



Error-mitigated fermionic classical shadows on noisy quantum devices

Bujiao Wu^{1,2,3} & Dax Enshan Koh⁴

Efficiently estimating fermionic Hamiltonian expectation values is vital for simulating various physical systems. Classical shadow (CS) algorithms offer a solution by reducing the number of quantum state copies needed, but noise in quantum devices poses challenges. We propose an error-mitigated CS algorithm assuming gate-independent, time-stationary, and Markovian (GTM) noise. For n -qubit systems, our algorithm, which employs the easily prepared initial state $|0^n\rangle$ assumed to be noiseless, efficiently estimates k -RDMs with $\tilde{O}(kn^k)$ state copies and $\tilde{O}(\sqrt{n})$ calibration measurements for GTM noise with constant fidelities. We show that our algorithm is robust against noise types like depolarizing, damping, and X -rotation noise with constant strengths, showing scalings akin to prior CS algorithms for fermions but with better noise resilience. Numerical simulations confirm our algorithm's efficacy in noisy settings, suggesting its viability for near-term quantum devices.

Assessing the properties of interacting fermionic systems constitutes one of the core tasks of modern physics, a task that has a wealth of applications in quantum chemistry¹, condensed matter physics², and materials science³. Notions of *quantum simulation* offer an alternative route to studying this important class of systems. In *analog simulation*, one prepares the system of interest under highly controlled conditions. However, any such effort makes sense only if one has sufficiently powerful *readout techniques* available that allow one to estimate properties. In fact, the read-out step constitutes a core bottleneck in many schemes for quantum simulation.

Fortunately, for natural fermionic systems, one often does not need to learn the full unknown quantum state; trying to do so regardless would be highly impractical, as the resources required for a full tomographic recovery would scale exponentially with the size of the system. Instead, what is commonly needed are the so-called k -particle reduced density matrices, abbreviated as k -RDMs. These are expectation values of polynomials of fermionic operators of the $2k$ -th degree. Naturally, the expectation value of any interaction fermionic Hamiltonian can be estimated using 2-RDMs only^{4,5}. Indeed, the adaptive *variational quantum algorithm* (VQE)⁶ also utilizes up to 4-RDMs to simulate many-body interactions in the ground and excited state^{7,8}. That is to say, meaningful methods of read-out often focus on estimating such fermionic reduced density matrices.

On the highest level, several approaches can be pursued when dealing with fermionic operators. One of those—and the one followed here—is to treat the fermionic system basically as a collection of spins. Then given spin

Hamiltonians $\{H_i\}_{i=1}^m$ and an unknown quantum state ρ , where $m = \mathcal{O}(\text{poly}(n))$, the *classical shadow* (CS) algorithm or its variants^{9–21} in qubit systems are among the most promising ways to calculate the expectations $\text{Tr}(\rho H_i)$, with the representation of the Hamiltonians $\{H_i\}_{i=1}^m$ in the Pauli basis, which invokes a *fermion-to-spin mapping* such as the *Jordan-Wigner*^{22,23} or *Bravyi-Kitaev encodings*^{24,25}. We define the classical shadow channel as \mathcal{M} , which involves operating the unitary channel \mathcal{U} uniformly randomly sampled from the Clifford group before measurements in the Z -basis and classical postprocessing operations on the measurement outcomes. By performing the inverse of the classical shadow channel \mathcal{M}^{-1} on the resulting state after performing \mathcal{M} on the initial state ρ , we obtain the classical shadow representation $\hat{\rho}$ of the quantum state ρ , allowing for the calculation of the expected values of observables $\{H_i\}_{i=1}^m$ with respect to ρ using classical methods.

While the classical shadow algorithm requires exponentially many copies even for some local interacting fermions due to the inefficient representation in the qubit system, recently, several classical shadow algorithms for fermionic systems without encoding of the Hamiltonians have been proposed^{26–28}. Zhao, Rubin, and Miyake²⁶ utilize the generalized CS method⁹ for fermionic systems, and proposed an algorithm that requires $\mathcal{O}\left(\binom{n}{k} k^{3/2} (\log n) / \epsilon^2\right)$ copies for the unknown quantum states to output all the elements of a k -RDM. Low²⁷ proves that all elements of the k -RDM can

¹Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany. ²Shenzhen Institute for Quantum Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China. ³Center on Frontiers of Computing Studies, Peking University, Beijing 100871, China.

⁴Institute of High Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), 1 Fusionopolis Way, #16-16 Connexis, Singapore 138632, Republic of Singapore. ✉e-mail: bujiaowu@gmail.com; dax_koh@ihpc.a-star.edu.sg

be estimated with $\binom{\eta}{k} (1 - \frac{\eta-k}{n})^k \frac{1+\eta}{1+\eta-k} / \epsilon^2$ number of copies of the quantum state, where η is the number of particles and n is the number of modes. These fermionic shadow estimation methods (along with the generic classical shadow formalism) do not account for noise in the system, which is an inevitability in real physical systems.

Since we are still in the noisy intermediate-scale quantum (NISQ) era, current quantum simulators are heavily affected by noise; hence, any characterization technique needs to be robust for these simulators to be useful. For qubit systems, robust shadow estimation was developed^{29,30} where Chen et al.²⁹ use techniques from randomized benchmarking to mitigate the effect of gate-independent time-stationary Markovian (GTM) noise channels on the procedure. Inane et al.³¹ proposed error-mitigated classical shadow with probabilistic error cancellation.

Utilizing the robust shadow estimation scheme and taking inspiration from the fermionic shadow estimation of Zhao et al.^{26,28}, we present an error-mitigated shadow estimation scheme for fermionic systems and demonstrate its feasibility for realistic noise channels. Note that akin to the fermionic CS approaches proposed in refs. 26,28, our error-mitigated CS method circumvents the need to encode the Hamiltonian using the qubit representation.

We sample our unitaries U_Q from the matchgate group³², a natural choice for our protocol as there is a one-to-one correspondence between two-qubit matchgates and free-fermionic evolution^{33,34}. We therefore design the classical postprocessing operations by leveraging the irreducible representation of the matchgate group. We successfully introduce an unbiased estimator \widehat{M} for the noisy classical shadow channel \mathcal{M} , where we require an additional calibration protocol to generate the estimator \widehat{M} with the assumption that the computational basis state $|0\rangle$ ($\mathbf{0}$) can be prepared noiselessly. Additionally, we demonstrate the efficacy of our protocol under conditions of constant noise strength by evaluating its performance across various common noise channels: depolarizing noise, generalized amplitude damping, X-rotation, and Gaussian unitary noise. The number of samples required for the estimation process of our protocol is in the same order as the noise-free matchgate classical shadow scheme^{26,28}.

We determine the effectiveness of our protocol with the above noise models by calculating the expectations of all elements of the k -particle reduced density matrix (k -RDM) when the noise strength is constant. The number of samples required for estimation, in this case, is $\mathcal{O}(kn^k \ln(n/\delta_c)/\epsilon_c^2)$ and for calibration is $\mathcal{O}(\sqrt{n} \ln n \ln(1/\delta_c)/\epsilon_c^2)$ with error $\epsilon_c + \epsilon_c$ and success probability $(1 - \delta_c)(1 - \delta_c)$.

We have extended the analysis of our error-mitigated fermionic shadow channel estimation to more general physical quantities inspired by the fermionic shadow analysis of Wan et al.²⁸, with more details in Supplementary Notes 5, 9. We list distinct classical shadow approaches in both noiseless and noisy qubit and fermionic systems in Table 1. Our error-mitigated fermionic classical shadow technique constitutes an extension of the work by Chen et al.²⁹, accommodating scenarios where the gate-set lacks (1) 3-design properties³⁵ and (2) the applicability of the randomized benchmarking scheme developed by Helsen et al.³⁶.

We tested the accuracy and efficacy of our protocol by performing numerical experiments to estimate $\text{Tr}(\rho \widetilde{\gamma}_S)$ (where $\widetilde{\gamma}_S = U_Q^\dagger \gamma_S U_Q$, where γ_S is the product of $|S|$ Majorana operators and plays a crucial role in computing k -RDMs) on a noisy quantum device subjected to various types of gate noise such as depolarizing, generalized amplitude damping, X-rotation, and Gaussian unitaries. Our numerical investigations confirm the potential of our methods in real-world experimental scenarios.

