Title: Quantum Advantages in Energy Minimization - VIRTUAL ONLY

Speakers: Leo Zhou

Series: Perimeter Institute Quantum Discussions

Date: February 28, 2024 - 11:00 AM

URL: https://pirsa.org/24020098

Abstract: Minimizing the energy of a many-body system is a fundamental problem in many fields. Although we hope a quantum computer can help us solve this problem better than classical computers, we have a very limited understanding of where a quantum advantage may be found. In this talk, I will present some recent theoretical advances that shed light on quantum advantages in this domain. First, I describe rigorous analyses of the Quantum Approximate Optimization Algorithm applied to minimizing energies of classical spin glasses. For certain families of spin glasses, we find the QAOA has a quantum advantage over the best known classical algorithms. Second, we study the problem of finding a local minimum of the energy of quantum systems. While local minima are much easier to find than ground states, we show that finding a local minimum under thermal perturbations is computationally hard for classical computers, but easy for quantum computers.

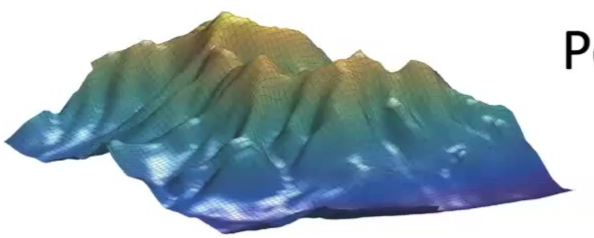
Zoom link TBA

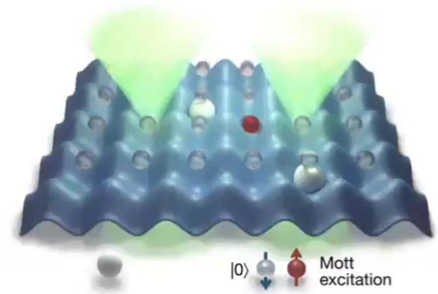
Quantum Advantages in Energy Minimization

Leo Zhou

Perimeter Institute Seminar

Feb 28, 2024





Energy minimization: a fundamental problem

Energy function $C(\boldsymbol{x}) : \mathcal{X} \to \mathbb{R}$

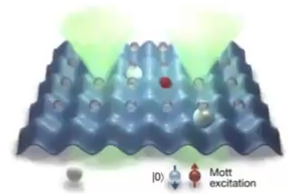
Want $\boldsymbol{x}^* \in \mathcal{X}$
so $C(\boldsymbol{x}^*)$ is minimized

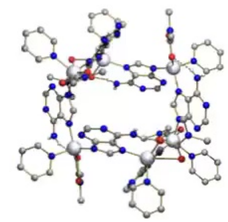
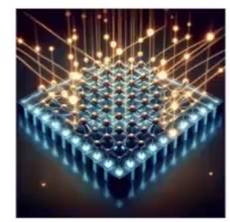
Classical systems



$$\mathcal{X} = \{\pm 1\}^n \text{ or } \mathbb{R}^n$$

Quantum systems



$$\mathcal{X} = \{|\psi\rangle\} \subseteq \mathbb{C}^{2^n} \text{ or } \{\boldsymbol{\rho}\}$$

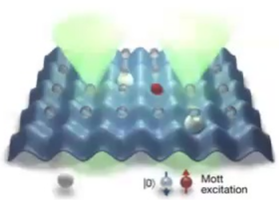
Minimize energy *better* with *quantum computers*?

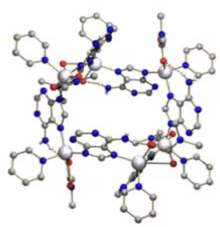
Classical computers are already (surprisingly) successful…

Classical systems $\{\pm 1\}^n$ or $\mathbb{R}^n$

Classical algorithms:
Linear Programming, SDP relaxation, Stochastic gradient descent, Bayesian optimization, Simulated annealing, …

Quantum systems $\mathbb{C}^{2^n}$



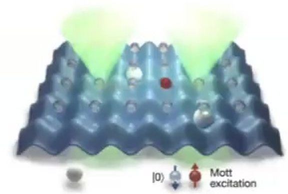


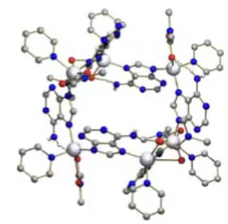
Classical algorithms:
Tensor network, DMRG, Quantum Monte-Carlo, Density Functional Theory, Neural network ansatz, …

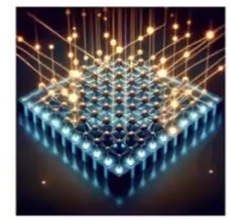
Today: Rigorous evidence of quantum advantages for minimizing energy of both classical and quantum systems!

Classical systems $\{\pm 1\}^n$ or $\mathbb{R}^n$

Quantum systems $\mathbb{C}^{2^n}$





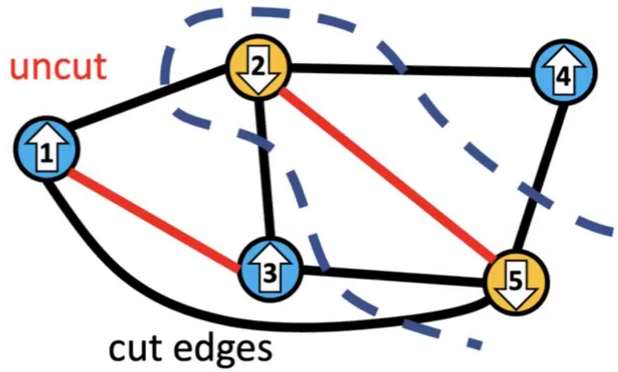


1. **qAdvantage** for finding near-ground states of spin glasses

2. **qAdvantage** for find local minima of quantum systems

A Classical Problem: **MaxCut** (diluted spin glass)

Goal: Find a **bipartition** of vertices that cuts the most edges



$$C = \sum_{\langle i,j \rangle} \frac{1}{2}(1 - z_i z_j)$$

Equivalent to minimizing $\tilde{C} = \sum_{\langle i,j \rangle} z_i z_j$

Random Guessing or **Greedy Search** guarantees for any graph,

$$C(\boldsymbol{z})/C_{\max} \geq 0.5$$

Semidefinite Programming (SDP) algorithm improves worst-case guarantee

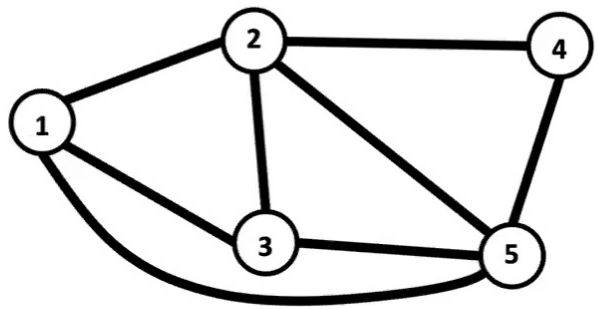
$$C(\boldsymbol{z})/C_{\max} \geq 0.878$$

[Goemans Williamson 1995]

Quantum Approximate Optimization Algorithm (QAOA)

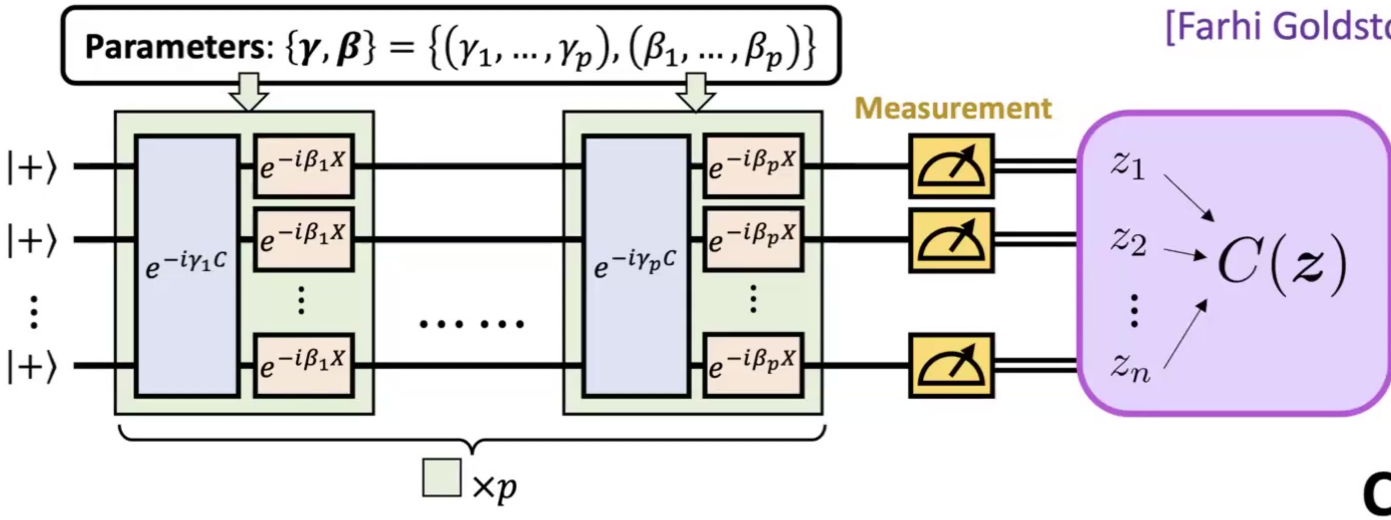
[Farhi Goldstone Gutmann 2014]

$$C(\boldsymbol{z}) = \sum_{\langle ij \rangle} \frac{1}{2}(1 - z_i z_j)$$



Quantum Approximate Optimization Algorithm (QAOA)

[Farhi Goldstone Gutmann 2014]



$$C = \sum_{\langle ij \rangle} \frac{1}{2}(1 - Z_i Z_j)$$

$$B = \sum_{i=1}^{n} X_i$$

$$|\boldsymbol{\gamma}, \boldsymbol{\beta}\rangle = e^{-i\beta_p B} e^{-i\gamma_p C} \dots e^{-i\beta_1 B} e^{-i\gamma_1 C} \; |+\rangle^{\otimes n}$$

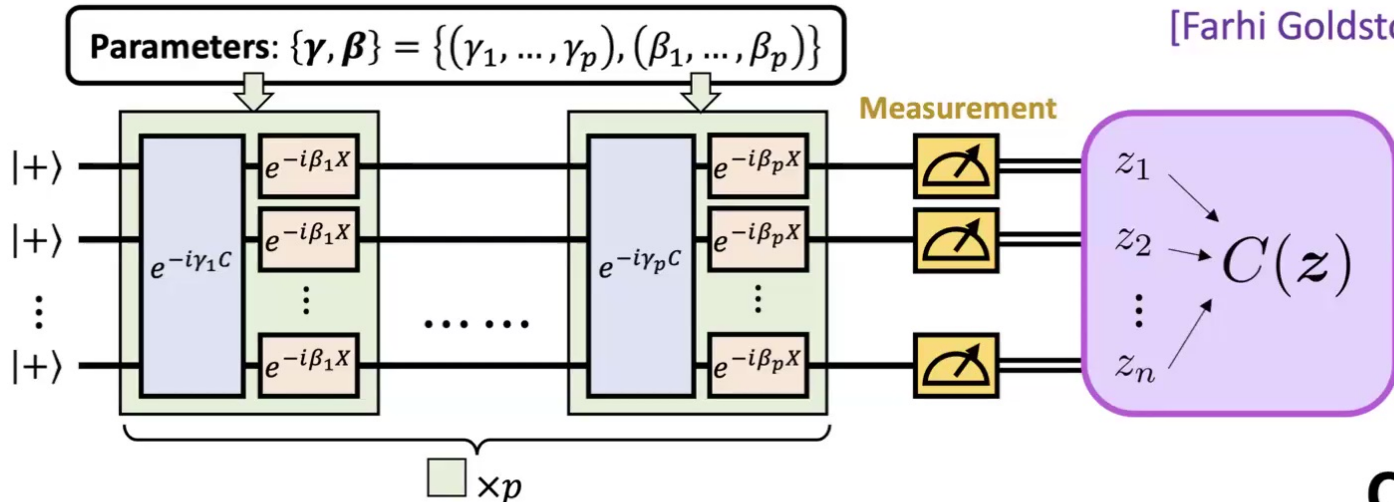
Choose $\gamma, \boldsymbol{\beta}$ to maximize $\langle C \rangle$

