PROCEEDINGS OF SCIENCE

Generalized HMC using Nambu mechanics

Erik Lundstrum[a,1,*]

[a] *Physics Department, Columbia University, New York, NY 10027, USA*

E-mail: ell2156@columbia.edu

We describe a generalization of the Hybrid Monte Carlo (HMC) algorithm with molecular dynamics (MD) steps that use Nambu's generalized Hamiltonian dynamics. Time evolution in Nambu mechanics is dictated by multiple Hamiltonian functions, only one of which is required for an accept/reject step. We choose the other Hamiltonian to contain non-local functions of the gauge links, the forces from which enter the MD evolution. In this way, the link variables can be updated with knowledge of the large-distance behavior of the gauge field. We hope that this structure will provide the large-distance communication required to reduce critical slowing down in lattice QCD simulations.

The 41st International Symposium on Lattice Field Theory (LATTICE2024)
Liverpool, UK
28 July - 3 August 2024

*Speaker

© Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

PoS(LATTICE2024)044

1. Introduction

Lattice QCD simulations approaching the continuum limit encounter the critical slowing down (CSD) phenomenon, increasing the computational effort required to produce independent configurations. This results from the fact that the dominant part of the QCD action is the discretized Yang-Mills action which only couples nearby links, so that it takes many local updates for changes to diffuse throughout the lattice and equilibrate long-distance fluctuations.

At the same time, the fermion term in the QCD action depends non-locally on gauge links and is accommodated by the Hybrid Monte Carlo (HMC) algorithm [1], in which the non-local fermion forces are applied in parallel to all gauge links in the lattice. While the fermion forces are much weaker than the gauge forces, one might naturally expect that the presence of such non-local forces in the HMC will result in changes diffusing more rapidly across the lattice, since the gauge links instantaneously feel the force from long-distance fluctuations. Recent RBC/UKQCD investigations into CSD seem to confirm that this is the case [2, 3]. Adopting a momentum space perspective in which long-distance fluctuations correspond to low-momentum wave modes, these investigations reveal that the density of forces from the fermion determinant are flat across the spectrum of the gauge-covariant Laplace operator, and even show a slight reverse-CSD behavior, see Figures 3 and 4 in Ref. [3].

This evidence suggests that an algorithm which adds non-local forces to the molecular dynamics with greater magnitude than the fermion force may help to diffuse changes more rapidly across the lattice and thereby reduce CSD. In these proceedings we discuss such an algorithm based on a generalization of Hamiltonian dynamics often referred to as Nambu mechanics [4]. In this formalism, a mechanical body is described by an $n-$dimensional canonical multiplet, whose dynamics is described by $n-1$ Hamiltonian functions. The $n=2$ case is the familiar Hamiltonian dynamics. We focus on the $n=3$ version, though the idea is applicable to any n. In this case, a dynamical body is described by a canonical triplet (p,q,r) whose time evolution is dictated by the Hamiltonians (H,G). The main idea is that one of the functions $H(p,q,r)$ enters the Metropolis accept/reject step and the auxiliary Hamiltonian $G(p,q,r)$ is chosen to contain non-local functions of the gauge links, the forces from which enter the MD. These long-distance functions might be chosen to be the relevant degrees of freedom in the continuum limit, and we hope that including these forces provides the global communication required to reduce CSD.

The layout of this talk is as follows: Section 2 describes Nambu mechanics. Section 3 shows how the path integral may be evaluated using a classical Nambu mechanics system. In Section 4, we describe the application to lattice gauge theory. Numerical demonstrations are given in Section 5. Section 6 provides a further discussion of the algorithm.

2. Nambu mechanics

First we describe Nambu mechanics, emphasizing the features required to construct a generalized HMC algorithm. In this formalism, a single dynamical body is described by a canonical triplet of real-valued variables $\vec{x}=(p,q,r)$ whose time evolution is determined by two Hamiltonian functions $H(p,q,r)$ and $G(p,q,r)$. A generalized Hamiltonian mechanics that conserves both Hamiltonian functions results from choosing the rate of change $d\vec{x}/dt$ to be perpendicular to

