The Quantum Approximate Optimization Algorithm and the Sherrington-Kirkpatrick Model at Infinite Size

Leo Zhou (Harvard University) with Edward Farhi, Jeffrey Goldstone, and Sam Gutmann
QIP – Feb 1, 2021
Combinatorial Optimization Problems

Cost function

\[ C(z) = \sum_{\alpha} C_{\alpha}(z) \]

\[ z = (z_1, \ldots, z_n) \in \{\pm 1\}^n \]

Want \( z^* \) so \( C(z^*) \) is maximized
Combinatorial Optimization Problems

**Goal**: find a **bipartition** of vertices that cut the maximum \# edges

**Cost function**

\[ C(z) = \sum_{\alpha} C_\alpha(z) \]

\[ z = (z_1, \ldots, z_n) \in \{\pm 1\}^n \]

Want \( z^* \) so \( C(z^*) \) is maximized

**MaxCut**

![Graph](image)
Combinatorial Optimization Problems

Cost function

\[ C(z) = \sum_\alpha C_\alpha(z) \]

\[ z = (z_1, \ldots, z_n) \in \{\pm 1\}^n \]

Want \( z^* \) so \( C(z^*) \) is maximized

**MaxCut**

**Goal**: find a **bipartition** of vertices that cut the maximum # edges
Combinatorial Optimization Problems

Cost function

\[ C(z) = \sum_{\alpha} C_\alpha(z) \]

\[ z = (z_1, \ldots, z_n) \in \{\pm 1\}^n \]

Want \( z^* \) so \( C(z^*) \) is maximized

**MaxCut**

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

**Goal:** find a **bipartition** of vertices that cut the maximum \# edges
Combinatorial Optimization Problems

**Cost function**

\[ C(z) = \sum_{\alpha} C_\alpha(z) \]

**Goal**: find a bipartition of vertices that cut the maximum # edges

\[ z = (z_1, \ldots, z_n) \in \{\pm 1\}^n \]

Want \( z^* \) so \( C(z^*) \) is maximized

**MaxCut**

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

**Sherrington-Kirkpatrick (SK) model**

Goal: find a **bipartition** of vertices that cut the maximum # edges
Combinatorial Optimization Problems

Cost function
\[ C(z) = \sum_{\alpha} C_{\alpha}(z) \]

\( z = (z_1, \ldots, z_n) \in \{\pm 1\}^n \)
Want \( z^* \) so \( C(z^*) \) is maximized

**MaxCut**
\[ C = \sum_{i,j} \frac{1}{2} (1 - z_i z_j) \]

Goal: find a **bipartition** of vertices that cut the maximum # edges

**Sherrington-Kirkpatrick (SK) model**
\[ C = \frac{1}{\sqrt{n}} \sum_{i<j} J_{ij} z_i z_j \]

\( J_{ij} \sim \text{Normal}(0, 1) \)
or \( J_{ij} \in \{\pm 1\} \)
Combinatorial Optimization Problems

**Goal**: find a **bipartition** of vertices that cut the maximum number of edges

Cost function

\[ C(z) = \sum_{\alpha} C_{\alpha}(z) \]

\[ z = (z_1, \ldots, z_n) \in \{\pm 1\}^n \]

Want \( z^* \) so \( C(z^*) \) is maximized

**MaxCut**

\[ C = \sum_{i,j} \frac{1}{2}(1 - z_i z_j) \]

**Sherrington-Kirkpatrick (SK) model**

\[ C = \frac{1}{\sqrt{n}} \sum_{i<j} J_{ij} z_i z_j \]

\( J_{ij} \sim \text{Normal}(0, 1) \)

or \( J_{ij} \in \{\pm 1\} \)

**MaxCut on Erdős-Rényi graphs = SK** (average case)
Quantum Approximate Optimization Algorithm (QAOA)

[Farhi Goldstone Gutmann 2014]
Quantum Approximate Optimization Algorithm (QAOA)

\[ |+\rangle \]
\[ |+\rangle \]
\[ \vdots \]
\[ |+\rangle \]

\[ |+\rangle \otimes n \]

[Farhi Goldstone Gutmann 2014]
Quantum Approximate Optimization Algorithm (QAOA)

\[ e^{-i\beta_1 B} e^{-i\gamma_1 C} |+\rangle \otimes n \]

\[ B = \sum_{i=1}^{n} X_i \]

[Farhi Goldstone Gutmann 2014]
Quantum \textbf{Approximate} Optimization Algorithm (QAOA)

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

\[ B = \sum_{i=1}^{n} X_i \]

[Farhi Goldstone Gutmann 2014]
Quantum Approximate Optimization Algorithm (QAOA)

