LEAST-SQUARES FINITE ELEMENT METHODS FOR QUANTUM ELECTRODYNAMICS[*]

J. BRANNICK[†], C. KETELSEN[‡], T. MANTEUFFEL[‡], AND S. MCCORMICK[‡]

Abstract. A significant amount of the computational time in large Monte Carlo simulations of lattice field theory is spent inverting the discrete Dirac operator. Unfortunately, traditional covariant finite difference discretizations of the Dirac operator present serious challenges for standard iterative methods. For interesting physical parameters, the discretized operator is large and ill-conditioned and has random coefficients. More recently, adaptive algebraic multigrid (AMG) methods have been shown to be effective preconditioners for Wilson's discretization [J. Brannick et al., *Lect. Notes Comput. Sci. Eng.*, 55 (2006), pp. 499–506], [J. Brannick et al., *Phys. Rev. Lett.*, 100 (2008), pp. 041601–041604] of the Dirac equation. This paper presents an alternate discretization of the two-dimensional Dirac operator of quantum electrodynamics (QED) based on least-squares finite elements. The discretization is systematically developed and, physical properties of the resulting matrix system are discussed. Finally, numerical experiments are presented that demonstrate the effectiveness of adaptive smoothed aggregation (αSA) multigrid as a preconditioner for the discrete field equations.

Key words. quantum chromodynamics, lattice, finite element, multigrid, smoothed aggregation

AMS subject classifications. 81V05, 65N30, 65N55

DOI. 10.1137/080729633

1. Introduction. Quantum chromodynamics (QCD) is the leading theory in the standard model of particle physics of the strong interactions between color charged particles (quarks) and the particles that bind them (gluons). Analogous to the way that electrically charged particles exchange photons to create an electromagnetic field, quarks exchange gluons to form a very strong color force field. Contrary to the electromagnetic force, the strong force binding quarks does not get weaker as the particles get farther apart. As such, at long distances (low energies), quarks have not been observed independently in experiment and, due to their strong coupling, perturbative techniques, which have been so successful in describing weak interactions in quantum electrodynamics (QED), diverge for the low-energy regime of QCD. Instead, hybrid Monte Carlo (HMC) simulations are employed in an attempt to numerically predict physical observables in accelerator experiments [5].

A main computational bottleneck in an HMC simulation is computation of the so-called *fermion propagator*, another name for the inverse of the discrete Dirac operator. This process accounts for a large amount of the overall simulation time. For realistic physical parameter values, the Dirac operator has random coefficients and is extremely ill-conditioned. The two main parameters of interest are the temperature (β) of the background gauge field and the fermion mass (m). For small temperature values ($\beta < 5$), the entries in the Dirac matrix become extremely disordered.

[†]Department of Mathematics, Pennsylvania State University, University Park, PA 16802 (brannick@psu.edu).

[‡]Department of Applied Mathematics, University of Colorado at Boulder, Boulder, CO 80309-0526 (ketelsen@colorado.edu, tmanteuf@colorado.edu, stevem@colorado.edu).

Moreover, as the fermion mass approaches its true physical value, performance of the community standard Krylov solvers degrades; a phenomenon known as *critical slowing down.* As a result, the development of sophisticated preconditioners for computing propagators has been a priority in the physics community for some time. Recently, multilevel preconditioners like algebraic multigrid have proved to be especially effective at speeding up simulation time [1], [2]. While these works have focused mainly on the task of developing better iterative methods for traditional discretizations of the continuous Dirac operator, it is also important to investigate alternate discretizations as a way to decrease the computational cost of simulation. In [9], a nonlocal approximation to the continuum normal equations is formulated using traditional finite difference techniques. In [6], the continuum equations are expanded in an infinite set of Bloch wavefunctions, and an approximation is obtained by restriction to the lowest mode wavefunctions, which are very similar to the finite element functions employed in this paper. This paper presents an alternate discretization of the Dirac operator based on least-squares finite elements. The discretization is systematically developed, and physical properties of the resulting matrix system are discussed.

In the remainder of this section, we introduce the general continuum Dirac equation and its specific extensions to both QED and the full QCD theory. We then formally describe the simplified two-dimensional (2D) Schwinger model of QED, which is the focus of the rest of this paper. In section 2, we discuss the challenges presented by the discrete Dirac equation, including traditional finite difference discretizations of the field equations and their undesirable properties. The least-squares discretization of the Dirac equations is developed, and several important properties of the resulting system are discussed, including gauge covariance of the propagator, chiral symmetry, and the problem of species doubling. In section 3, we describe the use of an adaptive algebraic multilevel method as a preconditioner for the solution process. Finally, in section 4, we make some concluding remarks.

1.1. The continuum Dirac operator. The Dirac equation is the relativistic analogue of the Schrödinger equation. Depending on the specific gauge theory, the operator can take on several forms, the most general of which is given by

$$(1.1) \qquad \mathcal{D}\psi = \sum_{\mu=1}^{d} \gamma_\mu \left(\partial_\mu - i\mathcal{A}_\mu \right) \psi + m\psi.$$

Here, d is the problem dimension, γ_μ is a matrix coefficient, ∂_μ is the usual partial derivative in the x_μ direction, m is the particle mass, and $\mathcal{A}_\mu(x)$ is the gauge field representing the force carriers. Operator $\mathcal{D}$ acts on $\psi : \mathbb{R}^d \mapsto \mathbb{C}^{n_s} \otimes \mathbb{C}^{n_c}$, a tensor field (multicomponent wavefunction) describing the particle, where n_s and n_c are the number of spin and color components, respectively. These symbols take on different values and dimensions depending on the gauge theory. In full QCD, $d = 4$ (one time and three spatial dimensions), γ_μ are the 4×4 anticommuting complex Dirac matrices, and $\mathcal{A}_\mu(x) \in \mathrm{su}(3)$ is the set of 3×3, traceless, Hermitian matrices that describe the gluon fields. The unknown ψ is a 12-component wavefunction describing a single fermion, with each component corresponding to a quark state of a specific color (red, green, or blue), handedness (right or left), and energy (positive or negative). Here, *handedness*, or *helicity*, is a characterization of a particle's angular momentum relative to its direction of motion [7]. Suppose that s represents, for instance, the state of a

quark being red, right-handed, and having positive energy. Then,

$$\int_V |\psi_s|^2 dV$$

is the probability that the particle is red, right-handed, has positive energy, and can be found in the space-time region V [7].

The Dirac equation is not restricted to the behavior of quarks. In general, it can describe the behavior of any fermion, including electrons. Because of the considerable complexity of the full physical model, when developing algorithms for QCD, it is common practice to consider the simplified 2D Schwinger model of QED [1], which models the interaction between electrons and photons. In this case, only two spatial directions are considered: the particle wavefunction ψ has only two components (right- and left-handed), and the photon field $\mathcal{A}_\mu(x)$ is a real-valued scalar. Although it is a substantial simplification, the discrete Dirac operator associated with the Schwinger model presents many of the same numerical difficulties found in the full physical model.

1.2. Model problem. Let the domain be $\mathcal{R} = [0,1] \times [0,1]$, and let $\mathcal{V}_\mathbb{c} \subset H^1(\mathcal{R})$ be the space of complex-valued, periodic functions on $\mathcal{R}$. We introduce the shorthand notation $\nabla_\mu = (\partial_\mu - i\mathcal{A}_\mu)$ for the μth covariant derivative. The continuum Dirac equation for the 2D Schwinger model with periodic boundary conditions is given by

$$(1.2) \qquad \mathcal{D}(\mathcal{A})\psi = [\gamma_1 \nabla_x + \gamma_2 \nabla_y + mI]\psi = f \qquad \text{in } \mathcal{R},$$
$$\psi(0,y) = \psi(1,y) \quad \forall y \in (0,1),$$
$$\psi(x,0) = \psi(x,1) \quad \forall x \in (0,1),$$

where $\mathcal{A}(x,y) = [\mathcal{A}_1(x,y), \mathcal{A}_2(x,y)]^t$ is the periodic real-valued gauge field and $\psi(x,y) = [\psi_R(x,y), \psi_L(x,y)]^t \in \mathcal{V}_\mathbb{c}^2$ is the fermion field with ψ_R and ψ_L representing the right- and left-handed particles, respectively. In 2D, the γ-matrices correspond to the Pauli spin matrices of quantum mechanics. They are

$$(1.3) \qquad \gamma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \gamma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}.$$

Note that (1.2) appears in matrix notation as

$$(1.4) \qquad \begin{bmatrix} mI & \nabla_x - i\nabla_y \\ \nabla_x + i\nabla_y & mI \end{bmatrix} \begin{bmatrix} \psi_R \\ \psi_L \end{bmatrix} = \begin{bmatrix} f_R \\ f_L \end{bmatrix}.$$

A word on notation. In this paper, we use three different types of objects: continuum functions, finite element functions, and discrete vectors. Continuum functions are represented by scripted and Greek symbols, as in $\mathcal{A}$, f, and ψ. Finite element functions are represented by the similar symbols, but with a superscript h, as in A^h, f^h, and ψ^h. Finally, discrete vectors appear with an underbar, as in $\underline{A}$, $\underline{f}$, and $\underline{\psi}$. Operators in the continuum are denoted by scripted symbols, as in $\mathcal{D}$, while discrete operators are represented by bold symbols, as in $\mathbb{D}$. In any case, the nature of the operator should always be clear from context.

