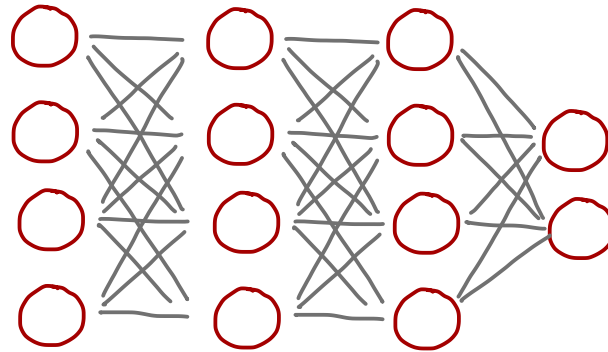


Quantum Kolmogorov - Arnold Networks

Petr Ivashkov , Po-Wei Huang, Kelvin Koor
Lirandë Pira , Patrick Rebentrost



[arxiv.org / abs / 2410.04435](https://arxiv.org/abs/2410.04435)

Talk outline

5 min { Kolmogorov - Arnold Theorem
Kolmogorov - Arnold Networks

10 min { Interlude : Quantum linear algebra

25 min { QKAN
o Definition & construction
o Solution models
o Parametrization
o Possible application

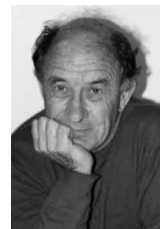
A brief history of Kolmogorov - Arnold

1900

1956

13th Problem

Kolmogorov - Arnold Theorem



Theorem: f continuous \Rightarrow

$$\exists \{\varphi_{pq}\}, \{g_q\} \text{ such that}$$

$$f(\vec{x}) = \sum_{q=1}^{2N+1} g_q \left(\sum_{p=1}^N \varphi_{pq}(x_p) \right)$$