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

\[
B = \sum_{i=1}^{n} X_i
\]

[Farhi Goldstone Gutmann 2014]
Quantum Approximate Optimization Algorithm (QAOA)

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

\[ B = \sum_{i=1}^{n} X_i \]

[Farhi Goldstone Gutmann 2014]
Quantum Approximate Optimization Algorithm (QAOA)

Parameters: \( \{\gamma, \beta\} = \{(\gamma_1, \ldots, \gamma_p), (\beta_1, \ldots, \beta_p)\} \)

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

\[
B = \sum_{i=1}^{n} X_i
\]

[Farhi Goldstone Gutmann 2014]
Quantum Approximate Optimization Algorithm (QAOA)

Parameters: $\{\gamma, \beta\} = \{(\gamma_1, \ldots, \gamma_p), (\beta_1, \ldots, \beta_p)\}$

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

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

As $p \to \infty$ QAOA can get the global optimum
Previous Results on the QAOA

• Analyze performance via “subgraphs”
  
  e.g. MaxCut on 3-regular graphs

\[ C = \sum_{\langle j,k \rangle} \frac{1}{2} (1 - Z_j Z_k) \]
Previous Results on the QAOA

- Analyze performance via “subgraphs”
  
  *e.g. MaxCut on 3-regular graphs*

\[
p = 1
\]

\[
C = \sum_{\langle j, k \rangle} \frac{1}{2} (1 - Z_j Z_k)
\]
Previous Results on the QAOA

- Analyze performance via “subgraphs”
  
  \( p = 1 \)

  \( |s\rangle = |+\rangle^\otimes n \)

  \( \langle s | e^{i\gamma C} e^{i\beta B} Z_j Z_k e^{-i\beta B} e^{-i\gamma C} |s\rangle \)

  e.g. MaxCut on 3-regular graphs
Previous Results on the QAOA

- Analyze performance via “subgraphs”

  e.g. MaxCut on 3-regular graphs

\[ p = 1 \quad |s\rangle = |+\rangle^\otimes n \]

\[ \langle s | e^{i\gamma C} e^{i\beta B} Z_j Z_k e^{-i\beta B} e^{-i\gamma C} |s\rangle \]

supported on 3 types of subgraphs
Previous Results on the QAOA

• Analyze performance via “subgraphs”

  
  e.g. MaxCut on 3-regular graphs

\[
p = 1 \quad |s\rangle = |+\rangle^\otimes n
\]

\[
\langle s | e^{i\gamma C} e^{i\beta B} Z_j Z_k e^{-i\beta B} e^{-i\gamma C} |s\rangle
\]

supported on 3 types of subgraphs
Previous Results on the QAOA

- Analyze performance via “subgraphs”

  *e.g. MaxCut on 3-regular graphs*

\[
p = 1 \quad |s\rangle = |+\rangle^\otimes n
\]

\[
|s\rangle = e^{i\gamma C} e^{i\beta B} Z_j Z_k e^{-i\beta B} e^{-i\gamma C} |s\rangle
\]

supported on 3 types of subgraphs
Previous Results on the QAOA

• Analyze performance via “subgraphs”

\textit{e.g. MaxCut on 3-regular graphs}

\[ |s\rangle = |+\rangle^\otimes n \]

\[ \langle s | e^{i\gamma C} e^{i\beta B} Z_j Z_k e^{-i\beta B} e^{-i\gamma C} |s\rangle \]

supported on 3 types of subgraphs

Worst case guarantee:

\[ \langle C \rangle / C_{\text{max}} \geq 0.6924 \ @ \ p = 1 \]

[Farhi Goldstone Gutmann 2014]
Previous Results on the QAOA

• Analyze performance via “subgraphs”

\[ p = 1 \quad \quad |s\rangle = |+\rangle^\otimes n \]

\[ \langle s | e^{i\gamma C} e^{i\beta B} Z_j Z_k e^{-i\beta B} e^{-i\gamma C} |s\rangle \]

-supported on 3 types of subgraphs

\[ \langle C \rangle / C_{\text{max}} \geq 0.6924 \quad @ \quad p = 1 \]

[Farhi Goldstone Gutmann 2014]

Difficult for higher \( p \) as the complexity of classical simulation grow as \( O(2^{2^p})! \)
Previous Results on the QAOA

• Analyze performance via “subgraphs”
  
  e.g. MaxCut on 3-regular graphs

  “Landscape-Independence”

  \[ F_G(\gamma, \beta) = \langle \gamma, \beta | G C G | \gamma, \beta \rangle_G \]

