



PASQAL

Extending Graph Transformers with Quantum Computed Aggregation

QTML 2023 21/11/2023

Slimane Thabet
slimane.thabet@pasqal.com

arXiv:2210.10610
arXiv:2310.20519



FRESNEL

Team



Mehdi Djellabi



Loïc Henriët



Sachin Kasture



Romain Fouilland



Louis-Paul Henry



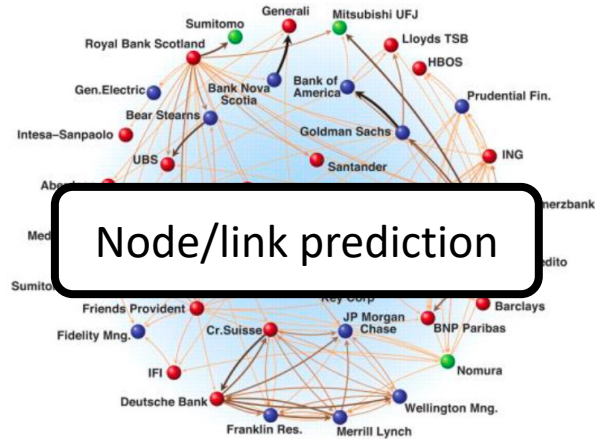
Slimane Thabet



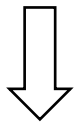
Igor Sokolov

Graph-structured data

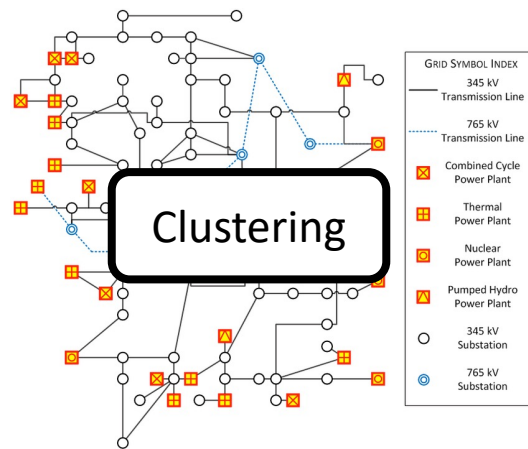
Machine learning tasks on graph-structured data



