# A linear query lower bound for graph-isomorphism testing at maximum degree three

Scott Schubr
Luminary Oddities

October 23, 2026

**Abstract**

For every sufficiently large even $N$, we exhibit existentially a known connected simple 3-regular $N$-vertex target whose randomized two-sided isomorphism testing requires $\Omega(N)$ incidence queries at proximity $1/175$, even with connected cubic inputs promised. We also give maximum-degree-three fixed targets at every sufficiently large integer order. These results sharpen Dror's maximum-degree-thirteen construction and reach the minimum maximum-degree threshold: degree two admits size-independent query bounds. We retain Dror's sparse-code and adaptive indistinguishability framework. A compact $17n$-vertex encoding admits the arbitrary-bijection bound $\mathrm{dist}_H(b, \mathrm{Im}\, A) \le (3/4)\,\mathrm{dist}_{\mathrm{iso}}(G_b, G_0)$: one intact variable cycle anchors an assignment bit, and unmatched source edges pay for unanchored variables and defective rows. Bounded paired triangle expansions handle every even order without a multiplicative distance loss. We also prove that the older capped completion family has normalized distance at most $4/429 < 0.01$ on every input.

## 1 Introduction and statement

In the bounded-degree graph model, a tester asks for a specified neighbor of a named vertex of an unknown graph. Reading the entire incidence description costs linearly many queries. The fixed-target isomorphism problem asks whether the unknown graph is a relabeling of a fully known target or is far from every such relabeling.

Dror [1] proved a linear two-sided query lower bound using targets of maximum degree thirteen. His Remark 1.3 asks for the smallest maximum degree supporting this phenomenon. Goldreich and Tauber [2] show that every property of degree-two graphs has a query bound depending only on proximity. We reach the remaining endpoint.

**Theorem 1.** *There are constants $\varepsilon_0, c > 0$ and graphs $H_N$ for all sufficiently large integers $N$ such that $H_N$ is simple, undirected, has $N$ vertices and maximum degree at most three, and every randomized two-sided $\varepsilon_0$-tester for isomorphism to the known $H_N$ makes at least $cN$ incidence queries. Target preprocessing and computation are unrestricted. One may take $\varepsilon_0 = 10^{-6}$.*

The following strengthening removes both connectedness and regularity qualifications, at an explicit larger fixed proximity.

**Theorem 2.** *For every sufficiently large even $N$, there is a known connected simple 3-regular $N$-vertex graph $J_N$ such that every randomized two-sided $(1/175)$-tester for isomorphism to $J_N$ requires $\Omega(N)$ incidence queries. The lower bound holds even if inputs are promised connected, simple and 3-regular. Target preprocessing and computation are unrestricted.*

Odd orders are necessarily excluded for cubic graphs. Both target constructions are existential. Goldreich's Open Problem 4.1 [3] explicitly asks about proximity 0.01. The square-root bounds for almost all regular targets [4] are compatible with exceptional hard targets.

**Contribution and ancestry.** The contribution is the complete subcubic encoding and distance transfer under arbitrary *uncolored* bijections, together with constant-query simulation. The sparse matrix, code-membership lower bound and reduction architecture are due to Dror [1]. Parity gadgets have CFI ancestry [5]; code-distance reductions to colored graph matching already appear in Arvind et al. [6]. Partial recovery of components under edge defects occurs in O'Donnell et al. [7]; intrinsic recognition of uncolored CFI components is studied by Chen, Flum and Liu [8]. These general principles are classical. The older code-distance reduction assumes color-preserving maps; bounded color-class size is not a degree bound. Standard cubic CFI nonisomorphism also does not itself give testing-far instances: one twist changes only constantly many edges. The encoding here provides the required conjunction of degree, symmetry, robustness, size and access. The subdivided-$K_4$ completion cap is standard; for example, Goetze, Jungeblut and Ueckerdt [9] use it to complete subcubic graphs. The connected-cubic result uses a compact cap-free encoding, exact short-cycle recognition and assignment anchoring on just one intact variable cycle. Source-only edge charging handles arbitrary bijections, including split or mixed components. Connected sparse-system selection and paired triangle expansion preserve the oracle and all even orders. The capped-family upper bound identifies an obstruction specific to that completion method, rather than a limitation of cubic graphs in general.

