Quantum Computing Quantum Monte Carlo

Jinzhao Sun

Imperial College London → University of Cambridge

Joint with Yukun Zhang & Xiao Yuan (Peking Univ), Yifei Huang & Dingshun Lv (ByteDance)

QTML November 24th, 2023

arXiv:2206.10431

Outline

Background

Given molecular Hamiltonian $H = \sum_i H_i$, find the ground state of H

Talk on Thursday, we discuss how to solve physical problems using quantum embedding
- Challenges: understanding the problem (interpretation, partition ,ansatz design)

Results: Qubit reduction
What about the circuit depth?

The other aspect: given the Hamiltonian, design quantum algorithms

Minimization problem: $\min_{\vec{\theta}} E(\vec{\theta})$ with $E(\vec{\theta}) = \langle \psi(\vec{\theta}) | H | \psi(\vec{\theta}) \rangle$, $|\psi(\vec{\theta})\rangle = U(\vec{\theta}) |\psi_{\mathrm{ref}}\rangle$ with VQE

Background

VQE approaches, in principle, have several drawbacks:

Limited expressivity:

- Due to the presence of various noises, the circuit depths are restricted. Thus, their expressiveness is limited.
- Barren plateau[1]: Once the ansatz is too expressive (i.e., reaches unitary 2-design), the gradients of the parameters vanish exponentially.

Challenge of converge and training issue:

- The training of the VQE algorithm is hard[3] in general.
- A super-polynomial number of local minimums is presented near the global minimum in the VQE loss landscape, avoiding convergence to the global minimum [2].

Intercorrelated

Can we further enhance the performance of NISQ devices (shallow circuits)?

Benefit from advanced classical approaches

1 McClean et al, Nat. Commun. 9, 4812 (2018).

2 Anschuetz and Kiani, Nat. Commun. 13, 7760 (2022).

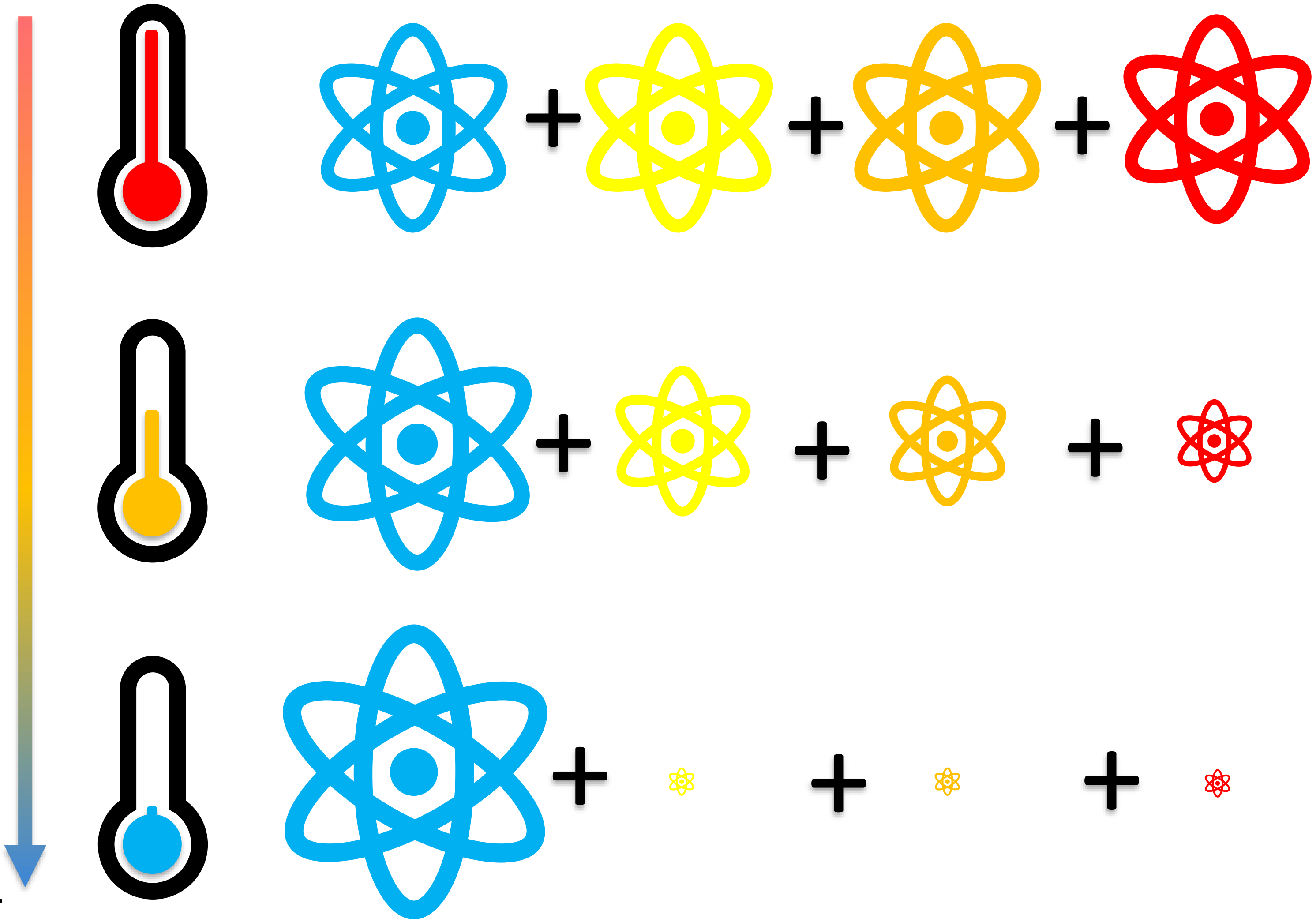
3 Bittel and Kliesch PhysRevLett.127.120502 (2021).

