npj | quantum information

www.nature.com/npjqi

ARTICLE OPEN

# Deterministic improvements of quantum measurements with grouping of compatible operators, non-local transformations, and covariance estimates

Tzu-Ching Yen[1], Aadithya Ganeshram[1] and Artur F. Izmaylov[1,2]✉

Obtaining the expectation value of an observable on a quantum computer is a crucial step in the variational quantum algorithms. For complicated observables such as molecular electronic Hamiltonians, one of the strategies is to present the observable as a linear combination of measurable fragments. The main problem of this approach is a large number of measurements required for accurate estimation of the observable's expectation value. We consider three previously studied directions that minimize the number of measurements: (1) grouping commuting operators using the greedy approach, (2) involving non-local unitary transformations for measuring, and (3) taking advantage of compatibility of some Pauli products with several measurable groups. The last direction gives rise to a general framework that not only provides improvements over previous methods but also connects measurement grouping approaches with recent advances in techniques of shadow tomography. Following this direction, we develop two measurement schemes that achieve a severalfold reduction in the number of measurements for a set of model molecules compared to previous state-of-the-art methods.

*npj Quantum Information* (2023)9:14 ; https://doi.org/10.1038/s41534-023-00683-y

## INTRODUCTION

Variational Quantum Algorithms (VQA) constitute one of the most promising class of applications for quantum computers in the noisy intermediate scale quantum era[1,2]. In VQAs, classically intractable optimization problems are represented as lowest eigenstates of $N_q$-qubit operators

$$\hat{H} = \sum_{n=1}^{N_P} c_n \hat{P}_n,\ \hat{P}_n = \otimes_{k=1}^{N_q} \hat{\sigma}_k \tag{1}$$

where $c_n$ are coefficients and $\hat{P}_n$ are tensor products of Pauli operators or identities, $\hat{\sigma}_k \in \{\hat{x}_k, \hat{y}_k, \hat{z}_k, \hat{1}_k\}$. VQAs then solve these problems by minimizing $E(\boldsymbol{\theta}) = \langle\psi(\boldsymbol{\theta})|\hat{H}|\psi(\boldsymbol{\theta})\rangle$, where the quantum computer prepares the trial wavefunction $|\psi(\boldsymbol{\theta})\rangle$ and is given a task to measure $E(\boldsymbol{\theta})$, while a classical optimizer determines the optimal $\boldsymbol{\theta}$. However, it was found that estimating $E(\boldsymbol{\theta})$ accurately for chemical systems requires large numbers of measurements that diminish VQA's advantage over classical alternatives[3].

Measuring $E(\boldsymbol{\theta})$ is indeed not a straightforward task since only $z$-Pauli operators can be measured on current digital quantum computers. A common approach to measuring the expectation value of the Hamiltonian is to present $\hat{H}$ as a sum of measurable fragments $\hat{H} = \sum_\alpha \hat{A}_\alpha$. The condition for selecting $\hat{A}_\alpha$ is that they can be easily rotated into polynomial functions of $z$-Pauli operators

$$\hat{A}_\alpha = \hat{U}_\alpha^\dagger \left[\sum_i a_{i,\alpha} \hat{z}_i + \sum_{ij} b_{ij,\alpha} \hat{z}_i \hat{z}_j + \ldots\right] \hat{U}_\alpha. \tag{2}$$

Then $\langle\psi(\boldsymbol{\theta})|\hat{H}|\psi(\boldsymbol{\theta})\rangle = \sum_\alpha \langle\psi(\boldsymbol{\theta})|\hat{A}_\alpha|\psi(\boldsymbol{\theta})\rangle$ where the latter can be obtained by measuring $z$-Pauli operators of $\hat{A}_\alpha$ for the rotated wavefunction $\hat{U}_\alpha|\psi(\boldsymbol{\theta})\rangle$.

Unfortunately, in general, the wavefunction $|\psi(\boldsymbol{\theta})\rangle$ is not an eigenstate of $\hat{A}_\alpha$, and thus each fragment requires a set of measurements to obtain an estimator $\bar{A}_\alpha$ for $\langle\psi(\boldsymbol{\theta})|\hat{A}_\alpha|\psi(\boldsymbol{\theta})\rangle$. The efficiency of the Hamiltonian measurement scheme is determined by the total number of measurements, $M$, needed to reach $\epsilon$ accuracy for $E(\boldsymbol{\theta})$. For a simple estimator of $E(\boldsymbol{\theta})$ as the sum of $\bar{A}_\alpha$ estimators, the error scales as $\epsilon = \sqrt{\sum_\alpha \mathrm{Var}_\psi(\hat{A}_\alpha)/m_\alpha}$, where $\mathrm{Var}_\psi(\hat{A}_\alpha) = \langle\psi|\hat{A}_\alpha^2|\psi\rangle - \langle\psi|\hat{A}_\alpha|\psi\rangle^2$ is the variance of each fragment, and $m_\alpha$ are the numbers of measurements allocated for each fragment, with the condition $\sum_\alpha m_\alpha = M$. The optimal distribution of measurements is $m_\alpha \sim \sqrt{\mathrm{Var}_\psi(\hat{A}_\alpha)}$, which gives the total estimator error as $\epsilon = \sum_\alpha \sqrt{\mathrm{Var}_\psi(\hat{A}_\alpha)}/\sqrt{M}$.

This consideration shows superiority of estimators operating with a set of measurable fragments that have the lowest sum over variance square roots. For practical use of this consideration, there are two difficulties in explicit optimization of the estimator error: (1) there is an overwhelming number of choices for measurable operator fragments and (2) variance estimates require knowledge of the wavefunction. The second problem can be addressed by introducing a classically efficient proxy for the quantum wavefunction (e.g. from Hartree-Fock or configuration interaction singles and doubles (CISD) methods in quantum chemistry problems) or by utilizing the measurement results from VQAs to gain empirical estimates if classical efficient proxy cannot be found for the trial wavefunction. Yet, the search space in the first problem is so vast that it has only been addressed heuristically in previous studies. The Hamiltonian partitioning has been done in qubit space[4–11] and in the original fermionic space with subsequent transfer of all operators into the qubit space[12,13]. An initial heuristic idea was to reduce the number of measurable fragments without accounting for variances. It was shown for several partitioning that the number of fragments is not a good

