**PAPER • OPEN ACCESS**

# Discontinuous Galerkin discretization for quantum simulation of chemistry

View the article online for updates and enhancements.

# New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft **DPG**

**IOP** Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

CrossMark

**OPEN ACCESS**

# Discontinuous Galerkin discretization for quantum simulation of chemistry
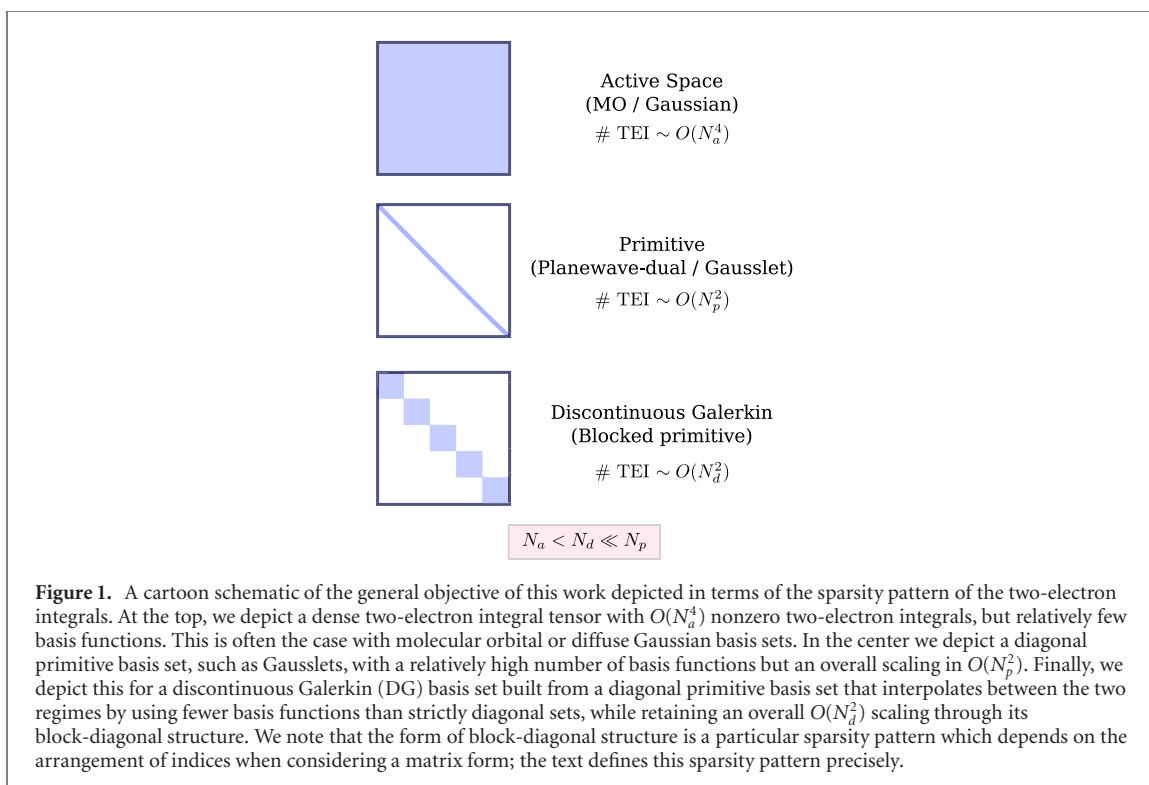
Jarrod R McClean[1,8], Fabian M Faulstich[2], Qinyi Zhu[3], Bryan O'Gorman[4,5], Yiheng Qiu[6], Steven R White[6], Ryan Babbush[1] and Lin Lin[3,7,8]

[1] Google Research, 340 Main Street, Venice, CA 90291, United States of America
[2] Hylleraas Centre for Quantum Molecular Sciences, Department of Chemistry, University of Oslo, Oslo, Norway
[3] Department of Mathematics, University of California, Berkeley, CA 94720, United States of America
[4] Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720, United States of America
[5] Quantum Artificial Intelligence Laboratory, NASA Ames Research Center, Moffett Field, CA 94035, United States of America
[6] Department of Physics and Astronomy, University of California, Irvine, CA 92697, United States of America
[7] Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, United States of America
[8] Authors to whom any correspondence should be addressed.

E-mail: linlin@math.berkeley.edu and jmcclean@google.com

## Abstract

All-electron electronic structure methods based on the linear combination of atomic orbitals method with Gaussian basis set discretization offer a well established, compact representation that forms much of the foundation of modern correlated quantum chemistry calculations—on both classical and quantum computers. Despite their ability to describe essential physics with relatively few basis functions, these representations can suffer from a quartic growth of the number of integrals. Recent results have shown that, for some quantum and classical algorithms, moving to representations with diagonal two-body operators can result in dramatically lower asymptotic costs, even if the number of functions required increases significantly. We introduce a way to interpolate between the two regimes in a systematic and controllable manner, such that the number of functions is minimized while maintaining a block-diagonal structure of the two-body operator and desirable properties of an original, primitive basis. Techniques are analyzed for leveraging the structure of this new representation on quantum computers. Empirical results for hydrogen chains suggest a scaling improvement from $O(N^{4.5})$ in molecular orbital representations to $O(N^{2.6})$ in our representation for quantum evolution in a fault-tolerant setting, and exhibit a constant factor crossover at 15 to 20 atoms. Moreover, we test these methods using modern density matrix renormalization group methods classically, and achieve excellent accuracy with respect to the complete basis set limit with a speedup of 1–2 orders of magnitude with respect to using the primitive or Gaussian basis sets alone. These results suggest our representation provides significant cost reductions while maintaining accuracy relative to molecular orbital or strictly diagonal approaches for modest-sized systems in both classical and quantum computation for correlated systems.

## 1. Introduction

Predicting properties of both molecular and extended systems from first principles has long been the goal of electronic structure in both correlated classical methods [1], including new approaches based on tensor networks [2–5], and now many approaches based on quantum computing [6–15], some of which have even been implemented on experimental devices [16–25]. A crucial aspect of such simulations is the representation of the problem in a tractable discretization scheme such as a finite-difference method, or (more commonly) a basis set, also known as a Galerkin discretization of the problem. The selection of the basis influences not only the accuracy of the calculation but also the fundamental scaling of the cost of the simulation as a function of the system size.

**Figure 1.** A cartoon schematic of the general objective of this work depicted in terms of the sparsity pattern of the two-electron integrals. At the top, we depict a dense two-electron integral tensor with $O(N_a^4)$ nonzero two-electron integrals, but relatively few basis functions. This is often the case with molecular orbital or diffuse Gaussian basis sets. In the center we depict a diagonal primitive basis set, such as Gausslets, with a relatively high number of basis functions but an overall scaling in $O(N_p^2)$. Finally, we depict this for a discontinuous Galerkin (DG) basis set built from a diagonal primitive basis set that interpolates between the two regimes by using fewer basis functions than strictly diagonal sets, while retaining an overall $O(N_d^2)$ scaling through its block-diagonal structure. We note that the form of block-diagonal structure is a particular sparsity pattern which depends on the arrangement of indices when considering a matrix form; the text defines this sparsity pattern precisely.

The design of basis sets for correlated electronic structure has a long and rich history, packing much of the essential physics and chemistry into very compact representations to exploit the power of existing methods. While these basis sets have ranged from general purpose to those optimized for individual computations, a common mainstay has been the use of Gaussian-based molecular orbitals [1, 26–28]. Molecular orbitals tend to offer compact representations of correlated problems and also present an energy ordering of orbitals that facilitates further reduction of the space through active-space methods. However, a side effect of this reduction in the number of orbitals is often a Hamiltonian with a quartic number of terms and orbitals that are delocalized in space (figure 1).

The utility of localizing orbitals in both the occupied and virtual space has been recognized in the development of many linear-scaling methods [29], such as those based on pair natural orbitals [30–32]. In these cases the localization helps not only in allowing the screening of some terms in the Hamiltonian, but also reducing the correlations that need to be treated to reach a certain level of accuracy. However, even the most localized orbitals that can be produced from transformations of a standard basis are often not strictly local in the sense that they have heavy tails that can extend throughout the system to enforce orthogonality between orbitals. While reasonably compatible with some methods, these tails negatively impact tensor-network methods such as the density matrix renormalization group (DMRG), where an area-law entanglement system that is solvable in modest polynomial time becomes a volume law system with exploding bond dimension in a completely delocalized basis [33]. This effect, in combination with consideration of the number of Hamiltonian terms, led to the recent development of Gausslet basis sets [34, 35], which use properties of wavelet transforms to maintain strict localization, orthogonality, and a diagonal Coulomb operator.

In the case of quantum computing-based approaches, while basis localization may impact the representational power of an ansatz-based variational approach, the structure and number of terms in the Hamiltonian represent the dominant cost factor [36–39]. Moreover, some traditional density-based truncations and localizations can be difficult to utilize in quantum computers, due to the need to measure the density and rotate the basis to maximize the benefits. Recent advances in quantum-computing approaches have shown that Hamiltonians with a diagonal structure with a quadratic number of terms allow algorithms to execute numerically exact real-time dynamics and perform full configuration interaction (assuming a reference state with nonvanishing overlap on the ground state can be prepared) with a cost in terms of basis size that scales as roughly $O(N_p^{1/3})$ in first quantization [40] or $O(N_p^2)$ in second quantization [41]. While the first work was done for periodic systems based on basis sets related to discrete variable representations (DVR) [36], these results apply to all basis sets with these properties. This is in contrast to the most advanced methods based on Gaussian molecular orbitals on quantum computers,

| | |
|---|---|
| $N_p$ | Primitive basis functions in diagonal basis |
| $N_a$ | Functions in compact active space basis |
| $N_b$ | Blocks in the DG basis |
| $n_\kappa$ | Functions in DG block $\kappa$ |
| $N_d$ | Total DG functions, $(\sum_\kappa n_\kappa)$ |

**Figure 2.** Compact description of the notation used throughout the paper in counting basis functions in different representations for the electronic structure problem. Here discontinuous Galerkin (DG) is the block basis we construct from primitive functions to represent the active-space orbitals with a block-diagonal Hamiltonian representation.

which have costs more like $O(N_a^3)$ in first quantization [42] or $O(N_a^4)$ in second quantization [38] (see figure 2 for an overview of set size abbreviations used throughout the article).

While the scaling advantages of basis sets with diagonal representations appear at reasonable system sizes, reduced representational flexibility can mean the cost of switching representations for equivalent accuracy is still higher than current quantum computers with limited resources can afford. Thus, it is desirable to be able to split the difference between strictly diagonal basis sets and molecular orbitals. The adaptive local basis set [43–47] was recently introduced to achieve this tradeoff in the setting of density functional theory (DFT). The adaptive local basis is constructed on-the-fly to capture atomic and environmental effects, by solving eigenvalue problems restricted to local domains called elements. Each basis function is only supported on one element and is discontinuous on the level of the global domain. These basis functions are 'glued' together to approximate the continuous electron density using the interior penalty discontinuous Galerkin (DG) formalism [48, 49]. Compared to typical numerical methods for solving DFT, the DG formulation includes extra correction terms to handle the discontinuities of the basis functions at the boundary of the elements. Numerics indicate that roughly 10 to 40 basis functions per atom are often sufficient to achieve chemical accuracy for DFT calculations using pseudopotentials [50, 51].

In this work, we introduce a variation of the DG approach for simulation on quantum computers, which has two main differences compared to previous works [43, 44]. First, our target is to model electronic structure at the correlated level instead of the DFT level, and hence we can afford to obtain the best local basis functions, e.g., by starting from molecular orbitals or approximate natural orbitals. This removes a significant step of approximation in the original DG approach due to the solution of local eigenvalue problems with certain artificial boundary conditions (such as the Dirichlet or periodic boundary conditions). Second, our basis functions are represented as linear combination of primitive basis functions, which means that the basis functions are no longer strictly discontinuous in real space. We demonstrate that the two-body operator can still maintain a block-diagonal structure for efficient quantum simulation, and for simplicity we will still refer to the basis set as the DG basis set. The use of the primitive basis functions removes the need for correction terms that account for the discontinuity, which is similar in spirit to the discrete discontinuous basis projection method [52] for DFT calculations. The DG basis set enables one to interpolate between cheap diagonal representations and compact nondiagonal basis sets with blocks of specified size.

We begin by introducing the standard discretization of the electronic structure problem in a basis, and define precisely what is meant by a diagonal basis representation and strictly localized functions. We then briefly review basis sets that exhibit spatial locality and diagonal interactions that have already been used in the context of DMRG and quantum computing. This allows us to highlight the cost advantages of these representations within each approach. Then, the DG approach is introduced as a general framework for maintaining these properties by using those basis sets as a primitive building block. The block diagonal structure of the resulting Hamiltonians leads us to new cost models based on swap networks for quantum algorithms. We show that the new approach both maintains accuracy in correlated calculations, and demonstrates a crossover to dramatically lower costs at modest system sizes between 15 and 20 atoms. We finish with an outlook on how this approach will influence quantum and classical approaches to correlated electronic structure alike.

## 2. Discretizing the electronic structure problem

A crucial aspect of essentially all algorithms for the simulation of electronic systems is the discretization of the system into some tractable representation. This step takes the electronic Hamiltonian that acts in some continuous space, and maps it to a discrete space. The continuous electronic structure Hamiltonian is in atomic units given by

$$\hat{H} = -\sum_i \frac{\nabla_{\mathbf{r}_i}^2}{2} - \sum_{I,j} \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_j|} + \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + E_{\mathrm{II}}, \tag{1}$$

where r$_i$ denotes the position of the *i*th electron, and we have assumed the Born–Oppenheimer approximation such that the positions of nuclei, denoted by $\mathbf{R}_I$ are constants (giving rise to a constant energy correction $E_{II}$ for the nuclear–nuclear interaction). In the case of Galerkin discretizations, where one chooses a basis set given by some set of orthonormal (single-particle) functions $\{\chi_i(\mathbf{r})\}$ (for simplicity we assume a spin-restricted formulation), and enforces the antisymmetry of electrons in the operators. Then, one may express the Hamiltonian in the standard second quantized form. It is possible to select basis functions from a number of complete sets that allow the electron–electron interaction to be represented in a way that is entirely diagonal under a Jordan–Wigner representation in the computational basis and also exhibits a diagonal property under matricization. In this representation, the second quantized Hamiltonian is given by

$$\hat{H}^{(p)} = \sum_{\mu,\nu=1}^{N_p} h_{\mu\nu}^{(p)} \hat{b}_\mu^\dagger \hat{b}_\nu + \frac{1}{2} \sum_{\mu,\nu=1}^{N_p} v_{\mu\nu}^{(p)} \hat{n}_\mu \hat{n}_\nu, \tag{2}$$

where $\hat{b}_\mu^\dagger$ is a creation operator in the primitive basis, and $\hat{n}_\mu = \hat{b}_\mu^\dagger \hat{b}_\mu$ is a number operator. We refer to Hamiltonians written in this form as 'diagonal Hamiltonians', and use such basis sets here as our 'primitive' basis (associated with the superscript $(p)$ in $\hat{H}^{(p)}$) to efficiently construct compact bases that partially retain this property.