Imaginary time evolution

$$|\psi(0)\rangle = \sum_i \alpha_i |E_i\rangle$$

$$|\psi(\tau)\rangle = \frac{e^{-H\tau}|\psi(0)\rangle}{\sqrt{\langle\psi(0)|e^{-2H\tau}|\psi(0)\rangle}}$$

$$|\psi(\tau \to \infty)\rangle = |E_0\rangle$$



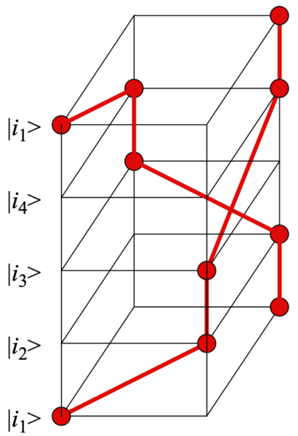
Challenge: non-unitarity.

Review on Quantum Monte Carlo

Key idea: stochastically propagating the *walkers* when implementing $e^{-\Delta\tau H}$

Walkers: states in a chosen (complete orthonormal) basis set.

Key challenge: *sign problem* causes exponential sample complexity of the walkers

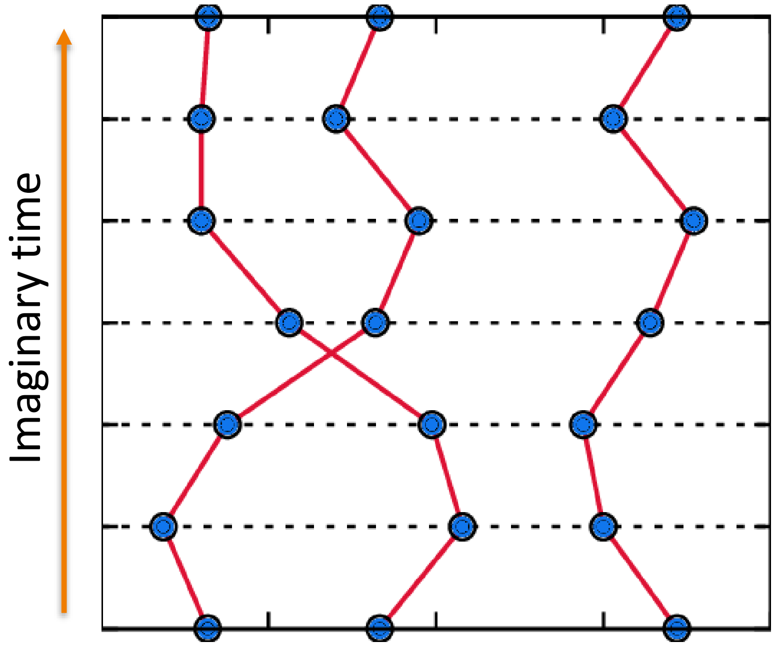


$$\langle O \rangle = \mathrm{Tr}(e^{-\tau H} O)/Z$$

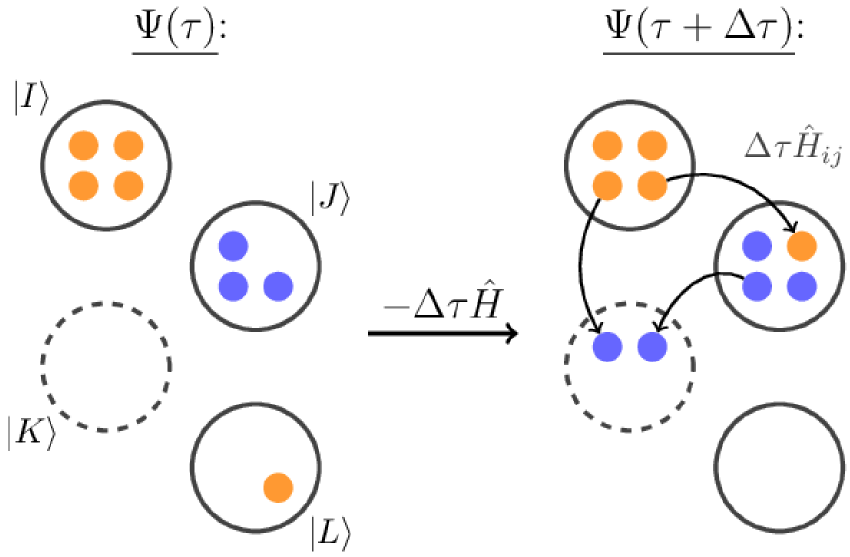
$$Z = \mathrm{Tr}(e^{-\tau H}) = \int d^{Nd}r \, \langle r | e^{-\tau H} | r \rangle$$
$$= \int dr_0 dr_1 \ldots dr_M \langle r_0 | e^{-\delta\tau H} | r_M \rangle \cdots \langle r_2 | e^{-\delta\tau H} | r_1 \rangle \langle r_1 | e^{-\delta\tau H} | r_0 \rangle$$

