Evolution of Approximation Algorithms: Local Hamiltonian Problem

Owen Bardeen[*], Jenny Chen[†], and Nathan Song[‡]

C191, Spring 2023

Abstract

In our research paper, we will first focus on the widely established QAOA and give an intuitive explanation behind how it brings one to the solution state. But rather than restricting ourselves to classical combinatorial problems, we transition into a class of algorithms which attempt to solve the more general local Hamiltonian problem. We will outline the various proposed approximation algorithms and offer insight into their functionality and drawbacks.

1 Introduction

Despite recent advancements in quantum computing hardware, we remain in the Noisy Intermediate Scale Quantum (NISQ) era where limitations on gate fidelity and scalability prevent us from being able to carry out algorithms like Shor's which provide exponential speedup. Thus we are restricted to a narrow domain of algorithms which are shallow depth and thus implementable on current hardware. They make up a class of hybrid quantum-classical algorithms which include a classical optimizer and some form of a quantum flavor included in components such as the cost function or ansatz. Well known examples include the quantum approximation optimization algorithm (QAOA) and variational quantum eigensolver (VQE) but their results, such as QAOA, fail to exceed that of the best known classical algorithms like semi-definite programming (SDP) on classical combinatorial problems such as Maxcut.

Thus it has prompted a greater interest towards solving a broader class of problems, namely the k-local Hamiltonian problem. It consists of solving for the ground state for some Hamiltonian within an accuracy ϵ, and the problem has been proven to be QMA-complete for $k \geq 3$ [12]. There are a variety of proposed approximation algorithms, including but not limited to using SDP

[*]Contributed to QAOA, its implementation and drawbacks

[†]Introduction, Product state approximation algorithms, and overall evolution into King's combined method

[‡]Anshu's shallow depth circuit algorithm and proofs of his first two theorems

relaxation to round to product states and applying shallow depth quantum circuits to arbitrary product states [9, 2].

2 QAOA

Initially a leading approximation algorithm for classical problems, QAOA was first introduced by Fahri and Goldstone in 2014. Its purpose is to solve classical combinatorial problems, namely Max-Cut. The setup outlined in Ref. [7] is as follows:

Let there be a graph $G = (V, E)$ for which we want to find its max cut. Let

$$\begin{aligned} C &= \sum_{\{i,j\}\in E} (\mathbb{I} - Z_i Z_j) \text{ be the cost Hamiltonian} \\ B &= \sum_{i\in V} (X_i) \text{ be the mixer Hamiltonian} \\ |\psi_0\rangle &= |+\rangle^{\otimes n} \text{ be the initial state} \\ |\psi(\gamma,\beta)\rangle &= e^{-i\gamma_1 C} e^{-i\beta_1 B} \dots e^{-i\gamma_p C} e^{-i\beta_p B} |\psi_0\rangle \text{ where } p \text{ is the depth} \end{aligned}$$

The cost Hamiltonian effectively counts the number of edges cut because the eigenstate of $\mathbb{I} - Z_i Z_j$ consists of the vectors in which $i \neq j$ and the mixer Hamiltonian changes the vertices' values.

The goal is to maximize the expectation value of the new state over the cost Hamiltonian, thereby maximizing the number of edges with opposing vertices:

$$\begin{aligned} F(\gamma,\beta) &= \langle\psi(\gamma,\beta)| C |\psi(\gamma,\beta)\rangle \\ &= \langle\psi_0| e^{i\gamma_p C} e^{i\beta_p B} \dots e^{i\gamma_1 C} e^{i\beta_1 B} C e^{-i\gamma_1 C} e^{-i\beta_1 B} \dots e^{-i\gamma_p C} e^{-i\beta_p B} |\psi_0\rangle \end{aligned}$$

We can treat it as a sort of "cost function" that we minimize where $C_{cost} = -F(\gamma,\beta)$ and to map out this cost landscape we choose a set of γ and β to measure the expectation value with respect to the cost Hamiltonian C. Then using a classical optimizer, such as gradient descent, we attempt to solve for the parameters which brings us the closest to the optimal max cut state.

For a brief explanation of why we need to alternately time evolve under each Hamiltonian, consider the difference between the purpose of each Hamiltonian. The cost Hamiltonian gives the "direction" for our classical gradient descent algorithm to go and alter our phases of eigenstates, since it is diagonal in our computational basis. The mixer Hamiltonian allows us to escape from local minima and actually alter the amplitudes of the eigenstates since it is not diagonal in the computational basis [17].

