Research Report: The application of QAOA

Yixian Qiu

August 7, 2019

1 MaxCut problem

1.1 QAOA for Max-Cut requires hundreds of qubits for quantum speed-up (Scientific reports,(2019) 9:6903)

Given a graph, assign one of two colors (e.g., black or white) to each vertex. An edge can be "cut" when it connects two vertices of different color. The problem is to assign the colors so that as many edges as possible can be cut at the same time. The objective function corresponds to the hermitian operator:

$$C = -\frac{1}{2}\sum_{(i,j)\in E} Z_i Z_j$$

The variational state prepared by the QAOA circuit has the form:

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

where $B = \sum_k X_k$ being the sum of the X Pauli matrices associated to each vertex of the graph. Here is a typical 8-qubit instance with average over 384 noise realizations.

The computational cost of both the classical approach and the quantum variational algorithm are showed in Fig. 1, the blue markers corresponding to datapoints for the classical AKMAXSAT, while the red and green markers for QAOA with p =4 and p=8 respectively.

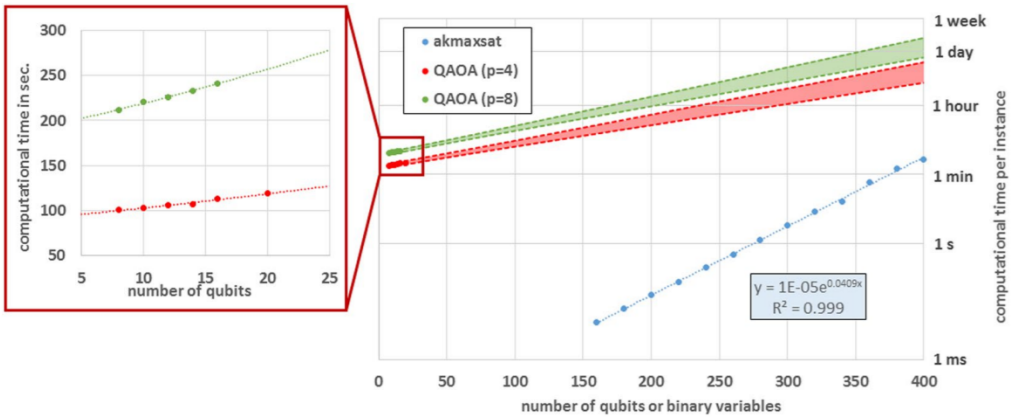


Figure 1: Main panel: Computational cost of solving a single Max-Cut instance on random 3-regular graphs. Blue markers correspond to the classical baseline (AKMAXSAT solver) while red and green marks correspond to the experimental time required by the quantum algorithm QAOA, with p=4 and p=8 respectively. The error bars for the single data points are smaller than the markers. Notice that in the time needed by QAOA to partition graphs with 20 vertices, AKMAXSAT partitions graphs about 20 times larger. The blue dashed line is the result of a fitting procedure with an exponential function. The red and green areas are associated with a 95% confidence interval for the prediction of the QAOA cost based on a linear regression of $log_{10}T$ as a function of the number of qubits (here T is the computational time per instance).

1.2 Performance of the Quantum Approximate Optimization Algorithm on the Maximum Cut Problem (arxiv: 1811.08419)

Key words: Automatic differentiation, Stochastic gradient descent.

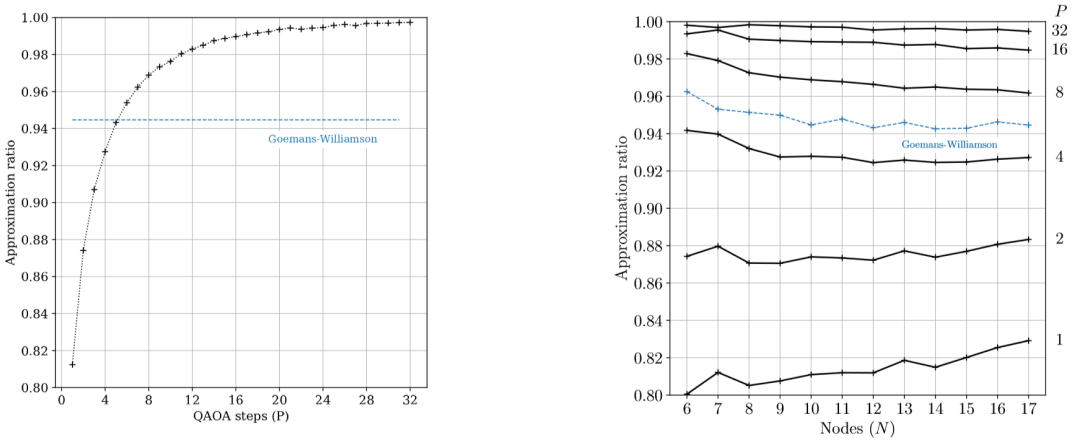


Figure 2: (The left) The average approximation ratio of QAOA on the MaxCut optimization problem for 10 node graphs, as a function of QAOA steps (P). Each point represents an independently optimized protocol obtained via stochastic gradient descent. The training data consists of 100 graphs from the E-R ensemble with edge probability 50%, and the test data an independently sampled collection of 100 graphs from the same ensemble. For this problem size, QAOA with 5 steps matches the average performance of the classical, polynomial time Goemans-Williamson algorithm on the same data set. (The right) The approximation ratio of QAOA on the MaxCut optimization problem, as a function of graph size (N) and QAOA steps (P). We also show the performance of the classical, polynomial time Goemans-Williamson algorithm on the same test sets. The performance of QAOA increases with circuit depth, and significantly exceeds that of GoemansWilliamson by P = 8. The performance of QAOA degrades as the graphs get bigger. However, quantum circuits with P >= 8 maintain their relative performance advantage over the classical algorithm.

