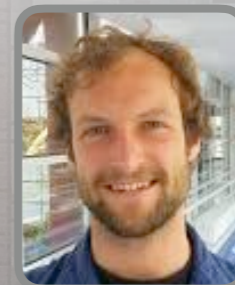
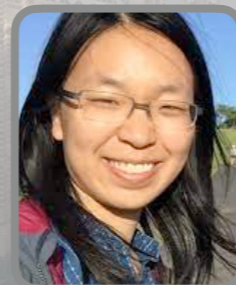
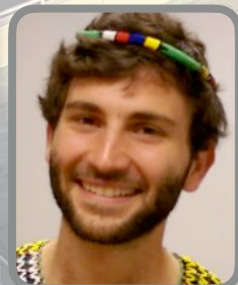




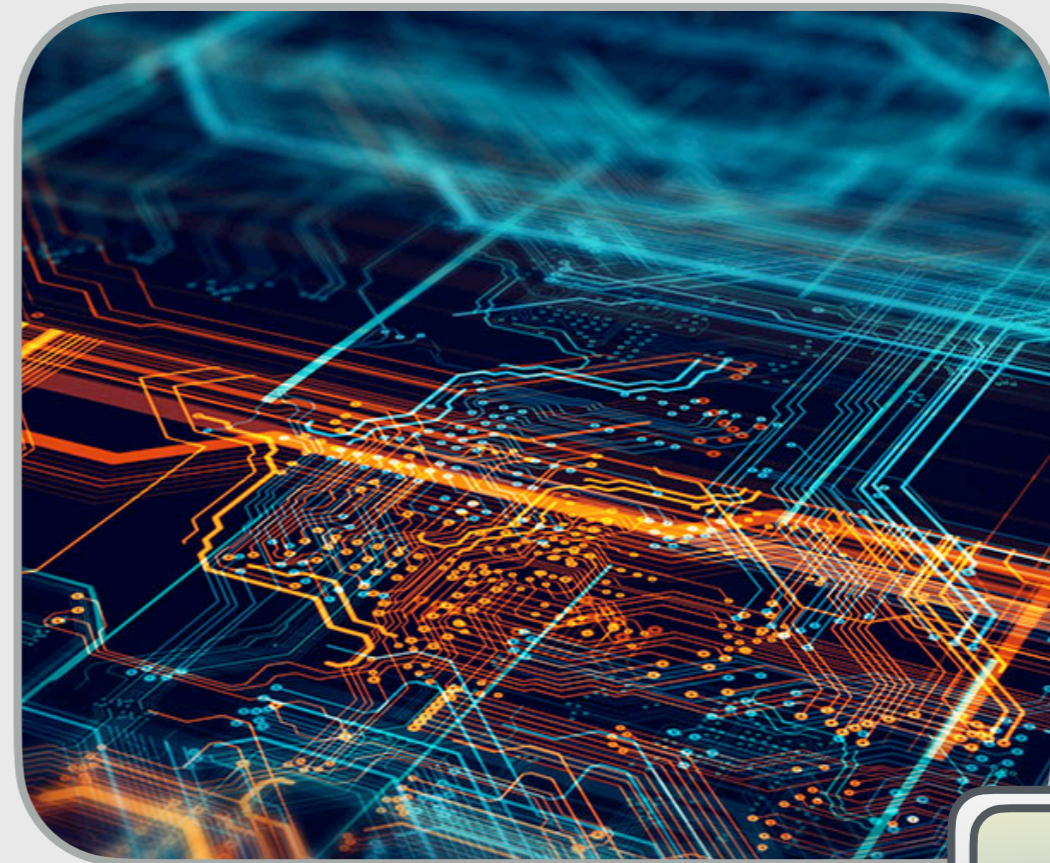
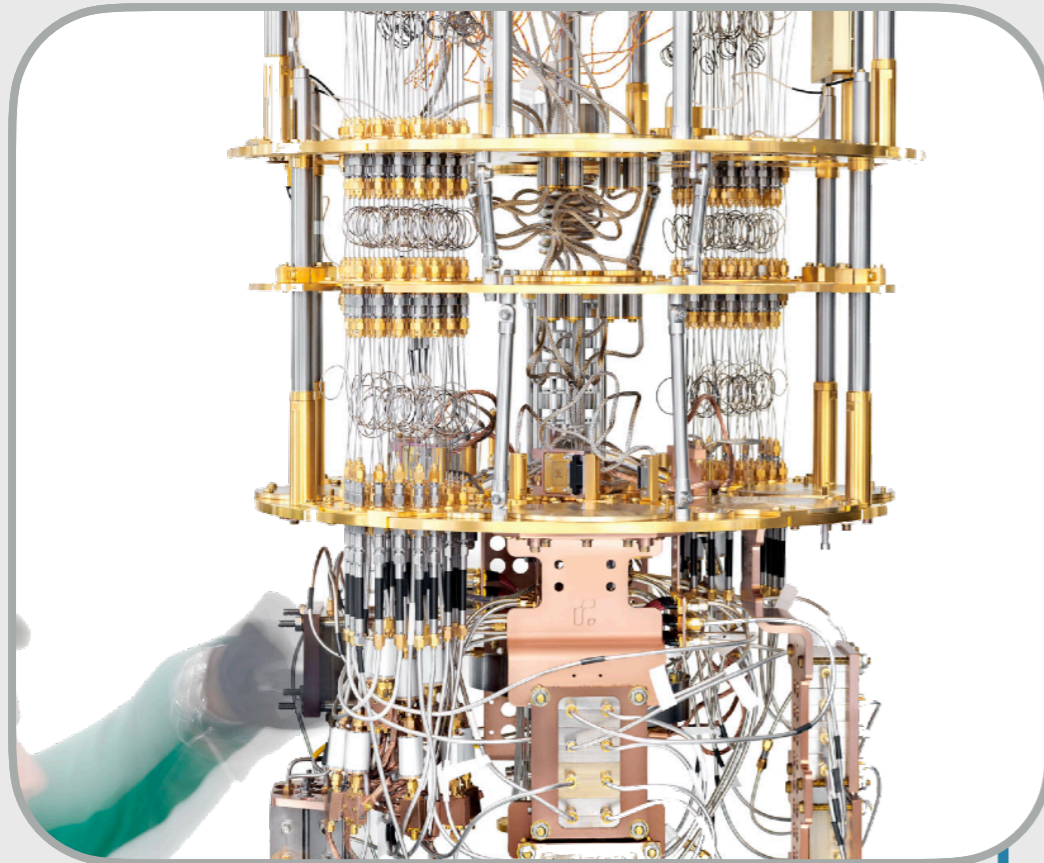
DO QUANTUM COMPUTERS HAVE APPLICATIONS IN MACHINE LEARNING AND COMBINATORIAL OPTIMIZATION? JENS EISERT, FU BERLIN

IPAM Workshop II: Mathematical Aspects of Quantum Learning, University of California, LA, October 2023



With R Sweke, M Hinsche, D Hangleiter, M Ioannou, Y Quek, J Haferkamp, A Nietner, J-P Seifert and others

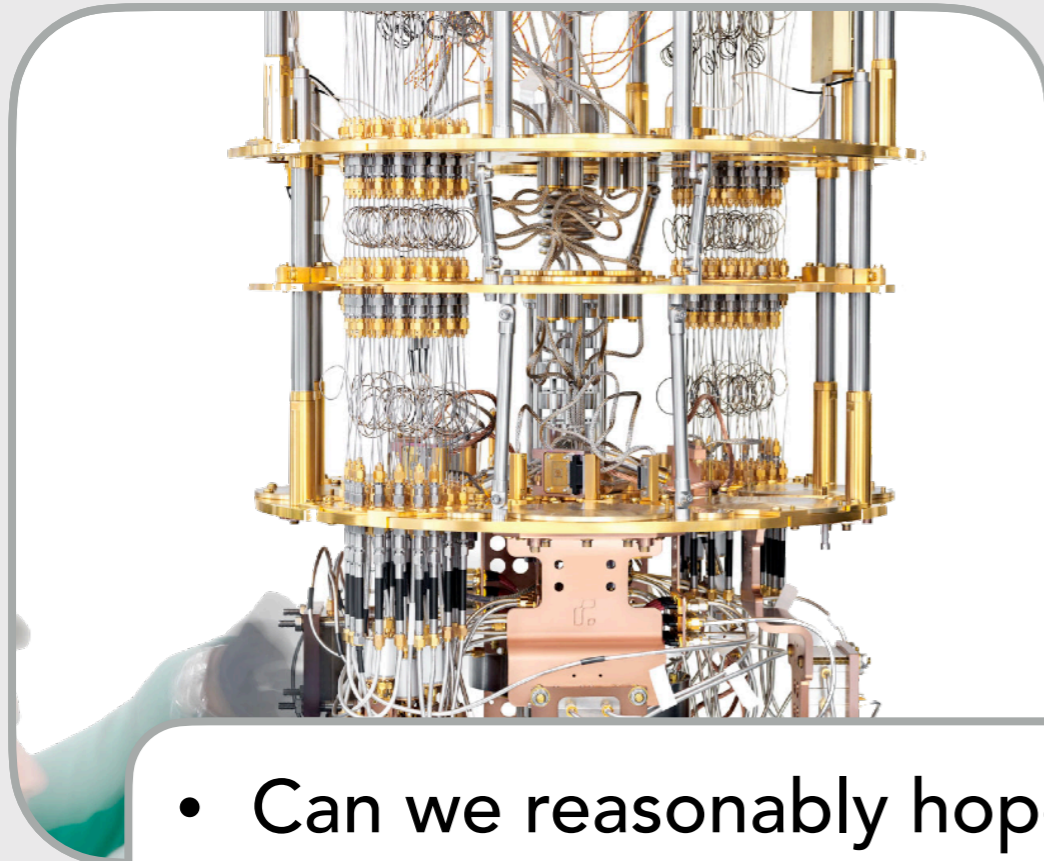
- **Quantum computers** promise advantages in computational tasks



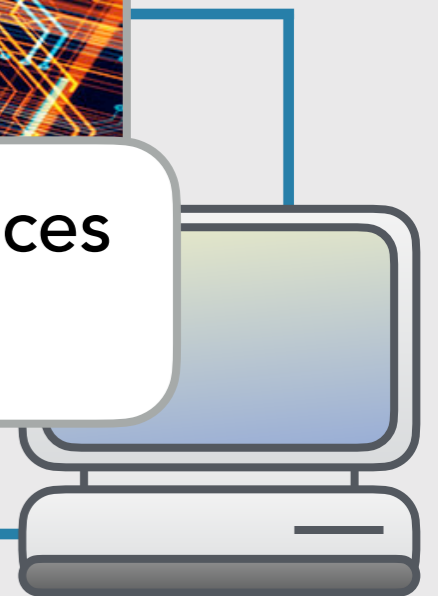
Arute, Arya, ..., Martinis, Nature 574, 505 (2019)

Wang, Qin, Ding, Chen, Chen, You, He, Jiang, Wang, You, Renema, Hoefling, Lu, Pan, Phys Rev Lett 123, 250503 (2019)

- **Quantum computers** promise advantages in computational tasks



- Can we reasonably hope noisy, realistic quantum devices to provide a **speedup over classical computers**?

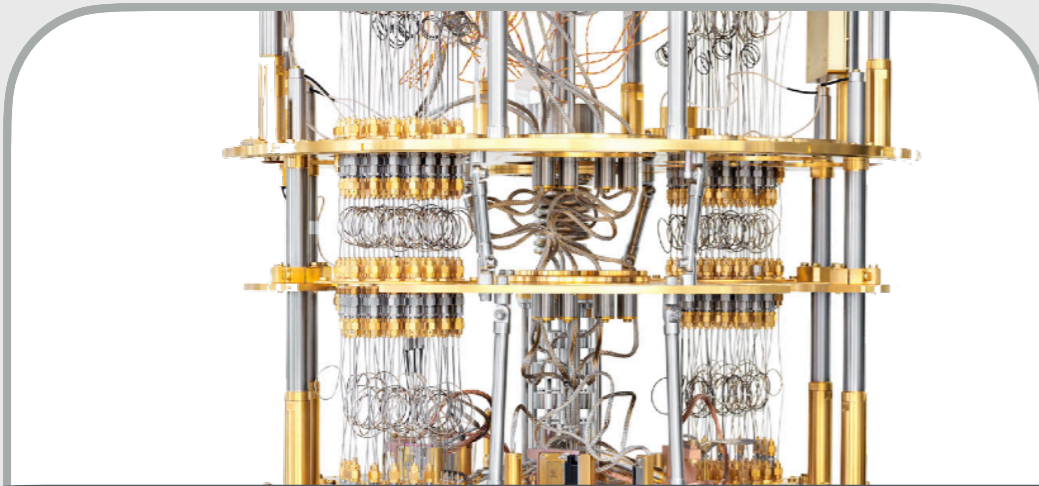


Arute, Arya, ..., Martinis, Nature 574, 505 (2019)

Wang, Qin, Ding, Chen, Chen, You, He, Jiang, Wang, You, Renema, Hoefling, Lu, Pan, Phys Rev Lett 123, 250503 (2019)



- **Quantum computers** promise advantages in computational tasks

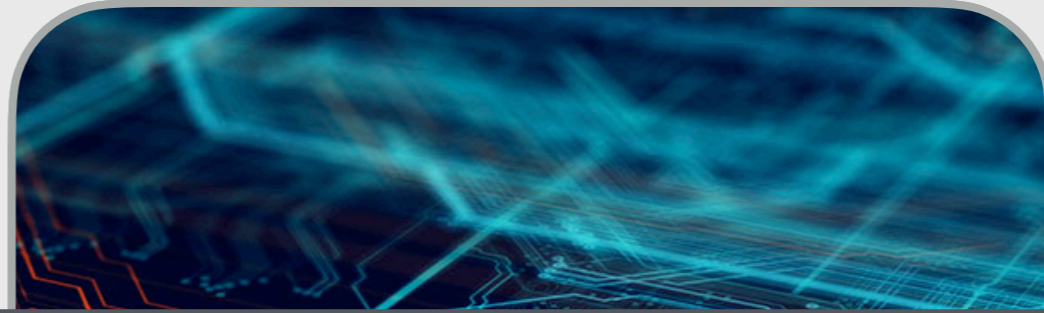
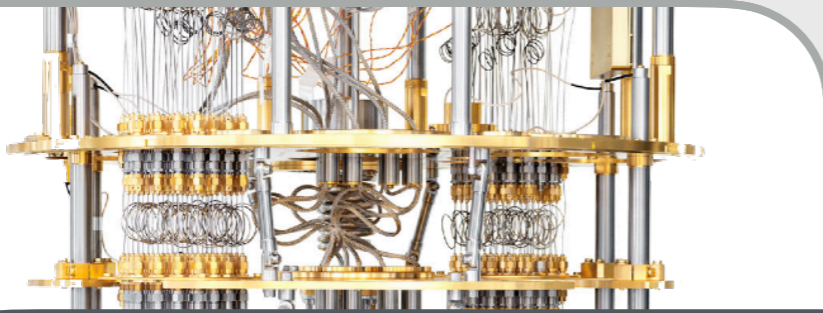


- **Quantum advantage** for paradigmatic sampling problems
- Sampling up to a constant error in $\|\cdot\|_{l_1}$ distance is **classically hard**

Arute, Arya, ..., Martinis, Nature 574, 505 (2019)

Wang, Qin, Ding, Chen, Chen, You, He, Jiang, Wang, You,
Renema, Hoeffling, Lu, Pan, Phys Rev Lett 123, 250503 (2019)

- **Quantum computers** promise advantages in computational tasks



- **Verification** is subtle: XEB benchmarking, no black-box verification

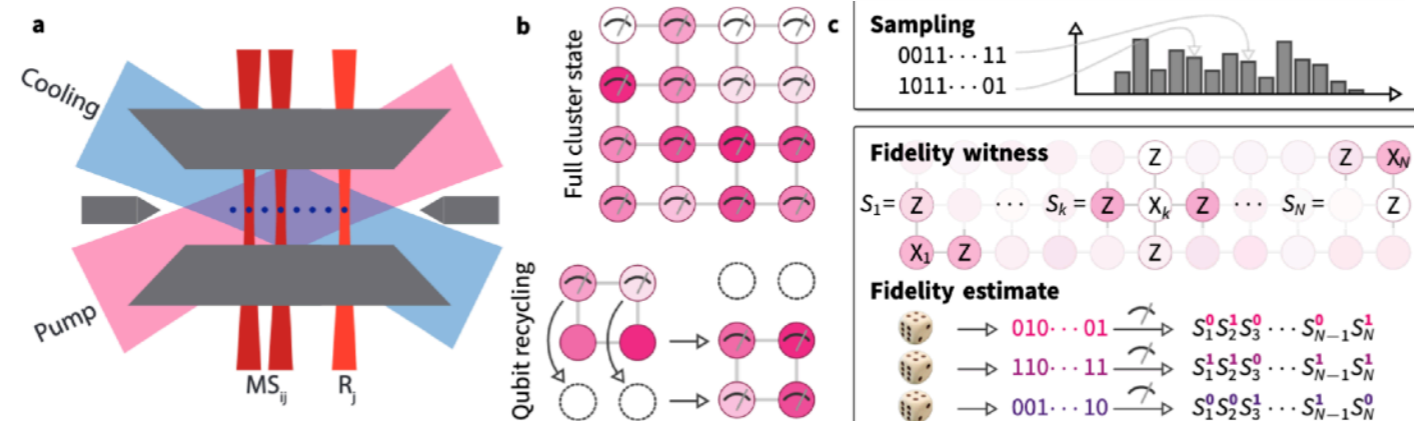
Hangleiter, Kliesch, Eisert, Gogolin, Phys Rev Lett 122, 210502 (2019)

