TOPOLOGICAL DATA ANALYSIS ON NOISY QUANTUM COMPUTERS

Ismail Yunus Akhalwaya[†*] **Shashanka Ubaru**[†*] **Kenneth L. Clarkson**[†]

Mark S. Squillante[†] **Vishnu Jejjala**[‡] **Yang-Hui He**[§] **Kugendran Naidoo**[‡]

Vasileios Kalantzis[†] **Lior Horesh**[†]

ABSTRACT

Topological data analysis (TDA) is a powerful technique for extracting complex and valuable shape-related summaries of high-dimensional data. However, the computational demands of classical algorithms for computing TDA are exorbitant, and quickly become impractical for high-order characteristics. Quantum computers offer the potential of achieving significant speedup for certain computational problems. Indeed, TDA has been purported to be one such problem, yet, quantum computing algorithms proposed for the problem, such as the original Quantum TDA (QTDA) formulation by Lloyd, Garnerone and Zanardi, require currently unavailable fault-tolerance. In this study, we present NISQ-TDA, a *fully implemented end-to-end* quantum machine learning algorithm needing only a short circuit-depth, that is applicable to high-dimensional classical data, and with provable asymptotic speedup for certain classes of problems. The algorithm neither suffers from the data-loading problem nor does it need to store the input data on the quantum computer explicitly. The algorithm was successfully executed on quantum computing devices, as well as on noisy quantum simulators, applied to small datasets. Preliminary empirical results suggest that the algorithm is robust to noise.

1 INTRODUCTION

With the advent of modern technology, the collection of information-rich, high-dimensional data has become prevalent. These high-dimensional datasets are typically characterized by multidimensional correlation structures that are difficult to uncover. Extracting and analyzing such structural information is crucial in machine learning as well as in accelerating scientific discovery. Topological data analysis (TDA) is a powerful unsupervised machine learning technique for the extraction of valuable shape-related features of large datasets (Zomorodian & Carlsson, 2005; Ghrist, 2008; Wasserman, 2018). It represents one of the few data analysis algorithms that can process high-dimensional datasets and reduce them to a small set of local and global signature values that are interpretable and laden with predictive and analytical value. TDA has been shown to be useful in various scientific applications, including machine learning and artificial intelligence (AI) for the analysis of deep neural network architectures (e.g., estimate the capacity (Guss & Salakhutdinov, 2018) and topological complexity (Naitzat et al., 2020) of neural networks); neuroscience (Giusti et al., 2015), where topology is used to reveal intrinsic geometric structures in neural correlations; cosmology (Cole & Shiu, 2018b), where TDA is used for detecting non-Gaussianity of the cosmic microwave background (CMB); and genetics (Rabadán et al., 2020; Mandal et al., 2020b), for predicting phenotypes from gene co-expression or raw genomics data. Despite such progress in some applications, the true potential of TDA has been severely limited because classical algorithms for

[*]These authors contributed equally to this work. Corresponding authors
[†]IBM Research, USA and South Africa
[‡]University of the Witwatersrand, South Africa
[§]Royal Institution, UK

TDA have proven to be computationally prohibitive, only mitigated to some extent by sampling or by limiting calculations to low-dimensional properties.

Quantum computers represent one potential approach to address these prohibitive computational requirements of TDA. The power of quantum computers lies in their ability to perform computations in large computational (Hilbert) spaces, accessed via relatively small physical systems (Deutsch, 1985; Lloyd, 1996). With the recognition of this novel computational power in the 1980s (Feynman, 1982), there has been an arduous search for algorithms that achieve significant computational speedups over classical algorithms (Shor, 1994; Grover, 1996; Nielsen & Chuang, 2010). Such quantum algorithms offer the potential to solve problems that can not be solved using conventional computers. Quantum computers outperforming current classical supercomputers has been termed *quantum advantage* in the literature (Bravyi et al., 2018; Arute et al., 2019; Deshpande et al., 2022; Rinott et al., 2022). However, this has not yet been achieved for any problem of practical value.

In a seminal paper, Lloyd et al. (2016) proposed Quantum TDA (QTDA), an algorithm that achieves an expected exponential speedup in solving an approximation of TDA. Recent works (Gyurik et al., 2020; Cade & Crichigno, 2021; Crichigno & Kohler, 2022; Schmidhuber & Lloyd, 2022) have studied the hardness of the approximation problem solved by QTDA, and discussed the conditions under which the algorithm provably enjoys speedup over classical algorithms. Furthermore, this speedup is not overshadowed by the data-loading cost (Aaronson, 2015), which plagues several other quantum algorithms (Harrow et al., 2009; Gilyén et al., 2019), especially those related to machine learning (Biamonte et al., 2017; Schuld et al., 2015). However, the QTDA algorithm still requires long-lasting quantum coherence and low computational error to store and process the loaded data. Indeed, it requires *fault-tolerant* quantum computing (Shor, 1996; Aaronson, 2015; Preskill, 2018), an error-corrected quantum computer needing a very large overhead in resources (number of low-noise qubits and operations) (Arute et al., 2019; Zhao et al., 2020). Many components of the Lloyd et al. (2016) algorithm require fault-tolerance: Grover's search (Grover, 1996), Quantum Phase Estimation (Nielsen & Chuang, 2010), and repeated access to the input data. However, fault-tolerance has not yet been achieved on currently available quantum devices, and is likely several years away from full realization (goo, 2023). Intriguingly, the qubit numbers and noise levels that are currently realized in hardware are not classically simulatable, which raises the question of whether some algorithm could make use of these non-fault-tolerant noisy devices (Noisy Intermediate-Scale Quantum (NISQ) (Preskill, 2018)) for quantum advantage?

In this paper we present a quantum algorithm for solving the same problem as QTDA with an improved runtime, shorter circuit depth, and without fault tolerance requirements. Our NISQ-TDA algorithm solves the principal problem of TDA, estimating the Betti numbers of the given data (Ghrist, 2008). The algorithm only requires pairwise distances of the n data-points as input and outputs an estimate for the (normalized) Betti numbers of the data, which are signature values that describe the shape of the data. However, the calculation of these Betti numbers by current methods requires operating on large exponential-sized matrices (details of TDA and Betti numbers are provided in the next section). The approximation problem solved by our algorithm is believed to be intractable classically (likely belonging to a class of problems called DQC1-hard (Morimae et al., 2014)) under certain settings (Gyurik et al., 2020; Cade & Crichigno, 2021; Crichigno & Kohler, 2022), and in this sense potentially enjoys super-polynomial to exponential speedup over classical algorithms for certain classes of problems (Schmidhuber & Lloyd, 2022). We present a theoretical error analysis for the proposed algorithm, establishing error guarantees for the estimated Betti numbers, and show that the algorithm requires only $O(n/\sqrt{\delta})$-depth circuit complexity. We then present preliminary empirical results from implementations on real hardware and quantum simulations that illustrate the noise resiliency of our algorithm. Our presented theoretical and numerical results demonstrate that NISQ-TDA has the potential to be the *first* generically useful NISQ algorithm.

2 PRELIMINARIES

We begin by introducing the key concepts of quantum computing, TDA and quantum TDA (QTDA).

