PAPER • OPEN ACCESS

Application of fermionic marginal constraints to hybrid quantum algorithms

To cite this article: Nicholas C Rubin *et al* 2018 *New J. Phys.* **20** 053020

View the article online for updates and enhancements.

New Journal of Physics

The open access journal at the forefront of physics

PAPER

CrossMark

Application of fermionic marginal constraints to hybrid quantum algorithms

Nicholas C Rubin[1], Ryan Babbush[2] and Jarrod McClean[2]

[1] Rigetti Computing, Berkeley, CA 94710, United States of America
[2] Google Inc., Venice, CA 90291, United States of America

E-mail: nick@rigetti.com, ryanbabbush@gmail.com and jarrod.mcc@gmail.com

Abstract

Many quantum algorithms, including recently proposed hybrid classical/quantum algorithms, make use of restricted tomography of the quantum state that measures the reduced density matrices, or marginals, of the full state. The most straightforward approach to this algorithmic step estimates each component of the marginal independently without making use of the algebraic and geometric structure of the marginals. Within the field of quantum chemistry, this structure is termed the fermionic *n*-representability conditions, and is supported by a vast amount of literature on both theoretical and practical results related to their approximations. In this work, we introduce these conditions in the language of quantum computation, and utilize them to develop several techniques to accelerate and improve practical applications for quantum chemistry on quantum computers. As a general result, we demonstrate how these marginals concentrate to diagonal quantities when measured on random quantum states. We also show that one can use fermionic *n*-representability conditions to reduce the total number of measurements required by more than an order of magnitude for medium sized systems in chemistry. As a practical demonstration, we simulate an efficient restoration of the physicality of energy curves for the dilation of a four qubit diatomic hydrogen system in the presence of three distinct one qubit error channels, providing evidence these techniques are useful for pre-fault tolerant quantum chemistry experiments.

1. Introduction

Quantum computers are expected to provide an exponential speedup in the solution of the electronic structure problem [1–6] in some cases. This could potentially revolutionize research in chemistry and material science by allowing a new mechanism for designing new materials, drugs, and catalysts. Accordingly, there is now a significant body of literature dedicated to developing new algorithms [7–20], tighter bounds and better implementation strategies [21–28], more desirable Hamiltonian representations [29–39], and experimental demonstrations [40–43] for this problem.

The rapid development of quantum hardware in recent years has spurred interest in the development of practical algorithms that do not require fault tolerant quantum computers. It has been conjectured that the leading candidates for first demonstrations of practical algorithms on a pre-threshold device are so-called quantum/classical hybrid algorithms, such as the variational quantum eigensolver (VQE) for chemistry [44–52] and the quantum approximate optimization algorithm (QAOA) for optimization problems [53]. Recent experimental results [50] have shown that the quantum–classical feedback does provide robustness against noise in the device, and simple extensions are possible that allow one to further damp the noise through additional measurements [54, 55].

A common step in almost all near-term implementations of these algorithms is the determination of an operator's expected value through a form of partial tomography of the quantum state. The statistical criteria relating the number of measurements from the distribution and the accuracy of the expected value are relatively

straightfoward to determine [23, 46, 52]. When studying chemical systems, the general benchmark for accuracy is $\langle H \rangle \pm \epsilon$ Hartree where $\epsilon = 1.6 \times 10^{-3}$ is known as chemical accuracy and $\pm \epsilon$ expresses a standard confidence interval. This stringent absolute accuracy is required for matching experimentally determined thermochemical properties such as heats of formation or ionization potentials [56]. Naturally, this accuracy requirement on the estimator for the Hamiltonian may require a large number of independent preparations of a state and measurements using current strategies.

Some routes to reduce the number of experiments by collecting Hamiltonian terms into commuting groups and dropping Hamiltonian terms with small coefficients have already been proposed [23, 46, 51, 57]. Wecker *et al* [48], estimated that for simulating the energy of ferrodoxin using this approach would require 10^{19} total measurements. Though this estimate was made with pessimistic assumptions, the sheer number of measurements motivates one to pursue techniques to accelerate the operator averaging step of VQE and other hybrid algorithms.

The strategy we adopt in order to ameliorate the burden of excessive measurements is to leverage known structure in the marginals of the fermionic density matrices which have not yet been taken advantage of within the field of quantum computation. The set of necessary conditions on marginals of density matrices are known as n-representability conditions and were originally developed to use the reduced density matrices (RDMs) of fermionic and bosonic systems as the main computational variable instead of the wavefunction [58–61]. While in this work we focus on fermionic systems with at most pair-wise interactions and thus use constraints on the 2-marginal ^2D, these ideas are more generally applicable to problems with local Hamiltonian objectives, such as many instances of QAOA. Measuring the marginals of a state ρ, specifically the fermionic two-particle reduced-density-matrix (2-RDM), provides a powerful extension to VQE. These quantities are useful for connecting VQE to other quantum chemical techniques such as multiconfigurational self-consistent field, embedding procedures such as density matrix embedding theory, and augmenting the accuracy of electronic structure methods with perturbation theory, which is thought to be required to apply these methods to complex systems larger than one may fit in the quantum computer alone.

In this work we explore the utility of measuring marginals instead of directly measuring the Hamiltonian, and how the n-representability constraints on the marginals provide (i) a program for reducing the norm of the Hamiltonian directly leading to fewer measurements and (ii) using the set of necessary, but not sufficient constraints, on the fermionic marginals, known as p-positivity constraints, to propose two computational procedures for projecting measured marginals into the set of allowed marginals. The variance reduction that is developed is examined for linear hydrogen chains and generally shows an order of magnitude reduction in the required number of measurements.

We also introduce a series of polynomial time post-processing techniques for certifying and projecting measured marginals. This computational procedure is similar to maximum likelihood tomography except on a RDM space, thus making it efficient. Similar techniques have been applied to corrupted process tomography [62] and reduced state tomography on the RDMs of the system [63]. We compare four projection techniques that balance enforcing n-representability with computational efficiency. The first two techniques are based on positive projection of the 2-RDM matrix—which should be positive semidefinite—but does not include any constraints beyond positivity of the 2-RDM and fixing the trace. The third and fourth techniques add approximate n-representability constraints implemented through a more expensive iterative procedure and semidefinite program (SDP).

The paper is structured as follows: sections 2 and 3 describe fermionic marginals and the n-representability conditions used in this work, sections 4 and 7 review the utility of marginals as they pertain to perturbation theory and a concentration bound for 2-RDMs that suggests a structured ansatz is required when parameterizing the space for hybrid quantum–classical algorithms, section 5 discusses the optimal bound on the number of samples required to measure the expected value of a Hamiltonian and variance reduction from equality n-representability constraints, section 6 discusses the 2-RDM projection procedures and how they can reduce the number of measurements required to measure to fixed accuracy ϵ and also restore physicality after state corruption by common error channels. We close with an outlook as to how n-representability techniques can further improve hybrid algorithms.

2. Marginals of the density operator

Here we introduce our terminology and set notation with respect to RDMs, or marginals of the full density, in both the qubit and fermionic setting. Generally speaking, a marginal of a multivariable probability distribution is the partial trace, or integration, of a subset of the variables leaving a distribution on a smaller set of variables. Given a general quantum state ρ on n qubits

$$\rho = \sum_i w_i |\psi_i\rangle \langle\psi_i|, \tag{1}$$

where $|\psi_i\rangle$ are pure states of n qubits and $\sum_i w_i = 1$, the set of p-qubit RDMs, or p-marginals, of the state are determined by integrating out q-qubits (such that $n - q = p$) of the joint distribution as

$$^p\rho_{m_1,...,m_p} = \mathrm{Tr}_{n_1,n_2,...,n_q}[\rho] \tag{2}$$

resulting in $\binom{n}{p}$ different marginals each of dimension $2^p \times 2^p$. The coefficients $n_1, ..., n_q$ on the trace operator indicate which qubits are integrated out of ρ and coefficients $m_1, ..., m_p$ label the subsystem marginal. The result of marginalization is a distribution on the state space of p-qubits. We will interchangeably refer to these objects as marginals and RDMs alluding to the fact that like the von Neumann density matrix the marginals can be expressed as a sum

$$^p\rho_{m_1,...,m_p} = \sum_j w_j |\phi_j\rangle \langle\phi_j|, \tag{3}$$

where $|\phi_j\rangle$ is a pure state on the subsystem of qubits that has not been traced out.

It is well known that this $\binom{n}{p}$ set of marginals is sufficient for calculating the expected value of a p-local Hamiltonian or other observable [64, 65]. A p-local Hamiltonian is one where any term in the Hamiltonian involves no more than p-qubits interacting. Naturally, such a description of a quantum system is an attractive polynomial size representation. However, in order to use the set of marginals as computational objects (e.g. minimize energy with respect to them) rather than simply measuring them, the RDMs must satisfy certain constraints to ensure physicality. These constraints are termed 'consistency' and are the requirement that all marginals satisfy equation (2) for the same initial state ρ. Despite the considerable structure on $^p\rho_{m_1,...,m_p}$, confirming a set of marginals are consistent was demonstrated to be QMA-complete [64]. Naturally, one may ask if there is some alternative form of consistency, or approximation, that makes working with marginal distributions a computationally attractive approach.

Analogous to the case of qubits, the requirement for calculating expected values of k-local operators describing interactions of indistinguishable particles, such as fermions or bosons, is that only the k-marginal is needed. Note that as a point of clarification, while p-local qubit operators refer to operators acting on at most p qubits, k-local fermionic operators refer to interactions that derive from k-body interactions, and generically act on $2k$ fermionic modes. This does not complicate the methods to be discussed, but is a common point of confusion when working between chemistry and quantum computation. As an example, in chemical systems the energy is a functional of the 1- and 2-local fermionic operators and the 2-marginal of the system [59, 66–68]. Consider a Fock space constructed with a single particle basis associated with a Hilbert space $\mathcal{H}$ of size m. We can represent a general state on this space

$$|\psi\rangle = \sum_{i_1,...,i_m=0}^{1} c_{i_1,...,i_m} (a_m^\dagger)^{i_m} (a_{m-1}^\dagger)^{i_{m-1}}...(a_1^\dagger)^{i_1} |\mathrm{vac}\rangle, \tag{4}$$

where each fermionic creation operator $\{a_m^\dagger\}$ is associated with a single particle basis state $|\phi_m\rangle$ and $|\mathrm{vac}\rangle$ is the vacuum. We may consider a state ψ with fixed particle number n enforced by restricting the coefficients i_m in equation (4) to satisfy $\sum_m i_m = n$. As noted in [69], once an ordering of fermions is selected the Fock space states can be mapped to a Hilbert space of m-distinguishable spin-$\frac{1}{2}$ particles. Marginals of the fermionic n-particle density matrix

$$^nD = |\psi\rangle \langle\psi| = {}^nD_{j_1,...,j_n}^{i_1,...,i_n} |i_n \cdots i_1\rangle \langle j_n \cdots j_1| \tag{5}$$

involve integrating out particles by a trace operation. For example, the 2-RDM ^2D is obtained from ^nD by integrating out particles 3 to n.

$$^2D = \mathrm{Tr}_{3\,...\,n}[{}^nD]. \tag{6}$$

By convention, in the quantum chemistry community the normalization constant for ^pD is typically scaled to $\binom{n}{p}$ when $i_1 < i_2 < ... < i_p$ or $n!/(n-p)!$ when i_k is allowed to be range over all values in $[1, m]$ [59, 61, 70]. In this work we choose the later of the normalizations for ease of computation.

From the study of n-representability theory, a number of efficiently implementable and necessary constraints on the one- and two-particle marginals (^1D and ^2D) are known. Specifically, defining

$$^1D = \sum_{ij} {}^1D_j^i |i\rangle \langle j|, \tag{7}$$

$$^2D = \sum_{pq,rs} {}^2D^{pq}_{rs}|pq\rangle\langle rs|, \tag{8}$$

where

$$^1D^i_j = \text{Tr}[a^\dagger_i a_j\ {}^ND] = \langle\psi|a^\dagger_i a_j|\psi\rangle, \tag{9}$$

$$^2D^{pq}_{rs} = \text{Tr}[a^\dagger_p a^\dagger_q a_s a_r\ {}^ND] = \langle\psi|a^\dagger_p a^\dagger_q a_s a_r|\psi\rangle \tag{10}$$

the simplest of these constraints may be enumerated as:

1. Hermiticity of the density matrices

$$^1D^j_i = ({}^1D^i_j)^*, \tag{11}$$

$$^2D^{pq}_{rs} = ({}^2D^{rs}_{pq})^*. \tag{12}$$

2. Antisymmetry of the 2-particle marginal

$$^2D^{pq}_{rs} = -{}^2D^{pq}_{sr} = -{}^2D^{qp}_{rs} = {}^2D^{qp}_{sr}. \tag{13}$$

3. The $(p-1)$-marginal is related to the p-marginal by contraction—e.g. the 2-marginal can be contracted to the 1-marginal

$$^1D^i_j = \frac{1}{n-1}\sum_k {}^2D^{ik}_{jk}. \tag{14}$$

4. The trace of each marginal is fixed by the number of particles in the system

$$\text{Tr}[^1D] = n, \tag{15}$$

$$\text{Tr}[^2D] = n(n-1). \tag{16}$$

5. The marginals are proportional to density matrices and are thus positive semidefinite

$$\{^1D,\ {}^2D\} \succeq 0. \tag{17}$$

Additional constraints based on the quantum numbers of S^2 and S^z operators can be derived for each marginal [71]. A short description of the form of the linear constraints inspired by fixed particle number $\langle n\rangle$, fixed total angular momentum $\langle S^2\rangle$, and fixed projected total angular momentum $\langle S_z\rangle$ are described in appendix B.

