

QAOA energy bounds at large interaction order: odd spins and sparse logarithmic depth

Yuxuan Zhang

Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL),
CH-1015 Lausanne, Switzerland

Department of Physics, Princeton University,
Princeton, New Jersey 08544, USA

26 October 2026 — Version 1

Abstract

We give a proof candidate for a large-interaction-order obstruction to the quantum approximate optimization algorithm (QAOA), including odd orders. In the ordered Gaussian q -spin model, with variance n at each assignment, the limiting expected energy per spin at every fixed number of layers is at most any prescribed $\alpha > 0$, once q exceeds a threshold depending only on α . Angles are deterministic and independent of the instance. The argument first treats a signed Poisson hypergraph model. Exact outside-marginal couplings realize a tree of QAOA outputs; a metric packing lemma selects forbidden outputs at possibly different interpolation times. A marked-Poisson calculation bounds these events without a Gaussian approximation. Quantitative locality and variance estimates give a sparse obstruction for at most $\kappa \log n$ layers. The established fixed-depth sparse–dense correspondence gives the Gaussian conclusion. The contribution is this particular mixed-time covariance and coupling argument, rather than a new general overlap-gap or quantum-locality principle.

Status. Author-directed, AI-assisted proof candidate. This version has received a sequential proof and source-interface audit, but no independent review of the complete manuscript, external refereeing, or formal verification. No priority or catalog-acceptance claim is made.

1 Models and main results

For $x \in \{-1, 1\}^n$, let

$$H_{n,q}^G(x) = n^{-(q-1)/2} \sum_{i_1, \dots, i_q=1}^n J_{i_1 \dots i_q} x_{i_1} \cdots x_{i_q}, \quad J_{i_1 \dots i_q} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1). \quad (1)$$

Ordered tuples and repeated coordinates are retained, so $\text{Var } H_{n,q}^G(x) = n$. Given a diagonal Hamiltonian H , QAOA prepares

$$|\psi_H\rangle = e^{-i\varphi_p B} e^{-i\theta_p H} \dots e^{-i\varphi_1 B} e^{-i\theta_1 H} |+\rangle^{\otimes n}, \quad B = \sum_j X_j. \quad (2)$$

Angles are independent of the sampled instance. Define

$$e_{p,q} = \sup_{\theta, \varphi \in \mathbb{R}^p} \lim_{n \rightarrow \infty} n^{-1} \mathbb{E} \langle \psi_H^G | H_{n,q}^G | \psi_H^G \rangle.$$

The fixed-angle limit exists by Basso, Gamarnik, Mei, and Zhou [1, Theorem 1]. It precedes optimization.

Theorem 1.1 (Gaussian model). *For every $\alpha > 0$ there is $q_0(\alpha)$ such that $e_{p,q} \leq \alpha$ for every integer $q \geq q_0(\alpha)$ and every fixed integer $p \geq 0$. In particular, setting $\alpha = \zeta\sqrt{2\log 2}$ proves the stated M26 target for all sufficiently large odd q , with threshold independent of p .*

The centered energy and the order of limits matter. Setting all angles to zero gives a zero-energy baseline consistent with the bound. The theorem limits substantial positive limiting energy within this QAOA class; it is not a statement about all quantum optimization algorithms or the exact optimum at a finite q .

For the sparse model, sample $M \sim \text{Pois}(dn)$, independent uniform ordered tuples $E_e \in [n]^q$, and independent signs $J_e \in \{-1, 1\}$. Set

$$H_{n,q,d}^S(x) = d^{-1/2} \sum_{e=1}^M J_e \chi_{E_e}(x), \quad \chi_E(x) = \prod_{j=1}^q x_{E_j}. \quad (3)$$

An occurrence is one term of the list, even when tuples repeat. Let X_H denote the actual Born output of (2), and write $v_{n,p,q,d} = \mathbb{E}[H(X_H)/n]$.

Theorem 1.2 (Sparse logarithmic depth). *For every $\alpha > 0$ there are $q_0(\alpha), d_0(\alpha)$ such that, for fixed $q \geq q_0, d \geq d_0$, there is $\kappa(\alpha, q, d) > 0$ satisfying*

$$v_{n,p,q,d} \leq \alpha + n^{-1/8}$$

for all sufficiently large n , all $p \leq \kappa \log n$, and all deterministic angle lists. The sufficiently large threshold is uniform over these depths and angles. The angles may depend on n, p, q, d , but not the instance. Both parities of q are included.

