Greedy Gradient-free Adaptive Variational Quantum Algorithms on a Noisy Intermediate Scale Quantum Computer

César Feniou[1,2], **Baptiste Claudon**[2], **Muhammad Hassan**[3], **Axel Courtat**[2], **Olivier Adjoua**[1], **Yvon Maday**[3,4]**, and Jean-Philip Piquemal**[1,2,*]

[1]Sorbonne Université, Laboratoire de Chimie Théorique (UMR-7616-CNRS), F-75005 Paris, France
[2]Qubit Pharmaceuticals, Advanced Research Department, Paris, France
[3]Sorbonne Université, CNRS, Université Paris Cité, Laboratoire Jacques-Louis Lions (LJLL), F-75005 Paris, France
[4]Institut Universitaire de France, 75005, Paris, France
[*]jean-philip.piquemal@sorbonne-universite.fr

ABSTRACT

Hybrid quantum-classical adaptive Variational Quantum Eigensolvers (VQE) already hold the potential to outperform classical computing for simulating quantum many-body systems. However, their practical implementation on current quantum processing units (QPUs) is very challenging due to the noisy evaluation of a polynomially scaling number of observables, undertaken for operator selection and optimisation of a high-dimensional cost function. To overcome this, we propose new techniques to execute adaptive algorithms on a 25-qubit error-mitigated QPU coupled to a GPU-accelerated HPC simulator. Targeting physics applications, we compute the ground state of a 25-body Ising model using the newly introduced Greedy Gradient-free Adaptive VQE (CGA-VQE) requiring only five circuit measurements per iteration, regardless of the number of qubits and size of the operator pool. Towards chemistry, we combine the GGA-VQE and Overlap-ADAPT-VQE algorithms to approximate a molecular system ground state. We show that the QPU successfully executes the algorithms and yields the correct choice of parametrised unitary operators. While the QPU evaluation of the resulting ansatz wave-function is polluted by hardware noise, a single final evaluation of the sought-after observables on a classical GPU-accelerated/noiseless simulator allows the recovery of the correct approximation of the ground state, thus highlighting the need for hybrid quantum-classical observable measurement.

1 Introduction

Quantum computing has gained considerable interest due to its potential to solve complex computational problems that are intractable on classical devices. Finding the ground state of a many-body quantum system is one such problem as it suffers from an exponentially scaling complexity in the system size[1]. Quantum computing provides an appealing solution as it allows, in principle, the encoding of the exponentially scaling many-body wave-function onto a linearly scaling qubit register. In the context of quantum chemistry, extensive efforts have been made to develop quantum algorithms for ground and excited state preparation of molecular systems with the goal of ultimately surpassing classical techniques[2,3]. In order to be able to take advantage of near-term quantum devices in the noisy intermediate-scale quantum (NISQ) era, emphasis has been placed on the development of hybrid quantum-classical algorithms such as the Variational Quantum Eigensolver (VQE) which incorporates a quantum subroutine within a classical optimization loop thereby reducing the quantum computer workload and mitigating the effects of hardware noise and measurement errors[4].

The core idea of the variational quantum eigensolver is to generate a parameterised wave-function, known as the ansatz, and then variationally tune this ansatz so as to minimise the expectation value of some relevant Hermitian operator, typically the system Hamiltonian. The fundamental challenge in implementing the VQE methodology on NISQ devices is thus to construct an ansatz wave-function that can accurately describe the ground-state of the Hermitian operator under study and, at the same time, can be represented on shallow quantum circuits which are not dominated by noise. Most commonly used VQEs for quantum chemistry are so-called "fixed-ansatz" methods wherein the ansatz wave-function consists of a predetermined product of parametrised unitary operators acting on an initial (usually Hartree-Fock reference) state[5–10]. Whether hardware-efficient or chemically-inspired, these "fixed" ansatz methods have a limited accuracy and do not provide a route for exact simulations of strongly correlated systems on near-term quantum hardware[6,8]. Since fixed ansätze are by definition system-agnostic, they are also likely to contain superfluous operators that do not contribute to a better approximation of the ground

state of the Hermitian operator under study[11]. Such redundant operators needlessly increase the length of the ansatz circuit as well as the number of variational parameters to be tuned, both of which are serious problems for NISQ-era devices.

Unsurprisingly therefore, recent works have proposed iterative VQE protocols that construct system-tailored ansätze using some kind of quasi-greedy strategy. The ADAPT-VQE algorithm[12] has made a notable impact in the field by demonstrating a significant reduction in the redundant terms in the ansatz circuits for a range of molecules, thus enhancing the accuracy and efficiency of the VQE. At it's core, the ADAPT-VQE algorithm consists of two steps:

Step 1 Given an initial parameterised ansatz wave-function $|\Psi^{(m)}\rangle := |\Psi^{(m)}(\theta'_m, \ldots, \theta'_1)\rangle$ at the mth iteration, a new ansatz wave-function is constructed by carefully selecting a parameterised unitary operator from a pre-selected operator pool and appending this parameterised unitary operator to the current parameterised ansatz wave-function $|\Psi^{(m)}(\theta'_m, \ldots, \theta'_1)\rangle$. More precisely, given a pool $\mathbb{U}$ of parameterised unitary operators, a Hermitian operator $\widehat{A}$ whose ground state is being prepared, the ADAPT-VQE selection criterion consists of identifying the unitary operator $\mathscr{U}^* \in \mathbb{U}$ such that

$$\mathscr{U}^* = \underset{\mathscr{U} \in \mathbb{U}}{\operatorname{argmax}} \left| \frac{d}{d\theta} \left\langle \Psi^{(m)}(\theta'_m, \ldots, \theta'_1) \middle| \mathscr{U}(\theta)^\dagger \widehat{A} \mathscr{U}(\theta) \middle| \Psi^{(m)}(\theta'_m, \ldots, \theta'_1) \right\rangle \Big|_{\theta=0} \right|. \tag{1}$$

This results in a new parametrised ansatz wave-function

$$|\widetilde{\Psi}^{(m+1)}\rangle := |\widetilde{\Psi}^{(m+1)}(\theta_{m+1}, \theta'_m, \ldots, \theta'_1)\rangle := \mathscr{U}^*(\theta_{m+1}) |\Psi^{(m)}(\theta'_m, \ldots, \theta'_1)\rangle$$

. Note that, at this stage, θ_{m+1} is a free parameter whose value has not been fixed.

Step 2 The new parametrised ansatz wave-function at iteration $m+1$ is then obtained by performing a <u>global optimisation</u> over all parameters $\theta_1, \ldots, \theta_{m+1}$ of the expectation value of the Hermitian operator under study. More precisely, we solve the $m+1$-dimensional optimisation problem

$$(\theta'_1, \ldots, \theta'_{m-1}, \theta'_m) := \underset{\theta_1, \ldots, \theta_{m-1}, \theta_m}{\operatorname{argmin}} \left\langle \widetilde{\Psi}^{(m+1)}(\theta_{m+1}, \theta_m, \ldots, \theta_1) \middle| \widehat{A} \middle| \widetilde{\Psi}^{(m+1)}(\theta_{m+1}, \theta_m, \ldots, \theta_1) \right\rangle,$$

and define the new ansatz wave-function at iteration $m+1$ as $|\Psi^{(m+1)}\rangle := |\Psi^{(m+1)}(\theta'_{m+1}, \theta'_m, \ldots, \theta'_1)\rangle$.

It can now readily be seen that the main computational bottleneck in the ADAPT-VQE methodology, which incidentally arises also "fixed-ansatz" methods, is the global optimisation step wherein the ansatz wave-function is variationally tuned to minimise the associated expectation value. Indeed, the cost function for this second step, since it arises from measurements on a NISQ device, is both high-dimensional and extremely noisy, thus often rendering the associated optimisation problem computationally intractable[13]. As a representative example, a single measurement of the expectation value of a molecular Hamiltonian represented in second quantised-form using N spin-orbitals requires $\mathcal{O}(N^4)$ individual measurements on the quantum device. It is thus hardly surprising that the total number of measurements required for a successful ADAPT-VQE procedure on even the smallest molecules requires tens of thousands of measurements (see Table 1 below).

Molecule	Noise (Y/N)	ADAPT iterations	ADAPT inner optimiser iterations	Gradient screening measurements	Total measurements
LiH	N	4	8+9+16+20 = 53	565,248	576,876
LiH	Y	4	22+39+148+70 = 279	565,248	642,252
HF	N	1	17	141,312	146,004
HF	Y	1	28	141,312	149,040

Table 1. Total number of measurements required for successful QEB-ADAPT-VQE[14] procedures that yield a chemically accurate (i.e., with an error lower than 10^{-3} Ha) ansatz wave-functions. For both molecules, the Hamiltonian was discretised using a minimal STO-3G basis set. Both resulting molecular Hamiltonians were composed of 276 terms. The operator pool for both molecules consists of 44 different unitary operators. The simulations with hardware noise involved 2500 shots. Both molecules posses a stretched geometry with a bong length of 2.5 Angstrom. All simulations used a classical BFGS optimiser.

In light of these difficulties, it is unsurprising that practical implementations of ADAPT-VQE-type algorithms on the current generation of quantum hardware are yet to be achieved. Interestingly, while a great deal of effort has been devoted to developing various improvements of the original ADAPT procedure, a majority of this research has focused on further compactifying the adaptively generated ansatz wave-functions, thus reducing the quantum circuit depth and the number of CNOT gates required to represent the ansätze (see[14–18]). Unfortunately, these improvements either do not deal with, or even worse,

come at the expense of further increasing the number of measurements on the quantum device required to perform the iterative procedure. Comparatively fewer articles in the literature have focused on decreasing the quantum measurement overhead required to implement ADAPT-VQE-type methods on current NISQ devices, and despite some notable advancements, (see, e.g., the energy evaluation schemes[19–21]), the gap between the quantum resources afforded by the current generation of quantum hardware and those required by these improved adaptive algorithms is yet to be bridged.

The goal of the current study is to present a noise-resistant and resource-efficient, greedy gradient-free adaptive variational quantum algorithm that can successfully be implemented on the quantum devices of today. The algorithm that we introduce is motivated by the gradient-free, analytical optimisation approaches that have been proposed in the VQE literature on 'fixed-ansatz' methods. Indeed, it is well-known that the expectation value of a Hermitian operator with respect to an ansatz wave-function $\Psi(\theta)$ parametrised by a single quantum gate is simply an elementary trigonometric function of θ[22–27]. Thus, for a wave-function parametrised by a single rotation gate, only two measurements of the expectation value for judiciously chosen values of θ, allow an exact reconstruction of the full expectation value as a parametrised function of θ. Using this insight, we have replaced the conventional ADAPT-VQE operator selection criterion with a newly developed gradient-free energy sorting approach that allows us to identify the locally optimal parametrised unitary operator and the associated optimal angle, which when appended to the current ansatz wave-function, will produce a new ansatz wave-function with the biggest drop in expectation value. In other words, in contrast to the ADAPT-VQE gradient-based criterion (1), our adaptive algorithm selects a locally optimal unitary operator $\mathscr{U}^*$ and associated optimal angle θ'_{m+1} that satisfy

$$\mathscr{U}^* = \operatorname*{argmin}_{\mathscr{U} \in \mathbb{U}} \min_{\theta \in [-\pi, \pi)} \langle \Psi^{(m)}(\theta'_m, \ldots, \theta'_1) | \mathscr{U}(\theta)^\dagger \widehat{A} \mathscr{U}(\theta) | \Psi^{(m)}(\theta'_m, \ldots, \theta'_1) \rangle \quad \text{and} \tag{2}$$

$$\theta'_{m+1} = \operatorname*{argmin}_{\theta \in [-\pi, \pi)} \min_{\mathscr{U} \in \mathbb{U}} \langle \Psi^{(m)}(\theta'_m, \ldots, \theta'_1) | \mathscr{U}(\theta)^\dagger \widehat{A} \mathscr{U}(\theta) | \Psi^{(m)}(\theta'_m, \ldots, \theta'_1) \rangle.$$

As we discuss in Section 2.3 below, the one-dimensional objective functions appearing in the optimisation problem (2), also known as landscape functions, can be expressed as analytical functions of θ for any unitary operator $\mathscr{U}$ belonging to popular choices of operator pools. Moreover, these analytical landscape functions can be determined explicitly using a fixed number of measurements that depends on the nature of the operator pool and the Hermitian operator $\widehat{A}$. Consequently, using a minimal number of measurements of the quantum device, we can immediately determine both the best unitary operator $\mathscr{U}^*$ and the optimal angle θ'_{m+1} that should be used to update the current ansatz wave-function. By iteratively growing an ansatz wave-function using only such locally optimal parametrised unitary operators and not re-optimising the "frozen-core" of the previous ansatz, we are able to eschew entirely the need for a global optimisation of a multi-dimensional noisy objective function. We refer to this adaptive algorithm as the greedy, gradient-free adaptive variational quantum eigensolver or GGA-VQE algorithm for short (see Section 2.3 below for a detailed description). Let us remark here that while gradient-free, analytical optimisation approaches for VQEs have been explored in the 'fixed-ansatz' literature[22–25] and energy-sorting algorithms to improve the ADAPT-VQE operator selection criterion have also been proposed[28, 29], to the best of our knowledge, our work is the first attempt to combine both approaches and develop a greedy gradient-free adaptive variational quantum VQE.

Equipped with this resource-efficient methodology, we explore practical implementations of such greedy gradient-free adaptive algorithms on quantum devices. For our first numerical experiment, we consider the ground state preparation of an open boundary, one-dimensional transverse-field Ising model. We show that for Ising Hamiltonians of this nature, using a minimal hardware-efficient operator pool[14], each iteration of the GGA-VQE algorithm requires measuring only five observables on quantum circuits, regardless of the system size (i.e., the number of qubits involved). As a proof of concept, we run the GGA-VQE algorithm for such an Ising model on a 25-qubit register on a state-of-the-art, trapped ion quantum computer and successfully achieve a ground state fidelity of over 98%.

Our second numerical experiment, on the same trapped ion quantum computer, pertains to the recently developed Overlap-ADAPT-VQE algorithm[18] that seeks to generate a compact approximation of a target wave-function through an iterative, adaptive overlap maximisation procedure. We consider a stretched hydrogen fluoride (HF) molecular system, and we take as the target wave-function, an approximate ground-state generated through a classical QEB-ADAPT-VQE procedure[15]. We then apply the Overlap-GGA-VQE algorithm (described in detail in Section 2.5) to progressively grow an ansatz wave-function that achieves an overlap of over 99% with the target wave-function. Since the Overlap-ADAPT-VQE algorithm requires measuring wave-function overlaps on quantum devices, we consider two possible methods– each requiring different quantum resources– that may be employed for this purpose. These are the so-called compute-uncompute approach, which requires a deeper circuit but no additional qubits to perform overlap measurements, and the Swap test method, which utilises a second qubit register to compute the wave-function overlaps but has the advantage of not increasing the circuit depth[30]. Thus, our work also provides an empirical investigation of the merits of each approach for overlap computations.

Finally, let us emphasise that the methodology that we introduce in this study, while motivated by the aforementioned goal of attaining practical realisations of adaptive variational algorithms on current NISQ devices, can also be used for the purpose

of <u>compactifying</u> adaptive ansatz wave-functions at the cost of increasing the measurement overhead on the quantum device. Such generalisations are briefly discussed in Section 2.3 below, and a study of the capabilities of our methodology in relation to other works on generating ultra-compact adaptive ansatz wave-functions will be the subject of future work.

