An efficient quantum algorithm for the time evolution of parameterized circuits

Stefano Barison, Filippo Vicentini, and Giuseppe Carleo

Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

We introduce a novel hybrid algorithm to simulate the real-time evolution of quantum systems using parameterized quantum circuits. The method, named "projected - Variational Quantum Dynamics" (p-VQD) realizes an iterative, global projection of the exact time evolution onto the parameterized manifold. In the small time step limit, this is equivalent to the McLachlan s variational principle. Our approach is efficient in the sense that it exhibits an optimal linear scaling with the total number of variational parameters. Furthermore, it is global in the sense that it uses the variational principle to optimize all parameters at once. The global nature of our approach then significantly extends the scope of existing efficient variational methods, that instead typically rely on the iterative optimisation of a restricted subset of variational parameters. Through numerical experiments, we also show that our approach is particularly advantageous over existing global optimisation algorithms based on the time-dependent variational principle that, due to a demanding quadratic scaling with parameter numbers, are unsuitable for large parameterized quantum circuits.

1 Introduction

In recent years, our ability to manipulate and measure quantum systems has improved tremendously. Among all experimental platforms, the number of addressable qubits has increased remarkably: both Google [1] and IBM [2] have reported superconducting circuits chips with ≈ 50 qubits and according to their public timelines they expect to build 1000-qubit devices within the next 2 years. Despite this impressive development, universal quantum-computing remains still far in the future. Algorithms such as Shor's [3], Quantum-Fourier Transform [4] or general Quantum Simulators [5, 6, 7] require a number of operations (gates) at least polynomial in the qubit number. However, in the absence of large-scale quantum error correction, the number of gates that can be applied is at present strongly limited by hardware noise and decoherence.

To circumvent the problem, current generations of quantum algorithms rely on a hybrid classical-quantum approach [8, 9, 10, 11, 12, 13]. A typical example is given by the classical optimisation of a quantum circuit encoding the solution of a given problem. Quantum circuits can be used as the model of a machine-learning problem [14, 15], such as a quantum classifiers [16, 17], quantum kernel machines [18, 19, 20], quantum Boltzmann machines [21] or convolutional networks [22] to name a few. Quantum circuits can also be used to directly approximate the state of an interacting quantum system [23, 24, 25]. In particular, several of the proposed hybrid algorithms extend the use of the variational method to quantum computing: a trial quantum state (ansatz state) with a tractable number of parameters and mild circuit depth is considered, then these parameters are optimized in order to approximate a target state as accurately as possible.

For quantum simulation, several variational, resource friendly alternatives to Trotterization [26, 27, 28, 29] have been proposed [9, 30, 31, 32, 33]. Among those, an interesting approach introduced in 2017 [9] is the The Time-Dependent Variational Algorithm (TDVA). This method, based on a reformulation of the Dirac-Frenkel and McLachlan variational principle [34, 35, 36], encodes the state into a variational circuit and iteratively updates the parameters by solving a Euler-Lagrange equation of motion. The approach is conceptually close to the study of quantum dynamics using classically parameterized many-body wave functions [37, 38, 39, 40]. At variance with the classical many-body counterparts, the computational cost of the TDVA approach is dominated by the stochastic estimation of the Quantum Geometric Tensor [41, 42] and its inversion, resulting in an expensive quadratic cost in the total number of variational parameters. To alleviate this issue, new variational methods based on partial, local optimisations of the variational parameters have been recently proposed [43, 44, 45].

In this paper we propose a hybrid quantum algorithm for approximating the real time evolution of an interacting quantum system, called projected-Variational Quantum Dynamics (p-VQD). This algorithm is both global – it optimizes all parameters at once – and

efficient – it scales linearly with the number of parameters. Moreover, it does not require auxiliary (ancilla) qubits and the depth of the circuit is constant throughout all the simulation. The structure of this paper is as follows: in Section 2 we present the algorithm in order to simulate real time evolution of quantum systems, while in Section 3 we benchmark our algorithm on a Transverse Field Ising model, assessing the accuracy of the method on an ideal quantum simulator performing single and multi shot executions. We conclude, in Section 4, with some considerations on the proposed algorithm.

2 Method

We address the simulation of a quantum system with Hamiltonian $\hat{H}$. For clarity of exposition we will assume that $\hat{H}$ is time independent, but it is not a requirement of the algorithm. The time evolution operator for a small time-step $\delta t \in \mathbf{R}$ is $e^{-i\hat{H}\delta t}$. Let $|\psi_{w(t)}\rangle$ be the parameterized ansatz state approximating the exact quantum state of the system $|\Psi(t)\rangle$ at time t, where $w(t) \in \mathbf{R}^p$ is the vector of its $p-$parameters. We assume $|\psi_{w(t)}\rangle$ is a unit vector and define the evolved state $|\phi(t + \delta t)\rangle = e^{-i\hat{H}\delta t}|\psi_{w(t)}\rangle$. From now on, we will indicate $w(t)$ as w to simplify the notation, implying that the parameters we are referring to are the one assigned to the ansatz at time t, except when explicitly indicated. Also, we will shorten $|\phi(t + \delta t)\rangle$ to $|\phi(\delta t)\rangle$.

In order to variationally approximate the time evolution of the system, we aim to maximize the overlap between the evolved parameterized state $|\phi(\delta t)\rangle$ and the state $|\psi_{w+dw}\rangle$, where $dw \in \mathbf{R}^p$ is a vector of parameter variations. We want to find dw such that

$$\arg\max_{dw\in\mathbb{R}^p} |\langle\phi(\delta t)|\psi_{w+dw}\rangle|^2 \quad . \qquad (1)$$

This intermediate overlap optimization is routinely used in classical simulations of the dynamics of quantum systems, and its application to quantum circuits has been recently suggested in the context of quantum tensor networks [46, 45] and symmetry-preserving variational wavefunction [47]. Similarly to [48], we derive the condition indicated in Eq. (1) defining the projected real-time evolution, as can be seen in Appendix A. The optimal update $dw^\star$ therefore minimizes the step-infidelity

$$L(dw, \delta t) = \frac{1 - |\langle\phi(\delta t)|\psi_{w+dw}\rangle|^2}{\delta t^2}, \qquad (2)$$

where the $\frac{1}{\delta t^2}$ factor has been added to make L independent from the time step size in the limit of $\delta t \to 0$. More details about the introduction of this factor can

be found in Appendix B. Given that the wave-function is encoded as the circuit $C(w)$, substituting into the loss function in Eq. (2) the definition $|\psi_w\rangle = C(w)|0\rangle$ we obtain

$$L(dw, \delta t) = \frac{1 - |\langle 0|C^\dagger(w)e^{i\hat{H}\delta t}C(w+dw)|0\rangle|^2}{\delta t^2}, \quad (3)$$

where the second term on the right-hand side is an expectation value which can be sampled on a quantum computer. Another method that can be considered to measure the overlap between $|\psi_{w+dw}\rangle$ and $|\phi(\delta t)\rangle$ is the SWAP test [49, 50], although it requires ancillary qubits and long range gates. For these reasons, we will not consider this method in the following discussion. The time evolution operator $e^{-i\hat{H}\delta t}$ is encoded in the form of a Trotter-Suzuki decomposition [26, 27].

