

Optimizing quantum-classical algorithmic resources for accuracy in VQE

Harshit Verma^{1,2}, Thomas Ayrat^{3,4}, Alexia Auffèves^{1,2}, and Robert Whitney⁵,
¹ MajuLab, CNRS-UCA-SU-NUS-NTU International Joint Research Laboratory, 117543 Singapore
² Centre for Quantum Technologies, National University of Singapore, 117543 Singapore
³ CPHT, CNRS, École Polytechnique, IP Paris, F-91128 Palaiseau, France
⁴ Eviden Quantum Laboratory, 78340 Les Clayes-sous-Bois, France
⁵ Université Grenoble Alpes, CNRS, LPMCM, 38000 Grenoble, France

Introduction. Variational Quantum Eigensolver (VQE) is a hybrid quantum-classical algorithm that works by refining the tunable parameters (through classical optimizer) in a variational quantum circuit, in order to minimize the measured energy of the prepared state [1, 2]. This work [3] aims at characterizing and minimizing the algorithmic resources – quantum and classical – that control the achievable accuracy of VQE algorithm.

We consider the number of gates as an algorithmic quantum resource. In the absence of noise, increasing the number of gates – through adding ansatz layers – enhances the expressibility, and thereby improves the achievable accuracy of the algorithm. On the classical side, we consider the number of iterations of the classical optimizer as an algorithmic resource. For a fixed circuit size, the accuracy eventually saturates with iterations as the optimization converges.

Noise – a characteristic of current quantum computers – upends the above simple relationship of algorithmic resources with accuracy. First, noise limits the circuit depth that can produce meaningful results, as errors accumulate with increasing depth [4, 5, 6], inducing bias in the estimated energy. Moreover, the scenario is further nuanced when the number of iterations is also taken into account. A key question then arises: does increasing the computational effort through additional iterations continue to improve accuracy, or does the accumulation of noise ultimately offset the marginal benefits gained from further optimization? This question also feeds into the over-arching resource optimization objective: What is required to minimize resources for a given (desired) accuracy of the algorithm?

Methodology and Results. In this work [3], we develop phenomenological scaling relations that captures the trade-offs between accuracy, circuit size, and iteration count in VQE algorithm; that enables us to either minimize the resource for a target accuracy or maximize the accuracy for a given resource cost. The scaling relations approach parallels recent developments in classical machine learning, where scaling laws relate model performance to compute, data, and parameters, albeit the element of noise is generally absent in this case and forms a distinguishing feature of quantum algorithms.

As a toy model to illustrate the scaling relations, we adopt the open one-dimensional Heisenberg chain, with the Hamiltonian given as: $H = J \sum_{i=1}^{n-1} (S_i^x S_{i+1}^x + S_i^y S_{i+1}^y + S_i^z S_{i+1}^z)$. We set $J = 1$ (unitless) and $n = 5$ spins, thus, the exact ground state energy – found using exact diagonalization – is $E_{\text{gs}} = -7.712$. The parametrized quantum state is prepared from $|0\rangle^{\otimes n}$ using the Hamiltonian Variational Ansatz, which decomposes H into non-commuting terms h_q applied as $e^{-i\theta_q h_q}$ in alternating even/odd sublayers (six parameters per layer), yielding a gate count N_g tunable through the number of layers (N_{layers}).

The circuit is transpiled to a hardware-agnostic gate set that includes explicit identity operations representing idle periods under an equal duration scheduling, since idling also accumulates error; single-qubit gates (including identities) are counted once and two-qubit gates twice to reflect their higher error rates and resource cost, giving $N_g = 150 N_{\text{layers}}$ for the toy model. Measurement-basis rotations are excluded from N_g as a small, implementation specific overhead, and since the shot count is assumed large enough that sampling error is negligible relative to circuit noise, both the number of shots and measurements are held fixed rather than treated as optimizable algorithmic resources. We also assume that the algorithm is warm-started in a fertile valley with non-vanishing gradients and therefore avoids barren plateaus.

With these specifications of the toy model, and the hardware-agnostic compilation of the ansatz, in following section we proceed to establish the phenomenological scaling relations based on simulations with 50 random initial parameter sets, fitting the mean energy over seeds at each iteration and at convergence.

