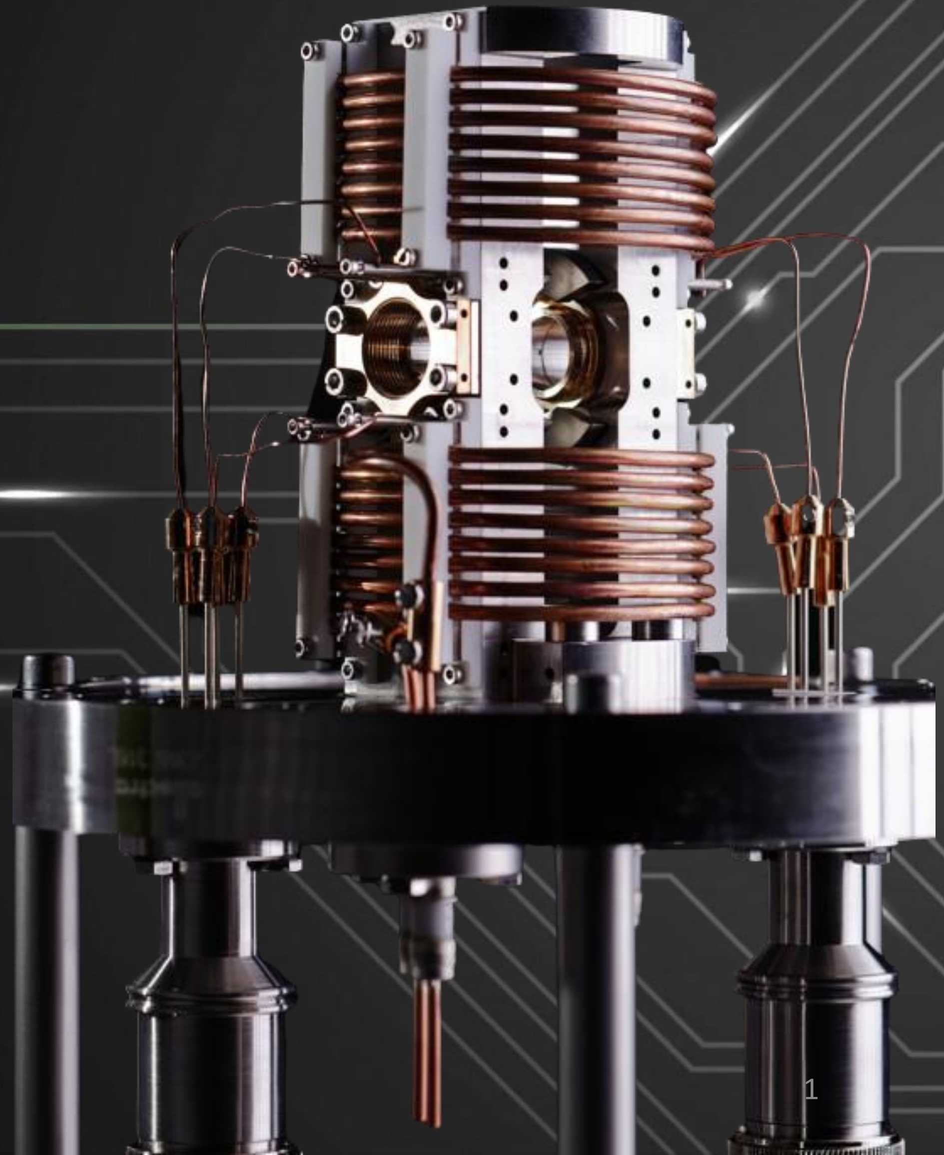
Graph machine learning using Pasqal's neutral atom quantum computer

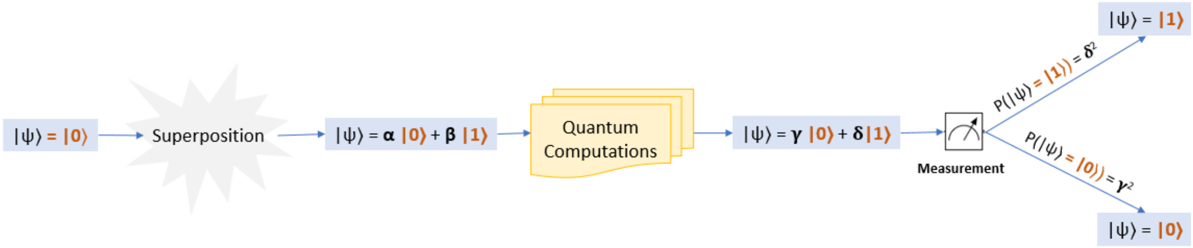


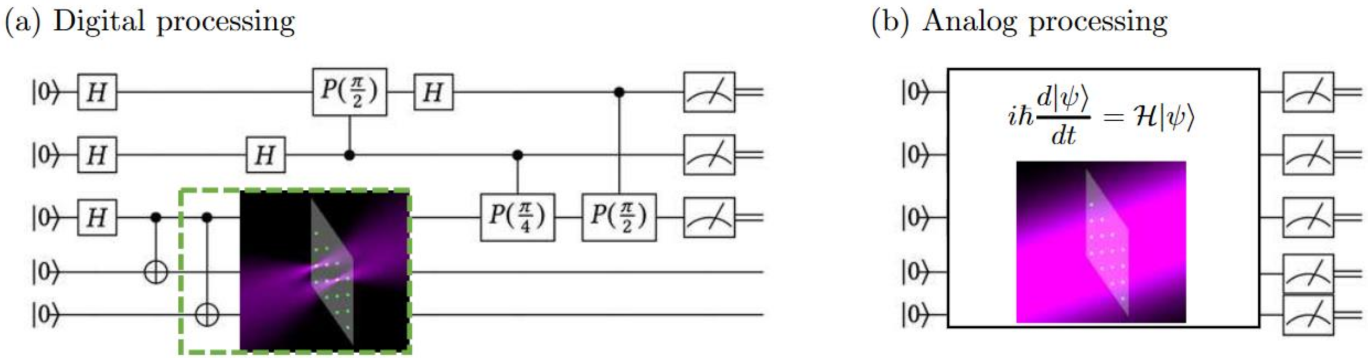
Victor Drouin-Touchette, ETS QML class, 21th March

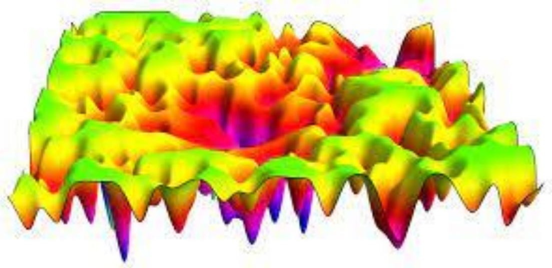
Table of Content

Re: Analog Algorithms

Analog quantum computing approaches



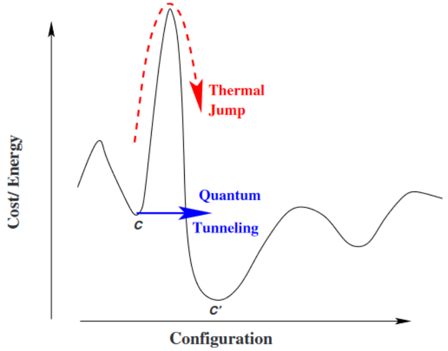


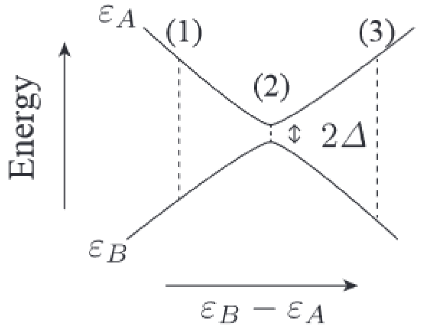


Operations applied on **entire** qubit register, let state time evolve.

