

From Problem-Structured QAOA to Problem-Agnostic Variational MaxCut Solvers: Representation, Redundancy, and Optimization Landscapes

Leonardo Lavagna * and Massimo Panella

Department of Information Engineering, Electronics and Telecommunications, Sapienza University of Rome, 00184 Rome, Italy
 * Correspondence: leonardo.lavagna@uniroma1.it

How To Cite: Lavagna, L.; Panella, M. From Problem-Structured QAOA to Problem-Agnostic Variational MaxCut Solvers: Representation, Redundancy, and Optimization Landscapes. *Applied Mathematics and Statistics* **2026**, *3*(2), 21. <https://doi.org/10.53941/ams.2026.100021>

Received: 30 August 2026
 Revised: 18 September 2026
 Accepted: 8 October 2026
 Published: 13 October 2026

Abstract: Variational quantum optimization provides a useful setting for examining how circuit structure, parameterization, and trainability interact in practice. Objectives: Among the most promising variational quantum algorithms, the Quantum Approximate Optimization Algorithm (QAOA) depends critically on circuit architecture, but greater circuit complexity does not necessarily lead to easier optimization or better solutions. Using the Maximum Cut (MaxCut) problem on graphs as a controlled setting, we investigate how variational behavior changes as problem-specific QAOA structure is progressively removed, with particular attention to representational capacity, optimization geometry, and finite-budget trainability. Methods: We analytically characterize a graph-agnostic product-state formulation of MaxCut and derive exact equivalences, redundancies, and local-landscape properties. These results motivate a controlled comparison among separable product ansätze, a representative graph-agnostic entangling circuit, and graph-structured QAOA at several depths, including a parameter-count-matched control. The numerical benchmark comprises 95 small unweighted graph instances optimized under a common objective-evaluation budget. Results: The product-state formulation preserves the exact discrete MaxCut optimum, while its strict local maxima correspond to classical cuts that are stable under single-vertex flips. Redundant phase parameters increase optimization effort without enlarging the objective-relevant family. Under the prescribed optimization budget, the product-state ansatz attains higher average solution quality than the tested QAOA depths while requiring substantially fewer objective evaluations. The graph-agnostic entangling comparator provides no consistent advantage over the separable baseline, and a simple classical local-search method remains highly competitive. Conclusion: In the regime considered here, objective-relevant parameterization and optimization geometry can be as important as circuit richness.

Keywords: quantum approximate optimization algorithm; MaxCut; variational quantum algorithms; product-state ansatz

1. Introduction

Variational quantum algorithms combine parameterized quantum circuits with classical optimization and provide a flexible framework for constructing low-depth quantum heuristics, with applications ranging from drug discovery to quantitative finance [1–3]. Among them, the Quantum Approximate Optimization Algorithm (QAOA) has become a central framework for combinatorial optimization and has generated a broad family of variants and parameter-setting strategies [4,5]. Standard QAOA is distinguished by an explicitly problem-structured ansatz: alternating evolutions generated by a cost Hamiltonian and a mixer Hamiltonian are optimized to increase the expectation value of a combinatorial objective. For the prototypical Maximum Cut (MaxCut) problem on graphs [4,6,7], the cost evolution

is built directly from the graph edges, so the input instance is encoded not only in the objective evaluated by the classical optimizer but also in the quantum dynamics itself. This architectural feature motivates a basic question that is often obscured when QAOA is compared with generic variational circuits: what part of its behavior follows from representational capacity, what part follows from problem-aligned circuit structure, and what part follows from the geometry of the resulting optimization landscape?

This distinction is important because increasing circuit flexibility does not necessarily make a variational problem easier to train. Standard QAOA uses only two global parameters per layer, whereas variants such as multi-angle QAOA enlarge the parameter space to obtain greater local flexibility and can improve approximation quality at reduced depth on selected MaxCut families [8]. Other approaches augment a shallow QAOA circuit with additional problem-independent gates [9], warm-start the variational search from a classical relaxation [10], or introduce problem-informed preprocessing into the ansatz itself [11]. These developments show that circuit depth, number of parameters, and degree of problem specificity are independent design choices rather than interchangeable proxies for expressivity. At the same time, numerical evidence indicates that QAOA training can become difficult even at relatively low depth because the landscape may contain a rapidly growing number of poor local minima [12]. More broadly, recent advances in barren-plateau theory further underscore the trainability challenges that can arise in parameterized quantum circuits [13, 14]. Consequently, comparing architectures only by their best attainable objective values can conflate two different questions: whether a family can represent an optimal solution and whether a finite-budget classical optimizer can locate it. In this context, earlier work on Enhanced QAOA (EQAOA) investigated an intermediate point in this design space by appending a compact multiparameter ansatz to a QAOA circuit and using blockwise optimization followed by a final joint optimization [15]. That construction retains the problem-dependent QAOA block while adding trainable degrees of freedom that are not tied directly to the graph topology. The present study examines the complementary limiting case suggested by this architecture: the problem-structured QAOA block is removed, and optimization is performed using a graph-agnostic variational circuit alone. We use this limit as a controlled setting, rather than as a replacement algorithm for QAOA, so that problem structure and other architectural properties can be varied separately.

This perspective is especially revealing for MaxCut because its standard quantum objective is diagonal in the computational basis. Every classical cut is therefore represented by a computational-basis product state, so the existence of an exact optimum within the product-state family does not, by itself, constitute a new quantum resource claim. Earlier work has shown more generally that exact MaxCut variational constructions can be obtained at the cost of difficult circuit design or classical variational optimization [11]. In particular, separable ansätze have been used successfully in quantum-imaginary-time approaches to weighted MaxCut [16]. Likewise, recent work has found that problem-agnostic entanglement in hardware-efficient ansätze can be detrimental for MaxCut, while QAOA behaves differently when its entanglement is generated by the problem Hamiltonian [17]. Relatedly, recent classical work has constructed smooth Ising relaxations with an explicit correspondence between continuous local minima and discrete one-flip optima [18]. These results make “more entanglement” or “more parameters” inadequate explanations by themselves and motivate a more controlled analysis of which variational degrees of freedom are actually visible to the objective and how the resulting continuous landscape relates to classical local search.

We address this issue through an analytical and numerical study centered on a tractable product-state limit and a small set of controlled comparators. The canonical graph-agnostic limit is an R_y product state with a rotation angle per vertex along the y direction. Likewise, an R_x product state is also included as an analytically equivalent phase control, while adding an R_z gate to each qubit creates nominal overparameterization without changing the MaxCut expectation, thereby isolating parameter redundancy. A fixed chain of two-qubit rotations R_{xx} is then used as a representative graph-agnostic entangling comparator, while standard QAOA provides the problem-structured reference. This architectural ladder anchors the numerical comparisons to analytically identifiable changes in objective-visible parameterization and problem alignment.

Within this framework, the product-state MaxCut objective reduces to a multilinear relaxation whose global optimum is attained at a classical cut in the graph, providing a natural structural baseline for the subsequent analysis. Simple changes in circuit parameterization can also be characterized exactly: matched R_x and R_y product ansätze generate the same MaxCut objective landscape, whereas terminal R_z rotations introduce flat directions without enlarging the objective-relevant variational family. The landscape further admits a direct connection with classical local search, since its strict local maxima correspond exactly to computational-basis cuts that are strict local optima under single-vertex flips, while the uniform point $\theta_i = \pi/2$ can be characterized through the weighted adjacency spectrum. Building on these analytical observations, we compare the product-state family with a representative graph-agnostic entangling circuit and with graph-structured QAOA at several depths, including a parameter-count-matched control, under a common finite optimization budget. This combined analytical and

numerical approach is intended to clarify how objective-visible degrees of freedom, finite-budget trainability, and problem-specific structure interact in variational MaxCut optimization, without attributing the observed behavior to quantum advantage or to entanglement alone.

2. Methods

Following the perspective outlined in Section 1, we now formalize the analytical and numerical framework used to disentangle representational capacity, optimization geometry, and problem-specific circuit structure in variational MaxCut optimization. We begin from the standard MaxCut formulation and problem-structured QAOA, which provide the common reference point for the study. We then introduce the graph-agnostic product-state limit and the related variational ansätze used to isolate objective-relevant parameterization, redundancy, and generic entanglement. Finally, we describe the graph ensembles, optimization protocol, performance metrics, statistical analysis, and validation procedures adopted for the controlled numerical comparison.

