Received 8 September 2024; revised 9 November 2025; accepted 7 January 2026; date of publication 16 January 2026;
date of current version 26 February 2026.

Digital Object Identifier 10.1109/TQE.2026.3654930

Equivariant Quantum Approximate Optimization Algorithm

BORIS TSVELIKHOVSKIY[1], ILYA SAFRO[2], AND YURI ALEXEEV[3] (Senior Member, IEEE)

[1]Department of Mathematics, University of California, Riverside, CA 92521 USA
[2]Department of Computer and Information Sciences, University of Delaware, Newark, DE 19716 USA
[3]NVIDIA Corporation, Santa Clara, CA 95051 USA

Corresponding author: Boris Tsvelikhovskiy (e-mail: borist@ucr.edu).

ABSTRACT Constructing effective mixer Hamiltonians is essential for enhancing the performance of the quantum approximate optimization algorithm (QAOA) in solving combinatorial optimization problems. In this work, we develop a systematic methodology for designing QAOA mixers that align with the symmetries of the classical objective function, with the goal of achieving values (mean, median, and minimum over multiple runs) that are closer to the true optimum. Our main idea is to design QAOA operators that are explicitly adapted to the action of symmetry groups on the Hilbert space. We focus on subgroups of the symmetric group S_d, where $d = 2^\ell$, to ensure compatibility with qudit-based quantum architectures. In particular, we construct QAOA mixers invariant under the full symmetric group S_d as well as its cyclic subgroup $\mathbb{Z}_d \subset S_d$. These constructions are natural in that they respect the decomposition of the Hilbert space into isotypic components under the symmetry group action. Notably, to the best of the authors' knowledge, the QAOA algorithm based on the $\mathbb{Z}_d$-invariant mixer provides the first example of a QAOA protocol whose dynamics (up to final measurement) are confined entirely within a nontrivial irreducible representation of a symmetry group of the objective function. Although our work does not investigate the benefits of exploiting such subspaces as computational resources, we think that the very realization of a variational algorithm whose evolution is restricted to a nontrivial symmetry-adapted subspace is of fundamental conceptual interest. We provide closed-form expressions for these mixers, together with explicit quantum circuit implementations. To empirically evaluate our approach, we compare QAOA variants employing the standard mixer $B = \sum X_i$ with those using our proposed Hamiltonians H_M and H_χ on edge coloring and graph partitioning problems. Across multiple graph instances, our symmetry-adapted mixers consistently yield objective values closer to the optimum, demonstrating statistically significant improvements over classical baselines.

INDEX TERMS Mixer Hamiltonians, quantum approximate optimization algorithm, warm-start quantum approximate optimization algorithm (QAOA).

I. INTRODUCTION

Variational quantum algorithms, a type of hybrid quantum-classical algorithm, are regarded as leading contenders for showcasing quantum advantage in fields like optimization and machine learning across diverse applications [13], [16], [21]. These algorithms consist of parameterized quantum circuits, with parameters updated through classical computation. The quantum approximate optimization algorithm (QAOA) [6] is a variational algorithm designed to solve combinatorial optimization problems. In the domain of optimization on graphs, it has been demonstrated on NP-hard problems, such as MaxCut [6], community detection [24], and partitioning [29], usually, by mapping these problems onto a classical spin-glass model (the Ising model) and minimizing the corresponding energy, which is itself an NP-hard task. Demonstrating the practical quantum advantage in future using QAOA class of algorithms depends on several factors, such as the quality of qubits, fidelity of quantum gates and entire circuits, circuit depth, connectivity of the quantum hardware architecture, and classical optimizer to find the best variational parameters to mention just a few. Recent studies have further investigated the technical capabilities of the algorithm through the lens of dynamical Lie algebras, providing a deeper understanding of its expressive power and limitations (see, e.g., [7] and [17] for details). One of the most significant factors is the way to construct the efficient mixer

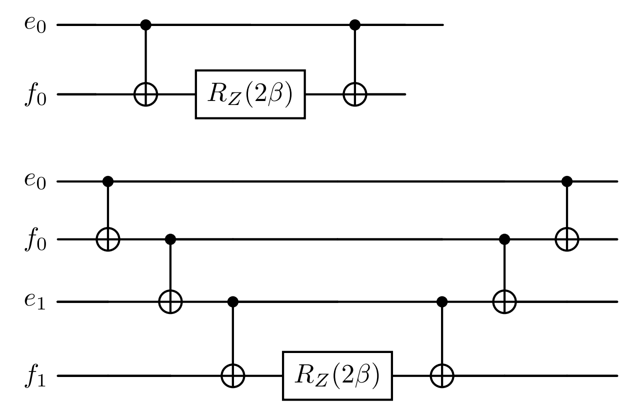

FIGURE 1. **Quantum circuits for $e^{-i\beta Z_{e,0}Z_{f,0}}$ and $e^{-i\beta Z_{e,0}Z_{f,0}Z_{e,1}Z_{f,1}}$.**

operator (will be explained later) to accelerate the variational algorithm. Constructing the efficient mixer and providing a theoretical explanation of its performance with respect to fundamental graph problems, as well as several aspects of popular quantum-classical heuristics, are the central goals of this article.

We consider the optimization problem of finding extremal values of a function $F : \mathbb{D}^n \to \mathbb{R}$, where $\mathbb{D}^n$ represents the set of n-element d-ary strings and $\mathcal{S}$ is the group of permutations acting on these d^n elements. The QAOA is a widely used approach for solving the quantum version of the optimization problem. Consequently, there is a growing interest to enhance its performance. To bridge the classical and quantum realms, one employs the following correspondences.

1) $\mathbb{D}^n \rightsquigarrow$ vector space W of dimension d^n with basis $|x\rangle$ indexed by elements $x \in \mathbb{D}^n$.
2) Objective function $F \rightsquigarrow$ linear operator H_P acting on W.
3) Minima of F on $\mathbb{D}^n \rightsquigarrow$ lowest energy states of H_P in W.

Here, the Hamiltonian H_P represents the objective function F, meaning it satisfies the equation $H_P(|x\rangle) = F(x)|x\rangle$ for any string $x \in \mathbb{D}^n$. Another important component of the QAOA approach is an operator referred to as the mixer Hamiltonian H_M. This operator plays a pivotal role in the optimization process, as it possesses an easily identifiable ground state, which aids in initializing the optimization process.

The QAOA algorithm involves a multistep transformation of H_M into H_P, aiming to obtain a lowest energy state for the latter Hamiltonian. This is achieved by alternately applying exponentials of H_M and H_P, with the number of iterations denoted by p (known as QAOA depth). We express this transformation as

$$\mathfrak{Q}_p = e^{-i\beta_1 H_M} e^{-i\gamma_1 H_P} \ldots e^{-i\beta_p H_M} e^{-i\gamma_p H_P}. \tag{1}$$

The algorithm concludes with a measurement of the resulting state in the standard basis.

While the problem Hamiltonian H_P is uniquely determined by the classical original problem (unless it is decided to be changed, e.g., by sparsification [18]), there is a flexibility in choosing the mixer Hamiltonian H_M. The convergence of QAOA is ensured by the adiabatic theorem if H_M satisfies certain conditions. For example, the assumptions outlined in the Perron–Frobenius theorem (see Theorem III.1) are sufficient.

A commonly used mixer Hamiltonian consists of Pauli X-gates, $B = \sum_{j=0}^{\ell-1} X_j$, where ℓ is the number of qubits required for the problem. However, this choice may not exploit problem-specific attributes. The choice of mixer Hamiltonian has been discussed in the literature. Hadfield et al. [11] introduced a quantum alternating operator ansatz to allow more general families of Hamiltonian operators. The mixers discussed in that article are particularly effective for optimization problems involving hard constraints that must always be satisfied, thereby defining a feasible subspace of W and soft constraints, which must be violated as little as possible.

In [10], it was experimentally verified (via numerical simulations) that linear combinations of X- and Y-Pauli gates as mixers can outperform the standard low depth QAOA. More examples can be found in [1], [8], [25], and [34] and subsequent references.

Constructing an effective mixer for QAOA Hamiltonian is crucial for enhancing the performance of QAOA in solving combinatorial optimization problems. effective mixers not only enforce hard constraints and align with the initial state for improved performance but also contribute to the universality and computational efficiency of QAOA, enabling the algorithm to exploit the structure of optimization problems for significant speed-ups and to adapt effectively to constrained problems. Here, are a few of the examples as follows.

1) *Enforcing hard constraints:* The application of QAOA to problems with constraints presents a notable challenge, especially for near-term quantum resources. Utilizing XY Hamiltonians as mixers has been shown to enforce hard constraints effectively. These mixers can be implemented without Trotter error in certain cases, and they demonstrate significant improvement in performance over traditional X mixers in solving graph-coloring problems, a known challenge for classical algorithms [31].
2) *Alignment with initial state:* The improvement in QAOA performance has been linked to better alignment between the initial state and the ground state of the mixing Hamiltonian. This alignment, mimicking the adiabatic algorithm's requirements, has been particularly beneficial in constrained portfolio optimization, showcasing that an effective mixer enhances results across different QAOA depths [12].
3) *Universality and computational efficiency.* The universality of QAOA with effective mixers extends its applicability across a broader spectrum of problems.