1. Adiabatic quantum computing/Quantum annealing
2. Quantum Approximate Optimization Algorithm
3. Others

$$|\psi_0> \xrightarrow{\theta(t)...} |\psi(t)>$$

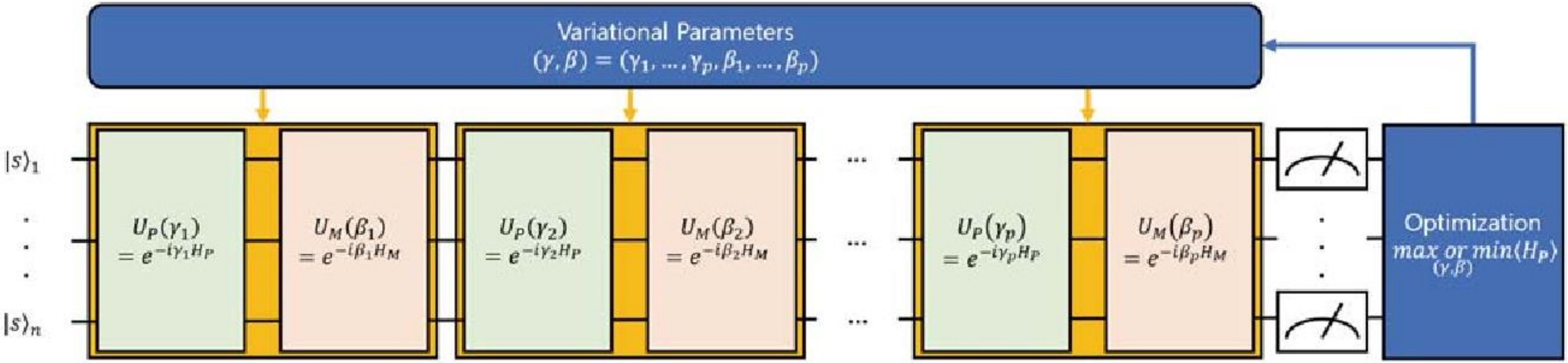




PASQAL

Quantum Approximate Optimization Algorithm

Maybe you don’t need *the* ground state, but rather a state with significant overlap with it.
Act a series of **register-wide** operations (i.e. pulses) and optimize those pulses with a set cost function.



Has been implemented on digital and analog platform.
As depth increases, results are more precise, but that needs increased qubit lifetimes.

The Quantum Evolution Kernel is inspired by QAOA and specifically applied to graph machine learning problems.

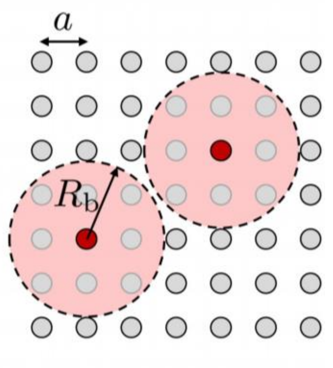
Re: Embedding graphs

Neutral atoms and Unit-Disk graphs

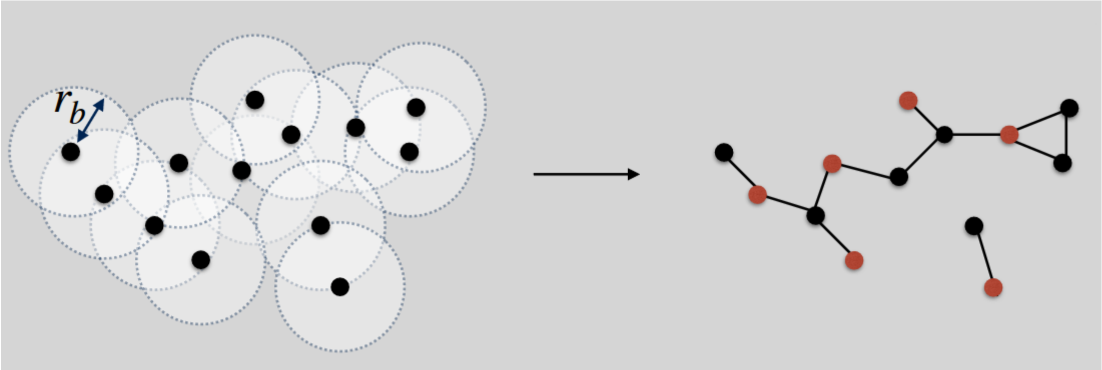
$$\mathcal{H}(t) = \frac{\hbar}{2}\Omega(t)\sum_j \sigma_j^x - \hbar\delta(t)\sum_j n_j + \sum_{i\neq j}\frac{C_6}{r_{ij}^6}n_i n_j,$$

$$R_b = \left(\frac{C_6}{\hbar\Omega}\right)^{1/6}$$

Van-der-Walls interaction is very strong at small distance ($r < R_b$) but decays fast after. Essentially a hard-core repulsion, where states $|01>$ and $|10>$ are favored.

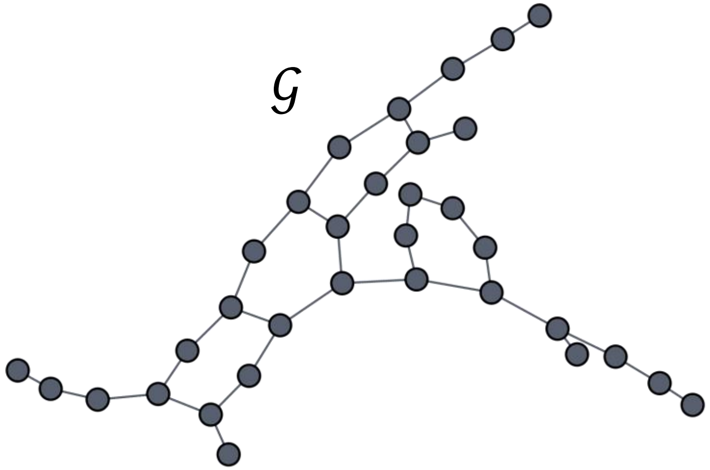


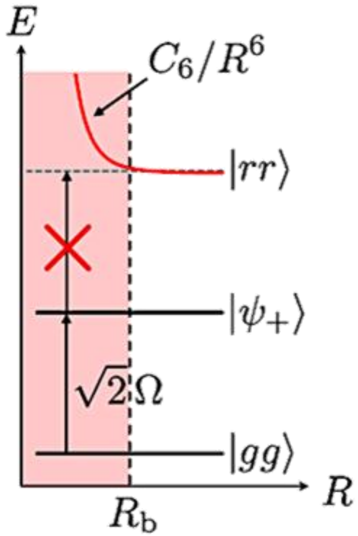
When atoms are placed in real space, we consider that an edge exists between a pair if $R_{ij} < R_b$, i.e. they are blockaded with one another.



PASQAL

1 node → 1 atom , edges → interactions





Dipole-dipole

$|r_S\rangle \leftrightarrow |r_S\rangle$

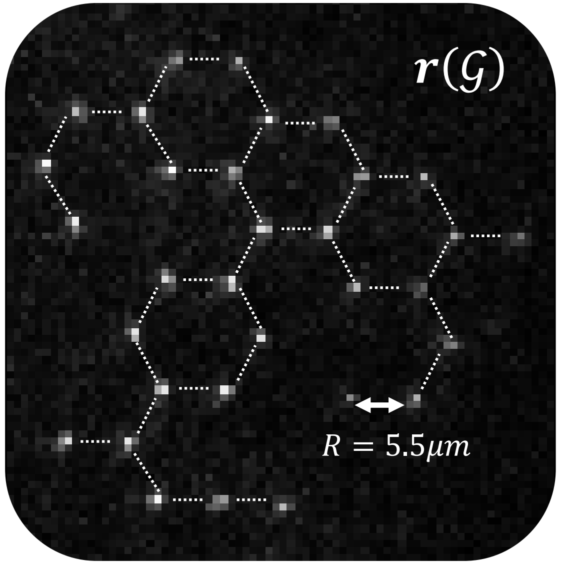
↓

Van der Waals interactions

$\propto 1/R^6$

↓

NN Blockade effect



Graph topology

$$H_{\mathcal{G}} = \sum_{(i,j)\,\in\, E(\mathcal{G})} n_i n_j$$

Interaction Hamiltonian

$$H_{dd}(\boldsymbol{r}(\mathcal{G})) \propto \sum_{i>j} \left(\frac{R}{R_{ij}}\right)^6 n_i n_j$$

Nearest Neighbour

$\Longrightarrow$

approximation

✓

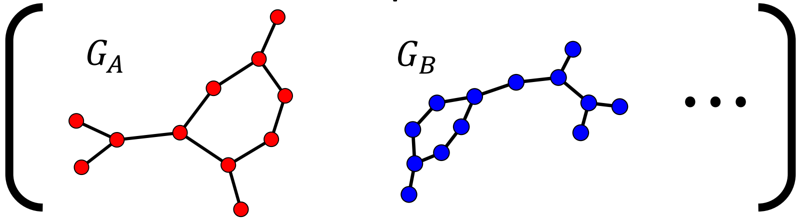
$$\boxed{H_{dd}(\boldsymbol{r}(\mathcal{G})) \propto H_{\mathcal{G}}}$$