QAOA can be thought of as a trotterized version of quantum annealing. While not pure trot-

terization, since the angles are altered for each iteration, QAOA is very similar to adiabatically evolving from the ground state of our mixer Hamiltonian to that of our cost Hamiltonian [15]. One beautiful aspect of it is that with infinite iterations, it is guaranteed to converge to the state that maximizes the cost Hamiltonian. This is analogous to the universal approximation theorem for neural networks. In fact, in many ways, QAOA is reminiscent of classical neural networks.

However, it soon has been discovered that QAOA has significant limitations. Considering noise effects on implementations of QAOA, the algorithm efficiently provides improved approximation ratios when the topology of the machine matches the topology of the graph in the problem. The issue here is that classical algorithms can usually solve these problems significantly more efficiently such as through semi-definite programming [11]. For problems where the topology of the machine does not match that of the problem's graph, even at low noise rates, noisy quantum computers do not exhibit any advantage over classical algorithms [8]. France (2021) concludes that even in a scenario where one has access to thousands of qubits and noise rates sufficient for fault tolerance quantum advantage is "unlikely."

3 Approximation Algorithms: Local Hamiltonian Problem

Due to the limited performance of QAOA, we transition into a broader class of approximation algorithms which attack the local Hamiltonian problem in the context of graphs. We will outline the evolution of these approximation algorithms and share some of our understandings of them and what they could bring.

The problem describes a graph $G = (V, E)$ with the Hamiltonian $H = \sum_{\{i,j\}\in E} h_{ij}$. Its goal is to find the ground state of H, which is a highly interesting problem because of its implications in finding solutions to quantum many-body systems where h_{ij} represent the interaction Hamiltonians. For instance, the Ising model where $h_{int} = -J\sigma_i\sigma_j$ and J is the spin-spin interaction, models spin in material [5]. And one will notice that classical max-cut is a specific application of such a problem where h_{ij} is $\mathbb{I} - Z_iZ_j$, thus the local Hamiltonian problem has more far-reaching implications. But the issue remains that solving this is QMA, which has prompted a variety of algorithms, namely Parkeh's SDP relaxation to product state method, Anshu's shallow quantum circuit applied to general product states, and recently King's combined approach of the two. The progression of such approximation algorithms prompt a discussion of which methods offer the better approach.

3.1 Product State Approx. Algorithms

Brandao and Harrow's results show that product states, which are tensor products of single qubits with no entanglement, can serve as good approximations in the following three specific settings: planar Hamiltonians, dense Hamiltonians, and those with low threshold rank [4]. In the following, we will focus on explaining how product states serve as good candidate solutions in dense

graphs, case two, before discussing how they are found using SDP relaxation.

In Theorem 13 of Ref. [4], the following is stated:

> For every $\epsilon > 0$, there exists a polynomial time algorithm which when given a k-local Hamiltonian, can compute x such that $e_0(H) \leq x \leq e_0(H) + \epsilon$.

It says that to find such an energy within ϵ of the ground state energy $e_0(H)$ is efficiently computable for *dense* graphs, which is a powerful statement to make considering this problem is QMA. Now the pressing questions are why do product states make such good approximations and how are they actually found. The underlying idea is a fundamental property called the monogamy of entanglement.

Monogamy of Entanglement

Monogamy of entanglement states that a system cannot be entangled with many other systems. This means that for dense graphs the relative entanglement is low [4]. Therefore, product states, which contain no entanglement, are not that far from ground state if chosen well.

We term this sharing of entanglement as shareability. In Yang's 2006 paper, a bipartite state, such as the Bell state, is defined as n-shareable if there is a density matrix $\rho_{AB_1B_2\ldots B_n}$ such that $\rho_{AB_1} = \rho_{AB_2} = \ldots = \rho_{AB_n}$ where $\rho_{AB_k} = tr_{B_k}\rho_{AB_1B_2\ldots B_n}$, and he provides an upper bound to n [18]. This statement means that tracing out any one of the systems does not change the resulting density matrix, thereby preserving the shared information between all parties.

