

Optimal Clifford Initial States for Ising Hamiltonians

Bikrant Bhattacharyya

Illinois Mathematics and Science Academy

bbhattacharyya@imsa.edu

Gokul Subramanian Ravi

University of Michigan

gsravi@umich.edu

Abstract—Modern day quantum devices are in the NISQ era, meaning that the effects of size restrictions and noise are essential considerations when developing practical quantum algorithms. Recent developments have demonstrated that Variational Quantum Algorithms (VQAs) are an appropriate choice for current era quantum devices. VQAs make use of classical computation to iteratively optimize parameters of a quantum circuit, referred to as the ansatz. These parameters are usually chosen such that a given objective function is minimized. Generally, the cost function to be minimized is computed on a quantum device, and then the parameters are updated using a classical optimizer. One of the most promising VQAs for practical hardware is the Variational Quantum Eigensolver (VQE). VQE uses the expectation value of a given Hamiltonian as the cost function, and the minimal value of this particular cost function corresponds to the minimal eigenvalue of the Hamiltonian.

Because evaluating quantum circuits is currently very noisy, developing classical bootstraps that help minimize the number of times a given circuit has to be evaluated is a powerful technique for improving the practicality of VQE. One possible such bootstrapping method is creating an ansatz which can be efficiently simulated on classical computers for restricted parameter values. Once the optimal set of restricted parameters is determined, they can be used as initial parameters for a VQE optimization which has access to the full parameter space. Stabilizer states are states which are generated by a particular group of operators called the Clifford Group. Because of the underlying structure of these operators, circuits consisting of only Clifford operators can be simulated efficiently on classical computers. Clifford Ansatz For Quantum Algorithms (CAFQA) is a proposed classical bootstrap for VQAs that uses ansatzes which reduce to clifford operators for restricted parameter values [1].

CAFQA has been shown to produce fairly accurate initialization for VQE applied to molecular Hamiltonians. Motivated by this result, in this paper we seek to analyze the Clifford states that optimize the cost function for a new type of Hamiltonian, namely Transverse Field Ising Hamiltonians. Our primary result is contained in theorem IV.1 which connects the problem of finding the optimal CAFQA initialization to a submodular minimization problem which in turn can be solved in polynomial time.

I. INTRODUCTION

Quantum computing (QC) represents a groundbreaking computational paradigm for solving specific problems that are traditionally impractical to tackle classically. It is anticipated that quantum computers (QCs) will wield significant advantages in areas of critical impact such as cryptography [2], chemistry [3], optimization [4], and machine learning [5].

During the ongoing Noisy Intermediate-Scale Quantum (NISQ) era, we are set to operate with quantum machines

equipped with hundreds to thousands of imperfect qubits [6]. In this era, these machines will grapple with limited connectivity and relatively short qubit lifetimes due to design constraints. Noise remains a major obstacle, preventing current quantum computers from outperforming classical computers in nearly all applications. Machines in the NISQ era will be incapable of executing extensive quantum algorithms like Shor Factoring [2] and Grover Search [7]. These algorithms necessitate error correction involving millions of qubits to establish fault-tolerant quantum systems [8]. However, a range of error mitigation approaches [9]–[16] have been proposed, enhancing execution fidelity on today’s quantum devices. Nonetheless, the achieved fidelity still falls short of the requirements for most practical applications.

In recent times, there has been a growing focus on leveraging classical computing support to elevate the applicability of NISQ applications and devices in the real world. This effort encompasses various enhancements, such as optimizations at the compiler level [17], advancements in classical optimizers [18], circuit segmentation with classical compensation [19], [20], among others. We are currently in the early stages of exploring this synergistic quantum-classical paradigm. There exists significant potential for employing sophisticated classical bootstrapping tailored to specific applications, thereby pushing the boundaries of NISQ capabilities forward.

Variational quantum algorithms (VQAs) are anticipated to align well with NISQ machines, demonstrating a broad array of applications, including electronic energy estimation for molecules [21] and approximations for MAXCUT [4]. The quantum circuit in a VQA is defined by a set of angles, which are fine-tuned by a classical optimizer across multiple iterations to reach a specific target objective representing the VQA problem. These algorithms exhibit greater suitability for present-day quantum devices due to their ability to adapt to the idiosyncrasies and noise profile of the quantum machine [21], [22]. Regrettably, the accuracy achieved by VQAs on existing NISQ machines, even with error mitigation strategies, frequently falls significantly short of the exacting accuracy demands, particularly in domains like molecular chemistry, especially when dealing with larger problem sizes [3], [13], [17].

To advance NISQ VQAs towards practical utility, it is crucial to carefully select the parameterized circuit (ansatz)

for a VQA and optimize its initial parameters classically to be as close to optimal as possible before venturing into quantum exploration. This approach holds the promise of enhancing accuracy and expediting algorithmic convergence on the quantum device, even in the presence of noise [23], [24]. While certain applications may derive advantages from domain-specific knowledge guiding the choice of particular parameterized circuits and initial parameters (e.g., UCCSD [25]), these choices are less appropriate for execution on contemporary quantum devices due to their substantial quantum circuit depth. Ansatz circuits tailored for today’s devices, often termed as “hardware efficient ansatz” [3], are typically agnostic to specific applications and stand to benefit significantly from judiciously chosen initial parameters. However, accurately estimating these parameters through classical means can be a challenging task.

Prominent prior work like CAFQA [1] focuses on initializing the VQA ansatz through classical simulation. In CAFQA, the initial parameters for VQA are selected through an efficient and scalable search within a classically simulable segment of the quantum space known as the Clifford space, employing Bayesian Optimization. CAFQA attains remarkable accuracy during initialization. Specifically, for the crucial chemistry application involving the estimation of the ground state energy of molecules, CAFQA restores up to 99.99% of the accuracy lost in previous state-of-the-art classical initialization, demonstrating mean improvements of 56x.

Motivated by this result, in this paper we seek to analyze the Clifford states that optimize the cost function for a new type of Hamiltonian, namely Transverse Field Ising Hamiltonians. Our primary result is contained in theorem IV.1 which connects the problem of finding the optimal Clifford initialization to a submodular minimization problem which in turn can be solved in polynomial time. This connection arises by mapping the problem of finding the Hamiltonian ground state to a graph theoretic optimization problem, which is submodular. Submodular functions satisfy a criterion formalized in Appendix B-C. As a result of this property, submodular functions can be minimized by considering related constrained convex optimization problems [26]. When running numerical experiments, we consistently obtain Clifford approximations of the ground state energies that have a relative error varying from roughly 0-25% when compared to the exact ground state energies. This error itself is problem dependent and is an indicator of the true ‘quantum-ness’ of the problem. Once the good Clifford initialization is found, VQE can be run on future quantum devices (with reasonably low error rates), to find (nearly) exact ground state energy estimates.

