Hybrid Monte Carlo algorithm with fat link fermion actions

Waseem Kamleh, Derek B. Leinweber, and Anthony G. Williams
Special Research Centre for the Subatomic Structure of Matter and Department of Physics, University of Adelaide, Adelaide, SA 5005, Australia
(Received 21 March 2004; published 23 July 2004)

The use of APE smearing or other blocking techniques in lattice fermion actions can provide many advantages. There are many variants of these fat link actions in lattice QCD currently, such as flat link irrelevant clover (FLIC) fermions. The FLIC fermion formalism makes use of the APE blocking technique in combination with a projection of the blocked links back into the special unitary group. This reunitarization is often performed using an iterative maximization of a gauge invariant measure. This technique is not differentiable with respect to the gauge field and thus prevents the use of standard Hybrid Monte Carlo simulation algorithms. The use of an alternative projection technique circumvents this difficulty and allows the simulation of dynamical fat link fermions with standard HMC and its variants. The necessary equations of motion for FLIC fermions are derived, and some initial simulation results are presented. The technique is more general however, and is straightforwardly applicable to other smearing techniques or fat link actions.

DOI: 10.1103/PhysRevD.70.014502 PACS number(s): 11.15.Ha, 12.38.Gc

Recent advances in computing power and in lattice QCD (in particular, overlap fermions [1]) have allowed simulations at sufficiently light quark masses to see that the behavior of quenched QCD can differ from the true theory both qualitatively and quantitatively in the chiral regime [2,3]. As it is in the chiral regime where the difference from the quenched approximation will be highlighted, we would like to simulate at light quark masses in dynamical QCD. This is an extremely computationally expensive endeavour. Ideally, this would be done using overlap fermions, although large scale dynamical overlap simulations challenge the limits of current computing power, to say the least.

Fat link irrelevant clover (FLIC) fermions have shown a number of promising advantages over standard actions, including improved convergence properties [4] and $O(a)$ improved scaling without the need for nonperturbative tuning [5,6]. Furthermore, a reduced exceptional configuration problem has allowed efficient access to the light quark mass regime in the quenched approximation [6], where recent studies have highlighted deviations from the true theory [2,3]. As interest shifts to focus on dynamical QCD, be it (truly) unquenched, or partially quenched, one might hope that the excellent behavior at light quark mass displayed by FLIC fermions will carry over from the quenched theory to the unquenched one. This brings us to the issue of generating dynamical gauge field configurations with the fermionic determinant being that of the FLIC action. Brief accounts of this work were presented last year [7,8]. Recently, an alternative proposal for another type of smearing scheme that is differentiable has also appeared [9].

I. HYBRID MONTE CARLO

The standard technique for simulating dynamical fermions has for some time now been hybrid Monte Carlo (HMC) [10]. It is exact, ergodic and is $O(V^{5/4})$ (using the standard leapfrog integration scheme), that is, it scales almost linearly with the lattice volume V (for other integration schemes see Refs. [11,12]). In order to introduce our notation and a framework for our technique, we briefly review the HMC algorithm for generating dynamical gauge field configurations.

We wish to generate an ensemble $\{U_i\}$ of (statistically independent) representative gauge fields distributed according to the probability distribution

$$\rho(U_i) = e^{-S_{\text{eff}}[U_i]}, \tag{1}$$

where the effective action for full QCD

$$S_{\text{eff}}[U] = S_{\text{g}}[U] - \ln\det D_{\text{f}}[U], \tag{2}$$

$$\det D_{\text{f}} = \int \mathcal{D}\overline{\psi}\mathcal{D}\psi e^{-\int d^4x \overline{\psi}(x) D_{\text{f}} \psi(x)} \tag{3}$$

is obtained from the standard action

$$S[U,\overline{\psi},\psi] = S_{\text{g}}[U] + S_{\text{f}}[U,\overline{\psi},\psi], \tag{4}$$

$$S_{\text{f}} = \int d^4x \overline{\psi}(x) D_{\text{f}}[U] \psi(x), \tag{5}$$

by integrating the fermionic degrees of freedom $\overline{\psi},\psi$ of the functional integral

$$\langle \mathcal{O} \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}U\mathcal{D}\overline{\psi}\mathcal{D}\psi \mathcal{O}[U,\overline{\psi},\psi] e^{-S[U,\overline{\psi},\psi]}. \tag{6}$$

For Wilson-like fermions in the physical region (away from exceptional configurations) $\det D_{\text{f}}$ is real and positive. Hence if $M = D_{\text{f}}^{\dagger} D_{\text{f}}$, then $\det M = \det D_{\text{f}}^2$. For complex fields ϕ we have

$$\frac{1}{\det M^{-1}} = \int \mathcal{D}\phi^{\dagger}\mathcal{D}\phi e^{-\int d^4x \phi^{\dagger}(x) M^{-1} \phi(x)}, \tag{7}$$

and as $\det M = 1/\det M^{-1}$, we thus obtain the fermion determinant for two flavor QCD given in terms of an auxiliary

pseudofermion field ϕ, so called because it is complex (bosonic) rather than Grassmannian. For $\det D_{\mathrm{f}}$ real and positive it is an identity that

$$\det D_{\mathrm{f}} = \sqrt{\det D_{\mathrm{f}} \det D_{\mathrm{f}}^{\dagger}} = \sqrt{\det M} = \det \sqrt{M}, \tag{8}$$

and hence to simulate an odd number of sea quark flavors, it is possible to use $\det\sqrt{M}$ [13].

For HMC, the four-dimensional quantum lattice theory is embedded in a classical five-dimensional system through the introduction of a fictitious (simulation) time, the (classical) evolution parameter τ. The gauge field U is associated with its (fictitious) conjugate momenta P, and the (classical) five-dimensional system is described by the Hamiltonian

$$\mathcal{H}[U,P] = \sum_{x,\mu} \frac{1}{2} \mathrm{Tr}\, P_{\mu}(x)^2 + S_{\mathrm{eff}}[U]. \tag{9}$$

For Gaussian distributed P the expectation value of an observable is unaffected by the 5D kinetic energy

$$\langle \mathcal{O} \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}P \mathcal{D}U \mathcal{O}[U] e^{-\mathcal{H}[U,P]}, \tag{10}$$

$$\mathcal{Z} = \int \mathcal{D}P \mathcal{D}U e^{-\mathcal{H}[U,P]}. \tag{11}$$

Given U, a new gauge field U' is generated by the update $U \to U'$, which consists of the following.