In this paper, they train on batches of graphs drawn from the same statistical ensemble, and find a protocol $(\beta^{opt}, \gamma^{opt})$ that is effective for an entire class of problem instances,

$$\left(\boldsymbol{\beta}^{\mathrm{opt}}, \boldsymbol{\gamma}^{\mathrm{opt}}\right) = -\arg\min_{\boldsymbol{\beta},\boldsymbol{\gamma}} \frac{1}{|\mathcal{T}|} \sum_{C \in \mathcal{T}} \langle \boldsymbol{\beta}, \boldsymbol{\gamma} \left| H_{\mathrm{C}} \right| \boldsymbol{\beta}, \boldsymbol{\gamma} \rangle$$

Here $\mathcal{T}$ in the training set of graph adjacency matrices, and $|\mathcal{T}|$ is the number of graphs in the training set. They initial QAOA parameters from a normal distribution, of which a standard deviation of 0.01 and a mean of 0.5 (to avoid a symmetry about zero in the parameter space) appears satisfactory.

They perform automatic differentiation of the performance metric with respect to the parameters of the quantum circuit, and can therefore train the quantum circuits using back-propagation and stochastic gradient descent, as shown in Fig 2. Typical optimal protocols for different QAOA steps are illustrated in Fig. 3.

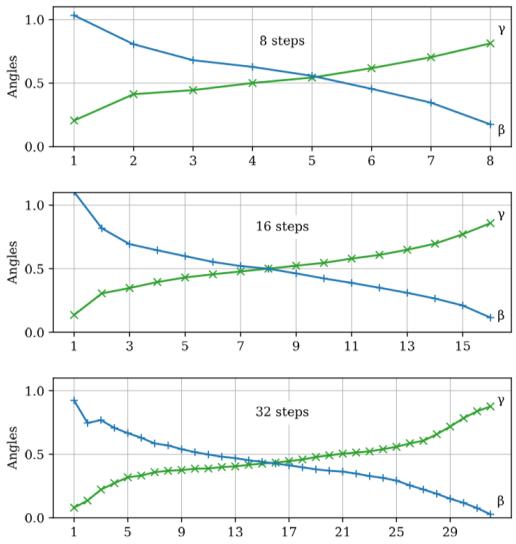


Figure 3: Examples of optimized protocols $(\beta^{opt}, \gamma^{opt})$ for MaxCut QAOA on 10 node graphs, with 8, 16, and 32 QAOA steps.

1.3 Quantum Annealing: a journey through Digitalitalization, Control, and hybrid Quantum Variational schemes (arxiv: 1906.08948)

Key words: Residual energy bound.

For QAOA residual energy bound for MaxCut on periodic 2-regular graphs, this amounts to searching for the minimum of $\widehat{H}_z = \sum_{\langle i,j\rangle\in\mathcal{E}} \left(\hat{\sigma}_i^z \hat{\sigma}_j^z - 1\right)$. The expectation value is

$$\epsilon_{\mathrm{P}}^{\mathrm{res}}(\gamma,\boldsymbol{\beta}) \geq \left(\frac{E_{\mathrm{gs}}^{(-)}}{2(2\mathrm{P}+2)} + 1\right) = \frac{1}{2\mathrm{P}+2}$$

where $E_{\mathrm{gs}}^{(\pm)}$ is the ground state energy of $\hat{\mathcal{H}}_z^{(\pm)}$ and the inequality follows from the standard variational principle. This establishes: when 2P < N(the variables 2P is smaller than the number of constraints N), one can give a lower bound for the residual energy is $\frac{1}{2\mathrm{P}+2}$, as shown in Fig. 4.

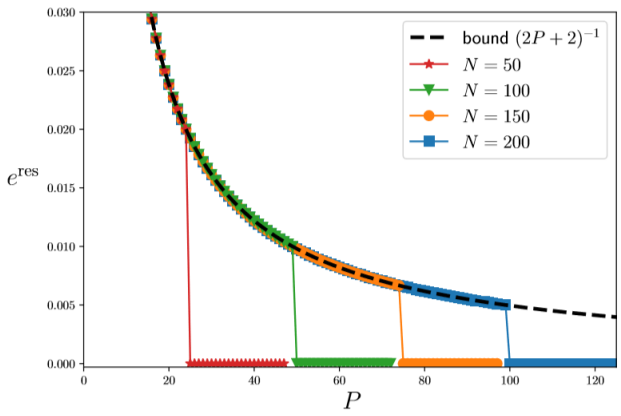


Figure 4: Optimal residual energies $\epsilon_{\mathrm{P}}^{\mathrm{res}}$ obtained applying the QAOA with P Trotter steps for various system sizes N = 50,100,150,200. The symbols represent the data obtained by numerical optimization while the dashed black line represents the theoretical bound obtained in equation above.

