Measurement reduction in variational quantum algorithms

Andrew Zhao,[1] Andrew Tranter,[2] William M. Kirby,[2] Shu Fay Ung,[3] Akimasa Miyake,[1] and Peter J. Love[2, *]

[1]*Center for Quantum Information and Control, Department of Physics and Astronomy,*
University of New Mexico, Albuquerque, NM 87131, USA
[2]*Department of Physics and Astronomy, Tufts University, Medford MA*
[3]*California Institute of Technology, Pasadena, CA, USA*

Variational quantum algorithms are promising applications of noisy intermediate-scale quantum (NISQ) computers. These algorithms consist of a number of separate prepare-and-measure experiments that estimate terms in the Hamiltonian. The number of terms can become overwhelmingly large for problems at the scale of NISQ hardware that may soon be available. We use unitary partitioning (developed independently by Izmaylov *et al.*) to define VQE procedures in which additional unitary operations are appended to the ansatz preparation to reduce the number of terms. This approach may be scaled to use all coherent resources available after ansatz preparation. We also study the use of asymmetric qubitization to implement the additional coherent operations with lower circuit depth. We investigate this technique for lattice Hamiltonians, random Pauli Hamiltonians, and electronic structure Hamiltonians. We find a constant factor speedup for lattice and random Pauli Hamiltonians. For electronic structure Hamiltonians, we prove that linear term reduction with respect to the number of orbitals, which has been previously observed in numerical studies, is always achievable. For systems represented on 10–30 qubits, we find that there is a reduction in the number of terms by approximately an order of magnitude. Applied to the plane-wave dual basis representation of fermionic Hamiltonians, however, unitary partitioning offers only a constant factor reduction. Finally, we show that noncontextual Hamiltonians are equivalent to commuting Hamiltonians by giving a reduction via unitary partitioning.

I. INTRODUCTION

Quantum simulation is a promising application of future quantum computers [1–4]. Applications in materials science, chemistry and high energy physics offer the prospect of significant advantages for simulation of quantum systems [5–7]. Calculations on quantum computers that would challenge the classical state of the art require large-scale, error-corrected quantum computers [8]. However, quantum hardware is entering the NISQ (Noisy Intermediate-Scale Quantum) era [9] in which the machines are still too small to implement error correction, but are already too large to simulate classically [10]. It is natural to ask whether NISQ computers can perform useful tasks in addition to demonstrations of quantum supremacy [11, 12].

The Variational Quantum Eigensolver (VQE) was developed to enable quantum estimation of ground state energies on noisy small-scale quantum computers [13]. VQE was developed as a method for quantum simulation of electronic structure [13] and concurrently as a simulation method for quantum field theory by cavity QED [14]. Contemporaneously, the Quantum Approximate Optimization Algorithm (QAOA) was developed as a variational approach to approximate solutions of classical optimization problems [15]. VQE has been widely implemented experimentally due to its simplicity and suitability for NISQ devices [13, 16–20].

VQE consists of preparation of a variational ansatz state by a low-depth parameterized quantum circuit fol-

lowed by estimation of the expectation values of the terms in the Hamiltonian, obtained by measuring each separately. This process is repeated until the statistical error on the expectation value of each term is less than some desired precision threshold. Thus, in VQE the long coherent evolutions of phase estimation are replaced by many independent and short coherent evolutions. However, the necessary number of independent measurements may become overwhelmingly large for problem sizes of ∼50 qubits, which may soon be accessible. Recently, there has been much activity in addressing this measurement problem, via numerous approaches [21–33]. In the present paper, we consider the use of extra coherent resources to reduce the number of separate Pauli terms whose expectation values must be estimated. We refer to this process as *term reduction*. Our methods are closely related to those introduced in [26, 30], which we discuss later.

We consider throughout a k-local Pauli Hamiltonian on n qubits:

$$H = \sum_{j=1}^{m} \alpha_j P_j, \qquad (1)$$

where the m terms $P_j \in \{X, Y, Z, I\}^{\otimes n}$ are tensor products of the Pauli matrices and the 2×2 identity, in which at most k tensor factors are not the identity. We refer to such tensor products as Pauli operators. Such Hamiltonians are called k-local, irrespective of any geometrical locality of the layout of the physical qubits.

The Hamiltonian H for $1 \leq k \leq n$ and $1 \leq m \leq 4^n$ can represent any qubit observable. Interesting cases occur for k a small constant ($2 \leq k \leq 4$) [15] and for k scaling logarithmically with n [34, 35]. Jordan–Wigner map-

pings of fermions to qubits generate Hamiltonians with $k \leq n$, albeit of a restricted form and in which m is still a polynomial in n [36]. Techniques to map interesting physical Hamiltonians to Pauli Hamiltonians on qubits are very well-developed, showing that the Hamiltonian H is expressive enough to represent problems in physics and chemistry ranging from condensed matter models to molecular electronic structure to quantum field theory. Restricting to Eq. (1) is therefore not a significant limitation on the applicability of our results to the simulation of quantum systems.

Let us assume that our NISQ computer is only capable of performing z-basis measurements on individual qubits. To simulate the terms of Eq. (1), it is necessary to map each P_j to a measurement in the computational basis. If our NISQ device has all-to-all pairwise connectivity (as is the case for ion trap NISQ devices) then we require $k - 1$ CNOT gates and up to k single qubit Clifford operations to reduce our measurement of a k-local Pauli operator P_j to a z-basis measurement [37]. If our NISQ computer has only nearest neighbor connectivity on the line we may require an additional $O(n)$ CNOT gates to swap the qubits into an adjacent set.

Any completely commuting set of Pauli operators S_C may be mapped to a set of Pauli words over Z and the identity [37]. This transformation is accomplished simply by the operation mapping the common eigenbasis of the set of operators S_C to the logical basis. Because the eigenbasis of S_C is a set of stabilizer states (with stabilizers given by elements of S_C up to a sign), this map is necessarily a Clifford operation. Clifford operators are known to lack transformation contextuality [38], i.e., they are describable by positive maps on Wigner functions. Furthermore, such a transformation requires an additional $O(n^2)$ gates, with numerical evidence for an $O(n)$ measurement count reduction [27, 29].

Because Clifford operations map single Pauli operators to single Pauli operators, transformations that effect term reduction for a Pauli Hamiltonian must have some non-Clifford structure, and hence must in general possess transformation contextuality. We describe two methods for term reduction based on such transformations. The first technique, unitary partitioning, was previously and independently obtained in [26, 30]. Our second technique provides a more efficient realization of the required transformations at the cost of some ancilla state preparation using asymmetric qubitization—an extension of the linear combination of unitaries model [39]—introduced in [40]. We present these two methods in Section II. In Section III, we evaluate the method for several classes of Hamiltonians. Section IV is devoted to analyzing electronic structure Hamiltonians in depth. We confirm and extend the previous numerical results of [26] that observe a linear term reduction with respect to the number of orbitals is available. We prove that this linear reduction can always be achieved. We also show in Section IV D that unitary partitioning offers a constant factor reduction in the number of terms of a fermionic Hamiltonian

expressed in the plane wave dual basis defined in [21]. Then, in Section V, we show that noncontextual Hamiltonians, defined in [41] (also studied in [42]) are equivalent to commuting Hamiltonians under unitary partitioning. We close the paper with discussion and directions for future work.

II. TERM REDUCTION FOR PAULI HAMILTONIANS

Given a Hamiltonian of the form Eq. (1), we wish to reduce the number of distinct expectation values to estimate in a VQE experiment. The resources available to us are the coherent operations of the quantum computer. Suppose that our ansatz $|\psi_A\rangle$ is prepared by a quantum circuit U from the state $|\psi_0\rangle \equiv |0\rangle^{\otimes n}$, so that

$$|\psi_A\rangle = U |\psi_0\rangle. \qquad (2)$$

Then our experiment estimates the expectation values

$$\langle P_j \rangle = \langle \psi_0 | U^\dagger P_j U |\psi_0\rangle. \qquad (3)$$

Suppose instead we rewrite our Hamiltonian in terms of a different set of Pauli operators $\{Q_l\}_{l=1}^{m_c}$ and unitary operations $\{R_l\}_{l=1}^{m_c}$ as follows:

$$H = \sum_{j=1}^{m} \alpha_j P_j = \sum_{l=1}^{m_c} \gamma_l R_l^\dagger Q_l R_l. \qquad (4)$$

Such decompositions give the correct variational estimate:

$$\langle \psi_A | H |\psi_A\rangle = \sum_{j=1}^{m} \alpha_j \langle \psi_A | P_j |\psi_A\rangle \qquad (5)$$

$$= \sum_{l=1}^{m_c} \gamma_l \langle \psi_A | R_l^\dagger Q_l R_l |\psi_A\rangle. \qquad (6)$$

Each term labeled by l is estimated by a separate prepare and measure ansatz which appends a different unitary R_l to the ansatz preparation. The unitary rotations R_l therefore represent the additional coherent resources required to reduce the number of separate expectations to be obtained.

Unlike the approach of [26], we do not estimate the unitary operators $R_l^\dagger Q_l R_l$ themselves. Instead, we propose to perform a set of m_c experiments in which the coherent operations R_l are appended to U, so that the expectation values are obtained by measuring P_j in the resultant state. In this case, the R_l may be made as simple or as complex as the coherent resources available after state preparation allow. Term reduction therefore allows the use of VQE for larger systems, by optimally using the increasing amount of coherent resources available in new devices.

A. Unitary partitioning

We will apply rotations in the adjoint representation of $\mathfrak{su}(2^n)$ with the goal of reducing the number of Pauli terms in the Hamiltonian. For classical algorithms the number of such terms is not a relevant variable as one must represent all the nonzero terms of the Hamiltonian in some way. There are some general constraints on the form of terms arising from a Pauli matrix by an adjoint unitary action. We now consider what resources the R_l operations require and give constructions that achieve term reduction. These ideas were previously presented in [26].

We may write:

$$R_l^\dagger Q_l R_l = \sum_j \beta_{lj} P_{f(l,j)}, \qquad (7)$$

where f is a relabelling of generalized Pauli matrices. Any unitary rotation of a generalized Pauli matrix is self-inverse, so $(R_l^\dagger Q_l R_l)^2 = \mathbb{1}$, which implies

$$\sum_j \beta_{lj}^2 = 1 \ \text{ and } \ \sum_{j<k} \beta_{lj}\beta_{lk}\{P_{f(l,j)}, P_{f(l,k)}\} = 0. \quad (8)$$

The first constraint can be satisfied for any subset of terms by scaling the coefficients β_{lj} by appropriately defining γ_l. The second constraint is the defining property of subsets of terms which can be combined into a single term by unitary rotation. We focus on the case where we divide the terms of the Hamiltonian into *completely anticommuting sets*: sets in which each pair of operators anticommutes. We then rescale these terms to satisfy the first constraint, and seek unitary operators that map each set to a single Pauli operator.

The *compatibility graph* associated to a set of Pauli operators is an undirected graph whose vertices are the operators in the set, and in which a pair of vertices is connected if the associated operators commute. Completely anticommuting sets of Pauli operators are independent sets of the compatibility graph. A partition of the operators into completely anticommuting sets is provided by a coloring of the vertices of the graph such that no two vertices connected by an edge have the same color. The number of sets is determined by the number of colors. Graph coloring is a well-known NP-complete problem; however, we only require the number of colors to be less than the number of vertices for our method to provide a reduction in the number of terms. A detailed study of the use of various heuristics for graph coloring for the compatibility graphs of Hamiltonians was performed in [26].

We now construct the rotation R that maps a completely anticommuting set to a single Pauli operator by conjugation. Let S be a set of Pauli operators appearing in the Hamiltonian, such that $\{P_j, P_k\} = 0$ $\forall P_j \neq P_k \in S$. Further, it will be useful to define $s = |S|$. The set of terms corresponding to S in the Hamiltonian

is then written

$$H_S = \sum_{P_j \in S} \beta_j P_j. \qquad (9)$$

We will assume for now that the coefficients satisfy

$$\sum_j \beta_j^2 = 1. \qquad (10)$$

We define the following Hermitian, self-inverse operators:

$$\mathcal{X}_{sk} = iP_s P_k, \quad 1 \leq k \leq s-1. \qquad (11)$$

It is straightforward to verify that $\mathcal{X}_{sk}$ commutes with all $P_j \in S$ for $j \neq s$, $j \neq k$, and that it anticommutes with P_k and P_s.

We define the adjoint rotation generated by $\mathcal{X}_{sk}$:

$$R_{sk} = \exp\left(-i\frac{\theta_{sk}}{2}\mathcal{X}_{sk}\right), \qquad (12)$$

whose action on the terms in H_S is given by

$$\begin{aligned} R_{sk}P_k R_{sk}^\dagger &= \cos\theta_{sk}P_k + \sin\theta_{sk}P_s, \\ R_{sk}P_s R_{sk}^\dagger &= -\sin\theta_{sk}P_k + \cos\theta_{sk}P_s. \end{aligned} \qquad (13)$$

That is, R_{sk} is an adjoint rotation acting in the space spanned by P_s and P_k.

If we act on H_S with R_{sk}, we obtain

$$\begin{aligned} R_{sk}H_S R_{sk}^\dagger &= (\beta_k \cos\theta_{sk} - \beta_s \sin\theta_{sk})P_k \\ &\quad + (\beta_k \sin\theta_{sk} + \beta_s \cos\theta_{sk})P_s \\ &\quad + \sum_{P_j \in S\backslash\{P_k, P_s\}} \beta_j P_j. \end{aligned} \qquad (14)$$

Choosing $\beta_k \cos\theta_{sk} = \beta_s \sin\theta_{sk}$ therefore gives a rotation of the Hamiltonian with the P_k term removed, and where the norm of the term P_s is increased from β_s to $\sqrt{\beta_s^2 + \beta_k^2}$. Defining the operator

$$R_S = R_{s(s-1)}(\theta_{s(s-1)})\cdots R_{s2}(\theta_{s2})R_{s1}(\theta_{s1}), \qquad (15)$$

where the angles θ_{sk} satisfy

$$\beta_1 \cos\theta_{s1} = \beta_s \sin\theta_{s1}, \qquad (16)$$

and, for $k > 1$,

$$\beta_k \cos\theta_{sk} = \sqrt{\left(\beta_s^2 + \sum_{j=1}^{k-1}\beta_j^2\right)}\sin\theta_{sk}, \qquad (17)$$

therefore gives

$$R_S H_S R_S^\dagger = P_s, \qquad (18)$$

where we used the fact that $\sum_{j=1}^s \beta_j^2 = 1$. Care must be taken when choosing θ_{sk} so as to obtain the positive root.

