Effective temperature in approximate quantum many-body states

Yu-Qin Chen[1,*] and Shi-Xin Zhang [2,†]

[1]*Graduate School of China Academy of Engineering Physics, Beijing 100193, China*
[2]*Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China*

In the pursuit of numerically identifying ground states of quantum many-body systems, approximate quantum wave function ansatzes are commonly employed. This study focuses on the spectral decomposition of these approximate quantum many-body states into exact eigenstates of the target Hamiltonian, which reflects the intricate physics at the interplay between quantum systems and numerical algorithms. Here, we examine various parametrized approximate quantum states constructed from neural networks, tensor networks, and quantum circuits, employing differentiable programming to numerically approximate ground states and imaginary-time evolved states. Our findings reveal a consistent exponential pattern in the energy eigenbasis decomposition contributions of approximate states across different ansatzes, optimization objectives, and quantum systems, characterized by remarkably small decay rates, i.e., *high* effective temperatures. This finding is counterintuitive for high-fidelity approximate ground states: While the total contribution from excited states can be made sufficiently small, the residual spectral weight does not decay rapidly with energy. This behavior is an intrinsic property of the variational ansatz, independent of the approximation's overall accuracy. The effective temperature is related to ansatz expressiveness and accuracy and shows phase transition behaviors in learning imaginary-time evolved states. The universal picture and unique features suggest the significance and potential of the effective temperature metric in characterizing approximate quantum states.

I. INTRODUCTION

Understanding the ground-state properties of quantum many-body systems is a central challenge in modern physics, with broad implications ranging from fundamental principles of quantum complexity theories to the design of materials and quantum technologies [1–4]. The exponentially large Hilbert space with the system size often precludes analytical solutions and exact numerical methods, necessitating the development of approximate numerical methods, and approximate quantum wave function ansatzes have proven to be powerful for their good trade-off between expressivity and complexity.

In this paper, we investigate a profound and previously unexplored aspect of approximate quantum states: the effective temperature that characterizes their spectral properties. We apply a concept analogous to temperature, arising from statistical mechanics [5], to characterize the spectral distribution of quantum many-body pure states, specifically the approximate states obtained through optimizing different ansatzes.

The success of variational methods is typically measured by energy accuracy or fidelity. However, less is known about

the spectral structure of the residual error that persists even in high-fidelity approximations. This work investigates the residue error structure and reveals a surprising universality in the effective temperatures of approximate quantum ground states, irrespective of the ansatz structure, objective function, training wellness, or underlying physical system. By decomposing the approximate states into the exact eigenstates of the system Hamiltonian, we observe a spectrum extending to the high-energy end with exponential decay of small decay factors, interpreted as inverse effective temperatures. The consistently high or even negative effective temperatures (inverse temperature $\tilde{\beta} < 0.3$ for most cases) imply a nearly flat spectrum contribution; namely, we cannot directly ignore the excited-state components at high-energy regimes and the consequent presence of spectral weight at high energies has practical implications for calculating physical observables such as high orders of the Hamiltonian operator, which challenges prevailing notions about variational approximation error. The effective temperature also shows an intriguing two-stage behavior when approximating pure states of finite "temperature.". The effective temperature offers a lens to assess the accuracy and reliability of different variational ansatzes and algorithms and the universal pattern could play an important role in understanding complex quantum systems or designing efficient methods.

II. SETUP

A wave function ansatz comprises a family of parametrized quantum states $|\psi(\theta)\rangle$ with tunable parameters θ. Such

*Contact author: yqchen@gscaep.ac.cn
†Contact author: shixinzhang@iphy.ac.cn

quantum states usually admit a compact and scalable representation on classical computers or quantum computers to circumvent the challenge of the exponentially large Hilbert space. The approximate state $|\psi\rangle$ within the manifold of a given ansatz can be obtained through variational optimization [6] $\min_\theta \mathcal{L}(|\psi(\theta)\rangle)$, where $\mathcal{L}$ is the objective function built on top of the wave function. Specifically, energy fidelity $\mathcal{L}_E = \langle\psi(\theta)|H|\psi(\theta)\rangle$ is often utilized to obtain the ground state of the system Hamiltonian H. The Hamiltonian has an exact energy spectral decomposition as $H = \sum_{i=0} \varepsilon_i |\varepsilon_i\rangle\langle\varepsilon_i|$, with $|\varepsilon_0\rangle$ being the exact ground state. Additionally, we also frequently utilize the fidelity objective $\mathcal{L}_F = 1 - |\langle\varepsilon_0|\psi(\theta)\rangle|^2$ for numerically approaching ground states. Unlike the energy objective, minimizing fidelity objective penalizes deviation from the ground state equally for all excited states, offering no inherent preference for suppressing higher-energy excitations more strongly than lower-energy ones. Throughout this work, we perform optimization via gradient descent, namely, the variational parameters are updated in each step according to $\Delta\theta \propto \partial\mathcal{L}/\partial\theta$, where the gradients can be efficiently evaluated using automatic differentiation [7–10].