2. The discrete Dirac operator. One computationally intensive part of a QCD simulation is the repeated solution of linear systems of the form

$$\mathbb{D}(\underline{A})\underline{\psi} = \underline{f},$$

where $\mathbb{D}$ is a matrix version of the Dirac operator. The dependence on the discrete gauge field $\underline{A}$ is emphasized here by the notation $\mathbb{D}(\underline{A})$. We omit showing this dependence below when it is clear. Solution of systems of this type are needed both for computing observables and for generating gauge fields with the correct probabilistic characteristics [2]. In these processes, $\mathbb{D}$ must be inverted numerous times with many different right-hand sides and gauge configurations. Because the background fields must be varied, the entries in the matrix themselves change throughout a simulation.

In the discrete setting, $\mathcal{R}$ is replaced by an $n \times n$ periodic lattice. Let $\mathcal{N}_{\mathbb{C}}$ be the space of discrete complex-valued vectors, with values associated with the sites (nodes) on the lattice. Then, the continuum wavefunction ψ becomes $\underline{\psi} = [\underline{\psi}_R, \underline{\psi}_L]^t \in \mathcal{N}_{\mathbb{C}}^2$, which specifies complex values of both the right- and left-handed components of the fermion field at each lattice site. Similarly, the source term f becomes $\underline{f} = [\underline{f}_R, \underline{f}_L]^t \in \mathcal{N}_{\mathbb{C}}^2$. Let $\mathcal{E}$ be the space of discrete real-valued vectors, with values associated with the lattice links. The continuum gauge field $\mathcal{A}$ becomes $\underline{A} = [\underline{A}_1, \underline{A}_2]^t \in \mathcal{E}$, where $\underline{A}_1$ specifies the values of the gauge field on the horizontal lattice links and $\underline{A}_2$ specifies the values of the gauge field on the vertical lattice links.

Traditional discretization methods for the Dirac operator are based on covariant finite differences (CoFD) [12]. Formulations of this type are problematic from a computational perspective because they frequently introduce numerical instabilities into the solution process, which are sometimes remedied by adding artificial stabilization terms. Furthermore, the resulting discrete operator is not usually Hermitian and positive definite. It is standard practice to solve the discrete form of the normal equations

$$\mathbb{D}^* \mathbb{D} \underline{\psi} = \mathbb{D}^* \underline{f}, \tag{2.1}$$

rather than treating the original system directly. This decreases the efficiency of the simulation, since $\mathbb{D}^* \mathbb{D}$ has a larger stencil than $\mathbb{D}$ and a larger condition number. The proposed discretization, based on least-squares finite elements, requires the solution of linear systems that are Hermitian positive definite (HPD) but have smaller stencils than CoFD produces.

2.1. The least-squares discretization. We begin by formulating the solution to (1.2) in terms of a minimization principle:

$$\psi = \arg \min_{\varphi \in \mathcal{V}_{\mathbb{C}}^2} \|\mathcal{D}\varphi - f\|_0^2, \tag{2.2}$$

where $\mathcal{V}_{\mathbb{C}}$ is the space of continuous, periodic, complex-valued H^1 functions defined previously. Equation (2.2) is equivalent to the weak form

$$\text{find } \psi \in \mathcal{V}_{\mathbb{C}}^2 \text{ such that (s.t.) } \langle \mathcal{D}\psi, \mathcal{D}v \rangle = \langle f, \mathcal{D}v \rangle \quad \forall v \in \mathcal{V}_{\mathbb{C}}^2, \tag{2.3}$$

where $\langle \cdot, \cdot \rangle$ is the usual L^2 inner product. If ψ is sufficiently smooth, (2.3) is formally equivalent to the weak form

$$\text{find } \psi \in \mathcal{V}_{\mathbb{C}}^2 \text{ s.t. } \langle \mathcal{D}^* \mathcal{D}\psi, v \rangle = \langle \mathcal{D}^* f, v \rangle \quad \forall v \in \mathcal{V}_{\mathbb{C}}^2.$$

Thus, we can think of the least-squares formulation of the problem as being loosely equivalent to solving the continuum normal equations $\mathcal{D}^* \mathcal{D}\psi = \mathcal{D}^* f$ by the Galerkin method. Looking at the *formal normal* operator $\mathcal{D}^* \mathcal{D}$ can often give insight into the potential success of the least-squares formulation:

$$\mathcal{D}^* \mathcal{D} = \begin{bmatrix} mI & -\nabla_x + i\nabla_y \\ -\nabla_x - i\nabla_y & mI \end{bmatrix} \begin{bmatrix} mI & \nabla_x - i\nabla_y \\ \nabla_x + i\nabla_y & mI \end{bmatrix}$$
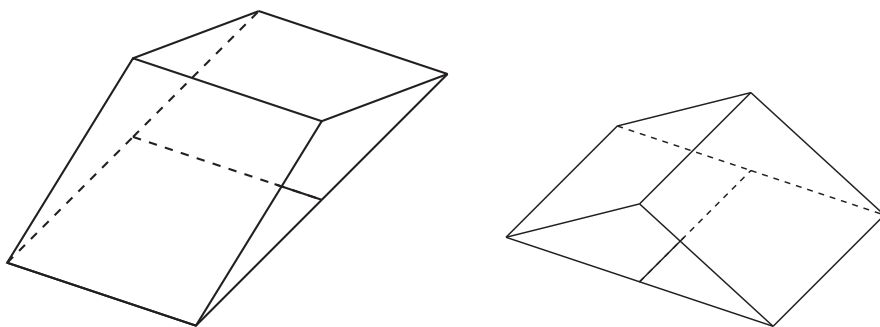
FIG. 2.1. *Nédélec elements associated with a horizontal lattice link (left) and a vertical lattice link (right).*

$$= \left[\begin{array}{cc} m^2 I - \nabla_x^2 - \nabla_y^2 - i\left[\nabla_x, \nabla_y\right] & 0 \\ 0 & m^2 I - \nabla_x^2 - \nabla_y^2 - i\left[\nabla_y, \nabla_x\right] \end{array} \right].$$

In the Schwinger case, the formal normal has uncoupled Laplacian-like operators on the main diagonal. The term $\nabla_x^2 + \nabla_y^2$ is known as the *gauge Laplacian*. Though these are not simple constant coefficient operators (because they include the random background fields), their HPD scalar character should lend themselves to a more efficient treatment by multigrid methods.

The least-squares solution is obtained by restricting the minimization problem in (2.2) and, thus, the weak form in (2.3), to a finite-dimensional space $\mathcal{V}_{\mathbb{C}}^h \subset \mathcal{V}_{\mathbb{C}}$. That is, our solution must satisfy the weak form

$$(2.4) \qquad \text{find } \psi^h \in \left(\mathcal{V}_{\mathbb{C}}^h\right)^2 \text{ s.t. } \left\langle \mathcal{D}\psi^h, \mathcal{D}v^h \right\rangle = \left\langle f^h, \mathcal{D}v^h \right\rangle \quad \forall v^h \in \left(\mathcal{V}_{\mathbb{C}}^h\right)^2.$$

In an analogy to the nodal setting, each elementary square on the lattice, or *plaquette*, is represented by a quadrilateral finite element. We equate any $\underline{f} \in \mathcal{N}_{\mathbb{C}}^2$ with the bilinear function $f^h \in \left(\mathcal{V}_{\mathbb{C}}^h\right)^2$, where $\mathcal{V}_{\mathbb{C}}^h = \text{span}\{\phi_j\}_{j=1}^{n^2}$ is taken to be the space of periodic bilinear finite element functions over the complex numbers. Here, ϕ_j is the usual bilinear nodal basis function associated with lattice site x_j. Similarly, we equate any $\underline{A} \in \mathcal{E}$ with $A^h \in \mathcal{W}^h$, where $\mathcal{W}^h$ is the Nédélec space over the real numbers. In this context, the x-component of the gauge field A_1^h is represented by a linear combination of edge functions associated with the horizontal lattice links. The corresponding basis functions are constant along the link and have support only in the elements above and below. They take on the constant value 1.0 on the link and are linear in y, decaying to 0 at the opposite horizontal links in their shared elements (see Figure 2.1). The basis for the y-component A_2^h is similar, but oriented on the vertical links (see Figure 2.1) [8].