Our decomposition strategy is therefore the following:

$$H = \sum_{j=1}^{m} \alpha_j P_j = \sum_{l=1}^{m_c} \gamma_l H_{S_l}, \qquad (19)$$

where

$$H_{S_l} = \sum_{P_j \in S_l} \beta_{lj} P_j \qquad (20)$$

has support on a set S_l of self-inverse operators for which $\{P_j, P_k\} = 0 \ \forall j \neq k$ and $\sum_j \beta_{lj}^2 = 1$. Each H_{S_l} can be obtained from a single Pauli operator by a unitary rotation as in Eq. (15), so we can rewrite Eq. (19) as

$$H = \sum_{l=1}^{m_c} \gamma_l R_{S_l}^\dagger P_{s_l} R_{S_l}, \qquad (21)$$

where the R_{S_l} operators are given for each set of pairwise anticommuting operators by Eq. (15).

For each H_S we must therefore append to our ansatz preparation the set of $s - 1$ operators R_{sk} (recall that $s = |S|$). For an l-local Hamiltonian each of these requires $O(l)$ CNOT and single qubit rotations to implement. Hence one exchanges s separate Pauli expectation value estimations for a single expectation value estimation, at the cost of $O(sl)$ additional coherent operations. Note that directly appending these transformations to the ansatz preparation results in a factor of two reduction in the required coherent resources as compared to [26], where both R and $R^\dagger$ must be implemented as controlled operations.

The decomposition given above and in [26] is the most direct implementation of the transformation of the Hamiltonian. Improvement can be made through the use of ancilla qubits and more coherent resources, as we now show in Section II B.

B. Low-depth implementation of the rotations

In Section II A and in [26] an ordered sequence of rotations is used to write a completely anticommuting set of Pauli operators as a single term. Here we will show how to use a single rotation to perform the same reduction, and show how to implement this rotation using the methods based on linear combinations of unitaries (LCU) [39].

We define a set of operators H_k for $1 \leq k \leq n$ such that $H_1 = P_1$, $H_n = \sin \phi_{n-1} H_{n-1} + \cos \phi_{n-1} P_n$. Each H_n is self-inverse, and we consider rotations of H_n around an axis that is Hilbert–Schmidt orthogonal to both H_{n-1} and P_n. The operator defining this axis is:

$$\mathcal{X} = \frac{i}{2} [H_{n-1}, P_n]. \qquad (22)$$

The operator $\mathcal{X}$ is self-inverse, anticommutes with H_n, and so $[\mathcal{X}, H_n] = 2\mathcal{X} H_n$. Furthermore, we may show that

$$\mathcal{X} H_n = i(-\sin \phi_{n-1} P_n + \cos \phi_{n-1} H_{n-1}). \qquad (23)$$

The operator $\mathcal{X}$ generates the rotation

$$R = \exp(-i\alpha \mathcal{X}/2) = \cos(\alpha/2)\mathbb{1} - i\sin(\alpha/2)\mathcal{X}. \qquad (24)$$

The adjoint action of R on H_n is given by:

$$RH_nR^\dagger = \sin(\phi_{n-1} - \alpha)H_{n-1} + \cos(\phi_{n-1} - \alpha)P_n. \qquad (25)$$

Choosing $\alpha = \phi_{n-1}$ therefore gives $RH_nR^\dagger = P_n$. This is a simple constructive demonstration that any self-inverse operator supported on a set of pairwise anticommuting operators S can be mapped to a single Pauli operator. (The details of these calculations can be found in Appendix A 1.)

The terms in the operator $\mathcal{X}$ all pairwise anticommute, and $\mathcal{X}$ squares to the identity. This yields the expression for R given in Eq. (24). This expression is a linear combination of Pauli operators (which are unitary). This naturally suggests implementation of R using LCU [39]. These methods can be combined with qubitization and quantum signal processing to reduce the required gate count [43–46]. However, $\mathcal{X}$ has coefficients that are L_2 normalized to one, whereas the original (symmetric) LCU methods naturally treat Hamiltonians with L_1 normalized coefficients. Fortunately this issue was already addressed in [40] in which an asymmetric LCU (ALCU) method was introduced. We propose the ALCU method for the implementation of R. Because R is equivalent to evolution under the Hamiltonian $\mathcal{X}$, the cost of asymmetric qubitization scales as the square root of the number of terms in $\mathcal{X}$, and hence the use of this method offers a quadratic speedup in asymptotic scaling compared to the methods of Section II A and [26].

ALCU requires $O(\log s)$ additional qubits (s being the maximum size of any of the anticommuting sets) and more complex gate operations than the method of Section II A and [26]. However, the use of these methods in the context of VQE provides a motivation to implement more sophisticated quantum algorithms on NISQ devices. It should be noted that implementation of ALCU for this purpose is much simpler than its use for direct simulation of time evolution under the original Hamiltonian. This is because the number of terms in $\mathcal{X}$ is only equal to the number of terms in an anticommuting set. As we will discuss in detail below, this can be made smaller in order to take advantage of any additional coherent resources available after state preparation.

C. Commuting terms

Requiring that the sets of terms to be combined anticommute, as in Sections II A and II B, is sufficient but not necessary to perform term reduction. If there is additional structure on the coefficients of the Hamiltonian, the second constraint in Eq. (8) may be satisfied without the individual terms all vanishing. Here we consider the possibility that for some l,

$$\sum_{j<k} \beta_{lj}\beta_{lk}\{P_j, P_k\} = 0 \qquad (26)$$

while the individual terms are nonzero (note that we have simplified the labelling of the Pauli terms). Because generalized Pauli matrices have the property that they either commute or anticommute we can restrict attention to the subset of the operators that commute. We then require that

$$\sum_{j<k} \beta_{lj}\beta_{lk}\{P_j, P_k\} = 2 \sum_{S(l,j,k)} \beta_{lj}\beta_{lk}P_jP_k = 0, \quad (27)$$

where $S(l, j, k)$ is the set of indices satisfying $j < k$ and $[P_j, P_k] = 0$. Each term here is nonzero so the condition must be enforced by cancellation of pairs, i.e., due to relations of the form:

$$\beta_{lj}\beta_{lk}P_jP_k + \beta_{ls}\beta_{lr}P_sP_r = 0. \quad (28)$$

This can only be true if $|\beta_{lj}\beta_{lk}| = |\beta_{ls}\beta_{lr}|$, and so this possibility of term reduction depends on the details of the coefficients more sensitively than simply requiring all terms to anticommute in a particular subset.

Supposing that the conditions on pairs of coefficients are satisfied, we also require that

$$P_jP_k \pm P_sP_r = 0 \quad (29)$$

(for $\beta_{lj}\beta_{lk} = \pm\beta_{ls}\beta_{lr}$). Suppose the pairs (j, k) and (s, r) have one operator in common, $j = s$. Then our requirement is $P_k = \pm P_r$, meaning that (j, k) and (s, r) are the same pair. Hence the pairs (j, k) and (s, r) must be completely distinct. This implies that $P_jP_k = P_t$ and $\pm P_sP_r = P_t$. This is perfectly possible: for example, if $P_k = IX$, $P_j = XI$, $P_r = ZZ$, and $P_s = YY$, then $P_kP_j = XX$ and $P_rP_s = -XX$. We leave further investigation of this possibility for term reduction to future work.

D. Total measurement cost estimates

Achieving precision ϵ in the estimate of the expectation value $\langle H \rangle$ requires a statistically significant sample of qubit measurements for each Pauli term in H. Naively, this requires approximately $|\alpha_j|^2/\epsilon^2$ measurements for the jth term, where α_j is its associated weight. However, it was proposed in [47], and formally proven in [22], that the optimal number of measurements per term is

$$M_j = \frac{|\alpha_j|\sigma_j}{\epsilon^2}\left(\sum_{k=1}^m |\alpha_k|\sigma_k\right), \quad (30)$$

where $\sigma_j^2 = \langle P_j^2 \rangle - \langle P_j \rangle^2$ is the operator variance of the jth term. Using $\sigma_j^2 \leq 1$ for all self-inverse operators, the upper bound for the total number of measurements to estimate the full Hamiltonian is [22]

$$M = \sum_{j=1}^m M_j = \left(\frac{1}{\epsilon}\sum_{j=1}^m |\alpha_j|\sigma_j\right)^2 \leq \frac{\Lambda^2}{\epsilon^2}, \quad (31)$$

where $\Lambda = \sum_{j=1}^m |\alpha_j|$ is the L_1 norm of the Hamiltonian weights.

Using the standard inequalities

$$\frac{1}{\sqrt{d}}||x||_1 \leq ||x||_2 \leq ||x||_1 \quad (32)$$

for any $x \in \mathbb{R}^d$, where $||\cdot||_p$ denotes the L_p norm, we may establish bounds for the value of Λ^2 after transforming the Hamiltonian via unitary partitioning. We reuse the notation of Eqs. (19) and (21), so that

$$H = \sum_{j=1}^m \alpha_j P_j \quad (33)$$

is the Hamiltonian as given, and

$$H = \sum_{l=1}^{m_c} \gamma_l R_{S_l} P_l R_{S_l}^\dagger \quad (34)$$

is its form after unitary partitioning. Note that the variances of the $R_{S_l}P_lR_{S_l}^\dagger$ terms satisfy the same upper bound as Pauli operators. Since the coefficients associated with each anticommuting set S_l must be L_2 normalized, the resulting weights γ_l are the L_2 norms of those coefficients:

$$\gamma_l^2 = \sum_{k \in S_l} \alpha_k^2. \quad (35)$$

By abuse of notation, here we use S_l to denote the index set on which its elements are supported.

Let Λ be the L_1 norm of the weights $\{\alpha_j\}_{j=1}^m$ as before, and Λ_c be the L_1 norm of $\{\gamma_l\}_{l=1}^{m_c}$. Then, using the right-hand inequality of Eq. (32), we have

$$\begin{aligned}
\Lambda_c = \sum_{l=1}^{m_c} |\gamma_l| &= \sum_{l=1}^{m_c} \sqrt{\sum_{k \in S_l} \alpha_k^2} \\
&\leq \sum_{l=1}^{m_c} \sum_{k \in S_l} |\alpha_k| \quad (36) \\
&= \sum_{j=1}^m |\alpha_j| = \Lambda.
\end{aligned}$$

Thus $\Lambda_c \leq \Lambda$, and in fact this bound is only saturated if no partitioning is performed at all.

Applying the left-hand inequality of Eq. (32) to the first line of Eq. (36) yields

$$\sum_{l=1}^{m_c}\left(\frac{1}{\sqrt{|S_l|}}\sum_{k \in S_l} |\alpha_k|\right) \leq \Lambda_c. \quad (37)$$

Let $s_{\max} = \max_l |S_l|$ be the size of the largest set in the partition. Then

$$\frac{1}{\sqrt{s_{\max}}}\sum_{l=1}^{m_c}\sum_{k \in S_l} |\alpha_k| = \frac{\Lambda}{\sqrt{s_{\max}}} \leq \Lambda_c. \quad (38)$$

Bounding all the set sizes by $s_{\max}$ is fairly tight if all the sizes are all roughly equal, which is both desirable (since the gate complexity scales with the size) and possible (one may take a large set and simply divide it into smaller ones, which remain fully anticommuting). The total number of measurements M_c under unitary partitioning can be thought of as being lower bounded by $M/s_{\max}$, although this is not technically correct, since Λ (resp. Λ_c) is itself an upper bound estimate for M (resp. M_c). Eq. (38) gives an rough sense for the maximum amount of measurement reduction possible by unitary partitioning, when taking into account the statistical repetitions.

It is worth noting that this lower bound is saturated when $|\alpha_j| = |\alpha_k| \ \forall j, k$. In fact, a weaker condition saturates the tighter bound of Eq. (37). There, we only require that $|\alpha_j| = |\alpha_k| \ \forall j, k \in S_l$ for each l. That is, if the coefficients associated with each set are equal merely among themselves, then the lower bound of Eq. (37) may be achieved. Supposing that this approximately holds, and again that all $|S_l|$ are roughly the same, yields $\Lambda_c \approx \Lambda/\sqrt{s_{\max}}$.

Thus partitioning with additional constraints respecting these coefficient conditions may result in even more measurement reduction, without requiring any additional coherent rotations. The partitioning algorithm would then require significantly more classical computational resources, as this is now a *weighted* graph coloring problem, but in principle these ideas may be implemented straightforwardly. For the analysis in the following section, we focus only on the number of unique Hamiltonian terms before and after partitioning as a rough estimate for the amount of measurement reduction achieved by our method.

III.   APPLICATIONS

A.   The Transverse field Ising model (TIM) in one-dimension

To give a simple realization of these ideas we consider the TIM on a 1D lattice with L sites and periodic boundary conditions:

$$H = \sum_{j=1}^{L} (Z_{j+1}Z_j + xX_j). \tag{39}$$

We can divide this Hamiltonian into sets of completely anticommuting operators by observing that no pair of Z terms and no pair of X terms can be in the same anticommuting set. We therefore choose pairs of anticommuting operators composed of $Z_{j+1}Z_j$ and X_{j+1}.

