Ergodicity of the SMD algorithm in lattice QCD

Martin Lüscher October 2016; revised June 2017

1. Introduction

An exact implementation of the stochastic molecular dynamics (SMD) algorithm [1] requires the inclusion of an accept-reject step at the end of the molecular dynamics trajectories [2]. If Hamilton's equations are integrated very accurately, using a 4th order symplectic integrator [3], for example, the effect of the accept-reject step may however be negligible with respect to the statistical errors.

In this note the long-time behaviour of the algorithm is examined and it is shown that the process (with or without accept-reject step) converges to a unique stationary state if the simulation step size ϵ is sufficiently small.

2. SMD algorithm

In order to keep the discussion as simple as possible, the pure SU(N) gauge theory is first considered. A finite lattice is assumed and the Hamilton function is taken to be of the standard separated form

$$H(\pi,U) = \tfrac{1}{2}(\pi,\pi) + S(U), \qquad (\pi,\pi) = \sum_{x,\mu} \pi^a(x,\mu)\pi^a(x,\mu), \tag{2.1}$$

with $S(U)$ being the gauge action. The exact form of the action does not matter, but it must be a smooth function of the link variables $U(x,\mu)$ with uniformly bounded derivatives, as is the case for the widely used actions. The (exact) SMD algorithm [1,2] simulates the distribution

$$P(\pi,U) \propto \mathrm{e}^{-H(\pi,U)} \tag{2.2}$$

by repeatedly applying three update steps (momentum rotation, molecular dynamics evolution, accept-reject step).

2.1 Momentum rotation

This operation transforms the current fields π, U according to

$$R_\epsilon : \ \pi \to c_1\pi + c_2\upsilon, \quad U \to U, \tag{2.3}$$

where υ is a Gaussian random momentum field with mean zero and unit variance. The coefficients

$$c_1 = \mathrm{e}^{-\gamma\epsilon}, \quad c_2 = (1 - c_1^2)^{1/2}, \tag{2.4}$$

depend on the molecular-dynamics time step ϵ and an adjustable parameter γ that controls the rotation angle. Typically γ is taken to be of order unity and $\gamma\epsilon$ is then a small number. Both ϵ and γ will be assumed to be positive.

2.2 Molecular dynamics evolution

The second operation, I_ϵ, approximately integrates Hamilton's equations

$$\partial_t\pi(x,\mu) = -\partial_{x,\mu}S(U), \qquad \partial_t U(x,\mu) = \pi(x,\mu)U(x,\mu), \tag{2.5}$$

from the current time t to $t+\epsilon$ using a reversible, symplectic integrator. A class of such integrators is obtained by combining the elementary steps

$$I_{\pi,\epsilon} : \ \pi \to \pi - \epsilon\partial S(U), \quad U \to U, \tag{2.6}$$

$$I_{U,\epsilon} : \ \pi \to \pi, \quad U \to \mathrm{e}^{\epsilon\pi}U. \tag{2.7}$$

The leapfrog integrator

$$I_\epsilon = I_{\pi,\epsilon/2}I_{U,\epsilon}I_{\pi,\epsilon/2}, \tag{2.8}$$

for example, integrates Hamilton's equations up to an error of order ϵ^3 per step. In this note, only integrators of this kind, i.e. short sequences of the steps (2.6) and (2.7) with varying step sizes of order ϵ, are considered.

2.3 Accept-reject step

The fields $(\tilde{\pi}, \tilde{U}) = I_\epsilon(\pi, U)$ obtained by integrating the molecular dynamics equations are accepted with probability

$$P_{\rm acc}(\pi, U) = \min\{1, \mathrm{e}^{-\Delta H(\pi,U)}\}, \qquad \Delta H(\pi, U) = H(\tilde{\pi}, \tilde{U}) - H(\pi, U). \tag{2.9}$$

Otherwise, i.e. if the proposed fields are not accepted, the algorithm sets the current fields to $-\pi, U$ [2].

A complete update cycle consists of a momentum rotation, followed by the molecular dynamics evolution and the accept-reject step. Note that the momentum rotation applied in the first step is not undone in the last step if the proposed fields are rejected.

3. Transition probabilities

The discussion of the long-time behaviour of the SMD algorithm later in these notes is based on an analysis of transition probabilities and the theory of discrete Markov processes (for an introduction, see ref. [4], for example).

3.1 Preliminaries

Let $\mathcal{X}$ be the phase space of the theory, equipped with the usual metric topology, σ-algebra of measurable sets and integration measure. The states of the system thus correspond to the points (π, U) of $\mathcal{X}$, and these will often be denoted by a single symbol such as s or s'. A discrete Markov process on $\mathcal{X}$ is described by a transition probability $T(s, A)$ with certain properties, where s and A run over the points and measurable subsets of $\mathcal{X}$, respectively.

The interpretation of $T(s, A)$ as transition probability is as follows. Suppose μ is a probability measure describing the initial distribution of the system and let A, B be measurable subsets of $\mathcal{X}$. Before the transition, the system is thus found in A with probability $\mu(A)$. If it was in A, the probability to find it in B after the transition is then

$$\int_A \mu(\mathrm{d}s)\, T(s, B). \tag{3.1}$$

Clearly, for this interpretation to be meaningful, $T(s, \cdot)$ must be a probability measure for each s and $T(\cdot, A)$ a real-valued bounded measurable function for each A.

Transition probabilities may be considered to be linear operators in the space of signed measures on $\mathcal{X}$, their action on a given signed measure ν being

$$(T\nu)(A) = \int_{\mathcal{X}} \nu(\mathrm{d}s)\, T(s, A). \tag{3.2}$$

Their transpose T^t with respect to the scalar product

$$(f, \nu) = \int_{\mathcal{X}} \nu(\mathrm{d}s) f(s) \tag{3.3}$$

of signed measures ν and bounded measurable functions f on $\mathcal{X}$ acts on the latter according to

$$(T^t f)(s) = \int_{\mathcal{X}} f(s') T(s, \mathrm{d}s'). \tag{3.4}$$

Since $T(s, \mathcal{X}) = 1$ for all s, T^t is a bounded linear transformation with norm 1.

Products of transition probabilities,

$$(T_1 T_2)(s, A) = \int_{\mathcal{X}} T_2(s, \mathrm{d}s') T_1(s', A), \tag{3.5}$$

are again transition probabilities according to the definition given here. The order of the factors on the right of eq. (3.4) is chosen so that $(T_1 T_2)\nu = T_1(T_2\nu)$ for any signed measure ν.

3.2 Transition probability associated with R_ϵ

The operation R_ϵ implements the transition probability

$$T_{R,\epsilon}(s, A) = \frac{1}{\mathcal{Z}_{c_2}} \int_A \mathrm{d}s' \, \exp\left\{ -\frac{1}{2c_2^2} (\pi' - c_1\pi, \pi' - c_1\pi) \right\} \delta(U', U), \tag{3.6}$$

where $s = (\pi, U)$, $s' = (\pi', U')$ and

$$\mathcal{Z}_c = \int \mathrm{d}\upsilon \, \mathrm{e}^{-\frac{1}{2c^2}(\upsilon, \upsilon)}. \tag{3.7}$$

This expression has all the required properties of a transition probability. Moreover, a straightforward calculation shows that

$$T_{R,\epsilon_1} T_{R,\epsilon_2} = T_{R,\epsilon_1 + \epsilon_2} \tag{3.8}$$

i.e. the momentum rotations, when considered to be transformations of probability measures, form a one-dimensional semigroup with group parameter ϵ.

