Higher-order topological kernels via quantum computation

Massimiliano Incudini
University of Verona

November 22, 2023

7th conference on Quantum Techniques in Machine Learning (QTML 2023)

Two different flavours of quantum kernels

On NISQ devices
Search for useful applications (e.g. Belis' talk 15:30 today)

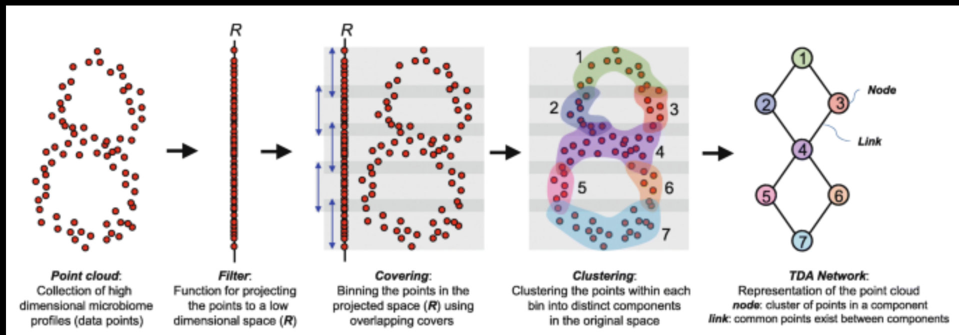
On fault-tolerant devices
Search for a provable speedup on artificial tasks (e.g. Liu et al 2021)

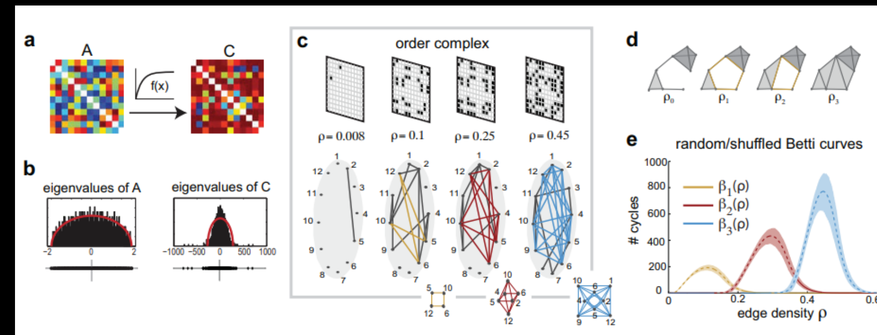
Motivation

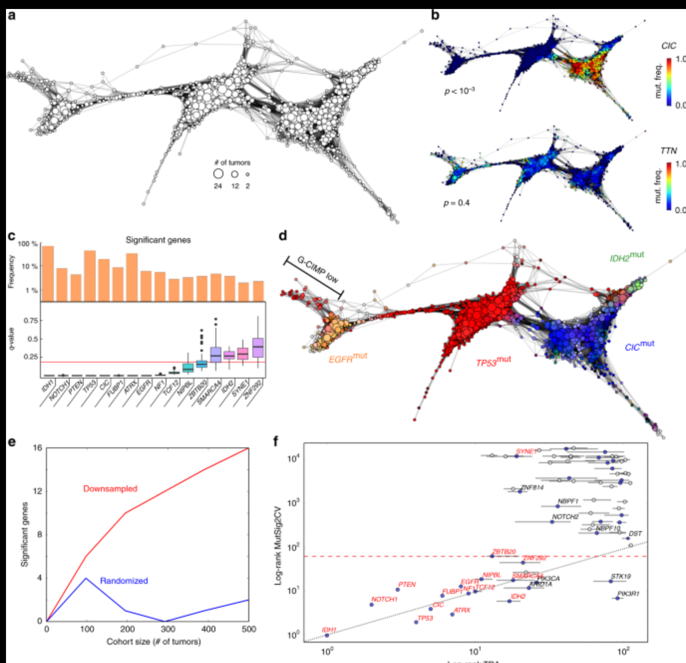
Can we kernelize any quantum algorithm
that has both theoretical guarantees
and practical applications in data analysis?

