2D Adjoint QCD

Patrick Oare

1 2D Adjoint QCD

Cite the old QCD_2 papers here:

2d adjoint QCD (QCD_2) is the theory of a Majorana fermion coupled to a $SU(N)$ gauge field in 2 spacetime dimensions. This has the following action in Minkowski space:

$$S = \int d^2x \, \text{Tr}\left[\frac{1}{2g^2} G_{\mu\nu}G^{\mu\nu} + \overline{\psi}(i\gamma^\mu D_\mu - m)\psi\right] \tag{1}$$

where $\psi = \psi^a_\alpha(x)$ is a Majorana adjoint fermion. Note here that the theory is often considered with a **massless** adjoint fermion, $m = 0$, and as will be discussed shortly, the confinement picture differs sharply when the fermion is massive or massless.

The confinement picture for this theory is the following [1]. For $N = 2$, calculations [**?**] have shown the following:

- $m = 0$: Deconfining (fundamental Wilson lines have a perimeter law). This should imply that $\langle P\rangle \neq 0$, which **may be impossible to see** with lattice calculations. This is a strange result that was puzzling for a while, since having adjoint quarks should mean there is nothing to screen the charged objects (fundamental Wilson lines). However, there are modern symmetry arguments that show this is indeed the case and should make sense, which I should read up on.

- $m > 0$: Confining (fundamental Wilson lines have an area law). This should imply that $\langle P\rangle = 0$. This is the sector that we can measure on the lattice; since we'll likely measure $\langle P\rangle = 0$, this means we won't be able to extrapolate results for the Polyakov loop to the massless limit.

For arbitrary N, things are more complicated. The portion of Aleksey's paper [5] when the four-fermion deformations are turned off should give us some knowledge about what to expect, and Ref. [1] also does a detailed analysis of this sector. They find that the results of the $SU(2)$ theory holds for arbitrary N:

- $m = 0$: Deconfining.

- $m > 0$: Confining.

Ross and Igor have studied this theory for a lot of different cases using lightcone quantization [2]. There are a variety of interesting results in this paper that we should think about trying to compute, including:

- The spectrum of the theory, as a function of N and the adjoint mass m_{adj}. There are results here for finite N, and for large N, so it would be interesting to compare.

- The lowest lying fermionic and bosonic masses, as a function of m_{adj} (see Section 4.2 of that paper). In particular, there's a nice plot that we could try to replicate.

John Donahue and Sergei Dubovsky have also studied this theory using string theory at large m_{adj}, in Refs. [3, 4]. Read these papers and investigate this

Quantities to compute:

- Polyakov loop correlator $\langle P \rangle$.
- String tension σ for Wilson loops in a variety of representations.
- Chiral condensate $\langle \overline{\psi}\psi \rangle$.
- (Harder) Fermion and bosonic spectrum ρ_f, ρ_b. If this is too hard, just do low-lying states.

Questions to ask:

- It would be very interesting to simulate the massless theory, but we can't naively do that with a LQCD calculation. I know in the Schwinger model, there are ways to tune the lattice mass so that it better corresponds to a specific value of the continuum mass, and in particular when the continuum mass is zero using this tuning can yield a non-zero lattice mass. Is there any analog of this for 2d adjoint QCD?
- What parameters (adjoint fermion mass m_{adj}, gauge coupling g) would be the most interesting ones to look at?
- Lightcone quantization vs Euclidean lattice: assuming it doesn't matter and they're both just computational ways to do it, but confirm that we can match quantities.

2 2D Adjoint QCD with four fermion deformations

2d adjoint QCD with four-fermion deformations is studied in Alexei's paper [5]. The main point is that this theory is very similar to the four-fermion deformed massless Schwinger model (with even N), in a number of ways:

- Adjoint QCD has a $\mathbb{Z}_2$ chiral symmetry when $m_q = 0$.
- Adjoint QCD admits two four-fermion deformations, just like the massless Schwinger model, which are consistent with chiral symmetry. When these deformations are turned off, the theory is deconfined, but when the theory turns on, the theory confines.

The 2d adjoint QCD theory is described by a single Majorana fermion coupled in the adjoint representation to an $SU(N)$ gauge field in 2d, with action:

$$S = \int d^2x \left\{ \frac{1}{2g^2} \operatorname{Tr} G_{\mu\nu} G^{\mu\nu} + \operatorname{Tr} \psi^T i\gamma^\mu D_\mu \psi + \frac{c_1}{N} \operatorname{Tr} \psi_+\psi_+\psi_-\psi_- + \frac{c_2}{N^2} \operatorname{Tr}[\psi_+\psi_-] \operatorname{Tr}[\psi_+\psi_-] \right\}. \quad (2)$$

This action is in Euclidean space, with Euclidean γ matrices in 2d are given by

$$\gamma^1 = \sigma_1 \qquad \gamma^2 = \sigma_3 \qquad \gamma = i\gamma^1\gamma^2 \quad (3)$$

with γ taking the role of γ_5. For $N > 2$, there are four symmetries of the theory that are unbroken by anomalies:

1. Center symmetry $\mathbb{Z}_N^{[1]}$, also just referred to as $\mathbb{Z}_N$ 1-form symmetry.

2. Charge conjugation $\mathbb{Z}_2^C$, $a_{ij}^\mu \mapsto -a_{ji}^\mu, \psi_{ij} \mapsto \psi_{ji}$, with $i, j = 1, ..., N$ being color indices for the adjoint representation. In $N = 2$, this transformation reduces to global $SU(2)$ symmetry, so in this case this is not an additional symmetry.

3. Fermion parity $\mathbb{Z}_2^F$, $\psi \mapsto -\psi$.

4. Chiral symmetry $\mathbb{Z}_2^\chi$, $\psi \mapsto \gamma\psi$.

The important symmetry to consider is the $\mathbb{Z}_N^{[1]}$ 1-form symmetry, as the spontaneous breaking of this symmetry signals confinement, since the expectation value of the Polyakov loop $\langle P(x) \rangle$ is the order parameter for confinement. **If center symmetry is unbroken, then the expectation value for large Wilson loops obeys an area law, which signals that we are in the confining phase**.

Ref. [6] showed that discrete 1-form symmetries in 2d are always unbroken, which would imply any system which has a 1-form symmetry must be confining, as the symmetry cannot break spontaneously. However, this paper shows that there is an exception to this conclusion when N is even: in this case, a representation with N-ality $N/2$ is **not confined**. This will be interesting to verify numerically.

2.1 't Hooft anomalies

Recall: A 't Hooft anomaly is an *obstruction* to gauging a global symmetry. It is defined as a mixed anomaly between 3 global symmetries. The global symmetry survives in the quantum theory (i.e. its associated conserved current is still conserved), but it cannot be coupled to a dynamical gauge field.

The key idea in this paper is to study the mixed 't Hooft anomalies between the chiral symmetry $\mathbb{Z}_2^\chi$ and the other discrete symmetries. Heuristically, they find that the partition function $\mathcal{Z}$ transforms under $\mathbb{Z}_2^\chi$ as $\mathcal{Z} \mapsto -\mathcal{Z}$, TODO

We will not go into the detail of how the anomalies are computed, but the important point is the 't Hooft anomaly matching conditions. Anomaly matching (Sec. 5) implies that the low-energy theory must have some combination of:

a) Intrinsic topological order.

b) Gapless excitations.

c) Spontaneous symmetry breaking. This would lead to deconfinement and / or screening.

Option (c) is the one that Alexei favored, but it would be good to have explicit verification of this. In particular, intrinsic topological order cannot appear in 2 dimensions [7]. Option (b) is harder to exclude; it can be excluded if the 4-fermion terms are not added to the theory, but if those terms are present in the theory, it can no longer be excluded. It would be good to study this in the computation.

3 Discussion

3.1 What do we want to compute?

- String tension σ as a function of N, for even N. We want to:

- Verify Fig. 2.
- Compute σ_1, $\sigma_{\max}$.
- Numerically verify $\sigma_{N/2} = 0$.
- Verify the string tension degeneracy for even N:

$$\sigma_q = \sigma_{N-q} = \sigma_{q+N/2} = \sigma_{q-N/2}. \tag{4}$$

Computing the string tension can be done by computing the free energy of the system in the confining phase, which is done by examining the Polyakov loop correlator. We have:

$$C_2^q(x) \equiv \langle \mathrm{Tr}[P(x)^q] \mathrm{Tr}[P^\dagger(0)^q] \rangle = e^{-F_q(x)} \tag{5}$$

where $P(x)^q$ is a charge q Polyakov loop, and $F_q(x)$ is the free energy of the system. The free energy scales as the string tension for charge q,

$$F_q(x) \sim x\beta\sigma_q, \tag{6}$$

hence performing an "effective mass" analysis with the correlator $C_2^q(x)$ will yield the string tension σ_q.

- Vary N and **verify the regions we expect to see the theory confine**. The general idea is that the $\mathbb{Z}_N^{[1]}$ center symmetry will spontaneously break to different subgroups, depending on the value of N. Depending on the subgroup, this will tell us which values of test charges will confine, and which ones will not. We should be able to compute the Polyakov loop in these different regimes to see if the theory is confining or not.
  - Even N: $\mathbb{Z}_N^{[1]} \to \mathbb{Z}_{N/2}^{[1]}$, which implies that test charges with N-ality $N/2$ are screened, not confined.
  - Odd N: Center symmetry is unbroken and test charges of all values confine.
  - TODO: Screening vs. confining: what's the difference, and how does the difference make itself manifest?
- Chiral symmetry: **when does chiral symmetry break?** The condensate

$$\langle \mathrm{Tr}[\psi^T i\gamma\psi] \rangle \sim \pm\Lambda \tag{7}$$

is an order parameter for the spontaneous breaking of chiral symmetry (like in 4d QCD). We can compute this and verify when chiral symmetry is broken; we expect it to be spontaneously broken for even N (pg. 23), and for odd N with $N = 4n + 3$ (pg. 26). For odd $N = 4n + 1$, chiral symmetry need not be spontaneously broken.

- Gapless excitations in the theory. Existence of gapless excitations would imply that we don't necessarily have spontaneous center symmetry breaking; although Alexei doesn't believe it's likely, it would be good to verify numerically that there are no gapless excitations in the spectrum (pg. 22).
- Spectrum: The density of states can be computed in the large N limit as

$$\rho(E) \sim m^\alpha e^{\beta_H E} \tag{8}$$

where α is an unknown parameter. It would be interesting to apply our spectral function tools to this, although I don't know if we would be able to really approach the large N limit with lattice calculations. Discuss tomorrow

3.2 What challenges might we face?

- Simulating Majorana fermions. I have to look into this more carefully, but I believe there's some form of a sign problem associated with the computation.
- Simulating massless fermions. Alexei mentions in the paper that one could work with sufficiently light fermions instead and examine the behavior as m gets smaller and smaller.
  - This shouldn't be too much of a problem, we can just do a chiral extrapolation as $am \to 0$. It would be helpful if we know how the string tension and other quantities of interest relate to m.

3.3 Additional questions / notes

- How does the theory change if we couple a Dirac fermion to $SU(N)$ adjoint, rather than a Majorana fermion? This would be a lot easier to implement.
- Do we have predictions for how observables like the string tensions behave when $m \to 0$? In the footnote in Eq. 25, the paper mentions that $\sigma_{N/2} \sim m\Lambda$ goes linearly with m. Can we predict this for the other representations, and if so, how?
- We should be precise when working with the base manifold: there's a section of the paper that deals with $\mathbb{R} \times S^1$, which approaches $\mathbb{R}^2$ in the infinite volume limit, but at finite T there might be finite volume effects to be wary of.
- Refs. to read: Igor's paper on screening vs. confinement [35], papers about the large N spectrum of 2d adjoint QCD [19-22].
- What does David Schaich mean when he says that the sign problem goes away in the continuum limit?