Results

Basic notations and concepts

Here we give the basic notations and concepts that will be used throughout this work.

Table 1 | Enumeration of the classical shadow protocols in noiseless and noisy settings, for qubit and fermionic systems respectively

	Clifford-based shadows	Fermionic shadows
Noiseless	Huang et al. ⁹	Zhao et al. ²⁶ ; Low ²⁷ ; Wan et al. ²⁸
Noisy	Chen et al. ²⁹ ; Koh and Grewal ³⁰	This work

Basic notations. The symbols X , Y , and Z denote the Pauli X , Y , and Z operators respectively. The operator $R_X(\theta) = \exp(-i\frac{\theta}{2}X)$ denotes the rotation operator around the x -axis. A Z -basis measurement is performed with respect to the basis of eigenstates of the Pauli- Z operator. We utilize \mathbb{I} to represent the identity operator on the full system. The set of linear operators on a vector space \mathcal{H} is denoted as $\mathcal{L}(\mathcal{H})$. We utilize the symbol $\widetilde{\mathcal{O}}$ to omit the logarithmic terms.

Superoperator. We denote the superoperator representation of a linear operator $O \in \mathcal{L}(\mathcal{H})$ as $|O\rangle\rangle := O/\sqrt{\text{Tr}(OO^\dagger)}$ and the scaled Hilbert-Schmidt inner product between linear operators as $\langle\langle O|R\rangle\rangle = \text{Tr}(O^\dagger R)/\sqrt{\text{Tr}(OO^\dagger)\text{Tr}(RR^\dagger)}$. The action of a channel $\mathcal{E} \in \mathcal{L}(\mathcal{L}(\mathcal{H}))$ operating on a linear operator $O \in \mathcal{L}(\mathcal{H})$ can hence be written as $\mathcal{E}|O\rangle\rangle = \mathcal{E}(O)/\sqrt{\text{Tr}(OO^\dagger)}$. The channel representation of a measurement with respect to the computational basis can be represented as $\mathcal{X} = \sum_{x \in \{0,1\}^n} |x\rangle\rangle \langle\langle x|$. We denote the unitary channel corresponding to the unitary operator U as $\mathcal{U}(\cdot) = U(\cdot)U^\dagger$.

Majorana operator. The Majorana operators γ_j for $1 \leq j \leq 2n$ describes the fermionic system with $\gamma_j = b_{(j+1)/2} + b_{(j+1)/2}^\dagger$ for odd j and $\gamma_j = -i(b_{j/2} - b_{j/2}^\dagger)$ for even j , where b_j and b_j^\dagger are the annihilation and creation operators, respectively, associated with the j -th mode. Let γ_S be the product of the Majorana operators indexed by the subset S , denoted as $\gamma_S = \gamma_{l_1} \cdots \gamma_{l_{|S|}}$ for $|S| > 0$ and $\gamma_\emptyset = \mathbb{I}$, where $S = \{l_1, \dots, l_{|S|}\}$ and $l_1 < l_2 < \dots < l_{|S|}$. It can be shown that γ_S forms the complete orthogonal basis for $\mathcal{L}(\mathcal{H})$ for $S \subseteq [2n]$. Let $\Gamma_k := \{\gamma_S | |S| = k\}$ be the subspace of γ_S with cardinality k . We denote the even subspace as $\Gamma_{\text{even}} := \bigoplus_l \Gamma_{2l}$. Also, we denote \mathcal{P}_k as the projector onto the subspace Γ_k , i.e.

$$\mathcal{P}_k = \sum_{S \in \binom{[2n]}{k}} |\gamma_S\rangle\rangle \langle\langle \gamma_S|, \tag{1}$$

where we have used the notation that for a set A and an integer k , $\binom{A}{k} = \{T \subseteq A : |T| = k\}$ denotes the set of subsets of A with cardinality k .

Gaussian unitaries. Matchgates are in a one-to-one correspondence with the fermionic Gaussian unitaries and can serve as a qubit representation for these unitaries. We denote \mathbb{M}_n as the matchgate group, and write its elements $U_Q \in \mathbb{M}_n$ in terms of rotation matrices Q belonging to the orthogonal group $\text{Orth}(2n)$ (see Supplementary Note 1 for details)^{33,37}. Following Wan et al.'s study²⁸, which demonstrated that the continuous matchgate group \mathbb{M}_n and the discrete subgroup $\mathbb{M}_n \cap \text{Cl}_n$ (where Cl_n represents the Clifford group) deliver equivalent performances for fermionic classical shadows, our findings remain applicable to both continuous and discrete matchgate circuits. Since $U_Q^\dagger \gamma_j U_Q = \sum_k Q_{jk} \gamma_k$, the matchgate U_Q transforms the product of Majorana operators γ_S in the $\Gamma_{|S|}$ subspace as $U_Q^\dagger \gamma_S U_Q = \sum_{S' \in \binom{[2n]}{|S|}} \det(Q_{|S|S'}) \gamma_{S'}$.

k-particle reduced density matrices (k-RDM). We denote a k -RDM as ${}^k\mathcal{D}$, which can be obtained by tracing out all but k particles. Here we

denote it as a tensor with $2k$ indices,

$${}^k\mathbf{D}_{j_1, \dots, j_k; l_1, \dots, l_k} = \text{Tr}(\rho b_{j_1}^\dagger \cdots b_{j_k}^\dagger b_{l_1} \cdots b_{l_k}), \quad (2)$$

where j_i and l_i are in $[n]$ for $i \in [k]$. The fermionic system can be equivalently described in the Majorana basis, in which case a tensor can be rewritten as the linear combinations of $\text{Tr}(\rho \gamma_S)$, and $|S| \leq 2k$. Hence all n^{2k} elements of the k -RDM can be obtained by calculating $\text{Tr}(\rho \gamma_S)$, for the scaling of $\mathcal{O}(n^k)$ different S with $|S| \leq 2k$ ³⁸.

Pfaffian function. The Pfaffian of a matrix $Q \in \mathbb{R}^{2n \times 2n}$ is defined as

$$\text{pf}(Q) = \frac{2^n}{n!} \sum_{\sigma \in S_{2n}} \text{sgn}(\sigma) \prod_{i=1}^n Q_{\sigma_{2i-1}, \sigma_{2i}}, \quad (3)$$

which can be calculated in $\mathcal{O}(n^3)$ time³⁹.

Ideal fermionic shadow (Wan et al.²⁸). Given an unknown quantum state ρ , the classical shadow method applies a unitary U_Q uniformly randomly sampled from matchgate group \mathbb{M}_n , followed by measuring the generated state in the computational basis. With the measurement result $|x\rangle$, we can generate the classical representation $\hat{\rho} = \mathcal{M}^{-1} \mathcal{U}_Q^\dagger |x\rangle$ for the unknown quantum state ρ , where the channel \mathcal{M} describing the overall process is defined as

$$\begin{aligned} \mathcal{M}(\rho) &= \int_Q d\mu(Q) \left[\sum_{x \in \{0,1\}^n} \langle x | U_Q \rho U_Q^\dagger | x \rangle U_Q^\dagger | x \rangle \langle x | U_Q \right] \\ &= \sum_k \binom{n}{k} \binom{2n}{2k} \mathcal{P}_{2k}(\rho). \end{aligned} \quad (4)$$

Noise assumptions. In this work, we assume that the noise is gate-independent, time-independent, and Markovian (a common assumption in randomized benchmarking (RB) abbreviated as the *GTM noise assumption*⁴⁰) and that the preparation noise for the easily prepared state $|\mathbf{0}\rangle \langle \mathbf{0}|$ is negligible. For the convenience of calculation, we utilize the left-hand side noisy representation for a noisy fermionic unitary $\tilde{U}_Q := \Lambda U_Q$. Here we define the average fidelity in Γ_{2k} subspace for noise channel Λ as

$$\mathcal{B}_k := \frac{(-i)^k}{2^n \binom{n}{k}} \sum_x \sum_{S \in \binom{[n]}{k}} (-1)^{x_S} \text{Tr}(|x\rangle \langle x| \Lambda(\gamma_{D(S)})), \quad (5)$$

where $D(S) = \{2j - 1, 2j | j \in S\}$, $0 \leq k \leq n$, $x_S = \sum_{j \in S} x_j$. It is easy to check that $\mathcal{B}_k = 1$ if $k = 0$. With some calculations, we have $\mathcal{B}_k = 1$ for the noise-free model where noise channel Λ equals the identity. In the following, we give the analysis of the simplified result for \mathcal{B}_k for several common noise models in the qubit system and fermionic system when $k > 1$. See more details for the analysis in Supplementary Note 3.

