Operator-Spectral Truncated Priors for Query-Efficient QAOA Parameter Search

Molena Huynh[1,*]

[1] *North Carolina State University, Raleigh, North Carolina 27695, USA*

(Dated: June 29, 2026)

We introduce OST-QAOA, a method that searches in operator space rather than raw parameter-vector space to overcome the true bottleneck of low-depth QAOA: the objective queries available to tune angles on each new graph. It maps each graph to a truncated symmetric operator on the $2p$-dimensional angle space, built from two noncommuting graph-derived generators, and converts its low-rank spectral decomposition into ranked collective search directions for a deterministic, auditable query policy. We prove the prior covariance is positive definite, truncation is Loewner-monotone, the construction reduces to per-angle coordinate refinement in its commutative limit, and truncation caps the effective search dimension; a validated scaling argument links truncation to query cost. On exact-statevector MaxCut over four graph families, a ten-seed study (160 paired held-out instances) beats a budget-matched TQA baseline by $+0.107 \pm 0.011$ mean approximation ratio (157/160 wins) and reaches 98% of the best ratio in 10.5 queries against 24.7, a 2.4× reduction; at the reference point (p=3, rank 4, weight 4.0) it reaches 0.820 ± 0.019, $+0.112$ over baseline (sign-test $p < 10^{-4}$). A truncation sweep shows the predicted interior optimum, and an ablation localizes the gain to noncommutative directions (-0.010 when restricted to diagonal ones). Shipping as the package `uq-qaoa`, where changing depth, budget, rank, weight, seed, or graph family regenerates every figure and query trace, the method is a reproducible operator-valued prior for query-efficient variational search on graphs.

I. INTRODUCTION

Many combinatorial graph-optimization problems—including instances arising in biomedical settings such as pathway-subset selection and patient-similarity partitioning—can be cast as graph objectives, with MaxCut a canonical example of the broad class of NP-hard problems admitting Ising formulations [1, 2], classically approached by the Goemans–Williamson semidefinite relaxation [3]. QAOA [4, 5] provides a variational quantum template for them, with MaxCut its canonical benchmark [6–8] and one instance of the broader variational-quantum-algorithm paradigm [9]. The alternating-operator structure generalizes beyond the original mixer [10], admits a fermionic analysis on MaxCut [11], and is itself computationally universal [12]; see [13] for a survey. In the noisy intermediate-scale regime [14], however, practical use is gated less by the circuit than by the *classical policy* that decides how to spend a small budget of expensive objective evaluations on each new instance—a classical optimization that is itself NP-hard in general [15] and sensitive to the choice of optimizer on noisy objectives [16–18]. A good warm-start point alone is not enough: under a fixed query budget the optimizer also needs to know *which angle directions* deserve exploration.

A large body of work shows that good QAOA angles transfer across instances. Objective values concentrate for fixed parameters on typical instances [19, 20], optimal parameters concentrate with system size [21], and parameters trained on one graph transfer to others [22, 23], an ingredient even in recent demonstrations of scaling advantage on hardware [24–26]. Trotterized–annealing (TQA) schedules supply strong instance-independent initializations [27], reflecting the adiabatic and quantum-annealing roots of the ansatz [28, 29]. Related warm-start and learned-initialization strategies start the optimizer from a relaxation solution [30], a graph-neural-network prediction [31], or a meta-learned policy [32]. These results motivate *transferring structure* between graphs; what they do not specify is a budgeted search geometry—the directions, and their relative scales, along which a few queries should be spent once a warm start is in hand.

The central idea of this work is to learn and search in *operator space* rather than only in parameter-vector space. For a graph G, OST-QAOA constructs two $2p \times 2p$ symmetric operators from the graph Laplacian spectrum, degree moments, and topology features. Their noncommuting product defines a commutator energy, and a truncated spectral decomposition of the resulting angle operator gives a low-rank covariance. The covariance is not used as a passive uncertainty estimate; it becomes the ranked list of search directions evaluated by a deterministic query policy that can be audited line by line.

This is motivated by, but distinct from, the spectral-truncation kernel framework of Hashimoto *et al.* [33], in which truncation of positive-definite C^*-algebraic kernels controls noncommutativity and trades representation power against model complexity. Here the output is not a supervised vector-valued regression kernel. It is a graph-conditioned prior over QAOA angles, together with a deterministic objective-query policy. We adopt the design idea—finite-rank noncommutativity as a tractable way to encode interaction—and ask what it buys in the operational currency that limits near-

* molena.huynh@jmp.com

term QAOA, namely the number of objective queries. The same kernel/feature-space viewpoint underlies quantum machine learning [34], and using machine-learned structure to guide combinatorial search is a broader theme [35]; the prospect of a quantum advantage from low-depth QAOA [36] makes the query cost of tuning it the decisive near-term quantity.

a. Scope. To keep every reported number exactly reproducible on a CPU, all experiments use small exact statevectors ($n \leq 10$) and brute-force MaxCut optima on standard synthetic graph families. This is a methods study of a search policy, not a claim of hardware-scale or domain (e.g. biomedical) deployment; section X states the limitations and the path to scale.

This paper makes four contributions. First, OST-QAOA constructs a truncated, noncommuting operator on QAOA angle space, rather than a scalar graph feature or a diagonal uncertainty vector, from two graph-derived generators $\mathcal{A}_G, \mathcal{B}_G$ that do not commute, and spectrally truncates it to a rank-n covariance. Second, the off-diagonal structure of the operator covariance is used operationally: the query policy probes the *collective* eigendirections of Σ_G rather than the raw angle coordinates, and restricting the search to the diagonal (commutative) directions removes a measurable part of the advantage (-0.010 mean ratio, section VIII A). Third, we establish a query-budget truncation tradeoff with theory and a stated assumption: we prove that the truncated prior is positive definite and Loewner-monotone, that it reduces exactly to the commutative diagonal prior in the $n \to 1$ limit, and that truncation caps the prior's effective dimension $d_{\text{eff}}(n)$ (sections III and IV); under an explicitly stated coverage model we give a *heuristic* scaling argument relating $d_{\text{eff}}(n)$ to objective-query cost, and the predicted interior optimum $n^\star$ is then observed empirically (fig. 3). Fourth, the artifact is reproducible and installable with a multi-seed evaluation: every figure, table, CSV, and query trace is regenerated by the package from a single seed, and the headline advantage is shown to hold across ten seeds ($+0.107 \pm 0.011$, 157/160 instance wins; fig. 4). The package installs with `pip install .` from `submission/code` and exposes both a CLI (`uqqaoa-reproduce`) and a Python API for retuning every experimental parameter.

This work continues a single-author program on spectral truncation of operators for query-efficient quantum optimization. Earlier entries develop graph-conditioned trust regions for query-efficient and uncertainty-calibrated QAOA [37–39], the measurement cost of warm-started low-depth QAOA and certified query budgets [40, 41], and topology-conditioned parameter transfer for budgeted graph optimization [42, 43]. The common thread is to fix a small objective-query budget and decide, from graph structure, where those queries are spent. Here we move that decision from a scalar trust region or a transferred point into a truncated, noncommuting operator on the full angle space, so the search geometry itself is conditioned on the graph and the truncation rank sets the number of collective directions probed.