4 The Lorentz group in $d = 2$ spacetime dimensions

What do Majorana spinors look like in 2 dimensions? Here our Majorana fermion ψ^a has 2 chiral components given by ψ_+^a and ψ_-^a. The Euclidean γ matrices[1] are the following:

$$\gamma^1 = \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \gamma^2 = \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \gamma_5 = i\gamma^1\gamma^2 = i\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \sigma_2 \tag{9}$$

We have the usual identities,

$$\{\gamma^\mu, \gamma^\nu\} = 2\delta^{\mu\nu} \qquad \{\gamma^\mu, \gamma_5\} = 0 \qquad \gamma_\mu^\dagger = \gamma_\mu, \gamma_5^\dagger = \gamma_5. \tag{10}$$

In this basis[2], we expand the components of ψ as

$$\psi^a = \begin{pmatrix} \psi_1^a \\ \psi_2^a \end{pmatrix}. \tag{11}$$

[1] The Minkowski space γ matrices can be found in David Tong's lecture notes [8] and are $\gamma^0 = \sigma_1$, $\gamma^1 = i\sigma^2$, and $\gamma_5 = -\gamma^0\gamma^1 = \sigma_3$

[2] Note that in this basis, the matrix γ is not diagonal. This is the $d = 2$ version of the Dirac basis, while the Weyl basis can be determined by diagonalizing γ, likely resulting in $\gamma = \sigma_3$.

These are related to the chiral components $\psi^a_\pm$ by the usual projection formulas,

$$P_\pm = \frac{1}{2}(1 \pm \gamma) \qquad \psi^a_\pm = P_\pm \psi^a. \tag{12}$$

The other important thing to consider is charge conjugation. Charge conjugation is defined to satisfy the relation,

$$C\gamma_\mu^T C^{-1} = -\gamma_\mu. \tag{13}$$

The simplest way to satisfy this is by choosing

$$C = \gamma_5 = \sigma_2. \tag{14}$$

Note that this also satisfies $C\gamma_5 C = \gamma_5$, since $\gamma_5^2 = 1$.

TODO there's a decent chance that $\psi_1 = \psi_+$ and $\psi_2 = \psi_-$, figure this out and a mass term $\mathrm{Tr}\,\psi^T\gamma\psi$ couples ψ_1 to ψ_2 and vice versa.

5 Simulating Majorana Fermions

Here are some notes [9] that explain the basic ideas that we will try to use. We'll work on a 2d Euclidean lattice of volume $L \times T$.

The major thing to worry about is a "sign problem" that can be circumvented with reweighting. The idea is that when you integrate out Majorana fermions, you get a fermion Pfaffian instead of a fermion determinant. This Pfaffian can have either a positive or negative sign: the theory incurs a sign problem when the sign is negative. So, one must keep track of the sign of the Pfaffian and monitor whether or not it will yield a sign problem. Formally, when evaluating the path integral we encounter the Pfaffian after integrating out the fermionic modes:

$$\mathcal{Z} = \int DU \int d\psi e^{-S_g[U] - \frac{1}{2}\psi^T \mathcal{D}\psi} = \int DU\, \mathrm{Pf}[\mathcal{D}] e^{-S_g[U]} \tag{15}$$

where we use the notation $\psi^T\mathcal{D}\psi \equiv \int d^2x\, d^2y\, \psi^T(x)\mathcal{D}(x,y)\psi(y)$. The Pfaffian is harder to treat than a determinant because it is allowed to be negative. We have $(\mathrm{Pf}\,\mathcal{D})^2 = \det \mathcal{D}$, so although $|\mathrm{Pf}\,\mathcal{D}|$ is determined, it can be positive or negative. Since we will use the measure

$$D\mathbb{P} = DU\, \mathrm{Pf}[\mathcal{D}] e^{-S_g[U]} \tag{16}$$

we will have problems if the Pfaffian is not positive. This means we have to monitor the sign of the Pfaffian as we do the calculation, in order to ensure the theory doesn't have a sign problem.

We will monitor the sign problem with a **spectral flow** method.

5.1 Discretizations

There are a bunch of formulations of Majorana fermions that people have tried: Wilson, twisted mass, overlap, and domain-wall. The easiest one to start with is Wilson, in which the 4d case is described in hep-lat/1802.07797. Note that none of these have been done for $d = 2$ spacetime dimensions and $Q = 2$-component Majorana spinors, the closest that has been simulated is $Q = 4$ in 2d (which is the $\mathcal{N} = 1$, $Q = 4$ supersymmetry).

Let's try Wilson fermions to start. We need to add the appropriate term to the action to cancel the doubler modes. The original 2d adjoint QCD action is:

$$S_0 = \int d^2x \left\{ \frac{1}{2g^2} \mathrm{Tr}\, G_{\mu\nu} G^{\mu\nu} + \mathrm{Tr}\, \psi^T i\gamma^\mu D_\mu \psi + m \mathrm{Tr}\, \psi^T \gamma \psi \right\} \tag{17}$$

where ψ is a two-component Majorana spinor in $d = 2$ spacetime dimensions that lives in the adjoint representation of $SU(N)$, i.e. $\psi = \psi^a t^a$, and transforms under gauge transformations $\Omega(x) \in SU(N)$ as

$$\psi(x) \mapsto \Omega(x)\psi(x)\Omega^\dagger(x). \tag{18}$$

We also add a mass term here, because it is likely easier to work with massive fermions and then take a chiral limit (note that $\psi^T\psi$ vanishes because ψ is Grassman-valued).

Let us work with the massive fermion action. We have:

$$S_F = \int d^2x \left\{ \mathrm{Tr}\, \psi^T i\gamma^\mu D_\mu \psi + m \mathrm{Tr}\, \psi^T \gamma \psi \right\} \tag{19}$$

We begin with a naïve discretization of this action. The usual action of D_μ in the adjoint representation is

$$D_\mu \psi = \partial_\mu \psi + i[a_\mu, \psi] \tag{20}$$

where a_μ is the gauge field.

Montvay [9] claims that the Wilson action for a single Majorana fermion (in $d = 4$) can be expanded as:

$$\begin{aligned} S_F^W &= \frac{1}{2}\psi^T \mathcal{D} \psi \\ &= \frac{1}{2}\sum_x \left\{ \overline{\psi}^a(x)\psi^a(x) - K\sum_{\mu=1}^{4} [\overline{\psi}^a(x+\hat{\mu}) V_\mu^{ab}(x)(1+\gamma_\mu)\psi^b(x) + \overline{\psi}^r(x)(V_\mu^{ab})^T(x)(1-\gamma_\mu)\psi^b(x+\hat{\mu}) \right\} \end{aligned} \tag{21}$$

where the V are the matrix elements of the link variables in the adjoint representation,

$$V_\mu^{ab}(x) = 2\mathrm{Tr}[U_\mu^\dagger(x) t^a U_\mu(x) t^b] = V_\mu^{ab}(x)^*. \tag{22}$$

We should verify this is the case in $d = 2$ dimensions by explicitly taking the naïve Dirac operator to momentum space and projecting out the doublers.

TODO: consider doing a domain wall discretization. Also, see if it's easy to port this over to GPT

6 Computing the Pfaffian: Rational Hybrid Monte Carlo (RHMC)

Incorporating the Pfaffian into an effective action will require the use of the identity

$$|\mathrm{Pf}[\mathcal{D}]| = |\det \mathcal{D}|^{1/2} = (\det[\mathcal{D}^\dagger\mathcal{D}])^{1/4} \propto \int D\Phi D\Phi^\dagger \exp\left[-\Phi^\dagger(\mathcal{D}^\dagger\mathcal{D})^{-1/4}\Phi\right] \equiv \int D\Phi\, D\Phi^\dagger e^{-\Phi^\dagger K^{-1/4}\Phi} \tag{23}$$

where here Φ and $\Phi^\dagger$ are bosonic pseudofermion fields with indices $\Phi_\alpha^a(n)$ (as the Dirac operator carries indices $\mathcal{D}_{\alpha\beta}^{ab}$ and $K \equiv \mathcal{D}^\dagger\mathcal{D}$. The assumption here that $|\mathrm{Pf}[\mathcal{D}]| = \mathrm{Pf}[\mathcal{D}]$ is not in general correct, and the Pfaffian can have an arbitrary phase

$$\mathrm{Pf}[\mathcal{D}[U]] = e^{i\alpha[U]}|\mathrm{Pf}[\mathcal{D}[U]]| \tag{24}$$

where $\alpha[U]$ is the phase associated with the configuration U. One can deal with the possibility of non-zero phase by reweighting, which absorbs the phase into the denominator. The idea is to

define the **phase-quenched path integral** as

$$\begin{aligned}\langle\mathcal{O}\rangle_{pq} &\equiv \frac{1}{\mathcal{Z}_{pq}}\int DU\,|\mathrm{Pf}[\mathcal{D}[U]]|e^{-S_g[U]}\mathcal{O} = \frac{1}{\mathcal{Z}_{pq}}\int DU\,D\Phi\,D\Phi^\dagger\,e^{-S_{\text{eff}}[U,\Phi]}\mathcal{O}\\ \mathcal{Z}_{pq} &\equiv \int DU\,|\mathrm{Pf}[\mathcal{D}[U]]|e^{-S_g[U]} = \int DU\,D\Phi\,D\Phi^\dagger\,e^{-S_{\text{eff}}[U,\Phi]},\end{aligned} \tag{25}$$

i.e. we use the norm of the Pfaffian instead of the complex Pfaffian. Here the effective action is the result of using the identity in Eq. (23),

$$S_{\text{eff}}[U,\Phi] = S_g[U] + \underbrace{\Phi^\dagger K[U]^{-1/4}\Phi}_{S_F[U,\Phi]}. \tag{26}$$

With this phase quenching, one can then evaluate a correlator as:

$$\langle\mathcal{O}\rangle = \frac{\langle\mathcal{O}e^{i\alpha}\rangle_{pq}}{\langle e^{i\alpha}\rangle_{pq}} \tag{27}$$

The key here is to measure the full partition function $\mathcal{Z} = \langle e^{i\alpha}\rangle_{pq}$ and make sure it never gets close to zero; if it does, the theory will have a sign problem! There are two ways we should monitor this:

- Monitor $\langle e^{i\alpha}\rangle_{pq}$ on each configuration that we compute. This test should verify that we never sample configurations that would render the correlation functions intractable. However, this misses a degenerate case: if there is a small number of configurations that contributes to the path integral in a dominant fashion. The issue with this is that these configurations will never be sampled, so we will be missing their contributions to the path integral.
- Use a Monte Carlo generator on the space of possible gauge field configurations to monitor $\mathrm{Pf}[\mathcal{D}[U]]$, and verify the phase of the Pfaffian is always close to 1. This will ensure that if there is a region of configuration space that has a low probability to be sampled in the path integral but a high contribution to the integral, then we still have a chance to sample it.

6.0.1 What are the necessary conditions to make this a valid Monte Carlo simulation?

The decomposition into

6.1 The rational approximation

We'll need to use RHMC to approximate the fermion Pfaffian, then compute $\langle e^{i\alpha}\rangle_{pq}$ on each ensemble to make sure our ensemble is OK. Note that for some extra notation, we'll often use Q to denote a Hermitian counterpart to the Dirac operator,

$$Q[U] \equiv \gamma_5\mathcal{D}[U]. \tag{28}$$

In Arthur's notes, he only defines Q for a Wilson-Dirac operator $\mathcal{D}$ because it is γ_5-hermitian. I should make sure that this works for the Dirac operators I am considering, both (a) a Wilson operator and (b) an overlap operator. The idea behind rational HMC lies in approximating $K^{-1/4}$ as a rational function r:

$$r(K) \approx K^{-1/4}. \tag{29}$$

The choice of this function r lies at the heart of RHMC. Rational approximations typically only work well when the eigenvalues λ of $K[U]$ fall within a given window $[\lambda_{\text{low}}, \lambda_{\text{high}}]$, so typically K is

scaled dynamically (note that $K[U]$ depends on the gauge field, so it is often scaled as a function of the configuration U) to make its eigenvalues fall within that range. For a given configuration U, its minimum and maximum eigenvalues of $K[U]$, $\lambda_{\min}(U)$ and $\lambda_{\max}(U)$, must therefore fall within the domain of convergence,

$$\lambda_{\text{low}} < \lambda_{\min}(U) \ll \lambda_{\max}(U) < \lambda_{\text{high}}, \tag{30}$$

and this must be monitored as the simulation progresses to make sure no dynamical rescaling is needed. This is cheaper to monitor than the Pfaffian, and in C can be carried out using the PReconditioned Iterative Multi-Method Eigensolver (PRIMME) library [10].