2 Boolean satisfiability problem (SAT)

2.1 Reachability Deficits in Quantum Approximate Optimization (arxiv: 1906.11259)

Key words: critical depth p^*.

There is an Ising Hamiltonian from a given SAT instance as $\mathcal{H}_{\mathrm{SAT}} = \sum_l \mathcal{P}(l)$ where l indexes each clause in the SAT instance and $\mathcal{P}(l)$ are rank-one projectors that penalize each unsatisfiable assignments with at-least 1 unit of energy. Satisfiable instances are characterized by a zero ground state energy, $E_g = 0$ and unsatisfiable instances with $E_g \geq 1$.

QAOA with standard settings, $\mathcal{H}_x = \sum_i \sigma_x^{(i)}$ and $\mathcal{H} = \mathcal{H}_{\mathrm{SAT}}$, can now be used to calculate the energy approximation E_g^{QAOA}, where

$$E_g^{\mathrm{QAOA}} = \min_{\boldsymbol{\alpha},\boldsymbol{\beta}} \langle\psi(\boldsymbol{\alpha},\boldsymbol{\beta})\,|\mathcal{H}_{\mathrm{SAT}}|\,\psi(\boldsymbol{\alpha},\boldsymbol{\beta})\rangle$$

as a function of clause density α, for p - depth QAOA circuit on randomly generated 3 - SAT and 2 - SAT instances, as shown in Fig. 5. If with a new driver $\mathcal{H}_x = |+\rangle\langle+|^{\otimes n}$. The modified version requires lower circuit depths for achieving similar performance as standard QAOA but still exhibits reachability deficits, as shown in Fig. 6.

For some critical depth p^*, QAOA returns the exact ground state energies for both 3 - SAT and 2 -SAT instances up-to a set tolerance. We fix the tolerance as a condition on the overlap

$$\eta = \sum_{i=1}^{d} |\langle\psi_p(\boldsymbol{\alpha},\boldsymbol{\beta})|\mathrm{gs}_i\rangle|^2 \tag{1}$$

where $\{|\mathrm{gs}_i\rangle\}$ be the d degenerate ground states of $\mathcal{H}_{\mathrm{SAT}}$. The critical depth is then defined as the minimum circuit depth for which the algorithm attains $\eta > 0.99$. The critical depth depends on the problem instance non-trivially. Based on the author's findings, QAOA suffers from reachability deficits for circuit depths $p < p^*$, where p^*, is dependent on the problem density (the ratio of an instances constraints to variables).

The next part about "Variational Grover Search" will be expanded in the other section.

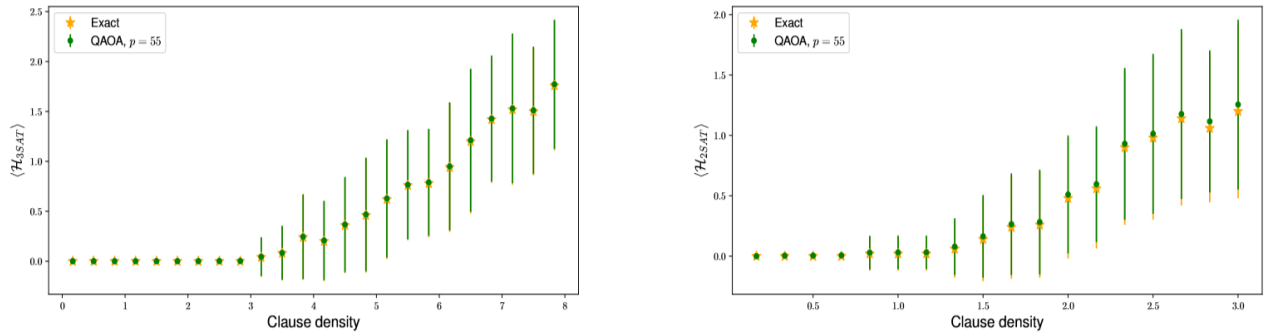


Figure 5: Energy versus clause density for 3-SAT (Left) and 2-SAT (Right) for n=6.

