

Online Resource 1: Supplementary Information for “Feasible-Space Dilution and Objective Alignment in Shallow Penalty-X QAOA: A Controlled RCSP Study”

Zhilin Chen^{1*}

^{1*}University of Copenhagen, Copenhagen, Denmark.

Corresponding author(s). E-mail(s): zwq152@alumni.ku.dk;

Abstract

This Online Resource contains supplementary methods, mathematical proofs, protocol and robustness details, additional figures and tables, and reproducibility information for the accompanying Original Research article.

Contents

S1	Mathematical Results and Access-Model Boundaries	2
S2	Task Construction	7
S3	Hamiltonian and Optimization Details	8
S4	Optimizer Attribution	9
S5	Objective-Alignment Details	10
S6	Held-Out Preregistration and Statistics	11
S7	Post-Hoc Finite-Shot Endpoint Robustness	12
S8	Reviewer-Robustness Experiments	14
S9	Scaling Analysis	21
S10	Reproducibility	22
S11	Extended Figures and Comparison Tables	23

S1 Mathematical Results and Access-Model Boundaries

The search-theoretic core uses established prior art. The membership-query results are average-case and access-model-specific; they do not lower-bound the rich-energy RCSP/CVaR experiment. Required assumptions are collected in [table S1](#).

S1.1 Fixed-query phase-sensitive dilution

Let $\Omega = \{1, \dots, N\}$ and let $\mathcal{F}$ be uniform over the size- M subsets of Ω , with $\phi = M/N$. Define

$$\Pi_{\mathcal{F}} = \sum_{x \in \mathcal{F}} |x\rangle\langle x|, \quad Q_{\mathcal{F}}(\gamma) = I + (e^{-i\gamma} - 1)(\Pi_{\mathcal{F}} \otimes I_A), \quad (\text{S1})$$

where A is an arbitrary ancilla register.

Theorem S1 (Phase-sensitive average bound) *Consider a normalized pure-state query algorithm*

$$|\psi_{\mathcal{F}}\rangle = V_q Q_{\mathcal{F}}(\gamma_q) \cdots V_1 Q_{\mathcal{F}}(\gamma_1) V_0 |\psi_{\text{in}}\rangle, \quad (\text{S2})$$

where the input, V_t , and γ_t are independent of the particular $\mathcal{F}$. If

$$P_{\mathcal{F}} = \langle \psi_{\mathcal{F}} | (\Pi_{\mathcal{F}} \otimes I_A) | \psi_{\mathcal{F}} \rangle, \quad (\text{S3})$$

then

$$\mathbb{E}_{\mathcal{F}} P_{\mathcal{F}} \leq \min \left\{ 1, \left(1 + \sum_{t=1}^q |e^{-i\gamma_t} - 1| \right)^2 \phi \right\}. \quad (\text{S4})$$

Proof Replace every query in [equation \(S2\)](#) by the identity and let $|\varphi_t\rangle$ be the normalized reference state immediately before slot t . It is independent of $\mathcal{F}$. Write $|\varphi_t\rangle = \sum_x |x\rangle |\alpha_{t,x}\rangle$. Then

$$\mathbb{E}_{\mathcal{F}} \|(\Pi_{\mathcal{F}} \otimes I_A) |\varphi_t\rangle\|^2 = \sum_x \Pr[x \in \mathcal{F}] \|\alpha_{t,x}\|^2 \quad (\text{S5})$$

$$= \frac{M}{N} \sum_x \|\alpha_{t,x}\|^2 = \phi. \quad (\text{S6})$$

Let $c_t = |e^{-i\gamma_t} - 1|$ and let $d_t(\mathcal{F})$ be the Euclidean distance between actual and reference states after slot t and the following $\mathcal{F}$ -independent unitary. Inserting and subtracting the query applied to the reference state, and using unitarity, gives

$$d_t(\mathcal{F}) \leq d_{t-1}(\mathcal{F}) + c_t \|(\Pi_{\mathcal{F}} \otimes I_A) |\varphi_t\rangle\|. \quad (\text{S7})$$

The $L^2(\mathcal{F})$ triangle inequality (Minkowski) and [equation \(S6\)](#) therefore yield

$$\|d_q\|_{L^2(\mathcal{F})} \leq \sqrt{\phi} \sum_{t=1}^q c_t. \quad (\text{S8})$$

The final identity-reference state is also independent of $\mathcal{F}$, so its average marked mass is ϕ . Projection contractivity gives pointwise

$$\sqrt{P_{\mathcal{F}}} \leq \|(\Pi_{\mathcal{F}} \otimes I_A) |\varphi_{q+1}\rangle\| + d_q(\mathcal{F}). \quad (\text{S9})$$

A second L^2 triangle inequality gives $\sqrt{\mathbb{E} P_{\mathcal{F}}} \leq (1 + \sum_t c_t) \sqrt{\phi}$. Squaring the nonnegative sides and applying the probability cap proves [equation \(S4\)](#). $\square$

The cases $M = 0$ and $M = N$ follow directly; the probability cap applies at all q .

S1.2 Adaptive hard cap, training, and expected queries

For the phase-flip membership oracle $O_{\mathcal{F}} = I - 2\Pi_{\mathcal{F}}$, $|e^{-i\pi} - 1| = 2$, so [theorem S1](#) immediately gives the coarse fixed-schedule form. The following explains the adaptive extension.

Proof of the membership-only dilution theorem in the main article Introduce registers for the query label, work space, classical transcript, random seed, discarded environment, halt flag, scratch query, and output. Replace each measurement/instrument by a Stinespring isometry and retain its outcome in an orthogonal transcript register. Purify classical randomness and make feed-forward coherently controlled on the transcript. Mixed initial states are covered by a purifying environment.

Pad every branch to q query slots. Once a branch halts, coherently swap an isolated scratch label into the query port, apply the ordinary oracle, and swap back. The scratch data never recouple to the output. Finally, reversibly copy the selected classical output label into a designated query register. The result is a pure q -slot algorithm with $\mathcal{F}$ -independent interquery unitaries.

For each slot, $\|(O_{\mathcal{F}} - I)|\varphi_t\rangle\| = 2\|(\Pi_{\mathcal{F}} \otimes I)|\varphi_t\rangle\|$. The identity-reference calculation in [equation \(S6\)](#) and the same two Minkowski steps give

$$\sqrt{\mathbb{E}P_{\text{success}}} \leq (2q + 1)\sqrt{\phi}. \quad (\text{S10})$$

Squaring and capping at one proves the theorem. Orthogonal histories also cover early stopping and feed-forward. A history-controlled direct sum counts as one query slot only under this explicit controlled membership interface. $\square$

The bit-query oracle $B_{\mathcal{F}}|x, b\rangle = |x, b \oplus \mathbf{1}[x \in \mathcal{F}]\rangle$ obeys $B_{\mathcal{F}} - I = \Pi_{\mathcal{F}} \otimes (X - I)$ and $\|X - I\| = 2$, so the same coarse theorem holds. If slot t instead uses a history-controlled phase $\gamma_t(h)$, the phase-sensitive result holds with $c_t = \sup_h |e^{-i\gamma_t(h)} - 1|$ under the corresponding counted direct-sum interface.

If a training procedure satisfies the same membership-only premise, regard all training circuits, sampled outputs, optimizer transcripts, feedback, and the final circuit as one adaptive interaction. A deterministic pathwise cap

$$q_{\text{total}} = q_{\text{training}} + q_{\text{sampling}} + q_{\text{feedback}} + q_{\text{final}} \quad (\text{S11})$$

is substituted for q . Classical storage of a query-generated transcript does not erase its calls. Free explicit coefficients or rich cost values would instead define a stronger access model.

S1.2.1 Expected-query truncation

Let Q be a nonnegative integer-valued number of membership calls and let $\bar{q} = \mathbb{E}Q$, with the expectation covering the subset prior and internal randomness. For any deterministic integer $T \geq 0$, convert executions with $Q > T$ into failures. The retained protocol has hard cap T , and Markov's inequality on the integer event gives

$$\Pr(Q > T) = \Pr(Q \geq T + 1) \leq \frac{\bar{q}}{T + 1}. \quad (\text{S12})$$

Therefore

$$\mathbb{E}P \leq \min \left\{ 1, \frac{\bar{q}}{T + 1} + \min\{1, (2T + 1)^2 \phi\} \right\}. \quad (\text{S13})$$

The substitution $q = \mathbb{E}Q$ in a hard-cap bound is invalid. If Q is sampled independently before the interaction, conditioning additionally yields a bound in terms of $\mathbb{E}(2Q+1)^2$; that statement is not transferred to transcript-dependent stopping.

S1.3 Raw-density counterexample and representation padding

Proof of the raw-density counterexample in the main article Let $s = v_0 \rightarrow v_1 \rightarrow \dots \rightarrow v_m = t$, assign objective cost and resource one to every edge, and set budget m . The only source–target path uses all m edges and is feasible and optimal. Exactly one of 2^m edge selections encodes it, so $\phi_{\text{state}} = 2^{-m}$. Starting from s , an adjacency-list algorithm follows the unique outgoing edge at each internal vertex, checks the additive totals, and writes the m edge identifiers in $O(m)$ time.

If constants $a, c > 0$ satisfied both the algorithmic upper bound $T(I_m) \leq am + a$ and a purported universal lower bound $T(I_m) \geq c\phi_{\text{state}}^{-1/2} = c2^{m/2}$, then $(am + a)/2^{m/2} \geq c$ for all m . The left side tends to zero, a contradiction for sufficiently large m . $\square$

Lemma S2 (Edge-subdivision representation padding) *Let I be an explicit directed RCSP instance with additive rational objective coefficients and nonnegative rational resource vectors. Replace one edge $e = (u, v)$ by a private serial path of $r \geq 1$ edges, and assign each segment $1/r$ of every coefficient of e . Then routes are in bijection, feasibility and objective ordering are preserved, and*

$$\phi_{\text{state}}(I^{(r)}) = 2^{-(r-1)}\phi_{\text{state}}(I). \quad (\text{S14})$$

Proof Every new internal vertex has indegree and outdegree one and is incident to no other edge. A path avoiding e is unchanged. A path using e expands to all r segments, while any path entering the serial chain must traverse it and contracts uniquely to e . These maps are mutual inverses.

For a route using e , the new contribution is $\sum_{k=1}^r c_e/r = c_e$ and likewise $\sum_{k=1}^r a_{e,j}/r = a_{e,j}$ for every resource j . Hence paired route totals, feasibility, objective comparisons, and tied optima are identical. The number of valid feasible route bitstrings is unchanged, while the raw edge-bit denominator gains $r - 1$ binary variables and hence a factor 2^{r-1} .