$$f(a, b, c) = g_1 \circ g_2(a, b) + g_3 \circ g_4(b, c) + \dots$$

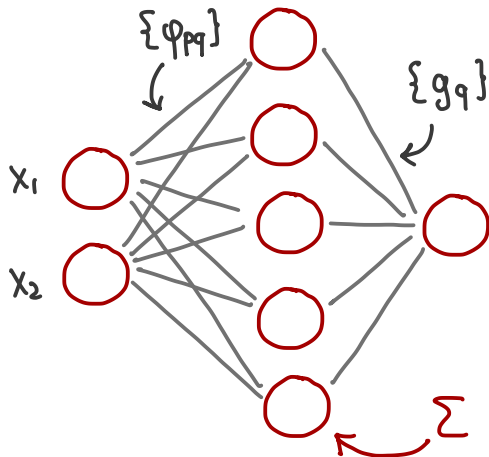
A brief history of Kolmogorov - Arnold

1956

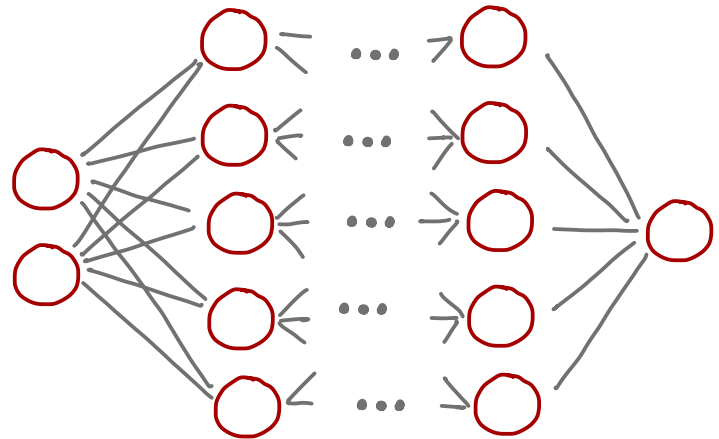
2024



Kolmogorov - Arnold
Theorem

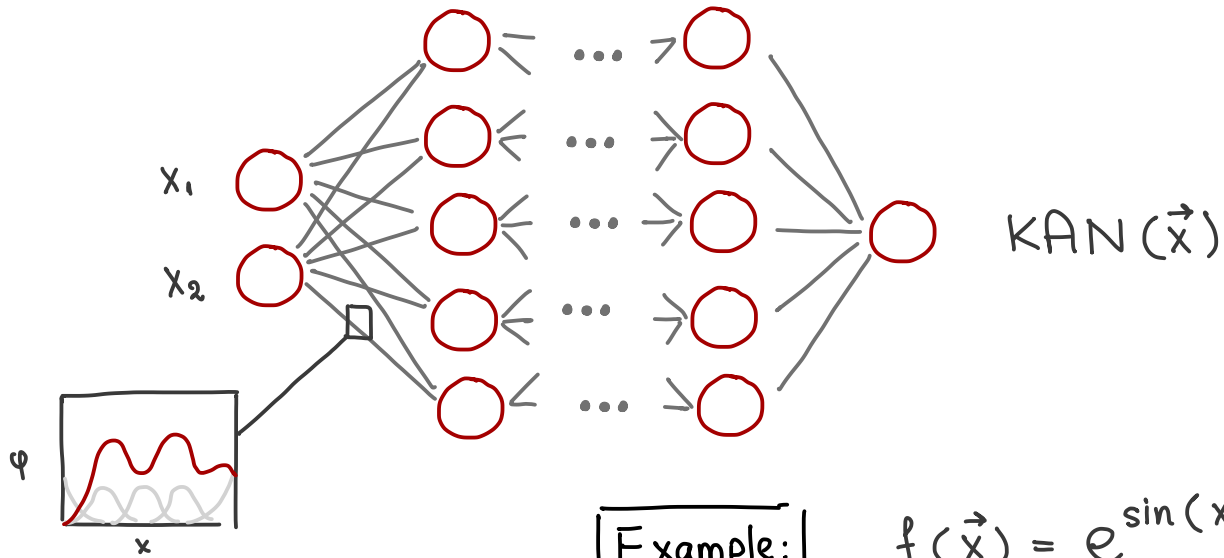


Kolmogorov - Arnold
Networks



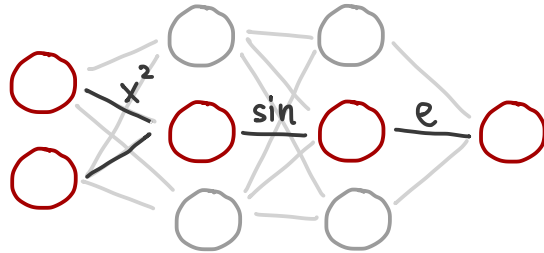
arxiv : 2404.19756 (Liu et al.)

Kolmogorov - Arnold Networks



Example: $f(\vec{x}) = e^{\sin(x_1^2 + x_2^2)}$

$$\varphi_{pq}(x) = \sum_{r=1}^d w_{pq}^{(r)} b^{(r)}$$



Interlude: Quantum linear algebra



- o Can we handle nonunitary matrices?
- o Yes, embed it into a larger unitary.

$$\underset{\text{nonunitary}}{A} \rightarrow \begin{matrix} \langle 0 | \\ \hat{0} \end{matrix} \left(\begin{array}{cc} A & * \\ * & * \end{array} \right) = \underset{\text{unitary}}{U_A} = |0\rangle\langle 0| \otimes A + \dots$$

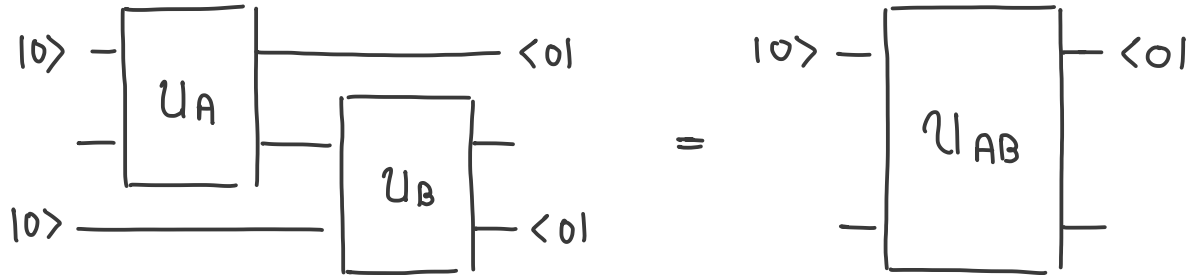
$$\text{shape}(A) = N \times N$$

$$\# \text{ qubits to represent } A : n = \log_2 N$$

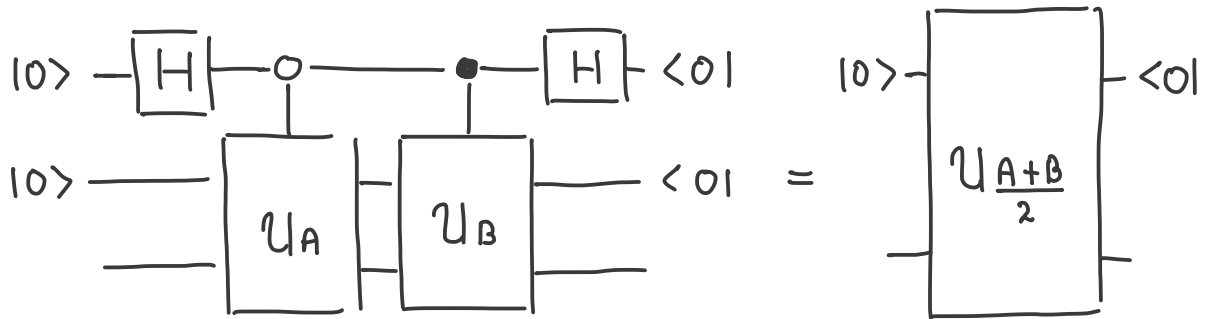
$$\# \text{ qubits to represent } U_A : n + a$$