PASQAL

Embedding: Batching atomic registers

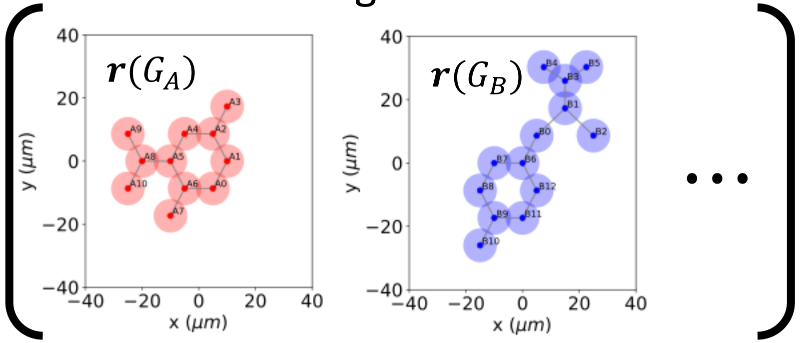
Changing trap layout is resource consuming → Do several registers with one layout

Graphs

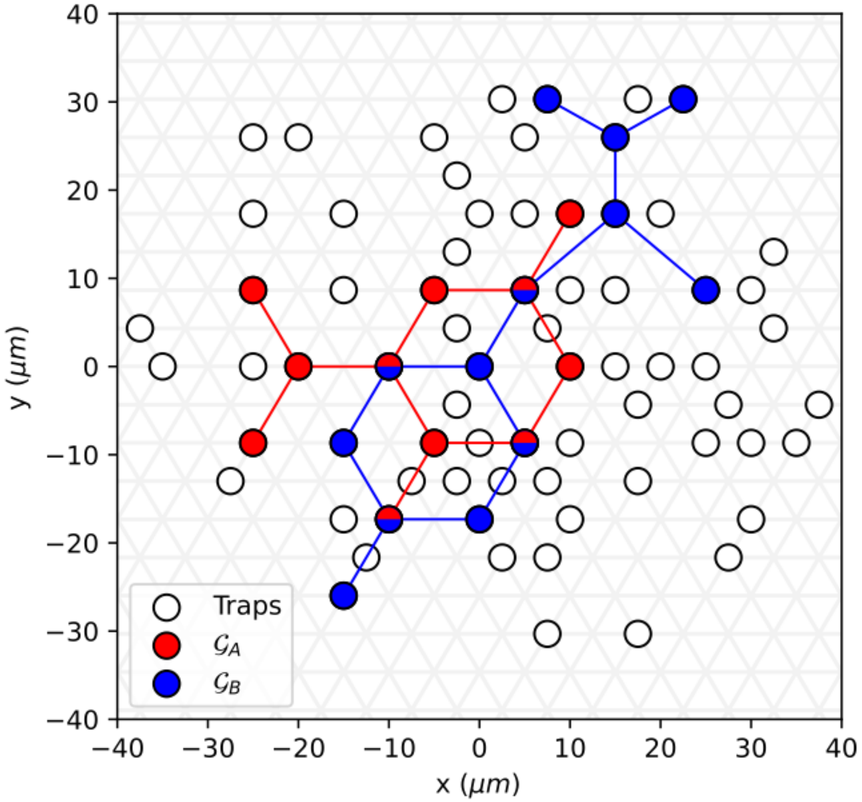


Mapping
to underlying
triangular pattern

Registers



Batching and keeping
only traps used at least once



PASQAL

Graph Machine Learning: QEK

Henry, Louis-Paul, Slimane Thabet, Constantin Dalyac, and Loïc Henriet. "Quantum evolution kernel: Machine learning on graphs with programmable arrays of qubits." *Physical Review A* 104, no. 3 (2021): 032416.

Machine Learning: Classification

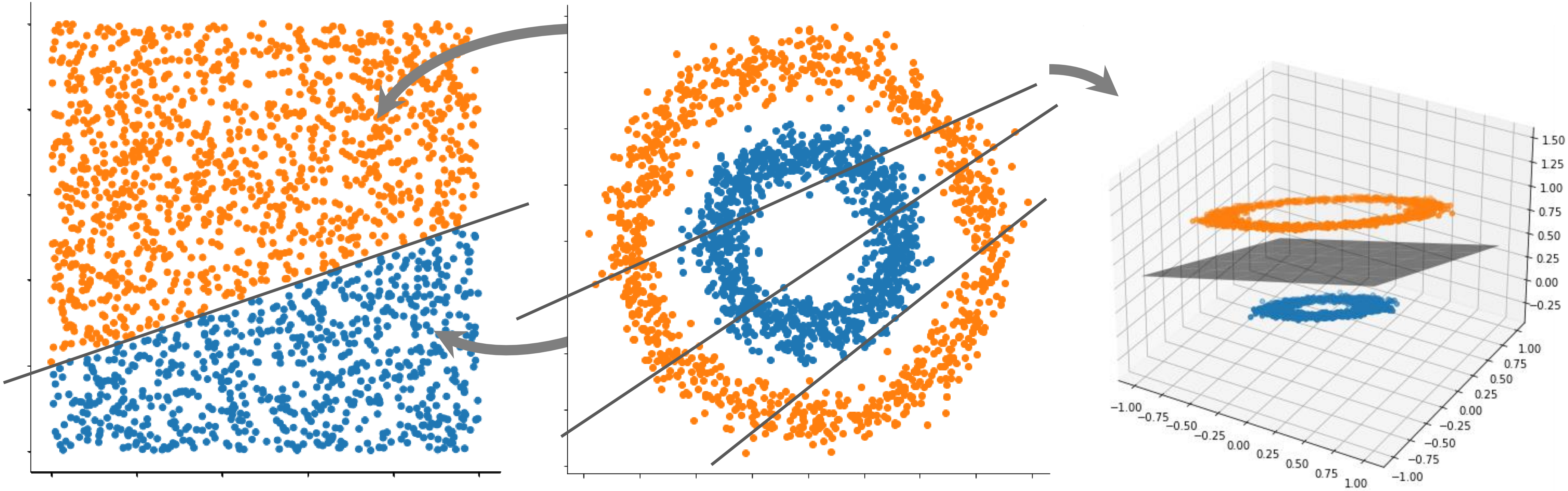
Given a data set

Split into two classes

Find best splitting hyperplane

What about non-linear data?

Embed in « feature space »

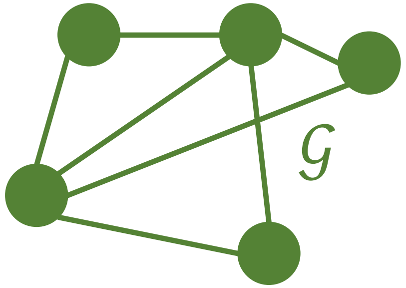


Schlköpf and Smola, *Learning with Kernels* (2001)

Machine learning with graphs: kernels

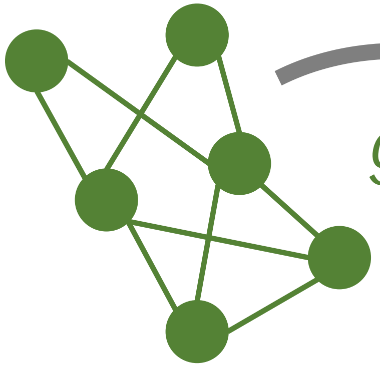
Feature vector

$$f_{\mathcal{G}} = [f_1, \ldots, f_d]$$



Kernel

$$K(\mathcal{G}, \mathcal{G}') = \langle f_{\mathcal{G}}, f_{\mathcal{G}'} \rangle$$