[1]Chemical Physics Theory Group, Department of Chemistry, University of Toronto, Toronto, ON M5S 3H6, Canada. [2]Department of Physical and Environmental Sciences, University of Toronto, Scarborough, Toronto, ON M1C 1A4, Canada. ✉email: artur.izmaylov@utoronto.ca

proxy for the total number of measurements, and the fragments' variances cannot be ignored[13,14]. The key element determining a particular set of measurable fragments is a class of unitary transformations $\hat{U}_\alpha$ in Eq. (2). Compared to single-qubit transformations, multi-qubit transformations are more flexible and therefore have a greater potential to minimize the total number of measurements by selecting fragments with lower variances. Yet, they also have a downside of an extra circuit overhead needed to perform the rotation before the measurement. Once the set of unitary transformations has been selected, empirically, it was found more beneficial for the estimator variance to use greedy algorithms for the Hamiltonian partitioning. In these algorithms one finds $\hat{A}_\alpha$ fragments sequentially by minimizing the norm of the difference between partial sum of $\hat{A}_\alpha$ and $\hat{H}$[13,14]. This can be rationalized considering that greedy algorithms produce first fragments with larger variances and later fragments with smaller variances. Such a distribution of variances makes sum of square roots somewhat smaller compare to the case where variances are distributed relatively equally over all fragments.

Fragmentation techniques in the qubit space are based on grouping mutually commuting Pauli products in each fragment $\hat{A}_\alpha$ [Eq. (2)]. Two types of commutativity between Pauli products are used: qubit-wise and full commutativity. The full commutativity (FC) is the regular commutativity of two operators[7], whereas the qubit-wise commutativity (QWC) for two Pauli products is a condition when corresponding single-qubit operators commute[5]. Using either commutativity to find $\hat{A}_\alpha$, one can efficiently identify unitary operators $\hat{U}_\alpha$ from the Clifford group that bring the fragments to the form of Eq. (2) for measurement. Only one-qubit Clifford gates are sufficient for $\hat{U}_\alpha$ of the qubit-wise commuting fragments[5], while $\hat{U}_\alpha$ for fully commuting fragments require also two-qubit Clifford gates[7].

Initial QWC- and FC-based schemes had $\hat{A}_\alpha$ consisting of disjoint (non-overlapping) sets of Pauli products. Generally, each Pauli product can belong to multiple $\hat{A}_\alpha$ as long as it commutes with all terms in these fragments. This follows from non-transitivity of both FC and QWC as binary relations: if $\hat{P}_1$ commutes with $\hat{P}_2$, and $\hat{P}_2$ commutes with $\hat{P}_3$, this does not lead to commutativity of $\hat{P}_1$ and $\hat{P}_3$. For the measurement problem, $\hat{P}_1$ and $\hat{P}_3$ form separate measurable groups while $\hat{P}_2$ can be measured within both of these groups. Here, $\hat{P}_2$ constitutes an overlapping element for the $\hat{P}_1$ and $\hat{P}_3$ groups (see Fig. 1 where $\hat{P}_1$, $\hat{P}_2$, and $\hat{P}_3$ are $\hat{z}_1$, $\hat{z}_1\hat{z}_2$, and $\hat{x}_1\hat{x}_2$ respectively). Recent developments based on shadow tomography[15–18] and grouping[19,20] techniques exploiting overlapping fragments found considerable reduction in the number of needed measurements over non-overlapping grouping schemes. However, all non-overlapping schemes used in those comparisons did not use the greedy approach. Since within qubit-based partitioning schemes there are multiple estimator improvement techniques, it is interesting to assess them all systematically.

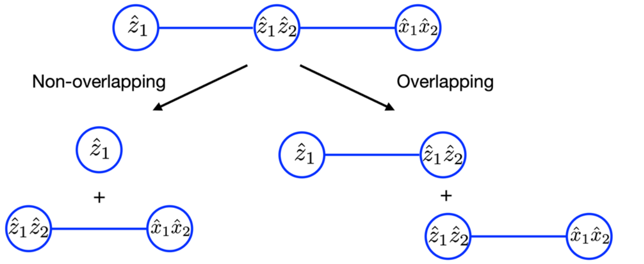

**Fig. 1 Illustration of non-overlapping and overlapping partitioning.** The graph is based on full commutativity for a model Hamiltonian, $\hat{H} = c_1\hat{z}_1 + c_2\hat{z}_1\hat{z}_2 + c_3\hat{x}_1\hat{x}_2$. Within the non-overlapping scheme the fragments are: $\hat{A}_1 = c_1\hat{z}_1$ and $\hat{A}_2 = c_2\hat{z}_1\hat{z}_2 + c_3\hat{x}_1\hat{x}_2$. For the overlapping scheme based on coefficient splitting (measurement allocation) the fragments are: $\hat{A}_1 = c_1\hat{z}_1 + c_2^{(1)}\hat{z}_1\hat{z}_2$ $(\hat{A}_1 = c_1\hat{z}_1 + c_2\hat{z}_1\hat{z}_2)$ and $\hat{A}_2 = c_2^{(2)}\hat{z}_1\hat{z}_2 + c_3\hat{x}_1\hat{x}_2$ $(\hat{A}_2 = c_2\hat{z}_1\hat{z}_2 + c_3\hat{x}_1\hat{x}_2)$.

In this work, we assess improvements in the total number of measurements from introducing a series of ideas: (1) grouping commuting operators using the greedy approach[14], (2) involving non-local (entangling) unitaries for measuring groups of fully commuting Pauli products[6–8,11], and (3) taking advantage of compatibility of some Pauli products with several measurable groups (i.e. overlapping grouping)[15–20]. It is shown that these ideas, used separately or combined, can give rise to schemes superior to prior art within grouping and shadow tomography techniques[16,19]. One of the most striking findings is that using only greedy non-overlapping grouping within the QWC approach can already surpass the performance of recent techniques that employed overlapping local frames. We do not consider fermionic-algebra-based techniques here because they do not allow overlapping grouping while all other improvements were already discussed for them[13]. Other measurement techniques that do not involve grouping of Hamiltonian terms are also outside of the scope of the current work[21–25].

## RESULTS

We assess the performance of the proposed approaches (IMA, GMA, and ICS) in comparison to prior works (LF, SI, and classical-shadow-based algorithms) in estimating energy expectation values for ground eigen-states of several molecular electronic Hamiltonians. The qubit Hamiltonians were generated using the STO-3G basis and the BK transformation. The nuclear geometries for the Hamiltonians are R(H–H) = 1Å ($H_2$), R(Li–H) = 1Å (LiH), R(Be–H) = 1Å with collinear atomic arrangement ($BeH_2$), R(O–H) = 1Å with $\angle HOH = 107.6^\circ$ ($H_2O$), and R(N–H) = 1Å with $\angle HNH = 107^\circ$ ($NH_3$). The overlapping groups ($\mathcal{P}_\alpha$) of the proposed methods are obtained from an extension of the sorted insertion technique (see Supplementary Note 1). The initial measurement allocations or coefficient splittings are derived from measurement allocations of the SI technique using exact or CISD wavefunctions.

