Self-learning eigenstates with a quantum processor

Andrey Kardashin,1 Alexey Uvarov,1 Dmitry Yudin,1 and Jacob Biamonte1

1Deep Quantum Laboratory, Skolkovo Institute of Science and Technology, Moscow 121205, Russia

(Dated: June 25, 2020)

Solutions to many-body problem instances often involve an intractable number of degrees of freedom and admit no known approximations in general form. In practice, representing quantum-mechanical states of a given Hamiltonian using available numerical methods, in particular those based on variational Monte Carlo simulations, become exponentially more challenging with increasing system size. Recently quantum algorithms implemented as variational models, have been proposed to accelerate such simulations. The variational ansatz states are characterized by a polynomial number of parameters devised in a way to minimize the expectation value of a given Hamiltonian, which is emulated by local measurements. In this study, we develop a means to certify the termination of variational algorithms. We demonstrate our approach by applying it to three models: the transverse field Ising model, the model of one-dimensional spinless fermions with competing interactions and the Schwinger model of quantum electrodynamics. By means of comparison, we observe that our approach shows better performance near critical points in these models. We hence take a further step to improve the applicability and to certify the results of variational quantum simulators.

Introduction.—Experimental advances have fostered the development of mid-sized quantum simulators—realizing prototypes of ideas dating back to celebrated proposals by Feynman and others [1–7]. Indeed, controllable quantum simulators emulate classes of Hamiltonians—mimicking Hamiltonian properties to replace traditional numerical methods [8–12]. The difficulty of numerical simulations of interacting quantum systems has resulted in advanced numerical methods, including variational quantum Monte Carlo methods [13] as well as different realizations of the renormalization group routine [14]. In limiting cases, these methods suffer from the exponential slowdown (and/or exponential memory overhead) with the size of a system.

Multiqubit quantum circuits can implement the so-called variational model of quantum computation [15–17], which extends certain methods of machine learning [18–20]. In the variational quantum circuits approach, one relies on an iterative control loop. A quantum state is prepared and measured: the measurement outcome(s) are used to prepare increasingly more optimal states with respect to minimization of a given objective function (given as a Hamiltonian). Variational algorithms emerged as a practically viable application of quantum computers with several dozen qubits and short decoherence times which would otherwise preclude the use of more traditional quantum algorithms [15–17, 21]. The results of measurement are used in a classical optimisation routine to update the prepared state so as to minimize an externally calculated objective function. The process is iterated and the states are prepared by varying over a family of low-depth circuits.

Although experimental realizations of variational algorithms [21, 22] were reported in the recent years [23, 24], theoretical estimates of their efficiency [17] are largely lacking. A particular example, the variational quantum eigensolver (VQE), prepares a family of states characterized by a polynomial number of parameters and minimizes the expectation value of a given Hamiltonian within this family [15, 25–27]. The key idea of VQE is based on decomposing the Hamiltonian into a sum of Pauli strings, i.e., tensor products of Pauli matrices, provided that each Pauli string can be measured separately on the quantum device. VQE can be applied to find ground states of small molecules and interacting spin systems [25, 27]. Scaling of such an approach could access simulations that are not possible to evaluate explicitly using traditional numerical methods; for example owing to the lack of memory or computational resources.

The performance of VQE crucially depends on the choice of the ansatz state. Typically, a common approach is to represent a rather cumbersome quantum state in terms of a variational state and estimate approximation quality, i.e., to explore how close the obtained solution is to the ground state of a given Hamiltonian. Knowing an exact solution drastically simplifies the analysis, otherwise the proximity to the global minimum cannot be guaranteed. Generally, minimization of Hamiltonians is QMA-hard, whereas its restriction to Ising spins is NP-hard. Lately, a way to estimate the quality of the solution by measuring the variance of the energy has been proposed in Ref. [28].

In the scope of this paper, we propose an alternative approach by simulating the Hamiltonian evolution. We clearly demonstrate that in this scenario the number of measurements can be dramatically reduced. We consider the two competing criteria as optimization problems on their own, aside from the VQE problem. We compare the convergence of the two algorithms and clarify the limits of applicability of our method, with a special focus on connection between computational complexity of Hamiltonians and the properties of their eigenstates that are parametrized in terms of the hardware efficient ansatz.