This fundamental property of entanglement, which limits n, can be more understood with a toy problem, consisting of three people: Alice, Bob, and Eve, which we will adapt from Ref. [14]. The setup is as follows:

Suppose we want to share the same entangled state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ between Alice and Bob, and Alice and Eve. The common naive proposal is the state $|\psi_{ABC}\rangle = \alpha\,|000\rangle + \beta\,|111\rangle)$. It is true that when we measure one of the qubits, for example Eve's, the state collapses to either $|00\rangle$ or $|11\rangle$ and thus it seems as if they are all entangled with each other due to all three measuring the same state. But the crucial point is that this means Alice and Bob no longer share the entangled state. We can observe this in the context of density matrices, where

$$\begin{aligned}
\rho_{ABE} &= (\alpha\,|000\rangle + \beta\,|111\rangle)(\alpha\,\langle 000| + \beta\,\langle 111|) \\
&= (\alpha^2\,|000\rangle\,\langle 000| + \alpha\beta\,|111\rangle\,\langle 000| + \alpha\beta\,|000\rangle\,\langle 111| + \beta^2\,|111\rangle\,\langle 111|) \\
tr_E(\rho_{ABE}) &= \alpha^2\,|000\rangle\,\langle 000| + \beta^2\,|111\rangle\,\langle 111| \\
&\neq (\alpha\,|00\rangle + \beta\,|11\rangle)(\alpha\,\langle 00| + \langle 11|)
\end{aligned}$$

This can be interpreted in the context of the above definition as 1-shareable because $\rho_{AB} \neq \rho_{AE}$. In taking the trace of a system, Eve, which is entangled to another system, it results in a mixed state, no longer the maximally entangled pure state we desire. So sharing entanglement beyond

two parties is already a challenging task.

This argument can be extended much further using information theory to provide an upper bound n which Yang in Ref. [18] proves as:

$$n \leq N = \lfloor \frac{S(\rho_A)}{G_{\leftarrow}(\rho_{AB})} \rfloor$$

$S(\rho_A)$ describes the von Neumann entropy and we observe that S is zero when ρ_A is a pure state and a nonzero value that grows as ρ_A becomes a maximally mixed state. Thus we observe that the bound loosens as the state shared is less entangled and vice versa, which corresponds with the monogamy property of entanglement. Thus in dense graphs with many particles, the relative entanglement present is low.

Now that we have a general intuition as to why product states may serve as good approximations, we will provide a general sense of how close product states are to ground state. The foundational result in Lieb's 1973 work shows that there exists a product state which has at least 1/9th the energy of the max energy (in our case, we look at ground state but $e_0(H) = e_{max}(-H)$) and recent results from Gossett have suggested the bounds are much tighter at $\frac{1}{3}$ with some product that "is well above that hardness threshold," albeit it is hard to compute exactly [10].

The 1/9th bound was obtained by special properties of the depolarizing channel in which the noisy state outputted has its energy decreased to δ^2 of the original energy where when $\delta = \frac{1}{3}$, the depolarizing channel has a special disentangling property [10]. Thus if we feed in the maximal energy eigenstate, the state disentangles to become a product state with at least 1/9th the energy. Though this bound is quite low, it holds regardless of the density of the graph, and thus the energy of good product states are on the order of the maximal energy, and this bound is expected to be much higher for dense graphs as per the property of monogamy of entanglement.

SDP Relaxation

Existing product state approximation methods then use semi-definite programming (SDP) relaxation to find good product states. SDP solves for parameters in a matrix that preserves the semi definite property of the matrix with the encoded problem. In the classical max-cut problem, the problem can be structured as finding the maximum value of $\sum_{\{i,j\}\in E} \frac{1-X_{ij}}{2}$ where $X \in \mathbb{R}^{n\times n}$, $X = [x_i^T x_j]_{ij}$ such that $X \succeq 0$ [6]. Because the solutions do not exactly yield -1 or 1, the algorithm uses a random hyperplane to cut through a unit sphere with solutions on it, with points on one side designated as -1 and the other as 1 [6]. In the local Hamiltonian problem, the process is similar and the value over which we are maximizing is based on the Hamiltonian of the graph. The crucial question remains: How is the solution rounded to a product state?

In Parekh's 2019 paper, the high level interpretation of the setup is as follows. First a solution matrix $M \in \mathbb{R}^{n \times n}$ where n denotes the number of vertices in the graph is found using SDP. The goal is to transform the matrix into $3n$ parameters which describe the coefficients of the Pauli matrices for the single qubit density matrix $\rho_i = (I + \alpha X + \beta Y + \gamma Z)/2$, all three parameters if the problem is quantum Max-Cut [9]. Then a randomly normally distributed matrix about zero is applied to vector entries of M, outputting a vector corresponding to the parameters α, β, and γ.