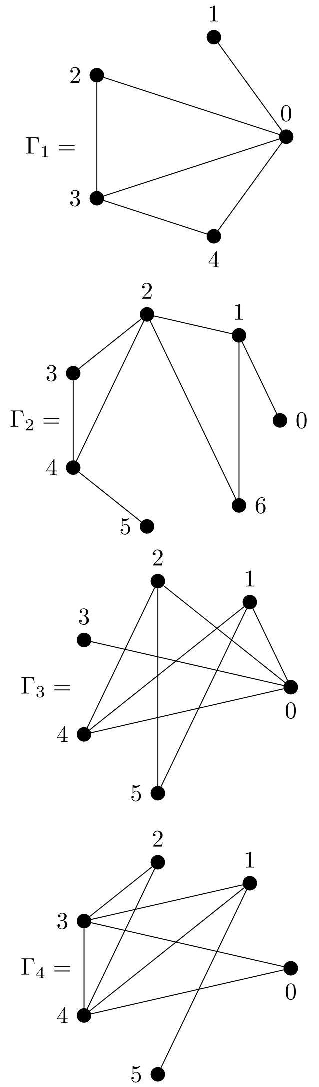

FIGURE 2. Graphs Γ_1, Γ_2, Γ_3, and Γ_4 considered for edge coloring problem.

effective mixers contribute to the quantum computational universality, enabling the solution of complex optimization problems with high efficiency and precision. This universality underpins QAOA's potential in leveraging quantum computing for practical applications [19].

4) *Exploiting problem structure for speed-up:* Recent studies have provided numerical evidence that QAOA, with appropriately chosen mixers and phase operators [18], can significantly outperform classical unstructured search algorithms in finding approximate solutions to (un)constrained optimization problems. This suggests that effective mixers are key to leveraging the structure of optimization problems for computational speed-up [9].
5) *Custom mixers for constrained problems:* For constrained optimization problems, particularly those related to network flows, specialized mixers based on

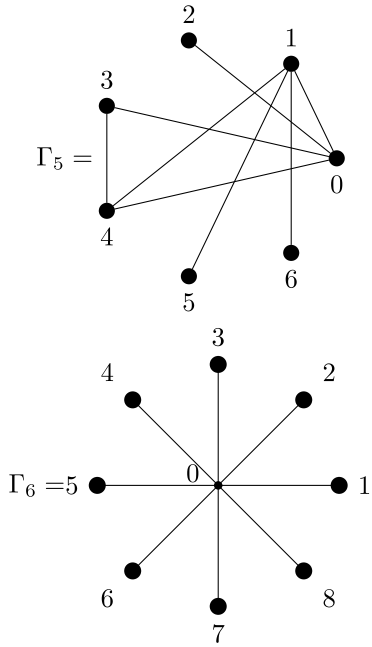

FIGURE 3. Graphs Γ_5 and Γ_6 considered for edge coloring problem.

quantum electrodynamics have been demonstrated to maintain flow constraints, resulting in an exponential reduction in the search space. This adaptation results in higher quality approximate solutions, underscoring the importance of mixer customization [33].

Let G denote the symmetry group of the problem Hamiltonian. Recent work (see [22] and [25]) has shown that incorporating the G-action can both enable reductions of QAOA to symmetry-invariant subspaces of W and improve fidelity. In this article, we extend various investigation into tailoring the mixer Hamiltonian to accommodate groups of classical symmetries inherent in the objective function. In particular, our exploration builds upon the groundwork laid out in [28], where we detailed the construction of mixer Hamiltonians, along with their corresponding ground states, designed for cases where the group of classical symmetries includes the symmetric group S_ℓ, encompassing permutations of string elements. While we presented compelling arguments advocating for the adoption of such mixer Hamiltonians over classical counterpart, practical validation was hindered by the challenge of implementing the suggested matrices as concrete quantum circuits. Our current focus is on cases where the group of classical symmetries involves a different symmetric group, S_d, acting by simultaneous permutation of all factors in $\mathbb{D}^n$

$$\sigma(d_1, d_2, \ldots, d_n) := (\sigma(d_1), \sigma(d_2), \ldots, \sigma(d_n)). \quad (2)$$

Remark I.1: We highlight a key distinction between the symmetry groups considered in [28] and in the present work. The mixers in [28] are tailored to the action of the symmetric

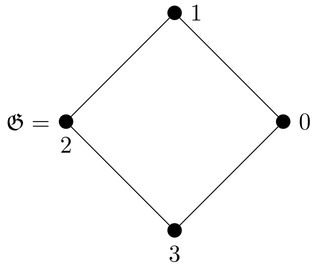

FIGURE 4. **Graph $\mathfrak{G}$ considered for graph partitioning problem.**

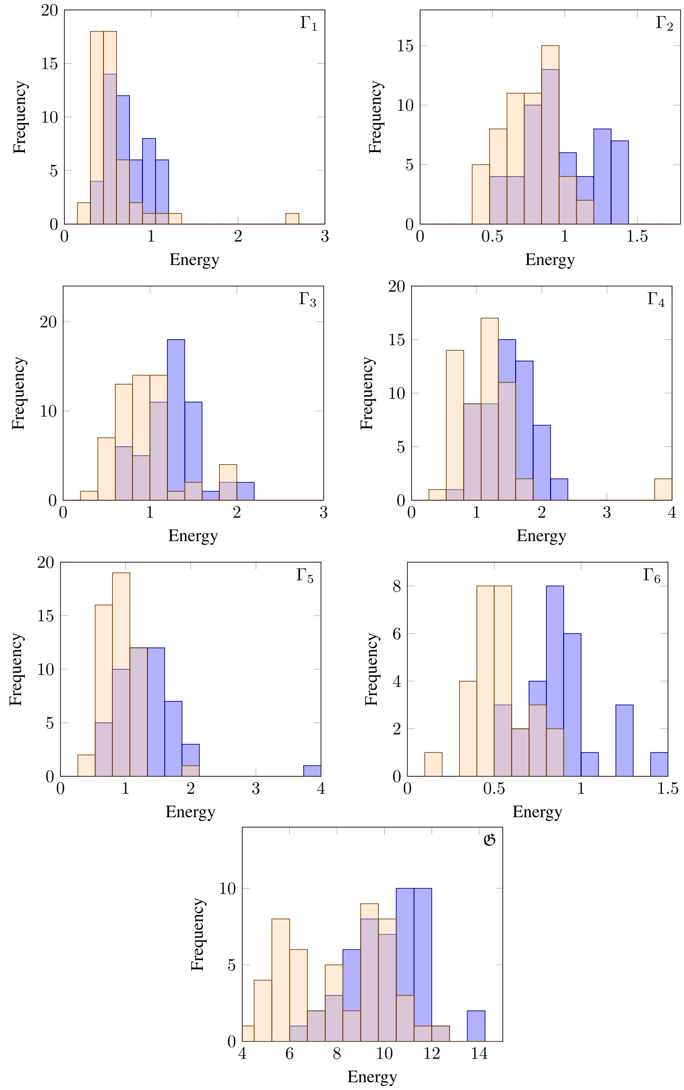

FIGURE 5. **Histograms illustrating the frequency distributions of $\mathcal{E}p$-values for algorithms utilizing mixers H_M and H_M.**

group S_ℓ, which permutes qubits (or qudits) among themselves. In contrast, our focus is on symmetry groups corresponding to simultaneous permutations of standard basis states within each individual qudit.

Considering such cases offers the following two significant advantages.

1) Many optimization problems exhibit these symmetries (e.g., several versions of graph coloring and partitioning).
2) We can construct a mixer Hamiltonian that commutes with the action of S_d on W, which can be easily implemented as a composition of basic quantum gates.

The rest of this article is organized as follows for clear and systematic exposition. Section II offers an overview of the main results to orient the reader. In Section III, a concise review of QAOA fundamentals relevant to this study is provided.

Section IV presents formulations of the main results and delineates properties concerning the newly proposed Hamiltonian. The subsequent Sections V and VI, respectively, explore the classical optimization problems under consideration and provide simulation results for three QAOA versions: one utilizing the classical mixer and the others employing the newly proposed mixers.

In Section VII, the impossibility of tailoring a mixer Hamiltonian that satisfies the Perron–Frobenius theorem within the context of warm-start QAOA is discussed. Appendix offers a conceptual overview of the construction process and provides rigorous verification of the claims made throughout this article.

II. MAIN RESULTS

In this section, we present a summary of our principal contributions. The focus lies on constructing symmetry-respecting mixer Hamiltonians for QAOA, analyzing their mathematical properties, and evaluating their empirical performance.

A. MIXER HAMILTONIANS WITH S_d SYMMETRY

We introduce a systematic approach for constructing a mixer Hamiltonian in QAOA that respects certain symmetries of the objective function in the underlying classical optimization problem. Specifically, we construct an operator that commutes with the action of the symmetric group S_d, as defined in (2).

We rigorously verify that the proposed operator H_M satisfies the conditions of the Perron–Frobenius theorem (see Theorem III.1), ensuring its suitability as a mixer Hamiltonian. We also provide both a concrete analytic formula and a corresponding quantum circuit for its implementation [see (6) and Fig. 7 for examples with $d = 2$ and 3].

In contrast, we show that the classical mixer B commutes only with a subgroup of S_d of order $2^\ell \cdot \ell!$ (when $d = 2^\ell$ is a power of two). This is significantly smaller than $d!$, the order of S_d (see Proposition B.4), thereby highlighting the stronger symmetry properties of our proposed mixer.

Remark II.1: The use of S_d symmetry in designing QAOA mixers for qudits appears to be novel, despite its natural appearance in the qudit setting.

B. $\mathbb{Z}_d$-INVARIANT MIXER H_χ AND CONFINEMENT OF QAOA TO A NONTRIVIAL REPRESENTATION

Next, we examine the cyclic subgroup $\mathbb{Z}_d \subset S_d$, generated by the element $g = (23 \dots d1)$, which acts by cyclically permuting the symbols: $1 \mapsto 2 \mapsto 3 \mapsto \cdots \mapsto d \mapsto 1$.

