




Quantum computation for periodic solids in second quantization

Aleksei V. Ivanov ^{1,*} Christoph Sünderhauf,¹ Nicole Holzmann ¹ Tom Ellaby,²
Rachel N. Kerber,² Glenn Jones,² and Joan Camps ¹

¹Riverlane, St Andrews House, 59 St Andrews Street, Cambridge CB2 3BZ, United Kingdom

²Johnson Matthey Technology Centre, Blounts Court, Sonning Common RG4 9NH, United Kingdom



(Received 14 October 2022; accepted 30 January 2023; published 23 March 2023)

In this work, we present a quantum algorithm for ground-state energy calculations of periodic solids on error-corrected quantum computers. The algorithm is based on the sparse qubitization approach in second quantization and developed for Bloch and Wannier basis sets. We show that Wannier functions require less computational resources with respect to Bloch functions because (i) the L_1 norm of the Hamiltonian is considerably lower and (ii) the translational symmetry of Wannier functions can be exploited in order to reduce the amount of classical data that must be loaded into the quantum computer. The resource requirements of the quantum algorithm are estimated for periodic solids such as NiO and PdO. These transition metal oxides are industrially relevant for their catalytic properties. We find that ground-state energy estimation of Hamiltonians approximated using 200–900 spin orbitals requires ca. 10^{10} – 10^{12} T gates and up to 3×10^8 physical qubits for a physical error rate of 0.1%.

DOI: [10.1103/PhysRevResearch.5.013200](https://doi.org/10.1103/PhysRevResearch.5.013200)

I. INTRODUCTION

Quantum mechanical simulation of molecules and materials is a promising application area of quantum computers [1–3] that will enable the calculation of key properties of chemical systems with controllable errors using physically accurate models. Following Feynman’s original idea of modeling quantum systems on quantum computers [4] and the first formalized procedures for carrying out such simulations [5–7], a plethora of quantum algorithms for calculating energies of molecular systems have been developed in recent years [8–34]. Similarly, but to a lesser degree, quantum algorithms taking into account the specifics of condensed matter applications have also been conceived. These include the development of different flavors of variational quantum eigensolvers (VQE) [35–38], the quantum imaginary time evolution algorithm [39], and fault-tolerant algorithms [24,34,40–43] for simulation of model Hamiltonians, such as the Hubbard model, as well as first-principles Hamiltonians.

Quantum computers can provide a computational advantage over classical computers only for hard classical problems. These include the simulation of so-called strongly correlated systems and, more practically, problems that are not solved with sufficient accuracy using classical methods with low computation cost—such as Kohn-Sham density functional theory (KS-DFT) [44–46] or coupled-cluster theory [47,48]. Notwithstanding varying definitions and interpretations of

“strong correlation”, and an ongoing debate regarding the extent to which KS-DFT can describe such systems [49,50], the general consensus is that molecular and solid-state systems with a large number of localized d or f electrons present a significant challenge for classical simulations. Examples of such systems include transition metal oxides such as NiO and PdO used in heterogeneous catalysis applications. The number of localized sites in such systems is formally infinite as the solids should be simulated at the thermodynamic limit. In practice, one restricts calculations to a periodic finite-sized cell (also referred to as supercell) with ca. 30–100 unique transition metal atoms; all other atoms in the solid are replicas of those in this computational cell.

The ability to accurately model the electronic structure of materials such as NiO and PdO would no doubt prove extremely useful in the study of heterogeneous catalysis, a field with no shortage of materials that are poorly described by DFT. It is often the case that the interpretation of calculated results (e.g., regarding trends in activity) must be presented with significant caveats regarding the underlying nature of the models used.

In this paper, we focus on the calculation of the ground-state energy of electrons in materials within the Born-Oppenheimer approximation [51]. This corresponds to finding the lowest eigenvalue of the electronic Hamiltonian for a fixed position of the nuclei. Two main families of quantum algorithms can perform such calculation: VQE [13] and quantum phase estimation (QPE) [52–54]. While VQE might have its merits in certain use cases, it appears the emerging consensus is that QPE has a superior scaling with the system size [1,55]. In order to estimate the eigenvalues of the Hamiltonian with QPE, one has to implement a unitary operator encoding the spectrum of the Hamiltonian. QPE requires deep quantum circuits, and as such it will need to run on error-corrected quantum computers. In such error-corrected implementations

*Corresponding author: aleksei.ivanov@riverlane.com

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

one must strive to minimize the number of T gates needed to encode the Hamiltonian, as these gates are the costliest to implement (see e.g., [56]). To this date, the most cost-efficient approaches for such encodings are based on the so-called qubitization technique [25–28,31,32,57]. Previous work on second quantized Hamiltonians for realistic solid state systems has mainly focused on the Trotterization approach [24,40]. In this paper, we adapt the sparse qubitization approach to the simulation of crystalline solids with QPE and estimate the resources required for calculating the ground-state energy of crystals in error-corrected quantum computers.

The quantum resources required to simulate a Hamiltonian strongly depend on the single-electron basis sets used to represent electron interactions. For crystalline solids, plane waves (PW) currently appear to be one of the most efficient basis sets both in first and second quantization [24,34]. An advantage of using PW basis sets is the sparse representation of the electronic Hamiltonian. This advantage is always exploited in classical computations such as KS-DFT [58,59]. The number of two-body terms in PW representation scales cubically with the size of the basis set. The main disadvantage of such basis sets, however, is that they require a large number of basis functions, especially in all-electron calculations. For crystalline solids one can exploit Bloch functions instead, which are plane waves times a periodic function with the periodicity of the unit cell. In the Bloch representation, the number of terms also scales cubically with the system size, and at the same time such a representation allows using localized atomic orbitals as the periodic constituent of the orbitals. The other commonly used representation in computational condensed matter physics is the Wannier representation, in which orbitals are localized in space [60–62]. Wannier orbitals can be related to Bloch functions through Fourier transformation, and can be localized using unitary optimization in order to produce maximally localized Wannier functions [60]. When the periodic function in the Bloch representation is a constant, the Wannier representation coincides with the PW dual representation introduced in the context of quantum computing in Ref. [24]. At the same time, Wannier orbitals can be spanned in the localized atomic orbital basis, which in turn can significantly reduce the size of the basis set for an accurate description of finite band-gap solids. In this paper, we investigate Bloch and Wannier representations in the context of qubitized QPE. We note that such basis sets have recently been investigated in the context of the VQE algorithm [36].

Quantum computation with qubitization-based QPE requires a large number of gates in a circuit. In order to perform large quantum computations, one has to encode a logical qubit using several physical qubits with a technique known as quantum error correction [63,64]. In order to estimate the total number of physical and logical qubits required for the implementation of quantum algorithms as well as their runtime, we have followed Litinski's approach [65]. This scheme operates the surface code [66] with lattice surgery [56,67], and compiles logical quantum circuits down to just multi-qubit T gates and multi-qubit measurements—all Clifford gates are commuted past the end of the circuit. In this way, runtime is directly related to T-gate count.

The article is organized as follows. In Sec. II, we first describe the relevance of modeling bulk materials such as

NiO and PdO for applications to heterogeneous catalysis—an area where quantum computation can provide high accuracy results when error-corrected quantum computers become available. In Sec. III, the Hamiltonian, basis sets, and quantum algorithms for modeling of crystalline solids are introduced. In Sec. IV, we discuss the performance of quantum algorithms and provide quantum resource estimations for several solid state systems. Finally, discussion and conclusions are presented in Sec. V. Detailed logical qubit and Toffoli gate counts of the sparse qubitization are provided in the Appendix.

II. MATERIALS AND HETEROGENEOUS CATALYSTS

Catalysts are used in practically every industrial chemical process, with applications in agriculture, transportation, and energy production, among many others. The function of a catalyst to ultimately reduce the energy requirements of a process to make it viable or more efficient means that catalytic processes are a key component for ensuring a sustainable future and reducing human impact on the environment. Transition metal oxide catalysts are essential components for many important industrial processes (such as refining and petrochemistry, fuel cells, hydrogen production, biomass conversion, photocatalysis) where they are used both directly, as the active material (providing the active site), and indirectly, as a support material (commonly as a reducible oxide taking a secondary role in the catalysis). The overall performance of the solid catalyst depends on many factors, including the particle size, particle shape, crystallinity, chemical composition, and all preparation and activation procedures. High catalytic efficiencies are achieved as the number of active surface sites grows, while the structural flexibility of supported metal catalysts (dynamic structural changes) is key for the catalytic reactivity when we consider that the surface sites repeatedly participate in adsorption/desorption cycles.

The systems considered in this paper, nickel oxide (NiO) and palladium oxide (PdO), both form the basis of industrially relevant catalyst materials. In the field of energy and environment, natural gas reforming is the most common process used in industry to produce H_2 from fossil fuels, known as methane steam reforming (MSR). Here NiO is reduced to Ni, which functions as a high-temperature catalyst. Despite its age and ubiquity, the MSR process still has many technical challenges, for instance around deactivation from carbon whisker formation and stability at high temperature. Under certain operating conditions, a local oxidizing environment can form within the reactor, leading to the deactivation of Ni due to NiO being present. Thermodynamics can predict the conditions at which this can occur [68]. However, in general, it is still a challenge to obtain reliable or accurate thermodynamic parameters for strongly correlated oxide materials, especially when they deviate from the bulk limit such as in nanoparticles.

Methane has an estimated greenhouse warming potential (GWP 100) of 27.9 [69], meaning its emissions contribute significantly to global warming and climate change; it is therefore necessary to reduce them wherever possible. Among the many different technologies for methane abatement, methane combustion catalysts based on palladium can be found. Such technologies include after-treatment for combustion of natural

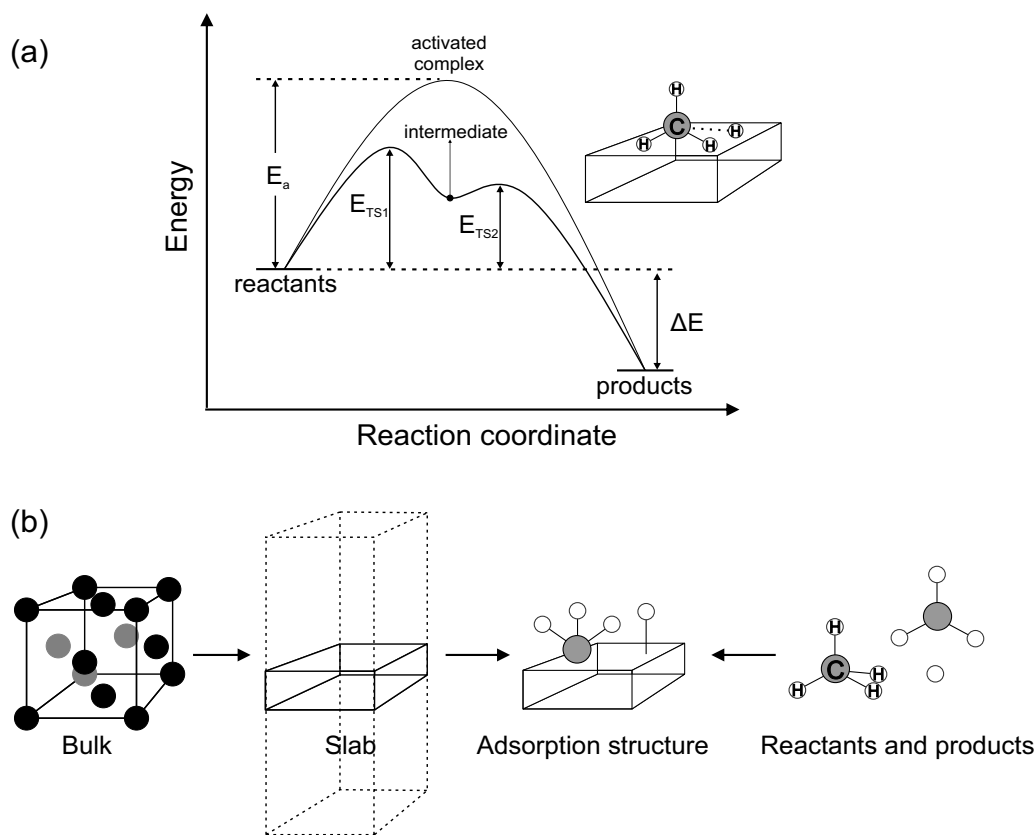


FIG. 1. (a) Schematic of a reaction pathway with and without a catalyst, and (b) diagram of the necessary calculation steps required to determine these properties. Relative energies between the reactants, products, and intermediates (if they are known) can be calculated via density functional theory or other computational approaches. Reaction barriers for a catalyzed reaction can be approximated via adsorption energy calculations in many cases [70], or by transition state searches, which require the calculation of forces as well as the energy.

gas engines (CNG) and diesel oxidation catalysts (DOC) as well as in mine ventilation systems.

In the above applications, methane is efficiently combusted over palladium (or alloyed) oxide catalyst to produce H_2O and CO_2 , with activity in this process influenced by a number of factors. A technical target in practical catalysis is to reduce the temperature at which this occurs, allowing for a lower operating temperature and more efficient handling of emissions. Partial oxidation can sometimes occur, and may indeed be desirable in the development of processes to produce precursors for more complex chemicals. The ability to simulate accurately not only the activity but also the selectivity, which is a measure of a catalyst's ability to promote the formation of the desired product(s) over other possibilities, is crucial to the prediction of new catalysts.

Figure 1(a) shows a schematic of a typical catalyzed reaction. The presence of a catalyst provides additional reaction coordinates, or reaction intermediates, with their own activation energies (E_{TS1} , E_{TS2}). For an effective catalyst, these energies are necessarily lower than the uncatalyzed activation energy E_a . In the case of heterogeneous catalysts, reaction intermediates are typically adsorption steps, where one or more of the reactants binds to a surface site of the catalyst. Depending on the complexity of the reaction mechanisms, there may be a large number of these intermediates as well as branches and side reactions that must be considered when studying a reaction in order to determine the key step(s).

It is often necessary to find these steps, which govern the activity and/or selectivity of a catalyst, as in doing so, the problem is reduced to fewer dimensions and descriptors that facilitate a more rapid study. For example, in a kinetic analysis the largest activation energy is usually of most interest, as this will be the rate determining step. Whilst this knowledge may be well established in well-known reactions, it can be necessary to perform many calculations in more novel applications. Furthermore, whilst the accuracy of current computational approaches may be good enough to predict trends in similar systems, obtaining chemical accuracy and absolute values for detailed kinetic studies remains a challenge. Figure 1(b) shows a typical set of model systems that would be used to estimate the energetics of a heterogeneous catalytic process.

When running simulations of a catalyst, consideration needs to be made of the question at hand and the level of accuracy that is needed. Broadly speaking, we are interested in activity, selectivity and stability. When simulating activity, we often need a kinetic model, which can provide rates or turn-over frequencies. If we are interested in screening for materials, it is often sufficient to correlate these rates with descriptors [71].

For example, following the Sabatier principle [72], which is employed primarily for materials screening, calculating the (heterogeneous) catalytic activity of a material is performed by determining the binding strengths of the reactants,

products and any important intermediates of a given reaction with the surface of that material. These binding strengths can be determined from energy calculations using a wide variety of models, each with their own trade-offs between accuracy, transferability and computational cost.

However, if we are interested in predicting reactor performance or process conditions then we need significantly greater precision in the simulated parameters. Likewise, simulating the often subtle differences in competing reactions (which result in different products) typically requires greater accuracy in calculations to predict selectivity.

Whilst the questions of activity and selectivity are crucial for a material's function as a catalyst, when looking for a technical solution, the question of stability becomes critical. Catalysts often need to operate over many years under harsh conditions (high temperature, pressure, contaminated conditions and, in the case of electrocatalysis, high potentials and corrosive environments). The simulation of stability introduces a whole range of other problems; for instance, predicting morphological changes and thermal degradation of a catalyst requires a large number of calculations, often of large model systems, to allow sintering of nanoparticles or ceramic supports to be conducted. Material complexity (e.g., simulation of realistic metal/ceramic interfaces), bridging time and length scales where accurate atomic-scale materials properties can be fed into multiscale models, are all open challenges in this area.

DFT is one of the most successful and widely used models for calculating the energies of molecular and solid state systems relevant to industrial processes. It is an *ab initio* method that uses functionals of the electronic density to calculate energy rather than attempt to deal directly with the many-body wavefunction. In KS-DFT, the electronic density is constructed using a fictitious set of noninteracting single electron wavefunctions and approximating an unknown correction term. This term, known as the exchange-correlation (XC) functional, includes exchange and correlation effects as well as discrepancy between the real and noninteracting kinetic energy. There are many choices, though all of them approximated, for its form.

Ultimately, it is the use of single-particle wavefunctions in DFT that leads to some of its most prominent shortcomings. In the case of NiO, and indeed most transition metal oxides, the strong electron-electron interactions of the d electrons in these materials is poorly described by approximate KS-DFT, leading to over-delocalization of these bands (and to the prediction of more metallic electronic structures than the reality). A Hubbard U [73] correction can be used alongside local density approximation (LDA) and generalized gradient approximation (GGA) XC functionals to mitigate this issue in some cases, although it is overly empirical in nature. While the use of hybrid XC functionals such as PBE0 [74] can sometimes perform better [75], due to the inclusion of Hartree-Fock exact exchange, the fraction of exact exchange to use can be varied (depending on the XC functional used), which again leads to empirical fitting. Hybrid functionals are also incomplete (and incorrect) in their description of the electronic structure, and are by no means a guaranteed improvement over GGA functionals in their prediction of transition metal oxide properties [76].