Relation to prior work. Chou, Love, Sandhu, and Shi [2] establish local quantum algorithm obstructions, marginal-sharing principles and concentration. Basso et al. [1] supply fixed-angle thermodynamic limits and the sparse–dense correspondence used here; their stated obstruction covers even orders. Goh [4] studies logarithmic depth and limit ordering for even-order models and discusses odd-order evidence. General odd-order ensemble overlap gaps already appear in [5]. Sharp random-CSP results [6] require checking their even-predicate hypotheses. We claim none of these general ingredients as new.

Mints and Anschuetz [3, Section 4] use interpolating paths and Ramsey extraction for a different quantum hypergraph Hamiltonian. That methodology motivates our coupling architecture. Here an exact covariance at *different* path times permits an elementary extraction with only m^2 branches. This is a model-specific composition. The source comparison does not establish exhaustive worldwide priority.

2 Exact output couplings on an interpolation tree

Fix Q and deterministic sets

$$A_\ell = \{1, \dots, \lfloor \ell n/Q \rfloor\}, \quad R_\ell = [n] \setminus A_\ell, \quad S_\ell = A_\ell \setminus A_{\ell-1}.$$

All instances share one Poisson graph. Each occurrence receives independent signs $J_e^0, \dots, J_e^T$. At node (ℓ, t) use

$$J_e^{(\ell, t)} = \begin{cases} J_e^0 & E_e \subseteq R_\ell, \\ J_e^t & E_e \cap A_\ell \neq \emptyset. \end{cases} \quad (4)$$

Every node has law (3). At time zero all nodes coincide. Endpoint signs are independent *conditional on the common graph*. Each node's static Hamiltonian is used at every cost layer, and the angles are the same at all nodes.

Let $B_r(S)$ be the radius- r vertex neighborhood, counting one occurrence crossing as one hypergraph step.

Lemma 2.1. *The joint output marginals at $(\ell - 1, t)$ and (ℓ, t) coincide outside $W_\ell = B_{p+1}(S_\ell)$. There is a coupling whose two strings agree there.*

Proof. All cost gates in one layer commute. In backward Heisenberg evolution, gates disjoint from the observable's support before that layer can be moved past the other cost gates and canceled. Thus each cost layer enlarges support by at most one hypergraph step; the mixer does not enlarge support. The changed occurrences all meet S_ℓ . Any observable on the complement of $B_{p+1}(S_\ell)$ therefore has identical backward evolution in the two circuits. This gives equality of the entire outside reduced state, hence its joint Born marginal. Sample that common outside string, then the two inside strings according to their conditional laws. Arbitrary definitions on zero-probability conditioning events cause no problem. $\square$

Condition on the *entire* instance collection. Sample a common root output and, on each edge, disintegrate the preceding coupling given the parent output to sample the child. Induction on the tree preserves the correct Born marginal at every node and the required edge coupling. The child conditional on its parent need not have its unconditional Born law. No output is conditioned to have high energy, and no coupling around cycles is required.

3 Mixed-time landscape estimates

Take m nodes with distinct nonroot branch labels and at most one root; their times ℓ_s need not agree. For fixed strings $x^1, \dots, x^m$, put $D_{st} = d_H(x^s, x^t)/n$ and $W_E = \sum_s J_E^{(\ell_s, t_s)} \chi_E(x^s)$. Two nodes share a sign precisely when $E \subseteq R_{\max(\ell_s, \ell_t)}$. Hence

$$\mathbb{E}W_E^2 = m + \sum_{s \neq t} \left(\frac{1}{n} \sum_{i \in R_{\max(\ell_s, \ell_t)}} x_i^s x_i^t \right)^q. \quad (5)$$

Repeated tuple coordinates are included exactly. For every subset R ,

$$-D_{st} \leq n^{-1} \sum_{i \in R} x_i^s x_i^t \leq 1 - D_{st}.$$

If $a \leq D_{st} \leq b < 1/2$, the absolute value is at most $1 - a$. Thus, for either parity,

$$V_2 := \mathbb{E}W_E^2 \leq m[1 + (m - 1)(1 - a)^q]. \quad (6)$$

At distinct endpoints, $V_2 = m$ for arbitrary strings. Negative covariance at odd q is harmless.

