Chemical Science

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

View Article Online
View Journal | View Issue

Cite this: *Chem. Sci.*, 2021, **12**, 3497

All publication charges for this article have been paid for by the Royal Society of Chemistry

Received 3rd December 2020
Accepted 7th January 2021

DOI: 10.1039/d0sc06627c

rsc.li/chemical-science

A feasible approach for automatically differentiable unitary coupled-cluster on quantum computers†

Jakob S. Kottmann, *[ab] Abhinav Anand [a] and Alán Aspuru-Guzik *[abcd]

We develop computationally affordable and encoding independent gradient evaluation procedures for unitary coupled-cluster type operators, applicable on quantum computers. We show that, within our framework, the gradient of an expectation value with respect to a parameterized n-fold fermionic excitation can be evaluated by four expectation values of similar form and size, whereas most standard approaches, based on the direct application of the parameter-shift-rule, come with an associated cost of $\mathcal{O}(2^{2n})$ expectation values. For real wavefunctions, this cost can be further reduced to two expectation values. Our strategies are implemented within the open-source package Tequila and allow blackboard style construction of differentiable objective functions. We illustrate initial applications through extended adaptive approaches for electronic ground and excited states.

1 Introduction

The proposition of using the quantum phase estimation algorithm to extract eigenenergies of electronic Hamiltonians with the help of quantum computers[1] resulted in various new research ideas for quantum chemistry on quantum computers. One such direction was the introduction and successful demonstration of the quantum variational eigensolver (VQE),[2,3] a new method for approximating eigenenergies with current and near-term quantum hardware in mind. Variational quantum algorithms apply the variational principle to expectation values of a parameterized quantum circuit $U(\theta)$ and a qubit Hamiltonian H by optimizing the parameters with a classical optimization algorithm

$$\min_{\theta}(\langle H\rangle_{U(\theta)}) \equiv \min_{\theta}(\langle 0|U^{\dagger}(\theta)HU(\theta)|0\rangle). \quad (1)$$

The original proposal of VQE inspired development of numerous new variational quantum algorithms for estimating energies of quantum chemical and many-body models on quantum computers (see recent reviews[4,5]) as well as in various other fields, including quantum machine learning,[6–13] combinatorial optimization,[14] and quantum optics.[15–17] These developments in combination with the recent availability of open access quantum computers,[18] and significant improvements of currently available quantum hardware[19–21] are currently clearing the path towards Feynman's original idea of simulating physics with quantum computers[22] leveraging this powerful tool to elucidate challenging chemical processes.[23]

Gradient based optimization methods are not always the canonical choice for variational quantum algorithms. Their applicability depends strongly on the number of used parameters and the form of the parameterized quantum gates. For specific types of gates, the evaluation of analytical gradients of expectation values becomes comparably cheap by applying the parameter-shift-rule,[24] where the expectation value has to be evaluated two times with shifted parameters, leading to elegant and computationally feasible implementations of automatic differentiation within quantum algorithms pioneered within PennyLane.[25]

While automatic differentiation is becoming a standard technique in quantum computing or classical machine learning algorithms, it started to gain attention in classical quantum chemistry as well. Examples are within the optimization of basis set parameters,[26] classical coupled-cluster amplitudes,[27] response properties,[28] nuclear derivatives[29–31] or new frontiers in excited state methods.[32] In this work, we extend the framework of automatically differentiable quantum algorithms to unitary coupled-cluster in its separated framework.[33–35] We develop procedures to calculate the gradients of expectation values, generated by arbitrary n-fold fermionic excitation operators, by compiling it into linear combinations of expectation values, that can be evaluated on quantum computers. Similar to the parameter-shift-rule,[24] those expectation values include the original excitation unitaries with shifted parameters but require an additional unitary generated by the nullspace projector of the corresponding excitation generator. If all underlying

[a]*Chemical Physics Theory Group, Department of Chemistry, University of Toronto, Toronto, Ontario M5S 3H6, Canada*

[b]*Department of Computer Science, University of Toronto, 214 College St, Toronto, ON M5T 3A1, Canada. E-mail: jakob.kottmann@utoronto.ca; aspuru@utoronto.ca*

[c]*Vector Institute for Artificial Intelligence, 661 University Ave. Suite 710, Toronto, Ontario M5G 1M1, Canada*

[d]*Canadian Institute for Advanced Research (CIFAR), 661 University Ave., Toronto, ON M5G 1M1, Canada*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d0sc06627c

wavefunctions are real we show that two of the four expectation values become equivalent, which reduces the computational cost to two. The main results are summarized in Table 1. All developed techniques are implemented in the open-source package Tequila[36] and allow blackboard style construction of automatically differentiable objectives constructed from fermionic excitation operators and other quantum gates. We present explicit code examples and initial application for electronic ground and excited states.

2 Methodology

The framework of unitary coupled-cluster allows the construction of quantum circuits as a product of unitaries that create fermionic excitations within the wavefunction. Those unitary operations can be described in terms of their hermitian generators G as

$$U(\theta) = \mathrm{e}^{-i\frac{\theta}{2}G} \tag{2}$$

where the generators for singles, doubles and n-fold fermionic excitations are given by

$$G_{\mathrm{pq}} = i(a_{\mathrm{p}}^{\dagger}a_{\mathrm{q}} - a_{\mathrm{q}}^{\dagger}a_{\mathrm{p}}), \tag{3}$$

$$G_{\mathrm{pqrs}} = i(a_{\mathrm{p}}^{\dagger}a_{\mathrm{q}}a_{\mathrm{r}}^{\dagger}a_{\mathrm{s}} - \mathrm{h.c.}), \tag{4}$$

$$G_{\mathbf{pq}} = i\left(\prod_{i=1}^{n} a_{\mathrm{p}_i}^{\dagger}a_{\mathrm{q}_i} - \mathrm{h.c.}\right), \tag{5}$$

with $a^{\dagger}$, a being the usual anti-commuting fermionic creation and annihilation operators.[4,37–40] Given the condition that the generator G of a parametrized unitary $U(\theta) = \mathrm{e}^{-i\theta G}$ has only two distinct eigenvalues $\pm r$, Schuld *et al.*[24] showed that the direct measurement of the analytical gradient of expectation values $\langle\Psi|\hat{O}|\Psi\rangle \equiv \langle O\rangle U_{\Psi}$, formed from unitaries U_{Ψ} including $U(\theta)$, and some hermitian operator $\hat{H}$, can be achieved as

$$\frac{\partial\langle H\rangle XU(\theta)Y}{\partial\theta} = r(\langle H\rangle XU(\theta+s)Y - \langle H\rangle XU(\theta-s)Y) \tag{6}$$

with $s = \dfrac{\pi}{4r}$ and where we used $U_{\Psi}(\theta) = XU(\theta)Y$ to illustrate the differentiable gate $U(\theta)$ within a larger abstracted unitary, symbolized by X and Y as arbitrary placeholders. Further improvements on this gradient evaluation technique in the context of stochastic sampling,[41] higher order derivatives,[42] classical simulation[43] and noisy evaluations[44] have been developed, making it a standard tool within variational quantum algorithms. If the condition of having only two distinct eigenvalues does not hold for the generator, the analytical gradient can still be obtained by more sophisticated techniques. One is the decomposition of $U(\theta)$ into a product of directly differentiable unitaries[45] and another involves the execution of the unitary controlled by an additional ancillary qubit.[24,46] The fermionic generators of eqn (3)–(5) do not fulfill the necessary condition to be directly differentiable but their gradients can still be evaluated by these extended approaches, where the latter has been demonstrated by Romero *et al.*[46] The decomposition approach can be realized in a straightforward way described in the following. In order to be executable on a general quantum computer, the n-fold excitation generators of eqn (5) are transformed into qubit operators by writing them as a linear combination of tensor-products of Pauli matrices $\boldsymbol{\sigma}_i$, often referred to as Pauli strings

$$G_{\mathbf{pq}} = \sum_i c_{i(\mathbf{p},\mathbf{q})}\boldsymbol{\sigma}_{i(\mathbf{p},\mathbf{q})} \tag{7}$$

where the length of the sum and the individual form of the coefficients c and Pauli strings $\boldsymbol{\sigma}$ depends on the generator and the chosen transformation. The Pauli strings arising from an individual generator commute amongst each other[46] so that the qubit unitary generated from those Pauli strings decomposes into a product of multi-Pauli rotations $\mathrm{e}^{i\frac{\theta}{2}G} \rightarrow \prod_k \mathrm{e}^{i\frac{\theta}{2}c_k\boldsymbol{\sigma}_k}$. Since individual Pauli strings $\boldsymbol{\sigma}$ are self-inverse they only have two distinct eigenvalues (± 1) making the unitaries generated by them directly differentiable. The gradient can then be evaluated by combining the product rule of calculus with the parameter-shift-rule of eqn (6). Usually the generators of n-fold fermionic excitation operators are transformed into $\mathcal{O}(2^{2n-1})$ Pauli strings by the Jordan–Wigner or Bravyi–Kitaev transformation.[47–50] This will lead to the evaluation of $\mathcal{O}(2^{2n})$ expectation values in order to obtain their gradients by decomposition into directly differentiable gates.