Various **heuristics** exist to efficiently choose good parameters

- Parameter pattern & interpolation between depths [***LZ*** Wang Choi Pichler Lukin 2018, PRX 2020]

Quantum Approximate Optimization Algorithm (QAOA)

[Farhi Goldstone Gutmann 2014]



$$C = \sum_{\langle ij \rangle} \frac{1}{2}(1 - Z_i Z_j)$$

$$B = \sum_{i=1}^{n} X_i$$

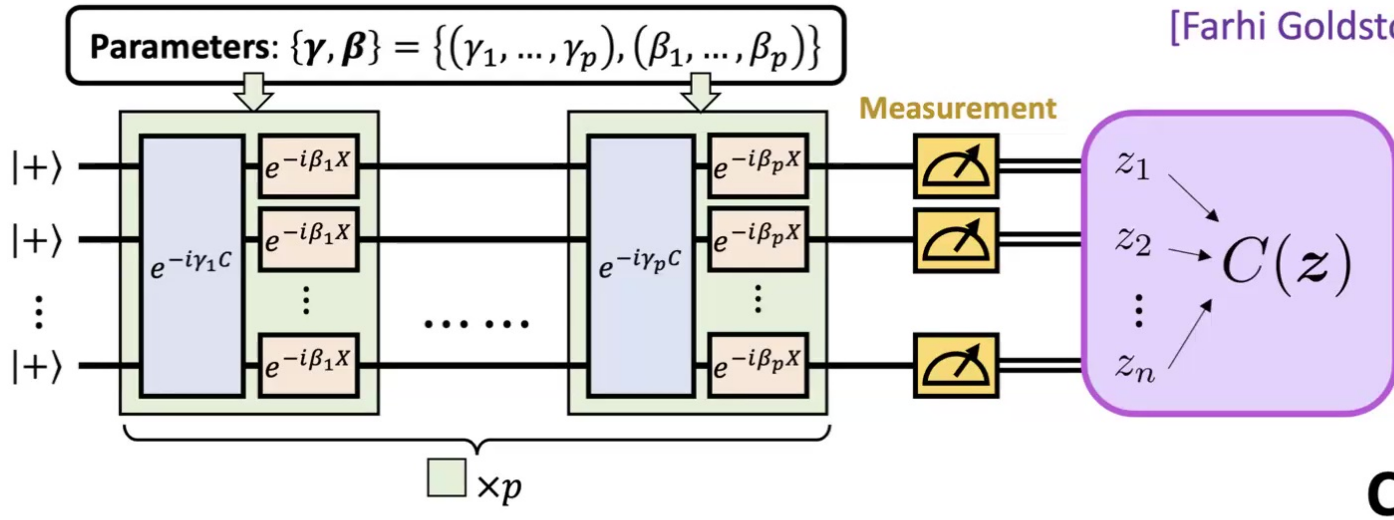
$$|\boldsymbol{\gamma}, \boldsymbol{\beta}\rangle = e^{-i\beta_p B} e^{-i\gamma_p C} \cdots e^{-i\beta_1 B} e^{-i\gamma_1 C} \ |+\rangle^{\otimes n}$$

Choose $\gamma, \boldsymbol{\beta}$ to maximize $\langle C \rangle$

Various **heuristics** exist to efficiently choose good parameters

- Parameter pattern & interpolation between depths [***LZ*** Wang Choi Pichler Lukin 2018, PRX 2020]
- Classical Warm Start [Egger *et al* 2020], Annealing-inspired [Sacks Serbyn 2021], Parameter Transfers [Galda *et al* 2021], Graph Neural Network [Jain *et al* 2022] ……

Quantum Approximate Optimization Algorithm (QAOA)



[Farhi Goldstone Gutmann 2014]

$$C = \sum_{\langle ij \rangle} \frac{1}{2}(1 - Z_i Z_j)$$

$$B = \sum_{i=1}^{n} X_i$$

$$|\boldsymbol{\gamma}, \boldsymbol{\beta}\rangle = e^{-i\beta_p B} e^{-i\gamma_p C} \ldots e^{-i\beta_1 B} e^{-i\gamma_1 C} \; |+\rangle^{\otimes n}$$

Choose γ, β to maximize $\langle C \rangle$

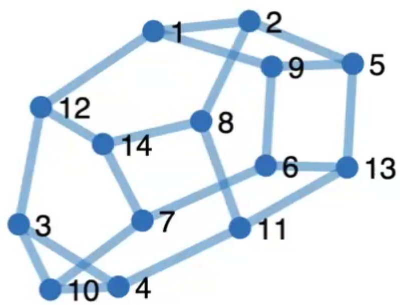
- Simple & easy implementation (e.g. ions (2019), superconductors (2020) and atoms (2022))
- Cannot classically sample ("supremacy") even @ p=1 [Farhi Harrow 2016] [Krovi 2022]
- ***Guaranteed to get $C_{\max}$ as depth $p \to \infty$!***

Performance guarantee of the QAOA

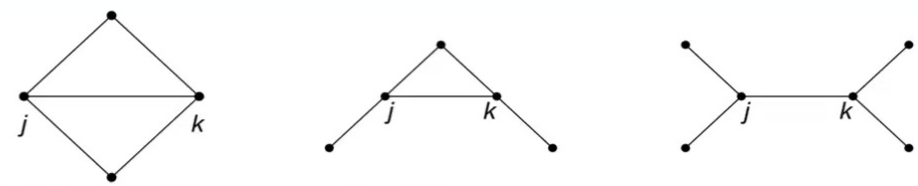
- Early approach: Analyze performance via “subgraphs”

$|s\rangle = |+\rangle^{\otimes n}$

<u>Example</u>: MaxCut on 3-regular graphs



$$p = 1 \quad \underbrace{\langle s| e^{i\gamma C} e^{i\beta B} Z_j Z_k e^{-i\beta B} e^{-i\gamma C} |s\rangle}_{\text{supported on 3 types of subgraphs}}$$



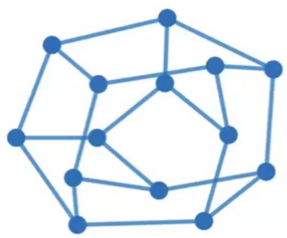
Worst case guarantee: [Farhi Goldstone Gutmann 2014]

$$\langle C \rangle / C_{\max} \geq 0.6924$$

$$C(\boldsymbol{z}_{\text{Greedy}})/C_{\max} \geq 0.5$$
$$C(\boldsymbol{z}_{\text{SDP}})/C_{\max} \geq 0.878$$

Difficult for higher p as (classical) analysis complexity grow as $2^{2^{O(p)}}$!

How about **average case**?



$$C = \sum_{\langle i,j \rangle} \frac{1}{2}(1 - z_i z_j)$$

For a large random D-regular graph, statistical theory predicts the maximum “**cut fraction**” is (w.h.p.)

$$\frac{C_{\max}}{\#\ \text{edges}} = \frac{1}{2} + \frac{V_{\max}(D)}{\sqrt{D}}$$

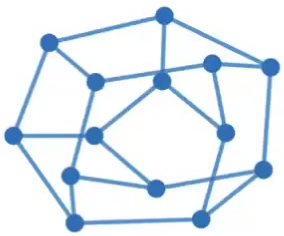
$$\lim_{D \to \infty} V_{\max}(D) = \Pi_* = 0.7631\ldots$$

(Parisi value)

[Parisi 1979] [Dembo Montanari Sen 2017]

Nobel Physics 2021

How about **average case**?



For a large random D-regular graph, statistical theory predicts the maximum "**cut fraction**" is (w.h.p.)

$$\frac{C_{\max}}{\#\ \text{edges}} = \frac{1}{2} + \frac{V_{\max}(D)}{\sqrt{D}}$$

$$C = \sum_{\langle i,j \rangle} \frac{1}{2}(1 - z_i z_j)$$

$$\lim_{D \to \infty} V_{\max}(D) = \Pi_* = 0.7631\ldots$$

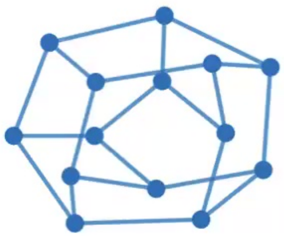
(Parisi value)

[Parisi 1979] [Dembo Montanari Sen 2017]

Nobel Physics 2021

We know the ground energy… BUT no algorithm to find the ground state bit string!

How about **average case**?



$C = \sum_{\langle i,j \rangle} \frac{1}{2}(1 - z_i z_j)$

For a large random D-regular graph, statistical theory predicts the maximum "**cut fraction**" is (w.h.p.)

$$\frac{C_{\max}}{\#\text{ edges}} = \frac{1}{2} + \frac{V_{\max}(D)}{\sqrt{D}}$$

$$\lim_{D \to \infty} V_{\max}(D) = \Pi_* = 0.7631\ldots$$

(Parisi value)

[Dembo Montanari Sen 2017]

Best *assumption-free* classical algorithms get:

$$V^{\mathrm{SDP}}(D) = 2/\pi + o_D(1)$$

$\hookrightarrow 0.6366\ldots$

[Montanari Sen 2015]
[Barak Marwaha 2021]
[Thompson Parekh Marwaha 2021]

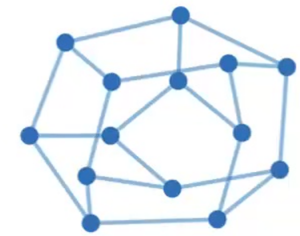
Assuming a "No OGP" conjecture, a **Local Message Passing algorithm** with ***p* rounds** gets:

$$V_p^{\mathrm{LMP}}(D) \geq \Pi_* - O_D(1/\sqrt{p})$$

$\hookrightarrow 0.7631\ldots$

[Alaoui Montanari Sellke 2021]

QAOA on MaxCut of D-regular graphs



- Let the **cut fraction** achieved by the QAOA be

$$\frac{\langle \boldsymbol{\gamma}, \boldsymbol{\beta} | C | \boldsymbol{\gamma}, \boldsymbol{\beta} \rangle}{\# \text{ edges}} = \frac{1}{2} + \frac{V_p^{\text{QAOA}}(D, \boldsymbol{\gamma}, \boldsymbol{\beta})}{\sqrt{D}}$$

and let $\bar{V}_p^{\text{QAOA}}(D) := \max_{\boldsymbol{\gamma}, \boldsymbol{\beta}} V_p^{\text{QAOA}}(D, \boldsymbol{\gamma}, \boldsymbol{\beta})$

- At $p = 1$, for triangle-free graphs, we know $\lim_{D \to \infty} \bar{V}_p^{\text{QAOA}}(D) \simeq 0.3033$

[Wang Hadfield Jiang Rieffel 2018]

- At $p = 2$, for graphs with girth > 5, we know $\lim_{D \to \infty} \bar{V}_p^{\text{QAOA}}(D) \simeq 0.4075$

[Marwaha 2021]

Note: **Large girth ≅ Random / Average case**

Performance of the QAOA on MaxCut for random or large-girth regular graphs

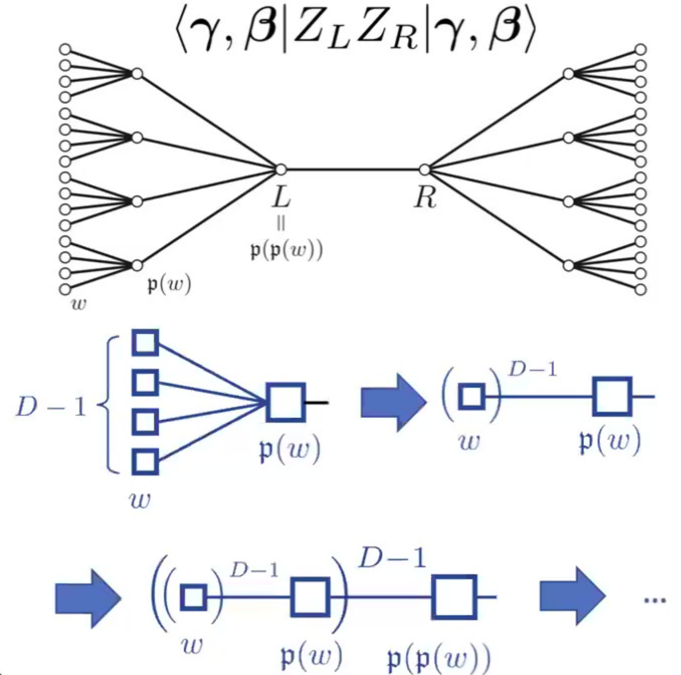
- For any D-regular graph with girth $> 2p+1$, we give an $\tilde{O}(16^p)$ time iteration using $O(4^p)$ memory to *exactly* compute

$$V_p^{\mathrm{QAOA}}(D, \boldsymbol{\gamma}, \boldsymbol{\beta})$$

- In the $D \to \infty$ limit, we give an $\tilde{O}(4^p)$ time iteration using $O(p^2)$ memory for

$$V_p^{\mathrm{QAOA}}(\boldsymbol{\gamma}, \boldsymbol{\beta}) := \lim_{D\to\infty} V_p^{\mathrm{QAOA}}(D, \boldsymbol{\gamma}, \boldsymbol{\beta})$$

