Title: Classical algorithms for quantum mean values

Speakers: David Gosset

Collection: Symmetry, Phases of Matter, and Resources in Quantum Computing

Date: November 27, 2019 - 9:30 AM

URL: http://pirsa.org/19110130

Abstract: Consider the task of estimating the expectation value of an n-qubit tensor product observable in the output state of a shallow quantum circuit. This task is a cornerstone of variational quantum algorithms for optimization, machine learning, and the simulation of quantum many-body systems. In this talk I will describe three special cases of this problem which are "easy" for classical computers. This is joint work with Sergey Bravyi and Ramis Movassagh.

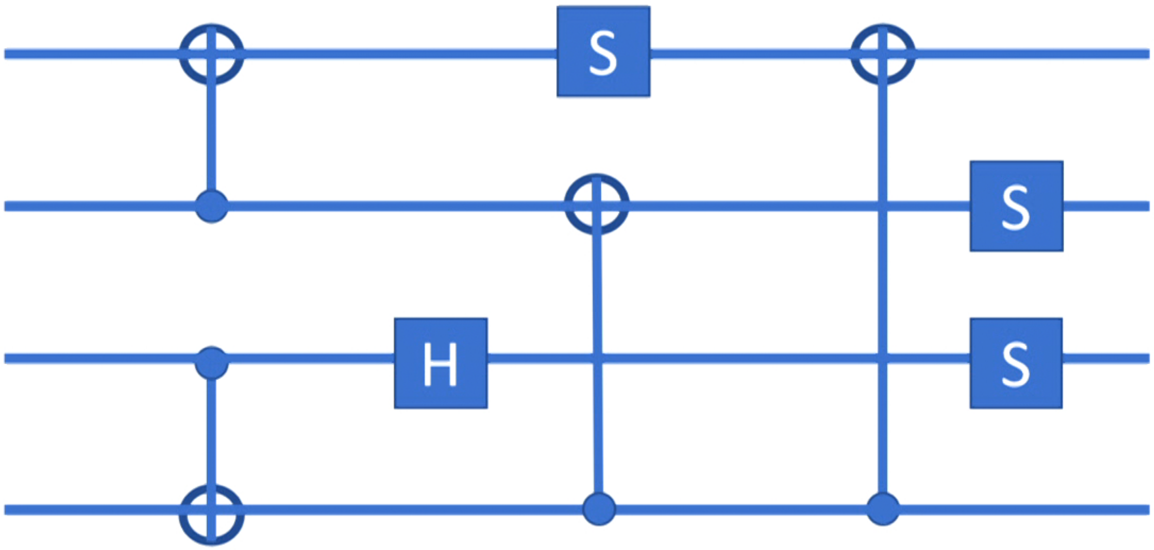
Classical algorithms for quantum mean values

Sergey Bravyi
David Gosset
Ramis Movassagh

arXiv:1909.11485

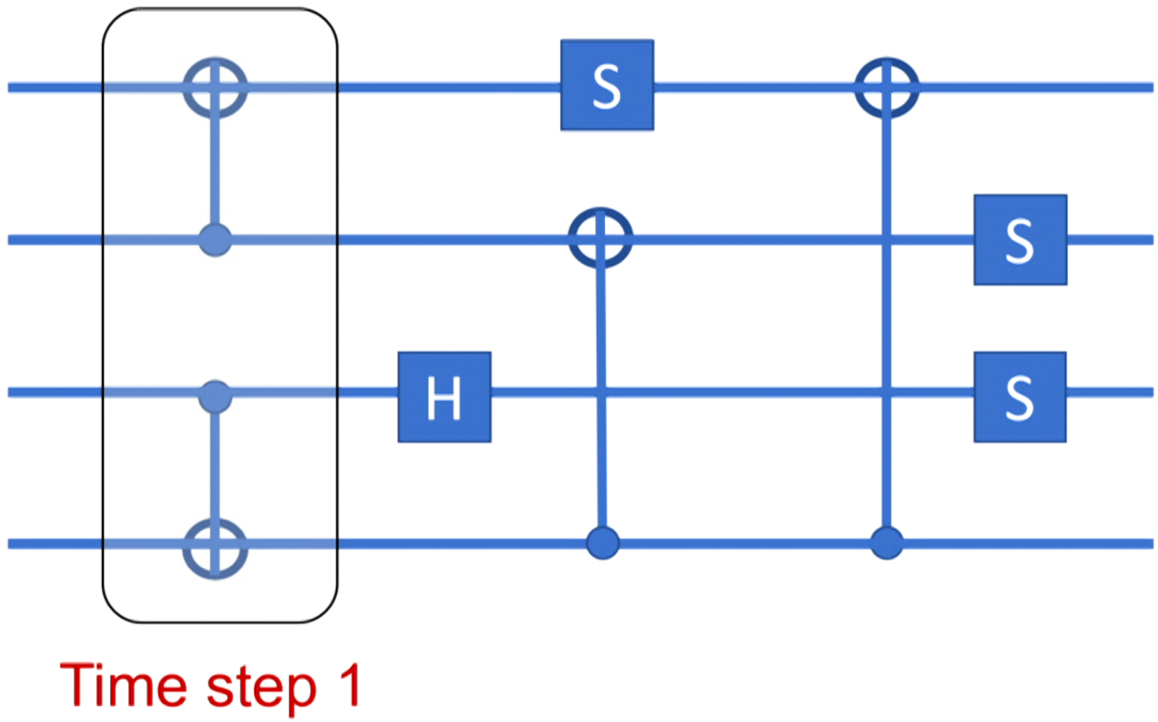
Circuit depth

Circuit depth is the number of time steps allowing for parallel gates.