II. PRELIMINARIES

A. QAOA and the angle-space objective

Let $G = (V, E)$ be an undirected graph with $n = |V|$. The MaxCut Hamiltonian is

$$H_C(G) = \sum_{(i,j)\in E} \frac{1 - Z_i Z_j}{2}. \tag{1}$$

At depth p, the QAOA parameter vector is $\theta = (\gamma_1, \ldots, \gamma_p, \beta_1, \ldots, \beta_p) \in [0, \pi]^{2p}$, and with mixer $H_M = \sum_i X_i$ the state and objective are

$$|\psi_G(\theta)\rangle = \prod_{\ell=1}^{p} e^{-i\beta_\ell H_M} e^{-i\gamma_\ell H_C(G)} |+\rangle^{\otimes n}, \tag{2}$$

$$f_G(\theta) = \langle \psi_G(\theta) | H_C(G) | \psi_G(\theta) \rangle. \tag{3}$$

We report the approximation ratio $r_G(\theta) = f_G(\theta)/C_G^\star$, where $C_G^\star$ is computed exactly by brute force for the small audit graphs used here. An algorithm receives a fixed query budget Q, where one query is one exact evaluation of $f_G(\theta)$; each method is measured by the best ratio found within the same Q. This query model is deliberately simple: it isolates the angle-search policy from finite-shot readout and hardware noise, which section X discusses as the next experimental layer. The $2p$-dimensional angle space $[0, \pi]^{2p}$ plays the role that the data function domain plays for an operator-valued kernel: it is the space on which our operators act.

B. Spectral truncation on the angle operator

Spectral truncation replaces an operator by the part of its spectrum carried by its leading directions. Given a symmetric operator $\mathcal{O}$ on angle space with eigendecomposition $\mathcal{O} = U \Lambda U^\mathsf{T}$, retaining the r leading eigenpairs and clipping their eigenvalues to be nonnegative yields a rank-r positive-semidefinite operator. In the C^*-algebraic setting this is the truncation operation controlled by a parameter that interpolates between commutative and noncommutative products [33]; here it is the truncation of the angle operator that we construct in section III.

Definition 1 (Spectral truncation on angle space)**.** *Let $\mathcal{O}_G = U \Lambda U^\mathsf{T}$. For truncation rank r, define $\mathcal{O}_{G,r} = U_r \Lambda_r U_r^\mathsf{T}$, where U_r contains the r leading eigenvectors after nonnegative clipping of the corresponding eigenvalues.*

The truncation parameter r is analogous in spirit to spectral truncation in C^*-algebraic kernels [33], but here it controls the number of QAOA angle directions that can be emphasized by the search covariance.

C. Existing angle-search priors

Existing classical policies for low-depth QAOA sit at two extremes, mirroring the separable and commutative ends of operator-valued kernels. A *fixed-schedule* prior such as the TQA schedule [27] supplies one global angle vector independent of the graph—the analogue of a separable kernel that ignores interactions across the domain. A *diagonal* prior, whether a per-coordinate uncertainty vector or a feature-neighbour (kNN) posterior built from transfer [22, 23], refines each angle coordinate independently—the analogue of a commutative kernel that captures only pointwise structure. Neither emphasizes *collective*, off-coordinate angle directions. Table I summarizes these priors and the proposed one along the properties that distinguish them: whether a graph-derived operator is used, whether its generators are noncommuting, whether the search probes off-diagonal (collective) directions, and whether the queries are ranked.

III. OPERATOR-SPECTRAL TRUNCATED PRIOR ON QAOA ANGLE SPACE

For graph G, let λ_i denote normalized Laplacian eigenvalues and let d_i denote centered degree values; these spectral quantities are the classical descriptors of graph structure in spectral graph theory [44, 45]. Define two symmetric angle-space operators,

$$\mathcal{A}_G = \mathrm{Toeplitz}(0.70\, m_\lambda + 0.30\, \phi(G))\,, \tag{4}$$

$$\mathcal{B}_G = \mathrm{sym}[\mathrm{roll}\,\mathrm{Toeplitz}(0.55\, m_d + 0.45\, \phi(G)_{\mathrm{rev}})]\,, \tag{5}$$

where m_λ and m_d are spectral and degree moments of length $2p$, and $\phi(G)$ is the normalized graph descriptor used by the package. The operator $\mathcal{A}_G$ carries *global* spectral information and $\mathcal{B}_G$ carries *local* degree information; they are designed not to commute, so that $\mathcal{C}_G = [\mathcal{A}_G, \mathcal{B}_G] = \mathcal{A}_G\mathcal{B}_G - \mathcal{B}_G\mathcal{A}_G \neq 0$ in general, encoding local/global structure that no single symmetric generator captures. The final raw angle operator is

$$\mathcal{O}_G = \mathrm{sym}\big(0.55\mathcal{A}_G + 0.45\mathcal{B}_G + \omega\, \mathcal{C}_G^{\mathsf{T}}\mathcal{C}_G\big)\,, \tag{6}$$

with $\omega = 4.0$ at the reference operating point, and is then spectrally truncated to $\mathcal{O}_{G,r}$ by definition 1. The mixing weights are fixed package defaults, held constant across all experiments; the only swept design parameters are the truncation rank and the commutator weight ω. Figure 1 gives an overview of this construction on a held-out graph: the truncated noncommuting operator, its retained spectrum, and the resulting collective search widths.

A. Connection between the proposed prior and existing priors

The truncation rank and the commutator weight interpolate between the proposed prior and the existing priors of section II C.

Proposition 1 (Commutative limit recovers the diagonal prior)**.** *The commutator weight ω parameterizes noncommutativity: at $\omega = 0$ the raw operator $\mathcal{O}_G = \mathrm{sym}(0.55\,\mathcal{A}_G + 0.45\,\mathcal{B}_G)$ drops the commutator $\mathcal{C}_G = [\mathcal{A}_G, \mathcal{B}_G]$ entirely; and restricting Σ_G to its diagonal makes the eigenframe V the coordinate axes, in which case the search of section V reduces exactly to per-angle coordinate refinement. The commutator-off and diagonal-covariance variants of section VIII are therefore the commutative special cases of* OST-QAOA*, and any gap between them and the full method measures the operational value of the noncommutative geometry.*

Thus the fixed-schedule and diagonal priors of table I are recovered as the $n \to 1$ and $\omega \to 0$ degenerations of the proposed construction, exactly as separable and commutative kernels are recovered as limits of the spectral-truncation kernel [33].

B. Convergence and interactions

As the truncation rank grows, the truncated operator converges monotonically to the positive-semidefinite part of the full operator, and the number of angle directions it activates—its effective dimension—grows accordingly. This is the operator analogue of how the truncation parameter of a spectral-truncation kernel controls the balance of local and global interactions.