In the following we will show how to evaluate the gradients of n-fold fermionic excitation operators with a constant cost factor of 4 instead of $\mathcal{O}(2^{2n})$. For real wavefunctions this cost factor can be lowered to 2, making it equivalent in cost as the simplest finite difference stencils. We will start with generalized operators – that don't have to be fermionic excitations – and afterwards show how those operators can be constructed in the fermionic representation. For the convenience of the reader, we summarize our findings in Table 1 and illustrate the implementation of the automatic differentiation within Tequila[36] in Fig. 1.

Table 1 Overview over the used fermionic generators in this work. The gradient cost is given as a factor with respect to the cost of the evaluation of the original expectation value. See eqn (5), (11) and (33) for the definition of the G and $G_{\pm}$ generators and how to construct them using the nullspace projector P_0

Generator form	Gradient cost	Strategy
$G_{\mathbf{pq}} = \sum_i c_i\boldsymbol{\sigma}_i$	$\mathcal{O}(2^{2n})$	Shift-rule (6)
$G_{\mathbf{pq}} = \frac{1}{2}(G_+ + G_-)$	4	Fermionic-shift (16)
Real wavefunctions		
$G_{\mathbf{pq}} = \frac{1}{2}(G_+ + G_-)$	2	Fermionic-shift (19)
Generator approximation		
$G_{\mathbf{pq}} \approx G_{\pm}$	2	Shift-rule (6)

2.1 Generator decomposition

The generators $G_{\mathbf{pq}}$ (5) of fermionic excitations act only non-trivially on states with all **p** orbitals empty and all **q** orbitals

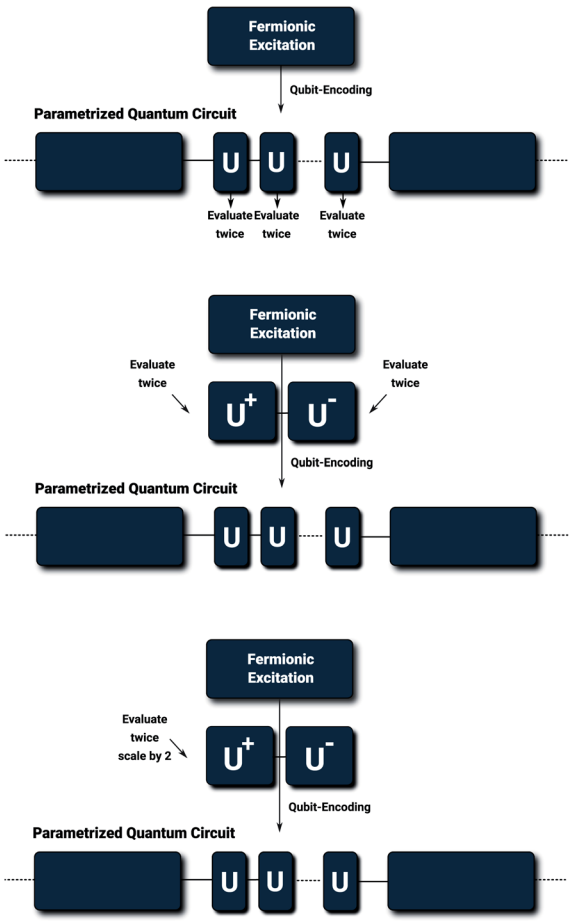

Fig. 1 Automatic differentiation for fermionic operators: schematic overview over standard approaches acting on the qubit level using eqn (6) (top) and the general gradient evaluation schemes on the fermionic level according to eqn (16) (middle) as well as for real wavefunctions according to eqn (19) (bottom). See also Table 1 for a general overview and eqn (18) for the definition of the fermionic shift gates $U_{\pm}^{\alpha}$.

occupied or *vice versa*. On all other states the generators act as zeroes, so that the corresponding unitary acts as identity operator. As a consequence, the generators have three distinct eigenvalues, ± 1 and 0 (see the ESI† for more details). We can formally write the generator as the sum over the projectors P_+, P_- and P_0, that project onto the spaces spanned by the eigenfunctions of G, multiplied by their corresponding eigenvalues ± 1 and 0. The generator can then formally be written as

$$G = P_+ - P_-. \tag{8}$$

Note that the projector P_0 is not included due to its zero eigenvalue. The individual projectors $P_\pm$ themselves have two distinct eigenvalues (0 and 1) and commute amongst themselves due to the orthogonality of the eigenstates. The generated unitary can then be split into two directly differentiable parts

$$e^{-i\frac{\theta}{2}G} = e^{-i\frac{\theta}{2}P_+}e^{+i\frac{\theta}{2}P_-} \tag{9}$$

reducing the gradient cost to a constant factor of 4 when the parameter-shift-rule is combined with the product rule. In Fig. 1 this procedure is schematically illustrated and compared with the standard approach that applies the parameter-shift-rule on the qubit level to each Pauli string individually. An alternative way to split the generator is by adding (subtracting) the nullspace projector P_0 to the generators

$$G = \frac{1}{2}(G_+ + G_-) \tag{10}$$

introducing the self-inverse generators

$$G_\pm = G \pm P_0. \tag{11}$$

$$G_\pm{}^2 = P_+ + P_- + P_0 = 1, \tag{12}$$

where the last equation holds due to the completeness and orthonormality of the eigenspace projectors. Both ways of splitting the generator will result in similar quantum circuits since the unitaries generated by $G_\pm$ just differ by a phase from the unitaries generated by $P_\pm$ (see eqn (25)) and we will mostly stick with the $G_\pm$ generators, since they share many properties with single qubit rotations and can be treated as generalized multi-qubit rotations

$$e^{-i\frac{\theta}{2}\frac{1}{2}G_\pm} = \cos\left(\frac{\theta}{4}\right) - i\sin\left(\frac{\theta}{4}\right)G_\pm. \tag{13}$$

Using this formula we can express how a unitary, generated by a fermionic excitation, acts in a closed analytical form (see the ESI† for a detailed derivation)

$$U(\theta) = e^{-i\frac{\theta}{2}G} = \cos\left(\frac{\theta}{2}\right) - i\sin\left(\frac{\theta}{2}\right)G + \left(1 - \cos\left(\frac{\theta}{2}\right)\right)P_0. \tag{14}$$

When acting on a specific electronic configuration $|\Phi\rangle$ the unitary acts as a unit operator if the configuration is in the nullspace of the generator, or, as a rotation between the original configuration and the n-fold excited configuration

$$U(\theta)|\Phi\rangle = \begin{cases} |\Phi\rangle, & P_0|\Phi\rangle = |\Phi\rangle \\ \left(\cos\left(\frac{\theta}{2}\right) - i\sin\left(\frac{\theta}{2}\right)G\right)|\Phi\rangle, & \text{else} \end{cases}. \tag{15}$$

Note that the generated superpositions are real due to the definition of the hermitian generators G in eqn (5).

2.2 Exact analytical gradients

With the splitting of the generators introduced in eqn (10), analytical gradients of n-fold fermionic excitations can be evaluated by combining the parameter-shift-rule with the product rule of calculus. Due to the product rule, this will lead

to a gradient cost factor of 4, meaning that 4 expectation values with similar cost to the original expectation value have to be evaluated

$$\frac{\partial\langle H\rangle XU(\theta)Y}{\partial\theta} = r\sum_{\alpha\in\{+,-\}}\left(\langle H\rangle XU_{+}^{\alpha}Y - \langle H\rangle XU_{-}^{\alpha}Y\right) \tag{16}$$

where $U_{\pm}^{\alpha}$ denotes the fermionic unitaries with shifted G_{α} part. Following ref. 24 the shift will be $s = \frac{\pi}{4r} = \pi$ with $r = \frac{1}{4}$.

