stopping condition is adopted for the solver.

To avoid this situation we decided to change the solver algorithm. Our strategy is to combine (1) a strict stopping condition, (2) applying the method of chronological guess, and (3) adopting a solver algorithm robust against near zero and negative eigenvalues. The use of strict stopping condition (1) gives us room to flexibly change the solver algorithm without the reversibility violation, although this adds an extra computational cost. A part of the extra cost can be reduced by optimizing the choice of the initial vector (the chronological guess method [15]). It is also required to adopt a solver algorithm which is robust and fast against the ill-conditioned case. We employ the inner-outer solver strategy and the deflation technique [65–67] aiming for speed up and taming the difficult eigenmodes.

a. Inner-outer strategy

The gap between the rapidly increasing floating point capability of processors and the memory bandwidth is spreading because of the rather slow development of memory speed. To fill the gap the mixed precision or the inner-outer nested solver strategy has been proposed [68]. The outer solver must have the property that the preconditioner can be changed from iteration to iteration. Since the preconditioner can be replaced by another iterative solver to make an approximation for the outer problem, the preconditioner can be called as the inner-solver for the outer solver. The inner-outer solver enables us the use of single precision which effectively doubles the memory bandwidth, data cache size, and processor registers [69]. The GCR-SAP solver proposed by Lüscher [29] is also along this strategy. If the solver parameters can be chosen such that most of the computational time is spent in the inner-solver, we receive a maximal benefit from the use of single-precision arithmetic [69].

In this work we developed a version of the BiCGStab algorithm which enables us to follow the inner-outer strategy [70]. The benefit of BiCGStab compared to GCR (or GMRES) type algorithms is that BiCGStab has a shorter recurrence iteration, small memory requirement, and no restarting. To make the BiCGStab solver flexible against substitutions of the preconditioner, we slightly modify the algorithm. The point of modification is the following.

Any solver algorithm which has the following update point for the solution and residual vector can be modified to take the inner-outer solver form. To solve $Ax = b$,

suppose that an algorithm has the lines [compute parameter α and pre-search vector p.],

$$q = Ap, \tag{B3a}$$
$$r = r - \alpha q, \tag{B3b}$$
$$x = x + \alpha p, \tag{B3c}$$

where the method to obtain α and p depends on the outer solver algorithm. To enable a flexible preconditioner replace these lines as [compute parameter α and pre-search vector p.],

$$v = Mp, \tag{B4a}$$
$$q = Av, \tag{B4b}$$
$$r = r - \alpha q, \tag{B4c}$$
$$x = x + \alpha v, \tag{B4d}$$

where M is a preconditioner and must be an approximation for A^{-1}. The extra vector v is required to hold an intermediate vector. In this modification the search vector q is produced for the equation $AMy = b$, while the solution still keeps the solution-residual relation $r = b - Ax$ of the unpreconditioned equation. The preconditioner M can be changed from iteration to iteration in the outer solver, as far as the solution-residual relation is kept intact. In this way, a flexible preconditioner can be introduced for the outer algorithm.

This modification is applicable to many solvers which have similar local update points (CG, MR, CGS, *etc.*). The iterative refinement or Richardson iteration [69] is the simplest example. Solvers of GMRES type is also modified along this strategy. Longer recurrence relations of those algorithms require a series of extra vectors such as the above vector v (for ex. GCR, FGMRES), however.

We implement this modification to the BiCGStab solver and replace two update points. The preconditioner M is replaced by the single precision solver for $Ax = b$ with the appropriate precision conversion interface (single to double and *vice versa*). The tolerance of the inner solver can be relaxed as the outer residual approaches the desired tolerance, and this also reduces the cost of the inner-solver. We use the following tolerance control method for the inner-solver.

$$\mathrm{tol}_{\mathrm{inner}} = \min\left(\max\left(\frac{\mathrm{err}_{\mathrm{outer}}}{\mathrm{tol}_{\mathrm{outer}}}, 10^{-6}\right), 10^{-3}\right), \tag{B5}$$

where $\mathrm{err}_{\mathrm{outer}}$ is the relative residual norm $|b - Ax|/|b|$ for the outer solver. When the residual gap to the desired tolerance is larger than 10^{-6}, the inner solver is called with 10^{-6} tolerance which is the limit of single-precision arithmetic. As the outer residual decreases the inner-solver tolerance is relaxed.