Proposition 2 (Monotone spectral truncation)**.** *Order the clipped eigenvalues of $\mathcal{O}_G$ as $\sigma_1 \geq \cdots \geq \sigma_{2p} \geq 0$. Then truncation is Loewner-monotone in the rank, $\mathcal{O}_{G,r} \preceq \mathcal{O}_{G,r+1}$ for every r, with $\mathrm{tr}\,\mathcal{O}_{G,r} = \sum_{j\leq r}\sigma_j$ and $\mathcal{O}_{G,2p}$ the PSD part of $\mathcal{O}_G$. The truncation parameter therefore interpolates between a one-direction prior (r=1) and the full operator prior (r=2p).*

Proof. $\mathcal{O}_{G,r+1} - \mathcal{O}_{G,r} = \sigma_{r+1}u_{r+1}u_{r+1}^{\mathsf{T}} \succeq 0$ since $\sigma_{r+1} \geq 0$; the trace identity follows from orthonormality of the u_j. □

Definition 2 (Effective search dimension)**.** *For a PSD covariance Σ with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_d$, the effective dimension is the participation ratio $d_{\mathrm{eff}}(\Sigma) = (\sum_j \lambda_j)^2 / \sum_j \lambda_j^2 \in [1, d]$: the number of directions carrying comparable variance.*

Proposition 3 (Truncation reduces effective dimension)**.** *In the operator-dominated regime $\beta \to 1$ of eq.* (10) *(up to the ϵ-floor), $d_{\mathrm{eff}}(\Sigma_G) \leq r + \mathcal{O}(\epsilon)$, because $\mathcal{O}_{G,r}$ has at most r nonzero eigenvalues. Spectral truncation thus caps the number of collective angle directions the query policy explores; fig. 2 shows d_{eff} increasing monotonically with r.*

Figure 2 traces d_{eff} as the truncation parameter n grows: it increases monotonically and then saturates

TABLE I. Summary of the existing and the proposed angle-search priors. The proposed OST-QAOA prior is the only one that builds a graph-derived operator from noncommuting generators and searches its off-diagonal collective directions under a ranked query budget; the diagonal and fixed-schedule priors are its commutative and separable special cases (proposition 1).

Prior / policy	Graph operator used	Noncommuting generators	Off-diagonal search	Query-ranked
Random	—	×	×	×
TQA	fixed schedule	×	×	×
TQA+coordinate	fixed schedule	×	×	✓
kNN+coordinate	graph features	×	×	✓
OST diagonal	$\mathcal{O}_{G,r}$ (diagonal)	✓	×	✓
OST-QAOA (proposed)	$\mathcal{O}_{G,r}$ (full)	✓	✓	✓

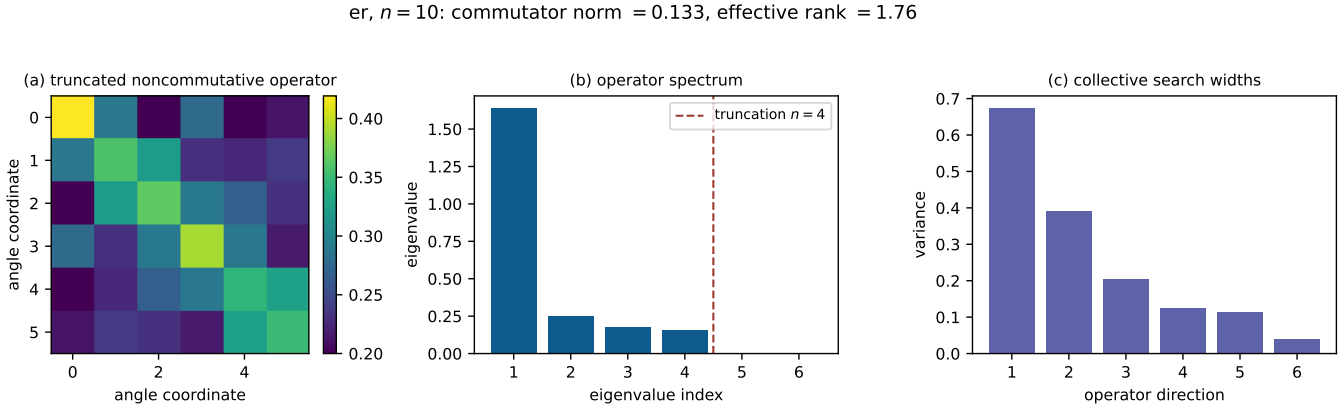

FIG. 1. Overview of the construction of the operator-spectral truncated prior $\mathcal{O}_{G,r} = U_r \Lambda_r U_r^{\mathsf{T}}$. A held-out graph G is mapped to (a) the truncated noncommuting angle operator $\mathcal{O}_{G,r}$, whose (b) spectrum is truncated at rank $r = n$ (here $n = 4$), giving (c) the collective search widths (directional variances) that the query policy probes.

as the truncated operator concentrates on a few collective directions, well below the no-concentration reference $d_{\text{eff}} = n$, instantiating proposition 3.

C. Positive definiteness

The construction yields a valid search frame: the prior covariance assembled from the truncated operator (eq. (10) below) is symmetric positive definite, so its eigendecomposition is an orthonormal basis with strictly positive directional variances.

Proposition 4 (Positive definiteness and a valid search frame)**.** *For any floor $\epsilon > 0$, commutator weight $\omega \geq 0$, and operator weight $\beta \geq 0$, the matrix Σ_G in eq. (10) is symmetric positive definite. Hence its eigendecomposition $\Sigma_G = V \Lambda V^{\mathsf{T}}$ yields an orthonormal frame V with strictly positive directional variances λ_j, so the query policy of section V is well defined.*

Proof. The neighbour scatter is a nonnegative combination of rank-one PSD terms and is PSD; the truncated operator $\mathcal{O}_{G,r} = U_r \Lambda_r U_r^{\mathsf{T}}$ is PSD by the nonnegative clipping of definition 1; and $\epsilon I \succ 0$. A sum of PSD matrices with a strictly positive-definite summand is positive definite; symmetry is explicit. □

IV. QUERY-BUDGET BOUND

Where the spectral-truncation kernel comes with a statistical generalization bound, the operator-spectral prior comes with an *operational* statement: the truncation parameter controls the number of objective queries needed to reach a target ratio, through the effective dimension of definition 2. In contrast to the exact propositions 2 to 4, this statement is *not* a theorem: it rests on a coverage assumption we make explicit, and we present it as a scaling heuristic to be checked against the data, not as a proved bound.

Assumption 1 (Coverage model)**.** *Write $m(r) \in [0, 1]$ for the fraction of the trust region's ε-optimal mass captured by the leading r operator directions, and assume (i) $m(r)$ is nondecreasing in r, and (ii) the refinement of section V reaches the ε-optimal subset once it has probed a number of collective directions proportional to $d_{\text{eff}}(\Sigma_G^{(r)})$, provided those directions retain the ε-optimal mass.*

Remark 1 (Query-budget scaling under assumption 1; heuristic). Under assumption 1, the expected queries to reach a target ratio τ scale as

$$\mathbb{E}[Q_G(\tau)] \lesssim \frac{c\, d_{\text{eff}}(\Sigma_G^{(r)})}{m(r)} \tag{7}$$

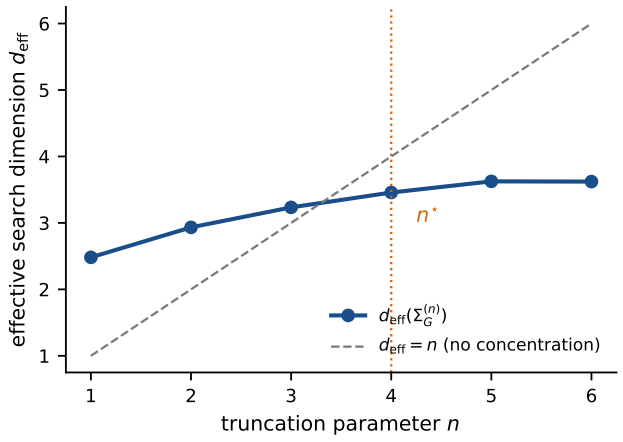

FIG. 2. Convergence and interactions. Effective search dimension $d_{\text{eff}}(\Sigma_G^{(n)})$ versus the truncation parameter n: as n grows the truncated operator $\mathcal{O}_{G,r}$ converges to the full operator (proposition 2) and d_{eff} rises monotonically toward a saturated value below the no-concentration reference $d_{\text{eff}} = n$ (dashed), the operator analogue of how a Féjer-type truncation controls the local/global interactions of the prior; $n^\star$ marks the sweep optimum of section VIII.

for a constant c independent of r. The numerator decreases with truncation (proposition 3) while $m(r)$ is nondecreasing; the minimizer $r^\star$ is therefore interior whenever aggressive truncation sheds near-optimal mass faster than it concentrates the search. This is the query-budget analogue of the representation-versus-complexity tradeoff of spectral-truncation kernels [33]: there the truncation parameter trades approximation against generalization; here it trades near-optimal coverage against the number of objective queries. We emphasize that eq. (7) is a scaling statement under assumption 1, with no sharp constant, and we treat its prediction—an interior $r^\star$—as a hypothesis to be tested.