The optimal $dw^\star$ are determined by iteratively descending along the steepest direction given by the gradient $\frac{\partial}{\partial dw_i}L(dw, \delta t)$, starting from an initial guess dw_0. In general, the gradient of a quantum expectation value involving parameterized quantum circuits can be determined using finite differences [51] or simultaneous perturbation techniques [52]. However, those techniques rely on an approximation of the real gradient. To improve the accuracy, we consider circuits of the general form

$$C(w) = V_p U_p(w_p) V_{p-1} \dots V_1 U_1(w_1) \qquad (4)$$

where the gates V_k do not depend on any parameter and the parameterized gates $U_j(w_j)$ are of the form

$$U_j(w_j) = e^{-\frac{i}{2}w_j G_j} = \cos(w_j/2)\mathbb{I} - i\sin(w_j/2)G_j, \qquad (5)$$

with $G_j^2 = \mathbb{I}$. We remark that in general G_j can be any tensor product of Pauli operators. In this case, the gradient can be computed exactly in a hardware-friendly way using the parameter shift rule [51, 53, 54, 55, 56, 57], obtaining

$$\frac{\partial}{\partial dw_i}L(dw, \delta t) =$$
$$= \frac{L(dw + se_i, \delta t) - L(dw - se_i, \delta t)}{2\sin(s)}, \quad (6)$$

where e_i is the versor in the i-th direction and $s \in \mathbb{R}$. We remark that our p-VQD algorithm does not require an ancilla qubit to perform measurements.

The optimisation of dw is then performed according to a standard gradient descent scheme

$$dw^{\text{new}} = dw^{\text{old}} - \eta\nabla_{dw}L(dw^{\text{old}}, \delta t), \qquad (7)$$

with learning rate $\eta \in \mathbf{R}^+$. A single optimisation step requires $O(p)$ measurements. The optimisation
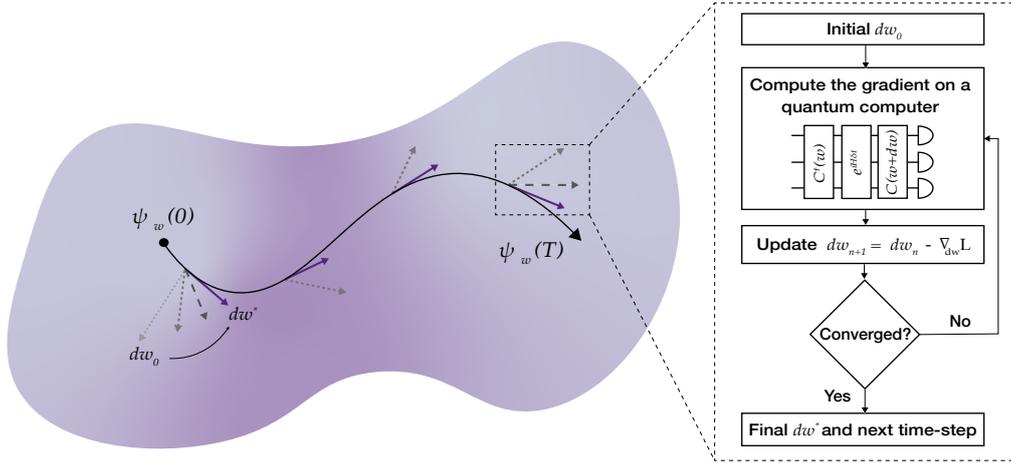
Figure 1: Sketch of the p-VQD algorithm. We follow the real time evolution of the ansatz state $\psi_w(t)$ in the Hilbert space by optimizing the parameter variation dw at every time step. The optimisation is performed through the gradient of the step-infidelity function $L(dw, \delta t)$, computed using a quantum computer.

continues until the loss function goes below the desired threshold ν. Finally, once $dw^\star$ is determined, the parameters at time $t + \delta t$ are obtained by

$$w(t + \delta t) = w(t) + dw^\star(t) \tag{8}$$

We highlight that the circuit width and depth are the same from the beginning to the end of the simulation. A sketch of the algorithm is shown in Fig 1, while an accurate study of the total hardware cost can be found in Section 3.

2.1 Relationship with other methods

In order to understand the connection of our approach with existing methods proposed in the quantum computing setting, it is conceptually interesting to explicitly take the limit of a vanishing time step. In this limit, the parameter updates obtained with p-VQD is $dw = \dot{w}\delta t$ and $\dot{w}$ is the solution of the equation

$$\sum_j \text{Re}\big[G_{kj}\big]\dot{w}_j = \text{Im}\Big[\langle \partial_k \psi_w | H | \psi_w \rangle\Big] +$$
$$+ i\langle \psi_w | H | \psi_w \rangle \langle \partial_k \psi_w | \psi_w \rangle \tag{9}$$

where G_{kj} is the Quantum Geometric Tensor (QGT)

$$G_{kj}(w) = \left\langle \frac{\partial \psi_w}{\partial w_k}, \frac{\partial \psi_w}{\partial w_j} \right\rangle - \left\langle \frac{\partial \psi_w}{\partial w_k}, \psi_w \right\rangle \left\langle \psi_w, \frac{\partial \psi_w}{\partial w_j} \right\rangle . \tag{10}$$

This expression for dw coincides with that given by the McLachlan's variational principle [36, 30] (considering also a time-dependent global phase on the ansatz state). Both McLachlan's and the time-dependent variational principle suffice to simulate real time dynamics of closed systems. Moreover, it is possible to prove that the two approaches are equivalent under certain assumptions [30]. However, on the practical

point of view, there are some differences between them due to numerical instabilities that we will analyse further in Sec. 3. Considering the expression for the QGT given in Eq. (10), the TDVA relies on the evaluation of its imaginary part, which is skew-symmetric, making its inversion unstable when the off-diagonal elements are close to 0. The McLachlan's variational principle, on the contrary, requires the real part, as can be seen in Eq. (9). For this reason, it is suggested that McLachlan's is the most consistent variational principle for quantum simulation [30]. A detailed derivation of Eq. (9) can be found in Appendix C.

Another method conceptually similar to pVQD is the Restarted Quantum Dynamics (RQD), independently proposed in a recent preprint [47]. This technique also solves an optimisation problem to obtain the time-evolved state. However, the RQD is based on the minimisation of the infidelity squared by varying the parameters of the ansatz state, which can lead to different optimisation dynamics.

2.2 Barren plateaus

It has recently been shown that the optimisation of shallow circuits can also be affected by so-called barren plateaus if the cost function is global [58]. This covers the case of operators in the form of

$$O = a\mathbf{I} + cO_1 \otimes O_2 \otimes \cdots \otimes O_i, \tag{11}$$

when O_i is a non-trivial projector ($O_i^2 = O_i \neq \mathbf{I}$) acting on subsystem S_i. The operator $O_G = \mathbf{I} - |0\rangle\langle 0|$ that we use to estimate the cost function in Eq. 2 belongs to this class, meaning that our cost function is global.

Barren plateaus are usually recognized by measuring the variance of the components of the gradient [58,

59, 60, 61, 62]. However, it is possible to show that both the fidelity between two states and the variance of its gradient have a non-vanishing lower bound (in the limit of a large number of qubits) if the two states differ by an infinitesimal transformation [60]. Therefore, p-VQD avoids such plateaus entirely because at every time step we optimise the infidelity between the current state $|\psi_{w(t)}\rangle$ and the infinitesimally evolved state $U(\delta t)|\psi_{w(t)}\rangle$.

A more general approach to avoid barren plateaus consists in replacing the global cost function by one that is local but has the same minimum [58]. The global operator O_G (in our case, encoding the infidelity) can be replaced by its local counterpart

$$O_L = \mathbf{I} - \frac{1}{N}\sum_{j=1}^{N}|0_j\rangle\langle 0_j| \otimes \mathbf{I}_{\bar{j}} \qquad (12)$$

where $\mathbf{I}_{\bar{j}}$ is the identity operator on all qubits except for j. Since $\langle O_L\rangle_w = 0 \Leftrightarrow \langle O_G\rangle_w = 0$, in the limit of small δt the generated dynamics will also be equivalent to the McLachlan's variational principle. In Sec. 3 we discuss some numerical results comparing the two approaches.