Insert complete basis

Projector QMC

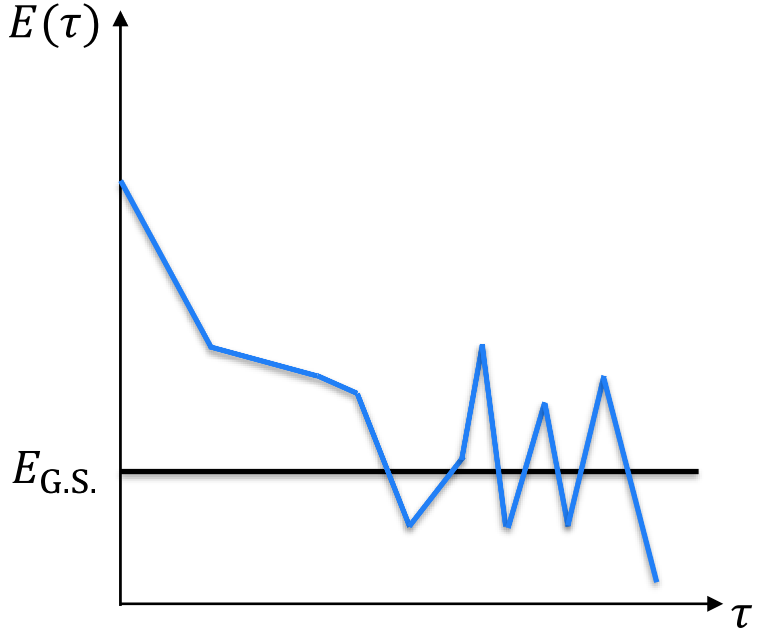


Path integral Monte Carlo (QMC): configurational space propagation.



Full Configuration interaction Quantum Monte Carlo (FCIQMC): Fock space propagation.

$\langle \mathrm{r}_k | e^{-\delta\tau H_j} | r_{k+1} \rangle$ may be negative



The sign problem causes great statistical fluctuation.

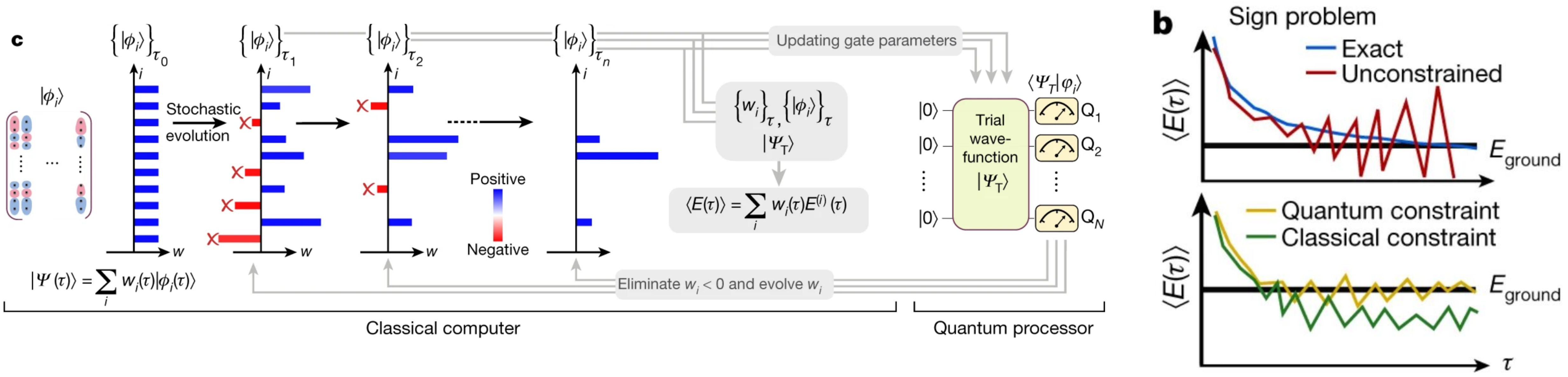
Current progress of QC-QMC

Key idea of Huggins et al.'s work

Introduce quantum ansatz

Original wave function vs. Importance sampled wave function

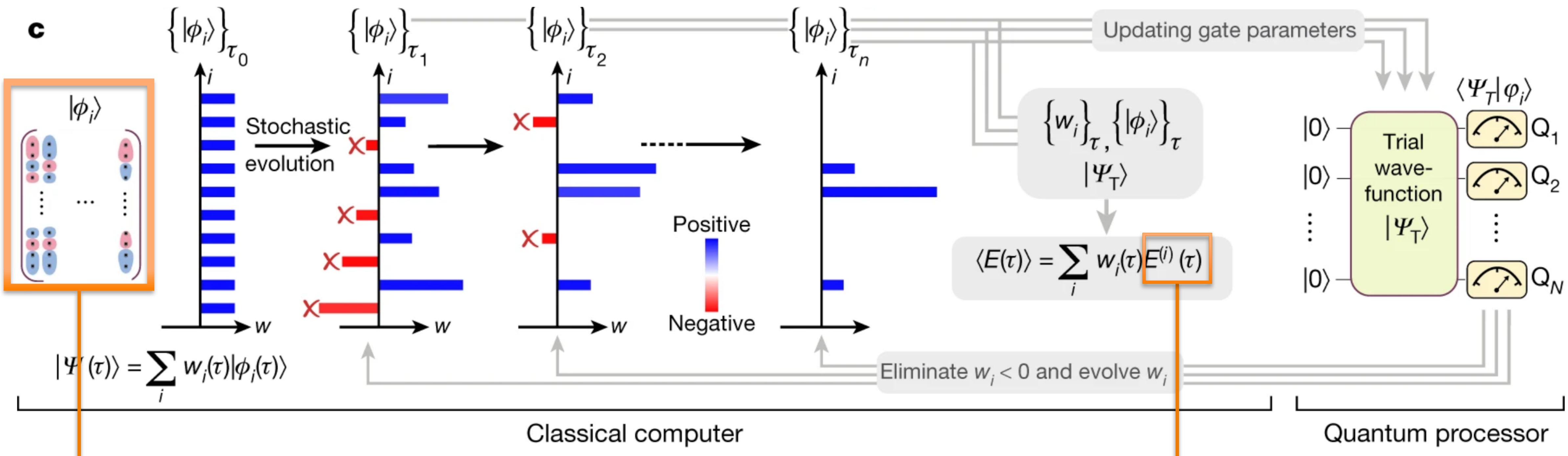
$$|\Psi\rangle(\tau) = \sum_i \omega_i(\tau)\,|\phi_i(\tau)\rangle \qquad |\Psi\rangle(\tau) = \sum_i \omega_i(\tau)\frac{|\phi_i(\tau)\rangle}{\langle\psi_T|\phi_i(\tau)\rangle} \qquad |\Psi\rangle(\tau\to\infty) = |\Psi_0\rangle$$