It follows from these remarks that repeated application of the momentum rotations at fixed U leads to the equilibrium distribution

$$\lim_{n\to\infty}(T_{R,\epsilon}^n)(s,A) = \frac{1}{\mathcal{Z}_1}\int_A \mathrm{d}s'\,\mathrm{e}^{-\frac{1}{2}(\pi',\pi')}\delta(U',U) \tag{3.9}$$

for any initial state $s = (\pi, U)$.

3.3 Transition probability associated with I_ϵ

The transition probability representing the molecular dynamics evolution is

$$T_{I,\epsilon}(s,A) = \int_A \mathrm{d}s'\,\delta(\pi'-\tilde{\pi})\delta(U',\tilde{U}) = \begin{cases} 1 & \text{if } (\tilde{\pi},\tilde{U})\in A, \\ 0 & \text{otherwise,} \end{cases} \tag{3.10}$$

where $(\tilde{\pi},\tilde{U}) = I_\epsilon(\pi,U)$ as before. If the accept-reject step is included in the transition, the associated transition probability becomes

$$T_{AI,\epsilon}(s,A) = \int_A \mathrm{d}s'\,\Big\{P_{\mathrm{acc}}(\pi,U)\delta(\pi'-\tilde{\pi})\delta(U',\tilde{U}) + (1-P_{\mathrm{acc}}(\pi,U))\delta(\pi'+\pi)\delta(U',U)\Big\}. \tag{3.11}$$

Noting

$$P_{\mathrm{acc}}(\pi,U)\mathrm{e}^{-H(\pi,U)} = P_{\mathrm{acc}}(-\tilde{\pi},\tilde{U})\mathrm{e}^{-H(-\tilde{\pi},\tilde{U})} \tag{3.12}$$

and that the change of variables from π, U to $-\tilde{\pi},\tilde{U}$ has unit Jacobian, it is straightforward to show that $T_{AI,\epsilon}$ preserves the equilibrium distribution (2.2). Since $T_{R,\epsilon}$ preserves the distribution as well, the same is the case for the transition probability $T_{AI,\epsilon}T_{R,\epsilon}$ representing a complete SMD update cycle.

3.4 SMD Markov chain

Repeated application of the SMD algorithm evolves a given initial probability measure μ_0 according to

$$\mu_0 \to \mu_1 \to \mu_2 \to \ldots, \qquad \mu_{n+1} = T_{AI,\epsilon}T_{R,\epsilon}\mu_n. \tag{3.13}$$

Eventually only the distributions of the gauge field are of interest, i.e. the measures on $\mathcal{X}$ reduced to the sets $P \times A$, where P denotes the full momentum space and A is any measurable subset of position space. Clearly, the same gauge-field distributions are obtained with the modified Markov chain

$$T_{R,\epsilon/2}\mu_n, \quad n = 0, 1, 2, \ldots, \tag{3.14}$$

which is generated by the transition probability

$$T = T_{R,\epsilon/2} T_{AI,\epsilon} T_{R,\epsilon/2}. \tag{3.15}$$

Use has here been made of the semigroup property (3.8) of the momentum rotations. In the following, the convergence properties of the Markov chains with the symmetric transition probabilities (3.15) and

$$\hat{T} = T_{R,\epsilon/2} T_{I,\epsilon} T_{R,\epsilon/2} \tag{3.16}$$

(where the accept-reject step is omitted) will be studied.

4. Convergence analysis: Boundedness

The Markov chains generated by T and $\hat{T}$ do not obviously converge to a stationary distribution. If the transition probability satisfies certain conditions, the convergence of such processes can be rigorously guaranteed. Here an elegant version of Harris' famous theorem [5] recently proved by Hairer and Mattingly [6] will be invoked to establish the convergence of the SMD algorithm.

One of the reasons the algorithm might fail to converge is the fact that the phase space is not compact and that the simulations might therefore run away (for illustration, a simple case where this happens is discussed in appendix A). In this section, the boundedness of the process is proved by constructing a so-called Lyapunov function.

4.1 Strategy

A Lyapunov function for the transition probability T is a measurable function $L : \mathcal{X} \to [0, \infty)$ that goes to infinity at large momenta and which satisfies

$$\int_{\mathcal{X}} T(s, \mathrm{d}s') L(s') \leq r L(s) + K \tag{4.1}$$

for all $s \in \mathcal{X}$ and some constants r, K in the range $0 < r < 1$, $K \geq 0$. In particular, the integral on the left of eq. (4.1) is required to be finite, which already shows that the measure $T(s, \cdot)$ must, in a certain sense, be small at large momenta.

If the simulation starts from some initial configuration s_0, the state after n update cycles with transition probability T is distributed according to the probability measure

$$\mu_n(A) = (T^n)(s_0, A). \tag{4.2}$$

Repeated application of the bound (4.1) then shows that

$$\int_{\mathcal{X}} \mu_n(\mathrm{d}s) L(s) \leq r^n L(s_0) + K \sum_{k=0}^{n-1} r^k. \tag{4.3}$$

The integral on the left thus remains bounded as $n \to \infty$ and there exists a constant C, independent of n, such that

$$\mu_n(\|\pi\| \geq R) \leq \frac{C}{\inf_{\|\pi\| \geq R} L(s)}, \tag{4.4}$$

where $\|\cdot\|$ is any suitable norm in the space of momentum fields. The probability to find the system at large momenta π thus goes to zero with a rate determined by the Lyapunov function. Clearly, since this result holds uniformly in n, the property is inherited by the limiting distribution if it exists.

4.2 Example of a Lyapunov function

Lyapunov functions for the transition probabilities T and $\hat{T}$ are not difficult to find. A particularly simple choice is

$$L(s) = (\pi, \pi) \quad \text{for all} \quad s = (\pi, U). \tag{4.5}$$

This function is measurable, non-negative and goes to infinity at large momenta. A short calculation furthermore shows that, for gauge group SU(N),

$$\int_{\mathcal{X}} T_{R,\epsilon/2}(s, \mathrm{d}s') L(s') = \bar{c}_1^2 L(s) + \bar{c}_2^2 (N^2 - 1) N_l, \tag{4.6}$$

where N_l denotes the number of (active) links on the lattice and $\bar{c}_1, \bar{c}_2$ coincide with c_1, c_2 with ϵ replaced by $\epsilon/2$. The function (4.5) is thus a Lyapunov function for the transition probability representing the momentum rotations.

In the case of the molecular-dynamics evolution,

$$\int_{\mathcal{X}} T_{I,\epsilon}(s, \mathrm{d}s')L(s') = (\tilde{\pi}, \tilde{\pi}), \tag{4.7}$$

where $(\tilde{\pi}, \tilde{U}) = I_\epsilon(\pi, U)$ and $s = (\pi, U)$ as before. The integrators considered are all such that

$$\tilde{\pi} = \pi - \sum_{k=0}^{m} \epsilon_k (\partial S)(U_k) \tag{4.8}$$

for some fixed (usually small) number m of integration steps and step sizes ϵ_k proportional to ϵ. Since the action was assumed to have uniformly bounded derivatives, there exists $M > 0$, independent of ϵ and the lattice volume, such that

$$(\tilde{\pi} - \pi, \tilde{\pi} - \pi) \leq d^2 = \epsilon^2 (N^2 - 1) N_l M \tag{4.9}$$

for all U. As a consequence, the bound

$$(\tilde{\pi}, \tilde{\pi}) \leq (\pi, \pi) + 2d(\pi, \pi)^{1/2} + d^2 \leq \frac{1}{\bar{c}_1^2}(\pi, \pi) + \frac{d^2}{\bar{c}_2^2} \tag{4.10}$$

is obtained.