To illustrate the relative performance of our methods, Table 1 presents the Hamiltonian estimator variances based on covariances calculated with the exact wavefunction (Supplementary

**Table 1.** Variances of the Hamiltonian estimators using exact wavefunction.

| Systems | LF | SI | IMA | GMA | ICS |
|---|---|---|---|---|---|
| Qubit-wise commutativity | | | | | |
| $H_2$ | 0.136 | 0.136 | 0.136 | 0.136 | 0.136 |
| LiH | 5.84 | 2.09 | 1.73 | 1.52 | 0.976 |
| $BeH_2$ | 14.3 | 6.34 | 5.60 | 5.26 | 4.29 |
| $H_2O$ | 116 | 48.6 | 27.9 | 18.8 | 13.5 |
| $NH_3$ | 352 | 97.0 | 83.3 | 62.1 | 44.8 |
| Full commutativity | | | | | |
| $H_2$ | 0.136 | 0.136 | 0.136 | 0.136 | 0.136 |
| LiH | 1.43 | 0.882 | 0.647 | 0.517 | 0.232 |
| $BeH_2$ | 5.18 | 1.11 | 1.02 | 0.974 | 0.459 |
| $H_2O$ | 43.4 | 7.59 | 5.88 | 4.27 | 1.50 |
| $NH_3$ | 78.7 | 18.8 | 13.6 | 9.35 | 3.32 |

Variances of the Hamiltonian estimators in different methods calculated with the exact wavefunction: largest first (LF), sorted insertion (SI), iterative measurement allocation (IMA), gradient-based measurement allocation (GMA), and iterative coefficient splitting (ICS). Covariances calculated with the exact wavefunction were used for finding optimal parameters in all methods.

**Table 2.** Number of optimization variables.

| Systems | $N_q$ | $N_P$ | QWC | | FC | |
|---|---|---|---|---|---|---|
| | | | MA | CS | MA | CS |
| $H_2$ | 4 | 15 | 3 | 4 | 2 | 6 |
| LiH | 12 | 631 | 155 | 3722 | 42 | 1466 |
| $BeH_2$ | 14 | 666 | 183 | 5946 | 36 | 1203 |
| $H_2O$ | 14 | 1086 | 334 | 11192 | 50 | 1823 |
| $NH_3$ | 16 | 3609 | 1359 | 61137 | 122 | 6138 |

The number of optimization variables in the measurement allocation (MA) and coefficient splitting (CS) methods for the full and qubit-wise commutativities (FC and QWC) and different molecular electronic Hamiltonians. $N_q$ is the number of qubits, and $N_P$ is the number of Pauli products.

Note 5 illustrates the connection of these variances with energy errors). Lower variances in SI compared to those in the largest first (LF) algorithm are consistent with earlier findings[14]. All proposed methods result in lower variances than those in SI. As the most flexible approach, the coefficient splitting method ICS achieves the lowest variances. GMA has a slight edge over IMA in estimator variances, but due to the computational cost of GMA, we will only consider IMA from here on.

Table 2 shows the number of optimization variables in the measurement allocation and coefficient splitting techniques. For the measurement allocation approaches (IMA and GMA) the number of variables is equal to the number of measurable groups. For the qubit-wise (full) commutativity, the number of such groups scales as ~$N_P/3$ ($\sim N_q^3$) since on average each group contains three ($N_q$) Pauli products. For relatively small molecules in our set (i.e. only few atoms), $N_P$ scales as $N_q^4$. In the coefficient splitting approach, the number of variables is a product of $N_P$ and an average number of measurable groups that are compatible with an average Pauli product. For our model systems, it was found empirically that the latter number grows as $\sim N_q^3$ for the qubit-wise commutativity, whereas for the full commutativity the number is within a range of [0.4, 2.3] and thus can be considered relatively constant. These considerations clarify why the measurement allocation techniques can be employed for both commutativities, but the coefficient splitting without extra constraints can be afforded only for full commutativity.

To compare the proposed methods to the classical shadow tomography techniques (Derand[16] and OGM[19]), we consider qubit-wise commuting (QWC) grouping methods that do not require non-local (entangling) transformations and use approximate covariances obtained from CISD wavefunction to choose optimal parameters for the other algorithms (Table 3). Unlike the original OGM treatment, we avoid deleting measurement bases to compare all methods on an equal footing. Comparison between the non-overlapping techniques (LF and SI) and classical shadow techniques reveals that only employing the greedy approach to QWC grouping in SI is already enough to surpass the classical shadow tomography techniques. Due to sensitivity of ICS optimization to inaccurate covariance estimates, we only optimize coefficients of Pauli products with the top 90% CISD variances. The remaining Pauli products have their coefficients frozen to that of the SI scheme. In accord with results of Table 1, both IMA and ICS outperform SI even when approximate covariances are used.

Similarly, switching to application of CISD variances to optimize grouping based on full commutativity clearly shows several times improvements in the number of measurements for IMA and ICS compared to non-overlapping techniques (Table 4).

To explore possible advantages of the IMA and ICS scheme in cases where approximate covariances cannot be obtained from classical wavefunction approximations, we consider a case of random wavefunctions. For all molecular systems corresponding wavefunctions were randomly generated by selecting their coefficients in computational basis from a uniform distribution and renormalizing. These wavefunctions were used to generate exact covariances needed for the overlapping grouping optimizations in IMA and ICS. Table 5 shows that using exact covariances IMA and ICS can improve the number of measurements even for randomly generated wavefunctions.

**Table 3.** Variances with qubit wise commuting fragments.

| Systems | LF | OGM | Derand | SI | IMA | ICS |
|---|---|---|---|---|---|---|
| $H_2$ | 0.136 | 0.173 | 0.144 | 0.136 | 0.136 | 0.136 |
| LiH | 5.84 | 3.50 | 3.74 | 2.09 | 1.73 | 1.07 |
| $BeH_2$ | 14.3 | 18.3 | 12.5 | 6.34 | 5.60 | 4.54 |
| $H_2O$ | 116 | 148 | 114 | 48.6 | 27.9 | 15.9 |
| $NH_3$ | 352 | 305 | 251 | 97.0 | 83.4 | 53.8 |

Variances of Hamiltonian estimators with qubit wise commuting fragments: largest first (LF), overlapped grouping measurement (OGM), derandomization (Derand), sorted insertion (SI), iterative measurement allocation (IMA), and iterative coefficient splitting (ICS). The LF, SI, IMA, and ICS algorithms utilize CISD wavefunctions for choosing parameters, but the final variances are computed using exact wavefunctions.

