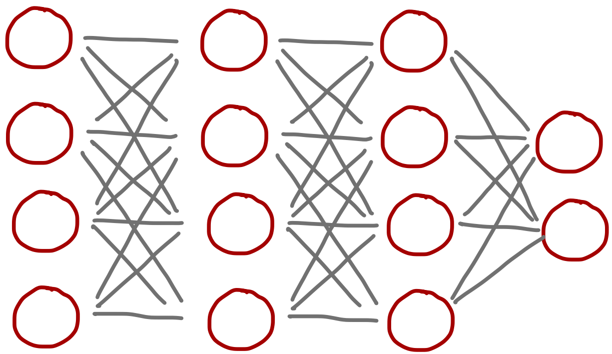
Quantum Kolmogorov-Arnold Networks

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



arxiv.org/abs/2410.04435

Talk outline

5 min
- Kolmorogov - Arnold Theorem
- Kolmorogov - Arnold Networks

10 min
- Interlude : Quantum linear algebra

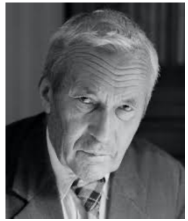
25 min
- QKAN
  - Definition & constuction
  - Solution models
  - Parametrization
  - Possible application

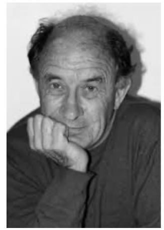
A brief history of Kolmogorov - Arnold

1900 — 1956 — t

$13^{\underline{th}}$ Problem

Kolmogorov - Arnold Theorem





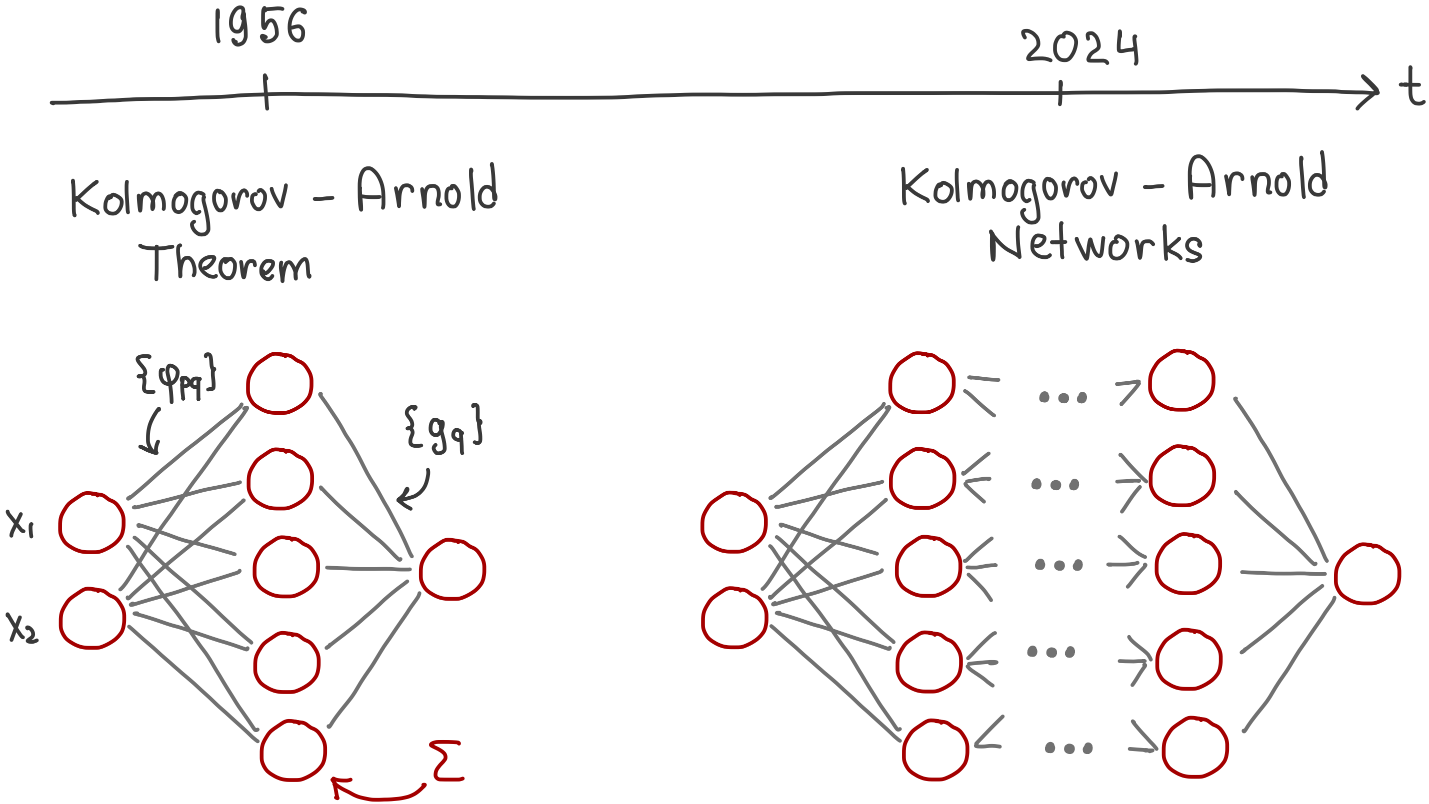
$$f(a,b,c) \overset{?}{=} g_1 \circ g_2(a,b) + g_3 \circ g_4(b,c) + \dots$$

