Problem-informed Graphical Quantum Generative Learning

Bence Bakó, Wigner RCP

arXiv:2405.14072

HUN REN Wigner

ELTE EÖTVÖS LORÁND UNIVERSITY

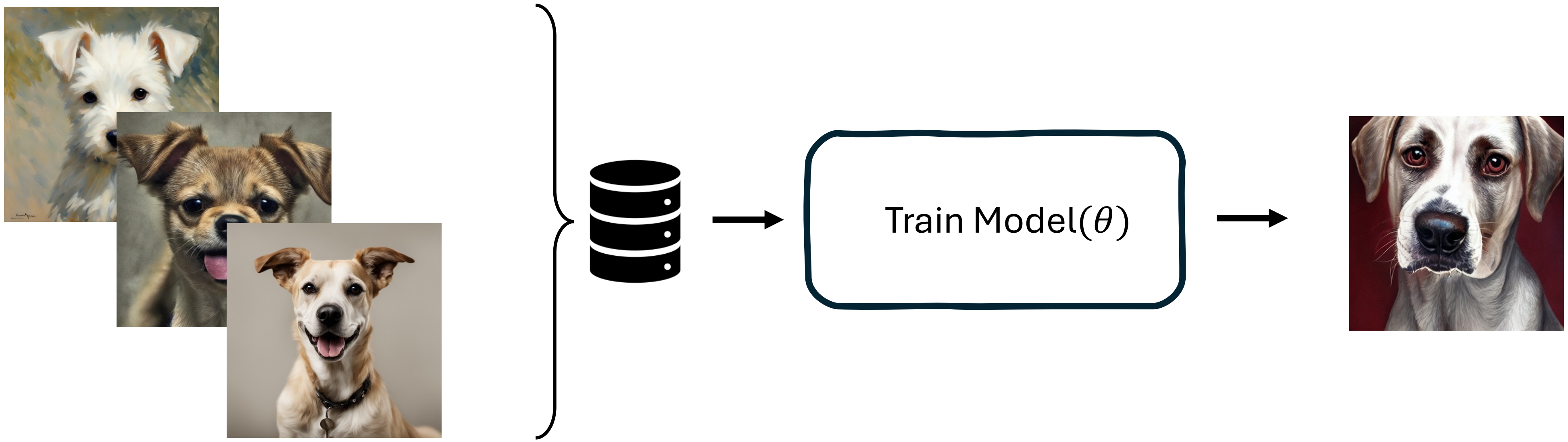
Quantum Information National Laboratory HUNGARY

GPU Day 2024

Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal

Generative modeling

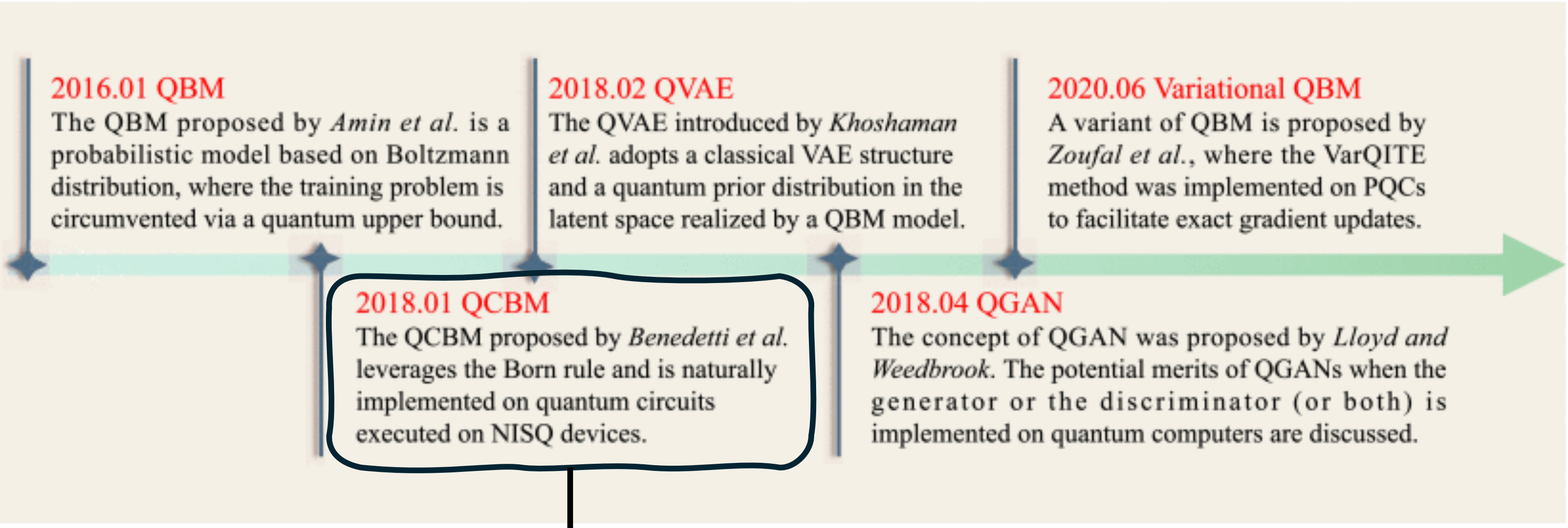
- Learn a representation of some **probability distribution** in order to **generate realistic samples**.



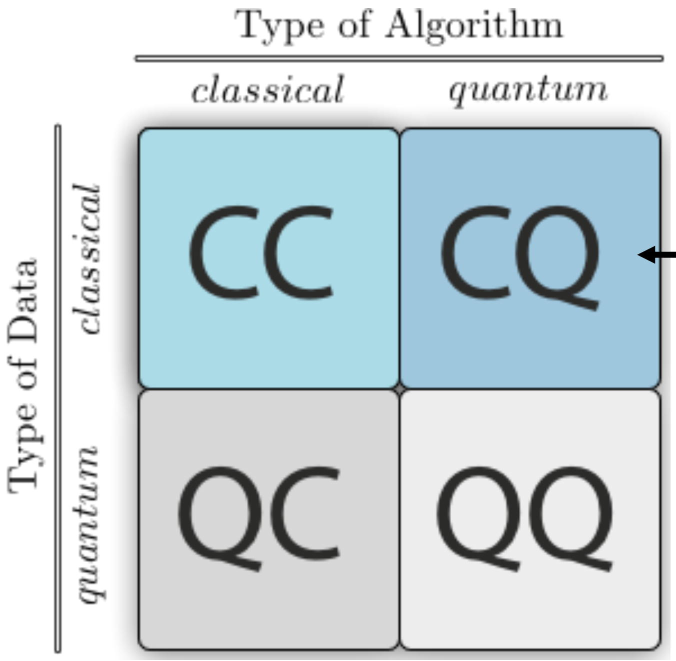
Generated with DeepAI image generator

Generative QML

“Natural” ML application for quantum computers.



