

Compressing Many-Body Fermion Operators under Unitary Constraints

Nicholas C. Rubin,* Joonho Lee,* and Ryan Babbush*



Cite This: <https://doi.org/10.1021/acs.jctc.1c00912>



Read Online

ACCESS |



Metrics & More

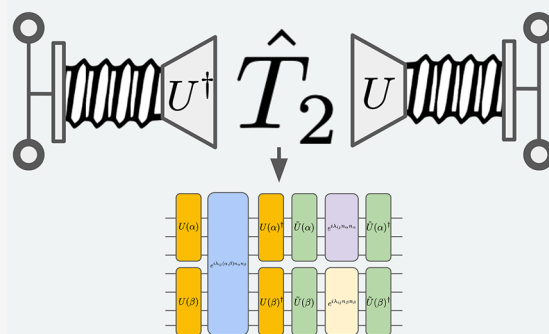


Article Recommendations



Supporting Information

ABSTRACT: The most efficient known quantum circuits for preparing unitary coupled cluster states and applying Trotter steps of the arbitrary basis electronic structure Hamiltonian involve interleaved sequences of Fermionic Gaussian circuits and Ising interaction-type circuits. These circuits arise from factorizing the two-body operators generating those unitaries as a sum of squared one-body operators that are simulated using product formulas. We introduce a numerical algorithm for performing this factorization that has an iteration complexity no worse than single particle basis transformations of the two-body operators and often results in many times fewer squared one-body operators in the sum of squares, compared to the analytical decompositions. As an application of this numerical procedure, we demonstrate that our protocol can be used to approximate generic unitary coupled cluster operators and prepare the necessary high-quality initial states for techniques (like ADAPT-VQE) that iteratively construct approximations to the ground state.



1. INTRODUCTION

Efficient quantum circuit compilation is an important task for performing quantum simulations on near-term and fault-tolerant quantum devices. Different approximation schemes can lead to vastly different circuit representations and, thus, varying runtimes and success probabilities.^{1–4} In this work, we focus on the general problem of circuit implementations of propagators generated by Fermionic many-body operators relevant to a wide variety of simulation such as time-dynamics and state preparation in electronic structure simulations of chemistry and condensed matter models.

The strategy we explore is the decomposition of a generic two-body operator into a sum-of-squares of normal operators where each term in the sum can be implemented exactly with Fermionic Gaussian unitaries (i.e., single-particle rotations) and unitaries generated by charge–charge (i.e., Ising) type interactions. Circuits of this form can be implemented exactly through Givens rotation networks⁵ and swap networks.^{3,6} Many recently proposed simulation strategies for Fermions leverage an analytical sum-of-squares many-body operator decomposition or use a sum-of-squares type ansatz for approximate ground states. Some examples for both near-term quantum computers and fault-tolerant quantum computers are the double factorized Trotter steps for chemical Hamiltonians,⁷ tensor-hypercontraction-based Hamiltonian dynamics,⁸ restricted models of generalized coupled-cluster,^{9–11} and compressed density fitting.¹² The sum-of-squares picture unifies these ansätze and suggests a numerical compilation strategy for determining a sum-of-squares operator decomposition with few terms. In the context of coherence limited near-term quantum computers, we

highlight how a numerical approach can lead to substantially shorter circuits for the implementation of many-body operators relevant to quantum chemistry.

A numerical optimization strategy for a nonorthogonal single particle basis representation of many-body operators has already been used in determining one-particle bases to measure chemical Hamiltonians¹³ and compressing many-body operators through numerical density fitting.¹² A classical analogue of these methods is matching pursuit, where the dictionary is a set of nonorthogonal single particle bases, which are obtained as the algorithm progresses.^{14,15} In the context of many-body operators, these methods are connected by the Lie algebraic perspective on operator decomposition into bases that maximize their Cartan subalgebra representations.¹³ Numerically determining a unitary that maximizes the Cartan subalgebra of a qubit operator has also been used to determine efficient compilations for two-local qubit Hamiltonians.¹⁶ A missing component of many of these proposals is an efficient computational scheme for determining the basis that maximizes the Cartan subalgebra components.

We propose an efficient local search (i.e., greedy) algorithm for recursively decomposing a Fermionic two-body operator

Received: September 10, 2021

into a sum-of-squares terms with an iteration cost that scales no worse than a one-particle basis rotation of the two-body operator $\mathcal{O}(n^5)$, where n is the number of spin-orbitals describing the problem. A numerical study on decomposition of the unitary coupled cluster operator indicates that a greedy search provides substantial improvements over analytical decompositions such as the singular value decomposition⁷ or Takagi¹⁷ decomposition and provides a computationally efficient alternative to nongreedy approaches to unitary tensor fitting.^{10,12}

To further highlight the utility of this approach, we demonstrate how compressed unitary coupled cluster doubles can serve as a starting point for iterative wave function methods that rely on the initial wave function overlap with the ground state to succeed. In classical quantum chemistry, iterative approaches based on similarity transformations of a state^{18,19} or reduced-density-matrix propagation²⁰ require high accuracy initial states if a high-quality approximations to the true ground state is desired. We validate that quantum versions of these techniques, such as ADAPT-VQE²¹ or the quantum anti-hermitian-contracted Schrödinger equation solver,²² can also be sensitive to initial states by studying the performance of ADAPT-VQE on finding a high-accuracy approximation to the ground state of O_2 . This example highlights how iterative circuit construction techniques can fail in the absence of high initial overlap with the target state, even when one introduces artificial symmetry breaking.

In Section 2, we describe the sum-of-squares decomposition of a generic Fermionic two-body operator and how this translates into the Gaussian unitary and Ising swap network circuit primitives under a Trotter approximation. In Section 3, we describe our greedy algorithm for performing the decomposition of a generic two-body operator into a sum-of-squares form. In Section 4, we compare the numerical sum-of-squares decomposition to analytical techniques for unitary coupled-cluster generators and electron–electron interaction operators, along with an application to iterative circuit constructions. We close with perspectives on the numerical compression and when it is most applicable in the context of quantum computing for simulating Fermions.

2. A SUM-OF-SQUARES NORMAL OPERATOR REPRESENTATION

2.1. Background on Simulating Two-Body Fermion Operators as a Sum-of-Squares One-Body Operators.

Starting from a generic antihermitian two-body operator,

$$G = \sum_{pqrs} A_{rs}^{pq} a_p^\dagger a_q^\dagger a_s a_r \quad (1)$$

where $\{p, q, r, s\}$ index Fermionic modes, the charge–charge form can be obtained by reordering ladder operators under the Fermionic anticommutation relations,

$$G = \sum_{pq,rs} A_{sr}^{pq} (a_p^\dagger a_s a_q^\dagger a_r - \delta_s^q a_p^\dagger a_r) \quad (2)$$

where $A \in \mathbb{C}^{\otimes n}$ and is antisymmetric in the upper and lower indices. A sum-of-squares of normal operator decomposition of G has the form

$$G = \sum_l Z_l^2 - \sum_{pr} S_{pr} a_p^\dagger a_r \quad Z_l = \sum_{pq} z(l)_{pq} a_p^\dagger a_q \quad (3)$$

where $z(l)_{pq}$ is a collection of coefficients such that Z_l is a normal operator and $S_{pr} = \sum_{qr} A_{rs}^{pq} \delta_s^q$. Ansatz generated by G can be viewed as a unitary form of the generalized coupled-cluster ansatz of Nooijen²³ or, if A is an imaginary antihermitian tensor with the correct symmetries, a generic quantum chemistry Hamiltonian evolution.