Huggins et al., Nature 603, 416–420 (2022).

Current progress of QC-QMC

Exponential challenge[1] is encountered even in sign-problem-free cases.



Slater determinant

$$E^{(i)}(\tau) = \frac{\langle \psi_T | H | \phi_i(\tau)\rangle}{\langle \psi_T | \phi_i(\tau)\rangle}.$$

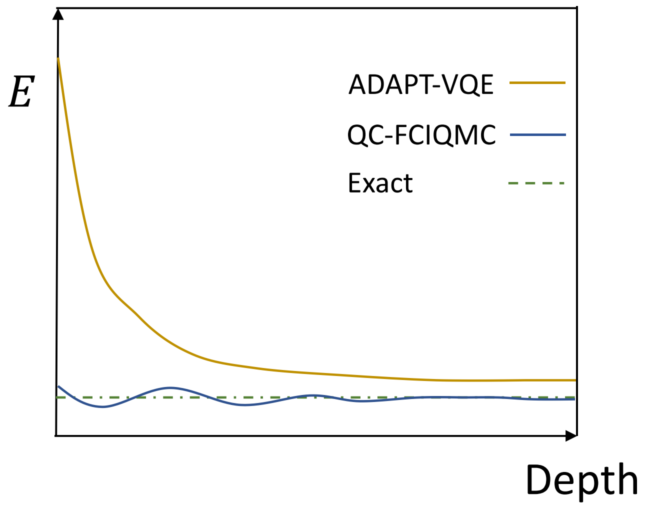
The overlap in the denominator could be exponentially small.
the number of samples for estimating the quantity could grow exponentially.

Mazzola and Carleo, Arxiv 2205.09203 (2022).

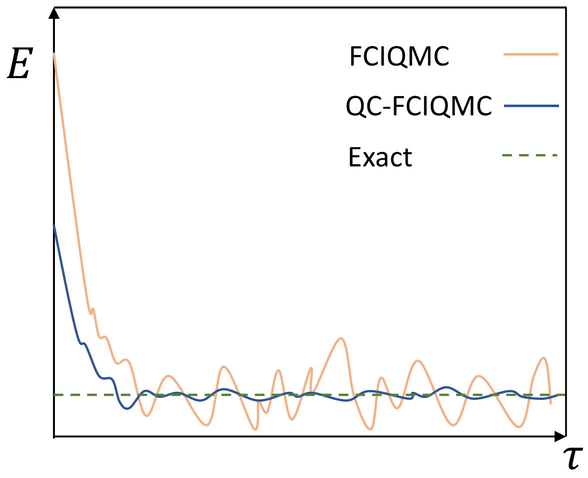
QC-QMC: Overview

Quantum Computing-Full Configuration Interaction QMC (QC-FCIQMC) algorithm.

- Improve both the VQE and classical QMC methods:
  - Provide systematic improvements over the given VQE-prepared state
  - A general way for mitigating the sign problem



Compared to VQE
Increasing depth



Compared to QMC
Small variance

The walkers are entangled states (opposed to QMC)

- Given the VQE-prepared state $|\psi(\vec{\theta})\rangle = U(\vec{\theta})|\psi_{\text{ref}}\rangle$, the basis is chosen to be $\{|\phi_i\rangle = U(\vec{\theta})|i\rangle, \forall i\}$, where each $|i\rangle$ is a classical product state.
- Each state (walker) in the basis is orthonormal since $\langle\phi_i|\phi_j\rangle = \delta_{ij}$.

