





**Coarse-grained intermolecular interactions on quantum processors**Lewis W. Anderson <sup>1,\*</sup>, Martin Kiffner <sup>2,1</sup>, Panagiotis Kl. Barkoutsos,<sup>3</sup> Ivano Tavernelli,<sup>3</sup>  
Jason Crain <sup>4,5</sup> and Dieter Jaksch <sup>6,1,2</sup><sup>1</sup>*Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom*<sup>2</sup>*Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543*<sup>3</sup>*IBM Quantum, IBM Research Zurich, Säumerstrasse 4, 8803 Rüschlikon, Switzerland*<sup>4</sup>*IBM Research Europe, Hartree Centre STFC Laboratory, Sci-Tech Daresbury, Warrington WA4 4AD, United Kingdom*<sup>5</sup>*Department of Biochemistry, University of Oxford, Oxford OX1 3QU, United Kingdom*<sup>6</sup>*Institut für Laserphysik, Universität Hamburg, 22761 Hamburg, Germany*

(Received 5 October 2021; accepted 10 May 2022; published 6 June 2022)

Variational quantum algorithms (VQAs) are increasingly being applied in simulations of strongly bound (covalently bonded) systems using full molecular orbital basis representations. The application of quantum computers to the weakly bound intermolecular and noncovalently bonded regime, however, has remained largely unexplored. In this work, we develop a coarse-grained representation of the electronic response that is ideally suited for determining the ground state of weakly interacting molecules using a VQA. We require qubit numbers that grow linearly with the number of molecules and derive scaling behavior for the number of circuits and measurements required, which compare favorably to traditional variational quantum eigensolver methods. We demonstrate our method on IBM superconducting quantum processors and show its capability to resolve the dispersion energy as a function of separation for a pair of nonpolar molecules—thereby establishing a means by which quantum computers can model Van der Waals interactions directly from zero-point quantum fluctuations. Within this coarse-grained approximation, we conclude that current-generation quantum hardware is capable of probing energies in this weakly bound but nevertheless chemically ubiquitous and biologically important regime. Finally, we perform experiments on simulated and real quantum computers for systems of three, four, and five oscillators as well as oscillators with anharmonic onsite binding potentials; the consequences of the latter are unexamined in large systems using classical computational methods but can be incorporated here with low computational overhead.

DOI: [10.1103/PhysRevA.105.062409](https://doi.org/10.1103/PhysRevA.105.062409)**I. INTRODUCTION**

Computer simulation of matter at the atomic and molecular scale has advanced to the point where it is now playing an increasingly significant role in the design and discovery of new functional materials. However, these simulations all face the common challenge of how to strike inevitable compromises between the physical realism of an underlying model and the practicality of its implementation across the deep hierarchy of intermolecular forces responsible for cohesion in condensed phases.

This hierarchy spans a continuum of energy scales ranging from greater than 1 eV for *intramolecular* bonds, extending down to  $10^{-1}$ – $10^{-2}$  eV for the intermolecular regime [1]. The weakest of these are the dispersion interactions arising from instantaneous quantum fluctuations of the electron distribution. These give rise to correlated induced multipole moments with the leading-order contribution coming from dipolar coupling, attenuated by the familiar  $1/R^6$  asymptotic decay with distance. Dispersion interactions therefore form part of the nonlocal correlation energy of the system and are not pair-

wise additive. Although generally weaker than electrostatic interactions or hydrogen bonds, their long-ranged character and ubiquity means that they exert influence to some extent in all noncovalent associations.

There is now clear evidence that dispersion effects extend beyond cohesive energy and have impact on structural and mechanical properties [2,3] as well as barrier heights and phase transitions [4]. They can therefore play governing roles in a wide range of condensed matter phenomena including liquid state physics, molecular crystal stability, low-dimensional and layered system stacking, hybrid organic-inorganic interface properties, and biophysical interactions such as drug-ligand binding. Moreover, as these interactions scale with system size, their importance is expected to grow for larger molecules and supramolecular assemblies [5].

Principally, there are two distinct strategies employed to capture this molecular interaction hierarchy in computer models: First, there are empirical potential or force-field methods. Here the system energy is typically decomposed into bonded (stretches, bends, and torsions) and nonbonded (electrostatic and dispersive) terms. These are parameterized to fit certain experimental reference quantities. The motion of the system can then be evolved by classical mechanics (molecular dynamics), and trajectories averaged to obtain static properties.

\*lewis.anderson@physics.ox.ac.uk

These methods have the advantage of extreme computational efficiency, meaning that they can be applied to relatively large systems. However, they generally lack explicit electronic responses and, consequently, transferability to thermodynamic or environmental conditions outside those of the parametrization range is limited.

Second, there are so-called first principles approaches in which the Schrödinger equation is solved to various levels of approximation. Here the prevailing methods for molecular dynamics applications are based on Kohn-Sham density functional theory. In these models, there are no empirical force laws: Instead, forces are computed “on-the-fly” directly from the electronic structure according to the Hellmann-Feynman [6,7] theorem with trajectories evolved on the Born-Oppenheimer surface [8]. These techniques are computationally resource-intensive and tend to be limited to relatively small systems. Also, common density functional approximations do not properly capture nonlocal correlations involved in dispersion forces, and there are active efforts to devise dispersion-corrected density functionals [9–12]. High-level quantum chemistry methods (such as configuration interaction (CI) [13–15] or canonical coupled cluster (CC) [16,17] with perturbative—single, double, triple—excitations) can account for dispersion interactions explicitly within various levels of approximation albeit with even more severe limitations imposed by system size scaling [18–20].

Quantum computing offers promising routes for simulations of correlated electronic systems at the first-principles level. Materials modeling at atomic and molecular scales is therefore one of its most anticipated application areas. One such route is to use variational quantum algorithms (VQAs) [21,22] for solving the Schrödinger equation for chemical systems [23]. VQAs have the benefits of requiring relatively short coherence times and modest numbers of qubits, which is achieved by using a shallow parameterized quantum circuit embedded within a classical optimization routine to find the ground state of a given problem Hamiltonian [24]. Since the first proposal to use variational methods—namely, the variational quantum eigensolver (VQE) for chemical systems [24], VQAs have seen application to a significant number of areas including combinatorial optimization [25,26], strongly correlated materials [27–29], and nonlinear partial differential equations [30]. Current capabilities to model molecular properties from first principles in large systems using VQAs are, however, limited by the prohibitive computational cost of the high-level, wavefunction methods typically deployed in quantum chemistry. For example, solution of the Schrödinger equation at the full configuration interaction (so-called “full CI”) level for the ground-state wave function for molecular systems exhibits factorial scaling [31].

Chemical applications of VQAs have consequently been confined to the strongly bound intramolecular regime of the interaction hierarchy and to small systems with examples now demonstrated for dissociation curves of small molecules [32,33], ground-state energies for intramolecular bonds [34], molecular excitation energies [31], covalent binding energies of hydrogen chains, isomerization barriers [35], molecular vibrations [36,37], and two-site DMFT calculations [38,39]. All except the last two examples consider

Hamiltonians and states within full electronic basis sets such as Slater-type or Gaussian-type orbitals [40]. For the sake of comparison with our method, we will refer to these as orbital-based VQE methods.

In contrast, the intermolecular (noncovalent) interaction regime has been largely unexplored using quantum algorithms and is a gap which will need to be filled in order to address realistic material science problems using near-term, noisy, quantum computers. Due to the size of the molecular complexes of interest and the necessary inclusion of core electrons for dispersive forces, the required number of electrons and orbitals implies a large number of basis functions and qubits [41,42]. This makes the direct scale up of current quantum algorithms for electronic structure calculations very impractical, and therefore strategies are needed to reduce the number of electrons and orbitals treated explicitly at the highest levels of accuracy.

To address this difficulty, we develop a coarse-grained model of electronic interactions that is ideally suited for VQA and can easily be extended to represent a wide range of possible interactions. Our approach is inspired by and extends a maximally coarse-grained model used in classical computations [43]. In this model the responsive (polarizable) part of the molecular charge distribution is modeled as a quantum harmonic oscillator embedded in the molecular frame and its properties can be tuned to reproduce reference values of polarisability or dispersion coefficients for real molecular species. Zero-point multipolar fluctuations are present by construction and so dispersion arises naturally. While the model parameters are defined empirically to fit molecular properties there are no interaction potentials or force laws between atoms or molecules defined *a priori*. These are instead determined from the coarse-grained electronic structure as in *ab initio* approaches.

This model blends therefore features from both empirical potential and first principles methodologies and has already been solved, with model limitations, using a path integral approach on classical processors. In particular, it has been successfully applied to noble gas liquids and solids (pure dispersion) and to water (where polarization and dispersion are both present) across its phase diagram where it has correctly predicted vaporization enthalpy, temperature of maximum density, temperature dependences of the dielectric constant and surface tension [43–45] and revealed molecular signatures of Widom line crossing in the supercritical fluid [46]. These classical methods have, however, been developed with the specific aim of deriving force laws and are restricted to harmonic confining potentials where certain order interactions are missing and others are related by simple algebraic relationships that hold only approximately for real systems [47].

By using quantum processors to model a coarse-grained representation, our method gives a more general framework than classical approaches. This allows for the extension to anharmonic oscillators, which would capture these missing interactions as well as allowing for the introduction of more complex physics through nonlinear interaction terms. Figure 1 shows a schematic representation of the algorithm and correlated charge fluctuations captured by the model.

Here we summarize our main findings: First, the number of qubits required to simulate the model on a quantum

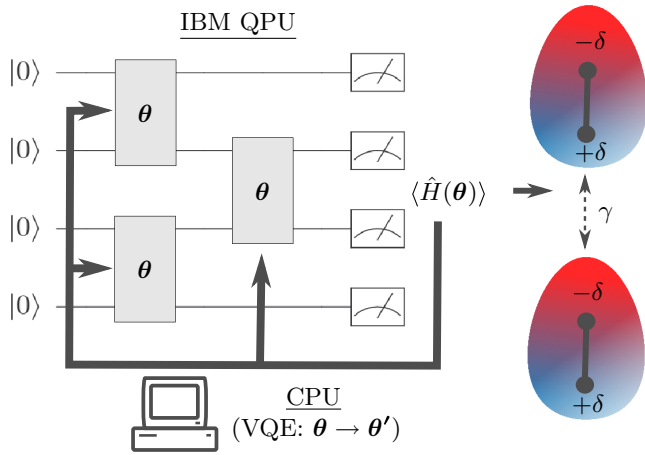


FIG. 1. Schematic of the VQE optimization procedure and output states showing instantaneous asymmetry of the molecular charge distribution. These correlated charge density fluctuations lead to dispersion interaction which, to the lowest order, is the familiar  $1/R^6$  London attraction.

computer grows linearly with the number of molecules. Second, the number of unique circuits required to measure our Hamiltonian with all-to-all dipolar interactions scales at most linearly with the number of oscillators. This compares favorably to the  $O(N^3)$  unique circuits that must be measured for a full orbital-based VQE Hamiltonian with  $N$  orbitals [48]. For our algorithm, relaxation of the harmonic approximation incurs negligible experimental overhead. Classical methods, such as path integral and Monte Carlo techniques, typically rely on the efficient sampling [49] of two-point correlators for Gaussian states generated by harmonic Hamiltonians. However, the non-Gaussian ground states of anharmonic Hamiltonians are much more costly to sample. Thus, anharmonic potentials have not been studied much using classical computational approaches.

We also present proof-of-principle experiments using IBM superconducting quantum processors demonstrating the solubility of our model on current generation hardware. We focus on two linear nonpolar molecules where the electronic responses are described by embedded quantum oscillators. Here the interactions arise from London dispersion forces for which the leading contribution comes from dipolar quantum fluctuations. Looking ahead to the main experimental results of this paper, Fig. 2 shows the result of our VQA; combined with a simple method to extract the binding energy; for a pair of these interacting oscillators with parameters chosen to represent an iodine ( $I_2$ ) molecular dimer. This demonstration successfully recreates the leading order London dispersion interaction, thereby giving the first demonstration of the use of quantum processors to model dispersion forces directly arising from quantum fluctuations.

Finally, we extend our proof-of-principle experiments to systems of three, four, and five oscillators using simulated quantum computers. We show how the algorithm performs as a function of ansatz size and demonstrate that it is able to find ground states of systems containing nontrivial many-body interactions. We also perform experiments using an IBM superconducting quantum processor in order for a system

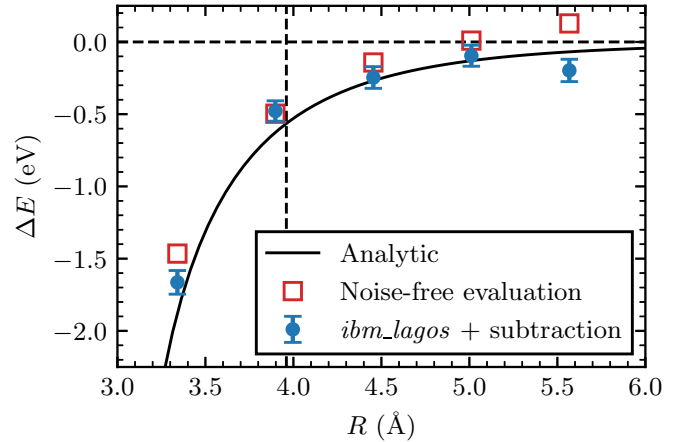


FIG. 2. London dispersion energy ( $\Delta E$ ) calculated using variational quantum algorithm for one-dimensional quantum Drude oscillators computing interacting  $I_2$  dimers. The black line gives the dispersion energy obtained analytically. Red squares correspond to noise-free evaluation (via state vector simulations) of the experimentally optimized coupled ground state. Blue circles correspond to energy directly obtained from experimentally optimized ground states (error bars represent one-standard deviation on the mean). The horizontal line marks zero dispersion interaction corresponding to uncoupled oscillators at  $R = \infty$  and vertical dashed line corresponds to the  $I_2$  van der Waals diameter.

including anharmonic onsite binding potentials. We show that the algorithm is again able to capture systems showing strong anharmonic interactions however device noise levels must be further improved in order to accurately measure the anharmonic contribution to the energy of the system, even when using state-of-the-art error mitigation techniques.

Our coarse-grained model allows for the use of quantum computers to efficiently model (with favorable scaling) the intermolecular interactions between large molecules, which are important in characterizing a wide range of chemical and biological systems. This fills an essential gap in the chemical interaction hierarchy and allows us to extend current coarse-grained classical methods to capture a richer set of physical behaviors with the necessary accuracy. Given the reduced resource requirements (i.e., number of qubits, gate operations, and circuit measurements) when compared to full orbital-based electronic structure calculations, the methodology presented here may be more readily scalable to demonstrating an application of quantum advantage.

The layout of this article is as follows. In Sec. II we describe the coarse-grained model and how it can be solved using a variational quantum algorithm. This includes Sec. II A describing the coarse-grained model of interacting oscillators; Sec. II B giving a description of how this system is represented on a quantum computer; Sec. II C, which makes use of results from graph theory and measure theory to give a scheme for grouping operators needed to measure the Hamiltonian as well as derive upper and lower bounds for the number of shots required to measure the Hamiltonian to a given accuracy; and Sec. II D, which describes the variational ansatz used in our experiments. In Sec. III we present proof-of-principle results modeling the London dispersion interaction of an  $I_2$  dimer and a technique to subtract the effect of device noise to

accurately measure the binding energy. In Sec. IV we describe how the one-dimensional model can be extended to three dimensions and to include anharmonic and nonlinear terms to model richer physics than those accessible to efficient classical methods. In Sec. V we present further experiments on real and simulated quantum processors. This includes Sec. VA for systems of anharmonic oscillators, as well as Sec. VB where we increase the size of the harmonic system to include three, four, and five oscillators. In Sec. VI we discuss our results and describe further research directions.

## II. COARSE-GRAINED MODELING ON DIGITAL QUANTUM COMPUTERS

### A. Coarse-grained model

In the coarse-grained model, collective electronic responses are described by a charged quantum Drude oscillator (QDO) tethered to an atomic nucleus or embedded in a molecular frame. For neutral species the Drude (quasi-)particle and the nucleus to which it is bound have charges  $-q$  and  $+q$ , respectively. The correlated displacements of these QDOs, which correspond to instantaneous charge density fluctuations in molecules, allows us to capture dispersive interactions between nonpolar molecules. It is an electronically coarse-grained model which nonetheless exhibits the full hierarchy of electronic responses at long range: Many-body polarization and dispersion to all orders arise naturally and, in principle, intermolecular forces can then be computed “on the fly” from the coarse-grained electron distribution of Coulomb-coupled oscillators without the need for *a priori* force laws or empirical interaction potentials. One further feature of this model, is that it obeys distinguishable particle statistics—Fermi statistics are not required. The model must therefore be augmented by a simple short-ranged repulsion term. The basic method has been demonstrated on classical processors in the case of noble gas solids and single-component liquids [50,51]. Using only the properties of the isolated molecule in the model parameters, emergent properties of these systems have been predicted across a range of their respective phase diagrams. For a recent review of these methods and results see [43].

For simplicity, for most of this work we will focus on the case of one-dimensional QDOs with dipole order coupling; however, much of the methodology developed here can be applied straightforwardly to systems of three-dimensional oscillators, and we give a brief description of this extension in Sec. IV A. Even the simple one-dimensional model recreates the London dispersion interaction proportional to  $R^{-6}$  and  $\alpha^2$ , where  $\alpha$  is the dipole polarizability and  $R$  the physical separation, up-to a constant factor for spherical nonpolar molecules and exactly for a pair of linear nonpolar molecules in the limit of infinitely anisotropic polarizability (see Appendix A for further details.)

The nondimensional Hamiltonian for  $N$  one-dimensional QDOs with quadratic coupling is

$$\hat{H} = \sum_{i=1}^N (\hat{x}_i^2 + \hat{p}_i^2) + \sum_{j>i} \gamma_{i,j} \hat{x}_i \hat{x}_j, \quad (1)$$

where  $\hat{x}_i$  and  $\hat{p}_i$  are the bosonic position and momentum operators for oscillator  $i$  and  $\gamma_{i,j} \in \mathbb{R}$  is the coupling between

oscillators  $i$  and  $j$ . Throughout this work we will implicitly work in energy units of  $\hbar\omega/2$  for simplicity of notation assuming that all oscillators share the same trapping frequency  $\omega$ . For each pair of identical one-dimensional oscillators aligned parallel to each other and along the interoscillator axis, separated by distance  $R_{i,j}$ , the coupling constant is given by  $\gamma_{i,j} = -2(q^2/\mu\omega^2)R_{i,j}^{-3} = -4\alpha R_{i,j}^{-3}$ , where the effective charge of the Drude particle  $q$ , the effective mass  $\mu$  and the oscillator frequency  $\omega$  are model parameters. The dipole polarizability is given by  $\alpha = q^2/\mu\omega^2$ . If instead the oscillators are aligned parallel to each other and perpendicular to the interoscillator axis, the coupling is given by  $\gamma_{i,j} = 2\alpha R_{i,j}^{-3}$ .

One advantage of this harmonically bound coarse-grained model is that efficient classical methods exist which can be used for benchmarking quantum implementations. However, the price paid for this harmonic approximation is that the relative strengths of interactions within the model are governed by Gaussian statistics meaning that there are simple algebraic relations relating interactions at different orders. This leads to the absence of terms, such as the second hyperpolarizability (dipole-dipole-dipole-dipole) [51], due to certain symmetries that emerge in the harmonic model.

In what follows we present a variational quantum algorithm for calculating intermolecular interactions between molecules using a one-dimensional QDO model and outline how to extend this to a full three-dimensional representation. Such a quantum algorithm is not limited to the inclusion of harmonic Hamiltonian terms and can be extended to include anharmonic trapping potentials as well as nonlinear interaction terms that may allow for a richer set of physics beyond Gaussian statistics and the study of more realistic chemical systems [51].