2 Methods

2.1 The Adaptive Derivative-Assembled Pseudo-Trotter Variational Quantum Eigensolver

The adaptive derivative-assembled pseudo-Trotter variational quantum eigensolver (ADAPT-VQE)[12] is a VQE-inspired algorithm designed to approximate the ground state wave-function and ground state energy of a given Hamiltonian. Unlike many other classical variational quantum eigensolvers such as the various flavours of trotterised unitary coupled cluster[5–10] however, ADAPT-VQE does not specify a fixed ansatz for the sought-after ground state at the beginning of the algorithm. Instead, ADAPT-VQE functions by first fixing a set of admissible Hermitian generators (the so-called operator pool). The ansatz wave-function is then grown iteratively by parametrically exponentiating a carefully selected Hermitian generator, appending this exponentiated generator to the previous ansatz wave-function, and then variationally tuning the new ansatz wave-function. Since the selection procedure is tailored to the specific Hamiltonian system under consideration (see below), one usually hopes to obtain a more compact ansatz than the one generated by non-adaptive VQEs while still retaining the practical advantages of the VQE for near-term quantum hardware.

The general workflow of the ADAPT-VQE algorithm is as follows. Given the qubit representation of an input Hamiltonian H, a pool of admissible Hermitian generators $\mathbb{P}$, and a stopping criterion:

1. Boot the qubits to an initial state $|\Psi^{(0)}\rangle$.

2. At the mth iteration, identify the Hermitian generator $B_m \in \mathbb{P}$ such that the action of the parameterised unitary operator $\exp(-\iota\theta_m B_m)$, $\theta_m \in [-\pi, \pi)$ on the current ansatz $|\Psi^{(m-1)}\rangle$ is <u>likely</u> to produce a new wave-function with the largest drop in energy. This identification is done by computing the gradient, at $\theta = 0$, of the expectation value of the Hamiltonian, i.e.,

$$B_m = \underset{B \in \mathbb{P}}{\arg\max} \left| \frac{\partial}{\partial \theta} \langle \Psi^{(m-1)} | \exp(\iota\theta B) H \exp(-\iota\theta B) | \Psi^{(m-1)} \rangle \big|_{\theta=0} \right|. \tag{3}$$

   Note that the criterion (3) is simply a heuristic, and there is no guarantee that the Hermitian generator B_m selected through this criterion will indeed lead to the parameterised unitary operator whose action on the current ansatz $|\Psi^{(m-1)}\rangle$ results in the largest drop in energy. This point will be the subject of further discussion in Section 2.3.

3. Exit the iterative process if the stopping criterion is met (see below for more explanation). Otherwise, append the resulting parametrised unitary operator to the left of the current ansatz wave-function $|\Psi^{(m-1)}\rangle$, i.e., define

$$|\widetilde{\Psi^{(m)}}\rangle := \exp(-\iota\theta_m B_m) |\Psi^{(m-1)}\rangle = \exp(-\iota\theta_m B_m) \exp(-\iota\theta'_{m-1} B_{m-1}) \ldots \exp(-\iota\theta'_1 B_1) |\Psi^{(0)}\rangle.$$

4. Run a classical VQE routine by optimising all parameters $\theta_m, \theta_{m-1}, \ldots, \theta_1$ in the new ansatz wave-function $|\widetilde{\Psi^{(m)}}\rangle$ so as to minimize the expectation value of the Hamiltonian, i.e., solve the optimisation problem

$$\vec{\theta}^{\text{opt}} := (\theta'_1, \ldots, \theta'_{m-1}, \theta'_m) \tag{4}$$

$$:= \underset{\theta_1, \ldots, \theta_{m-1}, \theta_m}{\arg\min} \left\langle \Psi^{(0)} \Big| \prod_{k=1}^{k=m} \exp(\iota\theta_k B_k) H \prod_{k=m}^{k=1} \exp(-\iota\theta_k B_k) \Big| \Psi^{(0)} \right\rangle,$$

   and define the new ansatz wave-function $|\Psi^{(m)}\rangle$ using the newly optimized parameters $\theta'_1, \ldots, \theta'_m$, i.e., define

$$|\Psi^{(m)}\rangle := \prod_{k=m}^{k=1} \exp(-\iota\theta'_k B_k) |\Psi^{(0)}\rangle.$$

   Let us emphasize that although we also denote the newly optimized parameters at the current m^{th} iteration by $\theta'_1, \ldots \theta'_m$, these optimized values are not necessarily the same as those used to define $|\Psi^{(m-1)}\rangle$ and referenced in Step 3 above.

5. Return to Step 2 with the updated ansatz $|\Psi^{(m)}\rangle$.

Let us remark here that a common choice of stopping criterion is to impose a pre-defined threshold tolerance $\varepsilon > 0$ on the magnitude of the gradients computed in Step 2 above, i.e., exit the ADAPT-VQE algorithm at iteration m if

$$\max_{B \in \mathbb{P}} \left| \frac{\partial}{\partial \theta} \left\langle \Psi^{(m-1)} \right| \exp(\iota \theta B) H \exp(-\iota \theta B) \left| \Psi^{(m-1)} \right\rangle \Big|_{\theta=0} \right| < \varepsilon.$$

An obvious alternative option is to impose a maximal iteration count on the number of ADAPT-VQE steps or a minimal decrease of the expectation value between two iterates

$$\left\langle \Psi^{(0)} \right| \prod_{k=1}^{k=m-1} \exp(\iota \theta_k' B_k) H \prod_{k=m-1}^{k=1} \exp(-\iota \theta_k' B_k) \left| \Psi^{(0)} \right\rangle - \left\langle \Psi^{(0)} \right| \prod_{k=1}^{k=m} \exp(\iota \theta_k' B_k) H \prod_{k=m}^{k=1} \exp(-\iota \theta_k' B_k) \left| \Psi^{(0)} \right\rangle < \varepsilon$$

Next, let us discuss some commonly used operator pools for ADAPT-VQE.

2.2 Operator Pools for ADAPT-VQE

As one may expect by studying the ADAPT-VQE workflow, the success of the algorithm is strongly impacted by the choice of the Hermitian generator pool $\mathbb{P}$. As an extreme example, if all generators in the operator pool commute with the Hamiltonian, then the algorithm will terminate at the first iteration thus resulting in no improvement of the initial guess. The goal of this section is to briefly present some popular operator pools. While a great variety of operator pools have been introduced in the literature[8,31,32], we will limit ourselves to a 'chemically-inspired' pool which is popular for simulating quantum chemical systems[15], a simplified version of this chemically-inspired pool which (empirically) leads to lower quantum gate counts[14], and finally a so-called <u>minimal</u> operator pool which possesses some useful mathematical properties. Before doing so however, let us first clarify some additional details concerning the ADAPT-VQE algorithm.

ADAPT-VQE was primarily developed for quantum chemistry applications, i.e., for application to molecular systems typically described by a second-quantized molecular Hamiltonian mapped to a qubit representation through the Jordan-Wigner transformation. In this setting, an obvious choice for the initial state in the ADAPT-VQE algorithm is the Hartree-Fock wave-function. Indeed, in the standard formalism where each qubit is used to represent a specific spin-orbital, we can write $|0\rangle_p$ and $|1\rangle_p$ to denote states corresponding to an <u>empty</u> and <u>occupied</u> spin-orbital p respectively. With this notation, the reference Hartree-Fock state for a system having n electrons in N spin-orbitals can be expressed as $|\Psi_{\text{HF}}\rangle := |1_0 \ldots 1_{n-1} 0_n \ldots 0_{N-1}\rangle$, which is straightforward to represent on quantum circuits. Note that this is an example of a case where we have access to an initial state that is both simple to represent on quantum architecture and also yields a wave-function having a reasonable overlap with the sought-after ground state wave-function. Of course, for arbitrary Hamiltonians, such an efficient choice for the initial state might not be possible, in which case we might have to rely on random initialisations, for instance.

The Qubit Excitation-based Pool[15]

The first commonly used operator pool is the Qubit excitation-based (QEB) pool which is inspired by the popular coupled cluster method from computational quantum chemistry. The QEB pool consists of so-called single-qubit and double-qubit excitation operators which take the form

$$A_{pq} = \frac{1}{2} \left(X_q Y_p - Y_p X_q \right). \tag{5}$$

and

$$A_{pqrs} = \frac{1}{8} \left(X_r Y_s X_p X_q + Y_r X_s X_p X_q + Y_r Y_s Y_p X_q + Y_r Y_s X_p Y_q - X_r X_s Y_p X_q - X_r X_s X_p Y_q - Y_r X_s Y_p Y_q - X_r Y_s Y_p Y_q \right). \tag{6}$$

Here p, q, r, and s denote qubit indices and X_p and Y_p are the usual one-qubit Pauli gates acting on qubit p. Thus, the single-qubit generator A_{pq} acts between the single qubits p and q while the double-qubit generator A_{pqrs} acts between the qubit pairs (p,q) and (r,s).

Given an N-qubit system, it is easy to see that the QEB operator pool a priori has $\mathscr{O}(N^4)$ elements. In practice however, not all possible excitation operators of the form (5)-(6) are included in the operator pool. Instead, the QEB operator pool is limited to those single-qubit and double-qubit excitation operators which preserve important symmetries in the system such as spin or the number of particles. Additionally, it is readily checked that the parametric exponentiation of a QEB operator is easy to calculate, and the resulting unitary operators have well-known CNOT-optimised circuits (see, e.g., Figure 11 in the appendix).

The Qubit Hardware-efficient Pool[14]

While the Qubit excitation-based pool provides excellent performance in numerical simulations on quantum simulators, the practical implementation of QEB-based ansatz wave-functions on near-term quantum hardware remains challenging. This is primarily due to the fact that the number of CNOT gates required to construct the associated QEB circuits, while significantly smaller than the CNOT counts for classical "fixed-ansatz" approaches, is still far too high. The so-called Qubit hardware-efficient pool[7] addresses this issue by considering instead a pool consisting of modified single and double excitation operators of the form

$$\widetilde{X}_{pq} = \frac{1}{2} X_q Y_p \tag{7}$$

and

$$
\begin{aligned}
\widetilde{X}_{pqrs}^{(1)} &= \frac{1}{8} X_r Y_s X_p X_q, \quad \widetilde{X}_{pqrs}^{(2)} = \frac{1}{8} Y_r X_s X_p X_q, \quad \widetilde{X}_{pqrs}^{(3)} = \frac{1}{8} Y_r Y_s Y_p X_q, \quad \widetilde{X}_{pqrs}^{(4)} = \frac{1}{8} Y_r Y_s X_p Y_q, \\
\widetilde{X}_{pqrs}^{(5)} &= \frac{1}{8} X_r X_s Y_p X_q, \quad \widetilde{X}_{pqrs}^{(6)} = \frac{1}{8} X_r X_s X_p Y_q, \quad \widetilde{X}_{pqrs}^{(7)} = \frac{1}{8} Y_r X_s Y_p Y_q, \quad \widetilde{X}_{pqrs}^{(8)} = \frac{1}{8} X_r Y_s Y_p Y_q,
\end{aligned}
\tag{8}
$$

where p, q, r, s again denote qubit indices and X_p and Y_p are one-qubit Pauli gates acting on qubit p. Note that not all modified double qubit-excitation operators of the form (8) are added to the operator pool since, for instance, the operator $\widetilde{X}_{pqrs}^{(1)} = X_r Y_s X_p X_q$ and $\widetilde{X}_{pqrs}^{(7)} = Y_r X_s Y_p Y_q$ are related by a global rotation[14]. Let us also emphasize here that the qubit hardware-efficient pool is not particle conserving and thus violates an important symmetry of the system.

Numerical experiments involving the qubit hardware-efficient pool demonstrate that the resulting ansatz wave-functions can indeed be represented on quantum architectures using a lower CNOT count than that required for the QEB-pool based ansätze. The tradeoff, however, is that the qubit hardware-efficient pool is several times larger than than QEB-pool, and this produces a significant computational overhead when evaluating operator gradients for use in the operator selection criterion. The next operator pool attempts to resolve this issue.

Minimal Hardware-efficient Pool[14]

It is well-known that when the Hamiltonian under study is real-valued, then the ground state eigenfunction can be expressed as a real linear combination of real basis vectors. For such settings therefore, it is possible to show the existence of a so-called minimal operator pool that allows the transformation of any real-valued wave-function (in particular, the Hartree-Fock reference state) to another real-valued wave-function (in particular, the sought-after ground state eigenfunction). More precisely, given an N-qubit system, we may define the operator pool

$$\mathbb{P} = \{Y_p\}_{p=0}^{N-2} \cup \{Z_p Y_{p+1}\}_{p=0}^{N-2},$$

where Y_p and Z_p are the usual one-qubit Pauli gates acting on qubit p. Then, for any two real-valued wave-functions $|\Phi\rangle$ and $|\Psi\rangle$, there exist $\theta_1, \ldots, \theta_M \in [-\pi, \pi)$ and Hermitian generators $B_1, \ldots, B_M \in \mathbb{P}$ such that

$$|\Phi\rangle = \prod_{k=M}^{1} \exp(-\iota \theta_k B_k) |\Psi\rangle.$$

In other words, for a precise choice of parameters and generators in the pool $\mathbb{P}$, we can construct the sought-after ground-state eigenfunction by applying the parametrised, exponentiated generators to the Hartree-Fock reference state.

The key advantage of this so-called minimal hardware-efficient pool $\mathbb{P}$ is that it consists of only $2N - 2$ elements, and the operators in this pool can be parametrically exponentiated using very simple circuits that require a minimal number of CNOT gates.

2.3 Resource Saving Enhancements and the GGA-VQE Algorithm

As described in detail in Section 2.1, a core step in the ADAPT-VQE algorithm is the selection of an optimal Hermitian generator B from the operator pool whose addition to the current ansatz can produce a new ansatz wave-function with the largest drop in energy. Current implementations of ADAPT-VQE make this choice through a heuristic criterion based on evaluating certain gradients of the expectation value of the Hamiltonian. More precisely, for a given pool of operators $\mathbb{P}$, at the mth iteration, one computes (c.f., Equation (3))

$$B_m = \underset{B \in \mathbb{P}}{\operatorname{argmax}} \left| \frac{\partial}{\partial \theta} \langle \Psi^{(m-1)} | \exp(\iota \theta B) H \exp(-\iota \theta B) | \Psi^{(m-1)} \rangle \big|_{\theta=0} \right|, \tag{9}$$

where $|\Psi^{(m-1)}\rangle$ denotes the ansatz wave-function at the $m-1$ iteration. Of course, Equation (9) is still a heuristic, and it is conceivable that there exists another Hermitian generator $\widetilde{B}_m$ such that

$$\min_{\theta \in [-\pi,\pi)} \langle \Psi^{(m-1)}|\exp(\iota\theta\widetilde{B}_m)H\exp(-\iota\theta\widetilde{B}_m)|\Psi^{(m-1)}\rangle < \min_{\theta \in [-\pi,\pi)} \langle \Psi^{(m-1)}|\exp(\iota\theta B_m)H\exp(-\iota\theta B_m)|\Psi^{(m-1)}\rangle$$

while

$$\left| \frac{\partial}{\partial\theta} \langle \Psi^{(m-1)}|\exp(\iota\theta\widetilde{B}_m)H\exp(-\iota\theta\widetilde{B}_m)|\Psi^{(m-1)}\rangle \big|_{\theta=0} \right| < \left| \frac{\partial}{\partial\theta} \langle \Psi^{(m-1)}|\exp(\iota\theta B_m)H\exp(-\iota\theta B_m)|\Psi^{(m-1)}\rangle \big|_{\theta=0} \right|.$$

A representative example of this situation is displayed in Figure 1.
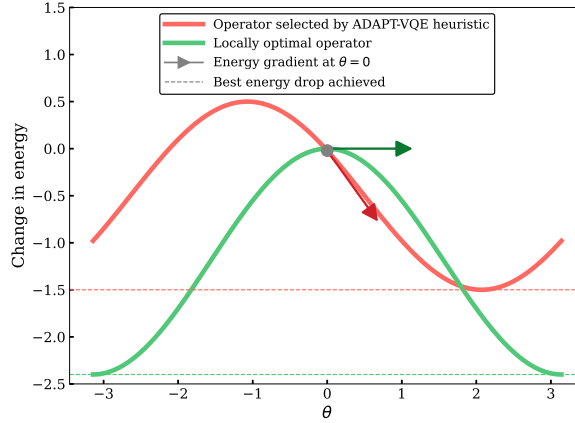

Figure 1. Illustration of a situation where the ADAPT-VQE selection criterion does not pick the optimal operator leading to the largest energy drop.