Under the exponential map, the operator G expressed in the form of eq 3 admits a simple compilation strategy through a first-order Trotter approximation:

$$e^G \approx e^{-S} \prod_l e^{Z_l^2} \quad (4)$$

By representing each Z_l in its eigenbasis, each Z_l^2 term in eq 4 can be exactly implemented as

$$e^{Z_l^2} = U_l e^{\sum_{pq} J_{pq}^{(l)} n_p n_q} U_l^\dagger \quad (5)$$

where U_l is a Fermionic Gaussian unitary rotating to the eigenbasis of Z_l and $J(l)_{pq}$ corresponds to the outer product of eigenvalues of $z(l)$. Equation 4 becomes

$$e^G \approx e^{-\sum_{pq} S_{pq} a_p^\dagger a_q} U_L^\dagger \prod_{l=1}^L e^{\sum_{pq} J_{pq}^{(l)} n_p n_q} \tilde{U}_l \quad (6)$$

where $\tilde{U}_l = U_l U_{l-1}^\dagger$ is the concatenation of single-particle basis rotations—we take U_0 to be identity. This concatenation can be performed classically and implemented on the quantum computer as a linear depth Givens rotation network.^{5,24–27} The charge–charge component of this sequence of unitaries $e^{\sum_{pq} J_{pq}^{(l)} n_p n_q}$ can be implemented with a linear depth swap network.^{3,6}

The Trotterized form of e^G is thus implemented as a sequence of linear depth circuits alternating between Givens rotations networks and charge–charge type networks, all of which require only nearest-neighbor connectivity between a linear array of qubits. Furthermore, these compilations are conjectured to be optimal for these decompositions. The scaling of this implementation is linear in the rank of the matrix A , which is the super matrix formed from reshaping A such that the row and column indices are labeled by pair indices ps and qr (see the Supporting Information for a detailed discussion on geminal ordering of the supermatrix).

2.2. Determining Normal Operators. There are a variety of methods for decomposing G into sum-of-squares normal operators as eq 3. In ref 7, the double factorization technique, similar to the Cholesky decomposition, is applied to the \tilde{A} super matrix to produce a sum-of-squares representation. The factors from the Cholesky decomposition can be reshaped and factorized again, via an eigenvalue decomposition, because of the 4-fold symmetry of the Hamiltonian coefficients coming from the two-electron integral coefficients. Reference 7 also demonstrated a general decomposition of a unitary coupled cluster operator that relies on a singular value decomposition (SVD) of \tilde{A} . In the Supporting Information, we derive the SVD decomposition for an arbitrary operator G without the structural requirements of a coupled cluster doubles operator. A slightly more efficient decomposition for unitary coupled-cluster doubles operators is pointed out by Mastsuzawa et al.¹⁰ in the context of implementing Jastrow inspired sum-of-squares many-body operators by leveraging the Takagi decomposition.¹⁷ The Takagi decomposition is applicable to complex symmetric matrices as the decomposition

$$\tilde{A} = U \text{diag}(\sigma_1, \dots, \sigma_m) U^T \quad (7)$$

where $\sigma_i \geq 0$ and U is unitary. Generally, the Takagi decomposition is not equivalent to a scaled SVD²⁸ as U is complex. To form a sum-of-squares decomposition from the Takagi decomposition, we follow ref 10 and reshape the columns of U into matrices $u(l)$, where l indexes the column and define

$$y(l) = \sqrt{\sigma(l)} u(l) \quad (8)$$

such that

$$A_{qr}^{ps} = \sum_l y(l)_{ps} y(l)_{qr} \quad (9)$$

Because $y(l)$ is not generally a normal matrix, we can form one by taking linear combinations

$$y(l)^\pm = y(l) \pm iy(l)^\dagger \quad (10)$$

such that

$$A_{qr}^{ps} = \frac{1}{4} \sum_l (y(l)_{ps}^+ y(l)_{qr}^+ + y(l)_{ps}^- y(l)_{qr}^-) \quad (11)$$

The coefficients of $y(l)^\pm$ can be thought of as the coefficients for the one-body operators $z(l)$ in eq 3. Thus, they can be diagonalized by a unitaries μ^\pm such that

$$A_{cd}^{ab} = \frac{1}{4} \sum_{l,p,q} (\mu^{+*}(l)_{ap} \mu_{cp}^+(l) J_{pq}^+(l) \mu_{bq}^{+*}(l) \mu_{dq}^+(l) + \mu^{-*}(l)_{ap} \mu_{cp}^-(l) J_{pq}^-(l) \mu_{bq}^{-*}(l) \mu_{dq}^-(l)) \quad (12)$$

where $J_{pq}^\pm(l) = \lambda_p^\pm(l) \lambda_q^\pm(l)$ and $\lambda_p^\pm(l)$ come from diagonalizing $y(l)^\pm$. A full derivation of this general form is described in the Supporting Information. Written more abstractly,

$$A_{cd}^{ab} = \sum_l \mu(l)_{ap}^* \mu(l)_{cp} J_{pq}(l) \mu(l)_{bq}^* \mu(l)_{dq} \quad (13)$$

where we have moved the \pm terms into the l sum. Both the Takagi decomposition and the SVD sum-of-squares decomposition form an operator decomposition like eq 13. It was previously noted that $J_{pq}(l)$ is a rank-one matrix, and thus carries little information.¹⁰ This justifies a numerical optimization of this doubles term, as is performed in the unitary cluster Jastrow factor ansatz¹⁰ (k -uCJ) and its variants.^{9,11} Reference 12 proposed an approach based on a gradient descent least-squares fitting of $\mu(l)$ and $J(l)$ for \tilde{A} under the constraint that the coefficients have the proper symmetry to represent a spin-free chemical Hamiltonian. Although ref 12 introduces a clever optimization strategy that alternates between one particle basis $\mu(l)$ optimization and $J(l)$ coefficient optimization, the overall convergence of the least-squares fitting is unknown and seems to be limited to six tensor factors before numerical difficulties make optimization challenging. In Section 3, we propose a fitting algorithm that determines a decomposition of A , according to eq 13, which has comparable performance but is substantially more computationally efficient.