It is now straightforward to show that (4.5) is a Lyapunov function for the product $T = T_{R,\epsilon/2} T_{I,\epsilon} T_{R,\epsilon/2}$, the constants in eq. (4.1) being

$$r = \bar{c}_1^2, \qquad K = \Big\{ 2\bar{c}_2^2 + \epsilon^2 M \frac{\bar{c}_1^2}{\bar{c}_2^2} \Big\} (N^2 - 1) N_l. \tag{4.11}$$

It is also a Lyapunov function for the transition probability $\hat{T}$, with the same r, K, as can be easily verified starting from eq. (3.11) and using the bound (4.10) again.

4.3 Synthesis

The discussion in this section shows that SMD simulations do not run away, independently of the initial state, the step size ϵ and whether the accept-reject step is included in the algorithm or not. More precisely, at large simulation times, the probability to find the system in a state where $(\pi, \pi) \geq R^2$ is asymptotically bounded by $K/(\bar{c}_2 R)^2$. Note that the proportionality constant

$$\frac{K}{\bar{c}_2^2} = \Big\{ 2 + \frac{M}{\gamma^2} \Big\} (N^2 - 1) N_l + \mathrm{O}(\epsilon) \tag{4.12}$$

remains bounded in the limit $\epsilon \to 0$ and that it scales proportionally to the number of links, just like the expectation value of (π, π) in the canonical distribution (2.2).

5. Convergence analysis: Minorization

In addition to the existence of a Lyapunov function, the version of Harris' theorem established in ref. [6] assumes a minorization condition to hold. More precisely, an integer $n \geq 1$, a radius $R > 2K/(1-r)$, a constant $\alpha > 0$ and a probability measure ν on $\mathcal{X}$ are assumed to exist such that

$$(T^n)(s, \cdot) \geq \alpha \nu(\cdot) \tag{5.1}$$

for all states $s \in \mathcal{X}$ satisfying $L(s) \leq R$. The goal in this section is to show that this condition is fulfilled with measure

$$\nu(A) = \frac{1}{\mathcal{Z}_c} \int_A \mathrm{d}s \, \mathrm{e}^{-\frac{1}{2c^2}(\pi,\pi)}, \qquad c = \bar{c}_2/\sqrt{2}, \tag{5.2}$$

if ϵ is not too large (i.e. such that Assumption 5.4 below holds).

5.1 Properties of the action

In the following, the norm

$$\|U\| = \sup_{x,\mu} \|U(x,\mu)\|_2 \tag{5.3}$$

is used for fields of matrices like the gauge and momentum fields, where $\|\cdot\|_2$ denotes the matrix norm (B.1).

Since the action S is assumed to be a smooth function of the link variables, and since the field space is compact, its first derivative

$$(\partial S)(x, \mu; U) = T^a \partial^a_{x,\mu} S(U) \tag{5.4}$$

satisfies

$$\|(\partial S)(\cdot, U)\| \leq M_1 \tag{5.5}$$

for some constant M_1 independent of U. Moreover, for any two fields U and U',

$$\|(\partial S)(\cdot, U') - (\partial S)(\cdot, U)\| \leq M_2 \|U' - U\|, \tag{5.6}$$

where M_2 is another finite constant. In the case of the Wilson action, for example, $M_1 = 8/g_0^2$ and $M_2 = 24/g_0^2$, where g_0 denotes the bare coupling.

5.2 Bounds on the molecular-dynamics evolution

Starting from any point (π, U) in phase space, the integrator I_ϵ evolves the momentum und gauge field alternatingly in, say, m steps. In the case of the momentum-first algorithms, for example, the steps leading to $(\tilde{\pi}, \tilde{U}) = I_\epsilon(\pi, U)$ are then

$$\pi_1 = \pi, \qquad U_1 = U, \tag{5.7}$$

$$\pi_{k+1}(x, \mu) = \pi_k(x, \mu) - \epsilon_k (\partial S)(x, \mu; U_k), \qquad k = 1, \ldots, m, \tag{5.8}$$

$$U_{k+1}(x, \mu) = \exp\{\tilde{\epsilon}_k \pi_{k+1}(x, \mu)\}\, U_k(x, \mu), \qquad k = 1, \ldots, m-1, \tag{5.9}$$

$$\tilde{\pi} = \pi_{m+1}, \qquad \tilde{U} = U_m, \tag{5.10}$$

where the step sizes $\epsilon_k, \tilde{\epsilon}_k$ are proportional to ϵ with proportionality constants such that

$$\sum_{k=1}^{m} \epsilon_k = \sum_{k=1}^{m-1} \tilde{\epsilon}_k = \epsilon. \tag{5.11}$$

The maximum $\bar{\epsilon}$ of the sums $\sum_{k=1}^{m} |\epsilon_k|$ and $\sum_{k=1}^{m-1} |\tilde{\epsilon}_k|$ may however be slightly larger than ϵ if some of the step sizes are negative.

Lemma 5.1: *The bounds*

$$\|\pi_k - \pi\| \leq \bar{\epsilon} M_1, \tag{5.12}$$

$$\|U_k - U\| \leq \bar{\epsilon} \|\pi\| + \bar{\epsilon}^2 M_1, \tag{5.13}$$

hold for any initial data.

Proof: Equations (5.7),(5.8) imply

$$\pi_k(x, \mu) = \pi(x, \mu) - \sum_{j=1}^{k-1} \epsilon_j (\partial S)(x, \mu; U_j). \tag{5.14}$$

Using the triangle inequality and the uniform boundedness (5.5) of the derivative of the action then immediately leads to the inequality (5.12).

Another application of the triangle inequality furthermore shows that

$$\|\pi_k\| \leq \|\pi\| + \bar{\epsilon} M_1. \tag{5.15}$$

Together with the bound

$$\|U_{k+1} - U_k\| \leq |\tilde{\epsilon}_k| \|\pi_{k+1}\|, \qquad k = 1, \ldots, m-1, \tag{5.16}$$

which is a consequence of the recursion (5.9) and the Lipschitz bound (B.2), this proves the second inequality (5.13). □

Lemma 5.2: *Let (π_k, U_k) and (π'_k, U'_k) be the intermediate fields generated by the recursion (5.7)–(5.10) for a given initial gauge field U and some initial momenta π and π', respectively. Then the bounds*

$$\|\pi'_k - \pi_k\| \leq M_3 \|\pi' - \pi\|, \tag{5.17}$$

$$\|U'_k - U_k\| \leq \bar{\epsilon} M_4 \|\pi' - \pi\|, \tag{5.18}$$

hold uniformly in U, where $M_3 = 1 + \mathrm{O}(\epsilon^2)$ and $\bar{\epsilon} M_4 = \sum_{k=1}^{m-1} |\tilde{\epsilon}_k| \{1 + \mathrm{O}(\epsilon^2)\}$.