The law of W_E is symmetric and $|W_E| \leq m$. The marked-Poisson exponential formula gives

$$\begin{aligned} \log \mathbb{E} e^{\lambda \sum_s H_{\ell_s, t_s}(x^s)} &= dn(\mathbb{E} e^{\lambda W_E / \sqrt{d}} - 1) \\ &\leq \frac{n\lambda^2 V_2}{2} + \frac{n\lambda^4 m^4}{24d} \cosh(\lambda m / \sqrt{d}). \end{aligned} \quad (7)$$

The inequality follows from the fourth-order remainder bound for $\cosh$. It includes common-graph fluctuations and is not a central-limit approximation.

Choose, in order,

$$\begin{aligned} m &\in \mathbb{N}, \quad m \geq \max\{2, 32 \log 2 / \alpha^2\}, \\ 0 < b < 1/4, \quad h(b) &\leq \alpha^2 / 64, \quad a = b / (4m), \quad \eta = a / 2, \\ (m-1)(1-a)^q &\leq 1/4, \quad d \geq \max\{1, \alpha^2 m^3\}, \end{aligned} \quad (8)$$

where $h(b) = -b \log b - (1-b) \log(1-b)$. With $\lambda = \alpha/2$, the last term of (7) is at most $nm\alpha^2/192$: indeed $\lambda m / \sqrt{d} \leq 1/(2\sqrt{m})$ and $\cosh(\lambda m / \sqrt{d}) \leq 2$. Since $V_2 \leq 5m/4$, Chernoff's inequality yields

$$\mathbb{P}\{H_{\ell_s, t_s}(x^s) \geq \alpha n \forall s\} \leq \exp(-65nm\alpha^2/192). \quad (9)$$

This applies to pair distances in $[a, b]$ and, at distinct endpoints, without a lower-distance condition.

The number of ordered m -tuples with pair distances at most bn is bounded by

$$2^n \left(\sum_{j \leq \lfloor bn \rfloor} \binom{n}{j} \right)^{m-1} \leq e^{n[\log 2 + (m-1)h(b)]} \leq e^{9nm\alpha^2/192}. \quad (10)$$

Counting only distances from the first string overcounts safely.

Lemma 3.1. *A specified admissible mixed-time node tuple admits m strings of energy at least αn with all pair distances in $[an, bn]$ with probability at most e^{-cn} , $c = 7m\alpha^2/24$. The same bound holds for a specified endpoint tuple with all pair distances at most bn . The probability of any such configuration on T branches and Q times is at most*

$$[(1 + TQ)^m + T^m] e^{-cn}. \quad (11)$$

Proof. Combine (9) and (10), noting $(65 - 9)/192 = 7/24$, and union bound over node tuples. This union occurs before any adaptive selection. $\square$

4 A deterministic mixed-time extraction

Lemma 4.1. *Let $T = m^2$ paths in a metric space have a common root and steps at most η . Suppose no m endpoints have all pair distances at most b . If $2a \leq b$ and $(m-1)(a+\eta) \leq b$, then the root and $m-1$ nodes on distinct branches can be selected with all pair distances in $(a, b]$.*

Proof. Start with the root as a set S of size $k = 1$. Maintain pair distances greater than a and diameter at most $(k-1)(a+\eta)$, using at most $k-1$ branches. If every unused endpoint belonged to the union of the k closed a -balls around S , then

$$T - (k-1) > k(m-1) \quad (k \leq m-1)$$

would put m endpoints in one ball by the pigeonhole principle. Their distances would be at most $2a \leq b$, a contradiction. Choose an unused branch ending outside the union. At its first exit, the new point is farther than a from every selected point and within $a+\eta$ of one of them. Its distance to any other selected point is at most $k(a+\eta)$. Add it and continue to $k = m$. Closed balls handle ties. The choices (8) give final diameter at most $3(m-1)b/(8m) < b$. $\square$

Different interpolation times are what make this economical extraction possible. Equation (5) allows the resulting tuple to be tested by Lemma 3.1.

5 Quantitative concentration and geometry

Constants in this section depend only on fixed q, d . Put $\mu = dq^2$ and $C_r = \sum_{j=0}^r \mu^j$.