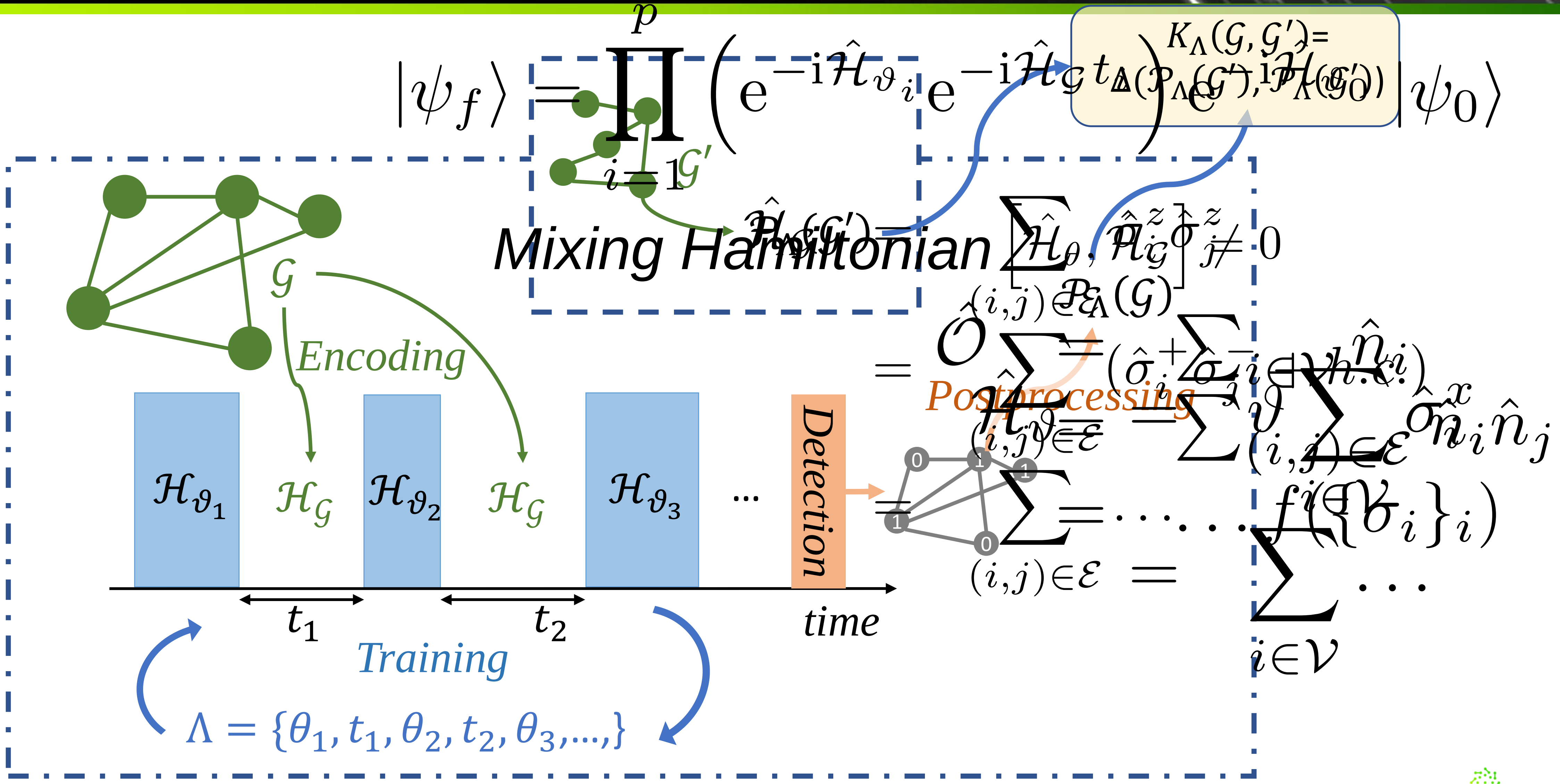


$$f_{\mathcal{G}'} = [f'_1, \ldots, f'_d]$$

Example of feature vectors/kernels:
graph subsampling kernel,
random walk kernel…

Kriege *et al.,* Applied NetworkScience, **5** (2020)

Quantum Evolution Kernel: Principle



PASQAL

Probability distributions

Varying sequence $\Lambda = \{\theta_1, t_1, \theta_2, t_2, \theta_3, \ldots,\}$

$$O_\Lambda = \langle \psi_f(\Lambda) | \, \hat{O} \, | \psi_f(\Lambda) \rangle$$

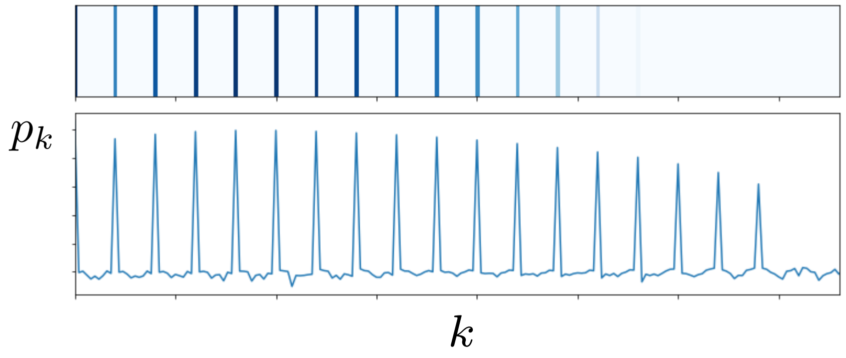
Repeat for different values of Λ

$$\{O_\Lambda\}_\Lambda$$

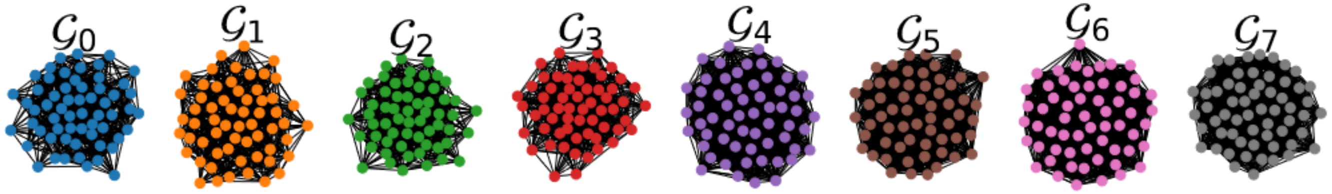
For example :
Different pulse total durations

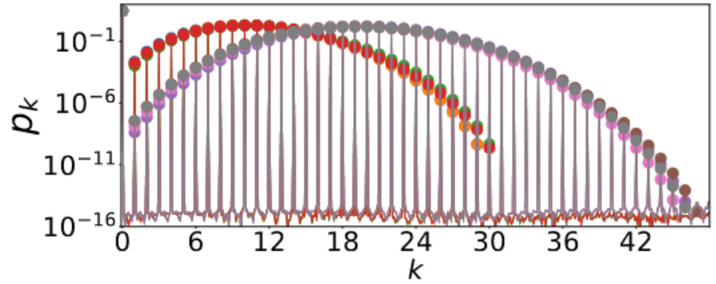
$$\Lambda \equiv t \longrightarrow p_k = \frac{1}{T} \left| \int_0^T dt \, \mathrm{e}^{-2\mathrm{i}\pi k t/T} \bar{o}(t) \right|$$

Power spectrum for graph G



Example:





PASQAL

Jensen-Shannon Divergence and the Kernel

For two probability distributions $\mathcal{P}$ and $\mathcal{P}'$, the Jensen-Shannon divergence is

$$\mathrm{JS}(\mathcal{P},\mathcal{P}') = H\left(\frac{\mathcal{P}+\mathcal{P}'}{2}\right) - \frac{H(\mathcal{P})}{2} - \frac{H(\mathcal{P}')}{2}$$

Properties:

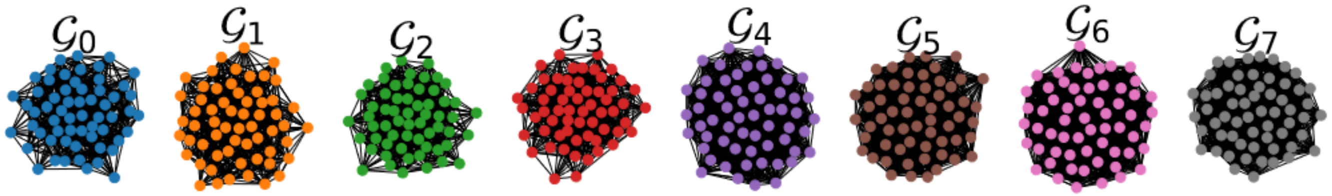
- $\log 2 \geq \mathrm{JS}(\mathcal{P},\mathcal{P}') \geq 0$
- $\mathrm{JS}(\mathcal{P},\mathcal{P}) = 0$
- If $\mathcal{P}$ and $\mathcal{P}'$ have disjoint support, $\mathrm{JS}(\mathcal{P},\mathcal{P}) = \log 2$