For an approximate state $|\psi\rangle$, we decompose it on the basis of the exact eigenstates $|\varepsilon_i\rangle$ of Hamiltonian H:

$$|\psi\rangle = \sum_i c_i |\varepsilon_i\rangle. \tag{1}$$

The spectral decomposition coefficients $|c_i|^2$ are the central focus of our study. We find that the pairs of $(\varepsilon_i, |c_i|^2)$ statistically follow an exponential decay for the approximate ground state, analogous to Boltzmann distribution,

$$|c_i|^2 \propto e^{-\tilde{\beta}\varepsilon_i}, \tag{2}$$

where the decay factor $\tilde{\beta}$ is defined as the inverse effective temperature for the pure state $|\psi\rangle$ under investigation.

To gain deeper insights into the spectrum pattern of approximate quantum states, we evaluate additional target states beyond the ground state $|\varepsilon_0\rangle$. This set of states is directly parametrized with an inverse effective temperature β as

$$|\phi(\beta)\rangle = \frac{1}{Z(\beta)} \sum_i e^{-\beta\varepsilon_i/2} |\varepsilon_i\rangle, \tag{3}$$

where $Z(\beta) = \sqrt{\sum_i e^{-\beta\varepsilon_i}}$ is the normalization factor. These states can be understood as imaginary-time evolved states with time β from the initial state $|\phi(0)\rangle = 1/Z(0) \sum_i |\varepsilon_i\rangle$. Therefore, we call the target states imaginary-time evolved states (ITES). Exact ITES by definition admits an exponential decay for spectral decomposition with inverse temperature β. We investigate how spectral properties change for approximate ITES $|\psi\rangle$ obtained by optimizing the fidelity objective $\mathcal{L}_F(\beta) = 1 - |\langle\phi(\beta)|\psi(\theta)\rangle|^2$. This objective reduces to the ground-state target in the $\beta \to \infty$ limit. We find that the approximate ITES spectrum shows an evident two-stage behavior for all ansatzes: $\tilde{\beta}$ from approximate ITES successfully matches β for small β (high-temperature regime) and shows strong deviation $\tilde{\beta} < \beta$ for large β (low-temperature regime). More importantly, the transition point β^* characterizes the lowest effective temperature a given ansatz can reach and implies the intrinsic power of the corresponding ansatz. ITES and their approximation offer a unique platform to explore the

interplay between spectral properties, ansatz expressiveness, and optimization hardness. Notably, the conclusion remains qualitatively the same when the coefficient for each eigenstate in ITES acquires an extra random phase for generality, i.e., $|\phi(0)\rangle = 1/Z(0) \sum_i e^{i\omega_i} |\varepsilon_i\rangle$.

III. METHODS

We employ various quantum wave function ansatzes to demonstrate the universality of our findings, including (1) tensor-network-based ansatz including matrix product states (MPS) [11–13] and projected entangled-pair states (PEPS) [14,15], (2) neural quantum states (NQS) built on top of neural networks [16,17], (3) output states from variational quantum circuits [18,19] such as variational quantum eigensolvers (VQE), [20,21], and (4) vector state ansatz (VEC), where each wave function component is directly modeled as a trainable parameter. We also comment on the relevance of our results to quantum approximate optimization algorithms (QAOA) [22] (see the Supplemental Material for details on these ansatz structures [23]). We apply these ansatzes to different system Hamiltonians on different lattice geometries and optimize both objectives $\mathcal{L}_E$ and $\mathcal{L}_F$. The main findings of this paper are summarized in Fig. 1.

IV. RESULTS FOR APPROXIMATE GROUND STATES

Since producing the exact spectrum requires a full diagonalization of the Hamiltonian, our numerical study focuses on systems up to size $L = 16$. We note that the effective temperature can be related to measurable quantities, enabling its estimation beyond exact diagonalization limits (see the Supplemental Material [23]). We use TensorCircuit-NG [24] for the numerical simulation of variational training and QuSpin [25] for the exact diagonalization of the Hamiltonian within full Hilbert space or specific charge sectors.

We focus on the two-dimensional (2D) XXZ model on a square lattice with spin-1/2 degrees of freedom and periodic boundary conditions:

$$H = \sum_{\langle ij\rangle} J_x X_i X_j + J_y Y_i Y_j + J_z Z_i Z_j, \tag{4}$$

where $X(Y, Z)_i$ are Pauli matrices on lattice site i and $\langle ij\rangle$ denotes the nearest-neighbor sites. The phase diagram of this model features $J_z^c = 1$ that separates the spin-flipping phase and the antiferromagnetic phase when $J_x = J_y = 1$ [26].