Given that at each iteration of ADAPT-VQE, we face the task of optimising a multi-dimensional objective function which is very noisy due to the quality of the current quantum hardware, the selection of the wrong operator to append to the current ansatz wave-function can be a costly mistake[33]. In this section, we introduce an energy sorting algorithm that allows the exact selection of the locally, optimal Hermitian generator from any of the three operator pools that we have introduced in Section 2.2. In other words, we show that it is possible, using a few measurements on the quantum device, to exactly solve the optimisation problem

$$B_m = \underset{B \in \mathbb{P}}{\mathrm{argmin}} \min_{\theta \in [-\pi,\pi)} \mathscr{L}(B,\theta,|\Psi^{(m-1)}\rangle) := \underset{B \in \mathbb{P}}{\mathrm{argmin}} \min_{\theta \in [-\pi,\pi)} \langle \Psi^{(m-1)}|\exp(\iota\theta B)H\exp(-\iota\theta B)|\Psi^{(m-1)}\rangle, \tag{10}$$

where, for notational convenience, we have introduced the objective function $\mathscr{L}(B,\theta,|\Psi^{(m-1)}\rangle)$, which, in some literature, is also referred to as a landscape function.

The main idea of our energy sorting algorithm is to take advantage of a result that is largely known in the VQE literature on 'fixed-ansatz' methods, namely that the expectation value of a Hermitian operator with respect to an ansatz wave-function $|\Psi(\theta)\rangle$ parametrised by a single quantum gate can be expressed in terms of elementary trigonometric functions of θ[22,23]. While this result is typically expressed for ansatz wave-function parametrised by rotation gates, the core arguments can be extended to wave-functions parametrised by Hermitian generators belonging to any of the three operator pools that we have introduced in Section 2.2. Indeed, taking advantage of the precise functional form of the Hermitian generators introduced in Section 2.2, a simple calculation shows that

1. For any generator B in the Qubit-Excitation-Based (QEB) operator pool, it holds that $B^3 = B$ with I denoting the identity matrix (see Appendix 5.2).

2. For any generator B in the Qubit hardware-efficient and minimal hardware-efficient pools, it holds that $B^2 = I$.

The above simple relations now imply (see also Appendix 5.2) that for any generator B in the Qubit-Excitation-Based (QEB) operator pool and any $\theta \in [-\pi, \pi)$, it holds that

$$\exp(-\iota\theta B) = I + (\cos(\theta) - 1)B^2 - \iota\sin(\theta)B, \tag{11}$$

and for any generator B in the Qubit hardware-efficient and minimal hardware-efficient pools and any $\theta \in [-\pi, \pi)$, it holds that

$$\exp(-\iota\theta B) = \cos(\theta)I - \iota\sin(\theta)B. \tag{12}$$

Using Equations (11) and (12), it is easy to establish that for any wave-function $|\phi\rangle$ and any $\theta \in [-\pi, \pi)$ the objective function $\mathscr{L}(B, \theta, |\phi\rangle)$ has the analytical form

$$\mathscr{L}(B, \theta, |\phi\rangle) = \begin{cases} \langle\phi|H|\phi\rangle + \big(\cos(\theta) - 1\big)\big(\langle\phi|\{H, B^2\}|\phi\rangle - 2\langle\phi|BHB|\phi\rangle\big) & \text{if } B^3 = B, \\ + \big(1 - \cos(\theta)\big)^2\big(\langle\phi|B^2HB^2|\phi\rangle - \langle\phi|BHB|\phi\rangle\big) \\ + \sin(\theta)(\cos(\theta) - 1)\langle\phi|\iota B[H, B]B|\phi\rangle \\ + \sin(\theta)\langle\phi|\iota[B, H]|\phi\rangle \\ \\ \cos^2(\theta)\langle\phi|H|\phi\rangle + \dfrac{\sin(2\theta)}{2}\langle\phi|\iota[B, H]|\phi\rangle & \text{if } B^2 = I. \\ + \sin^2(\theta)\langle\phi|BHB|\phi\rangle, \end{cases} \tag{13}$$

where $\{\cdot, \cdot\}$ and $[\cdot, \cdot]$ denote the anti-commutator and commutator respectively. A demonstration of this result can be found in Appendix 5.2.

Equation (13) implies that for any Hermitian generator B from our operator pools and any arbitrary wave-function $|\phi\rangle$, the objective function $\mathscr{L}(B, \theta, |\phi\rangle)$ can be expressed in terms of elementary trigonometric functions of θ. An important consequence of this expression is that, if we now evaluate the objective function $\mathscr{L}(B, \theta, |\phi\rangle)$ at certain well-chosen angles, we can obtain a linear system of equations for the unknown operator expectation values. More precisely,

For the QEB pool ($B^3 = B$):

Since $\langle\phi|H|\phi\rangle$ can be measured directly on the quantum device, four measurements of $\mathscr{L}(B, \theta, |\phi\rangle)$ at well-chosen $\theta = \theta^{(1)}, \theta^{(2)}, \theta^{(3)}, \theta^{(4)}$ yields a linear system for the four unknowns $\langle\phi|\{H, B^2\}|\phi\rangle - 2\langle\phi|BHB|\phi\rangle$, $\langle\phi|B^2HB^2|\phi\rangle - \langle\phi|BHB|\phi\rangle$, $\langle\phi|iB[H, B]B|\phi\rangle$ and $\langle\phi|\iota[B, H]|\phi\rangle$.

For the hardware efficient pools ($B^2 = I$):

Since $\langle\phi|H|\phi\rangle$ can be measured directly on the quantum device, two measurements of $\mathscr{L}(B, \theta, |\phi\rangle)$ at well-chosen $\theta = \theta^{(1)}, \theta^{(2)}$ yields a linear system for the two unknowns $\langle\phi|i[B, H]|\phi\rangle$ and $\langle\phi|BHB|\phi\rangle$.

In other words, using a minimal number of measurements and by solving a very small linear system, we can compute all terms involving $B, H,$ and $|\phi\rangle$ in the expression (13) for the objective function $\mathscr{L}(B, \theta, |\phi\rangle)$. Since the dependency of this function on θ is through elementary trigonometric functions, we can thus express the objective function $\mathscr{L}(B, \theta, |\phi\rangle)$ analytically for any generator B, any angle θ, and any wave-function $|\phi\rangle$. This allows us to solve the optimisation problem (10) up to arbitrary precision for any Hermitian generator B from our operator pool, and thereby obtain the <u>locally, optimal</u> generator that should be added to the current ansatz wave-function $|\Psi^{(m-1)}\rangle$. Note that if we assume an operator pool of size M, a total of $4M + 1$ measurements will be required to screen all Hermitian generators from the qubit-excitation based pool while a total of $2M + 1$ measurements will be required to screen all Hermitian generators from the two hardware efficient pools.

Two important remarks are now in order. First, as mentioned previously, analytical expressions such as (13) for landscape functions $\mathscr{L}(B, \theta, |\phi\rangle)$ of the form (10) are present in the existing literature on 'fixed-ansatz' methods[22–27], albeit not- to the best of our knowledge- for the specific operator pools that we have considered in the current study. However, these landscape functions are almost exclusively used to perform an operator-by-operator, analytical optimisation of a structurally fixed, parameterised quantum circuit. Indeed, the only numerical method we are aware of that extends the applicability of analytical landscapes functions beyond simple iterative optimisation of a fixed ansatz circuit is the Rotoselect algorithm[23], and

Rotoselect still assumes a fixed structure for the parametrised quantum circuit, in which only the choice of $X, Y,$ or Z rotation gate (and not the 'location') can be varied according to the associated landscape function.

Second, let us point out that the landscape function (10) assumes the addition of a single operator to the current ansatz wave-function at each iteration of the adaptive algorithm. If the pool of potential unitary operators is commutative, then the specific order in which operators are chosen is unimportant, and it is therefore sufficient to consider a sequential application of the representation (13) of $\mathscr{L}(B, \theta, |\phi\rangle)$ to determine, at each iteration, the optimal operator to append to the current ansatz. On the other hand, if the Hermitian generators belonging to the operator pool do not commute (which is often the case), then the ordering of the operators is important, and it is potentially useful to consider landscape functions based on the simultaneous addition of $d > 1$ operators to the current ansatz wave-function at each iteration. We now briefly discuss this generalisation.

Multi-dimensional analytical landscape functions

Let us consider an adaptive procedure in which d unitary operators, constructed using d Hermitian generators from a given operator pool $\mathbb{P}$ are to be appended to the current ansatz wave-function $|\Psi^{(m-1)}\rangle$ at iteration m. We are now interested in determining the ordered d-tuple of Hermitian generators $(B_{m_d}, \ldots, B_{m_1})$ such that

$$
\begin{aligned}
(B_{m_d}, \ldots, B_{m_1}) &= \underset{B_d, \ldots, B_1 \in \mathbb{P}}{\operatorname{argmin}} \quad \underset{\substack{\theta_d, \ldots \theta_1 \\ \in [-\pi, \pi)}}{\min} \mathscr{L}\Big((B_d, \theta_d), \ldots, (B_1, \theta_1), |\Psi^{(m-1)}\rangle\Big) \\
&:= \underset{B_d, \ldots, B_1 \in \mathbb{P}}{\operatorname{argmin}} \quad \underset{\substack{\theta_d, \ldots \theta_1 \\ \in [-\pi, \pi)}}{\min} \langle \Psi^{(m-1)}| \exp(\iota \theta_1 B_1) \ldots \exp(-\iota \theta_d B_d) H \exp(\iota \theta_d B_d) \ldots \exp(-\iota \theta_1 B_1)|\Psi^{(m-1)}\rangle.
\end{aligned} \tag{14}
$$

In order to obtain an analytical representation of the d-dimensional objective function $\mathscr{L}\Big((B_d, \theta_d), \ldots, (B_1, \theta_1), |\Psi^{(m-1)}\rangle\Big)$, the fundamental idea is to appeal once again to Equations (11) and (12) and expand each exponential in θ_j, $j \in \{1, \ldots, d\}$ as a sum of a sine and cosine function of θ_j. This expansion allows us to conclude that the d-dimensional objective function $\mathscr{L}\Big((B_d, \theta_d), \ldots, (B_1, \theta_1), |\Psi^{(m-1)}\rangle\Big)$ can be written in the general form

$$
\mathscr{L}\Big((B_d, \theta_d), \ldots, (B_1, \theta_1), |\Psi^{(m-1)}\rangle\Big) =
$$

$$
\begin{cases}
\left\langle \Psi^{(m-1)}\Big| \prod_{j=1}^{j=d}\Big(I + (\cos(\theta_j) - 1)B_j^2 + \iota \sin(\theta_j)B_j\Big) H \prod_{j=d}^{j=1}\Big(I + (\cos(\theta_j) - 1)B_j^2 - \iota \sin(\theta_j)B_j\Big) \Big| \Psi^{(m-1)}\right\rangle & \text{if } B^3 = B, \\[2em]
\left\langle \Psi^{(m-1)}\Big| \prod_{j=1}^{j=d}(\cos(\theta_j)I + \iota \sin(\theta_j)B_j) H \prod_{j=d}^{j=1}(\cos(\theta_j)I - \iota \sin(\theta_j)B_j) \Big| \Psi^{(m-1)}\right\rangle & \text{if } B^2 = I.
\end{cases}
$$

In other words, the d-dimensional objective function $\mathscr{L}\Big((B_d, \theta_d), \ldots, (B_1, \theta_1), |\Psi^{(m-1)}\rangle\Big)$ can be written as a polynomial of the variables $\{1, \cos(\theta_j), \sin(\theta_j)\colon j \in \{1, \ldots, d\}\}$ with the exact structure of the polynomial depending on the properties of the operator pool $\mathbb{P}$. As a representative example, the landscape function for hardware efficient pools with $d = 2$, after some simplifications, is of the form:

$$
\begin{aligned}
\mathscr{L}\big((B_2, \theta_2), (B_1, \theta_1), |\phi\rangle\big) = &\langle \phi|H|\phi\rangle \\
&+ \frac{\cos(2\theta_1)}{2}\Big(\langle \phi|H - B_1 H B_1|\phi\rangle + \frac{\cos(2\theta_2)}{2}\langle \phi|H - B_1 H B_1 - B_2 H B_2 + B_2 B_1 H B_1 B_2|\phi\rangle \\
&\qquad + \frac{\sin(2\theta_2)}{2}\langle \phi|i[B_2, H - B_1 H B_1]|\phi\rangle\Big) \\
&+ \frac{\sin(2\theta_1)}{2}\Big(\langle \phi|i[B_1, H]|\phi\rangle + \frac{\cos(2\theta_2)}{2}\langle \phi|i[B_1, H] - i B_2[B_1, H]B_2|\phi\rangle \\
&\qquad - \frac{\sin(2\theta_2)}{2}\langle \phi|[B_2, [B_1, H]]|\phi\rangle\Big)
\end{aligned} \tag{15}
$$

Consequently, a total of 7 measurements on a quantum device are required to deduce an analytical expression for the two-dimensional landscape function $\mathscr{L}\big((B_2, \theta_2), (B_1, \theta_1), |\phi\rangle\big)$ for any Hermitian generators B_1, B_2 belonging to either of the two

hardware efficient operator pools. Since the selection of the best two operator to append to the current ansatz wave-function requires comparing all pairs of Hermitian generators, we conclude that for an operator pool of size M, at most $6M^2 + 1$ measurements are required to determine the locally optimal pair of unitary operators that should be appended to the current ansatz wave-function at each iteration in order to achieve the largest drop in expectation value of the underlying Hermitian operator.

In the case of a general d-dimensional objective function $\mathscr{L}\left((B_d, \theta_d), \ldots, (B_1, \theta_1), |\Psi^{(m-1)}\rangle\right)$, similar arguments yield that

- for the Qubit-Excitation-Based (QEB) operator pool of size M, we require $\mathscr{O}(5^d M^d)$ measurements to determine the locally optimal d-tuple of unitary operators that should be appended to the current ansatz wave-function at each iteration in order to achieve the largest drop in expectation value of the underlying Hermitian operator;

- for the Qubit hardware-efficient and minimal hardware-efficient pools of size M, we require $\mathscr{O}(3^d M^d)$ measurements to determine the locally optimal d-tuple of unitary operators that should be appended to the current ansatz wave-function at each iteration in order to achieve the largest drop in expectation value of the underlying Hermitian operator.

