Resolving Goldreich's Open Problem 4.1: Linear Lower Bounds for Cubic Graph Isomorphism Testing

Scott Schubr
Luminary Oddities

Abstract

Testing whether an unknown bounded-degree graph is isomorphic to a fully known target can require reading a linear fraction of its incidence representation. At the literal proximity 0.01, we prove this for a connected simple asymmetric bridgeless cubic target at every sufficiently large even order, even when all inputs are promised to have a common positive edge-expansion bound and to be asymmetric. The linear lower bound holds both for randomized classical testers and for bounded-error quantum testers using coherent incidence queries. Reading all incidences gives the matching linear upper bound. Thus the classical result answers Goldreich's Open Problem 4.1 negatively at its specified degree and proximity. The central distance estimate controls every uncolored vertex relabeling of the cubic construction. A joint occupancy bound and median rounding turn edge disagreement into constraint violations. A simultaneous expansion argument and bounded surgeries preserve the strong input promise at all even orders. For quantum queries, affine dependence of each oracle matrix entry on one hidden bit gives an exact low-degree indistinguishability argument. The same construction covers all sufficiently large integer orders with at most one isolate. Despite rigidity, linearly many vertex pairs have identical rooted neighborhoods to radius $\log_2 N - O(1)$. Distance counts unordered edge changes, and target preprocessing and computation are unrestricted.

Keywords. graph property testing; graph isomorphism; cubic graphs; edge expansion; quantum query complexity.
2020 Mathematics Subject Classification. Primary 68Q25; Secondary 68R10, 81P68.

1 Introduction

In fixed-target graph-isomorphism testing, the reference graph is known in full and may be preprocessed without cost. Only access to the unknown input is restricted: a query asks for a specified neighbor of a specified vertex. The question is whether a small number of these local observations can distinguish a relabeling of the target from every graph far from all such relabelings.

Goldreich's Open Problem 4.1 [9] asks whether every fixed 3-regular target admits a sublinear two-sided tester at proximity 0.01. We prove a negative answer, with targets and inputs that are connected, simple, asymmetric bridgeless cubic expanders. We also prove linear hardness for the coherent quantum version of this incidence oracle. The targets are selected existentially; the complexity measure counts queries with unrestricted computation and target preprocessing.

Dror [7] established a linear lower bound at maximum degree thirteen and a sufficiently small constant proximity. His Remark 1.3 asks for the smallest possible degree, and Remark 3.8 already provides connected hard graphs. We use his sparse-linear-system and adaptive-query framework. The additional quantitative burden is to encode it in cubic graphs and retain enough distance

under *arbitrary uncolored bijections* to reach the specified one-percent threshold. The estimate proved below is

$$\operatorname{dist}_H(b, \operatorname{Im} A) \le \frac{2}{3} \operatorname{dist}_{\mathrm{iso}}(G_b, G_0), \qquad |V(G_b)| = 16n.$$

A joint source-row and target-occupancy budget pays for images that mix the apparent roles of the vertices. Median rounding then completes the reverse-distance estimate. Global expansion is proved simultaneously for the sparse system and transferred to every encoded word, so bottlenecks are not needed for the hard input distribution.

Related methods and the exact distinction. Parity constructions have CFI ancestry [5]; colored code-distance reductions appear in Arvind et al. [3]. Partial recovery under edge defects is studied in [14], and intrinsic recognition of uncolored CFI components in [6]. Occurrence-cycle cubic reductions, including local-oracle and distance transfers, appear in Adler–Köhler–Peng [1, Section 5], for a tested property with many isomorphism classes. Conditional approximate-isomorphism gaps in Makarychev–Manokaran–Sviridenko [12] have a different objective and completeness requirement. The result here combines exact completeness, a fixed target, cubic incidence access, arbitrary uncolored distance, and the literal numerical proximity.

The quantum argument uses the polynomial method of Beals et al. [4]. Limited independence already yields quantum graph-property-testing lower bounds: Apers et al. [2, Section 5] apply sparse-parity hardness to bounded-degree 3-colorability testing. We apply the method to the same fixed-target cubic isomorphism experiment, with its precise coherent multibit oracle and expansion promise.

Goldreich–Tauber's almost-all-target results [10] are compatible with an exceptional hard family. Their locating algorithms use local vertex-identification hypotheses in addition to a diameter bound; expansion by itself is not that hypothesis. Our targets are asymmetric, yet linearly many disjoint vertex pairs have identical rooted balls out to radius $\log_2 N - O(1)$. Full column rank excludes global assignment flips, while small row spans still permit local ones. We combine this elementary rank argument with the specified-degree bipartite asymmetry theorem of McKay–Wormald [13]. Canonical inverse parsing preserves rigidity under the all-even surgeries; bounded edit distance alone would not suffice. At maximum degree two, isomorphism testing has query bounds depending only on proximity [11], so degree three is the minimum possible degree threshold for the present linear hardness.

2 Model and main results

All graphs are finite, simple and undirected, with vertex set $[N]$. For graphs G, H on that set, define

$$\operatorname{dist}_{\mathrm{iso}}(G, H) = \min_{\phi:[N]\to[N] \text{ bijective}} |E(G) \triangle \phi^{-1}(E(H))|.$$

Deleted and inserted unordered edges both count. At degree bound three, an input is ε-far from the target's isomorphism class if $\operatorname{dist}_{\mathrm{iso}}(G, H) > \varepsilon(3N/2)$.

An incidence representation is a function $f : [N] \times [3] \to \{0, \dots, N\}$. The nonzero answers in a vertex's three slots list its distinct neighbors once each; zero denotes an empty slot, and lists are reciprocal. Slot ordering is arbitrary. A tester accepts every representation of an isomorphic promised input with probability at least $2/3$, and rejects every far promised input with probability at least $2/3$. Queries are bounded in the worst case. The target is fully specified and free to preprocess.

A classical query reads $f(v, s)$. For a quantum query, let $L = \lceil \log_2(N+1) \rceil$, encode answers in L bits, and use

$$O_f|v, s, z, w\rangle = |v, s, z \operatorname{XOR} \operatorname{enc}(f(v, s)), w\rangle. \tag{1}$$

Here w is arbitrary workspace. Set $f = 0$ on invalid binary address encodings, so this defines a unitary on the full register space; it is its own inverse. A quantum tester has input-independent, possibly target-dependent computation and advice, and the same bounded error requirements. Since correctness is required for every incidence representation, a lower bound may use particular fixed labels and slots.

For a graph G, put

$$h(G) = \min_{0<|S|\le|V(G)|/2} \frac{|\partial_G S|}{|S|},$$

where $\partial_G S$ is its set of unordered edges crossing the cut. We call this the edge Cheeger constant. A graph is cubic if it is 3-regular, bridgeless if it has no cut edge, and asymmetric (or rigid) if its only automorphism is the identity.

Theorem 2.1. *There are absolute constants $h_* > 0$ and $c > 0$ such that, for every sufficiently large even N, there exists a fully known connected simple asymmetric bridgeless cubic N-vertex target J_N with $h(J_N) \ge h_*$ for which every randomized classical or bounded-error quantum $(1/100)$-tester needs at least cN incidence queries. The lower bound holds even under the promise that the unknown input is connected, simple, asymmetric, bridgeless, cubic and has edge Cheeger constant at least h_*. Computation and target preprocessing are unrestricted.*

Reading all $3N$ slots gives a matching $O(N)$ upper bound with unrestricted computation, so the worst-case query order is $\Theta(N)$ in both models. The classical assertion gives the negative answer to Goldreich's precise question. Cubic orders are necessarily even.