We construct an operator H_χ whose action on W commutes with this $\mathbb{Z}_d$ subgroup (see (7) for the explicit formula). The operator H_χ admits a unique ground state

$$|\psi\rangle := |\underbrace{-+\ldots+}_{n\ell}\rangle$$

which lies in the subspace $W_{d/2}$ of the decomposition

$$W = \bigoplus_{j=0}^{d-1} W_j$$

under the $\mathbb{Z}_d$-action.

Importantly, during the execution of QAOA with H_χ as the mixer Hamiltonian, the evolution of $|\psi\rangle$ remains confined to $W_{d/2}$ until the final measurement (see Appendix A for the precise statements).

Remark II.2: This construction provides, to the best of the authors' knowledge, the first example of a QAOA algorithm that is executed entirely, aside from the final measurement, within an isotypic component (i.e., a direct sum of isomorphic irreducible representations) of a nontrivial symmetry group of the objective function.

C. BENCHMARKING: NUMERICAL PERFORMANCE COMPARISON

We evaluate the performance of three QAOA variants employing different mixer Hamiltonians: the conventional $B = \sum_i X_i$, and our proposed H_M and H_χ. The benchmarks include edge-coloring and graph-partitioning problems across diverse graph families.

For edge-coloring, we employ depth parameter $p = 9$, and for graph partitioning, $p = 7$. In each setting, we perform at least 50 independent trials. The results reveal statistically significant differences in mean objective values at the 1.5% level, with the proposed mixers consistently yielding lower means. Furthermore, both the median and minimum values observed with H_M and H_χ are substantially lower than those achieved with the conventional mixer (see Section VI for details).

SIGNIFICANCE AND POTENTIAL IMPACT

The results presented in this manuscript highlight how natural symmetry considerations can be leveraged to design fundamentally new QAOA mixers. On the practical side, our proposed mixers H_M and H_χ demonstrate tangible improvements in numerical experiments on standard combinatorial optimization problems. The empirical improvements we observe suggest that the proposed symmetry-respecting mixers are not merely of theoretical interest but may lead to more resource-efficient quantum algorithms that achieve better performance at shallower circuit depths.

Various QAOA implementations for qudit systems have recently appeared in the literature (see, e.g., [4] and [32]). While the practical realization of the constructions proposed in this article has so far been limited to $d = 2^\ell$ due to experimental constraints (see Remark IV.1), an interesting future direction is to explore their realization in more general qudit settings. In particular, one could investigate how QAOA circuits for arbitrary qudit dimensions, described in the aforementioned works, could be adapted to implement analogous symmetry-respecting mixers. Such implementations may enable more resource-efficient quantum algorithms that achieve better performance at shallower circuit depths, extending the empirical advantages we observed for H_M and H_χ in standard combinatorial optimization problems.

Beyond QAOA itself, the methodology developed here has the potential to impact a broader class of variational quantum algorithms. Consequently, our results offer not only immediate insights into QAOA but also a roadmap for symmetry-guided strategies in quantum algorithm design more generally.

D. OBSERVATIONS ON WARM-START QAOA

Finally, we address an intriguing observation regarding the subpar performance of warm-start QAOA variants, a phenomenon recently documented in the literature. Warm-start strategies involve initiating QAOA from a promising classical solution generated by a classical algorithm, with the aim of further refining it through quantum optimization. While this approach has garnered a lot of attention in recent studies [5], [20], [26], our investigation sheds light on its fundamental limitations.

In a recent study in [3], extensive numerical experiments across a range of problem sizes and depths uncovered a significant finding. Notably, when QAOA initializes from a single warm-start string, it demonstrates minimal progress. We provide a conceptual elucidation for this observation. Specifically, we identify the absence of an operator satisfying the assumptions of the Perron–Frobenius theorem while also possessing a superposition of classical states with identical objective function value as its ground states. This absence undermines the convergence guarantee of any warm-start QAOA variant to an optimal solution, even in the limit as the depth parameter approaches infinity ($p \to \infty$).

Consequently, the convergence of some warm-start QAOA variants to an optimal solution hinges entirely on the classical optimizer's ability to avoid being trapped in parameter sets.

III. OVERVIEW OF QAOA

Let $\mathbb{D}^n := \{0, 1, \dots, d-1\}^n$ be the set of n-element strings and $\mathcal{S}$ the group of permutations of these d^n elements. A classical optimization problem can be formulated as follows: given a function $F : \mathbb{D}^n \to \mathbb{R}$, find the elements in $\mathbb{D}^n$ on which it attains min (max) values. If a permutation $g \in \mathcal{S}$ is undetectable by F, i.e. $F(g(x)) = F(x)$ for any $x \in \mathbb{D}^n$, then g is symmetry of F. Such elements form a subgroup $G \subset \mathcal{S}$ and F is invariant with respect to this subgroup.

One of the widely employed algorithms for tackling the quantum version of the optimization problem is the QAOA, introduced in [6]. In the QAOA framework, the Hamiltonian H_F is commonly referred to as the *problem Hamiltonian* and is denoted by H_P (as per Farhi's et al. paper [6]). We will adopt this notation consistently.

Central to QAOA is the mixer Hamiltonian H_M, characterized by a distinct lowest energy state $|\xi\rangle \in W$ and adherence to the requirements of the Perron–Frobenius theorem (refer to Theorem III.1). The core idea behind the QAOA algorithm lies in iteratively transforming the mixer Hamiltonian H_M into the problem Hamiltonian. This process ensures that the image of the lowest energy vector from the preceding step becomes the lowest energy vector in the subsequent one.

The algorithm initiates by preparing the state $|\xi\rangle$, the ground state for the mixer Hamiltonian H_M, and then proceeds with multiple alternating applications of (certain exponents of) the problem and mixer Hamiltonians. The number of iterations is conventionally denoted by p (also known as QAOA depth), and we use $\mathfrak{Q}_p$ to express the entire composition of operators

$$\mathfrak{Q}_p := e^{-i\beta_1 H_M} e^{-i\gamma_1 H_P} \ldots e^{-i\beta_p H_M} e^{-i\gamma_p H_P}. \tag{3}$$

The final step of QAOA involves performing a measurement of the state obtained after applying $\mathfrak{Q}_p$ in the standard basis. For an in-depth description of the algorithm, we direct the reader to Section II and the references therein.

While the Hamiltonian H_P, representing the objective function, is uniquely determined by the classical problem, there is some flexibility in choosing the pair of mixer Hamiltonian and initial state. The convergence of QAOA to a classical state representing an element on which F attains a minimum value is guaranteed by the adiabatic theorem, provided the mixer Hamiltonian satisfies the conditions of the Perron-Frobenius theorem (see below and [14, Thm. 8.4.4]) and the initial state is the ground state for it.

Theorem III.1 (Perron–Frobenius): Let $M = (m_{ij}) \in \mathrm{Mat}_n(\mathbb{R})$ be an irreducible matrix with $m_{ij} \geq 0$.

1) Then, there is a positive real number r, such that r is an eigenvalue of M and any other eigenvalue λ (possibly complex) has $\mathrm{Re}(\lambda) < r$.
2) Moreover, there exists a unique real vector $v = (v_1, v_2, \ldots, v_n)$ such that $M(v) = rv$ and $v_1 + v_2 + \ldots + v_n = 1$. This vector is positive, i.e., all v_i are strictly greater than 0.

The standard and most common choice of mixer Hamiltonian involves Pauli X-gates and is given by $B = \sum_{0 \leq j \leq \ell-1} X_j$, where ℓ is the number of qubits needed for the (re)formulation of the original problem. The corresponding ground state is $|\xi\rangle = |+\rangle^{\otimes l}$. While this choice offers certain advantages, it does not consider any specific attributes of a given problem, in particular, the group of symmetries G.

IV. SYMMETRIES OF THE MIXERS

In this section, we offer a broad, high-level overview of our approach to selecting the mixer Hamiltonian based on symmetries inherent in the objective function of the optimization problem being addressed. A more comprehensive and conceptual discussion is deferred to the appendix.

When determining the symmetries, it is natural to start by considering the group $\mathcal{S}$ consisting of all permutations of the elements within the set of all d-element strings $\mathbb{D}^n$. This action naturally extends to an action on classical states, and by linearity, to the vector space W associated with $\mathbb{D}^n$. The group of classical symmetries for an optimization problem forms a subgroup G comprising elements $g \in \mathcal{S}$ that remain "undetectable" by F, meaning that $F(g(x)) = F(x)$ for any $x \in \mathbb{D}^n$. It is straightforward to observe that elements in this subgroup commute with the action of the problem Hamiltonian (representing F) on W. It is natural to seek a mixer Hamiltonian that satisfies the necessary technical requirements of the Perron–Frobenius theorem (see Theorem III.1), ensuring convergence as $p \to \infty$ and commuting with the largest subgroup of G, ideally encompassing the entire group G. Given that the latter condition implies that the corresponding unitary operator $\mathfrak{Q}_p$, which is the product of p alternating applications of mixer and problem Hamiltonian operators, commutes with G, it is natural to refer to the corresponding QAOA as G-equivariant.

Within $\mathcal{S}$, there exists a subgroup $S_d = \mathrm{Perm}(\mathbb{D})$, comprising permutations of elements within a single copy of the symbol set $\mathbb{D}$. This subgroup acts by simultaneously permuting elements of $\mathbb{D}^n$ in the same manner across all copies

$$g(d_1, d_2, \ldots, d_n) := (g \cdot d_1, g \cdot d_2, \ldots, g \cdot d_n).$$

In many optimization problems (as discussed in the following sections), the objective function contains S_d as a subgroup of its classical symmetries, i.e., $S_d \subseteq G$.