While this procedure can become computationally intractable for moderately large d, various simplifications are possible that can lead to more tractable gradient-free adaptive algorithms involving multi-operator selection and optimisation, and it is likely that such methods offer an advantage when formulating a greedy gradient-free adaptive VQE for a complex Hamiltonian using a non-commutative operator pool. As representative examples, given an ansatz wave-function $|\Psi^{(m-1)}\rangle$, at iteration m:

1. We may use the energy sorting algorithm based on <u>one-dimensional landscape functions</u> to classify, in descending order of importance, the best d Hermitian generators $(B_d, \ldots, B_1)$ whose addition to the current ansatz wave-function can result in the largest drops in the expectation value of the underlying Hermitian operator, i.e.,

$$\min_{\theta_d \in [-\pi, \pi)} \langle \Psi^{(m-1)} | \exp(\iota \theta B_d) H \exp(-\iota \theta B_d) |\Psi^{(m-1)}\rangle \geq \ldots \geq \min_{\theta_1 \in [-\pi, \pi)} \langle \Psi^{(m-1)} | \exp(\iota \theta B_1) H \exp(-\iota \theta B_1) |\Psi^{(m-1)}\rangle.$$

We can then switch to the analytical expression of d-dimensional landscape function $\mathscr{L}\left((B_d, \theta_d), \ldots, (B_1, \theta_1), |\Psi^{(m-1)}\rangle\right)$ in order to compute the optimal parameters $(\theta_d^*, \ldots, \theta_1^*)$ such that

$$
\begin{aligned}
(\theta_d^*, \ldots, \theta_1^*) =& \operatorname*{argmin}_{\substack{\theta_d, \ldots \theta_1 \\ \in [-\pi, \pi)}} \mathscr{L}\left((B_d, \theta_d), \ldots, (B_1, \theta_1), |\Psi^{(m-1)}\rangle\right) \\
=& \operatorname*{argmin}_{\substack{\theta_d, \ldots \theta_1 \\ \in [-\pi, \pi)}} \langle \Psi^{(m-1)} | \exp(\iota \theta_1 B_1) \ldots \exp(-\iota \theta_d B_d) H \exp(\iota \theta_d B_d) \ldots \exp(-\iota \theta_1 B_1) |\Psi^{(m-1)}\rangle
\end{aligned}
$$

In other words, we may use one-dimensional landscape functions to <u>identity</u> the best d Hermitian generators to add to the current ansatz wave-function and we may employ the d-dimensional landscape functions to perform the <u>analytical optimisation</u>. The number of quantum measurements required by this procedure scales as $3^d M$ (resp. $5^d M$) for a hardware efficient (resp. QEB) operator pool of size M.

2. Taking the newly obtained ansatz wave-function $|\Psi^{(m)}\rangle$ after iteration m as structurally fixed, we may perform Rotoselect-style[23] backwards and forwards optimisation sweeps over all parameterised unitary operators $\exp(\iota \theta_1^* B_d), \ldots \exp(\iota \theta_d^* B_1)$. In particular, thanks to the analytical expression (15) for the d-dimensional landscape function, each iteration in these optimisation sweeps can involve d parametrised unitary operators simultaneously. The number of quantum measurements required by a single sweep utilising d-dimensional landscape functions scales as $3^d M$ (resp. $5^d M$) for a hardware efficient (resp. QEB) operator pool of size M.

Such considerations will be the subject of numerical investigations in forthcoming works. Since the goal of the current study is the successful implementation of adaptive variational quantum algorithms on current generation NISQ devices, we will, for the moment, limit ourselves to the computationally cheap case of one-dimensional landscape functions, which suffices for the relatively simple Hamiltonians that we consider in the sequel. In this one-dimensional setting, the energy sorting algorithm that we have introduced is the basis of the following greedy gradient-free adaptive VQE, which we dub GGA-VQE.

The Greedy Gradient-free Adaptive Variational Quantum Eigensolver (GGA-VQE)

Given the qubit representation of an input Hamiltonian H, a pool of admissible Hermitian generators $\mathbb{P}$, and a stopping criterion:

1. Boot the qubits to an initial state $|\Psi^{(0)}\rangle$.

2. At the mth iteration, use the energy sorting algorithm detailed above to identify the Hermitian generator $B_m \in \mathbb{P}$ that solves the optimisation problem (10), i.e.,

$$B_m = \underset{B \in \mathbb{P}}{\operatorname{argmin}} \min_{\theta \in [-\pi,\pi)} \langle \Psi^{(m-1)} | \exp(\iota\theta\widetilde{B}) H \exp(-\iota\theta\widetilde{B}) | \Psi^{(m-1)} \rangle. \tag{16}$$

3. Exit the iterative process if the stopping criterion is met. Otherwise, append the resulting parametrised unitary operator to the left of the current ansatz wave-function $|\Psi^{(m-1)}\rangle$, i.e., define the new ansatz wave-function

$$|\Psi^{(m)}\rangle := \exp(-\iota\theta'_m B_m) |\Psi^{(m-1)}\rangle = \exp(-\iota\theta'_m B_m)\exp(-\iota\theta'_{m-1}B_{m-1})\ldots\exp(-\iota\theta'_1 B_1)|\Psi^{(0)}\rangle,$$

where the angle θ'_m is obtained in the process of solving the optimisation problem (16).

4. Return to Step 2 with the updated ansatz $|\Psi^{(m)}\rangle$.

It is important to emphasise that, in contrast to the classical ADAPT-VQE procedure, the GGA-VQE algorithm described above does not involve a global optimisation of all parameters in the current ansatz at each iteration. Instead, at each iteration, we use the energy sorting algorithm to identify the locally optimal Hermitian generator B as well as the optimal angle θ which should be used to construct the new ansatz wave-function– a process which involves the optimisation of one-dimensional, elementary trigonometric functions. In particular, in contrast to the classical ADAPT-VQE, we avoid entirely the need to optimise a multi-dimensional and extremely noisy cost function involving the system Hamiltonian. The resulting huge savings in quantum resources suggest that the GGA-VQE is particularly suited for implementation on near-term quantum devices. Let us also note that the d-dimensional landscape functions introduced in Section 2.3 above can be used to develop natural generalisations of the GGA-VQE algorithm, which we dub GGA-VQE(d), that are likely to be particularly suited for the ground state preparation of strongly correlated systems.

It can now readily be seen that the main computational bottle-neck in the GGA-VQE algorithm is the energy sorting procedure which requires $\mathcal{O}(M)$ measurements of the system Hamiltonian for an operator pool of size M. It is therefore natural to ask if the number of measurements required to perform the energy sorting can be further reduced, at least for certain types of Hamiltonians. In the next section, we answer this question affirmatively by showing that for a certain class of Ising Hamiltonians, it is possible to perform energy sorting using a number of measurements independent of both the size of the operator pool and the number of qubits.

2.4 GGA-VQE for a Transverse-field Ising Model

While the ADAPT-VQE algorithm is predominantly applied to compute the ground state energies of molecular systems, there is, in principle, no restriction in applying the method to obtain ground state energies for more general Hamiltonians[32]. The goal of this section is to describe in detail, the application of the GGA-VQE algorithm that we have introduced in Section 2.3 to an open boundary transverse-field Ising Hamiltonian[34]. Ising Hamiltonians of this type are of great importance in condensed-matter physics since they are among the simplest models capable of representing different phases of matter, depending on the value of various systems parameters[35]. As the Ising Hamiltonian is well-known theoretically, it also presents a good first test for computational experiments prior to tackling more complex molecular Hamiltonians.

Given an N-qubit register, we consider the transverse-field Ising Hamiltonian given by

$$H = h\sum_{p=0}^{N-1} X_p + J\sum_{p=0}^{N-2} Z_p Z_{p+1}, \tag{17}$$

where X_p and Z_p denote the usual Y and Z Pauli matrices acting on qubit p, and $h, J > 0$ are system parameters. The physical constant h models the intensity of a magnetic field directed along the x-axis, whereas the constant J models the strength of the nearest-neighbour interactions. If $J < 0$, neighbouring spins tend to align, and the opposite is true if $J > 0$. Note that in this model, each qubit represents a spin-state.

Since the Ising Hamiltonian is real valued, a natural choice of operator pool is the minimal hardware-efficient pool introduced in Section 2.2, which is given by

$$\mathbb{P} = \{Y_p\}_{p=0}^{N-1} \cup \{Z_p Y_{p+1}\}_{p=0}^{N-1}. \tag{18}$$

Let us now recall from Section 2.3 that implementing the GGA-VQE algorithm requires us to solve, at each iteration, a minimisation problem so as to identify the optimal Hermitian generator which should be used to construct the new ansatz wave-function. The objective function associated with this minimisation problem (see Equation (10)) is given by

$$\mathscr{L}(B, \theta, |\Psi^{(m-1)}\rangle) = \langle \Psi^{(m-1)} | \exp(\imath\theta B) H \exp(-\imath\theta B) | \Psi^{(m-1)} \rangle,$$

where $B \in \mathbb{P}$ is any Hermitian generator from our operator pool, the parameter $\theta \in [-\pi, \pi)$, and $|\Psi^{(m-1)}\rangle$ denotes the previous ansatz wave-function.

It can now be shown (see the Appendix for a detailed demonstration) that for the Ising Hamiltonian defined through Equation (17) and the minimal hardware-efficient pool $\mathbb{P}$ given by (18), the objective function $\mathscr{L}(B, \theta, |\Psi^{(m-1)}\rangle)$ has the following simple structure:

$$\mathscr{L}(B, \theta, |\Psi^{(m-1)}\rangle) = \begin{cases} \left\langle \Psi^{(m-1)} \middle| H \middle| \Psi^{(m-1)} \right\rangle & \text{if } B = Y_p, \\ + \sin(2\theta) \left\langle \Psi^{(m-1)} \middle| hZ_p - J(X_p Z_{p+1} + Z_{p-1} X_p \delta_{p>0}) \middle| \Psi^{(m-1)} \right\rangle \\ -2\sin^2(\theta) \left\langle \Psi^{(m-1)} \middle| hX_p + J(Z_p Z_{p+1} + Z_{p-1} Z_p \delta_{p>0}) \middle| \Psi^{(m-1)} \right\rangle. \\[2ex] \left\langle \Psi^{(m-1)} \middle| H \middle| \Psi^{(m-1)} \right\rangle & \text{if } B = Z_p Y_{p+1}. \\ + \sin(2\theta) \left\langle \Psi^{(m-1)} \middle| h(Z_p Z_{p+1} - Y_p Y_{p+1}) \middle| \Psi^{(m-1)} \right\rangle \\ - \sin(2\theta) \left\langle \Psi^{(m-1)} \middle| J(X_{p+1} + Z_p X_{p+1} Z_{p+2} \delta_{p+2<n}) \middle| \Psi^{(m-1)} \right\rangle \\ -2\sin^2(\theta) \left\langle \Psi^{(m-1)} \middle| hX_p + hX_{p+1} + JZ_p Z_{p+1} + JZ_{p+1} Z_{p+2} \delta_{p+2<n} \middle| \Psi^{(m-1)} \right\rangle. \end{cases} \quad (19)$$

A close study of the right-hand side of Equation (19) now indicates that many terms involving the expectation values of the Pauli matrices can be measured directly and simultaneously on the quantum device without the need to run over all possible Hermitian generators in $\mathbb{P}$.

- The terms containing only tensor products of Z operators can readily be measured in the computational basis.

- The terms containing only tensor products of X (resp. Y) operators can be measured by applying a Hadamard (resp. $S^\dagger \equiv \text{diag}(1, -\imath)$ and a Hadamard) gate on each qubit.

- The remaining terms are of the form $X_p Z_{p+1}$ or $Z_{p-1} X_p Z_{p+1}$. Terms of this form can be measured by applying a Hadamard gate on qubit p. The terms corresponding to p even commute and can therefore be measured simultaneously. The same holds true for the p odd terms which can thus also be measured simultaneously.

Consequently, it is possible, at each iteration of the Greedy-ADAPT-VQE algorithm to construct exactly five quantum circuits whose measurements allow us to recover an analytical expression for all objective functions $\mathscr{L}(B, \theta, |\Psi^{(m-1)}\rangle)$, $B \in \mathbb{P}$ in terms of elementary trigonometric functions of θ. We have thus achieved a radical reduction in the number of required measurements from $4N - 3$ (for the minimal hardware-efficient pool of size $2N - 2$) to <u>five</u>.

We end this section by noting that a simple choice of initial state for the energy-sorting frozen core ADAPT-VQE procedure is given by the ground-state of the non-interacting Hamiltonian $\sum_{p=0}^{N-1} X_p$, i.e.,

$$|\Psi^{(0)}\rangle = |-\rangle^{\otimes n}$$

Assuming that the system parameters satisfy $|h| > |J|$, it is not unreasonable to expect the ground state of the true <u>interacting</u> Hamiltonian to be a perturbation of $|\Psi^{(0)}\rangle$.

2.5 Overlap-GGA-VQE for Molecular Systems

Unfortunately, despite the various resource-saving enhancements that we have introduced, the application of the GGA-VQE algorithm to molecular systems of chemical interest seems to be out of reach on current quantum hardware. Indeed, in contrast to the Ising model introduced in Section 2.4 for which the expectation value of the Hamiltonian can be be computed through

measurements of just two quantum circuits, computing the expectation value of an *N*-spin orbital molecular Hamiltonian typically requires $\mathcal{O}(N^4)$ measurements– a process which introduces an overwhelming amount of hardware and measurement noise. Fortunately, we have at our disposal an alternative and considerably simpler adaptive algorithm that can be used to explore the limits of the current quantum hardware for the simulation of molecular systems. This is the so-called Overlap-ADAPT-VQE algorithm introduced in[18].

Overlap-ADAPT-VQE is a hybrid quantum/classical algorithm in the spirit of ADAPT-VQE which aims to construct compact approximations of target wave-functions through an iterative procedure. As discussed in an earlier contribution[18], Overlap-ADAPT-VQE seeks to improve the construction of adaptive ansatz wave-function for a VQE procedure in the following two ways:

- First, it can be used as a compression strategy to generate more compact ansatz wave-functions which can be represented using shallower quantum circuits and fewer CNOT gates. More precisely, given a target wave-function represented on a quantum device, the Overlap-ADAPT-VQE can generate a compact approximation of this target wave-function. This compact approximation can then be used as an initialisation for a subsequent second adaptive VQE procedure with the aim of further improving the quality of the ansatz whilst reducing the overall quantum circuit-depth.

- Second, Overlap-ADAPT-VQE can be used to generate a high-quality initialisation for a subsequent adaptive VQE algorithm on a quantum device. More precisely, by taking a moderately accurate, classically computed wave-function as the target, Overlap-ADAPT-VQE can produce a high-fidelity, compact approximation of this classical wave-function on a quantum device. This compact approximation can then be used as the initialisation for a subsequent adaptive VQE procedure which helps alleviate the issue of initial barren plateaus.

In contrast to ADAPT, the Overlap-ADAPT procedure does not require the measurement of the expectation value of the Hamiltonian. Instead, at each iteration of Overlap-ADAPT, we measure the <u>overlap</u> between the current ansatz wave-function and the target wave-function to be approximated– a measurement that is much simpler to achieve. To be more precise, the general workflow of the Overlap-ADAPT-VQE algorithm is as follows.

