No-$(k+1)$-in-line problem for large constant k

Alexandr Grebennikov and Matthew Kwan

Abstract. How many points can be placed in an $n \times n$ grid so that every (affine) line contains at most k points? We prove that for $n \geqslant k \geqslant 10^{37}$ the maximum number of points is exactly kn. Our proof builds on the recent work of Kovács, Nagy, and Szabó (who proved an analogous result when k is at least about $\sqrt{n \log n}$), incorporating ideas of Jain and Pham. Using the same approach, we also obtain new bounds for higher-dimensional extensions of this problem.

1. Introduction

The *no-three-in-line* problem, posed by Dudeney at the beginning of the 20th century [7], asks for the maximum number of points that can be placed on an $n \times n$ grid such that no three points are collinear. Despite significant attention over the years, the problem remains open: the best known upper bound $2n$ comes from the observation that each of the n horizontal lines can contain at most two points, while the best known lower bound $(1.5-o(1))n$ comes from the *modular hyperbola* construction of Hall, Jackson, Sudbery, and Wild [17]. For more history and background, see for example the surveys by Brass, Moser, and Pach [5] and Eppstein [8], and Green's list of open problems [14].

Several different conjectures were made concerning the asymptotic behaviour of the answer to the no-three-in-line problem, including suggestions that it could be roughly $1.5n$ [14], roughly $2n$ [5], or somewhere in between [8, 15]. For small values of n, examples attaining the trivial upper bound $2n$ are known [2, 10, 11].

A natural generalisation of this problem, first studied by Brass and Knauer [4] in a more general context, is to ask for the maximum size of a subset of the $n \times n$ grid containing at most k points on each line. Denoting this maximum by $f_k(n)$, the trivial upper bound is $f_k(n) \leqslant kn$. Lefmann [25, Proposition 2] proved that $f_k(n) = \Omega(kn)$ for every $k \geqslant 2$. Recently, Kovács, Nagy, and Szabó [24] showed that $f_k(n) = kn$ for each $k \geqslant C\sqrt{n \log n}$ (when $C > 12.5$ and n is sufficiently large in terms of C). They also noted [24, Section 4] that, conditionally on some strong results about random regular graphs (i.e., the bipartite version of the Kim–Vu sandwich conjecture [21]; see [3, 12, 22] for recent progress), their approach could yield an analogous statement for k of order at least $\log n$.

We strengthen these results by proving that $f_k(n) = kn$ for every $n \geqslant k \geqslant K_0$, for some absolute constant K_0 (our proof gives $K_0 = 10^{37}$, but we did not attempt to optimise it).

Theorem 1.1. *Let n, k be integers such that $10^{37} \leqslant k \leqslant n$. Then there exists a subset S of the $n \times n$ grid of size kn such that every line contains at most k points of S.*

In a different work, Kovács, Nagy, and Szabó [23] studied this problem in the regime when k is small compared to n, using randomised algebraic constructions. In particular, they demonstrated that $f_k(n) \geqslant (k-3)n$ when n is sufficiently large in terms of k, and conjectured that $f_k(n) = kn + o(n)$ when $k = k(n) = \omega(1)$ and $n \to \infty$. Theorem 1.1 confirms this conjecture in a strong sense.

Interestingly, all known linear-size constructions for the no-three-in-line problem are algebraic in nature (though there are random constructions [9, 13] of size about $n/\sqrt{\log n}$). In a related question, Green [14] asked whether every "large" no-three-in-line set reduces to an algebraic curve modulo some prime. Although Theorem 1.1

Both authors are supported by ERC Starting Grant "RANDSTRUCT" No. 101076777.

only gives no-$(k+1)$-in-line sets for large k, we remark that our constructions are purely probabilistic and do not exhibit any algebraic structure.

1.1. **Proof ideas.** Let $\mathcal{L}$ be the set of lines intersecting the $n \times n$ grid in at least two points. For each line $L \in \mathcal{L}$, define its *weight* as

$$w(L) := \frac{|[n]^2 \cap L|}{n} \in [2/n, 1].$$

The framework of Kovács, Nagy, and Szabó [24] allows us to reduce Theorem 1.1 to Theorem 1.2 below, which permits a $0.01k$ excess over the expected number of points on the lines that are not horizontal or vertical. Therefore, the main content of this paper is the proof of Theorem 1.2 (though, for the reader's convenience, we also present a deduction of Theorem 1.1 in Section 5).

Theorem 1.2. *Let n, k be integers such that $10^{36} \leqslant k \leqslant 0.9 \cdot n$. Then there exists a subset $S \subseteq [n]^2$ that contains exactly k points on each horizontal and each vertical line, and at most $k \cdot (w(L) + 0.01)$ points on every other line $L \in \mathcal{L}$.*

Kovács, Nagy, and Szabó [24] were able to prove a result along the lines of Theorem 1.2 under the much stronger assumption that k is at least about $\sqrt{n \log n}$. They did this by choosing the desired point set *randomly*, using concentration inequalities to show that the number of points on each line is unlikely to deviate much from its expectation. This works when k grows sufficiently fast as a function of n but breaks down for constant k: the main issue is that a union bound over all lines becomes too inefficient.

To overcome this issue, we build on ideas from the work of Jain and Pham [18] on optimal thresholds for Latin squares in random hypergraphs. Namely, we design an iterative subsampling procedure that uses the Lovász Local Lemma at each step, maintaining control over the number of points on each line (roughly speaking, we maintain tight control on the number of points on lines whose weight exceeds some threshold, gradually raising this threshold as the procedure continues and the light lines become increasingly unimportant). Moreover, at each step of the procedure, instead of constructing a single configuration of points, we construct a "spread" probability distribution over suitable configurations. We then use this spread property to show that a sample from the final distribution obtained from our procedure is likely to satisfy a certain Hall-type condition, which provides a subset containing exactly k points on each horizontal and each vertical line.