II. BACKGROUND

Here, we provide a brief overview of VQE. For a more in-depth overview of VQE see [27]. The *variational principle* from quantum mechanics is a statement regarding the expectation value of a Hamiltonian. Given a Hamiltonian H with ground state E_0 , any quantum state $|\psi\rangle$ will satisfy

$$E_0 \leq \langle \psi | H | \psi \rangle \quad (1)$$

Thus if the state can be parameterized by a vector parameter angles $\theta \in [0, 2\pi]^k$, then

$$E_0 \approx \min_{\theta \in \mathbb{R}^k} \langle \psi(\theta) | H | \psi(\theta) \rangle \quad (2)$$

Assuming that the parameterization of $|\psi\rangle$ is expressive enough to accurately predict the ground state of the Hamiltonian.

A common way to parameterize $|\psi\rangle$ is by using a parameterized quantum circuit which implements some parameterized unitary operator $U(\theta)$ acting on the initial state $|0\rangle^{\otimes n}$ where n is the number of qubits.¹ Using this parameterization,

$$|\psi\rangle = U(\theta)|0\rangle \quad (3)$$

and

$$\langle \psi | H | \psi \rangle = \langle 0 | U^\dagger(\theta) H U(\theta) | 0 \rangle \quad (4)$$

The circuit which implements the unitary operator U is called the *ansatz*, and the choice of *ansatz* is critical to the performance of VQAs [28].

The *Variational Quantum Eigensolver* (VQE) begins with some set of parameters θ_i and some ansatz $U(\theta)$. It then repeatedly computes the expectation value of the Hamiltonian on $U(\theta)|0\rangle$ using a quantum circuit derived from the ansatz. Between each computation, the parameters are updated using a classical optimizer. The benefit of using quantum devices to find expectation values is that the gate depth of such circuits (both the gates required to construct the ansatz and those required to compute the expectation value once the ansatz has been applied) grows polynomially with n whereas using purely classical simulation will require matrices of size $2^n \times 2^n$ at minimum.

CAFQA’s primary proposal is to make use of an ansatz which lends itself to a classical initial search before requiring the usage of a quantum computer [1]. It has been shown that circuits which are composed entirely of *Clifford Gates* can be simulated in polynomial time using the *stabilizer technique* [29] [30]. The states obtained from applying Clifford Gates are called *Stabilizer states* or *Clifford states*.²

From here on, we will assume that the ansatz under consideration will be able to reach every n -qubit clifford state. With this assumption, determining the best CAFQA initial value is the same as determining the best clifford state initial value.

From here on

- $|\psi\rangle$ will represent an arbitrary state
- $|\varphi\rangle$ will represent an arbitrary Clifford state.
- $|\psi_0\rangle$ will be the ground state for a given Hamiltonian H
- $|\varphi_0\rangle$ will represent the Clifford state that minimizes $\langle \varphi | H | \varphi \rangle$ over all Clifford states $|\varphi\rangle$.

III. MATHEMATICAL PREREQUISITES

Before moving onto determining the optimal clifford states for a given Hamiltonian, we require the development of some prerequisite theory.

¹From here on we use the shorthand $|0\rangle^{\otimes n} = |0\rangle$. Generally, context will make it clear how many qubits we are considering.

²We will use these two terms interchangeably.

A. Graph Theory

In this section we review the required graph theory concepts the notations that we will use. For a fuller overview of graph theory alongside proofs of some of the corollaries left unproven here, see [31].

Definition III.1. A *graph* G is given by a set of vertices V_G and a set of edges E_G . We write this as $G = (V_G, E_G)$.³

The particular graphs that we will use for computations later on are defined in Appendix A. For our purposes, $V_G = \{q_0, q_1, q_2, \dots, q_{N-1}\}$ with $N \geq 2$. This labelling provides a natural correspondence between node q_i and qubit q_i . An edge between nodes q_i and q_j will be denoted as $\langle q_i, q_j \rangle = \langle q_j, q_i \rangle$ ⁴. Furthermore, we only consider graphs with edges between distinct nodes.

Definition III.2. A *subgraph* $S = (V_S, E_S)$ of a given graph G is a new graph satisfying $V_S \subset V_G$ and $E_S \subset V_E$. We write $S \subset G$ to say that S is a subgraph of G .

There's a special type of subgraph that will play a special role later on.

Definition III.3. Given a set of nodes $\{q_{i_1}, q_{i_2}, \dots\}$ of a graph G , the subgraph S *induced* by $\{q_{i_1}, q_{i_2}, \dots\}$ is the subgraph with $V_S = \{q_{i_1}, q_{i_2}, \dots\}$ and E_S equal to the set of all edges in V_G between nodes in $\{q_{i_1}, q_{i_2}, \dots\}$. We will write $E(S)$ to denote the edges induced by vertex set S .

Given a graph G and two nodes q_i and q_j , we can define whether or not q_i or q_j are connected as follows.

Definition III.4. Nodes q_i and q_j are *G-connected* if either of the two following statements are true

- $\langle q_i, q_j \rangle \in E_G$
- There exists a sequence of nodes $q_{k_1}, q_{k_2}, \dots, q_{k_m}$ such that

$$\langle q_i, q_{k_1} \rangle, \langle q_{k_1}, q_{k_2} \rangle, \dots, \langle q_{k_m}, q_j \rangle \in E_G$$

A graph G is called *connected* if every element of V_G is G -connected to every other element in V_G .

Corollary III.5. The following two immediately follow from Definition III.4 and Definition.

- q_i being G -connected to q_j is equivalent to q_j being G -connected to q_i (reflexivity).
- If q_i is G -connected to q_j and q_j is G -connected to q_k then q_i is G -connected to q_k (transitivity).

Corollary III.6. Every graph G is the union of $k \geq 1$ disjoint connected subgraphs called *connected components*. In other words, there exists a set of graphs $C_1, C_2, \dots, C_k \subset G$ such that

- $V_G = \bigcup_{i=1}^k V_{C_i}$

³Graphs will always be denoted with capital letters.

⁴In general graph theory it's often useful to assign a real number weight to each edge. Furthermore, there are times where $\langle q_i, q_j \rangle \neq \langle q_j, q_i \rangle$ is a more natural choice. However, here we consider unweighted and undirected graphs, neglecting these alternate definitions.

- $E_G = \bigcup_{i=1}^k E_{C_i}$
- Each C_i is connected
- $V_{C_i} \cap V_{C_j} = \emptyset$ if $i \neq j$

Connectivity allows us to define spanning trees as follows.

Definition III.7. Given a connected graph S , we call any connected subgraph $M \subset S$ a *spanning tree* of S if M is connected and satisfies the following properties.

- $V_M = V_S$
- If $\langle q_i, q_j \rangle \in E_M$, then there exists no sequence $q_{k_1}, q_{k_2}, \dots, q_{k_m}$ with

$$\langle q_i, q_{k_1} \rangle, \langle q_{k_1}, q_{k_2} \rangle, \dots, \langle q_{k_m}, q_j \rangle \in E_G$$

If S is disconnected but the union of disjoint connected subgraphs C_i , then we define a spanning tree on S as

$$M = \left(V_S, \bigcup_i E_{M_i} \right) \quad (5)$$

Where M_i is a spanning tree for C_i .⁵

The following facts about spanning trees are well known [32].

