Machine learning and optical quantum information

K. Bartkiewicz[1,2,3], A. Černoch[2], C. Gneiting[3]
K. Jiráková[2], K. Lemr[2], F. Nori[3,4]
J. Roik[2], V. Trávníček[2]

[1]Institute of Spintronics and Quantum Information,
Faculty of Physics, Adam Mickiewicz University, Poland

[2]Joint Laboratory of Optics of Palacký University
and Institute of Physics of Czech Academy
of Sciences, Czech Republic

[3]Theoretical Quantum Physics Laboratory,
RIKEN Cluster for Pioneering Research, Japan

[4]Department of Physics, The University of Michigan, USA

QO X, Toruń, 10.09.2021

Timeline

1900
- Max Planck discovers quantization of black –body radiation

1905
- Albert Einstein discovers quantization of light

1923
- De Broglie postulates the duality of waves and particles

1925
- Heisenberg, Jordan, Born formulate matrix mechanics

1926
- Erwin Schrodinger formulates wave mechanics

1930-1932
- P. Dirac and J. von Neuman formulate modern quantum theory

TIME
- Quantum theory branches out into physical subdisciplines

1976
- R. S. Ingarden publishes "Quantum Information Theory"

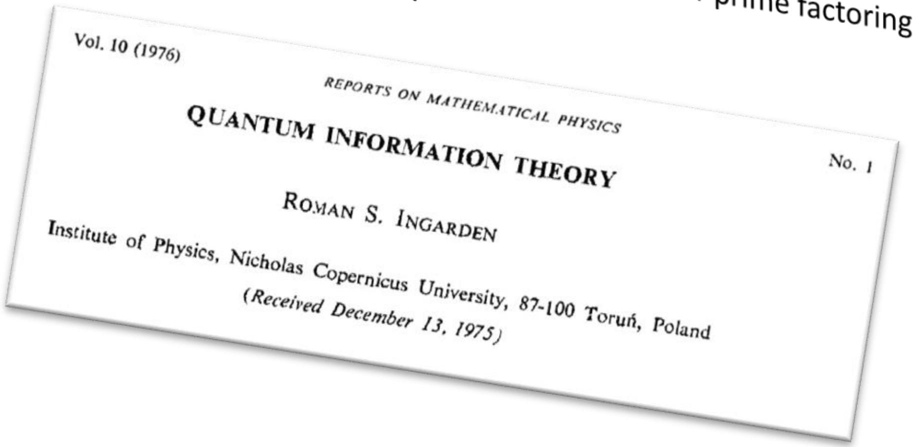
1994
- Peter Shor proposes a quantum algorithm for prime factoring

1996
- Luv Grover proposes a quantum algorithm for unordered search

2009
- Harrow, Hassidim & Lloyd propose a quantum algorithm to solve linear systems of equations

~2013
- Machine Learning starts to become very popular due to technological progress

TIME
- **Quantum Machine Learning becomes a subdiscipline of quantum theory**


Vol. 10 (1976) REPORTS ON MATHEMATICAL PHYSICS No. 1

QUANTUM INFORMATION THEORY

ROMAN S. INGARDEN

Institute of Physics, Nicholas Copernicus University, 87-100 Toruń, Poland

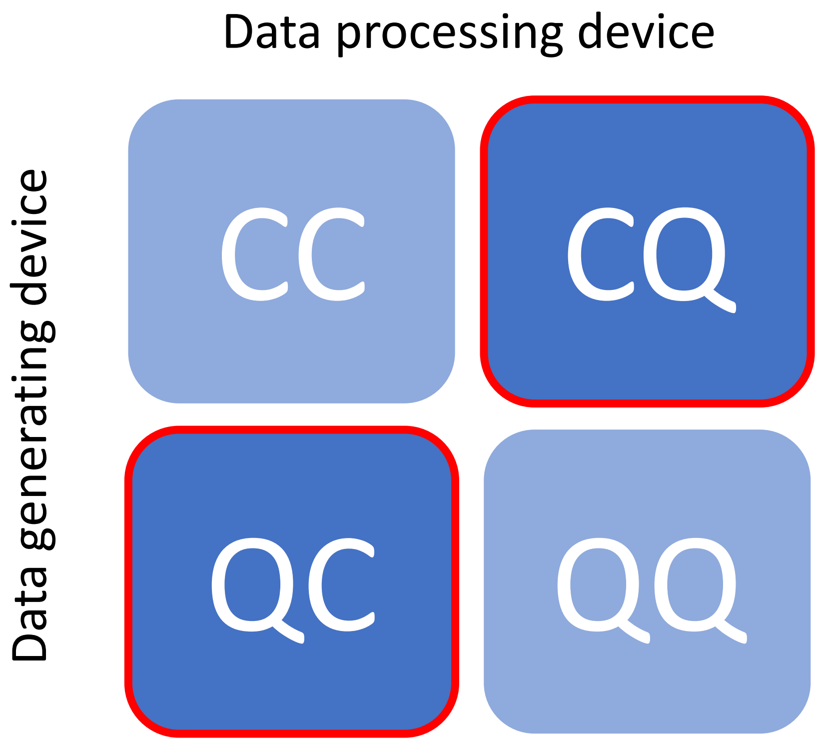
(Received December 13, 1975)

Context

- **ML:** route selection (Google maps), friend tagging suggestions (Facebook: deepface), ETA and personalization (Uber), iPhone: FaceId (Apple), self-driving car (Tesla), recommender systems at Netflix (75% success rate) and Amazon (35% of revenue), Google Translate (most statistically probable translation), KUKA (industrial robots – can learn ping-pong), IT security, …
- **Challenges in classical ML:** instability of generative models, unsupervised learning, explainability, confidence intervals, reasoning and probabilistic inference, energy consumption (M. Matuszewski)…
- **Better ML with quantum physics? Better quantum physics with ML?** (e.g. Google, IBM, Microsoft, Xanadu, Dwave, NASA, NTT, RIKEN, …)…

Typology of ML in quantum physics

- There are several approaches to **quantum machine learning (QML)** depending on the type of **data/algorithm**.
- Here, we focus on **CQ, QC, and classification, and clustering** problems.
- There are two main approaches to creating a highly nonlinear predictive model: **quantum kernels** and **quantum ANN.**



Classification: two main QML approaches

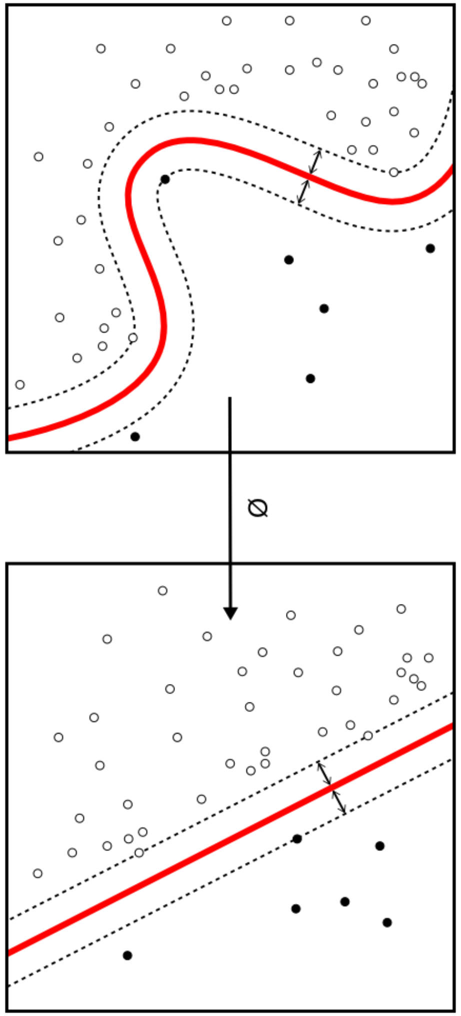
Quantum Artificial Neural Networks:

- F. Tacchino et al., npj Quantum Information **5**, 26 (2019)
- E. Farhi and H. Neven, arXiv:1802.06002 …

Kernel-based QML (SVM in larger space):

IBM
- V. Havlíček et al., Nature **567**, 209 (2019)
- KB et al., Sci. Rep. **10**, 12356 (2020), manuscript ready in early 2019.

Xanadu
- M. Schuld and N. Killoran, Phys. Rev. Lett. **122**, 040504 (2019)
- R. Chatterjee and T. Yu, Quantum Inf. Commun. **17**, 1292 (2017)



CQ

SVMs gained popularity in the 1990s, method of choice for many practical problems.