### B. Algorithm description

In order to find the ground state of the Hamiltonians of interest we use the common procedure for a Variational Quantum Eigensolver (VQE) algorithm [24]. The main goal of VQE is to produce quantum states  $|\psi(\theta)\rangle$  using a quantum circuit parameterized by real parameters  $\theta = \{\theta_0, \theta_1, \dots\}$  which can be varied. Throughout the algorithm, the parameters are updated using a classical feedback loop in order to minimise the expected values of the Hamiltonian on the particular quantum state. Guided from the variational principle the optimization procedure can provide a quantum state that approximates the exact ground state of the system by solving the minimization problem

$$\min_{\theta} \langle \psi(\theta) | \hat{H} | \psi(\theta) \rangle, \quad (2)$$

provided the variational ansatz is able to represent a state close to the ground state within the required precision. In the following we outline how we encode the states of harmonic QDOs into qubit states and how the Hamiltonian is measured efficiently on a quantum computer before defining our variational ansatz later.

We represent the state of the QDOs in the Fock basis which for bosonic oscillators form an infinite basis. We are limited to a finite number of qubits and so must restrict ourselves to the space of a finite subset of Fock states. This will naturally lead to solutions that approximately match the true states in full

Fock space. Since dispersive interaction energies are expected to be much smaller than the bare energy scale  $\hbar\omega$ , the ground states of realistic Hamiltonians will be dominated by the lowest energy Fock states, and thus the truncation of the Fock space will not severely impact the accuracy of solutions. We use a compact encoding [52], which was proposed in earlier work on vibrational modes on a quantum computer [36,53]. For this encoding, the first  $d$  Fock states are mapped onto  $m = \log_2 d$  qubits as

$$\mathcal{H}_{d\text{-Fock}} \rightarrow \mathcal{H}_2^{\otimes m} : |n\rangle \rightarrow |\text{bin}_m(n)\rangle, \quad (3)$$

where  $\text{bin}_m(n)$  denotes the  $m$ -length binary representation of integer  $n \in \{0, 1, \dots, d-1\}$ . Throughout this work, underlined integers within a  $|\cdot\rangle$  will denote a Fock state, and all other states will be written in the computational basis of the qubits. The basis for multiple QDOs is then given by the tensor product of the single QDO basis such that for a total of  $N$   $2^m$ -dimensional QDO we will use  $M = mN$  qubits.

Although we are restricted to the lowest  $d$  Fock states, we expect that the ground states required will be sufficiently close to the uncoupled ground state  $|0\rangle^{\otimes N}$  such that a finite number of basis states will give a good approximation to the ground-state energy, particularly in the weakly coupled

regime. Since we require logarithmically many qubits, this qubit encoding will be efficient provided the number of Fock states required for a chosen accuracy grows less than double exponentially.

To measure the cost function for a given variational state we must decompose each term in the Hamiltonian in (1), restricted to a finite Fock basis, into tensor product Pauli operators  $\hat{P} \in \{\mathbb{1}, X, Y, Z\}^{\otimes M}$  that can be measured on a quantum computer. For the noninteracting and coupling terms we show this for a single QDO and a single pair of QDOs respectively and note that the same decomposition can be used for all terms within the Hamiltonian just acting on different subsets of  $m$  and  $2m$  qubits. For the noninteracting contribution of a single QDO, there are  $m$  nonidentity  $Z$  operators that must be measured:

$$\hat{x}^2 + \hat{p}^2 = 2^m \mathbb{1}^{\otimes m} - \sum_{i=0}^{m-1} 2^i Z_i, \quad (4)$$

where  $Z_i$  is the Pauli  $Z$  operator acting on qubit  $i$ . For a pair of  $m$  qubit QDOs the coupled part of the Hamiltonian leads to  $(\frac{1}{2}d \log_2 d)^2$  Pauli terms. The form of the coupling terms for arbitrary  $d$ , as well as a proof for the number of terms, is given in Appendix C. As an example, the decompositions for  $m = 1$  and  $m = 2$  are given by

$$\hat{x}_1 \hat{x}_2 |_{m=1} = X_1 X_2, \quad (5)$$

$$\begin{aligned} \hat{x}_1 \hat{x}_2 |_{m=2} = & \frac{\sqrt{3}+2}{4} X_1 X_3 + \frac{\sqrt{3}+1}{4\sqrt{2}} X_1 X_2 X_3 + \frac{\sqrt{3}+1}{4\sqrt{2}} X_1 X_3 X_4 + \frac{1}{4} X_1 X_2 X_3 X_4 + \frac{\sqrt{3}+1}{4\sqrt{2}} Y_1 Y_2 X_3 + \frac{1}{4} Y_1 Y_2 X_3 X_4 \\ & + \frac{\sqrt{3}+1}{4\sqrt{2}} X_1 Y_3 Y_4 + \frac{1}{4} X_1 X_2 Y_3 Y_4 - \frac{\sqrt{3}-1}{4\sqrt{2}} Y_1 Y_2 X_3 Z_4 - \frac{\sqrt{3}-1}{4\sqrt{2}} X_1 Z_2 Y_3 Y_4 - \frac{1}{4} X_1 Z_2 X_3 \\ & - \frac{\sqrt{3}-1}{4\sqrt{2}} X_1 Z_2 X_3 X_4 - \frac{1}{4} X_1 X_3 Z_4 - \frac{\sqrt{3}-1}{4\sqrt{2}} X_1 X_2 X_3 Z_4 + \frac{2-\sqrt{3}}{4} X_1 Z_2 X_3 Z_4 + \frac{1}{4} Y_1 Y_2 Y_3 Y_4. \end{aligned} \quad (6)$$

The interaction Hamiltonian for  $N$  QDOs can then be constructed by combining these pairwise interactions for all  $N(N-1)/2$  pairs of QDOs with the correct coupling constant.

### C. Measurement cost

When combining the single-oscillator noninteracting and two-oscillator coupling terms together to measure the total Hamiltonian, if all  $\gamma_{i,j}$  are nonzero there will be a total of  $(\log_2 d)N + \frac{1}{8}d^2(\log_2 d)^2 N(N-1)$  nonidentity Pauli operators. Naively  $\hat{H}$  can be measured by measuring each of these operators using its own circuit, however the number of measurements required can be reduced significantly by measuring mutually commuting operators concurrently. Many heuristic methods exist for partitioning the terms into commuting subsets, see, for example, Refs. [48,58–63]. From the structure of our Hamiltonian we can derive an exact method for grouping terms into subsets that qubit-wise commute and therefore obtain an upper bound on the number of unique measurements that need to be performed. We find an overall upper bound of  $(N-1)(\frac{d}{2} \log_2 d)^2 + 1$  or  $N(\frac{d}{2} \log_2 d)^2 + 1$  circuits, for even

and odd  $N$  respectively, needed to measure  $\hat{H}$ . In general, for a Hamiltonian with all-to-all  $k_{\text{th}}$  order coupling, there is an upper bound of  $O((d \log_2 d)^k N^{k-1})$  scaling with the number of QDOs. Using heuristic sorting methods the scaling with respect to  $d$  can be improved beyond that given by these upper bounds, however an exact scaling is not known. For a description of the grouping method, proof of these scalings as well as examples of the result of using heuristic grouping algorithms, see Appendix C.

Since we expect that, for physically relevant systems, the ground states will be dominated by low-lying Fock states, we do not expect the need to increase  $d$  significantly to model nontrivial systems. Therefore, we are primarily concerned with the scaling with respect to the number of QDOs  $N$ . Appendix B shows that this is the case for the pair of interacting QDOs and that  $d = 4$  is a good approximation to the true ground state until very large and unphysical couplings.

Furthermore, using this known grouping, we can estimate the expected number of shots needed to measure  $\langle \hat{H} \rangle$  with absolute error  $\epsilon$ . Without knowing the ground state of an arbitrary Hamiltonian, we calculate the expected number of shots with respect to two different distributions of states in Hilbert

TABLE I. Number of unique circuits required using our exact grouping as well as expected number of shots required to achieve absolute measurement error  $\epsilon$  for a system of  $N$   $d$ -level one-dimensional QDOs. Here  $\gamma$  is a typical scale for the interoscillator couplings. The expectations for number of shots are calculated with respect to distribution over the uniform spherical measure of states (Uniform spherical) as well as the pure state of the uncoupled QDOs (Pure uncoupled). For ungrouped results all Pauli operators in the Hamiltonian are measured separately, for grouped results Pauli operators are measured in groups according to the method described in Appendix C Theorem 4.

Measure over states	Ungrouped		Grouped	
	Circuits	Shots	Circuits	Shots
Uniform spherical	$O(N^2(\log_2 d)^2 d^2)^a$	$O\left(\frac{\gamma^2 N^4 d^8}{\epsilon^2}\right)^b$	$O(Nd^2)^c$	$O\left(\frac{\gamma^2 N^3 (\log_2 d)^4 d^6}{\epsilon^2}\right)^d$
Pure uncoupled		$O\left(\frac{\gamma^2 N^4 d^4}{\epsilon^2}\right)^e$		$O\left(\frac{\gamma^2 N^3 d^4}{\epsilon^2}\right)^f$

For full expressions and proofs see Appendix D; <sup>a</sup>Theorem 1; <sup>b</sup>Theorem 7; <sup>c</sup>Theorem 4; <sup>d</sup>Theorem 6; <sup>e</sup>Theorem 9; <sup>f</sup>Theorem 8.

space. The first distribution is the uniform spherical measure on the full Hilbert space which assumes no knowledge of the distribution of possible states [[64], Chapter 7]. The second is the pure state of the uncoupled ground state  $|\psi_0\rangle = |0\rangle^{\otimes mN}$  which we expect to be close to the ground state of the system with small nonzero coupling. The scaling of the expected number of shots required for these two distributions is given in Table I for measuring all Pauli operators individually or by grouping them as described. We see that in both distributions, grouping the measurements improves the scaling from  $O(N^4)$  to  $O(N^3)$  with respect to  $N$  and significantly reduces the polynomial scaling with respect to  $d$  in the case of the uniform spherical measure. In all cases, the exact values as well as proofs for the number of shots required can be found in Appendix D (the number of shots is an upper bound in the case of the uniform spherical measure and tight for the pure uncoupled state).

To demonstrate the number of shots required for modeling real molecules of interest, Fig. 3 highlights the number of shots required to achieve a measurement error on  $\langle \hat{H} \rangle$  of  $\epsilon = 10\% \times \Delta E(R_{vdW})$ . Here we consider  $\Delta E(R_{vdW})$  to be the dispersion interaction between a pair of one-dimensional QDOs representing the molecule evaluated at  $R_{vdW}$  the Van der Waals diameter of the molecules. Assuming distributions of quantum states of the uniform spherical measure as well as a pure uncoupled state, we see that the grouping method derived in this work can reduce the required number of shots by multiple orders of magnitude. Further reductions may be possible if heuristic methods are used to find more optimal groups.

For the prototypical orbital-based VQE methods there are in general  $O(N^4)$  Pauli operators that must be measured where  $N$  is the size of the molecular orbital basis set used. The best known scaling for the number of commuting operator collections is  $O(N^3)$  [48] which, when compared to the order  $O(N)$  scaling in terms of the number of QDOs for the dipole coupled Hamiltonian here, suggests that the system we consider in this work may be significantly more scalable than those considered in orbital-based VQE methods. To our knowledge, no attempt has been made to find a scaling for the number of shots required for these.

#### D. Variational ansatz

For the results presented in this paper, we use a variational ansatz consisting of layers of two-qubit operators acting on

nearest-neighbor qubits. Since the problem Hamiltonian is positive, the ground state will be real-valued (up to a global phase) and we restrict ourselves to using real-valued, two-qubit operators based on the optimal construction of general SO(4) operators found by Vatan and Williams [65]. We removed two parameterized  $R_z$  gates from their construction to reduce the number of variational parameters. This naturally restricts the available operations implemented by each block to a subset of SO(4) but was found to still give reasonable results in simulation and experiment while reducing the number of circuits required for gradient measurements. The circuit used for two QDOs is shown in Fig. 4(a), which has a total of 12 variational parameters, the two-qubit circuit blocks are given in Fig. 4(b).

Although more costly than ansätze designed to make efficient use of the quantum hardware [33], quantum advantage may be achieved for *ab initio* VQE by using chemistry-inspired ansätze such as unitary coupled cluster (UCC) [66]. Analogous ansätze for bosonic systems can be constructed [36,37,67], namely, the unitary-vibrational-coupled-cluster (UVCC) ansatz (following the naming convention of McArdle *et al.* [36]). This is a unitary extension to the vibrational coupled cluster (VCC) used in classical computations [68,69] and, unlike the VCC ansatz, the UVCC ansatz obeys the variational principle and may deal better with strong static correlations [36].

The UVCC ansatz will lead to significantly larger circuits than our ansatz in Fig. 4 due to the nontrivial decomposition of bosonic operators into Pauli operators that can be implemented natively on quantum devices, thus we do not consider its application in this work. The analysis of commutation relations presented in Appendix C as well as various encoding methods [67] will be important in understanding the experimental cost of the Trotterization required for such an ansatz.

### III. EXPERIMENTAL RESULTS USING IBM QUANTUM COMPUTERS

To demonstrate the feasibility of calculating weak intermolecular interactions from the coarse-grained model we represent two linear, nonpolar molecules at distance  $R$  apart using two one-dimensional QDOs. We choose model parameters matching those of  $I_2$ , which is known to produce a dimer molecular liquid above its melting point of 114 °C in which dispersion forces dominate the cohesion.

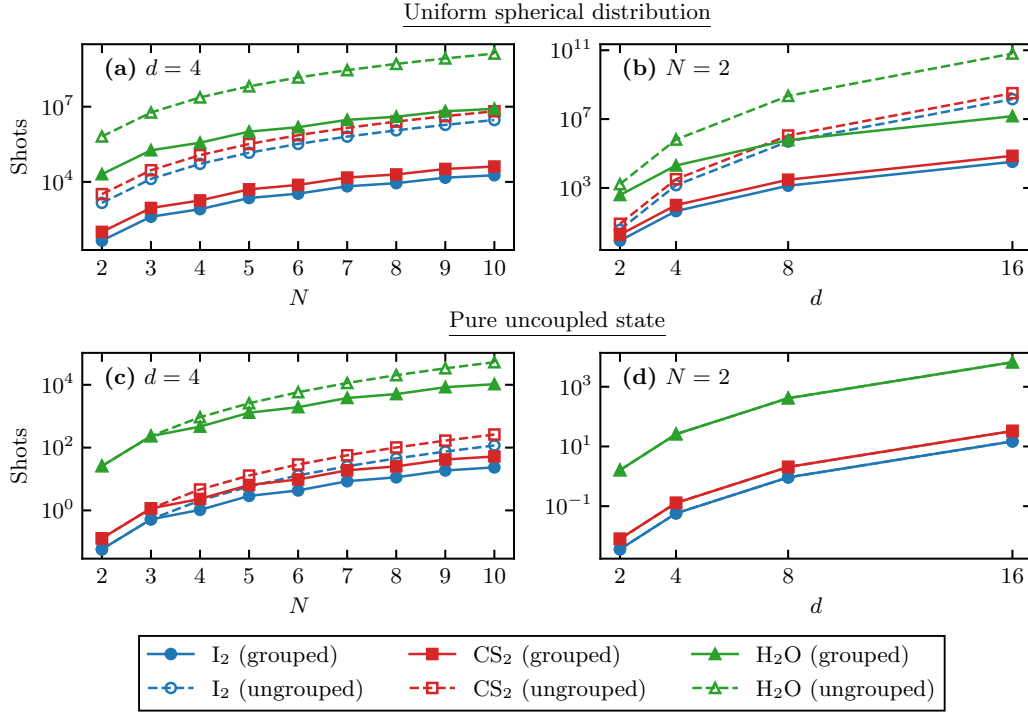


FIG. 3. Estimated number of shots required to measure  $\langle \hat{H}(R_{vdW}) \rangle$  with error  $\epsilon = 10\% \times \Delta E(R_{vdW})$  for a range of different molecules. Here  $R_{vdW}$  is the van der Waals diameter of the respective molecules. The expected number of shots is calculated with respect to the spherically uniform measure (a), (b) as well as with the uncoupled state (c), (d) with and without grouping commuting operators. The number of shots is shown both as a function  $N$  for fixed  $d = 4$  and as a function of  $d$  for fixed  $N = 2$ . Note that in (d), the grouped and ungrouped lines are the indistinguishable from one another. Values for  $R_{vdW}$  and  $\alpha$  are taken from Refs. [54–57].

For each QDO, we use a  $d = 4$  level Fock basis, leading to a total of four qubits required for the two-molecule system. We model the pair of  $I_2$  molecules, both aligned with their molecular axis along the intermolecular axis such that the one-dimensional QDO models their induced polarization along the molecular axis. We use a value for the polarizability of  $\alpha = \alpha_{zz} = 14.5 \text{ \AA}^3$  the polarizability coefficient along the molecular axis [54] and a harmonic trapping frequency of  $\hbar\omega = 9.61 \text{ eV}$ . For the coarse-grained electronic model, it is usual to choose the  $\omega$  parameter (or alternatively the point charge  $q$  and mass  $\mu$  if working with raw Drude particle parameters). This is done to capture a certain chosen property, such as the  $C_6$  London dispersion coefficient, and then further terms (such as  $C_8$  and  $C_9$  terms) arise naturally from the model. In this case, we choose  $\hbar\omega$  such that the full three-dimensional QDO in this setup would give  $C_6$  coefficient matching experimental measurements of  $I_2$  [70]. For the pair of interacting molecules, the full three-dimensional model and the total London dispersion energy for the pair of interacting molecules can be found from repetitions of the one-dimensional model using the correct possibilities for each remaining axis a discussion of which is deferred until Sec. IV A. As proof of principle, we consider polarizability only along the intermolecular axis. To generate the dispersion energy as a function of intermolecular separation  $R$ , we perform the VQE procedure to calculate the ground-state energy of the coupled QDOs for a range of possible coupling constants.

For the experimental optimization procedure we used, for  $R = \infty$ , the quantum computer *ibmq\_santiago* and, for all

other values of  $R$ , *ibmq\_montreal* (for further details about the hardware used in this work, see Appendix F). Each evaluation of  $\langle \hat{H} \rangle$  is performed by running ten circuits with 8192 shots each where the 21 Pauli operators needed to estimate  $\langle \hat{H} \rangle$  are collected into ten groups of operators that qubit-wise commute as described in Sec. II B. We evaluated gradients using the parameter shift rule [71] and performed the optimization using the ADAM optimizer (learning rate 0.25, decay rates 0.9, 0.99) [72]. For each value of  $R$ , the VQE procedure was run for 200 optimization steps with the final set of parameters taken as the output. For each experiment, the values of  $\theta$  are independently randomly initialized to values taken uniformly on the interval  $[-\pi/2, \pi/2]$ .