Lemma III.8. If M is a spanning tree of a graph S with k connected components, then $|E_M| = |V_S| - k$. Furthermore, every graph S contains at least one spanning tree M .

Given a graph $S = (V_S, E_S)$, there is an algorithm that runs in $O(|V_S| + |E_S|)$ that can always find the spanning tree for S .

The following function will be a useful theoretical tool later on.

Definition III.9. Given a graph G , let the sequence of sets \mathcal{Q}_n for $2 \leq n \leq N$ be defined as⁶

$$\mathcal{Q}_n = \{S \in P(V_G) \text{ such that } |S| \leq n\} \quad (6)$$

We define the following function as the *edge-function*

$$\mathcal{E}(n) = \max_{S \in \mathcal{Q}_n} |E(S)| \quad (7)$$

For completeness we define $\mathcal{E}(0) = \mathcal{E}(1) = 0$.

If S is a set of vertices that maximizes $|E(S)|$ over \mathcal{Q}_n , we can S an n -optimal vertex set.

Notice that for fixed n , there may be multiple n -optimal vertex sets. For completeness, we make the empty set the 0-optimal vertex set and $\{q_k\}$ for $0 \leq k < N$ the 1-optimal vertex sets.

B. Stabilizer States

The result of a Clifford Circuit will always be a stabilizer state, which can be equivalently defined as below:

Definition III.10. A *n-qubit stabilizer state* $|\varphi\rangle$ is a state which has a stabilizer group consisting of 2^n elements of the n -qubit Pauli group.

⁵Sometimes spanning trees for disconnected graphs are called spanning forests. Here we will not distinguish them.

⁶We use $P(S)$ to denote the power set of S .

In other words, every operator O satisfying $O|\varphi\rangle = |\varphi\rangle$ is an element of the n -qubit Pauli Group and there are 2^n such operators.

It's well known that the stabilizer group for state $|\varphi\rangle$, which will be denoted by \mathcal{S}_φ can always be generated by n operators as described in the Theorem below.

Theorem III.11. For any stabilizer state $|\varphi\rangle$, \mathcal{S}_φ can be written as $\langle \mathcal{G}_\varphi \rangle$ where $\mathcal{G}_\varphi = \{O_1, O_2, \dots, O_n\}$ consists of n elements of the n -qubit Pauli group, each of which satisfies

- $O_i \notin \langle \mathcal{G}_\varphi \setminus O_i \rangle$.
- $[O_i, O_j] = 0$ if $i \neq j$.

This theorem has the immediate corollary

Corollary III.12. Given qubit indices i, j, k with $i \neq j$, $\pm Z_i Z_j$ and $\pm X_k$ cannot simultaneously be in \mathcal{G}_φ for any state $|\varphi\rangle$ if $k \in \{i, j\}$.

Given distinct indices l and s , no sequence of operators of the form $Z_l Z_{k_1}, Z_{k_1} Z_{k_2}, Z_{k_2} Z_{k_3}, \dots, Z_{k_m} Z_s$ can be in \mathcal{G}_φ if $Z_l Z_s$ is and vice versa.⁷

The other well known result that will be used is the following.

Theorem III.13. Given an operator O in the n -qubit Pauli group and a n -qubit stabilizer state $|\psi\rangle$, the expectation value $\langle \psi | O | \psi \rangle$ is always 0, ± 1 , or $\pm i$. Furthermore, if O is Hermitian, then the only possible expectation values are 0 and ± 1 .

Proofs of Theorems III.11 and III.13 can be found in [29]. The following lemma is inspired by III.13.

Lemma III.14. If $|\chi\rangle$ is stabilized by X_j and $|\varphi\rangle$ is stabilized by $Z_i Z_j$, then

- $\langle \chi | Z_i Z_j | \chi \rangle = 0$
- $\langle \varphi | X_j | \varphi \rangle = 0$

Proof. The eigenspace of X_j is spanned by states of the form

$$|b_0 b_1 b_2 \dots b_{j-1}\rangle \otimes |+\rangle \otimes |b_{j+1} b_{j+2} \dots b_N\rangle$$

Applying $Z_i Z_j$ results in one of the following states depending on whether $i > j$ or $i < j$.

- $|b_0 b_1 \dots b_{i-1}\rangle \otimes |(1-b_i)\rangle \otimes |b_{i+1} b_{i+2} \dots b_{j-1}\rangle \otimes |-\rangle \otimes |b_{j+1} b_{j+2} \dots b_N\rangle$ if $i < j$
- $|b_0 b_1 \dots b_{j-1}\rangle \otimes |-\rangle \otimes |b_{j+1} b_{j+2} \dots b_{i-1}\rangle \otimes |(1-b_i)\rangle \otimes |b_{i+1} b_{i+2} \dots b_N\rangle$ if $i > j$

It immediately follows that for every basis state of the $+1$ eigenspace of X_j , $\langle \psi | Z_i Z_j | \psi \rangle = 0$. It follows that for any state $|\chi\rangle$ satisfying $X_j |\chi\rangle = |\chi\rangle$, $\langle \chi | Z_i Z_j | \chi \rangle = 0$.

The eigenspace of $Z_i Z_j$ is spanned by states of the form $|b_0 b_1 \dots b_N\rangle$ where $b_i = b_j$. Acting with X_j on this state flips b_j , and it follows that $\langle \psi | X_j | \psi \rangle = 0$ for any $|\psi\rangle$ in the basis for the 1 eigenspace of $Z_i Z_j$. Thus for any state $|\varphi\rangle$ satisfying $Z_i Z_j |\varphi\rangle = |\varphi\rangle$, $\langle \varphi | X_j | \varphi \rangle = 0$. \square

⁷Notice the edges in a spanning tree follow a similar requirement.

IV. THE TRANSVERSE ISING HAMILTONIAN

Given a graph G , the Transverse Ising Hamiltonian on graph G can be defined as the following N -qubit Hamiltonian,

$$H = -J \sum_{\langle q_i, q_j \rangle \in E_G} Z_i Z_j - h \sum_{q_i \in V_G} X_i \quad (8)$$

For our purposes we rescale this Hamiltonian as

$$H = - \sum_{\langle q_i, q_j \rangle \in E_G} Z_i Z_j - g \sum_{q_i \in V_G} X_i \quad (9)$$

for some non-negative dimensionless $g = \frac{h}{J}$. This rescaling has the effect of normalizing all eigenvalues by J . An overview of Transverse Ising models on L_n and P_n can be found in [33] while an overview of Transverse Ising Models on random graphs can be found in [34]. A more general overview can be found in [35].

Minimizing the expectation value of H acting on Clifford state $|\varphi\rangle$ then corresponds to minimizing the following,

$$\langle \varphi | H | \varphi \rangle = - \sum_{\langle q_i, q_j \rangle \in E_G} \langle \varphi | Z_i Z_j | \varphi \rangle - g \sum_{q_i \in V_G} \langle \varphi | X_i | \varphi \rangle \quad (10)$$

Because all the Pauli operators above are Hermitian, every expectation value in the above expression is either ± 1 or 0 . The following theorem defines a method for computing $\langle \varphi_0 | H | \varphi_0 \rangle$ by solving a related graph theoretic problem.