Scaling relations for the noiseless VQE. From the simulation, we find that in the noiseless case the energy decays toward its converged value as

$$E(N_{\text{it}}, N_g) = \alpha e^{-\mu(N_g)N_{\text{it}}} + E_{\infty}^{\text{noiseless}}(N_g), \quad (1)$$

where N_{it} is the iteration count, N_g the total transpiled gate count, and α, μ are fit parameters. We take μ (related to trainability) and $E_{\infty}^{\text{noiseless}}$ (related to expressibility) to depend on N_g and neglect the weak N_g -dependence of α . The converged energy $E_{\infty}^{\text{noiseless}}$ approaches the exact ground state as the ansatz gains expressibility with added layers,

$$E_{\infty}^{\text{noiseless}}(N_g) = \beta e^{-\kappa N_g} + E_{\text{gs}}, \quad (2)$$

while the convergence rate μ decreases with circuit size according to a power law with an empirical offset N_g^{th} ,

$$\mu(N_g) = \mu_0 (N_g - N_g^{\text{th}})^{-\lambda}, \quad (3)$$

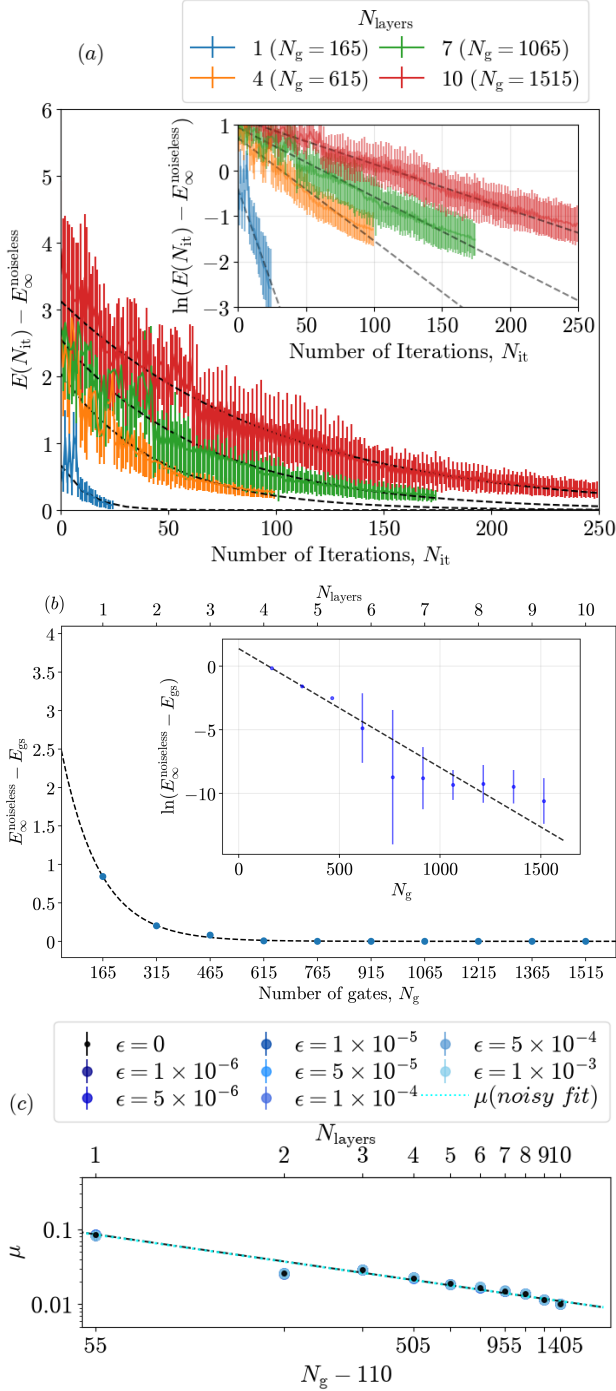


Figure 1: (a) Convergence of the VQE energy with the number of iterations N_{it} for varying numbers of layers (i.e. different N_g), fit to Eq. (1) (black dashed lines) to extract the convergence factor μ ; the inset shows the same fit on a log-linear scale. (b) Convergence of the converged noiseless energy ($N_{it} \rightarrow \infty$) to the exact value E_{gs} as a function of the gate count N_g , following the relation in Eq. (2); the inset shows the same on a log-linear scale. (c) Dependence of the convergence rate μ on the rescaled gate count $N_g - 110$ (log scale), exhibiting the modified power law of Eq. (3) (noiseless: black dot-dashed; noisy: cyan dashed). Datapoints in (c) are obtained by fitting Eq. (1) for each layer count using $\sim 25N_{\text{layers}}$ iterations. Error bars throughout represent the standard deviation over 50 random seeds. The fits yield $\mu_0 = 1.09 \pm 0.20$, $\lambda = 0.63 \pm 0.03$, $N_g^{\text{th}} = 110$, $\beta = 3.94$, and $\kappa = 0.01$, with $\alpha = 2$ taken as a fiducial value owing to its weak N_g -dependence.