## 2 Model and sparse-code input

All graphs are finite, simple and undirected. For $N$-vertex graphs set

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

For degree bound $d$, an input is $\varepsilon$-far from the isomorphism class of $H$ if this distance exceeds $\varepsilon dN/2$. An incidence oracle on $[N]\times[d]$ returns the indicated neighbor or 0 for a missing slot. Neighbor order is arbitrary; isomorphism is a property of the graph rather than of the slot names. A tester accepts every representation of an isomorphic input with probability at least $2/3$ and rejects every far input with probability at least $2/3$. We count worst-case queries.

The following is Dror's Lemma 2.3 [1]; Appendix A reconstructs its proof.

**Lemma 3.** *For some $\delta > 0$, every sufficiently large even $n$ admits a matrix $A \in \mathbb{F}_2^{n\times(n/2)}$ of row weight three and column weight six such that every set of at most $\lfloor \delta n \rfloor$ rows is independent.*

Fix such $A$, write $m = n/2$ and $C = \operatorname{Im} A$, and fix independently of $b$ an ordering $(v_{i1}, v_{i2}, v_{i3})$ of each row's three distinct variables. Order each variable's six occurrences by $j = 1,\dots,6$. All bit additions below are in $\mathbb{F}_2$. Put

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

**Lemma 4** (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 $b \in \mathbb{F}_2^n$. For every fixed $0 < \beta < 1/2$ with $H_2(\beta) < 1/2$, uniform $b$ satisfies $\operatorname{dist}_H(b,C) > \beta n$ with probability at least $3/4$ for all large $n$. In particular, $\beta = 1/100$ is admissible.*

*Proof.* Fix random coins and a realized decision-tree path. The distinct rows queried so far together with a new queried row are independent while their number is at most $\lfloor \delta n \rfloor$. Their coordinates of $Ax$ are uniform and independent, so the next new answer remains unbiased after conditioning on earlier answers. Repeated answers are determined. Induction proves the transcript assertion.

Since $|C| \le 2^{n/2}$, the entropy estimate for a Hamming ball, with $0 < \beta < 1/2$, gives

$$\Pr[\mathrm{dist}_H(b, C) \le \beta n] \le 2^{-n/2} \sum_{j=0}^{\lfloor \beta n \rfloor} \binom{n}{j} \le 2^{-n(1/2 - H_2(\beta))}.$$

For $\beta = 1/100$, $H_2(\beta) < 9/100$: use $\log_2 100 < 7$, $\log_2 e < 2$, and $-(1-\beta)\ln(1-\beta) \le \beta$. For every admissible fixed $\beta$ the exponent is strictly positive, so the probability is at most $1/4$ for large $n$. No full-column-rank assumption is needed. □

## 3 Subcubic encoding

We define $F_b$ on a fixed named vertex set independent of $b$. For each variable $v$ create

$$z(v, a, p), \qquad a \in \mathbb{F}_2, \quad p \in \mathbb{Z}/28\mathbb{Z}.$$

For each $a$, join the vertices in the cycle $p \leftrightarrow p + 1$. Add the single *rung* $\{z(v, 0, 0), z(v, 1, 0)\}$. Occurrence $j$ uses position $4j$ on each cycle. This variable component has 56 vertices and 57 internal edges. Interchanging the two cycles preserves every occurrence position; completeness does not require arbitrary socket permutations.

For each row $i$ create outer vertices $o(i, k, a)$, $k \in [3], a \in \mathbb{F}_2$, and middle vertices $q(i, u, w)$, $u, w \in \mathbb{F}_2$. Set

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

Join $q(i, u, w)$ to $o(i, k, t_k(i, u, w))$ for each $k$. If $v_{ik}$ has occurrence number $j$ in row $i$, add $\{o(i, k, a), z(v_{ik}, a, 4j)\}$ for $a \in \mathbb{F}_2$. The internal row graph is a once-subdivided $K_4$.

Designated cycle positions have degree three; other cycle vertices have degree two. Every row vertex has degree three in the full graph. Thus

$$|V(F_b)| = 56m + 10n = 38n, \qquad |E(F_b)| = 93n/2.$$

There are $21n$ degree-two and $17n$ degree-three vertices. Distinct variables in a row and distinct sockets for occurrences ensure simplicity. Variable and row roles are construction descriptions, not colors supplied to an isomorphism.