Remark IV.1 (Convention): The proposed construction of the mixer Hamiltonian applies to arbitrary values of d. However, for the sake of clarity in exposition and in view of practical realizations (such as circuit decompositions), we will henceforth assume that d is a power of two, i.e., $d = 2^\ell$ for some integer $\ell \geq 1$. This assumption simplifies implementation and simulation, while the underlying mathematical framework remains valid for general d.

In this case, two subgroups of S_d will play a fundamental role in our discussion. To describe them, it is convenient to consider the set $\mathbb{D}$ as a union of ℓ bits.

The group $K_\ell := \underbrace{\mathbb{Z}_2 \times \ldots \times \mathbb{Z}_2}_{\ell}$, which is a subgroup of $\mathcal{S}$, represents the bit flips for each of these bits. Meanwhile, $S_\ell \subset S_d$ is the subgroup responsible for permuting the bits.

In the appendix, we elaborate on the construction (and the reasoning behind it) of a mixer Hamiltonian H_M, whose action on W (the vector space corresponding to $\mathbb{D}^n$) commutes with the action of the entire group S_d. Importantly, H_M satisfies the assumptions of the Perron–Frobenius theorem.

Similar to the classical mixer Hamiltonian B, the operator H_M has a uniform superposition of all classical states

$$|\xi\rangle = \frac{1}{2^{n\ell}} H^{\otimes n\ell}(|\underbrace{00\ldots0}_{n\ell}\rangle) = |\underbrace{++\ldots+}_{n\ell}\rangle$$

as its unique ground state. However, we also highlight a significant difference between the two operators, B and H_M. Specifically, the action of the classical mixer Hamiltonian B only commutes with a smaller subgroup, which is the semidirect product of the groups K_ℓ and S_ℓ, and has an order of $2^\ell \cdot \ell!$. This is notably less than the order of S_d, which is $d! = 2^\ell!$, as demonstrated in Proposition B.4 and the subsequent Corollary B.6.

We now outline an explicit circuit-level implementation of the unitary operator $e^{-\beta H_M}$ corresponding to the mixer Hamiltonian H_M. For the derivation, justification, and a discussion of the relevant commutation properties, we refer the reader to the Appendix. A Qiskit implementation can be found in [27].

Step 1: Apply a Hadamard gate to each qubit, transforming the basis from standard to Hadamard.

Step 2: For each block of ℓ qubits encoding a single qudit (see Remark IV.1 for the convention $d = 2^\ell$), apply a controlled phase of -4β to the state $|0^{\otimes \ell}\rangle$ in the standard basis. This is done as follows.

- **2.1** Apply X gates to all ℓ qubits to map $|0\rangle$ to $|1\rangle$.
- **2.2** Apply a multicontrolled $P(-4\beta)$ gate to the ℓth qubit, controlled on the remaining $\ell - 1$ qubits being in the $|1\rangle$ state.
- **2.3** Apply X gates again to revert the basis to the original.

Step 3: Apply Hadamard gates to each qubit to return to the standard basis.

Remark IV.2: It follows from the construction above that the mixer Hamiltonian H_M can be implemented without any auxiliary (ancilla) qubits. The circuit fragment for applying a single instance of $e^{-\beta H_M}$ has depth 5 under the assumption that the multicontrolled phase gate on ℓ qubits is available as a primitive operation. However, in a setting where only single- and two-qubit gates are allowed, this gate must be decomposed into elementary operations, typically resulting in a circuit depth of order ℓ^2. In typical applications, the local dimension $d = 2^\ell$ is fixed, and thus ℓ is treated as a constant. In this case, the circuit depth required to implement $e^{-\beta H_M}$ remains $\mathcal{O}(1)$, even when decomposing multicontrolled phase gates into standard 1- and 2-qubit gates.

We proceed by examining the cyclic subgroup $\mathbb{Z}_d$ within S_d, generated by the element $g := (23\ldots n1)$, which cyclically shifts the elements from 1 to 2, 2 to 3, and n to 1. We then construct an operator H_χ whose action on W commutes with $\mathbb{Z}_d$ and has the state

$$|\psi\rangle := \frac{1}{2^{n\ell}} H^{\otimes n\ell}(|\underbrace{10\ldots0}_{n\ell}\rangle) = |\underbrace{-+\ldots+}_{n\ell}\rangle$$

as its unique ground state.

Remark IV.3: The ambient Hilbert space W admits a decomposition into a direct sum of subspaces

$$W = \bigoplus_{j=0}^{d-1} W_j$$

according to the $\mathbb{Z}_d$-action. It is interesting to note that the state vector $|\xi\rangle$ resides in W_0, while $|\psi\rangle$ is located in $W_{d/2}$. Moreover, the images of these vectors during the execution of their respective QAOAs remain within these subspaces prior to the final projection (see Remark A.3 for a precise statement).

Let us reiterate that we defer the verification of the existence of the operators H_M and H_χ satisfying the aforementioned properties to the appendix. Instead, our focus in the subsequent sections will be on demonstrating its practical advantages over the classical mixer.

V. OUTLINE OF THE TWO PROBLEMS

In this section, we elucidate two significant classical optimization problems and their reformulations within the framework of QAOA. These problems find numerous applications across various domains [2], [15].

A. PROBLEM 1: EDGE COLORING

One class of optimization problems with objective function having the aforementioned group of symmetries, S_d, is coloring of the vertices or edges of a graph in d colors.

Definition V.1: A *vertex coloring* of a graph $\Gamma = (V, E)$ is a map $\widetilde{C} : E \to \mathfrak{C}$, where $\mathfrak{C}$ is a set of colors with $|\mathfrak{C}| = d$. A coloring $\widetilde{C}$ is called *proper* if $\widetilde{C}(v_1) \neq \widetilde{C}(v_2)$ for any two adjacent vertices $v_1, v_2 \in V$.

Similarly, an *edge coloring* of a graph $\Gamma = (V, E)$ is a map $C : E \to \mathfrak{C}$. A coloring C is called *proper* if $C(e) \neq C(f)$ for any two adjacent edges $e, f \in E$.

To represent d colors, we employ $\ell = \ell og_2(d)$ bits through the following encoding:

$$\begin{aligned} \text{color}_0 &\longleftrightarrow 0\ldots00 \\ \text{color}_1 &\longleftrightarrow 0\ldots01 \\ &\ldots. \end{aligned}$$

In this section, we focus on the edge coloring. Each edge $e \in E$ is assigned ℓ bits $e_0, e_1, \ldots, e_{\ell-1}$ whose values uniquely determine the color of the edge. The characteristic function of a color $C \in \mathfrak{C}$ is defined as follows:

$$\chi_c(c') := \begin{cases} 1, & \text{if } c'_i \equiv c_i \text{ for all } i \in \{1, \ldots, \ell\} \\ 0, & \text{otherwise.} \end{cases}$$

This function, denoted as $\chi_c(C(e))$, is explicitly given by

$$\chi_c(C(e)) = \prod_{i=1}^{\ell}(c_i e_i + (1-c_i)(1-e_i)).$$

This defining property ensures that the characteristic function equals 1 on the specific color C and 0 on all other colors. We define the objective function

$$F_\Gamma(C) := \sum_{e\bullet f}\sum_{c\in\mathfrak{C}} \chi_c(C(e))\chi_c(C(f))$$

where the notation $e \bullet f$ represents adjacent edges. This function calculates the number of adjacent edges with the same color.

Remark V.2: A coloring C is proper if and only if $F_\Gamma(C) = 0$.

It is evident that the action of the group S_d, permuting the colors, preserves the values of the objective function

$$F_\Gamma(\sigma^{-1}(C)) = F_\Gamma(C) \quad \forall\sigma \in S_d, C \in \mathfrak{C}.$$

Definition V.3: The *chromatic index* χ_Γ of a graph Γ is the minimum number of colors needed for a proper coloring of Γ.

The following result was proved in [30].

Theorem V.4: Let Γ be a simple undirected graph with maximum degree $\triangle(\Gamma)$. Then, $\triangle(\Gamma) \le \chi(G) \le \triangle(\Gamma) + 1$.

Definition V.5: Graphs that can be colored with $\triangle(\Gamma)$ are called *class one* graphs. Graphs that require at least $\triangle(\Gamma) + 1$ colors are called *class two* graphs.

In order to resolve the dichotomy in Theorem V.4, whether the minimal proper coloring of edges involves k or $k+1$ colors, it suffices to find out if a proper k coloring exists.

The operator representing the characteristic function χ_c is given by

$$\widetilde{\chi}_c(e) := \begin{cases} e, & e_i \equiv c_i \;\; \forall i \in \{1,\dots,\ell\} \\ 0, & \text{otherwise}\end{cases}$$

and is expressed as

$$\widetilde{\chi}_c(e) = \frac{1}{2^\ell}\bigotimes_{i=1}^{\ell}(1 + (-1)^{c_i} Z_{e,i}).$$

In the case $\ell = 2$, this expression simplifies to

$$Z_{e,0}Z_{f,0}Z_{e,1}Z_{f,1} + Z_{e,0}Z_{f,0} + Z_{e,1}Z_{f,1} + \lambda 1.$$

Meanwhile, the problem Hamiltonian representing F_Γ is

$$H_P = \sum_{e\bullet f}\sum_{c\in\mathfrak{C}} \widetilde{\chi}_c(e)\widetilde{\chi}_c(f).$$

The building blocks for the quantum circuit representing the exponent of the latter operator

$$\begin{aligned} e^{-i\beta H_P} &= \prod_{e\bullet f} e^{-i\beta(Z_{e,0}Z_{f,0}Z_{e,1}Z_{f,1} + Z_{e,0}Z_{f,0} + Z_{e,1}Z_{f,1})} \\ &= \prod_{e\bullet f}(e^{-i\beta Z_{e,0}Z_{f,0}Z_{e,1}Z_{f,1}} e^{-i\beta Z_{e,0}Z_{f,0}} e^{-i\beta Z_{e,1}Z_{f,1}}) \end{aligned} \tag{4}$$

are presented in Fig. 1 below.