If a rational coefficient is represented in lowest terms as p/q , the unreduced segment coefficient is $p/(qr)$. Its binary encoding grows by at most $\lceil \log_2 r \rceil + O(1)$. The explicit topology still contains r segments and $r - 1$ vertices, so its listing and route-output overhead is $\Theta(r)$ rather than $O(\log r)$. $\square$

S1.4 Explicit-attribute-query RCSP

Proof of the explicit-attribute-query RCSP theorem in the main article Let the known topology contain K two-edge branches $P_i : s \rightarrow u_i \rightarrow t$. The first edge of each branch has resource one. The second has resource one when $z_i = 1$ and two when $z_i = 0$. With budget two,

$$P_i \text{ is feasible} \iff z_i = 1. \quad (\text{S15})$$

Equal fixed objective costs leak no marked index. One branch-feasibility check uses exactly one query to the attribute oracle defined in the main article, and conversely the feasibility bit is z_i . Finding any feasible route is thus multiple-marked search over K labels with M marks.

For a uniform size- M array, the membership-only dilution theorem gives

$$\mathbb{E}_z P_{\text{success}} \leq \min \left\{ 1, (2q + 1)^2 \frac{M}{K} \right\}. \quad (\text{S16})$$

Average target success τ requires $q \geq \frac{1}{2}(\sqrt{\tau K/M} - 1)$. A uniform worst-case guarantee implies the same average lower bound. In the nontrivial regime $M/K < \tau$, integrality converts this to $\Omega_\tau(\sqrt{K/M})$. A uniform superposition over branches plus amplitude amplification uses $O(\sqrt{K/M})$ attribute queries, giving the matching upper bound.

When $M/K \geq \tau$, a zero-query uniformly random branch already achieves the target, so the unqualified all- M expression would be false. For exact success, the order holds for $1 \leq M < K$, and $M = K$ needs zero queries. The input contains $\Theta(K)$ attribute cells and graph entries; no compact $O(\log K)$ input or RAM-time result is asserted. $\square$

S1.5 Posterior structure and finite classical advice

Condition on a supported advice value $S = s$ and define $\bar{\Pi}_s$, λ_s , and $\Lambda(S)$ as in the main article.

Lemma S3 (Posterior projector) *For every supported s ,*

$$\text{Tr } \bar{\Pi}_s = M, \quad \lambda_s = \max_x \Pr[x \in \mathcal{F} \mid S = s], \quad \phi \leq \lambda_s \leq 1. \quad (\text{S17})$$

Proof $\bar{\Pi}_s$ is diagonal in the label basis and its x th diagonal entry is the posterior inclusion probability. Its trace is the conditional expectation of $|\mathcal{F}| = M$. The maximum of N non-negative entries summing to M is at least M/N , and every inclusion probability is at most one. $\square$

Proof of the posterior structure theorem in the main article Fix s . Purify measurements, randomness, discarded data, feed-forward, and early stopping as above. Every identity-oracle reference state $|\psi_s\rangle$ may depend on s but is independent of the residual uncertainty in $\mathcal{F}$ conditional on s . Consequently,

$$\mathbb{E}[\langle \psi_s | (\Pi_{\mathcal{F}} \otimes I) | \psi_s \rangle \mid S = s] \quad (\text{S18})$$

$$= \langle \psi_s | (\bar{\Pi}_s \otimes I) | \psi_s \rangle \leq \lambda_s. \quad (\text{S19})$$

Repeating the two conditional $L^2(\mathcal{F} \mid s)$ Minkowski steps with per-slot phase coefficient $c_t(s)$ gives

$$\mathbb{E}[P \mid S = s] \leq \min \left\{ 1, \left(1 + \sum_{t=1}^q c_t(s) \right)^2 \lambda_s \right\}. \quad (\text{S20})$$

Averaging the already capped conditional expression is valid. Since $c_t(s) \leq 2$ and probabilities are at most one,

$$\mathbb{E}P \leq \mathbb{E}_S [\min\{1, (2q+1)^2 \lambda_S\}] \quad (\text{S21})$$

$$\leq \min\{1, (2q+1)^2 \Lambda(S)\}. \quad (\text{S22})$$

The cap is not moved through expectation in the phase-dependent statement; the displayed hard-cap form is its valid coarser consequence. $\square$

Lemma S4 (Finite advice alphabet) *If $|\text{supp } S| \leq 2^b$, then*

$$\Lambda(S) \leq \min\{1, 2^b \phi\}. \quad (\text{S23})$$

Proof For each supported s , choose $x_s \in \arg \max_x \Pr[x \in \mathcal{F} \mid S = s]$. Then

$$\Lambda(S) = \sum_s \Pr(S = s) \Pr(x_s \in \mathcal{F} \mid S = s) \quad (\text{S24})$$

$$= \sum_s \Pr(S = s, x_s \in \mathcal{F}) \quad (\text{S25})$$

$$\leq \sum_s \Pr(x_s \in \mathcal{F}) = |\text{supp } S| \phi \leq 2^b \phi. \quad (\text{S26})$$

Duplicate maximizing labels merely repeat a marginal term in an upper bound; they may loosen, but cannot invalidate, the inequality. The cap by one follows from $\lambda_s \leq 1$. $\square$

Combining the posterior structure theorem in the main article with [lemma S4](#) proves the finite-advice inequality stated there. If target success is $\tau > 0$, a necessary integer condition is

$$b \geq \max \left\{ 0, \left\lceil \log_2 \frac{\tau}{(2q+1)^2 \phi} \right\rceil \right\}. \quad (\text{S27})$$

For $b_{\text{eff}} = \log_2(\Lambda/\phi)$ and $K_{\text{eff}} = M/\Lambda$,

$$0 \leq b_{\text{eff}} \leq \log_2(1/\phi), \quad M \leq K_{\text{eff}} \leq N, \quad \frac{N}{K_{\text{eff}}} = 2^{b_{\text{eff}}}. \quad (\text{S28})$$

The target conditions

$$b_{\text{eff}} + 2 \log_2(2q+1) \geq \log_2(\tau/\phi), \quad K_{\text{eff}} \leq \frac{(2q+1)^2 M}{\tau} \quad (\text{S29})$$

are necessary only. A candidate set of size K gives $\lambda_s = M/K$ only when the conditional posterior is uniform over size- M subsets of that candidate set.

S1.6 One-shot finite-dimensional quantum advice

The following result concerns one pre-search advice state.

Proposition S5 (One-shot quantum-advice dimension bound) *Suppose one $\mathcal{F}$ -dependent density operator $\rho_{\mathcal{F}}$ on a total accessible advice register of Hilbert-space dimension d is supplied before an otherwise membership-only fixed-query interaction. Then*

$$\mathbb{E} P_{\text{success}} \leq \min \left\{ 1, \left(1 + \sum_t c_t \right)^2 d\phi \right\}, \quad (\text{S30})$$

and hence the coarse hard-cap factor is $(2q+1)^2 d\phi$.

Proof For an $\mathcal{F}$ -independent trace-preserving reference channel A acting before a query, $\rho_{\mathcal{F}} \preceq I_d$ implies

$$\mathbb{E}_{\mathcal{F}} \text{Tr}[(\Pi_{\mathcal{F}} \otimes I) A(\rho_{\mathcal{F}})] \leq \mathbb{E}_{\mathcal{F}} \text{Tr}[(\Pi_{\mathcal{F}} \otimes I) A(I_d)] \quad (\text{S31})$$

$$= \phi \text{Tr } A(I_d) = d\phi. \quad (\text{S32})$$

A canonical purification of each $\rho_{\mathcal{F}}$ preserves the marked projected norm; an inaccessible purifying reference does not enlarge the counted advice dimension. The phase-sensitive hybrid proof then replaces the reference overlap ϕ by $d\phi$. Multiple simultaneously supplied advice registers count by their joint accessible dimension. $\square$

Fresh copies, refreshed advice and an interactive $\mathcal{F}$ -dependent provider are outside [proposition S5](#).

Table S1 Assumptions that bound the interpretation of the formal results.

Result	Required assumptions
Fixed/adaptive dilution	Uniform fixed-cardinality prior; all remaining feasible-set dependence enters through counted membership queries; pathwise hard cap for adaptive protocols
Trained specialization	Training, sampling, feedback, and final membership calls are counted jointly
Expected-query form	Integer truncation with an explicit tail term; no substitution of $\mathbb{E}Q$ for a hard cap
Posterior structure	One classical pre-search channel; conditional reference circuit independent of residual feasible-set uncertainty
Classical advice	Advice support has cardinality at most 2^b ; b is not mutual information
Quantum advice	One accessible pre-search state of total dimension d ; no refreshing or interactive provider
Attribute-query RCSP	Explicit length- K attribute array accessed only through counted quantum random access

S2 Task Construction

S2.1 Layered graph generator

For a prescribed layer-width sequence $(1, w_1, \dots, w_L, 1)$, vertices receive stable integer identifiers in layer order. A deterministic coverage construction first adds adjacent-layer edges so every vertex has route-relevant incoming and outgoing connectivity. Remaining forward edges, including optional two-layer skips, are sampled without replacement until the target m is reached. Edges are sorted by endpoint before their stable indices, costs, and resources are assigned. A SHA-256-derived seed namespaces the master seed by size stratum and base index, avoiding dependence on language-level hash randomization.

The v2 strata use layer widths $(1, 2, 2, 1)$, $(1, 2, 2, 2, 1)$, $(1, 3, 3, 1)$, $(1, 2, 3, 3, 1)$, and $(1, 3, 3, 3, 1)$ at target edge counts 7, 10, 13, 16, and 19. Five base graphs are generated per stratum. Positive costs are sampled from $[1, 9]$ and resources from $[1, 1000]$ under the frozen master seed 20260827.

S2.2 Route and state enumeration

All simple source–target routes in a generated DAG are enumerated by a deterministic traversal. Each route record contains ordered edge identifiers, objective cost, and resource total. For budget B , the exact solver retains routes with resource total at most B and records all feasible routes tied at the minimum objective.

Separately, all 2^m binary edge selections are enumerated for the benchmark sizes. For each state, matrix multiplication computes objective cost, incidence residual, and resource total. A feasible state must have zero flow residual and meet the budget. This state-domain validator and the path-domain enumerator agree on the valid route encodings. Their denominators remain different: 2^m for ϕ_{state} and the number of candidate routes for ϕ_{path} .

S2.3 Duplicate audit and corrected stress schedule

The historical v1 construction requested seven ordinary quantiles at every base graph. Of 175 task rows, 53 repeated a feasible set. Ten base graphs had too few structural path-cardinality levels to supply seven distinct values; integer resource ties further reduced resolution. The audit retained these rows as historical evidence rather than modifying them.

For v2, let $R_1 < \dots < R_J$ be the distinct route-resource thresholds and let n_j be the number of routes with resource exactly R_j . The achievable feasible counts are cumulative sums $N_j = \sum_{k \leq j} n_k$. Preferred counts are reserved when exactly achievable. An unavailable preferred count u is mapped to the unused N_j minimizing $|\log(N_j/u)|$, with smaller cardinality as deterministic tie-breaker. Remaining capacity is filled by the unused count maximizing its minimum log-distance from selected counts. The sorted selected thresholds become D1, D2, and so forth; the construction emits at most seven levels and never repeats a feasible set.

The wider resource support yields 169 distinct route-resource totals across 169 route entries and zero resource-tie graphs. Ten graphs still have fewer than seven achievable sets and are labeled as insufficient for seven levels; their available distinct levels are retained without duplication. The final task counts and fractions are shown in [table S2](#).

