hexagon:2611.00977v1

A Linear Lower Bound for the Extensible No-Three-in-Line Problem

Samuel Korsky

November 19, 2026

Abstract

Answering a question of Erde, we construct one set of positive lattice points with no three on a line and at least cn in every sufficiently large square $[1, n]^2$, for an absolute constant $c > 0$.

1 Introduction

Dudeney's 1917 chessboard puzzle asks for sixteen pawns with no three on a straight line [1]. A set with this property is in *general position*. On the grid $[n]^2$, where $[n] = \{1, \dots, n\}$, two points per row give the upper bound $2n$, attained by the sixteen pawns when $n = 8$. For prime p, Erdős took the p points satisfying $y \equiv x^2 \pmod p$ in $[p]^2$ [2]. Choosing a prime with $p \le n$ and $p/n \to 1$ gives $n - o(n)$ points. Hall, Jackson, Sudbery, and Wild raised the bound to $(3/2 - o(1))n$ using modular hyperbolas [3].

Both constructions choose their prime to fit the board. In the *extensible* problem we seek a single set S in the positive lattice $\mathbb{N}^2 = \{1, 2, \dots\}^2$ as the board grows. The parabola $y = x^2$ is in general position but contributes only $\lfloor \sqrt{n} \rfloor$ points to $[n]^2$. Erde asked whether every S in general position must satisfy $\liminf_{n\to\infty} |S \cap [n]^2|/n = 0$, or whether one set can retain a positive fraction of $2n$ in every large square [4]. Green records a version with squares centered at the origin [5].

A first improvement comes from a greedy rule. Visit the grid square by square and keep each point that preserves general position. Every rejected point in $[n]^2$ then lies on a line through two retained points. Aichholzer, Eppstein, and Hainzl proved that covering the grid by such lines requires at least a constant times $n^{2/3}$ retained points [6]. Nagy, Nagy, and Woodroofe observed that this bound holds in every sufficiently large square for the same infinite greedy set [7].

The same authors placed pieces of Erdős's modular parabolas in separated squares to obtain a lower bound of order $n/(\log n)^{1+\varepsilon}$ for each $\varepsilon > 0$, while their experiments with two greedy constructions suggested linear growth [7]. Ghosal reached order $n/\sqrt{\log n}$ by sampling independently and removing points from collinear triples [8]. The following theorem removes the remaining logarithmic loss and answers Erde's question.

Theorem 1.1. *There exist a set $S \subseteq \mathbb{N}^2$ in general position and an absolute constant $c > 0$ such that, for every sufficiently large n,*

$$|S \cap [n]^2| \ge cn.$$

A related problem allows each line to hold up to $k(n)$ points in $[n]^2$. Gábriel, Jánosik, Melján, and Nádor showed how to construct sets obeying this rule from one set in general position with $|S \cap [n]^2| \geq cn$ for all large n [9, Theorem 4.3]. Theorem 1.1 supplies their starting set. For each k in their class of *sublinear limiters* [9, Definition 4.1], the resulting set obeys the line limit in every square and contains at least a constant times $nk(n)$ points in every sufficiently large $[n]^2$. This is the optimal order, since there are only n rows. In particular, a logarithmic allowance gives order $n \log n$ points in every large square.

Proof outline

We visit the positive lattice in growing squares. Each point is left blank or colored, with a fair coin choosing red or blue. Only the reds enter our set. Growing the square to side r adds $2r - 1$ points, so a starting chance proportional to $1/r$ of coloring each would give roughly the same expected number of reds in every layer. The obstacle is that just two reds can rule out a whole line.

Two points of the same color close their line to all later choices, so no three reds can be collinear. Consider a line whose only colored point is red. A second red closes it, but a blue is equally likely. In the blue case we double the chance of coloring each still eligible point on the line, up to a limit that keeps probabilities below one. Wherever doubling is possible, the gain in one outcome exactly balances the loss in the other.

