

Probing Nontrivial Betti Numbers with Variational Quantum Methods

Zherui Jerry Wang^{1,2}, Vedran Dunjko^{1,3}, Jordi Tura^{1,2} and Patrick Emonts^{1,2,4,5}

¹ $\langle aQa^L \rangle$ Applied Quantum Algorithms, Universiteit Leiden, The Netherlands

²Instituut-Lorentz, Universiteit Leiden, P.O. Box 9506, 2300 RA Leiden, The Netherlands

³LIACS, Universiteit Leiden, Niels Bohrweg 1, 2333 CA Leiden, The Netherlands

⁴Institute for Complex Quantum Systems, Ulm University, 89069 Ulm, Germany

⁵Center for Integrated Quantum Science and Technology (IQST), Ulm-Stuttgart, Germany

Abstract

Topological data analysis (TDA) extracts robust global features from high-dimensional data by representing data sets as simplicial complexes. A central challenge is to compute topological invariants of these complexes, especially the k^{th} Betti number, whose nonzero value indicates the presence of a k -dimensional hole. In this work, we study the decision problem of determining whether the k^{th} Betti number is nonzero for the clique complex of a graph G . Using a hard-core fermionic construction, we encode this problem into the ground-space structure of an associated fermionic Hamiltonian. Betti-number detection is thereby reduced to a ground-state problem in the $(k+1)$ -fermion sector. To search for low-energy states of this Hamiltonian, we introduce a progressive multi-stage optimization protocol based on variational quantum circuits. The method is benchmarked on island-graph families with analytically controlled Betti numbers and spectral gaps. Finally, we derive resource estimates for Erdős–Rényi random graphs in the regime where the associated clique complexes have nontrivial topology.

Introduction

Extracting valuable information from large-scale and high-dimensional datasets is increasingly important in modern data analysis and machine learning. Topological data analysis (TDA) has recently gained attention as a powerful framework for extracting robust global features from data using tools from algebraic topology [1–3]. One key topological invariant in TDA is the Betti number β_k , which counts the number of k -dimensional holes in a simplicial complex built from a graph constructed by connecting pairs of data points within a chosen distance threshold. However, the computational cost of classical algorithms for computing Betti numbers quickly becomes prohibitive as the homology dimension and the number of vertices increase [4–7]. In practice, the combinatorial growth of simplicial complexes leads to rapidly increasing memory and computational costs as the dimension increases, making the computation of high-dimensional Betti numbers intractable. This computational challenge has motivated a line of quantum algorithms that seek potential quantum advantage for TDA [8–15]. Further evidence for the hardness of Betti number computation comes from complexity-theoretic results, which show that even the detection of nontrivial Betti numbers is QMA₁-hard [16, 17]. This result is enabled by a mapping of Betti number computation to the problem of determining the ground-space dimension of a specific data-dependent fermionic Hamiltonian [18, 19].

This Hamiltonian formulation of Betti-number computation suggests a natural route toward near-term quantum approaches. Since a nonzero Betti number is witnessed by a zero-energy state of the Hamiltonian in the corresponding particle-number sector, the problem can be recast as a ground-state search problem. Variational quantum algorithms (VQAs) are therefore a natural candidate [19], as they have been widely developed for approximating low-energy states of Hamiltonians on near-term quantum devices [20–23].

Background and preliminaries

Topological data analysis (TDA) aims to extract global topological features from data by representing the data as a simplicial complex. In graph-based constructions, one first associates a point cloud with a graph G and then promotes this graph to its clique complex. More precisely, the clique complex $\text{Cl}(G)$ is the simplicial complex whose k -simplices are in one-to-one correspondence with the $(k+1)$ -cliques of G , i.e., complete subgraphs on $k+1$ vertices. Throughout this work, we assume that the input data have already been represented by a graph G and focus on the homology of this clique complex. The computational task studied here is the decision problem of determining whether its k^{th} Betti number is nonzero: $\beta_k(\text{Cl}(G)) > 0$.