Proof: From eqs. (5.6),(5.8),(5.9) and the Lipschitz bound (B.2) one infers that

$$\|\pi'_{k+1} - \pi_{k+1}\| \leq \|\pi'_k - \pi_k\| + |\epsilon_k| M_2 \|U'_k - U_k\|, \tag{5.19}$$

$$\|U'_{k+1} - U_{k+1}\| \leq |\tilde{\epsilon}_k| \|\pi'_{k+1} - \pi_{k+1}\| + \|U'_k - U_k\|. \tag{5.20}$$

Together with the initial values

$$\|\pi'_1 - \pi_1\| = \|\pi' - \pi\|, \qquad \|U'_1 - U_1\| = 0, \tag{5.21}$$

these inequalities recursively yield bounds on the deviations of the fields. The bounds (5.17),(5.18) are now implied by the fact that the ansatz

$$\|\pi'_k - \pi_k\| \leq \{1 + \mathrm{O}(\epsilon^2)\} \|\pi' - \pi\|, \tag{5.22}$$

$$\|U'_k - U_k\| \leq \sum_{j=1}^{k-1} |\tilde{\epsilon}_j| \{1 + \mathrm{O}(\epsilon^2)\} \|\pi' - \pi\|, \tag{5.23}$$

is preserved by the recursion. □

Lemma 5.3: *Let (π_k, U_k), (π'_k, U'_k) and the constants M_3, M_4 be as in lemma 5.2. The fields $\tilde{U}$ and $\tilde{U}'$ obtained at the end of the molecular-dynamics evolution then satisfy*

$$\|\tilde{U}' - \tilde{U}\| \geq \big\{\epsilon - \tfrac{1}{2}\bar{\epsilon}^2 M_3(\|\pi'\| + \|\pi\|) - \bar{\epsilon}^3(M_1 M_3 + M_2 M_4)\big\}\|\pi' - \pi\| \tag{5.24}$$

for any initial values π, π', U of the fields.

Proof: The recursion (5.9) implies

$$\tilde{U}(x,\mu) = \exp\{\tilde{\epsilon}_{m-1}\pi_m(x,\mu)\}\ldots\exp\{\tilde{\epsilon}_1\pi_2(x,\mu)\}U(x,\mu) \tag{5.25}$$

and $\tilde{U}'$ is given by the same formula with π_k replaced by π'_k. A straightforward application of the Lipschitz bound (B.8) then leads to

$$\|\upsilon\| \leq \|\tilde{U}' - \tilde{U}\| + \tfrac{1}{2}\sum_{k=1}^{m-1}\tilde{\epsilon}_k(\|\pi'_{k+1}\| + \|\pi_{k+1}\|)\sum_{j=1}^{m-1}\tilde{\epsilon}_j\|\pi'_{j+1} - \pi_{j+1}\|, \tag{5.26}$$

where

$$\begin{aligned}\upsilon(x,\mu) &= \sum_{k=1}^{m-1}\tilde{\epsilon}_k(\pi'_{k+1} - \pi_{k+1})(x,\mu) \\ &= \epsilon(\pi' - \pi)(x,\mu) - \sum_{k=1}^{m-1}\sum_{j=1}^{k}\tilde{\epsilon}_k\epsilon_j\big\{(\partial S)(x,\mu;U'_j) - (\partial S)(x,\mu;U_j)\big\}.\end{aligned} \tag{5.27}$$

Recalling eqs. (5.15) and (5.17), the last term in eq. (5.26) is seen to be bounded by

$$\big\{\bar{\epsilon}^2 M_1 + \tfrac{1}{2}\bar{\epsilon}(\|\pi'\| + \|\pi\|)\big\}\bar{\epsilon}M_3\|\pi' - \pi\|, \tag{5.28}$$

while

$$\|\upsilon\| \geq (\epsilon - \bar{\epsilon}^3 M_2 M_4)\|\pi' - \pi\|, \tag{5.29}$$

as can be quickly shown using the triangle inequality, (5.6) and subsequently (5.18). Taken together these estimates imply (5.24). □

Assumption 5.4: *The simulation step size ϵ is such that*

$$\bar{\epsilon}^3(2M_1M_3 + M_2M_4) < \epsilon, \tag{5.30}$$

where $M_1, \ldots, M_4$ are the constants appearing in eqs. (5.5), (5.6), (5.17) and (5.18), respectively.

The inequality (5.30) sets an upper limit on ϵ, which is, in practice, not far from the values of ϵ typically chosen in the simulations. If the accept-reject step is omitted, the step size must anyway be fairly small in order to suppress the systematic errors deriving from the inexact integration of the molecular-dynamics equations.

In the following, it will be taken for granted that Assumption 5.4 holds. The constant

$$\lambda = \epsilon - \bar{\epsilon}^3(M_1M_3 + M_2M_4) \tag{5.31}$$

then satisfies $\lambda > \bar{\epsilon}^3 M_1 M_3$ and is thus strictly positive. The symbol

$$\mathcal{B}_\delta = \Big\{ \pi \;\Big|\; \|\pi\| < \frac{\delta}{\bar{\epsilon}^2 M_3} \Big\} \tag{5.32}$$

will now often be used for the ball with radius $\delta/(\bar{\epsilon}^2 M_3)$ at the origin in momentum space.

Lemma 5.5: *Let $\tilde{U}$ be the gauge field obtained by applying the integrator I_ϵ to the initial fields π, U. At fixed U, and if restricted to $\mathcal{B}_\lambda$, the mapping $\pi \to \tilde{U}(\pi)$ is then a diffeomorphism.*

Proof: The map $\pi \to \tilde{U}(\pi)$ is one-to-one on $\mathcal{B}_\lambda$, since

$$\tilde{U}(\pi') = \tilde{U}(\pi), \quad \pi', \pi \in \mathcal{B}_\lambda, \tag{5.33}$$

implies, by Lemma 5.3,

$$0 = \|\tilde{U}(\pi') - \tilde{U}(\pi)\| \geq \big\{ \lambda - \tfrac{1}{2}\bar{\epsilon}^2 M_3 (\|\pi'\| + \|\pi\|) \big\} \|\pi' - \pi\|, \tag{5.34}$$

which can only be true if $\pi' = \pi$ (the expression in the curly bracket is positive).

Since the map $\pi \to \tilde{U}(\pi)$ is smooth, its inverse is smooth if the Jacobian matrix

$$J(x,\mu;y,\nu)^{ab} = -2\,\mathrm{tr}\Big\{ T^a \frac{\partial \tilde{U}(x,\mu)}{\partial \pi^b(y,\nu)} \tilde{U}(x,\mu)^{-1} \Big\} \tag{5.35}$$

is non-singular at all $\pi \in \mathcal{B}_\lambda$. In other words, what needs to be shown is that

$$(J\omega)(x,\mu)^a = \sum_{y,\nu,b} J(x,\mu;y,\nu)^{ab}\omega(y,\nu)^b \tag{5.36}$$

does not vanish, unless the momentum field ω is equal to zero. Noting

$$J\omega = \lim_{s\to 0}\frac{1}{s}\big\{\tilde{U}(\pi+s\omega)-\tilde{U}(\pi)\big\}\tilde{U}(\pi)^{-1} \tag{5.37}$$

and using Lemma 5.3 again, the bound

$$\|J\omega\| \geq \big\{\lambda - \bar{\epsilon}^2 M_3\|\pi\|\big\}\|\omega\| \tag{5.38}$$

is obtained. The Jacobian matrix thus has no zero modes if $\pi \in \mathcal{B}_\lambda$, which proves that the mapping $\pi \to \tilde{U}(\pi)$ restricted to $\mathcal{B}_\lambda$ is a diffeomorphism. □

Lemma 5.6: *Let $\tilde{U}$ be as in Lemma 5.5 and $\omega \in \mathcal{B}_\delta$ for some δ in the range $(0,\lambda)$. Then all gauge fields U' satisfying*

$$\|U' - \tilde{U}(\omega)\| < \frac{(\lambda-\delta)^2}{2\bar{\epsilon}^2 M_3} \tag{5.39}$$

are contained in the image $\mathcal{I}_\lambda(U)$ of $\mathcal{B}_\lambda$ under the mapping $\pi \to \tilde{U}(\pi)$.