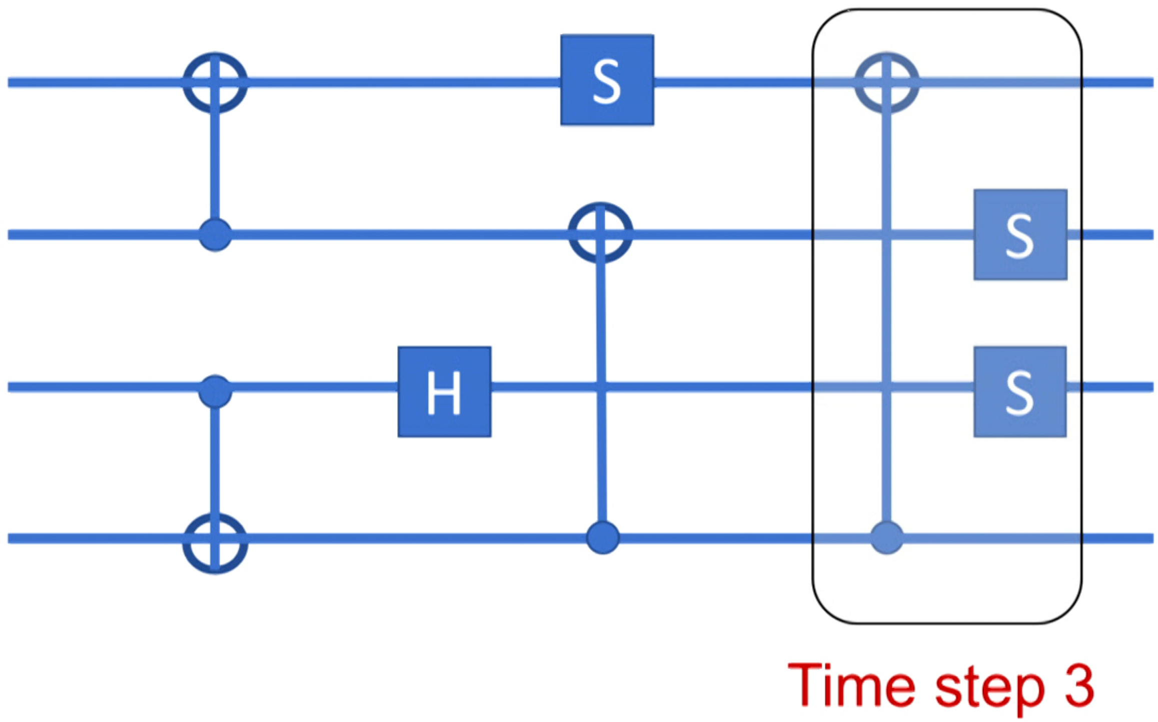
Circuit depth

Circuit depth is the number of time steps allowing for parallel gates.

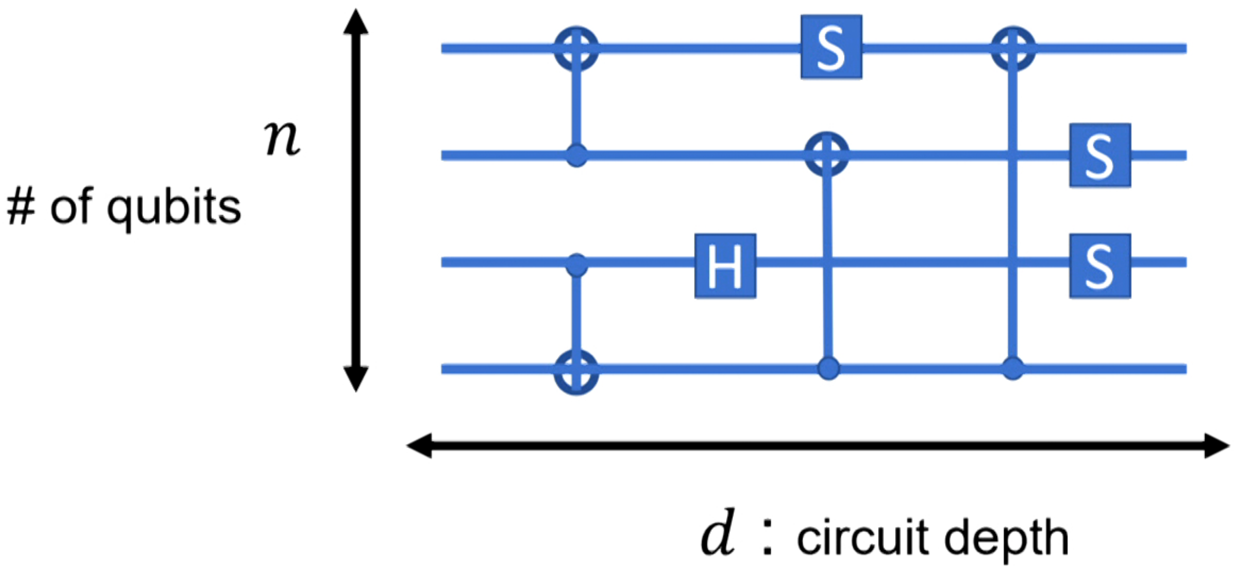


Circuit depth

Circuit depth is the number of time steps allowing for parallel gates.

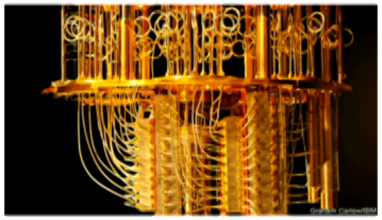


Shallow quantum circuits



We are interested in circuits with depth $d = O(1)$.

Why study shallow quantum circuits?

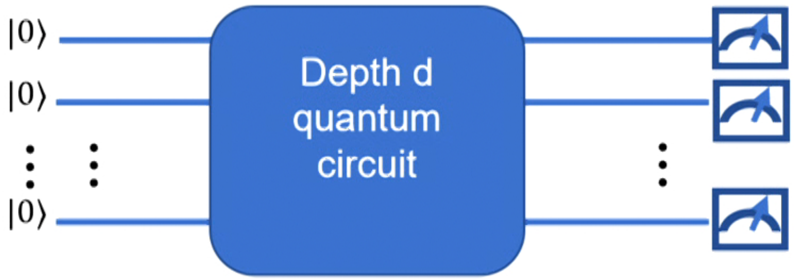


Small quantum computers: lack of error correction places limits on circuit size. So look at either few qubits (uninteresting) or low circuit depth.

Simplicity: a restricted model of quantum computation with structure that can be exploited.

Computational power…

What are shallow quantum circuits good for?

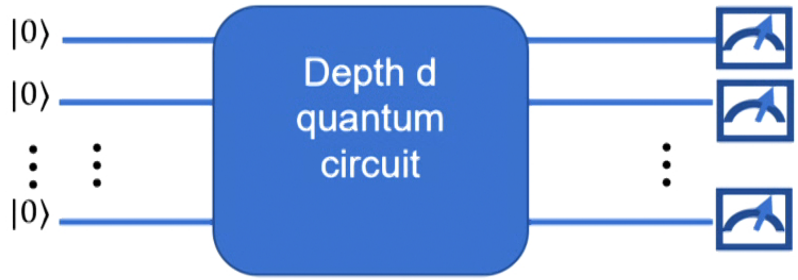


$z \in \{0,1\}^n$

Sample from classically inaccessible probability distributions

[Terhal Divincenzo 2002]
[Gao et al 17]
[Bermejo-Vega et al. 17]

What are shallow quantum circuits good for?

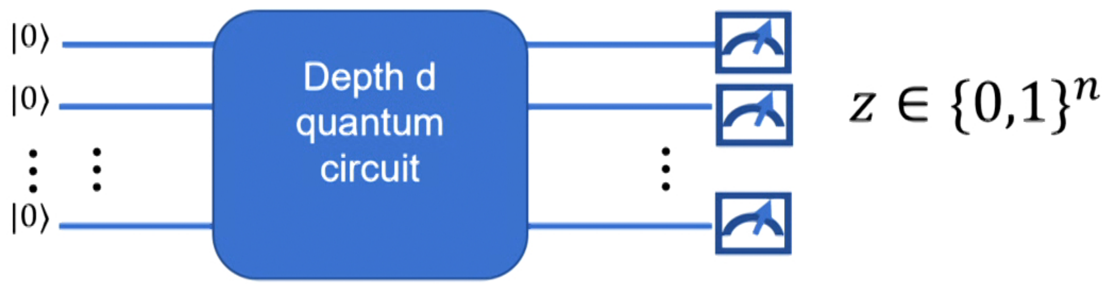


$z \in \{0,1\}^n$

Solve certain linear algebra problems faster than classical algorithms

[Bravyi, G., Koenig 18]
[Bene Watts, Kothari, Schaeffer, Tal 19]
[Bravyi, G., Koenig, Tomamichel 19]

What are shallow quantum circuits good for?



...Anything else?