The specifics of the approximation can be found in the Remez Algorithm [11], and they provide the coefficients for the following expansion of $K^{-1/4}$ in terms of P partial fractions:

$$r(K) \approx \alpha_0 + \sum_{i=1}^{P} \frac{\alpha_i}{K + \beta_i}. \tag{31}$$

Note here that this is an operator-valued equation, so β_i is really $\beta_i \text{id}$, with the same shape as K. One must also use a similar expansion for $K^{-1/8}$ to initialize the pseudofermions Φ. For a given range $[\lambda_{\text{low}}, \lambda_{\text{high}}]$, the Remez algorithm allows us to deterministically compute the (α_i, β_i) parameters, and provides a bound on the error of the rational approximation, provided the eigenvalues of $K[U]$ lie within the spectral bound $[\lambda_{\text{low}}, \lambda_{\text{high}}]$. There are also a number of other ways to express this by changing variables, as done in Lüscher's notes [12], which we cite for completeness:

$$r(K) = a_0 \left(1 + \sum_{k=1}^{P} \frac{r_{2k}}{K + a_{2k}}\right) = a_0 \prod_{k=1}^{P} \frac{K + a_{2k-1}}{K + a_{2k}} \tag{32}$$

Here for the first identification we have:

$$a_0 \equiv \alpha_0 \qquad\qquad a_{2k} \equiv \beta_k \qquad\qquad r_{2k} \equiv \frac{\alpha_k}{\alpha_0} \tag{33}$$

and for the second identification, r_{2k} can be related to the a_k coefficients as

$$r_{2k} = \frac{\prod_{\ell=1}^{P}(-a_{2k} - a_{2\ell-1})}{\prod_{\ell \neq k}^{P}(-a_{2k} - a_{2\ell})} \tag{34}$$

The expansion of Eq. (32) is called the **Zolotarev rational approximation**.

6.2 The explicit action

The most efficient way to deal with the Dirac operator is to exploit its sparse structure. For the 4D Wilson action case, from Montvay's paper, we have:

$$\begin{aligned} S_{\text{Wilson}} &= \frac{1}{2} \sum_{x,y \in \Lambda} \overline{\psi}^a_\alpha(x) (\mathcal{D}_W)^{ab}_{\alpha\beta}(x,y) \psi^b_\beta(y) \\ &= \frac{1}{2} \sum_{x \in \Lambda} \left\{ \overline{\psi}^a(x)\psi^a(x) - K \sum_{\mu=1}^{2} \left[\overline{\psi}^a(x+\hat{\mu}) V^{ab}_\mu(x)(1+\gamma_\mu)\psi^b(x) + \overline{\psi}^a (V^T)^{ab}_\mu(x)(1-\gamma_\mu)\psi^b(x+\hat{\mu}) \right] \right\} \end{aligned} \tag{35}$$

For the 2d case, we can read off the Wilson-Dirac operator $\mathcal{D}_W$ as:

$$(\mathcal{D}_W)^{ab}_{\alpha\beta}(x,y) = \delta^{ab}\delta_{\alpha\beta}\delta_{x,y} - K\sum_{\mu=1}^{2}\left[V^{ab}_{\mu}(y)(1+\gamma_\mu)_{\alpha\beta}\delta_{x,y+\hat{\mu}} + (V^T)^{ab}_{\mu}(x)(1-\gamma_\mu)_{\alpha\beta}\delta_{x,y-\hat{\mu}}\right] \tag{36}$$

Note here that $V^{ab}_{\mu}(x)$ is defined to be the gauge link in the adjoint representation,

$$V^{ab}_{\mu}(x) \equiv 2\mathrm{Tr}[U^{\dagger}_{\mu}(x)t^a U_\mu(x) t^b]. \tag{37}$$

which satisfy the identity

$$V^{ab}_{\mu}(x) = V^{ab}_{\mu}(x)^* = ((V^{-1}_{\mu})^T)^{ab}. \tag{38}$$

The first identity can be shown by taking the conjugate of the equation and using the fact that $\{t^a\}$ are Hermitian,

$$(V^{ab}_{\mu})^* = 2\mathrm{Tr}[(t^b)^\dagger U^\dagger_\mu (t^a)^\dagger U_\mu] = 2\mathrm{Tr}[U^\dagger_\mu t^a U_\mu t^b] = V^{ab}_{\mu} \tag{39}$$

The second identity results from the fact that V is unitary, $V^{-1} = V^\dagger$, along with the first identity, $V = V^* = (V^\dagger)^T = (V^{-1})^T$.

The other identity we wish to work with is the usual one when we move sites, i.e. that $U_{-\mu}(x) = U^\dagger_\mu(x-\hat{\mu})$. We have:

$$\begin{aligned} V^{ab}_{-\mu}(x) &= 2\mathrm{Tr}[U^\dagger_{-\mu}(x)t^a U_{-\mu}(x)t^b] = 2\mathrm{Tr}[U_\mu(x-\hat{\mu})t^a U^\dagger_\mu(x-\hat{\mu})t^b] \\ &= 2\mathrm{Tr}[U^\dagger_\mu(x-\hat{\mu})t^b U_\mu(x-\hat{\mu})t^a] = V^{ba}_{\mu}(x-\hat{\mu}) = (V^T_\mu)^{ab}(x). \end{aligned} \tag{40}$$

Combining this with the fact that the adjoint links are real, we see that

$$V_{-\mu}(x) = V^T_\mu(x-\hat{\mu}) = V^\dagger_\mu(x-\hat{\mu}) \tag{41}$$

which is the same identity that is satisfied by the fundamental links $U_\mu(x)$.

TODO Determine correct form of the operator. The above identities need to be resolved so that we can know what to do with $V^\dagger$. I think we should use this following Wilson-Dirac operator, which is likely equal to the one in the Montvay paper if we assume that V are real or something. Let's work with the usual form for the Wilson-Dirac operator from Gattringer and Lang [13], which is expressed as

$$(\mathcal{D}_W)^{ab}_{\alpha\beta}(x,y) = \delta^{ab}\delta_{\alpha\beta}\delta_{x,y} - K\sum_{\mu=\pm 1}^{2} V^{ab}_{\mu}(x)(1-\gamma_\mu)_{\alpha\beta}\delta_{x+\hat{\mu},y} \tag{42}$$

with $\gamma_{-\mu} \equiv -\gamma_\mu$. This reduces down to a version of the Wilson-Dirac operator that looks like what we had previously, modulo a few signs and conjugates, but more importantly agrees with Ref. [14]. This is because we can expand this action out and use our expression for $V_{-\mu}(x)$ to obtain:

$$\begin{aligned} (\mathcal{D}_W)^{ab}_{\alpha\beta}(x,y) &= \delta^{ab}\delta_{\alpha\beta}\delta_{x,y} - K\sum_{\mu=1}^{2}\left[V^{ab}_{\mu}(x)(1-\gamma_\mu)_{\alpha\beta}\delta_{x+\hat{\mu},y} + V^{ab}_{-\mu}(x)(1+\gamma_\mu)_{\alpha\beta}\delta_{x-\hat{\mu},y}\right] \\ &= \delta^{ab}\delta_{\alpha\beta}\delta_{x,y} - K\sum_{\mu=1}^{2}\left[V^{ab}_{\mu}(x)(1-\gamma_\mu)_{\alpha\beta}\delta_{x+\hat{\mu},y} + (V^\dagger_\mu)^{ab}(x-\hat{\mu})(1+\gamma_\mu)_{\alpha\beta}\delta_{x-\hat{\mu},y}\right] \\ &= \delta^{ab}\delta_{\alpha\beta}\delta_{x,y} - K\sum_{\mu=1}^{2}\left[V^{ab}_{\mu}(x)(1-\gamma_\mu)_{\alpha\beta}\delta_{x+\hat{\mu},y} + (V^\dagger_\mu)^{ab}(y)(1+\gamma_\mu)_{\alpha\beta}\delta_{x-\hat{\mu},y}\right] \end{aligned} \tag{43}$$

This is exactly the expression in Ref. [14]. Using the identification $V = V^*$, we can also relate this quite closely to Montvay's Dirac operator. They will end up being the same up to a phase convention on the adjoint field V, where we take $V \to V^\dagger = V^T$ and relabel $x \leftrightarrow y$:

$$\begin{aligned}(D_W)^{ab}_{\alpha\beta}(x,y) &= \delta^{ab}\delta_{\alpha\beta}\delta_{x,y} - K\sum_{\mu=1}^{2}\left[V^{ab}_\mu(x)(1-\gamma_\mu)_{\alpha\beta}\delta_{x+\hat{\mu},y} + (V^T_\mu)^{ab}(y)(1+\gamma_\mu)_{\alpha\beta}\delta_{x-\hat{\mu},y}\right] \\ &\xrightarrow{x\leftrightarrow y, V\to V^\dagger} \delta^{ab}\delta_{\alpha\beta}\delta_{x,y} - K\sum_{\mu=1}^{2}\left[(V^T_\mu)^{ab}(y)(1-\gamma_\mu)_{\alpha\beta}\delta_{y+\hat{\mu},x} + V^{ab}_\mu(x)(1+\gamma_\mu)_{\alpha\beta}\delta_{y,x+\hat{\mu}}\right].\end{aligned} \tag{44}$$

All this to say that the operators we've been studying are consistent. Let's use this form of the Wilson-Dirac operator from here on out:

$$\boxed{(D_W)^{ab}_{\alpha\beta} = \delta^{ab}\delta_{\alpha\beta}\delta_{x,y} - K\sum_{\mu=1}^{2}\left[V^{ab}_\mu(x)(1-\gamma_\mu)_{\alpha\beta}\delta_{x+\hat{\mu},y} + (V^T_\mu)^{ab}(y)(1+\gamma_\mu)_{\alpha\beta}\delta_{x-\hat{\mu},y}\right]} \tag{45}$$

We will also use the standard Wilson gauge action,

$$S_g = \beta\sum_{x\in\Lambda}\sum_{\mu<\nu}\left(1 - \frac{1}{N}\mathrm{Re}\,\mathrm{Tr}\,\mathcal{P}_{\mu\nu}(x)\right) \tag{46}$$

where $\mathcal{P}_{\mu\nu}$ is the plaquette in direction (μ,ν) at $x \in \Lambda$. Note that for $d = 2$, the sum is easy:

$$S_g = \beta\sum_{x\in\Lambda}\left(1 - \frac{1}{N}\mathrm{Re}\,\mathrm{Tr}\,\mathcal{P}(x)\right) \tag{47}$$

where $\mathcal{P}(x) \equiv \mathcal{P}_{01}(x)$ is the only direction of plaquette that can be formed on a 2d lattice. We think about $\mathcal{O}(a)$ improvement for this action later.

Additionally, we want to verify that the Dirac operator is γ_5-hermitian, i.e. that

$$\gamma_5\mathcal{D}\gamma_5 = \mathcal{D}^\dagger \iff (\gamma_5\mathcal{D}\gamma_5)^{ab}_{\alpha\beta}(x,y) = (\mathcal{D}^*)^{ba}_{\beta\alpha}(y,x). \tag{48}$$

Now, γ_5-hermicity implies that we can Hermitize the Dirac operator as $Q = Q^\dagger$, where

$$Q = \gamma_5\mathcal{D}. \tag{49}$$

However, the charge conjugation matrix is $C = \gamma_5$, so we note that Q also equals $C\mathcal{D}$. In Montvay's notes, he comments that $M = C\mathcal{D}$ is a skew-symmetric matrix that has a Pfaffian, hence we also expect that

$$Q^T = -Q. \tag{50}$$

This should be confirmed in the test cases. Note that because

$$\mathcal{D}^\dagger\mathcal{D} = \mathcal{D}^\dagger\gamma_5^\dagger\gamma_5\mathcal{D} = Q^\dagger Q \tag{51}$$

The operator K can also be considered as $K = Q^\dagger Q$, and the fermion Pfaffian should be considered as

$$\mathrm{Pf}(\mathcal{D}) \longrightarrow \mathrm{Pf}(Q) \tag{52}$$

since the Pfaffian is only defined for a skew-symmetric matrix.

