Title: Reductions from weak to strong simulation of quantum systems

Speakers: Sergey Bravyi

Series: Perimeter Institute Quantum Discussions

Date: October 25, 2023 - 11:00 AM

URL: https://pirsa.org/23100109

Abstract: Classical simulation techniques are widely used in quantum computation and condensed matter physics. In this talk I will describe algorithms for classically simulating measurement of an n-qubit quantum state in the standard basis, that is, sampling a bit string from the probability distribution determined by the Born rule. Our algorithms reduce the sampling task (known as weak simulation) to computing poly(n) amplitudes of n-qubit states (strong simulation). Two classes of quantum states are considered: output states of polynomial-size quantum circuits and ground states of local Hamiltonians with an inverse polynomial energy gap. We show that our algorithm can significantly accelerate quantum circuit simulations based on tensor network contraction and low-rank stabilizer decompositions. To sample ground state probability distributions we employ the fixed-node Hamiltonian construction, previously used in Quantum Monte Carlo simulations to address the fermionic sign problem. We implement the proposed sampling algorithm numerically and use it to sample from the ground state of Haldane-Shastry Hamiltonian with up to 56 qubits.

Joint work with Giuseppe Carleo, David Gosset, and Yinchen Liu

---

Zoom link https://pitp.zoom.us/j/93297869296?pwd=TVpRdVJmU3lWZjVQM3NNKzBucVVRUT09

Reductions from weak to strong simulation of quantum systems

Sergey Bravyi (IBM Quantum)

Giuseppe Carleo
(EPFL)

David Gosset
(Perimeter, IQC)

Yinchen (Calvin) Liu
(IQC)

PRL 128, 220503 (2022); arXiv:2207.07044

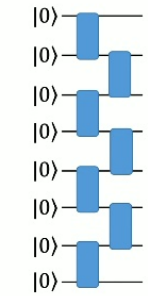
Perimeter Institute, Oct 25 (2023)

Success stories of classical simulation

Shallow circuits

Efficient simulation by tensor networks

Vidal 2003
Markov & Shi 2005
Aaronson & Chen 2016

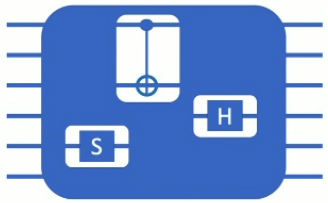


small depth

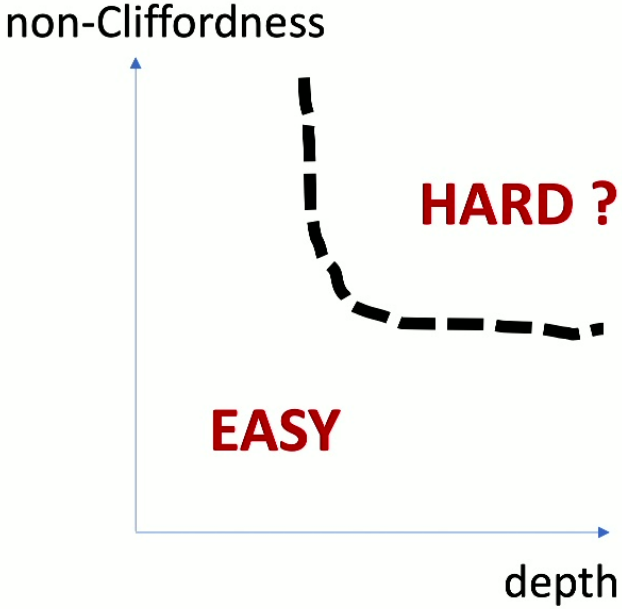
Near-Clifford circuits

Efficient simulation by the stabilizer formalism

Gottesman 1998
Aaronson & Gottesman 2003
SB & Gosset 2016



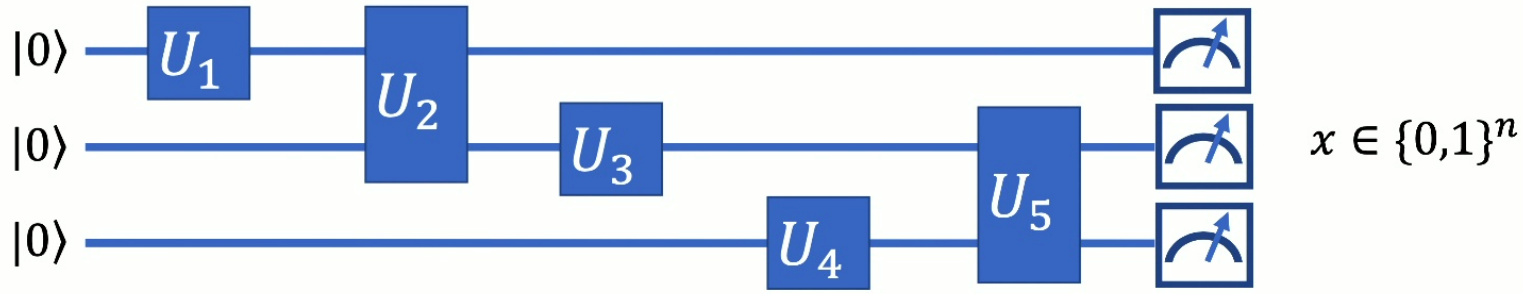
few non-Clifford gates



Strong vs Weak classical simulation

The goal of **strong simulation** is to **calculate** output amplitudes or probabilities