Variational quantum algorithms

A recent family of "near-term" algorithms which has attracted great interest:

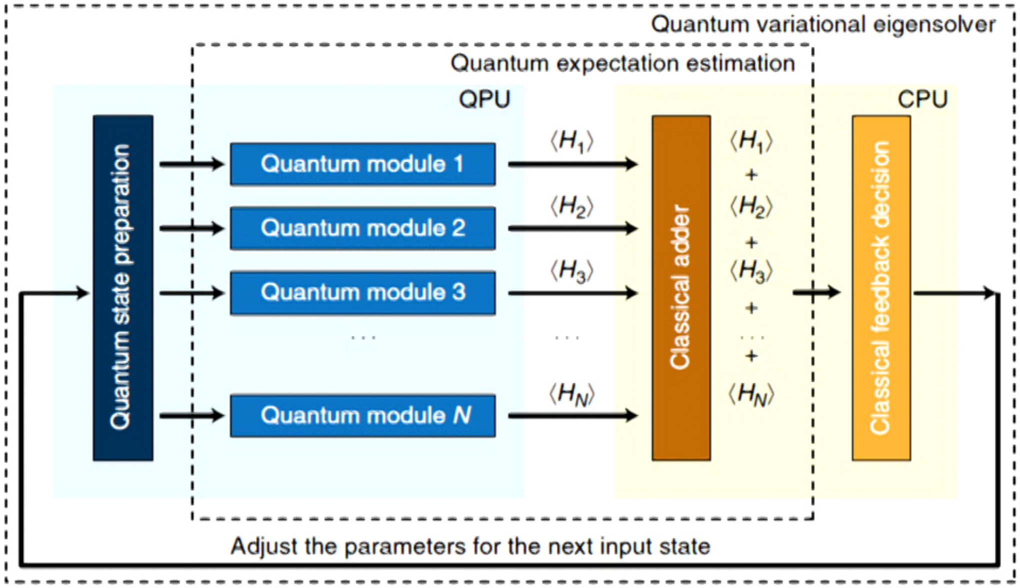


Image depicts the Variational Quantum Eigensolver paper, taken from [Peruzzo et al. 2013] …

Variational quantum algorithms

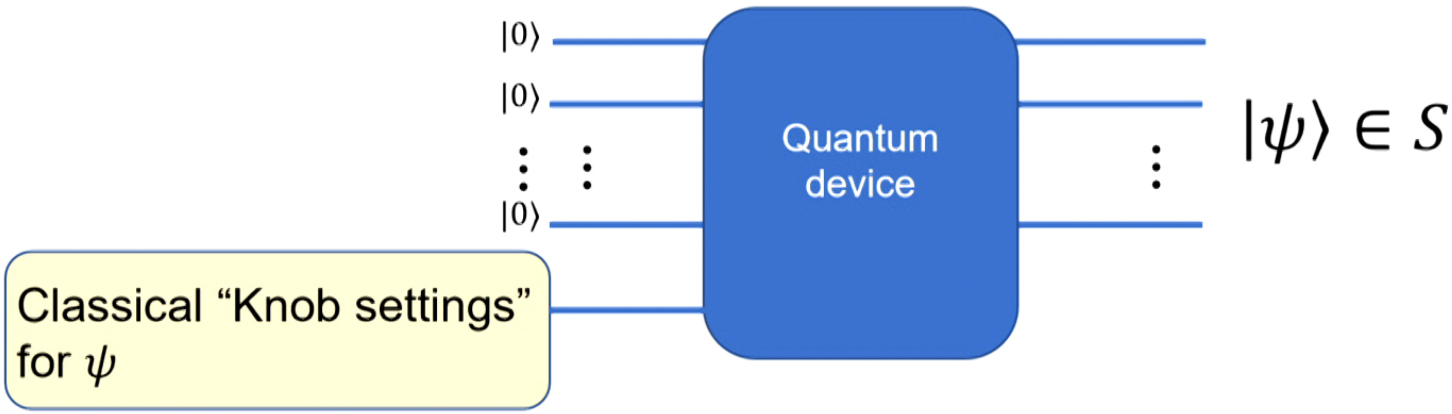
Goal: find the minimum energy of a given Hamiltonian.

$$H = \sum_i P_i \qquad\qquad E_{min} = \min_{\psi} \langle \psi | H | \psi \rangle$$

We will be interested in the case
where each term is an n-qubit Pauli operator

Variational quantum algorithms

Mild assumption #1: your quantum device can prepare a subset S of n-qubit states ψ



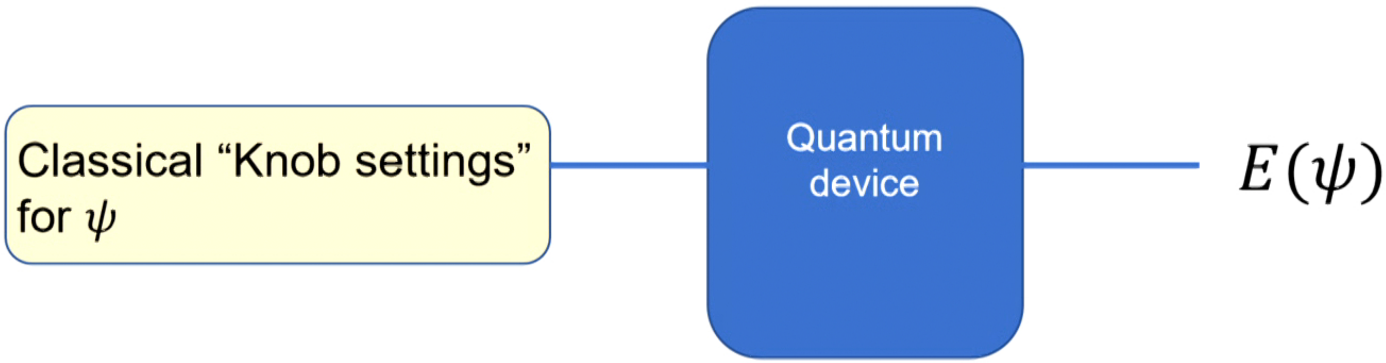
We will be interested in the case where S consists of states that can be prepared by constant-depth quantum circuits.

Variational quantum algorithms

Mild assumption #2: The device can be used to measure the energy of a given state $\psi \in S$

$$E(\psi) = \langle\psi|H|\psi\rangle = \sum \langle\psi|P_i|\psi\rangle$$

This can be achieved by computing each mean value $\langle\psi|P_i|\psi\rangle$ separately and then summing them.

