

Quantum machine learning with qudits

Sebastián Roca-Jerat

Juan Román-Roche, Fernando Luis, David Zueco

INMA - CSIC

31-05-23



Outline

- I. Introduction of single qudit processors**
- II. Learning with qudits
- III. Results
- IV. Conclusions

Introduction of single qudit processors

PHYSICAL REVIEW APPLIED 17, 064030 (2022)

Dispersive Readout of Molecular Spin Qudits

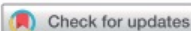
Álvaro Gómez-León^{1,*}, Fernando Luis² and David Zueco²

Chemical
Science



EDGE ARTICLE

View Article Online
View Journal | View Issue



Cite this: Chem. Sci., 2021, 12, 5621

All publication charges for this article have been paid for by the Royal Society of Chemistry

Broad-band spectroscopy of a vanadyl porphyrin: a model electronuclear spin qudit†

Ignacio Gimeno^{1,a}, Ainhoa Urtizberea^{1,ab}, Juan Román-Roche^{1,a}, David Zueco^{1,a}, Agustín Camón^{1,a}, Pablo J. Alonso^{1,a}, Olivier Roubeau^{1,*a} and Fernando Luis^{1,*a}

Accepted Paper

Featured in Physics

Blueprint for a molecular-spin quantum processor

Phys. Rev. Applied

A. Chiesa, S. Roca, S. Chicco, M. C. de Ory, A. Gómez-León, A. Gomez, D. Zueco, F. Luis, and S. Carretta

Accepted 24 April 2023

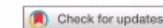
PHYSICAL REVIEW APPLIED 17, 064028 (2022)

Optimal Control of Molecular Spin Qudits

Alberto Castro^{1,2,*}, Adrián García Carrizo², Sebastián Roca^{3,4}, David Zueco^{3,4} and Fernando Luis^{3,4}

communications
physics

ARTICLE



<https://doi.org/10.1038/s42005-022-01017-8>

OPEN

High cooperativity coupling to nuclear spins on a circuit quantum electrodynamics architecture

Victor Rollano^{1,2}, Marina C. de Ory³, Christian D. Buch⁴, Marcos Rubín-Osanz^{1,2}, David Zueco^{1,2}, Carlos Sánchez-Azqueta⁵, Alessandro Chiesa^{6,7,8}, Daniel Granados⁹, Stefano Carretta^{6,7,8}, Alicia Gomez³, Stergios Piligkos⁴ & Fernando Luis^{1,2}

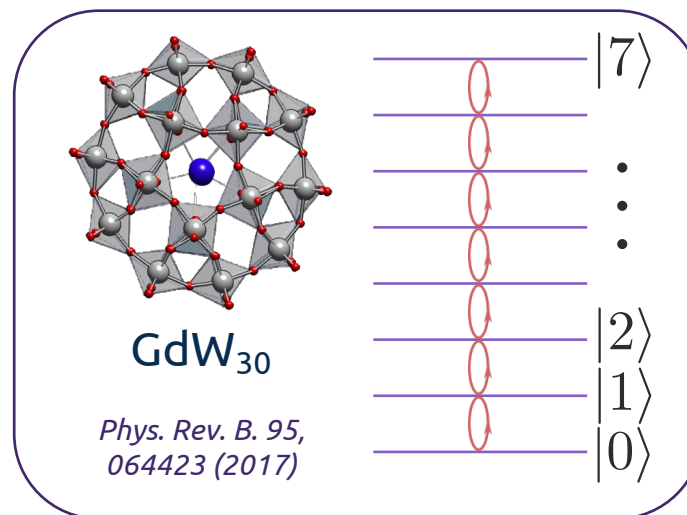


Quantum
Materials and Devices

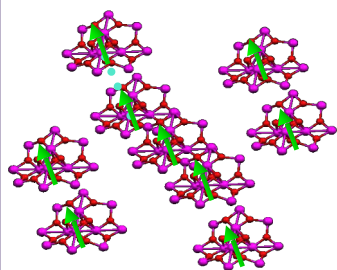
Introduction of single qudit processors

Quantum computing

d-level system → Qudit!



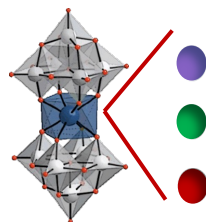
Reproducibility



Identical
microscopic
qudits

Nature chemistry 11 (4), 301-309 (2019)

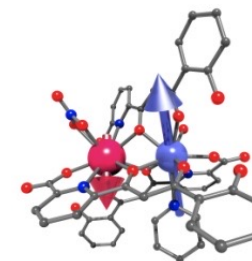
Tunability



Fine tuning
of qudit
properties

Phys. Rev. Lett. 108, 247213 (2012)

Scalability



Molecular
NISQs

Appl. Phys. Lett. 118, 240501 (2021)