2.3. S_z -Symmetry Adaptation. Both the SVD scheme and Takagi scheme can be implemented in such a fashion that all generators commute with S_z . Specifically, this can be achieved by applying the decompositions to \tilde{A} partitioned into the nonredundant spin components $\tilde{A}_{\alpha,\omega}^{\alpha,\alpha}$, $\tilde{A}_{\alpha,\beta}^{\alpha,\beta}$, and $\tilde{A}_{\beta,\beta}^{\beta,\beta}$. Consider the spin-indexed generator \tilde{G} ,

$$\tilde{G}_{p\sigma\sigma',q\tau\tau'} = a_{p\sigma}^\dagger a_{s\sigma'} a_{q\tau}^\dagger a_{r\tau'} - a_{p\sigma}^\dagger a_{r\tau} \delta_{q\tau}^{\sigma\sigma'} \quad (14)$$

We can view the spin-block structure in \tilde{A} in matrix form (via the same supermatrix formed in the Takagi and SVD decomposition):

$$\tilde{A} = \begin{pmatrix} \tilde{A}(\alpha, \alpha) & \tilde{A}(\alpha, \beta) \\ \tilde{A}(\beta, \alpha) & \tilde{A}(\beta, \beta) \end{pmatrix} = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \quad (15)$$

which is a complex symmetric matrix. $\tilde{A}(\tau, \sigma)$ indicates all terms with two-body operators of the form $a_{p\tau}^\dagger a_{q\sigma} a_{r\sigma}^\dagger a_{s\tau}$. By performing one of the sum-of-squares decomposition on the A and C blocks and then the larger matrix involving B and B^T , we are guaranteed the one-particle basis rotations for each sum-of-squares operator is restricted to a single spin-sector. The A and C blocks can be implemented simultaneously and can be merged with the single-particle basis transformation obtained by rearranging $a_p^\dagger a_q^\dagger a_s a_r \rightarrow a_p^\dagger a_s a_q^\dagger a_r$. Figure 1 shows an example of the circuit compilation for one of the l terms in the Takagi or SVD decomposition of \tilde{A} that commutes with S_z .

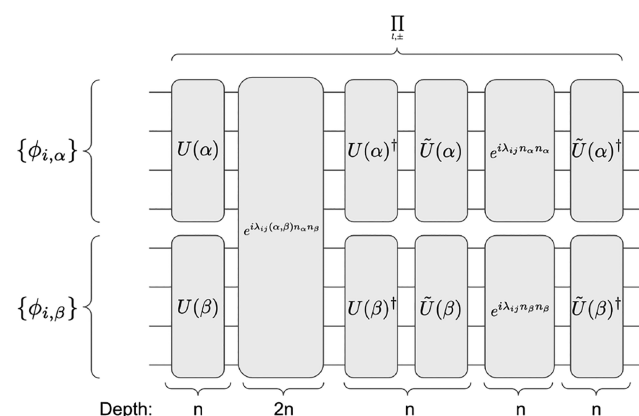


Figure 1. Depiction of a single Trotter slice of the sum-of-squares decomposition for each spin sector. Each $U(\tau)$ is implemented as a linear depth Givens rotation circuit and each $e^{i\sum_{ij} \lambda_{ij} n_\alpha n_\beta}$ is also implemented in a linear depth swap network pattern. Each swap network or basis change unitary consists of $\mathcal{O}(n^2)$ gates.

3. COMPRESSING MANY-BODY OPERATORS WITH SEQUENTIAL ORBITAL OPTIMIZATION

The previous section reviewed methods that allow for a coefficient tensor associated with a generic operator to be decomposed into a sum-of-squares of normal operators. It also explained how the evolution by these sum-of-squares operators can be efficiently implemented using Fermionic Gaussian unitaries and Ising interactions. Here, we describe a numerical algorithm to express A in a sum-of-squares normal matrices according to eq 13. The procedure is a greedy algorithm that sequentially finds a single-particle basis such that A has large coefficients for $n_p n_q$ terms. We then remove this component leaving a remainder tensor. The protocol is repeated until the remainder is numerically zero or the remainder norm is below a preset threshold. We note that this procedure leads to matrices $J_{pq}(l)$ (the coefficients associated with the $n_p n_q$ in eq 5) which are not restricted to being rank one. We numerically determine the single-particle basis that maximizes $J_{pq}(l)$ through gradient descent on the nonredundant generator coefficients. Derivatives of the basis rotation unitaries, with respect to the generating

coefficients, are provided in refs 29 and 30 and have been used in gradient optimization of the Hartree–Fock equations.^{27,30} The algorithm we propose is as follows:

- (1) Starting from the desired two-body operator G , we first maximize the coefficients associated with the $n_i n_j$ terms by orbital optimization.
- (2) Select out the $n_i n_j$ coefficients and store them along with the orbital rotation generator.
- (3) Rotate the operator represented by the diagonal coefficients back to the original basis just obtained from the optimization, and then subtract the tensor from the original, generating a remainder.
- (4) Repeat steps 1–3 until the norm of the subtraction remainder is below a predefined threshold.

Using this approach the iteration cost of our optimization is never more than the $O(n^5)$ cost of orbital rotation. Throughout the rest of this paper, this recursive fitting procedure will be referred to as the “unitary compression” technique. We also emphasize that the classical storage requirement for each sum of square tensor factor is quadratic in the system size making the total storage of the unitary compressed tensor $O(ln^2)$, where l is the number of compression terms and n is the basis set size. This is similar to Cholesky factorized storage requirements if l scaled linearly with system size. In the results section, we explore the scaling of l with various systems, in terms of tensor accuracy under the L_2 norm and energy of the reconstructed generator.

To derive a cost function for maximizing $J_{pq}(l)$ (coefficients associated with $n_i n_j$), we consider a one-body transform of a two-body operator,

$$\hat{T} = \sum_{ijkl} t_{ij,kl} a_i^\dagger a_j a_k^\dagger a_l \quad (16)$$

which generates unitary dynamics. Given the single particle basis transformation operator,

$$U(\kappa) = e^{\sum_{pq} \kappa_{pq} a_p^\dagger a_q} \quad (17)$$

the orbital rotated generator is represented below as \tilde{T} :

$$\tilde{T} = \hat{U}^\dagger(\kappa) \hat{T} \hat{U}(\kappa) \quad (18)$$

$$\tilde{T} = \sum_{pqrs} \tilde{t}_{pq,rs} a_p^\dagger a_q a_r^\dagger a_s \quad (19)$$

$$\tilde{t}_{pq,rs} = \sum_{ijkl} u_{pi}^* u_{qj} u_{rk}^* u_{sl} t_{ijkl} \quad (20)$$

such that the objective of maximizing the coefficients of the $n_i n_j$ component of \tilde{T} can be expressed as

$$\max_{\kappa} O(\kappa), \quad O(\kappa) = \sum_{xy} |\tilde{t}_{xx,yy}(\kappa)|^2 \quad (21)$$