Proof: The image $\mathcal{I}_\lambda(U)$ is an open set that contains $\tilde{U}(\omega)$. The set

$$\mathcal{N}_\kappa = \big\{U' \;\big|\; \|U'-\tilde{U}(\omega)\| < \kappa\big\} \tag{5.40}$$

is therefore contained in $\mathcal{I}_\lambda(U)$ if $\kappa > 0$ is sufficiently small. Let

$$\hat{\kappa} = \frac{(\lambda-\delta)^2}{2\bar{\epsilon}^2 M_3}, \tag{5.41}$$

$$\bar{\kappa} = \sup\big\{\kappa \leq \hat{\kappa} \;\big|\; \mathcal{N}_\kappa \subset \mathcal{I}_\lambda(U)\big\}. \tag{5.42}$$

Clearly, $0 < \bar{\kappa} \leq \hat{\kappa}$ and what needs to be proved is that $\bar{\kappa} = \hat{\kappa}$.

Any field $U' \in \mathcal{N}_{\bar{\kappa}}$ is equal to $\tilde{U}(\pi)$ for some $\pi \in \mathcal{B}_\lambda$. Lemma 5.3 then implies

$$\bar{\kappa} \geq \tfrac{1}{2}(\lambda-\delta)\|\pi-\omega\| \tag{5.43}$$

and thus, using the triangle inequality,

$$\|\pi\| \leq \frac{2\bar{\kappa}}{\lambda - \delta} + \frac{\delta}{\bar{\epsilon}^2 M_3}. \tag{5.44}$$

The expression on the right of this inequality is less than $\lambda/(\bar{\epsilon}^2 M_3)$ if $\bar{\kappa} < \hat{\kappa}$.

It follows from this remark that $\bar{\kappa}$ cannot be smaller than $\hat{\kappa}$, since $\mathcal{N}_{\bar{\kappa}}$ would then be contained in $\mathcal{I}_\rho(U)$ for some $\rho < \lambda$. Since $\pi \to \tilde{U}(\pi)$ is a diffeomorphism on $\mathcal{B}_\lambda$, this would imply that the closure of $\mathcal{N}_{\bar{\kappa}}$ and consequently $\mathcal{N}_\kappa$ for some $\kappa > \bar{\kappa}$ would be contained in $\mathcal{I}_\lambda(U)$, thus contradicting the definition of $\bar{\kappa}$. □

5.3 Lower bound on the transition probability

The transition probability T is explicitly given by

$$T(s, A) = \frac{1}{\mathcal{Z}_{\bar{c}_2}^2} \int_A \mathrm{d}s' \int \mathrm{d}\upsilon\, \delta(U', \tilde{U}) \\ \times \exp\Big\{ -\frac{1}{2\bar{c}_2^2} \big[(\pi' - \bar{c}_1\tilde{\upsilon}, \pi' - \bar{c}_1\tilde{\upsilon}) + (\upsilon - \bar{c}_1\pi, \upsilon - \bar{c}_1\pi) \big] \Big\}, \tag{5.45}$$

where $s = (\pi, U)$, $s' = (\pi', U')$ and $(\tilde{\upsilon}, \tilde{U}) = I_\epsilon(\upsilon, U)$. Since the integrand in this formula is non-negative, the restriction of the integration over the momentum υ to $\mathcal{B}_\lambda$ provides a lower bound on the integral. In view of Lemma 5.5, one can then change variables from υ to $\tilde{U}$, the new domain of integration being $\mathcal{I}_\lambda(U)$.

With the change of variables, the integration measure transforms according to

$$\mathrm{d}\upsilon \propto \mathrm{d}\tilde{U}\, |\det J|^{-1}, \tag{5.46}$$

where J is the Jacobian matrix (5.35) (with π replaced by υ). The proportionality constant in eq. (5.46) depends on the number of link variables and the gauge group, but not on the fields. Following the lines after (5.36), the bound (5.18) implies

$$\|J\omega\| \leq \bar{\epsilon} M_4 \|\omega\| \tag{5.47}$$

and thus a uniform lower bound on the Jacobian factor in eq. (5.46). As a result, one obtains

$$T(s, A) \geq \text{constant} \times \int_A \mathrm{d}s'\, \theta(U' \in \mathcal{I}_\lambda(U)) \\ \times \exp\Big\{ -\frac{1}{2\bar{c}_2^2} \big[(\pi' - \bar{c}_1\tilde{\upsilon}, \pi' - \bar{c}_1\tilde{\upsilon}) + (\upsilon - \bar{c}_1\pi, \upsilon - \bar{c}_1\pi) \big] \Big\}, \tag{5.48}$$

where $v \in \mathcal{B}_\lambda$ is such that $\tilde{U}(v) = U'$ (the function $\theta(\cdot)$ is equal to 1 if its argument is true and 0 otherwise). Since v and $\tilde{v}$ are bounded by a constant that does not depend on U or U', the exponent in eq. (5.48) may be simplified,

$$T(s, A) \geq \text{constant} \times \int_A \mathrm{d}s'\, \theta(U' \in \mathcal{I}_\lambda(U)) \exp\Big\{-\frac{1}{\bar{c}_2^2}(\pi', \pi') - \frac{\bar{c}_1^2}{\bar{c}_2^2}(\pi, \pi)\Big\}, \quad (5.49)$$

at the expense of a modification of the constant in front of the integral.

By substituting the inequality (5.49) in

$$(T^n)(s, A) = \int_{\mathcal{X}} T(s, \mathrm{d}s_1) T(s_1, \mathrm{d}s_2) \dots T(s_{n-1}, A) \quad (5.50)$$

and integrating over the intermediate momenta, the bound

$$\begin{aligned}(T^n)(s, A) \geq C_n \int_A \mathrm{d}s' \int \mathrm{d}V_1 \dots \mathrm{d}V_{n-1} \\ \times \theta(U' \in \mathcal{I}_\lambda(V_{n-1}))\theta(V_{n-1} \in \mathcal{I}_\lambda(V_{n-2})) \dots \theta(V_1 \in \mathcal{I}_\lambda(U)) \\ \times \exp\Big\{-\frac{1}{\bar{c}_2^2}(\pi', \pi') - \frac{\bar{c}_1^2}{\bar{c}_2^2}(\pi, \pi)\Big\} \quad (5.51)\end{aligned}$$

is obtained for the powers of the transition probability, where the constants $C_n > 0$ depend on n but not on s or A.

5.4 Proof of the minorization bound (5.1)

In view of the reversibility of the integrator I_ϵ, one expects that there is a non-zero probability for the SMD process to return to any initial state after two update steps. The following Lemma confirms this (any unexplained notation is as in the previous subsections).

Lemma 5.7: *There exist positive constants ρ and h such that*

$$\int_{\mathcal{X}} \mathrm{d}V\, \theta(U' \in \mathcal{I}_\lambda(V))\theta(V \in \mathcal{I}_\lambda(U)) \geq h\, \theta(\|U' - U\| < \rho) \quad (5.52)$$

for all U, U'.