TODO note that the x and y indices here are not the same as the boxed equation, I think because I was originally using a different definition Let us first consider the case of the identity gauge field, where the adjoint links are the identity $V_\mu^{ab}(x) = \delta^{ab3}$. In this case, we can use the fact that $\gamma_\mu^\dagger = \gamma_\mu$, $\{\gamma^\mu, \gamma_5\} = 0$ and $\gamma_5^2 = 1$ to show that:

$$\begin{aligned}
\gamma^5_{\rho\alpha}\mathcal{D}^{ab}_{\alpha\beta}(x,y)\gamma^5_{\beta\sigma} &= \delta^{ab}\gamma^5_{\rho\alpha}\left(\delta_{\alpha\beta}\delta_{x,y} - K\sum_{\mu=1}^{2}\left[(1-\gamma_\mu)_{\alpha\beta}\delta_{x,y+\hat{\mu}} + (1+\gamma_\mu)_{\alpha\beta}\delta_{x,y-\hat{\mu}}\right]\right)\gamma^5_{\beta\sigma} \\
&= \delta^{ab}\left(\delta_{\rho\sigma}\delta_{x,y} - K\sum_{\mu=1}^{2}\left[(1+\gamma_\mu)_{\rho\sigma}\delta_{x,y+\hat{\mu}} + (1-\gamma_\mu)_{\rho\sigma}\delta_{x,y-\hat{\mu}}\right]\right) \\
&= \delta^{ab}\left(\delta_{\rho\sigma}\delta_{x,y} - K\sum_{\mu=1}^{2}\left[(1+\gamma_\mu)^\dagger_{\rho\sigma}\delta_{y,x-\hat{\mu}} + (1-\gamma_\mu)^\dagger_{\rho\sigma}\delta_{y,x+\hat{\mu}}\right]\right) \\
&= \mathcal{D}^{ba}_{\sigma\rho}(y,x)^*
\end{aligned} \tag{53}$$

as desired (i.e. γ_5-hermiticity means that $\gamma_5\mathcal{D}\gamma_5 = \mathcal{D}^\dagger$, where the transpose on $\mathcal{D}$ in $\mathcal{D}^\dagger$ is over **all spin, color, and spacetime indices**).

For the case of an arbitrary gauge field, this is also not too difficult:

$$\begin{aligned}
\gamma^5\mathcal{D}(x,y)\gamma^5 &= \gamma^5\left(1_s1_c\delta_{x,y} - K\sum_{\mu=1}^{2}\left[V_\mu(x)(1_s-\gamma_\mu)\delta_{x+\hat{\mu},y} + V_\mu^T(y)(1_s+\gamma_\mu)\delta_{x-\hat{\mu},y}\right]\right)\gamma^5 \\
&= 1_s1_c\delta_{x,y} - K\sum_{\mu=1,2}\left[V_\mu(x)(1_s+\gamma_\mu)\delta_{x+\hat{\mu},y} + V_\mu^T(y)(1_s-\gamma_\mu)\delta_{x-\hat{\mu},y}\right] \\
&= 1_s1_c\delta_{x,y} - K\sum_{\mu=1,2}\left[(V_\mu^T)^\dagger(x)(1_s+\gamma_\mu)^\dagger\delta_{x+\hat{\mu},y} + V_\mu^\dagger(y)(1_s-\gamma_\mu)^\dagger\delta_{x-\hat{\mu},y}\right] \\
&= \left(1_s1_c\delta_{x,y} - K\sum_{\mu=1,2}\left[V_\mu^T(x)(1_s+\gamma_\mu)\delta_{y,x+\hat{\mu}} + V_\mu(y)(1_s-\gamma_\mu)\delta_{y,x-\hat{\mu}}\right]\right)^\dagger \\
&= \left(1_s1_c\delta_{x,y} - K\sum_{\mu=1,2}\left[V_\mu(y)(1_s-\gamma_\mu)\delta_{y+\hat{\mu},x} + V_\mu^T(x)(1_s+\gamma_\mu)\delta_{y-\hat{\mu},x}\right]\right)^\dagger \\
&= \mathcal{D}(y,x)^\dagger.
\end{aligned} \tag{54}$$

Here note that a full $\dagger$ acts on all indices $(a, b, \alpha, \beta, x, y)$, but to avoid confusion in the above calculation (Eq. (54)) we only use $\dagger$ to refer to the color and spin indices.

6.3 The Dirac operator and sparse matrices

We need to treat the Dirac operator $\mathcal{D}$ in an efficient way; to do this, we'll use a sparse matrix data structure, since most of its entries are zero. This is clearly seen in Eq. (36), since we the $\delta_{x,y\pm\hat{\mu}}$ means the field only interacts with its nearest neighbors. The best way to construct the Dirac operator is therefore in blocks. Consider $\mathcal{D}^{ab}_{\alpha\beta}(x,y)$ as a $d_{N_c}N_sLT \times d_{N_c}N_sLT$ matrix, where :

$$D_{ij} = \mathcal{D}^{ab}_{\alpha\beta}(x,y) \tag{55}$$

[3]Note that by the definition of the adjoint link, this is equivalent to the fundamental link field $U_\mu(x)$ being the identity as well.

where i and j are multi-indices encoding (a, α, x^μ) and (b, β, y^μ), respectively. For a natural spacetime blocking, these multi-indices traverse first in color, then in spin, then in the spatial dimension, and finally in the temporal direction, as is clear in Eq. (58). For a multi-index (a, α, x, t) (where here x is a scalar; we have made the spacetime coordinates explicit, with $x^\mu = (x, t)$), we can flatten it with:

$$i = a + d_{N_c} a + d_{N_c} N_s x + d_{N_c} N_s L t \tag{56}$$

where $d_{N_c} = (N_c^1 - 1)$ is the dimension of the adjoint representation of $SU(N_c)$. Likewise, given an index $i \in \{0, 1, ..., d_{N_c} N_s L T\}$, we can get back to the original multi-index (a, α, x, t) with the iterative algorithm:

$$\begin{aligned}
&\mathrm{t} \;\to\; \mathrm{i} \;//\; (d_{N_c} N_s L) \\
&\mathrm{i} \;=\; \mathrm{i} \;-\; \mathrm{t} \;*\; (d_{N_c} N_s L) \\
&\mathrm{x} \;\to\; \mathrm{i} \;//\; (d_{N_c} N_s) \\
&\mathrm{i} \;=\; \mathrm{i} \;-\; \mathrm{x} \;*\; (d_{N_c} N_s) \\
&\alpha \;\to\; \mathrm{i} \;//\; d_{N_c} \\
&\mathrm{a} \;\to\; \mathrm{i} \;\%\; d_{N_c}
\end{aligned}$$

This essentially just keeps dividing out by the correct block sizes for each step in the process.

We will use the bsr_matrix implementation in scipy. Consider the following construction:

```
indptr = np.array([0, 2, 3, 6])
indices = np.array([0, 2, 2, 0, 1, 2])
data = np.array([1, 2, 3, 4, 5, 6]).repeat(4).reshape(6, 2, 2)
mat = bsr_matrix((data,indices,indptr), shape=(6, 6))
```

Here, data is a list of 6 2D arrays (blocks), each of which are 2×2 and filled with the corresponding value $1, 2, 3, 4, 5, 6$. The matrix it constructs is a 6×6 matrix from these blocks 2×2 blocks, given as follows:

$$\mathrm{mat} = \begin{pmatrix} 1 & 1 & 0 & 0 & 2 & 2 \\ 1 & 1 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 & 3 & 3 \\ 4 & 4 & 5 & 5 & 6 & 6 \\ 4 & 4 & 5 & 5 & 6 & 6 \end{pmatrix} \tag{57}$$

Here indptr tells you where in indices to start. For each j, the interval [indptr[j], indptr[j + 1]) tells you which parts of the data are in the j^{th} row. Concretely, the non-zero elements (or blocks, if you're blocking) of row j are given by data[indptr[j] : indptr[j + 1]], with each one corresponding to column indices[indptr[j] : indptr[j + 1]].

Our implementation as an array will look like the following (here we use the example of $SU(2)$, so the adjoint indices $a \in \{0, 1, 2\}$, and the spinor indices $\alpha, \beta \in \{0, 1\}$, and let $\vec{1} \equiv \hat{0} = (1, 0)$):

$$D = \begin{pmatrix} D_{00}^{00}(\vec{0},\vec{0}) & D_{00}^{01}(\vec{0},\vec{0}) & D_{00}^{02}(\vec{0},\vec{0}) & D_{01}^{00}(\vec{0},\vec{0}) & D_{01}^{01}(\vec{0},\vec{0}) & D_{01}^{02}(\vec{0},\vec{0}) & D_{00}^{00}(\vec{0},\vec{1}) & ... \\ D_{00}^{10}(\vec{0},\vec{0}) & D_{00}^{11}(\vec{0},\vec{0}) & D_{00}^{12}(\vec{0},\vec{0}) & D_{01}^{10}(\vec{0},\vec{0}) & D_{01}^{11}(\vec{0},\vec{0}) & D_{01}^{12}(\vec{0},\vec{0}) & D_{00}^{10}(\vec{0},\vec{1}) & ... \\ D_{00}^{20}(\vec{0},\vec{0}) & D_{00}^{21}(\vec{0},\vec{0}) & D_{00}^{22}(\vec{0},\vec{0}) & D_{01}^{20}(\vec{0},\vec{0}) & D_{01}^{21}(\vec{0},\vec{0}) & D_{01}^{22}(\vec{0},\vec{0}) & D_{00}^{20}(\vec{0},\vec{1}) & ... \\ D_{10}^{00}(\vec{0},\vec{0}) & D_{10}^{01}(\vec{0},\vec{0}) & D_{10}^{02}(\vec{0},\vec{0}) & D_{11}^{00}(\vec{0},\vec{0}) & D_{11}^{01}(\vec{0},\vec{0}) & D_{11}^{02}(\vec{0},\vec{0}) & D_{10}^{00}(\vec{0},\vec{1}) & ... \\ D_{10}^{10}(\vec{0},\vec{0}) & D_{10}^{11}(\vec{0},\vec{0}) & D_{10}^{12}(\vec{0},\vec{0}) & D_{11}^{10}(\vec{0},\vec{0}) & D_{11}^{11}(\vec{0},\vec{0}) & D_{11}^{12}(\vec{0},\vec{0}) & D_{10}^{10}(\vec{0},\vec{1}) & ... \\ D_{10}^{20}(\vec{0},\vec{0}) & D_{10}^{21}(\vec{0},\vec{0}) & D_{10}^{22}(\vec{0},\vec{0}) & D_{11}^{20}(\vec{0},\vec{0}) & D_{11}^{21}(\vec{0},\vec{0}) & D_{11}^{22}(\vec{0},\vec{0}) & D_{10}^{20}(\vec{0},\vec{1}) & ... \\ D_{00}^{00}(\vec{1},\vec{0}) & D_{00}^{01}(\vec{1},\vec{0}) & D_{00}^{02}(\vec{1},\vec{0}) & D_{01}^{00}(\vec{1},\vec{0}) & D_{01}^{01}(\vec{1},\vec{0}) & D_{01}^{02}(\vec{1},\vec{0}) & D_{00}^{00}(\vec{1},\vec{1}) & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \end{pmatrix} \tag{58}$$

Note that between any two sites x and y, we have blocks of size $(N_c^2-1)N_s$ (6, in our case of $N_c=2$), given by:

$$D_{\text{block}}(x,y) = \begin{pmatrix} D_{00}^{00} & D_{00}^{01} & D_{00}^{02} & D_{01}^{00} & D_{01}^{01} & D_{01}^{02} \\ D_{00}^{10} & D_{00}^{11} & D_{00}^{12} & D_{01}^{10} & D_{01}^{11} & D_{01}^{12} \\ D_{00}^{20} & D_{00}^{21} & D_{00}^{12} & D_{01}^{20} & D_{01}^{21} & D_{01}^{22} \\ D_{10}^{00} & D_{10}^{01} & D_{10}^{02} & D_{11}^{00} & D_{11}^{01} & D_{11}^{02} \\ D_{10}^{10} & D_{10}^{11} & D_{10}^{12} & D_{11}^{10} & D_{11}^{11} & D_{11}^{12} \\ D_{10}^{20} & D_{10}^{21} & D_{10}^{12} & D_{11}^{20} & D_{11}^{21} & D_{11}^{22} \end{pmatrix} \tag{59}$$

where the indices (x,y) have been suppressed.