Multiplication & addition

$A \cdot B$



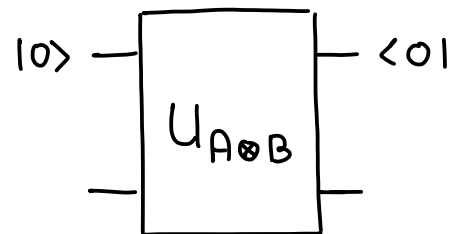
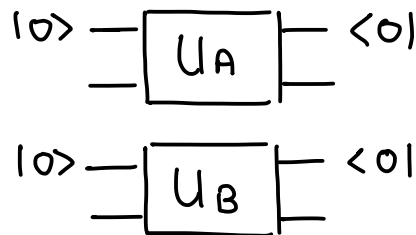
$A + B$
(LCU)



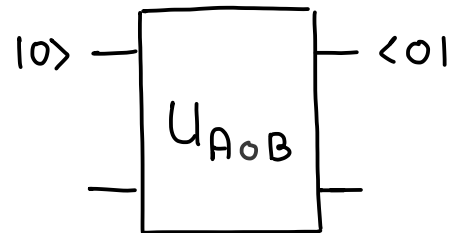
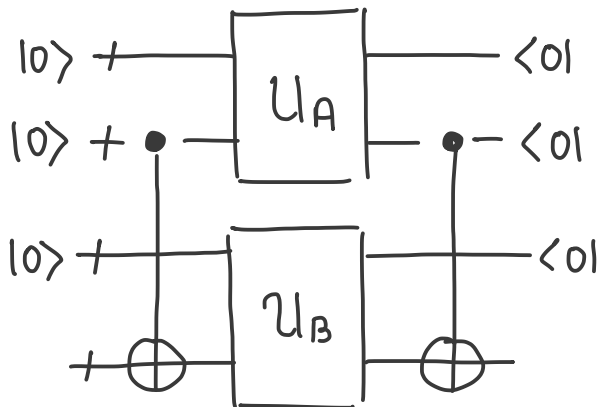
arxiv: 1806.01838
(Gilyén et al.)

Tensor & Hadamard product

$A \otimes B$



$A \circ B$



arxiv: 2402.16714
(Guo et al.)

General matrix functions $f(A)$

Rotation matrix $O(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$

$$O^r(\theta) = \begin{pmatrix} \cos r\theta & -\sin r\theta \\ \sin r\theta & \cos r\theta \end{pmatrix} \quad \boxed{\begin{array}{l} x := \cos \theta \\ \theta = \arccos x \end{array}}$$

$$O^r(x) = \begin{pmatrix} \cos(r \cdot \arccos x) & * \\ * & * \end{pmatrix} = \begin{pmatrix} T_r(x) & * \\ * & * \end{pmatrix}$$

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_2(x) = 2x^2 - 1$$

arxiv: 2201.08309
(Lin Lin)

↑ a few Chebyshev polynomials

General matrix functions $f(A)$

Let U_x be a (Hermitian) block-encoding of x :

$$U_x = \begin{pmatrix} x & * \\ * & * \end{pmatrix} \quad \& \quad U_x = U_x^\dagger$$

Then, $U_x |0\rangle = x |0\rangle + \sqrt{1-x^2} |1\rangle$

and
$$\begin{aligned} U_x^\dagger U_x |0\rangle &= x U_x^\dagger |0\rangle + \sqrt{1-x^2} U_x^\dagger |1\rangle \\ &= x^2 |0\rangle + x \sqrt{1-x^2} |1\rangle \\ &\quad + \sqrt{1-x^2} U_x |1\rangle = |0\rangle \end{aligned}$$