* andrey.kardashin@skoltech.ru
Variational eigenvector search.—The problem we solve is somewhat complementary to VQE. Given a Hamiltonian, defined by its Hermitian matrix, find an eigenvector of this Hamiltonian. The apparent simplicity of determining the eigenvectors of a given matrix nevertheless obscures its computational complexity. It can be done either by means of exact diagonalization, e.g., leveraging Lanczos algorithm, or variational-ansatz-based simulations, both being computationally demanding [28]. We will fix the standard computational basis throughout.

Consider the problem of finding an eigenvector of a Hermitian matrix using a variational quantum algorithm approach. VQE, at its core, relies on preparing an ansatz state $|\psi(\theta)\rangle$ by applying an adjustable sequence of quantum gates $U(\theta)$ to the quantum register $|0\rangle^\otimes n$ of $n$ qubits and sampling the expectation value of a given matrix $\mathcal{H}$ relative to this state. This is followed by a classical optimizer to minimize the energy, $\langle \psi(\theta) | \mathcal{H} | \psi(\theta) \rangle$. The circuit is parametrized by $\theta \in [0, 2\pi)^{\otimes p}$ with $p$ being the number of parameters. Assume that within VQE our best guess is $|\psi\rangle \equiv |\psi(\theta)\rangle$. In Ref. [23], to quantify the accuracy of the variational solution $|\psi\rangle$ it was proposed to employ the mean squared deviation, $\delta = \langle \hat{\mathcal{H}}^2 \rangle - \langle \hat{\mathcal{H}} \rangle^2$ (note that we make use of notation $\langle A \rangle = \langle |A| \psi \rangle$ below). In fact, let the eigenenergy $\lambda_0$ be the closest to the initial trial $\mathcal{E}$, then the energy error is upper bound by $\delta$,

$$|\mathcal{E} - \lambda_0| \leq \sqrt{\delta}. \tag{1}$$

The vector $|\psi\rangle$ is an eigenvector of the Hermitian $\mathcal{H}$ if and only if the mean squared variance $\delta$ is zero. Alternatively, a unitary matrix $Q = e^{-i\mathcal{H}t}$, possesses eigenvalues lying on the unit circle, so that $|\psi\rangle$ is an eigenvector of $\mathcal{H}$ as long as $\langle |\psi| Q |\psi\rangle = 1$.

In numerical simulations, we choose the unitary $U(\theta)$ to be parametrized in terms of three-layered hardware-efficient ansatz as depicted in Fig. 1. By construction, the hardware efficient ansatz — first introduced in Ref. [20] — consists of an array of universal one-qubit gates and an entangling block. The universal one-qubit gates are represented in the $X$-$Z$ decomposition while the entangling block is composed of subsequent controlled $Y$ rotations. The $m$-layered $n$-qubit ansatz would have $4mn$ parameters for $n \geq 3$. In this study, we use a four-qubit ansatz with $m = 3$ layers, and therefore, 48 free parameters.

The parameters $\theta$ are updated by means of the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [29], which is a gradient-based method that uses an approximation of the Hessian matrix.

Model systems.—In the following, we address the convergence properties of physically relevant systems. We consider a one-dimensional quantum Ising chain of $n$ spins, which corresponds to the number of qubits,

$$\mathcal{H}_{\text{TFIM}} = \sum_{j=1}^{n} (Z_j Z_{j+1} + h X_j), \tag{2}$$

in the presence of transverse magnetic field $h \mathcal{I}$ with $J$ specifying the strength of exchange interaction [20]. Note that $(1, X_j, Y_j, Z_j)$ stands for the vector of Pauli matrices at the $j$-th site equipped with a $2 \times 2$ unity matrix. In the thermodynamic limit, $n \to \infty$, the system undergoes the phase transition from a collinearly ordered to a disordered phase at $h = 1$, which will be discussed in the follow-up analysis. Quite interestingly, recent analysis based on neural networks machinery in the form of single [31] and multilayer perceptron [32] demonstrated its efficiency in studying phase transition for the model of Eq. 2.