[ LZ et al. 2018 ]
[ Brandão et al. 2018 ]
Previous Results on the QAOA

• Analyze performance via “subgraphs”
  e.g. MaxCut on 3-regular graphs

“Landscape-Independence”

\[ F_G(\gamma, \beta) = \langle \gamma, \beta | C_G | \gamma, \beta \rangle_G \]

\[ F(\gamma, \beta) \times \left[ 1 + O \left( \frac{1}{\sqrt{n}} \right) \right] \]

[Brandão et al. 2018]

[Let al. 2018]
Previous Results on the QAOA

• Analyze performance via “subgraphs”
  
  *e.g. MaxCut on 3-regular graphs*

  \[
  F_G(\gamma, \beta) = \langle \gamma, \beta \rvert_G C_G \rvert \gamma, \beta \rangle_G
  \]

  \[
  F(\gamma, \beta) \times \left[ 1 + O \left( \frac{1}{\sqrt{n}} \right) \right]
  \]

  \[\text{[LZ et al. 2018] [Brandão et al. 2018]}\]

• Low-depth QAOA don’t see the whole graph $\rightarrow$ limited performance

  \[\text{[Bravyi Kliesch Koenig Tang 2019]}\]
  \[\text{[Farhi Gamarnik Gutmann 2020]}\]
Previous Results on the QAOA

• Analyze performance via “subgraphs”
  
  *e.g. MaxCut on 3-regular graphs*

  \[ F_G(\gamma, \beta) = \langle \gamma, \beta | G | \gamma, \beta \rangle_G \]

  \[ n \to \infty \rightarrow F(\gamma, \beta) \times \left[ 1 + O \left( \frac{1}{\sqrt{n}} \right) \right] \]

  \[ \mathcal{L}_{\text{Landscape-Independence}} \]

  \[ \frac{\partial^2 F}{\partial \gamma \partial \beta} \]

  \[ \mathcal{L}_{\text{Landscape-Independence}} \]

• Low-depth QAOA don’t see the whole graph \( \to \) limited performance

On \( d \)-regular graphs, mostly see trees when \( p < \log_{d-1} n \)

[Bravyi Kliesch Koenig Tang 2019]
[Farhi Gamarnik Gutmann 2020]
Previous Results on the QAOA

• Analyze performance via “subgraphs”
  e.g. MaxCut on 3-regular graphs

“Landscape-Independence”

\[
F_G(\gamma, \beta) = \langle \gamma, \beta | G | \gamma, \beta \rangle_G
\]

\[
\lim_{n \to \infty} F(\gamma, \beta) \times \left[ 1 + O\left(\frac{1}{\sqrt{n}}\right)\right]
\]

• Low-depth QAOA don’t see the whole graph \(\rightarrow\) limited performance

On \(d\)-regular graphs, mostly see trees when \(p \ll \log_{d-1} n\)

[Bravyi Kliesch Koenig Tang 2019]
[Farhi Gamarnik Gutmann 2020]

Cannot distinguish bipartite vs. typical (frustrated) graphs

[Brandão et al. 2018]
[Bravyi Kliesch Koenig Tang 2019]
[Farhi Gamarnik Gutmann 2020]
The Sherrington-Kirkpatrick model

\[ C_J = \frac{1}{\sqrt{n}} \sum_{i<j} J_{ij} Z_i Z_j \quad J_{ij} \sim \text{Normal}(0, 1) \]
The Sherrington-Kirkpatrick model

\[ C_J = \frac{1}{\sqrt{n}} \sum_{i < j} J_{ij} Z_i Z_j \quad J_{ij} \sim \text{Normal}(0, 1) \]

- Unbounded vertex degree \( \Rightarrow \) QAOA sees the whole graph at \( p = 2 \)
The Sherrington-Kirkpatrick model

\[ C_J = \frac{1}{\sqrt{n}} \sum_{i<j} J_{ij} Z_i Z_j \quad J_{ij} \sim \text{Normal}(0, 1) \]

- Unbounded vertex degree \( \rightarrow \) QAOA sees the whole graph at \( p = 2 \)

- **Worst case**: NP-hard to approximate within \( O(1/\log^c(n)) \) factor [Arora et al. 2005]
The Sherrington-Kirkpatrick model

\[ C_J = \frac{1}{\sqrt{n}} \sum_{i<j} J_{ij} Z_i Z_j \quad J_{ij} \sim \text{Normal}(0, 1) \]

- Unbounded vertex degree \( \Rightarrow \) QAOA sees the whole graph at \( p = 2 \)

- **Worst case:** NP-hard to approximate within \( O(1/\log^c(n)) \) factor [Arora et al. 2005]

- **Typical case:** Famously, Parisi (1979) predicted and Talagrand (2006) proved that

\[
\lim_{n \to \infty} \frac{1}{n} \max_z C_J(z) = \Pi_* = 0.763166…
\]
Complexity of solving a **typical** SK instance?