The canonical maps between members of the discrete spaces $\mathcal{N}_{\mathbb{C}}$ and $\mathcal{E}$ and the finite element spaces $\mathcal{V}_{\mathbb{C}}^h$ and $\mathcal{W}^h$ are straightforward. To see this, let

$$\underline{f} = \left[\begin{array}{c} f_1 \\ \vdots \\ f_j \\ \vdots \\ f_{n^2} \end{array} \right] \quad \text{and} \quad f^h = \sum_{j=1}^{n^2} b_j \phi_j.$$

Note that f_j is the value of the discrete field at the jth lattice site, and the finite element field f^h takes on the value b_j at the jth lattice site. For the two field descriptions to be consistent, we must have $f_j = b_j$, $j = 1, \ldots, n^2$. Thus, the canonical mapping between $\mathcal{N}_\mathbb{C}$ and $\mathcal{V}_\mathbb{C}^h$ is simply the bijective identity map between the entries of the nodal vector and the coefficients of the finite element function. A similar analysis shows that the same relationship holds between the gauge field edge values of $\underline{A} \in \mathcal{E}$ and the coefficients of the Nédélec representation of the gauge field $A^h \in \mathcal{W}^h$.

We wish to use the least-squares formalism described above to approximate the solution of (2.1). This process should accept source data $\underline{f}$ defined on the nodes and gauge field data $\underline{A}$ prescribed on the lattice links, and return the discrete wavefunction $\underline{\psi}$, defined at the nodes. We do this by mapping $\underline{f}$ and $\underline{A}$ into their respective finite element spaces, solving the weak formulation (2.4), and mapping the resulting finite element solution back to $\mathcal{N}_\mathbb{C}^2$. This process is summarized in Algorithm 1.

ALGORITHM 1 (least-squares Dirac solve).
- Input: *Gauge field $\underline{A}$, source term $\underline{f}$.*
- Output: *Wavefunction $\underline{\psi}$.*
1. *Map $\underline{A} \mapsto A^h \in \mathcal{W}^h$.*
2. *Map $\underline{f} \mapsto f^h \in \left(\mathcal{V}_\mathbb{C}^h \right)^2$.*
3. *Find $\psi^h \in \left(\mathcal{V}_\mathbb{C}^h \right)^2$ s.t. $\left\langle \mathcal{D}\psi^h, \mathcal{D}v^h \right\rangle = \left\langle f^h, \mathcal{D}v^h \right\rangle \quad \forall v^h \in \left(\mathcal{V}_\mathbb{C}^h \right)^2$, where $\mathcal{A} = A^h$.*
4. *Map $\psi^h \mapsto \underline{\psi} \in \mathcal{N}_\mathbb{C}^2$.*

It is not immediately obvious how to best implement the weak form (2.4), which appears in step 3 of Algorithm 1. Using the nodal basis for $\mathcal{V}_\mathbb{C}^h$, we can establish the following matrix equation for this process:

$$\mathbb{L}u = \mathbb{G}b,$$

where the entries in vectors u and b are the coefficients in the expansions of ψ^h and f^h, respectively, and the elements of the matrices are given by

$$[\mathbb{L}]_{j,k} = \left\langle \mathcal{D}\phi_k, \mathcal{D}\phi_j \right\rangle,$$
$$[\mathbb{G}]_{j,k} = \left\langle \phi_k, \mathcal{D}\phi_j \right\rangle.$$

Then, step 3 in Algorithm 1 can be replaced by computing

$$u = \mathbb{L}^{-1}\mathbb{G}b$$

and setting

$$\psi^h = \sum_{j=1}^{n^2} u_j \phi_j.$$

Recalling the relationship between the entries of $\underline{\psi}$ and $\underline{f}$ and the coefficients in the expansion of ψ^h and f^h, we see that steps 2–4 in Algorithm 1 can be replaced by

(2.5) $$\underline{\psi} = \mathbb{L}^{-1}\mathbb{G}\underline{f}.$$

It is easy to see that, for $m > 0$, both $\mathbb{L}$ and $\mathbb{G}$ are nonsingular. For $\mathbb{L}$, note that by construction, $\mathbb{L}$ is Hermitian positive semidefinite, and if it were singular, then the

original Dirac operator would be singular on some element of $\left(\mathcal{V}_{\mathbb{C}}^h\right)^2$. Note also that $\mathbb{L}$ is block diagonal. Specifically, $\mathbb{L}$ and $\mathbb{G}$ have the form

$$(2.6) \quad \mathbb{L} := \begin{bmatrix} \mathbb{L}_{xx} + \mathbb{L}_{yy} + i(\mathbb{L}_{xy} - \mathbb{L}_{yx}) + m^2\mathbb{M} & 0 \\ 0 & \mathbb{L}_{xx} + \mathbb{L}_{yy} - i(\mathbb{L}_{xy} - \mathbb{L}_{yx}) + m^2\mathbb{M} \end{bmatrix},$$

$$(2.7) \quad \mathbb{G} := \begin{bmatrix} m\mathbb{M} & \mathbb{B}_x - i\mathbb{B}_y \\ \mathbb{B}_x + i\mathbb{B}_y & m\mathbb{M} \end{bmatrix},$$

where

$$\begin{aligned}
(\mathbb{L}_{xx})_{j,k} &= \langle \nabla_x \phi_k, \nabla_x \phi_j \rangle & (\mathbb{M})_{j,k} &= \langle \phi_k, \phi_j \rangle \\
(\mathbb{L}_{yy})_{j,k} &= \langle \nabla_y \phi_k, \nabla_y \phi_j \rangle & (\mathbb{B}_x)_{j,k} &= \langle \phi_k, \nabla_x \phi_j \rangle \\
(\mathbb{L}_{xy})_{j,k} &= \langle \nabla_x \phi_k, \nabla_y \phi_j \rangle & (\mathbb{B}_y)_{j,k} &= \langle \phi_k, \nabla_y \phi_j \rangle \\
(\mathbb{L}_{yx})_{j,k} &= \langle \nabla_y \phi_k, \nabla_x \phi_j \rangle. &
\end{aligned}$$

$\mathbb{G}$ is a skew-Hermitian matrix shifted by mI. Thus, all eigenvalues of $\mathbb{G}$ are of the form $m + is$ for some $s \in \mathbb{R}$.

2.2. Gauge covariance of the fermion propagator. A desirable property of any QED (or QCD) theory is that the fermion propagator must transform covariantly under local gauge transformations. These local transformations can be thought of as redefining the coordinate system of the background gauge field at different points in space. In full QCD, for instance, applying a gauge transformation to wavefunction ψ at position x changes the color reference frame at that particular point. A trivial example would be if the roles of blue and red particles where switched at one or several points in the domain.

Suppose we have a fermion field ψ, defined in a color reference frame $\mathcal{C}$. Now suppose we are given a gauge transformation $\Omega(x) \in \mathrm{SU}(3)$, the set of 3×3, unitary matrices, with determinant one. Suppose the field is transformed into a new reference frame $\widetilde{\mathcal{C}}$, according to $\psi \mapsto \Omega(x)\psi$. Propagator $\mathcal{D}^{-1}$ transforms covariantly if, given $\Omega(x)$, it is possible to specify a modified propagator $\widetilde{\mathcal{D}}^{-1}$ such that applying $\widetilde{\mathcal{D}}^{-1}$ to a field in $\widetilde{\mathcal{C}}$ is equivalent to applying the original propagator to the field in $\mathcal{C}$ and then transforming the result to $\widetilde{\mathcal{C}}$. In other words, given $\Omega(x)$, we must be able to specify $\widetilde{\mathcal{D}}^{-1}$ such that

$$\widetilde{D}^{-1}\Omega(x)\psi = \Omega(x)D^{-1}\psi.$$

It should not be surprising that the correct transformation of $\mathcal{D}^{-1}$ requires modifying the background gauge fields that the Dirac operator is built upon. It is helpful to look at an example of this concept in the 2D Schwinger model of QED, where the gauge transformation comes from U(1), that is, $\Omega(x, y)$ is a complex scalar with unit magnitude.