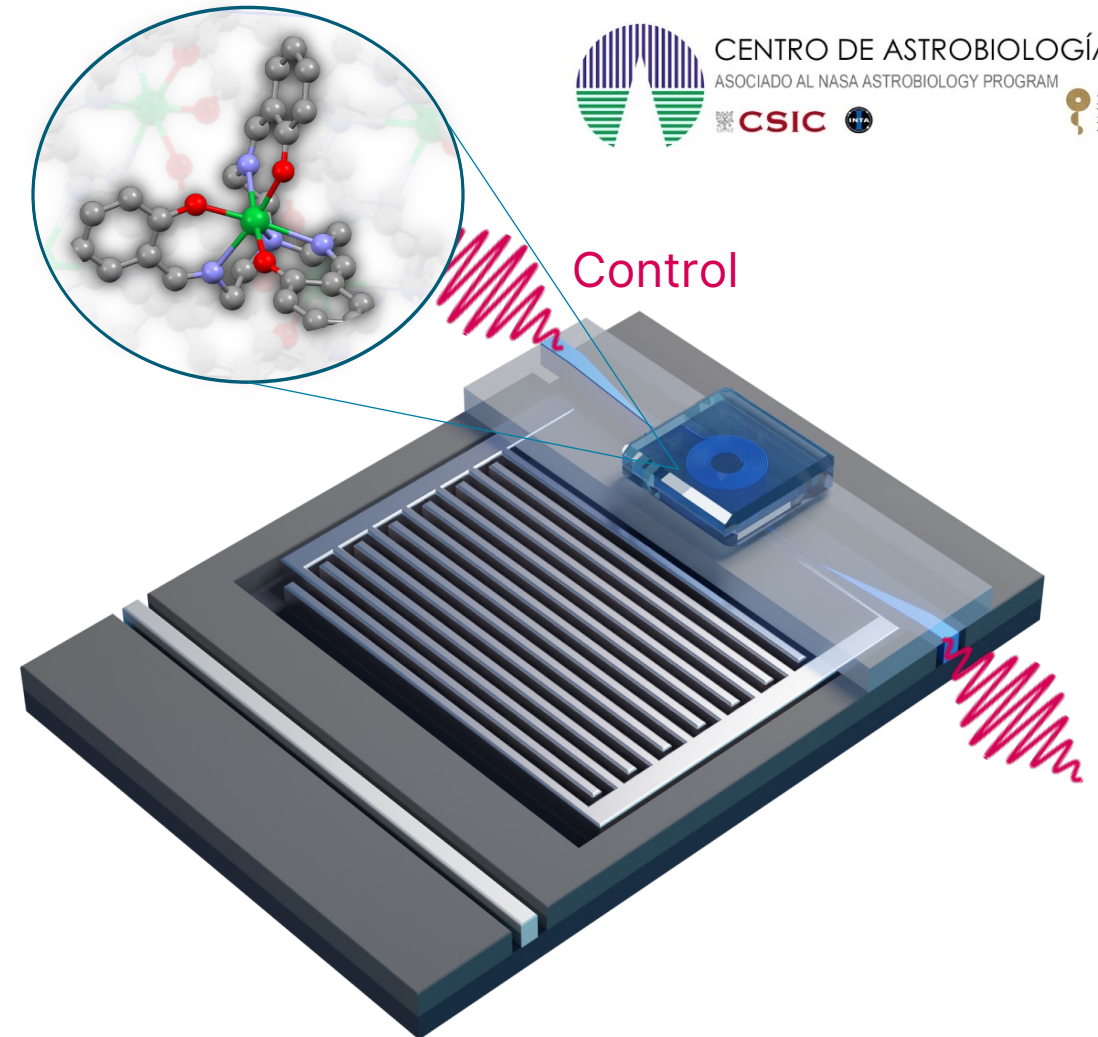
Introduction of single qudit processors

- Coupling molecules to superconducting circuits: **Lumped Element Resonators**

Tunable resonant frequency to match spin levels

$$\omega_r = \frac{1}{\sqrt{LC}}$$

- Coupling via microwave electromagnetic fields



CENTRO DE ASTROBIOLOGÍA · CAB
ASOCIADO AL NASA ASTROBIOLOGY PROGRAM



EXCELENCIA
MARIA
DE MAEZTU

M. Jenkins *et al.* *NJP* 15, 095007 (2013)

V. Rollano *et al.* *Comm. Phys.* 5, 246 (2022)



Outline

I. Introduction of single qudit processors

II. Learning with qudits

III. Results

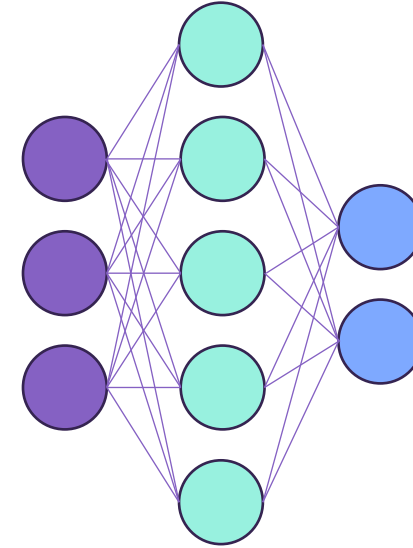
IV. Conclusions

Learning with qudits

Motivation: supervised learning tasks

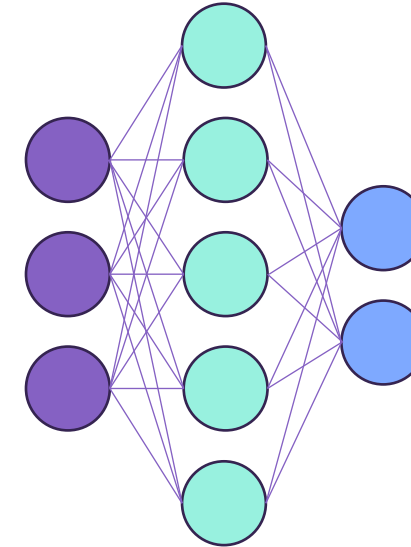
- Digital paradigm

→ Usual neural network models, random forests, Xgboost...



Learning with qudits

Motivation: supervised learning tasks

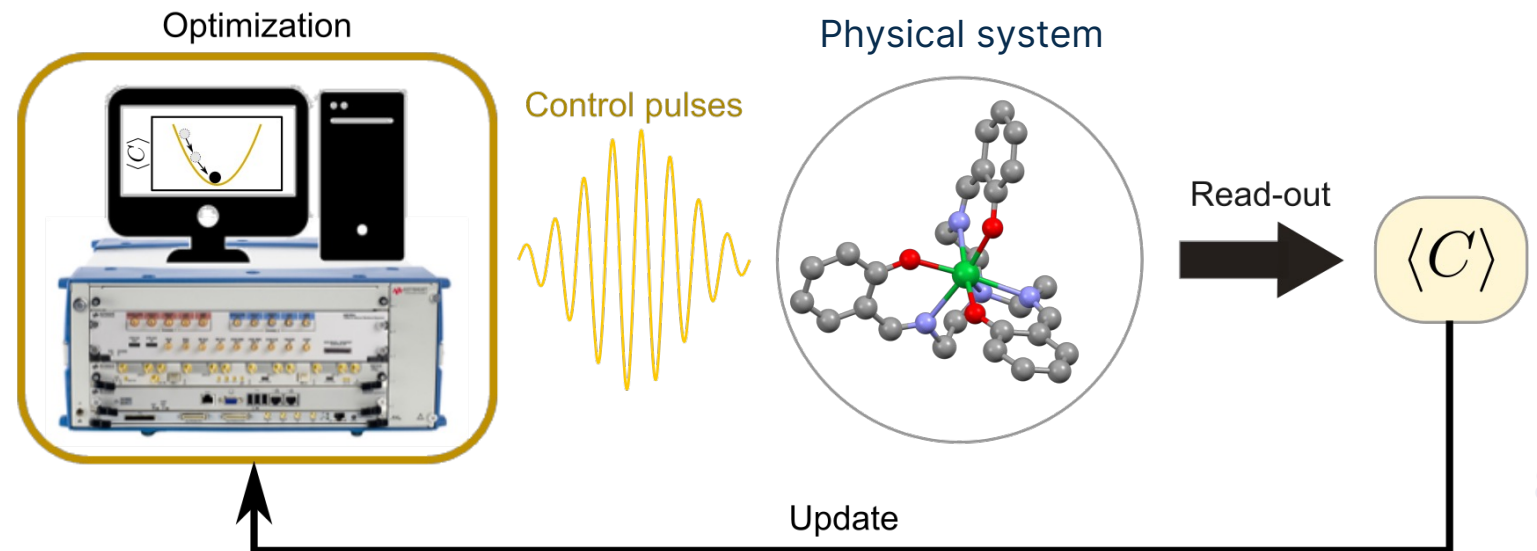


- Digital paradigm

→ Usual neural network models, random forests, Xgboost...

- Digital-analog perspective

→ Ising machines, reservoir computing, etc



Learning with qudits

Motivation: supervised learning tasks

Data re-uploading for a universal quantum classifier

Adrián Pérez-Salinas^{1,2}, Alba Cervera-Lierta^{1,2}, Elies Gil-Fuster³, and José I. Latorre^{1,2,4,5}

¹Barcelona Supercomputing Center

²Institut de Ciències del Cosmos, Universitat de Barcelona, Barcelona, Spain

³Dept. Física Quàntica i Astrofísica, Universitat de Barcelona, Barcelona, Spain.

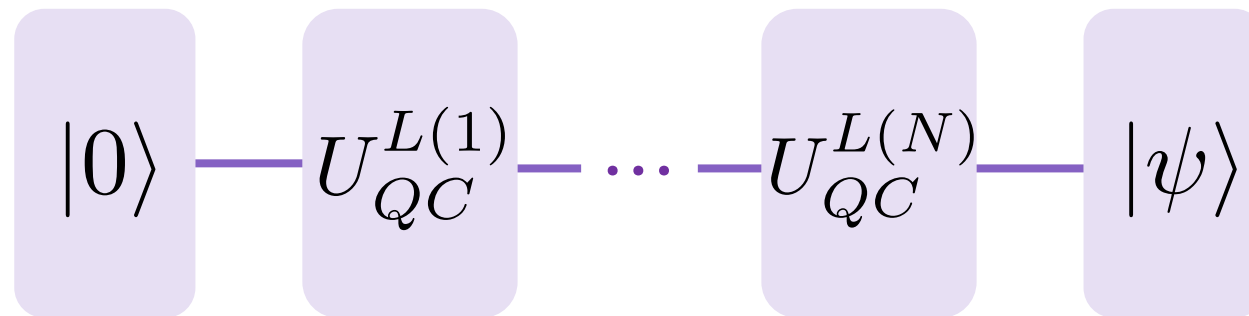
⁴Nikhef Theory Group, Science Park 105, 1098 XG Amsterdam, The Netherlands.

⁵Center for Quantum Technologies, National University of Singapore, Singapore.

Quantum 4, 226 (2020)

$$\mathcal{L}(\vec{x}; \vec{\varphi}) = \sum_{n=1}^N (1 - |\langle \psi(\vec{x}_n; \vec{\varphi}_n) | \psi_n^R \rangle|^2)$$

$$|\psi\rangle = \prod_{l=1}^N U_{QC}^l(\vec{x}, \vec{\phi}_l) |\psi_0\rangle \quad \max \{ |\langle \psi | \psi_i^R \rangle| \} \rightarrow \text{class}$$



Layers to increase non-linearity in the model

Learning with qudits

Motivation: supervised learning tasks

- Generalization to more than two levels?
Any advantage? Any new paradigm?

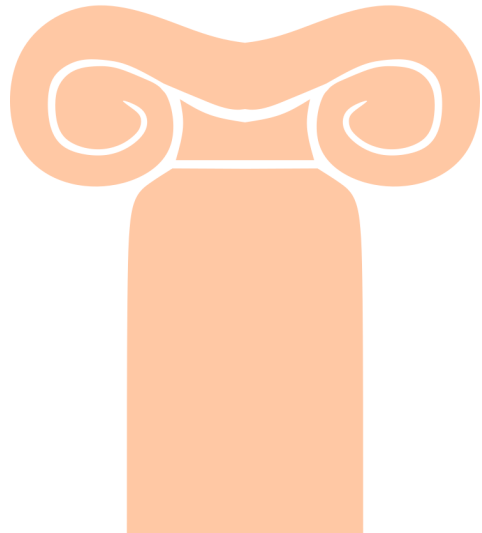
Learning with qudits

Motivation: supervised learning tasks

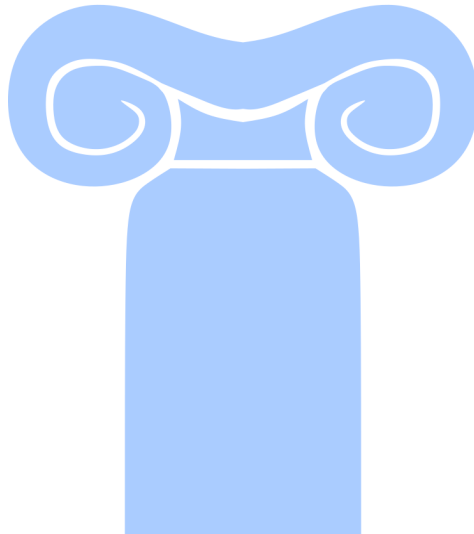
- Generalization to more than two levels?
Any advantage? Any new paradigm?
- What is the model really “learning”?