2.1. MaxCut and Problem-Structured QAOA

Let $G = (V, E, \mathbf{w})$ be an undirected graph with $n = |V|$ and edge weight matrix $\mathbf{w} = (w_{ij})$ with non-negative entries $w_{ij} \geq 0$. A classical graph cut can be encoded by spin variables $z_i \in \{-1, +1\}$, with opposite signs corresponding to vertices assigned to different sides of the partition. In this formulation the weighted cut value is

$$C(\mathbf{z}) = \frac{1}{2} \sum_{\{i,j\} \in E} w_{ij} (1 - z_i z_j). \quad (1)$$

The corresponding diagonal MaxCut Hamiltonian is obtained by promoting each spin variable z_i to the Pauli- Z operator acting on qubit i . Denoting $Z = \text{diag}(1, -1)$ and by Z_i its action on qubit i , the Hamiltonian is

$$H_C = \frac{1}{2} \sum_{\{i,j\} \in E} w_{ij} (I - Z_i Z_j), \quad (2)$$

so that $\langle \mathbf{z} | H_C | \mathbf{z} \rangle = C(\mathbf{z})$ for every computational-basis state. Throughout this work the objective is maximized, and we denote by C^* the exact MaxCut value:

$$C^* = \max_{\mathbf{z} \in \{-1, +1\}^n} C(\mathbf{z}). \quad (3)$$

Standard depth- p QAOA starts from the uniform superposition state $|+\rangle^{\otimes n}$ and alternates the cost evolution with a transverse-field mixer [4, 19],

$$|\psi_p(\gamma, \beta)\rangle = \prod_{k=1}^p e^{-i\beta_k H_M} e^{-i\gamma_k H_C} |+\rangle^{\otimes n}, \quad H_M = \sum_{i=1}^n X_i. \quad (4)$$

where the overall ansatz contains $2p$ trainable angles and X_i is the X -Pauli gate acting on the i -th qubit. The graph appears explicitly in H_C , and therefore in every cost layer $e^{-i\gamma_k H_C}$. We use the term problem-structured in this architectural sense. By contrast, a circuit is called graph agnostic below when its gate pattern and entangling connectivity are fixed independently of E . Notice that this does not mean that the optimization problem is graph independent: the same graph-dependent objective $\langle H_C \rangle$ is evaluated for every architecture.

The formulation above provides the problem-structured reference against which the subsequent variational models are compared. To isolate the role of this structure, we next consider the opposite limit, in which the graph enters only through the objective function while the circuit architecture itself remains independent of the edge set. This graph-agnostic setting leads naturally to a product-state formulation of MaxCut, whose analytical tractability allows us to separate representational properties of the variational family from the optimization behavior induced by its parameterization.

2.2. Product-State Reduction

Consider an arbitrary pure product state $|\psi\rangle = \bigotimes_{i=1}^n |\psi_i\rangle$ and define the local magnetizations

$$m_i = \langle \psi_i | Z_i | \psi_i \rangle \in [-1, 1]. \quad (5)$$

The product structure gives $\langle Z_i Z_j \rangle = m_i m_j$, and hence

$$F(\mathbf{m}) = \langle \psi | H_C | \psi \rangle = \frac{1}{2} \sum_{\{i,j\} \in E} w_{ij} (1 - m_i m_j), \quad \mathbf{m} \in [-1, 1]^n. \quad (6)$$

Conversely, every value $m_i \in [-1, 1]$ can be realized as the Z expectation of a pure single-qubit state. The formula in Equation (6) therefore characterizes the complete MaxCut expectation family attainable by product states; related product-state formulations of variational MaxCut optimization have also been considered in previous work [16, 17].

Since the objective in Equation (6) is multilinear in the local magnetizations, a standard property of multilinear optimization over a hypercube [20] applied to the present MaxCut formulation, yields the following structural result.

Proposition 1 (Product-state preservation of the MaxCut optimum). *For the Hamiltonian in Equation (2),*

$$\max_{\mathbf{m} \in [-1, 1]^n} F(\mathbf{m}) = \max_{\mathbf{z} \in \{-1, +1\}^n} C(\mathbf{z}) = C^*. \quad (7)$$

Proof. For fixed values of all coordinates except m_k , the function F is affine in m_k . Therefore, for any point $\mathbf{m} \in [-1, 1]^n$, the coordinate m_k can be replaced by one of the endpoints -1 or $+1$ without decreasing the objective. Applying this argument successively to all coordinates yields a vertex $\mathbf{z} \in \{-1, +1\}^n$ with $F(\mathbf{z}) \geq F(\mathbf{m})$. Hence a global maximizer exists at a hypercube vertex. At those vertices Equation (6) reduces exactly to Equation (1), proving Equation (7). $\square$

Notice that Proposition 1 is a statement about representation rather than computational complexity. It shows that entanglement is unnecessary to represent a state attaining the exact MaxCut objective, but it does not provide an efficient procedure for locating such a state. In particular, the continuous objective remains nonconvex and retains the combinatorial structure of the original problem.

2.3. R_x – R_y Equivalence and Objective-Irrelevant Phases

The product-state reduction discussed in the previous subsection characterizes the MaxCut objective independently of any specific circuit parameterization. To connect this general description with the variational architectures used in the numerical study, we next consider explicit single-qubit rotation ansätze that realize the local magnetizations m_i . This allows us to determine which changes in circuit parameterization genuinely modify the objective landscape and which instead introduce only equivalent or redundant degrees of freedom. We take the R_y product family as the canonical one-parameter-per-vertex realization of Equation (6),

$$|\psi_y(\boldsymbol{\theta})\rangle = \bigotimes_{i=1}^n R_y(\theta_i) |0\rangle. \quad (8)$$

Since $\langle Z_i \rangle = \cos \theta_i$, its objective is

$$F_y(\boldsymbol{\theta}) = \frac{1}{2} \sum_{\{i,j\} \in E} w_{ij} (1 - \cos \theta_i \cos \theta_j). \quad (9)$$

The choices $\theta_i \in \{0, \pi\}$ recover the two classical assignments of vertex i , so Equation (8) contains every computational-basis cut while providing a continuous interpolation between them.

Proposition 2 (R_x – R_y objective equivalence). *Let $|\psi_x(\boldsymbol{\theta})\rangle = \bigotimes_{i=1}^n R_x(\theta_i) |0\rangle$. For matched angles and the MaxCut Hamiltonian in Equation (2), $F_x(\boldsymbol{\theta}) = F_y(\boldsymbol{\theta})$ for every $\boldsymbol{\theta}$.*

Proof. We use standard single-qubit rotation identities. In particular, the states $R_x(\theta) |0\rangle$ and $R_y(\theta) |0\rangle$ differ only in their computational-basis phases, while both satisfy $\langle Z \rangle = \cos \theta$. Since the corresponding many-qubit states are product states with identical local Z expectations, Equation (6) proves the result. $\square$

The R_x family is therefore not introduced as a competing solver with independent representational power. It serves as a theory-driven control: any difference between matched R_x and R_y objective evaluations must arise from numerical or implementation effects rather than from the exact MaxCut landscape.

Consider now the nominally larger separable family

$$|\psi_{yz}(\boldsymbol{\theta}, \boldsymbol{\phi})\rangle = \bigotimes_{i=1}^n R_z(\phi_i) R_y(\theta_i) |0\rangle. \quad (10)$$

Again, with standard quantum-gate computations, we get the following result.

Proposition 3 (R_z redundancy). *For the ansatz in Equation (10), we have $F_{yz}(\theta, \phi) = F_y(\theta)$ for every θ and ϕ .*

Proof. The terminal $R_z(\phi_i)$ rotations leave the computational-basis probabilities of each qubit unchanged. Equivalently, they leave $\langle Z_i \rangle = \cos \theta_i$ unchanged. Since H_C is diagonal in the computational basis, its expectation depends only on those populations, and the thesis follows. $\square$

Thus Equation (10) contains $2n$ nominal trainable parameters but only n objective-relevant directions. The remaining n directions are globally flat for the MaxCut expectation. This construction provides a controlled way to ask whether objective-irrelevant overparameterization changes the behavior of a derivative-free classical optimizer even when the attainable objective family is unchanged.