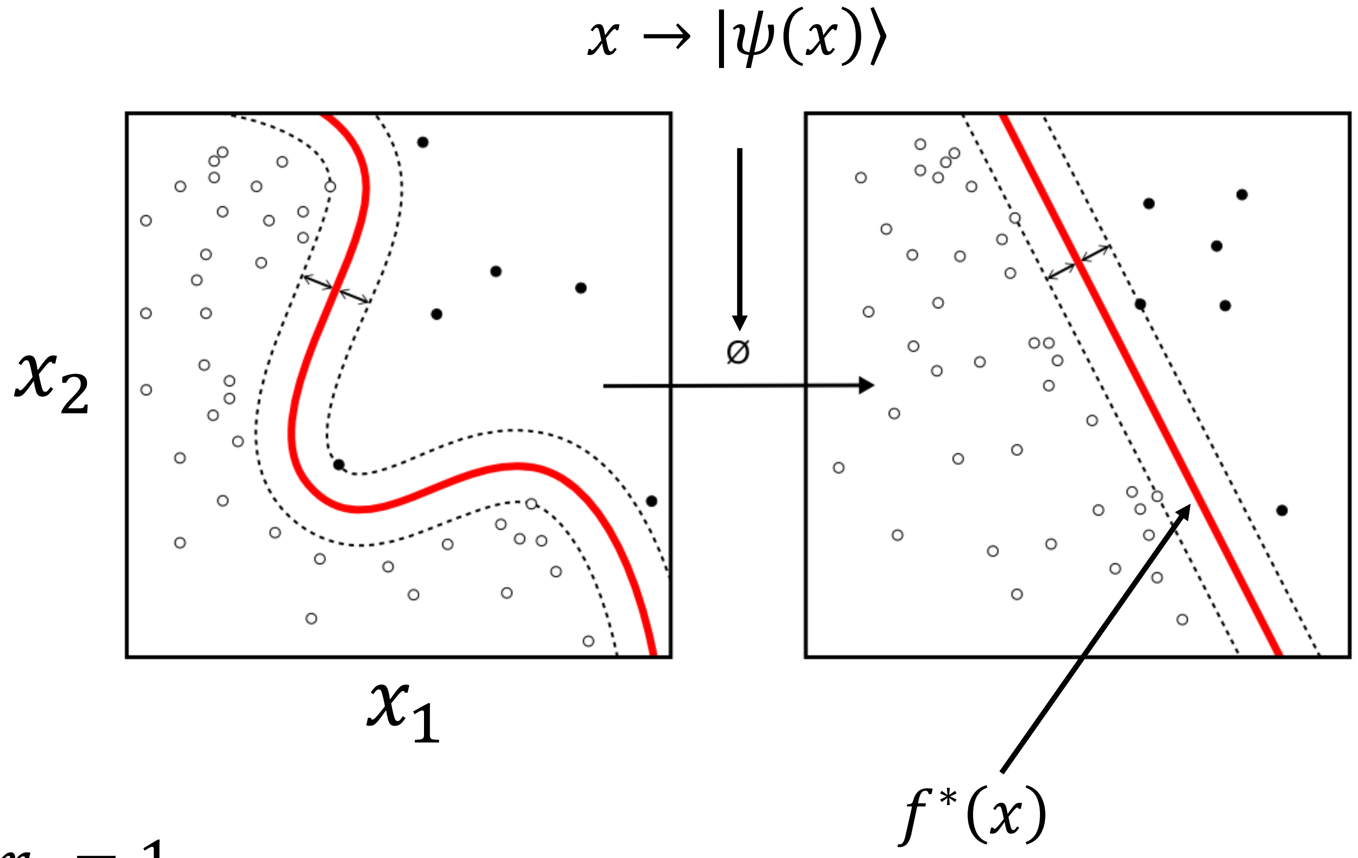
Kernel-based QML

- Representer theorem (kernel trick):

$$f^*(x) = \sum_{m=1}^{M} a_m \kappa(x, x^{(m)})$$

- Feature map (FM), 1D

$$x \to |\psi(x)\rangle = \sum_{n=0}^{N} \sqrt{r_n}\, e^{i2\pi n x} |n\rangle, \sum_{n=0}^{N} r_n = 1$$



$$\kappa(x, x') = |\langle \psi(x') | \psi(x) \rangle|^2$$

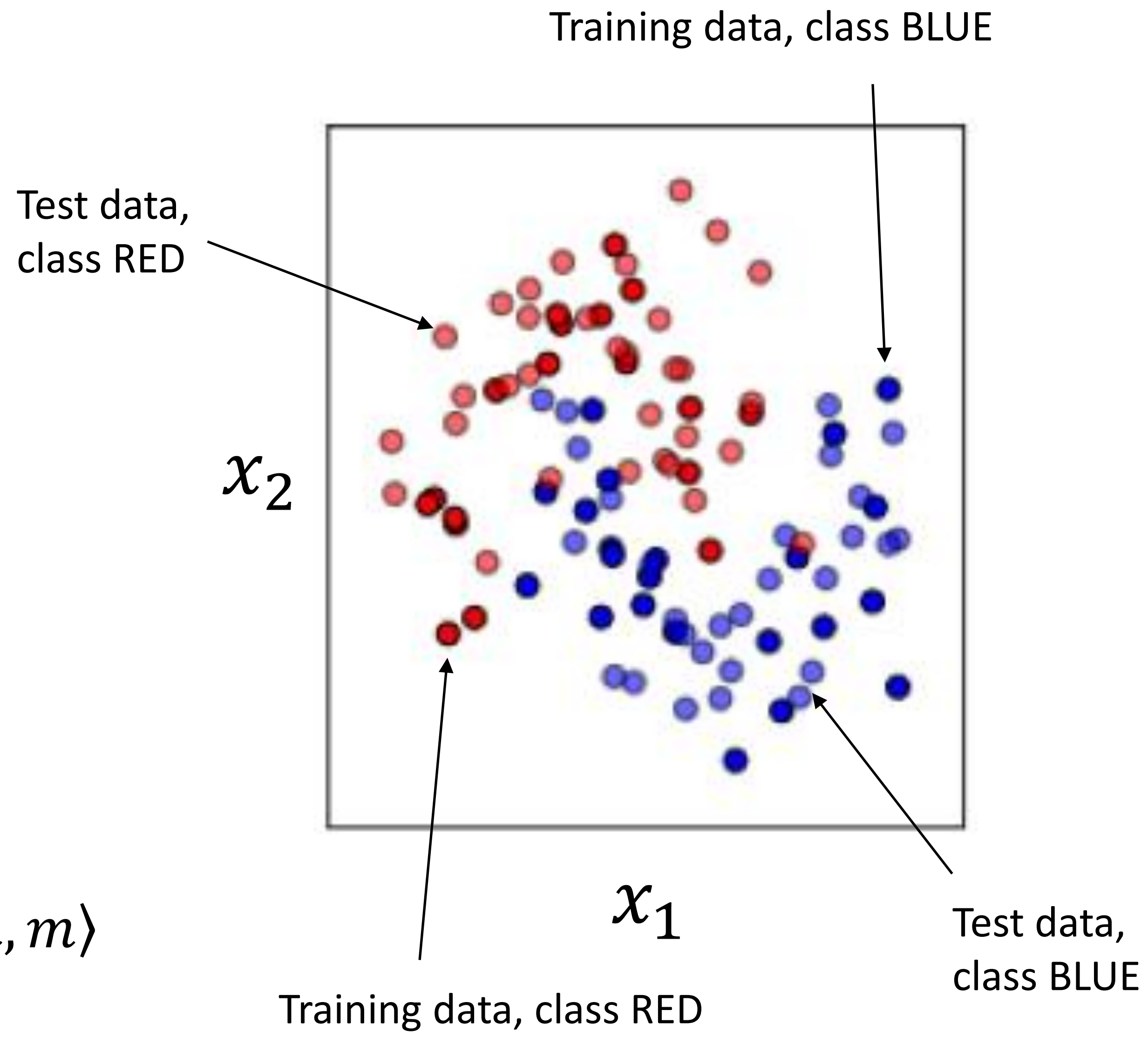
Kernel-based QML

- Representer theorem (kernel trick):

$$f^*(x) = \sum_{m=1}^{M} a_m \kappa(x, x^{(m)})$$

- Feature map (FM), 2D

$$x \to |\psi(x)\rangle = \sum_{n=0}^{N} \sqrt{r_n r_m}\, e^{i2\pi(nx_1+mx_2)} |n, m\rangle$$