Corollary 2.2. *For every sufficiently large integer N, there is a fully known maximum-degree-three N-vertex target for which randomized classical and bounded-error quantum two-sided $(1/100)$-testing requires $\Omega(N)$ incidence queries. The target and hard inputs each consist of a connected simple asymmetric bridgeless cubic expander and at most one isolate, with no isolate when N is even. These graphs are themselves asymmetric.*

Proof structure. A binary linear system gives a codeword experiment and a uniform-word experiment with identical low-order marginals. We define fixed incidence representations for cubic graphs, prove exact isomorphisms for codewords, and derive the reverse-distance inequality for every uncolored bijection, including isolate mixing. Expansion is verified uniformly over the entire word family. Bounded equivariant surgeries handle all even residues; the distance estimate is applied before surgery, while canonical recovery separately proves rigidity. The classical transcript argument and the quantum polynomial argument then use the same distance and promise guarantees.

3 Sparse linear systems and exact indistinguishability

Lemma 3.1. *For some $\delta > 0$ and $h_B > 0$, every sufficiently large n divisible by three admits $A \in \mathbb{F}_2^{n\times(n/3)}$ of row weight three and column weight nine, of full column rank, with asymmetric bipartite incidence support of edge Cheeger constant at least h_B and, for every nonempty row set S of size at most $\lfloor \delta n \rfloor$, $|N(S)| > 7|S|/4$. Every such row set is independent, and its row span contains no coordinate vector of $\mathbb{F}_2^{n/3}$.*

Its complete configuration-model proof, including the simultaneous global expansion, rank and asymmetry properties, appears in Appendix A. Fix such a matrix, put $m = n/3$ and $C = \operatorname{Im} A$, and order its row coordinates. Write (v_{i1}, v_{i2}, v_{i3}) for the three distinct variables of row i, with this order independent of b. All bit additions are in $\mathbb{F}_2$. Define

$$\operatorname{dist}_H(b, C) = \min_{x \in \mathbb{F}_2^m} |\{i : (Ax)_i \neq b_i\}|.$$

Order each variable's nine occurrences by positions $j \in \mathbb{Z}/9\mathbb{Z}$, independently of b.

Lemma 3.2 (Adaptive indistinguishability, following Dror)**.** *Algorithms making at most $\lfloor \delta n \rfloor$ coordinate queries have identical transcript laws on Ax, for uniform $x \in \mathbb{F}_2^m$, and on uniform $b \in \mathbb{F}_2^n$. Moreover, for all sufficiently large n, uniform b has $\operatorname{dist}_H(b, C) > n/6$ with probability at least $3/4$.*

Proof. Fix the algorithm's random coins and a realized decision-tree path. Every collection of at most $\lfloor \delta n \rfloor$ distinct queried rows is independent. Its coordinates of Ax are therefore uniform independent bits. A newly queried coordinate remains unbiased given the preceding answers; a repeated answer is determined. Induction on the path length proves the transcript assertion, including adaptive queries.

Since $|C| \leq 2^{n/3}$, the entropy bound for Hamming balls gives

$$\Pr[\operatorname{dist}_H(b, C) \leq n/6] \leq 2^{-n(2/3 - H_2(1/6))}.$$

Here H_2 is binary entropy. The exponent is positive, since $H_2(1/6) < 2/3$ is equivalent to $6^6 < 2^4 5^5$, or $46656 < 50000$. The probability is at most $1/4$ for all sufficiently large n. No full-column-rank assumption is needed. □

4 The connected cubic encoding

For every variable v, create vertices $z(v, a, j)$ with $a \in \mathbb{F}_2$ and $j \in \mathbb{Z}/9\mathbb{Z}$, and join each fixed-bit set in a 9-cycle. For every row i, create outer vertices $o(i, k, a)$ with $k \in [3]$ and $a \in \mathbb{F}_2$, and middle vertices $q(i, u, w)$ with $u, w \in \mathbb{F}_2$. Put

$$t(i, u, w) = (u, w, b_i + u + w),$$

and join $q(i, u, w)$ to $o(i, k, t_k(i, u, w))$ for each k. The internal row graph is a once-subdivided K_4: the four parity-b_i tuples are its branch vertices, and each outer vertex subdivides the edge between the two tuples sharing its coordinate and value. Indeed, two distinct tuples of the same parity differ in exactly two coordinates, so they share exactly one coordinate and value. These are the six unordered pairs of the four tuples. For a parity-zero row the middle tuples are

(u, w)	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
(t_1, t_2, t_3)	000	011	101	110.

This description distinguishes the three pairs of outer vertices by their disjoint middle-neighbor sets.

If row i uses $v = v_{ik}$ at its occurrence position j, join $o(i, k, a)$ to $z(v, a, j)$ for both bits a. Call the resulting graph G_b. It has

$$|V(G_b)| = 18m + 10n = 16n, \qquad |E(G_b)| = 24n.$$

Every vertex has degree three. Distinct variables in a row and distinct occurrence positions ensure simplicity. Each connected row gadget ties both bit cycles of every incident variable; connected incidence support therefore makes G_b connected. Variable, outer and middle roles are construction descriptions, not colors supplied to the tester or bijection.

Lemma 4.1 (Completeness)**.** *For every b, x, $G_b \cong G_{b+Ax}$. In particular, if $b \in C$, then $G_b \cong G_0$.*

Proof. Use the bijection

$$z(v, a, j) \mapsto z(v, a + x_v, j),$$
$$o(i, k, a) \mapsto o(i, k, a + x_{v_{ik}}),$$
$$q(i, u, w) \mapsto q(i, u + x_{v_{i1}}, w + x_{v_{i2}}).$$

Its third target bit is $(b_i + (Ax)_i) + (u + x_{v_{i1}}) + (w + x_{v_{i2}}) = b_i + u + w + x_{v_{i3}}$. All cycle, row and external edges are consequently preserved. □

Lemma 4.2 (One-bit simulation)**.** *A fixed labeling and fixed incidence slots have the following property: for every query address a, the answer is constant or is a fixed function $F_a(b_{i(a)})$ of a single coordinate, with both $i(a)$ and F_a independent of b. In particular each classical answer reads at most one coordinate of b.*

Proof. At a variable vertex, slots one and two return predecessor and successor, and slot three its external outer neighbor, all independently of b. At an outer vertex, slot one returns its variable neighbor. The other two slots enumerate middle neighbors: for $k = 1$ fix $u = a$ and enumerate w; for $k = 2$ fix $w = a$ and enumerate u; for $k = 3$ enumerate u and set $w = b_i + a + u$. At a middle vertex, slot k returns $o(i, k, t_k(i, u, w))$. Only row bit b_i can be needed. The lists are reciprocal and have no repeated neighbors; all vertex and slot names are fixed before b is known. □

Lemma 4.3 (Bridgelessness)**.** *Every graph G_b is bridgeless.*

Proof. Each internal row edge lies on a cycle of the subdivided K_4, and each variable-cycle edge lies on its designated 9-cycle. Consider an external edge at an occurrence of variable v in row i. Choose a second occurrence of v, in a distinct row i'. Inside each of these two row gadgets, join the opposite-bit outer vertices for v by a simple path. In each bit cycle of v, join the two occurrence positions by a simple path. These four paths, together with the four external edges at the two occurrences, form a simple cycle containing the chosen edge: the row gadgets are distinct, the two bit cycles are disjoint, and the four paths have disjoint interiors. Thus every edge lies on a cycle, proving the claim. □

5 Distance under arbitrary uncolored bijections

Recognizing uncolored CFI components using short cycles has explicit ancestry in Chen–Flum–Liu [6, Lemma 7.12]. Here we need the following exact fact.