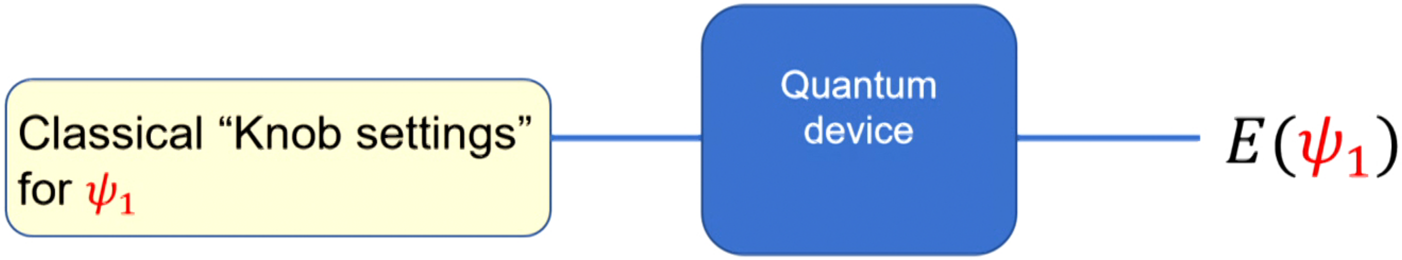


Variational quantum algorithms

A variational algorithm aims to compute the minimum energy **over states in S**

$$\min_{\psi \in S} \langle \psi | H | \psi \rangle$$

The algorithm uses the quantum device to compute energies and a classical computer to choose the knob settings:

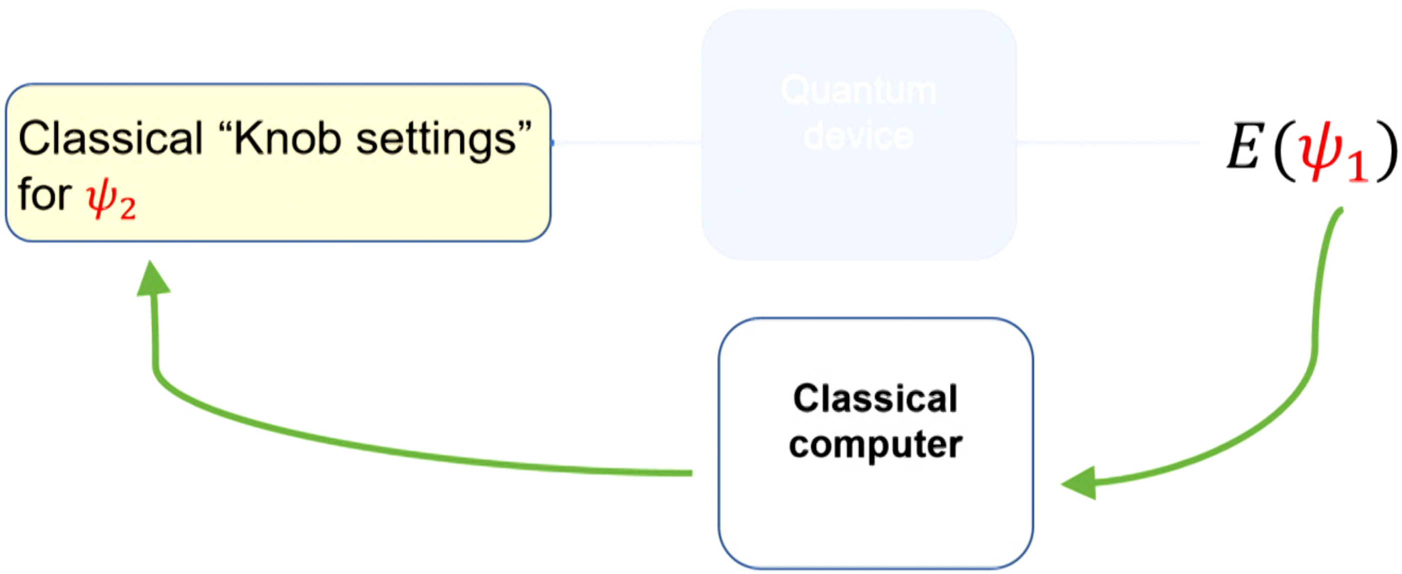


Variational quantum algorithms

A variational algorithm aims to compute the minimum energy **over states in S**

$$\min_{\psi \in S} \langle \psi | H | \psi \rangle$$

The algorithm uses the quantum device to compute energies and a classical computer to choose the knob settings:



What are variational quantum algorithms good for?

International Workshop on Quantum Technology and Optimization Problems
QTOP 2019: Quantum Technology and Optimization Problems pp 74-85 | Cite as

Variational Quantum Factoring

Authors | Authors and affiliations

Eric Anschuetz, Jonathan Olson, Alán Aspuru-Guzik, Yudong Cao

PHYSICAL REVIEW LETTERS

Highlights Recent Accepted Collections Authors Referees Search Pre

Quantum Machine Learning in Feature Hilbert Spaces

Maria Schuld and Nathan Killoran
Phys. Rev. Lett. **122**, 040504 – Published 1 February 2019

nature
International journal of science

Letter | Published: 13 September 2017

Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets

Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Markus Brink, Jerry M. Chow & Jay M. Gambetta

nature
International journal of science

Letter | Published: 13 March 2019

Supervised learning with quantum-enhanced feature spaces

Vojtěch Havlíček, Antonio D. Córcoles, Kristan Temme, Aram W. Harrow, Abhinav Kandala, Jerry M. Chow & Jay M. Gambetta

Lack of performance guarantees

Unfortunately, variational algorithms don't have performance guarantees as they are challenging to analyze:

Challenge #1: Is $\min_{\psi \in S} \langle \psi | H | \psi \rangle$ **close to** $\min_{\psi} \langle \psi | H | \psi \rangle$ **?**

Challenge #2: Is the output of the algorithm close to $\min_{\psi \in S} \langle \psi | H | \psi \rangle$ **?**

Do these algorithms really have any algorithmic speedup over classical computers…

Do we really need a quantum computer?

The quantum computer is only used to compute mean values of observables at the output of a quantum computation

$$\langle 0^n | U^\dagger O U | 0^n \rangle$$

How hard is this problem? Could we use a classical computer instead?

The mean value problem

Let U be a depth $d = O(1)$ quantum circuit.

Let O be a tensor product of single-qubit Hermitian operators

$$O = O_1 \otimes O_2 \otimes \cdots \otimes O_n$$

Assume $\|O_j\| \leq 1$

We are interested in estimating the mean value

$$\mu = \langle 0^n | U^\dagger O \, U | 0^n \rangle$$

The mean value problem

$$\mu = \langle 0^n | U^\dagger O\, U | 0^n \rangle$$

Interesting special case:

$$O = |x_1\rangle\langle x_1| \otimes |x_2\rangle\langle x_2| \otimes \cdots \otimes |x_n\rangle\langle x_n|$$

Then the mean value is an output probability of the quantum circuit

$$\mu = |\langle x|U|0^n\rangle|^2$$

The mean value problem

$$O = O_1 \otimes O_2 \otimes \cdots \otimes O_n$$

$$\mu = \langle 0^n \middle| U^\dagger O\, U \middle| 0^n \rangle$$

Additive error mean value problem

Given $\epsilon = \frac{1}{poly(n)}$, compute an estimate $\tilde{\mu}$ such that

$$|\tilde{\mu} - \mu| < \epsilon$$

The additive error mean value problem can be solved efficiently on a quantum computer.

The mean value problem

$$O = O_1 \otimes O_2 \otimes \cdots \otimes O_n$$

$$\mu = \langle 0^n \middle| U^\dagger O\, U \middle| 0^n \rangle$$

Relative error mean value problem

Given $\epsilon = \frac{1}{poly(n)}$, compute an estimate $\tilde{\mu}$ such that

$$|\tilde{\mu} - \mu| < \epsilon\mu$$

The relative error mean value problem is $\#P$-hard.

