

# Improving Generalization and Trainability of Quantum Eigensolvers via Graph Neural Encoding

Jungyun Lee<sup>1</sup> and Daniel K. Park<sup>\*1,2,3</sup>

<sup>1</sup>Department of Statistics and Data Science, Yonsei University, Seoul, Republic of Korea

<sup>2</sup>Department of Applied Statistics, Yonsei University, Seoul, Republic of Korea

<sup>3</sup>Department of Quantum Information, Yonsei University, Seoul, Republic of Korea

## Abstract

Determining Hamiltonian ground states is central to physics, chemistry, and combinatorial optimization, but is often classically intractable due to the exponential growth of Hilbert space. While convergence-guaranteed quantum eigensolvers require high-overlap initial states, standard variational quantum eigensolver (VQE) used to prepare such states requires instance-specific optimization and often suffers from barren plateaus. To address these limitations, we propose an end-to-end VQE framework that combines a graph autoencoder with a classical neural network to generate ansatz parameters that generalize across Hamiltonian families. Numerical experiments demonstrate improved generalization and trainability, with reduced test error and a significantly milder decay of gradient variance.

## 1 Introduction

Determining the ground state of a many-body Hamiltonian is a central problem across physics, chemistry, and combinatorial optimization, yet it is often classically intractable due to the exponential growth of Hilbert space with system size. These classical limitations have motivated quantum algorithms with convergence guarantees for ground-state energy estimation, such as quantum phase estimation (QPE) and quantum subspace methods (QSMs) [1–5]. However, an efficient implementation of these methods typically requires an initial state with at least inverse-polynomial ground-state overlap,  $|\langle\psi_{\text{initial}}|\psi_{\text{ground}}\rangle|^2$ . For generic initial states, this overlap typically decays exponentially with the number of qubits,  $n$ , requiring exponentially many repetitions and undermining the expected quantum advantage. Thus, high-quality initial-state preparation is essential.

Variational quantum eigensolvers (VQEs) [6,7] are natural candidates for preparing such initial states. However, their practical deployment is limited by two major challenges. First, standard VQEs lack generalization capability and must be retrained for each Hamiltonian instance, even when instances share substantial structural similarities. Second, VQEs can suffer from barren plateaus (BPs), where gradients of the cost function vanish rapidly with system size or circuit depth, severely impeding optimization [8,9]. These limitations create substantial overhead when solving many related Hamiltonians.

To address these limitations, we develop an end-to-end representation-learning-assisted VQE framework that generalizes across diverse families of one- and two-local Hamiltonians by integrating a graph autoencoder (GAE) with a classical neural network (NN) into VQE training. Once trained, the model directly generates VQE parameters for unseen Hamiltonians, producing high-overlap initial states without instance-specific optimization. By capturing both interaction topology and Hamiltonian coupling structure, our approach achieves stronger generalization than standard VQE and other prior learning-based VQE methods, such as neural-network-encoded VQE (NNVQE) [10] and exhibits a significantly milder decay of gradient variance, indicating enhanced robustness to BP effects.

---

<sup>\*</sup>dkd.park@yonsei.ac.kr

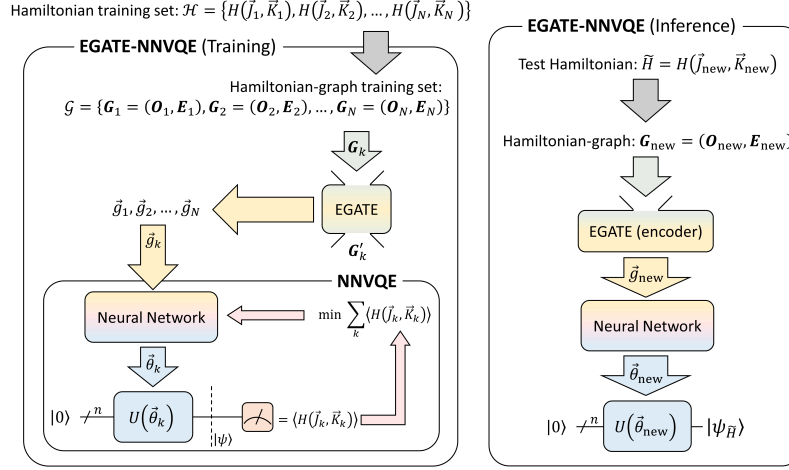


Figure 1: Schematic overview of EGATE-NNVQE for training (left) and inference (right).

## 2 Method

Figure 1 illustrates the workflow of the proposed method, which proceeds in three steps.