To allow for the limit, the proof counts a smaller fraction of a point's chance with each doubling it uses. A line carrying both colors can still close without a compensating gain, but that requires three colored points on the same line. And although one toss can affect many future points, the cases where this increases the estimated loss require a triangle with an extra point on each of its three side lines. Counting these patterns at the chosen coloring rate shows that a fixed fraction of the linear total survives on average.

For squares of side $2, 4, 8, \ldots$, the probabilities of falling short sum to less than one beyond some scale. Thus one run keeps a linear number of reds at every sufficiently large power of two, which implies the result.

2 How to choose the points

Start with a blank grid. When we visit a point, we either leave it blank or put a dot there. If we put down a dot, we toss a fair coin for its color, either red or blue with equal probability. Both colors affect later decisions, and each decision is permanent. The set S consists of the points that receive red dots.

We visit all of $[1]^2$, then the new points of $[2]^2$, then the new points of $[3]^2$, and so on. The points added at each stage form a *layer*. The layer number of $x = (i, j)$ is $r_x = \max(i, j)$, the side length of the first square containing it. Within a layer we visit points in order of increasing i, breaking ties by increasing j. The first four visits are to $(1, 1)$, $(1, 2)$, $(2, 1)$ and $(2, 2)$.

1. **Decide whether a dot is allowed.** Look along each line joining the current point x to an

earlier dot. If any such line contains two red dots or two blue dots, leave x blank and move on.

2. **Decide whether to place the dot.** If x passes the first test, each of those lines contains either a single dot or one dot of each color. Start with a small probability of placing a dot at x. Double that probability once for each line with a red–blue pair, stopping after a prescribed number of doublings. Place a dot with the resulting probability, using a fresh random draw, and toss a separate fair coin for its color if it is placed.

Example. Consider the visit to $(4, 4)$ when the only earlier dots are among the three points shown below. All three lie on the diagonal through $(4, 4)$.

$(1,1)$	$(2,2)$	$(3,3)$	Decision at $(4,4)$
red	blank	blank	use the starting probability
red	red	blank	leave blank
red	blue	blank	double the starting probability
red	blue	either color	leave blank

In the last row the dot at $(3, 3)$ matches one of the first two dots, whichever color the coin gives it.

We can already see why the red points are in general position. If a line ever received a third red dot, its first two red dots would have forced us to leave that point blank in step 1.

It remains to specify the two numbers used in step 2. At an allowed point x, let μ_x be the chance of placing a dot when no line gives a doubling, and let K_x count how many doublings we allow. With $\gamma = 10^{-4}$, take

$$\mu_x = \frac{\gamma}{r_x}, \qquad K_x = 1 + \lfloor \gamma \log_2 r_x \rfloor .$$

There are $2r - 1$ points in layer r, so their starting probabilities add up to nearly 2γ in every layer. If m lines through an allowed point x contain a red–blue pair, its chance of receiving a dot is $\mu_x 2^{\min(m, K_x)}$. For instance, the third row of the example gives probability $2(\gamma/4) = \gamma/2$. Even if all K_x doublings occur, the probability is at most

$$a_x = 2^{K_x} \mu_x \le 2 r_x^{\gamma} \mu_x = 2\gamma r_x^{\gamma - 1} < 1. \tag{1}$$

Why double the probability?

Consider a future point x at any finite stage of the construction. Write p_x for the chance the rule would give it if we visited it next. Suppose that a line through x has one red dot, that x is still allowed, and that a doubling is available. If a second dot appears on the line, a red dot changes the chance at x to zero and a blue dot changes it to $2p_x$. The two outcomes are equally likely, so their average is p_x.

To account for the limit on doubling, we give each future point a score that also records how many doublings remain. Suppose x is still allowed and has used $d = \min(m, K_x)$ doublings. Multiply its current probability by the fraction of its doublings still available to get

$$H_x = p_x \cdot \frac{K_x - d}{K_x}, \qquad p_x = \mu_x 2^d. \tag{2}$$

A point already ruled out has probability and score zero. The score starts at μ_x and always lies between zero and p_x. We use it in the proof because its average loss has a simple formula.