Using different ansatzes and variational optimization against $\mathcal{L}_E$ or $\mathcal{L}_F$, we obtain different approximate ground states. The typical spectral decomposition patterns $(\varepsilon_i, |c_i|^2)$ of these states are shown in Figs. 2(a) and 2(b) (see the SupplementalMaterial for more spectral decomposition results from other variational ansatzes [23]). We find a robust exponential relation in the spectral decomposition across different ansatzes and training stages. The exponential decay factor $\tilde{\beta}$ represents inverse effective temperature. The training dynamics of $\tilde{\beta}$ for different ansatzes and objectives is shown in Fig. 3. We find that the effective temperature $1/\tilde{\beta}$ is high during training, and this is particularly true for optimized approximate ground states where $\tilde{\beta} < 0.3$ in most cases. Such a small $\tilde{\beta}$ unexpectedly contrasts sharply with imaginary-
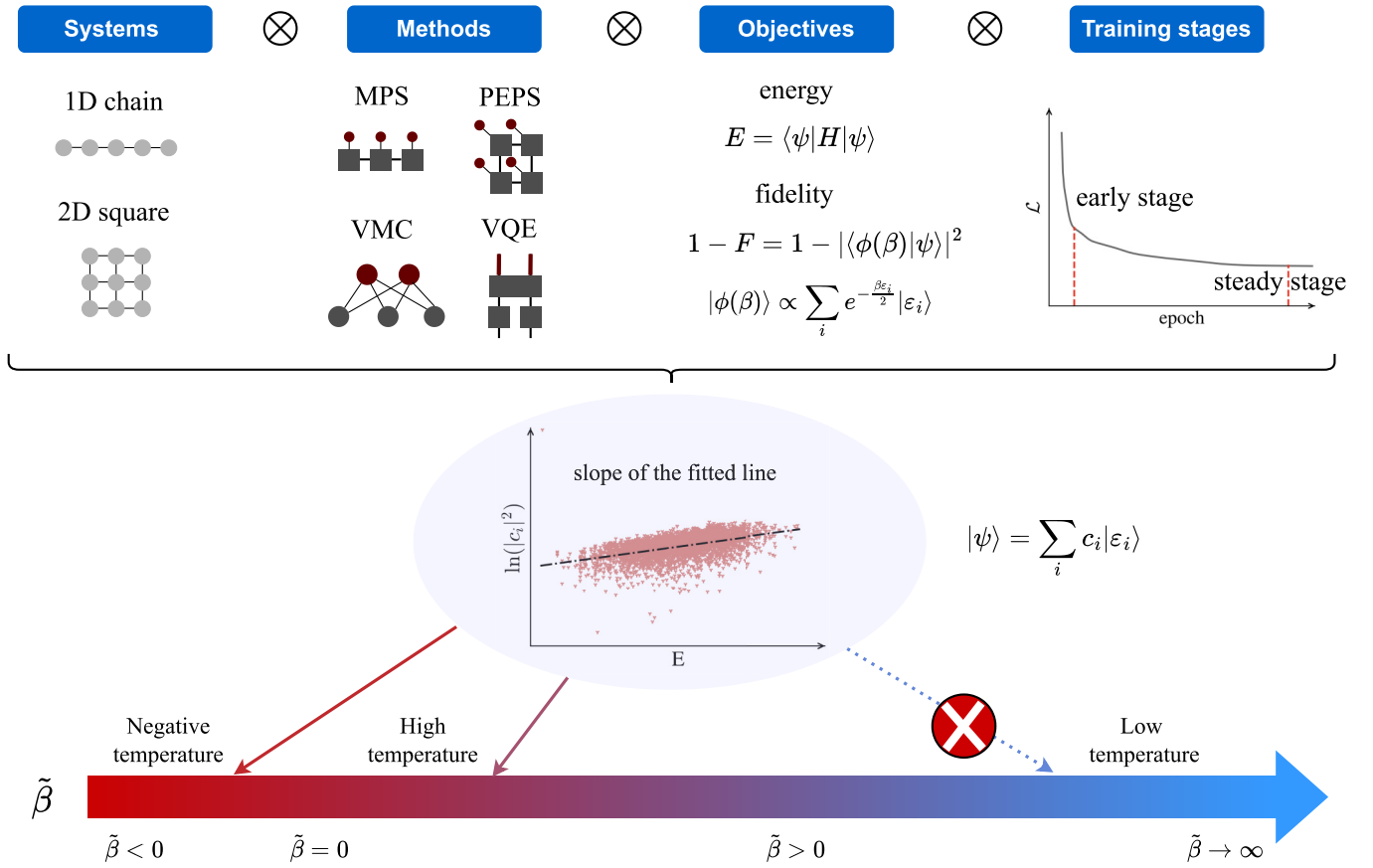
FIG. 1. Sketch of the effective temperature as a metric characterizing different systems and numerical methods. The effective temperature $(1/\tilde{\beta})$ can be extracted from the estimated slope of the approximate quantum state spectral decomposition (center inset). The metric varies with different systems, methods, objectives, and training wellness. However, universal features remain the same and are valuable in diagnosing the capacity of different numerical methods. In general, $\tilde{\beta}$ is small and can be even negative, indicating a nearly flat excited-state spectrum of approximate ground states.