Example 2.1. Consider the continuum 2D Schwinger model. From (1.1), the Dirac operator is

$$\mathcal{D} = [\gamma_1 \nabla_x + \gamma_2 \nabla_y + mI] = \begin{bmatrix} mI & \nabla_x - i\nabla_y \\ \nabla_x + i\nabla_y & mI \end{bmatrix},$$

where ∇_x and ∇_y are the covariant derivative in the x and y directions, respectively. Let $\Omega(x) = e^{i\theta(x,y)}$ be a transformation from the gauge group U(1). Here, θ is a real-valued, periodic, continuous function in H^1. We denote the space of such functions by

$\mathcal{V}_\mathbb{R} \subset \mathcal{V}_\mathbb{C}$. We want to show that, given transformation Ω, we can modify the covariant derivative operators ∇_x, ∇_y so that the propagator $\mathcal{D}^{-1}$ transforms covariantly. To see this, set

$$e^{i\theta} \left[\gamma_1 \nabla_x + \gamma_2 \nabla_y + mI \right]^{-1} \psi = \zeta,$$

implying

$$
\begin{aligned}
\psi &= \left[\gamma_1 \nabla_x + \gamma_2 \nabla_y + mI \right] e^{-i\theta} \zeta \\
&= \gamma_1 \nabla_x \left(e^{-i\theta} \zeta \right) + \gamma_2 \nabla_y \left(e^{-i\theta} \zeta \right) + m e^{-i\theta} \zeta \\
&= \gamma_1 \left(\partial_x - i\mathcal{A}_1 \right) \left(e^{-i\theta} \zeta \right) + \gamma_2 \left(\partial_y - i\mathcal{A}_2 \right) \left(e^{-i\theta} \zeta \right) + m e^{-i\theta} \zeta \\
&= e^{-i\theta} \left[\gamma_1 \left(\partial_x - i\{\mathcal{A}_1 + \theta_x\} \right) + \gamma_2 \left(\partial_y - i\{\mathcal{A}_2 + \theta_y\} \right) + mI \right] \zeta,
\end{aligned}
$$

where $\theta_x = \partial_x \theta$ and $\theta_y = \partial_y \theta$. Thus,

$$\left[\gamma_1 \widetilde{\nabla}_x + \gamma_2 \widetilde{\nabla}_y + mI \right]^{-1} e^{i\theta} \psi = \zeta,$$

implying

$$\left[\gamma_1 \widetilde{\nabla}_x + \gamma_2 \widetilde{\nabla}_y + mI \right]^{-1} e^{i\theta} \psi = e^{i\theta} \left[\gamma_1 \nabla_x + \gamma_2 \nabla_y + mI \right]^{-1} \psi.$$

This shows that if fermion field ψ is transformed according to $\psi \mapsto e^{i\theta(x,y)}\psi$, a necessary and sufficient condition for obtaining covariance is that the gauge field transforms according to $\mathcal{A} \mapsto \mathcal{A} + \nabla\theta$.

A simple consequence of these facts is the following. Suppose we are given continuum data $\mathcal{A}$ and f. Then we define the related gauge field and rotated source term $\widetilde{\mathcal{A}} = \mathcal{A} + \nabla\theta$ and $\widetilde{f} = e^{i\theta} f$. It is easy to check, using the principle of gauge covariance, that if ψ is the solution to the continuum Dirac equation with data $\mathcal{A}$ and f, then the solution with the modified data should be $\widetilde{\psi} = e^{i\theta}\psi$. We use this fact as a basis for a test of the gauge covariance of our discrete algorithm.

Example 2.2. Consider the continuum Dirac equation with gauge field $\mathcal{A}$, which we write as

$$(2.8) \qquad\qquad \mathcal{D}(\mathcal{A})\psi = f.$$

Consider a Helmholtz decomposition of the gauge field $\mathcal{A}$:

$$\mathcal{A} = \mathcal{A}_0 + \nabla\omega,$$

where $\mathcal{A}_0$ is divergence free and $\omega \in \mathcal{V}_\mathbb{R}$. Then (2.8) becomes

$$(2.9) \qquad\qquad \mathcal{D}(\mathcal{A}_0 + \nabla\omega)\psi = f,$$

to which the solution is

$$(2.10) \qquad\qquad \psi = [\mathcal{D}(\mathcal{A}_0 + \nabla\omega)]^{-1} f.$$

Rewriting the source function as $f = e^{i\omega} g$ for some $g \in \mathcal{V}_\mathbb{C}$, then (2.10) becomes

$$(2.11) \qquad\qquad \psi = [\mathcal{D}(\mathcal{A}_0 + \nabla\omega)]^{-1} e^{i\omega} g.$$

But, from gauge covariance of the propagator, we know that

$$\psi = e^{i\omega} \left[\mathcal{D} \left(\mathcal{A}_0 \right) \right]^{-1} g,$$

implying

$$\psi = e^{i\omega} \left[\mathcal{D} \left(\mathcal{A}_0 \right) \right]^{-1} e^{-i\omega} f.$$

Now, suppose that we wish to solve the same problem but with rotated data. In this case, the Dirac equation becomes

$$\mathcal{D}(\widetilde{\mathcal{A}})\widetilde{\psi} = \widetilde{f}.$$

The Helmholtz decomposition of $\widetilde{\mathcal{A}}$ is

$$\widetilde{\mathcal{A}} = \mathcal{A}_0 + \nabla \left(\omega + \theta \right),$$

and the Dirac equation becomes

$$\mathcal{D} \left(\mathcal{A}_0 + \nabla \left(\omega + \theta \right) \right) \widetilde{\psi} = \widetilde{f}.$$

Writing the source term as $\widetilde{f} = e^{i(\omega+\theta)}\widetilde{g}$, the solution becomes

$$\widetilde{\psi} = \left[\mathcal{D} \left(\mathcal{A}_0 + \nabla \left(\omega + \theta \right) \right) \right]^{-1} e^{i(\omega+\theta)}\widetilde{g}.$$

Again, by gauge covariance, the solution becomes

$$\widetilde{\psi} = e^{i(\omega+\theta)} \left[\mathcal{D} \left(\mathcal{A}_0 \right) \right]^{-1} \widetilde{g},$$

implying

$$\begin{aligned}\widetilde{\psi} &= e^{i(\omega+\theta)} \left[\mathcal{D} \left(\mathcal{A}_0 \right) \right]^{-1} e^{-i(\omega+\theta)} \widetilde{f} \\ &= e^{i\theta}\{e^{i\omega} \left[\mathcal{D} \left(\mathcal{A}_0 \right) \right]^{-1} e^{-i\omega} f\}.\end{aligned}$$

Thus, $\widetilde{\psi} = e^{i\theta}\psi$, as desired.

The key to retaining this property in the discrete setting is that the quark propagator, computed in both cases, is constructed with the same gauge field $\mathcal{A}_0$ and the same source term $e^{-i\omega} f$. As such, we must be able to efficiently compute a discrete Helmholtz decomposition of the gauge field A^h. Fortunately, the choice to represent the gauge field by Nédélec elements makes this fairly easy. Given $A^h \in \mathcal{W}^h$, there exists a unique $q^h \in \mathcal{V}^h_{\mathbb{R}}$ such that

$$A^h = A_0^h + \nabla q^h,$$

where $q^h \in \mathcal{V}^h_{\mathbb{R}}$ is a bilinear function and $A_0^h \in \mathcal{W}^h$ is characterized by the property that

(2.12) $$\left\langle A_0^h, \nabla v^h \right\rangle = 0 \ \ \forall v^h \in \mathcal{V}^h_{\mathbb{R}}.$$

A vector in $\mathcal{W}^h$ that satisfies (2.12) is known as a weak curl [8]. The decomposition can be computed by solving the least-squares problem

$$q^h = \arg \min_{v^h \in \mathcal{V}^h_{\mathbb{R}}} \| A^h - \nabla v^h \|_0^2,$$

which is equivalent to the weak form

$$\left\langle \nabla q^h, \nabla v^h \right\rangle = \left\langle A^h, \nabla v^h \right\rangle \quad \forall v^h \in \mathcal{V}_{\mathbb{R}}^h.$$

This weak formulation yields a linear system that is equivalent to that involved in the solution of Poisson's equation with periodic boundary conditions using a Galerkin finite element method. It is easily solved by standard geometric multigrid methods.

We are now ready to construct a new discrete algorithm that is gauge covariant. First, given $\underline{q} \in \mathcal{N}_{\mathbb{R}} \subset \mathcal{N}_{\mathbb{C}}$ defined on the lattice sites, let $\underline{\Omega}_q$ be the $n^2 \times n^2$ complex-valued matrix with e^{iq_j} in the jth diagonal position and 0 elsewhere. Notice that $\underline{\Omega}_{\underline{q}}^*$ is also diagonal with e^{-iq_j} in the jth diagonal position. Both $\underline{\Omega}_q$ and $\underline{\Omega}_{\underline{q}}^*$ are unitary matrices.