2.4. Local Curvature and Classical Single-Vertex Flips

The product-state formulation of the previous subsections exposes a direct connection between the continuous landscape and a standard local operation on classical cuts, as discussed here. In particular, differentiating Equation (9) gives

$$\frac{\partial F_y}{\partial \theta_k} = \frac{1}{2} \sin \theta_k \sum_{j: \{k, j\} \in E} w_{kj} \cos \theta_j. \quad (11)$$

At a computational-basis point $\theta_i \in \{0, \pi\}$, define $z_i = \cos \theta_i \in \{-1, +1\}$. In this setting, the gradient vanishes, all mixed second derivatives vanish, and

$$\left. \frac{\partial^2 F_y}{\partial \theta_k^2} \right|_{\theta \in \{0, \pi\}^n} = \frac{1}{2} z_k \sum_{j: \{k, j\} \in E} w_{kj} z_j. \quad (12)$$

Let $\mathbf{z}^{(k)}$ denote the assignment obtained by flipping only vertex k and define

$$\Delta_k C := C(\mathbf{z}^{(k)}) - C(\mathbf{z}). \quad (13)$$

Only edges incident on k change their contribution, yielding

$$\Delta_k C = z_k \sum_{j: \{k, j\} \in E} w_{kj} z_j. \quad (14)$$

Combining Equations (12) and (14) gives

$$\left. \frac{\partial^2 F_y}{\partial \theta_k^2} \right|_{\theta \in \{0, \pi\}^n} = \frac{1}{2} \Delta_k C. \quad (15)$$

A computational-basis cut that admits an improving one-vertex flip therefore has a positive-curvature direction in the continuous R_y landscape. Conversely, a strict one-flip local maximum has negative curvature along every coordinate direction. The correspondence can be sharpened to an exact characterization of strict local maxima.

Proposition 4 (Strict local maxima and one-flip optimality). *A point θ^* is a strict local maximum of F_y if and only if, modulo the 2π periodicity of each angle, $\theta_i^* \in \{0, \pi\}$ for every i and its associated spin assignment $z_i = \cos \theta_i^*$ is a strict one-vertex-flip local maximum of the weighted MaxCut objective, i.e. $\Delta_k C < 0$ for every vertex k .*

Proof. Any strict local maximum is stationary. Suppose that $\sin \theta_k^* \neq 0$ for some coordinate k . Equation (11) then requires $\sum_j w_{kj} \cos \theta_j^* = 0$. With all other angles fixed, Equation (9) depends on θ_k only through the term $-\frac{1}{2} \cos \theta_k \sum_j w_{kj} \cos \theta_j^*$, which is identically constant under this condition. Hence the objective is locally flat along coordinate k , and the stationary point cannot be a strict local maximum. Every strict local maximum must therefore satisfy $\sin \theta_k^* = 0$ for all k , so each θ_k^* is an integer multiple of π and is equivalent, modulo 2π , to either 0 or π . At such a point the mixed second derivatives vanish and the diagonal Hessian entries are $\Delta_k C/2$ by Equation (15). If $\Delta_k C < 0$ for every k , the Hessian is negative definite and the point is a strict local maximum. Conversely, if $\Delta_k C > 0$ for some k , the corresponding coordinate has positive curvature; if $\Delta_k C = 0$, the objective is exactly flat along that coordinate with the remaining angles fixed. Thus strict local maximality is equivalent to $\Delta_k C < 0$ for all k , which is the thesis. $\square$

Proposition 4 is related to recent work on local-minima-preserving continuous relaxations of Ising problems, which constructs additional attractor terms so that continuous local minima correspond to discrete one-flip

optima [18]. Here no auxiliary attractor is introduced: the correspondence arises directly from the trigonometric product-state parameterization of the MaxCut expectation. The result motivates the single-vertex-flip baseline used in the numerical study below because it probes the same local combinatorial structure that governs strict local maxima of the variational landscape.

A second useful reference point is the uniform product state obtained by setting $\theta_i = \pi/2$ for all vertices. Equation (9) gives $F_y = \frac{1}{2} \sum_{\{i,j\} \in E} w_{ij}$, equal to the expected weighted cut value of an independently random assignment, and Equation (11) gives zero gradient. At this point the Hessian has entries

$$\left. \frac{\partial^2 F_y}{\partial \theta_i \partial \theta_j} \right|_{\theta = \frac{\pi}{2} \mathbf{1}} = -\frac{1}{2} W_{ij}, \quad (16)$$

where W is the weighted adjacency matrix of G . Hence

$$\nabla^2 F_y \left(\frac{\pi}{2} \mathbf{1} \right) = -\frac{1}{2} W. \quad (17)$$

For any graph with at least one positive-weight edge, W is a nonzero real symmetric matrix with zero trace; its spectrum therefore contains both positive and negative eigenvalues. The Hessian in Equation (17) is consequently indefinite, so the uniform point is a saddle. This observation is relevant when importing initialization intuitions from QAOA: the $|+\rangle^{\otimes n}$ -like population profile is natural for QAOA, but its direct product-state analogue is a stationary saddle for the continuous MaxCut objective considered here.

2.5. Controlled Ansatz Ladder

The analytical results developed above provide a natural basis for constructing the variational models used in the numerical study. Rather than comparing a collection of unrelated circuits, we organize the benchmark as a controlled architectural ladder in which each ansatz is chosen to probe a specific structural feature identified in the preceding analysis. The product-state families provide the analytically tractable reference, the redundant parameterization tests whether additional objective-irrelevant degrees of freedom affect optimization, and the graph-agnostic entangling circuit introduces nonseparable dynamics without encoding the input graph. Standard QAOA then completes the comparison by restoring explicit problem structure through the cost Hamiltonian. This organization allows the numerical experiments to be interpreted directly in light of the theoretical properties established in the previous subsections. Table 1 summarizes the role of each model. The R_y product state is the canonical graph-agnostic baseline because it uses one objective-relevant angle per vertex and contains every classical cut. The matched R_x family tests Proposition 2. The $R_y R_z$ product family tests Proposition 3 while exposing the optimizer to exact flat directions. A fourth architecture provides a representative entangling circuit whose connectivity is fixed independently of the input graph. QAOA then introduces graph-aligned cost evolution and serves as the problem-structured reference.

Table 1. Architectural roles in the controlled comparison.

Ansatz	Separable	Extra Phases	Entangled	Uses E	Scientific Role
R_y	yes	no	no	no	Canonical product-state baseline.
R_x	yes	no	no	no	Pointwise-equivalence control.
$R_y R_z$	yes	yes	no	no	Exact redundant-parameter control.
$R_y R_z + R_{xx}$	no	yes	yes	no	Representative graph-agnostic entangling comparator.
QAOA	no	–	yes	yes	Problem-structured reference.

The column “uses E ” refers to circuit construction, not to evaluation of the common graph-dependent objective.

To introduce nonseparable states into the comparison, we consider the graph-agnostic entangling ansatz

$$|\psi_A\rangle = \left[\prod_{i=1}^{n-1} R_{xx}^{(i,i+1)}(\alpha_i) \right] \left[\bigotimes_{i=1}^n R_z(\phi_i) R_y(\theta_i) \right] |0\rangle^{\otimes n}, \quad (18)$$

with $3n - 1$ trainable parameters, where $R_{xx}(\alpha) := \exp(-i \frac{\alpha}{2} X \otimes X)$ is a two-qubit rotation gate of angle α acting on adjacent qubits via a combination of X -Pauli gates. Vertices are mapped to qubits with a fixed ordering and the R_{xx} gates always follow the chain from 1 to n , irrespective of E . The purpose of this model is not to represent a privileged hardware architecture or to isolate entanglement as a single causal variable, but to provide a reproducible graph-agnostic entangling comparator that is deliberately unaligned with the MaxCut graph. This fixed ordering

also means that the ansatz is not permutation invariant; we retain a deterministic vertex ordering throughout and treat this dependence as a limitation rather than optimizing the ordering per instance.

For QAOA we use Equation (4) with the standard mixer at fixed reference depths $p \in \{1, 3, 5\}$, corresponding to 2, 6, and 10 trainable parameters, respectively. We additionally include a size-dependent parameter-count-matched control $p_{\text{match}} = n/2$, so that QAOA has $2p = n$ trainable parameters, equal to the R_y product ansatz. For the even sizes used here this adds $p = 2, 4$, and 6 at $n = 4, 8$, and 12, while $p = 3$ at $n = 6$ and $p = 5$ at $n = 10$ are already contained in the fixed-depth set. The numerical implementation applies the cost phase exactly from the MaxCut spectrum and the mixer as independent $R_x(2\beta_k)$ rotations. This convention is checked against independent dense-matrix exponentiation on small graphs before benchmark execution.