Lemma 5.1. *The only simple 9-cycles of G_b are its designated variable cycles. Every simple 6- or 8-cycle lies in one row gadget.*

Proof. The internal variable graphs are 9-cycles and the internal row graphs are once-subdivided K_4's, whose cycles have lengths six or eight. An external cycle uses an even number of external edges. With exactly two, its internal pieces would lie in one row and one variable. A row uses a variable once, so these endpoints are the opposite values of one coordinate; their variable vertices belong to disconnected bit cycles. This is impossible. With $2k \geq 4$ external edges, the k internal row runs have length at least two, and the k internal variable runs length at least one. These bounds hold even on revisiting gadgets: outer vertices are nonadjacent and each socket has just one external edge. Its length is at least $5k \geq 10$. □

5.1 Row rigidity and occupancy

Fix any vertex bijection $\phi : G_b \to G_0$. Let K be its unordered-edge symmetric difference. A source edge is unmatched if its image is absent. Write $q_{\mathrm{cyc}}, q_{\mathrm{row}}, q_{\mathrm{ext}}$ for the unmatched counts by source cycle, internal-row and external roles, without assumptions on images. Equal edge cardinalities give

$$q_{\mathrm{cyc}} + q_{\mathrm{row}} + q_{\mathrm{ext}} = K/2. \tag{2}$$

A source row is intact if all twelve internal edges match. The following consequence uses no assumption that the full bijection preserves vertex roles.

Lemma 5.2 (Intact rows)**.** *An intact source row fills one whole target row by an internal isomorphism. Its three outer pairs are permuted, with value flips $y(i,k)$ satisfying*

$$\sum_{k=1}^{3} y(i,k) = b_i. \tag{3}$$

Proof. The four subdivided triangles of the source K_4 are overlapping 6-cycles. Their images are target-row cycles by Lemma 5.1; overlaps force all four into the same target row, since distinct rows have disjoint vertex sets. Their ten-vertex union fills that row. The twelve matched edges then exhaust its twelve internal edges, giving an internal isomorphism. Within this internal graph, degrees distinguish middles from outers. Two outer vertices have disjoint middle-neighbor sets exactly when they form one of the three opposite pairs. These pairs are therefore permuted. Write $y(i,k)$ for the bit flip on the source coordinate pair k under that permutation. Every source middle tuple has parity b_i, while its target middle tuple has parity zero. Permuting coordinates does not change parity, so the sum of the three flips is b_i. □

Lemma 5.3. *A nonintact source row has at least two unmatched internal edges. If it has exactly two, it retains a matched* 6- *or* 8*-cycle.*

Proof. Here a row leg means one of the two actual edges replacing an underlying K_4 edge. Suppose exactly one row leg is unmatched, over underlying K_4 edge uv. Write c, d for the other branch vertices. The two triangles avoiding uv give intact 6-cycles sharing the subdivided cd path. Their images are distinct 6-cycles in one target row. Their union contains nine vertices: all four middle vertices and five outers. The shared path is the common subdivided edge of two target triangles, identifying its middle endpoints. Each of the other two paths between those endpoints has four edges, whose central vertex identifies one of the other middles. Thus every source middle maps to a target middle. The omitted outer over uv has one still-matched leg. Its image must be the remaining target outer: it neighbors a target middle and the other nine vertices are occupied. Its other middle neighbor is then precisely the image of the other endpoint of uv, contradicting the single unmatched edge. Here the complementary edge of a subdivided K_4 is determined by the five already identified edges.

With exactly two unmatched row legs, remove their at most two underlying K_4 edges. Two adjacent deletions leave a triangle; two opposite deletions leave a 4-cycle. Fewer underlying deletions also leave a triangle. The corresponding subdivided cycle has all edges matched. □

Let B count nonintact source rows and a those with exactly two defects. Their disjoint internal edge sets give

$$q_{\mathrm{row}} \geq 2a + 3(B - a) = 3B - a. \tag{4}$$

The $n - B$ intact rows fill distinct whole target rows. Call the remaining B target rows unoccupied. Every target-row image of a source variable vertex lies in an unoccupied row, by injectivity. For

each of the a two-defect rows choose the cycle from Lemma 5.3. Its image is in an unoccupied target row. These a target rows are distinct: source cycles from distinct rows are vertex-disjoint, while two disjoint cycles of length at least six cannot fit in one ten-vertex row. Each occupies at least three target middles. We use only this cycle occupancy, not a type-preserving map on a defective row.

5.2 Adaptive middle colors and median rounding

Proposition 5.4. *For every b, $\operatorname{dist}_H(b, C) \le (2/3)\operatorname{dist}_{\mathrm{iso}}(G_b, G_0)$. The same inequality holds after adding any equal number of isolates.*

Proof. For the arbitrary ϕ above, color target variable vertices by their bit, and target outer vertices by their value. Call this binary color χ. It agrees at both ends of every target variable-cycle or external edge. Let P consist of images of matched source cycle edges. Its maximum degree is two. At each target middle choose its color to minimize disagreements on its incident P-edges. Their other endpoints are target outers with fixed colors. With two differing neighbor bits exactly one disagreement is unavoidable; with one neighbor or two equal bits choose that bit, and with no neighbor choose zero. Let c count unavoidable disagreements.

The target parity-zero row has middle tuple 000, which contributes zero, and three other middles, each contributing at most one. An occupied target row contains no source-variable image and hence contributes zero. Each of the a cycle-anchored unoccupied rows has at least three middles occupied by source-row images, so contributes at most one. Therefore the complete accounting is

target-row kind	number	maximum contribution to c
filled by an intact source row	$n - B$	0
unoccupied, with a two-defect cycle anchor	a	1
other unoccupied row	$B - a$	3

The middle vertices occupied by source-row images cannot be endpoints of an image of a source variable-cycle edge. Combining this table with the source defect bound (4) gives

$$c \le 3B - 2a, \qquad c + B \le 4B - 2a \le \tfrac{4}{3}(3B - a) \le \tfrac{4}{3} q_{\mathrm{row}}.$$

This coloring may depend on ϕ; it is a decoding existence argument, not oracle advice. No matched-external-edge or full-star condition is imposed on the imported middle vertices.

Along all source variable cycles the total color variation is at most $q_{\mathrm{cyc}} + c$: unmatched cycle edges cost at most one, and matched ones change color only on the c selected target-middle cuts. Define

$$\alpha(v, a, j) = a + \chi(\phi(z(v, a, j))) \pmod 2, \qquad s(v, j) = \frac{\alpha(v, 0, j) + \alpha(v, 1, j)}{2}.$$

Let t_v be a median of the nine real values $s(v, j)$, and choose x_v independently with $\Pr[x_v = 1] = t_v$. For nine real cyclic values, their median t satisfies

$$\sum_j |s_j - t| \le 4(\max s_j - \min s_j) \le 2 \sum_j |s_j - s_{j+1}|.$$

Pair the four sorted values below the median with the four above it; each pair contributes at most the range, giving the first inequality; cyclic variation is at least twice the range. Also

$$\sum_j |s(v, j) - s(v, j+1)| \le \tfrac{1}{2} \sum_{a,j} |\alpha(v, a, j) - \alpha(v, a, j+1)|.$$

The added bit a is constant along each cycle. Summing gives

$$\sum_{v,j} |s(v,j) - t_v| \le q_{\mathrm{cyc}} + c. \tag{5}$$

For an intact row, its flips satisfy (3); violation implies $x_{v_{ik}} \neq y(i,k)$ for some coordinate. Its probability is at most $\sum_k |t_{v_{ik}} - y(i,k)|$. For its occurrence (v,j) and bit $y = y(i,k)$,

$$|t_v - y| \le |t_v - s(v,j)| + |s(v,j) - y|, \qquad |s(v,j) - y| = \tfrac{1}{2} \sum_a |\alpha(v,a,j) - y|.$$

If the source external edge at bit a matches, then $\alpha(v,a,j) = y$. Indeed its outer image has bit $a + y$. If its other image endpoint is a target variable, external adjacency forces that bit. If it is a row vertex, it cannot belong to the already fully occupied intact target row; distinct target rows have no connecting edges, so such an edge cannot match. Thus every failed bit equality costs a distinct unmatched source external edge. Summed over intact rows, the $|s - y|$ terms cost at most $q_{\mathrm{ext}}/2$; occurrence positions are used once each in (5). The B nonintact rows cost at most B. Some assignment consequently has

$$\mathrm{dist}_H(b, C) \le q_{\mathrm{cyc}} + c + q_{\mathrm{ext}}/2 + B \le q_{\mathrm{cyc}} + \tfrac{4}{3} q_{\mathrm{row}} + q_{\mathrm{ext}}/2 \le \tfrac{4}{3}(q_{\mathrm{cyc}} + q_{\mathrm{row}} + q_{\mathrm{ext}}) = \tfrac{2}{3} K.$$

This holds for every ϕ, with no restriction on roles or occurrence images.