Theorem IV.1. Given

$$V = \arg \min_{S \in P(V_G)} -|E(S)| - g|V_G \setminus S| \quad (11)$$

we can use V to find φ_0 which satisfies

$$\langle \varphi_0 | H | \varphi_0 \rangle = -|E(V)| - g|V_G \setminus V| \quad (12)$$

This theorem will follow from the proofs of Lemmas IV.2 and IV.3 below.

Lemma IV.2. Given an arbitrary vertex set $V \in P(V_G)$, we can always find a state φ such that

$$\langle \varphi | H | \varphi \rangle = -|E(V)| - g|V_G \setminus V| \quad (13)$$

Proof. Let S be the subgraph induced by V . By Lemma III.8, there must be some spanning tree M with $|E_M| < |S|$.

By definition, there cannot exist a sequence of edges of the form $\langle q_i, q_{k_1} \rangle, \langle q_{k_1}, q_{k_2} \rangle, \dots, \langle q_{k_m}, q_j \rangle$ if $\langle q_i, q_j \rangle \in E_M$ and vice versa. Thus it's possible to have a state φ with $Z_i Z_j \in \mathcal{G}_\varphi$ for all $\langle q_i, q_j \rangle \in E_M$ without violating Corollary III.12.

If we include all these $Z_i Z_j$ terms, then we can also include $X_i \in \mathcal{G}_\varphi$ if $q_i \notin V_M$ without violating Corollary III.12.

Totalalling the number of $Z_i Z_j$ and X_i terms gives $|E_M| + |V_G \setminus S|$ many terms, which is clearly less than N . This suggests that we can compute a state $|\varphi\rangle$ such that

- $X_i \in \mathcal{G}_\varphi$ if $q_i \notin V_M$
- $Z_i Z_j \in \mathcal{G}_\varphi$ for all $\langle q_i, q_j \rangle \in E_M$

It can be verified via direct computation that the state

$$|\varphi\rangle = \prod_{q_j \notin V} R_Y \left(\frac{\pi}{2} \right)_j |0\rangle \quad (14)$$

Satisfies the desired operators being in G_φ .

Furthermore, this state $|\varphi\rangle$ satisfies the following (which can again be verified by direct computation).

- $\langle\varphi|Z_iZ_j|\varphi\rangle = 1$ if $\langle q_i, q_j \rangle \in E(V)$
- $\langle\varphi|Z_iZ_j|\varphi\rangle = 0$ if $\langle q_i, q_j \rangle \notin E(V)$
- $\langle\varphi|X_i|\varphi\rangle = 1$ if $q_i \notin V$
- $\langle\varphi|X_i|\varphi\rangle = 0$ if $q_i \in V$

Thus, $\langle\varphi|H|\varphi\rangle = -|E(V)| - g|V_G \setminus V|$

- If $g \rightarrow \infty$, then

$$H \approx H_\infty = -g \sum_{q_i \in V_G} X_i \quad (20)$$

$|+\rangle^{\otimes N}$ is clearly a ground state of H_∞ .

For convenience, let

$$C(n) = \mathcal{E}(n) + g(N - n) \quad (21)$$

Since $N - n$ controls the number of X_i terms in G_φ , we expect that for small g , $n = N$ will be optimal while for large g we expect $n = 0$ to be optimal.

The following Lemma verifies this behavior.

Lemma IV.5.

$$g \leq \min_{0 \leq n < N} \frac{\mathcal{E}(N) - \mathcal{E}(n)}{N - n} \quad (22)$$

If any only if N maximizes $C(n)$.

On the other hand,

$$g \geq \max_{0 < n \leq N} \frac{\mathcal{E}(n)}{n} \quad (23)$$

If any only if 0 maximizes $C(n)$.

Proof. Suppose that

$$g \leq \min_{0 \leq n < N} \frac{\mathcal{E}(N) - \mathcal{E}(n)}{N - n} \quad (24)$$

This is equivalent to the following holding for all $0 \leq n < N$

$$g \leq \frac{\mathcal{E}(N) - \mathcal{E}(n)}{N - n} \quad (25)$$

$$\mathcal{E}(N) \geq g(N - n) + \mathcal{E}(n) \quad (26)$$

$$C(N) \geq C(n) \quad (27)$$

Which is then equivalent to $n = N$ being optimal.

Now suppose that

$$g \geq \max_{0 < n \leq N} \frac{\mathcal{E}(n)}{n} \quad (28)$$

This is equivalent to the following holding for all $0 < n \leq N$

$$g \geq \frac{\mathcal{E}(n)}{n} \quad (29)$$

$$gN \geq g(N - n) + \mathcal{E}(n) \quad (30)$$

$$C(0) \geq C(n) \quad (31)$$

Which is then equivalent to, $n = 0$ being optimal. \square

This Lemma has the following corollary

Corollary IV.6. If $\frac{\mathcal{E}(N)}{N} \geq \frac{\mathcal{E}(n)}{n}$ for $0 < n < N$, then

$$\min_{0 \leq n < N} \frac{\mathcal{E}(N) - \mathcal{E}(n)}{N - n} = \frac{\mathcal{E}(N)}{N} \quad (32)$$

and furthermore,

$$\langle\varphi_0|H|\varphi_0\rangle = \begin{cases} -\mathcal{E}(N) & \text{if } g \leq \frac{\mathcal{E}(N)}{N} \\ -gN & \text{if } g \geq \frac{\mathcal{E}(N)}{N} \end{cases} \quad (33)$$

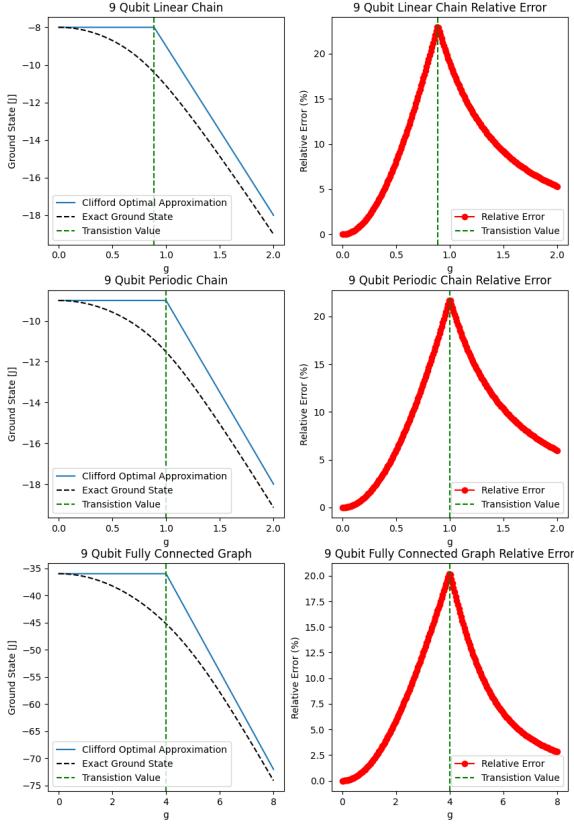


Fig. 1. Comparison of the optimal Clifford state to the true ground state expectation for the selected 9 qubit graphs. Notice that the maximal errors line up with the transition values and that all of these graphs are two segmented.