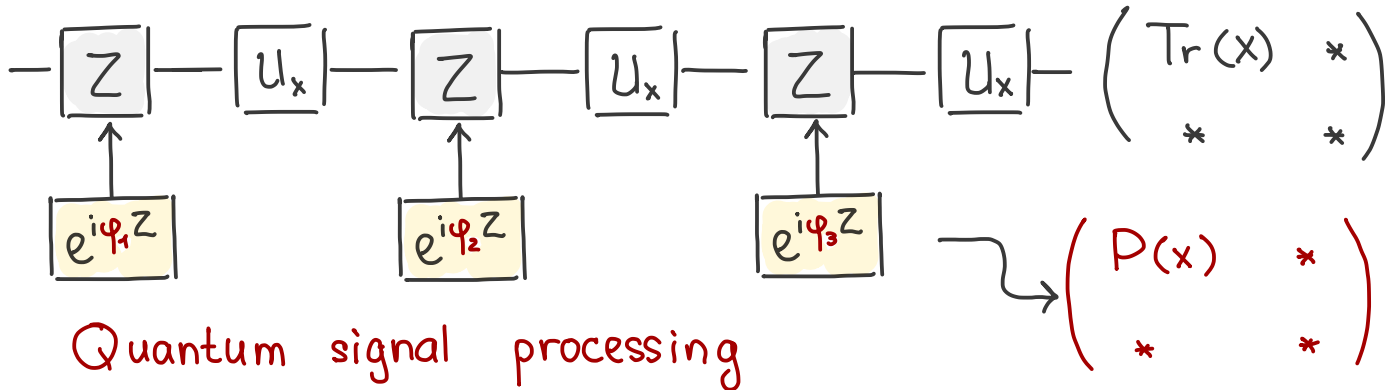
$$\Rightarrow U_x |1\rangle = \sqrt{1-x^2} |0\rangle - x |1\rangle$$

General matrix functions $f(A)$

$$U_x |0\rangle = x |0\rangle + \sqrt{1-x^2} |1\rangle$$

$$U_x |1\rangle = \sqrt{1-x^2} |0\rangle - x |1\rangle$$

$$\Rightarrow U_x = \begin{pmatrix} x & \sqrt{1-x^2} \\ \sqrt{1-x^2} & -x \end{pmatrix} \neq \begin{pmatrix} \overbrace{x}^{\cos} & \overbrace{-\sqrt{1-x^2}}^{-\sin} \\ \underbrace{\sqrt{1-x^2}}_{\sin} & \underbrace{x}_{\cos} \end{pmatrix}$$



General matrix functions $f(A)$

Idea: Replace x by A and

$$f(A) := \sum_{\lambda} f(\lambda) |\lambda\rangle\langle\lambda|$$

Theorem: **QSVT** Let U_A be a block-encoding of A . Let $P(x) \in \mathbb{R}[x]$ of degree d . Then, there is a sequence of angles $\tilde{\varphi}_i$ such that

$$(-i)^d e^{i\tilde{\varphi}_0 Z_\pi} \prod_{j=1}^{d/2} (U_A^\dagger e^{i\tilde{\varphi}_{2j-1} Z_\pi} U_A e^{i\tilde{\varphi}_{2j} Z_\pi})$$

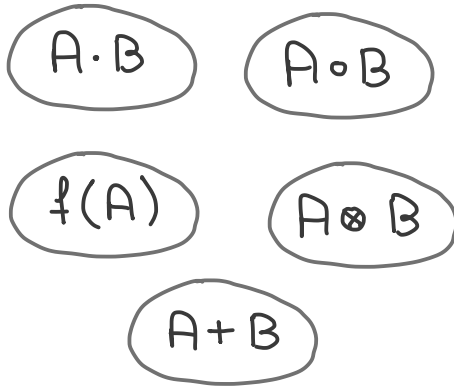
is a block-encoding of $P(A)$.

General matrix functions $f(A)$

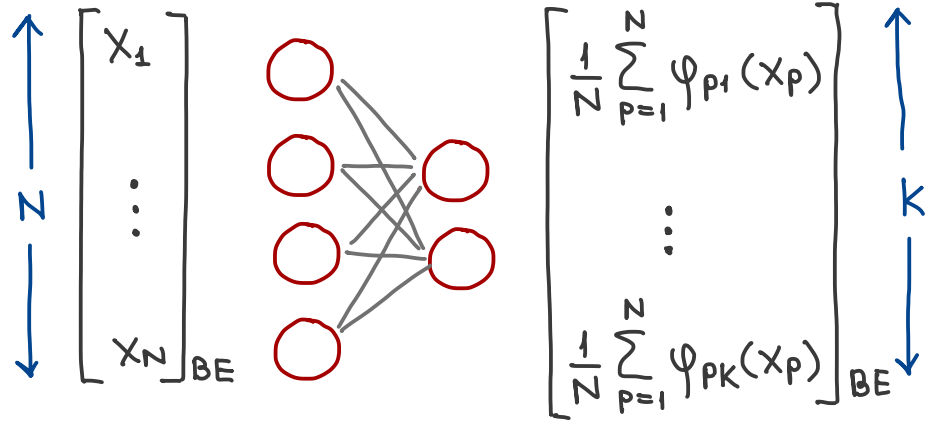
Examples:

- $f(H) = e^{-iHt}$: Hamiltonian simulation
- $f(A) = A^{-1}$: Linear system solver
($A|x\rangle = |b\rangle \Rightarrow |x\rangle = A^{-1}|b\rangle$)
- $f(H) = e^{-\beta H}$: Gibbs state preparation