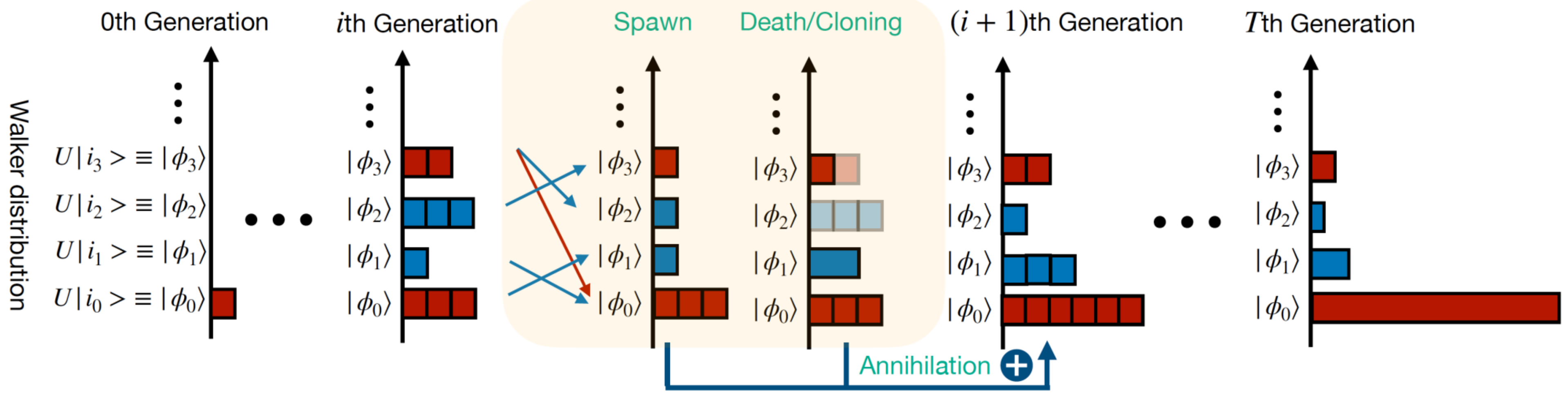
QC-FCIQMC: Implementation

Key idea: effectively realize the diagonal and off-diagonal terms of the ITE operator in a probabilistic way

$$e^{-\Delta\tau(H-E_s)} \approx \mathbb{I} - \Delta\tau(H - E_s)$$

Given $|\phi_0\rangle = U(\vec{\theta})|0\rangle\,|0\rangle \equiv |\psi_{\text{ref}}\rangle$. We may initialize the walker set to be $\{|\phi_0\rangle\}$, and repeat

- **Spawn**: For each walker $|\phi_i\rangle$ in the current walker set → walkers $|\phi_j\rangle$ with probability
  $p(j|i) = |-\Delta\tau\langle\phi_j|H|\phi_i\rangle| / \sum_j |-\Delta\tau\langle\phi_j|H|\phi_i\rangle|, \forall j \neq i$ adjusting the sign of the corresponding walker
- **Death/Cloning**: For each $|\phi_i\rangle$ in the old generation, if $p(i) = \Delta\tau(\langle\phi_i|H|\phi_i\rangle - E_s) > 0$, then kill the walker with probability $p(i)$; Otherwise, clone with $-p(i)$
- **Annihilation**: Sum up the new and old generation of walkers



Red: walkers with +1 sign;
Blue: walkers with −1 sign.

QC-FCIQMC: Implementation

Key techniques: *For walker $|\phi_i\rangle$, sample the $|\phi_j\rangle$s according to $|H_{ji}|^2, \forall j$.*

Denote $\Pi_j = |j\rangle\langle j|$, and let $H = \sum_k h_k P_k$. We have

$$|H_{ji}|^2 = \sum_{kk'} h_k h_{k'} p^i_{kk'}(j), \text{ s.t. } \sum_j p^i_{kk'}(j) \leq 1,$$

$$\text{where } p^i_{kk'}(j) = \mathrm{Re}\langle i|U^\dagger P_k U \Pi_j U^\dagger P_{k'} U|i\rangle.$$

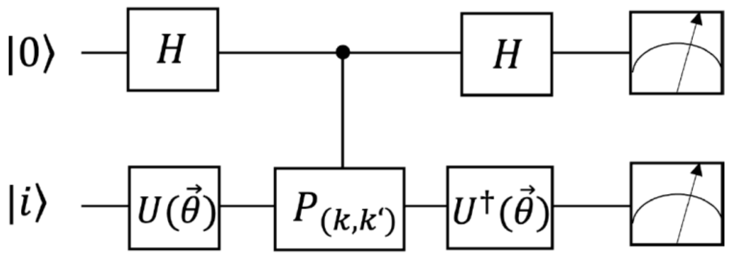


Figure 1: Circuit samples $p^i_{kk'}(j)$

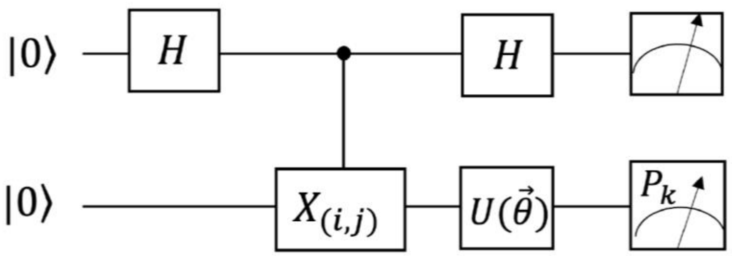


Figure 2: Circuit estimates H_{ji}.

Sign problem

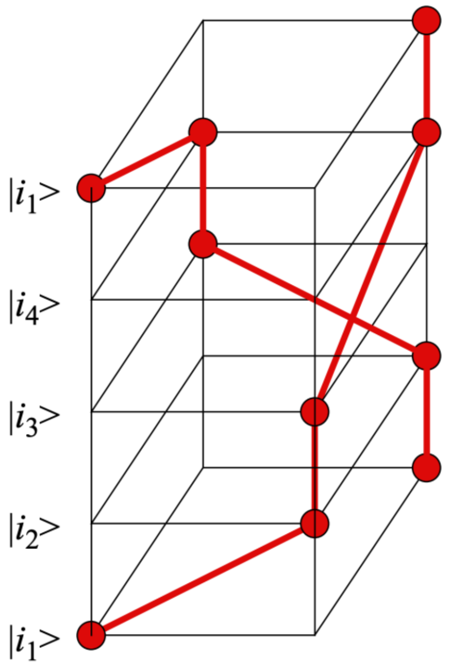
An example:

Consider the partition function of the thermal state in the form of path integrals,

$$Z = \mathrm{Tr}\left(e^{-\beta H}\right) = \sum_{n=0}^{\infty} \frac{(-\beta)^n}{n!} \mathrm{Tr}\left(H^n\right)$$

$$= \sum_{n=0}^{\infty} \sum_{i_1,\cdots,i_n} \frac{(-\beta)^n}{n!} \langle i_1|H|i_2\rangle \langle i_2|H|i_3\rangle \cdots \langle i_n|H|i_1\rangle$$

$$= \sum_{n=0}^{\infty} \sum_{i_1,\cdots,i_n} p\left(i_1,\cdots,i_n\right) = \sum_c p(c)$$



Each path integral forms a closed path

The sign problem:

- Presents whenever $p(c)$ contains negative terms. E.g, fermions, negative weights arise from the Pauli exclusion principle under particle exchange.
- Basis dependent.

Troyer and Wiese, PhysRevLett.94.170201 (2005)

Origin of the sign problem

For arbitrary observable A, the standard way for evaluation is to sample w.r.t. the bosonic system with $|p(c)|$ and $s(c) = \text{sign}\,(p(c))$, we have

$$\langle A\rangle = \frac{\sum_c A(c)s(c)|p(c)|/\sum_c |p(c)|}{\sum_c s\,(c)|p(c)|/\sum_c |p(c)|} = \frac{\langle A\rangle'}{\langle s\rangle'},$$

The average sign decays exponentially with the free energy difference Δf,

$$\langle s\rangle = \frac{z}{z'} = \exp(-\beta N\Delta f)$$

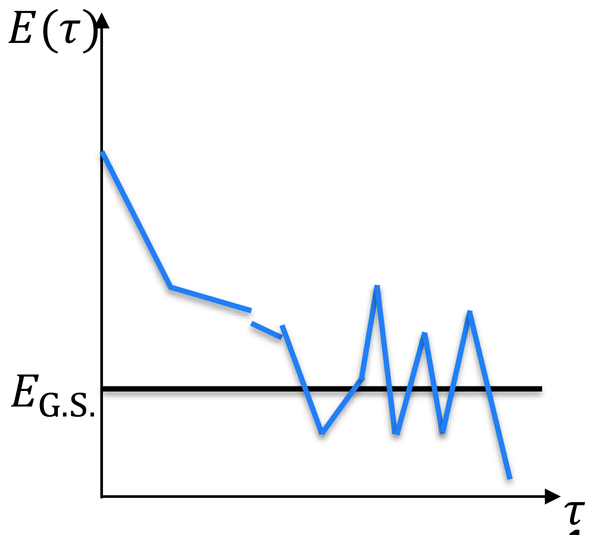
Fermionic and corresponding bosonic system

$$z = \sum_c p(c)\,, z' = \sum_c |p(c)|$$

Thus, the relative error grows exponentially:

$$\frac{\Delta s}{\langle s\rangle} = \frac{\sqrt{\langle s^2\rangle - \langle s\rangle^2/M}}{\langle s\rangle} = \frac{\sqrt{1-\langle s\rangle^2}}{\sqrt{M}\langle s\rangle} \sim \frac{e^{\beta N\Delta f}}{M}$$

M is the number of samples.



Stoquasticity and mitigation of the sign problem

Definition (Stoquastic Hamiltonian[a])

[a] Bravyi et al. 2006.

For any Hamiltonian H, s.t. $H_{ij} \leq 0, \forall i \neq j$. These types of Hamiltonian shall not present sign problem, i.e., all terms in the path integrals are positive.

Proof idea: Denote $G = \alpha I - H$ with $\alpha = \max_i H_{ii}$,

$Z = \mathrm{Tr}(e^{-\beta H}) = e^{-\beta\alpha}\mathrm{Tr}(e^{\beta G})$ $(G_{ij} \geq 0)$

Thus, stoquastic Hamiltonians are sign-problem-free

Mitigation: The sign problem is basis dependent. A universal approach for suppressing the sign problem is by similarity transformation of H, s.t, to approach **stoquastic Hamiltonians**

$$H(U) = U^{\dagger} H U$$

Remark: By expanding the wave function in the VQE-unitary-rotated basis set, *our method effectively implements the similarity transformation with $U = U(\vec{\theta})$ prepared by VQE*

Measure of the sign problem

Relative deviation from the stoquastic Hamiltonians

Definition (Non-Stoquastic Indicator)

Denoting bosonic form of H as $\tilde{H}$:

$$\tilde{H}_{ij} = H_{ij} \text{ if } i = j \text{ or } i \neq j \text{ and } H_{ij} < 0;$$
$$\tilde{H}_{ij} = -H_{ij} \text{ if } i \neq j \text{ and } H_{ij} > 0.$$

The NSI is defined as

$$S(H) = \frac{\mathrm{Tr}[e^{-\beta \tilde{H}}] - \mathrm{Tr}[e^{-\beta H}]}{\mathrm{Tr}[e^{-\beta H}]}.$$

Bosonic form: stoquastic Hamiltonians

Remark: The NSI is computationally non-trivial.

Measure of the sign problem

Theorem (Non-Stoquastic Indicator upper bound)

Define

$$(H_-)_{ij} = H_{ij} \text{ if if } i = j \text{ or } i \neq j \text{ and } H_{ij} < 0; \quad \text{wanted}$$

$$(H_+)_{ij} = H_{ij} \text{ if } i \neq j \text{ and } H_{ij} > 0. \quad \text{Unwanted}$$

The NSI is upper-bounded by

$$S(H) \leq 2e^{\beta\|(\alpha - H_-)\|_{L_1}} \sinh(\beta\|H_+\|_{L_1}).$$

Here, $\|M\|_{L_1} := \sum_{i,j} |M_{ij}|$ *is the matrix norm for matrix* M.