Once the optimal set of parameters was found, we extracted the binding energy from the noisy quantum computer using a simple scheme to subtract the effect of device noise. The binding energy due to the London dispersion interaction at a certain  $R$  is then given by

$$\Delta E(R) = E(R) - E(R = \infty), \quad (7)$$

where  $E(R)$  is the ground-state energy at  $R$  and with  $E(R = \infty)$  corresponding to the ground-state energy of the completely uncoupled state. By modeling the output of the noisy quantum computers as a unitary channel depending on variational parameters  $\theta$  along with depolarizing noise with error rate  $\lambda$ , the output state for a given set of variational parameters is

$$\rho(\theta) = (1 - \lambda)|\psi(\theta)\rangle\langle\psi(\theta)| + \frac{\lambda}{2^{mN}}\mathbb{1}, \quad (8)$$

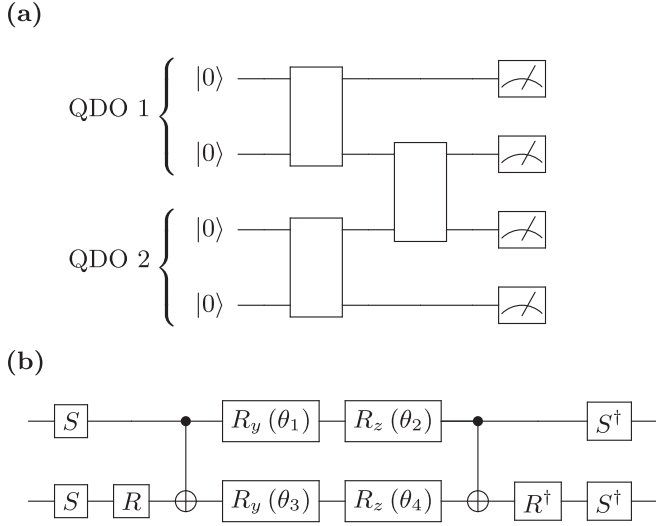


FIG. 4. Circuits used in experiments. (a) Four qubit variational ansatz representing the state of two  $d = 4$  QDOs. The upper (lower) qubit within the two-qubit register for each QDO corresponds to the most (least) significant bit in the binary encoding of the corresponding QDO Fock states. Each two-qubit block consists of a four-parameter real-valued operator. (b) Composition of the four-parameter real-valued operators.  $R_z$  and  $R_y$  gates in the center are Pauli Z and Y rotation gates, each one contains a single variational parameter. The remaining gates are fixed with  $S = R_z(\frac{\pi}{2})$  and  $R = R_y(\frac{\pi}{2})$ . This block is based on the general SO(4) operator from Ref. [65] with two parameterized  $R_z$  gates removed.

where  $|\psi(\boldsymbol{\theta})\rangle$  is the output state from a noiseless ansatz with parameters  $\boldsymbol{\theta}$ . Assuming  $|\psi(\boldsymbol{\theta})\rangle$  is the required ground state, this corresponds to a noisy expectation value for the Hamiltonian

$$\langle \hat{H}(R) \rangle_{\rho(\boldsymbol{\theta})} = (1 - \lambda)E(R) + \frac{\lambda}{2^{mN}} \text{Tr}[\hat{H}(R)]. \quad (9)$$

Therefore, assuming  $\lambda$  is independent of the variational parameters and does not change between experiments, the dispersion energy is given by

$$\Delta E(R) = \frac{\langle \hat{H}(R) \rangle_{\rho(\boldsymbol{\theta})} - \langle \hat{H}(R = \infty) \rangle_{\rho(\boldsymbol{\theta}_\infty)}}{(1 - \lambda)}, \quad (10)$$

where  $\boldsymbol{\theta}_\infty$  are variational parameters found for the ground state of  $\hat{H}(R = \infty)$ , here we have used that  $\text{Tr}[\hat{H}(R)] = \text{Tr}[\hat{H}_{\text{nonint}}] = dN \forall R$ . We use (10) to calculate the dispersion energy from noisy measurements, estimating  $\epsilon$  by measuring the fidelity of the uncoupled ground state on the noisy computer  $\mathcal{F} = \text{Tr}[\rho(\boldsymbol{\theta}_\infty)|0\rangle\langle 0|^{\otimes mN}] = (1 - \lambda) + \frac{\lambda}{2^{mN}}$ . This can be done without additional cost by reusing the counts measured when calculating  $\langle \hat{H} \rangle_{\rho(\boldsymbol{\theta}_\infty)}$ . We find that, for the devices and circuits we use in this work, this method to subtract the device noise gives accurate values for the dispersion energy.

To measure the dispersion energy for the optimized states we use the *ibm\_lagos* quantum computer. One practical consideration when performing this subtraction, is that due to the significant time dependence of the noise levels of the device, it is important to evaluate the expectations for different values of  $R$  in (10) with the same qubits on the same device as close to

simultaneously as possible. For the results presented here we took the mean of 200 repetitions of each measurement of  $\langle \hat{H} \rangle$  (with 8192 shots per circuit), performing the measurements for each  $R$  sequentially before moving onto the next of the 200 repetitions. As many measurements were performed as possible in a single job and all jobs were submitted simultaneously to minimise the time between successive jobs. Figure 5 shows how the level of noise varies significantly throughout the 200 evaluations but the drift seen in the expectation values is approximately matched for each of the different values of  $R$  when measured in this way.

The spectrum achieved by the optimization procedure in this proof-of-principle experiment is shown in Fig. 2. From this, we see that by using the variational procedure combined with our ansatz, optimal parameters can be found using noisy gradients calculated on the real device that correspond to energy values close to the true ground state when evaluated without device noise. Furthermore, although the raw experimental energies are much larger than even the uncoupled energy of the system, using the simple subtraction method described above, we are able to successfully measure dispersion energies on the device closely matching the exact London dispersion interaction. From Fig. 2 we see that for all values of  $R$  tested, the binding energy using the quantum processor is within two standard deviations of the exact value. It is worth noting that the value of the ground-state energy with the same variational parameters but evaluated using noiseless simulation [shown in Fig. 5(b)] are all slightly above the true value given by the curve. This could be due to either imperfect or unconverged optimization or systematic errors in the rotation gates applied by the device which are accounted for as over- or underrotations in the variational parameters learnt by the algorithm and not reflected in the error-free simulation. We expect that these effects will be reduced by improvements in hardware performance and by using better gradient descent schemes or better fine tuning of the settings for the ADAM optimization used.

## IV. EXTENDING THE MODEL

### A. Three-dimensional QDOs

As described in Sec. II A, for a complete description of the dipole-dipole interaction and higher order terms, full three-dimensional QDOs are required. Here we demonstrate how our representation and variational algorithm can be extended to three-dimensional QDO systems.

The Hamiltonian for such a system is given by

$$\hat{H} = \sum_{i=1}^N \left( \frac{1}{2\mu} \hat{\mathbf{p}}_i^2 + \frac{1}{2} \mu \omega^2 \hat{\mathbf{x}}_i^2 \right) + q^2 \sum_{j>i} \frac{3(\hat{\mathbf{x}}_i \cdot \mathbf{R}_{i,j})(\hat{\mathbf{x}}_j \cdot \mathbf{R}_{i,j}) - \hat{\mathbf{x}}_i \cdot \hat{\mathbf{x}}_j R_{i,j}^2}{R_{i,j}^5}, \quad (11)$$

where  $\hat{\mathbf{x}}_i = (\hat{x}_i, \hat{y}_i, \hat{z}_i)$  and  $\hat{\mathbf{p}}_i = (\hat{p}_{x_i}, \hat{p}_{y_i}, \hat{p}_{z_i})$  are the position and momentum operators for QDO  $i$  in each of the three Cartesian coordinates.  $\mathbf{R}_{i,j}$  is the vector separating the centers of the pairs of QDOs  $i$  and  $j$  and  $\mu$ ,  $\omega$ , and  $q$  represent the effective mass, natural frequency, and charge for the QDO



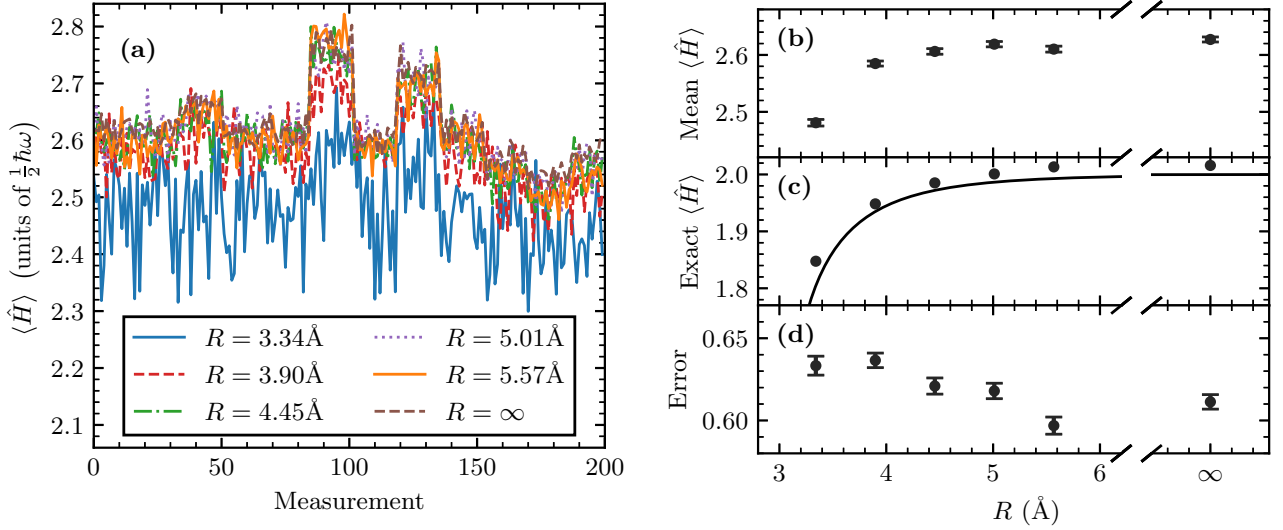


FIG. 5. Experimental measurements used for error subtraction. (a) Measurements of  $\langle \hat{H}(R) \rangle_{\rho(\theta_R)}$  where  $\theta_R$  are the optimal parameters found for each separation  $R$  evaluated on *ibm\_lagos* quantum computer. As described in the main text, each of the 200 experiments is performed for each value of  $R$  before moving onto the next one. (b) Mean values of the measurements for each  $R$ . (c) Exact calculation of  $\hat{H}$  with no noise for the optimized parameters. Solid curve shows the true ground state. (d) Absolute difference between exact value of  $\langle \hat{H} \rangle$  and mean of noisy measurements. This shows that, although the effect of noise varies with each experiment, by measuring each  $R$  in the order described the effect of noise is approximately equal for each data point. Points at  $R = \infty$  correspond to the uncoupled Hamiltonian and parameters found using the optimization procedure.

particles, which, for simplicity, are assumed to be isotropic and equal for all oscillators.

To represent the state of these three-dimensional oscillators on a quantum computer, we treat each of the three spatial dimensions as a separate harmonic oscillator with position and momentum operators  $\{(\hat{x}, \hat{p}_x), (\hat{y}, \hat{p}_y), (\hat{z}, \hat{p}_z)\}$ . As in Sec. II B, each of these one-dimensional oscillators can be represented using a binary encoding in  $m$  qubits. Thus, the state of each three-dimensional oscillator state is encoded in a register of  $3m$  qubits:

$$\mathcal{H}_{d\text{-Fock}}^{\otimes 3} \rightarrow \mathcal{H}_2^{\otimes m} \otimes \mathcal{H}_2^{\otimes m} \otimes \mathcal{H}_2^{\otimes m} : \quad (12)$$

$$|n_x, n_y, n_z\rangle \rightarrow |\text{bin}_m(n_x), \text{bin}_m(n_y), \text{bin}_m(n_z)\rangle.$$

Following Sec. II B, an upper bound for the number of circuits that are required to measure  $H$  is  $(N-1)(\frac{d}{2} \log_2 d)^2 + 1$  or  $N(\frac{d}{2} \log_2 d)^2 + 1$ , for even and odd  $N$ , respectively. See Appendix C for details.

Extending the two-oscillator system to three dimensions, the Hamiltonian simplifies to

$$\hat{H} = \sum_{u \in \{x, y, z\}} (\hat{u}_1^2 + \hat{p}_{u_1}^2) + (\hat{u}_2^2 + \hat{p}_{u_2}^2) + \gamma_{uu} \hat{u}_1 \hat{u}_2, \quad (13)$$

where  $\gamma_{xx} = 2\alpha_{xx}/R^3$ ,  $\gamma_{yy} = 2\alpha_{yy}/R^3$  and  $\gamma_{zz} = -4\alpha_{zz}/R^3$ ,  $\alpha_{xx}$ ,  $\alpha_{yy}$ ,  $\alpha_{zz}$  represent the dipole polarisability along each of the three axes (which we now allow to be anisotropic, corresponding to  $q$  being an anisotropic tensor) and the intermolecular axis is in the  $z$  direction, along which the physical separation is  $R$ . We have ignored dimensions to match the structure of (1). We see from this that the Hamiltonian is equivalent to three independent one-dimensional oscillators with the Hamiltonian matching (1) for  $N = 2$  with the appropriate choices of coupling constants. Therefore, in the previous section where we modeled the interaction of two linear molecules aligned

along an axis, an extension to a full three-dimensional system follows trivially.

## B. Induction, anharmonic, and nonlinear terms

The quadratic Hamiltonians that we have presented so far can easily be extended to include higher-order and anharmonic terms. Here we briefly describe three extensions of the model giving rise to additional interactions and phenomena. We also discuss how the decomposition and grouping strategies described in Sec. II C can be applied to these richer systems.

The cheapest and simplest modification is the inclusion of induction effects by applying a uniform electric field  $E$  to the oscillators. This simple modification can, for two or more oscillators, lead to nonadditive effects [43]. For one-dimensional oscillators this leads to an interaction of the form  $\hat{H}_{E\text{-field}} = \sum_{i=1}^N q_i E \hat{x}_i$  for  $N$  Drude oscillators with charges  $q_i$ . These can be measured using the same decompositions for  $\hat{x}$  as were used previously since these interactions are guaranteed to qubit-wise commute with terms of the form  $\hat{x} \otimes \hat{x}$  incurring no additional circuits to measure the total Hamiltonian.

For more realistic induction models, we can instead consider oscillators in a Coulomb potential caused by one or more charges. In this case, the Hamiltonian will have the form  $\hat{H}_{\text{Coulomb}} = \sum_{i=0}^N \sum_{j=0}^M \phi_j(\hat{x}_i)$  for  $N$  oscillators in Coulomb potentials represented by  $\phi_j$  caused by  $M$  fixed charges. To measure a Hamiltonian of this form, each of the Coulomb potentials must be Taylor expanded up to some finite order  $w$ :

$$\phi_j(\hat{x}_i) \sim \frac{1}{|r_j - \hat{x}_i|} \approx \frac{1}{r_j} + \frac{\hat{x}_i}{r_j^2} + \frac{\hat{x}_i^2}{r_j^3} + \dots + \frac{\hat{x}_i^{(w-1)}}{r_j^w}, \quad (14)$$

where  $r_j \gg |\hat{x}_i|$  is the position of the point charge relative to the center of the QDO. Again, since each term of the form  $\hat{x}_i^w$  acts on distinct oscillators they can all be measured simultaneously. The expected contribution to the number of shots will be  $O(\sqrt{N})$  with respect to  $N$  thus leaving the asymptotic scaling unchanged when added to the linear dipolar Hamiltonian of (1).

Using a similar addition to the Hamiltonian we are able to relax the assumption of a harmonic trapping potential and consider anharmonic trapping potentials that have not been accessible for large systems using classical computation methods [51]. Additional terms of the form  $\hat{H}_{\text{anharm}} = \omega_3 \hat{x}_i^3 + \dots + \omega_w \hat{x}_i^w$  with appropriate model parameters defining coefficients  $\{\omega_3, \omega_4, \dots, \omega_w\}$  can be used to choose a desired, potentially asymmetric, trapping potential. The measurement cost will scale in the same way as terms in (14). Relaxing the harmonic assumption is required for inclusion of some higher order terms (such as four-dipole hyperpolarizability [51,73]) and leads to nonlinear optical effects such as the Kerr effect, second- and third-harmonic generation, and forms of optical mixing [74].

Finally, we introduce nonlinear interaction terms. As an example we consider interactions of the form  $\hat{H}_{\text{non-lin}} = \sum_{i,j,k=0}^N \beta_{(i,j,k)} \hat{x}_i \hat{x}_j \hat{x}_k$ . Terms similar to this have been considered for use in quantum algorithms for vibrational spectra, noting that these are difficult to deal with classical computational methods [36,75]. If all  $O(N^3)$  possible terms in  $\hat{H}_{\text{non-lin}}$  are nonzero, using the exact grouping that we have described will contribute an  $O(\frac{d^9 N^5}{\epsilon^2})$  cost to the number of shots required. In practice the number of shots may be considerably lower as heuristic sorting methods will make use of additional qubit-wise commutation not considered in the grouping method we have used in this work. For higher order interactions, these commuting pairs will likely be more common. Also, many physical systems may not include all-to-all interactions or will have interactions that decay strongly with distance such that many of the interaction terms can be ignored. Terms of this form, as well as higher order nonlinear interaction terms will allow for a large amount of flexibility in the model and will allow for the investigation of a wide range of physical and chemical phenomena.

## V. FURTHER EXPERIMENTAL RESULTS

### A. Anharmonic systems

To test the performance of the algorithm for a system of anharmonic oscillators, we investigate two systems with 1D anharmonic QDOs on both simulated and real quantum computers.

We first consider the system of a single oscillator in an external potential, modifying the onsite potential to include a quartic term. In this case, the Hamiltonian is given by  $\hat{H} = \hat{x}^2 + \hat{p}^2 + \frac{1}{5}\hat{x}^4 + E\hat{x}$ , where parameter  $E$  is a uniform external field which is varied. We used two qubits to give a truncated Fock space of  $d = 4$  and an ansatz consisting of the single four parameter block given in Fig. 4(b).

Optimization was performed using a simulated version of the *ibmq\_montreal* device and using the same optimizer and

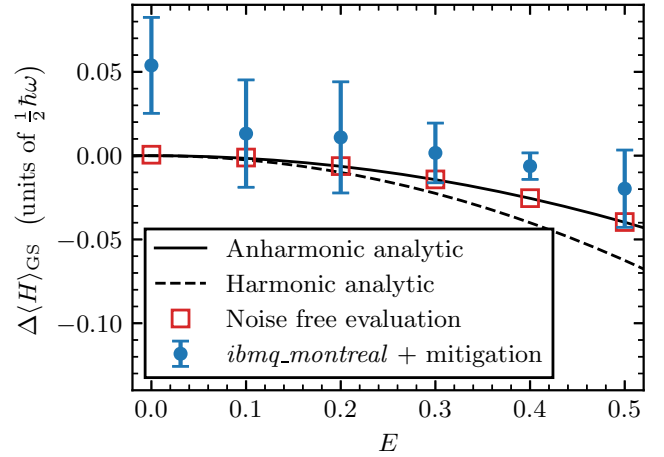


FIG. 6. Deviation from zero field ground-state energy of 1D anharmonic oscillator in external field  $E$ . Blue circles correspond to evaluating the energy on the real *ibmq\_montreal* quantum processor using error mitigation techniques described in the main text. Red square data points correspond to evaluating the final energy with no device or shot noise. Both sets of data points use the same variational parameters which were found through optimization performed on a simulated version of *ibmq\_montreal*. Black solid line shows the exact spectrum found by numerically diagonalizing the Hamiltonian. Black dashed line shows the exact spectrum for the analogous harmonic system (setting  $x^4$  term to zero).

settings as in Sec. III for 200 optimization steps. As with the experiments in Sec. VB, for the first data point at  $E = 0$ ,  $\theta$  were initialized to values taken randomly on the uniform interval  $[-\pi/2, \pi/2]$ . Proceeding data points were then run in order of increasing  $E$  initializing  $\theta$  to the optimized values linearly extrapolating the optimized parameters from the previous data points with respect to the two corresponding values of  $E$ .

Once the optimization was completed, we evaluated the final energy value using the optimized parameters on the real *ibmq\_montreal* device to see how real device noise affected the accuracy of the energy measured. Since the system was no longer harmonic, the zero field ground state was not known to be the zero state and so the subtraction method given in Eqn. (10) could not be used to remove device noise from the measured value. Instead two error mitigation techniques—zero noise extrapolation [76,77] and one based on the Lanczos method [78]—were used in combination to measure the final ground-state energy. See Appendix E for details on these two methods, as well as experimental settings used in this work.