Generally, the coefficients in equation (2) are given by the following integrals,

$$h_{\mu\nu}^{(p)} = \int d\mathbf{r}\, \overline{\chi}_\mu(\mathbf{r}) \left( \frac{-\nabla_\mathbf{r}^2}{2} - \sum_I \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}|} \right) \chi_\nu(\mathbf{r}),$$

$$v_{\mu\sigma\gamma\nu} = \int d\mathbf{r}d\mathbf{r}'\, \frac{\overline{\chi}_\mu(\mathbf{r})\overline{\chi}_\sigma(\mathbf{r}')\chi_\gamma(\mathbf{r})\chi_\nu(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \tag{3}$$

and the defining property of our primitive basis sets may be written as $v_{\mu\sigma\gamma\nu} \to v_{\mu\nu}^{(p)} \delta_{\mu,\gamma} \delta_{\sigma,\nu}$, i.e., $v$ becomes a diagonal matrix when we view $(\mu, \sigma)$ as the row index and $(\gamma, \nu)$ as the column index, respectively. Note that the relation between $v_{\mu\sigma\gamma\nu}$ and $v_{\mu\nu}^{(p)} \delta_{\mu,\gamma} \delta_{\sigma,\nu}$ is not necessarily an equality: what one requires is that solutions to the Schrödinger equation using the two different forms of the interaction can systematically approach each other as one approaches the complete basis set limit. A grid, defined via finite differences, has the diagonal property, although there is no underlying basis. Here we consider only basis sets, but the basis functions with the diagonal property are naturally associated with a uniform or nonuniform grid, with typical spacings between grid points much smaller than the interatomic distances.

For a basis set, a sufficient condition for equality of these expressions is that one has functions with strictly disjoint support, such that (formally) $\chi_\mu(\mathbf{r})\chi_\sigma(\mathbf{r}) = 0$ for all $\mu \neq \sigma$. However, such a requirement would be unrealistic because upon careful inspection it would imply a simultaneously diagonal kinetic and potential operator if evaluated in the Galerkin formulation in equation (3) (in contrast to a finite-difference or overlapping finite-element approach). Classical finite-element methods can produce near-diagonality by allowing overlapping elements between only spatially neighboring sites, which retains the favorable scaling, but generically introduces nonorthogonality in the basis [53–56]. Accordingly, methods requiring a return to orthogonality, such as quantum-computing methods, suffer a transformation to $O(N_p^4)$ terms on orthogonalization, the introduction of unnecessary orthogonality tails, or both. More recent developments in classical discretization have extended this idea of disjoint cells to allow for truly disjoint basis sets, but at the cost of the introduction of additional surface terms and discontinuity penalties to reintroduce physicality into the problem. These methods are known as discontinuous Galerkin (DG) methods [48, 49] and were first introduced to electronic structure in the context of density functional theory [43–46]. They represent a general and rigorous framework for constructing problem representations that have this property of strict (block) locality, and we develop a variation of these methods in this work to achieve this goal for correlated electronic structure methods without the need for the introduction of surface terms.

While the condition of spatial disjointness is a sufficient condition to obtain a Hamiltonian with $O(N_p^2)$ terms, it is not necessary. Two basis sets that have the diagonal property without the equality of the individual matrix elements, and thus without the strict spatial disjointness property, are the plane-wave dual basis [36] and Gausslet basis [34]. The plane-wave dual basis is related to periodic sinc functions and discrete variable representations (DVR) [36] and is constructed from an aliased Fourier transform of plane waves for a given box to yield a diagonal Hamiltonian. Here the diagonality originates from the plane wave dual functions (also called the periodic sinc functions) that are Lagrange interpolation functions on a uniform grid. It has the advantage that the basis is naturally periodic, and thus is well suited for the treatment of materials and other condensed phase systems. Moreover, a modification using a truncated Coulomb interaction can enable the treatment of isolated systems. Two key downsides of this approach are

that it inherently reflects a uniform discretization in space of the problem, lacking the ability to adaptively refine sharp features such as the electron–nuclear cusp, and that it has long tails responsible for maintaining the orthogonality of the basis. This first property is reflected in a large overhead for representing atomic systems to a level of accuracy similar to Gaussians, and the second makes the method difficult to use with geometric entanglement based approaches such as tensor networks.

When integrated with smooth functions, Gausslets behave like $\delta$ functions, which results in the diagonal property. The Gausslets are obtained as linear combinations of arrays of Gaussians using wavelet transforms. Specific moment properties of the wavelet transforms make the Gausslets integrate polynomials up to a certain order like a $\delta$ function. The $\delta$-function property survives under smooth coordinate transformations, so unlike plane-wave dual bases, Gausslets can have variable resolutions, with more degrees of freedom near the nuclei to represent the electron–nuclear cusp. Also in contrast to the plane-wave dual basis functions, Gausslets have strong localization characteristics, avoiding long tails. The lack of tails means these basis sets are naturally suited for tensor-network methods and, relatedly, a variational quantum ansatz may have more expressive power at shorter circuit depths. The ability to more naturally represent inhomogeneous features means the representational overhead is expected to be modest relative to plane-wave representations.

We note, however, that both methods can benefit from the introduction of pseudopotentials, and that the representational power of all single-particle basis sets are expected to be limited by the same asymptotic scaling in the limit of a very large number of basis functions. This means that for very large basis set sizes approaching the complete basis set limit, the representational overheads are expected to be negligible, and the determining factors are the other properties of the basis sets. Despite this, however, there is considerable interest in treating systems before reaching this limit, where the representational overhead for basis sets can differ considerably.

In order to meet the demands of compactness one could start from a more generic basis set and use the expressions in equation (3) to determine the Hamiltonian. However, an approach that will prove fruitful here is to use the fact that it is equivalent to start from a complete primitive basis set, such as those above, and project into a compact 'active-space' Hamiltonian. Gaussian, molecular orbitals, or other active-space constructions can been seen as a specific case of the active space we refer to here.

Since the number of primitive basis functions is typically large compared to the number of electrons, quantum chemistry calculations, especially at the correlated level, commonly utilize more physically motivated basis sets that result in a more compact description (e.g. correlation-consistent basis sets in the LCAO approach). Let $\{\varphi_p(\mathbf{r})\}_{p=1}^{N_a}$ be a set of orthonormal single-particle functions; we refer to them as active-space orbitals (for instance, canonical Hartree–Fock orbitals, or natural orbitals), which can be expanded using a primitive basis set as

$$\varphi_p(\mathbf{r}) = \sum_\mu \chi_\mu(\mathbf{r}) \Phi_{\mu p}. \tag{4}$$

Here $\Phi \in \mathbb{C}^{N_p \times N_a}$ is a matrix with orthogonal columns. While this projection step represents an approximation depending on the nature of the primitive and active bases, it is one that can be systematically controlled and understood by increasing the size of the underlying primitive basis without significantly increasing the size of the resulting block basis that we will construct subsequently. For flexibility and accuracy, we will allow quite general definitions of the active-space basis set in this work. It will pertain to traditional Hartree–Fock canonical orbitals as well as Gaussian basis sets such as the dunning cc-pVDZ basis set [26], which allows us to use Gaussians with strictly block diagonal properties by going through the primitive basis set using point sampling through the approximate delta function properties of the primitive basis set. It has been shown previously that for the Gausslet primitive basis set we use, point sampling is extremely accurate due to the delta function property of the basis [34]. We will also make use of this construction to build hybrid active spaces, where we use a weighted density matrix from multiple basis sets to define $\Phi$ through its most important natural orbitals. In particular, we will combine the expressive power of a large primitive basis, such as Gausslets, to capture static correlations through an unrestricted Hartree–Fock defined active space, while including a Gaussian basis empirically refined to express dynamic electronic correlation effects, e.g. cc-pVDZ through weighting. In this work we make use of a joined set of density matrices built from a unrestricted Hartree–Fock (UHF) solution $D^{\mathrm{UHF}}$ and Gaussian orbitals $\alpha D^{\mathrm{Gaussian}}$, to form $D = D^{\mathrm{UHF}} + \alpha D^{\mathrm{Gaussian}}$ in the primitive basis, and use a natural-orbital truncation of $D$ to define $\Phi$. Empirically we use a value of $\alpha \approx 0.01$ later in this work when we combine these basis sets, which appears to give an excellent improvement in accuracy. A more detailed description of this procedure and refinement of the value $\alpha$ is left to a future work. We term this the hybrid active-space approach. As we only use this hybrid approach in conjunction with the DG-blocking procedure, we offload concerns about orthogonality in the projected basis to the singular value decomposition (SVD) used in the DG procedure.

Taking as granted the construction of the matrix $\Phi$, we define a rotated set of creation and annihilation operators in the active space as

$$\hat{a}_p^\dagger = \sum_{\mu=1}^{N_p} \hat{b}_\mu^\dagger \Phi_{\mu p}, \quad \hat{a}_p = \sum_{\mu=1}^{N_p} \hat{b}_\mu \overline{\Phi}_{\mu p}, \tag{5}$$

where $\overline{\Phi}_{\mu p}$ denotes the complex conjugate and we may project the Hamiltonian as

$$\hat{H}^{(a)} = \sum_{p,q=1}^{N_a} h_{pq}^{(a)} \hat{a}_p^\dagger \hat{a}_q + \frac{1}{2} \sum_{p,q,r,s=1}^{N_a} v_{pqrs}^{(a)} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r \hat{a}_s, \tag{6}$$

which we refer to as the active-space Hamiltonian.

Generally, we see that our primitive basis sets have favorable scaling in number of terms in the Hamiltonian ($O(N_p^2)$ vs $O(N_a^4)$), which often corresponds to better scaling algorithms. While $N_p$ and $N_a$ have the same asymptotic scaling, for modest sized calculations it is often observed that $N_p \gg N_a$ in order to achieve comparable accuracy. Here we will seek a way to split the difference between these two regimes by forming a more compact basis that partially retains the diagonal properties, i.e., the resulting Hamiltonian is block-diagonal.

## 3. Discontinuous Galerkin discretization

At a high level, we construct the block-diagonal basis by fitting spatially connected blocks of the primitive basis set to the active basis set, while preserving the properties of the primitive basis set. We therewith interpolate between the primitive basis set and the active basis set. We will refer to the general class of basis sets that achieve both completeness in some limit and have the diagonal property as primitive basis sets.

Our goal is to systematically compress the active basis set $\{\varphi_p(\mathbf{r})\}_{p=1}^{N_a}$ into a set of orthonormal basis functions partitioned into elements (groups), so that basis functions associated with different elements have mutually disjoint support. Assume that the index set $\Omega = \{1, \ldots, N_p\}$ can be partitioned into $N_b$ nonoverlapping index sets

$$\mathcal{K} = \{\kappa_1, \kappa_2, \ldots, \kappa_{N_b}\}, \tag{7}$$

so that $\cup_{\kappa \in \mathcal{K}} \kappa = \Omega$. Then the matrix $\Phi$ can be partitioned into $N_b$ blocks $\Phi_\kappa := [\Phi_{\mu p}]_{\mu \in \kappa}$ for $\kappa \in \mathcal{K}$ and $p \in \{1, \ldots, N_a\}$. Performing the singular value decomposition for $\Phi_\kappa$,

$$\Phi_\kappa \approx U_\kappa S_\kappa V_\kappa^\dagger, \tag{8}$$

where $U_\kappa$ is a matrix with orthonormal columns corresponding to the leading $n_\kappa$ singular values up to some truncation tolerance $\tau$, we obtain our compressed basis

$$\phi_{\kappa,j}(\mathbf{r}) = \sum_{\mu \in \kappa} \chi_\mu(\mathbf{r})(U_\kappa)_{\mu,j}. \tag{9}$$

The basis set is adaptively compressed with respect to the given set of basis functions, and are locally supported (in a discrete sense) only on a single index set $\kappa$. In the absence of SVD truncation, we clearly have $\text{span}\{\varphi_p\} \subseteq \text{span}\{\phi_{\kappa,j}\}$. We refer to this basis set $\{\phi_{\kappa,j}\}$ as the DG basis set. Note that each DG basis function $\phi_{\kappa,j}$ is a linear combination of primitive basis functions which are themselves continuous, so $\phi_{\kappa,j}$ is also technically continuous in real space. In fact, $\phi_{\kappa,j}$ might not be locally supported in real space if each primitive basis function $\chi_p$ is delocalized. When the primitive basis functions are localized, $\phi_{\kappa,j}$ can be very close to a discontinuous function. (See figure 3 for an example.) When computing the projected Hamiltonian, we do not need to evaluate the surface terms in the DG formalism. If we form a block-diagonal matrix
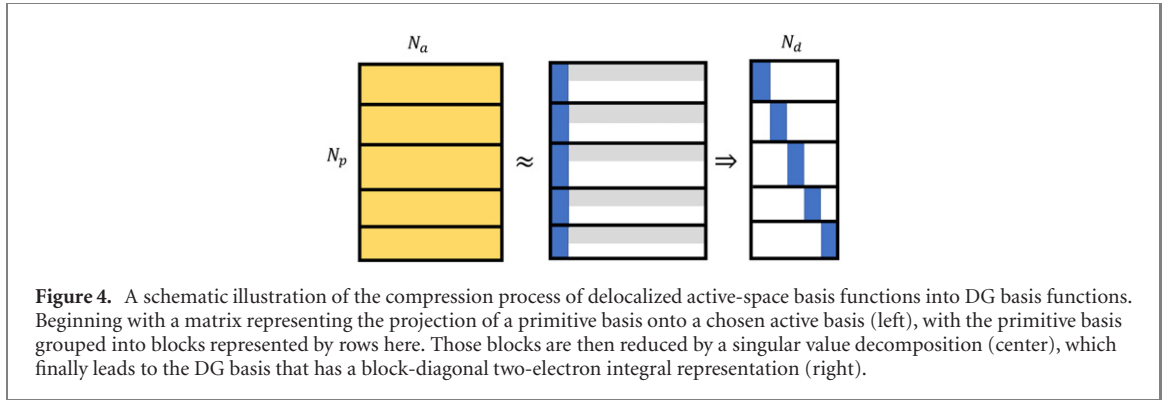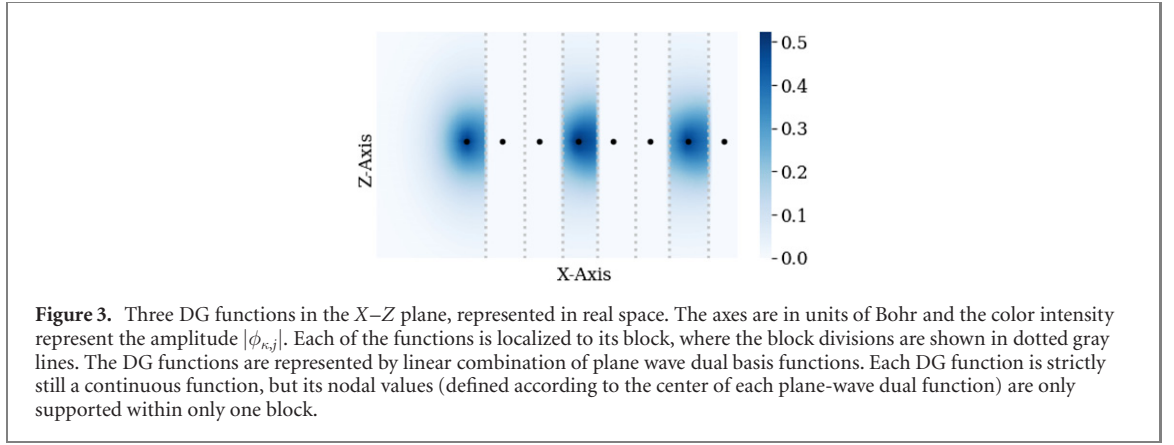
$$U = \text{diag}[U_1, \ldots, U_{N_b}], \tag{10}$$

the total number of basis functions is thus $N_d := \sum_{\kappa \in \mathcal{K}} n_\kappa$. We remark that the number of basis functions $n_\kappa$ can be different across different elements.