Using an automatically differentiable framework, as for example offered in Tequila,[36] this gradient evaluation procedure can be implemented in a straightforward way. Using the most straightforward realization of this scheme, the explicit implementation of the unitary that is being differentiated will require approximately twice the number of native quantum gates as in the original expectation value since the two generators $G_{\pm}$ decompose into a similar number of Pauli strings as the original generator G. Note however, that this only holds for the unitary that is being differentiated, so the overall gate count for the gradients will just grow by a small constant that can be mitigated by applying more advanced compiling and gate fusion techniques. We will now show one simplification that can already be done on the fermionic level and that will result in the original unitary plus a unitary generated by it's nullspace projector P_0. Consider the first part of the product rule in eqn (16), where the parameter-shift is performed on the G_{+} generator, *i.e.* $\alpha = +$. The shifted unitary is then

$$U_{\pm}^{+} = \mathrm{e}^{-\mathrm{i}\frac{1}{4}(\theta\pm\pi)G_{+}}\mathrm{e}^{-\mathrm{i}\frac{\theta}{4}G-} = \mathrm{e}^{-\mathrm{i}\frac{1}{2}\left(\theta\pm\frac{\pi}{2}\right)G}\mathrm{e}^{-\frac{\mathrm{i}}{2}\left(\pm\frac{\pi}{2}\right)P_0}, \tag{17}$$

and if we shift the G_{-} part, we will arrive at the same expression with inverted sign on the P_0 dependent part. In general this fermionic shift operator can be written as a product of the shifted fermionic gate $U_{\pm} = U\left(\theta\pm\frac{\pi}{2}\right)$ and an additional, parameter independent, gate U_0^{α} generated by the nullspace projector of the generator

$$U_{\pm}^{\alpha}(\theta) = U_{\pm}(\theta)U_0^{\alpha} = \mathrm{e}^{-\mathrm{i}\frac{1}{2}\left(\theta\pm\frac{\pi}{2}\right)G}\mathrm{e}^{-\alpha\frac{\mathrm{i}}{2}\left(\pm\frac{\pi}{2}\right)P_0}, \quad \alpha\in\{+1,-1\}. \tag{18}$$

2.3 Real wavefunctions

In the following we will present a strategy to lower the cost of the exact gradient formula in eqn (16) to only two expectation values under the condition that the involved wavefunctions are real. In particular we demand that for a general quantum circuit $XU(\theta)Y$ the wavefunctions $Y|0\rangle$ are real and that the part denoted by X only generates real superpositions. For pure unitary coupled-cluster type circuits this reduces to the requirement of a real reference wavefunction. Under this conditions, two parts of the sum in eqn (16) become equivalent and the gradient can be evaluated as

$$\frac{\partial\langle H\rangle XUY}{\partial\theta} = \frac{1}{2}\left(\langle H\rangle XU_{+}^{\alpha}Y - \langle H\rangle XU_{-}^{\alpha}Y\right) \tag{19}$$

where α can be freely chosen to be either + or −. Note that the shift in the shifted fermionic gate $U_{\pm}$ is $s = \frac{\pi}{2}$ making the evaluation scheme of eqn (19) similar to the parameter-shift-rule for single qubit rotations (eqn (6) with $r = \frac{1}{2}$) with the only difference being the U_0^{α} gate after the shifted unitaries.

Lets start with the exact expression (16) for the analytical gradient derived in the last section. We can formally decompose the individual expectation values by inserting the identity as $1 = P_0 + (1 - P_0)$ and using the properties

$$XU_{\pm}^{\alpha}(1 - P_0)Y = XU_{\pm}Y|0\rangle \tag{20}$$

$$XU_{\pm}^{\alpha}P_0Y = \mathrm{e}^{-\alpha\mathrm{i}\frac{1}{2}\frac{\pi}{2}}XP_0Y|0\rangle \tag{21}$$

where we used the idempotency ${P_0}^2 = P_0$ of the projector and $GP_0 = 0$ resulting from P_0 being the nullspace projector of G. As in eqn (18), $U_{\pm} = \mathrm{e}^{-\mathrm{i}\frac{1}{2}\left(\theta\pm\frac{\pi}{2}\right)G}$ denotes the shifted fermionic unitary. The analytical gradient becomes then

$$\begin{aligned}\frac{\partial\langle H\rangle XUY}{\partial\theta} &= \frac{1}{4}\sum_{\alpha\in\{+,-\}}\Big(\langle H\rangle XU_{+}Y - \langle H\rangle XU_{-}Y\\ &+\left(\mathrm{e}^{-\alpha\mathrm{i}\frac{\pi}{4}}\langle 0|Y^{\dagger}U_{+}^{\dagger}X^{\dagger}HXP_0Y|0\rangle + \mathrm{h.c.}\right)\\ &-\left(\mathrm{e}^{\alpha\mathrm{i}\frac{\pi}{4}}\langle 0|Y^{\dagger}U_{-}^{\dagger}X^{\dagger}HXP_0Y|0\rangle + \mathrm{h.c.}\right)\Big)\\ &= \frac{1}{2}\left(\langle H\rangle XU_{+}Y - \langle H\rangle XU_{-}Y\right) + R\end{aligned} \tag{22}$$

where the first part is the same formula as for single qubit rotations and the second part denotes the residual R that can be written as

$$\begin{aligned}R = &\frac{1}{4}\left(cd_{+} + c^{*}d_{+}^{*} - cd_{-} - c^{*}d_{-}^{*}\right)\\ &+\frac{1}{4}\left(c^{*}d_{+} + cd_{+}^{*} - c^{*}d_{-} - cd_{-}^{*}\right),\end{aligned} \tag{23}$$

using $d_{\pm} = \langle 0|Y^{\dagger}U_{+}^{\dagger}X^{\dagger}HXP_0Y|0\rangle$ and $c = \mathrm{e}^{-\mathrm{i}\frac{\pi}{4}}$. Under the assumptions made above, the numbers $d_{\pm}$ are real numbers resulting in the two terms in the residual R to become identical. Note however, that the residue does not vanish. This means, that for real wavefunctions the two parts of the α sum in eqn (16) become identical, hence it is sufficient to evaluate only one part of the product-rule induced sum and scale it by a factor of two leading to the expression in eqn (19).

2.4 Approximations

If complex wavefunctions are involved, the gradient evaluation scheme of eqn (19) becomes an approximation. It will however

still be exact in cases where the wavefunction $Y|0\rangle$ (using the notation of the previous section) has no overlap with the nullspace of the fermionic generator that is being differentiated. Other types of approximations could be made for example by approximating the whole generator by either G_+ or G_- (or equivalently as $P_\pm$) which will lead to unitaries that are directly differentiable by the original shift rule of eqn (6). The generated unitaries will act in the same way as the unitaries generated from the original generators but will introduce phase factors to all nullspace elements of the wavefunction (see eqn (13)). If the wavefunction on which the unitaries act are not supported on the nullspace of the original generator G, *i.e.* $P_0Y|0\rangle = 0$, the unitaries generated by $G_\pm$ will act identical as unitaries generated by G. This leaves us with three choices for complex wavefunctions: either using eqn (16) resulting in exact gradients for the exact fermionic generators with an associated cost factor of 4, or, using eqn (19) to approximate the gradient with an associated cost factor of 2, or, approximate the fermionic generators with either G_+ or G_- where the exact gradient of the approximated generator can be obtained with eqn (6) with an associated cost factor of 2.

2.5 Operator construction

In the previous sections we showed how to obtain gradients of fermionic excitations with an overall cost factor of 4 (eqn (16)) or a cost factor of 2 (eqn (19)) for real wavefunctions. The derivation holds in general for operators with 3 distinct eigenvalues $\{-r,0,+r\}$. In eqn (10), the original generators were formally split into two self-inverse parts $G_\pm$ which are later recombined to result in a fermionic shift gate (18), combining the shifted fermionic excitation gate and an additional unitary generated by the nullspace projector P_0 of the fermionic generator. In order to construct the fermionic shift gate in eqn (18), one only needs the nullspace projector P_0 along with the original generator G, which for fermionic excitations is given in eqn (5). Note that the explicit construction of the $G_\pm$ generators is never necessary, but they might be used in alternative implementations of the fermionic shift gate. Given the original generator G and it's nullspace projector P_0, the $G_\pm$ as well as the original eigenspace projectors $P_\pm$ can be constructed as

$$G_\pm = G \pm P_0 \tag{24}$$

$$P_\pm = \pm\frac{1}{2}(G_\pm \pm 1), \tag{25}$$

where the completeness of the three eigenspace projectors was used. In the following we will show how the corresponding nullspace projector can be constructed for fermionic excitation generators. We will start with an intuitive illustration in the qubit representation and give the fermionic construction afterwards.