Source: J. Tian, et al., IEEE Transactions on Pattern Analysis and Machine Intelligence, Oct. 2023.



Source: Wikipedia

Quantum Circuit Born Machine

- Paradigmatic quantum generative model
- Inherits the Born rule

Data

Target distribution

Model distribution

Quantum Circuit Born Machines

General task:

- Learn a representation of the (target) probability distribution over binary random variables.
- Access to explicit distribution (not realistic) OR a **limited number of samples**.

QCBM task:

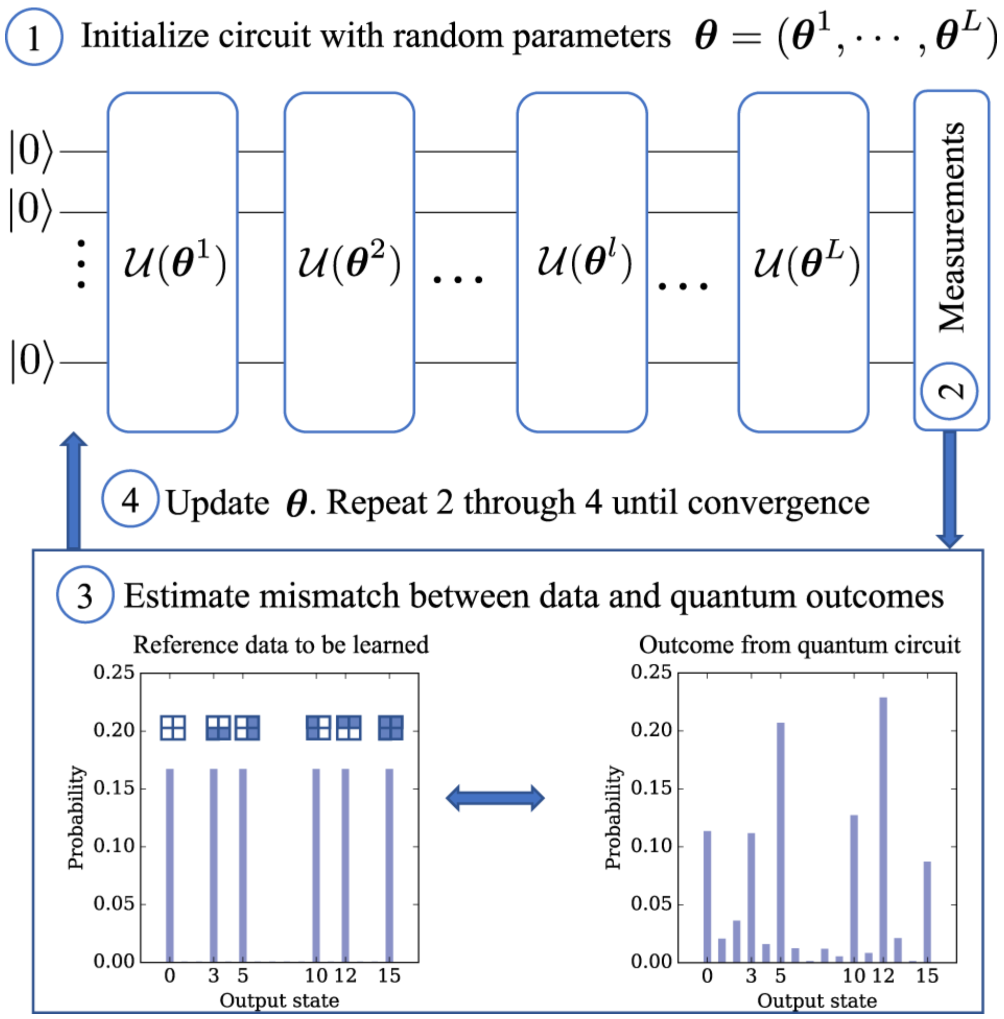
- Learn a quantum state via optimizing the parameters of a variational quantum circuit s. t.

$$|\psi\rangle = U(\boldsymbol{\theta})|\psi_0\rangle$$

$$P_{\boldsymbol{\theta}}(\boldsymbol{x}) = |\langle \boldsymbol{x}|\psi\rangle|^2$$

$$d(P_{\boldsymbol{\theta}}, P^*) \leq \varepsilon$$

where $\varepsilon \in (0, 1)$

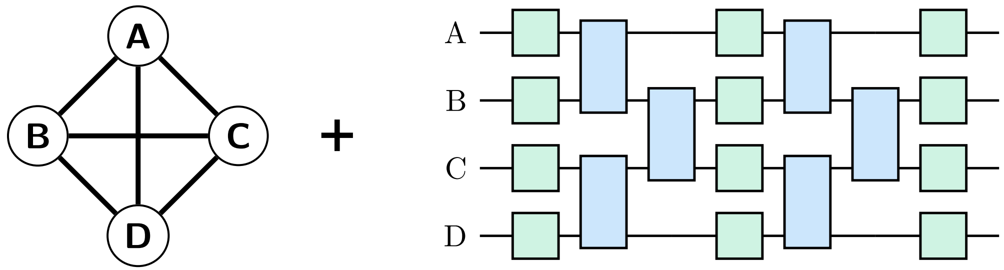


Source: M. Benedetti, et al., npj QI, vol. 5, no. 1, p. 45, 2019.

General-purpose QCBMs

Task: learn P(A,B,C,D) – joint probability distribution of correlated (binary) random variables

No problem-specific knowledge

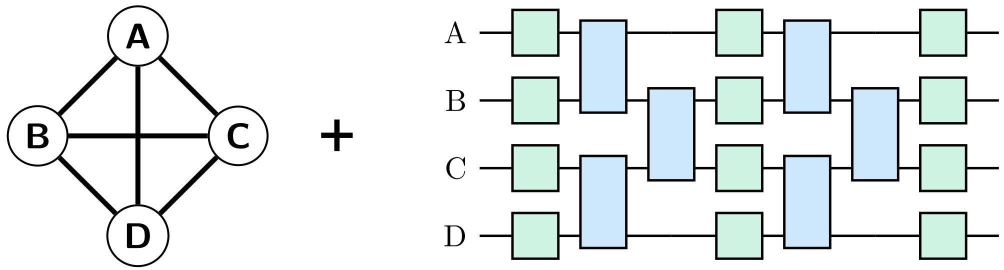


General-purpose Ansatz

General-purpose QCBMs

Task: learn P(A,B,C,D) – joint probability distribution of correlated (binary) random variables