Lemma 2.1. *When a line through a future point x receives its second dot, the expected loss of score at x, averaged over the color of that dot, is at most p_x/K_x. A third dot on the line costs at most p_x.*

Proof. Suppose x is allowed and $k = K_x - d \geq 1$ doublings remain. If the second dot matches the first, the new score is zero. If the colors differ, the probability doubles and $k-1$ doublings remain. By (2), the mean of the two new scores is

$$\frac{1}{2} \cdot 0 + \frac{1}{2} \cdot \frac{2p_x(k-1)}{K_x} = \frac{p_x k}{K_x} - \frac{p_x}{K_x} = H_x - \frac{p_x}{K_x}.$$

If all doublings are used or the point is already ruled out, the score stays zero. Three dots on one line include two of the same color, so the third dot can erase at most the current score $H_x \leq p_x$. □

The gain from the factor $1/K_x$ is easiest to see by comparison with independent sampling. Suppose we choose each point in $[n]^2$ independently with probability $1/n$, so we expect to choose n points. Guy and Kelly showed that the number of collinear triples in the square is asymptotic to $(3/\pi^2)n^4 \log n$ [10]. Each triple survives with probability $1/n^3$, so the expected number of surviving triples grows like a constant times $n \log n$. In the next section the factor $1/K_x$ will cancel the extra logarithm in the triple count. Related sequential sampling methods appear in work of Dhawan [11].

3 Counting red points in one shell

Fix a power of two $N \geq 2$. We count red points in the part of the square between layers N and $2N-1$, which we call the *shell*

$$T_N = \{x : N \leq r_x < 2N\}.$$

Every decision that affects this shell takes place in the finite square $B = [2N-1]^2$. Our goal is to show that the chance of finding fewer than $\gamma N/2$ red points in this shell becomes small as N grows. Section 5 will use that estimate for all the shells at once.

Write $y < u$ if we visit y before u, and let I_y be 1 if y receives a dot of either color and 0 if it stays blank. When we speak of a *conditional* mean at u, we hold all earlier decisions fixed and average over the possible outcomes at u.

Consider what a visit to u can do to the score at a later point $x \in T_N$. A blank at u leaves the score unchanged, as does the first dot on the line through u and x. If the line already contains one dot, a new dot is its second, so Lemma 2.1 bounds the mean loss by $p_u p_x/K_x$. Here the factor p_u is the chance of putting down a dot at all. If the line already contains two dots, the mean loss is at most $p_u p_x$. A line with three or more dots has already ruled out x.

We can therefore bound the total mean loss at later shell points by summing two kinds of terms, one for each collinear triple and one for each collinear quadruple with last point $x \in T_N$. Their respective contributions are

$$I_y p_u p_x/K_x \quad (y < u < x), \qquad I_y I_v p_u p_x \quad (y < v < u < x). \tag{3}$$

The indicators make a term zero unless the required earlier dots are present. If u itself is in the shell, we replace its score by 2 when it receives a red dot and by 0 otherwise. The expected replacement is $2(p_u/2) = p_u \geq H_u$. After the last visit, the sum of these replacements is twice the number of red points in the shell.

Keeping track of the possible losses

Each term in (3) contains probabilities that change as we visit the grid. To bound those terms, we assign a product to each triple or quadruple and follow it until its next-to-last point is visited. For a tuple ending in u, x, the visit to u is precisely when the score at x can suffer the loss in (3).

For a tuple on a line L and an unvisited point w of the tuple, evaluate the rule at w using the actual earlier dots on every line except L. Let q_w^L be twice the probability so obtained. Restoring the dots on L can rule out w, leave its probability unchanged, or supply at most one doubling. With $a_w = 2^{K_w}\mu_w$ denoting the upper bound from (1), we have

$$p_w \leq q_w^L \leq 2a_w. \tag{4}$$

Omitting L ensures that a new dot on L leaves q_w^L unchanged at every other unvisited point w of the tuple.