The flexible BiCGStab algorithm is applied to both the IR and the UV problems. We solve

$$\tilde{D}(Mz) = w, \tag{B6}$$

with a flexible preconditioner M for the IR problem Eq. (A30b) (heavy mass κ' version is also modified), and

$$(\hat{D}_{EE})_{\mathrm{SSOR}}(Mz) = d, \tag{B7}$$

with a flexible preconditioner M for the UV problem Eq. (A26b). With these setup the flexible BiCGStab calls the inner solver 3 to 5 times to obtain the double precision solution.

b. Inner solver and deflation technique

Because the outer solver is well preconditioned by an inner-solver, the residual stagnation or convergence failure should not take place. However the problem of the near zero modes still remains and is left to the inner-solver to handle.

As explained in the main text we use the combination of BiCGStab and GCRO-DR [16] solvers. The inner solver usually uses BiCGStab. If residual stagnation or breakdown is detected the solver restarts with the GCRO-DR algorithm. The use of GCRO-DR is the key point to handle the ill-conditioned problem in our algorithm. The GCRO-DR incorporates the so-called deflation technique which removes or deflates the ill-conditioned eigenmodes from the matrix spectrum as has been described in the literature [65–67].

GCRO-DR solver has the following properties; (1) solves a linear equation and its eigensubspace simultaneously, (2) deflates the eigenmodes from the coefficient matrix and reduces the condition number, (3) can recycle the eigenmodes for another linear equation with the same or perturbed coefficient matrix but different right-hand vectors.

Since the inner solver is to be called several times by the outer solver and the outer solver is to be called many times during the MD evolution, the property (2) might largely help to solve the ill-conditioned problem. The property (3) opens the possibility of reusing the deflation subspace among the MD evolution steps for possible further speedup.

Unfortunately the performance of GCRO-DR algorithm highly depends on the problem to be solved, and we observed that the overhead is large compared to normal BiCGStab for well conditioned cases. One may consider reusing the deflation subspace generated by GCRO-DR for the so-called deflated BiCGStab (D-BiCGStab) algorithm. However, for well conditioned cases the overhead is still rather large and no improvement is observed. Moreover we observed that the rate of the occurrence of ill-conditioned cases is low. We, therefore, use the normal BiCGStab algorithm for a first attack, and switch the solver to GCRO-DR only when the stagnation or breakdown is detected as described above, otherwise continue to use the undeflated BiCGStab. Once the inner solver is switched to the GCRO-DR solver, GCRO-DR is kept being used until the outer iteration converges. If there is another linear equation with the same ill-conditioned coefficient matrix

in the MD force calculation, GCRO-DR continues with recycling the deflation subspace.

The actual equation to be solved by the inner solver is as follows. For the IR problem Eq. (B6) to obtain $t = Mz \sim (\tilde{D})^{-1}z$, we use

$$\tilde{D}Ks = z, \tag{B8a}$$

$$t = Ks, \tag{B8b}$$

where Eq. (B8a) is solved by BiCGStab or GCRO-DR algorithms, and computation are entirely done with single precision. The deflation subspace is spanned for $\tilde{D}K$ when switching occurs. Similarly we solve

$$(\hat{D}_{EE})_{\mathrm{SSOR}}t = z, \tag{B9}$$

for the UV problem Eq. (B7) to obtain $t = Mz \sim (\hat{D}_{EE})^{-1}_{\mathrm{SSOR}}z$.

The parameters for the GCRO-DR algorithm is the maximal dimension of Krylov subspace N_{KV} and the dimension of deflation/recycling subspace N_{REC}. The initial value is set to $(N_{\mathrm{KV}}, N_{\mathrm{REC}}) = (40, 20)$, and it is automatically enlarged when slow convergences are observed.