**Table 4.** Variances of Hamiltonian estimators with fully commuting fragments.

| Systems | LF | SI | IMA | ICS |
|---|---|---|---|---|
| $H_2$ | 0.136 | 0.136 | 0.136 | 0.136 |
| LiH | 1.43 | 0.882 | 0.647 | 0.295 |
| $BeH_2$ | 5.19 | 1.11 | 1.02 | 0.543 |
| $H_2O$ | 43.4 | 7.59 | 5.89 | 2.21 |
| $NH_3$ | 78.8 | 18.8 | 13.7 | 4.95 |

Variances of Hamiltonian estimators with fully commuting fragments: largest first (LF), sorted insertion (SI), iterative measurement allocation (IMA), and iterative coefficient splitting (ICS). All algorithms utilize CISD wavefunctions for choosing parameters, but the final variances are computed using exact wavefunctions.

For considering a more realistic scenario where covariances cannot be evaluated because the wavefunction is not known, it was assumed that covariances can be obtained through accumulated measurement results for any trial wavefunction. A modest 1000 measurements were considered for each fragment to estimate covariances between simultaneously measured Pauli products: $\hat{P}_i$, $\hat{P}_j$. We simulated such measurements to obtain approximate $\langle \hat{P}_i \rangle$, $\langle \hat{P}_j \rangle$ and $\langle \hat{P}_i \hat{P}_j \rangle$. If a Pauli product appears in multiple fragments, measurements in all fragments contribute to the expectation value estimate. The approximate expectation values allow us to estimate covariances between $\hat{P}_i$ and $\hat{P}_j$, which are then used to obtain results shown in Table 6. The results reaffirm that IMA and ICS are the most efficient measurement methods among the presented even with approximate covariances. Note that incorporating measured covariances into measurement optimization can be done more efficiently, as detailed in ref. [20].

Interestingly, the advantage of ICS over IMA diminishes when we use random wavefunctions. This suggests that the extra degrees of freedom in optimizing $c_k^{(\alpha)}$ is not more beneficial to reducing estimator variance than the simple choice

**Table 5.** Average estimator variances with random wavefunction and exact covariances.

| Systems | Derand (QWC) | SI (QWC) | IMA (QWC) | ICS (QWC) | SI (FC) | IMA (FC) | ICS (FC) |
|---|---|---|---|---|---|---|---|
| $H_2$ | 0.241 | 0.233 | 0.226 | 0.219 | 0.202 | 0.185 | 0.177 |
| LiH | 13.6 | 11.2 | 8.64 | 8.59 | 7.43 | 6.18 | 6.13 |
| $BeH_2$ | 45.5 | 38.7 | 29.3 | 29.2 | 24.0 | 21.2 | 21.1 |
| $H_2O$ | 799 | 715 | 517 | 505 | 478 | 410 | 406 |
| $NH_3$ | 865 | 657 | 392 | 391 | 324 | 249 | 246 |

Average variances of the Hamiltonian estimators in methods using qubit-wise and full commutativity (QWC and FC) calculated from 4 random wavefunctions for each system. Optimal parameters for sorted insertion (SI), iterative measurement allocation (IMA), iterative coefficient splitting (ICS) are obtained using the exact covariances.

**Table 6.** Average number of measurements with random wavefunction and approximate covariances.

| Systems | Derand (QWC) | SI (QWC) | IMA (QWC) | ICS (QWC) | SI (FC) | IMA (FC) | ICS (FC) |
|---|---|---|---|---|---|---|---|
| $H_2$ | 0.241 | 0.236 | 0.230 | 0.222 | 0.204 | 0.187 | 0.179 |
| LiH | 13.6 | 11.4 | 8.80 | 8.81 | 7.47 | 6.22 | 6.23 |
| $BeH_2$ | 45.5 | 38.9 | 29.5 | 29.7 | 24.0 | 21.2 | 21.3 |
| $H_2O$ | 799 | 715 | 517 | 510 | 480 | 410 | 407 |
| $NH_3$ | 865 | 658 | 394 | 393 | 324 | 249 | 249 |

Average number of measurements in millions that are required to have $\epsilon = 10^{-3}$ a.u. accuracy in the true expectation values of 4 random wavefunctions for each system. Note that due to the choice of $\epsilon$ and use of millions as units, the numbers here are similar to those in Table 5. These numbers include measurements used for estimating covariances for sorted insertion (SI), iterative measurement allocation (IMA), and iterative coefficient splitting (ICS). The obtained approximate covariances were employed to determine optimal parameters.

$c_k^{(\alpha)} = c_k m_\alpha / M$. Indeed, in the case of random wavefunctions, any Pauli product $\hat{P}_k$ tends to not correlate with fragments consisting of many Pauli products, whose covariances with $\hat{P}_k$ are distributed symmetrically about zero. In such case, it makes intuitive sense to choose $c_k^{(\alpha)}$ to be proportional to the number of times that $\hat{P}_k$ is measured in each group.

## DISCUSSION

We assessed multiple ideas for reduction of the number of measurements required to accurately obtain the expectation value of any operator that can be written as a sum of Pauli products. Since these ideas can be used separately or combined, our main goal was to understand the impact on the number of measurements and incurred computational cost of each idea. Exploring the idea of Pauli products' compatibility led to the realization that the coefficient splitting framework is the most general implementation of this idea for the grouping methods.

Among previously suggested measurement allocation approaches[15,16,19,20] only ref. [20] went beyond QWC fragments and utilized their FC counterparts for the first time. In addition, in ref. [20] analytical formulas for the measurement error were derived and the measurement shots were distributed according to the knowledge on the covariances. Although these techniques have shown performance superior to that of the non-overlapping measurement scheme based on graph-coloring algorithms, by employing a greedy heuristic the non-overlapping scheme can already outperform the Derand and OGM techniques. Thus, for future developments, new approaches need to be compared with greedy grouping-based algorithms rather than with grouping approaches that try to minimize the overall number of measurable groups (e.g. LF).