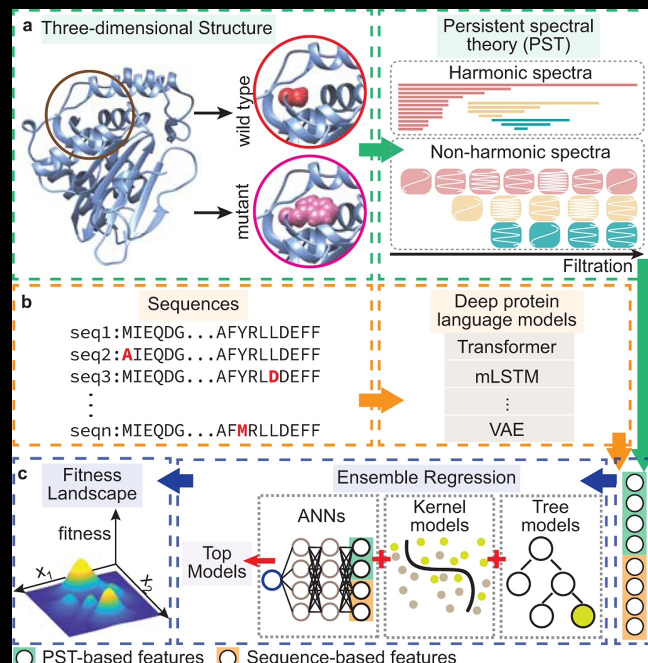
Yes! One choice is *topological data analysis*.

Motivation







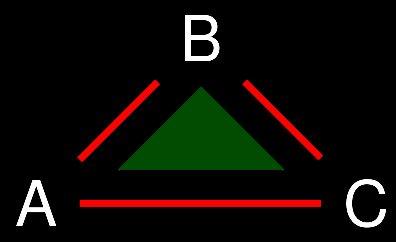


(1) Liao et al. Genome Biol 2019.
(2) Giusti et al. PNAS 2015.
(3) Rabadán et al. Nat Comm 2020.
(4) Qiu et al. Nat Comput Sci 2023.

Topological data analysis

Algebraic topology

Abstract simplicial complex
$=$ collection of subsets of S
closed under $\subset$



$$\Sigma = \{\{A\}, \{B\}, \{C\},$$
$$\{A, B\}, \{A, C\}, \{B, C\},$$
$$\{A, B, C\}\}$$
$$\Sigma = \Sigma_0 \cup \Sigma_1 \cup \Sigma_2 \cup \dots$$

k-th chain group
$$\mathcal{C}_k = \mathrm{span}\{\Sigma_k\}$$

boundary map $\partial_k : \mathcal{C}_k \to \mathcal{C}_{k-1}$,
$$\partial_k \sigma = \sum_{j=0}^{k} (-1)^j (\sigma \setminus \{\sigma_j\})$$
Note that $\partial_k \partial_{k+1} = 0$

k-th cycle group
$$Z_k = \mathrm{span}\{\sigma \in \Sigma_k \mid \partial\sigma = 0\} \subset \mathcal{C}_k$$

k-th boundary group
$$B_k = \mathrm{span}\{\sigma \in \Sigma_k \mid \partial\tau = \sigma, \tau \in \mathcal{C}_{k+1}\} \subset Z_k$$

k-th Betti numbers $\beta_k = \dim\ker(Z_k) - \dim\ker(B_k)$

Algorithms to estimate the Betti numbers

We can study the topological properties of Σ via the Combinatorial Laplacian operator, Δ_k,

$$\Delta_k = \partial_k^\dagger \partial_k + \partial_{k+1} \partial_{k+1}^\dagger$$

for which

$$\dim \ker \Delta_k = \beta_k .$$

Algorithms to estimate the Betti numbers

Hard to have the explicit representation of Δ_k on a classical computer, but easy on a quantum one (supposing efficient sampling of k-cliques).

We can estimate an ϵ-additive approximation for the normalized Betti number, $\beta_k/|\Sigma_k|$,

$$\text{Randomly sampled eigenvector} \longrightarrow \Delta_k \longrightarrow \text{Randomly sampled eigenvalue}$$

and estimate the % of zero eigenvalues.