Proof: Assumption 5.4 guarantees that there exists a real number σ in the range

$$\bar{\epsilon}^3 M_1 M_3 < \sigma < \lambda. \quad (5.53)$$

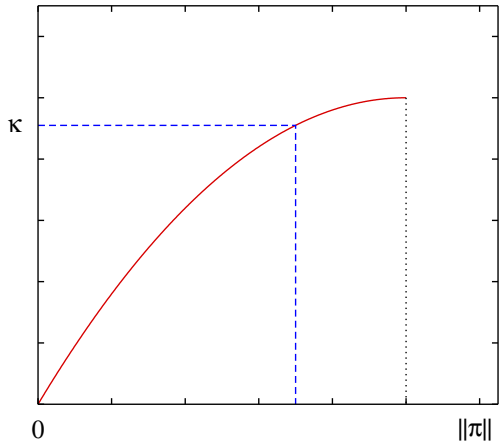

Fig. 1. Plot of the expression on the right of the inequality (5.56) as a function of $\|\pi\|$. The constant κ is always below the maximum of the expression, which is reached at $\|\pi\| = \lambda/(\bar{\epsilon}^2 M_3)$ (dotted line).

Lemma 5.6 then implies that all gauge fields V satisfying

$$\|V - \tilde{U}(0)\| < \kappa, \qquad \kappa = \frac{\lambda^2 - \sigma^2}{2\bar{\epsilon}^2 M_3}, \tag{5.54}$$

are contained in $\mathcal{I}_\lambda(U)$. The integral

$$\int_{\mathcal{X}} \mathrm{d}V\, \theta(U' \in \mathcal{I}_\lambda(V))\theta(\|V - \tilde{U}(0)\| < \kappa). \tag{5.55}$$

thus provides a lower bound on the integral (5.52).

For any field V in the range (5.54) there exists a unique momentum field $\pi \in \mathcal{B}_\lambda$ such that $V = \tilde{U}(\pi)$. Actually, by Lemma 5.3, the bound

$$\kappa \geq \left\{\lambda - \tfrac{1}{2}\bar{\epsilon}^2 M_3 \|\pi\|\right\}\|\pi\| \tag{5.56}$$

holds and consequently

$$\|\pi\| \leq \frac{\lambda - \sigma}{\bar{\epsilon}^2 M_3} \tag{5.57}$$

(see fig. 1).

Now if $(\tilde{\pi}, \tilde{U}) = I_\epsilon(\pi, U)$, as before, the combination of Lemma 5.1 and the bound (5.57) implies $\tilde{\pi} \in \mathcal{B}_\delta$, where δ is any constant in the range

$$\lambda - \sigma + \bar{\epsilon}^3 M_1 M_3 < \delta < \lambda \tag{5.58}$$

The choice (5.53) of σ guarantees that the range (5.58) is not empty and that δ is positive. Setting

$$\rho = \frac{(\lambda - \delta)^2}{2\bar{\epsilon}^2 M_3} \tag{5.59}$$

and invoking Lemma 5.6 again, it then follows that any U' satisfying

$$\|U' - \tilde{V}(-\tilde{\pi})\| < \rho \tag{5.60}$$

is contained in $\mathcal{I}_\lambda(V)$. Moreover, the reversibility of the integrator I_ϵ implies

$$\tilde{V}(-\tilde{\pi}) = U, \tag{5.61}$$

i.e. the range (5.60) of U' is the same for all V.

The integral (5.56) is thus seen to be bounded from below by

$$\int_{\mathcal{X}} \mathrm{d}V \, \theta(\|U' - U\| < \rho)\theta(\|V - \tilde{U}(0)\| < \kappa). \tag{5.62}$$

After passing to the integration variables $W = V\tilde{U}(0)^{-1}$, the integral (5.62) decomposes into a product of link integrals over the range $\|W(x,\mu)\|_2 < \kappa$ and is therefore equal to some constant $h > 0$. □

The minorization bound (5.1) may now be established by recalling (5.51) and choosing $n = 2m$ to be even. Lemma 5.7 allows the integrations over $V_1, V_3, \ldots, V_{2m-1}$ to be performed, the new lower bound being

$$\begin{aligned}(T^n)(s, A) \geq h^m C_n \int_A \mathrm{d}s' \int \mathrm{d}V_2 \mathrm{d}V_4 \ldots \mathrm{d}V_{2m-2} \\ \times \theta(\|U' - V_{2m-2}\| < \rho) \ldots \theta(\|V_2 - U\| < \rho) \\ \times \exp\Big\{-\frac{1}{\bar{c}_2^2}(\pi', \pi') - \frac{\bar{c}_1^2}{\bar{c}_2^2}(\pi, \pi)\Big\}.\end{aligned} \tag{5.63}$$

According to Lemma B.2, one can choose $m \geq 1$ so that, for any U and U', there exists a sequence $\hat{V}_0, \hat{V}_2, \ldots, \hat{V}_{2m}$ of gauge fields satisfying

$$\hat{V}_0 = U, \qquad \hat{V}_{2m} = U', \tag{5.64}$$

$$\|\hat{V}_{2k} - \hat{V}_{2k-2}\| < \delta \quad \text{for all} \quad k = 1, \ldots, n, \tag{5.65}$$

where $\delta = \rho/3$. Now let $W_2, \ldots, W_{2m-2}$ be independent gauge fields satisfying

$$\|W_{2k} - 1\| < \delta \tag{5.66}$$

and set

$$V_k = W_k \hat{V}_k, \quad k = 2, 4, \ldots, 2m-2. \tag{5.67}$$

Another application of the triangle inequality and the Lipschitz bound (B.2) shows that

$$\|V_2 - U\| < \rho, \qquad \|U' - V_{2m-2}\| < \rho, \tag{5.68}$$

$$\|V_k - V_{k-2}\| < \rho \quad \text{for all} \quad k = 2, 4, \ldots, 2m-2. \tag{5.69}$$

and thus leads to the bound

$$(T^n)(s, A) \geq h^m C_n \int_A \mathrm{d}s' \prod_{k=1}^{m-1} \int_{\mathcal{X}} \mathrm{d}W_{2k} \theta(\|W_{2k} - 1\| < \delta)$$

$$\times \exp\Big\{-\frac{1}{\bar{c}_2^2}(\pi', \pi') - \frac{\bar{c}_1^2}{\bar{c}_2^2}(\pi, \pi)\Big\}. \tag{5.70}$$

The integrals over $W_2, \ldots, W_{2m-2}$ do not depend on U or U' and therefore merely contribute to the constant on the right of the inequality. Finally, since $(\pi, \pi) = L(s) \leq R$, the minorization bound (5.1) follows.

5.5 Remarks

According to Harris' theorem (Theorem 1.2 in ref. [6]), the convergence of the stochastic process with transition probability T^n to a unique stationary distribution is guaranteed by the existence of the Lyapunov function (4.5) and the minorization

bound (5.1). The theorem also yields an upper bound on the exponential autocorrelation time, but as is often the case with rigorous estimates, the bound is likely to be a poor estimate of the true autocorrelation time.

It is straightforward to show that the process with transition probability T converges to the same stationary distribution as the one with transition probability T^n. Moreover, the minorization bound (5.1) also holds for the transition probability $\hat{T}^n$ that includes the acceptance-rejection step.

6. SMD algorithm in QCD

The results obtained in the previous sections extend to QCD with any number and type of quarks. In particular, convergence to a unique stationary distribution is again guaranteed if Assumption 5.4 holds, with the constants $M_1, \ldots, M_4$ determined by the gauge action and the integration scheme for the molecular-dynamics equations. In order to show this, it suffices to consider the theory with two flavours of mass-degenerate quarks and the simplest possible choice for the pseudo-fermion action as representative case.