2.5.1 Qubit perspective. In the Jordan–Wigner encoding N spin-orbitals are directly mapped to N qubits. The computational basis-states of the N qubits correspond directly to the occupation number vectors in second-quantization and annihilation/creation operators are mapped to $\sigma_\pm = \frac{1}{2}(\sigma_x \pm i\sigma_y)$ qubit operators as well as σ_z operators on other qubits

$$a_k = \quad 1^{\otimes k-1}\sigma_k^+\sigma_z^{\otimes N-k}, \tag{26}$$

$$a_k^\dagger = \quad 1^{\otimes k-1}\sigma_k^-\sigma_z^{\otimes N-k}. \tag{27}$$

The transformed generators (5) of fermionic excitations are acting with $\sigma_\pm$ operations onto qubits where electrons are excited from/to. Leaving potential phase changes introduced by σ_z operators aside, they are acting as

$$\tilde{G}_{\mathbf{pq}} = i|o_{\mathbf{p}}^{\mathbf{q}}\rangle\langle o_{\mathbf{q}}^{\mathbf{p}}| - i|o_{\mathbf{q}}^{\mathbf{p}}\rangle\langle o_{\mathbf{p}}^{\mathbf{q}}| \tag{28}$$

where $o_{\mathbf{p}}^{\mathbf{q}}$ denotes all configurations where the $\mathbf{p}$ orbitals are occupied and the $\mathbf{q}$ orbitals are unoccupied with all combinations allowed for all other orbitals. Note, that the operators in eqn (28) neglect possible σ_z operations between the indexed qubits (that would, for example, be present in the single excitations of eqn (3)), those neglected operators will however not interfere with the structure of the nullspace and we will ignore them here for the sake of readability. The nullspace of those generators is formed by the collection of configurations which are neither denoted by $o_{\mathbf{p}}^{\mathbf{q}}$ nor $o_{\mathbf{q}}^{\mathbf{p}}$. A convenient way to construct the nullspace is over it's complement, which is given by all configuration onto which the generator acts non-trivial. The projector $1 - P_0$ onto the complement of the nullspace can then be constructed directly and rearranging leads to the expression for the nullspace projector as

$$\tilde{P}_{0;\mathbf{pq}} = 1 - |o_{\mathbf{p}}^{\mathbf{q}}\rangle\langle o_{\mathbf{p}}^{\mathbf{q}}| - |o_{\mathbf{q}}^{\mathbf{p}}\rangle\langle o_{\mathbf{q}}^{\mathbf{p}}|. \tag{29}$$

As an example, consider a two electron excitation generator that excites two electrons from the same spatial orbital to another spatial orbital, in a four spin-orbital representation this generator is $G = i|0011\rangle\langle 1100| - i|1100\rangle\langle 0011|$ with the nullspace projector $\tilde{P}_0 = 1 - |1100\rangle\langle 1100| - |0011\rangle\langle 1100| = 1 - Q_- \otimes Q_- \otimes Q_+ \otimes Q_+ - Q_+ \otimes Q_+ \otimes Q_- \otimes Q_-$, and $Q_\pm = \frac{1 \pm \sigma_z}{2}$. Within a larger basis, there will be more qubits, but the generator will act trivial (or introducing a phase) on them, leading to all possible combinations in the nullspace projector $\sum_{xy\ldots}|1100xy\ldots\rangle\langle 1100xy\ldots|$ that will sum up to unit operations as $\sum_{xy\ldots}|xy\ldots\rangle\langle xy\ldots| = 1$ ending up in the same expression as before. In general, the nullspace projector for an n-fold excitation can be constructed as

$$\tilde{P}_{0;\mathbf{pq}} = 1 - \prod_{i=1}^{n}(Q_-(\mathrm{p}_i)Q_+(\mathrm{q}_i) + Q_+(\mathrm{p}_i)Q_-(\mathrm{q}_i)). \tag{30}$$

2.5.2 Fermionic perspective. Constructing all generators directly in their fermionic representation will result in auto-

matic gradient evaluation schemes independent of the underlying qubit encoding, allowing flexible adaption of new encodings[51,52] and improved compiling strategies into quantum gates.[53] In analogy to the qubit construction in eqn (29) the nullspace projector of a general single, double or n-fold fermionic excitation (5) can be constructed as

$$P_{0;\mathrm{pq}} = 1 - N_\mathrm{p}\tilde{N}_\mathrm{q} - N_\mathrm{q}\tilde{N}_\mathrm{p} \tag{31}$$

$$P_{0;\mathrm{pqrs}} = 1 - N_\mathrm{p}\tilde{N}_\mathrm{q}N_\mathrm{r}\tilde{N}_\mathrm{s} - N_\mathrm{s}\tilde{N}_\mathrm{r}N_\mathrm{q}\tilde{N}_\mathrm{p} \tag{32}$$

$$P_{0;\mathbf{pq}} = 1 - N_{\mathrm{p}_0}\tilde{N}_{\mathrm{q}_0}\ldots N_{\mathrm{p}_\mathrm{n}}\tilde{N}_{\mathrm{q}_\mathrm{n}} - N_{\mathrm{q}_0}\tilde{N}_{\mathrm{p}_0}\ldots N_{\mathrm{q}_\mathrm{n}}\tilde{N}_{\mathrm{p}_\mathrm{n}}. \tag{33}$$

using fermionic particle and hole number operators $N_\mathrm{pq} = a_\mathrm{p}^\dagger a_\mathrm{q}$ and $\tilde{N}_\mathrm{pq} = 1 - N_\mathrm{pq} = a_\mathrm{p}a_\mathrm{q}^\dagger$. Note that using the Jordan–Wigner transformation (27), those operators get transformed into $Q_\pm$ (see also the next section) and eqn (33) gets transformed into eqn (30).

2.6 Additional cost in quantum gates

Compared to the original circuit, an implementation of the fermionic shift gate (18) will require additional gates, resulting from the U_0^α unitary that is generated by the nullspace projector of the fermionic generator. The individual cost of additional native quantum gates will depend on the qubit encoding of the fermionic algebra and individual properties of the underlying hardware, like their connectivity and native operations. See, for example, ref. 54 for an estimate of the resources required for different variants of unitary coupled-cluster. It can however be expected, that those details will affect the U_0^α unitary similar ways as the other fermionic unitaries. We will do a first estimate by analysing the P_0 projector (33) in the Jordan–Wigner encoding (27) given in eqn (30) and consider the P_0 projectors resulting from single and double excitations explicitly. For the single excitation P_0 projector we get the encoding

$$\tilde{P}_{0;\mathrm{pq}} = 1 - Q_{-,\mathrm{p}}Q_{+,\mathrm{q}} - Q_{-,\mathrm{q}}Q_{+,\mathrm{p}} = 1 - \frac{1}{2}(1 - \sigma_z(\mathrm{p})\sigma_z(q)). \tag{34}$$

The U_0^α gate for single excitations can then be implemented as a two qubit gate generated by $\sigma_z(\mathrm{p})\sigma_z(\mathrm{q})$. In the same way, the P_0 corresponding to a doubles excitation will result into 6 non trivial Pauli strings consisting of two σ_z operations and one assembled from four σ_z operations. Similar to the original fermionic generators, the P_0 projector of an n-fold fermionic excitation will decompose into $\mathcal{O}(2^{2n-1})$ individual Pauli strings which are in this case build up solely from σ_z operations with the largest one having σ_z on all $2n$ qubits. The gate cost for implementing the U_0^α unitary will therefore always be cheaper than for the associated fermionic excitation and we can upper bound it by the cost of implementing those.

3 Applications and examples

In the following we will illustrate potential initial applications for automatically differentiable unitary coupled-cluster. Note, that the techniques developed in this work allow convenient implementation of those techniques with computationally cheap gradients but can of course not guarantee that gradient based optimization schemes converge. The proposed applications are initial demonstrations of the automatic differentiation techniques developed within this work, applied to specific examples and should not be viewed as benchmarks or fully defined methods. Our aim is rather to provide a generalized toolbox for automatic differentiable unitary coupled-cluster, employable for the development of new methods. Our implementation is available within the free to use and open-source Tequila[36] package. The improved gradient evaluation schemes are automatically applied to already implemented methods like UpCCGSD[54] that was employed in previously developed basis-set-free methods[55] and as an application for the meta-VQE[56] approach. Apart from already existing implementations, the developed schemes can be employed for the development of new unitary coupled-cluster approaches in a blackboard style fashion. In the next sections we will illustrate this with explicit examples.

