Erratum: Progress toward favorable landscapes in quantum combinatorial optimization [Phys. Rev. A 104, 032401 (2021)]

Juneseo Lee, Alicia B. Magann, Herschel A. Rabitz, and Christian Arenz

(Received 27 February 2024; published 22 March 2024)

DOI: 10.1103/PhysRevA.109.039903

Recently, Leontica and Amaro found a counterexample [1] to our claim in Lemma 1 of our paper that the optimization landscape for the MaxCut problem with an $\mathbb{X}$-*Ansatz* is free of local optima. We describe and investigate this counterexample. We confirm that the central Theorem (2) of our paper, i.e., that establishes the equivalence between the solution sets of the $\mathbb{X}$-*Ansatz* and a purely classical algorithm, remains valid.

In Appendix A 2 b of our paper, we distinguished between different cases (a)–(c) when the critical points condition in Eq. (A27) can be satisfied, yielding different types of critical points. We have found that the conclusion presented in case (a) is false, as Eq. (A28) only holds *at* a critical point. For example, consider the vector $\boldsymbol{\theta} + \frac{x}{2} v_k$ for some real scalar x, and v_k being the unit vector with a 1 in the kth position and 0 everywhere else. Using the critical point condition in Eq. (A27), we obtain

$$J\left(\boldsymbol{\theta} + \frac{x}{2} v_k\right) - J(\boldsymbol{\theta}) = [\cos(x) - 1][\cos(2\theta_k) S_k + \sin(2\theta_k) T_k],$$

which shows that $J(\boldsymbol{\theta} + \frac{x}{2} v_k) - J(\boldsymbol{\theta})$ is either always nonpositive or always non-negative, depending on the sign of $[\cos(2\theta_k) S_k + \sin(2\theta_k) T_k]$, which is independent of the scalar x. As a result, we have shown that in this example, k does not lead to $\boldsymbol{\theta}$ being a saddle. Consider now any critical point corresponding to an eigenstate of H_p, and choose any k corresponding to some cut element H_k. From the discussion of case (a) we conclude that if we vary θ_k along the *entire* interval $[0, \pi)$ and the objective function value does not change, we obtain a class of critical points that are not saddles.

We illustrate this phenomenon and allude to a broader class of critical points by expanding upon the counterexample by Leontica *et al.* based on a 4-qubit star. Consider the 4-qubit star described by the Hamiltonian [1],

$$H_p = Z_0(Z_1 + Z_2 + Z_3), \tag{1}$$

and an $\mathbb{X}$-*Ansatz* consisting of all 1- and 2-qubit elements. We collect the corresponding variational parameters in the set $\{\theta_S\}$, where $S \in \{0, 1, 2, 3, 01, 02, 03, 12, 13, 23\}$ denotes the different vertex subsets. The global optimum (i.e., corresponding to the smallest eigenvalue -3 of H_p) is obtained by "cutting out" the center vertex. Leontica *et al.* point out the following:

(1) Set $\theta_0 = \theta_2 = \theta_3 = \theta_{02} = \theta_{03} = \theta_{23} = \frac{\pi}{2}$, i.e., effectively "doing nothing" since $\cos(\frac{\pi}{2}) = 0$.

(2) Set $\theta_{01} = \pi$ while setting $\theta_{12} = \theta_{13} = 0$, which effectively "decouples" vertex 1 while selecting the central vertex (which is now connected to vertices 2 and 3 giving a cut value of -2).

(3) Now, θ_1 can be arbitrarily varied without changing the objective function, which is constant at -2. Thus, for all values of θ_1 the objective function is identically -2, the Jacobian is identically 0, and the Hessian row/column corresponding to vertex 1 is identical zero. As such, the "line" that is described by θ_1 corresponds to critical points that are local optima.

This observation is a specific illustration of the aforementioned class of critical points. Vertex 1 has effectively been "decoupled" from the graph, and the resultant line of local minima are in fact global minima of the decoupled subgraph. On the other hand, consider the following construction:

(1) Set $\theta_0 = \theta_1 = \theta_2 = \theta_3 = \theta_{13} = \frac{\pi}{2}$, effectively "doing nothing" for all the one-body terms (as before) and the 1,3 interaction term.

(2) Further set $\theta_{03} = -\frac{\pi}{2}$ and $\theta_{02} = \theta_{12} = 0$, but allow both θ_{01} and θ_{23} to be freely varied. This choice yields $J(\boldsymbol{\theta}) = -[1 + \sin(\theta_{23} - \theta_{01})]$, while the Jacobian elements are all 0 except for the terms $\cos(\theta_{23} - \theta_{01})$ and $-\cos(\theta_{23} - \theta_{01})$ that correspond to θ_{01} and θ_{23}, respectively.

(3) Consider the line $\theta_{23} = \frac{\pi}{2} + \theta_{01}$. On this line, the objective function is identically -2 and the Jacobian is identically 0. However, the Hessian along this line does not have 0′s in the corresponding columns/rows, although it is still degenerate. Instead, the Hessian has a kernel of dimension 1 defined by the $\{\theta_{01}, \theta_{23}\}$ line above. All other Hessian eigenvalues are positive. As such, this line is also a "line of local minima."