To model the bulk properties of materials effectively, the use of periodic boundary conditions (PBCs) is required, allowing for a simulation box to include only the primitive unit cell in highly ordered systems. Even in disordered systems, periodicity is still imposed (on a larger unit cell), as the approximation still provides more representative models than any nonperiodic alternative, without extending the system far beyond practical limits.

The study of heterogeneous catalysis primarily concerns the properties of surfaces, so slab models are often used. These are also periodic, albeit in two dimensions rather than three. Bulk calculations are also required in order to determine the surface energies of the facets of a material, which, for example, allow for the prediction of the expected shape of nanoparticles, as well as which facets are most predominant and relevant for catalysis. The stability of a material is another important aspect that can be predicted by energy calculations on bulk systems.

III. METHODOLOGY

A. Hamiltonian for periodic systems

The Hamiltonian of interacting electrons in the Born-Oppenheimer approximation can be written as follows:

$$\hat{H} = \hat{H}^{(0)} + \hat{H}^{(1)} + \hat{H}^{(2)}, \quad (1)$$

where $\hat{H}^{(0)}$ is a constant term describing nuclear repulsion, $\hat{H}^{(1)}$ and $\hat{H}^{(2)}$ are one-body and two-body terms, respectively [[77], p. 32],

$$\hat{H}^{(1)} = \iint d\mathbf{x} d\mathbf{x}' \hat{\psi}^\dagger(\mathbf{x}) h(\mathbf{x}, \mathbf{x}') \hat{\psi}(\mathbf{x}'), \quad (2)$$

$$\hat{H}^{(2)} = \frac{1}{2} \iint d\mathbf{x} d\mathbf{x}' \hat{\psi}^\dagger(\mathbf{x}) \hat{\psi}^\dagger(\mathbf{x}') g(\mathbf{x}, \mathbf{x}') \hat{\psi}(\mathbf{x}') \hat{\psi}(\mathbf{x}), \quad (3)$$

\mathbf{x} denotes position and spin, (\mathbf{r}, σ) , of an electron and the integration domain is over the volume of the macroscopic crystal, V . In this paper, we do not consider the external magnetic field or spin-orbit coupling and therefore, the one- and two-body kernels are diagonal with respect to spin degrees of freedom. The spatial part of one-body kernel is

$$h(\mathbf{r}) = -\frac{1}{2} \nabla^2 + U(\mathbf{r}), \quad (4)$$

where $U(\mathbf{r})$ is the nuclei potential

$$U(\mathbf{r}) = \sum_{a \in V} \frac{Z_a}{|\mathbf{r} - \mathbf{P}_a|}, \quad (5)$$

Z_a and \mathbf{P}_a are the nuclear charge and position of nucleus a . The spatial part of two-body kernel is

$$g(\mathbf{r}, \mathbf{r}') = g(|\mathbf{r} - \mathbf{r}'|) = \frac{1}{|\mathbf{r} - \mathbf{r}'|}. \quad (6)$$

We assume Born-von-Kármán periodic boundary conditions at the boundaries of the macroscopic crystal, which is defined by the vectors $\mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3$,

$$A(\mathbf{r} + \mathbf{L}_\alpha) = A(\mathbf{r}), \quad A = \hat{\psi}, h, U, g \quad \alpha = 1, 2, 3. \quad (7)$$

In this case, the external potential and two-body kernel are defined in terms of their Fourier series

$$U(\mathbf{r}) = \sum_{a \in V} \sum_{\mathbf{K}} \frac{4\pi Z_a e^{-i\mathbf{K}\mathbf{P}_a}}{V\mathbf{K}^2} e^{i\mathbf{K}\mathbf{r}}, \quad (8)$$

$$g(\mathbf{r}) = \sum_{\mathbf{K}} \frac{4\pi}{V\mathbf{K}^2} e^{i\mathbf{K}\mathbf{r}}, \quad (9)$$

where \mathbf{K} satisfies

$$\mathbf{K}\mathbf{L}_\alpha = 2\pi m_\alpha, \quad m_\alpha \in \mathbb{Z}, \quad \alpha = 1, 2, 3. \quad (10)$$

Crystalline solids consist of unit cells and each unit cell V_{uc} is defined by translation lattice vectors, $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$. Each unit cell can be labeled with \mathbf{R} indicating a node of the Bravais lattice

$$\mathbf{R} = A^T \mathbf{n}; \quad \mathbf{n} = (n_1, n_2, n_3)^T, \quad n_\alpha \in \mathbb{Z}; \quad A = [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3]^T. \quad (11)$$

Let $N_\alpha \geq 1$ be the number of unit cells along \mathbf{a}_α , $n_\alpha = 0, 1, \dots, N_\alpha - 1$ and thus, the total number of unit cells, which spans the whole finite macroscopic crystal is $N = N_1 N_2 N_3$. We also introduce the reciprocal lattice, which is defined as

$$\mathbf{G} = B\mathbf{n}; \quad \mathbf{n} = (n_1, n_2, n_3)^T, \quad n_\alpha \in \mathbb{Z}; \\ B = [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3] = 2\pi A^{-1}. \quad (12)$$

The vectors $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ and $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ satisfy the following relations:

$$\mathbf{a}_\alpha \mathbf{b}_\beta = 2\pi \delta_{\alpha\beta}. \quad (13)$$

In the case of crystalline solids, the external potential can also be rewritten in terms of reciprocal lattice vectors, because it has periodicity of the lattice,

$$U(\mathbf{r}) = \sum_{a \in V_{uc}} \sum_{\mathbf{G}} \frac{4\pi Z_a e^{-i\mathbf{G}\mathbf{P}_a}}{V_{uc}\mathbf{G}^2} e^{i\mathbf{G}\mathbf{r}}. \quad (14)$$

This is similar to Eq. (8) but written for unit-cell periodicity instead of periodicity within macroscopic crystal.

In order to perform practical calculations, one can choose a single-particle basis, which is suitable for the problem of interest,

$$\hat{\psi}(\mathbf{x}) = \sum_{p,\sigma} \psi_p(\mathbf{r}) \hat{a}_{p\sigma}, \quad (15)$$

where p can be a set of numbers describing a single-particle state such as wave-vector index and band index, for example. The Hamiltonian in the new basis set can be written as

$$\hat{H} = H^{(0)} + \sum_{pq\sigma} \left(h_{pq} - \frac{1}{2} \sum_r g_{prrq} \right) \hat{a}_{p\sigma}^\dagger \hat{a}_{q\sigma} \\ + \frac{1}{2} \sum_{pqrs;\sigma,\tau} g_{pqrs} \hat{a}_{p\sigma}^\dagger \hat{a}_{q\sigma} \hat{a}_{r\tau}^\dagger \hat{a}_{s\tau}, \quad (16)$$

where two-body matrix elements (often referred to as the electron-repulsion integrals or just Coulomb integrals) are

$$g_{pqrs} = \iint_V d^3\mathbf{r} d^3\mathbf{r}' \psi_p^*(\mathbf{r}) \psi_q(\mathbf{r}) \psi_r^*(\mathbf{r}') \psi_s(\mathbf{r}') g(|\mathbf{r} - \mathbf{r}'|). \quad (17)$$

From this definition, g_{pqrs} obeys the fourfold symmetry relations

$$g_{pqrs} = g_{rspq} = g_{qpsr}^* = g_{srqp}^*. \quad (18)$$

If coefficients are real then g_{pqrs} obeys the eightfold symmetry relations

$$g_{pqrs} = g_{pqsr} = g_{qprs} = g_{qpsr} = g_{rspq} = g_{srpq} = g_{rsqp} = g_{srqp}. \quad (19)$$

The set of orbitals $\{\psi_p(\mathbf{r})\}$ should be an orthonormal basis set, which satisfies the periodic boundary conditions. Below we describe two basis sets, which are commonly used in computational condensed matter physics. Further in the text, the number of spatial orbitals per unit cell is denoted as M (the number of bands) while the total number of spatial orbitals in the crystal is $P = MN$.

1. Bloch Functions as a basis set

Bloch functions are the solution of a mean-field problem in a periodic potential. The Bloch functions can be written as follows [[78], p.167]:

$$\psi_{\mathbf{k}j}(\mathbf{r}) = e^{i\mathbf{k}\mathbf{r}} u_{\mathbf{k}j}(\mathbf{r}), \quad (20)$$

where $u_{\mathbf{k}}(\mathbf{r})$ is periodic function with periodicity of the unit cell, $u_{\mathbf{k}}(\mathbf{r}) = u_{\mathbf{k}}(\mathbf{r} + \mathbf{R})$, j is the band index, $0 \leq j \leq M$ and $\mathbf{k} = (k_1, k_2, k_3)^T$ is the wave vector belonging to the first Brillouin zone, which can be defined as

$$k_\alpha = \sum_{\beta=1,2,3} B_{\alpha\beta} \frac{g_\beta}{N_\beta}, \quad (21)$$

where g_β is an integer such that

$$-\frac{N_\alpha}{2} < g_\alpha \leq \frac{N_\alpha}{2}. \quad (22)$$

The larger the size of the macroscopic crystal, the larger the number of k points in the Brillouin zone as can be seen from (22). Using Bloch functions as the basis set

$$\hat{\psi}(\mathbf{x}) = \sum_{\mathbf{k}j} \psi_{\mathbf{k}j}(\mathbf{r}) \hat{c}_{\mathbf{k}j\sigma}, \quad (23)$$

the Hamiltonian can be written as

$$\hat{H} = H^{(0)} + \sum_{\mathbf{k}ij\sigma} \tilde{h}_{\mathbf{k}ij} \hat{c}_{\mathbf{k}i\sigma}^\dagger \hat{c}_{\mathbf{k}j\sigma} \\ + \frac{1}{2N} \sum_{\sigma,\sigma'} \sum_{ijlm} \sum_{\mathbf{k}\mathbf{q}\mathbf{k}'\mathbf{q}'} \left(\sum_{\mathbf{G}} \delta_{\mathbf{q}+\mathbf{q}'-\mathbf{k}-\mathbf{k}',\mathbf{G}} \right) \tilde{g}_{\mathbf{k}i,\mathbf{q}j,\mathbf{k}'l,\mathbf{q}'m} \\ \times \hat{c}_{\mathbf{k}i\sigma}^\dagger \hat{c}_{\mathbf{k}'l\sigma'}^\dagger \hat{c}_{\mathbf{q}'m\sigma'} \hat{c}_{\mathbf{q}j\sigma}, \quad (24)$$

where we used the convention that the Bloch functions are normalized in the unit cell. Due to $\sum_{\mathbf{G}} \delta_{\mathbf{q}+\mathbf{q}'-\mathbf{k}-\mathbf{k}',\mathbf{G}}$, the number of two-body terms scales as $O(N^3)$, the same as in the plane-wave basis set [24]. One- and two-body matrix elements are defined as

$$\tilde{h}_{\mathbf{k}ij} = \int_{V_{uc}} d^3\mathbf{r} \psi_{\mathbf{k}i}^*(\mathbf{r}) h(\mathbf{r}) \psi_{\mathbf{k}j}(\mathbf{r}) \quad (25)$$

and

$$\begin{aligned} \tilde{g}_{\mathbf{k}i,\mathbf{q}j,\mathbf{k}'l,\mathbf{q}'m} &= \int_{V_{uc}} d^3\mathbf{r} \psi_{\mathbf{k}i}^*(\mathbf{r}) \psi_{\mathbf{q}j}(\mathbf{r}) \\ &\times \int_{V_{uc}} d^3\mathbf{r}' \psi_{\mathbf{k}'l}^*(\mathbf{r}') \psi_{\mathbf{q}'m}(\mathbf{r}') g_{\mathbf{q}'-\mathbf{k}'}(\mathbf{r}-\mathbf{r}'), \end{aligned} \quad (26)$$

where

$$g_{\mathbf{q}-\mathbf{k}}(\mathbf{r}-\mathbf{r}') = \sum_{\mathbf{G}} \frac{4\pi}{(\mathbf{q}-\mathbf{k}+\mathbf{G})^2 V_{uc}} e^{i(\mathbf{q}-\mathbf{k}+\mathbf{G})(\mathbf{r}-\mathbf{r}')} \quad (27)$$

and V_{uc} is the volume of the unit cell. Coefficients in the Hamiltonian are usually complex and the two-body terms obey fourfold symmetry (18). The composite index from Sec. III A indicates both band and wave vector, $p = (\mathbf{k}, i)$, $q = (\mathbf{q}, j)$, $r = (\mathbf{k}', l)$, $s = (\mathbf{q}', m)$.

2. Wannier functions as a basis set

Wannier functions are the set of localized orbitals, which obey the translational symmetry of the crystal,

$$\phi_{\mathbf{R}j}(\mathbf{r}-\mathbf{R}') = \phi_{\mathbf{R}+\mathbf{R}'j}(\mathbf{r}). \quad (28)$$

Such localized orbitals can be obtained by carrying out a localization procedure in the supercell or applying a Fourier transformation to the Bloch orbitals [62],

$$\phi_{\mathbf{R}j}(\mathbf{r}) = \frac{1}{N} \sum_{\mathbf{k}} e^{-i\mathbf{k}\mathbf{R}} \sum_m \psi_{\mathbf{k}m}(\mathbf{r}) U_{mj}^{(\mathbf{k})}, \quad (29)$$

where unitary matrices can be chosen according to localization criteria such as, for example, Foster-Boys [79] (maximally localized Wannier orbitals [60]) or Pipek-Mezey [80–82]. Contrary to Bloch functions, these functions can be chosen to be real valued [62]. In the Wannier basis set

$$\hat{\psi}(\mathbf{x}) = \sum_{\mathbf{R}j} \phi_{\mathbf{R}j}(\mathbf{r}) \hat{a}_{\mathbf{R}j\sigma}, \quad (30)$$

the Hamiltonian is

$$\begin{aligned} \hat{H} &= H^{(0)} + \sum_{\mathbf{R}\mathbf{R}'ij\sigma} h_{\mathbf{R}p,\mathbf{R}'q} \hat{a}_{\mathbf{R}i\sigma}^\dagger \hat{a}_{\mathbf{R}'j\sigma} \\ &+ \frac{1}{2} \sum_{\sigma,\sigma'} \sum_{ijlm} \sum_{\mathbf{R}\mathbf{R}'\mathbf{T}\mathbf{T}'} g_{\mathbf{R}i,\mathbf{T}j,\mathbf{R}'l,\mathbf{T}'m} \hat{a}_{\mathbf{R}i\sigma}^\dagger \hat{a}_{\mathbf{R}'l\sigma'}^\dagger \hat{a}_{\mathbf{T}'m\sigma'} \hat{a}_{\mathbf{T}j\sigma} \end{aligned} \quad (31)$$

with matrix elements

$$h_{\mathbf{R}i,\mathbf{R}'j} = \int_V d^3\mathbf{r} \phi_{\mathbf{R}i}^*(\mathbf{r}) h(\mathbf{r}) \phi_{\mathbf{R}'j}(\mathbf{r}) \quad (32)$$

and

$$\begin{aligned} g_{\mathbf{R}i,\mathbf{T}j,\mathbf{R}'l,\mathbf{T}'m} &= \int_V d^3\mathbf{r} \phi_{\mathbf{R}i}^*(\mathbf{r}) \phi_{\mathbf{T}j}(\mathbf{r}) \\ &\times \int_V d^3\mathbf{r}' \phi_{\mathbf{R}'l}^*(\mathbf{r}') \phi_{\mathbf{T}'m}(\mathbf{r}') g(\mathbf{r}-\mathbf{r}'), \end{aligned} \quad (33)$$

which satisfy the following relations:

$$h_{\mathbf{R}i,\mathbf{R}'j} = h_{0i,\mathbf{R}'-\mathbf{R}j}, \quad (34)$$

$$v_{\mathbf{R}i\mathbf{T}j\mathbf{R}'l\mathbf{T}'m} = v_{0i,\mathbf{T}-\mathbf{R}j,\mathbf{R}'-\mathbf{R}l,\mathbf{T}'-\mathbf{R}m}. \quad (35)$$

This is due to the fact that Wannier functions obey Eq. (28). In this paper we do not construct Wannier functions from Bloch orbitals but rather choose natural atomic orbitals in the supercell calculations (see Sec. III C for computational details). Since the two-body term is real, it satisfies Eq. (19). The composite index from Sec. III A indicates both band and unit-cell indices, $p = (\mathbf{R}, j)$, $q = (\mathbf{T}, j)$, $r = (\mathbf{R}', l)$, $s = (\mathbf{T}', m)$.