No problem-specific knowledge



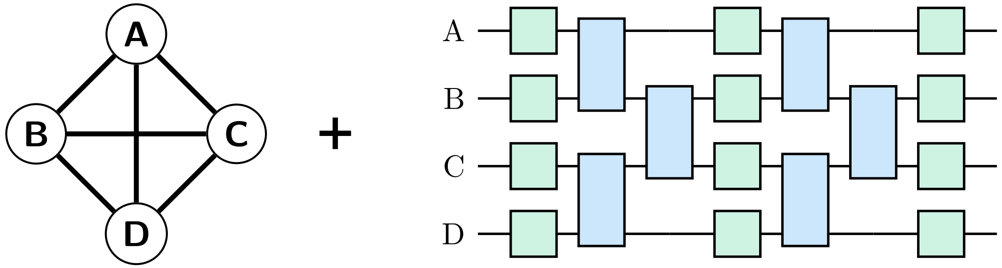
General-purpose Ansatz

- Trainability issues (e.g., barren plateaus)
- Poor average performance (no-free-lunch theorem)

General-purpose QCBMs

Task: learn P(A,B,C,D) – joint probability distribution of correlated (binary) random variables

No problem-specific knowledge



General-purpose Ansatz

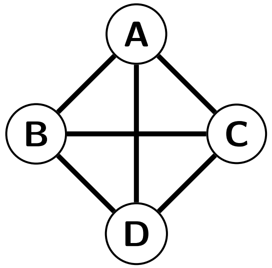
- Trainability issues (e.g., barren plateaus)
- Poor average performance (no-free-lunch theorem)

NO “one model to rule them all!”

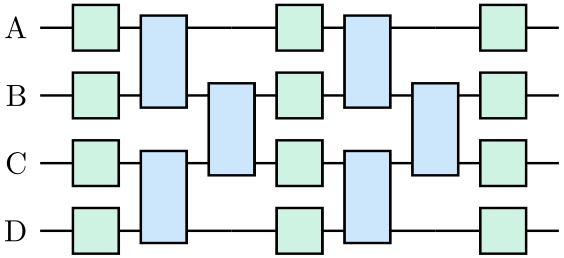
General-purpose QCBMs

Task: learn P(A,B,C,D) – joint probability distribution of correlated (binary) random variables

No problem-specific knowledge



+



General-purpose Ansatz

- Trainability issues (e.g., barren plateaus)
- Poor average performance (no-free-lunch theorem)

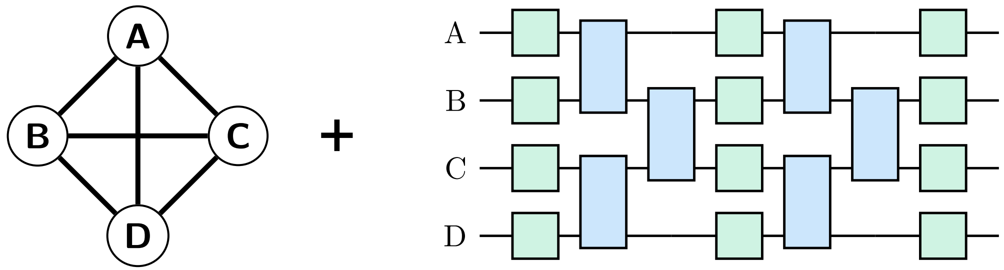
Insufficient inductive bias!

How to incorporate problem-specific knowledge?

General-purpose vs Problem-informed

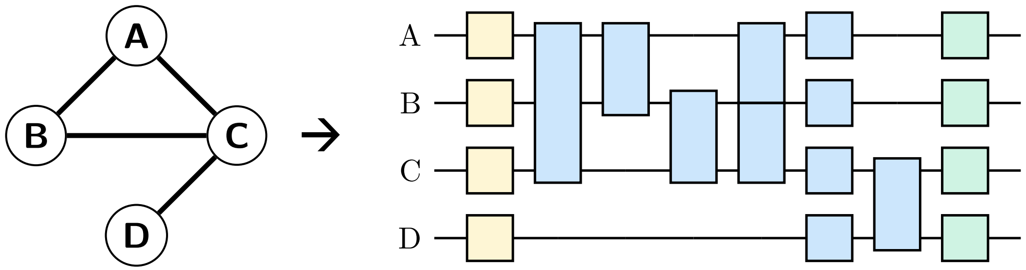
Task: learn P(A,B,C,D) – joint probability distribution of correlated (binary) random variables

No problem-specific knowledge



General-purpose Ansatz

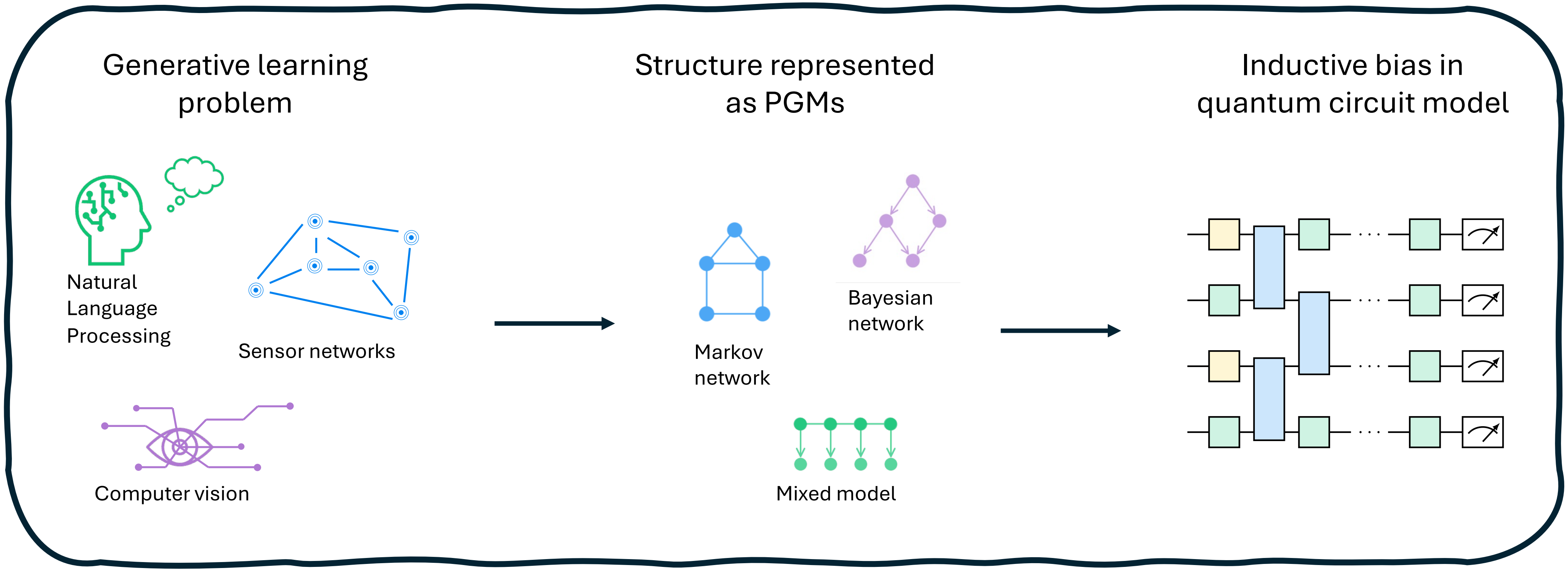
Independencies: $(A \perp D \mid C)$ $(B \perp D \mid C)$