Quantum computing: Quantum computing is characterized by operations on the quantum state of n quantum bits or qubits, representing a vector in 2^n dimensional complex vector (Hilbert) space. The quantum operations or measurements correspond to multiplying the quantum state vector by certain $2^n \times 2^n$ matrices. Quantum circuits represent these operations in terms of a set of quantum

gates operating on the qubits. The number of these gates and the depth of the circuit define the circuit complexity of a given quantum algorithm. Quantum computers are difficult to build (preparing and maintaining the quantum states is extremely hard) and are very noisy. Therefore, the principles of quantum error correction were proposed to protect the quantum system from information loss and other damages (Gottesman, 2010). A (large-scale) quantum computer with many qubits is said to be fault-tolerant if the device is capable of such quantum error correction. However, realization of such fault-tolerant quantum systems is likely several years away. Currently available quantum computers are termed "Noisy Intermediate-Scale Quantum" (NISQ) (Preskill, 2018), and these devices are prone to considerable error rates and are limited in size by the number of logical qubits available in the system. In order to obtain results with reasonable accuracies on a NISQ device, the quantum circuit implementing a given algorithm needs to be of short depth.

Topological data analysis: TDA represents one of the few data analysis methodologies that can process high-dimensional datasets and reduce them to a small set of local and global signature values that are interpretable and laden with predictive and analytical value. Given a set of n data-points $\{x_i\}_{i=0}^{n-1}$ in some space together with a distance metric $\mathcal{D}$, a Vietoris-Rips (Ghrist, 2008) simplicial complex is constructed by selecting a resolution/grouping scale ε that defines the "closeness" of the points with respect to the distance metric $\mathcal{D}$, and then connecting the points that are a distance of ε from each other (i.e., connecting points x_i and x_j whenever $\mathcal{D}(x_i, x_j) \leq \varepsilon$, forming a so-called 1-skeleton). A k-simplex is then added for every subset of $k+1$ data-points that are pair-wise connected (i.e., for every k-clique, the associated k-simplex is added).

Let S_k denote the set of k-simplices in the Vietoris–Rips complex $\Gamma = \{S_k\}_{k=0}^{n-1}$, with $s_k \in S_k$ written as $\{j_0, \ldots, j_k\}$ where j_i is the ith vertex of s_k. Let $\mathcal{H}_k$ denote an $\binom{n}{k+1}$-dimensional Hilbert space, with basis vectors corresponding to each of the possible k-simplices (all subsets of size $k+1$). Further let $\tilde{\mathcal{H}}_k$ denote the subspace of $\mathcal{H}_k$ spanned by the basis vectors corresponding to the simplices in S_k, and let $|s_k\rangle$ denote the basis state corresponding to $s_k \in S_k$. Then, the n-qubit Hilbert space $\mathbb{C}^{2^n}$ is given by $\mathbb{C}^{2^n} \cong \bigoplus_{k=0}^{n} \mathcal{H}_k$. The boundary map (operator) on k-dimensional simplices $\partial_k : \mathcal{H}_k \to \mathcal{H}_{k-1}$ is a linear operator defined by its action on the basis states as follows:

$$\partial_k |s_k\rangle = \sum_{l=0}^{k-1} (-1)^l |s_{k-1}(l)\rangle , \tag{1}$$

where $|s_{k-1}(l)\rangle$ is the *lower* simplex obtained by leaving out vertex l (i.e., s_{k-1} has the same vertex set as s_k except without j_l), and s_{k-1} is $k-1$-dimensional, a dimension less than s_k. The factor $(-1)^l$ produces the *oriented* (Ghrist, 2008) sum of boundary simplices, which keeps track of neighbouring simplices so that $\partial_{k-1}\partial_k |s_k\rangle = 0$, given that the boundary of the boundary is empty.

The boundary map $\tilde{\partial}_k : \tilde{\mathcal{H}}_k \to \tilde{\mathcal{H}}_{k-1}$ restricted to a given Vietoris–Rips complex Γ is given by $\tilde{\partial}_k = \partial_k \tilde{P}_k$, where $\tilde{P}_k$ is the projector onto the space S_k of k simplices in Γ. The full boundary operator on the fully connected complex (the set of all subsets of n points) is the direct sum of the k-dimensional boundary operators, namely $\partial = \bigoplus_k \partial_k$. The *$k$-homology group* is the quotient space $\mathbb{H}_k := \ker(\tilde{\partial}_k)/\mathrm{img}(\tilde{\partial}_{k+1})$, representing all k-holes which are not "filled-in" by $k+1$ simplices and counted once when connected by k simplices (e.g., the two holes at the ends of a tunnel count once). Such global structures moulded by local relationships is what is meant by the "shape" of data. The *kth Betti Number* β_k is the dimension of this k-homology group, namely $\beta_k := \dim \mathbb{H}_k$.

These Betti numbers therefore count the number of holes at scale ε, as described above. By computing the Betti numbers at different scales ε, we can obtain the *persistence barcodes/diagrams* (Ghrist, 2008), i.e., a set of powerful interpretable topological features that account for different scales while being robust to small perturbations and invariant to various data manipulations. These stable persistence diagrams not only provide information at multiple resolutions, but they also help identify, in an unsupervised fashion, the resolutions at which interesting structures exist. The *Combinatorial Laplacian*, or Hodge Laplacian, of a given complex is defined as $\Delta_k := \tilde{\partial}_k^\dagger \tilde{\partial}_k + \tilde{\partial}_{k+1} \tilde{\partial}_{k+1}^\dagger$. From the Hodge theorem (Friedman, 1998; Lim, 2019), we can compute the kth Betti number as

$$\beta_k := \dim \ker(\Delta_k). \tag{2}$$

Therefore, computing Betti numbers for TDA can be viewed as a rank estimation problem (i.e., $\beta_k = \dim \tilde{\mathcal{H}}_k - \mathrm{rank}(\Delta_k)$). Additional TDA details can be found in Appendix A.2. The problem

of normalized Betti number estimation (BNE) is defined as (Gyurik et al., 2020): Given a set of n points, its corresponding Vietoris–Rips complex Γ, an integer $0 \leq k \leq n-1$, and the parameters $(\epsilon, \eta) \in (0, 1)$, find the value $\chi_k \in [0, 1]$ that satisfies with probability $1 - \eta$ the condition

$$\left| \chi_k - \frac{\beta_k}{|S_k|} \right| \leq \epsilon, \tag{3}$$

where $|S_k|$ is the the number of k-simplices $S_k \in \Gamma$ or $\dim \tilde{\mathcal{H}}_k$, the dimension of the Hilbert space spanned by the set of k-simplices in the complex.

Quantum TDA: Lloyd et al. (2016) proposed Quantum TDA (QTDA), an algorithm for solving an approximation of TDA in polynomial time for a class of simplicial complexes. Recent works have shown, e.g., (Gyurik et al., 2020; Schmidhuber & Lloyd, 2022), that the problem QTDA solves approximately is intractable classically for certain classes of complexes. The TDA problem of computing Betti numbers exactly has been shown to be intractable for even quantum computers as decision clique homology has been proven to be QMA1-hard (Crichigno & Kohler, 2022) for clique complexes; and promise weighted clique homology has been shown to be QMA1-hard and contained in QMA (King & Kohler, 2023). The approximative version that QTDA actually solves involves a different computational class: DQC1-hard. This normalized Betti number estimation problem has been shown to be DQC1-hard for general chain complexes (Cade & Crichigno, 2021) and is conjectured to hold for clique complexes (Cade & Crichigno, 2021; King & Kohler, 2023).