(i) *Refreshment*. Sample P from a Gaussian ensemble, $\rho(P) \propto e^{-(1/2)\mathrm{Tr}\, P^2}$. Generate a pseudofermionic background field ϕ according to $\rho(\phi) \propto e^{-\phi^{\dagger} M^{-1} \phi}$.

(ii) *Molecular dynamics trajectory*. Integrate Hamilton's equations of motion to deterministically evolve (U,P) along a phase space trajectory to (U',P').

(iii) *Metropolis step*. Accept or reject the new configuration (P',U') with probability $\rho(U \to U') = \min(1, e^{-\Delta\mathcal{H}})$, $\Delta\mathcal{H} = \mathcal{H}[U'] - \mathcal{H}[U]$.

The discretized equations of motion are derived by requiring that the Hamiltonian be conserved along the trajectory $\mathrm{d}\mathcal{H}/\mathrm{d}\tau = 0$. The following discretized equations of motion then approximately conserve $\mathcal{H}$ for small step sizes $\Delta\tau$:

$$U_{\mu}(x, \tau + \Delta\tau) = U_{\mu}(x,\tau) \exp(\mathrm{i}\Delta\tau P_{\mu}(x,\tau)), \tag{12}$$

$$P_{\mu}(x, \tau + \Delta\tau) = P_{\mu}(x,\tau) - U_{\mu}(x,\tau) \frac{\delta S_{\mathrm{eff}}}{\delta U_{\mu}(x,\tau)}. \tag{13}$$

In our implementation, we evaluate the matrix exponential directly through diagonalization, rather than expanding it.

FLIC fermions [4] are clover-improved fermions where the irrelevant operators are constructed using APE smeared links [14,15]. As with other efficient updating algorithms, HMC makes use of the variation of the action with respect to the links $\delta S/\delta U$ in order that the proposed configurations have high acceptance rates. Previously, it has not been clear how to perform HMC with fermion actions that make use of the APE blocking technique in combination with a projection of the blocked links back into the special unitary group. This reunitarization is often performed using an iterative maximization of a gauge invariant measure, and this choice of reuniterization is the source of the difficulty. The problem is that the iterative technique is not differentiable with respect to the gauge field and thus it is not possible to calculate $\delta S/\delta U$, which is necessary for the equations of motion above. In the next section we consider an alternative technique and show that is does not suffer from this problem, allowing the simulation of dynamical fat link fermions with standard HMC (and its variants).

II. SU(3) PROJECTION

The APE smeared links $U_{\mu}^{(n)}(x)$ present in the FLIC fermion action are constructed from $U_{\mu}(x)$ by performing n smearing sweeps, where in each sweep we first perform an APE blocking step

$$V_{\mu}^{(j)}(x)[U^{(j-1)}] = (1-\alpha)$$  $$+ \frac{\alpha}{6} \sum_{\nu \neq \mu}$$,

(14)

followed by a projection back into $SU(3)$, $U_{\mu}^{(j)}(x) = \mathcal{P}[V_{\mu}^{(j)}(x)]$. Frequently, the projection is performed using an algorithm which updates $U^{(j)}$ iteratively in order to maximize the following gauge invariant measure:

$$U_{\mu}^{(j)}(x) \in \{ U' \in SU(3) \,|\, \mathrm{Re}\,\mathrm{Tr}(U' V_{\mu}^{(j)\dagger}(x)) \ \text{is maximal} \}. \tag{15}$$

We refer to this projection technique as MaxReTr projection. While this projection minimizes the local action [16], as we mentioned earlier it is not differentiable with respect to $U_{\mu}(x)$ and hence not suitable for use in HMC.

Now, given any matrix X, then $X^{\dagger}X$ is hermitian and may be diagonalized. Then it is possible (for $\det X \neq 0$) to define a matrix

$$W = X \frac{1}{\sqrt{X^{\dagger}X}} \tag{16}$$

whose spectrum lies on the complex unit circle and is hence unitary [$w(z) = z/z^{*}z$ is the complex version of the sign function]. Furthermore, W possesses the same gauge transformation properties as X. This is easily seen. Let $X_{\mu}(x)$ transform as

$$X_{\mu}(x) \to G(x) X_{\mu}(x) G^{\dagger}(x + e_{\mu}), \tag{17}$$

then

$$X_{\mu}^{\dagger}(x) \to G(x + e_{\mu}) X_{\mu}^{\dagger}(x) G^{\dagger}(x) \tag{18}$$

and hence

$$X_{\mu}^{\dagger}(x) X_{\mu}(x) \to G(x + e_{\mu}) X_{\mu}^{\dagger}(x) X_{\mu}(x) G^{\dagger}(x + e_{\mu}). \tag{19}$$

TABLE I. The mean link u_0 for a single configuration as a function of number of APE smearing sweeps at $\alpha=0.7$, for the two different projection methods. The boldface indicates significant digits which match. The configuration is a dynamical gauge field with DBW2 glue and FLIC sea fermions, at $\beta=8.0$, $\kappa=0.1280$ (a $=0.17$ fm, $m_q=200$ MeV).

Sweep	Unit Circle	MaxReTr
0	**0.866138301214314**	**0.866138301214314**
1	**0.9603**13394813806	**0.9603**48747275940
2	**0.980**735000838119	**0.980**751346847750
3	**0.9883**84926461589	**0.9883**93707639555
4	**0.99210**3013943516	**0.99210**7844842705
5	**0.99418**2852413813	**0.99418**5532052157
6	**0.99545**7365275018	**0.99545**8835653863
7	**0.99629**3668622924	**0.99629**4454083006
8	**0.996878**305318083	**0.996878**710433084

Noting that $[X^\dagger_\mu(x)X_\mu(x)]^{-1/2}$ has the same transformation properties as $X^\dagger_\mu(x)X_\mu(x)$ it is then straightforward to see that

$$W_\mu(x)\to G(x)W_\mu(x)G^\dagger(x+e_\mu), \tag{20}$$

as required.