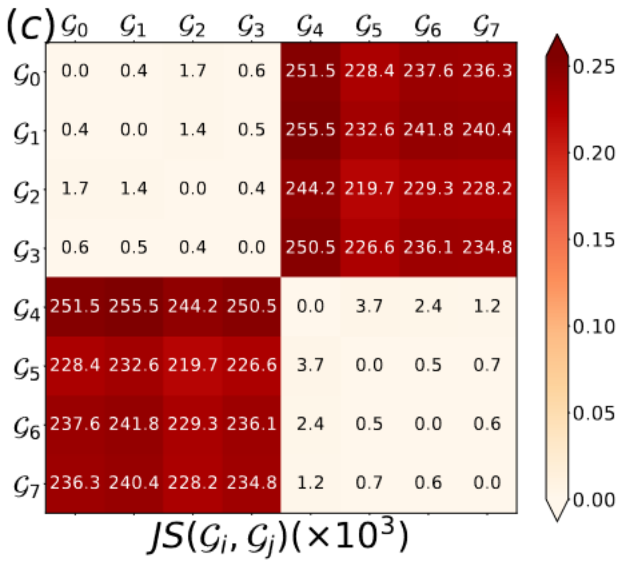
With $H(\mathcal{P})$ the Shannon entropy of $\mathcal{P} = \{p_k\}_k$

$$H(\mathcal{P}) = -\textstyle\sum_k p_k \log p_k$$

The kernel is then defined as

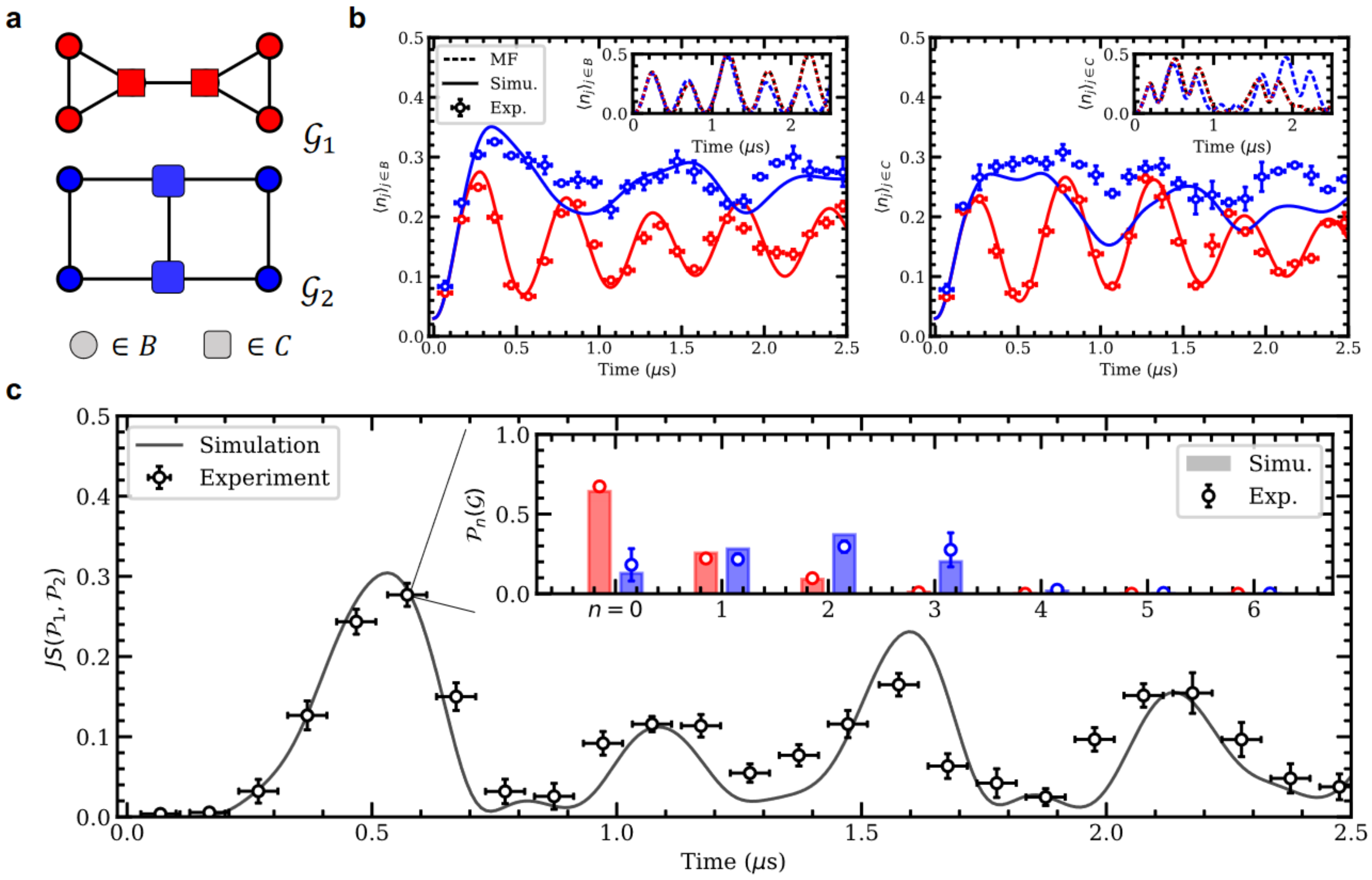
$$K_\mu(\mathcal{P},\mathcal{P}') = e^{-\mu\, JS(\mathcal{P},\mathcal{P}')} \in [2^{-\mu}, 1]$$

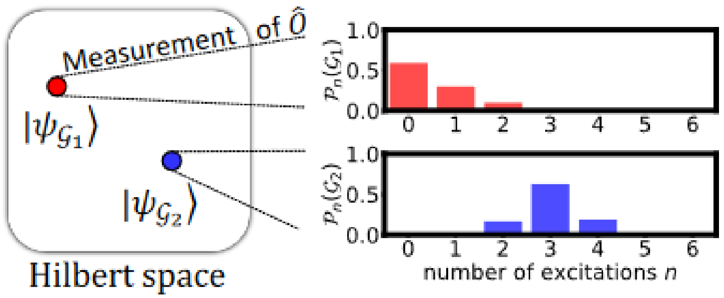




PASQAL

Hardware example

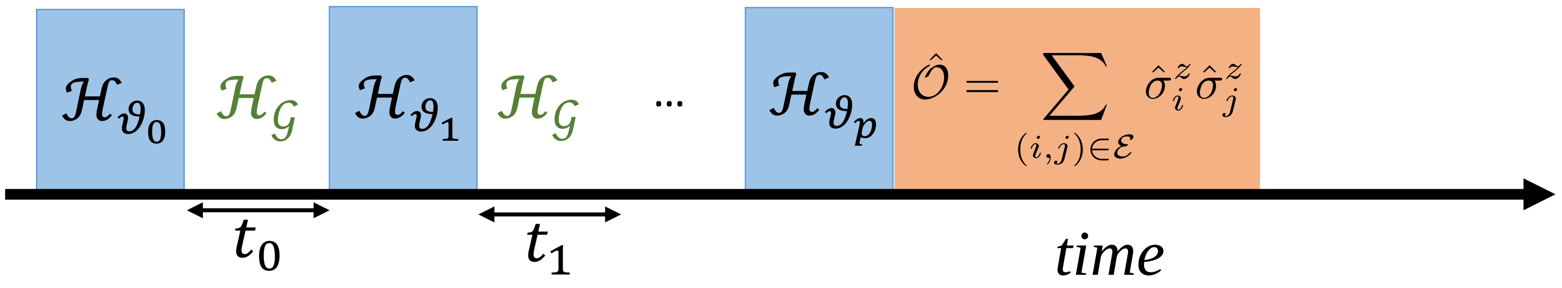




The two graphs are locally identical, yet, the use of the QEK allows to differentiate their time evolution, therefore leveraging non-local quantities to differentiate the graphs.