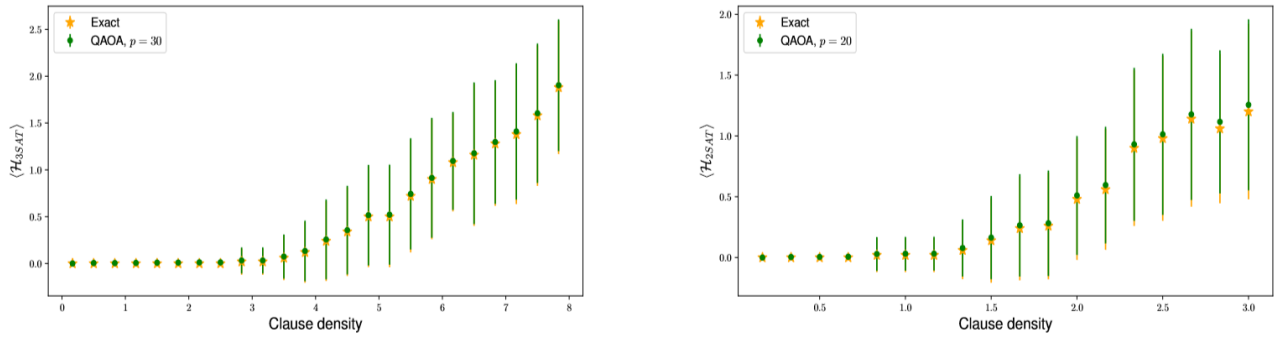


Figure 6: Energy versus clause density for 3-SAT (Left) and 2-SAT (Right) for n=6,with driver Hamiltonian $\mathcal{H}_x = |+\rangle\langle+|^{\otimes n}$.

3 State Transfer problem

3.1 Optimizing QAOA: Success Probability and Runtime Dependence on Circuit Depth (arxiv: 1905.12134)

The state transfer problem of interest is defined in a one-dimensional qubit chain of length N. The two Hamiltonians used for QAOA iteration to be

$$\hat{H}_C = |\overline{N}\rangle\langle\overline{N}| = \frac{1}{2}\left(\sigma_N^z + I_N\right)$$

$$\hat{H}_B = \sum_{i=1}^{N}\left(\sigma_i^x\sigma_{i+1}^x + \sigma_i^y\sigma_{i+1}^y\right)$$

The task of state transfer is to realize a unitary transformation U such that:

$$|\psi_f\rangle = U\,|\psi_i\rangle = U(\alpha|\overline{1}\rangle + \beta|\overline{0}\rangle) = \alpha|\overline{N}\rangle + \beta|\overline{0}\rangle$$

where $|\overline{N}\rangle$ satisfies that $\sigma_n^z|\overline{n}\rangle = |\overline{n}\rangle$, $\sigma_{i,i\neq n}^z|\overline{n}\rangle = -|\overline{n}\rangle$, e.g. $|\overline{n}\rangle = |0\rangle_1|0\rangle_2\cdots|0\rangle_{n-1}|1\rangle_n|0\rangle_{n+1}\cdots|0\rangle_N$.

A depth p QAOA realizes the following unitary transformation:

$$U_p = \prod_{k=1}^{p} U\left(\hat{H}_C, \delta_k^C\right) U\left(\hat{H}_B, \delta_k^B\right)$$

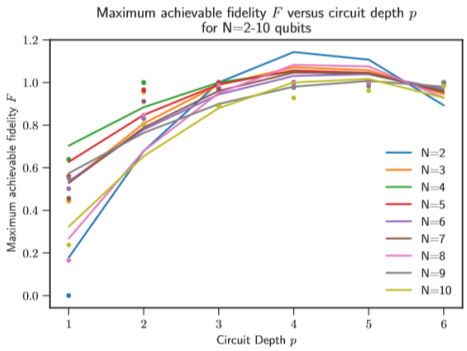
To simplify the analysis, the evolution resembles a Grover oracle:

$$U_p = \left(e^{-i\pi|\overline{N}\rangle\langle\overline{N}|}U\left(\hat{H}_B, \delta\right)\right)^p$$

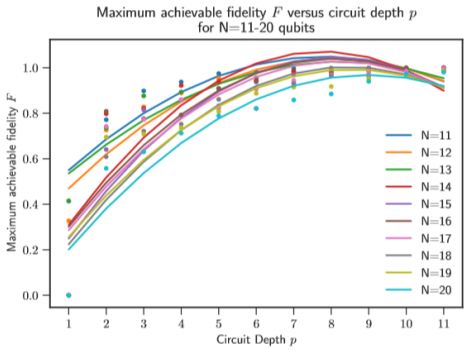
With this ansatz, the result can be connected to the scaling analysis in **conventional Grover search** and thus serves as a **lower bound** on the success probability for the optimized QAOA to be discussed in the subsequent sections.

In the large depth limit, the term with the largest power of p dominates:

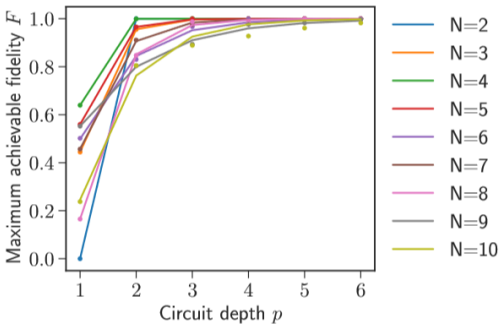
$$\lim_{p\to\infty} P_{\text{succ}}\,(p) \propto p^{4p+2}$$



