

## QUANTUM TOPOLOGICAL DATA ANALYSIS WITH CONTINUOUS VARIABLES

GEORGE SIOPSIS

Department of Physics and Astronomy  
The University of Tennessee  
Knoxville, Tennessee 37996-1200, USA

**ABSTRACT.** I introduce a continuous-variable quantum topological data algorithm. The goal of the quantum algorithm is to calculate the Betti numbers in persistent homology which are the dimensions of the kernel of the combinatorial Laplacian. I accomplish this task with the use of qRAM to create an oracle which organizes sets of data. I then perform a continuous-variable phase estimation on a Dirac operator to get a probability distribution with eigenvalue peaks. The results also leverage an implementation of continuous-variable conditional swap gate.

**1. Introduction.** Extracting useful information from data sets is a difficult task, and in large cases it can be impossible on a classical computer. It is an ongoing field of research to produce quantum algorithms which can analyze data at large scales [27, 23, 5, 11, 20]. Topological methods for data analysis allow for general useful features of the data to be revealed, and these features do not depend of the representation of the data or any additional noise. This makes topological techniques a powerful analytical tool [45, 37, 13, 7, 12, 46, 8, 9, 2, 3, 4, 34, 19, 32]. These methods classically scale with exponential computing time, but have been shown to be a great example of the power of quantum algorithms [27, 11, 40, 20].

In this work, I follow and build upon the results in [27], focusing on finding the Betti numbers in persistent homology. Persistent homology is a topological method that revolves around representing a space in terms of a simplicial complex and examining the application of a scaled boundary operator. The simplicial complex is built on a set of points sampled from the space and has to satisfy certain theoretical properties dictated by topological inference, e.g., it has to approximate the homology of the space. In particular, the Vietoris-Rips complex [36], which is discussed here, is appealing because it can be computed efficiently. One deduces the Betti numbers which represent features of the data, such as the number of connected components, holes, and voids. In order to compute the Vietoris-Rips complex and determine the Betti numbers, I use a quantum algorithm which employs the tools of continuous-variable (CV) quantum computation. More specifically, I use quantum principal component analysis (QPCA) to resolve a spectrum of Betti numbers [23].

---

2010 *Mathematics Subject Classification.* Primary: 62-07; Secondary: 81P68, 81P70.

*Key words and phrases.* Quantum algorithm, quantum random access memory, continuous-variable quantum computation, topological data analysis, persistent homology, Vietoris-Rips complex, combinatorial Laplacian, Betti numbers.

A CV quantum system is one that utilizes an infinite-dimensional Hilbert space, where the measurement of variables produces a continuous result. This is a subject that is being studied extensively, and shown to have applications in generating entanglement, quantum cryptography, quantum teleportation, and quantum computation [39, 6, 26, 30, 44, 31, 41, 35, 38, 18, 1]. The use of a continuous system provides advantages over a qubit, or discrete-variable (DV) system, such as low cost of optical components, less need for environmental control, and potentially better scaling to larger problems [23, 39, 26, 35, 38, 43, 42, 28, 18]. The CV substrate has also been demonstrated to be more useful in situations with high rate of information transfer such as computing on encrypted data [29, 43, 38, 42]. These advantages can be very useful when analyzing and extracting information on large volumes of classical data, and as a result make CV quantum computing the natural choice in this setting.

The body of this work starts by discussing persistent homology in a general way, and by setting up the mathematical background of the algorithm, including some useful definitions such as the combinatorial Laplacian and the  $k$ th Betti number. I then define the use of quantum Random Access Memory (qRAM) which allows a mapping of classical data into a set of quantum states [15, 16, 10]. Similar to RAM in a classical computer, qRAM consists of a memory array and input(address)/output registers. Unlike RAM, the registers in qRAM are quantum (superpositions of states), which makes physical implementation rather challenging due to fast decoherence of complex quantum registers.

In addition, I outline the process of exponentiation of a Hermitian operator, and arrive at the construction of an oracle that returns the elements of the  $k$ th Vietoris-Rips complex. Finally, I provide the CV quantum algorithm which uses the process of QPCA, utilizing an implementation of a hybrid [25] exponential conditional swap, to determine the Betti numbers of the system.

The discussion is organized as follows. In Section 2, I discuss in general the steps involved in persistent homology as well as some basics in topological data analysis. In Section 3, I introduce pertinent mathematical background needed to set up the algorithm. Section 4 describes the usage of qRAM, exponentiation, and the oracle. Section 5 outlines the algorithm. I offer a discussion and a conclusion in Section 6.

**2. Background.** The final goal of topological data analysis, along with the algorithm introduced here, is to determine interesting features of a data set. In this case, the indicator of structure is the Betti numbers, which are a count of topological features. The Betti numbers distinguish between topological spaces based on their connectivity, and are grouped based on dimension. The common notation for Betti numbers is  $\beta_k$ , where for  $k = 0, 1, 2$  one has the Betti numbers that correspond to connected components, one-dimensional holes, and two-dimensional voids, respectively.

As an example, see Figure 1, where I consider some simple topological surfaces in one to three dimensions, and list the values of the  $k = 0, 1, 2$  Betti numbers. The algorithm introduced in this work allows one to find the Betti numbers after representing some given data in a space of vertices with connecting edges. Starting with the data, one can create this representation in the following steps.

- (a) Start by allowing each data point to represent a position vector, and place one point at the end of each of these vectors. These points are referred to as the vertices.

- (b) Next, for a given  $\epsilon$  diameter, draw a circle around each vertex in the space.
- (c) Between every two vertices which have contacting circles (and as such are less than  $\epsilon$  distance apart), draw a connecting line. These connections are edges of  $n$ -dimensional shapes called simplices, and the space of simplices is called a simplicial complex.