1.2. **Higher dimensions.** A further generalisation of this problem, introduced by Brass and Knauer [4], is to ask for the maximum size of a subset of the d-dimensional grid $[n]^d$ that contains at most k points in each affine subspace of dimension t (where $1 \leqslant t \leqslant d-1$). Denoting this maximum by $f_{k,d,t}(n)$, the trivial upper bound is $f_{k,d,t}(n) \leqslant kn^{d-t}$.

Brass and Knauer [4, Lemma 8] proved that $f_{k,d,t}(n) = \Omega(n^{d-t-(d(t+1))/k})$, which was slightly sharpened by Lefmann [25, Lemma 2] to $\Omega(n^{d-t-(t(d+1))/k})$. Later, Sudakov and Tomon [29, Theorem 1.4] established the optimal bound $f_{k,d,t}(n) = \Omega_d(n^{d-t})$ when k is sufficiently large in terms of d, and Ghosal, Goenka, and Keevash [13, Theorem 1.1] extended this result to all $k \geqslant d+1$.

However, all existing lower bounds differ from the trivial upper bound by a multiplicative factor of order k. We close this gap asymptotically as $k \to \infty$ by showing that in this regime $f_{k,d,t}(n) = (1 - o_d(1))kn^{d-t}$.

Theorem 1.3. *For an integer d and $\varepsilon > 0$, there exists $K_1 = K_1(\varepsilon, d)$ such that the following holds. For all integers n, k, t such that $K_1 \leqslant k \leqslant n^t$ and $1 \leqslant t \leqslant d-1$, there exists a subset $S \subseteq [n]^d$ that contains at most k points in each t-dimensional affine subspace of $\mathbb{R}^d$, and contains at least $(1-\varepsilon)k$ points in each axis-aligned t-dimensional affine subspace of $\mathbb{R}^d$ that intersects $[n]^d$. As a consequence,*

$$f_{k,d,t}(n) \geqslant (1-\varepsilon)kn^{d-t}.$$

The proof of Theorem 1.3 follows the same overall scheme as the proof of Theorem 1.2, but is significantly simpler (because it does not aim for an exact bound). So, the rest of the paper is organised as follows. After reviewing several known results in Section 2, we present the short proof of Theorem 1.3 in Section 3. Then, we prove Theorem 1.2 in Section 4, use it to derive Theorem 1.1 in Section 5, and discuss related questions and further directions in Section 6.

Acknowledgements. We would like to thank Huy Pham for insightful discussions about the proof approach.

2. Preliminaries

Notation. For a positive integer n, we write $[n] = \{1, \ldots, n\}$. For non-negative reals a, b we use the shorthand $a \pm b$ to denote the interval $[a-b, a+b]$. For an event $\mathcal{E}$ in some probability space, we denote its complement by $\overline{\mathcal{E}}$. For functions $f = f(n)$ and $g = g(n)$, we write $f = O(g)$ to mean that there is a constant C such that $|f| \leqslant C|g|$, $f = \Omega(g)$ to mean that there is a constant $c > 0$ such that $f(n) \geqslant c|g(n)|$ for sufficiently large n. Subscripts on asymptotic notation indicate quantities that should be treated as constants. For a (hyper)graph G, we write $V(G)$ and $E(G)$ for its vertex and edge sets, respectively. For a graph G and subsets $A, B \subseteq V(G)$, we write $E_G(A, B)$ for the set of edges between A and B in G, and $e_G(A, B) = |E_G(A, B)|$. For a vertex v of a graph G, we write $\deg_G(v)$ for its degree in G.

We use the following standard version of the Chernoff bound.

Proposition 2.1. *Let $X_1, \ldots, X_n$ be independent* $\mathrm{Ber}(p)$ *random variables. Then, for every $\delta \in (0, 1)$,*

$$\mathbb{P}[X_1 + \ldots + X_n \notin (1 \pm \delta)pn] \leqslant 2\exp\left(-\frac{\delta^2 pn}{3}\right).$$

The following version of the Lovász Local Lemma (also employed in [16, 18]) provides control over an ensemble of independent random variables, conditioned on the non-occurrence of a family of "bad events". Its proof follows from the standard inductive proof of the Lovász Local Lemma (see [1, Chapter 5.1]).

Proposition 2.2 ([18, Proposition 6])**.** *Given independent random variables $\{\xi_i\}_{i \in I}$ and events $\{\mathcal{E}_j\}_{j \in J}$, where each event $\mathcal{E}_j$ depends on a subset $X_j \subseteq I$ of variables. Assume each event has probability at most q, and that each set X_j intersects at most Δ other sets from $\{X_{j'}\}_{j' \in J}$. If $4q\Delta \leqslant 1$, then*

1. *$\mathbb{P}\big[\bigcap_{j \in J} \overline{\mathcal{E}_j}\,\big] > 0$, and*
2. *for every event $\mathcal{E}$ depending on a subset of variables $X \subseteq I$ which intersects at most M of the sets $\{X_j\}_{j \in J}$,*

$$\mathbb{P}\Big[\mathcal{E} \,\Big|\, \bigcap_{j \in J} \overline{\mathcal{E}_j}\Big] \leqslant \mathbb{P}[\mathcal{E}] \cdot \exp(6qM).$$

Also, we recall the classical Ore–Ryser criterion for the existence of a k-regular spanning subgraph in a bipartite graph (see e.g. [26, §7, Ex. 16]), which also follows from the max-flow–min-cut theorem.

Proposition 2.3. *A bipartite graph G with parts A_0 and B_0 of equal size has a k-regular spanning subgraph if and only if for all subsets $A \subseteq A_0$ and $B \subseteq B_0$, we have*

$$e_G(A, B_0 \setminus B) \geqslant k(|A| - |B|).$$

Finally, we need the following facts about the number of integer points on lines in the plane (Proposition 2.4), and, more generally, on t-dimensional affine subspaces in $\mathbb{R}^d$ (Proposition 2.5).