Although these product state approximation algorithms are natural to think about and employ advanced classical algorithms such as SDP, they are plagued by an upper bound on their approximation ratios, which are 2/3 for the XY Model and 1/2 for the quantum Heisenberg Model (Quantum Max-Cut) [9]. Thus it has prompted a different approach, where instead of using classical algorithms to provide quantum solutions one uses quantum circuits to improve an ansatz.

3.2 Shallow depth circuit Approx. Algorithms

In a new set of approximation algorithms, Anshu, Gossett, and Morenz describe a procedure exceeding the theoretical upper bound of the approximation ratio 1/2 without needing SDP or a density requirement, which in that sense makes it a more general and powerful algorithm [2]. Instead it relies on applying a shallow depth quantum circuit on not a good product state, but any product state. Their main results can be summarized by Theorem 4 of Ref. [2] which states that for a k-regular graph, there exists an efficiently computable $(k+1)$ depth unitary U such that $U(G)\,|0^n\rangle$ provides a better solution than any product state (including the maximal energy product state).

Improving General Product States

Anshu further develops his proposed algorithm in the following year, where he shows that there is am efficiently computable unitary U such that when applied to an initial product state its energy improvement is proportional to the number of edges [3].

The setup to their results is as follows:

Define $G = (V, E)$ of maximum degree $d \geq 2$ to be a graph with the Hamiltonian $H = \sum_{\{i,j\} \in E} h_{ij}$. G can be generalized to a d-regular graph whereby adding some local terms $h_{ij} = 0$ can become any graph. And let $|v\rangle$ be some product state given by $|v_1\rangle \otimes |v_2\rangle \otimes \ldots \otimes |v_n\rangle$.

We will be focusing on Anshu's first main two theorems, stated as the following:

Theorem 1: Given a product state $|v_i\rangle$, we can efficiently compute a depth-$(d+1)$

quantum circuit U such that the state $|\psi\rangle = U|v\rangle$ satisfies:

$$\langle\psi| H |\psi\rangle \geq \langle v| H |v\rangle + \Omega\left(\frac{Var_v(H)^2}{d^2|E|}\right)$$

Theorem 2: Let $|v\rangle$ be a product state and $|\psi\rangle = V(\theta)|v\rangle$ be the state prepared by the quantum circuit. Define the positive real parameter α by:

$$\alpha = \mathbb{E}_{\{i,j\}\in E} |\langle v_i, v_j| [P_i P_j, h_{ij}] |v_i, v_j\rangle|$$

There is an efficient classical algorithm to select parameters θ satisfying

$$\langle\psi| H |\psi\rangle \geq \langle v| H |v\rangle + \Omega\left(\frac{|E|\alpha^2}{d}\right)$$

The overall picture of these two theorems is that improvements from any product state is achievable and is proportional to how far the state is from the maximal energy state due to the dependence on the variance of H in the first theorem [1].

In the following, our goal is to lay out Anshu's proofs, which were provided in the supplemental material [3], rederiving the stated rigorous lower bound and demonstrating that existing-algorithms based on tensor products can be improved under certain conditions.

First we define $P_1, P_2, \ldots, P_n$ as the collection of single-qubit operators such that $1 \geq \|P_i\|$ and $\langle v_k| P_k |v_k\rangle = 0$ for all k, v_i is the vector component, h_{ij} are the Hamiltonian components, and $\mathbb{E}\{i,j\}$ is the expectation value. Interestingly, Anshu uses the result of Theorem 2 in his proof for Theorem 1, so we delve into 2 first.