This process is visualized in Figure 2. Now, in order to begin to analyze this representation of the data on a quantum computer, they must be first encoded in a quantum state. This is the subject of the next section.





Point	Circle	Sphere	Torus
			
$\beta_0 = 1$	$\beta_0 = 1$	$\beta_0 = 1$	$\beta_0 = 1$
$\beta_1 = 0$	$\beta_1 = 1$	$\beta_1 = 0$	$\beta_1 = 2$
$\beta_2 = 0$	$\beta_2 = 0$	$\beta_2 = 1$	$\beta_2 = 1$

FIGURE 1. The Betti numbers  $\beta_{0,1,2}$  for four example shapes (point, circle, spherical shell, and torus). They are the number of connected components, one-dimensional holes (also called tunnels or handles), and two-dimensional voids, respectively.

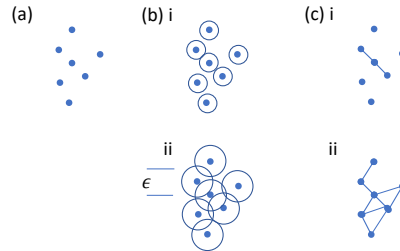


FIGURE 2. (a) Given data represented by points. (b) For a given distance  $\epsilon$ , a circle is drawn around each point. (c) Between every two points with contacting circles a line is drawn. These connections are edges of  $n$ -dimensional shapes (simplices), and the space of simplices in (c) is called a simplicial complex. For two different values of  $\epsilon$ , as in (b) i, ii, and (c) i, ii, one can get more or less connections between the data points resulting in different topologies. Therefore Betti numbers depend on the initial choice of  $\epsilon$ . It is useful to vary  $\epsilon$  to find interesting structures.

**3. Initialization of algorithm.** To start, we are given  $n$  points in a  $d$ -dimensional space at position vectors  $\mathbf{v}_i$ ,  $i = 1, 2, \dots, n$ . For simplicity, assume that all points are on the unit sphere,  $|\mathbf{v}_i| = 1$ . More general sets of points can be considered by extending the discussion in a straightforward, albeit somewhat tedious, manner. For each vector, construct the quantum state

$$|\mathbf{v}_i\rangle = \sum_{j=0}^{d-1} (\mathbf{v}_i)_j |j\rangle, \quad (1)$$

using  $\log_2 d$  qubits. This is analogous to the first part of Figure 2 where data are represented with a series of dots of variable distance from one another.

A  $k$ -simplex  $s_k$  is defined as a simplex consisting of  $k + 1$  vertices at points  $\mathbf{v}_{i_0}, \dots, \mathbf{v}_{i_k}$  connected with  $k(k + 1)/2$  edges. The first four  $k$ -simplices are shown in Figure 3. We wish to map the simplex  $s_k$  that can be represented by  $n$  classical bits to a quantum state of  $n$  qubits that will be named  $|s_k\rangle$ . To this end, note that a simplex can be represented by a string of  $n$  bits consisting of  $k + 1$  1s at positions  $i_0, \dots, i_k$ , and 0s otherwise (for an example of this representation see Figure 4). Let  $s_k$  denote the number written as this string of  $n$  bits in binary notation ( $s_k = 0, 1, \dots, 2^n - 1$ , if all values of  $k$  are considered). The state  $|s_k\rangle$  is constructed using  $n$  qubits. Thus, simplices are mapped onto basis vectors  $|s_k\rangle$ .

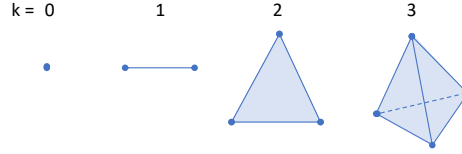


FIGURE 3. The  $k$ -simplices for  $k = 0, 1, 2, 3$ . These are a vertex, an edge, a triangle, and a tetrahedron, respectively.

Next, define the diameter  $\mathcal{D}(s_k)$  as the maximum distance between two vertices of the simplex,

$$\mathcal{D}(s_k) = \max_{i_l, i_{l'}} |\mathbf{v}_{i_l} - \mathbf{v}_{i_{l'}}|. \quad (2)$$

This allows one to define the Vietoris-Rips complex  $S_k^\epsilon$  as the complex consisting of all  $k$ -simplices with diameter  $\mathcal{D} \leq \epsilon$ , for a given scale  $\epsilon$ . The construction of the Vietoris-Rips complex is equivalent to the latter pair of steps in Figure 2, where circles of diameter  $\epsilon$  are drawn around each vertex, and then the vertices of contacting circles are connected with edges. The objective of persistent homology is to continuously vary the scale  $\epsilon$  until one finds a value which gives the space an interesting structure, as determined by the Betti numbers. The word persistent comes from this varying of  $\epsilon$ , whereas homology is the algebraic tool that measures the structure of the complex. For the algorithm in the Hilbert space of  $n$  qubits, one can define the projection operator

$$P_k^\epsilon = \sum_{s_k \in S_k^\epsilon} |s_k\rangle\langle s_k|, \quad (3)$$

onto  $S_k^\epsilon$ . Evidently, for  $\epsilon \geq 2$ , all  $k$ -simplices are included in  $S_k^\epsilon$ , so one need only consider  $\epsilon \in (0, 2)$ . The parameter  $\epsilon$  can be encoded using  $m$  qubits as  $|x\rangle$ ,  $x = 0, 1, \dots, 2^m - 1$ , where  $\epsilon = x/2^{m-1}$ .

