A solution for infinite variance problem of fermionic observables

Hyunwoo Oh
in collaboration with Andrei Alexandru, Paulo Bedaque, Andrea Carosso

based on “Infinite variance problem in fermion models”
Phys. Rev. D 107, 094502 (2022)

August 2 2023, Lattice 2023
Auxiliary Field for Fermionic Monte Carlo

• For fermionic quartic interactions, we use the Hubbard-Stratonovich (HS) transformation for Monte Carlo calculation

\[
\int d\bar{\eta}d\eta \ e^{-\bar{\eta}M\eta} = \det M \rightarrow \text{No more Grassmann}
\]

• Two ways:

1. Discrete

\[
\exp[-\hat{n}_\uparrow\hat{n}_\downarrow] \propto \sum_{\sigma=\pm 1} \exp[2a\sigma(\hat{n}_\uparrow - \hat{n}_\downarrow) - \frac{1}{2}(\hat{n}_\uparrow + \hat{n}_\downarrow)]
\]

2. Continuous

\[
\exp[\frac{1}{2}\hat{X}^2] \propto \int d\phi \exp[-\frac{1}{2}\phi^2 - \phi\hat{X}]
\]

Continuous Auxiliary Field is Needed

Hybrid Monte Carlo

- Higher acceptance rate
- Minimize correlation

![Graph showing curves of constant acceptance rate](image)

Fig 2. Curves of constant acceptance rate are shown as a function of the integration step size $\delta t$ on the lattices of size $4^4$, $8^4$ and $12^4$ in quenched QED at a coupling of $\beta = 0.97$.

Sign Problem

- Contour deformation approach

A. Alexandru et al, Rev. Mod. Phys. 94, 015006 (2022)
Hubbard Model Hamiltonian with half-filling at $\mu = 0$:

$$H = -\kappa \sum_{\langle x,y \rangle} (\hat{\psi}_x^\dagger \hat{\psi}_{y,x} + \hat{\psi}_{x,y}^\dagger \hat{\psi}_{x,y}) + U \sum_x (\hat{\psi}_{x,x}^\dagger \hat{\psi}_{x,x} - \frac{1}{2})(\hat{\psi}_{x,x}^\dagger \hat{\psi}_{x,x} - \frac{1}{2}) - \mu \sum_x (\hat{\psi}_{x,x}^\dagger \hat{\psi}_{x,x} + \hat{\psi}_{x,x}^\dagger \hat{\psi}_{x,x} - 1)$$

- We will first consider the $\mu = 0$ case, sign problem-free.
Path Integral Formulation

- Fermion coherent state formalism:

\[
\text{Tr}(e^{-\beta H}) = \int d\bar{\eta} d\eta \ e^{-\bar{\eta}\eta} \langle -\eta \mid e^{-\beta H} \mid \eta \rangle \\
= \int [d\bar{\eta} d\eta] \ e^{-\bar{\eta}\eta} \langle \eta_N \mid e^{-\epsilon H} \mid \eta_{N-1} \rangle \cdots \langle \eta_1 \mid e^{-\epsilon H} \mid \eta_0 \rangle \text{ with } \beta = \epsilon N
\]

- The effective action:

\[
S = S_0(\phi) - \log \det \ M_1(\phi) M_2(\phi) \text{ with } S_0(\phi) = -\beta \sum_{i=1}^{N} \cos(\phi_i) \\
M_a(\phi) = 1 + B_a(\phi_N) \cdots B_a(\phi_1)
\]
Monte Carlo Method

• With the importance sampling,

\[ \bar{\mathcal{O}} = \int D\phi \, \mathcal{O}(\phi) \frac{e^{-S[\phi]}}{\int D\phi \, e^{-S[\phi]}} \approx \frac{1}{n} \sum_{a=1}^{n} \mathcal{O}(\phi_a) \]

• As the number of samples increases,

\[ \bar{\mathcal{O}} = \langle \mathcal{O} \rangle + O\left( \frac{1}{\sqrt{n}} \right) \]
Infinite Variance Problem

- Error bar doesn’t decrease / Variance doesn’t converge

near exceptional configurations
Determinant and Diverging Variance

- Monte Carlo sampling cannot choose $\det M = 0$ configuration, but the neighborhood of zero points is dangerous.
Infinite Variance Problem