Table S2 Distinct-cardinality task strata used to define the controlled universe.

Stratum	Edges	Graphs	Tasks	$\phi_{\min}$	$\phi_{\max}$
S1	7	5	15	0.00781	0.0234
S2	10	5	20	0.000977	0.00391
S3	13	5	35	0.000122	0.000854
S4	16	5	35	1.53e-05	0.000153
S5	19	5	35	1.91e-06	2.29e-05

S2.4 Construction scope

The controlled schedule deliberately increases within-graph resolution. It does not assert that the preferred cardinalities occur naturally, that all graphs have comparable combinatorial structure, or that D alone captures their difficulty. Task generation and selection never use a QAOA result.

S3 Hamiltonian and Optimization Details

S3.1 Statewise Hamiltonian evaluation

For edge-bit row vector x , edge-cost vector c , resource vector r , edge-vertex incidence matrix A , and demand vector d , the implementation computes

$$C(x) = xc, \quad f_{\text{raw}}(x) = \|xA - d\|_2^2, \quad u(x) = \max\{0, xr - B\}. \quad (\text{S33})$$

The scale-controlled functions and normalized energy then follow the controlled-energy definition in the main article. Arrays are evaluated in deterministic chunks; the diagonal energy vector has length 2^m and is not stored outside the active simulation run.

The controlled penalty contract was audited before QAOA. Both the historical raw and prospective controlled contracts returned an exact original RCSP optimum on 140/140 tasks. The controlled absolute span was 490.51–744.17, compared with 3.42698×10^7 – 1.82405×10^9 for the historical raw form. Normalization by 172 gives pilot spans 2.97815–4.32657. Ground-state correctness does not imply perfect phase-scale invariance; the remaining within-base relative range has median 6.8% in the selected pilot tasks.

S3.2 Exact statevector circuit

The initial vector has amplitude $2^{-m/2}$ on every basis state. A cost layer multiplies amplitude x by $e^{-i\gamma H_C(x)}$. Each factor $e^{-i\beta X_e}$ is applied by pairwise amplitude updates on bit e , so a full mixer implements $e^{-i\beta \sum_e X_e}$. Parameter arrays store all γ angles followed by all β angles. State norm is checked after simulation; no raw statevector is persisted.

All arms use the same H_C for the cost phase. O1, O2, and O3 do not replace the quantum Hamiltonian; they replace only the scalar classical loss computed from the terminal probability distribution. Exact weighted CVaR sorts the energy/probability pairs stably, accumulates probability until $\alpha = 0.10$, and uses only the fractional mass necessary at the cutoff.

S3.3 COBYLA accounting

COBYLA is called with initial radius 0.5 and constraint tolerance 10^{-8} . Pilot random starts sample γ in $[0, 2\pi]$ and β in $[0, \pi]$. These are initialization domains only: subsequent optimizer coordinates are unbounded, and the implementation applies neither a bound nor modulo- 2π wrapping to γ . Because the normalized Hamiltonian generally has a noninteger spectrum, no fundamental 2π periodicity in γ is assumed. The pilot uses 120 objective evaluations per independently initialized run, continuation uses 120 from the embedded point, and the matched objective and held-out $p = 3$ arms use 240. Diagnostic start and terminal evaluations remain separate from the optimizer’s counted calls.

For the pilot’s primary metric view, each metric is summarized by its median over three fixed seeds. A secondary multistart view selects the seed only by minimum terminal mean energy and subsequently reports that seed’s operational metrics. Feasibility and optimality never participate in seed selection. The objective-comparison arms share the selected $p = 2$ initialization.

S4 Optimizer Attribution

S4.1 Nested identity and certified failures

The exact $p = 2 \rightarrow 3$ embedding defined in the main article was evaluated for every one of the 56 tasks and three $p = 2$ seeds. Maximum statevector, probability, and objective differences were all zero at stored precision. The original $p = 3$ terminal energy exceeded the embedded point by more than 10^{-10} in 29 of 168 runs. The median signed gap over all runs was -0.199582 ; this overall median does not diminish the certification of the 29 positive-gap cases.

Continuation improved the mean objective for 29/29 certified failures and G_{feas} for 27/29. Across all 168 runs, continuation classifications were 151 “objective and feasibility improve” and 17 “objective improves, feasibility worsens.” Terminal COBYLA points matched the tracked best evaluated objective within 10^{-10} in every continuation run.

S4.2 Budget and optimizer controls

For one objective-selected seed per task, independent continuations at 120, 240, and 480 evaluations tested budget response. Median objective gains were 0.03197 from 120 to 240 and 0.01386 from 240 to 480; corresponding median G_{feas} changes were 0.08196 and 0.02410. Objective improved while G_{feas} worsened in 4/56 and 10/56 transitions, respectively. Thus additional budget did not remove objective–feasibility non-equivalence.

A matched 120-evaluation Nelder–Mead control had a median objective difference of -8.95×10^{-5} relative to COBYLA and a median G_{feas} difference of -0.1168 . It reached lower energy on 29/56 tasks; 15 of those also had lower G_{feas} . This control is diagnostic, not an optimizer leaderboard, because only one configuration per method was tested.

The complementary three-seed, three-optimizer and $p \leq 4$ controls are reported in [section S8](#).

S4.3 Decomposition and response slices

For every matched state, the stored mean objective equals the sum of routing, flow-penalty, and resource-penalty components within numerical tolerance. The four mutually exclusive mass categories—valid and resource-feasible, flow-invalid/resource-feasible, flow-valid/resource-violating, and both violating—sum to one. The medians reported in the main article’s optimizer-attribution section are computed from seed-level paired differences, not differences of separately aggregated medians.

Fifteen deterministic two-dimensional response slices varied the added (γ_3, β_3) while freezing earlier layers. On a 41×41 grid, median objective gain over the embedded point was 0.05648 and median maximum G_{feas} gain was 0.33887. These slices visualize available local response but cannot certify the global six-parameter optimum.

S5 Objective-Alignment Details

S5.1 Discovery objective summaries

The four-objective comparison contains 224 successful task–objective cells: 56 historical O0 cells reused from the matched 240-evaluation continuation and 168 newly evaluated O1/O2/O3 cells. The common statevector ansatz and initialization make taskwise contrasts meaningful, while the discovery status precludes confirmatory interpretation.

Discovery metric medians are given in [table S3](#).

The taskwise median O2–O0 gap is 0.2192983546 decades, with interquartile range 0.2928. The O2–O1 and O2–O3 residual medians are 0.2417 and 0.006897. Because medians do not distribute over ratios, the reported 2.8% and 97.9% gap closures are medians of per-task closure fractions, not ratios of the displayed median gaps.

Table S3 Discovery-stage metric medians over 56 tasks.

Arm	P_{feas}	G_{feas}	P_{opt}	$P_{\text{opt} \text{feas}}$	$\mathbb{E}[C \mid \text{feas}]$	median κ
O0 mean energy	0.021678	2.1161	0.009563	0.4373	13.0425	0.0777
O1 expected penalty	0.023357	2.1728	0.009509	0.4347	13.0131	0.0457
O2 exact feasibility	0.039915	2.3209	0.016458	0.4345	12.9871	0.1885
O3 CVaR-0.10	0.042787	2.4275	0.016696	0.4082	13.0241	0.1726

The exploratory mechanistic tags were frozen before evaluation and are nonexclusive: 44/56 tasks met the objective-limited criterion, 7/56 the capacity-limited criterion, 2/56 penalty-surrogate success, and 7/56 partial CVaR alignment. Continuous gaps and taskwise metrics remain primary; these tags are not latent-class estimates.

S5.2 Energy-tail checks

For each task, the maximum energy over all valid feasible states was compared with the minimum energy over all other states. Strict separation held on 140/140 v2 tasks. The O3 tail check compares the Boolean event $P_{\text{feas}} \geq \alpha$ with whether every probability contribution to the exact lowest-energy α tail is feasible, including the fractional cutoff mass. It held on 56/56 discovery and 84/84 held-out O3 cells.

S6 Held-Out Preregistration and Statistics

S6.1 Frozen hypotheses and analysis unit

For graph g , let L_g denote its full set of planned dilution levels and define

$$\Delta_{1g} = |L_g|^{-1} \sum_{\ell \in L_g} [G_{\text{feas}}(\text{O3}; g, \ell) - G_{\text{feas}}(\text{O0}; g, \ell)], \quad (\text{S34})$$

$$\Delta_{2g} = |L_g|^{-1} \sum_{\ell \in L_g} [G_{\text{feas}}(\text{O3}; g, \ell) - G_{\text{feas}}(\text{O2}; g, \ell)]. \quad (\text{S35})$$

All 15 graphs were complete. The population estimands are arithmetic means of Δ_{1g} and Δ_{2g} over graph units. H1 has null margin zero. H2 has non-inferiority margin -0.10 decades and is tested after shifting Δ_{2g} by $+0.10$.

The exact randomization-style p -value enumerates all sign vectors $s \in \{-1, +1\}^{15}$ and compares $15^{-1} \sum_g s_g \Delta_g$ with the observed mean in the one-sided direction. The two resulting values enter Holm’s two-hypothesis step-down correction. The grouped bootstrap samples 15 graph identifiers with replacement 10,000 times under seed 20261017 and takes the fifth percentile of the bootstrap mean as its one-sided lower bound. Individual task rows are never resampled as independent units.

The frozen inferential contract is summarized in [table S4](#).

S6.2 Confirmatory and secondary outcomes

The H1 estimate is 0.3547149490180106, with bootstrap standard error 0.07472582111285479 and lower bound 0.23737853648218638. Its raw sign-flip value is $p = 0.0001220703125$, and its Holm-adjusted value is $p = 0.000244140625$. The H2 estimate is -0.008595149938274194 , with standard error 0.016676047763212343

Table S4 Frozen held-out inferential contract.

Item	Frozen specification
Analysis unit	Base graph; arithmetic mean over all planned dilution levels
Primary family	H1 superiority and H2 non-inferiority only
Uncertainty	10,000 grouped bootstrap resamples; one-sided 5th-percentile lower bound
Tests	Exact $2^{15} = 32768$ graph-level sign flips
Multiplicity	Holm correction over the two-hypothesis family at $\alpha = 0.05$
H2 margin	-0.10 decades in G_{feas}
Separation	84 tasks and 15 graphs; zero overlap with discovery

and lower bound -0.03604874990402986 . Its raw shifted sign-flip value is $p = 0.000152587890625$, with the same Holm-adjusted value. These are the only members of the preregistered confirmatory multiplicity family.

Secondary task-level results include 80 O3-over-O0 P_{opt} wins, four losses, and no ties; 79 tasks improve both P_{feas} and P_{opt} . O3 ends at a higher mean Hamiltonian energy than O0 on 81/84 tasks, which is expected to be possible because O3 optimizes CVaR rather than the mean. Median feasible-probability change from the common initialization is 0.004496 for O0 and 0.023135 for O3. These values contextualize the primary graph-level endpoint but do not enlarge its hypothesis family.