(1) The depolarizing noise with channel representation $\Lambda_d(A) = (1 - p)A + p \text{Tr}(A) \frac{1}{2^n}$ for any n -qubit linear operator A , where the noise strength $p \in [0, 1]$, and $\mathcal{B}_k = 1 - p$.

(2) The generalized amplitude-damping noise with the Kraus representation

$$\Lambda_a(\cdot) = \sum_{u, v \in \{0, 1\}^n} E_{uv}(\cdot) E_{uv}^\dagger + E_0(\cdot) E_0^\dagger, \quad (6)$$

Table 2 | Comparison of average noise fidelity F_{ave} , Z-basis average noise fidelity F_Z , and average noise fidelity in Γ_{2k} subspace \mathcal{B}_k

Noise type	Λ_d	Λ_a	Λ_r
Λ_{avg}	$1 - p/2$	$2/3 - (\rho_0 + \rho_1)/6 + \sqrt{(1 - \rho_0)(1 - \rho_1)}/3$	$(\cos(\theta) + 2)/3$
F_Z	$1 - p/2$	$1 - (\rho_0 + \rho_1)/2$	$\cos^2(\theta/2)$
\mathcal{B}_1	$1 - p$	$1 - (\rho_0 + \rho_1)$	$\cos \theta$

where $E_{uv} = \sqrt{\bar{p}_u} |v\rangle \langle u|$ for $u \neq v \in \{0, 1\}^n$ and $E_0 = \sqrt{1 - \sum_{u, v \in \{0, 1\}^n, u \neq v} E_{uv}^\dagger E_{uv}}$,

where the probabilities \bar{p}_u satisfy $(2^n - 1)\bar{p}_u \leq 1$ for any $u \in \{0, 1\}^n$. The average fidelity $\mathcal{B}_k = 1 - \sum_{u \in \{0, 1\}^n} \bar{p}_u$ if $k \neq 0$. We let $\sum_{u \in \{0, 1\}^n} \bar{p}_u$ denote the noise strength.

(3) The X -rotation noise with the channel representation

$$\Lambda_r(\cdot) = R_X(\theta)(\cdot)R_X(-\theta) \quad (7)$$

where $R_X(\theta) = \exp(-i \sum_{l=1}^n \theta_l X_l / 2)$, where the noise strengths θ_l are some real numbers. By some calculations, we have

$$\mathcal{B}_k = \binom{n}{k}^{-1} \sum_{S \in \binom{[n]}{k}} \prod_{l \in S} \cos \theta_l. \text{ Hence } \min_l \cos^k \theta_l \leq \mathcal{B}_k \leq \max_l \cos^k \theta_l.$$

(4) Noise that is a Gaussian unitary^{41,42}, where we assume that the noise has no coherence with the environment. This noise channel is denoted as

$$\Lambda_g(\cdot) = U_Q(\cdot)U_Q^\dagger, \quad (8)$$

where U_Q is a Gaussian unitary. By selecting the noise model to be Λ_g , we get

$$\mathcal{B}_k = \binom{n}{k}^{-1} \sum_{S, S' \in \binom{[n]}{k}} \det(Q|_{D(S), D(S')}).$$

Note that for the noise models defined in (1–3), the average fidelity $\mathcal{B}_k \in [0, 1]$ is close to one when the noise strengths are close to zero.

For comparison, it is worth noting that the standard average noise fidelity⁴³ $F_{\text{avg}} = \int_\psi d\psi \langle \psi | \Lambda(|\psi\rangle \langle \psi|) | \psi \rangle$ where ψ is drawn from the Haar measure, and the Z-basis average noise fidelity defined in Chen et al.²⁹, $F_Z = \frac{1}{2^n} \sum_b \langle \langle b | \Lambda | b \rangle \rangle$ are not equivalent to \mathcal{B}_k under the same noise model. We present a comparison of these three quantities for $\Lambda_d, \Lambda_a, \Lambda_r$ for a single qubit, as depicted in Table 2. They are closely aligned, with \mathcal{B}_1 slightly smaller than F_{avg} and F_Z . We give a more detailed analysis in Supplementary material X.

Mitigation algorithm and error analysis

Let $\widehat{\mathcal{M}} := \sum_{k=0}^n \hat{f}_{2k} \mathcal{P}_{2k}$ be the estimator for the noisy shadow channel $\widetilde{\mathcal{M}} = \sum_{k=0}^n f_{2k} \mathcal{P}_{2k}$. In the Methods Section, we provide an explicit expression and efficiency for the calculation of \hat{f}_{2k} . Using the estimated noisy channel $\widehat{\mathcal{M}}$, we can now obtain an estimate for $\{\text{Tr}(\rho H_j)\}_{j=1}^m$, where ρ represents certain quantum states and H_j denotes certain observables. If $f_{2k} = 0$, the channel $\widetilde{\mathcal{M}}$ becomes non-invertible. Consequently, the effectiveness of the fermionic CS method diminishes, and we cannot retrieve $\text{tr}(\rho H)$ using it. In this study, we operate under the assumption that the noise is permissible and the fermionic CS channel is consistently invertible. Additionally, we provide a scenario in Supplementary Note 3 where the extreme noise channel occurs, i.e., $f_k = 0$. However, it is anticipated that such an extreme case will rarely occur. Algorithm 1 demonstrates the method for mitigated estimation.

Algorithm 1. Error-mitigated estimation for noisy fermionic classical shadows

- 1: **Input** Quantum state ρ , observables H_1, \dots, H_m , integers N_c, K_c, N_e, K_e .
- 2: **Output** \hat{v}_i for $i \in [m]$.
- 3: $R_c := N_c K_c$;
- 4: **for** $j \leftarrow 1$ to R_c **do**
- 5: Prepare state $\rho_0 = |0^n\rangle$, uniformly sample $Q \in \text{Orth}(2n)$ or $\text{Perm}(2n)$, implement the associated noisy Gaussian unitary \hat{U}_Q on ρ_0 , and measure in the Z_{-1} -basis with outcomes x ;
- 6: Let $\hat{f}_{2k}^{(j)} := 2^n (-1)^k \binom{n}{k} \langle \mathbf{0} | \mathcal{P}_{2k} \mathcal{U}_Q^\dagger | x \rangle, \forall k \in [n]$;
- 7: **end for**
- 8: $\hat{f}_{2k} := \text{MedianOfMeans} \left(\left\{ \hat{f}_{2k}^{(j)} \right\}_{j=1}^{R_c}, N_c, K_c \right) \forall k \in [n]$;
- 9: $\hat{\mathcal{M}} := \sum_{k=1}^n \hat{f}_{2k} \mathcal{P}_{2k}$;
- 10: $R_e := N_e K_e$;
- 11: **for** $i \leftarrow 1$ to m **do**
- 12: **for** $j \leftarrow 1$ to R_e **do**
- 13: Prepare ρ , uniformly sample $Q \in O(2n)$ or $\text{Perm}(2n)$, implement the associated noisy Gaussian unitary \hat{U}_Q on ρ , and measure in the Z -basis with outcomes x ;
- 14: Generate estimation $\hat{v}_i^{(j)} := \text{Tr} \left(H_i \hat{\mathcal{M}}^{-1} \mathcal{U}_Q^\dagger(x) \right)$;
- 15: **end for**
- 16: $\hat{v}_i := \text{MedianOfMeans} \left(\left\{ \hat{v}_i^{(j)} \right\}_{j=1}^{R_e}, N_e, K_e \right)$;
- 17: **end for**
- 18: **return** $\{\hat{v}_i\}_{i=1}^m$

Incorporating the **MedianOfMeans** sub-procedure, as explained in Ref. 9, guarantees that the number of quantum state copies needed relies on the logarithm of the number of observables. We included the **MedianOfMeans** sub-procedure in Supplementary Note 6 to ensure the completeness and consistency of this paper. Our error analysis will involve selecting appropriate values for the number of calibrations and estimation samplings N_c, N_e, K_c , and K_e to estimate the coefficients \hat{f}_{2k} associated with the noisy channel.