Key idea: efficient analytical contraction of a tree tensor network



[Basso Farhi Marwaha Villalonga **LZ**, TQC 2022]

Performance of the QAOA on MaxCut for random or large-girth regular graphs

- For any D-regular graph with girth $> 2p+1$, we give an $\tilde{O}(16^p)$ time iteration using $O(4^p)$ memory to *exactly* compute

$$V_p^{\mathrm{QAOA}}(D, \boldsymbol{\gamma}, \boldsymbol{\beta})$$

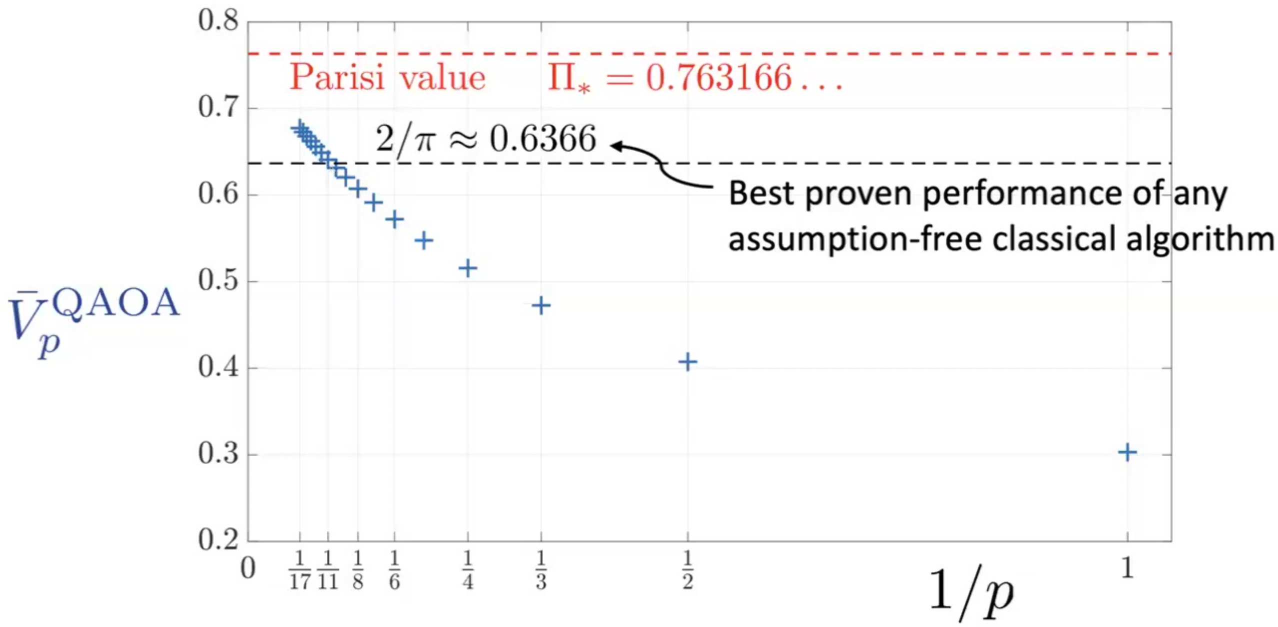
- In the $D \to \infty$ limit, we give an $\tilde{O}(4^p)$ time iteration using $O(p^2)$ memory for

$$V_p^{\mathrm{QAOA}}(\boldsymbol{\gamma}, \boldsymbol{\beta}) := \lim_{D \to \infty} V_p^{\mathrm{QAOA}}(D, \boldsymbol{\gamma}, \boldsymbol{\beta})$$

p	Best V_p^{QAOA}	
1	0.3033	optimized
2	0.4075	optimized
⋮	⋮	optimized
17	0.6773	optimized
18	0.6813	lower bounds
19	0.6848	lower bounds
20	0.6879	lower bounds

[Basso Farhi Marwaha Villalonga **LZ**, TQC 2022]

Performance of the QAOA on MaxCut for random or large-girth regular graphs



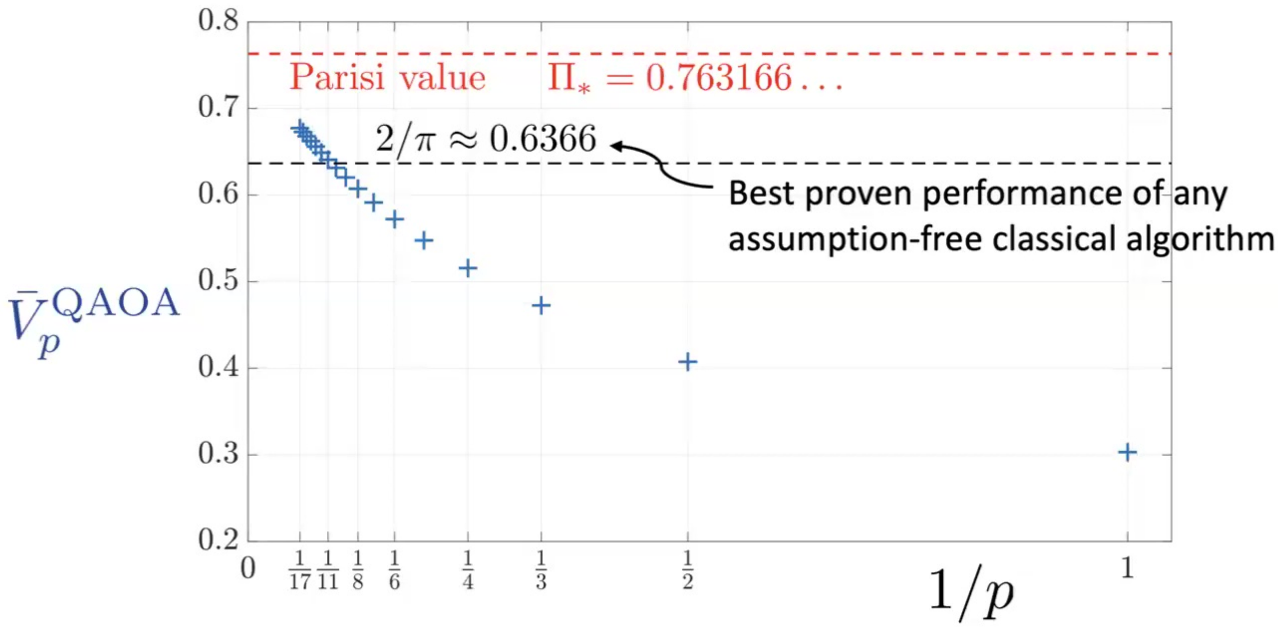
p	Best V_p^{QAOA}
1	0.3033
2	0.4075
⋮	⋮
17	0.6773
18	0.6813
19	0.6848
20	0.6879

optimized (p = 1–17)

lower bounds (p = 18–20)

[Basso Farhi Marwaha Villalonga **LZ**, TQC 2022]

Performance of the QAOA on MaxCut for random or large-girth regular graphs



p	Best V_p^{QAOA}
1	0.3033
2	0.4075
⋮	⋮
17	0.6773
18	0.6813
19	0.6848
20	0.6879

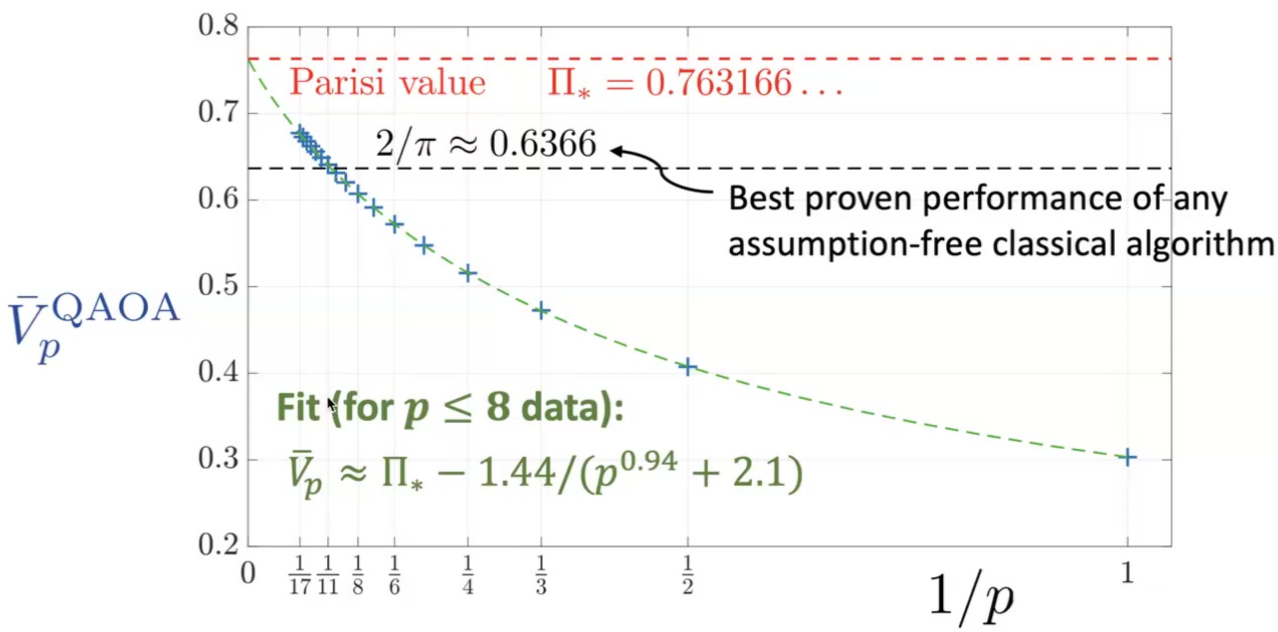
optimized (p = 1–17)

lower bounds (p = 18–20)

Classical state-of-the-art: Assuming "No OGP" conjecture, the [AMS21] Local Message Passing algorithm after p rounds gets $V_p^{\mathrm{LMP}} \geq \Pi_* - O(1/\sqrt{p})$

[Basso Farhi Marwaha Villalonga **LZ**, TQC 2022]