Unlike previous classical shadow techniques that focus on qubit-wise commuting groups, we also considered measuring techniques involving non-local (entangling) transformations that allow one to measure groups of fully commuting Pauli products. An efficient implementation of these non-local transformations using Clifford gates was proposed by Gottesman[26] and would introduce only $O(N_q^2/\log N_q)$ CNOT gates. The schemes based on fully commuting groups outperform their qubit-wise commuting counterparts up to a factor of seven in variances of the expectation value estimators. Even accounting for increase of the number of measurements related to uncertainties from a lower fidelity of CNOT gates, fully commuting grouping schemes require fewer numbers of measurements than their qubit-wise commuting counterparts[27].

Taking advantage of compatibility of some Pauli products with members of multiple measurable groups (i.e. overlapping groups idea) can be generally presented as augmenting the measurable groups with all Pauli products compatible with initial members of these groups. Then the coefficients of Pauli products entering multiple groups are optimized to lower the estimator variance, with the constraint that the sum over coefficients in different groups for each Pauli product is equal to the coefficient of the Pauli product in the Hamiltonian. This coefficient splitting approach incorporates as a special case a heuristic technique of optimizing measurement allocation for overlapping measurable groups.

Even though the coefficient splitting variance minimization provides the lowest variances among all studied approaches, it requires optimizing a large number of variables: $\sim N_q^4$ ($\sim N_q^7$) for full (qubit-wise) commutativity. Due to certain restrictions, the measurement allocation approach is much more economical in the number of optimization variables: $\sim N_q^3$ ($\sim N_q^4$) for full (qubit-wise) commutativity. Another contributor of the computational cost of these techniques is calculation of the variance gradients. To reduce the computational cost of this part we proposed iterative schemes, the ICS method converges to true extrema, while the IMA scheme deviates from extrema. IMA and ICS provide up to forty and eighty percent reduction in the number of measurements required compared to corresponding best non-overlapping techniques.

Both IMA and ICS use approximate covariances between Pauli products to lower the estimator variance. Use of CISD wavefunction for obtaining these covariances for physically relevant states generally show improvements comparable to those obtained using the exact covariances. In cases where classically efficient approximate wavefunctions are not available, approximate covariances can be obtained via quantum measurements.

## METHODS

### Estimator for non-overlapping Pauli groups

All measurable fragments $\hat{A}_\alpha$ are linear combinations of mutually commuting or qubit-wise commuting Pauli products

$$\hat{A}_\alpha = \sum_k c_k \hat{P}_k,\ \hat{P}_k \in \mathcal{P}_\alpha, \tag{3}$$

where $\mathcal{P}_\alpha$ are disjoint sets of Pauli products measured as parts of corresponding $\hat{A}_\alpha$, and $c_k$ are coefficients of $\hat{P}_k$ in the Hamiltonian. The commutativity between Pauli products within $\mathcal{P}_\alpha$ implies a common eigen-basis $\mathbf{B}_\alpha$, where one can measure all the members of $\mathcal{P}_\alpha$. Initial proposals to find these fragments aim to minimize the total number of fragments using graph coloring algorithms, such as the largest first (LF) algorithm[5,7]. But later the sorted insertion (SI) algorithm employing the greedy approach was found to produce better groups in terms of the number of measurements[14].

Let $\overline{H}$ denotes the estimator for $\langle\psi|\hat{H}|\psi\rangle$; it is a sum of estimators for its parts

$$\overline{H} = \sum_{\alpha=1}^{L} \overline{A}_\alpha. \tag{4}$$

Each $\overline{A}_\alpha$ comes from $m_\alpha$ repeated measurements of $\hat{A}_\alpha$,

$$\overline{A}_\alpha = \frac{1}{m_\alpha}\sum_{i=1}^{m_\alpha} A_{\alpha,i}, \tag{5}$$

where $A_{\alpha,i}$ is the $i$-th measurement result of $\hat{A}_\alpha$. The variance of $\overline{H}$ is

$$\mathrm{Var}\left(\overline{H}\right) = \sum_{\alpha=1}^{L} \mathrm{Var}\left(\overline{A}_\alpha\right), \tag{6}$$

where $\mathrm{Var}\left(\overline{A}_\alpha\right)$ are variances of estimators characterizing differences between $\overline{A}_\alpha$ and the true expectation values $\langle\psi|\hat{A}_\alpha|\psi\rangle$. Note that covariances between different fragments $\mathrm{Cov}(\overline{A}_\alpha, \overline{A}_\beta)$ are zero because measurements of different fragments are done independently. $\mathrm{Var}\left(\overline{A}_\alpha\right)$ can be evaluated using quantum operator variances $\mathrm{Var}_\psi\left(\hat{A}_\alpha\right)$, $\mathrm{Var}\left(\overline{A}_\alpha\right) = \mathrm{Var}_\psi\left(\hat{A}_\alpha\right)/m_\alpha$, which leads to the Hamiltonian estimator variance as

$$\mathrm{Var}\left(\overline{H}\right) = \sum_{\alpha=1}^{L} \frac{1}{m_\alpha}\mathrm{Var}_\psi\left(\hat{A}_\alpha\right). \tag{7}$$

Using the constraint $M = \sum_\alpha m_\alpha$ one can minimize $\mathrm{Var}\left(\overline{H}\right)$ with respect to $m_\alpha$[14,28] which gives

$$\mathrm{Var}\left(\overline{H}\right)_{\min} = \frac{1}{M}\left(\sum_\alpha \sqrt{\mathrm{Var}_\psi\left(\hat{A}_\alpha\right)}\right)^2. \tag{8}$$

with

$$m_\alpha^{(\min)} = \sqrt{\mathrm{Var}_\psi\left(\hat{A}_\alpha\right)}\frac{\sum_\beta \sqrt{\mathrm{Var}_\psi\left(\hat{A}_\beta\right)}}{\mathrm{Var}\left(\overline{H}\right)} \tag{9}$$

Note that this minimization gives generally non-integer $m_\alpha^{(\min)}$. Here and in what follows we will always assume taking the integer approximation $\lfloor m_\alpha \rfloor$ for obtained $m_\alpha$ if $m_\alpha$ are used as integer quantities. In case of large $M$, the difference between $m_\alpha$ and $\lfloor m_\alpha \rfloor$ in the estimator variance is negligible.

The minimum variance in Eq. (8) is generally lower if there is an uneven distribution of $\mathrm{Var}_\psi\left(\hat{A}_\alpha\right)$. This motivates the sorted insertion (SI) algorithm to employ the greedy approach to achieve an uneven distribution of norms of coefficients in fragments, which was found to produce the lowest variances for the energy estimators out of all non-overlapping grouping techniques[14].