To compute the Pfaffian of $\mathcal{D}$, we need to consider an antisymmetric Dirac operator, since the Pfaffian is only defined for skew-symmetric matrices of even dimensionality. The way to construct such a matrix is to use the charge conjugation matrix C, through which the skew-symmetric Dirac operator $M=-M^T$ is defined as

$$M = C\mathcal{D} \tag{60}$$

The charge conjugation matrix is defined through the γ-matrices as

$$C^{-1}\gamma_\mu C = -\gamma_\mu^T. \tag{61}$$

The simplest choice for this is $C=\gamma_5$, which means that charge conjugation and chirality are the same in $d=2$.

6.4 Implementing RHMC

Important thing to consider: what gauge variables do we want to update? We have three choices: first, the fundamental gauge field $U_\mu(n)$; second, the adjoint gauge field $V_\mu^{ab}(n)$; third, the components of the fundamental, $U_\mu(n)=e^{i\omega_\mu^a(n)t^a}$

Now that we have a rational approximation to $K^{-1/4}$ (Eq.(31))[4], we need to determine how to use it. We start with careful definitions of all the variables in the system. The fields we are interested in are the pseudofermion fields $\Phi(n)$, and the adjoint gauge fields and their conjugate momenta $\Pi_\mu^a(n)$: TODO this is wrong...

$$U_\mu(n) = U_\mu^a(n)t^a \in SU(N) \qquad\qquad \Pi_\mu(n) = \Pi_\mu^a(n)t^a \in SU(N) \tag{62}$$

For HMC, we are interested in updating the system via the effective Hamiltonian,

$$H = \frac{1}{2}\sum_{n\in\Lambda}\text{Tr}[\Pi_\mu^2(n)] + S_{\text{eff}}[U,\Phi] = \frac{1}{2}\sum_n \text{Tr}[\Pi_\mu^2(n)] + S_g[U] + \Phi^\dagger K^{-1/4}\Phi. \tag{63}$$

Note that K can be defined in terms of $Q=\gamma_5\mathcal{D}$, the Hermitian Dirac operator, as $Q^\dagger Q$, or as the conventional $\mathcal{D}^\dagger\mathcal{D}$ we've been using. Each field is initialized at computer time $s=0$ as follows:

- $\Pi_\mu(n;s=0)$: This should have dimensions equal to the gauge field $U_\mu(x)$, since it is its conjugate momentum, and is distributed to a random Gaussian,

$$f(\Pi_\mu) = e^{-\text{Tr}[\Pi_\mu^2]} \tag{64}$$

[4]Later we will see that we also need a rational approximation to $K^{+1/8}$; this is done in the same way as $K^{-1/4}$, it just requires different coefficients.

Here $\Pi_\mu^2 \equiv \sum_{n\in\Lambda} \Pi_\mu(n)^2$. The easiest way to sample these coefficients is to generate a random element of the Lie algebra $\Pi_\mu^a(n)$, and to set $\Pi_\mu(n) = \Pi_\mu^a(n)t^a$, where now $P_\mu(n)$ has fundamental gauge indices. Using basic properties of t^a, we have:

$$\mathrm{Tr}[\Pi_\mu^2] = \sum_{n\in\Lambda} \mathrm{Tr}[(\Pi_\mu^a(n)t^a)(\Pi_\mu^b(n)t^b)] = \frac{1}{2}\sum_{n\in\Lambda}(\Pi_\mu^a(n))^2. \tag{65}$$

Note that the easiest way to do something like this is to initialize $\Pi_\mu(n)$ on each link, i.e. iterate over (n,μ) and generate $d_{N_c} = N_c^2 - 1$ random real numbers according to

$$f(\Pi_\mu(n)) = e^{-\mathrm{Tr}[\Pi_\mu(n)^2]} = e^{-\frac{1}{2}(\Pi_\mu^a(n))^2} \tag{66}$$

i.e. generate them according to a d_{N_c} dimensional normal distribution with mean 0 and covariance $1_{d_{N_c}}$.

- $U_\mu(n; s=0)$: Either perform a **hot start** and initialize $U_\mu(n)$ to a random $SU(N)$ gauge field, or do a **cold start** and set $U_\mu(n)$ equal to the identity at each link.

- Φ: The pseudofermion field is initialized to

$$\Phi(n; s=0) = K^{1/8}g(n) = r_{1/8}(K)g(n) \tag{67}$$

where $g(n)$ are random vectors drawn from a $2\times(N_c^2-1)$-dimensional Gaussian distribution TODO make sure this is correct. The reason for this is that the pseudofermions are distributed according to

$$P[\Phi] \propto e^{-\Phi^\dagger K^{-1/4}\Phi}. \tag{68}$$

Hence if we define

$$g(\Phi) = K^{-1/8}[U]\Phi \qquad\qquad \Phi(g) = K^{1/8}[U]g \tag{69}$$

then we see that

$$\mathbb{P}[g] = \mathbb{P}[\Phi]\left(\det\frac{\partial\Phi}{\partial g}\right)^{-1} = \det\left(K^{-1/8}[U]\right)e^{-g^\dagger g}. \tag{70}$$

Since $K^{1/8}[U]$ does not depend on our auxiliary variable g, we see that we can sample g from the distribution

$$\mathbb{P}[g] \propto e^{-g^\dagger g} \sim \mathcal{N} \tag{71}$$

where $\mathcal{N}$ is a Gaussian distribution, then simply transform back to Φ by

$$\Phi = K^{1/8}g = r_{1/8}(K)g. \tag{72}$$

This is why we needed a rational approximation to $K^{1/8}$ as well as $K^{-1/4}$.

TODO make sure the factor of 1/2 is correct and all derivatives are taken correctly We can derive the molecular dynamics (MD) time evolution equations from Hamilton's equations for the system, which are:

$$\begin{aligned}\delta_t\Pi_\mu(n) &= -\delta_{U_\mu(n)}H = -\frac{\delta S_g[U]}{\delta U_\mu(n)} - \frac{\delta}{\delta U_\mu(n)}\left(\Phi^\dagger K^{-1/4}\Phi\right)\\ \delta_t U_\mu(n) &= \delta_{\Pi_\mu(n)}H = \Pi_\mu(n).\end{aligned} \tag{73}$$

6.5 Updating with ω_μ

It is often easier to write out Hamilton's equations by updating the real coefficients in the algebra that go into computing $U_\mu(n)$. That is, we expand

$$U_\mu(n) = \exp\left(i\omega_\mu^a t^a\right) \tag{74}$$

for real-valued $\{\omega_\mu^a\}$, and use $\omega_\mu^a(n)$ as the dynamical variables to be updated in our system instead of $U_\mu(n)$. The conjugate momenta to $\omega_\mu^a(n)$ are defined as $P_\mu^a(n)$. We will interchangeably use the real $P_\mu^a(n)$ coefficients, which are the coordinates of the conjugate field in the Lie algebra, and an actual algebra-valued field $P_\mu(n) \in \mathfrak{su}(N_c)$, defined as

$$P_\mu(n) \equiv P_\mu^a(n) t^a \in \mathfrak{su}(N_c). \tag{75}$$

Note that $P_\mu(n)$ is **not conjugate** to $U_\mu(n)$, since P_μ is an element of the algebra and U_μ is an element of the group. Rather, $e^{iP_\mu(n)}$ is the conjugate variable to $U_\mu(n)$.

These equations must be integrated to perform MD evolution of the theory. To do this, we need to first compute the relevant forces that we will update our theory with, then write out the update steps. To initialize the fields, we'll initialize the pseudofermions in the same way, but the gauge fields will be initialized slightly differently:

- $\omega_\mu^a(n; s=0)$: Either perform a **hot start** and intialize $\omega_\mu^a(n)$ to a random field, or do a **cold start** and set $\omega_\mu(n) = 0$ completely (which corresponds to $U_\mu(n) = 1$).
- $P_\mu^a(n)$: Here we want to initialize these as part of a normal distribution, TODO

We want to update the system here with the Hamiltonian (here summing on μ and a is implied):

$$H = \frac{1}{2}\sum_{n\in\Lambda} P_\mu^a(n) P_\mu^a(n) + S_{\text{eff}}[U,\Phi] = \sum_n \text{Tr}[P_\mu(n)^2] + S_g[U] + \Phi^\dagger K^{-1/4}\Phi. \tag{76}$$

where note that U and K are functions of our "position" variables, $\omega_\mu^a(n)$, and the equality uses that $\text{Tr}[t^a t^b] = \frac{1}{2}\delta^{ab}$. Hamilton's equations hence yield:

$$\begin{aligned}
\delta_t P_\mu^a(n) &= -\delta_{\omega_\mu^a(n)} H = -\left(\frac{\delta S_g[\omega]}{\delta\omega_\mu^a(n)} + \Phi^\dagger \frac{\delta K[\omega]^{-1/4}}{\delta\omega_\mu^a(n)}\Phi\right) \equiv -F_\mu^a(n)[\omega,\Phi] \\
\delta_t \omega_\mu^a(n) &= \delta_{P_\mu^a(n)} H = P_\mu^a(n)
\end{aligned} \tag{77}$$

where the derivative is the force driving the conjugate momenta update,

$$F_\mu^a(n)[\omega,\Phi] = \frac{\delta S_g[\omega]}{\delta\omega_\mu^a(n)} + \Phi^\dagger \frac{\delta K[\omega]^{-1/4}}{\delta\omega_\mu^a(n)}\Phi \equiv (F_g)_\mu^a(n)[\omega] + (F_{\text{pf}})_\mu^a(n)[\omega,\Phi]. \tag{78}$$

Note that these forces depend functionally on ω and Φ, but in terms of indices they have an adjoint color, Lorentz, and spacetime index (a,μ,n). We will compute these forces in the next section for our choices of action.

We can also take these coordinates into the algebra if we wish,

$$\begin{aligned}
\delta_t P_\mu(n) &= -\delta_{\omega_\mu(n)} H = -\left(\frac{\delta S_g[\omega]}{\delta\omega_\mu(n)} + \Phi^\dagger \frac{\delta K[\omega]^{-1/4}}{\delta\omega_\mu(n)}\Phi\right) \equiv -F_\mu(n)[\omega,\Phi] \\
\delta_t \omega_\mu(n) &= \delta_{P_\mu(n)} H = P_\mu(n)
\end{aligned} \tag{79}$$

where here $P_\mu(n) = P_\mu^a(n)t^a \in \mathfrak{su}(N)$ and $\omega_\mu(n) = \omega_\mu^a(n)t^a \in \mathfrak{su}(N)$. The algebra representation is more compact in certain cases, and we will specify when this is the case. The force has the following structure,

$$(F_g)_\mu(n) = (F_g)_\mu^a(n)t^a \qquad\qquad (F_{\rm pf})_\mu(n) = (F_{\rm pf})_\mu^a(n)t^a \tag{80}$$

where each force is implied to be a functional of ω or (ω, Φ).

7 Gauge and pseudofermion forces

7.1 Gauge force (updating with V_μ^{ab})

We'll begin with the gauge force, since it's much easier to take a derivative of than the pseudofermion piece.