The results of this procedure are shown in Fig. 6. We see that values of the ground-state energy evaluated without noise agree well with the true value found by directly diagonalizing the Hamiltonian matrix. As in previous experiments, this shows that, despite the simulated device noise in the optimization procedure, a good set of variational parameters are found by the algorithm and that the ansatz is able to accurately generate the required ground states. We see that, when combined with error mitigation, the interaction energy of the system shows some agreement with the true value; however, the uncertainties in the values mean that we are not

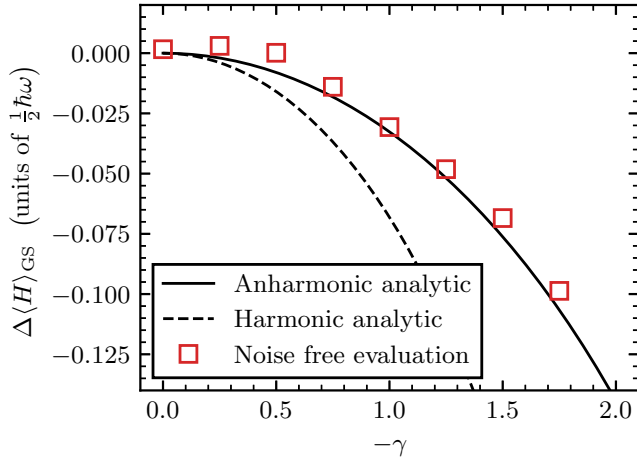


FIG. 7. Deviation from uncoupled ground-state energy for system of two coupled anharmonic oscillators with optimization performed on simulated *ibmq\_montreal* device. Data points show ground-state energy relative to true ground-state energy of uncoupled system, evaluated without device or shot noise. Solid line gives exact spectrum of the anharmonic system calculated by direct matrix diagonalization. Dashed line gives exact spectrum of the corresponding harmonic system.

able to resolve the difference in interaction energies between the harmonic and anharmonic systems. We note that in this system, the energy resolution required is significantly smaller than the dispersion interaction of Sec. III.

An increase in the measurement uncertainty is a consequence of both error mitigation techniques. This uncertainty can be decreased, as well as the overall reduction in the error, by using a larger number of shots (thereby increasing runtime) or by improving error rates on the quantum computer.

We additionally perform experiments for a system of two linearly coupled oscillators (the system using in Sec. III with additional quartic onsite potential terms). For such a system the Hamiltonian is given by  $\hat{H} = \hat{x}_1^2 + \hat{p}_1^2 + \frac{1}{5}\hat{x}_1^4 + \hat{x}_2^2 + \hat{p}_2^2 + \frac{1}{5}\hat{x}_2^4 + \gamma\hat{x}_1\hat{x}_2$ . We again use two qubits per oscillator and use the ansatz shown in Fig. 4(a) and the same optimization settings as in Sec. III. Again, values of  $\theta$  were randomly initialized for  $\gamma = 0$  data point and initial values of  $\theta$  for remaining data points were chosen by extrapolating from previous optimized values. Optimization was performed using a simulated version of the *ibmq\_montreal* device and we calculated the final values for the ground-state energy found using exact simulation, removing the effect of device or shot noise. The results of these experiments are given in Fig. 7. These show a good agreement between the true ground-state energy and results found through the optimization procedure.

### B. Many oscillators

We also test our variational algorithm for more than two-body systems, namely,  $N = 3, 4$ , and 5 oscillators. Using a Hamiltonian of the form Eq. (1) again, coupling constants  $\gamma_{i,j}$  were chosen to represent systems of 1D QDOs arranged in regular  $N$ -sided polygons and allowed to os-

cillate perpendicular to the interoscillator plane. For these systems,  $\gamma_{i,j} = 2\alpha/R_{i,j}^3 \propto 2\alpha/D^3$  where  $R_{i,j}$  are the distances between oscillators  $i$  and  $j$ ,  $D$  gives the diameter of the circle that effectively encloses the  $N$  vertex polygons. By varying the ratio of  $2\alpha/D^3$  (effectively varying the size of the polygons) we generate the spectra shown in the results of Fig. 8.

For the results shown in Fig. 8, we simulate the performance of our VQA using a range of different ansätze depths. We use an exact simulation for the optimization procedure and final energy measurement so there are no contributions from finite number of shots or device noise. As a variational ansatz, we use a similar construction to that of the ansätze used in Sec. III with the slight modification of using cyclic connectivity such that there are additional SO(4) circuit blocks between the first and last qubits in even numbered layers. We represent each QDO with two qubits such that the local dimension of the oscillators is  $d = 4$  and experiments require six, eight, and ten qubits for  $N = 3, 4$ , and 5, respectively. For each data point, 500 optimization steps and the same optimization settings as in Sec. III. For the  $\alpha/D^3 = 0$  data points,  $\theta$  were initialized to values taken uniformly on the interval  $[-\pi/2, \pi/2]$ . Proceeding data points were then run in order of increasing  $\alpha/D^3$  initializing  $\theta$  to the optimized values for the previous data point (in the case for the second data point) or by linearly extrapolating the optimized parameters from the previous data points with respect to the two corresponding values of  $\alpha/D^3$  (for all other data points).

From the results in Fig. 8, we see that the variational algorithm is able to successfully recreate the energy spectrum of the three, four, and five QDO systems. Of particular interest is the spectrum of the three-oscillator case in which the three-body contribution, similar to the full Axilrod-Teller interaction energy [79], makes a significant contribution to the dispersion energy and is accurately recreated by the  $d = 4$  truncated Fock space as well as the variational algorithm. Across all three systems, over the coupling parameter values considered, the truncated spectrum corresponding to the truncated Fock space matches well the exact ground-state energy calculated through symplectic methods (see Appendix G for details on how this was calculated). As the system size increases, a deeper ansatz circuit with more variational parameters is needed to capture reasonable estimates of the ground-state energy.

## VI. DISCUSSION AND CONCLUSIONS

In this work we have developed a variational quantum algorithm for modeling intermolecular interactions using a coarse-grained framework. We have introduced an encoding scheme, variational ansatz, and variational algorithm for digital quantum computers assessing measurement cost and scaling. We presented an exact method for grouping mutually commuting operators that can be measured simultaneously which gave bounds on the number of different circuits and number of shots required to measure a system with dipole-dipole interactions. We have shown evidence that this scaling may be more favorable than for existing VQE methods. We have also presented a pathway for adding anharmonic and higher order terms that are not easily accessible to classical

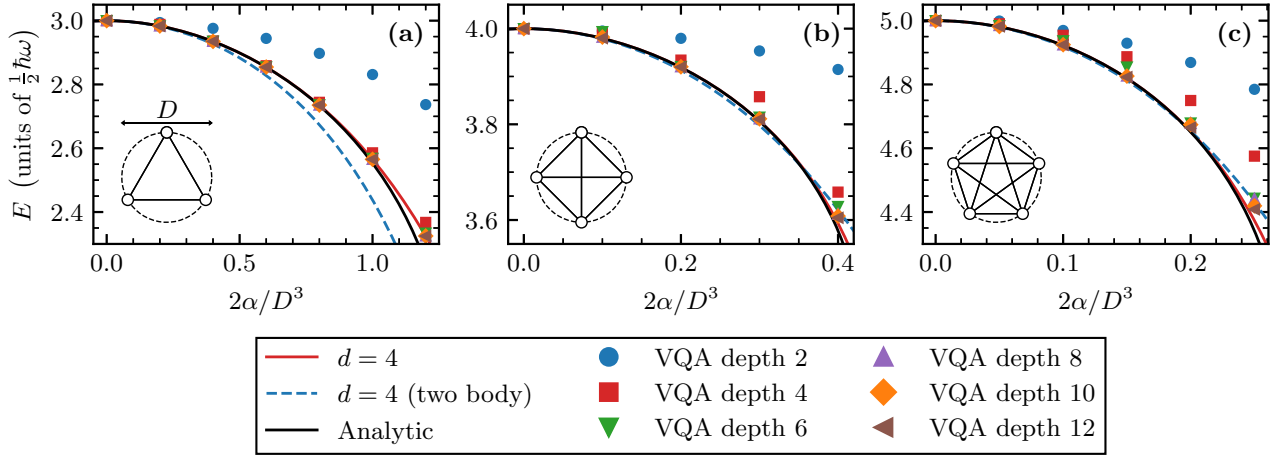


FIG. 8. Simulation of VQA for systems of (a) three, (b) four, and (c) five QDOs arranged in regular polygons using ansätze of varying depth. Insets show respective oscillator geometries, labeling the diameter  $D$  of the circle enclosing the polygons which was effectively varied to generate the spectra. Red solid line (labeled  $d = 4$ ) corresponds to numerically diagonalizing Hamiltonian of oscillators of  $d = 4$  energy levels. Blue dashed line corresponds to the same system but only considering the two body pairwise interactions. Black solid line (labeled analytic) corresponds to using symplectic representation of many oscillator system (as described in Appendix G) which effectively corresponds to  $d = \infty$ . Data points correspond to outputs of VQA algorithm for varying depth ansätze.

computational methods. This, combined with the favorable scaling, suggests that our algorithm may be a more promising candidate for showing quantum advantage than other variational methods.

As with other applications of VQE methods, the global scaling of the algorithm presented—namely, how many optimization steps are required to find the ground state within a given accuracy—is not yet known. That the classical optimization required for variational quantum algorithms is known to be NP-hard [80] and the occurrence of barren plateaus [81–84]—exponentially vanishing gradients that mean ansätze cannot be optimized efficiently—may limit the scalability of the variational approach. As seen in the field of classical machine learning however, worst case complexity theoretic arguments [85] do not necessarily preclude the use of such methods for practical applications, particularly as algorithmic improvements can be made to mitigate such issues. There is already a significant body of work on methods to address the issue of barren plateaus. These include investigating ansatz structure [86–88] and parameter initialization strategies [89–91] among others [92–94]. We expect that many methods found to avoid barren plateaus for VQAs in a general setting or for specific cost functions will likely be applicable to the work presented here. For the problem presented here, a more physically inspired ansatz (e.g., UVCC described in Sec. IID) may be needed to avoid barren plateaus induced by approximate 2-designs [81] within the ansatz that we have used. Additionally, the cost function we use in this work is local and may be more resilient to barren plateaus than other applications of VQAs [82,95].

We have experimentally demonstrated that van der Waals dispersion energies are resolvable on current generation quantum computing hardware and that they lie within physically realistic ranges found in molecular materials. This represents the first use of existing quantum computing capability to access the weak but influential intermolecular energy

scales arising from noncovalent interactions (including correlated quantum fluctuations) relevant in condensed states of molecular assemblies. While the specific model used in this experiment can also be solved on classical processors, the impetus to develop and demonstrate an efficient and well characterized quantum implementation includes the following:

First, it expands the palette of problems that can be accessed on quantum computing systems by providing a further level of electronic coarse graining while preserving a complete set of responses, fluctuations and correlation phenomena at long range. This methodology could therefore be combined with other electronic structure strategies allowing only selected electronic subsystems to be treated at the highest level of accuracy. Such combined hybrid approaches are likely to be required to treat complex or multiscale problems even if full CI computations can be realized in polynomial time complexity. Having established that dispersion interactions are detectable, extensions to the current implementation can be included to move up the interaction hierarchy to capture induction and many-body polarization thereby opening opportunities for versatile molecular simulation problems on quantum architecture.

Second, this framework may also be used to enrich existing simple models such as bead based descriptions of protein and branched polymer folding which can approach classically intractable combinatorial complexity and are already being actively explored for implementation in quantum architecture [96,97]. Including coarse-grained representations of the long-ranged interactions will result in more realistic treatments of model folding problems.

Finally, we have shown that the aforementioned extension to the model, in which anharmonic on-site binding potentials can be introduced with low experimental overhead, can already be solved within our VQA and how the measured anharmonic energies are affected by device noise on superconducting processors. This paves the way for using quantum

algorithms for anharmonic systems not easily accessible to classical algorithms and should lead to a better understanding of how anharmonic effects can be used to extend coarse-grained models for molecules as well as how new physics may arise in anharmonic many-body systems.

### ACKNOWLEDGMENTS

L.W.A. would like to thank Christopher Willby for helpful discussions about symplectic methods for quadratic systems. L.W.A., M.K., and D.J. acknowledge support from the EPSRC National Quantum Technology Hub in Networked Quantum Information Technology (EP/M013243/1) and the EPSRC Hub in Quantum Computing and Simulation (EP/T001062/1). M.K. and D.J. acknowledge financial support from the National Research Foundation, Prime Minister's Office, Singapore, and the Ministry of Education, Singapore, under the Research Centres of Excellence program. The authors would like to acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility in carrying out this work [98]. This work was also supported by

the Hartree National Centre for Digital Innovation program, funded by the Department for Business, Energy and Industrial Strategy, United Kingdom.

### APPENDIX A: LONDON DISPERSION IN LINEAR MOLECULES AND ONE-DIMENSIONAL QUANTUM DRUDE OSCILLATORS

This Appendix shows how the interaction between two axially symmetric linear molecules can be captured by one-dimensional QDOs. The results and derivations given here can be found in the literature but are combined and included here for the reader.

In the framework developed by London [99], the interaction between two anisotropic linear molecules that are axially symmetric about their respective principle intramolecular bond can be described by ellipsoids of charge. For a pair of identical linear molecules that share components of the dipole polarizability  $\alpha_{\parallel}$  ( $\alpha_{\perp}$ ) and characteristic angular frequencies  $\omega_{\parallel}$  ( $\omega_{\perp}$ ) parallel (perpendicular) to their intramolecular axis, the London dispersion energy at large separation  $R$  is

$$\Delta E = -\frac{1}{R^6} \{ (C_{\parallel,\parallel} + C_{\perp,\perp} - C_{\parallel,\perp}) [\sin^2(\theta_a) \sin^2(\theta_b) \cos(\phi_a - \phi_b) - 2 \cos(\theta_a) \cos(\theta_b)]^2 + 3(C_{\parallel,\perp} - C_{\perp,\perp}) [\cos^2(\theta_a) + \cos^2(\theta_b)] + 2(C_{\parallel,\perp} + 4C_{\perp,\perp}) \}, \quad (\text{A1})$$

where  $\theta_a, \phi_a, \theta_b, \phi_b$  describe the orientation of molecules  $a$  and  $b$  relative to the intramolecular axis.  $C_{\cdot,\cdot}$  coefficients are given by

$$C_{\parallel,\parallel} = \frac{1}{8} \alpha_{\parallel}^2 \hbar \omega_{\parallel}, \quad C_{\perp,\perp} = \frac{1}{8} \alpha_{\perp}^2 \hbar \omega_{\perp}, \\ C_{\parallel,\perp} = \frac{1}{4} \alpha_{\parallel} \alpha_{\perp} \hbar \frac{\omega_{\parallel} \omega_{\perp}}{\omega_{\parallel} + \omega_{\perp}}. \quad (\text{A2})$$

For further details about the geometry and derivation see Ref. [100]. When the two molecules are aligned parallel to each other and the intermolecular axis  $\theta_a = \theta_b = 0$  and  $\phi_a = \phi_b$  (which can be set to 0 without loss of generality) in the completely anisotropic limit  $\alpha_{\parallel} \gg \alpha_{\perp}$  the London dispersion interaction is given by

$$\Delta E = -\frac{\alpha_{\parallel}^2 \hbar \omega_{\parallel}}{2R^6}. \quad (\text{A3})$$

We now show how two interacting one-dimensional QDOs can recreate this dispersion interaction. The Hamiltonian for pair of one-dimensional QDO is

$$\hat{H} = \frac{\hat{p}_1^2}{2\mu} + \frac{\hat{p}_2^2}{2\mu} + \frac{1}{2} \mu \omega^2 (\hat{x}_1^2 + \hat{x}_2^2 + \gamma \hat{x}_1 \hat{x}_2), \quad (\text{A4})$$

where  $\gamma = -4 \frac{q^2}{\mu \omega^2} \frac{1}{R^3}$ . Alternatively  $\gamma = -4 \frac{\alpha}{R^3}$  when written in terms of the dipole polarizability  $\alpha = \frac{q^2}{\mu \omega^2}$ . Defining the transformed variables

$$\hat{p}_{\pm} = \frac{1}{\sqrt{2}} (\hat{p}_1 \pm \hat{p}_2), \quad \hat{x}_{\pm} = \frac{1}{\sqrt{2}} (\hat{x}_1 \pm \hat{x}_2), \quad (\text{A5})$$

the Hamiltonian can be written in this uncoupled basis as

$$\hat{H} = \left[ \frac{\hat{p}_+^2}{2\mu} + \frac{1}{2} \mu \omega^2 \left( 1 + \frac{\gamma}{2} \right) \hat{x}_+^2 \right] + \left[ \frac{\hat{p}_-^2}{2\mu} + \frac{1}{2} \mu \omega^2 \left( 1 - \frac{\gamma}{2} \right) \hat{x}_-^2 \right]. \quad (\text{A6})$$

Therefore the ground-state energy is given by

$$E = \frac{1}{2} \hbar \omega \left( \sqrt{1 + \frac{\gamma}{2}} + \sqrt{1 - \frac{\gamma}{2}} \right) \\ = \frac{1}{2} \hbar \omega \left( 2 - \frac{\gamma^2}{16} + O(\gamma^4) \right). \quad (\text{A7})$$

Using the definition  $\gamma = -\frac{4\alpha}{R^3}$  gives the  $R^6$  term corresponding for the oscillator pairs

$$\Delta E = -\frac{1}{32} \gamma^2 \hbar \omega = -\frac{1}{2} \frac{\alpha^2}{R^6} \hbar \omega. \quad (\text{A8})$$

### APPENDIX B: TRUNCATED FOCK SPACE

For our one-dimensional QDO system, we want to be sure that using a finite-dimensional Fock space for each QDO still gives a reasonable approximation to the ground-state energy of the true infinite-dimensional system. Figure 9 shows how the ground-state energy of the system deviates for different values of the local QDO dimension  $d$ . We can see that even for  $d = 4$ , the ground-state energy is approximated well until very large couplings  $\gamma$ . For reference,  $\gamma = 2$  corresponds to a Hamiltonian that is no longer positive-semidefinite, at this point the charged Drude particles dissociate from the nuclei

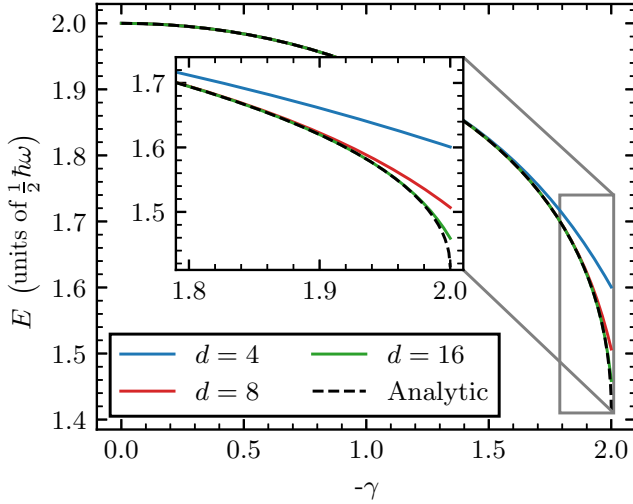


FIG. 9. Ground-state energy of one-dimensional QDO pair as a function of coupling constant  $\gamma$  for various values of  $d$ ; the dimensionality of the truncated Fock space. Analytic refers to exact energy given by (A7).

and the system breaks down. For physically relevant systems, we expect to require value of  $\gamma$  much lower than 2, and we can therefore limit ourselves to small  $d$ . For the real device experiments in the main text, the smallest separation  $R$  considered corresponds to the largest coupling value of  $\gamma = 1.55$ .