If one removes the  $l$ th vertex from the  $k$ -simplex  $s_k$ , one obtains the  $(k - 1)$ -simplex  $s_{k-1}(l)$ ,  $l = 0, 1, \dots, k$ . Evidently,

$$|s_{k-1}(l)\rangle = X_{i_l} |s_k\rangle, \quad (4)$$

where  $X_{i_l}$  is the first Pauli matrix,  $X = |0\rangle\langle 1| + |1\rangle\langle 0|$ , acting on the  $i_l$ th qubit. Let us now probe the space to determine whether or not any interesting features are present. This probing is done by the boundary map acting on the space. Define

the boundary map  $\partial_k$  by

$$\partial_k |s_k\rangle = \sum_{l=0}^k (-)^l |s_{k-1}(l)\rangle = \sum_{l=0}^k (-)^l X_{i_l} |s_k\rangle. \quad (5)$$

One easily deduces  $\partial_k \partial_{k+1} = 0$ . To restrict to Vietoris-Rips complexes, introduce

$$\tilde{\partial}_k^\epsilon = P_{k-1}^\epsilon \partial_k P_k^\epsilon. \quad (6)$$

The action of the boundary operator on a simplex, as well as an example encoding into bits is visualized in Figure 4.

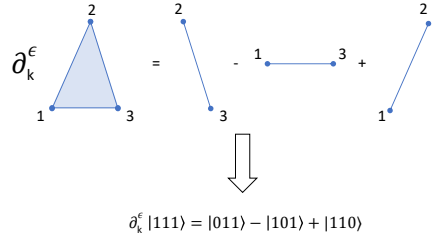


FIGURE 4. The action of the boundary operator is shown on a  $k = 2$  simplex. A visual representation of a simplex being broken down into its boundary is depicted above. Its boundary consists of simplices of  $k - 1 = 1$ . Below is the encoded representation of the boundary operator acting on the 2-simplex. In this encoding a 1 represents a vertex in the corresponding position in the string of bits. The boundary sum is represented by a clockwise rotation around the original simplex, and the negative sign in the result alternates as in Eq. (5).

The entire Hilbert space of  $n$  qubits is split into  $n + 1$  subspaces labeled by  $k$ . To keep track of this splitting, I introduce a register of  $\log_2 n$  qubits to store the state  $|k\rangle$  and map

$$|s_k\rangle \mapsto |k\rangle |s_k\rangle, \quad (7)$$

in parallel. This can be done in  $n$  steps as follows. Start with the state  $|0\rangle$  for the register. Apply the permutation  $\mathcal{P} : |0\rangle \rightarrow |1\rangle \rightarrow \cdots \rightarrow |n-1\rangle \rightarrow |0\rangle$  for each digit of  $|s_k\rangle$  equal to 1 (using the qubit corresponding to each digit as control). The permutation  $\mathcal{P}$  is a 1-sparse matrix and can be implemented efficiently. Thus one applies  $\mathcal{P}^k$ , so  $|0\rangle \rightarrow \mathcal{P}^k |0\rangle = |k\rangle$ , as desired.

A general state can be written as

$$|\Psi\rangle = \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} |k\rangle |\Psi_k\rangle = \frac{1}{\sqrt{n}} \begin{pmatrix} |\Psi_0\rangle \\ |\Psi_1\rangle \\ \vdots \\ |\Psi_n\rangle \end{pmatrix}, \quad (8)$$

where  $|\Psi_k\rangle$  is in the span of  $\{|s_k\rangle\}$ .

Define the Dirac operator as the Hermitian matrix

$$\tilde{B}^\epsilon = \begin{pmatrix} 0 & \tilde{\partial}_1^\epsilon & 0 & & \\ \tilde{\partial}_1^{\epsilon\dagger} & 0 & \tilde{\partial}_2^\epsilon & & \\ 0 & \tilde{\partial}_2^{\epsilon\dagger} & 0 & & \\ & & & \ddots & \\ & & & & 0 & \tilde{\partial}_n^\epsilon \\ & & & & \tilde{\partial}_n^{\epsilon\dagger} & 0 \end{pmatrix} \quad (9)$$

One easily obtains

$$(\tilde{B}^\epsilon)^2 = \text{diag}(\Delta_0^\epsilon, \Delta_1^\epsilon, \dots, \Delta_n^\epsilon), \quad (10)$$

where  $\Delta_0^\epsilon = \tilde{\partial}_1^{\epsilon\dagger} \tilde{\partial}_1^\epsilon$ ,  $\Delta_n^\epsilon = \tilde{\partial}_n^{\epsilon\dagger} \tilde{\partial}_n^\epsilon$ , and

$$\Delta_k^\epsilon = \tilde{\partial}_k^{\epsilon\dagger} \tilde{\partial}_k^\epsilon + \tilde{\partial}_{k+1}^{\epsilon\dagger} \tilde{\partial}_{k+1}^\epsilon, \quad k = 1, \dots, n-1 \quad (11)$$

is the combinatorial Laplacian of the  $k$ th simplicial complex. The output of a  $k$ th combinatorial Laplacian being zero tells us exactly that we have found a space which is boundary less and also not a boundary itself. The number of these features in our data is the Betti number. Therefore, one can say that the dimension of the kernel of  $\Delta_k^\epsilon$  is the  $k$ th Betti number,

$$\beta_k = |\ker \Delta_k^\epsilon|. \quad (12)$$

An example of determining a void in a  $k = 2$  complex is shown in Figure 5. The total number of voids is the Betti number  $\beta_2$ .

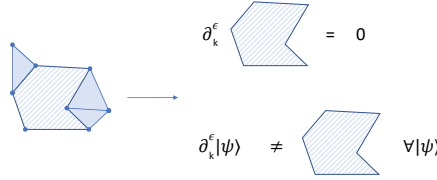


FIGURE 5. Consider the  $k = 2$  complex on the left, for a given value of  $\epsilon$ . In order to show that the striped area is a void, it itself must be boundary-less, and not a boundary for any part of the complex. Fulfillment of these two properties is equivalent to the combinatorial Laplacian (11) applied to the stripped area returning zero. Therefore this area would be part of the kernel of the combinatorial Laplacian for  $k = 2$  contributing to the  $\beta_2$  Betti number.