Toolbox



Goal: Quantum KAN



$$\psi_{pq}(x) = \frac{1}{d+1} \sum_{r=0}^d w_{pq}^{(r)} T_r(x)$$

Chebyshev polynomials: $T_r(x) := \cos(r \arccos(x))$

QKAN construction

①
$$\begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}_{BE} \otimes \mathbb{1}_{\log K} = \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_1 \\ \vdots \\ x_N \\ \vdots \\ x_N \end{bmatrix}_{BE} \right\}_K$$

Diagram illustrating the construction of the input state for the QKAN circuit. The input is a vector $\begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}_{BE}$ of size n (indicated by a red bracket). This is tensored with the identity $\mathbb{1}_{\log K}$ to produce a set of K identical copies of the input vector, each of size $n + k$ (indicated by a red bracket). The diagram shows two such copies: the top one is labeled $|0\rangle +$ and the bottom one is labeled $k +$. Both are inputs to a unitary U_x , which produces outputs $\langle 0|$.

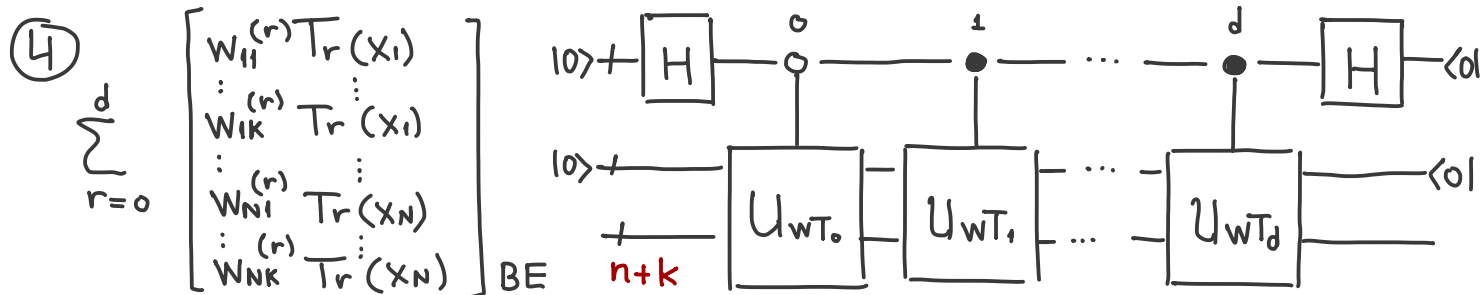
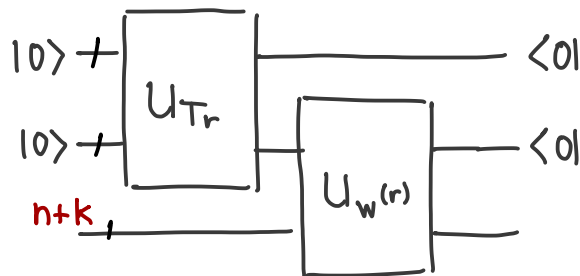
②
$$\begin{bmatrix} x_1 \\ \vdots \\ x_1 \\ \vdots \\ x_N \\ \vdots \\ x_N \end{bmatrix}_{BE} \rightarrow \begin{bmatrix} \text{Tr}(x_1) \\ \vdots \\ \text{Tr}(x_1) \\ \vdots \\ \text{Tr}(x_N) \\ \vdots \\ \text{Tr}(x_N) \end{bmatrix}_{BE}$$

Diagram illustrating the construction of the input state for the QKAN circuit. The input is a vector $\begin{bmatrix} x_1 \\ \vdots \\ x_1 \\ \vdots \\ x_N \\ \vdots \\ x_N \end{bmatrix}_{BE}$ of size n (indicated by a red bracket). This is transformed into a vector of traces $\begin{bmatrix} \text{Tr}(x_1) \\ \vdots \\ \text{Tr}(x_1) \\ \vdots \\ \text{Tr}(x_N) \\ \vdots \\ \text{Tr}(x_N) \end{bmatrix}_{BE}$ of size r (indicated by a red bracket). The diagram shows the input state $|0\rangle$ (indicated by a red bracket) entering a circuit. The circuit consists of a $-Z$ gate followed by a unitary U_x . The output is a vector of size r (indicated by a red bracket) with elements $\langle 0|$.