We then write:

$$Z_{j+1}Z_j + xX_{j+1}I_j = \sqrt{1+x^2} \left(\frac{Z_{j+1}Z_j + xX_{j+1}I_j}{\sqrt{1+x^2}} \right). \tag{40}$$

From Eqs. (11) and (12) we define the operator

$$R_j = \exp\left(\frac{i\theta}{2} Y_{j+1} Z_j \right) \tag{41}$$

$$= \mathrm{CNOT}_{(j+1,j)} \times \exp\left(\frac{i\theta}{2} Y_{j+1} I_j \right), \tag{42}$$

where θ is given from Eq. (17):

$$\frac{x}{\sqrt{1+x^2}} \cos\theta = \frac{1}{\sqrt{1+x^2}} \sin\theta \tag{43}$$

Our final Hamiltonian decomposition is then:

$$H = \sum_{j=1}^{L} \left[R_j^\dagger Z_j Z_{j+1} R_j \right] \tag{44}$$

whereas our initial Hamiltonian had $2L$ terms, our final Hamiltonian has L terms. While this is only a constant factor in the reduction of terms, it demonstrates that term reduction is possible by this method.

B.   Transverse Ising models on arbitrary graphs

If we consider transverse Ising Hamiltonians defined on arbitrary graphs our analysis does not change very much. The maximum size of a totally anticommuting set is still two, independent of the graph. This is because once a single local X_i is included in the set, one can include only one Z_iZ_j term in the set before exhausting all possibilities. Hence the number of terms in a transverse Ising Hamiltonian on a general graph with vertex set V and edges E can be reduced from $|E| + |V|$ to $|E|$. This cannot change the asymptotic scaling of the number of terms as a function of the number of vertices. In particular, for regular graphs with degree q the number of edges is $|V|q/2$ and the number of terms in the transverse Ising model Hamiltonian is $|V|(1 + q/2)$ which can be reduced to $|V|q/2$, a constant factor improvement of $q/(q + 2)$. Note that this case includes lattice models. The relative lack of performance here is due to the presence of little anticommutative structure in the operators of the transverse field Ising model.

C.   Compatibility graphs of random Hamiltonians

Randomly choosing Pauli terms from the complete set of n-qubit Pauli observables corresponds to selecting a subset of the vertices of the full compatibility graph of all Pauli observables. The resulting compatibility graphs can only be subgraphs of this graph, which has a finite geometric structure considered in [48]. Therefore randomly sampling Pauli terms, resulting in an edge in the compatibility graph with given probability, say p, does not result in Erdős–Rényi random graphs given by populating edges with probability p. The constraint that

the graphs arising be subgraphs of the full compatibility graph of all Pauli operators causes this deviation.

However, for large numbers of qubits, fixed locality of operators and a number of Pauli terms scaling polynomially with the number of qubits the probability that a randomly sampled pair of Pauli operators commutes should approach 1 with increasing n. In this limit the compatibility graph will be closely approximated by a polynomially sized complete subgraph of the exponentially large compatibility graph of all Pauli operators on n qubits, with a few edges missing. Asymptotically we expect that the number of colors required will tend to the chromatic number of the complete graph, which is equal to the number of vertices.

As we we shall see in Sections III D and III E, for any fixed random k-local Hamiltonian we may write the probability that a randomly sampled pair of terms commute as:

$$p_c \simeq 1 - \frac{\text{const}}{n}. \tag{45}$$

The chromatic number of almost all such graphs will be proportional to n and hence we expect at most a constant factor reduction in the number of terms [49]. The problem of finding commuting cliques of related graphs was discussed in [25]. Here, we study the problem from the context of finding anticommuting sets for unitary partitioning.

D. Random 2-local Pauli Hamiltonians

Consider a 2-local Pauli Hamiltonian defined on an Erdős–Rényi random interaction graph with n vertices and $|E|$ edges. A term in the Hamiltonian corresponds to an edge in the set E, and a sample drawn uniformly at random from $\{X, Y, Z\}^{\otimes 2}$. We choose Hamiltonians with only one term per edge. Two terms corresponding to edges e_1 and e_2 from such a Hamiltonian anticommute if:

1. $e_1 \neq e_2$

2. $|e_1 \bigcap e_2| = 1$

What is the probability that $|e_1 \bigcap e_2| = 1$? There are $n - 2$ vertices connected to each vertex of e_1 that form edges with $|e_1 \bigcap e_2| = 1$. There are therefore $2(n - 2)$ of the $n(n - 1)/2$ possible edges that give $|e_1 \bigcap e_2| = 1$ for any given e_1. The probability of such an incidence is therefore $p_e = 4(n - 2)/(n(n - 1))$.

What is the probability that two terms intersect on one qubit and do not commute? There are nine operators that can be associated with an edge. Examination of this set gives a probability of 2/3 that tensor factors incident on the same vertex disagree.

Given a pair of edges from the interaction graph, i.e., a pair of terms in the Hamiltonian, the probability that the associated operators anticommute is therefore:

$$p_a = \frac{8}{3n} \frac{n - 2}{n - 1}. \tag{46}$$

We now analyze the coloring of an Erdős–Rényi random graph in which edges are populated independently with probability p [50]. As noted above, the compatibility graphs of random Pauli Hamiltonians cannot be Erdős–Rényi but in the limit of large numbers of qubits we expect these results to be asymptotically correct. Our procedure for defining a random two-local Pauli Hamiltonian has given us a probability $1 - p_a$ that an edge is present in the compatibility graph, because Pauli operators either commute or anticommute.

Almost every random graph with m vertices drawn from an ensemble where the probability of an edge between any pair of vertices is $1 - p_a$ has chromatic number [49]:

$$\chi = \left(\frac{1}{2} + o(1) \right) \log \frac{1}{p_a} \frac{m}{\log m}. \tag{47}$$

This immediately enables us to characterize the performance of our method on random 2-local Hamiltonians. Suppose the number of terms rises as a power τ of the number of qubits $m = O(n^\tau)$. Then the fractional improvement m_c/m in the number of terms in the Hamiltonian will be

$$\frac{m_c}{m} = \left(\frac{1}{2} + o(1) \right) \log \frac{3n(n - 1)}{8(n - 2)} \frac{1}{\tau \log n}. \tag{48}$$

This implies that we should expect a reduction in the number of the terms in the Hamiltonian by a constant factor of about 2τ.

E. Random k-local Hamiltonians

To choose a random interaction hypergraph of a k-local Hamiltonian we choose m independent k-tuples of qubit labels between 1 and n. We then uniformly randomly assign one of the 3^k Pauli operators of weight k to that k-tuple. Let S_1 and S_2 be two sets of k qubits. Given tuple S_1 there are

$$N_I = \binom{n - k}{k - I} \binom{k}{I} \tag{49}$$

tuples S_2 with $I \leq |S_1 \cup S_2|$, where $0 \leq I \leq k$. Summing over I recovers all k-tuples, by the Chu–Vandermonde identity. The probability of tuples S_1 and S_2 intersecting on I qubits is therefore:

$$p_I = \binom{n}{k}^{-1} \binom{n - k}{k - I} \binom{k}{I}. \tag{50}$$

Given that the tuples S_1 and S_2 intersect on I qubits, what is the probability that they commute? Let the Pauli

factors of S_1 and S_2 be identical on a subset of their intersection of size σ and otherwise every pair of tensor factors in the intersection disagree. The total number of pairs of Pauli operators on the intersection is 9^I. The number of Pauli operators identical on σ qubits is:

$$t_{I,\sigma} = 3^\sigma \binom{I}{\sigma} 3^{I-\sigma} 2^{I-\sigma} \tag{51}$$

which is obtained by multiplying the 3^σ Pauli operators common to the subset of σ qubits by the number of subsets of size σ and the number of distinct assignments to pairs of tensor factors in the complement of the subset of size σ. The total number of Pauli operators is then given by:

$$9^I = \sum_{\sigma=0}^{I} \frac{6^I}{2^\sigma} \binom{I}{\sigma}. \tag{52}$$

In order that a pair of operators commutes the size of the complement of the identical set must be even. That is:

$$p_c^{(I)} = \left(\frac{2}{3}\right)^I \sum_{I-\sigma \text{ even}} \frac{1}{2^\sigma} \binom{I}{\sigma} = \frac{1}{2}\left(1 + \frac{1}{3^I}\right). \tag{53}$$

The overall probability that a pair of tuples commute is therefore:

$$p_c = \sum_I p_I p_c^{(I)} = \sum_I \frac{p_I}{2}\left(1 + \frac{1}{3^I}\right). \tag{54}$$

For $k = 2$ we recover Eq. (46). For $k = 3$ we obtain:

$$p_c = 1 - \frac{1}{n(n-1)(n-2)}\left(3n^2 - 13n - \frac{134}{3}\right). \tag{55}$$

Higher values of k can be obtained from Eq. (54). The expression in Eq. (54) justifies the use of coloring bounds for Erdős–Rényi random graphs for large numbers of qubits when the expression Eq. (54) limits to Eq. (45).

IV. ELECTRONIC STRUCTURE HAMILTONIANS

Quantum chemistry simulations are expected to be an important use of variational quantum algorithms [51]. The goal is to find the eigenvalues and eigenvectors of the molecular electronic Hamiltonian

$$H = \sum_{p,q} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{p,q,r,s} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s, \tag{56}$$

where $a_p^\dagger$ and a_p are fermionic creation and annihilation operators acting on the space spanned by molecular spin orbitals χ_p. For computational purposes, this basis set is truncated to the first N orbitals. The fermionic operators satisfy the canonical anticommutation relations:

$$\{a_p^\dagger, a_q^\dagger\} = \{a_p, a_q\} = 0,$$
$$\{a_p, a_q^\dagger\} = \delta_{pq}\mathbb{1}. \tag{57}$$

The weights h_{pq} and h_{pqrs} are defined as

$$h_{pq} = \delta_{\sigma_p \sigma_q} \int d^3r\, \chi_p^*(\mathbf{r})\left(-\frac{\nabla^2}{2} - \sum_I \frac{\zeta_I}{|\mathbf{r} - \mathbf{R}_I|}\right)\chi_q(\mathbf{r}), \tag{58}$$

$$h_{pqrs} = \delta_{\sigma_p \sigma_s}\delta_{\sigma_q \sigma_r} \int d^3r_1 d^3r_2 \frac{\chi_p^*(\mathbf{r}_1)\chi_q^*(\mathbf{r}_2)\chi_r(\mathbf{r}_2)\chi_s(\mathbf{r}_1)}{|\mathbf{r}_1 - \mathbf{r}_2|}, \tag{59}$$

where $\mathbf{r}$ denotes the electronic spatial coordinates, $\sigma_p \in \{\uparrow, \downarrow\}$ is the spin value of the pth orbital, and $\{\mathbf{R}_I\}_I$ and $\{\zeta_I\}_I$ are the molecule's classical nuclear positions and their associated charges, respectively. These spatial integrals can be efficiently pre-computed on a classical computer. For use in a quantum algorithm, the Hamiltonian is then transformed to a weighted sum of Pauli strings using a fermion-to-qubit encoding, such as the Jordan–Wigner [52], Bravyi–Kitaev [53–55], or other similar [56] mappings. For the former two encodings, the number n of qubits is the same as the number N of molecular spin orbitals. The expectation value of each Pauli string is measured independently. The power of this approach stems from the ability to prepare ansatz states that cannot be efficiently constructed on a classical computer; these are typically derived from a unitary coupled-cluster ansatz [57–59]. This allows for efficient computation of high-precision eigenvalues, which has importance when considering calculations that require such precision, such as reaction kinetics and dynamics.

Implementation of this procedure for chemical systems at the desired accuracy is challenging. For chemistry, the required precision is typically considered to be a constant $1\,\text{kcal/mol}$, or 1.6 mHa. This level of precision is roughly commensurate with that obtained by experimental techniques in thermochemistry. Recall from Eq. (30) that the number of independent measurements that must be performed to estimate the expectation value of a single term with weight h to precision ϵ is $O(|h|/\epsilon^2)$. For chemical accuracy, this means that each term requires on the order of hundreds of thousands of independent measurements, each of which requires a separate ansatz preparation stage. This must be repeated for each step of the variational optimisation, for each of the $O(N^4)$ terms in the molecular Hamiltonian (noting that using the Jordan–Wigner transformation requires up to 16 Pauli strings for each term). As such, this quantum chemistry problem has recently garnered much interest with regards to reducing these VQE measurement costs [21, 23, 26–30]. The term reduction strategy discussed in Section II appears a promising way to reduce the overall resources required, by utilising available coherent computational resources subsequent to ansatz preparation.

In the absence of restrictions on the length of circuits that can be performed coherently, the term reduction strategy reduces the number of expectation values that must be independently estimated, going from the number of Hamiltonian terms to the number of fully anticommuting sets of terms. The main task is therefore to partition the Hamiltonian into such sets. The effectiveness of this term reduction strategy can be quantified by examining the number of fully anticommuting sets for a given Hamiltonian, with respect to both the number of orbitals and the total number terms in the unmodified Hamiltonian. In Section IV B, we show that it is always possible to reduce the number of terms from $O(N^4)$ to at most $O(N^3)$ for any electronic structure Hamiltonian. In Section IV C, we perform numerical studies using specific molecules and compare the results to our analytic construction. We also consider how the constraint of circuit size affects one's ability to construct such partitions. Finally, in Section IV D, we examine such Hamiltonians in the plane wave dual basis introduced in [21] and observe a constant factor reduction of terms by unitary partitioning.