3 Results

To demonstrate a practical application of the p-VQD algorithm, we consider the Transverse Field Ising Model on an open chain,

$$H = J\sum_{i=0}^{N-1}\sigma_i^z\sigma_{i+1}^z + h\sum_{i=0}^{N}\sigma_i^x \quad . \qquad (13)$$

The first term accounts for interaction between spins while the latter represents a local and uniform magnetic field along the transverse direction x. For our simulations, we considered $J = \frac{1}{4}$, $h = 1$ and, except when explicitly indicated, $N = 3$ spins. We compare the time-evolution obtained through the p-VQD against the more-established TDV-Algorithm. We analyse the ideal case of a simulation in which we have access to the state vector produced by the quantum circuit (state-vector simulation) and the case in which to gain information about the quantum state we have to repeatedly measure the qubits (multi-shot simulation). The two simulations coincide in the limit of infinite samples. However, when the number of samples is finite, statistical fluctuations produce a noise on the results which we will refer to as shot noise. For both state-vector and multi-shot simulations we use IBM's open-source library for quantum computing, Qiskit [63].

We consider a circuit ansatz of the form

$$C(w) = \prod_{l=1}^{d} c_l(w_l) \qquad (14)$$

$$= \prod_{l=1}^{d}\left[\prod_{i=1}^{N} R_\alpha^{(i)}(w_{i,l})\right]\left[\prod_{j=1}^{N-1} e^{-iw_{j,l}\sigma_j^z\sigma_{j+1}^z}\right], \qquad (15)$$

where $R_\alpha^{(i)}(w_{i,l}) = e^{-iw_{i,l}\sigma_i^\alpha}$ is a single qubit rotation around the $\alpha = \{x, y\}$ axis. Every block $c_l(w_l)$ has a layer of single qubits rotations followed by a layer of entangling two-qubits gates. The total number of blocks, or depth, is d. When $\alpha = x$, a block c_l is equivalent to the Trotter-Suzuki approximation of the unitary operator $e^{-i\hat{H}\delta t}$. A more general parametrization is obtained by alternating rotations around the x and y axis ($\alpha = x$ iff l is odd, $\alpha = y$ otherwise). The representation power and the number of variational parameters can be increased by making the ansatz deeper. For the chosen system, the two ansätze perform similarly when shot noise is neglected. However, alternating blocks of x and y rotations proved to be more stable in presence of shot noise. In the following we will be studying the latter unless otherwise noted.

As a first comparison, we consider the integrated infidelity $\Delta_F(T)$ achieved by both algorithms with respect to the exact simulation of the system over an entire time evolution from $t = 0$ to $t = T$. This quantity can be expressed as

$$\Delta_F(T) = \int_0^T \left(1 - |\langle\Psi(t)|\psi_{w(t)}\rangle|^2\right) dt \quad . \qquad (16)$$

We have performed several simulations both with p-VQD and the TDVA for different shots per circuit evaluation (total number of samples) and report the mean performance in Fig. 2.

At a fixed number of samples, the integrated infidelity Δ_F for the p-VQD is up to an order of magnitude below the TDVA one for the same number of time step. In Appendix C it is shown that with small time steps and in the limit of infinite samples (ideal measures) the parameter variation found by p-VQD coincide with the solution of the Euler-Lagrange equation generated by the McLachlan's variational principle. The matrices of the Euler-Lagrange equations often show high condition number [64], in particular those generated by the time-dependent variational principle [30]. These ill-conditioned matrices are likely to produce large variations on the results when subject to small variations of the coefficients, as those produced by shot noise. We remark that shot noise is unavoidable for quantum computers, even for fault-tolerant devices. Improving iteratively the solution of the Euler-Lagrange equation leads to the estimate of
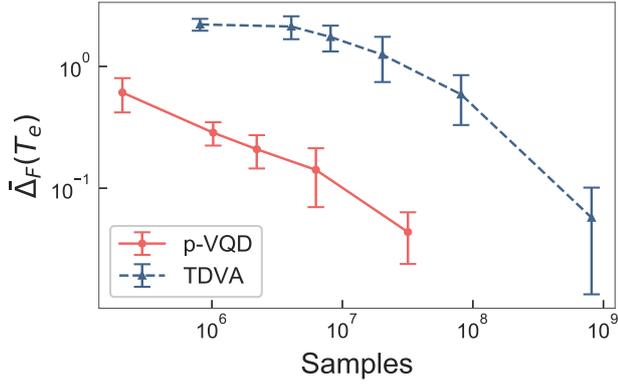
Figure 2: Mean error on fidelity accumulated over an entire time evolution for the two different methods. The plot shows, as a function of the total samples required, the fidelity error accumulated by the algorithm over an entire time evolution. The total time of evolution is $T_e = 3$, the number of time steps is $n_{\text{steps}} = 60$, the number of spins is $N = 3$ and the ansatz chosen is the custom one with $d = 3$ for both p-VQD and TDVA. For each number of samples the experiment has been iterated $n_{\text{iter}} = 10$ times and the data shown represent mean and standard deviation of those values.

the parameters' variation, but does not require matrix inversion. As a result, we obtain a more stable algorithm against shot noise.

In Fig. 3 we compare the expectation values of the total magnetization along the $\hat{x}$ and $\hat{z}$ axis for the states evolved according to the two algorithms. The results are reported for a different number of measurements per circuit.

The magnetization along the $\hat{x}$ axis proved to be the most difficult to capture for both the algorithms, suggesting that the problem could be the ansatz choice. We note that both the algorithms converge to the exact simulation values as the number of shots increases. In general, p-VQD shows comparable results with TDVA using one order of magnitude fewer shots.

To further analyse the performance of our algorithm, we studied the behaviour of the step-infidelity $L(dw, \delta t)$. In Fig. 4 we show the number of optimisation steps required as a function of time and the consequent decrease of the cost function.

We remark that choosing the previous $dw^\star$ as an initial guess often leads to convergence in only one iteration. More details about the convergence of the step-infidelity in a single time step can be found in Appendix D.

We characterise the hardware cost of p-VQD when the number of variational parameters increases. Since the number of measures required by p-VQD scales linearly ($O(p)$) with the number of variational parameters, it has a lower asymptotic cost with respect to TDVA, which is quadratic. We can provide an accurate estimation of the hardware cost of p-VQD as an
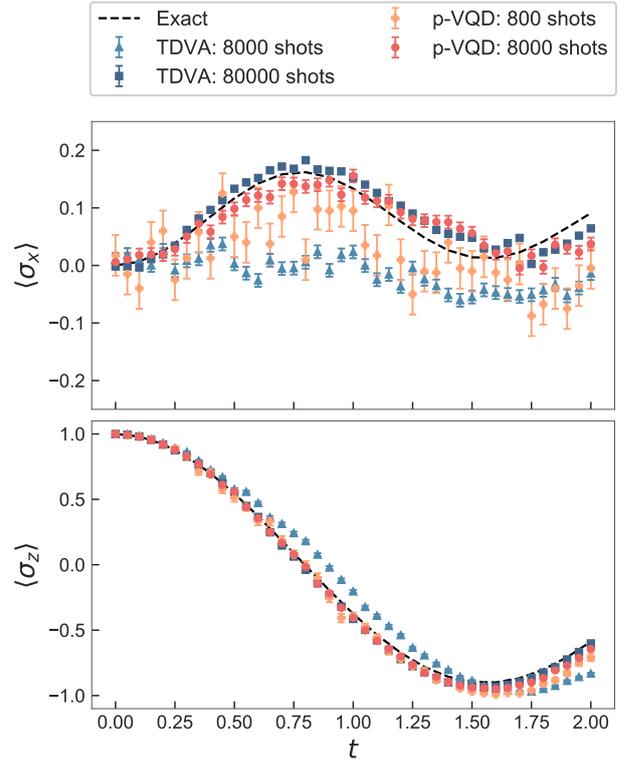