QTDA involves two main steps, namely: (a) repeatedly constructing the simplices in the given simplicial complex as a mixed quantum state using Grover's search algorithm (Boyer et al., 1998); and (b) projecting this onto the eigenspace of Δ_k in order to calculate the *Betti numbers* of the complex, using quantum phase estimation (QPE) (Nielsen & Chuang, 2010) (details are provided in the Appendix A). The computational complexity is $O(n^5/(\delta_k\sqrt{\zeta_k}))$ where n is the number of data points, δ_k denotes the smallest nonzero eigenvalue of Δ_k, and ζ_k is the fraction of all simplices of order k in the given complex, resulting in significant speedup over known classical algorithms. However, QTDA requires long-lasting quantum coherence to store the loaded data for the length of the long-depth circuits thus requiring fault-tolerant quantum computing. In particular, Grovers and QPE require precise phase information where any errors would accumulate multiplicatively.

3 NISQ-TDA

We now present our proposed quantum algorithm, NISQ-TDA, for estimating the (normalized) Betti numbers of datasets (simplicial complexes) defined through vertices and edges. The algorithm involves three key components, namely: (a) an efficient representation of the full boundary operator as a sum of Pauli operators; (b) a quantum rejection sampling technique to project onto the data-defined simplicial complex; and (c) a stochastic rank estimation method to estimate the output signature *Betti* numbers. In order to calculate the Betti numbers, the first of two major tasks is to construct a quantum circuit that applies the data-defined Laplacian to *any* input set of simplices. In our algorithm, this involves three main sub-components.

The first is a quantum representation of the complete (not data-defined) boundary map operator (say B), called the **Fermionic boundary operator** (Cade & Crichigno, 2021; Akhalwaya et al., 2022). It acts on all possible simplices with n points and returns their corresponding boundary simplices. The representation involves only unitary operators written as a sum of Pauli (fermionic) operators. The Hermitian boundary operator B is written as $B = \sum_{i=0}^{n-1} a_i + a_i^\dagger$, where the a_i are the Jordan-Wigner (Jordan & Wigner, 1928) Pauli embeddings corresponding to the n-spin fermionic annihilation operators. The implementation of this fermionic boundary operator B on a quantum computer requires only n qubits, $O(n^2)$ gates, and an $O(n)$-depth circuit; see Appendix B for details.

The second sub-component, which we call **Projection onto simplices**, consists of constructing the simplicial complex (Γ) corresponding to the given data by implementing the projector (P_Γ) onto Γ as qubit gates and measurements. A series of multi-qubit control-NOT gates, one for each edge in the data, checks if the edges of the input simplices (in superposition) are actually present in the data. The result is stored in a *flag* register which is then measured. Since there are $\binom{n}{2} \sim O(n^2)$ potential edges, this seems to require $O(n^2)$ depth. Fortunately, the checks can be run in parallel and in batches using a round-robin procedure, reusing the same flag register through the power of

mid-circuit measurement. By repeating the entire circuit until the register measurements read the 'all-in' flag, the input simplices are projected onto Γ.

Given the classical encoding of the ε-close pairs, we systematically entangle the simplices with an $n/2$-qubit flag register. The $n/2$ qubits are used to process $n/2$ pairs of vertices at a time in $n-1$ rounds, thereby covering all $\binom{n}{2}$ potential ε-close pairs of vertices. The projection begins by creating a uniform superposition over all simplices (or over all k-simplices). We check $n/2$ pairs at a time, for all simplices in the superposition (hence the quantum speedup). The $n/2$ pairs are chosen such that the C-C-NOT (Toffoli) gates, controlling on pairs of vertex qubits targeting the flag register, are executed in parallel. In each round, we measure the flag register and proceed only if we receive all zeros. This collapses the simplex superposition into those simplices that have pairs which are not missing from the adjacency graph. The procedure succeeds when repeated $\frac{1}{\zeta_k}$ times, where ζ_k is the fraction of all possible simplices of order k that are in Γ. The 'all-orders' data-defined Laplacian can thus be expressed as $\Delta = P_\Gamma B P_\Gamma B P_\Gamma$.

Although this simple linear-depth circuit implementation of P_Γ suggests a requirement of quantum computers with all-to-all connectivity (as used in our experiments), we can indeed implement it on quantum computers with only linear qubit connectivity using a sorting network approach in $O(n)$ depth (Beals et al., 2013; O'Gorman et al., 2019). The network uses nearest-neighbor SWAP gates and with n layers of such 'qubit swaps', all $\binom{n}{2}$ pair of qubits become nearest-neighbors at some layer; see O'Gorman et al. (2019) for details. Moreover, if we use $\binom{n}{2}$ ancilla qubits, one per edge, we can measure these only once at the end of the $n-1$ rounds (instead of $n/2$ flag registers and $n-1$ measure and resets), and if we measure all zeros, then the projection is successful.

Most importantly, the ability to write the Laplacian in terms of a circuit that does not require accessing stored quantum data is one of the key enabling innovations of NISQ-TDA. The input edge data is not stored on the quantum computer but enters through the presence or absence of the multi-qubit control gates of the projector. Every time the complex projection is called, the data is freshly and accurately injected into the quantum computer. This suggests that NISQ-TDA is partially self-correcting, and under noise presence, the last application of P_Γ mitigates the noise. When noise-levels only allow for one coherent application of P_Γ, this application meaningfully represents the data and can be used for alternate machine learning tasks.

The third sub-component, which we call **Projection to a simplicial order**, is the construction of the projector (P_k) onto the k-simplex subspace. The circuit is a sequence of control-'add one' sub-circuits that conditions on each vertex qubit of the simplex register and increments a $\log(n)$-sized count register. The operation is equivalent to implementing conditional-permutation, and can be efficiently implemented using diagonalization (Shende et al., 2006) in the Fourier basis. Finally, the projection is completed and fully realized as a non-unitary operation by measuring the count register. The cost in depth is only $O(\log^2 n)$. The data-defined Laplacian corresponding to simplicial order k can thus be written as $\Delta_k = P_k \Delta P_k$.

The second major part of the NISQ-TDA algorithm, which we call the **Stochastic Chebyshev method**, consists of using the above quantum circuit in a larger classically controlled framework, making NISQ-TDA a hybrid quantum-classical algorithm. The classical framework is a stochastic *rank* estimation using the *Chebyshev* polynomials (Ubaru & Saad, 2016; Ubaru et al., 2017). Once we obtain the rank of the Laplacian, we have the Betti numbers $\beta_k = \dim(\ker(\Delta_k)) = |S_k| - \mathrm{rank}(\Delta_k)$, where $S_k \subseteq \Gamma$ is the set of k-simplices in the given complex Γ. Stochastic rank estimation recasts the eigen-decomposition problem into the estimation of the matrix function trace.

Assuming the smallest nonzero eigenvalue of $\tilde{\Delta}_k = \Delta_k / n$ is greater than or equal to δ, we have

$$\mathrm{rank}(\Delta_k) \overset{def}{=} \mathrm{trace}(h(\tilde{\Delta}_k)), \text{ where } h(x) = \begin{cases} 1 & \text{if } x > \delta \\ 0 & \text{otherwise} \end{cases}.$$

Supposing $\tilde{\Delta}_k = \sum_i \lambda_i |u_i\rangle\langle u_i|$ is the eigen-decomposition, we have $h(\tilde{\Delta}_k) = \sum_i h(\lambda_i)|u_i\rangle\langle u_i|$, where the step function $h(\cdot)$ takes a value of 1 above the threshold $\delta > 0$ and the eigenvalues of $\tilde{\Delta}_k$ are in the interval $\{0\} \cup [\delta, 1]$. Next, $h(\tilde{\Delta}_k)$ is approximated using a truncated Chebyshev polynomial series (Trefethen, 2019) as $h(\tilde{\Delta}_k) \approx \sum_{j=0}^{m} c_j T_j(\tilde{\Delta}_k)$, where $T_j(\cdot)$ is the jth-degree Chebyshev polynomial of the first kind and c_j are the coefficients with closed-form expressions. The trace is approximated using the stochastic trace estimation method (Hutchinson, 1990) given by $\mathrm{trace}(A) \approx$