$\boxed{\text{Theorem:}}$ f continuous $\Rightarrow$

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

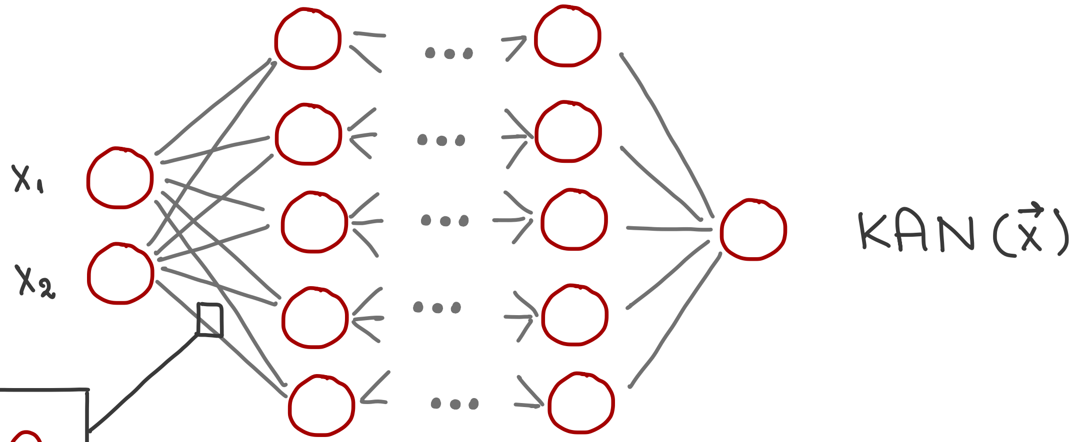
$$f(\vec{x}) = \sum_{q=1}^{2N+1} g_q\left(\sum_{p=1}^{N} \varphi_{pq}(x_p)\right)$$

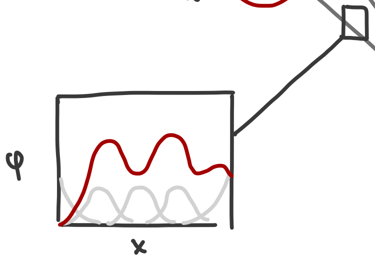
A brief history of Kolmogorov - Arnold



arxiv : 2404.19756 (Liu et al.)