• Parisi *et al.*’s result does not *construct* the solution!

\[ \Pi_* = 0.763166 \ldots \]
Complexity of solving a typical SK instance?

• Parisi et al.’s result does not construct the solution!

• Known results of typical-case complexity:
  1. Simulated Annealing is believed to fail for this problem [Parisi]
  2. Semi-Definite Programming obtains $C/n = 2/\pi \approx 0.6366$ [Montanari Sen 2016]
  3. Assuming the conjecture that the SK model has no "overlap gap property" (OGP), Andrea Montanari’s algorithm (2018) outputs $\hat{z}$ with

$$C/n \geq (1 - \epsilon) \Pi_*$$

in time $O(n^2 / \epsilon^k)$
Main Result 1: Performance of the QAOA applied to the SK model

We give an $O(16^p)$-time method to evaluate

$$V_p(\gamma, \beta) = \lim_{n \to \infty} \frac{1}{n} \mathbb{E}_J[\langle \gamma, \beta | C | \gamma, \beta \rangle]$$

Much better than $O(2^{2^p})$-time subgraph method
Main Result 1: Performance of the QAOA applied to the SK model

We give an $O(16^p)$-time method to evaluate

$$V_p(\gamma, \beta) = \lim_{n \to \infty} \frac{1}{n} \mathbb{E}_J[\langle\gamma, \beta|C|\gamma, \beta\rangle]$$

Much better than $O(2^{2^p})$-time subgraph method

$$\bar{V}_p = \max_{\gamma, \beta} V_p(\gamma, \beta)$$
Main Result 1:
Performance of the QAOA applied to the SK model

We give an $O(16^p)$-time method to evaluate

$$V_p(\gamma, \beta) = \lim_{n \to \infty} \frac{1}{n} \mathbb{E}_J [\langle \gamma, \beta | C | \gamma, \beta \rangle]$$

Much better than $O(2^{2^p})$-time subgraph method

$$\bar{V}_p = \max_{\gamma, \beta} V_p(\gamma, \beta)$$

QAOA beats SDP @ $p=11$
Main Result 2: Concentration of QAOA on the SK model

• We also prove, for any fixed depth $p$: 

$$\lim_{n \to \infty} \mathbb{E}_J[\langle (C/n)^2 \rangle] = \lim_{n \to \infty} \mathbb{E}_J^2[\langle C/n \rangle]$$
Main Result 2: Concentration of QAOA on the SK model

- We also prove, for any fixed depth $p$:

$$\lim_{n \to \infty} \mathbb{E}_J[\langle (C/n)^2 \rangle] = \lim_{n \to \infty} \mathbb{E}^2_J[\langle C/n \rangle]$$

Concentration over instances ("Landscape-Independence")

\[
\langle C/n \rangle
\]

\[
\gamma
\]

$J_1$ $J_2$
Main Result 2: Concentration of QAOA on the SK model

- We also prove, for any fixed depth $p$:

$$\lim_{n \to \infty} \mathbb{E}_J[\langle (C/n)^2 \rangle] = \lim_{n \to \infty} \mathbb{E}_J^2[\langle C/n \rangle]$$

**Concentration over instances**
(“Landscape-Independence”)

Concentration over measurements

\[
\langle C/n \rangle
\]
Main Result 2: Concentration of QAOA on the SK model

- We also prove, for any fixed depth $p$:

  $$\lim_{n \to \infty} \mathbb{E}_J[\langle (C/n)^2 \rangle] = \lim_{n \to \infty} \mathbb{E}_J^2[\langle C/n \rangle]$$

Concentration over instances
("Landscape-Independence")

Concentration over measurements

- With probability $\to 1$ as $n \to \infty$, applying QAOA and measuring will give us a bit string $z$ which has

  $$C(z)/n \approx \langle C/n \rangle \approx V_p$$
Key Idea: **Average over instances**

- Parisi’s formalism requires delicate tricks
  - A replica-symmetry-breaking ansatz for the free energy:

$$\mathbb{E}_J[\log Z_J] \quad Z_J(T) = \text{tr}(e^{C_J/T})$$
Key Idea: **Average over instances**

- Parisi’s formalism requires delicate tricks
  - A replica-symmetry-breaking ansatz for the free energy:

\[
\mathbb{E}_J[\log Z_J] \neq \log \mathbb{E}_J[Z_J] \quad \quad \quad Z_J(T) = \text{tr}(e^{C_J/T})
\]
Key Idea: **Average over instances**