Neil, ..., Martinis, Science, 360, 195 (2018)

Boixo, Isakov, Smelyanskiy, Babbush, Ding, Jiang, Bremner, Martinis, Neven, Nature Phys 14, 595 (2018)

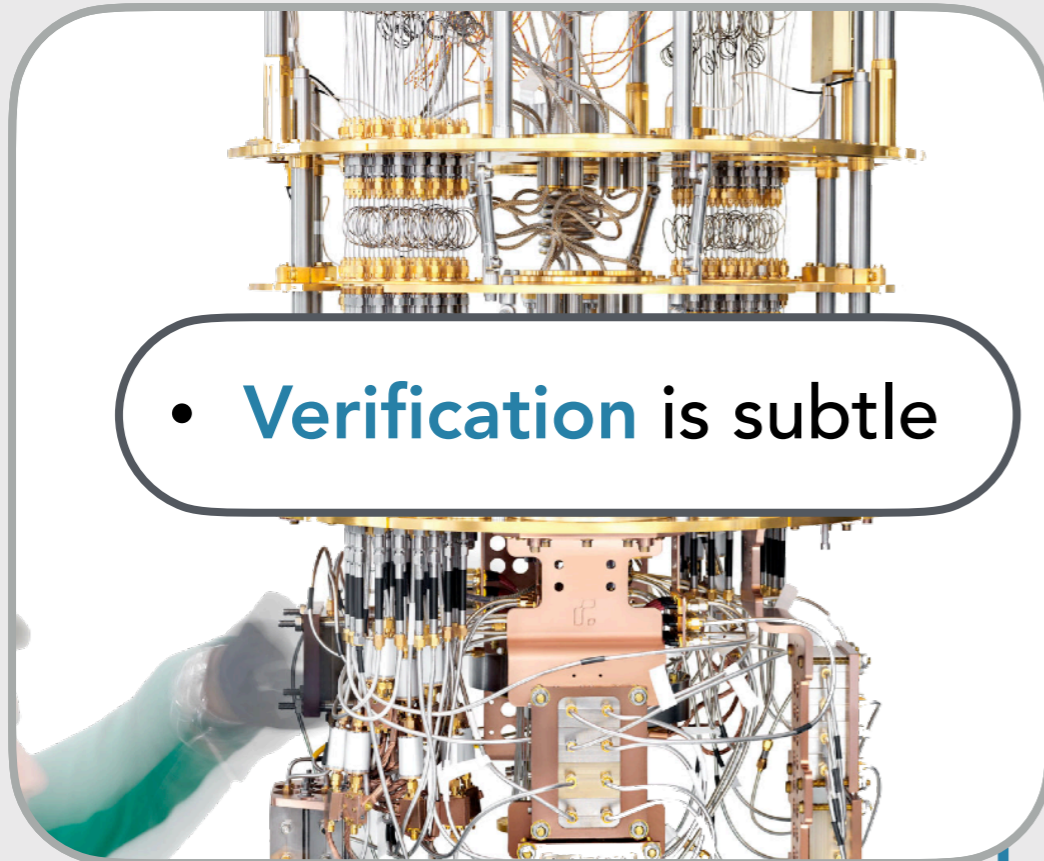
Deshpande, Mehta, Vincent, Quesada, Hinsche, Ioannou, Madsen, Lavoie, Qi, Eisert, Hangleiter, Fefferman, Dhand, Science Advances 8, eabi7894 (2022)

- **Quantum** verification: Trapped ions (with Ringbauer-Blatt-Monz)

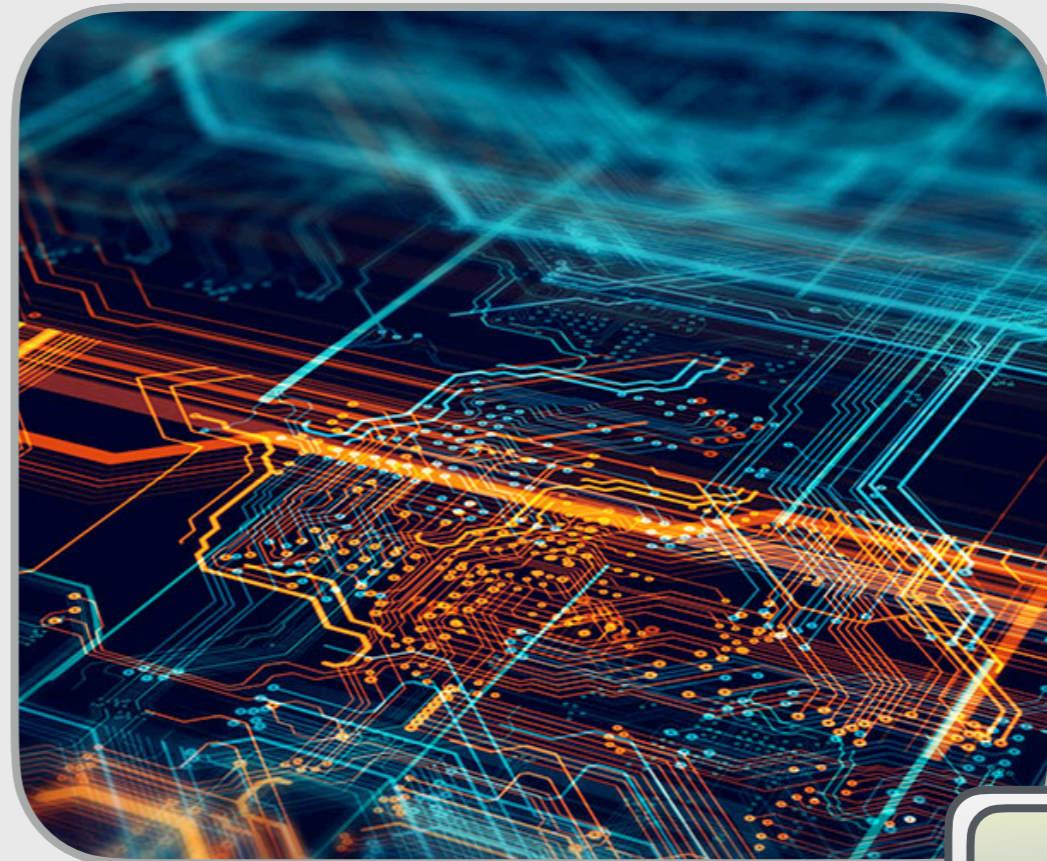


Ringbauer, Hinsche, Feldker, Faehrmann, Bermejo-Vega, Edmunds, Stricker, Marciniak, Meth, Pogorelov, Postler, Blatt, Schindler, Eisert, Monz, Hangleiter, arXiv:2307.14424 (2023)

- **Quantum computers** promise advantages in computational tasks



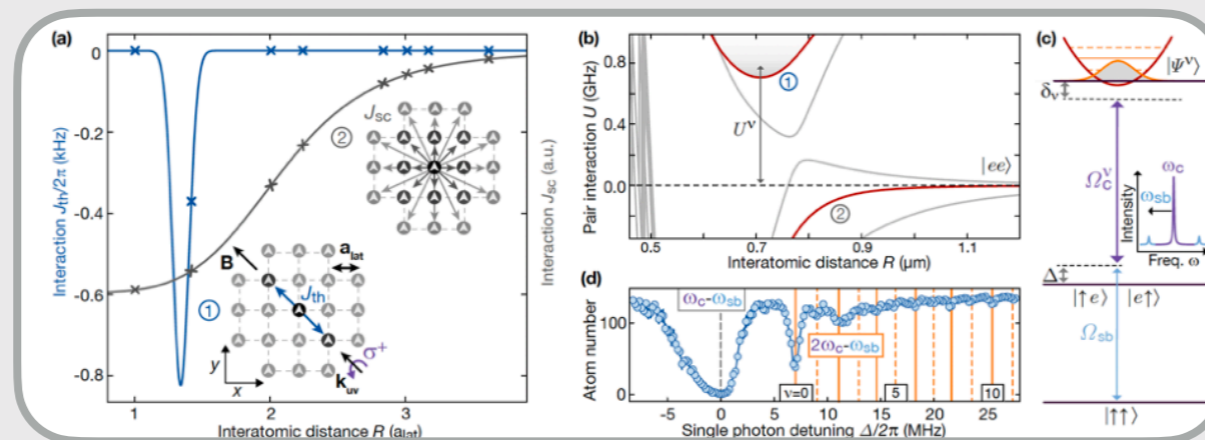
- **Verification** is subtle



M Caro, "Classical verification of quantum learning"

- Encouraging - but what **next?**

- **Programmable** quantum simulators



Hollerith, Srakaew, Wei, Rubio-Abadal, Adler, Weckesser, Kruckenhauser, Walther, Bijnen, Rui, Gross, Bloch, Zeiher, Phys Rev Lett 128, 113602 (2022)

- **Variational** quantum algorithms

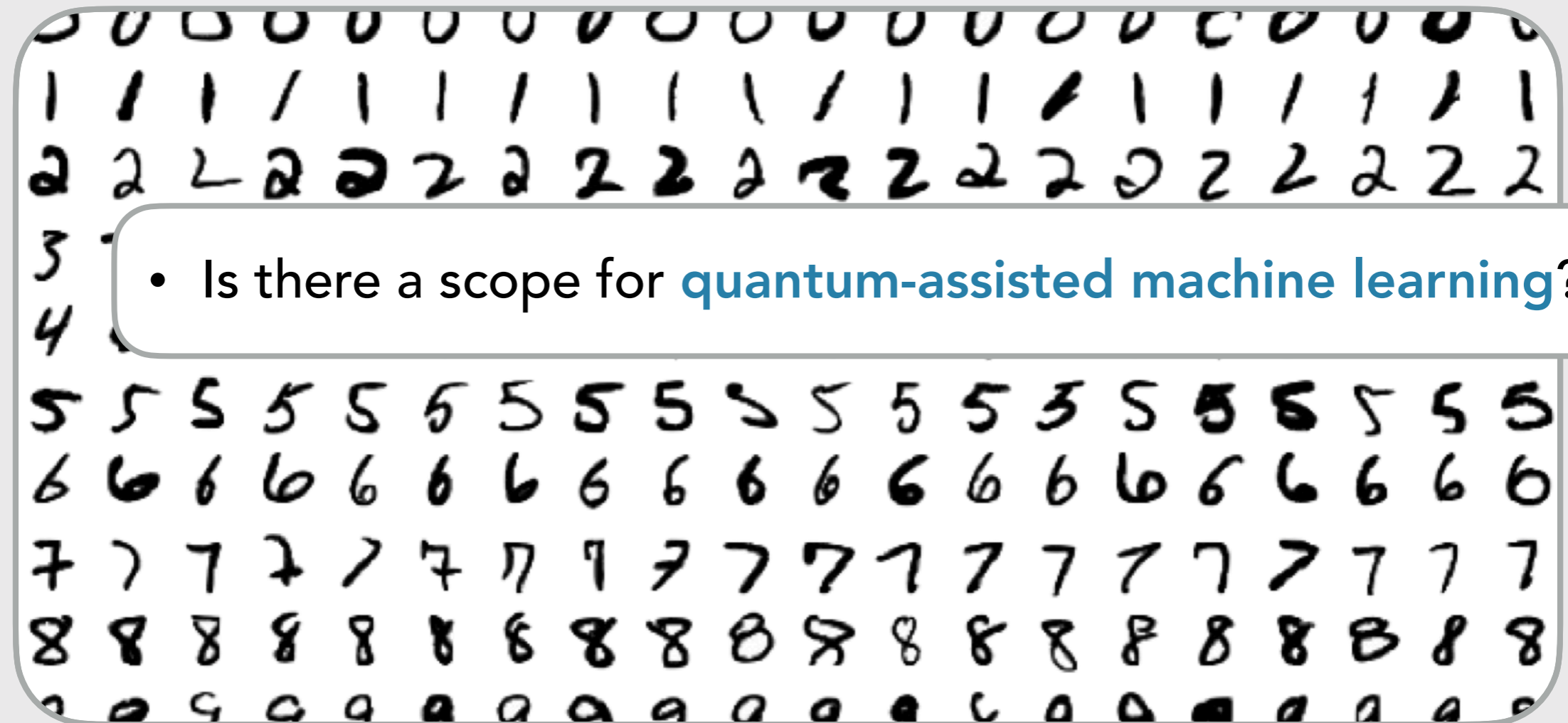
McClean, Romero, Babbush, Aspuru-Guzik, New J Phys 18, 023023 (2016)

Farhi, Goldstone, Gutmann, arXiv:1411.4028 (2014)

- Instances of quantum-assisted **machine learning**



- **Machine learning** has changed the world we live in
 - Chat-GPT, DALL-E 2
 - Machine learning in **study of materials** and **quantum chemistry**



Dunjko, Briegel, Rep Prog Phys 81, 074001 (2018)

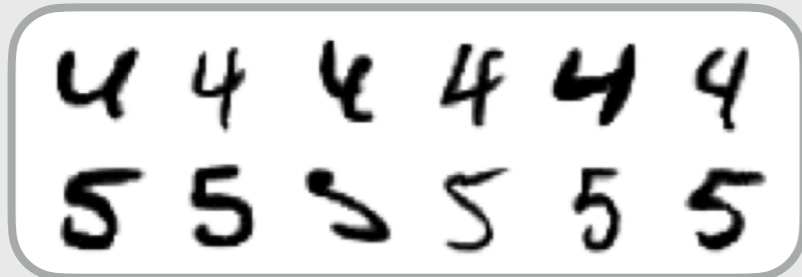
Biamonte, Wittek, Pancotti, Rebentrost, Wiebe, Lloyd, Nature 549, 195 (2017)

Arunachalam, de Wolf, arXiv:1701.06806 (2017)

QUANTUM ADVANTAGES IN LEARNING TASKS?



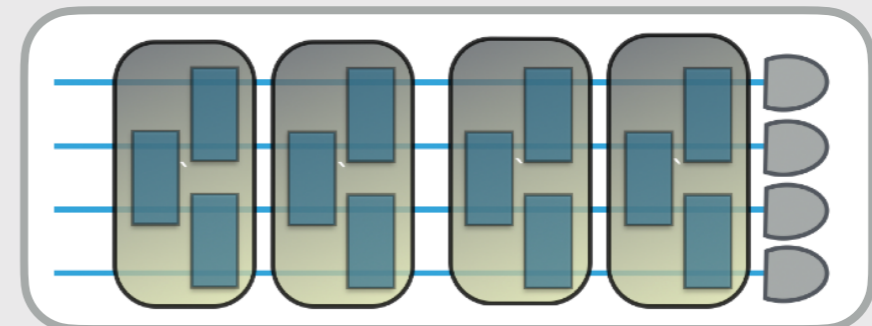
- We would like to have **rigorous guarantees** for **state-of-the-art learning** algorithms applied on **real-world datasets**



- **Data set**
 - Pictures of cats and dogs
 - Protein configurations
 - Stock market data



- **Algorithm**
 - Stochastic gradient descent
 - Expectation maximization



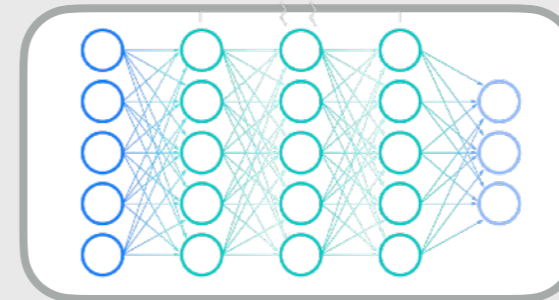
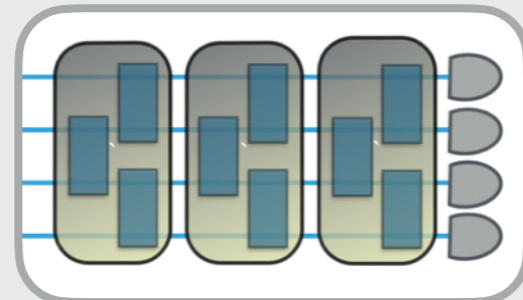
- **Model**
 - Neural networks
 - Parameterised quantum circuits
 - Probabilistic graphical models

- **Sample complexity, computational complexity, generalisation bounds**

QUANTUM ADVANTAGES IN LEARNING TASKS?



- Even better would be “classical lower bounds” and “quantum upper bounds”



- **Algorithm**

- Stochastic gradient descent

- **Model**

- Parametrized quantum circuits

- **All classical learners**

- All models
- All algorithms

- **Quantum**

- **Classical**



- Can we make substantial progress for **fine-tuned data sets**?