To optimize, we use L-BFGS-B using the following equations to construct the gradient optimally. The gradient for an arbitrary continuous cost function, depending on $\tilde{t}_{pq,rs}$ can be obtained by differentiating eq 21 with respect to κ_{ab} via the chain rule:

$$\frac{\partial O(\kappa)}{\partial \kappa_{ab}} = \sum_{cd} \frac{\partial O(\kappa)}{\partial u_{cd}} \left(\frac{\partial u}{\partial \kappa_{ab}} \right)_{cd} + \sum_{cd} \frac{\partial O(\kappa)}{\partial u_{cd}^*} \left(\frac{\partial u^*}{\partial \kappa_{ab}} \right)_{cd} \quad (22)$$

where

$$\begin{aligned} \frac{\partial O(\kappa)}{\partial u_{cd}} &= 2 \sum_{xy} \operatorname{Re} \left[\tilde{t}_{xx,yy}^* \frac{\partial \tilde{t}_{xx,yy}}{\partial u_{cd}} \right] \\ &= 2 \sum_{yikl} \operatorname{Re} [t_{ickl}^* u_{id}^* u_{ky} u_{ly} \tilde{t}_{dd,yy}] + 2 \sum_{xijk} \operatorname{Re} [t_{ijck}^* u_{ix}^* u_{jx} u_{ld}^* \tilde{t}_{xx,dd}] \end{aligned} \quad (23)$$

$$\begin{aligned} \frac{\partial O(\kappa)}{\partial u_{cd}^*} &= 2 \sum_{xy} \operatorname{Re} \left[\tilde{t}_{xx,yy}^* \frac{\partial \tilde{t}_{xx,yy}}{\partial u_{cd}^*} \right] \\ &= 2 \sum_{yikl} \operatorname{Re} [t_{cjk}^* u_{jd}^* u_{ky}^* u_{ly} \tilde{t}_{dd,yy}] + 2 \sum_{xijl} \operatorname{Re} [t_{ijcl}^* u_{ix}^* u_{jx} u_{ld} \tilde{t}_{xx,dd}] \end{aligned} \quad (24)$$

which, for all $\{c, d\}$, is obtained in $O(n^5)$ once per gradient call. The n^2 intermediates are reused for each κ_{ab} derivative. Thus, the overall scaling is $O(n^5)$. Using the Wilcox identity,²⁹ the $\{c, d\}$ element of the partial derivative of the unitary $\left(\frac{\partial u}{\partial \kappa_{a,b}} \right)_{cd}$ or its Hermitian conjugate is

$$\left(\frac{\partial u}{\partial \kappa_{ab}} \right)_{cd} = \sum_r W_{cr}^{ab} u_{rd} \quad (25)$$

$$\left(\frac{\partial u^*}{\partial \kappa_{ab}} \right)_{cd} = \sum_r -W_{rc}^{ab} u_{rd}^* \quad (26)$$

where the matrix W^{ab} are the coefficients for the anti-Hermitian operator obtained from the Wilcox formula,²⁹ which is an analytical expression for the derivative of a unitary, with respect to its generating parameter $\kappa_{a,b}$ (see Appendix G of ref 27 for a full derivation). This analytical formula merely requires diagonalizing the generator matrix κ and requires no truncation of the matrix exponential Taylor expansion.

4. RESULTS

We examine the performance of the SVD, Takagi, and unitary compression decomposition for the two-body components of unitary coupled-cluster operators and two-electron integral tensors. In this analysis, we compare the number of tensor factors versus maximum absolute error and L_2 -norm variation from the true tensor. We found that, for cluster operator compilations, the greedy unitary compression technique requires very few tensors to reach sub-millihartree accuracy but suffers a substantial slowdown in optimization, because of the increasing rank of the residual. In the two-electron integral case, we demonstrate unitary compression optimized through a global least-squares and the greedy approach both fail to recover high-quality factorizations for a simple system. All calculations are accomplished with PySCF,³¹ OpenFermion,³² the Fermionic Quantum Emulator,³³ and a custom implementation of the tensor decomposition schemes.

4.1. Factorization of Coupled Cluster Doubles. The first system we examine is the analytical (via SVD and Takagi) and numerical (unitary compression and global least-squares) decomposition of the two-body generators constructed from classical coupled-cluster singles and doubles (CCSD) solutions. Given the T_2 CCSD operator,

$$T_2 = \frac{1}{4} \sum_{i,j,a,b} t_{i,j,a,b} a_a^\dagger a_i a_b^\dagger a_j \quad (27)$$

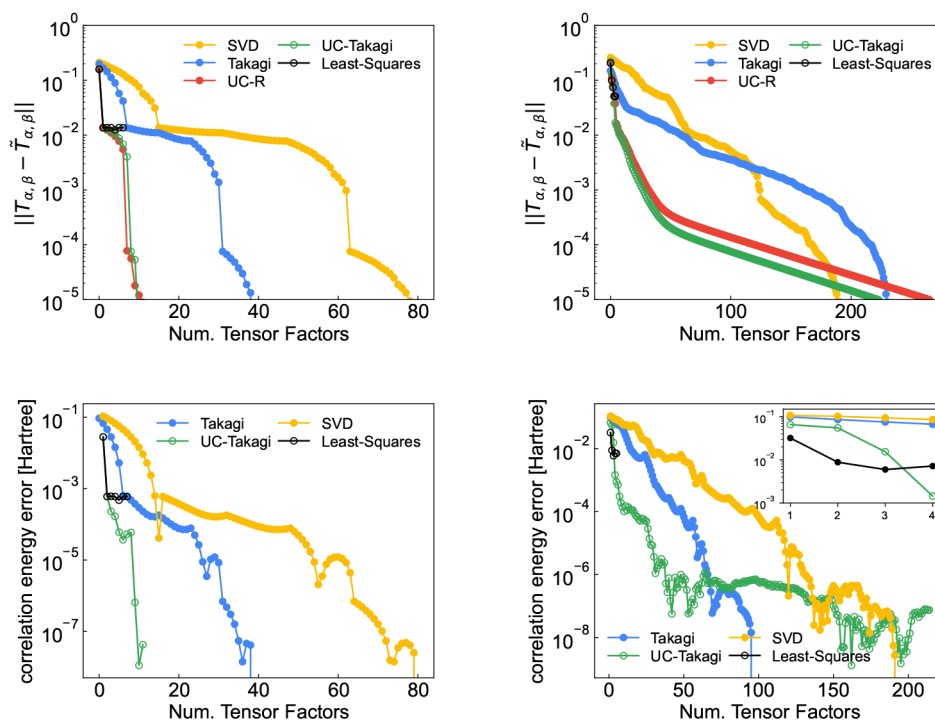


Figure 2. Compression of the $\alpha\beta$ two-body cluster operator for hydrogen fluoride (left) in a minimal basis (12 orbitals) and H_4 -linear (right) in a 6-31G basis (16 orbitals). Unitary compression is labeled with UC. UC-R in red is the unitary compression with random initial starting states for the orbital optimization. The UC-Takagi in green is unitary compression, starting from the Takagi decomposition of the remainder matrix. The SVD decomposition and Takagi decomposition are described in the [Supporting Information](#). Least-squares optimization is labeled in black. The starting parameters for the least-squares optimization are the Takagi decomposition κ matrix and J matrices along with zero-mean and 0.1 variance Gaussian noise added to each variable. Least-squares only has less than five data points due to excessively long computational run times. Each tensor factor can be implemented with $O(n)$ depth, where n is the number of orbitals.