Performance of the QAOA on MaxCut for random or large-girth regular graphs

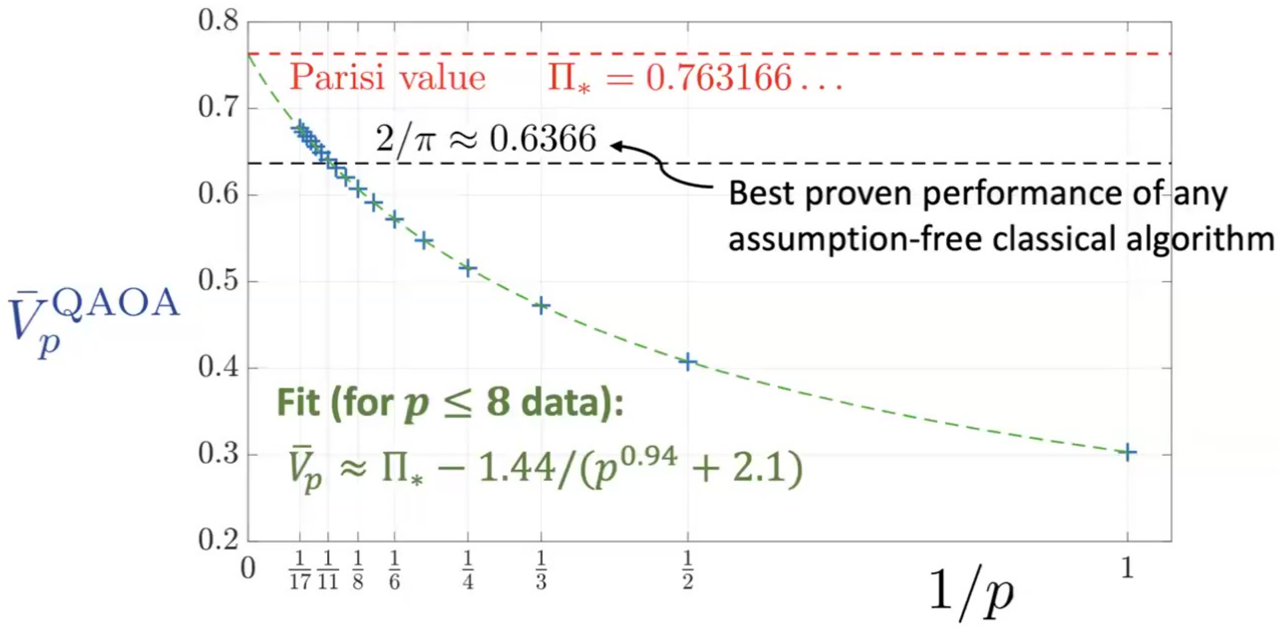


p	Best V_p^{QAOA}	
1	0.3033	optimized
2	0.4075	optimized
⋮	⋮	optimized
17	0.6773	optimized
18	0.6813	lower bounds
19	0.6848	lower bounds
20	0.6879	lower bounds

Classical state-of-the-art: Assuming "No OGP" conjecture, the [AMS21] Local Message Passing algorithm after p rounds gets $V_p^{\mathrm{LMP}} \geq \Pi_* - O(1/\sqrt{p})$

[Basso Farhi Marwaha Villalonga **LZ**, TQC 2022]

Performance of the QAOA on MaxCut for random or large-girth regular graphs



p	Best V_p^{QAOA}
1	0.3033
2	0.4075
⋮	⋮
17	0.6773
18	0.6813
19	0.6848
20	0.6879

optimized (rows 1–17); lower bounds (rows 18–20)

Quadratic Quantum Advantage

Classical state-of-the-art: Assuming "No OGP" conjecture, the [AMS21] Local Message Passing algorithm after p rounds gets $V_p^{\text{LMP}} \ge \Pi_* - O(1/\sqrt{p})$

[Basso Farhi Marwaha Villalonga **LZ**, TQC 2022]

More general problems

$$C_J(\boldsymbol{z}) = \sum_{i_1,\ldots,i_q=1}^{n} J_{i_1,i_2,\ldots,i_q} z_{i_1} z_{i_2} \cdots z_{i_q}$$

- Example: fully connected Gaussian spin glasses $J_{i_1,\ldots,i_q} \sim \mathcal{N}(0, 1/n^{q+1})$
- For J drawn from any "reasonable" distribution with mean zero, we give an ***exact*** formula for

$$\lim_{n\to\infty} \mathbb{E}_J\left[\,\langle \boldsymbol{\gamma}, \boldsymbol{\beta} | C_J | \boldsymbol{\gamma}, \boldsymbol{\beta}\rangle\,\right] = V_p^{\mathrm{QAOA}}(\boldsymbol{\gamma}, \boldsymbol{\beta})$$

- Based on proving a "**generalized multinomial theorem**", a rigorous version of a "path integral" calculation

$$\langle \hat{F} \rangle = \sum_{\text{paths } \boldsymbol{y}} \langle s | e^{i\gamma_1 C_J} | y_1 \rangle \langle y_1 | e^{i\beta_1 B} | y_2 \rangle \langle y_2 | \cdots = \sum_{\{m_j\}} \binom{n}{\{m_j\}} \prod_j Q_j^{m_j} e^{nP(\boldsymbol{m}/n)} f(\boldsymbol{m}/n)$$

$$\approx \int d\boldsymbol{\mu} e^{nS(\boldsymbol{\mu})} f(\boldsymbol{\mu}) \xrightarrow{n\to\infty} f(\boldsymbol{\mu}^*)$$

[Basso Gamarnik Mei **LZ**, FOCS 2022]

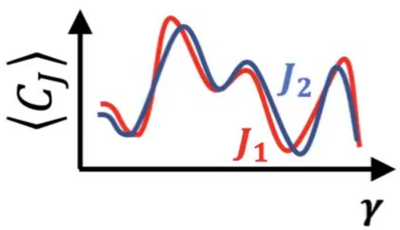
More general problems

$$C_J(\boldsymbol{z}) = \sum_{i_1,\dots,i_q=1}^{n} J_{i_1,i_2,\dots,i_q} z_{i_1} z_{i_2} \cdots z_{i_q}$$

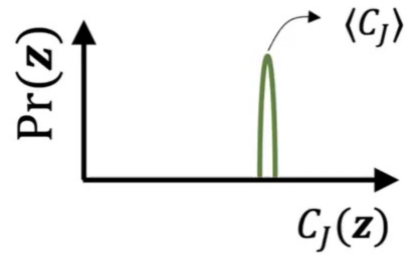
- Can also show **concentration** via second moment calculation

$$\lim_{n\to\infty} \mathbb{E}_J \left[\langle \boldsymbol{\gamma}, \boldsymbol{\beta} | C_J^2 | \boldsymbol{\gamma}, \boldsymbol{\beta} \rangle \right] = \lim_{n\to\infty} \Big(\mathbb{E}_J \left[\langle \boldsymbol{\gamma}, \boldsymbol{\beta} | C_J | \boldsymbol{\gamma}, \boldsymbol{\beta} \rangle \right] \Big)^2$$

Concentrate over instances



Concentrate over measurements



- Can reuse optimized parameters for similar instances!
- This is true even in nonlocal regime (e.g. all-to-all connected graph)!!

[Basso Gamarnik Mei **LZ**, FOCS 2022]

Summary: Minimizing classical spin glasses

- Rigorous average-case analysis of QAOA to high depths is possible
  - $2^{O(p)}$ vs previous $2^{2^{O(p)}}$ cost for worst-case analysis
- **A quantum advantage** in solving MaxCut on random graph

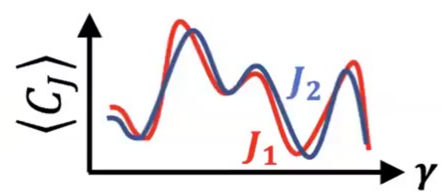
	Best Classical	QAOA
Proven Performance	$2/\pi \simeq 0.6366$	≥ 0.6879
Conjectured	$0.763166 - 1/\sqrt{p}$	$0.763166 - 1/p$

[Basso Farhi Marwaha Villalonga **LZ**, TQC 2022]

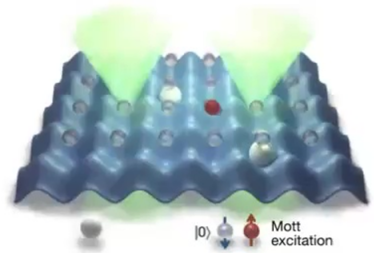
- Can analyze harder general problems with path integrals

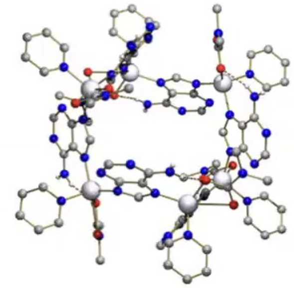
$$C_J(\boldsymbol{z}) = \sum_{i_1,\ldots,i_k=1}^{n} J_{i_1,i_2,\ldots,i_k} z_{i_1} z_{i_2} \cdots z_{i_k}$$

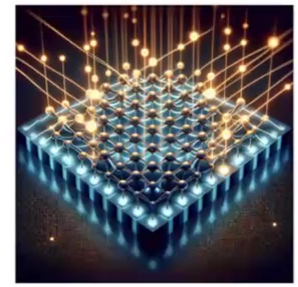
Concentrate over instances



[Basso Gamarnik Mei **LZ**, FOCS 2022]







Part 2

Minimizing Energy of *Quantum* Systems

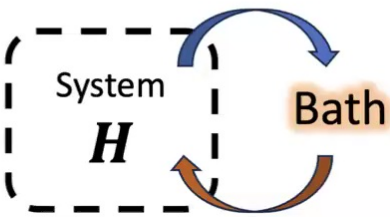
Based on [Chen Huang Preskill **LZ**, STOC 2024] *arXiv:2204.10306*

Anthony Chen

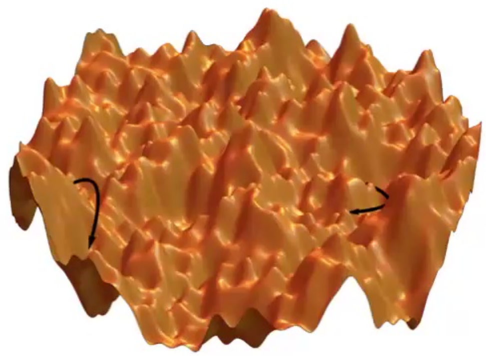
Robert Huang

John Preskill

Ground states (global minima) of quantum systems are hard to find in general, even for Nature!



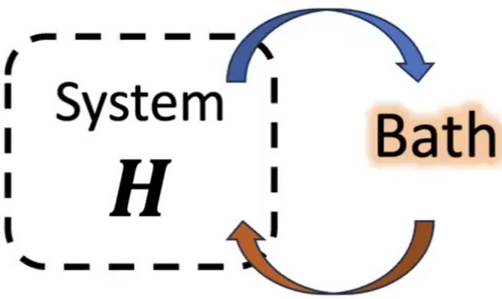
Nature cools system to a *local minimum*



e.g., spin glass

Local minima are more *physical* than global minima

How tractable is the problem of ***finding a local minimum*** *using classical and quantum computers?*



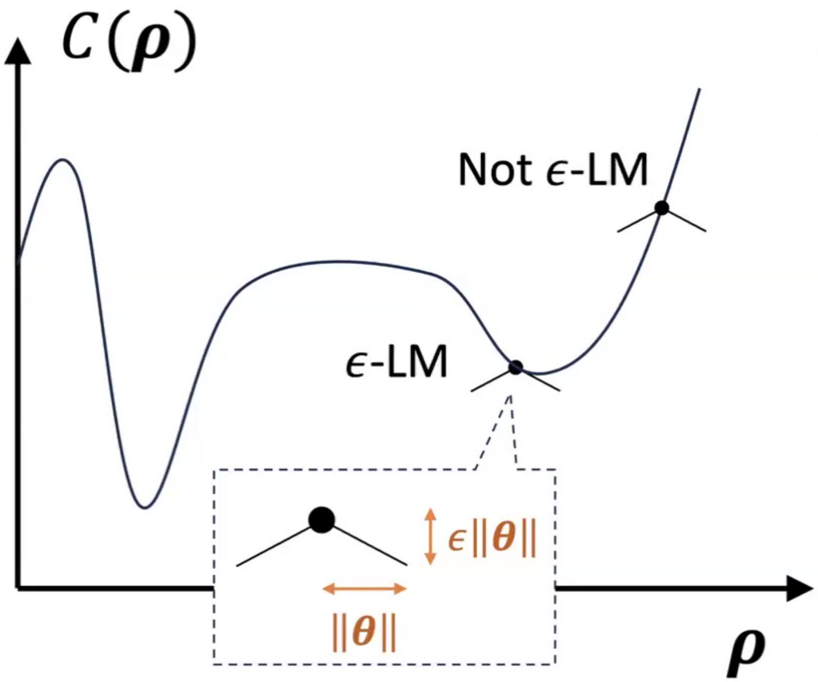
[Chen Huang Preskill **LZ,** STOC 2024]

Our Main Result

Cooling to a local minimum is universal for quantum computation.

⇒ Finding a local minimum of quantum system is **classically HARD** and **quantumly EASY**!

What is a local minimum?