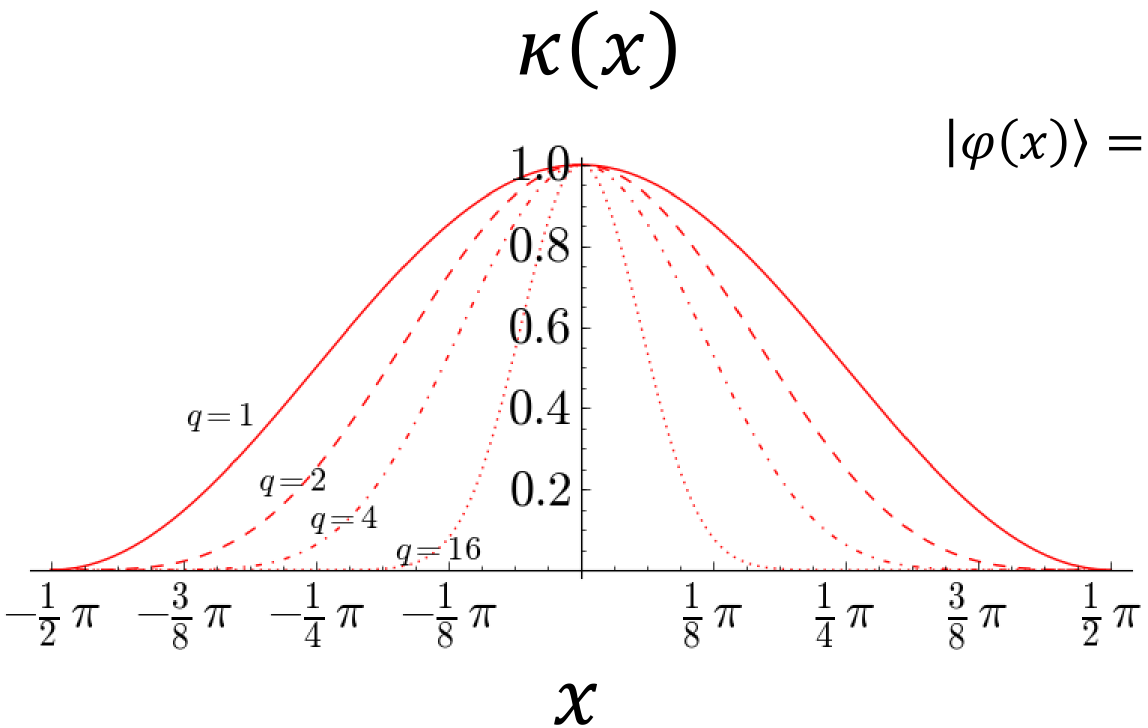


Kernel-based QML: cosine kernel

- Definition:

$$\kappa(x', x) = |\langle \varphi(x') | \varphi(x) \rangle|^2 = \prod_{n=1}^{D} \cos^{2N}(x'_n - x_n)$$

- Resolution of a naive FM (angle encoding):



$$|\varphi(x)\rangle = \bigotimes_{m=1}^{D} \bigotimes_{n=1}^{N} \sum_{k=0}^{1} \sin^k(x_n) \cos^{1-k}(x_n) \; |k\rangle_{n,m}$$

$$q = DN$$

Number of qubits

NOTATION in 1D: $\kappa(x', x) = \kappa(x' - x)$

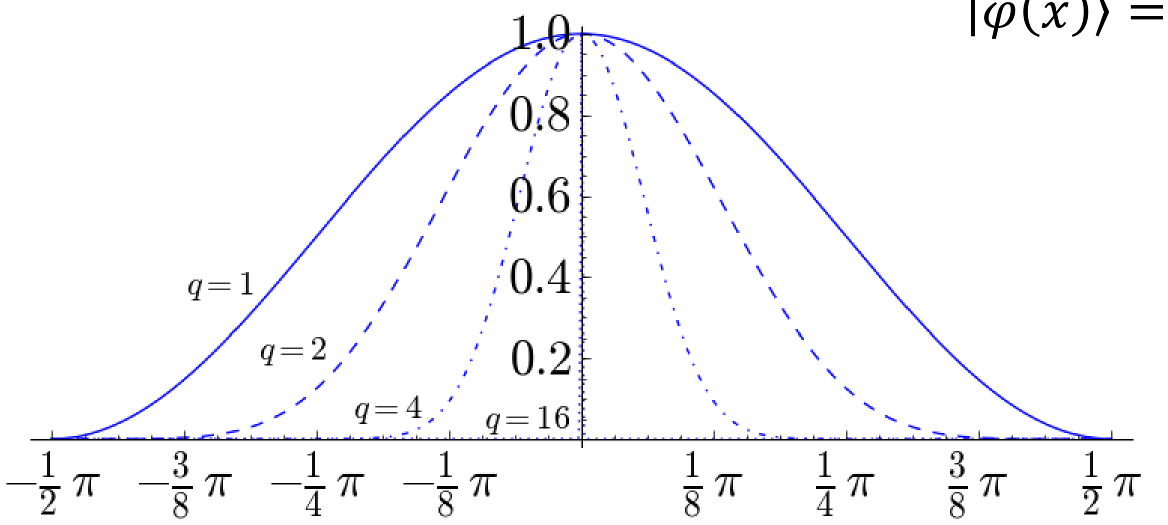
Kernel-based QML: cosine kernel

- Definition:

$$\kappa(x',x) = |\langle\varphi(x')|\varphi(x)\rangle|^2 = \prod_{n=1}^{D} \cos^{2N}(x'_n - x_n)$$

- Resolution enhancement:

$\kappa(x)$



x

$$|\varphi(x)\rangle = \bigotimes_{m=1}^{D} \sum_{k=0}^{N} \sqrt{\binom{N}{k}} \sin^k(x_n) \cos^{N-k}(x_n)\ |k\rangle_m$$

$$q = D\lceil \log_2(N+1) \rceil$$

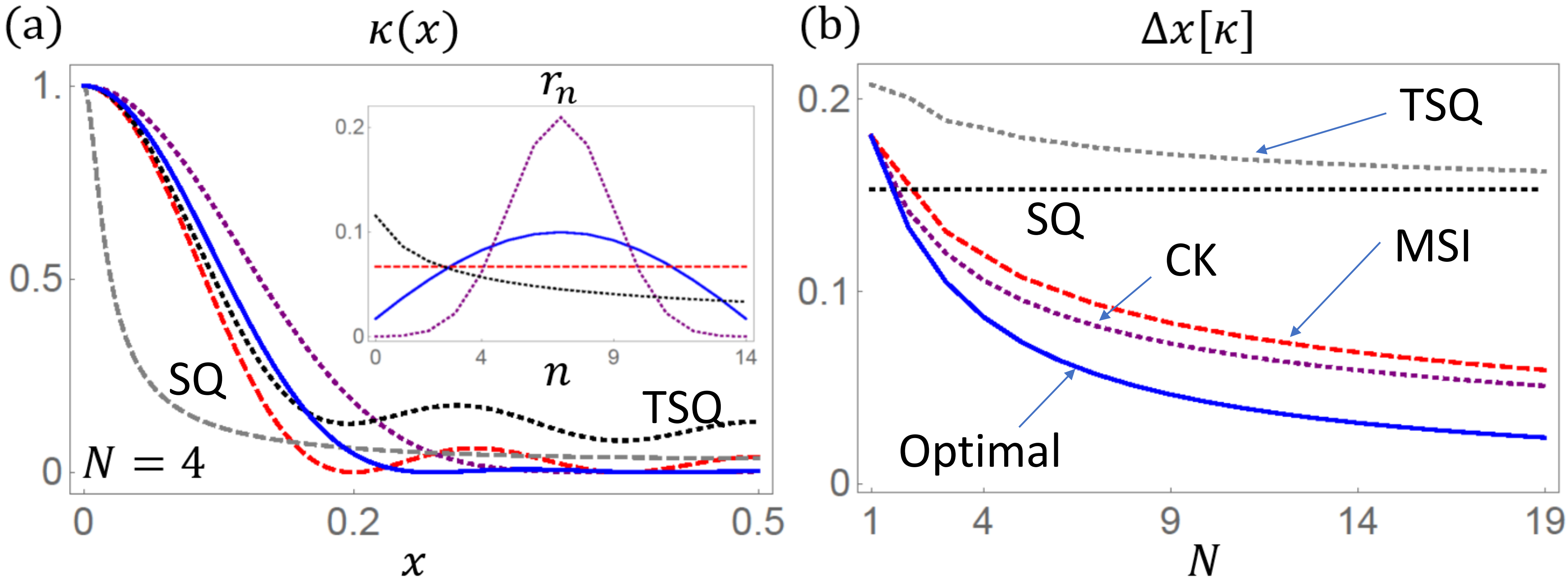
Multiple FMs give raise to the same kernel!

Exponential enhancement!