Benchmark - training

$$|\psi_f(\Lambda)\rangle = \prod_{i=1}^{p} \left(\mathrm{e}^{-\mathrm{i}\hat{\mathcal{H}}_{\theta_i}} \mathrm{e}^{-\mathrm{i}\hat{\mathcal{H}}_{\mathcal{G}} t_i} \right) \mathrm{e}^{-\mathrm{i}\hat{\mathcal{H}}_{\theta_0}} |\psi_0\rangle$$



$$\hat{\mathcal{H}}_{\mathcal{G}} = \begin{cases} \hat{\mathcal{H}}_{\text{Ising}} = \sum_{(i,j)\in\mathcal{E}} \hat{\sigma}_i^z \hat{\sigma}_j^z \\ \hat{\mathcal{H}}_{\text{XY}} = \sum_{(i,j)\in\mathcal{E}} (\hat{\sigma}_i^+ \hat{\sigma}_j^- + h.c.) \end{cases}$$

$$\hat{\mathcal{H}}_{\vartheta} = \vartheta \sum_{i\in\mathcal{V}} \hat{\sigma}_i^y$$

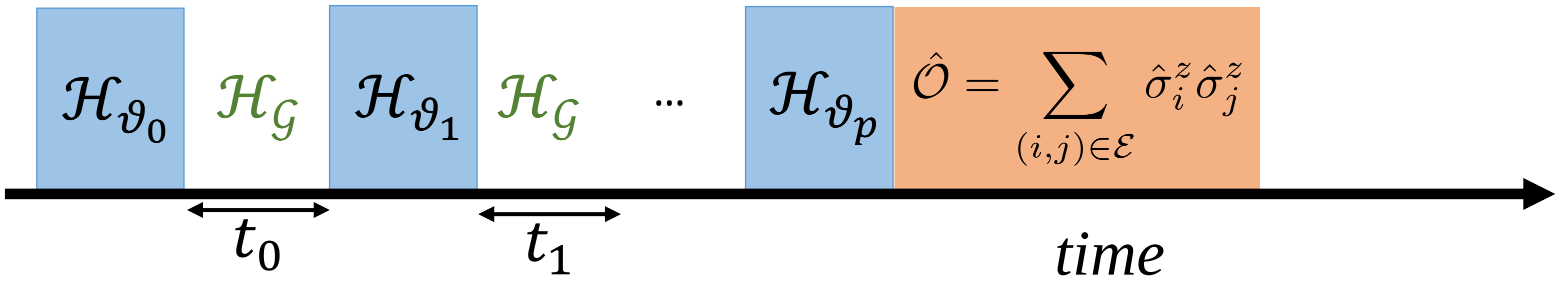
$\Lambda = \{\theta_1, t_1, \theta_2, t_2, \theta_3, \ldots,\}$ trained by bayesian optimisation of the accuracy (the % of graph properly labeled)

Information about the graphs is obtained through controlled quantum evolution of a register of atoms (**using our resource)** into easily differentiable states.

Dataset	samples	classes	samples per classes
IMDB-MULTI	1185	3	371, 403, 411
IMDB-BIN	499	2	239, 260
PTC_FM	234	2	135, 99
PROTEINS	307	2	82, 225
NCI1	361	2	282, 79
Fingerprint	1467	3	515, 455, 597

Benchmark - results

$$|\psi_f(\Lambda)\rangle = \prod_{i=1}^{p} \left(\mathrm{e}^{-\mathrm{i}\hat{\mathcal{H}}_{\theta_i}} \mathrm{e}^{-\mathrm{i}\hat{\mathcal{H}}_{\mathcal{G}} t_i} \right) \mathrm{e}^{-\mathrm{i}\hat{\mathcal{H}}_{\theta_0}} |\psi_0\rangle$$



$$\hat{\mathcal{H}}_{\mathcal{G}} = \begin{cases} \hat{\mathcal{H}}_{\text{Ising}} = \sum_{(i,j)\in\mathcal{E}} \hat{\sigma}_i^z \hat{\sigma}_j^z \\ \hat{\mathcal{H}}_{\text{XY}} = \sum_{(i,j)\in\mathcal{E}} (\hat{\sigma}_i^+ \hat{\sigma}_j^- + h.c.) \end{cases}$$

$$\hat{\mathcal{H}}_{\vartheta} = \vartheta \sum_{i\in\mathcal{V}} \hat{\sigma}_i^y$$

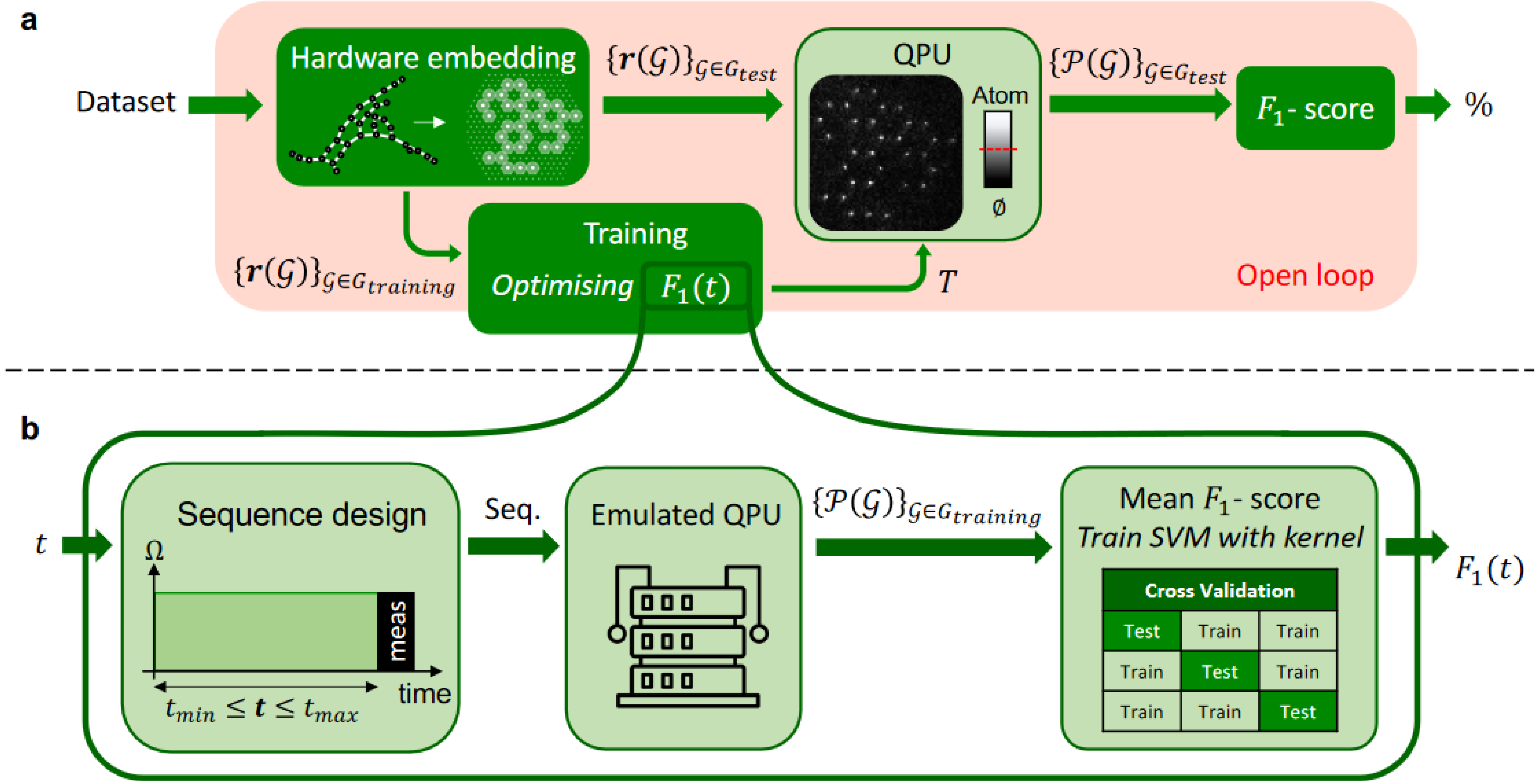
Accuracy

Dataset	Ising_1 (150)	Ising_4 (2000)	Ising_8 (6000)	XY_4 (2000)	GS	RW
IMDB-MULTI	46.8 ± 4.4	$\mathbf{48.1 \pm 4.4}$	47.7 ± 4.4	47.5 ± 4.5	40.9 ± 3.5	45.2 ± 3.4
IMDB-BIN	69.0 ± 6.1	71.6 ± 5.7	$\mathbf{71.8 \pm 5.4}$	70.6 ± 5.6	66.5 ± 5.9	67.8 ± 6.5
PTC_FM	62.5 ± 7.9	65.8 ± 7.9	$\mathbf{66.0 \pm 7.6}$	65.2 ± 8.2	61.5 ± 8.9	59.4 ± 7.8
PROTEINS	73.3 ± 1.2	74.5 ± 2.6	$\mathbf{76.0 \pm 5.3}$	74.8 ± 3.7	73.3 ± 1.2	73.3 ± 1.2
NCI1	78.1 ± 0.8	78.6 ± 3.2	$\mathbf{80.1 \pm 3.5}$	78.8 ± 4.8	78.1 ± 0.8	78.1 ± 0.8
Fingerprint	58.6 ± 2.0	$\mathbf{60.2 \pm 3.2}$	60.1 ± 3.3	60.1 ± 3.3	57.9 ± 3.3	59.9 ± 2.2