B. PROBLEM 2: GRAPH PARTITIONING

The second optimization problem explored in this article is the balanced graph partitioning problem. This problem appears in numerous applications [2] and has been a subject of several investigations in QAOA and other frameworks [23], [29]. Given a graph Γ and a fixed integer k that divides the number of vertices in Γ, the objective is to find a partition of the vertices: $V = V_0 \sqcup V_1 \sqcup \dots \sqcup V_{k-1}$ into k disjoint subsets of equal cardinality that minimizes the total number of cut edges. A cut edge is defined as an edge with endpoints in different subsets. The requirement of exact equality of sizes of V_i's for all i is often referred to as perfectly balanced graph partitioning.

This problem bears some resemblance to the vertex coloring problem for k colors. Specifically, we can refer to vertices in subset V_i as colored with the ith color. However, unlike the coloring problem where we aim to minimize the number of adjacent vertices with the same color, here we seek to maximize this number. In addition, we must account for the restriction on the cardinalities of the V_i's.

We will examine examples with $k = 4$ and the number of vertices in the graph being a multiple of 4. As before, we encode the four colors using 2 bits. We define the objective function $F(C)$ as follows:

$$\begin{aligned} F(C) = &-\sum_{v-v'}\sum_{c\in\mathfrak{C}}\chi_c(C(v))\chi_c(C(v')) + \left(2E\sum_{v\in V}(v_0 - 0.5)\right)^2 \\ &+ \left(2E\sum_{v\in V}(1-v_0)(v_1 - 0.5)\right)^2 \\ &+ \left(2E\sum_{v\in V}v_0(v_1 - 0.5)\right)^2 \end{aligned}$$

where the notation $v - v'$ is used for adjacent vertices. The first sum evaluates the number of pairs of adjacent vertices belonging to different subsets of the partition, while the remaining three ensure that $|V_0| = |V_1| = |V_2| = |V_3| = \frac{|V|}{4}$. Specifically, $(\sum_{v\in V}(v_0 - 0.5))^2$ equals zero if and only if the numbers of vertices with the first color bit equal to 0 and 1 coincide; otherwise, it is positive. Similarly, $(\sum_{v\in V}(1 - v_0)(v_1 - 0.5))^2$ and $(\sum_{v\in V} v_0(v_1 - 0.5))^2$ equal zero if and only if the numbers of vertices with the second color bit equal to 0 and 1 coincide, respectively, for the first color bit being fixed at 0 and 1.

The corresponding problem Hamiltonian is given by

$$H_P = -\sum_{v-v'}\sum_{c\in\mathfrak{C}} \widetilde{\chi}_v(c)\widetilde{\chi}_{v'}(c) + E\left(\sum_{v\in V} Z_{v,0}\right)^2 + E\left(\sum_{v\in V}(1-Z_{v,0})Z_{v,1}\right)^2 + E\left(\sum_{v\in V}(1-Z_{v,0})Z_{v,1}\right)^2$$

and is equivalent to

$$-\sum_{v-v'}\sum_{c\in\mathfrak{C}} \widetilde{\chi}_v(c)\widetilde{\chi}_{v'}(c) + 2E\sum_{v,v'\in V}(Z_{v,0}Z_{v',0} + Z_{v,1}Z_{v',1}) + 2E\sum_{v,v'\in V} Z_{v,0}Z_{v',0}Z_{v,1}Z_{v',1}.$$

VI. DUEL: EQUIVARIANT H_M, H_χ VS CLASSICAL B

In this section, we contrast the performance of QAOA algorithms using different mixer Hamiltonians: the classical one, $B = \sum X_i$, and the newly introduced equivariant H_M and H_χ. We analyze their effectiveness on the problems discussed in the preceding section, primarily comparing H_M and H_χ with H_B. We implement the algorithms iteratively. The algorithms begin by establishing the initial state

$$|\xi\rangle = \frac{1}{2^{n\ell}} H^{\otimes n\ell}(|\underbrace{00\ldots0}_{n\ell}\rangle) = |\underbrace{++\ldots+}_{n\ell}\rangle$$

for H_M and H_B, or

$$|\psi\rangle = \frac{1}{2^{n\ell}} H^{\otimes n\ell}(|\underbrace{10\ldots0}_{n\ell}\rangle) = |\underbrace{-+\ldots+}_{n\ell}\rangle$$

for H_χ. The initial pair of parameters (β_1, γ_1) is randomly selected from the uniform distribution on the set $[0, 0.25\pi] \times [0, 2\pi]$. Subsequently, the algorithm iterates through runs: after completing the $p = 1$ run, optimal values (β_1^*, γ_1^*) are determined with the aid of a classical optimizer. The subsequent QAOA run is then executed with starting parameters $(\beta_1^*, \gamma_1^*, 0, 0)$ for $p = 2$, and this process continues iteratively. The objective of the classical optimizer is to minimize the *energy*, which is defined as the average value of the objective function on the states output by the algorithm over multiple runs

$$\mathcal{E}_p := \frac{\sum_{i=1}^m F_\Gamma(\mathfrak{Q}_p(|s_i\rangle))}{m}. \quad (5)$$

Remark VI.1: In case of the edge coloring problem, if the energy $\mathcal{E}_p < 1$, it implies that at least one of the obtained values $F_\Gamma(\mathfrak{Q}_p(|s_i\rangle))$ is zero. Consequently, the corresponding coloring is proper, indicating that Γ is a class one graph.

On each successive step, the starting parameters consist of the values converged by the classical optimizer on the preceding step, complemented by two zeros for the additional angles that did not appear in the previous step. This deliberate choice ensures that the energy values $\mathcal{E}_1, \mathcal{E}_2, \ldots$ obtained in subsequent steps are nonincreasing, as outlined in [6]. The algorithm's depth for the edge coloring problem was set at $p = 9$ and for the graph partitioning problem at $p = 7$. We conducted multiple independent simulations, ranging from 50 to 56, for various graphs (see Figs. 2–4) using the qiskit codes available at [27].

TABLE 1. QAOA Performance Comparison for Edge Coloring Problem

Graph	Mean	Median	Min	$\mathcal{E}_9 < 1$
Γ_1, B	0.726	0.7056	0.3584	41/50
Γ_1, H_M	0.5692	0.4673	0.1923	47/50
Γ_1, H_χ	0.5726	0.5142	0.1621	47/50
Γ_2, B	0.9696	0.9316	0.4814	33/56
Γ_2, H_M	0.7437	0.7388	0.3691	51/56
Γ_2, H_χ	0.8688	0.7148	0.3964	47/56
Γ_3, B	1.2495	1.2417	0.6533	11/56
Γ_3, H_M	0.9344	0.8857	0.3691	35/56
Γ_3, H_χ	0.7334	0.6763	0.2598	50/56
Γ_4, B	1.4857	1.5313	0.7382	6/56
Γ_4, H_M	1.1959	1.1074	0.5117	21/56
Γ_4, H_χ	1.2415	1.1489	0.4395	20/56
Γ_5, B	1.3469	1.3066	0.6162	14/50
Γ_5, H_M	0.9149	0.9507	0.3516	30/50
Γ_5, H_χ	0.94123	0.9375	0.2939	27/50
Γ_6, B	0.8726	0.8569	0.502	23/28
Γ_6, H_M	0.5227	0.5073	0.17	28/28

TABLE 2. QAOA Performance Comparison for Graph Partitioning Problem

Graph	Mean	Median	Min
$\mathfrak{G}, B$	10.08135	10.34229	6.24414
$\mathfrak{G}, H_M$	8.05236	8.11035	4.47656
$\mathfrak{G}, H_\chi$	9.039	8.888	4.94434

TABLE 3. Table of *p*-Values for Student's *t*-Test

Graph/Mixer	H_M	H_χ
Γ_1	0.01252	0.007
Γ_2	$3.054 \cdot 10^{-7}$	0.237
Γ_3	$1.9807 \cdot 10^{-6}$	$4.0636 \cdot 10^{-15}$
Γ_4	0.0033	0.0169
Γ_5	$8.1731 \cdot 10^{-7}$	$5.9511 \cdot 10^{-5}$
Γ_6	$1.2231 \cdot 10^{-8}$	
$\mathfrak{G}$	$9.125 \cdot 10^{-7}$	0.0145

The main characteristics of the outcomes are summarized in Table 1 and Table 2, while the histograms displaying the average $\mathcal{E}_p$-values across sample runs of the equivariant algorithms, as compared to the classical one for each graph, are depicted in Figs. 5 and 6.

Based on Table 3, which presents the Student's t-test values for testing the hypothesis that the means of energy values for the two algorithms are equal, we reject this hypothesis at a significance level of $\alpha = 1.5\%$ for all graphs analyzed (with the exception of graph Γ_2 for the mixer H_χ). This indicates a statistically significant difference in the energy values between the algorithms across all examined graphs. Furthermore, we consistently observe lower median and minimal energy values (columns 3 and 4) for the newly proposed mixers.