We call graphs for which $\frac{\mathcal{E}(N)}{N} \geq \frac{\mathcal{E}(n)}{n}$ for $0 < n < N$ *two-segmented* because of the result in Corollary IV.6. For two segmented graphs, we call $\frac{\mathcal{E}(N)}{N}$ the transition value.

It turns out, computing the maximal value of $\frac{\mathcal{E}(n)}{n}$ alongside the maximal vertex set associated with this maximum is possible in polynomial time as discussed in Appendix B.

V. NUMERICAL EXPERIMENTS

In this section we compare the optimal Clifford solution via our proposed method to ground states of Ising Hamiltonians obtained via exact diagonalization.

A. Linear Chains and Fully Connected Graphs

Figure 1 plots a comparison between $\langle \varphi_0 | H | \varphi_0 \rangle$ and $\langle \psi_0 | H | \psi_0 \rangle$ for the Hamiltonians obtained from L_9, P_9 , and K_9 for different values of g .

Figure 2 plots the relative error for L_n, P_n, K_n with $4 \leq n \leq 20$ at the transition value of g .

B. Other Graphs

Figure 4 plots the value of $\langle \psi_0 | H | \psi_0 \rangle$ versus $\langle \varphi_0 | H | \varphi_0 \rangle$ for G_1, G_2, G_3 for a range of g values.

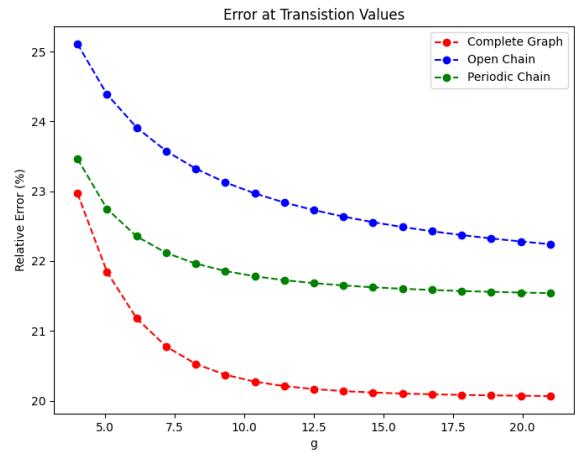


Fig. 2. Plot of the relative error for L_n, P_n , and K_n for a range of values for n . Notice that all of the relative errors decrease with n and eventually begin plateau.

C. Randomized Graphs

Figure 3 plots the average relative error for random graphs. For $4 \leq N \leq 17$, 100 N vertex graphs were generated by randomly placing a node between every pair of vertices with probability 0.5. If the graph with no edges was obtained, this distribution was resampled.

VI. FURTHER RESEARCH

A. Other Hamiltonians

Here we studied a specific spin-hamiltonian. There are multiple generalizations that could have been considered.

1) *Weighted Ising Hamiltonians:* The most natural generalization of the Ising Hamiltonian considered here is the following

$$H = - \sum_{\langle q_i, q_j \rangle \in E_G} J_{ij} Z_i Z_j - \sum_{q_i \in V_G} J_i X_i$$

Assuming that J_{ij} and J_i are both positive for all values of i, j , we can once again tie the minimization of $\langle \varphi | H | \varphi \rangle$ to the minimization of a submodular function by incorporating these terms into the cost function.

Corollary VI.1. *For a set of coefficients J_{ij} and J_i ,*

$$\langle \varphi_0 | H | \varphi_0 \rangle = \min_{V \in P(V_G)} - \sum_{\langle q_i, q_j \rangle \in E(V)} J_{ij} - \sum_{q_i \in V_G \setminus V} J_i$$

Where the function on the right side is a submodular function of V .

An example of an application of this modified Transverse Ising Hamiltonian, which accounts for different strength spin couplings can be found in [36]. In the case where J_{ij}, J_i are arbitrary, more careful analysis is required to handle negative coefficients because these can cause the corresponding graph theoretic function to no longer be submodular.

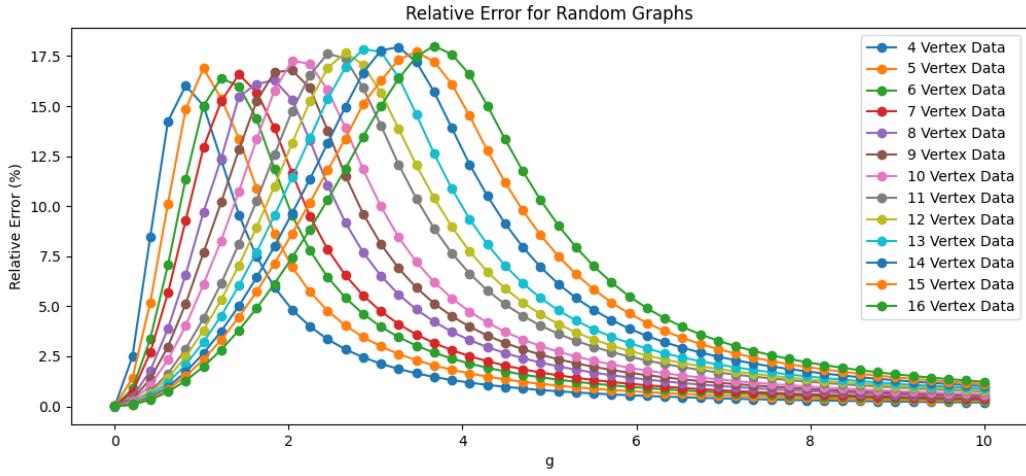


Fig. 3. Plot of the mean relative error for different amounts of vertices. Notice that the value of g for which the maximum mean relative error is obtained increases with N as expected.

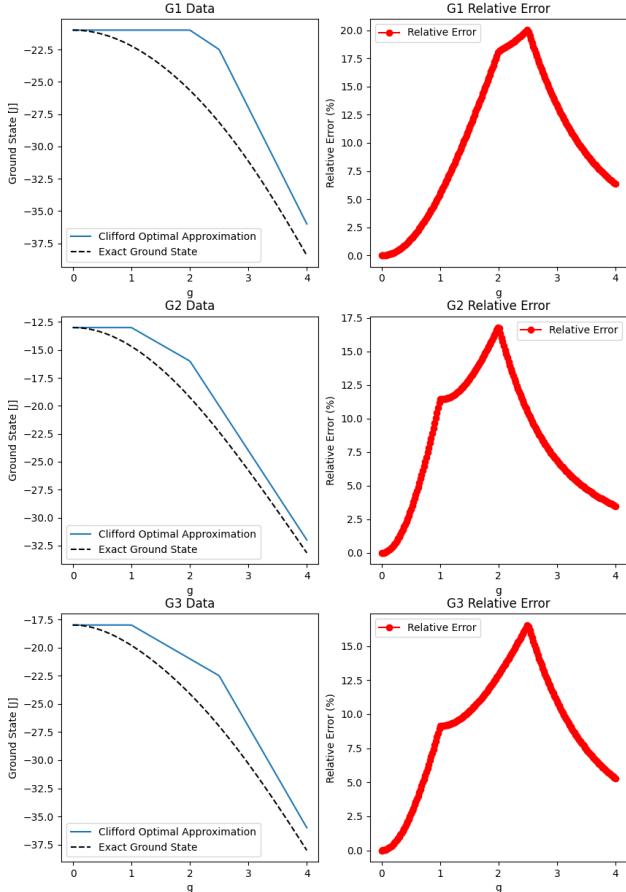


Fig. 4. Comparison of the optimal Clifford state to the true ground state expectation for the selected graphs. Notice that none of these graphs are two segmented.