The prediction of remark 1 is instantiated in table II (mean queries-to-target $\bar{Q}_{0.98}$) and in the truncation sweep of section VIII (figs. 2 and 3), where the effective dimension rises monotonically with n and the observed interior optimum $r^\star$=4 appears.

V. QUERY-EFFICIENT ANGLE SEARCH WITH THE TRUNCATED PRIOR

We illustrate the effect of the proposed prior by applying it to the query-efficient angle search. The package builds a deterministic offline library of small training graphs. For each training graph, a budgeted reference optimizer records a high-quality angle vector. For a test graph G, the vectorized upper triangle of $\mathcal{O}_{G,r}$, graph features, commutator norm, and effective rank form the operator signature $s(G)$. The k nearest library entries receive Gaussian weights $w_i(G) \propto \exp[-\|z(G) - z(G_i)\|_2^2/2h_G^2]$, where $z(\cdot)$ is the standardized signature and h_G is the median distance among the selected neighbours. The prior mean blends the weighted library mean with the fixed TQA schedule,

$$\mu_G = (1-\alpha)\sum_i w_i(G)\theta_i + \alpha\,\theta_{\text{TQA}}, \qquad \alpha = 0.25. \quad (8)$$

This neighbour-weighted prior mean plays the role of the mean function in a Gaussian-process prior over angles [46], while the covariance below supplies the search geometry probed by the policy, in the spirit of model-based and Bayesian optimization of expensive objectives [47, 48]. The covariance is operator-dominated: the truncated operator supplies the leading search geometry and the neighbour scatter

$$\Sigma_G^{\text{nbr}} = \sum_i w_i(G)(\theta_i - \mu_G)(\theta_i - \mu_G)^{\mathsf{T}} \quad (9)$$

a data-driven floor,

$$\Sigma_G = \beta\,\frac{\mathcal{O}_{G,r}}{\text{tr}(\mathcal{O}_{G,r})} + \left(1 - \tfrac{\beta}{2}\right)\frac{\Sigma_G^{\text{nbr}}}{\text{tr}(\Sigma_G^{\text{nbr}})} + \epsilon I, \qquad \beta = 0.8,\ \epsilon = 0.025, \quad (10)$$

so that the spectral-truncation parameter r and the noncommuting operator geometry, rather than the diagonal neighbour variances, set the directions the query policy probes.

a. Query-policy steps. The search evaluates a small set of anchors and then probes the leading eigendirections of Σ_G. The anchors are TQA, the operator posterior mean, the nearest operator-neighbour angle, and the global midpoint prior. The remaining queries evaluate positive and negative perturbations around the incumbent best angle, with step sizes proportional to the square root of the directional variance. The executable policy is:

1. build $\mathcal{O}_{G,r}$ from eq. (6) and compute (μ_G, Σ_G) from eq. (10);
2. evaluate the anchors $\{\theta_{\text{TQA}}, \mu_G, \theta_{\text{nearest}}, \theta_{\text{global}}\}$;
3. diagonalize $\Sigma_G = V\Lambda V^{\mathsf{T}}$;
4. for scales $s \in \{1.0, 0.55, 0.28\}$, evaluate incumbent perturbations $\pm s\sqrt{\lambda_j}v_j$ in descending variance order until the query budget is exhausted;

5. return the best angle and full query trace.

b. Computational cost. Building $\mathcal{O}_G$ and its truncation costs $\mathcal{O}(p^3)$ per graph from the $2p \times 2p$ operator, and is incurred once; each subsequent query is a single statevector evaluation of f_G. Because the prior ranks directions before spending queries, the dominant cost—objective evaluations—scales with the effective dimension $d_{\text{eff}}(\Sigma_G)$ of remark 1 rather than with the ambient $2p$, which is the source of the query savings reported in section VIII.

VI. EXTENSION TO A COMBINED MODEL

This section is for readers interested in a more flexible, learnable variant. As discussed in section III A, the truncation parameter n controls the global and local angle directions emphasized by the prior. To capture both at once, consider a combined prior assembled from L truncated operators $\mathcal{O}_{G,n_1}, \ldots, \mathcal{O}_{G,n_L}$ at different ranks with learnable weights $\beta_1, \ldots, \beta_L \geq 0$:

$$\Sigma_G^{\text{comb}} = \sum_{j=1}^{L} \beta_j \frac{\mathcal{O}_{G,n_j}}{\text{tr}(\mathcal{O}_{G,n_j})} + \epsilon I. \tag{11}$$

Setting a different n_j for each component lets the prior emphasize different levels of collective angle structure simultaneously: a low-rank component concentrates the budget on the dominant global direction, while a higher-rank component retains finer local directions. Each summand is PSD by definition 1, so Σ_G^{comb} is positive definite by the same argument as proposition 4, and its effective dimension is bounded by $\sum_j n_j$. The weights β_j can be tuned on the offline library, so the combined prior is a strict generalization of the single-rank prior of section V (recovered at L=1). This mirrors the combined spectral-truncation model, whose product of kernels with different truncation parameters captures multiple interaction scales at once [33]; learning the weights end to end is left to future work (section X).

VII. APPLICATIONS: EXTRACTING LOCAL AND GLOBAL GRAPH STRUCTURE

There are many potential applications of the operator-spectral truncated prior. The construction is useful precisely when the angle objective depends on both *local* graph structure (degrees, neighbourhoods) and *global* graph structure (spectral gap, connectivity), because the noncommuting generators $\mathcal{A}_G$ (global spectrum) and $\mathcal{B}_G$ (local degree) encode the two jointly. We list three examples; these are motivating use cases for the construction, not experiments reported here (section X).

a. Pathway and module selection. Selecting a maximally informative subset of a biological pathway or co-expression module is a graph-partition objective in which both the local degree of a node and the global community structure of the network matter. A purely diagonal (commutative) prior captures only the per-coordinate effect; the proposed prior can emphasize the collective directions that couple local hubs to global modules.

b. Patient-similarity partitioning. Partitioning a patient-similarity graph for stratification depends on local neighbourhood density and on the global spectral geometry of the cohort. The noncommuting operator lets the angle search trade these off rather than treating each angle coordinate independently.

c. Drug-combination and interaction pruning. Screening drug combinations or pruning molecular-interaction graphs balances local pairwise interactions against global network reachability. As with operator-learning tasks for spectral-truncation kernels, the proposed prior extracts both by ranking the collective angle directions; a numerical study on weighted biomedical graphs is the natural next experiment.