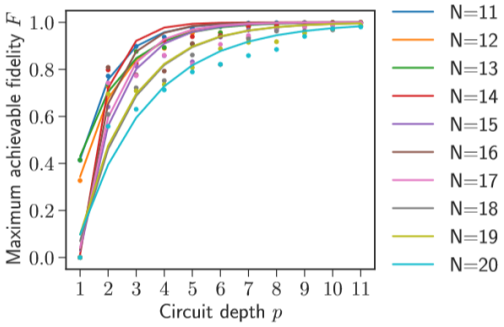
(a) $N = 2 \rightarrow 10$ qubits



(b) $N = 11 \rightarrow 20$ qubits



(a) $N = 2 \rightarrow 10$ qubits



(b) $N = 11 \rightarrow 20$ qubits

Figure 7: (The left) Maximum achievable fidelity F using QAOAs as a function of the circuit depth p with no constraints on t_f for N $= 2 \rightarrow 20$ qubits. The dots are numerical points. We fit the results with quadratic function $F(p) = ap2 + bp + c$, as represented by the lines. We observe that the fidelity grows at low circuit depth, and then slowly converges to 1.0. (The right) Maximum achievable fidelity F as a function of the circuit depth p with no constraints on tf for N $= 2 \rightarrow 20$ qubits.The dots are numerical points. We fit the results with inverted exponential function $F(p) = 1 - exp(-a(p - b))$. We can observe fidelity grows rapidly at low circuit depth, and then the fidelity slowly converge to unity.

In the low-depth limit, only the lowest order of δ terms dominates:

$$\lim_{p \to 1} P_{\text{succ}}(p) \propto F(N)^2 \delta p^2$$

Figure 7 shows maximum achievable fidelity F as a function of circuit depth p with no constraints on t_f for N $= 2 \rightarrow 20$ qubits, which agree with the analysis above.

As a simple example, Fig. 8 plot the control landscape for N = 3 qubits with p = 2,4, which shows that the QAOA ansatz with p = 2 is uncontrollable, as there are many local minima.

Now discuss the optimized QAOA performance with a fixed physical run time t_f. The authors numerically study the maximum achievable fidelity F versus the circuit depth p for a fixed t_f in Figs. 9 to 10.

Figure 11 shows the performances of the QAOA with a fixed physical runtime. We identify three different temporal dependencies of fidelity as predicted by the Lieb-Robinson bound, as depicted in Fig. 12: exponentially suppressed region; exponentially growing region; and steady growing region.

The three regions of the growth region become more apparent as circuit depth increase in Fig. 13. In Fig. 14, we fit the minimum required run time t_f for achieving fidelity F = 0.99 as a function of the number of qubits (N $= 2 \rightarrow 19$).

The linear dependence from the Lieb-Robinson bound is seen with $t_f \sim 2.439N$. Given the same amount of run time, as shown in Fig. 12. The Lieb-Robinson bound gives a prediction about the size of the exponentially suppression region: $t_s \sim N/(6eJ) = 0.03N$. plot the exponentially suppressed time as a function of the number of qubits in Fig. 13.

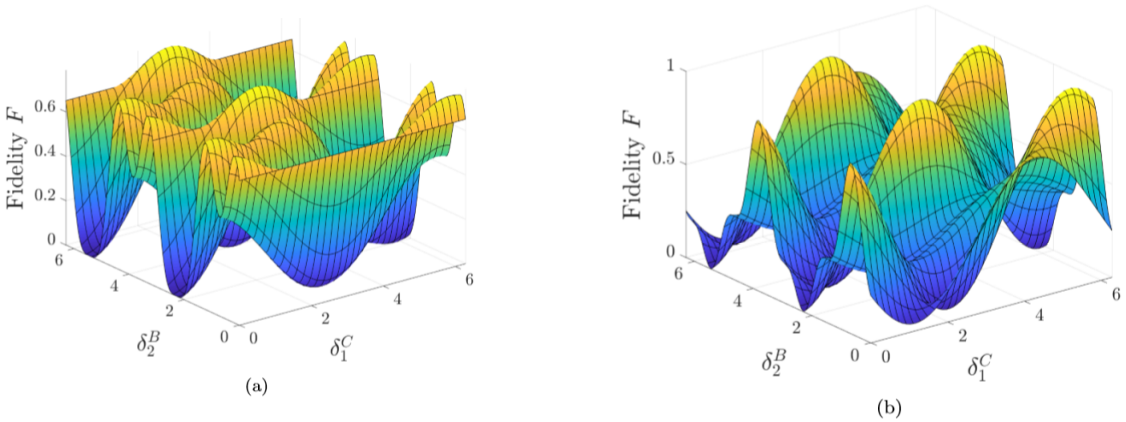
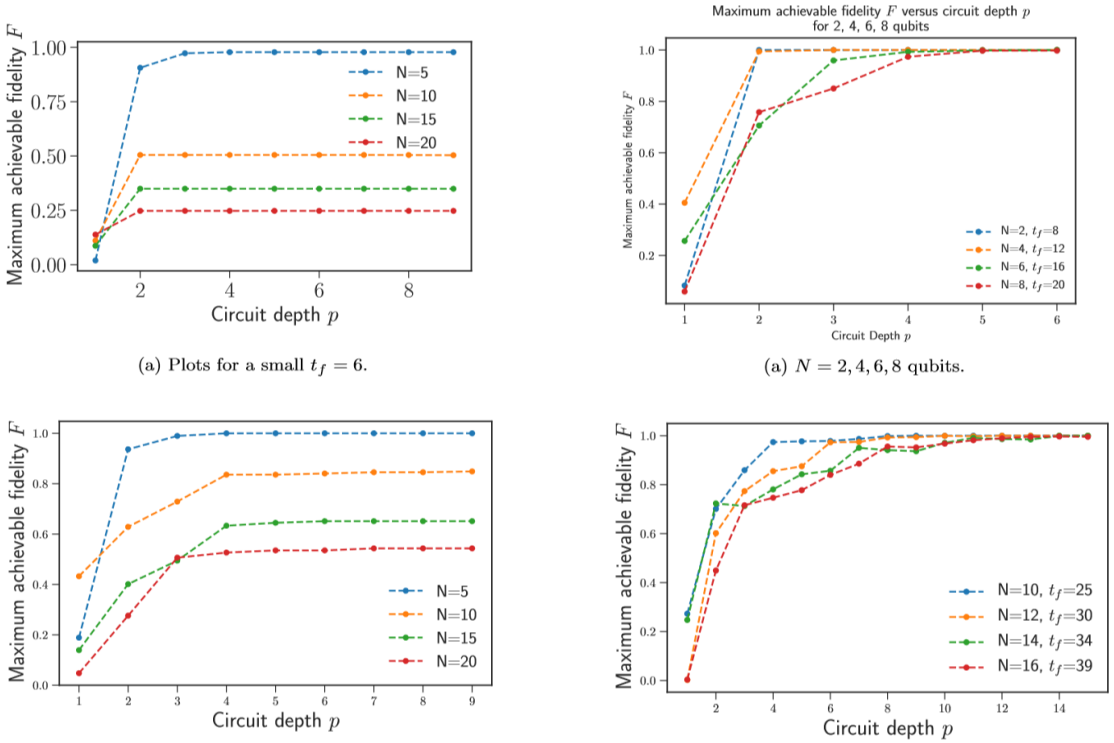