PASQAL

Results on datasets

Predictive Toxicity Challenge on Female Mice (PTC-FM)

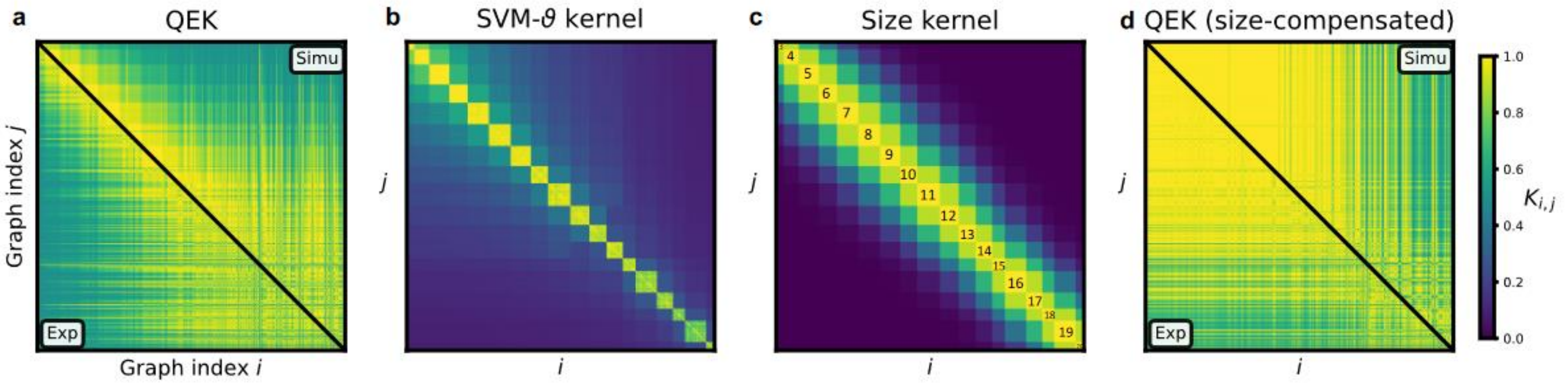


Kernel	F_1-score (%)
QEK	60.4 ± 5.1
QEK (size-compensated)	45.1 ± 3.7
SVM-ϑ	58.2 ± 5.5
Size	56.7 ± 5.6
Graphlet Sampling	56.9 ± 5.0
Random Walk	55.1 ± 6.9
Shortest Path	49.8 ± 6.0

TABLE I. F_1-score reached experimentally on the PTC-FM dataset by QEK ($\pm$ std. on the splits). In addition, the scores reached numerically by the classical kernels SVM$-\vartheta$, Size, Graphlet Sampling, Random Walk and Shortest-Path. The values reported are the average over a 5-fold cross-validation repeated 10 times.

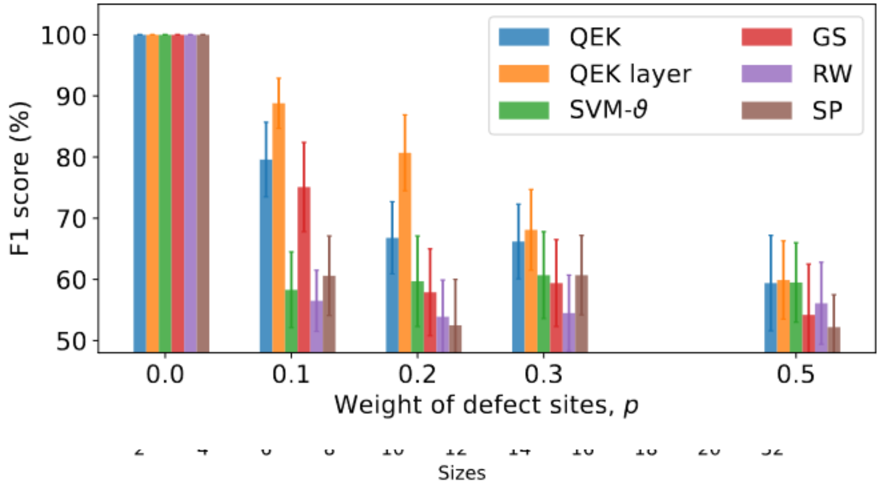
Albrecht, B., Dalyac, C., Leclerc, L., Ortiz-Gutiérrez, L., Thabet, S., D'Arcangelo, M., Cline, J.R., Elfving, V.E., Lassablière, L., Silvério, H. and Ximenez, B., 2023. Quantum feature maps for graph machine learning on a neutral atom quantum processor. *Physical Review A*, *107*(4), p.042615.

Predictive Toxicity Challenge on Female Mice (PTC-FM)



Kernel	F_1-score (%)
QEK	60.4 ± 5.1
QEK (size-compensated)	45.1 ± 3.7
SVM-ϑ	58.2 ± 5.5
Size	56.7 ± 5.6
Graphlet Sampling	56.9 ± 5.0
Random Walk	55.1 ± 6.9
Shortest Path	49.8 ± 6.0

TABLE I. F_1-score reached experimentally on the PTC-

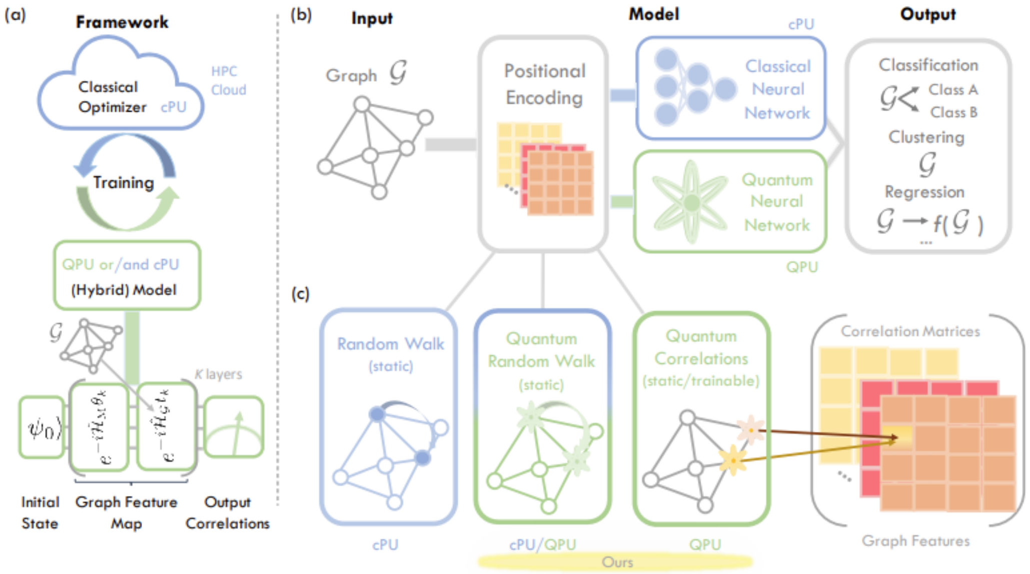


Mmm: QEK is not amazing if you remove the dependence on graph size. On artifical datasets (lattice with defects), the quantum kernel performs well

Albrecht, B., Dalyac, C., Leclerc, L., Ortiz-Gutiérrez, L., Thabet, S., D'Arcangelo, M., Cline, J.R., Elfving, V.E., Lassablière, L., Silvério, H. and Ximenez, B., 2023. Quantum feature maps for graph machine learning on a neutral atom quantum processor. *Physical Review A*, *107*(4), p.042615.