In practice, quantum variances $\mathrm{Var}_\psi\left(\hat{A}_\alpha\right)$ are not known a priori. They can be evaluated using covariances between Pauli products,

$$\mathrm{Var}_\psi\left(\hat{A}_\alpha\right) = \sum_{jk} c_j c_k \mathrm{Cov}_\psi\left(\hat{P}_j, \hat{P}_k\right) \tag{10}$$

$$\mathrm{Cov}_\psi\left(\hat{P}_j, \hat{P}_k\right) = \langle\psi|\hat{P}_j\hat{P}_k|\psi\rangle - \langle\psi|\hat{P}_j|\psi\rangle \times \langle\psi|\hat{P}_k|\psi\rangle, \tag{11}$$

where $\hat{P}_j,\ \hat{P}_k \in \mathcal{P}_\alpha$. The covariances for different Pauli products are generally non-zero because all of these Pauli products are measured together within the same fragment. The covariances can be approximated for molecular Hamiltonians using approximate wavefunctions obtained on a classical computer. Configuration interaction singles and doubles (CISD) is one example for obtaining approximation for $|\psi\rangle$ that will be used in the current work. Alternatively, the measurements results obtained from measurement basis $\mathbf{B}_\alpha$ can help estimate the covariances between Pauli products of $\mathcal{P}_\alpha$ during VQA cycles.

### Optimization by coefficient splitting

Many Pauli products in the Hamiltonian can be measured in multiple fragments because of their compatibility with other members of those fragments. The coefficient splitting approach, briefly mentioned in ref. [14], takes advantage of this opportunity by splitting coefficients of Pauli products that are compatible with multiple fragments

$$\hat{A}_\alpha = \sum_k c_k^{(\alpha)} \hat{P}_k,\ \hat{P}_k \in \mathcal{P}_\alpha \tag{12}$$

$$c_k = \sum_{\alpha \in \mathcal{I}_k} c_k^{(\alpha)} \tag{13}$$

where $\mathcal{I}_k$ is a set of group indices $\alpha$ corresponding to fragments $\hat{A}_\alpha$ whose members are compatible with $\hat{P}_k$ (see Fig. 1 for an example). To find fragments $\hat{A}_\alpha$ and to establish compatibility relations between their members we developed an extension of the SI algorithm detailed in Supplementary Note 1. The SI algorithm was taken as the basis of this extension because it produces fragments with a lowest estimator variance among all non-overlapping grouping techniques. From here on, we assume use of the extension for methods proposed in this work.

Note that the equations for the estimator variance and the optimal measurement distribution remain the same [Eqs. (9) and (8)]. However, freedom in the coefficient splitting approach [Eq. (13)] can be used to minimize the Hamiltonian estimator variance [Eq. (8)].

A straightforward approach to minimization of $\mathrm{Var}\left(\overline{H}\right)$ with respect to $c_k^{(\alpha)}$ is to use analytical gradients $\partial \mathrm{Var}\left(\overline{H}\right)/\partial c_k^{(\alpha)}$. The gradients are non-linear functions of $c_k^{(\alpha)}$ and computing them becomes computationally expensive as the number of $c_k^{(\alpha)}$ grows with the size of the system. As a computationally more efficient alternative, we propose an iterative heuristic that quickly converges to a zero gradient solution.

*Iterative coefficient splitting (ICS)*. Given a particular choice of $c_k^{(\alpha)}$ and its optimal $m_\alpha$, the procedure consists of iteratively applying two steps: (1) optimizing $c_k^{(\alpha)}$ with fixed $m_\alpha$ and (2) updating $m_\alpha$ for evaluated $c_k^{(\alpha)}$ using Eq. (9). For step 1, we solve a linear system of equations originating from the $\frac{\partial \mathrm{Var}\left(\overline{H}\right)}{\partial c_k^{(\alpha)}} = 0$ condition (see Supplementary Note 2 for details).

If the number of $c_k^{(\alpha)}$ overcomes computationally affordable limits, one can always limit the minimization to a selected subset of $c_k^{(\alpha)}$. The criteria for the suitable subset can be the $\hat{P}_k$ variances, which correlate with magnitudes of their covariances and therefore the importance of their coefficients for $\mathrm{Var}\left(\overline{H}\right)$.

## Optimization by measurement allocation

Another approach to reducing the Hamiltonian estimator variance is to measure each Pauli product as a member of as many compatible measurable fragments as possible. This idea was used in classical shadow tomography methods based on local transformations for measurement of Pauli products[15,16] and grouping techniques for qubit-wise commuting[19] and fully commuting[20] fragments. First, for a particular Pauli product $\hat{P}_k$, one finds a set of measurement bases $\mathbf{B}_\alpha$ where $\hat{P}_k$ can be measured (see Fig. 1 for an example, by a measurement group this method considers a set of compatible Pauli products). Then, all measurement results for $\hat{P}_k$ obtained in $\mathbf{B}_\alpha$ are used to estimate $\overline{P}_k$:

$$\overline{P}_k = \frac{1}{M_k}\sum_{\alpha\in\mathcal{I}_k}\sum_{i=1}^{m_\alpha} P_{k,i}^{(\alpha)}, \tag{14}$$

where $P_{k,i}^{(\alpha)}$ is the $i$-th measurement result of $\hat{P}_k$ measured in basis $\mathbf{B}_\alpha$, and $M_k = \sum_{\alpha\in\mathcal{I}_k} m_\alpha$ is the total number of times $\hat{P}_k$ is measured. $\overline{P}_k$ are used in the Hamiltonian estimator as $\overline{H} = \sum_k c_k \overline{P}_k$. The variance of $\overline{H}$ is

$$\mathrm{Var}\left(\overline{H}\right) = \sum_{jk} c_j c_k \,\mathrm{Cov}\left(\overline{P}_j, \overline{P}_k\right) \tag{15}$$

$$= \sum_{jk}\frac{c_j c_k}{M_j M_k}\sum_{\substack{\alpha\in\mathcal{I}_j,\\ \beta\in\mathcal{I}_k}}\sum_{u=1}^{m_\alpha}\sum_{v=1}^{m_\beta}\mathrm{Cov}\left(P_{j,u}^{(\alpha)}, P_{k,v}^{(\beta)}\right) \tag{16}$$