$\frac{1}{\mathrm{n_v}} \sum_{l=1}^{\mathrm{n_v}} \langle v_l|A|v_l\rangle$, where $|v_l\rangle, l = 1, \ldots, \mathrm{n_v}$, are random vectors with zero mean and uncorrelated coordinates. It can be shown that a set of random columns of the Hadamard matrices works well as a choice for $|v_l\rangle$, both in theory and practice (see the supplementary material). Sampling a random Hadamard state vector in a quantum computer can be conducted with a short-depth circuit. Given an initial state $|0\rangle$, we randomly flip the n qubits (by applying a NOT gate as determined by a random n-bit binary number generated classically). Thereafter, we apply the n-qubit Hadamard gate to produce a state corresponding to a random column of the $2^n \times 2^n$ Hadamard matrix. Therefore, the rank of Δ_k can be approximately estimated as $\mathrm{rank}(\Delta_k) \approx \frac{1}{\mathrm{n_v}} \sum_{l=1}^{\mathrm{n_v}} \left[\sum_{j=0}^{m} c_j \langle v_l|T_j(\tilde{\Delta}_k)|v_l\rangle\right]$, where the c_j are Chebyshev coefficients for approximating the step function. Given a circuit that block-encodes $\tilde{\Delta}_k$, we can block-encode a j-degree Chebyshev polynomial $T_j(\tilde{\Delta}_k)$ using the idea of qubitization (Low & Chuang, 2019; Gilyén et al., 2019). Details are given in Appendix B.

NISQ-TDA Algorithm: We now have all the ingredients to present our NISQ-TDA algorithm:

Algorithm 1 NISQ-TDA Algorithm

Input: Pairwise distances of n data points and encoding of the ε-close pairs; parameters ϵ, δ, and $\mathrm{n_v} = O(\epsilon^{-2})$; and $\mathrm{n_v}$ n-bit random binary numbers.
Output: Betti number estimates $\chi_k,\ k = 0, \ldots, n-1$.
for $l = 1, \ldots, \mathrm{n_v} = O(\epsilon^{-2})$ **do**
  for $j = 0, \ldots, m = O(\log(1/\epsilon))$ **do**
    1. Prepare a random Hadamard state vector $|v_l\rangle$ from $|0\rangle$ using the l-th random number.
    2. Use the circuits for P_k, P_Γ, and $\tilde{B} = B/\sqrt{n}$ to compute
      $|\phi_l\rangle = |0^q\rangle\, \tilde{\Delta}_k\, |v_l\rangle + \left|\tilde{\perp}\right\rangle$, where $q = \#$ancilla qubits needed for projections.
    3. Use qubitization to form: $\left|\psi_l^{(j)}\right\rangle = \left|0^{q+1}\right\rangle T_j(\tilde{\Delta}_k)\, |v_l\rangle + |\perp\rangle$ from $|\phi_l\rangle$.
    4. Compute the Chebyshev moments $\theta_l^{(j)} = \langle v_l|\, T_j(\tilde{\Delta}_k)\, |v_l\rangle$ from $\left|\psi_l^{(j)}\right\rangle$.
  end for
  For $j = 0$, estimate $|S_k|$ using the average norm of the $P_\Gamma P_k\, |v_l\rangle$.
end for
Estimate $\chi_k = 1 - \frac{1}{\mathrm{n_v}} \sum_{l=1}^{\mathrm{n_v}} \left[\sum_{j=0}^{m} c_j \theta_l^{(j)}\right]$.
Repeat for $k = 0, \ldots, n-1$.

Analyses: Our *NISQ-TDA* algorithm returns the estimates χ_k for the normalized Betti numbers $\beta_k/|S_k|$, for each order $k = 0, \ldots, n-1$, where $|S_k|$ is the number of k-simplices in the given Γ. We discuss potential scientific machine learning and AI applications of NISQ-TDA in the Appendix. The remainder of this section focuses on theoretical analyses of our *NISQ-TDA* algorithm, with the formal details and proofs provided in Appendix C. We begin with the following main result.

Theorem 1. *Assume we are given the pairwise distances of any n data points and the encoding of the corresponding ε-close pairs, together with an integer $0 \le k \le n-1$ and the parameters $(\epsilon, \delta, \eta) \in (0, 1)$. Further assume the eigenvalues of the scaled Laplacian $\tilde{\Delta}_k$ are in the interval $\{0\} \cup [\delta, 1]$, and choose $\mathrm{n_v}$ and m such that*

$$\mathrm{n_v} = O\left(\frac{\log(1/\eta)}{\epsilon^2}\right) \qquad \textit{and} \qquad m > \frac{\log(1/\epsilon)}{\sqrt{\delta}}.$$

Then, the Betti number estimation $\chi_k \in [0, 1]$ by NISQ-TDA, with probability at least $1 - \eta$, satisfies

$$\left|\chi_k - \frac{\beta_k}{|S_k|}\right| \le \epsilon.$$

Our analysis accounts for errors due to (a) polynomial approximation of the step function; (b) stochastic trace estimator; and (c) also shot noise, i.e., errors in Chebyshev moments estimation and their propagation in classical computation; for details, see Appendix C.

We next discuss the circuit and computational complexities of our proposed algorithm and show that it is NISQ implementable under certain conditions, such as the requirement for simplices-dense com-

plexes, which commonly occur for large resolution scale. The main quantum component of the algorithm comprises the computation of $\theta_l^{(j)} = \langle v_l | T_j(\tilde{\Delta}_k) | v_l \rangle$, for $j = 0, \ldots, m \sim O(\log(1/\epsilon)/\sqrt{\delta})$, with $\mathrm{n_v} \sim O(\epsilon^{-2})$ random Hadamard vectors. The random Hadamard state preparation requires n single-qubit Hadamard gates in parallel and $O(1)$ time. For a given k, constructing $\tilde{\Delta}_k$ involves implementing the boundary operator $\tilde{B}$ and the projectors P_Γ and P_k. The operator B, involving the sum of n Pauli operators, can be implemented using a circuit with $O(n)$ gates. Constructing P_k requires $O(n \log^2 n)$ gates, and this succeeds for a random order k. Then, for P_Γ, we need to find all the simplices that are in the complex Γ. This is achieved using $n/2$ qubits in parallel and $n-1$ operations, and thus the time complexity remains $O(n)$. The number of gates required will be $O(n^2 \bar{\zeta}_k)$, where $\bar{\zeta}_k := \min\{1-\zeta_k, \zeta_k\}$. When we use the measure and reset approach (e.g. for the first projection onto the complex), the procedure of applying the projector succeeds when repeated $1/\zeta_k$ times. The projectors together require $O(n^2)$ gates, while the depth remains $O(n)$, and the time complexity for a projection will be $O\left(\frac{n}{\zeta_k}\right)$ (for the first projection) and $O(n)$ when we use $O(n^2)$ ancilla qubits for qubitization. Therefore, the total time complexity of our algorithm is

$$O\left(\frac{n \log(1/\epsilon)}{\sqrt{\delta}\epsilon^2 \zeta_k}\right).$$

Supposing δ_k is the spectral gap of Δ_k and $\tilde{\Delta}_k = \frac{\Delta_k}{n}$, then $\delta = \frac{\delta_k}{n}$. The best-known classical algorithm for Betti number estimation of order k has a time complexity of $O(\mathrm{poly}(n^k))$ (Gyurik et al., 2020) or $O(n^{1/\delta \log(1/\epsilon)})$ (Apers et al., 2022). Thus, the QTDA algorithms can achieve super-polynomial to exponential speedups over the best-known classical algorithms whenever we have:

- **Simplices/Clique dense complexes** – the given complex Γ is simplices/clique dense, i.e., ζ_k is large or $|S_k| \sim O(\mathrm{poly}(n))$;
- $O(1/\mathbf{poly}(n))$ **spectral gap** – the spectral gap between zero and nonzero eigenvalues of Δ_k is not exponentially small, i.e., δ of $\tilde{\Delta}_k$ is $O(1/\mathrm{poly}(n))$ (Apers et al., 2022); and
- **Large Betti number** – the Betti number β_k (and the ratio $\beta_k/|S_k|$) needs to be large so that a large ϵ suffices to estimate it to a reasonable precision.