Learning with qudits

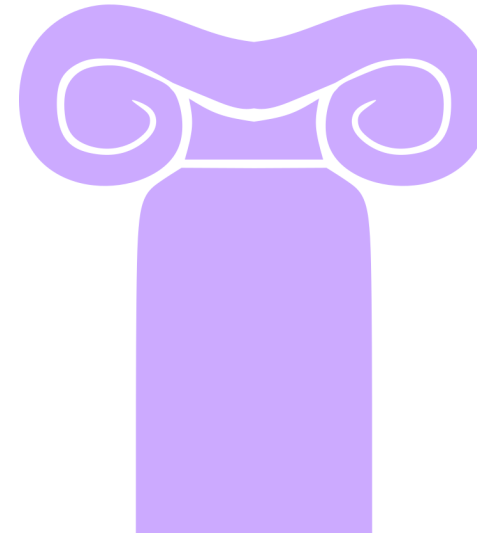
Variational
Algorithms



Encoding
strategy



Metric
Learning



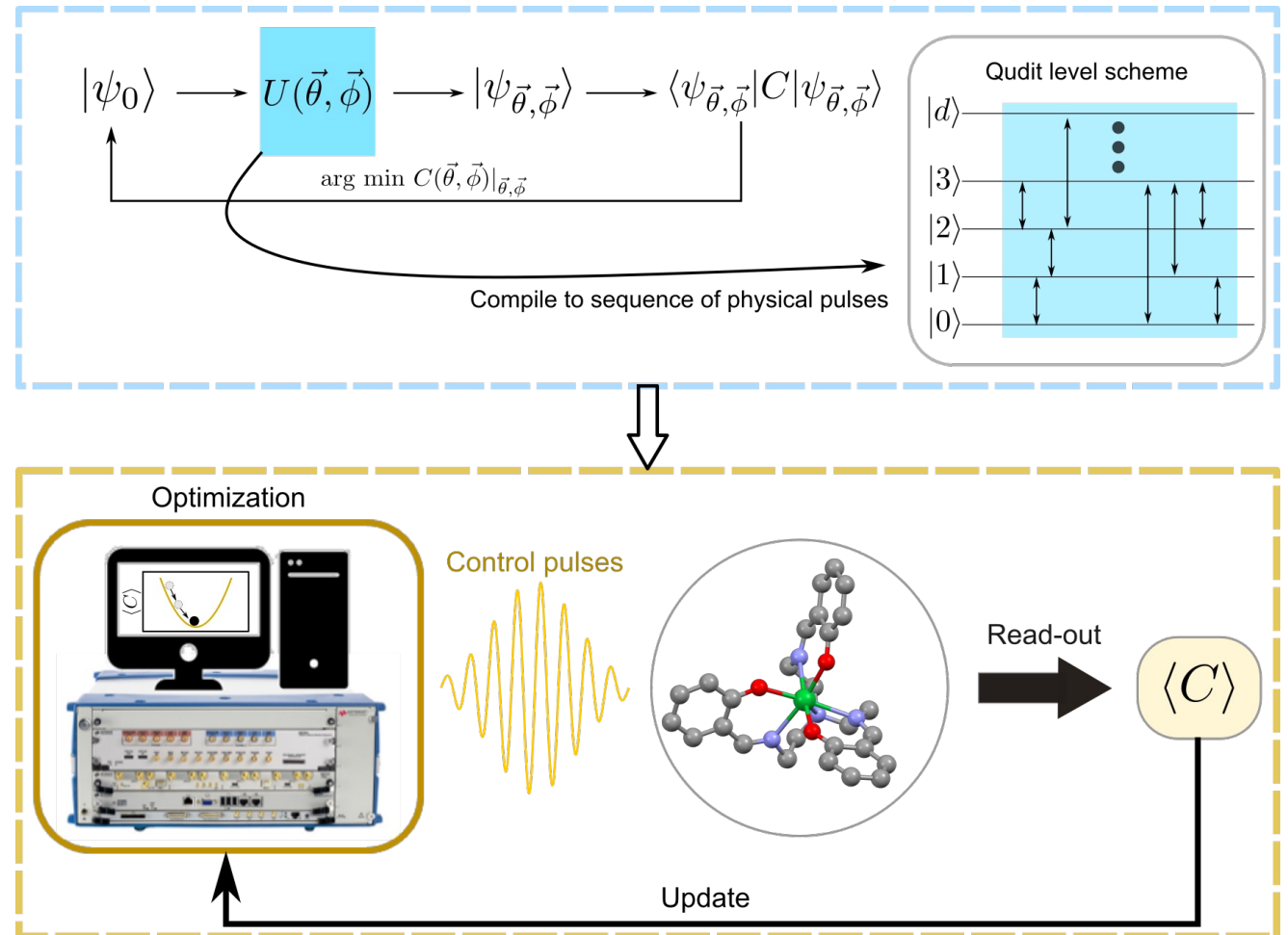
Learning with qudits

Variational algorithms

➤ Hybrid quantum-classical algorithm:
classical optimizer + quantum processor

➤ Finding ground states, dynamical simulations, error correction, machine learning...

A. Peruzzo *et al.* *Nat Comm* 5, 4213 (2014)
 A. Kandala *et al.* *Nature* 549, 242-246 (2017)
 A. Chiesa *et al.* *Nat Phys* 15, 455-459 (2019)
 M. Cerezo *et al.* *Nat. Rev. Phys.* 3, 625-644 (2021)

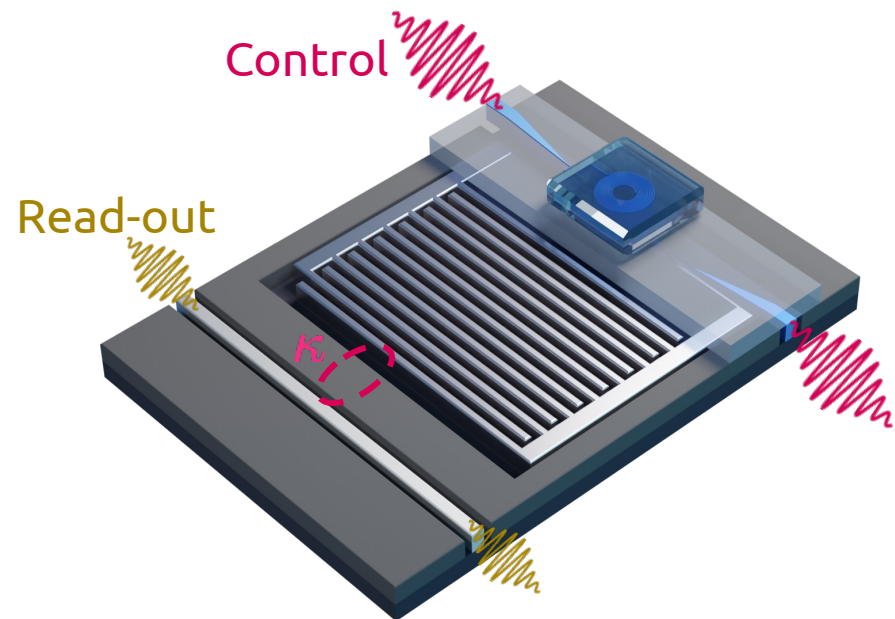


Learning with qudits

Variational algorithms: ansatz

Our *ansatz*: monochromatic RF pulses \rightarrow Rotations in the XY-plane of the two levels involved

$$\mathcal{H}_d = -g\mu_B \cos(\omega t + \phi) \vec{b}_{\text{rf}} \cdot \vec{S}$$

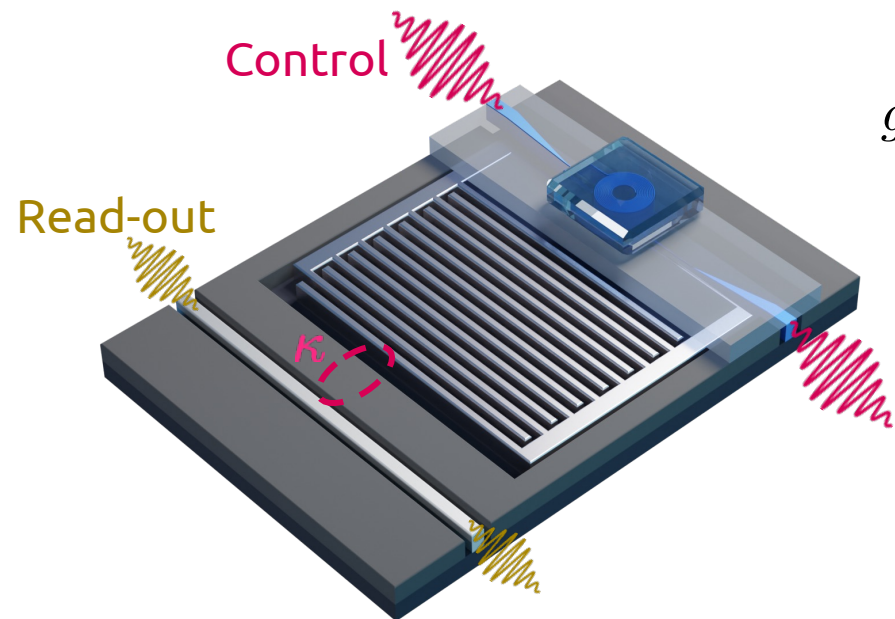


Learning with qudits

Variational algorithms: ansatz

Our *ansatz*: monochromatic RF pulses \rightarrow Rotations in the XY-plane of the two levels involved

$$\mathcal{H}_d = -g\mu_B \cos(\omega t + \phi) \vec{b}_{\text{rf}} \cdot \vec{S} \longrightarrow R_{j,k}(\theta, \varphi) = g_{j,k}(\theta, \varphi) \oplus \mathbb{I}_{\overline{jk}}$$



$$g_{j,k}(\theta, \varphi) = \begin{pmatrix} \cos(\theta/2) & -i \sin(\theta/2) e^{i\varphi} \\ -i \sin(\theta/2) e^{-i\varphi} & \cos(\theta/2) \end{pmatrix}$$

$$\theta = g\mu_B |M_{j,k}| t$$

$$\varphi = \arg M_{j,k} + \phi$$

$$M_{j,k} = \langle j | \vec{b}_{\text{rf}} \cdot \vec{S} | k \rangle$$

Learning with qudits

Encoding strategy

Our *ansatz*: monochromatic RF pulses \rightarrow Rotations in the XY-plane of the two levels involved

$$\mathcal{H}_d = -g\mu_B \cos(\omega t + \phi) \vec{b}_{\text{rf}} \cdot \vec{S} \longrightarrow R_{j,k}(\theta, \varphi) = g_{j,k}(\theta, \varphi) \oplus \mathbb{I}_{\overline{jk}}$$