We now introduce the fermionic hard-core model [24], which relates clique homology to quantum many-body physics. For a graph $G = (V, E)$, the hard-core Hamiltonian is

$$H_{\text{hard-core}}^{\text{f}} = \sum_{(i,j) \in E} P_i a_i^\dagger a_j P_j + \sum_{i \in V} P_i, \quad (1)$$

where $a_i^\dagger$ and a_i are the fermionic creation and annihilation operators on the i th mode, respectively, and $P_i = \prod_{j \in \mathcal{N}(i)} (1 - n_j)$. Its zero-energy states in the fermion-number- $(k+1)$ sector correspond to the k -dimensional homology classes of the independence complex of the underlying graph [16, 18, 19, 25, 26]. Since the independence complex of the complement graph $G_c = (V, E_c)$ is the clique complex of G , we apply the hard-core construction to G_c . For the variational formulation, we use

$$H_G^f = \sum_{(i,j) \in E_c} P_i a_i^\dagger a_j P_j + \sum_{i \in V} P_i + \sum_{(i,j) \in E_c} \frac{1}{2} n_i n_j. \quad (2)$$

The first two terms are the hard-core Hamiltonian on G_c . The last term is added because the HWP ansatz fixes only the fermion number and does not explicitly enforce the hard-core constraint. It penalizes configurations in which two vertices connected by a complement edge are simultaneously occupied, allowing the variational search to take place over the full fermion-number- $(k+1)$ sector while leaving the zero-energy hard-core subspace unchanged. Thus Betti-number detection is reduced to a ground-state problem: $\beta_k(\text{Cl}(G)) > 0$ if and only if $H_{G,k+1}^f$ has a zero-energy state.

Methods

In this section, we introduce a variational approach for detecting nontrivial Betti numbers. The method uses a Hamming-weight-preserving (HWP) parametrized circuit to search for low-energy states of the graph-dependent Hamiltonian H_G^f in the relevant sector [27]. Let $\lambda_{\min}(G, k+1)$ denote the smallest nonzero eigenvalue of $H_{G,k+1}^f$. Finding a state with energy below this threshold certifies the presence of a zero-energy state, and hence indicates that the k^{th} Betti number is nontrivial.

We use layered parametrized quantum circuits with periodic structures of the form $U(\boldsymbol{\theta}) = \prod_{l=1}^L U_l(\boldsymbol{\theta}^{(l)})$. Each layer $U_l(\boldsymbol{\theta}^{(l)})$ is built from nearest-neighbor two-qubit gates arranged in a HWP gate fabric structure. To implement the variational procedure on qubit-based devices, we map the fermionic Hamiltonian H_G^f to a Pauli operator using the Jordan–Wigner transformation, which we denote by $\pi(\cdot)$. Under this mapping, the w -fermion sector is identified with the qubit subspace of Hamming weight w . We initialize the variational circuit with the weight- w Dicke state $|D_w^n\rangle = \binom{n}{w}^{-1/2} \sum_{x \in \{0,1\}^n, \text{wt}(x)=w} |x\rangle$.

To study the k^{th} Betti number of the clique complex of a graph $G = (V, E)$, we work in the $(k+1)$ -excitation sector. If the optimized variational energy satisfies

$$\mathcal{C}_{G,k+1}(\boldsymbol{\theta}) = \langle D_{k+1}^n | U^\dagger(\boldsymbol{\theta}) \pi(H_G^f) U(\boldsymbol{\theta}) | D_{k+1}^n \rangle < \lambda_{\min}(G, k+1), \quad (3)$$

then the variational state must have nonzero overlap with the zero-energy subspace, which certifies the existence of a nontrivial k -dimensional homology class.

We minimize the cost function $\mathcal{C}_{G,k+1}(\boldsymbol{\theta})$ using a progressive multi-stage optimization (PMO) protocol, inspired by layerwise optimization strategies for mitigating training difficulties of parametrized quantum circuits [28–30]. Rather than optimizing the full-depth circuit from the beginning, PMO activates the ansatz progressively and reuses parameters obtained at earlier optimization stages. Inspired by these developments, we introduce a PMO protocol tailored to our Hamming-weight-preserving ansatz. Beyond standard layerwise growth, PMO combines several consecutive refinement stages: the circuit is first grown and optimized layer by layer, then improved through sweep-based local updates, followed by a gradual unfreezing of previously fixed parameters, and finally a global optimization over all variational parameters.