Start the product with one factor q_w^L for every point of the tuple and include an additional factor $1/K_x$ for a triple. When an early point is visited, replace its factor by its indicator. Remove the whole product just before the visit to the next-to-last point. For a triple $y < u < x$, writing $q_w = q_w^L$, the successive expressions are

$$\frac{q_y q_u q_x}{K_x} \quad \longrightarrow \quad \frac{I_y q_u q_x}{K_x} \quad \longrightarrow \quad 0.$$

The first arrow is the visit to y and the second is the removal at u. Between these visits, every remaining q_w is evaluated using the current history.

Let W be the sum of the scores at unvisited shell points plus twice the number of red points already seen in the shell. Subtract all the products still present to form

$$\Phi = W - \sum_{P \text{ still present}} P.$$

When a product disappears, its value is added to Φ and covers the corresponding mean loss in (3) by (4). Only increases in the conditional means of the remaining products can lower the conditional mean of Φ.

Lemma 3.1. *For all sufficiently large N, the initial and final values satisfy*

$$\Phi_{\text{start}} \geq \tfrac{3}{2}\gamma N, \qquad \Phi_{\text{end}} = 2|S \cap T_N|.$$

Proof. Initially $q_w^L = 2\mu_w$. Give each tuple the product of its starting weights, and let W_d be the sum over collinear d-tuples in B listed in visit order.

Fix the first point b and let $r = r_b$. Every lattice line through b has a direction vector $v = (v_1, v_2)$ with relatively prime coordinates, unique up to sign. Call $h = \max(|v_1|, |v_2|)$ its height. The boundary of the square $[-h, h]^2$ has $8h$ lattice points, so identifying opposite vectors gives at most $4h$ directions of height h. Every later point on that line has the form $b + jv$ for a nonzero integer j. Its layer number is at least r by visit order. It is also at least $|j|h$ because the later square contains both points and one of their coordinates differs by $|j|h$. With $R = 2N$, the total weight of possible partners in that direction is at most

$$2\gamma \sum_{j=1}^{\lfloor R/h \rfloor} \frac{1}{\max(r, jh)} \le 2\gamma \int_0^{R/h} \frac{dt}{\max(r, th)} = \frac{2\gamma}{h}(1 + \log(R/r)).$$

Use this bound for each of the $d - 1$ later points, allowing repeated partners to obtain an upper bound. The first points in any one layer have total weight at most 2γ. Summing $(1 + \log(R/r))^{d-1}$ over layers $r \le R$ gives $O_d(R)$ by comparison with its integral. For triples, the remaining sum over directions is bounded by a constant times $\sum_{h \le R} 1/h \le 1 + \log R$. For quadruples it involves $\sum_{h \le R} 1/h^2 \le 2$. This gives

$$W_3 = O(\gamma^3 N \log(2N)), \qquad W_4 = O(\gamma^4 N).$$

Each triple product comes with the factor $1/K_x$, and $K_x \ge \gamma \log_2 N$ throughout the shell. This is where the saving from the fair coin cancels the logarithm in the triple count. Since each initial factor is $2\mu_w$, the sum of all initial products is at most

$$\frac{8W_3}{\gamma \log_2 N} + 16W_4 = O((\gamma^2 + \gamma^4)N).$$

For our choice of γ, this is less than $\gamma N/3$ once N is sufficiently large. The initial shell scores total $\sum_{r=N}^{2N-1}(2r-1)\gamma/r \ge 2\gamma N - \gamma$, giving the first bound after subtraction. At the end the products are gone and $W = 2|S \cap T_N|$. □

4 When one visit affects two points

Fix a product that is still present after we visit z, and let L be the line of its tuple. First suppose z lies on L. The other unvisited factors ignore that line, so the only possible change is the replacement of q_z^L by I_z if z belongs to the tuple. The mean of I_z is $p_z \le q_z^L$, so this change cannot increase the product in conditional mean.