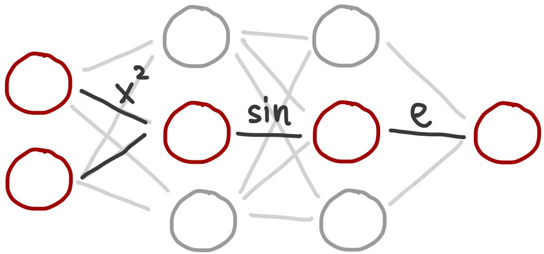
Kolmogorov - Arnold Networks





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

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



Interlude: Quantum linear algebra

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

$$
\underset{\text{nonunitary}}{A} \longrightarrow
\begin{array}{cc}
 & |0\rangle \\
\langle 0| & \begin{pmatrix} A & * \\ * & * \end{pmatrix}
\end{array}
= \underset{\text{unitary}}{U_A} = |0\rangle\langle 0| \otimes A + \dots
$$

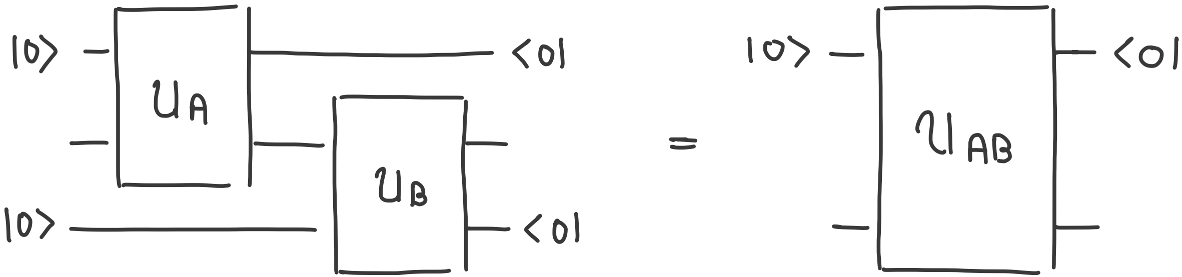
$\text{shape}(A) = N \times N$

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

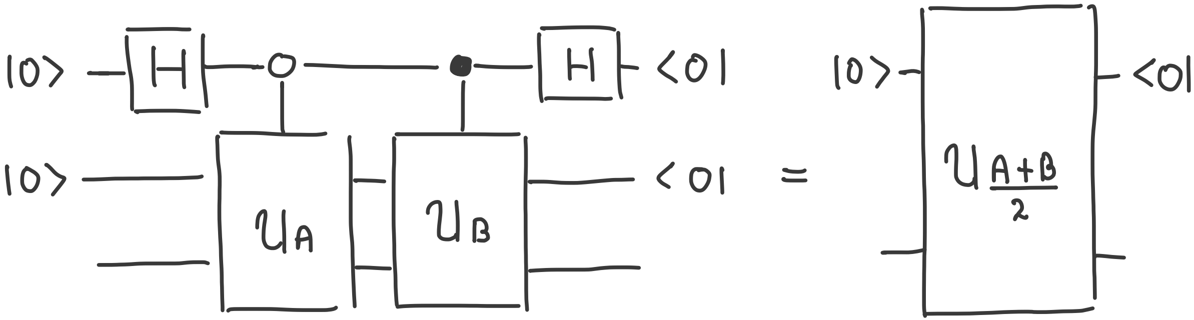
\# qubits to represent U_A : $n + a$

Multiplication & addition

$A \cdot B$



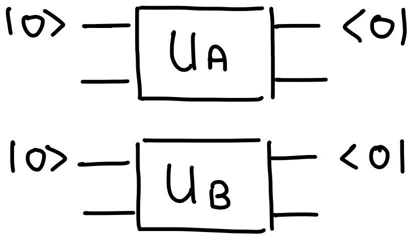
$A + B$
(LCU)

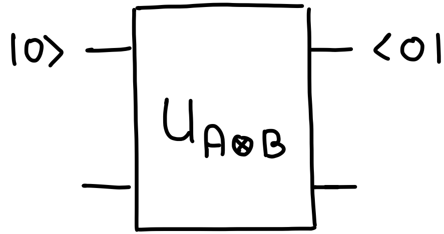


arxiv: 1806.01838
(Gilyén et al.)