Figure 3: Total magnetization measured on simulated time evolution. The plot shows the expectation values of Pauli operators for systems simulated with p-VQD and TDVA at different shot number. The magnetization is evaluated on a chain of $N = 3$ spins and we used the custom ansatz with $d = 3$ for both the algorithms. We indicated the number of shots for circuit evaluation instead of total samples: $800, 8000$ and 80000 shots are equivalent to $\sim 10^6, 10^7, 10^8$ total samples for the p-VQD and $\sim 10^7, 10^8, 10^9$ for the TDVA, respectively.

upper-bound to the total number of samples needed per simulation. We indicate with n_s and n_t the number of shots and time steps, respectively. In the spirit of the iterative methods, we suppose that the procedure will find a solution $dw^\star$ in at most M steps. In this case, we have that the total number of hardware measurements is

$$N_{\text{samples}} \leq 2Mn_t n_s p \qquad (17)$$

Some terms in Eq 17 are not free parameters: the number of shots n_s depends on the accuracy required, while the number M of optimisation steps depends also on the ansatz and on the Hamiltonian of the system considered. More details about the role of n_s can be found in the Appendix E. To characterise the dependence of M upon the number of parameters p we performed different simulations increasing the depth of our ansatz and reported the results in Fig 5. From the figure we deduce that $M \sim O(1)$, therefore it does not scale up by increasing the number of parameters in the ansatz circuit. Under these circumstances,
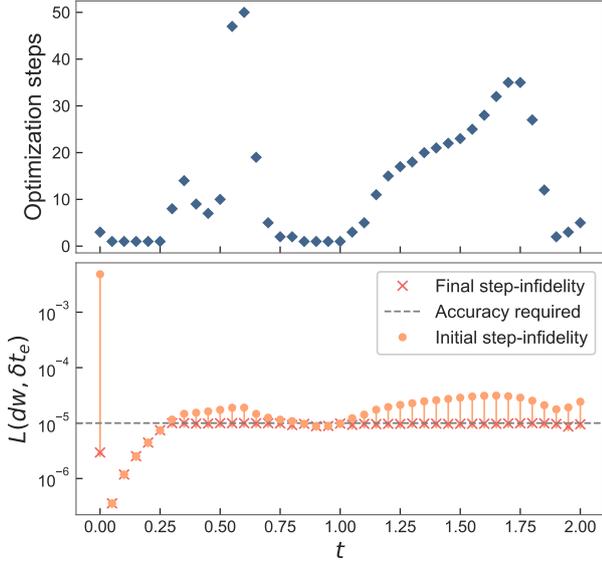
Figure 4: Number of optimisation steps required (top) to decrease accuracy error per step (bottom). The algorithm optimizes the parameters shift dw until the step-in delity $L(dw, \delta t_e)$ is below the threshold at 10^{-5}. The time step chosen is $\delta t_e = 0.05$. The analysis is performed on a state-vector simulation of a system of $N = 3$ spins, using the custom ansatz with depth $d = 3$.

the total dependence of the computational cost of the algorithm upon the number of parameters is indeed $O(p)$.

All the simulations have been done fixing a number of shots n_{shots} per circuit evaluation. For p-VQD, we made multiple simulations and then reported the mean values of our results. For TDVA the number of circuit evaluation is fixed and can be estimated a priori. This comparison does not consider the fidelities of the results obtained with the exact simulation, we know from what we showed in Fig. 2 that p-VQD has a lower error when the total number of samples is comparable with TDVA. We note that the number of samples required scales approximately linearly with the number of parameters, with fluctuations due to different optimisation steps required. The lower the step-infidelity required, the higher will be the number of shots and possibly of the optimisation steps, resulting in greater fluctuations on the total number of samples. In this case, we remark that more advanced methods like the use of an adaptive learning rate η can be used to improve convergence performance of p-VQD.

Finally, we compare numerically the performance of a global and a local cost function. As discussed in Sec. 2, both O_G (Eq. 2) and O_L (Eq. 12) can be used. In Fig. 6 we show the minimisation of both cost functions for different time steps and increasing number of spins $N \in [3, 11]$, obtained using the exact state-vector simulator.
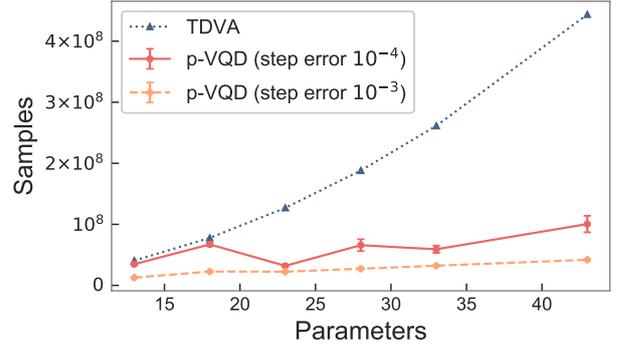

Figure 5: Measurements required as a function of circuit parameters. The plot shows the total number of circuits created and measured using p-VQD and TDVA. The number of required measurements by the TDVA is xed, while in the p-VQD it depends on the optimisation of the parameter variation. We xed the number of spins at $N = 3$ and increased the depth of the custom ansatz from $d = 2$ to $d = 8$; the parameters varied accordingly. We considered $n_{\text{shots}} = 8000$ per circuit evaluation for both methods and performed multiple simulations for p-VQD, showing mean values and standard deviations of the results. Di erent error on the accuracy per step are showed for p-VQD.
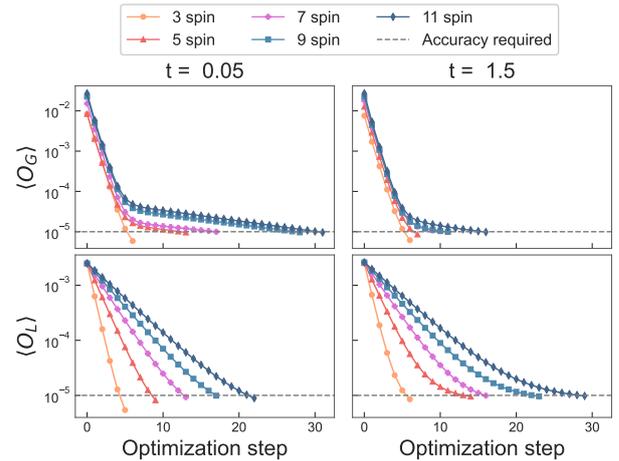

Figure 6: Optimisation steps required per time step when increasing the qubits number. The plot compares the use of the global and the local cost function in the optimisation process of two di erent time steps. The analysis is performed using the state-vector simulator and an accuracy per step of 10^{-5}. The ansatz has $\alpha = x$ at every layer and $d = 3$. The number of parameters varies accordingly. The two cost functions show similar performances in the number of iterations required.

All the curves are obtained by setting the initial condition $dw_0 = 0$. We highlight that our numerical findings point to an efficient scaling with the system size. We remark that the first optimisation steps of O_G depend very weakly on the number of qubits. However, we also note that in general barren plateaus during optimisation may not show up until even larger systems are considered [58, 62]. For this reason, one should benchmark which cost function behaves best for their

specific system and size.