3.1 Using the implementation within Tequila

The techniques developed within this work are implemented in the free and open-source package Tequila[36] that operates on abstract expectation values of quantum circuits and operators, which can themselves be transformed and combined in an intuitive black-board style way. Our implementation is inspired by various open-source packages such as MADNESS,[57] PennyLane,[25] and DiffiQult,[26] that are not focused on unitary coupled-cluster, but offer intuitive application programming interfaces, exposing, often highly specialized, numerical algorithms to a broad audience of interested scientists. Tequila leverages state of the art high-performance simulators,[58,59] quantum chemistry software[57,60,61] and JAX[62] for extended automatic differentiation techniques. Given that access rights are permitted, Tequila can furthermore access state of the art quantum computers. Note however, that the quantum circuits of this work are still too deep in order to produce accurate results on current day hardware. Combined with the qubit-compressed, low-depth approaches of ref. 55 successful demonstration on emerging quantum computers can however be anticipated. For the results of this work we interfaced Qulacs[58] as simulation backend, the BFGS implementation of SciPy,[63] molecular integrals from Psi4 (ref. 64) and qubit encodings from OpenFermion.[60]

As an illustration of how our implementation can be employed we will give the explicit code that sequentially solves for ground and excited states of the hydrogen molecule in a minimal representation where we have two electrons in four spin-orbitals – STO-3G(2,4). The results in the following sections are obtained in similar ways. The ground and excited states are solved using eqn (37) and, in the explicit code example, we restrict the unitary to a single double excitation generated by $G_{(0,2),(1,3)}$ where two electrons are transferred

between the two spin-up orbitals (0 and 2) and spin-down orbitals (1 and 3).

```
import tequila as tq

# initialize molecular data
geom = "H 0.0 0.0 0.0\nH 0.0 0.0 0.7"
mol = tq.chemistry.Molecule(geometry=geom,
    basis_set="sto-3g",
    transformation="jordan_wigner")

H = mol.make_hamiltonian()

# index pairs for the double excitation
idx = [(0,2),(1,3)]

# construct ground-state objective
U0 = mol.prepare_reference()
U0 += mol.make_excitation_gate(idx, angle="a")
E0 = tq.ExpectationValue(H=H, U=U0)

# evaluate the objective at point 1.0
energy = tq.simulate(E0, {"a":1.0})

# compile and use as abstract function
energy_function = tq.compile(E0)
energy = energy_function({"a":1.0})

# optimize ground-state objective
result = tq.minimize("bfgs", E0)
variables=result.variables
energy0 = result.energy

# construct excited-state objective
U1 = mol.prepare_reference()
U1 += mol.make_excitation_gate(idx, angle="b")

Qp = tq.paulis.Qp(qubit=H.qubits)
E1 = tq.ExpectationValue(H=H, U=U1)
S2 = tq.ExpectationValue(H=Qp,U=U1+U0.dagger())

E = E1 - energy0*S2

# optimize excited-state objective
variables["b"] = 0.0
result = tq.minimize(method="bfgs",
    objective=E,
    variables=["b"],
    initial_values=variables)
energy1 = result.energy
```

since usually unitary coupled-cluster wavefunctions are real, the default gradient evaluation is done according to eqn (19). For complex wavefunctions the exact gradients can be demanded through further keywords in the make_-excitation_gate function. As our explicit code example illustrates, the gradient evaluation procedures are handled automatically by Tequila. Both states could also be optimized by minimizing the square gradient of the energy expectation value $\min_{\theta}\left(\frac{\partial\langle H\rangle U(\theta)}{\partial\theta}\right)^2$, a simplified strategy inspired by recent developments in classical mean-field theory.[32] Using structures from the code block above, the Tequila code block to directly deal with gradients of objectives looks like

```
E = tq.ExpectationValue(U=U0, H=H)
dE = tq.grad(E, "a")
dE2 = dE**2
```

and the dE2 objective can be used within the minimize function in the same way as illustrated above. The optimization will then converge to the ground or excited state depending on the optimization method and initial values. Note that gradient based optimization methods will then actually evaluate the gradient of $\left(\frac{\partial\langle H\rangle U(\theta)}{\partial\theta}\right)^2$ without further specifications necessary. The abstract objective for this gradient could however be obtained in the same way as before using the tq.grad operation. In Fig. 2 we plot the surfaces of those objectives explicitly. In the example code above we have illustrated how to evaluate the abstract objectives at specific angles. More details can be found in the online tutorials of Tequila.[36] Note that the direct minimization of the square of the gradient will in general not result in accurate results for excited state calculations due to the presence of flat plateaus in the parameterized objective.[65] More sophisticated optimization protocols are necessary to develop robust methods for this task providing good initial guesses and leveraging more generalized objective functions. It works for this basic illustration because of the simple shape of the one dimensional potential energy surface (see Fig. 2) and we are using it as a first example where the gradient directly enters the objective function in the hope that in can be the first step for future method development towards alternatives to amplitude amplification based approaches.[66]

3.2 An Illustrative example

We will use a simple example to further illustrate some the concepts of the last section. In particular we will demonstrate that the simplified gradient evaluation scheme of eqn (19) is not exact for complex wavefunctions. As a toy model we will use a specific circuit for a hydrogen molecule in 4 spatial orbitals (6-31G basis) – similar to the last section, just with more orbitals in

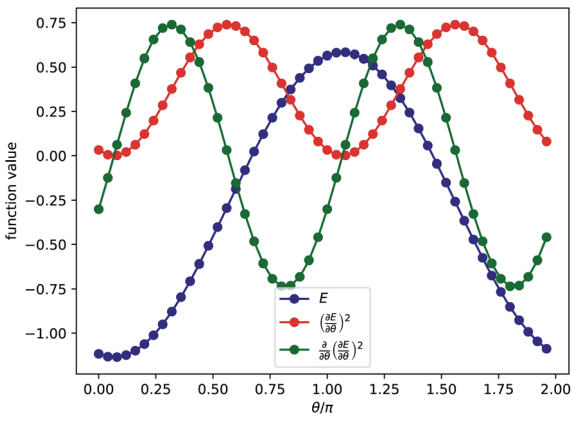

Fig. 2 Energy and gradients of a toy model: expectation value and gradients of the hydrogen molecule in a minimal representation with a unitary restricted to a single double excitation. See the main text for explicit illustration of how to construct the abstract expectation values and objectives with Tequila.

the model. Consider the fermionic excitation generator G_{0213} that will excite electrons from configurations $|0011\cdots\rangle$ to $|1100\cdot\rangle$, where the $\cdot$ can be arbitrary combinations of 0 and 1. The nullspace of this generator is spanned by all configurations whose bitstring representations do not start with 0011 or 1100. Consider the quantum circuit

$$U = \sigma_x(0)\sigma_x(1)e^{-i\frac{1}{2}G_{0213}}e^{-i\frac{1}{2}G_{0411}}e^{-i\frac{1}{2}A_{0213}}, \quad (35)$$

with $A \in \{G,G_+\}$. This circuit will prepare a superposition of the configurations $|11000000\rangle$, $|00110000\rangle$ and $|00001100\rangle$ and depending on the choice of the generator A the wavefunction will be real ($A = G$) or complex ($A = G_+$) since the unitary generated from $A = G_+$ will introduce a complex phase (see eqn (13)) to all configurations that are part of its nullspace (here $|00001100\rangle$). Now we will add a parameterized unitary generated by G_{0213} and take the expectation value with respect to the electronic Hamiltonian

$$\langle H\rangle U(\theta), \quad U(\theta) = Ue^{-i\frac{\theta}{2}G_{0213}}. \quad (36)$$

The gradients of this expectation value computed with different methods are shown in Fig. 3. For the complex wavefunction the approximate scheme where the two parts of the sum in eqn (16) are considered to be equivalent leads to slight deviations. This individual contributions to the exact gradient are also shown in Fig. 3. For the real wavefunction the individual contributions of $G_\pm$ are identical and the approximation becomes exact.

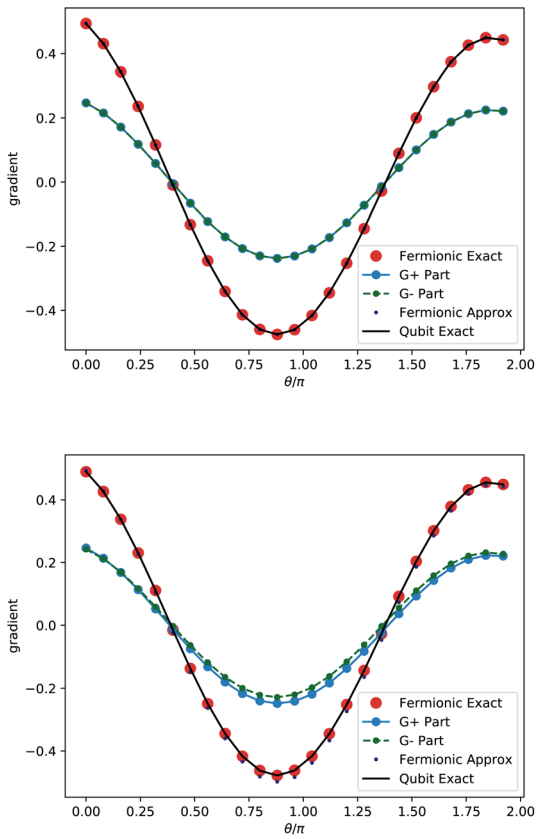

Fig. 3 An illustrative example: we show the gradients of the energy expectation value for the hydrogen molecule in 4 spatial orbitals (6-31G) with respect to a specific generator, preparing the wavefunctions in a way that ensures overlap with the nullspace of the generator (see main text). The underlying wavefunctions are real (top) and complex (bottom) and the gradients are computed according to eqn (16) (fermionic exact, Fig. 1 middle), where the two parts of the sum (G_+ and G_- corresponding to $\alpha \in \{+1,-1\}$) are also shown in the plots. The approximated fermionic gradient (Fig. 1 bottom) is according to eqn (19). The exact gradient in the qubit representation is computed by automatic differentiation of the individual rotational gates in the compiled circuit (Fig. 1 top).