- Parisi’s formalism requires delicate tricks
  - A replica-symmetry-breaking ansatz for the free energy:

\[
\mathbb{E}_J[\log Z_J] = \lim_{k \to 0} \frac{1}{k} \log \mathbb{E}[Z_J^k]
\]

\[
Z_J(T) = \text{tr}(e^{C_J/T})
\]

- \(k\) replicas of \(J\)
Key Idea: **Average over instances**

- Parisi’s formalism requires delicate tricks
  - A replica-symmetry-breaking ansatz for the free energy:
    \[
    \mathbb{E}_J[\log Z_J] = \lim_{k \to 0} \frac{1}{k} \log \mathbb{E}[Z^{k}_J] \quad Z_J(T) = \text{tr}(e^{C_J}/T)
    \]
  - For QAOA, averaging over $J$ is easier
    \[
    \frac{1}{n} \mathbb{E}_J[\langle C \rangle] = \mathbb{E}_J[\langle s | e^{i\gamma C} e^{i\beta B} \frac{C}{n} e^{-i\beta B} e^{-i\gamma C} | s \rangle]
    \]
Key Idea: **Average over instances**

- Parisi’s formalism requires delicate tricks
  - A replica-symmetry-breaking ansatz for the free energy:
    \[
    \mathbb{E}_J[\log Z_J] = \lim_{k \to 0} \frac{1}{k} \log \mathbb{E}[Z_J^k] \quad Z_J(T) = \text{tr}(e^{C_J/T})
    \]

- For QAOA, averaging over \( J \) is easier
  \[
  \frac{1}{n} \mathbb{E}_J[\langle C \rangle] = \mathbb{E}_J\left[ \langle s | e^{i\gamma C} e^{i\beta B} \frac{C}{n} e^{-i\beta B} e^{-i\gamma C} | s \rangle \right]
  \]
  \[
  = \frac{1}{2^n} \sum_{z^1, z^m, z^2} \mathbb{E}_J\left[ e^{i\gamma C(z^1)} \langle z^1 | e^{i\beta B} | z^m \rangle \frac{C(z^m)}{n} \langle z^m | e^{-i\beta B} | z^2 \rangle e^{-i\gamma C(z^2)} \right]
  \]
Key Idea: **Average over instances**

- Parisi’s formalism requires delicate tricks
  - A replica-symmetry-breaking ansatz for the free energy:
    \[
    \mathbb{E}_J[\log Z_J] = \lim_{k \to 0} \frac{1}{k} \log \mathbb{E}[Z_J^k] \quad \Rightarrow \quad Z_J(T) = \text{tr}(e^{C_J/T})
    \]

- For QAOA, averaging over \( J \) is easier

\[
\frac{1}{n} \mathbb{E}_J[\langle C \rangle] = \mathbb{E}_J \left[ \langle s| e^{i\gamma C} e^{i\beta B} \frac{C}{n} e^{-i\beta B} e^{-i\gamma C} |s \rangle \right] = \frac{1}{2^n} \sum_{z^1 z^m z^2} \mathbb{E}_J \left[ e^{i\gamma C(z^1)} \langle z^1 | e^{i\beta B} | z^m \rangle \frac{C(z^m)}{n} \langle z^m | e^{-i\beta B} | z^2 \rangle e^{-i\gamma C(z^2)} \right]
\]
Key Idea: **Average over instances**

- Parisi’s formalism requires delicate tricks
  - A replica-symmetry-breaking ansatz for the free energy:
    \[
    \mathbb{E}_J[\log Z_J] = \lim_{k \to 0} \frac{1}{k} \log \mathbb{E}[Z_J^k]
    \]
    \[
    Z_J(T) = \text{tr}(e^{C_J/T})
    \]
- For QAOA, averaging over \( J \) is easier
  \[
  \frac{1}{n} \mathbb{E}_J[\langle C \rangle] = \mathbb{E}_J[\langle s | e^{i\gamma C} e^{i\beta B} \frac{C}{n} e^{-i\beta B} e^{-i\gamma C} | s \rangle]
  \]
Key Idea: **Average over instances**

- Parisi’s formalism requires delicate tricks
  - A replica-symmetry-breaking ansatz for the free energy:
    \[
    \mathbb{E}_J[\log Z_J] = \lim_{k \to 0} \frac{1}{k} \log \mathbb{E}[Z_J^k] \quad Z_J(T) = \text{tr}(e^{C_J/T})
    \]
    
