QA(O)A: How to exploit symmetries?

Boris Tsvelikhovskiy
University of California, Riverside

Joint work with

Ilya Safro (University of Delaware) and Yuri Alexeev (Argonne National Laboratory)

References to our preprints:
https://arxiv.org/abs/2309.13787
https://arxiv.org/abs/2405.07211

The problem

Let $\mathbb{D}^n := \{0, 1, \ldots, d-1\}^n$ be the set of n-element strings and $\mathcal{S}$ the group of permutations of these d^n elements.

Goal: given a function $F : \mathbb{D}^n \to \mathbb{R}$, find the elements in $\mathbb{D}^n$ on which it attains min (max) values.

The symmetries: level 1

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 called a **symmetry** of F.
Such symmetries form a subgroup $G \subseteq \mathcal{S}$. The set $\mathbb{D}^n$ can be written as a disjoint union of G-orbits:

$$\mathbb{D}^n = \bigsqcup_{j=1}^{m} \mathcal{O}_j.$$

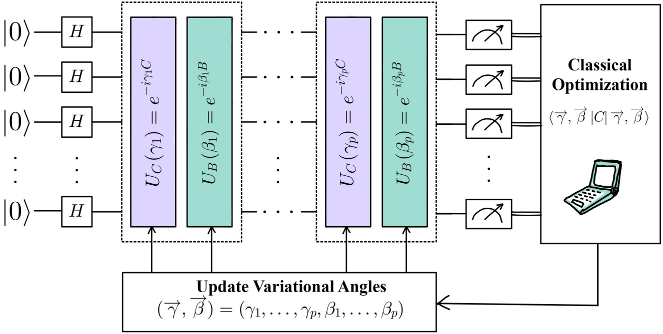
If the strings x and y are in the same G-orbit, then $F(x) = F(y)$.

Classical $\rightsquigarrow$ quantum

Let W be a vector space of dimension d^n with basis indexed by elements of $\mathbb{D}^n$, *standard basis.*
The Hamiltonian H_F is said to **represent** a function $F : \mathbb{D}^n \to \mathbb{R}$ if it satisfies $H_F(v_x) = F(x)v_x$ for any $x \in \mathbb{D}^n$.

- $\mathbb{D}^n \rightsquigarrow W$
- $F \rightsquigarrow$ linear operator H_F acting on W
- Minima of F on $\mathbb{D}^n \rightsquigarrow$ lowest energy states of H_F in W.

QAOA



$|0\rangle$
H
$U_C(\gamma_1) = e^{-i\gamma_1 C}$
$U_B(\beta_1) = e^{-i\beta_1 B}$
$U_C(\gamma_p) = e^{-i\gamma_p C}$
$U_B(\beta_p) = e^{-i\beta_p B}$
Classical Optimization
$\langle \vec{\gamma}, \vec{\beta} | C | \vec{\gamma}, \vec{\beta} \rangle$
Update Variational Angles
$(\vec{\gamma}, \vec{\beta}) = (\gamma_1, \ldots, \gamma_p, \beta_1, \ldots, \beta_p)$

QAOA: a closer look

While the Hamiltonian $\mathbf{H_P}$, which encodes the objective function, is **uniquely determined by the classical problem**, there is some **flexibility in selecting the mixer Hamiltonian**.

The most common choice of mixer Hamiltonian is $B = \sum_{0 \leq j \leq \ell - 1} X_j$.

The ground state for this mixer is the uniform superposition state $|\xi\rangle = |++\ldots+\rangle$.

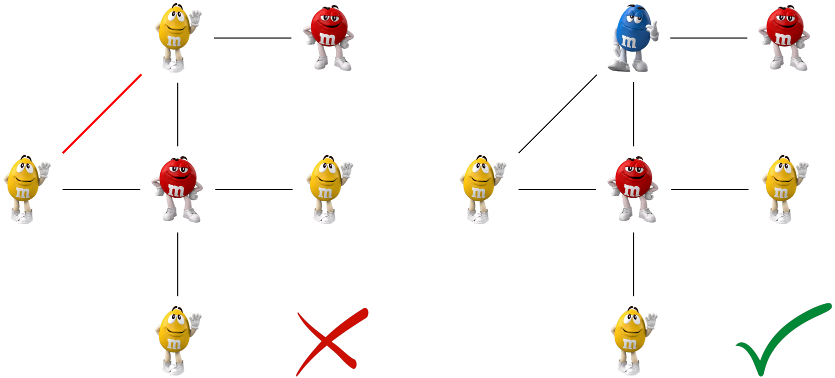
In many optimization problems, the objective function exhibits invariance under the action of the symmetric group $S_d = W(U_d)$,which acts collectively on all qudits as follows:

$$\sigma(x_1, x_2, \ldots, x_n) := (\sigma(x_1), \sigma(x_2), \ldots, \sigma(x_n)).$$

The **standard mixer Hamiltonian does not exhibit commutativity with the entire symmetric group** S_d, but only with a noticeably smaller subgroup. With this in mind, we explore an **alternative mixer Hamiltonian**, H_M, which maintains **commutativity** with the action of the **whole group**.