Given the unitary matrix W, we can construct another matrix

$$W'=\frac{1}{\sqrt[3]{\det W}}W, \tag{21}$$

which is special unitary. Earlier work [17] has incorrectly omitted the cube root. As there are three different complex roots, we have a $\mathbb{Z}_3$ ambiguity which we break by choosing the principal value of the cube root.[1] In selecting the principal value, the projected matrices lie closest to those given by the MaxReTr method, and are hence smoother. The mean plaquette is closer to unity thus minimizing the action. We refer to this technique for projecting $X_\mu(x)$ into the special unitary group as unit circle projection.

The two methods produce smeared links that are different but numerically close [according to the usual matrix norm $\|A\|=\sqrt{\lambda_{\max}(A^\dagger A)}$]. Using the mean link as a measure of the smoothness of the smeared gauge field, Table I indicates that the two methods presented here produce equally smooth gauge fields.

While numerically the two methods may be nearly equivalent, unit circle projection possesses a significant advantage over MaxReTr projection. The matrix inverse square root function can be approximated by a rational polynomial (whose poles lie on the imaginary axis) [18,19], $W[X]\approx W_k[X]$,

$$W_k[X]\equiv d_0X(X^\dagger X+c_{2n})\sum_{l=1}^{k}\frac{b_l}{X^\dagger X+c_l}, \tag{22}$$

where the formula for the coefficients d_0,b_l,c_l can be found Ref. [19]. This approximation is differentiable in a matrix sense for all X for which the inverse square root can be defined. This means that we can construct $\delta S/\delta U$ for fermion actions which involve unit circle projection, and hence it is a reuniterization method which is compatible with HMC.

III. EQUATIONS OF MOTION

Having now defined the APE smearing prescription (with projection) in a differentiable closed form, we proceed to derive the equations of motion necessary for the use of the HMC algorithm with FLIC fermions.

A. Mathematical preliminaries

The equations of motion are derived using multi-variate calculus. To make the derivation simple and provide an understanding of how best to implement the equations efficiently, we develop some appropriate mathematical tools. Using index notation, we define a (minimal) set of tensor operations (including differentiation) such that we can perform the derivation in an index free language.

The derivative of a real-valued function $f[A]$ with respect to the matrix A is a rank 2 type (1,1) tensor (distinguishing contravariant and covariant indices)

$$\left[\frac{\partial f}{\partial A}\right]^i_{\ j}=\frac{\partial}{\partial A^j_{\ i}}f[A]. \tag{23}$$

The derivative of a matrix-valued function $M[A]$ with respect to the matrix A is a rank 4 type (2,2) tensor

$$\left[\frac{\partial M}{\partial A}\right]^{i\ \ k}_{\ j\ \ l}=\frac{\partial}{\partial A^j_{\ k}}M[A]^i_{\ l}. \tag{24}$$

The set of type (m,n) tensors $\mathcal{T}^m_n$ forms a vector space. We define the outer product $\otimes$: $\mathcal{T}^1_1\times\mathcal{T}^1_1\to\mathcal{T}^2_2$ as

$$(A\otimes B)^{i\ k}_{\ j\ l}=A^i_{\ j}B^k_{\ l}. \tag{25}$$

Noting carefully the index ordering, define the "direct" product $\oplus$: $\mathcal{T}^1_1\times\mathcal{T}^1_1\to\mathcal{T}^2_2$ as

$$(A\oplus B)^{i\ k}_{\ j\ l}=A^k_{\ j}B^i_{\ l}. \tag{26}$$

Given a scalar function $f[B]$ and a matrix function $B[A]$ the (scalar-matrix) chain rule states

$$\frac{\partial f}{\partial A}=\frac{\partial f}{\partial B}\star\frac{\partial B}{\partial A}, \tag{27}$$

where we define the contraction induced by the chain rule as the (rank 2) star product $\star$: $\mathcal{T}^1_1\times\mathcal{T}^2_2\to\mathcal{T}^1_1$ with

$$(A\star T)^i_{\ l}=A^j_{\ k}T^{i\ k}_{\ j\ l}. \tag{28}$$

[1]For complex z, the principal value of the cube root satisfies $-\pi/3<\arg\sqrt[3]{z}<\pi/3$. For purely real z, we choose $\sqrt[3]{z}$ to be real.

Given two matrix functions $M[B]$ and $B[A]$, the (matrix-matrix) chain rule states

$$\frac{\partial M}{\partial A} = \frac{\partial M}{\partial B} \star \frac{\partial B}{\partial A}, \tag{29}$$

where we define the contraction induced by this chain rule as the (rank 4) star product $\star$: $\mathcal{T}_2^2 \times \mathcal{T}_2^2 \to \mathcal{T}_2^2$ with

$$(S \star T)^{i\ k}_{\ j\ l} = S^{i\ \ n}_{\ m\ \ l} T^{m\ \ k}_{\ j\ \ n}. \tag{30}$$

It is interesting to note that the star product induces an algebra structure on the vector space of type (2,2) tensors, that is, $(\mathcal{T}_2^2, +, \star)$ is an algebra with multiplicative identity $I \otimes I$.

We define juxtaposition for $A \in \mathcal{T}_1^1, T \in \mathcal{T}_2^2$ by the contractions

$$(AT)^{i\ k}_{\ j\ l} = A^i_{\ m} T^{m\ k}_{\ j\ l}, \tag{31}$$

$$(TA)^{i\ k}_{\ j\ l} = T^{i\ k}_{\ j\ m} A^m_{\ l}. \tag{32}$$

All our derivatives will be derived from the basic matrix differentiation rule. Given matrices M, A, B, C then for $M = ABC$ we have

$$\frac{\partial M}{\partial B} = A \otimes C. \tag{33}$$

An immediate consequence of this is that

$$\frac{\partial M}{\partial M} = I \otimes I. \tag{34}$$

The (scalar-matrix) product rule is

$$\frac{\partial}{\partial A}(fM) = \frac{\partial f}{\partial A} \oplus M + f \frac{\partial M}{\partial A}. \tag{35}$$

The (matrix-matrix) product rule is

$$\frac{\partial}{\partial A}(XY) = X \frac{\partial Y}{\partial A} + \frac{\partial X}{\partial A} Y, \tag{36}$$

which is easily shown to imply the identity

$$\frac{\partial X^{-1}}{\partial A} = -X^{-1} \frac{\partial X}{\partial A} X^{-1}. \tag{37}$$

In the following sections we will make use of the identity

$$A \star (B \oplus C) = (A \cdot B) C, \tag{38}$$

where $A \cdot B = A^i_{\ j} B^j_{\ i}$. Additionally, of particular numerical importance is the identity

$$A \star (B \otimes C) = BAC, \tag{39}$$

which has two major benefits. It allows us to evaluate two matrix multiplications instead of an outer product (computational saving), hence enabling us to implement the equations of motion without having to store any tensor fields (memory saving).