Proof of Theorem 2:

$$\langle\psi| H |\psi\rangle \geq \langle v| H |v\rangle + \Omega\left(\frac{|E|\alpha^2}{d}\right)$$

Define U as $V(\vec{\theta}) = \bigotimes_{\{i,j\}\in E} e^{i\theta_{ij}P_iP_j} = e^{i\sum_{\{i,j\}\in E}\theta_{ij}P_iP_j}$. Observe that by varying $\vec{\theta}$ we are not affecting the right-hand side of the inequality – we are purely changing the original state $|\psi\rangle$ in order to establish a good upper bound for our energy expectation value. Now considering the energy of a single element of the Hamiltonian H:

$$\langle\psi| h_{ij} |\psi\rangle = \langle v| V(\vec{\theta})^\dagger h_{ij} V(\vec{\theta}) |v\rangle$$

This simply uses the definition of $|\psi\rangle$ to expand into $V(\theta)|v\rangle$. The quantum circuit $V(\theta)$ is made

up of gates associated with the edges of our graph. As we are only considering a single component h_{ij} of our Hamiltonian, all gates with edges not incident to $\{i, j\}$ commute. The energy is then rewritten as $\langle v|V_{ij}^{\dagger}h_{ij}V_{ij}|v\rangle$ where $V_{ij} = \prod_{\{k,\ell\}\in N_{ij}} e^{i\theta_{k\ell}P_kP_\ell}$.

The crucial point is that now we can now represent the expectation value $\langle\psi| H |\psi\rangle$ over $|\psi\rangle$ as the *difference* of the expectation value of the product state $\langle v| H |v\rangle$ and the states over the edges not included in the maximal energy eigenstate. These combinations of all those unwanted states can be expressed using an m-nested commutator. The equation then becomes:

$$\langle\psi| h_{ij} |\psi\rangle = \langle v| h_{ij} |v\rangle + \sum_{m=2}^{\infty} \frac{i^m}{m!} \langle v| \left[\sum_{\{k,\ell\}\in N_{ij}} -\theta_{k\ell}P_kP_\ell, h_{ij} \right]_m |v\rangle$$

Note that the negative sign is inside the summation. Knowing that $\langle v_k| P_k |v_k\rangle$ by construction, the expression is simplified for the $m = 1$ case and an upper bound is set for the $m > 1$ cases. This yields:

$$m = 1 \qquad \sum_{\{k,\ell\}\in N_{ij}} -i\theta_{k\ell} \langle v| [1] |v\rangle = -i\theta_{ij} \langle v| [P_iP_j, h_{ij}] |v\rangle$$

$$m > 1 \qquad \left| \langle v| \left[\sum_{\{k,\ell\}\in N_{ij}} -\theta_{k\ell}P_kP_\ell, h_{ij} \right]_m |v\rangle \right| \leq \sum_{\substack{\{k_1,\ell_1\},\{k_2,\ell_2\},\ldots \\ \{k_m,\ell_m\}\in N_{ij}}} \theta^m |\langle v| [1] |v\rangle|$$

These are some very intimidating equations, but we can break the meaning down. We see that for the $m = 1$ case, the summation of the unwanted terms simplifies to all combinations of edges in N_{ij}. This is expressed notationally as θ_{ij} and later as θ. For $m > 1$, the upper bound is set as the magnitude of our subtraction term. Thus we examine the net energy of the terms we are throwing out and observe it must be less than or equal to the summation of energy of m sets of distinct edges in N_{ij}. More clearly, they select edges such that no vertex appears exactly once. Once again, the yet-to-be determined θ is utilized to simplify the terms:

Then Claim 1, stated here, is used to establish an upper bound on the $m > 1$ case:

$$\theta^m |\langle v| [P_{k_1}P_{\ell_1}, [\ldots, [P_{k_m}P_{\ell_m}, h_{ij}]]] |v\rangle| \leq (2\theta)^m$$

Consequently, this means that that sum of all terms in the $m > 1$ case of the unwanted edge energies can be set as:

$$\sum_{m=2}^{\infty} \frac{1}{m!}(4m\sqrt{d})^m\theta^m \leq \sum_{m=0}^{\infty}(4e\sqrt{d}\theta)^{m+2} \leq 32e^2d\theta^2$$

Here an upper bound of $m^m/m! \leq e^m$ is utilized and assuming that the choice of θ such that

$\theta \leq \frac{1}{8e\sqrt{d}}$ is true, everything put together yields the final result:

$$\langle\psi| H |\psi\rangle \geq \langle v| H |v\rangle + |E|(\theta\alpha - 32e^2 d\theta^2)$$

Now θ may be efficiently computed to achieve a good lower bound using the equation $(\theta\alpha - 32e^2 d\theta^2)$. In summary, given a desired wave function, we can represent the Hamiltonian's expectation value as the sum of that Hamiltonian acting on a product state approximation across all edges minus the sum of the Hamiltonian acting on the product state approximation across some unwanted edges. Those unwanted edges are defined precisely as the non-interacting terms for any individual Hamiltonian term h_{ij}. Setting an upper bound on the value of those unwanted term's expectation value in turn gives us a lower bound on our energy expectation value.