An interesting variation of this Hamiltonian is the version with the X_i terms each replaced with Z_i . Such Hamiltonians often arise from QUBOs [37] [38] [39]. These Hamiltonians are generally constructed so that all of their eigenvalues are Clifford, which generally implies that there is no non-NP-hard algorithm to find the optimal Clifford State for these Hamiltonians.

2) *Heisenberg Hamiltonians*: The general Heisenberg Hamiltonian can be defined as follows [40]

$$H = - \sum_{\langle q_i, q_j \rangle} J_x X_i X_j + J_y Y_i Y_j + J_z Z_i Z_j$$

For arbitrary coefficients J_x, J_y, J_z .

In the case where $J_x = J_y = J_z = 1$, we obtain the XXX model, and in the case where $J_x = J_y = 1$ we obtain the XXZ model.

Using methods similar to those in [41], it can be shown analytically that the ground state of the XXX model and the XXZ model with $J_z > 1$ is always the Clifford state $|0\rangle$ regardless of the underlying graph G .

Finding the optimal Clifford state for a generalized Heisenberg Hamiltonian on an arbitrary graph will likely correspond to minimizing some constrained submodular function because of the additional restrictions on which terms in the Hamiltonian can simultaneously be in a given stabilizer. Unfortunately, this means that finding the optimal Clifford state for these problems may be NP-Hard [42].

VII. CONCLUSION

In this work we outlined an efficient method for finding the optimal Clifford ground state of a given transverse field Ising Hamiltonian on an arbitrary graph. Our numerical estimates suggest that these optimal Clifford ground states are generally fairly accurate (with errors always less than 25%), suggesting that CAFQA-inspired approaches to VQE problems are useful initialization techniques for these Hamiltonians.

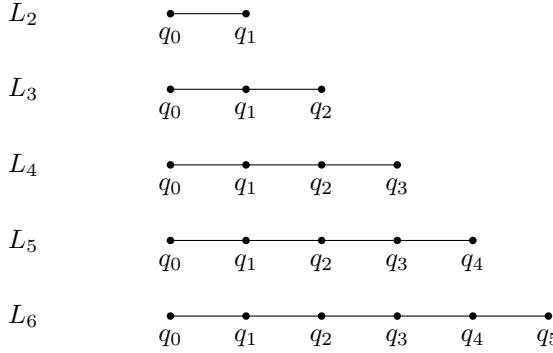


Fig. 5. First 5 open linear chains

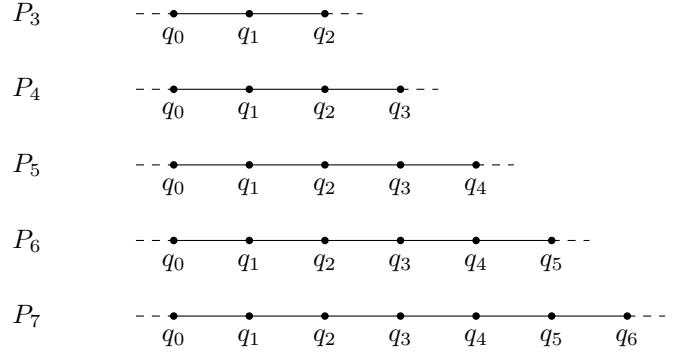


Fig. 6. First 5 periodic linear chains

ACKNOWLEDGEMENT

This work is funded in part by EPiQC, an NSF Expedition in Computing, under award CCF-1730449; in part by STAQ under award NSF Phy-1818914; in part by NSF award 2110860; in part by the US Department of Energy Office of Advanced Scientific Computing Research, Accelerated Research for Quantum Computing Program; and in part by the NSF Quantum Leap Challenge Institute for Hybrid Quantum Architectures and Networks (NSF Award 2016136) and in part based upon work supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers. This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.

APPENDIX A GRAPH LIST

The following graphs are used throughout for computations. Notice every graph here is connected.

A. Linear Chains

The n -vertex open linear chains, L_n , with $n \geq 2$ will be defined as the graphs with nodes $\{q_0, q_1, \dots, q_{n-1}\}$ and only edges from q_i to q_{i+1} for $1 \leq i < n$. Figure 5 shows the first 5 open linear chains. The n -vertex periodic linear chains, P_n , with $n \geq 3$ will be defined as L_n with the additional edge from q_{n-1} to q_1 . The first 5 periodic chains are drawn in Figure 6.

B. Fully Connected Graphs

The fully connected graph with K_4 with $n \geq 4$ will be defined as the graph with vertices $\{q_0, q_1, \dots, q_{n-1}\}$ and an edge between every pair of vertices. Figure 7 shows the first 3 fully connected graphs.

C. Miscellaneous Graphs

The following graphs (Figure 8) are designed to have subgraphs which are more dense than the entire graph, and thus are all not two-segmented. These graphs will be referred as G_1 , G_2 , and G_3 .

APPENDIX B

OVERVIEW OF RELEVANT ALGORITHMS

In this appendix we provide a brief overview of the relevant algorithms used in this paper. All relevant code can be found at [43].

A. Edge Functions (DkS)

The *densest k-subgraph* (DkS) problem can be defined as the problem of finding the vertex set $V \subset V_G$ with $|V| = k$ for fixed k such that $\frac{|E(V)|}{|V|}$ is maximized. This algorithm is known to be NP-hard [44].

Clearly, computing $\mathcal{E}(n)$ for $n \in \{0, 1, \dots, N\}$ is equivalent to solving DkS for $k \in \{0, 1, \dots, N\}$, meaning that computing all of the values for $\mathcal{E}(n)$ is an NP-hard problem.

Currently, the best known approximate algorithm for DkS was introduced in [45], and computes an $O(n^{1/4+\epsilon})$ approximation in $O(n^{1/\epsilon})$.

Although edge functions are useful theoretical tools, we generally don't need to consider specific values of \mathcal{E} when searching for the optimal Clifford state, meaning we can avoid this NP-Hard problem.

B. Two-Segmented Testing (DSP)

The *densest subgraph problem* (DSP) can be defined as the problem of finding the vertex set $V \subset V_G$ such that $\frac{|E(V)|}{|V|}$ is maximized. It is well known that the Densest Subgraph Problem can be solved in polynomial time [44].