S7 Post-Hoc Finite-Shot Endpoint Robustness

S7.1 Status and Design

This post hoc analysis samples the frozen O0 and O3 $p = 3$ terminal distributions for all 84 held-out tasks; it does not rerun optimization, change terminal parameters, or enter the preregistered H1/H2 family. The 168 endpoint distributions were regenerated deterministically from the task generator, penalty contract, simulator convention, and stored terminal angles. The maximum absolute difference from the authoritative exact endpoint metrics was 7.33×10^{-15} .

Each endpoint was sampled at 1,000, 10,000, and 100,000 shots with 20 deterministic replicates. A seed is the SHA-256-derived identity of task, objective, shot count, and replicate. Every replicate estimates P_{feas} , P_{opt} , mean normalized energy, and empirical lowest-energy CVaR-0.10. The CVaR implementation accumulates probability mass in ascending energy order and fractionally weights the cutoff observation when αN_{shots} is noninteger; for the three reported shot counts the cutoff is integral.

S7.2 Estimator Summary

For O0/O3, respectively, CVaR mean absolute error decreases from 0.0275/0.0352 at 1,000 shots to 0.0027/0.0034 at 100,000 shots. The corresponding P_{feas} mean absolute errors decrease from 0.0041/0.0056 to 0.00043/0.00054. Across task-replicate comparisons, the sampled sign of the O3-versus-O0 P_{feas} contrast agrees with the exact sign in 94.8%, 99.3%, and 100.0% of cases at 1,000, 10,000, and 100,000 shots. Full bias, standard deviation, RMSE, percentile, relative-error, and ordering summaries are retained in the post-hoc result directory. Ordering recovery is conditioned on a non-tied exact comparison, and exact ties are reported separately.

Figure S1 displays the estimator errors and ordering recovery.

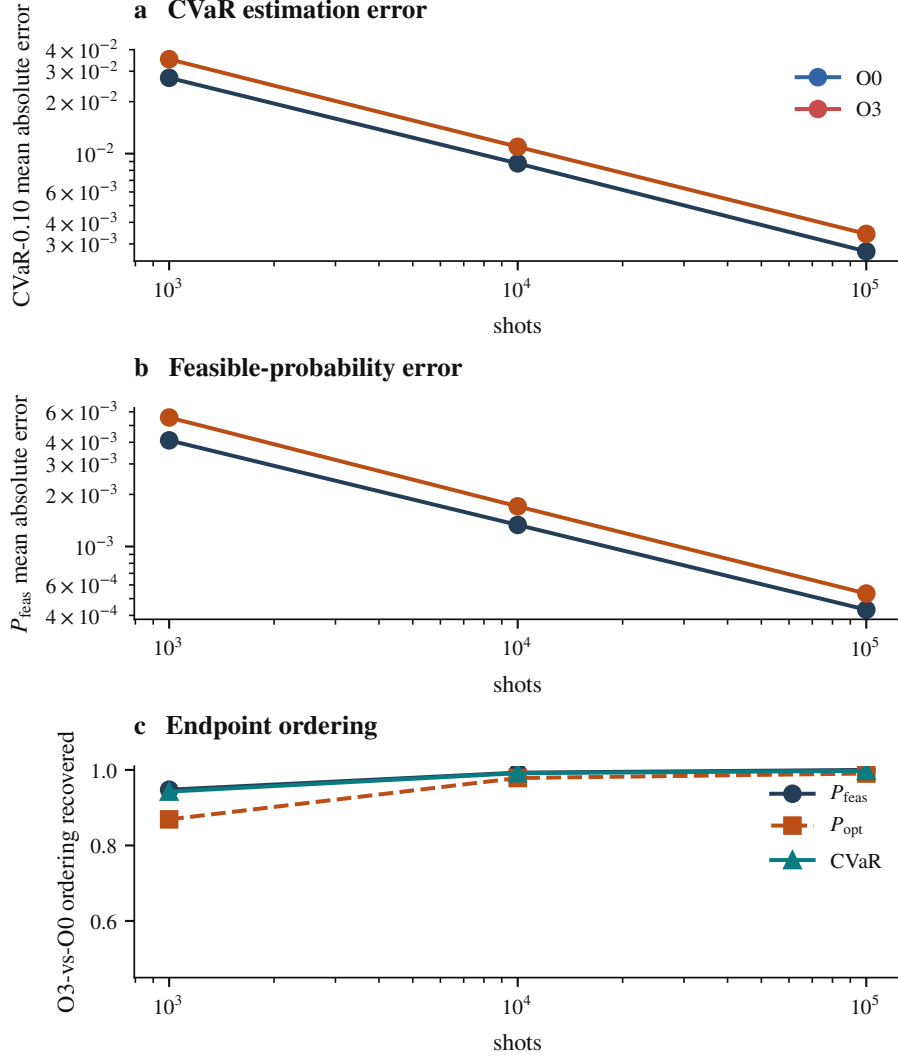


Fig. S1 Post-hoc finite-shot endpoint robustness. **a**, empirical CVaR-0.10 mean absolute error. **b**, feasible-probability mean absolute error. **c**, fraction of sampled O3-versus-O0 orderings matching the exact endpoint ordering, conditioned on non-tied exact comparisons. These fixed-endpoint simulator results do not evaluate finite-shot training, hardware noise, or device robustness

S8 Reviewer-Robustness Experiments

All analyses in this section are post-hoc robustness/sensitivity analyses and do not modify the preregistered H1/H2 family. The original 56-task discovery set, 84-task held-out set and $\alpha = 0.10$ selection remain fixed. Intervals aggregate paired task effects within graph and resample graphs with equal weight, consistent with the graph-level inference in [section S6](#). Task selections, seeds and factors were specified before these analyses were executed.

The complete protocol grid is listed in [table S5](#).

Table S5 Frozen post-hoc reviewer-robustness protocols. All task selections used graph and structural variables only. Additional tail fractions and all robustness intervals are exploratory and do not enter the preregistered H1/H2 family.

Study	Scope	Frozen factors
Optimizer	56 discovery tasks, 10 graphs	COBYLA, Nelder–Mead, SLSQP; O0/O3; $p = 2, 3$; three seeds; 120 actual objective calls
Tail fraction	56 discovery tasks, 10 graphs	$\alpha \in \{.02, .05, .10, .25, .50, 1\}$; COBYLA; $p = 3$; three seeds; 120 calls
Depth–budget	24 discovery tasks, 10 graphs	$p \in \{2, 3, 4\}$; 120/240/480-call prefixes; O0/O3; COBYLA; three seeds
Finite-shot training	10 held-out tasks, 10 graphs	O0/O3; $p = 3$; 1,000/10,000 shots per call; five seeds; 240-call cap
Classical context	All 140 canonical tasks, 25 graphs	Exact instrumented label-setting; repeated descriptive timing

S8.1 Optimizer Robustness

The formal matrix contains 2,016 runs: 56 tasks, three seeds, two depths, two objectives, and three local optimizers. Each run has a strict cap of 120 actual objective calls. Wrapper and SciPy counts agree on every one of the 1,512 runs for which SciPy returns an `nfev`; the 504 SLSQP runs stop at the wrapper cap while numerical finite differences are being evaluated, so the absent SciPy terminal count is retained as missing rather than imputed. Termination status, message, iteration count, wall time, and both counters are stored per run.

Nested failures occur in both O0 and O3 under every optimizer ([table S6](#)). The dual-PASS restriction then retains only task–seed pairs for which neither arm has a positive signed nested regret above 10^{-10} . The positive graph intervals and frequent O0 lower-energy/lower-feasibility pairs persist in that restricted population. The post-hoc COBYLA/O0 count of 28/168 is a separately frozen strict-cap rerun, not a redefinition of the historical 29/168 observation, which remains unchanged. This supports empirical separation of optimizer inadequacy and objective misalignment under the diagnostic, not a universal statement about either optimizer or loss landscape.

Table S6 Optimizer robustness on 56 discovery tasks and three fixed seeds. Nested failures compare the best $p = 3$ point with the exactly embedded $p = 2$ point under 120 actual objective calls. The last two columns retain only pairs for which both O0 and O3 pass that diagnostic; the interval is a 95% graph-cluster bootstrap CI for O3-minus-O0 G_{feas} . Configurations are diagnostic, not an optimizer ranking.

Optimizer	O0 nested failure	O3 nested failure	dual-PASS pairs	O0 lower energy, lower P_{feas}	PASS-only ΔG_{feas} graph CI
COBYLA	28/168	38/168	113	103/113	[0.5770, 0.9971]
Nelder–Mead	54/168	61/168	72	65/72	[0.5796, 1.2473]
SLSQP	46/168	72/168	70	58/70	[0.6096, 1.3458]

Figure S2 visualizes the failure rates and PASS-only effects.

S8.2 CVaR Tail-Fraction Sensitivity

The discovery scan completes 1,008/1,008 planned runs with no failed records. The response peaks at the historically frozen $\alpha = 0.10$ in the equal-weight graph mean, but the nearby 0.05 and 0.25 responses differ by only -0.0386 and -0.0329 decades; both graph-cluster intervals for these differences include zero. By contrast, the optimized response is not assumed to interpolate monotonically, and 54/56 task trajectories are non-monotone. Additional held-out scans at $\alpha = 0.05, 0.25, 0.50, 1.00$ are labelled POST_HOC_SENSITIVITY; none is used to select a new primary tail fraction.

The $\alpha = 1$ control covers 222 executed discovery and post-hoc held-out cells and gives a maximum direct objective discrepancy of 2.42×10^{-13} from mean energy. For the 168 discovery cells matched to the closest historical O0 control, task, depth, seed, exact embedded- $p = 2$ initial parameters, nominal budget, optimizer family, initial trust-region radius, and constraint tolerance agree. The initial objective discrepancy is at most 2.02×10^{-13} . The historical and post-hoc wrappers and SciPy versions differ, however, and complete comparable trajectories were not persisted; their terminal parameters need not coincide. This comparison establishes objective equivalence only.

Table S7 reports the numerical tail-fraction summary.

Table S7 Post-hoc CVaR tail-fraction sensitivity on 56 discovery tasks. Values are equal-weight means over ten graph aggregates. The frozen confirmatory choice remains $\alpha = 0.10$; the final column is descriptive.