For equal added isolates, give target isolates color zero. Edge counts and source roles in (2) are unchanged. Intact row cycles cannot map to isolates, so the occupied-row counts remain valid. A middle whose preimage is a source isolate has no incident P-edge and contributes nothing to c. Matched source edges cannot land at an isolate; a variable endpoint mapped there has an unmatched external edge. All variation, external-error and median estimates therefore remain valid, even when ϕ mixes isolates with nonisolated vertices. □

6 Expansion for every encoded word

The incidence support in Lemma 3.1 has a common positive edge Cheeger constant. We transfer it without treating a variable's two disconnected bit cycles as one connected gadget.

Lemma 6.1 (The twin-variable macrograph)**.** *Replace each variable vertex v of the bipartite incidence support by two vertices $(v,0), (v,1)$, both adjacent to every row formerly incident to v. Call this macrograph M. Then*

$$h(M) \ge \frac{1}{2/h_B + 1/9}.$$

Proof. For a cut S of M, let F count variable pairs split between the two sides and let $K_M = |\partial_M S|$. Every split pair contributes exactly nine crossing edges, so $9F \le K_M$. Coarsen to a cut T of the original support by retaining each row's side, assigning an unsplit variable its common side and assigning a split variable either side. Unsplit-variable boundary edges occur twice in M, while split variables can contribute at most $9F$ coarse boundary edges. Thus

$$|\partial T| \le (K_M - 9F)/2 + 9F \le K_M.$$

Lift T back by placing both variable copies on their coarse side. The lift differs from S in exactly F vertices, and each coarse vertex produces at most two vertices. Consequently

$$\min\{|S|, |S^c|\} \le 2 \min\{|T|, |T^c|\} + F \le (2/h_B + 1/9) K_M.$$

This holds for every cut. □

Lemma 6.2 (Connected-gadget transfer)**.** *Suppose each vertex of a macrograph of edge Cheeger constant $h_M > 0$ and maximum degree at most $\Delta \geq 1$ is replaced by a connected gadget of at most L vertices, with one edge for each macroedge between its prescribed ports. The resulting graph has edge Cheeger constant at least $1/(L\Delta/h_M + L)$.*

Proof. For a cut of boundary size $K = K_{\text{int}} + K_{\text{ext}}$, let B count gadgets meeting both sides. Connectedness gives $B \leq K_{\text{int}}$. Label each unmixed gadget by its side and each mixed gadget arbitrarily. The induced macrocut T has at most $K_{\text{ext}} + \Delta B \leq \Delta K$ crossing edges. The actual cut differs from the selected union of whole gadgets in at most LB vertices. Hence its smaller side has size at most

$$L \min\{|T|, |T^c|\} + LB \leq (L\Delta/h_M + L)K.$$

Empty or full macrocuts are covered by the same inequality, with their minimum side size equal to zero. □

Each row gadget is a connected ten-vertex subdivided K_4, and each individual bit cycle is a connected nine-vertex gadget. They replace exactly the vertices of M, whose degrees are six at rows and nine at bit variables. Lemmas 6.1–6.2 give, uniformly for every word b,

$$h(G_b) \geq h_G := \frac{1}{180/h_B + 20} > 0. \tag{6}$$

Changing row parity does not affect this estimate.

We will also use the following bounded-edit consequence.

Lemma 6.3 (Padding and bounded surgery)**.** *Suppose G has edge Cheeger constant at least h_G, and a connected graph P on $N = |V(G)| + r$ vertices differs by at most 161 unordered-edge changes from $G \sqcup rK_1$, under a named labeling, where $0 \leq r \leq 46$. Then*

$$h(P) \geq h_* := \min\left\{\frac{h_G}{2}, \frac{1}{93 + 322/h_G}\right\} > 0. \tag{7}$$

Proof. For a cut with smaller side size s, both sides contain at least $s - r$ base vertices. A changed edge alters the boundary count by at most one, so

$$|\partial_P S| \geq h_G(s - r) - 161.$$

If $s \geq 2r + 322/h_G$, this is at least $h_G s/2$. Otherwise connectedness gives at least one crossing edge. All cuts with $s \geq 92 + 322/h_G$ are in the first case, and below that threshold the ratio is larger than $1/(93 + 322/h_G)$. □

7 Proof of the cubic theorem at proximity 0.01

For even N, choose the largest positive multiple of three n with $16n \leq N$. Then $r = N - 16n \in \{0, 2, \ldots, 46\}$. Put $\eta = 0$ when $r \equiv 0 \pmod 4$, and $\eta = 2$ otherwise. If $\eta = 2$, subdivide the edge between positions zero and one on both bit cycles of one fixed variable and join the two new vertices. This adds two vertices and preserves simplicity, cubicity and connectedness. A bit flip exchanges the two new vertices. Relative to the old graph plus two isolates it costs at most seven edits: delete two edges, insert four path edges and one joining edge.

Use $(r - \eta)/4 \leq 11$ paired triangle expansions (a standard cubic operation; see [8]) at original variable sites, avoiding those surgery endpoints. There are $9m = 3n$ paired positions, enough for

large n. Replace each selected vertex by a triangle, with ports fixed by predecessor, successor and external roles. Adjacent selected vertices are allowed: their edge attaches to the appropriate port at both ends. Each pair adds four vertices; each individual expansion costs at most seven edits relative to the old vertex and two isolates, by retaining one incidence, rerouting two, and adding three triangle edges. This bound also follows by sequential expansion when sites are adjacent.

The resulting P_b is simple, connected, cubic and has exactly N vertices. Bridgelessness also survives these operations. Subdividing the two chosen cycle edges preserves cycles through every old edge; deleting the new joining edge leaves a connected subdivided copy of G_b, so that edge also lies on a cycle. In a triangle expansion, each old cycle extends through the appropriate triangle ports and every new internal edge lies on the triangle. Sequential expansion covers adjacent sites. Lemma 4.3 therefore shows that every P_b is bridgeless.

All choices are independent of b and paired bit flips extend to full isomorphisms. New variable-site incidences require no bit query; row answers only translate the returned variable label to its fixed port. At triangle vertices two slots return triangle neighbors and the third its assigned old incidence. At a subdivision vertex the three neighbors are the old endpoints and the opposite subdivision vertex. Thus fixed names and slots still simulate each graph answer with at most one coordinate query. On each side the edit cost relative to $G_b \sqcup rK_1$ is at most $7 + 11 \cdot 14 = 161$. By the triangle inequality,

$$\mathrm{dist}_{\mathrm{iso}}(G_b \sqcup rK_1, G_0 \sqcup rK_1) \le \mathrm{dist}_{\mathrm{iso}}(P_b, P_0) + 322. \tag{8}$$

The distance argument uses short-cycle recognition only before these surgeries. Section 10 separately proves that canonical surgery recovery preserves asymmetry for every word. The graphs are connected, so Lemma 6.3 also gives $h(P_b) \ge h_*$ for every b. The fixed incidence formulas after surgery retain the stronger single-coordinate property of Lemma 4.2: a row address still uses only its own row bit, and translation to a surgery port is independent of b.