3.3 Adapt VQE

Adaptive approaches iteratively increase the chain of excitations in the total unitary by adding operators from a operator pool through a, usually gradient based, screening process (see Fig. 4 for a high level overview). These type of algorithms have been successfully applied in different flavors like qubit-coupled-cluster[67–70] or adapt-vqe,[35,71] which mostly differ in the way they screen and construct operators. In these approaches, the commutator between the Hamiltonian and the generator of potential excitations is used in the screening process to compute the gradient. In order to perform the actual optimization with gradient based methods on a quantum computer, the commutator approach would only work for the gradient of the operator added last to the unitary circuit. With an automatically differentiable framework screening as well as optimization can be treated in the same way. This allows for more generalized adaptive growth procedures, where the adaptive part is not restricted to be the trailing part of the quantum circuit. In Fig. 5 we show some initial demonstrations combining static and adaptive blocks (see Fig. 4 for an illustration), using a restricted set of fermionic single and double excitations for the static blocks, where the double excitations are restricted to pair excitations in the same way as in the UpCCGSD approach.[54] As model systems we chose the H_4/STO-3G system as in ref. 54, with varying distance between the individual H_2 molecules, and the BeH_2/STO-3G molecule, with varying Be–H distance. By varying the corresponding distances we generate different problem instances that are representative of electronic structure problems without being too specialized.

3.4 Excited adapt VQE

As another application we use a modified version of the adaptive ground-state algorithm in order to optimize excited states, following the strategy applied in combination with the k-UpCCGSD[54] model of unitary coupled-cluster. A variational quantum algorithm for bound excited states (states with negative energies) can be achieved sequentially by projecting out previously solved solutions (see Fig. 4 for an illustration). Since it is known how to prepare previously found solutions with the unitaries U_i, the variational preparation of the target excited state becomes equal to the minimization of

$$E = \langle H\rangle_{U(\theta)} - \sum_i E_i\langle \mathbf{Q}_+\rangle_{U_i^\dagger U(\theta)}, \quad (37)$$

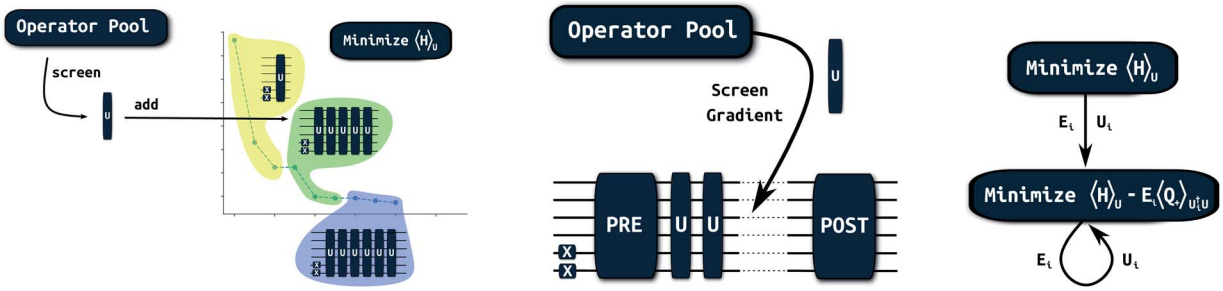

Fig. 4 Initial applications for automatically differentiable unitary coupled-cluster: basic procedure for adaptive circuit construction (left). Generalized framework combining adaptive and static blocks (middle). Sequential solver for excited states (right).

where $\langle O \rangle_U$ denotes the expectation value of the operator O with respect to the wave function prepared by the unitary U and the operator $\mathbf{Q}_+$ denotes the projector on the all-zero state

$$Q_+ = |0\ldots0\rangle\langle 0\ldots0| = \overset{N}{\otimes} Q_+. \tag{38}$$

The second term of eqn (37) computes the square of the overlap between the two wave functions, scaled by the energy of the previous state, and ensures orthogonality to all previously found solutions (see ESI†). Note that unbound states can in principle also be found with this approach by replacing the energies with large positive factors. Estimating overlaps like this is an alternative approach to the computationally costly SWAP test based strategies, originally proposed for excited state solvers in ref. 72 and successfully applied for similar systems as in this work by using imaginary time evolution in ref. 73. The same strategy for overlap estimation could be applied successfully in the optimization of quantum optical setups.[15] One important property of the $\mathbf{Q}_+$ operator is that, other than for the Hamiltonian H, all of its components naturally commute, allowing simultaneous sampling of all terms within a single run. Compared to the Hamiltonian H the additional measurements coming from $\mathbf{Q}_+$ are negligible.

We combine the sequential strategy with adaptive solvers by simply replacing the original objective function, which was just the expectation value of the Hamiltonian, with eqn (37) and solve sequentially for low lying excited states. In Fig. 6 we show some results for H_4 and BeH_2. In the case of BeH_2 we used the dominant contribution of the lowest configuration interaction singles solution as reference for the excited state calculation (see the ESI† for explicit circuits) and for H_4 we used the Hartree–Fock

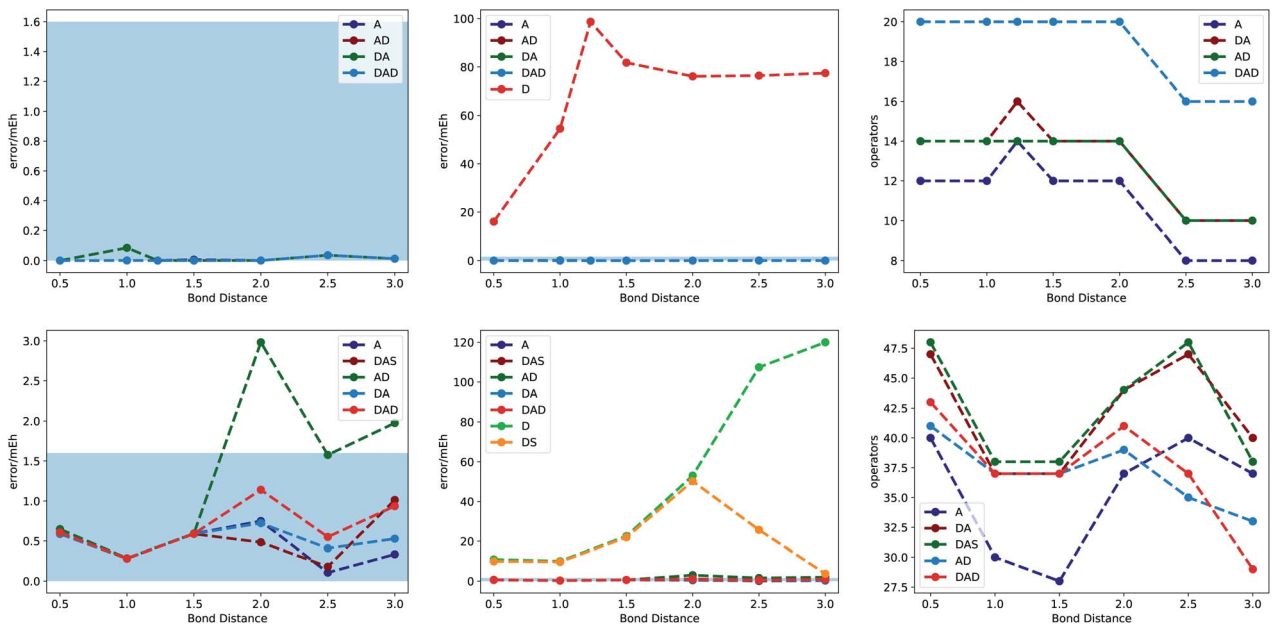

Fig. 5 Combining adaptive and static methods: combination of adaptively growing unitaries (A) with static blocks of generalized single (S) and double (D) excitations. Results are shown for H_4/STO-3G(4,8) (top) and BeH_2/STO-3G(6,14) (bottom). We used generalized doubles restricted to pair excitations for the D blocks (6 operators for H_4 and 15 for BeH_2). For the DAD results the generalized doubles block was split up into standard UpCCD doubles and the residual set of operators. In the DAS approach the trailing doubles are replaced with generalized single excitations. Excitations involving the 1s orbital of Be were not included into the static blocks D and S. The plots in the center show in addition the results without the adaptive blocks.