We have $2(d-1)$ parameters to tune in a qudit \rightarrow data of dimension $2(d-1)$

However, we can encode higher dimensional data by dividing each point into n vectors of dimension $2(d-1)$:

$$U_{QC}^{L(i)} = \prod_{j=1}^n \prod_{k=0}^{d-2} R_{k,k+1}(\theta_k^{(j)}, \varphi_k^{(j)}) \quad (\vec{\theta}, \vec{\varphi}) = \vec{f}(\vec{x}, \vec{\phi})$$

Learning with qudits

Encoding strategy

$$\mathcal{E} : \mathbb{R}^{\dim(\vec{x})} \rightarrow \mathcal{H}^d$$

$$\vec{x} \mapsto |\psi\rangle$$

$$g : \mathbb{R}^{D_x} \rightarrow \mathbb{R}^{D_{x'}}$$

$$\vec{x} \mapsto \vec{x}' = \bar{\omega}\vec{x} + \vec{b}$$

$$(\theta_i, \phi_i) = (x'_{2i}, x'_{2i+1})$$

$$|\psi(\vec{x}; \vec{\varphi})\rangle = \hat{U}(\vec{x}; \vec{\varphi})|\psi_0\rangle$$

Learning with qudits

Metric learning: implicit vs explicit

In supervised learning, one of the most powerful approaches is Metric Learning.

We want to map our data to a feature space where points belonging to the same class are near to each other while being as far as possible from points belonging to other classes: learning some metric.

Learning with qudits

Metric learning: implicit vs explicit

In supervised learning, one of the most powerful approaches is Metric Learning.

We want to map our data to a feature space where points belonging to the same class are near to each other while being as far as possible from points belonging to other classes: learning some metric.

We do so in QM! → Our data-points are mapped into the Hilbert space of our qudit.

We seek to map points belonging to the same class to a quantum state that is close to the reference state defined for that class and as far as possible (maximally orthogonal) to the other reference states.

V. Havlíček *et al.* *Nature* 567, 209-212 (2019)

M. Schuld & N. Killoran *Phys. Rev. Lett.* 122, 040504 (2019)

S. Lloyd *et al.* *Arxiv:2001.03622v2* (2022)

Learning with qudits

Metric learning: implicit vs explicit

In supervised learning, one of the most powerful approaches is Metric Learning.

We want to map our data to a feature space where points belonging to the same class are near to each other while being as far as possible from points belonging to other classes: learning some metric.

We do so in QM! → Our data-points are mapped into the Hilbert space of our qudit. We seek to map points belonging to the same class to a quantum state that is close to the reference state defined for that class and as far as possible (maximally orthogonal) to the other reference states.

V. Havlíček *et al.* *Nature* 567, 209-212 (2019)

M. Schuld & N. Killoran *Phys. Rev. Lett.* 122, 040504 (2019)

S. Lloyd *et al.* *Arxiv:2001.03622v2* (2022)

Implicit: you fix the reference states

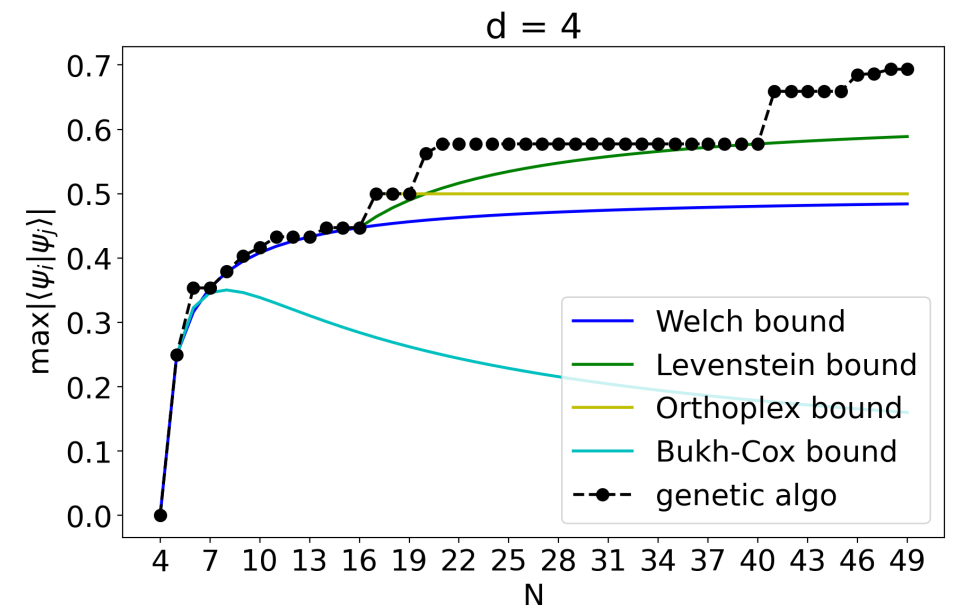
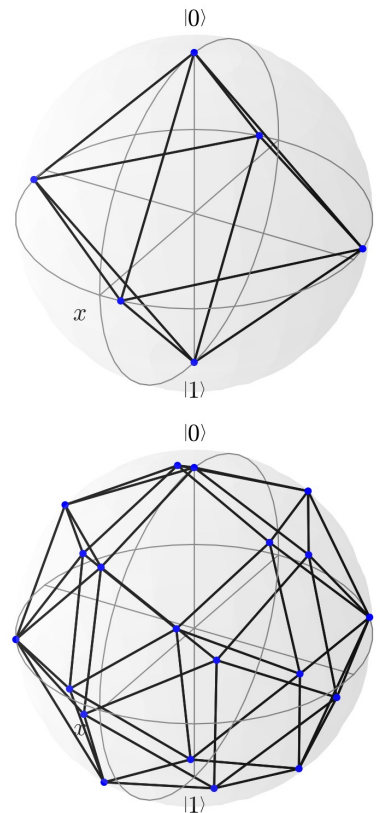
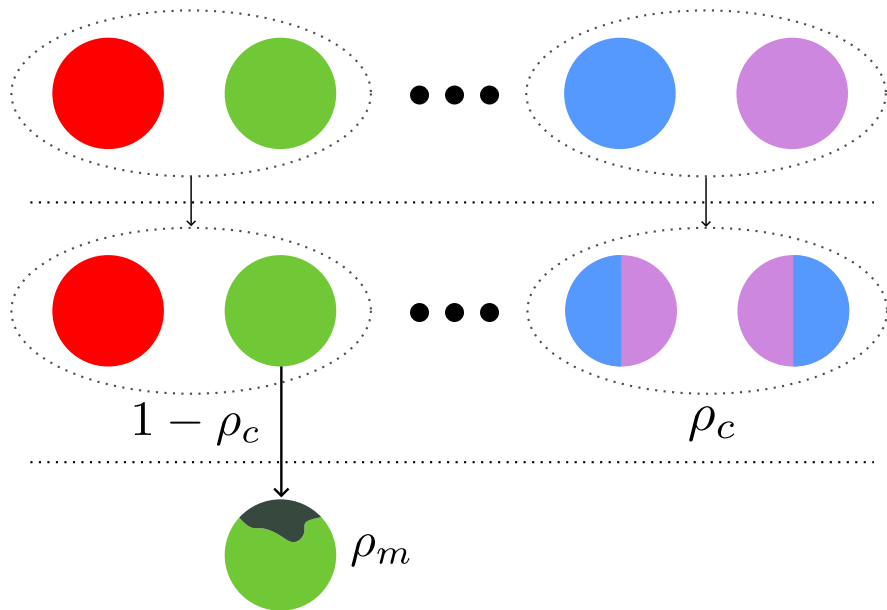
Explicit: you include these states as parameters to be found in the optimization process

N. A. Nghiem *et al.* *Phys. Rev. Res.* 3, 033056 (2021)

Learning with qudits

Maximally orthogonal states: More classes than levels?

