

## Trainable, self-verifying variational quantum eigensolver using adiabatic warm starts

Bojan Žunkovič<sup>1,2,\*</sup>, Marco Ballarín<sup>3</sup>, Lewis Wright<sup>3</sup>, and Michael Lubasch<sup>3,†</sup><sup>1</sup>*Faculty of Computer and Information Science, University of Ljubljana, Večna pot 113, 1000 Ljubljana, Slovenia*<sup>2</sup>*Rudolfovo, Science and Technology Center, Novo mesto, Slovenia*<sup>3</sup>*Quantinuum, Partnership House, Carlisle Place, London SW1P 1BX, United Kingdom*

(Received 30 April 2026; accepted 22 September 2026; published 29 October 2026)

We study an adiabatic variant of the variational quantum eigensolver (VQE) in which VQE is performed iteratively for a sequence of Hamiltonians along an adiabatic path. We derive the conditions under which gradient-based optimization successfully prepares the adiabatic ground states. These conditions define a trainable tracking region. Additionally, when combined with a gap lower bound, energy-standard-deviation measurements certify eigenstate accuracy at run-time and verify convergence to the global optimum.

DOI: [10.1103/m538-45pk](https://doi.org/10.1103/m538-45pk)

**Introduction.** The variational quantum eigensolver (VQE) is a promising algorithm for ground-state preparation on current noisy intermediate-scale quantum (NISQ) computers [1–3]. In its traditional formulation, however, VQE faces significant obstacles, including challenging optimization landscapes characterized by barren plateaus [4] and numerous local optima [5], as well as the absence of an intrinsic mechanism to verify the correctness of the obtained solutions [6].

An alternative to VQE, which can avoid these obstacles, is adiabatic ground-state preparation. In its standard circuit-model implementation [7,8], however, this approach often generates quantum circuits that are too deep to be feasible on NISQ devices. Various NISQ-compatible schemes have been proposed—variational path-following and Hamiltonian-ansatz-based methods [9–16], digitized and counterdiabatic protocols [17–23], and compression or meta-learning approaches [24–27].

Closely related to adiabatic methods is the concept of warm start, which has recently gained attention in variational quantum optimization [28–31]. Several adiabatically inspired methods already reuse parameters along a path, whereas these recent works study warm-start initialization and landscape guarantees more generally. The central idea is to optimize the variational ansatz incrementally along a predefined optimization path. At each point along this path, the optimizer initializes the ansatz with the parameters from the previous point, thereby warm starting the optimization. It has been shown that warm-start strategies can provably avoid the barren plateau problem in variational quantum simulation [28],

multivariate state preparation [30], and VQE optimization using adiabaticity [31].

In this Letter, we analyze the warm-start paradigm in an adiabatic VQE (AVQE), where VQE is executed sequentially along the Hamiltonians of a discretized adiabatic path. We determine the conditions under which gradient descent tracks the instantaneous ground states, introduce a run-time verification procedure based on the energy standard deviation, provide gap-conditional correctness certificates, present numerical experiments, and analyze robustness against measurement shot noise. Figure 1 visualizes our key findings. Our results establish that, under the stated gap, representability, and regularity assumptions, AVQE can be formulated as a trainable, self-verifying algorithm for ground-state computations compatible with NISQ-style circuits.

**Setup.** We consider a linear adiabatic schedule defined by the Hamiltonian

$$H(\lambda) = (1 - \lambda)H_i + \lambda H_f, \quad \lambda \in [0, 1], \quad (1)$$

where the initial Hamiltonian  $H_i$  is chosen to have an easy-to-prepare ground state and the final Hamiltonian  $H_f$  encodes the desired target ground state. We denote the instantaneous ground state by  $|\psi_0(\lambda)\rangle$ , the spectral gap by

$$\Delta(\lambda) = E_1(\lambda) - E_0(\lambda), \quad (2)$$

where  $E_0(\lambda)$  is the ground and  $E_1(\lambda)$  the first-excited-state energy at  $\lambda$ , and we denote the minimum gap along the path by

$$\Delta_{\min} = \min_{\lambda \in [0, 1]} \Delta(\lambda). \quad (3)$$

We use the variational ansatz

$$|\psi(\boldsymbol{\theta})\rangle = U_{M-1} \cdots U_0 |0\rangle, \quad U_j = e^{-iP_j \theta_j}, \quad (4)$$

where bold symbols denote vectors,  $P_j \in \{I, X, Y, Z\}^{\otimes n}$ , and  $M$  is the number of variational parameters. For a general smooth ansatz, analogous conclusions follow with ansatz-specific smoothness and curvature bounds, but without the explicit dependence on the number of variational parameters.