Given a target wave-function $|\Psi_{\text{ref}}\rangle$, a pool of admissible Hermitian generators $\mathbb{P}$, and a maximal operator count *p*:

1. Boot the qubits to an initial state $|\Psi^{(0)}\rangle$.

2. At the mth iteration, identify the Hermitian generator $B_m \in \mathbb{P}$ such that the action of the parameterised unitary operator $\exp(\iota\theta_m B_m)$, $\theta_m \in [-\pi, \pi)$ on the current ansatz $|\Psi^{(m-1)}\rangle$ is <u>likely</u> to produce a new wave-function having the largest overlap with the target wave-function. This identification is done by maximising a specific gradient involving the current ansatz wave-function at $\theta_m = 0$, i.e.,

$$B_m = \arg\max_{B \in \mathbb{P}} \left| \frac{\partial}{\partial\theta} \left| \left\langle \exp(\iota\theta B)\Psi^{(m-1)} \middle| \Psi_{\text{ref}} \right\rangle \right|^2 \bigg|_{\theta=0} \right|. \tag{20}$$

Note that, as in the classical ADAPT-VQE procedure, the criterion (20) is a heuristic, and there is no guarantee that the Hermitian generator B_m selected through this criterion will indeed lead to the parametrised unitary operator whose action on the current ansatz $|\Psi^{(m-1)}\rangle$ results in the greatest increase in overlap.

3. Append the resulting parametrised unitary operator to the left of the current ansatz wave-function $|\Psi^{(m-1)}\rangle$, i.e., define

$$\begin{aligned} |\widetilde{\Psi^{(m)}}\rangle &:= \exp(-\iota\theta_m B_m) |\Psi^{(m-1)}\rangle \\ &= \exp(-\iota\theta_m B_m)\exp(-\iota\theta'_{m-1}B_{m-1})\ldots\exp(-\iota\theta'_1 B_1) |\Psi^{(0)}\rangle. \end{aligned}$$

4. Run a classical VQE routine by optimising all parameters $\theta_m, \theta_{m-1}, \ldots, \theta_1$ in the new ansatz wave-function $|\widetilde{\Psi^{(m)}}\rangle$ so as to maximise its overlap with the target wave-function i.e., solve the optimisation problem

$$\begin{aligned} \vec{\theta}^{\text{opt}} &:= (\theta'_1, \ldots, \theta'_{m-1}, \theta'_m) \\ &:= \underset{\theta_1,\ldots,\theta_{m-1},\theta_m}{\arg\max} \left| \left\langle \prod_{k=m}^{k=1} \exp(-\iota\theta_k B_k)\Psi^{(0)} \middle| \Psi_{\text{ref}} \right\rangle \right|^2, \end{aligned} \tag{21}$$

and define the new ansatz wave-function $|\Psi^{(m)}\rangle$ using the newly optimised parameters $\theta'_1, \ldots, \theta'_m$, i.e., define

$$|\Psi^{(m)}\rangle := \prod_{k=m}^{k=1} \exp(-\imath\theta'_k B_k)|\Psi^{(0)}\rangle\,.$$

Let us emphasise that although we also denote the newly optimised parameters at the current m^{th} iteration by $\theta'_1, \ldots \theta'_m$, these optimised values are not necessarily the same as those used to define $|\Psi^{(m-1)}\rangle$ and referenced in Step 3 above.

5. If the total number of operators in the updated ansatz is equal to p, exit the iterative process. Otherwise go to Step 2 with the updated ansatz $|\Psi^{(m)}\rangle$.

It is not difficult to see that the Overlap-ADAPT-VQE procedure can also be viewed as finding the <u>maximizer</u> of a Hamiltonian H given by $H = |\Psi\rangle_{\text{ref}}\langle\Psi|_{\text{ref}}$. Thus, the formalism developed in Sections 2.1-2.3 can readily be adapted to fit the framework of Overlap-ADAPT-VQE, and in particular, we can define an Overlap-GGA-VQE algorithm. In order to be able to take advantage of the energy sorting algorithm in this setting, we will, additionally, describe how to compute the expectation of this type of Hamiltonian, i.e., how to compute the overlap between two arbitrary states.

The Compute-Uncompute Method

One method to compute the overlap between two states represented on N-qubit quantum registers is to use the so-called compute-uncompute method. Indeed, assume we have knowledge of two quantum circuits U_Ψ and U_Φ such that $U_\Psi|0\rangle = |\Psi\rangle$ and $U_\Phi|0\rangle = |\Phi\rangle$, where $|0\rangle$ denotes the initial (usually Hartree-Fock) state. Then, the overlap $|\langle\Phi|\Psi\rangle|^2$ can be computed as the expectation value of the projector on the zero state $|0\rangle\langle0| = \left(\frac{I+Z}{2}\right)^{\otimes n}$ with respect to the state $U_\Phi^\dagger|\Psi\rangle$. Indeed, we have

$$|\langle\Phi|\Psi\rangle|^2 = |\langle0|U_\Phi^\dagger U_\Psi|0\rangle|^2 = \langle0|U_\Psi^\dagger U_\Phi|0\rangle\langle0|U_\Phi^\dagger U_\Psi|0\rangle\,.$$

The compute-uncompute method has the advantage of not requiring any additional qubits beyond those required to represent the circuits U_Ψ and U_Φ. It does, however, require combining the individual circuits U_Φ and U_Ψ into a single quantum circuit which therefore has twice the depth of the initial circuits.

The Hadamard SWAP-Test

The overlap $|\langle\Phi|\Psi\rangle|^2$ may also be computed through the so-called SWAP test method for which the associated circuit is shown below. The essential idea of this method is to construct a circuit containing an ancillary qubit such that the probability $p(0)$ of measuring 0 on the ancillary qubit is related to the overlap through the relation

$$p(0) = \frac{1+|\langle\Phi|\Psi\rangle|^2}{2}\,.$$

The SWAP test circuit has the advantage of having the same circuit depth as that of the individual circuits U_Ψ and U_Φ representing the states $|\Psi\rangle$ and $|\Phi\rangle$ respectively. On the other hand, this efficiency in circuit depth comes at the cost of doubling the number of qubits and requiring N SWAP gates.
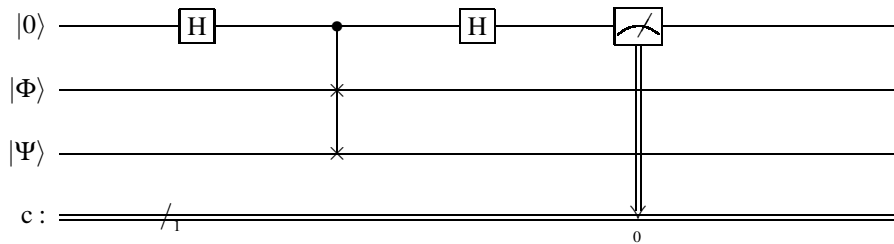

Figure 2. Hadamard SWAP-Test circuit to compute the overlap $|\langle\Phi|\Psi\rangle|^2$

.

2.6 Hybrid Measurements of Quantum Simulations

Before proceeding to the actual numerical results, let us briefly describe the specific outcomes that we wish to explore using QPU implementations of these adaptive algorithms. The primary objective of such adaptive procedures is to yield a wave-function ansatz that accurately represents the ground state of the physical system under study. Our goal in executing such adaptive algorithms on the QPU therefore is to obtain an ordered set of operators (together with corresponding optimal parameters) whose application to the initial state, yield a state that exhibits a high fidelity to the true ground state of the physical system under study. Note that the fidelity of two quantum states $|\psi\rangle$ and $|\phi\rangle$ is defined as the overlap squared of these two states, i.e., $F(|\psi\rangle, |\phi\rangle) = |\langle\psi|\phi\rangle|^2$.

To achieve the sought-after state preparation, the GGA-VQE algorithm minimises the variational energy of the ansatz wave-function, while the Overlap-GGA-VQE algorithm maximises the overlap (fidelity) of the ansatz with an accurate target state. In order to evaluate the success of our QPU implementations, we must therefore measure the fidelities of the ansatz wave-functions generated by these adaptive algorithms and compare the results obtained on the QPU to those obtained from an HPC simulator. Large differences between the results obtained from the QPU implementations and those obtained from the classical HPC simulations would naturally indicate that the measurement and hardware noise on the quantum device is too great for our algorithm to succeed.

It is crucial to note, however, that any divergence between the QPU and HPC simulator results, while ultimately due to device noise, can nevertheless arise from two different sources:

1. The divergence could be due to algorithmic failure. In other words, the noise on the quantum device results in the adaptive algorithm either selecting the wrong parametrised unitary operators to append to the current ansatz, or the wrong 'optimal' angles, or both.

2. The divergence could be due to ansatz evaluation errors. In other words, while the adaptive algorithm has succeeded (at least to a certain degree) in selecting appropriate parametrised unitary operators and corresponding optimal angles, the actual measurement of the fidelity on the quantum device is a failure due to hardware and measurement noise.

In our opinion, the ansatz evaluation error is largely a reflection of the limits of the current QPU hardware and not indicative of the failure of our adaptive algorithms. Consequently, we will primarily focus on evaluating possible algorithmic failures in the QPU implementations of our adaptive algorithms. An obvious strategy to make this evaluation, is to retrieve the ansatz wave-function yielded by the QPU-implemented GGA-VQE or Overlap-GGA-VQE methods, represent this ansatz wave-function using classical methods on a HPC simulator, and measure the sought-after observables. We refer to this approach as 'hybrid' observable evaluation in the sequel. It is important to emphasise that the construction of the ansatz wave-function, i.e., the choice and order of unitary operators as well as the corresponding angles are all determined by calculations on the QPU. It is only the final measurement of the ansatz wave-function that takes place on the simulator. Of course, it is also of interest to determine the degree of ansatz evaluation errors and we will, where possible, evaluate the energy or fidelity of the generated ansatz wave-function directly on the quantum computer.
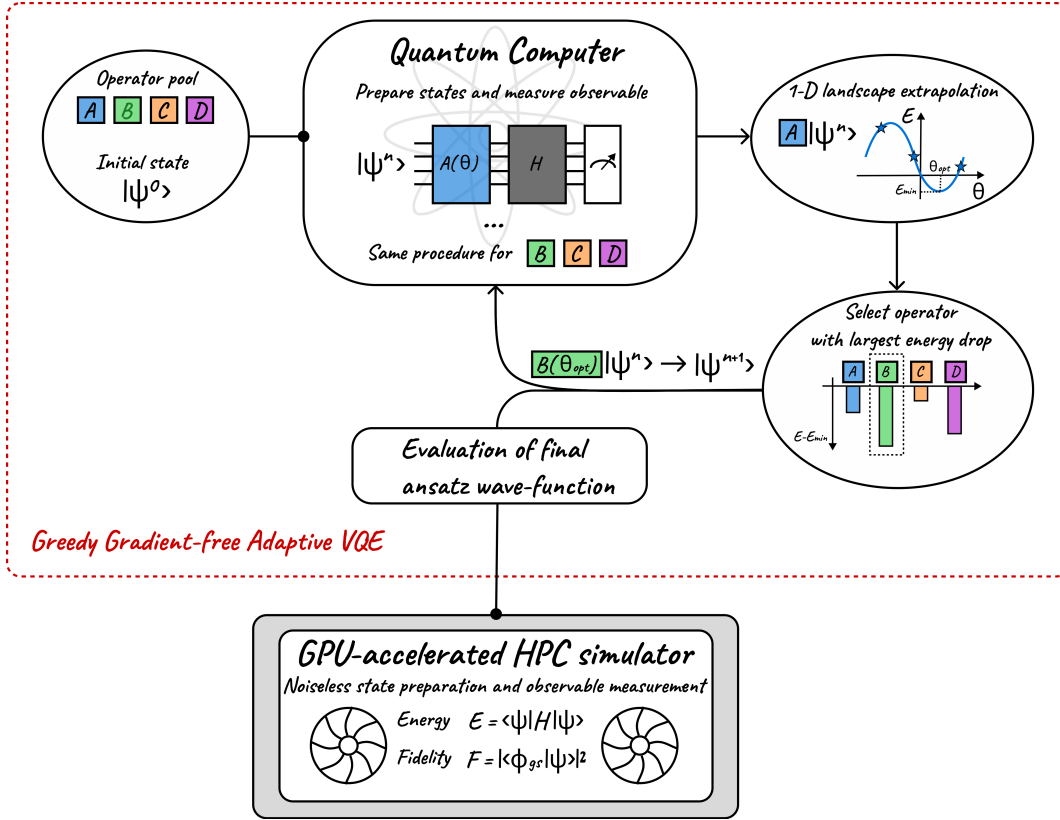
Figure 3. A schematic overview of the GGA-VQE algorithm with hybrid observable measurement. We emphasise that the HPC simulator is used <u>only</u> to evaluate the final ansatz wave-function obtained by executing GGA-VQE on the QPU.

3 Results

The algorithmic procedures in this research have been performed using an in-house code developed within the Amazon Braket SDK. All quantum computations were performed using Amazon Braket and have been executed on an IonQ Aria 25-qubits trapped-ion quantum computer which incorporates built-in error mitigation techniques, and each observable is evaluated using 2500 shots. Classical simulations were conducted using our internal Hyperion multi-GPU-accelerated quantum simulator[36] with all simulator computations being run on a single NVIDIA DGX A100 node.

3.1 The GGA-VQE algorithm applied to the Ising Model

For our first set of numerical experiments, we apply the GGA-VQE algorithm described in Section 2.3 to the transverse-field Ising model described in Section 2.4. We set the system parameters of the Ising Hamiltonian to $h = 0.5$, $J = 0.2$, which ensures that the two-body interactions in this Ising Hamiltonian play an important role.
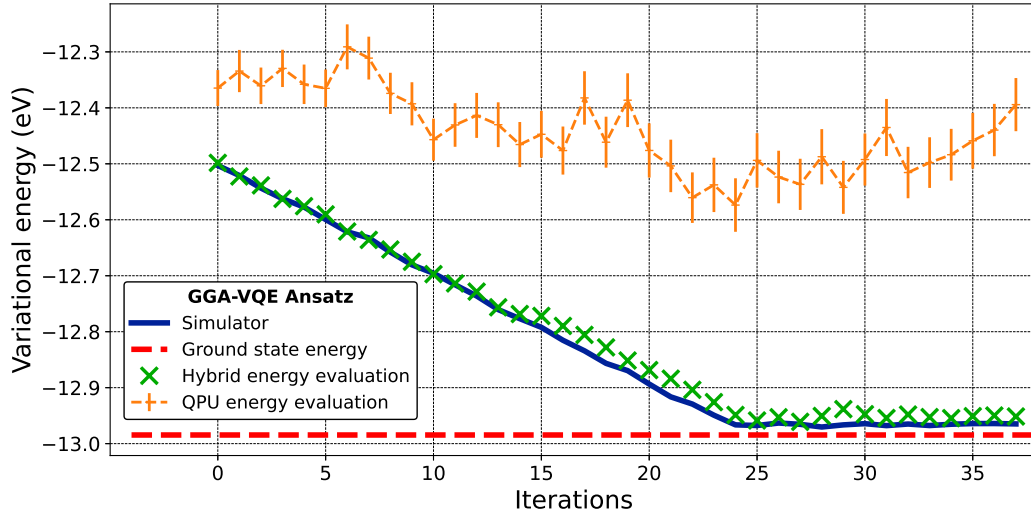
Figure 4. Energy convergence of the GGA-VQE algorithm with respect to the number of iterations. The blue reference curve denotes the energy of a classically simulated ansatz. The green and orange curves denote the hybrid and QPU energy evaluations of the GGA-VQE ansatz wave-function respectively. Note that the hybrid evaluation is carried out by retrieving the GGA-VQE ansatz wave-function generated by the QPU, re-implementing it on the Hyperion HPC simulator, and then evaluating the variational energy.