2.6. Graph Ensembles and Exact Reference Solutions

Although the analytical results above allow nonnegative edge weights, the controlled benchmark is restricted to the unweighted case $w_{ij} = 1$ so that architectural effects are not mixed with edge-weight parameterization. Node counts are $n \in \{4, 6, 8, 10, 12\}$. Three graph families are used. Complete graphs provide highly symmetric deterministic controls. Erdős–Rényi graphs with edge probability $\frac{1}{2}$ probe heterogeneous random connectivity, with ten independently generated instances for each n . Random 3-regular graphs probe fixed-degree sparse structure, again with ten instances for each admissible node count; the $n = 4$ case is omitted because the unique simple 3-regular graph is the complete graph with $n = 4$, and would thus duplicate the complete-graph control. The resulting benchmark contains 95 unique graph instances.

All stochastic graphs are generated reproducibly from a master seed, with graph identifiers and generation seeds stored alongside every result. Because $n \leq 12$, the exact MaxCut value C^* and the full set of optimal computational-basis assignments are obtained by exhaustive enumeration. Exact enumeration is used only to evaluate the small controlled benchmark and is not part of any claimed scalable solution procedure.

2.7. Optimization Protocol and Initialization

Every variational model is optimized with the derivative-free COBYLA method under a common cap of 1000 objective evaluations. Five independent restarts are performed for each graph–architecture pair. The optimizer is deliberately held fixed across architectures so that the numerical experiment isolates changes associated with circuit parameterization and structure rather than introducing optimizer choice as an additional experimental variable. The aim is therefore not to identify the individually best classical optimizer for each ansatz. Optimizer dependence in QAOA has already been investigated comparatively in the literature, with substantial variation across optimization methods and problem settings [5,21]. In particular, introducing gradient-based methods such as Adam together with parameter-shift gradients would require a different accounting of computational effort, since constructing a gradient generally entails multiple shifted circuit evaluations rather than a single objective call. The comparison is therefore budgeted by objective evaluations rather than by optimizer iterations, because an objective evaluation is the most directly comparable unit of classical optimization effort across the parameterizations in the exact-simulation setting. This is not an equal-computational-cost budget: one evaluation can require different statevector work, circuit depth, or measurement resources across architectures, so wall-clock and quantum-resource equivalence are not implied.

All angles in the product and graph-agnostic circuits are initialized independently from $U(0, 2\pi)$ (the uniform distribution). QAOA cost angles are initialized from $U(0, 2\pi)$ and mixer angles from $U(0, \pi)$. A deterministic seed is derived from the master seed, graph identifier, and restart index. For a fixed graph and restart, the same seed is reused across architectures; where parameterizations overlap, this yields matched initial local-angle draws for the controlled comparisons. The optimizer’s termination status is stored separately from the scientific success criterion defined below, so numerical termination is not conflated with attainment of a high-quality cut.

All objective values in the primary benchmark are exact and noiseless. For the separable R_x , R_y , and $R_y R_z$ controls, optimization uses the closed-form product-state expectation in Equation (9); the corresponding final states are still constructed explicitly to evaluate P_{opt} (cf., Equation (20)). The graph-agnostic entangling circuit and QAOA are evaluated by exact statevector simulation. Exploiting the analytical product-state formula does not change its optimization landscape, but it makes explicit that those controls are efficiently classically evaluable. The purpose of the benchmark is to study architectural and classical-optimization effects rather than finite-shot or hardware effects, and no claim about noise robustness, device-level advantage, or near-term hardware efficiency is made from these simulations.

2.8. Metrics and Classical Contextual Baselines

For optimized parameters ϑ^* , the primary metric is the expected approximation ratio

$$\text{AR}_{\text{exp}} = \frac{\langle \psi(\vartheta^*) | H_C | \psi(\vartheta^*) \rangle}{C^*}, \quad (19)$$

which lies in $[0, 1]$ up to numerical precision. In paired architectural comparisons below, the subscript of AR identifies the corresponding architecture or QAOA depth; for example, AR_{R_y} denotes the expected approximation ratio obtained with the R_y product-state ansatz, while $\text{AR}_{p=5}$ denotes that obtained with QAOA at depth $p = 5$. Because a high expected value does not imply that an exact optimum is sampled with high probability, we separately record

$$P_{\text{opt}} = \sum_{\mathbf{z}: C(\mathbf{z})=C^*} |\langle \mathbf{z} | \psi(\vartheta^*) \rangle|^2, \quad (20)$$

the total probability mass assigned to all optimal bitstrings. We also store the complete objective-evaluation history, the number of evaluations consumed, runtime, nominal parameter count, and analytically known objective-relevant parameter count. A variational run is classified as target-successful when $\text{AR}_{\text{exp}} \geq 0.99$; this threshold is a reporting convention and is not used to terminate optimization.

Two simple classical references are included for interpretation rather than as state-of-the-art benchmarks. First, the expectation of a uniformly random cut is $|E|/2$, corresponding to the stationary product-state point analyzed in Equation (17). Second, a best-improvement single-vertex-flip local search starts from deterministic random assignments and repeatedly applies the improving flip with largest $\Delta_k C$ until no positive-gain move remains. This baseline is included because Equation (15) gives it a direct analytical connection to the coordinate curvature of the R_y landscape. It is not intended to replace stronger classical MaxCut methods such as the Goemans–Williamson semidefinite-programming approximation [22].

2.9. Aggregation and Statistical Reporting

Having defined the graph ensembles, optimization protocol, and performance measures used for each individual run, we next specify how the resulting observations are aggregated across repeated optimizations and graph instances. This distinction is important because variability arises at two different levels: from the stochastic initialization and optimization of a fixed graph, and from the choice of graph instance itself. The statistical procedure is therefore designed to preserve this hierarchical structure while enabling consistent comparisons among the variational architectures.

The proposed experimental evaluation has two nested sources of variability: optimizer restarts and random graph instances. To avoid treating repeated optimizations of the same graph as independent graph samples, restart-level outcomes are first summarized within each graph. For Erdős–Rényi and random 3-regular families, family-level summaries are then computed across graph instances. Means are reported together with 95% bootstrap confidence intervals obtained from 10,000 resamples, using the 2.5th and 97.5th percentiles of the bootstrap distribution. Because node count is a fixed design factor and the graph distributions differ with n , graph instances are resampled independently within each node-count stratum while preserving the original number of graphs per size in every bootstrap replicate. Pairwise architectural comparisons on stochastic families use the same graph instances and are therefore summarized through paired graph-level differences with the same size-stratified bootstrap procedure. Bootstrap resampling is reproducible from a fixed random seed of 20260815 (the date of the last run). Complete graphs are reported separately because there is only one graph at each node count; their variability therefore reflects optimizer restarts rather than an ensemble of graph instances. No benchmark records were missing after the integrity checks. We report effect sizes and confidence intervals rather than null-hypothesis significance tests, so no multiple-testing correction is applied. The parameter-count-matched QAOA comparison and the convergence analysis across evaluation checkpoints provide complementary robustness checks on the main finite-budget comparison.

Convergence is analyzed as a function of objective-evaluation budget using stored histories at common checkpoints. This allows a circuit that terminates early to be distinguished from one that is still improving near the evaluation cap. The principal comparisons therefore concern both final solution quality and the amount of classical optimization required to obtain it.

2.10. Validation and Reproducibility

The numerical implementation is validated against the analytical results before benchmark outputs are accepted. For randomly generated graphs and matched parameter values, the statevector implementation of the R_x and R_y product ansätze is required to reproduce the pointwise objective equivalence established in Proposition 2. Likewise,

the $R_y R_z$ implementation is checked against the exact phase redundancy derived above, and the simulated R_y expectation is compared with its closed-form expression. Exact MaxCut enumeration is independently verified on graph families with analytically known optima, including complete and bipartite graphs.

Additional tests verify the correspondence between the local curvature of the R_y landscape and single-vertex-flip gains, the Hessian at the uniform point, and the ability of the fixed R_{xx} block to generate entanglement for nontrivial parameter choices. The QAOA implementation is checked against independent dense-matrix exponentiation on small graphs. Reproducibility tests additionally verify deterministic graph generation and restart initialization, matched seeds across architectures where required by the controlled comparisons, enforcement of the common objective-evaluation cap, and preservation of complete optimization histories. These checks are performed independently of the benchmark analysis so that the numerical results reported below are accepted only after the corresponding analytical and implementation-level consistency tests are satisfied.