Lemma 5.1. *There is $C \geq 2 + \mu$ such that radius- r occurrence neighborhoods, rooted at a Palm occurrence, have first and second moments at most C^{2r+4} . For deterministic S , $\mathbb{E}|B_r(S)| \leq |S|C_r$.*

Proof. An unexplored vertex exposes a Poisson number of unexposed incident occurrences with conditional mean at most dq . Each occurrence introduces at most q vertices. Ignoring collisions dominates exploration by a branching forest with offspring $Y = q \text{Pois}(dq)$. Conditioning on previously exposed absence reduces intensity. A Palm occurrence introduces at most q roots and leaves the remaining Poisson process unchanged.

For the single-root population through generation r , write $m_r = \mathbb{E}T_r$ and $s_r = \mathbb{E}T_r^2$. Then

$$m_r = 1 + \mu m_{r-1}, \quad s_r = 1 + 2\mu m_{r-1} + \mu s_{r-1} + \mathbb{E}[Y(Y-1)]m_{r-1}^2.$$

Since $m_{r-1}^2 \leq s_{r-1}$, a fixed C_0 gives $s_r \leq C_0 s_{r-1}$. Several roots cost their squared number by Cauchy–Schwarz. One additional exploration generation bounds occurrence counts. Enlarge C to absorb these fixed factors. The first-moment path count from S is $|S|C_r$. No maximum-degree bound is imposed. $\square$

Lemma 5.2. *Uniformly over deterministic angles, n and p ,*

$$\text{Var}[H(X_H)/n] \leq C^{4p+10}/n. \quad (12)$$

Proof. Conditional on the instance, occurrence observables with disjoint backward light cones have zero connected covariance, since the input is a product. All other covariances have absolute value at most 2. Overlapping cones put the occurrences within $2p$ steps, up to fixed endpoint slack. The Poisson Mecke identity [7] and Lemma 5.1 bound the expected number of such ordered pairs by dnC^{4p+8} . The d^{-1} Hamiltonian normalization gives

$$\mathbb{E} \text{Var}[H(X_H) \mid H] \leq nC^{4p+9}.$$

For disorder fluctuations, let $F(H) = \mathbb{E}[H(X_H) \mid H]$. Adding one marked occurrence contributes at most $d^{-1/2}$ directly and changes each old term by at most $2d^{-1/2}$. Only old terms whose backward cone meets the added occurrence can change. If K counts them,

$$|D_e F| \leq (1 + 2K)/\sqrt{d}, \quad \mathbb{E}(1 + 2K)^2 \leq C^{2p+8}.$$

The Poisson add-one Poincaré inequality [7] gives $\text{Var} F \leq \int \mathbb{E}(D_e F)^2 \nu(de) \leq nC^{2p+8}$, where ν has mass dn . Total variance and division by n^2 prove (12), after enlarging C . $\square$

Lemma 5.3. *For $n \geq Q$ and the slices in Section 2,*

$$\mathbb{P}\{\exists \ell : |B_r(S_\ell)| > 3C_r n/Q\} \leq K Q^3 C^{2r+8} / (C_r^2 n) \quad (13)$$

with fixed K .

Proof. Since $|S_\ell| \leq n/Q + 1$, its neighborhood mean is at most $2C_r n/Q$. Adding an occurrence changes this neighborhood only inside the radius- $(r+1)$ neighborhood of the added occurrence: every new path traverses it. The same Poisson Poincaré inequality gives variance at most nC^{2r+8} . Chebyshev with deviation $C_r n/Q$, then a union over Q slices, proves the claim. Slice neighborhoods need not be independent. $\square$

6 Completion of the proofs

Choose (8), then take

$$q_0 = \max \left\{ 3, \left\lceil \frac{\log(4(m-1))}{-\log(1-a)} \right\rceil \right\}, \quad d_0 = \max\{1, \alpha^2 m^3\}.$$

These do not depend on p . Fix $q \geq q_0, d \geq d_0$ and a constant C covering the preceding lemmas. Set

$$T = m^2, \quad Q = \lceil 12C_{p+1}/a \rceil + 1, \quad \kappa = (40 \log C)^{-1}.$$

For $p \leq \kappa \log n$, $C^p \leq n^{1/40}$ and $Q \leq KC^{p+2} \leq K'n^{1/40}$, so $n \geq Q$ eventually. Lemma 5.3 with $r = p + 1$ bounds geometric failure by

$$KQ^3 C^{2p+10}/n = O(n^{-7/8}).$$

On success the coupled normalized Hamming steps are at most $3C_{p+1}/Q < a/4 < \eta$.