By Lemma 3.2, uniform b has code distance greater than $n/6$ with probability at least $3/4$ for all sufficiently large n. Proposition 5.4, including its isolate extension, and (8) give $\mathrm{dist}_{\mathrm{iso}}(P_b, P_0) > n/4 - 322$ for every such word. The literal $1/100$ far threshold obeys

$$\frac{1}{100}\frac{3N}{2} \le \frac{6n}{25} + \frac{69}{100}, \qquad \frac{n}{4} - 322 - \left(\frac{6n}{25} + \frac{69}{100}\right) = \frac{n - 32269}{100} > 0$$

for $n > 32269$. Combine this arithmetic threshold with the asymptotic matrix and entropy thresholds.

Simulating a tester with at most $\lfloor \delta n \rfloor$ graph queries uses at most that many coordinate queries. Codewords yield isomorphic connected cubic graphs and acceptance at least $2/3$. Uniform words give far connected cubic graphs with probability at least $3/4$, so acceptance is at most $(3/4)(1/3) + 1/4 = 1/2$. Identical adaptive transcript laws contradict this. Since $n = \Theta(N)$, this proves the classical assertion of Theorem 2.1 at literal proximity 0.01, with the stated expansion, asymmetry and bridgelessness promises. The quantum assertion is proved next using the same graph family.

8 Coherent quantum incidence queries

We prove the quantum assertion directly for the multibit oracle (1). This also specifies why the classical one-bit simulation has the needed coherent meaning.

Lemma 8.1 (Acceptance polynomial)**.** *For the fixed representations f_b of P_b, the acceptance probability of a T-query quantum algorithm is a real multilinear polynomial in $b_1, \dots, b_n$ of degree at most $2T$.*

Proof. At a fixed address a, Lemma 4.2 and the fixed surgery translation give a constant answer or an answer $F_a(b_{i(a)})$. The corresponding oracle block is

$$(1 - b_{i(a)})X_{\mathrm{enc}(F_a(0))} + b_{i(a)}X_{\mathrm{enc}(F_a(1))},$$

where X_t means XOR by t on the answer register. Constant and invalid-address blocks are fixed permutations. Every matrix entry of the entire oracle is therefore affine in the hidden bits, even though different address blocks may use different coordinates.

Purify input-independent randomness and mixed advice, and defer intermediate measurements. Branches that stop early can address a fixed invalid address during their remaining steps. Initially every amplitude has degree zero in b. An input-independent operation preserves degree, and an oracle matrix with affine entries raises it by at most one. After T queries every amplitude is a complex polynomial of degree at most T. Squared moduli and the final accepting measurement give a real polynomial of degree at most $2T$. On the Boolean cube replace b_i^r by b_i for $r \geq 1$; this makes it multilinear without increasing degree.

This is the polynomial method of [4], with the affine answer-register step made explicit. Target-dependent preprocessing or advice is input independent and changes only the degree-zero initial data and between-query operations. □

If $2T \leq \lfloor \delta n \rfloor$, every monomial of this polynomial has the same expectation on Ax, for uniform x, and on uniform b. Indeed, every set of at most $\lfloor \delta n \rfloor$ coordinates is uniform by row independence, so a monomial on S has expectation $2^{-|S|}$ in both experiments. Thus the two average acceptance probabilities are equal. This is an equality of expectations, requiring no assertion that complete quantum states or classical transcripts agree.

All codeword graphs are isomorphic to P_0, so their average acceptance is at least $2/3$. At least $3/4$ of uniform words give far inputs, and every word gives a promised expanding bridgeless cubic graph. Their average acceptance is therefore at most $1/2$, exactly as in the classical proof. This contradiction gives

$$T > \tfrac{1}{2}\lfloor \delta n \rfloor = \Omega(N),$$

completing Theorem 2.1. The mechanism of matching limited independence to quantum acceptance degree is established; the graph testing application in [2, Section 5] is a close antecedent. Here it applies to the exact fixed-target cubic experiment and its literal one-percent metric.

9 All integer orders

Proof of Corollary 2.2. For a sufficiently large integer N, put

$$n = 3\lfloor N/48 \rfloor, \quad r = N - 16n \in \{0, \ldots, 47\}, \quad h = N \bmod 2.$$

Apply the even-order surgeries to $N - h = 16n + (r - h)$, where $r - h \in \{0, 2, \ldots, 46\}$, and add h isolates. Call the resulting graph R_b. Its nonisolated component is connected, simple, bridgeless, cubic and has the same positive edge-expansion bound. By Lemma 10.2 this component is rigid, and the at most one isolated vertex is uniquely isolated, so the whole graph is asymmetric. Each side differs from $G_b \sqcup rK_1$ in at most 161 edits. The isolate version of Proposition 5.4 therefore gives, on the same far-word event,

$$\mathrm{dist}_{\mathrm{iso}}(R_b, R_0) > n/4 - 322.$$

This exceeds $3N/200$ eventually, since

$$\frac{n}{4} - 322 - \left(\frac{6n}{25} + \frac{141}{200}\right) = \frac{2n - 64541}{200} > 0 \qquad (n > 32270.5).$$

Combine this arithmetic threshold with the asymptotic source and entropy thresholds. Isolate slots always return zero. The other slots retain the same fixed single-coordinate dependence, and codeword isomorphisms extend while fixing the isolate. The classical transcript proof and Lemma 8.1 therefore apply unchanged. Since $n = \Theta(N)$, both query lower bounds follow. □

10 Rigidity of every hard input

The rank and support-asymmetry properties in Lemma 3.1 make the local ambiguities proved below strictly different from global symmetries. We first identify all base-graph automorphisms, then verify that the size-control surgeries have a canonical inverse.

Lemma 10.1. *If the support of A is asymmetric, then for every b, $\mathrm{Aut}(G_b) \cong \ker_{\mathbb{F}_2} A$. In particular full column rank makes every G_b asymmetric.*

Proof. An arbitrary uncolored automorphism permutes the designated nine-cycles, by the short-cycle recognition proved above. Each row's four overlapping six-cycles map into a single row and cover its ten vertices. The internal row isomorphism preserves middle and outer roles and permutes the three opposite outer pairs, as in the exact intact-row argument.

The opposite outer pair for an occurrence of variable v has its two external neighbors on the two designated bit cycles of v. Its image is an opposite pair in the image row, whose two external neighbors lie on the two bit cycles of one variable w. Thus the two entire cycles of a variable remain paired; they cannot map to unrelated variable pairs. This induces permutations of rows and variables preserving the incidence support. By support asymmetry both permutations are identities.

Let x_v record whether the two cycles of v are swapped. The nine occurrence rows of each variable are distinct. Since those rows are fixed, external adjacency fixes every occurrence position j, giving

$$z(v, a, j) \longmapsto z(v, a + x_v, j).$$

No rotation or reflection of a cycle remains. The external neighbor then forces $o(i, k, a) \mapsto o(i, k, a + x_{v_{ik}})$. The three support variables of a row are fixed, so its coordinate pairs are fixed as well. Each middle vertex is uniquely determined by its three outer neighbors; preservation of their parity requires

$$x_{v_{i1}} + x_{v_{i2}} + x_{v_{i3}} = 0$$

for every row. Hence $Ax = 0$. Conversely every such x gives the assignment-flip automorphism of Lemma 4.1. Composition corresponds to addition, proving the group identification. □

Lemma 10.2. *Every graph P_b constructed in Section 7 is asymmetric.*

Proof. Let H_b be the graph before triangle expansion: either G_b, or G_b with the paired subdivision and joining rung. We claim that H_b is triangle-free and that every six- or eight-cycle is contained in a row gadget. Without external row-to-variable edges a cycle is inside a row, an ordinary nine-cycle, or a selected ten-cycle: the rung is a bridge in the variable-only induced graph. With at least four external edges, the alternating run bound gives length at least ten, since each row run has length at least two and each variable run at least one. With exactly two external edges, both row

endpoints concern the same variable at its unique occurrence in that row. They are opposite-bit outer vertices and have row distance four, since they subdivide disjoint edges of the internal K_4. They cannot connect through ordinary bit cycles. In the selected variable a connecting path must use the rung and has length at least three: socket to subdivision vertex, rung, then subdivision vertex to opposite-bit socket. The resulting cycle has length at least $4+3+2=9$. New nine-cycles may therefore appear; we do not use nine-cycle recognition after this surgery.