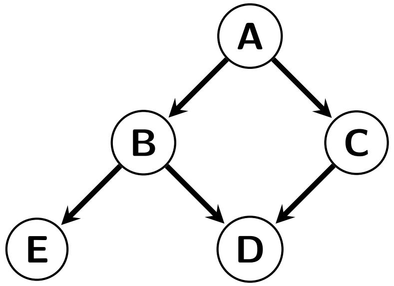
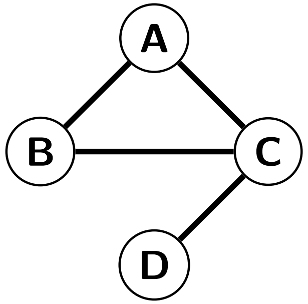
Problem-informed Ansatz

Use **Probabilistic Graphical Models (PGMs)**

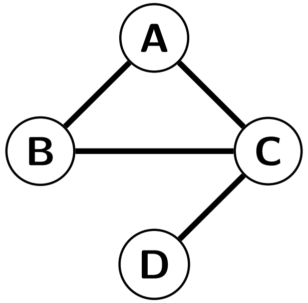
Problem-informed Generative QML Framework



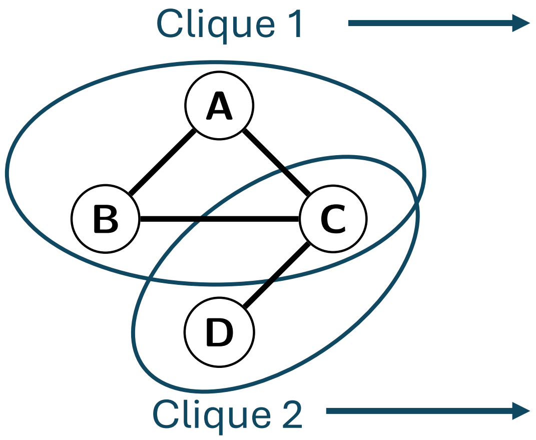
Probabilistic Graphical Models

Graph	Independencies	Factorization
 Bayesian networks	$(B \perp C \mid A)$ $(D \perp A \mid B, C)$ $(E \perp C, D \mid B)$	$P(A,B,C,D,E) = P(A)P(B\|A)P(C\|A)P(D\|B,C)P(E\|B)$
 Markov networks	$(A \perp D \mid C)$ $(B \perp D \mid C)$	$P(A,B,C,D) = \frac{1}{Z}\phi_1(A,B,C)\phi_2(C,D)$

Markov Networks (MN)



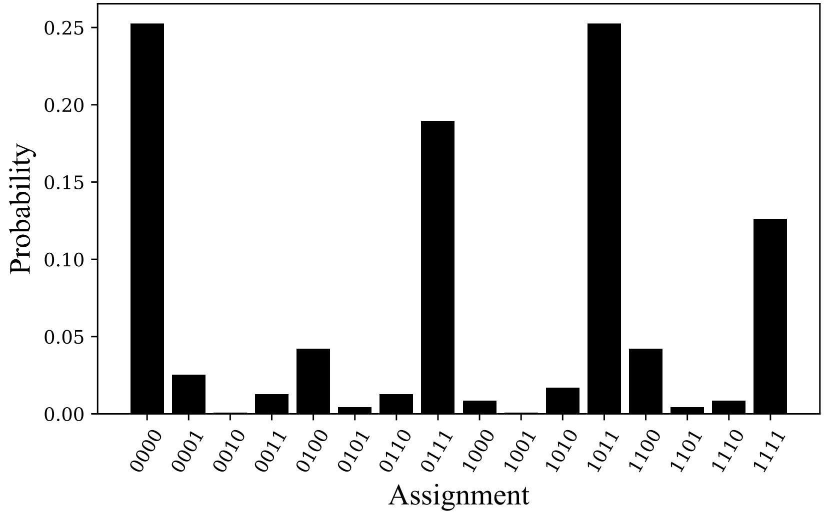
Maximal clique factorization



$\phi_1(A,B,C)$			
a^0	b^0	c^0	30
a^0	b^0	c^1	1
a^0	b^1	c^0	5
a^0	b^1	c^1	15
a^1	b^0	c^0	1
a^1	b^0	c^1	20
a^1	b^1	c^0	5
a^1	b^1	c^1	10

$\times$

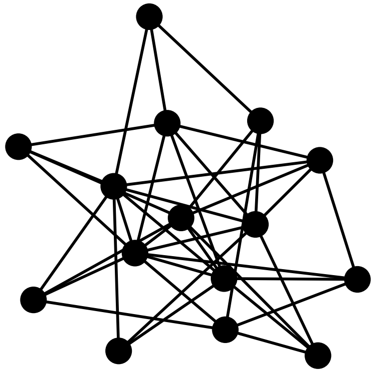
$\phi_2(C,D)$		
c^0	d^0	10
c^0	d^1	1
c^1	d^0	1
c^1	d^1	15



$$P(A,B,C,D) = \frac{1}{Z}\phi_1(A,B,C)\phi_2(C,D)$$

We use this framework for benchmark construction.

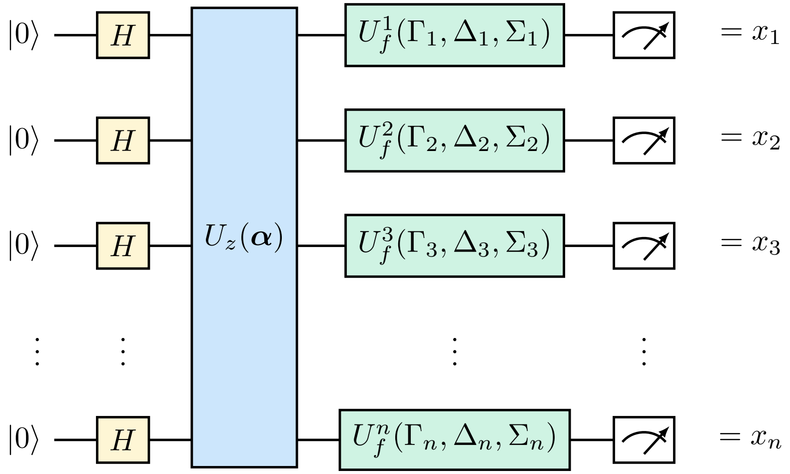
Quantum Circuit Markov Random Field (QCMRF)