Proposition 2.4. *Let $\mathcal{L}$ be the set of lines in $\mathbb{R}^2$ intersecting the grid $[n]^2$ in at least two points. Then, for every point $(x_1, x_2) \in [n]^2$ and $\alpha \in (0, 1]$,*

$$\big|\big\{L \in \mathcal{L} : (x_1, x_2) \in L,\ |[n]^2 \cap L| \geqslant \alpha n\big\}\big| \leqslant 18\alpha^{-2}.$$

Proof. If $\alpha \leqslant 2/n$, then the total number of lines in $\mathcal{L}$ containing (x_1, x_2) is at most $n^2 \leqslant 4\alpha^{-2}$. Otherwise, let $\ell := \lceil 2\alpha^{-1} \rceil \leqslant n$. If (x_1, x_2) is the only integer point within the rectangle $[x_1, x_1+\ell-1] \times [x_2-(\ell-1), x_2+(\ell-1)]$ on some line $L \in \mathcal{L}$, then

$$w(L) = \frac{|[n]^2 \cap L|}{n} < \frac{n/\ell + 1}{n} \leqslant \frac{2n/\ell}{n} \leqslant \alpha.$$

Therefore, since every two points determine a unique line, the desired number of lines is bounded by the number of integer points in this rectangle, which is at most $2\ell^2 \leqslant 2(2\alpha^{-1}+1)^2 \leqslant 18\alpha^{-2}$. □

Proposition 2.5. *Let n, d, t be integers such that $1 \leqslant t \leqslant d-1$, and let $\mathcal{F}_{d,t}$ denote the set of t-dimensional affine subspaces of $\mathbb{R}^d$. Then, for every point $x \in [n]^d$ and $\alpha \in (0,1]$,*

$$\left|\left\{F \subseteq [n]^d : x \in F,\ |F| \geqslant \alpha n^t,\ F = F' \cap [n]^d \text{ for some } F' \in \mathcal{F}_{d,t}\right\}\right| = O_d(\alpha^{-d}). \tag{1}$$

Proof. Consider a set $F = F' \cap [n]^d$ satisfying the conditions in (1), and let $\Gamma \subseteq \mathbb{Z}^d$ be the lattice generated by $F - x$. Note that $F \subseteq (\Gamma + x) \cap [n]^d \subseteq F' \cap [n]^d = F$, and hence $F = (\Gamma + x) \cap [n]^d$. In particular, F is uniquely determined by Γ, and

$$|\Gamma \cap [-n,n]^d| \geqslant |F| \geqslant \alpha n^t. \tag{2}$$

Denote the rank of Γ by $t' \leqslant t$, and let $v_1, \dots, v_{t'} \in F - x \subseteq [-n,n]^d$ be some generating vectors of Γ. Also let $V = \operatorname{span}_{\mathbb{R}}(v_1, \dots, v_{t'})$, let $D = \left\{\sum_{i=1}^{t'} c_i v_i : c_1, \dots, c_{t'} \in [0,1)\right\} \subseteq V$ be the fundamental domain of Γ, and let $\det \Gamma = \operatorname{vol}_{t'}(D)$ be the determinant[1] of Γ. Then, for every two distinct points $y_1, y_2 \in \Gamma \cap [-n,n]^d$, the translates $y_1 + D$ and $y_2 + D$ are disjoint and contained in $[-(t'+1)n, (t'+1)n]^d \cap V$. Hence, $\bigsqcup_{y \in \Gamma \cap [-n,n]^d} (y + D)$ is contained in the t'-dimensional Euclidean ball in V of radius $\sqrt{d}(t'+1)n$ centred at the origin. Therefore, as $t' \leqslant t \leqslant d$,

$$|\Gamma \cap [-n,n]^d| \leqslant \frac{\operatorname{vol}_{t'}(B(0, \sqrt{d}(t'+1)n))}{\operatorname{vol}_{t'}(D)} = O_d\left(\frac{n^t}{\det \Gamma}\right).$$

Combining this with (2), we conclude that $\det \Gamma = O_d(\alpha^{-1})$. The statement now follows from a result of Schmidt [28, Theorem 2], which states that for every $M \geqslant 0$, the number of lattices $\Gamma \subseteq \mathbb{Z}^d$ with $\det \Gamma \leqslant M$ is $O_d(M^d)$. □

3. Proof of Theorem 1.3

It might be possible to prove the bound $f_{k,d,t} \geqslant (1-\varepsilon) k n^{d-t}$ (for k sufficiently large in terms of ε, d) using some type of probabilistic deletion method, in the spirit of [13]. However, we take a different approach based on repeated applications of the Lovász Local Lemma. This gives the stronger "almost regularity" property in the statement of Theorem 1.3 (that every axis-aligned subspace has about k points), and serves as a warm-up for the proof of Theorem 1.2.

Our proof of Theorem 1.3 does not rely heavily on the specific structure of affine subspaces intersecting the integer grid. The following definition captures the property of main importance to us.

Definition 3.1 (polynomial tails)**.** We say that a hypergraph $\mathcal{H}$ (of mixed uniformity) *has polynomial tails (with parameters $N, C \in \mathbb{N}$)*, if every edge of $\mathcal{H}$ contains at most N vertices, and for every vertex $v \in V(\mathcal{H})$ and $\alpha \in (0,1]$,

$$|\{F \in E(\mathcal{H}) : v \in F,\ w(F) \geqslant \alpha\}| \leqslant C\alpha^{-C},$$

where $w(F) := |F|/N$ is the *weight* of an edge F.