## APPENDIX C: DECOMPOSITION AND MEASUREMENT OF HAMILTONIAN

### 1. Scaling in the total number of Pauli operators

*Theorem 1.* For a  $d$ -level harmonic QDO, there are  $(\frac{1}{2}d \log_2 d)^2$  nonzero Pauli terms in the binary encoding of the  $\hat{x} \otimes \hat{x}$  interaction.

*Proof.* The following single-qubit operators can be represented in the Pauli basis

$$\begin{aligned} |0\rangle\langle 0| &= \frac{1}{2}(\mathbb{1} + Z), & |1\rangle\langle 0| &= \frac{1}{2}(X - iY), \\ |1\rangle\langle 1| &= \frac{1}{2}(\mathbb{1} - Z), & |0\rangle\langle 1| &= \frac{1}{2}(X + iY). \end{aligned} \quad (\text{C1})$$

We first consider a  $\hat{x} = \frac{1}{\sqrt{2}}(\hat{a}^\dagger + \hat{a})$  for a single QDO, where the ladder operators  $\hat{a}^\dagger$  and  $\hat{a}$  for the  $d$ -level Fock space are given by

$$\begin{aligned} \hat{a}^\dagger &= \sum_{n=0}^{d-1} \sqrt{n+1} |n+1\rangle\langle n|, \\ \hat{a} &= \sum_{n=0}^{d-1} \sqrt{n+1} |n\rangle\langle n+1|, \end{aligned} \quad (\text{C2})$$

which follow the usual bosonic commutation relations. We begin by considering a single  $|n+1\rangle\langle n|$  which given in terms of  $k$  single-qubit operators is

$$\begin{aligned} |n+1\rangle\langle n| &= \left( \bigotimes_{i=k+1}^{m-1} |\text{bin}(n)_i\rangle\langle \text{bin}(n)_i| \right) \\ &\times \otimes |1\rangle\langle 0| \otimes (|0\rangle\langle 1|)^{\otimes k}, \end{aligned} \quad (\text{C3})$$

where  $\text{bin}(n)_i \in \{0, 1\}$  is the binary digit at position  $i$  of the binary representation of  $n$  and  $k(n) = \max\{i \mid \text{bin}(n)_i = 0\}$  is the position of the rightmost 0. Using the definitions in (C1) this is equal to

$$\begin{aligned} |n+1\rangle\langle n| &= \left[ \bigotimes_{i=k+1}^{m-1} (\mathbb{1} + (-1)^{\text{bin}(n)_i} Z) \right] \\ &\times \otimes (X + iY) \otimes (X - iY)^{\otimes k}. \end{aligned} \quad (\text{C4})$$

Expanding (C4) will give  $2^m$  unique Pauli terms of the form  $\hat{P} = c\sigma_i^{\otimes m}$  where  $\sigma_i^{\otimes m} \in \{\mathbb{1}, X, Y, Z\}^m$  and  $c \in \{\pm 1, \pm i\}$ . Each  $\hat{P}$  will either be hermitian or anti-Hermitian, depending on whether  $c = \pm 1$  or  $c = \pm i$ .

For each term in  $\hat{x}$  like  $(|n+1\rangle\langle n| + |n\rangle\langle n+1|)$ , we note that  $|n+1\rangle\langle n| = (|n\rangle\langle n+1|)^\dagger$  and their sum is Hermitian. Therefore the Pauli terms within this summed term are simply two times the Hermitian terms in the expansion of (C4) (anti-Hermitian terms cancel). Of the  $2^m$  different  $\hat{P}$  within (C4), half will be Hermitian ( $c = \pm 1$ ) and half will be anti-Hermitian ( $c = \pm i$ ). Therefore, there are a total of  $\frac{1}{2}2^m$  unique Pauli operators in the compact representation of  $(|n+1\rangle\langle n| + |n\rangle\langle n+1|)$ .

Finally, in the summation

$$\hat{x} = \sum_{n=0}^{d-1} \sqrt{n+1} (|n+1\rangle\langle n| + |n\rangle\langle n+1|), \quad (\text{C5})$$

there are  $\log_2 d = m$  unique values of  $k(n)$ , each of which will contribute  $\frac{2^m}{2}$  unique Pauli operators to the binary encoding of  $\hat{x}$ . Thus there are a total of  $(\frac{1}{2}d \log_2 d)^2$  Pauli operators within  $\hat{x} \otimes \hat{x}$ , where we have used  $m = \log_2 d$ . ■

### 2. Scaling of Number of Circuits Required

*Lemma 2.* The  $(\frac{1}{2}d \log_2 d)^2$  Pauli operators for the  $\hat{x} \otimes \hat{x}$  interaction in the binary encoding can be measured using  $(d-1)^2$  circuits.

*Proof.* We define the cover of an  $M$  qubit Pauli operator  $\hat{P} \in \{\mathbb{1}, X, Y, Z\}^M$  as

$$\begin{aligned} \text{Cover}(\hat{P}) &= \{\hat{B} \in \{\mathbb{1}, X, Y, Z\}^{\otimes M} \mid \hat{B}^{(i)} = \hat{P}^{(i)} \text{ when } \hat{P}^{(i)} \neq \mathbb{1}\}, \end{aligned} \quad (\text{C6})$$

where  $\hat{P}^{(i)}$  is Pauli operator acting on the  $i$ th qubit. For a collection of Pauli operators  $\{\hat{P}_1, \hat{P}_2, \hat{P}_3, \dots\}$  that qubit-wise commute, there exists at least one basis  $\hat{B}$  which covers all of the operators:

$$\hat{B} \in \text{Cover}(\hat{P}_1) \cup \text{Cover}(\hat{P}_2) \cup \text{Cover}(\hat{P}_3) \cup \dots \quad (\text{C7})$$

The expectation values of every operator in  $\{\hat{P}_1, \hat{P}_2, \hat{P}_3, \dots\}$  can be measured using a single circuit in which qubit  $i$  is measured in the  $\hat{B}^{(i)}$  basis, where the  $\pm 1$  measurement value is used if  $\hat{P}^{(i)} \neq \mathbb{1}$ , and discarded otherwise. Here we called bases independent if they do not qubit-wise commute.

We show that there are  $(d-1)$  independent covers for terms in  $\hat{x}$  and therefore  $(d-1)^2$  covers needed to measure  $\hat{x} \otimes \hat{x}$ .

We begin by considering the expansion of  $|n+1\rangle\langle n|$  given in (C4). Expanding out the  $\bigotimes_{i=k+1}^{m-1} [1 + (-1)^{\text{bin}(n)_i} Z]$  term on the first  $(m-k-1)$  qubits will give tensor products of only  $1$  and  $Z$ . The covers for this  $2^{(m-k-1)}$  operator subspace will always contain the basis  $\bigotimes_{i=k+1}^{m-1} Z^{(i)}$ , which can be used to measure them all. Then expanding the  $(X+iY) \otimes (X-iY)^{\otimes k}$  term on the remaining  $(m-k+1)$  qubits gives  $2^{(k+1)}$  terms that do not qubit-wise commute and so each correspond to a unique basis.

Thus, there are  $2^{(k+1)}$  independent bases for the terms in  $|n+1\rangle\langle n|$ . Using the argument that  $|n+1\rangle\langle n| = |n\rangle\langle n+1|^\dagger$  as in Theorem 4, only the Hermitian terms remain when summing the two. This halves the number of unique covers and so  $2^k$  bases are required for  $|n+1\rangle\langle n| + |n\rangle\langle n+1|$ .

Again, for the sum of these terms given in (C5), there are  $m$  unique values for  $k(n) \in \{0, 1, \dots, m-1\}$  when written in the form of (C1). For each unique value of  $k(n)$ , the value of  $\text{bin}(n)$  does not affect the coverings required [since the first  $(m-k-1)$  qubits are always covered with  $Z^{\otimes(m-k-1)}$ ], so each unique  $k(n)$  contributes a unique set of  $2^k$  coverings to the summation. Therefore, the total number of coverings needed for the Pauli terms in  $\hat{x}$  are  $\sum_{k=0}^{m-1} 2^k = 2^m - 1$ , where we have used a standard result for the summation.

Finally, the number of unique coverings for the  $\hat{x} \otimes \hat{x}$  interaction is then  $(2^m - 1)^2 = (d - 1)^2$ . ■

**Definition 1** (Hypergraphs [101]). A hypergraph is a generalization of a graph in which edges can connect arbitrary number (i.e., not just a pair) of vertices. Formally, a hypergraph  $G = (V, E)$  consists of  $V$  a finite set of vertices and  $E$  a set of nonempty subsets of  $V$  called hyperedges.

**Definition 2** (Complete hypergraphs [101]). For  $0 \leq r \leq N$ , a complete hypergraph  $K_N^r = (V, E)$  is a hypergraph with vertices  $V$  and set of hyperedges  $E$  such that  $|V| = N$  and  $E$  contains all  $r$ -subsets of  $V$ .

**Lemma 3** (Baranyai's Theorem [102]). For a complete hypergraph  $K_N^r$  where  $2 \leq r < N$  and  $r$  divides  $N$ , the  $N$  vertices can be partitioned into subsets (called 1-factorization) of  $r$  vertices in  $\binom{N-1}{r-1}$  different ways, such that each  $r$ -element subset (hyperedge) appears in exactly one partition.

*Proof.* See [[102], Chapter 38]. ■

For example, the complete hypergraph  $K_4^2$  contains four vertices, each connected by two-element hyperedges (i.e., traditional edges). Baranyai's theorem states that the vertices can be partitioned into  $\binom{4-1}{2-1} = 3$  partitions.  $K_4^2$  and its corresponding partitions are shown in Fig. 10.

**Theorem 4.** The number of circuits required for measurement of the Hamiltonian of the form

$$\hat{H} = \sum_{i=1}^N (\hat{x}_i^2 + \hat{p}_i^2) + \sum_{j>i} \gamma_{i,j} \hat{x}_i \hat{x}_j, \quad (\text{C8})$$

for  $N \geq 2$   $d$ -level QDOs in the binary encoding is upper bounded by

$$\begin{aligned} (d-1)^2(N-1) + 1; \quad N \text{ even} \\ (d-1)^2N + 1; \quad N \text{ odd,} \end{aligned} \quad (\text{C9})$$

which is  $O(d^2N)$ .

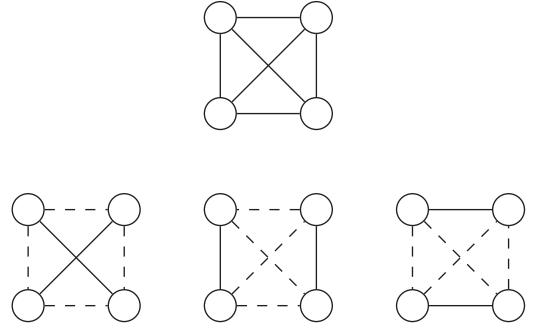


FIG. 10. Fully connected graph  $K_4^2$  along with the three corresponding partitions of edges.

*Proof.* All  $(\hat{x}^2 + \hat{p}^2) = \sum_{n=0}^{d-1} n |n\rangle\langle n|$  terms within the Hamiltonian can be decomposed into Pauli operators given in (4) of the main text. All of these terms commute within each QDO register as well between QDO registers and can thus be measured using the single covering  $Z^{\otimes Nm}$ .

For the pairwise coupling terms, from Theorem 2, a single  $\hat{x} \otimes \hat{x}$  term between two QDOs can be measured with  $(d-1)^2$  circuits. The bases required to measure pairwise coupling terms between disjoint pairs of QDOs are guaranteed to commute since they act on different qubit registers. Therefore, by measuring each of the bases concurrently for collections of mutually disjoint interactions, the number of measurement bases can be reduced in a predictable way.

We represent interaction terms as a graph in which each vertex corresponds to one of the  $N$  QDOs and edges between vertices correspond represent a nonzero pairwise coupling. For all-to-all interactions this will be a fully connected graph. For even  $N$  the number of partitions needed to separate interaction edges into disjoint sets is equal by Baranyai's theorem (Lemma 3) for the graph  $K_N^2$ . For odd  $N$  the edges cannot be partitioned exactly—there will always be one unpaired vertex when collecting disjoint edges. Considering a dummy vertex where an edge between that and a QDO vertex gives no contribution to the Hamiltonian, the number of partitions is given by Baranyai's theorem for  $K_{N+1}^2$ .

Using these collections, the number of bases that will need to be measured of the pairwise interaction terms is

$$\begin{aligned} (d-1)^2 \binom{N-1}{1} &= (d-1)^2(N-1); \quad N \text{ even,} \\ (d-1)^2 \binom{N}{1} &= (d-1)^2N; \quad N \text{ odd.} \end{aligned} \quad (\text{C10})$$

Combining this with the single basis needed to measure the noninteraction term gives an upper bound for the number of bases that must be measured and thus the result. ■

This upper bound is plotted as a function of  $N$  and  $d$  in Fig. 11 as well as the number of circuits required when using a heuristic sorting algorithm, namely largest-degree first coloring [60]. We see that for  $N = 1$  both curves agree as there are no interaction terms and that the heuristic sorted saturates the upper bound for  $N = 2$  QDOs. For  $N > 2$  we see that the sorting significantly reduces the number of circuits below the upper bound due to additional qubit wise commutation between nondisjoint interaction terms.

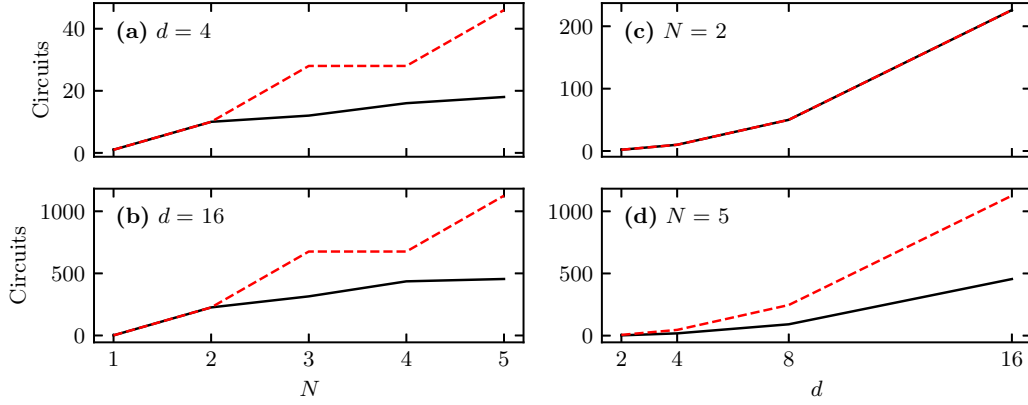


FIG. 11. Number of circuits required to measure Hamiltonian as (a) and (b) a function of  $N$  for two fixed values of  $d$  as well as (c) and (d) a function of  $d$  for two fixed values of  $N$ . Red dashed line: upper bound scaling given by (C9), black solid line: using largest-degree first coloring for grouping Paulis. In (c), the two lines coincide.

*Corollary 4.1.* The number of circuits required for measurement of the Hamiltonian for  $N \geq r$   $d$ -level QDOs with  $r$ -order coupling in the binary encoding is  $O(d^2 N^{r-1})$ .

*Proof.* The proof is the same as that for Theorem 4 except using  $r > 2$  for the  $\hat{x}^{\otimes r}$  interaction in Lemma 2 and the corresponding hypergraph  $K_N^r$  in Lemma 3. This gives an upper bound on the total number of groups of

$$(d-1)^r \binom{N-1}{r-1} + 1, \quad (\text{C11})$$

assuming  $r$  divides  $N$ , which is  $O(d^r N^{(r-1)})$ . If  $r$  does not divide  $N$ , the  $O(d^r N^{(r-1)})$  scaling will remain as dummy vertices can be used to give the number of vertices that is divisible for  $r$ . ■

### 3. Number of circuits for three-dimensional QDOs

*Theorem 5.* The number of circuits required for measurement of the Hamiltonian of the form

$$H = \sum_{i=1}^N \left( \frac{\hat{\mathbf{P}}_i^2}{2\mu} + 2\mu\omega\hat{\mathbf{x}}_i^2 \right) + q^2 \sum_{j>i} \frac{3(\hat{\mathbf{x}}_i \cdot \mathbf{R}_{i,j})(\hat{\mathbf{x}}_j \cdot \mathbf{R}_{i,j}) - \hat{\mathbf{x}}_i \cdot \hat{\mathbf{x}}_j R_{i,j}^2}{R_{i,j}^5}, \quad (\text{C12})$$

for  $N$  three-dimensional QDOs with  $d$  levels per spatial dimension in the binary encoding

$$(d-1)^2(N-1) + 1; \quad N \text{ even}, \quad (d-1)^2N + 1; \quad N \text{ odd}. \quad (\text{C13})$$

*Proof.* The uncoupled terms within (C12) will correspond to operators  $\sum_{n=0}^{2^m-1} n |\underline{n}\rangle \langle \underline{n}|$  for each of the three spatial dimensions for each QDO. As before, these require just  $Z$  measurements on each of the qubits and can thus be measured using a single circuit.

Written in terms of operators for each one-dimensional component, the interaction part of (C12) will contain all nine terms of the form  $\{\hat{x}_i, \hat{y}_i, \hat{z}_i\} \otimes \{\hat{x}_j, \hat{y}_j, \hat{z}_j\}$ , where across each  $3m$  qubit register for a single QDO, the operators making up  $\hat{\mathbf{x}}_i$  are given by  $\hat{x}_i = \hat{x} \otimes \mathbb{1}^{\otimes m} \otimes \mathbb{1}^{\otimes m}$ ,  $\hat{y}_i = \mathbb{1}^{\otimes m} \otimes \hat{x} \otimes \mathbb{1}^{\otimes m}$ , and  $\hat{z}_i = \mathbb{1}^{\otimes m} \otimes \mathbb{1}^{\otimes m} \otimes \hat{x}$ , following a similar pattern for  $\hat{\mathbf{x}}_j$ ,

$\hat{\mathbf{p}}_i$ , and  $\hat{\mathbf{p}}_j$ . Since every single term of the form  $\hat{x}\hat{x}$ ,  $\hat{x}\hat{y}$ ,  $\hat{y}\hat{z}$ , etc., in the coupling interactions will consist of the same decomposition of Pauli operators either acting on the same or different QDO subregisters. These all qubit-wise commute and each term each of the  $(\frac{d}{2} \log_2 d)^2$  Pauli terms for all oscillator-oscillator interaction terms can be measured together. Thus for each pair of coupled three-dimensional QDOs, only  $(\frac{d}{2} \log_2 d)^2$  circuits are needed; the same number required for a pairwise interaction between one-dimensional QDOs.

Finally, using Lemma 3 and following the same logic as Theorem 4, an upper bound for the number of circuits needed to measure in (C12) is  $(N-1)(\frac{d}{2} \log_2 d)^2 + 1$  or  $N(\frac{d}{2} \log_2 d)^2 + 1$ , for even and odd  $N$ , respectively. ■

### APPENDIX D: SHOT REQUIREMENTS

In this section, we derive an upper bound for an estimate of the number of shots required to achieve a given statistical error on the measurement of  $\langle \hat{H} \rangle$  for a system of 1D coupled QDOs using the grouping of measurements described in Theorem 4.