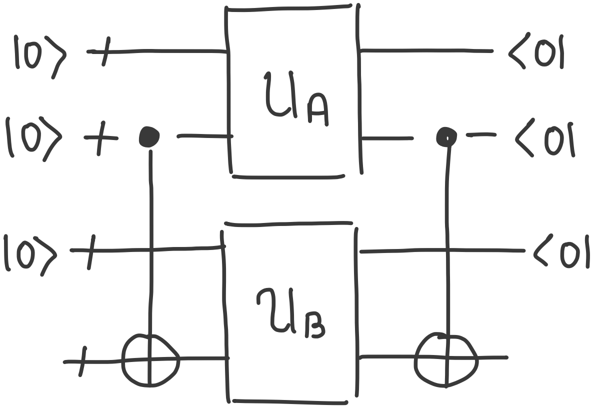
Tensor & Hadamard product

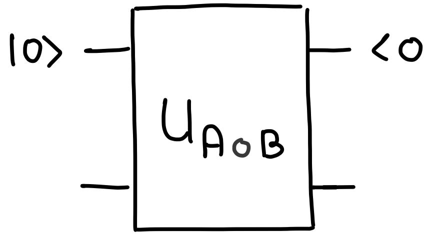
$A \otimes B$





$A \circ B$





arxiv: 2402.16714
(Guo et al.)

General matrix functions $f(A)$

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

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

$$\boxed{\begin{aligned} x &:= \cos\theta \\ \theta &= \arccos x \end{aligned}}$$

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

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

a few Chebyshev polynomials

arxiv: 2201.08309
(Lin Lin)

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

and

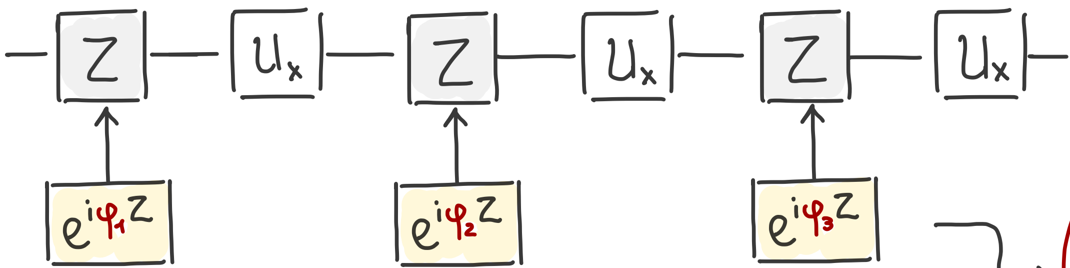
$$\begin{aligned} U_x^\dagger U_x|0\rangle &= x\,U_x^\dagger|0\rangle + \sqrt{1-x^2}\,U_x^\dagger|1\rangle \\ &= x^2|0\rangle + x\sqrt{1-x^2}\,|1\rangle \\ &\quad + \sqrt{1-x^2}\,U_x|1\rangle = |0\rangle \end{aligned}$$

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

General matrix functions $f(A)$

$$\mathcal{U}_x |0\rangle = x|0\rangle + \sqrt{1-x^2}\,|1\rangle$$
$$\mathcal{U}_x |1\rangle = \sqrt{1-x^2}\,|0\rangle - x|1\rangle$$

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



$$\begin{pmatrix} \mathrm{Tr}(X) & * \\ * & * \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} P(x) & * \\ * & * \end{pmatrix}$$

Quantum signal processing

General matrix functions $f(A)$

Idea: Replace x by A and

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

$\boxed{\text{Theorem:}}$ 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

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

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

General matrix functions $f(A)$

$\boxed{\text{Examples:}}$

- $f(H) = e^{-iHt}$: Hamiltonian simulation
- $f(A) = A^{-1}$: Linear system solver

  $\left(A|x\rangle = |b\rangle \Rightarrow |x\rangle = A^{-1}|b\rangle \right)$
- $f(H) = e^{-\beta H}$: Gibbs state preparation