PoS(LATTICE2024)044

the gradients of these two functions

$$\frac{d\vec{x}}{dt} = \vec{\nabla}H \times \vec{\nabla}G, \tag{1}$$

with respect to the coordinate system $\vec{x}$. This equation is the cross product of two conservative vector fields and satisfies $\vec{\nabla}\cdot(\vec{\nabla}H\times\vec{\nabla}G) = 0$. This means that $d\vec{x}/dt$ describes incompressible flows in phase space, so that Eq. (1) preserves the volume of phase space exactly. We can generalize this to a system of N coupled Nambu canonical triplets (p_i, q_i, r_i), $1 \le i \le N$, for which the generalized form of Eq. (1) gives the rate of change of any function $F(p, q, r)$ as

$$\frac{dF}{dt} = \sum_{i=1}^{N} \frac{\partial(F, H, G)}{\partial(q_i, p_i, r_i)} \equiv [F, H, G]\,. \tag{2}$$

Each term in the sum is calculated as a Jacobian, and in the final equality we have defined the generalized Poisson bracket called the Nambu bracket. We consider Hamiltonian functions satisfying $H(p, q, r) = H(-p, q, r)$ and $G(p, q, r) = G(-p, q, r)$, in which case the reverse trajectory is obtained by reflecting the set of variables $\{p_i\} \to \{-p_i\}$, $1 \le i \le N$.

3. Extending HMC phase space

This section describes how to generalize the classical MD component of the HMC algorithm to instead use a classical Nambu mechanics system. For the HMC, we can add fictitious variables to the Feynman path integral without changing the physical content of the theory

$$\langle A\rangle = \frac{1}{Z}\int [dU][d\pi]A(U)e^{-S(U)-\pi^2/2}, \tag{3}$$

since the contribution from the π integral cancels in the normalization. If we interpret the sum $H(U, \pi) = S(U) + \pi^2/2$ as the Hamiltonian of a classical mechanics system, we can evaluate path integral using a Markov chain Monte Carlo driven by evolution in a fictitious MD time [5]. For $n = 3$ Nambu mechanics, we add two fictitious variables

$$\langle A\rangle = \frac{1}{Z'}\int [dU][d\pi][d\rho]A(U)e^{-S(U)-\pi^2/2-\rho^2/2}, \tag{4}$$

and interpret the sum $H(p, q, r) = S(U) + \pi^2/2 + \rho^2/2$ as one of the Hamiltonians of the Nambu mechanics system. We can freely choose the form of the auxiliary Hamiltonian $G(p, q, r)$. We evaluate the path integral by integrating the classical Nambu evolution equations in Eq. (2). At the end of the trajectory the new configuration is accepted with probability $P_A = \min(1, e^{-\Delta H})$. Provided the integrator for the classical evolution equations is volume-preserving and reversible, this algorithm will satisfy detailed balance for *any* choice of the auxiliary Hamiltonian [6].

4. Application to lattice gauge theory

Here, we develop the finite-time-step updates needed to apply Nambu mechanics to lattice gauge theory. We consider first a single gauge link $U \in SU(N)$. This is an $N^2 - 1$ dimensional

PoS(LATTICE2024)044

manifold which, when needed, is viewed as a function of N^2-1 real variables q_a which parameterize the group in the vicinity of a constant group element U as

$$U' = \exp\left(-\sum_{a=1}^{N^2-1} q_a T_a\right) U. \tag{5}$$

Here T_a are the anti-hermitian Lie group generators and a is an adjoint representation index. It is convenient to introduce derivatives with respect to the link variables, defined abstractly as $\boldsymbol{e}_a$, in terms of the variables q_a. For a function $F(U)$, the derivative with respect to the link variable is

$$\boldsymbol{e}_a F(U) = \left.\frac{\partial F(U')}{\partial q_a}\right|_{q=0}. \tag{6}$$

For the gauge link U this means

$$\boldsymbol{e}_a U = -T_a U. \tag{7}$$

The minus sign is present in Eq. (5) so that the derivative operator in Eq. (6) corresponds to a basis in the Lie algebra of right-invariant vector fields that satisfy $[\boldsymbol{e}_i, \boldsymbol{e}_j] = c_{ij}^k \boldsymbol{e}_k$ [7].