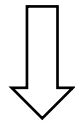
Economic networks



Anticipating new partnerships



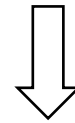
Power grid



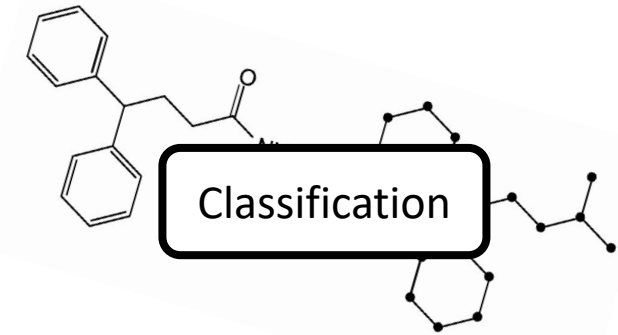
Segmenting highly-connected zones



Social networks



Marketing to like-minded users



Molecules



Predicting toxicity of compounds

Graph Machine Learning with Quantum Computers

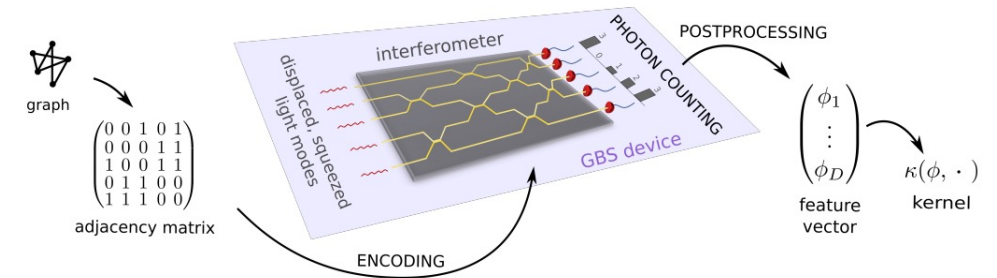
Theoretical protocols

- Verdon, G., McCourt, T., Luzhnica, E., Singh, V., Leichenauer, S., & Hidary, J. (2019). Quantum graph neural networks. *arXiv preprint arXiv:1909.12264*.
- Schuld, M., Brádler, K., Israel, R., Su, D., & Gupt, B. (2020). Measuring the similarity of graphs with a Gaussian boson sampler. *Physical Review A*, 101(3), 032314.
- Henry, L. P., Thabet, S., Dalyac, C., & Henriët, L. (2021). Quantum evolution kernel: Machine learning on graphs with programmable arrays of qubits. *Physical Review A*, 104(3), 032416.

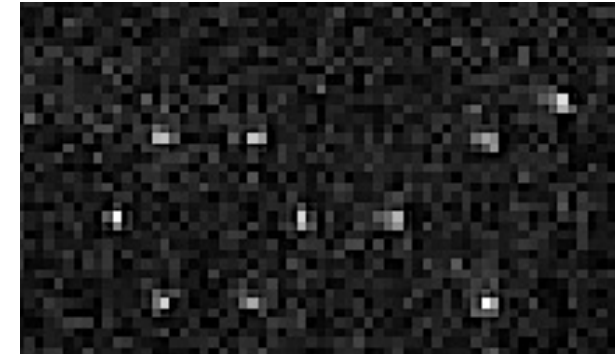
Experimental implementation on neutral atoms platforms

- Albrecht, B., Dalyac, C., Leclerc, L., Ortiz-Gutiérrez, L., Thabet, S. et al. (2023). Quantum feature maps for graph machine learning on a neutral atom quantum processor. *Physical Review A*, 107(4), 042615.

These protocols work for graph level tasks (classification or regression)



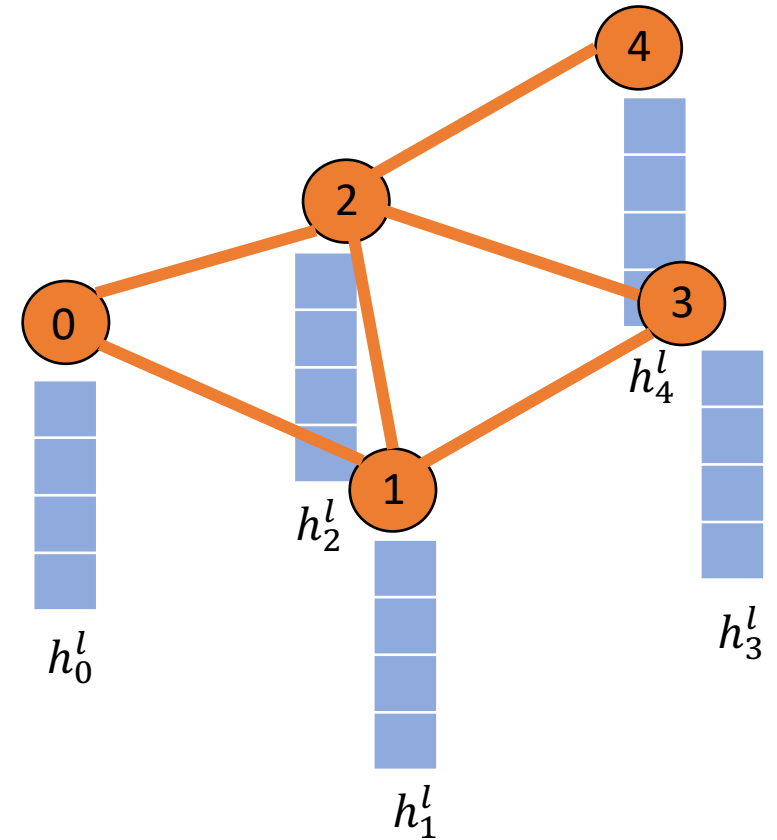
Schuld et al.