VII. MIXER HAMILTONIANS FOR WARM-START QAOA

The standard QAOA typically begins in the uniform superposition of all classical bit strings. Its primary objective is to

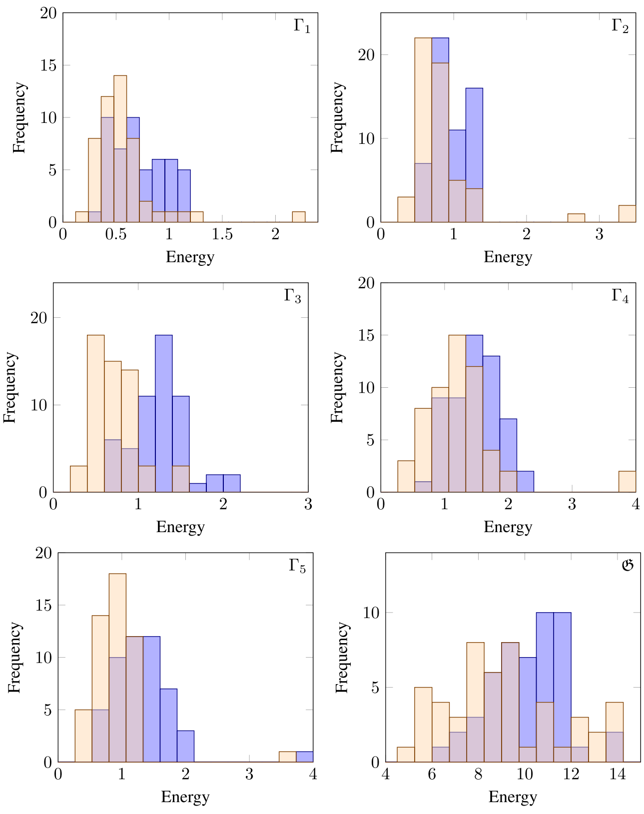

FIGURE 6. **Histograms illustrating the frequency distributions of $\mathcal{E}p$-values for algorithms utilizing mixers H_M and H_χ.**

enhance the objective function's value beyond the expected value in this initial state. A natural extension involves running a classical algorithm to generate a promising string (i.e., a good solution for certain practical goals, e.g., optimizing time/quality tradeoff), then initializing the QAOA in the corresponding computational basis state to seek further improvement. This approach, known as warm-start QAOA, has been explored in various studies.

In a recent paper [3], extensive numerical experiments involving both small and large instances at varying depths revealed a notable observation. Specifically, when the QAOA commences from a single warm-start string (or superposition of strings with equal energy level), it exhibits negligible progress. The authors further emphasize that these results hold even when QAOA is initialized with a single classical string, and the unitary operators used in QAOA are independent of the specific initial string chosen.

We aim to further explore this topic by offering additional insights into the limitations of warm-start QAOA. We start by observing that any mixer Hamiltonian with a nontrivial spectral gap possesses a one-dimensional eigenspace corresponding to the smallest eigenvalue λ. Let $|s\rangle$ denote a state spanning this subspace. It follows that $|s\rangle$ cannot be an eigenvector for the problem Hamiltonian H_P. If it were, both Hamiltonians would merely scale $|s\rangle$, and executing the corresponding QAOA starting with the state $|s\rangle$ would result in the identical state (up to a phase). Subsequently, measuring in the standard basis would yield a standard state with the same energy as the starting one.

In addition, as emphasized in the second assertion of Theorem III.1, every irreducible matrix with nonnegative values ensures that the vector corresponding to the highest eigenvalue has coordinates that are all nonzero (positive) in the standard basis.

This crucially implies that the ground state for a mixer Hamiltonian, satisfying the assumptions of the Perron–Frobenius theorem, *must be* a superposition of all classical states with nonzero amplitudes. Consequently, the standard argument for guaranteeing the convergence of QAOA as $p \to \infty$ to an optimal classical solution is inapplicable unless the initial state is a superposition of all classical states with nonzero amplitudes.

APPENDIX
PROOFS AND TECHNICAL DETAILS

In this section, we introduce and detail the construction, matrix representation, and quantum circuit for the newly introduced mixer Hamiltonian H_M, which is employed throughout this article. We demonstrate that H_M satisfies the Perron–Frobenius theorem, ensuring the convergence of the corresponding QAOA as the number of iterations, p, tends to infinity. In addition, we revisit essential definitions and provide a concrete representation of the result concerning subgroups of $\mathcal{S}$ that commute with the actions of the operators H_M and B on W.

A. NEW MIXERS:H_M AND H_χ

The symmetric group S_d discussed in previous sections is also known as $W(U_d)$, the Weyl subgroup of the unitary group acting collectively on all qudits. A mixer Hamiltonian H_M, which commutes with this group's action, can be constructed as follows.

Consider the sum of all transpositions, $\zeta = \sum_{1\leq i<j\leq n}(ij) \in \mathbb{C}[S_d]$, where $\mathbb{C}[S_d]$ denotes the group algebra of S_d. The group algebra is a vector space with a basis indexed by group elements, where multiplication is defined by the group operation of the underlying group. This element ζ commutes with all permutations and, therefore, resides in the center of the group algebra. The matrix representation of ζ in the standard basis of a vector space representing a single qudit is given by $\widehat{H}_{M_{ij}} = 1$ for $i \neq j$, and $\widehat{H}_{M_{ii}} = \binom{d-1}{2}$. A notable practical observation is that, in the Hadamard basis, this matrix becomes diagonal

$$H^{\otimes \ell}\widehat{H}_M H^{\otimes \ell} = \mathrm{diag}\left(\frac{d(d-1)}{2}, \nu, \dots, \nu\right)$$

where $\nu := \frac{(d-1)(d-2)}{2} - 1$, or, ignoring the addition of a scalar $(\frac{(d-1)(d-2)}{2} - 1) \cdot \mathrm{Id}$ operator

$$H^{\otimes \ell}\widehat{H}_M H^{\otimes \ell} = \mathrm{diag}(d, 0, \dots, 0)$$

resulting in $e^{-\beta H^{\otimes \ell}\widehat{H}_M H^{\otimes \ell}} = \mathrm{diag}(e^{-d\beta}, 1, \dots, 1)$.

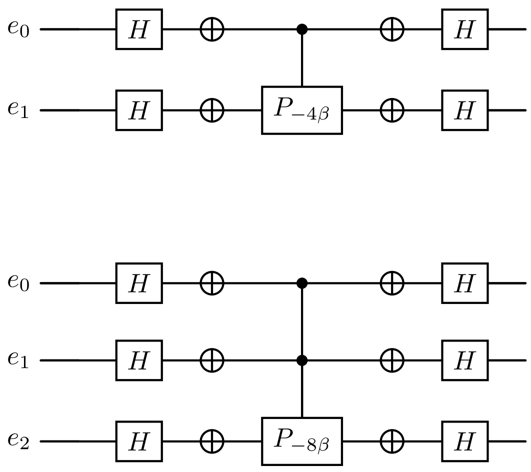

FIGURE 7. **Quantum circuit for $e^{-\beta H^{\otimes\ell}\widehat{H}_M H^{\otimes\ell}}$ with $d = 4$ and $d = 8$.**

Example A.1: The quantum circuits for $e^{-\beta H^{\otimes\ell}\widehat{H}_M H^{\otimes\ell}}$ with $d = 4$ and $d = 8$ are illustrated on Fig. 7.

We define the mixer Hamiltonian H_M as

$$H_M := \sum_{i=1}^{n} \widehat{H}_M^i \tag{6}$$

the sum of individual terms $\widehat{H}_M^i$, where

$$\widehat{H}_M^i := Id \otimes \ldots \otimes Id \otimes \widehat{H}_M \otimes Id \otimes \ldots \otimes Id$$

represents the action of H_M on the i^{th} copy of the vector space V corresponding to the ith copy of the set $\mathbb{D}$. We would like to remind the reader that $W = V \otimes \ldots \otimes V$ is the n-fold tensor product of such vector spaces.

It is straightforward to verify that the operator H_M defined in this manner satisfies the assumptions of the Perron-Frobenius theorem (see Theorem III.1), thereby qualifying as a mixer Hamiltonian. Notably, its ground state $|\xi\rangle$ coincides with that of the classical mixer B.

This time, we start with the element

$$\eta = \sum_{1\le i<j\le n} (-1)^{i+j}(ij) \in \mathbb{C}[S_d].$$

We then examine the cyclic subgroup $\mathbb{Z}_d \subset S_d$, generated by the element $g = (23\ldots n1)$, which cyclically permutes the elements from 1 to n.

Lemma A.2: The element η commutes with the group $\mathbb{Z}_d$.

Proof: To demonstrate this, we calculate $g\eta g^{-1} = \sum_{1\le i<j\le n}(-1)^{i+j}(g(i)g(j)) = \sum_{1\le i<j\le n}(-1)^{i+j}((i+1) \bmod d)((j+1) \bmod d) = \sum_{1\le i<j\le n}(-1)^{i+j}(ij) = \eta$, where $i + j + 2 \equiv i + j \pmod 2$. Hence, $g\eta g^{-1} = \eta$ implies $g\eta = \eta g$, indicating that η and g commute. □

The matrix representation of η in the standard basis of a vector space representing a single qudit is given by $\widehat{H}_{\chi_{ij}} = (-1)^{i+j}$ for $i \neq j$, and $\widehat{H}_{\chi_{ii}} = \binom{d-1}{2}$. Furthermore, in the Hadamard basis, this matrix becomes diagonal

$$H^{\otimes\ell}\widehat{H}_\chi H^{\otimes\ell} = \mathrm{diag}\left(\nu, \ldots, \nu, \frac{d(d-1)}{2}, \nu, \ldots, \nu\right)$$

or, ignoring the addition of a scalar $(\frac{(d-1)(d-2)}{2} - 1) \cdot \mathrm{Id}$ operator

$$H^{\otimes\ell}\widehat{H}_\chi H^{\otimes\ell} = \mathrm{diag}(0, \ldots, 0, d, 0, \ldots, 0)$$

acting with multiplication by d on the 1-D vector space spanned by the vector $|-\underbrace{+\ldots++}_{\ell-1}\rangle$ Resulting in $e^{-\beta H^{\otimes\ell}\widehat{H}_\chi H^{\otimes\ell}} = \mathrm{diag}(1, \ldots, 1, e^{-d\beta}, 1, \ldots, 1)$.