6.1 Simulation algorithm

As usual the quarks may be included in the theory by introducing a pseudo-fermion field ϕ with action

$$S_{\mathrm{pf}}(U, \phi) = (D^{-1}\phi, D^{-1}\phi), \tag{6.1}$$

where D denotes the (massive) lattice Dirac operator. The Dirac operator, its inverse D^{-1} and the derivatives of D^{-1} with respect to the gauge field are assumed to be bounded in the norm by constants that do not depend on the latter. This condition is not satisfied in the case of the Wilson–Dirac operator, but can be met by including a small twisted mass term in the quark action as infrared regulator [7].

The SMD algorithm now essentially proceeds as described in sect. 2, with action

$$S(U, \phi) = S_{\mathrm{G}}(U) + S_{\mathrm{pf}}(U, \phi) \tag{6.2}$$

in place of the pure gauge action S_{G}. A new element of the algorithm is the rotation

$$R_{\phi,\epsilon} : \ \phi \to c_1\phi + c_2 D\eta \tag{6.3}$$

of the pseudo-fermion field, which is applied together with the momentum rotation. In eq. (6.3), η is a Gaussian random pseudo-fermion field with mean zero and unit variance. For simplicity, the friction parameter γ and thus the coefficients c_1 and c_2 are here taken to be the same as in the momentum rotation (2.3), but the algorithm remains correct, and its performance may improve, if a different value is chosen.

In the molecular-dynamics step, the momentum and the gauge fields are updated as described in subsect. 2.2, while the pseudo-fermion field is held fixed. For the integrator one can choose a nested one, where the different driving forces are integrated with different speeds (see ref. [7], for example).

6.2 Lyapunov function

Since the space of pseudo-fermion fields is non-compact, a useful Lyapunov function must depend on the pseudo-fermion field. In the following lines, the functions

$$L_1(s) = (\pi, \pi), \tag{6.4}$$

$$L_2(s) = (\phi, \phi)^2, \tag{6.5}$$

$$L_3(s) = (\phi, \phi), \tag{6.6}$$

are considered and a Lyapunov function is constructed in the form of a linear combination of these with positive coefficients.

Proceeding as in subsect. 4.2, it is straightforward to derive the bounds

$$\int_{\mathcal{X}} T(s, \mathrm{d}s') L_i(s') \leq \sum_{j=1}^{3} A_{ij} L_j(s) + K_i, \qquad K_i \geq 0, \tag{6.7}$$

where the matrix A is of the triangular form

$$A = \begin{pmatrix} c_1 & A_{12} & A_{13} \\ 0 & c_1^4 & A_{23} \\ 0 & 0 & c_1^2 \end{pmatrix}. \tag{6.8}$$

with positive coefficients A_{12}, A_{13}, A_{23}. A has a left eigenvector $v = (v_1, v_2, v_3)$ with components

$$v_1 = 1, \qquad v_2 = \frac{A_{12}}{c_1 - c_1^4}, \qquad v_3 = \frac{A_{13} + v_2 A_{23}}{c_1 - c_1^2}, \tag{6.9}$$

and eigenvalue c_1. The function

$$L(s) = \sum_{i=1}^{3} v_i L_i(s) \tag{6.10}$$

is hence a Lyapunov function, i.e. satisfies eq. (4.1) with $r = c_1$ and $K = \sum_{i=1}^{3} v_i K_i$.

Since the coefficients v_i are all positive, the inequality $L(s) \leq R$ implies uniform upper bounds on π and ϕ. The SMD algorithm thus cannot run away to arbitrarily large momentum or pseudo-fermion fields.

6.3 Minorization bound

The minorization bound (5.1) remains valid in full QCD with

$$\nu(A) \propto \int_A \mathrm{d}s \, \exp\Big\{-\frac{1}{2c^2}(\pi,\pi) - \frac{d^2}{c^2}(\phi,\phi)\Big\}, \tag{6.11}$$

where c is as before and

$$d = \sup_U \|D^{-1}\|_2. \tag{6.12}$$

As before, the proof of the bound requires Assumption 5.4, *with $M_1, \ldots, M_4$ determined in the pure gauge theory*, to hold. The convergence of the SMD algorithm to a unique stationary state then follows from Harris' theorem.

First note that the transition probability corresponding to the rotation (6.3) is given by

$$\begin{aligned} T_{R_\phi,\epsilon}(s,A) = \frac{1}{\mathcal{Z}_{\phi,c_2}(U)} \int_A \mathrm{d}s' \, \delta(\pi' - \pi)\delta(U',U) \\ \times \exp\Big\{-\frac{1}{c_2^2}\big(D^{-1}(\phi' - c_1\phi), D^{-1}(\phi' - c_1\phi)\big)\Big\}, \end{aligned} \tag{6.13}$$

where $s = (\pi, U, \phi)$, $s' = (\pi', U', \phi')$ and

$$\mathcal{Z}_{\phi,c}(U) = \int \mathrm{d}\phi \, \mathrm{e}^{-\frac{1}{c^2}(D^{-1}\phi, D^{-1}\phi)} \tag{6.14}$$

is the pseudo-fermion partition function. In QCD the integrand in eq. (5.45) thus gets modified by the factor

$$\frac{1}{\mathcal{Z}_{\phi,\bar{c}_2}(U')\mathcal{Z}_{\phi,\bar{c}_2}(U)} \exp\Big\{-\frac{1}{\bar{c}_2^2}\big[(D^{-1}(\phi' - \bar{c}_1\chi), D^{-1}(\phi' - \bar{c}_1\chi)) + (D^{-1}(\chi - \bar{c}_1\phi), D^{-1}(\chi - \bar{c}_1\phi))\big]\Big\} \tag{6.15}$$

and the intermediate pseudo-fermion field χ is to be integrated over.

By replacing the pseudo-fermion partition functions by a suitable constant and by restricting the integral over χ to the ball

$$\|\chi\|_2 < \bar{\rho} \tag{6.16}$$

for some constant $\bar{\rho} > 0$, a lower bound on the transition probability is obtained. Moreover, the exponential factor in eq. (6.15) is then bounded from below by

$$\exp\Big\{-\frac{d^2}{c^2}\big[(\phi', \phi') + \bar{c}_1^2(\phi, \phi)\big]\Big\} \tag{6.17}$$

times a constant. One is free to set $\bar{\rho}$ to a very small value, in which case the fermion forces practically have no effect on the molecular-dynamics evolution that leads from U to $\tilde{U}$ in eq. (5.45). In particular, the Lemmas proved in sect. 5 continue to hold with constants $M_1, \ldots, M_4$ slightly modified by terms of order $\bar{\rho}^2$. From here on the proof of the minorization bound proceeds as in the pure gauge theory (subsect. 5.4).

Appendix A

SMD simulations of the standard ϕ^4 lattice theory, using integrators of the kind described in subsect. 2.2, are unstable if the acceptance-rejection step is omitted. The full algorithm is probably stable, but it can stagnate over very long periods of simulation time.

A.1 One-dimensional model

In order to show this, it suffices to consider a single coordinate q with momentum p and Hamilton function

$$H(p,q) = \tfrac{1}{2}p^2 + S(q), \tag{A.1}$$

$$S(q) = \tfrac{1}{2}q^2 + \tfrac{1}{4}q^4. \tag{A.2}$$

With the leapfrog integrator, the SMD algorithm amounts to applying the update cycle

$$p \to c_1 p + c_2 \upsilon - \tfrac{1}{2}\epsilon S'(q), \tag{A.3}$$

$$q \to q + \epsilon p, \tag{A.4}$$

$$p \to p - \tfrac{1}{2}\epsilon S'(q), \tag{A.5}$$

where υ is a Gaussian random variable with mean zero and unit variance.