• Quantum

• Classical

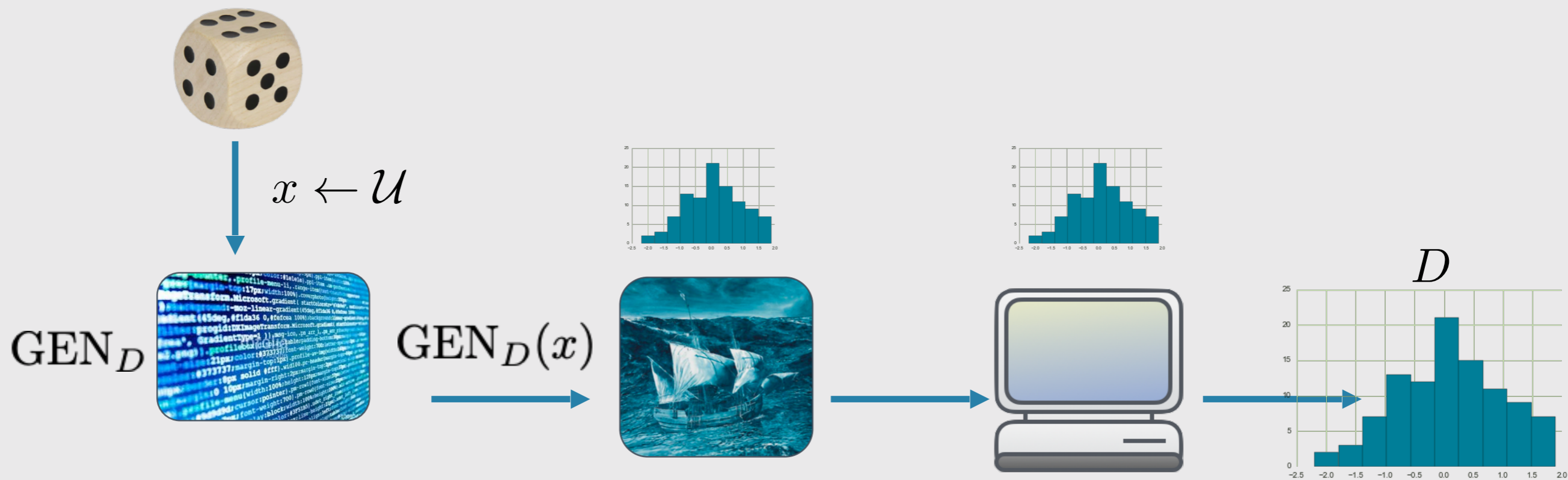


PROVEN QUANTUM ADVANTAGES IN LEARNING TASKS?

Sweke, Seifert, Hangleiter, Eisert, Quantum 5, 417 (2021)

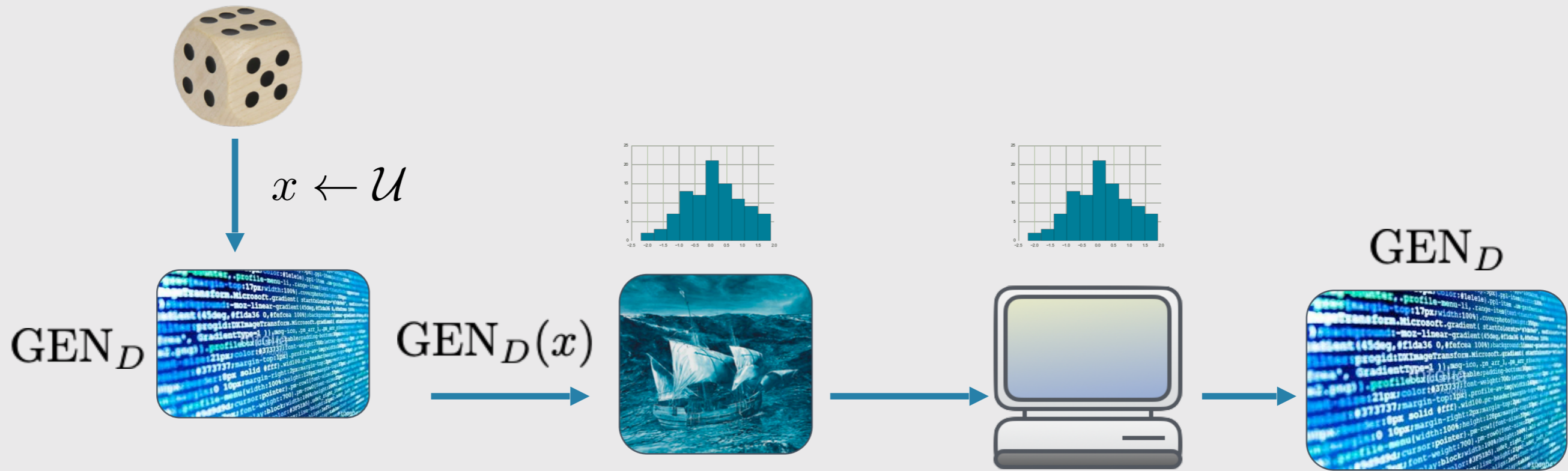
Pirnay, Sweke, Eisert, Seifert, Phys Rev A 107, 042416 (2023)

Liu, Liu, Liu, Ye, Alexeev, Eisert, Liang, arXiv:2303.03428 (2023)



All distributions

Concept class

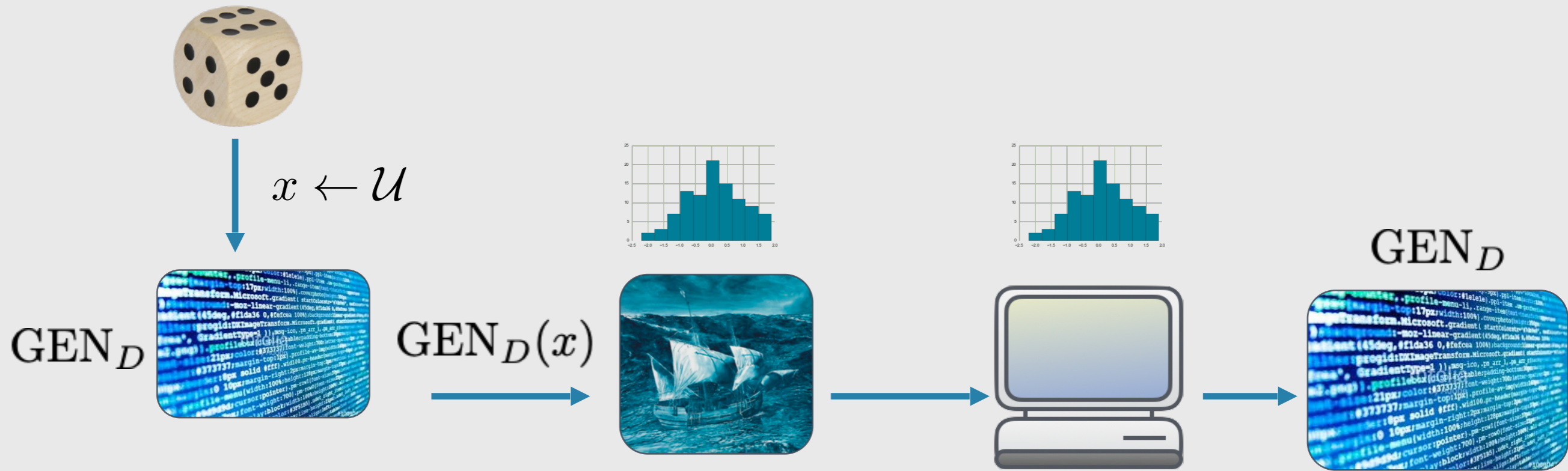


- **Generator modelling**
Create samples

All distributions

- **Task:** Learn a generator GEN_D of a distribution D

Concept class



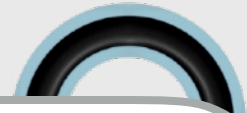
• Can there be a **quantum advantage**? Are there distributions that

• are **efficiently** quantum generator learnable?

• are **not efficiently** classical generator learnable?

• Quantum

• Classical



- “Probably approximately correct” learning of distribution classes
- A distribution class \mathcal{C} is efficiently PAC learnable w.r.t. distance d if there is an algorithm \mathcal{A} which for every $D \in \mathcal{C}$ and every $\epsilon, \delta > 0$ given access to an oracle $O(D)$, outputs in time $\text{poly}(|D|, 1/\epsilon, 1/\delta)$
- with probability at least $1 - \delta$ (“probably”) a generator $\text{GEN}_{D'}$ of a distribution D' such that

$$d(D, D') < \epsilon$$

(“approximately correct”)

GEN_D

• Quantum

• Classical



- Can there be a **quantum advantage**? Are there distributions that
 - are **efficiently** quantum generator learnable?
 - are **not efficiently** classical generator learnable?
- For the Kullback Leibler divergence and the natural **sample oracle**?

YES!*

* under the decisional Diffie-Hellman assumption for the group family of quadratic residues

• **Quantum**

• **Classical**



- It is **quantum** easy

- Is it **classical** hard

• **Quantum**

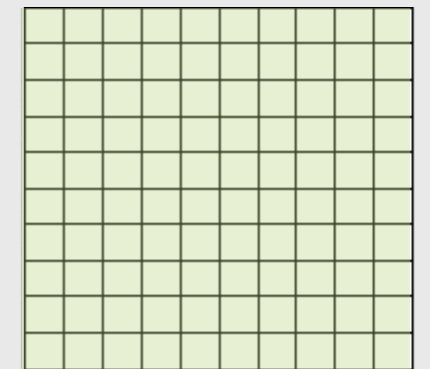
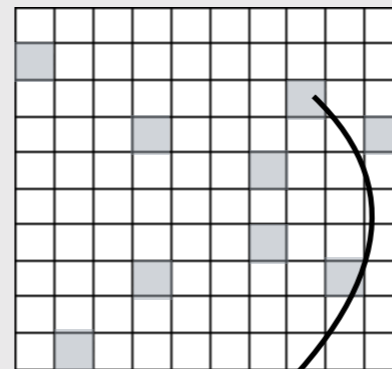
• **Classical**



- Pseudorandom function
 - A collection of keyed functions $\{F_k : D \rightarrow D'\}_{k \in \mathcal{K}}$ that cannot be distinguished from uniformly random functions $\{F : D \rightarrow D'\}$ by any polynomial time algorithm \mathcal{A}

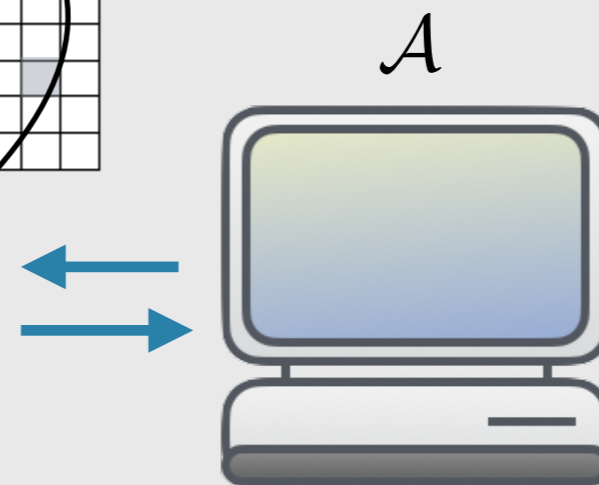
$$\{F_k : D \rightarrow D'\}_{k \in \mathcal{K}}$$

$$\{F : D \rightarrow D'\}$$



F_k

F



• Quantum

• Classical



- **Pseudorandom function**

- A collection of keyed functions $\{F_k : D \rightarrow D'\}_{k \in \mathcal{K}}$ that cannot be distinguished from uniformly random functions $\{F : D \rightarrow D'\}$ by any polynomial time algorithm \mathcal{A}

$$\{F_k : D \rightarrow D'\}_{k \in \mathcal{K}}$$

$$\{F : D \rightarrow D'\}$$

- **Theorem**

- Given a classical secure pseudorandom function $\{F_k\}_k$, the distribution class $\{D_k\}_k$ defined by the "Kearns generator"

$$\text{KGEN}_k(x) = x || F_k(x)$$

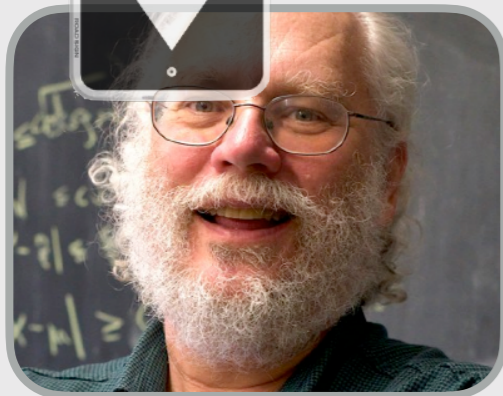
cannot be efficiently classically generator learned

HINT AT THE PROOF: QUANTUM PART



• \exists one-way function

ONE WAY

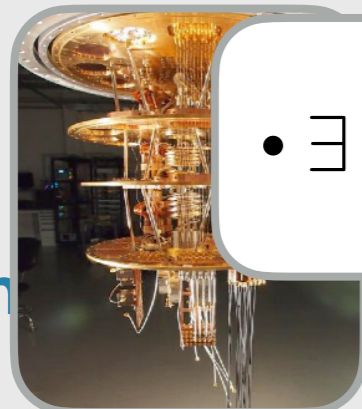


• \exists pseudorandom generator

• \exists pseudorandom function

• \exists hard-to-learn distribution class

• Quantum



• Classical

Sweke, Seifert, Hangleiter, Eisert, Quantum 5, 417 (2021)

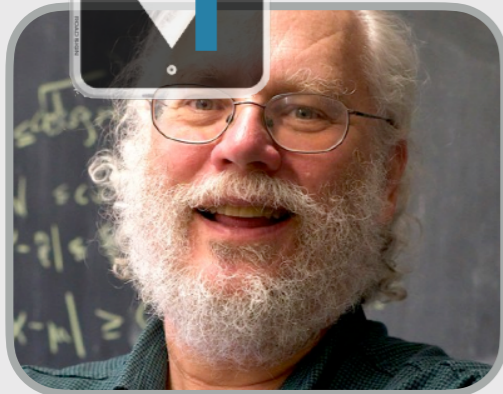
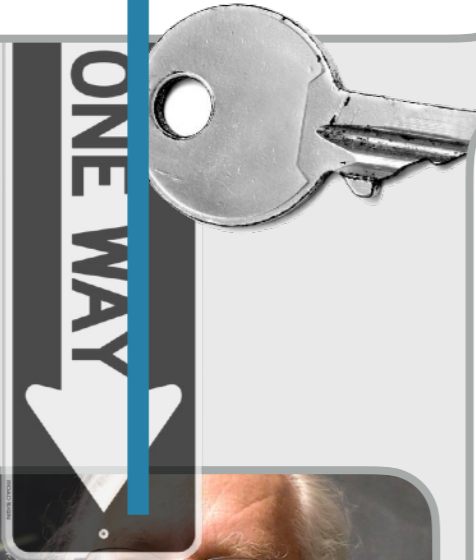
Shor, SIAM J Comp 26, 1484 (1997)

HINT AT THE PROOF: QUANTUM PART



• \exists one-way function

• \exists pseudorandom generator



Input: (p, g, g^a) , key $b \in \mathbb{Z}_q^*$

$\text{modexp}_{p, g^{x_1 a}}$

f_p

$\text{modexp}_{p, g^{x_2 a}}$

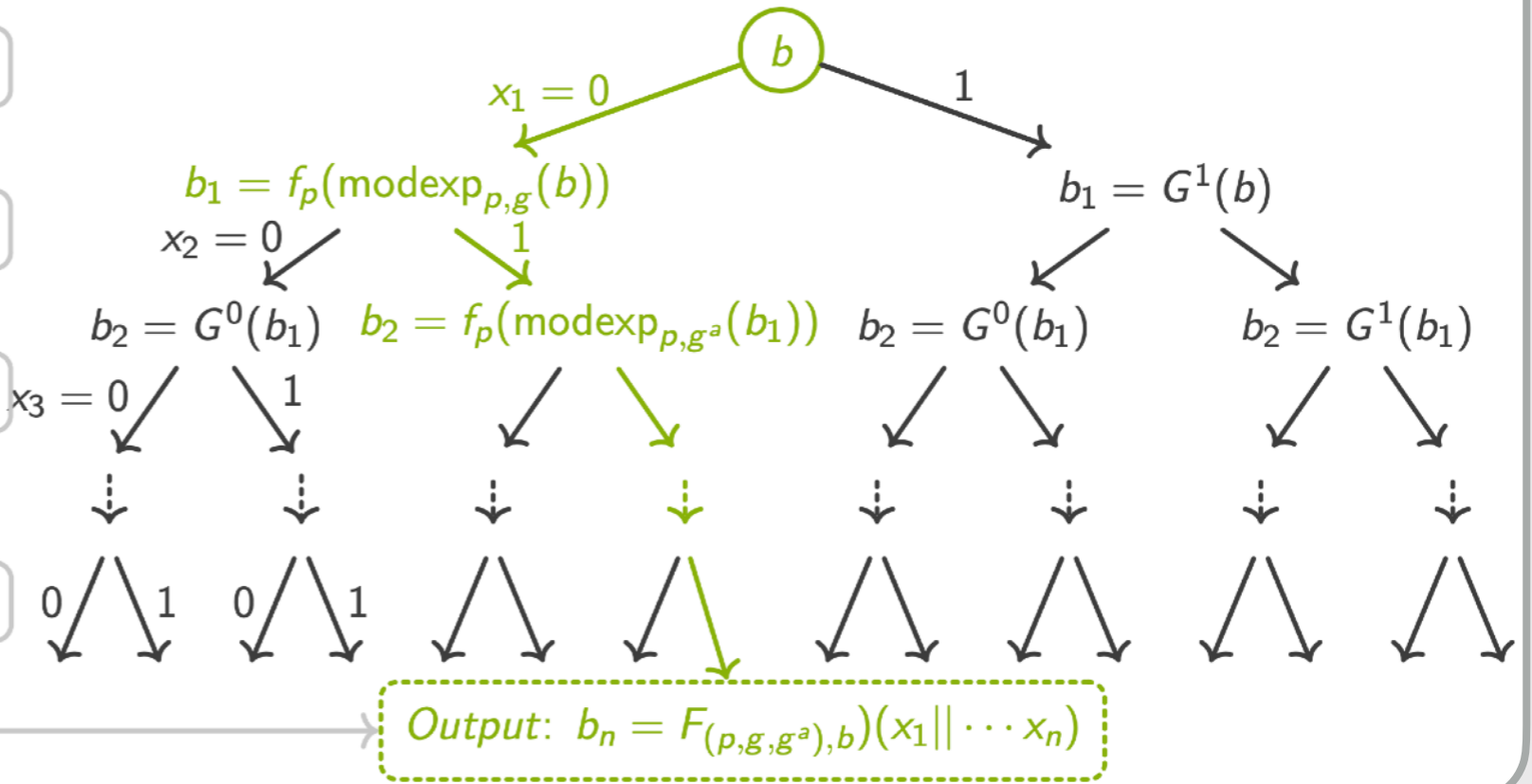
f_p

$\text{modexp}_{p, g^{x_3 a}}$

f_p

$\text{modexp}_{p, g^{x_n a}}$

f_p



• Cracking hard-to-learn distribution classes with a QC

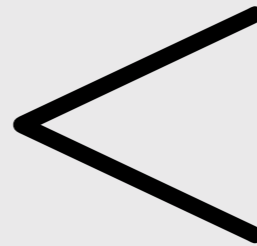
• Quantum

• Classical

Sweke, Seifert, Hangleiter, Eisert, Quantum 5, 417 (2021)