We write a Hamiltonian in the form

$$\hat{H} = \sum_{g \in \mathcal{G}} \sum_{i \in g} a_i \hat{P}_i, \quad (\text{D1})$$

where  $\mathcal{G} = \{g \mid [\hat{P}_i, \hat{P}_j] = 0, \forall i, j \in g\}$  is the set of all collections for the Pauli operators  $\hat{P}_i$  in which all operators in a collection qubit-wise commute. A standard method [103] using Lagrange multipliers to optimally distribute measurements between circuits gives a total number experimental shots of

$$S = \frac{1}{\epsilon^2} \left[ \sum_{g \in \mathcal{G}} \sqrt{\sum_{i \in g} |a_i|^2 \text{Var} \hat{P}_i} \right]^2 \quad (\text{D2})$$

to achieve an expected measurement error of  $\epsilon$ , assuming that  $\text{Cov}[\hat{P}_i, \hat{P}_j] = 0$  if  $i \neq j$ . The variance of a Pauli operator is given by  $\text{Var} \hat{P}_i = 1 - \langle \hat{P}_i \rangle$ . In the absence of knowledge about the state that the expectation value is taken with respect to, we take the expectation of this variance with respect to a spherical



measure (see [[64], Chapter 7] and following a similar method to Ref. [59]):

$$\mathbb{E}[\text{Var}\hat{P}_i] = \begin{cases} 0; & \hat{P}_i = \mathbb{1} \\ 1 - \frac{1}{2^{n_i+1}}; & \text{otherwise} \end{cases}, \quad (\text{D3})$$

where  $n_i$  is the number of nonidentity Pauli operators in  $\hat{P}_i$ . Additionally, in the uniform spherical measure,  $\mathbb{E}[\text{Cov}(\hat{P}_i, \hat{P}_j)] = 0$  if  $i \neq j$ . Thus the expected number of shots  $S$  to achieve error  $\epsilon$  is upper bounded by

$$\mathbb{E}[S] = \frac{1}{\epsilon^2} \left[ \sum_{g \in \mathcal{G}} \sqrt{\sum_{i \in g} |a_i|^2 \left(1 - \frac{1}{2^{n_i+1}}\right)} \right]^2. \quad (\text{D4})$$

$$S \leq \frac{1}{\epsilon^2} \left( \frac{\sqrt{2}}{3} \sqrt{N} \sqrt{d^2 - 1} + \left( \sum_{F \in \text{Part}(G)} \sqrt{\sum_{(i,j) \in F} \gamma_{i,j}^2} \right) \left\{ 2d - (4 + 2 \log_2 d)d^2 + \left[ 2 + 2 \log_2 d + \frac{1}{2}(\log_2 d)^2 \right] d^3 \right\} \right)^2, \quad (\text{D6})$$

when taken over the spherically symmetric distribution of states, where  $\text{Part}(G)$  is the set of coverings for a given 1-factorization of the interaction graph  $G$ . Assuming  $G$  is fully connected and that values of  $\gamma_{i,j}$  are independent of both  $N$  and  $d$  and randomly distributed between different  $F \in \text{Part}(G)$ , then  $S \in O\left(\frac{\gamma^2 N^3 (\log_2 d)^4 d^6}{\epsilon^2}\right)$ , where  $\gamma$  is a typical scale for  $\gamma_{i,j}$ .

*Proof.* We consider the groupings described in Theorem 4. Since interacting and noninteracting parts of the Hamiltonian have different forms and groupings, we consider each one's contribution to the total shots separately:

$$\epsilon^2 S = (\epsilon \tilde{S}_{\text{non-int}} + \epsilon \tilde{S}_{\text{int}})^2. \quad (\text{D7})$$

Starting with the uncoupled part of the Hamiltonian:

$$\sum_{i=1}^N (\hat{x}_i^2 + \hat{p}_i^2) = \sum_{i=1}^N 2^m \mathbb{1}^{\otimes mN} - \sum_{j=0}^{m-1} 2^j Z^{(im+j)}, \quad (\text{D8})$$

where  $Z^{(im+j)}$  is the Z operator acting on the  $j$ th qubit of QDO  $i$ , all terms form a single collection, which gives

$$\epsilon \tilde{S}_{\text{non-int}} = \sqrt{N \sum_{j=0}^{m-1} 2^{2j} \text{Var}Z} = \sqrt{\frac{2}{9} N (2^{2m} - 1)}, \quad (\text{D9})$$

where we have used that  $\mathbb{E}[\text{Var}Z] = \frac{2}{3}$  from (D3).

For the interacting terms, the collection to which a single Pauli operators given by (C4) (i.e., which  $g \in \mathcal{G}$  it is sorted into) is determined by three things:

*Theorem 6* (Expected shots over uniform measure when grouping measurements). Given a Hamiltonian

$$\hat{H} = \sum_{i=1}^N (\hat{x}_i^2 + \hat{p}_i^2) + \sum_{j>i} \gamma_{i,j} \hat{x}_i \hat{x}_j, \quad (\text{D5})$$

for  $N \geq 2$   $d$ -level QDOs in the binary encoding and measuring commuting operators according to the procedure described in Theorem 4, the expected number of shots  $S$  required to achieve a constant measurement error  $\epsilon$  is upper bounded by

(1) Which of the  $(N-1)$  (even  $N$ ) or  $N$  (odd  $N$ ) 1-factorizations of the interaction graph the corresponding interaction term belongs to.

(2) The value of  $k(n)$ , the position of the rightmost zero in the binary representation of integer  $n$  labeling the Fock state, for each of the two the QDOs involved in an interaction term.

(3) For a given pair of  $k(n)$ , which of the possible  $2^{k+1}$  of  $X$  and  $Y$  operators is applied to the right most  $(k+1)$  qubits. As before, the anti-Hermitian terms in (C4) are not needed in the sum  $|n+1\rangle \langle n| + |n\rangle \langle n+1|$ .

Separating the terms within  $\hat{H}$  according to these criteria (their contributions to the equation labeled with curly braces), the expected number of shots required is

$$\epsilon \tilde{S}_{\text{int}} = \frac{1}{2} \underbrace{\sum_{k_1, k_2=0}^{m-1}}_1 \underbrace{\sum_{F \in \text{Part}(G)}}_2 \underbrace{\sum_{\substack{\hat{P}_1^{(k_1+1)} \in \{X, Y\}^{\otimes(k_1+1)} \\ \hat{P}_2^{(k_2+1)} \in \{X, Y\}^{\otimes(k_2+1)}}}}_3 \Gamma(F, k_1, k_2), \quad (\text{D10})$$

where  $\text{Part}(G)$  is the set of the interaction graph  $G$ . From Lemma 3, if  $G$  is fully connected then  $|G| = (N-1)$  or  $N$  (for  $N$  even or odd). We have defined  $\mathcal{N}_{k_1} = \{n \mid k(n) = k_1, n \in \mathbb{Z}, 0 \leq n < 2^m\}$  as the set of integers satisfying a given  $k(n)$  (similarly for  $\mathcal{N}_{k_2}$ ) and  $\hat{P}_1^{(k_1+1)}$  is a  $(k_1+1)$  qubit Pauli operator (similarly for  $\hat{P}_2^{(k_2+1)}$ ). The 1/2 prefactor results from the cancellation of anti-Hermitian terms which make up half of the operators. The value of  $\Gamma$  is given by

$$\Gamma(F, k_1, k_2) = \left[ \sum_{(i,j) \in F} \sum_{\substack{n_1 \in \mathcal{N}_{k_1} \\ n_2 \in \mathcal{N}_{k_2}}} \sum_{\substack{\hat{P}_1^{(k_1+1:m)} \in \{\mathbb{1}, Z\}^{\otimes(m-k_1-1)} \\ \hat{P}_2^{(k_2+1:m)} \in \{\mathbb{1}, Z\}^{\otimes(m-k_2-1)}}} (2\gamma_{i,j} \sqrt{n_1+1} \sqrt{n_2+1})^2 \mathbb{E}[\text{Var}\hat{P}_1 \otimes \hat{P}_2] \right]^{\frac{1}{2}}, \quad (\text{D11})$$

where,  $\hat{P}_1^{(k_1+1:m)}$  is an  $(m - (k_1 + 1))$  qubit Pauli operator and  $\hat{P}_1 = \hat{P}_1^{(k_1+1)} \otimes \hat{P}_1^{(k_1+1:m)}$  (similarly for  $\hat{P}_2^{(k_2+1:m)}$ ). This is the equivalent of the square root term in (D4); the contribution to the variance from a single collection of commuting operators. In

the above,  $\langle i, j \rangle$  correspond to edges within 1-factorization  $F$ ,  $n_1$  and  $n_2$  correspond to Fock terms satisfying the requirement of  $k(n)$ ; the position of the rightmost X or Y operator. Using the expectation on the spherical measure from (D3) gives

$$\mathbb{E}[\text{Var}\hat{P}_1 \otimes \hat{P}_2] = \left(1 - \frac{1}{2^{k_1+k_2+\alpha+\beta} + 1}\right), \quad (\text{D12})$$

where integers  $0 < \alpha_1 < m - k_1$  and  $0 < \alpha_2 < m - k_2$  count the number of Z operators in  $\hat{P}_1^{(k_1+1:m)}$  and  $\hat{P}_2^{(k_2+1:m)}$  respectively.

By recognizing that when using the spherical measure, many of the expected variances are equivalent and only depend on the total number of X, Y, or Z operators, the expected number of shots simplifies to

$$\epsilon_{\tilde{S}_{\text{int}}} = \frac{1}{2} \sum_{k_1, k_2=0}^{m-1} \sum_{F \in \text{Part}(K_N^2)} 2^{k_1+1} 2^{k_2+1} \Gamma(F, k_1, k_2) \quad (\text{D13})$$

with

$$\Gamma(F, k_1, k_2) = \left[ \sum_{(i,j) \in F} \sum_{\substack{n_1 \in \mathcal{N}_{k_1} \\ n_2 \in \mathcal{N}_{k_2}}} \sum_{\alpha=0}^{m-k_1-1} \sum_{\beta=0}^{m-k_2-1} \binom{m-k_1-1}{\alpha} \binom{m-k_2-1}{\beta} (2\gamma_{i,j} \sqrt{n_1+1} \sqrt{n_2+1})^2 \left(1 - \frac{1}{2^{k_1+k_2+\alpha+\beta} + 1}\right) \right]^{\frac{1}{2}}. \quad (\text{D14})$$

Now we aim to find a simple upper bound for  $S$ . We start by grouping terms:

$$\begin{aligned} \epsilon_{\tilde{S}_{\text{int}}} &= \left( \sum_{F \in \text{Part}(G)} \sqrt{\sum_{(i,j) \in F} \gamma_{i,j}^2} \right) \sum_{k_1, k_2=1}^m 2^{k_1+1} 2^{k_2+1} \sqrt{\sum_{\substack{n_1 \in \mathcal{N}_{k_1} \\ n_2 \in \mathcal{N}_{k_2}}} (n_1+1)(n_2+1)} \\ &\quad \times \sqrt{\sum_{\alpha=0}^{m-k_1-1} \sum_{\beta=0}^{m-k_2-1} \binom{m-k_1-1}{\alpha} \binom{m-k_2-1}{\beta} \left(1 - \frac{1}{2^{k_1+k_2+\alpha+\beta} + 1}\right)}. \end{aligned} \quad (\text{D15})$$

We note that

$$\sum_{\substack{n_1 \in \mathcal{N}_{k_1} \\ n_2 \in \mathcal{N}_{k_2}}} (n_1+1)(n_2+1) = \left(2^{k_1} \sum_{\gamma=0}^{2^{m-k_1-1}} (2\gamma+1)\right) \left(2^{k_2} \sum_{\delta=0}^{2^{m-k_2-1}} (2\delta+1)\right) = 2^{-k_1-2} (2^{k_1+1} + 2^m)^2 \times 2^{-k_2-2} (2^{k_2+1} + 2^m)^2 \quad (\text{D16})$$

and

$$\begin{aligned} \sum_{\alpha=0}^{m-k_1-1} \sum_{\beta=0}^{m-k_2-1} \binom{m-k_1-1}{\alpha} \binom{m-k_2-1}{\beta} \left(1 - \frac{1}{2^{k_1+k_2+\alpha+\beta} + 1}\right) &\leq \sum_{\alpha=0}^{m-k_1-1} \binom{m-k_1-1}{\alpha} \sum_{\beta=0}^{m-k_2-1} \binom{m-k_2-1}{\beta} \\ &= 2^{m-k_1-1} \times 2^{m-k_2-1}, \end{aligned} \quad (\text{D17})$$

where we have used a standard result for the summation of binomial coefficients. Substituting (D16) and (D17) into (D15) gives

$$\begin{aligned} \epsilon_{\tilde{S}_{\text{int}}} &\leq \left( \sum_{F \in \text{Part}(G)} \sqrt{\sum_{(i,j) \in F} \gamma_{i,j}^2} \right) \sum_{k_1, k_2=0}^{m-1} 2^{m-1} (4 \times 2^{k_1+k_2} + 2^{m+1} \times 2^{k_1} + 2^{m+1} \times 2^{k_2} + 2^{2m}) \\ &= \left( \sum_{F \in (G)} \sqrt{\sum_{(i,j) \in F} \gamma_{i,j}^2} \right) (2d - [4 + 2 \log_2 d]d^2 + [2 + 2 \log d + \frac{1}{2}(\log_2 d)^2]d^3), \end{aligned} \quad (\text{D18})$$

using standard results for the summations and substitution  $d = 2^m$ . Combining this with (D9) gives the result.  $\blacksquare$

*Theorem 7* (Expected shots over uniform measure when not grouping measurements.). Given a Hamiltonian

$$\hat{H} = \sum_{i=1}^N (\hat{x}_i^2 + \hat{p}_i^2) + \sum_{j>i} \gamma_{i,j} \hat{x}_i \hat{x}_j, \quad (\text{D19})$$

for  $N \geq 2$   $d$ -level QDOs in the binary encoding and measuring each corresponding Pauli operator separately, the expected number of shots required to achieve a constant measurement error  $\epsilon$  is given by the upper bound

$$S \leq \frac{1}{\epsilon^2} \left\{ N \frac{\sqrt{6}}{3} (d-1) + \left( \sum_{(i,j) \in G} |\gamma_{i,j}| \right) \left[ (3 + 2\sqrt{2})d^2 - (2 + \sqrt{2})d^{\frac{5}{2}} - \left( \frac{7}{2} + 3\sqrt{2} \right) d^3 + (1 + \sqrt{2})d^{\frac{7}{2}} + \left( \frac{3}{2} + \sqrt{2} \right) d^4 \right] \right\}^2, \quad (\text{D20})$$

when taken over the spherically symmetric distribution of states, where  $G$  is the interaction graph. Assuming  $G$  is fully connected and, that values of  $\gamma_{i,j}$  are independent of both  $N$  and  $d$ , then  $S \in O\left(\frac{\gamma^2 N^4 d^8}{\epsilon^2}\right)$  where  $\gamma$  is a typical scale for  $\gamma_{i,j}$ .

*Proof.* As in Theorem 6, we consider contributions from noninteraction and interaction terms separately. Starting with the uncoupled part of the Hamiltonian

$$\epsilon \tilde{S}_{\text{non-int}} = N \sum_{j=0}^{m-1} \sqrt{2^{2j} \text{Var} Z} = N \frac{\sqrt{6}}{3} (2^m - 1). \quad (\text{D21})$$

For the interaction part of the Hamiltonian

$$\begin{aligned} \epsilon \tilde{S}_{\text{int}} &= \frac{1}{2} \sum_{(i,j) \in G} \sum_{k_1, k_2=0}^{m-1} \sum_{\substack{\hat{P}_1 \in \{X, Y\}^{\otimes k_1+1} \otimes \{\mathbb{1}, Z\}^{\otimes m-k_1-1} \\ \hat{P}_2 \in \{X, Y\}^{\otimes k_2+1} \otimes \{\mathbb{1}, Z\}^{\otimes m-k_2-1}}} \left[ \sum_{\substack{n_1 \in \mathcal{N}_{k_1} \\ n_2 \in \mathcal{N}_{k_2}}} 4\gamma_{i,j}^2 (n_1+1)(n_2+1) \mathbb{E}(\text{Var} \hat{P}_1 \otimes \hat{P}_2) \right]^{\frac{1}{2}} \\ &= \sum_{(i,j) \in G} |\gamma_{i,j}| \sum_{k_1, k_2=0}^{m-1} 2^{k_1+1} 2^{k_2+1} \sqrt{\sum_{\substack{n_1 \in \mathcal{N}_{k_1} \\ n_2 \in \mathcal{N}_{k_2}}} (n_1+1)(n_2+1)} \\ &\quad \times \sum_{\alpha=0}^{m-k_1-1} \sum_{\beta=0}^{m-k_2-1} \binom{m-k_1-1}{\alpha} \binom{m-k_2-1}{\beta} \sqrt{\left(1 - \frac{1}{2^{k_1+k_2+\alpha+\beta} + 1}\right)}. \end{aligned} \quad (\text{D22})$$

Using (D17) and (D16) again and writing  $d = 2^m$ , we get the inequality

$$\begin{aligned} \epsilon \tilde{S}_{\text{int}} &\leq \sum_{(i,j) \in G} |\gamma_{i,j}| \sum_{k_1, k_2=0}^{m-1} (2^{k_1+1} \times 2^{k_2+1}) (2^{m-k_1-1} \times 2^{m-k_2-1}) \left[ 2^{-\frac{k_1}{2}-1} (2^{k_1+1} + 2^m) \times 2^{-\frac{k_2}{2}-1} (2^{k_2+1} + 2^m) \right] \\ &= \left( \sum_{(i,j) \in G} |\gamma_{i,j}| \right) \left[ (3 + 2\sqrt{2})d^2 - (2 + \sqrt{2})d^{\frac{5}{2}} - \left( \frac{7}{2} + 3\sqrt{2} \right) d^3 + (1 + \sqrt{2})d^{\frac{7}{2}} + \left( \frac{3}{2} + \sqrt{2} \right) d^4 \right]. \end{aligned} \quad (\text{D23})$$

Combining this with (D21) gives the result.  $\blacksquare$

The preceding results assume no knowledge about the states that are being measured and use an expectation over all of Hilbert space respect to the spherically symmetric measure. Here we make a second estimate for the number of shots required using the pure state  $|\psi_0\rangle = |0\rangle^{\otimes mN}$ .  $|\psi_0\rangle$  corresponds to the product state of all QDOs in their lowest Fock state and is assumed to be reasonably close to the states that we seek using the VQE procedure.

When not using the spherical measure, contributions to the measurement error from the covariance of terms within the same collection is nonzero may be nonzero. In this case, the expected number of shots is given by

$$S = \frac{1}{\epsilon^2} \left[ \sum_{g \in \mathcal{G}} \sqrt{\sum_{i,j \in g} a_i a_j \text{Cov}[\hat{P}_i, \hat{P}_j]} \right]^2, \quad (\text{D24})$$

where  $\text{Cov}[\hat{P}_i, \hat{P}_j] = \langle \hat{P}_i \hat{P}_j \rangle_{|\psi_0\rangle} - \langle \hat{P}_i \rangle_{|\psi_0\rangle} \langle \hat{P}_j \rangle_{|\psi_0\rangle}$  with subscripts indicating the state that expectation values are taken with respect to.

*Theorem 8* (Expected shots for uncoupled state when grouping measurements.). Given a Hamiltonian

$$\hat{H} = \sum_{i=1}^N (\hat{x}_i^2 + \hat{p}_i^2) + \sum_{j>i} \gamma_{i,j} \hat{x}_i \hat{x}_j, \quad (\text{D25})$$

for  $N \geq 2$   $d$ -level QDOs in the binary encoding and measuring commuting operators according to the procedure described in Theorem 4, the expected number of shots  $S$  required to achieve a constant measurement error  $\epsilon$  for state  $|0\rangle^{\otimes mN}$  is upper bounded by

$$S = \frac{1}{\epsilon^2} \frac{1}{16} \left( \sum_{F \in \text{Part}(G)} \sqrt{\sum_{(i,j) \in F} \gamma_{i,j}^2} \right)^2 d^4, \quad (\text{D26})$$

where  $\text{Part}(G)$  is the set of coverings for a given 1-factorization of the interaction graph  $G$ . Assuming  $G$  is fully connected and that values of  $\gamma_{i,j}$  are independent of both  $N$  and  $d$  and randomly distributed between different  $F \in \text{Part}(G)$ , then  $S \in O(\frac{\gamma^2 N^3 d^4}{\epsilon^2})$ , where  $\gamma$  is a typical scale for  $\gamma_{i,j}$ .