α	Graph-mean G_{feas}	Difference from $\alpha = 0.10$
0.02	1.8982	-0.0907
0.05	1.9503	-0.0386
0.10	1.9889	+0.0000
0.25	1.9560	-0.0329
0.50	1.8445	-0.1444
1.00	1.6867	-0.3022

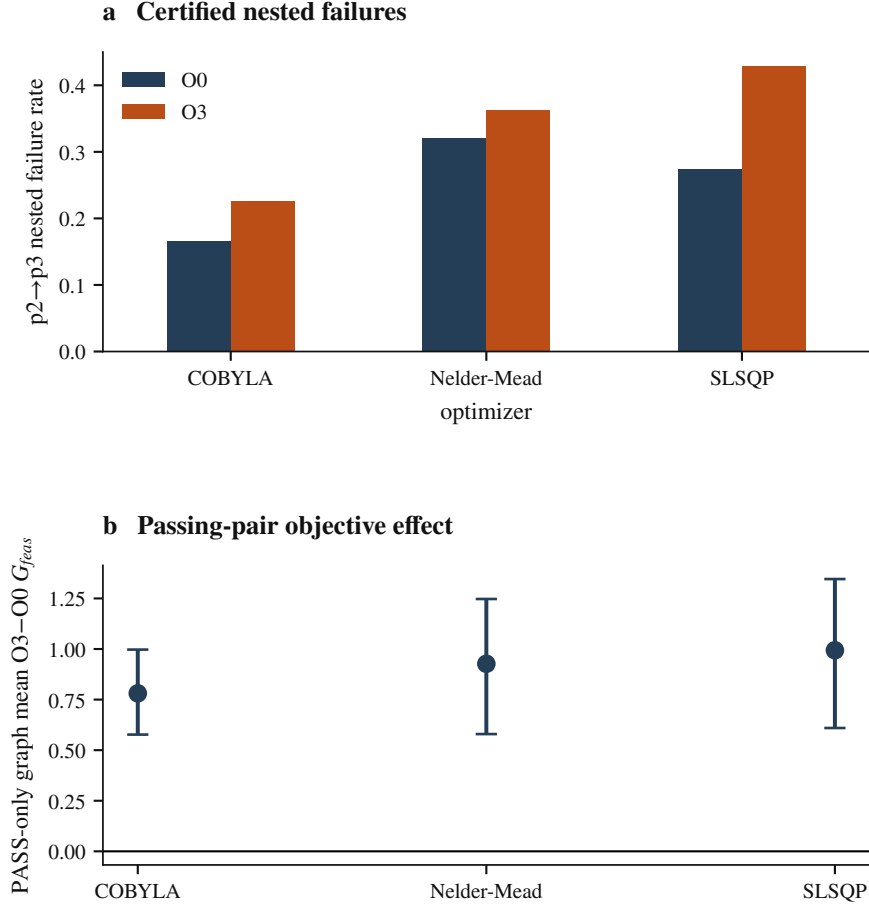


Fig. S2 Optimizer robustness. Nested-failure rates are shown separately for O0 and O3; paired graph-level O3-minus-O0 feasibility effects are evaluated on runs passing both diagnostics. All configurations use the same 120-call resource unit, but the figure is not an optimizer leaderboard

S8.3 Depth–Budget Ablation

The 24 tasks were selected using graph, size, feasible fraction, and dilution, without reference to O0/O3 outcomes. The formal run stores one deterministic 480-call trajectory after verification confirmed that its first 120 and 240 objective calls match separately capped runs exactly. Each reported checkpoint is the best evaluated incumbent in that genuine prefix, rather than necessarily the solver’s last evaluated point.

[Figure S4](#) retains the complete depth–budget panels.

Verification covered 16 structural checks and 24 stratified PASS/FAIL cases evaluated from saved parameters. Maximum absolute discrepancies were 1.55×10^{-15} for the deeper objective, 2.22×10^{-16} for the embedded objective, and 1.42×10^{-15} for signed regret; both embedding identities were exact at stored precision. The conclusion is limited to the tested subset, COBYLA, O0/O3, depths, seeds, and budgets: increasing the tested classical evaluation budget did not monotonically reduce certified nested-ansatz failures. The positive $p = 4 - p = 3$ feasibility effects in [table S8](#) show why that diagnostic must not be conflated with final-state quality.

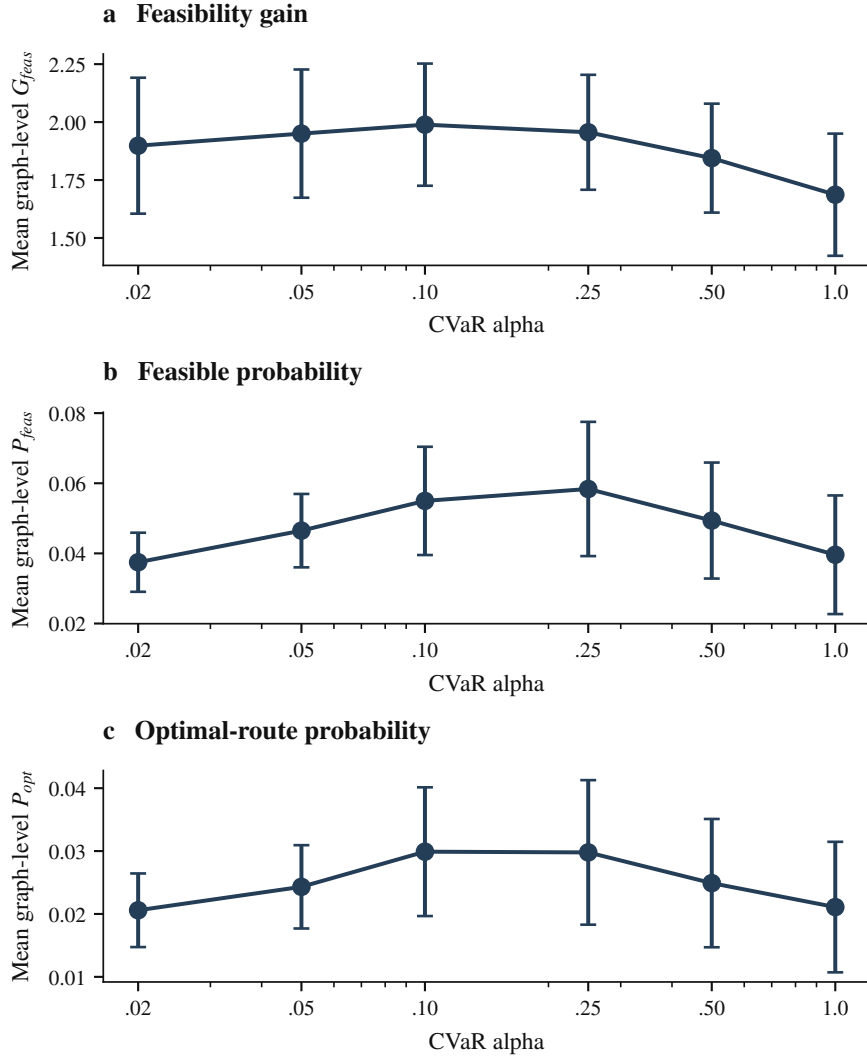


Fig. S3 Tail-fraction sensitivity: discovery task summaries and equal-weight graph effects across the prespecified grid. Error bars are standard errors across discovery graphs; $\alpha = 0.10$ remains the confirmatory choice

Table S8 Depth-budget ablation on the pre-frozen 24-task subset. Failure counts pool O0 and O3 over three seeds. $p = 4 - p = 3$ effects are equal-weight graph means in G_{feas} ; the last column gives positive-graph counts for O0; O3. Checkpoints are the best evaluated incumbents in verified trajectory prefixes.

N_{fev}	$p2 \rightarrow p3$ failure	$p3 \rightarrow p4$ failure	$p4 - p3$ O0 G_{feas}	$p4 - p3$ O3 G_{feas}	positive graphs O0; O3
120	30/144 (20.8%)	45/144 (31.2%)	0.7206	0.2021	10/10; 10/10
240	31/144 (21.5%)	45/144 (31.2%)	0.7843	0.2074	10/10; 10/10
480	38/144 (26.4%)	47/144 (32.6%)	0.8683	0.1970	10/10; 10/10

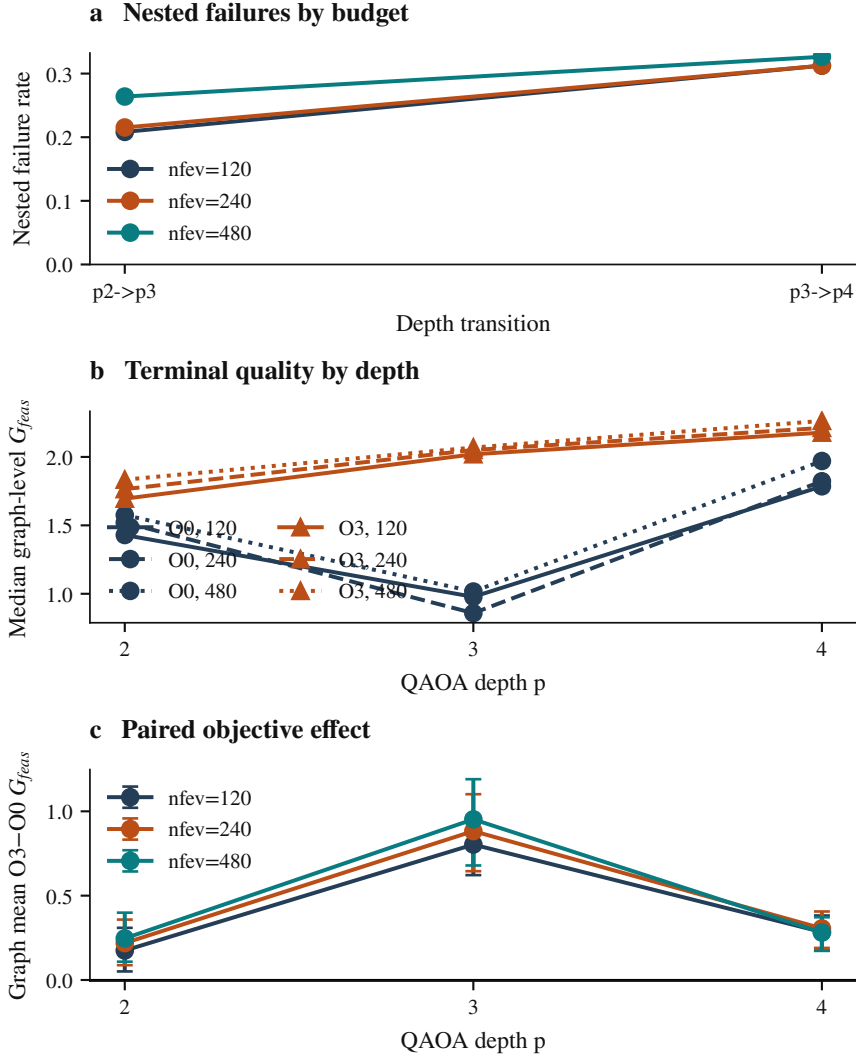


Fig. S4 Depth-budget robustness: pooled nested failures, terminal graph quality by depth and objective, and paired O3–O0 graph effects with 95% graph-bootstrap intervals. The depth-gain and budget-boundary view appears in the main article

S8.4 Finite-Shot Estimation

The all-held-out fixed-endpoint protocol and its 20 replicates per cell are detailed in [section S7](#). A separate outcome-blind 18-task subset uses 50 sampling seeds to cross $\alpha \in \{0.02, 0.05, 0.10, 0.25, 0.50\}$ with $N_{\text{shots}} \in \{10^3, 10^4, 10^5\}$. The mean of the separately computed O0 and O3 CVaR-0.10 RMSE values falls from 0.04839 to 0.01570 and 0.00486 as shots increase. The heat map in [figure S5](#) shows the larger sampling cost of more aggressive lower tails. No quantum optimization is rerun for this estimator analysis.