Shor, SIAM J Comp 26, 1484 (1997)

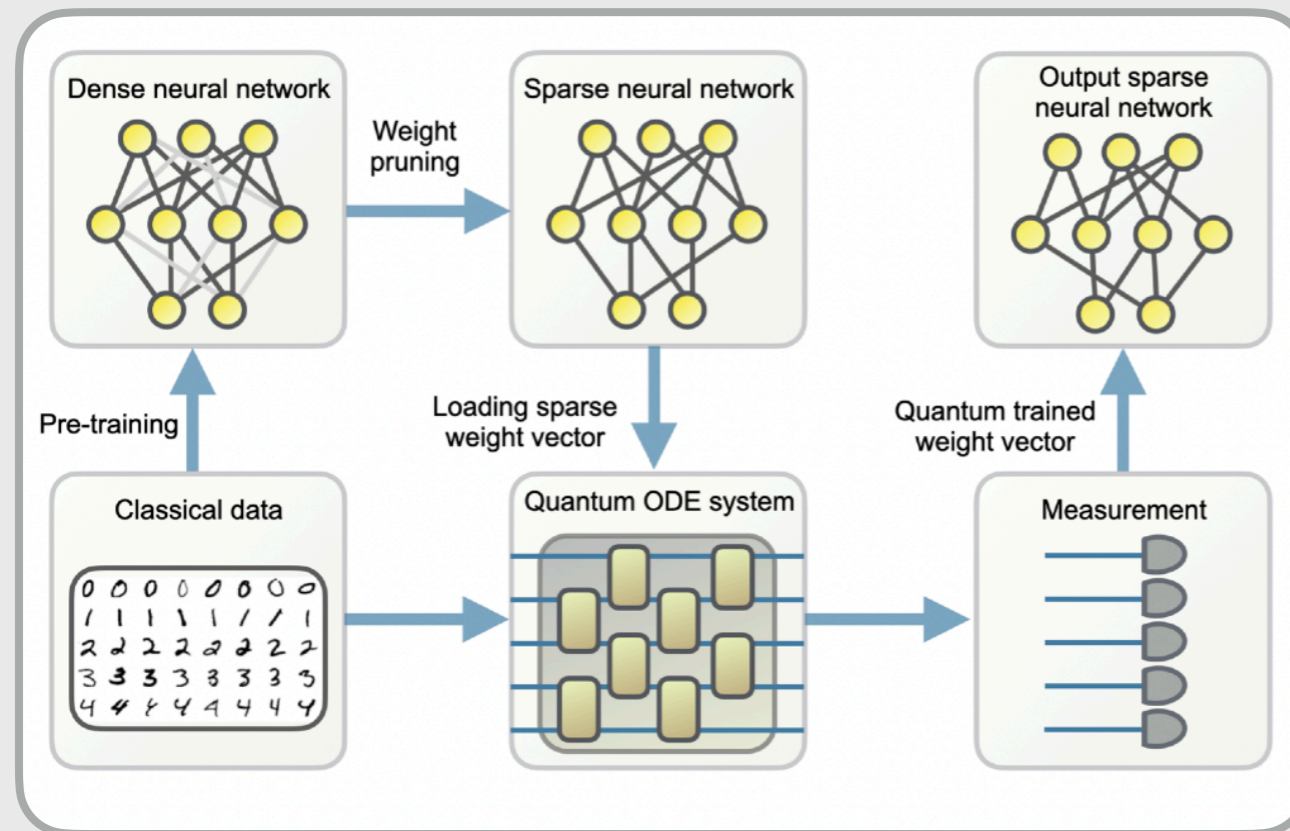
- There is a proven **quantum generator learning advantage**:
"Quantum computers learn (exponentially) more efficiently"



Sweke, Seifert, Hangleiter, Eisert, Quantum 5, 417 (2021)
Pirnay, Sweke, Eisert, Seifert, Phys Rev A 107, 042416 (2023)
Liu, Arunachalam, Temme, Nature Phys 17, 1013 (2021)

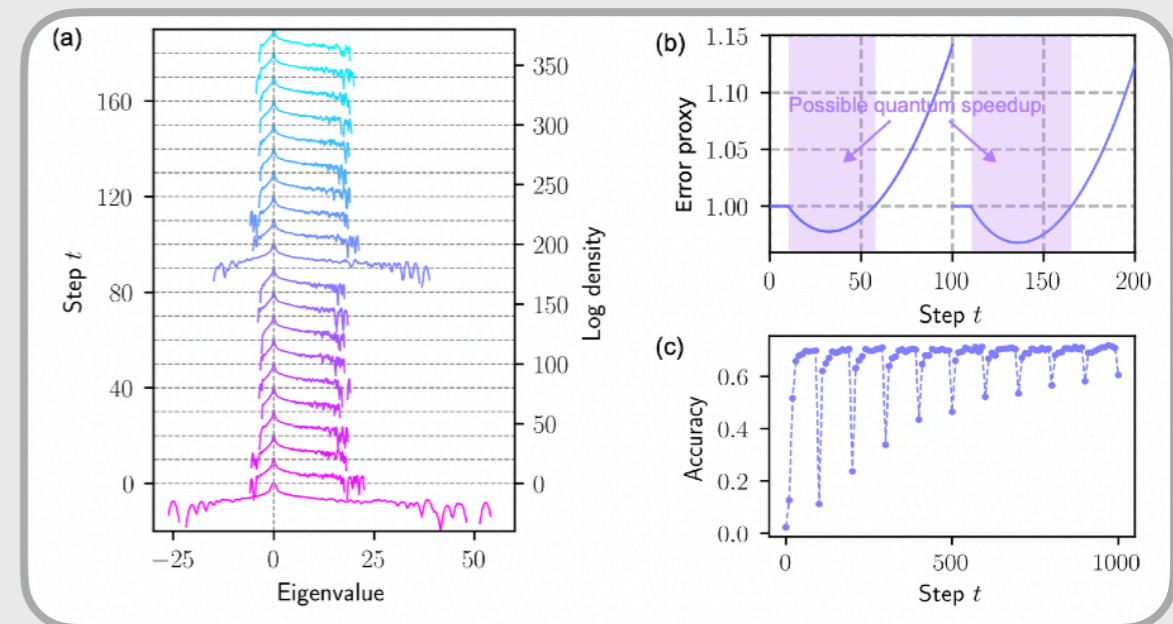
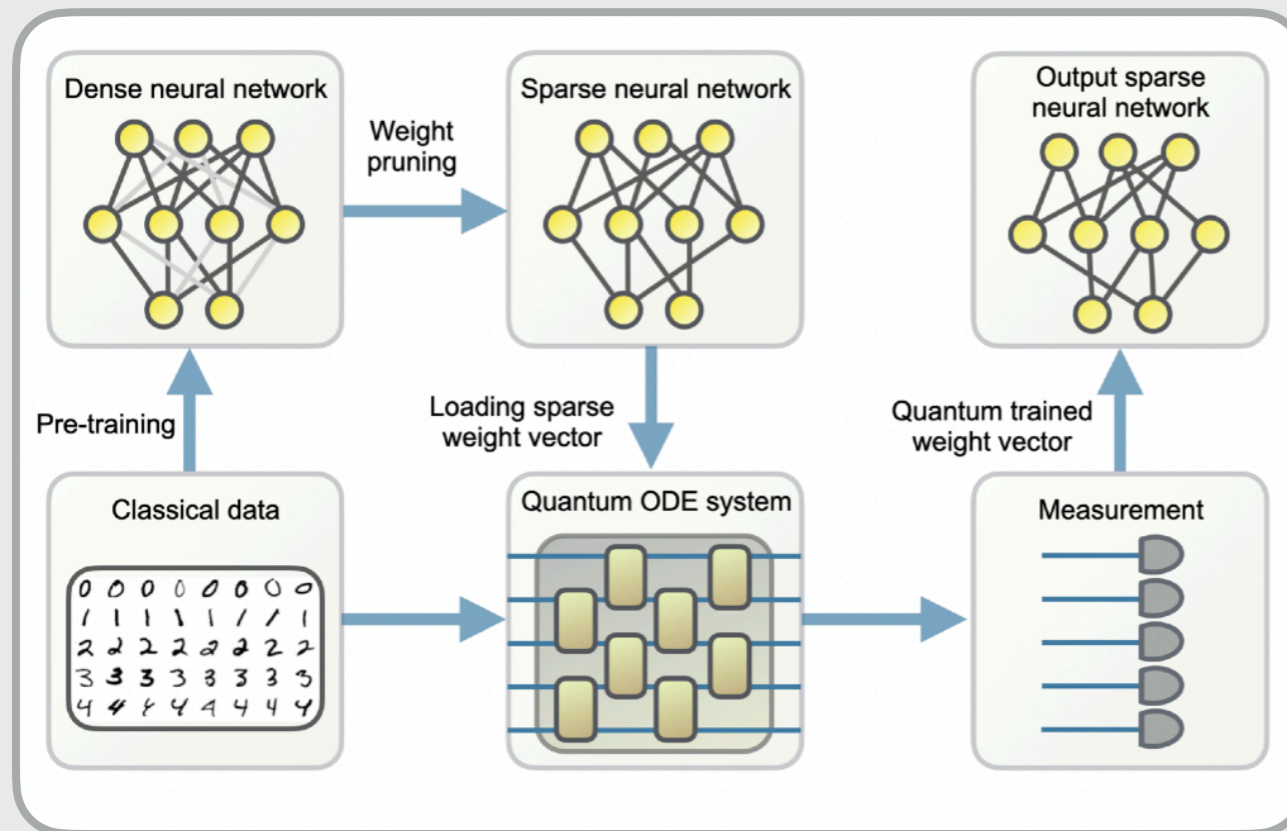


- Can **classical networks** be better trained with quantum computers?





- Can **classical networks** be better trained with quantum computers?



- Devise variants of **HHL** for training sparse (pruned) classical networks

Theorem 1 (Informal). For a sparse machine learning model with model size n , running T iterations, with the algorithm being fully dissipative with small learning rates (whose formal definition is given in the supplemental material), there is a quantum algorithm that runs in

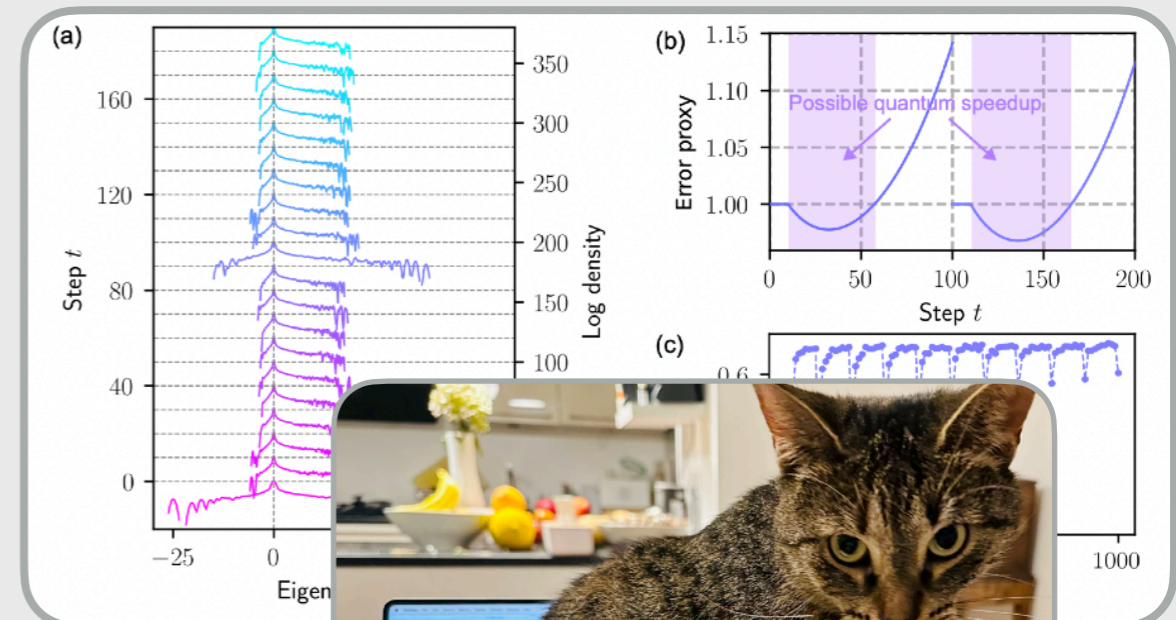
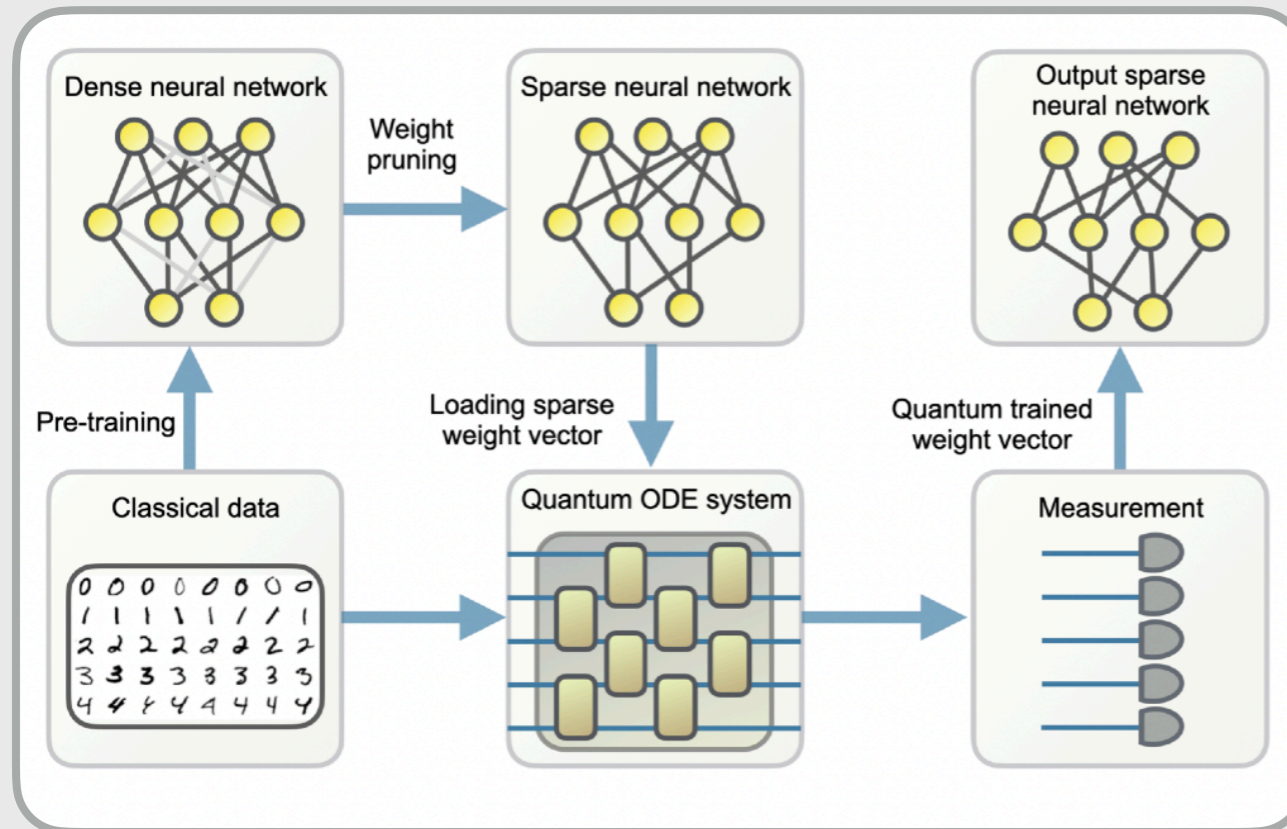
$$\mathcal{O}\left(T \times \text{poly}\left(\log n, \frac{1}{\epsilon}\right)\right) \quad (1)$$

time with precision $\epsilon > 0$. The sparsity condition also ensures the efficiency of uploading and downloading quantum states towards classical processors.