Toolbox

$A \cdot B$

$A \circ B$

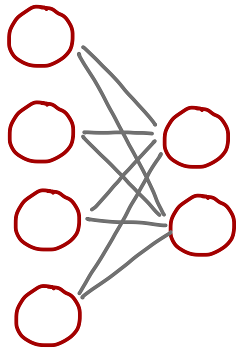
$f(A)$

$A \otimes B$

$A + B$

Goal: Quantum KAN

$$\left[\begin{matrix} x_1 \\ \vdots \\ x_N \end{matrix}\right]_{BE} \updownarrow N$$



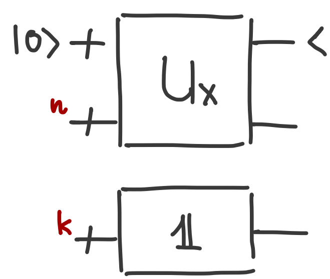
$$\left[\begin{matrix} \frac{1}{N}\sum_{p=1}^{N} \varphi_{p1}(x_p) \\ \vdots \\ \frac{1}{N}\sum_{p=1}^{N} \varphi_{pK}(x_p) \end{matrix}\right]_{BE} \updownarrow K$$

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

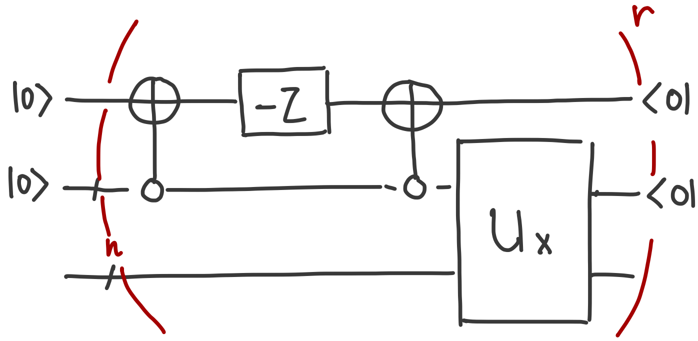
Chebyshev polynomials: $T_r(x) := \cos(r \arccos(x))$

QKAN construction

① $$\begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix}_{BE} \otimes \mathbb{1}_{\log K} = \begin{bmatrix} \left.\begin{matrix} X_1 \\ \vdots \\ X_1 \end{matrix}\right\} K \\ \vdots \\ X_N \\ \vdots \\ X_N \end{bmatrix}_{BE}$$

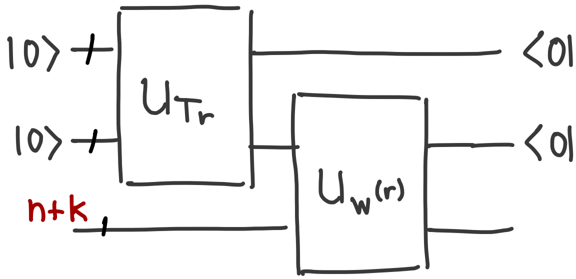


② $$\begin{bmatrix} X_1 \\ \vdots \\ X_1 \\ \vdots \\ X_N \\ \vdots \\ X_N \end{bmatrix}_{BE} \rightarrow \begin{bmatrix} T_r(X_1) \\ \vdots \\ T_r(X_1) \\ \vdots \\ T_r(X_N) \\ \vdots \\ T_r(X_N) \end{bmatrix}_{BE}$$

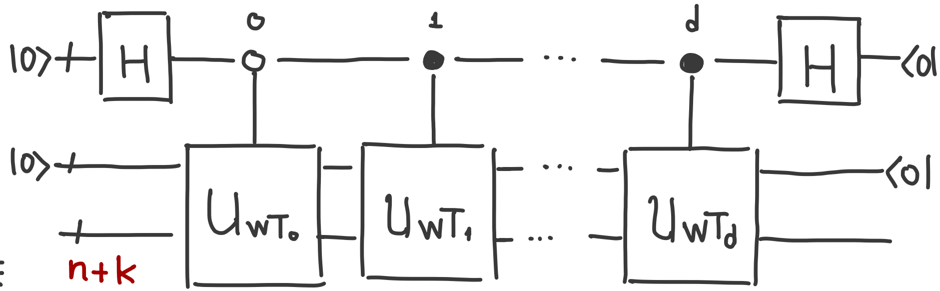


QKAN construction

③ $$\begin{bmatrix} T_r(x_1) \\ \vdots \\ T_r(x_1) \\ \vdots \\ T_r(x_N) \\ \vdots \\ T_r(x_N) \end{bmatrix}_{BE} \cdot \begin{bmatrix} W_{11}^{(r)} \\ \vdots \\ W_{1K}^{(r)} \\ \vdots \\ W_{N1}^{(r)} \\ \vdots \\ W_{NK}^{(r)} \end{bmatrix}_{BE}$$