A. Majorana operators

The approach we take here will be agnostic to the choice of qubit encoding. However, in order to partition the terms into completely anticommuting sets, it will be convenient to express them using Majorana operators. This is because they place all the fermionic operators on an equal footing, are Hermitian and unitary, and obey a single anticommutation relation. Here, we briefly review the properties of these operators essential for our analysis. The single-mode Majorana operators are defined from the fermionic modes as

$$
\begin{aligned}
\gamma_{2p} &= a_p + a_p^\dagger, \\
\gamma_{2p+1} &= -i(a_p - a_p^\dagger).
\end{aligned}
\tag{60}
$$

In this formalism, the anticommutation relations of Eq. (57) become

$$
\{\gamma_j, \gamma_k\} = 2\delta_{jk}\mathbb{1}.
\tag{61}
$$

These $2N$ single-mode operators generate a basis (up to phase) for the full algebra of Majorana operators via arbitrary products, i.e.

$$
\gamma_A = \prod_{j \in A} \gamma_j,
\tag{62}
$$

where $A \subseteq \{0, \ldots, 2N-1\}$ is the support of γ_A. From Eq. (61), it is straightforward to show that the anticommutator between two arbitrary Majorana operators γ_A and γ_B is determined by the parities of their individual supports and of their overlap:

$$
\{\gamma_A, \gamma_B\} = \left[1 + (-1)^{|A||B|+|A \cap B|}\right]\gamma_A\gamma_B.
\tag{63}
$$

This relation provides a clear picture of how to construct fully anticommuting sets of fermionic operators. Since the electronic Hamiltonian contains only terms of quadratic and quartic order, we restrict our attention to even parity products. In this setting, we only need to examine the overlap of the Majorana operators' supports: if $|A \cap B|$ is odd (i.e. the two operators share an odd number of single-mode indices), then they anticommute.

B. Linear reduction in terms

Since there are no spin interaction terms in our Hamiltonian, we can always choose molecular orbital basis functions χ_p which are real-valued. With this, it follows that $h_{pq}, h_{pqrs} \in \mathbb{R}$, and in particular, we have the permutational symmetries

$$
h_{pq} = h_{qp},
\tag{64}
$$

$$
h_{pqrs} = h_{sqrp} = h_{prqs} = h_{srqp}.
\tag{65}
$$

Furthermore, the canonical anticommutation relations give $a_p^\dagger a_q^\dagger a_r a_s = a_q^\dagger a_p^\dagger a_s a_r$, which implies that

$$
h_{pqrs} = h_{qpsr},
\tag{66}
$$

for a total of eight permutational symmetries in the two-body integrals. Using these symmetries and the generalized anticommutation relation Eq. (63), one can rewrite the Hamiltonian using Majorana operators as (see Appendix A 2 for details of the derivation)

$$
\begin{aligned}
H = \tilde{h}\mathbb{1} &+ \sum_{p,q} \tilde{h}_{pq} i\gamma_{2p}\gamma_{2q+1} \\
&+ \frac{1}{2} \sum_{\substack{p,q,r,s \\ p \neq q, r \neq s}} \tilde{h}_{pqrs} \gamma_{2p}\gamma_{2q}\gamma_{2r+1}\gamma_{2s+1}.
\end{aligned}
\tag{67}
$$

The redefined weights $\tilde{h}$, $\tilde{h}_{pq}$, and $\tilde{h}_{pqrs}$ are given by Eq. (A31) in the Appendix. For our present analysis, the only relevant detail here is that each term features an equal number of even and odd indices in its support, and in principle, any such combination may appear in the Hamiltonian. In this form, it becomes clear that there are up to N^2 quadratic terms and $\binom{N}{2}^2$ quartic terms.

Furthermore, since the single-mode Majorana operators are Hermitian, there is a one-to-one correspondence between Majorana operators and the respective Pauli strings obtained after a fermion-to-qubit transformation (for encodings that preserve the number of orbitals as the number of qubits). For instance, in the Jordan–Wigner encoding, we have

$$
\begin{aligned}
\gamma_{2p} &= X_p Z_{p-1} \cdots Z_0, \\
\gamma_{2p+1} &= Y_p Z_{p-1} \cdots Z_0.
\end{aligned}
\tag{68}
$$

Since the single-mode Majorana operators simply become Pauli strings, arbitrary products of them remain

single Pauli strings. In contrast, if one were to deal with the fermionic operators directly, a single $a_p^\dagger a_q^\dagger a_r a_s$ term would generate a linear combination of up to 16 unique Pauli strings. By writing the Hamiltonian in terms of Majorana operators, we have not circumvented this constant overhead, but rather have explicitly dealt with it, while remaining encoding agnostic. In particular, many cancellations and simplifications occur between those 16 terms, yielding the expression given above in Eq. (67). Also note that, since we aim to partition terms by anticommuting sets, such partitions in the Majorana formalism remain valid after any qubit transformation, since the anticommutation relations are preserved.

Recall from Eq. (63) that we had determined that every pair of terms anticommutes if and only if their supports intersect an odd number of times. This fact, along with the specific form of the terms appearing in Eq. (67), is crucial for showing that it is always possible to partition this Hamiltonian into at most $O(N^3)$ completely anticommuting sets.

We note that very recent results have made similar findings. In [30], it was observed that at least $\Omega(N^3)$ sets would be necessary to divide the set of *all* quartic Majorana operators, rather than the specific terms appearing in electronic structure Hamiltonians. Meanwhile, in [32], an algorithm was presented which partitions electronic structure terms into $O(N^3)$ completely *commuting* sets. The analysis presented there specifies the Jordan–Wigner encoding, but does not assume any of the permutational symmetries in the coefficients h_{pq}, h_{pqrs}.

We now prove our claim by providing an explicit construction of such a partition.

Theorem 1. *Let*

$$\mathcal{M} = \{\gamma_{2p}\gamma_{2q}\gamma_{2r+1}\gamma_{2s+1} \mid p < q \text{ and } r < s\} \qquad (69)$$

be the set of all possible quartic Majorana operators appearing in the electronic structure Hamiltonian. For each triple $(q, r, s) \in \{0, \ldots, N-1\}^3$ satisfying $r < s$, define

$$S_{(q,r,s)} = \{\gamma_{2p}\gamma_{2q}\gamma_{2r+1}\gamma_{2s+1} \mid p < q\}. \qquad (70)$$

These sets are completely anticommuting, and they partition $\mathcal{M}$. Furthermore, there are $O(N^3)$ such sets.

Proof. By construction, all elements of $S_{(q,r,s)}$ share support on exactly three indices, hence they all pairwise anticommute, per Eq. (63). It is also straightforward to see that these sets form an exact cover of $\mathcal{M}$:

$$\left|S_{(q,r,s)} \cap S_{(q',r',s')}\right| = q\,\delta_{qq'}\delta_{rr'}\delta_{ss'}, \qquad (71)$$

$$\bigcup_{\substack{q,r,s \\ r<s}} S_{(q,r,s)} = \mathcal{M}. \qquad (72)$$

By a simple counting argument, there are $\binom{N}{2}$ values that the pair (r, s) can take and $N-1$ values that q can take ($q = 0$ yields the empty set, which we ignore). A slight optimization arises from the observation that the union

$S_{(1,r,s)} \cup S_{(2,r,s)}$ remains a completely anticommuting set. Hence there are a total of $\binom{N}{2}(N-2) = O(N^3)$ such sets. $\qquad\square$

We refer the reader to Appendix B for further details of the above proof. Although there are only $O(N^2)$ quadratic terms, hence not affecting the asymptotic scaling of Theorem 1, they can in fact be included in the above construction with no additional overhead. Intuitively, since there at most N^2 such operators which need to be placed into $O(N^3)$ sets, one has a great deal of freedom in how to allocate them. As one example, consider the set

$$T_p = \{i\gamma_{2p}\gamma_{2q+1} \mid 0 \le q \le N-1\} \qquad (73)$$

for some fixed p. Then all the elements of T_p anticommute with all of some $S_{(p,r,s)}$, except for those with $q = r$ or $q = s$. The new completely anticommuting set then becomes

$$S_{(p,r,s)} \cup T_p \setminus \{i\gamma_{2p}\gamma_{2r+1}, i\gamma_{2p}\gamma_{2s+1}\}, \qquad (74)$$

and those two excluded operators can be placed with any other $S_{(p,r',s')}$, where all of r, r', s, and s' are different:

$$S_{(p,r',s')} \cup \{i\gamma_{2p}\gamma_{2r+1}, i\gamma_{2p}\gamma_{2s+1}\}. \qquad (75)$$

Since there are N such sets T_p, this procedure combines all possible N^2 quadratic operators with only $2N$ of the pre-existing sets of quartic operators.

We emphasize that the partition presented here is not an optimal solution to the problem. Rather, it demonstrates that, even in the worst-case, one can always achieve term reduction by at least a factor of $O(N)$. For a practical demonstration, we now move to numerical studies of specific molecular Hamiltonians.

C. Pauli-level colouring and numerics

The above analysis demonstrates a reduction in difficulty of VQE by considering the number of fully anticommuting sets of terms in the electronic Hamiltonian. Equivalently, we may consider fully anticommuting sets of terms at the level of Pauli strings, i.e. subsequent to transforming the electronic Hamiltonian with, for example, the Jordan–Wigner or Bravyi–Kitaev mappings. This approach could hold advantage by allowing the combination of duplicate strings, and allowing the combination of anticommuting Pauli subterms between different fermionic terms. However, once the fermion-to-qubit mapping is applied, the natural symmetries of the spatial molecular orbital integrals are embedded into a complex structure. Moreover, the anticommutativity structure of the resulting Pauli terms is difficult to predict. As such, we turn to numerical methods.

The key metric here is the number of fully anticommuting sets in the Pauli Hamiltonian. As discussed in Section II A, this is equivalent to a colouring of the compatibility graph—the graph composed of nodes corresponding to terms, with edges drawn where terms commute.

Optimal graph colouring is an NP-hard problem [60], but many approximate algorithms exist [61]. While minimising the number of sets is advantageous for reducing the number of measurements needed, an approximate solution is sufficient, and diminishing returns are obtained from improving the quality of the approximation.

In order to assess whether this strategy is viable for molecular Hamiltonians, we generated colouring schemes for 65 Hamiltonians (previously used in References [62, 63] and described in Appendix C). Geometry specifications were obtained from the NIST CCBDB database [64]. Molecular orbital integrals in the Hartree–Fock basis were gathered using the Psi4 package [65] and OpenFermion [66]. Our code was then used to generate Jordan–Wigner and Bravyi–Kitaev Hamiltonians, which were divided into anticommuting subsets using the NetworkX Python package [67] and the greedy independent sets strategy [61]. As our focus was on quantifying whether the term reduction technique is viable, alternative colouring strategies were not considered – such an analysis is performed in [26]. Our colouring strategy here is relatively computationally expensive, limiting our analysis to a maximum of 36 spin-orbitals, with only three systems involving 30 or more. While our code is unoptimised and can likely be improved upon, this does indicate that it would be difficult to extend this approach to larger systems. The Majorana-based scheme of Section IV B was also used to partition the Hamiltonians. In contrast to the greedy colouring strategy, this does not require extensive classical computational resources.

Figure 1 shows the results of this analysis. Both colouring schemes prove to be consistently effective. Beyond the smallest Hamiltonians, a roughly linear trend between the number of sets found and the number of Hamiltonian terms is observed, demonstrating that the asymptotic improvement discussed in Section IV B can be achieved when using numerical approaches to colouring Pauli Hamiltonians. The numerical Majorana results, and the greedy colouring strategy, consistently outperform the analytic upper bound. This is primarily due to sparsity in the h_{pq} and h_{pqrs} weights, due to molecular geometric symmetries and the locality of basis functions.

The ratio of the number of terms to the number of sets also appears to increase linearly with the number of spin-orbitals (albeit with high variance), in agreement with the scaling properties discussed in Section IV B. The number of fully anticommuting sets is approximately an order of magnitude less than the number of terms. The choice of Jordan–Wigner and Bravyi–Kitaev mapping does not appear to meaningfully affect the number of fully anticommuting sets found, as the anticommutativity structure is dependent on the underlying molecular Hamiltonian. Encouragingly, the agreement demonstrated here by Figure 1 suggests that the greedy independent set strategy is finding close-to-optimal colourings.

The greedy colouring scheme yields roughly a factor of 10 improvement over the numerical Majorana scheme,
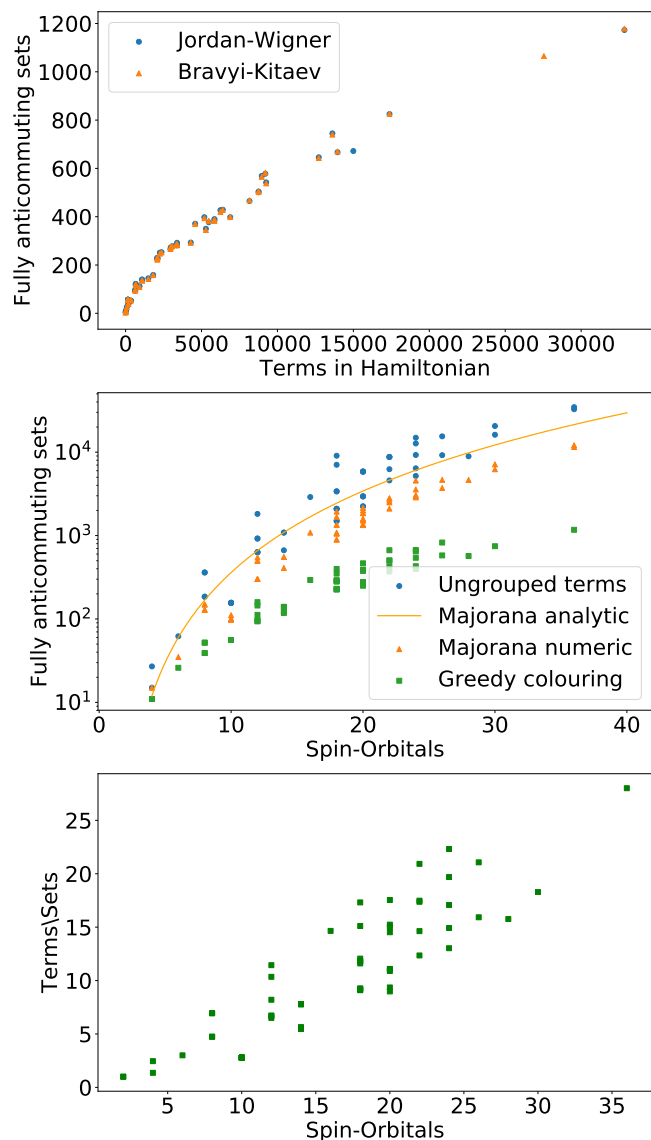

FIG. 1. Number of fully anticommuting sets for electronic structure Hamiltonians. Upper: Versus the number of terms in the full Pauli Hamiltonian, using the greedy independent sets strategy. The number of fully anticommuting sets is at least an order of magnitude less than the number of terms. The Jordan–Wigner and Bravyi–Kitaev mappings perform almost equivalently. Middle: Versus the number of spin orbitals (equivalent to that of qubits), on a logarithmic scale, using the Jordan–Wigner mapping. The "Majorana analytic" curve is the $\binom{N}{2}(N-2)$ upper bound obtained from Theorem 1 for generic Hamiltonians of Eq. (67). The "Majorana numeric" data points correspond to partitions of the same molecules under the strategy of Section IV B without further optimization. Notably, since many molecules feature additional symmetries that drastically reduce the number of inherent Hamiltonian terms (compared to the set of all possible terms), such an upper bound generally overestimates the number of sets achievable. Lower: Ratio of the number of terms to the number of anticommuting sets, with respect to the number of spin orbitals (for systems with more than 5 spin orbitals). The Jordan–Wigner mapping is used. A roughly linear trend is observed, in agreement with the analytic scaling discussed in Section IV B.

suggesting that it may be of substantial use in NISQ VQE experiments. However, it should be emphasised that the substantial classical computing resources required may inhibit its use for systems with larger numbers of spin-orbitals. The Majorana-based scheme demonstrates the same scaling for substantially reduced classical overhead.