**Lemma 5** (Completeness)**.** *If $b = Ax$, then $F_0 \cong F_b$.*

*Proof.* The map

$$\begin{aligned} z(v, a, p) &\mapsto z(v, a + x_v, p), \\ o(i, k, a) &\mapsto o(i, k, a + x_{v_{ik}}), \\ q(i, u, w) &\mapsto q(i, u + x_{v_{i1}}, w + x_{v_{i2}}) \end{aligned}$$

is bijective. Its third row bit is correct because $b_i = x_{v_{i1}} + x_{v_{i2}} + x_{v_{i3}}$. Every cycle edge, rung, row edge and external edge is preserved. □

**Lemma 6** (One-bit simulation)**.** *A fixed labeling and fixed incidence slots permit each answer about $F_b$ to read at most one coordinate of $b$.*

*Proof.* Cycle slots one and two return predecessor and successor. Slot three returns the rung neighbor at $p = 0$, the appropriate outer neighbor at $p = 4j$, or 0 otherwise. All are independent of $b$. At an outer vertex, slot one is the external neighbor. The other two slots enumerate middle neighbors as follows: 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 the corresponding row bit can be needed. These are reciprocal adjacency lists without repeated neighbors, and all vertex and slot names are fixed before $b$ is known. □

# 4 Distance preservation under arbitrary bijections

**Proposition 7.** *For every $b$,*

$$\mathrm{dist}_H(b, C) \leq 48\ \mathrm{dist}_{\mathrm{iso}}(F_b, F_0).$$

*This also holds after adding any equal number of isolates to both graphs.*

*Proof.* Call a vertex of *variable type* if it has degree two, or has degree three and a degree-two neighbor. Exactly variable vertices satisfy this predicate: every designated cycle position has two degree-two neighbors, whereas no row vertex has a degree-two neighbor. The induced variable-type components are precisely the 56-vertex variable components. The induced row-type components are the 10-vertex row components. With padding, degree-zero vertices form a third type and are excluded from row type.

Fix any bijection $\phi : V(F_b) \to V(F_0)$, including padding. Pull target adjacency back to the source set and let

$$K = |E(F_b) \triangle \phi^{-1}(E(F_0))|.$$

Let $D$ contain all endpoints of discrepant edges, including edges present only in the pulled target. Then $|D| \leq 2K$. Set $W = D \cup N_{F_b}(D)$; degree at most three gives $|W| \leq 8K$. A source variable or row component is *bad* if it meets $W$. Their counts $B_v, B_c$ obey

$$B_v + B_c \leq 8K, \tag{1}$$

since these components are disjoint.

At each vertex outside $D$, the whole adjacency list agrees with its pulled-target list. At a vertex outside $W$, this holds at all source neighbors too. Its degree and every neighbor's degree agree with their image degrees. Hence intrinsic type is preserved outside $W$. Extra target neighbors cannot evade this observation, since the whole list at the vertex agrees.

A component avoiding $W$, called *good*, has exact internal adjacency, and its connected image has one intrinsic type. Its image lies in a single target component of that type. Equal component sizes, 56 or 10, and injectivity force the image to be that whole target component. Distinct good components have distinct images.

The rung is the unique bridge of a variable component. A good variable component map therefore permutes its two complete cycles. For each good source variable $v$, define $x_v$ so that its bit-$a$ cycle maps to the target bit-$(a + x_v)$ cycle. Set $x_v$ arbitrarily on bad variables. Independent cycle reflections and changes of occurrence positions do not change this global bit interchange.

Now consider a good source row $i$ with all three incident variable components good. It maps onto an entire target row component. Internal degree distinguishes its middle and outer vertices, with degrees three and two. The two source outer vertices for one coordinate have external edges into the same good source variable component. Their images have external edges into the same good target

variable component. Distinct source variables map to distinct target variable components, and the target row uses three distinct variables. Thus source coordinate pairs map to target coordinate pairs, permuted. External-edge preservation shows that source value $a$ becomes $a + x_v$ in that target pair.

Each source middle tuple has parity $b_i$, and its image is a target middle tuple of parity zero. Coordinate permutation preserves parity, so

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

This row is satisfied by $x$. No alignment of source and target row names is assumed. If independent cycle reflections give incompatible occurrence images, the external edges of this good row cannot all be preserved, so that attempted exception does not satisfy the hypotheses.