[1] S. Aoki *et al.* (CP-PACS Collaboration), Phys. Rev. Lett. **84**, 238 (2000); Phys. Rev. D **67**, 034503 (2003).
[2] A. Ali Khan *et al.* (CP-PACS Collaboration), Phys. Rev. Lett. **85**, 4674 (2000); Phys. Rev. D**65**, 054505 (2002).
[3] S. Aoki *et al.* (JLQCD Collaboration), Phys. Rev. D **68**, 054502 (2003).
[4] S. Aoki *et al.* (CP-PACS and JLQCD Collaborations), Phys. Rev. D **73**, 034501 (2006).
[5] Y. Iwasaki, Report No. UTHEP-118, 1983 (unpublished).
[6] T. Ishikawa *et al.* (CP-PACS and JLQCD Collaborations), Proc. Sci. LAT2006 (2006) 181.
[7] T. Ishikawa *et al.* (CP-PACS and JLQCD Collaborations), arXiv:hep-lat/0704.193.
[8] S. Aoki *et al.* (PACS-CS Collaboration), Proc. Sci. LAT2005 (2005) 111.
[9] A. Ukawa *et al.* (PACS-CS Collaboration), Proc. Sci. LAT2006 (2006) 039.
[10] T. Boku *et al.*, in *Cluster Computing and the Grid, IEEE International Symposium on* (IEEE Computer Society, Los Alamitos 2006), p. 233.
[11] M. Lüscher, J. High Energy Phys. 05 (2003) 052; Comput. Phys. Commun. **165**, 199 (2005).
[12] A. Kennedy, Nucl. Phys. B, Proc. Suppl. **140**, 190 (2005).
[13] M. Hasenbusch, Phys. Lett. B **519**, 177 (2001).
[14] M. Hasenbusch and K. Jansen, Nucl. Phys. **B659**, 299 (2003).
[15] R. Brower, T. Ivanenko, A. Levi, and K. Orginos, Nucl. Phys. **B484**, 353 (1997).
[16] M. Parks *et al.*, SIAM J. Sci. Comput. **28**, 1651 (2006).
[17] Ph. de Forcrand and T. Takaishi, Nucl. Phys. B, Proc. Suppl. **53**, 968 (1997).
[18] R. Frezzotti and K. Jansen, Phys. Lett. B **402**, 328 (1997); Nucl. Phys. **B555**, 395 (1999); **B555**, 432 (1999).
[19] S. Aoki *et al.* (JLQCD Collaboration), Phys. Rev. D **65**, 094507 (2002).
[20] C. Alexandrou, P. de Forcrand, M. D'Elia, and H. Panagopoulos, Phys. Rev. D **61**, 074503 (2000); Nucl. Phys. B, Proc. Suppl. **83**, 765 (2000); P. de Forcrand, Nucl. Phys. B, Proc. Suppl. **73**, 822 (1999).
[21] K-I. Ishikawa *et al.* (PACS-CS Collaboration), Proc. Sci. LATTICE 2006 (2006) 027.
[22] Y. Kuramashi, Proc. Sci. LATTICE 2007 (2007) 017.
[23] N. Ukita *et al.* (PACS-CS Collaboration), Proc. Sci. LATTICE 2007 (2007) 138.
[24] D. Kadoh *et al.* (PACS-CS Collaboration), Proc. Sci. LATTICE 2007 (2007) 109.
[25] L. Del Debbio *et al.*, J. High Energy Phys. 02 (2006) 011.
[26] L. Del Debbio *et al.*, J. High Energy Phys. 02 (2007) 056; 02 (2007) 082.
[27] J. C. Sexton and D. H. Weingarten, Nucl. Phys. **B380**, 665 (1992).
[28] A. Boriçi and P. de Forcrand, Nucl. Phys. **B454**, 645 (1995); A. Borrelli, P. de Forcrand, and A. Galli, Nucl. Phys. **B477**, 809 (1996); P. de Forcrand and A. Galli, arXiv:hep-lat/9603011; A. Galli and P. de Forcrand, Nucl. Phys. B, Proc. Suppl. **53**, 956 (1997).
[29] M. Lüscher, Comput. Phys. Commun. **156**, 209 (2004).
[30] A. Ukawa, Nucl. Phys. B, Proc. Suppl. **106**, 195 (2002).
[31] N. Madras and A. D. Sokal, J. Stat. Phys. **50**, 109 (1988).
[32] See, for example, G. P. Lepage, in *Proceedings of TASI'89 Summer School*, edited by T. DeGrand and D. Toussaint (World Scientific, Singapore, 1990), p. 97.
[33] T. Kaneko *et al.* (CP-PACS/JLQCD and ALPHA Collaborations), J. High Energy Phys. 04 (2007) 092.
[34] S. Aoki *et al.*, Phys. Rev. D **58**, 074505 (1998).
[35] Y. Taniguchi and A. Ukawa, Phys. Rev. D **58**, 114503 (1998).
[36] A. Aoki, R. Frezzotti, and P. Weisz, Nucl. Phys. **B540**, 501 (1999).
[37] S. Aoki, O. Bär, T. Ishikawa, and S. Takeda, Phys. Rev. D **73**, 014511 (2006); S. Takeda, Doctor Thesis (unpublished).
[38] J. Gasser and H. Leutwyler, Nucl. Phys. **B250**, 465 (1985).
[39] G. Colangelo, S. Dürr, and C. Haefeli, Nucl. Phys. **B721**, 136 (2005).
[40] G. Amorós, J. Bijnens, and P. Talavera, Nucl. Phys. **B602**, 87 (2001).
[41] C. Allton *et al.* (RBC and UKQCD Collaborations), Phys. Rev. D **78**, 114509 (2008).
[42] C. Bernard *et al.*, Proc. Sci., LAT2007 (2007) 090 [arXiv:0710.1118].