3. Majorana representation

Majorana operators represent a convenient choice for working with quantum computing algorithms. The reason is that each Majorana operator is Hermitian and can be mapped onto one Pauli string using a qubit representation and, as a result, any unique product of Majorana operators is one Pauli string. The actual qubit representation depends on the choice of transformation such as the Jordan-Wigner [7,83] or Bravyi-Kitaev [11,84]. However, some properties of the Hamiltonian, which do not depend on the choice of qubit mapping can conveniently be obtained in Majorana representation. We will use this representation in order to generalize the sparse qubitization approach on Hamiltonians with complex coefficients. Majorana operators are defined as

$$\hat{\gamma}_{p\sigma,0} = \hat{a}_{p\sigma} + \hat{a}_{p\sigma}^\dagger, \quad (36)$$

$$\hat{\gamma}_{p\sigma,1} = -i(\hat{a}_{p\sigma} - \hat{a}_{p\sigma}^\dagger), \quad (37)$$

with an additional binary index specifying the Majorana type.

They satisfy the following anticommutation relations:

$$\{\hat{\gamma}_{p\sigma,k}, \hat{\gamma}_{q\tau,j}\} = 2\delta_{pq}\delta_{\sigma\tau}\delta_{kj}. \quad (38)$$

Following Refs. [32,85], the constant, one-body and two-body terms of the Hamiltonian (16) in Majorana representation can be written as follows:

$$\tilde{H}_0 = H_0 + \sum_p h_{pp} + \frac{1}{2} \sum_{pq} \left(g_{ppqq} - \frac{1}{2} g_{ppqp} \right), \quad (39)$$

$$\hat{H}^{(1)} = \frac{i}{2} \sum_{p \leq q} \text{Re} \left(h_{pq} + \frac{1}{2} \sum_r (2g_{pqrr} - g_{prrq}) \right) \left(\frac{1}{2} \right)^{\delta_{pq}} \sum_{\sigma:j \neq k} (-1)^j \hat{\gamma}_{p\sigma,j} \hat{\gamma}_{q\sigma,k} \quad (40)$$

$$+ \frac{i}{2} \sum_{p < q} \text{Im} \left(h_{pq} + \frac{1}{2} \sum_r (2g_{pqrr} - g_{prrq}) \right) \sum_{\sigma:j} \hat{\gamma}_{p\sigma,j} \hat{\gamma}_{q\sigma,j}, \quad (41)$$

$$\hat{H}^{(2)} = \frac{1}{4} \sum_{\substack{p \leq q, r \leq s \\ (p,q) \leq (r,s)}} B_{pqrs} \left(\frac{1}{2}\right)^{\delta_{pq} + \delta_{rs} + \delta_{(p,q),(r,s)}} \sum_{\sigma \neq \tau, j \neq k, l \neq m} \hat{\gamma}_{p\sigma,j} \hat{\gamma}_{q\sigma,k} \hat{\gamma}_{r\tau,l} \hat{\gamma}_{s\tau,m} (-1)^{j+m} \tag{42}$$

$$+ \frac{1}{4} \sum_{\substack{p < q, r < s \\ (p,q) \leq (r,s)}} F_{pqrs} \left(\frac{1}{2}\right)^{\delta_{(p,q),(r,s)}} \sum_{\sigma \neq \tau; j,k} \hat{\gamma}_{p\sigma,j} \hat{\gamma}_{q\sigma,j} \hat{\gamma}_{r\tau,k} \hat{\gamma}_{s\tau,k} (-1)^{j+k} \tag{43}$$

$$+ \frac{1}{4} \sum_{p \leq q, r < s} (-1) C_{pqrs} \left(\frac{1}{2}\right)^{\delta_{pq}} \sum_{\sigma \neq \tau; j \neq k; l} \hat{\gamma}_{p\sigma,j} \hat{\gamma}_{q\sigma,k} \hat{\gamma}_{r\tau,l} \hat{\gamma}_{s\tau,l} (-1)^j \tag{44}$$

$$+ \frac{1}{4} \sum_{\substack{p < q, r < s \\ (p,q) \leq (r,s)}} (F_{pqrs} + B_{prqs} - B_{psqr}) \left(\frac{1}{2}\right)^{\delta_{(p,q),(r,s)}} \sum_{\sigma, j \neq k} \hat{\gamma}_{p\sigma,j} \hat{\gamma}_{q\sigma,j} \hat{\gamma}_{r\sigma,k} \hat{\gamma}_{s\sigma,k} \tag{45}$$

$$+ \frac{1}{4} \sum_{p < q < r < s} (F_{pqrs} - F_{prqs} + F_{psqr}) \sum_{\sigma, j} \hat{\gamma}_{p\sigma,j} \hat{\gamma}_{q\sigma,j} \hat{\gamma}_{r\sigma,j} \hat{\gamma}_{s\sigma,j} \tag{46}$$

$$+ \frac{1}{4} \sum_{p, q < r < s} (C_{pqrs} - C_{prqs} + C_{psqr}) \sum_{\sigma, j \neq k} \hat{\gamma}_{p\sigma,j} \hat{\gamma}_{q\sigma,k} \hat{\gamma}_{r\sigma,k} \hat{\gamma}_{s\sigma,k} (-1)^j, \tag{47}$$

where the tensors

$$B_{pqrs} = \frac{\text{Re}(g_{pqrs} + g_{pqsr})}{2} = \frac{g_{pqrs} + g_{qpsr} + g_{pqsr} + g_{qprs}}{4}, \tag{48}$$

$$C_{pqrs} = \frac{\text{Im}(g_{pqrs} - g_{pqsr})}{2} = \frac{g_{pqrs} - g_{qpsr} - g_{pqsr} + g_{qprs}}{4i}, \tag{49}$$

$$F_{pqrs} = \frac{\text{Re}(g_{pqrs} - g_{pqsr})}{2} = \frac{g_{pqrs} + g_{qpsr} - g_{pqsr} - g_{qprs}}{4}. \tag{50}$$

B_{pqrs} is symmetric with respect to (p, q) and symmetric with respect to (r, s) , C_{pqrs} is symmetric with respect to (p, q) and antisymmetric with respect to (r, s) , F_{pqrs} is antisymmetric with respect to (p, q) and antisymmetric with respect to (r, s) , B and F are symmetric with respect to interchange of pairs (p, q) , (r, s) , while C_{pqrs} is not.

The reverse relation is

$$g_{pqrs} = B_{pqrs} + F_{pqrs} + i(C_{rspq} + C_{pqrs}). \tag{51}$$

This representation of the Hamiltonian is valid for both Bloch and Wannier orbitals and the only differences are indices labeling states and the value of coefficients. For real-valued Wannier functions, only the B_{pqrs} tensor is nonzero while for complex-valued orbitals, such as Bloch functions, all tensors need to be taken into account as coefficients of the Hamiltonian in Eq. (16) can be complex. The Hamiltonian in Majorana representation (39)–(47) provides a decomposition in a linear combination of unitaries (LCU [86]) and the entry-wise L_1 norm of such a Hamiltonian, which is the same in any qubit representation [85], can be written as

$$\alpha = \lambda_0 + \lambda = \lambda_0 + \lambda_1 + \lambda_2 \tag{52}$$

where

$$\lambda_0 = \left| H_0 + \sum_p h_{pp} + \frac{1}{4} \sum_{pr} (2g_{pprr} - g_{prrp}) \right|, \tag{53}$$

$$\lambda_1 = \sum_{pq} \left| \text{Re} \left(h_{pq} + \frac{1}{2} \sum_r (2g_{pqrr} - g_{prrq}) \right) \right| + 2 \sum_{p < q} \left| \text{Im} \left(h_{pq} + \frac{1}{2} \sum_r (2g_{pqrr} - g_{prrq}) \right) \right|, \tag{54}$$

$$\lambda_2 = \frac{1}{4} \sum_{pqrs} |B_{pqrs}| + \sum_{p < q, r < s} |F_{pqrs}| + \sum_{p, q, r < s} |C_{pqrs}| + \frac{1}{2} \sum_{p < q, r < s} |F_{pqrs} + B_{prqs} - B_{psqr}| + \tag{55}$$

$$+ \sum_{p < q < r < s} |F_{pqrs} - F_{prqs} + F_{psqr}| + \sum_{p, q < r < s} |C_{pqrs} - C_{prqs} + C_{psqr}|. \tag{56}$$

In the context of quantum computation, the magnitude of $\lambda = \lambda_1 + \lambda_2$ defines the number of repetitions of controlled-unitary in QPE as discussed below.

B. Quantum algorithms

QPE [52,54] allows to determine the phase of a unitary operator, which is implemented in the quantum circuit. Originally, Hamiltonian simulation with the choice of unitary

$$U = e^{iHt} \tag{57}$$

was used, as the time evolution operator can be implemented with Trotterization [87,88] and its phases $E_j t$ are directly related to the system's energies E_j . Other methods for Hamiltonian simulation have since been developed, including Taylor series [20,23,24] and randomized methods [89–91].

Instead of the time evolution operator (57), one can apply QPE to a unitary *walk operator* W with eigenvalues $e^{\pm i \arccos(E_j/\lambda)}$, from whose phases the Hamiltonian energies can readily be retrieved. This is the approach of more recent qubitization methods [25,26]. The normalization factor λ is the L_1 norm of Hamiltonian's coefficients in an LCU decomposition, and ensures that all energies correspond to phases between 0 and 2π . The walk operator is constructed from a reflection along with a *block encoding* of H/λ . The starting point for a block encoding is an LCU decomposition of the Hamiltonian

$$H = \sum_{i=1}^d w_i H_i, \quad w_i \in \mathbb{R}^+ \quad (58)$$

into simpler unitary operators H_i that can be readily implemented on a quantum computer. In our case, the LCU decomposition of the Hamiltonian is given in Sec. III A 3, and the H_i consist of strings of up to four Majorana fermions that can be implemented in the Jordan–Wigner representation with a ranged operation [27]. The two operators

$$\begin{aligned} \text{SELECT } |i\rangle|\psi\rangle &= |i\rangle H_i |\psi\rangle \text{ and PREPARE } |0\rangle \\ &= \sum_{i=1}^d \sqrt{\frac{w_i}{\lambda}} |i\rangle \text{ with } \lambda = \sum_{i=1}^d |w_i| \end{aligned} \quad (59)$$

facilitate a block encoding of H/λ because

$$\langle 0 | (\text{UNPREPARE} \otimes \mathbb{1}) \text{SELECT} (\text{PREPARE} \otimes \mathbb{1}) | 0 \rangle = H/\lambda \quad (60)$$

with $\text{UNPREPARE} = \text{PREPARE}^{-1}$.

The leading order cost of qubitization algorithms is typically proportional to $O(\lambda\sqrt{d}/\epsilon)$. A multiplicative λ is needed to maintain a fixed energy accuracy ϵ in phase estimation despite the normalization of the Hamiltonian by λ . Meanwhile, the \sqrt{d} stems from data loading. The values of the d coefficients of the LCU w_i can be loaded with Toffoli cost of order $O(\sqrt{d})$ with a select-swap network [92] (also dubbed QROAM [28]) at the expense of $O(\sqrt{d})$ ancilla qubits.

Various implementations of the qubitization walk operator [27,28,31,32] factorize the Hamiltonian to find alternative LCUs with lower d and/or λ . They also show bespoke implementations of PREPARE and SELECT matching the choice of LCU. We base our method on the sparse qubitization method [28,31], which does not factorize the Hamiltonian, but instead exploits sparsity. In the next section, we will give an overview of that method. Then we will describe how we have adapted and extended the method to deal with Bloch and Wannier basis sets. We focus on the asymptotically dominant costs and confine detailed Toffoli and qubit number costings to Appendix. The number of T gates is four times the number of Toffoli gates [27,93].

1. Overview of sparse qubitization

In this section we outline the original sparse qubitization algorithm, which has been used for simulation of real Hamiltonians, which satisfy eightfold symmetry relations (19). It was first developed in [28] and improved in [31], Appendix, on which we base our exposition. Our modifications to this

algorithm are presented in the Secs. III B 2 and III B 3. For simplicity of this summary, we focus on the Hamiltonian's two-body terms, of which there are significantly more than one-body terms. The main insight the sparse qubitization algorithm [28,31] uses is that the Hamiltonian's LCU decomposition into unitaries Q ,

$$H = \frac{1}{8} \sum_{pqrs,\sigma,\tau} V_{pqrs} Q_{pq\sigma} Q_{rs\tau} \quad (61)$$

is very sparse with respect to the orbital indices (p, q, r, s), and spin indices (σ, τ), especially if small coefficients V_{pqrs} are approximated to zero. Thus one can save on quantum resources required for the data loading: Instead of loading the coefficients for all values of p, q, r, s, σ, τ , only a unique set of nonzero coefficients is loaded onto the quantum computer and the rest of the Hamiltonian can be restored using symmetry restoration circuits. Let these nonzero coefficients be indexed by $l = 1 \dots d$ in an arbitrary way, such that we must load only d data items. While the scaling of d with the total number of spatial orbitals P is still expected to be the same as the full number of electron repulsion integrals,

$$d \sim O(P^4), \quad (62)$$

one can truncate small coefficients and reduce d . Each data item to be loaded consists of the value of the coefficient V_{pqrs} as well as the corresponding indices p, q, r, s, σ, τ that allow to apply the correct unitary $Q_{pq\sigma} Q_{rs\tau}$. QROAM allows to load these as qubit bitstrings. However, PREPARE requires the coefficient values as amplitudes of the state, not bitstrings. This gap is bridged by instead loading so-called “keep-probabilities” and performing coherent alias sampling [27]. Figure 2 shows a sketch of the PREPARE operator.

The number of data items d to load can be reduced further by leveraging symmetries of the Hamiltonian that cause multiple identical coefficients. First of all, coefficient values are independent of spin. Thus we must only load one coefficient, and the other identical terms can be restored in the quantum circuit. Likewise, for a given permutation of orbital indices, eight coefficients, which possess eightfold symmetries [see Eqs. (19)] can also be restored in the quantum circuit using only one set of orbital indices. This reduces the number of terms to be loaded d by approximately a factor of 8.

The PREPARE operator is implemented in the following steps illustrated in Fig. 2 (see [28,31] for details):

a. Equal superposition state. Prepare $\frac{1}{\sqrt{d}} \sum_{l=0}^{d-1} |l\rangle$, where d is the number of nonzero LCU terms (up to eightfold and spin symmetries). This uses ancillas for amplitude amplification not shown in the figure.

b. Data loading. A QROAM loads data of width m qubits. In principle, these m qubits include the value of the coefficient indices $p(l), q(l), r(l), s(l)$ along with the value $w_l = |V_{p(l)q(l)r(l)s(l)}|/8$ and its sign θ .

However, in practice, in order to perform coherent alias sampling, slightly different data items must be loaded [27]. Instead of the coefficient value V_{pqrs} , a data field of \aleph qubits (so-called “keep-probability”) is needed. \aleph determines the accuracy with which the coefficients V_{pqrs} are ultimately loaded and can be computed with (A1). Further, coherent alias sampling requires two values of the other data to be loaded

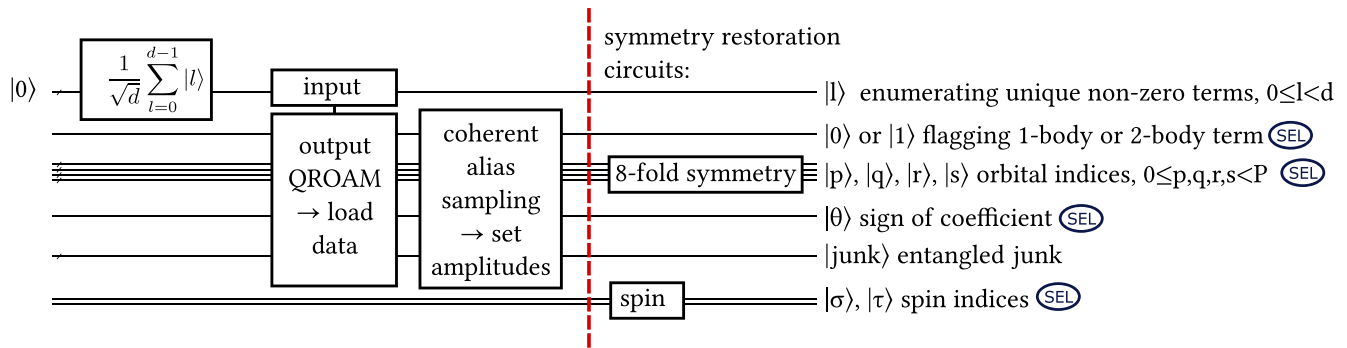


FIG. 2. PREPARE operator for the original sparse qubitization method [28,31], outlined in Sec. III B 1. It prepares a superposition of product states describing the Hamiltonian’s LCU: The amplitudes in the superposition are the (square root modulus of the) LCU’s coefficients $\propto V_{pqrs}$, whilst the corresponding product state specifies the unitary operator $Q_{pq\sigma} Q_{rs\tau}$. The state is entangled to an additional junk register, such as the additional qubits needed for coherent alias sampling. Only the qubits indicated with (SEL) are used as controls by the SELECT operator. At the vertical dashed line (red), the prepared state describes just one symmetry sector with respect to spin symmetry and eightfold symmetry. This reduces the amount of data to be loaded, which is the most expensive step in the circuit. The superposition is then enlarged to more product states covering all symmetry sectors with the symmetry restoration circuits. Note that ancilla qubits (such as in the QROAM or eightfold symmetry restoration circuit) are not depicted. Despite becoming entangled with the other qubits, like |junk>, they do not affect the rest of the circuit.