Higher-order Ising Hamiltonian
Duplicate terms and identities discarded

$$\rightarrow \quad H'(\boldsymbol{\beta}) = \sum_{C\in\mathcal{C}} \bigotimes_{v\in C} \beta_{C,v}(I+Z_v) \quad \rightarrow \quad H(\boldsymbol{\alpha}) \quad \rightarrow \quad U_Z(\boldsymbol{\alpha}) = e^{-iH(\boldsymbol{\alpha})}$$

$\downarrow$

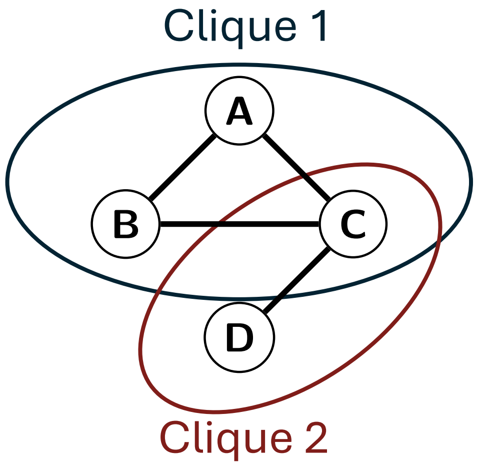


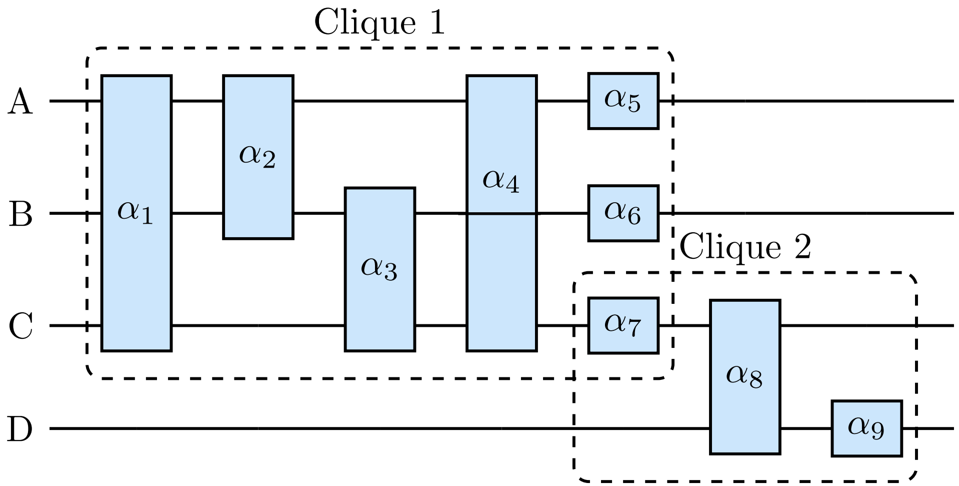
Quantum Circuit Ising Born Machine (QCIBM)

- Similar, but problem-agnostic Ansatz
- Only 2-local interactions
- All-to-all connectivity

B. Coyle, et al., npj QI, vol. 6, no. 1, p. 60, 2020.

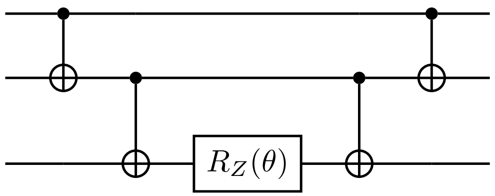
QCMRF example





$$H(\boldsymbol{\alpha}) = \alpha_1 Z_A Z_B Z_C + \alpha_2 Z_A Z_B + \alpha_3 Z_B Z_C + \alpha_4 Z_A Z_C + \alpha_5 Z_C Z_D + \alpha_6 Z_A + \alpha_7 Z_B + \alpha_8 Z_C + \alpha_9 Z_D$$

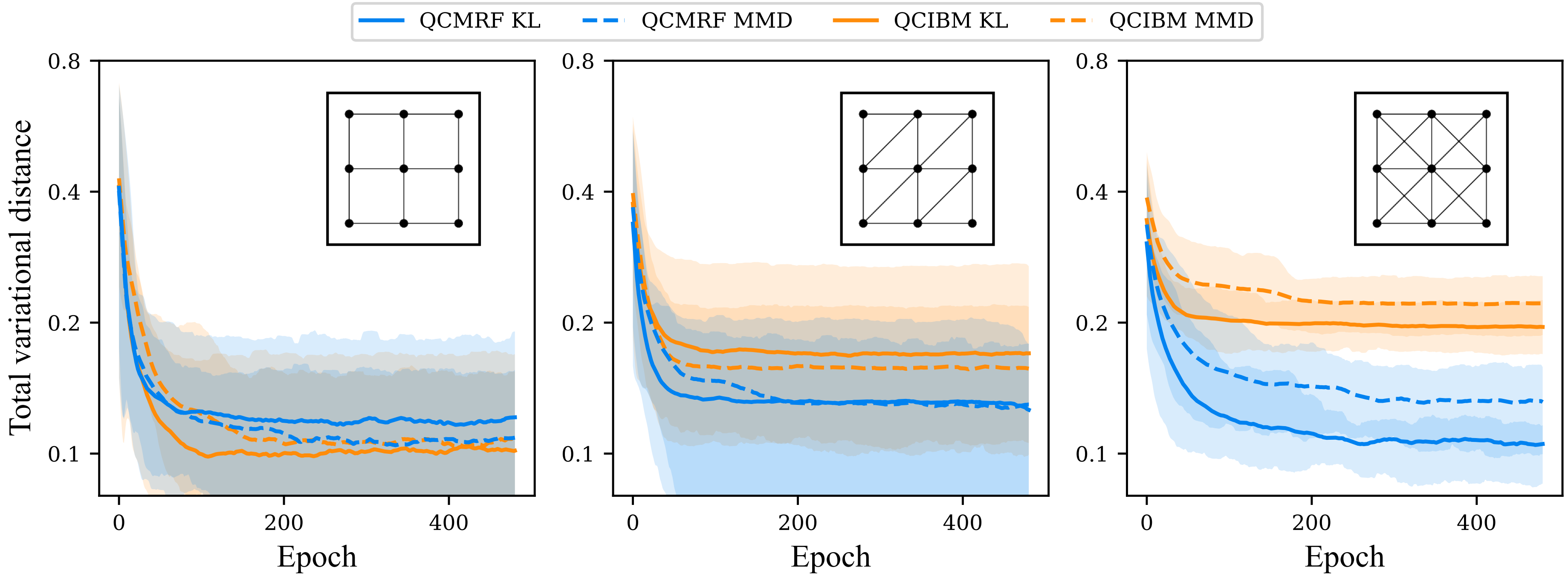
$$\exp\left(-i\theta Z_0 Z_1 Z_2 / 2\right) =$$



- MNs can represent **ANY probability distribution**!