Let $\hat{v} := \text{Tr}(\hat{\mathcal{M}}^{-1} \mathcal{U}^\dagger(|x\rangle\langle x|)H)$ be an estimation of $\text{Tr}(\rho H)$ for some observables H in the even subspace Γ_{even} and quantum states ρ , where x follows the distribution $\text{Tr}(|x\rangle\langle x| \hat{\mathcal{U}}(\rho))$ for $x \in \{0, 1\}^n$, then we have

$$\begin{aligned}
 |\hat{v} - \text{Tr}(\rho H)| &\leq |\hat{v} - \text{Tr}(\hat{\mathcal{M}}^{-1} \hat{\mathcal{M}}(\rho)H)| \\
 &\quad + |\text{Tr}(\hat{\mathcal{M}}^{-1} \hat{\mathcal{M}}(\rho)H) - \text{Tr}(\rho H)| \\
 &= \varepsilon_e + \varepsilon_c,
 \end{aligned}
 \tag{9}$$

where $\varepsilon_e := |\hat{v} - \text{Tr}(\hat{\mathcal{M}}^{-1} \hat{\mathcal{M}}(\rho))|$ is the estimation error and $\varepsilon_c := |\text{Tr}(\hat{\mathcal{M}}^{-1} \hat{\mathcal{M}}(\rho)H) - \text{Tr}(\rho H)|$ is the calibration error. Therefore, by determining the necessary number of samples N_e and K_e to achieve the desired level of estimation error ε_e (as well as N_c and K_c to account for the calibration error ε_c), we can obtain an estimation \hat{v} with an overall error of $\varepsilon = \varepsilon_e + \varepsilon_c$, using $N_e K_e$ copies of the input state ρ .

Theorem 1. Let ρ be an unknown quantum state and $\{H_i\}_{i=1}^m$ be a set of observables in the even subspace Γ_{even} . Consider Algorithm 1 with the number of estimation samplings

$$\begin{aligned}
 R_e &= \frac{68(1+\varepsilon_e)^2 \ln(2m/\delta_e)}{2^{2n} \varepsilon_e^2} \sum_{0 \leq l_1+l_2+l_3 \leq n} g(l_1, l_2, l_3) \\
 &\sum_{\substack{S_1, S_2, S_3 \text{ disjoint} \\ |S_i|=2l_i, \forall i \in [3]}} \text{Tr}(\tilde{\gamma}_{S_1} \tilde{\gamma}_{S_2} H_0) \text{Tr}(\tilde{\gamma}_{S_2} \tilde{\gamma}_{S_3} H_0) \text{Tr}(\tilde{\gamma}_{S_3} \tilde{\gamma}_{S_1} \rho),
 \end{aligned}$$

where

$$g(l_1, l_2, l_3) = \frac{\binom{-1}{l_1+l_2+l_3} \binom{n}{l_1, l_2, l_3}_p \binom{2n}{2l_1+2l_2} \binom{2n}{2l_2+2l_3} \mathcal{B}_{l_1+l_3}}{\binom{2n}{2l_1, 2l_2, 2l_3}_p \binom{n}{l_1+l_2} \binom{n}{l_2+l_3} \mathcal{B}_{l_1+l_2} \mathcal{B}_{l_2+l_3}}, \text{ and}$$

$H_0 = \max_i (H_i - \text{Tr}(H_i) \frac{\mathbb{1}}{2^n})$, and the number of calibration samplings

$$R_c = \mathcal{O} \left(\frac{\mathcal{B}_{\max} \sqrt{n} \ln n \ln(1/\delta_c)}{\mathcal{B}_{\min}^2 \varepsilon_c^2} \right),$$

where $\mathcal{B}_{\max} = \max_k |\mathcal{B}_k|$ and $\mathcal{B}_{\min} = \min_k |\mathcal{B}_k|$. Then, the outputs $\{v_i\}_{i=1}^m$ of the algorithm approximate $\{\text{Tr}(\rho H_i)\}_{i=1}^m$ with error $\varepsilon_e + \varepsilon_c$ and success probability $1 - \delta_e - \delta_c$ under the assumption that $\|H_i\|_\infty = \mathcal{O}(1)$, where $\|H_i\|_\infty$ is the spectral norm of H_i .

We observe that the sampling for estimation we obtained is consistent with Wan et al.²⁸ in the absence of noise. In the following, we will provide an analysis of the necessary number of measurements to compute $\langle^k \mathbf{D}\rangle$ using Algorithm 1. To calculate the representation of each element $\langle^k \mathbf{D}\rangle_{j_1, \dots, j_k; l_1, \dots, l_k}$, where j_b, l_i are in the range $[n]$ for $i \in [k]$, we need to calculate $m = \mathcal{O}(n^k)$ expectations for different $\tilde{\gamma}_S$, where $|S| = 2k$. By choosing the observable $H_j = \tilde{\gamma}_S$ where $|S| = 2k$ and $m = \mathcal{O}(n^k)$ in Theorem 1, with the number of estimation samplings

$$R_e = \mathcal{O} \left(\frac{kn^k \ln(n/\delta_e)}{\mathcal{B}_k^2 \varepsilon_e^2} \right) \tag{10}$$

and the number of calibration samplings

$$R_c = \mathcal{O} \left(\frac{\mathcal{B}_{\max} \sqrt{n} \ln n \ln(1/\delta_c)}{\mathcal{B}_{\min}^2 \varepsilon_c^2} \right), \tag{11}$$

the estimation error can be bounded to $\varepsilon_e + \varepsilon_c$. The equations for R_e and R_c can be simplified to $R_e = \mathcal{O}(\frac{kn^k \ln(n/\delta_e)}{\varepsilon_e^2})$ and $R_c = \mathcal{O}(\frac{\sqrt{n} \ln n \ln(1/\delta_c)}{\varepsilon_c^2})$ for the general noises with constant average fidelity \mathcal{B}_k in subspace Γ_{2k} for any $k \in \{0, \dots, n\}$. We give more details for the calculations in Supplementary Note 8. However, some types of noise channels, such as certain Gaussian unitary channels present in the related $2n \times 2n$ matrix Q , cannot be mitigated with our mitigation algorithm. In particular, there exists a signed permutation matrix Q for which $f_{2k} = 0$, resulting in complete loss of projection for the observable onto the subspace Γ_{2k} . As a result, it is impossible to calculate $\text{Tr}(\rho \tilde{\gamma}_S)$ for any set S containing $2k$ elements. We anticipate that the noise in the quantum device will differ significantly from the U_Q which belongs to the intersection of the matchgate and Clifford groups.

Numerical results

Here we give the numerical results for the mitigated shadow estimation in the fermionic systems. Since the elements of a k -RDM can be expressed in the form $\text{Tr}(\rho \gamma_S)$, we give the numerical results of the errors of the estimators for the expectation value of local fermionic observables $\text{Tr}(\rho \tilde{\gamma}_S)$.

The estimator \hat{v} for $\text{Tr}(\rho \tilde{\gamma}_S)$ can be represented as in Eq. (17). Here we choose the number of qubits $n = 4$ and $S = \{1, 2\}$, and $\tilde{\gamma}_S = U_Q^\dagger \gamma_S U_Q$ and Q is uniformly randomly chosen from $\text{Perm}(2n)$. The quantum state ρ is a uniformly randomly generated 4-qubit pure state. As shown in Fig. 1, we depict the estimations of classical shadow estimators²⁸ and our mitigated Algorithm 1, with the changes of noise strength (Fig. 1a–d) and the changes of the number of samples (Fig. 1e–h). Here we numerically test the estimations with respect to depolarizing, amplitude damping, X -rotation, and Gaussian unitaries.