Every violated row is consequently bad itself or meets a bad variable. Each variable occurs six times, and (1) gives

$$\operatorname{dist}_H(b, C) \le |\{i : (Ax)_i \ne b_i\}| \le B_c + 6B_v \le 6(B_c + B_v) \le 48K.$$

Minimize over $\phi$. Padding does not change the proof: outside $D$, degrees agree, so no good nonisolated component maps into isolates. □

## 5 Linear query hardness

Take $\beta = 1/100$ and $\varepsilon_0 = \beta/10000$. If $\operatorname{dist}_H(b, C) > \beta n$, Proposition 7 gives $\operatorname{dist}_{\text{iso}}(F_b, F_0) > \beta n/48$. For arbitrary sufficiently large $N$ put

$$n = 2\lfloor N/76 \rfloor, \qquad r = N - 38n \in \{0, \dots, 75\}.$$

Define $H_N$ as $F_0$ plus $r$ isolates, and use the same padding on $F_b$. Since $N \le 113n$ for $n \ge 1$,

$$\varepsilon_0(3N/2) \le 339\beta n/20000 < \beta n/48.$$

Each far vector therefore gives an $\varepsilon_0$-far graph. Codewords give isomorphic graphs, and graph queries use at most one bit query each.

If a graph tester made at most $\lfloor \delta n \rfloor$ queries, it would accept uniform $Ax$ with probability at least $2/3$. It would accept uniform $b$ with probability at most

$$(3/4)(1/3) + (1/4) \cdot 1 = 1/2,$$

using Lemma 4 and soundness. The identical transcript laws of Lemma 4 contradict this. As $n = \Theta(N)$, Theorem 1 follows. The full description has $3N$ slots, giving the matching trivial linear-query upper bound.

## 6 Connected cubic targets at proximity $1/175$

We prove Theorem 2 with a more compact graph than $F_b$. Choose $A$ with connected bipartite incidence support using Lemma 12. Keep its fixed row and occurrence orderings.

### 6.1 A compact cubic encoding

For each variable $v$ create

$$\zeta(v, a, p), \qquad a \in \mathbb{F}_2, \quad p \in \mathbb{Z}/7\mathbb{Z}.$$

Join each bit's positions in a 7-cycle and add the rung $\{\zeta(v,0,0), \zeta(v,1,0)\}$. Occurrence $j \in \{1, \dots, 6\}$ uses position $j$ on each cycle. Use new copies of the same ten-vertex row gadgets as before, with middle neighbor tuple $(u, w, b_i+u+w)$, and join $o(i,k,a)$ to $\zeta(v_{ik}, a, j)$ at its occurrence. Denote this graph by $G_b$. Every vertex has degree three, and

$$|V(G_b)| = 14m + 10n = 17n, \qquad |E(G_b)| = 15m + 18n = 51n/2.$$

Distinct row variables and unique sockets ensure simplicity. The variable and row clouds are connected; contracting them gives the connected incidence support. Hence every $G_b$ is connected.

The formulas of Lemma 5, with $z$ replaced by $\zeta$, give $G_0 \cong G_b$ when $b = Ax$. More generally they give $G_b \cong G_{b+Ax}$: the third target value is

$$(b_i + (Ax)_i) + (u + x_{v_{i1}}) + (w + x_{v_{i2}}) = b_i + u + w + x_{v_{i3}}.$$

Cycle slots return predecessor, successor and the rung or occurrence neighbor. The row slots are exactly those of Lemma 6. All names and slots are fixed before $b$, and each graph answer still reads at most one coordinate of $b$.

Recognizing uncolored CFI components through short cycles has explicit ancestry in Chen–Flum–Liu [8, Lemma 7.12]. We prove the exact cycle classification needed for this encoding.

**Lemma 8** (Exact short-cycle recognition)**.** *The only simple 7-cycles of $G_b$ are its designated variable cycles. Every simple 6-cycle lies inside a row cloud, and every row vertex lies on such a cycle.*

*Proof.* Inside a variable cloud the rung is a bridge, so its only simple cycles are the two designated 7-cycles. A row cloud is a subdivided $K_4$; its cycles have lengths six or eight, and its four subdivided triangles cover all ten vertices.

A cycle using external edges has an even number of them. If there are exactly two, its internal parts lie in one variable cloud and one row cloud. A row uses a variable only once, so the outer endpoints are the two values of one coordinate. Their middle neighborhoods are disjoint, giving internal row distance four. Their variable endpoints have opposite bits at the same nonzero occurrence; the path between them uses the rung and has length at least three. The whole cycle therefore has length at least $2 + 4 + 3 = 9$. With $2k \geq 4$ external edges, there are $k$ row-internal runs of length at least two and $k$ variable-internal runs of length at least one. Outer vertices are not adjacent, and a socket has only one external edge, so these lower bounds apply even if clouds are revisited. The cycle has length at least $5k \geq 10$. Thus no external cycle has length six or seven. □