3. The n-representability problem

In the previous section we provided a brief introduction to the notation and setup of problems formulated through their marginal distributions. Here, we review a concise and elegant theoretical framework that allows one to derive the full set of representability conditions for fermionic systems, from which one may select a subset to form efficient approximations. The polynomial size of the p-marginals makes them attractive candidates for use as the representation of quantum systems. This was originally noticed by Coulson and Coleman [72, 73] in the context of quantum chemical Hamiltonians where the energy operator is 2-local and thus a linear functional of the 2-RDM. The characterization and structure of valid 2-marginals arising from the integration of a fermionic density matrix led to the field of n-representability [59, 73].

A significant amount of progress has been made in tackling the n-representability problem; most notably, the original works by Erdahl [74], Percus and Garrod [75, 76], and Mazziotti [67, 77–79] formalize the n-representability problem by specification of an approximate set of constraints by parameterizing the polar cone of the set of 2-RDMs. Recently, Mazziotti formalized the complete structure of the ensemble n-representability constraints [79] and gave the structure of pure-state constraints for states with fixed particle number [78]. n-Representability has also been greatly influenced by quantum information: Liu [64] demonstrated that the n-representability problem is QMA-complete and Bravyi [65] and Klyacho [80] enumerated n-representability constraints for marginals of a pure-state.

3.1. *n*-Representability by characterizing the polar cone

We begin with a geometric picture of the constraints within the space of fermionic density matrices. The formal characterization of the *n*-representable set of ^2D operators relies on characterizing the polar cone of the 2-marginals [59, 60, 68]. Consider the convex set of 2-marginals $^2\mathcal{D}$ acting on the antisymmetric two-fermion space $\wedge^2\mathcal{H}$, which may be defined in terms of its basis vectors $\{a \wedge b = a \otimes b - b \otimes a$ for $a, b \in \mathcal{H}\}$. The polar cone is defined as the subset of Hermitian operators that satisfy the positive projection condition,

$$^2\tilde{\mathcal{D}} = \{^2B \in \wedge \ ^2\mathcal{H}|\langle^2B|^2D\rangle \geqslant 0 \ \forall \ ^2D \in \ ^2\mathcal{D}\}. \tag{18}$$

Operators of the polar cone $^2B \in \ ^2\tilde{\mathcal{D}}$ are positive operators with respect to the 2-RDM, which implies their non-negativity with respect to all fermionic density matrices ND when lifted to the *n*-particle space. In quantum information this lifting procedures is accomplished by taking the tensor product of the operator with identities. For fermions, the lifting procedure involves the tensor product with the appropriate antisymmeterization operations.

The bipolar theorem states that elements of $^2\mathcal{D}$ are completely characterized by the polar cone $^2\tilde{\mathcal{D}}$

$$^2\mathcal{D} = \{^2D \in \ ^2\mathcal{H}|\langle^2B|^2D\rangle \geqslant 0 \ \forall \ ^2B \in \ ^2\tilde{\mathcal{D}}\}. \tag{19}$$

Though specification of inequalities with elements of the polar cone provides a characterization of $^2\mathcal{D}$, it has the major drawbacks that (a) there are infinitely many possible ^2B operators to check and (b) checking ^2B requires checking if an exponentially large operator is positive [81].

The key approximation within *n*-representability is constructing a polynomial size approximation to the polar cone $^2\tilde{\mathcal{D}}_a$ and deriving conditions on the 2-marginal through equation (18). This is achieved by selecting a *k*th-order (where $k < n$) operator basis for $^2\tilde{\mathcal{D}}_a$. Given that the $^2\tilde{\mathcal{D}}_a \subset \ ^2\tilde{\mathcal{D}}$ any representability conditions derived from $^2\tilde{\mathcal{D}}_a$ will be an approximate set of representability conditions. By duality, the polar of $^2\tilde{\mathcal{D}}_a$ implies $^2\mathcal{D} \subset \ ^2\mathcal{D}_a$. This naturally explains why variational calculations using the RDM and approximate *n*-representability constraints are strictly a lower bound to the true energy [70, 77, 82–87].

In the *n*-representability literature related to simulating fermions, fermionic operators up to a particular order—e.g. $\{a_i, a_i^\dagger, a_i a_j, a_i^\dagger a_j^\dagger, \dots\}$—are used as the operator basis for the approximate polar cone. Given a rank-*k* operator basis for the polar cone, we can define a real linear space of Hermitian operators $O_k^\dagger O_k$ where

$$O_k = \sum_{k=1}^{N} \prod_{j=1, o \in \{1, \dagger\}}^{k} c_{k_j} \hat{a}_{k_j}^o \tag{20}$$

that when constrained to be non-negative (implied by equation (18)) form a necessary set of conditions on the *p*-marginals of the von Neumann density matrix.

As an example, we will derive the famous 2-positivity conditions by restricting the rank of the monomials in the operator basis to rank less than 2 as

$$O_2 = \sum_i c_i^a \hat{a}_i + \sum_i c_i^b \hat{a}_i^\dagger + \sum_{ij} c_{ij}^c \hat{a}_i \hat{a}_j + \sum_{ij} c_{ij}^d \hat{a}_i \hat{a}_j^\dagger + \sum_{ij} c_{ij}^e \hat{a}_i^\dagger \hat{a}_j + \sum_{ij} c_{ij}^f \hat{a}_i^\dagger \hat{a}_j^\dagger, \tag{21}$$

setting

$$M_2 = O_2^\dagger O_2, \tag{22}$$

and requiring that $M_2 \succcurlyeq 0$. The $\{c\}$ coefficients in equation (21) specify an arbitrary element of the approximate polar cone $^2\tilde{\mathcal{D}}_a$ in a similar fashion to how a sum-of-squares polynomial can be expressed as a quadratic form $c^T A c$ where elements of A represent various products of monomials. Considering symmetries of the system, such as fixed particle number, reduces the large matrix M_2 to a block diagonal matrix [82]. In this work we consider spinless fermionic Hamiltonians that commute with the number operator of our system and thus we can decompose M_2 into blocks where monomials correspond to number preserving operators—i.e. $\{a_i^\dagger a_j, a_j^\dagger a_i, \dots\}$. Restricting the operator M_2 to be non-negative against the 1-RDM and 2-RDM for all values of c yields the following inequalities

$$\sum_{ij} c_i c_j^* \langle\psi|a_j^\dagger a_i|\psi\rangle \geqslant 0, \tag{23}$$

$$\sum_{ij} c_i c_j^* \langle\psi|a_j a_i^\dagger|\psi\rangle \geqslant 0, \tag{24}$$

$$\sum_{ij,kl} c_{ij} c_{kl}^* \langle\psi|a_i^\dagger a_j^\dagger a_l a_k|\psi\rangle \geqslant 0, \tag{25}$$

$$\sum_{ij,kl} c_{ij} c_{kl}^* \langle\psi|a_i a_j a_l^\dagger a_k^\dagger|\psi\rangle \geqslant 0, \tag{26}$$

$$\sum_{ij,kl} c_{ij} c_{kl}^* \langle \psi | a_i^\dagger a_j a_l^\dagger a_k | \psi \rangle \geqslant 0, \tag{27}$$

where ψ is an arbitrary state. These conditions imply that the following matrices are positive semidefinite

$$^1D = \langle \psi | a_j^\dagger a_i | \psi \rangle \succcurlyeq 0, \tag{28}$$

$$^1Q = \langle \psi | a_j a_i^\dagger | \psi \rangle \succcurlyeq 0, \tag{29}$$

$$^2D = \langle \psi | a_i^\dagger a_j^\dagger a_l a_k | \psi \rangle \succcurlyeq 0, \tag{30}$$

$$^2Q = \langle \psi | a_i a_j a_l^\dagger a_k^\dagger | \psi \rangle \succcurlyeq 0, \tag{31}$$

$$^2G = \langle \psi | a_i^\dagger a_j a_l^\dagger a_k | \psi \rangle \succcurlyeq 0. \tag{32}$$

The positivity of these matrices canonically known as $\{^1D,\ ^1Q,\ ^2D,\ ^2Q,\ ^2G\}$ form a set of necessary conditions that the 2-marginal must obey. Clearly, as the 2-marginal is included in our set, the positivity of this operator naturally appears when building constraints starting from a polar cone picture. Though these constraints are formulated with pure states it is simple to show these conditions hold for mixed states as well.

The positivity of the operators on $\mathcal{H}$–$\{^1D,\ ^1Q\}$—and $\wedge^2 \mathcal{H}$–$\{^2D,\ ^2Q,\ ^2G\}$—are constrained to live in the space defined by equalities obtained by rearranging the fermionic ladder operators according to their anticommutation rules. A full enumeration of the equality constraints is contained in appendix C. In this work we use the positivity constraints and the linear constraints from the anticommutation relationships as a set of constraints that 2-RDMs measured from a quantum resource must satisfy. This enables us to enhance the accuracy of estimation of quantities through employing basic physical relations.

4. Concentration of measure in p-RDMS

Hybrid quantum–classical schemes depend both on the ability to perform the partial tomography that has been discussed as well as some parameterization of the quantum state space. Recent experimental proposals have considered the use of quantum states that are constructed from unitaries that are uniformly random with respect to the Haar measure acted upon a well defined initial state in order to demonstrate so-called 'quantum supremacy' over classical devices [88]. A natural question is to ask whether these states can be harnessed as a resource within hybrid schemes. However, while these states are highly entangled, they demonstrate a number of surprising properties related to results on concentration of measure in high dimensional spaces. From the discussion above, we know it is sufficient to characterize a local fermionic system by its reduced marginals, so we investigate these states in that setting. In a colloquial sense, concentration results for random quantum states show that for many local observables, typical measurement results will yield the average value with overwhelming probability. Here, specifically we investigate the implications of this for p-particle RDMs.

A key result we will leverage from the theory of measure concentration is Levy's lemma [89], which relates to the concentration of Lipshitz-continuous functions. Levy's lemma is as follows: consider a Lipschitz continuous function $f\colon S^{(2n-1)} \to \mathcal{R}$ with Lipschitz constant C, i.e. $|f(x) - f(y)| < C \| x - y \|$ for all $x, y \in S^{(2n-1)}$ where $\|.\|$ is the Euclidean norm in the surrounding space $\mathcal{R}^{2n} \supset S^{(2n-1)}$, and $S^{(2n-1)}$ is the unit sphere in $\mathcal{R}^{2n}$ which would correspond to a quantum state of $\log_2(n)$ qubits. Drawing a point $x \in S^{(2n-1)}$ at random with respect to the uniform measure on the sphere yields

$$\mathrm{Prob}[|f(x) - \langle f \rangle| \geqslant \epsilon] \leqslant 2 \exp\left(-\frac{n\epsilon^2}{9\pi^3 C^2} \right). \tag{33}$$

We will leverage this lemma together with the fact that the expectation value on a normalized quantum state of any Hermitian operator A with bounded spectrum is Lipschitz continuous, with a Lipschitz constant that may be bounded by the norm of the operator.