- For QAOA, averaging over $J$ is easier
  \[
  \frac{1}{n} \mathbb{E}_J[\langle C \rangle] = \mathbb{E}_J \left[ \langle s | e^{i \gamma C} e^{i \beta B} C \frac{n}{n} e^{-i \beta B} e^{-i \gamma C} | s \rangle \right]
  \]
  \[
  = \frac{1}{2^n} \sum_{z^1 z^m z^2} \mathbb{E}_J \left[ e^{i \gamma C(z^1)} \langle z^1 | e^{i \beta B} | z^m \rangle \frac{C(z^m)}{n} \langle z^m | e^{-i \beta B} | z^2 \rangle e^{-i \gamma C(z^2)} \right]
  \]

For $\phi$ small, use

\[
\mathbb{E}_J[e^{i J \phi}] = 1 - \frac{1}{2} \phi^2 + \cdots \quad \mathbb{E}_J[J e^{i J \phi}] = i \phi + \cdots
\]
Key Idea: Average over instances

\[
\frac{1}{n} \mathbb{E}_J[\langle C \rangle] \approx \frac{i}{n^{3/2}} \sum_{z^1, z^2} \left[ \langle z^1 | e^{i\beta B} | 1 \rangle \langle 1 | e^{-i\beta B} | z^2 \rangle \sum_{k < \ell} \phi_{k\ell} \prod_{i < j} (1 - \frac{1}{2} \phi_{ij}^2) \right]
\]

\[
\phi_{ab} = \frac{\gamma}{\sqrt{n}} (z_a^1 z_b^1 - z_a^2 z_b^2)
\]
Key Idea: Average over instances

\[
\frac{1}{n} \mathbb{E}_J[\langle C \rangle] \approx \frac{i}{n^{3/2}} \sum_{z^1, z^2} \left[ \langle z^1 | e^{i\beta B} | 1 \rangle \langle 1 | e^{-i\beta B} | z^2 \rangle \sum_{k < \ell} \phi_{k\ell} \prod_{i < j} \left( 1 - \frac{1}{2} \phi_{ij}^2 \right) \right]
\]

Permutation symmetry → configuration basis

\[
\phi_{ab} = \frac{\gamma}{\sqrt{n}} (z_a^1 z_b^1 - z_a^2 z_b^2)
\]
Key Idea: Average over instances

$$\frac{1}{n} \mathbb{E}_J[\langle C \rangle] \approx \frac{i}{n^{3/2}} \sum_{z^1,z^2} \left[ \langle z^1 | e^{i\beta B} | 1 \rangle \langle 1 | e^{-i\beta B} | z^2 \rangle \sum_{k<\ell} \phi_{k\ell} \prod_{i<j} \left( 1 - \frac{1}{2} \phi_{ij}^2 \right) \right]$$

Permutation symmetry $\Rightarrow$ configuration basis

$$\phi_{ab} = \frac{\gamma}{\sqrt{n}} (z^1_a z^1_b - z^2_a z^2_b)$$

$$\sum_a n_a = n$$
Key Idea: Average over instances

\[
\frac{1}{n} \mathbb{E}_J \langle C \rangle \approx \frac{i}{n^{3/2}} \sum_{z^1, z^2} \left[ \langle z^1 | e^{i \beta B} | 1 \rangle \langle 1 | e^{-i \beta B} | z^2 \rangle \sum_{k < \ell} \phi_{k \ell} \prod_{i < j} \left( 1 - \frac{1}{2} \phi_{ij}^2 \right) \right]
\]

Permutation symmetry $\rightarrow$ configuration basis

\[
\phi_{ab} = \frac{\gamma}{\sqrt{n}} (z^1_a z^1_b - z^2_a z^2_b)
\]

\[
\sum_{a} n_a = n
\]

\[
\sum_{z^1, z^2} = \sum_{\{n_a\}} \binom{n}{n_{++}, n_{+-}, n_{-+}, n_{--}}
\]

$n$-bit strings \quad configurations
Key Idea: Average over instances

\[ \frac{1}{n} \mathbb{E}_J[\langle C \rangle] \approx \frac{i}{n^{3/2}} \sum_{z^1, z^2} \left[ \langle z^1 | e^{i\beta B} | 1 \rangle \langle 1 | e^{-i\beta B} | z^2 \rangle \sum_{k \neq \ell} \phi_{k\ell} \prod_{i < j} (1 - \frac{1}{2} \phi_{ij}^2) \right] \]

Permutation symmetry $\rightarrow$ configuration basis

\[ \phi_{ab} = \frac{\gamma}{\sqrt{n}} (z_a^1 z_b^1 - z_a^2 z_b^2) \]

\[ \sum_a n_a = n \]