Likewise, we examine our method to find the eigenstate of the massive Schwinger Hamiltonian,

$$\mathcal{H}_{\text{Sch}} = \sum_{j=1}^{n} \left[ \sigma_j^+ \sigma_{j+1}^- + \sigma_j^- \sigma_{j+1}^+ + m_c \frac{1}{2} (-1)^j Z_j + L_j^2 \right], \tag{3}$$

provided that $L_j = -\frac{1}{2} \sum_{i=1}^{j} [Z_i + (-1)^i \mathbb{1}_i]$ and $\sigma_j^\pm = (X_j \pm i Y_j) / 2$. The model [3] has remained in the focus of research activity as it allows one to capture intriguing properties of quantum chromodynamics. In a nutshell, the Schwinger model represents quantum electrodynamics in two-dimensional space-time [33] and can be addressed in a seemingly related approach of matrix product states, see e.g. [34]. In our simulations, we put $m_c = -0.7$ that corresponds to criticality of this model.

Finally we consider a system of one-dimensional spinless fermions with competing interactions,

$$\mathcal{H}_{\text{el}} = -t \sum_{\langle i,j \rangle} c_i^\dagger c_j + U_1 \sum_{j=1}^{n} n_j n_{j+1} + U_2 \sum_{j=1}^{n} n_j n_{j+2}, \tag{4}$$

FIG. 1. Hardware efficient ansatz for 4 qubits with $R_\sigma(\theta) = e^{-i\theta \sigma}$, $\sigma \in \{X, Y, Z\}$. The entire layer can be repeated several times as needed.
where $n_j = c_j^\dagger c_j$ is the number of electrons at $j$-th site and summation over nearest neighbors $\langle \ldots \rangle$ is implied. In this model, $t$ is the hopping energy, while $U_1$ and $U_2$ stand for matrix elements of Coulomb repulsion between electrons residing on two neighboring and next-neighboring sites respectively. The model \cite{1} represents a versatile still rather simple playground to study effects of frustration in interacting systems \cite{35–38}. With the fixed ratio $U_1/U_2 = 2$, this model is expected to exhibit a metallic behavior \cite{35}. In contrast to the models of Eqs. (2) and (3), the Hamiltonian of interacting electrons \cite{1} is written in terms of second-quantized operators fermionic annihilation ($c_j$) and creation ($c_j^\dagger$) operators, which requires spin-fermion mapping to be implemented. We utilize the Jordan-Wigner transformation to represent these operators as

$$c_j^\dagger = \left( \prod_{k=1}^{j-1} Z_k \right) \otimes \sigma_j^+, \quad c_j = \left( \prod_{k=1}^{j-1} Z_k \right) \otimes \sigma_j^-.$$  

**Cost function.** —To make a direct comparison with the results of the previous studies \cite{23} and discuss the range of applicability of our method we consider the performance of two cost functions determined by

$$F_H(\psi) = \langle H^2 \rangle - \langle H \rangle^2,$$

$$F_Q(\psi) = 1 - |\langle Q \rangle|^2,$$

respectively. Notably, both functions return zero if and only if $|\psi\rangle$ is an eigenstate of the matrix $H$. To control the efficiency of both methods we apply the gain characteristic as a quantitative measure \cite{59},

$${\mathcal G} = \mathbb{E}_F \left[ \frac{F_{\text{conv}} - F_{\text{init}}}{F_{\text{opt}} - F_{\text{init}}} \right],$$

representing the mean variance of the cost function over all instances and written in terms of the value of the objective function at the start of optimization ($F_{\text{init}}$) and at the end of convergence ($F_{\text{conv}}$), as well as the optimal value of ($F_{\text{opt}}$), i.e., global minimum or maximum. Likewise, we elaborate on gain of the overlap $O$ between the variational state $|\psi\rangle$ and an exact eigenstate $|\phi\rangle$ of a target Hamiltonian, i.e.,

$$O(\theta) = |\langle \psi | \phi \rangle|^2.$$  

It is worth noting that we measure the performance of the functions $F_H$ and $F_Q$ by the convergence rate, i.e. the percent of problem instances which converged to the values of the overlap less than or equal to some $\gamma \in [0, 1]$. In our numerical experiments, we set $\gamma = 0.999$.