The goal of **weak simulation** is to **sample** output probability distribution

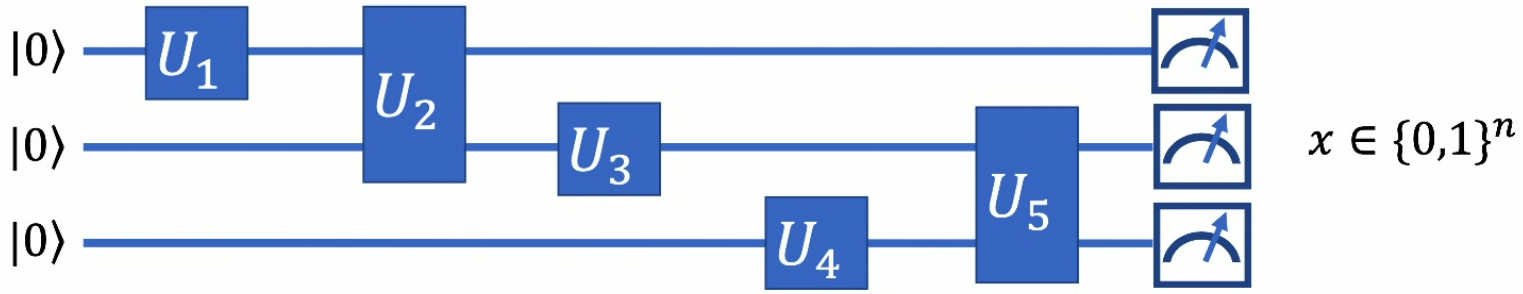


Strong vs Weak classical simulation

The goal of **strong simulation** is to **calculate** output amplitudes or probabilities

The goal of **weak simulation** is to **sample** output probability distribution

this talk



Strong vs Weak classical simulation

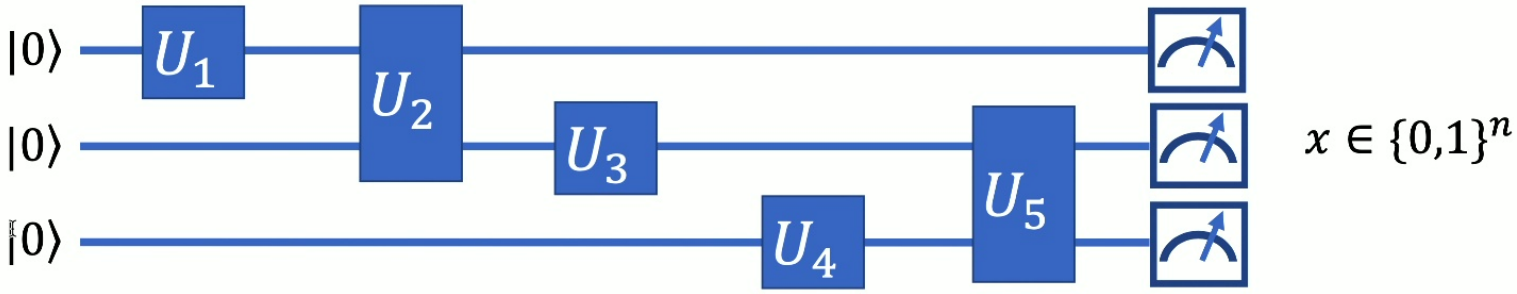
The goal of **strong simulation** is to **calculate** output amplitudes or probabilities

The goal of **weak simulation** is to **sample** output probability distribution 

- Almost all known classical simulation methods require a subroutine for calculating amplitudes or probabilities (strong simulation)
- Sampling requires an **efficient reduction from weak to strong simulation**
- Popular weak-to-strong reductions: Metropolis sampling, "qubit-by-qubit sampling"

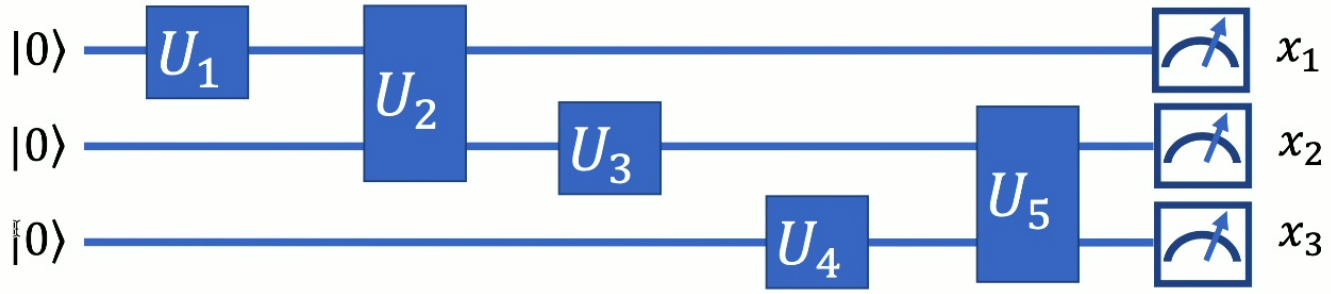
Qubit-by-qubit sampling

How to sample a bit string x from the output distribution of a quantum circuit ?



Qubit-by-qubit sampling

How to sample a bit string x from the output distribution of a quantum circuit ?



Strong simulator is called n times to calculate **marginal probabilities**

Marginal probabilities are expensive to compute using tensor network or Clifford-based methods. **Can we sample without computing marginals ?**

This talk: new efficient reductions

from weak to strong simulation

Subroutine for computing amplitudes

Sampling

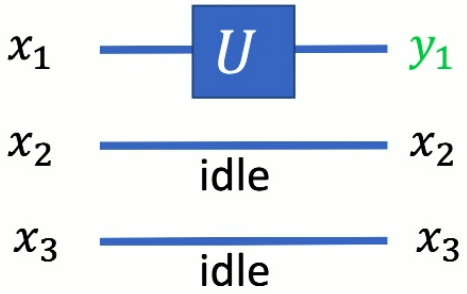
Gate-by-gate sampling

Sample a bit string x from a distribution $P_t(x)$ describing the first t gates

Sample x from $P_1(x) = |\langle x|U_1|000\rangle|^2$

Gate-by-gate sampling

- A unitary gate does not change the reduced density matrix of idle qubits that do not participate in the gate.
- Copy measurement outcomes on idle qubits from the previous time step
- Sample measurement outcomes from scratch only for non-idle qubits



Gate-by-gate sampling

Input: An n-qubit quantum circuit $U = U_m \cdots U_2 U_1$.
Output: $x \in \{0,1\}^n$ with probability $|\langle x|U|0^n\rangle|^2$.