\*Contact author: bojan.zunkovic@fri.uni-lj.si

†Contact author: michael.lubasch@quantinuum.com

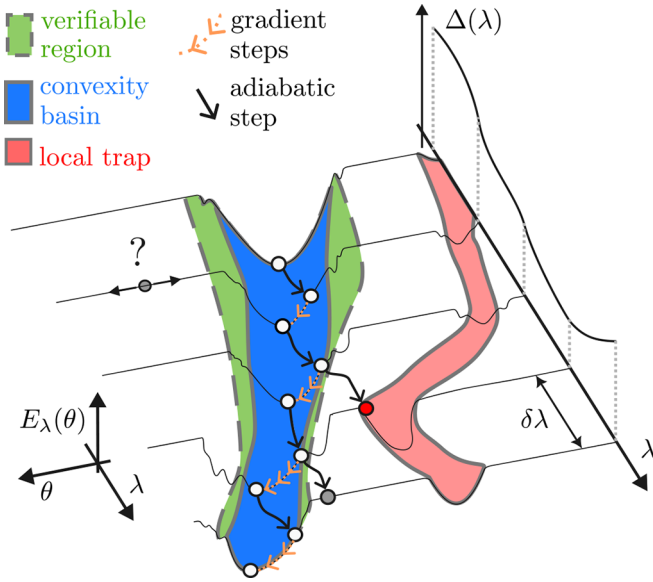


FIG. 1. AVQE optimization landscape. Consider a Hamiltonian  $H(\lambda)$ , where  $\lambda \in [0, 1]$  defines an adiabatic path, a variational ansatz parametrized by  $\theta$ , and the associated energy-based cost function  $E_\lambda(\theta)$ . Starting from the ground state at  $\lambda = 0$ , there exists a basin (blue; Theorem 1) that follows the adiabatic path and guarantees convexity. The size of the convexity basin decreases at most linearly with the instantaneous gap  $\Delta(\lambda)$  (black curve, right-hand two-dimensional plot), which denotes the energy difference between the ground and first excited states at  $\lambda$ . Outside the convexity basin, the optimization landscape may exhibit barren plateaus (white) and local minima (red), which can hinder gradient-based optimization. In the verifiable region (green; Theorem 2), we cannot assume convexity, but we can identify a unique nearby eigenstate. Ground-state initialization and stepwise tracking identify it with the ground-state branch. This enables adaptation of the adiabatic step width  $\delta\lambda$  to ensure provable convergence to the target ground state at  $\lambda = 1$ .

The variational energy at fixed  $\lambda$  is

$$E_\lambda(\theta) = \langle \psi(\theta) | H(\lambda) | \psi(\theta) \rangle, \quad (5)$$

and we denote by  $\theta^*(\lambda)$  a global minimizer satisfying  $|\psi(\theta^*(\lambda))\rangle = |\psi_0(\lambda)\rangle$ .

The variational manifold is equipped with the quantum metric, defined as the real part of the quantum geometric tensor:

$$g_{\mu\nu}(\theta(\lambda)) = \text{Re}(\langle \partial_\mu \psi | \partial_\nu \psi \rangle - \langle \partial_\mu \psi | \psi \rangle \langle \psi | \partial_\nu \psi \rangle), \quad (6)$$

which determines the local behavior of the variational optimization landscape.

We discretize the adiabatic parameter into  $T$  slices,

$$\lambda_t = t \delta\lambda, \quad \delta\lambda = \frac{1}{T}, \quad t = 0, 1, \dots, T, \quad (7)$$

and perform optimization sequentially along the path using warm starts. We initialize  $\theta_0 = \theta^*(0)$ . For each slice  $t = 1, \dots, T$ , we run  $K$  steps of gradient-based optimization on  $E_{\lambda_t}$ , starting from the previous parameter values  $\theta_{t-1}$ :

$$\theta_t^{(0)} = \theta_{t-1}^{(K)}, \quad (8)$$

$$\theta_t^{(k+1)} = \theta_t^{(k)} - \eta \mathcal{G}_t(\theta_t^{(k)}), \quad k = 0, 1, \dots, K-1, \quad (9)$$

$$|\psi_t\rangle = |\psi(\theta_t^{(K)})\rangle, \quad (10)$$

where  $\eta > 0$  is the learning rate and  $-\mathcal{G}_t(\theta)$  denotes the chosen descent direction for the objective at  $\lambda = \lambda_t$ . In the simplest case,  $\mathcal{G}_t(\theta) = \nabla_\theta E_{\lambda_t}(\theta)$  (vanilla gradient descent); Theorem 1 covers only this case, while natural-gradient and momentum methods [32,33] require separate stability analyses. Overall, the procedure performs  $T$  parameter shifts in  $\lambda$  and  $K$  optimization steps per shift, leading to a total of  $N_{\text{updates}} = TK$  optimization updates.

For any operator  $A$ ,  $\|A\|_{\text{op}}$  represents its operator norm. If  $A$  depends on  $\lambda$ , we define  $\|A\|_{\text{op}} = \max_\lambda \|A(\lambda)\|_{\text{op}}$ . For a Hermitian operator  $A$ , its standard deviation with respect to  $|\psi_t\rangle$  is defined as

$$\sigma_{\psi_t}(A) = \sqrt{\langle \psi_t | A^2 | \psi_t \rangle - (\langle \psi_t | A | \psi_t \rangle)^2}. \quad (11)$$

Throughout the Letter, we use the following assumptions.

**Assumption 1 (Spectral gap).** The instantaneous ground state of  $H(\lambda)$  is nondegenerate and  $\Delta_{\min} > 0$ .