Results

In this section, we present numerical and theoretical results for using the PMO protocol to detect nontrivial Betti numbers. We consider two classes of graph instances. First, we numerically benchmark the method on the so-called island graphs, introduced in Ref. [14]. For these graph families, the relevant Betti numbers and spectral gaps are analytically controlled, enabling a direct assessment of both optimization performance and measurement-resource requirements. Second, we study Erdős–Rényi random graphs, where we derive theoretical resource estimates and analyze landscape-variance diagnostics for barren-plateau behavior [31, 32].

We begin with island graphs. Let $K(m, k)$ denote the complete k -partite graph with m vertices in each partition. The graph consists of k islands of size m , with no edges within each island and all possible edges between distinct islands. These instances provide a controlled testbed for assessing whether the PMO objective resolves the low-energy subspace associated with the target Betti number. The corresponding

optimization curves and shot-scaling results are shown in Figure 1.

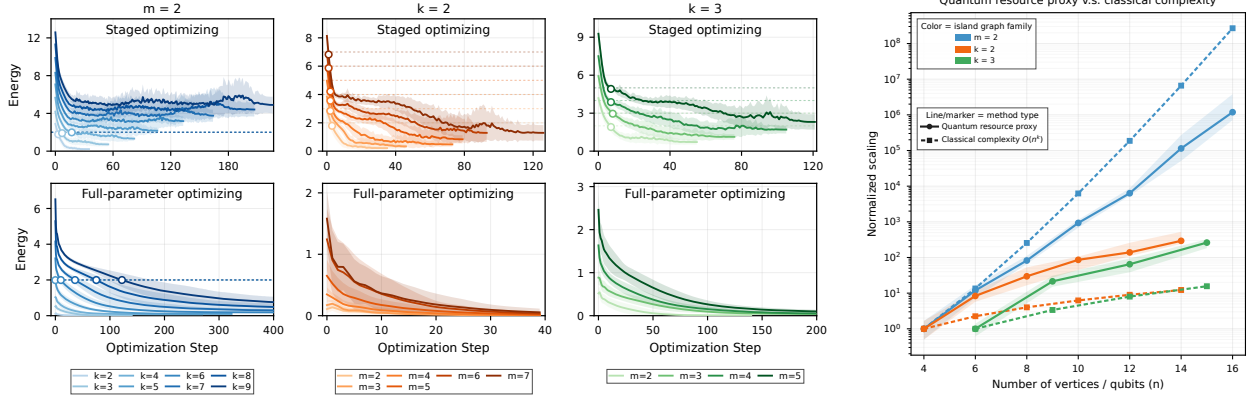


Figure 1: PMO performance on island-graph families. **Left:** Optimization trajectories under the PMO protocol for three parameter sweeps: fixed $m = 2$, fixed $k = 2$, and fixed $k = 3$. The circuit depth is set as $L = \lfloor n/2 \rfloor + 1$ for $n = mk$ qubits. The top row shows the staged optimization procedure, while the bottom row shows the subsequent full-parameter fine-tuning. Curves and shaded regions indicate the mean and standard deviation over 100 independent trials. Dashed horizontal lines mark the spectral threshold $\lambda_{\min}$, and open circles indicate the first step at which the mean energy falls below this threshold. **Right:** Normalized resource scaling for the same graph families. Solid curves report the quantum resource proxy, defined as the accumulated post hoc measurement-cost estimate along the noiseless reference optimization trajectory, evaluated up to the first crossing of the certification threshold $\lambda_{\min}$. Dashed curves show the complexity of the best-known classical algorithm for Betti number computation, which scales $\mathcal{O}(n^k)$. Both quantum and classical quantities are normalized within each family by their respective smallest-instance values.