ALGORITHM 2 (gauge covariant least-squares Dirac solve).
- Input: *Gauge field $\underline{A}$, source term $\underline{f}$.*
- Output: *Wavefunction $\underline{\psi}$.*
1. *Map $\underline{A} \mapsto A^h \in \mathcal{W}^h$.*
2. *Compute $A^h = A_0^h + \nabla q^h$.*
3. *Map $q^h \to \underline{q}$.*
4. *Set $\underline{g}_R = \underline{\Omega}_{\underline{q}}^{\overline{*}} \underline{f}_R$ and $\underline{g}_L = \underline{\Omega}_q^* \underline{f}_L$.*
5. *Build $\mathbb{L}$, $\mathbb{G}$ based on $\mathcal{A} = A_0^h$.*
6. *Solve $\mathbb{L}\underline{\phi} = \mathbb{G}\underline{g}$ for $\underline{\phi}$.*
7. *Set $\underline{\psi}_R = \underline{\Omega}_q \underline{\phi}_R$ and $\underline{\psi}_L = \underline{\Omega}_q \underline{\phi}_L$.*

Note also that steps 5–6 can then be replaced by the familiar matrix operation

$$\underline{\phi} = \mathbb{L}^{-1}\mathbb{G}\underline{g}, \tag{2.13}$$

where matrices $\mathbb{L}$ and $\mathbb{G}$ were constructed using the grad-free gauge field A_0^h.

Through our development of the gauge covariant algorithm, we have shown a natural relationship between two gauge fields that differ only by a gradient. That is, they represent the same physical system, but in a different color reference frame. Thus, any gauge field is physically equivalent to an infinite number of other fields. Rather than consider a specific gauge field, we instead consider the equivalence class that it belongs to. Formally, given $\underline{q} \in \mathcal{N}_{\mathbb{R}}$, define $[\Delta_x \underline{q}, \Delta_y \underline{q}]^T \in \mathcal{E}$ such that

$$\left[\Delta_x \underline{q}\right]_{(k+1/2,l)} = \frac{q_{(k+1,l)} - q_{(k,l)}}{h},$$

where $q_{(k,l)}$ is the value of $\underline{q}$ associated with the kth lattice site in the x-direction and the lth lattice site in the y-direction. Subscript $(k+1/2, l)$ indicates that the value is associated with the lattice link between the lattice sites (k, l) and $(k+1, l)$. Similarly, define

$$\left[\Delta_y \underline{q}\right]_{(k,l+1/2)} = \frac{q_{(k,l+1)} - q_{(k,l)}}{h}.$$

DEFINITION 2.3. *We say that pairs $(\underline{\psi}, \underline{A})$ and $(\widetilde{\underline{\psi}}, \widetilde{\underline{A}})$ are in the same equivalence class if there exist $\underline{q} \in \mathcal{N}_{\mathbb{R}}$ such that*

$$\widetilde{\underline{\psi}} = \left[\widetilde{\underline{\psi}}_R, \ \widetilde{\underline{\psi}}_L\right]^t = \left[\underline{\Omega}_q \underline{\psi}_R, \ \underline{\Omega}_q \underline{\psi}_L\right]^t$$

and

$$\widetilde{\underline{A}} = \left[\widetilde{\underline{A}}_1, \ \widetilde{\underline{A}}_2\right]^t = \left[\underline{A}_1 + \Delta_x \underline{q}, \ \underline{A}_2 + \Delta_y \underline{q}\right]^t.$$

THEOREM 2.4. *Suppose $(\underline{f}, \underline{A})$ and $(\widetilde{\underline{f}}, \widetilde{\underline{A}})$ are in the same equivalence class. Then Algorithm 2 yields $\underline{\psi}$ and $\widetilde{\underline{\psi}}$ such that*

$$\underline{\psi} = \left[\underline{\psi}_R, \underline{\psi}_L\right]^t = \left[\underline{\Omega}_q \, \widetilde{\underline{\psi}}_R, \, \underline{\Omega}_q \, \widetilde{\underline{\psi}}_L\right]^t.$$

Proof. The proof follows directly from Definition 2.3 and the development of Algorithm 2. □

2.3. Chiral symmetry. In the broadest sense, chiral symmetry is a global symmetry property that, in the massless case, independent transformations of the right- and left-handed fields do not change the physics of the model. This property is manifested mathematically by the property that, when $m = 0$, the inner product $< \psi, \gamma_1 \mathcal{D}\psi >$ remains invariant under transformations of the form $\hat{\psi} = \Omega\psi$, where γ_1 is as defined in (1.3) and

$$(2.14) \qquad \Omega = \begin{bmatrix} e^{i\theta_R} & 0 \\ 0 & e^{i\theta_L} \end{bmatrix}$$

for $\theta_R, \theta_L \in \mathbb{R}$. Thus, we require that, in the massless case,

$$(2.15) \qquad < \hat{\psi}, \gamma_1 \mathcal{D}\hat{\psi} > = < \psi, \gamma_1 \mathcal{D}\psi > .$$

It is important to note the differences between the requirements of chiral symmetry and that of gauge covariance. First, chiral symmetry is a *global* symmetry, which is why θ_R and θ_L in Ω do not have spatial dependence. All right- and left-handed fields are rotated by the same transformation at each point. Second, we are not permitted to alter $\mathcal{D}$ to make (2.15) hold.

A little algebra shows that a sufficient condition for (2.15) is that the term $\gamma_1 \mathcal{D}\hat{\psi}$ transforms as $\Omega\gamma_1 \mathcal{D}\psi$. This, in turn, is equivalent to the statement that if $m = 0$, $\mathcal{D}\psi = f$, and $\hat{\psi} = \Omega\psi$, then $\mathcal{D}\hat{\psi} = \hat{f}$, where $\hat{f} = \gamma_1 \Omega\gamma_1 f$. To see this, let

$$\gamma_1 \mathcal{D}\hat{\psi} = \Omega\gamma_1 \mathcal{D}\psi,$$

implying that

$$\gamma_1 \hat{f} = \Omega\gamma_1 f,$$

and the result immediately follows. We make the following definition of chiral symmetry.

DEFINITION 2.5. *$\mathcal{D}$ preserves chiral symmetry if, for any θ_R, $\theta_L \in \mathbb{R}$ used to define Ω in (2.14) and any $\psi, f \in \mathcal{V}_{\mathbb{C}}^2$ such that*

$$\mathcal{D}\psi = f$$

for $m = 0$, then

$$\hat{\psi} = \Omega\psi, \qquad \hat{f} = \gamma_1 \Omega\gamma_1 f,$$

satisfy

$$\mathcal{D}\hat{\psi} = \hat{f}.$$

It is clear from (1.4) that this holds for the continuum Dirac operator. In the discrete case, operators Ω and γ_1 become matrices, which we denote by

$$\underline{\Omega} = \left[\begin{array}{cc} e^{i\theta_R} I & 0 \\ 0 & e^{i\theta_L} I \end{array} \right],$$

$$\underline{\Gamma}_1 = \left[\begin{array}{cc} 0 & I \\ I & 0 \end{array} \right],$$

where I is the identity operator on $\mathcal{N}_c$. We state Chiral symmetry of the least-squares operator in the following lemma.

LEMMA 2.6 (chiral symmetry for the discrete least-squares operator). *Given any* θ_R, $\theta_L \in \mathbb{R}$, *and any* $\underline{\psi}$, $\underline{f} \in \mathcal{N}_c^2$ *such that*

$$\mathbb{L}\underline{\psi} = \mathbb{G}\underline{f}$$

for $m = 0$, *then*

$$\underline{\hat{\psi}} = \underline{\Omega}\underline{\psi}, \qquad \underline{\hat{f}} = \underline{\Gamma}_1 \underline{\Omega}\underline{\Gamma}_1 \underline{f},$$

satisfy

$$\mathbb{L}\underline{\hat{\psi}} = \mathbb{G}\underline{\hat{f}}.$$

Proof. Recalling (2.6) and (2.7), it is easy to see that $\mathbb{L}$ and $\mathbb{G}$ are of the form

(2.16) $$\mathbb{L} = \left[\begin{array}{cc} \mathbb{L}_{11} & 0 \\ 0 & \mathbb{L}_{22} \end{array} \right],$$

(2.17) $$\mathbb{G} = \left[\begin{array}{cc} m\mathbb{M} & \mathbb{G}_{12} \\ -\mathbb{G}_{12}^* & m\mathbb{M} \end{array} \right].$$

From (2.16), we see that, for any m,

$$\mathbb{L}\underline{\Omega} = \underline{\Omega}\mathbb{L}.$$

From (2.17), we also see that, for $m = 0$, we have

$$\mathbb{G}\underline{\Gamma}_1 \underline{\Omega}\underline{\Gamma}_1 = \underline{\Omega}\mathbb{G}.$$