Figure 4 illustrates the convergence of the hybrid energy evaluations of the GGA-VQE ansatz wave-function with respect to the number of algorithm iterations. We remind the reader that these hybrid energy evaluations are obtained by first running the GGA-VQE algorithm on the IonQ Aria QPU, retrieving the resulting ansatz wave-function and re-implementing it on the Hyperion HPC simulator, and then evaluating the variational energy on the HPC simulator. For reference, we also plot the corresponding energy curve obtained by executing the GGA-VQE ansatz directly on the HPC Hyperion simulator using 10^6 samples per measured circuit. In addition, as an indication of the measurement and hardware noise, we also plot the GGA-VQE energies obtain by direct measurement on the QPU.

Figure 5 clearly indicates that QPU-implemented GGA-VQE procedure successfully provides an ansatz wave-function that closely matches the ground state. Moreover, the QPU implementation and HPC simulator implementation of the GGA-VQE algorithm seem highly consistent despite significant noise in the quantum evaluation of observables, as noticeable from the QPU energy evaluation curve in Figure 4. Indeed, the greedy, gradient-free operator selection procedure that we have introduced in this study, which relies on a function extrapolation using five noisy evaluations on the QPU, is able to build an ansatz with an energy error below 2.50×10^{-2} eV and a fidelity exceeding 98% with the exact ground state (see Figure 5).
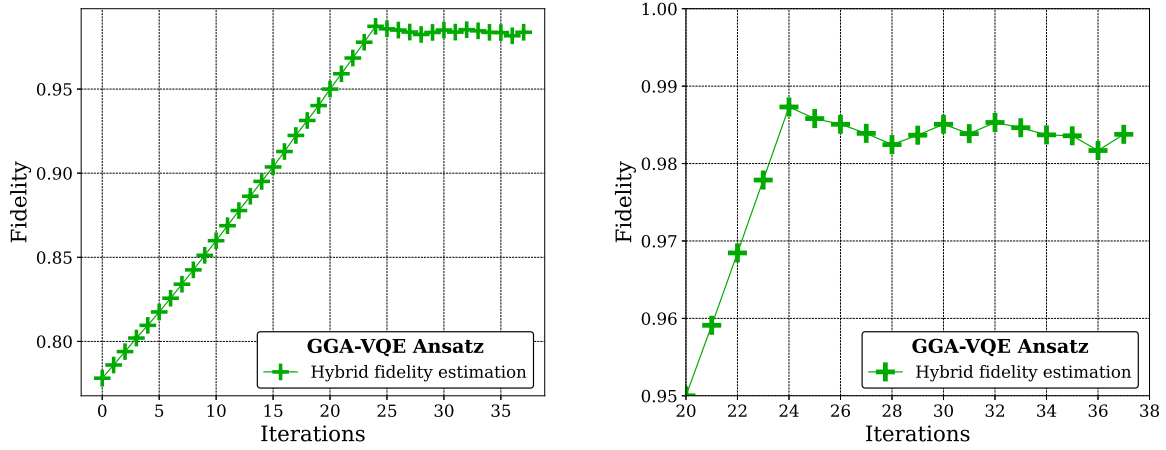
Figure 5. A convergence plot of the fidelity of the GGA-VQE ansatz wave-function produced by the QPU and re-implemented in the Hyperion HPC simulator (hybrid evaluation approach) with the exact ground state of this Ising model obtained using a diagonalisation procedure on the Hyperion simulator. The figure on the right is a zoomed-in version of the figure on the left to better appreciate the fidelity of the GGA-VQE ansatz wave-function.
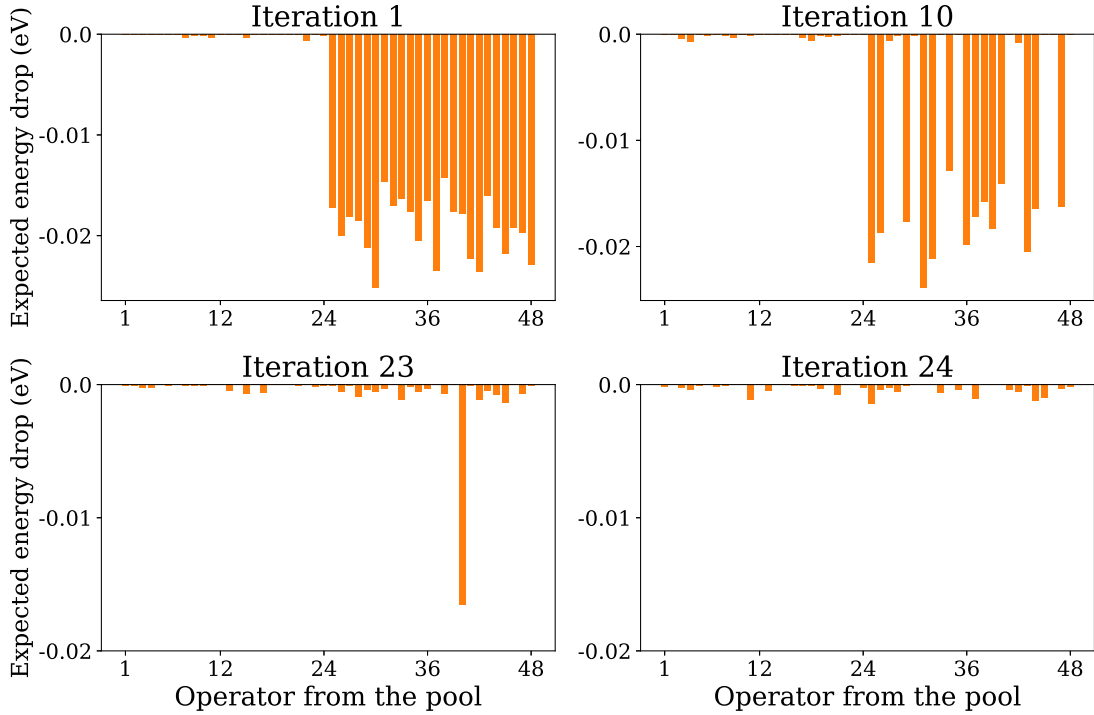
Figure 6. Expected energy drop of each Hermitian generator from the operator pool during the GGA-VQE iterative procedure on the QPU. The minimal pool operators, numbered from 1 to 48, are listed in the same order as defined in Equation (18). All energy differences are expressed in eV.

To better illustrate the outstanding robustness of the GGA-VQE procedure with respect to QPU noise, we depict in Figure 6 the expected maximal energy drop of each Hermitian generator from the chosen minimal operator pool throughout the iterative procedure. We observe maximal energy drops of approximately 1.5×10^{-2} eV for the first 24 iterations followed by a great decrease in the potential energy drops from iteration 25 onwards. This is consistent with the energy curve displayed

in Figure 4 which steadily decreases for the first 24 iterations and then reaches a plateau. It is important to note that while Figure 4 displays a decrease in the hybrid evaluation of energy between the 24th and 25th iterations, in our opinion, this decrease cannot be attributed to the newly introduced operator. Rather, it is simply a fortuitous consequence of measurement noise. Indeed, the reference simulator curve does not demonstrate any such energy decrease after the 24th iteration. Note that Figure 6 also suggests a possible explanation for the noise-resilience of the GGA-VQE algorithm. Indeed, we see that the pool of potential unitary operators that can be appended to the current ansatz wave-function can broadly be divided into two camps, namely, operators whose addition to the current ansatz will lead to an approximate energy drop of 1.5×10^{-2} eV, and operators whose addition to the current ansatz will not meaningfully lower the expectation energy. Thus, the presence of hardware noise can result in a <u>less optimal</u> operator from the first camp being appended to the current ansatz but it will likely not yield a useless operator from the second camp.

For further confirmation, we examine the energy landscapes associated with certain Hermitian generators from the operator pool, extrapolated using five noisy measurements on the QPU. We compare these noisy QPU-based landscapes with the reference landscapes obtained from the Hyperion HPC simulator. Our results, displayed in Figure 7, indicate a nearly perfect match for an operator that enables an energy drop (Z_0Y_1), as well as for an operator that does not improve the ansatz (Y_0) at the first GGA-VQE iteration. This finding explains the remarkable resilience to QPU noise of operator selection in the GGA-VQE procedure, and suggests that the QPU-implemented algorithm can consistently pick the optimal operator and associated parameter, resulting in a gradual reduction of the variational energy of the ansatz wave-function and thus convergence towards the ground state.
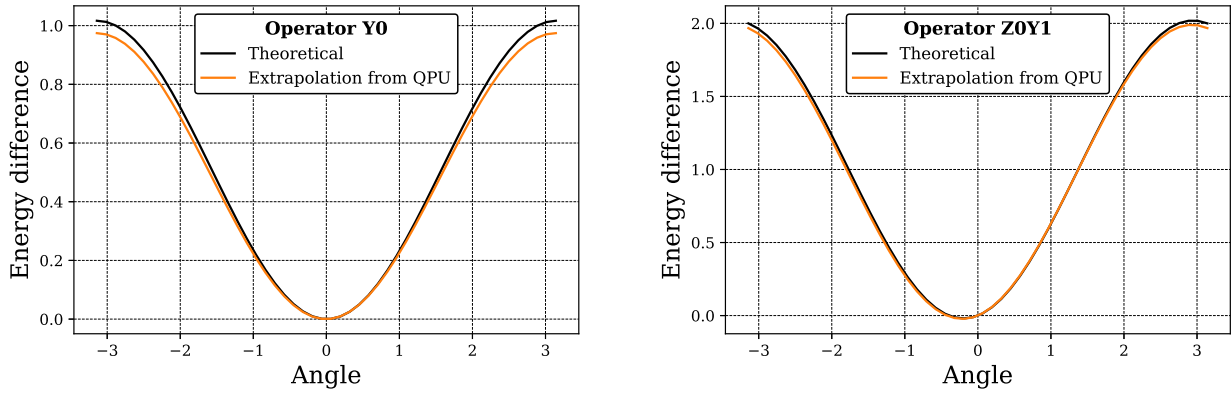

Figure 7. One-dimensional energy landscapes of the operators Y_0 and Z_0Z_1 when applied to the initial state. The orange curve is extrapolated from five noisy circuit evaluations following the method described in 2.4. The black curve is the exact energy landscape obtained from an HPC simulation. The energy differences are expressed in eV, and the angles are given in radians.

3.2 Overlap-GGA-VQE for a Stretched HF molecule

For our next set of numerical experiments, we consider the application of the Overlap-GGA-VQE algorithm for the approximation of the ground state eigenfunction of the HF molecule at a bond distance of 2.5 Å. We consider an active space of 8 electrons in 10 spin orbitals in the minimal STO-3g basis set, thus freezing the lowest 1s orbital as doubly occupied. The Hartree-Fock state can therefore be represented as $|\phi\rangle = |1111111100\rangle$, which requires 10 qubits. The target wave-function for the Overlap-GGA-VQE process is obtained using a QEB-ADAPT-VQE procedure carried out on the Hyperion HPC simulator until convergence at the chemical accuracy level. The resulting QEB-ADAPT-VQE target wave-function, which has an error of about 1.4 mHa, is constructed using four generators from the Qubit Excitation-based pool (see Section 2.2) leading to a total CNOT circuit count of 32. The purpose of applying the Overlap-GGA-VQE algorithm is to obtain a high-fidelity approximation of this target wave-function using fewer CNOT gates.

We employ a subset $\mathbb{P}$ of the qubit hardware-efficient pool introduced in Section 2.2. More precisely, we define an index set P for pairs of qubits given by

$$P = \{(4,0),(8,0),(5,1),(9,1),(5,0),(7,0),(7,1)\}.$$

Corresponding to this index set P, we define the sub-pool $\mathbb{P}$ of qubit hardware-efficient operators as

$$\mathbb{P} = \{X_{\mathbf{p}} = \frac{1}{2}X_{p_2}X_{p_1} : \mathbf{p} = (p_1, p_2) \in P\}.$$

In other words, $\mathbb{P}$ consists of a collection of <u>single excitation</u> qubit hardware-efficient operators (recall Equation (7)). Equipped with the operator pool $\mathbb{P}$, we apply the Overlap-GGA-VQE algorithm to the target QEB-ADAPT-VQE wave-function. It is important to note that, for the current HF system, the initial Hartree-Fock state exhibits no overlap with the QEB-ADAPT-VQE target.

(a) QPU
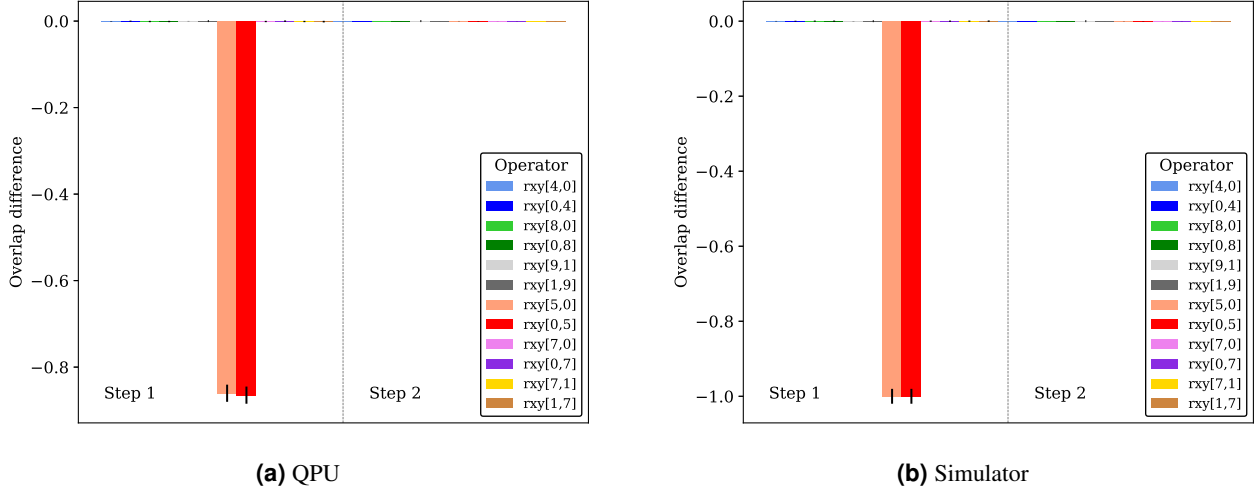
(b) Simulator

Figure 8. Overlap difference induced by each operator from the operator pool $\mathbb{P}$ when applied to the ansatz in the QPU execution (Fig. 8a), and in a classical simulation (Fig. 8b). The steps correspond to the iterations in the Overlap-GGA-VQE procedure. At step 2, none of the operators in the pool can significantly improve the overlap with the target state, indicating that convergence has been achieved.
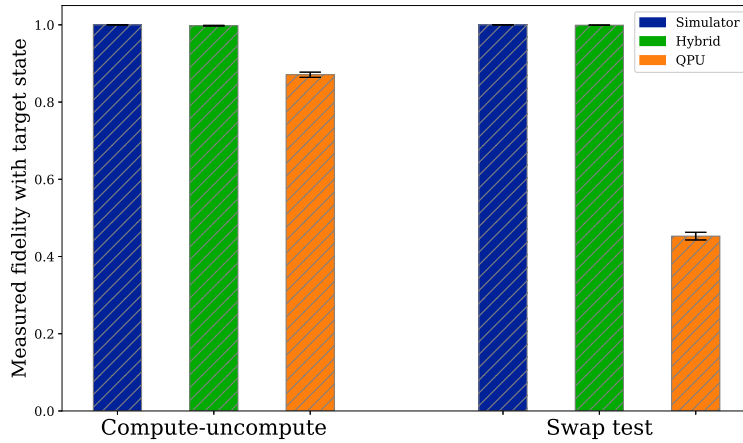
Figure 9. Ansatz fidelity of the Overlap-GGA-VQE ansatz with the target state. The blue bars correspond to the fidelity of a classically simulated ansatz with the target state. The green and orange bars denote the hybrid and QPU fidelity evaluations respectively. Note that the hybrid evaluation is carried out by retrieving the Overlap-GGA-VQE ansatz wave-function generated by the QPU, re-implementing it on the Hyperion HPC simulator, and then evaluating the variational energy.