**Assumption 2 (Exact representability).** There exists a continuously differentiable parameter branch  $\theta^*(\lambda)$ , with  $\theta^*(0) = \theta_0$ , such that, for all  $\lambda \in [0, 1]$ ,

$$|\psi(\theta^*(\lambda))\rangle = |\psi_0(\lambda)\rangle. \quad (12)$$

**Assumption 3 (Nondegenerate geometry).** The entire branch  $\theta^*(\lambda)$  is represented in one fixed subset of nonredundant Pauli-rotation angles, with all remaining angles held fixed, and optimization is performed in that subset. In these coordinates,

$$g(\theta^*(\lambda)) \succeq \gamma I \quad (13)$$

for some  $\gamma > 0$  and all  $\lambda \in [0, 1]$ .

As defined above,  $M$  is the total number of Pauli-rotation angles, so using  $M$  in the bounds is conservative if fewer coordinates are retained. In these coordinates, the curvature bound proved in Supplemental Material (SM) [34] is

$$\nabla_\theta^2 E_\lambda(\theta^*) \succeq 2\Delta(\lambda) g(\theta^*) \succeq 2\gamma \Delta_{\min} I, \quad (14)$$

which is, together with the above assumptions, the basis for proving two theorems that enable gap-conditioned, verifiable, variational ground-state tracking.

**Adiabatic tracking via gradient descent.** Let us first give a brief, informal summary of the central result of our adiabatic optimization approach: Gradient descent, as outlined in the preceding section, tracks the instantaneous ground state along the adiabatic path with  $N_{\text{updates}}$  that depends polynomially on  $\Delta_{\min}^{-1}$ ,  $\gamma^{-1}$ , the number of variational parameters  $M$ , and the relevant operator norms.

**Theorem 1 (Adiabatic tracking).** Assume that Assumptions 1–3 hold and that the optimization protocol described above is used with the Pauli-rotation ansatz of Eq. (4) containing  $M$  parameters and vanilla gradient descent with  $\eta = 1/L$ , where  $L = 4\|H\|_{\text{op}}M$ . Then, the following holds.

If the adiabatic discretization satisfies

$$\delta\lambda \leq \delta\lambda_A = \frac{\gamma^2 \Delta_{\min}^2}{16M^2 \|H\|_{\text{op}} \|\partial_\lambda H\|_{\text{op}}} \quad (15)$$

and the number of optimization steps per slice satisfies

$$K \geq \frac{4 \ln 2 M \|H\|_{\text{op}}}{\gamma \Delta_{\min}}, \quad (16)$$

then, at every slice  $t = 1, \dots, T$ , all iterates  $k = 0, \dots, K$  remain inside the local tracking region,

$$\|\theta_t^{(k)} - \theta^*(\lambda_t)\|_2 \leq \frac{\gamma \Delta_{\min}}{8 M^{3/2} \|H\|_{\text{op}}}. \quad (17)$$

Because the proof is stepwise, the same conclusion holds for any monotone partition  $0 = \lambda_0 < \dots < \lambda_T = 1$  with  $\lambda_t - \lambda_{t-1} \leq \delta \lambda_A$  at every slice.

The local convexity and drift bounds are optimizer independent, but the stated step count and complexity are proved only for vanilla gradient descent; other optimizers require separate local stability and contraction guarantees.

In SM [34], we prove Theorem 1 using the Polyak-Łojasiewicz (PL) condition [35] and show that continued gradient descent at each slice converges to the ground state and requires only logarithmic overhead to reach a specified fidelity. More precisely, if  $\delta E_{\text{init}}$  is the initial energy error at a slice and  $0 < \zeta < 1$ , fidelity at least  $1 - \zeta$  is reached after

$$K_{\text{fid}} \geq \begin{cases} \frac{L}{\gamma \Delta_{\min}} \ln \frac{\delta E_{\text{init}}}{\zeta \Delta_{\min}}, & \delta E_{\text{init}} > \zeta \Delta_{\min}, \\ 0, & \delta E_{\text{init}} \leq \zeta \Delta_{\min}. \end{cases} \quad (18)$$

We also present a complementary imaginary-time time-dependent variational principle perspective [36].

Choosing the largest admissible slice size (up to integer rounding) and the smallest admissible  $K$ , the total number of updates scales as

$$N_{\text{updates}} = \mathcal{O}\left(\frac{\|H\|_{\text{op}}^2 M^3 \|\partial_\lambda H\|_{\text{op}}}{\gamma^3 \Delta_{\min}^3}\right). \quad (19)$$

Notably, the scaling with the gap is worse than the expected  $\mathcal{O}(1/\Delta_{\min}^2)$  scaling observed in the adiabatic theorem [37]. For the parameter choice in Theorem 1, the observed  $\mathcal{O}(\Delta_{\min}^{-3})$  scaling admits an intuitive interpretation: Adiabatic passage slows down as  $T \propto \Delta_{\min}^{-2}$ , while variational tracking requires repeated relaxation toward the instantaneous ground state whose cooling time is determined by the gap and scales as  $\mathcal{O}(\Delta_{\min}^{-1})$ .