*Proof.* Again we consider noninteracting and interacting contributions separately. For the noninteracting contribution, we note that for  $|\psi\rangle = |0\rangle^{\otimes mN}$ ,  $\langle Z^{(i)}\rangle = \langle Z^{(j)}\rangle = \langle Z^{(i)}Z^{(j)}\rangle = 1 \forall i, j$ . Therefore,

$$\epsilon \tilde{S}_{\text{non-int}} = \sqrt{N \sum_{i,j=0}^{m-1} 2^i 2^j \text{Cov}[Z^{(i)}, Z^{(j)}]} = 0. \quad (\text{D27})$$

The contribution from the coupling terms is

$$\epsilon \tilde{S}_{\text{int}} = \frac{1}{2} \sum_{k_1, k_2=0}^{m-1} \sum_{F \in \text{Part}(K_N^2)} \sum_{\substack{\hat{P}_1^{(k_1)} \in \{X, Y\}^{\otimes(k_1+1)} \\ \hat{P}_2^{(k_2)} \in \{X, Y\}^{\otimes(k_2+1)}}} \Gamma(F, k_1, k_2), \quad (\text{D28})$$

where  $\text{Part}(K_N^2)$  is the set of the  $(N-1)$  or  $N$  (for  $N$  even or odd) partitions of the interaction graph  $K_N^2$ ,  $\mathcal{N}_{k_1} = \{n \mid k(n) = k_1, n \in \mathbb{Z}, 0 \leq n < 2^m\}$  is the set of integers satisfying a given  $k(n)$  (similarly for  $\mathcal{N}_{k_2}$ ) and  $\hat{P}_1^{(k_1+1)}$  is a  $(k_1+1)$  qubit Pauli operator (similarly for  $\hat{P}_2^{(k_2+1)}$ ).  $\Gamma$  is given by

$$\begin{aligned} & \Gamma(G, k_1, k_2) \\ &= \left[ \sum_{(i,j) \in F} \sum_{(i',j') \in F} \sum_{\substack{n_1 \in \mathcal{N}_{k_1} \\ n_2 \in \mathcal{N}_{k_2}}} \sum_{\substack{n'_1 \in \mathcal{N}_{k_1} \\ n'_2 \in \mathcal{N}_{k_2}}} \right. \\ & \quad \times 2^2 \gamma_{i,j} \gamma_{i',j'} \sqrt{n_1+1} \sqrt{n_2+1} \sqrt{n'_1+1} \sqrt{n'_2+1} \\ & \quad \left. \times \text{Cov}[\hat{A}_{(i,j)}(k_1, k_2, n_1, n_2), \hat{A}_{(i',j')}(k_1, k_2, n'_1, n'_2)] \right]^{\frac{1}{2}}, \quad (\text{D29}) \end{aligned}$$

where

$$\begin{aligned} & \hat{A}_{(i,j)}(k_1, k_2, n_1, n_2) \\ &= \left( \hat{P}_1^{(k_1)} \otimes \bigotimes_{\alpha=k_1+1}^{m-1} [\mathbb{1} + (-1)^{\text{bin}(n_1)_\alpha} Z] \right)_{\text{QDO } i} \\ & \quad \times \left( \hat{P}_2^{(k_2)} \otimes \bigotimes_{\beta=k_2+1}^{m-1} [\mathbb{1} + (-1)^{\text{bin}(n_2)_\beta} Z] \right)_{\text{QDO } j} \quad (\text{D30}) \end{aligned}$$

acts on the subregisters corresponding to QDOs  $i$  and  $j$ .

We note that  $\hat{P}_1^{(k_1)} \in \{\pm 1\} \times \{X, Y\}^{\otimes(k_1+1)}$  (recalling only Hermitian terms remain in  $\hat{H}$ ) and so  $\langle \hat{P}_1^{(k_1)} \rangle_{|\psi_0\rangle} = 0 \forall k_1$  and similarly for  $\hat{P}_2^{(k_2)}$ , from which it follows

$$\begin{aligned} & \langle i, j \rangle \neq \langle i', j' \rangle \\ & \Rightarrow \text{Cov}[\hat{A}_{(i,j)}(k_1, k_2, n_1, n_2), \hat{A}_{(i',j')}(k_1, k_2, n'_1, n'_2)] = 0 \quad (\text{D31}) \end{aligned}$$

and

$$\begin{aligned} & \text{Cov}[\hat{A}_{(i,j)}(k_1, k_2, n_1, n_2), \hat{A}_{(i,j)}(k_1, k_2, n'_1, n'_2)] \\ &= \langle \hat{A}_{(i,j)}(k_1, k_2, n_1, n_2) \hat{A}_{(i,j)}(k_1, k_2, n'_1, n'_2) \rangle_{|\psi_0\rangle} \\ &= \langle \hat{P}_1^{(k_1)} \hat{P}_1^{(k_1)} \rangle_{|\psi_0\rangle} \langle \hat{P}_2^{(k_2)} \hat{P}_2^{(k_2)} \rangle_{|\psi_0\rangle} \\ & \quad \times \left\langle \bigotimes_{\alpha=k_1+1}^{m-1} (\mathbb{1} + (-1)^{\text{bin}(n_1)_\alpha} Z) (\mathbb{1} + (-1)^{\text{bin}(n'_1)_\alpha} Z) \right\rangle_{|\psi_0\rangle} \\ & \quad \times \left\langle \bigotimes_{\alpha=k_1+1}^{m-1} (\mathbb{1} + (-1)^{\text{bin}(n_2)_\alpha} Z) (\mathbb{1} + (-1)^{\text{bin}(n'_2)_\alpha} Z) \right\rangle_{|\psi_0\rangle}, \quad (\text{D32}) \end{aligned}$$

where all expectation values are with respect to  $|\psi_0\rangle$ . Trivially  $\langle \hat{P}_1^{(k_1)} \hat{P}_1^{(k_1)} \rangle_{|\psi_0\rangle} = \langle \hat{P}_2^{(k_2)} \hat{P}_2^{(k_2)} \rangle_{|\psi_0\rangle} = 1$ . It can be seen that

$$\begin{aligned} & \left\langle \bigotimes_{\alpha=k_1+1}^{m-1} [\mathbb{1} + (-1)^{\text{bin}(n_1)_\alpha} Z] [\mathbb{1} + (-1)^{\text{bin}(n'_1)_\alpha} Z] \right\rangle_{|\psi_0\rangle} \\ &= \begin{cases} 4^{m-k_1-1}, & \text{if } n_1 = n'_1 = 0, \\ 0, & \text{otherwise,} \end{cases} \quad (\text{D33}) \end{aligned}$$

and similarly for  $n_2$  and  $n'_2$ . Therefore

$$\begin{aligned} & \text{Cov}[\hat{A}_{(i,j)}(k_1, k_2, n_1, n_2), \hat{A}_{(i',j')}(k_1, k_2, n'_1, n'_2)] \\ &= 2^{4m-2k_1-2k_2-4} \delta_{i,i'} \delta_{j,j'} \delta_{n_1,0} \delta_{n_2,0} \delta_{n'_1,0} \delta_{n'_2,0}, \quad (\text{D34}) \end{aligned}$$

where  $\delta$  is the Kronecker delta. We note that  $\delta_{n_1,0} = \delta_{n_1,0} \delta_{k_1,0} \forall k_1 : n_1 \in \mathcal{N}_{k_1}$  and similarly for  $\delta_{n_2,0}$ . Then substituting this into (D29) gives

$$\Gamma(G, k_1, k_2) = \sum_{(i,j)} 2^2 \gamma_{i,j}^2 2^{4m-k_1-k_2-4} \delta_{k_1,0} \delta_{k_2,0}, \quad (\text{D35})$$

which substituting into (D28) gives

$$\epsilon \tilde{S}_{\text{int}} = \frac{1}{4} \sum_{F \in \text{Part}G} \sqrt{\sum_{(i,j) \in F} \gamma_{i,j}^2 d^4}, \quad (\text{D36})$$

from which the result follows.  $\blacksquare$

*Theorem 9* (Expected shots for uncoupled state when not grouping measurements.). Given a Hamiltonian

$$\hat{H} = \sum_{i=1}^N (\hat{x}_i^2 + \hat{p}_i^2) + \sum_{j>i} \gamma_{i,j} \hat{x}_i \hat{x}_j, \quad (\text{D37})$$

for  $N \geq 2$   $d$ -level QDOs in the binary encoding and measuring each corresponding Pauli operator separately, the expected number of shots  $S$  required to achieve a constant measurement

error  $\epsilon$  for state  $|0\rangle^{\otimes mN}$  is upper bounded by

$$S = \frac{1}{\epsilon^2} \frac{1}{16} \left( \sum_{(i,j) \in G} |\gamma_{i,j}| \right)^2 d^4, \quad (\text{D38})$$

where  $G$  is the interaction graph. Assuming  $G$  is fully connected and that values of  $\gamma_{i,j}$  are independent of both  $N$  and  $d$  and randomly distributed between different  $F \in \text{Part}(G)$ , then  $S \in O(\frac{\gamma^2 N^4 d^4}{\epsilon^2})$ , where  $\gamma$  is a typical scale for  $\gamma_{i,j}$ .

*Proof.* Again we consider noninteracting and interacting contributions separately. As in Theorem 8, the uncoupled contribution is zero.

The contribution from the coupling terms is

$$\begin{aligned} \epsilon \tilde{S}_{\text{int}} &= \frac{1}{2} \sum_{k_1, k_2=0}^{m-1} \sum_{(i,j) \in G} \sum_{\substack{\hat{P}_1^{(k_1)} \in \{X,Y\}^{\otimes(k_1+1)} \\ \hat{P}_2^{(k_2)} \in \{X,Y\}^{\otimes(k_2+1)}}} \\ &\times \sum_{\substack{n_1 \in \mathcal{N}_{k_1} \\ n_2 \in \mathcal{N}_{k_2}}} \sqrt{2^2 \gamma_{i,j}^2 \text{Var}[\hat{A}_{(i,j)}(k_1, k_2, n_1, n_2)]}, \quad (\text{D39}) \end{aligned}$$

where

$$\begin{aligned} \hat{A}_{(i,j)}(k_1, k_2, n_1, n_2) &= \left( \hat{P}_1^{(k_1)} \otimes \bigotimes_{\alpha=k_1+1}^{m-1} [\mathbb{1} + (-1)^{\text{bin}(n_1)_\alpha} Z] \right)_{\text{QDO } i} \\ &\times \otimes \left( \hat{P}_2^{(k_2)} \otimes \bigotimes_{\beta=k_2+1}^{m-1} [\mathbb{1} + (-1)^{\text{bin}(n_2)_\beta} Z] \right)_{\text{QDO } j} \quad (\text{D40}) \end{aligned}$$

acts on the subregisters corresponding to QDOs  $i$  and  $j$ . Using (D34), this simplifies to

$$\epsilon \tilde{S}_{\text{int}} = \frac{1}{4} \left( \sum_{(i,j) \in G} |\gamma_{i,j}| \right) d^2, \quad (\text{D41})$$

from which the result follows.  $\blacksquare$

## APPENDIX E: NOISE MITIGATION FOR ANHARMONIC EXPERIMENTS

Unlike for the harmonic systems, for anharmonic systems it is not possible to perform the subtraction method given in (10). This is because we do not have a good way to estimate  $\lambda$  as the ground state of the zero field system is not simply given by the zero Fock state. Therefore, we use more costly error mitigation techniques in order to reduce the effect of device noise in the final energy measurements. We choose to use a combination of two error mitigation schemes, namely, zero-noise extrapolation [76,77] and an algorithmic error mitigation scheme based on Lanczos extrapolation [78].

To reduce the effect of noise on the measurement of  $\langle \hat{H} \rangle$  using zero-noise extrapolation, the expectation value is measured multiple times with the noise in the system artificially increased to various levels. The amount that the noise is amplified in the system is parameterized by a variable which we call  $\Lambda$  such that we have values  $\{\langle H \rangle_\Lambda\}$  evaluated at

each chosen value of  $\Lambda$ . A curve is then fitted to the values of  $\Lambda$  and  $\langle \hat{H} \rangle_\Lambda$  which can then be extrapolated back to  $\Lambda = 0$  to hopefully remove the effect of noise on the system. Zero noise extrapolation requires no more qubits than the original system and can be performed fairly easily. The difficulty lies in finding a good method to amplify the noise to enough values of  $\Lambda$  such that a good fit can be found, as well as choosing a good fitting function for your device noise model.

In our experiments we boost the device noise by replacing each two-qubit controlled-NOT (CNOT) gate by an odd number of CNOT gates. Since in the noiseless setting an odd number of CNOT gates has the same effect as a single CNOT gate, this has the effect of artificially boosting the two-qubit gate noise—which is expected to be the main source of noise in the superconducting system. We use three different values of  $\Lambda$  which correspond to using one, three, and five CNOTs in place of each CNOT in the circuit. A linear fit was found to match values of  $\{\langle H \rangle_\Lambda\}$  well and used to extrapolate to the  $\Lambda = 0$  to achieve a zero noise estimate.

For the further mitigation based on the Lanczos algorithm, we follow the method described in Ref. [78] to perform an  $m = 2$  Krylov subspace expansion of the Hamiltonian. In the  $m = 2$  case, this requires measuring  $\langle H \rangle$ ,  $\langle \hat{H}^2 \rangle$ , and  $\langle \hat{H}^3 \rangle$  on the noisy device to create the Krylov basis. Minimization in this basis (performed by exact matrix diagonalization) is then used to find a minimum value for  $\langle \hat{H} \rangle$  which satisfies the variational principle.

The zero-noise extrapolation and Lanczos algorithm were combined as in Ref. [78] to give an increased level of noise reduction. First zero noise extrapolation was used on expectation values of  $\langle \hat{H} \rangle$ ,  $\langle \hat{H}^2 \rangle$ , and  $\langle \hat{H}^3 \rangle$  which were then used in the Lanczos mitigation method. As with measurements of the cost function  $\langle H \rangle$  in the main text, the Pauli strings for each operator were grouped into collections that qubit-wise

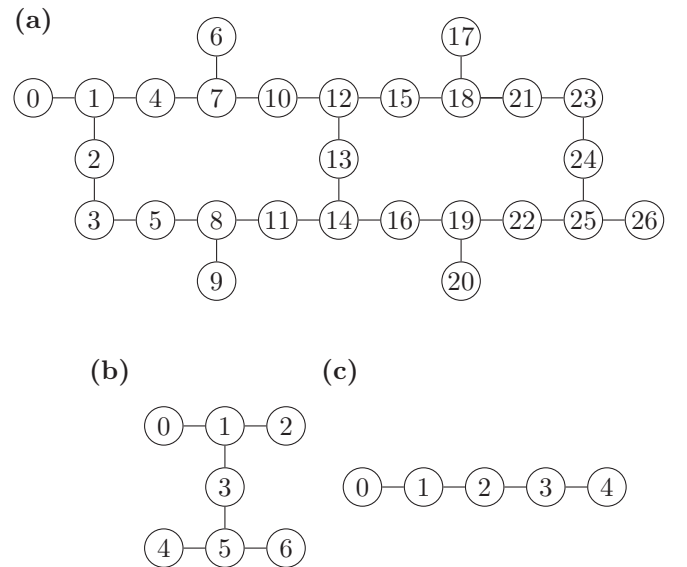


FIG. 12. Coupling map of the IBM Quantum devices used to produce the results of the paper. The devices used are (a) *ibmq\_montreal*, (b) *ibmq\_lagos*, and (c) *ibmq\_santiago* with 27, 7, and 5 qubits, respectively.

commuted and each group measured with approximately 16 million shots. Operators  $\hat{H}$ ,  $\hat{H}^2$ , and  $\hat{H}^3$  consist of 6, 7, and 8 Pauli strings which were grouped into 2, 3, and 3 commuting groups respectively.

#### APPENDIX F: IBM QUANTUM HARDWARE

For practical reasons related to device availability and queue-times we used three different quantum processors to evaluate our algorithm. All three processors were provided by IBM via the Qiskit [104] open source framework. We used *ibmq\_montreal* (27 qubits, Processor type Falcon r4), *ibmq\_lagos* (7 qubits, processor type Falcon r5.11H) and *ibmq\_santiago* (5 qubits, processor type Falcon r4L). The reported quantum volume of the aforementioned devices is 128, 32 and 32 respectively.

The experiments in Sec. III use 4 qubits on every processor, namely qubits 13, 14, 16, and 19 in *ibmq\_montreal*, qubits 1, 3, 4, and 5 in *ibmq\_lagos* and qubits 1, 2, 3, and 4 in *ibmq\_santiago* following the coupling map show in Fig. 12. The experiments were performed in a time frame of 2 months and the noise parameters (and qubit properties) were varied along the period of the experiments. The experiments in Sec. VA used qubits 14 and 16 on *ibmq\_montreal* and were performed over a period of two days.

#### APPENDIX G: SYMPLECTIC EIGENVALUES OF HARMONIC OSCILLATORS WITH LINEAR COUPLING

In the case of purely harmonic oscillators with linear coupling given by the Hamiltonian in Eqn. (1), the exact

ground-state energy (without any truncation of the Fock space) can be found by calculating the eigenspectrum of a  $2N \times 2N$  matrix.

The Hamiltonian in (1) can be rewritten as

$$\hat{H} = \hat{\mathbf{q}}^T \mathbf{H} \hat{\mathbf{q}}, \quad (\text{G1})$$

where we have defined the vector operator  $\hat{\mathbf{q}}^T = (\hat{q}_1, \hat{q}_2, \dots, \hat{q}_{2N-1}, \hat{q}_{2N}) := (\hat{x}_1, \hat{p}_1, \dots, \hat{x}_N, \hat{p}_N)$  and correlation matrix  $\mathbf{H} \in \mathbb{R}^{2N \times 2N}$  which has elements  $H_{i,j}$  equal to the coefficients of the  $\hat{q}_i \hat{q}_j$  operator in  $\hat{H}$ . In this representation, the bosonic commutation relation can be written as elements of a matrix  $[\hat{q}_i, \hat{q}_j] = \Omega_{i,j}$  where

$$\Omega = \bigoplus_{k=1}^N \omega \in \mathbb{R}^{2N \times 2N}, \quad \omega = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (\text{G2})$$

The eigenspectrum of the matrix  $|i\Omega\mathbf{H}|$  (where the absolute value is defined in the operatorial sense [105]) is  $\{\epsilon_1, \epsilon_2, \dots, \epsilon_{2N}\}$ , where  $\{\epsilon_i\}$  are known as the symplectic eigenvalues. The ground-state energy of the system is given by  $E_0 = \frac{1}{2} \sum_{i=1}^{2N} \epsilon_i$ . We consider systems of increasing coupling strength up to the point at which  $\min_i \epsilon_i = 0$  and the system becomes critical. For further details see [105,106].

We note that this method for calculating the ground-state energy works only for Hamiltonians that can be written in the quadratic form of (G1) for which ground states and thermal states are Gaussian.