One solution method uses a linear programming problem in $N + |E_G|$ variables, which was introduced in [46]. Another polynomial time solution was introduced in [47] and uses maximum-flow computations.

Starting with an arbitrary graph G , we can always run a solution to DSP on it, and if the obtained vertex set V satisfies $\frac{|E(V)|}{|V|} = \frac{|E(N)|}{|N|}$, we immediately know that G must be two segmented.

C. Submodular Minimization

The most important algorithm used here is the algorithm to minimize the function

$$f(V) = -|E(V)| - g|V_G \setminus V|$$

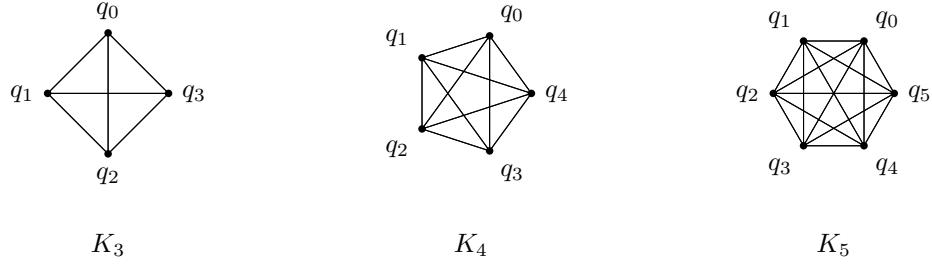


Fig. 7. First 3 fully connected graphs

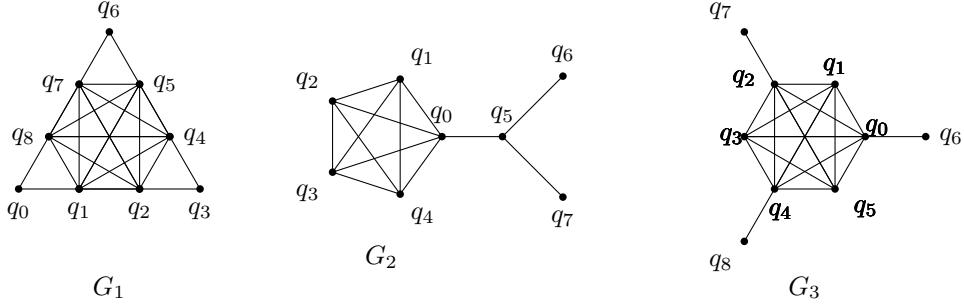


Fig. 8. Various non two-segmented graphs.

The following corollary immediately follows from the definition f ,

Corollary B.1. *For any sets $A, B \subset V_G$,*

$$f(A) + f(B) \leq f(A \cap B) + f(A \cup B)$$

Functions f which satisfy corollary B.1 are called *submodular* functions. A comprehensive overview of submodular functions can be found in [48].

One of the most important properties of submodular functions is that they can always be minimized in polynomial time. The first algorithm to do so was introduced in [49] and [50], but this algorithm is seemingly too slow for practical applications. Another independent algorithm was introduced in [51], which uses an algorithm to minimize the norm of a point in a polyhedra introduced in [52] and theory developed in [53]. An approximation for the performance of this algorithm was obtained in [26], which provided a pseudopolynomial bound on the runtime of this algorithm.

REFERENCES

- [1] G. S. Ravi, P. Gokhale, Y. Ding, W. M. Kirby, K. N. Smith, J. M. Baker, P. J. Love, H. Hoffmann, K. R. Brown, and F. T. Chong, “Cafqa: A classical simulation bootstrap for variational quantum algorithms,” 2022.
- [2] P. W. Shor, “Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer,” *SIAM Journal on Computing*, vol. 26, no. 5, p. 1484–1509, Oct 1997. [Online]. Available: <http://dx.doi.org/10.1137/S0097539795293172>
- [3] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, “Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets,” *Nature*, vol. 549, no. 7671, pp. 242–246, 2017.
- [4] N. Moll, P. Barkoutsos, L. S. Bishop, J. M. Chow, A. Cross, D. J. Egger, S. Filipp, A. Fuhrer, J. M. Gambetta, M. Ganzhorn *et al.*, “Quantum optimization using variational algorithms on near-term quantum devices,” *Quantum Science and Technology*, vol. 3, no. 3, p. 030503, 2018.
- [5] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, “Quantum machine learning,” *Nature*, vol. 549, no. 7671, pp. 195–202, 2017.
- [6] J. Preskill, “Quantum computing in the nisq era and beyond,” *Quantum*, vol. 2, p. 79, 2018.
- [7] L. K. Grover, “A fast quantum mechanical algorithm for database search,” in *ANNUAL ACM SYMPOSIUM ON THEORY OF COMPUTING*. ACM, 1996, pp. 212–219.
- [8] J. O’Gorman and E. T. Campbell, “Quantum computation with realistic magic-state factories,” *Physical Review A*, vol. 95, no. 3, Mar 2017. [Online]. Available: <http://dx.doi.org/10.1103/PhysRevA.95.032338>
- [9] P. Czarnik, A. Arrasmith, P. J. Coles, and L. Cincio, “Error mitigation with clifford quantum-circuit data,” 2020.
- [10] E. Rosenberg, P. Ginsparg, and P. L. McMahon, “Experimental error mitigation using linear rescaling for variational quantum eigensolving with up to 20 qubits,” *Quantum Science and Technology*, Nov 2021. [Online]. Available: <http://dx.doi.org/10.1088/2058-9565/ac3b37>
- [11] G. S. Barron and C. J. Wood, “Measurement error mitigation for variational quantum algorithms,” 2020.
- [12] L. Botelho, A. Glos, A. Kundu, J. A. Miszczak, Özlem Salehi, and Z. Zimborás, “Error mitigation for variational quantum algorithms through mid-circuit measurements,” 2021.
- [13] S. Wang, P. Czarnik, A. Arrasmith, M. Cerezo, L. Cincio, and P. J. Coles, “Can error mitigation improve trainability of noisy variational quantum algorithms?” 2021.
- [14] K. Temme, S. Bravyi, and J. M. Gambetta, “Error mitigation for short-depth quantum circuits,” *Physical review letters*, vol. 119, no. 18, p. 180509, 2017.
- [15] Y. Li and S. C. Benjamin, “Efficient variational quantum simulator incorporating active error minimization,” *Phys. Rev. X*, vol. 7, p. 021050, Jun 2017. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevX.7.021050>

[16] T. Giurgica-Tiron, Y. Hindy, R. LaRose, A. Mari, and W. J. Zeng, “Digital zero noise extrapolation for quantum error mitigation,” in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*. IEEE, 2020, pp. 306–316.

[17] G. S. Ravi, K. N. Smith, P. Gokhale, A. Mari, N. Earnest, A. Javadi-Abhari, and F. T. Chong, “Vaqem: A variational approach to quantum error mitigation,” 2021.

[18] W. Lavrijsen, A. Tudor, J. Müller, C. Iancu, and W. de Jong, “Classical optimizers for noisy intermediate-scale quantum devices,” in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*, 2020, pp. 267–277.