Genetic algorithm for finding maximally orthogonal states for any configuration (d, N)



A. Pérez-Salinas *et al.* *Quantum* 4, 226 (2020)

S. Smale. *The mathematical intelligencer* 20.2 pp. 7-15 (1998)

J. R. Morris *et al.* *Phys. Rev. B* 53, R1740(R) (1996)

Learning with qudits

Metric learning: implicit vs explicit

$$\mathcal{L}' = 1 - \frac{T}{N} \sum_m \mathcal{F}(\sigma_m, \rho_m) \quad \mathcal{F}(\sigma, \rho) = \|\sqrt{\sigma}\sqrt{\rho}\|^2$$

$$\mathcal{F}(\sigma_m, \rho_m) \geq (1 - \epsilon) \quad \|\sigma_m - \rho_m\| \leq 2\sqrt{\epsilon}$$

$$\|\rho_m - \rho_{m'}\| \leq \|\rho_m - \sigma_m\| + \|\rho_{m'} - \sigma_{m'}\| + \|\sigma_{m'} - \sigma_m\|$$

$$\underline{\|\rho_m - \rho_{m'}\|} \leq 4\sqrt{\epsilon} + \underline{\|\sigma_{m'} - \sigma_m\|}$$

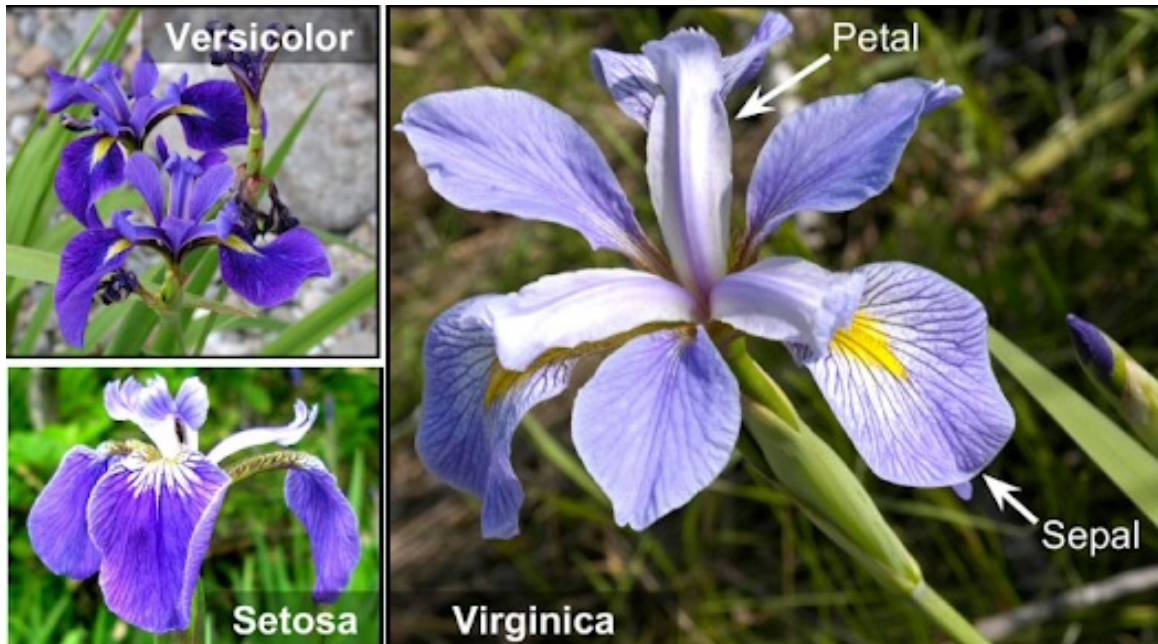
To maximize the distance between data-points ensembles we have to obtain the furthest possible centres:
Maximally orthogonal states!

Outline

- I. Introduction of single qudit processors
- II. Learning with qudits
- III. Results**
- IV. Conclusions

Results

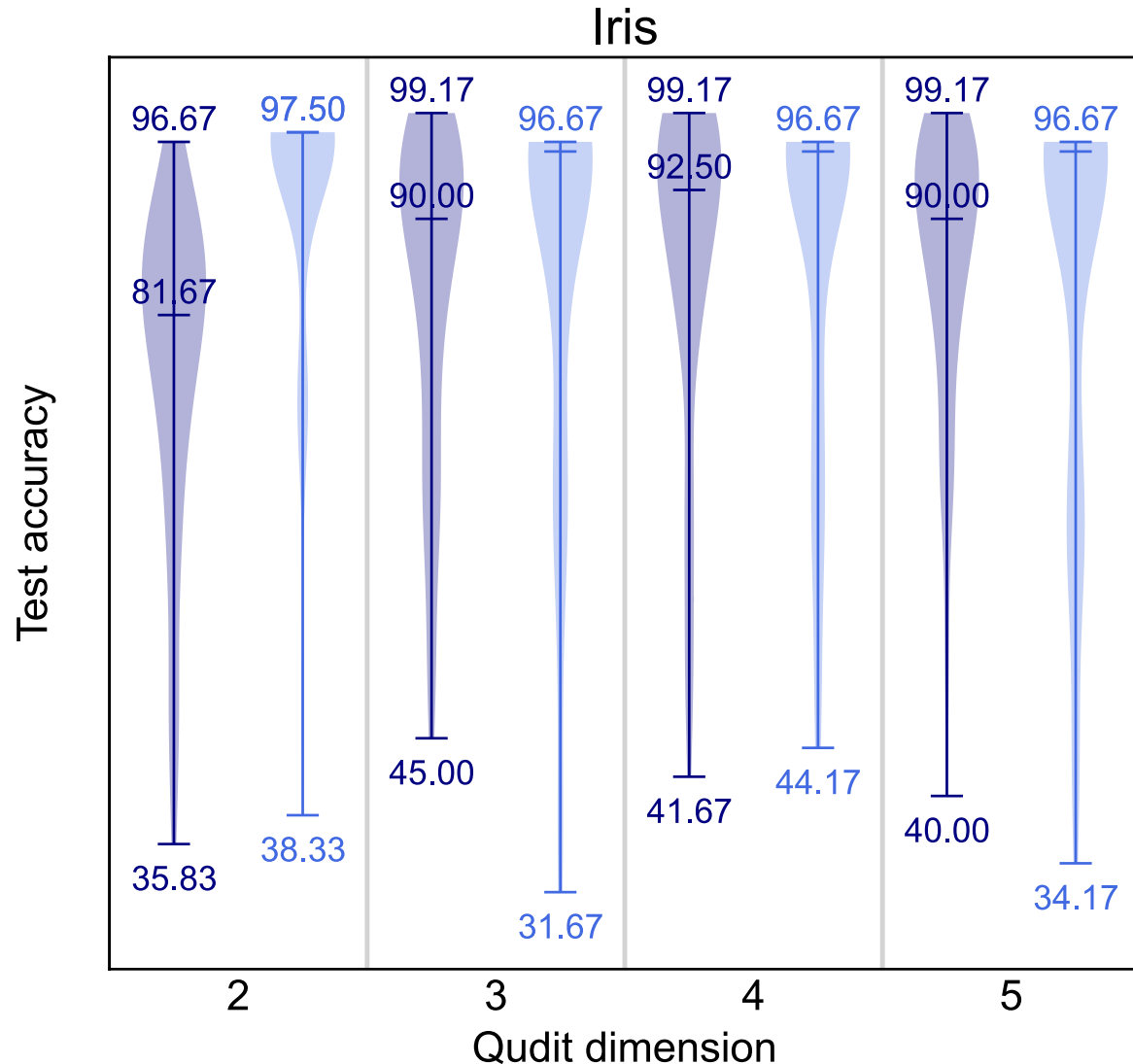
- Iris dataset: 4 features & 3 classes



- One of the simplest datasets that we can use to benchmark our model
- Data dimension and number of classes low

Results

○ Implicit vs explicit in Iris dataset



- The system is able to find better solutions by itself in some cases: higher accuracies for the implicit method
- Test accuracy saturates at $d = 3$

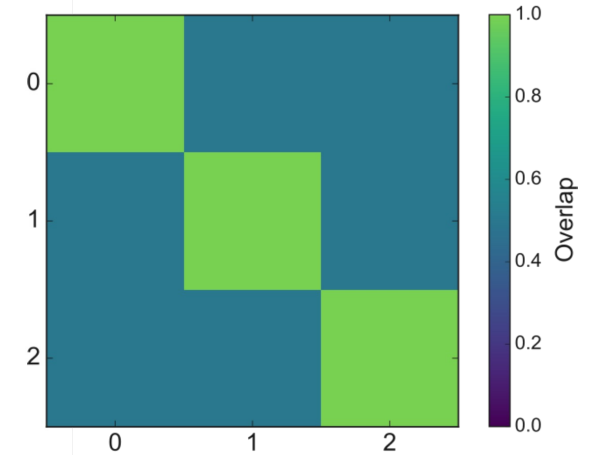
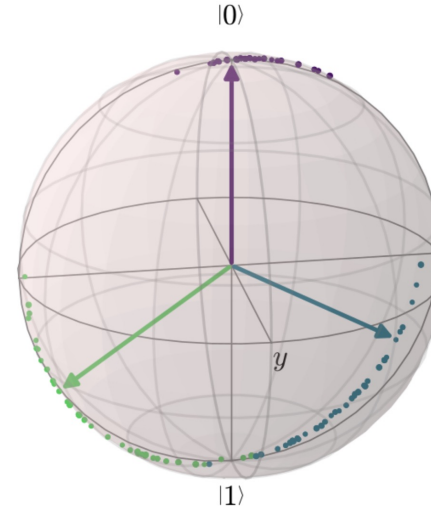
Implicit