The calibration samples in Algorithm 1 for the numerics are $N_c = 4,000$ and $K_c = 20$ for all noise models. The number of samples for the classical shadow method is set as $N_e = 4,000$, $K_e = 10$, and for Algorithm 1 are set as $N_e = 4,000/(1 - p_{\text{noise}})$ and $K_e = 10$, where p_{noise} is the noise strength varying for different noise settings:

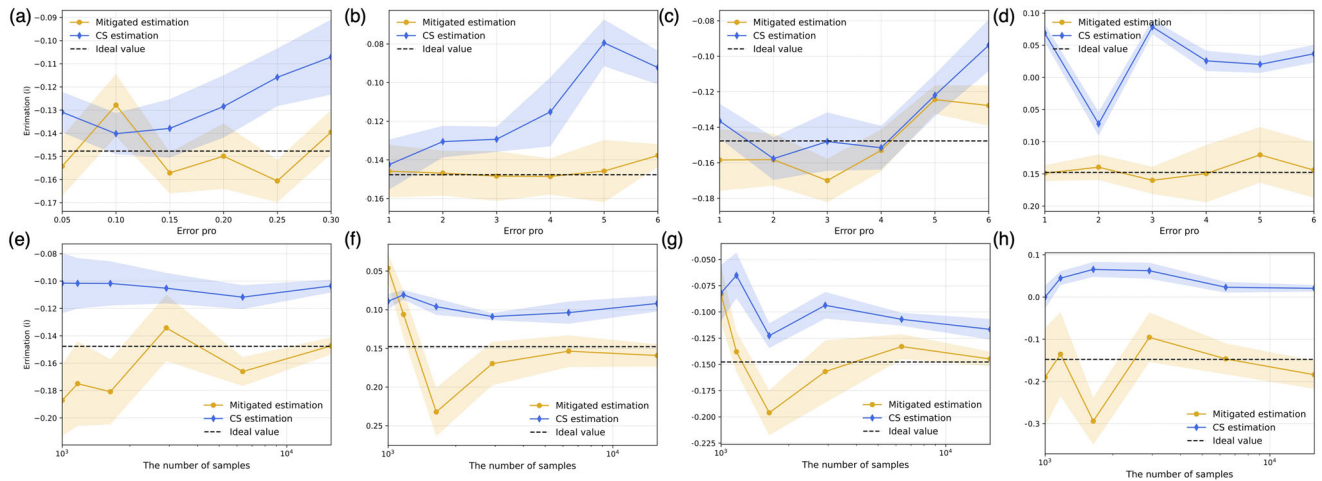


Fig. 1 | The estimations for the expectation values of Majorana operators. The estimations with the changes of (a–d) the noise strength and (e–h) the number of samples for fixed noise parameters, where (a, e) are associated with depolarizing noise, (b, f) are associated with generalized amplitude damping noise, (c, g) are associated with X-rotation noise, (d, h) are associated with Gaussian unitary noise. The error bar is the estimation of the standard deviation by repeating the procedure for $R = 10$ rounds, and it is estimated to be $\sqrt{\sum_{i=1}^R (\hat{v}_i - \bar{v})^2 / R}$, where \bar{v} is the

expectation of the corresponding estimations. Specially, $\bar{v} = \frac{f_{|S|}}{g} \text{Tr}(\rho \gamma_S)$, $g = \hat{f}_{|S|}$ for Mitigated estimation, and $g = \binom{n}{|S|} \binom{2n}{k}$ for CS estimation. In (a–d), the x-axis represents the labels for different noise parameters. The dotted horizontal line indicates the value $\text{Tr}(\rho \tilde{\gamma}_S)$.

- (1) Depolarizing noise $\Lambda_d(\rho) = (1 - p)\rho + p \frac{\mathbb{I}}{2^n}$, where ρ is any quantum state and p is the depolarizing noise strength. In Fig. 1a, p varies from 0.05 to 0.3 ($p = 0.05j$ for x -axis equals j where $j \in [6]$), and in Fig. 1(b), $p = 0.2$.
- (2) Generalized amplitude damping noise Λ_a with representation

$$\Lambda_a(\rho) = \sum_{\substack{u, v \in \{0, 1\}^n \\ u \neq v}} E_{uv} \rho E_{uv}^\dagger + E_0 \rho E_0^\dagger, \tag{12}$$

where $E_{uv} = \sqrt{p_{uv}} |v\rangle \langle u|$ for $u \neq v \in \{0, 1\}^n$ and $E_0 = \sqrt{\mathbb{I} - \sum_{u \neq v} p_{uv} E_{uv}^\dagger E_{uv}}$. Note that Eq. (12) is a generalization of Eq. (6), which connects to Eq. (6) by setting $p_{uv} = \bar{p}_u$ for any $u, v \in [n]$. Here we uniformly randomly choose p_{uv} in $[\frac{0.1+j-1}{6 \times 2^{n+1}}]$, and labeled it as j in the x -axis of Fig. 1 where $j \in [6]$, and choose the generated damping errors for $j = 5$ case in Fig. 1 (b) as the damping errors for Fig. 1f.

- (3) X-rotation noise Λ_r defined in Eq. (7) with noise parameters $\theta_j = \frac{\pi}{2(8-j)}$ for $j \in [6]$, and the noise strength is chosen as $1 - \cos \theta$ in Fig. 1c, and Fig. 1g depicts the errors in the X-rotation noise with noise parameter θ_6 . The x -axis of the mitigation results of X-rotation noise in Fig. 1c denotes the label for noise parameters $\cos \theta_j$ for j range from 1 to 6.
- (4) The Gaussian unitary noise channel \mathcal{U}_Q is chosen such that Q is sampled from the signed permutation group, ensuring that the coefficient f_1 for the noise channel is non-zero. The associated numerical results are shown in Fig. 1d, h. The number of estimation samplings $N_e = 8000$, $K_e = 10$ for Fig. 1d. Fig. 1h choose the same noise parameter with the fifth noise parameter in Fig. 10d. From the figure, we see that without mitigation, the error is enormous with the CS algorithm.

In Fig. 1e–h the number of samples ranges from $\lfloor 900 + 100 \exp(j) \rfloor$ for $j \in \{0, \dots, 5\}$. From Fig. 1, we see that with the increase of the noise strength, the classical shadow method with depolarizing, amplitude damping, and X-rotation noise all gradually diverge to the expected value $\text{Tr}(\rho \tilde{\gamma}_S)$, while our error-mitigated estimation protocol in Alg. 1 gives an

expected value that is close to the noiseless value. Based on the numerical results depicted in Fig. 1e–h, it is evident that as the number of samples increases, the estimation outcomes generated by Algorithm 1 approaches the expected value $\text{Tr}(\rho \tilde{\gamma}_S)$, while the convergent value for the classical shadow method is far from the expected value with depolarizing, amplitude damping, and X-rotation noises. Conversely, the error bar associated with CS estimations is observed to be smaller compared to mitigated estimations, using the same number of samplings, as illustrated in Fig. 1e–h. This is due to the variance of the mitigated estimations associated with the average noise fidelities in Γ_{2k} subspace \mathcal{B}_k .

Discussion

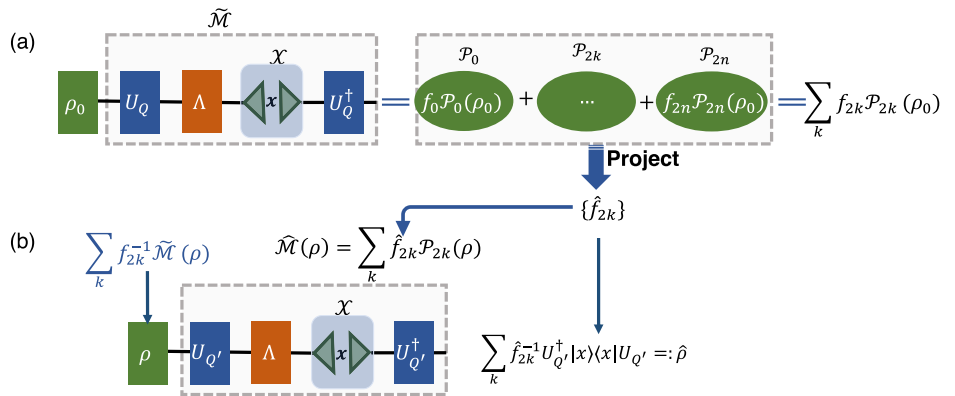
We present an error-mitigated classical shadow algorithm for noisy fermionic systems, thereby extending matchgate classical shadows for noiseless systems^{26,28}. With our method, the calibration process requires a number of copies of the classical state $|0\rangle \langle 0|$ that scales logarithmically with the number of qubits. Assuming a constant average noise fidelity for the noise channel, our algorithm requires the same order of estimation copies as the matchgate classical shadow without error mitigation²⁸. Our algorithm is applicable for efficiently calculating all the elements of a given k -RDM.