Consider first the case of 1-RDMs on a space consisting of any number of particles between 0 and M in M spin-orbitals. We are interested in the average value of a matrix element $[^1D]_{ij} = \langle \psi | a_i^\dagger a_j | \psi \rangle$ where $|\psi\rangle$ is a randomly selected pure state. These states may be represented by the density matrix of all possible equally likely occupations,

$$\rho = \frac{I}{d}, \tag{34}$$

where I is the identity matrix on the space of 2^M possible occupations and d is the dimension of this matrix, $d = 2^n$. To evaluate a trace in this case, it suffices to choose the basis of all determinants ranging from 0 to M occupied spin-orbitals, created in a standardized ordering $|\psi\rangle_S = \prod_{k \in S} a_k^\dagger | \rangle$ where $| \rangle$ is the standard fermi vacuum state. To compute the average value of this operator, we thus need to evaluate

$$\langle a_i^\dagger a_j \rangle = \mathrm{Tr}[\rho a_i^\dagger a_j]. \tag{35}$$

On expanding this trace in the determinant basis, one realizes that the terms vanish unless $i = j$, unless this index appears in the wavefunction, and there is at least 1 particle in the wavefunction. The trace is given by

$$
\begin{aligned}
\langle a_i^\dagger a_j \rangle &= \frac{\delta_{ij}}{2^M} \sum_{n=0}^{M} \binom{M-1}{n-1} \\
&= \frac{\delta_{ij}}{2^M} \sum_{n=0}^{M} \frac{n}{M} \binom{M}{n} \\
&= \frac{\delta_{ij}}{M 2^M} (M 2^{M-1}) \\
&= \frac{\delta_{ij}}{2}.
\end{aligned}
\tag{36}
$$

Thus the average 1-RDM is a diagonal matrix with entries $1/2$, corresponding to a system with an average number of particle of $M/2$. Note that for the operator $a_i^\dagger a_j$, the Lipschitz constant can be safely bounded by 1, as independent of the distance in space, this expectation value can differ by at most 1 due to the spectrum of the operator. As a result, Levy's lemma informs us how large we expect typical deviations to be from this average matrix element as

$$\mathrm{Prob}[|\langle \psi | a_i^\dagger a_j | \psi \rangle - \langle a_i^\dagger a_j \rangle| \geqslant \epsilon] \leqslant 2 \exp\left(-\frac{2^M \epsilon^2}{9\pi^3} \right). \tag{37}$$

Examining the case of 2-RDMs now using the same techniques, we find that

$$
\begin{aligned}
[{}^2D]_{kl}^{ij} &= \frac{1}{2} \langle a_i^\dagger a_j^\dagger a_l a_k \rangle = \frac{1}{2^{M+1}} (\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk})(1 - \delta_{ij})(1 - \delta_{kl}) \sum_{n=0}^{M} \binom{M-2}{n-2} \\
&= \frac{1}{2^{M+1} M(M-1)} (\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk})(1 - \delta_{ij})(1 - \delta_{kl}) \sum_{n=0}^{M} (n^2 - n)\binom{M}{n} \\
&= \frac{1}{2^{M+1} M(M-1)} (\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk})(1 - \delta_{ij})(1 - \delta_{kl})((M + M^2)2^{M-2} - M 2^{M-1}) \\
&= \frac{1}{8}(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk})(1 - \delta_{ij})(1 - \delta_{kl}).
\end{aligned}
\tag{38}
$$

From this, we see that the 2-RDM average also represents a generalization of a diagonal matrix to a 4-tensor with signs reflecting the antisymmetry properties of the electrons. Moreover, it is easy to see from the operator $\frac{1}{2} a_i^\dagger a_j^\dagger a_l a_k$ that the same concentration results hold for each individual matrix element as in the 1-RDM case, except that the Lipzschitz constant is modified by the normalization factor $1/2$. Using the same process, one may derive similar results for higher particle RDMs, and conclude that all higher RDMs concentrate towards diagonal matrices at a similar rate, with modified Lipschitz constants due to normalization.

Now we consider the case where one restricts to random states within an n-particle subspace. Such a state can be represented by a density matrix of the form

$$\rho_R = \frac{I_R}{d_R}, \tag{39}$$

where I_R is the identity operator on the full space of 2^M spin-orbitals subject to the restriction R to the space of n particles, and d is the dimension of that space. To evaluate this trace, we consider the basis of n-particle determinants. Using the same machinery as above, but restricting the sum to the case of only n particles, we find

$$
\begin{aligned}
\langle a_i^\dagger a_j \rangle_R &= \frac{\delta_{ij}}{d_R} \binom{M-1}{n-1} \\
&= \delta_{ij} \binom{M-1}{n-1} \binom{M}{n}^{-1} \\
&= \frac{n}{M}.
\end{aligned}
\tag{40}
$$

Thus the average 1-RDM in a space of randomly generated n-particle states in M orbitals is the diagonal matrix with equally probable occupations on all sites. The exact convergence results of Levy's lemma depend on the spherical geometry of states, so direct application of the concentration results would require a modification of the lemma. However, we may embed the allowable quantum states into the space of $\lceil \log_2(d_R) \rceil$ qubits and leveraging the fact that the norm of the operator of interest, and thus Lipschitz constant will remain the same. If one generates random quantum states within this embedded space, we can see that the concentration result holds with the same Lipschitz constant but modified dimension $\binom{M}{n}$.

A consequence of these results is that random quantum states generate p-particle marginals that are effectively trivial, concentrating exponentially quickly to their average value as a function of system size. This would mean that one can evaluate the expectation value to specified precision of any p-particle observable (where p is held fixed as system size grows) efficiently on a classical computer for a random quantum state. Thus, we conclude meaningful explorations of the space of quantum states must be structured, whether it be in the path of time evolution or the design of a variational ansatz. If an ansatz and method used for hybrid quantum–classical methods cannot easily exit the space of Haar random states, the above analysis dictates it is doomed to give trivial observables for fermionic systems at relatively small system sizes.

5. Reducing operator sample variance using n-representability constraints

5.1. Optimal operator averaging

Any L-sparse Hermitian operator on Hilbert space can be expressed as

$$H = \sum_{\ell=0}^{L-1} w_\ell H_\ell \qquad \text{s.t.} \quad w_\ell \in \mathbb{R} \qquad H_\ell^2 = \mathbf{1}, \tag{41}$$

where w_ℓ are real scalars and H_ℓ are 1-sparse self-inverse operators which act on qubits. In second quantized formulations of electronic structure, the H_ℓ are typically a special case of 1-sparse operators that have particularly convenient properties for measurement, namely they are strings of Pauli operators. Very often one is interested in estimating $\langle H \rangle$ by making measurements on M independent copies of a state $|\psi\rangle$. For instance, the typical procedure in variational algorithms is to estimate the energy $\langle H \rangle$ by repeatedly preparing a state and performing projective measurements onto the eigenstates of Pauli operators $\langle H_\ell \rangle$. Since the H_ℓ are self-inverse, the intrinsic variance of these projective measurements is computed as

$$\sigma_\ell^2 = \langle H_\ell^2 \rangle - \langle H_\ell \rangle^2 = 1 - \langle H_\ell \rangle^2 \leqslant 1. \tag{42}$$

As sample variance is due to statistical fluctuations which are uncorrelated from sample to sample, the total variance of $\langle H \rangle$ scales as

$$\epsilon = \sqrt{\sum_{\ell=0}^{L-1} \frac{w_\ell^2 \sigma_\ell^2}{M_\ell}} = \sqrt{\sum_{\ell=0}^{L-1} w_\ell^2 \left(\frac{1 - \langle H_\ell \rangle^2}{M_\ell} \right)}. \tag{43}$$

The real question is how to choose the number of samples for each term in the Hamiltonian M_ℓ in order to minimize ϵ for the fewest overall measurements $M = \sum_\ell M_\ell$.

In [48], it is suggested that one choose $M_\ell \propto |w_\ell|$ with no guarantee of optimality. Here we prove that this choice is optimal by application of the Lagrange conditions (no proof was provided in [48]). We start with the Lagrangian

$$\mathcal{L} = \sum_{\ell=0}^{L-1} M_\ell + \lambda \left(\sum_{\ell=0}^{L-1} \frac{w_\ell^2 \sigma_\ell^2}{M_\ell} - \epsilon^2 \right), \tag{44}$$

where the constant λ is the Lagrangian multiplier. Our goal will be to solve the following expression for M_ℓ,

$$\min_{M_\ell} \max_{\lambda} \mathcal{L} = \min_{M_\ell} M. \tag{45}$$

Accordingly, we take the derivative of $\mathcal{L}$ with respect to M_ℓ to find,

$$\frac{\partial \mathcal{L}}{\partial M_\ell} = \sum_{\ell=0}^{L-1} \left(1 - \lambda \frac{w_\ell^2 \sigma_\ell^2}{M_\ell^2} \right) = 0 \qquad \rightarrow \qquad M_\ell = \sqrt{\lambda} |w_\ell| \sigma_\ell. \tag{46}$$

Plugging this back into equation (43), we find exactly that

$$\epsilon^2 = \frac{1}{\sqrt{\lambda}} \sum_{\ell=0}^{L-1} |w_\ell| \sigma_\ell \qquad \rightarrow \qquad \sqrt{\lambda} = \frac{1}{\epsilon^2} \sum_{\ell=0}^{L-1} |w_\ell| \sigma_\ell. \tag{47}$$

Therefore,

$$M = \sum_{\ell=0}^{L-1} M_\ell = \left(\frac{1}{\epsilon} \sum_{\ell=0}^{L-1} |w_\ell| \sigma_\ell \right)^2 \leqslant \frac{\Lambda^2}{\epsilon^2} \qquad \Lambda = \sum_{\ell=0}^{L-1} |w_\ell|. \tag{48}$$

If we insist on getting an asymptotic bound then we assume $\sigma_\ell = \mathcal{O}(1)$ and this leads us to confirm the optimality of the suggestion of [48].

5.2. Reducing variance using *n*-representability

After mapping RDM elements to Pauli matrices (e.g. by using fermionic transforms such as the Jordan–Wigner or Bravyi–Kitaev transformation), one can express all equality *n*-representability constraints in the notation of equation (41). Assuming a list of K equality constraints, we will express the kth constraint C_k as

$$C_k = \sum_{\ell=0}^{L-1} c_{k,\ell} \langle H_\ell \rangle = 0 \qquad c_{k,\ell} \in \mathbb{R}, \tag{49}$$

where we can always choose to have a constant term in the set (e.g. $H_0 = \mathbf{1}$) so that the equality sums to zero[3]. These constraints provide extra information about the relationships between expectation values which, in principle, should allow us to make fewer measurements. One very straightforward way to exploit this information in order to make fewer measurements is to add these constraints to the operator of interest in order to minimize the associated Λ from equation (48). Specifically, we have that

$$\widetilde{H} = H + \sum_{k=0}^{K-1} \beta_k C_k = \sum_{\ell=0}^{L-1} \left(w_\ell + \sum_{k=0}^{K-1} \beta_k c_{k,\ell} \right) H_\ell \qquad \langle H \rangle = \langle \widetilde{H} \rangle \quad \forall \quad \beta_k \in \mathbb{R}, \tag{50}$$

where the relation $\langle H \rangle = \langle \widetilde{H} \rangle$ follows from the observation that $C_k = 0$ for *n*-representable states due to the definition of C_k in equation (49). When we do this, from equations (48) and (50) we can see that the number of measurements required is expected to scale as

$$\widetilde{M} = \left(\frac{1}{\epsilon} \sum_{\ell=0}^{L-1} \left(w_\ell + \sum_{k=0}^{K-1} \beta_k c_{k,\ell} \right) \right)^2 \leqslant \frac{\widetilde{\Lambda}^2}{\epsilon^2} \qquad \widetilde{\Lambda} = \sum_{\ell=0}^{L-1} \left| w_\ell + \sum_{k=0}^{K-1} \beta_k c_{k,\ell} \right|. \tag{51}$$

In order to minimize measurements then, the strategy is to compute

$$\beta^* = \underset{\beta}{\operatorname{argmin}} \left(\sum_{\ell=0}^{L-1} \left| w_\ell + \sum_{k=0}^{K-1} \beta_k c_{k,\ell} \right| \right) \qquad \text{or} \qquad \beta^* = \underset{\beta}{\operatorname{argmin}}(\widetilde{\Lambda}), \tag{52}$$

depending on whether or not one has any meaningful prior on the expectation values $\langle H_\ell \rangle$ (which would provide a meaningful prior on σ_ℓ via equation (42)).

We can easily recast this optimization problem in a form amenable to efficient solution by common numerical methods. To do this, we think of the original Hamiltonian H as being expressed as a vector v_H where each element of a vector represents a different fermionic operator; for example, we could map term $a_p^\dagger a_q$ to vector element $1 + p + q\,N$ and map $a_p^\dagger a_q^\dagger a_r a_s$ to $1 + N^2 + p + q\,N + r\,N^2 + s\,N^3$. The coefficients of the vector correspond to the coefficients of the term. Likewise, we can represent all of the constraints in a matrix C of dimension $K \times L$ where each constraint C_k is a row of the matrix vectorized in the same way as v_H. Then, we see that the optimization task at hand can be expressed as

$$\beta^* = \underset{\beta}{\operatorname{argmin}} \; \| v_H - C^\top \beta \|_1 \tag{53}$$

where β is a vector of dimension K. We can see now that this is a convex L_1 minimization. Such minimizations can be solved efficiently using simplex methods. We can cast L_1 minimization as the linear program:

$$\text{minimize} \quad \mathbf{1}^\top q \qquad \text{subject to} \quad -q \leqslant v_H - C^\top \beta \leqslant q, \tag{54}$$

where q is an auxiliary vector variable. We provide freely available source code that generates the equality constraints and performs this optimization in the open source project OpenFermion [90]. Our code uses GLPK (GNU Linear Programming Kit) for the linear programming component via a Python wrapper known as CVXOPT. We show results of several numerical experiments which demonstrate the effectiveness of our method in figure 1. These numerical experiments involved linear *n*-representability constraints coming from the 1-RDM trace, 1-RDM Hermiticity, 2-RDM trace, 2-RDM Hermiticity, 2-RDM to 1-RDM contraction, and the mappings between the 2-RDM and the other marginals in the 2-positive set. The set of mappings are described in appendix C. These techniques often reduce the required measurements by an order of magnitude or more.