Explicit

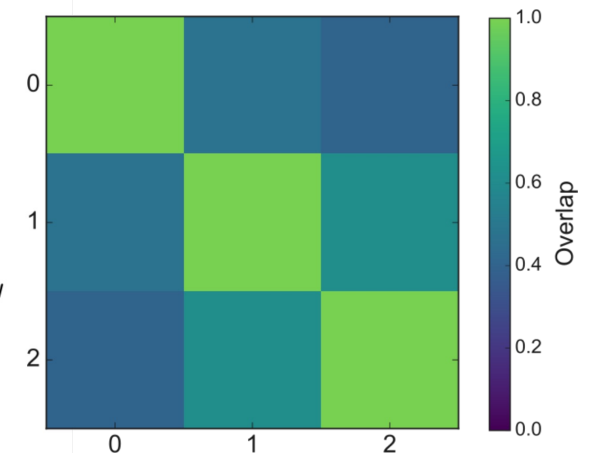
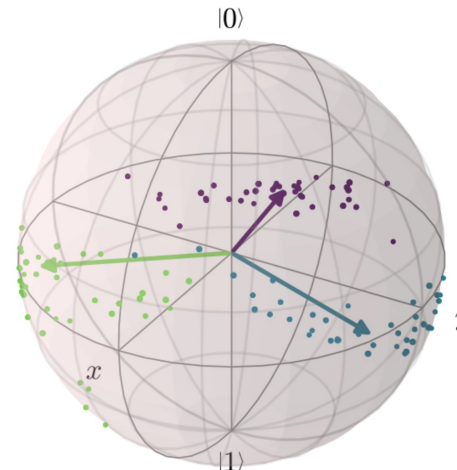
Results

Implicit vs explicit in Iris dataset

$$\mathcal{L}_{\text{explicit}} = 1 - \frac{1}{K} \sum_k \text{Tr}(\rho_k \sigma_k)$$

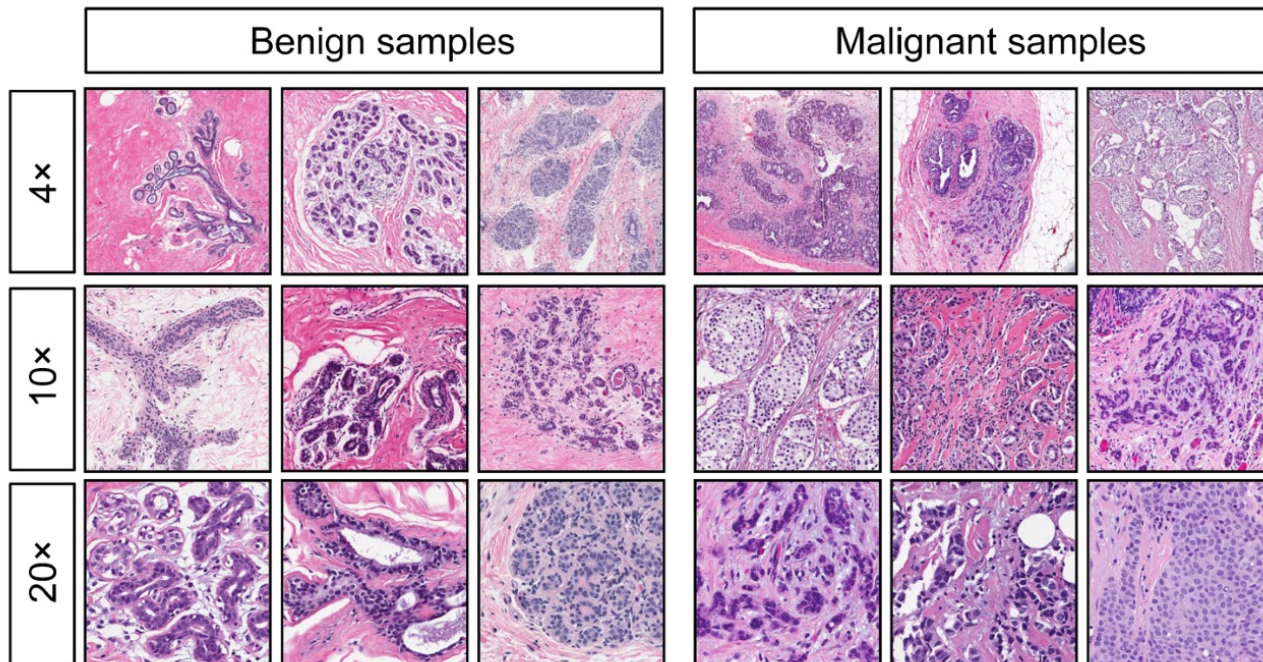


$$\mathcal{L}_{\text{implicit}} = 1 - \frac{1}{K} \sum_k \text{Tr}(\rho_k^2) + \frac{2}{K} \sum_{k < l} \text{Tr}(\rho_k \rho_l)$$



Results

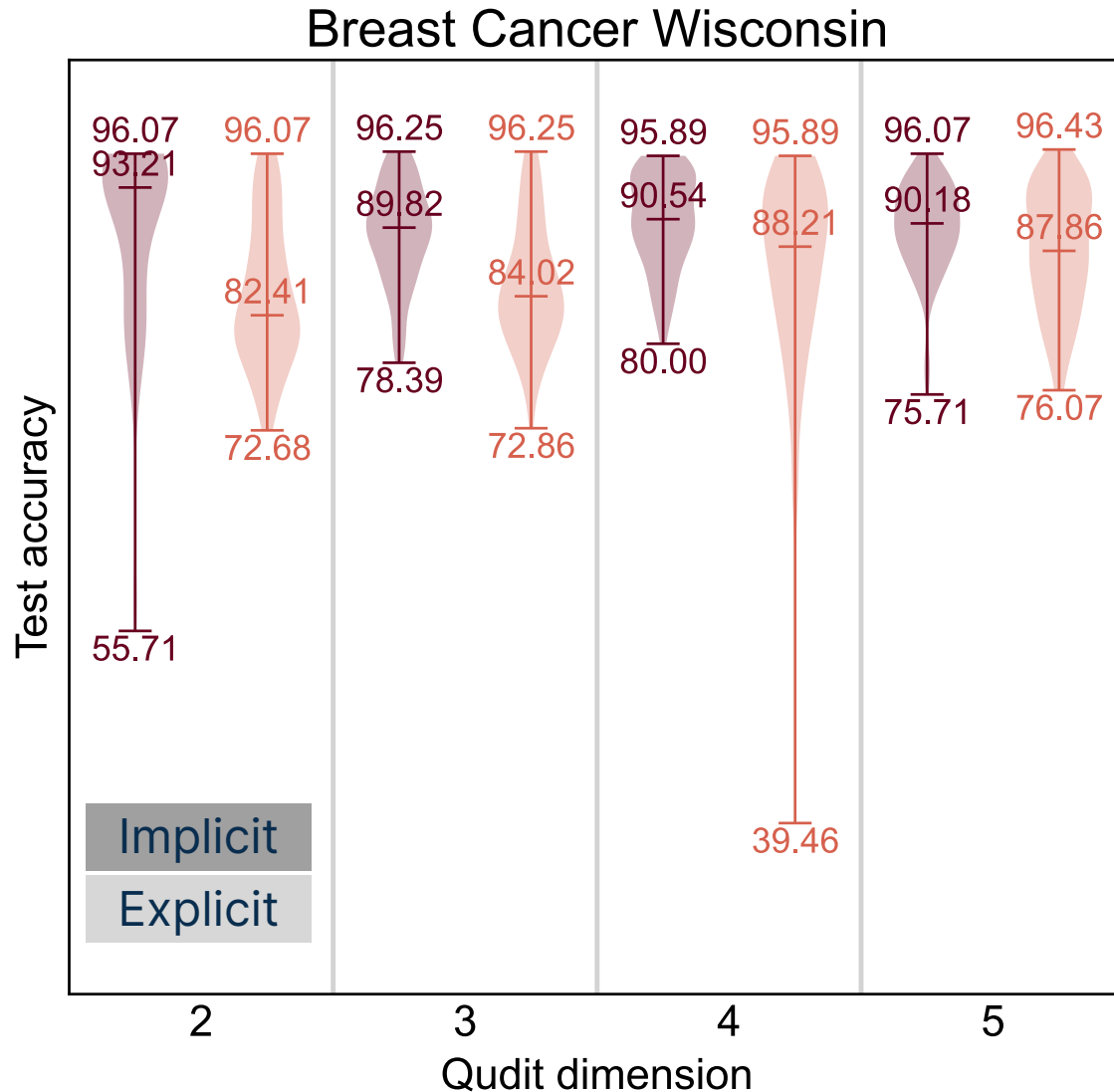
- Breastcancer Wisconsin dataset: 10 features & 2 classes