NOTATION in 1D: $\kappa(x',x) = \kappa(x'-x)$

NOTATION in 1D: $\kappa(x', x) = \kappa(x' - x)$

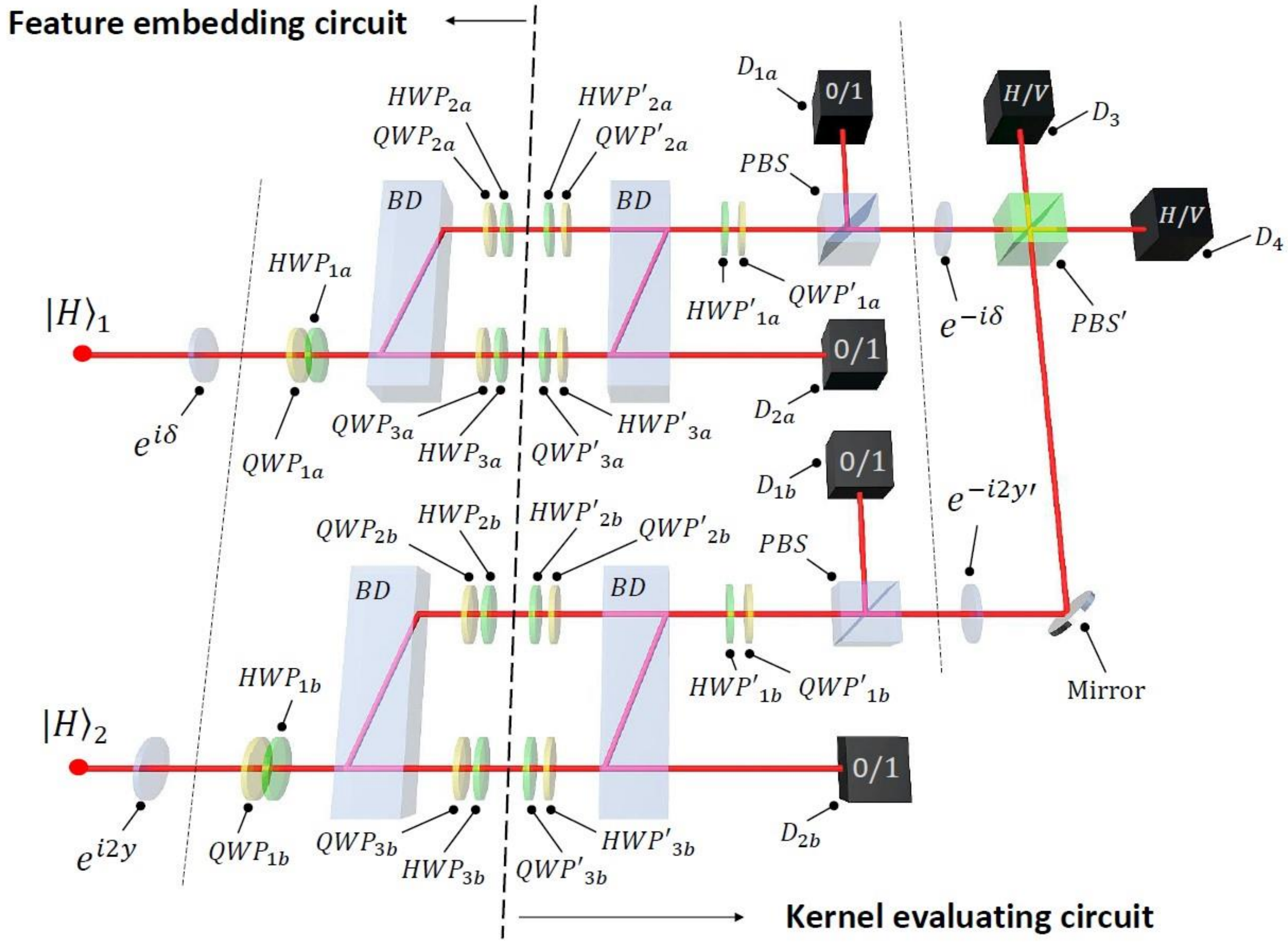
Kernel-based QML: finite HS and resolution (1D)

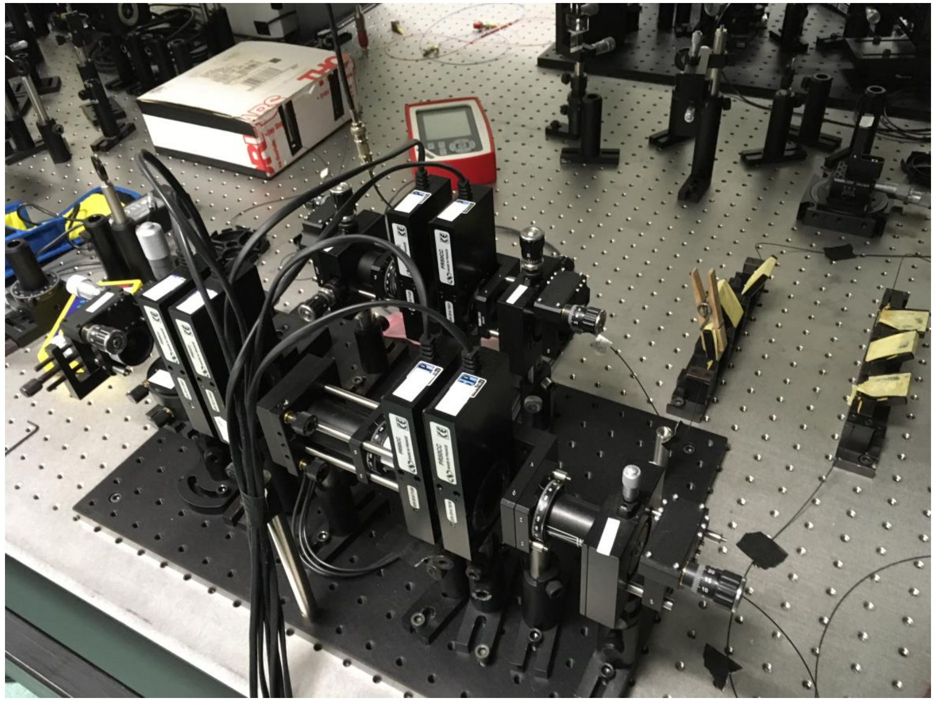


Truncated squeezed states (**TSQ**), Multi-slit interference states (**MSI**), Cosine kernel (**CK**)

Optimal map can be well approximated as a product of CK and MSI maps, where the joint size of the HS is N.

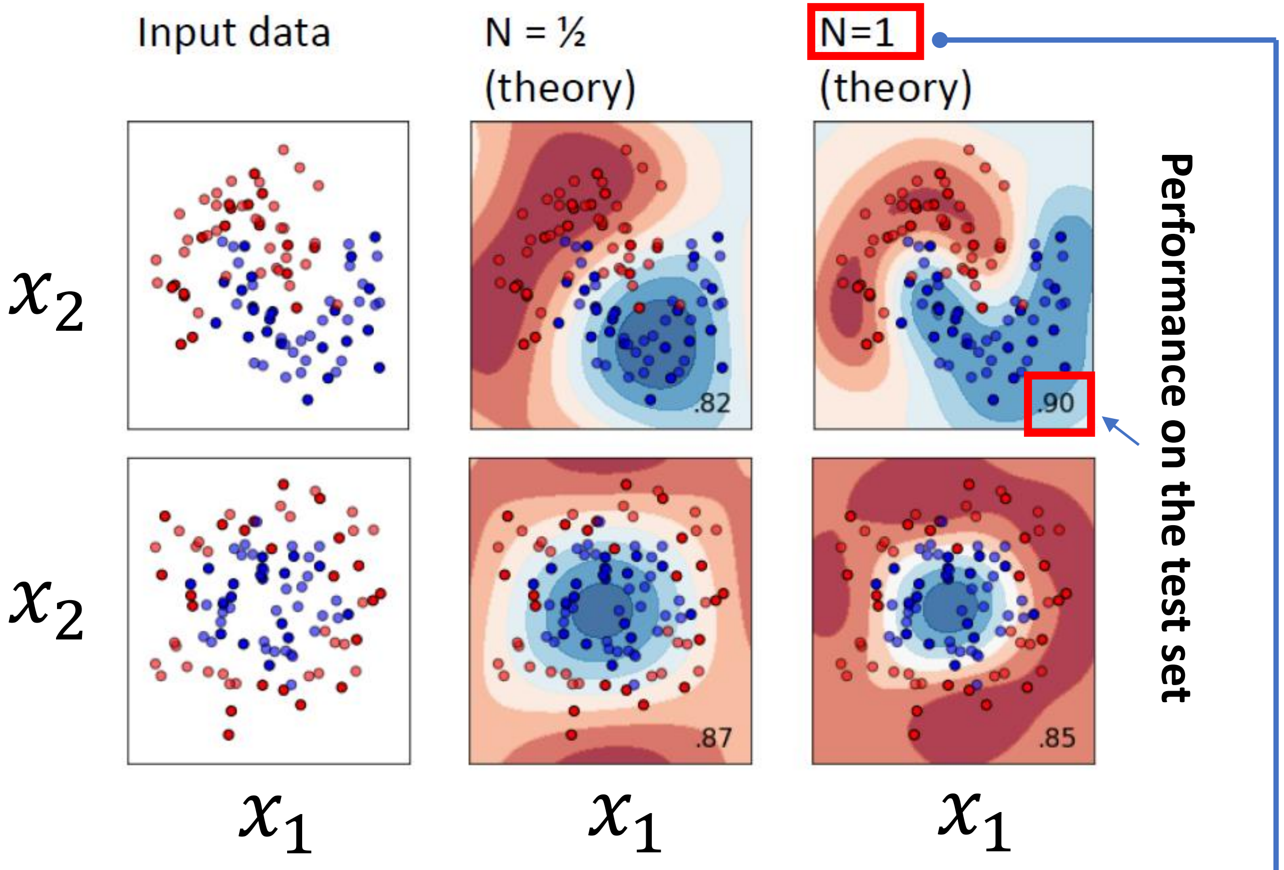
KQML: two-photon setup (2D)