The vertices lying on six-cycles are consequently exactly the row vertices. Their overlapping six-cycles identify each row intrinsically. Delete these vertices from H_b. Without subdivision the remainder is a union of nine-cycles. With subdivision it is a union of ordinary nine-cycles and two ten-cycles joined by the rung. Exactly two remaining vertices have internal degree three: the new subdivision vertices. They are adjacent. Delete their joining edge and suppress those two now degree-two vertices. This canonically recovers G_b. Every automorphism of H_b therefore descends to an automorphism of G_b, which is the identity by Lemma 10.1. In the subdivision case the two new vertices have different pairs of original neighbors and so are fixed individually. Thus H_b is asymmetric in both cases.

Finally, the only triangles of P_b are the inserted ones. A triangle using three original-vertex cells would project to a triangle in H_b; a triangle using two cells would require two edges between their corresponding vertices in H_b. Both are impossible because H_b is simple and triangle-free. The inserted triangles are thus intrinsically identifiable and disjoint. Contracting all of them recovers H_b. Every automorphism fixes each quotient vertex. Each triangle's three ports lead to three distinct quotient neighbors, so they are fixed individually. If a neighbor is also an expanded cell, the unique edge between those fixed cells still fixes the relevant port. Singleton cells are fixed already, proving that P_b is asymmetric. □

11 Paired local neighborhoods in the hard expanders

The same source choice gives a direct explanation of persistent local ambiguity. Let $B_\ell^G(u)$ mean the induced graph on vertices within graph distance ℓ of u, with u distinguished.

Proposition 11.1. *There are absolute constants $c_0 > 0$ and $C > 0$ such that the even-order targets in Theorem 2.1 can be chosen with at least $c_0 N$ disjoint pairs of vertices having isomorphic rooted radius-ℓ neighborhoods, although the targets have no nonidentity automorphisms, for every nonnegative integer $\ell \le \lfloor \log_2 N - C \rfloor$.*

Proof. First consider the base graph G_b. For a row set R of size at most $\lfloor \delta n \rfloor$, Lemma 3.1 excludes every coordinate vector e_v from the row span of A_R. Finite-dimensional orthogonality therefore supplies

$$A_R x = 0, \qquad x_v = 1 : \tag{9}$$

otherwise e_v would annihilate $\ker A_R$ and belong to its row span.

Fix paired roots $z(v,0,j)$ and $z(v,1,j)$. Let R contain every row gadget meeting either of their radius-ℓ balls. In a cubic graph,

$$|B_\ell(u)| \le 1 + 3(2^\ell - 1) = 3 \cdot 2^\ell - 2,$$

so $|R| \le 6 \cdot 2^\ell - 4$. If $6 \cdot 2^\ell \le \delta n$, choose x as in (9). Let F_x be the assignment-flip permutation displayed in Lemma 4.1. It is an involution interchanging the two roots. In the same graph G_b, it preserves every cycle and external edge, and every internal edge of a row in R, since $(Ax)_i = 0$ there.

Every path of length at most ℓ from either root uses internal row edges only in rows in R. Applying F_x therefore maps paths from one root to equally long paths from the other. Apply this in both directions and use the involution to obtain a bijection between the two ball vertex sets. Edges within the induced balls are preserved for the same reason, and the inverse gives preservation of nonedges. Thus the balls are isomorphic as ordinary uncolored rooted graphs. The assignment is selected for these two balls.

Under the named all-even surgery, let U be the base vertices incident to changed edges. The edit bound gives $|U| \le 322$. A base root at distance greater than ℓ from U has exactly the same induced radius-ℓ ball after surgery: a path using a changed edge would first reach a base endpoint in U by unchanged base edges. This also excludes shortcuts through the former isolates.

There are $9(n/3) = 3n$ disjoint paired variable roots. At most

$$322(3 \cdot 2^{\ell} - 2) \le 966 \cdot 2^{\ell}$$

pairs have an endpoint within distance ℓ of U. If

$$2^{\ell} \le cn, \qquad c = \min\{\delta/12, 1/1000\},$$

the base-ball condition holds and fewer than n pairs are spoiled. More than $2n$ pairs survive in P_b. For all large N, $N = 16n + r$, $r \le 46$, implies $n \ge N/32$. Thus a fixed choice of C with $2^{-C} \le c/32$, and for example $c_0 = 1/16$, gives the stated conclusion, in particular for the target $b = 0$. Choose the surviving pairs at the largest stated radius; the same pairs work for every smaller radius. □

The proposition concerns rooted induced graphs, rather than port or vertex labels. It is consistent with global edge expansion: many local views remain paired even though the diameter is logarithmic. This distinguishes the family from the local-identification hypotheses used by the fixed-target algorithms in [10]. The parity-flip method has CFI ancestry; the proposition is a structural consequence for the same hard targets.

12 Analytic constructions and finite examples

The source existence argument, universal distance estimate, padding inequalities and exact acceptance calculations prove all stated asymptotic claims. Each finite encoding and its incidence representation is specified by the displayed formulas and fixed occurrence orders. Finite implementations can check adjacency, assignment flips, selected bijections and surgery residues directly from those definitions. The proofs apply to every bijection and every encoded word, independently of such finite examples.

A Expanding $(3, 9)$ systems

We prove Lemma 3.1 by the sparse-configuration method underlying Dror's input [7], with the degree-nine parameters made explicit. Write $(z)_k = z(z-1)\cdots(z-k+1)$. Give each of n left vertices three stubs and each of $n/3$ right vertices nine; match the $3n$ stubs uniformly.

A candidate parallel pair specifies a left vertex, two of its stubs, a right vertex, two of its stubs and one of their two pairings. There are

$$n\binom{3}{2}(n/3)\binom{9}{2}2 = 72n^2$$

candidates, each with probability $1/(3n)_2$. If X counts them, then simplicity is exactly $X = 0$, including triple edges which contribute three candidate events. For each fixed k, ordered tuples with distinct left and right vertices number $(72n^2)^k + O_k(n^{2k-1})$ and prescribe $2k$ pairings; their contribution to $\mathbb{E}(X)_k$ tends to 8^k. For a compatible exceptional tuple group its prescribed pairings into cells. Each cell contains at least two pairings; left degree three permits at most one cell at each left vertex. With q cells, $\ell = q$ left vertices, $s \le q$ right vertices and $e \ge 2q$ pairings. Equality $\ell + s - e = 0$ forces two pairings in each cell and distinct right vertices, the principal case. Every other pattern has exponent at most -1, so its $O_k(n^{\ell+s})$ realizations of probability $O_k(n^{-e})$ contribute $O_k(1/n)$. Incompatible tuples contribute zero. Thus $\mathbb{E}(X)_k \to 8^k$. For each fixed h, finite Bonferroni bounds give

$$\sum_{j=0}^{2h+1} \frac{(-1)^j \mathbb{E}(X)_j}{j!} \le \Pr[X = 0] \le \sum_{j=0}^{2h} \frac{(-1)^j \mathbb{E}(X)_j}{j!}.$$

First let $n \to \infty$ and then $h \to \infty$. Both bounds tend to $e^{-8} > 0$.