To provide a clearer demonstration of the average fidelity of common noises, we consider depolarizing, amplitude damping, and X-rotation noises. The average fidelity of depolarizing and amplitude damping noises are given by $(1 - p)$ and $(1 - \sum_u \bar{p}_u)$ respectively, where p, \bar{p}_u are the noise parameters. For X-rotation noise, the average fidelity lies between $[\min_\theta \cos^k \theta, \max_\theta \cos^k \theta]$ where θ is the rotated angles. To evaluate the effectiveness of our algorithm in mitigating these noises, we compare its performance with the matchgate CS algorithm to calculate the expectations of $\tilde{\gamma}_S$, where $|S| = 2k$, which is crucial for calculating k -RDMs. Our numerical results show good agreement with the theory, validating the effectiveness of our algorithm.

While our algorithm demonstrates good performance in the presence of common types of noise in near-term quantum devices, further investigations are required to explore its potential limitations and improvements. Some of the open questions that can be addressed in future research include:

- (1) Is it possible to extend our algorithm to handle other types of noise, such as time-dependent, non-Markovian and environmental noise⁴⁴,

Fig. 2 | Schematic diagram of the error-mitigated matchgate classical shadow algorithm. a Protocol to learn the noisy classical shadow channel, where $\tilde{\mathcal{M}} = \mathcal{U}_Q^\dagger \mathcal{X} \Delta \mathcal{U}_Q(\cdot) = \sum_{k=0}^n f_{2k} \mathcal{P}_{2k}(\cdot)$, where \mathcal{U}_Q is uniformly randomly sampled from the matchgate group. The estimation \hat{f}_{2k} for f_{2k} is obtained by projecting $\tilde{\mathcal{M}}$ to \mathcal{P}_{2k} subspace with input state $\rho_0 = |\mathbf{0}\rangle \langle \mathbf{0}|$. b The shadow estimation $\hat{\rho}$ for input state ρ with the noisy quantum circuit and classical post-processing with the approximated noisy classical shadow channel in (a), where $\mathcal{U}_{Q'}$ is uniformly randomly sampled from the matchgate group.



or more generally, noise that does not satisfy the GTM assumption? If so, how would these different types of noise impact the performance of our algorithm?

- (2) We provide numerical results under the assumption of Gaussian unitary noise, a common noise model in the fermionic platform. An intriguing unanswered question pertains to the performance of our algorithm in the presence of more typical noise channels inherent to fermionic platform.
- (3) The number of gates required by the matchgate circuit is $\mathcal{O}(n^2)^{45}$. As a result, the accumulation of noise significantly increases the error mitigation threshold^{46,47}, which raises the intriguing question of whether it is feasible to provide an error-mitigated classical shadow using a shallower circuit. This may be compared with Bertoni et al.⁴⁸, who propose a shallower classical shadow approach for qubit systems.
- (4) In addition, we have included in the Supplementary materials the analysis and numerical results regarding the overlap between a Gaussian state and any quantum states, as well as the inner product between a Slater determinant and any pure state. This prompts the question of whether our algorithm can be utilized to calculate other physical, chemical, or material properties beyond the scope of this paper.

Exploring these questions would enhance our comprehension of the potential and limitations of our algorithm, and could potentially pave the way for advancements in the estimation of fermionic Hamiltonian expectation values with near-term quantum devices.

Note added

Following the completion of our manuscript, we became aware of recent independent work by Zhao and Miyake⁴⁹, who also study ways to counteract noise in the fermionic shadows protocol.

Methods

Noisy fermionic channel representation

Here we present an unbiased estimation approach for the noisy representation of the fermionic shadow channel, which utilizes a protocol similar to the matchgate benchmarking protocol⁵⁰. According to representation theory⁵¹ (see details in Supplementary Note 1), the noisy fermionic shadow channel can be represented as $\tilde{\mathcal{M}} = \sum_{k=0}^n f_{2k} \mathcal{P}_{2k}$. Since $\text{Tr}(H\tilde{\mathcal{M}}(\rho)) = \text{Tr}(\rho\mathcal{M}(H))$, with the pre-knowledge of f_{2k} we can calculate $\text{Tr}(H\tilde{\mathcal{M}}(\rho))$ for any observable H in the even subspace. To learn the $2(n+1)$ coefficients, we begin with the easily prepared state $\rho_0 = |\mathbf{0}\rangle \langle \mathbf{0}|$ and apply a noisy unitary channel $\tilde{\mathcal{U}}$ with \mathcal{U} sampled from the matchgate group. We then perform a Z-basis measurement \mathcal{X} with measurement outcomes $x \in \{0, 1\}^n$, followed by classically operating the unitary channel \mathcal{U}^\dagger on $|x\rangle \langle x|$. The generated state has expected value $\sum_{k=0}^n f_{2k} \mathcal{P}_{2k}(\rho_0)$, and f_{2k} is obtained by projecting the final state to the \mathcal{P}_{2k} subspace with some classical post-processing. We illustrate the learning process of the noisy channel in Fig. 2a. The following theorem provides an unbiased estimation of the noisy fermionic classical shadow.

Theorem 2. The noisy fermionic shadow channel can be represented as $\tilde{\mathcal{M}} = \sum_k f_{2k} \mathcal{P}_{2k}$, where \mathcal{P}_{2k} is defined in Eq. (1), and $\hat{f}_{2k} = 2^n \langle \langle \mathbf{0} | \mathcal{P}_{2k} \mathcal{U}_Q^\dagger |x\rangle \rangle / \binom{n}{k}$ is an unbiased estimator of $f_{2k} \in \mathbb{R}$, where $|x\rangle$ is the measurement outcome from the noisy shadow protocol obtained by starting from the input state $|\mathbf{0}\rangle$ and applying a noisy quantum circuit $\tilde{\mathcal{U}}_Q$ followed by a Z-basis measurement, where \mathcal{U}_Q is uniformly randomly picked from the matchgate group.

The representation of the noisy fermionic channel, denoted by $\tilde{\mathcal{M}} = \sum_k f_{2k} \mathcal{P}_{2k}$, where $f_{2k} \in \mathbb{C}$, can be obtained by the irreducible representation of the Gaussian unitary. A detailed proof of this theorem is provided in Supplementary Note 4. We claim that the coefficients f_{2k} can be efficiently calculated with the following lemma.

Lemma 1. \hat{f}_{2k} is the coefficient of x^k in the polynomial $p_Q(x)$, where

$$p_Q(x) = \binom{n}{k}^{-1} \text{pf}(C_{|\mathbf{0}\rangle}) \text{pf}(-C_{|\mathbf{0}\rangle}^{-1} + xQ^T C_{|x\rangle} Q), \quad (13)$$

where $C_{|x\rangle} = \bigoplus_{i=1}^n \begin{pmatrix} 0 & (-1)^{x_i} \\ (-1)^{x_i+1} & 0 \end{pmatrix}$ is the covariance matrix of $|x\rangle$.

This lemma can be obtained by Proposition 1 in Ref. 28. For the completeness of this paper, we also give the proof of this lemma in Supplementary Note 4. The coefficients can be calculated with the polynomial interpolation method in polynomial time. With Theorem 2, we can give an unbiased estimation $\tilde{\mathcal{M}}$ for $\tilde{\mathcal{M}}$.

By the definition of \hat{f}_{2k} , and the twirling properties of $\int_Q d\mu(Q) \mathcal{U}_Q^{\otimes 2}$, the expectation value for the estimation \hat{f}_{2k} can be formulated as

$$f_{2k} = \binom{2n}{2k}^{-1} \binom{n}{k} \mathcal{B}_k. \quad (14)$$

We postpone the details of this proof to Supplementary Note 4. It implies that in the noiseless scenario, $\tilde{\mathcal{M}}$ degenerates into \mathcal{M} as defined in Equation (4). Combined with the definition of \mathcal{B}_k in Eq. (5), f_{2k} is close to $\binom{2n}{2k}^{-1} \binom{n}{k}$ if the average noise fidelity in Γ_{2k} subspace is close to one. Recall that \mathcal{B}_k is a constant in the depolarizing, amplitude-damping, and X-rotation noises with a constant noise strength, which implies that these noises can be efficiently mitigated with our algorithm.