which defines the optimization landscape. Optimizing Δ is practically useful: for a target accuracy it identifies the most efficient balance between circuit size and iteration count, reducing runtime and energy cost (provided a hardware specific energy model is available) and noise exposure while preserving performance. Using Eq. (5), we map the error, $\delta E = E(N_{it}, N_g) - E_{gs}$ onto the $\Delta = N_{it}N_g$ heatmap in Fig. 2 (a), with the fit parameters from phenomenological scaling relations (see Fig. 1 caption) and noise fixed at $\epsilon = 10^{-6}$.

reflecting the slower optimization as more variational parameters are introduced. The phenomenological scaling parameters – μ , κ , N_g^{th} , β – are obtained by maximizing the fits in Eqs. (1)–(3). While their specific values (provided in Fig. 1 caption) are system-dependent, we expect similar scaling forms across a broad class of variational algorithms.

Together, Eqs. (1)–(3) express the noiseless energy in terms of the algorithmic resources N_g and N_{it} . These are shown for our toy model in Fig. 1.

Scaling relations for the noisy VQE . Gate noise causes errors that accumulate with circuit size. We model incoherent noise with a global depolarizing channel, $\mathcal{D}(\rho) = (1 - \epsilon)\rho + \epsilon\mathbb{I}/2$, which approximates local gate noise in sufficiently random circuits. After N depolarizing events the circuit state is $\tilde{\rho} = (1 - \epsilon)^N U(\theta)\rho U(\theta)^\dagger + [1 - (1 - \epsilon)^N] \frac{\mathbb{I}}{2^n}$, with fidelity $\mathcal{F} = (1 - \epsilon)^N$. The corresponding energy obtained from VQE is $E_{\text{noisy}}(\theta) = (1 - \epsilon)^N \text{Tr}[H U(\theta)\rho U(\theta)^\dagger] + [1 - (1 - \epsilon)^N] \frac{\text{Tr}(H)}{2^n}$, whose second term vanishes for traceless Hamiltonians such as the Heisenberg model, so noise simply rescales the noiseless energy by $(1 - \epsilon)^N$. Setting $N = N_g$ couples noise directly to the number of gates as algorithmic resource, giving a converged energy

$$E_{\infty}^{\text{noisy}}(N_g) = (1 - \epsilon)^{N_g} E_{\infty}^{\text{noiseless}}(N_g), \quad (4)$$

which exposes a trade-off: added gates (through ansatz layers) first improve accuracy by increasing the expressibility but noise eventually dominates, defining an optimal circuit size $N_g^{\text{optimal}}(\epsilon)$. This is shown in Fig. 2(a). As global depolarizing noise uniformly rescales the cost landscape rather than deforming it, μ is unchanged from the noiseless case and independent of ϵ (see Fig. 1 (c)). Combining the above with Eqs. (1)–(3) yields the energy as a function of the two algorithmic resources – N_{it} and N_g – and the noise parameter ϵ :

$$E(N_{it}, N_g) = (1 - \epsilon)^{N_g} \left[\alpha e^{-\mu_0 N_{it} (N_g - N_g^{\text{th}})^{-\lambda}} + \beta e^{-\kappa N_g} + E_{gs} \right], \quad (5)$$

where the first term captures optimizer convergence, the latter two the growth and saturation of the ansatz's expressivity, and the prefactor $(1 - \epsilon)^{N_g}$ the cumulative noise factor that reflects error accumulation due to each gate added. This unified form is the basis for the efficiency and resource-optimization analysis that follows.

Resource optimization. We employ the phenomenological scaling relations from the previous section to minimize resource consumption under noise. Treating the iteration count N_{it} and gate count N_g as the primary algorithmic resources, the cumulative cost is the total number of gate operations, $\Delta = N_{it}N_g$,

In Fig. 2 (a), each iso-error contour (magenta dotted) has a point of minimum Δ (red stars). Our plots show that for a fixed metric the total resources can be cut by a factor of two or more (e.g. at $\delta E = 0.01$), yielding significant speed and resource savings. Thus Eq. (5) provides a minimal yet complete link between the resources (N_g, N_{it}), the noise ϵ , and the accuracy, underpinning the resource-constrained analysis below and enabling analytic extraction of the optimal gate count.