[1] Recall that the determinant $\det \Gamma$ of a rank-t' lattice Γ is the t'-dimensional volume of its fundamental domain. Equivalently, $\det \Gamma = \sqrt{\det G}$, where $G = (\langle v_i, v_j \rangle)_{1 \leqslant i,j \leqslant t'}$ is the Gram matrix of some generating vectors $v_1, \dots, v_{t'}$ of Γ.

Definition 3.2. Let n, d, t be integers such that $1 \leqslant t \leqslant d-1$, and let $\mathcal{F}_{d,t}$ denote the set of t-dimensional affine subspaces of $\mathbb{R}^d$. Define $\mathcal{H}_{n,d,t}$ as the hypergraph on the vertex set $[n]^d$ with

$$E(\mathcal{H}_{n,d,t}) = \{F \subseteq [n]^d : F = F' \cap [n]^d \text{ for some } F' \in \mathcal{F}_{d,t}\}.$$

Since every t-dimensional affine subspace of $\mathbb{R}^d$ intersects $[n]^d$ in at most n^t points, Proposition 2.5 immediately implies that $\mathcal{H}_{n,d,t}$ has polynomial tails with $N = n^t$ and some $C = C(d)$.

Definition 3.3 (good set)**.** Let $\mathcal{H}$ be a hypergraph having polynomial tails (with parameters N and C). For a real number $m \in [1, N]$, we say that a set $S \subseteq V(\mathcal{H})$ is *m-good* if it satisfies the following conditions:

(1) (concentration for "heavy" edges) For every edge $F \in E(\mathcal{H})$ with $w(F) \in (m^{-2/3}, 1]$, we have

$$|S \cap F| \in (1 \pm m^{-1/12}) m \cdot w(F).$$

(2) (upper bound for "medium" edges) For every edge $F \in E(\mathcal{H})$ with $w(F) \in (m^{-2}, m^{-2/3}]$, we have

$$|S \cap F| \leqslant (1 + m^{-1/12}) m^{1/3}.$$

(3) (very few vertices in "light" edges) For every edge $F \in E(\mathcal{H})$ with $w(F) \leqslant m^{-2}$, we have

$$|S \cap F| \leqslant 6C + 3.$$

Theorem 3.4. *Let $\mathcal{H}$ be a hypergraph having polynomial tails (with parameters N and C). Then there is $K = K(C)$ such that for every m with $K \leqslant m \leqslant N$, there exists an m-good subset of $V(\mathcal{H})$.*

Proof of Theorem 1.3 assuming Theorem 3.4. Recall from Definition 3.2 that the hypergraph $\mathcal{H}_{n,d,t}$ has polynomial tails with parameters $N = n^t$ and some $C = C(d)$. We may assume that $\varepsilon < 1/2$, and that k is sufficiently large in terms of ε, d, C. Then, by Theorem 3.4, there exists an $(1 - \varepsilon/2)k$-good subset of $S \subseteq [n]^d$.

We claim that S has the desired properties. Indeed, if F' is a t-dimensional affine subspace, then $F = F' \cap [n]^d$ is an edge of $\mathcal{H}_{n,d,t}$. Thus, by conditions (1), (2), (3) from Definition 3.3,

$$|S \cap F| \leqslant \max\Big((1 + ((1 - \varepsilon/2)k)^{-1/12})(1 - \varepsilon/2)k, 6C + 3\Big) \leqslant k.$$

For the lower bound, we note that if F' is an axis-aligned t-dimensional affine subspace intersecting $[n]^d$, then the corresponding edge $F = F' \cap [n]^d$ of $\mathcal{H}_{n,d,t}$ has weight 1. Therefore, by condition (1) from Definition 3.3,

$$|S \cap F| \geqslant (1 - ((1 - \varepsilon/2)k)^{-1/12})(1 - \varepsilon/2)k \geqslant (1 - \varepsilon)k.$$

This completes the proof. □

To prove Theorem 3.4, we consider a sequence of real numbers $m_{(1)} < m_{(2)} < \ldots < m_{(r)}$ defined as

$$m_{(1)} = m, \quad m_{(i+1)} = m_{(i)}^3, \quad N^{1/3} < m_{(r)} \leqslant N. \tag{3}$$

Then, we prove the existence of an $m_{(i)}$-good set for each $i = r, r-1, \ldots, 1$ via a downward induction, applying the Lovász Local Lemma (Proposition 2.2(1)) at each step. This way, Theorem 3.4 easily follows from Lemma 3.5 below.

Lemma 3.5. *Let $\mathcal{H}$ be a hypergraph having polynomial tails (with parameters N and C). Let m_0, m be real numbers such that $K \leqslant m \leqslant m_0 \leqslant N$ for some $K = K(C)$. Suppose that either*

(a) *(induction step) $m_0 = m^3$ and S_0 is an m_0-good subset of $V(\mathcal{H})$, or*

(b) *(base case) $m_0 = N \leqslant m^3$ and $S_0 = V(\mathcal{H})$ (which is trivially m_0-good).*

Then there exists an m-good set $S \subseteq S_0$.

Proof. Let $\delta := m^{-1/12}$. We will assume that m is sufficiently large in terms of C (and, in particular, that $\delta \leqslant 0.01$). We claim that for an edge $F \in E(\mathcal{H})$,

$$|S_0 \cap F| \in (1 \pm \delta^3)m_0 \cdot w(F) \quad \text{ if } w(F) \in (m^{-2}, 1]; \tag{4}$$

$$|S_0 \cap F| \leqslant (1 + \delta^3)m_0 \cdot m^{-2} \quad \text{ if } w(F) \in (m_0^{-2}, m^{-2}]. \tag{5}$$

Indeed, in case (a), since $m_0 = m^3$ and $\delta^3 = m_0^{-1/12}$, this follows from conditions (1) and (2) from the definition of an m_0-good set. On the other hand, in case (b) we have $m_0 = N$ and $S_0 = V(\mathcal{H})$, and hence $|S_0 \cap F| = m_0 \cdot w(F)$ for every edge $F \in E(\mathcal{H})$ by the definition of weight.