7.2 Pseudofermion force (updating with V_μ^{ab})

The difficult part here is to differentiate $\Phi^\dagger K^{-1/4}\Phi$ with respect to U_μ. We use the rational approximation and apply the chain rule for Lie derivatives:

$$\begin{aligned}
\frac{\delta}{\delta U_\mu(x)}\Phi^\dagger K^{-1/4}\Phi &= \frac{\delta}{\delta U_\mu(x)}\Phi^\dagger\left(\alpha_0 + \sum_{i=1}^P \frac{\alpha_i}{K+\beta_i}\right)\Phi = \sum_{i=1}^P \alpha_i \Phi^\dagger\left(\frac{\delta}{\delta U_\mu(x)}\frac{1}{K+\beta_i}\right)\Phi \\
&= \sum_{i=1}^P \alpha_i \left((K+\beta_i)^{-1}\Phi\right)^\dagger \frac{\delta K}{\delta U_\mu(x)}\left((K+\beta_i)^{-1}\Phi\right) \\
&\equiv \sum_{i=1}^P \alpha_i \psi_i^\dagger \frac{\delta K}{\delta U_\mu(x)}\psi_i
\end{aligned} \tag{81}$$

We see that we will need to determine two things: the Lie derivative $\delta K/\delta U_\mu(x)$, and the solution $\psi_i = (K+\beta_i)^{-1}\Phi$ to the equation $(K+\beta_i)\psi = \Phi$.

Simply applying $r_{-1/4}(K)$ or $r_{1/8}(K)$ to a vector Φ requires the same inversion (i.e., to initialize the pseudofermion field as $r_{1/8}(K)\Phi$). In this case, we have:

$$r(K)\Phi = \alpha_0\Phi + \sum_{i=1}^P \alpha_i \underbrace{(K+\beta_i)^{-1}\,\Phi}_{\psi_i} \tag{82}$$

We have the same situation where we need to solve the equation $(K+\beta_k)\psi_i = \Phi$ (no sum on i) for $\psi_i = (K+\beta_i)^{-1}\Phi$. We'll solve this equation by using the built-in CG solver in scipy.sparse.linalg.

To evaluate the Lie derivative, we'll work in terms of the Hermitian Dirac operator $Q = \gamma_5\mathcal{D}$, as the chain rule simplifies down:

$$\frac{\delta K}{\delta U_\mu(x)} = \frac{\delta}{\delta U_\mu(x)}(Q^\dagger Q) = \frac{\delta Q^\dagger Q}{\delta V_\nu^{ab}(y)}\frac{\delta V_\nu^{ab}(y)}{\delta U_\mu(x)} = 2Q\frac{\delta Q}{\delta V_\nu^{ab}}\frac{\delta V_\nu^{ab}(y)}{\delta U_\mu(x)} \tag{83}$$

Appealing to the definition of V, we can take the derivative with respect to $U_\mu(x)$:

$$V_\nu^{ab}(y) = 2\mathrm{Tr}[U_\nu^\dagger(y)t^a U_\nu(y)t^b], \tag{84}$$

We see that:

$$\frac{\delta V_\nu^{ab}(y)}{\delta U_\mu^{ij}(x)} = 2\left(t^b U_\mu^\dagger(x) t^a\right)_{ji} \delta_{xy}\delta_{\mu\nu} \iff \frac{\delta V_\nu^{ab}(y)}{\delta U_\mu(x)} = 2\left(t^a U_\mu(x) t^b\right)^* \delta_{xy}\delta_{\mu\nu} \tag{85}$$

where $i, j = 1, 2, ..., N_c$ are fundamental color indices for U_μ, using that $(t^a)^t = (t^a)^*$.

To differentiate Q, we recall the definition of the hermitian Wilson-Dirac operator as

$$Q_{\alpha\beta}^{ab}(x,y) = \delta^{ab}\delta_{\alpha\beta}\delta_{x,y} - K\sum_{\mu=1}^{2}\left[V_\mu^{ab}(x)[\gamma_5(1-\gamma_\mu)]_{\alpha\beta}\delta_{x+\hat{\mu},y} + V_\mu^{ba}(y)[\gamma_5(1+\gamma_\mu)]_{\alpha\beta}\delta_{x-\hat{\mu},y}\right] \tag{86}$$

where as usual we use $V = V^*$. Let's differentiate this with respect to $V_\nu^{cd}(z)$. We have:

$$\begin{aligned}\frac{\delta}{\delta V_\mu^{cd}(z)} Q_{\alpha\beta}^{ab}(x,y) &= -K\sum_{\nu=1}^{2}\delta_{\mu\nu}\left[\delta_{xz}\delta_{ca}\delta_{db}[\gamma_5(1-\gamma_\nu)]_{\alpha\beta}\delta_{x+\hat{\nu},y} + \delta_{zy}\delta_{cb}\delta_{da}[\gamma_5(1+\gamma_\nu)]_{\alpha\beta}\delta_{x-\hat{\nu},y}\right] \\ &= -K\left[\delta_{xz}\delta_{ca}\delta_{db}(1-\gamma_\mu)_{\alpha\beta}\delta_{x+\hat{\mu},y} + \delta_{yz}\delta_{cb}\delta_{da}(1+\gamma_\mu)_{\alpha\beta}\delta_{x-\hat{\mu},y}\right].\end{aligned} \tag{87}$$

So, we want to contract this with $Q^\dagger$ and the derivative of V_μ. Note that in terms of indices, this is:

$$K_{\alpha\beta}^{ab}(x,y) = \sum_z Q_{\alpha\gamma}^{ac}(x,z) Q_{\gamma\beta}^{cb}(z,y) \tag{88}$$

and we see the derivative looks like:

$$\frac{\delta}{\delta U_\mu^{ij}(z)} K_{\alpha\beta}^{ab}(x,y) = 2\sum_w Q_{\alpha\gamma}^{ac}(x,w)\frac{\delta Q_{\gamma\beta}^{cb}(w,y)}{\delta V_\mu^{de}(v)}\frac{\delta V_\mu^{de}(v)}{\delta U_\mu^{ij}(z)} \tag{89}$$

The way this is typically dealt with is to absorb the conjugating spinors into $\delta K/\delta U$, i.e. to evaluate $(Q\psi_i)^\dagger$ and ψ_i and put them on the left and right of the derivative. For each i, we have:

$$\begin{aligned}\psi_i^\dagger \frac{\delta K}{\delta U_\mu^{k\ell}(z)}\psi_i &= 2\sum_{\nu,\xi}(Q\psi_i)^\dagger \frac{\delta Q}{\delta V_\nu(\xi)}\frac{\delta V_\nu(\xi)}{\delta U_\mu^{k\ell}(z)}\psi_i \\ &= 2\sum_{x,y}\underbrace{\left(\sum_{x'}\left(Q(x,x')_{\alpha\beta}^{ab}(\psi_i)_\beta^b(x')\right)^\dagger\right)}_{(Q\psi_i)^\dagger(x)_\alpha^a}\underbrace{\left(\sum_{\nu,\xi}\frac{\delta Q(x,y)}{\delta V_\nu(\xi)}\frac{\delta V_\nu(\xi)}{\delta U_\mu^{k\ell}(z)}\right)_{\alpha\beta}^{ab}}_{\equiv(\mathcal{M}_\mu)_{\alpha\beta;k\ell}^{ab}(x,y;z)}(\psi_i)_\beta^b(y) \\ &= 2\sum_{a,\alpha,x}(Q\psi_i)^\dagger(x)_\alpha^a(\mathcal{M}_\mu\psi_i)_{\alpha;k\ell}^a(x;z)\end{aligned} \tag{90}$$

where we've made the indices explicit in the second equation. Forming the spinors on the left and right sides is easy to do, just by evaluating ψ_i with a shifted CG solver, then evaluating $Q\psi_i$. I'll make the sums explicit for the manipulations, because there are a lot of indices. Let's evaluate

$\mathcal{M}_\mu$:

$$\begin{aligned}
(\mathcal{M}_\mu)^{ab}_{\alpha\beta;k\ell}(x,y;z) &= \sum_{c,d,\nu}\sum_{\xi\in\Lambda}\frac{\delta Q^{ab}_{\alpha\beta}(x,y)}{\delta V^{cd}_\nu(\xi)}\frac{\delta V^{cd}_\nu(\xi)}{\delta U^{k\ell}_\mu(z)} \\
&= -K\sum_{c,d,\nu}\sum_{\xi\in\Lambda}[\delta_{x\xi}\delta_{ca}\delta_{db}(1-\gamma_\nu)_{\alpha\beta}\delta_{x+\hat{\nu},y}+\delta_{y\xi}\delta_{cb}\delta_{da}(1+\gamma_\nu)_{\alpha\beta}\delta_{x-\hat{\nu},y}]\left(2\left(t^cU_\mu(z)t^d\right)_{\ell k}\delta_{\nu\mu}\delta_{z\xi}\right) \\
&= -2K\sum_{c,d}[\delta_{xz}\delta_{ca}\delta_{db}(1-\gamma_\mu)_{\alpha\beta}\delta_{x+\hat{\mu},y}+\delta_{yz}\delta_{cb}\delta_{da}(1+\gamma_\mu)_{\alpha\beta}\delta_{x-\hat{\mu},y}]\left(t^cU_\mu(z)t^d\right)_{\ell k} \\
&= -2K\left[\delta_{xz}(1-\gamma_\mu)_{\alpha\beta}\left(t^aU_\mu(z)t^b\right)_{\ell k}\delta_{x+\hat{\mu},y}+\delta_{yz}(1+\gamma_\mu)_{\alpha\beta}\left(t^bU_\mu(z)t^a\right)_{\ell k}\delta_{x-\hat{\mu},y}\right]
\end{aligned} \tag{91}$$

So, we want to form $\mathcal{M}_\mu\psi$, which again has a lot of indices[5]. We have:

$$\begin{aligned}
(\mathcal{M}_\mu\psi_i)^a_{\alpha;k\ell}(x;z) &= \sum_{b,\beta}\sum_{y\in\Lambda}\mathcal{M}^{ab;\mu}_{\alpha\beta;k\ell}(x,y;z)(\psi_i)^b_\beta(y) \\
&= -2K\left[\delta_{xz}\left(t^aU_\mu(z)t^b\right)_{\ell k}(1-\gamma_\mu)_{\alpha\beta}(\psi_i)^b_\beta(z+\hat{\mu})+\delta_{x,z+\hat{\mu}}\left(t^bU_\mu(z)t^a\right)_{\ell k}(1+\gamma_\mu)_{\alpha\beta}(\psi_i)^b_\beta(z)\right]
\end{aligned} \tag{92}$$

7.3 Gauge force and ΔS (updating with $\omega^a_\mu(n)$)

Here, we need to compute the forces as derivatives in the coordinates $\omega^a_\mu(n)$, which are defined in Eq. (78). For the gauge force, recall we use the Wilson gauge action,

$$S_g = \beta\sum_{x\in\Lambda}\left(1-\frac{1}{N}\mathrm{Re}\,\mathrm{Tr}\,\mathcal{P}(x)\right) = \beta|\Lambda| - \frac{\beta}{2N}\sum_{x\in\Lambda}\mathrm{Tr}\left[\mathcal{P}(x)+\mathcal{P}^\dagger(x)\right] \tag{93}$$

where the plaquette $\mathcal{P}(x)=\mathcal{P}_{01}(x)=U_0(x)U_1(x+\hat{0})U_0^\dagger(x+\hat{1})U_1^\dagger(x)$. The derivative of this will simply be a staple in the corresponding direction that we differentiate with respect to.

$$\begin{aligned}
(F_g)^a_\mu(n)[\omega] = \frac{\partial S_g}{\partial\omega^a_\mu(n)} &= -\frac{\beta}{2N}\frac{\partial}{\partial\omega^a_\mu(n)}\sum_{x\in\Lambda}\mathrm{Tr}\left[\mathcal{P}(x)+\mathcal{P}(x)^\dagger\right] \\
&= -\frac{\beta}{2N}\frac{\partial}{\partial\omega^a_\mu(n)}\sum_{x\in\Lambda}\mathrm{Tr}\left[U_\rho(x)A_\rho(x)+A^\dagger_\rho(x)U^\dagger_\rho(x)\right] \\
&= -\frac{\beta}{2N}\mathrm{Tr}\left[(it^a)U_\mu(n)A_\mu(n)+A^\dagger_\mu(n)U^\dagger_\mu(n)(-it^a)\right] \\
&= -\frac{i\beta}{2N}\mathrm{Tr}\left[t^aU_\mu(n)A_\mu(n)-A^\dagger_\mu(n)U^\dagger_\mu(n)t^a\right]
\end{aligned} \tag{94}$$

where here $A_\mu(n)$ is the staple formed from removing the link $U_\mu(n)$ from the sum on plaquettes $\sum_{x\in\Lambda}\mathcal{P}(x)$, which is given by Eq. 4.20 in Gattringer,

$$\begin{aligned}
A_\mu(n) &= [U_\nu(n+\hat{\mu})U_{-\mu}(n+\hat{\mu}+\hat{\nu})U_{-\nu}(n+\hat{\nu})+U_{-\nu}(n+\hat{\mu})U_{-\mu}(n+\hat{\mu}-\hat{\nu})U_\nu(n-\hat{\nu})]\Big|_{\nu\neq\mu} \\
&= \left[U_\nu(n+\hat{\mu})U^\dagger_\mu(n+\hat{\nu})U^\dagger_\nu(n)+U^\dagger_\nu(n+\hat{\mu}-\hat{\nu})U^\dagger_\mu(n-\hat{\nu})U_\nu(n-\hat{\nu})\right]\Big|_{\nu\neq\mu}.
\end{aligned} \tag{95}$$

[5]Note that the indices (μ,z,k,ℓ) are from the differentiation with respect to $U_\mu(z)$ (and (k,ℓ) are fundamental $SU(N)$ indices), so they are carried through the whole computation.