Proof of Theorem 1

$$\langle\psi| H |\psi\rangle \geq \langle v| H |v\rangle + \Omega\left(\frac{Var_v(H)^2}{d^2|E|}\right)$$

The necessary and sufficient conditions for this to be true are:

(1) Our variance of energy satisfies $Var_v(H) = \Omega(|E|)$ where $Var_v(H) = \langle v| H^2 |v\rangle - \langle v| H |v\rangle^2$

(2) The state $|v\rangle$ is a product state.

Observe that Condition I is not satisfied in a purely classical case, as $Var_v(H) = 0$ due to $|v\rangle$ being an eigenstate of H in the classical computational basis. Condition I is, however, satisfied in the quantum setting. Toy examples from Anshu et. al. show that neither condition on its own is sufficient to guarantee the existence of a state with better approximation than $|v\rangle$ obtained from a typical variational method.

Before proceeding, *Proposition 1* was defined as follows:

> Let $Q2$ be the projector onto computational basis states with Hamming weight 2. We can efficiently choose operators $\{P_i\}$, with i being elements of V, such that $\alpha \geq \frac{1}{|E|}\langle 0^n| HQ_2H |0^n\rangle$.

We'll take this for granted here, since a detailed proof on why this is is provided by Anshu. Proceeding, Q_t is set as the projector onto computational basis states with Hamming weights t being either 1 or 2. Because of this, the variance of H is set to:

$$Var(H) = \langle 0^n|HQ_1H|0^n\rangle + \langle 0^n|HQ_2H|0^n\rangle$$

where the variance can be viewed as the distance from Q_1 and Q_2. Anshu then observes that since

both terms must be greater than or equal to 0, $\langle 0^n| HQ_tH |0^n\rangle \geq Var(H)/2$. The following analysis is split into cases, one where the Hamming weight $t = 1$ and the other $t = 2$

Starting with $t = 2$, Proposition 1 is used to obtain an upper bound on $Var(H)$. Observe that $2|E| \max\{\alpha_1, \alpha_2\} \geq \text{Var}(H)$ is an immediate consequence, obtained by multiplying both sides with the maximal energy expectation value and applying the upper bound on Var(H). Plugging this into Theorem 2:

$$\begin{aligned}\langle\psi| H |\psi\rangle &\geq \langle v| H |v\rangle + \Omega\left(\frac{|E|\alpha^2}{d}\right) \\ &\geq \langle v| H |v\rangle + \Omega\left(\frac{|E|(Var(H)/2|E|)^2}{d}\right) \\ &\geq \langle v| H |v\rangle + \Omega\left(\frac{Var(H)^2}{d|E|}\right)\end{aligned}$$

Then for $t = 1$, the math is much more intensive but ultimately results in a similar lower bound improvement but scaled with $1/d^2$ rather than $1/d$ in the $t = 2$ case:

$$\langle\psi| H |\psi\rangle \geq \langle v| H |v\rangle + \Omega\left(\frac{Var(H)^2}{d^2|E|}\right)$$

Since this improvement is the smaller of the two, it serves as the lower bound. It seems like the $t = 1$ case yields a worse lower bound because there are more degrees of freedom to account for and thus tuning the parameters is more difficult.

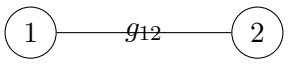
3.3 Combining the Two Algorithms

The natural question after those two methods was: what if we could combine the two strengths of each to yield an even better approximation? And this is what King did shortly after in '22 in which he combined aspects of the two previous methods to achieve an approximation ratio of $1/\sqrt{2} \approx 0.7$ when applied to the EPR Hamiltonian [13]. The general approach consists of two main parts, first using SDP relaxation to round to a good product state and then applying some commuting entangling gates to introduce global entanglement into the product state [13].