CAN QUANTUM COMPUTERS HELP IN CLASSICAL TRAINING?



- Can **classical networks** be better trained with quantum computers?



Theorem 1 (In...
with model size...
being fully diss...
definition is gi...
quantum algo...

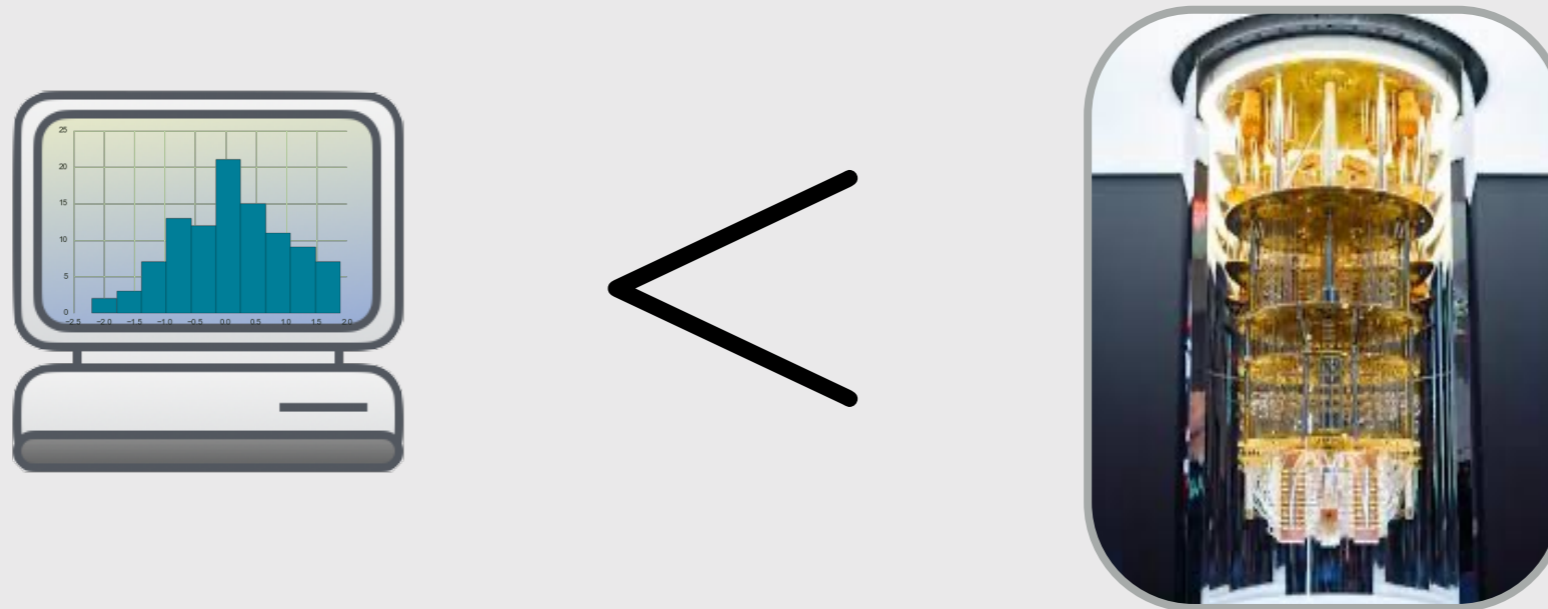
del
hm
mal
is a

- Devise variants of **HHL** for training sparse (pruned) classical networks

(1)

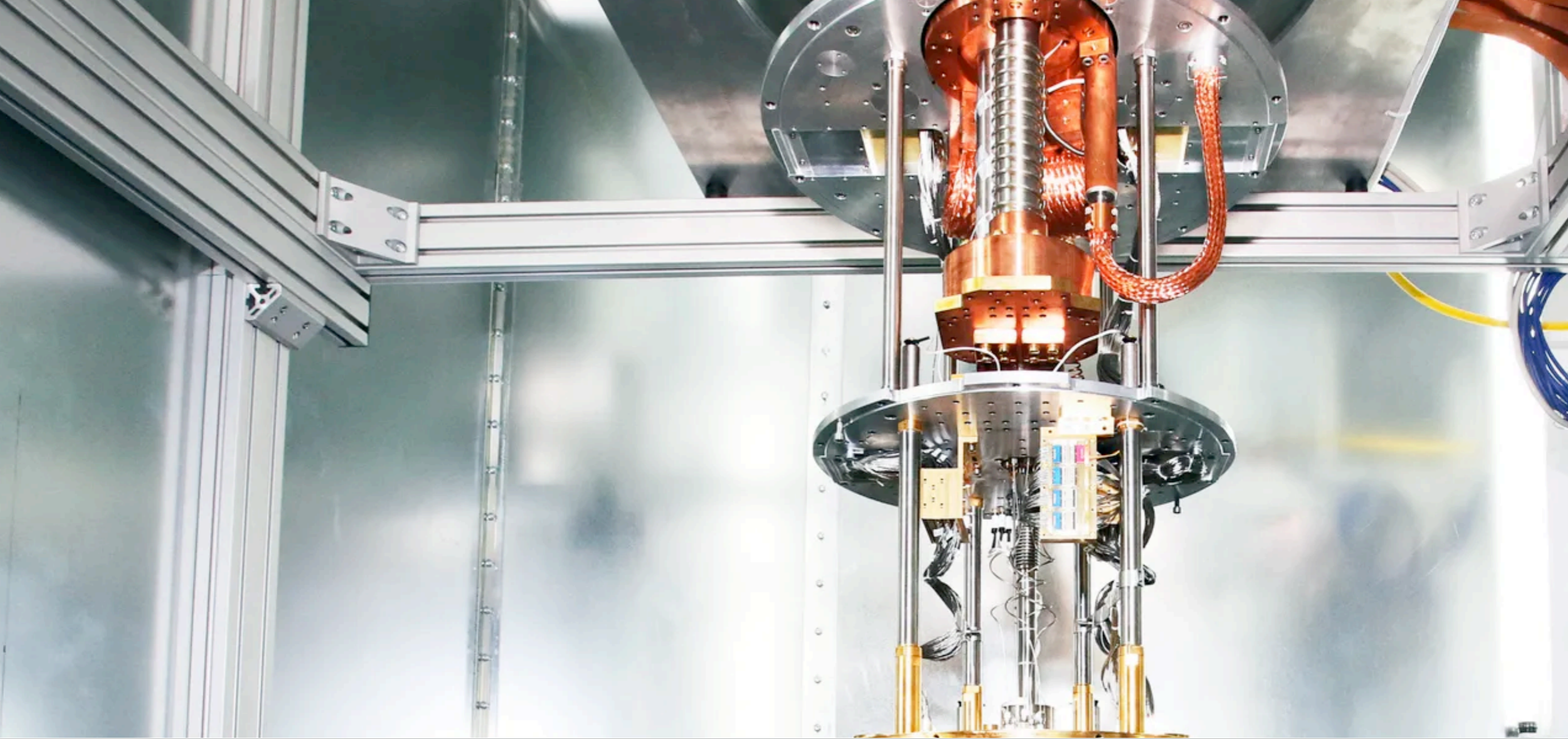
asures
ates

• Quantum algorithms can have superpolynomial speedups over stochastic gradient descent in **training** pruned classical networks



- For unstructured data, but still **fault tolerant quantum computers**

- Quantum algorithms can have superpolynomial speedups over stochastic gradient descent in **training** pruned classical networks



LEARNING THE OUTPUT DISTRIBUTION OF QUANTUM CIRCUITS

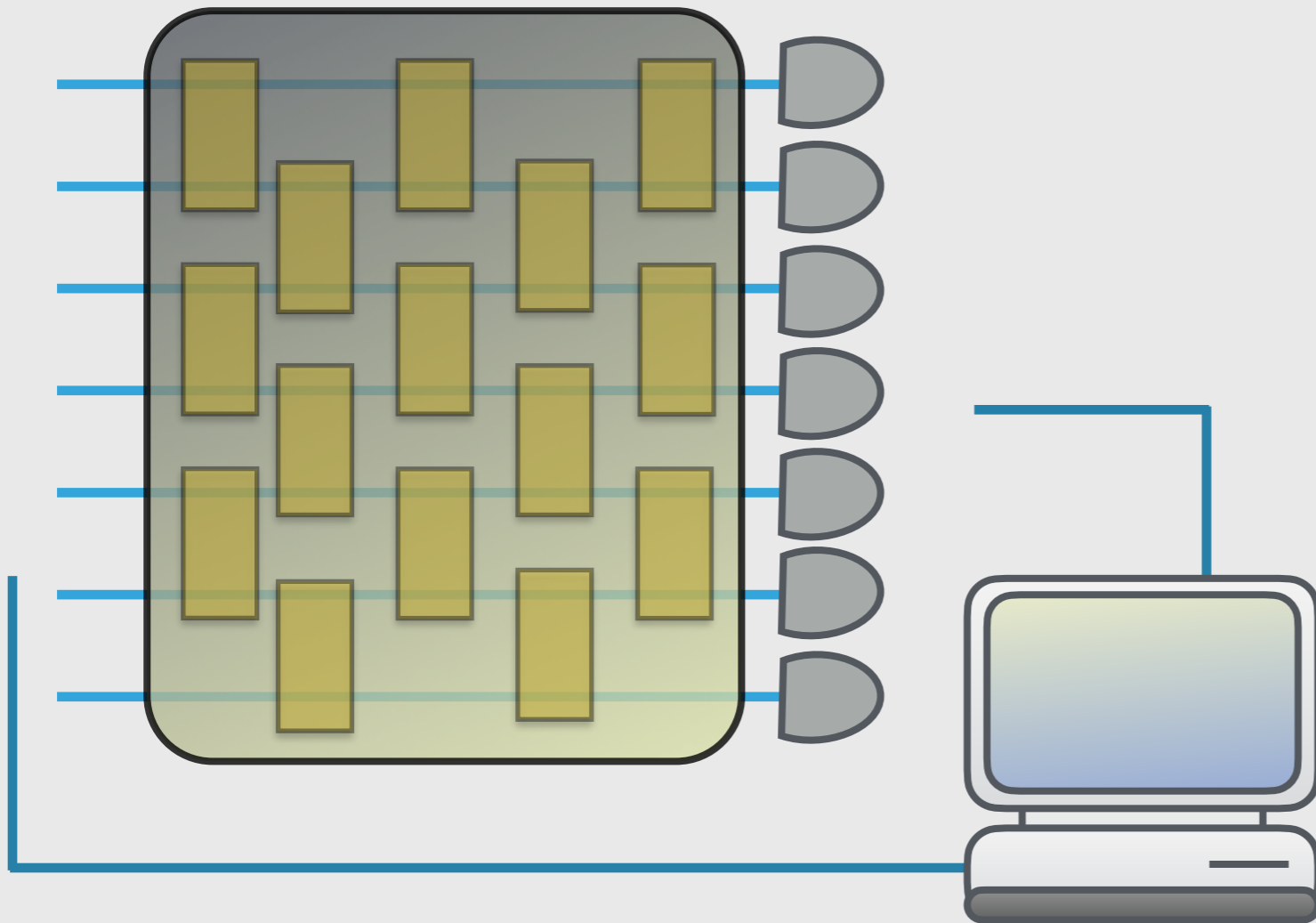
Hinsche, Ioannou, Nietner, Haferkamp, Quek, Hangleiter, Seifert, Eisert, Sweke, Phys Rev Lett 130, 240602 (2023)

Haferkamp, Montealegre-Mora, Heinrich, Eisert, Gross, Roth, Comm Math Phys 397, 995-1041 (2023)

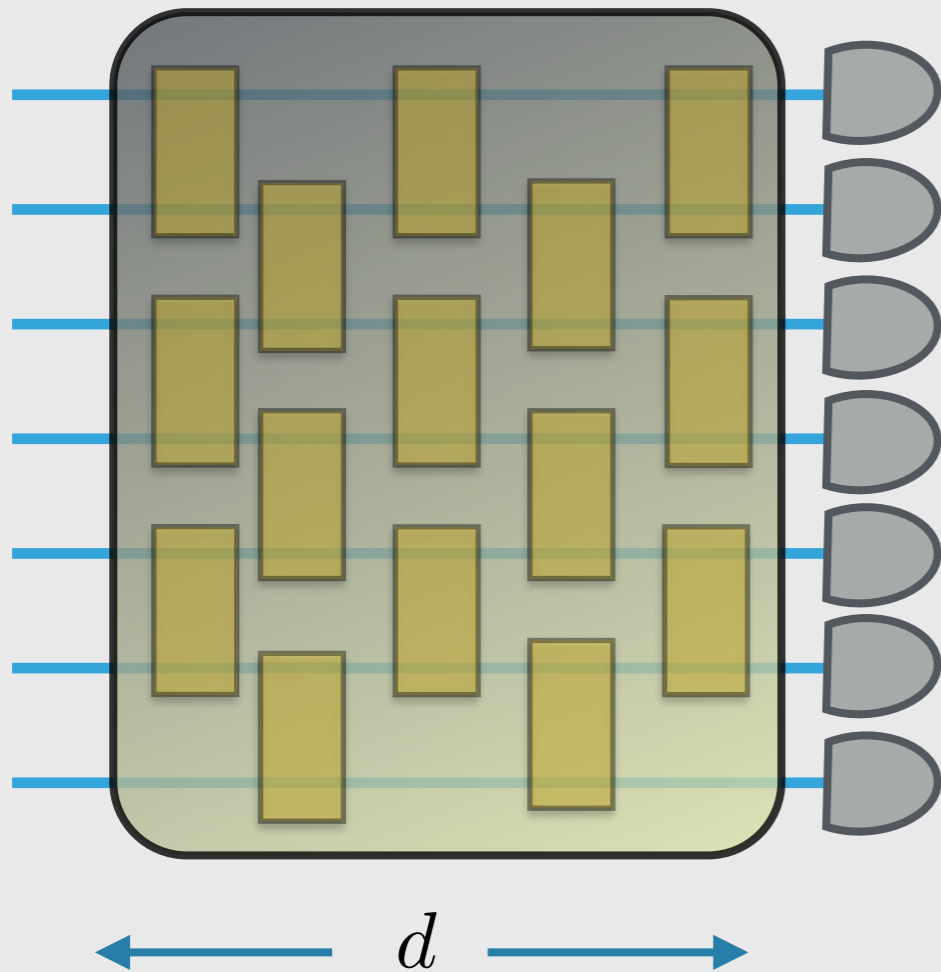
Nietner, Ioannou, Sweke, Kueng, Eisert, Hinsche, Haferkamp, arXiv:2305.05765 (2023)

Hinsche, Ioannou, Nietner, Haferkamp, Quek, Hangleiter, Seifert, Eisert, Sweke, arXiv:2110.05517 (2021)

- **Question:** Can we find a **near-term quantum-classical separation**?



- Take **output distributions of quantum circuits themselves**



- Hard to classically simulate, but to learn?
 - $d = 1$: product distributions
 - $d \rightarrow \infty$: all distributions

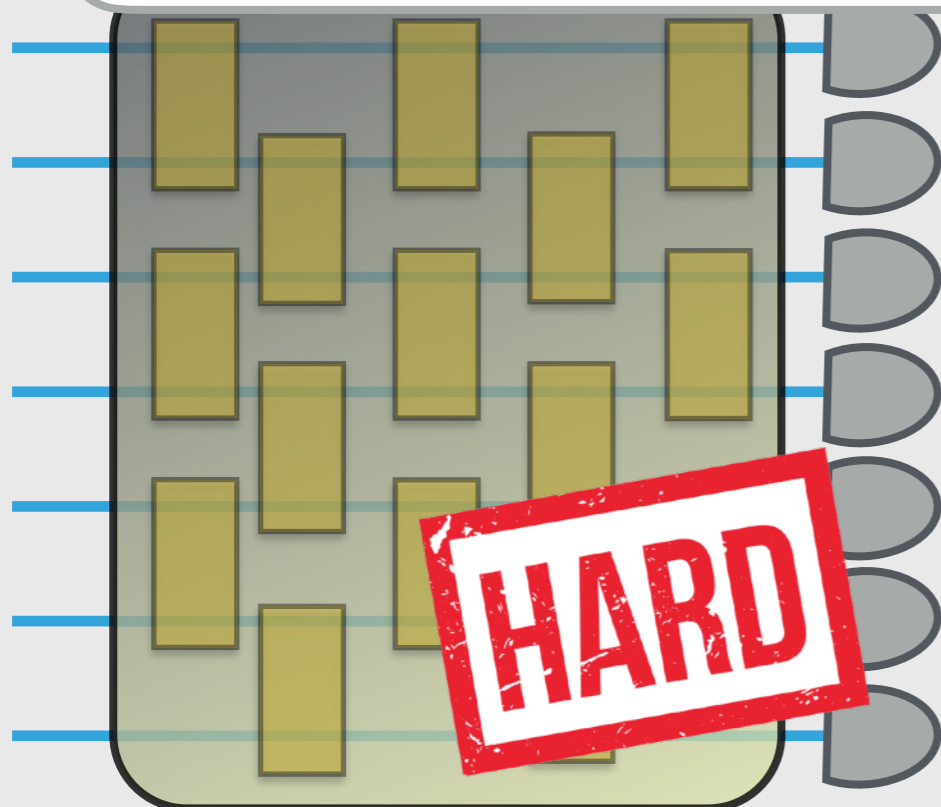
- When does **hardness** set in?