To proceed further, it is important to distinguish covariances between Pauli products measured within the same fragment and in different fragments. The former correspond to $\alpha = \beta$ and $u = v$ in Eq. (16) and generally are non-zero, while the latter ($\alpha \neq \beta$ or $u \neq v$) are zero

$$\begin{aligned}\mathrm{Var}\left(\overline{H}\right) &= \sum_{jk}\frac{c_j c_k}{M_j M_k}\sum_{\substack{\alpha\in\mathcal{I}_j,\\ \beta\in\mathcal{I}_k}}\sum_{u=1}^{m_\alpha}\sum_{v=1}^{m_\beta}\delta_{\alpha\beta}\delta_{uv}\mathrm{Cov}_\psi\left(\hat{P}_j, \hat{P}_k\right)\\ &= \sum_{jk}\frac{c_j c_k}{M_j M_k}\sum_{\alpha\in\mathcal{I}_j\cap\mathcal{I}_k} m_\alpha \mathrm{Cov}_\psi\left(\hat{P}_j, \hat{P}_k\right).\end{aligned} \tag{17}$$

Note that the key element in deriving this Hamiltonian estimator variance is the consideration that if a Pauli product is measured as a part of a certain group, all members of this group contribute to the average and to the variance. Thus, the variance of each group gives rise to covariances between its members. Since the covariances in different groups are different in magnitude, placing a particular Pauli product in all compatible groups can be sub-optimal for the total variance of the Hamiltonian estimator (an example illustrating this phenomenon is given in Supplementary Note 4).

Dependencies of $M_j$ and $M_k$ on $m_\alpha$ in $\mathrm{Var}\left(\overline{H}\right)$ [Eq. (17)] make finding the optimal measurement allocation in the analytic form infeasible. To minimize $\mathrm{Var}\left(\overline{H}\right)$ with respect to $m_\alpha$ in Eq. (17) one can numerically optimize $m_\alpha$ as positive variables with restriction $\Sigma_\alpha m_\alpha = M$. We will refer to this strategy as the measurement allocation approach.

Interestingly, the measurement allocation technique is equivalent to a restricted coefficient splitting optimization with $c_k^{(\alpha)} = c_k m_\alpha / M_k$. Indeed, substituting $c_k^{(\alpha)}$ for $m_\alpha$ in $\hat{A}_\alpha$ and using Eq. (7), we obtain $\mathrm{Var}\left(\overline{H}\right)$ as

$$\begin{aligned}\mathrm{Var}\left(\overline{H}\right) &= \sum_\alpha \frac{1}{m_\alpha}\sum_{jk:\alpha\in\mathcal{I}_j\cap\mathcal{I}_k}\mathrm{Cov}_\psi\left(\frac{m_\alpha}{M_j}c_j\hat{P}_j, \frac{m_\alpha}{M_k}c_k\hat{P}_k\right)\\ &= \sum_{jk}\frac{c_j c_k}{M_j M_k}\sum_{\alpha\in\mathcal{I}_j\cap\mathcal{I}_k} m_\alpha \mathrm{Cov}_\psi\left(\hat{P}_j, \hat{P}_k\right),\end{aligned} \tag{18}$$

which agrees with Eq. (17).

One can formulate approximation for gradients of $\mathrm{Var}\left(\overline{H}\right)$ with respect to continuous proxy of $m_\alpha$ (see Supplementary Note 3), which leads to a gradient descent scheme that we will refer to as gradient-based measurement allocation (GMA). Yet, a computationally more efficient, non-gradient iterative scheme was found and detailed below.

*Iterative measurement allocation (IMA).* Given an initial guess for $m_\alpha^{(0)}$ and resulting $M_k^{(0)}$, the corresponding coefficient splitting partitioning of the Hamiltonian is

$$\hat{H} = \sum_\alpha^L \hat{A}_\alpha^{(0)}, \tag{19}$$

where

$$\hat{A}_\alpha^{(0)} = \sum_k \frac{m_\alpha^{(0)}}{M_k^{(0)}} c_k \hat{P}_k,\ \hat{P}_k \in \mathcal{P}_\alpha. \tag{20}$$

Recall that the optimal measurement allocation for any coefficient splitting is given by Eq. (9). Thus, we use this optimal allocation to update $m_\alpha^{(i)}$ as

$$m_\alpha^{(i)} \to m_\alpha^{(i+1)} \propto \sqrt{\mathrm{Var}_\psi\left(\hat{A}_\alpha^{(i)}\right)}, \tag{21}$$

which leads to the update in measurable groups

$$\hat{A}_\alpha^{(i)} \to \hat{A}_\alpha^{(i+1)} = \sum_k \frac{m_\alpha^{(i+1)}}{M_k^{(i+1)}} c_k \hat{P}_k,\ \hat{P}_k \in \mathcal{P}_\alpha \tag{22}$$

Since there is no guarantee that each iteration will necessarily lower $\mathrm{Var}\left(\overline{H}\right)$ in Eq. (17), we repeat these steps multiple times and choose $m_\alpha$ that result in the lowest estimator variance. Empirically, the procedure finds the best measurement allocation in first few cycles.

## Method summary

Conceptually, there are three approaches described above: non-overlapping grouping, coefficient splitting, and measurement allocation. For all of them expectation value of the Hamiltonian is a sum of estimators for expectation values of fragments $\overline{H} = \sum_\alpha \overline{H}_\alpha$, and the variance for the $\overline{H}$ estimator is given by Eq. (7). The differences between three methods are in the fragment definitions: non-overlapping grouping use fragments with original Hamiltonian coefficients $c_k$ for Pauli products and each Pauli products entering only a single fragment, coefficient splitting and measurement allocation allow Pauli products to enter multiple groups with coefficients defined by the optimization procedure for Eq. (12) and $c_k^{(\alpha)} = c_k m_\alpha / M_k$ (cf. Eq. (22)), respectively. Variables that are optimized to obtain the lowest estimator variance are the numbers of measurements $m_\alpha$ for measurement allocation and $c_k^{(\alpha)}$ and $m_\alpha$ for coefficient splitting. The main advantage of the measurement allocation approach is a much lower number of optimization variables ($m_\alpha$) compared to that of the coefficient splitting scheme ($c_k^{(\alpha)}$). Yet, note that positivity of $m_\alpha$ imposes not only a limitation of measurement allocation with respect to coefficient splitting but also with respect to non-overlapping grouping. In non-overlapping grouping, $c_k^{(\alpha)}$ are either 0 or $c_k$, but in measurement allocation, $c_k^{(\alpha)}$ cannot be zero.

## DATA AVAILABILITY

The data that support the findings of this study are available on https://github.com/tymcr/VQE_Hamiltonian.

## CODE AVAILABILITY

Some part of the code that supports the findings of this study is available in the OpenFermion[29] and PySCF[30] libraries. The implementation of ICS is available at Tequila[31] (https://github.com/tequilahub/tequila). The rest of the code is available from the corresponding author upon request.