ENHANCING GRAPH NEURAL NETWORKS WITH QUANTUM COMPUTED ENCODINGS

Slimane Thabet*, Romain Fouilland, Mehdi Djellabi, Igor Sokolov, Sachin Kasture
Louis-Paul Henry & Loïc Henriet
PASQAL, Massy, France
{slimane.thabet, loic}@pasqal.com



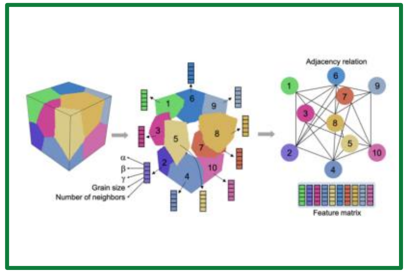
Potentially many more, its a vast field of study…

PASQAL

Conclusion

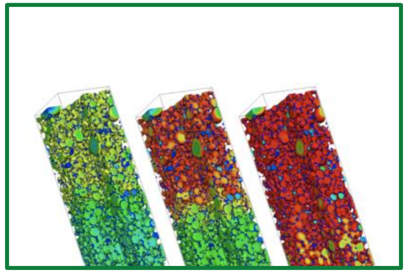
Classes of problems

Graph neural networks



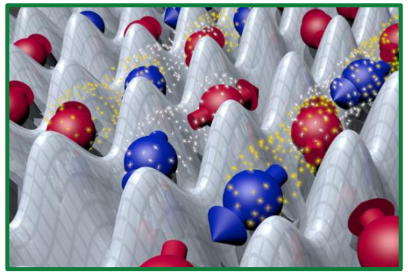
- Quantum Evolution Kernel (QEK)

Differential Equations



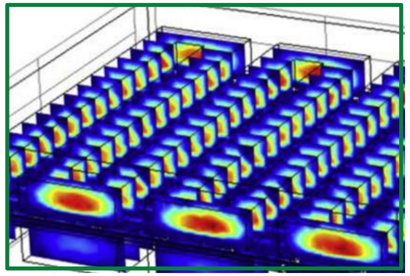
Differential Quantum Circuits (DQC)

Materials-science simulations



Hamiltonian simulation

Optimization problems

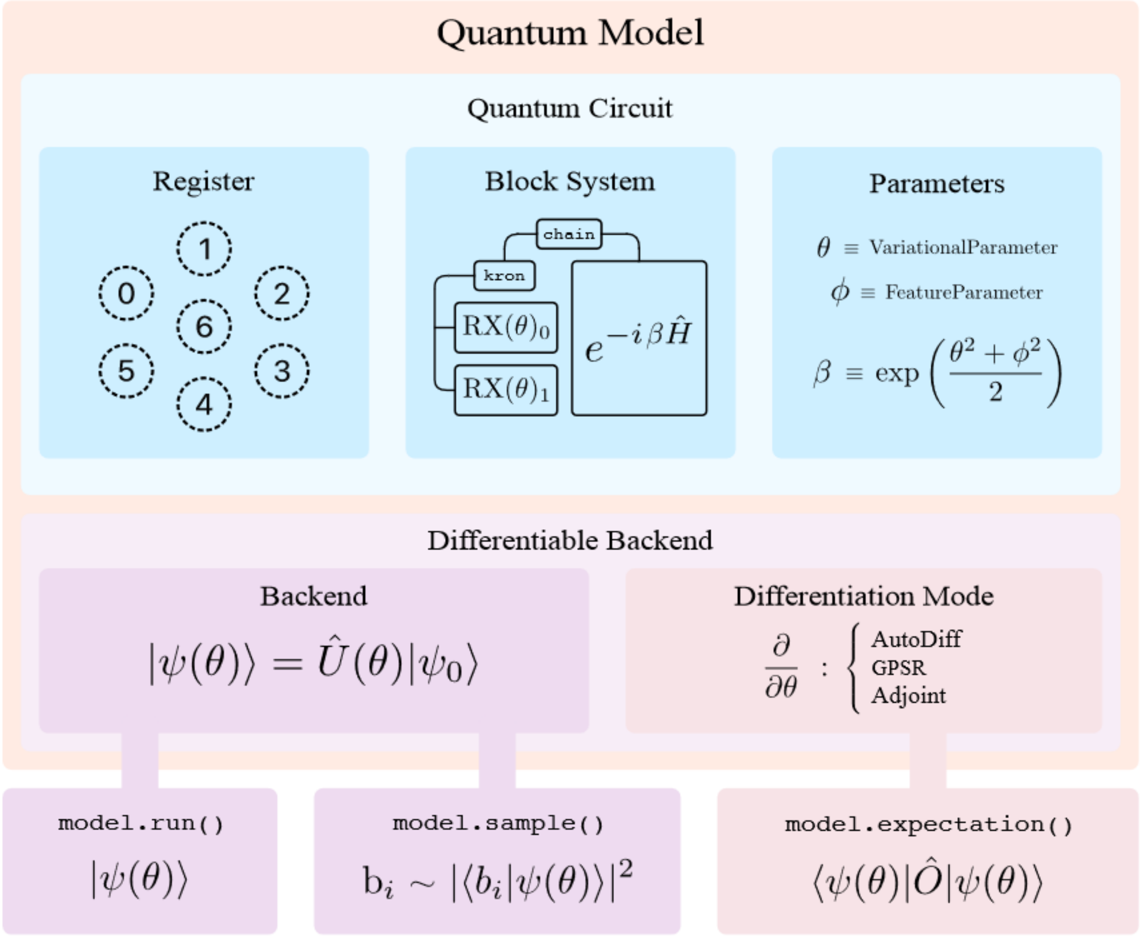


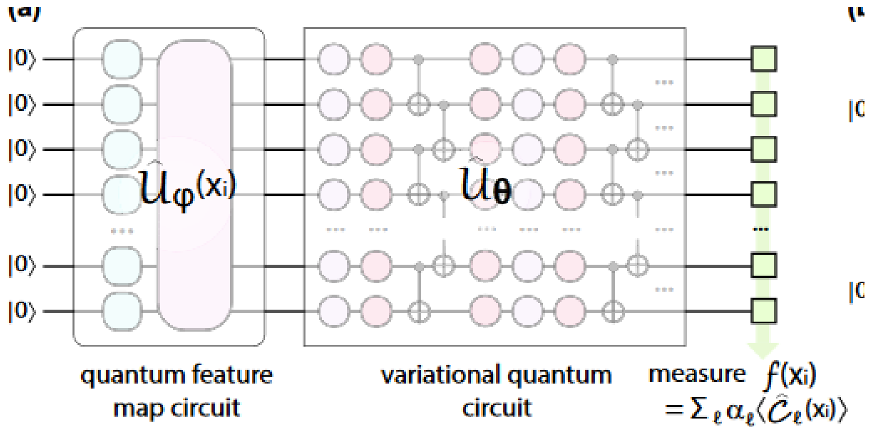
Quantum Extremal Learning (QEL)

- We saw one type of approach based on graphs
- But the analog device can also encode efficient <u>feature maps</u>
  - Useful for DQC and QEL…
  - Digital-analog approaches

PASQAL

Qadence: a good playground for neutral atom analog QML



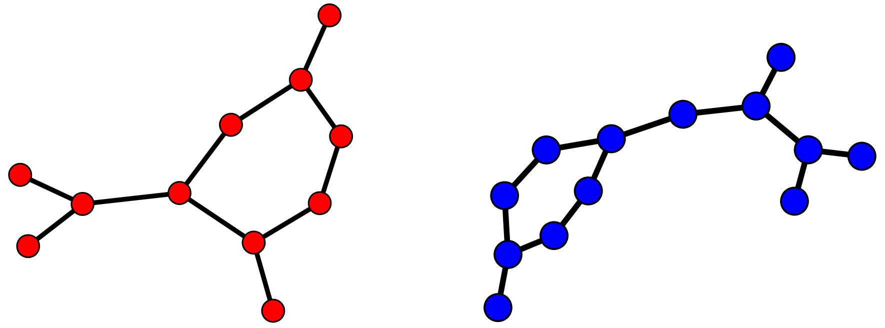


See [Qadence - Qadence (pasqal-io.github.io)]
Ex. Kyriienko, Oleksandr, Annie E. Paine, and Vincent E. Elfving. "Solving nonlinear differential equations with differentiable quantum circuits." *Physical Review A* 103, no. 5 (2021): 052416.

PASQAL

Conclusion

- Rydberg atoms are a promising route for analog quantum computing, with a high degree of control of the quantum states.
- Rydberg blockade interaction leads to a natural graph structure to explore
- Variational algorithms can help create efficient graph kernels
- There are many more open directions for research.



THANK YOU

QUESTIONS ?

If you are interested by PASQAL → Visit pasqal.com for more infos & opportunities

If you are interested in work done at UdeS and the Institut Quantique – contact me.

PASQAL