All distributions

$$\mathcal{C} = \{P_U : \text{depth} \leq d, \text{gates} \in \mathcal{G}\}$$



- **Theorem 2:** If “standard-secure PRFs” exist, then depth $d = n^{\Omega(1)}$ is **hard**, for generative and density modelling



$$d = n^{\Omega(1)}$$

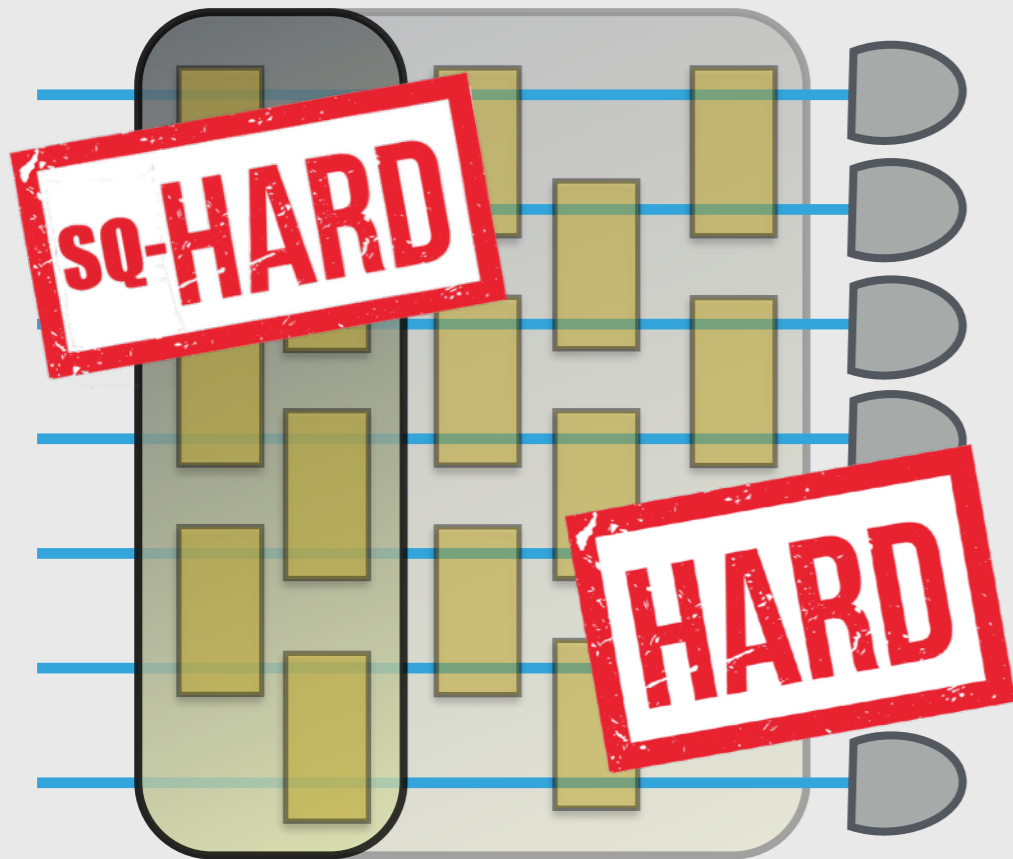
- Any (**quantum and classical**) algorithm!
- This time, PRFs cannot even be broken by quantum algorithms: **No advantage**

$$\mathcal{C} = \{P_U : d = n^{\Omega(1)}, \text{ gates universal}\}$$

Uniform over $x || \text{PRF}(x)$



- **Theorem 3:** Depth $d = \omega(\log(n))$ is **hard**, from **statistical queries**



- Think of **"generic"** learning algorithms

$d = \omega(\log(n))$ $d = n^{\Omega(1)}$

$\mathcal{C} = \{P_U : d = \omega(\log(n)), \text{ gates universal}\}$

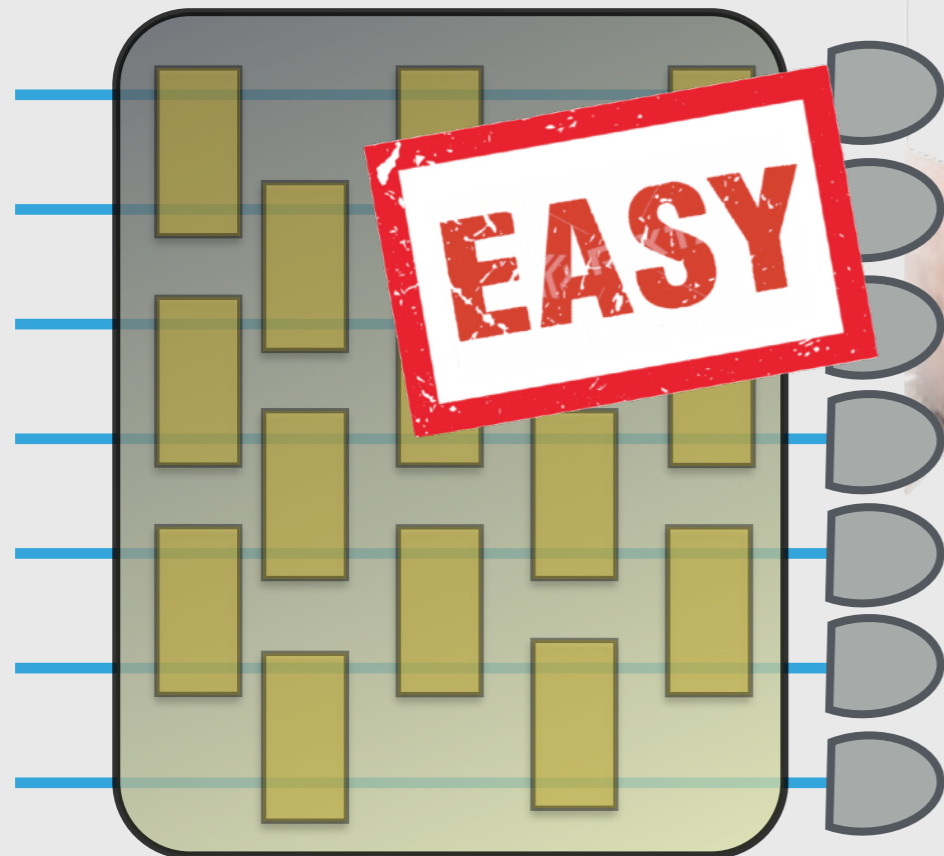
Parity distributions

Hinsche, Ioannou, Nietner, Haferkamp, Quek, Hangleiter, Seifert, Eisert, Sweke, Phys R
Hinsche, Ioannou, Nietner, Haferkamp, Quek, Hangleiter, Seifert, Eisert, Sweke, arXiv:2110

- How are **simulation** and **learning** related?



- Theorem 4:



Clifford circuits, any depth

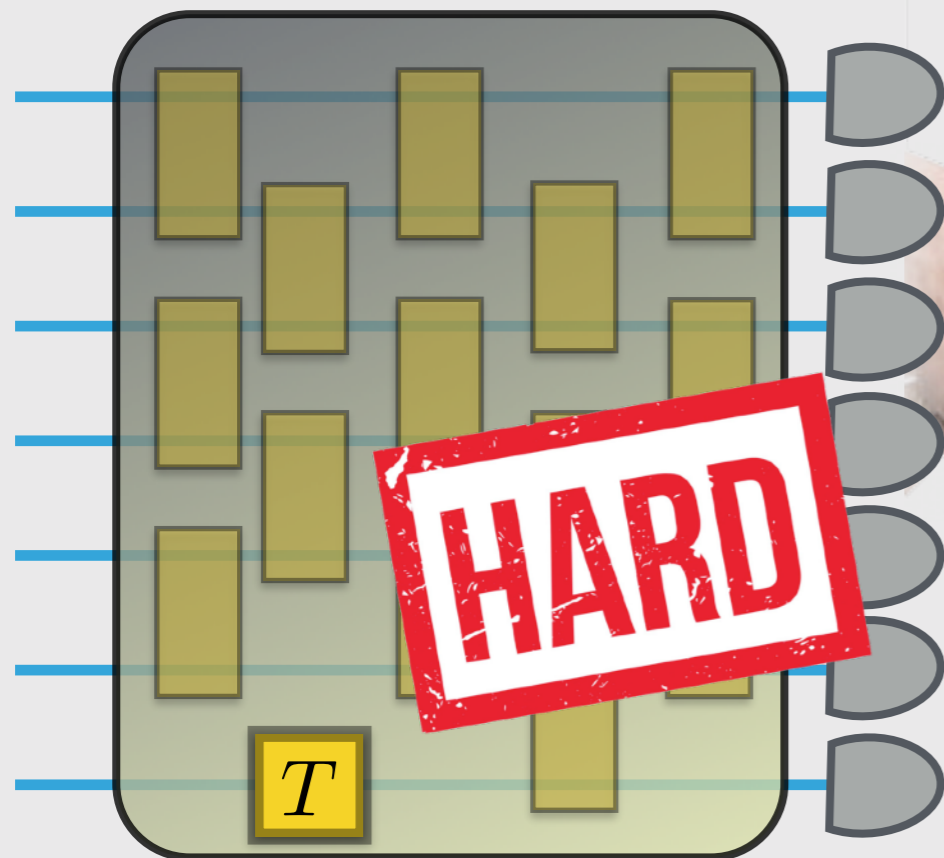


Simulation

Learning

- Clifford-circuit output distributions are uniform over affine subspaces of the finite n -dimensional vector space \mathbb{F}_2^n

- **Theorem 5:** A single T -gate renders learning hard



Simulation

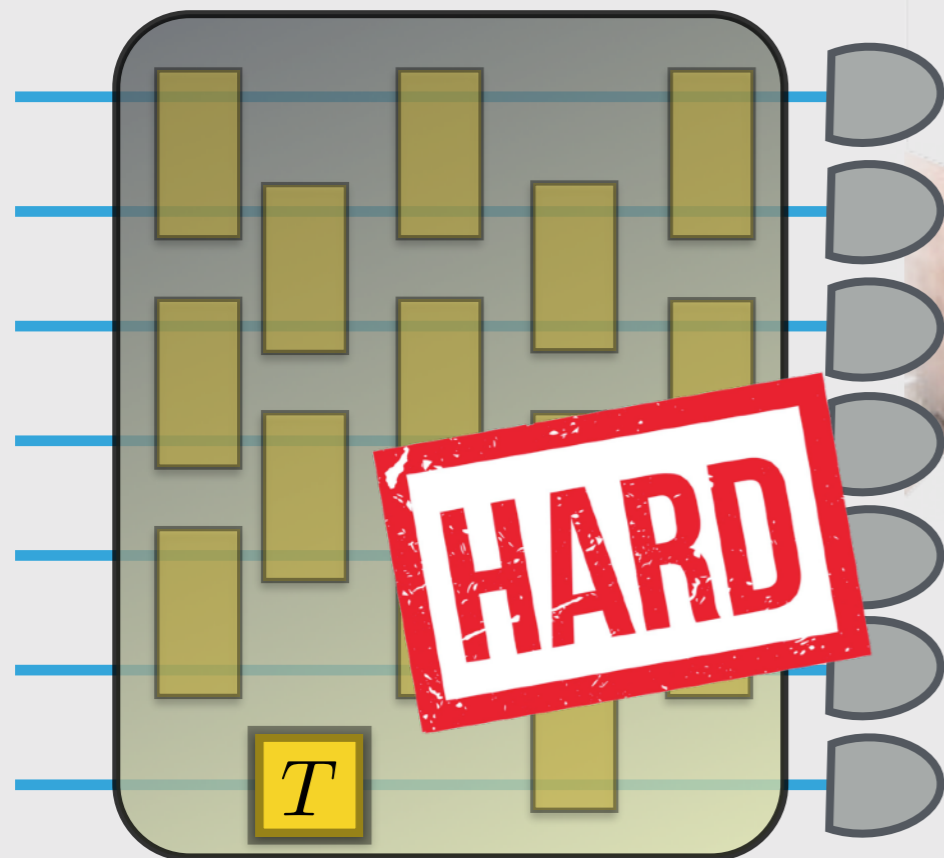
Learning

- But circuits with logarithmically many T -gates can be efficiently simulated

Clifford circuits and a single T -gate, linear depth



- **Theorem 5:** A single T -gate renders learning hard



Clifford circuits and a single T -gate, linear depth



Simulation

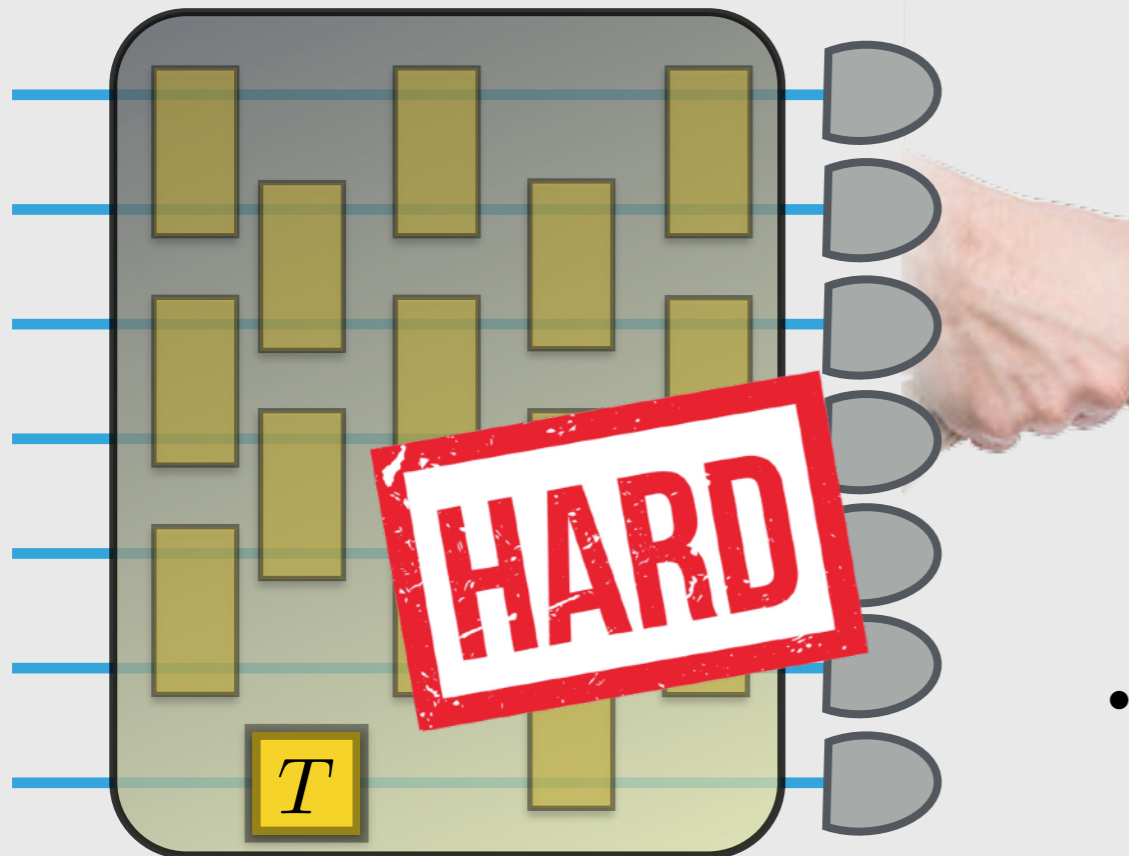
Learning



R Sweke: "Should we use parametrized quantum circuits for machine learning?"