To facilitate the complexity count below we may, without loss of generality, assume that $n_\kappa$ is a constant and that $N_d = N_b n_\kappa$ (for a schematic illustration of the subsequent procedure see figure 4). Then we have defined a new set of creation and annihilation operators

$$\hat{c}_{\kappa,j}^\dagger = \sum_\mu \hat{b}_\mu^\dagger (U_\kappa)_{\mu j}, \quad \hat{c}_{\kappa,j} = \sum_\mu \hat{b}_\mu (\overline{U}_\kappa)_{\mu j}, \tag{11}$$

with $\kappa = 1, \ldots, N_b$ and $j = 1, \ldots, n_\kappa$ that correspond to the DG basis set.

**Figure 3.** Three DG functions in the $X–Z$ plane, represented in real space. The axes are in units of Bohr and the color intensity represent the amplitude $|\phi_{\kappa,j}|$. Each of the functions is localized to its block, where the block divisions are shown in dotted gray lines. The DG functions are represented by linear combination of plane wave dual basis functions. Each DG function is strictly still a continuous function, but its nodal values (defined according to the center of each plane-wave dual function) are only supported within only one block.



**Figure 4.** A schematic illustration of the compression process of delocalized active-space basis functions into DG basis functions. Beginning with a matrix representing the projection of a primitive basis onto a chosen active basis (left), with the primitive basis grouped into blocks represented by rows here. Those blocks are then reduced by a singular value decomposition (center), which finally leads to the DG basis that has a block-diagonal two-electron integral representation (right).

Unlike equation (5), the basis set rotation in equation (11) is restricted to each element $\kappa$. We readily obtain the projected Hamiltonian in the DG basis as

$$\hat{H}^{(d)} = \sum_{\kappa,\kappa';j,j'} h^{(d)}_{\kappa,\kappa';j,j'} \hat{c}^{\dagger}_{\kappa,j} \hat{c}_{\kappa',j'}$$

$$+ \frac{1}{2} \sum_{\kappa,\kappa';i,i',j,j'} v^{(d)}_{\kappa,\kappa';i,i',j,j'} \hat{c}^{\dagger}_{\kappa,i} \hat{c}^{\dagger}_{\kappa',i'} \hat{c}_{\kappa',j'} \hat{c}_{\kappa,j}. \tag{12}$$

The matrix elements are

$$h^{(d)}_{\kappa,\kappa';j,j'} = \sum_{\mu\nu} (\overline{U}_{\kappa})_{\mu j} h^{(p)}_{\mu\nu} (U_{\kappa'})_{\nu j'}, \tag{13}$$

and

$$v^{(d)}_{\kappa,\kappa';i,i',j,j'} = \sum_{\mu\nu} (\overline{U}_{\kappa})_{\mu i} (\overline{U}_{\kappa'})_{\nu i'} v^{(p)}_{\mu\nu} (U_{\kappa})_{\mu j} (U_{\kappa'})_{\nu j'}. \tag{14}$$

In general, the one-body matrix $h^{(d)}$ can be a full dense matrix, but the two-body tensor $v^{(d)}$ always takes a 'block diagonal' form in the following sense (it has a specific sparsity pattern). In principle, the two-body interaction in the DG basis set should take the form

$$\frac{1}{2} \sum_{\kappa,\kappa',\lambda,\lambda';i,i',j,j'} v_{\kappa,i;\kappa',i';\lambda,j;\lambda',j'} \hat{c}^{\dagger}_{\kappa,i} \hat{c}^{\dagger}_{\kappa',i'} \hat{c}_{\lambda',j'} \hat{c}_{\lambda,j}. \tag{15}$$

Compared to equation (12), we find that

$$v_{\kappa,i;\kappa',i';\lambda,j;\lambda',j'} = v^{(d)}_{\kappa,\kappa';i,i',j,j'} \delta_{\kappa\lambda} \delta_{\kappa'\lambda'}. \tag{16}$$

In other words, $v$ can be viewed as a block-diagonal matrix with respect to the grouped indices $(\kappa\kappa', \lambda\lambda')$.

We remark that the convergence of the DG basis set is independent of the choice of the primitive basis set so long as the primitive basis has sufficient degrees of freedom to form a good approximation to the active-space functions of interest. At the end of this adaptive procedure, we expect the number of elements in the Hamiltonian to scale as $O(N_b^2 n_\kappa^4)$. However, as we expect the number of basis functions required to reach a fixed accuracy within a block (i.e., $n_\kappa$) to be bounded by a constant as system size grows, and the

scaling, with system size becomes $O(N_d^2)$. We substantiate the rapid asymptotic convergence of $n_\kappa$ for real systems later in this work, however, simple arguments from spatial locality and basis set completeness lead to the same conclusion.

## 4. Quantum simulation with a DG basis

Here we introduce the methods used to exploit the properties of the DG basis on a quantum computer for evolution under the Hamiltonian. We first describe a method that implements evolution using a Suzuki–Trotter decomposition through swap networks, taking advantage of recent advances in quantum swap networks. An alternative Trotter approach based on low-rank decompositions is detailed in appendix A. We then describe a method for time evolution based on the linear combinations of unitaries approach, which will provide the required background for determining the degree of advantage for using DG basis sets for fault-tolerant quantum computations of chemistry.

### 4.1. Swap networks for block diagonal Hamiltonians

In the first work using a strictly diagonal basis in quantum computing for chemistry [36], the ability for quantum computers to perform fast Fourier transforms on quantum wavefunctions was exploited to capitalize on the representational advantages of being in either the plane-wave basis or its Fourier-transformed dual. That method was originally restricted to Hamiltonians with that particular structure in the coefficients, similar to split-operator Fourier transform methods used in classical simulation of quantum systems. However, it was soon realized that the structure of any diagonal Hamiltonian could be similarly exploited. This generalization used a linear, fermionic swap network to achieve perfect parallelization of a Trotter step with depth that scales linearly in the number of orbitals [57], even when gates are restricted to act on nearest neighbors of a line of qubits.

Fermionic swap networks are analogous to sorting networks from traditional computer science except built upon the primitive of the fermionic-swap operation,

$$\hat{f}_{\text{swap}}^{pq} = 1 + \hat{a}_p^\dagger \hat{a}_q + \hat{a}_q^\dagger \hat{a}_p - \hat{a}_p^\dagger \hat{a}_p - \hat{a}_q^\dagger a_q, \tag{17}$$
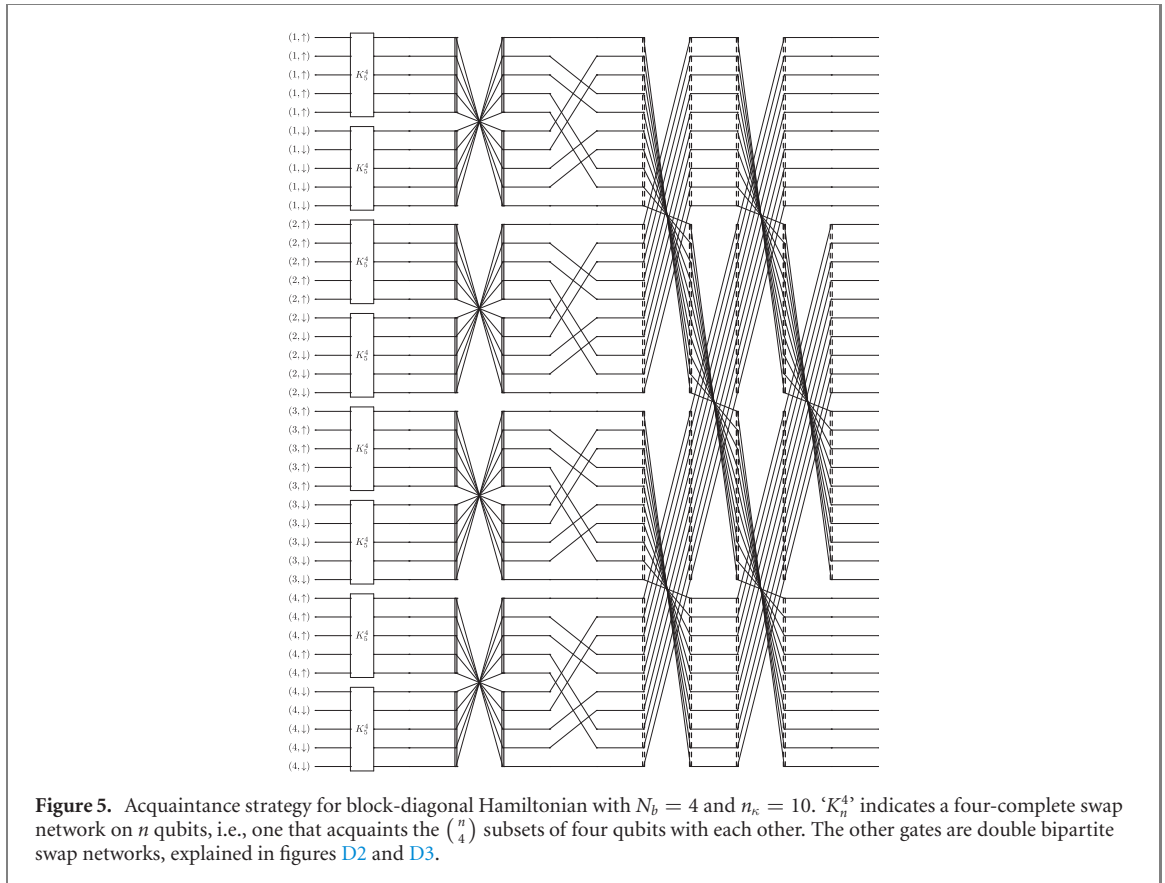
$$\hat{f}_{\text{swap}}^{pq} \hat{a}_p^\dagger (\hat{f}_{\text{swap}}^{pq})^\dagger = \hat{a}_q^\dagger, \tag{18}$$

where $\hat{f}_{\text{swap}}^{pq}$ is the fermionic swap that swaps the labeling of modes. The fermionic-swap operation was introduced in [58] and also studied in the context of tensor networks. The difference between such a swap and a traditional swap is by swapping fermionic modes instead of assignments to qubits, nonlocal parity strings used to enforce the fermionic anticommutation relations can be avoided.

The basic idea of the linear swap network is to fermionic swap all neighboring qubits, interact them with their current neighbors, and repeat until all qubits have interacted with each other. Since the introduction and use of these linear fermionic swap networks in quantum algorithms, they have been generalized for use in nondiagonal Hamiltonians with some overhead. For example, the quantum chemistry Hamiltonian can be decomposed into a sum of diagonal Hamiltonians (each in a rotated basis) using techniques similar to Cholesky or density-fitting methods, where each diagonal Hamiltonian can then be implemented in sequence [39]. We show how to use this method in the DG representation in appendix A.

For the general Hamiltonian in quantum chemistry, which is nondiagonal, a generalized swap network that works directly with such Hamiltonians was developed [59]. This network implements time steps for generic $O(N_a^4)$ Hamiltonians in a time that scales as $O(N_a^3)$, and we take advantage of it here with specializations for the block diagonal structure. To implement a Trotter step of the Hamiltonian, the swap network dynamically updates the mapping from qubits to orbitals so that for each term in the Hamiltonian, the involved orbitals are mapped to adjacent qubits. We say that a swap network 'acquaints' a set of orbitals when it brings them together at some point in this way, and represent that point by an empty box in the circuit diagrams, which acts as a placeholder for the logical gate to be executed there. Prior work utilizing swap networks has applied them to two extremal regimes with respect to the structure of the two-electron terms in the Hamiltonian: the strictly diagonal case, which can be implemented with $O(N_p)$ depth [57]; and the fully general case, which can be implemented in $O(N_a^3)$ [59]. Here we show how to interpolate between these to achieve $O(N_b n_\kappa^3) = O(N_d n_\kappa^2)$ depth for block-diagonal Hamiltonians. (For simplicity, in this section we will assume that all blocks have the same size $n_\kappa$, but the techniques generalize in a straightforward way to nonuniform block sizes.)

We focus on how to implement the quartic terms in the Hamiltonian (i.e., two-electron terms involving four distinct spin orbitals). The lower-order terms can be addressed with negligible additional resources by

**Figure 5.** Acquaintance strategy for block-diagonal Hamiltonian with $N_b = 4$ and $n_\kappa = 10$. '$K_n^4$' indicates a four-complete swap network on $n$ qubits, i.e., one that acquaints the $\binom{n}{4}$ subsets of four qubits with each other. The other gates are double bipartite swap networks, explained in figures D2 and D3.

incorporating them into the quartic terms. The quartic terms in the block-diagonal Hamiltonian satisfy the following properties:

(a) Two orbitals are from one block $\kappa$ and two orbitals are from another block $\kappa'$ (or all four from the same block when $\kappa = \kappa'$).

(b) The orbital spins have even parity (i.e., all up, all down, or two and two).

We will exploit both of these properties in constructing our swap network, which uses primitives originally designed for implementing unitary coupled cluster [59]. Figure 5 shows the overall swap network. Initially, the orbitals are arranged on the line in lexicographical ordering; only the block index $\kappa$ and spin are indicated for concision. The logic of the strategy is as follows:

(a) The first layer acquaints all sets of four spin orbitals within each block in which all four orbitals have the same spin. This is achieved by a 'four-complete' swap network on each half block of orbitals, denoted by $K_{n_\kappa/2}^4$ because the sets of orbitals it acquaints correspond to the edges of a complete four-uniform hypergraph; it has depth $O(n_\kappa^3)$. Note that the edges of the complete $k$-uniform hypergraph $K_n^k$ on $n$ are the $\binom{n}{k}$ sets of $k$ vertices. The 'uniform' qualifier indicates that all of the hyperedges have the same number of vertices. See [59] for details.

(b) The second layer acquaints all sets of four spin orbitals within each block in which two orbitals have spin up and the other two have spin down. This is achieved by a 'double bipartite' swap network on each block in depth $O(n_\kappa^3)$; see figure D2.

(c) The third layer permutes, in $O(n_\kappa)$ depth, the orbitals within each block in preparation for the inter-block acquaintances to follow.

(d) The rest of the strategy consists of $N_b$ alternating layers that acquaint pairs of parts. In each layer, each block of qubits is paired up with an adjacent one and a 'balanced double bipartite' swap network is executed on the pair of blocks; see figure D3. Each balanced double bipartite swap network acquaints the sets of four orbitals containing two from each block and with even ('balanced') spin parity. This also has the effect of swapping the blocks, so overall a balanced double bipartite swap network is applied to every pair of blocks. Each double bipartite swap network has depth $O(n_\kappa^3)$.

Overall, the depth is $O(N_b n_\kappa^3)$, dominated by the latter swap networks that effect the inter-block interactions. The components of this approach are explained in more detail in appendix D and an alternative approach that may have asymptotic advantages in some regimes is discussed in appendix A.

### 4.2. LCU approaches for simulation

The quantum simulation algorithms discussed in the previous section and in appendix A are useful for implementing Trotter steps of the chemistry Hamiltonian. Such Trotter steps can be repeated to perform time-evolution for modeling dynamics or for preparing eigenstates via the phase estimation algorithm [60, 61], but they can also serve as an ansatz for composing quantum variational algorithms [15, 62]. This in conjunction with these Trotter steps requiring only minimal (linear) connectivity makes them attractive algorithms for near-term quantum computing. However, within cost models appropriate for a fault-tolerant quantum computer, Trotter steps are not the most competitive technique for chemistry simulation. In such a cost model, the key resource to minimize is the number of T gates required by the algorithm. This is because in most practical error-correcting codes (e.g., the surface code), T gates require many physical qubits for their distillation and are orders of magnitude slower to implement than Clifford gates.