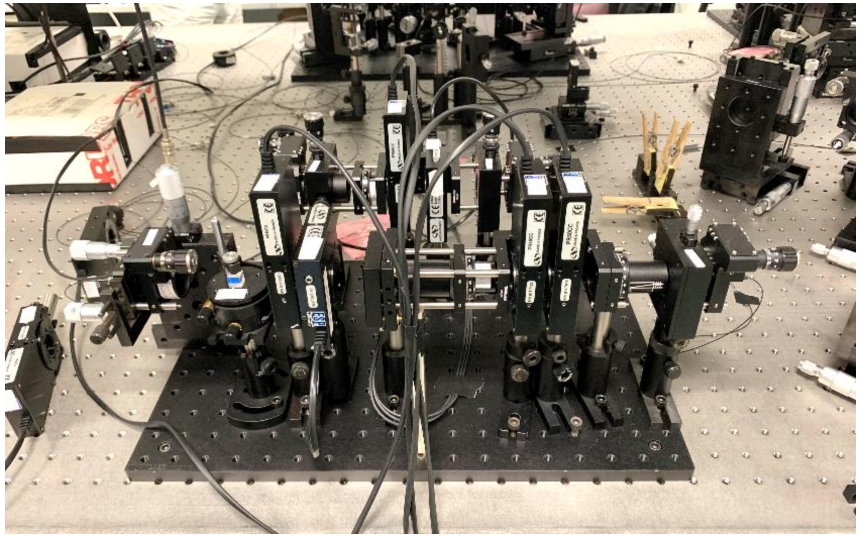
$$|\varphi(x)\rangle = \bigotimes_{n=1}^{2}\left(\cos^3(x_n)\,|HT\rangle_n + \sqrt{3}\sin(x_n)\cos^2(x_n)\,|HB\rangle + \sqrt{3}\cos(x_n)\sin^2(x_n)\,|VB\rangle_n + \sin^2(x_n)\,|VT\rangle_n\right)$$

KQML: experimental implementation

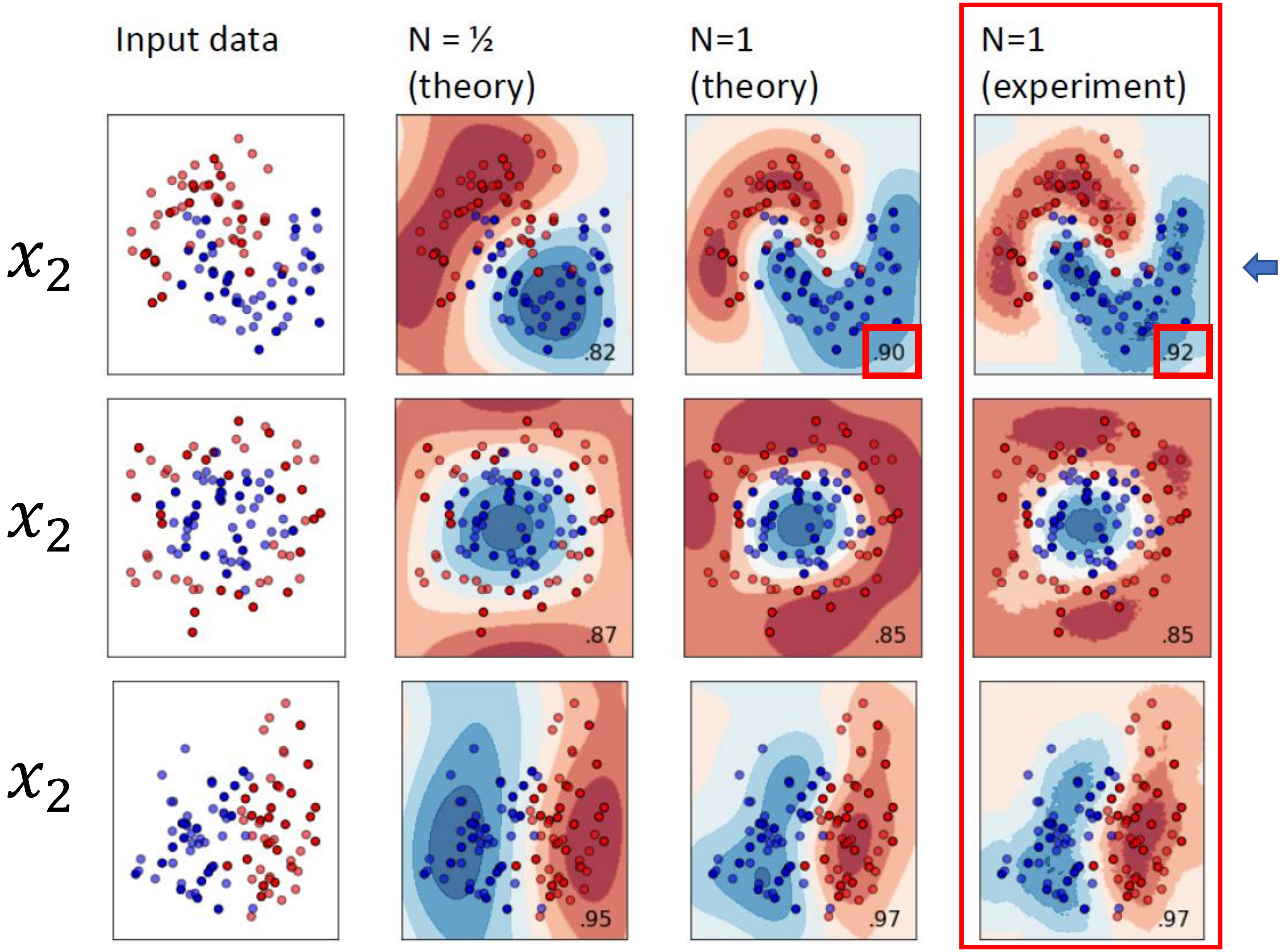


$$\kappa(x', x) = |\langle\phi(x')|\phi(x)\rangle|^2 = \prod_{n=1}^{D} \cos^{2N}(x'_n - x_n)$$

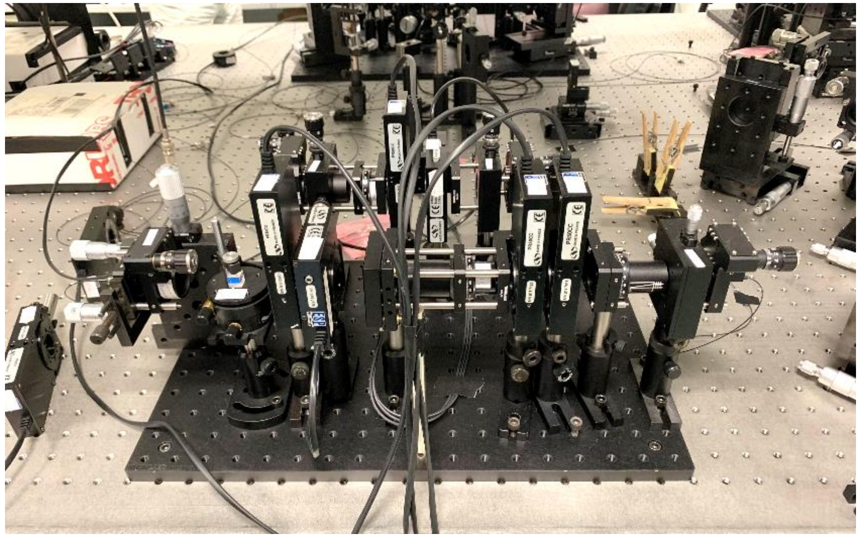
Can experimental KQML outperform a classical computer?
(**scikit-learn SVC in Python**)



KQML: experimental implementation



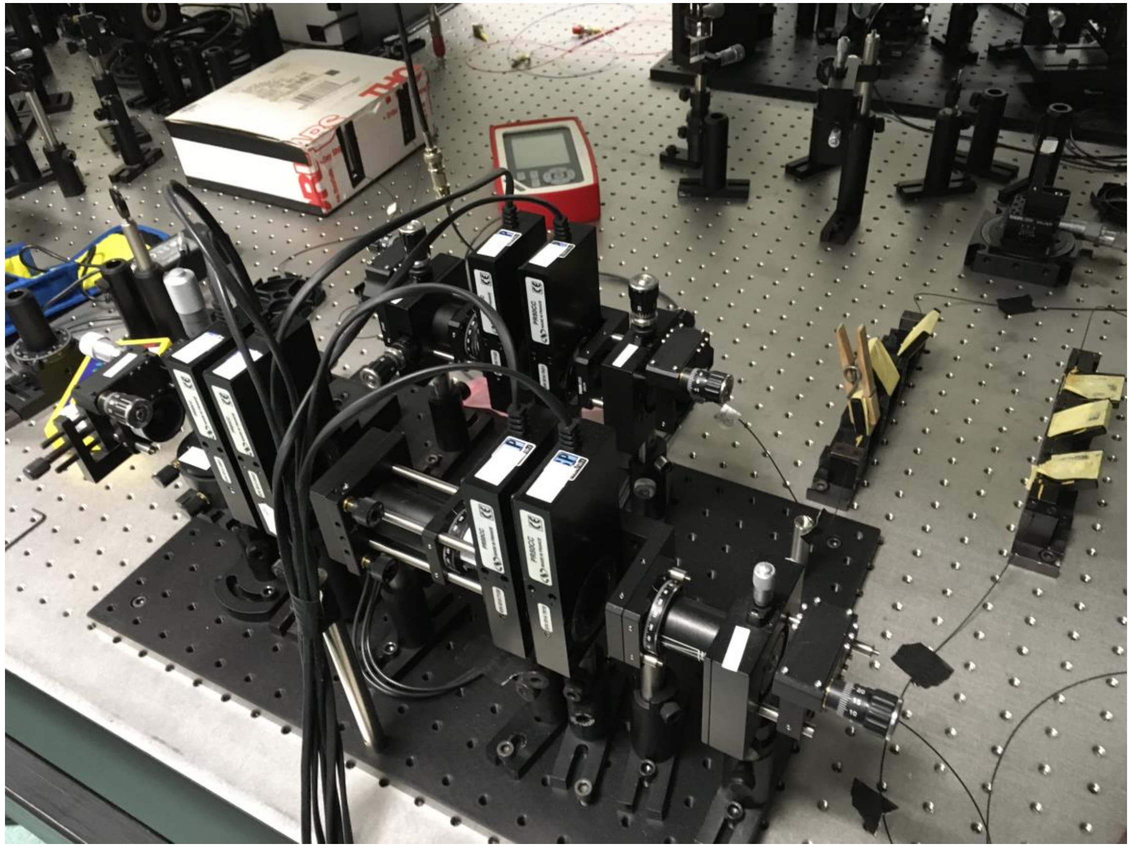
Better performance on the test set
(by a data point close to the decision boundary)