- When is this **representation useful** (for our model)?

1. When does it outperform problem-agnostic?

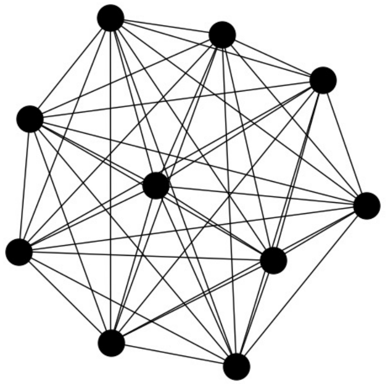


Performance decreases!

Can capture higher-order correlations!

Problem complexity

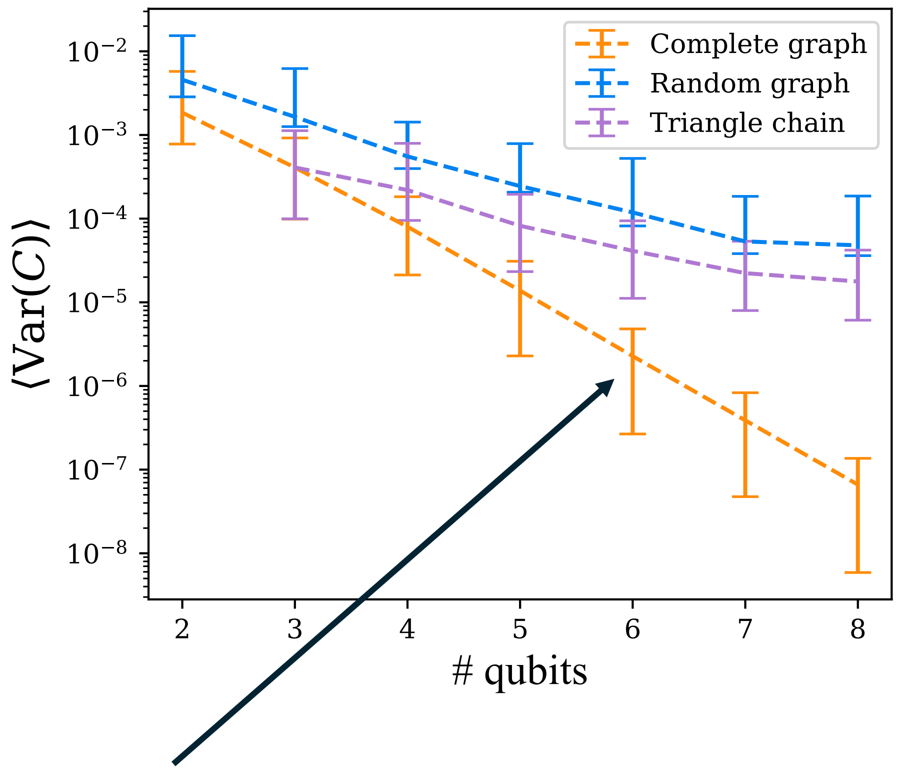
2. What problems should we consider?



Complete graph with
maximal clique factorization

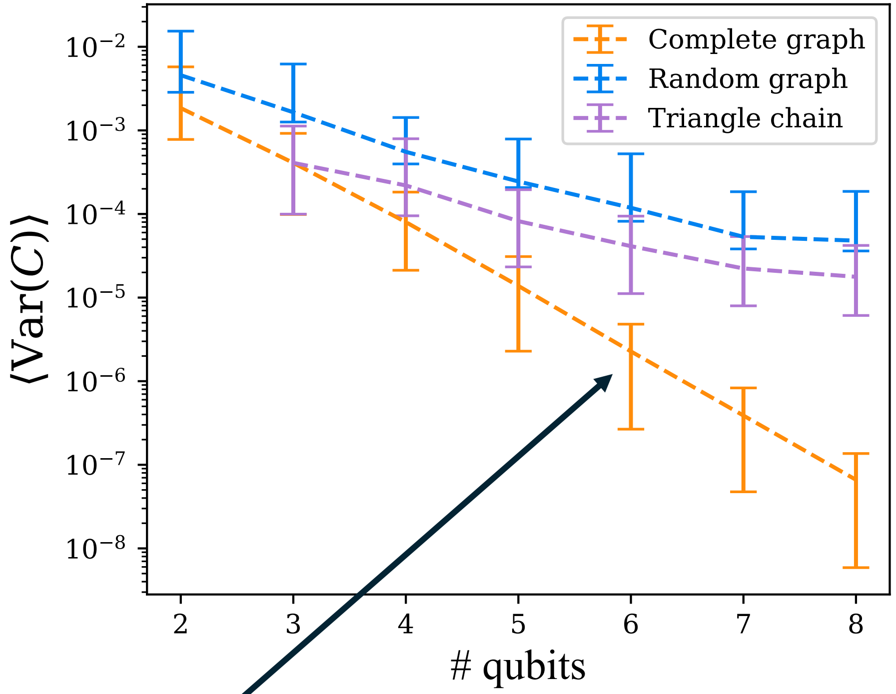
$\mathcal{O}(2^n)$ degrees of freedom

Can this be trainable?