Note that the method discussed here is actually quite a bit more general than presented. In particular, we have found an interesting method for transforming the Hamiltonian in a way that leaves its spectrum invariant in the fixed particle number sector. One might postulate that this method could also be used to optimize other simulation metrics, for instance, to reduce Trotter errors which are well known to be related to the norm of the Hamiltonian. Note that all constraints C_k will be either Hermitian or anti-Hermitian operators. In particular, the Hermiticity constraints take the form of constraining anti-Hermitian components of the density matrix to be zero (thus those C_k are themselves anti-Hermitian operators). So after applying the procedure here, one may end

---

[3] Throughout this section, the H_ℓ should be the full set of terms that map to qubits from the 2-RDM. Terms that do not appear in the operator of interest (for instance, 2-RDM elements that are not in the Hamiltonian) simply have a coefficient of zero.
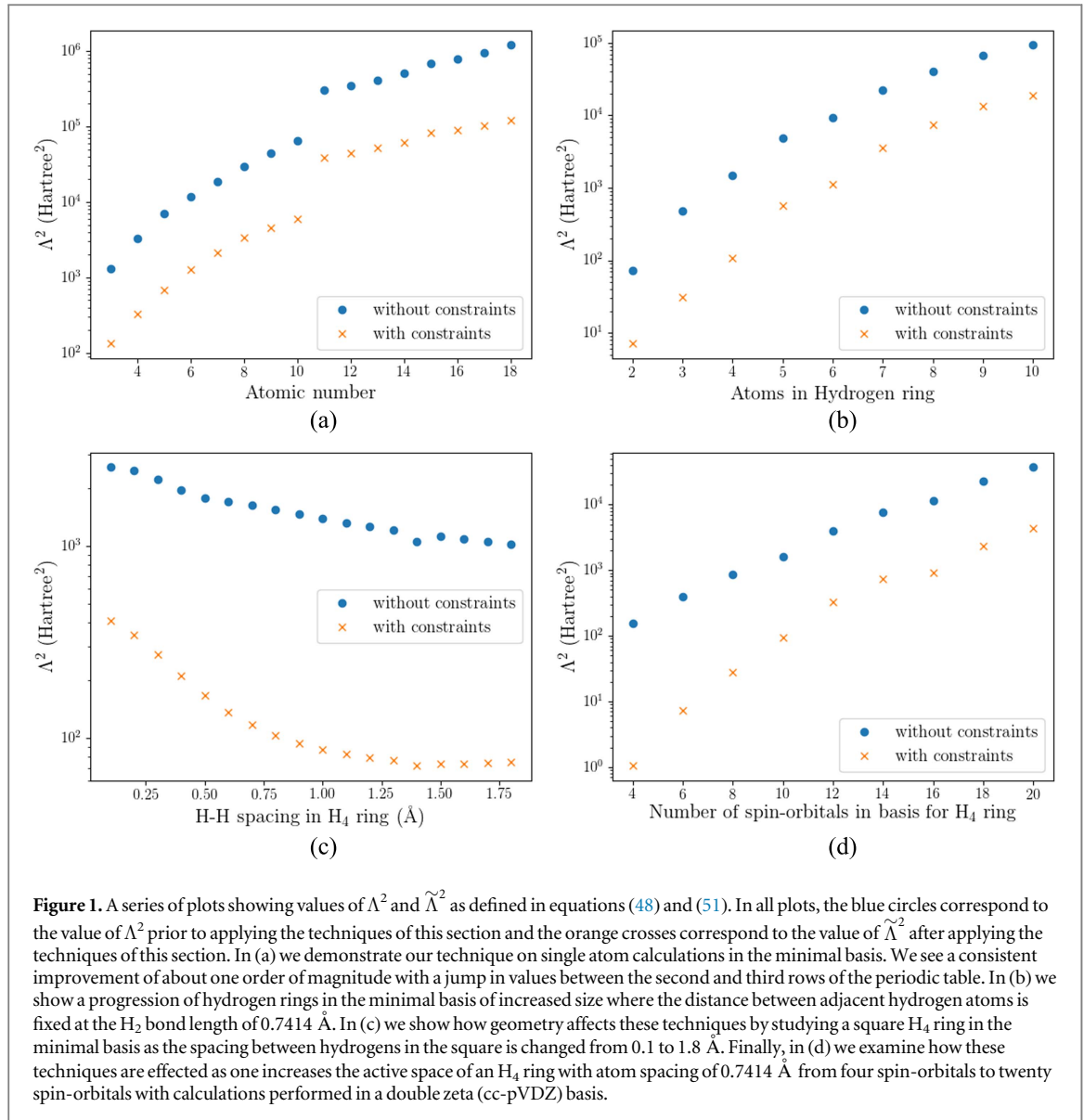
Figure 1. A series of plots showing values of Λ^2 and $\widetilde{\Lambda}^2$ as defined in equations (48) and (51). In all plots, the blue circles correspond to the value of Λ^2 prior to applying the techniques of this section and the orange crosses correspond to the value of $\widetilde{\Lambda}^2$ after applying the techniques of this section. In (a) we demonstrate our technique on single atom calculations in the minimal basis. We see a consistent improvement of about one order of magnitude with a jump in values between the second and third rows of the periodic table. In (b) we show a progression of hydrogen rings in the minimal basis of increased size where the distance between adjacent hydrogen atoms is fixed at the H_2 bond length of 0.7414 Å. In (c) we show how geometry affects these techniques by studying a square H_4 ring in the minimal basis as the spacing between hydrogens in the square is changed from 0.1 to 1.8 Å. Finally, in (d) we examine how these techniques are effected as one increases the active space of an H_4 ring with atom spacing of 0.7414 Å from four spin-orbitals to twenty spin-orbitals with calculations performed in a double zeta (cc-pVDZ) basis.

up with a $\widetilde{H}$ that is not Hermitian. Fortunately, one can restore Hermiticity without changing the value of $\widetilde{\Lambda}$ by creating a new Hamiltonian, $H^\star = (\widetilde{H} + \widetilde{H}^\dagger)/2$. $H^\star$ will be isospectral to H in the n-electron manifold and will have the same $\widetilde{\Lambda}$ as $\widetilde{H}$.

It remains an open question if this variance reduction technique can be applied to other hybrid algorithms such as QAOA or quantum spin Hamiltonians. In the QAOA case, spin-marginals with significantly less structure than fermionic marginals must be considered. Linear constraints outside of the consistency of the marginals with overlapping support are generally unknown for an arbitrary Hamiltonian encoding a combinatorial optimization problem as is done in QAOA. For spin Hamiltonians constraints generated by fixed values of $\langle S^2 \rangle$ can also be considered. Further investigation of the consistency constraints and eigenvalue constraints in the form of pure-state constraints may provide additional variance reduction for problems described by fermionic and spin Hamiltonians.

6. *n*-Representability informed projection of 2-RDMs

In this section we discuss the possibility of using n-representability conditions to improve 2-RDMs sampled from a quantum device. Errors in the 2-RDM measured from a quantum state can appear in multiple ways: (1) stochastic errors associated with the operator averaging techniques used to measure expected values and (2) device errors such as unexpected measurement correlations. We explored the utility of 2-marginal reconstruction schemes using n-representability rules to remove stochastic errors associated with sampling and state errors corresponding to noise of the device corrupting the intended state.

The general strategy is to design a process that projects a 2-RDM back into the set of n-representable 2-RDMs while balancing data-collection time and classical post-processing time. In this section we first discuss two simple purification procedures: positive-semidefinite projection of the measured 2-RDM with and without fixed-trace. These simple projection techniques are compared against procedures involving projections with knowledge of representability constraints.

6.1. Positive-semidefinite projection and positive-semidefinite projection with fixed trace

The simplest of the n-representability rules enforce the 2-RDM to be Hermitian and non-negative with fixed trace. Given a measured 2-RDM we can define a computational procedure that determines the closest positive-semidefinite matrix

$$\min ||^2D - {}^2D_{\text{measured}}||_2 \tag{55}$$

$$\text{s.t. } {}^2D \succcurlyeq 0 \ \text{Tr}[^2D] = n(n-1). \tag{56}$$

The normalization is fixed by the particle number of the system. Without the trace condition the 2-RDM that minimizes the objective in equation (55) is the marginal constructed from the non-negative eigenvalues and eigenvectors of ${}^2D_{\text{measured}}$ [91]. The procedure for finding a fixed-trace positive-semidefinite projection have appeared in contexts such as tomography [92], iterative purification of 2-RDMs from response theory [93], and finding positive-semidefinite correlation matrices [91, 94]. This projection procedure benefits from computational simplicity but suffers from the lack of information about representability conditions. Therefore, given a sufficiently corrupted ${}^2D_{\text{measured}}$, physicality is not guaranteed after projection.

6.2. RDM reconstruction with approximate representability constraints

In order improve the projection criteria we add additional n-representability constraints to the minimization procedure outlined in equation (55). Given a collection of 2-RDM elements at some unknown precision, or possibly missing crucial elements, our reconstruction scheme seeks to minimize the Frobenius norm of the difference between the reconstructed 2-RDM and the set of known measurements subject to approximate n-representability constraints. Denoting E to be the difference between the reconstructed 2-RDM and ${}^2D_{\text{measured}}$ the minimization procedure can be formulated as the following non-convex optimization problem:

$$\min ||E||_F^2 \tag{57}$$

$$\text{s.t.} \text{Tr}[^2D] = n(n-1)$$
$$\{^1D, {}^1Q, {}^2D, {}^2Q, {}^2G\} \succcurlyeq 0$$
$$A_1(^1D) \rightarrow {}^1Q, \ A_2(^2D) \rightarrow {}^1D$$
$$A_3(^2D) \rightarrow {}^2Q, \ A_4(^2D) \rightarrow {}^1G, \tag{58}$$

where A_i is the map from one marginal to another required by the fermionic ladder operator algebra. The details of these mappings can be found in appendix A. The squared Frobenius norm of the error $||E||_F$ is quadratic in 2-RDM. The optimization problem specified in equation (57) can be relaxed to a SDP by taking the Schur complement in the identity block of the large matrix M

$$M = \begin{pmatrix} I & E \\ E^\dagger & F \end{pmatrix} \succcurlyeq 0 \tag{59}$$

constrained to be positive-semidefinite. In M, I is the identity matrix, F is a matrix of free variables, and E is the error between the reconstructed ${}^2D$ and the ${}^2D_{\text{measured}}$. Taking the Schur complement in the identity block of M gives

$$F - E^\dagger E \succcurlyeq 0 \tag{60}$$

or

$$F \succcurlyeq E^\dagger E. \tag{61}$$

Noting that the Frobenius norm of a matrix A, $||A||_F$, is given as $\sqrt{\text{Tr}[A^\dagger A]}$, taking the trace of equation (61) gives the semidefinite relaxation of minimizing the Frobenius norm

$$\text{Tr}[F] \succcurlyeq \text{Tr}[E^\dagger E], \tag{62}$$

$$||E||_F^2 = \text{Tr}[E^\dagger E]. \tag{63}$$

We can now formulate the non-convex RDM reconstruction scheme in equation (57) in terms of a SDP:

$$\min \text{Tr}[F] \tag{64}$$

$$\text{s.t.}\{^1D,\ ^1Q,\ ^2D,\ ^2Q,\ ^2G,\ M\} \succcurlyeq 0$$
$$A_1(^1D) \rightarrow\ ^1Q,\quad A_2(^2D) \rightarrow\ ^1D$$
$$A_3(^2D) \rightarrow\ ^2Q,\quad A_4(^2D) \rightarrow\ ^1G$$
$$A_5(^2D) \rightarrow\ M,\quad A_6(^2D) \rightarrow n(n-1)$$
$$A_7(^2D) \rightarrow\ ^2D, \tag{65}$$

where A_i are the linear maps between ^2D and the other matrices $\{^1D,^1Q,^2D,^2Q,^2G,M\}$ along with the trace constraint and antisymmetry constraint on ^2D. These maps are described explicitly in appendix C and are used in the section 6.4 for the SDP reconstruction program.

6.3. Iterative procedure for projecting noisy 2-RDMs into the approximate n-representable subspace

Although the SDP projection procedure can be extended to include better approximate n-representability conditions, it suffers from the requirement of solving a SDP. Despite the fact that an SDP can be solved in polynomial time with respect to the total number of variables and constraints, the high-order polynomial scaling of SDP algorithms makes the SDP-based project method infeasible for on-the-fly or online projections. An alternative to the SDP projection combines the faster projection techniques based on fixed-trace positive projection with augmented n-representability conditions. The projection technique is an iterative procedure that was originally developed to enforce approximate n-representability on 2-RDMs obtained through a response formalism [93]. The algorithm involves sequentially mapping ^2D to ^2Q to ^2G and enforcing the positivity and trace constraints at each of the operators. The algorithm's main drawback is that any representable 2-RDM is a valid fixed point. As a result, linear constraints on the 2-RDM preserving projected spin and total spin expectation values are not enforced and there is no guarantee that the ^2D obtained from the iterative procedure is any closer to the true 2-RDM. Therefore, it is likely required that the input 2-RDM measured from the quantum resources is sufficiently close to the true 2-RDM for this procedure to be most successful.