A few examples for simplicial complexes that satisfy these conditions are discussed in the Appendix. Further examples and discussions on the potential speedups for quantum TDA algorithms are presented in (Schmidhuber & Lloyd, 2022).

We wish to remark that known examples of simplicial complexes with exponentially many holes (Betti number) are limited. An example family of graphs with exponentially many high-dimensional holes are presented in (Fendley & Schoutens, 2005). More importantly, our algorithm still likely achieves exponential advantage over known classical approaches in efficiently answering the question: *does the given simplicial complex have exponentially many holes or not?* In that regard, our algorithm is indeed applicable to *non-handcrafted high-dimensional classical data*.

From a different point of view, the Chebyshev moments capture the spectral information of Δ_k and have even more information than the Betti numbers. This therefore opens the door for these (P_Γ-corrected) noisy moments to be used directly as input features in downstream contexts such as machine learning classification, further relieving the depth and noise requirements of NISQ-TDA.

4 Experimental Results

With the theory promising short depths, it remains to demonstrate that NISQ-TDA is sufficiently noise-robust for quantum advantage to be achieved for the actual depths in realizable hardware and under realistic noise levels. Currently optimized classical TDA algorithms cannot compute all Betti numbers for 64 generic vertices (we have empirically verified with a popular public package called *GUDHI* (Maria et al., 2014)). Hence, quantum advantage could possibly be achieved when running NISQ-TDA on 64 vertices. Such large NISQ-TDA circuits are also beyond what is simulatable classically (Pednault et al., 2017; 2019).

We first present the actual depths needed in the form of a depth versus number of vertices plot, which also empirically confirms that circuit depth grows linearly with the number of vertices. Figure 1

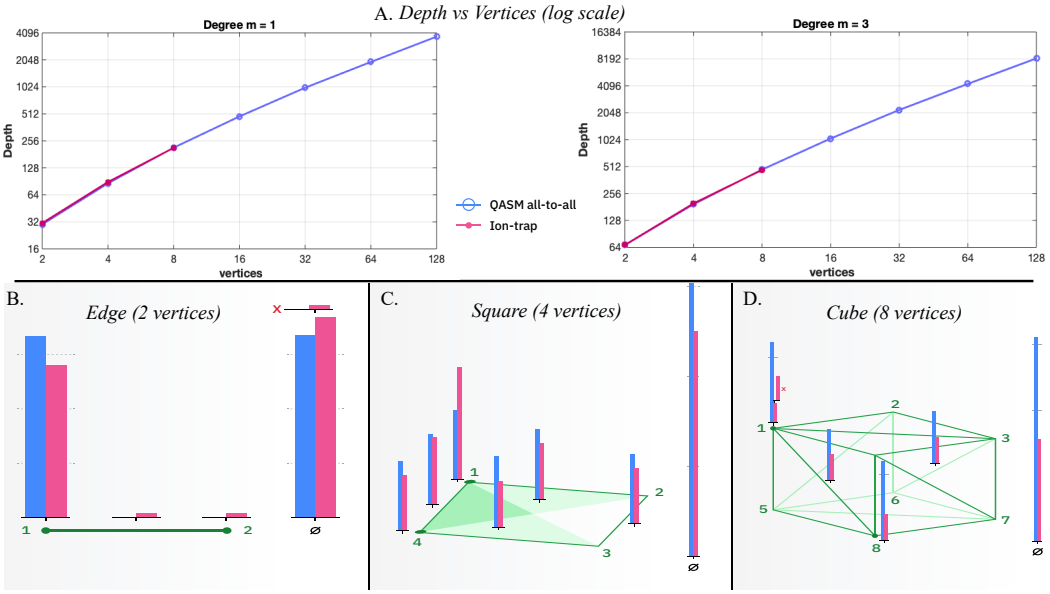

Figure 1: Results from real hardware of Laplacian applications (using measure and reset projections): A. Circuit depth versus the number of vertices for degree $m = 1$ and 3; (B., C. and D.) Histograms of the probability measurements as obtained from the hardware (right, magenta bars) and from a simulator (left, blue bars) for three different datasets namely, an edge (2 vertices), a square (4 vertices), and a cube (8 vertices). ϕ defines the null state, and 'X' denotes the probability mass with incorrect flag readings.

shows depths for both actual quantum hardware circuits and generic all-to-all quantum simulator circuits. For the quantum hardware, we employed the public-cloud accessible 'H1' 12-qubit trapped-ion quantum computer from Quantinuum (powered by Honeywell) (Honeywell, 2022). We selected the most conservative number of edges to cover the worst-case depth scenario. The magenta solid points of sub-figure A correspond to Laplacian circuit depths obtained from Quantinuum's own native compiler, and the blue circled points correspond to those obtained from a quantum simulator. We observe that the circuit depth for the Laplacian scales *linearly* with respect to the number of data points. As discussed earlier, our algorithm can be implemented on quantum computers with *linear qubit connectivity* as well, using the sorting network approach (Beals et al., 2013).

The remaining three sub-figures (B, C, D) present the histograms of the top probability measurements for different numbers of vertices (2, 4, 8) for both hardware runs (right, magenta bars) and simulation runs (left, blue bars). These measurements are the raw outputs of the quantum circuit before being converted into expectations (where flag values play a role). The respective complexes chosen correspond to easily understandable shapes (edge, square, cube) represented by the (green) edges. The input simplex set corresponds to a uniform superposition over all simplices (including not shown triangles, tetrahedrons, and all higher-order polytopes). Due to projection onto the specified complexes and interference (correctly eliminating boundaries, sending mass to the null state ϕ), not all simplices will appear/remain after the application of the Laplacian, demonstrating that the hardware is truly performing a coherent quantum calculation. These sub-figures clearly show that there is agreement between hardware and noise-free simulations on which simplices receive the top probability measurements. Three types of errors are, however, visible by the hardware: reduced probability mass for correct simplices, some small probability mass on incorrect simplices (also not shown are the non-top measurements), and correct simplex mass but incorrect flag readings (marked with a red 'X' in sub-figures B and D). Even with these errors, Figure 1 unequivocally demonstrates that at real-world noise-levels there is sufficient coherence to reproduce the correct interference at the depths of these circuits.