QKAN construction

③

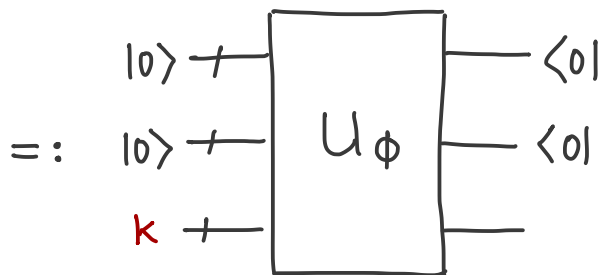
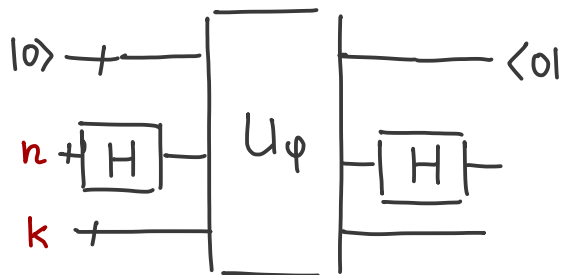
$$\begin{bmatrix} \text{Tr}(x_1) \\ \vdots \\ \text{Tr}(x_i) \\ \vdots \\ \text{Tr}(x_N) \\ \vdots \\ \text{Tr}(x_N) \end{bmatrix}_{\text{BE}} \cdot \begin{bmatrix} W_{11}^{(r)} \\ \vdots \\ W_{1K}^{(r)} \\ \vdots \\ W_{N1}^{(r)} \\ \vdots \\ W_{NK}^{(r)} \end{bmatrix}_{\text{BE}}$$



QKAN construction

⑤

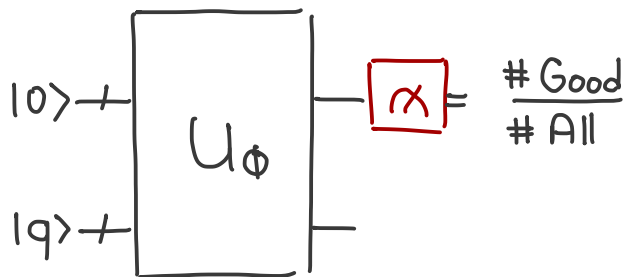
$$\begin{bmatrix} \psi_{11}(x_1) \\ \vdots \\ \psi_{1K}(x_1) \\ \vdots \\ \psi_{N1}(x_N) \\ \vdots \\ \psi_{NK}(x_N) \end{bmatrix}_{BE} \rightarrow \begin{bmatrix} \frac{1}{N} \sum_{p=1}^N \psi_{p1}(x_p) \\ \vdots \\ \frac{1}{N} \sum_{p=1}^N \psi_{pK}(x_p) \end{bmatrix}_{BE} =: \Phi(\vec{x})$$



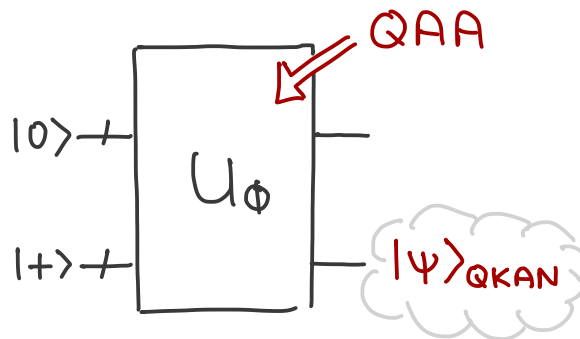
Solution models

How do I get something useful
out of the block-encoding?

Estimate $\bigoplus (\vec{x})_q \pm \epsilon$



Prepare $|\psi\rangle_\Phi$



Solution models

I Classical output : Estimate $\Phi(\vec{x})_q \pm \varepsilon$

$$U_\Phi |0\rangle |q\rangle = \underbrace{\Phi(\vec{x})_q}_{\text{Good}} |0\rangle |q\rangle + \underbrace{\sqrt{1 - \Phi^2(\vec{x})}}_{\text{Bad}} |1\rangle |q\rangle$$

- $P[\text{Good}] = |\Phi(\vec{x})_q|^2 \longrightarrow O\left(\frac{1}{\varepsilon^2}\right)$ samples
- What about the sign? Hadamard test
- QAE to improve from $O\left(\frac{1}{\varepsilon^2}\right)$ to $O\left(\frac{1}{\varepsilon}\right)$