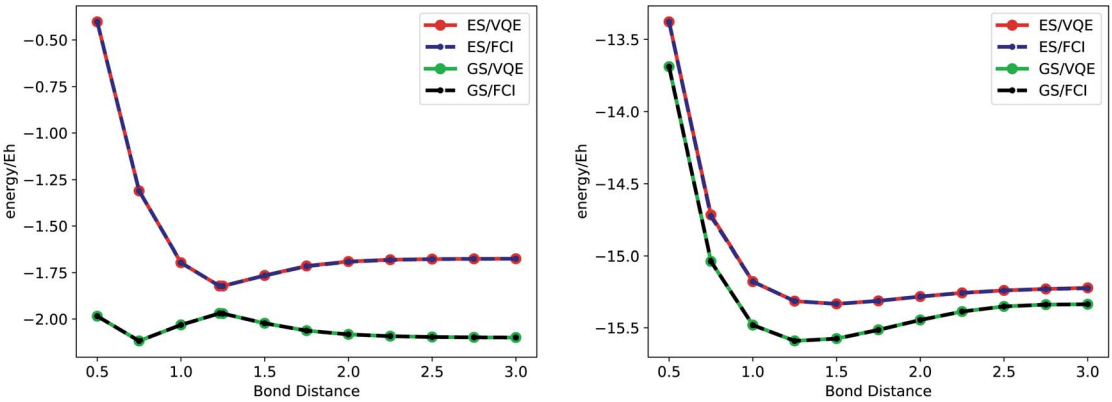

Fig. 6 Adapt-VQE for ground and excited states: adapt-VQE results using automatic differentiation for the screening and the optimization for ground and first excited state energies of H_4/STO-3G(4,8) (left) and BeH_2/STO-3G(6,14) (right). Except for the last point of the BeH_2 excited state, all points agree to millihartree accuracy with the corresponding exact solution (FCI) in the given basis set. See also Fig. 4. We included the special point at distance 1.23 Å with a square configuration.

reference for both calculations. The reason for this being that the lowest lying electronic singlet state of H_4 has the same irreducible representation as the ground-state which is not the case for BeH_4. So this particular excited state of BeH_4 could have also been found with a ground-state algorithm with a symmetry restricted pool of excitations. In both cases, we restricted the adaptive optimization to stay within the singlet manifold, in order to avoid collapsing to the lower lying triplet states. Note, that commutator based screening techniques are not possible in this case since the adaptive circuit is not the trailing part of the overlap expectation value. In order to make them work, the $U_i^\dagger$ unitaries would have to be folded into the P_0 operator, increasing the number of measurements significantly (see ref. 15 for a similar argument). Such folding techniques are achieved in an iterative version of qubit-coupled-cluster,[68] using the properties of self-inverse qubit generators. Eqn (14) could be employed to develop similar strategies in the fermionic representation.

4 Conclusion and outlook

Variational algorithms are currently one of the most promising applications on current and future quantum computers. Quantum chemistry is one of the target fields of those algorithms and expectations are high for new types of methods developed within that framework. Analytical gradients for unitary coupled-cluster type approaches can in principle be computed on quantum computers, they come however with high computational costs using standard techniques on the qubit level. We developed the necessary techniques in order to evaluate analytical gradients of general n-fold fermionic excitation operators with a cost factor of 4 in general and factor of 2 for real wavefunctions. Our strategies to compile gradients can be done entirely in the fermionic representation making it independent of the used qubit mapping. The developed techniques combined with Tequilas automatic differentiation framework provide a testbed for quantum chemistry on quantum computers where new ideas, like low-depth approaches based on pair-natural orbitals[55] or Krylov subspaces,[74,75] can be prototyped and demonstrated in a blackboard fashion. Our implementation provides an easy to use, automatically differentiable framework for unitary-coupled cluster, that leverages state of the art high performance simulators[58,59] and is ready for emerging quantum computers. We demonstrated initial applications for ground and excited state calculations for small model systems where we extended adaptive circuit construction schemes and, for the first time, applied them to excited state optimization.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We would like to thank Philipp Schleich and Sumner Alperin-Lea for providing valuable suggestions and comments on the manuscript. A. A.-G. acknowledges the generous support from Google, Inc. in the form of a Google Focused Award. This work was supported by the U.S. Department of Energy under Award No. DE-SC0019374 and the U.S. Office of Naval Research (ONS-506661). A. A.-G. also acknowledges support from the Canada Industrial Research Chairs Program and the Canada 150 Research Chairs Program. Computations were performed on the niagara supercomputer at the SciNet HPC Consortium.[76,77] SciNet is funded by: the Canada Foundation for Innovation; the Government of Ontario; Ontario Research Fund – Research Excellence; and the University of Toronto. We thank the generous support of Anders G. Frøseth.

Notes and references

1 A. Aspuru-Guzik, A. D. Dutoi, P. J. Love and M. Head-Gordon, *Science*, 2005, **309**, 1704–1707.
2 A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik and J. L. O'brien, *Nat. Commun.*, 2014, **5**, 4213.

3 J. R. McClean, J. Romero, R. Babbush and A. Aspuru-Guzik, *New J. Phys.*, 2016, **18**, 023023.
4 Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. Sawaya, *et al.*, *Chem. Rev.*, 2019, **119**, 10856–10915.
5 S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin and X. Yuan, *Rev. Mod. Phys.*, 2020, **92**, 015003.
6 M. Schuld, R. Sweke and J. J. Meyer, *The effect of data encoding on the expressive power of variational quantum machine learning models*, 2020.
7 J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe and S. Lloyd, *Nature*, 2017, **549**, 195–202.
8 A. Pérez-Salinas, A. Cervera-Lierta, E. Gil-Fuster and J. I. Latorre, *Quantum*, 2020, **4**, 226.
9 M. Schuld and N. Killoran, *Phys. Rev. Lett.*, 2019, **122**, 040504.
10 M. Schuld, A. Bocharov, K. M. Svore and N. Wiebe, *Phys. Rev. A*, 2020, **101**, 032308.
11 J. Romero, J. P. Olson and A. Aspuru-Guzik, *Quantum Sci. Technol.*, 2017, **2**, 045001.
12 A. Anand, J. Romero, M. Degroote and A. Aspuru-Guzik, arXiv preprint, 2020, arXiv:2006.01976.
13 L. B. Kristensen, M. Degroote, P. Wittek, A. Aspuru-Guzik and N. T. Zinner, *An Artificial Spiking Quantum Neuron*, 2019.
14 E. Farhi, J. Goldstone and S. Gutmann, arXiv preprint, 2014, arXiv:1411.4028.
15 J. S. Kottmann, M. Krenn, T. H. Kyaw, S. Alperin-Lea and A. Aspuru-Guzik, arXiv preprint, 2020, arXiv:2006.03075.
16 J. Carolan, M. Mohseni, J. P. Olson, M. Prabhu, C. Chen, D. Bunandar, M. Y. Niu, N. C. Harris, F. N. C. Wong, M. Hochberg, S. Lloyd and D. Englund, *Nat. Phys.*, 2020, **16**, 322–327.
17 N. P. Sawaya, T. Menke, T. H. Kyaw, S. Johri, A. Aspuru-Guzik and G. G. Guerreschi, *npj Quantum Inf.*, 2020, **6**, 1–13.
18 See for example https://quantum-computing.ibm.com/for IBM's open access quantum computers.
19 F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell, *et al.*, *Nature*, 2019, **574**, 505–510.
20 X.-L. Wang, Y.-H. Luo, H.-L. Huang, M.-C. Chen, Z.-E. Su, C. Liu, C. Chen, W. Li, Y.-Q. Fang, X. Jiang, *et al.*, *Phys. Rev. Lett.*, 2018, **120**, 260502.
21 T. Rudolph, *APL Photonics*, 2017, **2**, 030901.
22 R. P. Feynman, *Int. J. Theor. Phys.*, 1982, **21**, 467–488.
23 M. Reiher, N. Wiebe, K. M. Svore, D. Wecker and M. Troyer, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 7555–7560.
24 M. Schuld, V. Bergholm, C. Gogolin, J. Izaac and N. Killoran, *Phys. Rev. A*, 2019, **99**, 032331.
25 V. Bergholm, J. Izaac, M. Schuld, C. Gogolin, C. Blank, K. McKiernan and N. Killoran, arXiv preprint, 2018, arXiv:1811.04968.
26 T. Tamayo-Mendoza, C. Kreisbeck, R. Lindh and A. Aspuru-Guzik, *ACS Cent. Sci.*, 2018, **4**, 559–566.
27 F. Pavošević and S. Hammes-Schiffer, *Automatic differentiation for coupled cluster methods*, 2020.
28 U. Ekström, L. Visscher, R. Bast, A. J. Thorvaldsen and K. Ruud, *J. Chem. Theory Comput.*, 2010, **6**, 1971–1980.
29 R. Steiger, C. H. Bischof, B. Lang and W. Thiel, *Future Gener. Comput. Syst.*, 2005, **21**, 1324–1332.
30 S. Sorella and L. Capriotti, *J. Chem. Phys.*, 2010, **133**, 234111.
31 C. Song, T. J. Martínez and J. B. Neaton, 2020.
32 J. A. Shea, E. Gwin and E. Neuscamman, *J. Chem. Theory Comput.*, 2020, **16**, 1526–1540.
33 F. A. Evangelista, G. K.-L. Chan and G. E. Scuseria, *J. Chem. Phys.*, 2019, **151**, 244112.
34 A. F. Izmaylov, M. Díaz-Tinoco and R. A. Lang, *Phys. Chem. Chem. Phys.*, 2020, **22**, 12980–12986.
35 H. R. Grimsley, S. E. Economou, E. Barnes and N. J. Mayhall, *Nat. Commun.*, 2019, **10**, 1–9.
36 J. S. Kottmann, S. Alperin-Lea, T. Tamayo-Mendoza, A. Cervera-Lierta, C. Lavigne, T.-C. Yen, V. Verteletskyi, P. Schleich, A. Anand, M. Degroote, S. Chaney, M. Kesibi, A. F. Izmaylov and A. Aspuru-Guzik, arXiv preprint, 2020, arXiv:2011.03057.
37 T. Helgaker, P. Jorgensen and J. Olsen, *Molecular electronic-structure theory*, John Wiley & Sons, 2014.
38 I. Shavitt and R. J. Bartlett, *Many-body methods in chemistry and physics: MBPT and coupled-cluster theory*, Cambridge University Press, 2009.
39 P. R. Surján, *Second quantized approach to quantum chemistry: an elementary introduction*, Springer Science & Business Media, 2012.
40 P. Jørgensen, *Second quantization-based methods in quantum chemistry*, Elsevier, 2012.
41 L. Banchi and G. E. Crooks, *Measuring Analytic Gradients of General Quantum Evolution with the Stochastic Parameter Shift Rule*, 2020.
42 A. Mari, T. R. Bromley and N. Killoran, *Estimating the gradient and higher-order derivatives on quantum hardware*, 2020.
43 T. Jones and J. Gacon, *Efficient calculation of gradients in classical simulations of variational quantum algorithms*, 2020.
44 J. J. Meyer, J. Borregaard and J. Eisert, *A variational toolbox for quantum multi-parameter estimation*, 2020.
45 G. E. Crooks, arXiv preprint, 2019, arXiv:1905.13311.
46 J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love and A. Aspuru-Guzik, *Quantum Sci. Technol.*, 2018, **4**, 014008.
47 S. B. Bravyi and A. Y. Kitaev, *Ann. Phys.*, 2002, **298**, 210–226.
48 J. T. Seeley, M. J. Richard and P. J. Love, *J. Chem. Phys.*, 2012, **137**, 224109.
49 A. Tranter, S. Sofia, J. Seeley, M. Kaicher, J. McClean, R. Babbush, P. V. Coveney, F. Mintert, F. Wilhelm and P. J. Love, *Int. J. Quantum Chem.*, 2015, **115**, 1431–1441.
50 A. Tranter, P. J. Love, F. Mintert and P. V. Coveney, *J. Chem. Theory Comput.*, 2018, **14**, 5617–5630.
51 K. Setia and J. D. Whitfield, *J. Chem. Phys.*, 2018, **148**, 164104.
52 R. W. Chien and J. D. Whitfield, *Custom fermionic codes for quantum simulation*, 2020.
53 Y. S. Yordanov and C. H. W. Barnes, *Efficient quantum circuits for quantum computational chemistry*, 2020.
54 J. Lee, W. J. Huggins, M. Head-Gordon and K. B. Whaley, *J. Chem. Theory Comput.*, 2018, **15**, 311–324.
55 J. S. Kottmann, P. Schleich, T. Tamayo-Mendoza and A. Aspuru-Guzik, *J. Phys. Chem. Lett.*, 2021, **12**, 663–673.