Let $p := m/m_0$, and let S be a random subset of S_0 obtained by including each vertex independently with probability p. We will show that S is m-good with positive probability, through an application of the Lovász Local Lemma (Proposition 2.2(1)). First, we identify the bad events and estimate their probabilities.

Heavy edges. For an edge $F \in E(\mathcal{H})$ with $w(F) \in (m^{-2/3}, 1]$, let $\mathcal{E}_1(F)$ denote the bad event that S does not satisfy condition (1) for this edge F: that is, $|S \cap F| \notin (1 \pm \delta)m \cdot w(F)$. By (4), for such edges F we have $|S_0 \cap F| \in (1 \pm \delta^3)m_0 \cdot w(F)$. Thus, if $\mathcal{E}_1(F)$ occurs, then

$$|S \cap F| \notin (1 \pm \delta/2)p|S_0 \cap F|.$$

Therefore, by the Chernoff bound (Proposition 2.1),

$$\mathbb{P}[\mathcal{E}_1(F)] \leqslant 2\exp\left(-\frac{(\delta/2)^2 p|S_0 \cap F|}{3}\right) \leqslant 2\exp\left(-\frac{(\delta/2)^2(1-\delta^3)m \cdot w(F)}{3}\right) \leqslant 2\exp\left(-m^{1/6}/13\right).$$

Medium edges. For an edge $F \in E(\mathcal{H})$ with $w(F) \in (m^{-2}, m^{-2/3}]$, let $\mathcal{E}_2(F)$ denote the bad event that S does not satisfy condition (2) for this edge F. By (4), for such edges F we have

$$|S_0 \cap F| \leqslant m' := \lfloor (1+\delta^3)m_0 \cdot m^{-2/3} \rfloor.$$

Let X be a sum of m' independent $\mathrm{Ber}(p)$ random variables. Then, by the Chernoff bound (Proposition 2.1),

$$\begin{aligned}\mathbb{P}[\mathcal{E}_2(F)] = \mathbb{P}\Big[|S \cap F| > (1+\delta)m^{1/3}\Big] &\leqslant \mathbb{P}\Big[X > (1+\delta)m^{1/3}\Big] \leqslant \mathbb{P}[X \notin (1 \pm \delta/2)pm'] \\ &\leqslant 2\exp\left(-\frac{(\delta/2)^2 pm'}{3}\right) \leqslant 2\exp\left(-\frac{(\delta/2)^2(m^{1/3}-1)}{3}\right) \leqslant 2\exp\left(-m^{1/6}/13\right).\end{aligned}$$

Light edges. Note that for edges F with $w(F) \leqslant m_0^{-2}$, condition (3) for S is implied by condition (3) for S_0, so it suffices to consider only edges with $w(F) \in (m_0^{-2}, m^{-2}]$. Let $\mathcal{E}_3(F)$ denote the bad event that S does not satisfy condition (3) for such an edge F. By (5), we have $|S_0 \cap F| \leqslant (1+\delta^3)m_0 m^{-2}$. Therefore, denoting $C_1 := 6C + 4$,

$$\mathbb{P}[\mathcal{E}_3(F)] = \mathbb{P}[|S \cap F| \geqslant C_1] \leqslant \binom{|S_0 \cap F|}{C_1} \cdot p^{C_1} \leqslant \big((1+\delta^3)m_0 m^{-2} \cdot m/m_0\big)^{C_1} \leqslant (2/m)^{C_1}.$$

Since m is sufficiently large in terms of C, the maximum probability q of a bad event then satisfies

$$q \leqslant \max\Big(2\exp(m^{-1/6}/13),\ (2/m)^{6C+4}\Big) = (2/m)^{6C+4}.$$

Degree of the dependency graph. So, we have exactly one bad event ($\mathcal{E}_1$, $\mathcal{E}_2$, or $\mathcal{E}_3$) for each edge $F \in E(\mathcal{H})$ with $w(F) \in (m_0^{-2}, 1]$. Note that each of these events depends only on the Bernoulli random variables associated with the underlying vertices of $S_0 \cap F$, and the number of such vertices satisfies $|S_0 \cap F| \leqslant (1+\delta^3)m_0$ by (4) and (5).

Recall that $\mathcal{H}$ has polynomial tails (with parameters N and C). Therefore, for every vertex $v \in V(\mathcal{H})$ we can bound the number of edges with large weight containing v as follows:

$$\big|\{F \in E(\mathcal{H}) : v \in F,\ w(F) \geqslant m_0^{-2}\}\big| \leqslant C m_0^{2C}.$$

As $m_0 \leqslant m^3$, the maximum degree Δ of the dependency graph of our bad events then satisfies

$$\Delta \leqslant (1+\delta^3)m_0 \cdot C m_0^{2C} \leqslant 2Cm^{6C+3}, \quad \text{ and thus } \quad 4q\Delta \leqslant C2^{6C+7}/m.$$

Since m is sufficiently large in terms of C, Proposition 2.2(1) completes the proof. □

Proof of Theorem 3.4. Consider a sequence of real numbers $m = m_{(1)} < m_{(2)} < \ldots < m_{(r)}$ satisfying (3). By Lemma 3.5(b), there exists an $m_{(r)}$-good subset of $V(\mathcal{H})$. Iteratively applying Lemma 3.5(a), we conclude that there exists an $m_{(1)}$-good subset of $V(\mathcal{H})$, as required. □

4. Proof of Theorem 1.2

In this section, our goal is to prove Theorem 1.2. Here we do not use the language of hypergraphs with polynomial tails: though some parts of our argument can be translated to that setting, other parts substantially rely on the fact that the grid $[n]^2$ can be interpreted as the complete bipartite graph $K_{n,n}$. Definition 4.1 below provides a suitable analogue of good sets from Definition 3.3.