We next consider the Erdős–Rényi random graph ensemble $G(n, p)$, where the graph has n vertices and each edge is included independently with probability p . For this ensemble, we provide a theoretical analysis of the resources required for estimating the gradients using parameter-shift rules. As summarized in Proposition 1, we first estimate the number of circuit runs needed to resolve a gradient component in the weight- $(k+1)$ sector, which is the sector associated with the k^{th} Betti number. We then analyze the variance of the PMO cost landscape in an expressive reference regime, as summarized in Proposition 2.

Proposition 1 (Expected gradient-shot complexity). *Let $G \sim G(n, p)$ be an Erdős–Rényi random graph with $0 < p < 1$, and let H_G be the corresponding Hamiltonian used to probe the nontriviality of $\beta_k(\text{Cl}(G))$ in the fermion-number- $(k+1)$ sector. Given a variational circuit evaluated on this sector, the expected number of circuit runs required to estimate a single gradient component using parameter-shift rules to additive precision ε scales as*

$$\mathcal{O}\left(\frac{(n + k^2)^2}{\varepsilon^2}\right). \quad (4)$$

Proposition 2 (Expected landscape variance in the homological regime). *Fix an integer $k \geq 1$ and let $G \sim G(n, p)$ with $p = n^\alpha$, where $-\frac{1}{k} < \alpha < -\frac{1}{k+1}$. In this regime, the k^{th} Betti number of the random clique complex $\text{Cl}(G)$ is asymptotically nonzero with high probability. Assume that the HWP ansatz forms a unitary 2-design on the weight- $(k+1)$ sector $\mathcal{H}_{k+1}$ and that its dynamical Lie algebra is subspace-controllable on this sector. Then, the expected landscape variance scales as*

$$\mathbb{E}_{G \sim G(n, p)} [\text{Var}_{\boldsymbol{\theta}} [\mathcal{C}_{G, k+1}(\boldsymbol{\theta})]] = \Theta\left(\frac{k^4}{\binom{n}{k+1}}\right). \quad (5)$$

Summary and Outlook

In this work, we develop a VQA framework for detecting nontrivial Betti numbers of the clique complex of a given graph. We apply this framework to island-graph instances, where the relevant Betti numbers and spectral gaps are analytically controlled, and numerically demonstrate its ability to certify nontrivial topology. We also introduce a progressive multi-stage optimization protocol and use it to assess the optimization performance and measurement-resource requirements of the algorithm. Furthermore, we analytically study the measurement-resource requirements for gradient-based optimization and discuss the implications for barren plateaus in nontrivial Erdős–Rényi random graph instances. Overall, our work offers a pathway for applying quantum methods to persistent homology and examines its feasibility and limitations.