To summarize, when starting with a data set, the following steps are needed in order to perform persistent homology using the quantum algorithm presented in this paper.

- Start with  $n$  points called vertices in a  $d$ -dimensional space defined by a set of position vectors. For each vector, the quantum state (1) is constructed.
- Using the diameter  $\mathcal{D}(s_k)$  defined in (2), the vertices are connected to form the Vietoris-Rips complex  $S_k^\epsilon$  consisting of all  $k$ -simplices of diameter  $\leq \epsilon$ .
- The space is then split into subspaces labeled by  $k$ , consisting of all  $k$ -simplices. A general state is then constructed which spans all of these subspaces, and is given in Eq. (8).

- (d) In order to probe the space for interesting structures, use the boundary map (5). To determine if a region of the space is one of the features which are tallied up to become the Betti numbers (Holes, Voids, etc.), this region must be boundary-less and also not a boundary of any other part of the space, as shown in Fig. 5. These two properties are satisfied when the action of the combinatorial Laplacian (11) returns zero. The combinatorial Laplacians of  $k$ th order are the elements of the diagonal matrix which is the square of the Dirac operator (9).
- (e) In order to apply this operator in the CV algorithm presented here, and to construct some of the states mentioned above, some additional mathematical tools are required, and outlined in the following section.

**4. Mathematical tools.** In this section, the tools needed in the CV quantum algorithm are outlined, and the way they are used and implemented is discussed.

In order to complete the quantum algorithm, We assume that we are equipped with a qRAM [15, 16, 10] which, given an input state  $|i\rangle|0\rangle$ , produces the output state  $|i\rangle|v_i\rangle$  in quantum parallel,

$$\text{qRAM} : \sum_{i=0}^{n-1} a_i |i\rangle |0\rangle \longrightarrow \sum_{i=0}^{n-1} a_i |i\rangle |v_i\rangle. \quad (13)$$

Another useful tool which will be used in the algorithm is the exponentiation of an operator. Given a Hermitian operator  $\mathbf{A}$ , and a resource qumode of quadratures  $(q_{\mathcal{R}}, p_{\mathcal{R}})$ , it is necessary to apply

$$e^{ip_{\mathcal{R}}\mathbf{A}t}, \quad (14)$$

in parallel. To this end, I will use the exponential swap operator

$$e^{i\theta p_{\mathcal{R}}\mathcal{S}}, \quad (15)$$

where  $\mathcal{S}$  is the swap operator. Its implementation is discussed in appendix A (and differs from the one given in [23]). While the body of this work uses a CV quantum algorithm, the method of implementing the exponential conditional swap also uses single photon qubits in a hybrid approach [25]. The latter can also be implemented using CV systems in the dual rail representation.

Then we form the state  $\rho_{\mathbf{A}} = \frac{\mathbf{A}}{\text{tr}\mathbf{A}}$  (assuming  $\text{tr}\mathbf{A} \neq 0$ ), and apply (15) on the combined system of  $|\Psi\rangle$  and  $\rho_{\mathbf{A}}$ , for a short time  $\delta t$ . After tracing out the auxiliary mode  $\rho_{\mathbf{A}}$ , we obtain

$$\begin{aligned} & \text{tr}_{\rho_{\mathbf{A}}} (e^{i\delta t p_{\mathcal{R}}\mathcal{S}} |\Psi\rangle\langle\Psi| \otimes \rho_{\mathbf{A}} e^{-i\delta t p_{\mathcal{R}}\mathcal{S}}) \\ &= e^{i\delta t p_{\mathcal{R}}\rho_{\mathbf{A}}} |\Psi\rangle\langle\Psi| e^{-i\delta t p_{\mathcal{R}}\rho_{\mathbf{A}}} + \mathcal{O}((\delta t)^2) \\ &= e^{i\delta t(\text{tr}\mathbf{A})p_{\mathcal{R}}\mathbf{A}} |\Psi\rangle\langle\Psi| e^{-i\delta t(\text{tr}\mathbf{A})p_{\mathcal{R}}\mathbf{A}} + \mathcal{O}((\delta t)^2). \end{aligned} \quad (16)$$

By repeating this  $\frac{t}{\delta t(\text{tr}\mathbf{A})}$  times, an approximation to the desired operator (14) is obtained.

We also assume we are in possession of a quantum oracle  $\mathcal{O}_k^\epsilon$  that acts on  $|s_k\rangle|\psi\rangle$  in parallel, flipping the last qubit if  $s_k \in S_k^\epsilon$ , otherwise doing nothing,

$$\mathcal{O}_k^\epsilon |s_k\rangle|\psi\rangle = \begin{cases} |s_k\rangle \otimes X|\psi\rangle & , \quad s_k \in S_k^\epsilon \\ |s_k\rangle|\psi\rangle & , \quad s_k \notin S_k^\epsilon \end{cases} \quad (17)$$

This oracle can be implemented in  $\mathcal{O}(k^2)$  steps. If we choose  $|\psi\rangle = |-\rangle$ , where  $X|\pm\rangle = \pm|\pm\rangle$ , then the last qubit decouples and the oracle is a unitary acting on

$|s_k\rangle$  as

$$\mathcal{O}_k^\epsilon |s_k\rangle = \begin{cases} -|s_k\rangle & , \quad s_k \in S_k^\epsilon \\ |s_k\rangle & , \quad s_k \notin S_k^\epsilon \end{cases} \quad (18)$$