**Hamiltonian-to-Graph Mapping:** We consider one- and two-local Pauli Hamiltonians of the form

$$H = \sum_{\langle i,j \rangle} \sum_{\alpha, \beta \in \{x,y,z\}} J_{ij}^{\alpha\beta} \sigma_i^\alpha \sigma_j^\beta + \sum_i \sum_{\alpha \in \{x,y,z\}} K_i^\alpha \sigma_i^\alpha, \quad (1)$$

where  $J_{ij}^{\alpha\beta}$  denotes two-local interaction strengths and  $K_i^\alpha$  denotes one-local field strengths. Each Hamiltonian can be represented as a Hamiltonian-graph (H-graph), with qubits as nodes, two-local Pauli interaction coefficients as edge features, and one-local field strengths as node features. This representation captures both the interaction topology and the physical coupling parameters of the Hamiltonian, while keeping the graph size polynomial in the number of qubits.

**Graph Representation Learning:** Building on this representation, we introduce an edge-feathered graph attention autoencoder (EGATE), adapted from edge-feathered graph attention networks [11], to learn compact H-graph embeddings in an unsupervised manner. The encoder applies an attention mechanism to jointly learn node and edge features and pools them into a latent vector, while the decoder reconstructs the original node and edge features using an H-graph reconstruction loss.

**Parameter prediction using NNVQE:** After training EGATE, we discard its decoder and use the encoder to map each H-graph to a compact latent vector. This vector is fed into a classical NN, which outputs VQE ansatz parameters and is trained by minimizing the corresponding energy expectation value. Thus, the original NNVQE [10] procedure is retained, with Hamiltonian parameters,  $J_{ij}^{\alpha\beta}$  and  $K_i^\alpha$ , replaced by a learned H-graph embedding.

At inference level, a new Hamiltonian is first converted into an H-graph and encoded by the trained EGATE encoder. The trained NNVQE [10] model then directly predicts circuit parameters, which define the quantum state. This state can then be used as an initial state for subsequent quantum algorithms such as QPE or QSMs.

## 3 Results and Impact

We evaluate EGATE-NNVQE on the 1D XXZ Heisenberg spin chain without an external field,

$$H_{\text{XXZ}} = \sum_{i=1}^n (\sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + J^{zz} \sigma_i^z \sigma_{i+1}^z), \quad (2)$$

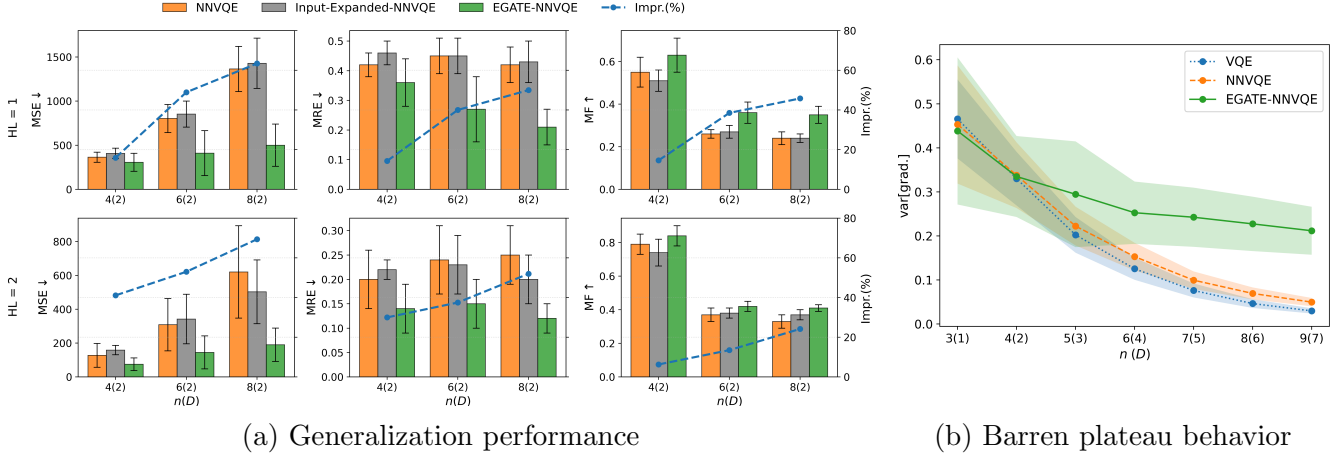


Figure 2: Numerical results for EGATE-NNVQE on the  $H_{XXZ}$ , Eq. 2. (a) EGATE-NNVQE improves generalization over NNVQE in MSE, MRE, and MF, while input-expanded NNVQE does not close the gap; HL denotes the number of hidden layers in NNVQE. (b) EGATE-NNVQE mitigates barren plateau behavior more effectively than standard VQE and NNVQE by preserving larger gradient variance as system size increases.

where periodic boundary conditions are assumed. We compare EGATE-NNVQE with the original NNVQE baseline, which does not use graph-based representation learning, and an input-expanded NNVQE which repeats the scalar Hamiltonian parameter to match the EGATE latent dimension.