(indices p, q, r, s and θ and a qubit not mentioned here to distinguish between one- and two-body terms) [27]. Thus

$$m = \aleph + 2(4\lceil \log P \rceil + 2), \quad (63)$$

where P is the number of spatial orbitals ($2P$ is the number of spin orbitals).

The QROAM is the asymptotically most expensive step with Toffoli cost

$$\lceil d/\kappa \rceil + m(\kappa - 1). \quad (64)$$

Adjusting κ (which must be a power of 2) leads to a tradeoff between Toffoli cost and ancilla qubit count [92]

$$m\kappa + \lceil \log(d/\kappa) \rceil. \quad (65)$$

Choosing $\kappa \sim \sqrt{d}$ to optimize Toffoli cost, both Toffoli and ancilla count of the data lookup asymptotically follow (dropping logarithmic factors)

$$\sim O(\sqrt{d}) \sim O(P^2). \quad (66)$$

While the QROAM lookup will also have to be uncomputed in UNPREPARE, the cost is lower because it does not depend on the size m of the data items when using a measurement-based uncomputation scheme [28].

Coherent alias sampling. From the information thus loaded, coherent alias sampling then creates the state

$$\sum_{l=0}^{d-1} \sqrt{\frac{w_l}{\lambda}} |l\rangle |0/1\rangle |p\rangle |q\rangle |r\rangle |s\rangle |\theta\rangle |\text{junk}_l\rangle, \quad (67)$$

which is entangled to some $|\text{junk}_l\rangle$ that is not relevant. The second register, 0 or 1, flags one- or two-body terms, respectively.

Symmetry restoration. Now, the spin symmetry and eightfold symmetry must be restored. Two new qubits encoding spin $|\sigma\rangle$ and $|\tau\rangle$ in the state $(|0\rangle + |1\rangle)/\sqrt{2}$ are added as further tensor product factors. When the tensor product is expanded, it quadruples the number of states in the superposition (67).

To restore eightfold symmetry, similarly three qubits $(|0\rangle + |1\rangle)/\sqrt{2}$ for each of the symmetries $p \leftrightarrow q, r \leftrightarrow s, (p, q) \leftrightarrow (r, s)$ are added as tensor product factors. Swaps controlled on these qubits then swap p, q, r, s registers depending on the symmetry.

The result is a state describing the full LCU in (61). A subtlety of symmetry restorations is that slightly different values of w_i must be loaded by the QROAM, because the symmetry restoration accumulates factors of $1/\sqrt{2}$. Yet, this does not affect the overall subnormalization λ of the Hamiltonian.

Next, a SELECT operator selects the correct unitary for the $p, q, r, s, \sigma,$ and τ indices. The UNPREPARE operator uncomputes PREPARE. Using measurement based uncomputation, this is much more efficient than PREPARE [28]. The total leading order Toffoli cost is

$$\left\lceil \frac{\pi \lambda}{2\epsilon_{\text{QPE}}} \right\rceil \cdot \left(\left\lceil \frac{d}{\kappa} \right\rceil + m(\kappa - 1) \right), \quad (68)$$

the product of the QROAM cost for data loading (64) with the number of iterations of the walk operator $\lceil \pi \lambda / (2\epsilon_{\text{QPE}}) \rceil$. The normalization factor λ together with the desired accuracy ϵ_{QPE} determine the number of iterations of the walk operator required for phase estimation.

2. Generalization of sparse qubitization for Bloch basis functions

While the original sparse qubitization method supports Hamiltonians with real electron repulsion integrals only, Bloch orbitals usually lead to complex coefficients. We generalise the sparse method to complex Hamiltonians by expanding the Hamiltonian in Majorana strings (Sec. III A 3) and instead of working with eightfold symmetry restoration circuits, we introduce Majorana type restoration circuits. For a real Hamiltonian the expansion (40)–(47) only contains the terms (40), (42), and (45). The other terms arise for complex Hamiltonians. The coefficients of the Majorana strings are all real (or purely imaginary for the one-body terms) due to the Hamiltonian’s Hermiticity.

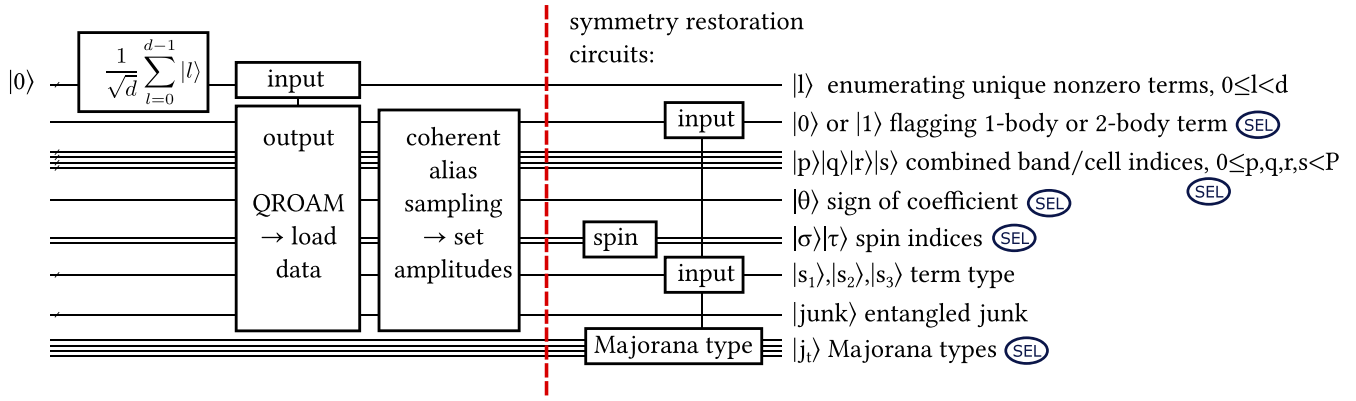


FIG. 3. PREPARE operator for the Bloch basis algorithm. In addition to the PREPARE operator of the original sparse qubitization (see Fig. 2), it has Majorana type indices that are used by SELECT to apply the correct Majorana types for the Majorana string described by a product state. $j_i = 0, 1$ and $t = 1, 2, 3, 4$ enumerates the positions of the Majorana operators in a Majorana string. The circuit for Majorana type symmetry restoration is given in the circuit (71); its input qubits must be loaded by QROAM. The spin symmetry restoration circuit is given in circuit (75). The Hamiltonian has been written such that no fourfold or eightfold symmetry is present any more (see Sec. III A 3). To simplify the sketch we have depicted that both spin qubits are loaded by QROAM [see the circuit (76) for details].

We use a SELECT operator [27,32] that allows to select Majorana strings based on: the indices p, q, r, s ; spin indices σ, τ ; and four Majorana type indices j_1 through j_4 —as they appear in the Hamiltonian (Sec. III A 3). In addition, the correct sign of the LCU coefficient is selected based on a $|\theta\rangle$ qubit. In Fig. 3, the control qubits for SELECT are indicated by **SEL**. The indices p, q, r, s index the Bloch basis functions (20), and as such they are composite indices, each consisting of band index and k -wave-vector index. For the Bloch basis we do not need to split up the composite indices and arbitrarily enumerate them as

$$p, q, r, s \in \{0, \dots, P - 1\}. \tag{69}$$

A main benefit of using Bloch basis functions even in classical methods is that momentum conservation causes many terms to be zero [see (20)]. Therefore we can expect the number of nonzero terms to scale as

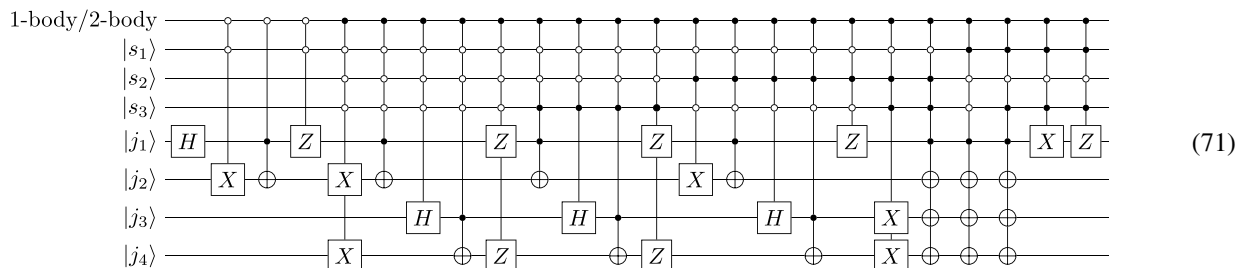
$$d \sim O(M^4 N^3), \tag{70}$$

for an LCAO basis set, while for PW basis sets M^4 can be reduced to M^3 . Since the number of bands M is defined per unit cell, the scaling of the algorithm with the system size N is cubic.

The PREPARE operator is sketched in Fig. 3. The electron repulsion integrals in the original sparse method possess eightfold symmetry in the p, q, r, s indices, and this is restored in PREPARE with controlled SWAPS (see Sec. III B 1).

In our case, at first, the electron repulsion integrals merely have fourfold symmetry (18) because the basis functions are complex. Once the Hamiltonian is expanded in Majorana strings (Sec. III A 3), this results in different types of symmetry for different coefficients as explained in Sec. III A 3. Instead of restoring these symmetries with controlled SWAPS, we rewrite the LCU decomposition of the Hamiltonian such that the symmetries are not explicitly present anymore, see Sec. III A 3. The sums can be restricted to one branch of the symmetry by instead summing over the Majorana type. For example, the coefficient B_{pqrs} in (42) has eightfold symmetry. Yet the sum is restricted to $p \leq q, r \leq s, (p, q) \leq (r, s)$, such that only one branch of the symmetry is present in the LCU, and it does not have repeated coefficients for eightfold permutations of p, q, r, s . The “missing” terms are compensated by summing over multiple values of the Majorana type indices j_1, \dots, j_4 . The resulting symmetry in the Majorana type can then be more easily restored, because it does not involve CSWAPS of multi-qubit p, q, r, s registers. Further, even if we did not restrict the sums in this fashion, some Majorana type symmetry would be still present and have to be restored anyway due to the complex nature of the Hamiltonian.

The LCU has multiple repeated coefficients for different values of j_1, \dots, j_4 of the Majorana type indices, up to signs like $(-1)^{j_1+j_4}$. Rather than repeatedly loading the same coefficient with a different sign multiple times, we restore the Majorana type in PREPARE with the following symmetry restoration circuit:



The circuit needs three input qubits $|s_1\rangle|s_2\rangle|s_3\rangle$ along with the qubit flagging one-body or two-body terms that together distinguish all of the types of terms and thereby Majorana type symmetry in the LCU. The four output qubits $|j_i\rangle$ are initialized as $|0\rangle$ and will indicate Majorana type to be used by SELECT (see Fig. 3). The initial values of the $|s_i\rangle$ qubits must be loaded from QROAM [94]. They identify terms in the Hamiltonian as follows:

term	$ s_1\rangle$	$ s_2\rangle$	$ s_3\rangle$
1-body Re	$ 0\rangle$	$ 0\rangle$	$ 0\rangle$
1-body Im	$ 1\rangle$	$ 0\rangle$	$ 0\rangle$
B	$ 0\rangle$	$ 0\rangle$	$ 0\rangle$
F	$ 0\rangle$	$ 0\rangle$	$ 1\rangle$
C	$ 0\rangle$	$ 1\rangle$	$ 0\rangle$
$F - B + B$	$ 0\rangle$	$ 1\rangle$	$ 1\rangle$
$F - F + F$	$ 1\rangle$	$ 0\rangle$	$ 0\rangle$
$C - C + C$	$ 1\rangle$	$ 0\rangle$	$ 1\rangle$

Let us give an example of Majorana type symmetry restoration for a term of type B , i.e., (42). The table above shows that the qubits are loaded such that at the vertical dashed line (red) in Fig. 3 the input qubits are

$$|1 \text{ or } 2 \text{ body}\rangle|s_1\rangle|s_2\rangle|s_3\rangle = |1\rangle|0\rangle|0\rangle|0\rangle. \quad (73)$$

The symmetry restoration circuit (71) acting on these then results in the state

$$|1000\rangle \otimes \frac{1}{2}(-|0101\rangle + |1001\rangle + |0110\rangle - |1010\rangle) \\ = |1000\rangle \otimes \frac{1}{2} \sum_{j_1 \neq j_2, j_3 \neq j_4} |j_1 j_2 j_3 j_4\rangle (-1)^{j_1 + j_4}. \quad (74)$$

This selects the correct Majorana type indices and signs for the B -type term in (42). Note that when inverting the circuit for UNPREPARE, one must omit the CZs. Otherwise the -1 in the Majorana symmetries of the coefficients [like in (74)] introduced in both PREPARE and UNPREPARE would cancel. The cost of this circuit is subleading, except to the extent that the necessary data loading increases m . Its cost can be reduced by using a unary iteration over the $|s_i\rangle$ qubits.

As in the original sparse method, the spin symmetry is also restored on the quantum computer. Identical coefficients corresponding to different spin configurations must only be loaded once. Here we have two spin qubits $|\sigma\rangle$ and $|\tau\rangle$ for the first two and second two Majoranas, respectively. Because we have more possible combinations than in the original sparse method, we need a short spin symmetry restoration circuit



Similarly to Majorana restoration, $|\sigma\rangle$ is initialized as $|0\rangle$ and different initial values for $|\tau\rangle$ must be loaded for different coefficient types,

$$|0\rangle \text{ for 1-body terms and 2-body terms with } \sum_{\sigma} \text{ and} \quad (76)$$

$$|1\rangle \text{ for 2-body terms with } \sum_{\sigma \neq \tau}.$$

Note that (contrary to the simplified Fig. 3), the qubit $|\sigma\rangle$ is not loaded from QROAM but initialized as zero.

As in the original sparse qubitization algorithm, the dominant cost is (68), the product of the number of iterations in phase estimation and the QROAM cost of data loading. Here, the parameters are as follows:

- (i) ϵ_{QPE} is the error budget for the phase estimation.
- (ii) λ is the normalization factor of the Hamiltonian, i.e., the L_1 norm of the LCU (Sec. III A 3). It depends on the specific material under consideration.
- (iii) d is the number of nonzero coefficients of the LCU to load from QROAM. (Up to spin symmetry and Majorana type symmetry, which are restored with circuits as discussed above.)
- (iv) m is the size of each of the d data items that need to be loaded. Specifically, we have

$$m = \aleph + 2(4\lceil \log(P) \rceil + 6). \quad (77)$$

The values of the coefficients V_{pqrs} are effectively encoded in the \aleph qubits and restored with coherent alias sampling [27]. The other qubits correspond to the necessary indices and further qubits to be loaded. For technical reasons of coherent alias sampling, they must be loaded with two values each, explaining the factor of 2 in above equation. First, $4 \log(P)$ is the total width of the four basis function indices p, q, r, s . A small number of qubits are needed for Majorana type restoration, spin symmetry restoration, distinguishing one- and two-body terms, and the sign of the coefficient.

(v) κ is a power of 2, and can be tuned to achieve a tradeoff between Toffoli and ancilla count (65). We choose it such that Toffoli count is minimized. Then the overall cost (dropping logarithmic factors) is

$$\sim O(\lambda \sqrt{d} / \epsilon_{\text{QPE}}) \sim O(\lambda M^2 N^{3/2} / \epsilon_{\text{QPE}}). \quad (78)$$

3. Generalization of sparse qubitization for Wannier basis functions

The algorithm closely follows the one for Bloch states. However, since Wannier functions can be chosen to be real, only the terms (40), (42), and (45) in LCU expansion from Sec. III A 3 are nonzero. Additionally, these terms possess the same translational symmetry as ERIs (35). This is taken into account through a translational symmetry restoration circuit, which further reduces the cost of quantum computation. We sketch the PREPARE operator in Fig. 4.