Compared to standard quantum annealing [38,39], Theorem 1 has larger gap complexity. However, AVQE can offer greater control over the schedule and additional options for hybrid quantum-classical optimization that are not available in real-time analog quantum annealing. In particular, the gap does not determine the coherence time of the variational ansatz but rather the number of optimization steps, which can be beneficial in noisy devices, assuming that the required circuit depth is not too big. Moreover, we can adjust the step size  $\delta \lambda$  based on an approximation of the gap. Also, we can offload the initial, easy part of the annealing evolution to a classical device and use the quantum device only if the classical variational ansatz is insufficiently expressive.

As a side result, the tracking region is trainable in the following precise sense. In SM [34], we prove that, under the product-uniform perturbations defined there, each retained

gradient component satisfies

$$\text{Var}_X[\partial_k E_\lambda(\theta^* + X)] \geq \frac{\gamma^4 \Delta_{\min}^4}{192 \|H\|_{\text{op}}^2 M^4}. \quad (20)$$

For comparison, the same perturbation distribution gives the following lower bound on the variance of the energy values over the parameter samples (not the quantum variance of an energy measurement):

$$\text{Var}_X[E_\lambda(\theta^* + X)] \geq \frac{\gamma^6 \Delta_{\min}^6}{184320 \|H\|_{\text{op}}^4 M^8}. \quad (21)$$

These statements apply only to the tracking region and do not exclude barren plateaus elsewhere in parameter space.

*Run-time verification via standard deviation.* Theorem 1 establishes sufficient conditions for tracking the instantaneous ground state along the adiabatic path. However, it does not provide an *a posteriori* certification procedure: It offers no mechanism to assess whether the variational state obtained at a given step is close to an exact eigenstate, nor to verify that the final state coincides with the target ground state.

In this section, we introduce a run-time verification criterion based on the energy standard deviation, inspired by Ref. [40]. This criterion localizes the state to a unique eigenvalue. Combined with exact ground-state initialization, the tracking induction, and a gap lower bound, it certifies that the eigenvalue belongs to the ground-state branch.

*Theorem 2 (Run-time verification).* Consider the parameter trajectory initialized in the exact ground state at  $\lambda_0 = 0$ , and evolved according to Theorem 1. Assume that there exists a constant  $\Delta_c > 0$  such that  $\Delta_c \leq \Delta_{\min}$ . If, for all  $t = 0, 1, \dots, T$ ,

$$\sigma_{\psi_t}(H(\lambda_t)) \leq \frac{\Delta_c}{4} \quad (22)$$

and, for every  $t < T$ , the adiabatic step size satisfies

$$\delta \lambda < \delta \lambda_V = \frac{\Delta_c/2 - \sigma_{\psi_t}(H(\lambda_t))}{\sigma_{\psi_t}(H_f - H_i)}, \quad (23)$$

then the state  $|\psi_t\rangle$  is uniquely associated with the ground-state eigenbranch of  $H(\lambda_t)$  and obeys the fidelity bound  $|\langle \psi_0(\lambda_t) | \psi_t \rangle|^2 \geq \frac{8}{9}$ .

For a monotone adaptive partition, the step condition is imposed separately with  $\delta \lambda = \lambda_{t+1} - \lambda_t$  for every  $t < T$ .

The theorem tells us that, if a gap lower bound  $\Delta_c \leq \Delta_{\min}$  is available, then the standard deviation condition  $\sigma_{\psi_t}(H(\lambda_t)) \leq \Delta_c/4$  provides a sufficient, run-time-verifiable certificate that the variational state approximates the actual ground state with fidelity at least 8/9. Starting from the exact ground state, the half-gap step-size condition preserves the branch across slices; PL descent remains in the ground-state basin, and the output test certifies the optimized state. The proof of Theorem 2 is presented in SM [34]. With finite shots, we replace each standard deviation by the one-sided upper confidence bound defined in Sec. S7 of the SM [34] and add a measured energy test. The certificate involves no curvature or metric constants: Its only spectral input is the gap bound  $\Delta_c$ . Moreover, the trainability constants entering  $\delta \lambda_A$  and  $K$  need not be known at run-time—an adaptive halving schedule retains the certification guarantees unconditionally

and, under Assumptions 1–3, terminates after logarithmically many restarts (Secs. S4 and S7 of the SM [34])—so they enter only the complexity analysis.

Energy-based approximate representability with error  $\epsilon$  still yields the verification guarantees, provided the certificate holds. A sufficient condition for the variational optimum to satisfy the quarter-gap test is  $\epsilon < \Delta_c^2/(16W)$ , where  $W = \max_\lambda (E_{\max}(\lambda) - E_0(\lambda))$ . In contrast, tracking requires a generalized version of Assumption 3 that does not follow from the energy error alone: a smooth branch of variational optima with local curvature  $2\mu_{\text{opt}}$ ; the tracking theorem then applies unchanged with  $\gamma\Delta_{\min}$  replaced by  $\mu_{\text{opt}}$ . For  $\epsilon \leq \gamma^2\Delta_{\min}^2/(1600M^2W)$ , metric nondegeneracy along the optimum branch implies  $\mu_{\text{opt}} = \gamma\Delta_{\min}/2$  (see SM [34]).

**Algorithm.** The tracking and verification components combine naturally into an adiabatic variational quantum optimization protocol. Algorithm 1 explicitly implements the guarantees of Theorems 1 and 2: The optimization steps enforce adiabatic tracking, and the energy-standard-deviation measurements provide run-time verification (i.e., we are tracking an eigenstate with a lower bound on the fidelity) and gap-conditional certification (i.e., if the standard deviation is sufficiently small relative to the gap, then the tracked state is the ground state).