VIII. NUMERICAL RESULTS

The reference operating point is $p = 3$, $Q = 24$, truncation rank $r = 4$, commutator weight $\omega = 4.0$, and seed 260424803. The training library contains six graphs per family from {Erdős–Rényi [49], random-regular, Watts–Strogatz [50], Barabási–Albert [51]} with $n \in \{8, 10\}$, generated and manipulated with `NetworkX` [52] on top of the `NumPy/SciPy` numerical stack [53, 54]; the test set contains four held-out graphs per family, for 16 paired instances per seed. All objectives are exact statevector expectations and all MaxCut normalizers are exact brute-force optima.

To test robustness to the random seed—which fixes the library/test split and all stochastic baseline draws—we repeat the entire matched-budget benchmark across ten seeds, giving 160 paired held-out instances (section VIII B). The truncation sweep, ablation, and per-family breakdown are reported at the reference operating point. Where a baseline performs local refinement, it stands in for the gradient-free classical optimizers—coordinate descent, Nelder–Mead simplex [55], COBYLA [56], and SPSA [57]—commonly paired with QAOA, and for Bayesian-optimization angle search [58]. The baselines are: (i) **Random**, uniform random angles under the same query budget; (ii) TQA, one fixed trotterized quantum annealing schedule [27]; (iii) TQA+**coordinate**, a TQA anchor followed by isotropic coordinate refinement (the paired baseline); (iv) **kNN+coordinate**, a feature-neighbour posterior mean with diagonal coordinate refinement; (v) **OST diagonal**, the same operator posterior as OST-QAOA but with off-diagonal covariance removed (the commutative special case of proposition 1); and (vi) OST-QAOA, the full operator covariance with eigendirection search.

A. Experiment with synthetic graph families

OST-QAOA is the strongest method at the reference operating point (table II), reaching 0.820 ± 0.019 mean ratio, +0.112 over TQA+coordinate with a paired bootstrap interval $[+0.083, +0.142]$, winning on all 16 paired test graphs (sign-test $p < 10^{-4}$). The diagonal operator variant is also strong at 0.810 ± 0.021, but the full covariance improves the mean by 0.010, showing that the off-diagonal operator geometry is measurable in this configuration. Beyond the final ratio, the operator prior is markedly more query-efficient: it reaches 98% of the best observed ratio in a mean of 10.1 objective queries against 25.0 for TQA+coordinate and 13.1 for the diagonal variant ($\bar{Q}_{0.98}$)—the operationally relevant quantity of remark 1.

The truncation sweep (fig. 3) shows the interior optimum predicted by remark 1: rank $n^\star$=4 is best, aggressive truncation (n=1) underfits the angle geometry, and no truncation (n=2p=6) over-diffuses the covariance; the effective dimension rises monotonically from 2.50 at n=1 to 3.64 at n=6 (fig. 2), confirming proposition 3. The ablation and the sweep attribute the improvement: restricting the covariance to its diagonal—the commutative special case of proposition 1—drops the mean ratio by 0.010, so the off-diagonal, noncommuting search directions are operationally used. The *explicit* squared-commutator energy term is a distinct matter: in the sweep the commutator-on minus commutator-off gap Δ_{nc} is +0.003 at the operating point but oscillates in sign across ranks (e.g. −0.006 at n=2), placing it within run-to-run noise. The operationally relevant noncommutativity therefore already resides in the truncated $\mathcal{A}_G, \mathcal{B}_G$ geometry and the eigenbasis search (the diagonal restriction costs 0.010), not in the squared-commutator energy. This clean negative result sharpens, rather than weakens, the localization of the gain.

B. Experiment across seeds and graph families

Crucially, the advantage is not a single-seed artifact. Repeating the entire matched-budget benchmark across ten seeds (160 paired held-out instances), the paired gain of OST-QAOA over TQA+coordinate is $+0.107 \pm 0.011$, with a per-seed range of $[+0.087, +0.126]$—positive on every seed—and 157/160 instance wins; the diagonal variant averages 0.811 ± 0.009 across seeds, with the full covariance consistently better by 0.007–0.010. Across seeds OST-QAOA reaches the target in a mean of 10.5 queries against 24.7 for TQA+coordinate and 11.5 for the diagonal variant—a 2.4× reduction. Figure 4 shows that this ordering is stable: (a) the approximation-ratio ranking is preserved across seeds, with OST-QAOA and the diagonal variant on top; (b) the queries-to-target $\bar{Q}_{0.98}$ is likewise stable, the two operator methods reaching the target in roughly half the queries of the baselines; and (c) every graph family has a positive paired mean

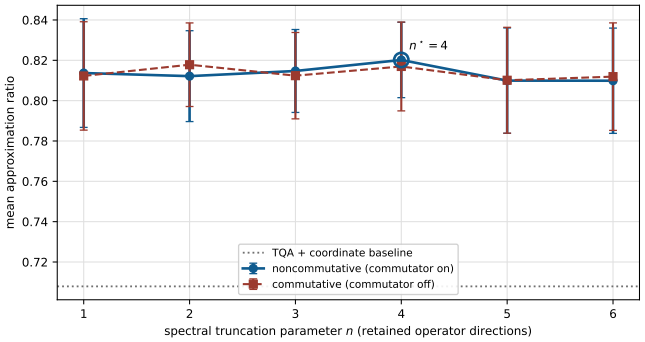

FIG. 3. Approximation ratio of the synthetic-graph angle-search task with different values of the truncation parameter n, the central experiment. Mean approximation ratio versus $n \in \{1, \ldots, 2p\}$, with the noncommuting operator (commutator on) and its commutator-off counterpart, against the TQA+coordinate baseline. The curve has an interior optimum $n^\star$=4: aggressive truncation (n=1) and no truncation (n=2p) are both worse, the query-budget analogue of the representation–complexity tradeoff of spectral-truncation kernels [33].

improvement, largest on Barabási–Albert (+0.190) and smallest on random-regular (+0.035), where the TQA baseline is already strong. Figure 5 resolves this per instance: across all 16 held-out graphs the operator methods (bottom rows) attain visibly higher ratios than the fixed-schedule and diagonal baselines (top rows).

C. Experiment on query trajectories

Finally we examine the realised search trajectory query by query, the operator-search analogue of comparing a predicted solution against the true one. Figure 6 plots the mean best-so-far approximation ratio against the objective-query budget: the operator anchors lift the incumbent early, and the eigendirection refinement preserves and widens the gap as the budget is spent, approaching the exact brute-force optimum ($r = 1$, the true solution). Figure 7 reports the complementary pointwise optimality gap $1 - r_G(\theta)$: for OST-QAOA it falls fastest and lowest, while the fixed-schedule and random policies retain a large residual gap throughout. The gain is therefore not only a final-query artifact: the operator mean and nearest-neighbour anchor improve early incumbent quality, while covariance eigendirections preserve gains as the budget is spent.

IX. GENERALIZATION OF THE OPERATOR-SPECTRAL TRUNCATED PRIOR

The proposed prior has potential possibilities of construction for more general situations.

TABLE II. Approximation ratio and query cost of the angle-search task with OST-QAOA and the baselines at the reference operating point ($p = 3$, $Q = 24$, rank 4, seed 260424803). Values are mean approximation ratio with 95% confidence intervals across 16 held-out graphs. Δ is the paired difference against TQA+coordinate with a bootstrap 95% interval; $\bar{Q}_{0.98}$ is the mean number of objective queries to reach 98% of the best observed ratio (remark 1).