Translational symmetry can be leveraged to avoid loading repeated coefficient values. In order to restore it with a symmetry restoration circuit, the compound indices p, q, r, s must be split into an orbital index i_t and cell index $\mathbf{n}_t = (n_{t,1}, n_{t,2}, n_{t,3})^T$ as indicated in Fig. 4, $t = 1, 2, 3, 4$ enumerates the positions of the Majorana operators in a Majorana string. The figure also shows that SELECT must now be controlled on all qubits constituting the compound indices. While $i_t \in \{0, \dots, M - 1\}$ for the number of orbitals per cell M , the cell index $n_{t,i} \in \{0, \dots, N_i - 1\}$ for each spatial direction. Even though the total number of nonzero terms is larger than in the Bloch basis set, the translational symmetry reduces the number of unique nonzero terms d to the same asymptotic

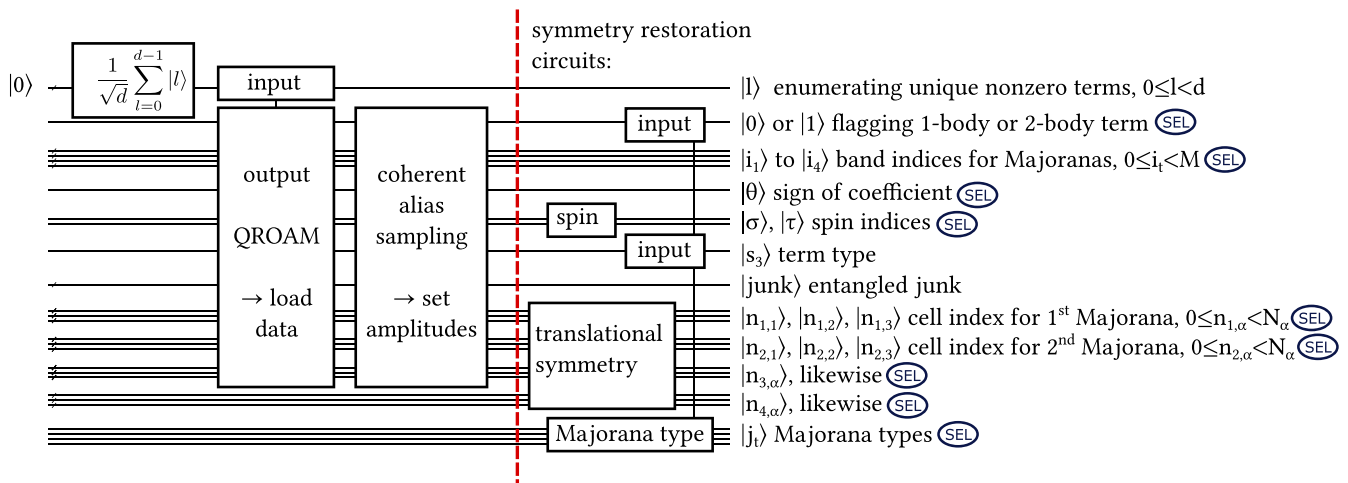


FIG. 4. PREPARE operator for the Wannier basis algorithm. We make use of translational symmetry of the LCU’s coefficients (80) to reduce data loading. This necessitates splitting the combined indices p, q, r, s into a band index i_t and a cell index $\mathbf{n}_t = (n_{t,1}, n_{t,2}, n_{t,3})$. $t = 1, 2, 3, 4$ enumerates the positions of the Majorana operators in a Majorana string. $j_t = 0, 1$. The translational symmetry restoration circuit is given in the circuit (81).

scaling

$$d \sim O(M^4 N^3) \tag{79}$$

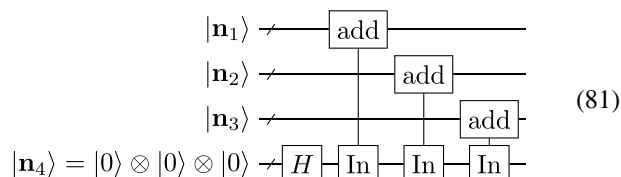
as in the Bloch basis (70).

The spin restoration circuits work identically to circuits used for the Bloch representation, while the Majorana type symmetry restoration circuit (71) can be simplified: Since the only terms appearing are the one-body term Eq. (40), and two two-body terms Eq. (42) and Eq. (45), we can remove all of the gates from (71) that are controlled for other contribution in the Hamiltonian. Further, we can remove the qubits $|s_1\rangle|s_2\rangle$ keeping only $|s_3\rangle$.

The translational symmetry of the LCU coefficients B from Eq. (42) can be expressed as

$$\begin{aligned} B_{i_1 \mathbf{n}_1, i_2 \mathbf{n}_2, i_3 \mathbf{n}_3, i_4 \mathbf{n}_4} &= B_{i_1 (\mathbf{n}_1 + \mathbf{v}), i_2 (\mathbf{n}_2 + \mathbf{v}), i_3 (\mathbf{n}_3 + \mathbf{v}), i_4 (\mathbf{n}_4 + \mathbf{v})} \\ &= B_{i_1 (\mathbf{n}_1 - \mathbf{n}_4), i_2 (\mathbf{n}_2 - \mathbf{n}_4), i_3 (\mathbf{n}_3 - \mathbf{n}_4), i_4 \mathbf{0}}, \quad \forall \mathbf{v} \in \tilde{\mathbb{Z}} \end{aligned} \tag{80}$$

with additions taken mod $(N_1, N_2, N_3)^T$ independently for each spatial component of \mathbf{n} . (The same holds for one-body and the other two-body coefficients.) We reduce the data loading by only loading one of these coefficients, and restoring the others with the following translational symmetry restoration circuit [95]:



The in-place additions [93] are to be performed separately for the three spatial directions of \mathbf{n} . In this circuit, \mathbf{n}_4 is an ancilla and the index must not be loaded from QROAM, leading to a reduction in m compared to the Bloch basis.

The dominant cost of the algorithm in the Wannier basis has the same formula (68) and scaling (78) as in the Bloch basis. Yet, m is lower in the Wannier basis. This reduces the dominant Toffoli (64) and qubit costs (65), which stem from data loading. Further, d is expected to be lower (if only by a factor) because of the two-body terms, only Eq. (42) and Eq. (45) are nonzero for real orbitals, and, as we will find in Sec. III C 1, more small coefficients can be truncated. The dependence of the normalization factor, λ , which has a large effect on the dominant cost, on the choice of the basis set is discussed in Sec. IV.

4. Error correction scheme

The details of the error correction implementation we used in this paper can be found in Sec. 4 of Ref. [55], but for completeness a short description is also provided here.

We estimate the error correction overheads following Litinski’s game of surface codes scheme [65]. The computational qubits are arranged in *fast blocks* [65], that enable the consumption of one magic state per time step—the clock rate of an error corrected quantum processor. We select magic state factories from a subset of the ones described in [65,96]. We allow ourselves as many magic state factories as necessary, so that computational qubits do not idle between magic state injections. Magic state factories are arranged around the computational qubits, accounting for the routing space necessary to move the distilled T states to the computational area.

The code distance d and magic state factory types are determined by allocating a failure rate budget of 1% to error correction; this budget is divided between magic state distillation failure (0.1%) and probability of logical circuit failure (0.9%). Logical circuit failure is estimated using the conventional formula for error rate in the surface code [56],

$$P_{\text{logical}} = 0.1 \times (100 p)^{\frac{d+1}{2}}, \tag{82}$$

TABLE I. The magnitude of the translational symmetry error, ϵ_t , in one-body terms arising from generation of localized orbitals in the supercell calculations. $\epsilon_t = \max_{i,j;\mathbf{R}\mathbf{R}'} |h_{\mathbf{R}i,\mathbf{R}'j} - h_{0i,\mathbf{R}-\mathbf{R}j}|$.

System	Error (Ha)
H (16 atoms)	1.53×10^{-11}
H (54 atoms)	1.70×10^{-6}
H (128 atoms)	3.53×10^{-10}
H (250 atoms)	2.43×10^{-10}
LiH (64 atoms)	3.08×10^{-9}
NiO (64 atoms)	9.76×10^{-8}
PdO (16 atoms)	5.11×10^{-11}
PdO (72 atoms)	2.41×10^{-5}

where p is the probability of physical error, that we take to be 10^{-3} or 10^{-4} . The former is an accepted rate for superconducting systems, while the latter is considered optimistic in these platforms.

To obtain runtime from T-gate count, we assume serial implementation and multiply the number of T gates by the duration of the time step, which in the surface code is d times the duration of the code cycle. We take the code cycle (the time needed to carry out one round of stabiliser measurements) to be $1\mu\text{s}$ long—this is a conventional figure for superconducting qubits [97,98].

C. Classical computational details

In order to generate ERIs, we have carried out restricted Hartree-Fock calculations in the Gaussian basis set using the PySCF software [99–102]. The Fermi-Dirac distribution of occupation numbers has been used. In order to numerically study the scaling properties of the algorithms, we carry out calculations on a model hydrogen (H) crystal in body-centre cubic (BCC) structure. Quantum resource estimations have also been applied to realistic solids such as lithium hydride (LiH), NiO, and PdO. In the case of H and LiH, all-electron calculations have been carried out with STO-3G [103–107] basis set, while for NiO and PdO the GTH pseudopotentials [108,109] with GTH-SZV basis set have been used [110,111]. Gaussian density fitting has been employed for the calculation of ERIs [112]. The pseudopotential setups with 18 valence electrons for Ni and Pd, and six valence electrons for O have been used. The geometry of H has been optimized with GPAW software [113–115], the geometry of PdO has been optimized with Quantum Espresso [116,117], while experimental structures have been used for LiH [[118], p.7] and NiO [119]. We provide geometry files on Zenodo [120]. In order to obtain spatially localized orbitals satisfying translational symmetry (28), the natural atomic orbitals [121] in a supercell have been constructed. We have then validated that the translational symmetry [(34), (35)] of the integrals is satisfied (see Table I). Calculations in the supercell have been carried out using cubic cells of different sizes for H, LiH, and NiO, while a tetragonal cell has been used for PdO. For calculations using Bloch functions, we used primitive 2-atom cells for LiH and NiO with varying k -point meshes (up to a $3\times 3\times 3$ mesh). The unit cell used for H consists of two atoms while the tetragonal unit cell of PdO has been chosen such that it contains four atoms.

I. Errors and truncation strategies

The total additive error of the energy estimation can be split into three errors [31],

$$\epsilon = \epsilon_{\text{QPE}} + \epsilon_{\text{trunc}} + \epsilon_{\text{prep}}, \quad (83)$$

where ϵ_{QPE} is the error due to limited precision of the QPE, ϵ_{trunc} is the error due to discarding Hamiltonian coefficients with a certain precision, and ϵ_{prep} is the error due to finite precision of state preparation amplitudes. Given the accuracy ϵ we distribute the total error as

$$\epsilon_{\text{QPE}} = \frac{10}{16}\epsilon, \quad \epsilon_{\text{trunc}} = \frac{3}{16}\epsilon, \quad \epsilon_{\text{prep}} = \frac{3}{16}\epsilon. \quad (84)$$

Since in condensed matter calculations one is often interested in the energy of the crystal per formula unit (f.u.), we allow the total error ϵ to grow proportionally to the system size, that is the error in energy estimation per f.u. is fixed when the size of the supercell increases. The required total accuracy of calculations strongly depends on the problem of interest. For example, the band gap of transition metal oxides is on the order of a few eV, while the energy difference between different crystal structures can be up to a few hundred meV/f.u. [122]. Therefore, we carry out truncation with different accuracy of 50 meV/f.u. (1.8 mHa/f.u., which is close to chemical accuracy), 5 meV/f.u., and 0.5 meV/f.u.

One of the bottlenecks in the quantum algorithm is the number of data items d that must be loaded into a quantum computer. The number of data items can be reduced by truncating small terms in the Hamiltonian. In order to estimate the error ϵ_{trunc} due to truncation of the Hamiltonian coefficients, one usually exploits classical heuristics based on the norm of the Hamiltonian's coefficients [32] or classical quantum chemistry calculations such as coupled-cluster theory [28,31]. For large systems, such as crystalline solids, high-order coupled-cluster calculations require large computational efforts and, therefore, we have used a simpler truncation methods based on the LCU's L_2 norm following Ref. [32]. Namely, if $H = \sum_i w_i U_i$ as in Eq. (58) then we define the truncated Hamiltonian $H_{\text{trunc}} = \sum_i w_i^{\text{trunc}} U_i$ based on the following equation:

$$\sqrt{\sum_i |w_i - w_i^{\text{trunc}}|^2} < \epsilon_{\text{trunc}}. \quad (85)$$

One can also estimate the error based on L_1 -norm truncation and even though such an estimation is rigorous, it does not allow truncating many coefficients. As shown in Fig. 5, there are many two-body terms in the Wannier representation with magnitude less than 10^{-8} Ha, and as a result, the error estimate based on L_1 -norm would already be on the order of 0.01 Ha if all these terms are neglected. The L_2 norm allows truncating much more coefficients but it is not a rigorous error estimate and, therefore, it might provide optimistic resource estimations. Instead of using L_1 -norm-based truncation, we will also present results for a truncation threshold of 10^{-9} Ha in order to understand how much L_2 -norm truncation reduces resource as compared to an accurate representation of the Hamiltonian. We use this procedure for moderate size systems. However, large systems such as NiO and PdO with supercells made of 64 and 72 atoms with around 900 spin orbitals require large

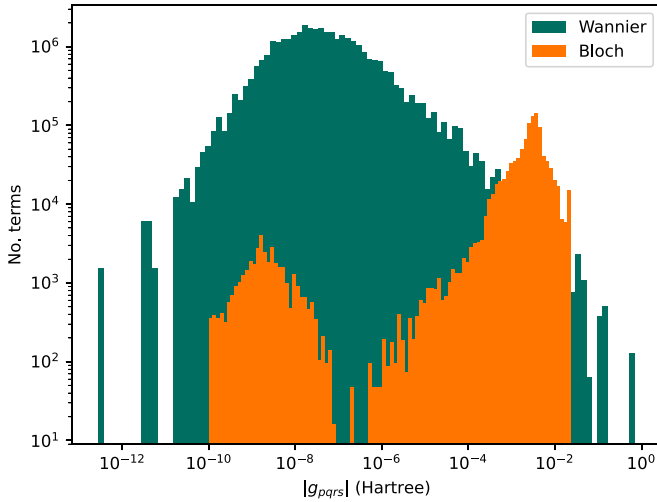


FIG. 5. The absolute value distribution of two-body matrix elements for H using Bloch and Wannier functions. The Hamiltonian terms were truncated at 10^{-16} Ha. The k -point mesh of $(4, 4, 4)$ is used for Bloch functions with a 2-atom unit cell, and a 128-atom supercell is used for constructing Wannier functions.

amount of memory. In this case, we save only those integrals whose absolute value is less than 10^{-7} Ha and then the Hamiltonian coefficients in Majorana representation are truncated with the threshold (1.81×10^{-6} Ha and 3.82×10^{-7} Ha for NiO-64 and PdO-72, respectively) obtained from calculations on smaller supercells at a total accuracy, ϵ , of 5 meV/f.u.

IV. RESULTS

A. Scaling

Both Wannier and Bloch functions provide the same spectrum of the Hamiltonian since they are related to each other

through unitary transformation. However, the Hamiltonian have different properties in different basis sets and the main task is to chose such basis functions, which minimize the total cost of quantum computation. It is proportional to the square root of the number of Hamiltonian coefficients d loaded into the quantum computer times the L_1 norm, λ , of the Hamiltonian divided by the precision of QPE simulation ϵ_{QPE} [see Eq. (78)]. With the example of model H in a BCC lattice structure, one can see that the number of nonzero terms in the Hamiltonian in Bloch representation scales as a cubic power of the system size, which is almost an order of magnitude better than in Wannier representation if no symmetries are taken into account [see Fig. 6(a)]. However, the trend is opposite for the L_1 norm, λ , and the Hamiltonian has a much smaller norm using Wannier functions [see Fig. 6(b)], which in turn significantly reduces the number of controlled-unitary steps in Hamiltonian simulation. This is on par with molecular calculations where the set of localized orbitals reduces the L_1 norm [85]. The number of T gates shown on Fig. 7 is significantly lower in Wannier representation and has a better asymptotic scaling. This is achieved by (i) loading only unique Hamiltonian terms into the quantum computer as discussed in Sec. III B 3, and (ii) providing a lower L_1 norm, λ , than in Bloch representation. Further reduction in T count can be accomplished by truncating the Hamiltonian coefficients using the L_2 -norm based truncation as shown on Fig. 7. After applying the L_2 -norm truncation, the number of T gates in model H systems scales as $O(N^{1.5})$ with the system size. Such a low asymptotic scaling can also be explained by the fact that the permissible error grows proportionally to the system size and the number of terms, remained after truncation in such model systems with Wannier functions, scales as $O(N^3)$. In molecular systems, one can expect that localized orbitals produce Hamiltonians with quadratic scaling with respect to number of nonzero terms [123]. We can expect the same for

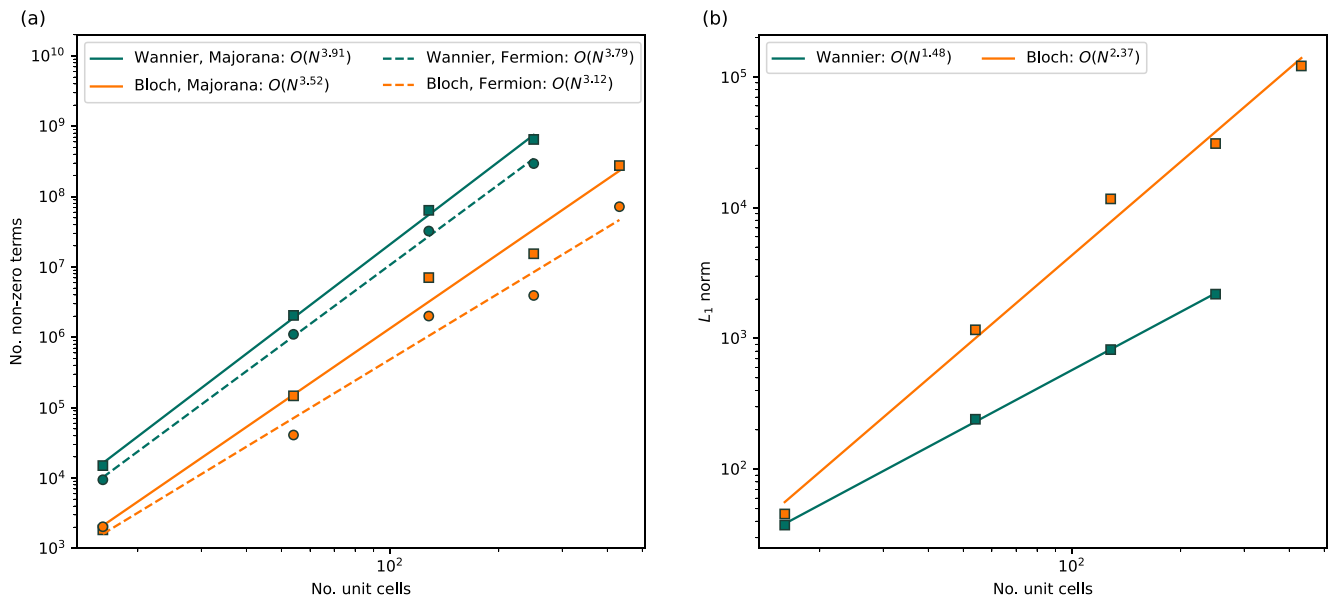


FIG. 6. (a) The total number of nonzero terms before taking into account translational symmetry of Wannier functions or symmetry of Brillouin zone and (b) the L_1 norm of the Hamiltonian in Majorana representations. Hydrogen in BCC lattice structure with varying supercell size or k -point sampling (x axis) is used. The Hamiltonian terms were truncated at 10^{-9} Ha.