The safety factor  $\kappa$  enforces the strict step-size inequality of Eq. (23). Algorithm 1 is stated with exact standard deviations; in the finite-shot setting, each measured standard deviation is replaced by the one-sided upper confidence bound  $\bar{\sigma}$  of Sec. S7 of the SM [34], the output test of line 11 is also applied at  $\lambda = 0$  and supplemented by a measured energy test, and line 12 also doubles the shot counts. The algorithmic parameters  $K$ ,  $\delta\lambda_A$ , and  $\eta$  may be chosen according to Theorem 1 or treated as hyperparameters and selected via empirical optimization; only the theorem choices—or the adaptive restart schedule of line 12 with  $\eta = 1/L$  (Sec. S4 of the SM [34])—carry the stated guarantees.

---



---

ALGORITHM 1. Self-verifying AVQE.

---



---

**Input:** Initial Hamiltonian  $H_i$ , final Hamiltonian  $H_f$ , initial parameters  $\theta_0$ , adiabatic tracking step size  $\delta\lambda_A$ , learning rate  $\eta$ , number of gradient steps  $K$ , estimated gap bound  $\Delta_c$ , safety factor  $\kappa \in (0, 1)$

**Output:** Final variational state  $|\psi(\theta)\rangle$

```

1: Prepare  $|\psi(\theta_0)\rangle$  as the ground state of  $H(0)$ 
2:  $\lambda \leftarrow 0$ 
3: Measure  $\sigma_\psi(H(\lambda))$ 
4: While  $\lambda < 1$  do
5:   Measure  $\sigma_\psi(H_f - H_i)$ 
6:    $\delta\lambda_V \leftarrow \kappa \frac{\Delta_c/2 - \sigma_\psi(H(\lambda))}{\sigma_\psi(H_f - H_i)}$ 
7:    $\delta\lambda \leftarrow \min\{\delta\lambda_A, \delta\lambda_V, 1 - \lambda\}$ 
8:    $\lambda \leftarrow \lambda + \delta\lambda$ 
9:   Perform  $K$  gradient-descent steps on  $E_\lambda(\theta)$ 
10:  Measure  $\sigma_\psi(H(\lambda))$ 
11:  If  $\sigma_\psi(H(\lambda)) > \Delta_c/4$  then
12:     $\delta\lambda_A \leftarrow \delta\lambda_A/2$ ,  $K \leftarrow 2K$ , go to line 1
13:  end If
14: end While
15: return  $|\psi(\theta)\rangle$ 

```

---



---

The algorithm assumes knowledge of a lower bound of the gap. In the absence of such a lower bound, one possibility is to estimate  $\Delta_c$  and run the protocol starting from that estimate, successively reducing it until the energy converges. This heuristic does not itself certify that the estimate is a lower bound on the gap.

**Numerical experiments.** To illustrate the protocol, we consider the one-dimensional transverse-field Ising model with longitudinal field on  $n$  qubits:

$$H(\lambda) = \lambda J \sum_{\ell=1}^{n-1} Z_\ell Z_{\ell+1} + g \sum_{\ell=1}^n X_\ell + h \sum_{\ell=1}^n Z_\ell, \quad (24)$$

for  $J = 1$ ,  $g = 0.5$ , and  $h = 0.25$ . The variational quantum state  $|\psi(\theta)\rangle$  is defined by a Hamiltonian-informed quantum circuit of 64 layers, each containing  $n$   $X$ ,  $n$   $Y$ ,  $n$   $Z$ , and  $n - 1$  nearest-neighbor  $ZZ$  rotations, and a total of  $M = 64(4n - 1)$  parameters (see SM [34]). We investigate system sizes  $n \in \{7, 11, 15\}$ . We divide the adiabatic path  $\lambda \in [0, 1]$  into slices and optimize each slice with gradient descent, using  $K = 50\,000$  gradient steps, a learning rate  $\eta = 0.1$ , and momentum  $\alpha = 0.9$ . Due to computational limitations, we choose  $K = 10\,000$  gradient steps for  $n = 15$ . The momentum update is empirical and is not covered by the vanilla gradient descent complexity proof. The simulations use noiseless gradients evaluated by the classical simulator; shot-based parameter-shift gradients are included only in the resource analysis of Sec. S7 of the SM [34]. The size of each slice,  $\delta\lambda$ , is determined in accordance with Algorithm 1, i.e.,  $\delta\lambda \leftarrow \min\{\delta\lambda_A, \delta\lambda_V, 1 - \lambda\}$ , and is recomputed every five adiabatic slices—we assume that the optimization remains within the convex region between computations of  $\delta\lambda$ . This recomputation schedule is an empirical shortcut and is not covered by Theorem 1, which requires the step-size cap at every accepted slice.