[19] W. Tang, T. Tomesh, M. Suchara, J. Larson, and M. Martonosi, “Cutqc: using small quantum computers for large quantum circuit evaluations,” *Proceedings of the 26th ACM International Conference on Architectural Support for Programming Languages and Operating Systems*, Apr 2021. [Online]. Available: <http://dx.doi.org/10.1145/3445814.3446758>

[20] Y. Zhang, L. Cincio, C. F. Negre, P. Czarnik, P. Coles, P. M. Anisimov, S. M. Mniszecki, S. Tretiak, and P. A. Dub, “Variational quantum eigensolver with reduced circuit complexity,” *arXiv preprint arXiv:2106.07619*, 2021.

[21] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O’Brien, “A variational eigenvalue solver on a photonic quantum processor,” *Nature communications*, vol. 5, p. 4213, 2014.

[22] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, “The theory of variational hybrid quantum-classical algorithms,” *New Journal of Physics*, vol. 18, no. 2, p. 023023, 2016.

[23] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, “Barren plateaus in quantum neural network training landscapes,” *Nature Communications*, vol. 9, no. 1, Nov 2018. [Online]. Available: <http://dx.doi.org/10.1038/s41467-018-07090-4>

[24] S. Wang, E. Fontana, M. Cerezo, K. Sharma, A. Sone, L. Cincio, and P. J. Coles, “Noise-induced barren plateaus in variational quantum algorithms,” *arXiv preprint arXiv:2007.14384*, 2020.

[25] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. Love, and A. Aspuru-Guzik, “Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz,” 2018.

[26] D. Chakrabarty, P. Jain, and P. Kothari, “Provable submodular minimization using Wolfe’s algorithm,” 2014.

[27] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth, and J. Tennyson, “The variational quantum eigensolver: A review of methods and best practices,” *Physics Reports*, vol. 986, pp. 1–128, nov 2022. [Online]. Available: <https://doi.org/10.1016%2Fphsrep.2022.08.003>

[28] Y. Du, Z. Tu, X. Yuan, and D. Tao, “Efficient measure for the expressivity of variational quantum algorithms,” *Physical Review Letters*, vol. 128, no. 8, feb 2022. [Online]. Available: <https://doi.org/10.1103%2Fphysrevlett.128.080506>

[29] S. Aaronson and D. Gottesman, “Improved simulation of stabilizer circuits,” *Physical Review A*, vol. 70, no. 5, nov 2004. [Online]. Available: <https://doi.org/10.1103%2Fphysreva.70.052328>

[30] D. Gottesman, “The heisenberg representation of quantum computers,” *arXiv preprint quant-ph/9807006*, 1998.

[31] R. J. Wilson, *Introduction to Graph Theory*, 4th ed. Addison-Wesley Longman Limited, 1996.

[32] R. Scheffler, “Graph search trees and their leaves,” 2023.

[33] G. B. Mbeng, A. Russomanno, and G. E. Santoro, “The quantum ising chain for beginners,” 2020.

[34] G. Grimmett, *Probability on Graphs*, 2nd ed. Cambridge University Press, 2018.

[35] J. Strecka and M. Jascur, “A brief account of the ising and ising-like models: Mean-field, effective-field and exact results,” 2015.

[36] M. P. Kaicher, D. Vodola, and S. B. Jäger, “Mean-field treatment of the long-range transverse field ising model with fermionic gaussian states,” *Physical Review B*, vol. 107, no. 16, apr 2023. [Online]. Available: <https://doi.org/10.1103%2Fphysrevb.107.165144>

[37] E. Farhi, J. Goldstone, and S. Gutmann, “A quantum approximate optimization algorithm,” 2014.

[38] K. Blekos, D. Brand, A. Ceschini, C.-H. Chou, R.-H. Li, K. Pandya, and A. Summer, “A review on quantum approximate optimization algorithm and its variants,” 2023.

[39] J. Nüßlein, T. Gabor, C. Linnhoff-Popien, and S. Feld, “Algorithmic QUBO formulations for k -SAT and hamiltonian cycles,” in *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. ACM, jul 2022. [Online]. Available: <https://doi.org/10.1145%2F3520304.3533952>

[40] B. J. Powell, “An introduction to effective low-energy hamiltonians in condensed matter physics and chemistry,” 2010.

[41] M. V. Rakov and M. Weyrauch, “Spin-1/2 xxz heisenberg chain in a longitudinal magnetic field,” *Physical Review B*, vol. 100, no. 13, oct 2019. [Online]. Available: <https://doi.org/10.1103%2Fphysrevb.100.134434>

[42] K. Nagano, Y. Kawahara, and K. Aihara, “Size-constrained submodular minimization through minimum norm base,” in *Proceedings of the 28th International Conference on International Conference on Machine Learning*, ser. ICML’11. Madison, WI, USA: Omnipress, 2011, p. 977–984.

[43] B. Bhattacharyya, “OptimalCliffordIsing.” [Online]. Available: <https://github.com/Bhattacharyya1729/OptimalCliffordIsing>

[44] T. Lanciano, A. Miyauchi, A. Fazzone, and F. Bonchi, “A survey on the densest subgraph problem and its variants,” 2023.

[45] A. Bhaskara, M. Charikar, E. Chlamtac, U. Feige, and A. Vijayaraghavan, “Detecting high log-densities – an $o(n^{1/4})$ approximation for densest k -subgraph,” 2010.

[46] M. Charikar, “Greedy approximation algorithms for finding dense components in a graph,” in *Proceedings of the Third International Workshop on Approximation Algorithms for Combinatorial Optimization*, ser. APPROX ’00. Berlin, Heidelberg: Springer-Verlag, 2000, p. 84–95.

[47] A. V. Goldberg, “Finding a maximum density subgraph,” EECS Department, University of California, Berkeley, Tech. Rep. UCB/CSD-84-171, 1984. [Online]. Available: <http://www2.eecs.berkeley.edu/Pubs/TechRpts/1984/5956.html>

[48] S. Fujishige, *Submodular Functions and Optimization*, 2nd ed. Elsevier Science, 2005.

[49] S. Iwata, L. Fleischer, and S. Fujishige, “A combinatorial, strongly polynomial-time algorithm for minimizing submodular functions,” 2000.

[50] A. Schrijver, “A combinatorial algorithm minimizing submodular functions in strongly polynomial time,” *Journal of Combinatorial Theory, Series B*, vol. 80, no. 2, pp. 346–355, 2000. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0095895600919890>

[51] S. Fujishige and S. Iwata, “A submodular function minimization algorithm based on the minimum-norm base,” *Pacific Journal of Optimization*, vol. 7, 04 2009.

[52] P. Wolfe, “Finding the nearest point in a polytope,” *Mathematical Programming*, vol. 11, pp. 128–149, 1976. [Online]. Available: <https://api.semanticscholar.org/CorpusID:17825447>

[53] J. Edmonds, *Submodular Functions, Matroids, and Certain Polyhedra*. Berlin, Heidelberg: Springer-Verlag, 2003, p. 11–26.