S8.5 Finite-Shot Training

The finite-shot training matrix completes 200/200 runs with no failed records. Each objective evaluation uses a fresh reproducible sample; final parameters are evaluated

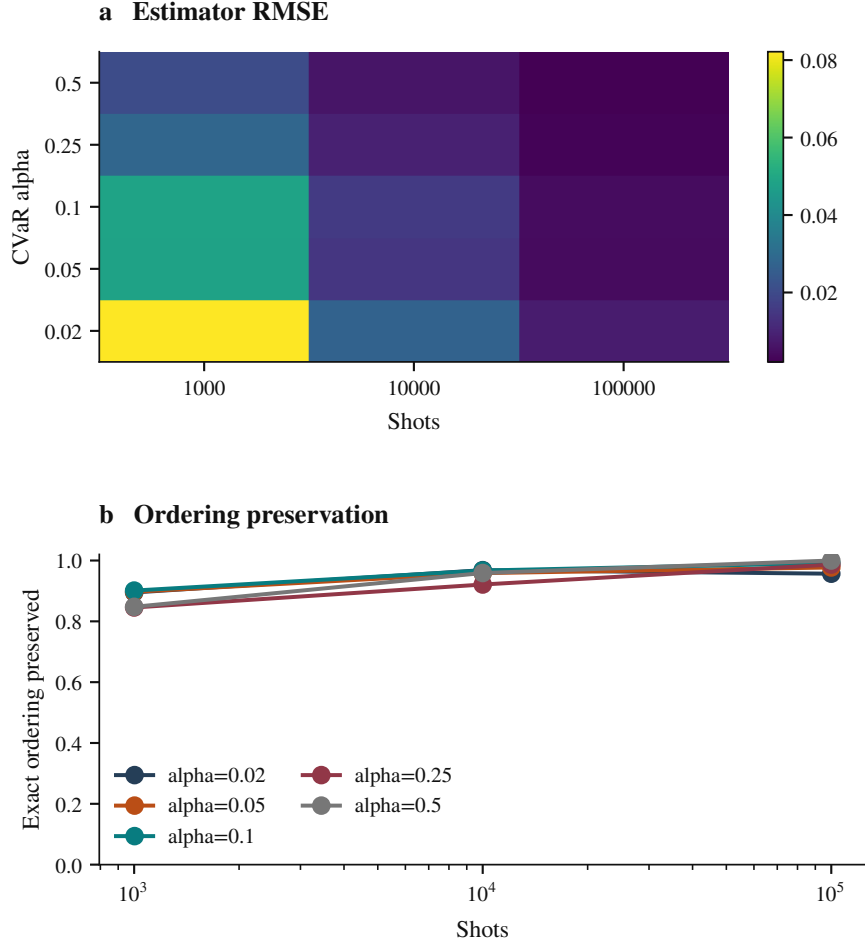


Fig. S5 CVaR sampling error across tail fraction and shot count at fixed exact-trained endpoints. Smaller tails generally require more shots

exactly. The two shot-trained point estimates preserve the exact O3-minus-O0 direction, but both graph-cluster intervals cross zero (table S9). This is mixed evidence about noisy training and stronger evidence about fixed-endpoint estimation; it is not a hardware or device-noise experiment.

Table S9 O3-minus-O0 finite-shot training effects on ten pre-frozen tasks from ten held-out graphs. Final parameters are evaluated exactly; intervals are 95% graph-cluster bootstrap CIs. The exact row uses the matched canonical endpoints for the same tasks.

Training regime	Positive graphs	Mean ΔG_{feas}	Graph CI
Exact training	9/10	0.3599	[0.1641, 0.5697]
10,000-shot training	8/10	0.1706	[-0.0106, 0.3862]
1,000-shot training	6/10	0.1246	[-0.0050, 0.3351]

Figure S6 displays graph intervals and terminal-state distributions.

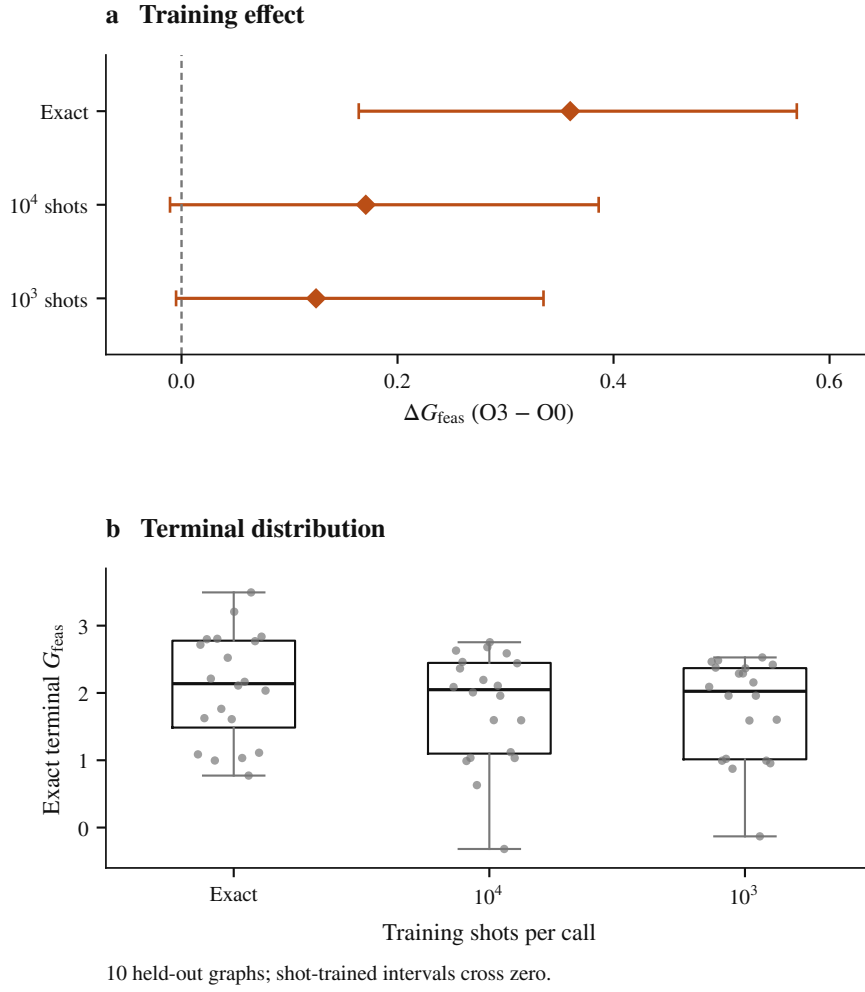


Fig. S6 Exact-terminal evaluation after exact or finite-shot training on ten held-out graphs. **a**, paired O3–O0 graph effects and two-sided 95% graph-bootstrap intervals; both shot-trained intervals cross zero. **b**, terminal quality for the same three training regimes, pooling 20 task–objective summaries per regime. Boxes show quartiles and medians, whiskers extend to the most extreme observations within 1.5 interquartile ranges, and dots show all summaries. These are task summaries, not individual noisy-training runs

S8.6 Classical RCSP Context

An instrumented exact label-setting solver reproduces the frozen optimum on 140/140 canonical tasks. Across repeated high-resolution solves, the median, 95th-percentile, and maximum per-task medians are 50.0, 95.4, and 188.1 microseconds; generated-label counts have median 13 and maximum 30. These small descriptive values confirm the intended context: the instances are classically tractable mechanism probes. They do not define a quantum–classical runtime comparison.

[Table S10](#) gives the compact classical-solver summary.

Table S10 Exact classical RCSP context. Times are descriptive medians of repeated high-resolution measurements and are not compared with statevector runtime or hardware performance.

Matches	Median (μ s)	p95 (μ s)	Max (μ s)	Median labels	Maximum labels
140/140	50.0	95.4	188.1	13	30

S9 Scaling Analysis

S9.1 Frozen universe and splits

The scaling universe contains 30 base graphs and 180 tasks. Candidate-route counts range from 6 at $m = 12$ to 17 at $m = 22$; each graph contributes six distinct feasible-cardinality levels. The global range is $\phi_{\text{state}} = 2.3842 \times 10^{-7}$ to 1.4648×10^{-3} , or $D = 2.8342\text{--}6.6227$. Phase-0–2 rows are excluded from task generation, model selection, and coefficient fitting.

The scale-controlled coefficient changes from 172 to 199 because the declared maximum routing cost changes to $22 \times 9 = 198$. The normalized ansatz, objectives, $p = 2$ preparation, $p = 3$ continuation, 240-evaluation budget, and CVaR $\alpha = 0.10$ otherwise follow the held-out protocol.

S9.2 Resource guard

Before optimization, representative fixed-parameter forward passes were timed at each m . A size was permitted only if a numerically valid statevector was obtained, predicted worker memory was at most 40% of available RAM, and predicted optimization time did not exceed 1200 seconds. The guard passed at $m = 12, 14, 16, 18, 20$ and failed at $m = 22$ because the predicted optimization time was 2249.3 seconds. The $m = 22$ cells were therefore administratively censored by the preregistered computational budget before optimization. The observed resource table is retained in the frozen artifact; no result-dependent code rewrite or GPU extension was allowed.

S9.3 Model selection and response summaries

Four prespecified response forms were evaluated on development graphs with leave-one-base-graph-out cross-validation. The one-standard-error rule selected the simplest linear model defined in the main article for every objective. Development cross-validated RMSE was 0.33769 for O0, 0.16922 for O2, and 0.19391 for O3. The graph-level exponent distributions, rather than only their means, are shown in [figure S11](#).

At the completed $m = 20$ extrapolation split, the preregistered mean η_{O0} hypothesis was supported, but the intended O3-versus-O0 objective contrast was not; its observed sign reversed. The scientific verdict is MIXED OBJECTIVE SCALING; $m = 22$ supplies no QAOA outcome and no positive scaling result. Historical Phase-1/2 exponent summaries are read-only comparisons and never enter the frozen model fit.

S9.4 Optimization adequacy

Twelve selected development cells compared 240 and 480 objective evaluations. The prospectively defined high-size optimizer-limited ceiling did not trigger: no O0/O2/O3 cell at $m = 16$ or 18 exceeded the 0.10-decade substantial ΔG_{feas} threshold. Several $m = 12$ and $m = 14$ controls were budget-sensitive, so this check does not prove global optimization. It only prevents using a prespecified high-size budget artifact as an explanation for the $m = 20$ objective reversal.

The retained execution and censoring census is summarized in [table S11](#).

Table S11 Retained execution and censoring census. A zero denotes a retained planned category with no observed event.

Stage	Success	Timeout/ OOM	Other scientific failure	Resource- censored
Pilot optimized runs	504	0	0	0
Held-out $p = 2$ preparation	252	0	0	0
Held-out $p = 3$ comparison	252	0	0	0
Scaling $p = 2$ and $p = 3$	1080	0	0	180 planned cells

Online Resource 2 supplies the complete execution census, including zero-count failure categories and every scaling split/size/depth/objective cell.