The algorithm starts by enforcing Hermiticity of the given 2-RDM matrix by averaging

$$^2D^s = \frac{1}{2}(^2D_{\text{meas}} +\ ^2D^\dagger_{\text{meas}}), \tag{66}$$

followed by a positive projection with fixed trace according to the procedure in [94]. A detailed description of the algorithm for fixed-trace positive projection can also be found in the appendix of [93]. Given a system with r spin orbitals, n particles, and $\eta = r - n$ holes, the iterative projection algorithm is as follows:

1. Enforce Hermiticity of ^2D and project to positive set with trace $n(n-1)$.

2. Map ^2D to the ^2Q.

3. Enforce Hermiticity of ^2Q and project to positive set with trace $\eta(\eta-1)$ where η is the number of holes.

4. Map ^2Q to ^2G.

5. Enforce Hermiticity of ^2G and project to the positive set with trace $n(\eta+1)$.

6. Check the stopping condition associated with fixed trace for ^2D, ^2Q, and ^2G and positivity of their eigenvalues.
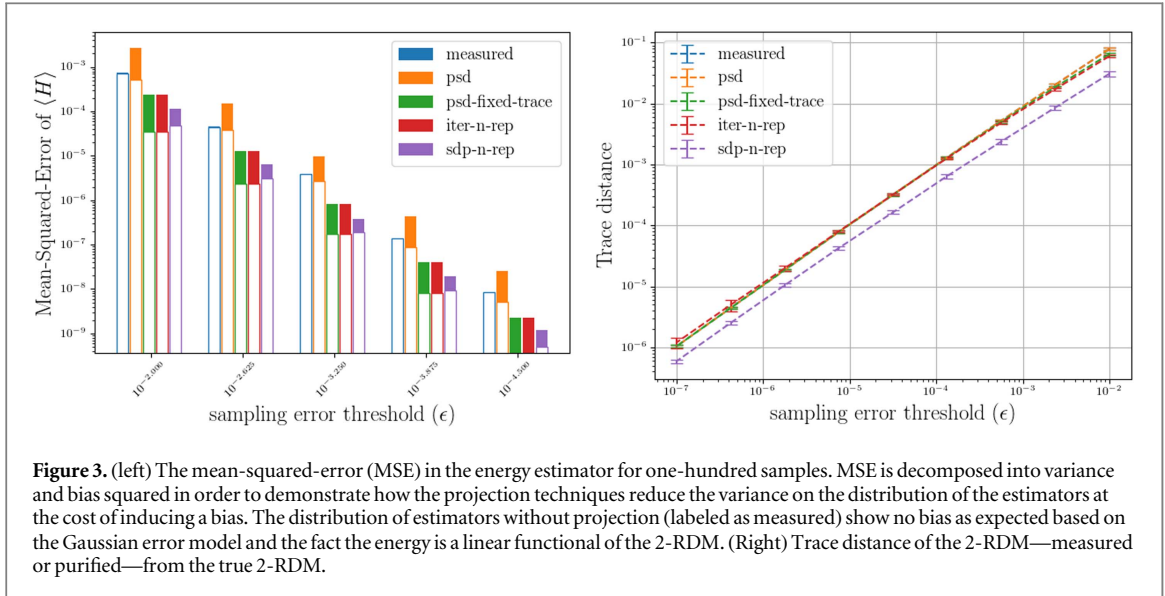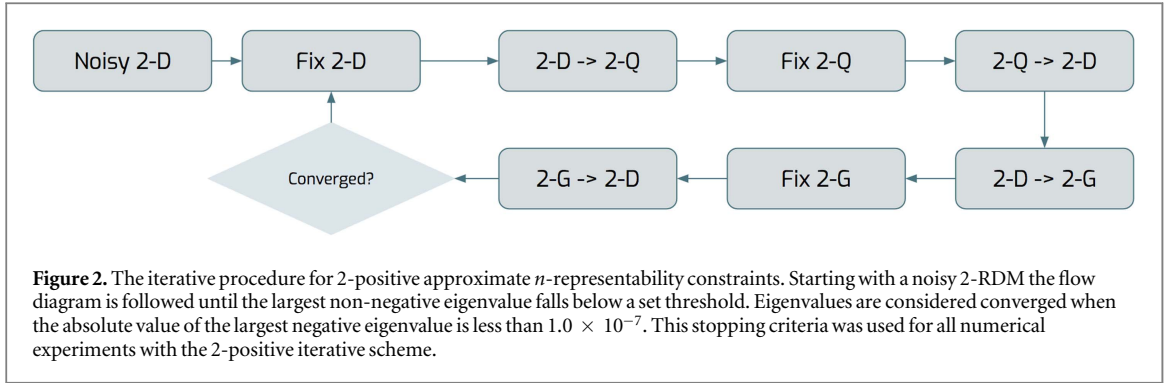
The iterative procedure is considered converged when the largest negative eigenvalue of any marginal in the 2-positive set is below a set threshold. The total algorithm is depicted in figure 2.

6.4. Reconstruction results

6.4.1. Reconstruction of small systems

To probe the utility of the n-representability inspired reconstruction schemes, we examined the accuracy of the energy and chemical properties obtained from 2-RDMs with simulated sampling noise for diatomic hydrogen and a linear four-hydrogen chain. All experiments involved corrupting the elements of a pure-state 2-RDM with Gaussian noise proportional to the amount of samples used in operator averaging, followed by reconstructing the corrupted marginal with the four projection procedures outline above. The accuracy and precision of the reconstructed energies, particle-number, projected spin expectation $\langle S_z \rangle$, and total spin $\langle S^2 \rangle$ are compared to provide the noise tolerance and precision of the various reconstruction schemes.

We obtained the Hamiltonian and ground state wavefunction for diatomic hydrogen and a linear four-hydrogen chain with a bond length of 0.75 Å described with an STO-3G basis using the OpenFermion [90] and the OpenFermion-Psi4 plugin [95]. One hundred different corrupted RDMs were constructed by applying zero-mean Gaussian noise with variance ϵ^2
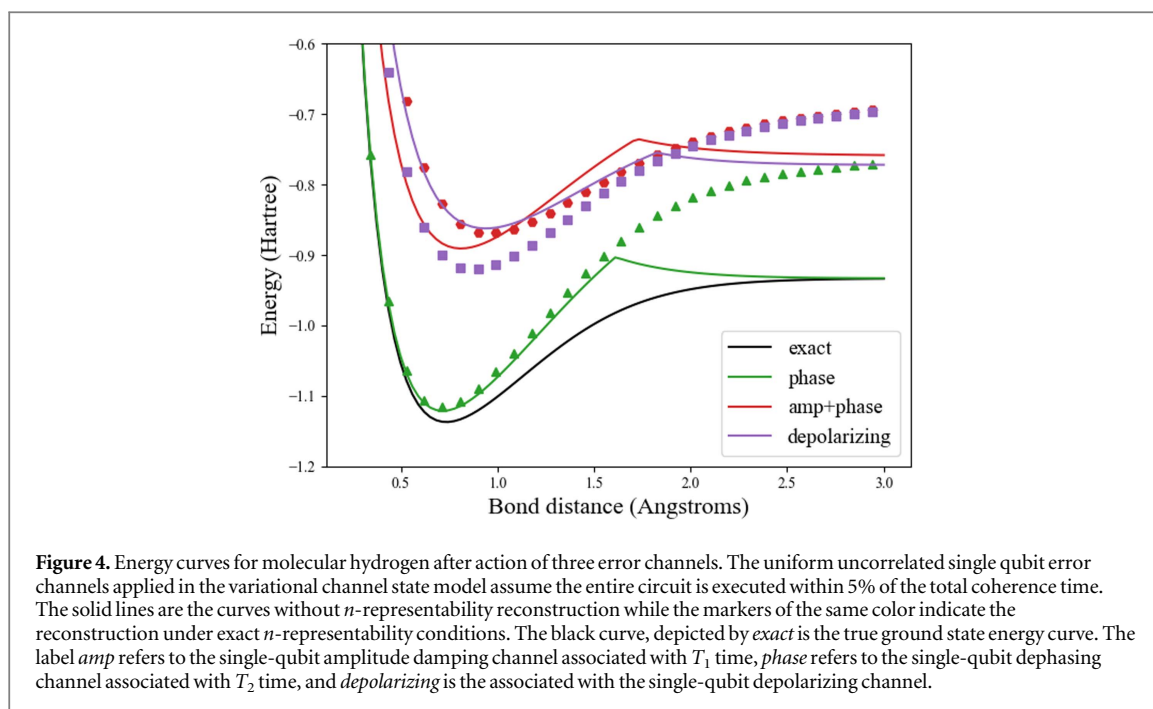
Figure 2. The iterative procedure for 2-positive approximate *n*-representability constraints. Starting with a noisy 2-RDM the flow diagram is followed until the largest non-negative eigenvalue falls below a set threshold. Eigenvalues are considered converged when the absolute value of the largest negative eigenvalue is less than 1.0×10^{-7}. This stopping criteria was used for all numerical experiments with the 2-positive iterative scheme.

Figure 3. (left) The mean-squared-error (MSE) in the energy estimator for one-hundred samples. MSE is decomposed into variance and bias squared in order to demonstrate how the projection techniques reduce the variance on the distribution of the estimators at the cost of inducing a bias. The distribution of estimators without projection (labeled as measured) show no bias as expected based on the Gaussian error model and the fact the energy is a linear functional of the 2-RDM. (Right) Trace distance of the 2-RDM—measured or purified—from the true 2-RDM.

$$^{2}D^{pq}_{rs\,\text{measured}} = \,^{2}D^{pq}_{rs} + \mathcal{N}(0, \epsilon^{2}). \tag{67}$$

This error model is bias-free because the energy is linearly proportional to the 2-RDM and has variance proportional to the error added to the 2-RDM elements. For each of the one hundred corrupted density matrices we solve for a projected 2-RDM with the positive projection, positive projection with fixed-trace, SDP *n*-representability reconstruction, and iterative *n*-representability projection techniques. For each method we find the mean-square-error (MSE) of the aforementioned observables over the projected 2-RDMs as a function of the noise parameter ϵ.

Figure 3 contains a plot of the MSE of the energy estimator for H_2 decomposed into its variance and bias components, and a plot of the average trace distance of the reconstructed 2-RDMs from the true 2-RDM of H_4. The solid bars in the MSE plot are the squared bias component of the MSE while the transparent bars are the variance component. In general, the *n*-representability inspired projection techniques decrease the variance of the energy estimator but introduce a bias. Similar MSE plots (figures 5 and 6) are shown for $\langle S^2 \rangle$, $\langle S_z \rangle$, and $\langle n \rangle$ in appendix E. The expected value for S_z, S^2, and n shows zero MSE for the SDP-based projection technique because these values are added as constraints to the SDP. We refer to the correction of the three aforementioned expected values as restoration of physicality—i.e. the particle number expectation is what is expected for an isolated system. The reduced trace distance for the SDP projected 2-RDMs for H_4 indicates that the physicality constraints are important for removing errors from 2-RDMs measured from a quantum resource.

6.4.2. Reconstruction of marginals from the variational channel state model

One appealing application of projection techniques based on *n*-representability is purification of states corrupted by an error channel. The SDP *n*-representability reconstruction procedure ensures physicality of the states by ensuring known symmetries are preserved by formulating the projection as a constrained optimization. To verify this we used the SDP *n*-representability method in conjunction with the variational channel state error models presented in [54] to demonstrate the restoration of physicality. The variational channel state model is implemented by corrupting a pure state $|\psi\rangle$ with a channel described in Kraus operator form. We consider uniform uncorrelated single-qubit error channels associated with dephasing, amplitude damping and

Figure 4. Energy curves for molecular hydrogen after action of three error channels. The uniform uncorrelated single qubit error channels applied in the variational channel state model assume the entire circuit is executed within 5% of the total coherence time. The solid lines are the curves without n-representability reconstruction while the markers of the same color indicate the reconstruction under exact n-representability conditions. The black curve, depicted by *exact* is the true ground state energy curve. The label *amp* refers to the single-qubit amplitude damping channel associated with T_1 time, *phase* refers to the single-qubit dephasing channel associated with T_2 time, and *depolarizing* is the associated with the single-qubit depolarizing channel.

dephasing, and depolarizing noise. The dephasing and amplitude damping channels are parameterized with the assumption that 5% of the coherence time has elapsed with respect to T_1 and T_2. For the depolarizing channel, the Kraus operators are constructed assuming 5% of the dephasing time T_2 has elapsed. For each point along the binding curve of diatomic hydrogen the action of the channel on the pure-state is calculated as

$$\rho_{\text{channel}} = \sum_{i=1}^{M} K_i \rho_{\text{pure}} K_i^{\dagger}. \tag{68}$$

Each ensemble 2-RDM is then reconstructed with 2-positive n-representability conditions using the SDP projection technique where the error matrices are set as the spin-adapted components of the 2-RDM associated with ρ_{channel}. Spin adapting eliminates the need to explicitly enforce the antisymmetry of the 2-RDM elements and thus reduces the total number of constraints in the SDP.