Method	Mean ratio	Δ vs. TQA+coord. [95% CI]	Wins	Sign p	$\bar{Q}_{0.98}$
Random	0.743 ± 0.028	+0.035 [+0.001, +0.070]	11/16	0.2101	24.1
TQA	0.614 ± 0.030	−0.093 [-0.156, -0.040]	0/16	$< 10^{-4}$	25.0
TQA+coordinate	0.708 ± 0.038	0.000	0/0	1.0000	25.0
kNN+coordinate	0.751 ± 0.040	+0.043 [+0.019, +0.068]	13/16	0.0213	23.6
OST diagonal	0.810 ± 0.021	+0.102 [+0.073, +0.131]	16/16	$< 10^{-4}$	13.1
OST-QAOA	0.820 ± 0.019	+0.112 [+0.083, +0.142]	16/16	$< 10^{-4}$	10.1

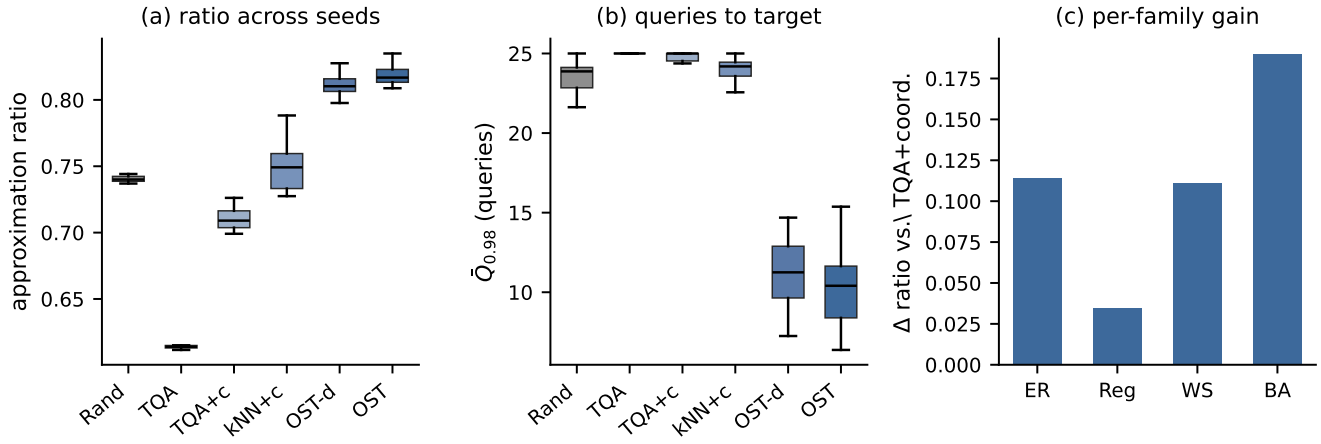

FIG. 4. Robustness of the angle-search task across seeds and families. (a) Approximation ratio of each method across ten independent seeds. (b) Queries to reach 98% of the best observed ratio, $\bar{Q}_{0.98}$, across seeds. (c) Per-family paired improvement Δ of OST-QAOA over the TQA+coordinate baseline; every family is positive.

A. Generalization to deeper circuits

The construction generalizes from p to larger depth in a straightforward manner: at depth p the angle space is $[0, \pi]^{2p}$ and the operators $\mathcal{A}_G, \mathcal{B}_G, \mathcal{O}_G$ are $2p \times 2p$, so definition 1 and propositions 2 to 4 hold verbatim with $2p$ in place of the ambient dimension. The query-budget heuristic of remark 1 then governs how the effective dimension—not the ambient $2p$—sets the query cost as depth grows.

B. Generalization to weighted and constrained graph objectives

The unweighted MaxCut Hamiltonian $H_C(G)$ can be replaced by a weighted or constrained cost Hamiltonian by using the weighted graph Laplacian in $\mathcal{A}_G$ and weighted degree moments in $\mathcal{B}_G$; the operator map and its truncation are unchanged. This is the route to the weighted biomedical graphs of section VII.

C. Generalization to multivariate and multi-objective search

Several objectives or several graphs can be handled jointly by stacking their angle operators into a block-diagonal operator on the product angle space and truncating the stack; the same participation-ratio bound (definition 2) applies to the joint covariance, so the query policy ranks directions across objectives at once.

D. Generalization to other operator constructions

The generators $\mathcal{A}_G, \mathcal{B}_G$ need not be the Laplacian-and-degree pair used here. Any pair of graph-derived symmetric operators that do not commute—for example normalized adjacency, heat-kernel, or PageRank operators—can replace them in eq. (6) while keeping the truncation machinery and all of sections III and IV intact.

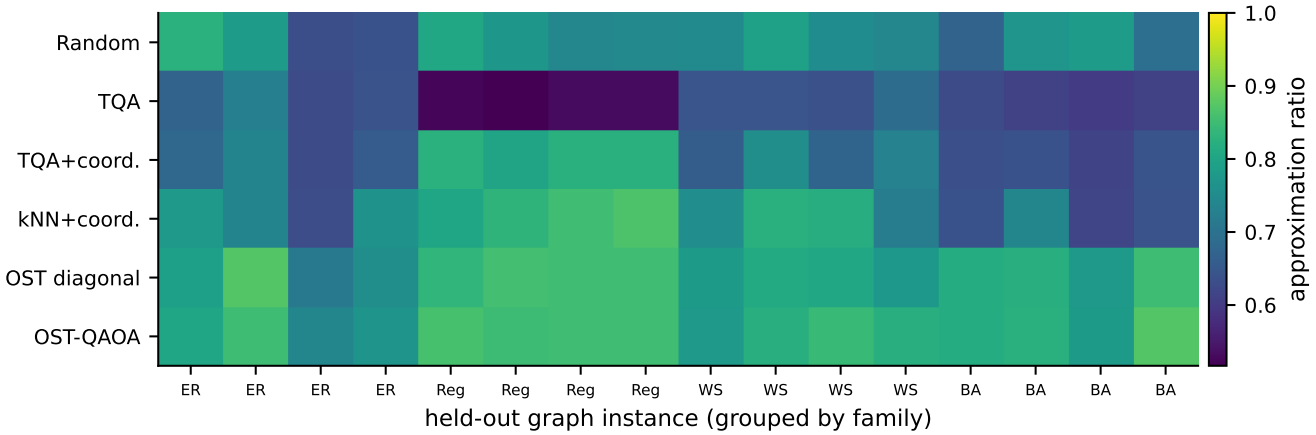

FIG. 5. Per-instance output of the angle-search task with different methods. Approximation ratio of every method (rows) on each of the 16 held-out graphs (columns, grouped by family). The operator methods attain consistently higher ratios than the fixed-schedule and diagonal baselines.

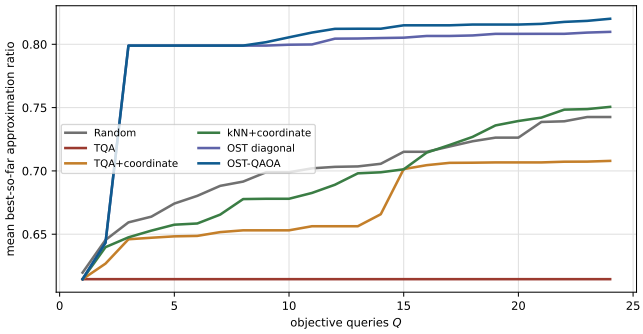

FIG. 6. Realised search trajectory of the angle-search task. Mean best-so-far approximation ratio versus the objective-query budget Q for OST-QAOA and the baselines (the predicted trajectory), approaching the exact brute-force optimum $r = 1$ (the true solution). Operator anchors improve the early budget and eigendirection refinement widens the gap thereafter.