Now suppose z lies off L. Each factor separately has conditional mean at most its old value. Indeed, a blank changes nothing, and whenever one color can raise a factor, it can at most double it while the other color makes it zero. If at most one factor can rise, bound all the other factors by their old values for every outcome. Averaging the one remaining factor shows that the whole product has conditional mean at most its old value.

For the mean of a product to increase, there must therefore be two different unvisited points x, y of the tuple whose factors can rise at z. A doubling at x requires an earlier dot on the line zx, and a

doubling at y requires an earlier dot on zy. The tuple itself supplies a third point on its line xy. We have found a triangle with one extra point on each of its three side lines, as in Figure 1. These six points are distinct because different side lines meet only at a vertex. A quadruple contributes a seventh point on the base line xy.

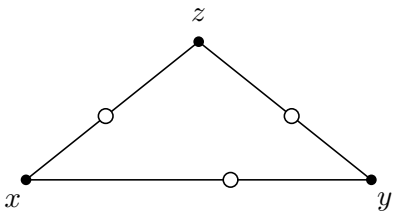


Figure 1: The current visit is at z. The open circles on zx and zy represent earlier dots, and the open circle on xy is another point of the tuple. An extra point can lie beyond either end of its side.

To count these triangles we need an upper bound for the probability that several specified earlier points all receive dots. The decisions at different points depend on one another, but the random draws used to make them are independent. For each point w, draw a number U_w uniformly from $[0, 1]$, independently of the other draws and of the color coins. Place a dot there if $U_w < p_w$. Since $p_w \le a_w$, the indicator I_w of a dot is at most $J_w = \mathbf{1}_{\{U_w < a_w\}}$. The upper bounds a_w are fixed in advance, so the J_w are independent. For every set A of distinct points this gives

$$\mathbb{E}\left[\prod_{w\in A} I_w\right] \le \mathbb{E}\left[\prod_{w\in A} J_w\right] = \prod_{w\in A} a_w. \tag{5}$$

A blank at $z \notin L$ leaves a product P unchanged, and a dot can at most double each of its four or fewer factors. Its possible gain in conditional mean is at most $(2^4 - 1)p_z P = 15p_z P$. Sum this bound over the choices of x, y and the earlier dots on zx and zy whenever a gain is possible. Use a_z at the current point, $2a_w$ for each unvisited point of the tuple, and (5) for the required earlier dots. After averaging, each of the six or seven distinct points contributes one factor a_w, up to a constant for the factors of 2 and the possible assignments of roles.

Write Δ_z for the change in Φ at the visit to z and $\mathbb{E}_z[\,\cdot\,]$ for expectation given the past just before that visit. For a real number s, write $s_+ = \max(s, 0)$, so that $(-\mathbb{E}_z[\Delta_z])_+$ measures a fall in conditional mean. The products removed at z cover the losses in (3). It follows that $\mathbb{E}\,[\sum_z (-\mathbb{E}_z[\Delta_z])_+]$ is bounded by a constant times the sum over all such configurations in the grid, with weight a_w at each point. We will first count them with the simpler starting weights μ_w. By (1), replacing μ_w by a_w on m distinct points of B increases a bound by at most $[2(2N)^\gamma]^m$.

Counting partners on a line

Let $M = \sum_{w\in B} \mu_w$ be the total starting weight in the square. For a point $x \in B$ and a line L through it, let $m_x(L)$ be the total starting weight of the points of B on L other than x. The quantity $m_x(L)^2$ counts the weight of two ordered choices on that line, allowing the same point twice. To bound this count uniformly in x, put

$$D = 1 + \max_{x\in B} \sum_{L\ni x} m_x(L)^2.$$

Only lines that contain another point of B contribute to the sum.