[43] J. Gasser and H. Leutwyler, Ann. Phys. (N.Y.) **158**, 142 (1984).
[44] J. Noaki *et al.* Phys. Rev. Lett. **101**, 202004 (2008).
[45] C. Urbach, Proc. Sci. LATTICE 2007 (2007) 022.
[46] G. Colangelo, J. Gasser, and H. Leutwyler, Nucl. Phys. **B603**, 125 (2001).
[47] A. Roessl, Nucl. Phys. **B555**, 507 (1999).
[48] G. Colangelo and S. Dürr, Eur. Phys. J. C **33**, 543 (2004).
[49] W.-M. Yao *et al.*, J. Phys. G **33**, 1 (2006).
[50] E. Jenkins, A. Manohar, and M. Wise, Phys. Rev. Lett. **75**, 2272 (1995).
[51] E. Jenkins and A. Manohar, Phys. Lett. B **255**, 558 (1991).
[52] B. Borasoy and U.-G. Meissner, Phys. Lett. B **365**, 285 (1996); J. Bijnens, P. Gosdzinsky, and P. Talavera, Nucl. Phys. **B501**, 495 (1997).
[53] M. Lüscher, Commun. Math. Phys. **105**, 153 (1986); Nucl. Phys. **B354**, 531 (1991); **B364**, 237 (1991).
[54] S. Aoki *et al.* (CP-PACS Collaboration), Phys. Rev. D **76**, 094506 (2007).
[55] G. S. Bali and K. Schilling, Phys. Rev. D **46**, 2636 (1992); G. S. Bali, K. Schilling, and C. Schlichter, *ibid.* **51**, 5165 (1995).
[56] CP-PACS and JLQCD Collaborations (unpublished).
[57] R. Gupta, G. W. Kilcup, and S. R. Sharpe, Phys. Rev. D **38**, 1278 (1988).
[58] T. Takaishi and Ph. de Forcrand, Phys. Rev. E **73**, 036706 (2006).
[59] I. P. Omelyan, I. M. Mryglod, and R. Folk, Comput. Phys. Commun. **151**, 272 (2003).
[60] C. Jung (RBC and UKQCD Collaborations), Proc. Sci. LATTICE2007 (2007) 037.
[61] N. Eicker, W. Bietenholz, A. Frommer, H. Hoeber, T. Lippert, and K. Schilling, Nucl. Phys. B, Proc. Suppl. **63**, 955 (1998); S. Fischer, A. Frommer, U. Glassner, T. Lippert, G. Ritzenhofer, and K. Schilling, Comput. Phys. Commun. **98**, 20 (1996).
[62] B. Joo *et al.* (UKQCD Collaboration), Phys. Rev. D **62**, 114501 (2000).
[63] Y. Namekawa *et al.* (CP-PACS Collaboration), Phys. Rev. D **70**, 074503 (2004).
[64] T. Takaishi, Comput. Phys. Commun. **133**, 6 (2000).
[65] For recent review on the deflation technique for lattice QCD, see: W. M. Wilcox, Proc. Sci. LATTICE 2007 (2007) 025; and references there in.
[66] A. M. Abdel-Rehim, R. B. Morgan, and W. Wilcox, Proc. Sci. LATTICE 2007 (2007) 026; R. B. Morgan, SIAM J. Sci. Comput. **24**, 20 (2002); R. B. Morgan and W. Wilcox, Nucl. Phys. B, Proc. Suppl. **106**, 1067 (2002); arXiv:math-ph/0405053; arXiv:0707.0505.
[67] M. Lüscher, J. High Energy Phys. 12 (2007) 011; 07 (2007) 081.
[68] A. Buttari, J. Dongarra, J. Kurzak, P. Luszczek, and S. Tomov, ACM Trans. Math. Softw. 34.
[69] S. Dürr *et al.*, arXiv:0802.2706.
[70] T. Sakurai and H. Tadano (private communication); H. Tadano and T. Sakurai, LSSC'07 **4818**, 721 (2007).