To construct the oracle, first we construct the state

$$\frac{1}{n} \sum_{i,j=0}^{n-1} |i\rangle |j\rangle, \quad (19)$$

by making two copies of the state  $\frac{1}{\sqrt{n}} \sum_i |i\rangle$ . To this state we attach  $|\mathbf{0}\rangle \in \mathbb{C}^d$  as well as a qubit in the state  $|+\rangle$ . We then query qRAM to obtain

$$\frac{1}{n} \sum_{i,j=0}^{n-1} |i\rangle |j\rangle |\mathbf{v}_i\rangle |+\rangle. \quad (20)$$

Then we use the last qubit as control to apply the swap operator and obtain

$$\frac{1}{\sqrt{2}n} \sum_{i,j=0}^{n-1} |i\rangle |j\rangle (|\mathbf{v}_i\rangle |0\rangle + |\mathbf{v}_j\rangle |1\rangle). \quad (21)$$

Then we measure  $X$  on the last qubit. If the outcome is  $-1$ , the state collapses to (unnormalized)

$$\sum_{i,j=0}^{n-1} |i\rangle |j\rangle (|\mathbf{v}_i\rangle - |\mathbf{v}_j\rangle) |-\rangle. \quad (22)$$

We then add ancillae and copy the labels  $i$  and  $j$  on them, respectively. We obtain the state (ignoring the last qubit which has decoupled)

$$\sum_{i,j=0}^{n-1} |i\rangle |j\rangle |i\rangle_A |j\rangle_A (|\mathbf{v}_i\rangle - |\mathbf{v}_j\rangle). \quad (23)$$

Tracing out the ancillae and the last qubit, we obtain the (unnormalized) state

$$H = \sum_{i,j=0}^{n-1} |\mathbf{v}_i - \mathbf{v}_j|^2 |i\rangle \langle i| \otimes |j\rangle \langle j|. \quad (24)$$

It is a Hermitian operator that can be implemented as  $e^{itH}$ , as discussed above. Its eigenvalues are the distances between points. Also any function of  $H$  can be implemented; in particular, the step function  $\theta(\epsilon^2 - H)$ , that tests whether  $s_k \in S_k^\epsilon$ ; hence the oracle.

**5. Quantum topological data analysis for CVs.** The algorithm requires a register of  $m$  qubits to record  $\epsilon$  as  $|x\rangle$  with  $\epsilon = x/2^{m-1} \in (0, 2)$ . Suppose  $\epsilon$  is fixed (a condition that can be relaxed to include a filtration). Let us start with the initial state that includes all  $k$ -simplices equally weighted,

$$|\Psi^{(0)}\rangle = \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} |k\rangle |\Psi_k^{(0)}\rangle = \frac{1}{\sqrt{n}} \begin{pmatrix} |\Psi_0^{(0)}\rangle \\ |\Psi_1^{(0)}\rangle \\ \vdots \\ |\Psi_n^{(0)}\rangle \end{pmatrix}, \quad (25)$$



where

$$|\Psi_k^{(0)}\rangle = \frac{1}{\sqrt{\binom{n}{k+1}}} \sum_{s_k} |s_k\rangle. \quad (26)$$

The initial state can be constructed from the state  $|0\rangle|s\rangle$ , where the register  $|0\rangle$  consists of  $\log_2 n$  qubits, and

$$|s\rangle = \frac{1}{2^{n/2}} \sum_{y=0}^{2^n-1} |y\rangle, \quad (27)$$

consists of  $n$  qubits. By using each qubit in the state  $|s\rangle$  as control to apply the permutation  $\mathcal{P}$  on the register, we arrive at the desired initial state (25).

From the initial state (25), we construct an approximation to the state

$$|\Psi^\epsilon\rangle = \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} |k\rangle |\Psi_k^\epsilon\rangle = \frac{1}{\sqrt{n}} \begin{pmatrix} |\Psi_0^\epsilon\rangle \\ |\Psi_1^\epsilon\rangle \\ \vdots \\ |\Psi_n^\epsilon\rangle \end{pmatrix}, \quad (28)$$

where

$$|\Psi_k^\epsilon\rangle = \frac{1}{\sqrt{|S_k^\epsilon|}} \sum_{s_k^\epsilon \in S_k^\epsilon} |s_k^\epsilon\rangle, \quad (29)$$

using Grover's search algorithm [17], with the aid of the oracle.

Notice that the action of the Dirac operator (9) simplifies, because all projection operators  $P_k^\epsilon$  act as the identity on  $|\Psi_k^\epsilon\rangle$ . Therefore, one could instead consider the simpler operator

$$B = \begin{pmatrix} 0 & \partial_1 & 0 & & \\ \partial_1^\dagger & 0 & \partial_2 & & \\ 0 & \partial_2^\dagger & 0 & & \\ & & & \ddots & \\ & & & & 0 & \partial_n \\ & & & & \partial_n^\dagger & 0 \end{pmatrix} \quad (30)$$

My goal is to compute the eigenvalues of  $B$ . For Betti numbers, I am interested in the frequency of occurrence of the zero eigenvalue which yields the dimension of the kernel of the combinatorial Laplacian. I will compute the eigenvalues using QPCA, as discussed in [23], which is a more specific implementation than the original work in [27] which cites the use of general Hamiltonian simulation.

Let us attach a squeezed resource qumode in the state (unnormalized)

$$\int dp_{\mathcal{R}} e^{-p_{\mathcal{R}}^2/(2s)} |p_{\mathcal{R}}\rangle \quad (31)$$

and apply the unitary

$$e^{i\gamma p_{\mathcal{R}} B} \quad (32)$$

where  $\gamma$  is a parameter that can be adjusted at will. This unitary is of the form (14), except that  $\text{tr} B = 0$ . We need to regulate  $B$ , by adding  $\alpha \mathbb{I}$ , where  $\alpha$  is arbitrary. The eigenvalues are shifted by  $\alpha$ , and  $\text{tr}(B + \alpha \mathbb{I}) \neq 0$ .