Thus,

$$\mathbb{L}\underline{\hat{\psi}} = \mathbb{L}\underline{\Omega}\underline{\psi} = \underline{\Omega}\mathbb{L}\underline{\psi},$$
$$\mathbb{G}\underline{\hat{f}} = \mathbb{G}\underline{\Gamma}_1 \underline{\Omega}\underline{\Gamma}_1 \underline{f} = \underline{\Omega}\mathbb{G}\underline{f},$$

which yields the result. □

2.4. Species doubling. A concern in the numerical analysis of the field equations for QCD is the problem of *species doubling*. We illustrate this phenomenon by turning to the one-dimensional (1D) Schwinger model. In CoFD formulations, the discrete Dirac operator is given by

$$\mathbb{D} = \gamma_1 \otimes \nabla_x^h + mI,$$

where I is the $2n^2 \times 2n^2$ identity matrix. The so-called *naive* discretization corresponds to approximating covariant derivative ∇_x using central differences. In the absence of a gauge field, this becomes

$$\nabla_x^h = \frac{\psi(x+h) - \psi(x-h)}{2h}.$$

We write the discrete Dirac operator $\mathbb{D}_N$ as

$$\mathbb{D}_N = \begin{bmatrix} mI & \frac{1}{h}\mathbb{N} \\ \frac{1}{h}\mathbb{N} & mI \end{bmatrix},$$

where $\mathbb{N}$ is the periodic Toeplitz matrix with stencil $[-1/2 \ \ 0 \ \ 1/2]$. Assume that the lattice has $n \times n$ cells, meaning that it has $(n+1) \times (n+1)$ periodic lattice sites and that n is even. The eigenvalues of $\mathbb{N}$ are given by

$$(2.18) \qquad \nu_k = i \sin\left(\frac{2\pi k}{n}\right)$$

for $k = -(n/2 - 1), \ldots, n/2$. Note that ν_k and ν_{-k}, for $k = 1, \ldots, n/2$, are complex conjugate pairs. From the form of $\mathbb{D}_N$, we see that the eigenvalues of the discrete propagator $\mathbb{D}_N^{-1}$ are given by

$$\kappa_k = \frac{h}{mh \pm i \sin(2\pi k/n)},$$

with corresponding eigenvectors

$$(2.19) \quad \underline{v}_k = \begin{cases} [1, 1, \ldots, 1, 1]^t & k = 0, \\ [\ldots, \cos(2\pi k\ell/n) \pm \sin(2\pi k\ell/n), \ldots]^t & k = \pm 1, \ldots, n/2 - 1, \\ [1, -1, \ldots, 1, -1]^t & k = n/2, \end{cases}$$

where $\ell = 1, \ldots, n$. Notice the symmetry of κ_k. For every low-frequency eigenvector, there is a corresponding high-frequency eigenvector that shares the same eigenvalue. The physics community is especially concerned with the correspondence between the eigenvalues of the $k = 0$ and $k = n/2$ modes. In the naive discretization, the eigenvalues of the propagator $\mathbb{D}_N^{-1}$ associated with these modes both approach ∞ as $m \to 0$. Loosely speaking, this represents two particles of different momenta with the same energy, which is impossible. Hence, this phenomenon is referred to as *species doubling*.

In the applied mathematics community, doubling is known as a red/black instability. There are a number of successful approaches for handling the issue [11]. However, the issue is not only removal of the spurious high-frequency components in the discrete solution, but overall accuracy of the discretization process. The addition of the complex gauge field further complicates the situation. The traditional remedy in the physics community is to add an artificial stabilization term to $\mathbb{D}_N$, which we demonstrate below. Later in this section, we demonstrate that the least-squares approach also eliminates this difficulty.

The addition of the artificial stabilization term to $\mathbb{D}_N$ is the basis for the Dirac–Wilson operator, which is given in the 1D gauge-free case by

$$(2.20) \qquad \mathbb{D}_W = \mathbb{D}_N - I \otimes \frac{h}{2}\Delta^h,$$

where Δ^h is the usual 1D Laplacian operator and I is the 2×2 identity matrix. In block form, we have

$$\mathbb{D}_W = \left[\begin{array}{cc} \dfrac{1}{2}\mathbb{H} + mI & \dfrac{1}{h}\mathbb{N} \\[2ex] \dfrac{1}{h}\mathbb{N} & \dfrac{1}{2}\mathbb{H} + mI \end{array} \right],$$

where $\mathbb{N}$ is as before and $\mathbb{H}$ is the periodic Toeplitz matrix constructed via the 3-point, periodic Laplacian stencil $\frac{1}{h}[-1 \ \ 2 \ \ -1]$. The eigenvalues of $\mathbb{H}$ are

$$(2.21) \qquad\qquad \alpha_k = \frac{2}{h}\left[1 - \cos\left(\frac{2\pi k}{n}\right)\right].$$

Consequently, the eigenvalues of the propagator $\mathbb{D}_W^{-1}$ are given by

$$\begin{aligned} \lambda_k &= \left[m + \frac{1}{2}\alpha_k \pm \frac{1}{h}\nu_k\right]^{-1} \\ &= \left[m + \frac{1}{h}\left\{1 - \cos\left(\frac{2\pi k}{n}\right) \pm i\,\sin\left(\frac{2\pi k}{n}\right)\right\}\right]^{-1}, \end{aligned}$$

with the corresponding eigenvectors again given by (2.19). Note that, in this formulation, the eigenvalue corresponding to the lowest frequency mode λ_0 still goes to ∞ as $m \to 0$, but the eigenvalue corresponding to the highest frequency $\lambda_{n/2}$ now approaches h. Thus, the Dirac–Wilson operator does not suffer from species doubling. This comes at a high price, however. To avoid doubling, a nonphysical term was added to the operator. Furthermore, the additional term appears on the main diagonal of $\mathbb{D}_W$, which breaks chiral symmetry.

Species doubling does not occur with the least-squares discretization. To see this, consider the general form of (2.5) in one dimension. The effective least-squares discrete operator is given by

$$\mathbb{D}_{LS} = \mathbb{G}^{-1}\mathbb{L}.$$

In the 1D, gauge-free case, we have

$$\mathbb{L} = \left[\begin{array}{cc} \mathbb{H} + m^2\mathbb{M} & 0 \\ 0 & \mathbb{H} + m^2\mathbb{M} \end{array} \right],$$

$$\mathbb{G} = \left[\begin{array}{cc} m\mathbb{M} & \mathbb{N} \\ \mathbb{N} & m\mathbb{M} \end{array} \right],$$

where $\mathbb{N}$ and $\mathbb{H}$ are as defined above and $\mathbb{M}$ is the periodic Toeplitz matrix with stencil $h\,[1/6 \ \ 2/3 \ \ 1/6]$. The eigenvalues of $\mathbb{M}$ are given by

$$(2.22) \qquad\qquad \mu_k = \frac{h}{3}\left[2 + \cos\left(\frac{2\pi k}{n}\right)\right].$$

Using Fourier analysis as above, we see that the eigenvalues of the least-squares propagator $\mathbb{D}_{LS}^{-1}$ are given by

$$(2.23) \qquad\qquad \tau_k = \frac{m\mu_k \pm \nu_k}{m^2\mu_k + \alpha_k}.$$

Substituting (2.18), (2.21), and (2.22) for μ_k, ν_k, and α_k into (2.23) and simplifying, we have

$$\tau_k = \frac{m\,h^2\,[2 + \cos\,(2\pi k/n)] \pm 3i\,h\,\sin\,(2\pi k/n)}{m^2 h^2\,[2 + \cos\,(2\pi k/n)] + 6\,[1 - \cos\,(2\pi k/n)]}.$$

Again, we are concerned with the lowest frequency mode τ_0 and the highest frequency mode $\tau_{n/2}$. Setting $k = 0$ and taking the limit as $m \to 0$, we see that $\tau_0 \to \infty$, as expected. Then, setting $k = n/2$ and taking the limit to the massless case, we see that $\tau_{n/2}$ approaches 0, not ∞, as in the naive case. Thus, the least-squares formulation for the 1D Dirac operator does not suffer from species doubling. The generalization of this analysis to the 2D case is straightforward.

3. Numerical experiments. In this section, we explore the use of a multilevel iterative method for solving the matrix system (2.13) appearing in step 6 of Algorithm 2. To avoid working in complex arithmetic, we solve the equivalent real formulation of (2.13):

$$(3.1) \qquad \left[\begin{array}{cc} \mathbb{X} & -\mathbb{Y} \\ \mathbb{Y} & \mathbb{X} \end{array}\right] \left[\begin{array}{c} \underline{x} \\ \underline{y} \end{array}\right] = \left[\begin{array}{c} \underline{a} \\ \underline{b} \end{array}\right],$$

where $\mathbb{X}, \mathbb{Y}$ are real-valued matrices satisfying $\mathbb{L} = \mathbb{X} + i\mathbb{Y}$, $\underline{\phi} = \underline{x} + i\underline{y}$, and $\mathbb{G}\underline{g} = \underline{a} + i\underline{b}$. Note that $\mathbb{Y}$ is skew-Hermitian so that (3.1) is a symmetric real system. Moreover, since the complex matrix is HPD, the real system is SPD.