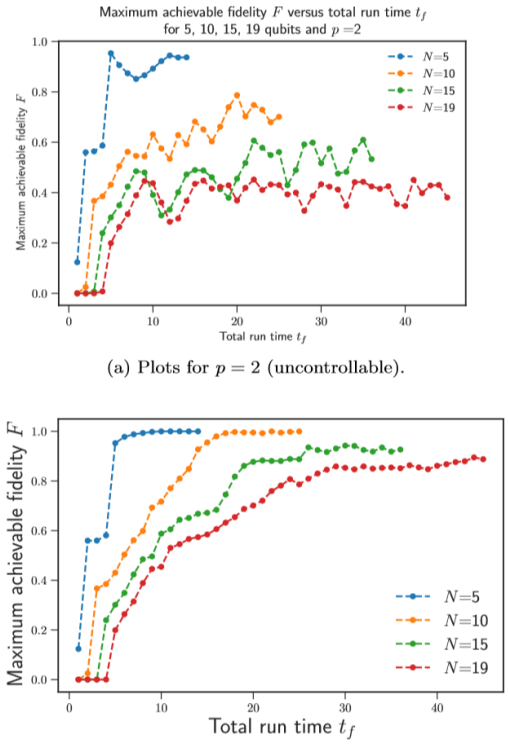
Figure 8: (a) The control landscape for two chosen variables of a depth-2 QAOA, which admits a maximum achievable fidelity of 0.787. As we observed many local minima, the system is uncontrollable. (b) The control landscape for two chosen variables of a depth-4 QAOA, which admits a maximum achievable fidelity of 1.000. Since all local minima are global minima, the system is controllable.



(a) Plots for a small $t_f = 6$.

(a) $N = 2, 4, 6, 8$ qubits.

Figure 9: (The left) For fixed t_f, we observed that there exists a circuit depth p beyond which there will be no improvement of fidelity. We find a depth-3 circuit is sufficient for $t_f = 6$ while a depth-4 circuit is needed for $t_f = 13$. (The right) Maximum achievable fidelity F versus circuit depth p with fixed t_f. The physical run time t_f is chosen through a hyper-parameter grid search separately.



(a) Plots for $p = 2$ (uncontrollable).