CLASSICAL SIMULATION AND LE

- **Theorem 5: A single T -gate**



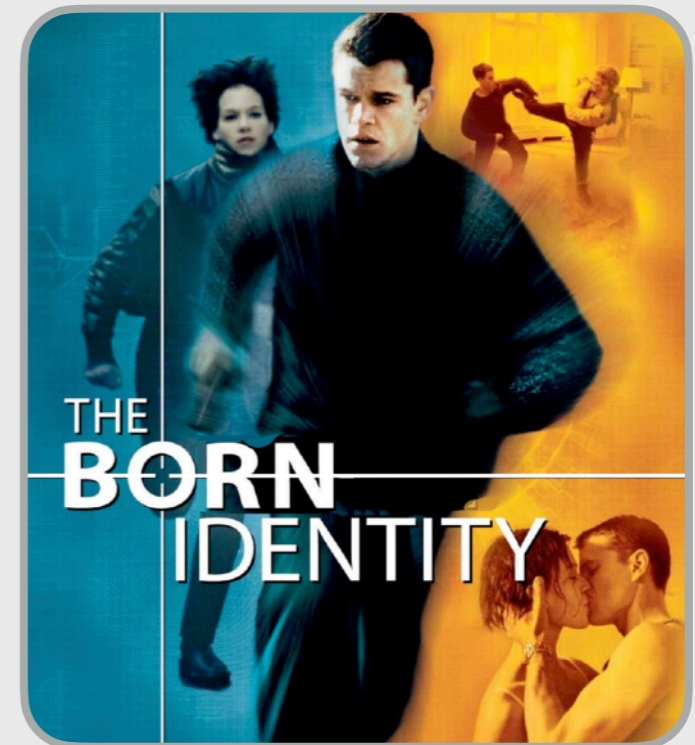
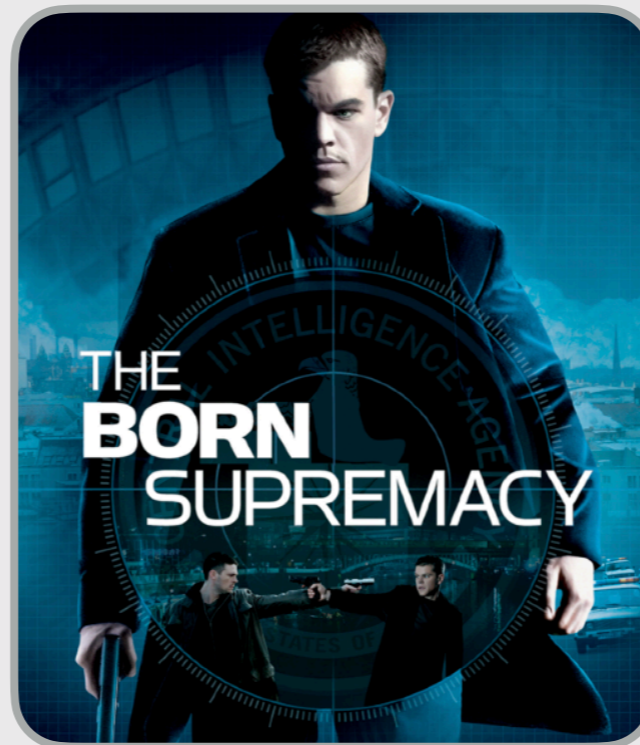
Clifford circuits and a single T -gate, linear depth

- **Random Clifford circuits** are unitary 3-designs
- T -gates uplift then to (approximate) arbitrary order designs
- **Theorem 6: A constant (!) number of T -gates is sufficient**

Haferkamp, Montaelegre-Mora, Heinrich, Eisert, Gross, Roth, Commun Math Phys 397, 995-1041 (2023)



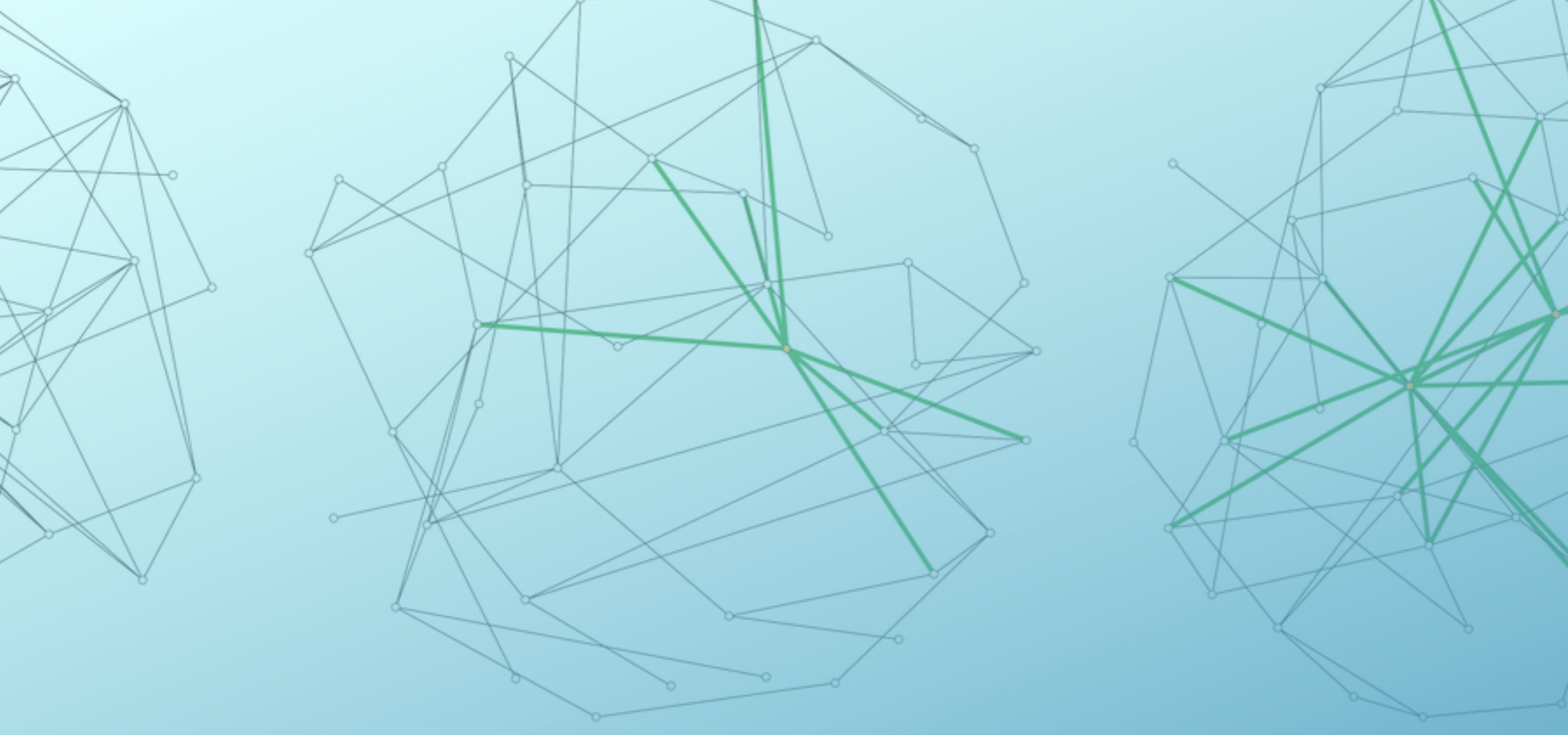
- The connection between **simulation** and **learning** is intricate
- **Obstruction** to proving **quantum advantages**
- First rigorous insights into learnability of outputs of **quantum circuits**



Hinsche, Ioannou, Nietner, Haferkamp, Quek, Hangleiter, Seifert, Eisert, Sweke, Phys Rev Lett 130, 240602 (2023)

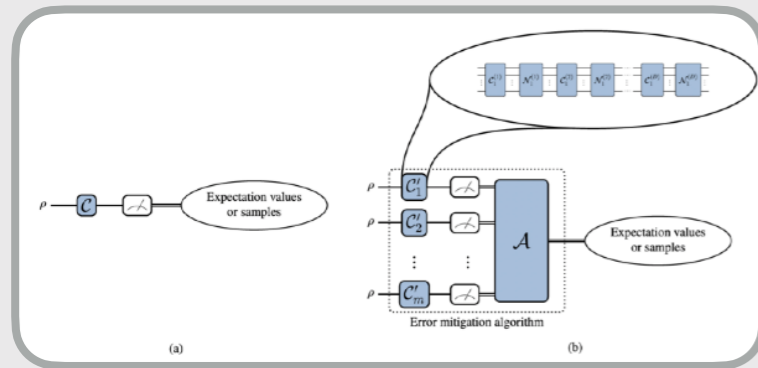
Hinsche, Ioannou, Nietner, Haferkamp, Quek, Hangleiter, Seifert, Eisert, Sweke, arXiv:2110.05517 (2021)

Compare Coyle, Mills, Danos, Kashefi, npj Quant Inf 6, 1 (2020)



CODA ON QUANTUM ADVANTAGES FOR COMBINATORIAL OPTIMIZATION

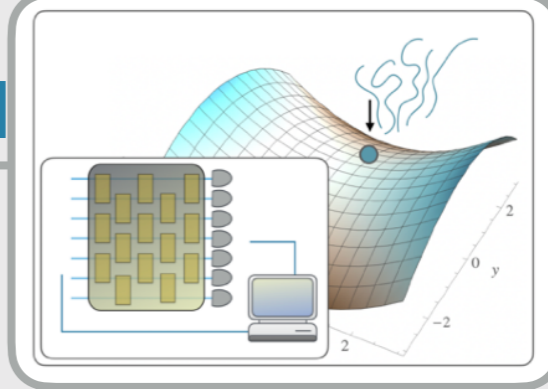
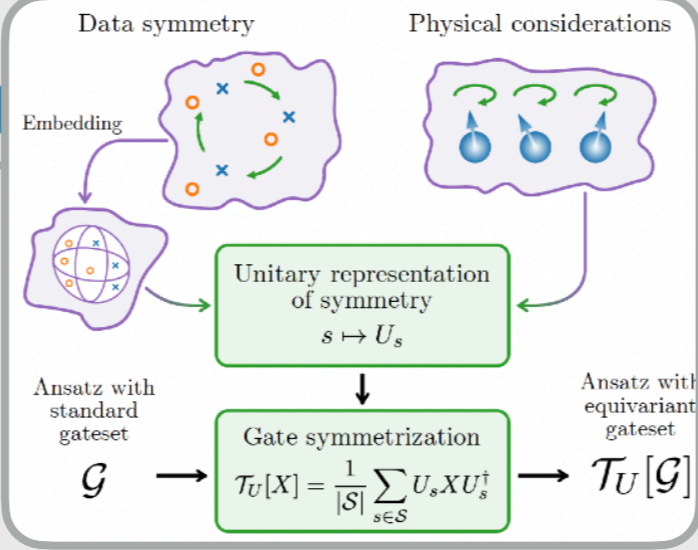
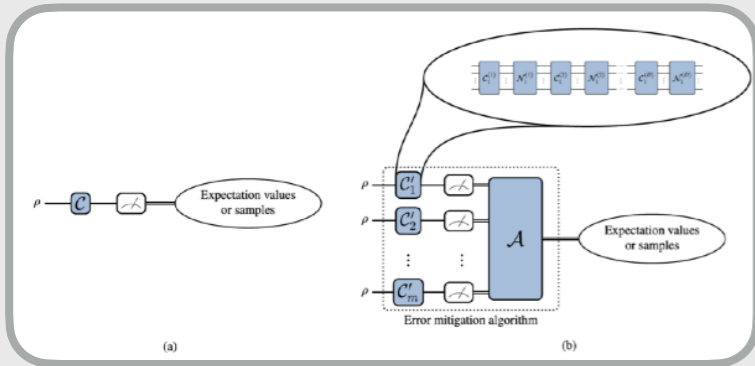
Pirnay, Ulitzsch, Wilde, Eisert, Seifert, arXiv:2212.08678 (2022)



Y Quek: "The signal and the noise:
Learning with random quantum circuits
and other agents of noise"

- Strong **limitations** for **quantum error mitigation**

Quek, Franca, Khatri, Meyer, Eisert,
arXiv:2210.11505 (2022)



- **Noise** can help in variational quantum algorithms

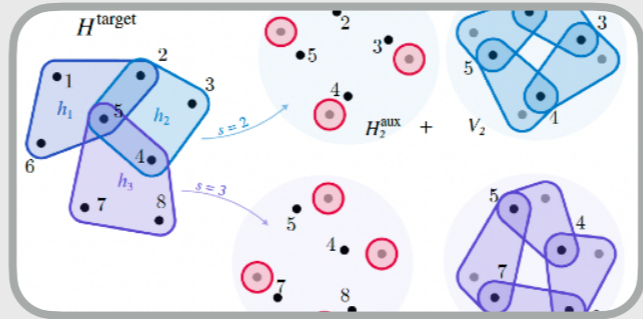
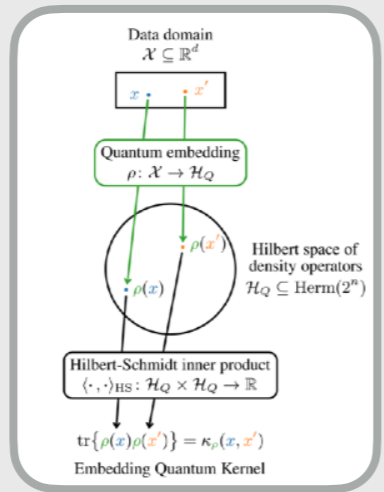
Liu, Wilde, Mele, Jiang, Eisert, arXiv:2210.06723 (2022)

- **Strong limitations** for quantum error mitigation

Quek, Franca, Khatri, Meyer, Eisert, arXiv:2210.11505 (2022)

- **Symmetry** helps

Meyer, Mularski, Gil-Fuster, Mele, Arzani, Wilms, Eisert, PRX Quantum 4, 010328 (2023)

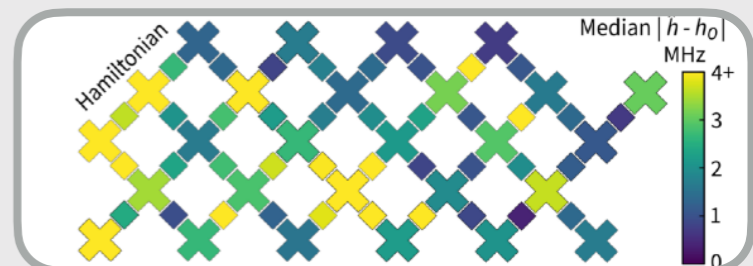


- **Expressivity** of embedding quantum kernels

Gil-Fuster, Eisert, Dunjko, arXiv:2303.14419 (2023)

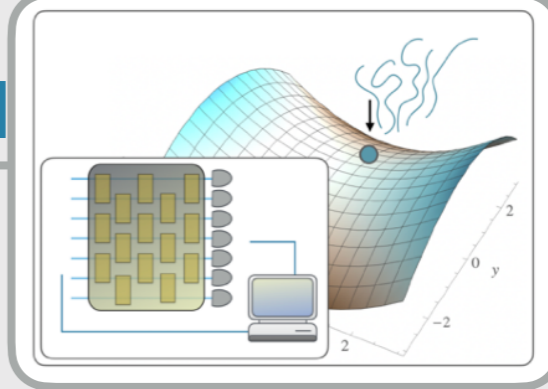
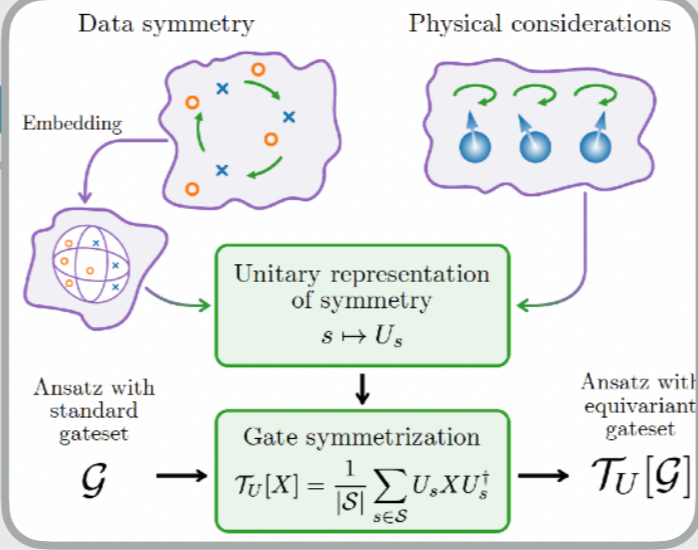
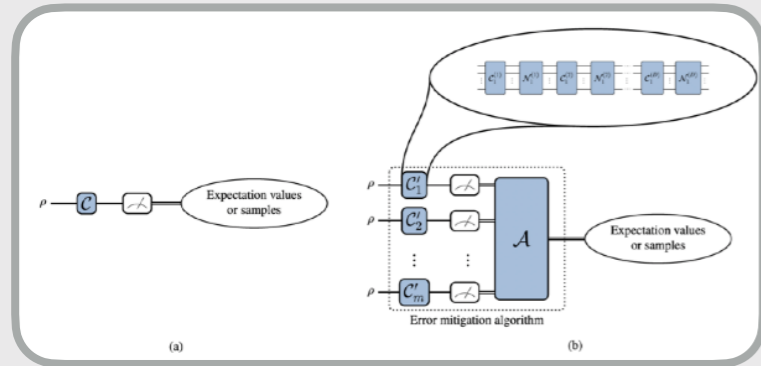
- **Gadgets** against barren plateaus

Simon, Faehrmann, Khatri, Eisert, arXiv:2210.03099 (2022)



- **Hamiltonian** learning

Wilde, Kshetrimayum, Roth, Hangleiter, Sweke, Eisert, arXiv:2209.14328 (2022)
Hangleiter, Roth, Eisert, Roushan, arXiv:2108.08319 (2021)



- **Noise** can help in variational quantum algorithms

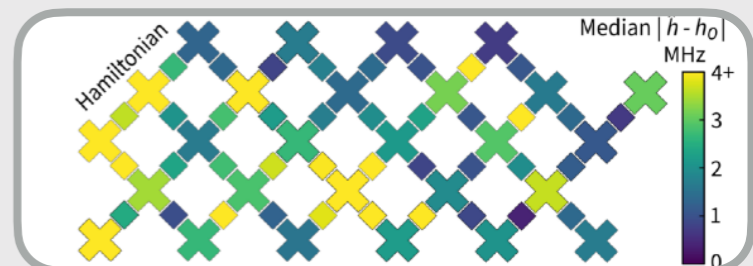
Liu, Wilde, Mele, Jiang, Eisert, arXiv:2210.06723 (2022)

- **Strong limitations** for quantum error mitigation

Quek, Franca, Khatri, Meyer, Eisert, arXiv:2210.11505 (2022)