1: $x \leftarrow 0^n$
2: **for** $t = 1$ to m **do**
3: $\quad A \leftarrow \{1, 2, \ldots, n\} \setminus \mathrm{supp}(U_t)$
4: $\quad S \leftarrow \{y \in \{0,1\}^n : y_A = x_A\}$
5: $\quad$ Sample $x \in S$ from the probability distribution $P_t(x) / \sum_{y \in S} P_t(y)$
6: **end for**
7: **return** x

Gate-by-gate sampling

Input: An n-qubit quantum circuit $U = U_m \cdots U_2 U_1$.
Output: $x \in \{0,1\}^n$ with probability $|\langle x|U|0^n\rangle|^2$.

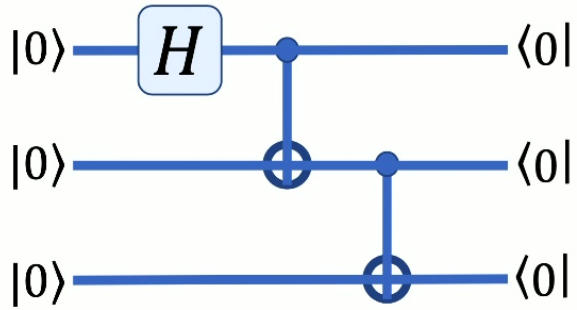
1: $x \leftarrow 0^n$
2: **for** $t = 1$ to m **do**
3: $A \leftarrow \{1, 2, \ldots, n\} \setminus \mathrm{supp}(U_t)$ idle qubits
4: $S \leftarrow \{y \in \{0,1\}^n : y_A = x_A\}$ candidates for the next sampled bit string
5: Sample $x \in S$ from the probability distribution $P_t(x)/\sum_{y \in S} P_t(y)$ call strong simulator to compute each probability
6: **end for**
7: **return** x

Gate-by-gate sampling

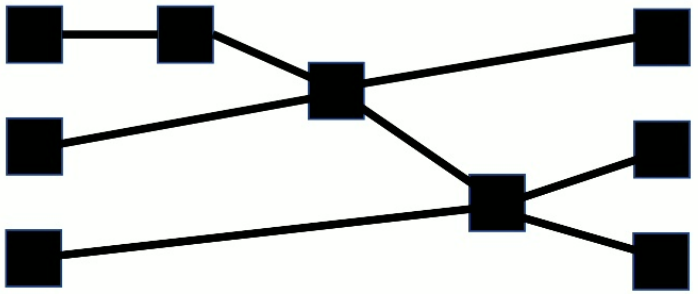
- One needs to call the strong simulator only for "superposition-creating" gates such as the Hadamard or X-rotation.
- The same algorithm can simulate adaptive circuits with mid-circuit measurements and classical feedforward.
- The algorithm is robust to small errors in the strong simulator. It suffices to compute amplitudes of a state ϵ-close to $U_t \cdots U_1|0^n\rangle$. Final simulation error $O(\epsilon m)$.

Tensor network simulators

- Pioneered by [Vidal 2003] and [Markov & Shi 2005]
- Accomplish a strong simulation
- Space complexity can be small (no need to store the entire state vector)



Amplitude of a quantum circuit



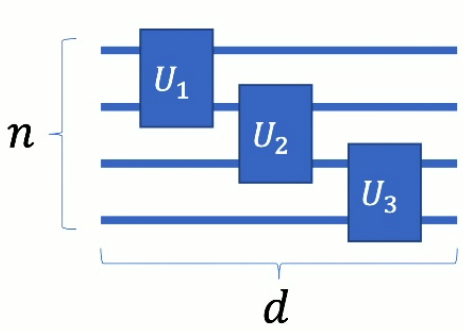
Contraction of a tensor network

Tensor network simulators

Let $f(n, d)$ be the runtime for computing a single amplitude of an n-qubit depth-d circuit.

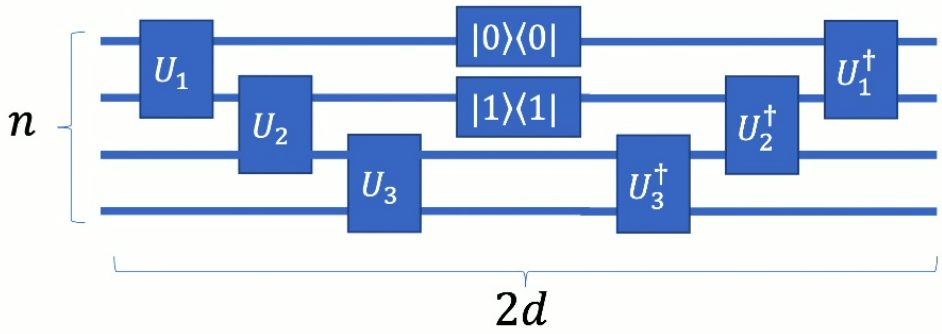
Gate-by-gate sampling:
strong simulator is called
to compute **amplitudes**

Runtime: $f(n, d)$



Qubit-by-qubit sampling:
strong simulator is called
to compute **marginal probabilities**

Runtime: $f(n, 2d)$



Tensor network simulators

Let $f(n,d)$ be the runtime for computing a single amplitude of an n-qubit depth-d circuit.

Gate-by-gate sampling:
strong simulator is called
to compute **amplitudes**

Runtime: $f(n,d)$

Qubit-by-qubit sampling:
strong simulator is called
to compute **marginal probabilities**

Runtime: $f(n,2d)$

State vector simulator: space 2^n, runtime $f(n,d) \sim nd2^n$.
Only 2X speedup.

Tensor network simulators

Let $f(n,d)$ be the runtime for computing a single amplitude of an n-qubit depth-d circuit.

Gate-by-gate sampling:
strong simulator is called
to compute **amplitudes**

Runtime: $f(n,d)$

Qubit-by-qubit sampling:
strong simulator is called
to compute **marginal probabilities**

Runtime: $f(n,2d)$

State vector simulator: space 2^n, runtime $f(n,d) \sim nd2^n$.
Only 2X speedup.