Lemma 4.1. *The total weight satisfies $M = O(N)$, and the line count satisfies $D = O((\log(2N))^3)$.*

Proof. Each of the $2N-1$ layers in B has weight at most 2γ, which proves the bound for M. For the second bound, fix x and order the other points of each line through x by increasing layer number. On a line of height h, the jth such point has layer number at least $(j-1)h$. Indeed, the square reached at the jth point must contain the first j points, and in one coordinate these points are spaced at least h apart. Set aside the first point on each line. Since $\mu_w \le 1/r_w$, the remaining points of a line have total weight at most $(1+\log(2N))/h$ by the harmonic sum.

The points set aside on different lines are distinct, so the sum of their squared weights is at most $\sum_{r<2N}(2r-1)/r^2 = O(\log(2N))$. To put each of these points back, use $(s+t)^2 \le 2s^2+2t^2$. There are at most $4h$ directions of height h, so the contribution from the remaining points is at most a constant times

$$(1+\log(2N))^2 \sum_{h<2N} \frac{4h}{h^2} = O((\log(2N))^3).$$

□

For distinct x, y, let e_{xy} be the weight of all points on their line other than x and y, and set $e_{xx} = 0$. The definition of D gives the three bounds

$$e_{xy} \le \sqrt{D}, \qquad \sum_y \mu_y e_{xy} \le D, \qquad \sum_x \mu_x^2 \le D. \tag{6}$$

The first follows because one line contributes at most D to the sum of squares. For the second, group the possible y by their line through x, so that each line contributes at most $m_x(L)^2$. For the third, fix any root point and expand the line squares there. They contain the squared weight of every other point, and the 1 in the definition of D covers the squared weight of the root.

Lemma 4.2. *The configurations shown in Figure 1, including those with an additional point on the base line, have total starting weight $O(N^{2/3}D^5)$. As a consequence,*

$$\mathbb{E}\left[\sum_{z\in B}(-\mathbb{E}_z[\Delta_z])_+\right] = O(N^{3/4}). \tag{7}$$

Proof. Fix three noncollinear vertices x, y, z, and write $a = e_{xy}$, $b = e_{xz}$ and $c = e_{yz}$. These are exactly the weights of the choices for an extra point on each side line. Thus the contribution of the six points, for these fixed vertices, is $\mu_x\mu_y\mu_z\, abc$. The three sides play identical roles in this count, so we may assume that $\mu_z \le \min(\mu_x, \mu_y)$ and multiply the resulting bound by three.

We estimate three sums, always giving a vertex triple weight $\mu_x\mu_y\mu_z$. First consider ab, which counts two sides through their common vertex x. Allowing all triples of vertices and using (6) gives

$$\sum_{x,y,z} \mu_x\mu_y\mu_z\, ab = \sum_x \mu_x \left(\sum_y \mu_y e_{xy}\right)^2 \le MD^2.$$

The same argument bounds the sum for ac by MD^2, using y as the common vertex.

For the third sum, consider ab^2c^2 and fix x, y. An off-base vertex z is determined by a line through x and a line through y, since those lines meet in at most one point off xy. We can thus sum over all pairs of lines to get

$$\sum_{z \notin xy} b^2c^2 \le \left(\sum_{L \ni x} m_x(L)^2\right)\left(\sum_{L' \ni y} m_y(L')^2\right) \le D^2.$$

Using $\mu_z \le \mu_x$ and then (6), the weighted sum of ab^2c^2 is at most $D^2 \sum_{x,y} \mu_x^2 \mu_y e_{xy} \le D^4$.

To recover the product abc, take the arithmetic mean of three quantities whose geometric mean is abc. For every $t > 0$ the arithmetic–geometric mean inequality gives

$$3abc \le t\,ab + t\,ac + t^{-2}ab^2c^2.$$

Summing and using the three estimates just proved bounds the six-point total by $O(tMD^2 + t^{-2}D^4)$. Take $t = N^{-1/3}$, which balances tN and t^{-2} at $N^{2/3}$. Since $M = O(N)$ and $D \ge 1$, the six-point total is $O(N^{2/3}D^4)$. A seventh point on the original base line adds at most a factor of $\sqrt{D}$, giving the stated bound for both counts.