Currently, the lowest T complexity quantum algorithms for simulating chemistry are all based on linear combinations of unitaries (LCU) methods [63]. This family includes Taylor series methods [64] (applied to chemistry in [14, 42]), qubitization [65] (applied to chemistry in [38, 65, 66]) and Hamiltonian simulation in the interaction picture [41] (applied to chemistry in [40, 41]). What all LCU methods have in common is that they involve simulating the Hamiltonian from a representation where it can be accessed as a linear combination of unitaries,

$$H = \sum_{\ell=1}^{L} \omega_\ell U_\ell, \quad \lambda = \sum_{\ell=1}^{L} |\omega_\ell|, \tag{19}$$

where $U_\ell$ are unitary operators, $\omega_\ell$ are scalars, and $\lambda$ is a parameter that determines the complexity of these methods. The second-quantized Hamiltonians discussed in this paper satisfy this requirement once mapped to qubits (e.g., under the Jordan–Wigner transformation) since strings of Pauli operators are unitary.

LCU methods perform quantum simulation in terms of queries to two oracle circuits defined as

$$\text{SELECT}|\ell\rangle|\psi\rangle \mapsto |\ell\rangle U_\ell|\psi\rangle, \tag{20}$$

$$\text{PREPARE}|0\rangle^{\otimes \log L} \mapsto \sum_{\ell=1}^{L} \sqrt{\frac{\omega_\ell}{\lambda}}|\ell\rangle, \tag{21}$$

where $|\psi\rangle$ is the system register and $|\ell\rangle$ is an ancilla register which usually indexes the terms in equation (19) in binary and thus contains $\log L$ ancillae. Up to log factors in precision and other system parameters, LCU methods can perform time-evolution with gate complexity scaling as

$$O\left((C_S + C_P)\lambda t\right), \tag{22}$$

where $C_S$ and $C_P$ are the gate complexities of SELECT and PREPARE respectively, and $t$ is time.

To implement the LCU oracles one must be able to coherently translate the index $\ell$ into the associated $U_\ell$ and $\omega_\ell$. In the quantum chemistry context the $U_\ell$ are related to the second-quantized fermion operators (e.g., $\hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r \hat{a}_s$) and the $\omega_\ell$ are related to the molecular integrals. While the $U_\ell$ have a structure that is straightforward to unpack in a quantum circuit (see [38, 66] for explicit implementations), the $\omega_\ell$ are typically challenging to compute directly from this index (unless one pursues the highly impractical strategy of computing the integrals on-the-fly, as in [14]). As a consequence, with state-of-the-art LCU methods for simulating chemistry the primary bottleneck has been implementation of PREPARE rather than SELECT [38, 66].

In the most performant LCU approaches for chemistry (see [66] for plane waves and [38] for arbitrary basis sets), PREPARE is implementing (and bottlenecked) by using a data-lookup routine referred to as QROM (quantum read-only memory) to load the $\omega_\ell$ values into superposition. Using this approach, the cost of PREPARE is a function of the number of unique coefficients $\omega_\ell$ in the Hamiltonian. Using the QROM of [67] (which improves on the original concept from [66]), one can (up to log factors) implement PREPARE with T complexity scaling as $O(L/g + g)$ where $L$ is the number of molecular integrals and $g$ is a free parameter so long as at least $g$ dirty ancilla are available during the implementation of PREPARE. As PREPARE acts only on the ancilla register, there are typically at least $N_d$ dirty ancilla available (from the $|\psi\rangle$ register). Thus (assuming $L > N_d^2$) the scaling is $O(L/N_d)$ without ancilla or $\mathcal{O}(\sqrt{L})$ with $O(\sqrt{L})$ ancilla (often a reasonable tradeoff within error-correction).

When simulating in a molecular orbital basis with $L = O(N_a^4)$, one can (up to log factors) evolve the Hamiltonian for some time $t$ and achieve T complexity $O(N_a^2 \lambda t)$ [38]. If a straightforward extension of the approach in [38] is applied to the DG basis Hamiltonian with $L = n_\kappa^4 N_b^2$ then this scaling would be reduced to $O(n_\kappa^2 N_b \lambda t)$. However, if one can index the $\omega_\ell$ in a way that exploits symmetries in the coefficients, then one can further reduce the effective value of '$L$' in both of these expressions. For instance, in [66] it is recognized that while there are $N_p^2$ coefficients $\omega_\ell$ in the Hamiltonian of equation (2) (one for each value of

$\nu$ and $\mu$), there are only $N_p$ *unique* coefficients due to the translational invariance of the Coulomb operator (essentially, one for each unique value of the index $\mu-\nu$), and the cost of the algorithm is reduced accordingly. As a rough rule of thumb, the scaling becomes $O(J/g + g)$ where $J$ can be thought of as the number of unique scalars needed to completely describe the Hamiltonian without recomputing the molecular integrals. In [38], it is shown that because only $J = O(N_a^3)$ numbers are needed to describe the low-rank factorized Coulomb operator in an arbitrary basis [39], one can improve the scaling to $O(N_a^{3/2}\lambda t)$. By tailoring such techniques to symmetries such as periodic crystalline symmetry or other redundancies in the DG Hamiltonian (e.g., those exploited in the low-rank representation of appendix A) one can also further improve the $\mathcal{O}(n_\kappa^2 N_b \lambda t)$ scaling.

Finally, we note that the $\lambda$ value associated with the molecular orbital basis Hamiltonian is likely larger than the $\lambda$ value associated with the DG Hamiltonian, and we quantify the extent to which that is the later in this paper. This is yet another way in which the DG representation should lead to even more efficient implementation of LCU methods for chemistry simulation. Note also, that the value of $\lambda$ is important for the scaling of quantum simulation in a near-term cost model as well. In particular, for quantum variational algorithms the number of measurements (corresponding to the number of circuit repetitions) required to estimate the energy of a Hamiltonian to within error $\epsilon$ scales as $O(\lambda^2/\epsilon^2)$ [68].

## 5. Numerical results

To understand the performance of different discretization schemes, we examine the costs and associated accuracy for each of the methods, using hydrogen chains of increasing lengths as test systems. We show that the crossover point occurs around 15 to 20 atoms, above which the DG representation with the block-diagonal Hamiltonian structure has significant lower costs. For two-dimensional and three-dimensional hydrogen lattices, we expect that the number of basis functions will increase slightly from our experience at the DFT level [43, 51].

The subsequent calculations are twofold. We begin with numerical simulations designed to exhibit the expected crossover behavior. We investigate hydrogen chains of increasing length (up to $H_{30}$) in a Gaussian cc-pVDZ basis [26] (the active space to be fit), where we choose the primitive diagonal basis set to be plane wave dual functions with refinement built to match the accuracy of the Gaussian basis set to a specified tolerance at the level of density functional theory with the Perdew–Burke–Ernzerhof (PBE) exchange–correlation functionals [69]. We are not aware of any electronic structure software package using a plane-wave basis set for practical all-electron DFT calculations. Hence we use the pseudopotential formulation based on the ONCV pseudopotential [70]. As we fit to the span of the basis set itself, rather than a density matrix generated with a fixed level of theory, these results should be representative of performance across different levels of theory including, but not limited to, correlated methods. This is done to observe the crossover without extrapolation, as by the size the crossover occurs, the more accurate correlated calculations become prohibitively expensive using traditional classical methods.

In the second set of calculations, we instead fit to a natural-orbital active space using a Gausslet basis set as the primitive diagonal basis set [34, 35] to demonstrate that this representation maintains accuracy in correlated calculations. This is possible as the properties of a Gausslet basis allow accurate DMRG calculations for comparison on these systems. Similar DMRG calculations are expected to be more expensive in basis sets with heavy tails, such as the plane-wave dual basis, which makes the one-particle Hamiltonian a dense matrix, and result in excessive bond dimension requirements for accurate descriptions. These calculations exhibit the versatility of this method and its ability to achieve accuracy, even in a correlated active space built fit to a highly nonlocal active space. Furthermore, we demonstrate that by fitting the DG basis simultaneously to two active basis sets (molecular orbitals at the UHF level and the cc-pVDZ basis functions), the DG basis set can obtain accurate results when compared to quantum Monte Carlo calculations at the complete basis set limit, while achieving a 1–2 orders of magnitude reduction of computational cost over the primitive Gausslet basis or the Gaussian basis set.

### 5.1. Scaling crossover in a DG plane-wave dual basis

In order to demonstrate the performance of the DG approach for a long hydrogen chain, we use the DFT-based MATLAB toolbox KSSOLV [71]. KSSOLV uses a plane-wave discretization in combination with pseudopotentials to make larger systems computationally tractable. Figure 6 shows that chemical accuracy is achieved when the kinetic-energy cutoff ($E_{cut}$) is set to around 20 hartree for $H_{10}$ using the ONCV pseudopotential. The kinetic-energy cutoff needs to be much larger for all-electron calculations using the plane-wave basis set. With this kinetic-energy cutoff, we can perform calculations on hydrogen chains up to $H_{30}$. For $H_{30}$ the box size is $10 \times 10 \times 118\ a_0^3$, and the grid size is $20 \times 20 \times 238$, or 95 200 plane-wave dual basis functions (around 3000 basis functions per atom).
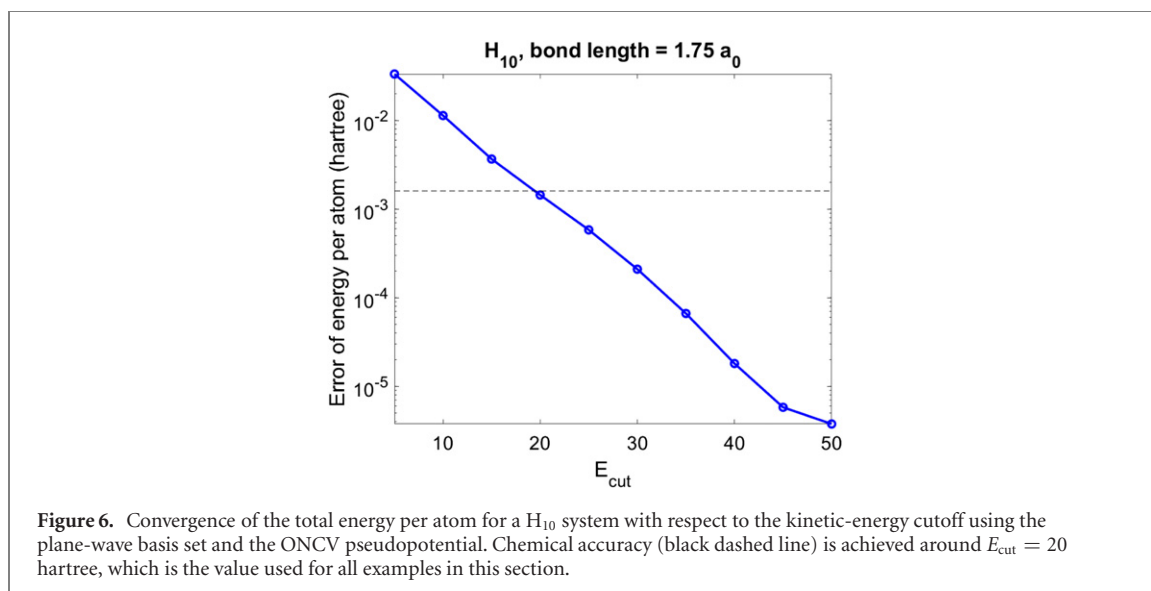
**Figure 6.** Convergence of the total energy per atom for a $H_{10}$ system with respect to the kinetic-energy cutoff using the plane-wave basis set and the ONCV pseudopotential. Chemical accuracy (black dashed line) is achieved around $E_{cut} = 20$ hartree, which is the value used for all examples in this section.
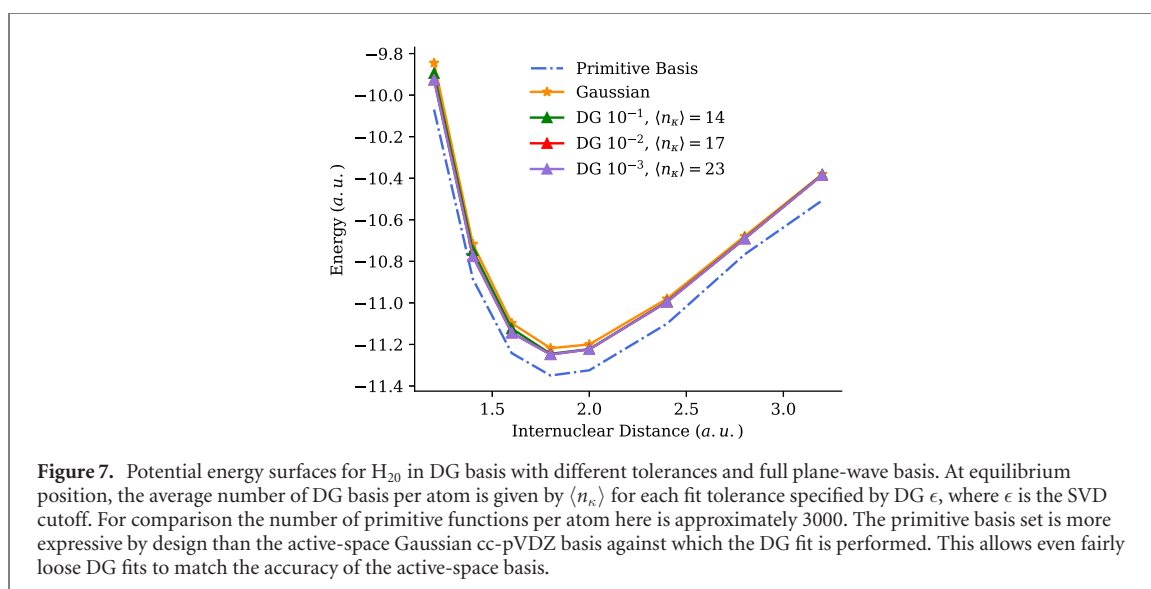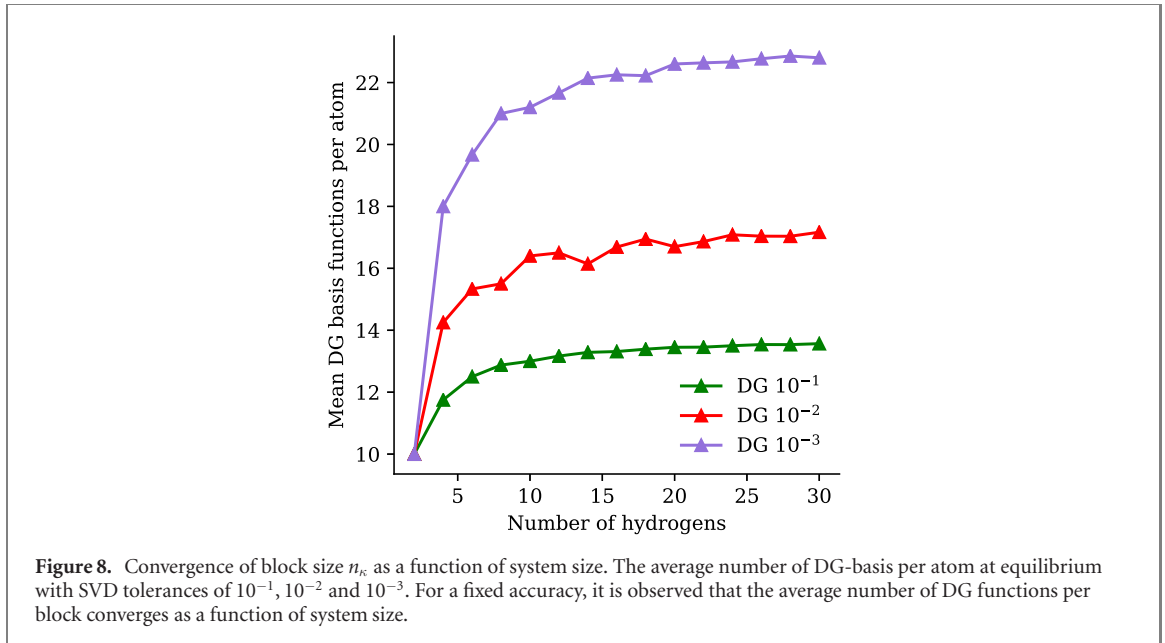


**Figure 7.** Potential energy surfaces for $H_{20}$ in DG basis with different tolerances and full plane-wave basis. At equilibrium position, the average number of DG basis per atom is given by $\langle n_\kappa \rangle$ for each fit tolerance specified by DG $\epsilon$, where $\epsilon$ is the SVD cutoff. For comparison the number of primitive functions per atom here is approximately 3000. The primitive basis set is more expressive by design than the active-space Gaussian cc-pVDZ basis against which the DG fit is performed. This allows even fairly loose DG fits to match the accuracy of the active-space basis.