Although these results are promising, they do not consider the difficulty of performing the additional coherent operations required for the term recombination procedure. In principle, our analytic construction of anticommuting sets in Section IV B requires only $O(N)$ depth circuits under the Jordan–Wigner mapping. This can be shown using well-known gate-compiling techniques [68, 69]. However, near-term quantum devices are likely to be heavily constrained in the number of operations that can be performed coherently. As such, it is likely that it will not be possible to combine entire sets of anticommuting terms. Crucially, however, the term recombination procedure can be applied to *subsets* of the fully anticommuting sets. Provided the available coherent resources can be quantified prior to execution of the circuits, subsets of terms can be found to maximally use such resources to reduce the overall number of measurements required. This yields a hardware-dependent tunable parameter—for example, the number of gates that can be implemented coherently subsequent to ansatz preparation—introduced at compile-time. This parameter allows for optimal use of the quantum resources provided by a given hardware option.

In order to assess the implications of varying such a parameter, we generated circuits corresponding to the implementation of the term reduction procedure for each Hamiltonian, introducing a maximum post-ansatz preparation gate count parameter. For simplicity, these circuits used the standard method of implementing exponentiated Pauli strings given in Section II A, rather than the ALCU circuits of Section II B. Where circuits exceeded this length, the corresponding anticommuting set was split in half and new circuits were generated. This binary splitting process was iterated until sufficiently short circuits were found. Adjacent self-inverse gates were cancelled, moving through commuting gates where necessary [69]. For verification purposes, we calculated the expectation values with the true ground state of the Hamiltonians predicted by the circuits for systems with less than 10 qubits. As the results presented in Figure 1 suggest that there is little difference between Jordan–Wigner and Bravyi–Kitaev circuits, we consider only Jordan–Wigner circuits.

Figure 2 shows the results of this process. Using a maximum circuit length of 10000 gates subsequent to ansatz preparation allows all anticommuting sets, in all Hamiltonians, to be combined. Allowing only 10 gates removes any possibility of term recombination. Encouragingly, allowing 100 gates does not dramatically impede term recombination. Even for the longest circuit considered, using 100 gates allows for a reduction in terms by a factor of over 2. Allowing 1000 post-ansatz gates similarly per-
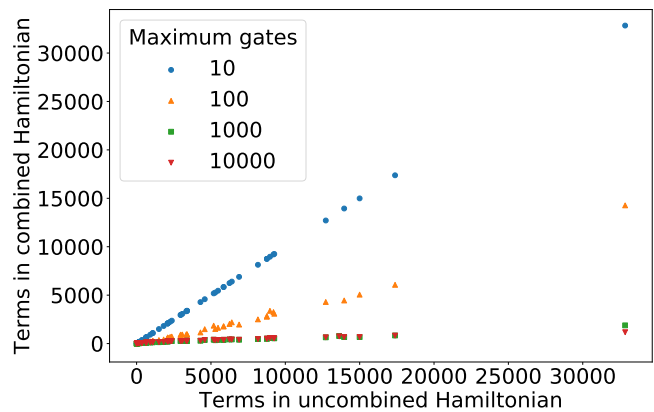

FIG. 2. Reduction of number of required independent expectation values given restrictions on maximum individual circuit length. With highly restricted circuit length, term recombination is largely impossible. However, 1000 post-ansatz preparation gates allows for the full anticommuting sets to be combined in all bar the single longest circuit. Encouragingly, reducing the number of allowed gates by an order of magnitude still allows for dramatic term reduction. Failed trials (indicated in Appendix C) are not plotted.

forms as well as full anticommuting set recombination, in all systems apart from the bromine atom; in this instance, the difference between the 1000 and 10000 gate decompositions is minor.

Our choice of allowable circuit length here is intended to be illustrative of the practicality of the term recombination procedure. In a true simulation, the maximum post-ansatz gates parameter should be set to a value that is empirically determined by the ability of the hardware, and should not be restricted to an integer power of ten. Given the relatively low gate counts required for substantial improvement with regards to the number of terms, the results here strongly suggest that this approach is an effective way of reducing the overall runtime of variational quantum algorithms for electronic structure theory.

D. Electronic structure Hamiltonians in the plane wave dual basis

The use of plane wave bases is well established for condensed matter systems. The plane wave and plane wave dual basis was recently used in the context of quantum simulation of quantum chemistry to define a Hamiltonian with a number of terms scaling quadratically with the number of basis functions [21]. Note that, while suitable for periodic systems, the plane wave dual basis requires a constant factor of additional spin orbitals to achieve the same accuracy as Gaussian-type orbitals for non-periodic systems such as molecules. Thus the choice of basis to use depends highly on the system under consideration, especially for near-term applications.

The qubit Hamiltonian obtained from the Jordan–Wigner transformation is (see Eq. (9) in [21]):

$$
\begin{aligned}
H = &\sum_{\substack{p,\sigma \\ \nu \neq 0}} \left(\frac{\pi}{\Omega\, k_\nu^2} - \frac{k_\nu^2}{4\,N} + \frac{2\pi}{\Omega} \sum_j \zeta_j \frac{\cos\left[k_\nu\cdot(R_j - r_p)\right]}{k_\nu^2} \right) Z_{p,\sigma} \\
&+ \frac{\pi}{2\,\Omega} \sum_{\substack{(p,\sigma)\neq(q,\sigma') \\ \nu \neq 0}} \frac{\cos\left[k_\nu\cdot r_{p-q}\right]}{k_\nu^2} Z_{p,\sigma} Z_{q,\sigma'} \\
&+ \frac{1}{4\,N} \sum_{\substack{p\neq q \\ \nu,\sigma}} k_\nu^2 \cos\left[k_\nu\cdot r_{q-p}\right] X_{p,\sigma} Z_{p+1,\sigma}\cdots Z_{q-1,\sigma} X_{q,\sigma} \\
&+ \frac{1}{4\,N} \sum_{\substack{p\neq q \\ \nu,\sigma}} k_\nu^2 \cos\left[k_\nu\cdot r_{q-p}\right] Y_{p,\sigma} Z_{p+1,\sigma}\cdots Z_{q-1,\sigma} Y_{q,\sigma} \\
&+ \sum_{\nu\neq 0} \left(\frac{k_\nu^2}{2} - \frac{\pi\,N}{\Omega\,k_\nu^2} \right) \mathbb{1}.
\end{aligned}
\tag{76}
$$

The labels p run over N basis functions and so by inspection we see that the number of terms in the Hamiltonian is $O(N^2)$. Again by inspection we can identify a set of N^2 commuting operators $Z_{p,\sigma} Z_{q,\sigma'}$. Thus, we can immediately conclude that unitary partitioning cannot reduce the asymptotic number of terms in this Hamiltonian.

However, we may use unitary partitioning to reduce the number of terms by a constant factor. We can identify sets of anticommuting terms from Eq. (76) as follows. Define the sets:

$$
\begin{aligned}
A_p = &\{Z_p\} \cup \{X_{p-1}X_p\} \cup \{Y_p Y_{p+1}\} \\
&\cup \{Y_p Z_{[p+1,p+l+1]} Y_{p+l+2} \mid 0 \leq l \leq N-p-3\} \\
&\cup \{X_l Z_{[l+1,p-1]} X_p \mid 0 \leq l \leq p-2\}.
\end{aligned}
\tag{77}
$$

There are N operators in each set A_p, all of which pairwise anticommute. All sets A_p are distinct and so unitary partitioning can reduce each set A_p to a single term. This results in a fractional reduction in the number of terms of $(2N+1)/(4N-1)$ giving an asymptotic reduction in the number of terms by a factor of two.

V. NONCONTEXTUAL HAMILTONIANS

In [41], contextuality of a qubit Hamiltonian is defined as the condition under which it is impossible to consistently assign values to the Pauli terms in the Hamiltonian. Contextuality, if present, is a manifestation of nonclassicality of the Hamiltonian. In [41], contextuality of a Hamiltonian is shown to be determined by the following criterion on the set $\mathcal{S}$ of Pauli terms. First, obtain the largest subset $\mathcal{T} \subseteq \mathcal{S}$ such that every operator in $\mathcal{T}$ anticommutes with some other operator in $\mathcal{S}$ (in other words, $\mathcal{S} \setminus \mathcal{T}$ is the subset of $\mathcal{S}$ that commutes with all operators in $\mathcal{S}$.) Then $\mathcal{S}$ is noncontextual if and only if commutation is an equivalence relation on $\mathcal{T}$. In other

words, if and only if $\mathcal{S}$ is noncontextual, $\mathcal{T}$ partitions into a union of disjoint cliques $C_1, C_2, \ldots, C_N$ such that operators in different cliques anticommute, while operators in the same clique commute (so in the graph-theoretic sense these are cliques in the compatibility graph).

We now show that, using the term reduction technique presented above, we can map any noncontextual Hamiltonian to a commuting Hamiltonian. First, as shown in [41], we can check that the Hamiltonian is noncontextual in $O(|\mathcal{S}|^3)$ time. Given that the Hamiltonian is noncontextual, we know that it has the structure described above: we can find the cliques C_i as well as the fully-commuting set, which we may denote $\mathcal{T}'$, in $O(|\mathcal{S}|^2)$ time.

To map these terms to a commuting set, find a largest clique, and without loss of generality let it be C_1. Then construct a set D_1 by selecting exactly one element from each of the C_i (D_1 is a minimal hitting set on the C_i). Similarly, construct D_2 by selecting exactly one element from each of the C_i after removing the elements in D_1, and so forth, until we have covered $\mathcal{T}$ with disjoint sets $D_1, D_2, \ldots, D_M$, where $M = |C_1|$ Letting C_{ij} denote the jth element of C_i, we may visualize D_j as

$$
\begin{array}{c|cccc}
 & D_1 & D_2 & \cdots & D_M \\
\hline
C_1 & C_{11} & C_{12} & \cdots & C_{1M} \\
C_2 & C_{21} & C_{22} & \cdots & \cdots \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
C_N & C_{N1} & \cdots & \cdots & \cdots
\end{array}
\tag{78}
$$

Since C_1 is a largest clique, it is the only clique that is guaranteed to have nonempty intersection with all of the sets D_j. Elements of different cliques anticommute, so each of the D_j is completely anticommuting. Therefore, we can use the techniques described in Sections II A and II B to construct Pauli rotations R_{D_j} that map the operators in each D_j to the operator C_{1j}, the shared operator in D_j and C_1. The resulting Hamiltonian has terms $\mathcal{T}' \cup C_1$, which commute, since by definition the operators in $\mathcal{T}'$ commute with all operators in $\mathcal{S}$, and the operators in C_1 also commute with each other. Thus any noncontextual Hamiltonian may be mapped to a commuting Hamiltonian, using as a resource only the ability to append the additional Pauli rotations R_{D_j} to the state preparations as in Eq. (21).

VI. CONCLUSIONS

In this paper we have discussed unitary partitioning—the technique for using anticommuting sets of Hamiltonian terms to reduce the number of measurements needed when performing variational quantum algorithms. Applying this technique to transverse Ising models and random Hamiltonians resulted in a constant factor improvement in the number of independent expectation value estimations required. However, applying the technique to electronic structure Hamiltonians yielded greater benefit.

The dramatic growth in the number of independent expectation values which must be determined is a key problem in the use of variational quantum algorithms for quantum chemistry in the NISQ era. Due to the nature of the plane wave dual basis representation, in which one defines a basis which yields only $O(N^2)$ non-zero electronic Hamiltonian weights, we observed only a constant factor reduction in terms with unitary partitioning. However, using generic molecular orbital basis sets, we were able to obtain a reduction which scales linearly. We proved this result in Section IV B and confirmed its practicality by numerics in Section IV C.

We report two strategies for partitioning the electronic structure Hamiltonian into fully anticommuting subsets. The first of these, based on expressing the fermionic Hamiltonian using Majorana operators, demonstrates the favourable scaling properties, and can be rapidly performed for even large numbers of spin orbitals. Conversely, using a greedy colouring scheme is relatively expensive with regards to classical computational resources, but demonstrates an order of magnitude reduction in the number of independent expectation values that must be determined, even for relatively small systems (less than 30 qubits). The latter of these schemes is likely to be useful in NISQ applications where systems are small and greedy solutions can be feasibly classically computed. The former yields the same scaling, and is not restricted to this regime, but suffers from a constant factor increase in the number of fully anticommuting sets compared to the greedy colouring. The availability of post-ansatz coherent resources, and the relative difficulty of the classical partitioning step, may determine which scheme is favoured.