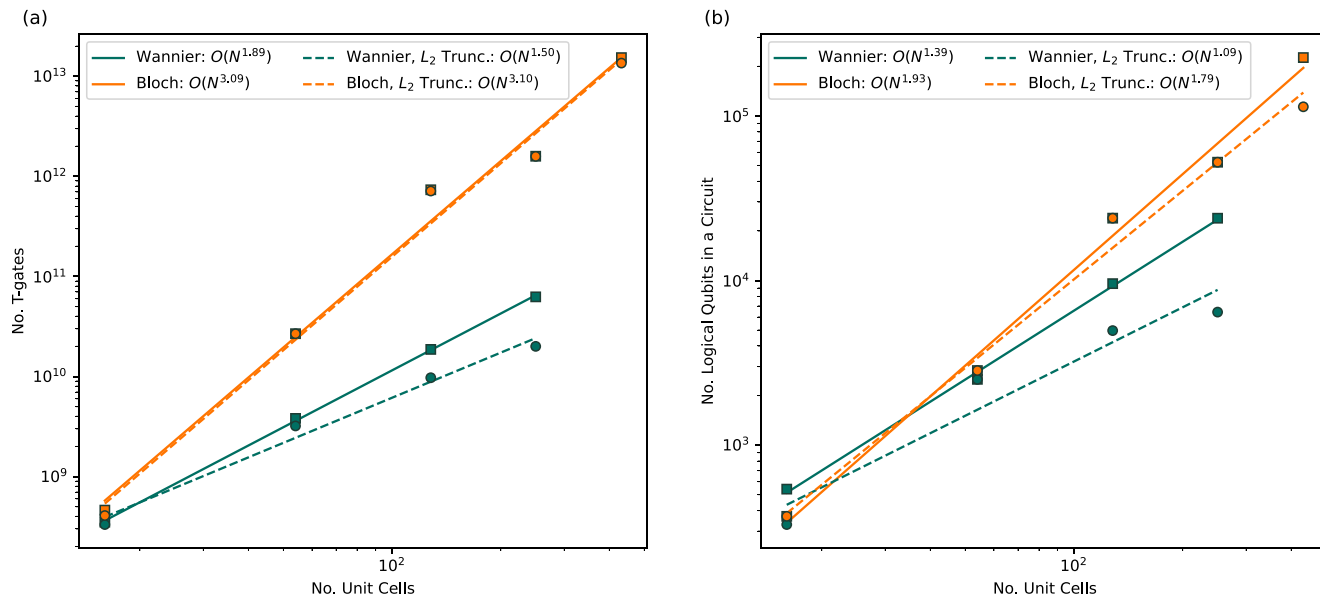


FIG. 7. (a) The number of T gates (b) The number of logical qubits in the circuit; no error correction has been applied here. Hydrogen in BCC lattice structure with varying supercell size or k -point sampling (x axis) is used. Resource estimations carried out so that the total accuracy ϵ is 0.5 meV/f.u.

materials with large band gaps while the model H system in our calculations have zero-gap at both PBE and Hartree-Fock level of theory.

B. Resource estimations

The number of T gates required for a single shot of the qubitized QPE circuit is presented in Fig. 8. When Wannier

functions are used as a basis set and the targeted accuracy is 50 meV/f.u., the number of T gates in the circuit is less than 10^{11} for all materials, system sizes and truncation strategies used in this paper. As one can see, reducing the total permissible error of the Hamiltonian simulation by a factor of 10 increases the number of T gates by approximately an order of magnitude in agreement with Eq. (78). The L_2 -norm truncation strategy reduces the number of T gates but not much except for H in

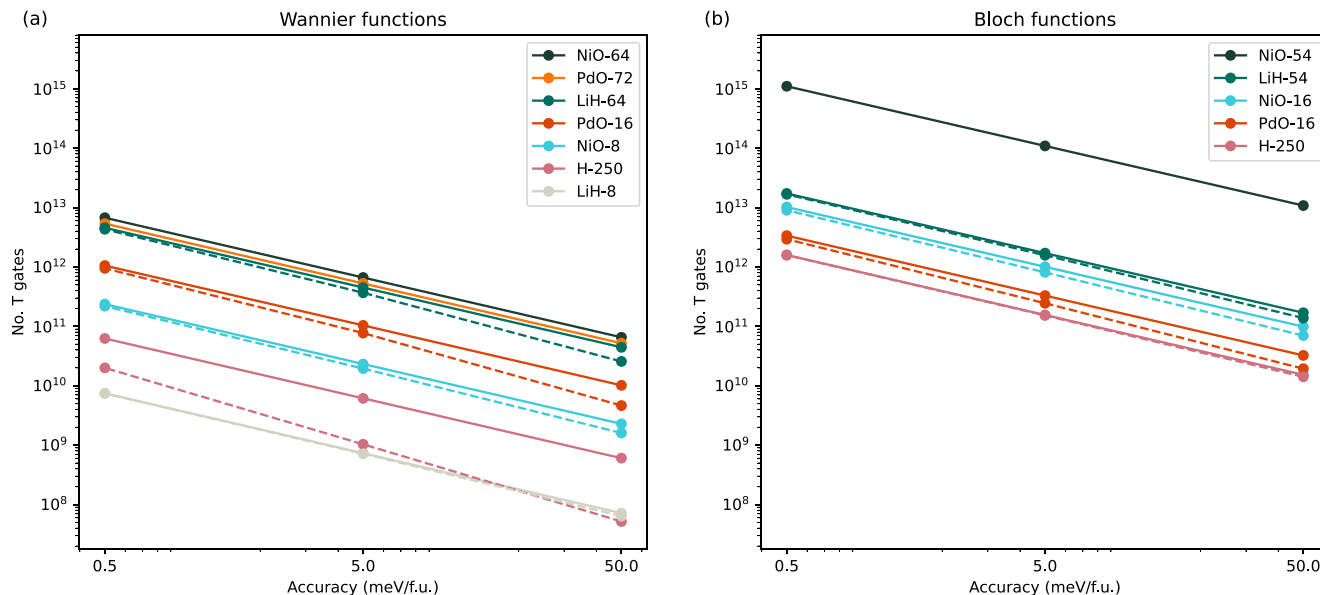


FIG. 8. The number of T gates as a function of the total accuracy of the Hamiltonian simulation ϵ using different basis functions: (a) Wannier functions. (b) Bloch functions. Solid lines have been obtained for Hamiltonians in which coefficients are truncated at 10^{-9} Ha except for NiO-64 and PdO-72, where truncation was chosen based on smaller cell calculations due to high memory requirements. Dashed lines were obtained using adaptive truncation threshold based on L_2 norm as explained in the main text. Legends indicate the name of the materials and the number next to the name is the total number of atoms in the supercell or the size of the k -point mesh for Wannier and Bloch functions, respectively.

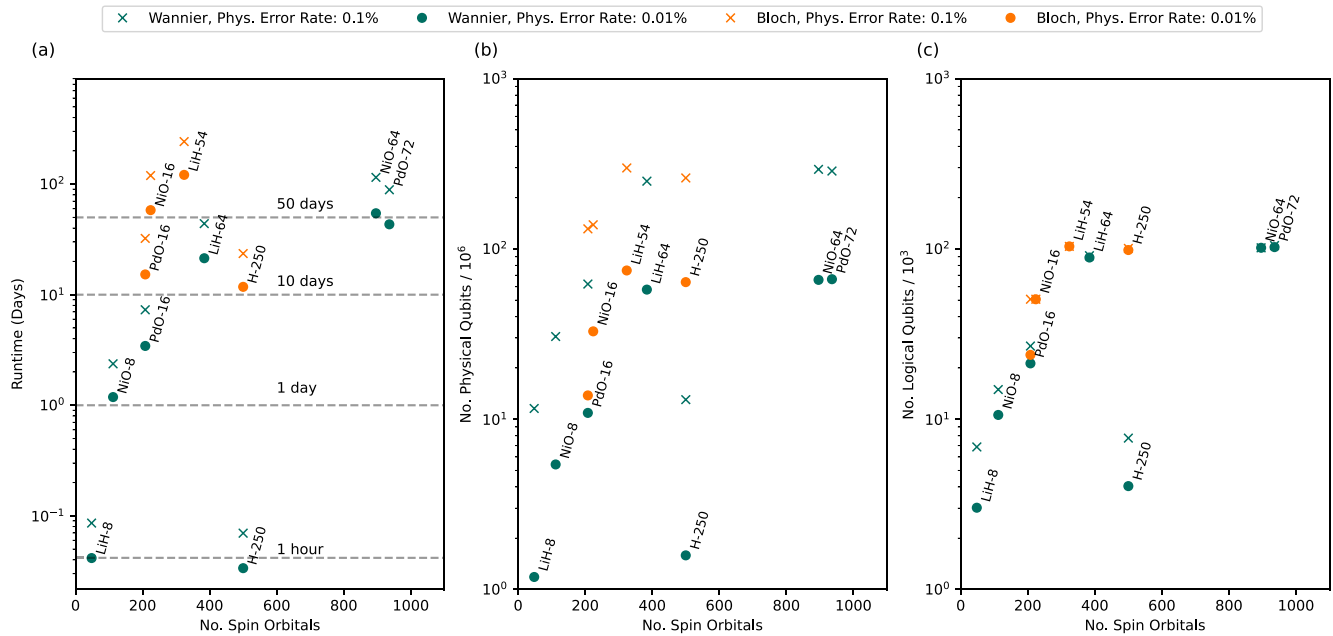


FIG. 9. Resource estimation for calculation of the ground-state energy of several solid state systems using Wannier and Bloch functions. The results have been obtained based on L_2 -norm truncation and the accuracy of Hamiltonian simulations is 50 meV/f.u. (1.8 mHa/f.u.). The code cycle duration is 10^{-6} s. The physical error rates are assumed to be 0.1% and 0.01%. (a) Runtime in days. (b) The number of physical qubits. (c) The number of logical qubits. x axis is the total number of spin orbitals. The number next to the name of the crystal indicates how many atoms in the supercell contains and how large the k -point mesh has been used for Wannier and Bloch basis sets, respectively.

Wannier basis, where the truncation of coefficients leads to an order of magnitude reduction of T gates. Similar trends are observed when Bloch functions are employed. However, in this basis set, the T-gate count is consistently higher than the T-gate count obtained with Wannier functions: for H the number of T gates is 25 times larger, for LiH with 54 k -points the T count is four times larger than for LiH with 64 atoms in the supercell whereas for NiO the difference between the two approaches is almost two orders of magnitudes. We can see that more than an order of magnitude comes from the fact that the L_1 norm in Wannier representation is smaller and the other improvement comes from truncation of small coefficients and symmetry considerations.

Figure 9(a) demonstrates the runtime required for a single shot of the QPE circuit with permissible total error of 50 meV/f.u. As can be seen, small-unit-cell simulations of NiO and PdO consisting of 8 and 16 atoms, respectively, can be performed within less than 10 days. Systems with larger computational cells such as LiH with 64 atoms in the cell can be simulated within 50 days even if the physical error rate of gate operations is 0.1%. Materials such as NiO with 64 atoms and PdO with 72 atoms in the supercell require a runtime of about 100 days when the physical error rate is 0.1%. Reducing the physical error rate by an order of magnitude to 0.01% leads to the reduction of the runtime by approximately a factor of 2 for all systems considered in this paper.

The number of physical and logical qubits for simulations described above are shown in Figs. 9(b) and 9(c). As can be seen from Fig. 9(b), the smallest simulations will require few million physical qubits if the physical error rate reaches 0.01%, while the largest simulations of NiO and PdO need about 65 million physical qubits. For the error rate of 0.1%,

quantum error correction requires the number of physical qubits to be 4–5 times larger. The number of logical qubits [see Fig. 9(c)] required for the simulation of small cells is around few thousands while large super cells would need around 10^5 logical qubits. We note that the improvement in the physical error rate occasionally reduces the number of logical qubits. This is because the number of logical qubits is the sum of the computational qubits and the magic state factory qubits. The computational qubit count is a feature of the system we study, and does not depend on the error rate of the quantum computer. The number of logical qubits dedicated to magic state distillation could in principle change when the error rate changes—because the fidelity with which we need to distill magic states depends on the error rate. Yet, many of our resource estimations just require the highest fidelity factory (225-to-1) in our list [96] for either error rates (0.1% or 0.01%), and hence the logical qubit count does not change. For much smaller error rates, smaller factories would suffice and we would see such changes.

V. DISCUSSION AND CONCLUSIONS

The estimation of the ground-state energy of crystalline solids with a supercell of ca. 50–70 atoms requires ca. 10^{10} – 10^{12} T gates when the size of the basis set is ca. 300–500 spatial orbitals. This is comparable to the T-gate count required for the estimation of the energy of molecular systems within the active space of several tens of orbitals. For example, simulation of a Ru complex with 65 spatial orbitals and using double factorization (DF) requires around 4.6×10^{10} T gates [32] while simulation of cytochrome P450 with 58 spatial orbitals and using tensor hypercontraction (THC) [31] requires

around 7.8×10^9 T gates [124]. Thus, if molecular Hamiltonians can be simulated within a reasonable time then so can a Hamiltonian describing crystalline solids. However, the number of logical qubits is larger for solids, 10^4 – 10^5 , which is due to the fact that the number of orbitals considered in this paper is larger almost by an order of magnitude as compared to molecular resource estimates.

We have considered the use of minimal Gaussian basis sets. However, for realistic solids one would need to use at least DZP or TZP basis sets, which would lead to a higher T-gate count. However, numerical studies for molecular systems indicate that methods like THC provide the best asymptotic scaling with respect to the number of orbitals for molecules [124]. Thus, using such approaches one might still obtain a reasonable resource estimates for crystalline solids. In order to carry out such estimations, one would have to (i) adapt such methods for periodic systems by also taking the translation symmetry into the consideration and (ii) develop classical electronic structure software for efficient generation of factorized Hamiltonians. For example, in order to generate ERIs for NiO with 64 atoms in the supercell and using GTH-DZP basis set one would need to use several TB of memory. Another approach, which would allow to perform useful simulations of solids on error-corrected quantum computers with larger basis sets is by choosing the active space within the size of several hundred of orbitals or by using quantum embedding methods.

In this paper, we have investigated how the translational symmetry of the Hamiltonian can be exploited in order to reduce the quantum resources. Similarly, other symmetries such as point group symmetry can be taken into account. However, we do not expect that this will lead to an order of magnitude reduction of T gates as qubitization-based algorithms scale as the square root of the number of terms. The use of Brillouin zone symmetry can also reduce the cost of quantum algorithms in Bloch basis set but we expect that Wannier representation will provide a better resource estimates for moderate-size systems. For example, 27 k points were used for NiO and even if each k point provided the same coefficients it would reduce the cost by a factor of 5.

The efficiency of the QPE to estimate the ground-state energy depends also on the overlap between an initial state such as Hartree-Fock state and the true ground-state wave function. We have not investigated this in this paper. For molecular systems, this overlap can appear to be either sufficient [32,124] or small [125] and more investigation needs to be carried out for crystalline solids. In this paper, we have focused only on the single-shot cost of the total QPE circuit.

In classical computations of ground-state energy of periodic solids, using Bloch functions is currently considered to be the most efficient approach in both KS-DFT [126] and wave function methods such as coupled-cluster theory [102]. Wannier functions are often used as a postprocessing tool for calculation of properties such as conductivity or band structure interpolation [62]. In quantum computing, however, the Wannier functions represent an efficient choice as a basis set for the ground-state energy calculations. Other areas, such as linear-scaling DFT also uses the Wannier functions [61,127] as the primary basis set.