For a fermionic observable $\mathcal{O}$,

$$\mathcal{O}(\phi) \sim \frac{1}{\det M} f(M_{ij}(\phi))$$

large near $\det M = 0$

Expectation value is finite:

$$\langle \mathcal{O} \rangle \sim \int D\phi \ e^{-S_0(\phi)} \det M \frac{1}{\det M} f(M_{ij}(\phi))$$

Variance can be divergent:

$$\langle \mathcal{O}^2 \rangle \sim \int D\phi \ e^{-S_0(\phi)} \det M \frac{1}{(\det M)^2} f^2(M_{ij}(\phi)) \rightarrow \infty$$
Fermionic Observable without Diverging Variance

- Hubbard Model: \( \det M = \det M_1 \det M_2 \)

Density \( \mathcal{O}(\phi) = \frac{1}{V} \sum_x \langle n^\uparrow(x) + n^\downarrow(x) \rangle_F = \frac{1}{V} \sum_x [M_2^{-1}(\phi)_{x,x} - M_1^{-1}(\phi)_{x,x}] \)

Only one of M's
Reweighting and Overlap Problem

- **Reweighting:** \[ \langle \mathcal{O} \rangle = \int D\phi \ \mathcal{O}(\phi)p(\phi) = \frac{\langle \mathcal{O} \ p/q \rangle_q}{\langle p/q \rangle_q} \]

- If a new distribution is **too different** from the original, it cannot sample well

\[
\langle x \rangle = \frac{\int dy \ y \ e^{-(y-\mu)^2}}{\int dy \ e^{-(y-\mu)^2}} = \mu
\]

\[
= \frac{\int dy \ y \ e^{-(y-\mu)^2} \ e^{-y^2}}{\int dy \ e^{-y^2} \ e^{-y^2}}
\]
Extra Time Slice Method

• We want to use a different distribution from the original, but not too much

• Add one more auxiliary field along time direction:

\[
Z = \int [d\phi]_N e^{-S_0(\phi)} \det M_N(\phi) \frac{F(\phi)}{F(\phi)}^{(=1)}
\]

\[
= \int [d\phi]_N d\phi^* \; R(\phi) \; e^{-S_0(\phi,\phi^*)} \det M_{N+1}(\phi, \phi^*)
\]

with \( F(\phi) \equiv \int d\phi^* \; e^{-S_0(\phi^*)} \det M_{N+1}(\phi, \phi^*) \)

with \( R(\phi) = \det M_N(\phi)/F(\phi) \)

Expansion Method

- Integrate $F(\phi)$ analytically and expand it ($\epsilon = \beta/N$)

\[
F(\phi) \equiv \int d\phi^* \ e^{-S_0(\phi^*)} \det M_{N+1}(\phi, \phi^*)
\]

**BSS Formula**

\[
= \text{Tr}[e^{-\tilde{H}_2}e^{-\epsilon H_4}B(\phi_N)\cdots B(\phi_1)] = \text{Tr}[(1 - \epsilon H)B(\phi_N)\cdots B(\phi_1)] + O(\epsilon^2)
\]

\[
= (1 - \epsilon H(\phi)) \det M_N(\phi) + O(\epsilon^2)
\]

- Disadvantage:

1. Perturbative error
2. $F(\phi)$ can be 0 or negative
3. Difficult to calculate at higher orders

Sub-Monte Carlo Method

- Estimate the integral using Monte Carlo

\[ F(\phi) \equiv \int d\phi^* \ e^{-S_0(\phi^*)} \det M_{N+1}(\phi, \phi^*) \approx \frac{1}{n_{sub}} \sum_{i=1}^{n_{sub}} \det M_{N+1}(\phi, \phi_i^*) \]

\[ M_a(\phi) = 1 + B_a(\phi_{N+1})B_a(\phi_N)\cdots B_a(\phi_1) \]

- Advantage:
  1. No approximation
  2. \( F(\phi) > 0 \) since \( \det M \geq 0 \)

A. Alexandru, P. Bedaque, A. Carosso, HO, Phys. Rev. D 107, 094502 (2022)
Biased Estimation

- We need to calculate $1/F(\phi)$, not $F(\phi)$ ($R(\phi) \equiv \det M_N(\phi)/F(\phi)$)