B. Standard derivatives

The equations of motion for FLIC fermions are derived starting from the equations for the standard clover fermion action [20,21]. We divide the effective action into its gauge part and pseudofermionic part

$$S_{\text{eff}} = S_{\text{g}} + S_{\text{pf}}. \tag{40}$$

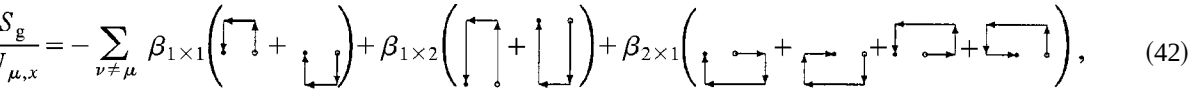
We reformulate some standard results in terms of the mathematics of the previous section. We will adopt a more convenient notation for quantities with a lattice site index x, using a subscript $U_{\mu,x}$ rather than $U_\mu(x)$. The matrix products of link variables are often denoted diagrammatically.

For a plaquette plus rectangle improved gauge action

$$S_{\text{g}} = \sum_{x,\mu<\nu} \text{Re Tr}\{\beta_{1\times1}[1 - U_{\mu\nu}(x)] + \beta_{2\times1}[1 - R^{2\times1}_{\mu\nu}(x)] + \beta_{1\times2}[1 - R^{1\times2}_{\mu\nu}(x)]\}, \tag{41}$$

we have

$$\frac{\partial S_{\text{g}}}{\partial U_{\mu,x}} = -\sum_{\nu\neq\mu} \beta_{1\times1}\left(\quad + \quad\right) + \beta_{1\times2}\left(\quad + \quad\right) + \beta_{2\times1}\left(\quad + \quad + \quad + \quad\right), \tag{42}$$



where the filled circles indicate the point x. The coefficients $\beta_{m\times n} = \frac{1}{3}\beta c_{m\times n}$ depend on the choice of gauge action. For Lüscher-Weisz glue [22], $c_{1\times1} = \frac{5}{3}, c_{2\times1} = c_{1\times2} = -\frac{1}{12}$. For DBW2 glue [23–26], we choose coefficients that are normalized such that $c_{1\times1} = 1$.

The pseudofermionic action is $S_{\text{pf}} = -\Sigma_x \phi_x^\dagger \eta_x$, where $\eta = (D^\dagger D)^{-1}\phi$, hence by equations (36) and (37) we have

$$\frac{\partial S_{\text{pf}}}{\partial U_{\mu,x}} = \phi^\dagger (D^\dagger D)^{-1} \left(D^\dagger \frac{\partial D}{\partial U_{\mu,x}} + \frac{\partial D^\dagger}{\partial U_{\mu,x}} D \right) (D^\dagger D)^{-1} \phi. \tag{43}$$

Setting $\chi = D\eta$, we obtain

$$\frac{\partial S_{\rm pf}}{\partial U_{\mu,x}} = \chi^\dagger \frac{\partial D}{\partial U_{\mu,x}} \eta + \eta^\dagger \frac{\partial D^\dagger}{\partial U_{\mu,x}} \chi. \tag{44}$$

Now, the FLIC action is explicitly given by

$$\begin{aligned}(D_{\rm flic}\psi)_x = &-\frac{1}{2}\sum_\mu \left(\frac{U^{\rm fl}_{\mu,x}}{u^{\rm fl}_0} - \gamma_\mu \frac{U_{\mu,x}}{u_0} \right) \psi_{x+\mu} \\ &+ \left(\frac{U^{\rm fl\dagger}_{\mu,x}}{u^{\rm fl}_0} + \gamma_\mu \frac{U^\dagger_{\mu,x}}{u_0} \right) \psi_{x-\mu} \\ &+ \left(4 + m - \frac{1}{4u^{\rm fl4}_0} \sigma_{\mu\nu} F^{\rm cl}_{\mu\nu,x} \right) \psi_x \end{aligned} \tag{45}$$

and contains three terms, the Dirac term (constructed with standard links), the Wilson term and the clover term (using fat links for $F^{\rm cl}_{\mu\nu,x}$, and setting $\sigma_{\mu\nu} = \frac{1}{2}[\gamma_\mu, \gamma_\nu]$). Hence we may decompose the pseudofermionic derivative into three terms also. The first comes from the Dirac term

$$\frac{\partial S^{\rm d}_{\rm pf}}{\partial U_{\mu,x}} = \frac{1}{2u_0} {\rm Tr}_{\rm spin}(\eta^\dagger_x \otimes \gamma_\mu \chi_{x+\mu} + \chi^\dagger_x \otimes \gamma_\mu \eta_{x+\mu}), \tag{46}$$

while the Wilson and clover terms only explicit dependence is on the smeared links

$$\frac{\partial S^{\rm w}_{\rm pf}}{\partial U^{\rm fl}_{\mu,x}} = -\frac{1}{2u^{\rm fl}_0} {\rm Tr}_{\rm spin}(\eta^\dagger_x \otimes \chi_{x+\mu} + \chi^\dagger_x \otimes \eta_{x+\mu}), \tag{47}$$