For a left set of size $s \ge 2$, put $t = \lfloor 7s/4 \rfloor$. Failure of $|N(S)| > 7s/4$ puts all its stubs in a right set of size t. For $t \le n/3$, the union bound is

$$\binom{n}{s}\binom{n/3}{t}\frac{(9t)_{3s}}{(3n)_{3s}} \le \binom{n}{s}\binom{n/3}{t}(3t/n)^{3s} \le [C(s/n)^{1/4}]^s.$$

Indeed with $z = t/s \in [3/2, 7/4]$, elementary binomial estimates give $[e^{1+z}(3z)^{3-z}(s/n)^{2-z}]^s$. Choose $\delta < 4/21$ small enough that $C\delta^{1/4} \le 1/2$. Splitting $2 \le s \le \delta n$ at $\sqrt{n}$ bounds the first part by a geometric tail with base $Cn^{-1/8}$ and the second by a tail of $\sum_s 2^{-s}$. Both vanish. In a simple configuration, singleton sets also have three distinct neighbors. For a nonempty small set S of size s, let $u(S)$ count neighbors incident to exactly one edge from S. Counting its $3s$ edges gives

$$u(S) \ge 2|N(S)| - 3s > s/2.$$

For $s \ge 2$, at least two columns are uniquely incident; for a singleton there are three. Thus every nonempty sum of at most $\lfloor \delta n \rfloor$ rows has Hamming weight at least two. It is neither zero nor a coordinate vector. Applying this to every subset of a small row set proves both independence and the stated exclusion from its row span.

A.1 Global expansion of all cuts

We prove that for some absolute $\alpha > 0$, with probability $1 - o(1)$, the configuration multigraph satisfies

$$|\partial W| > \alpha \min\{\mathrm{vol}(W), \mathrm{vol}(W^c)\} \tag{10}$$

for every nontrivial vertex cut. Volume is the sum of degrees, whose total is $6n$. Choose the smaller-volume side and write it as $W = S \cup T$, with $s = |S|$ left vertices and $t = |T|$ right vertices. Thus $v = 3s + 9t \le 3n$. A bad cut has boundary $b \le \alpha v$, which implies

$$|3s - 9t| \le b \le \alpha(3s + 9t).$$

For $0 < \alpha \le 1/100$, a nonempty bad cut has $s, t > 0$, and

$$\frac{1-\alpha}{3(1+\alpha)} \le r := \frac{t}{s} \le \frac{1+\alpha}{3(1-\alpha)}, \qquad \frac{3}{10} \le r \le \frac{7}{20}, \qquad s < \frac{3n}{5}. \tag{11}$$

If one side type is absent, the boundary equals its volume and is not bad.

For fixed vertex sets S, T, the images X of the $3s$ stubs of S form a uniform $3s$-subset of the $3n$ right stubs. If T_{st} denotes the $9t$ stubs at T, then $b = |X \triangle T_{\mathrm{st}}|$. Symmetric difference with T_{st} is injective. With $B = \lfloor \alpha v \rfloor$, the probability of this fixed bad cut is at most

$$\frac{\sum_{j=0}^{B} \binom{3n}{j}}{\binom{3n}{3s}}. \tag{12}$$

Multiply by $\binom{n}{s}\binom{n/3}{t}$ and split the sizes as follows.

Small cuts. By (11), $B \le 7\alpha s$. For $B \ge 1$,

$$\sum_{j \le B} \binom{3n}{j} \le (B+1)(3en/B)^B \le (6en/B)^B \le [6en/(7\alpha s)]^{7\alpha s}.$$

The last step uses monotonicity of $x \log(6en/x)$ for $x < 6n$. For $B = 0$ its final bound is at least one and still applies. The elementary binomial bounds give

$$\frac{\binom{n}{s}\binom{n/3}{t}}{\binom{3n}{3s}} \le [e^{1+r}(3r)^{-r}(s/n)^{2-r}]^s.$$

On $3/10 \le r \le 7/20$, the first factor is less than 10: $e^{1+r} < e^2 < 9$ and $(3r)^{-r} \le 10/9$. The function $(6e/(7\alpha))^{7\alpha}$ increases for $0 < \alpha \le 1/100$ and is less than 3, since at $1/100$ it is bounded by $(3^6)^{7/100} = 3^{42/100} < 3$. Also $2 - r - 7\alpha > 3/2$. Thus the union bound for a fixed s, t is at most

$$[30(s/n)^{3/2}]^s.$$

There are at most $s+1 \le 2^s$ possible t, so their sum replaces 30 by 60. Put $\gamma = 1/100$. For $s \le \sqrt{n}$, sum a geometric tail with base $60n^{-3/4}$; for $\sqrt{n} < s \le \gamma n$, use $60\gamma^{3/2} = 3/50 < 1/2$. Both totals tend to zero, uniformly over $0 < \alpha \le 1/100$.

Macroscopic cuts. Set $z = s/n$ and $w = 3t/n$. Then

$$\gamma \le z < 3/5, \qquad |z - w| \le \alpha(z + w) \le \alpha, \qquad z + w \le 1.$$

Let $H(x) = -x \log x - (1-x)\log(1-x)$, continuously extended to $[0,1]$. The binomial entropy bounds

$$\frac{e^{MH(k/M)}}{M+1} \le \binom{M}{k} \le e^{MH(k/M)}$$

show that the union bound for fixed s, t is at most

$$(3n+1)^2 \exp\{n[H(w)/3 - 2H(z) + 3H(\alpha)]\}. \tag{13}$$

For the numerator in (12), use $B/(3n) \le \alpha < 1/2$: each of its at most $3n+1$ terms is at most $e^{3nH(\alpha)}$. The lower entropy bound follows, for example, because k is a mode of Binomial$(M, k/M)$ and its mass is at least $1/(M+1)$.

Put $h_0 = \min_{\gamma \le z \le 3/5} H(z) > 0$. Uniform continuity permits choosing a fixed $0 < \alpha \le 1/100$ so small that

$$\frac{1}{3} \sup_{|x-y| \le \alpha} |H(x) - H(y)| + 3H(\alpha) \le 5h_0/6.$$

The bracket in (13) is then at most $-5h_0/6$. Summing over at most n^2 size pairs still tends to zero. This proves (10).

Every degree is at least three, so

$$\min\{\mathrm{vol}(W), \mathrm{vol}(W^c)\} \geq 3 \min\{|W|, |W^c|\}.$$

Thus the successful support has edge Cheeger constant at least $h_B = 3\alpha$ and in particular is connected. Simplicity has probability $e^{-8} + o(1)$, while failure of the small-row or global-expansion properties has probability $o(1)$. Their simultaneous success therefore has positive probability for all sufficiently large n divisible by three, without assuming independence. This establishes the expansion assertions of Lemma 3.1; full column rank and support asymmetry are proved next.

A.2 Full column rank and asymmetric support

We first prove full column rank in the same unconditioned configuration model. For a nonzero vector supported on t right vertices, the preimage of its $9t$ stubs is uniform among the $9t$-subsets of the $3n$ left stubs. It is a kernel vector precisely when every row has zero or two selected stubs. Thus t must be even and $t \leq 2n/9$. The expected number of kernel vectors with that support size is exactly

$$E_t = \binom{n/3}{t} \binom{n}{9t/2} 3^{9t/2} / \binom{3n}{9t}. \tag{14}$$

Here the middle factor chooses the rows with two selected stubs, and $3^{9t/2}$ chooses their pairs. For other t the expectation is zero. Elementary binomial estimates give

$$E_t \leq [C(t/n)^{7/2}]^t, \qquad C = e^{11/2} 2^{9/2} 3^{7/2}.$$

Choose $\gamma > 0$ with $C\gamma^{7/2} \leq 1/2$. The sum for $2 \leq t \leq \sqrt{n}$ is bounded by a geometric tail with base $Cn^{-7/4}$, and for $\sqrt{n} < t \leq \gamma n$ by the tail of $\sum_t 2^{-t}$; both tend to zero.