Finally, in Section V we studied the class of noncontextual Hamiltonians, as defined in [41]. The presence of contextuality in a quantum system provides a barrier to a classical description of the system. Here, we have shown that any noncontextual Hamiltonian (which lacks this separation from a classical Hamiltonian) may be transformed into a Hamiltonian of fully commuting terms, using only the rotations developed in Section II A. This result helps us further understand the connection between noncontextual Hamiltonians and commuting Hamiltonians, and it adds support to the notion that VQE experiments should focus on contextual Hamiltonians [41].

Our analysis of circuits for implementing the unitary partitioning procedure indicates that relatively modest additional coherent resources are required, compared to those typically needed for ansatz preparation. Crucially, this optimisation is tunable, allowing for optimal use of coherent resources by hardware-dependent parameterisation at compile-time. It is also likely that unitary partitioning is compatible with other aspects of VQE optimisation. For instance, while we have remained agnostic to the choice of the parametrised ansatz for this study, the form of the unitaries required to perform term reduction matches those of popular ansatz choices, such as the unitary coupled cluster and related methods [58, 59, 70]. Thus with proper circuit compilation, one may significantly reduce the effective number of post-ansatz operations in practice, instead incorporating their rotation angles into the appropriate ansatz parameters. For these reasons, we believe that unitary partitioning could substantially aid in the use of variational quantum algorithms for studying classically intractable systems.

[1] R. P. Feynman, International Journal of Theoretical Physics **21**, 467 (1982).

[2] S. Lloyd, Science , 1073 (1996).

[3] D. S. Abrams and S. Lloyd, Physical Review Letters **79**, 2586 (1997).

[4] D. S. Abrams and S. Lloyd, Physical Review Letters **83**, 5162 (1999).

[5] L.-A. Wu, M. Byrd, and D. Lidar, Physical Review Letters **89**, 057904 (2002).

[6] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, Science **309**, 1704 (2005).

[7] S. P. Jordan, K. S. M. Lee, and J. Preskill, Science **336**, 1130 (2012), arXiv:1111.3633 [quant-ph].

[8] R. Babbush, C. Gidney, D. W. Berry, N. Wiebe, J. McClean, A. Paler, A. Fowler, and H. Neven, Phys. Rev. X **8**, 041015 (2018).

[9] J. Preskill, Quantum **2**, 79 (2018).

[10] F. Arute *et al.*, Nature **574**, 505 (2019).

[11] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, J. M. Martinis, and H. Neven, arXiv preprint arXiv:1608.00263 (2016).

[12] A. W. Harrow and A. Montanaro, Nature **549**, 203 (2017).

[13] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'brien, Nature communications **5**, 4213 (2014).

[14] S. Barrett, K. Hammerer, S. Harrison, T. E. Northup, and T. J. Osborne, Physical review letters **110**, 090501 (2013).

[15] E. Farhi, J. Goldstone, and S. Gutmann, arXiv preprint arXiv:1411.4028 (2014).

[16] Y. Wang, F. Dolde, J. Biamonte, R. Babbush, V. Bergholm, S. Yang, I. Jakobi, P. Neumann, A. Aspuru-Guzik, J. D. Whitfield, and J. Wrachtrup, ACS Nano **9**, 7769 (2015).

[17] P. J. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly,

P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, and J. M. Martinis, Phys. Rev. X **6**, 031007 (2016).

[18] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Nature **549**, 242 (2017).

[19] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, and C. F. Roos, Phys. Rev. X **8**, 031022 (2018).

[20] E. F. Dumitrescu, A. J. McCaskey, G. Hagen, G. R. Jansen, T. D. Morris, T. Papenbrock, R. C. Pooser, D. J. Dean, and P. Lougovski, Phys. Rev. Lett. **120**, 210501 (2018).

[21] R. Babbush, N. Wiebe, J. McClean, J. McClain, H. Neven, and G. K.-L. Chan, Physical Review X **8**, 011044 (2018).

[22] N. C. Rubin, R. Babbush, and J. McClean, New Journal of Physics **20**, 053020 (2018).

[23] D. Wang, O. Higgott, and S. Brierley, Physical Review Letters **122**, 140504 (2019).

[24] V. Verteletskyi, T.-C. Yen, and A. F. Izmaylov, arXiv preprint arXiv:1907.03358 (2019).

[25] A. Jena, S. Genin, and M. Mosca, arXiv preprint arXiv:1907.07859 (2019).

[26] A. F. Izmaylov, T.-C. Yen, R. A. Lang, and V. Verteletskyi, arXiv preprint (2019), arXiv:1907.09040 [quant-ph].

[27] T.-C. Yen, V. Verteletsky, and A. F. Izmaylov, arXiv preprint arXiv:1907.09386 (2019).

[28] W. J. Huggins, J. McClean, N. Rubin, Z. Jiang, N. Wiebe, K. B. Whaley, and R. Babbush, arXiv preprint arXiv:1907.13117 (2019).

[29] P. Gokhale, O. Angiuli, Y. Ding, K. Gui, T. Tomesh, M. Suchara, M. Martonosi, and F. T. Chong, arXiv preprint arXiv:1907.13623 (2019).

[30] X. Bonet-Monroig, R. Babbush, and T. E. O'Brien, arXiv preprint arXiv:1908.05628 (2019).

[31] O. Crawford, B. van Straaten, D. Wang, T. Parks, E. Campbell, and S. Brierley, arXiv preprint arXiv:1908.06942 (2019).

[32] P. Gokhale and F. T. Chong, arXiv preprint arXiv:1908.11857 (2019).

[33] G. Torlai, G. Mazzola, G. Carleo, and A. Mezzacapo, arXiv preprint arXiv:1910.07596 (2019).

[34] S. B. Bravyi and A. Y. Kitaev, Annals of Physics **298**, 210 (2002).

[35] J. T. Seeley, M. J. Richard, and P. J. Love, Journal of Chemical Physics **137**, 224109 (2012).

[36] R. Somma, G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, Phys. Rev. A **65**, 042323 (2002).

[37] M. A. Nielsen and I. Chuang, "Quantum computation and quantum information," (2002).

[38] L. Kocia and P. Love, Phys. Rev. A **96**, 062134 (2017).

[39] A. M. Childs and N. Wiebe, Quantum Info. Comput. **12**, 901 (2012).

[40] R. Babbush, D. W. Berry, and H. Neven, Physical Review A **99**, 040301(R) (2019).

[41] W. M. Kirby and P. J. Love, Phys. Rev. Lett. **123**, 200501 (2019).

[42] R. Raussendorf, J. Bermejo-Vega, E. Tyhurst, C. Okay, and M. Zurel, arXiv preprint arXiv:1905.05374 (2019).

[43] G. H. Low and I. L. Chuang, Quantum **3**, 163 (2019).

[44] G. H. Low and I. L. Chuang, Phys. Rev. Lett. **118**, 010501 (2017).

[45] D. Poulin, A. Kitaev, D. S. Steiger, M. B. Hastings, and M. Troyer, Phys. Rev. Lett. **121**, 010501 (2018).

[46] R. Babbush, C. Gidney, D. W. Berry, N. Wiebe, J. McClean, A. Paler, A. Fowler, and H. Neven, Phys. Rev. X **8**, 041015 (2018).

[47] D. Wecker, M. B. Hastings, and M. Troyer, Phys. Rev. A **92**, 042303 (2015).

[48] M. Planat and M. Saniga, Quantum Information and Computation **8**, 0127 (2008).

[49] B. Bollobás, Combinatorica **8**, 49 (1988).

[50] P. Erdős and A. Rényi, Publ. Math. Inst. Hungar. Acad. Sci **5**, 17 (1960).

[51] J. Olson, Y. Cao, J. Romero, P. Johnson, P.-L. Dallaire-Demers, N. Sawaya, P. Narang, I. Kivlichan, M. Wasielewski, and A. Aspuru-Guzik, arXiv preprint arXiv:1706.05413 (2017).

[52] P. Jordan and E. Wigner, Z. Phys. **47**, 631 (1928).

[53] S. B. Bravyi and A. Y. Kitaev, Annals of Physics **298**, 210 (2002).

[54] J. T. Seeley, M. J. Richard, and P. J. Love, The Journal of Chemical Physics **137**, 224109 (2012).

[55] A. Tranter, S. Sofia, J. Seeley, M. Kaicher, J. Mcclean, R. Babbush, P. Coveney, F. Mintert, F. Wilhelm, and P. Love, International Journal of Quantum Chemistry **115** (2015), 10.1002/qua.24969.

[56] K. Setia and J. D. Whitfield, The Journal of Chemical Physics **148**, 164104 (2018).

[57] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, New Journal of Physics **18**, 023023 (2016).

[58] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik, Quantum Science and Technology **4**, 014008 (2018).

[59] J. Lee, W. J. Huggins, M. Head-Gordon, and K. B. Whaley, Journal of Chemical Theory and Computation **15**, 311 (2019).

[60] M. R. Garey, D. S. Johnson, and L. Stockmeyer, in *Proceedings of the Sixth Annual ACM Symposium on Theory of Computing*, STOC '74 (ACM, New York, NY, USA, 1974) pp. 47–63.

[61] A. Kosowski and K. Manuszewski, in *Contemporary Mathematics*, Vol. 352, edited by M. Kubale (American Mathematical Society, Providence, Rhode Island, 2004) pp. 1–19.

[62] A. Tranter, P. J. Love, F. Mintert, and P. V. Coveney, Journal of Chemical Theory and Computation **14**, 5617 (2018).

[63] A. Tranter, P. J. Love, F. Mintert, N. Wiebe, and P. V. Coveney, Entropy **21**, 1218 (2019).

[64] R. D. Johnson III, *NIST Computational Chemistry Comparison and Benchmark Database NIST Standard Reference Database Number 101 Release 18* (2016).

[65] R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, J. F. Gonthier, A. M. James, H. R. McAlexander, A. Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill, Journal of Chemical Theory and Computation

13, 3185 (2017).

[66] J. R. McClean, I. D. Kivlichan, K. J. Sung, D. S. Steiger, Y. Cao, C. Dai, E. S. Fried, C. Gidney, B. Gimby, T. Häner, T. Hardikar, V. Havlíček, C. Huang, Z. Jiang, M. Neeley, T. O'Brien, I. Ozfidan, M. D. Radin, J. Romero, N. Rubin, N. P. D. Sawaya, K. Setia, S. Sim, M. Steudtner, W. Sun, F. Zhang, and R. Babbush, arXiv:1710.07629 [physics, physics:quant-ph] (2017), arXiv:1710.07629.

[67] A. A. Hagberg, D. A. Schult, and P. J. Swart, in *Proceedings of the 7th Python in Science Conference*, edited by G. Varoquaux, T. Vaught, and J. Millman (Pasadena, CA USA, 2008) pp. 11 – 15.

[68] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, Molecular Physics **109**, 735 (2011), https://doi.org/10.1080/00268976.2011.552441.

[69] M. B. Hastings, D. Wecker, B. Bauer, and M. Troyer, Quantum Information & Computation **15**, 1 (2015).

[70] I. G. Ryabinkin, T.-C. Yen, S. N. Genin, and A. F. Izmaylov, Journal of chemical theory and computation **14**, 6317 (2018).

Appendix A: Calculational details

In this section we give some derivations of the algebraic results used in the text.

1. Computation of $\mathcal{X}$ for the ALCU method

We now derive the results that follow Eq. (22). The operator $\mathcal{X}$ is given by

$$
\begin{aligned}
\mathcal{X} &= \frac{i}{2}\left[H_{n-1}, P_n\right] \\
&= \frac{i}{2}\sum_{k=1}^{n-1}\beta_k\left[P_k, P_n\right] \\
&= i\sum_{k=1}^{n-1}\beta_k P_k P_n,
\end{aligned}
\tag{A1}
$$

where we wrote $H_{n-1} = \sum_{k=1}^{n-1}\beta_k P_k$ with $\sum_{k=1}^{n-1}\beta_k^2 = 1$. Then we can compute:

$$
\begin{aligned}
\mathcal{X}^2 &= -\sum_{k=1}^{n-1}\sum_{j=1}^{n-1}\beta_k\beta_j P_k P_n P_j P_n \\
&= -\sum_{k=1}^{n-1}\beta_k^2 P_k P_n P_k P_n - \sum_{k<j}^{n-1}\beta_k\beta_j\{P_k P_n, P_j P_n\} \\
&= -\sum_{k=1}^{n-1}\beta_k^2 P_k P_n P_k P_n \\
&= \sum_{k=1}^{n-1}\beta_k^2 P_k^2 P_n^2 \\
&= \mathbb{1}.
\end{aligned}
\tag{A2}
$$

Now consider the commutator of $\mathcal{X}$ and H_n. We can use $\mathcal{X} = iH_{n-1}P_n$ to write

$$
\begin{aligned}
\mathcal{X}H_n &= iH_{n-1}P_n H_n \\
&= iH_{n-1}P_n(\sin\phi_{n-1}H_{n-1} + \cos\phi_{n-1}P_n) \\
&= i(\sin\phi_{n-1}H_{n-1}P_n H_{n-1} + \cos\phi_{n-1}H_{n-1}P_n^2).
\end{aligned}
\tag{A3}
$$

Using $\{H_{n-1}, P_n\} = 0$ and $P_n^2 = 1$ we have

$$
\mathcal{X}H_n = i(-\sin\phi_{n-1}P_n + \cos\phi_{n-1}H_{n-1}),
\tag{A4}
$$

so that

$$
[\mathcal{X}, H_n] = 2i(-\sin\phi_{n-1}P_n + \cos\phi_{n-1}H_{n-1}).
\tag{A5}
$$