### 6.2 Anchoring one intact cycle

**Proposition 9.** *For every $b$ and every arbitrary vertex bijection $\phi : G_b \to G_0$ with $K$ unordered-edge discrepancies,*

$$\operatorname{dist}_H(b, C) \leq \frac{3}{4}K.$$

*The same holds after adding any equal number of isolates to both graphs.*

*Proof.* Call a source edge unmatched if its image is not a target edge. Partition their counts as $q_{\mathrm{cyc}}, q_{\mathrm{rung}}, q_{\mathrm{row}}, q_{\mathrm{ext}}$ according to source cycle, rung, row-internal and external roles. No role-preservation assumption is made about $\phi$. Source and target have equal edge cardinalities, so

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

Call a source variable *anchored* if at least one of its two 7-cycles has all seven edges matched. Choose one such cycle. It maps injectively to a simple target 7-cycle, hence to one entire variable bit cycle by Lemma 8. If its source bit is $a$ and target bit is $c$, set $x_v = a + c$. Every occurrence vertex on that chosen cycle then has image bit $a + x_v$. No agreement with the other cycle, its image variable, or its occurrence positions is required. Even two intact source cycles may give inconsistent candidate bits; only the selected one is used.

Call the other variables *free*, and let $B$ count them. Each has an unmatched source edge on each of its two disjoint cycles. Those cycle edge sets are also disjoint between variables, so

$$2B \le q_{\mathrm{cyc}}. \tag{3}$$

A source row is *intact* if all twelve internal edges match. Its four overlapping 6-cycles map into target row clouds by Lemma 8. Their overlaps force the same target cloud. Their union has ten distinct vertices, so the image is the whole row. The twelve matched edges exhaust its twelve internal edges, giving an internal isomorphism. Internal degrees distinguish middle and outer vertices. The two outer vertices of each coordinate have disjoint middle neighborhoods; these three complementary pairs are preserved. Thus the coordinate pairs are permuted, with value flips $y(i,k)$. The middle parities give

$$\sum_{k=1}^{3} y(i,k) = b_i. \tag{4}$$

If $B_c$ counts nonintact rows, their disjoint internal edge sets give $B_c \le q_{\mathrm{row}}$.

Fix the chosen bits on anchored variables and give free variables independent uniform bits. Partition rows into three disjoint classes. A row meeting a free variable has violation probability exactly $1/2$, since its variables are distinct. At most $6B$ rows meet free variables; their expected violations are at most $3B \le (3/2)q_{\mathrm{cyc}}$. Among other rows, those with nonintact interiors contribute at most $B_c \le q_{\mathrm{row}}$ violations.

It remains to consider intact rows whose variables are all anchored. If such a row is violated, (4) gives some coordinate $k$ with $y(i,k) \neq x_{v_{ik}}$. Use its external edge at the bit $a$ of that variable's chosen intact cycle. Its mapped outer endpoint has bit $a + y(i,k)$, while the mapped variable endpoint has bit $a + x_{v_{ik}}$. These differ. A target edge between an outer and a variable vertex is external and has equal endpoint bits, so this source edge is unmatched. Incorrect target variable or occurrence labels cannot help; if the socket maps to target rung position zero, there is no external edge there at all. Distinct violated rows have distinct source outer vertices, so their chosen edges are distinct. These violations number at most $q_{\mathrm{ext}}$.

By linearity of expectation, some choice of free bits has at most

$$\frac{3}{2}q_{\mathrm{cyc}} + q_{\mathrm{row}} + q_{\mathrm{ext}} \le \frac{3}{2}(q_{\mathrm{cyc}} + q_{\mathrm{rung}} + q_{\mathrm{row}} + q_{\mathrm{ext}}) = \frac{3}{4}K$$

violated rows. This proves the claim for every $\phi$. Independence between different rows is not needed. Equal added isolates create no cycles, change no source edge roles, and preserve the equal total edge count in (2). An intact cycle cannot map into isolates. Hence every step also holds with isolates, even if $\phi$ mixes them with nonisolated vertices. Taking the minimum over $\phi$ gives the corresponding inequality for isomorphism distance. □

### 6.3 Every sufficiently large even order

For even $N$, choose the largest positive even $n$ such that $17n \le N$ and $17n \equiv N \pmod 4$. Eligible $n$ form one class modulo four, so

$$r = N - 17n \in \{0, 4, 8, \dots, 64\}.$$

Select $r/4 \le 16$ pairs $\zeta(v, 0, p), \zeta(v, 1, p)$ in a fixed order independently of $b$. There are $7n/2$ such pairs, enough for large $n$. At each selected vertex replace that vertex by a triangle, attaching its three former edges at distinct triangle vertices labeled by the predecessor, successor and third-incidence roles. Call the result $P_b$. Each pair adds four vertices, so $P_b$ has exactly $N$ vertices and is simple, connected and cubic. Adjacent selected vertices are handled by attaching their former connecting edge to the appropriate port at each end.