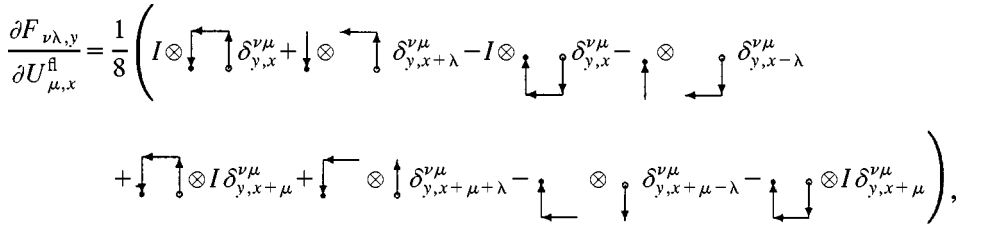
$$\begin{aligned}\frac{\partial S^{\rm cl}_{\rm pf}}{\partial U^{\rm fl}_{\mu,x}} = &-\frac{1}{4u^{\rm fl4}_0} {\rm Tr}_{\rm spin}\Big(\eta^\dagger_y \sigma_{\nu\lambda} \frac{\partial F_{\nu\lambda,y}}{\partial U^{\rm fl}_{\mu,x}} \chi_y \\ &+ \chi^\dagger_y \sigma_{\nu\lambda} \frac{\partial F_{\nu\lambda,y}}{\partial U^{\rm fl}_{\mu,x}} \eta_y \Big), \end{aligned} \tag{48}$$

where the vector outer product defines a matrix $(\eta \otimes \chi^\dagger)^i{}_j = \eta^i \chi^*_j$. The one loop clover term is given by $F_{\nu\lambda,y} = \frac{1}{8}(C_{\nu\lambda,y} - C^\dagger_{\nu\lambda,y})$, where

$$\begin{aligned}C_{\nu\lambda,y} = &U_{(+\nu)(+\lambda),y} + U_{(+\lambda)(-\nu),y} + U_{(-\lambda)(+\nu),y} \\ &+ U_{(-\nu)(-\lambda),y}, \end{aligned} \tag{49}$$

and $U_{(\pm\nu)(\pm\lambda),y}$ indicates the plaquette starting at y, oriented in the $\nu-\lambda$ plane, with the first (second) link in the direction indicated by the first (second) index. When differentiating with respect to $U^{\rm fl}_{\mu,x}$, any terms where y lies further away from x than $x \pm \nu \pm \lambda$ will be zero. Further, noting that the derivative is zero unless either $\nu = \mu$ or $\lambda = \mu$ and $\nu \neq \lambda$ we can without loss of generality choose $\mu = \nu$. Letting $\mu = \nu$ be in the horizontal direction and λ be in the transverse (vertical) direction, the contribution to the derivative due to the clover term is

$$\frac{\partial F_{\nu\lambda,y}}{\partial U^{\rm fl}_{\mu,x}} = \frac{1}{8}\Bigg(I \otimes \delta^{\nu\mu}_{y,x} + \otimes \delta^{\nu\mu}_{y,x+\lambda} - I \otimes \delta^{\nu\mu}_{y,x} - \otimes \delta^{\nu\mu}_{y,x-\lambda} + \otimes I\, \delta^{\nu\mu}_{y,x+\mu} + \otimes \delta^{\nu\mu}_{y,x+\mu+\lambda} - \otimes \delta^{\nu\mu}_{y,x+\mu-\lambda} - \otimes I\, \delta^{\nu\mu}_{y,x+\mu} \Bigg), \tag{50}$$



where the filled circles indicate the point x, and the point y is located at the start (end) of the diagrams that lie on the left (right) side of the outer product, as can be deduced from the Kronecker δ's.

C. Smeared link derivatives

Now, having constructed the explicit derivatives of $S_{\rm pf}$ with respect to the thin and fat links, the total derivative with respect to the thin links is

$$\frac{{\rm d}S_{\rm pf}}{{\rm d}U_{\mu,x}} = \frac{\partial S_{\rm pf}}{\partial U_{\mu,x}} + \frac{\partial S_{\rm pf}}{\partial U^{\rm fl}_{\nu,y}} \star \frac{{\rm d}U^{\rm fl}_{\nu,y}}{{\rm d}U_{\mu,x}}. \tag{51}$$

If we have performed n sweeps of APE smearing to form the fat links, then the right-hand term is constructed through n applications of the chain rule

$$\frac{{\rm d}S}{{\rm d}U^{(j-1)}_{\mu,x}} = \frac{\partial S}{\partial U^{(j)}_{\nu,y}} \star \frac{\partial U^{(j)}_{\nu,y}}{\partial U^{(j-1)}_{\mu,x}} + \frac{\partial S}{\partial U^{(j)\dagger}_{\nu,y}} \star \frac{\partial U^{(j)\dagger}_{\nu,y}}{\partial U^{(j-1)}_{\mu,x}}, \tag{52}$$

until we arrive at ${\rm d}S/{\rm d}U^{(0)}_{\mu,x}$. We note here that the partial derivative with respect to a forward (backward) link only picks up terms that contain the forward (backward) link, and not its conjugate [that is, U and $U^\dagger$ are considered independent with regard to partial differentiation, see Eqs. (14),(66), (67)]. For the sake of both computational efficiency and simplicity, this chain rule is itself composed of several chain rules, and hence evaluated in several steps. Each step corre-

sponds to a step in the APE smearing process, but we go through them in reverse order.

The final step in the APE smearing process (with unit circle projection) is

$$U^{(n)}_{\mu,x} = \frac{1}{\det W^{(n)\ -1/3}_{\mu,x}} W^{(n)}_{\mu,x}. \tag{53}$$

Therefore the first chain rule corresponds to this step

$$\begin{aligned}\frac{\partial S}{\partial W_{\mu,x}} &= \frac{\partial S}{\partial U^{(n)}_{\mu,x}} \star \frac{\partial U^{(n)}_{\mu,x}}{\partial W_{\mu,x}} \\ &= \frac{\partial S}{\partial U^{(n)}_{\mu,x}} \star \Bigg(-\frac{1}{3} \det W^{-4/3}_{\mu,x} \\ &\quad \times \frac{\partial \det W_{\mu,x}}{\partial W_{\mu,x}} \oplus W_{\mu,x} + \det W^{-1/3}_{\mu,x} I \otimes I \Bigg),\end{aligned} \tag{54}$$

where

$$\det A = \epsilon^{ijk} A^1{}_i A^2{}_j A^3_k \tag{55}$$

and hence denoting the permutations $\pi_i = (i \bmod 3) + 1, \pi_i^2 = \pi_{\pi_i}$,

$$\frac{\partial \det A}{\partial A^i{}_j} = \epsilon_{jlm} A^{\pi_i}{}_l A^{\pi_i^2}{}_m. \tag{56}$$

There are several chain rules that correspond to

$$W^{(n)}_{\mu,x} = V^{(n)}_{\mu,x} (V^{(n)\dagger}_{\mu,x} V^{(n)}_{\mu,x})^{-1/2}. \tag{57}$$