Finally, each configuration has at most seven points, so replacing its starting weights by the upper bounds a_w introduces a factor of $O(N^{7\gamma})$. By Lemma 4.1, the result is $O(N^{2/3+7\gamma}(\log(2N))^{15})$. Since $2/3 + 7\gamma < 3/4$, this proves (7). □

5 One set for all squares

We have bounded the loss in the conditional mean of Φ. An actual run can also fall below its mean through the randomness of individual visits. We will bound the variance of these changes and use it to estimate how often a shell has too few red points.

Lemma 5.1. *Start with one point and adjoin at most j collinear triples of distinct points, each triple sharing at least one point with those already chosen. For fixed j, all such configurations have total starting weight $O_j(MD^j)$, with each distinct point weighted once.*

Proof. First specify which positions in the triples represent the same point. There are only a bounded number of such specifications when j is fixed. Summing over the first point contributes M. If a triple adds two points at a known point x, the sum of their weights in pairs is at most $\sum_{L \ni x} m_x(L)^2 \le D$. If a triple adds one point, the two known points determine its line and the new weight sums to at most $\sqrt{D}$. A triple whose points are all known adds a factor of 1. Each triple thus contributes a factor at most D, proving the bound. □

Lemma 5.2. *Let Var_z denote variance conditional on the past just before the visit to z. Then*

$$\mathbb{E}\left[\sum_{z \in B} \mathrm{Var}_z(\Delta_z)\right] = O(N^{3/2}).$$

Proof. Fix the past and compare the three possible outcomes of the visit to z. Write Φ_{blank}, Φ_{red} and Φ_{blue} for the values immediately after these outcomes. Variance is at most the mean squared distance from any fixed number, so use the value after a blank as the reference. Only a dot can give a nonzero distance, and its probability is at most a_z. Because the value of Φ before the visit is fixed, this gives

$$\operatorname{Var}_z(\Delta_z) \le a_z \max_{c\in\{\text{red},\text{blue}\}} |\Phi_c - \Phi_{\text{blank}}|^2. \tag{8}$$

If $p_z = 0$, the variance is zero.

Each term of Φ involves either one shell point or one collinear triple or quadruple. To bound the difference in (8), sum the absolute differences of the individual terms. There are two ways a term can change.

- If the term contains z, a shell point at z contributes at most 2. For a product, the difference is bounded by a constant times the bounds on its other factors. Use $2a_w$ for each future point and its actual indicator I_w for each already visited point.
- If the term omits z, some future point v of the term must see an earlier dot at a point s on the line zv. Otherwise a new dot at z would be the first on that line and would have no effect at v. Sum over the choices of v, s and include the factor I_s for this required dot. The earlier indicators in the term stay fixed, and the future factors are bounded by $2a_w$ both before and after the visit. These bounds control the change up to a constant because there are at most four factors.

In the second case the triple z, v, s connects the term to the starting point z. Within either kind of term, four collinear points a, b, c, d are covered by the two triples a, b, c and a, b, d, where we take a to be the point already connected to z. Thus each summand in the bound for the absolute difference involves at most three collinear triples attached to z. Squaring that bound pairs two summands, giving at most six triples. There are at most $1 + 2 \cdot 6 = 13$ distinct points because a new triple adds at most two.

The factor a_z in (8) gives the weight for z itself. Use (5) to average the indicators of all required earlier dots. A repeated indicator has the same value as its square, and a repeated upper bound satisfies $a_w^2 \le a_w$. Thus, up to a constant, we get one factor a_w per distinct point. By (1), replacing this product by a bound in terms of starting weights costs at most a factor of $O(N^{13\gamma})$. Lemma 5.1 with $j = 6$ now bounds the expected sum of variances by

$$O(N^{13\gamma} M D^6) = O(N^{1+13\gamma}(\log(2N))^{18}) = O(N^{3/2}). \qquad \square$$