King's Improved Approximation Algorithm

To understand how these sequence of steps could bring us to the maximal energy eigenstate, we interpret the toy problem in the context of the EPR Hamiltonian given in King's '22 paper:

Consider the simplest possible graph with two vertices and one edge connecting them like so:



The Hamiltonian g_{12} describing the edge is given by $g = \frac{1}{2}(\mathbb{I} + X_1 \otimes X_2 - Y_1 \otimes Y_2 + Z_1 \otimes Z_2)$. The "good" product state would be $|0_1 0_2\rangle$ in this scenario where:

$$\begin{aligned}\langle 00| g_{12} |00\rangle &= \langle 00| \frac{1}{2}(\mathbb{I} + X_1 \otimes X_2 - Y_1 \otimes Y_2 + Z_1 \otimes Z_2) |00\rangle \\ &= \langle 00| (|00\rangle + |11\rangle + |11\rangle + |00\rangle)/2 \\ &= 1\end{aligned}$$

In the context of more complicated graphs, this good product state $|00\rangle$ would be found using SDP relaxation, which was detailed in the first generation of these approximation algorithms. But we know that $|00\rangle$ is not the maximal energy state. Instead it is an entangled state, more specifically the Bell state $|\psi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ as we can see here:

$$\begin{aligned}\langle \psi| g_{12} |\psi\rangle &= |\psi\rangle (4 |00\rangle + 4 |11\rangle)/2\sqrt{2} \\ &= (\langle 00| + \langle 11|)(|00\rangle + |11\rangle) \\ &= 2\end{aligned}$$

So there needs to be some entanglement in this product state, which is expected in most other Hamiltonians, which deal with complex interactions. Thus King applies some commuting entangling gates on the product state, such as $|00\rangle$ in this case, to gain energy improvements. This is done by choosing a superposition of Pauli operators:

$$P_k = \frac{X_k - Y_k}{\sqrt{2}}$$

where when applied to our edge $\{1, 2\}$ becomes:

$$e^{i\theta_{ij} P_1 \otimes P_2} |00\rangle = \cos\theta_{ij} |00\rangle + \sin\theta_{ij} |00\rangle$$

and the optimal angle to rotate $|00\rangle$ such to achieve the maximal energy is $\theta = \pi/4$

$$= \frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

This parallels the second part of King's process, which is applying specific gates onto the product states and optimizing over θ to maximize energy. The second portion is where King's method shines over product state approximations, especially seen in the toy example where $|00\rangle$ yields an energy $1/2$ that of the maximal energy, because in most cases, we can expect some level of entanglement in these product states. Though this improvement may not be as significant as graphs become more dense. Furthermore, his results yield an approximation ratio of roughly 0.582 on triangle-free graphs, higher than that of Anshu's 0.53 [13]. But King was unable to prove this for the general case, so there is reason to suspect that the approximation ratio has not been pushed beyond this limit. But King's idea is an interesting one and represent the hybrid quantum-classical algorithms

that characterize today's NISQ era.

4 Conclusion

We have introduced the various approximation algorithms throughout the years, starting with Brandao and Harrow's work with product state approximations to Anshu's shallow depth circuit application to a more recent work by King which draws inspiration from both. Product state approximations haved reached their peak in utility, especially because of their relatively low theoretical bound of 0.5. Thus new methods are being developed to in a sense, use quantum to solve these complex problems rather than solely sticking to classical algorithms. Quantum circuits have the nice advantage of being able to introduce entanglement and easily transform product states. And a nice figure below captures the evolution of the landscape of these algorithms and how they have improved steadily:

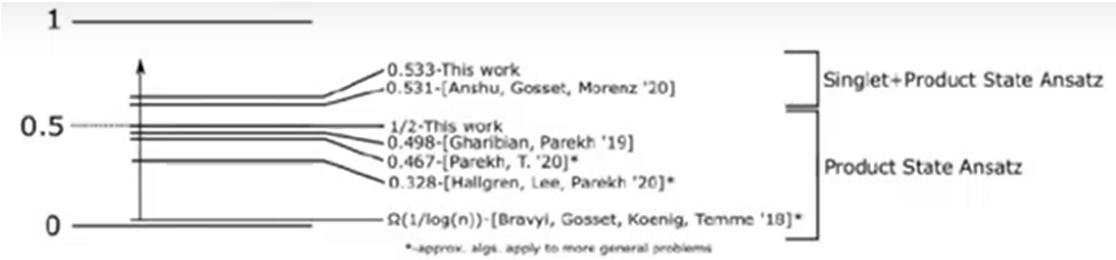


Figure 1: Approximation ratios of algorithms applied to quantum Max-Cut[16]

But we observe that there has not been significant improvement beyond 0.5, so there is much more needed progress ahead of us if we ever want to utilize these approximation algorithms to solve some of maths', physics', and chemistry's most challenging problems.