We start by constructing an active basis set using the cc-pVDZ basis with the given molecular geometry, and sample it with an underlying grid of plane-wave dual functions. This yields a real-space discretization of the cc-pVDZ basis, which is then transformed into a block-diagonal form by means of a DG blocking procedure, also implemented in KSSOLV. As a technical note, the partitioning boundaries in the DG approach are important. The hydrogen chains that are the subject of this numerical investigation are quasi 1D-problems. This and the fact that we use a real-space discretization suggests that the ideal partitioning of the basis in terms of a DG-scheme (see section 3) corresponds to a nonuniform partitioning strategy, so that each hydrogen atom is approximately located at the center of each element. We remark that by construction the DG-blocking procedure is able to produce accurate results even if the partition is nonideal, e.g. when an hydrogen atom is located near the boundary of an element. However, this can require a larger number of basis functions per atom compared to the ideal partitioning strategy.

We first verify the accuracy of our DG-basis in figure 7 by comparing potential energy surfaces (PES) in DG discretizations of different tolerances with the corresponding PES in the primitive basis. The primitive basis has been chosen to be much more expressive than the Gaussian active-space basis, showing a lower absolute energy. This allows even a coarse DG fit to reliably match the Gaussian active-space accuracy, while being far more compact than the original primitive basis set. In fact, the energies obtained from the DG basis are slightly lower than those from the Gaussian basis set. This is because as $n_\kappa$ increases, the span of the Gaussian basis set becomes approximately a subspace of the span of the DG basis set, and hence the DG basis can possibly yield lower energies due to the variational principle. Also due to the variational principle, the energies obtained from the DG basis are noticeably higher than those from the primitive plane-wave

**Figure 8.** Convergence of block size $n_\kappa$ as a function of system size. The average number of DG-basis per atom at equilibrium with SVD tolerances of $10^{-1}$, $10^{-2}$ and $10^{-3}$. For a fixed accuracy, it is observed that the average number of DG functions per block converges as a function of system size.

basis set, and the main limiting factor is the cc-pVDZ basis set to which DG is fitted. The average number of functions per atom $\langle n_\kappa \rangle$ is shown for each tolerance and is approximately 20, as compared to the primitive basis which is built from roughly 3000 functions per atom. The results suggest that the overall energetic accuracy is relatively insensitive even to rather aggressive singular value truncations in the DG-blocking procedure.

We then plot the average number of DG-basis functions per atom in figure 8 for fixed SVD thresholds in the DG-blocking procedure. The data confirms that, as the system size grows, $\langle n_\kappa \rangle$ approaches a constant value for a fixed SVD-truncation threshold. The combination of the energetic insensitivity beyond a kinetic-energy cutoff of 20 hartree and the asymptotically constant $\langle n_\kappa \rangle$ strongly supports the existence of a crossover regime where a DG discretization is more cost effective than other representations. We confirm the crossover directly by examining the quantities most relevant for fault-tolerant cost models of chemistry on a quantum computer, i.e., the number of nonzero two-electron integrals in each representation and the $\lambda$ factor equation (B1) in figures 9 and 10, respectively (for $\lambda$ factors for different bond lengths see appendix B, figure B1). A threshold of $10^{-6}$ is used to count an individual integral as 0 when calculating these quantities empirically on the systems of interest. We choose to illustrate the cost crossover for the bond length $1.7a_0$ since we here detected the largest deviation with respect to the SVD-truncation threshold.

Considering the cost-model in the fault-tolerant setting outlined in section 4.2, one can already observe a scaling advantage for the DG representation over simple Gaussian based active-space representations. Recall from that section that the cost using an LCU method to evolve for some time $t$ is roughly $O(\sqrt{L}\lambda t)$ where $L$ is the number of nonzero terms in the Hamiltonian. We consider here only the two-electron integrals as they represent the dominant cost contribution in most cases. For molecular orbital representations, this is generally $O(N_a^2 \lambda t)$, which was subsequently improved to $O(N_a^{3/2} \lambda t)$ using low-rank structure in the problem. For the hydrogen chains examined here in the Gaussian basis set, we see empirically over the system sizes considered $L \propto N_h^4$ and $\lambda \propto N_h^{2.5}$, where $N_h$ is the number of hydrogen atoms in the chain, leading to an expected cost scaling of $O(N_h^{4.5}t)$ when not exploiting further low rank structure. In the DG representation for the same problem, we see that $L \propto 2.25$ and $\lambda \propto 1.5$, which suggests an empirical scaling for this physical system of $O(N_h^{2.6}t)$. The same calculation for the primitive basis suggests a scaling of $O(N_h^{1.8})$, which is the lowest seen, however, the simulation cost and qubit counts required are many orders of magnitude higher here as seen in the figures. In both quantities we observe a constant factor crossover where the DG representation has strictly lower unique two-electron integrals and $\lambda$ factor when compared to the Gaussian active-space basis at around 15 to 20 hydrogen atoms. This suggests that the DG representation is not only advantageous in the asymptotic scaling, but also for modest finite system sizes for fault-tolerant implementations.

## 5.2. Correlated calculations in a DG Gausslet basis

Here we demonstrate that the performance of the DG basis set for a vastly different regime, which uses a natural-orbital active space from an exact DMRG calculation fit to Gausslet primitive functions in a correlated calculation. The Gausslet basis set is a recent approach to improve the discretization of quantum
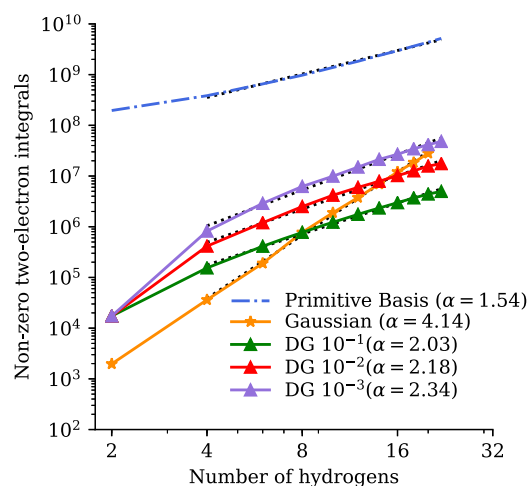
**Figure 9.** The number of nonzero two-electron integrals in different representations, for equilibrium and dissociation bond lengths with SVD-truncation tolerances of $10^{-1}$, $10^{-2}$ and $10^{-3}$, plotted on a log–log scale. We fit a trendline plotted with black dots from the second point onward to extract the scaling as a function of system size as $N^{\alpha} + c$ for some constant $c$, and list the exponent $\alpha$ beside each representation in the legend. As predicted, for these system sizes the number of two-electron integrals lies between the primitive and active-space representations, tending closer to the $O(N^2)$ scaling of the primitive representation, requiring fewer functions. Note that for the Gaussian basis set, certain elements of the two-electron integrals can vanish due to the symmetry of the atomic configuration of the hydrogen chain. This has a larger impact for small system than for large systems, and therefore the scaling is observed to be slightly larger than $O(N_a^4)$.
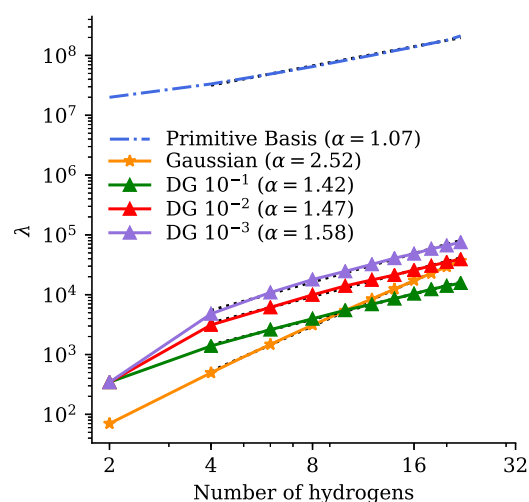


**Figure 10.** $\lambda$ value for Gaussian and DG basis. A core quantity in determining the cost in quantum algorithms, $\lambda$, is plotted as a function of system size for different representations. The notation DG $\epsilon$ indicates an SVD truncation threshold of $\epsilon$ in the blocking procedure. We observe an advantageous crossover before or around $H_{20}$ in all cases with respect to an actual value. We fit a trendline plotted with black dots from the second point onward to extract the scaling as a function of system size as $N^{\alpha} + c$ for some constant $c$, and list the exponent $\alpha$ beside each representation in the figure legend. We see the scaling for the DG basis is significantly better in all cases than for the active-space basis as well.

chemical problems [34, 35]. It has a special focus on sparsity, spatial locality, and orthonormality, to fulfill the needs of strong-correlation methods like the DMRG, while keeping the number of basis functions lower than other grid-based bases. Our calculations illustrate the accuracy of the coupled-cluster method with single and double excitations (CCSD) method in a DG-basis with an underlying primitive Gausslet basis set with respect to the original active space. The block-diagonal form of the Hamiltonian in a DG-basis yields an asymptotic improvement in the number of nonzero two-electron integrals, which is directly related to the circuit depth and size. The calculations presented here are limited to short hydrogen chains due to the size of the Gausslet basis set and expense of correlated calculations. The goal of these calculations is to demonstrate that the DG approach maintains accuracy for correlated calculations due to its construction as a fit to the span of the active basis.

Specifically, we compare CCSD-energy results in an active-space basis set with calculations in a DG-basis fit to the same active-space basis. This serves to quantify the overhead in restricting the basis to have block
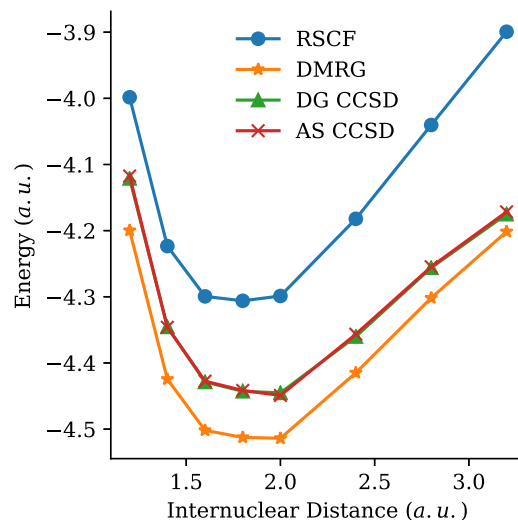
**Figure 11.** Correlated potential-energy surfaces for $H_8$. We plot the potential-energy surface computed using the Gausslet basis as the primitive basis and a subset of the exact DMRG natural orbitals as an active space (AS). The CCSD energy is calculated in the active space as well as the DG fit to that active space, showing excellent agreement with the active-space energetics in a correlated calculation. The exact DMRG energy in the more expressive, full primitive basis (Gausslets) is shown for reference.

locality versus the totally delocalized natural-orbital active space. The calculations are performed for $H_2$, $H_4$, $H_6$ and $H_8$ with varying symmetrically stretched bond lengths. Figure 11 shows the calculation for $H_8$ with optimal DG partitioning of the Gausslet basis set, the respective figures for $H_2$, $H_4$, and $H_6$ are presented in appendix C (figures C1–C3). From these plots one may see that enforcing a block-diagonal structure of the Hamiltonian does not reduce the accuracy further from the active-space approximation.

The Gausslet basis for computations resulting in figure 11 consists of 1336 Gausslets corresponding to an average of 167 functions per atom. This is significantly lower than the number of basis functions per atom in the plane-wave dual basis which reflects the variable resolution available with the Gausslet basis. Averaging over all values along the PES, we find that the full two-electron integral tensor for this basis has $\sim 2\,310\,318$ nonzero elements. Using a DG basis it is possible to reduce this number to $\sim 511\,449$ nonzero two-electron integrals without losing accuracy compared to the respective active-space approximation. This suggests cost reductions for correlated calculations that are similar to the observations in section 5.1. Note, however, that $H_8$ is a small system compared to the systems considered in section 5.1, consequently, the number of nonzero two-electron integrals in the DG-basis is still larger than in the active-space approximation ($\sim 52\,346$). This aligns with the results from the previous section, which showed that the improvement of the nonzero two-electron integral count for the DG-basis becomes observable for 15 to 20 atoms, depending on the imposed truncation tolerance. A naive extrapolation of the nonzero two-electron integral count in table C2 (see appendix C figure C4) suggests a crossover around 25 atoms for the coupled-cluster calculations, which is again in agreement with the computations in section 5.1 with a low truncation tolerance (see appendix B table B1). We conclude that the trial calculations for small systems performed here together with results from section 5.1 suggest that the DG approach maintains the accuracy of active-space approximations for correlation methods but with a more efficient representation on quantum devices once a certain system size is reached.

**5.3. DMRG calculations in a DG basis**

Here we examine the power of the DG technique to compactly and cheaply represent problems for classical, correlated DMRG calculation by constructing an active space that takes static correlation from a UHF calculation done in a flexible primitive basis set such as Gausslets, combined with contributions from a basis set that has been empirically refined to capture dynamic correlations. These calculations have the dual purpose of demonstrating the power of the DG approach even in highly-correlated calculations more generally, including calculations on a quantum computer. Here we demonstrate a hybrid active-space approach for cheaply finding a balanced active-space. Specifically, after performing a UHF calculation in a Gausslet basis, we project both the cc-pVDZ basis set and the UHF orbitals onto the primitive Gausslet basis. The UHF calculation can be performed cheaply. By including the UHF orbitals in our active-space, we ensure that there are no HF-level errors in the basis. The only lack of completeness is associated with correlation beyond HF. To (partially) capture correlation in a compact way, we include contributions from the empirical Gaussian basis with a slightly smaller factor, $\alpha = 0.01$ here, then perform the DG blocking
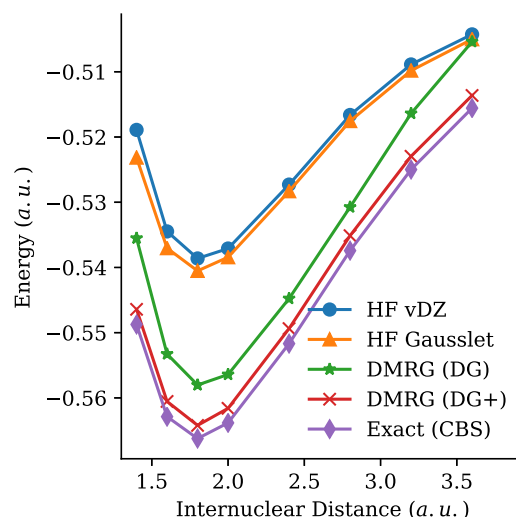
**Figure 12.** Potential-energy surfaces for $H_{10}$ constructed with the hybrid approach. We plot the potential-energy surface for unrestricted Hartree–Fock (here denoted HF) calculations in both a cc-pVDZ (vDZ) Gaussian basis and a Gausslet basis, from which the DG basis sets are derived to perform a DMRG calculation. We then contrast this with a calculation done in the DG basis using the first 7 virtuals from the HF–DG basis (DG), and finally showcase the power of the hybrid approach to cost-effectively reintroduce dynamic correlation without the cost of the original Gaussian basis by adding a point sampling of the cc-pVDZ basis into the DG basis (DG+). This hybrid basis attains nearly the exact solution with respect to the complete basis set limit with a fraction of the cost of using either the Gausslet or Gaussian basis set directly.

procedure to develop the basis as before. The resulting basis maintains the flexibility of the HF solution in the Gausslet basis for the low-lying orbitals, while now including the refined features of the Gaussian orbitals without losing the block-diagonal structure. Although it is difficult to compare the efficiency of primitive and DG bases in a precise manner, we find that this approach can achieve a 1–2 orders of magnitude reduction of computational cost over either the primitive Gausslet basis or Gaussian basis set, while achieving excellent accuracy with respect to the complete basis set limit.