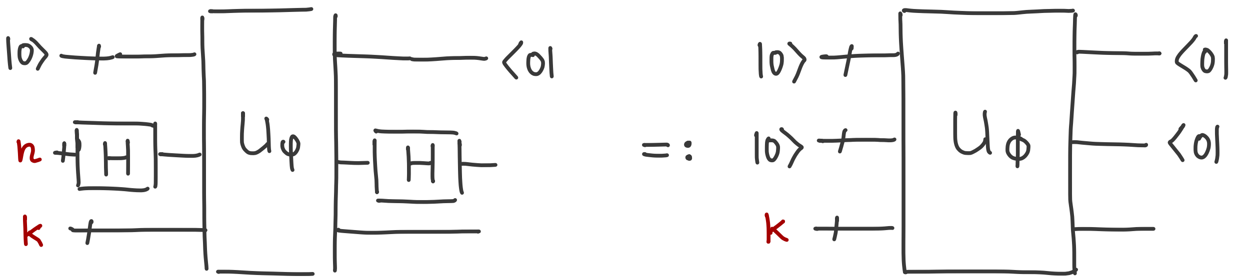
④ $$\sum_{r=0}^{d} \begin{bmatrix} W_{11}^{(r)} T_r(x_1) \\ \vdots \qquad \vdots \\ W_{1K}^{(r)} T_r(x_1) \\ \vdots \qquad \vdots \\ W_{N1}^{(r)} T_r(x_N) \\ \vdots \qquad \vdots \\ W_{NK}^{(r)} T_r(x_N) \end{bmatrix}_{BE}$$



QKAN construction

5

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

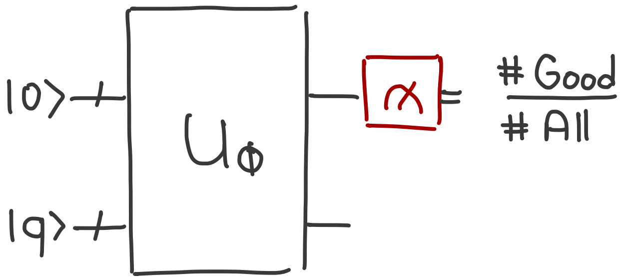


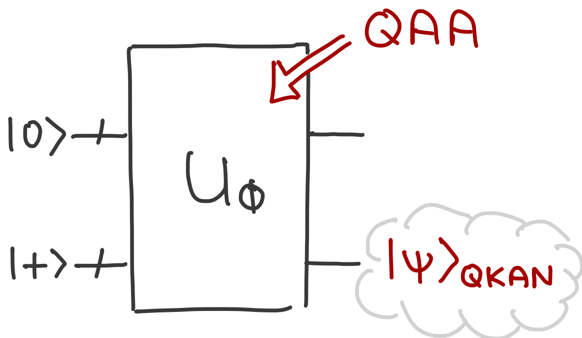
Solution models

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

Estimate $\Phi(\vec{x})_q \pm \varepsilon$

Prepare $|\psi\rangle_\Phi$





Solution models

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

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

- $P[\text{Good}] = |\Phi(\vec{x})_q|^2 \longrightarrow O\left(\frac{1}{\varepsilon^2}\right)$ 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 : $\boxed{\text{Prepare } |\psi\rangle_\phi}$