The plots of the overlap gains and convergence rates for the Hamiltonians of Eqs. (2)–(4) are shown in Fig. 3. As can be visually confirmed, the solution converges suboptimally for both cost functions, but $F_Q$ has a bit higher efficiency in finding the eigenstates of the considered Hamiltonians. Note that $F_Q$ is more suitable in dealing with the physical Hamiltonians, specified by Eqs. (2)–(4), at criticality. Particularly, this is justified by addressing the dependence of optimization performance on the value of the parameter $h$ in the Ising Hamiltonian. It is clearly visible in Fig. 3 that both functions $F_H$ and $F_Q$ begin to perform better after $h = 0.7$ corresponding to highly correlated state(s). However, at $h = 1.1$ the efficiency of $F_H$ drops significantly. On the other hand, the function $F_Q$ exhibits a decreased performance for $h = 1.2, 1.3, 1.4$.

**Evolution time as a hyperparameter.** —The authors of \cite{40} showed the importance of tuning the hyperparameters of different optimizers applied for solving various problems. Since for the objective function $F_Q$ we can control the evolution time $t$, we could use it as a hyperparameter, making the idea of using $t$ in such a way for the function $F_Q$ to outperform the function $F_H$ in terms of overlap gain or convergence rate viable. For certain types of Hamiltonians discussed above there exists $t$ which gives the best performance for $F_Q$. However, as discussed in the Supplemental Material, we did not find any considerable benefit from adjusting the evolution time.

**Discussion and conclusion.** —The two methods based on minimizing objective functions (5) and (6) have substantially different resource costs. To estimate the variance, one has to know both $\langle H \rangle$ and $\langle H^2 \rangle$. Evaluating $\langle H \rangle$ on a quantum processor requires decomposing a given Hamiltonian $H$ into the sum of Pauli strings,

$$H = \sum J_{\alpha\beta...\gamma}^i \sigma_{\alpha}^i \sigma_{\beta}^j \ldots \sigma_{\gamma}^k,$$

and calculating the expectation value of each term separately. In Eq. (9) upper indices of the real-valued tensor $J$ denote the qubit number, while the lower indices stand for a specific Pauli operator $\sigma \in \{I, X, Y, Z\}$. Let us then assume that we need $m$ measurements per a Pauli string to achieve predetermined accuracy. If $H$ contains $k$ Pauli terms, then $H^2$ contains $k^2$ terms at worst. Thus, we need to run the preparation and measurement circuit about $m(k^2 + k)$ times. This number may be decreased by a smart choice of measurements provided commuting Pauli strings are evaluated simultaneously \cite{41,42}. For the needs of quantum chemistry, this approach reduces the number of measurements by an order of $n$, the number of qubits. The number $m$ also has to scale with the number of terms. For one term, the error scales with $1/\sqrt{m}$, so that $k$ terms would add up to $k/\sqrt{m}$. Thus, to keep the error value fixed, $m$ must scale with the number of Pauli strings, making the number of measurements to be of the order of $k^4/n$. If we assume that $k$ is at least linear with $n$, the total number of measurements scales as $\mathcal{O}(n^3)$ versus the number of Pauli strings.

The second method applied to $Q$, on the other hand, requires performing only one set of measurements. The downside is that the quantum circuit is at least twice as deep as that for the variance estimator. On top of that, one needs to be able to implement the Hamiltonian evolution. In the gate model of quantum computation, this can be done using the Suzuki–Trotter formula, which introduces its own error. This means that this technique
should require fewer measurements, but also a higher degree of gate fidelity. We also note that in order to apply
the first method in VQE one has to know the decomposition of the target Hamiltonian. At the same time, the
second method offers greater utility in the sense that the unitary Q can be given as a black box quantum circuit.

Finally, using these criteria as optimization targets on their own can be helpful for VQE as well. If the solution
gets close to some eigenstate, but this state is known not to be the ground state, one can minimize the eigenvalue
criteria to get close to that state, and then exclude that eigenstate from the search by penalizing overlap with it.

To summarize, using hybrid quantum-classical algorithms remains one of the most promising applications
of near-term quantum computers. Within such an approach one executes as much calculations as possible with
classical hardware. VQE is one of the most reliable ways of finding the lowest energy eigenstate of a given matrix.
It was recently proposed to make use of mean square deviation to quantify the accuracy of the VQE. In the
meantime, such an approach seems to be computationally heavy. In this paper, we proposed a way around this
with an objective function which is determined by the evolution operator, or more specifically, a one-parametric
unitary group.