time evolution-based approaches where a large $\tilde{\beta}$ is expected. During variational training, a general trend emerges where $\tilde{\beta}$ first increases and then decreases, implying an upper bound for the possible $\tilde{\beta}$ of a given ansatz. The initial increase in $\tilde{\beta}$ corresponds to the optimizer quickly finding the correct low-energy manifold of the Hilbert space. The subsequent decrease in $\tilde{\beta}$ reflects the difficulty of fine-grained optimization within that manifold, where further reducing the energy often involves complex, nonlocal reconfigurations of the wave function that can slightly "heat up" the spectral tail, rendering the residual spectral error more complex and "flatter." Besides, training dynamics of $\tilde{\beta}$ show similar patterns for both objectives within the same ansatz, while $\tilde{\beta}$ for optimizing energy objectives $\mathcal{L}_E$ is typically higher due to the unequal penalty on different excited states. Conversely, fidelity objective equally penalizes all excited states and the nonzero $\tilde{\beta}$ in this case is highly nontrivial, which can be attributed to the physical prior encoded in the ansatz structures. This understanding is validated by the near-zero $\tilde{\beta}$ for VEC ansatz, which contains the least physical prior in the structure. Furthermore, the $\tilde{\beta}$ trends are more closely aligned within the same family of ansatzes, reflecting its potential power to diagnose intrinsic properties in these ansatzes. The universality and validness of the results are also confirmed with other system sizes, models,

and lattice geometries (see the Supplemental Material for details [23]).

A crucial question is whether these findings are independent of the specific optimization algorithm used. To address this, we performed additional simulations using the standard imaginary time evolution algorithm time-evolving block decimation (TEBD). Crucially, even states prepared via TEBD—a method specifically designed to suppress high-energy components—exhibit a persistent high effective temperature (see Supplemental Material for details [23]). This provides strong evidence that the slow spectral decay is not an optimization flaw but a fundamental characteristic of the variational ansatz itself, where the restricted manifold leaves a "hot" imprint on the spectral properties.