4 Discussion

In this work, we presented a new algorithm for the efficient variational simulation of the real-time evolution of quantum systems. We have shown that it is asymptotically more hardware efficient than the time-dependent variational algorithm (TDVA), while retaining an higher accuracy. Considering the projected time evolution, we avoid numerical instabilities due to the matrix inversion combined with statistical fluctuations due to finite shot measurements. We have numerically investigated the absence of barren plateaus, paving the way towards the simulation of larger quantum systems. One possible application of our approach is to study the dynamical properties of two-dimensional interacting systems, a notoriously difficult problem in classical computational physics. Similarly to all other variational algorithms, the choice of the right parametrization is fundamental for the algorithm to succeed. In this sense, having an efficient quantum algorithm to perform variational time evolution is essential to compare to classical results obtained with variational states either based on tensor networks [65, 66, 67], or neural networks [68, 69]. A possible improvement to further enhance the efficiency of our approach concerns the estimation of the gradient. At present, a drawback of our method is that the circuit constructed on the quantum device is approximatively twice as deep as the ansatz used to represent the system. However, by suitably controlling the number of two-qubits gates in the chosen ansatz, representing the major source of circuit error, we believe that p-VQD can already be used to simulate small quantum systems on available devices.

Data availability

The code used to run the simulations in this article has been written in Python using Qiskit [63]. It is open source and can be found on GitHub [70].

Acknowledgments

References

[1] Frank Arute et al. Quantum supremacy using a programmable superconducting processor . In: *Nature* 574.7779 (Oct. 2019), pp. 505 510. DOI: 10.1038/s41586-019-1666-5.

[2] LeeAnn M. Sager, Scott E. Smart, and David A. Mazziotti. Preparation of an exciton condensate of photons on a 53-qubit quantum computer . In: *Physical Review Research* 2.4 (Nov. 2020). DOI: 10.1103/physrevresearch.2.043205.

[3] P.W. Shor. Algorithms for quantum computation: discrete logarithms and factoring . In: *Proceedings 35th Annual Symposium on Foundations of Computer Science.* IEEE Comput. Soc. Press, 1994. DOI: 10.1109/sfcs.1994.365700.

[4] D. Coppersmith. *An approximate Fourier transform useful in quantum factoring.* June 1994. arXiv: quant-ph/0201067v1 [quant-ph].

[5] Giuseppe E Santoro and Erio Tosatti. Optimization using quantum mechanics: quantum annealing through adiabatic evolution . In: *Journal of Physics A: Mathematical and General* 39.36 (Aug. 2006), R393 R431. DOI: 10.1088/0305-4470/39/36/r01.

[6] Ivan Kassal et al. Polynomial-time quantum algorithm for the simulation of chemical dynamics . In: *Proceedings of the National Academy of Sciences* 105.48 (2008), pp. 18681 18686. ISSN: 0027-8424. DOI: 10.1073/pnas.0808245105.

[7] I. M. Georgescu, S. Ashhab, and Franco Nori. Quantum simulation . In: *Rev. Mod. Phys.* 86 (1 Mar. 2014), pp. 153 185. DOI: 10.1103/RevModPhys.86.153.

[8] Alberto Peruzzo et al. A variational eigenvalue solver on a photonic quantum processor . In: *Nature Communications* 5.1 (July 2014). ISSN: 2041-1723. DOI: 10.1038/ncomms5213.

[9] Ying Li and Simon C. Benjamin. E cient Variational Quantum Simulator Incorporating Active Error Minimization . In: *Phys. Rev. X* 7 (2 June 2017), p. 021050. DOI: 10.1103/PhysRevX.7.021050.

[10] Pauline J. Ollitrault et al. Quantum equation of motion for computing molecular excitation energies on a noisy quantum processor . In: *Physical Review Research* 2.4 (Oct. 2020). ISSN: 2643-1564. DOI: 10.1103/physrevresearch.2.043140.

[11] Mario Motta et al. Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution . In: *Nature Physics* 16.2 (Sept. 2019), 205 210. ISSN: 1745-2481. DOI: 10.1038/s41567-019-0704-4.

[12] M. Cerezo et al. *Variational Quantum Algorithms.* 2020. arXiv: 2012.09265 [quant-ph].

[13] Jacob L. Beckey et al. *Variational Quantum Algorithm for Estimating the Quantum Fisher Information*. Oct. 2020. arXiv: 2010.10488v1 [quant-ph].

[14] Jacob Biamonte et al. Quantum machine learning . In: *Nature* 549.7671 (Sept. 2017), 195 202. ISSN: 1476-4687. DOI: 10.1038/nature23474.

[15] Sima E. Borujeni et al. *Quantum circuit representation of Bayesian networks*. 2020. arXiv: 2004.14803 [quant-ph].

[16] Jonathan Romero, Jonathan P Olson, and Alan Aspuru-Guzik. Quantum autoencoders for e cient compression of quantum data . In: *Quantum Science and Technology* 2.4 (Aug. 2017), p. 045001. DOI: 10.1088/2058-9565/aa8072.

[17] Iordanis Kerenidis et al. q-means: A quantum algorithm for unsupervised machine learning . In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach et al. Vol. 32. Curran Associates, Inc., 2019, pp. 4134 4144. URL: https://proceedings.neurips.cc/paper/2019/file/16026d60ff9b54410b3435b403afd226-Paper.pdf.

[18] Maria Schuld and Nathan Killoran. Quantum Machine Learning in Feature Hilbert Spaces . In: *Phys. Rev. Lett.* 122 (4 Feb. 2019), p. 040504. DOI: 10.1103/PhysRevLett.122.040504.

[19] Maria Schuld et al. Circuit-centric quantum classi ers . In: *Physical Review A* 101.3 (Mar. 2020). DOI: 10.1103/physreva.101.032308.

[20] Vojt ch Havlí ek et al. Supervised learning with quantum-enhanced feature spaces . In: *Nature* 567.7747 (Mar. 2019), pp. 209 212. DOI: 10.1038/s41586-019-0980-2.

[21] Mohammad H. Amin et al. Quantum Boltzmann Machine . In: *Phys. Rev. X* 8 (2 May 2018), p. 021050. DOI: 10.1103/PhysRevX.8.021050.

[22] Iris Cong, Soonwon Choi, and Mikhail D. Lukin. Quantum convolutional neural networks . In: *Nature Physics* 15.12 (Aug. 2019), pp. 1273 1278. DOI: 10.1038/s41567-019-0648-8.

[23] P. J. J. O Malley et al. Scalable Quantum Simulation of Molecular Energies . In: *Phys. Rev. X* 6 (3 July 2016), p. 031007. DOI: 10.1103/PhysRevX.6.031007.

[24] Abhinav Kandala et al. Hardware-e cient variational quantum eigensolver for small molecules and quantum magnets . In: *Nature* 549.7671 (Sept. 2017), 242 246. ISSN: 1476-4687. DOI: 10.1038/nature23879. URL: http://dx.doi.org/10.1038/nature23879.

[25] Bela Bauer et al. Quantum Algorithms for Quantum Chemistry and Quantum Materials Science . In: *Chemical Reviews* 120.22 (Oct. 2020), pp. 12685 12717. DOI: 10.1021/acs.chemrev.9b00829.

[26] H. F. Trotter. On the product of semi-groups of operators . In: *Proc. Amer. Math. Soc.* 10 (1959), pp. 545 551. DOI: 10.1090/S0002-9939-1959-0108732-6.

[27] Masuo Suzuki. General theory of fractal path integrals with applications to many-body theories and statistical physics . In: *Journal of Mathematical Physics* 32.2 (1991), pp. 400 407. DOI: 10.1063/1.529425.

[28] Daniel S. Abrams and Seth Lloyd. Simulation of Many-Body Fermi Systems on a Universal Quantum Computer . In: *Phys. Rev. Lett.* 79 (13 Sept. 1997), pp. 2586 2589. DOI: 10.1103/PhysRevLett.79.2586.

[29] G. Ortiz et al. Quantum algorithms for fermionic simulations . In: *Phys. Rev. A* 64 (2 July 2001), p. 022319. DOI: 10.1103/PhysRevA.64.022319.