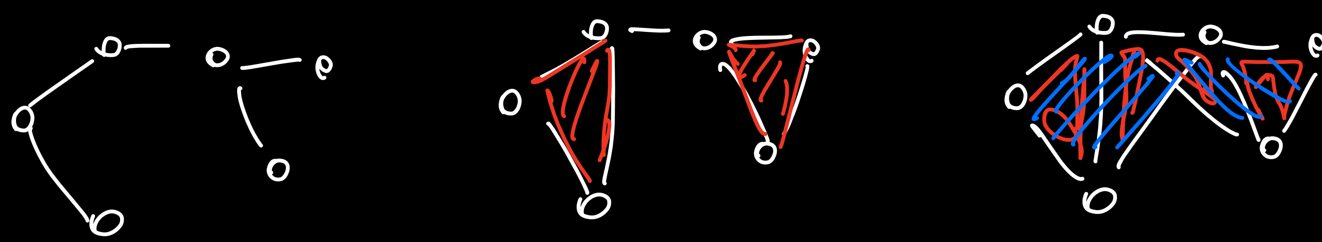
Algorithms to estimate the Betti numbers (more)

1. Encoding
   - $\sigma \in \Sigma$ over n vertices
   - $|\sigma\rangle = |b_1 \dots b_n\rangle$
   - $b_j = 1 \iff v_j \in \sigma$
2. Uniform distribution of eigenvectors
   - $\rho = \sum_{\sigma \in \Sigma_k} |\sigma\rangle\langle\sigma|$
   - you are able to sample k-cliques efficiently
3. Efficient construction of Δ_k
   - $\Delta = \sum_j Z^{\otimes(j-1)} \otimes X \otimes I^{(n-j)}$ for complete abs
   - Project onto the subspace of Σ's simplices
4. Eigenvalues estimation
   - no lower bound on the spectral gap

- Lloyd et al. Nat. Comm. 7(1), 2016.
- Ubaru et al. arXiv:2108.02811, 2021.
- Gyurik et al. Quantum 6, 2022.
- Hayakawa. Quantum 6, 2022.
- McArdle et al. arXiv:2209.12887, 2022.
- Apers et al. arXiv:2211.09618, 2023.
- Berry et al. arXiv:2209.13581, 2023.
- ...

Kernelize BNE

Intuition



κ

β_κ

$$h(\cdot) = \sum_j \alpha_j \, \kappa(\cdot, x^j)$$

TRAINING SET ELEMENTS

Multidimensional Betti curves

Let $\mathcal{F} = (\Sigma_1, \ldots, \Sigma_q), \Sigma_i \subset \Sigma_j$ for $i < j$ be a filtration.

The Betti curve $b \in \mathbb{R}^q$ is defined by

$$[b]_j = \beta_1(\Sigma_j).$$

Classically, only low-degree topological features are estimated due to the computational cost.

Does not use persistent features
(e.g. holes surviving from one abs to another of the filtration)

This approach is still relevant in the literature, even for non-machine learning tasks (e.g. Giusti el al. PNAS 2015).

Multidimensional Betti curves

We can extend the concept of the Betti curve, and adapt to the normalized value we can efficiently retrieve, on many orders of Betti numbers. This leads to $B \in \mathbb{R}^{q \times m}$,

$$[B]_{j,k} = \frac{\beta_k(\Sigma_j)}{|\Sigma_j|}.$$

The mapping $\phi : \mathrm{Filtration} \to \mathbb{R}^{q \times m}$ is a *feature map*.

Multidimensional Betti curves

$$\kappa(\mathcal{F}_1, \mathcal{F}_2) = \langle \phi(\mathcal{F}_1), \phi(\mathcal{F}_2) \rangle_F = \mathrm{Tr}[\phi(\mathcal{F}_1)^\dagger \phi(\mathcal{F}_2)] \tag{1}$$

$$\kappa(\mathcal{F}_1, \mathcal{F}_2) = \exp\Big(-\gamma \cdot \|\phi(\mathcal{F}_1) - \phi(\mathcal{F}_2)\|_F \Big) \tag{2}$$

Note that:

- Gaussian smoothing (eq. 2) leads to smoother prediction functions;
- Johnson and Jung (2021) derived single-dimensional Betti curves from persistent diagrams and showed that the formulation is stable with respect to the 1-Wasserstein metric on persistent diagrams (robust to small perturbation of the persistent diagrams).
  We are not arguing about stability in our approach.