This enables us to compute the adjoint action generated by $\mathcal{X}$ on H_n. Using the identity (for any operators A and B, where $A^2 = \mathbb{1}$ so that $e^{-i(\alpha/2)A} = \cos(\alpha/2)\mathbb{1} - i\sin(\alpha/2)A$)

$$
\begin{aligned}
e^{-i(\alpha/2)A}Be^{i(\alpha/2)A} &= \cos^2(\alpha/2)B + \sin^2(\alpha/2)ABA \\
&\quad + i\sin(\alpha/2)\cos(\alpha/2)[A, B],
\end{aligned}
\tag{A6}
$$

we have ($R = e^{-i(\alpha/2)\mathcal{X}}$):

$$
\begin{aligned}
RH_nR^\dagger &= \cos^2(\alpha/2)H_n + \sin^2(\alpha/2)\mathcal{X}H_n\mathcal{X} \\
&\quad + i\sin(\alpha/2)\cos(\alpha/2)[\mathcal{X}, H_n] \\
&= (\cos^2\alpha/2 - \sin^2\alpha/2)H_n \\
&\quad + (i/2)2\sin(\alpha/2)\cos(\alpha/2)[\mathcal{X}, H_n] \\
&= \cos\alpha H_n \\
&\quad - \sin\alpha(-\sin\phi_{n-1}P_n + \cos\phi_{n-1}H_{n-1}) \\
&= \cos\alpha(\cos\phi_{n-1}P_n + \sin\phi_{n-1}H_{n-1}) \\
&\quad - \sin\alpha(-\sin\phi_{n-1}P_n + \cos\phi_{n-1}H_{n-1}) \\
&= (\cos\alpha\cos\phi_{n-1} + \sin\alpha\sin\phi_{n-1})P_n \\
&\quad + (\cos\alpha\sin\phi_{n-1} - \sin\alpha\cos\phi_{n-1})H_{n-1} \\
&= \cos(\phi_{n-1} - \alpha)P_n \\
&\quad + \sin(\phi_{n-1} - \alpha)H_{n-1}.
\end{aligned}
\tag{A7}
$$

Choosing $\alpha = \phi_{n-1}$ gives $RH_nR^\dagger = P_n$. Given this role for R, which is generated by $\mathcal{X}$, we wish to know the commutation relations amongst the terms of $\mathcal{X}$. Because $\mathcal{X} = 2iP_n H_{n-1}$, the terms of $\mathcal{X}$ have the form $2iP_n P_j$ for $j < n$. The commutation relations between any pair of terms are:

$$
\begin{aligned}
[P_n P_j, P_n P_k] &= P_n P_j P_n P_k - P_n P_k P_n P_j \\
&= -(P_n P_n P_j P_k - P_n P_n P_k P_j) \\
&= -[P_j, P_k] \\
&= 2P_k P_j.
\end{aligned}
\tag{A8}
$$

2. Electronic structure Hamiltonian using Majorana operators

Here, we derive the form of the Hamiltonian given in Eq. (67). Since the single-mode Majorana operators are

linear combinations of the fermionic ladder operators, we have the identities

$$a_p = \frac{\gamma_{2p} + i\gamma_{2p+1}}{2}, \ a_p^\dagger = \frac{\gamma_{2p} - i\gamma_{2p+1}}{2}. \qquad (A9)$$

Furthermore, recall the permutational symmetries in the coefficients, given by Eqs. (64) to (66), and the anti-commutation relation for arbitrary Majorana operators, Eq. (63). These are the only properties we use, but they allow for considerable simplification to the structure of the Hamiltonian terms. For brevity, we shall make use of such properties freely and often without comment.

First, consider the one-body terms, which are quadratic in fermionic operators. Using Majorana operators, they become

$$\sum_{p,q} h_{pq} a_p^\dagger a_q = \frac{1}{4} \sum_{p,q} h_{pq} (\gamma_{2p}\gamma_{2q} + \gamma_{2p+1}\gamma_{2q+1} \qquad (A10)$$
$$+ i\gamma_{2p}\gamma_{2q+1} - i\gamma_{2p+1}\gamma_{2q}).$$

This expression can be simplified by separating the summation into diagonal and off-diagonal terms, a technique which we employ heavily throughout this derivation. The sum over the $\gamma_{2p}\gamma_{2q}$ and $\gamma_{2p+1}\gamma_{2q+1}$ terms simply yields a multiple of the identity:

$$\sum_{p,q} h_{pq} (\gamma_{2p}\gamma_{2q} + \gamma_{2p+1}\gamma_{2q+1})$$
$$= \sum_p h_{pp} \left(\gamma_{2p}^2 + \gamma_{2p+1}^2\right)$$
$$+ \sum_{\substack{p,q \\ p<q}} h_{pq} \left(\{\gamma_{2p}, \gamma_{2q}\} + \{\gamma_{2p+1}, \gamma_{2q+1}\}\right)$$
$$= 2\sum_p h_{pp} \mathbb{1}. \qquad (A11)$$

The remaining terms simplify but do not cancel or reduce in order: by relabeling the indices (another trick which we make frequent use of), we see that $\sum_{p,q} h_{pq} i\gamma_{2p}\gamma_{2q+1} = \sum_{p,q} h_{pq} i\gamma_{2q}\gamma_{2p+1}$, hence

$$\sum_{p,q} h_{pq} a_p^\dagger a_q = \frac{1}{2} \left(\sum_p h_{pp} \mathbb{1} + \sum_{p,q} h_{pq} i\gamma_{2p}\gamma_{2q+1} \right). \qquad (A12)$$

Next, we consider the two-body interaction terms, which feature the quartic order operators. Any such term is written as a linear combination of 16 Majorana operators. To do so, define

$$\Gamma_{pqrs}^{\mathbf{x}} = i^{|\mathbf{x}|}(-1)^{x_1+x_2} \gamma_{2p+x_1}\gamma_{2q+x_2}\gamma_{2r+x_3}\gamma_{2s+x_4}, \quad (A13)$$

where $\mathbf{x} = x_1 x_2 x_3 x_4 \in \{0,1\}^4$ is a binary string encoding the parity of each index and $|\mathbf{x}|$ is its Hamming weight. Then, from Eq. (A9), a straightforward algebraic expansion gives the following expression for each two-body term:

$$a_p^\dagger a_q^\dagger a_r a_s = \frac{1}{16} \sum_{\mathbf{x} \in \{0,1\}^4} \Gamma_{pqrs}^{\mathbf{x}}. \qquad (A14)$$

Consider the set $B_1 = \{0011, 1100, 0101, 1010\}$. These strings correspond to the quartic Majorana operators appearing in Eq. (67), and as we will see, they are the only such terms which do not vanish. Also, note that since $a_j^2 = (a_j^\dagger)^2 = 0$, we impose the trivial constraints in the summations that $p \neq q$ and $r \neq s$. Specifying these conditions explicitly will be useful once we relabel the indices. We rewrite these terms as

$$\sum_{p,q,r,s} h_{pqrs} \Gamma_{pqrs}^{1100} = - \sum_{\substack{p,q,r,s \\ p\neq q; r\neq s}} h_{pqrs} \gamma_{2p+1}\gamma_{2q+1}\gamma_{2r}\gamma_{2s}$$
$$= - \sum_{\substack{p,q,r,s \\ p\neq q; r\neq s}} h_{pqrs} \gamma_{2r}\gamma_{2s}\gamma_{2p+1}\gamma_{2q+1}$$
$$= - \sum_{\substack{p,q,r,s \\ p\neq q; r\neq s}} h_{pqrs} \gamma_{2p}\gamma_{2q}\gamma_{2r+1}\gamma_{2s+1},$$
$$(A15)$$

and, for $x, y \in \{0,1\}$ such that $x \neq y$,

$$\sum_{p,q,r,s} h_{pqrs} \Gamma_{pqrs}^{xyxy} = \sum_{\substack{p,q,r,s \\ p\neq q; r\neq s}} h_{pqrs} \gamma_{2p+x}\gamma_{2q+y}\gamma_{2r+x}\gamma_{2s+y}$$
$$= - \sum_{\substack{p,q,r,s \\ p\neq q; r\neq s}} h_{pqrs} \gamma_{2p+x}\gamma_{2r+x}\gamma_{2q+y}\gamma_{2s+y}$$
$$= - \sum_{\substack{p,q,r,s \\ p\neq r; q\neq s}} h_{pqrs} \gamma_{2p+x}\gamma_{2q+x}\gamma_{2r+y}\gamma_{2s+y}.$$
$$(A16)$$

Thus we obtain

$$\sum_{p,q,r,s} h_{pqrs} \left(\sum_{\mathbf{x} \in B_1} \Gamma_{pqrs}^{\mathbf{x}} \right) \qquad (A17)$$
$$= -2 \left(\sum_{\substack{p,q,r,s \\ p\neq q; r\neq s}} + \sum_{\substack{p,q,r,s \\ p\neq r; q\neq s}} \right) h_{pqrs} \gamma_{2p}\gamma_{2q}\gamma_{2r+1}\gamma_{2s+1}.$$

Since we would like to completely separate the quadratic terms from the quartic terms, we observe that if $p = q$ or $r = s$ in the above expression, then those terms reduce to quadratic order (or the identity, if both equalities hold). The first summation automatically excludes such reduction, so we analyze the second one, again separating the diagonal and off-diagonal summands with respect to each

pair (p, q) and (r, s):

$$\sum_{\substack{p,q,r,s \\ p\neq r; q\neq s}} h_{pqrs}\gamma_{2p}\gamma_{2q}\gamma_{2r+1}\gamma_{2s+1}$$

$$= \sum_{\substack{p,q,r,s \\ p\neq r; q\neq s \\ p\neq q; r\neq s}} h_{pqrs}\gamma_{2p}\gamma_{2q}\gamma_{2r+1}\gamma_{2s+1} + \sum_{\substack{p,r \\ p\neq r}} h_{pprr}\mathbb{1}$$

$$+ \sum_{\substack{p,q,r \\ p\neq r; q\neq r \\ p\neq q}} h_{pqrr}\gamma_{2p}\gamma_{2q} + \sum_{\substack{p,r,s \\ p\neq r; p\neq s \\ r\neq s}} h_{pprs}\gamma_{2r+1}\gamma_{2s+1}$$

$$= \sum_{\substack{p,q,r,s \\ p\neq r; q\neq s \\ p\neq q; r\neq s}} h_{pqrs}\gamma_{2p}\gamma_{2q}\gamma_{2r+1}\gamma_{2s+1} + \sum_{\substack{p,r \\ p\neq r}} h_{pprr}\mathbb{1}$$

$$+ \sum_{\substack{p,q,r \\ p\neq r; q\neq r \\ p\neq q}} h_{pqrr}\{\gamma_{2p}, \gamma_{2q}\} + \sum_{\substack{p,r,s \\ p\neq r; p\neq s \\ r<s}} h_{pprs}\{\gamma_{2r+1}, \gamma_{2s+1}\}$$

$$= \sum_{\substack{p,q,r,s \\ p\neq r; q\neq s \\ p\neq q; r\neq s}} h_{pqrs}\gamma_{2p}\gamma_{2q}\gamma_{2r+1}\gamma_{2s+1} + \sum_{\substack{p,q \\ p\neq q}} h_{ppqq}\mathbb{1}. \quad \text{(A18)}$$

So we see that these quadratic terms in fact vanish due to anticommutation.

Now we show that the remaining 12 cases yield the same operators as those already obtained in Eq. (A12). Let $B_2 = \{0000, 0110, 1001, 1111\}$ and $x, y \in \{0, 1\}$:

$$\sum_{p,q,r,s} h_{pqrs}\Gamma^{xyyx}_{pqrs} = \sum_{\substack{p,q,r,s \\ p\neq q; r\neq s}} h_{pqrs}\gamma_{2p+x}\gamma_{2q+y}\gamma_{2r+y}\gamma_{2s+x}$$

$$= \sum_{\substack{p,q,r,s \\ p\neq q; r\neq s \\ p\neq s}} h_{pqrs}\gamma_{2p+x}\gamma_{2q+y}\gamma_{2r+y}\gamma_{2s+x}$$

$$+ \sum_{\substack{p,q,r \\ p\neq q; r\neq p}} h_{pqrp}\gamma_{2q+y}\gamma_{2r+y}. \tag{A19}$$

The second sum simplifies to

$$\sum_{\substack{p,q,r \\ p\neq q; r\neq p}} h_{pqrp}\gamma_{2q+y}\gamma_{2r+y} = \sum_{\substack{p,q \\ p\neq q; r\neq p \\ q<r}} h_{pqrp}\{\gamma_{2q+y}, \gamma_{2r+y}\}$$

$$+ \sum_{\substack{p,q \\ p\neq q}} h_{pqqp}\mathbb{1}$$

$$= \sum_{\substack{p,q \\ p\neq q}} h_{pqqp}\mathbb{1}. \tag{A20}$$

The first sum depends on whether x and y are the same

or not. If $x \neq y$, then

$$\sum_{\substack{p,q,r,s \\ p\neq q; r\neq s \\ p\neq s}} h_{pqrs}\gamma_{2p+x}\gamma_{2q+y}\gamma_{2r+y}\gamma_{2s+x}$$

$$= \sum_{\substack{p,q,r,s \\ p\neq q; r\neq s \\ p<s}} h_{pqrs}(\gamma_{2p+x}\gamma_{2q+y}\gamma_{2r+y}\gamma_{2s+x}$$

$$+ \gamma_{2s+x}\gamma_{2q+y}\gamma_{2r+y}\gamma_{2p+x})$$

$$= \sum_{\substack{p,q,r,s \\ p\neq q; r\neq s \\ p<s}} h_{pqrs}(\gamma_{2p+x}\gamma_{2q+y}\gamma_{2r+y}\gamma_{2s+x}$$

$$- \gamma_{2p+x}\gamma_{2q+y}\gamma_{2r+y}\gamma_{2s+x})$$

$$= 0. \tag{A21}$$

If $x = y$, we first observe that if $p \neq r$ and $q \neq s$, then the sum vanishes, as demonstrated above. Therefore we have the three remaining cases ($p \neq r$ and $q = s$, $p = r$ and $q \neq s$, and $p = r$ and $q = s$):

$$\sum_{\substack{p,q,r,s \\ p\neq q; r\neq s \\ p\neq s}} h_{pqrs}\gamma_{2p+x}\gamma_{2q+x}\gamma_{2r+x}\gamma_{2s+x}$$

$$= -\sum_{\substack{p,q \\ p\neq q}} h_{pqpq}\mathbb{1} - \sum_{\substack{p,q,r \\ p\neq q; r\neq q \\ p\neq r}} h_{pqrq}\gamma_{2p+x}\gamma_{2r+x}$$

$$- \sum_{\substack{p,q,s \\ p\neq q; p\neq s \\ q\neq s}} h_{pqps}\gamma_{2q+x}\gamma_{2s+x}$$

$$= -\sum_{\substack{p,q \\ p\neq q}} h_{pqpq}\mathbb{1} - \sum_{\substack{p,q,r \\ p\neq q; r\neq q \\ p<r}} h_{pqrq}\{\gamma_{2p+x}, \gamma_{2r+x}\}$$

$$- \sum_{\substack{p,q,s \\ p\neq q; p\neq s \\ q<s}} h_{pqps}\{\gamma_{2q+x}, \gamma_{2s+x}\}$$

$$= -\sum_{\substack{p,q \\ p\neq q}} h_{pqpq}\mathbb{1}. \tag{A22}$$

Altogether, the terms corresponding to B_2 are just the identity operator:

$$\sum_{p,q,r,s} h_{pqrs}\left(\sum_{\mathbf{x}\in B_2} \Gamma^{\mathbf{x}}_{pqrs}\right) = \sum_{\substack{p,q \\ p\neq q}} (4h_{pqqp} - 2h_{pqpq})\mathbb{1}. \tag{A23}$$