Complexity of the mean value problem

Quantum circuit U	**Observables O_j**	**Relative error**	**Additive error**
Polynomial size	Pos. semidefinite	#P-hard [16]	BQP-complete
Constant depth	Close to I	P [Thm. 1]	P [Thm. 1]
Constant depth	Pos. semidefinite	#P-hard [15, 16]	BQP Subexp. classical [Thm. 4]
2D Constant depth	Hermitian	#P-hard [15, 16] Subexp. classical [17]	BPP [Thm. 5]

Complexity of the mean value problem

Quantum circuit U	Observables O_j	Relative error	Additive error
Polynomial size	Pos. semidefinite	#P-hard [16]	BQP-complete
Constant depth	Close to I	P [Thm. 1]	P [Thm. 1]
Constant depth	Pos. semidefinite	#P-hard [15, 16]	BQP Subexp. classical [Thm. 4]
2D Constant depth	Hermitian	#P-hard [15, 16] Subexp. classical [17]	BPP [Thm. 5]

In the rest of the talk I will describe these 3 classical simulation algorithms…

Case 1: Single-qubit observables are each close to the identity

Quantum circuit U	Observables O_j	Relative error	Additive error
Polynomial size	Pos. semidefinite	#P-hard [16]	BQP-complete
Constant depth	Close to I	P [Thm. 1]	P [Thm. 1]
Constant depth	Pos. semidefinite	#P-hard [15, 16]	BQP Subexp. classical [Thm. 4]
2D Constant depth	Hermitian	#P-hard [15, 16] Subexp. classical [17]	BPP [Thm. 5]

Restricted family of tensor product observables

Suppose U is a depth-d quantum circuit and consider an observable

$$O = O_1 \otimes O_2 \otimes \cdots \otimes O_n$$

where

$$\|O_j - I\| \leq \frac{0.001}{2^{5d}}$$

Closeness to identity depends only on the depth d

For 2D circuits we can replace the RHS with $O(d^{-4})$

In this part of the talk we will be interested in obtaining a (highly demanding) relative error approximation to the mean value

$$\mu = \langle 0^n | U^\dagger O \, U | 0^n \rangle.$$

Restricted family of tensor product observables

$$\mu = \langle 0^n | U^\dagger O\, U | 0^n \rangle \qquad O = O_1 \otimes O_2 \otimes \cdots \otimes O_n$$

$$\|O_j - I\| \le \frac{0.001}{2^{5d}}$$

Example

Suppose we consider an output probability of a **noisy quantum circuit**

$$\mu' = \langle 0^n | \mathcal{E}^{\otimes n}(U^\dagger |0\rangle\langle 0|^n)\, U | 0^n \rangle.$$

$$\mathcal{E}(\rho) = (1-p)\rho + pX\rho X$$

Flip each bit with probability p

The noisy mean value is proportional to an ideal mean value:

$$\mu' = \frac{1}{2^n}\mu$$ with single-qubit observables $O_j = I + (1-2p)Z$

The above restriction is satisfied in a high noise regime $p \ge \frac{1}{2} - O(2^{-5d})$

Main result

$\mu = \langle 0^n | U^\dagger O\, U | 0^n \rangle$

$O = O_1 \otimes O_2 \otimes \cdots \otimes O_n$

$\|O_j - I\| \leq \frac{0.001}{2^{5d}}$

Theorem

Let $\delta \in \left(0, \frac{1}{2}\right)$ be given. There is a deterministic classical algorithm which outputs an estimate $\tilde{\mu}$ satisfying

$$|\log(\tilde{\mu}) - \log(\mu)| < \delta$$

The runtime of the algorithm is $(n\delta^{-1})^{c \cdot 2^d}$.

Solves the relative error mean value problem for this restricted set of observables
Runtime can be improved for 2D geometrically local circuits

The algorithm is based on a polynomial interpolation method due to Barvinok…

Classical simulation by polynomial interpolation

Define a polynomial

$$f(\epsilon) = \langle 0^n \big| U^\dagger O(\epsilon)\, U \big| 0^n \rangle \qquad O(\epsilon) = O_1(\epsilon) \otimes O_2(\epsilon) \otimes \cdots \otimes O_n(\epsilon)$$

$$O_j(\epsilon) = (1-\epsilon)I + \epsilon O$$

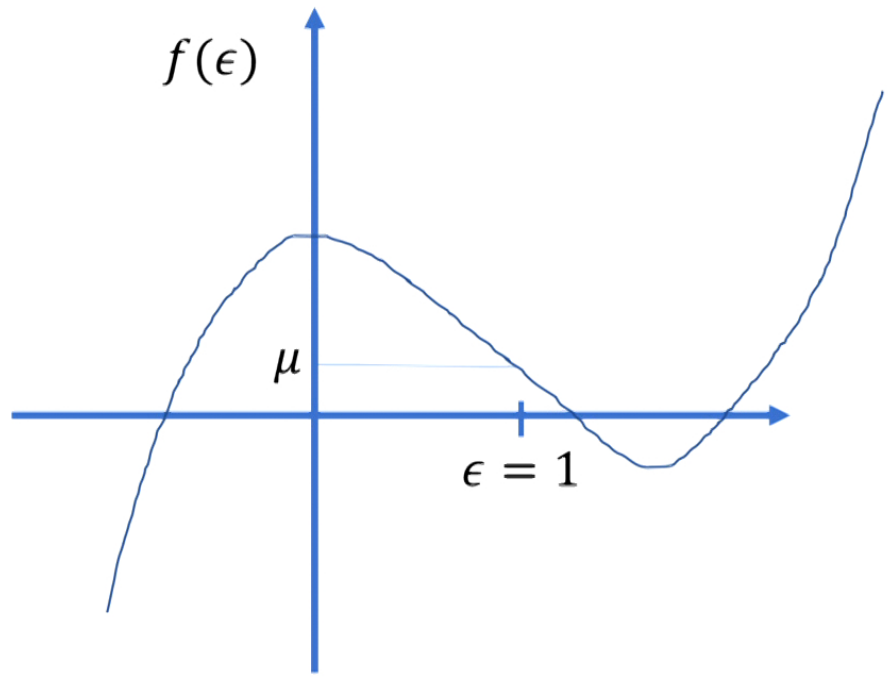
Note that $f(0) = 1$ and we aim to compute $\mu = f(1)$

Also note that derivatives $f^{(k)}(0)$ can be computed efficiently for small k

e.g.,

$$f^{(1)}(0) = \sum_{j=1}^{n} \langle 0^n | \underbrace{U^\dagger (O_j - I) U}_{\text{Acts nontrivially on } \leq 2^d \text{ qubits}} | 0^n \rangle$$

Classical simulation by polynomial interpolation



Since we know the function value and can compute derivatives at $\epsilon = 0$, it is natural to try to use a Taylor series approximation.

Barvinok: use Taylor series for the function $g(\epsilon) = \log\big(f(\epsilon)\big)$ instead…

Classical simulation by polynomial interpolation

Approximate the log by its truncated Taylor series

$$g(\epsilon) = \log f(\epsilon)$$

We want to compute $g(1)$

$$T_p(\epsilon) = g(0) + \sum_{k=1}^{p} \frac{\epsilon^k}{k!} g^{(k)}(0)$$

Theorem [Barvinok]

If the polynomial $f(\epsilon)$ is zero-free on the disk $|\epsilon| \leq 2$ then

$$|T_p(\epsilon) - g(\epsilon)| \leq \frac{n}{(p+1)2^p} \qquad |\epsilon| \leq 1$$

To achieve error δ we need only take $p = O(\log(n\delta^{-1}))$

Classical simulation by polynomial interpolation

$$g(\epsilon) = \log f(\epsilon)$$

To use Barvinok's method we need two ingredients:

1) We need to compute derivatives

$$g^{(1)}(0), \ldots, g^{(p)}(0) \qquad\qquad p = O(\log(n\delta^{-1}))$$

These can be computed efficiently from the derivatives $f^{(1)}(0), \ldots, f^{(k)}(0)$.