V. RESULTS FOR APPROXIMATE IMAGINARY-TIME EVOLVED STATES

To better understand the spectrum pattern and the underlying mechanism of effective temperature in approximate ground states, we evaluate approximate ITES whose targets have strictly exponential decay spectrum patterns. For different target β's, we identify a two-stage behavior with a putative phase transition or crossover in between. For the
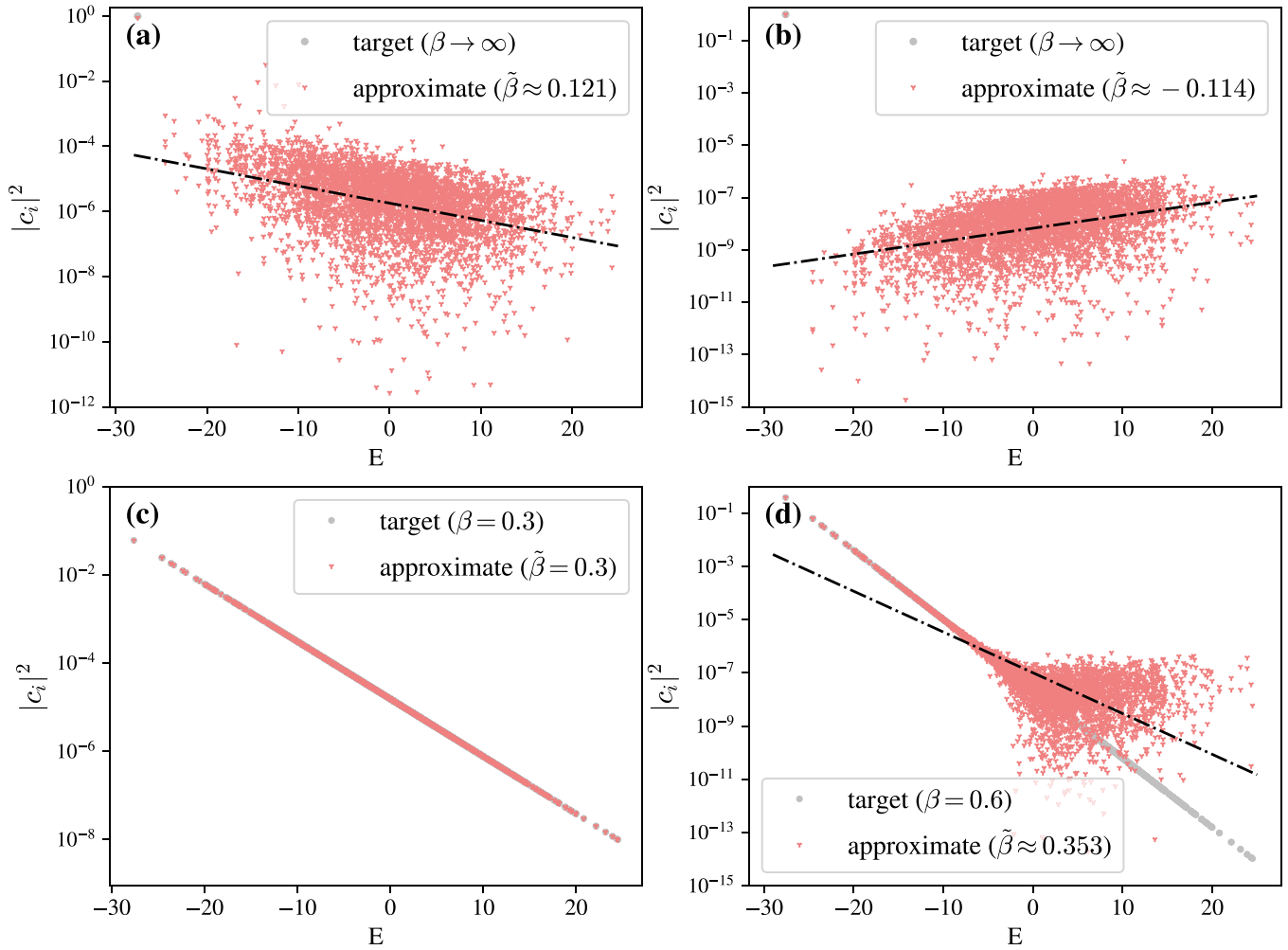
FIG. 2. Spectral decomposition of approximate states. The 2D XXZ model on a 4×3 square lattice with $J_x = J_y = 1$, $J_z = 0.8$ and MPS ansatz with bond dimension $\chi = 32$ are employed. The optimization objective is the infidelity between approximate states and the target states. The target states are chosen as the ground states for panels (a) and (b) and imaginary-time evolved states $|\phi(\beta)\rangle$ with $\beta = 0.3$ (high temperature) (c) and $\beta = 0.6$ (low temperature) (d). Overlaps with different energy eigenstates are depicted as scatter plots in red and gray for approximate and target states, respectively. The optimization is at the early stage for panel (a) and converged for panels (b)–(d) at the steady stage. The logarithmic overlaps with excited eigenstates show a linear pattern during training with varying fitted slope $\tilde{\beta}$, as known as. inverse effective temperature. The fitted line is shown in dashdot. When the target state is ITES, there is a phase transition in the spectrum patterns of the approximate state. For small β in panel (c), the optimized approximate state exhibits a near-perfect correspondence with the target state overlap, with a fitted slope $\tilde{\beta}$ matching $\beta = 0.3$. Conversely, for large β in panel (d), the overlaps from approximate states exhibit distinct behaviors—an exponential decay in the lower-energy regime and a plateau in the higher-energy regime. This behavior leads to a poor linear fit, characterized by a deviating slope $\tilde{\beta} < \beta$.

high-temperature regime of small β, the approximate ITES conforms to the strict exponential decay spectrum pattern with $\tilde{\beta} = \beta$, as shown in Fig. 2(c). On the contrary, for the low-temperature regime of large β, the spectral decomposition follows the exponential decay only for the low-energy part, while a nearly flat spectrum emerges from high-energy contributions. If we attempt to extract a single exponential decay factor $\tilde{\beta}$ in this case, its value is significantly smaller than the exact β, as shown in Fig. 2(d). It is worth noting that the nearly flat spectrum deviation is not due to numerical machine precision, as points $|c_i|^2$ of the same magnitude conform to the exponential decay for small β but fail to do so for large β (see the Supplemental Material for details [23]).

The results of $\tilde{\beta}$ of approximate ITES for different target β's are shown in Fig. 4(a), where the two-stage behavior is evident. For each ansatz, there is a transition point β^* when $\tilde{\beta}$ begins to deviate from β and the associated spectrum pattern moves from Fig. 2(c) to Fig. 2(d). β^* is a figure of merit as it marks the critical temperature that the ansatz can accurately represent and sets an upper bound on $\tilde{\beta}$ that an ansatz can reach. More importantly, β^* is also correlated with the expressive capacity of the ansatz, as a higher β^* generally implies a better fidelity. In the low-temperature limit $\beta \to \infty$, the results for approximate ITES are reduced to the results for approximate ground states by optimizing $\mathcal{L}_f$.