Exponential decay → deterministic barren plateaus

2. Efficient MN representation



$poly(n)$ parameters (still classically hard) → better trainability

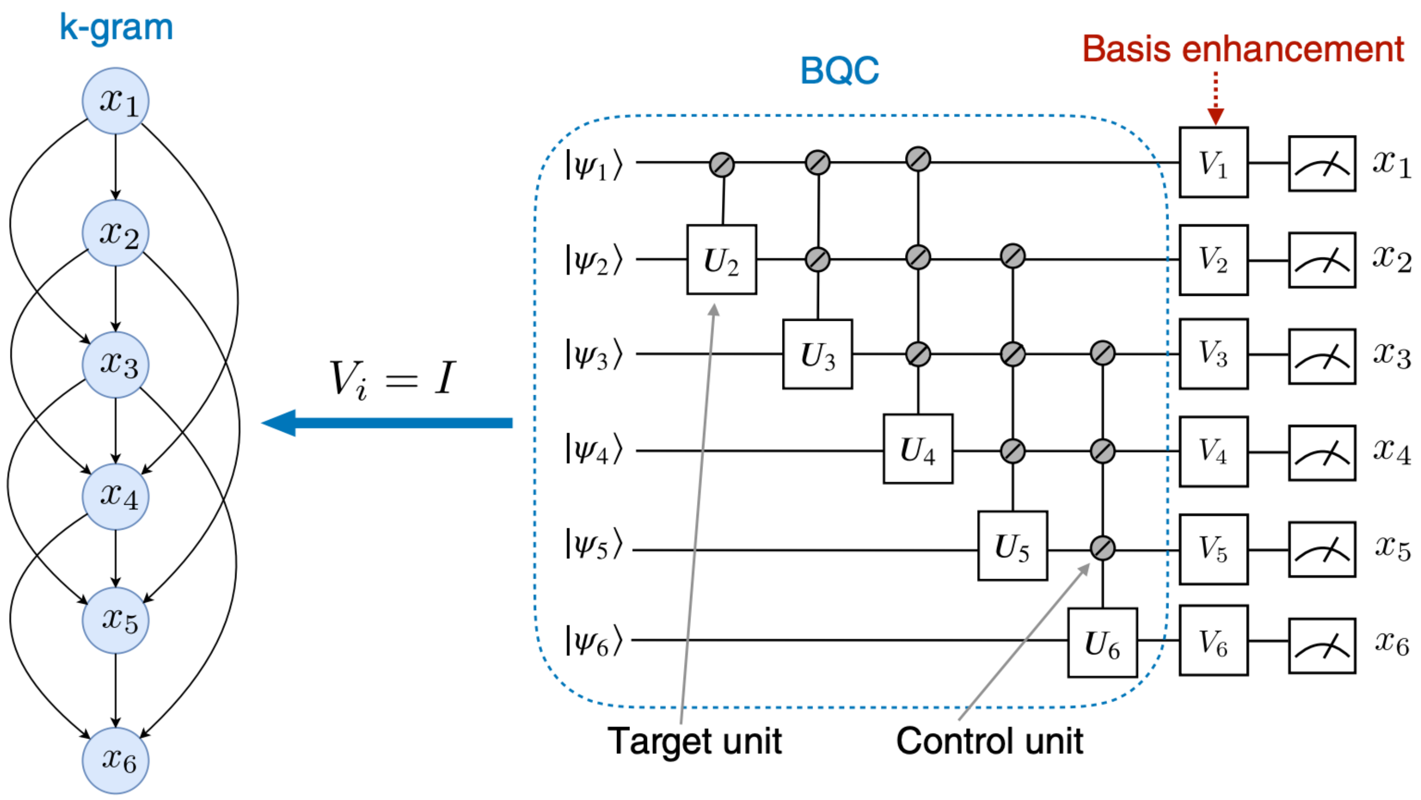
⬇

Efficient MN representation

Exponential decay → deterministic barren plateaus

3. Basis-enhanced Bayesian Quantum Circuit

Bayesian Network

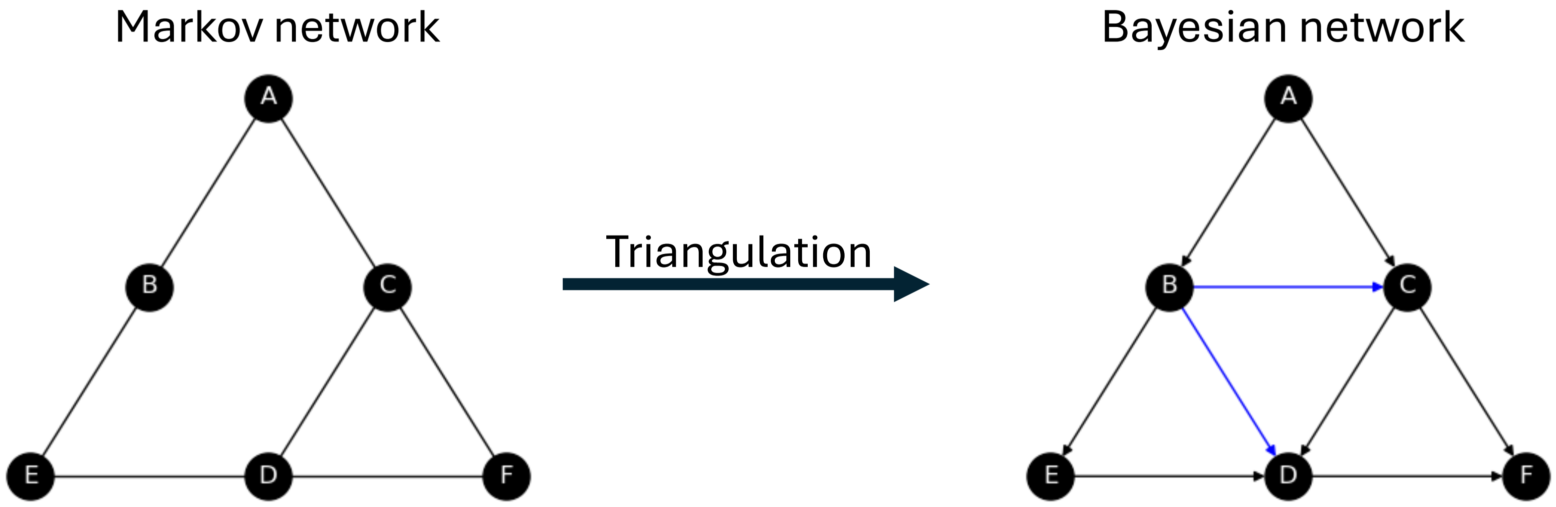


- Can represent exactly the probability distribution induced by the BN!
- Basis-enhancement makes BBQC more expressive than the corresponding BN!

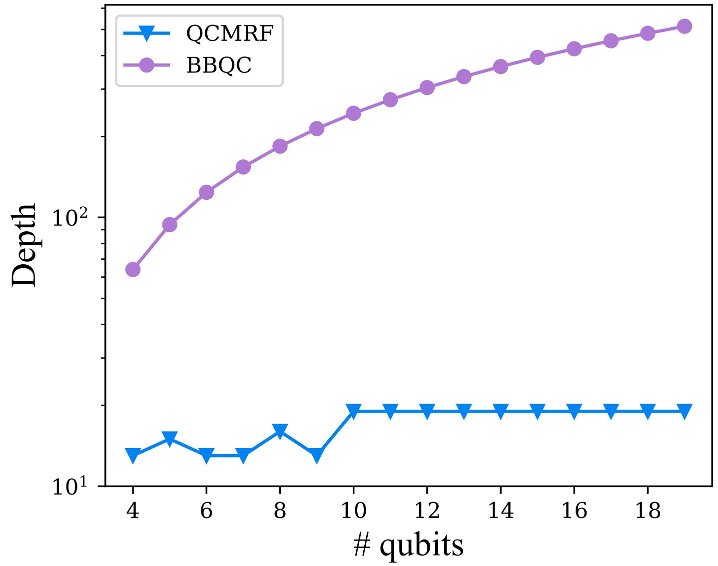
X. Gao, et al. PRX, vol. 12, no. 2, p. 021037, 2022.