We define the mixer Hamiltonian H_χ by the formula

$$H_\chi := \widehat{H}_\chi \otimes Id \otimes \ldots \otimes Id + \sum_{i=2}^{n} \widehat{H}_M^i. \tag{7}$$

The operator H_χ, defined in this manner, does not meet the assumptions of the Theorem III.1. Nevertheless, it is straightforward to verify that it possesses a 1-D eigenspace with the minimal eigenvalue, and therefore, a nonzero spectral gap. This eigenspace is spanned by the state $|\psi\rangle := |-\underbrace{+\ldots++}_{n\ell-1}\rangle$.

Let $\zeta := e^{\frac{2\pi i}{d}}$ be the primitive dth root of unity. Under the action of the cyclic group $\mathbb{Z}_d$, the Hilbert space W decomposes into a direct sum of vector spaces: $W = \bigoplus_{j=0}^{d-1} W_j$, where each W_j has dimension d^{n-1}. The action of $\mathbb{Z}_d$ on W_j is given by the equation $g \cdot w_j = \zeta^j w_j$, for all $w_j \in W_j$.

It is worth noting that $\zeta^{d/2} = e^{\pi i} = -1$, and the vector $|\psi\rangle$ resides in $W_{d/2}$.

Remark A.3: Since the actions of both operators H_P and H_χ on W commute with that of the group $\mathbb{Z}_d$, it follows that the operator $\mathfrak{Q}_p$ preserves each subspace W_j, meaning $\mathfrak{Q}_p(W_j) \subseteq W_j$. Specifically, this implies $\mathfrak{Q}_p(|\xi\rangle) \subseteq W_0$ and $\mathfrak{Q}_p(|\psi\rangle) \subseteq W_j$.

B. GROUP ACTIONS AND MIXERS

Recall that the symmetric group $\mathcal{S}$ acts on the set of all states $\mathbb{D}^n$ by permutations. This action can be uniquely extended to a linear action on the state vector space W. Said differently, there is a homomorphism $\varphi : \mathcal{S} \to GL(W)$.

We elucidate key properties concerning the interaction of S_d, K_ℓ, and S_ℓ (see Section III for the definitions of these groups) with the objective function F and the Hamiltonians H_P, B, and H_M.

Suppose $A : V \to V$ is a linear operator. We denote by $Z_{S_d}(A)$ the subgroup of elements in S_d whose action on V commutes with that of A.

Proposition B.4: Suppose the objective function $F : \mathbb{D}^n \to \mathbb{R}$ is invariant with respect to the action of symmetric group S_d.

a) $Z_{S_d}(H_P) = S_d$.
b) $Z_{S_d}(H_M) = S_d$.
c) $Z_{S_d}(B) = K_\ell \rtimes S_\ell$, where $K_\ell \lhd Z_{S_{2^\ell}}(B)$ is normal.

Proof: The statement in (a) is an immediate consequence of the initial assumption. The assertion in (b) follows from the fact that the element $\sum_{1\le i<j\le d}(ij)$ is in the center of the group algebra $\mathbb{C}[S_d]$. We focus on justification for the statement in (c).

As the action of S_d is identical on all qudits, it suffices to verify the assertion for $n = 1$. In this case the standard mixer Hamiltonian can be written as

$$B = \sum_{i=1}^{\ell} X_i = \sum_{i=1}^{\ell} \varphi(\sigma_i)$$

where $\sigma_i \in K_\ell$ denotes the involution that flips the ith bit in a binary string $b \in \mathbb{B}^\ell = \mathbb{D}$, i.e., $\sigma_i(b) = b_{\hat{i}}$, where $b_{\hat{i}}$ is b with the ith bit flipped.

Let $g \in \mathcal{S}$ be a permutation of $\mathbb{B}^\ell$ such that $\varphi(g)$ commutes with B. Then

$$\varphi(g)B\varphi(g)^{-1} = B \quad \iff \quad \sum_{i=1}^{\ell} \varphi(g\sigma_i g^{-1}) = \sum_{i=1}^{\ell} \varphi(\sigma_i).$$

This implies that the set $\{g\sigma_i g^{-1}\}_{i=1}^{\ell}$ must coincide with $\{\sigma_i\}_{i=1}^{\ell}$. In particular, each conjugate $g\sigma_i g^{-1}$ must again act as a single-bit flip on some bit.

To see this, note that the action of σ_i on $\mathbb{B}^\ell$ is an involution swapping each string b with $b_{\hat{i}}$, differing from b in exactly one bit. Since σ_i is a bit flip and conjugation preserves cycle structure of permutations, $g\sigma_i g^{-1}$ is the involution that swaps $g(b)$ with $g(b_{\hat{i}})$. Moreover, as $g\sigma_i g^{-1}$ is in $\{\sigma_i\}_{i=1}^{\ell}$, it must also flip exactly one bit of its input. Hence, $g(b)$ and $g(b_{\hat{i}})$ must differ in exactly one bit. Therefore, conjugation by g permutes the σ_i, i.e.,

$$g\sigma_i g^{-1} = \sigma_{\pi(i)}$$

for some permutation $\pi \in S_\ell$, and hence g belongs to the normalizer of the subgroup K_ℓ, generated by $\{\sigma_i\}$.

Define $\tau := \pi^{-1} \in S_\ell$, and let $h := \tau \cdot g \in S_d$, then

$$h\sigma_i h^{-1} = \tau g\sigma_i g^{-1}\tau^{-1} = \tau\sigma_{\pi(i)}\tau^{-1} = \sigma_{(\tau\circ\pi)(i)} = \sigma_i.$$

Hence, $h \in \mathrm{Stab}_{S_d}(\sigma_i)$ is in the stabilizer of σ_i, which is precisely $K_\ell \rtimes S_\ell$, since all elements of S_d that fix every σ_i under conjugation must preserve the bit structure and act by bit flips and permutations.

Therefore, $g = \tau^{-1}h \in S_\ell \cdot (K_\ell \rtimes S_\ell) = K_\ell \rtimes S_\ell$, proving that the normalizer is contained in $K_\ell \rtimes S_\ell$. The reverse inclusion is clear, since $K_\ell \rtimes S_\ell$ normalizes K_ℓ by construction. Thus

$$\mathrm{Norm}_{S_d}(K_\ell) = K_\ell \rtimes S_\ell.$$

Finally, K_ℓ is normal in $K_\ell \rtimes S_\ell$, as the symmetric group S_ℓ acts on K_ℓ by permuting the bit indices. We conclude that the centralizer of B in S_d is the subgroup $K_\ell \rtimes S_\ell$. □

Remark B.5: It is interesting to point out that $K_\ell \rtimes S_\ell$ is $W(B_\ell)$, the Weyl group for root system of type B_ℓ.

Corollary B.6: The subgroup of symmetries of the mixer H_M surpasses that of B. For instance, when $\ell = 2$, $W(B_2)$ equals the dihedral group of order $|D_4| = 8$, while $|S_4| = 24$. For $\ell = 3$, $|W(B_3)| = 48$, and $|S_8| = 8! = 40320$. As ℓ increases, the disparity in orders becomes more pronounced: $|W(B_\ell)| = 2^\ell \cdot \ell!$ and $|S_d| = 2^\ell!$.

REFERENCES

[1] E. Bourreau, G. Fleury, and P. Lacomme, "Mixer Hamiltonian with QAOA for max k-coloring: Numerical evaluations," 2022, *arXiv:2207.11520*, doi: 10.48550/arXiv.2207.11520.

[2] A. Buluç, H. Meyerhenke, I. Safro, P. Sanders, and C. Schulz, "Recent advances in graph partitioning," in *Algorithm Engineering: Selected Results and Surveys* (Lecture Notes in Computer Science, 9220). Berlin, Germany: Springer, 2016, pp. 117–158, doi: 10.1007/978-3-319-49487-6_4.

[3] M. Cain, E. Farhi, S. Gutmann, D. Ranard, and E. Tang, "The Qaoa gets stuck starting from a good classical string," 2022, *arXiv:2207.05089*, doi: 10.48550/arXiv.2207.05089.

[4] Y. Deller et al., "Quantum approximate optimization algorithm for qudit systems," *Phys. Rev. A*, 107, 2023, Art. no. 062410, doi: 10.1103/PhysRevA.107.062410.

[5] D. Egger, J. Mareček, and S. Woerner, "Warm-starting quantum optimization," *Quantum*, vol. 5, pp. 479, 2021, doi: 10.22331/q-2021-06-17-479.

[6] E. Farhi, J. Goldstone, and S. Gutmann, "A quantum approximate optimization algorithm," 2014, *arXiv:1411.4028*, doi: 10.48550/arXiv.1411.4028.

[7] E. Fontana et al., "Characterizing barren plateaus in quantum ansätze with the adjoint representation," *Nature Commun.*, vol. 15, 2024, Art. no. 7171, doi: 10.1038/s41467-024-49910-w.

[8] J. Golden, A. Bärtschi, D. O'Malley, and S. Eidenbenz, "Numerical evidence for exponential speed-up of QAOA over unstructured search for approximate constrained optimization," in *Proc. 2023 IEEE Int. Conf. Quantum Comput. Eng. (QCE)*, 2023, doi: 10.1109/QCE57702.2023.00063.

[9] J. K. Golden, A. Bärtschi, D. O'Malley, and S. Eidenbenz, "Numerical evidence for exponential speed-up of QAOA over unstructured search for approximate constrained optimization," in *Proc. 2022 IEEE Int. Conf. Quantum Comput. Eng.*, 2022, pp. 496–505, doi: 10.1109/QCE57702.2023.00063.