Recall that $\mathcal{L}$ denotes the set of lines in $\mathbb{R}^2$ intersecting the grid $[n]^2$ in at least two points, and that for every line $L \in \mathcal{L}$,

$$w(L) = \frac{|[n]^2 \cap L|}{n}.$$

Definition 4.1 (nice set)**.** For real number $m \in [1, n]$, we say that a set $S \subseteq [n]^2$ is *m-nice* if it satisfies the following conditions:

(1) (concentration for "heavy" lines) For every line $L \in \mathcal{L}$ with $w(L) \in (m^{-2/3}, 1]$, we have
$$|S \cap L| \in (1 \pm m^{-1/12})m \cdot w(L).$$

(2) (upper bound for "medium" lines) For every line $L \in \mathcal{L}$ with $w(L) \in (m^{-2}, m^{-2/3}]$, we have
$$|S \cap L| \leqslant (1 + m^{-1/12})m^{1/3}.$$

(3) (very few points on "light" lines) For every line $L \in \mathcal{L}$ with $w(L) \leqslant m^{-2}$, we have
$$|S \cap L| \leqslant C := 14.$$

(4) (quasirandomness) For every pair of subsets $I, J \subseteq [n]$ of size at least $n/10$, we have
$$|S \cap (I \times J)| \in (1 \pm m^{-1/12})|I||J| \cdot m/n.$$

For the rest of this section, we fix $K := 10^{36}$ and $\varepsilon := 0.005$.

Conditions (1), (2), (3) from Definition 4.1 imply that, for $k \geqslant K$, a $(1+\varepsilon)k$-nice set contains at most $k \cdot (w(L) + 0.01)$ points on each line $L \in \mathcal{L}$. To prove Theorem 1.2, we will take such a set and "regularise" it: that is, we will find a subset containing exactly k points on each horizontal and each vertical line.

This regularisation step is the reason we need the quasirandomness condition in Definition 4.1. However, this quasirandomness condition is not enough on its own, and we need a further idea of Jain and Pham [18]. Indeed, instead of constructing a single nice set, we construct a sufficiently "spread" probability distribution supported on nice sets, and prove that a sample from this distribution can be regularised with positive probability.

Formally, we say that a probability distribution on subsets of $[n]^2$ is *p-spread* if, for every subset $T \subseteq [n]^2$ and a sample S from this distribution,

$$\mathbb{P}[T \subseteq S] \leqslant p^{|T|}.$$

Our proof of Theorem 1.2 follows the same general strategy as the proof of Theorem 1.3. Namely, given $n \geqslant k \geqslant K$, we consider a sequence of real numbers $m_{(1)} < m_{(2)} < \ldots < m_{(r)}$ such that

$$m_{(1)} = (1+\varepsilon)k, \quad m_{(i+1)} = m_{(i)}^3, \quad n^{1/3} < m_{(r)} \leqslant n, \tag{6}$$

and find $(m_{(i)} + 1)/n$-spread distributions supported on $m_{(i)}$-nice sets for $i = r, r-1, \ldots, 1$ via a downward induction. We handle the base case and the induction step of this procedure in Lemma 4.2(b) and Lemma 4.2(a), respectively. Then, in Lemma 4.3, we perform the final regularisation step: specifically, we show that a sample from the $(m_{(1)}+1)/n$-spread distribution supported on $m_{(1)}$-nice sets has a subset containing exactly k points on each horizontal and each vertical line with positive probability.

Lemma 4.2. *Let n be an integer and m_0, m be real numbers such that $K \leqslant m \leqslant m_0 \leqslant n$. Suppose that either*

(a) *$m_0 = m^3$ and there exists an $(m_0+1)/n$-spread distribution supported on m_0-nice sets, or*

(b) *$m_0 = n \leqslant m^3$ (in this case, we have the trivial 1-spread distribution supported on the m_0-nice set $[n]^2$).*

Then there exists an $(m+1)/n$-spread distribution supported on m-nice sets.

Proof. Let $\delta := m^{-1/12} \leqslant 0.001$. Let S_0 be a random m_0-nice set sampled from the "original distribution" (provided by (a) or (b)). We claim that for each line $L \in \mathcal{L}$,

$$|S_0 \cap L| \in (1 \pm \delta^3)m_0 \cdot w(L) \quad \text{if } w(L) \in (m^{-2}, 1]; \tag{7}$$

$$|S_0 \cap L| \leqslant (1 + \delta^3)m_0 \cdot m^{-2} \quad \text{if } w(L) \in (m_0^{-2}, m^{-2}]. \tag{8}$$

Indeed, in case (a), since $m_0 = m^3$ and $\delta^3 = m_0^{-1/12}$, this follows from conditions (1) and (2) from the definition of an m_0-nice set. On the other hand, in case (b), we have $m_0 = n$ and $S_0 = [n]^2$, and hence $|S_0 \cap L| = m_0 \cdot w(L)$ for every line L by the definition of weight.

Let $p := m/m_0$, and let S be a random subset of S_0 obtained by including each point with probability p (independently of each other and of the randomness of S_0). To apply the Lovász Local Lemma (Proposition 2.2), we first identify the bad events and estimate their probabilities (these calculations are almost identical to those in the proof of Lemma 3.5).