Figure 8a shows the maximal overlap increases that can be achieved by the application of different operators from the pool on the initial Hartree-Fock state. These overlap values are obtained through extrapolation of the one-dimensional objective function defined in Equation (13) for each operator from the pool $\mathbb{P}$. The extrapolations require, for each operator, two noisy

overlap evaluations in the QPU. It is readily seen however, that despite the possible hardware and measurement noise on the QPU, the QPU extrapolations are nearly identical to the extrapolations obtained using an HPC simulator implementation of the Overlap-GGA-VQE algorithm displayed in Figure 8b. This numerical result supports the noise robustness of our operator selection approach. Indeed, we see that in both the QPU implementation and the simulator implementation, the Overlap-GGA-VQE algorithm correctly identifies the operators acting on the qubit pair (5, 0) as the ones that should lead to the largest increase in the overlap, and thus the largest increase in fidelity with the target. For this simple example, convergence is reached after a single iteration as, at the second step, no generator from the operator pool can further improve the overlap with the target state.

The final fidelities of the QPU-implemented and simulator-implemented Overlap-GGA-VQE ansatz wave-functions are plotted in Figure 9. We remind the reader that the term 'hybrid fidelity evaluation' refers to the classically recomputed fidelity of the QPU-generated ansatz (see the discussion at the start of this section). This hybrid fidelity evaluation precisely matches the value of the fidelity obtained through the pure HPC simulator implementation of Overlap-GGA-VQE, regardless of the choice of overlap measurement technique.

For completeness, we have also plotted the fidelities obtained from the QPU implementation of the Overlap-GGA-VQE procedure through direct measurement on the QPU. In this case, we see that the Swap test exhibits a higher noise level than the compute-uncompute method which indicates that– at least for this hardware– a deeper circuit involving only 10 qubits is less affected by device noise than a shorter circuit that requires gates and connectivity across 20 qubits. Note, however, that in both cases, the hybrid evaluation of the Overlap-GGA-VQE ansatz is highly accurate as indicated in the green bars in Figure 9. Indeed, the QPU implementation of the Overlap-GGA-VQE procedure manages to provided an ansatz wave-function that achieves a fidelity of over 99% with a chemically accurate target wave-function while using only 2 CNOT gates.

4 Discussion

In this study, we have developed new resource-saving strategies to execute, for the first time, adaptive variational quantum algorithms on a state-of-the-art, 25-qubit trapped ion, error-mitigated quantum computer. While a great deal of effort has recently been devoted to developing adaptive variational quantum algorithms that yield ultra-compact ansatz wave-functions, the actual realization of such adaptive algorithms on the current generation of NISQ devices has received less attention. Since the main bottleneck in practical implementations on NISQ devices is the multi-dimensional optimization of a highly noise cost function, we have introduced a new noise-resistant and resource-efficient, greedy gradient-free variational quantum algorithm that relies on operator-by-operator local optimizations using only a small number of measurements on the quantum device.

As a physics application, we have used the novel greedy gradient-free adaptive variational quantum eigensolver (GGA-VQE) introduced in this paper to successfully compute the ground state of an open boundary 25-qubit transverse-field Ising Hamiltonian, achieving a ground state fidelity of over 98%. The GGA-VQE algorithm that we have developed for the Ising model is also highly scalable since each iteration of this method requires a fixed number of circuit measurements, regardless of the number of qubits or the size of the operator pool. Ising models have already been studied using various methods in quantum regimes that claim to surpass the memory capacity of classical computers[37,38], and these studies, in combination with ours, demonstrate promising results in the potential of useful quantum computation before the era of fault-tolerance.

As an additional application targeted at chemistry applications, we have combined our greedy approach with the Overlap-ADAPT-VQE algorithm introduced in[18] to compute compact approximations of a target wave-function through an iterative, adaptive overlap maximisation procedure. We have applied this novel Overlap-GGA-VQE algorithm to a stretched 10-qubit hydrogen fluoride (HF) molecular system and shown that the algorithm is able to generate a highly compact approximation of a target approximate ground-state that achieves a fidelity of over 99%. The target ground-state for this numerical experiment was generated through a QEB-ADAPT-VQE[15] procedure on a classical simulator while the wave-function overlaps– required by the Overlap-GGA-VQE procedure– were measured on the QPU using two different methods: the compute-uncompute method and the Swap test.

For both the Ising model and the stretched HF molecule, we have demonstrated that, despite the high level of device noise in observable quantum measurements, our noise-resistant GGA-VQE procedure can select a sequence of unitary operators and corresponding optimal angles that can be used to construct an accurate approximation of the ground state. Indeed, our greedy operator selection relies on an extrapolation of the associated objective function using a minimal number of noisy quantum measurements, and this extrapolation technique seems resilient to device noise, as evidenced by the close alignment between the QPU-extrapolated objective function and the HPC simulator-extrapolated objective function. Moreover, because we utilize extrapolated objective functions for the VQE portion of the algorithm, our greedy, gradient-free protocols do not require any multi-dimensional noisy optimization at all, thus bypassing the main bottleneck of QPU implementations of adaptive VQEs[13].

Let us emphasize that the energy sorting procedure for optimal operator selection that we have developed in this study is easily extendable to multi-operator selection and optimization at the cost of a higher number of measurements on the quantum device, and some preliminary ideas in this direction have been presented in Section 2.3. While the ground state preparation of the relatively simple Hamiltonians considered in this study could be effectively carried out by appending one locally optimal operator at a time to the current ansatz wave-function, it is likely that the multi-operator generalizations of our energy sorting procedure (see Section 2.3) will be effective in the ground state preparation of strongly correlated systems such as stretched linear chains of hydrogen atoms. Similarly, the extensions of the Ising model GGA-VQE algorithm that we have developed (see Section 5.4) can easily be applied to other spin-chain systems such as the Hubbard model. Further research in both these directions will be the subject of future work.

In our opinion, the successful implementation of adaptive variational algorithms on the quantum hardware of today indicates the suitability of these algorithms for approximate state preparation that can be used as the basis of a more accurate Quantum Phase Estimation (QPE) procedure[5, 39, 40] to evaluate the ground state energy of a given Hamiltonian. Since the probability of success for QPE is directly proportional to the fidelity between the approximate eigenstate and the true eigenstate, accurate, adaptive hybrid algorithms can play an important role in the pre-processing step for quantum phase estimation. Independent of such a pre-processing application, let us also point out that certain interesting studies have demonstrated potential applications of adaptive algorithms to dynamic simulation problems[41, 42].

Data availability

Data generated during the study is available upon request from the authors (E-mail: jean-philip.piquemal@sorbonne-universite.fr).

Code availability

The code used during the study is available upon request from the authors (E-mail: jean-philip.piquemal@sorbonne-universite.fr).

Acknowledgements

5 Appendix

5.1 Quantum circuits for qubit-excitation operators

For the sake of completeness, we present a few key quantum circuits used in the hardware experiments carried out for this study. The circuit for a single-qubit excitation is given in Figure 10 whereas the circuit for a double-qubit excitation is displayed in Figure 11. Note that both qubit excitations correspond to the qubit-excitation based (QEB) pool introduced in Section 2.2, and, as explained in[43], the circuits displayed here are the most hardware-efficient implementations of these operators.
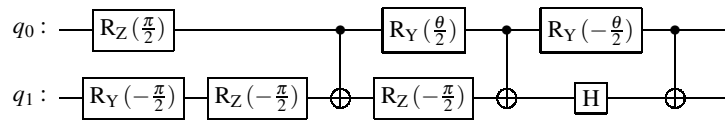

Figure 10. A quantum circuit performing a generic single-qubit evolution[43].

Remark: when displaying quantum circuits, H denotes the Hadamard gate and not a physical Hamiltonian.

Figure 11. A quantum circuit performing a generic double-qubit evolution[43].

5.2 Periodicity of QEB operator pool and involutory property of hardware-efficient pools

Recall the definition of the qubit excitation-based (QEB) pool given in Section 2.2 and A_{pqrs} denote a double-qubit excitation generator between the pairs of qubits (p,q) and (r,s) as defined through Equation (6), i.e.,

$$A_{pqrs} = \frac{1}{8}\left(X_r Y_s X_p X_q + Y_r X_s X_p X_q + Y_r Y_s Y_p X_q + Y_r Y_s X_p Y_q - X_r X_s Y_p X_q - X_r X_s X_p Y_q - Y_r X_s Y_p Y_q - X_r Y_s Y_p Y_q\right).$$

A direct calculation shows that A_{pqrs} can be written in the equivalent form

$$A_{pqrs} = i(Q_p^\dagger Q_q^\dagger Q_r Q_s - Q_r^\dagger Q_s^\dagger Q_p Q_q),$$

where for any qubit index a, we define $Q_a = \frac{1}{2}(X_a + iY_a)$.

Using this representation, we can easily show that $A_{pqrs}|1_p 1_q 0_r 0_s\rangle = i|0_p 0_q 1_r 1_s\rangle$, $A_{pqrs}|0_p 0_q 1_r 1_s\rangle = -i|1_p 1_q 0_r 0_s\rangle$ while the action of A_{qprs} on all other states is zero. Consequently, the subspace spanned by $e_1 \equiv |1_p 1_q 0_r 0_s\rangle$ and $e_2 \equiv |0_p 0_q 1_r 1_s\rangle$ is an invariant subspace of A_{pqrs}, and in the basis $\{e_1, e_2\}$ of this invariant subspace, A_{pqrs} has the representation

$$\begin{pmatrix} 0 & -\iota \\ \iota & 0 \end{pmatrix},$$

which is the well-known Y Pauli matrix. We thus conclude that A_{pqrs} has eigenvalues $0, \pm 1$ and satisfies $A_{pqrs}^3 = A_{pqrs}$.

A similar demonstration can be carried out for single-qubit generators from the QEB pool and generators from the qubit hardware efficient and minimal hardware efficient pools, which shows that these generators are involutory, i.e., they satisfy $B^2 = I$. For the sake of brevity, we do not provide a detailed argument.

The above observation motivates further investigation and leads to the following result.

Theorem 1. *Let H denote an N-qubit Hamiltonian and let $\mathbb{P}$ denote any of the operator pools introduced in Section 2.2. Then define, for any N-qubit wave-function $|\phi\rangle$, any Hermitian generator $B \in \mathbb{P}$ and any $\theta \in [-\pi, \pi)$ the landscape function*

$$\mathscr{L}(B, \theta, |\phi\rangle) = \langle\phi|\exp(\iota\theta B)H\exp(-\iota\theta B)|\phi\rangle.$$

Then it holds that

$$\mathscr{L}(B, \theta, |\phi\rangle) = \begin{cases} \langle\phi|H|\phi\rangle + \big(\cos(\theta)-1\big)\big(\langle\phi|\{H,B^2\}|\phi\rangle - 2\langle\phi|BHB|\phi\rangle\big) & \text{if } B^3 = B, \\ + \big(1-\cos(\theta)\big)^2\big(\langle\phi|B^2 H B^2|\phi\rangle - \langle\phi|BHB|\phi\rangle\big) \\ + \sin(\theta)(\cos(\theta)-1)\langle\phi|\iota B[H,B]B|\phi\rangle \\ + \sin(\theta)\langle\phi|\iota[B,H]|\phi\rangle \\ \\ \cos^2(\theta)\langle\phi|H|\phi\rangle + \dfrac{\sin(2\theta)}{2}\langle\phi|\iota[B,H]|\phi\rangle & \text{if } B^2 = I. \\ + \sin^2(\theta)\langle\phi|BHB|\phi\rangle, \end{cases}$$

where $\{\cdot,\cdot\}$ and $[\cdot,\cdot]$ denote the anti-commutator and commutator respectively.

Proof. We consider first the case $B^3 = B$. For such a Hermitian generator, we can use the Taylor series expansion of the exponential to deduce that

$$
\begin{aligned}
\exp(-\iota\theta B) &= \sum_{k=0}^{\infty} \frac{(-i\theta B)^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{(-i\theta B)^{2k+1}}{(2k+1)!} \\
&= I + (\cos(\theta) - 1)B^2 - \iota\sin(\theta)B.
\end{aligned}
\tag{22}
$$

Plugging in the expression (22) into the definition of the landscape function $\mathscr{L}(B, \theta, |\phi\rangle)$ now yields the desired result. The case $B^2 = I$ is simply a special case. □

5.3 Analytical expressions of GGA-VQE objective functions for the Ising Hamiltonian

Throughout this section, we use the setting and notation of Section 2.4. Our goal now is to demonstrate that for the Ising Hamiltonian defined through Equation (17) and the minimal hardware-efficient pool $\mathbb{P}$ given by (18), the objective function $\mathscr{L}(B, \theta, |\Psi^{(m-1)}\rangle)$ has the following simple structure:

$$
\mathscr{L}(B, \theta, |\Psi^{(m-1)}\rangle) =
\begin{cases}
\left\langle \Psi^{(m-1)} \middle| H \middle| \Psi^{(m-1)} \right\rangle & \text{if } B = Y_p, \\
+ \sin(2\theta)\left\langle \Psi^{(m-1)} \middle| hZ_p - J(X_pZ_{p+1} + Z_{p-1}X_p\delta_{p>0}) \middle| \Psi^{(m-1)} \right\rangle & \\
- 2\sin^2(\theta)\left\langle \Psi^{(m-1)} \middle| hX_p + J(Z_pZ_{p+1} + Z_{p-1}Z_p\delta_{p>0}) \middle| \Psi^{(m-1)} \right\rangle. & \\
\\
\left\langle \Psi^{(m-1)} \middle| H \middle| \Psi^{(m-1)} \right\rangle & \text{if } B = Z_pY_{p+1}. \\
+ \sin(2\theta)\left\langle \Psi^{(m-1)} \middle| h(Z_pZ_{p+1} - Y_pY_{p+1}) \middle| \Psi^{(m-1)} \right\rangle & \\
- \sin(2\theta)\left\langle \Psi^{(m-1)} \middle| J(X_{p+1} + Z_pX_{p+1}Z_{p+2}\delta_{p+2<n}) \middle| \Psi^{(m-1)} \right\rangle & \\
- 2\sin^2(\theta)\left\langle \Psi^{(m-1)} \middle| hX_p + hX_{p+1} + JZ_pZ_{p+1} + JZ_{p+1}Z_{p+2}\delta_{p+2<n} \middle| \Psi^{(m-1)} \right\rangle.
\end{cases}
\tag{23}
$$

To show that Equation (23) indeed holds, we first recall the definition of the objective function $\mathscr{L}(B, \theta, |\Psi^{(m-1)}\rangle)$ which is given by

$$
\mathscr{L}(B, \theta, |\Psi^{(m-1)}\rangle) = \langle \Psi^{(m-1)} | \exp(\iota\theta B) H \exp(-\iota\theta B) | \Psi^{(m-1)} \rangle,
\tag{24}
$$

where $B \in \mathbb{P}$ is any Hermitian generator from the minimal hardware-efficient operator pool, the parameter $\theta \in [-\pi, \pi)$, and $|\Psi^{(m-1)}\rangle$ denotes the previous ansatz wave-function.