Numerical calculations for the $H_{10}$ system are shown in figure 12. For comparison, we perform an unrestricted Hartree–Fock calculation in both a traditional Gaussian basis set (cc-pVDZ), and a multi-sliced Gausslet basis set. The Gaussian basis set contains 5 spatial orbitals per atom, totaling to 50 spatial orbitals with the associated nondiagonal two-electron integrals as one would expect. The Gausslet basis is formed adaptively according to pre-determined cutoffs, and the number of functions ranges from 7000–10 000 spatial orbitals for these calculations, while retaining the diagonal property, making calculations at the UHF level relatively straightforward, even with such a formidable number of basis functions. This large primitive basis gives UHF results near the complete basis set limit, well beyond the accuracy of this Gaussian basis. This high accuracy in the HF comes at little cost; the UHF is still fast compared to the correlated calculations and the large number of primitive functions do not strongly affect the size of the DG basis. For larger systems the accuracy could be reduced to keep the HF manageable.

To construct the DG basis, we construct 10 spatial blocks, 1 around each atom. We make use of the UHF calculation density matrix and keep 7 total orbitals per block, yielding 70 total spatial orbitals, a number almost identical to the number of Gaussians in the cc-pVDZ basis, but maintaining the block diagonal property of the two-electron integrals. By construction, UHF in this basis can accurately match the UHF results of the Gausslet basis set. The introduction of correlation through DMRG on this basis shows improvement as expected but a relative offset from the exact answer due to UHF's focus on static correlation. By using the weigthing procedure, in order to include UHF orbitals in the DG construction, we find excellent agreement with calculations done in the complete basis set limit. To that end, a weighting factor of $\alpha = 0.01$ was imposed. The number of functions kept is 15 per block, yielding 150 functions in total, which is about three times larger than the cc-pVDZ basis. However, the structure of the interactions and spatially local construction of the DG functions allows the DMRG calculation to be done with a 1–2 order of magnitude reduction in computational cost as compared to using the DG or Gaussian basis sets alone for the current implementation. This suggests that the hybrid active-space approach with the DG-blocking procedure is a powerful technique for recovering both static and dynamic correlation in a cost-effective manner.

To elaborate on the scaling of DMRG, $m$ be the bond dimension for the state required for the desired accuracy. Then the computational cost of the Gaussian basis sets is expected to be $O(N_a^3 m^3)$. In contrast, for Gausslets and DG representations based on Gausslets, using matrix product operator (MPO) compression

and let $D$ be the MPO dimension, one expects the asymptotic cost to be $O(N_p D m^3)$ and $O(N_d D m^3)$, respectively. Due to the localization properties, $D$ is expected to depend weakly on length, and be comparable for both the Gausslet and DG block representations. One can see this from considering the MPO decomposition in the Gausslet representation then transformed to the DG representation. While some expansion of the bond dimension could happen within a block, it is bounded by the Gausslet dimension of the block from the properties of a Schmidt decomposition, and no inter-block mixing occurs. As a result, the bond dimension will be comparable. Hence, it is the massive reduction in number of basis functions, from 10 000 to 150 that reduces the cost by several orders of magnitude while maintaining excellent accuracy. The Gaussian basis set, while advantageous in number of functions, suffers from lack of spatial locality, but the cost of this nonlocality in terms of the required bond dimension for equivalent accuracy has not been studied, and likely depends significantly on the completeness of the basis. It is easier to compare the scaling of the Gaussian DMRG due to the lack of diagonality, resulting in a much larger set of two-electron integrals. In this case, the block diagonality of the DG basis typically results in approximate linear scaling in the number of atoms, versus cubic dependence with Gaussian DMRG. Hence in both cases, the DG approach offers a significant reduction in computational cost for an accuracy that nearly approaches the complete basis set limit, which was obtained through accurate quantum Monte-Carlo calculations [35, 72].

## 6. Conclusion

The discretization problem is a crucial aspect determining the cost and effectiveness of quantum chemistry methods both for classical and quantum methods. The most popular basis sets for correlated calculations, Gaussian based molecular orbitals, are notably more compact than alternatives and have a well developed set of tools for their use. Unfortunately, in cost models for quantum computation, the overhead of a quartic number of terms in the Hamiltonian leads to poor scaling with system size. On the other side of the spectrum, representations that achieve a strictly quadratic number of terms and a diagonal representation, such as Gausslets or plane-wave dual functions exhibit excellent scaling with system size, but have overheads that make them undesirable for modest size implementations.

Here we introduced a systematic method for interpolating between the two regimes through the use of a blocking procedure, motivated from the discontinuous Galerkin (DG) method. This method is able to use any primitive basis with the diagonal property to represent a delocalized active-space basis, while maintaining the diagonal property between blocks. By choosing a plane-wave dual primitive basis and Gaussian active space, we were able to show how one can adaptively interpolate between these two regimes to attain both a scaling and constant factor advantage over the target active space.

When these empirical results are put into the context of known costs for exact quantum algorithms for chemistry, we observed a scaling improvement over Gaussian basis sets from $O(N_h^{4.5})$ to $O(N_h^{2.6})$ with a constant factor crossover around 15 to 20 hydrogen atoms. This suggests that for modest sized systems, such as those just beyond the classically tractable regime, this representation will be the optimal choice for quantum algorithms. Moreover, we showed that for high accuracy DMRG calculations, one may take advantage of this representation to achieve a high accuracy calculation with a cost reduction that is over an order of magnitude with respect to traditional representations. In all cases, one may use this methodology to scale between a compact representation and one with superior integral scaling depending on the requirements of a particular method.

## Acknowledgments

## Appendix A. Trotter step by low-rank factorization

As an alternative to the fixed swap networks used in the main text to evolve under the two-electron integral term in the DG basis, one may use the low-rank factorization strategy in [39], but applied to the block diagonal matricized tensor $v_{\kappa,\kappa',\lambda,\lambda';i,i',j,j'}^{(d)}$. For such a matrix, we know the maximum number of Cholesky factors is given by the dimension of the matrix, $N_d^2$, however, because of the block structure, each of the
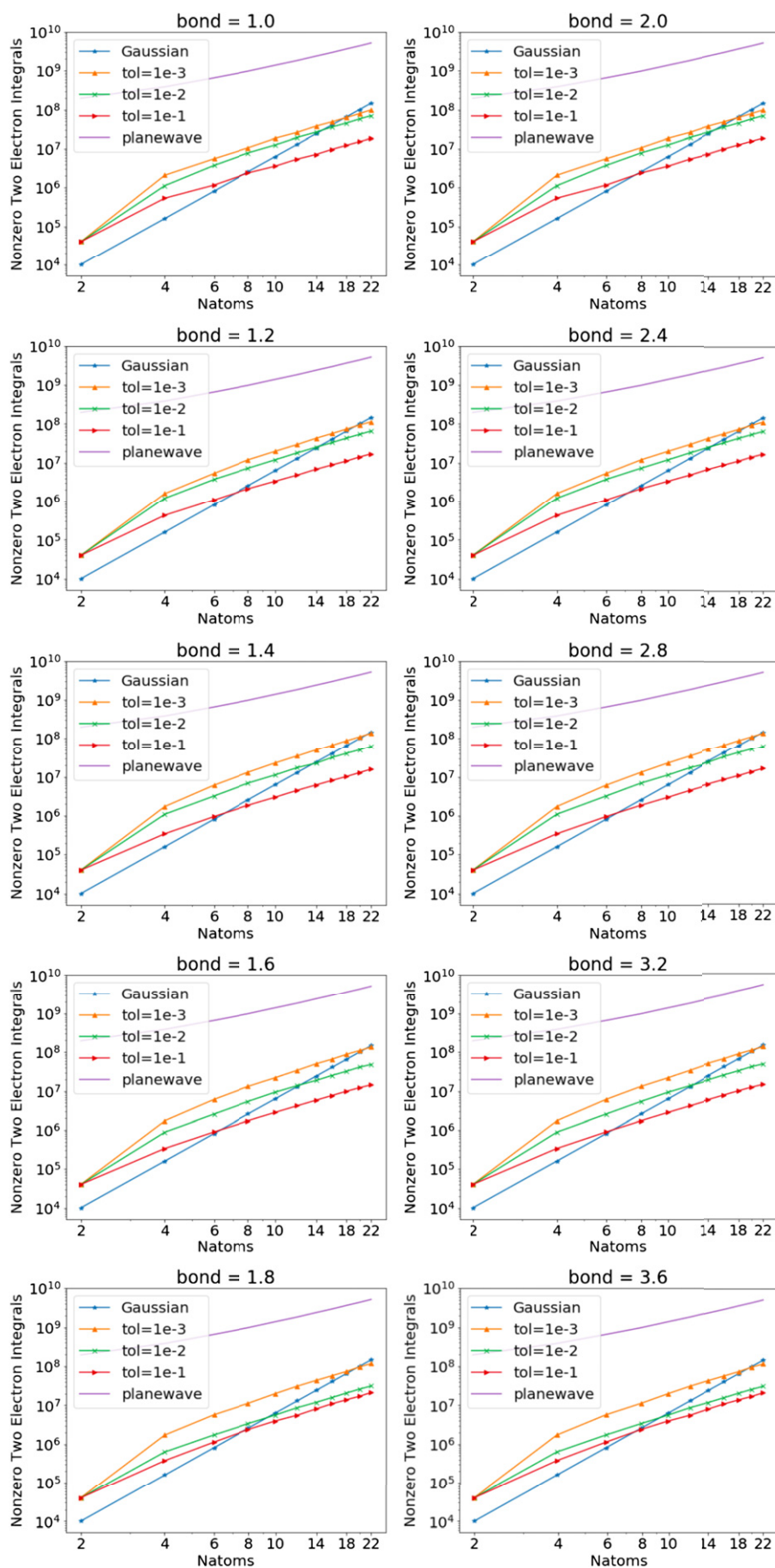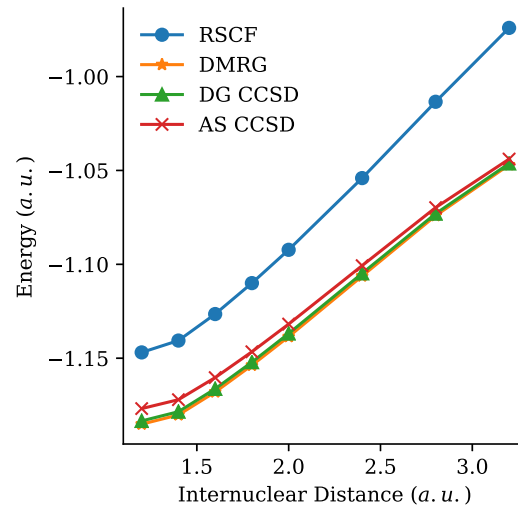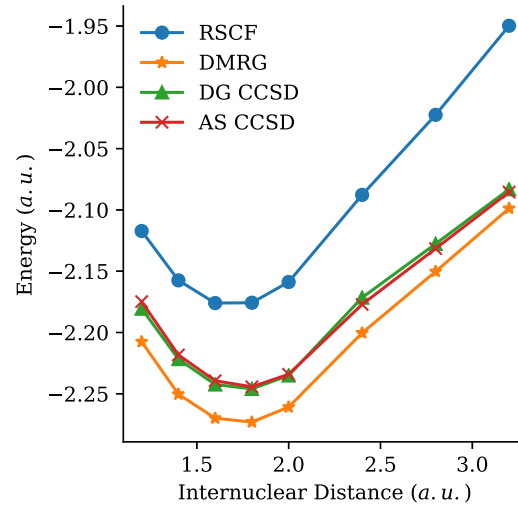
**Figure B1.** The number of non-zero two-electron integrals with bond lengths = 1.0 to 3.6 in atomic units, tol = $10^{-1}$ to $10^{-3}$, plotted on log–log scale. The crossover appears before or around $H_{22}$ in all cases.

**Table B1.** Crossover regions for different bond lengths and different truncation tolerances.

| Tolerance | $10^{-3}$ | $10^{-2}$ | $10^{-1}$ |
|---|---|---|---|
| $b = 1.0$ | $H_{16}-H_{18}$ | $H_{14}-H_{16}$ | $H_8-H_{10}$ |
| $b = 1.2$ | $H_{18}-H_{20}$ | $H_{14}-H_{16}$ | $H_6-H_8$ |
| $b = 1.4$ | $H_{20}-H_{22}$ | $H_{14}-H_{16}$ | $H_6-H_8$ |
| $b = 1.6$ | $H_{20}-H_{22}$ | $H_{12}-H_{14}$ | $H_6-H_8$ |
| $b = 1.7$ | $H_{20}-H_{22}$ | $H_{10}-H_{12}$ | $H_6-H_8$ |
| $b = 1.8$ | $H_{18}-H_{20}$ | $H_8-H_{10}$ | $H_8-H_{10}$ |
| $b = 2.0$ | $H_{16}-H_{18}$ | $H_8-H_{10}$ | $H_8-H_{10}$ |
| $b = 2.4$ | $H_{10}-H_{12}$ | $H_8-H_{10}$ | $H_8-H_{10}$ |
| $b = 2.8$ | $H_8-H_{10}$ | $H_8-H_{10}$ | $H_8-H_{10}$ |
| $b = 3.0$ | $H_8-H_{10}$ | $H_8-H_{10}$ | $H_8-H_{10}$ |
| $b = 3.2$ | $H_8-H_{10}$ | $H_8-H_{10}$ | $H_6-H_8$ |
| $b = 3.6$ | $H_8-H_{10}$ | $H_8-H_{10}$ | $H_4-H_6$ |



**Figure C1.** Potential energy surfaces for $H_2$ in a Gausslet basis computed by restricted SCF, (DG-)CCSD and benchmarked with DMRG results. The averaged number of non-zero two-electron integrals per atom (DG-element $\kappa$) is 4144. The averaged number of non-zero two-electron integrals of the CAS calculations is 205.



**Figure C2.** Potential energy surfaces for $H_4$ in a Gausslet basis computed by restricted SCF, (DG-)CCSD and benchmarked with DMRG results. The averaged number of non-zero two-electron integrals per atom (DG-element $\kappa$) is 16 580. The averaged number of non-zero two-electron integrals of the CAS calculations is 3269.