The next task would be to demonstrate that these errors, which inevitably enter, do not dramatically disturb the downstream Betti number calculation. For this we chose complexes with large eigenvalue gaps, and sufficiently many random vectors and shots. The Chebyshev parameters we selected

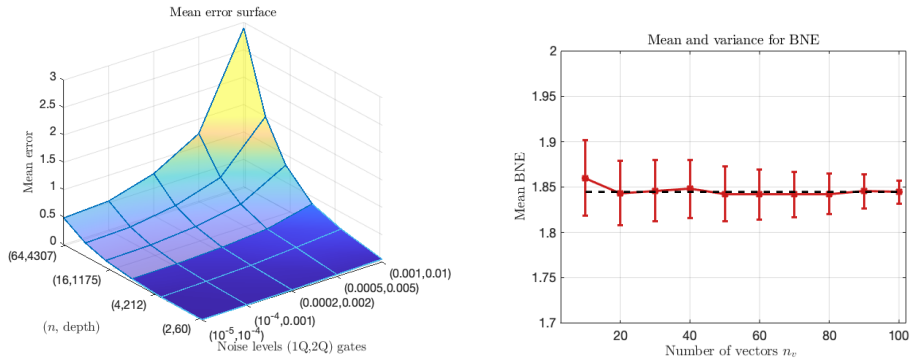

Figure 2: Results from noisy simulations: A. Mean error surface as a function of the noise levels in (1-qubit, 2-qubits) gates and (number of vertices n, circuit depth). B. Mean and the variance of the Betti number estimated as a function of the number of random vectors $\mathrm{n_v}$ with $n = 8$ vertices, degree $m = 5$ and the noise-level: $(0.001, 0.01)$.

are such that, in the noise-free scenario, the algorithm would calculate the Betti number almost perfectly (i.e., with a mean error of close to zero). Thus, any Betti-number errors involving the noisy simulations are due mainly to the quantum noise and only minimally on downstream classical approximations. In this setup, the error that can naturally be considered tolerable is 0.5, since any error less than 0.5 rounds to exactly the correct Betti number. In Figure 2, we present results from extensive noisy quantum simulations of the non-qubitized version of the algorithm. The right plot shows the mean and the variance (as error bars) of the Betti number estimated as a function of the number of random vectors $\mathrm{n_v}$. We note that the mean converges to ~ 1.84 (the true Betti number is $\beta_0 = 2$) and, most importantly, the variance reduces as we increase $\mathrm{n_v}$. This variance-reduction mitigates errors due to shot noise and randomness in the trace estimation, illustrating the precision-versus-number-of-trials benefit of NISQ-TDA. In the left figure, we present the mean error surface plot for Betti number estimation, as a function of noise levels (chosen triples of measurement, one and two-qubit gate errors) and number of vertices (with concomitant circuit depth). The first number of the listed noise-level pair corresponds to the one-qubit error probability. The measurement and the two-qubit error probabilities are both set to the second value. In the surface plot, the solid region (for $n = 2$ to $n = 8$) corresponds to actual noisy simulations and the translucent region (from 16 to 64 vertices) corresponds to an extrapolation of the surface for larger n, which we cannot simulate classically (even $n = 16$ was not simulatable using a large classical machine with 2 GPUs). The surface plot extrapolations provide the minimum noise-level requirements for NISQ-TDA to successfully run on future larger NISQ devices. See Appendix D for additional results, including preliminary results on cosmic microwave background (CMB) data.

5 Conclusions

The true potential of TDA for machine learning has been severely limited because of the computationally prohibitive requirements of classical algorithms. To address this critical issue and revive the potential of TDA as a viable machine learning approach, we presented a new quantum algorithm for Betti number estimation with comprehensive error and complexity analyses. This is one of the *first* quantum machine learning algorithms with short depth and potential significant speedup under certain assumptions. Our algorithm neither suffers from the data-loading problem nor does it likely require fault-tolerant coherence for even mid-size datasets. The algorithm fits the hybrid quantum-classical scheme but within a recently developed randomized-approximation framework. The implementation and successful execution of the entire algorithm on real quantum hardware and noisy simulations was demonstrated, illustrating noise-resiliency at realistic noise-levels. These advantages imply that this algorithm may be one of the few noise-robust quantum algorithms capable of performing an important and useful AI task on near-term (non-fault tolerant) quantum devices, beyond the reach of classical computation. Possible future research directions include: improvements to the algorithm in order to efficiently deploy it on sparsely connected quantum devices; achieving substantial asymptotic speedups under more general settings; and identifying interesting domain problems for which NISQ-TDA can be employed for practical purposes.

ACKNOWLEDGEMENTS

This research was supported in part by the Air Force Research Laboratory (AFRL) grant number FA8750-C-18-0098, and in part by IBM Research, South Africa under the Equity Equivalent Investment Programme (EEIP) of the government of South Africa. VJ is supported in part by the South African Research Chairs Initiative of the National Research Foundation, grant number 78554. Firstly, we wish to thank Brian Neyenhuis, Jennifer Strabley, Chad Edwards, and Tony Uttley from the Quantinuum quantum team for generously providing us credits to access their quantum computers and assisting with the experimentation. Secondly, we would like to thank Adam Connolly and Julien Sorci for pointing out a flaw in the algorithm of an earlier version of the paper. Next, we would like to acknowledge Tal Kachman for the suggestion to use controlled-increment to entangle the simplices with the count register. The authors would also like to thank Scott Aaronson, Paul Alsing, Ryan Babbush, Sergey Bravyi, Chris Cade, Marcos Crichigno, Aram Harrow, Gil Kalai, and Seth Lloyd for valuable discussions.

REFERENCES

Suppressing quantum errors by scaling a surface code logical qubit. *Nature*, 614(7949):676–681, 2023.

Scott Aaronson. Read the fine print. *Nature Physics*, 11(4):291–293, 2015.

Kartick Adhikari, Robert J Adler, Omer Bobrowski, and Ron Rosenthal. On the spectrum of dense random geometric graphs. *arXiv preprint arXiv:2004.04967*, 2020.

Robert J Adler, Omer Bobrowski, and Shmuel Weinberger. Crackle: The homology of noise. *Discrete & Computational Geometry*, 52(4):680–704, 2014.

N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters. *Astron. Astrophys.*, 641:A6, 2020. doi: 10.1051/0004-6361/201833910. [Erratum: Astron.Astrophys. 652, C4 (2021)].

Ismail Yunus Akhalwaya, Yang-Hui He, Lior Horesh, Vishnu Jejjala, William Kirby, Kugendran Naidoo, and Shashanka Ubaru. Representation of the fermionic boundary operator. *Physical Review A*, 106(2):022407, 2022.

Andris Ambainis and Joseph Emerson. Quantum t-designs: t-wise independence in the quantum world. In *Twenty-Second Annual IEEE Conference on Computational Complexity (CCC'07)*, pp. 129–140. IEEE, 2007.

Simon Apers, Sander Gribling, Sayantan Sen, and Dániel Szabó. A (simple) classical algorithm for estimating betti numbers. *arXiv preprint arXiv:2211.09618*, 2022.

Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando GSL Brandao, David A Buell, et al. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019.

Haim Avron and Sivan Toledo. Randomized algorithms for estimating the trace of an implicit symmetric positive semi-definite matrix. *Journal of the ACM (JACM)*, 58(2):1–34, 2011.

Frank Bauer, Jürgen Jost, and Shiping Liu. Ollivier-Ricci curvature and the spectrum of the normalized graph Laplace operator. *arXiv preprint arXiv:1105.3803*, 2011.

Robert Beals, Stephen Brierley, Oliver Gray, Aram W Harrow, Samuel Kutin, Noah Linden, Dan Shepherd, and Mark Stather. Efficient distributed quantum computing. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 469(2153):20120686, 2013.

Orr Beit-Aharon and Roy Meshulam. Spectral expansion of random sum complexes. *Journal of Topology and Analysis*, 12(04):989–1002, 2020.

Matteo Biagetti, Alex Cole, and Gary Shiu. The Persistence of Large Scale Structures I: Primordial non-Gaussianity. *JCAP*, 04:061, 2021. doi: 10.1088/1475-7516/2021/04/061.

Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.

Michel Boyer, Gilles Brassard, Peter Høyer, and Alain Tapp. Tight bounds on quantum searching. *Fortschritte der Physik: Progress of Physics*, 46(4-5):493–505, 1998.

Zvika Brakerski and Omri Shmueli. (pseudo) random quantum states with binary phase. In *Theory of Cryptography Conference*, pp. 229–250. Springer, 2019.

Sergey Bravyi, David Gosset, and Robert König. Quantum advantage with shallow circuits. *Science*, 362(6412):308–311, 2018.

Chris Cade and P Marcos Crichigno. Complexity of supersymmetric systems and the cohomology problem. *arXiv preprint arXiv:2107.00011*, 2021.

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.

Nadiia Chepurko, Kenneth L Clarkson, Lior Horesh, and David P Woodruff. Quantum-inspired algorithms from randomized numerical linear algebra. *arXiv preprint arXiv:2011.04125*, 2020.

Nai-Hui Chia, András Gilyén, Tongyang Li, Han-Hsuan Lin, Ewin Tang, and Chunhao Wang. Sampling-based sublinear low-rank matrix arithmetic framework for dequantizing quantum machine learning. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pp. 387–400, 2020.

Alex Cole and Gary Shiu. Persistent Homology and Non-Gaussianity. *JCAP*, 03:025, 2018a. doi: 10.1088/1475-7516/2018/03/025.

Alex Cole and Gary Shiu. Persistent homology and non-Gaussianity. *Journal of Cosmology and Astroparticle Physics*, 2018(03):025, 2018b.

Marcos Crichigno and Tamara Kohler. Clique homology is qma1-hard. *arXiv preprint arXiv:22209.11793*, 2022.

Abhinav Deshpande, Arthur Mehta, Trevor Vincent, Nicolás Quesada, Marcel Hinsche, Marios Ioannou, Lars Madsen, Jonathan Lavoie, Haoyu Qi, Jens Eisert, et al. Quantum computational advantage via high-dimensional gaussian boson sampling. *Science advances*, 8(1):eabi7894, 2022.

David Deutsch. Quantum theory, the Church–Turing principle and the universal quantum computer. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 400(1818): 97–117, 1985.

David P DiVincenzo. The physical implementation of quantum computation. *Fortschritte der Physik: Progress of Physics*, 48(9-11):771–783, 2000.

Job Feldbrugge, Matti van Engelen, Rien van de Weygaert, Pratyush Pranav, and Gert Vegter. Stochastic Homology of Gaussian vs. non-Gaussian Random Fields: Graphs towards Betti Numbers and Persistence Diagrams. *JCAP*, 09:052, 2019a. doi: 10.1088/1475-7516/2019/09/052.

Job Feldbrugge, Matti Van Engelen, Rien van de Weygaert, Pratyush Pranav, and Gert Vegter. Stochastic homology of Gaussian vs. non-Gaussian random fields: graphs towards betti numbers and persistence diagrams. *Journal of Cosmology and Astroparticle Physics*, 2019(09):052, 2019b.

Paul Fendley and Kareljan Schoutens. Exact results for strongly correlated fermions in 2+ 1 dimensions. *Physical review letters*, 95(4):046403, 2005.

Richard P Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6/7), 1982.

Paraskevi Fika and Christos Koukouvinos. Stochastic estimates for the trace of functions of matrices via Hadamard matrices. *Communications in Statistics-Simulation and Computation*, 46(5):3491–3503, 2017.

Joel Friedman. Computing Betti numbers via combinatorial Laplacians. *Algorithmica*, 21(4):331–346, 1998.

Robert Ghrist. Barcodes: the persistent topology of data. *Bulletin of the American Mathematical Society*, 45(1):61–75, 2008.

András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pp. 193–204, 2019.

Chad Giusti, Eva Pastalkova, Carina Curto, and Vladimir Itskov. Clique topology reveals intrinsic geometric structure in neural correlations. *Proceedings of the National Academy of Sciences*, 112 (44):13455–13460, 2015.

Timothy E Goldberg. Combinatorial Laplacians of simplicial complexes. *Senior Thesis, Bard College*, 2002.

Daniel Gottesman. An introduction to quantum error correction and fault-tolerant quantum computation. In *Quantum information science and its contributions to mathematics, Proceedings of Symposia in Applied Mathematics*, volume 68, pp. 13–58, 2010.

Lov K Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pp. 212–219, 1996.

Anna Gundert and Uli Wagner. On eigenvalues of random complexes. *Israel Journal of Mathematics*, 216(2):545–582, 2016.

Sam Gunn and Niels Kornerup. Review of a quantum algorithm for Betti numbers. *arXiv preprint arXiv:1906.07673*, 2019.

William H Guss and Ruslan Salakhutdinov. On characterizing the capacity of neural networks using algebraic topology. *arXiv preprint arXiv:1802.04443*, 2018.

Aldo Guzman-Saenz, Niina Haiminen, Basum Saugata, and Laxmi Parida. Signal enrichment with strain-level resolution in metagenomes using topological data analysis. *BMC Genomics*, 20:25–34, 2019.

Casper Gyurik, Chris Cade, and Vedran Dunjko. Towards quantum advantage for topological data analysis. *arXiv preprint arXiv:2005.02607*, 2020.

Aram W Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Physical review letters*, 103(15):150502, 2009.

Vojtěch Havlíček, Antonio D Córcoles, Kristan Temme, Aram W Harrow, Abhinav Kandala, Jerry M Chow, and Jay M Gambetta. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567(7747):209–212, 2019.

Alexander Semenovich Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problemy Peredachi Informatsii*, 9(3):3–11, 1973.

Honeywell. Quantinuum H1 series trapped-ion quantum processor powered by honeywell. , 2022. URL https://www.quantinuum.com/products/h1.

Danijela Horak and Jürgen Jost. Spectra of combinatorial Laplace operators on simplicial complexes. *Advances in Mathematics*, 244:303–336, 2013.

Hsin-Yuan Huang, Michael Broughton, Jordan Cotler, Sitan Chen, Jerry Li, Masoud Mohseni, Hartmut Neven, Ryan Babbush, Richard Kueng, John Preskill, et al. Quantum advantage in learning from experiments. *Science*, 376(6598):1182–1186, 2022.

MF Hutchinson. A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines. *Communications in Statistics-Simulation and Computation*, 19(2):433–450, 1990.

P Jordan and E Wigner. Über das paulische äquivalenzverbot. *Zeitschrift für Physik*, 47(9):631–651, 1928.

Matthew Kahle. Random simplicial complexes. *arXiv preprint arXiv:1607.07069*, 2016.

Robbie King and Tamara Kohler. Promise clique homology on weighted graphs is QMA1-hard and contained in QMA. *arXiv preprint arXiv:2311.17234*, 2023.

Emanuel Knill and Raymond Laflamme. Power of one bit of quantum information. *Physical Review Letters*, 81(25):5672, 1998.

Antti Knowles and Ron Rosenthal. Eigenvalue confinement and spectral gap for random simplicial complexes. *Random Structures & Algorithms*, 51(3):506–537, 2017.

Alan Lew. Spectral gaps, missing faces and minimal degrees. *Journal of Combinatorial Theory, Series A*, 169:105127, 2020a.

Alan Lew. The spectral gaps of generalized flag complexes and a geometric Hall-type theorem. *International Mathematics Research Notices*, 2020(11):3364–3395, 2020b.

Lek-Heng Lim. Hodge Laplacians on graphs. *arXiv:1507.05379 [cs, math]*, August 2019. URL http://arxiv.org/abs/1507.05379. arXiv: 1507.05379.