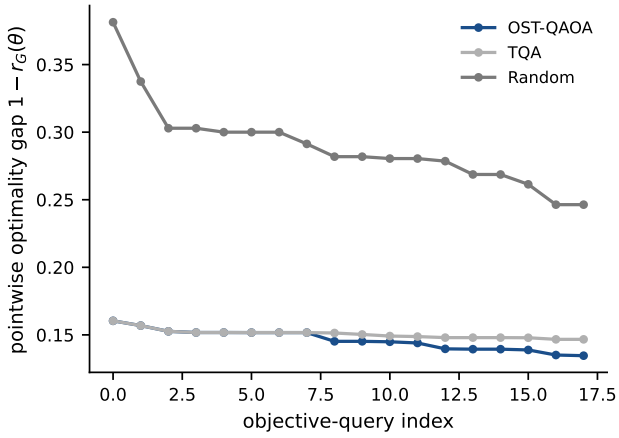

FIG. 7. Pointwise optimality gap of the angle-search task. Pointwise gap $1 - r_G(\theta)$ of the incumbent angle versus the objective-query index for OST-QAOA, TQA, and Random. The proposed policy drives the gap down fastest and to the smallest residual.

E. Generalization using other graph descriptors

The signature descriptor $\phi(G)$ that feeds $\mathcal{A}_G, \mathcal{B}_G$ can be replaced by other graph descriptors—graph-neural-network embeddings or Weisfeiler–Lehman features—without changing the prior construction, the query policy, or the heuristic; only the offline library signatures change.

X. CONCLUSION AND DISCUSSION

OST-QAOA replaces a diagonal uncertainty heuristic with a rank-controlled, noncommuting operator prior on QAOA angle space. The method uses the off-diagonal operator geometry to choose collective search directions, ships as an installable Python package, and regenerates every manuscript result from tunable parameters. In the exact-statevector benchmark the operator-spectral policy outperforms budget-matched TQA coordinate refinement by $+0.107 \pm 0.011$ across ten seeds (positive on every seed; 157/160 instance wins) and reaches comparable quality in roughly 2.4× fewer queries, while the truncation sweep exhibits the interior optimum predicted by the query-budget heuristic (remark 1) and the ablation localizes the advantage to the spectral-truncation parameter and the off-diagonal operator directions rather than to the squared-commutator energy.

a. Relationship to spectral-truncation kernels. Hashimoto *et al.* [33] introduce positive-definite,

noncommutative C^*-algebraic kernels whose truncation parameter trades representation power against model complexity, validated on supervised vector- and function-valued regression. OST-QAOA takes the same design idea—finite-rank noncommutativity as a tractable way to encode interaction—into a different setting, in three respects. *(i) Domain.* The truncated noncommuting operator is not a regression kernel but a graph-conditioned prior over variational-quantum parameters, converted into a deterministic, auditable objective-query policy. *(ii) An operational reading of the tradeoff.* Where their tradeoff is statistical (a generalization bound), ours is operational: remark 1 ties the truncation parameter to the *number of objective queries* through $d_{\text{eff}}(n)$, the quantity that limits near-term QAOA, and the predicted interior optimum $n^\star$ is observed (fig. 3). We are careful to label this a heuristic under assumption 1, not a transported theorem. *(iii) A tested mechanism.* The diagonal (commutative) restriction of proposition 1 is measurably worse than the full noncommuting geometry, so the value of operating off the coordinate axes is exhibited, not assumed.

b. Relationship to QAOA parameter transfer. A complementary line of work transfers good angles across instances, exploiting objective and parameter concentration [19, 21] and graph-to-graph transferability [22, 23]. OST-QAOA can be read as supplying what transfer alone does not: a budgeted *search geometry* around the transferred point. The TQA and kNN baselines here are exactly transfer-style warm starts; the operator covariance adds the directions and scales along which the remaining queries are spent, which is where the measured advantage comes from.

c. Limitations. The boundaries of these claims are as follows. *(i) Scale.* Experiments use exact statevectors with $n \leq 10$ and brute-force MaxCut optima, chosen so every number regenerates on a CPU; we do not claim hardware-scale behavior, and the operator dimension $2p$ is small at p=3. Quantum advantage for MaxCut may itself require hundreds to thousands of qubits [59], and at scale trainability is further constrained by barren plateaus [60] and by structural obstacles such as symmetry protection [61], reachability deficits [62], and the competitiveness of recursive and classical relaxations [63]; empirical small-graph QAOA performance baselines [64] bound what is achievable in this regime. *(ii) Problem class.* All instances are unweighted MaxCut on four synthetic graph families; weighted, biomedical, or non-MaxCut objectives are future work, not evidence presented here. *(iii) Query model.* One query is one exact expectation; finite-shot estimation and hardware noise—where query efficiency matters most—are deliberately excluded and remain to be tested. *(iv) Mechanism.* The explicit squared-commutator term is near-neutral at the studied operating point (fig. 3); the demonstrated value of noncommutativity is in the truncated-operator eigenbasis search, not in that term. *(v) Theory.* The query-budget statement (remark 1) is a heuristic under assumption 1, validated empirically, not a proved bound. The natural next steps are weighted biomedical graphs derived from transcriptomic, drug-target, and pathway data, larger n via tensor-network or sampled estimation, finite-shot/hardware-noise robustness, and learning the combined-model weights of section VI end to end.

ACKNOWLEDGMENTS

The author thanks the open-source scientific Python community. No external funding was used for this work. All code required to regenerate the figures, tables, CSV files, and query traces is in `submission/code`: the package command `uqqaoa-reproduce` writes the benchmark, sweep, and trace outputs, `multiseed_robustness.py` writes the per-seed robustness CSV, and `make_jmlr_mirror_artifacts.py` writes the four added display items and Table I from those outputs.

[1] R. M. Karp, in *Complexity of Computer Computations*, edited by R. E. Miller and J. W. Thatcher (Plenum Press, New York, 1972) pp. 85–103.
[2] A. Lucas, Front. Phys. **2**, 5 (2014).
[3] M. X. Goemans and D. P. Williamson, J. ACM **42**, 1115 (1995).
[4] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm (2014), arXiv:1411.4028 [quant-ph].
[5] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm applied to a bounded occurrence constraint problem (2014), arXiv:1412.6062 [quant-ph].
[6] G. E. Crooks, Performance of the quantum approximate optimization algorithm on the maximum cut problem (2018), arXiv:1811.08419 [quant-ph].
[7] J. Wurtz and P. J. Love, Phys. Rev. A **103**, 042612 (2021).
[8] J. Wurtz and D. Lykov, Phys. Rev. A **104**, 052419 (2021).
[9] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Nat. Rev. Phys. **3**, 625 (2021).
[10] S. Hadfield, Z. Wang, B. O'Gorman, E. G. Rieffel, D. Venturelli, and R. Biswas, Algorithms **12**, 34 (2019).
[11] Z. Wang, S. Hadfield, Z. Jiang, and E. G. Rieffel, Phys. Rev. A **97**, 022304 (2018).
[12] S. Lloyd, Quantum approximate optimization is computationally universal (2018), arXiv:1812.11075 [quant-ph].
[13] K. Blekos, D. Brand, A. Ceschini, C.-H. Chou, R.-H. Li, K. Pandya, and A. Summer, Phys. Rep. **1068**, 1 (2024).
[14] J. Preskill, Quantum **2**, 79 (2018).
[15] L. Bittel and M. Kliesch, Phys. Rev. Lett. **127**, 120502 (2021).