Acknowledgments.—AK and JB acknowledge support from the research project, *Leading Research Center on Quantum Computing* (Agreement No. 014/20). The work of AU and DY was supported by the Russian Foundation for Basic Research Project No. 19-31-90159.

---

[1] R. P. Feynman, *Int. J. Theor. Phys.* **21**, 467 (1982)
[2] S. Lloyd, *Science* **273**, 1073 (1996)
[3] I. Buluta and F. Nori, *Science* **326**, 108 (2009)
[4] K. L. Brown, W. J. Munro, and V. M. Kendon, *Entropy* **12**, 2268 (2010)
[5] P. Hauke, F. M. Cucchietti, L. Tagliacozzo, I. Deutsch, and M. Lewenstein, *Rep. Prog. Phys.* **75**, 082401 (2012)
[6] T. Schaeetz, C. R. Monroe, and T. Esslinger, *New J. Phys.* **15**, 085009 (2013)
[7] I. M. Georgescu, S. Ashhab, and F. Nori, *Rev. Mod. Phys.* **86**, 153 (2014)
[8] M. F. Parsons, A. Mazurenko, C. S. Chiu, G. Ji, D. Greif, and M. Greiner, *Science* **353**, 1253 (2016)
[9] M. J. Brenner, A. Montanaro, and D. J. Shepherd, *Phys. Rev. Lett.* **117**, 080501 (2016)
[10] X. Gao, S.-T. Wang, and L.-M. Duan, *Phys. Rev. Lett.* **118**, 040502 (2017)
[11] J. Bermejo-Vega, D. Hangleiter, M. Schwarz, R. Raussendorf, and J. Eisert, *Phys. Rev. X* **8**, 021010 (2018)
[12] M. Giuza, T. Schweigler, B. Rauer, C. Krumnow, J. Schmiedmayer, and J. Eisert, *Commun. Phys.* **3**, 12 (2020)
[13] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, *Rev. Mod. Phys.* **75**, 33 (2001)
[14] R. Bulla, T. A. Costi, and T. Pruschke, *Rev. Mod. Phys.*
FIG. 3. Number of instances (out of 30) that converge to $\mathcal{O} \geq 0.999$ (upper panel), overlap with the closest eigenstate (middle panel) after convergence, and the number of iterations needed for the BFGS optimizer to terminate (lower panel) after minimizing the functions $F_{\text{H}}$ (red) and $F_{\text{Q}}$ (blue) for the Ising Hamiltonian. In the middle and lower plots, for each value of the model parameter $h$, there are 30 data points, one for each set of randomly generated initial parameters for the ansatz. The color intensity of the circles corresponds to the density of data points. The solid lines connect the average values for the corresponding functions.

80, 395 (2008)
[15] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O’Brien, Nat. Commun. 5, 4213 (2014)
[16] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, New Journal of Physics 18, 023023 (2016)
[17] V. Akshay, H. Philathong, M. Morales, and J. Biamonte, Physical Review Letters 124, 10.1103/physrevlett.124.090504 (2020).
[18] Y. LeCun, Y. Bengio, and G. Hinton, Nature 521, 436 (2015)
[19] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Nature 549, 195 (2017)
[20] M.-H. Yung, J. Casanova, A. Mezzacapo, J. McClean, L. Lamata, A. Aspuru-Guzik, and E. Solano, Sci. Rep. 4, 3589 (2014).
[21] J. Preskill, Quantum 2, 79 (2018)
[22] N. Moll, P. Barkoutsos, L. S. Bishop, J. M. Chow, A. Cross, D. J. Egger, S. Filipp, A. Fowler, J. M. Gambetta, and M. Ganzhorn, Quantum Sci. Technol. 3, 030505 (2018).
[23] C. Kokail, C. Maier, R. van Bijnen, T. Brydges, M. K. Joshi, P. Jurcevic, C. A. Muschik, P. Silvi, R. Blatt, C. F. Roos, and P. Zoller, Nature 569, 355 (2019)
[24] R. LaRose, A. Tikku, E. ONeel-Judy, L. Cincio, and P. J. Coles, npj Quantum Inf. 5, 57 (2019)
[25] P. J. J. O’Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C.Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, and J. M. Martinis, Phys. Rev. X 6, 031007 (2016).
[26] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Nature 549, 242 (2017).
[27] Y. Shen, X. Zhang, S. Zhang, J.-N. Zhang, M.-H. Yung, and K. Kim, Phys. Rev. A 95, 020501 (2017).
[28] G. H. Golub and C. F. van Loan, Matrix Computations (Johns Hopkins University Press, 1996).
[29] J. Nocedal and S. J. Wright, Numerical Optimization (Springer, New York, 2006).
[30] A. Dutta, G. Aeppli, B. K. Chakrabarti, U. Divakaran, T. F. Rosenbaum, and D. Sen, *Quantum phase transitions in transverse field spin models: From statistical physics to quantum information* (Cambridge University Press, Cambridge, 2015).