Solution models

II Quantum output : Prepare $|\psi\rangle_\Phi$

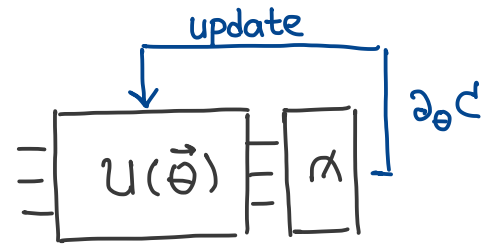
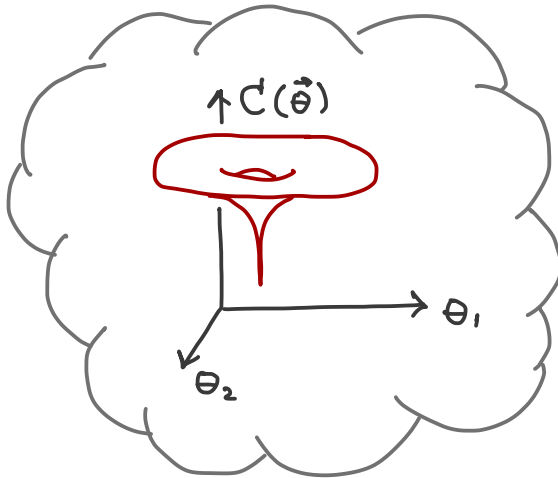
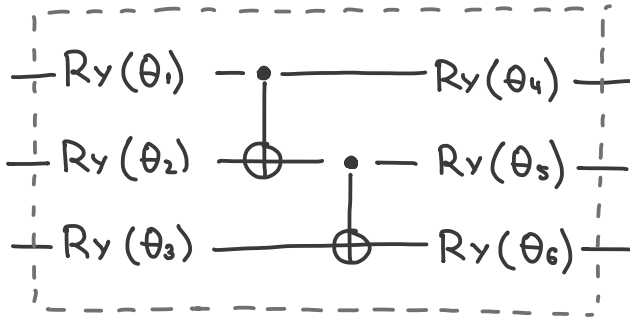
$$|\psi\rangle_\Phi := \frac{1}{N} \sum_{q=1}^K \left(\frac{1}{N} \sum_{p=1}^N \psi_{pq}(x_p) \right) |q\rangle$$

normalization $\rightarrow N^2 := \sum_{q=1}^K \left(\frac{1}{N} \sum_{p=1}^N \psi_{pq}(x_p) \right)^2$

$$\begin{aligned} U_\Phi |0\rangle |+\rangle &= \frac{1}{\sqrt{K}} \sum_{q=1}^K \left(\frac{1}{N} \sum_{p=1}^N \psi_{pq}(x_p) \right) \underbrace{|0\rangle |q\rangle}_{\text{Good}} + \underbrace{\dots}_{\text{Bad}} \\ &= \frac{N}{\sqrt{K}} |0\rangle |\psi\rangle_\Phi + \dots \quad \text{QAA} \quad O\left(\frac{\sqrt{K}}{N}\right) \end{aligned}$$

Training

Parametrized circuit



Observable

$$f_x(\vec{\theta}) = \langle \psi_{\theta} | M | \psi_{\theta} \rangle$$

Cost-function

$$C(\vec{\theta}) \stackrel{\text{e.g.}}{=} \| f_x(\vec{\theta}) - y \|^2$$

Barren plateaus

$$\text{Var}_{\theta} [C(\vec{\theta})] \in \mathcal{O}(b^{-n})$$



Training

- Diagonal block-encoding of amplitudes

$$|0\rangle \xrightarrow{U_w} |w\rangle \quad \text{s.t.} \quad |w\rangle = \sum_{i=1}^{NK} w_i |i\rangle \rightarrow \begin{bmatrix} w_1 \\ \vdots \\ w_{NK} \end{bmatrix}_{BE}$$

\nwarrow parametrize $U_w(\theta)$ s.t. $w_i = w_i(\theta)$

- Hadamard product $U_w \circ U_T =$

$$= \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix} \circ \begin{pmatrix} T & & \\ & T & \\ & & T \end{pmatrix}$$

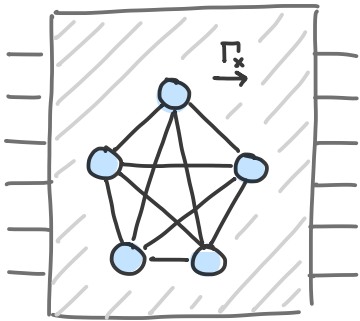
arxiv: 2309.09839

(Rathew & Rebentrost)

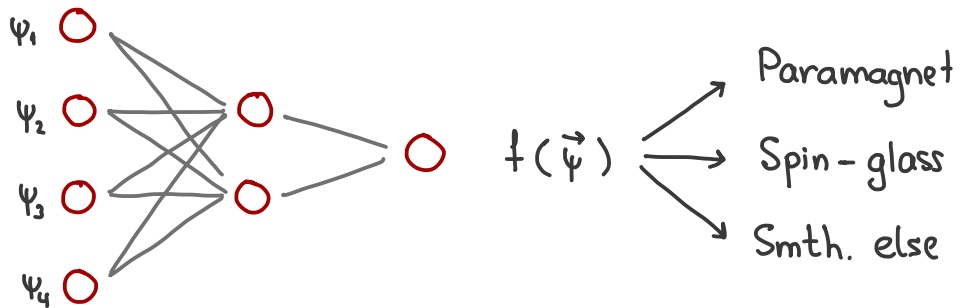
Potential application for QKAN

- o Where can QKAN be useful ?
 - Depends on availability of efficient block-encodings
 - Depends on trainability