[16] W. Lavrijsen, A. Tudor, J. Müller, C. Iancu, and W. de Jong, in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, 2020) pp. 267–277.

[17] A. Pellow-Jarman, I. Sinayskiy, A. Pillay, and F. Petruccione, Quantum Inf. Process. **20**, 202 (2021).

[18] J. Larkin, M. Jonsson, D. Justice, and G. G. Guerreschi, Quantum Sci. Technol. **7**, 045014 (2022).

[19] F. G. S. L. Brandão, M. Broughton, E. Farhi, S. Gutmann, and H. Neven, For fixed control parameters the quantum approximate optimization algorithm's objective function value concentrates for typical instances (2018), arXiv:1812.04170 [quant-ph].

[20] L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, Phys. Rev. X **10**, 021067 (2020).

[21] V. Akshay, D. Rabinovich, E. Campos, and J. Biamonte, Phys. Rev. A **104**, L010401 (2021).

[22] A. Galda, X. Liu, D. Lykov, Y. Alexeev, and I. Safro, in *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, 2021) pp. 171–180, arXiv:2106.07531 [quant-ph].

[23] R. Shaydulin, P. C. Lotshaw, J. Larson, J. Ostrowski, and T. S. Humble, ACM Trans. Quantum Comput. **4**, 1 (2023).

[24] R. Shaydulin, C. Li, S. Chakrabarti, M. DeCross, D. Herman, N. Kumar, J. Larson, D. Lykov, P. Minssen, Y. Sun, Y. Alexeev, J. M. Dreiling, J. P. Gaebler, T. M. Gatterman, J. A. Gerber, K. Gilmore, D. Gresh, N. Hewitt, C. V. Horst, S. Hu, J. Johansen, M. Matheny, T. Mengle, M. Mills, S. A. Moses, B. Neyenhuis, P. Siegfried, R. Yalovetzky, and M. Pistoia, Sci. Adv. **10**, eadm6761 (2024).

[25] M. P. Harrigan, K. J. Sung, M. Neeley, K. J. Satzinger, F. Arute, K. Arya, J. Atalaya, J. C. Bardin, R. Barends, S. Boixo, *et al.*, Nat. Phys. **17**, 332 (2021).

[26] G. Pagano, A. Bapat, P. Becker, K. S. Collins, A. De, P. W. Hess, H. B. Kaplan, A. Kyprianidis, W. L. Tan, C. Baldwin, L. T. Brady, A. Deshpande, F. Liu, S. Jordan, A. V. Gorshkov, and C. Monroe, Proc. Natl. Acad. Sci. U.S.A. **117**, 25396 (2020).

[27] S. H. Sack and M. Serbyn, Quantum **5**, 491 (2021).

[28] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, Science **292**, 472 (2001).

[29] T. Kadowaki and H. Nishimori, Phys. Rev. E **58**, 5355 (1998).

[30] D. J. Egger, J. Mareček, and S. Woerner, Quantum **5**, 479 (2021).

[31] N. Jain, B. Coyle, E. Kashefi, and N. Kumar, Quantum **6**, 861 (2022).

[32] S. Khairy, R. Shaydulin, L. Cincio, Y. Alexeev, and P. Balaprakash, in *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34 (2020) pp. 2367–2375.

[33] Y. Hashimoto, A. Hafid, M. Ikeda, and H. Kadri, Spectral truncation kernels: Noncommutativity in c^*-algebraic kernel machines (2024), revised 2026, arXiv:2405.17823 [stat.ML].

[34] M. Schuld and N. Killoran, Phys. Rev. Lett. **122**, 040504 (2019).

[35] Y. Bengio, A. Lodi, and A. Prouvost, Eur. J. Oper. Res. **290**, 405 (2021).

[36] E. Farhi and A. W. Harrow, Quantum supremacy through the quantum approximate optimization algorithm (2016), arXiv:1602.07674 [quant-ph].

[37] M. Huynh, Query-efficient quantum approximate optimization via graph-conditioned trust regions (2026), arXiv:2604.24803, arXiv:2604.24803 [quant-ph].

[38] M. Huynh, Graph-conditioned trust regions for uncertainty-calibrated quantum approximate optimization (2026), manuscript in preparation.

[39] M. Huynh, Graph-conditioned trust regions for uncertainty-calibrated quantum approximate optimization (extended) (2026), manuscript in preparation.

[40] M. Huynh, The measurement cost of warm-started low-depth qaoa (2026), manuscript in preparation.

[41] M. Huynh, Certified query budgets for the quantum approximate optimization algorithm (2026), manuscript in preparation.

[42] M. Huynh, Topology-conditioned qaoa parameter transfer for budgeted graph optimization (2026), manuscript in preparation.

[43] M. Huynh, Topology-conditioned qaoa parameter transfer for budgeted graph optimization (companion) (2026), manuscript in preparation.

[44] F. R. K. Chung, *Spectral Graph Theory*, CBMS Regional Conference Series in Mathematics, Vol. 92 (American Mathematical Society, Providence, RI, 1997).

[45] M. Fiedler, Czechoslovak Math. J. **23**, 298 (1973).

[46] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning* (MIT Press, Cambridge, MA, 2006).

[47] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas, Proc. IEEE **104**, 148 (2016).

[48] K. J. Sung, J. Yao, M. P. Harrigan, N. C. Rubin, Z. Jiang, L. Lin, R. Babbush, and J. R. McClean, Quantum Sci. Technol. **5**, 044008 (2020).

[49] P. Erdős and A. Rényi, Publ. Math. Debrecen **6**, 290 (1959).

[50] D. J. Watts and S. H. Strogatz, Nature **393**, 440 (1998).

[51] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).

[52] A. A. Hagberg, D. A. Schult, and P. J. Swart, in *Proceedings of the 7th Python in Science Conference (SciPy 2008)*, edited by G. Varoquaux, T. Vaught, and J. Millman (2008) pp. 11–15.

[53] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant, Nature **585**, 357 (2020).

[54] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, *et al.*, Nat. Methods **17**, 261 (2020).

[55] J. A. Nelder and R. Mead, Comput. J. **7**, 308 (1965).

[56] M. J. D. Powell, in *Advances in Optimization and Numerical Analysis*, edited by S. Gomez and J.-P. Hennart (Springer Netherlands, Dordrecht, 1994) pp. 51–67.

[57] J. C. Spall, IEEE Trans. Autom. Control **37**, 332 (1992).

[58] S. Tibaldi, D. Vodola, E. Tignone, and E. Ercolessi, IEEE Trans. Quantum Eng. **4**, 1 (2023).

[59] G. G. Guerreschi and A. Y. Matsuura, Sci. Rep. **9**, 6903 (2019).

[60] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Nat. Commun. **9**, 4812 (2018).

[61] S. Bravyi, A. Kliesch, R. Koenig, and E. Tang, Phys. Rev. Lett. **125**, 260505 (2020).

[62] V. Akshay, H. Philathong, M. E. S. Morales, and J. D. Biamonte, Phys. Rev. Lett. **124**, 090504 (2020).
[63] S. Bravyi, A. Kliesch, R. Koenig, and E. Tang, Quantum **6**, 678 (2022).
[64] P. C. Lotshaw, T. S. Humble, R. Herrman, J. Ostrowski, and G. Siopsis, Quantum Inf. Process. **20**, 403 (2021).