S9.5 Censoring interpretation

Resource censoring is deterministic under the frozen guard and is visible in all summaries. It is not treated as random missingness, and no imputation, survival model, or extrapolated $m = 22$ QAOA value is reported. The censoring records this study’s computational-budget limitation; it is not a claim that 22 qubits are intrinsically unsimulable, a statevector-memory limit, or a principled barrier for the quantum algorithm.

S10 Reproducibility

S10.1 Code, data and software

Code, benchmark definitions, task manifests, compact result summaries and reproduction documentation are available at <https://github.com/ZhilinChen02/Penalty-X-QAOA-A-Controlled-RCSP-Study>. The repository provides environment specifications and separate instructions for reconstructing the reported summaries from retained result rows. Raw statevectors and individual samples were not persisted. The public release does not include raw optimizer traces or the complete set of experimental working files.

S10.2 Provenance and numerical records

Machine-readable provenance and audit records are supplied in Online Resource 2. Its README describes the claim/evidence mapping, numerical audit, execution census and task-stratum summary. The artifact manifest records selected source fingerprints, summarized in [table S12](#).

Table S12 Selected frozen-input SHA-256 fingerprints (prefixes shown; complete values are in `artifact_manifest.txt` in Online Resource 2).

Asset	SHA-256 prefix
Task-universe manifest	33a2c1b689065951...
Held-out preregistration	87b8c3d814d74098...
Held-out statistics	bb4d24d65cd2cf9c...
Scaling freeze	6c0df9df4f7a1c00...
Synthesis claim matrix	4dc908667e31774f...

S10.3 Figure reproduction

The main figures use the v3 renderings from `paper_scripts/redesign_main_figures_v3.py`. Supplementary result plots use `redraw_supplement.py` with the original plotting inputs, groups, display summaries and stored intervals; the numerical-array checks are recorded in `REDRAW_DATA_AUDIT.json`. The conceptual flow diagrams in [figures S7](#) and [S13](#) are deterministic vector schematics and contain no empirical plot arrays. Reproduction details concern the displayed evidence and do not extend the study’s scientific scope: no quantum advantage or global scaling claim is made.

S11 Extended Figures and Comparison Tables

S11.1 Access models and literature comparisons

[Figures S7](#) and [S8](#) summarize access models and resource categories; [table S13](#) records the interpretation boundaries.

Representation, query access and structure

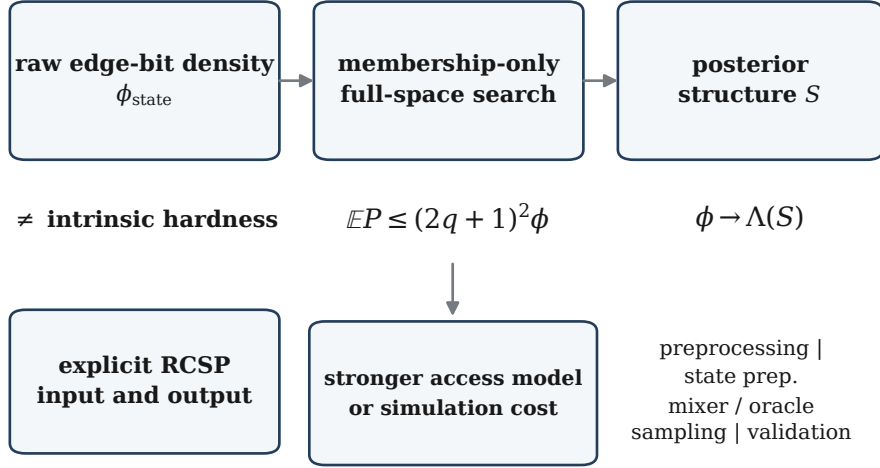


Fig. S7 Access-model boundary: raw density describes the encoding. Membership bounds require all residual feasible-set information to enter through counted queries; structure/advice costs must be recorded

Structure-related cost categories					
	Information	Classical preprocess	State preparation	Mixer / compilation	Query / validation
Penalty-X	cost	cost	cost	--	cost
CVaR Penalty-X	cost	cost	--	--	cost
Warm start	cost	cost	cost	--	--
Feasible-state prep.	cost	cost	cost	--	cost
Path-exchange mixer	cost	cost	--	cost	cost
Explicit path basis	cost	cost	cost	cost	--
Corridor restriction	cost	cost	--	--	cost

Cost categories in the method ledger;
not scalar runtime estimates.

Fig. S8 Structure-cost relocation. Filled cells identify resource categories associated with each method, not measured magnitudes. The diagram is an accounting taxonomy rather than a runtime comparison

Table S13 Theory statements are tied to their information-access model. They are interpretation boundaries, not an end-to-end runtime theorem for the experiments.

Result	Model and conclusion	Explicit nonclaim
Membership dilution	Uniform random size- M marked set; counted membership queries; $\mathbb{E}P \leq \min\{1, (2q+1)^2\phi\}$	Not pointwise for every set; not rich-energy RCSP
Raw-density counterexample	Explicit chain has $\phi_{\text{state}} = 2^{-m}$ and $O(m)$ output time	Raw density is not intrinsic explicit-RCSP hardness
Attribute-query RCSP	Length- K counted random-access array; $\Theta_\tau(\sqrt{K/M})$ in the nontrivial target regime	Not explicit RAM time or compact-input exponential hardness
Posterior structure	Classical side information enters through $\Lambda(S)$	Advice concentration is not implementation time
Finite advice	At most b classical support bits multiply the coarse bound by 2^b	Necessary condition only; not sufficient

All works in [tables S14](#) and [S15](#) are cited and discussed in the main article, which supplies their bibliographic details.

Table S14 Representative constrained-QAOA and objective-design comparators. Publication status was verified on 4 September 2026

Work	Constraint domain	Initial state	Mixer/dynamics	Training objective	Principal distinction
Hadfield et al.	Problem-defined feasible subset	Feasible state	Constraint-preserving alternating operators	Problem cost	General framework; changes support and dynamics
Barkoutsos et al.	Diagonal combinatorial cost	Ansatz dependent	VQE/QAOA examples	CVaR	Introduces CVaR; attribution is not the endpoint
Onah–Firt–Michielsen (preprint)	One-hot product manifold	Block W states	Block- XY	Energy/sampling design	Encoded structured construction with preparation cost
Bucher et al. (NJP)	One-hot constraints	Iterative warm start	Warm-start XY	Mean energy	Jointly changes initialization and mixer
Onah–Michielsen (preprint)	Spectrally filtered constrained model	Model dependent	Fejér-filtered evolution	Target spectral weight	Finite-depth/shot guarantee under explicit spectral assumptions
Ferrari et al. (preprint)	Multiple penalized constraints	Standard ramp start	Scheduled penalty phases	Feasibility loss	Optimizes constraint schedules and loss together
This work	Edge-bit RCSP penalty	Uniform $ +\rangle$	Transverse X	O0, O1, O2, O3	Holds Hamiltonian and ansatz fixed for attribution

Table S15 Validation-design comparison. “No” means the item is not a study-level attribution design in the cited work, not a quality judgment

Work	Nested optimizer certificate?	Graph-held-out objective test?	Exact-feasibility capacity control?	Relation to this study
Hadfield et al.	No	No	No	Supplies the structured alternating-operator framework
Barkoutsos et al.	No	No	No	Supplies the established CVaR objective
Onah–Firt–Michielsen	No	No	No	Structured one-hot competitor under a different access model
Bucher et al.	No	No	No	Warm-start/XY intervention, not a loss-only contrast
Onah–Michielsen	No	No	No	Analytic finite-resource model, not Penalty-X RCSP attribution
Ferrari et al.	No	No	No	Feasibility loss plus penalty scheduling, unlike fixed- H_C O3
This work	$p = 2 \rightarrow 3$; post-hoc $p = 3 \rightarrow 4$	15 graphs, frozen H1/H2	O2, mechanistic only	Separates optimizer adequacy, capacity, and loss

S11.2 Objective and scaling diagnostics

Figure S9 shows the exploratory objective comparison; table S16 records the experimental stages.

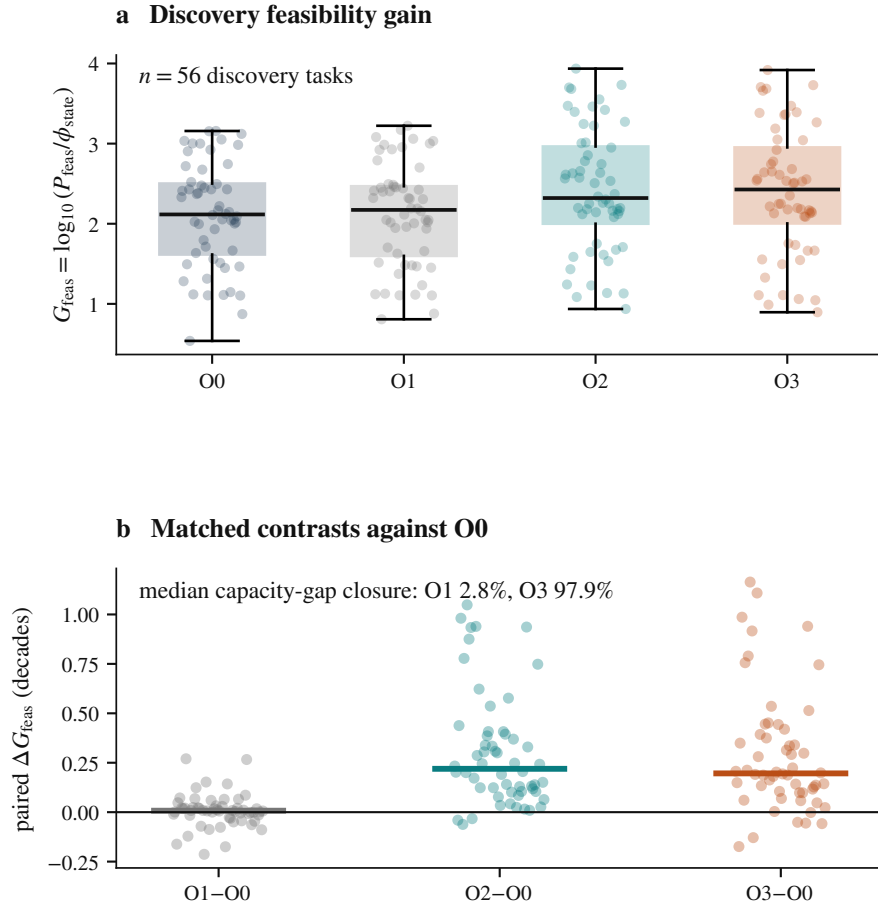


Fig. S9 Objective discovery on 56 tasks. **a**, feasibility gain under four classical losses applied to the same $p = 3$ ansatz. **b**, paired contrasts against O0; horizontal segments mark medians. O2 is an exact-feasibility capacity control, not a deployable method. This figure is exploratory and was used to select O3 before held-out analysis

Table S16 Experimental stages and protocol roles. Discovery and held-out rows are never pooled for inference.