In conclusion, we have considered ground-state energy estimation of crystalline solids on error-corrected quantum computers using qubitization based QPE. We present two materials such as NiO and PdO, which are known to be challenging systems for electronic structure methods on classical computers and are relevant for heterogeneous catalysis. The investigation of properties beyond ground-state energy calculations will be addressed in the future; this paper is a first step towards practical algorithms for simulation of crystalline solids on error-corrected quantum computers. We have adapted the qubitization algorithm to solid-state systems by taking into account the symmetries of the integrals in Wannier representation and generalized sparse qubitization for use with complex Hamiltonians, which are needed when Bloch functions are employed. Realistic resources estimations have been carried out and presented for error-corrected quantum computers. The simulation of crystalline solids in the minimal basis set on a quantum computer with the approach presented in this paper would require an order of 10–100 millions of physical qubits and 10^{10} – 10^{12} number of T gates. We expect that these numbers can be reduced further by using, for example, different qubitization techniques such as DF [32] or THC [31] adapted for solid-state systems.

The geometry and integral files for all systems considered in this paper are available at Zenodo [120].

ACKNOWLEDGMENTS

We thank Robert Íszak for valuable discussion and comments on the manuscript. We also thank Nick Blunt for helpful discussion as well as Earl Campbell for useful discussion on algorithms and error correction.

APPENDIX: DETAILED COSTINGS

In this Appendix, we give detailed costings for our Bloch and Wannier algorithms that include subdominant contributions. Since the two algorithms are very similar, we present them together and will note wherever there is a difference. All logarithms are base 2 as they refer to necessary qubit count for an integer range. The costings of standard circuit elements follow the costings of the sparse algorithm in Appendix A of Ref. [31].

1. Parameters

(i) Error parameters ϵ_{QPE} , ϵ_{prep} , ϵ_{trunc} are discussed in Sec. III C 1.

(ii) d is the number of nonzero coefficients in the Hamiltonian's LCU, after truncation and up to the symmetries restored. This means in the Wannier basis, one has to divide the number of nonzero coefficients (not taking into account translation symmetry) by the total number of cells N .

(iii) λ is the L_1 norm of the LCU, which is the the normalization factor of the block-encoded Hamiltonian. See (53).

(iv) M is the number of spatial orbitals per unit cell.

(v) N_1 , N_2 , and N_3 are the number of unit cells in each direction of the lattice as described in the main text. Therefore, the total number of spatial orbitals in the crystal is $P = MN_1N_2N_3$ and the number of spin orbitals is $2P$.

(vi) \aleph is the size of the “keep” register in coherent alias sampling. Coherent alias sampling effectively sets the amplitudes of the state based on the \aleph register. Therefore it affects the accuracy of the prepared state; following (A12) [31],

$$\aleph = \left\lceil \log \frac{\lambda}{2\epsilon_{prep}} \right\rceil. \quad (\text{A1})$$

(vii) m is the output size (qubit number) of the QROAM, the data loading in PREPARE. For Bloch functions, it is

$$m = \aleph + 2(4\lceil \log(P) \rceil + 6) \quad (\text{A2})$$

and for Wannier functions,

$$m = \aleph + 2(4\lceil \log M \rceil + 3 \sum_{i=1,2,3} \lceil \log N_i \rceil + 4), \quad (\text{A3})$$

where the compound index register must be split into its constituents. The +6 (Bloch) or +4 (Wannier) is made up of one qubit to distinguish one- and two-body terms, one qubit for the sign of the coefficient, three qubits (Bloch) or one qubit (Wannier) identifying term types for Majorana type restoration and one qubit for spin restoration. (Note that contrary to the simplified depiction in Figs. 3 and 4, the $|\sigma\rangle$ qubit in (76) is added as an ancilla rather than loaded.) The factor of 2 stems from needing “ind” and “alt” values for coherent alias sampling.

(viii) κ_1 and κ_2 are powers of 2. They determine the space-time tradeoff in the QROAM and QROAM uncompute. We choose them to minimize Toffoli cost.

(ix) $\mathcal{I} = \lceil \frac{\pi\lambda}{2\epsilon_{QPE}} \rceil$ the number of repetitions of the walk operator for quantum phase estimation.

(x) 2^n is the maximal power of 2 that's a factor of d .

(xi) b_r are bits of precision for the equal state preparation, we take it 7 as suggested in Ref. [31].

2. Toffoli count

The total Toffoli count is the product of \mathcal{I} (the number of iterations in phase estimation) and the number of Toffolis needed to construct the walk operator. The walk operator consists of the following circuit elements:

(i) PREPARE and UNPREPARE. The PREPARE operator is sketched in Figs. 3 and 4, while UNPREPARE uncomputes the state. We state the total cost of both, leading to a factor of two in most items.

(1) Equal state superposition over d basis states via amplitude amplification:

$$2(3\lceil \log d \rceil - 3\eta + 2b_r - 9) \quad (\text{A4})$$

(2) Data lookup via QROAM. This is the asymptotically dominant contribution to the walk operator. Uncomputing can be significantly simplified using a measurement based uncomputation scheme, such that the total cost is

$$\lceil d/\kappa_1 \rceil + m(\kappa_1 - 1) + \lceil d/\kappa_2 \rceil + \kappa_2. \quad (\text{A5})$$

(3) Coherent alias sampling. It can be uncomputed without Toffolis, giving

$$\aleph + (m - \aleph - 2)/2. \quad (\text{A6})$$

The \aleph is for an inequality test, and the register sizes to be swapped are $(m - \aleph)/2$. Swapping the sign qubits can be done without Toffolis, leading to the -2 .

(4) Majorana type symmetry restoration circuit (71) for Bloch basis,

$$2(5 + 25). \quad (\text{A7})$$

This cost includes a unary iteration, controlled on the 1-body/2-body qubit and iterating over six values of the $|s_i\rangle$ qubits (5 Toffolis). For the Wannier basis, the circuit can be simplified and the cost is only

$$2(1 + 7). \quad (\text{A8})$$

A controlled Hadamard can be implemented with a single Toffoli using a catalytic T state, see Fig. 17 in Ref. [31].

(5) Spin symmetry restoration: 0

(6) Translational symmetry restoration (Wannier basis only). The cost of computing and uncomputing the additions in (81) is

$$2 \left(\sum_{i=1,2,3} 3(\log N_i - 1) \right). \quad (\text{A9})$$

For this formula we have assumed that N_1, N_2, N_3 are all powers of 2. Then the in-place additions (or in the uncomputation, subtractions) modulo N_i can be performed with $\log N_i - 1$ Toffolis each [93]. If N_i is not a power of 2, the cost will be higher to perform the correct modular arithmetic. Yet this symmetry restoration circuit is a subleading contribution and for simplicity we use (A9) even if the cell numbers are not powers of 2.

(ii) SELECT. We have two ranged operations that are uncontrolled and two that are controlled on the qubit flagging one- or two-body terms and thereby the number of Majoranas in the unitary. Each of these unary iterations is over $4P$ values. These are spin, Majorana type, and the indices specifying each Majorana (see Figs. 3 and 4). The total cost is

$$2(4P - 2) + 2(4P - 1). \quad (\text{A10})$$

(iii) Reflection. The walk operator is built from the block-encoded Hamiltonian along with a reflection. This is implemented by a multicontrolled Z controlled on a number of qubits. These are $\lceil \log d \rceil$ for the QROAM index, one further qubit from preparing the equal superposition state, \aleph for the equal superposition state in coherent alias sampling, one qubit from spin, four qubits from Majorana type, and (only for Wannier) $\lceil \log N_1 \rceil + \lceil \log N_2 \rceil + \lceil \log N_3 \rceil$. The gate controlled on c qubits can be implemented with $c - 1$ Toffolis, giving

$$\lceil \log d \rceil + \aleph + 5 + (\text{Wannier only: } \sum_{i=1,2,3} \lceil \log N_i \rceil). \quad (\text{A11})$$

Toffolis.

(iv) Two more Toffolis for each step (to make the reflection controlled, and for the unary iteration in the phase estimation)

Note we do not include the cost of initial state preparation for the Heisenberg-limited phase estimation or of the inverse QFT. These are small additive costs that are not multiplied by $\mathcal{I} \propto \lambda/\epsilon_{QPE}$ in contrast to all the above contributions.

3. Qubit count

- (i) For phase estimation and unary iteration circuit:
 $\lceil \log(\mathcal{Z} + 1) \rceil + \lceil \log(\mathcal{Z} + 1) \rceil - 1$
- (ii) System qubits (on which the Majoranas act): $2P$
- (iii) QROAM input state: $\lceil \log d \rceil$
- (iv) $\aleph + 2 + b_r$ for equal superposition state and coherent alias sampling
- (v) QROAM qubits including output and ancillas: $m\kappa_1 + \lceil \log(d/\kappa_1) \rceil$
- (vi) Further ancillary qubits are not needed as QROAM ancilla qubits can be reused.
-
- [1] H. Liu, G. H. Low, D. S. Steiger, T. Häner, M. Reiher, and M. Troyer, Prospects of quantum computing for molecular sciences, *Mater. Theory* **6**, 11 (2022).
- [2] B. Bauer, S. Bravyi, M. Motta, and G. K.-L. Chan, Quantum algorithms for quantum chemistry and quantum materials science, *Chem. Rev.* **120**, 12685 (2020).
- [3] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, Quantum computational chemistry, *Rev. Mod. Phys.* **92**, 015003 (2020).
- [4] R. P. Feynman, Simulating physics with computers, *Int. J. Theor. Phys.* **21**, 467 (1982).
- [5] D. S. Abrams and S. Lloyd, Simulation of Many-Body Fermi Systems on a Universal Quantum Computer, *Phys. Rev. Lett.* **79**, 2586 (1997).
- [6] D. S. Abrams and S. Lloyd, Quantum Algorithm Providing Exponential Speed Increase for Finding Eigenvalues and Eigenvectors, *Phys. Rev. Lett.* **83**, 5162 (1999).
- [7] G. Ortiz, J. E. Gubernatis, E. Knill, and R. Lafamme, Quantum algorithms for fermionic simulations, *Phys. Rev. A* **64**, 022319 (2001).
- [8] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, Simulated quantum computation of molecular energies, *Science* **309**, 1704 (2005).
- [9] I. Kassal, S. P. Jordan, P. J. Love, M. Mohseni, and A. Aspuru-Guzik, Polynomial-time quantum algorithm for the simulation of chemical dynamics, *Proc. Natl. Acad. Sci. USA* **105**, 18681 (2008).
- [10] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, Simulation of electronic structure Hamiltonians using quantum computers, *Mol. Phys.* **109**, 735 (2011).
- [11] J. T. Seeley, M. J. Richard, and P. J. Love, The Bravyi-Kitaev transformation for quantum computation of electronic structure, *J. Chem. Phys.* **137**, 224109 (2012).
- [12] B. Toloui and P. J. Love, Quantum algorithms for quantum chemistry based on the sparsity of the CI-matrix, [arXiv:1312.2579](https://arxiv.org/abs/1312.2579).
- [13] 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).
- [14] D. Wecker, B. Bauer, B. K. Clark, M. B. Hastings, and M. Troyer, Gate-count estimates for performing quantum chemistry on small quantum computers, *Phys. Rev. A* **90**, 022305 (2014).
- [15] D. Poulin, M. B. Hastings, D. Wecker, N. Wiebe, A. C. Doherty, and M. Troyer, The trotter step size required for accurate quantum simulation of quantum chemistry, [arXiv:1406.4920](https://arxiv.org/abs/1406.4920).
- [16] M. B. Hastings, D. Wecker, B. Bauer, and M. Troyer, Improving quantum algorithms for quantum chemistry, [arXiv:1403.1539](https://arxiv.org/abs/1403.1539).
- [17] J. R. McClean, R. Babbush, P. J. Love, and A. Aspuru-Guzik, Exploiting locality in quantum computation for quantum chemistry, *J. Phys. Chem. Lett.* **5**, 4368 (2014).
- [18] R. Babbush, J. McClean, D. Wecker, A. Aspuru-Guzik, and N. Wiebe, Chemical basis of Trotter-Suzuki errors in quantum chemistry simulation, *Phys. Rev. A* **91**, 022311 (2015).
- [19] D. Wecker, M. B. Hastings, and M. Troyer, Progress towards practical quantum variational algorithms, *Phys. Rev. A* **92**, 042303 (2015).
- [20] R. Babbush, D. W. Berry, I. D. Kivlichan, A. Y. Wei, P. J. Love, and A. Aspuru-Guzik, Exponentially more precise quantum simulation of fermions in second quantization, *New J. Phys.* **18**, 033032 (2016).
- [21] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, The theory of variational hybrid quantum-classical algorithms, *New J. Phys.* **18**, 023023 (2016).
- [22] I. D. Kivlichan, N. Wiebe, R. Babbush, and A. Aspuru-Guzik, Bounding the costs of quantum simulation of many-body physics in real space, *J. Phys. A: Math. Theor.* **50**, 305301 (2017).
- [23] R. Babbush, D. W. Berry, Y. R. Sanders, I. D. Kivlichan, A. Scherer, A. Y. Wei, P. J. Love, and A. Aspuru-Guzik, Exponentially more precise quantum simulation of fermions in the configuration interaction representation, *Quantum Sci. Technol.* **3**, 015006 (2018).
- [24] R. Babbush, N. Wiebe, J. McClean, J. McClain, H. Neven, and G. K.-L. Chan, Low-Depth Quantum Simulation of Materials, *Phys. Rev. X* **8**, 011044 (2018).
- [25] D. Poulin, A. Kitaev, D. S. Steiger, M. B. Hastings, and M. Troyer, Quantum Algorithm for Spectral Measurement with Lower Gate Count, *Phys. Rev. Lett.* **121**, 010501 (2018).
- [26] D. W. Berry, M. Kieferová, A. Scherer, Y. R. Sanders, G. H. Low, N. Wiebe, C. Gidney, and R. Babbush, Improved techniques for preparing eigenstates of fermionic Hamiltonians, *npj Quantum Inf.* **4**, 22 (2018).
- [27] R. Babbush, C. Gidney, D. W. Berry, N. Wiebe, J. McClean, A. Paler, A. Fowler, and H. Neven, Encoding Electronic Spectra in Quantum Circuits with Linear T Complexity, *Phys. Rev. X* **8**, 041015 (2018).
- [28] D. W. Berry, C. Gidney, M. Motta, J. R. McClean, and R. Babbush, Qubitization of arbitrary basis quantum chemistry leveraging sparsity and low rank factorization, *Quantum* **3**, 208 (2019).
- [29] O. Higgott, D. Wang, and S. Brierley, Variational quantum computation of excited states, *Quantum* **3**, 156 (2019).