- More sophisticated problem in terms of number of points and features

Results

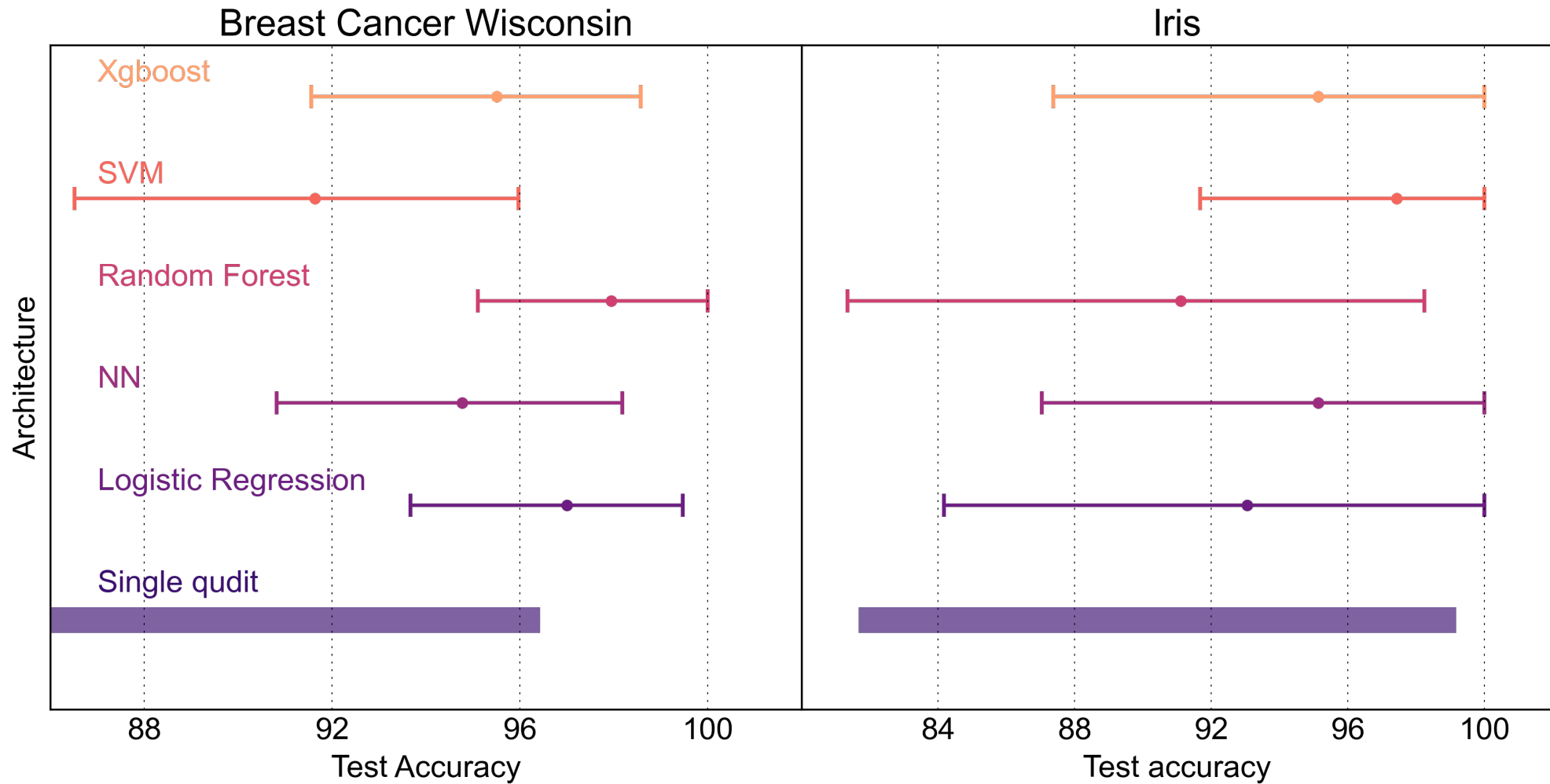
○ Implicit vs explicit in Breast Cancer dataset



- Lower accuracies than for the Iris dataset
- Both methods offer similar performance
- Qudit dimension does not play a crucial role

Results

Comparison with best classical models

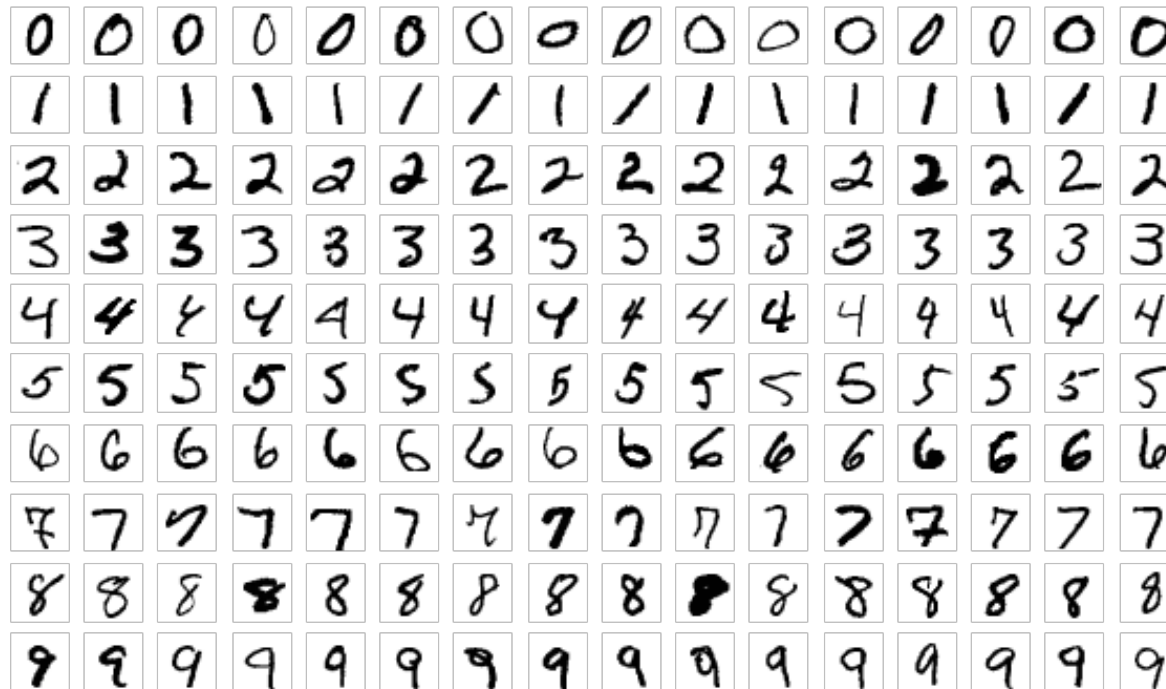


- What happens when the data dimension is much higher than the qudit dimension?

$$\text{Dim}(\vec{x}) \gg d$$

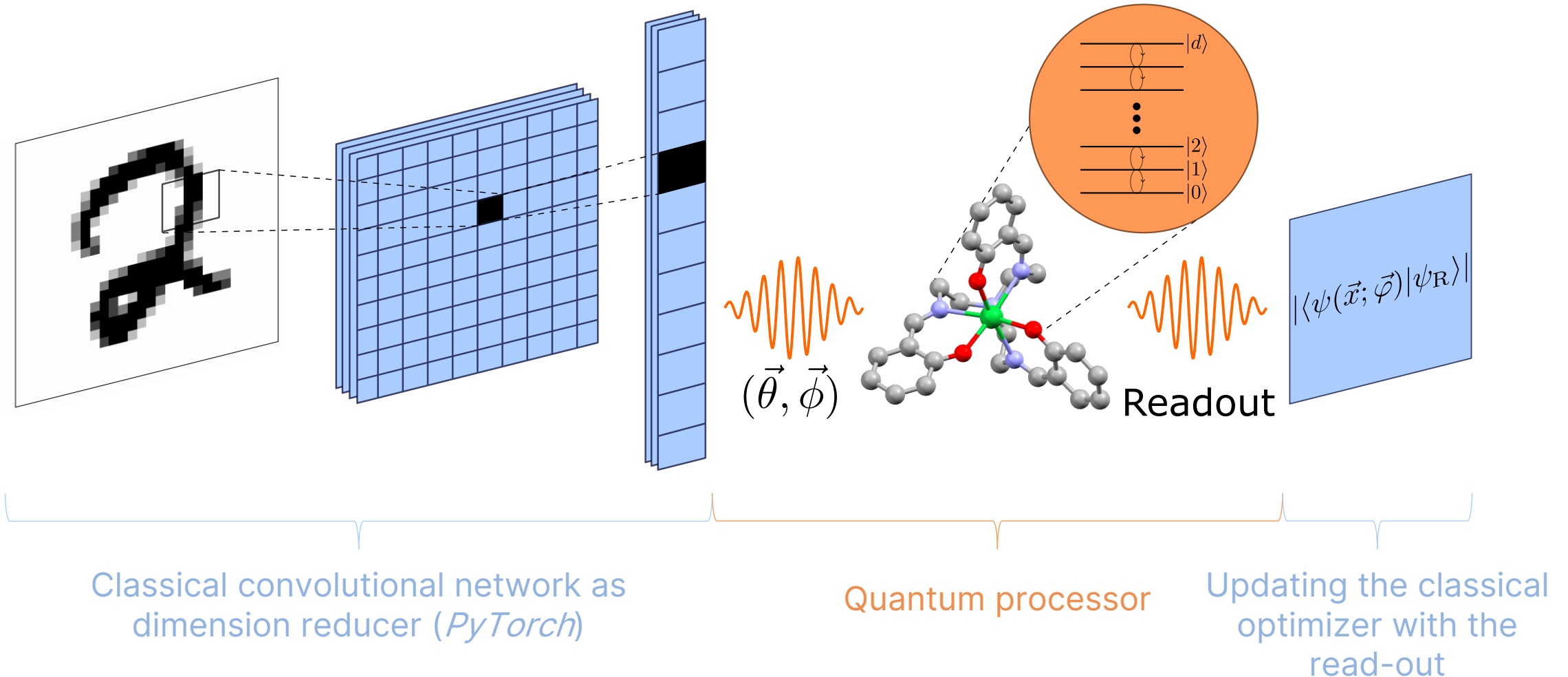
Results

- $\text{Dim}(\vec{x}) \gg d$
 - Image classification: MNIST digits 28x28 features and 10 classes