5 Acknowledgments

We would like to thank Prof. Anurag Anshu for his feedback on the Improved Approximation Algorithms presented in this paper, along with Yunchao Liu for providing us direction to pursue this topic, Andris Huang for evaluating and reviewing the structure of this paper, and all the course staff for helping us understand the material through this C191 Quantum Information course.

References

[1] Anurag Anshu. *"Improved Approximation Algorithms for Bounded-degree Local Hamiltonians"*. URL: https://www.youtube.com/watch?v=aFVnWq3RHYU.

[2] Anurag Anshu, David Gosset, and Karen Morenz. "Beyond Product State Approximations for a Quantum Analogue of Max Cut". en. In: Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. DOI: 10.4230/LIPICS.TQC.2020.7. URL: https://drops.dagstuhl.de/opus/volltexte/2020/12066/.

[3] Anurag Anshu et al. "Improved Approximation Algorithms for Bounded-Degree Local Hamiltonians". In: *Physical Review Letters* 127.25 (Dec. 2021). DOI: 10.1103/physrevlett.127.250502. URL: https://doi.org/10.1103%2Fphysrevlett.127.250502.

[4] Fernando G.S.L. Brandao and Aram W. Harrow. "Product-State Approximations to Quantum Ground States". In: New York, NY, USA: Association for Computing Machinery, 2013. ISBN: 9781450320290. DOI: 10.1145/2488608.2488719. URL: https://doi.org/10.1145/2488608.2488719.

[5] Jeffrey Chang. *The Ising Model*. URL: https://stanford.edu/~jeffjar/statmech/intro4.html.

[6] Visually Explained. *Goemans-Williamson Max-Cut Algorithm — The Practical Guide to Semidefinite Programming (4/4)*. URL: https://www.youtube.com/watch?v=aFVnWq3RHYU.

[7] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. *A Quantum Approximate Optimization Algorithm*. 2014. arXiv: 1411.4028 [quant-ph].

[8] Daniel Stilck França and Raul García-Patrón. "Limitations of optimization algorithms on noisy quantum devices". In: 17 (Oct. 2021). DOI: 10.1038/s41567-021-01356-3. URL: https://doi.org/10.1038/s41567-021-01356-3.

[9] Sevag Gharibian and Ojas Parekh. "Almost Optimal Classical Approximation Algorithms for a Quantum Generalization of Max-Cut". en. In: (2019). DOI: 10.4230/LIPICS.APPROX-RANDOM.2019.31. URL: http://drops.dagstuhl.de/opus/volltexte/2019/11246/.

[10] David Gosset. *Approximation algorithms for quantum many-body problems*. URL: https://www.youtube.com/watch?v=ReQOLDWjiLk.

[11] F. Hadlock. "Finding a Maximum Cut of a Planar Graph in Polynomial Time". In: *SIAM Journal on Computing* 4.3 (1975), pp. 221–225. DOI: 10.1137/0204019. eprint: https://doi.org/10.1137/0204019. URL: https://doi.org/10.1137/0204019.

[12] Julia Kempe, Alexei Kitaev, and Oded Regev. *The Complexity of the Local Hamiltonian Problem*. 2005. arXiv: quant-ph/0406180 [quant-ph].

[13] Robbie King. *An Improved Approximation Algorithm for Quantum Max-Cut*. 2022. arXiv: 2209.02589 [quant-ph].

[14] TUDelft OpenCourseWare. *TUDQCXQC2016-V005600*. URL: https://www.youtube.com/watch?v=pFK-3SNN2kM.

[15] Michal Stecly. *Quantum Approximate Optimization Algorithm explained*. Disqus. May 2020.

[16] Kevin Thomson. *QIP 2022 — Application of the Level-2 Quantum Lassere Hierarchy in Quantum Approximation Algorithms*. URL: https://www.youtube.com/watch?v=TsL2I-_z8nE.

[17] Isaac De Vlugt. *QAOA: A different perspective — PennyLane Tutorial*. URL: https://www.youtube.com/watch?v=VTFJjDYXPOA.

[18] Dong Yang. "A simple proof of monogamy of entanglement". In: *Physics Letters A* 360.2 (Dec. 2006), pp. 249–250. DOI: 10.1016/j.physleta.2006.08.027. URL: https://doi.org/10.1016%2Fj.physleta.2006.08.027.