Next, we recall from Equation (13) that the involutory property of the Hermitian generators from the minimal hardware-efficient pool yields the following simplification of Equation (24):

$$
\mathscr{L}(B, \theta, |\Psi^{(m-1)}\rangle) = \cos^2(\theta)\langle \phi|H|\phi\rangle + \frac{\sin(2\theta)}{2}\langle \phi|\iota[B,H]|\phi\rangle + \sin^2(\theta)\langle \phi|BHB|\phi\rangle.
\tag{25}
$$

Consequently, in order to arrive at Equation (23), we have to simplify each term involving the Ising Hamiltonian and minimal hardware-efficient generator B appearing in Equation (25).

To do so, recall that we denote the total number of qubits (i.e., the size of the quantum register) by $N \in \mathbb{N}$, and fix an index $p \in \{0, \ldots, N-2\}$. Using now the commutation relations of the Pauli matrices, a direct calculation reveals that

$$
[Y_p, H] = -2h\iota Z_p + 2\iota J(X_pZ_{p+1} + Z_{p-1}X_p\delta_{p>0})
$$

and

$$
[Z_pY_{p+1}, H] = 2\iota h(Y_pY_{p+1} - Z_pZ_{p+1}) + 2\iota J(X_{p+1} + Z_pX_{p+1}Z_{p+2}\delta_{p<N-2}).
\tag{26}
$$

A similar calculation utilising once again the commutation relations of the Pauli matrices further yields that

$$
Y_pHY_p = H - 2hX_p - 2J(Z_pZ_{p+1} + Z_{p-1}Z_j\delta_{p>0}),
$$

and

$$ZZ_{p+1}HZ_pZ_{p+1} = \sum_{q=0}^{N-1} Z_pZ_{p+1}X_qZ_pZ_{p+1} = H - 2h(X_p + X_{p+1}). \tag{27}$$

The result now follows by plugging in Equations (26) and (27) into Equation (25).

5.4 Reducing the computational complexity of the energy sorting algorithm for general spin chains

As demonstrated in Section 2.4, the specific structure of transverse-field Ising Hamiltonian leads to a huge reduction in the computational cost of the energy sorting step of the GGA-VQE algorithm. Indeed, while the energy sorting step a priori requires $\mathscr{O}(M)$ measurements for a general system Hamiltonian and an operator pool of size M, the number of required measurements reduces to just <u>five</u> in the case of the one-dimensional transverse field Hamiltonian. The goal of this section is to briefly describe similar reductions in the computational complexity of the energy sorting algorithm for Ising spin-chain Hamiltonians with local magnetic fields and couplings in all three spatial directions, i.e., Hamiltonians of the form

$$H = \sum_{k=0}^{N-1} h_k^x X_k + \sum_{k=0}^{N-1} h_k^z Z_k + \sum_{k=0}^{N-2} J_k^x X_k X_{k+1} + \sum_{k=0}^{N-2} J_k^y Y_k Y_{k+1} + \sum_{k=0}^{N-2} J_k^z Z_k Z_{k+1}. \tag{28}$$

Here, h_k^x and h_k^z denote constants that model the intensity of the magnetic field along the x and z directions while J_k^x, J_k^y and J_k^z are constants that model the strength of the nearest-neighbour interactions in the $x, y,$ and z directions respectively.

Tables 2 and 3 list the terms of interest that appear in the one-dimensional GGA-VQE landscape functions that are used to perform the energy sorting step. Comparing the terms that appear in Tables 2 and 3 with the simpler expressions for the transverse-field Ising Hamiltonian from Section 2.4, we see that the only new terms that arise are of the $Z_{p-2}Z_{p-1}X_p$ and $Y_{p-1}X_p$. As before, we can simultaneously measure such operators acting on a disjoint set of qubits– a process that will require an additional five quantum circuits at each step. Consequently, applying the GGA-VQE algorithm to general Ising Hamiltonians of the form (28) will require constructing and measuring at most ten quantum circuits, irrespective of the number of qubits and the size of the minimal operator pool.

$\mathbf{H}$	$\sum \mathbf{h_k X_k}$	$\sum \mathbf{h_k Y_k}$	$\sum \mathbf{h_k Z_k}$
$[\mathbf{Y_i, H}]$	$-2ih_iZ_i$	0	$2ih_iX_i$
$\mathbf{Y_iHY_i}$	$H - 2h_iX_i$	H	$H - 2Z_i$
$[\mathbf{Z_iY_{i+1}, H}]$	$-2ih_iZ_iZ_{i+1} + 2ih_{i+1}Y_iY_{i+1}$	$-2ih_iX_iY_{i+1}$	$2ih_{i+1}Z_iX_{i+1}$
$\mathbf{Z_iY_{i+1}HZ_iY_{i+1}}$	$H - 2h_{i+1}X_{k+1} - 2h_iX_i$	$H - 2h_{i+1}Y_{i+1}$	$H - 2h_iZ_i - 2h_{i+1}Z_{i+1}$

Table 2. Commutators involving generators from the minimal operator pool and the local magnetic field terms.

$\mathbf{H}$	$\sum \mathbf{J_k X_k X_{k+1}}$	$\sum \mathbf{J_k Y_k Y_{k+1}}$	$\sum \mathbf{J_k Z_k Z_{k+1}}$
$[\mathbf{Y_i, H}]$	$-2iJ_iZ_iX_{i+1} - 2iJ_{i+1}X_iZ_{i+1}$	0	$2iJ_iX_iZ_{i+1} + 2iJ_{i-1}Z_{i-1}X_i$
$\mathbf{Y_iHY_i}$	$H - 2J_iX_iX_{i+1} - 2J_{i+1}X_{i-1}X_i$	H	$H - 2J_iZ_iZ_{i+1} - 2J_{i-1}Z_{i-1}Z_i$
$[\mathbf{Z_iY_{i+1}, H}]$	$-2iJ_{i+1}Z_iZ_{i+1}X_{i+2} - 2iJ_{i-1}X_{i-1}Y_iY_{i+1}$	$-2iJ_iX_i - 2iJ_{i-1}Y_{i-1}X_iY_{i+1}$	$2iJ_iX_{i+1} + 2iJ_{i+1}Z_iX_{i+1}Z_{i+2}$
$\mathbf{Z_iY_{i+1}HZ_iY_{i+1}}$	$H - 2J_{i-1}X_{i-1}X_i$	$H - 2J_iY_iY_{i+1} - 2J_{i-1}Y_{i-1}Y_i$	$H - 2J_{i+1}Z_{i+1}Z_{i+2} - 2J_iZ_iZ_{i+1}$

Table 3. Commutators involving generators from the minimal operator pool and the interaction terms in each direction.

Finally, let us remark that we expect similar but likely less drastic simplifications to also hold for Hamiltonians arising from other physical models.

References

1. Feynman, R. P. Simulating physics with computers. In <u>Feynman and computation</u>, 133–153 (CRC Press, 2018).

2. Cao, Y. <u>et al.</u> Quantum chemistry in the age of quantum computing. <u>Chem. Rev.</u> **119**, 10856–10915 (2019).

3. McArdle, S., Endo, S., Aspuru-Guzik, A., Benjamin, S. C. & Yuan, X. Quantum computational chemistry. <u>Rev. Mod. Phys.</u> **92**, 015003 (2020).

4. Peruzzo, A. et al. A variational eigenvalue solver on a photonic quantum processor. <u>Nat. Commun.</u> **5**, 1–7 (2014).

5. Tilly, J. <u>et al.</u> The variational quantum eigensolver: a review of methods and best practices. <u>Phys. Reports</u> **986**, 1–128 (2022).

6. Romero, J. <u>et al.</u> Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. <u>Quantum Sci. Technol.</u> **4**, 014008 (2018).

7. Ryabinkin, I. G., Yen, T.-C., Genin, S. N. & Izmaylov, A. F. Qubit coupled cluster method: a systematic approach to quantum chemistry on a quantum computer. <u>J. Chem. Theory Comput.</u> **14**, 6317–6326 (2018).

8. Kandala, A. <u>et al.</u> Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. <u>Nature</u> **549**, 242–246 (2017).

9. Lee, J., Huggins, W. J., Head-Gordon, M. & Whaley, K. B. Generalized unitary coupled cluster wave functions for quantum computation. <u>J. Chem. Theory Comput.</u> **15**, 311–324 (2018).

10. Mizukami, W. <u>et al.</u> Orbital optimized unitary coupled cluster theory for quantum computer. <u>Phys. Rev. Res.</u> **2**, 033421 (2020).

11. Grimsley, H. R., Claudino, D., Economou, S. E., Barnes, E. & Mayhall, N. J. Is the trotterized uccsd ansatz chemically well-defined? <u>J. Chem. Theory Comput.</u> **16**, 1–6 (2019).

12. Grimsley, H. R., Economou, S. E., Barnes, E. & Mayhall, N. J. An adaptive variational algorithm for exact molecular simulations on a quantum computer. <u>Nat. Commun.</u> **10**, 3007 (2019).

13. Astrakhantsev, N., Mazzola, G., Tavernelli, I. & Carleo, G. Phenomenological theory of variational quantum ground-state preparation (2022). `2205.06278`.

14. Tang, H. L. <u>et al.</u> Qubit-ADAPT-VQE: An adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor. <u>PRX Quantum</u> **2**, 020310 (2021).

15. Yordanov, Y. S., Armaos, V., Barnes, C. H. & Arvidsson-Shukur, D. R. Qubit-excitation-based adaptive variational quantum eigensolver. <u>Commun. Phys.</u> **4**, 1–11 (2021).

16. Anastasiou, P. G., Chen, Y., Mayhall, N. J., Barnes, E. & Economou, S. E. TETRIS-ADAPT-VQE: An adaptive algorithm that yields shallower, denser circuit ansätze (2022). `2209.10562`.

17. Burton, H. G., Marti-Dafcik, D., Tew, D. P. & Wales, D. J. Exact electronic states with shallow quantum circuits from global optimisation. <u>npj Quantum Inf.</u> **9**, 75 (2023).

18. Feniou, C. <u>et al.</u> Overlap-ADAPT-VQE: Practical quantum chemistry on quantum computers via overlap-guided compact ansätze. <u>Commun. Phys.</u> **6**, 192 (2023).

19. Choi, S., Yen, T.-C. & Izmaylov, A. F. Improving quantum measurements by introducing "ghost" pauli products. <u>J. Chem. Theory Comput.</u> **18**, 7394–7402 (2022).

20. Loaiza, I., Khah, A. M., Wiebe, N. & Izmaylov, A. F. Reducing molecular electronic hamiltonian simulation cost for linear combination of unitaries approaches. <u>Quantum Sci. Technol.</u> **8**, 035019 (2023).

21. Yen, T.-C., Ganeshram, A. & Izmaylov, A. F. Deterministic improvements of quantum measurements with grouping of compatible operators, non-local transformations, and covariance estimates. <u>npj Quantum Inf.</u> **9**, 14 (2023).

22. Nakanishi, K. M., Fujii, K. & Todo, S. Sequential minimal optimization for quantum-classical hybrid algorithms. <u>Phys. Rev. Res.</u> **2**, 043158 (2020).

23. Ostaszewski, M., Grant, E. & Benedetti, M. Structure optimization for parameterized quantum circuits. <u>Quantum</u> **5**, 391 (2021).

24. Wada, K. <u>et al.</u> Simulating time evolution with fully optimized single-qubit gates on parametrized quantum circuits. <u>Phys. Rev. A</u> **105**, 062421 (2022).

25. Watanabe, H. C., Raymond, R., ya Ohnishi, Y., Kaminishi, E. & Sugawara, M. Optimizing parameterized quantum circuits with free-axis selection (2023). `2104.14875`.

26. Crooks, G. E. Gradients of parameterized quantum gates using the parameter-shift rule and gate decomposition (2019). `1905.13311`.

27. Izmaylov, A. F., Lang, R. A. & Yen, T.-C. Analytic gradients in variational quantum algorithms: Algebraic extensions of the parameter-shift rule to general unitary transformations. Phys. Rev. A **104**, 062443 (2021).

28. Yordanov, Y. S., Barnes, C. H. & Arvidsson-Shukur, D. R. Molecular-excited-state calculations with the qubit-excitation-based adaptive variational quantum eigensolver protocol. Phys. Rev. A **106**, 032434 (2022).

29. Fan, Y. et al. Circuit-depth reduction of unitary-coupled-cluster ansatz by energy sorting (2021). `2106.15210`.

30. Nielsen, M. A., Bremner, M. J., Dodd, J. L., Childs, A. M. & Dawson, C. M. Universal simulation of hamiltonian dynamics for quantum systems with finite-dimensional state spaces. Phys. Rev. A **66**, 022317 (2002).

31. Wiersema, R. et al. Exploring entanglement and optimization within the hamiltonian variational ansatz. PRX Quantum **1**, 020319 (2020).

32. Romero, A., Engel, J., Tang, H. L. & Economou, S. E. Solving nuclear structure problems with the adaptive variational quantum algorithm. Phys. Rev. C **105**, 064317 (2022).

33. Grimsley, H. R., Barron, G. S., Barnes, E., Economou, S. E. & Mayhall, N. J. Adaptive, problem-tailored variational quantum eigensolver mitigates rough parameter landscapes and barren plateaus. npj Quantum Inf. **9**, 19 (2023).

34. Pfeuty, P. The one-dimensional ising model with a transverse field. ANNALS Phys. **57**, 79–90 (1970).

35. SáBarreto, F. & Fittipaldi, I. Thermodynamical properties of the transverse Ising model. Phys. A: Stat. Mech. its Appl. **129**, 360–373 (1985).

36. Adjoua, O., Feniou, C. & et al. Sorbonne Université and Qubit Pharmaceuticals (2023).

37. Gacon, J., Zoufal, C., Carleo, G. & Woerner, S. Stochastic approximation of variational quantum imaginary time evolution (2023). `2305.07059`.

38. Kim, Y. et al. Evidence for the utility of quantum computing before fault tolerance. Nature **618**, 500–505 (2023).

39. Abrams, D. S. & Lloyd, S. Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors. Phys. Rev. Lett. **83**, 5162 (1999).

40. Aspuru-Guzik, A., Dutoi, A. D., Love, P. J. & Head-Gordon, M. Simulated quantum computation of molecular energies. Science **309**, 1704–1707 (2005).

41. Yao, Y.-X. et al. Adaptive variational quantum dynamics simulations. PRX Quantum **2**, 030307 (2021).

42. Gomes, N. et al. Adaptive variational quantum imaginary time evolution approach for ground state preparation. Adv. Quantum Technol. **4**, 2100114 (2021).

43. Yordanov, Y. S., Arvidsson-Shukur, D. R. M. & Barnes, C. H. W. Efficient quantum circuits for quantum computational chemistry. Phys. Rev. A **102**, 062612 (2020).