Let  $\lambda_\star$  denote the point at which the gap is minimal. In Fig. 2(a), we report the energy variance  $\text{Var}E_\lambda(\theta)$ , where  $\theta$  is sampled uniformly from a hypercube of size  $\delta\theta$  centered at  $\theta^\star$ , at the sampled slice closest to  $\lambda_\star$ . The gap-dependent factors in our bounds are least favorable near this point along the adiabatic path. We identify two regions in Fig. 2(a): For small hypercube sizes, the energy variance increases with  $\delta\theta$  and does not decrease with  $n$ ; for large hypercube sizes, it decreases over the three sizes studied. These finite-size data are illustrative and do not establish asymptotic scaling. We also show the approximate size of the convex region,  $r_{\text{PL}}$ , for different qubit counts. Details on how this quantity is computed can be found in Secs. S1 and S6 of the SM [34]. The peak of the variance roughly aligns with  $r_{\text{PL}}$ . Our results for the other sampled slices show the same qualitative behavior (see SM [34]).

Figure 2(b) illustrates that, in these simulations, the squared standard deviation (11) of the energy (energy variance) is closely related to the relative energy error. It is therefore a useful run-time indicator of the accuracy of the eigenstates in these examples.

**Shot-noise robustness.** In realistic implementations, gradients and observables are estimated from a finite number of measurement shots. Under the shot-noise model and shot budget specified in SM [34], the resulting stochastic gradients

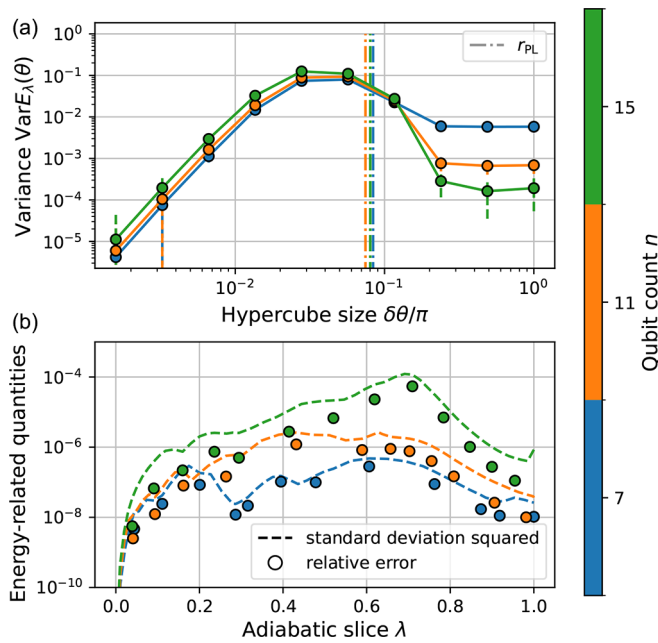


FIG. 2. AVQE numerical experiments. (a) Energy variance with respect to the hypercube size  $\delta\theta/\pi$  at the sampled slice closest to  $\lambda_*$ , where the gap is minimal. We also plot the approximate size of the convex region,  $r_{\text{PL}}$ , for different qubit counts. (b) Relative energy error and squared standard deviation (11) of the energy as a function of the adiabatic parameter  $\lambda$ .

remain unbiased and concentration bounds imply that, with high probability, the optimization dynamics stay within the tracking region identified in Theorem 1. In this regime, finite-shot noise yields a controlled optimization error floor.

The standard-deviation-based verification remains conservative under finite-shot noise when the one-sided confidence bound proved in Sec. S7 of the SM [34] is used. For a variance margin of order  $\Delta_c^2$ ,  $\mathcal{O}(\Delta_c^{-4})$  shots per measured Pauli term suffice, up to coefficient-norm and logarithmic confidence factors. The ungrouped cost per slice additionally includes every Pauli term in  $H$ ,  $H^2$ ,  $(H_f - H_i)$ , and  $(H_f - H_i)^2$ , as well as all parameter-shift gradient evaluations.

**Discussion.** Our results establish a principled paradigm for variational quantum algorithms. The resulting algorithm admits provable correctness guarantees. Under the stated assumptions and run-time tests, the protocol certifies both successful tracking of the target eigenstate and, conditioned on a gap bound, preparation of the actual ground state. It therefore supplements heuristic convergence diagnostics

with a rigorous run-time certificate. The guarantees remain conditional on the scaling of  $M$ ,  $\gamma^{-1}$ ,  $\Delta_{\text{min}}^{-1}$ , Hamiltonian norms, measurement coefficients, and representability resources; no favorable system-size scaling of these quantities is proved here.

From an optimization perspective, the framework directly tackles two central challenges in variational quantum optimization. First, adiabatic warm starting keeps the optimization in a continuously connected, trainable tracking region. Second, the combination of path-following and standard-deviation-based verification provides a test against accepting spurious local minima by enforcing consistency with an underlying physical ground-state branch throughout the evolution.

More broadly, the framework provides a rigorous foundation for developing new classes of hybrid quantum-classical algorithms. An interesting next step is to reformulate the tools developed here for adiabatic time evolution to enable us to track eigenstates other than the ground state. In this context, there are numerous important applications, ranging from excited-state computations to the solution of linear systems of equations [41–43]. Another interesting next step is to explore how experimental noise and gate errors affect our algorithm and whether we can develop noise-mitigation techniques tailored to our approach. By replacing heuristic stopping rules with principled run-time verification and gap-conditional certification, our approach opens a systematic route toward reliable variational algorithms in the NISQ and early fault-tolerant era.