All experiments and statistical analyses were carried out in Python 3.13.5 using NumPy 2.3.5, SciPy 1.17.0, NetworkX 3.6.1, pandas 2.2.3, Matplotlib 3.10.8, and PyYAML 6.0.3. The complete experimental configuration, deterministic seeds, optimization histories, and analysis scripts are retained in the project repository and can be provided by the corresponding author upon reasonable request.

3. Results

Having established the analytical controls, architectural comparisons, and experimental protocol, we now examine how these distinctions manifest in the numerical results. We first verify that the benchmark reproduces the exact equivalences and redundancies derived in Section 2, thereby linking the implementation directly to the theoretical framework. We then compare the variational architectures in terms of solution quality and optimization effort, with particular attention to the effects of redundant parameterization, graph-agnostic entanglement, QAOA depth, and parameter-count matching. Finally, the stored optimization histories and classical local-search reference are used to place the observed finite-budget behavior in a broader optimization context.

3.1. Validation of the Analytical Controls

The frozen benchmark produced 3590 variational runs with complete optimization histories on 95 graph instances, together with 475 single-vertex-flip local-search runs. All records satisfied the implementation integrity checks described in Section 2. In particular, the matched R_x and R_y controls were numerically identical in all 475 graph-restart pairs: the initial objective, final objective, P_{opt} , and number of objective evaluations agreed to machine precision. This behavior is consistent with Proposition 2 over the complete optimization trajectory; independent statevector-level unit tests separately verify the pointwise objective equivalence for matched angles.

The $R_y R_z$ control provides a complementary test of objective-irrelevant overparameterization. Since the additional R_z variables do not change the exact objective, any difference from R_y is attributable to the behavior of COBYLA in the enlarged parameter space. On Erdős–Rényi graphs, the graph-level paired difference $\text{AR}_{R_y R_z} - \text{AR}_{R_y}$ was -0.0003 with a 95% bootstrap confidence interval (CI) of $[-0.0032, 0.0024]$; on random 3-regular graphs it was 0.0010 with CI $[-0.0017, 0.0040]$. The enlarged parameterization therefore produced no detectable systematic gain in the attained objective. In contrast, the redundant parameterization required on average 108.8 additional objective evaluations on Erdős–Rényi graphs (95% CI $[95.4, 121.9]$) and 131.4 additional evaluations on 3-regular graphs (95% CI $[125.8, 137.1]$). Figure 1 shows how this optimization-cost penalty grows with graph size.

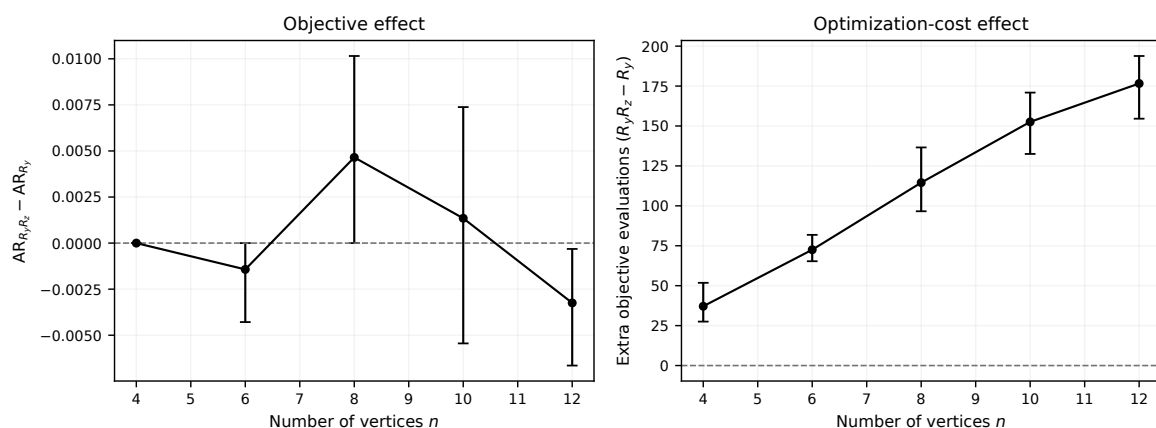


Figure 1. Effect of the exactly redundant R_z parameters. **Left:** graph-level paired difference in final expected approximation ratio between $R_y R_z$ and R_y . **Right:** paired difference in COBYLA objective evaluations. Points denote means over stochastic graph instances at each size; error bars denote 95% bootstrap confidence intervals.

3.2. Performance Across Graph Families and Sizes

With the analytical controls and implementation checks established, we next turn to the overall behavior of the variational architectures across the full benchmark. The aim at this stage is not to isolate any single architectural mechanism, but to provide a common quantitative overview of solution quality, optimal-solution probability, success rate, and optimization effort under the prescribed finite evaluation budget. This aggregate comparison provides the reference point for the more detailed graph-family- and architecture-specific analyses that follow. In particular, Table 2 summarizes the complete benchmark after first averaging the five optimizer restarts within each graph. The R_y product ansatz achieved a mean expected approximation ratio of 0.9784 and mean optimal-solution probability of 0.7832, with 78.3% of runs meeting the reporting threshold $\text{AR}_{\text{exp}} \geq 0.99$. The representative graph-agnostic entangling ansatz reached a slightly higher overall mean approximation ratio, 0.9811, with $P_{\text{opt}} = 0.8021$, but required more than three times as many objective evaluations on average. The aggregate mean QAOA performance improved monotonically across the tested depths, from a mean approximation ratio of 0.7954 at $p = 1$ to 0.8908 at $p = 3$ and 0.9202 at $p = 5$.

The separation between the product-state and QAOA results is concentrated in the stochastic graph families rather than the complete-graph controls. Complete graphs constitute a particularly symmetric setting for QAOA-MaxCut, and recent Lie-algebraic analysis has shown that, for sufficiently deep circuits in the corresponding unitary 2-design regime, the loss-function variance scales linearly rather than exponentially with system size, excluding barren-plateau behavior in that regime [23]. Although this asymptotic result does not directly characterize the small systems and finite QAOA depths considered here, it further emphasizes that complete graphs possess a distinctive optimization structure and should not be interpreted interchangeably with the stochastic graph ensembles. On complete graphs, the mean approximation ratio was essentially one for R_y and the agnostic entangling circuit, while QAOA reached 0.9955 at $p = 5$. On Erdős–Rényi graphs, the corresponding means were 0.9755 for R_y , 0.9817 for the graph-agnostic entangling comparator, and 0.9057 for QAOA $p = 5$. On random 3-regular graphs they were 0.9793, 0.9780, and 0.9290, respectively. Figure 2 shows the size dependence for the two stochastic families.

The graph-level paired comparison between R_y and QAOA $p = 5$ gives a mean difference $\text{AR}_{R_y} - \text{AR}_{p=5} = 0.0698$ on Erdős–Rényi graphs (95% CI [0.0604, 0.0793]) and 0.0504 on 3-regular graphs (95% CI [0.0423, 0.0588]). The corresponding paired differences in P_{opt} were 0.2743 (95% CI [0.2171, 0.3281]) and 0.2595 (95% CI [0.1988, 0.3172]). The finite-budget separation becomes more pronounced at the larger graph sizes in the main QAOA-depth comparison, even though QAOA itself improves as p increases.