For the first, we define $H_{\mu,x} = V^\dagger_{\mu,x} V_{\mu,x}$. Then

$$\frac{\partial S}{\partial H_{\mu,x}} = \frac{\partial S}{\partial W_{\mu,x}} \star \frac{\partial W_{\mu,x}}{\partial H_{\mu,x}}. \tag{58}$$

Using

$$W_{\mu,x} \approx d_0 V_{\mu,x} (H_{\mu,x} + c_0) \sum_{l=1}^{k} \frac{b_l}{H_{\mu,x} + c_l}, \tag{59}$$

we have

$$\begin{aligned}\frac{\partial W_{\mu,x}}{\partial H_{\mu,x}} &= d_0 V_{\mu,x} \Bigg(I \otimes \sum_{l=1}^{k} \frac{b_l}{H_{\mu,x} + c_l} - (H_{\mu,x} + c_0) \\ &\quad \times \sum_{l=1}^{k} b_l \frac{1}{H_{\mu,x} + c_l} \otimes \frac{1}{H_{\mu,x} + c_l} \Bigg).\end{aligned} \tag{60}$$

We can then construct

$$\begin{aligned}\frac{\partial S}{\partial V_{\mu,x}} &= \frac{\partial S}{\partial W_{\mu,x}} \star \frac{\partial W_{\mu,x}}{\partial V_{\mu,x}} + \frac{\partial S}{\partial H_{\mu,x}} \star \frac{\partial H_{\mu,x}}{\partial V_{\mu,x}} \\ &= \frac{\partial S}{\partial W_{\mu,x}} \star (I \otimes H^{-1/2}_{\mu,x}) + \frac{\partial S}{\partial H_{\mu,x}} \star (V^\dagger_{\mu,x} \otimes I)\end{aligned} \tag{61}$$

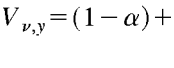
and also

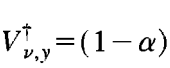
$$\begin{aligned}\frac{\partial S}{\partial V^\dagger_{\mu,x}} &= \frac{\partial S}{\partial W^\dagger_{\mu,x}} \star \frac{\partial W^\dagger_{\mu,x}}{\partial V^\dagger_{\mu,x}} + \frac{\partial S}{\partial H_{\mu,x}} \star \frac{\partial H_{\mu,x}}{\partial V^\dagger_{\mu,x}} \\ &= \frac{\partial S}{\partial W^\dagger_{\mu,x}} \star \Big(H^{-\frac{1}{2}}_{\mu,x} \otimes I \Big) + \frac{\partial S}{\partial H_{\mu,x}} \star (I \otimes V_{\mu,x}).\end{aligned} \tag{62}$$

Last, we make use of the chain rule

$$\frac{\partial S}{\partial U^{(n-1)}_{\mu,x}} = \frac{\partial S}{\partial V_{\nu,y}} \star \frac{\partial V_{\nu,y}}{\partial U^{(n-1)}_{\mu,x}} + \frac{\partial S}{\partial V^\dagger_{\nu,y}} \star \frac{\partial V^\dagger_{\nu,y}}{\partial U^{(n-1)}_{\mu,x}}, \tag{63}$$

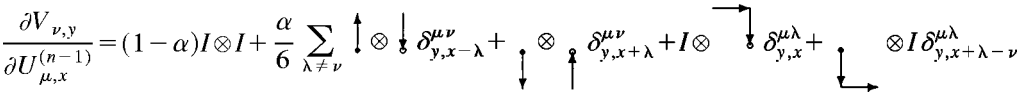
where

$$V_{\nu,y} = (1-\alpha) + \frac{\alpha}{6} \sum_{\lambda \neq \nu}$$  (64)

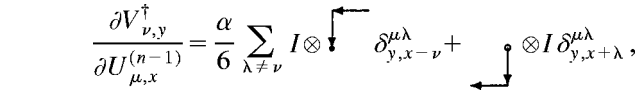
$$V^\dagger_{\nu,y} = (1-\alpha) + \frac{\alpha}{6} \sum_{\lambda \neq \nu}$$  (65)

and y is indicated by the filled circle.

It is then straightforward to show that

$$\frac{\partial V_{\nu,y}}{\partial U^{(n-1)}_{\mu,x}} = (1-\alpha) I \otimes I + \frac{\alpha}{6} \sum_{\lambda \neq \nu}$$  (66)

and

$$\frac{\partial V^\dagger_{\nu,y}}{\partial U^{(n-1)}_{\mu,x}} = \frac{\alpha}{6} \sum_{\lambda \neq \nu} I \otimes$$  (67)

where in these diagrams the filled circle indicates the point y. Hence,

$$\frac{\partial S}{\partial U^{(n-1)}_{\mu,x}} = (1-\alpha)\frac{\partial S}{\partial V_{\mu,x}} + \frac{\alpha}{6}\sum_{\nu\neq\mu}\frac{\partial S}{\partial V_{\mu,x-\nu}}\star \otimes + \frac{\partial S}{\partial V_{\mu,x+\nu}}\star \otimes$$
$$+ \frac{\partial S}{\partial V_{\nu,x}}\star I\otimes + \frac{\partial S}{\partial V_{\nu,x+\mu-\nu}}\star \otimes I + \frac{\partial S}{\partial V^{\dagger}_{\mu,x-\nu}}\star I\otimes + \frac{\partial S}{\partial V^{\dagger}_{\mu,x+\mu}}\star \otimes I, \tag{68}$$

where now x is indicated by the filled circle.

Having constructed the total derivative of the action with respect to $U_{\mu,x}$, we can calculate the variation of S with respect to the gauge field (noting $U^{\dagger}=U^{-1}$),

$$\frac{\delta S}{\delta U} = \frac{\mathrm{d}S}{\mathrm{d}U} + \frac{\mathrm{d}S}{\mathrm{d}U^{\dagger}}\star\frac{\partial U^{\dagger}}{\partial U} = \frac{\mathrm{d}S}{\mathrm{d}U} - U^{\dagger}\frac{\mathrm{d}S}{\mathrm{d}U^{\dagger}}U^{\dagger}, \tag{69}$$

and hence the necessary equations of motion (12) and (13). It is numerically efficient to make use of the fact that

$$\frac{\mathrm{d}S}{\mathrm{d}U^{\dagger}} = \left(\frac{\mathrm{d}S}{\mathrm{d}U}\right)^{\dagger}. \tag{70}$$