Let $B_3 = \{0010, 0100, 1011, 1101\}$. These strings give rise to the same terms, since for $x \in \{0, 1\}$,

$$\sum_{p,q,r,s} \Gamma^{x01x}_{pqrs} = \sum_{p,q,r,s} h_{pqrs}i\gamma_{2p+x}\gamma_{2q}\gamma_{2r+1}\gamma_{2s+x}$$

$$= -\sum_{p,q,r,s} h_{pqrs}i\gamma_{2p+x}\gamma_{2q+1}\gamma_{2r}\gamma_{2s+x} \tag{A24}$$

$$= \sum_{p,q,r,s} \Gamma^{x10x}_{pqrs}.$$

We simplify the sum using the same type of manipulations as in Eq. (A21):

$$\sum_{p,q,r,s} h_{pqrs} i\gamma_{2p}\gamma_{2q}\gamma_{2r+1}\gamma_{2s}$$

$$= \sum_{\substack{p,q,r,s \\ p\neq q; r\neq s \\ p<s; s\neq q}} h_{pqrs} i(\gamma_{2p}\gamma_{2q}\gamma_{2r+1}\gamma_{2s} - \gamma_{2p}\gamma_{2q}\gamma_{2r+1}\gamma_{2s})$$

$$- \sum_{\substack{p,q,r \\ p\neq q; r\neq q}} h_{pqrq} i\gamma_{2p}\gamma_{2r+1} + \sum_{\substack{p,q,r \\ p\neq q; r\neq p}} h_{pqrp} i\gamma_{2q}\gamma_{2r+1}$$

$$= \sum_{\substack{p,q,r \\ p\neq r; q\neq r}} (h_{prrq} - h_{prqr}) i\gamma_{2p}\gamma_{2q+1}. \tag{A25}$$

Thus we obtain

$$\sum_{p,q,r,s} h_{pqrs} \left(\sum_{\mathbf{x}\in B_3} \Gamma^{\mathbf{x}}_{pqrs} \right)$$

$$= 4 \sum_{\substack{p,q,r \\ p\neq r; q\neq r}} (h_{prrq} - h_{pqrr}) i\gamma_{2p}\gamma_{2q+1}. \tag{A26}$$

The last set is $B_4 = \{0001, 0111, 1000, 1110\}$. Again, all four strings correspond to the same terms. We show this by evaluating, for $w, x, y \in \{0,1\}$ with $w \neq y$,

$$\sum_{p,q,r,s} \Gamma^{wxxy}_{pqrs} = (-1)^{w+1} \sum_{p,q,r,s} h_{pqrs} i\gamma_{2p+w}\gamma_{2q+x}\gamma_{2r+x}\gamma_{2s+y}$$

$$= (-1)^{w+1} \sum_{\substack{p,q,r,s \\ p\neq q; r\neq s \\ q<r}} h_{pqrs} i(\gamma_{2p+w}\gamma_{2q+x}\gamma_{2r+x}\gamma_{2s+y}$$

$$- \gamma_{2p+w}\gamma_{2q+x}\gamma_{2r+x}\gamma_{2s+y})$$

$$+ (-1)^{w+1} \sum_{\substack{p,q,s \\ p\neq q; q\neq s}} h_{pqqs} i\gamma_{2p+w}\gamma_{2s+y}$$

$$= (-1)^{w+1} \sum_{\substack{p,q,r \\ p\neq r; r\neq q}} h_{prrq} i\gamma_{2p+w}\gamma_{2q+y}. \tag{A27}$$

If we order the Majorana product such that the even index appears first, then the sign of $(-1)^{w+1}$ cancels with that of swapping γ_{2p+w} with γ_{2q+y}, and so we have

$$\sum_{p,q,r,s} h_{pqrs} \left(\sum_{\mathbf{x}\in B_4} \Gamma^{\mathbf{x}}_{pqrs} \right) = 4 \sum_{\substack{p,q,r \\ p\neq r; q\neq r}} h_{prrq} i\gamma_{2p}\gamma_{2q+1}. \tag{A28}$$

Finally, we collect all the terms from Eqs. (A17), (A23), (A26) and (A28), along with the slight simplification in

Eq. (A18), to write the two-body terms as

$$\frac{1}{2} \sum_{p,q,r,s} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s = \frac{1}{32} \sum_{p,q,r,s} h_{pqrs} \left(\sum_{\mathbf{x}\in\{0,1\}^4} \Gamma^{\mathbf{x}}_{pqrs} \right)$$

$$= \frac{1}{8} \sum_{\substack{p,q \\ p\neq q}} (h_{pqqp} - h_{pqpq}) \, \mathbb{1}$$

$$+ \frac{1}{8} \sum_{\substack{p,q,r \\ p\neq r; q\neq r}} (2h_{prrq} - h_{pqrr}) i\gamma_{2p}\gamma_{2q+1}$$

$$- \frac{1}{16} \left(\sum_{\substack{p,q,r,s \\ p\neq q; r\neq s}} + \sum_{\substack{p,q,r,s \\ p\neq q; r\neq s \\ p\neq r; q\neq s}} \right) h_{pqrs}\gamma_{2p}\gamma_{2q}\gamma_{2r+1}\gamma_{2s+1}. \tag{A29}$$

Including the one-body terms, Eq. (A12), we express the full electronic structure Hamiltonian in terms of Majorana operators:

$$H = \frac{1}{2} \left[\sum_p h_{pp} + \frac{1}{4} \sum_{\substack{p,q \\ p\neq q}} (h_{pqqp} - h_{pqpq}) \right] \mathbb{1}$$

$$+ \sum_{p,q} \left[\frac{1}{2} h_{pq} + \sum_{\substack{r \\ p\neq r; q\neq r}} \left(\frac{1}{4} h_{prrq} - \frac{1}{8} h_{pqrr} \right) \right] i\gamma_{2p}\gamma_{2q+1}$$

$$- \frac{1}{16} \left(\sum_{\substack{p,q,r,s \\ p\neq q; r\neq s}} + \sum_{\substack{p,q,r,s \\ p\neq q; r\neq s \\ p\neq r; q\neq s}} \right) h_{pqrs}\gamma_{2p}\gamma_{2q}\gamma_{2r+1}\gamma_{2s+1}. \tag{A30}$$

Defining new coefficients as

$$\tilde{h} = \frac{1}{2} \sum_p h_{pp} + \frac{1}{8} \sum_{\substack{p,q \\ p\neq q}} (h_{pqqp} - h_{pqpq}),$$

$$\tilde{h}_{pq} = \frac{1}{2} h_{pq} + \sum_{\substack{r \\ p\neq r; q\neq r}} \left(\frac{1}{4} h_{prrq} - \frac{1}{8} h_{pqrr} \right), \tag{A31}$$

$$\tilde{h}_{pqrs} = -\frac{1}{8}[1 + (1-\delta_{pr})(1-\delta_{qs})]h_{pqrs}.$$

we obtain the Hamiltonian as presented in the main text, Eq. (67).

Appendix B: Proof details for Theorem 1

To see why Eqs. (71) and (72) hold, we first examine the structure of our anticommuting partition $\{S_{(q,r,s)}\}$. Although we have the choice of matching either one or three indices in each term's support, here we only use the
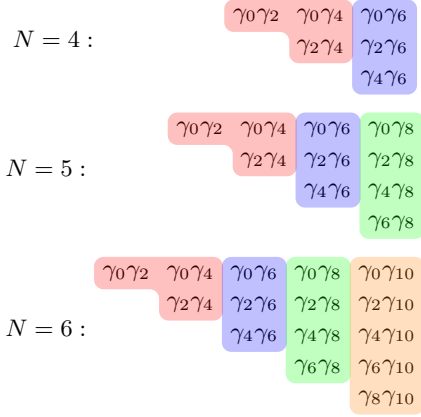
FIG. 3. Partitioning of electronic structure terms. Finding an anticommuting partition of the quartic terms can be reduced to finding an anticommuting partition of quadratic terms with exclusively even (equiv. odd) indices. Each highlighted bin is such an anticommuting set. Excluding the red bin, each set shares one common index $2q$ for $3 \le q \le N-1$. Although only 3 values of N are depicted, the induction of this diagram is straightforward for arbitrary N. Thus one obtains $N-3$ bins of size q each and 1 "red bin" of size 3.

condition of three matches. This amounts to matching exactly one even index, since the other two must be odd (or vice versa, by symmetry). In this sense, the problem reduces to finding an anticommuting partition of the set of all quadratic Majorana operators with only even indices in their support. Taking products with the set of all quadratic operators with only odd indices in their support then generates all the relevant quartic operators, $\mathcal{M}$ (up to phase factors).

One may readily check from the definition of $S_{(q,r,s)}$ that they do indeed cover $\mathcal{M}$ and are all pairwise disjoint. However, since we have reduced the problem to considering simply quadratic operators, we may provide a visual argument which clearly demonstrates the partitioning scheme, Figure 3. Note that for $N=2$, there is only one unique quartic term, and for $N=3$, all the even quadratics already anticommute (i.e. the red bin in the figure). From the figure, we immediately see the disjointness property satisfied, with each set of common index $2q$ having size q. The exception, again, is the red bin, which corresponds to the union $S_{(1,r,s)} \cup S_{(2,r,s)}$ as mentioned in the main text. Hence there are $N-2$ anticommuting sets of even-index quadratic operators, and taking products with all $\binom{N}{2}$ odd-index quadratic operators yields the desired $O(N^3)$ result.

Appendix C: Electronic structure systems

TABLE I. The systems examined in our numerical analysis. Geometries were obtained from the NIST CCBDB database [64], and molecular orbital integrals in the Hartee–Fock basis obtained from Psi4 [65] and OpenFermion [66].

System	Charge	Multiplicity	Basis	Qubits
Ar_1	0	1	STO-3G	18
B_1	0	2	STO-3G	10
Be_1	0	1	STO-3G	10
Br_1	0	2	STO-3G	36
C_1O_1	0	1	STO-3G	20
C_1O_2[a]	0	1	STO-3G	30
C_1	0	3	STO-3G	10
Cl_1	0	2	STO-3G	18
Cl_1	-1	1	STO-3G	18
F_1	0	2	STO-3G	10
F_2	0	1	STO-3G	20
H_1Cl_1	0	1	STO-3G	20
H_1F_1	0	1	3-21G	22
H_1F_1	0	1	STO-3G	12
H_1He_1	0	1	3-21G	8
H_1He_1	0	1	6-311G**	24
H_1He_1	0	1	6-311G	12
H_1He_1	0	1	6-31G**	20
H_1He_1	0	1	6-31G	8
H_1He_1	0	1	STO-3G	4
$H_1Li_1O_1$	0	1	STO-3G	22
H_1Li_1	0	1	3-21G	22
H_1Li_1	0	1	STO-3G	12
H_1Na_1	0	1	STO-3G	20
H_1O_1	-1	1	STO-3G	12
H_1	0	2	STO-3G	2
H_2Be_1	0	1	STO-3G	14
$H_2C_1O_1$	0	1	STO-3G	24
H_2C_1	0	3	3-21G	26
H_2C_1	0	3	STO-3G	14
H_2C_1	0	3	STO-3G	14
H_2C_2	0	1	STO-3G	24
H_2Mg_1	0	1	STO-3G	22
H_2O_1	0	1	STO-3G	14
H_2O_2[b]	0	1	STO-3G	24
H_2S_1	0	1	STO-3G	22
H_2	0	1	3-21G	8
H_2	0	1	6-311G**	24
H_2	0	1	6-311G	12
H_2	0	1	6-31G**	20
H_2	0	1	6-31G	8
H_2	0	1	STO-3G	4
H_3N_1	0	1	STO-3G	16
H_3	0	1	3-21G	12
H_3	1	1	STO-3G	6

[a] Circuit generation failed for 10 and 100 maximum gates.
[b] Failed for Jordan-Wigner mapping.

H_4C_1	0	1 STO-3G	18
H_4C_2	0	1 STO-3G	28
H_4N_1	1	1 STO-3G	18
He_1	0	1 STO-3G	2
K_1	0	2 STO-3G	26
Li_1	0	2 STO-3G	10
Mg_1	0	1 STO-3G	18
N_1O_2 [a]	0	2 STO-3G	30
N_1	0	4 STO-3G	10
N_2	0	1 STO-3G	20
Na_1	0	2 STO-3G	18
Ne_1	0	1 STO-3G	10
O_1	0	3 STO-3G	10
O_2	0	1 STO-3G	20
O_2	0	3 STO-3G	20
P_1	0	4 STO-3G	18
S_1	0	3 STO-3G	18
Si_1	0	3 STO-3G	18

[a] Failed for Bravyi-Kitaev mapping.