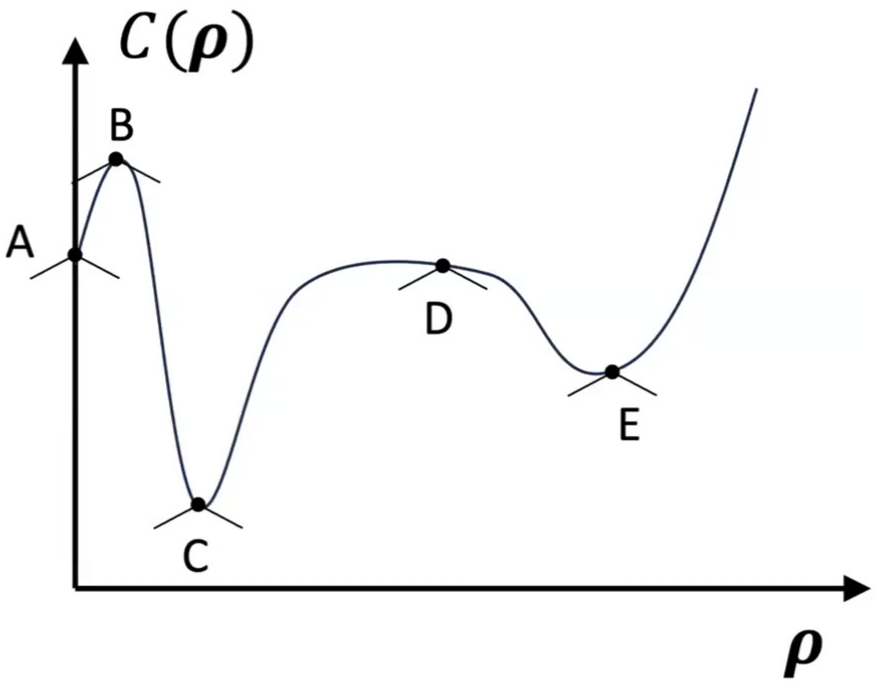


- Domain: n-qubit state $\boldsymbol{\rho}$
- Energy function: $C(\boldsymbol{\rho}) = \mathrm{tr}(\boldsymbol{H}\boldsymbol{\rho})$
- A family of perturbations: $\boldsymbol{\rho} \to \mathcal{P}_{\boldsymbol{\theta}}[\boldsymbol{\rho}]$
- $\boldsymbol{\rho}$ is an ***ϵ-approximate local minimum*** if

$$C(\boldsymbol{\rho}) \leq C(\mathcal{P}_{\boldsymbol{\theta}}[\boldsymbol{\rho}]) + \epsilon\|\boldsymbol{\theta}\|$$

for all small enough $\boldsymbol{\theta}$

What is a local minimum?



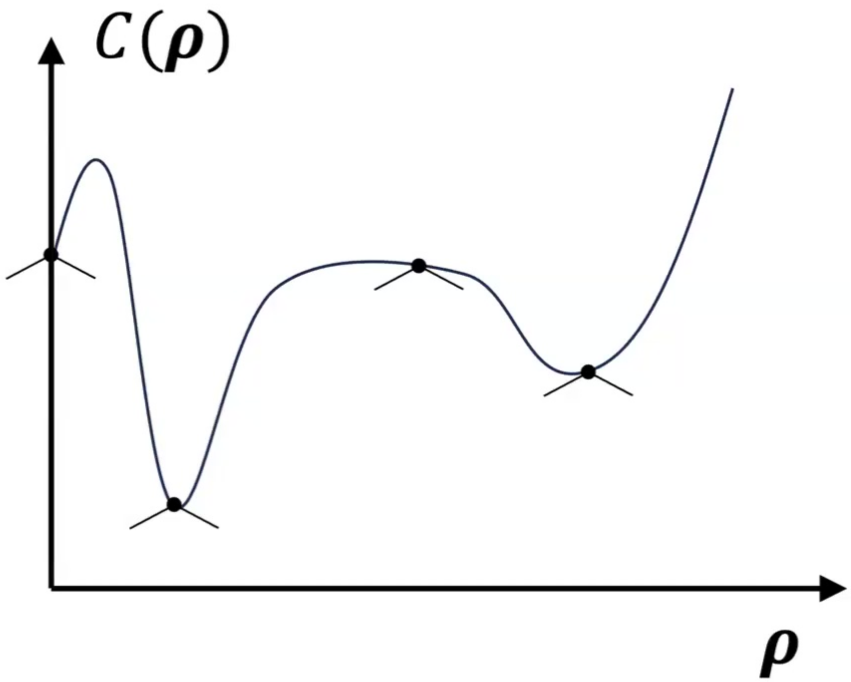
- Domain: n-qubit state $\boldsymbol{\rho}$
- Energy function: $C(\boldsymbol{\rho}) = \mathrm{tr}(\boldsymbol{H\rho})$
- A family of perturbations: $\boldsymbol{\rho} \to \mathcal{P}_{\boldsymbol{\theta}}[\boldsymbol{\rho}]$
- $\boldsymbol{\rho}$ is an ***ϵ-approximate local minimum*** if

$$C(\boldsymbol{\rho}) \leq C(\mathcal{P}_{\boldsymbol{\theta}}[\boldsymbol{\rho}]) + \epsilon \|\boldsymbol{\theta}\|$$

for all small enough $\boldsymbol{\theta}$

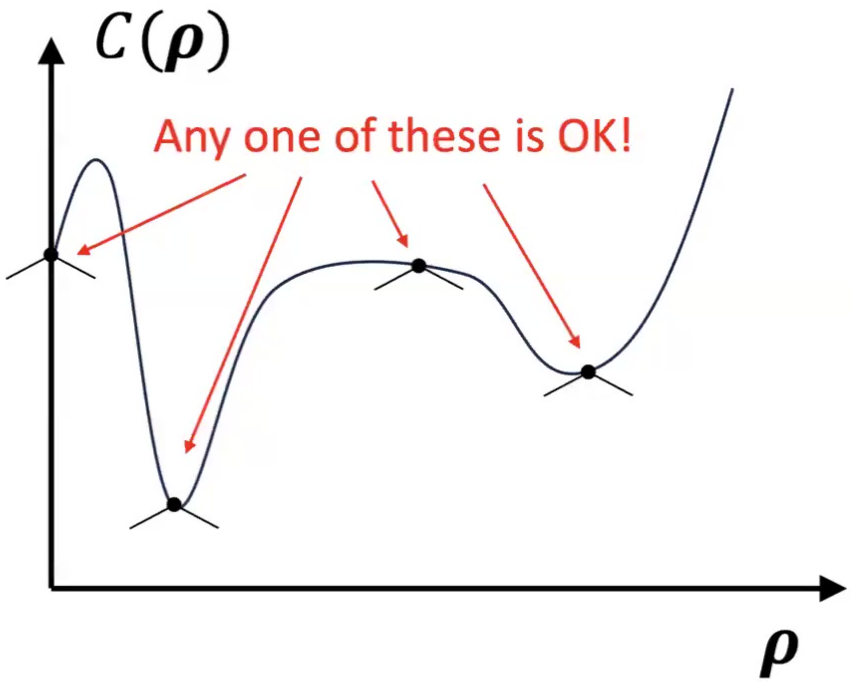
A, C, D, E are ϵ-approx LM
B is not

The problem of finding a local minimum



- **Input**:
  1. $\boldsymbol{H}$, where $\|\boldsymbol{H}\| = \mathrm{poly}(n)$
  2. a family of perturbation $\{\mathcal{P}_{\boldsymbol{\theta}}\}_{\boldsymbol{\theta}}$
  3. some $\epsilon > 1/\mathrm{poly}(n)$
  4. a (local) observable $\boldsymbol{O}$
- **Problem**: Output estimated $\mathrm{Tr}(\boldsymbol{O}\boldsymbol{\rho}^*)$ within ϵ error for *any* ϵ-approx local minimum $\boldsymbol{\rho}^*$ under the perturbations.

The problem of finding a local minimum



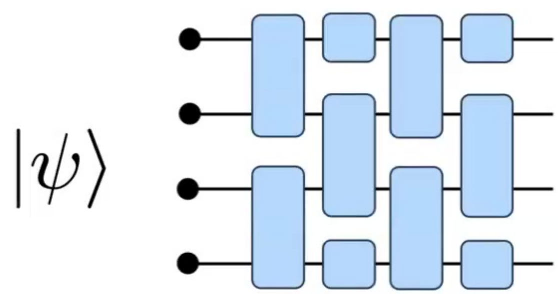
- **Input**:
  1. $\boldsymbol{H}$, where $\|\boldsymbol{H}\| = \mathrm{poly}(n)$
  2. a family of perturbation $\{\mathcal{P}_{\boldsymbol{\theta}}\}_{\boldsymbol{\theta}}$
  3. some $\epsilon > 1/\mathrm{poly}(n)$
  4. a (local) observable $\boldsymbol{O}$
- **Problem**: Output estimated $\mathrm{Tr}(\boldsymbol{O}\boldsymbol{\rho}^*)$ within ϵ error for *any* ϵ-approx local minimum $\boldsymbol{\rho}^*$ under the perturbations.

Note: purely classical input + output

Example: Local unitary perturbations

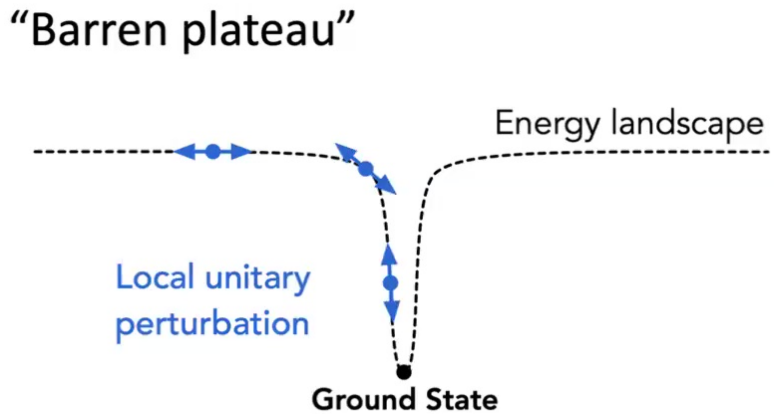
$$|\psi\rangle \to \mathcal{P}_{\boldsymbol{\theta}}[|\psi\rangle] = e^{-i\sum_a \theta_a \boldsymbol{h}_a}|\psi\rangle \qquad \theta_a \in \mathbb{R}$$

where $\{\boldsymbol{h}_a\}_a$ = {one- and two-qubit Pauli operators}.



e.g., Adapt-VQE,
QC as geometry

Finding a local minimum under local unitary perturbations is classically easy



- For any local Hamiltonian $\boldsymbol{H}$, any random state is a local minima for any $\epsilon \geq 1/\mathrm{poly}(n)$.
- There are $\exp(\exp(n))$ such local minima!
- The local minimum problem becomes **TOO EASY**! Classically just output $\mathrm{Tr}(\boldsymbol{O}\boldsymbol{\rho}^*) = \mathrm{Tr}(\boldsymbol{O})/2^n$

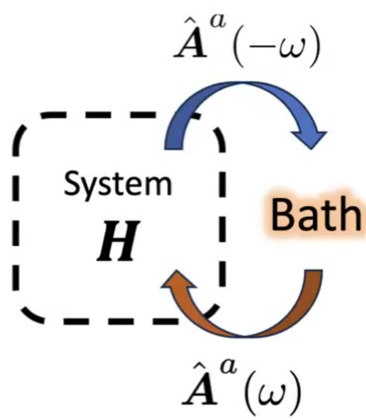
Nature-inspired Thermal perturbations

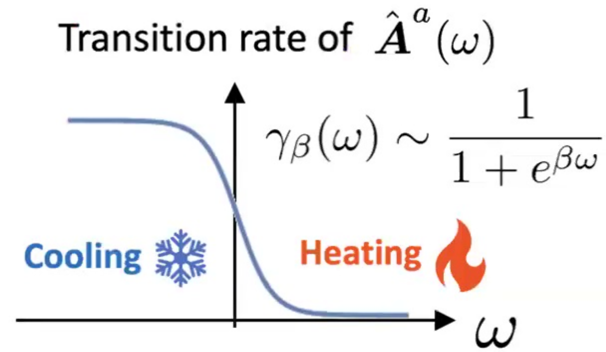
$$\boldsymbol{\rho} \to \mathcal{P}_{\boldsymbol{\theta}}[\boldsymbol{\rho}] = e^{\sum_a \theta_a \mathcal{L}_a}[\boldsymbol{\rho}] \qquad \theta_a \in \mathbb{R}_{\geq 0}$$

where $\mathcal{L}_a$ is a thermal Lindbladian for the system weakly coupled to a bath