It is meaningful to analyze VQE algorithm under a fixed total resource budget Δ , representing experimental limits from coherence time, measurement budget, or available compute. Fixing Δ eliminates N_{it} from Eq. (5),

$$E_{\Delta}(N_g) = (1 - \epsilon)^{N_g} \left[\alpha e^{-\mu_0 \Delta / [N_g(N_g - 110)^\lambda]} + \beta e^{-\kappa N_g} + E_{gs} \right], \quad (6)$$

which is equivalent to moving along the iso-resource lines (black dashed) of Fig. 2 (a) and reading off the minimum error. To quantify performance per unit cost we introduce the efficiency $\eta = \frac{\mathcal{M}}{\Delta} = \frac{1 - \delta E}{\Delta} = \frac{1 - [E(N_{it}, N_g) - E_{gs}]}{N_{it} N_g}$, chosen to be monotonic, so a higher η follows from a larger metric (lower error).

Maximizing the metric $\mathcal{M}$ over N_g at fixed Δ yields $\max(\mathcal{M}_{\Delta})$, plotted against Δ in Fig. 2 (b). The best achievable metric depends strongly on ϵ and saturates with Δ , showing that resource investment becomes useless beyond a point, and that a high cost is incurred to reach a practical target such as $\delta E = 0.01$. The corresponding efficiency (inset in Fig. 2 (b)) first rises with Δ , denoting a sharp metric gain, then decreases with added resources, denoting no further metric improvement. Both the peak efficiency and its location depend on ϵ , defining an efficiency frontier that requires low ϵ for reasonable and efficient computation. Furthermore, we can also extract a numerical solution based on stationarity condition (leading to a transcendental equation) from Eq.(6); although we do not show it here for brevity.

Conclusions and outlook. We have shown that the accuracy of a noisy VQE algorithm follows scaling relations in its two main resources – the circuit size and the number of iterations – from which the resource allocation for a target accuracy follows directly, cutting the total cost by a factor of two or more. We also tested these relations on a *fake* IBM Nighthawk backend, where noise is not global depolarizing and error mitigation is routinely applied (results omitted here for brevity). Without mitigation, the runs are in the noise dominated regime and resource optimization is not feasible; with zero-noise extrapolation the mitigated data are well described by our global-depolarizing form at an effective rate $\epsilon_{\text{eff}} \approx 3 \times 10^{-4}$, and the scaling relations hold. Therefore, this work [3] establishes a holistic way to minimize the quantum and classical resources in VQE and it survives on real quantum processors. It extends the machine-learning notion of scaling laws to the noisy quantum setting.