Received: 26 January 2022; Accepted: 11 February 2023;
Published online: 22 February 2023

## REFERENCES

1. Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).
2. Peruzzo, A. et al. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **5**, 4213 (2014).
3. Gonthier, J. F. et al. Measurements as a roadblock to near-term practical quantum advantage in chemistry: Resource analysis. *Phys. Rev. Res.* **4**, 033154 (2022).
4. Izmaylov, A. F., Yen, T.-C. & Ryabinkin, I. G. Revising the measurement process in the variational quantum eigensolver: is it possible to reduce the number of separately measured operators? *Chem. Sci.* **10**, 3746 (2019).
5. Verteletskyi, V., Yen, T.-C. & Izmaylov, A. F. Measurement optimization in the variational quantum eigensolver using a minimum clique cover. *J. Chem. Phys.* **152**, 124114 (2020).
6. Jena, A., Genin, S. & Mosca, M. Optimization of variational-quantum-eigensolver measurement by partitioning Pauli operators using multiqubit Clifford gates on noisy intermediate-scale quantum hardware. *Phys. Rev. A* **106**, 042443 (2022).
7. Yen, T.-C., Verteletskyi, V. & Izmaylov, A. F. Measuring all compatible operators in one series of a single-qubit measurements using unitary transformations. *J. Chem. Theory Comput.* **16**, 2400 (2020).
8. Gokhale, P. et al. $O(N^3)$ measurement cost for variational quantum eigensolver on molecular hamiltonians. *IEEE Trans. Quantum Eng.* **1**, 1 (2020).
9. Izmaylov, A. F., Yen, T.-C., Lang, R. A. & Verteletskyi, V. Unitary partitioning approach to the measurement problem in the variational quantum eigensolver method. *J. Chem. Theory Comput.* **16**, 190 (2020).
10. Zhao, A. et al. Measurement reduction in variational quantum algorithms. *Phys. Rev. A* **101**, 062322 (2020).
11. Hamamura, I. & Imamichi, T. Efficient evaluation of quantum observables using entangled measurements. *npj Quantum Inf.* **6**, 56 (2020).
12. Huggins, W. J. et al. Efficient and noise resilient measurements for quantum chemistry on near-term quantum computers. *npj Quantum Inf.* **7**, 23 (2021).
13. Yen, T.-C. & Izmaylov, A. F. Cartan subalgebra approach to efficient measurements of quantum observables. *PRX Quantum* **2**, 040320 (2021).
14. Crawford, O. et al. Efficient quantum measurement of Pauli operators in the presence of finite sampling error. *Quantum* **5**, 385 (2021).
15. Hadfield, C., Bravyi, S., Raymond, R. & Mezzacapo, A. Measurements of quantum hamiltonians with locally-biased classical shadows. *Commun. Math. Phys.* **391**, 951 (2022).
16. Huang, H.-Y., Kueng, R. & Preskill, J. Efficient estimation of pauli observables by derandomization. *Phys. Rev. Lett.* **127**, 030503 (2021).
17. Hillmich, S., Hadfield, C., Raymond, R., Mezzacapo, A. & Wille, R. Decision diagrams for quantum measurements with shallow circuits. *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)*, (2021).
18. Hadfield, C. Adaptive Pauli shadows for energy estimation. Preprint at https://arxiv.org/abs/2105.12207 (2021).
19. Wu, B., Sun, J., Huang, Q. & Yuan, X. Overlapped grouping measurement: a unified framework for measuring quantum states. *Quantum* **7**, 896 (2023).
20. Shlosberg, A. et al. Adaptive estimation of quantum observables. *Quantum* **7**, 906 (2023).
21. Radin, M. D. & Johnson, P. Classically-boosted variational quantum eigensolver. Preprint at https://arxiv.org/abs/2106.04755 (2021).
22. Bespalova, T. A. & Kyriienko, O. Hamiltonian operator approximation for energy measurement and ground-state preparation. *PRX Quantum* **2**, 030318 (2021).
23. Wang, G., Koh, D. E., Johnson, P. D. & Cao, Y. Minimizing estimation runtime on noisy quantum computers. *PRX Quantum* **2**, 010346 (2021).
24. Torlai, G., Mazzola, G., Carleo, G. & Mezzacapo, A. Precise measurement of quantum observables with neural-network estimators. *Phys. Rev. Res.* **2**, 022060 (2020).
25. García-Pérez, G. et al. Learning to measure: adaptive informationally complete generalized measurements for quantum algorithms. *PRX Quantum* **2**, 040342 (2021).
26. Aaronson, S. & Gottesman, D. Improved simulation of stabilizer circuits. *Phys. Rev. A* **70**, 052328 (2004).
27. Bansingh, Z. P., Yen, T.-C., Johnson, P. D. & Izmaylov, A. F. Fidelity overhead for non-local measurements in variational quantum algorithms. *J. Phys. Chem. A* **126**, 7007 (2022).
28. Rubin, N. C., Babbush, R. & McClean, J. Application of fermionic marginal constraints to hybrid quantum algorithms. *N. J. Phys.* **20**, 053020 (2018).
29. McClean, J. R. et al. OpenFermion: the electronic structure package for quantum computers. *Quantum Sci. Technol.* **5**, 034014 (2020).
30. Sun, Q. et al. Pyscf: the python-based simulations of chemistry framework. *WIREs Comput. Mol. Sci.* **8**, e1340 (2018).
31. Kottmann, J. S. et al. Tequila: a platform for rapid development of quantum algorithms. *Quantum Sci. Technol.* **6**, 024009 (2021).

## ACKNOWLEDGEMENTS

T.-C.Y. is grateful to Hsin-Yuan Huang and Bujiao Wu for providing details of their respective algorithms in refs. [16] and[19]. A.F.I. acknowledges financial support from the Google Quantum Research Program, Early Researcher Award, and Zapata Computing Inc. This research was enabled in part by support provided by Compute Ontario and Compute Canada.

## AUTHOR CONTRIBUTIONS

T.-C.Y. and A.F.I. conceptualized the project and wrote most of the paper. A.G. developed IMA and collected all the data, except for calculations of LF (in Table 1), SI (in Table 1), GMA, and OGM that T.-C.Y. performed. T.-C.Y., A.G., and A.F.I. participated in discussions that developed the theory as well as the GMA and ICS methods. T.-C.Y. and A.G. share co-first authorship.

## 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/s41534-023-00683-y.

**Correspondence** and requests for materials should be addressed to Artur F. Izmaylov.

**Reprints and permission 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.