Next, we formulate the Nambu mechanics classical system. To do so, to each of the variables $\{q_a\}$ in Eq. (5) we associate the two additional variables $\{p_a, r_a\}$, $1 \le a \le N^2-1$. These form the Nambu canonical triplets for the link. For the main Hamiltonian $H(p,U,r)$ against which we accept/reject, we choose

$$H = \frac{1}{2}\sum_{a=1}^{N^2-1} p_a^2 + \frac{1}{2}\sum_{a=1}^{N^2-1} r_a^2 + S(U). \tag{8}$$

Here $S(U)$ is the target action which dictates field dynamics, for example the Wilson gauge action in pure SU(3) gauge theory. Dynamical fermions may be included in the usual manner. According this form of H, at the beginning of each trajectory the variables p and r are drawn from a Gaussian distribution of unit variance. For the auxiliary Hamiltonian $G(p,U,r)$, we consider the separable form

$$G(p,q,r) = g_1(p) + g_2(r) + g_3(U), \tag{9}$$

subject to the reversibility condition $g_1(p) = g_1(-p)$. No restrictions are placed on the remaining functions $g_2(r)$ and $g_3(U)$.

Next, we construct the finite-time-step updates for each of the classical variables. To make a gauge link update, we can use Eq. (2) to calculate the rate of change of U

$$\dot{U} = \sum_{a=1}^{N^2-1} \frac{\partial(H,G)}{\partial(p_a,r_a)} \boldsymbol{e}_a U = \sum_{a=1}^{N^2-1} \frac{\partial(H,G)}{\partial(p_a,r_a)} (-T_a) U. \tag{10}$$

For H and G which are separable functions of the phase space variables, the factor multiplying the gauge link U is independent of U and this differential equation has a simple solution. The form of the discrete update can thus be written

$$U^{n+1} = \exp\left(-\tau \sum_{a=1}^{N^2-1} \frac{\partial(H,G)}{\partial(p_a,r_a)} T_a\right) U^n. \tag{11}$$

PoS(LATTICE2024)044

This is a left-multiplication by an SU(N) group element independent of U, under which the Haar measure is invariant. The updates for the real-valued phase space variables $\{p_a, r_a\}$ are linear

$$\begin{aligned} p_a^{n+1/2} &= p_a^n + \frac{\tau}{2}\Big[\boldsymbol{e}_a G \cdot \frac{\partial H}{\partial r_a} - \boldsymbol{e}_a H \frac{\partial G}{\partial r_a}\Big], \\ r_a^{n+1/2} &= r_a^n + \frac{\tau}{2}\Big[\boldsymbol{e}_a H \cdot \frac{\partial G}{\partial p_a} - \boldsymbol{e}_a G \frac{\partial H}{\partial p_a}\Big]. \end{aligned} \tag{12}$$

As in the usual HMC, we string these updates together in symmetric arrangements to craft reversible integrators. One such combination is a PRURP scheme

$$\begin{aligned} p_a^{n+1/2} &= p_a^n + \frac{\tau}{2}\Big[\boldsymbol{e}_a G \cdot \frac{\partial H}{\partial r_a} - \boldsymbol{e}_a H \cdot \frac{\partial G}{\partial r_a}\Big], \\ r_a^{n+1/2} &= r_a^n + \frac{\tau}{2}\Big[\boldsymbol{e}_a H \cdot \frac{\partial G}{\partial p_a} - \boldsymbol{e}_a G \cdot \frac{\partial H}{\partial p_a}\Big], \\ U^{n+1} &= \exp\left(-\tau \sum_{a=1}^{N^2-1} \frac{\partial(H,G)}{\partial(p_a, r_a)} T_a\right) U^n, \\ r_a^{n+1} &= r_a^{n+1/2} + \frac{\tau}{2}\Big[\boldsymbol{e}_a H \cdot \frac{\partial G}{\partial p_a} - \boldsymbol{e}_a G \cdot \frac{\partial H}{\partial p_a}\Big], \\ p_a^{n+1} &= p_a^{n+1/2} + \frac{\tau}{2}\Big[\boldsymbol{e}_a G \cdot \frac{\partial H}{\partial r_a} - \boldsymbol{e}_a H \cdot \frac{\partial G}{\partial r_a}\Big], \end{aligned} \tag{13}$$

All update steps are *explicit*, meaning that the rates of change are evaluated at the current instant in the MD time. The reversibility of this scheme is easily verified. Since the updating scheme is reversible, the first errors in the conservation of H and G occur at order τ^2.