**Acknowledgments.** B.Ž. acknowledges financial support from the Slovenian Research and Innovation Agency (ARIS) through Project No. J2-60034 and the UL VIP project (KTTK21) under Contract No. SN-ZRD/22-27/510, and computational resources by SLING, a Slovenian national supercomputing network. The authors thank Conor Mc Keever and Matthias Rosenkranz for carefully reading the manuscript and providing helpful feedback. We used the AI tools DeepSeek 4, Kimi 3, Claude Opus 5, Claude Fable 5, Astra, and Sol to check the text for missing definitions and inconsistencies and to assist in improving the proofs, in particular, the finite-shot analysis. The authors directed all such use, verified every resulting statement and derivation, and take full responsibility for the content of the Letter.

**Data availability.** There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

- [1] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Variational quantum algorithms, *Nat. Rev. Phys.* **3**, 625 (2021).
- [2] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik,

Noisy intermediate-scale quantum algorithms, *Rev. Mod. Phys.* **94**, 015004 (2022).

- [3] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth, and J. Tennyson, The variational quantum eigensolver: A review of methods and best practices, *Phys. Rep.* **986**, 1 (2022).

- [4] M. Larocca, S. Thanasilp, S. Wang, K. Sharma, J. Biamonte, P. J. Coles, L. Cincio, J. R. McClean, Z. Holmes, and M. Cerezo, Barren plateaus in variational quantum computing, *Nat. Rev. Phys.* **7**, 174 (2025).
- [5] E. R. Anschuetz and B. T. Kiani, Quantum variational algorithms are swamped with traps, *Nat. Commun.* **13**, 7760 (2022).
- [6] A. Gheorghiu, T. Kapourniotis, and E. Kashefi, Verification of quantum computation: An overview of existing approaches, *Theory Comput. Syst.* **63**, 715 (2019).
- [7] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, Quantum computation by adiabatic evolution, [arXiv:quant-ph/0001106](https://arxiv.org/abs/quant-ph/0001106).
- [8] T. Albash and D. A. Lidar, Adiabatic quantum computation, *Rev. Mod. Phys.* **90**, 015002 (2018).
- [9] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, [arXiv:1411.4028](https://arxiv.org/abs/1411.4028).
- [10] D. Wecker, M. B. Hastings, and M. Troyer, Progress towards practical quantum variational algorithms, *Phys. Rev. A* **92**, 042303 (2015).
- [11] A. Garcia-Saez and J. I. Latorre, Addressing hard classical problems with adiabatically assisted variational quantum eigensolvers, [arXiv:1806.02287](https://arxiv.org/abs/1806.02287).
- [12] M.-C. Chen, M. Gong, X. Xu, X. Yuan, J.-W. Wang, C. Wang, C. Ying, J. Lin, Y. Xu, Y. Wu, S. Wang, H. Deng, F. Liang, C.-Z. Peng, S. C. Benjamin, X. Zhu, C.-Y. Lu, and J.-W. Pan, Demonstration of adiabatic variational quantum computing with a superconducting quantum coprocessor, *Phys. Rev. Lett.* **125**, 180501 (2020).
- [13] S. M. Harwood, D. Trenev, S. T. Stober, P. Barkoutsos, T. P. Gujarati, S. Mostame, and D. Greenberg, Improving the variational quantum eigensolver using variational adiabatic quantum computing, *ACM Trans. Quantum Comput.* **3**, 1 (2022).
- [14] B. F. Schiffer, J. Tura, and J. I. Cirac, Adiabatic spectroscopy and a variational quantum adiabatic algorithm, *PRX Quantum* **3**, 020347 (2022).
- [15] I. Kolotouros, I. Petrongonas, M. Prokop, and P. Wallden, Simulating adiabatic quantum computing with parameterized quantum circuits, *Quantum Sci. Technol.* **10**, 015003 (2025).
- [16] C. Sanavio, F. Mascherpa, A. Marruzzo, A. Amendola, and S. Succi, Variational-adiabatic quantum solver for systems of linear equations with warm starts, [arXiv:2505.24285](https://arxiv.org/abs/2505.24285).
- [17] N. N. Hegade, K. Paul, Y. Ding, M. Sanz, F. Albarrán-Arriagada, E. Solano, and X. Chen, Shortcuts to adiabaticity in digitized adiabatic quantum computing, *Phys. Rev. Appl.* **15**, 024038 (2021).
- [18] Z. Zhan, C. Run, Z. Zong, L. Xiang, Y. Fei, Z. Sun, Y. Wu, Z. Jia, P. Duan, J. Wu, Y. Yin, and G. Guo, Experimental determination of electronic states via digitized shortcut to adiabaticity and sequential digitized adiabaticity, *Phys. Rev. Appl.* **16**, 034050 (2021).
- [19] P. Chandarana, N. N. Hegade, K. Paul, F. Albarrán-Arriagada, E. Solano, A. del Campo, and X. Chen, Digitized-counterdiabatic quantum approximate optimization algorithm, *Phys. Rev. Res.* **4**, 013141 (2022).