The assignment flips preserve these three variable-edge roles and the paired selection. They therefore extend to isomorphisms of the expanded graphs for codewords. All variable adjacencies and all expansion choices are independent of $b$. At a triangle vertex two slots return its fixed triangle neighbors; the third returns the old neighbor assigned to its port. Such answers are independent of $b$, since only variable vertices are expanded. Old row answers translate a neighbor label to its fixed triangle port when necessary, preserving the one-bit incidence simulation. This also applies when both endpoints of an edge are expanded.

For the distance comparison, place $G_b$ and $r$ isolates on the same $N$ named vertices as $P_b$, retaining one triangle vertex as each expanded old vertex. An individual expansion preserves one incident edge, reroutes two, and adds three triangle edges, costing at most $2 + 2 + 3 = 7$ edge modifications. Sequential expansions give the same bound when selected vertices are adjacent. With $s = r/2 \le 32$ individual expansions on each side, the triangle inequality yields

$$\operatorname{dist}_{\mathrm{iso}}(G_b \sqcup rK_1, G_0 \sqcup rK_1) \le \operatorname{dist}_{\mathrm{iso}}(P_b, P_0) + 14s \le \operatorname{dist}_{\mathrm{iso}}(P_b, P_0) + 448. \tag{5}$$

Here $rK_1$ denotes $r$ isolated vertices. We apply Proposition 9 to the unexpanded graphs with isolates; no preservation of their short-cycle predicates after expansion is assumed.

Take $\beta = 11/100$. The exact integer inequality

$$100^{100} < 2^{50} 11^{11} 89^{89}$$

is equivalent to $H_2(\beta) < 1/2$. Lemma 4 makes uniform $b$ satisfy $\operatorname{dist}_H(b, C) > \beta n$ with probability at least $3/4$ for large $n$. Proposition 9 and (5) then give

$$\operatorname{dist}_{\mathrm{iso}}(P_b, P_0) > 11n/75 - 448.$$

For $\varepsilon_1 = 1/175$ the far threshold is at most

$$\varepsilon_1(3N/2) \le 51n/350 + 96/175.$$

The slope margin is $11/75 - 51/350 = 1/1050$. It dominates the additive losses whenever $n > 470976$; in particular $n \ge 500000$ suffices, in addition to the asymptotic matrix and entropy thresholds. Thus every far word gives an $\varepsilon_1$-far graph. Codewords give isomorphic graphs, and both distributions consist of connected simple cubic inputs. Their adaptive transcript laws are identical through $\lfloor \delta n \rfloor$ graph queries, by the one-bit simulation and Lemma 4. The acceptance contradiction of Section 5 proves $\Omega(n) = \Omega(N)$ queries and hence Theorem 2.

## 7 An obstruction to capped completion

A natural alternative is to make the original $F_b$ cubic by attaching a constant-size cap at each degree-two site. We show that this particular completion family contains no 0.01-far inputs, regardless of how its lower distance estimate is sharpened.

Let $Q_5$ be $K_4$ on $u, v, c, d$ with edge $uv$ subdivided by a port $p$. The port has degree two and all other vertices degree three. Let $Q_7$ replace $u$ by a triangle, attaching its former neighbors $p, c, d$ at distinct triangle vertices. Both caps are simple and connected with exactly one degree-two port. The subdivided-$K_4$ completion cap and triangle expansion are standard constructions; see [9]. Attach one disjoint $Q_5$ or $Q_7$ by its port to every degree-two vertex of $F_b$. Use identical cap types at the paired sites $z(v, 0, p), z(v, 1, p)$, with every choice fixed independently of $b$. Write $\widehat{F}_b$ for this cubic completion and $M$ for its order. Since there are $21n$ sites, $M \geq 38n + 5(21n) = 143n$.