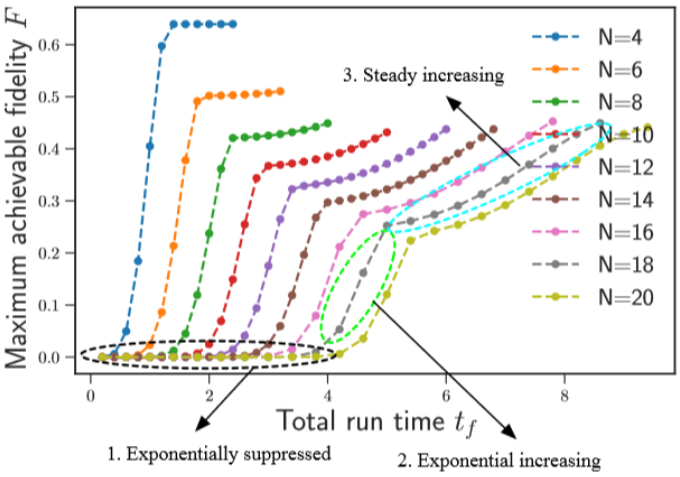
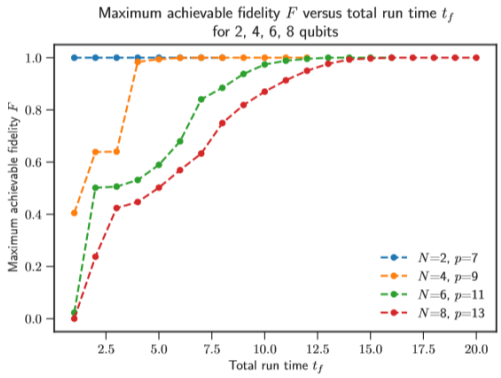
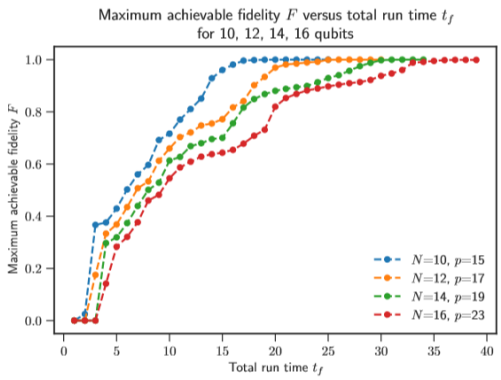


Figure 10: (The left) The oscillating behavior is due to the low circuit depth; such behaviour disappears when the circuit depth p grows larger (equivalently, the system gets more controllable). (The right) Maximum achievable fidelity F versus small physical run time t_f with a sufficiently large circuit depth p for N = 4, 6, 8, 10, 12, 14, 16, 18, 20 qubits.



(a) $N = 2, 4, 6, 8$ qubits.



(b) $N = 10, 12, 14, 16$ qubits.

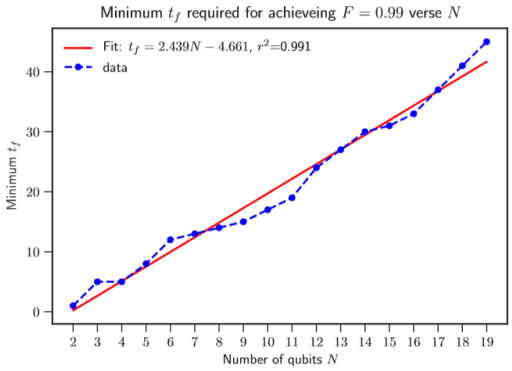
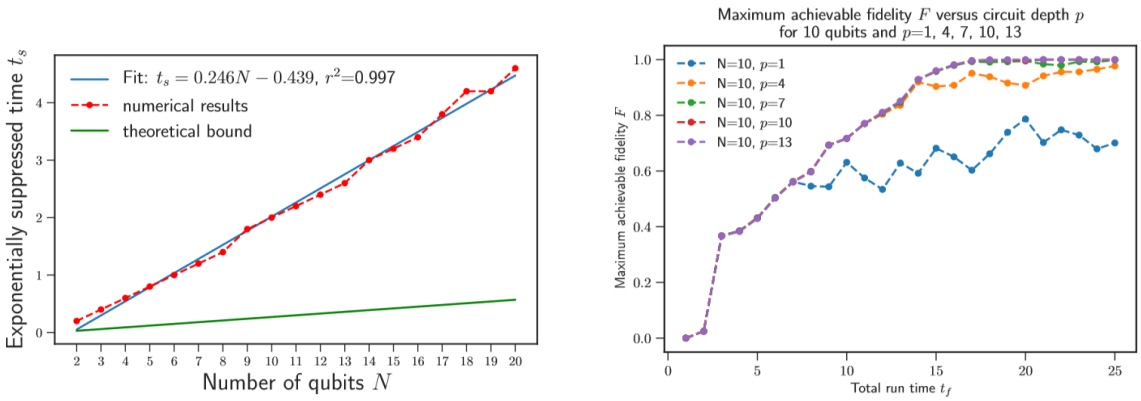
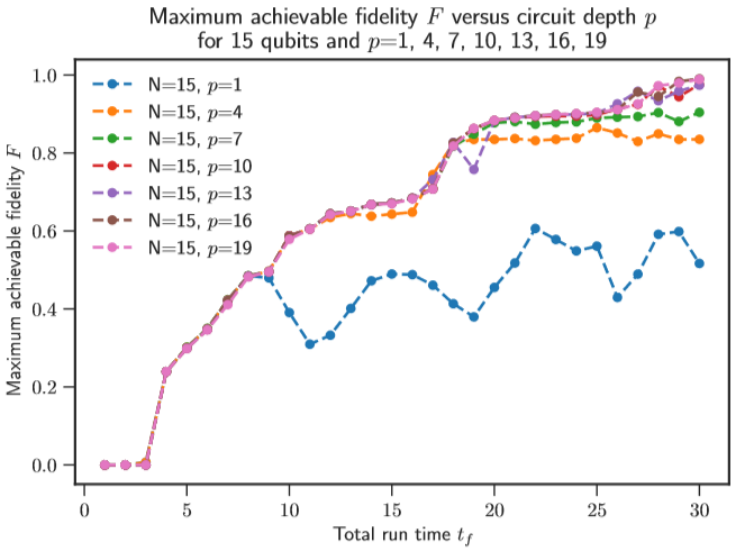