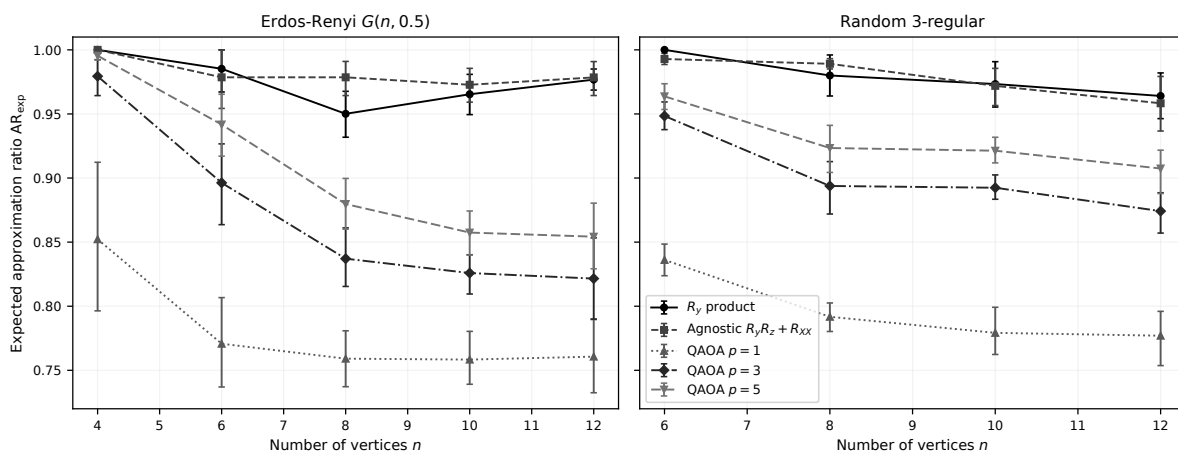


Figure 2. Expected approximation ratio versus graph size for the two stochastic graph families. Values are first averaged over five restarts within each graph; points denote graph-instance means and error bars denote 95% bootstrap confidence intervals across graph instances. The R_x and $R_y R_z$ controls are omitted for readability because their objective families are analytically equivalent to the R_y control.

Table 2. Overall benchmark summary.

Method	AR_{exp}	P_{opt}	Target (%)	Evaluations
R_y product	0.9784	0.7832	78.3	216.3
R_x product	0.9784	0.7832	78.3	216.3
$R_y R_z$ product	0.9787	0.7832	78.3	334.2
Agnostic $R_y R_z + R_{xx}$	0.9811	0.8021	76.6	696.7

Table 2. Cont.

Method	AR_{exp}	P_{opt}	Target (%)	Evaluations
QAOA $p = 1$	0.7954	0.2101	2.1	42.4
QAOA $p = 3$	0.8908	0.4284	9.1	730.7
QAOA $p = 5$	0.9202	0.5278	19.4	929.2
Single-flip local search	0.9684	–	73.3	29.5

Values are arithmetic means over the 95 graph-level summaries, after averaging five independent restarts for each graph. “Target” is the fraction of runs with $AR_{\text{exp}} \geq 0.99$. The single-flip baseline has no quantum sampling probability P_{opt} .

3.3. Representative Graph-Agnostic Entanglement

Turning now to the effect of entanglement, the graph-agnostic $R_y R_z + R_{xx}$ comparator did not exhibit a uniform advantage over the product-state landscape. Relative to R_y , its mean approximation-ratio change on Erdős–Rényi graphs was +0.0062 (equivalently $AR_{R_y} - AR_A = -0.0062$, 95% CI $[-0.0124, 0.0002]$), whereas on 3-regular graphs it was -0.0013 (95% CI for $AR_{R_y} - AR_A$: $[-0.0080, 0.0102]$). The paired P_{opt} differences likewise had confidence intervals containing zero for both families. Thus the small increase in the overall mean reported in Table 2 is not a consistent graph-family-level improvement.

The optimization-cost difference is clearer. Compared with R_y , the entangling comparator required on average 449.7 additional objective evaluations on Erdős–Rényi graphs (95% CI $[420.9, 476.5]$) and 517.9 additional evaluations on 3-regular graphs (95% CI $[497.1, 539.1]$). Moreover, the R_y target-success rate exceeded that of the entangling comparator by 9.5 percentage points on 3-regular graphs (95% CI $[1.5, 17.5]$), while the difference on Erdős–Rényi graphs was not resolved by the present sample. These observations apply only to the fixed-chain comparator of Equation (18); they do not establish a general ordering over problem-agnostic entangling ansätze.

3.4. Parameter-Count-Matched QAOA

Regarding parameter count, a direct parameter-count control was obtained by setting $p = n/2$, so that QAOA uses the same number of trainable angles, $2p = n$, as the R_y product ansatz. Across the complete benchmark, parameter-matched QAOA reached mean $AR_{\text{exp}} = 0.9109$, $P_{\text{opt}} = 0.4871$, and a target-success rate of 9.5%, with a mean of 856.5 objective evaluations. The corresponding R_y values were 0.9784, 0.7832, 78.3%, and 216.3 evaluations.

On Erdős–Rényi graphs the graph-level paired advantage of R_y was 0.0872 in approximation ratio (95% CI $[0.0760, 0.0987]$), and on 3-regular graphs it was 0.0509 (95% CI $[0.0439, 0.0578]$). Figure 3 shows the size-resolved paired differences. At $n = 12$, where the matched control corresponds to QAOA $p = 6$, QAOA used the full 1000-evaluation budget in every run and reached a mean approximation ratio of 0.9012, compared with 0.9718 for R_y . Equal parameter count therefore does not explain the finite-budget difference observed here. It should not, however, be interpreted as equalizing circuit depth, quantum resources, or the intrinsic cost of an objective evaluation.

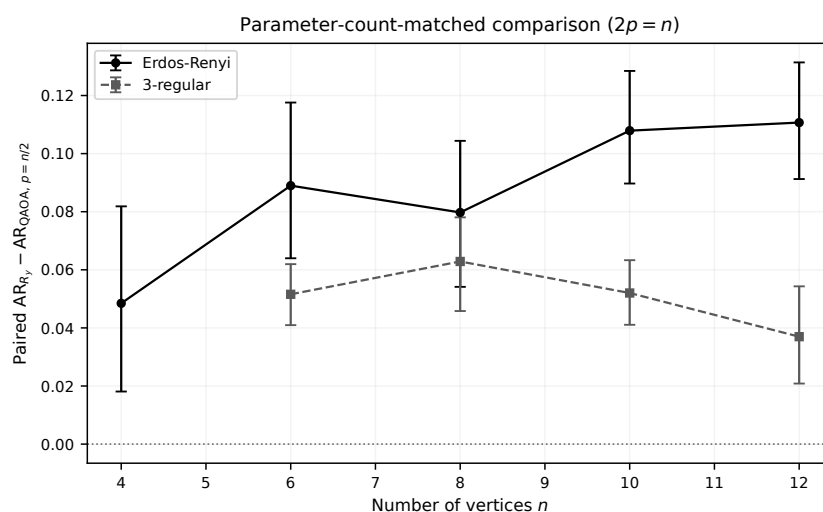


Figure 3. Parameter-count-matched comparison between the R_y product ansatz and QAOA with $p = n/2$, so that both have n trainable parameters. Positive values favor the R_y ansatz. Points denote mean graph-level paired differences; error bars denote 95% bootstrap confidence intervals.

3.5. Finite-Budget Convergence and Classical Context

The final experimental results concern convergence. In particular, the common evaluation cap exposes substantial differences in optimization dynamics. Figure 4 reports the mean approximation ratio after common evaluation checkpoints, carrying forward the final value of runs that terminated before a checkpoint. The R_y product ansatz reached a mean ratio of 0.9439 after 50 evaluations, 0.9724 after 100, and 0.9781 after 250, after which the aggregate curve changed little. The graph-agnostic entangling comparator started lower (0.8263 at 50 evaluations) but continued improving to 0.9811 at the 1000-evaluation endpoint. QAOA $p = 3$ increased from 0.8596 to 0.8908 and QAOA $p = 5$ from 0.8598 to 0.9202 over the same checkpoints.

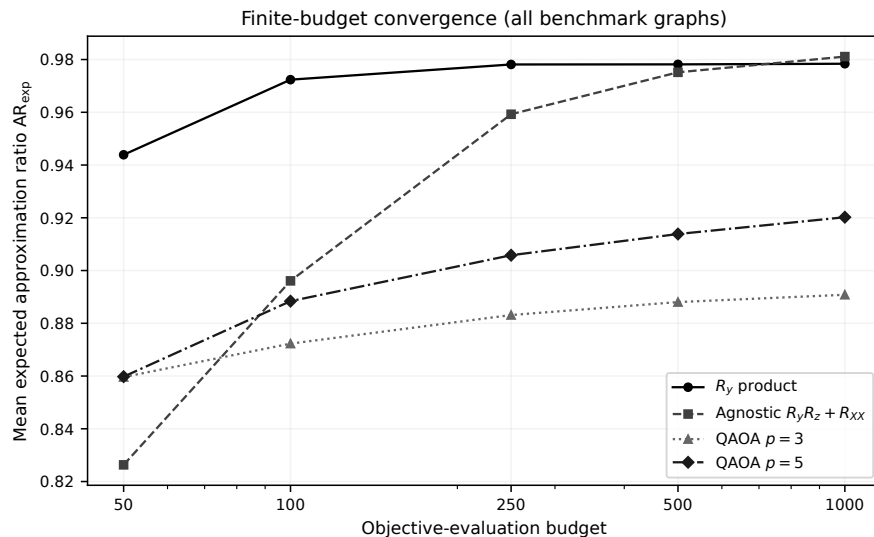


Figure 4. Mean expected approximation ratio as a function of the common objective-evaluation budget across all 95 benchmark graphs. For runs terminating before a checkpoint, the final attained value is carried forward. The continuing rise of the deeper QAOA and graph-agnostic entangling curves near the cap is part of the finite-budget interpretation and should not be read as an asymptotic performance ordering.