Supervised KQML: Advantage

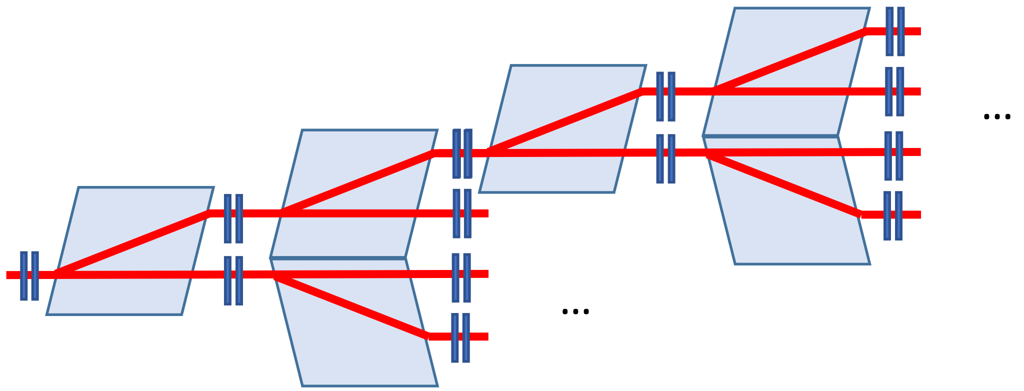
The quantum advantage:
classical power kernel: $O[\log(N)]$
(recursive classical algorithm)
vs. quantum kernel: $O(1)$**.**
[Grover-Rudolph-like speedup:
factor of $1/\log(N)$]

We (and other groups) work on observing **quantum advantage in the expressive power** of quantum kernels for some datasets.



KQML: Scaling to higher dimensions

- **Higher dimensions** by properly **stacking layers of BDs and waveplates.**
- We **did NOT use layers** to demonstrate the main advantage (simplicity) of KQML over QANN.
- For larger setups it is useful to parametrize kernels using **Tensor Networks** (MPS).
- **Number of steps to perform FM**: $O(DN)/O(\log(DN))$ (Grover-Rudolph 2002), but not always [Phys. Rev. Lett. **107**, 120501 (2011)]



Gate depth for optical q. computing setups:
12 for 24 spatial modes (www.xanadu.ai)

- D. Garcia, F. Verstraete, M. M. Wolf, J. I. Cirac, *Quantum Info. Comput.* **7**, 401 (2007)
- Y. Liu, X. Zhang, M. Lewenstein, S.-J. Ran, arXiv:1803.09111 (2018)
- D. Liu, S.-J. Ran, P. Wittek, C. Peng, R. B. García, G. Su, M. Lewenstein, *New J. Phys.* **21,** 073059 (2019)
- Conference papers on ANNs

Unsupervised CQ ML: K-means clustering

JanWei Pan

Amplitude encoding & SWAP test:

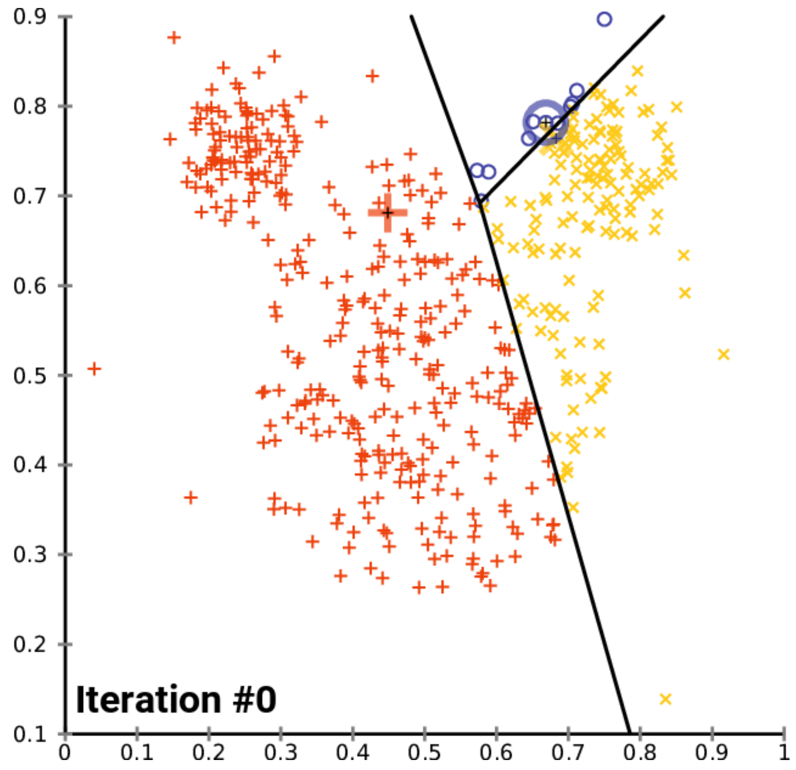
- X.-D. Cai et al., Phys. Rev. Lett. **114,** 110504 (2015)

Dense amplitude encoding & HSD:

- V. Trávníček, KB, A. Černoch, K. Lemr, Phys. Rev. Lett. **123**, 260501 (2019)

CQ

$$D_{HS}\left(\hat{\rho}_1, \hat{\rho}_2\right) \equiv \sqrt{\mathrm{Tr}[(\hat{\rho}_1 - \hat{\rho}_2)^2]}$$

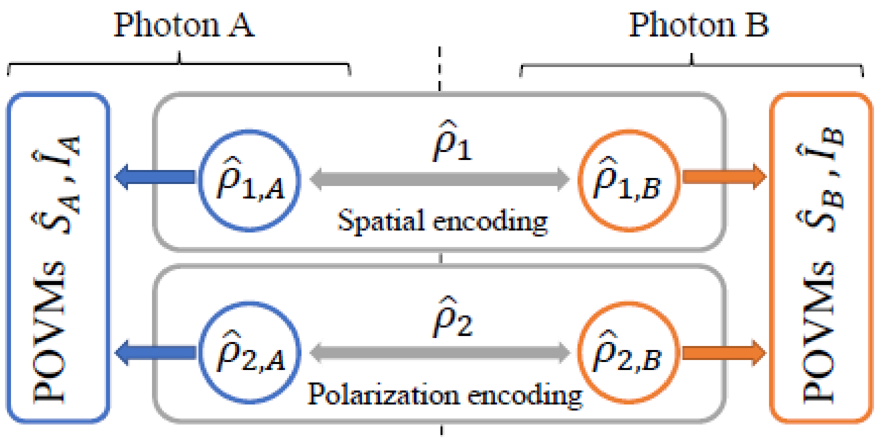


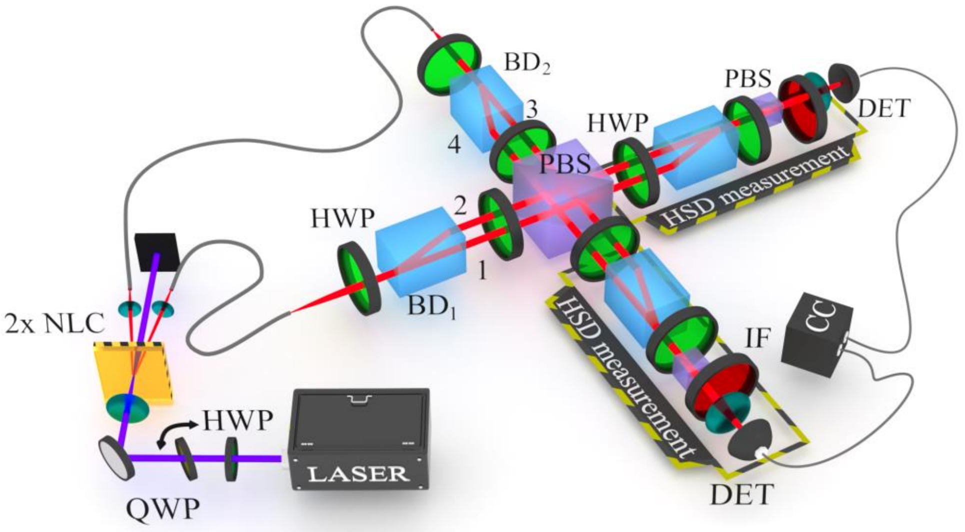
https://upload.wikimedia.org/wikipedia/commons/e/ea/K-means_convergence.gif