- [30] D. Wang, O. Higgott, and S. Brierley, Accelerated Variational Quantum Eigensolver, *Phys. Rev. Lett.* **122**, 140504 (2019).
- [31] J. Lee, D. W. Berry, C. Gidney, W. J. Huggins, J. R. McClean, N. Wiebe, and R. Babbush, Even more efficient quantum computations of chemistry through tensor hypercontraction, *PRX Quantum* **2**, 030305 (2021).
- [32] V. von Burg, G. H. Low, T. Häner, D. S. Steiger, M. Reiher, M. Roetteler, and M. Troyer, Quantum computing enhanced computational catalysis, *Phys. Rev. Res.* **3**, 033055 (2021).
- [33] W. J. Huggins, B. A. O’Gorman, N. C. Rubin, D. R. Reichman, R. Babbush, and J. Lee, Unbiasing fermionic quantum Monte Carlo with a quantum computer, *Nature (London)* **603**, 416 (2022).
- [34] Y. Su, D. W. Berry, N. Wiebe, N. Rubin, and R. Babbush, Fault-tolerant quantum simulations of chemistry in first quantization, *PRX Quantum* **2**, 040332 (2021).
- [35] D. Z. Manrique, I. T. Khan, K. Yamamoto, V. Wichitwechkarn, and D. M. Ramo, Momentum-space unitary coupled cluster and translational quantum subspace expansion for periodic systems on quantum computers, [arXiv:2008.08694](https://arxiv.org/abs/2008.08694).
- [36] L. Clinton, T. Cubitt, B. Flynn, F. M. Gambetta, J. Klassen, A. Montanaro, S. Piddock, R. A. Santos, and E. Sheridan, Towards near-term quantum simulation of materials, [arXiv:2205.15256](https://arxiv.org/abs/2205.15256).
- [37] D. Song, N. P. Bauman, G. Prawiroatmodjo, B. Peng, C. Granade, K. M. Rosso, G. H. Low, M. Roetteler, K. Kowalski, and E. J. Bylaska, Periodic plane-wave electronic structure calculations on quantum computers, [arXiv:2208.04444](https://arxiv.org/abs/2208.04444).
- [38] N. Yoshioka, T. Sato, Y. O. Nakagawa, Y.-y. Ohnishi, and W. Mizukami, Variational quantum simulation for periodic materials, *Phys. Rev. Res.* **4**, 013052 (2022).
- [39] M. Motta, C. Sun, A. T. K. Tan, M. J. O’Rourke, E. Ye, A. J. Minnich, F. G. S. L. Brandão, and G. K.-L. Chan, Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution, *Nat. Phys.* **16**, 205 (2020).
- [40] I. D. Kivlichan, C. Gidney, D. W. Berry, N. Wiebe, J. McClean, W. Sun, Z. Jiang, N. Rubin, A. Fowler, A. Aspuru-Guzik *et al.*, Improved fault-tolerant quantum simulation of condensed-phase correlated electrons via trotterization, *Quantum* **4**, 296 (2020).
- [41] E. T. Campbell, Early fault-tolerant simulations of the Hubbard model, *Quantum Sci. Technol.* **7**, 015007 (2021).
- [42] S. Kanno, S. Endo, T. Utsumi, and T. Tada, Resource estimations for the Hamiltonian simulation in correlated electron materials, *Phys. Rev. A* **106**, 012612 (2022).
- [43] S. Flannigan, N. Pearson, G. Low, A. Buyskikh, I. Bloch, P. Zoller, M. Troyer, and A. Daley, Propagation of errors and quantitative quantum simulation with quantum advantage, [arXiv:2204.13644](https://arxiv.org/abs/2204.13644).
- [44] P. Hohenberg and W. Kohn, Inhomogeneous electron gas, *Phys. Rev.* **136**, B864 (1964).
- [45] W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, *Phys. Rev. A* **140**, 1133 (1965).
- [46] W. Kohn, Nobel lecture: Electronic structure of matter—wave functions and density functionals, *Rev. Mod. Phys.* **71**, 1253 (1999).
- [47] 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).
- [48] I. Shavitt and R. J. Bartlett, *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory* (Cambridge University Press, Cambridge, 2009).
- [49] D. Cremer, Density functional theory: Coverage of dynamic and non-dynamic electron correlation effects, *Mol. Phys.* **99**, 1899 (2001).
- [50] J. P. Perdew, A. Ruzsinszky, J. Sun, N. K. Nepal, and A. D. Kaplan, Interpretations of ground-state symmetry breaking and strong correlation in wavefunction and density functional theories, *Proc. Natl. Acad. Sci. USA* **118**, e2017850118 (2021).
- [51] M. Born and R. Oppenheimer, Zur Quantentheorie der Molekeln, *Ann. Phys. (Leipzig)* **389**, 457 (1927).
- [52] A. Y. Kitaev, Quantum measurements and the Abelian stabilizer problem, [arXiv:quant-ph/9511026](https://arxiv.org/abs/quant-ph/9511026).
- [53] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, Quantum algorithms revisited, *Proc. R. Soc. London, Ser. A* **454**, 339 (1998).
- [54] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, 10th ed. (Cambridge University Press, Cambridge, 2010).
- [55] N. S. Blunt, J. Camps, O. Crawford, R. Izsák, S. Leontica, A. Mirani, A. E. Moylett, S. A. Scivier, C. Sünderhauf, P. Schopf *et al.*, Perspective on the current state-of-the-art of quantum computing for drug discovery applications, *J. Chem. Theory Comput.* **18**, 7001 (2022).
- [56] A. G. Fowler and C. Gidney, Low overhead quantum computation using lattice surgery, [arXiv:1808.06709](https://arxiv.org/abs/1808.06709).
- [57] G. H. Low and I. L. Chuang, Hamiltonian simulation by qubitization, *Quantum* **3**, 163 (2019).
- [58] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Iterative minimization techniques for *ab initio* total-energy calculations: Molecular dynamics and conjugate gradients, *Rev. Mod. Phys.* **64**, 1045 (1992).
- [59] G. Kresse and J. Furthmüller, Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set, *Phys. Rev. B* **54**, 11169 (1996).
- [60] N. Marzari and D. Vanderbilt, Maximally localized generalized Wannier functions for composite energy bands, *Phys. Rev. B* **56**, 12847 (1997).
- [61] C.-K. Skylaris, A. A. Mostofi, P. D. Haynes, O. Diéguez, and M. C. Payne, Nonorthogonal generalized Wannier function pseudopotential plane-wave method, *Phys. Rev. B* **66**, 035119 (2002).
- [62] N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt, Maximally localized Wannier functions: Theory and applications, *Rev. Mod. Phys.* **84**, 1419 (2012).
- [63] P. W. Shor, Scheme for reducing decoherence in quantum computer memory, *Phys. Rev. A* **52**, R2493 (1995).
- [64] J. Roffe, Quantum error correction: An introductory guide, *Contemp. Phys.* **60**, 226 (2019).
- [65] D. Litinski, A game of surface codes: Large-scale quantum computing with lattice surgery, *Quantum* **3**, 128 (2019).
- [66] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Surface codes: Towards practical large-scale quantum computation, *Phys. Rev. A* **86**, 032324 (2012).

- [67] C. Horsman, A. G. Fowler, S. Devitt, and R. V. Meter, Surface code quantum computing by lattice surgery, *New J. Phys.* **14**, 123011 (2012).
- [68] Edited by M. V. Twigg, *Catalyst Handbook*, 2nd ed. (Routledge, New York, 2018).
- [69] IPCC, 2021: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2021), pp. 2391, https://report.ipcc.ch/ar6/wg1/IPCC_AR6_WGI_FullReport.pdf.
- [70] J. Cheng, P. Hu, P. Ellis, S. French, G. Kelly, and C. M. Lok, Brønsted-Evans-Polanyi relation of multistep reactions and volcano curve in heterogeneous catalysis, *J. Phys. Chem. C* **112**, 1308 (2008).
- [71] H. Toulhoat and P. Raybaud, Kinetic interpretation of catalytic activity patterns based on theoretical chemical descriptors, *J. Catal.* **216**, 63 (2003).
- [72] A. J. Medford, A. Vojvodic, J. S. Hummelshøj, J. Voss, F. Abild-Pedersen, F. Studt, T. Bligaard, A. Nilsson, and J. K. Nørskov, From the Sabatier principle to a predictive theory of transition-metal heterogeneous catalysis, *J. Catal.* **328**, 36 (2015).
- [73] H. J. Kulik, Perspective: Treating electron over-delocalization with the DFT+U method, *J. Chem. Phys.* **142**, 240901 (2015).
- [74] C. Adamo and V. Barone, Toward reliable density functional methods without adjustable parameters: The PBE0 model, *J. Chem. Phys.* **110**, 6158 (1999).
- [75] S. Mandal, K. Haule, K. M. Rabe, and D. Vanderbilt, Systematic beyond-DFT study of binary transition metal oxides, *npj Comput Mater* **5**, 115 (2019).
- [76] J. E. Coulter, E. Manousakis, and A. Gali, Limitations of the hybrid functional approach to electronic structure of transition metal oxides, *Phys. Rev. B* **88**, 041107(R) (2013).
- [77] G. Stefanucci and R. van Leeuwen, *Nonequilibrium Many-body Theory of Quantum Systems* (Cambridge University Press, Cambridge, 2013).
- [78] C. Kittel, *Introduction to Solid State Physics*, 8th ed. (Wiley, Hoboken, NJ, 2005).
- [79] J. M. Foster and S. F. Boys, Canonical configurational interaction procedure, *Rev. Mod. Phys.* **32**, 300 (1960).
- [80] J. Pipek and P. G. Mezey, A fast intrinsic localization procedure applicable for *ab initio* and semiempirical linear combination of atomic orbital wave functions, *J. Chem. Phys.* **90**, 4916 (1989).
- [81] S. Lehtola and H. Jónsson, Pipek–mezey orbital localization using various partial charge estimates, *J. Chem. Theory Comput.* **10**, 642 (2014).
- [82] E. Ö. Jónsson, S. Lehtola, M. Puska, and H. Jónsson, Theory and applications of generalized Pipek–Mezey Wannier functions, *J. Chem. Theory Comput.* **13**, 460 (2017).
- [83] P. Jordan and E. Wigner, Über das paulische äquivalenzverbot, *Z. Angew. Phys.* **47**, 631 (1928).
- [84] S. B. Bravyi and A. Y. Kitaev, Fermionic quantum computation, *Ann. Phys.* **298**, 210 (2002).
- [85] E. Koridon, S. Yalouz, B. Senjean, F. Buda, T. E. O'Brien, and L. Visscher, Orbital transformations to reduce the 1-norm of the electronic structure hamiltonian for quantum computing applications, *Phys. Rev. Res.* **3**, 033127 (2021).
- [86] A. M. Childs and N. Wiebe, Hamiltonian simulation using linear combinations of unitary operations, *Quantum Inf. Comput.* **12**, 901 (2012).
- [87] M. Suzuki, General theory of fractal path integrals with applications to many-body theories and statistical physics, *J. Math. Phys.* **32**, 400 (1991).
- [88] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, and Y. Su, Toward the first quantum simulation with quantum speedup, *Proc. Natl. Acad. Sci. USA* **115**, 9456 (2018).
- [89] E. Campbell, Random Compiler for Fast Hamiltonian Simulation, *Phys. Rev. Lett.* **123**, 070503 (2019).
- [90] I. D. Kivlichan, C. E. Granade, and N. Wiebe, Phase estimation with randomized Hamiltonians, [arXiv:1907.10070](https://arxiv.org/abs/1907.10070).
- [91] K. Wan, M. Berta, and E. T. Campbell, Randomized Quantum Algorithm for Statistical Phase Estimation, *Phys. Rev. Lett.* **129**, 030503 (2022).
- [92] G. H. Low, V. Kliuchnikov, and L. Schaeffer, Trading T-gates for dirty qubits in state preparation and unitary synthesis, [arXiv:1812.00954](https://arxiv.org/abs/1812.00954).
- [93] C. Gidney, Halving the cost of quantum addition, *Quantum* **2**, 74 (2018).
- [94] The symmetry restoration circuit will add some factors of $1/\sqrt{2}$, similarly to spin restoration. In the Hamiltonian LCU (Sec. III A 3), the coefficients in between the two \sum s are the values that must be loaded and prepared by coherent alias sampling. The factor in front of the first \sum and the sign at the back of the terms are recovered along with the various symmetry restoration circuits.
- [95] In order to cancel the $1/\sqrt{N}$ coming from the Hadamards, the coefficients with $n_4 = 0$ should that are loaded should be multiplied N times their value. The total subnormalization of the Hamiltonian does not change.
- [96] Specifically, we consider 15-to-1, 14-to-2, 20-to-4, 116-to-12, and 225-to-1 factories. The largest calculations all require the highest fidelity factory in this list: 225-to-1.
- [97] S. Krinner, N. Lacroix, A. Remm, A. Di Paolo, E. Genois, C. Leroux, C. Hellings, S. Lazar, F. Swiadek, J. Herrmann *et al.*, Realizing repeated quantum error correction in a distance-three surface code, *Nature (London)* **605**, 669 (2022).
- [98] R. Acharya, I. Aleiner, R. Allen, T. I. Andersen, M. Ansmann, F. Arute, K. Arya, A. Asfaw, J. Atalaya, R. Babbush *et al.*, Suppressing quantum errors by scaling a surface code logical qubit, [arXiv:2207.06431](https://arxiv.org/abs/2207.06431).
- [99] Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui *et al.*, Recent developments in the pyscf program package, *J. Chem. Phys.* **153**, 024109 (2020).
- [100] Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, and G. K. Chan, Pyscf: The Python-based simulations of chemistry framework, *WIREs Computational Molecular Science* **8**, 1340 (2018).
- [101] Q. Sun, Libcint: An efficient general integral library for Gaussian basis functions, *J. Comput. Chem.* **36**, 1664 (2015).
- [102] J. McClain, Q. Sun, G. K.-L. Chan, and T. C. Berkelbach, Gaussian-based coupled-cluster theory for the ground-state

- and band structure of solids, *J. Chem. Theory Comput.* **13**, 1209 (2017).
- [103] W. J. Hehre, R. F. Stewart, and J. A. Pople, Self-consistent molecular-orbital methods. I. Use of Gaussian expansions of slater-type atomic orbitals, *J. Chem. Phys.* **51**, 2657 (1969).
- [104] W. J. Hehre, R. Ditchfield, R. F. Stewart, and J. A. Pople, Self-consistent molecular orbital methods. IV. Use of Gaussian expansions of slater-type orbitals. extension to second-row molecules, *J. Chem. Phys.* **52**, 2769 (1970).
- [105] D. Feller, The role of databases in support of computational chemistry calculations, *J. Comput. Chem.* **17**, 1571 (1996).
- [106] K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, and T. L. Windus, Basis set exchange: A community database for computational sciences, *J. Chem. Inf. Model.* **47**, 1045 (2007).
- [107] B. P. Pritchard, D. Altarawy, B. Didier, T. D. Gibbsom, and T. L. Windus, A new basis set exchange: An open, up-to-date resource for the molecular sciences community, *J. Chem. Inf. Model.* **59**, 4814 (2019).
- [108] S. Goedecker, M. Teter, and J. Hutter, Separable dual-space Gaussian pseudopotentials, *Phys. Rev. B* **54**, 1703 (1996).
- [109] C. Hartwigsen, S. Goedecker, and J. Hutter, Relativistic separable dual-space Gaussian pseudopotentials from H to Rn, *Phys. Rev. B* **58**, 3641 (1998).
- [110] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach, *Comput. Phys. Commun.* **167**, 103 (2005).
- [111] J. VandeVondele and J. Hutter, Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases, *J. Chem. Phys.* **127**, 114105 (2007).
- [112] Q. Sun, T. C. Berkelbach, J. D. McClain, and G. K.-L. Chan, Gaussian and plane-wave mixed density fitting for periodic systems, *J. Chem. Phys.* **147**, 164119 (2017).
- [113] J. J. Mortensen, L. B. Hansen, and K. W. Jacobsen, Real-space grid implementation of the projector augmented wave method, *Phys. Rev. B* **71**, 035109 (2005).
- [114] J. Enkovaara, C. Rostgaard, J. J. Mortensen, J. Chen, M. Duřak, L. Ferrighi, J. Gavnholt, C. Glinsvad, V. Haikola, H. A. Hansen *et al.*, Electronic structure calculations with GPAW: A real-space implementation of the projector augmented-wave method, *J. Phys.: Condens. Matter* **22**, 253202 (2010).
- [115] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Duřak, J. Friis, M. N. Groves, B. Hammer, C. Hargus *et al.*, The atomic simulation environment—A python library for working with atoms, *J. Phys.: Condens. Matter* **29**, 273002 (2017).
- [116] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo *et al.*, QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials, *J. Phys.: Condens. Matter* **21**, 395502 (2009).
- [117] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. Buongiorno Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli *et al.*, Advanced capabilities for materials modelling with Quantum ESPRESSO, *J. Phys.: Condens. Matter* **29**, 465901 (2017).
- [118] C. E. Messer, A survey report on lithium hydride, *Web* [10.2172/4062196](https://doi.org/10.2172/4062196) (1960).
- [119] S. Sasaki, K. Fujino, and Y. Takéuchi, X-ray determination of electron-density distributions in oxides, MgO, MnO, CoO, and NiO, and atomic scattering factors of their constituent atoms, *Proc. Jpn. Acad., Ser. B* **55**, 43 (1979).
- [120] A. V. Ivanov, C. Sünderhauf, N. Holzmann, T. Ellaby, R. N. Kerber, G. Jones, and J. Camps, Geometry and integral files for the manuscript “Quantum computation for periodic solids in second quantization” [Data set]., Zenodo, <https://doi.org/10.5281/zenodo.7148994>.
- [121] A. E. Reed, R. B. Weinstock, and F. Weinhold, Natural population analysis, *J. Chem. Phys.* **83**, 735 (1985).
- [122] H. Peng and J. P. Perdew, Synergy of van der Waals and self-interaction corrections in transition metal monoxides, *Phys. Rev. B* **96**, 100101(R) (2017).
- [123] I. Sabzevari and S. Sharma, Improved speed and scaling in orbital space variational Monte Carlo, *J. Chem. Theory Comput.* **14**, 6276 (2018).
- [124] J. J. Goings, A. White, J. Lee, C. S. Tautermann, M. Degroote, C. Gidney, T. Shiozaki, R. Babbush, and N. C. Rubin, Reliably assessing the electronic structure of cytochrome P450 on today’s classical computers and tomorrow’s quantum computers, *Proc. Natl. Acad. Sci. USA* **119**, e2203533119 (2022).
- [125] S. Lee, J. Lee, H. Zhai, Y. Tong, A. M. Dalzell, A. Kumar, P. Helms, J. Gray, Z.-H. Cui, W. Liu *et al.*, Is there evidence for exponential quantum advantage in quantum chemistry? [arXiv:2208.02199](https://arxiv.org/abs/2208.02199).
- [126] G. Kresse and J. Furthmüller, Efficiency of *ab-initio* total energy calculations for metals and semiconductors using a plane-wave basis set, *Comput. Mater. Sci.* **6**, 15 (1996).
- [127] A. A. Mostofi, C.-K. Skylaris, P. D. Haynes, and M. C. Payne, Total-energy calculations on a real space grid with localized functions and a plane-wave basis, *Comput. Phys. Commun.* **147**, 788 (2002).