The energy of the the H_2 system under the action of each separate channel is plotted in figure 4 along with the energy computed as a functional of the reconstructed 2-RDM. The kinks in the dephasing and amplitude + dephasing curves are associated with a spin-symmetry breaking, where the channels produced a mixed state dominated by a triplet state. The discontinuity in the depolarizing channel curve is associated with the channel state switching to be a mixed state with a large component of singlet character. The markers without a line in figure 4 are the reconstructed 2-RDM with $\langle S_z \rangle = 0$ and $\langle S^2 \rangle = 0$ imposed by linear constraints on the 2-RDM associated with S-representability [96]. Naturally the binding curves are now smooth as a function of bond distance. Though physicality is recovered by projecting onto the closest marginal with fixed symmetries, the energy increases at distances greater than 1.5 Å for each error model and the potential energy minimum is shifted by -0.03 Å for the dephasing channel, 0.062 Å for the dephasing and relaxation channel, and 0.186 Å for the depolarizing channel. The energy increase is due to constraining the projected 2-RDM to have the correct spin-symmetry when the the error channel, applied through the variational channel state model, has switched the ground state symmetry from a singlet to a triplet [54]. The increase of energy upon restoration of a symmetry is a well known effect in chemical systems and condensed matter systems. However, the qualitative improvement in the nuclear potential energy surface and the implications for forces derived from such a surface are more important than the energy increase that is incurred. The relative error between corrected 2-RDM and the uncorrected 2-RDM at the maximum bond distance considered in this work (3.0 Å) is 18% for the dephasing channel, 8.7% for the dephasing and relaxation channel, and 9.9% for the depolarizing channel.

7. Use cases of RDMs in augmenting energy expectation through perturbation theory

In this section, we briefly mention some of the additional use cases within chemistry for correcting the RDMs beyond simple estimation of the energy, and review techniques that allow one to utilize them with only the 2-RDM. In traditional approaches to electronic structure on classical computers, the solution of the electronic structure problem within an active space often lacks the contributions from so-called dynamical correlation.

This part of the correlation stems largely from the electronic cusp contributions and represents a low-rank interaction in many high-lying orbitals. Multi-reference perturbation theory techniques have been found to offer a good balance between cost and accuracy for including these contributions, and many approaches have been developed in this regard including complete active space second order perturbation theory (CASPT2) [97–99], multi-reference Moller–Plesset theory [100], n-electron valence perturbation theory (NEVPT2) [101–103], canonical transformation theory (CT) [104–106], perturbative explicitly correlated corrections ($[2]_{R12}$) [107, 108], and corrected multi-reference CI (MRCI+Q) [109, 110].

In understanding how these techniques may be utilized within the quantum domain, one is most interested in those that are compatible with the measurements one expects to be able to feasible make on a quantum computer. This means that explicit knowledge of the determinant decomposition of a wavefunction cannot be required, and instead one prefers methods that require only the k-RDM of the electronic system, where hopefully k is small. Of those discussed, NEVPT2, $[2]_{R12}$, and CT can be applied in this way.

For example, in NEVPT2, the effective equations may be derived entirely with the use of the 4-RDM, and it does not require knowledge of determinant decomposition or full electronic wavefunction. However the 4-RDM is a relatively expensive quantity to estimate, and thus cumulant based approximations to the 4-RDM have been developed such that only the 2-RDM is required to correct the energy perturbatively [106, 111].

The cumulant expansions decompose the RDMs into their non-separable (connected) components and separable unconnected components, and are quite useful for both developing approximations and enhancing understanding. A convenient notation for expressing these expansions is given by the Grassmann wedge product defined generally by

$$a \wedge b = \left(\frac{1}{N!}\right)^2 \sum_{\pi, \sigma} \epsilon(\pi) \epsilon(\sigma) \, \pi \, \sigma \, a \otimes b, \tag{69}$$

where π and σ are permutations on the upper and lower indices of the tensor $a \otimes b$ and ϵ denotes the parity of each permutation. As an example one might consider the wedge product of a cumulant matrix with itself

$$[^1\Delta \wedge {}^1\Delta]^{i_1 i_2}_{j_1 j_2} = \frac{1}{2}({}^1\Delta^{i_1}_{j_1} {}^1\Delta^{i_2}_{j_2} - {}^1\Delta^{i_1}_{j_2} {}^1\Delta^{i_2}_{j_1}). \tag{70}$$

With this notation, the RDMs up to $k = 4$ may be expressed in terms of the cumulant expansions:

$$^1D = {}^1\Delta, \tag{71}$$

$$^2D = {}^2\Delta + {}^1\Delta \wedge {}^1\Delta, \tag{72}$$

$$^3D = {}^3\Delta + 3\,{}^2\Delta \wedge {}^1\Delta + {}^1\Delta \wedge {}^1\Delta \wedge {}^1\Delta, \tag{73}$$

$$^4D = {}^4\Delta + 4\,{}^3\Delta \wedge {}^1\Delta + 3\,{}^2\Delta \wedge {}^2\Delta \\ + 6\,{}^2\Delta \wedge {}^1\Delta \wedge {}^1\Delta + {}^1\Delta \wedge {}^1\Delta \wedge {}^1\Delta \wedge {}^1\Delta. \tag{74}$$

These methods that neglect contributions from the 3- and 4-cumulant dramatically reduce the number of samples that would be required to estimate the energy, however the approximations introduce some error in the corrections. An additional factor is the consideration of the impact of measurement noise on the RDM in these numerical procedures. While the purification techniques suggested in this draft are expected to mitigate some problems in this regard, the effect on an iterative procedure may be dramatic. For this reason, NEVPT2 and its approximations may be the preferred method for use with quantum computers. The strongly contracted equations that explicitly define the corrections to the energy are given in appendix A of [102], and cumulant reconstruction methods may be used directly to form approximations as dictated in [111]. The above equations can be used to derive cumulant based approximations for up to the 4-RDM from using only the 2-RDM by setting the irreducible three- and four-particle components Δ^3 and Δ^4 to 0. Alternatively, more sophisticated approximation schemes have been developed in the context of RDM theory within traditional quantum chemistry [112, 113].

8. Conclusion

Reducing the number of experiments required in the partial tomography step of VQE and other hybrid algorithms is necessary for hybrid classical/quantum algorithms to become useful simulation tools. In this work we proposed using representability conditions on fermionic marginals as a route towards reducing the number of measurements required in a VQE operator averaging step. Directly measuring the marginals provides the information required to integrate the results from hybrid quantum algorithms with classical quantum simulation methods.

Fermionic representability conditions were used in two majors ways: (1) re-expressing the Pauli sum Hamiltonian in a form where the total variance is minimized given a fixed state and (2) designing projection techniques based on necessary conditions on 2-marginals. Both techniques have shown significant promise

toward minimizing the total number of measurements and reducing stochastic noise seen from sampling the quantum resources.

The SDP projection techniques are especially attractive because they are constructed to return physical states. We have observed the restoration of physicality when a pure-state is corrupted by single-qubit error channels. More significant representability conditions based on positivity of elements of the 3-RDM may provide more accurate reconstruction under noise. Another area that is especially exciting is the application of pure state constraints in the reconstruction procedure. The conditions discussed in this work do not constrain the marginal to be integrated from a pure state. As a result, the positivity constraints likely do little for the preservation of physicality as compared to the equality representability constraints, such as fixed number operator and constrained total spin. Pure-states could potentially further reduce systematic errors associated with gates.

We fully expect that further investigation of representability conditions with realistic systems, more realistic error models, and performant numerical implementations will demonstrate the utility of measuring marginals for quantum simulation using hybrid classical/quantum algorithms.

Acknowledgments

Appendix A. Structure of RDMs

The SDP for reconstructing a noisy 2-RDM has significant block diagonal structure for chemical problems. The block structure of the 2-RDM reflects the symmetries of the Hamiltonian. Therefore, the 2-RDM can be blocked according to $\langle \hat{S}_z \rangle$, $\langle S^2 \rangle$, $\langle n \rangle$, and any spatial symmetry groups. For example, time-reversal symmetry implies spin-adapted 1-RDM when a position space basis is used and provides additional constraints when momentum is a good quantum number. For Gaussian basis sets commonly used in quantum chemistry the time-reversal invariant spin-free quantum chemical Hamiltonian implies the following block structure: the total 1-RDM and 1-Hole-RDM can be blocked into α and β spin blocks

$$^1D = \begin{pmatrix} ^1D_\alpha & 0 \\ 0 & ^1D_\beta \end{pmatrix} \tag{A1}$$

and

$$^1Q = \begin{pmatrix} ^1Q_\alpha & 0 \\ 0 & ^1Q_\beta \end{pmatrix} \tag{A2}$$

while the two particle matrices can be blocked as follows

$$^2D = \begin{pmatrix} ^2D_{\alpha,\alpha}^{\alpha,\alpha} & 0 & 0 & 0 \\ 0 & ^2D_{\beta,\beta}^{\beta,\beta} & 0 & 0 \\ 0 & 0 & ^2D_{\alpha,\beta}^{\alpha,\beta} & 0 \\ 0 & 0 & 0 & ^2D_{\beta,\alpha}^{\beta,\alpha} \end{pmatrix}, \tag{A3}$$

where the blocks $^2D_{\alpha,\alpha}^{\alpha,\alpha}$, $^2D_{\beta,\beta}^{\beta,\beta}$, $^2D_{\alpha,\beta}^{\alpha,\beta}$ and $^2D_{\beta,\alpha}^{\beta,\alpha}$ have linear sizes of r_s^2. In practice we can reduce the size of the $^2D_{\alpha,\alpha}^{\alpha,\alpha}$ and $^2D_{\beta,\beta}^{\beta,\beta}$ blocks by noting that these tensors are spanned by (r_s choose 2) basis antisymmetric functions instead of r_s^2 symmetric functions. The $^2D_{\beta,\alpha}^{\beta,\alpha}$ block is removed as it can be mapped in a one-to-one fashion to the $^2D_{\alpha,\beta}^{\alpha,\beta}$ block rendering it redundant. The corresponding blocks in the Hamiltonian correspond to the antisymmeterized integrals. The ^2Q matrix has the same block structure. The ^2G matrix has slightly less block structure

$$
{}^2G = \begin{pmatrix} {}^2G^{\alpha,\alpha}_{\alpha,\alpha} & {}^2G^{\beta,\beta}_{\alpha,\alpha} & 0 & 0 \\ {}^2G^{\alpha,\alpha}_{\beta,\beta} & {}^2G^{\beta,\beta}_{\beta,\beta} & 0 & 0 \\ 0 & 0 & {}^2G^{\alpha,\beta}_{\alpha,\beta} & 0 \\ 0 & 0 & 0 & {}^2G^{\beta,\alpha}_{\beta,\alpha} \end{pmatrix},
\tag{A4}
$$

where the block sizes are $2r_s^2$, r_s^2, and r_s^2 respectively.

Appendix B. Observables from the 2-RDM and 1-RDM

In this work we make use of the fact that a number of important observables are linear functionals of the 2-RDM and 1-RDM. In this section we enumerate these relationships for clarity.

The energy can of a chemical Hamiltonian can be expressed as a linear function of the 1-RDM and 2-RDM. Consider a general chemical Hamiltonian in second quantization

$$
H = \sum_{ij} h_{ij} a_i^\dagger a_j + \frac{1}{2} \sum_{pqrs} V_{pqrs} a_p^\dagger a_q^\dagger a_s a_r,
\tag{B1}
$$

where h_{ij} and V_{pqrs} are the one- and two-electron integral tensors. When evaluating the expected value of the Hamiltonian $\langle H \rangle$ the dependence on the 1- and 2-RDM naturally emerges

$$
\begin{aligned}
\langle H \rangle &= \sum_{ij} h_{ij} \langle a_i^\dagger a_j \rangle + \frac{1}{2} \sum_{pqrs} V_{pqrs} \langle a_p^\dagger a_q^\dagger a_s a_r \rangle \\
&= \sum_{ij} h_{ij} \, {}^1D^i_j + \frac{1}{2} \sum_{pqrs} V_{pqrs} \, {}^2D^{pq}_{rs}.
\end{aligned}
\tag{B2}
$$

For the calculations depicted in figure 4 the total angular momentum S^2, projected angular momentum S_z, and the particle number n operators are used as linear constraints in the SDP. Just like the energy, these operators are linear functionals of the 1- and 2-RDMs. To see this we express each component of the aforementioned operators as sums of fermionic operators resulting in polynomials of rank-4 and rank-2

$$
n = \sum_{i=1}^m a_i^\dagger a_i,
\tag{B3}
$$

$$
S_z = \frac{1}{2} \sum_{i=1}^{m/2} (a_{i,\alpha}^\dagger a_{i,\alpha} - a_{i,\beta}^\dagger a_{i,\beta}),
\tag{B4}
$$

$$
S^2 = S^- S^+ + S_z^2 + S_z,
\tag{B5}
$$

where

$$
S^- = \sum_{i=1}^{m/2} a_{i,\beta}^\dagger a_{i,\alpha},
\tag{B6}
$$

$$
S^+ = \sum_{i=1}^{m/2} a_{i,\alpha}^\dagger a_{i,\beta},
\tag{B7}
$$

m is the total number of spin orbitals, and α (β) denotes the two eigenfunctions of the z-angular momentum operator for a single fermion. The expected value of each operator can be determined by summing over the indicated elements of the 2-RDM and 1-RDM.