The optimizer-status records support the same interpretation. COBYLA reported successful termination in 99.2% of R_y runs, compared with 67.8% for the graph-agnostic entangling comparator, 55.2% for QAOA $p = 3$, and only 10.9% for QAOA $p = 5$; most unsuccessful deep-QAOA runs exhausted the evaluation cap. Accordingly, the reported QAOA differences characterize performance under the predefined finite classical-optimization budget, not the best values attainable with unlimited optimization or specialized QAOA parameter strategies.

The single-vertex-flip classical baseline provides an additional scale for interpreting the product-state results. It obtained a mean approximation ratio of 0.9684 with 73.3% target success using an average of 29.5 explicit cut-value evaluations, counted as one evaluation of the initial assignment and one evaluation for each candidate single-vertex flip inspected during the search. This count provides an internal measure of effort for the classical heuristic and should not be interpreted as computationally equivalent to the objective evaluations used by the variational optimizations. Relative to local search, R_y improved the graph-level mean ratio by 0.0196 on Erdős–Rényi graphs (95% CI [0.0088, 0.0319]), but the difference on 3-regular graphs was -0.0008 with CI $[-0.0076, 0.0055]$. The competitiveness of this elementary classical heuristic is consistent with Proposition 4: the product-state landscape inherits the same one-flip local structure rather than providing evidence of a quantum computational advantage.

4. Discussion

Taken together, the analytical and numerical results support a clear distinction between representational capacity and finite-budget trainability. For MaxCut, product states are sufficient to represent every classical assignment and therefore an exact optimum. This fact alone is elementary and does not make the product-state ansatz an efficient solver. What is useful in the present setting is that the induced continuous objective can be characterized exactly: its strict local maxima coincide with strict one-vertex-flip local optima, R_x and R_y are objective-equivalent, and terminal R_z phases generate exact flat directions. The numerical study then uses these identities as controls for understanding optimizer behavior rather than as claims of a new quantum resource.

The redundancy experiment illustrates why nominal parameter count can be a poor proxy for task-relevant expressivity. Adding n gates of type R_z to the R_y product ansatz leaves the MaxCut expectation unchanged for

every parameter value, yet COBYLA systematically required more evaluations as n increased. In this case the additional parameters are not merely unnecessary at the optimum; they are globally invisible to the objective. The result is therefore a controlled example in which overparameterization increases the dimension explored by a derivative-free optimizer without enlarging the objective family.

The entangling comparison leads to an equally qualified conclusion. The fixed-chain R_{xx} architecture achieved an overall mean approximation ratio close to, and slightly above, the product-state result, but the paired graph-family comparisons did not show a consistent improvement and the optimization budget was substantially larger. This is compatible with the broader observation that entanglement must be interpreted together with its alignment to the problem structure [17], but the present experiment is intentionally narrower: one fixed, ordering-dependent graph-agnostic architecture cannot establish that agnostic entanglement is beneficial or detrimental in general. Its role is to demonstrate that introducing a generic entangling block does not automatically translate into easier finite-budget optimization on these instances.

Problem alignment, however, does not guarantee easy classical training. Increasing QAOA depth improved both AR_{exp} and P_{opt} , and on complete graphs QAOA $p = 5$ approached the exact optimum. Nevertheless, on the stochastic graph families the deeper QAOA models frequently consumed the full evaluation budget and remained below the product-state baseline. The parameter-count-matched control shows that this gap cannot be attributed solely to QAOA using fewer trainable angles. At the same time, the continuing improvement of QAOA near 1000 evaluations prevents a stronger asymptotic conclusion. Alternative initialization, parameter transfer, layerwise strategies, or different classical optimizers could change this finite-budget ordering. Optimizer dependence has been studied extensively for QAOA, and comparative analyses indicate that the relative behavior of classical optimization methods depends on the problem, parameterization, and computational budget rather than identifying a universally preferred optimizer [5, 21]. QAOA-specific structure can also be exploited independently of optimizer choice. In particular, concentration of optimal parameters with increasing problem size has been identified analytically and numerically at low QAOA depths [24], while parameter-transfer strategies have been shown to reuse parameters optimized on small donor graphs for substantially larger MaxCut instances under suitable structural-similarity conditions [25]. Moreover, recent dimensionality-reduction approaches indicate that structure in the QAOA parameter space can be exploited to reduce the number of circuit evaluations required for optimization, including through nonlinear low-dimensional representations at increasing circuit depth [26]. These results emphasize that random initialization followed by a common black-box optimizer, as deliberately adopted here, represents one controlled training regime rather than the best attainable QAOA optimization strategy. The present result is therefore specifically that problem-structured dynamics did not yield the easiest landscape for a common black-box optimizer under the predefined budget, rather than that QAOA is intrinsically more difficult to optimize under specialized parameter-selection strategies.

The classical local-search comparison is essential to the interpretation. The R_y objective is efficiently evaluable in closed form and its strict local maxima are classical one-flip optima. A simple best-improvement local search was already highly competitive while using far fewer objective evaluations. Among the explicitly classical methods considered in the benchmark, the single-vertex-flip heuristic therefore provides the strongest direct classical comparator shown. Consequently, the strong performance of the product-state ansatz should not be described as quantum advantage, nor does the present work compare against state-of-the-art classical MaxCut algorithms. The product-state circuit instead serves primarily as an analytically transparent bridge between a variational quantum parameterization and a classical continuous relaxation of the same discrete objective.

Several limitations define the scope of these conclusions. Most importantly for the numerical comparison, the experiments use exact noiseless simulation and graphs with at most 12 vertices. The restriction to $n \leq 12$ prevents the observed finite-budget performance ordering from being empirically extrapolated to graphs with tens or hundreds of vertices. The product-state representational result itself is independent of this restriction: every computational-basis cut, including an exact optimum, remains contained in the product-state family for arbitrary n . This representational statement does not imply size-independent trainability, however. The R_y parameterization contains n continuous angles and its landscape retains the one-vertex-flip local-optimum structure characterized above, so its optimization behavior under a fixed evaluation budget may change substantially as n grows. Conversely, standard QAOA at fixed depth p retains $2p$ global variational parameters independently of n , and parameter-concentration and transferability phenomena can facilitate training across problem sizes in some graph families [24, 25]. Increasing p can in turn introduce additional circuit and optimization costs. These competing effects provide no basis for predicting a monotonic continuation of the small-instance performance gap observed here. The analytical results allow nonnegative weights, but the numerical benchmark is unweighted. COBYLA is the only variational optimizer, and a cap on objective evaluations equalizes classical optimizer calls rather than wall-clock cost, circuit depth,

or measurement resources. Only one fixed-chain graph-agnostic entangling architecture is tested, and its vertex ordering is not permutation invariant. The stochastic ensembles contain ten graph instances per size, and the classical comparison is deliberately limited to a simple local search rather than stronger approximation algorithms. Finally, the main QAOA sweep uses $p = 1, 3, 5$, with additional depths only for the $2p = n$ parameter-matched control. These choices make the experiment reproducible and interpretable, but they also mean that the numerical ordering should not be extrapolated to large-scale, noisy, hardware-constrained, or highly optimized QAOA regimes.

Future work should test whether this landscape-based interpretation persists for weighted instances, alternative graph families, and larger systems accessible through tensor-network simulation or quantum hardware, together with different classical optimizers and QAOA-specific initialization or parameter-transfer schemes. The analytical controls also suggest a broader design principle for variational algorithms: before increasing ansatz size, it is useful to identify which parameter directions are visible to the objective and which structural properties change the optimization landscape rather than only the state-space expressivity.