Parameters

β	inverse temperature
τ	coarse-grain timescale
$\{\boldsymbol{A}^a\}_a$	local jump operators





Based on rigorous version of Davies equation [Mozgunov Lidar 2020]

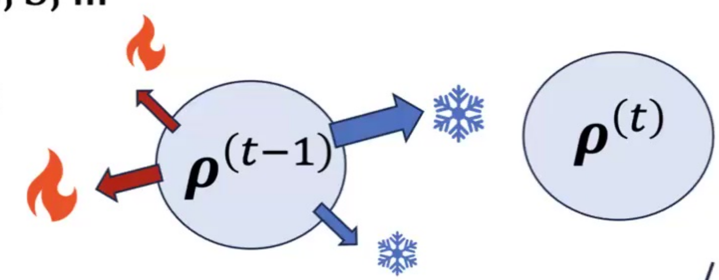
Quantum Thermal Gradient Descent can find local minima under thermal perturbations

- Consider any n-qubit Hamiltonian $\boldsymbol{H}$ where $\|\boldsymbol{H}\| \leq B$
- To find an ϵ-local minimum under thermal perturbations by $\{\mathcal{L}_a\}_{a=1}^{m}$:

Initialize at any state $\boldsymbol{\rho}^{(0)}$, and for each step $t = 1, 2, 3, \ldots$

① Estimate $g_a = \partial_a \langle \boldsymbol{H} \rangle = \mathrm{tr}\left(\boldsymbol{\rho} \mathcal{L}_a^{\dagger}[\boldsymbol{H}]\right)$ to 0.01ϵ precision

② If all $g_a > -0.99\epsilon$, STOP. Otherwise evolve $\boldsymbol{\rho}^{(t)} = \exp(\sum_a \theta_a \mathcal{L}_a)\, [\boldsymbol{\rho}^{(t-1)}]$ where $\theta_a = -\min(0, g_a)/9B^2$

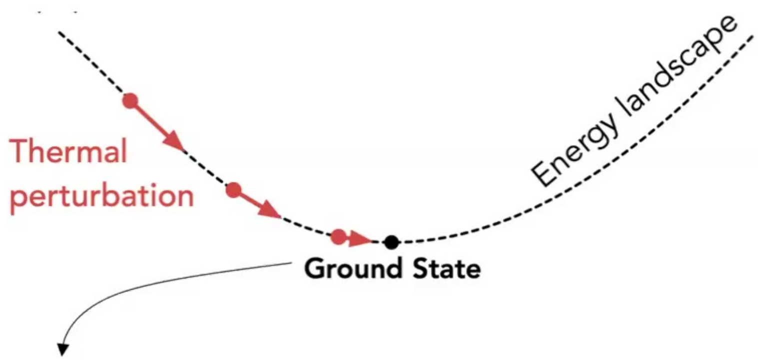


This **provably converges** within $O(B^3/\epsilon^2)$ steps!

Finding a local minimum is **quantumly easy**!

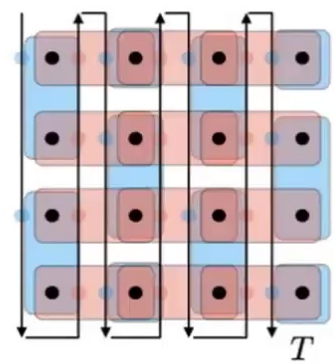
Finding a local minimum under thermal perturbations is classically hard

For any circuit $\boldsymbol{U}_C = \boldsymbol{U}_T \cdots \boldsymbol{U}_1$



$$|\mathrm{GS}\rangle = \sum_{t=0}^{T} \sqrt{\xi_t} (\boldsymbol{U}_t \cdots \boldsymbol{U}_2 \boldsymbol{U}_1 |0^n\rangle) \otimes |t\rangle$$

- **Theorem**: Certain 2D Hamiltonians whose ground states encode universal quantum computation have ***no suboptimal local minima*** when $\epsilon^{-1}, \beta, \tau \geq \mathrm{poly}(n)$.



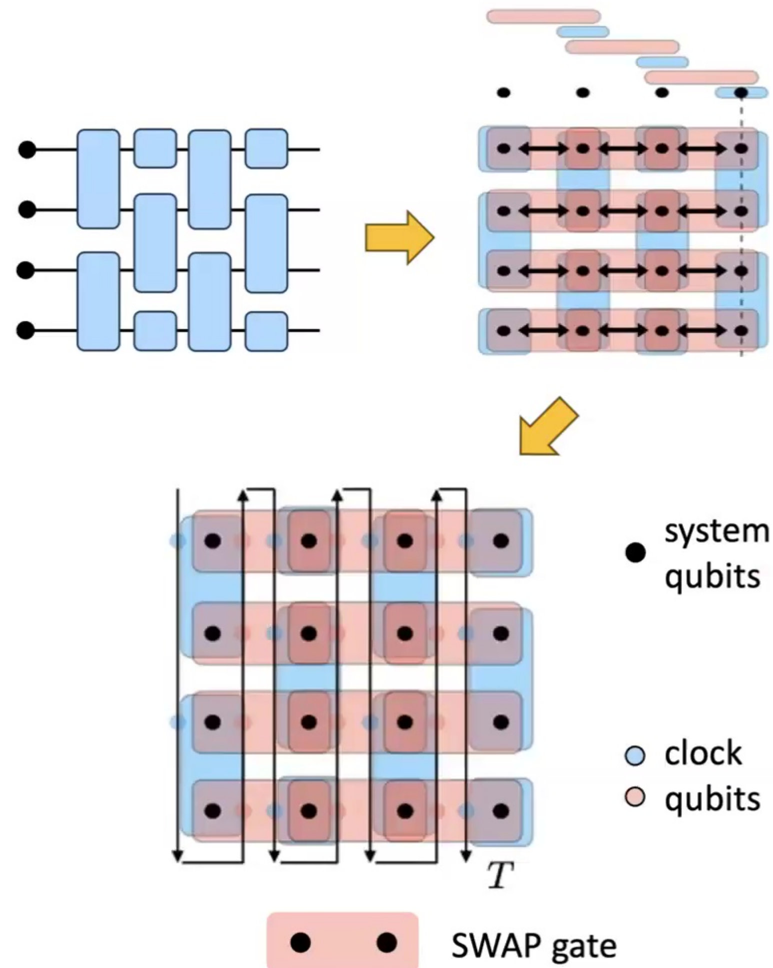
$$\boldsymbol{H}_C = \sum_t U_t \otimes |011\rangle\langle 001| + \cdots$$

BQP-hard Hamiltonian

- Write any circuit as a sparse 2D circuit $\boldsymbol{U}_C = \boldsymbol{U}_T \cdots \boldsymbol{U}_1$
- BQP-hard Hamiltonian on $n+T$ qubits:

$$\boldsymbol{H}_C = \boldsymbol{H}_{\text{clock}} + \boldsymbol{H}_{\text{in}} + \boldsymbol{H}_{\text{prop}}$$

- $\boldsymbol{H}_{\text{clock}} = \sum_t f_t \cdot \boldsymbol{I} \otimes |01\rangle\langle 01|_{t,t+1}$
  → clock qubits are in states $|t\rangle = |1^t 0^{T-t}\rangle$
- $\boldsymbol{H}_{\text{in}} = \sum_{i=1}^{n} g_i \, |1\rangle\langle 1|_i \otimes |10\rangle\langle 10|_{t_i}$ → Fix input at $|0^n\rangle$
- $\boldsymbol{H}_{\text{prop}} = -\sum_t h_t (\boldsymbol{U_t} \otimes |t+1\rangle\langle t| + \boldsymbol{U}_t^{\dagger} \otimes |t\rangle\langle t+1|)$



See e.g. [Kitaev Shen Yu 2002] [Oliveira Terhal 2008]

BQP-hard Hamiltonians have good gradients

- Goal: show $\boldsymbol{H}_C$ has no suboptimal local minima
- Negative gradient condition:

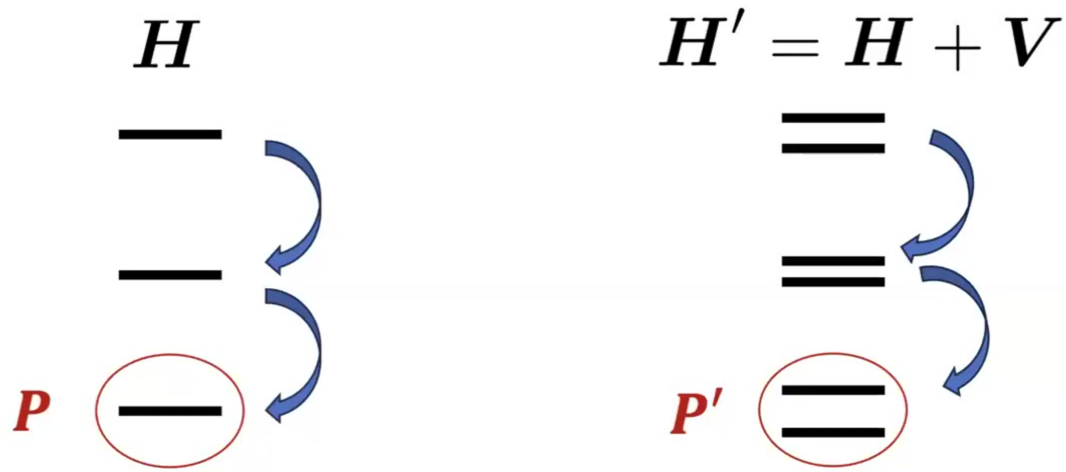
$$\mathcal{L}^\dagger[\boldsymbol{H}] \preceq -r(\boldsymbol{I} - \boldsymbol{P}_{\mathrm{GS}})$$

"any excited state must have good gradients"

$$\frac{d}{dt}\langle \boldsymbol{H} \rangle = \langle \mathcal{L}^\dagger[\boldsymbol{H}] \rangle \leq -r(1 - \langle \boldsymbol{P}_{\mathrm{GS}} \rangle)$$

- Proving this for *all* arbitrary superposition of excited states of $\boldsymbol{H}_C$ seems daunting… how?

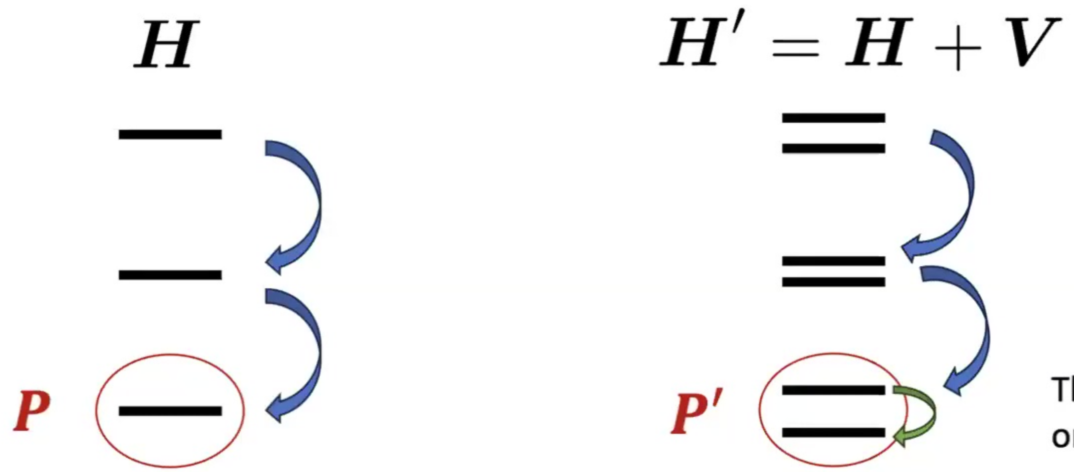
Key proof idea: Gradient of excited states robust under perturbation



Lemma

$$\mathcal{L}^\dagger[\boldsymbol{H}] \preceq -r(\boldsymbol{I} - \boldsymbol{P}) \implies \mathcal{L}'^\dagger[\boldsymbol{H}'] \preceq -r(\boldsymbol{I} - \boldsymbol{P}')$$