- [20] N. N. Hegade, X. Chen, and E. Solano, Digitized counterdiabatic quantum optimization, *Phys. Rev. Res.* **4**, L042030 (2022).
- [21] E. Karacan, C. Mc Keever, M. Foss-Feig, D. Hayes, and M. Lubasch, Filter-enhanced adiabatic quantum computing on a digital quantum processor, *Phys. Rev. Res.* **7**, 033153 (2025).
- [22] S. V. Romero, A.-M. Visuri, A. G. Cadavid, A. Simen, E. Solano, and N. N. Hegade, Bias-field digitized counterdiabatic quantum algorithm for higher-order binary optimization, *Commun. Phys.* **8**, 348 (2025).
- [23] B. A. Bhargava, S. Kumar, A.-M. Visuri, P. A. Erdman, E. Solano, and N. N. Hegade, Constant depth digital-analog counterdiabatic quantum computing, [arXiv:2601.01154](https://arxiv.org/abs/2601.01154).
- [24] A. Cervera-Lierta, J. S. Kottmann, and A. Aspuru-Guzik, Meta-variational quantum eigensolver: Learning energy profiles of parameterized Hamiltonians for quantum simulation, *PRX Quantum* **2**, 020329 (2021).
- [25] P. Chandarana, P. S. Vieites, N. N. Hegade, E. Solano, Y. Ban, and X. Chen, Meta-learning digitized-counterdiabatic quantum optimization, *Quantum Sci. Technol.* **8**, 045007 (2023).
- [26] J. Miao, C.-Y. Hsieh, and S.-X. Zhang, Neural-network-encoded variational quantum algorithms, *Phys. Rev. Appl.* **21**, 014053 (2024).
- [27] C. Mc Keever and M. Lubasch, Towards adiabatic quantum computing using compressed quantum circuits, *PRX Quantum* **5**, 020362 (2024).
- [28] R. Puig, M. Drudis, S. Thanasilp, and Z. Holmes, Variational quantum simulation: A case study for understanding warm starts, *PRX Quantum* **6**, 010317 (2025).
- [29] H. Mhiri, R. Puig, S. Lerch, M. S. Rudolph, T. Chotibut, S. Thanasilp, and Z. Holmes, A unifying account of warm start guarantees for patches of quantum landscapes, [arXiv:2502.07889](https://arxiv.org/abs/2502.07889).
- [30] M. Ballarin, J. J. García-Ripoll, D. Hayes, and M. Lubasch, Efficient quantum state preparation of multivariate functions using tensor networks, [arXiv:2511.15674](https://arxiv.org/abs/2511.15674).
- [31] R. Puig, B. Casas, A. Cervera-Lierta, Z. Holmes, and A. Pérez-Salinas, Warm starts, cold states: Exploiting adiabaticity for variational ground-states, [arXiv:2602.06137](https://arxiv.org/abs/2602.06137).
- [32] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).
- [33] J. Stokes, J. Izaac, N. Killoran, and G. Carleo, Quantum natural gradient, *Quantum* **4**, 269 (2020).
- [34] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/m538-45pk> for additional information on the proofs of Theorems 1 and 2, trainability within the tracking region, extension to  $\epsilon$ -exact representability, numerical analysis, and measurement cost and stability under shot noise, which include Refs. [44,45].
- [35] Y. Nesterov, *Lectures on Convex Optimization*, Springer Optimization and Its Applications Vol. 137 (Springer, Cham, 2018).
- [36] L. Hackl, T. Guaita, T. Shi, J. Haegeman, E. Demler, and J. I. Cirac, Geometry of variational methods: Dynamics of closed quantum systems, *SciPost Phys.* **9**, 048 (2020).
- [37] B. Žunkovič, P. Torta, G. Pecci, G. Lami, and M. Collura, Variational ground-state quantum adiabatic theorem, *Phys. Rev. Lett.* **134**, 130601 (2025).
- [38] T. Kadowaki and H. Nishimori, Quantum annealing in the transverse Ising model, *Phys. Rev. E* **58**, 5355 (1998).
- [39] P. Hauke, H. G. Katzgraber, W. Lechner, H. Nishimori, and W. D. Oliver, Perspectives of quantum annealing: Methods and implementations, *Rep. Prog. Phys.* **83**, 054401 (2020).

- [40] 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, Self-verifying variational quantum simulation of lattice models, *Nature (London)* **569**, 355 (2019).
- [41] Y. Subaşı, R. D. Somma, and D. Orsucci, Quantum algorithms for systems of linear equations inspired by adiabatic quantum computing, *Phys. Rev. Lett.* **122**, 060504 (2019).
- [42] P. C. S. Costa, D. An, Y. R. Sanders, Y. Su, R. Babbush, and D. W. Berry, Optimal scaling quantum linear-systems solver via discrete adiabatic theorem, *PRX Quantum* **3**, 040303 (2022).
- [43] P. C. S. Costa, D. An, R. Babbush, and D. Berry, The discrete adiabatic quantum linear system solver has lower constant factors than the randomized adiabatic solver, *Quantum* **9**, 1887 (2025).
- [44] W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Am. Stat. Assoc.* **58**, 13 (1963).
- [45] D. Hsu, S. Kakade, and T. Zhang, A tail inequality for quadratic forms of subgaussian random vectors, *Electron. Commun. Probab.* **17**, 1 (2012).