Appendix C. Mapping conditions and trace conditions

The linear constraints in the SDP-projection SDP include a trace constraint on the 2-RDM, mappings between the 1-RDM, 1-hole-RDM, 2-RDM and 2-hole-RDM, and the 2-RDM to the 2-particle-hole-RDM. These mappings between the marginals can be derived by rearranging the fermionic ladder operators resulting in the following matrix element equalities:

$$
{}^1D^p_q + {}^1Q^q_p = \delta_{p,q},
\tag{C1}
$$

$$
\begin{aligned}
{}^2D^{p,q}_{r,s} = {}^1D^p_r \delta^q_s + {}^1D^q_s \delta^p_r - ({}^1D^p_s \delta^q_r + {}^1D^q_r \delta^p_s) \\
- \delta^p_s \delta^q_r + \delta^p_r \delta^q_s + {}^2Q^{r,s}_{p,q},
\end{aligned}
\tag{C2}
$$

$$
{}^2G^{p,q}_{r,s} = \delta^q_s \, {}^1D^p_r - {}^2D^{p,s}_{q,r}.
\tag{C3}
$$

The contraction relation between the elements of the 2-RDM and 1-RDM

$$\sum_i {}^2D_{r,i}^{p,i} = (n-1)\, {}^1D_r^p \qquad (C4)$$

is included in the set of linear constraints.

Appendix D. Computational implementation of the reconstruction problem

The reconstruction problem outlined in section 6.2 is formulated as a SDP. Unique to this program is the sparsity of each constraint relative to the total number of variables in the program. A class of SDP solvers using the augmented Lagrangian technique have been shown to efficiently solve SDPs of this form in quantum chemistry and condensed matter [82, 85, 87, 114–117]. The primal SDP is mathematically stated as

$$\min \langle C, X \rangle \qquad (D1)$$

$$\text{s.t} \langle A, X \rangle = b;\ X \succcurlyeq 0, \qquad (D2)$$

where C is in the space of symmetric matrices ($C \in \mathcal{S}_n$), X is in the space of positive semidefinite matrices ($X \in \mathcal{S}_n^+$), $\langle \cdot, \cdot \rangle$ is defined as the trace inner product $\text{Tr}[C \cdot X]$, b is a vector in $\mathcal{R}^m$, and A is the matrix of constraints. The conjugate dual of the primal

$$\max b^T y \qquad (D3)$$

$$\text{s.t}\ S = C - y^T A;\ S \succcurlyeq 0 \qquad (D4)$$

the matrix X is the primal variable and the pair (y, S) are the dual variables. In this work we use the boundary point method to solve the augmented Lagrangian dual to the SDP [117]. The total boundary point algorithm is as follows

Repeat until $\delta_{\text{outer}} < \epsilon_{\text{outer}}$

Repeat until $\delta_{\text{inner}} < \epsilon_{\text{inner}}$

solve for y^k: $A(A^T y) = A(Z^k + C + \frac{1}{\sigma}X^k) - \frac{1}{\sigma}b$

Positive projection step: $W = A^T y^k - C - \frac{1}{\sigma}X^k$; $Z^k = W_+$; $V^k = W_-$

$\delta_{\text{inner}} = ||\langle A, V^k \rangle - b||$

$X^{k+1} = V^k$

$k = k + 1$; $\delta_{\text{outer}} = ||S^k - A^T y^k + C||$

update σ.

Here the positive and negative projections W_+ and W_- are determined by the minimization

$$W_+ = \text{argmin}_{U \succcurlyeq 0}||W - U|| \qquad (D5)$$
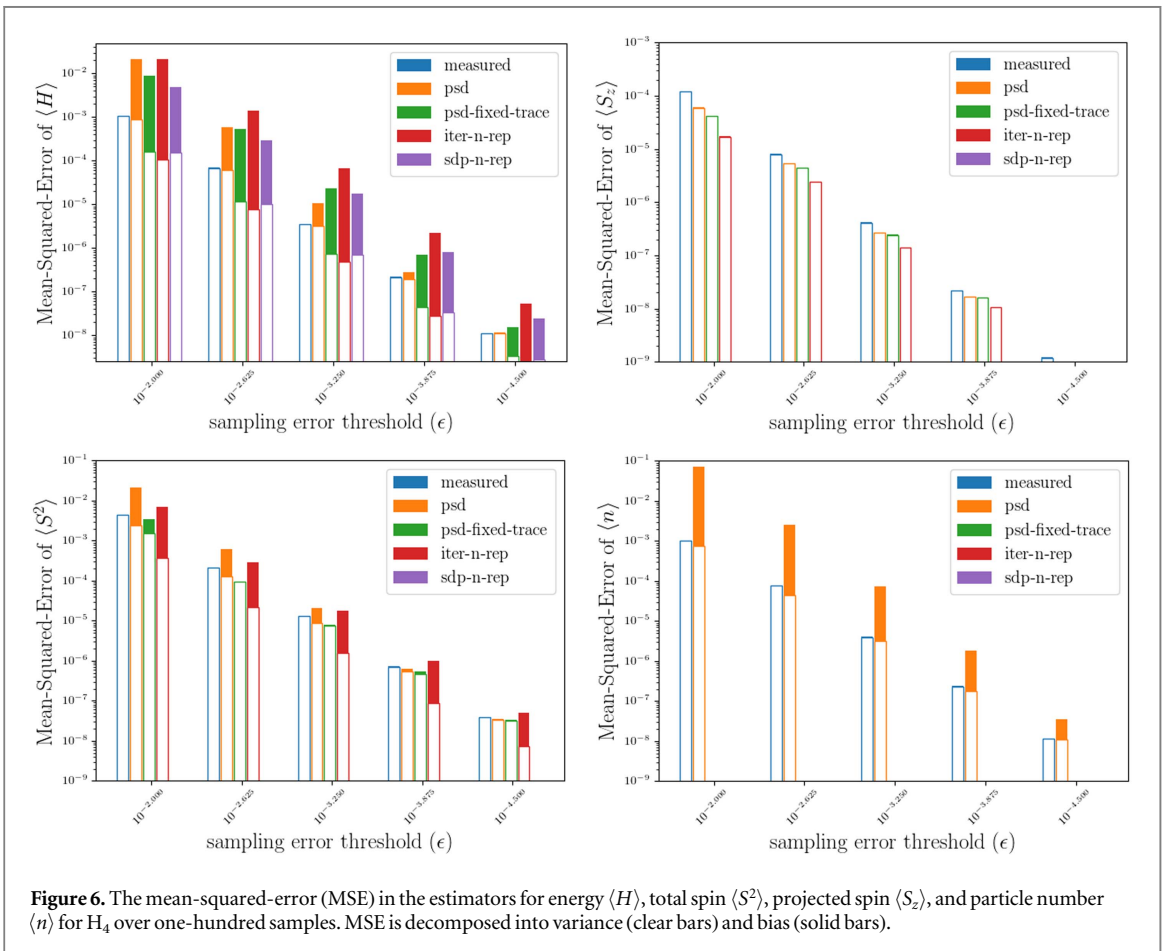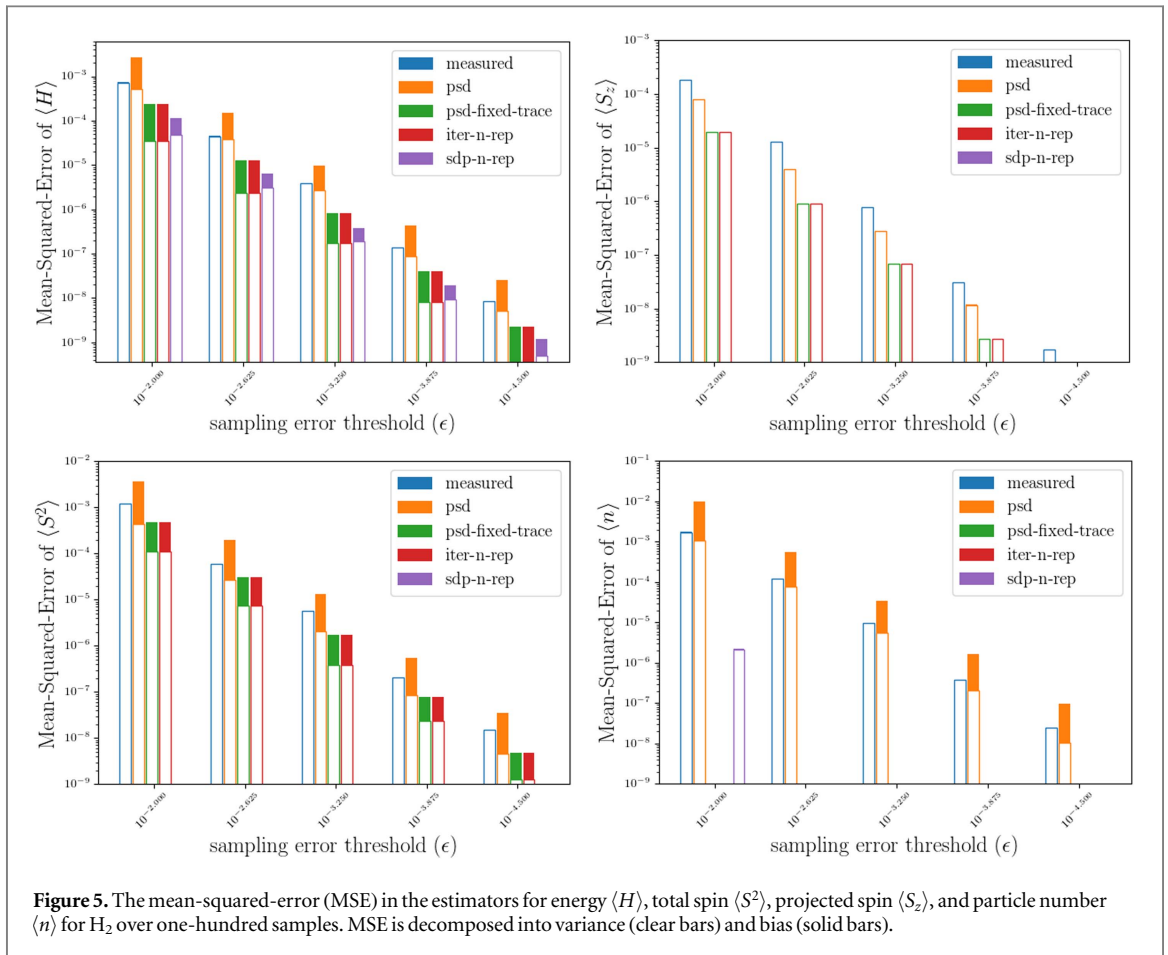
which corresponds to generating the W_+ by an eigenvalue decomposition and selecting positive eigenvalues along with their associated eigenvectors to generate the positive projection

$$W_+ = \sum_i \lambda_i^+ |\phi_i\rangle \langle\phi_i| \quad W_- = \sum_j \lambda_j^- |\phi_j\rangle \langle\phi_j|. \qquad (D6)$$

The computationally expensive task is the determination of y in the inner minimization problem. As AA^T does not change its Cholesky decomposition, it can be formed prior to the calculation and then used to backsolve for y^k. Using the backsolve method for the inner loop requires only one step. For larger problems, we can solve the inner loop with the conjugate gradient method and thus must set the inner stopping ϵ_{inner} condition significantly below the outer stopping condition ϵ_{outer}. For all SDPs, we use the *L2*-norm of the primal error $||\langle A|X\rangle - b||_2$ as δ_{outer}. All SDPs were stopped when δ_{outer} fell below 1.0×10^{-8} or the number of outer iterations reached five-thousand.

Appendix E. Constrained observables

To further examine the effects of the four projection techniques proposed we examined the MSE as a function of noise in the Gaussian error model and type of projection procedure used. The Gaussian error model does not preclude a positive semidefinite 2-RDM and thus restoration of physical symmetries such as positive-

Figure 5. The mean-squared-error (MSE) in the estimators for energy $\langle H \rangle$, total spin $\langle S^2 \rangle$, projected spin $\langle S_z \rangle$, and particle number $\langle n \rangle$ for H_2 over one-hundred samples. MSE is decomposed into variance (clear bars) and bias (solid bars).

Figure 6. The mean-squared-error (MSE) in the estimators for energy $\langle H \rangle$, total spin $\langle S^2 \rangle$, projected spin $\langle S_z \rangle$, and particle number $\langle n \rangle$ for H_4 over one-hundred samples. MSE is decomposed into variance (clear bars) and bias (solid bars).

semidefinite-ness, constrained spin-, and particle-numbers are expected to increase the observed energy of the 2-RDM with respect to the uncorrected noisy 2-RDM.