Finally, we compare the computational cost of applying a multilevel iterative method to both (2.13) and the 2D analogue of the Dirac–Wilson matrix described in (2.20), which, in 2D, becomes

$$(3.2) \qquad \mathbb{D}_W = \sum_{\mu=1}^{2} \gamma_\mu \otimes \nabla_\mu^h - I \otimes \frac{h}{2}\Delta^h,$$

where ∇_μ^h and Δ^h are the CoFD representations of the μth covariant derivative and the gauge Laplacian, respectively, and I is the 2×2 identity matrix. For a more indepth description of $\mathbb{D}_W$, see, for example, [1], [2], or [12]. A difficulty in working with $\mathbb{D}_W$ directly is that it is non-Hermitian. To apply standard multilevel techniques to the inversion of $\mathbb{D}_W$, we appeal to the discrete normal equations, with $\mathbb{D}_W^* \mathbb{D}_W$ as the coefficient matrix. Furthermore, $\mathbb{D}_W$ has complex entries, so we formulate the normal equations in an equivalent real way:

$$(3.3) \qquad \left[\begin{array}{cc} \mathbb{U} & -\mathbb{V} \\ \mathbb{V} & \mathbb{U} \end{array}\right] \left[\begin{array}{c} \underline{u} \\ \underline{v} \end{array}\right] = \left[\begin{array}{c} \underline{c} \\ \underline{d} \end{array}\right],$$

where $\mathbb{U}, \mathbb{V}$ are real-valued matrices satisfying $\mathbb{D}_W{}^*\mathbb{D}_W = \mathbb{U} + i\mathbb{V}$, $\underline{\phi} = \underline{u} + i\underline{v}$, and $\mathbb{D}_W{}^*\underline{g} = \underline{c} + i\underline{d}$.

3.1. Smoothed aggregation multigrid. In this section, we compare the performance of adaptive smoothed aggregation (αSA) applied to (3.1) and (3.3). Detailed results for system (3.3) appear in [1].

Multigrid methods rely on two complementary processes to reduce the error in each successive iterate. Relaxation is a *local* process that reduces a large portion of the error in a relatively inexpensive way. Error that relaxation fails to adequately reduce is called *algebraically smooth*. Coarse-grid correction is a *global* process that is
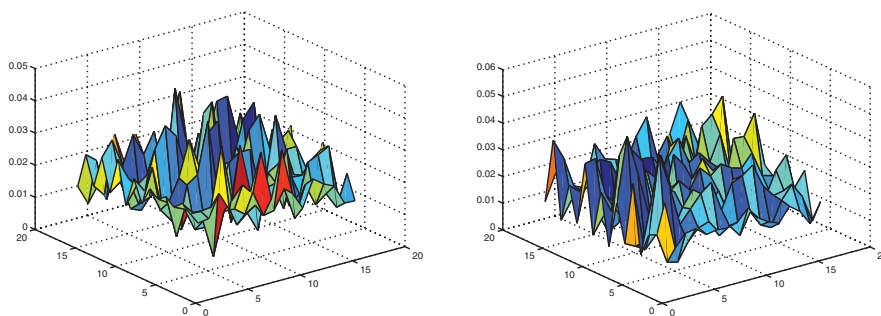
FIG. 3.1. *Real and complex components of algebraically smooth error of* $\mathbb{L}$ *for* $m = 0.1$, $\beta = 2$, *and* $N = 16$. *Error was computed using* 100 *iterations of Gauss-Seidel on* $\mathbb{L}\underline{\phi} = 0$ *with a random initial guess.*

designed to complement relaxation by reducing the algebraically smooth error. These two processes work in tandem, with relaxation performed on the fine grid until only algebraically smooth error remains, allowing coarse-grid correction to be effective. The coarse-grid approximation to the error is then taken back up to the fine grid through an interpolation process and used to correct the approximate there. The success of the coarse-grid correction process depends on how accurately algebraically smooth error modes can be represented on the coarse-grid.

For many problems in the physical sciences, the algebraically smooth error modes are geometrically smooth as well. Standard geometric multigrid methods are usually very effective at solving these problems. Unfortunately, due to the random nature of the background gauge fields in field theory, the algebraically smooth error modes are not at all geometrically smooth. In Figure 3.1, we see that both the real and imaginary components of the algebraically smooth error are highly oscillatory. These plots were obtained by applying 100 iterations of Gauss-Seidel on the problem $\mathbb{L}\underline{\phi} = 0$ with a random initial guess and rescaling the result. This process exposes the algebraically smooth error associated with $\mathbb{L}$.

Smoothed aggregation (SA) multigrid (see [10]) is a multilevel solver that is based on algebraic smoothness as an abstraction of the property of geometric smoothness used in conventional multilevel algorithms. Given prototype representations of algebraically smooth error, SA automatically builds intergrid transfer operators that attempt to represent all algebraically smooth error modes on coarser grids, regardless of their geometric smoothness. Unfortunately, this requires a priori knowledge of the prototype modes. Randomness of the background fields in field theory applications causes the nature of the algebraically smooth error to vary widely between different gauge configurations and, in any case, little is known about their local character. We turn instead to αSA multigrid (see [3]), which uses a setup procedure to first expose these problematic error components and then builds a multigrid process to effectively reduce them.

3.2. Results. Table 3.1 reports average convergence factors of a conjugate gradient (CG) iteration preconditioned by αSA applied to the homogeneous version of (3.1) for various values of the particle mass m and gauge field temperature β. In each case, averages were taken over 20 distinct gauge fields. The αSA preconditioner was based on V(2,2)-cycles with 3 grid levels and an algebraic aggregation process. The relaxation scheme was nodal Gauss-Seidel, meaning that the lattice sites were swept

TABLE 3.1

Average convergence factors for αSA-preconditioned CG and diagonally preconditioned CG applied to (2.13) on 64×64 (top) and 128×128 (bottom) lattices with varying choices of mass parameter m and temperature β.

β/m	.01	.1	.3
2	.16 / .98	.18 / .94	.17 / .94
3	.17 / .98	.17 / .95	.16 / .94
5	.15 / .97	.15 / .94	.18 / .93

β/m	.01	.1	.3
2	.16 / .98	.15 / .96	.15 / .94
3	.16 / .99	.16 / .95	.17 / .95
5	.17 / .96	.15 / .94	.17 / .92

TABLE 3.2

Average convergence factors for αSA-preconditioned CG applied to the least-squares formulation (left) and the normal equations of the Dirac–Wilson operator (right) on a 64×64 lattice with varying choices of mass parameter m and temperature β.

β/m	.01	.1	.3
2	.16 / .33	.15 / .31	.15 / .31
3	.16 / .42	.16 / .40	.17 / .31
5	.17 / .31	.15 / .29	.17 / .28

over in a lexicographic fashion, and all unknowns on a lattice site were updated simultaneously. Finally, 8 prototype error components were found during the adaptive setup process and used to define the intergrid transfer operators in the V-cycle. A single V(2,2)-cycle was used as the preconditioner in the CG solve. For comparison, convergence factors for diagonally preconditioned CG are also provided in Table 3.1. Notice that, as mass parameter m is decreased, the performance of the αSA solver remains static. This is an important result because it means that the problem of critical slowing down has been eliminated.

Table 3.2 compares the performance of αSA-preconditioned CG applied to the equivalent real formulation of the least-squares problem, given in (3.1), and the discrete normal equations of the Dirac–Wilson operator given in (3.3). Numerical results presented for (3.3) were taken from [1]. In both cases, 8 prototype error components were found during the adaptive setup process and used to define the intergrid transfer operators in the V-cycle.

Note that the average convergence factors are significantly better for the least-squares formulation. Data illustrate that αSA-preconditioned CG achieves slightly more accuracy per multigrid cycle for the proposed formulation than for the normal equations of CoFDs.

We must also recognize that the computational cost is significantly less for the least-squares problem because it avoids the added complexity of the discrete normal equations. Our least-squares approach does form normal equations, but more effectively on the continuum Dirac operator only, without the additional stabilization term. Discretizing the continuum normal equations in this way results in a stencil that has only the nearest-neighbor connections typical of second-order operators, in contrast to the wider and more complicated stencils for the normal equations of the Dirac–Wilson matrix. As a result, the least-squares matrix is more compact and has about 53% as many nonzeros as the normal equations of the Dirac–Wilson matrix.

