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 N, we exhibit existentially a known simple N-vertex graph of maximum degree at most three for which randomized two-sided isomorphism testing in the incidence-query model requires $\Omega(N)$ queries at one absolute positive proximity parameter. This sharpens Dror's maximum-degree-thirteen construction and resolves the minimum maximum-degree threshold raised there: degree two admits size-independent query bounds. We retain Dror's sparse binary-code and adaptive indistinguishability framework and give a constant-size subcubic encoding with a one-bit incidence oracle. The main new step is an arbitrary-bijection distance bound, proved by local recognition of variable and parity components and partial decoding outside edge defects. The targets need not be regular or connected.

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 ε_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 targets are existential and need not be connected or regular. 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.

2 Model and sparse-code input

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

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

For degree bound d, an input is ε-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 2. *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

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

Lemma 3 (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 $\beta = 1/100$, uniform b satisfies $\mathrm{dist}_H(b,C) > \beta n$ with probability at least $3/4$ for all large n.*

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 $-(1-\beta)\ln(1-\beta) \le \beta$. 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 4 (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 5 (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 6. *For every b,*

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

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 $|D| \le 2K$. Set $W = D \cup N_{F_b}(D)$; degree at most three gives $|W| \le 8K$. A source variable or row component is *bad* if it meets W. Their counts B_v, B_c obey

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

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 ϕ. 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 $\mathrm{dist}_H(b, C) > \beta n$, Proposition 6 gives $\mathrm{dist}_{\mathrm{iso}}(F_b, F_0) > \beta n/48$. 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 ε_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 3 and soundness. The identical transcript laws of Lemma 3 contradict this. As $n = \Theta(N)$, Theorem 1 follows. The full description has $3N$ slots, giving the matching trivial linear-query upper bound.

6 Finite verification

The accompanying standard-library programs check finite consequences of the construction. The encoding verifier checks all assignments on four small $(3, 6)$ incidence matrices, exact image distances, fixed-slot oracle answers, actual partial decoding, exhaustive single-transposition perturbations of selected maps, structured maps, cycle splices and padding. It rejects deliberately corrupted completeness maps, type predicates, oracle read counts and decoded bits. A separate program enumerates small configuration models, compares collision moments with independent event-intersection counts, and checks exact expansion, parity and padding arithmetic.

These computations check the specified finite encoding and decoding examples. The constant 48 has substantial slack on small nonzero-distance examples; eligible-row decoding and intermediate defect charging are more diagnostic than the final numerical bound. Reader instructions describe normal and optimized runs and recorded scope.

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 7. *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 ℓ, 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 7, simple configurations without a failing set of size 2 through δn 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 2.

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.