5. Conclusions

We used MaxCut as a controlled setting to separate objective-visible parameterization, generic entanglement, and problem-specific QAOA structure. The product-state limit yields an exact and transparent landscape: the discrete optimum is preserved, R_x and R_y are pointwise equivalent for the diagonal objective, terminal R_z phases are exactly redundant, and strict local maxima correspond precisely to strict one-vertex-flip classical optima. Under a common finite COBYLA budget on 95 small graphs, the simple R_y landscape was optimized more efficiently than the tested QAOA depths and a representative graph-agnostic entangling circuit, while parameter redundancy incurred a measurable optimization penalty without increasing objective-relevant expressivity. QAOA performance improved with depth and remained in active improvement near the evaluation cap, so these results characterize trainability under the chosen budget rather than asymptotic algorithmic superiority. Together with the strong single-flip classical baseline, the findings support a conservative conclusion: for the small exact-simulation regime considered here, the geometry induced by task-relevant parameterization can matter at least as much as nominal circuit richness, while no quantum computational advantage is implied.

Author Contributions

L.L.: Conceptualization, methodology, software, validation, formal analysis, investigation, visualization, writing—original draft preparation; M.P.: Conceptualization, supervision, validation, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The source code, frozen experiment configuration, raw benchmark outputs, statistical-analysis scripts, and figure-generation scripts supporting the numerical results of this study are publicly available at <https://github.com/leonardoLavagna/maxcut-vqa-architectures>. The repository contains the deterministic seeds and stored experiment records required to reproduce the core numerical analyses, tables, and figures reported in the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

Use of AI and AI-Assisted Technologies

During the preparation of this work, the authors used ChatGPT (OpenAI) to assist with language editing and code review. No generative-AI tools were used for experimental design, data analysis, interpretation of the results,

or the creation or alteration of figures, tables, or other scientific results. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the article.

References

1. McClean, J.R.; Romero, J.; Babbush, R.; et al. The Theory of Variational Hybrid Quantum-Classical Algorithms. *New J. Phys.* **2016**, *18*, 023023. <https://doi.org/10.1088/1367-2630/18/2/023023>.
2. Cerezo, M.; Arrasmith, A.; Babbush, R.; et al. Variational quantum algorithms. *Nat. Rev. Phys.* **2021**, *3*, 625–644. <https://doi.org/10.1038/s42254-021-00348-9>.
3. Wang, Z.; Huang, J.; Ye, R. A Review of Variational Quantum Algorithms: Insights into Fault-Tolerant Quantum Computing. *arXiv* **2026**, arXiv:2604.07909.
4. Farhi, E.; Goldstone, J.; Gutmann, S. A Quantum Approximate Optimization Algorithm. *arXiv* **2014**, arXiv:1411.4028.
5. Blekos, K.; Brand, D.; Ceschini, A.; et al. A Review on Quantum Approximate Optimization Algorithm and Its Variants. *Phys. Rep.* **2024**, *1068*, 1–66. <https://doi.org/10.1016/j.physrep.2024.03.002>.
6. Korte, B.; Vygen, J. *Combinatorial Optimization: Theory and Algorithms*, 6th ed.; Springer: Berlin/Heidelberg, Germany, 2018.
7. Wang, Z.; Hadfield, S.; Jiang, Z.; et al. Quantum approximate optimization algorithm for MaxCut: A fermionic view. *Phys. Rev. A* **2018**, *97*, 022304. <https://doi.org/10.1103/physreva.97.022304>.
8. Herrman, R.; Lotshaw, P.C.; Ostrowski, J.; et al. Multi-angle Quantum Approximate Optimization Algorithm. *Sci. Rep.* **2022**, *12*, 6781. <https://doi.org/10.1038/s41598-022-10555-8>.
9. Chalupnik, M.; Melo, H.; Alexeev, Y. Augmenting QAOA Ansatz with Multiparameter Problem-Independent Layer. In Proceedings of the IEEE International Conference on Quantum Computing and Engineering (QCE), Broomfield, CO, USA, 18–23 September 2022; pp. 97–103.
10. Egger, D.J.; Mareček, J.; Woerner, S. Warm-starting Quantum Optimization. *Quantum* **2021**, *5*, 479. <https://doi.org/10.22331/q-2021-06-17-479>.
11. Wurtz, J.; Love, P. Classically Optimal Variational Quantum Algorithms. *IEEE Trans. Quantum Eng.* **2021**, *2*, 3104107. <https://doi.org/10.1109/tqe.2021.3122568>.
12. Rajakumar, J.; Golden, J.; Bärttschi, A.; et al. Trainability Barriers in Low-Depth QAOA Landscapes. In Proceedings of the CF '24: 21st ACM International Conference on Computing Frontiers, Naples, Italy, 7–9 May 2024; pp. 199–206. <https://doi.org/10.1145/3649153.3649204>.
13. Larocca, M.; Thanasilp, S.; Wang, S.; et al. Barren plateaus in variational quantum computing. *Nat. Rev. Phys.* **2025**, *7*, 174–189. <https://doi.org/10.1038/s42254-025-00813-9>.
14. Mao, R.; Yuan, P.; Allcock, J.; et al. QAOA-MaxCut has barren plateaus for almost all graphs. *arXiv* **2025**, arXiv:2512.24577.
15. De Falco, F.; Lavagna, L.; Ceschini, A.; et al. Enhancing QAOA Ansatz via Multi-Parameterized Layer and Blockwise Optimization. In Proceedings of the ICASSP 2026—2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 3–8 May 2026. <https://doi.org/10.1109/icassp55912.2026.11461345>.
16. Bauer, N.M.; Alam, R.; Siopsis, G.; et al. Combinatorial Optimization with Quantum Imaginary Time Evolution. *Phys. Rev. A* **2024**, *109*, 052430. <https://doi.org/10.1103/physreva.109.052430>.
17. Rohe, T.; Baumann, M.; Harjes Ruiloba, F. Detrimental Agnostic Entanglement: The Case Against Hardware-Efficient Ansätze for Combinatorial Optimization. *arXiv* **2026**, arXiv:2605.19827.
18. Banerjee, D.; Mahapatra, S.; Chaudhury, K.N. Local-Minima-Preserving Continuous Relaxation of Ising Problems. In Proceedings of the International Conference on Machine Learning, Seoul, Republic of Korea, 6–11 July 2026.
19. Zhou, L.; Wang, S.T.; Choi, S.; et al. Quantum Approximate Optimization Algorithm: Performance, Mechanism, and Implementation on Near-Term Devices. *Phys. Rev. X* **2020**, *10*, 021067. <https://doi.org/10.1103/physrevx.10.021067>.
20. Sherali, H.D. Convex Envelopes of Multilinear Functions over a Unit Hypercube and over Special Discrete Sets. *Acta Math. Vietnam.* **1997**, *22*, 245–270.
21. Fernández-Pendás, M.; Combarro, E.F.; Vallecorsa, S.; et al. A Study of the Performance of Classical Minimizers in the Quantum Approximate Optimization Algorithm. *J. Comput. Appl. Math.* **2022**, *404*, 113388. <https://doi.org/10.1016/j.cam.2021.113388>.
22. Goemans, M.X.; Williamson, D.P. Improved Approximation Algorithms for Maximum Cut and Satisfiability Problems Using Semidefinite Programming. *J. ACM* **1995**, *42*, 1115–1145. <https://doi.org/10.1145/227683.227684>.
23. Allcock, J.; Yuan, P.; Zhang, S. The Dynamical Lie Algebra of QAOA-MaxCut on the Complete Graph. *arXiv* **2026**, arXiv:2607.00945.
24. Akshay, V.; Rabinovich, D.; Campos, E.; et al. Parameter Concentrations in Quantum Approximate Optimization. *Phys. Rev. A* **2021**, *104*, L010401. <https://doi.org/10.1103/physreva.104.L010401>.
25. Galda, A.; Gupta, E.; Falla, J.; et al. Similarity-Based Parameter Transferability in the Quantum Approximate Optimization Algorithm. *Front. Quantum Sci. Technol.* **2023**, *2*, 1200975. <https://doi.org/10.3389/frqst.2023.1200975>.
26. Brahmandam, S.; Ramkumar, V. Dimensionality Reduction of QAOA Parameter Space with Kernel PCA for Max-Cut. *arXiv* **2026**, arXiv:2606.23718.