We further remark on the fidelity results shown in Fig. 4(b) with varying methods and β's. For low-accuracy ansatzes,
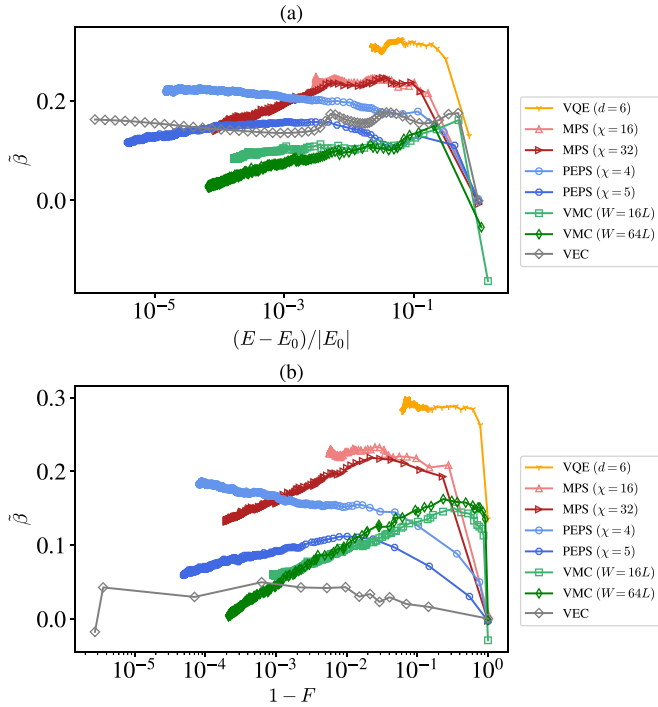
FIG. 3. Effective temperature during variational training targeting the ground states. The objectives are energy $\mathcal{L}_E$ (a) and fidelity $\mathcal{L}_F$ (b), respectively. The results are obtained from the 2D XXZ model on a 4×4 square lattice with $J_x = J_y = 1$, $J_z = 0.8$. During training, points move to the left toward higher accuracy and $\tilde{\beta}$ first increases and then decreases in general. This nonmonotonic "rise-and-fall" trajectory for the inverse temperature is a key feature of the training dynamics, reflecting the difficulty of perfectly suppressing all high-energy components within a restricted manifold.

FIG. 4. Effective temperature and fidelity of approximate ITES with different β's. The results are obtained from the 2D XXZ model on a 4×3 square lattice with $J_x = J_y = 1$, $J_z = 0.8$. (a) The inverse effective temperature $\tilde{\beta}$ shows a two-stage behavior with varying β. β^* (insets) marks the deviation of $\tilde{\beta}$ from target β, separating the two stages. (b) The fidelity of approximate ITES with different β's. For high-accuracy methods, the fidelity decreases with increasing β due to the increased optimization hardness. For low-accuracy methods, the fidelity increases with increasing β as the target states are less entangled and easier to approximate.

the fidelity gets improved with increasing β as $|\phi(\beta)\rangle$ of larger β is closer to the ground state and has smaller entanglement. Such less entangled states can be better suited in ansatzes of limited expressiveness. Conversely, for high-accuracy ansatzes, the expressive capacity is sufficient for targeting any ITES, and the bottleneck is the optimization difficulty. Since the number of optimization steps is fixed in our simulation, a slower optimization speed results in a worse final fidelity with increasing β. In other words, when the target is closer to the ground-state manifold, the optimization landscape is harder to navigate (see quantitative analysis for the expressiveness and optimization hardness in Supplemental Material [23]).

VI. DISCUSSIONS AND CONCLUSION

Some hints of the general picture presented in this paper are previously reported in QAOA [27,28] and MPS cases [29], where only the optimization of $\mathcal{L}_E$ is considered. In the QAOA case, the output states were found to approximate Gibbs distribution with $\tilde{\beta} \approx 0.2-0.3$ [27], consistent with the findings in this study, as QAOA can be regarded as a special limit of VQE ansatz. In the MPS case, a nearly flat spectrum is identified [29]. Given that the density of states was effectively considered in their work, the spectrum appeared flatter as the
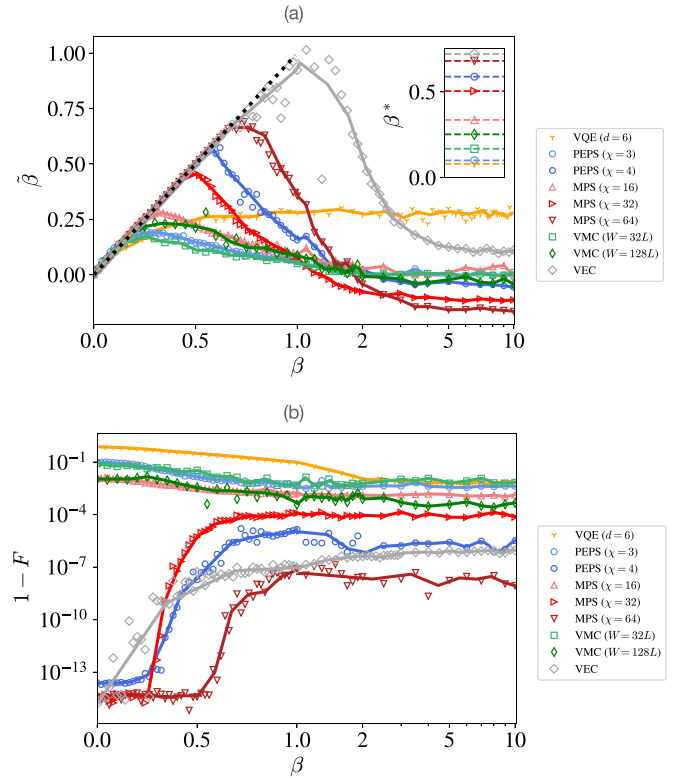
exponential decay of small positive $\tilde{\beta}$ was compensated with the high density of states in the middle-energy regime.