Feynman's sum-over-paths simulator: space $poly(n,d)$, runtime $f(n,d) \sim 2^{d\sqrt{n}}$
Aaronson and Chen, arXiv:1612.05903

Tensor network simulators

Let $f(n,d)$ be the runtime for computing a single amplitude of an n-qubit depth-d circuit.

Gate-by-gate sampling:
strong simulator is called
to compute **amplitudes**

Runtime: $f(n,d)$

Qubit-by-qubit sampling:
strong simulator is called
to compute **marginal probabilities**

Runtime: $f(n,2d)$

State vector simulator: space 2^n, runtime $f(n,d) \sim nd2^n$.
Only 2X speedup.

Feynman's sum-over-paths simulator: space $poly(n,d)$, runtime $f(n,d) \sim 2^{d\sqrt{n}}$
Aaronson and Chen, arXiv:1612.05903

Gate-by-gate achieves a quadratic speedup over qubit-by-qubit !

Tensor network simulators

Empirical comparison of gate-by-gate and qubit-by-qubit sampling algorithms for simulation of random 2D depth-16 circuits on a 7×7 grid of qubits.

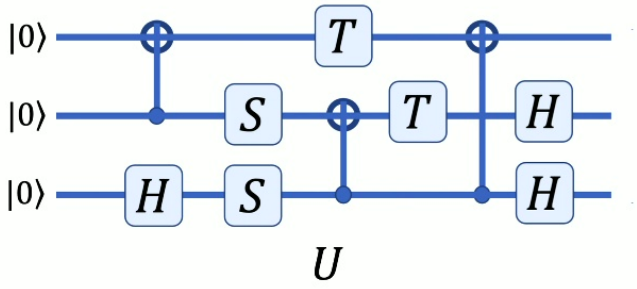
max degree of intermediate tensors

	29	**31**	**33**	**35**
log2(gate-by-gate FLOP count)	58	58	58	58
log2(qubit-by-qubit FLOP count)	75	73	71	69
speedup	10^5	3×10^4	7×10^3	2×10^3

Simulations were performed using CoTenGra software Gray & Kourtis 2021

Clifford simulators

The output state of a Clifford+T circuit with a few T-gates can be expressed as a linear combination of a few stabilizer states.



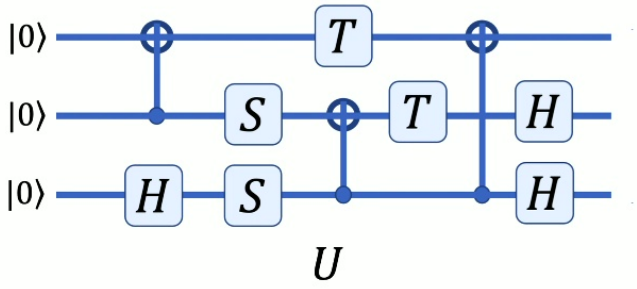
$$U|0^n\rangle = \sum_{a=1}^{\chi} c_a |\phi_a\rangle$$

stabilizer states

m_T non-Clifford gates $T = e^{i(\pi/8)Z}$

Clifford simulators

The output state of a Clifford+T circuit with a few T-gates can be expressed as a linear combination of a few stabilizer states.



m_T non-Clifford gates $T = e^{i(\pi/8)Z}$

$$U|0^n\rangle = \sum_{a=1}^{\chi} c_a \, |\phi_a\rangle$$

stabilizer states

number of stabilizer states

T-count m_T	χ
2	2
3	3
4	4
5	6
6	6

$$\chi \leq 2^{0.4\, m_T}$$

SB, Smith, Smolin 2015; SB and Gosset 2016; Qassim, Pashayan, Gosset 2021

Clifford simulators

Gate-by-gate sampling: strong simulator is called to compute **amplitudes**

$$\langle x|U_m \cdots U_2 U_1|0^n\rangle = \sum_{a=1}^{\chi} c_a \langle x|\phi_a\rangle$$

Runtime: $O(n^3 \cdot \chi)$ $\approx 2^{0.4\, m_T}$

cost of computing each amplitude $\langle x|\phi_a\rangle$

number of amplitudes

Runtime is dominated by simulating the largest subcircuits with $\approx m_T$ T-gates

Clifford simulators

Gate-by-gate sampling: strong simulator is called to compute **amplitudes**

$$\langle x|U_m \cdots U_2 U_1|0^n\rangle = \sum_{a=1}^{\chi} c_a \langle x|\phi_a\rangle$$

Runtime: $O(n^3 \cdot \chi)$ $\approx 2^{0.4\, m_T}$

Qubit-by-qubit sampling: strong simulator is called to compute **marginal probabilities**

$$\langle 0^n|U^\dagger(|y\rangle\langle y|\otimes I)U|0^n\rangle$$
$$= \sum_{a,b=1}^{\chi} c_a^* c_b \langle \phi_a|(|y\rangle\langle y|\otimes I)|\phi_b\rangle$$