Alternatively, we have a counterexample in Supplementary Note 3 that illustrates the limitations of our mitigation algorithm. Specifically, if the noise follows a Gaussian unitary channel \mathcal{U}_Q where Q is a signed permutation matrix (associated with a discrete Gaussian unitary), then \mathcal{B}_k may become zero. Hence, $f_{2k} = 0$, rendering our estimation approach unsuitable.

Recall that our goal is to estimate $\{\text{Tr}(\rho H_i)\}_{i=1}^m$ using a noisy quantum device and polynomial classical cost, where ρ is an n -qubit quantum state and H_i is an observable in the even subspace Γ_{even} . Here we visualize the estimation process with the guarantee in Theorem 2. We uniformly randomly sample a matchgate U_Q from the matchgate group and apply it to the quantum state ρ , and then measure in the Z -basis to get outcomes x . We define the estimator

$$\hat{v} = \text{Tr}\left(H\widehat{\mathcal{M}}^{-1}\left(U_Q^\dagger|x\rangle\langle x|U_Q\right)\right) \tag{15}$$

$$= \sum_{k=0}^n \hat{f}_{2k}^{-1} \text{Tr}\left(H\widehat{\mathcal{P}}_{2k}\left(U_Q^\dagger|x\rangle\langle x|U_Q\right)\right). \tag{16}$$

It is easy to show that $\text{tr}(H\widehat{\mathcal{M}}^{-1}\widehat{\mathcal{M}}(\rho))$ is an unbiased estimation of $\text{tr}(\rho H)$ when $\widehat{\mathcal{M}}$ is invertible, specifically when $\hat{f}_{2k} \neq 0$ for any k , and H belongs to the even subspace Γ_{even} . Given that $\mathbb{E}[\widehat{\mathcal{M}}] = \widehat{\mathcal{M}}$ and \hat{v} serves as an unbiased estimator of $\text{tr}(H\widehat{\mathcal{M}}^{-1}\widehat{\mathcal{M}}(\rho))$, it implies that the estimation error $\varepsilon := |\hat{v} - \text{tr}(\rho H)|$ is bounded by $|\hat{v} - \text{tr}(H\widehat{\mathcal{M}}^{-1}\widehat{\mathcal{M}}(\rho))| + |\text{tr}(H\widehat{\mathcal{M}}^{-1}\widehat{\mathcal{M}}(\rho)) - \text{tr}(\rho H)|$, which can be minimized with the increasing of the number of samplings for the estimations \hat{v} and $\text{tr}(H\widehat{\mathcal{M}}^{-1}\widehat{\mathcal{M}}(\rho))$, as shown in Eq. (9). Note that the estimator defined in Eq. (16) is not always efficient for all states ρ and observables H . Here we claim that with this estimator, we can efficiently calculate substantial physical quantities such as the expectation value of k -RDM, which not only serves the variational quantum algorithm (VQE) of a fermionic system with up to k particle interactions^{52,53}, but also provide supports to the calculations of derivatives of the energy^{54,55} and multipole moments⁵⁶. It is also an indispensable resource for the error mitigation technique^{8,57}. It also serves to calculate the overlap between a Gaussian state and any quantum state, and the inner product between a Slater determinant and any pure state inspired by the fermionic shadow analysis of Wan et al.²⁸. We postpone the details to Supplementary Note 5.

Note that all elements of k -RDMs can be derived through $\text{Tr}(\rho \gamma_S)$ for a total of $\mathcal{O}(n^k)$ sets S with $|S| = 2k$. In an expansion of this concept, we now focus on evaluating $\text{Tr}(\rho \tilde{\gamma}_S)$ for $\mathcal{O}(n^k)$ different S with $|S| = 2k$. To calculate the expectation value $\text{Tr}(\rho \tilde{\gamma}_S)$, we set the input quantum state to be ρ and the observable to be $H = \tilde{\gamma}_S$ in the estimation formula of Eq. (16), which can then be simplified to

$$\hat{v} = i^k \hat{f}_{2k}^{-1} \text{Pf}(Q_1 Q^T C_{|x\rangle} Q Q_1^T |_S), \tag{17}$$

where $\tilde{\gamma}_j = \sum_{l \in [2n]} Q_1(j, l) \gamma_l$, $Q_1(j, l)$ is the (j, l) -th element of Q_1 , and $C_{|x\rangle} = \bigoplus_{i=1}^n \begin{pmatrix} 0 & (-1)^{x_i} \\ (-1)^{x_i+1} & 0 \end{pmatrix}$ is the covariance matrix of $|x\rangle\langle x|$. Here, $A|_S$ refers to the submatrix obtained by taking the columns and rows of the matrix A that are indexed by S . The simplified quantity can be calculated in polynomial time since \hat{f}_c and the Pfaffian function can be calculated efficiently. We give a detailed proof of the simplification process in Supplementary Note 5.

Data availability

The datasets produced during the current study are available at <https://github.com/GillianOoO/Error-mitigated-fermionic-classical-shadow.git>.

Received: 1 December 2023; Accepted: 3 April 2024;

Published online: 16 April 2024

References

1. Cramer, C. J. *Essentials of computational chemistry: Theories and models* (John Wiley and Sons, Chichester, 2002).
2. Sachdev, S. *Quantum Phase Transitions* (Cambridge University Press, Cambridge, 2011).

3. Kaxiras, E. & Joannopoulos, J. D. *Quantum Theory of Materials* (Cambridge University Press, 2019).
4. Schwerdtfeger, C. A., DePrince III, A. E. & Mazziotti, D. A. Testing the parametric two-electron reduced-density-matrix method with improved functionals: Application to the conversion of hydrogen peroxide to oxywater. *J.Chem. Phys.* **134**, 174102 (2011).
5. Peterson, M. R. & Nayak, C. More realistic Hamiltonians for the fractional quantum hall regime in GaAs and graphene. *Phys. Rev. B* **87**, 245129 (2013).
6. Cerezo, M. et al. Variational quantum algorithms. *Nat. Rev. Phys.* **3**, 625–644 (2021).
7. Parrish, R. M., Hohenstein, E. G., McMahon, P. L. & Martínez, T. J. Quantum computation of electronic transitions using a variational quantum eigensolver. *Phys. Rev. Lett.* **122**, 230401 (2019).
8. Takeshita, T. et al. Increasing the representation accuracy of quantum simulations of chemistry without extra quantum resources. *Phys. Rev. X* **10**, 011004 (2020).
9. Huang, H.-Y., Kueng, R. & Preskill, J. Predicting many properties of a quantum system from very few measurements. *Nat. Phys.* **16**, 1050–1057 (2020).
10. Hadfield, C., Bravyi, S., Raymond, R. & Mezzacapo, A. Measurements of quantum Hamiltonians with locally-biased classical shadows. *Commun. Math. Phys.* **391**, 951–967 (2022).
11. Huang, H.-Y., Kueng, R. & Preskill, J. Efficient estimation of Pauli observables by derandomization. *Phys. Rev. Lett.* **127**, 030503 (2021).
12. Wu, B., Sun, J., Huang, Q. & Yuan, X. Overlapped grouping measurement: A unified framework for measuring quantum states. *Quantum* **7**, 896 (2023).
13. Hadfield, C. Adaptive Pauli shadows for energy estimation *arXiv preprint arXiv:2105.12207* (2021).
14. Hu, H.-Y., Choi, S. & You, Y.-Z. Classical shadow tomography with locally scrambled quantum dynamics. *Phys. Rev. Res.* **5**, 023027 (2023).
15. Acharya, A., Saha, S. & Sengupta, A. M. Shadow tomography based on informationally complete positive operator-valued measure. *Phys. Rev. A* **104**, 052418 (2021).
16. Bu, K., Koh, D. E., Garcia, R. J. & Jaffe, A. Classical shadows with Pauli-invariant unitary ensembles. *npj Quantum Inf.* **10**, 6 (2024).
17. Grier, D., Pashayan, H. & Schaeffer, L. Sample-optimal classical shadows for pure states. *arXiv preprint arXiv:2211.11810* (2022).
18. Ippoliti, M. Classical shadows based on locally-entangled measurements. *Quantum* **8**, 1293 (2024).
19. Zhou, Y. & Liu, Q. Performance analysis of multi-shot shadow estimation. *Quantum* **7**, 1044 (2023).
20. Garcia, R. J., Zhou, Y. & Jaffe, A. Quantum scrambling with classical shadows. *Phys. Rev. Res.* **3**, 033155 (2021).
21. Zhou, Y. & Liu, Z. A hybrid framework for estimating nonlinear functions of quantum states. *arXiv preprint arXiv:2208.08416* (2023).
22. Jordan, P. & Wigner, E. P. *Über das paulische äquivalenzverbot* (Springer, 1993).
23. Nielsen, M. A. The fermionic canonical commutation relations and the Jordan-Wigner transform. *School Phys. Sci. Univ. Queensland* **59**, 75 (2005).
24. Bravyi, S. B. & Kitaev, A. Y. Fermionic quantum computation. *Ann. Phys.* **298**, 210–226 (2002).
25. Tranter, A. et al. The Bravyi-Kitaev transformation: Properties and applications. *Int. J. Quantum Chem.* **115**, 1431–1441 (2015).
26. Zhao, A., Rubin, N. C. & Miyake, A. Fermionic partial tomography via classical shadows. *Phys. Rev. Lett.* **127**, 110504 (2021).
27. Low, G. H. Classical shadows of fermions with particle number symmetry *arXiv preprint arXiv:2208.08964* (2022).
28. Wan, K., Huggins, W. J., Lee, J. & Babbush, R. Matchgate shadows for fermionic quantum simulation. *Commun. Math. Phys.* **404**, 629–700 (2023).
29. Chen, S., Yu, W., Zeng, P. & Flammia, S. T. Robust shadow estimation. *PRX Quantum* **2**, 030348 (2021).