We now consider the full lattice of link variables $U(x,\mu)$ residing at lattice location x and Euclidean direction μ. Each link is assigned a set of $N^2 - 1$ Nambu canonical triplets in the manner described above, so that the classical Nambu phase space assumes the form $\big\{U(x,\mu), \{p_a(x,\mu), r_a(x,\mu)\}_{1\le a\le N^2-1}\big\}$. The derivatives which enter the classical evolution equations are extended to include lattice location and direction indices

$$\frac{\partial}{\partial p_a} \to \frac{\partial}{\partial p_a^{x,\mu}}, \quad \frac{\partial}{\partial r_a} \to \frac{\partial}{\partial r_a^{x,\mu}}, \quad \boldsymbol{e}_a \to \boldsymbol{e}_a^{x,\mu}. \tag{14}$$

Each of these derivatives return zero when acting on a variable with different values of x or μ. The main Hamiltonian $H(p,U,r)$ now includes the sum over lattice locations and directions:

$$H = \frac{1}{2}\sum_{x,\mu}\sum_{a=1}^{N^2-1} p_a(x,\mu)^2 + \frac{1}{2}\sum_{x,\mu}\sum_{a=1}^{N^2-1} r_a(x,\mu)^2 + S(U). \tag{15}$$

We've managed to craft reversible and volume-preserving updates without any restrictions on the functions of r and U entering the auxiliary Hamiltonian G. As such, we can choose it to contain non-local functions of the gauge links, the forces from which enter the discrete momentum updates in Eq. (13).

At this point two things should be noted: First, the functions of the gauge links are constant between adjacent p and r updates and thus need not be reevaluated between these steps. This

PoS(LATTICE2024)044

means that when dynamical fermions are included, this algorithm requires no additional fermion force evaluations than does the usual HMC. For this reason, the Nambu HMC is not expected to be much more expensive than the HMC. Second, when examining the form of the gauge link update in Eq. (11), we find that it is a rather awkward sum over the product of three objects with adjoint representation indices a. This suggests that the updates do not obey a simple gauge symmetry as in the standard HMC. This is not a problem since the algorithm preserves the Haar measure and has the gauge-invariant statistical weight as its fixed point. One could even choose the auxiliary Hamiltonian G to be an explicitly gauge-non-invariant function of the gauge links and still recover the desired gauge-invariant distribution, which we have confirmed numerically.

5. Numerical tests

We now present some preliminary tests of the algorithm in pure SU(3) gauge theory to examine how the auxiliary Hamiltonian affects the sampling efficiency. We use the Wilson action in the main Hamiltonian in Eq. (15). Tests are performed on a 14^4 lattice using periodic boundary conditions. We have numerically confirmed the correctness of this algorithm by performing high-precision plaquette measurements for several choices of the auxiliary Hamiltonian G. Some choices for gauge link functions in G included Polyakov loops, Wilson loops and the gauge-non-invariant function $\mathrm{ReTr}U(x,\mu)$. In our tests, we found that Wilson loops provided the greatest benefit of the functions listed above, so these are the tests which we present. There are many non-local functions to try with this algorithm, and it is likely some will provide greater benefit than Wilson loops.

For our tests, we choose the auxiliary Hamiltonian to be linear in the variable $r_a(x,\mu)$:

$$G(U,r) = \gamma \sum_{x,\mu} \sum_{a=1}^{N^2-1} r_a(x,\mu) - \kappa f(U). \tag{16}$$

Here γ and κ are constant parameters which can be varied to optimize the algorithm. The virtue of this auxiliary Hamiltonian is that for $\kappa = 0$ and $\gamma = 1$, the variables $r_a(x,\mu)$ become non-dynamic and the classical evolution equations for the remaining variables reduce to those of the familiar Hamiltonian mechanics. As such, this auxiliary Hamiltonian provides the "minimal deformation" of the HMC. Here, we choose the function $f(U)$ in Eq. (16) to be the sum of all 3×3 Wilson loops. We arbitrarily chose the parameters $\gamma = 1$ and $\kappa = 3$.