References

[1] Lloyd S 1996 *Science* **273** 1073
[2] Abrams D S and Lloyd S 1997 *Phys. Rev. Lett.* **79** 4
[3] Abrams D S and Lloyd S 1999 *Phys. Rev. Lett.* **83** 5162
[4] Ortiz G, Gubernatis J, Knill E and Laflamme R 2001 *Phys. Rev. A* **64** 22319
[5] Somma R D, Ortiz G, Gubernatis J E, Knill E and Laflamme R 2002 *Phys. Rev. A* **65** 17
[6] Aspuru-Guzik A, Dutoi A D, Love P J and Head-Gordon M 2005 *Science* **309** 1704
[7] Kassal I, Jordan S P, Love P J, Mohseni M and Aspuru-Guzik A 2008 *Proc. Natl Acad. Sci.* **105** 18681
[8] Ward N J, Kassal I and Aspuru-Guzik A 2008 *J. Chem. Phys.* **130** 194105
[9] Whitfield J D, Biamonte J and Aspuru-Guzik A 2011 *Mol. Phys.* **109** 735
[10] Cody Jones N, Whitfield J D, McMahon P L, Yung M-H, Meter R V, Aspuru-Guzik A and Yamamoto Y 2012 *New J. Phys.* **14** 115023
[11] Toloui B and Love P J 2013 arXiv:1312.2579
[12] Hastings M B, Wecker D, Bauer B and Troyer M 2015 *Quantum Inf. Comput.* **15** 1
[13] Babbush R, Berry D W, Kivlichan I D, Wei A Y, Love P J and Aspuru-Guzik A 2018 *Quantum Sci. Technol.* **3** 15006
[14] Babbush R, Berry D W, Kivlichan I D, Wei A Y, Love P J and Aspuru-Guzik A 2016 *New J. Phys.* **18** 33032
[15] Babbush R, Love P J and Aspuru-Guzik A 2014 *Sci. Rep.* **4** 1
[16] Kivlichan I D, Wiebe N, Babbush R and Aspuru-Guzik A 2017 *J. Phys. A: Math. Theor.* **50** 305301
[17] Sugisaki K, Yamamoto S, Nakazawa S, Toyota K, Sato K, Shiomi D and Takui T 2016 *J. Phys. Chem. A* **120** 6459
[18] Babbush R, Wiebe N, McClean J, McClain J, Neven H and Chan G K-L 2018 *Phys. Rev. X* **8** 011044
[19] Motzoi F, Kaicher M P and Wilhelm F K 2017 *Phys. Rev. Lett.* **119** 160503
[20] Kivlichan I D, McClean J, Wiebe N, Gidney C, Aspuru-Guzik A, Chan G K-L and Babbush R 2018 *Phys. Rev. Lett.* **120** 110501
[21] Veis L and Pittner J 2010 *J. Chem. Phys.* **133** 194106
[22] Veis L and Pittner J 2014 *J. Chem. Phys.* **140** 1
[23] McClean J R, Babbush R, Love P J and Aspuru-Guzik A 2014 *J. Phys. Chem. Lett.* **5** 4368
[24] Wecker D, Bauer B, Clark B K, Hastings M B and Troyer M 2014 *Phys. Rev. A* **90** 1
[25] Poulin D, Hastings M B, Wecker D, Wiebe N, Doherty A C and Troyer M 2015 *Quantum Inf. Comput.* **15** 361
[26] Babbush R, McClean J, Wecker D, Aspuru-Guzik A and Wiebe N 2015 *Phys. Rev. A* **91** 22311
[27] Reiher M, Wiebe N, Svore K M, Wecker D and Troyer M 2017 *Proc. Natl Acad. Sci.* **114** 7555
[28] Berry D W, Kieferová M, Scherer A, Sanders Y R, Low G H, Wiebe N, Gidney C and Babbush R 2018 *npj Quantum Inf.* **4** 22
[29] Seeley J T, Richard M J and Love P J 2012 *J. Chem. Phys.* **137** 224109
[30] Whitfield J D 2013 *J. Chem. Phys.* **139** 21105
[31] Tranter A, Sofia S, Seeley J, Kaicher M, McClean J, Babbush R, Coveney P V, Mintert F, Wilhelm F and Love P J 2015 *Int. J. Quantum Chem.* **115** 1431
[32] Moll N, Fuhrer A, Staar P and Tavernelli I 2016 *J. Phys. A: Math. Theor.* **49** 295301
[33] Whitfield J D, Havlíček V and Troyer M 2016 *Phys. Rev. A* **94** 030301
[34] Barkoutsos P K, Moll N, Staar P W J, Mueller P, Fuhrer A, Filipp S, Troyer M and Tavernelli I 2017 arXiv:1706.03637
[35] Havlíček V, Troyer M and Whitfield J D 2017 *Phys. Rev. A* **95** 32332
[36] Bravyi S, Gambetta J M, Mezzacapo A and Temme K 2017 arXiv:1701.08213
[37] Zhu G, Subasi Y, Whitfield J D and Hafezi M 2018 *npj Quantum Inf.* **4** 16
[38] Setia K and Whitfield J D 2017 arXiv:1712.00446
[39] Steudtner M and Wehner S 2017 arXiv:1712.07067
[40] Lanyon B P *et al* 2010 *Nat. Chem.* **2** 106
[41] Li Z, Yung M-H, Chen H, Lu D, Whitfield J D, Peng X, Aspuru-Guzik A and Du J 2011 *Sci. Rep.* **1** 1
[42] Wang Y *et al* 2015 *ACS Nano* **9** 7769
[43] Santagati R *et al* 2018 *Sci. Adv.* **4** (http://advances.sciencemag.org/content/4/1/eaap9646.full.pdf)
[44] Peruzzo A, McClean J, Shadbolt P, Yung M-H, Zhou X-Q, Love P J, Aspuru-Guzik A and O'Brien J L 2014 *Nat. Commun.* **5** 1
[45] Yung M-H, Casanova J, Mezzacapo A, McClean J, Lamata L, Aspuru-Guzik A and Solano E 2014 *Sci. Rep.* **4** 9
[46] McClean J R, Romero J, Babbush R and Aspuru-Guzik A 2016 *New J. Phys.* **18** 23023
[47] Shen Y, Zhang X, Zhang S, Zhang J-N, Yung M-H and Kim K 2017 *Phys. Rev. A* **95** 020501
[48] Wecker D, Hastings M B and Troyer M 2015 *Phys. Rev. A* **92** 42303
[49] Sawaya N P D, Smelyanskiy M, McClean J R and Aspuru-Guzik A 2016 *J. Chem. Theory Comput.* **12** 3097
[50] O'Malley P J J *et al* 2016 *Phys. Rev. X* **6** 31007
[51] Kandala A, Mezzacapo A, Temme K, Takita M, Chow J M and Gambetta J M 2017 *Nature* **549** 242
[52] Romero J, Babbush R, McClean J, Hempel C, Love P and Aspuru-Guzik A 2017 arXiv:1701.02691
[53] Farhi E, Goldstone J and Gutmann S 2014 arXiv:1411.4028
[54] McClean J R, Schwartz M E, Carter J and de Jong W A 2017 *Phys. Rev. A* **95** 42308
[55] Colless J I, Ramasesh V V, Dahlen D, Blok M S, Kimchi-Schwartz M E, McClean J R, Carter J, de Jong W A and Siddiqi I 2018 *Phys. Rev. X* **8** 011021
[56] Pople J A 1999 *Rev. Mod. Phys.* **71** 1267
[57] Raeisi S, Wiebe N and Sanders B C 2012 *New J. Phys.* **14** 103017
[58] Coulson C A 1960 *Rev. Mod. Phys.* **32** 170
[59] Coleman A J and Yukalov V I 2000 *Reduced Density Matrices: Coulsons Challenge* vol 72 (Berlin: Springer)
[60] Coleman A J 1978 *Int. J. Quantum Chem.* **13** 67
[61] Cioslowski J 2000 *Many-Electron Densities and Reduced Density Matrices* (Berlin: Springer)
[62] Foley J J and Mazziotti D A 2012 *Phys. Rev. A* **86** 012512
[63] Cramer M, Plenio M B, Flammia S T, Somma R, Gross D, Bartlett S D, Landon-Cardinal O, Poulin D and Liu Y-K 2010 *Nat. Commun.* **1** 149

[64] Liu Y-K 2006 *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques* (Berlin: Springer) pp 438–49
[65] Bravyi S 2004 *Quantum Inf. Comput.* **4** 12
[66] Szabo A and Ostlund N S 1996 *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Mineola, NY: Dover)
[67] Mazziotti D A 2002 *Phys. Rev. A* **65** 062511
[68] Coleman A J and Absar I 1980 *Int. J. Quantum Chem.* **18** 1279
[69] Verstraete F and Cirac J I 2005 *J. Stat. Mech.* P09012
[70] Nakata M, Nakatsuji H, Ehara M, Fukuda M, Nakata K and Fujisawa K 2001 *J. Chem. Phys.* **114** 8282
[71] Alcoba D R and Valdemoro C 2005 *Int. J. Quantum Chem.* **102** 629
[72] Löwdin P O 1987 Some aspects on the development of the theory of reduced density matrices and the representability problem (*Density Matrices and Density Functionals: Proceedings of the A. John Coleman Symposium*) ed R Erdahl and V H Smith (Dordrecht: Springer) pp 21–49
[73] Coleman A J 1963 *Rev. Mod. Phys.* **35** 668
[74] Erdahl R M 1978 *Int. J. Quantum Chem.* **13** 697
[75] Percus J K 1978 *Int. J. Quantum Chem.* **13** 89
[76] Garrod C and Percus J K 1964 *J. Math. Phys.* **5** 1756
[77] Mazziotti D A and Erdahl R M 2001 *Phys. Rev. A* **63** 042113
[78] Mazziotti D A 2016 *Phys. Rev. A* **94** 032516
[79] Mazziotti D A 2012 *Phys. Rev. Lett.* **108** 263002
[80] Klyachko A A 2006 *J. Phys.: Conf. Ser.* **36** 72
[81] Coleman A J 1977 *Int. J. Quantum Chem.* **11** 907
[82] Baumgratz T and Plenio M B 2012 *New J. Phys.* **14** 023027
[83] Fosso-Tande J, Nascimento D R and DePrince E III 2016 *Mol. Phys.* **114** 423
[84] Zhao Z, Braams B J, Fukuda M, Overton M L and Percus J K 2004 *J. Chem. Phys.* **120** 2095
[85] Rubin N C and Mazziotti D A 2015 *J. Phys. Chem. C* **119** 14706
[86] Hammond J R and Mazziotti D A 2005 *Phys. Rev. A* **71** 062503
[87] Mazziotti D A 2011 *Phys. Rev. Lett.* **106** 083001
[88] Boixo S, Isakov S V, Smelyanskiy V N, Babbush R, Ding N, Jiang Z, Bremner M J, Martinis J M and Neven H 2016 arXiv:1608.00263
[89] Milman V D and Schechtman G 2009 *Asymptotic Theory of Finite Dimensional Normed Spaces: Isoperimetric Inequalities in Riemannian Manifolds* vol 1200 (Berlin: Springer)
[90] McClean J R *et al* 2017 arXiv:1710.07629
[91] Higham N J 1988 *Linear Algebr. Appl.* **103** 103
[92] Smolin J A, Gambetta J M and Smith G 2012 *Phys. Rev. Lett.* **108** 070502
[93] Lanssens C, Ayers P W, Neck D V, Baerdemacker S D, Gunst K and Bultinck P 2018 *J. Chem. Phys.* **148** 084104
[94] Higham N J 2002 *IMA J. Numer. Anal.* **22** 329
[95] Parrish R M *et al* 2017 *J. Chem. Theory Comput.* **13** 3185
[96] Gidofalvi G and Mazziotti D A 2005 *Phys. Rev. A* **72** 052505
[97] Andersson K, Malmqvist P A, Roos B O, Sadlej A J and Wolinski K 1990 *J. Phys. Chem.* **94** 5483
[98] Andersson K, Malmqvist P-Å and Roos B O 1992 *J. Chem. Phys.* **96** 1218
[99] Celani P and Werner H-J 2000 *J. Chem. Phys.* **112** 5546
[100] Hirao K 1992 *Chem. Phys. Lett.* **190** 374
[101] Angeli C, Cimiraglia R, Evangelisti S, Leininger T and Malrieu J-P 2001 *J. Chem. Phys.* **114** 10252
[102] Angeli C, Cimiraglia R and Malrieu J-P 2002 *J. Chem. Phys.* **117** 9138
[103] Angeli C, Pastore M and Cimiraglia R 2006 *Theor. Chem. Acc.* **117** 743
[104] Yanai T and Chan G K-L 2006 *J. Chem. Phys.* **124** 194106
[105] Yanai T and Chan G K-L 2007 *J. Chem. Phys.* **127** 104107
[106] Neuscamman E, Yanai T and Chan G K-L 2010 *J. Chem. Phys.* **132** 024106
[107] Torheyden M and Valeev E F 2009 *J. Chem. Phys.* **131** 171103
[108] Kong L and Valeev E F 2010 *J. Chem. Phys.* **133** 174126
[109] Werner H-J and Reinsch E-A 1982 *J. Chem. Phys.* **76** 3144
[110] Werner H-J and Knowles P J 1988 *J. Chem. Phys.* **89** 5803
[111] Zgid D, Ghosh D, Neuscamman E and Chan G K-L 2009 *J. Chem. Phys.* **130** 194107
[112] Mazziotti D A 2003 *Phys. Rev. A* **68** 052501
[113] Greenman L and Mazziotti D A 2008 *J. Chem. Phys.* **128** 114109
[114] Rubin N C and Mazziotti D A 2014 *Theor. Chem. Acc.* **133** 1492
[115] Burer S and Monteiro R D 2003 *Math. Program.* **95** 329
[116] Burer S and Monteiro R D 2005 *Math. Program.* **103** 427
[117] Povh J, Rendl F and Wiegele A 2006 *Computing* **78** 277