[10] L. C. G. Govia, C. Poole, M. Saffman, and H. K. Krovi, "Freedom of the mixer rotation axis improves performance in the quantum approximate optimization algorithm," *Phys. Rev. A*, vol. 104, no. 6, 2021, Art. no. 062428, doi: 10.1103/PhysRevA.104.062428.

[11] S. Hadfield, Z. Wang, B. O'Gorman, E. G. Rieffel, D. Venturelli, and R. Biswas, "From the quantum approximate optimization algorithm to a quantum alternating operator ansatz," *Algorithms*, vol. 12, no. 34, 2019, pp. 1–12, doi: 10.3390/a12020034.

[12] Z. He et al., "Alignment between initial state and mixer improves QAOA performance for constrained portfolio optimization," *npj Quantum Inform.*, vol. 9, 2023, Art. no. 121, doi: 10.1038/s41534-023-00787-5.

[13] D. Herman et al., "Quantum computing for finance," *Nature Rev. Phys.*, 5, no. 8, pp. 450–465, 2023, doi: 10.1016/j.revip.2019.100028.

[14] R. Horn and C. Johnson, *Matrix Analysis*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2013, doi: 10.1017/CBO9780511810817.

[15] T. R. Jensen and B. Toft, *Graph Coloring Problems*. Hoboken, NJ, USA: Wiley, 2011, doi: 10.1017/CBO9780511810817.

[16] I. Joseph et al., "Quantum computing for fusion energy science applications," *Phys. Plasmas*, vol. 30, no. 1, 2023, pp. 10501, doi: 10.1063/5.0123765.

[17] S. Kazi, M. Larocca, M. Farinati, P. Coles, M. Cerezo, and R. Zeier, "Analyzing the quantum approximate optimization algorithm: Ansätze, symmetries, and lie algebras," *PRX Quantum*, vol. 6, 2025, Art. no. 040345, doi: 10.1103/yfwq-yqmk.

[18] X. Liu, R. Shaydulin, and I. Safro, "Quantum approximate optimization algorithm with sparsified phase operator," in *Proc. IEEE Int. Conf. Quantum Comput. Eng.*, 2022, pp. 133–141, doi: 10.1109/QCE53715.2022.00032.

[19] M. E. S. Morales, J. Biamonte, and Z. Zimbor'as, "On the universality of the quantum approximate optimization algorithm," *Quantum Inf. Process.*, vol. 19, 2020, Art. no. 291, doi: 10.1007/s11128-020-02748-9.

[20] K. Okada, H. Nishi, T. Kosugi, and Y. Matsushita, "Systematic study on the dependence of the warm-start quantum approximate optimization algorithm on approximate solutions," *Sci. Rep.*, vol. 14,, no. 1, 2024, Art. no. 1167, doi: 10.1038/s41598-023-50406-8.

[21] C. Outeiral, M. Strahm, J. Shi, G. Morris, S. Benjamin, and C. Deane, "The prospects of quantum computing in computational molecular biology," *Wiley Interdiscipl. Rev.: Comput. Mol. Sci.*, 11, no. 1, 2021, Art. no. e1481, doi: 10.1002/wcms.1481.

[22] R. Shaydulin, S. Hadfield, T. Hogg, and I. Safro, "Classical symmetries and the quantum approximate optimization algorithm," *Quantum Inf. Process.*, vol. 20, 2021, Art. no. 359, doi: 10.1007/s11128-021-03298-4.

[23] R. Shaydulin, I. Safro, and J. Larson, "Multistart methods for quantum approximate optimization," in *Proc. IEEE High Perform. Extreme Comput. Conf.*, 2019, pp. 1–8, doi: 10.1109/HPEC.2019.8916288.

[24] R. Shaydulin, H. Ushijima-Mwesigwa, I. Safro, S. Mniszewski, and Y. Alexeev, "Network community detection on small quantum computers," *Adv. Quantum Technol.*, vol. 2,, no. 9, 2019, Art. no. 1900029, doi: 10.1002/qute.201900029.

[25] R. Shaydulin and S. Wild, "Exploiting symmetry reduces the cost of training QAOA," *IEEE Trans. Quantum Eng.*, 2, 2021, Art. no. 3101409, doi: 10.1109/TQE.2021.3066275.

[26] V. Sridhar, Y. Chen, B. Gard, E. Barnes, and S. Economou, "Adapt-QAOA with a classically inspired initial state," 2023, doi: 10.48550/arXiv.2310.09694.

[27] B. Tsvelikhovskiy, "QAOA mixer hamiltonians for optimization problems with S_d symmetries," (n.d.). [Online]. Available: https://github.com/BorisTsv/QAOA-Mixer-Hamiltonians-for-Optimization-Problems-with-S_d-Symmetries/tree/main

[28] B. Tsvelikhovskiy, I. Safro, and Y. Alexeev, "Symmetries and dimension reduction in quantum approximate optimization algorithm," 2023, *arXiv:2309.13787v2*, doi: 10.48550/arXiv.2309.13787.

[29] H. Ushijima-Mwesigwa, R. Shaydulin, C. Negre, S. Mniszewski, Y. Alexeev, and I. Safro, "Multilevel combinatorial optimization across quantum architectures," *ACM Trans. Quantum Comput.*, vol. 2, no. 1, pp. 1–29, 2021, doi: 10.1145/342560.

[30] V. G. Vizing, "On an estimate of the chromatic class of a p-graph," *Diskret. Analiz*, vol. 3, pp. 25–30, 1964.

[31] Z. Wang, N. Rubin, Jason M. Dominy, and E. Rieffel, "XY-mixers: Analytical and numerical results for the quantum alternating operator ansatz," *Phys. Rev. A*, 101, 2020, Art. no. 012320, doi: 10.1103/PhysRevA.101.012320.

[32] J R. Weggemans, A. Urech, and A. Rausch et al., "Solving correlation clustering with QAOA and a Rydberg qudit system: A full-stack approach," *Quantum*, vol. 6, 2022, Art. no. 687, doi: 10.22331/q-2022-04-13-687.

[33] Y. Zhang, R. Zhang, and A. Potter, "QED driven QAOA for network-flow optimization," *Quantum*, vol. 5, 2020, Art. no. 510, doi: 10.22331/q-2021-07-27-510.

[34] L. Zhu, H. Lun Tang, F. A. Calderon-Vargas, N. Mayhall, E. Barnes, and S. Economou, "Adaptive quantum approximate optimization algorithm for solving combinatorial problems on a quantum computer," *Phys. Rev. Res.*, vol. 4, no. 3, 2022, Art. no. 033029, doi: 10.1103/PhysRevResearch.4.033029.

Boris Tsvelikhovskiy received the B.S. and M.Sc. degrees in mathematics from the National Research University, Higher School of Economics, Moscow, Russia, and the Ph.D. degree in mathematics from Northeastern University, Singapore, in 2020.

He is a Visiting Assistant Professor with the Department of Mathematics, University of California, Riverside, CA, USA. He subsequently completed a Postdoctoral Fellowship with the University of Pittsburgh, Pittsburgh, PA, USA. His research interests include quantum computing, representation theory, and algebraic geometry.

Dr. Tsvelikhovskiy is also an active Referee for prominent academic journals, including the *Arnold Mathematical Journal* and IEEE TRANSACTIONS ON QUANTUM ENGINEERING.

Ilya Safro received the B.S. degree in mathematics and computer science from the Ben-Gurion University of Negev, Beersheba, Israel, and the M.S. and Ph.D. degrees in applied mathematics and computer science from the Weizmann Institute of Science, Rehovot, Israel, in 2008.

He was a Postdoctoral Researcher and Argonne Scholar with the Argonne National Laboratory and a faculty member with the School of Computing, Clemson University, Clemson, SC, USA. He is currently an Associate Professor and Associate Chair with the Department of Computer and Information Sciences, University of Delaware, Newark, DE, USA. He also has a joint affiliation with the Department of Physics and Astronomy. He has authored more than 130 articles and book chapters. His research interests include quantum computing, graph algorithms, large-scale optimization, and machine learning.

Dr. Safro is on the editorial boards of such journals as IEEE TRANSACTIONS ON QUANTUM ENGINEERING and *SIAM Journal on Scientific Computing*. His research is primarily funded by the National Science Foundation, Defense Advanced Resarch Projects Agency, Department of Energy, National Institutes of Health, and industry. He is a Cofounder of Academentory.com, a portal to connect students with research advisers and facilitate Ph.D. position search.

Yuri Alexeev (Senior Member, IEEE) received the Ph.D. degree in physical chemistry from Iowa State University, Ames, IA, USA while a graduate student in Mark Gordon's quantum chemistry group, in 2002.

He is a Principal Project Specialist with the Argonne National Laboratory and a Senior Scientist with the Department of Computer Science, University of Chicago, Chicago, IL, USA. He is also a Principal Investigator with the National Quantum center Q-NEXT. He became a Postdoctoral Fellow with Pacific Northwest National Laboratory and worked in the NWChem group led by Dr. Theresa Windus and later, he joined the Nobel Prize winner Dr. Martin Karplus' group with Harvard University, Cambridge, MA, USA, and Université de Strasbourg, Strasbourg, France. He has authored more than 100 research articles and book chapters. His research interests include development of quantum computing algorithms, error correction/mitigation techniques, and numerical simulators of quantum systems using high-performance computing on next-generation high-performance supercomputers.

Dr. Alexeev recent projects include the development of quantum chemistry and combinatorial optimization quantum algorithms for NISQ quantum computers. He is a member of the Chicago Quantum Exchange. He serves on the editorial board for IEEE TRANSACTIONS ON QUANTUM ENGINEERING. His research is primarily funded by the National Science Foundation, Defense Advanced Research Projects Agency, Department of Energy, and industry.