[30] Xiao Yuan et al. Theory of variational quantum simulation . In: *Quantum* 3 (Aug. 2019), p. 191. ISSN: 2521-327X. DOI: 10.22331/q-2019-10-07-191.

[31] Cristina Cîrstoiu et al. Variational fast forwarding for quantum simulation beyond the coherence time . In: *npj Quantum Information* 6.1 (Sept. 2020). ISSN: 2056-6387. DOI: 10.1038/s41534-020-00302-0.

[32] Benjamin Commeau et al. *Variational Hamiltonian Diagonalization for Dynamical Quantum Simulation*. 2020. arXiv: 2009.02559 [quant-ph].

[33] Kishor Bharti and Tobias Haug. *Quantum Assisted Simulator*. 2020. arXiv: 2011.06911 [quant-ph].

[34] P. A. M. Dirac. Note on Exchange Phenomena in the Thomas Atom . In: *Mathematical Proceedings of the Cambridge Philosophical Society* 26.3 (1930), 376 385. DOI: 10.1017/S0305004100016108.

[35] Jacov Frenkel. *Wave Mechanics: Advanced General Theory*. London: Oxford University Press, 1934.

[36] A.D. McLachlan. A variational solution of the time-dependent Schrodinger equation . In: *Molecular Physics* 8.1 (1964), pp. 39 44. DOI: 10.1080/00268976400100041.

[37] Jutho Haegeman et al. Time-Dependent Variational Principle for Quantum Lattices . In: *Phys. Rev. Lett.* 107 (7 Aug. 2011), p. 070601. DOI: 10.1103/PhysRevLett.107.070601.

[38] Jutho Haegeman et al. Unifying time evolution and optimization with matrix product states . In: *Phys. Rev. B* 94 (16 Oct. 2016), p. 165116. DOI: 10.1103/PhysRevB.94.165116.

[39] Giuseppe Carleo et al. Localization and Glassy Dynamics Of Many-Body Quantum Systems . In: *Scienti c Reports* 2 (Feb. 2012), p. 243. DOI: 10.1038/srep00243.

[40] Giuseppe Carleo et al. Light-cone e ect and supersonic correlations in one- and two-dimensional bosonic super uids . In: *Phys. Rev. A* 89.3 (Mar. 2014), p. 031602. DOI: 10.1103/PhysRevA.89.031602.

[41] Michael Kolodrubetz et al. Geometry and non-adiabatic response in quantum and classical systems . In: *Physics Reports* 697 (June 2017), 1 87. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2017.07.001.

[42] Marin Bukov, Dries Sels, and Anatoli Polkovnikov. Geometric Speed Limit of Accessible Many-Body State Preparation . In: *Phys. Rev. X* 9 (1 Feb. 2019), p. 011034. DOI: 10.1103/PhysRevX.9.011034.

[43] Marcello Benedetti, Mattia Fiorentini, and Michael Lubasch. *Hardware-e cient variational quantum algorithms for time evolution.* 2020. arXiv: 2009.12361 [quant-ph].

[44] Lucas Slattery, Benjamin Villalonga, and Bryan K. Clark. *Unitary Block Optimization for Variational Quantum Algorithms.* 2021. arXiv: 2102.08403 [quant-ph].

[45] F. Barratt et al. Parallel quantum simulation of large systems on small NISQ computers . In: *npj Quantum Information* 7.1 (May 2021). ISSN: 2056-6387. DOI: 10.1038/s41534-021-00420-3.

[46] Sheng-Hsuan Lin et al. Real- and Imaginary-Time Evolution with Compressed Quantum Circuits . In: *PRX Quantum* 2.1 (Mar. 2021). ISSN: 2691-3399. DOI: 10.1103/prxquantum.2.010342.

[47] Matthew Otten, Cristian L. Cortes, and Stephen K. Gray. *Noise-Resilient Quantum Dynamics Using Symmetry-Preserving Ansatzes.* 2019. arXiv: 1910.06284 [quant-ph].

[48] James Stokes et al. Quantum Natural Gradient . In: *Quantum* 4 (May 2020), p. 269. ISSN: 2521-327X. DOI: 10.22331/q-2020-05-25-269.

[49] Daniel Gottesman and Isaac Chuang. *Quantum Digital Signatures.* May 2001. arXiv: quant-ph/0105032 [quant-ph].

[50] Harry Buhrman et al. Quantum Fingerprinting . In: *Physical Review Letters* 87.16 (Sept. 2001). ISSN: 1079-7114. DOI: 10.1103/physrevlett.87.167902.

[51] Maria Schuld et al. Evaluating analytic gradients on quantum hardware . In: *Phys. Rev. A* 99 (3 Mar. 2019), p. 032331. DOI: 10.1103/PhysRevA.99.032331.

[52] J.C. Spall. Implementation of the simultaneous perturbation algorithm for stochastic optimization . In: *IEEE Transactions on Aerospace and Electronic Systems* 34.3 (1998), pp. 817 823. DOI: 10.1109/7.705889.

[53] K. Mitarai et al. Quantum circuit learning . In: *Physical Review A* 98.3 (Sept. 2018). ISSN: 2469-9934. DOI: 10.1103/physreva.98.032309.

[54] Robert M. Parrish et al. *Hybrid Quantum/-Classical Derivative Theory: Analytical Gradients and Excited-State Dynamics for the Multistate Contracted Variational Quantum Eigensolver.* 2019. arXiv: 1906.08728 [quant-ph].

[55] Gavin E. Crooks. *Gradients of parameterized quantum gates using the parameter-shift rule and gate decomposition.* 2019. arXiv: 1905.13311 [quant-ph].

[56] Andrea Mari, Thomas R. Bromley, and Nathan Killoran. Estimating the gradient and higher-order derivatives on quantum hardware . In: *Physical Review A* 103.1 (Jan. 2021). ISSN: 2469-9934. DOI: 10.1103/physreva.103.012405.

[57] Leonardo Banchi and Gavin E. Crooks. Measuring Analytic Gradients of General Quantum Evolution with the Stochastic Parameter Shift Rule . In: *Quantum* 5 (Jan. 2021), p. 386. ISSN: 2521-327X. DOI: 10.22331/q-2021-01-25-386.

[58] M. Cerezo et al. Cost function dependent barren plateaus in shallow parametrized quantum circuits . In: *Nature Communications* 12.1 (Mar. 2021). ISSN: 2041-1723. DOI: 10.1038/s41467-021-21728-w.

[59] Jarrod R. McClean et al. Barren plateaus in quantum neural network training landscapes . In: *Nature Communications* 9.1 (Nov. 2018). ISSN: 2041-1723. DOI: 10.1038/s41467-018-07090-4.

[60] Tobias Haug and M. S. Kim. *Optimal training of variational quantum algorithms without barren plateaus.* 2021. arXiv: 2104.14543 [quant-ph].

[61] Edward Grant et al. An initialization strategy for addressing barren plateaus in parametrized quantum circuits . In: *Quantum* 3 (Dec. 2019), p. 214. ISSN: 2521-327X. DOI: 10.22331/q-2019-12-09-214.

[62] Carlos Bravo-Prieto et al. *Variational Quantum Linear Solver.* 2020. arXiv: 1909.05820 [quant-ph].

[63] Héctor Abraham et al. *Qiskit: An Open-source Framework for Quantum Computing.* 2019. DOI: 10.5281/zenodo.2562110.

[64] J. Demmel. On condition numbers and the distance to the nearest ill-posed problem. In: *Numerische Mathematik* 51 (1987), pp. 251 289. DOI: 10.1007/BF01400115.