A concrete application

Consider the problem of coloring the vertices (edges) of a graph. A coloring is considered *proper* if no adjacent vertices (edges sharing a vertex) have the same color.



A concrete application

We will focus on the edge coloring problem. To each edge $e \in E$, one associates ℓ bits $e_0, e_1, \ldots, e_{\ell-1}$, the values of which uniquely determine its color.

The function χ_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 serves as the characteristic function of a color: it has value 1 on color c and 0 on all other colors.

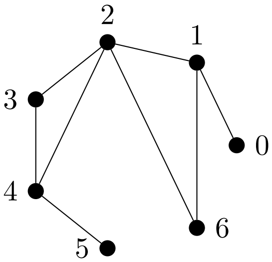
The objective function $F_\Gamma(C) := \sum_{e \bullet f} \sum_{c \in \mathfrak{C}} \chi_c(C(e))\chi_c(C(f))$ computes the number of adjacent edges of coinciding color.

Remark

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

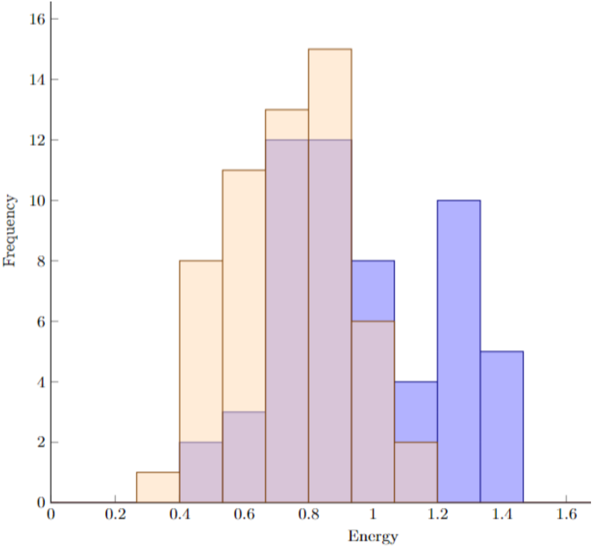
We would like to showcase a performance **comparison between two versions of QAOA** (standard and the newly proposed one) in determining appropriate edge colorings for the graph. Both algorithms are configured iteratively with a depth parameter of $p = 9$. Through over 50 independent trials for each scenario, we observe **statistically significant differences in mean values at the** 1.5% **significance level**, with the **new variant consistently demonstrating lower means**. Moreover, we note **considerably lower median and minimal values** in the experiments utilizing the **newly introduced mixer Hamiltonian** compared to the classical one.

Graph 1 (4 colors)

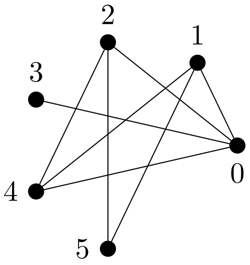


	Mean	Median	Min	Energy < 1
QAOA	0.9696	0.9316	0.4814	33/56
QAOA_{new}	0.7437	0.7388	0.3691	51/56

t-test p-value is $3.053993311768478 \cdot 10^{-7}$

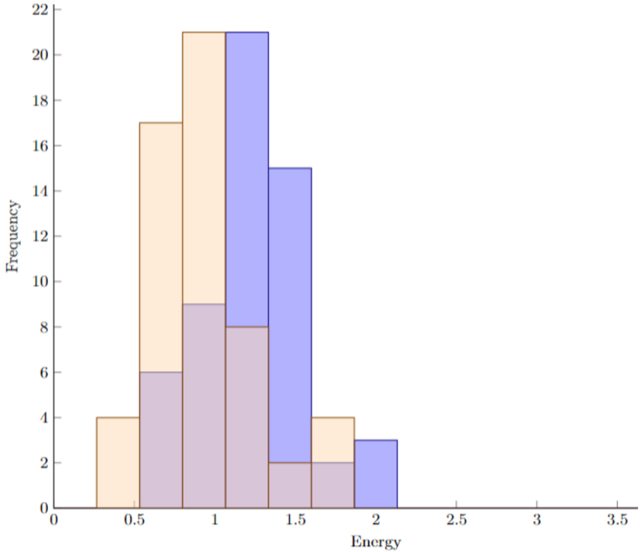

16
14
12
10
8
6
4
2
0
Frequency
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
Energy