Cholesky factors need only have non-trivial support only within a $\kappa, \kappa'$ block. If $n_\kappa$ bounds the larger of $n_\kappa$, $n_{\kappa'}$, then the dimension of one of these blocks is $O(n_\kappa^2)$. It is easy to see that the total matricized tensor has

**Figure C3.** Potential energy surfaces for $H_6$ in a Gausslet basis computed by restricted SCF, (DG-)CCSD and benchmarked with DMRG results. The averaged number of non-zero two-electron integrals per atom (DG-element $\kappa$) is 287 762. The averaged number of non-zero two-electron integrals of the CAS calculations is 16 551.
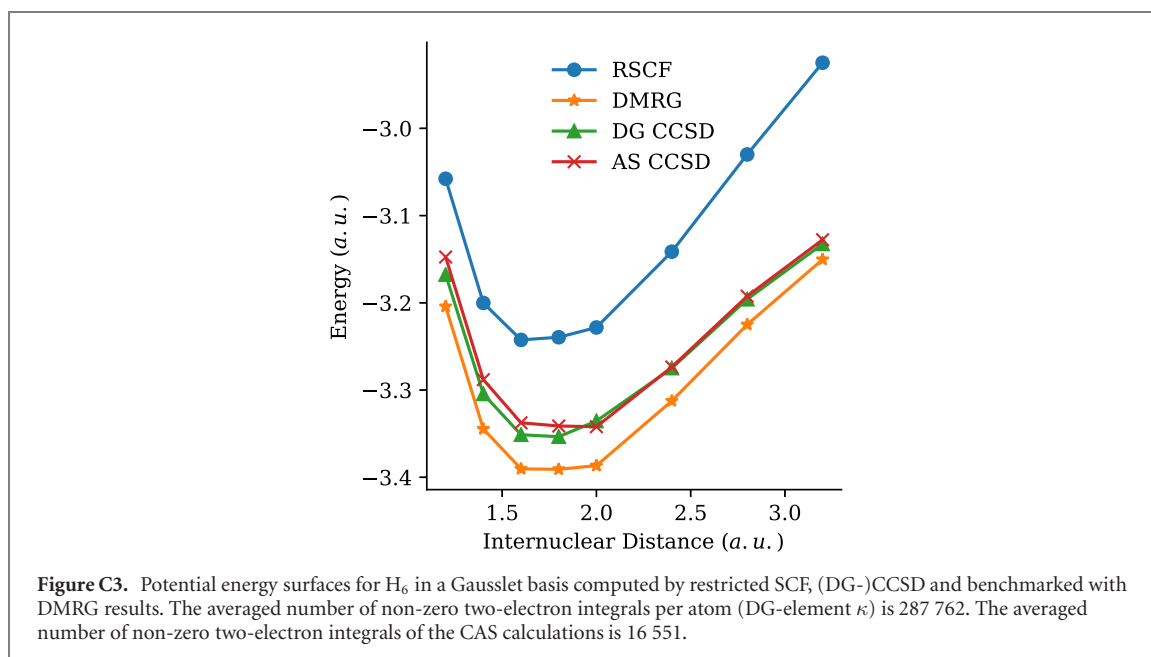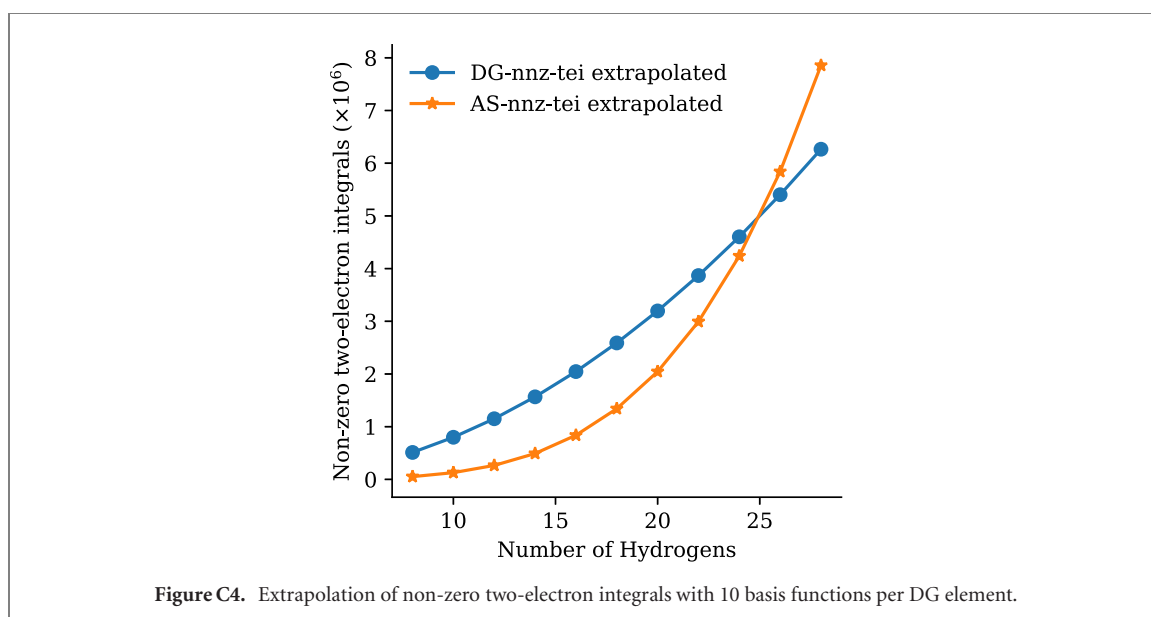
**Table C1.** Averaged numbers Gausslets used to discretize the molecular system.

| Atoms | $H_2$ | $H_4$ | $H_6$ | $H_8$ |
|---|---|---|---|---|
| Tot. no. of Gausslet | 529 | 801 | 1066 | 1336 |
| Gausslets in $z$-axis | 11−13 | 15−21 | 17−27 | 21−33 |
| Gausslets in $x$-resp. $y$-axis | 6.33−7.40 | 6.92−7.53 | 7.31−7.78 | 7.27−8.01 |

**Table C2.** Averaged numerical values of non-zero two-electron integrals for $H_2$, $H_4$, $H_6$ and $H_8$.

| Atoms | $H_2$ | $H_4$ | $H_6$ | $H_8$ |
|---|---|---|---|---|
| No of Gausslets | 529 | 801 | 1066 | 1336 |
| nnz-tei Gausslet | 351 084 | 816 232 | 1 459 901 | 2 310 318 |
| DG per elem. | 6 | 6 | 10 | 10 |
| nnz-tei DG | 4144 | 16 580 | 287 762 | 511 449 |
| CAS | 4 | 8 | 12 | 16 |
| nnz-tei CAS | 205 | 3269 | 16 551 | 52 346 |



**Figure C4.** Extrapolation of non-zero two-electron integrals with 10 basis functions per DG element.

**Figure D1.** Notation and decomposition for a $\mathcal{P}$-swap network with partition sizes (1, 2, 1, 2, 2). A $\mathcal{P}$-swap network for a partition $(P_1, P_2, \ldots, P_{|\mathcal{P}|})$ of the qubits $\bigcup_i P_i$ acquaints every union of a pair of parts, i.e., $\{P \cup P' | P, P' \in \mathcal{P}\}$. At a high level, the structure is similar to that of the simple linear swap network, except that instead of single qubits being swapped, groups are (i.e., the parts of the partition $\mathcal{P}$). There are $|\mathcal{P}|$ layers of generalized swap gates, each of which swaps sets of qubits. For more details, see [59].



**Figure D2.** Construction of the double bipartite swap network, with parts of size 4. The top half of the top circuit contains the same swap gates as a linear swap network but with additional acquaintance opportunities. In the bottom half of the top circuit are 4 linear swap networks in a row, one for each acquaintance layer of the linear swap network in the top half, which is copied for each acquaintance layer of the bottom half. Overall, for every set of of four orbitals consisting of two from the top part and two from the bottom part, there is a layer in the circuit in which both pairs are simultaneously acquainted. The bottom circuit, depicting the double bipartite swap network, is formed by replacing each such acquaintance layer in the top circuit with a $\mathcal{P}$-swap network, where a pair of qubits acquainted in the top circuit corresponds to a part of the partition $\mathcal{P}$. The $\mathcal{P}$-swap network acquaints the union of each pair of pairs; see figure D1. The final gate ensures that overall effect is to shift the parts.

dimension $O(n_\kappa^2 N_b^2)$, and a number of non-zero entries scaling as $O(n_\kappa^4 N_b^2)$. Hence it matches the primitive limit as $n_\kappa \to 1$ and $N_b \to N_d$ and the active space limit as $n_\kappa \to N_d$ and $N_b \to 1$.
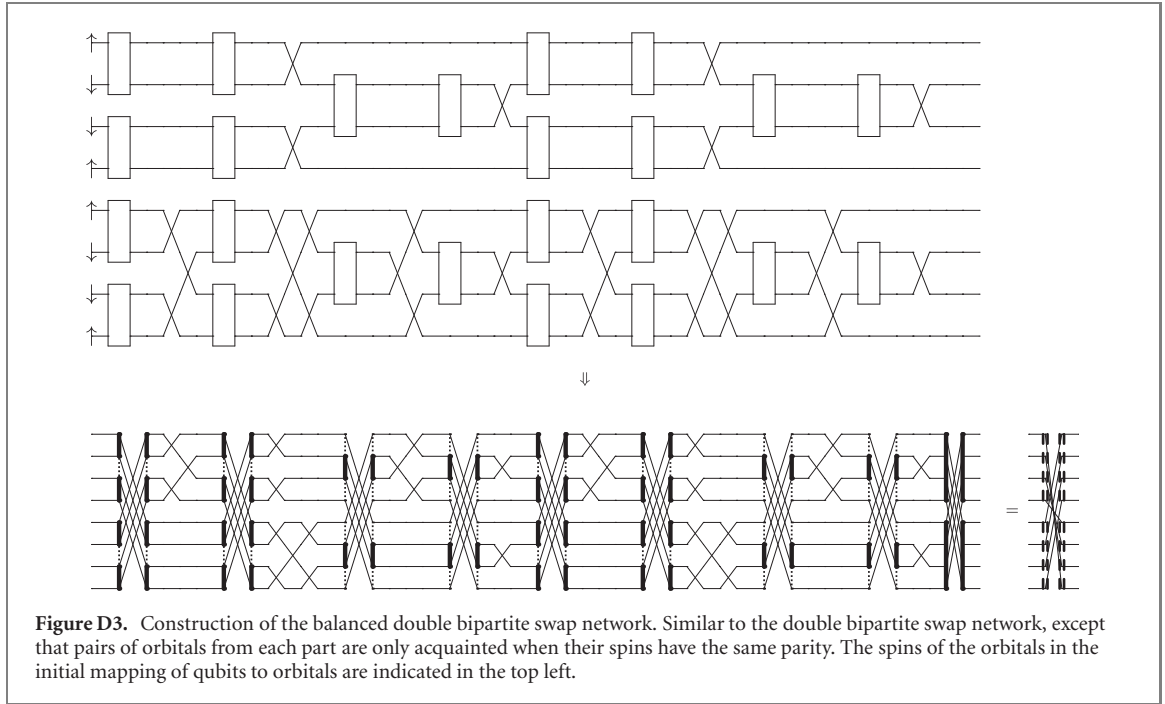
To execute a single Trotter step of the two-electron part of the Hamiltonian, one may start from a factorization that is a product over $\kappa, \kappa'$ blocks as

$$\exp\left[-\mathrm{i}\Delta t \sum_{\kappa,\kappa';i,i',j,j'} v^{(d)}_{\kappa,\kappa';i,i',j,j'} \hat{c}^\dagger_{\kappa,i}\hat{c}^\dagger_{\kappa',i'}\hat{c}_{\kappa',j'}\hat{c}_{\kappa,j}\right] \approx \prod_{\kappa,\kappa'} \exp\left[-\mathrm{i}\Delta t \sum_{i,i',j,j'} v^{(d)}_{\kappa,\kappa';i,i',j,j'} \hat{c}^\dagger_{\kappa,i}\hat{c}^\dagger_{\kappa',i'}\hat{c}_{\kappa',j'}\hat{c}_{\kappa,j}\right]$$

$$\equiv \prod_{\kappa\kappa'} R_{\kappa\kappa'} \tag{A1}$$

Within a non-trivial block $(v^{(d)}_{\kappa,\kappa',\lambda,\lambda';i,i',j,j'} = v^{(d)}_{\kappa,\kappa';i,i',j,j'})$ we expect the following decomposition

$$v^{(d)}_{\kappa,\kappa';i,i',j,j'} \approx \sum_{\mu\nu\ell}\left(U^{(\ell,\kappa,\kappa')}_{i\mu} U^{(\ell,\kappa,\kappa')}_{i'\mu} U^{(\ell,\kappa,\kappa')}_{j\nu} \cdot U^{(\ell,\kappa,\kappa')}_{j'\nu} \lambda^{(\ell,\kappa,\kappa')}_\mu \lambda^{(\ell,\kappa,\kappa')}_\nu\right). \tag{A2}$$

Here $\ell$ comes from Cholesky factorization, and $\mu, \nu$ comes from a second eigenvalue decomposition. As for a single block, $\kappa, \kappa'$, we assume $n_\kappa$ is independent of the system size, so are the index ranges of $\mu, \nu, \ell$. For each $\ell, \kappa, \kappa'$, $U^{(\ell,\kappa,\kappa')}$ is a matrix with orthogonal columns. This decomposition allows us to apply $R_{\kappa\kappa'}$ using the low-rank decomposition technique. The maximum rank of these factors is $L$, which empirically for Gaussian basis sets we expect to scale as $O(n_\kappa)$, but due to the lack of empirical data for large DG basis sets we assume the worst case, $O(n_\kappa^2)$. For each of these factors, the second eigenvalue decomposition will have

**Figure D3.** Construction of the balanced double bipartite swap network. Similar to the double bipartite swap network, except that pairs of orbitals from each part are only acquainted when their spins have the same parity. The spins of the orbitals in the initial mapping of qubits to orbitals are indicated in the top left.

maximal rank $\rho_\ell = O(n_\kappa)$. The depth of such circuits using a fermionic swap network scales as $\sum_\ell \rho_\ell < L n_\kappa$, which in empirical studies on molecular orbital basis sets scaled at least as $\Omega(n_\kappa \log n_\kappa)$ and in the worst case $O(n_\kappa^3)$. In any case, for a given error tolerance, from locality we know that $n_\kappa$ converges to a constant as a function of system size, and hence each $R_{\kappa\kappa'}$ may be executed in constant depth for large system using the low-rank Trotter step, which simultaneously executes a fermionic swap network.

Following this idea further, if one constructs a system mapped to qubits in 1D, where the mapping to orbitals is grouped by $\kappa$, then the $\kappa\kappa'$ block may be executed for $\kappa' = \kappa + 1$, and at the same time, the two blocks may be swapped with this technique. Moreover, as this adjacent $\kappa\kappa'$ block is disjoint from the qubits outside of $\kappa\kappa'$ this operation may be parallelized across pairs of adjacent blocks. This corresponds exactly to lifting the original fermionic swap network for time evolution to the level of blocks, where the individual fermionic simulation operations (as they were referred to in the original work) are performed via the low-rank method. From this, we see that the total required depth scales as $O(N_b)$ as desired. Although we have assumed here that $n_\kappa$ approaches a constant as system size grows, if we account for this in the worst case, we expect a depth of $O(N_b n_\kappa^3)$, which matches the fixed swap networks in the main body of the text. Hence, determining the optimal choice of implementation will be done to constant factors and requires further empirical study.