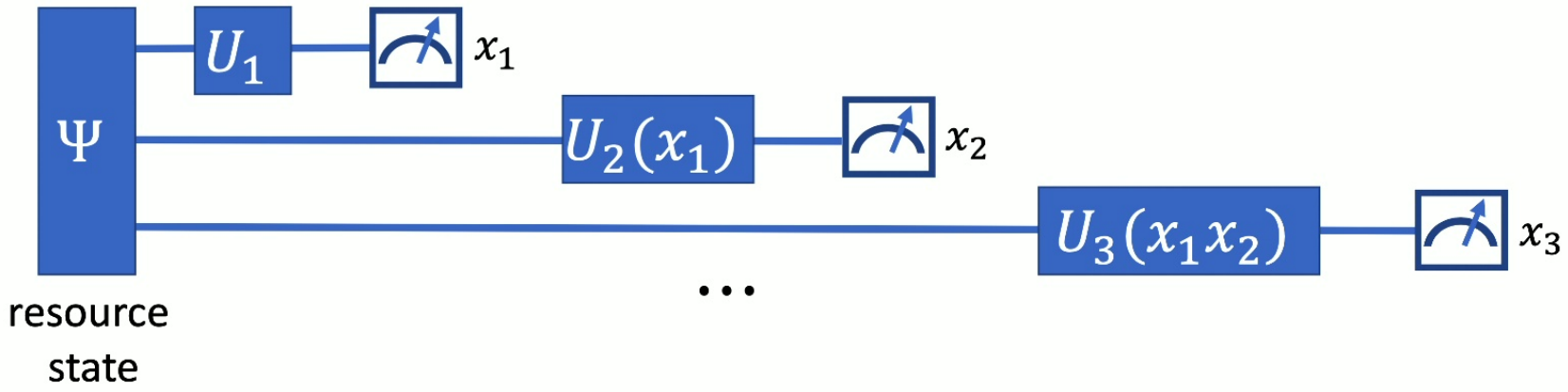
χ^2 inner products

Runtime: $O(n \cdot n^3 \cdot \chi^2)$ $\approx 2^{0.8\, m_T}$

number of inner products

Can we find examples when gate-by-gate algorithm runs in polynomial time while qubit-by-qubit algorithm requires exponential time ?

Yes, but this requires measurement-based quantum computation (MBQC)



Hardness of classical simulation for n-qubit MBQC with two types of resource states

	2D cluster state	
Amplitudes	#P-hard [1]	
Marginal probabilities	#P-hard [1]	
Sampling	BQP-complete [2]	

[1] Terhal & DiVincenzo 2002; Aaronson 2004
[2] Raussendorf & Briegel 2001

Hardness of classical simulation for n-qubit MBQC with two types of resource states

	2D cluster state	Surface code state
Amplitudes	#P-hard [1]	$O(n^3)$ [3]
Marginal probabilities	#P-hard [1]	
Sampling	BQP-complete [2]	

via fermionic linear optics

[1] Terhal & DiVincenzo 2002; Aaronson 2004
[2] Raussendorf & Briegel 2001
[3] SB & Raussendorf 2007

Hardness of classical simulation for n-qubit MBQC with two types of resource states

	2D cluster state	Surface code state
Amplitudes	#P-hard [1]	$O(n^3)$ [3]
Marginal probabilities	#P-hard [1]	#P-hard
Sampling	BQP-complete [2]	

via fermionic linear optics

reduction to counting perfect matchings in 3-regular graphs (new)

MBQC with the surface code: qubit-by-qubit algorithm has to solve #P-hard problems to compute marginals.

Part 2

weak vs strong simulation in quantum many-body physics

Our setup

Goal: sample a probability distribution $\pi(x)$ associated with the ground state ψ of a k-local Hamiltonian H describing n qubits with k-qubit interactions

$$\pi(x) = |\langle x|\psi\rangle|^2 \qquad H|\psi\rangle = E_0|\psi\rangle$$

- Non-degenerate ground state ψ
- Non-negligible spectral gap γ
- Subroutine for computing ratios of ground state amplitudes $\frac{\langle y|\psi\rangle}{\langle x|\psi\rangle}$
- H has real matrix elements in the standard basis (can be relaxed)

A solution that does not work

A naive solution is the standard Metropolis-Hastings Markov Chain Monte Carlo

- State space $S = \{x : \pi(x) > 0\}$
- Given a current state $x \in S$, propose a new state $y \in S$ by flipping a random subset of at most k bits of x
- Move to y with the probability $\min(1, \pi(y)/\pi(x))$

A solution that does not work

A naive solution is the standard Metropolis-Hastings Markov Chain Monte Carlo

Theorem [SB, Gosset, Liu 2021]

The Metropolis-Hastings MCMC has a mixing time $T \leq \frac{O(n^k)}{\gamma} \underbrace{\max_{x \neq y} \left| \frac{\langle y|H|x\rangle\langle x|\psi\rangle}{\langle y|\psi\rangle} \right|}_{s}$

Here s is a "sensitivity" parameter quantifying how much amplitudes of ψ can change upon flipping at most k bits. Rapid mixing only if $s \leq poly(n)$.

A solution that does not work

A naive solution is the standard Metropolis-Hastings Markov Chain Monte Carlo

Theorem [SB, Gosset, Liu 2021]

The Metropolis-Hastings MCMC has a mixing time $T \leq \frac{O(n^k)}{\gamma} \underbrace{\max_{x \neq y} \left| \frac{\langle y|H|x\rangle\langle x|\psi\rangle}{\langle y|\psi\rangle} \right|}_{s}$

Here s is a "sensitivity" parameter quantifying how much amplitudes of ψ can change upon flipping at most k bits. Rapid mixing only if $s \leq poly(n)$.

We prove that $s \leq poly(n)$ for any stoquastic Hamiltonian (sign-problem-free) ✅

No rapid mixing for general (non-stoquastic) Hamiltonians ❌

Sign problem

Reminder: a Hamiltonian H is said to be **stoquastic** (sign-problem-free) if it has real non-positive matrix elements in the standard basis.

$$\langle x|H|y\rangle \leq 0 \;\; \forall\, x \neq y$$

Main properties:

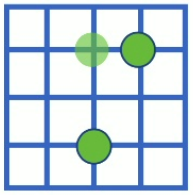
Ground state amplitudes are real and non-negative (Perron-Frobenius theorem)

The thermal Gibbs state $\rho = \frac{e^{-H/T}}{\mathrm{Tr}\, e^{-H/T}}$ is a real non-negative matrix.