Suppose that the eigenvalue problem of  $B + \alpha \mathbb{I}$  is

$$(B + \alpha \mathbb{I})|e_i\rangle = \lambda_i|e_i\rangle \quad (33)$$

and the state is expanded as

$$|\Psi^\epsilon\rangle = \sum_i a_i |e_i\rangle \quad (34)$$

Then we obtain

$$\sum_i a_i \int dp_{\mathcal{R}} e^{i\gamma p_{\mathcal{R}} \lambda_i} e^{-p_{\mathcal{R}}^2/(2s)} |e_i\rangle |p_{\mathcal{R}}\rangle \quad (35)$$

A measurement of the quadrature  $q_{\mathcal{R}}$  of the resource qumode with homodyne detection projects the state onto

$$\sum_i a_i e^{-s(\gamma \lambda_i - q_{\mathcal{R}})^2/2} |e_i\rangle |q_{\mathcal{R}}\rangle \quad (36)$$

with the probability distribution

$$P(q_{\mathcal{R}}) \propto \sum_i |a_i|^2 e^{-s(\gamma \lambda_i - q_{\mathcal{R}})^2} \quad (37)$$

consisting of peaks at the eigenvalues. If one is interested in distinguishing between eigenvalues, one ought to choose sufficiently large parameters  $s$  and  $\gamma$  so that the width of each peak,  $1/(\gamma\sqrt{s})$ , is narrow enough. From this probability distribution, one can deduce all Betti numbers.

**6. Discussion and conclusion.** In this work, I discussed a quantum algorithm for topological data analysis using the method of persistent homology. The algorithm was designed using qRAM as well as a continuous-variable substrate to take advantage of a continuous output from which Betti numbers can be calculated. I also examined the use of a continuous-variable exponential conditional swap operation which is outlined in more detail in the Appendix. As in the discrete-variable case [27], although the matrix (30) is exponentially large ( $O(2^n)$ ) the size of the required qRAM is small. This provides an advantage over other algorithms that require a large qRAM [19, 14, 21].

In general, the use of discrete-variable quantum algorithms for topological data analysis is something that has been used before [27, 11, 40, 20]. The work presented here provides new tools for continuous-variable systems as well as a direct circuit implementation of one of those tools.

In the algorithm presented here, I used a subset of phase estimation called principal component analysis in order to determine the eigenvalues of the exponential operator. This method is a natural fit for the continuous-variable framework discussed here, but there are other methods which have been examined. For example, a hybrid approach which uses a qumode as well as a mixed state of qubits [24]. There is also the well-understood purely-qubit phase estimation [33], but this approach would require many copies of the unitary (32), greatly increasing any resource costs as a result. Note that the algorithm used in this work is largely an adaptation of the exponentiation and phase estimation of ref. [23]. The discussion of resource costs in that work can then be sufficiently translated to the present algorithm.

**Appendix A. Exponential conditional swap.** Here I discuss the implementation of the exponential swap operator (15) conditioned on the quadrature  $p_{\mathcal{R}}$  of a resource mode. Let us concentrate on qubits labeled as  $i$  and  $j$  that we wish to swap,  $S_{ij}|x\rangle_i|y\rangle_j = |y\rangle_i|x\rangle_j$ . Each qubit consists of a pair of qumodes in single-photon states,  $|10\rangle = a^\dagger|00\rangle$ ,  $|01\rangle = b^\dagger|00\rangle$ , which are identified with the computational basis vectors  $|0\rangle$ ,  $|1\rangle$ , respectively.

Let us introduce the controlled rotations

$$\text{CR}_{ij}^X(\theta) = \frac{\mathbb{I}_i + Z_i}{2} + \frac{\mathbb{I}_i - Z_i}{2} e^{i\theta X_j} = e^{i\theta b_i^\dagger b_i (a_j^\dagger b_j + a_j b_j^\dagger)} \quad (38)$$

where  $\text{CR}_{ij}^X(\frac{\pi}{2})$  is the CNOT gate with control (target) the  $i$ th ( $j$ th) qubit,

$$\text{CR}_{ij}^Y(\theta) = \frac{\mathbb{I}_i + Z_i}{2} + \frac{\mathbb{I}_i - Z_i}{2} e^{i\theta Y_j} = e^{\theta b_i^\dagger b_i (a_j^\dagger b_j - a_j b_j^\dagger)} \quad (39)$$

and

$$\text{CR}_{ij}^Z(\theta) = \frac{\mathbb{I}_i + Z_i}{2} + \frac{\mathbb{I}_i - Z_i}{2} e^{i\theta Z_j} = e^{i\theta b_i^\dagger b_i (a_j^\dagger a_j - b_j^\dagger b_j)} \quad (40)$$

where  $\text{CR}_{ij}^Z(\frac{\pi}{2})$  is the CZ gate. They can be constructed using quartic phase gates [23].