**Proposition 10.** *For every $b$ and every such paired cap assignment,*

$$\operatorname{dist}_{\mathrm{iso}}(\widehat{F}_b, \widehat{F}_0) \leq 4 \operatorname{dist}_H(b, C) \leq 2n, \qquad \frac{\operatorname{dist}_{\mathrm{iso}}(\widehat{F}_b, \widehat{F}_0)}{3M/2} \leq \frac{4}{429} < \frac{1}{100}.$$

*Proof.* The assignment-flip formulas extend to $F_b \cong F_{b+Ax}$ for every $b, x$, by the same third-coordinate calculation used for $G_b$. Paired cap types allow identical extensions over all cap vertices. Choose $x$ minimizing the weight of $c = b + Ax$.

Map $\widehat{F}_c$ to $\widehat{F}_0$ by fixing all variable, middle and cap vertices and interchanging $o(i, 3, 0)$ with $o(i, 3, 1)$ exactly when $c_i = 1$. Every internal row edge is preserved: the third source bit $c_i + u + w$ becomes $u + w$. At each such row the two straight external edges are replaced by two crossed ones. These are four distinct unordered edges in their symmetric difference. Different rows have distinct outer vertices and occurrence sockets, so the discrepancy sets are disjoint. All other edges agree, giving exactly $4 \operatorname{wt}(c)$. This proves the first upper bound after composition with the assignment flip.

Every row of $A$ has odd weight, so $A\mathbf{1} = \mathbf{1}$ and both $0$ and $\mathbf{1}$ lie in $C$. Consequently $\operatorname{dist}_H(b, C) \leq \min(\operatorname{wt}(b), n - \operatorname{wt}(b)) \leq n/2$. Finally $M \geq 143n$ gives the displayed normalized bound. □

The obstruction concerns this capped input family, including paired $Q_7$ size adjustments. It does not exclude other cubic encodings. The compact construction of Section 6 avoids that size penalty, but its proved parameter $1/175$ remains below 0.01.

## 8 Finite verification

The accompanying standard-library checks test finite consequences of the analytic arguments. For the original encoding they include exact code image distances, all assignments on small $(3, 6)$ matrices, fixed-slot oracle answers, partial decoding and selected arbitrary-map perturbations. Separate finite configuration enumerations authenticate collision moments, component probabilities and exact expansion arithmetic.

The compact checks enumerate simple 6- and 7-cycles on small graphs, test one-bit oracles and completeness, and authenticate the capped-family $4 \operatorname{dist}_H$ repair. Selected splices, coordinate permutations and nongadget vertex maps exercise the anchor cases, not all vertex bijections. Padding checks construct all allowed residual sizes, including adjacent selected vertices and rung endpoints, and compare every incidence answer, full flip map and at-most-seven-edit expansion with the actual graph.

The edge-charging checks classify actual source-only defects, extract intact rows and single-cycle anchors, enumerate every free-variable assignment in the finite examples, and verify half expectations and actual unmatched external-edge witnesses before the final metric bound. Mutated maps, parsers, bit readers, partitions and extracted flips must be rejected by the same positive acceptance gates. Reader instructions record coverage, dependency identities, and normal and optimized runs.

These computations check the specified finite encoding and decoding examples. Neither finite checks nor the present 3/4 theorem settle proximity 0.01.

# A Reconstruction of the sparse-matrix input

We reconstruct the matrix argument of Dror [1], including an elementary proof of positive simplicity probability.

Give each of $n$ left vertices three labeled stubs and each of $n/2$ right vertices six. Match the $3n$ left stubs to right stubs uniformly. Let $(z)_k = z(z-1)\cdots(z-k+1)$.

**Lemma 11.** *The probability that the configuration has no parallel edges tends to $e^{-5}$.*

*Proof.* A candidate collision specifies a left vertex, two of its stubs, a right vertex, two of its stubs, and one of their two pairings. There are $n\binom{3}{2}(n/2)\binom{6}{2}\,2 = 45n^2$ candidates, each occurring with probability $1/(3n)_2$. Let $X$ count occurring candidates. Double edges contribute one event, triple edges three; simplicity is $X = 0$.

Fix $k$. In the expansion of $\mathbb{E}(X)_k$ over ordered distinct candidate events, tuples with distinct left vertices and distinct right vertices number $(45n^2)^k + O_k(n^{2k-1})$ and prescribe $2k$ distinct stub pairings. Their contribution tends to $5^k$.