[31] A. Berezutskii, M. Beketov, D. Yudin, Z. Zimborás, and J. Biamonte, arXiv:2005.02104 (2020).

[32] S. Arai, M. Ohzeki, and K. Tanaka, J. Phys. Soc. Jpn. **87**, 033001 (2018).

[33] C. J. Hamer, Z. Weihong, and J. Oitmaa, Phys. Rev. D **56**, 55 (1997).

[34] B. Buyens, F. Verstraete, and K. Van Acoleyen, Phys. Rev. D **94**, 085018 (2016).

[35] A. K. Zhuravlev and M. I. Katsnelson, Phys. Rev. B **61**, 15534 (2000).

[36] C. Karrasch and J. E. Moore, Phys. Rev. B **86**, 155156 (2012).

[37] M. Hohenadler, S. Wessel, M. Daghofer, and F. F. Assaad, Phys. Rev. B **85**, 195115 (2012).

[38] A. Uvarov, J. Biamonte, and D. Yudin, arXiv:2005.00544 (2020).

[39] M. Wilson, S. Stromswold, F. Wudarski, S. Hadfield, N. M. Tubman, and E. Rieffel, arXiv:1908.03185 (2019).

[40] K. J. Sung, M. P. Harrigan, N. C. Rubin, Z. Jiang, R. Babbush, and J. R. McClean, arXiv:2005.11011 (2020).

[41] V. Verteletskyi, T.-C. Yen, and A. F. Izmaylov, arXiv:1907.03358 (2019).

[42] T.-C. Yen, V. Verteletskyi, and A. F. Izmaylov, J. Chem. Theory Comput. **16**, 2400 (2020).

[43] O. Higgott, D. Wang, and S. Brierley, Quantum **3**, 156 (2019).
A. Stability of variational solution and impact of spectral gap on convergence

In the following we demonstrate that the functions $F_H$ and $F_Q$ have the same performance in dealing with Hamiltonians with small intereigenvalue distance. First, we generate 300 random Hamiltonians $H$ and 300 random sets of initial parameters for the ansatz. We next obtain the convergence rates and the overlap gains for the generated Hamiltonians in three variants: (a) multiplied by $0.1$, (b) the original ones, and (c) multiplied by $10$. The corresponding plots are illustrated in Fig. 4. We note that for the function $F_Q$, multiplying the Hamiltonian by a real number is equivalent to setting the evolution time $t$ since $Q = e^{iHt}$.

Using the spectral theorem for the target Hamiltonian, $H = \sum_{j=0}^{N} \lambda_j |\lambda_j\rangle \langle \lambda_j|$, we can rewrite the functions (5) and (6) as follows:

$$F_H = \sum_{ij} \lambda_i (\lambda_i - \lambda_j) |c_i|^2 |c_j|^2,$$

$$F_Q = 2 \sum_{ij} |c_i|^2 |c_j|^2 \sin^2 \left( \frac{(\lambda_i - \lambda_j)t}{2} \right),$$

where $c_j = \langle \lambda_j | \psi \rangle$. For the sake of simplicity, we assume $0 < \lambda_0 < \lambda_1 < \lambda_2 < \ldots < \lambda_N$. Looking at the equations above, one may expect that the smaller the average distance between the eigenvalues of $H$, $(\lambda_i - \lambda_k)$, the less the functions differ from each other. Therefore, provided that the target Hamiltonian has small intereigenvalue distances, the functions $F_H$ and $F_Q$ show the same efficiency in finding an eigenvector. The VQE solution in the neighborhood of the state $|\lambda_0\rangle$ may be written as