Stage	Scientific role	Tasks	Graphs	QAOA setting
Task audit v1	Expose duplicate feasible-set levels	175	25	No QAOA
Distinct-cardinality v2	Controlled universe and scale audit	140	25	No QAOA
Depth pilot	Exploratory dilution response	56	10	Penalty-X, $p = 1, 2, 3$
Optimizer diagnostic	Nested-ansatz and continuation attribution	56	10	Penalty-X, $p = 3$
Objective discovery	Select the held-out objective	56	10	O0–O3, $p = 3$
Held-out evaluation	Preregistered graph-level confirmation	84	15	O0/O2/O3, $p = 3$
Scaling boundary	Prespecified size-conditioned response; $m = 22$ administratively censored	180 planned	30	O0/O2/O3, $p = 3$

Figure S10 decomposes the mean-energy/feasibility mismatch.

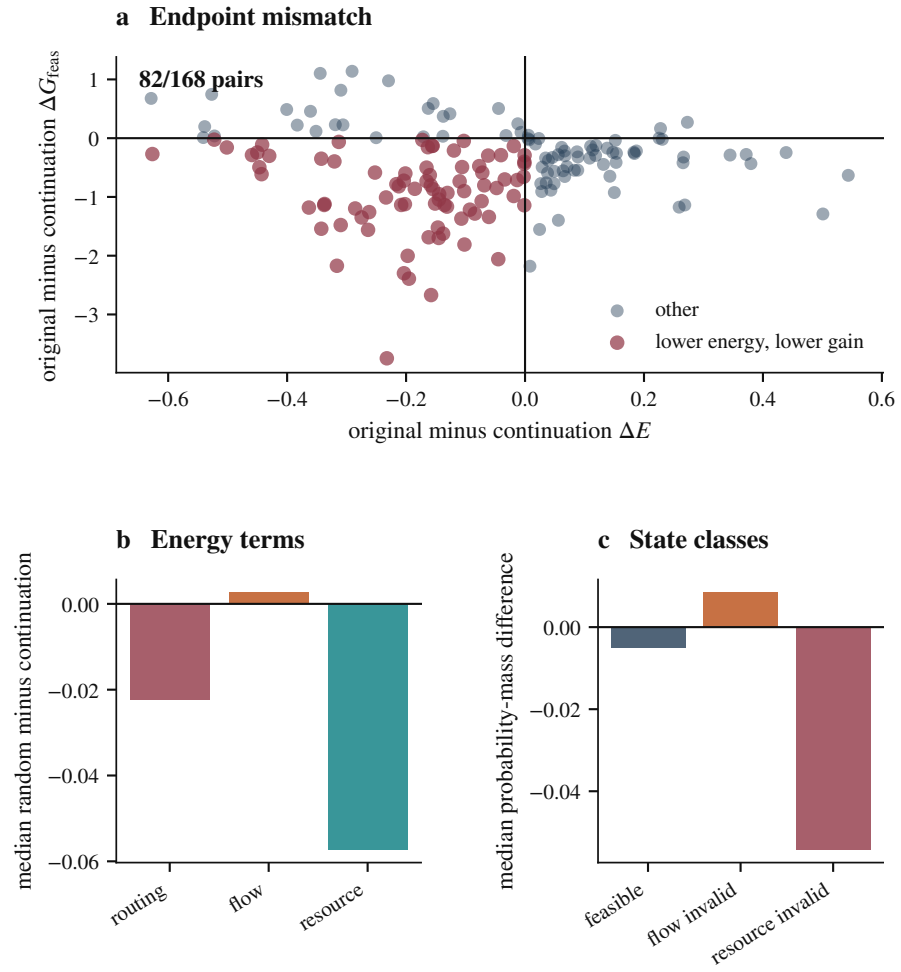


Fig. S10 Mean energy and feasible mass are distinct targets. **a**, paired original-minus-continuation differences. **b**, median energy-component differences. **c**, median state-class mass differences

Figures S11 and S12 show size-conditioned responses and the held-out feasible-entry decomposition.

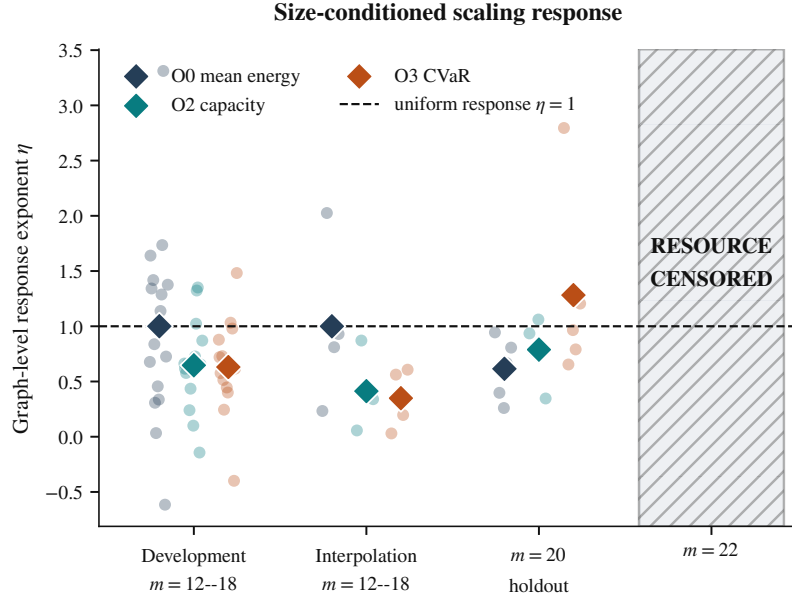


Fig. S11 Graph-level response exponents from protocol-defined splits. Small points are graphs, diamonds are split means, and the dashed line marks uniform-state response $\eta = 1$. The $m = 22$ column is administratively censored under the preregistered computational budget; no outcome or extrapolation is drawn

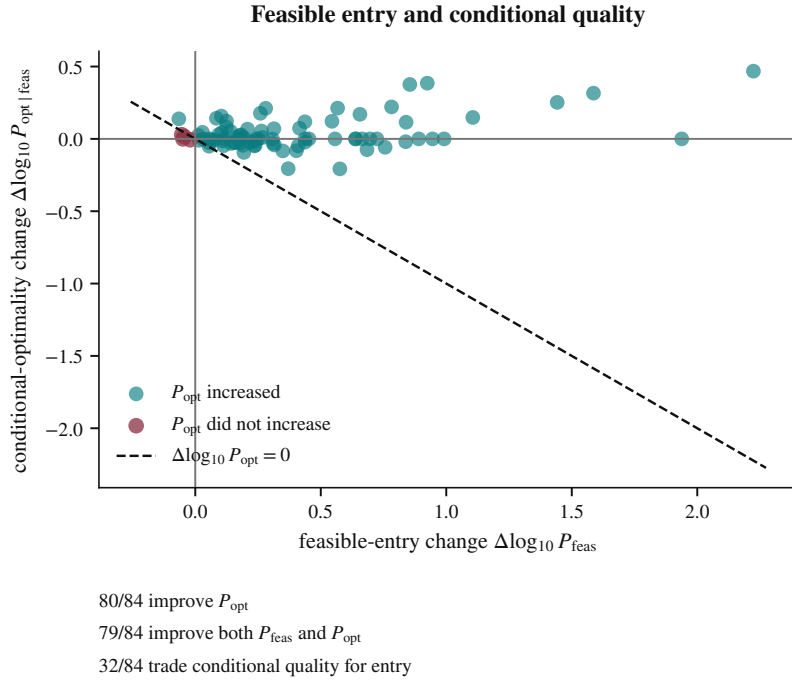


Fig. S12 Held-out feasible-entry and conditional-optimality decomposition. Each point is an O3-versus-O0 task contrast. Above the dashed line, $\Delta \log_{10} P_{\text{opt}} > 0$ by $P_{\text{opt}} = P_{\text{feas}} P_{\text{opt}|\text{feas}}$. Positive feasible-entry effects can coexist with heterogeneous conditional route quality

S11.3 Training objectives and attribution

Table S17 compares the four classical losses; figure S13 summarizes the attribution sequence.

Table S17 Classical training losses applied to the same $p = 3$ Penalty-X ansatz and the same normalized cost-phase Hamiltonian.

ID	Name	Loss minimized	Role
O0	Mean energy	$\mathbb{E}_\theta[H_C]$	Standard variational baseline
O1	Expected penalty	$\mathbb{E}_\theta[P_{\text{flow}} + P_{\text{resource}}]$	Mechanistic feasibility surrogate
O2	Exact feasibility	$1 - P_{\text{feas}}(\theta)$	Statevector-only ansatz-capacity control; not deployable
O3	CVaR-0.10	Mean of the lowest-energy probability mass $\alpha = 0.10$	Tail-focused rich-energy objective

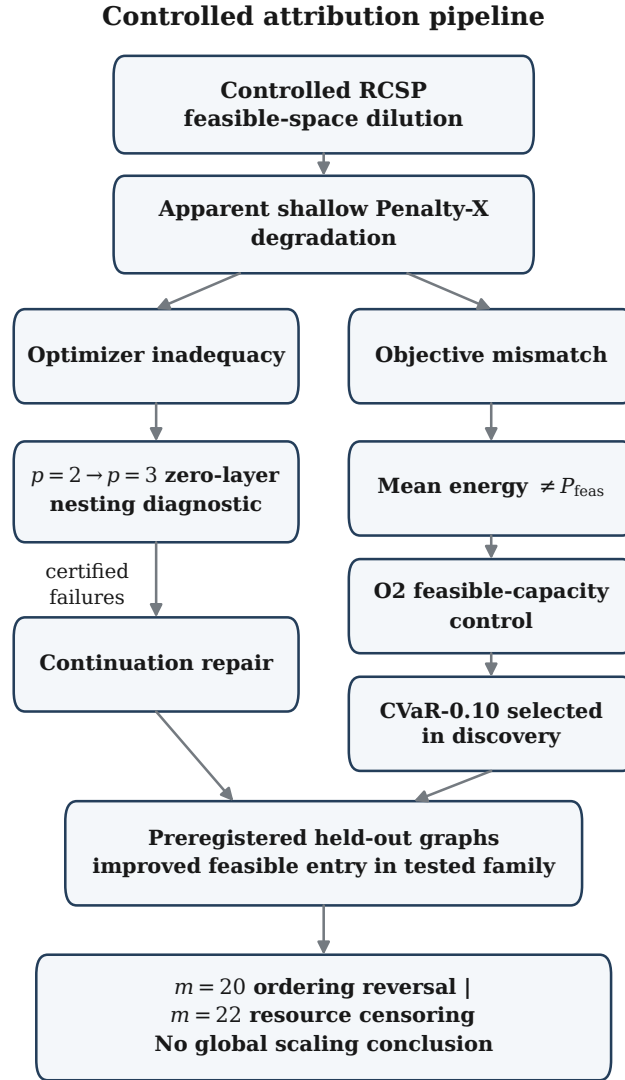


Fig. S13 Controlled attribution: nested diagnostics assess optimizer failure; matched losses assess alignment. Discovery-selected CVaR is tested on held-out graphs. The $m = 20$ reversal and $m = 22$ censoring limit scaling conclusions; post hoc tests do not modify confirmation