The effective temperature has profound and practical implications for numerical methods and many-body systems. For instance, the presence of a significant high-energy tail in the spectral distribution, a direct consequence of small $\tilde{\beta}$, can lead to unexpected fragility in variance estimation as mentioned in Ref. [29]. The variance H^2 is just an example for operator $f(H)$, with expectation $\langle\psi|f(H)|\psi\rangle = \sum_i |c_i|^2 f(\varepsilon_i)$. The operator expectation is highly sensitive to tiny high-energy components when $f(\varepsilon)$ is significantly larger for high-energy inputs. The expectation of $f(H) = e^{\alpha H}$ not only exhibits experimentally observable effects but also provides a practical route to estimate $\tilde{\beta}$ in a scalable way beyond exact diagonalization sizes. As elaborated in Supplemental Material [23], the effective temperature $\tilde{\beta}$ leaves an observable signature on the expectation value of operators such as $e^{\alpha H}$. Specifically, the function $\langle\psi|e^{\alpha H}|\psi\rangle$ as a function of α is minimized at $\alpha = \tilde{\beta}$. Through this scalable proxy, we successfully give the estimate on effective temperature for approximate matrix product states obtained from density matrix renormalization group in large-scale systems ($L = 128$). We found that the

resulting high-fidelity ground states still possess an extremely small inverse effective temperature ($\tilde{\beta} < 0.01$), confirming our findings well beyond exact diagonalization regime.

The persistence of the exponential decay spectrum, particularly under the fidelity objective that lacks an intrinsic bias against high-energy states, suggests that the effective temperature is a fundamental characteristic imprinted by the variational ansatz manifold rather than solely by the target Hamiltonian's energy landscape. The effective temperature also has strong relevance to the expressive capacity of ansatzes as β^* values differ for different ansatzes. Consequently, the metric and the methodology explored in this paper serve as a guiding principle to help design better approximate quantum ansatzes and variational algorithms.

The effective temperature metric opens up several intriguing future research directions. We can quantitatively investigate the scaling of effective temperature $\tilde{\beta}$ and its critical value $\tilde{\beta}^*$ with respect to varying system sizes, Hamiltonian parameters, and ansatz structure parameters, to gain deeper insights into the physics of the system and the intrinsic patterns in numerical methods. The ITES proposed in this paper are also of independent academic interest, as they provide a platform to explore the trade-off between expressiveness and optimization hardness, as well as host a putative phase transition at β^*, the nature and universal behavior of which merit future exploration. It is also crucial to validate the general picture of spectrum patterns on more ansatzes, including mean-field ansatzes, physics-inspired ansatzes [30], and ad-

vanced hybrid variational ansatzes [31,32], and more models, including integrable systems, many-body localized systems, and fermionic systems.

In this paper, we have identified a universal pattern in the spectrum of approximate ground states. We define the effective temperature $\tilde{\beta}$ to characterize the slow exponential decay in the spectrum and investigate the behavior of $\tilde{\beta}$ with different systems, ansatzes, objectives, and training steps. To better understand the underlying mechanism for this metric, we further propose ITES targets and identify the two-stage behaviors in approximating ITES with the figure of merit critical inverse effective temperature β^*. We believe that the universal picture presented here provides a fresh and powerful perspective on benchmarking numerical algorithms and understanding quantum many-body systems.

DATA AVAILABILITY

The data that support the findings of this article are openly available [33].

[1] A. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation* (American Mathematical Society, Providence, RI, 2002).

[2] J. Kempe, A. Kitaev, and O. Regev, The complexity of the local Hamiltonian problem, SIAM J. Comput. **35**, 1070 (2006).

[3] B.-X. Zheng, C.-M. Chung, P. Corboz, G. Ehlers, M.-P. Qin, R. M. Noack, H. Shi, S. R. White, S. Zhang, and G. K.-L. Chan, Stripe order in the underdoped region of the two-dimensional Hubbard model, Science **358**, 1155 (2017).

[4] S. Lee, J. Lee, H. Zhai, Y. Tong, A. M. Dalzell, A. Kumar, P. Helms, J. Gray, Z.-H. Cui, W. Liu, M. Kastoryano, R. Babbush, J. Preskill, D. R. Reichman, E. T. Campbell, E. F. Valeev, L. Lin, and G. K.-L. Chan, Evaluating the evidence for exponential quantum advantage in ground-state quantum chemistry, Nat. Commun. **14**, 1952 (2023).