The swap gate can be transformed into a CZ gate using

$$\text{SWAP}_{ij} = U_{ij} \cdot \text{CZ} \cdot U_{ij}^\dagger, \quad U_{ij} = \text{CNOT}_{ji} \cdot \text{CR}_{ij}^Y(\frac{\pi}{4}) \quad (41)$$

The CZ gate can be conveniently implemented by introducing an ancillary qubit in the state  $|0\rangle_A \equiv a_A^\dagger |00\rangle_A$ , and using

$$\text{CZ}_{ij}|0\rangle_A = \text{CCNOT}_{ij,A} \cdot Z_A \cdot \text{CCNOT}_{ij,A}|0\rangle_A \quad (42)$$

We deduce

$$\begin{aligned} e^{itp_{\mathcal{R}} \text{SWAP}_{ij}}|0\rangle_A &= U_{ij} \cdot \text{CCNOT}_{ij,A} \cdot e^{itp_{\mathcal{R}} \cdot (a_A^\dagger a_A - b_A^\dagger b_A)} \\ &\quad \cdot \text{CCNOT}_{ij,A} \cdot U_{ij}^\dagger |0\rangle_A \end{aligned} \quad (43)$$

If the system contains multiple qubits, then the total swap operator  $\mathcal{S}$  is a product of swap operators for individual qubits,

$$\mathcal{S} = \prod_i \text{SWAP}_{ii'} \quad (44)$$

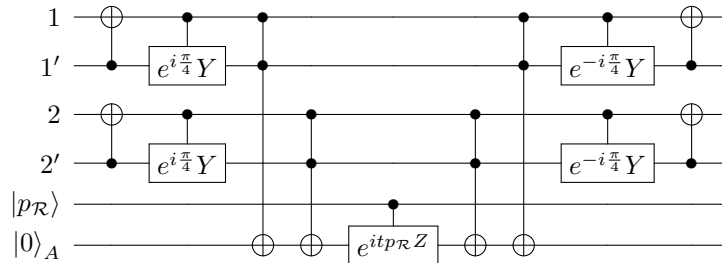
The above result can be straightforwardly extended. We obtain

$$e^{itp_{\mathcal{R}} \mathcal{S}}|0\rangle_A = \mathcal{U} \cdot e^{itp_{\mathcal{R}} \cdot (a_A^\dagger a_A - b_A^\dagger b_A)} \cdot \mathcal{U}^\dagger |0\rangle_A \quad (45)$$

where

$$\mathcal{U} = \prod_i U_{ii'} \cdot \text{CCNOT}_{ii',A} \quad (46)$$

The circuit for the exponential conditional swap operator (45) for a system of two qubits is shown below.



**Acknowledgments.** I would like to thank T. Kalajdzievski, S. Lloyd, P. Reben-trost, and C. Weedbrook for interesting discussions and helpful suggestions.

## REFERENCES

- [1] R. N. Alexander, S. C. Armstrong, R. Ukai and N. C. Menicucci, Noise analysis of single-mode Gaussian operations using continuous-variable cluster states, *Phys. Rev. A*, **90** (2014), 062324.
- [2] S. Basu, [On bounding the Betti numbers and computing the Euler characteristic of semi-algebraic sets](#), *Discret. Comput. Geom.*, **22** (1999), 1–18.
- [3] S. Basu, [Different bounds on the different Betti numbers of semi-algebraic sets](#), *Discret. Comput. Geom.*, **30** (2003), 65–85.
- [4] S. Basu, [Computing the top Betti numbers of semialgebraic sets defined by quadratic inequalities in polynomial time](#), *Found. Comput. Math.*, **8** (2008), 45–80.
- [5] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe and S. Lloyd, Quantum machine learning, *Nature*, **549** (2017), 195.
- [6] S. L. Braunstein and P. van Loock, [Quantum information with continuous variables](#), *Rev. Mod. Phys.*, **77** (2005), 513–577.
- [7] G. Carlsson, A. Zomorodian, A. Collins and L. Guibas, Persistence barcodes for shapes, *Int. J. Shape Model*, **11** (2005), 149.
- [8] F. Chazal and A. Lieutier, [Stability and computation of topological invariants of solids in  \$\mathbb{R}^n\$](#) , *Discret. Comput. Geom.*, **37** (2007), 601–617.
- [9] D. Cohen-Steiner, H. Edelsbrunner and J. Harer, [Stability of persistence diagrams](#), *Discret. Comput. Geom.*, **37** (2007), 103–120.
- [10] F. De Martini, V. Giovannetti, S. Lloyd, L. Maccone, E. Nagali, L. Sansoni and F. Sciarrino, Quantum random access memory, *Phys. Rev. A*, **80** (2009), 010302.
- [11] R. Dridi and H. Alghassi, Homology Computation of Large Point Clouds using Quantum Annealing, [arXiv:1512.09328](#).
- [12] H. Edelsbrunner, D. Letscher and A. Zomorodian, [Topological Persistence and Simplification](#), *Discret. Comput. Geom.*, **28** (2002), 511–533.
- [13] P. Frosini and C. Landi, Size theory as a topological tool for computer vision, *Pattern Recognit. Image Anal.*, **9** (1999), 596.
- [14] R. Ghrist, [Barcodes: The persistent topology of data](#), *Bull. Am. Math. Soc. (N.S.)*, **45** (2008), 61–75.
- [15] V. Giovannetti, S. Lloyd and L. Maccone, [Quantum random access memory](#), *Phys. Rev. Lett.*, **100** (2008), 160501, 4pp.
- [16] V. Giovannetti, S. Lloyd and L. Maccone, Architectures for a quantum random access memory, *Phys. Rev. A*, **78** (2008), 052310.
- [17] L. K. Grover, [A fast quantum mechanical algorithm for database search](#), *Annual ACM Symposium on the Theory of Computing*, (1996), 212–219.
- [18] M. Gu, C. Weedbrook, N. C. Menicucci, T. C. Ralph and P. van Loock, Quantum computing with continuous-variable clusters, *Phys. Rev. A*, **79** (2009), 062318.
- [19] S. Harker, K. Mischaikow, M. Mrozek and V. Nanda, [Discrete Morse theoretic algorithms for computing homology of complexes and maps](#), *Found. Comput. Math.*, **14** (2014), 151–184.
- [20] H. L. Huang, X. L. Wang, P. P. Rohde, Y. H. Luo, Y. W. Zhao, C. Liu, L. Li, N. L. Liu, C. Y. Lu and J. W. Pan, Demonstration of topological data analysis on a quantum processor, *Optica*, **5** (2018), 193.
- [21] D. Kozlov, *Combinatorial Algebraic Topology*, Algorithms and Computation in Mathematics, 21. Springer, Berlin, 2008.
- [22] H. K. Lau, R. Pooser, G. Siopsis and C. Weedbrook, [Quantum machine learning over infinite dimensions](#), *Phys. Rev. Lett.*, **118** (2017), 080501, 6pp.
- [23] H. K. Lau, R. Pooser, G. Siopsis and C. Weedbrook, [Quantum machine learning over infinite dimensions](#), *Phys. Rev. Lett.*, **118** (2017), 080501, 6pp.
- [24] N. Liu, J. Thompson, C. Weedbrook, S. Lloyd, V. Vedral, M. Gu and K. Modi, [Power of one qumode for quantum computation](#), *Phys. Rev. A*, **93** (2016), 052304, 10pp.
- [25] S. Lloyd, Hybrid quantum computing, preprint, [arXiv:quant-ph/0008057](#).
- [26] S. Lloyd and S. L. Braunstein, [Quantum computation over continuous variables](#), *Phys. Rev. Lett.*, **82** (1999), 1784–1787.
- [27] S. Lloyd, S. Garnerone and P. Zanardi, Quantum algorithms for topological and geometric analysis of data, *Nat. Commun.*, **7** (2016), 10138.
- [28] P. van Loock, C. Weedbrook and M. Gu, Building Gaussian cluster states by linear optics, *Phys. Rev. A*, **76** (2007), 032321.