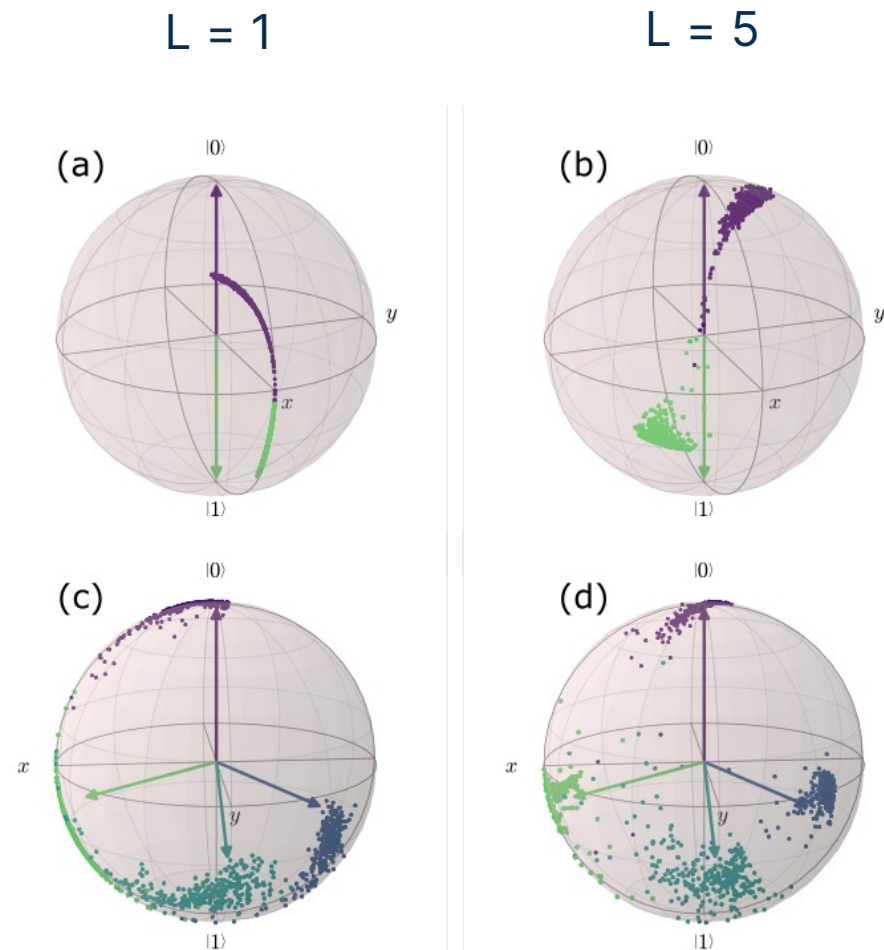
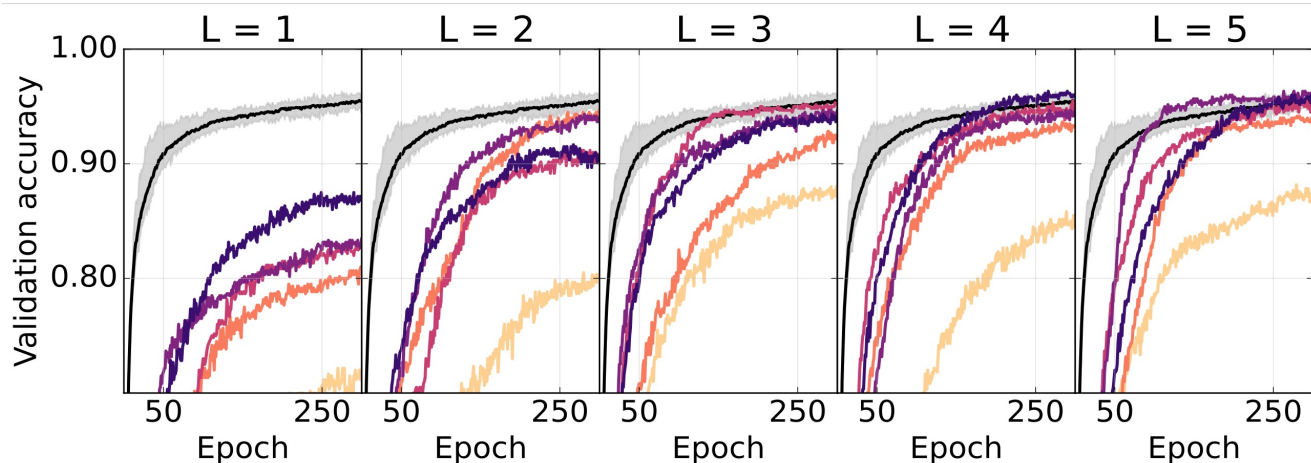
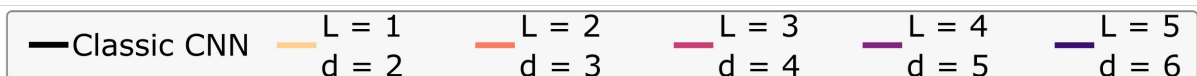
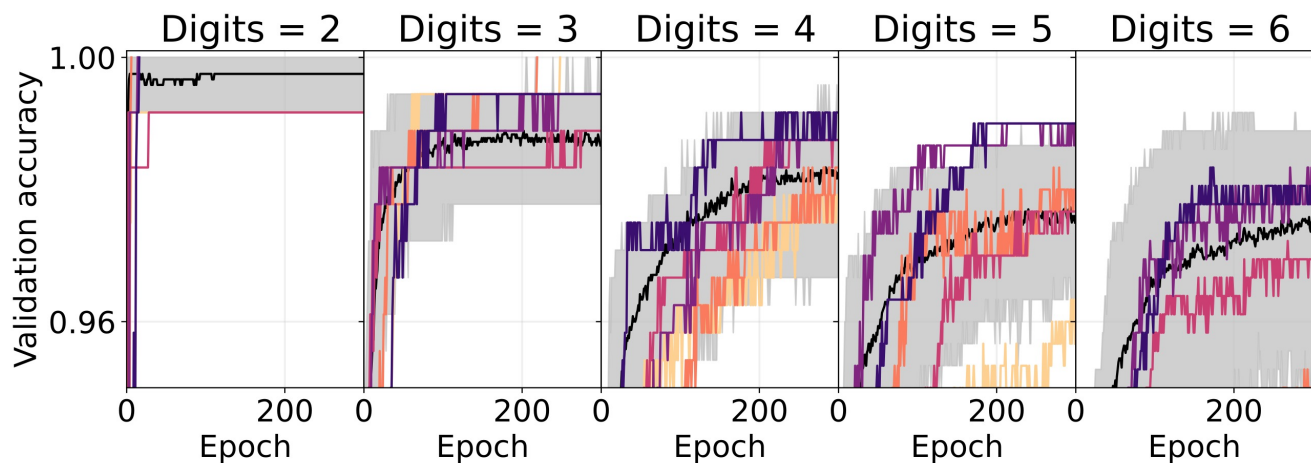
Results

Hybrid Convolutional Network



Results

Hybrid Convolutional Network



Outline

- I. Introduction of single qudit processors
- II. Learning with qudits
- III. Results

IV. Conclusions

Conclusions

- Increasing the number of levels offers advantages in terms of information management and processing.

Conclusions

- Increasing the number of levels offers advantages in terms of information management and processing.
- We have developed tools to deal with any kind of dataset (number of levels and data dimension) in supervised learning tasks with a single unit of information.

Conclusions

- Increasing the number of levels offers advantages in terms of information management and processing.
- We have developed tools to deal with any kind of dataset (number of levels and data dimension) in supervised learning tasks with a single unit of information.
- These tools are hardware efficient: it can be easily implemented in any experimental architecture.

Conclusions

- Increasing the number of levels offers advantages in terms of information management and processing.
- We have developed tools to deal with any kind of dataset (number of levels and data dimension) in supervised learning tasks with a single unit of information.
- These tools are hardware efficient: it can be easily implemented in any experimental architecture.
- Moreover, we can extract a geometrical interpretation of what does it mean to learn for this particular kind of model and strategy.

Acknowledgments



Juan Román



David Zueco



Fernando Luis



Funding and support



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 862893 (FATMOLS).