2) We need to show that $f(\epsilon)$ is zero-free on the disk $|\epsilon| \leq 2$ …

Zero-free region

$$f(\epsilon) = \langle 0^n | U^\dagger O(\epsilon)\, U | 0^n \rangle$$

$$O(\epsilon) = O_1(\epsilon) \otimes O_2(\epsilon) \otimes \cdots \otimes O_n(\epsilon)$$

$$O_j(\epsilon) = (1-\epsilon)I + \epsilon O$$

Theorem

Suppose $\|O_j - I\| \leq \gamma$. The polynomial f has no zeros in the disk

$$|\epsilon| \leq \frac{0.001}{\gamma 2^{5d}}$$

→ Depth of U

Choosing $\gamma = 0.001 \cdot 2^{-5d-1}$ suffices to make the disk radius equal to 2.

Proof sketch (zero-free region)

$$f(\epsilon) = \langle 0^n \middle| U^\dagger O_1(\epsilon) \otimes \cdots O_n(\epsilon)\, U \middle| 0^n \rangle$$

Write each 2×2 operator $O_j(\epsilon)$ as the upper left block of a 4×4 **unitary** $\boldsymbol{B_j}(\epsilon)$

$$f(\epsilon) = \langle 0^{2n} \middle| (U^\dagger \otimes I) B_1(\epsilon) \otimes \cdots B_n(\epsilon)(U \otimes I) \middle| 0^{2n} \rangle$$

Define $V_j(\epsilon) = (U^\dagger \otimes I) B_j(\epsilon)(U \otimes I)$

The $V_j(\epsilon)$ each act on 2^{d+1} qubits

Then $f(\epsilon) = \langle 0^{2n} | V_1(\epsilon) V_2(\epsilon) \ldots V_n(\epsilon) | 0^{2n} \rangle$

A constant depth circuit
Each gate is close to identity

Proof sketch (zero-free region)

$f(\epsilon) = \langle 0^{2n} | V(\epsilon) | 0^{2n} \rangle$ $\qquad V(\epsilon) = V_1(\epsilon) V_2(\epsilon) \ldots V_n(\epsilon)$

A constant-depth circuit
Each gate is close to identity

Now consider a probability distribution over $2n$-bit strings defined by

$$p_\epsilon(z) = \left| \langle z | V(\epsilon) | 0^{2n} \rangle \right|^2$$

Our goal is to show that $p_\epsilon(0^{2n}) > 0$ for all ϵ in the disk...

Proof sketch (zero-free region)

$$p_\epsilon(z) = \left|\langle z|V(\epsilon)|0^{2n}\rangle\right|^2$$

Let E_j be the event that the jth bit is 1. We show that each event $\{E_j\}_{1\le j\le 2n}$ occurs with a small probability $\boldsymbol{q = O(2^d \gamma|\epsilon|)}$ and is independent of most $\boldsymbol{D = O(2^{4d})}$ of the others…

$$\Pr_{p_\epsilon}[E_j] = \langle 0^{2n}|V(\epsilon)^\dagger|1\rangle\langle 1|_j\, V(\epsilon)\left|0^{2n}\right\rangle = \langle 0^{2n}\left|A_j\right|0^{2n}\rangle$$

All gates in $V(\epsilon)$ except $O(2^d)$ of them can be cancelled here.

Supported on $O(2^{2d})$ qubits

Proof sketch (zero-free region)

$$p_\epsilon(z) = \left|\langle z|V(\epsilon)|0^{2n}\rangle\right|^2$$

Let E_j be the event that the jth bit is 1. We show that each event $\{E_j\}_{1\le j\le 2n}$ occurs with a small probability $\boldsymbol{q = O(2^d\gamma|\epsilon|)}$ and is independent of most $\boldsymbol{D = O(2^{4d})}$ of the others…

The **Lovasz Local Lemma** then implies $p_\epsilon(0^{2n}) > 0$ as long as

$$\exp(1)\cdot q\cdot D < 1 \quad\Longrightarrow\quad |\epsilon| < O(1)\cdot 2^{-5d}\gamma^{-1}$$

Can the bound on zero-free radius be improved?

In the worst case it is possible for the zero-free radius to be exponentially small in the depth:

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle^{2^d} + |1\rangle^{2^d}\right) = U|0\rangle^{2^d}$$

Depth d

$$O_j(\epsilon) = I + \epsilon Z_j \qquad f(\epsilon) = \langle\psi|O_1(\epsilon)\otimes\cdots\otimes O_{2^d}(\epsilon)|\psi\rangle$$

$$= \frac{1}{2}\left((1+\epsilon)^{2^d} + (1-\epsilon)^{2^d}\right)$$

Has a root at

$$\epsilon_0 \approx \frac{i\pi}{2^{d+1}}$$

For random circuits the zero free radius is typically much larger…

Zero-free radius for random unitaries

Consider observables diagonal in the Z-basis:

$$O_j(\epsilon) = I + \epsilon Z_j \qquad f(\epsilon) = \langle 0^n | U^\dagger O_1(\epsilon) \otimes \cdots \otimes O_n(\epsilon) U | 0^n \rangle$$

Theorem

Let U be a random quantum circuit drawn from a unitary 2-design
The polynomial f has no zeros in a disk

$$|\epsilon| \leq 1 - O(n^{-1}\log(n))$$

Proof idea

Write $f(\epsilon) = 1 + \sum_{k=1}^{n} c_k \epsilon^k$ 2-design property gives $\mathbb{E}[|c_k|^2] \leq \frac{1}{2^n}\binom{n}{k}$

Use to show that w.h.p for ϵ in the disk we have $|f(\epsilon) - 1| \leq 1/2$

Case 2: Positive semidefinite observables

Quantum circuit U	**Observables O_j**	**Relative error**	**Additive error**
Polynomial size	Pos. semidefinite	#P-hard [16]	BQP-complete
Constant depth	Close to I	P [Thm. 1]	P [Thm. 1]
Constant depth	Pos. semidefinite	#P-hard [15, 16]	BQP Subexp. classical [Thm. 4]
2D Constant depth	Hermitian	#P-hard [15, 16] Subexp. classical [17]	BPP [Thm. 5]

Subexponential time classical algorithm

$$O = O_1 \otimes O_2 \otimes \cdots \otimes O_n \qquad \qquad \|O_j\| = 1$$

Theorem

Let $\delta \in \left(0, \frac{1}{2}\right)$ be given. There is a deterministic classical algorithm which outputs an estimate $\tilde{\mu}$ satisfying

$$\left|\tilde{\mu} - |\langle 0^n | U^\dagger O U | 0^n \rangle|\right| < \delta$$

The runtime of the algorithm is $e^{\tilde{O}(4^d \sqrt{n \cdot \log(\delta^{-1})})}$.