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

normalization

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

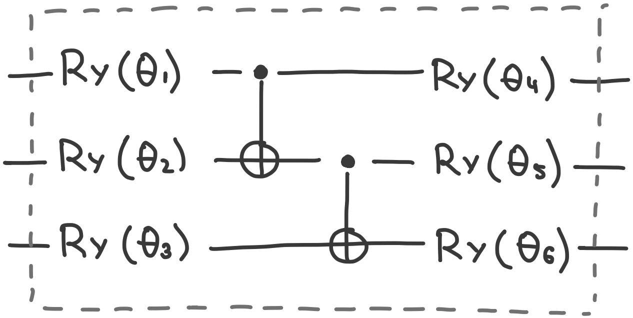
$$U_\phi |0\rangle |+\rangle = \frac{1}{\sqrt{K}} \sum_{q=1}^{K} \left(\frac{1}{N} \sum_{p=1}^{N} \varphi_{pq}(x_p) \right) \underbrace{|0\rangle |q\rangle}_{\text{Good}} + \underbrace{\dots}_{\text{Bad}}$$

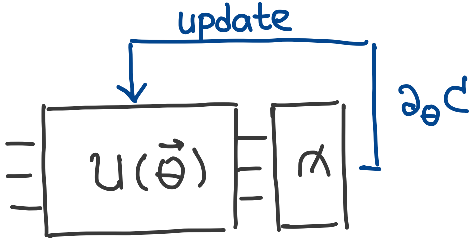
$$= \frac{\mathcal{N}}{\sqrt{K}} |0\rangle |\psi\rangle_\phi + \dots$$

QAA $O\left(\frac{\sqrt{K}}{\mathcal{N}}\right)$

Training

Parametrized circuit



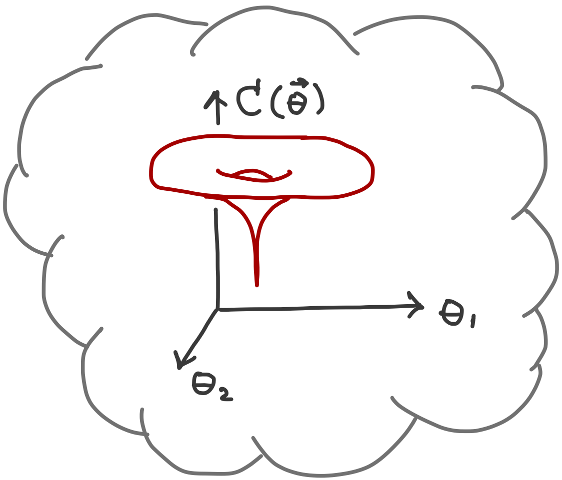


Observable

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

Cost-function

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



Barren plateaus

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

Training

- Diagonal block-encoding of amplitudes

$$|0\rangle \;—\; \boxed{U_w} \;—\; |w\rangle \qquad \text{s.t.} \qquad |w\rangle = \sum_{i=1}^{NK} w_i\,|i\rangle \;\rightarrow\; \begin{bmatrix} w_1 & & \\ & \ddots & \\ & & w_{NK} \end{bmatrix}_{BE}$$