Key proof idea: Gradient of excited states robust under perturbation



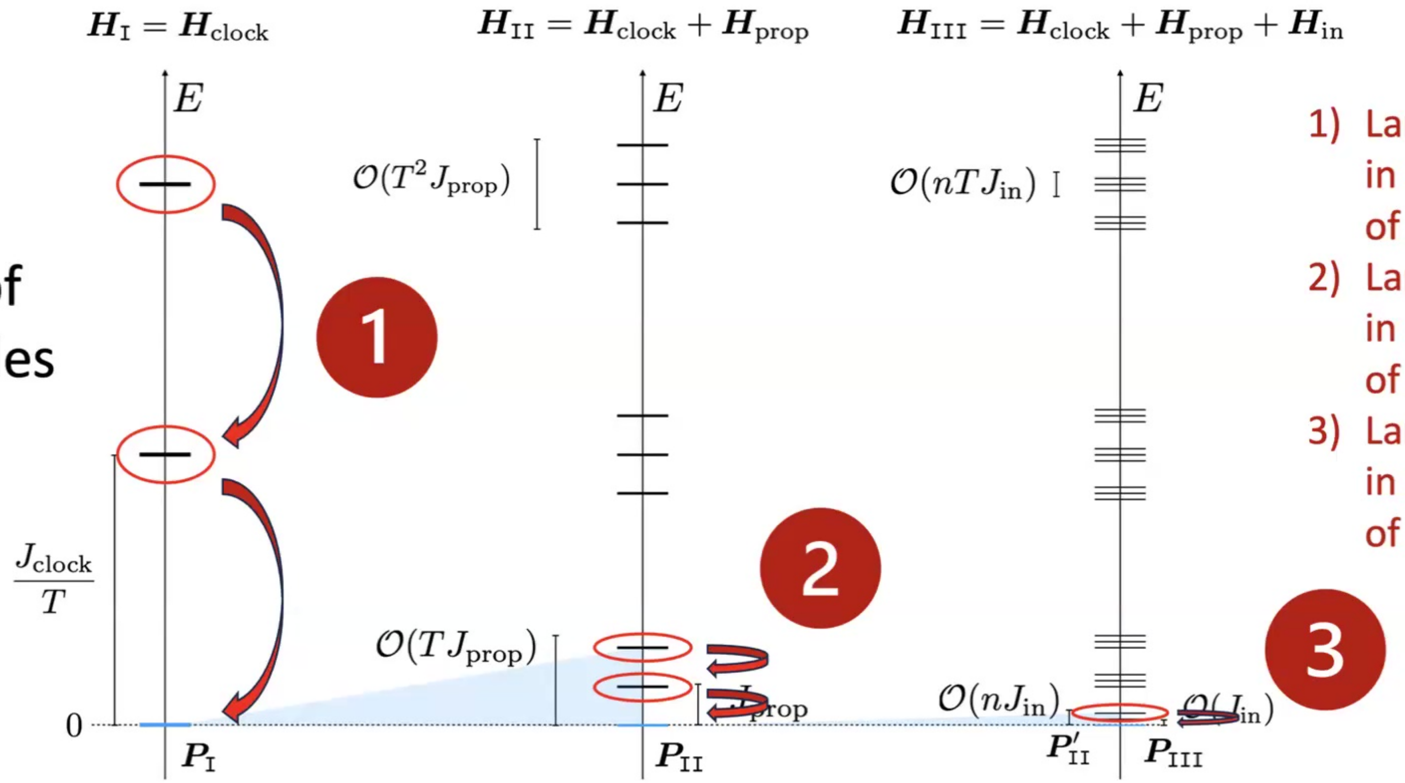
Then only need to focus on gradients inside P'

Lemma $\mathcal{L}^{\dagger}[H] \preceq -r(I - P) \implies \mathcal{L}'^{\dagger}[H'] \preceq -r(I - P')$

*Standard perturbative argument doesn't work – Errors in thermal Lindbladians are suppressed *not* by **spectral gap** ($\min |E_i - E_j|$) but by "**Bohr frequency gap**" ($\min \left| |E_i - E_j| - |E_k - E_l| \right|$)

No local minima in excited states

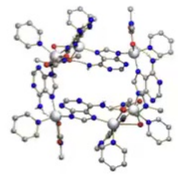
Exploit a hierarchy of energy scales



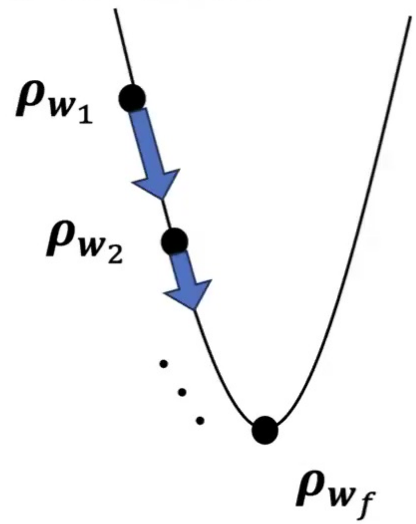
1) Large gradients in excited states of H_{I}
2) Large gradients in excited states of $P_{\mathrm{I}} H_{\mathrm{II}} P_{\mathrm{I}}$
3) Large gradients in excited states of $P_{\mathrm{II}} H_{\mathrm{III}} P_{\mathrm{II}}$

Quantum advantage in cooling to local minima

A possible method to detect quantum advantage



Energy landscape of ***classical ansatz***



Evaluate gradient

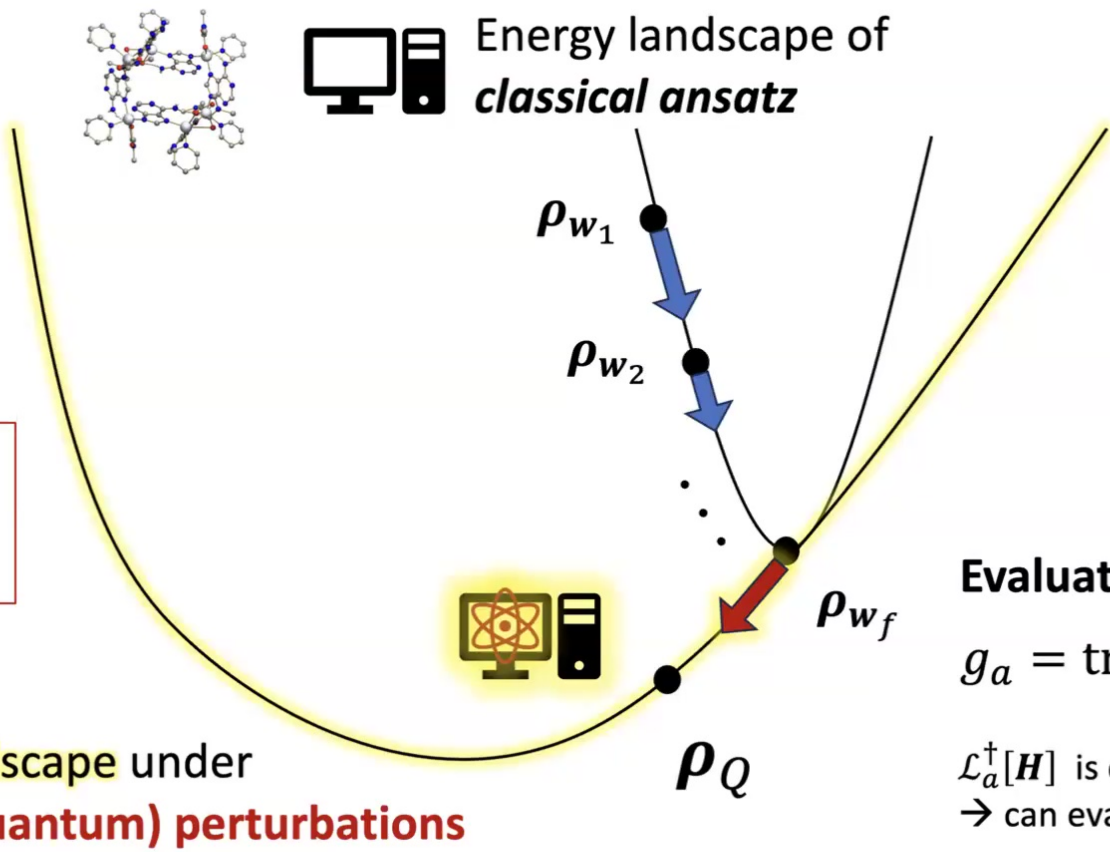
$$g_a = \mathrm{tr}\left(\mathcal{L}_a^\dagger[\boldsymbol{H}]\boldsymbol{\rho}_{\boldsymbol{w_f}}\right)$$

$\mathcal{L}_a^\dagger[\boldsymbol{H}]$ is often quasi-local
→ can evaluate classically

Quantum advantage in cooling to local minima

A possible method to detect quantum advantage

$g_a < -\epsilon$ → quantum advantage!



Evaluate gradient

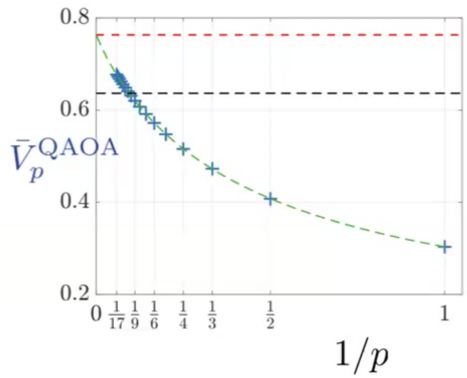
$$g_a = \mathrm{tr}\left(\mathcal{L}_a^\dagger[H]\rho_{w_f}\right)$$

$\mathcal{L}_a^\dagger[H]$ is often quasi-local
→ can evaluate classically

Summary: Quantum Advantages in Energy Minimization

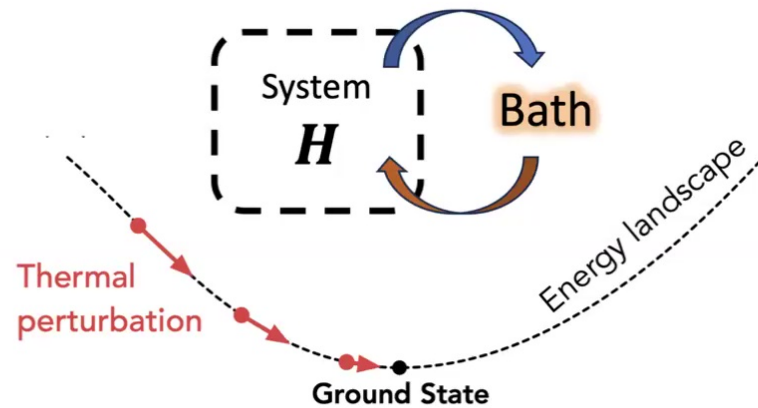
Classical systems

$$\max_{\boldsymbol{z}} \sum_{\langle ij \rangle} -z_i z_j$$



	Best Classical	QAOA
Proven	$2/\pi \simeq 0.6366$	≥ 0.6879
Conjectured	$0.763166 - 1/\sqrt{p}$	$0.763166 - 1/p$

Quantum systems



$$\langle \mathrm{GS}|\boldsymbol{Z}_j|\mathrm{GS}\rangle \approx \langle 0^n|\boldsymbol{U}_C^\dagger\, \boldsymbol{Z}_j\, \boldsymbol{U}_C|0^n\rangle$$

- No suboptimal local minima
- Finding a local minimum is ***classically HARD*** and ***quantumly EASY***!