$$|\psi\rangle = \frac{1}{\sqrt{1 + \epsilon^2}} |\lambda_0\rangle + \frac{\epsilon}{\sqrt{1 + \epsilon^2}} |\phi\rangle,$$

with $\epsilon \ll 1$. The vector $|\phi\rangle = \sum_{j>0} c_j |\lambda_j\rangle$ is normalized, so the squares of absolute values of $c_j$ sum to unity. Consider the variance of $H$ as given by Eq. (5).
\[ \delta \equiv F(\mathcal{H}) = (\lambda_q^2 + d^2 - 2e\lambda_q)e^2 - (d^2 + e^2 + 2\lambda_q^2 - 4e\lambda_q)e^4 + o(e^4), \quad (11) \]

on condition that \( d^2 = \langle \varphi | \mathcal{H}^2 | \varphi \rangle = \sum_{j>0} |c_j|^2 \lambda_j^2 \), and \( e = \langle \varphi | \mathcal{H} | \varphi \rangle = \sum_{j>0} |c_j|^2 \lambda_j \). Conversely, one can treat Eq. \( (11) \) as an implicit function \( \epsilon = \epsilon(\delta) \). By considering the derivatives of this function in the vicinity of \( \epsilon = 0, \delta = 0 \), we arrive at

\[
\epsilon^2 = \frac{\delta}{\lambda_q^2 + d^2 - 2e\lambda_q} + \frac{3(d^2 + e^2 + 2\lambda_q^2 - 4e\lambda_q)}{(\lambda_q^2 + d^2 - 2e\lambda_q)^3} \delta^2 + o(\delta^2),
\]

or, alternatively,

\[
\delta = \epsilon^2 \sum_{j>0} |c_j|^2 (\lambda_j - \lambda_0)^2 + o(\epsilon^2).
\]

Suppose that we are looking for an eigenvector of a unitary \( Q = e^{-i\mathcal{H}t} \) for some Hermitian \( \mathcal{H} \) and real \( t \), and we can implement this \( \mathcal{H} \)-evolution. Assume that \( |\psi\rangle \) is sufficiently close to an eigenvector, and \( |\langle \psi | e^{-i\mathcal{H}t} | \psi \rangle| = 1 - \delta \), where

\[
\delta = 2\epsilon^2 \left( 1 - \sum_{j>0} |c_j|^2 \cos \left[ (\lambda_j - \lambda_0)t \right] \right) + o(\epsilon^2).
\]

Notice that \( \delta \) is a function of time \( t \). Using this dependence, we can extract some extra properties of the target Hamiltonian.

---

**FIG. 5.** Number of instances (out of 50) converged to \( \mathcal{O} \geq 0.999 \) (upper), final overlap with the closest eigenstate (middle) and the number of iterations needed for the BFGS optimizer to terminate (lower) after minimizing the function \( F_\mathcal{Q} \) for the Ising Hamiltonian. In the middle and lower plots, for each evolution time, there are 50 data points, one for each set of randomly generated initial parameters for the ansatz. The color intensity of the circles corresponds to the density of data points. The solid line connects the average values.
B. Evolution time as a hyperparameter

In the main text, we emphasized that there could be an optimal evolution time for the function \( F_Q \) for certain problem instances. To provide a quantitative arguments we provide BFGS optimization performance versus the evolution time in Fig. 5. To illustrate our findings, we consider the Hamiltonian of transverse field Ising model given by Eq. (2) in the main text at criticality \( h = 1 \),

\[
H_{\text{TFIM}} = \sum_j (Z_j Z_{j+1} + X_j),
\]

as the target. Our numerical experiments do not support the idea that any significant advantage can be achieved by tuning the evolution time. However, some values of \( t \) (e.g., 0.1\( \pi \), 0.5\( \pi \) or 0.9\( \pi \)) allow to achieve a slightly better performance. We also note that the best results are obtained for \( t = 0.3\pi \approx 1 \).