Heavy lines. For a line $L \in \mathcal{L}$ with $w(L) \in (m^{-2/3}, 1]$, let $\mathcal{E}_1(L)$ denote the bad event that S does not satisfy condition (1) from Definition 4.1 for this line L: that is, $|S \cap L| \notin (1 \pm \delta)m \cdot w(L)$. By (7), for such lines L we have $|S_0 \cap L| \in (1 \pm \delta^3)m_0 \cdot w(L)$. Thus, if $\mathcal{E}_1(L)$ occurs, then

$$|S \cap L| \notin (1 \pm \delta/2)p|S_0 \cap L|.$$

Therefore, by the Chernoff bound (Proposition 2.1),

$$\mathbb{P}[\mathcal{E}_1(L)] \leqslant 2\exp\left(-\frac{(\delta/2)^2 p|S_0 \cap L|}{3}\right) \leqslant 2\exp\left(-\frac{(\delta/2)^2(1-\delta^3)m \cdot w(L)}{3}\right) \leqslant 2\exp\left(-m^{1/6}/13\right).$$

Medium lines. For a line $L \in \mathcal{L}$ with $w(L) \in (m^{-2}, m^{-2/3}]$, let $\mathcal{E}_2(L)$ denote the bad event that S does not satisfy condition (2) for this line L. By (7), for such lines L we have

$$|S_0 \cap L| \leqslant m' := \lfloor (1+\delta^3)m_0 \cdot m^{-2/3} \rfloor.$$

Let X be a sum of m' independent $\mathrm{Ber}(p)$ random variables. Then, by the Chernoff bound (Proposition 2.1),

$$\begin{aligned}\mathbb{P}[\mathcal{E}_2(L)] &= \mathbb{P}\Big[|S \cap L| > (1+\delta)m^{1/3}\Big] \leqslant \mathbb{P}\Big[X > (1+\delta)m^{1/3}\Big] \leqslant \mathbb{P}[X \notin (1 \pm \delta/2)pm'] \\ &\leqslant 2\exp\left(-\frac{(\delta/2)^2 pm'}{3}\right) \leqslant 2\exp\left(-\frac{(\delta/2)^2(m^{1/3}-1)}{3}\right) \leqslant 2\exp\left(-m^{1/6}/13\right).\end{aligned}$$

Light lines. Note that for lines with $w(L) \leqslant m_0^{-2}$, condition (3) for S is implied by condition (3) for S_0, so it suffices to consider only lines with $w(L) \in (m_0^{-2}, m^{-2}]$. Let $\mathcal{E}_3(L)$ denote the bad event that S does not satisfy condition (3) for such line L. By (8), we have $|S_0 \cap L| \leqslant (1+\delta^3)m_0 m^{-2}$. Therefore, recalling that $C = 14$,

$$\mathbb{P}[\mathcal{E}_3(L)] = \mathbb{P}[|S \cap L| \geqslant C+1] \leqslant \binom{|S_0 \cap L|}{C+1} \cdot p^{C+1} \leqslant \left(\frac{e \cdot (1+\delta^3)m_0 m^{-2}}{C+1} \cdot p\right)^{C+1} \leqslant \left(\frac{1}{5m}\right)^{15}.$$

(Here we used the standard estimate $\binom{n}{k} \leqslant (en/k)^k$.)

Since $m \geqslant K = 10^{36}$, the maximum probability q of a bad event then satisfies

$$q \leqslant \max(2\exp(-m^{1/6}/13), (1/(5m))^{15}) = (1/(5m))^{15}.$$

Passing to the conditional distribution. So, we have exactly one bad event ($\mathcal{E}_1$, $\mathcal{E}_2$, or $\mathcal{E}_3$) for each line L with weight in $(m_0^{-2}, 1]$. Note that each of the events $\mathcal{E}_1(L)$, $\mathcal{E}_2(L)$, and $\mathcal{E}_3(L)$ depends only on the Bernoulli random variables associated with the underlying points of $S_0 \cap L$, and the number of such points satisfies $|S_0 \cap L| \leqslant (1+\delta^3)m_0$ by (7) and (8). Furthermore, by Proposition 2.4, there are at most $D := 18m_0^4$ lines with weight larger than m_0^{-2} passing through each point of the grid $[n]^2$. Therefore, since $m_0 \leqslant m^3$, the maximum degree Δ of the dependency graph satisfies

$$\Delta \leqslant (1+\delta^3)m_0 \cdot D \leqslant 20m_0^5 \leqslant 20m^{15}, \quad \text{and thus} \quad q\Delta \leqslant (1/(5m))^{15} \cdot 20m^{15} < 1/4.$$

Let $\mathcal{E}$ denote the event that none of the bad events occur. Then $\mathbb{P}[\mathcal{E}] > 0$ by Proposition 2.2(1), and we can consider the conditional distribution of S given $\mathcal{E}$.

Quasirandomness. If $\mathcal{E}$ occurs, then S clearly satisfies conditions (1), (2), and (3). Next, we check that, in the conditional probability space given $\mathcal{E}$, our random set S also satisfies condition (4) with high probability (denote the event that condition (4) fails by $\mathcal{E}_4$).

Since each of the n horizontal lines contains at most $(1+\delta^3)m_0$ points of S_0, the total number of events N in our application of the local lemma satisfies

$$N \leqslant n \cdot (1+\delta^3)m_0 \cdot D \leqslant 20m^{15}n, \quad \text{and thus} \quad qN \leqslant (1/(5m))^{15} \cdot 20m^{15}n < n.$$

In case (a), condition (4) for S_0 gives that

$$|S_0 \cap (I \times J)| \in (1 \pm \delta^3)|I||J| \cdot m_0/n$$

for every pair of sets $I, J \subseteq [n]$ of size at least $n/10$. On the other hand, in case (b) we trivially have $|S_0 \cap (I \times J)| = |I||J|$. Either way, by the Chernoff bound (Proposition 2.1),

$$\begin{aligned}\mathbb{P}\big[|S \cap (I \times J)| \notin (1 \pm \delta)|I||J| \cdot m/n\big] &\leqslant \mathbb{P}\big[|S \cap (I \times J)| \notin (1 \pm \delta/2)p \cdot |S_0 \cap (I \times J)|\big] \\ &\leqslant 2\exp\left(-\frac{(\delta/2)^2 p |S_0 \cap (I \times J)|}{3}\right) \leqslant 2\exp\left(-\frac{(\delta/2)^2(1-\delta^3)mn}{3 \cdot 10^2}\right) \leqslant 2\exp\left(-\frac{m^{5/6}n}{1300}\right).\end{aligned}$$