IV. SIMULATION RESULTS

We have implemented the equations above within a standard two-flavor HMC, with multiple time scales. The standard leapfrog integration scheme is used. The BiCGStab algorithm is used for matrix inversion. Expensive pseudofermion momenta updates are performed at a larger step size $\Delta\tau=\Delta\tau_{\mathrm{pf}}$ and the cheaper gauge momenta updates are performed more often, $\Delta\tau_{\mathrm{g}}=(1/n)\,\Delta\tau_{\mathrm{pf}}$, for some integer n. Molecular dynamics trajectories are of unit length, $n_{\mathrm{md}}\Delta\tau=1$. In particular, we have implemented a modified version of the Ritz algorithm to diagonalise arrays of 3×3 matrices in parallel. This routine is used in the SU(3) projection step, and is also used to calculate the matrix exponentials that are needed in other parts of the algorithm, avoiding the need to use polynomial approximations to the exponential. This means that the accuracy of the exponential in Eq. (12) does not depend upon the step size $\Delta\tau$.

An eighth order Zolotarev approximation to the inverse square root is used to approximate $W_{\mu,x}$ in unit circle projection. We find that the spectral range at this order is ample. In smooth gauge backgrounds it is easily shown that unit circle projection is well defined, that is, $\det V^{\dagger}_{\mu,x}V_{\mu,x}>0$.

If we assume a smoothness condition $\|1-U_{\mu\nu,x}\|\leqslant\epsilon\ \forall x,\mu,\nu$, then we have a lower bound

$$V^{\dagger}_{\mu,x}V_{\mu,x}\geqslant 1-2\alpha\epsilon-\alpha^2\epsilon^2.$$

To prove this, we note that APE blocking may be written in terms of the plaquette field

$$V_{\mu,x} = U_{\mu,x}\left(1-\frac{\alpha}{6}\sum_{\pm\nu\neq\mu}(1-U^{\dagger}_{\mu\nu,x})\right). \tag{71}$$

Define $Z=\Sigma_{\pm\nu\neq\mu}(1-U^{\dagger}_{\mu\nu,x})$, then

$$V^{\dagger}_{\mu,x}V_{\mu,x} = 1-\frac{\alpha}{6}(Z+Z^{\dagger})+\frac{\alpha^2}{36}Z^{\dagger}Z. \tag{72}$$

As $\|Z\|\leqslant 6\epsilon$, we then have

$$\lambda_{\min}(V^{\dagger}_{\mu,x}V_{\mu,x})\geqslant 1-2\alpha\epsilon-\alpha^2\epsilon^2, \tag{73}$$

which is strictly positive for small enough ϵ.

While the smeared link equations of motions are complex, our implementation evaluates them efficiently due to the optimizations that can be performed through the calculus we constructed earlier. At large sea quark masses the code already spends over 90% of its time in the BiCGStab inversion required to calculate $\eta=(D^{\dagger}D)^{-1}\phi$, and as the quark mass decreases this fraction increases. So as is standard, the generation of dynamical gauge fields is dominated by the matrix inversion.

Simulation results are presented in Table II. Simulations for both Lüscher-Weisz and DBW2 glue are performed, at different β and κ values to conduct an initial exploration of the parameter space. Pion masses and lattice spacings are obtained from 20 configurations for each β and κ given. We observe that to obtain similar lattice spacings to our quenched lattices, we must work at smaller β values for our unquenched gauge fields. This is in accord with the expected renormalization of the strong coupling constant confirmed in previous dynamical fermion studies. As a result the mean link u_0 is significantly further from unity for unquenched lattices than quenched lattices, although the effect is more pronounced for the Lüscher-Weisz glue than the DBW2 glue. Our results also confirm the well known fact that as β and the quark mass are decreased one must reduce the step size $\Delta\tau$ in order to maintain a constant acceptance rate. Although an exact comparison is difficult, for a given step size and quark mass, our acceptance rates obtained compare well with standard simulations (see, for example, Ref. [27]).

It is a simple exercise to apply our results to generate gauge fields with dynamical FLIC overlap quarks, although this would be extremely computationally intensive. The availability of HMC as a simulation algorithm for dynamical FLIC fermions is significant, as it scales almost linearly with the lattice volume V, whereas previously there were only

TABLE II. Simulation parameters and results for various dynamical simulations. The parameters are the gauge coupling, hopping parameter, gauge action, step size, and psuedofermion to gauge step size ratio. The results given are the mean link, lattice spacing (in fm, obtained from r_0 via the static quark potential) and pion mass (in GeV). Two degenerate flavors of FLIC sea quarks are used, with either Lüscher-Weisz (IMP) glue or DBW2 glue. These results are obtained from 20 $12^3\times 24$ configurations. Simulations are done using multiple time step HMC with trajectories of unit length.

β	κ	S_{gauge}	$\Delta\tau$	$\Delta\tau_{\text{pf}}/\Delta\tau_{\text{g}}$	ρ_{acc}	u_0	a	m_π
3.6	0.1347	IMP	0.0143	2	0.55	0.8226	0.247(9)	0.702
3.7	0.1340	IMP	0.0147	2	0.64	0.8338	0.218(4)	0.680
3.8	0.1332	IMP	0.0151	2	0.65	0.8443	0.180(2)	0.738
3.9	0.1310	IMP	0.0200	2	0.66	0.8534	0.153(2)	0.834
3.9	0.1325	IMP	0.0156	2	0.55	0.8540	0.146(2)	0.702
4.0	0.1301	IMP	0.0200	2	0.66	0.8614	0.132(2)	0.906
4.0	0.1318	IMP	0.0161	2	0.64	0.8625	0.121(2)	0.799
4.1	0.1283	IMP	0.0200	2	0.75	0.8680	0.114(1)	1.088
4.1	0.1305	IMP	0.0166	2	0.70	0.8685	0.104(1)	0.668
4.2	0.1246	IMP	0.0200	2	0.86	0.8736	0.107(1)	1.496
4.2	0.1266	IMP	0.0200	2	0.80	0.8738	0.097(1)	1.346
4.3	0.1253	IMP	0.0200	2	0.83	0.8788	0.091(1)	1.574
4.4	0.1255	IMP	0.0200	2	0.88	0.8836	0.086(1)	1.411
4.5	0.1253	IMP	0.0200	2	0.83	0.8878	0.075(1)	1.657
4.6	0.1254	IMP	0.0200	2	0.84	0.8916	0.072(1)	1.617
7.0	0.1315	DBW2	0.0152	2	0.74	0.8344	0.252(6)	0.780
7.0	0.1345	DBW2	0.0156	2	0.68	0.8352	0.233(8)	0.673
7.5	0.1310	DBW2	0.0156	2	0.79	0.8516	0.206(3)	0.779
8.0	0.1305	DBW2	0.0161	2	0.73	0.8663	0.168(2)	0.764
8.5	0.1300	DBW2	0.0166	3	0.71	0.8774	0.134(1)	0.782
9.0	0.1224	DBW2	0.0200	2	0.79	0.8858	0.137(3)	1.412
9.0	0.1296	DBW2	0.0200	2	0.78	0.8865	0.115(1)	0.753
9.5	0.1228	DBW2	0.0200	2	0.82	0.8934	0.109(2)	1.576
10.0	0.1234	DBW2	0.0200	2	0.83	0.9000	0.099(2)	1.502
10.5	0.1236	DBW2	0.0200	2	0.79	0.9056	0.093(1)	1.567
11.0	0.1239	DBW2	0.0200	2	0.81	0.9110	0.086(1)	1.473