An important note concerns the cost comparison between this algorithm and the HMC. The trajectory length and acceptance ratio are free parameters which, in both algorithms, can be tuned to achieve optimal sampling efficiency; these likely take different values for each of the algorithms. Here, both the trajectory length and the separation of two Markov chain samples will be described in units of MD time and by the number of Wilson force evaluations performed. We do not attempt to tune the trajectory length but instead the choose the HMC and Nambu HMC trajectories to contain the same number of Wilson force evaluations and adjust the discrete time step separately for each algorithm to achieve the desired acceptance ratio. This is done with the view in mind that the fermion force evaluation dominates the computing cost per update, and so the additional overhead from evaluating gradients of the non-local function of the gauge links likely won't add much overhead to realistic simulations. To test the sampling efficiency we measure the normalized

PoS(LATTICE2024)044

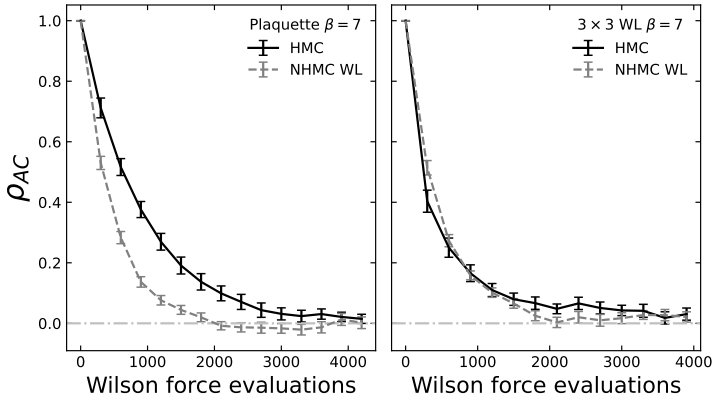

Figure 1: Comparison at $\beta = 7$ of the normalized autocorrelation ρ_{AC} decreasing as a function of the number of Wilson force evaluations for the HMC and Nambu HMC with 3×3 Wilson loops (NHMC WL). The left and right plots show the autocorrelation of the plaquette and 3×3 Wilson loop, respectively.

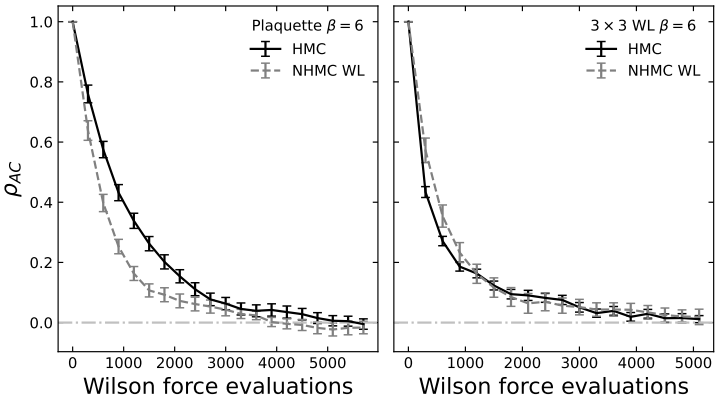

Figure 2: Comparison at $\beta = 6$ of the normalized autocorrelation ρ_{AC} decreasing as a function of the number of Wilson force evaluations for the HMC and Nambu HMC with 3×3 Wilson loops (NHMC WL). The left and right plots show the autocorrelation of the plaquette and 3×3 Wilson loop, respectively.

autocorrelation of observables. For a set of measurements $\{O\}$ with mean $\overline{O}$, the autocorrelation R_{AC}, normalized autocorrelation ρ_{AC} and integrated autocorrelation τ_{int} are defined as

$$R_{\mathrm{AC}}(k) = \frac{1}{N-k} \sum_{i=0}^{N-k-1} \Big(O(i) - \overline{O}\Big)\Big(O(i+k) - \overline{O}\Big), \tag{17}$$

$$\rho_{\mathrm{AC}}(k) = \frac{R_{\mathrm{AC}}(k)}{R_{\mathrm{AC}}(0)}, \qquad \tau_{\mathrm{int}} = \frac{1}{2} + \sum_{k=0}^{\infty} \rho_{AC}(k). \tag{18}$$

In practice we truncate the sum in τ_{int} when we find $\rho_{AC}(k)$ is consistent with zero, within errors. We make our tests at two values of β, plotting the decrease in the normalized autocorrelation with an increasing number of Wilson force evaluations for both the Nambu HMC and HMC. In these tests the HMC uses a standard leapfrog integrator. We haven't made an detailed tuning effort here to optimize the parameters entering each algorithm.