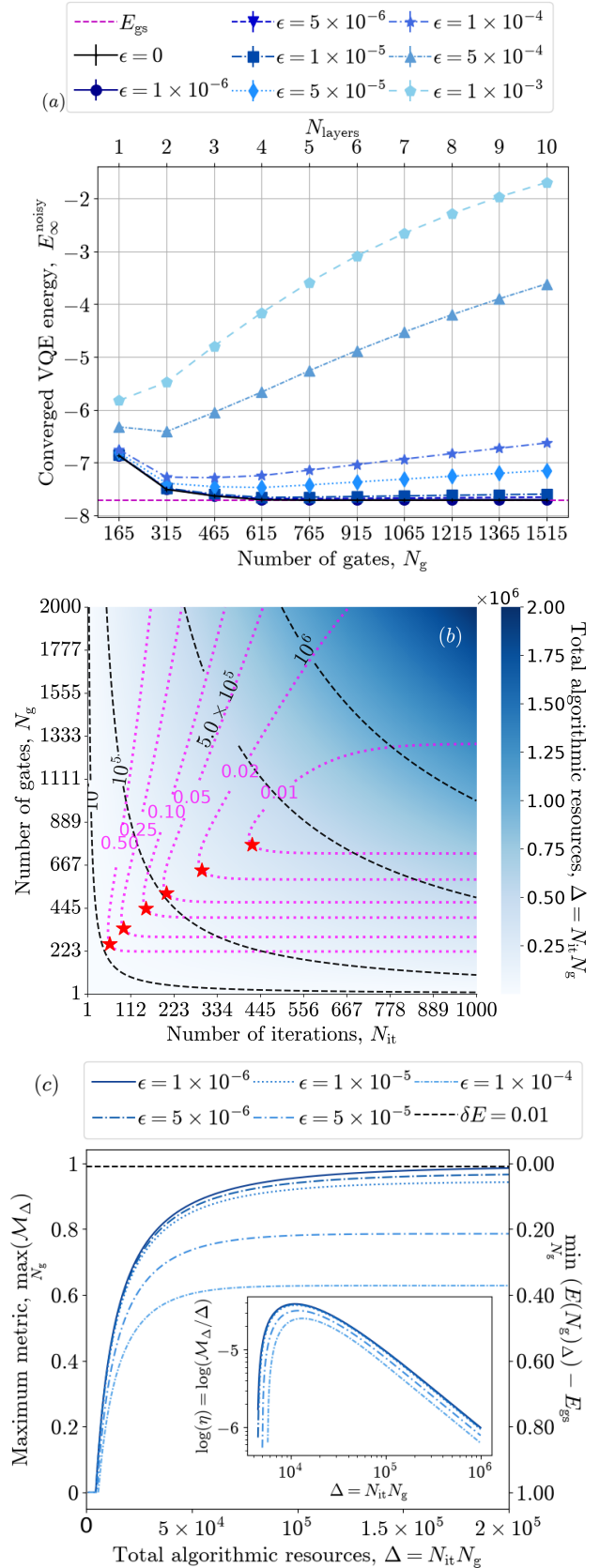


Figure 2: (a) Converged energy $E_{\infty}^{\text{noisy}}$ from Eq. (4) versus gate count N_g , showing a tradeoff between expressivity and error accumulation. (b) Heatmap of the total resource $\Delta = N_{it} N_g$ at $\epsilon = 10^{-6}$, obtained from Eq. (5), with iso-error contours of δE (magenta dotted) and iso- Δ lines (black dashed). Red stars mark the minimum Δ on each iso-error contour. This plot enables resource minimization for a target accuracy or resource-constrained accuracy analysis (c) Best achievable metric $\max_{N_g}(\mathcal{M}_{\Delta})$ (left axis) and minimum error (right axis) versus Δ , from Eq. (6). Inset: efficiency $\eta = \mathcal{M}/\Delta$, which peaks at intermediate Δ . All the plots are based on phenomenological scaling fit parameters mentioned in Fig. 1 caption.

References

- [1] Kishor Bharti, Alba Cervera-Lierta, Thi Ha Kyaw, Tobias Haug, Sumner Alperin-Lea, Abhinav Anand, Matthias Degroote, Hermann Heimonen, Jakob S. Kottmann, Tim Menke, Wai-Keong Mok, Sukin Sim, Leong-Chuan Kwek, and Alán Aspuru-Guzik. Noisy intermediate-scale quantum algorithms. *Reviews of Modern Physics*, 94:015004, 2 2022. ISSN 0034-6861. doi: 10.1103/RevModPhys.94.015004.
- [2] Jules Tilly, Hongxiang Chen, Shuxiang Cao, Dario Picozzi, Kanav Setia, Ying Li, Edward Grant, Leonard Wossnig, Ivan Rungger, George H. Booth, and Jonathan Tennyson. The variational quantum eigensolver: A review of methods and best practices. *Physics Reports*, 986:1–128, 11 2022. ISSN 03701573. doi: 10.1016/j.physrep.2022.08.003.
- [3] Harshit Verma, Thomas Ayril, Alexia Auffèves, and Robert Whitney. Optimizing resource allocation for accuracy in noisy variational quantum algorithms, 2026. URL <https://arxiv.org/abs/2606.20153>.
- [4] Enrico Fontana, Nathan Fitzpatrick, David Muñoz Ramo, Ross Duncan, and Ivan Rungger. Evaluating the noise resilience of variational quantum algorithms. *Phys. Rev. A*, 104(2):022403, aug 2021. doi: 10.1103/physreva.104.022403.
- [5] Giacomo De Palma, Milad Marvian, Cambyse Rouzé, and Daniel Stilck França. Limitations of variational quantum algorithms: A quantum optimal transport approach. *Phys. Rev. X Quantum*, 4(1):010309, January 2023. ISSN 2691-3399. doi: 10.1103/prxquantum.4.010309. URL <http://dx.doi.org/10.1103/PRXQuantum.4.010309>.
- [6] Samson Wang, Enrico Fontana, M. Cerezo, Kunal Sharma, Akira Sone, Lukasz Cincio, and Patrick J. Coles. Noise-induced barren plateaus in variational quantum algorithms. *Nature Communications*, 12(1):6961, Nov 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-27045-6. URL <https://doi.org/10.1038/s41467-021-27045-6>.