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

- Hadamard product $U_w \circ U_T =$

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

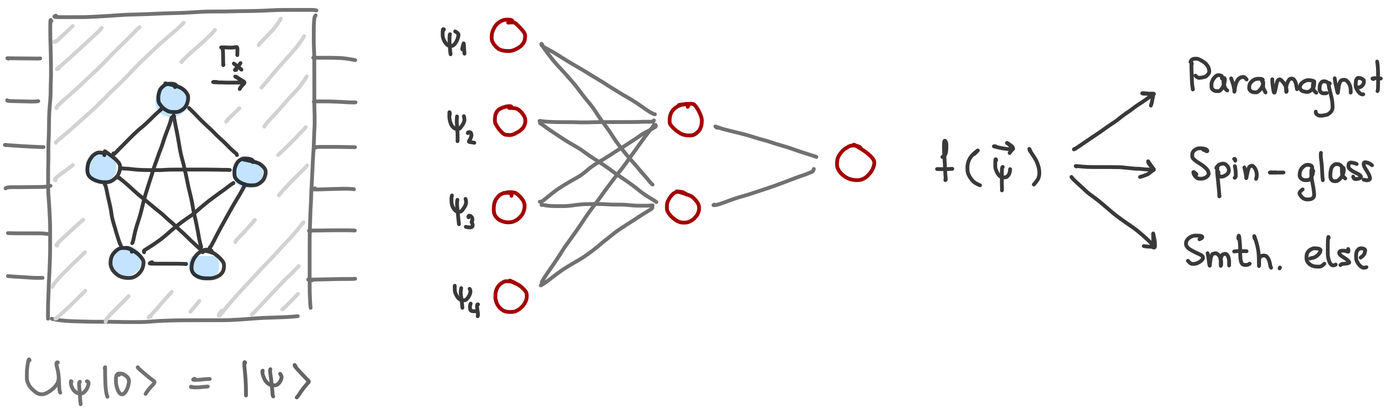
arxiv: 2309.09839

(Rathew & Rebentrost)

Potential application for QKAN

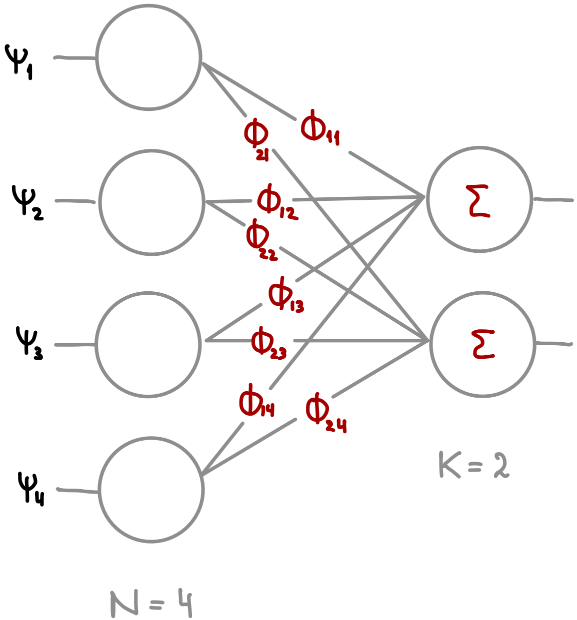
- Where can QKAN be useful ?
  - Depends on availability of efficient block-encodings
  - Depends on trainability

Proposal: Quantum phase classification



Thank you!

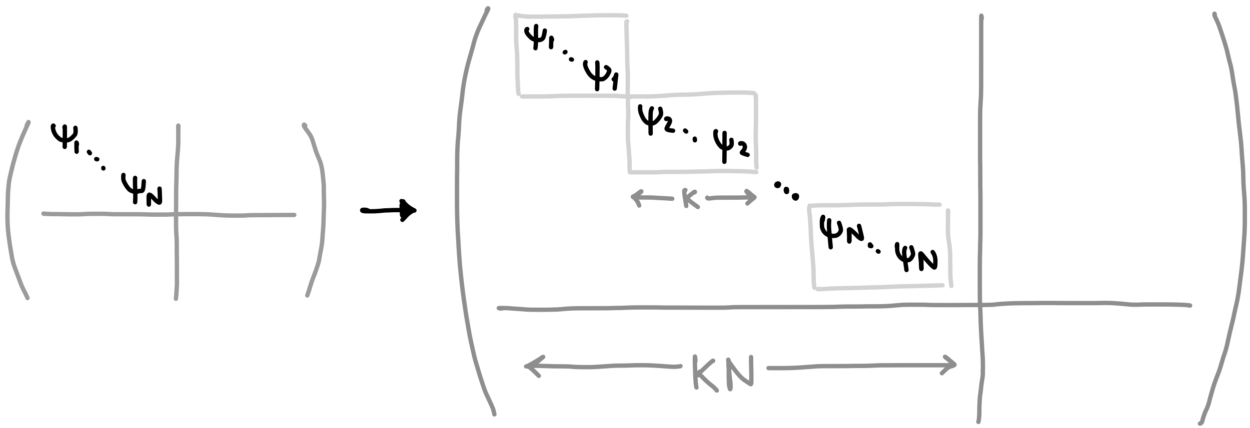
Backup slides



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

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

Backup slides



Backup slides