Unsupervised CQ ML : Distance measurement

$$D_{HS}(\hat{\rho}_1,\hat{\rho}_2)=\sqrt{\mathrm{Tr}[\hat{\rho}_1\hat{\rho}_1]+\mathrm{Tr}[\hat{\rho}_2\hat{\rho}_2]-2\mathrm{Tr}[\hat{\rho}_1\hat{\rho}_2]}$$

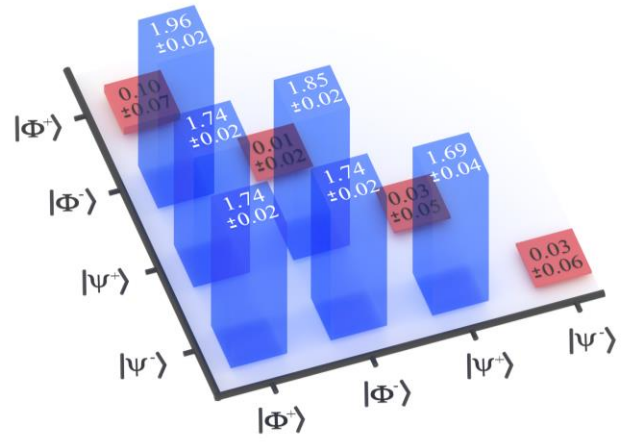
- 4 POVMs per observable for 2 qubits ($N=15$) (# linear in number of qubits)
- HS distance measurement: $O[\log(N=D^2-1)]$
- Classical distance computation: $O[\mathrm{Poly}(N)]$

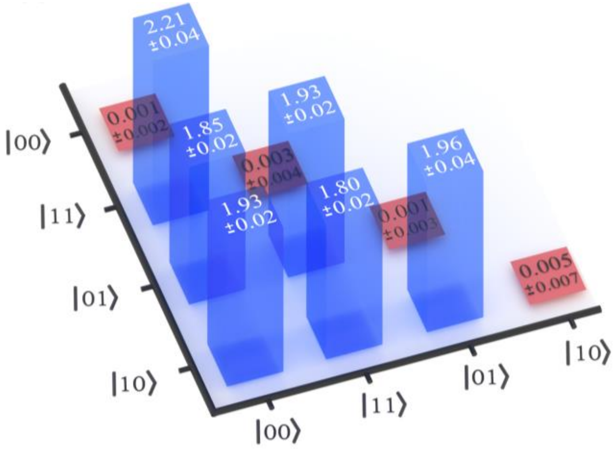


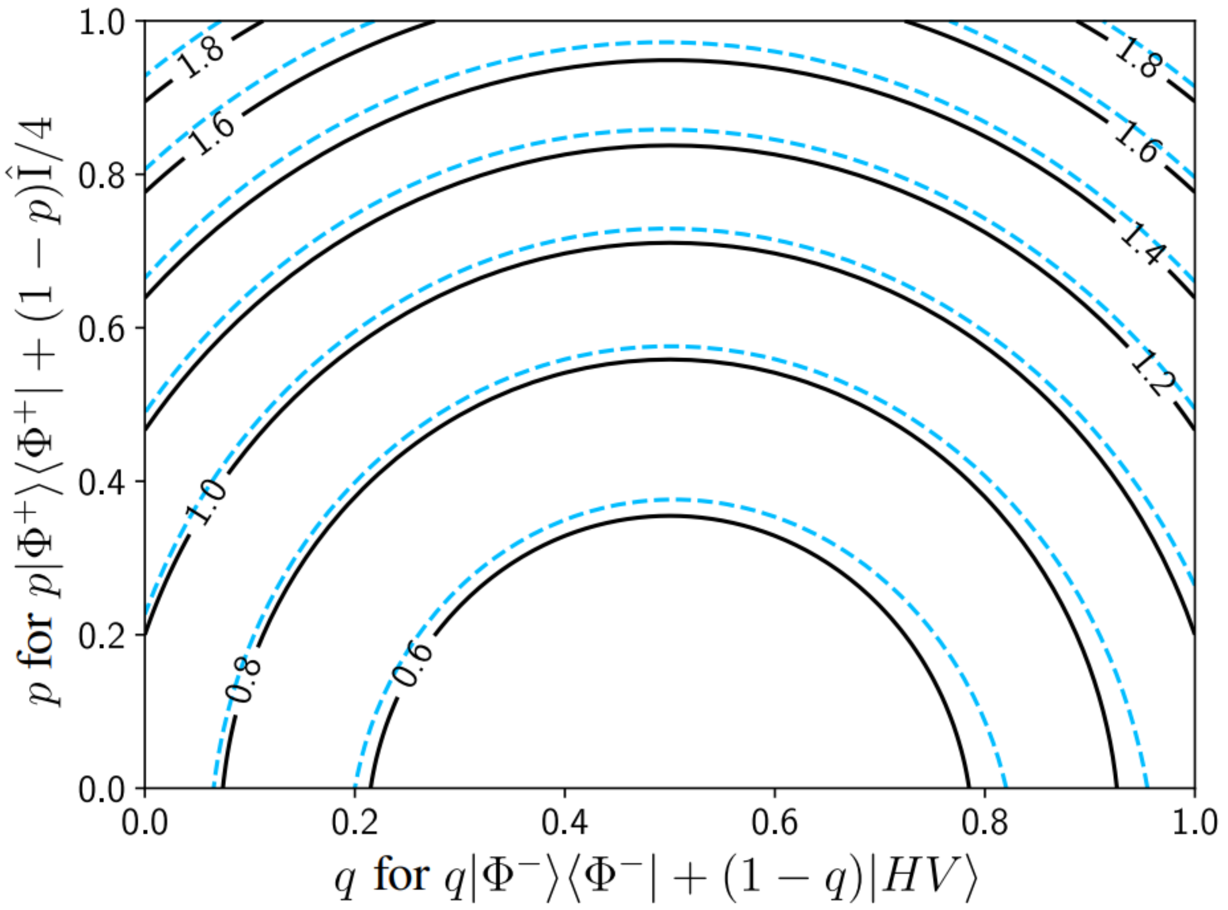


- ❖ **Proof of Strong Asymptotic Quantum Speedup**
- ❖ **Quadratic enhancement over the SWAP test**

Unsupervised CQ ML: Measured data







$$D_{HS}^2(\hat{\rho}_1, \hat{\rho}_2) = \mathrm{Tr}[\hat{\rho}_1\hat{\rho}_1] + \mathrm{Tr}[\hat{\rho}_2\hat{\rho}_2] - 2\mathrm{Tr}[\hat{\rho}_1\hat{\rho}_2]$$

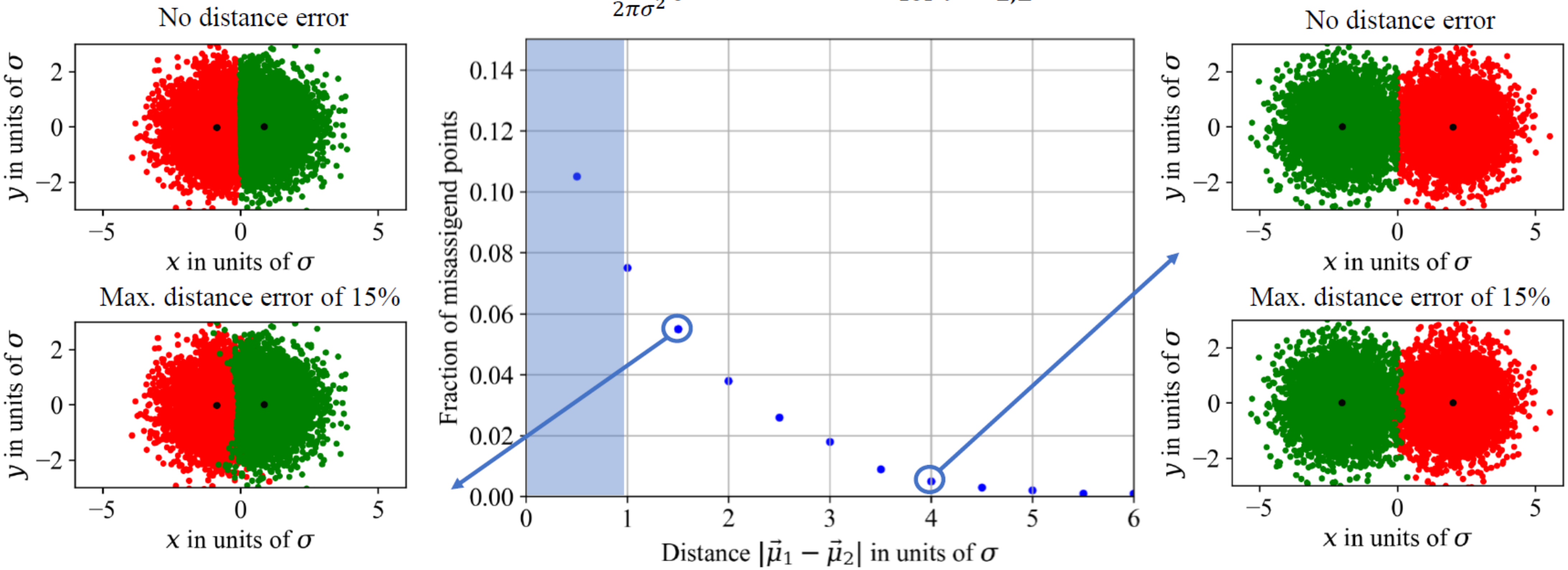
Unsupervised CQ ML: Error robustness

k-means results for $|\vec{\mu}_1 - \vec{\mu}_2| = 1.5$

Fraction of misassigned points for two sets of 10^4 points sampled from 2D symmetric Gaussians