## Appendix B. Scaling crossover in a DG plane wave dual basis

Besides the total number of two-electron integrals, the cost of certain quantum algorithms such as the LCU method depends also on $\lambda$, the sum of the absolute value of the two-electron integrals. Here $\lambda$ is computed as

$$\lambda = \sum_{\{\kappa,\kappa';pqrs\}/\{p=r,q=s\}} |v_{\kappa,\kappa';p,q,r,s}|. \tag{B1}$$

Notice that the terms $\{p = r, q = s\}$ nominally diverge in a plane wave basis set with periodic boundary conditions. These values can be computed by including correction terms to the periodic boundary condition [73, 74]. However, the number of such terms in the two-body interaction scales the same as that in the one-body interaction in the Hamiltonian, so we omit such low-order terms directly for simplicity. In figure 10, we observe that the crossover point occurs around $H_8$ when the tolerance is set to 0.1, and around $H_{18}$ when the tolerance is 0.01.

## Appendix C. Correlated calculations in a DG Gausslet basis

Figures C1–C3 show the DG calculations for $H_2$, $H_4$ and $H_6$ using Gausslets as the primitive basis set. Table C1 shows information on the real space discretization for the used Gausslet basis set.

Table C2 compares the non-zero two-electron integral count for the DG-Gausslet basis and the active space basis. Figure C4 shows a naive extrapolation of the non-zero two-electron integral count presented in table C2.

## Appendix D.  Swap network sub-circuits

In this section, we give some more detail about the components of the swap network described in section 4.1. Recall the structure of the overall swap network:

(a) A four-complete swap network within each half-block. This acquaints all sets of 4 orbitals with the same spin and within each block.

(b) A double bipartite swap network on each block. This acquaints all sets of 4 orbitals with no net spin and within each block. Details of the construction are given in figure D2.

(c) A permutation within each block. This changes the orbital to qubit mapping in preparation for the next stage.

(d) Alternating layers of balanced double bipartite swap networks. Each balanced double bipartite swap network acquaints, for some pair of blocks, all sets of 4 orbitals with an even number of each spin and with two orbitals from each block. The $N_b$ alternating layers ensure that every pair of blocks is involved together in some balanced double bipartite swap network. Details of the construction of a double bipartite swap network are given in figure D3.

## Reference

[1]  Helgaker T, Jorgensen P and Olsen J 2014 *Molecular Electronic-Structure Theory* (New York: Wiley)
[2]  White S R 1992 Density matrix formulation for quantum renormalization groups *Phys. Rev. Lett.* **69** 2863
[3]  Chan G K-L and Sharma S 2011 The density matrix renormalization group in quantum chemistry *Annu. Rev. Phys. Chem.* **62** 465
[4]  Nakatani N and Chan G K-L 2013 Efficient tree tensor network states (TTNS) for quantum chemistry: generalizations of the density matrix renormalization group algorithm *J. Chem. Phys.* **138** 134113
[5]  Szalay S, Pfeffer M, Murg V, Barcza G, Verstraete F, Schneider R and Legeza Ö 2015 Tensor product methods and entanglement optimization for ab initio quantum chemistry *Int. J. Quantum Chem.* **115** 1342
[6]  Abrams D S and Lloyd S 1997 Simulation of many-body Fermi systems on a universal quantum computer *Phys. Rev. Lett.* **79** 4
[7]  Ortiz G, Gubernatis J, Knill E and Laflamme R 2001 Quantum algorithms for fermionic simulations *Phys. Rev.* A **64** 022319
[8]  Aspuru-Guzik A, Dutoi A D, Love P J and Head-Gordon M 2005 Simulated quantum computation of molecular energies *Science* **309** 1704
[9]  Wecker D, Bauer B, Clark B K, Hastings M B and Troyer M 2014 Gate-count estimates for performing quantum chemistry on small quantum computers *Phys. Rev.* A **90** 022305
[10]  Hastings M B, Wecker D, Bauer B and Troyer M 2015 Improving quantum algorithms for quantum chemistry *Quantum Inf. Comput.* **15** 1
[11]  Babbush R, McClean J, Wecker D, Aspuru-Guzik A and Wiebe N 2015 Chemical basis of Trotter-Suzuki errors in chemistry simulation *Phys. Rev.* A **91** 022311
[12]  McClean J R, Babbush R, Love P J and Aspuru-Guzik A 2014 Exploiting locality in quantum computation for quantum chemistry *J. Phys. Chem. Lett.* **5** 4368
[13]  McClean J R, Schwartz M E, Carter J and de Jong W A 2017 Hybrid quantum-classical Hierarchy for mitigation of decoherence and determination of excited states *Phys. Rev.* A **95** 042308
[14]  Babbush R, Berry D W, Kivlichan I D, Wei A Y, Love P J and Aspuru-Guzik A 2016 Exponentially more precise quantum simulation of Fermions in second quantization *New J. Phys.* **18** 033032
[15]  McClean J R, Romero J, Babbush R and Aspuru-Guzik A 2016 The theory of variational hybrid quantum-classical algorithms *New J. Phys.* **18** 23023
[16]  Lanyon B P *et al* 2010 Towards quantum chemistry on a quantum computer *Nat. Chem.* **2** 106
[17]  Du J, Xu N, Peng X, Wang P, Wu S and Lu D 2010 NMR implementation of a molecular hydrogen quantum simulation with adiabatic state preparation *Phys. Rev. Lett.* **104** 030502
[18]  Peruzzo A, McClean J, Shadbolt P, Yung M-H, Zhou X-Q, Love P J, Aspuru-Guzik A and O'Brien J L 2014 A variational eigenvalue solver on a photonic quantum processor *Nat. Commun.* **5** 4213
[19]  Shen Y, Zhang X, Zhang S, Zhang J-N, Yung M-H and Kim K 2015 Quantum implementation of unitary coupled cluster for simulating molecular electronic structure (arXiv:1506.00443)
[20]  Santagati R *et al* 2016 Quantum simulation of Hamiltonian spectra on a silicon chip (arXiv:1611.03511)
[21]  O'Malley P J J *et al* 2016 Scalable quantum simulation of molecular energies *Phys. Rev.* X **6** 031007
[22]  Colless J, Ramasesh V, Dahlen D, Blok M, Kimchi-Schwartz M, McClean J, Carter J, De Jong W and Siddiqi I 2018 Computation of molecular spectra on a quantum processor with an error-resilient algorithm *Phys. Rev.* X **8** 011021
[23]  Kandala A, Mezzacapo A, Temme K, Takita M, Brink M, Chow J M and Gambetta J M 2017 Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets *Nature* **549** 242
[24]  Dumitrescu E F, McCaskey A J, Hagen G, Jansen G R, Morris T D, Papenbrock T, Pooser R C, Dean D J and Lougovski P 2018 Cloud quantum computing of an atomic nucleus *Phys. Rev. Lett.* **120** 210501
[25]  Hempel C *et al* 2018 Quantum chemistry calculations on a trapped-ion quantum simulator *Phys. Rev.* X **8** 031022
[26]  Dunning T H Jr 1989 Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen *J. Chem. Phys.* **90** 1007
[27]  Kendall R A, Dunning T H Jr and Harrison R J 1992 Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions *J. Chem. Phys.* **96** 6796

[28] Woon D E and Dunning T H Jr 1993 Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon *J. Chem. Phys.* **98** 1358

[29] Goedecker S 1999 Linear scaling electronic structure methods *Rev. Mod. Phys.* **71** 1085

[30] Ahlrichs R, Lischka H, Staemmler V and Kutzelnigg W 1975 Pno–ci (pair natural orbital configuration interaction) and cepa–pno (coupled electron pair approximation with pair natural orbitals) calculations of molecular systems. I. Outline of the method for closed-shell states *J. Chem. Phys.* **62** 1225

[31] Riplinger C, Sandhoefer B, Hansen A and Neese F 2013 Natural triple excitations in local coupled cluster calculations with pair natural orbitals *J. Chem. Phys.* **139** 134101

[32] Riplinger C and Neese F 2013 An efficient and near linear scaling pair natural orbital based local coupled cluster method *J. Chem. Phys.* **138** 034106

[33] Stoudenmire E M and White S R 2017 Sliced basis density matrix renormalization group for electronic structure *Phys. Rev. Lett.* **119** 046401

[34] White S R 2017 Hybrid grid/basis set discretizations of the schrödinger equation *J. Chem. Phys.* **147** 244102

[35] White S R and Stoudenmire E M 2019 Multisliced gausslet basis sets for electronic structure *Phys. Rev.* B **99** 081110

[36] Babbush R, Wiebe N, McClean J, McClain J, Neven H and Chan G K-L 2018 Low-depth quantum simulation of materials *Phys. Rev. X* **8** 011044

[37] Babbush R, Gidney C, Berry D W, Wiebe N, McClean J, Paler A, Fowler A and Neven H 2018 Encoding electronic spectra in quantum circuits with linear T complexity *Phys. Rev. X* **8** 041015

[38] Berry D, Gidney C, Motta M, McClean J and Babbush R 2019 Qubitization of arbitrary basis quantum chemistry by low rank factorization (arXiv:1902.02134)

[39] Motta M, Ye E, McClean J R, Li Z, Minnich A J, Babbush R and Chan G K-L 2018 Low rank representations for quantum simulation of electronic structure (arXiv:1808.02625)

[40] Babbush R, Berry D W, McClean J R and Neven H 2018 Quantum simulation of chemistry with sublinear scaling to the continuum (arXiv:1807.09802)

[41] Low G H and Wiebe N 2018 Hamiltonian simulation in the interaction picture (arXiv:1805.00675)

[42] Babbush R, Berry D W, Sanders Y R, Kivlichan I D, Scherer A, Wei A Y, Love P J and Aspuru-Guzik A 2018 Exponentially more precise quantum simulation of fermions in the configuration interaction representation *Quantum Sci. Technol.* **3** 015006

[43] Lin L, Lu J, Ying L and E W 2012 Adaptive local basis set for Kohn-Sham density functional theory in a discontinuous Galerkin framework I: total energy calculation *J. Comput. Phys.* **231** 2140

[44] Hu W, Lin L and Yang C 2015 DGDFT: a massively parallel method for large scale density functional theory calculations *J. Chem. Phys.* **143** 124110

[45] Banerjee A S, Lin L, Hu W, Yang C and Pask J E 2016 Chebyshev polynomial filtered subspace iteration in the discontinuous galerkin method for large-scale electronic structure calculations *J. Chem. Phys.* **145** 154101

[46] Banerjee A S, Lin L, Suryanarayana P, Yang C and Pask J E 2018 Two-level Chebyshev filter based complementary subspace method for pushing the envelope of large-scale electronic structure calculations *J. Chem. Theory Comput.* **14** 2930

[47] Li Y and Lin L 2019 Globally constructed adaptive local basis set for spectral projectors of second order differential operators *Multiscale Model. Simul.* **17** 92

[48] Arnold D N 1982 An interior penalty finite element method with discontinuous elements *SIAM J. Numer. Anal.* **19** 742

[49] Cockburn B, Karniadakis G E and Shu C-W 2000 The development of discontinuous Galerkin methods *Discontinuous Galerkin Methods* (Berlin: Springer) pp 3–50

[50] Hu W, Lin L and Yang C 2015 Edge reconstruction in armchair phosphorene nanoribbons revealed by discontinuous Galerkin density functional theory *Phys. Chem. Chem. Phys.* **17** 31397

[51] Zhang G, Lin L, Hu W, Yang C and Pask J E 2017 Adaptive local basis set for Kohn–Sham density functional theory in a discontinuous galerkin framework II: force, vibration, and molecular dynamics calculations *J. Comput. Phys.* **335** 426

[52] Xu Q, Suryanarayana P and Pask J E 2018 Discrete discontinuous basis projection method for large-scale electronic structure calculations *J. Chem. Phys.* **149** 094104

[53] Pask J E and Sterne P A 2005 Finite element methods in ab initio electronic structure calculations *Modelling Simul. Mater. Sci. Eng.* **13** R71

[54] Chen H, Dai X, Gong X, He L and Zhou A 2014 Adaptive finite element approximations for Kohn–Sham models *Multiscale Model. Simul.* **12** 1828

[55] Kanungo B and Gavini V 2017 Large-scale all-electron density functional theory calculations using an enriched finite-element basis *Phys. Rev.* B **95** 035112

[56] Lehtola S 2019 A review on non-relativistic, fully numerical electronic structure calculations on atoms and diatomic molecules *Int. J. Quantum Chem.* **0** e25968

[57] Kivlichan I D, McClean J, Wiebe N, Gidney C, Aspuru-Guzik A, Chan G K-L and Babbush R 2018 Quantum simulation of electronic structure with linear depth and connectivity *Phys. Rev. Lett.* **120** 110501

[58] Bravyi S and Kitaev A 2002 Fermionic quantum computation *Ann. Phys., NY* **298** 210

[59] O'Gorman B, Huggins W J, Rieffel E G and Whaley K B 2019 Generalized swap networks for near-term quantum computing (arXiv:1905.05118)

[60] Kitaev A Y 1995 Quantum measurements and the Abelian stabilizer problem (arXiv:quant-ph/9511026)

[61] Abrams D S and Lloyd S 1999 Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors *Phys. Rev. Lett.* **83** 5162

[62] Wecker D, Hastings M B and Troyer M 2015 Progress towards practical quantum variational algorithms *Phys. Rev. A* **92** 42303

[63] Childs A M and Wiebe N 2012 Hamiltonian simulation using linear combinations of unitary operations *Quantum Inf. Comput.* **12** 901

[64] Berry D W, Childs A M, Cleve R, Kothari R and Somma R D 2015 Simulating Hamiltonian dynamics with a truncated Taylor series *Phys. Rev. Lett.* **114** 90502

[65] Low G H and Chuang I L 2016 Hamiltonian simulation by qubitization (arXiv:1610.06546)

[66] Babbush R, Gidney C, Berry D, Wiebe N, McClean J, Paler A, Fowler A and Neven H 2018 Encoding electronic spectra in quantum circuits with linear T complexity *Phys. Rev. X* **8** 041015

[67] Low G H, Kliuchnikov V and Schaeffer L 2018 Trading T-gates for dirty qubits in state preparation and unitary synthesis (arXiv:1812.00954)

[68] Rubin N, Babbush R and McClean J 2018 Application of fermionic marginal constraints to hybrid quantum algorithms *New J. Phys.* **20** 053020

[69] Perdew J P, Burke K and Ernzerhof M 1996 Generalized gradient approximation made simple *Phys. Rev. Lett.* **77** 3865

[70] Hamann D 2013 Optimized norm-conserving vanderbilt pseudopotentials *Phys. Rev.* B **88** 085117

[71] Yang C, Meza J C, Lee B and Wang L-W 2009 Kssolva matlab toolbox for solving the Kohn-Sham equations *ACM Trans. Math. Softw.* **36** 10

[72] Motta M *et al* Simons Collaboration on the Many-Electron Problem 2017 Towards the solution of the many-electron problem in real materials: equation of state of the hydrogen chain with state-of-the-art many-body methods *Phys. Rev.* X **7** 031059

[73] Gygi F and Baldereschi A 1986 Self-consistent Hartree-Fock and screened-exchange calculations in solids: application to silicon *Phys. Rev.* B **34** 4405

[74] Bylaska E J, Tsemekhman K, Baden S B, Weare J H and Jonsson H 2011 Parallel implementation of $\gamma$-point pseudopotential plane-wave DFT with exact exchange *J. Comput. Chem.* **32** 54