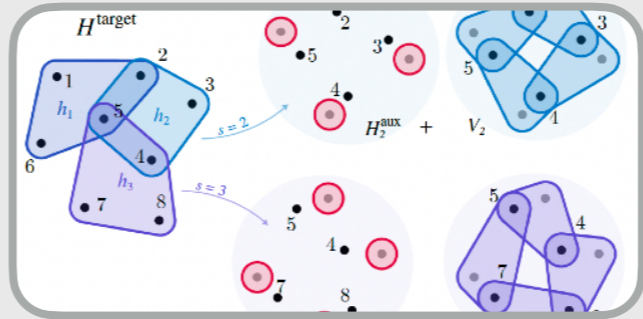
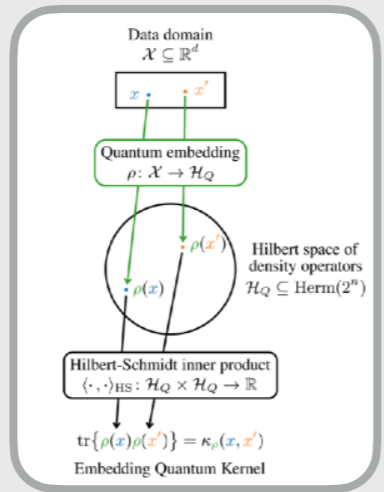
- **Symmetry** helps

Meyer, Mularski, Gil-Fuster, Mele, Arzani, Wilms, Eisert, PRX Quantum 4, 010328 (2023)



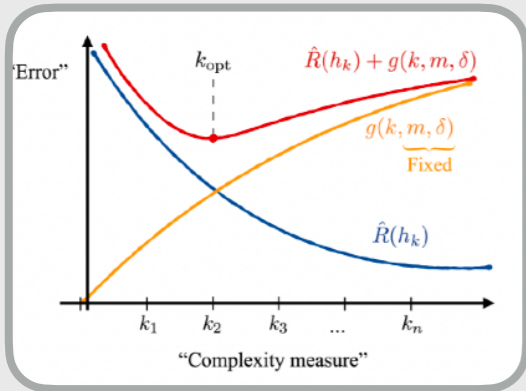
- **Hamiltonian** learning

Wilde, Kshetrimayum, Roth, Hangleiter, Sweke, Eisert, arXiv:2209.14328 (2022)
 Hangleiter, Roth, Eisert, Roushan, arXiv:2108.08319 (2021)



- **Gadgets** against barren plateaus

Simon, Faehrmann, Khatri, Eisert, arXiv:2210.03099 (2022)

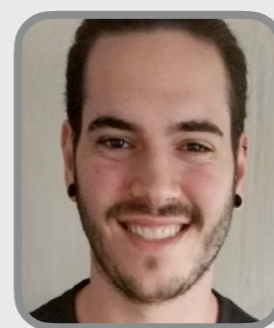


- **Expressivity** of embedding quantum kernels

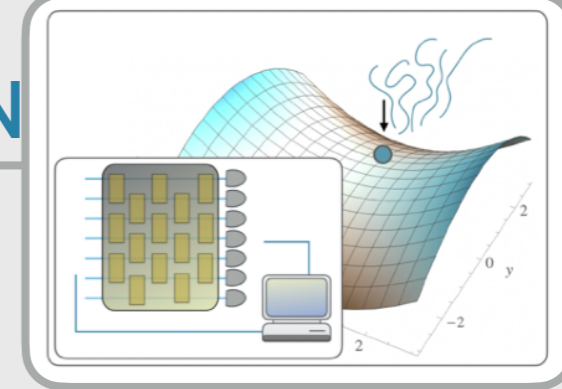
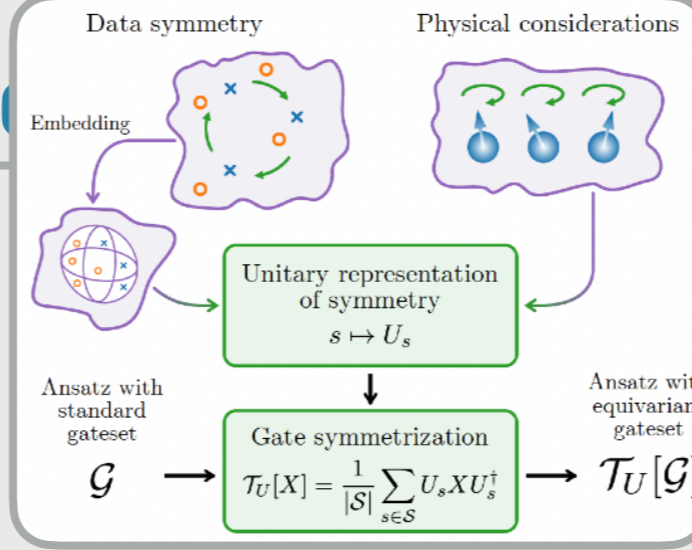
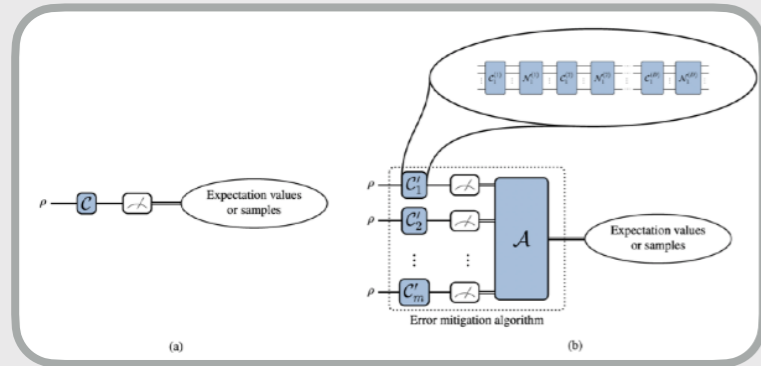
Gil-Fuster, Eisert, Dunjko, arXiv:2303.14419 (2023)

- **Generalization** bounds for parametrized quantum circuits

Caro, Gil-Fuster, Meyer, Eisert, Sweke, Quantum 5, 582 (2021)
 Gil-Fuster, Eisert, Bravo-Prieto, arXiv:2306.13461 (2023)



C Bravo Prieto: "Understanding quantum machine learning also requires rethinking generalization"



- **Noise** can help in variational quantum algorithms

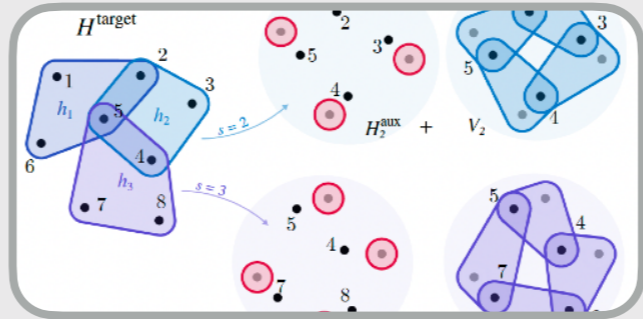
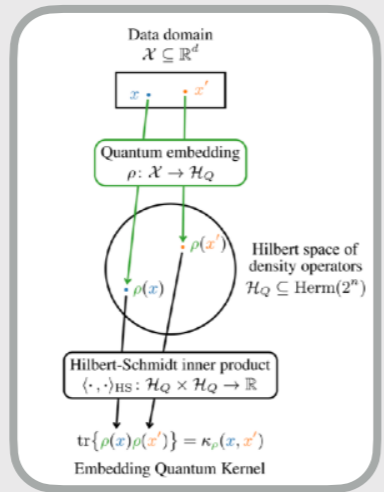
Liu, Wilde, Mele, Jiang, Eisert, arXiv:2210.06723 (2022)

- **Strong limitations** for quantum error mitigation

Quek, Franca, Khatri, Meyer, Eisert, arXiv:2210.11505 (2022)

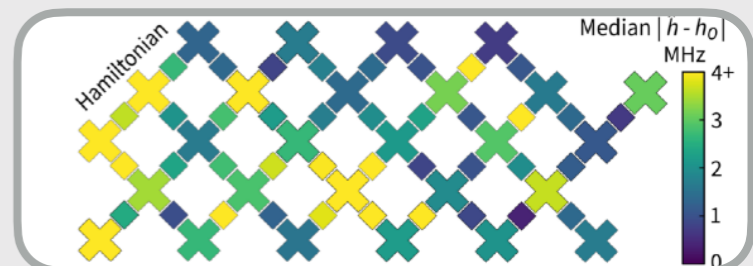
- **Symmetry** helps

Meyer, Mularski, Gil-Fuster, Mele, Arzani, Wilms, Eisert, PRX Quantum 4, 010328 (2023)



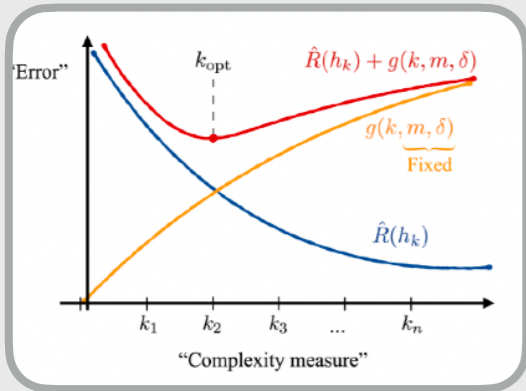
- **Gadgets** against barren plateaus

Simon, Faehrmann, Khatri, Eisert, arXiv:2210.03099 (2022)



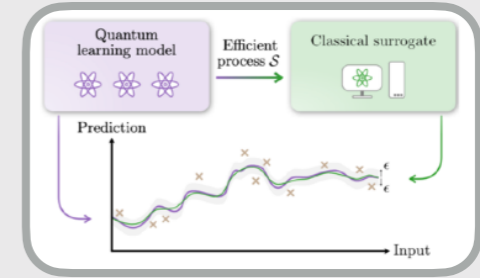
- **Hamiltonian** learning

Wilde, Kshetrimayum, Roth, Hangleiter, Sweke, Eisert, arXiv:2209.14328 (2022)
Hangleiter, Roth, Eisert, Roushan, arXiv:2108.08319 (2021)



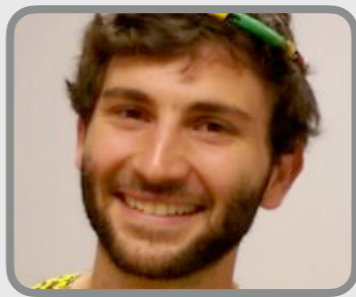
- **Expressivity** of embedding quantum kernels

Gil-Fuster, Eisert, Dunjko, arXiv:2303.14419 (2023)



- **Classical surrogates** for quantum learning models

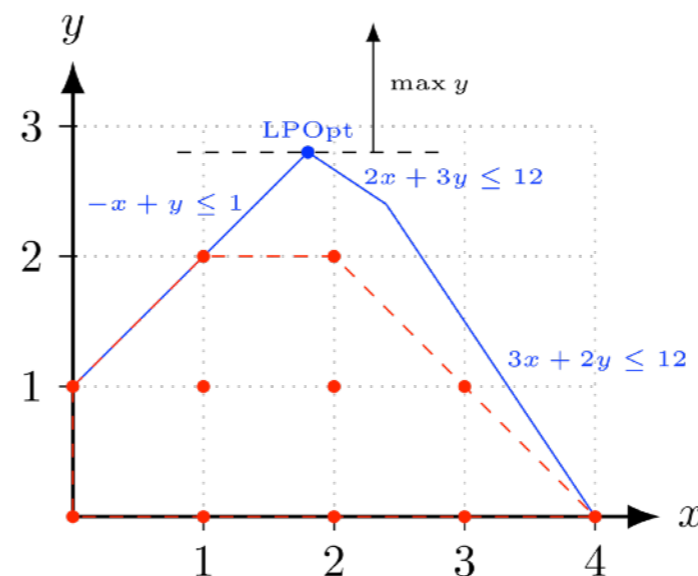
Schreiber, Eisert, Meyer, Phys Rev Lett 131, 100803 (2023)
Sweke, Recio, Jerbi, Gil-Fuster, Fuller, Eisert, Meyer, arXiv:2309.11647 (2023)





- Can we hope to see advantages for **combinatorial optimization**?
- Commonly **NP-hard** in worst case complexity
- E.g., **integer programming**

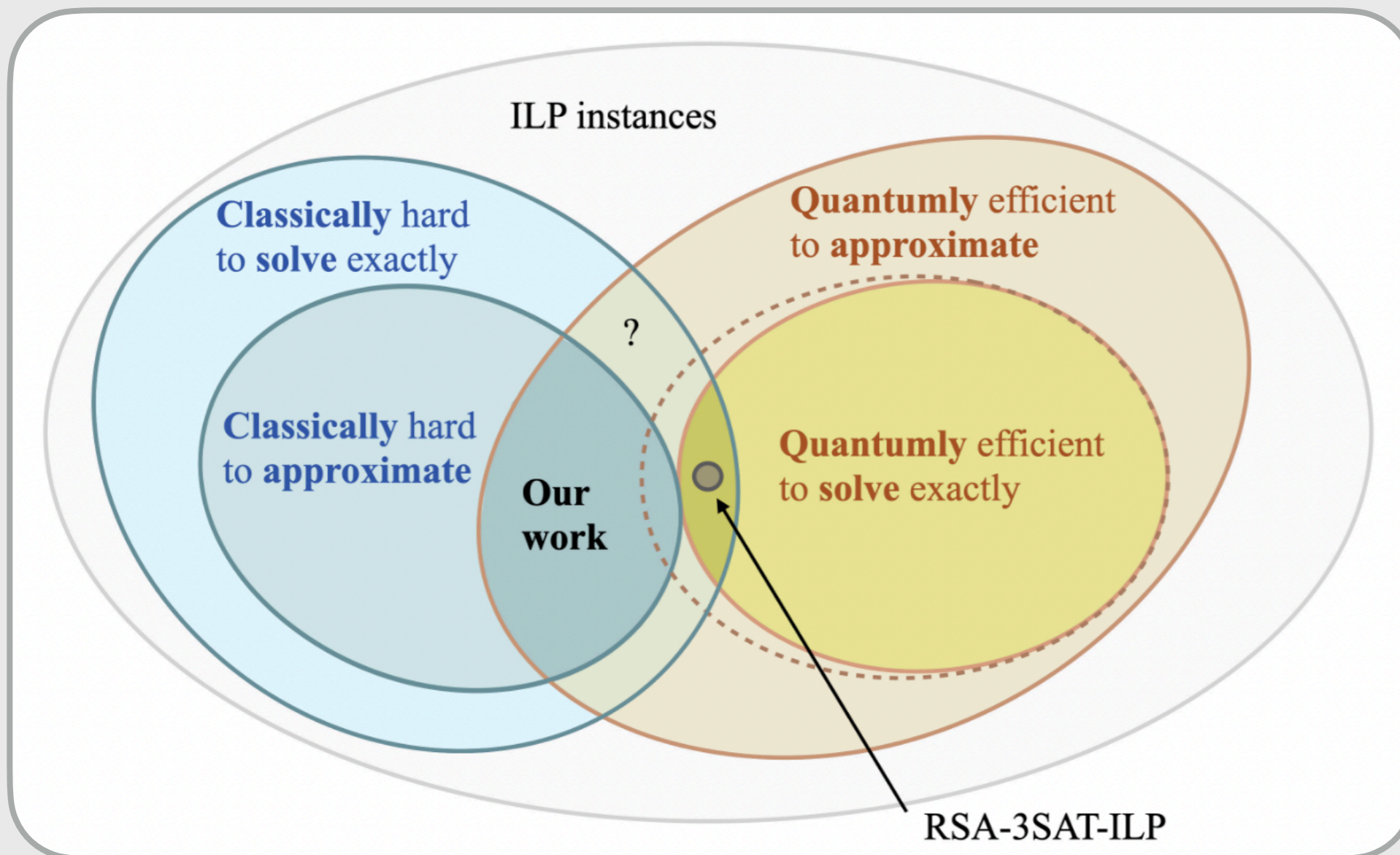
maximize $\mathbf{c}^T \mathbf{x}$
subject to $A\mathbf{x} \leq \mathbf{b}$,
 $\mathbf{x} \geq \mathbf{0}$,
and $\mathbf{x} \in \mathbb{Z}^n$



- In what sense can there be **quantum advantages**?



- A quantum computer can **approximate hard instances**, originating from the hardness of inverting the RSA function, in polynomial time





- A quantum computer can **approximate hard instances**, originating from the hardness of inverting the RSA function, in polynomial time



- So there is a **super-polynomial speed-up!**
- Can be translated to variational **Hamiltonian problems?**
- **Constructive:** Find classically "hard instances"
- Connection to **QAOA?**



OUTLOOK

OUTLOOK



- Can we hope noisy, realistic quantum devices to provide a **speedup over classical computers**?



- Yes, find **super-polynomial advantages in PAC-learning!**

- A little noise can be a **good thing!**

- We have found **super-polynomial advantages in optimization!**



- Well, but how much **structure** do we need?

- A **single-T gate** renders learning hard

- But **error mitigation already fails** for log-log deep circuits

- Let us see ...

OUTLOOK



- Can we hope noisy, realistic quantum devices to provide a **speedup over classical computers?**

Definitely Maybe

- Yes, advanced

- A little noise can be a **good thing!**

- We have found **super-polynomial**

- But **error mitigation already fails** for log-log deep circuits

THANKS FOR YOUR ATTENTION

WE ARE HIRING