For general $p$, there are $2^{2p}$ configurations

\[ \exp(O(p)) \text{ complexity} \]
Performance of the QAOA on the SK model

• Turn the crank, we get at $p = 1$

$$V_1 = \lim_{n \to \infty} \frac{1}{n} \mathbb{E}_J[\langle C \rangle] = \gamma e^{-2\gamma^2} \sin 4\beta$$

Optimum @ $\beta = \frac{\pi}{8}, \gamma = \frac{1}{2}$

$$\Rightarrow \max_{\gamma, \beta} V_1 = \frac{1}{\sqrt{4e}} \approx 0.303$$
Performance of the QAOA on the SK model

• Turn the crank, we get at $p = 1$

$$V_1 = \lim_{n \to \infty} \frac{1}{n} \mathbb{E}_J[\langle C \rangle] = \gamma e^{-2\gamma^2} \sin 4\beta$$

⇒ max $V_1 = \frac{1}{\sqrt{4e}} \approx 0.303$

Optimum @ $\beta = \frac{\pi}{8}, \gamma = \frac{1}{2}$

• Can also show concentration

$$\lim_{n \to \infty} \mathbb{E}_J[\langle (C/n)^2 \rangle] = \lim_{n \to \infty} \mathbb{E}^2_J[\langle C/n \rangle]$$
Performance of the QAOA on the SK model

- Turn the crank, we get at \( p = 1 \)
  \[
  V_1 = \lim_{n \to \infty} \frac{1}{n} \mathbb{E}_J[\langle C' \rangle] = \gamma e^{-2\gamma^2} \sin 4\beta
  \]
  Optimum @ \( \beta = \frac{\pi}{8}, \gamma = \frac{1}{2} \)
  \[\Rightarrow \max_{\gamma, \beta} V_1 = \frac{1}{\sqrt{4e}} \approx 0.303\]

- Can also show concentration

- Generic vs. optimized QAOA
Performance of the QAOA on the SK model

• Turn the crank, we get at $p = 1$

$$V_1 = \lim_{n \to \infty} \frac{1}{n} \mathbb{E}_J[\langle C \rangle] = \gamma e^{-2\gamma^2} \sin 4\beta$$

Optimum @ $\beta = \frac{\pi}{8}, \gamma = \frac{1}{2}$

$$\Rightarrow \max_{\gamma,\beta} V_1 = \frac{1}{\sqrt{4e}} \approx 0.303$$

• Can also show concentration

$$\lim_{n \to \infty} \mathbb{E}_J[\langle (C/n)^2 \rangle] = \lim_{n \to \infty} \mathbb{E}_J^2[\langle C/n \rangle]$$

• Generic vs. optimized QAOA

Suppose $J_{ij} \in \{\pm 1\}$

Then $\gamma$ is periodic on $[0, \sqrt{n}\pi)$

$\beta$ is periodic on $[0, \pi/2)$
Performance of the QAOA on the SK model

• Turn the crank, we get at \( p = 1 \)

\[
V_1 = \lim_{n \to \infty} \frac{1}{n} \mathbb{E}_J[\langle C \rangle] = \gamma e^{-2\gamma^2} \sin 4\beta
\]

Optimum @ \( \beta = \frac{\pi}{8}, \gamma = \frac{1}{2} \)

\[
\Rightarrow \max_{\gamma, \beta} V_1 = \frac{1}{\sqrt{4e}} \approx 0.303
\]

• Can also show concentration

\[
\lim_{n \to \infty} \mathbb{E}_J[\langle (C/n)^2 \rangle] = \lim_{n \to \infty} \mathbb{E}_J^2[\langle C/n \rangle]
\]

• Generic vs. optimized QAOA

Suppose \( J_{ij} \in \{ \pm 1 \} \)

Then \( \gamma \) is periodic on \([0, \sqrt{n}\pi]\)

\( \beta \) is periodic on \([0, \pi/2]\)

Generic QAOA state has \( \langle C \rangle = e^{-O(n)} \) !!
Performance of the QAOA on the SK model

• Higher $p$ : our current method uses $O(4^p)$ memory and $O(16^p)$ time

\[ V_p(\gamma, \beta) = \lim_{n \to \infty} \frac{1}{n} E_J[\langle \gamma, \beta | C | \gamma, \beta \rangle] \]
Performance of the QAOA on the SK model

• Higher $p$: our current method uses $O(4^p)$ memory and $O(16^p)$ time

$$V_p(\gamma, \beta) = \lim_{n \to \infty} \frac{1}{n} \mathbb{E}_J[\langle \gamma, \beta | C | \gamma, \beta \rangle]$$