Figure 11: (The left) Maximum achievable fidelity F versus physical run time t_f with a sufficiently large circuit depth p. (The right) The Lieb-Robinson bound gives a lower bound of approximately $t_f = 0.03N + c$.



(a) $N = 10$ qubits.

Figure 12: Maximum achievable fidelity F versus physical run time t_f using QAOA as a function of different circuit depth p. The lines with oscillating behaviors are uncontrollable.



(b) $N = 15$ qubits.

Figure 13: The Lieb-Robinson bound gives a bound approximately of $t_f = 0.03N + c$. The exponentially suppressed time is defined as the time needed to make fidelity higher than 0.01.

4 Variational Grover Search

4.1 Reachability Deficits in Quantum Approximate Optimization (arxiv: 1906.11259)

(The theory about p^* is on the section (SAT).)

Variational Grover search can be thought of as QAOA with the following setting, $\mathcal{H} = |\omega\rangle|\omega|$ and $\mathcal{H}_x = |+\chi+|^{\otimes n}$, where $|\omega\rangle \in C_2^{\otimes n}$ is the objective state we are searching for. Hence, the minimization of the expected value of the Hamiltonian is

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta}} \langle (1 - |\omega\rangle\langle\omega|)\rangle_{|\psi(\boldsymbol{\alpha},\boldsymbol{\beta})\rangle} = \min_{\boldsymbol{\alpha},\boldsymbol{\beta}} \left(1 - |\langle\omega|\psi(\boldsymbol{\alpha},\boldsymbol{\beta})\rangle|^2\right)$$

And the approximated energy will be $E_g^{\text{QAOA}} = 1 - |B_p|^2$. Minimization of equation above is numerically computed as a funcion of circuit-depth. To establish the dependence of p^* on the problem density, the author increase n, the size of the search space and recover a Grover scaling, $\mathcal{O}(\sqrt{N})$ for p^*, in Fig 14.

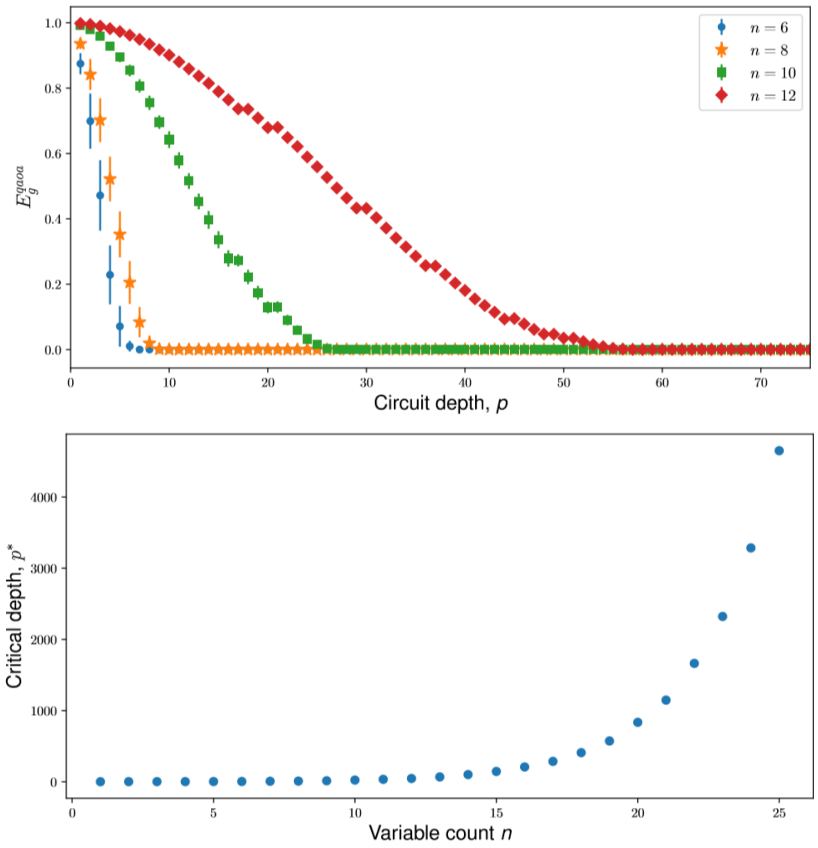


Figure 14: Top: Convergence to exact ground state energy as a function of QAOA circuit depth for the variational Grover search model on search space size, n = 6,8,10 and 12. Bottom: Scaling of critical depth p^* with variable count, n.