In general, the algorithm estimates the absolute value of the mean.

Solves the additive error MVP for pos. semidefinite observables.

Case 3: 2D shallow circuits

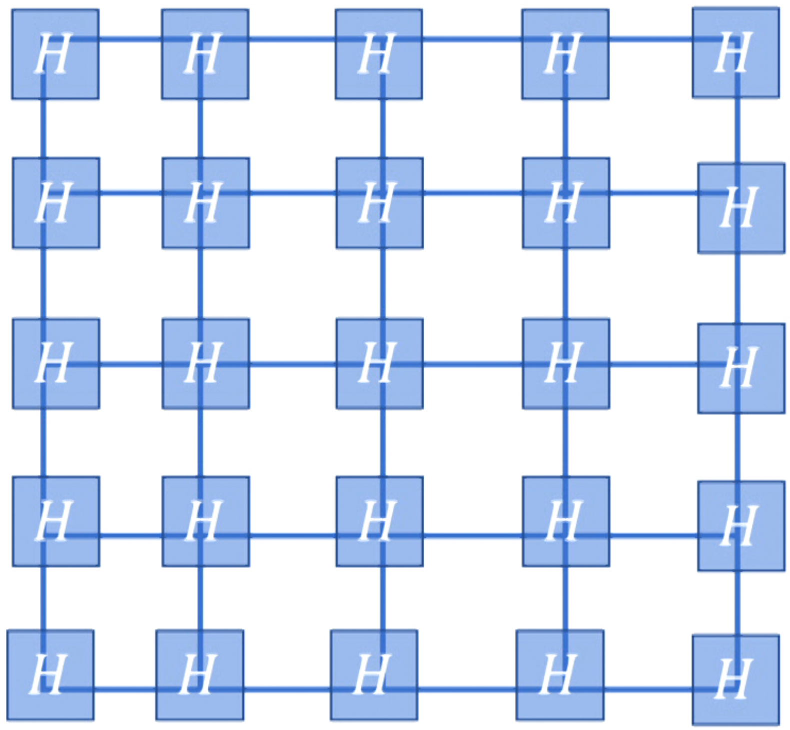
Quantum circuit U	Observables O_j	Relative error	Additive error
Polynomial size	Pos. semidefinite	#P-hard [16]	BQP-complete
Constant depth	Close to I	P [Thm. 1]	P [Thm. 1]
Constant depth	Pos. semidefinite	#P-hard [15, 16]	BQP Subexp. classical [Thm. 4]
2D Constant depth	Hermitian	#P-hard [15, 16] Subexp. classical [17]	BPP [Thm. 5]

2D shallow circuits

Suppose the qubits are located at the vertices of a 2D grid, and U is a depth d quantum circuit where each gate acts between nearest-neighbors.

Example:

$$U = \left(\prod_{(i,j)\in E} CZ_{ij} \right) H^{\otimes n} |0^n\rangle$$

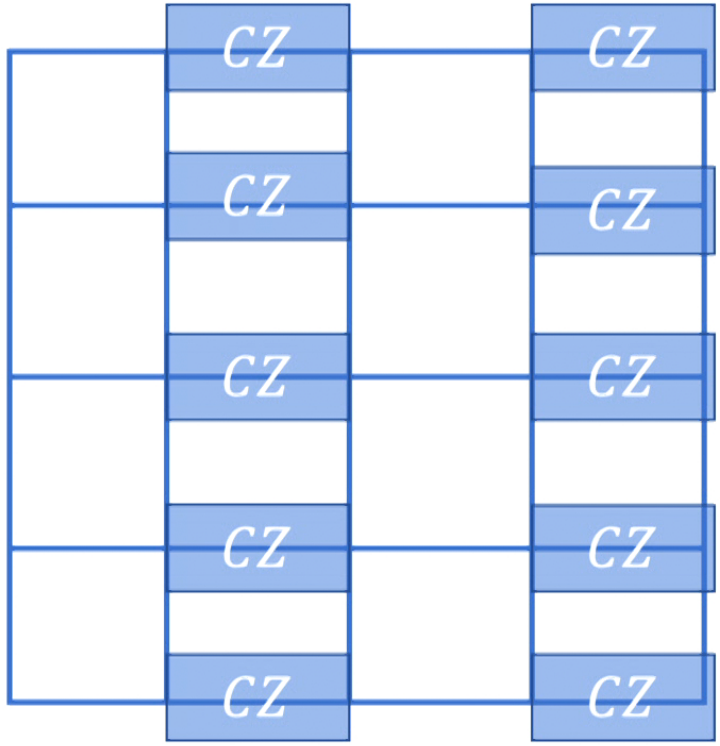


2D shallow circuits

Suppose the qubits are located at the vertices of a 2D grid, and U is a depth d quantum circuit where each gate acts between nearest-neighbors.

Example:

$$U = \left(\prod_{(i,j) \in E} CZ_{ij} \right) H^{\otimes n} |0^n\rangle$$

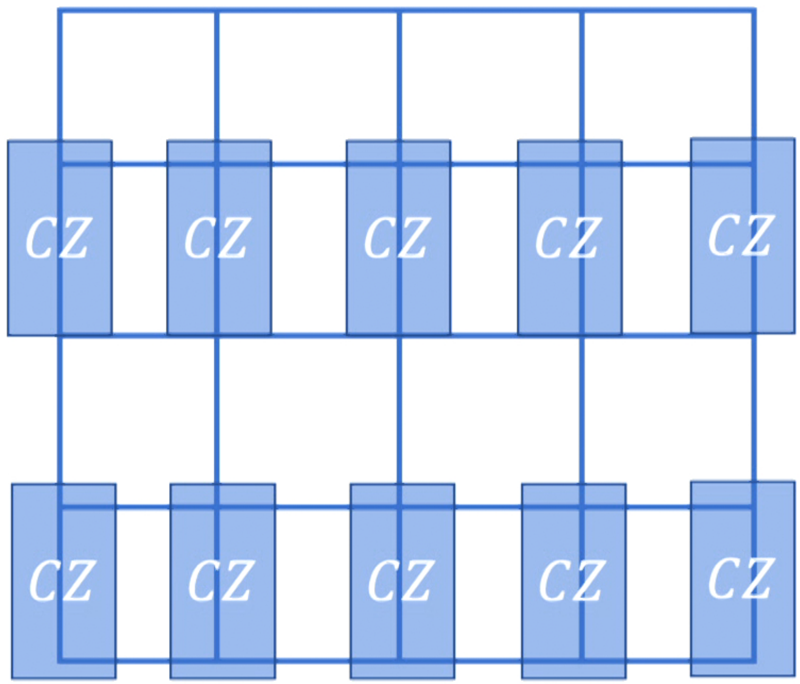


2D shallow circuits

Suppose the qubits are located at the vertices of a 2D grid, and U is a depth d quantum circuit where each gate acts between nearest-neighbors.

Example:

$$U = \left(\prod_{(i,j)\in E} CZ_{ij} \right) H^{\otimes n} |0^n\rangle$$



2D shallow circuits

Suppose the qubits are located at the vertices of a 2D grid, and U is a depth d quantum circuit where each gate acts between nearest-neighbors.

$$O = O_1 \otimes O_2 \otimes \cdots \otimes O_n \qquad\qquad \mu = \langle 0^n \big| U^\dagger O\, U \big| 0^n \rangle$$

Theorem

Let $\delta \in \left(0, \frac{1}{2}\right)$ be given. There is a randomized classical algorithm which, with probability at least 2/3, outputs an estimate $\tilde{\mu}$ satisfying

$$|\mu - \tilde{\mu}| \leq \delta$$

The runtime is $O(n\delta^{-2}2^{O(d^2)})$. Linear time!