For a remaining compatible tuple, group its distinct prescribed pairings by occupied vertex pair, or cell. Each cell contains at least two pairings. A left vertex has three stubs and hence belongs to at most one such cell. With $q$ cells, the numbers $\ell, s$ of distinct left and right vertices and the number $e$ of pairings satisfy $\ell = q$, $s \le q$, $e \ge 2q$. Equality $\ell + s - e = 0$ forces two pairings per cell and distinct right vertices. Each cell then determines one candidate; distinct tuple entries put this case in the principal class already counted. Every exceptional compatible pattern has $\ell + s - e \le -1$. There are finitely many patterns for fixed $k$, each with $O_k(n^{\ell+s})$ realizations of probability $1/(3n)_e = O_k(n^{-e})$. Their contribution is $O_k(n^{-1})$. Incompatible tuples have probability zero, so $\mathbb{E}(X)_k \to 5^k$ for every fixed $k$.

Finite Bonferroni bounds give, for fixed $h$,

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

Let $n \to \infty$ first and then $h \to \infty$. Both bounds converge to $e^{-5}$. □

For a left set $S$ with $s = |S| \ge 2$ put $t = \lfloor 3s/2 \rfloor$. When $t \le n/2$, failure of $|N(S)| > 3s/2$ forces all $3s$ matched stubs into some $t$-vertex right set. The fixed-set probability is at most

$$(6t)_{3s}/(3n)_{3s} \le (2t/n)^{3s}.$$

Thus the union bound at size $s$ is at most

$$\binom{n}{s}\binom{n/2}{t}(2t/n)^{3s} \le [C\sqrt{s/n}]^s$$

for an absolute $C$. To see the last estimate, write $z = t/s \in [1, 3/2]$ and apply the standard binomial estimates to obtain $[e^{1+z}(2z)^{3-z}(s/n)^{2-z}]^s$. Choose $\delta < 1/3$ with $C\sqrt{\delta} \le 1/2$. For $s \le \sqrt{n}$ these summands are bounded by $(Cn^{-1/4})^s$, whose sum from $s = 2$ is $o(1)$; for $s > \sqrt{n}$ they are at most $2^{-s}$ and again sum to $o(1)$.

By Lemma 11, simple configurations without a failing set of size 2 through $\delta n$ have positive probability for large $n$. Singletons have three distinct neighbors in a simple graph. For every nonempty such $S$, its $3|S|$ edges and $|N(S)| > 3|S|/2$ force a neighbor incident to exactly one edge from $S$. The binary incidence matrix consequently has no zero sum of a nonempty set of at most $\lfloor \delta n \rfloor$ rows. This proves Lemma 3.

### A.1 Connected support

**Lemma 12.** *The matrix in Lemma 3 may additionally be chosen with connected bipartite incidence support, for every sufficiently large even $n$.*

*Proof.* In the same $(3, 6)$ configuration model, a component with $s$ left vertices has $s/2$ right vertices by stub balance. If disconnected, some component has even $s$ with $2 \le s \le n/2$. For fixed vertex sets of these sizes, all $3s$ left stubs match all $3s$ selected right stubs with probability $1/\binom{3n}{3s}$. Hence disconnection has probability at most

$$\sum_{\substack{2 \le s \le n/2 \\ s \text{ even}}} \frac{\binom{n}{s}\binom{n/2}{s/2}}{\binom{3n}{3s}}.$$

The summand is at most $[e^{3/2}(s/n)^{3/2}]^s$ by elementary binomial bounds. Choose $0 < \alpha < 1/2$ with $e^{3/2}\alpha^{3/2} \le 1/2$. For $s \le \sqrt{n}$ the sum is bounded by a geometric tail from $s = 2$ with base $e^{3/2}n^{-3/4}$; for $\sqrt{n} < s \le \alpha n$ it is bounded by a tail of $\sum_s 2^{-s}$. Both vanish.

For $\alpha n \le s \le n/2$, natural-log entropy $H$ gives the bound

$$(3n + 1) \exp[-(3n/2)H(s/n)] \le (3n + 1) \exp[-(3n/2)H(\alpha)].$$

Here we used $e^{MH(k/M)}/(M + 1) \le \binom{M}{k} \le e^{MH(k/M)}$; the lower bound follows because $k$ is a mode of Binomial$(M, k/M)$ and its modal mass is at least $1/(M + 1)$. There are at most $n$ terms, so this range also vanishes. Connectedness therefore fails with probability $o(1)$. Intersecting it with the preceding $o(1)$ expansion failure and simplicity probability tending to $e^{-5}$ still leaves positive probability, without assuming independence. The earlier row-independence argument then applies unchanged. □

## References


[1] 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/.

[2] 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.

[3] 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.

[4] 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.

[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] 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/.

[7] 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.

[8] 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.

[9] 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.