56 A. Cervera-Lierta, J. S. Kottmann and A. Aspuru-Guzik, arXiv preprint, 2020, arXiv:2009.13545.
57 R. J. Harrison, G. Beylkin, F. A. Bischoff, J. A. Calvin, G. I. Fann, J. Fosso-Tande, D. Galindo, J. R. Hammond, R. Hartman-Baker, J. C. Hill, *et al.*, *SIAM J. Sci. Comput.*, 2016, **38**, S123–S142.
58 Y. Suzuki, Y. Kawase, Y. Masumura, Y. Hiraga, M. Nakadai, J. Chen, K. M. Nakanishi, K. Mitarai, R. Imai, S. Tamiya, T. Yamamoto, T. Yan, T. Kawakubo, Y. O. Nakagawa, Y. Ibe, Y. Zhang, H. Yamashita, H. Yoshimura, A. Hayashi and K. Fujii, *Qulacs: a fast and versatile quantum circuit simulator for research purpose*, 2020.
59 S. Efthymiou, S. Ramos-Calderer, C. Bravo-Prieto, A. Pérez-Salinas, D. García-Martín, A. Garcia-Saez, J. I. Latorre and S. Carrazza, *Qibo: a framework for quantum simulation with hardware acceleration*, 2020.
60 J. R. McClean, K. J. Sung, I. D. Kivlichan, Y. Cao, C. Dai, E. S. Fried, C. Gidney, B. Gimby, P. Gokhale, T. Häner, T. Hardikar, V. Havlíček, O. Higgott, C. Huang, J. Izaac, Z. Jiang, X. Liu, S. McArdle, M. Neeley, T. O'Brien, B. O'Gorman, I. Ozfidan, M. D. Radin, J. Romero, N. Rubin, N. P. D. Sawaya, K. Setia, S. Sim, D. S. Steiger, M. Steudtner, Q. Sun, W. Sun, D. Wang, F. Zhang and R. Babbush, *OpenFermion: The Electronic Structure Package for Quantum Computers*, 2017.
61 R. M. Parrish, L. A. Burns, D. G. Smith, A. C. Simmonett, A. E. DePrince III, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, *et al.*, *J. Chem. Theory Comput.*, 2017, **13**, 3185–3197.
62 J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin and S. Wanderman-Milne, *JAX: composable transformations of Python+NumPy programs*, 2018, http://github.com/google/jax.
63 P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. Jarrod Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. Carey, İ. Polat, Y. Feng, E. W. Moore, J. Vand erPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt and S. Contributors, *Nat. Methods*, 2020, **17**, 261–272.
64 P. Virtanen, R. Gommers, T. E. Oliphant, et al., *Nat. Methods*, 2020, **17**, 261–272.
65 J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush and H. Neven, *Nat. Commun.*, 2018, **9**, 1–6.
66 P. W. K. Jensen, L. B. Kristensen, J. S. Kottmann and A. Aspuru-Guzik, *Quantum Sci. Technol.*, 2020, **6**, 015004.
67 I. G. Ryabinkin, T.-C. Yen, S. N. Genin and A. F. Izmaylov, *J. Chem. Theory Comput.*, 2018, **14**, 6317–6326.
68 I. G. Ryabinkin, R. A. Lang, S. N. Genin and A. F. Izmaylov, *J. Chem. Theory Comput.*, 2020, **16**, 1055–1063.
69 R. A. Lang, I. G. Ryabinkin and A. F. Izmaylov, arXiv preprint, 2020, arXiv:2002.05701.
70 Z.-J. Zhang, T. H. Kyaw, J. S. Kottmann, M. Degroote and A. Aspuru-Guzik, arXiv preprint, 2020, arXiv:2008.07553.
71 H. L. Tang, E. Barnes, H. R. Grimsley, N. J. Mayhall and S. E. Economou, arXiv preprint, 2019, arXiv:1911.10205.
72 O. Higgott, D. Wang and S. Brierley, *Quantum*, 2019, **3**, 156.
73 T. Jones, S. Endo, S. McArdle, X. Yuan and S. C. Benjamin, *Phys. Rev. A*, 2019, **99**, 062304.
74 W. J. Huggins, J. Lee, U. Baek, B. O'Gorman and K. B. Whaley, *A Non-Orthogonal Variational Quantum Eigensolver*, 2019.
75 N. H. Stair, R. Huang and F. A. Evangelista, *A Multireference Quantum Krylov Algorithm for Strongly Correlated Electrons*, 2019.
76 M. Ponce, R. van Zon, S. Northrup, D. Gruner, J. Chen, F. Ertinaz, A. Fedoseev, L. Groer, F. Mao and B. C. Mundim, *et al.*, in *Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (learning)*, 2019, pp. 1–8.
77 C. Loken, D. Gruner, L. Groer, R. Peltier, N. Bunn, M. Craig, T. Henriques, J. Dempsey, C.-H. Yu, J. Chen, *et al.*, *J. Phys.: Conf. Ser.*, 2010, 012026.