The first test is performed at $\beta = 7$. The results can be found in Fig. 1. Plotted errors are calculated using the jackknife method. This test utilizes 5000 trajectories. Trajectories contain 300

PoS(LATTICE2024)044

Wilson force evaluations; for the Nambu HMC, this corresponds to $t = 3.0$ MD time units at an acceptance ratio of 0.766 and, for the HMC, $t = 3.5$ MD time units at an acceptance ratio 0.877. We find that the Nambu HMC provides a more rapid plaquette decorrelation, with an integrated autocorrelation time which is $66 \pm 3\%$ of that produced by the HMC. We also find a slightly more rapid 3×3 Wilson loop decorrelation, with an integrated autocorrelation time $93 \pm 3\%$ of that produced by the HMC.

The second test is performed at $\beta = 6$. The results can be found in Fig. 2. Here again the test utilize 5000 trajectories, each of which contains 300 Wilson force evaluations. For the Nambu HMC this corresponds to $t = 3.8$ MD time units with acceptance ratio 0.69, and for the HMC $t = 4.3$ MD time units with acceptance ratio 0.82. We again find that the Nambu HMC more rapidly decorrelates the plaquette; the integrated autocorrelation time is $74 \pm 2\%$ that of the HMC. The Nambu HMC performs more poorly with the Wilson loop integrated autocorrelation, which is $108 \pm 3\%$ that of the HMC.

6. Conclusions

In this talk we've described a generalization of the HMC algorithm which utilizes Nambu's generalized Hamiltonian mechanics. The virtue of this formalism is that the classical evolution equations used in the MD portion of the update can include forces from non-local functions of the gauge links. We hope to choose these non-local functions to be the relevant degrees of freedom in the continuum limit, thereby reducing CSD. In preliminary tests of the algorithm we find that it consistently produces shorter autocorrelation times than the HMC for the the plaquette, while producing comparable autocorrelation times for the 3×3 Wilson loop. In these tests we did not make a careful tuning to optimize the Nambu HMC and used a function with a modest amount of non-locality. The results here indicate that with a more effective non-local function and a careful tuning of the algorithm, the Nambu HMC could very well out perform the HMC.

The Nambu HMC can easily be extended to include dynamical fermions, and requires no more fermion force evaluations than the usual HMC. Fermions contribute the majority of the cost per link update, and so there isn't much additional overhead from evaluating the gradients of the additional non-local functions in this algorithm.

Clearly, there are a vast number of choices one could make for the second auxiliary Hamiltonian G and it is not immediately obvious which choice would most effectively provide the long-distance communication required to overcome CSD. A particularly attractive option is to include the fermion determinant in the auxiliary Hamiltonian with a large coefficient, as this adds non-locality with essentially no added computational cost compared to the HMC. Tests of additional non-local functions and simulations including dynamical fermions are the subject of current study.

Acknowledgments

Thank you to my RBC/UKQCD and Columbia colleagues and in particular to Norman Christ for helpful feedback and discussions. This work was supported in part by the U.S. Department of Energy (DOE) Grant No. DE-SC0011941 and in part by the Exascale Computing Project (17-SC-20-SC).

PoS(LATTICE2024)044

References

[1] S. Duane, A. Kennedy, B.J. Pendleton and D. Roweth, *Hybrid Monte Carlo*, *Phys. Lett. B* **195** (1987) 216.

[2] T. Nguyen, P. Boyle, N.H. Christ, Y.-C. Jang and C. Jung, *Riemannian manifold Hybrid Monte Carlo in Lattice QCD*, *PoS* **LATTICE2021** (2022) 582 [2112.04556].

[3] C. Jung and N.H. Christ, *Riemannian manifold HMC with fermions*, *PoS* **LATTICE2023** (2024) 009 [2401.13226].

[4] Y. Nambu, *Generalized hamiltonian dynamics*, *Phys. Rev. D* **7** (1973) 2405.

[5] D.J.E. Callaway and A. Rahman, *Microcanonical ensemble formulation of lattice gauge theory*, *Phys. Rev. Lett.* **49** (1982) 613.

[6] E. Lundstrum, *Generalized HMC using Nambu mechanics for lattice QCD*, 2409.18958.

[7] A.D. Kennedy, P.J. Silva and M.A. Clark, *Shadow Hamiltonians, Poisson Brackets, and Gauge Theories*, *Phys. Rev. D* **87** (2013) 034511 [1210.6600].

PoS(LATTICE2024)044