30. Koh, D. E. & Grewal, S. Classical shadows with noise. *Quantum* **6**, 776 (2022).
31. Jnane, H., Steinberg, J., Cai, Z., Nguyen, H. C. & Koczor, B. Quantum error mitigated classical shadows. *PRX Quantum* **5**, 010324 (2023).
32. Valiant, L. G. Expressiveness of matchgates. *Theoretical Computer Sci.* **289**, 457–471 (2002).
33. Knill, E. Fermionic linear optics and matchgates. *arXiv preprint quant-ph/0108033* (2001).
34. Terhal, B. M. & DiVincenzo, D. P. Classical simulation of noninteracting-fermion quantum circuits. *Phys. Rev. A* **65**, 032325 (2002).
35. Zhu, H., Kueng, R., Grassl, M. & Gross, D. The clifford group fails gracefully to be a unitary 4-design. *arXiv preprint arXiv:1609.08172* (2016).
36. Helsen, J., Xue, X., Vandersypen, L. M. & Wehner, S. A new class of efficient randomized benchmarking protocols. *npj Quant. Inf.* **5**, 1–9 (2019).
37. DiVincenzo, D. P. & Terhal, B. M. Fermionic linear optics revisited. *Foundations Phys.* **35**, 1967–1984 (2005).
38. Bonet-Monroig, X., Babbush, R. & O'Brien, T. E. Nearly optimal measurement scheduling for partial tomography of quantum states. *Phys. Rev. X* **10**, 031064 (2020).
39. Wimmer, M. Algorithm 923: Efficient numerical computation of the pfaffian for dense and banded skew-symmetric matrices. *ACM Trans. Math Softw. (TOMS)* **38**, 1–17 (2012).
40. Flammia, S. T. & Wallman, J. J. Efficient Estimation of Pauli Channels. *ACM Transac. Quantum Comput.* **1**, 1–32 (2020).
41. Campbell, E. T. Decoherence in Open Majorana Systems. In Beigi, S. & Koenig, R. (eds.) *10th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2015)*, vol. 44 of *Leibniz International Proceedings in Informatics (LIPIcs)*, 111–126 (Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2015).
42. Onuma-Kalu, M., Grimmer, D., Mann, R. B. & Martín-Martínez, E. A classification of markovian fermionic gaussian master equations. *J. Phys. A: Math. Theoretical* **52**, 435302 (2019).
43. Magesan, E., Gambetta, J. M. & Emerson, J. Scalable and robust randomized benchmarking of quantum processes. *Phys. Rev. Lett.* **106**, 180504 (2011).
44. Van Etten, W. C. *Introduction to random signals and noise* (John Wiley & Sons, New York, NY, USA, 2006).
45. Jiang, Z., Sung, K. J., Kechedzhi, K., Smelyanskiy, V. N. & Boixo, S. Quantum algorithms to simulate many-body physics of correlated fermions. *Phys. Rev. Appl.* **9**, 044036 (2018).
46. Deshpande, A. et al. Tight bounds on the convergence of noisy random circuits to the uniform distribution. *PRX Quant.* **3**, 040329 (2022).
47. Quek, Y., França, D. S., Khatri, S., Meyer, J. J. & Eisert, J. Exponentially tighter bounds on limitations of quantum error mitigation. *arXiv preprint arXiv:2210.11505* (2022).
48. Bertoni, C. et al. Shallow shadows: Expectation estimation using low-depth random clifford circuits. *arXiv preprint arXiv:2209.12924* (2022).
49. Zhao, A. & Miyake, A. Group-theoretic error mitigation enabled by classical shadows and symmetries. *arXiv preprint arXiv:2310.03071* (2023).
50. Helsen, J., Nezami, S., Reagor, M. & Walter, M. Matchgate benchmarking: Scalable benchmarking of a continuous family of many-qubit gates. *Quantum* **6**, 657 (2022).
51. Fulton, W. & Harris, J. *Representation theory: a first course*, vol. 129 (Springer Science & Business Media, New York, NY, USA, 2013).
52. Malone, F. D. et al. Towards the simulation of large scale protein–ligand interactions on nisq-era quantum computers. *Chem. Sci.* **13**, 3094–3108 (2022).
53. Liu, J., Li, Z. & Yang, J. An efficient adaptive variational quantum solver of the schrödinger equation based on reduced density matrices. *J. Chem. Phys.* **154**, 244112 (2021).
54. Overy, C. et al. Unbiased reduced density matrices and electronic properties from full configuration interaction quantum monte carlo. *J. Chem. Phys.* **141**, 244117 (2014).
55. O'Brien, T. E. et al. Calculating energy derivatives for quantum chemistry on a quantum computer. *npj Quantum Inf.* **5**, 113 (2019).
56. Rubín, N. C., Babbush, R. & McClean, J. Application of fermionic marginal constraints to hybrid quantum algorithms. *New J. Phys.* **20**, 053020 (2018).
57. McClean, J. R., Kimchi-Schwartz, M. E., Carter, J. & de Jong, W. A. Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. *Phys. Rev. A* **95**, 042308 (2017).

Acknowledgements

We express our appreciation for the valuable input and feedback given by Jens Eisert, Janek Denzler, and Ellen Derbyshire. We are also thankful for the enlightening conversations held with Xiao Yuan and Yukun Zhang. Furthermore, we extend our gratitude to Janek Denzler and Ellen Derbyshire for their assistance with certain numerical aspects. We also thank Guang Hao Low for their comments on an earlier version of this manuscript. BW acknowledges funding support from the Bundesministerium für Bildung und Forschung (FermiQP, DAQC), Bundesministerium für Wirtschaft und Klimaschutz (EniQMA), and the Munich Quantum Valley (K-8). DEK acknowledges funding support from the Agency for Science, Technology and Research (A*STAR) Central Research Fund (CRF) Award, A*STAR C230917003, and the National Research Foundation, Singapore and A*STAR under its Quantum Engineering Programme (NRF2021-QEP2-02-P03).

Author contributions

Bujiao Wu and Dax Enshan Koh conducted, discussed, and wrote the paper together. Bujiao Wu performed the numerical simulation.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41534-024-00836-7>.

Correspondence and requests for materials should be addressed to Bujiao Wu or Dax Enshan Koh.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024