Sign problem

Reminder: a Hamiltonian H is said to be **stoquastic** (sign-problem-free) if it has real non-positive matrix elements in the standard basis.

$$\langle x|H|y\rangle \leq 0 \;\; \forall\, x \neq y$$

Stoquastic		**Non-stoquastic**
Particle in a potential		Particle in a magnetic field
Hopping and interacting bosons		Interacting fermions
Some quantum ferromagnets		Quantum anti-ferromagnets

Curing the sign problem

Fixed-node Hamiltonian construction
Ceperley et al cond-mat/9412037

k-local Hamiltonian H → stoquastic Hamiltonian F

- Hamiltonians H and F have the same ground state ψ
- The energy gap of F is at least as large as the energy gap of H
- The matrix of F in the standard basis is sparse and efficiently computable, given a subroutine for computing ground state amplitudes

Curing the sign problem

Fixed-node Hamiltonian construction
Ceperley et al cond-mat/9412037

k-local Hamiltonian H → stoquastic Hamiltonian F

Informal definition of F:

- Set to zero all off-diagonal elements of H that "create the sign problem"
- Renormalize diagonal elements of H to preserve the ground state

Curing the sign problem

Fixed-node Hamiltonian construction

Ceperley et al cond-mat/9412037

k-local Hamiltonian H → stoquastic Hamiltonian F

$$S^+ = \{(x,y) \,:\, x \neq y \quad \text{and} \quad \langle\psi|x\rangle\langle x|H|y\rangle\langle y|\psi\rangle > 0\}$$
$$S^- = \{(x,y) \,:\, x \neq y \quad \text{and} \quad \langle\psi|x\rangle\langle x|H|y\rangle\langle y|\psi\rangle \leq 0\}$$

$$\langle x|F|y\rangle = \begin{cases} 0 & \text{if} \quad (x,y) \in S^+ \\ \langle x|H|y\rangle & \text{if} \quad (x,y) \in S^- \\ \langle x|H|x\rangle + \sum_{(x,z)\in S^+} \langle x|H|z\rangle \dfrac{\langle z|\psi\rangle}{\langle x|\psi\rangle} & \text{if} \quad x = y. \end{cases}$$

Curing the sign problem

Fixed-node Hamiltonian construction
Ceperley et al cond-mat/9412037

k-local Hamiltonian H → stoquastic Hamiltonian F

Recent work on curing/easing the sign problem :

Marvian, Lidar, and Hen	arXiv:1802.03408
Hangleiter, Roth, Nagaj, Eisert	arXiv:1906.02309
Klassen, Marvian, Piddock et al	arXiv:1906.08800
Ioannou, Piddock, Marvian et al	arXiv:2007.11964

These authors use a local unitary change of basis to cure or ease the sign problem for restricted classes of Hamiltonians. No need to compute ground state amplitudes.

Fixed-node works for all Hamiltonians but requires access to ground state amplitudes.

Our Markov Chain

We use a continuous-time version of the Metropolis-Hastings MCMC based on the fixed-node stoquastic Hamiltonian associated with H

$$\pi(x) = |\langle x|\psi\rangle|^2 = \lim_{t\to\infty} \pi_t(x) \qquad \frac{d\pi_t(x)}{dt} = \sum_y \langle x|G|y\rangle \pi_t(y)$$

Here G is a generator matrix such that $\langle x|G|y\rangle \geq 0$ is the rate of transitions $y \to x$

$$\langle x|G|y\rangle = E_0\delta_{x,y} - \langle x|F|y\rangle \frac{\langle x|\psi\rangle}{\langle y|\psi\rangle}$$

fixed-node Hamiltonian

Our Markov Chain

We use a continuous-time version of the Metropolis-Hastings MCMC based on the fixed-node stoquastic Hamiltonian associated with H

$$\pi(x) = |\langle x|\psi\rangle|^2 = \lim_{t\to\infty} \pi_t(x) \qquad \frac{d\pi_t(x)}{dt} = \sum_y \langle x|G|y\rangle \pi_t(y)$$

Theorem 1 (Rapid Mixing)

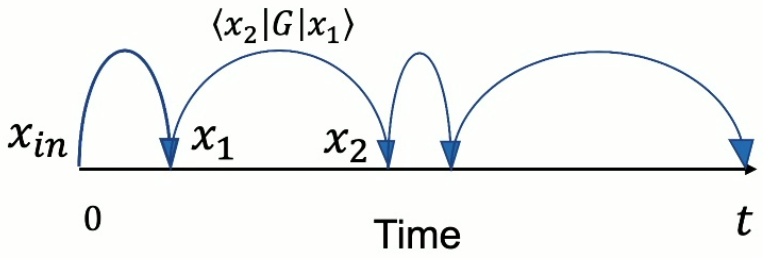
Suppose the chain is evolved for time t starting from a state $x_{in} \in \{0,1\}^n$

Then $\|\pi - \pi_t\|_1 \leq \dfrac{e^{-\gamma_F t}}{\sqrt{\pi(x_{in})}}$

Here $\gamma_F \geq \gamma$ is the energy gap of the fixed-node Hamiltonian.

Our Markov Chain

We simulate a continuous time Markov Chain using the standard Gillespie algorithm



The simulation outputs a random trajectory composed of easy-to-implement transitions

A transition $x \to y$ advances the time by a random variable δt sampled from $e^{-\delta t|\langle x|G|x\rangle|}$

We have to prove that the number of transitions required to reach the final time t is not too big (with a high probability).

Our Markov Chain

Theorem 2 (weak simulation): the number of calls to the subroutine computing ratios of ground state amplitudes required to sample $\pi(x)$ with an error ϵ is at most

$$T = \frac{\tilde{O}\left(\epsilon^{-1} n^{2k} \|H\|\right)}{\gamma_F}$$

Here $\gamma_F \geq \gamma$ is the energy gap of the fixed-node Hamiltonian.

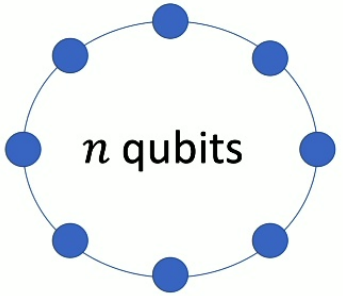
Caveat: our algorithm requires a "warm start" $x_{in} \in \{0,1\}^n$

We prove that a warm start always exists.

The algorithm can efficiently check whether a given initial state x_{in} is a warm start.