the unitary generator τ_2 can be formed by subtracting the Hermitian conjugate

$$\tau_2 = T_2 - T_2^\dagger \quad (28)$$

which we decompose into the sum-of-squares form. We apply the Takagi, SVD, and unitary compression decompositions to two systems, hydrogen fluoride in a minimal basis and linear H_4 in a 6-31G basis both with bond lengths of 1.6 Å. In [Figure 2](#), we plot the L_2 -norm difference of the $\alpha\beta$ -amplitudes tensor and the unitary compressed tensor, as a function of the number of tensor factors considered. For the unitary compression, we consider random initialization for the basis rotation coefficients or coefficients obtained from the top vector of the Takagi decomposition. We also plot the convergence of the correlation energy, as a function of tensor factors. For these two systems, there is little difference between the unitary compression seeded with the top eigenvector of the Takagi decomposition and random one-body unitaries. The residual L_2 -norm for the unitary compression technique drops quickly but eventually converges with slower scaling than the initial steps. Despite this slowdown, the initial tensors from unitary compression capture enough information such that the correlation energy converges to sub-millihartree levels, with respect to CCSD well before the Takagi decomposition and the SVD decomposition.

In [Figure 2](#), we also plot the L_2 -norm and correlation energy convergence of tensor factors determined by simultaneous optimization of all sum-of-squares variational parameters. We call this technique the global least-squares unitary compression. The objective function and gradient derivation can be found in the [Supporting Information](#). As reported in ref 12, convergence

of the simultaneous variation of all parameters is extremely slow. We found that only 4–5 tensor factors could be optimized within a reasonable amount of gradient calls ($O(1000)$) for these small systems when the convergence tolerance was set to 1.0×10^{-5} . Furthermore, least-squares simultaneous optimization was extremely sensitive to initial parameters. When seeded with random starts gradient descent on all parameters commonly stopped at a local minimum; producing correlation energies well above Takagi. In [Figure 2](#), we see the simultaneous optimization with Takagi vectors corrupted zero mean Gaussian random noise with a variance of 0.1. Random noise is added to help escape local minimum. The computational difficulties confirm the results of ref 12 and demonstrate the necessity of unitary compression. The inset plot of the bottom right panel of [Figure 2](#) demonstrates that unitary compression and full simultaneous optimization have similar performance for the first four tensors. Although we expect full least-squares optimization to always be lower than the greedy unitary compression, the higher value at the fourth point of the inset of [Figure 2](#) is due to difficulty in finding good starting parameters.

The cause of the convergence slowdown for greedy unitary compression is likely due to the fact that the nuclear norm of the residual is not being minimized in the greedy procedure. To show this, we plot the rank of the residual being fit by the unitary compression procedure and compare against the residual in the Takagi and SVD case. The residual in the Takagi and SVD cases are simply the true amplitude tensor minus the reconstructed tensor with a given number of tensor factors. The rank of the residual is the number of tensor factors determined using the

analytical decomposition on the residual. The rank in the greedy unitary compression case is the number of factors resulting from the Takagi decomposition. Using all the tensor factors produces an exact amplitude tensor and, thus, the residual rank for the Takagi and SVD decomposition asymptotes to zero. In Figure 3,

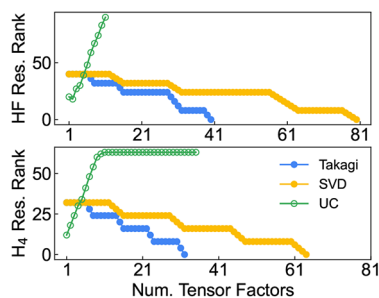


Figure 3. Rank of the residual being fit with Takagi, SVD, and unitary compression (UC). The Takagi and SVD residual ranks are determined by reconstructing the tensor with a given number of tensor factors (x -axis) and computing the rank of the remainder using the Takagi decomposition. The residual rank for unitary compression is the Takagi rank of the residual tensor for the next fitting iteration.

we show the rank of the residuals for hydrogen fluoride (HF) molecule and linear H_4 molecule as a stacked plot. As expected for Takagi and SVD decompositions, the rank of the residual (or remainder tensor) goes to zero. For the unitary compression, the rank quickly rises to its maximal value. Once the rank is maximized, the convergence of the unitary tensor fitting slows substantially, as seen in previous plots. Effectively, the unitary compression approach partitions the τ_2 coefficient tensor into a low-rank component that captures the majority of the correlation energy and a maximal rank residual component with many small amplitudes. This suggests to either include the nuclear norm of the residual in unitary compression objective or consider a hybrid scheme where unitary compression is used until the residual rank is maximal and then switch to a Takagi decomposition on the remainder tensor.

4.2. Electron Repulsion Interaction Decomposition.

We consider unitary compression on the two-electron integral tensor for the π -system of naphthalene computed in a cc-pVDZ basis at the geometry from ref 34. We demonstrate that unitary compression, both greedy and global least-squares, present numerical issues when determining high-quality tensor decompositions. We modify the greedy unitary compression scheme to fit Hermitian operators and optimize over the space of spatial orbital rotations. In Figure 4, we plot the maximum absolute deviation of the unitary compressed two-electron integral tensor. The greedy unitary compression protocol is performed until the residual L_2 -norm is $<1.0 \times 10^{-5}$. This bound is selected based on the convergence of second-order Møller–Plesset perturbation theory (MP2) and coupled-cluster with singles and doubles (CCSD) with truncated low-rank factorization of the two-electron integrals for the large metal–organic catalyst FeMoCo.⁸ The performance of greedy unitary compression is compared against a Cholesky factorization of the two-electron integral matrix and global least-squares optimization of unitary compression fitting. In the least-squares unitary compression, we use the basis rotation and charge–charge parameters determined from the greedy procedure as a starting guess and allow a maximum of 500 iterations with a stopping criteria of 1.0×10^{-4} in the BFGS gradient. The Cholesky decomposition is performed via an SVD. Greedy unitary compression succeeds in