Proof of Theorem 1.1. At each visit subtract the conditional mean from the actual change, writing $X_z = \Delta_z - \mathbb{E}_z[\Delta_z]$. Then $\mathbb{E}_z[X_z] = 0$. If $u < z$, the value of X_u is already known before the visit to z, so $\mathbb{E}_z[X_u X_z] = X_u \mathbb{E}_z[X_z] = 0$. All cross terms therefore vanish when we average the square of the sum. By Lemma 5.2,

$$\mathbb{E}\left[\left(\sum_z X_z\right)^2\right] = \mathbb{E}\left[\sum_z \operatorname{Var}_z(\Delta_z)\right] = O(N^{3/2}).$$

Cauchy–Schwarz gives $\mathbb{E}\big[\big|\sum_z X_z\big|\big] = O(N^{3/4})$. Adding the possible decreases in conditional mean from (7) gives

$$\mathbb{E}\left[(\Phi_{\text{start}} - \Phi_{\text{end}})_+\right] \le \mathbb{E}\left[\left|\sum_z X_z\right|\right] + \mathbb{E}\left[\sum_z (-\mathbb{E}_z[\Delta_z])_+\right] = O(N^{3/4}).$$

Call a shell unsuccessful if it has fewer than $\gamma N/2$ red points. On this event Lemma 3.1 gives $\Phi_{\text{start}} \ge 3\gamma N/2$ and $\Phi_{\text{end}} < \gamma N$. Thus an unsuccessful shell requires a drop of at least $\gamma N/2$. Markov's inequality bounds the probability of such a drop by its mean positive size divided by $\gamma N/2$. For some absolute constant C and every sufficiently large power of two N, we obtain

$$\mathbb{P}\{T_N \text{ is unsuccessful}\} \le CN^{-1/4}.$$

All these shells belong to one run of the same construction. The probability that at least one shell T_{2^j} with $j \ge J$ is unsuccessful is at most $C\sum_{j\ge J} 2^{-j/4}$ by the union bound. The tail of this geometric series tends to zero. Choose J large enough that the tail bound is less than one, and fix an outcome in which every shell from J onward succeeds.

For every sufficiently large n, choose the power of two N satisfying $2N \le n < 4N$. The entire shell T_N lies inside $[n]^2$, so the same set S satisfies

$$|S \cap [n]^2| \ge |S \cap T_N| \ge \frac{\gamma N}{2} \ge \frac{\gamma}{8} n.$$

This proves the theorem with $c = \gamma/8 = 1/80{,}000$. □

References

[1] H. E. Dudeney, *Amusements in Mathematics*, Thomas Nelson and Sons, 1917, Problem 317, pp. 94 and 222.

[2] K. F. Roth, On a problem of Heilbronn, *J. London Math. Soc.* **26** (1951), 198–204, appendix.

[3] R. R. Hall, T. H. Jackson, A. Sudbery and K. Wild, Some advances in the no-three-in-line problem, *J. Combin. Theory Ser. A* **18** (1975), 336–341.

[4] J. Erde, No three points on a line, Problem 5 in *A collection of open problems in celebration of Imre Leader's 60th birthday*, arXiv:2310.18163, 2023.

[5] B. Green, *100 open problems*, manuscript, Problem 72.

[6] O. Aichholzer, D. Eppstein, and E.-M. Hainzl, Geometric dominating sets—a minimum version of the No-Three-In-Line Problem, *Comput. Geom.* **108** (2023), 101913.

[7] D. T. Nagy, Z. L. Nagy and R. Woodroofe, The extensible No-Three-In-Line problem, *European J. Combin.* **114** (2023), 103796.

[8] A. Ghosal, A note on the extensible no-three-in-line problem, arXiv:2605.07000, 2026.

[9] T. Gábriel, M. Jánosik, D. Melján, and B. Nádor, The extensible no-$(k(n)+1)$-in-line problem, arXiv:2606.02843, 2026.

[10] R. K. Guy and P. A. Kelly, The no-three-in-line problem, *Canad. Math. Bull.* **11** (1968), 527–531.

[11] A. Dhawan, Fractional coloring via entropy, arXiv:2603.17730, 2026.