- 
- [1] K. Saito, *Chemical Physics of Molecular Condensed Matter* (Springer, Singapore, 2020).
- [2] A. M. Reilly and A. Tkatchenko, Van der Waals dispersion interactions in molecular materials: Beyond pairwise additivity, *Chem. Sci.* **6**, 3289 (2015).
- [3] Y. V. Shtogun and L. M. Woods, Many-body van der Waals interactions between graphitic nanostructures, *J. Phys. Chem. Lett.* **1**, 1356 (2010).
- [4] E. Pastorczak, J. Shen, M. Hapka, P. Piecuch, and K. Pernal, Intricacies of van der Waals interactions in systems with elongated bonds revealed by electron-groups embedding and high-level coupled-cluster approaches, *J. Chem. Theory Comput.* **13**, 5404 (2017).
- [5] J. Hermann, R. A. DiStasio Jr., and A. Tkatchenko, First-principles models for van der Waals interactions in molecules and materials: Concepts, theory, and applications, *Chem. Rev.* **117**, 4714 (2017).
- [6] H. Hellmann, *Einführung in die Quantenchemie* (Franz Deuticke, Leipzig, 1937), p. 285.
- [7] R. P. Feynman, Forces in molecules, *Phys. Rev.* **56**, 340 (1939).
- [8] M. Born and R. Oppenheimer, Zur Quantentheorie der Molekeln, *Ann. Phys.* **389**, 457 (1927).
- [9] R. H. French, V. A. Parsegian, R. Podgornik, R. F. Rajter, A. Jagota, J. Luo, D. Asthagiri, M. K. Chaudhury, Y. M. Chiang, S. Granick *et al.*, Long range interactions in nanoscale science, *Rev. Mod. Phys.* **82**, 1887 (2010).
- [10] O. A. Von Lilienfeld, I. Tavernelli, U. Rothlisberger, and D. Sebastiani, Optimization of Effective Atom Centered Potentials for London Dispersion Forces in Density Functional Theory, *Phys. Rev. Lett.* **93**, 153004 (2004).
- [11] A. Tkatchenko and M. Scheffler, Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data, *Phys. Rev. Lett.* **102**, 073005 (2009).
- [12] S. Grimme, Density functional theory with London dispersion corrections, *WIREs Comput. Mol. Sci.* **1**, 211 (2011).
- [13] J. C. Slater, The theory of complex spectra, *Phys. Rev.* **34**, 1293 (1929).
- [14] S. F. Boys, Electronic wave functions—I. A general method of calculation for the stationary states of any molecular system, *Proc. R. Soc. London A* **200**, 542 (1950).
- [15] P. G. Szalay, T. Muller, G. Gidofalvi, H. Lischka, and R. Shepard, Multiconfiguration self-consistent field and multireference configuration interaction methods and applications, *Chem. Rev.* **112**, 108 (2012).
- [16] J. Čížek, On the correlation problem in atomic and molecular systems. calculation of wavefunction components in

- UrSELL-type expansion using quantum-field theoretical methods, *J. Chem. Phys.* **45**, 4256 (1966).
- [17] R. J. Bartlett and M. Musiał, Coupled-cluster theory in quantum chemistry, *Rev. Mod. Phys.* **79**, 291 (2007).
- [18] P. A. Pieniazek, S. A. Arnstein, S. E. Bradforth, A. I. Krylov, and C. D. Sherrill, Benchmark full configuration interaction and equation-of-motion coupled-cluster model with single and double substitutions for ionized systems results for prototypical charge transfer systems: Noncovalent ionized dimers, *J. Chem. Phys.* **127**, 164110 (2007).
- [19] V. Oliveira and E. Kraka, Systematic coupled cluster study of noncovalent interactions involving halogens, chalcogens, and pnictogens, *J. Phys. Chem. A* **121**, 9544 (2017).
- [20] G. Bistoni, Finding chemical concepts in the Hilbert space: Coupled cluster analyses of noncovalent interactions, *WIREs Comput. Mol. Sci.* **10**, e1442 (2020).
- [21] J. Preskill, Quantum computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018).
- [22] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio *et al.*, Variational quantum algorithms, *Nat. Rev. Phys.* **3**, 625 (2021).
- [23] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, *Rev. Mod. Phys.* **92**, 015003 (2020).
- [24] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, *Nat. Commun.* **5**, 4213 (2014).
- [25] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, [arXiv:1411.4028](https://arxiv.org/abs/1411.4028).
- [26] E. Farhi and A. W. Harrow, Quantum supremacy through the quantum approximate optimization algorithm, [arXiv:1602.07674](https://arxiv.org/abs/1602.07674).
- [27] B. Bauer, D. Wecker, A. J. Millis, M. B. Hastings, and M. Troyer, Hybrid Quantum-Classical Approach to Correlated Materials, *Phys. Rev. X* **6**, 031045 (2016).
- [28] J. M. Kreula, L. García-Álvarez, L. Lamata, S. R. Clark, E. Solano, and D. Jaksch, Few-qubit quantum-classical simulation of strongly correlated lattice fermions, *EPJ Quantum Technology* **3**, 11 (2016).
- [29] J. M. Kreula, S. R. Clark, and D. Jaksch, Non-linear quantum-classical scheme to simulate non-equilibrium strongly correlated fermionic many-body dynamics, *Sci. Rep.* **6**, 32940 (2016).
- [30] M. Lubasch, J. Joo, P. Moinier, M. Kiffner, and D. Jaksch, Variational quantum algorithms for nonlinear problems, *Phys. Rev. A* **101**, 010301(R) (2020).
- [31] P. J. Ollitrault, A. Kandala, C.-F. Chen, P. K. Barkoutsos, A. Mezzacapo, M. Pistoia, S. Sheldon, S. Woerner, J. M. Gambetta, and I. Tavernelli, Quantum equation of motion for computing molecular excitation energies on a noisy quantum processor, *Phys. Rev. Research* **2**, 043140 (2020).
- [32] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, An adaptive variational algorithm for exact molecular simulations on a quantum computer, *Nat. Commun.* **10**, 3007 (2019).
- [33] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, *Nature (London)* **549**, 242 (2017).
- [34] Y. Nam, J.-S. Chen, N. C. Pienti, K. Wright, C. Delaney, D. Maslov, K. R. Brown, S. Allen, J. M. Amini, J. Apisdorf *et al.*, Ground-state energy estimation of the water molecule on a trapped-ion quantum computer, *npj Quantum Inf.* **6**, 33 (2020).
- [35] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley, D. A. Buell *et al.*, Hartree-Fock on a superconducting qubit quantum computer, *Science* **369**, 1084 (2020).
- [36] S. McArdle, A. Mayorov, X. Shan, S. Benjamin, and X. Yuan, Digital quantum simulation of molecular vibrations, *Chem. Sci.* **10**, 5725 (2019).
- [37] P. J. Ollitrault, A. Baiardi, M. Reiher, and I. Tavernelli, Hardware efficient quantum algorithms for vibrational structure calculations, *Chem. Sci.* **11**, 6842 (2020).
- [38] I. Rungger, N. Fitzpatrick, H. Chen, C. H. Alderete, H. Apel, A. Cowtan, A. Patterson, D. M. Ramo, Y. Zhu, N. H. Nguyen *et al.*, Dynamical mean field theory algorithm and experiment on quantum computers, [arXiv:1910.04735](https://arxiv.org/abs/1910.04735).
- [39] B. Jaderberg, A. Agarwal, K. Leonhardt, M. Kiffner, and D. Jaksch, Minimum hardware requirements for hybrid quantum-classical DMFT, *Quantum Sci. Technol.* **5**, 034015 (2020).
- [40] T. Helgaker, P. Jorgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (John Wiley & Sons, Chichester, 2014).
- [41] C. D. Sherrill, Computational scaling of the configuration interaction method with system size, Quantum Chemistry Lecture Notes, School of Chemistry and Biochemistry, Georgia Institute of Technology, 1996.
- [42] C. D. Sherrill, Wavefunction theory approaches to noncovalent interactions, in *Non-Covalent Interactions in Quantum Chemistry and Physics*, edited by A. O. de la Roza and G. A. DiLabio (Elsevier, Amsterdam, 2017), pp. 137–168.
- [43] F. S. Cipcigan, J. Crain, V. P. Sokhan, and G. J. Martyna, Electronic coarse graining: Predictive atomistic modeling of condensed matter, *Rev. Mod. Phys.* **91**, 025003 (2019).
- [44] F. Cipcigan, V. Sokhan, G. Martyna, and J. Crain, Structure and hydrogen bonding at the limits of liquid water stability, *Sci. Rep.* **8**, 1718 (2018).
- [45] V. P. Sokhan, A. P. Jones, F. S. Cipcigan, J. Crain, and G. J. Martyna, Signature properties of water: Their molecular electronic origins, *Proc. Natl. Acad. Sci. USA* **112**, 6341 (2015).
- [46] V. P. Sokhan, A. Jones, F. S. Cipcigan, J. Crain, and G. J. Martyna, Molecular-Scale Remnants of the Liquid-Gas Transition in Supercritical Polar Fluids, *Phys. Rev. Lett.* **115**, 117801 (2015).
- [47] F. S. Cipcigan, V. P. Sokhan, J. Crain, and G. J. Martyna, Electronic coarse graining enhances the predictive power of molecular simulation allowing challenges in water physics to be addressed, *J. Comput. Phys.* **326**, 222 (2016).
- [48] P. Gokhale, O. Angiuli, Y. Ding, K. Gui, T. Tomesh, M. Suchara, M. Martonosi, and F. T. Chong,  $O(N^3)$  measurement cost for variational quantum eigensolver on molecular Hamiltonians, *IEEE Trans. Quantum Eng.* **1**, 1 (2020).
- [49] M. J. E. Westbroek, P. R. King, D. D. Vvedensky, and S. Dürr, User's guide to Monte Carlo methods for evaluating path integrals, *Am. J. Phys.* **86**, 293 (2018).

- [50] A. Jones, F. Cipcigan, V. P. Sokhan, J. Crain, and G. J. Martyna, Electronically Coarse-Grained Model for Water, *Phys. Rev. Lett.* **110**, 227801 (2013).
- [51] A. P. Jones, J. Crain, V. P. Sokhan, T. W. Whitfield, and G. J. Martyna, Quantum Drude oscillator model of atoms and molecules: Many-body polarization and dispersion interactions for atomistic simulation, *Phys. Rev. B* **87**, 144103 (2013).
- [52] N. P. D. Sawaya, T. Menke, T. H. Kyaw, S. Johri, A. Aspuru-Guzik, and G. G. Guerreschi, Resource-efficient digital quantum simulation of  $d$ -level systems for photonic, vibrational, and spin- $s$  Hamiltonians, *npj Quantum Inf.* **6**, 49 (2020).
- [53] E. F. Dumitrescu, A. J. McCaskey, G. Hagen, G. R. Jansen, T. D. Morris, T. Papenbrock, R. C. Pooser, D. J. Dean, and P. Lougovski, Cloud Quantum Computing of an Atomic Nucleus, *Phys. Rev. Lett.* **120**, 210501 (2018).
- [54] G. Maroulis, C. Makris, U. Hohm, and D. Goebel, Electrooptical properties and molecular polarization of iodine,  $I_2$ , *J. Phys. Chem. A* **101**, 953 (1997).
- [55] S. Yamamoto, Y. Ishibashi, Y. Inamura, Y. Katayama, T. Mishina, and J. Nakahara, Pressure dependence of local structure in liquid carbon disulfide, *J. Chem. Phys.* **124**, 144511 (2006).
- [56] B. Schropp and P. Tavan, The polarizability of point-polarizable water models: Density functional theory/molecular mechanics results, *J. Phys. Chem. B* **112**, 6233 (2008).
- [57] Y. Huang, X. Zhang, Z. Ma, W. Li, Y. Zhou, J. Zhou, W. Zheng, and C. Q. Sun, Size, separation, structural order and mass density of molecules packing in water and ice, *Sci. Rep.* **3**, 3005 (2013).
- [58] A. Jena, S. Genin, and M. Mosca, Pauli partitioning with respect to gate sets, [arXiv:1907.07859](https://arxiv.org/abs/1907.07859).
- [59] O. Crawford, B. van Straaten, D. Wang, T. Parks, E. Campbell, and S. Brierley, Efficient quantum measurement of Pauli operators in the presence of finite sampling error, *Quantum* **5**, 385 (2021).
- [60] V. Verteletskyi, T.-C. Yen, and A. F. Izmaylov, Measurement optimization in the variational quantum eigensolver using a minimum clique cover, *J. Chem. Phys.* **152**, 124114 (2020).
- [61] A. F. Izmaylov, T.-C. Yen, R. A. Lang, and V. Verteletskyi, Unitary partitioning approach to the measurement problem in the variational quantum eigensolver method, *J. Chem. Theory Comput.* **16**, 190 (2020).
- [62] A. Zhao, A. Tranter, W. M. Kirby, S. F. Ung, A. Miyake, and P. J. Love, Measurement reduction in variational quantum algorithms, *Phys. Rev. A* **101**, 062322 (2020).
- [63] T.-C. Yen, V. Verteletskyi, and A. F. Izmaylov, Measuring all compatible operators in one series of single-qubit measurements using unitary transformations, *J. Chem. Theory Comput.* **16**, 2400 (2020).
- [64] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, 2018).
- [65] F. Vatan and C. Williams, Optimal quantum circuits for general two-qubit gates, *Phys. Rev. A* **69**, 032315 (2004).
- [66] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik, Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz, *Quantum Sci. Technol.* **4**, 014008 (2018).
- [67] M. Majland and N. T. Zinner, Resource-efficient encoding algorithm for variational bosonic quantum simulations, [arXiv:2102.11886](https://arxiv.org/abs/2102.11886).
- [68] O. Christiansen, A second quantization formulation of multi-mode dynamics, *J. Chem. Phys.* **120**, 2140 (2004).
- [69] O. Christiansen, Vibrational coupled cluster theory, *J. Chem. Phys.* **120**, 2149 (2004).
- [70] J. Kendrick, F. J. J. Leusen, and M. A. Neumann, Empirical van der Waals corrections to solid-state density functional theory: Iodine and phosphorous containing molecular crystals, *J. Comput. Chem.* **33**, 1615 (2012).
- [71] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, Quantum circuit learning, *Phys. Rev. A* **98**, 032309 (2018).
- [72] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).
- [73] A. Jones, Quantum Drude oscillators for accurate many-body intermolecular forces, Ph.D. thesis, University of Edinburgh, 2010.
- [74] R. W. Boyd, *Nonlinear Optics* (Academic Press, 2020).
- [75] N. P. D. Sawaya and J. Huh, Quantum algorithm for calculating molecular vibronic spectra, *J. Phys. Chem. Lett.* **10**, 3586 (2019).
- [76] K. Temme, S. Bravyi, and J. M. Gambetta, Error Mitigation for Short-Depth Quantum Circuits, *Phys. Rev. Lett.* **119**, 180509 (2017).
- [77] Y. Li and S. C. Benjamin, Efficient Variational Quantum Simulator Incorporating Active Error Minimization, *Phys. Rev. X* **7**, 021050 (2017).
- [78] P. Suchsland, F. Tacchino, M. H. Fischer, T. Neupert, P. K. Barkoutsos, and I. Tavernelli, Algorithmic error mitigation scheme for current quantum processors, *Quantum* **5**, 492 (2021).
- [79] B. M. Axilrod and E. Teller, Interaction of the van der Waals type between three atoms, *J. Chem. Phys.* **11**, 299 (1943).
- [80] L. Bittel and M. Kliesch, Training Variational Quantum Algorithms Is NP-Hard, *Phys. Rev. Lett.* **127**, 120502 (2021).
- [81] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Barren plateaus in quantum neural network training landscapes, *Nat. Commun.* **9**, 4812 (2018).
- [82] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Cost function dependent barren plateaus in shallow parametrized quantum circuits, *Nat. Commun.* **12**, 1791 (2021).
- [83] S. Wang, E. Fontana, M. Cerezo, K. Sharma, A. Sone, L. Cincio, and P. J. Coles, Noise-induced barren plateaus in variational quantum algorithms, *Nat. Commun.* **12**, 6961 (2021).
- [84] C. Ortiz Marrero, M. Kieferová, and N. Wiebe, Entanglement-induced barren plateaus, *PRX Quantum* **2**, 040316 (2021).
- [85] A. L. Blum and R. L. Rivest, Training a 3-node neural network is NP-complete, *Neural Networks* **5**, 117 (1992).
- [86] A. Pesah, M. Cerezo, S. Wang, T. Volkoff, A. T. Sornborger, and P. J. Coles, Absence of Barren Plateaus in Quantum Convolutional Neural Networks, *Phys. Rev. X* **11**, 041011 (2021).
- [87] K. Bharti and T. Haug, Iterative quantum-assisted eigensolver, *Phys. Rev. A* **104**, L050401 (2021).
- [88] R. Wiersema, C. Zhou, Y. de Sereville, J. F. Carrasquilla, Y. B. Kim, and H. Yuen, Exploring entanglement and optimization within the Hamiltonian variational ansatz, *PRX Quantum* **1**, 020319 (2020).



- [89] E. Grant, L. Wossnig, M. Ostaszewski, and M. Benedetti, An initialization strategy for addressing barren plateaus in parametrized quantum circuits, *Quantum* **3**, 214 (2019).
- [90] G. Verdon, M. Broughton, J. R. McClean, K. J. Sung, R. Babbush, Z. Jiang, H. Neven, and M. Mohseni, Learning to learn with quantum neural networks via classical neural networks, [arXiv:1907.05415](https://arxiv.org/abs/1907.05415).
- [91] H.-Y. Liu, T.-P. Sun, Y.-C. Wu, Y.-J. Han, and G.-P. Guo, A parameter initialization method for variational quantum algorithms to mitigate barren plateaus based on transfer learning, [arXiv:2112.10952](https://arxiv.org/abs/2112.10952).
- [92] A. Anand, M. Degroote, and A. Aspuru-Guzik, Natural evolutionary strategies for variational quantum computation, *Mach. Learn.: Sci. Technol.* **2**, 045012 (2021).
- [93] T. L. Patti, K. Najafi, X. Gao, and S. F. Yelin, Entanglement devised barren plateau mitigation, *Phys. Rev. Research* **3**, 033090 (2021).
- [94] S. H. Sack, R. A. Medina, A. A. Michailidis, R. Kueng, and M. Serbyn, Avoiding barren plateaus using classical shadows, [arXiv:2201.08194](https://arxiv.org/abs/2201.08194).
- [95] A. V. Uvarov and J. D. Biamonte, On barren plateaus and cost function locality in variational quantum algorithms, *J. Phys. A: Math. Theor.* **54**, 245301 (2021).
- [96] M. Fingerhuth, T. Babej, and C. Ing., A quantum alternating operator ansatz with hard and soft constraints for lattice protein folding, [arXiv:1810.13411](https://arxiv.org/abs/1810.13411).
- [97] A. Robert, P. K. Barkoutsos, S. Woerner, and I. Tavernelli, Resource-efficient quantum algorithm for protein folding, *npj Quantum Inf.* **7**, 38 (2021).
- [98] <https://doi.org/10.5281/zenodo.22558>.
- [99] F. London, On centers of van der Waals attraction. *J. Phys. Chem.* **46**, 305 (1942).
- [100] J. O. Hirschfelder, C. F. Curtiss, R. B. Bird, and M. G. Mayer, *Molecular Theory of Gases and Liquids* (Wiley, New York, 1964).
- [101] V. I. Voloshin, *Introduction to Graph and Hypergraph Theory* (Nova Science, 2009).
- [102] J. H. Van Lint and R. M. Wilson, *A Course in Combinatorics* (Cambridge University Press, 2001).
- [103] D. Wecker, M. B. Hastings, and M. Troyer, Progress towards practical quantum variational algorithms, *Phys. Rev. A* **92**, 042303 (2015).
- [104] M. S. Anis *et al.*, Qiskit/qiskit: Qiskit 0.36.2 (2022), <https://doi.org/10.5281/zenodo.6560959>.
- [105] S. Pirandola, A. Serafini, and S. Lloyd, Correlation matrices of two-mode bosonic systems, *Phys. Rev. A* **79**, 052327 (2009).
- [106] N. Schuch, J. I. Cirac, and M. M. Wolf, Quantum states on harmonic lattices, *Commun. Math. Phys.* **267**, 65 (2006).