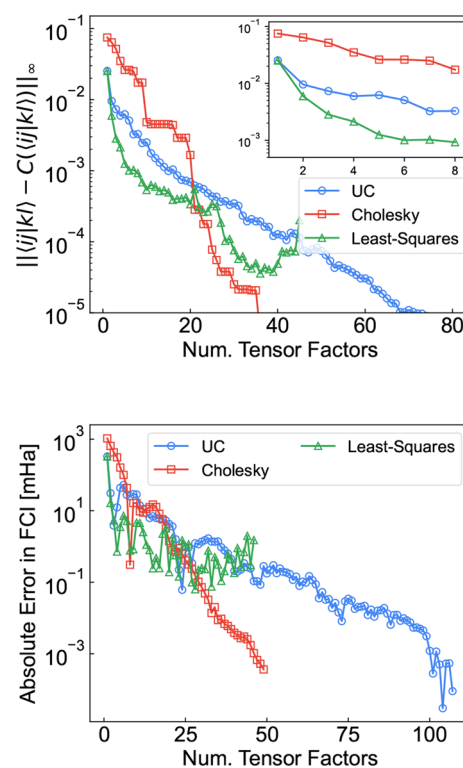


Figure 4. (Left) Maximum absolute deviation (MAD) of the unitary compressed two-electron integral tensor of naphthalene (10 orbital π system) obtained from canonical Hartree–Fock orbitals in a cc-pVDZ basis compared against the Cholesky factorized two-electron integral tensor. Inset plot is the MAD for the first eight tensor factors. (Right) Absolute difference from FCI energy computed with untruncated two-electron integrals. In both the left and right plots, we restrict the compression to real unitary rotations.

lowering the maximum absolute deviation over the Cholesky decomposition but only up until 20 tensor factors. On the right-hand side of Figure 4, we plot the absolute error in exact diagonalization (FCI) energies computed with truncated two-electron integral operators. The vastly different energies from FCI indicate that, although both greedy and least-squares global unitary compression lower the maximum absolute deviation of the two-electron integral tensor, the resulting Hamiltonian is very different from the original. This brings into question whether unitary compression is a useful technique for compressing operators derived from the Coulomb kernel.

4.3. Many-Body Starting States for Iterative Wave Function Construction.

To further illustrate the utility of unitary compression, we consider starting states for the iterative circuit construction technique ADAPT.²¹ In many classical and quantum algorithms, the initial state can vastly change the success probability of the algorithm. Here, we demonstrate a system where ADAPT converges to a state substantially higher in energy than the ground state when initialized with a symmetry preserving Hartree–Fock state and succeeds in finding a low-energy state when an approximation, through unitary compression, to a CCSD initial state is used. All ADAPT calculations used operator pools of S_2 -adapted two-body operators and numerical optimization was performed with BFGS.³⁵ Gradients were obtained through the dynamic programming approach described in ref 36. All numerics were performed with the Fermionic quantum emulator.³³

The system we consider is the triplet ground state of O_2 in a minimal STO-3G basis with a bond distance of 2.55 Å. Internal stability analysis is performed on all self-consistent-field (SCF) calculations to confirm that the SCF solution is not a saddle point. The restricted-openshell Hartree–Fock (ROHF) wave function has less than 1×10^{-5} overlap with the full configuration interaction (FCI) wave function and the unrestricted Hartree–Fock (UHF) wave function has $\sim 35\%$ overlap with the FCI wave function. This is because UHF should produce a state that is locally a singlet and triplet on each respective O atom. Thus, the overlap for the full triplet O_2 should be nontrivial. In all cases, ADAPT-VQE and CCSD calculations are performed on the full space of 10 orbitals and 16 electrons.

In Figure 5, we show that unitary compression can be used to approximate CCSD as a starting state with very few tensor

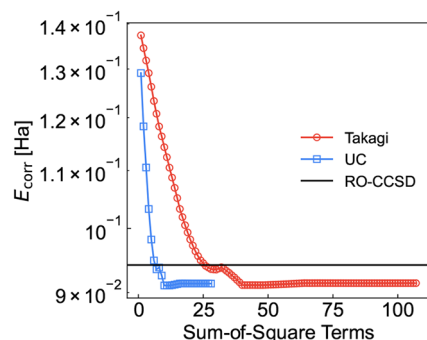


Figure 5. Correlation energy convergence of compiled unitary coupled cluster as a starting state for ADAPT. Blue is unitary compression and red is the Takagi decomposition.

factors. We compare convergence of the correlation energy captured by unitary compression and Takagi on the α, β spin blocks of τ_2 defined in eq 28, as a function of the number of tensor factors considered. Similar to the results of the previous section, unitary compression substantially outperforms the Takagi decomposition, in terms of circuit depth. To achieve a correlation energy similar to CCSD, unitary compression requires six factors whereas Takagi requires 25 factors.

On the left panel of Figure 6, we plot the progress of the ADAPT algorithm starting from a Hartree–Fock starting point

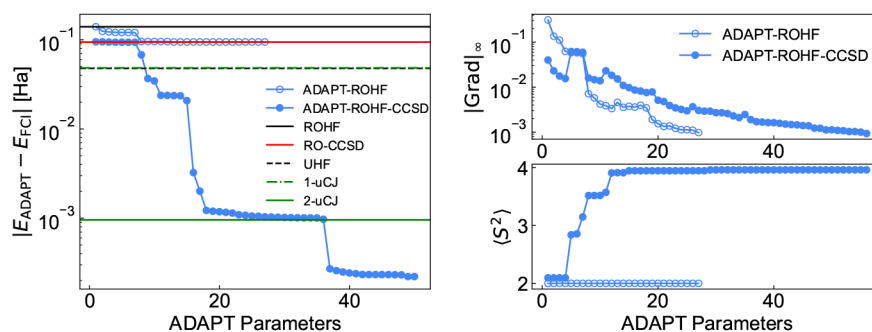


Figure 6. (Left) Convergence of ADAPT for different starting states. Open blue circles correspond to an ROHF starting point and solid blue circles correspond to a RO–CCSD starting point approximated with six unitary compression factors. Energies for various methods are also shown as reference: ROHF (solid black), UHF (dashed black), RO–CCSD (red), $k = 1$ unitary cluster–Jastrow¹⁰ (k -uCJ) (dash-dot green), and $k = 2$ k -uCJ (solid green). (Right, upper panel) Maximum absolute deviation of the two-body gradient tensor $g_{pqrs} = \langle \psi | [a_p^\dagger a_q^\dagger a_r a_s, H] | \psi \rangle$ after each ADAPT iteration for ROHF starting states and an approximate RO–CCSD starting state. (Right, lower panel) S^2 expectation value after each adapt iteration demonstrating symmetry breaking in ADAPT steps.

(ROHF) and a coupled-cluster starting point (RO–CCSD) approximated with six unitary compression factors. In blue open circles, we show how ADAPT starting from an ROHF wave function fails to converge to a ground state due to the difficulty of finding high-quality rotations with a gradient based approach when the system has almost zero overlap the exact ground state. Starting from RO–CCSD approximated with six unitary compression tensors (solid blue circles) ADAPT can succeed, but with substantial circuit depth. We note here that UHF as a starting point for ADAPT succeeds but with substantial symmetry breaking. The varying performance of ADAPT depending on starting state symmetry breaking hints at the importance of symmetry breaking for the algorithm overall. For reference, we also plot the performance of the $k = 1$ and $k = 2$ unitary cluster Jastrow ansatz¹⁰ (k -uCJ) where the generalized singles term (restricted such that rotations in the α and β spin sectors are equivalent) is implemented separately in the pair doubles term in a similar compilation described in Figure 1. For O_2 , in a minimal basis, k -uCJ has $k100$ parameters.