Suppose the common mean node energy per spin exceeded $\alpha + \varepsilon_n$, with $\varepsilon_n = n^{-1/8}$. Correct marginal laws at all nodes, Lemma 5.2 and a union bound give

$$\mathbb{P}\{\text{some node energy} < \alpha n\} \leq \frac{(1 + TQ)C^{4p+10}}{n\varepsilon_n^2} \leq KC^{5p+12}n^{-3/4} = O(n^{-5/8}).$$

Meanwhile (11) tends to zero exponentially. With positive probability there would be high-energy paths with small steps, no close endpoint tuple, and no forbidden mixed-time tuple. Lemma 4.1 constructs exactly such a forbidden tuple. This contradiction proves Theorem 1.2. All estimates are uniform in the deterministic angle list; no simultaneous angle-net event is claimed.

The ordered tensor couplings corresponding to (3) are

$$J_{i_1 \dots i_q} = (P_{i_1 \dots i_q}^+ - P_{i_1 \dots i_q}^-)/\sqrt{d}, \quad P_{i_1 \dots i_q}^\pm \sim \text{Pois}(d/(2n^{q-1})).$$

Their scaled characteristic exponent is

$$n^{q-1} \log \mathbb{E} e^{i\lambda J_{i_1 \dots i_q}} = d[\cos(\lambda/\sqrt{d}) - 1] \longrightarrow -\lambda^2/2,$$

with derivative tending to $-\lambda$. These are exactly the sparse–dense hypotheses and normalization of [1, Theorem 2]; they do not require even q . At every fixed depth and fixed angle list, $V_{p,q}^G = \lim_{d \rightarrow \infty} V_{p,q,d}^S$. Every term for $d \geq d_0$ is at most α . Passing to the density limit and then taking the supremum over fixed angles proves Theorem 1.1. No supremum or depth limit has been exchanged with the thermodynamic limit.

7 Scope and verification record

The dense result does not include $q = 3$, instance-trained angles, arbitrary mixers or initial states, or depth growing with n . The sparse logarithmic-depth result does not give a dense logarithmic-depth result. The constants are conservative and are not a device-level performance prediction.

This paper consolidates M26 records from rounds 94, 95 and 98. The mixed-time argument replaces the earlier equal-time Ramsey extraction. Old records in the source package provide provenance, not independent evidence. The sequential audit checked normalization, repeated coordinates, the maximum-time covariance mask, marginal preservation, entropy constants and the order of limits. Independent examination of the whole manuscript and its originality remains outstanding.

Human and AI contributions. The author set the research directions, designed the harness coordinating model tasks, tools, memory and checking, supplied physical and mathematical intuitions and proof skeletons, and guided development and revision. AI systems assisted derivations, primary-source comparison, computational checks, and manuscript drafting and revision. The assembling agent is not an independent reviewer.

References

- [1] J. Basso, D. Gamarnik, S. Mei, and L. Zhou, *Performance and limitations of the QAOA at constant levels on large sparse hypergraphs and spin glass models*, [arXiv:2204.10306v2](#) (2022), Theorems 1–2.
- [2] C.-N. Chou, P. J. Love, J. S. Sandhu, and J. Shi, *Limitations of local quantum algorithms on random Max-k-XOR and beyond*, [arXiv:2108.06049v3](#) (2022).
- [3] M. Mints and E. R. Anschuetz, *The quantum overlap gap property and algorithmic hardness for the quantum hypergraph Max-Cut problem*, [arXiv:2609.10838v1](#) (2026).
- [4] M. Goh, *The overlap gap property limits limit swapping in the QAOA*, Quantum Information & Computation **25**, 329–343 (2025), [doi:10.2478/qic-2025-0018](#).
- [5] D. Gamarnik, A. Jagannath, and E. C. Kızıldağ, *Shattering in the Ising pure p -spin model*, [arXiv:2307.07461](#) (2023).
- [6] A. Chen, N. Huang, and K. Marwaha, *Local algorithms and the failure of log-depth quantum advantage on sparse random CSPs*, [arXiv:2310.01563](#) (2023).
- [7] G. Last and M. Penrose, *Lectures on the Poisson Process*, Cambridge University Press (2017), the Mecke identity and Poisson Poincaré inequality.