Also, the *operator complexity* σ is significantly better for least-squares than it is for CoFDs. σ is defined to be the ratio of the total number of degrees of freedom on all grids in the multigrid hierarchy to the number of degrees of freedom on the finest grid. This number indicates how much work has to be done on the coarse grids compared to that on the fine grid. For the lattice sizes that were tested in these experiments, operator complexity stayed below 1.5, while the operator complexities applied to (3.3) were bounded by 3.0 [1].

The improved operator complexity of the proposed formulation is due to the reduced sparsity of the fine-grid operator. In the least-squares formulation, the fine-grid operator involves only nearest-neighbor coupling of the unknowns. Applying the SA methodology to a system of this form yields a coarse-grid operator that retains the sparsity structure of the original system, resulting in a small operator complexity.

In contrast, the fine-grid operator associated with (3.3) involves second nearest-neighbor coupling. As a result, the coarsening between the fine grid and the first coarse grid fails to significantly reduce the number of nonzeros in the coarse-grid operator. Aggressive coarsening on the remaining coarse levels does keep rapid stencil growth from occuring, but results in an overall larger operator complexity than the least-squares formulation [1].

A useful way to compare the performance of two methods, which differ both in convergence rate and in computational complexity, is to look at a measure of the number of work units necessary to gain one unit of accuracy in the approximate solution. This measure takes into consideration convergence factors, operator complexity, and sparsity of the original system. We define one work unit to be the cost of performing one relaxation sweep on the finest grid of the least-squares matrix $\mathbb{L}$. The normal equations of the Dirac–Wilson operator $\mathbb{D}_W^* \mathbb{D}_W$ has approximately 88% more nonzeros than that of $\mathbb{L}$. Thus, one relaxation sweep on $\mathbb{D}_W^* \mathbb{D}_W$ costs approximately 1.88 work units. To quantify this performance factor, we introduce a variant on a formula developed in [4]. Define η to be the work units necessary to improve the current iterate by one digit of accuracy. Then,

$$(3.4) \qquad \eta = \delta \, \sigma \, (\nu_1 + \nu_2 + 1) \, \frac{\log .1}{\log \rho},$$

where σ is the operator complexity, ν_1 and ν_2 are the number of prerelaxation and postrelaxation steps performed in the V-cycle, ρ is the usual convergence factor reported in Table 3.2, and δ is a scale factor that allows us to quantify the cost of a work unit relative to the specific discretization. For the least-squares discretization, we set $\delta = 1$ and, for the Dirac–Wilson system, $\delta = 1.88$. The values ν_1 and ν_2 are both 2 for all of the experiments, indicating that a V(2,2)-cycle was used. As stated above, we have $\sigma = 1.5$ for the least-squares formulation and $\sigma = 3.0$ for the Dirac–Wilson formulation. Table 3.3 gives the η-values for the experiments described previously in Table 3.2. Taking the ratio of η-values of the Dirac–Wilson operator to those of the least-squares formulation, we get an estimate of the speedup obtained from one discretization over the other. These ratios are given in Table 3.4.

Note that αSA-preconditioned CG applied to the least-squares formulation obtains between 5 and 7 times the accuracy per computational cost than that of the same iterative method applied to the normal equations of the Dirac–Wilson matrix.

The decision to build the intergrid transfer operators using eight prototype error components was convenient because it allowed direct comparison to the results obtained for the Dirac–Wilson operator. It does, however, beg the question whether it is more computationally efficient to use fewer error components. Using fewer com-

TABLE 3.3

Average η-value for αSA-preconditioned CG applied to the least-squares formulation (left) and the normal equations of the Dirac–Wilson operator (right) on a 64×64 lattice with varying choices of mass parameter m and temperature β.

β/m	.01	.1	.3
2	9.4 / 58.6	10.1 / 55.4	9.7 / 55.4
3	9.7 / 72.8	9.7 / 70.9	9.4 / 55.4
5	9.1 / 55.4	9.1 / 52.5	9.7 / 51.0

TABLE 3.4

Average speedup factors for αSA-preconditioned CG applied to the least-squares formulation over the normal equations of the Dirac–Wilson operator on a 64×64 lattice with varying choices of mass parameter m and temperature β.

β/m	.01	.1	.3
2	6.2	5.5	5.7
3	7.5	7.3	5.9
5	6.1	5.8	5.2

TABLE 3.5

Average convergence factors and operator complexities for αSA-preconditioned CG applied to the least-squares formulation using varrying number of prototype error components (K) on a 128×128 lattice. Here $\beta = 2$.

K/m	.01	.1	.3	σ
4	.80	.81	.80	1.12
5	.66	.64	.65	1.19
6	.40	.39	.40	1.28
7	.25	.24	.24	1.38
8	.17	.17	.15	1.50

TABLE 3.6

Average η-value for αSA-preconditioned CG applied to the least-squares formulation using varrying number of prototype error components (K) on a 128×128 lattice. Here $\beta = 2$.

K/m	.01	.1	.3
4	57.8	61.2	57.8
5	32.9	30.7	31.8
6	16.1	15.7	16.1
7	11.5	11.1	11.1
8	9.1	9.1	9.7

ponents leads to smaller operator complexities and thus cheaper computation cost on coarser grids. Table 3.5 shows the average convergence factors for a CG iteration preconditioned by an SA V(2,2)-cycle built on a varrying number of prototype error components. Operator complexities for each cycle are also reported.

We can again compute the number of work units required to improve the current iterate by one digit of accuracy. These values are given in Table 3.6. The data indicates that using eight prototype error components yields the most computationally efficient method.

4. Conclusions. We described a discretization of the continuous Dirac equation for the 2D Schwinger model based on least-squares finite elements. The formulation avoids several pitfalls of traditional discretizations based on covariant finite differences by producing a discrete operator that is Hermitian, positive definite, and extremely

sparse. We formulated our solution process in a gauge covariant way and argued that it retains a sense of global chiral symmetry. We showed that our method avoids the need for stabilization terms and that it does not suffer from species doubling. Furthermore, we showed that the resulting discrete system can be handled quite effectively by CG with αSA multigrid as a preconditioner. It is not immediately clear if the least-squares methodology can be extended to the full QCD model. The noncommutative nature of the gauge field makes the QCD problem much more complicated. As a result, further investigation is warranted.

REFERENCES

[1] J. BRANNICK, M. BREZINA, D. KEYES, O. LIVNE, I. LIVSHITS, S. MACLACHLAN, T. MANTEUFFEL, S. MCCORMICK, J. RUGE, AND L. ZIKATANOV, *Adaptive smoothed aggregation in lattice QCD*, Lect. Notes Comput. Sci. Eng., 55 (2006), pp. 499–506.
[2] J. BRANNICK, R.C. BROWER, M.A. CLARK, J.C. OSBORN, AND C. REBBI, *Adaptive multigrid algorithm for lattice QCD*, Phys. Rev. Lett., 100 (2008), pp. 041601–041604.
[3] M. BREZINA, R. FALGOUT, S. MACLACHLAN, T. MANTEUFFEL, S. MCCORMICK, AND J. RUGE, *Adaptive smoothed aggregation (αSA)*, SIAM J. Sci. Comput., 25 (2004), pp. 1896–1920.
[4] M. BREZINA, T. MANTEUFFEL, S. MCCORMICK, J. RUGE, AND G. SANDERS, *Towards adaptive smoothed aggregation (αSA) for nonsymmetric problems*, SIAM J. Sci. Comput., to appear.
[5] T. DEGRAND AND C. DETAR, *Lattice Methods for Quantum Chromodynamics*, World Scientific, River Edge, NJ, 2006.
[6] R. FRIEDBERG AND T.D. LEE, *Noncompact lattice formulation of gauge theories*, Phys. Rev. D (3), 52 (1995), pp. 4053–4081.
[7] D. GRIFFITHS, *Introduction to Elementary Particles*, Wiley-VCH, Weinheim, 2004.
[8] J.C. NÉDÉLEC, *A new family of mixed finite elements in $\mathbb{R}^3$*, Numer. Math., 50 (1986), pp. 57–81.
[9] C. REBBI, *Chiral-invariant regularization of fermions on the lattice*, Phys. Lett. B, 186 (1987), pp. 200–204.
[10] P. VANEK, J. MANDEL, AND M. BREZINA, *Algebraic multigrid by smooth aggregation for second and fourth order elliptic problems*, Computing, 56 (1996), pp. 179–196.
[11] P. WESSELING, *Principles of Computational Fluid Dynamics*, Springer, Heidelberg, 2001.
[12] K. WILSON, *Confinement of quarks*, Phys. Rev. D (3), 10 (1974), pp. 2445–2459.