- Just taking an inverse is biased:

$$\langle \frac{1}{A} \rangle = \frac{1}{\langle A \rangle} - \left( \frac{\overline{A} - \langle A \rangle}{\langle A^2 \rangle} \right) + \left( \frac{(\overline{A} - \langle A \rangle)^2}{\langle A^3 \rangle} \right) + \ldots$$

$$= \frac{1}{\langle A \rangle} + O\left(\frac{1}{n}\right) \quad \Rightarrow \quad \text{wrong value}$$
Unbiased Estimation

- Unbiased estimator of $1/\langle A \rangle$:

$$\hat{\xi}_A \equiv \frac{w}{q_n} \prod_{i=1}^{n} (1 - wA_i)$$

- Optimized choice of $w$ and $q_n$ (with respect to cost minimization):

$$w = \min \left\{ \frac{1}{k\bar{A}}, \frac{\bar{A}}{A^2}, \frac{1}{A_{\max}} \right\},$$

$$p = 1 - \left[ 1 - 2w\bar{A} + w^2\bar{A}^2 \right]^{\frac{1}{2}}, \quad q_n = p(1 - p)^n$$

where $n$ is from the geometric distribution with $p$
Result
Finite Variance

- Variance is not diverging

\[ D(\phi) = \frac{1}{V} \sum_x \langle n_1(x)n_1(x) \rangle_F \]
\[ = \frac{1}{V} \sum_x [1 - M_1^{-1}(\phi)_{x,x}] M_2^{-1}(\phi)_{x,x} \]

4x4 lattice, \( U/\kappa = 8, T/\kappa = 0.5, \mu = 0 \)
Finite Variance

- Exceptional configurations are reweighted well

4x4 lattice, $U/\kappa = 8$, $T/\kappa = 0.5$, $\mu = 0$
Comparison between Two Methods

- Consistent behavior at large $\epsilon = \beta/N$

4x4 lattice, $U/\kappa = 8$, $T/\kappa = 0.5$, $\mu = 0$

A. Alexandru, P. Bedaque, A. Carosso, HO, Phys. Rev. D 107, 094502 (2022)
Unbiased Estimator

- One can use less samples
- Error comes from two factors: full-MC and sub-MC

\[ U/\kappa = 8, \ T/\kappa = 0.5, \ \mu = 0 \]
Conclusion

• Summary

1. Discussed the infinite variance problem of fermion on the lattice using its determinant
2. Suggested a way to solve the diverging variance problem:
   - Extra time slice with sub-Monte Carlo method
   - Unbiased estimation of reweighting factor \( \propto 1/F(\phi) \) using sub-MC
3. Showed that the new method has a better behavior at large \( \epsilon \)

• Future study

1. Extra time slice method with Hybrid Monte Carlo
2. Infinite variance problem with the sign problem
Backup
Path Integral Formulation

• BSS formula:

\[
\begin{vmatrix}
1 & 0 & \cdots & B_N \\
B_1 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & B_{N-1} & 1
\end{vmatrix}
= \det \left( 1 - (-1)^N B_N \cdots B_1 \right)
\]

• Conventional action:

\[
B_a(\phi_t)_{x,y} = (Z_a - 1 - \varepsilon_a i \sin \phi_{t,x})\delta_{x,y} - \varepsilon \kappa \delta_{(x,y)}
\]

• Improved action:

\[
B^a(\phi_t)_{x,y} = e^{-\tilde{H}_2^a}e^{-\tilde{H}_4^a(\phi_t)}
\]

with

\[
(\tilde{H}_2^a)_{x,y} = \kappa \varepsilon \delta_{(x,y)} + \varepsilon_a \varepsilon \mu \delta_{x,y}
\]

\[
\tilde{H}_4^a(\phi_t)_{x,y} = - i \varepsilon_a \sin \phi_{t,x} \delta_{x,y}
\]
Determinant & Singular Value
Reweighting

\[-\log R(\phi) = \log \frac{F(\phi)}{\det M_N(\phi)}\]
Diagonal Discretization

\[ B^a(\phi_t)_{x,y} = - \left[ e^{\hbar \delta_{x,y}} \right]_{x,y} e^{i \phi_{x,t}} \]

Diagonal Discretization

\[ \epsilon = 1/32 \]

\[ D \]

\[ \sigma \]

\[ 0 \]

\[ 200,000 \]

\[ 400,000 \]

\[ 600,000 \]

\[ 800,000 \]

\[ 1 \times 10^6 \]