$$\frac{1}{2\pi\sigma^2} e^{-\frac{(\mu_{ix}-x)^2+(\mu_{iy}-y)^2}{2\sigma^2}} \quad \text{for } i = 1,2$$

k-means results for $|\vec{\mu}_1 - \vec{\mu}_2| = 4$



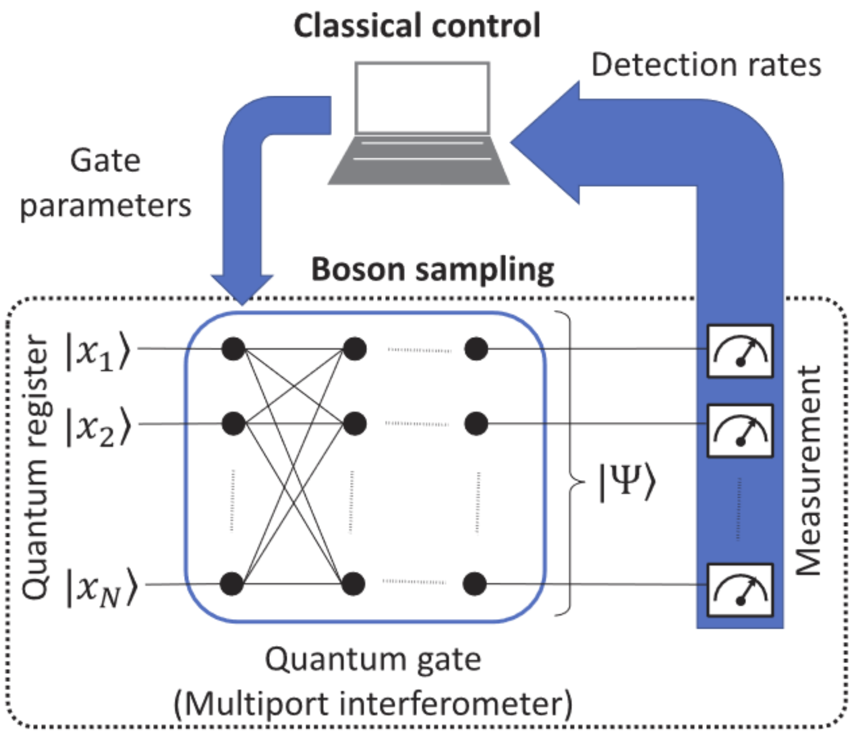
From CQ to QC

Reinforcement QC ML

- J. Jašek, K. Jiráková, KB, A. Černoch, T. Fürst, K. Lemr, Opt. Express, **27**, 32454 (2019)

Driving research in q. physics with deep ANNs

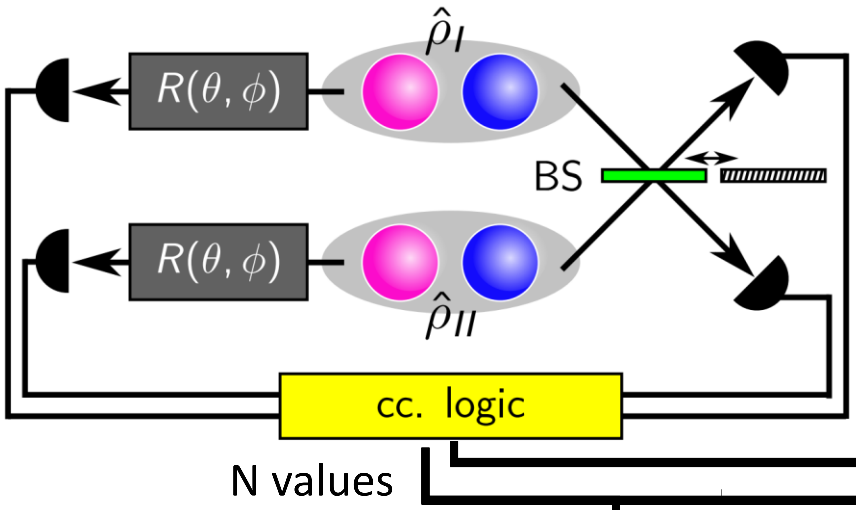
- J. Roik, KB, A. Černoch, K. Lemr, Phys. Rev. Applied, **15**, 054006 (2021)
- S. Ahmed, C. Munoz, F. Nori, A. Kockum, Phys. Rev. Research, in press (2021)



Variational quantum circuit

Driving research in q. physics with ANNs

Entanglement or no entanglement (without tomography)



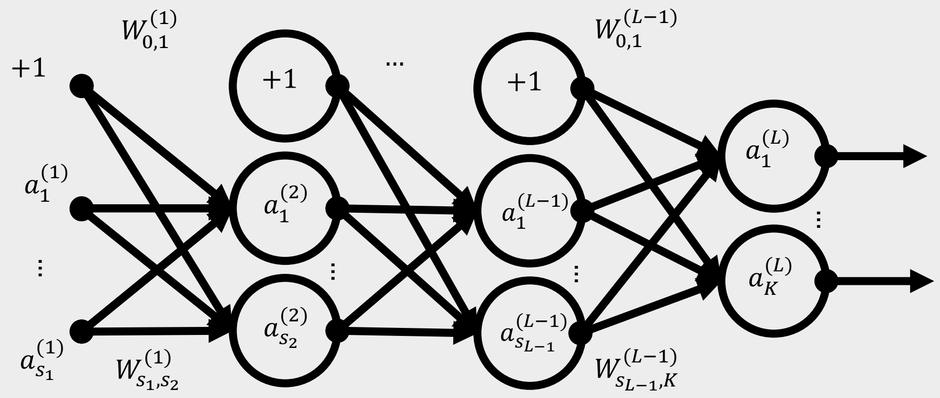
We can learn something from the ANN!

ANN
N = 3
Type-I error: ~1%
Type-II error: 33%
Success rate: 66%

ANN
N = 5
Type-I error: ~1%
Type-II error: 21%
Success rate: 78%

Collectibility
N = 5
Type-I error: 0%
Type-II error: 63%
Success rate: 37%
PRL 107, 150502 (2011)

ANN



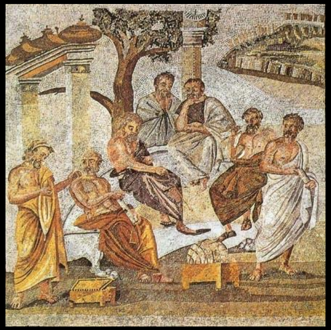
N = 12
Type-I error: ~0%
Type-II error: 8%
Success rate: 92%

Fully Ent. Frac.
N = 12
Type-I error: 0%
Type-II error: 14%
Success rate: 86%
PRA 54, 3824 (1996)

Lin. Entropy
N = 12
Type-I error: 0%
Type-II error: 42%
Success rate: 58%
PRL 95, 240407 (2005)

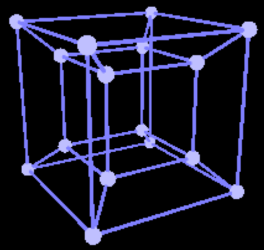
CHSH
N = 12
Type-I error: 0%
Type-II error: 54%
Success rate: 46%
PLA 223, 1 (1996)

Group meetings during the pandemic



Plato's Academy mosaic — from the Villa of T. Siminius Stephanus in Pompeii.

Athens c. 387 BC



Funding

Kernel based quantum machine learning in optical circuits, GAČR (CZ) 19-19002S

Fundamental problems and implementations of dissipative quantum engineering, NCN (PL) 2019/34/A/ST2/00081

Conclusions

C Q

Exponential kernel FM resolution enhancement in comparison to standard approach

Many q. kernels can be evaluated much faster (**Grover-Rudolph-like scaling instead of polynomial complexity**)

C Q

Strong asymptotic quantum speedup for HSD-based CQ ML (from polynomial to logarithmic)

Quadratic enhancement better than Swap Test thanks to dissipative state preparation

Q C

We combined **boson sampling and classical control** in our demonstration of reinforcement ML applied to VQC

Deep ANN demonstrated classification of quantum states can be improved

http://bark.home.amu.edu.pl/qmlg.html

https://dml.riken.jp/pub/ai_meets_qp/

Recruiting for a postdoc in QML!