On the right panel of Figure 6, we plot the maximum absolute value of the two-body gradient set $\{ \langle \psi | [a_p^\dagger a_q^\dagger a_r a_s, H] | \psi \rangle \}$, as a function of ADAPT iteration. At each iteration, a single spin-adapted two-body operator is added to the wave function. Thus, S^2 symmetry can be broken by ADAPT. We use S_z spin-adapted operators because singlet and triplet two-body operator implementation require Trotterization, depending on implementation strategy. With only S_z spin-adapted function, there is no way to fix S^2 expectation values through the ADAPT protocol, unless a penalty term is added. To demonstrate symmetry breaking as a mechanism for ADAPT successfully finding a stationary state, we also plot S^2 as a function of ADAPT iteration. With the improved starting state of approximate RO–CCSD 17 ADAPT iterations are required to come within 1 millihartree of the $k = 2$ k -uCJ energy, which is 1 millihartree from the true FCI value. This highlights that with improved starting states ADAPT can find high-quality ground states with very few parameters: 17 vs 100.

5. DISCUSSION

We have explored a variety of compilation techniques for implementing many-body Fermion dynamics and draw connections between known implementation strategies, such as the SVD or Takagi decomposition, circuit ansätze like k -uCJ,

and a sum-of-squares representation of a generic operator. All the referenced strategies seek to decompose a many-body operator into a sum-of-squares of normal operators which can be implemented with Trotter error, using interleaved Givens rotation networks and Ising interaction networks. The many-body decomposition schemes based on SVD or Takagi suffer from a rank deficiency in the Ising interaction matrix,¹⁰ which can be partially alleviated through full variational relaxation.

We proposed a strategy for determining a sum-of-squares decomposition of a general two-body operator by numerically searching for a low-depth nonorthogonal one-particle basis expansion of the operator. The greedy numerical decomposition has iteration complexity no worse than a single particle basis transformation. The decomposition can be applied in energy measurement schemes as in ref 13 or for time evolution or ansatz construction.

The numerical sum-of-squares decomposition clearly outperformed analytical decompositions for approximating a unitarized projected coupled-cluster generator (28), resulting in substantial circuit depth reduction. Thus, we can recommend unitary compression as a compilation strategy when the goal is to implement a many-body operator of unitary coupled-cluster form. Unitary compression can also be applied to many-body interaction terms of higher rank with an appropriate increase in iteration complexity mirroring single particle basis rotation costs. We demonstrated the use of coupled-cluster compilation via unitary compression as a starting state for ADAPT-VQE. Without approximate coupled-cluster initial state ADAPT fails to converge to the ground state for triplet O₂. Although unitary compression presents a nice starting point, it is important to note that $k = 1$ k -uCJ, which is an instance of a Fermionic non-Gaussian state, is efficiently simulatable^{37,38} and also provides a low-depth route for improving the starting state for ADAPT-VQE. We further note that it may be beneficial to study unitary compression for implementing dynamics of the generate determined from the gradient estimation step of ADAPT-VQE, which would implement a many-body step instead of implementing a single ADAPT term, which scales linearly with arbitrary qubit topology, or many-adapt terms, which would require Trotterization. The linear scaling of a single ADAPT term compared to the linear scaling circuit depth of a unitary compression term is an area of potential further investigation.

The numerical sum-of-squares decomposition is a generic and useful tool for translating low-symmetry many-body operators into quantum circuits and a unifying framework for many of today's algorithms for simulating Fermions on near-term and future quantum computers.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.1c00912>.

Decompositions of the doubles generator; least-squares optimization of sum-of-squares (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Nicholas C. Rubin – Google Quantum AI, San Francisco, California 94105, United States; orcid.org/0000-0003-3963-1830; Email: nickrubin@google.com

Joonho Lee – Google Quantum AI, San Francisco, California 94105, United States; Department of Chemistry, Columbia

University, New York, New York 10027, United States;

orcid.org/0000-0002-9667-1081; Email: linusjoonho@gmail.com

Ryan Babbush – Google Quantum AI, San Francisco, California 94105, United States; Email: ryanbabbush@gmail.com

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jctc.1c00912>

Notes

The authors declare no competing financial interest.

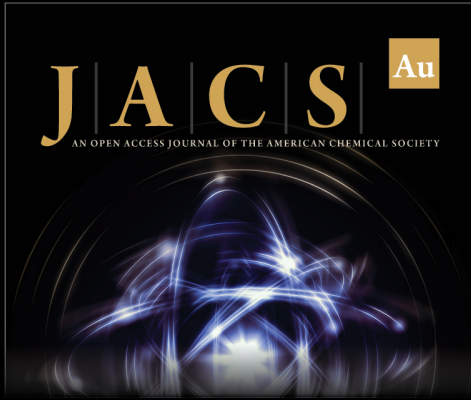
■ ACKNOWLEDGMENTS

J.L. thanks David Reichman for support. We thank William J. Huggins for a detailed reading of the manuscript and discussions.