3. Learning MNs with BBQCs

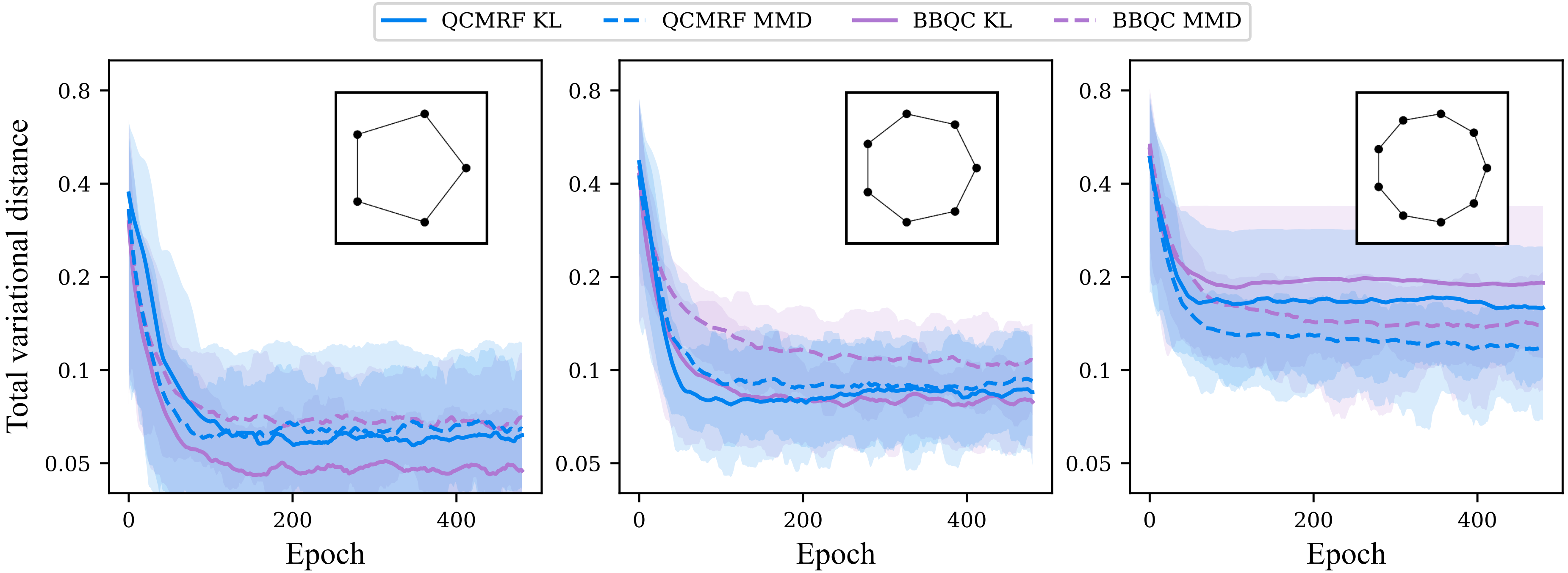
- MN graph has to be triangulated first → computationally intensive task!
- BBQCs also require significantly more quantum resources (than QCMRFs)!



3. QCMRF vs BBQC

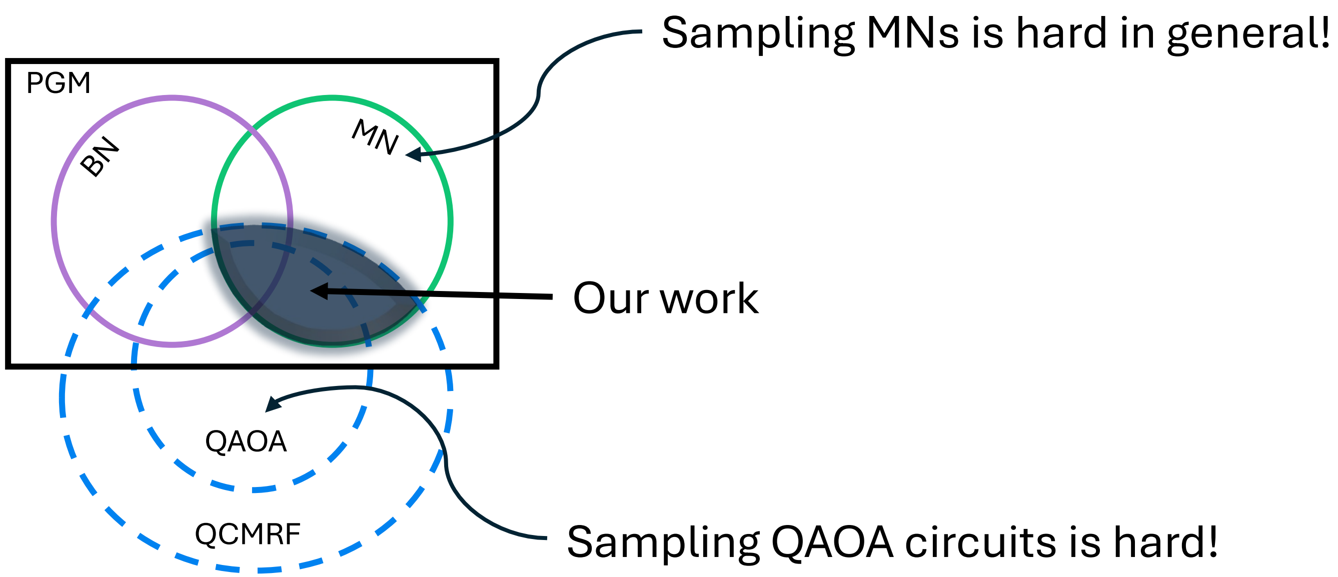


Similar performance at much lower costs!



4. Potential for Quantum Advantage?

1. Quantum learning advantage in
   - Accuracy
   - Learning speed
   - Sample complexity



Farhi, Harrow. arXiv:1602.07674, 2019.
Krovi. arXiv:2206.05642, 2022.

2. Quantum advantage in sampling the unknown target distribution:
- Target distribution is learnable (up to given error) by both a classical and a quantum model
- Sampling the trained quantum circuit is more efficient

Most promising problems

1. Higher-order correlations → **To outperform problem-agnostic**
2. Polynomial number of parameters → **Better trainability**
3. Hard to convert to BNs → **More cost-effective**
4. Hard to sample classically → **Potential for Q advantage**

Problem-informed Graphical Quantum Generative Learning

Bence Bakó,[1,2] Dániel T. R. Nagy,[1,2] Péter Hága,[3] Zsófia Kallus,[3] and Zoltán Zimborás[1,2,4]

[1] *Eötvös Loránd University, Budapest, Hungary*
[2] *HUN-REN Wigner Research Centre for Physics, Budapest, Hungary*
[3] *Ericsson Research, Budapest, Hungary*
[4] *Algorithmiq Ltd, Kanavakatu 3C, Helsinki, 00160, Finland*

arXiv:2405.14072