Taking this expression to the algebra can be done with the following trick. Suppose we are expanding $c^a t^a \in \mathfrak{su}(N)$, where c^a are coefficients given by $c^a = \mathrm{Tr}[t^a \zeta]$, where $\zeta \in \mathfrak{su}(N)$ is some other element of the algebra. We can expand $\zeta = \zeta^b t^b$ and insert this into the expression,

$$c^a t^a = \mathrm{Tr}[t^a \zeta] t^a = \mathrm{Tr}[\zeta^b t^a t^b] t^a = \left(\frac{1}{2} \zeta^b \delta^{ab} \right) t^a = \frac{1}{2} \zeta. \tag{96}$$

We see that bringing the coefficients c^a back to the algebra after the trace just induces a factor of $\frac{1}{2}$. After cycling the t^a factors to the front of Eq. (94), we have an expression for the algebra-valued force,

$$(F_g)_\mu(n)[\omega] = (F_g)^a_\mu(n)[\omega] t^a = -\frac{i\beta}{4N} \left(U_\mu(n) A_\mu(n) - A^\dagger_\mu(n) U^\dagger_\mu(n) \right). \tag{97}$$

We also need an expression for the change in action if we update a single link so that we can do the accept-reject step. Here we assume that we began with the initial gauge field U, and we updated the link at (μ, n) to $U'_\mu(n)$, so the configurations U' and U only differ by a single link. We have:

$$\begin{aligned} \Delta S_g[U, U'] &= S_g[U'] - S_g[U] \\ &= -\frac{\beta}{N} \mathrm{Re}\, \mathrm{Tr} \left[(U'_\mu(n) - U_\mu(n)) A_\mu(n) \right] \end{aligned} \tag{98}$$

where again here $U_\mu(n)$ is the updated link.

7.4 Pseudofermion force and ΔS (updating with $\omega^a_\mu(n)$)

Here's the hard part. We need to compute the derivative of the pseudofermion action

$$S_{\mathrm{pf}}[\omega, \Phi] = \Phi^\dagger K[\omega]^{-1/4} \Phi \qquad\qquad (F_{\mathrm{pf}})^a_\mu(n) = \Phi^\dagger \frac{\delta K[\omega]^{-1/4}}{\delta \omega^a_\mu(n)} \Phi \tag{99}$$

where recall we approximate the rational function with the Zolotarev approximation,

$$K^{-1/4}[\omega]\Phi \approx r_{-1/4}(K[\omega])\Phi = \alpha_0 \Phi + \sum_{i=1}^{P} \alpha_i (K + \beta_i)^{-1} \Phi \tag{100}$$

Using the fact that for a Lie derivative,

$$\frac{\partial}{\partial \omega} A^{-1} = -A^{-1} \frac{\partial A}{\partial \omega} A^{-1}, \tag{101}$$

we have

$$\frac{\partial}{\partial \omega^a_\mu(n)} K^{-1/4} \Phi = \frac{\partial}{\partial \omega^a_\mu(n)} \left(\alpha_0 \Phi + \sum_{i=1}^{P} \alpha_i (K + \beta_i)^{-1} \Phi \right) = -\sum_{i=1}^{P} \alpha_i (K + \beta_i)^{-1} \frac{\partial K}{\partial \omega^a_\mu(n)} (K + \beta_i)^{-1} \Phi. \tag{102}$$

Now, we need to evaluate $\partial K / \partial \omega$, which can be done by exploiting the hermiticity of Q:

$$\begin{aligned} \frac{\partial K}{\partial \omega^a_\mu(n)} &= \frac{\partial}{\partial \omega^a_\mu(n)} Q^\dagger Q = \frac{\partial Q^\dagger}{\partial \omega^a_\mu(n)} Q + Q^\dagger \frac{\partial Q}{\partial \omega^a_\mu(n)} \\ &= \frac{\partial Q}{\partial \omega^a_\mu(n)} Q + Q^\dagger \frac{\partial Q^\dagger}{\partial \omega^a_\mu(n)} \\ &= 2\mathrm{Re} \left[Q \frac{\partial Q}{\partial \omega^a_\mu(n)} \right]. \end{aligned} \tag{103}$$

Putting this together, we have

$$\begin{aligned}(F_{\rm pf})^a_\mu(n) &= -2\sum_{i=1}^{P}\alpha_i {\rm Re}\left[\psi_i^\dagger Q[\omega]\frac{\partial Q[\omega]}{\partial\omega^a_\mu(n)}\psi_i\right] \\ &= -2\sum_{i=1}^{P}\alpha_i {\rm Re}\left[(Q[\omega]\psi_i)^\dagger \frac{\partial Q[\omega]}{\partial\omega^a_\mu(n)}\psi_i\right] \\ &= -2\sum_{i=1}^{P}\alpha_i {\rm Re}\left[(D[\omega]\psi_i)^\dagger \frac{\partial D[\omega]}{\partial\omega^a_\mu(n)}\psi_i\right]\end{aligned} \tag{104}$$

where (note that $(K+\beta_i)^{-1}$ is Hermitian)

$$\psi_i = (K+\beta_i)^{-1}\Phi \tag{105}$$

may be computed with a CG solver.

The remaining piece to compute is $\partial D[\omega]/\partial\omega^a_\mu(n)$. For the Wilson Dirac operator D_W, we will use the fact that

$$\begin{aligned}\frac{\partial}{\partial\omega^a_\mu(n)}V^{bc}_\nu(x) = 2\frac{\partial}{\partial\omega^a_\mu(n)}{\rm Tr}\left(U^\dagger_\nu(x)t^bU_\nu(x)t^c\right) &= 2i\delta_{nx}\delta_{\mu\nu}{\rm Tr}\left(-t^aU^\dagger_\nu(x)t^bU_\nu(x)t^c + U^\dagger_\nu(x)t^bt^aU_\nu(x)t^c\right) \\ \frac{\partial}{\partial\omega^a_\mu(n)}(V^{bc}_\nu)^T(y) &= 2i\delta_{ny}\delta_{\mu\nu}{\rm Tr}\left(-t^aU^\dagger_\nu(y)t^cU_\nu(y)t^b + U^\dagger_\nu(y)t^ct^aU_\nu(y)t^b\right)\end{aligned} \tag{106}$$

We have:

$$\begin{aligned}\frac{\partial}{\partial\omega^a_\mu(n)}(D_W)^{bc}_{\beta\gamma}(x,y) &= -K\sum_{\nu=1}^{2}\frac{\partial}{\partial\omega^a_\mu(n)}\left[V^{bc}_\nu(x)(1-\gamma_\nu)_{\beta\gamma}\delta_{x+\hat\nu,y} + (V^T_\nu)^{bc}(y)(1+\gamma_\nu)_{\beta\gamma}\delta_{x-\hat\nu,y}\right] \\ &= 2iK\Big[\delta_{nx}{\rm Tr}\left(t^ct^aU^\dagger_\mu(x)t^bU_\mu(x) - t^bt^aU_\mu(x)t^cU^\dagger_\mu(x)\right)(1-\gamma_\mu)_{\beta\gamma}\delta_{x+\hat\mu,y} \\ &\quad + \delta_{ny}{\rm Tr}\left(t^bt^aU^\dagger_\mu(y)t^cU_\mu(y) - t^ct^aU_\mu(y)t^bU^\dagger_\mu(y)\right)(1+\gamma_\mu)_{\beta\gamma}\delta_{x-\hat\mu,y}\Big]\end{aligned} \tag{107}$$

We can apply this to ψ_i to determine

$$\begin{aligned}\left(\frac{\partial D}{\partial\omega^a_\mu(n)}\psi_i\right)^b_\beta(x) &= \sum_y\frac{\partial}{\partial\omega^a_\mu(n)}(D_W)^{bc}_{\beta\gamma}(x,y)\psi^c_\gamma(y) \\ &= 2iK\Big[\delta_{nx}{\rm Tr}\left(t^a\left(U^\dagger_\mu(n)t^bU_\mu(n)t^c - U_\mu(n)t^cU^\dagger_\mu(n)t^b\right)\right)(1-\gamma_\mu)_{\beta\gamma}\psi^c_\gamma(n+\hat\mu) \\ &\quad + \delta_{n+\hat\mu,x}{\rm Tr}\left(t^a\left(U^\dagger_\mu(n)t^cU_\mu(n)t^b - U_\mu(n)t^bU^\dagger_\mu(n)t^c\right)\right)(1+\gamma_\mu)_{\beta\gamma}\psi^c_\gamma(n)\Big]\end{aligned} \tag{108}$$

We can almost simplify this down using the definition $V^{ab}_\mu(n) = 2{\rm Tr}[U^\dagger_\mu(n)t^aU_\mu(n)t^b]$, but the t^a in front of the pieces inside the trace makes this impossible. However, we can simplify it a bit. We define the tensor

$$\mathcal{W}^{ab}_\mu \equiv U^\dagger_\mu(n)t^aU_\mu(n)t^b - U_\mu(n)t^bU^\dagger_\mu(n)t^a \qquad {\rm Tr}\,\mathcal{W}^{ab}_\mu = \frac{1}{2}(V^{ab}_\mu - V^{ab}_\mu) = 0. \tag{109}$$

Note that $\mathcal{W}_\mu^{ab}$ also has fundamental color indices, $(\mathcal{W}_\mu^{ab})_{ij}(n)$. This definition allows us to simplify down our expression for $\partial D/\partial\omega$:

$$\begin{aligned}\left(\frac{\partial D}{\partial\omega_\mu^a(n)}\psi_i\right)_\beta^b(x) = 2iK\Big[&\delta_{nx}\mathrm{Tr}\left(t^a\mathcal{W}_\mu^{bc}(n)\right)(1-\gamma_\mu)_{\beta\gamma}\psi_\gamma^c(n+\hat{\mu}) \\ &+\delta_{n+\hat{\mu},x}\mathrm{Tr}\left(t^a\mathcal{W}_\mu^{cb}(n)\right)(1+\gamma_\mu)_{\beta\gamma}\psi_\gamma^c(n)\Big].\end{aligned} \tag{110}$$

The best way to compute this is thus to precompute the traceless quantity $\mathcal{W}_\mu$, then form the appropriate tensor contractions. TODO this should be something we can test with autodiff

We conclude this calculation with a bit of bookkeeping, combining our previous results and making indices explicit:

$$\begin{aligned}(F_{\mathrm{pf}})_\mu^a(n) &= -2\sum_{i=1}^P\alpha_i\mathrm{Re}\left[(D[\omega]\psi_i)^\dagger\frac{\partial D[\omega]}{\partial\omega_\mu^a(n)}\psi_i\right] \\ &= -2\sum_{i=1}^P\alpha_i\sum_x\mathrm{Re}\left[\left((D[\omega]\psi_i)^\dagger\right)_\beta^b(x)\left(\frac{\partial D[\omega]}{\partial\omega_\mu^a(n)}\psi_i\right)_\beta^b(x)\right] \\ &= -4iK\sum_{i=1}^P\alpha_i\sum_x\mathrm{Re}\Big[\left((D[\omega]\psi_i)^\dagger\right)_\beta^b(x)\Big[\delta_{nx}\mathrm{Tr}\left(t^a\mathcal{W}_\mu^{bc}(n)\right)(1-\gamma_\mu)_{\beta\gamma}\psi_\gamma^c(n+\hat{\mu}) \\ &\qquad +\delta_{n+\hat{\mu},x}\mathrm{Tr}\left(t^a\mathcal{W}_\mu^{cb}(n)\right)(1+\gamma_\mu)_{\beta\gamma}\psi_\gamma^c(n)\Big]\Big].\end{aligned} \tag{111}$$

Simplifying the sum yields the following result for the pseudofermion force:

$$\boxed{\begin{aligned}(F_{\mathrm{pf}})_\mu^a(n) = -4iK\sum_{i=1}^P\alpha_i\mathrm{Re}\Big[&\left((D\psi_i)^\dagger\right)^b(n)\,\mathrm{Tr}\left(t^a\mathcal{W}_\mu^{bc}(n)\right)(1-\gamma_\mu)\psi^c(n+\hat{\mu}) \\ &+\left((D\psi_i)^\dagger\right)^b(n+\hat{\mu})\,\mathrm{Tr}\left(t^a\mathcal{W}_\mu^{cb}(n)\right)(1+\gamma_\mu)\psi^c(n)\Big]\end{aligned}} \tag{112}$$

TODO we also want the change in the action if we change 1 link

7.5 RHMC and MD evolution

Now that we've computed the pseudofermion forces, we need to do the leapfrog update. Let ϵ be the step size, and suppose we wish to perform n steps.