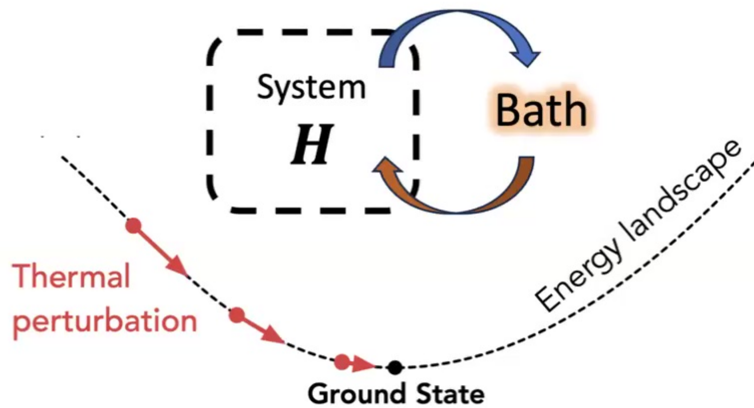
Summary: Quantum Advantages in Energy Minimization

Classical systems

Outlook

- Prove convergence to the Parisi value
- Can we analyze QAOA when $p > \log n$?
- Quantum algorithms for classical problems over random structure

Quantum systems



$$\langle \mathrm{GS} | \boldsymbol{Z}_j | \mathrm{GS} \rangle \approx \langle 0^n | \boldsymbol{U}_C^\dagger \, \boldsymbol{Z}_j \, \boldsymbol{U}_C | 0^n \rangle$$

- No suboptimal local minima
- Finding a local minimum is ***classically HARD*** and ***quantumly EASY***!

Summary: Quantum Advantages in Energy Minimization

Classical systems

Outlook

- Prove convergence to the Parisi value
- Can we analyze QAOA when $p > \log n$?
- Quantum algorithms for classical problems over random structure

Quantum systems

Outlook

- Evaluate thermal gradient of classical ansatzes
- If gradients are bad, how about mixing time?
- General energy functions
- Energy landscapes → classify quantum phases?

BQP-hard Hamiltonian

- Write any circuit as a sparse 2D circuit $\boldsymbol{U}_C = \boldsymbol{U}_T \cdots \boldsymbol{U}_1$
- BQP-hard Hamiltonian on $n + T$ qubits:

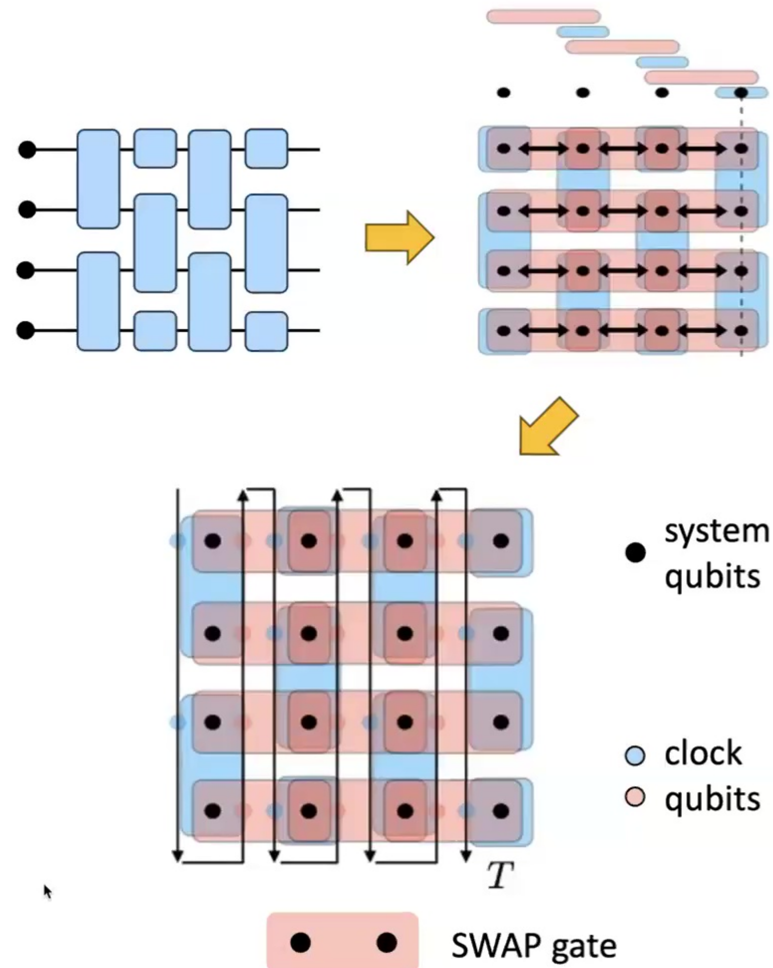
$$\boldsymbol{H}_C = \boldsymbol{H}_{\text{clock}} + \boldsymbol{H}_{\text{in}} + \boldsymbol{H}_{\text{prop}}$$

- $\boldsymbol{H}_{\text{clock}} = \sum_t f_t \cdot \boldsymbol{I} \otimes |01\rangle\langle 01|_{t,t+1}$

  → clock qubits are in states $|t\rangle = |1^t 0^{T-t}\rangle$
- $\boldsymbol{H}_{\text{in}} = \sum_{i=1}^{n} g_i \, |1\rangle\langle 1|_i \otimes |10\rangle\langle 10|_{t_i}$ → Fix input at $|0^n\rangle$
- $\boldsymbol{H}_{\text{prop}} = -\sum_t h_t(\boldsymbol{U_t} \otimes |t+1\rangle\langle t| + \boldsymbol{U}_t^\dagger \otimes |t\rangle\langle t+1|)$
- $\boldsymbol{H}_C$ has unique ground state

  $|\text{GS}\rangle = \sum_{t=0}^{T} \sqrt{\xi_t}(\boldsymbol{U}_t \cdots \boldsymbol{U}_2 \boldsymbol{U}_1 |0^n\rangle) \otimes |1^t 0^{T-t}\rangle$

See e.g. [Kitaev Shen Yu 2002] [Oliveira Terhal 2008]



BQP-hard Hamiltonian

- Write any circuit as a sparse 2D circuit $\boldsymbol{U}_C = \boldsymbol{U}_T \cdots \boldsymbol{U}_1$
- BQP-hard Hamiltonian on $n + T$ qubits:

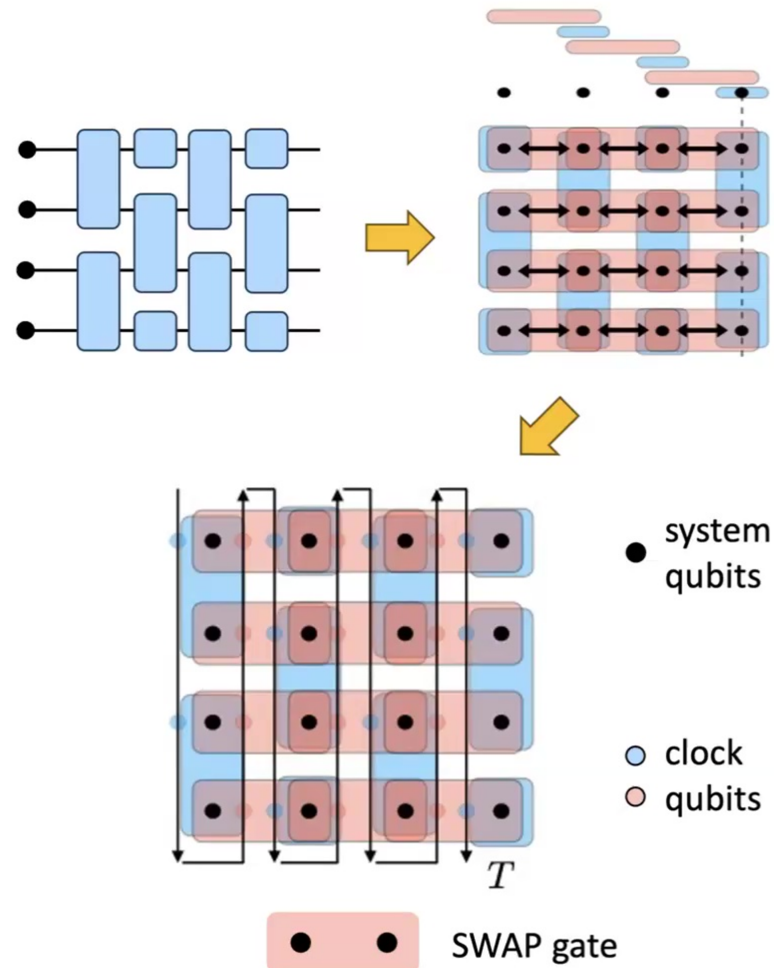
$$\boldsymbol{H}_C = \boldsymbol{H}_{\text{clock}} + \boldsymbol{H}_{\text{in}} + \boldsymbol{H}_{\text{prop}}$$

- $\boldsymbol{H}_{\text{clock}} = \sum_t f_t \cdot \boldsymbol{I} \otimes |01\rangle\langle 01|_{t,t+1}$

  → clock qubits are in states $|t\rangle = |1^t 0^{T-t}\rangle$
- $\boldsymbol{H}_{\text{in}} = \sum_{i=1}^{n} g_i\, |1\rangle\langle 1|_i \otimes |10\rangle\langle 10|_{t_i}$ → Fix input at $|0^n\rangle$
- $\boldsymbol{H}_{\text{prop}} = -\sum_t h_t(\boldsymbol{U_t} \otimes |t+1\rangle\langle t| + \boldsymbol{U}_t^\dagger \otimes |t\rangle\langle t+1|)$
- $\boldsymbol{H}_C$ has unique ground state

  $|\text{GS}\rangle = \sum_{t=0}^{T} \sqrt{\xi_t}(\boldsymbol{U}_t \cdots \boldsymbol{U}_2 \boldsymbol{U}_1 |0^n\rangle) \otimes |1^t 0^{T-t}\rangle$

See e.g. [Kitaev Shen Yu 2002] [Oliveira Terhal 2008]

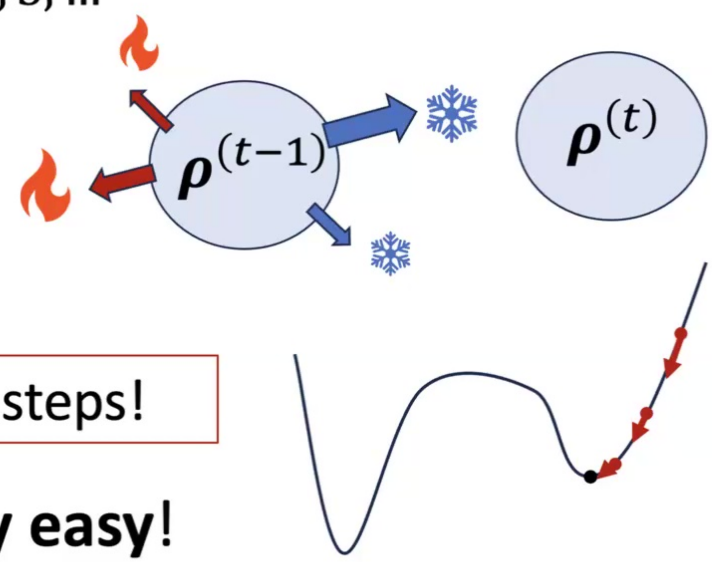


Quantum Thermal Gradient Descent can find local minima under thermal perturbations

- Consider any n-qubit Hamiltonian $\boldsymbol{H}$ where $\|\boldsymbol{H}\| \leq B$
- To find an ϵ-local minimum under thermal perturbations by $\{\mathcal{L}_a\}_{a=1}^{m}$:

Initialize at any state $\boldsymbol{\rho}^{(0)}$, and for each step $t = 1, 2, 3, \ldots$

1. Estimate $g_a = \partial_a \langle \boldsymbol{H} \rangle = \mathrm{tr}\big(\boldsymbol{\rho} \mathcal{L}_a^\dagger[\boldsymbol{H}]\big)$ to 0.01ϵ precision
2. If all $g_a > -0.99\epsilon$, STOP. Otherwise evolve $\boldsymbol{\rho}^{(t)} = \exp(\sum_a \theta_a \mathcal{L}_a)\,[\boldsymbol{\rho}^{(t-1)}]$ where $\theta_a = -\min(0, g_a)/9B^2$



This **provably converges** within $O(B^3/\epsilon^2)$ steps!

Finding a local minimum is **quantumly easy**!

Summary: Quantum Advantages in Energy Minimization

Classical systems

Outlook

- Prove convergence to the Parisi value
- Can we analyze QAOA when $p > \log n$?
- Quantum algorithms for classical problems over random structure

Quantum systems

Outlook

- Evaluate thermal gradient of classical ansatzes
- If gradients are bad, how about mixing time?
- General energy functions
- Energy landscapes → classify quantum phases?