Graph 2 (4 colors)

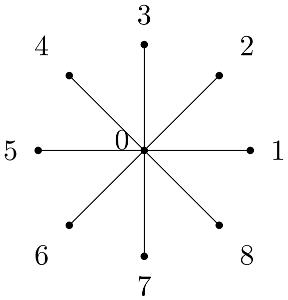


	Mean	Median	Min	Energy < 1
QAOA	1.2495	1.2417	0.6533	$11/56$
QAOA_{new}	0.9344	0.8857	0.3691	$35/56$

t-test p-value is $1.9806919304846427 \cdot 10^{-6}$

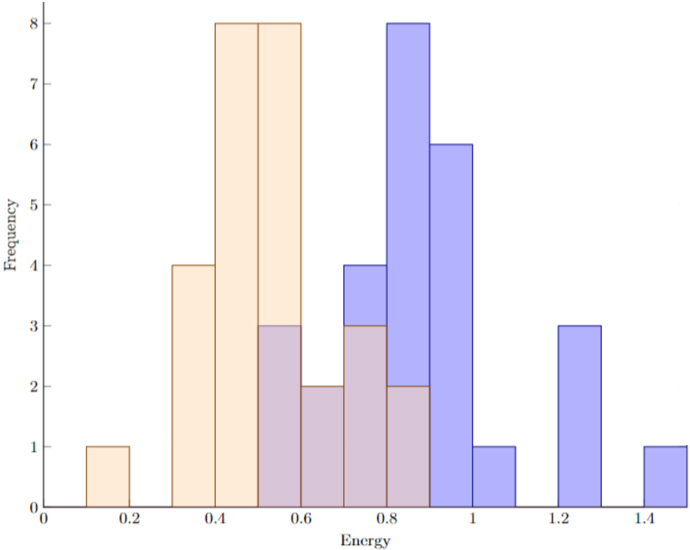

22
20
18
16
14
12
10
8
6
4
2
0
Frequency
0
0.5
1
1.5
2
2.5
3
3.5
Energy

Graph 3 (8 colors)



	Mean	Median	Min
QAOA	0.8726	0.8569	0.502
QAOA_{new}	0.5227	0.5073	0.17

t-test p-value is $1.2230598272375008 \cdot 10^{-8}$


Frequency
Energy
0
1
2
3
4
5
6
7
8
0
0.2
0.4
0.6
0.8
1
1.2
1.4

The symmetries: level 2

The action of G on the set $\mathbb{D}^n$ extends to an action on W.
Actions of G and H_P on W commute:

$$H_P(g(w)) = g(H_P(w)) \ \forall w \in W, \ \forall g \in G.$$

According to the G-action, W can be written as a direct sum of subspaces:

$$W = \bigoplus W_i.$$

The decomposition is preserved by H_P:

$$H_P(W_i) \subseteq W_i.$$

Here one of the subspaces is the **subspace of G-invariants**:

$$W^G = \{w \in W \mid g(w) = w, \forall g \in G\}.$$

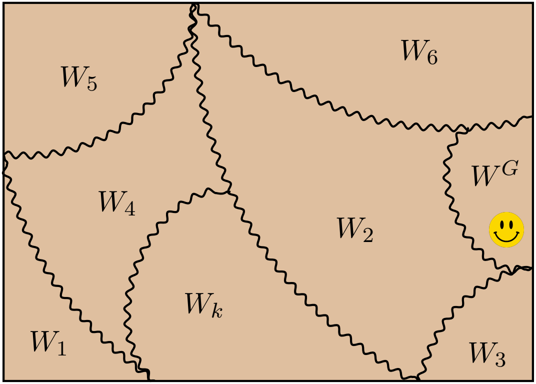


Figure: Decomposition of $W = V^{\otimes n}$

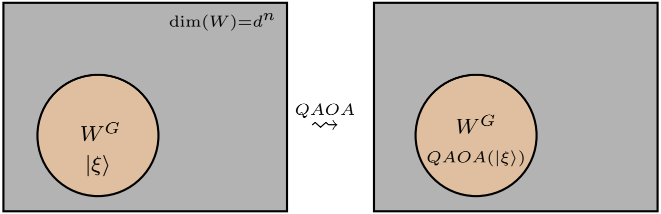
Where does QAOA 'live'?

Remark

Notice that the uniform superposition $|\xi\rangle = |++\ldots+\rangle$ is inside W^G for any $G \subseteq \mathcal{S}$.