[5] W. Greiner, L. Neise, and H. Stöcker, *Thermodynamics and Statistical Mechanics* (Springer, New York, NY, 1995).

[6] D. Wu, R. Rossi, F. Vicentini, N. Astrakhantsev, F. Becca, X. Cao, J. Carrasquilla, F. Ferrari, A. Georges, M. Hibat-Allah, M. Imada, A. M. Läuchli, G. Mazzola, A. Mezzacapo, A. Millis, J. Robledo Moreno, T. Neupert, Y. Nomura, J. Nys, O. Parcollet *et al.*, Variational benchmarks for quantum many-body problems, Science **386**, 296 (2024).

[7] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Learning representations by back-propagating errors, Nature (London) **323**, 533 (1986).

[8] M. Bartholomew-Biggs, S. Brown, B. Christianson, and L. Dixon, Automatic differentiation of algorithms, J. Comput. Appl. Math. **124**, 171 (2000).

[9] H.-J. Liao, J.-G. Liu, L. Wang, and T. Xiang, Differentiable programming tensor networks, Phys. Rev. X **9**, 031041 (2019).

[10] S.-X. Zhang, Z.-Q. Wan, and H. Yao, Automatic differentiable Monte Carlo: Theory and application, Phys. Rev. Res. **5**, 033041 (2023).

[11] S. R. White, Density matrix formulation for quantum renormalization groups, Phys. Rev. Lett. **69**, 2863 (1992).

[12] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, Ann. Phys. **326**, 96 (2011).

[13] E. Stoudenmire and S. R. White, Studying two-dimensional systems with the density matrix renormalization group, Annu. Rev. Condens. Matter Phys. **3**, 111 (2012).

[14] F. Verstraete and J. I. Cirac, Renormalization algorithms for quantum-many body systems in two and higher dimensions, arXiv:cond-mat/0407066.

[15] F. Verstraete and J. I. Cirac, Valence-bond states for quantum computation, Phys. Rev. A **70**, 060302 (2004).

[16] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, Science **355**, 602 (2017).

[17] D.-L. Deng, X. Li, and S. Das Sarma, Quantum entanglement in neural network states, Phys. Rev. X **7**, 021021 (2017).

[18] 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).

[19] 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).

[20] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, Nat. Commun. **5**, 4213 (2014).

[21] 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).

[22] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, arXiv:1411.4028.

[23] See Supplemental Material at http://link.aps.org/supplemental/10.1103/zpjv-bm5c for an introduction to the different approximate quantum state ansatzes, the training dynamics towards ground state and imaginary-time evolved states, results from different optimization methods, a discussion on the observable consequences and scalable measurement of effective temperature, and the hyperparameters used for variational optimization.

[24] S.-X. Zhang, J. Allcock, Z.-Q. Wan, S. Liu, J. Sun, H. Yu, X.-H. Yang, J. Qiu, Z. Ye, Y.-Q. Chen, C.-K. Lee, Y.-C. Zheng, S.-K. Jian, H. Yao, C.-Y. Hsieh, and S. Zhang, TensorCircuit: A quantum software framework for the NISQ era, Quantum **7**, 912 (2023).

[25] P. Weinberg and M. Bukov, QuSpin: A python package for dynamics and exact diagonalisation of quantum many body systems. Part I: Spin chains, SciPost Phys. **2**, 003 (2017).

[26] S. Yunoki, Numerical study of the spin-flop transition in anisotropic spin-1/2 antiferromagnets, Phys. Rev. B **65**, 092402 (2002).

[27] P. Díez-Valle, D. Porras, and J. J. García-Ripoll, Quantum approximate optimization algorithm pseudo-Boltzmann states, Phys. Rev. Lett. **130**, 050601 (2023).

[28] P. C. Lotshaw, G. Siopsis, J. Ostrowski, R. Herrman, R. Alam, S. Powers, and T. S. Humble, Approximate Boltzmann distributions in quantum approximate optimization, Phys. Rev. A **108**, 042411 (2023).

[29] J. M. Silvester, G. Carleo, and S. R. White, Unusual energy spectra of matrix product states, arXiv:2408.13616.

[30] H. Yokoyama and H. Shiba, Variational Monte-Carlo studies of Hubbard model. I, J. Phys. Soc. Jpn. **56**, 1490 (1987).

[31] S.-X. Zhang, Z.-Q. Wan, C.-K. Lee, C.-Y. Hsieh, S. Zhang, and H. Yao, Variational quantum-neural hybrid eigensolver, Phys. Rev. Lett. **128**, 120502 (2022).

[32] X. Yuan, J. Sun, J. Liu, Q. Zhao, and Y. Zhou, Quantum simulation with hybrid tensor networks, Phys. Rev. Lett. **127**, 040501 (2021).

[33] https://github.com/sxzgroup/et.