[65] Guifré Vidal. E cient Simulation of One-Dimensional Quantum Many-Body Systems. In: *Physical Review Letters* 93.4 (July 2004), p. 040502. DOI: 10.1103/PhysRevLett.93.040502.

[66] A. J. Daley et al. Time-dependent density-matrix renormalization-group using adaptive e ective Hilbert spaces. In: *Journal of Statistical Mechanics-Theory and Experiment* (Apr. 2004), P04005. DOI: 10.1088/1742-5468/2004/04/P04005.

[67] Steven R. White and Adrian E. Feiguin. Real-Time Evolution Using the Density Matrix Renormalization Group. In: *Physical Review Letters* 93.7 (Aug. 2004), p. 076401. DOI: 10.1103/PhysRevLett.93.076401.

[68] Giuseppe Carleo and Matthias Troyer. Solving the quantum many-body problem with arti cial neural networks. In: *Science* 355.6325 (Feb. 2017), pp. 602 606. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.aag2302.

[69] Markus Schmitt and Markus Heyl. Quantum Many-Body Dynamics in Two Dimensions with Arti cial Neural Networks. In: *Physical Review Letters* 125.10 (Sept. 2020). Publisher: American Physical Society, p. 100503. DOI: 10.1103/PhysRevLett.125.100503.

[70] Stefano Barison et al. *Github repository*. 2021. URL: https://github.com/StefanoBarison/p-VQD.

A The projected real time evolution

We aim to simulate the evolution of a quantum system by acting on the parameters of the variational ansatz that approximates the real state of the system. The parameter variation has to satisfy Eq. (1). We are going to see how this condition can be derived defining the projected real-time evolution. Consider the projector $P_w = |\psi_w\rangle\langle\psi_w|$ and define the squared distance between the evolved parameterized state and its projection on the subspace spanned by $|\psi_{w+dw}\rangle$

$$|||\phi(\delta t)\rangle - P_{w+dw}|\phi(\delta t)\rangle||^2 \quad . \tag{18}$$

To find the best approximation for the evolved state, we want to minimize this distance finding the optimal dw. Imposing this condition we obtain

$$
\begin{aligned}
\underset{dw\in\mathbb{R}^p}{\arg\min} |||\phi(\delta t)\rangle - P_{w+dw}|\phi(\delta t)\rangle||^2 &= \underset{dw\in\mathbb{R}^p}{\arg\min} \left(\langle\phi(\delta t)| - \langle\phi(\delta t)|P_{w+dw} \right)\left(|\phi(\delta t)\rangle - P_{w+dw}|\phi(\delta t)\rangle \right) \\
&= \underset{dw\in\mathbb{R}^p}{\arg\min} \left[1 - 2\langle\phi(\delta t)|P_{w+dw}|\phi(\delta t)\rangle + \langle\phi(\delta t)|P_{w+dw}^2|\phi(\delta t)\rangle \right] \\
&= \underset{dw\in\mathbb{R}^p}{\arg\min} \left[1 - \langle\phi(\delta t)|P_{w+dw}|\phi(\delta t)\rangle \right] \\
&= \underset{dw\in\mathbb{R}^p}{\arg\max} |\langle\phi(\delta t)|\psi_{w+dw}\rangle|^2
\end{aligned}
\tag{19}
$$

where we used the idempotence property of the projector $P_{w+dw}^2 = P_{w+dw}$. With the final equivalence we see that this condition is equivalent to Eq. (1).

B Relationship of step-in delity with time step and parameter variation

In this Appendix we explain the introduction of the factor $\frac{1}{\delta t^2}$ in Eq. (2). Consider the overlap contained in the step-infidelity definition in Eq. (2) and Taylor expand it in δt and dw. First, expand to the second order the time evolution operator $e^{-iH\delta t} \sim \mathbf{I} - iH\delta t - \frac{H^2}{2}(\delta t)^2$ to get

$$\langle\psi_{w+dw}|(\mathbf{I} - iH\delta t - \frac{H^2}{2}(\delta t)^2)|\psi_w\rangle\langle\psi_w|(\mathbf{I} + iH\delta t - \frac{H^2}{2}\delta t^2)|\psi_{w+dw}\rangle =$$

$$= |\langle\psi_{w+dw}|\psi_w\rangle|^2 - i\delta t\langle\psi_{w+dw}|H|\psi_w\rangle\langle\psi_w|\psi_{w+dw}\rangle + i\delta t\langle\psi_{w+dw}|\psi_w\rangle\langle\psi_w|H|\psi_{w+dw}\rangle + (\delta t)^2|\langle\psi_{w+dw}|H|\psi_w\rangle|^2 -$$

$$- \frac{\delta t^2}{2}\langle\psi_{w+dw}|H^2|\psi_w\rangle\langle\psi_w|\psi_{w+dw}\rangle - \frac{\delta t^2}{2}\langle\psi_{w+dw}|\psi_w\rangle\langle\psi_w|H^2|\psi_{w+dw}\rangle \quad (20)$$

Now we expand to first order in dw, obtaining $|\psi_{w+dw}\rangle = |\psi_w\rangle + \sum_j dw_j|\partial_j\psi_w\rangle$. The first order contribution in δt vanishes, but also the first order contribution in dw, then we have

$$\sum_j dw_j \left[\langle\partial_j\psi_w|\psi_w\rangle - \langle\partial_j\psi_w|\psi_w\rangle\right] = 0 \qquad \forall dw_j \quad (21)$$

The final result reads

$$|\langle\psi_{w+dw}|e^{-iH\delta t}|\psi_w\rangle|^2 \sim 1 + \delta t^2\Bigg[\sum_j \frac{dw_j}{\delta t}[2i\langle H\rangle_w\langle\partial_j\psi_w|\psi_w\rangle - 2\mathrm{Im}[\langle\psi_w|H|\partial_j\psi_w\rangle]] -$$

$$- \mathrm{Var}_w(H) + \sum_{j,k} \frac{dw_j}{\delta t}\frac{dw_k}{\delta t}\langle\partial_j\psi_w|\psi_w\rangle\langle\psi_w|\partial_k\psi_w\rangle\Bigg] \quad . \quad (22)$$

As $\delta t \to 0$, also $dw_j \to 0 \quad \forall j$ and their ratio remains constant. Therefore, the addition of $\frac{1}{\delta t^2}$ factor in Eq. (2) makes it independent of the time step size in the limit of $\delta t \to 0$.

C Equivalence between the p-VQD and the McLachlan s variational principle

In this Appendix, we show that in the limit of small time step the parameter variation that fulfils the request in Eq. (1) is the same obtained by applying the McLachlan's variational principle [36, 30].

We start Taylor expanding the overlap to the second order: using the substitution

$$|\psi_{w+dw}\rangle = |\psi_w\rangle + \sum_k dw_k|\partial_k\psi_w\rangle + \frac{1}{2}\sum_{k,j} dw_k dw_j|\partial_k\partial_j\psi_w\rangle + o(dw^3) \quad (23)$$

we obtain

$$|\langle\phi(\delta t)|\psi_{w+dw}\rangle|^2 = |\langle\phi(\delta t)|\psi_w\rangle|^2 + \sum_k \Bigg[\langle\phi(\delta t)|\psi_w\rangle\langle\partial_k\psi_w|\phi(\delta t)\rangle + \langle\phi(\delta t)|\partial_k\psi_w\rangle\langle\psi_w|\phi(\delta t)\rangle\Bigg]dw_k +$$

$$+ \sum_{k,j}\Bigg[\langle\phi(\delta t)|\partial_k\psi_w\rangle\langle\partial_j\psi_w|\phi(\delta t)\rangle + \frac{1}{2}\langle\phi(\delta t)|\partial_k\partial_j\psi_w\rangle\langle\psi_w|\phi(\delta t)\rangle + \langle\phi(\delta t)|\psi_w\rangle\langle\partial_k\partial_j\psi_w|\phi(\delta t)\rangle\Bigg]dw_k dw_j + o(dw^3)$$

$$(24)$$

where we used $|\phi(\delta t)\rangle = e^{-iH\delta t}|\psi_w\rangle$ as in the main text.