References

- [1] Edelsbrunner, Letscher, and Zomorodian. Topological persistence and simplification. *Discrete & computational geometry*, 28(4):511–533, 2002.
- [2] Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. In *Proceedings of the twentieth annual symposium on Computational geometry*, pages 347–356, 2004.
- [3] Nina Otter, Mason A Porter, Ulrike Tillmann, Peter Grindrod, and Heather A Harrington. A roadmap for the computation of persistent homology. *EPJ Data Science*, 6:1–38, 2017.
- [4] Joel Friedman. Computing betti numbers via combinatorial laplacians. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of Computing*, pages 386–391, 1996.
- [5] Saugata Basu. On bounding the betti numbers and computing the euler characteristic of semi-algebraic sets. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pages 408–417, 1996.
- [6] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. In *Proceedings of the twenty-first annual symposium on Computational geometry*, pages 263–271, 2005.
- [7] Frédéric Chazal and André Lieutier. Stability and computation of topological invariants of solids in. *Discrete & Computational Geometry*, 37(4):601–617, 2007.
- [8] Seth Lloyd, Silvano Garnerone, and Paolo Zanardi. Quantum algorithms for topological and geometric analysis of data. *Nature communications*, 7(1):10138, 2016.
- [9] Sam Gunn and Niels Kornerup. Review of a quantum algorithm for betti numbers. *arXiv preprint arXiv:1906.07673*, 2019.
- [10] Ryu Hayakawa. Quantum algorithm for persistent betti numbers and topological data analysis. *Quantum*, 6:873, 2022.
- [11] Stefano Scali, Chukwudubem Umeano, and Oleksandr Kyriienko. Quantum topological data analysis via the estimation of the density of states. *Physical Review A*, 110(4):042616, 2024.
- [12] Nhat A Nghiem, Xianfeng David Gu, and Tzu-Chieh Wei. Quantum algorithm for estimating betti numbers using cohomology approach. *Quantum*, 9:1955, 2025.
- [13] Sam McArdle, András Gilyén, and Mario Berta. A streamlined quantum algorithm for topological data analysis with exponentially fewer qubits. *Quantum*, 10:2058, 2026.
- [14] Dominic W Berry, Yuan Su, Casper Gyurik, Robbie King, Joao Basso, Alexander Del Toro Barba, Abhishek Rajput, Nathan Wiebe, Vedran Dunjko, and Ryan Babbush. Analyzing prospects for quantum advantage in topological data analysis. *PRX Quantum*, 5(1):010319, 2024.
- [15] Casper Gyurik, Alexander Schmidhuber, Robbie King, Vedran Dunjko, and Ryu Hayakawa. Quantum computing and persistence in topological data analysis. *arXiv preprint arXiv:2410.21258*, 2024.
- [16] Marcos Crichigno and Tamara Kohler. Clique homology is qma 1-hard. *Nature Communications*, 15(1):9846, 2024.
- [17] Robbie King and Tamara Kohler. Gapped clique homology on weighted graphs is qma 1-hard and contained in qma. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 493–504. IEEE, 2024.
- [18] Edward Witten. Supersymmetry and morse theory. *Journal of differential geometry*, 17(4):661–692, 1982.
- [19] Chris Cade and P Marcos Crichigno. Complexity of supersymmetric systems and the cohomology problem. *Quantum*, 8:1325, 2024.

- [20] 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*, 5:4213, 2014. doi: 10.1038/ncomms5213.
- [21] Jarrod R. McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, 18(2):023023, 2016. doi: 10.1088/1367-2630/18/2/023023.
- [22] 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. *Nature Reviews Physics*, 3(9): 625–644, 2021. doi: 10.1038/s42254-021-00348-9.
- [23] Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Markus Brink, Jerry M. Chow, and Jay M. Gambetta. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature*, 549(7671):242–246, 2017. doi: 10.1038/nature23879.
- [24] Paul Fendley, Kareljan Schoutens, and Jan de Boer. Lattice models with $n=2$ supersymmetry. *Physical review letters*, 90(12):120402, 2003.
- [25] Liza Huijse and Kareljan Schoutens. Superfrustration of charge degrees of freedom. *The European Physical Journal B*, 64(3):543–550, 2008.
- [26] Jakob Jonsson. Certain homology cycles of the independence complex of grids. *Discrete & Computational Geometry*, 43(4):927–950, 2010.
- [27] Gian-Luca R Anselmetti, David Wierichs, Christian Gogolin, and Robert M Parrish. Local, expressive, quantum-number-preserving vqe ansätze for fermionic systems. *New Journal of Physics*, 23(11):113010, 2021.
- [28] Andrea Skolik, Jarrod R McClean, Masoud Mohseni, Patrick Van Der Smagt, and Martin Leib. Layerwise learning for quantum neural networks. *Quantum Machine Intelligence*, 3(1):5, 2021.
- [29] Xiaoyuan Liu, Anthony Angone, Ruslan Shaydulin, Ilya Safro, Yuri Alexeev, and Lukasz Cincio. Layer vqe: A variational approach for combinatorial optimization on noisy quantum computers. *IEEE Transactions on Quantum Engineering*, 3:1–20, 2022.
- [30] Ernesto Campos, Daniil Rabinovich, Vishwanathan Akshay, and J Biamonte. Training saturation in layerwise quantum approximate optimization. *Physical Review A*, 104(3):L030401, 2021.
- [31] 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.
- [32] Martin Larocca, Supanut Thanasilp, Samson Wang, Kunal Sharma, Jacob Biamonte, Patrick J Coles, Lukasz Cincio, Jarrod R McClean, Zoë Holmes, and Marco Cerezo. Barren plateaus in variational quantum computing. *Nature Reviews Physics*, 7(4):174–189, 2025.