A.2 Long-time instability

In the first step (A.3) of the update cycle, the momentum p is not protected from becoming very large. The probability for this to happen depends on the integration step size and is exponentially suppressed at small ϵ. However, exceptionally large momenta do occur in sufficiently long runs.

An exceptionally large momentum shifts the coordinate q in the second step (A.4) to some large value and the momentum is then amplified in the last step (A.5) by a term proportional to the third power of q. Clearly, p and q are rapidly driven to infinity in this way.

Large momenta are made less probable by the acceptance-rejection step but are not rigorously excluded to occur. Once a large momentum (and consequently large coordinate) is accepted, the algorithm stagnates for a very long time, because the energy deficit ΔH will always be unacceptably large.

A.3 Remarks

In theories with non-compact state space and/or actions with unbounded first derivatives, the SMD algorithm may not be a viable choice of algorithm for the reasons given in this appendix.

Although the simulation will be less likely to run away or to stagnate, the problem cannot be cured by decreasing the integration step size or by replacing the leapfrog integrator by a higher-order integrator. The Langevin algorithm incidentally runs away too if the well-known explicit integrators are used. Implicit integrators such as the symplectic mid-point rule however behave differently and may yield stable algorithms.

Appendix B

B.1 Matrix norm

Starting from the standard L_2 norm $\|v\|_2$ of vectors v with n complex components, the norm of any complex $n \times n$ matrix A is defined by

$$\|A\|_2 = \sup_{\|v\|_2=1} \|Av\|_2. \tag{B.1}$$

This norm has the usual properties of a decent matrix norm. Moreover, $\|U\|_2 = 1$ and $\|AU\|_2 = \|UA\|_2 = \|A\|_2$ if U is unitary.

B.2 Properties of the exponential map

All result derived in this subsection apply to the gauge group SU(N), $N \geq 2$, and its Lie algebra $\mathfrak{su}(N)$.

Lemma B.1: *The Lipschitz bounds*

$$\left\|\mathrm{e}^X - \mathrm{e}^Y\right\|_2 \leq \|X - Y\|_2, \tag{B.2}$$

$$\left\|\mathrm{e}^X - \mathrm{e}^Y\right\|_2 \geq \left\{1 - \tfrac{1}{2}(\|X\|_2 + \|Y\|_2)\right\} \|X - Y\|_2, \tag{B.3}$$

hold for any $X, Y \in \mathfrak{su}(N)$.

Proof: First note that

$$\left\|\mathrm{e}^X - \mathrm{e}^Y\right\|_2 = \left\|\mathrm{e}^X \mathrm{e}^{-Y} - 1\right\|_2 \tag{B.4}$$

since e^Y is unitary. The identity

$$\mathrm{e}^X \mathrm{e}^{-Y} - 1 = \int_0^1 \mathrm{d}s\, \mathrm{e}^{sX}(X - Y)\mathrm{e}^{-sY} \tag{B.5}$$

and the subadditivity of the matrix norm then lead to the bound (B.2). Similarly, (B.3) is obtained using

$$X - Y = \mathrm{e}^X \mathrm{e}^{-Y} - 1 \\ - \int_0^1 \mathrm{d}s \left\{ (\mathrm{e}^{sX} - 1)(X - Y)\mathrm{e}^{-sY} + (X - Y)(\mathrm{e}^{-sY} - 1) \right\}, \tag{B.6}$$

the triangle inequality and the bound (B.2). □

A straightforward generalization of the arguments used in the proof of Lemma B.1 lead to the bounds

$$\left\| \mathrm{e}^{X_1} \dots \mathrm{e}^{X_n} - \mathrm{e}^{Y_1} \dots \mathrm{e}^{Y_n} \right\|_2 \leq \sum_{k=1}^n \|X_k - Y_k\|_2 \tag{B.7}$$

and

$$\left\| \mathrm{e}^{X_1} \dots \mathrm{e}^{X_n} - \mathrm{e}^{Y_1} \dots \mathrm{e}^{Y_n} \right\|_2 \geq \|Z\|_2 - \tfrac{1}{2} \sum_{k=1}^n (\|X_k\|_2 + \|Y_k\|_2) \sum_{j=1}^n \|X_j - Y_j\|_2, \tag{B.8}$$

$$Z = \sum_{k=1}^n (X_k - Y_k), \tag{B.9}$$

which hold for arbitrary sequences $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$ of elements of $\mathfrak{su}(N)$.

Lemma B.2: *Let $\delta > 0$ and $U \in \mathrm{SU}(N)$. Then there exists an integer $n \geq 1$ and a sequence $U_0, U_1, \dots, U_n$ of group elements such that*

$$U_0 = 1, \quad U_n = U, \tag{B.10}$$

$$\|U_k - U_{k-1}\|_2 < \delta \quad \text{for all} \quad k = 1, \dots, n. \tag{B.11}$$

Moreover, n can be chosen to be the same for all U.

Proof: U can be diagonalized with eigenvalues

$$\lambda_k = \mathrm{e}^{i\theta_k}, \quad |\theta_k| \leq \pi, \quad \sum_{k=1}^N \theta_k = 2\pi m, \tag{B.12}$$

where m is an integer, which may, without loss of generality, be assumed to be non-negative. At least m of the angles θ_k must then be positive and after subtracting 2π from m of them, the sum of the angles will vanish and the magnitude of all angles is less than 2π.

There thus exists $X \in \mathfrak{su}(N)$ such that

$$U = \mathrm{e}^X, \quad \|X\|_2 < 2\pi. \tag{B.13}$$

Setting

$$n = \lceil 2\pi/\delta \rceil, \tag{B.14}$$

$$U_k = \mathrm{e}^{\frac{k}{n}X}, \quad k = 0, 1, \ldots, n, \tag{B.15}$$

the bounds (B.2) and (B.13) show that all statements made in the lemma hold with this choice of n and sequence $U_0, \ldots, U_n$ of group elements. □

References

[1] A. M. Horowitz, *Stochastic quantization in phase space*, Phys. Lett. 156B (1985) 89; *The second order Langevin equation and numerical simulations*, Nucl. Phys. B280 [FS18] (1987) 510; *A generalized guided Monte Carlo algorithm*, Phys. Lett. B268 (1991) 247

[2] K. Jansen, C. Liu, *Kramers equation algorithm for simulations of QCD with two flavors of Wilson fermions and gauge group SU(2)*, Nucl. Phys. B453 (1995) 375 [E: *ibid.* B459 (1996) 437]

[3] I. P. Omelyan, I. M. Mryglod, R. Folk, *Symplectic analytically integrable decomposition algorithms: classification, derivation, and application to molecular dynamics, quantum and celestial mechanics simulations*, Comp. Phys. Commun. 151 (2003) 272

[4] M. Hairer, *Convergence of Markov processes*, lecture notes (2016) [available at `http://www.hairer.org/Teaching.html`]

[5] T. E. Harris, *The existence of stationary measures for certain Markov processes*, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954-1955, vol II. 113, University of California Press, Berkeley and Los Angeles, 1956 [`https://projecteuclid.org/euclid.bsmsp/1200502009`]

[6] M. Hairer, J. C. Mattingly, *Yet another look at Harris' ergodic theorem for Markov chains*, in: Seminar on Stochastic Analysis, Random Fields and Applications VI, eds. R. C. Dalang et al., Progr. Probab. 63, Spring Science & Business Media, 2011 [arXiv:0810.2777]

[7] M. Lüscher, S. Schaefer, *Lattice QCD with open boundary conditions and twisted-mass reweighting*, Comput. Phys. Commun. 184 (2013) 519