7.6 Computing the Pfaffian

We need to compute the Pfaffian of the skew-symmetric matrix $Q = \gamma_5\mathcal{D}$. The typical method to compute the Pfaffian of a large skew-symmetric matrix is to consider its LU decomposition. For a skew-symmetric matrix, one can show that the LU decomposition reduces to

$$Q = PJP^T \tag{113}$$

where P is lower triangular (note any lower triangular matrix satisfies $P^{-1} = P^T$) and J is a tridiagonal[6] matrix with trivial Pfaffian, $\mathrm{Pf}[J] = 1$. Passing to the Pfaffian, we see that given this

[6] A tridiagonal matrix is one with non-zero elements only on its diagonal, its super diagonal (elements above the diagonal) and its subdiagonal (elements below its diagonal).

decomposition, this provides us with a simple way to compute Pf$[Q]$

$$\mathrm{Pf}[Q] = \mathrm{Pf}[PJP^T] = \mathrm{Pf}[P]^2 \underbrace{\mathrm{Pf}[J]}_{1} = \det[P] = \prod_i P_{ii} \tag{114}$$

If we know P, the Pfaffian of Q is simply the product of the diagonal elements of P. The explicit algorithm, with an example performed for Wilson fermions, is laid out in Ref. [15].

So, the problem of computing Pf$[Q]$ is equivalent to determining the LU decomposition of Q. The algorithm to perform the LU decomposition of $Q = PJP^T$ is performed as follows. Let $p_{ij} = (P)_{ij}$ be the elements of the lower-triangular matrix P. The idea here is that we can simply solve a system of equations to constrain P, but naively this system of equations is over-determined. In order to make it a valid system of equations, we can constrain P by, for each odd $i = 1, 3, 5, ..., N-1$, setting the diagonal element and the element below it to zero:

$$p_{ii} = 1 \qquad\qquad p_{i+1,i} = 0. \tag{115}$$

Furthermore, we also need the constraint that $p_{ij} = 0$ for each $i < j$, i.e. that p is a lower-triangular matrix. So, the matrix P we are aiming to construct has the following form, where $*$ denotes an undetermined value that is computed in the algorithm:

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & ... \\ 0 & * & 0 & 0 & 0 & 0 & ... \\ * & * & 1 & 0 & 0 & 0 & ... \\ * & * & 0 & * & 0 & 0 & ... \\ * & * & * & * & 1 & 0 & ... \\ * & * & * & * & 0 & * & ... \end{pmatrix}. \tag{116}$$

We see that :

- In odd columns $(p_{ji})_j$ with $i \in \{1, 3, 5, ..., N-1\}$, the diagonal element p_{ii} is set to 1, and the next element $p_{i+1,i}$ is set to 0. All other elements are undetermined.
- The odd columns $(p_{ji})_j$ with $i \in \{2, 4, 6, ..., N\}$ are completely unspecified, up to the requirement that P be a lower triangular matrix.

The algorithm proceeds in pairs of columns. We let $\sum_{k=1}^{m\prime}$ denote a sum over only odd values $k = 1, 3, 5, ..., m$, and we iterate over odd $i = 1, 3, 5, ..., N-1$. For each $i = 1, 3, 5, ..., N$, we have three steps:

1. Set the $i+1$ (right) column. For each $j = i+1, i+2, ..., N$, compute:

$$p_{j,i+1} = a_{ij} - \sum_{k=1}^{i-1\prime} \left(p_{ik} p_{j,k+1} - p_{i,k+1} p_{jk} \right). \tag{117}$$

2. If $p_{i+1,i+1}$ is zero, pivot the columns. The point of this is that in the next step, $p_{i+1,i+1}$ (the diagonal even element which has just been determined) may be zero. In order to prevent this, one needs to pick a column in which this is not true, and pivot the columns. To do this, we search over all the entires of $(p_{j,i+1})_j$ that have just been determined, and denote $j_{\max}$ as the index with the maximum value of this quantity:

$$j_{\max} = \underset{j}{\mathrm{argmax}} |p_{j,i+1}|. \tag{118}$$

The idea now is to permute the columns of the matrix to swap the $j_{\max}$ row of Q with the $i+1$ row of Q. This will guarantee that, after permutation, the matrix element $|p'_{i+1,i+1}|$ is non-zero, where $'$ denotes that this matrix has been permuted.

Formally, let τ be the transposition $\tau = (i+1, j_{\max})$.

3. Set the ith (left) column. Note that the first two elements are already determined to be 1 and 0. For each $j = i+2, i+3, ..., N$, we hence compute

$$(-p_{i+1,i+1})p_{ji} = a_{i+1,j} - \sum_{k=1}^{i-2\prime} \left(p_{i+1,k}p_{j,k+1} - p_{i+1,k+1}p_{jk}\right). \tag{119}$$

Here we use the assumption that $p_{i+1,i+1}$ is non-zero. In the case where it is, we pivot the columns. If it turns out that $p_{i+1,i+1} = 0$ for any permutation, then it can be shown that the original matrix A must be singular, hence if A is antisymmetric and non-singular, then this algorithm may always be used.

From this construction, one can show that the trivial antisymmetric matrix J, defined componentwise for each $i \in \{0, 1, ..., N-1\}$ as

$$J_{i,i-(-1)^i} = -(-1)^i \implies J = \text{diag}\left[\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, ..., \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right], \tag{120}$$

satisfies the defining relation that $Q = PJP^T$.

8 The path forward

We want to work up to simulating Aleksey's 2d adjoint QCD theory. There are a few clear steps:

- Write down a discretized action for the fermionic part of the theory, with interactions turned off ($c_1 = c_2 = 0$). Here we'll want to consider any of the following: Wilson, Wilson-Clover, twisted mass, domain wall, overlap. Wilson is the simplest, but it does not seem that a Wilson action with one Majorana fermion in 2d has been written down (it might not play nice with the Majorana properties). It looks like there is a no-go theorem for removing the doublers in $d = 8k, 8k+1$ dimensions [16], but that shouldn't affect us in 2d.

- Use RHMC to simulate gauge field ensembles for the theory with no interactions ($c_1 = c_2 = 0$), which is standard 2d adjoint QCD. Monitor the sign of the Pfaffian using spectral flow to ensure there is no sign problem. Run this at a variety of different N and see how the time scales (the Pfaffian is harder to simulate than the determinant).

- Compute observables in the standard 2d adjoint QCD theory. Study confinement order parameters and chiral symmetry breaking.

- Write down and implement the overlap action for a Majorana fermion. We'll want to use this to have a semblance of chiral symmetry.

- Add four-fermion interactions to the theory and simulate the wish list.

obvious first step is to simulate the theory with parameters $c_1 = c_2 = 0$, i.e. with all four-fermion interactions turned off.

References

[1] Zohar Komargodski, Kantaro Ohmori, Konstantinos Roumpedakis, and Sahand Seifnashri. Symmetries and strings of adjoint QCD_2. *JHEP*, 03:103, 2021. doi: 10.1007/JHEP03(2021)103.

[2] Ross Dempsey, Igor R. Klebanov, Loki L. Lin, and Silviu S. Pufu. Adjoint Majorana QCD_2 at finite N. *JHEP*, 04:107, 2023. doi: 10.1007/JHEP04(2023)107.

[3] John C. Donahue and Sergei Dubovsky. Confining Strings, Infinite Statistics and Integrability. *Phys. Rev. D*, 101(8):081901, 2020. doi: 10.1103/PhysRevD.101.081901.

[4] John C. Donahue and Sergei Dubovsky. Classical Integrability of the Zigzag Model. *Phys. Rev. D*, 102(2):026005, 2020. doi: 10.1103/PhysRevD.102.026005.

[5] Aleksey Cherman, Theodore Jacobson, Yuya Tanizaki, and Mithat Ünsal. Anomalies, a mod 2 index, and dynamics of 2d adjoint QCD. *SciPost Phys.*, 8(5):072, 2020. doi: 10.21468/SciPostPhys.8.5.072.

[6] Davide Gaiotto, Anton Kapustin, Nathan Seiberg, and Brian Willett. Generalized Global Symmetries. *JHEP*, 02:172, 2015. doi: 10.1007/JHEP02(2015)172.

[7] Xie Chen, Zheng-Cheng Gu, and Xiao-Gang Wen. Classification of gapped symmetric phases in one-dimensional spin systems. *Physical Review B*, 83(3), jan 2011. doi: 10.1103/physrevb.83.035107. URL `https://doi.org/10.1103%2Fphysrevb.83.035107`.

[8] David Tong. David tong: Lectures on gauge theory, 2018. URL `https://www.damtp.cam.ac.uk/user/tong/gaugetheory.html`.

[9] Istvan Montvay. Majorana fermions on the lattice. 8 2001.

[10] Andreas Stathopoulos and James R. McCombs. Primme: Preconditioned iterative multimethod eigensolver?methods and software description. *ACM Trans. Math. Softw.*, 37(2), apr 2010. ISSN 0098-3500. doi: 10.1145/1731022.1731031. URL `https://doi.org/10.1145/1731022.1731031`.

[11] M. A. Clark and A. D. Kennedy. Accelerating dynamical fermion computations using the rational hybrid Monte Carlo (RHMC) algorithm with multiple pseudofermion fields. *Phys. Rev. Lett.*, 98:051601, 2007. doi: 10.1103/PhysRevLett.98.051601.

[12] Martin Luscher. Computational Strategies in Lattice QCD. In *Les Houches Summer School: Session 93: Modern perspectives in lattice QCD: Quantum field theory and high performance computing*, pages 331–399, 2 2010.

[13] Christof Gattringer and Christian B. Lang. *Quantum chromodynamics on the lattice*, volume 788. Springer, Berlin, 2010. ISBN 978-3-642-01849-7, 978-3-642-01850-3. doi: 10.1007/978-3-642-01850-3.

[14] Georg Bergner, Juan Camilo Lopez, Stefano Piemonte, and Ivan Soler Calero. Lattice simulations of adjoint QCD with one Dirac overlap fermion. *Phys. Rev. D*, 106(9):094507, 2022. doi: 10.1103/PhysRevD.106.094507.

[15] Jurgen Rubow and Ulli Wolff. A Factorization algorithm to compute Pfaffians. *Comput. Phys. Commun.*, 182:2530–2532, 2011. doi: 10.1016/j.cpc.2011.07.010.

[16] Hiroshi Suzuki. A No-go theorem for the Majorana fermion on the lattice. *Prog. Theor. Phys.*, 112:855–861, 2004. doi: 10.1143/PTP.112.855.