Then, we expand the time evolution operator to the first order $e^{-iH\delta t} = \mathbf{I} - iH\delta t + o(\delta t)$ and partially differentiate both sides of the normalization condition $||\psi_w||^2 = 1$ with respect to parameters w_i and w_j to obtain the two important properties

$$\langle\psi_w|\partial_k\psi_w\rangle = -\langle\partial_k\psi_w|\psi_w\rangle \quad (25)$$

$$\langle\psi_w|\partial_k\partial_j\psi_w\rangle + \langle\partial_k\partial_j\psi_w|\psi_w\rangle = -\langle\partial_k\psi_w|\partial_j\psi_w\rangle - \langle\partial_j\psi_w|\partial_k\psi_w\rangle \quad . \quad (26)$$

Substituting in Eq. (24) we have

$$|\langle\phi(\delta t)|\psi_{w+dw}\rangle|^2 = |\langle\phi(\delta t)|\psi_w\rangle|^2 +$$
$$+ i\sum_k \left[\langle\psi_w|H|\psi_w\rangle\langle\partial_k\psi_w|\psi_w\rangle - \langle\psi_w|H|\psi_w\rangle\langle\psi_w|\partial_k\psi_w\rangle + \langle\partial_k\psi_w|H|\psi_w\rangle - \langle\psi_w|H|\partial_k\psi_w\rangle\right]dw_k\delta t +$$
$$+ \frac{1}{2}\sum_{k,j}\left[2\langle\psi_w|\partial_k\psi_w\rangle\langle\partial_j\psi_w|\psi_w\rangle - \langle\partial_k\psi_w|\partial_j\psi_w\rangle - \langle\partial_j\psi_w|\partial_k\psi_w\rangle\right]dw_k dw_j \quad (27)$$

where we neglected the third order contribution in dw and δt. We notice that the second order term in $dw_k dw_j$ is the real part of the Quantum Geometric Tensor (QGT) , as expressed in Eq. (10). Finally, in the p-VQD algorithm we aim to find the dw that maximizes the overlap in Eq.(1) , thus we impose the first order optimality condition

$$\frac{\partial}{\partial dw_k}|\langle\phi(\delta t)|\psi_{w+dw}\rangle|^2 = 0 \quad \forall k \tag{28}$$

that in the limit $\delta t \to 0$ gives the equation

$$\sum_j \text{Re}\big[G_{kj}\big]\dot{w}_j = \text{Im}\left[\langle\partial_k\psi_w|H|\psi_w\rangle\right] + i\langle\psi_w|H|\psi_w\rangle\langle\partial_k\psi_w|\psi_w\rangle \tag{29}$$

where we indicated with G the QGT, as in the main text. This is the same evolution equation for parameters w that is obtained when the McLachlan's variational principle is used to simulate real time dynamics of closed systems with pure quantum states. For an extensive review of the variational principles for time evolution of pure and mixed states, see [30].

D   Optimisation routine for a single time-step

In this Appendix, we further analyse the optimisation routine of our algorithm. This optimisation is performed on the parameter variation dw for every time step. In this case we focus on a state-vector simulation of the Transverse Field Ising Model on an open chain (see main text), considering only a single time step (without loss of generality, the third time step of the simulation).
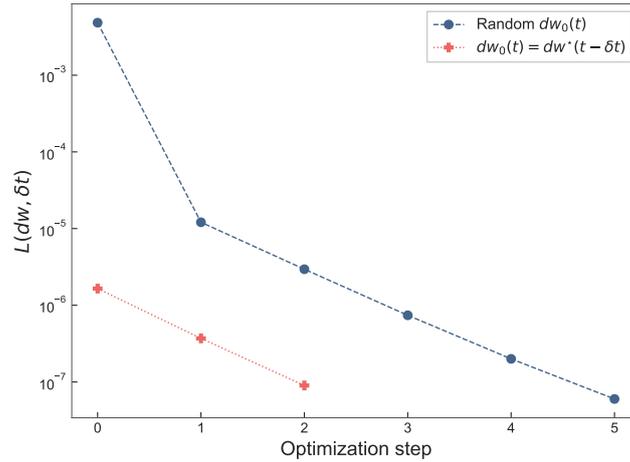

Figure 7: Infidelity as a function of the optimisation step. The optimisation procedure requires only a few optimisation steps to greatly decrease the step-infidelity. Moreover, we compare the random choice of an initial dw_0 of order $O(\delta t)$ to the educated guess $dw_0 = dw^\star(t - \delta t)$.

As Fig. 7 shows, at every optimisation steps the step-infidelity is reduced of nearly a order of magnitude. The initial choice $dw_0 = dw^\star(t - \delta t)$ is shown to reduce the number of steps required.

E Relationship between step in delity and number of shots

In this Appendix, we investigate the relationship between the step infidelity to reach in a p-VQD step and the number of shots n_s required. Given that the fidelity is equivalent to the probability of measuring a string of 0s, with n_s shots the minimum non-zero infidelity we can resolve is $\frac{1}{n_s}$. However, we stress that since we are optimising the infidelity by measuring the gradient, it is possible to converge to a state with a smaller step infidelity than this bound, depending on the variance of the gradient and on the optimisation algorithm employed.

To analyse the effects of the shots on the infidelity optimisation we considered a single time step of the algorithm and optimised the ansatz parameters using a gradient sampled with n_s shots, while measuring the infidelity exactly using the state-vector. For each number of shots n_s, we optimised the step infidelity and measured its convergence value. Due to the presence of shot noise, we repeated the optimisation multiple times and plotted the mean and the standard deviation of those calculations in Fig. 8. We performed the experiment using two different optimizers, Stochastic Gradient Descent and Adam, to study how this choice affects the infidelity minimisation. We also fit the points to a function $f = k n_s^\gamma$ with k and γ free parameters, to estimate the exponent γ. The expected n_s^{-1} dependence is shown as a visual guide for clarity.
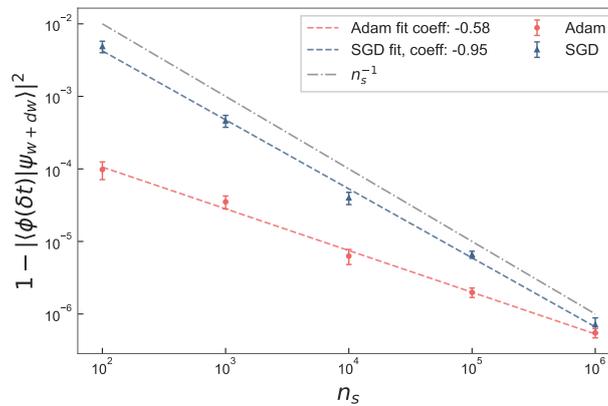

Figure 8: Asymptotic in delity reachable using shots to sample the gradient. This plots illustrates the in delity between the target and the ansatz state as the number of shots to sample the gradient increases. The values reported are the mean and the standard deviations of multiple runs made both with the Stochastic Gradient Descent and Adam as optimizers. The dashed lines are obtained tting the function $f = k n_s^\gamma$. The grey dashdotted line indicates the expected behaviour n_s^{-1}. In this simulation we considered $N = 3$, $d = 3$ and $M = 150$.

We can see that, when a small number of shots is considered, the Adam optimizer is able to minimise the infidelity up to two order of magnitude below the minimum resolvable with shots. When the number of shots increases, the performances of the two optimisers become comparable.