Suppose we employ a QAOA with the initial state $|\xi\rangle$, and the objective function has a group of symmetries G. **If the mixer Hamiltonian commutes with G, then the algorithm will operate within the subspace W^G prior to the final measurement.**

In case $G = S_d$ acts as described before, we have $\dim(W^G) \approx \frac{d^n}{d!}$.



Where does QAOA 'live'?

Question

Is it possible to pick an initial state and mixer Hamiltonian so that QAOA 'runs' in a different $W_i \neq W^G$?

If there exists a mixer Hamiltonian $H_{M,i}$ that meets the following criteria:

- the lowest energy eigenspace of $H_{M,i}$ is one-dimensional and is spanned by $|\xi_i\rangle \in W_i$,
- $H_{M,i}$ preserves the direct sum decomposition of W (suffices to commute with G-action),

then one can establish a *reduced* QAOA with the same problem Hamiltonian H_P, mixer Hamiltonian $H_{M,i}$ and initial state $|\xi_i\rangle$. Thus defined QAOA operates within the subspace W_i prior to the final measurement.

Where does QAOA 'live'?

However, there is something to keep in mind...

Remark

The Perron-Frobenius theorem states that for a non-negative, irreducible matrix (in the standard basis), the Perron-Frobenius vector is a linear combination of all basis vectors with positive coefficients. Consequently, the conventional argument for ensuring the convergence of QAOA as $\mathbf{p} \to \infty$ to an optimal classical solution is **not applicable** unless the initial state is a superposition of all classical states with **positive amplitudes**.

Among the subspaces W_i, the only subspace containing vectors that satisfy the condition of positive amplitudes across all classical states is W^G.

Advantages of W^G

We would like to summarize the arguments in favor of choosing the *reduced* QAOA on the subspace of invariants $W^G \subseteq W$ over any other W_i, when considering COPs with classical symmetry groups that include S_d.

Advantages of W^G

We would like to summarize the arguments in favor of choosing the *reduced* QAOA on the subspace of invariants $W^G \subseteq W$ over any other W_i, when considering COPs with classical symmetry groups that include S_d.

1. The subspace W_0 has the smallest dimension for a sufficiently large number of qudits, n.

Advantages of W^G

We would like to summarize the arguments in favor of choosing the *reduced* QAOA on the subspace of invariants $W^G \subseteq W$ over any other W_i, when considering COPs with classical symmetry groups that include S_d.

1. The subspace W_0 has the smallest dimension for a sufficiently large number of qudits, n.
2. It is impossible to ensure the convergence of reduced QAOA_i (operating on W_i) as the number of iterations p approaches infinity with the help of P-F Theorem for any other W_i.

Advantages of W^G

We would like to summarize the arguments in favor of choosing the *reduced* QAOA on the subspace of invariants $W^G \subseteq W$ over any other W_i, when considering COPs with classical symmetry groups that include S_d.

1. The subspace W_0 has the smallest dimension for a sufficiently large number of qudits, n.
2. It is impossible to ensure the convergence of reduced QAOA_i (operating on W_i) as the number of iterations p approaches infinity with the help of P-F Theorem for any other W_i.
3. The uniform superposition $\xi = |++\ldots+\rangle$, which is used as the initial vector in the standard QAOA resides in W^G.

However, there might be hope for other W_i's

Graph	Mean energy	Median energy	Min energy	Share of outcomes with $E < 1$
Γ_1, W	0.726	0.7056	0.3584	41/50
Γ_1, W^G	0.5692	0.4673	0.1923	47/50
Γ_1, W_1	0.5726	0.5142	0.1621	47/50
Γ_2, W	0.9696	0.9316	0.4814	33/56
Γ_2, W^G	0.7437	0.7388	0.3691	51/56
Γ_2, W_1	0.8688	0.7148	0.3964	47/56
Γ_3, W	1.2495	1.2417	0.6533	11/56
Γ_3, W^G	0.9344	0.8857	0.3691	35/56
Γ_3, W_1	0.7334	0.6763	0.2598	50/56
Γ_4, W	1.4857	1.5313	0.7382	6/56
Γ_4, W^G	1.1959	1.1074	0.5117	21/56
Γ_4, W_1	1.2415	1.1489	0.4395	20/56
Γ_5, W	1.3469	1.3066	0.6162	14/50
Γ_5, W^G	0.9149	0.9507	0.3516	30/50
Γ_5, W_1	0.94123	0.9375	0.2939	27/50

Reduced QAOA energy for edge coloring with $p = 9$ (on W, W^G and W_1)