Remark:

- Mitigate the sign problem needs minimize H_+ an H_- simultaneously.
- When H_+ vanishes completely, $S(H) = 0$ as expected for stoquastic Hamiltonians.
- NSI does not sufficiently guarantee the performance

Proof sketch

1. Expand the NSI into path integral form

$$S(H) = \mathrm{Tr}(e^{-\beta\tilde{H}}) - \mathrm{Tr}(e^{-\beta H})$$
$$= \mathrm{Tr}(e^{-\beta(\alpha I - \tilde{G})}) - \mathrm{Tr}(e^{-\beta(\alpha I - G)})$$
$$= e^{-\beta\alpha}\left(\sum_{k=0}^{\infty}\frac{\beta^k}{k!}\sum_{\{|\phi\rangle\}}\left(|\langle\phi_0|G|\phi_1\rangle\langle\phi_1|G|\phi_2\rangle\cdots\langle\phi_{k-1}|G|\phi_0\rangle| - \langle\phi_0|G|\phi_1\rangle\langle\phi_1|G|\phi_2\rangle\cdots\langle\phi_{k-1}|G|\phi_0\rangle\right)\right)$$

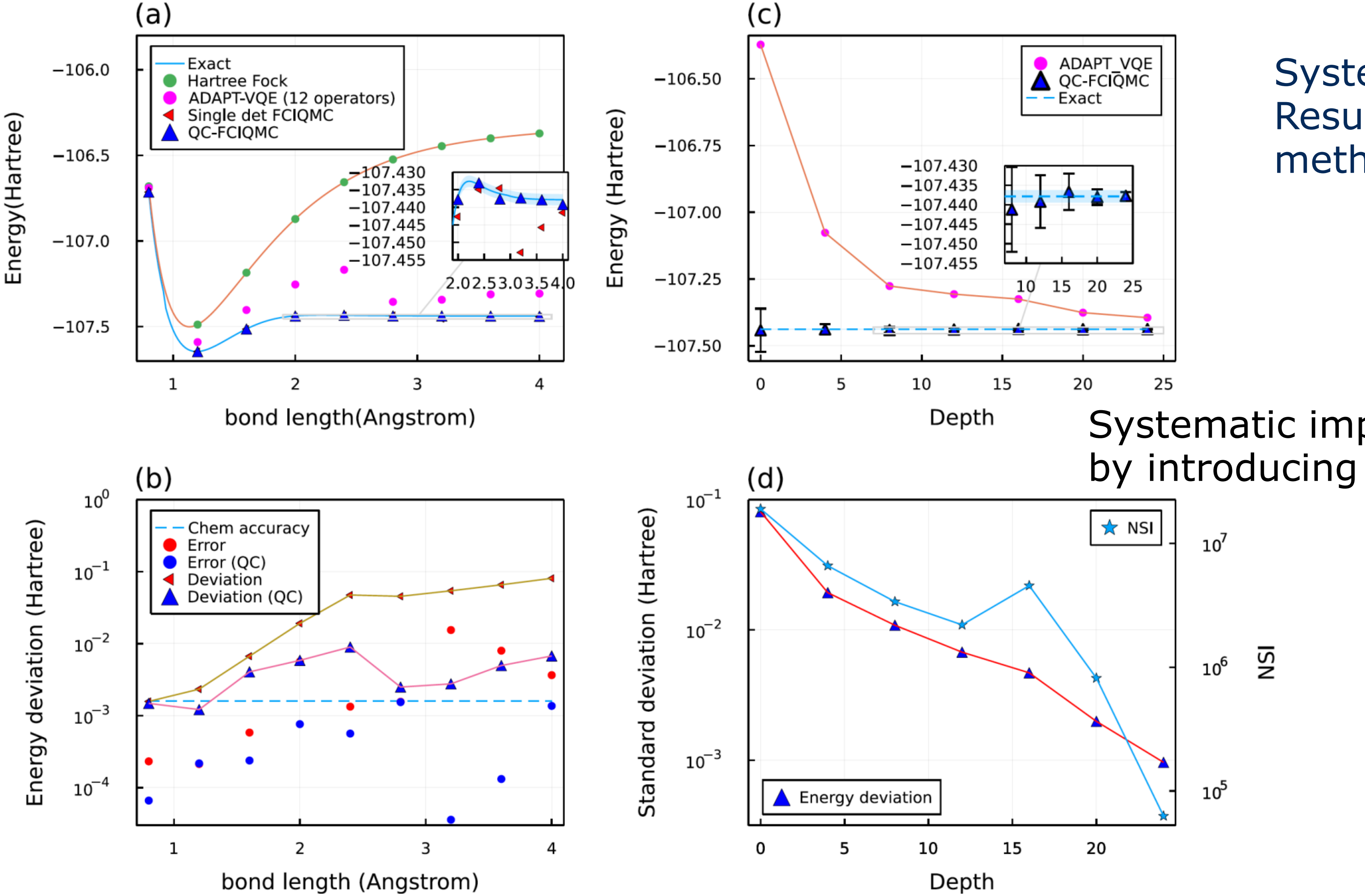
2. Define $G_+ = \alpha - H_+$ and $G_- = \alpha - H_-$. The above subtraction is supposed to be -2-fold of sum of all negative path integrals. all possible negative terms can be enumerated by combination of terms in G_+ and G_-:

$$=2\sum_{k=0}^{\infty}\frac{(\beta)^{2k+1}}{(2k+1)!}\left(\binom{2k+1}{1}\|G_-\|_{L_1}\|G_+^{2k}\|_{L_1} + \binom{2k+1}{3}\|G_-^3\|_{L_1}\|G_+^{2k-2}\|_{L_1} + \cdots + \binom{2k+1}{2k+1}\|G_-^{2k+1}\|_{L_1}\right)$$
$$+2\sum_{k=0}^{\infty}\frac{(\beta)^{2k}}{(2k)!}\left(\binom{2k}{1}\|G_-\|_{L_1}\|G_+^{2k-1}\|_{L_1} + \binom{2k}{3}\|G_-^3\|_{L_1}\|G_+^{2k-3}\|_{L_1} + \cdots + \binom{2k}{2k-1}\|G_-^{2k-1}\|_{L_1}\|G_+\|_{L_1}\right).$$

Group by odd and even terms

1. By property of L1 norm such that $\|A^m\|_{L_1} \le \|A\|_{L_1}^m$, we can relax each term above by $\|G_-^a\|_{L_1}\left\|G_+^b\right\|_{L_1} \le \|G_-\|_{L_1}^a\|G_+\|_{L_1}^b$.
2. Finally, by the binomial theo $\frac{1}{2}((b+a)^n - (b-a)^n) = \sum_{i=1,\mathrm{odd}}^{n}\binom{n}{i}a^x b^{n-x}$, we arrive at our results.

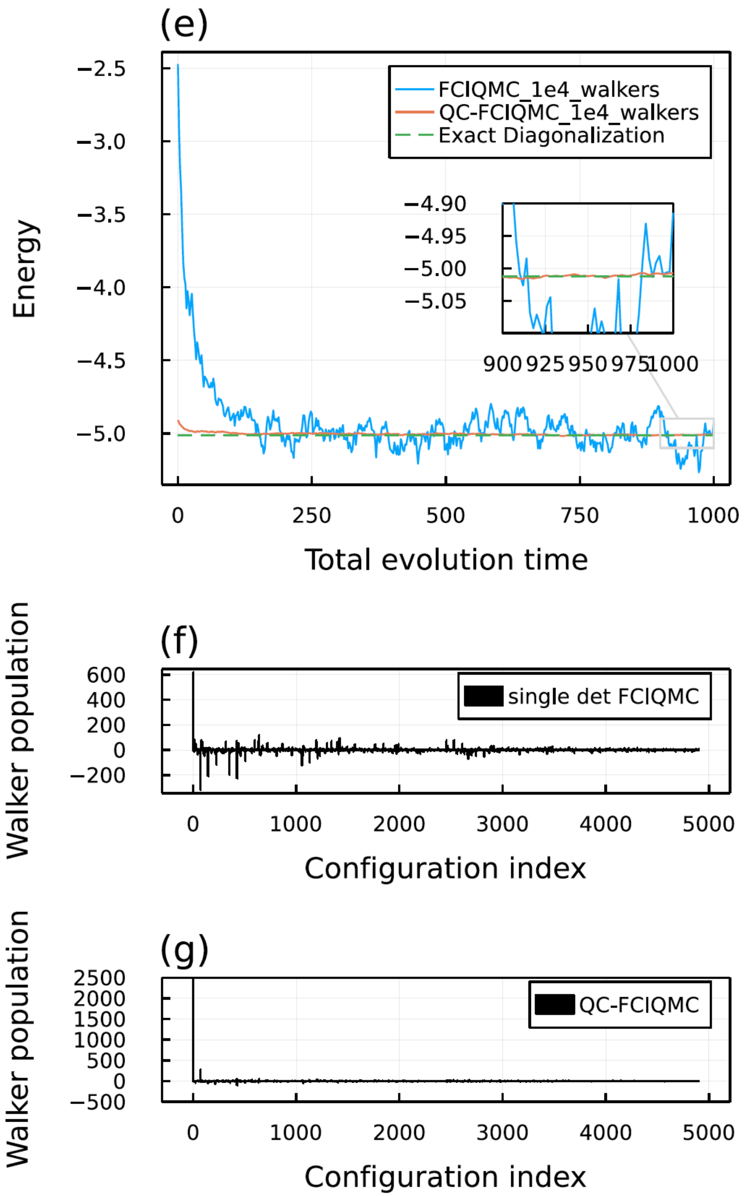
Verification by numerics



System: N2 molecule (12 qubits). Results vs shallow depth VQE methods.

Systematic improvement of the sign problem by introducing entangled bases

Verification by numerics



Results vs classical FCIQMC algorithms.

Distribution of the walkers become more concentrated to the initial walker $|\phi_0\rangle$ as it get closer to the ground state

- fewer walkers

Summary & outlook

Summary

- Introduce the QC-FCIQMC algorithm that improves the ability of shallow-depth quantum circuits and each state (walker) could be prepared by shallow depth circuit available by NISQ devices.
- Systematic suppression of the sign problem is achieved if the basis (set of walkers) is refined.
- NSI to characterize severity of the sign problem (not sufficient for performance guarantee)
- Limitation: our method demands a huge number of measurements for sampling the walkers that could be challenging for NISQ devices.

Possible future works

- The upper bound on NSI can serve as a low-cost loss function for easing the sign problem for classical QMC
- Explore other kinds of unitary construction for mitigating the sign problem, one with performance guarantee
- Performance analysis under certain noise channels
- Compatibility with classical shadows to reduce measurement costs