Algorithm is based on an MPS representation and Monte Carlo method…

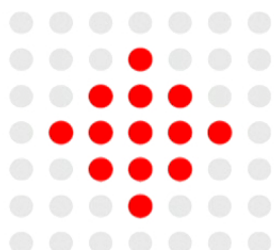
2D shallow circuit simulation

Express mean value as amplitude of a 2D constant depth circuit with **commuting gates**

$$\mu = \langle 0^n | U^\dagger O_1 \otimes O_2 \otimes \cdots \otimes O_n \, U | 0^n \rangle$$

$$= \langle 0^n | Q_n Q_{n-1} \dots Q_1 | 0^n \rangle \qquad Q_n = U^\dagger O_j \mathrm{U}$$

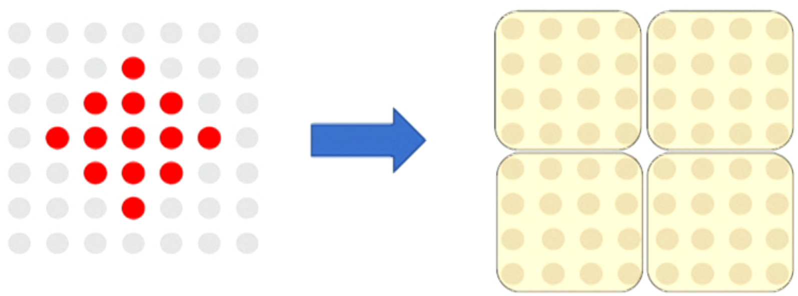
Each gate Q_j is supported on a $2d \times 2d$ square region centred at qubit j



2D shallow circuit simulation

$$\mu = \langle 0^n | Q_n Q_{n-1} \ldots Q_1 | 0^n \rangle$$

Coarse-grain: group the qubits into supersites of size $2d \times 2d$



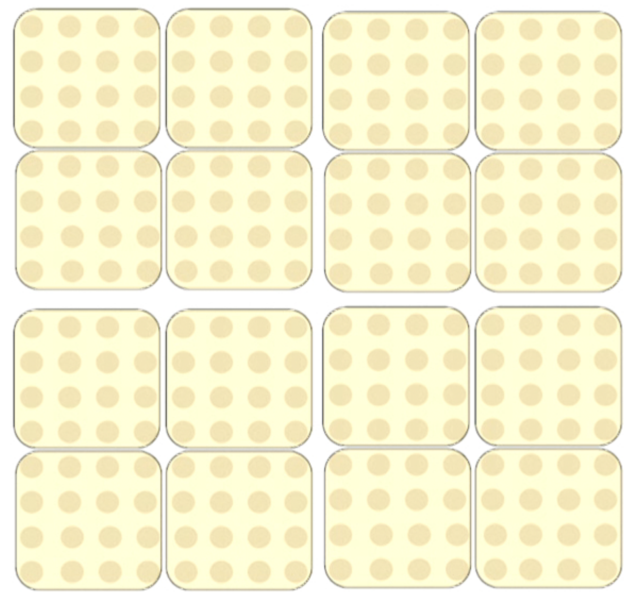
Each gate now acts nontrivially on 1 plaquette consisting of 4 supersites

2D shallow circuit simulation

Express mean value as inner product between two Matrix Product states

$$\mu = \langle 0^n | Q_n Q_{n-1} \dots Q_1 | 0^n \rangle$$

$$= \langle \Phi_{even} | \Phi_{odd} \rangle$$

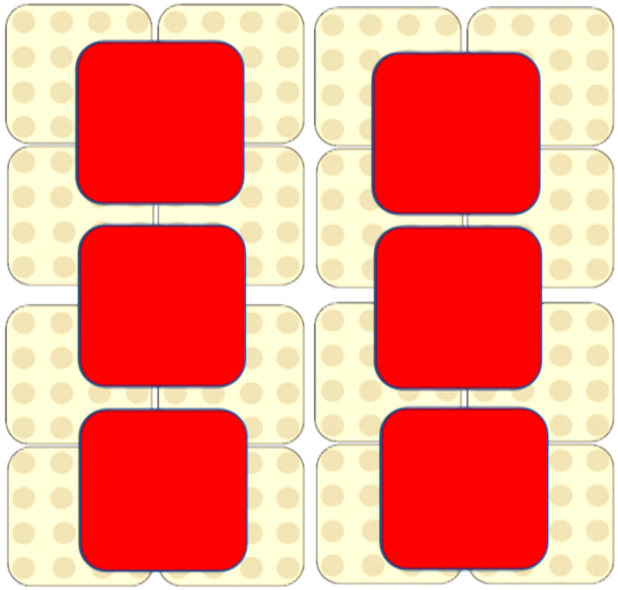


2D shallow circuit simulation

Express mean value as inner product between two Matrix Product states

$$\mu = \langle 0^n | Q_n Q_{n-1} \dots Q_1 | 0^n \rangle$$

$$= \langle \Phi_{even} | \Phi_{odd} \rangle$$

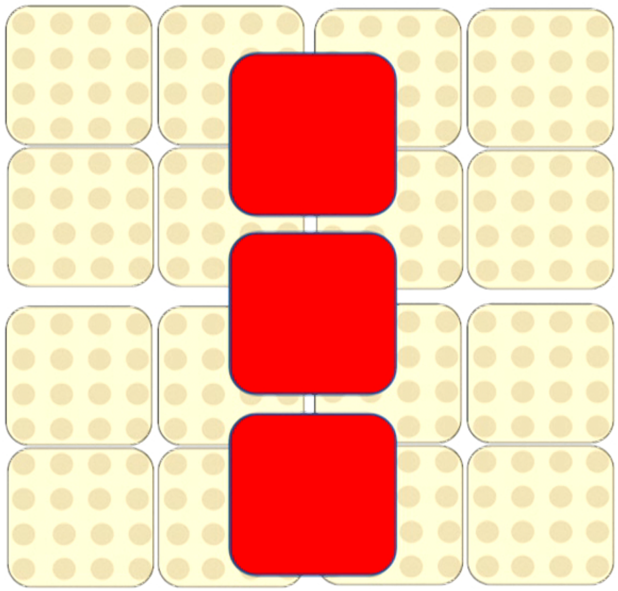


2D shallow circuit simulation

Express mean value as inner product between two Matrix Product states

$$\mu = \langle 0^n | Q_n Q_{n-1} \ldots Q_1 | 0^n \rangle$$

$$= \langle \Phi_{even} | \Phi_{odd} \rangle$$



Inner product between MPS can be estimated in polynomial time using a Monte Carlo method [Van den Nest 2009]

Open problems

Big question: what is the complexity of the additive-error mean value problem for constant-depth circuits?

Can the subexponential-time algorithm be generalized to the case of observables which may not be positive semidefinite?

Can the 2D algorithm be generalized to higher dimensional lattices?

Other applications of the zero-free region for quantum circuits?