$O(V^2)$ alternatives [28]. Furthermore, the method we have described is general and can be straightforwardly applied to any fermion action with reuniterization, including overlap fermions with a fat link kernel [29–32], or other types of fatlink actions [33] that may involve alternative smearing techniques [34]. Additionally, any of the variants of HMC can be also be used, in particular polynomial HMC [35] or rational HMC [36] which allow for the simulation of odd numbers of sea quark flavors.

ACKNOWLEDGMENTS

The authors would like to thank Herbert Neuberger for valuable correspondence. This work was supported by the Australian Research Council (ARC). Computations were performed on facilities provided by the Australian Partnership for Advanced Computing (APAC), the South Australian Partnership for Advanced Computing (SAPAC), and the National Computing Facility for Lattice Gauge Theory (NCFLGT).

[1] H. Neuberger, Phys. Lett. B **417**, 141 (1998).
[2] J.M. Zanotti, D.B. Leinweber, A.G. Williams, and J.B. Zhang, Nucl. Phys. B (Proc. Suppl.) **129**, 287 (2004).
[3] R.D. Young, D.B. Leinweber, and A.W. Thomas, Nucl. Phys. B (Proc. Suppl.) **129**, 290 (2004).
[4] CSSM Lattice, J.M. Zanotti *et al.*, Phys. Rev. D **65**, 074507 (2002).
[5] D.B. Leinweber *et al.*, nucl-th/0211014.
[6] W. Kamleh, D. B. Leinweber, A. G. Williams, and J. M. Zanotti (in preparation).
[7] W. Kamleh, D.B. Leinweber, and A.G. Williams, Nucl. Phys. B (Proc. Suppl.) **129**, 826 (2004).
[8] W. Kamleh, D.B. Leinweber, and A.G. Williams, Nucl. Phys. B (Proc. Suppl.) **128**, 96 (2004).
[9] C. Morningstar and M. Peardon, Phys. Rev. D **69**, 054501 (2004).
[10] S. Duane, A.D. Kennedy, B.J. Pendleton, and D. Roweth, Phys. Lett. B **195**, 216 (1987).

[11] M. Campostrini and P. Rossi, Nucl. Phys. **B329**, 753 (1990).
[12] J.C. Sexton and D.H. Weingarten, Nucl. Phys. **B380**, 665 (1992).
[13] T. Takaishi and P. de Forcrand, Int. J. Mod. Phys. C **13**, 343 (2002).
[14] M. Falcioni, M.L. Paciello, G. Parisi, and B. Taglienti, Nucl. Phys. **B251**, 624 (1985).
[15] APE, M. Albanese *et al.*, Phys. Lett. B **192**, 163 (1987).
[16] F.D.R. Bonnet, D.B. Leinweber, A.G. Williams, and J.M. Zanotti, Phys. Rev. D **65**, 114510 (2002).
[17] Y. Liang, K.F. Liu, B.A. Li, S.J. Dong, and K. Ishikawa, Phys. Lett. B **307**, 375 (1993).
[18] H. Neuberger, Phys. Rev. Lett. **81**, 4060 (1998).
[19] T.-W. Chiu, T.-H. Hsieh, C.-H. Huang, and T.-R. Huang, Phys. Rev. D **66**, 114502 (2002).
[20] X.-Q. Luo, Comput. Phys. Commun. **94**, 119 (1996).
[21] K. Jansen and C. Liu, Comput. Phys. Commun. **99**, 221 (1997).
[22] M. Luscher and P. Weisz, Commun. Math. Phys. **97**, 59 (1985).
[23] QCD-TARO, P. de Forcrand *et al.*, Nucl. Phys. **B577**, 263 (2000).
[24] A. Borici and R. Rosenfelder, Nucl. Phys. B (Proc. Suppl.) **63**, 925 (1998).
[25] QCD-TARO, P. de Forcrand *et al.*, Nucl. Phys. B (Proc. Suppl.) **53**, 938 (1997).
[26] T. Takaishi, Phys. Rev. D **54**, 1050 (1996).
[27] CP-PACS, A. Ali Khan *et al.*, Phys. Rev. D **65**, 054505 (2002).
[28] A. Alexandru and A. Hasenfratz, Phys. Rev. D **66**, 094502 (2002).
[29] W. Kamleh, D.H. Adams, D.B. Leinweber, and A.G. Williams, Phys. Rev. D **66**, 014501 (2002).
[30] MILC, T. DeGrand, Phys. Rev. D **63**, 034503 (2001).
[31] W. Bietenholz, Nucl. Phys. **B644**, 223 (2002).
[32] T.G. Kovacs, Phys. Rev. D **67**, 094501 (2003).
[33] M. Stephenson, C. DeTar, T. DeGrand, and A. Hasenfratz, Phys. Rev. D **63**, 034501 (2001).
[34] A. Hasenfratz and F. Knechtli, Phys. Rev. D **64**, 034504 (2001).
[35] R. Frezzotti and K. Jansen, Nucl. Phys. **B555**, 395 (1999).
[36] A.D. Kennedy, I. Horvath, and S. Sint, Nucl. Phys. B (Proc. Suppl.) **73**, 834 (1999).