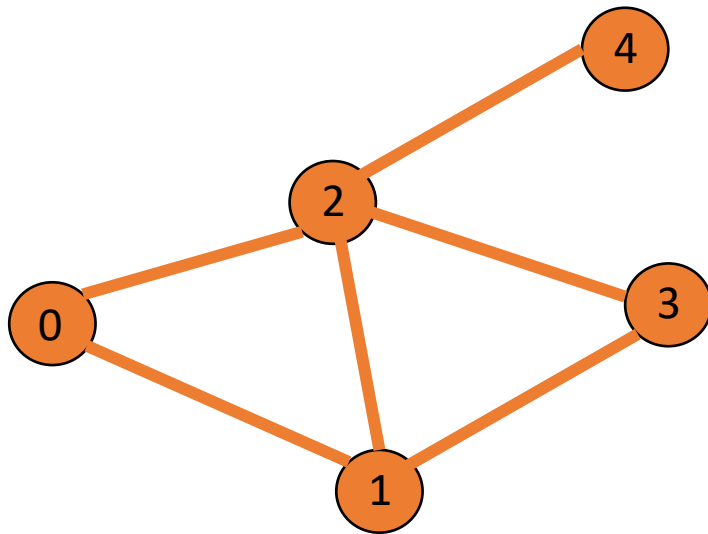
Albrecht et al.

Graph Neural Networks (classical)

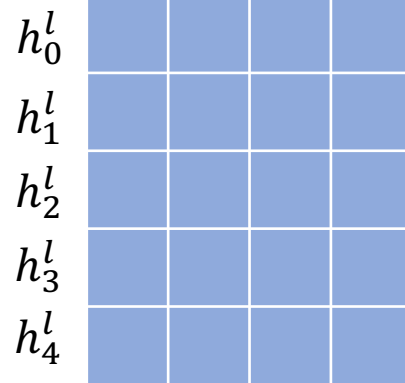
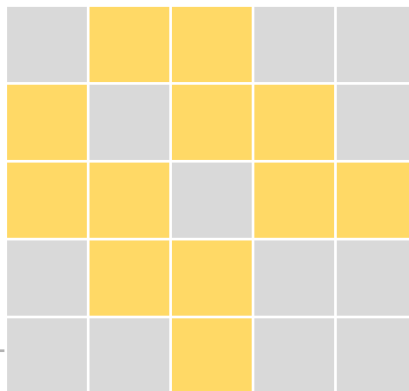
- Objective: learn representation of nodes (or edges) in a graph, with the use of neural networks.
- Each node is associated at the end with a vector of dimension d .
- These vectors can be the input of a machine learning algorithm. Can be used on many tasks.
- Can be done on one big graph, or several small graphs.