Example: Haldane-Shastry Hamiltonian

$$H = \sum_{1 \le i < j \le n} \frac{X_i X_j + Y_i Y_j + Z_i Z_j}{4\left(\frac{n}{\pi}\sin\frac{\pi(i-j)}{n}\right)^2}$$

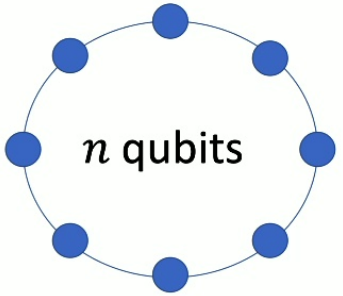


- Exactly solvable model
- Heisenberg anti-ferromagnetic $1/r^2$ power law interactions
- Universality class of $SU(2)_1$ Wess-Zumino-Witten CFT
- Non-stoquastic, not equivalent to free fermions

Haldane, PRL 60, 635 (1988); Shastry, PRL 60, 639 (1988)

Example: Haldane-Shastry Hamiltonian

$$H = \sum_{1 \le i < j \le n} \frac{X_i X_j + Y_i Y_j + Z_i Z_j}{4\left(\frac{n}{\pi}\sin\frac{\pi(i-j)}{n}\right)^2}$$



- Non-degenerate ground state with amplitudes

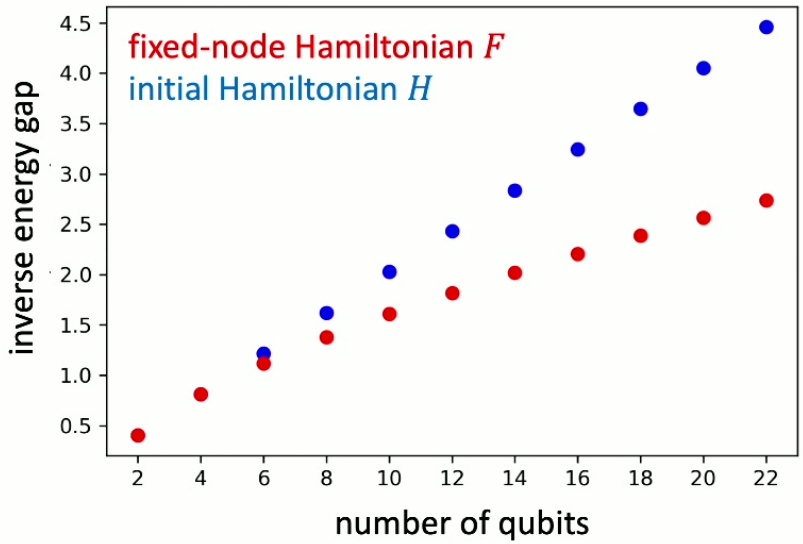
$$\langle x|\psi\rangle \propto \prod_{j=1}^{n}(-1)^{(j-1)x_j} \prod_{1 \le i < j \le n}\left(\sin\frac{\pi(i-j)}{n}\right)^{2x_i x_j} \qquad |x| = n/2$$

- Energy gap $\gamma \propto 1/n$
- Analytic formula for the ground state expected values such as $\langle\psi|Z_i Z_j|\psi\rangle$
  Stephan and Pollmann, arXiv:1608.06856

Example: Haldane-Shastry Hamiltonian

Reminder: the runtime of our sampling algorithm is proportional to the inverse energy gap $1/\gamma_F$ of the fixed-node Hamiltonian F

We found that the mapping to fixed-node Hamiltonian **amplifies the energy gap**



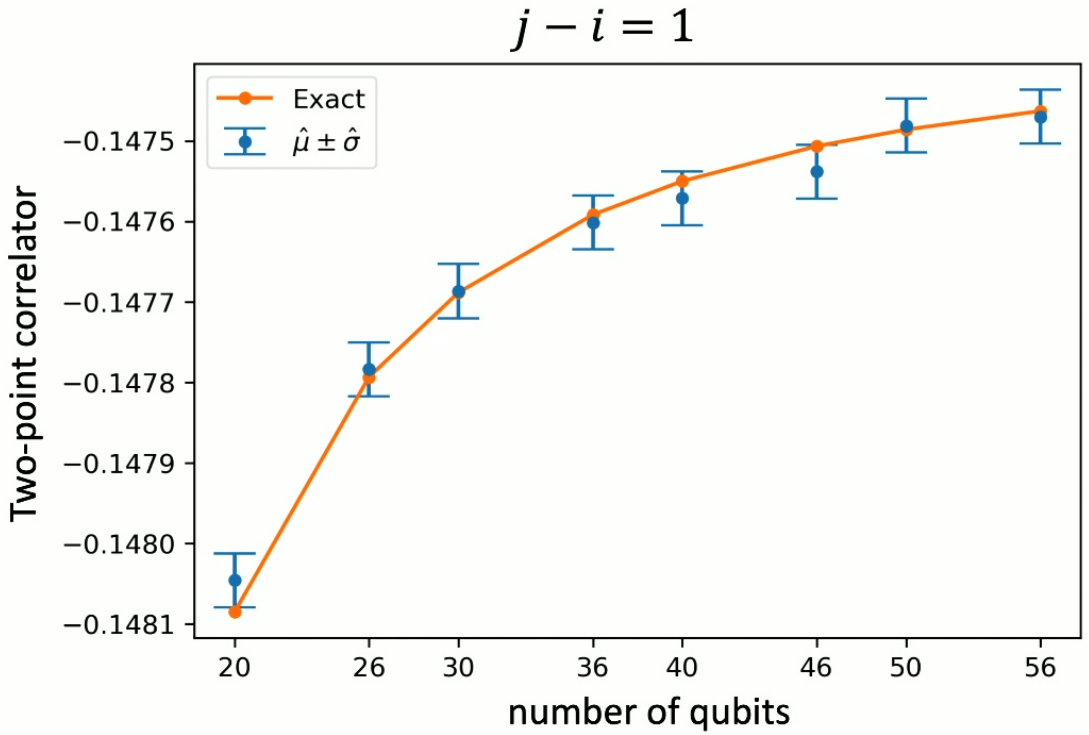
Theory: $\gamma_F \geq \gamma$

Numerics:

$\gamma \propto 1/n$ $\gamma_F \propto 1/n^{0.7}$

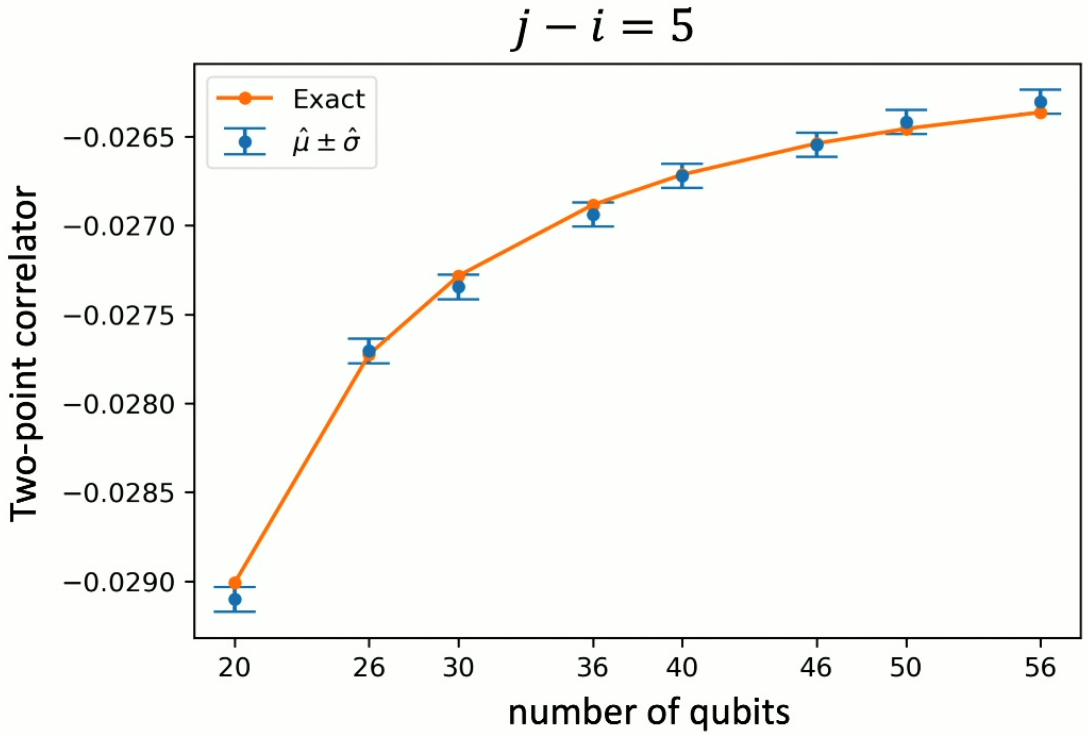
Example: Haldane-Shastry Hamiltonian

Comparison between the analytic formula for $\langle\psi|Z_iZ_j|\psi\rangle$ and our sampling algorithm with 5×10^6 transitions



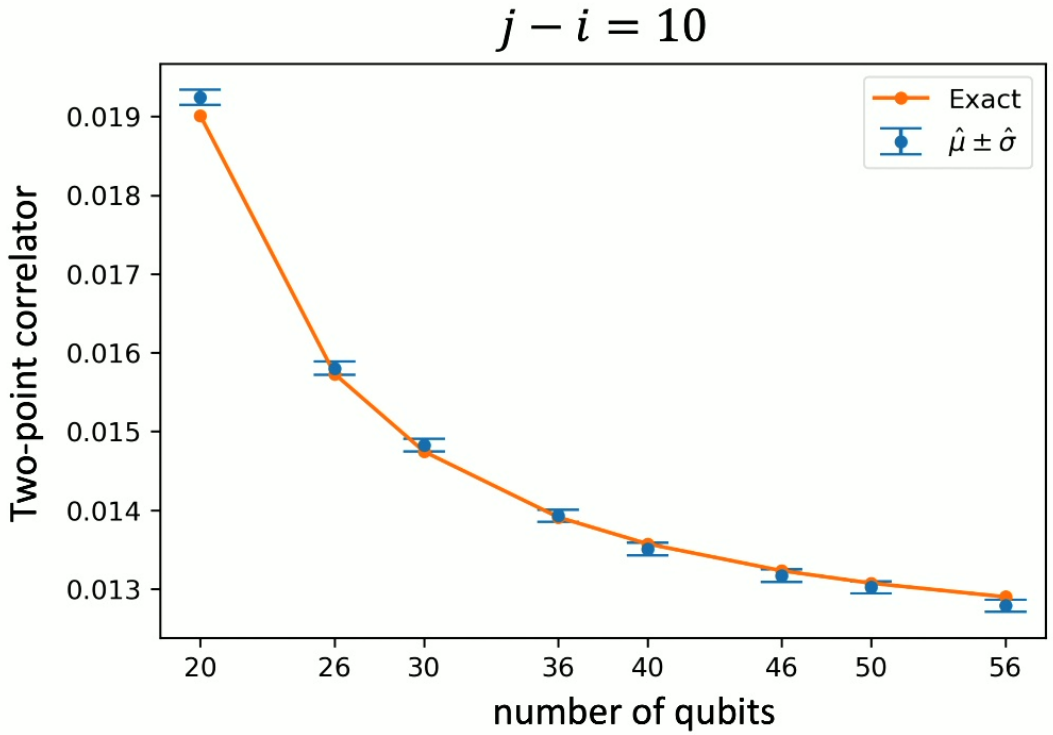
Example: Haldane-Shastry Hamiltonian

Comparison between the analytic formula for $\langle\psi|Z_iZ_j|\psi\rangle$ and our sampling algorithm with 5×10^6 transitions



Example: Haldane-Shastry Hamiltonian

Comparison between the analytic formula for $\langle\psi|Z_iZ_j|\psi\rangle$ and our sampling algorithm with 5×10^6 transitions



Open questions

- Other applications of the gate-by-gate sampling algorithm ?
  Example: approximating "forrelation" of degree-3 Boolean polynomials.

- Are there other weak-to-strong reductions that exploit more structure of ψ ?

- Can one perform a weak simulation directly skipping strong simulation ?
  Potential example:

  Michael Zurel, Cihan Okay, and Robert Raussendorf. Hidden variable model for universal quantum computation with magic states on qubits. *Physical Review Letters*, 125(26):260404, 2020.