For larger supports set $y = 3t/n \in (0, 2/3]$. Entropy bounds in (14) give

$$E_t \leq (3n+1) e^{nf(y)}, \qquad f(y) = H(3y/2) + (3y/2)\log 3 - \tfrac{8}{3} H(y),$$

where H is the natural-logarithm entropy used above. We verify its sign, rather than relying on a numerical maximization. In the interior,

$$f'(y) = \tfrac{3}{2} \log \frac{2-3y}{y} - \tfrac{8}{3} \log \frac{1-y}{y}, \qquad f''(y) = \frac{7-15y}{3y(1-y)(2-3y)}.$$

The derivative increases strictly up to $7/15$, then decreases strictly. Its limit at zero is $-\infty$, and $f'(1/2) = 0$, so its value at $7/15$ is positive. There is exactly one earlier critical point, a local minimum, and $1/2$ is the unique later local maximum. Since $f(0) = 0$, the function is negative before the minimum. Thereafter its maximum is

$$f(1/2) = H(3/4) + \tfrac{3}{4} \log 3 - \tfrac{8}{3} \log 2 = -\tfrac{2}{3} \log 2 < 0.$$

After $1/2$ it decreases, including the continuous endpoint $2/3$. Thus $f < 0$ on $(0, 2/3]$, uniformly bounded away from zero on $[3\gamma, 2/3]$. Summing at most n entropy bounds gives $o(1)$. Together with the small-support estimate and Markov's inequality this proves $\Pr[\ker A \neq \{0\}] = o(1)$. Before conditioning, incidence multiplicities in this argument are interpreted modulo two; after simplicity this is exactly the binary matrix in the lemma.

For support asymmetry we use the following classical input of McKay–Wormald [13, Corollary 3.6 and its partitioned extension]. For the uniform simple graph class with a prescribed degree sequence, bounded maximum degree and minimum degree $d \geq 3$, the expected number of nonidentity automorphisms is $O(v^{2-d})$. The stated partitioned extension includes simple bipartite graphs: the vertex partition has two cells, within-cell edges are forbidden, and across-cell multiplicity is at most one. Applied to cell sizes $n, n/3$ and degrees $3, 9$, with $v = 4n/3$, it gives expectation $O(1/n)$, and hence probability $O(1/n)$ of a nonidentity automorphism. Distinct side degrees force all uncolored automorphisms to preserve the cells, so forgetting the partition colors does not weaken the conclusion.

Conditioned on simplicity our configuration model is uniform on exactly this bipartite class: each simple support has $(3!)^n(9!)^{n/3}$ representing stub matchings. Simplicity has probability bounded away from zero. Therefore the $o(1)$ failures of small-row expansion, global expansion and full column rank remain $o(1)$ after conditioning. Their intersection with support asymmetry has probability $1 - o(1)$ under the simple law. No independence is needed. This completes all assertions of Lemma 3.1 for every sufficiently large n divisible by three.

Disclosure of AI assistance

Generative AI tools, including OpenAI Codex, assisted with mathematical exploration, literature comparison, proof checking and manuscript preparation. The author takes responsibility for the mathematical claims and references.

Statements and Declarations

Funding. This research received no external grant funding.

Competing interests. The author owns Luminary Oddities.

Data and materials availability. The results are analytic. The graph, oracle, surgery and probabilistic constructions are specified in the text.

Author contributions. Scott Schubr is the sole author and is responsible for the research and final manuscript. AI assistance is disclosed above.

References

[1] I. Adler, N. Köhler and P. Peng, *On Testability of First-Order Properties in Bounded-Degree Graphs and Connections to Proximity-Oblivious Testing*, SIAM Journal on Computing 53(4) (2024), pp. 825–883. doi:10.1137/23M1556253. Accepted manuscript: `https://eprints.whiterose.ac.uk/id/eprint/212370/1/SubmissionSICOMP-13.pdf`.

[2] S. Apers, F. Magniez, S. Sen and D. Szabó, *Quantum Property Testing in Sparse Directed Graphs*, APPROX/RANDOM 2025, LIPIcs 353, article 32, pp. 32:1–32:24. doi:10.4230/LIPIcs.APPROX/RANDOM.2025.32.

[3] V. Arvind, J. Köbler, S. Kuhnert and Y. Vasudev, *Approximate Graph Isomorphism*, MFCS 2012, LNCS 7464, Springer, 2012, pp. 100–111. Full version: `https://eccc.weizmann.ac.il/report/2012/078/`.

[4] R. Beals, H. Buhrman, R. Cleve, M. Mosca and R. de Wolf, *Quantum lower bounds by polynomials*, Journal of the ACM 48(4) (2001), pp. 778–797. doi:10.1145/502090.502097.

[5] J.-Y. Cai, M. Fürer and N. Immerman, *An optimal lower bound on the number of variables for graph identification*, Combinatorica 12 (1992), pp. 389–410.

[6] Y. Chen, J. Flum and M. Liu, *Some remarks on the uncolored versions of the original CFI-graphs*, arXiv:2507.01459, 2025. `https://arxiv.org/abs/2507.01459`.

[7] A. Dror, *A Linear Lower Bound on Testing Graph Isomorphism in the Bounded-Degree Model*, ECCC TR26-268, 2026. `https://eccc.weizmann.ac.il/report/2026/268/`.

[8] M. Goetze, P. Jungeblut and T. Ueckerdt, *Recognition Complexity of Subgraphs of k-Connected Planar Cubic Graphs*, Algorithmica 88 (2026), article 11. doi:10.1007/s00453-025-01350-8.

[9] O. Goldreich, *Testing Isomorphism in the Bounded-Degree Graph Model*, Computational Complexity and Local Algorithms, LNCS 15700, Springer, 2025, pp. 293–323. doi:10.1007/978-3-031-88946-2_16. Author manuscript (Open Problem 4.1): `https://www.wisdom.weizmann.ac.il/~oded/COL3/bdg-iso.pdf`.

[10] O. Goldreich and L. Tauber, *On Testing Isomorphism to a Fixed Graph in the Bounded-Degree Graph Model*, Computational Complexity and Local Algorithms, LNCS 15700, Springer, 2025, pp. 324–337. doi:10.1007/978-3-031-88946-2_17.

[11] O. Goldreich and L. Tauber, *Testing in the Bounded-Degree Graph Model with Degree Bound Two*, Computational Complexity and Local Algorithms, LNCS 15700, Springer, 2025, pp. 373–382. doi:10.1007/978-3-031-88946-2_20.

[12] K. Makarychev, R. Manokaran and M. Sviridenko, *Maximum Quadratic Assignment Problem: Reduction from Maximum Label Cover and LP-based Approximation Algorithm*, arXiv:1403.7721, 2014. `https://arxiv.org/abs/1403.7721`.

[13] B. D. McKay and N. C. Wormald, *Automorphisms of random graphs with specified degrees*, Combinatorica **4** (1984), 325–338. `https://users.cecs.anu.edu.au/~bdm/papers/AutomSpecDeg.pdf`.

[14] R. O'Donnell, J. Wright, C. Wu and Y. Zhou, *Hardness of Robust Graph Isomorphism, Lasserre Gaps, and Asymmetry of Random Graphs*, SODA 2014, pp. 1659–1677. doi:10.1137/1.9781611973402.120.