- [29] K. Marshall, C. S. Jacobsen, C. Schafermeier, T. Gehring, C. Weedbrook and U. L. Andersen, Continuous-variable quantum computing on encrypted data, *Nat. Comm.*, **7** (2016), 13795.
- [30] N. C. Menicucci, Fault-tolerant measurement-based quantum computing with continuous-variable cluster states, *Phys. Rev. Lett.*, **112** (2014), 120504.
- [31] N. C. Menicucci, P. van Loock, M. Gu, C. Weedbrook, T. C. Ralph and M. A. Nielsen, Universal quantum computation with continuous-variable cluster states, *Phys. Rev. Lett.*, **97** (2006), 110501.
- [32] K. Mischaikow and V. Nanda, [Morse theory for filtrations and efficient computation of persistent homology](#), *Discret. Comput. Geom.*, **50** (2013), 330–353.
- [33] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.
- [34] P. Niyogi, S. Smale and S. Weinberger, [A topological view of unsupervised learning from noisy data](#), *SIAM J. Comput.*, **40** (2011), 646–663.
- [35] M. Pysher, Y. Miwa, R. Shahrokhshahi, R. Bloomer and O. Pfister, Parallel generation of quadripartite cluster entanglement in the optical frequency comb, *Phys. Rev. Lett.*, **107** (2011), 030505.
- [36] H. Reitberger, Leopold Vietoris (1891–2002), *Notices of the American Mathematical Society*, **49** (2002), 1232–1236.
- [37] V. Robins, Towards computing homology from finite approximations, *Topol. Proc.*, **24** (1999), 503–532.
- [38] S. Takeda, T. Mizuta, M. Fuwa, J.-i. Yoshikawa, H. Yonezawa and A. Furusawa, Generation and eight-port homodyne characterization of time-bin qubits for continuous-variable quantum information processing, *Phys. Rev. A*, **87** (2013), 043803.
- [39] C. Weedbrook, S. Pirandola, R. Garcia-Patron, N. J. Cerf, T. C. Ralph, J. H. Shapiro and S. Lloyd, Gaussian quantum information, *Rev. Mod. Phys.*, **84** (2012), 621.
- [40] C. R. Wie, A Quantum Circuit to Construct All Maximal Cliques Using Grover Search Algorithm, [arXiv:1711.06146](#).
- [41] S. Yokoyama, R. Ukai, S. C. Armstrong, J.-i. Yoshikawa, P. van Loock and A. Furusawa, Demonstration of a fully tunable entangling gate for continuous-variable one-way quantum computation, *Phys. Rev. A*, **92** (2015), 032304.
- [42] S. Yokoyama, R. Ukai, S. C. Armstrong, C. Sornphiphatphong, T. Kaji, S. Suzuki, J. i. Yoshikawa, H. Yonezawa, N. C. Menicucci and A. Furusawa, Ultra-large-scale continuous-variable cluster states multiplexed in the time domain, *Nature Photonics*, **7** (2013), 982.
- [43] J. i. Yoshikawa, S. Yokoyama, T. Kaji, C. Sornphiphatphong, Y. Shiozawa, K. Makino and A. Furusawa, Generation of one-million-mode continuous-variable cluster state by unlimited time-domain multiplexing, *Applied Phys. Lett. Photonics*, **1** (2016), 060801.
- [44] J. Zhang and S. L. Braunstein, Continuous-variable Gaussian analog of cluster states, *Phys. Rev. A*, **73** (2006), 032318.
- [45] A. Zomorodian and G. Carlsson, [Computing persistent homology](#), *Discret. Comput. Geom.*, **33** (2005), 249–274.
- [46] A. Zomorodian, *Algorithms and Theory of Computation Handbook*, 2nd edition, Ch. 3, section 2. Chapman and Hall/CRC, 2009.

*E-mail address:* [siopsis@tennessee.edu](mailto:siopsis@tennessee.edu)