<table>
<thead>
<tr>
<th>$p$</th>
<th>Best known $V_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.303265</td>
</tr>
<tr>
<td>2</td>
<td>0.407545</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>8</td>
<td>0.607266</td>
</tr>
<tr>
<td>11$^\dagger$</td>
<td>0.639311 \textbf{beats SDP!}</td>
</tr>
<tr>
<td>12$^\dagger$</td>
<td>0.646557</td>
</tr>
</tbody>
</table>

$^\dagger$ unoptimized
Performance of the QAOA on the SK model

- Higher $p$: our current method uses $O(4^p)$ memory and $O(16^p)$ time

$$V_p(\gamma, \beta) = \lim_{n \to \infty} \frac{1}{n} \mathbb{E}_J[\langle \gamma, \beta | C | \gamma, \beta \rangle]$$

<table>
<thead>
<tr>
<th>$p$</th>
<th>Best known $V_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.303265</td>
</tr>
<tr>
<td>2</td>
<td>0.407545</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>8</td>
<td>0.607266</td>
</tr>
<tr>
<td>11$^\dagger$</td>
<td>0.639311</td>
</tr>
<tr>
<td>12$^\dagger$</td>
<td>0.646557</td>
</tr>
</tbody>
</table>

$^\dagger$ unoptimized

beats SDP!
Performance of the QAOA on the SK model

- Higher $p$: our current method uses $O(4^p)$ memory and $O(16^p)$ time

$$V_p(\gamma, \beta) = \lim_{n \to \infty} \frac{1}{n} \mathbb{E}_J[\langle \gamma, \beta | C | \gamma, \beta \rangle]$$

<table>
<thead>
<tr>
<th>$p$</th>
<th>Best known $V_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.303265</td>
</tr>
<tr>
<td>2</td>
<td>0.407545</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>8</td>
<td>0.607266</td>
</tr>
<tr>
<td>11†</td>
<td>0.639311</td>
</tr>
<tr>
<td>12†</td>
<td>0.646557</td>
</tr>
</tbody>
</table>

\[\text{† unoptimized} \]

- Beats SDP!
Performance of the QAOA on the SK model

- Higher $p$: our current method uses $O(4^p)$ memory and $O(16^p)$ time

$$V_p(\gamma, \beta) = \lim_{n \to \infty} \frac{1}{n} E_J[\langle \gamma, \beta | C | \gamma, \beta \rangle]$$

<table>
<thead>
<tr>
<th>$p$</th>
<th>Best known $V_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.303265</td>
</tr>
<tr>
<td>2</td>
<td>0.407545</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>8</td>
<td>0.607266</td>
</tr>
<tr>
<td>11†</td>
<td>0.639311</td>
</tr>
<tr>
<td>12†</td>
<td>0.646557</td>
</tr>
</tbody>
</table>

† unoptimized

If $\lim_{p \to \infty} V_p = \Pi_*$, then a power law fit of optimized $V_p$ yields

$$\tilde{V}_p \approx \Pi_* - \frac{1.2}{(p + 2)^{0.9}}$$

If $\lim_{p \to \infty} V_p = \Pi_*$, then a power law fit of optimized $V_p$ yields $\tilde{V}_p$ beats SDP!
Summary

• We *analytically* obtain a formula for **typical case performance** of the QAOA on the SK model at high $p$
  • Evaluation takes $O(16^p)$ currently but may be improvable

• QAOA **beats** Semi-Definite Programming at $p = 11$

• **Concentration** over instances and measurements

https://github.com/leologist/QAOA-SK
Outlook

• Show convergence of QAOA as \( p \to \infty \)?

\[
\lim_{p \to \infty} \lim_{n \to \infty} \frac{?}{\lim_{n \to \infty} \lim_{p \to \infty}}
\]
Outlook

• Show convergence of QAOA as $p \to \infty$?
\[
\lim_{p \to \infty} \lim_{n \to \infty} \frac{C}{n} = \lim_{n \to \infty} \lim_{p \to \infty} \frac{C}{n}
\]

• Average over instances for harder problems for provable speedup?

$q$-spin model

\[
C = \sum_{i_1 < \cdots < i_q} J_{i_1 \cdots i_q} Z_{i_1} \cdots Z_{i_q}
\]

Provably hard for classical algorithms due to their “Overlap Gap Property”

[Gamarnik Jagannath 2019]
[Gamarnik Jagannath Wein 2020]

Montanari’s algorithm stuck at 98.4%
approximation ratio for $q=3$

[Alaoui Montanari 2020]

QAOA @ $p=1$
gets 33% for $q=3$