Multidimensional Betti curves

Difference with classical approaches:

- Classical approaches are usually based on persistent features
  - Non-persistent features are still relevant
  - Can immediately exploit the majority of BNE quantum algorithms
  - Quantum algorithms for persistent BN but require even more resources
- We are not necessarily estimating Betti numbers (zero eigenvalues) but the low-lying portion of the spectrum of Δ_k (zero or close to zero)
  - Do we know the spectral gap?
  - Are we okay with estimating the low-lying portion of the spectrum?
- We are using *normalized* Betti numbers
  - $\beta_k/|\Sigma_k| = \beta'_k/|\Sigma'_k| \not\Rightarrow \beta_k = \beta'_k$

Open points

- Straightforward extension of quantum algorithms for Betti number estimation
  - can we have a more general approach?
- Can we characterize the expressibility of the multidimensional normalized Betti curve kernel?
  - Exponential concentration of kernel values? (Thanasilp et al)
- Which are the “nearest” term applications?

(Appendices, or catch up later!)

Take-away messages

Take-away messages

1. The kernelization of quantum TDA algorithms can lead to both theoretical guarantees and practical applications.
2. Such kernels can differ from topological kernels in the classical domain in many aspects
3. They require fault-tolerant quantum hardware and many resources, few simple use cases might come sooner than the most general and interesting data analysis use case.

Thank you!

Higher-order topological kernels via quantum computation.

M. Incudini, F. Martini and A. Di Pierro.

To appear at the Proceeding of the 2023 IEEE International Conference on Quantum Computing and Engineering (QCE).

Appendices $\rightarrow$

Appendix 1: open points

Spectral similarity

Let n be the vertices in the abs, and m be the precision of QPE. After QPE we end up with the state:

$$|\psi\rangle = \sum_j \alpha_j \left|\lambda_j\right\rangle_m \left|v_j\right\rangle_n$$

For BNE we estimate α_0 (corr. $\lambda_0 = 0$) and we embed this piece information of information into a kernel function.

We can generalize the approach to compare portions of the spectrum of two operators:

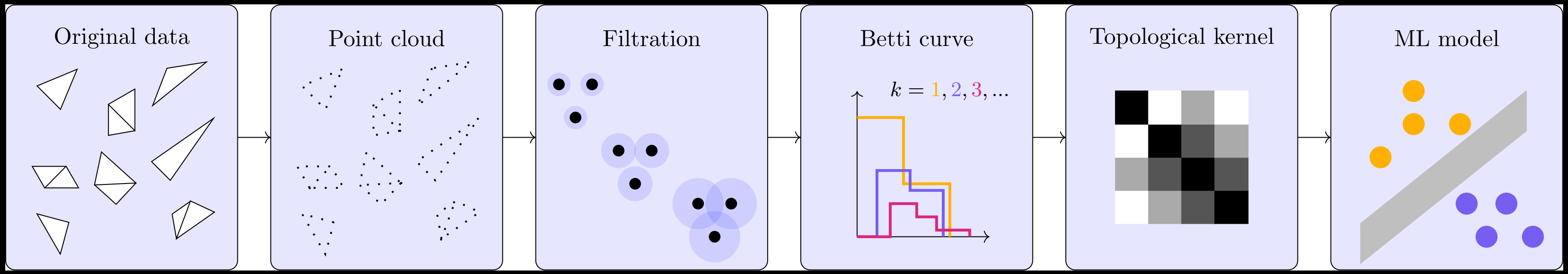
$$\left|\lambda_j'\right\rangle = (I \otimes \langle 0|) \left|\lambda_j\right\rangle$$