Lin Lin. Lecture notes on quantum algorithms for scientific computation. *arXiv preprint arXiv:2201.08309*, 2022.

Yunchao Liu, Srinivasan Arunachalam, and Kristan Temme. A rigorous and robust quantum speed-up in supervised machine learning. *Nature Physics*, 17(9):1013–1017, 2021.

Seth Lloyd. Universal quantum simulators. *Science*, pp. 1073–1078, 1996.

Seth Lloyd, Silvano Garnerone, and Paolo Zanardi. Quantum algorithms for topological and geometric analysis of data. *Nature communications*, 7:10138, 2016.

Guang Hao Low and Isaac L Chuang. Optimal Hamiltonian simulation by quantum signal processing. *Physical review letters*, 118(1):010501, 2017.

Guang Hao Low and Isaac L Chuang. Hamiltonian simulation by qubitization. *Quantum*, 3:163, 2019.

Lars S Madsen, Fabian Laudenbach, Mohsen Falamarzi Askarani, Fabien Rortais, Trevor Vincent, Jacob FF Bulmer, Filippo M Miatto, Leonhard Neuhaus, Lukas G Helt, Matthew J Collins, et al. Quantum computational advantage with a programmable photonic processor. *Nature*, 606(7912): 75–81, 2022.

Sayan Mandal, Aldo Guzman-Saenz, Niina Haiminen, Saugata Basu, and Laxmi Parida. A Topological Data Analysis Approach on Predicting Phenotypes from Gene Expression Data. In *Algorithms for Computational Biology*, pp. 178–187. Springer, 2020a.

Sayan Mandal, Aldo Guzmán-Sáenz, Niina Haiminen, Saugata Basu, and Laxmi Parida. A topological data analysis approach on predicting phenotypes from gene expression data. *Algorithms for Computational Biology*, 12099:178, 2020b.

Clément Maria, Jean-Daniel Boissonnat, Marc Glisse, and Mariette Yvinec. The gudhi library: Simplicial complexes and persistent homology. In *International congress on mathematical software*, pp. 167–174. Springer, 2014.

Vasileios Maroulas, Farzana Nasrin, and Christopher Oballe. A bayesian framework for persistent homology. *SIAM Journal on Mathematics of Data Science*, 2(1):48–74, 2020.

Tomoyuki Morimae, Keisuke Fujii, and Joseph F Fitzsimons. Hardness of classically simulating the one-clean-qubit model. *Physical review letters*, 112(13):130502, 2014.

Cameron Musco and Christopher Musco. Randomized block krylov methods for stronger and faster approximate singular value decomposition. *Advances in neural information processing systems*, 28, 2015.

Gregory Naitzat, Andrey Zhitnikov, and Lek-Heng Lim. Topology of deep neural networks. *J. Mach. Learn. Res.*, 21(184):1–40, 2020.

Michael A Nielsen and Isaac L Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2010.

Bryan O'Gorman, William J Huggins, Eleanor G Rieffel, and K Birgitta Whaley. Generalized swap networks for near-term quantum computing. *arXiv preprint arXiv:1905.05118*, 2019.

Edwin Pednault, John A Gunnels, Giacomo Nannicini, Lior Horesh, Thomas Magerlein, Edgar Solomonik, Erik W Draeger, Eric T Holland, and Robert Wisnieff. Pareto-efficient quantum circuit simulation using tensor contraction deferral. *arXiv preprint arXiv:1710.05867*, 2017.

Edwin Pednault, John A Gunnels, Giacomo Nannicini, Lior Horesh, and Robert Wisnieff. Leveraging secondary storage to simulate deep 54-qubit sycamore circuits. *arXiv preprint arXiv:1910.09534*, 2019.

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):1–7, 2014.

John Preskill. Quantum computing in the NISQ era and beyond. *Quantum*, 2:79, 2018.

Raúl Rabadán, Yamina Mohamedi, Udi Rubin, Tim Chu, Adam N Alghalith, Oliver Elliott, Luis Arnés, Santiago Cal, Álvaro J Obaya, Arnold J Levine, et al. Identification of relevant genetic alterations in cancer using topological data analysis. *Nature communications*, 11(1):1–10, 2020.

Patrick Rebentrost, Masoud Mohseni, and Seth Lloyd. Quantum support vector machine for big data classification. *Physical review letters*, 113(13):130503, 2014.

Christian Reiher. The clique density theorem. *Annals of Mathematics*, pp. 683–707, 2016.

Yosef Rinott, Tomer Shoham, and Gil Kalai. Statistical aspects of the quantum supremacy demonstration. *Statistical Science*, 37(3):322–347, 2022.

Alexander Schmidhuber and Seth Lloyd. Complexity-theoretic limitations on quantum algorithms for topological data analysis. *arXiv preprint arXiv:2209.14286*, 2022.

Maria Schuld and Nathan Killoran. Quantum machine learning in feature Hilbert spaces. *Physical review letters*, 122(4):040504, 2019.

Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. An introduction to quantum machine learning. *Contemporary Physics*, 56(2):172–185, 2015.

Vivek V Shende, Stephen S Bullock, and Igor L Markov. Synthesis of quantum-logic circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 25(6):1000–1010, 2006.

Peter W Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th annual symposium on foundations of computer science*, pp. 124–134. Ieee, 1994.

Peter W Shor. Fault-tolerant quantum computation. In *Proceedings of 37th Conference on Foundations of Computer Science*, pp. 56–65. IEEE, 1996.

Ewin Tang. Quantum principal component analysis only achieves an exponential speedup because of its state preparation assumptions. *Physical Review Letters*, 127(6):060503, 2021.

Lloyd N Trefethen. *Approximation Theory and Approximation Practice, Extended Edition*. SIAM, 2019.

Shashanka Ubaru and Yousef Saad. Fast methods for estimating the numerical rank of large matrices. In *Proceedings of The 33rd International Conference on Machine Learning*, pp. 468–477, 2016.

Shashanka Ubaru, Yousef Saad, and Abd-Krim Seghouane. Fast estimation of approximate matrix ranks using spectral densities. *Neural computation*, 29(5):1317–1351, 2017.

Larry Wasserman. Topological data analysis. *Annual Review of Statistics and Its Application*, 5: 501–532, 2018.

John Watrous. Quantum computational complexity. In *Computational Complexity: Theory, Techniques, and Applications*, pp. 2361–2387. Springer New York, 2012.

Taiki Yamada. Spectrum of the Laplacian on simplicial complexes by the Ricci curvature. *arXiv preprint arXiv:1906.07404*, 2019.

Andrew Zhao, Andrew Tranter, William M. Kirby, Shu Fay Ung, Akimasa Miyake, and Peter J. Love. Measurement reduction in variational quantum algorithms. *Phys. Rev. A*, 101:062322, Jun 2020. doi: 10.1103/PhysRevA.101.062322. URL https://link.aps.org/doi/10.1103/PhysRevA.101.062322.

Han-Sen Zhong, Hui Wang, Yu-Hao Deng, Ming-Cheng Chen, Li-Chao Peng, Yi-Han Luo, Jian Qin, Dian Wu, Xing Ding, Yi Hu, et al. Quantum computational advantage using photons. *Science*, 370(6523):1460–1463, 2020.

Leo Zhou, Sheng-Tao Wang, Soonwon Choi, Hannes Pichler, and Mikhail D Lukin. Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices. *Physical Review X*, 10(2):021067, 2020.

Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. *Discrete & Computational Geometry*, 33(2):249–274, 2005.