Proposal: Quantum phase classification

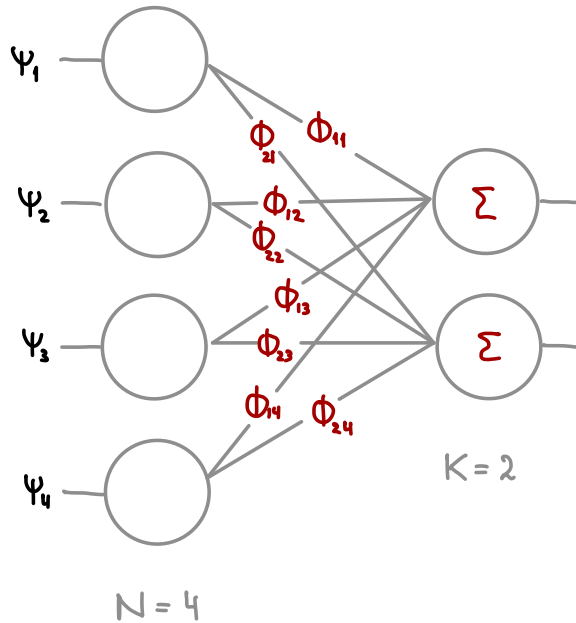


$$U_\psi |0\rangle = |\psi\rangle$$



Thank you !

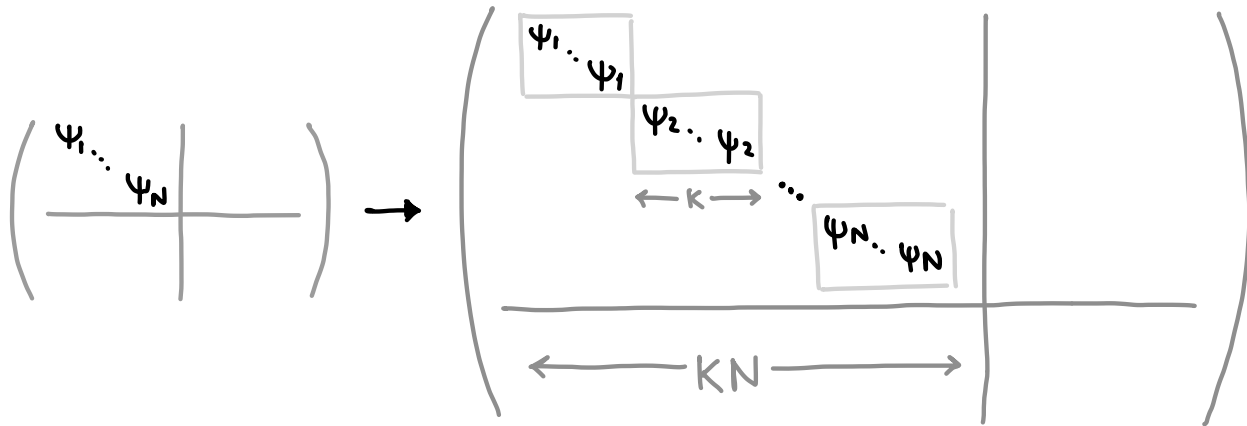
Backup slides



$$\frac{1}{4} \left(\phi_{11}(\psi_1) + \phi_{12}(\psi_2) + \phi_{13}(\psi_3) + \phi_{14}(\psi_4) \right) =: \psi_1^{(1)}$$

$$\frac{1}{4} \left(\phi_{21}(\psi_1) + \phi_{22}(\psi_2) + \phi_{23}(\psi_3) + \phi_{24}(\psi_4) \right) =: \psi_2^{(1)}$$

Backup slides



Backup slides

The diagram illustrates an equality between a unitary operation and a matrix element of a sum of unitaries.

On the left, a vertical rectangle represents a unitary operation U_ϕ . It has two input lines on the left and two output lines on the right. The bottom-left input line is labeled with the index k .

In the center is an equals sign $=$.

On the right, a large pair of parentheses (\quad) encloses a matrix structure. A horizontal line divides the interior into two parts. Above the line, there are two boxed expressions, one above the other, separated by an ellipsis \dots . The top box contains the expression $\frac{1}{N} \sum_{j=1}^N \varphi_{j1}(\psi_j)$. The bottom box contains the expression $\frac{1}{N} \sum_{j=1}^N \varphi_{jk}(\psi_j)$. Below the horizontal line, a double-headed arrow spans the width of the boxes and is labeled with the index k . A vertical line is positioned to the right of the boxed expressions, and the entire structure is enclosed in large parentheses.