■ REFERENCES

- (1) Hastings, M. B.; Wecker, D.; Bauer, B.; Troyer, M. Improving quantum algorithms for quantum chemistry. *arXiv Preprint* **2014**, arXiv:1403.1539.
- (2) Babbush, R.; McClean, J.; Wecker, D.; Aspuru-Guzik, A.; Wiebe, N. Chemical basis of Trotter-Suzuki errors in quantum chemistry simulation. *Phys. Rev. A* **2015**, *91*, 022311.
- (3) Kivlichan, I. D.; McClean, J.; Wiebe, N.; Gidney, C.; Aspuru-Guzik, A.; Chan, G. K.-L.; Babbush, R. Quantum simulation of electronic structure with linear depth and connectivity. *Phys. Rev. Lett.* **2018**, *120*, 110501.
- (4) others, et al. Postponing the orthogonality catastrophe: efficient state preparation for electronic structure simulations on quantum devices. *arXiv preprint* **2018**, arXiv:1809.05523.
- (5) Kivlichan, I. D.; McClean, J.; Wiebe, N.; Gidney, C.; Aspuru-Guzik, A.; Chan, G. K.-L.; Babbush, R. Quantum Simulation of Electronic Structure with Linear Depth and Connectivity. *Phys. Rev. Lett.* **2018**, *120*, 110501.
- (6) O’Gorman, B.; Huggins, W. J.; Rieffel, E. G.; Whaley, K. B. Generalized swap networks for near-term quantum computing. *arXiv Preprint* **2019**, arXiv:1905.05118.
- (7) Motta, M.; Ye, E.; McClean, J. R.; Li, Z.; Minnich, A. J.; Babbush, R.; Chan, G. K.-L. Low rank representations for quantum simulation of electronic structure. *arXiv Preprint* **2018**, arXiv:1808.02625.
- (8) Lee, J.; Berry, D. W.; Gidney, C.; Huggins, W. J.; McClean, J. R.; Wiebe, N.; Babbush, R. Even more efficient quantum computations of chemistry through tensor hypercontraction. *arXiv Preprint* **2020**, arXiv:2011.03494.
- (9) Lee, J.; Huggins, W. J.; Head-Gordon, M.; Whaley, K. B. Generalized unitary coupled cluster wave functions for quantum computation. *J. Chem. Theory Comput.* **2019**, *15*, 311–324.
- (10) Matsuzawa, Y.; Kurashige, Y. Jastrow-type Decomposition in Quantum Chemistry for Low-Depth Quantum Circuits. *J. Chem. Theory Comput.* **2020**, *16*, 944–952.
- (11) Kottmann, J. S.; Aspuru-Guzik, A. Optimized Low-Depth Quantum Circuits for Molecular Electronic Structure using a Separable Pair Approximation. *arXiv Preprint* **2021**, arXiv:2105.03836.
- (12) Cohn, J.; Motta, M.; Parrish, R. M. Quantum Filter Diagonalization with Double-Factorized Hamiltonians. *arXiv Preprint* **2021**, arXiv:2104.08957.
- (13) Yen, T.-C.; Izmaylov, A. F. Cartan sub-algebra approach to efficient measurements of quantum observables. *arXiv Preprint* **2020**, arXiv:2007.01234.
- (14) Wu, Y.; Batista, V. S. Matching-pursuit for simulations of quantum processes. *J. Chem. Phys.* **2003**, *118*, 6720–6724.
- (15) McClean, J. R.; Aspuru-Guzik, A. Compact wavefunctions from compressed imaginary time evolution. *RSC Adv.* **2015**, *5*, 102277–102283.


- (16) Kökcü, E.; Steckmann, T.; Freericks, J.; Dumitrescu, E. F.; Kemper, A. F. Fixed depth hamiltonian simulation via Cartan decomposition. *arXiv Preprint* **2021**, arXiv:2104.00728.
- (17) Cariolaro, G.; Pierobon, G. Bloch-Messiah reduction of Gaussian unitaries by Takagi factorization. *Phys. Rev. A* **2016**, *94*, 062109.
- (18) Yanai, T.; Chan, G. K.-L. Canonical transformation theory for multireference problems. *J. Chem. Phys.* **2006**, *124*, 194106.
- (19) White, S. R. Numerical canonical transformation approach to quantum many-body problems. *J. Chem. Phys.* **2002**, *117*, 7472–7482.
- (20) Mazziotti, D. A. Anti-Hermitian Contracted Schrödinger Equation: Direct Determination of the Two-Electron Reduced Density Matrices of Many-Electron Molecules. *Phys. Rev. Lett.* **2006**, *97*, 143002.
- (21) Grimsley, H. R.; Economou, S. E.; Barnes, E.; Mayhall, N. J. An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nat. Commun.* **2019**, *10*, 1–9.
- (22) Smart, S. E.; Mazziotti, D. A. Quantum Solver of Contracted Eigenvalue Equations for Scalable Molecular Simulations on Quantum Computing Devices. *Phys. Rev. Lett.* **2021**, *126*, 070504.
- (23) Nooijen, M. Can the Eigenstates of a Many-Body Hamiltonian Be Represented Exactly Using a General Two-Body Cluster Expansion? *Phys. Rev. Lett.* **2000**, *84*, 2108–2111.
- (24) Reck, M.; Zeilinger, A.; Bernstein, H. J.; Bertani, P. Experimental realization of any discrete unitary operator. *Phys. Rev. Lett.* **1994**, *73*, 58–61.
- (25) Clements, W. R.; Humphreys, P. C.; Metcalf, B. J.; Kolthammer, W. S.; Walsmley, I. A. Optimal design for universal multipoint interferometers. *Optica* **2016**, *3*, 1460–1465.
- (26) Jiang, Z.; Sung, K. J.; Kechedzhi, K.; Smelyanskiy, V. N.; Boixo, S. Quantum Algorithms to Simulate Many-Body Physics of Correlated Fermions. *Phys. Rev. Appl.* **2018**, *9*, 044036.
- (27) Arute, F.; et al. Hartree-Fock on a superconducting qubit quantum computer. *Science* **2020**, *369*, 1084–1089.
- (28) Hahn, T. Routines for the diagonalization of complex matrices. *arXiv Preprint* **2006**, physics/0607103.
- (29) Wilcox, R. M. Exponential operators and parameter differentiation in quantum physics. *J. Math. Phys.* **1967**, *8*, 962–982.
- (30) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; John Wiley & Sons, 2014; pp 104–106.
- (31) Sun, Q.; Zhang, X.; Banerjee, S.; Bao, P.; Barbry, M.; Blunt, N. S.; Bogdanov, N. A.; Booth, G. H.; Chen, J.; Cui, Z.-H.; et al. Recent developments in the PySCF program package. *J. Chem. Phys.* **2020**, *153*, 024109.
- (32) McClean, J. R.; Rubin, N. C.; Sung, K. J.; Kivlichan, I. D.; Bonet-Monroig, X.; Cao, Y.; Dai, C.; Fried, E. S.; Gidney, C.; Gimby, B.; et al. OpenFermion: the electronic structure package for quantum computers. *Quantum Sci. Technol.* **2020**, *5*, 034014.
- (33) Rubin, N. C.; Shiozaki, T.; Throssell, K.; Chan, G. K.-L.; Babbush, R. The Fermionic Quantum Emulator. *arXiv Preprint* **2021**, arXiv:2104.13944.
- (34) Mullinax, J. W.; Epifanovsky, E.; Gidofalvi, G.; DePrince, A. E., III Analytic energy gradients for variational two-electron reduced-density matrix methods within the density fitting approximation. *J. Chem. Theory Comput.* **2019**, *15*, 276–289.
- (35) Nocedal, J.; Wright, S. *Numerical Optimization*; Springer Science & Business Media, 2006.
- (36) Crooks, G. E. Gradients of parameterized quantum gates using the parameter-shift rule and gate decomposition. *arXiv Preprint* **2019**, arXiv:1905.13311.
- (37) Shi, T.; Demler, E.; Cirac, J. I. Variational study of fermionic and bosonic systems with non-Gaussian states: Theory and applications. *Ann. Phys.* **2018**, *390*, 245–302.
- (38) Kaicher, M. P.; Jäger, S. B.; Wilhelm, F. K. Algorithm for initializing a generalized fermionic Gaussian state on a quantum computer. *arXiv Preprint* **2021**, arXiv:2105.13047.




JACS Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY



Editor-in-Chief
Prof. Christopher W. Jones
Georgia Institute of Technology, USA

Open for Submissions 

pubs.acs.org/jacsau  ACS Publications
Most Trusted. Most Cited. Most Read.