Figure 2(a) shows the generalization performance on 1,000 test points in  $J^{zz} \in [-10, 10]$  for  $n = 4, 6, 8$  qubits with models trained on 20 uniformly spaced values of  $J^{zz} \in [-3, 3]$ . Performance is measured by the mean squared error (MSE) and mean relative error (MRE) of the predicted ground-state energy, where lower is better, and the mean fidelity (MF) with the exact ground state, where higher is better. The blue dashed line reports the relative improvement of EGATE-NNVQE (green) over the NNVQE (orange) baseline. Across the displayed hidden-layer (HL) settings, EGATE-NNVQE consistently improves over NNVQE in every metric. Notably, these improvements generally increase with the number of qubits, indicating that the benefit of graph representation learning becomes more pronounced as the system size increases. To test whether these gains come merely from increased input dimensionality, we also compare with input-expanded-NNVQE (gray). This baseline does not close the performance gap, indicating that the improvement comes from the informativeness of the EGATE latent representation rather than input dimension alone.

We further examine trainability using the barren plateau (BP) diagnostic in Fig. 2(b). For each system size  $n$  and ansatz block depth  $D$ , we fix  $J^{zz} = 1$  and compute the variance of ansatz-parameter gradients over 500 random initializations. Although all methods show decreasing gradient variance as  $n$  and  $D$  increase, the NNVQE baseline is nearly indistinguishable from standard VQE, whereas EGATE-NNVQE consistently exhibits larger gradient variance and a markedly slower decay. This suggests that our graph neural encoding efficiently mitigates BP effects.

Taken together, the results in Fig. 2 suggest that the advantage of EGATE-NNVQE is closely tied to its H-graph latent embedding, which preserves physical interaction structure in a compact form. This structured representation enables more accurate ansatz-parameter prediction for unseen Hamiltonians and improves trainability, thereby providing a more reliable warm start for quantum state preparation across Hamiltonian families.

*A full version of this work is available as an arXiv preprint [12].*

## References

- [1] A. Yu. Kitaev. Quantum measurements and the abelian stabilizer problem, 1995.
- [2] Daniel S Abrams and Seth Lloyd. Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors. *Physical Review Letters*, 83(24):5162, 1999.
- [3] Ethan N Epperly, Lin Lin, and Yuji Nakatsukasa. A theory of quantum subspace diagonalization. *SIAM Journal on Matrix Analysis and Applications*, 43(3):1263–1290, 2022.
- [4] Nobuyuki Yoshioka, Mirko Amico, William Kirby, Petar Jurcevic, Arkopal Dutt, Bryce Fuller, Shelly Garion, Holger Haas, Ikko Hamamura, Alexander Ivrii, et al. Krylov diagonalization of large many-body hamiltonians on a quantum processor. *Nature Communications*, 16(1):5014, 2025.
- [5] Jeffery Yu, Javier Robledo Moreno, Joseph T Iosue, Luke Bertels, Daniel Claudino, Bryce Fuller, Peter Groszkowski, Travis S Humble, Petar Jurcevic, William Kirby, et al. Quantum-centric algorithm for sample-based krylov diagonalization. *arXiv preprint arXiv:2501.09702*, 2025.
- [6] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J Love, Alán Aspuru-Guzik, and Jeremy L O’Brien. A variational eigenvalue solver on a photonic quantum processor. *Nature communications*, 5(1):4213, 2014.
- [7] Marco Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, et al. Variational quantum algorithms. *Nature Reviews Physics*, 3(9):625–644, 2021.
- [8] Jarrod R McClean, Sergio Boixo, Vadim N Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nature communications*, 9(1):4812, 2018.
- [9] Martin Larocca, Supanut Thanasilp, Samson Wang, Kunal Sharma, Jacob Biamonte, Patrick J Coles, Lukasz Cincio, Jarrod R McClean, Zoë Holmes, and M Cerezo. Barren plateaus in variational quantum computing. *Nature Reviews Physics*, pages 1–16, 2025.
- [10] Jiaqi Miao, Chang-Yu Hsieh, and Shi-Xin Zhang. Neural-network-encoded variational quantum algorithms. *Physical Review Applied*, 21(1):014053, 2024.
- [11] Ziming Wang, Jun Chen, and Haopeng Chen. Egat: Edge-featured graph attention network. In *Artificial Neural Networks and Machine Learning–ICANN 2021: 30th International Conference on Artificial Neural Networks, Bratislava, Slovakia, September 14–17, 2021, Proceedings, Part I 30*, pages 253–264. Springer, 2021.
- [12] Jungyun Lee and Daniel K Park. Improving generalization and trainability of quantum eigensolvers via graph neural encoding. *arXiv preprint arXiv:2602.19752*, 2026.