(4) Unlike the original example of Leontica *et al.*, however, the nondegenerate subspace does not correspond to a subset of the original graph.

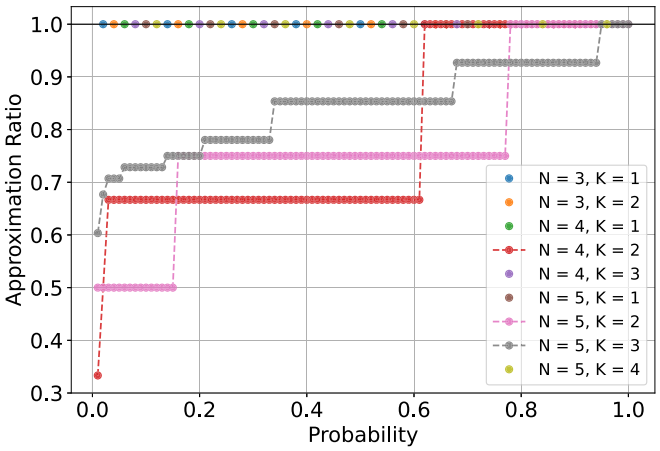

FIG. 1. Cumulative distribution function of the approximation ratio of an N-qubit star with K-body *Ansatz* elements, averaged over 100 randomly chosen initial variational parameters and optimized using the L-BFGS-B gradient algorithm for each (N, K) tuple. We see that for six of the nine tuples (corresponding to $K = 1$ and $K = N - 1$ for each of $N \in \{3, 4, 5\}$) the gradient algorithm always converges to the global optimum. The $(N, K) \in \{(4, 2); (5, 2); (5, 3)\}$ instances show convergence to local optima, e.g., the pink dots indicate that 62% of the time, $(N, K) = (5, 2)$ converges to an approximation ratio of 0.75.

In summary, we have described two lines of local minima: one defined by a single "free" *Ansatz* element (and therefore it can be reduced to an equivalent subgraph by decoupling) and another defined by a pair of "free" *Ansatz* elements with a constraint on their difference (i.e., still a single degree of freedom, but with no subgraph correspondence).

We numerically examine whether a gradient algorithm will get stuck in the local optima described above for an N-qubit star graph with $N \leqslant 5$. The numerical simulations are shown in Fig. 1, where we studied the cumulative distribution function (CDF) of the approximation ratio for different numbers of vertices N and K-body depth K. We find that the pair $(N, K) = (5, 3)$ achieves higher approximation ratios than $(N, K) = (5, 2)$. Furthermore, we also see that $K = 2$ instances perform strictly worse than $K = 1$ instances for small $N \leqslant 5$. Importantly, for a sufficiently large K-body depth, which translates into sufficiently many variational parameters, convergence to the global optimum can be obtained with high probability.

We conclude by restating the major theorems from the original paper following Lemma 1, with comments on which portions remain true, and which require modifications:

Theorem 1. The optimization landscape associated with an $\mathbb{X}$-*Ansatz* for which $\mathcal{A}$ consists of all nonsymmetric $2^{n-1} - 1$ vertex subsets exhibits no local optima.

Since this theorem rests on Lemma 1, it does not hold anymore.

Theorem 2. For any graph G and an $\mathbb{X}$-*Ansatz* $\mathcal{A}$ with size $|\mathcal{A}|$, there exists a purely classical algorithm with $O(|\mathcal{A}|)$ resources that has the same solution set as the set of local optima that satisfy conditions Eq. (17) and Eq. (18) in the original paper.

This theorem holds. The classical algorithm utilized in its proof is effectively independent from the prior results in the paper, and is thus unaffected by the existence of degenerate critical points in Lemma 1.

Corollary 2.1. Unless P = NP, there does not exist a class of $\mathbb{X}$-*Ansätze* with $|\mathcal{A}| = poly(n)$ that yields a landscape free from local optima for generic graphs.

This first corollary to Theorem (2) holds as well.

Corollary 2.2. Given a fixed approximation ratio α, and any $\mathbb{X}$-*Ansatz* with $|\mathcal{A}| = poly(n)$, even an algorithm that can escape saddle points cannot provide a superpolynomial advantage over a purely classical polynomial-time α-approximation scheme for MaxCut.

The original proof of this corollary sidestepped the need to consider the value of the cost function at saddle points by focusing only on approximation ratios derived from local optima. However, we now need to consider whether degenerate critical points can provably or numerically show faster convergence and/or better approximation ratios, which remains an open question.

We thank S. Leontica and D. Amaro for pointing out the counterexample to us. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under Contract No. DE-NA0003525. This article has been authored by an employee of National Technology & Engineering Solutions of Sandia, LLC under Contract No. DE-NA0003525 with the U.S. Department of Energy (DOE). The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this article or allow others to do so, for United States Government purposes. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan [2]. This paper describes objective technical results and analysis. Any subjective

views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government. SAND2024-01659O.

[1] S. Leontica and D. Amaro, Exploring the neighborhood of 1-layer QAOA with instantaneous quantum polynomial circuits, Phys. Rev. Res. **6**, 013071 (2024).

[2] https://www.energy.gov/downloads/doe-public-access-plan.