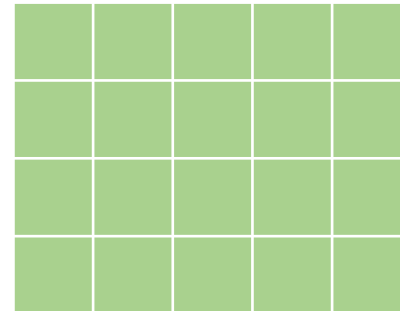
Graph Convolution Networks layer



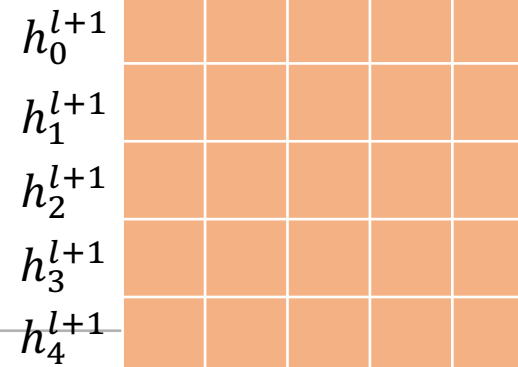
A adjacency matrix



H node features

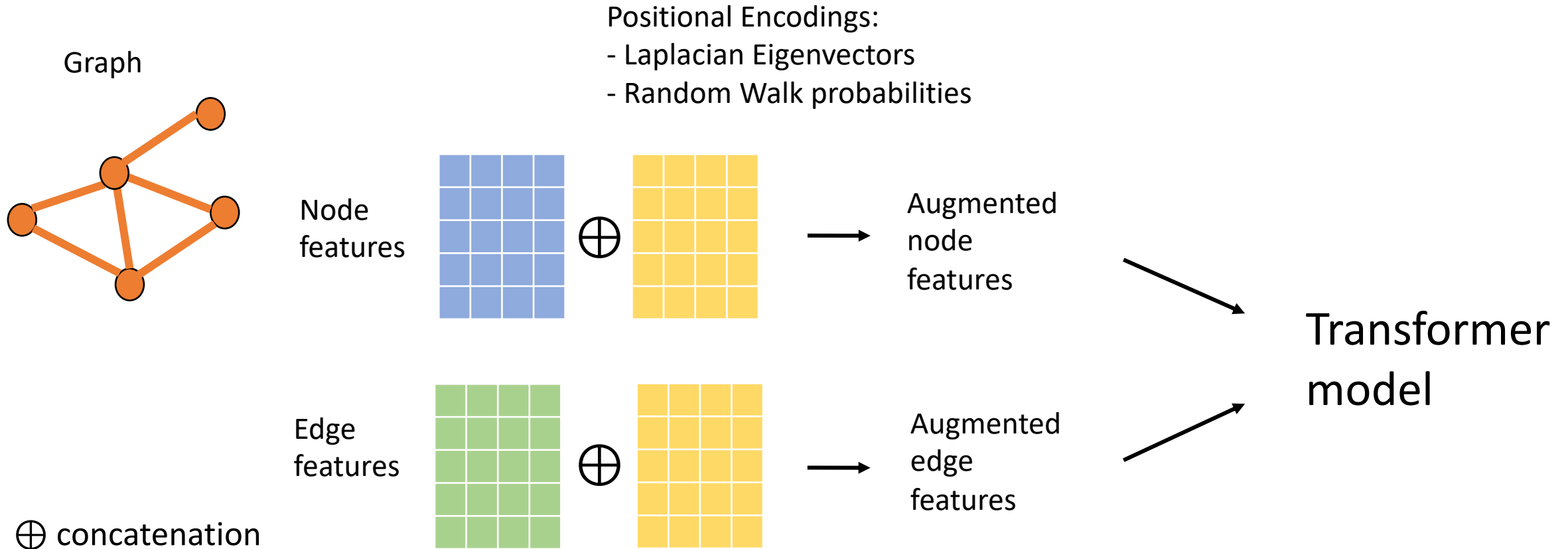


W weight matrix



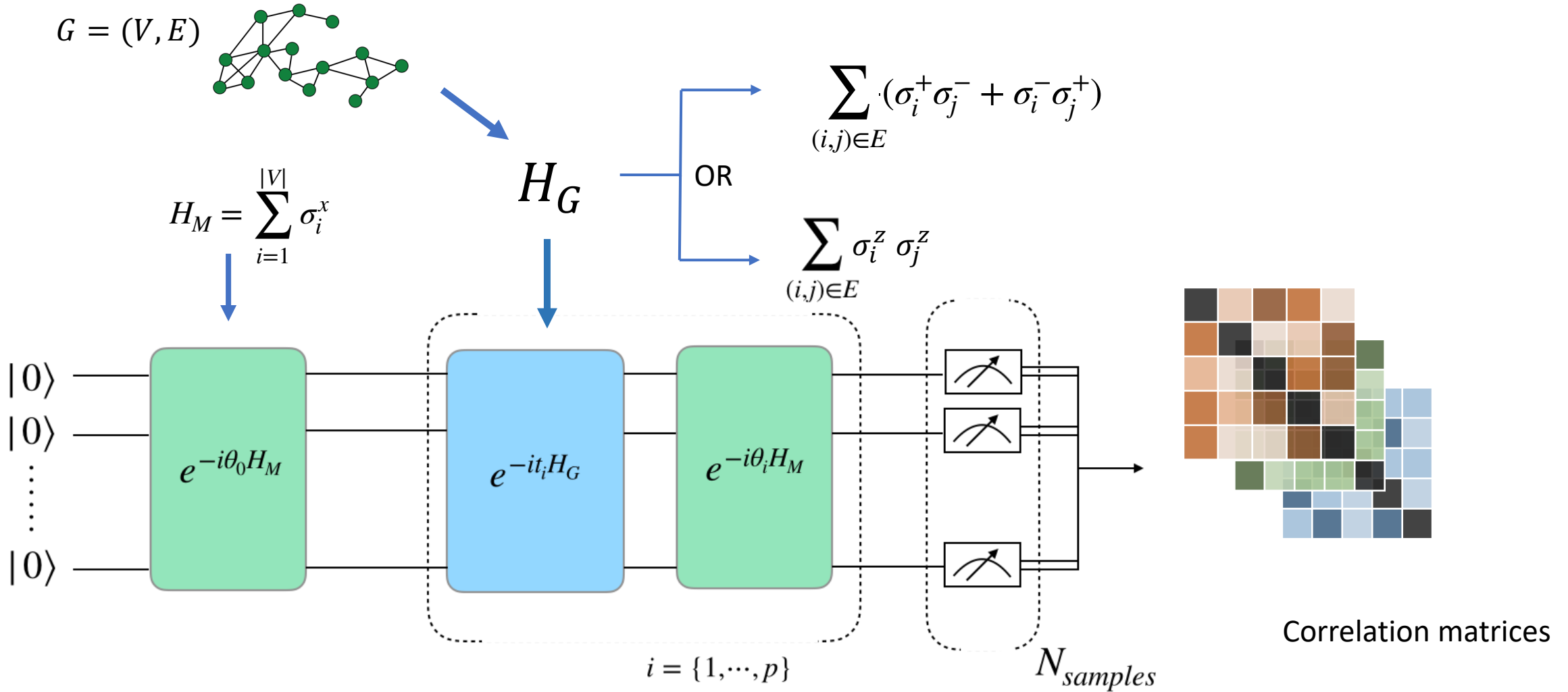
$$= \sigma(AHW)$$

Graph Transformers

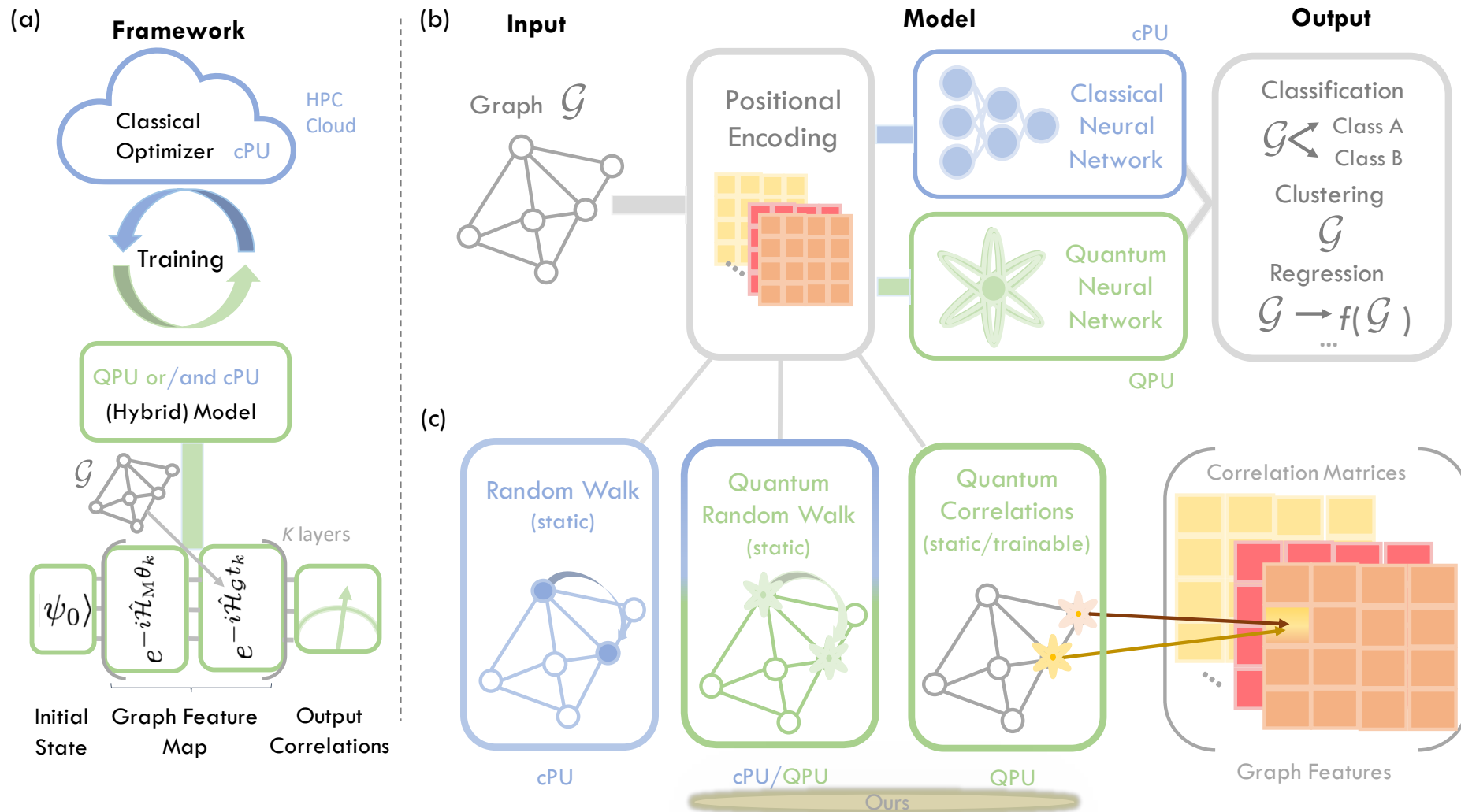


- Kreuzer, D., Beaini, D., Hamilton, W., Létourneau, V., & Tossou, P. (2021). Rethinking graph transformers with spectral attention. *NeurIPS 2021*
- Ying, C., Cai, T., Luo, S., Zheng, S., Ke, G., He, D. et al. Do transformers really perform bad for graph representation? *arXiv 2021. arXiv:2106.05234.*
- Rampášek, L., Galkin, M., Dwivedi, et al. (2022). Recipe for a general, powerful, scalable graph transformer. *NeurIPS 2022*
- Ma, L., Lin, C., Lim, D., Romero-Soriano, A. et al. (2023). Graph Inductive Biases in Transformers without Message Passing. *arXiv:2305.17589.*

Quantum correlations as graph features



Quantum position encodings and aggregation



Properties of our method

- Learn a representation that can be reused on many tasks, graph level, node level etc
- Naturally equivariant by permuting the labeling of the nodes
- One can compute intractable features e.g correlation on ground state
- The quantum features can be adapted with new classical algorithms