$$\left|\nu_j'\right\rangle = (I \otimes \langle 0|) \left|\nu_j\right\rangle$$

$$\kappa = \left\langle \nu_j' \middle| \lambda_j' \right\rangle$$

Appendix 2: nearest-term toy tasks

Toy problem



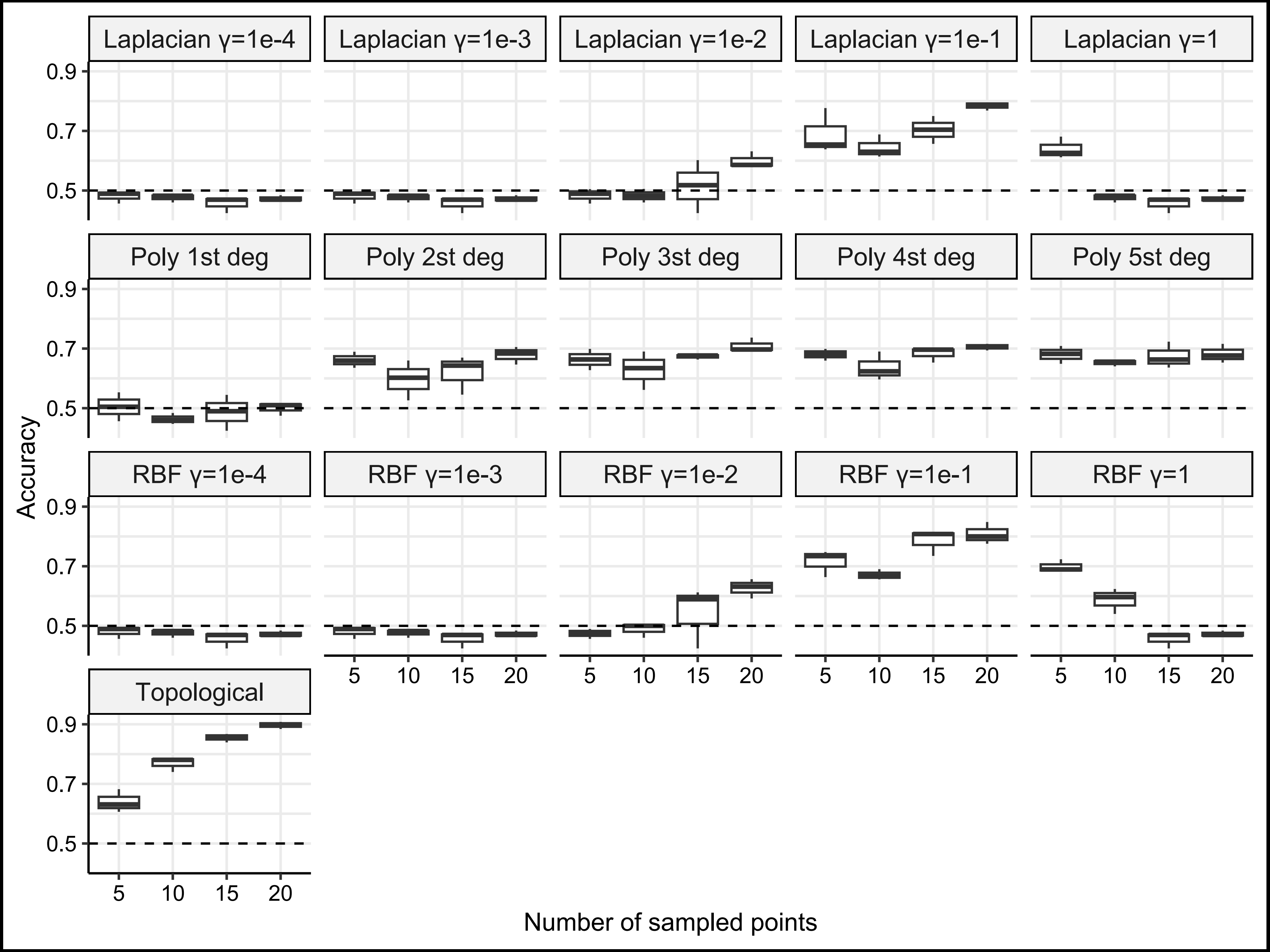
1. Distinguish triangles from squares cut in half (different topological features);
2. Create deformed figures, samples points on the border, construct a Vietoris-Rips filtration;
3. Estimate the multidimensional Betti curves for the filtrations;
4. Calculate the kernel matrix and fed it to the kernel machine.

Toy problem

Important points:

- Δ_k has been described as the sum of Pauli strings, which has polynomially many terms only for certain classes of ABS
- Allows to use Trotter or qDrift for Hamiltonian Simulation, minimizing the number of qubits required.

Some preliminary results



Some preliminary results