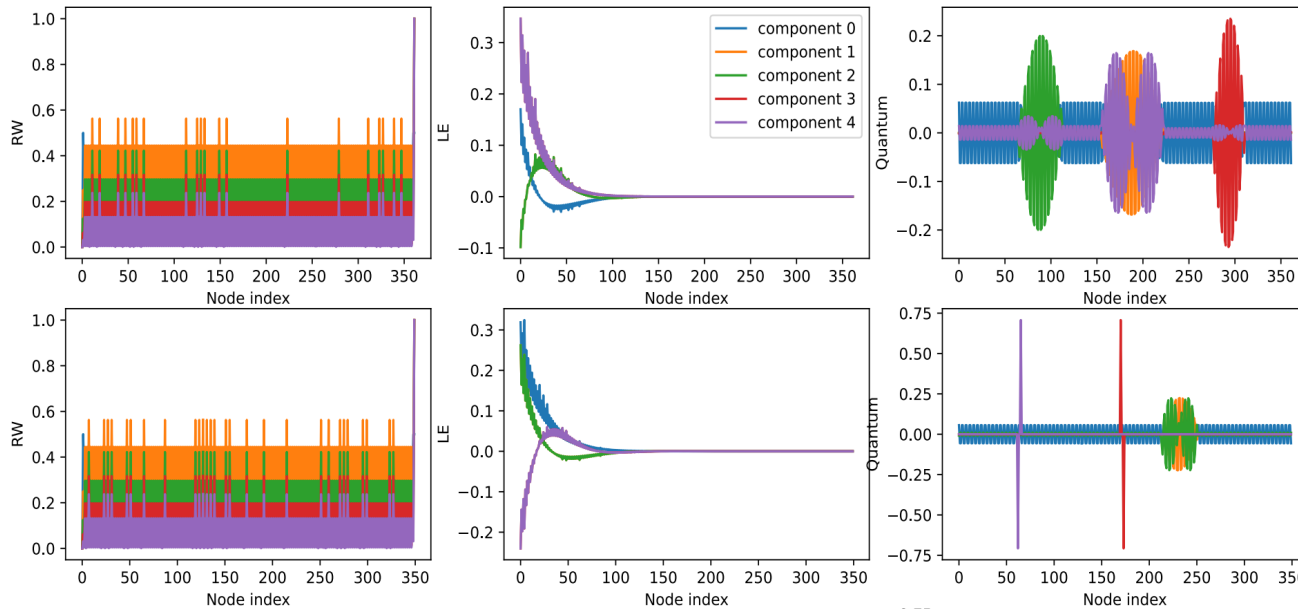
Open questions:

- Are there learning tasks that for which quantum features bring a significant advantage ?
- Is this advantage worth the cost of running a quantum computer ?

Experiments

Theorem 1 *GD-WL test with RRWP embedding fails to distinguish non isomorphic strongly regular graphs. 2 particle-QRW can distinguish some of them.*

Class 0



Class 1

Features on artificial datasets

| Model | ZINC-full (MAE ↓) | PCQM4Mv2 (MAE ↓) |
|----------------------------|----------------------|---------------------|
| GIN | 0.088 ± 0.002 | 0.1195 |
| GraphSAGE | 0.126 ± 0.003 | — |
| GAT | 0.111 ± 0.002 | — |
| GCN | 0.113 ± 0.002 | 0.1195 |
| SignNet | 0.024 ± 0.003 | — |
| Graphormer | 0.052 ± 0.005 | 0.0864 |
| Graphormer-URPE | 0.028 ± 0.002 | — |
| Graphormer-GD | 0.025 ± 0.004 | — |
| GPS-medium | — | 0.0858 |
| GRIT (our run) | 0.025 ± 0.002 | 0.0842 |
| GRIT 1-CQRW (ours) | 0.025 ± 0.003 | 0.0947* |
| GRIT 2-QiQRW (ours) | 0.023 ± 0.002 | 0.0838 |

Benchmarks on real datasets

Thanks for your attention