Therefore, by the union bound and Proposition 2.2(2),

$$\begin{aligned}\mathbb{P}[\mathcal{E}_4 \mid \mathcal{E}] &\leqslant \mathbb{P}[\mathcal{E}_4] \cdot \exp(6qN) \leqslant 2^{2n} \cdot 2\exp(-m^{5/6}n/1300) \cdot \exp(6qN) \\ &\leqslant \exp(n \cdot (2 + 1 - (m^{5/6}/1300) + 6)) \leqslant \exp(-n).\end{aligned} \tag{9}$$

(Here we again used that $m \geqslant K = 10^{36}$.)

Spreadness. We claim that the conditional distribution of S given $\mathcal{E}$ and $\overline{\mathcal{E}_4}$ has the desired properties. It is supported on m-nice sets by construction, hence it only remains to check that it is $(m+1)/n$-spread.

Consider a non-empty set $T \subseteq [n]^2$. In both cases (a) and (b), our original distribution is $((1+m^{-3})m_0/n)$-spread, and thus

$$\mathbb{P}[T \subseteq S] = p^{|T|} \cdot \mathbb{P}[T \subseteq S_0] \leqslant \big(p \cdot (1+m^{-3})m_0/n\big)^{|T|}.$$

The event "$T \subseteq S$" depends on at most $D \cdot |T|$ bad events. Therefore, by Proposition 2.2(2),

$$\begin{aligned}\mathbb{P}[T \subseteq S \mid \mathcal{E}] &\leqslant \mathbb{P}[T \subseteq S] \cdot \exp(6qD|T|) \leqslant \big(p \cdot (1+m^{-3})m_0/n \cdot \exp(6qD)\big)^{|T|} \\ &\leqslant \big((1+m^{-3}) \cdot \exp(m^{-3}/2) \cdot m/n\big)^{|T|} \leqslant \big((1+m^{-3})^2 \cdot m/n\big)^{|T|}.\end{aligned}$$

Since $n \geqslant m$, combining this with (9) gives that

$$\begin{aligned}\mathbb{P}\big[T \subseteq S \mid \overline{\mathcal{E}_4} \cap \mathcal{E}\big] &\leqslant \frac{\mathbb{P}[T \subseteq S \mid \mathcal{E}]}{1 - \mathbb{P}[\mathcal{E}_4 \mid \mathcal{E}]} \leqslant \frac{((1+m^{-3})^2 \cdot m/n)^{|T|}}{1 - \exp(-n)} \\ &\leqslant \big((1 + 2\exp(-m)) \cdot (1+m^{-3})^2 \cdot m/n\big)^{|T|} \leqslant \big((m+1)/n\big)^{|T|}.\end{aligned}$$

This means precisely that the conditional distribution of S given $\mathcal{E}$ and $\overline{\mathcal{E}_4}$ is $(m+1)/n$-spread, completing the proof. $\square$

In Lemma 4.3 below, we carry out the final regularisation step. The case analysis in its proof is similar to that in [18, Proposition 11].

Lemma 4.3. *Let n, k be integers such that $K \leqslant k \leqslant n$. Consider a $((1+\varepsilon)k+1)/n$-spread distribution supported on $(1+\varepsilon)k$-nice sets, and let $S \subseteq [n]^2$ be a random sample from this distribution. Then, with positive probability, there exists a subset $S' \subseteq S$ that contains exactly k points on each horizontal and each vertical line.*

Proof. Our random subset $S \subseteq [n]^2$ naturally corresponds to a random bipartite graph G with two parts of size n. Call these two parts A_0 and B_0. We need to prove that G has a k-regular spanning subgraph. By Proposition 2.3, it suffices to check that, with positive probability, for every choice of subsets $A \subseteq A_0, B \subseteq B_0$, we have

$$e_G(A, B_0 \setminus B) \geqslant k(|A| - |B|). \tag{10}$$

We cover all possible choices of $A \subseteq A_0$ and $B \subseteq B_0$ by seven "types" (a pair of sets (A, B) may have more than one type):

- (A, B) is of *type 0* if $|B| \geqslant |A|$;
- (A, B) is of *type 1* if $|A| > 2|B|$;
- (A, B) is of *type 2* if $|A| \geqslant n/10$ and $|B| \leqslant n/2$;
- (A, B) is of *type 3* if $|B| < |A| \leqslant 2|B|$ and $|B| \leqslant n/10$;
- (A, B) is of *type 1** if $|B_0 \setminus B| > 2|A_0 \setminus A|$;
- (A, B) is of *type 2** if $|B_0 \setminus B| \geqslant n/10$ and $|A_0 \setminus A| \leqslant n/2$;
- (A, B) is of *type 3** if $|A_0 \setminus A| < |B_0 \setminus B| \leqslant 2|A_0 \setminus A|$ and $|A_0 \setminus A| \leqslant n/10$.

First, we check that every pair of sets has at least one type. Indeed, if a pair of sets (A, B) is not of type 0 then $|B| < |A|$, and hence $|B| \leqslant n/2$ or $|A_0 \setminus A| \leqslant n/2$. Consider the case when $|B| \leqslant n/2$: if $|A| \geqslant n/10$ then this pair is of type 2; otherwise, $|B| < |A| < n/10$, and thus this pair is of type 1 or of type 3. Similarly, in the case when $|A_0 \setminus A| \leqslant n/2$, this pair is of type 1*, 2*, or 3*.

Note that pairs of sets (A, B) of type 0 trivially satisfy (10). We will show that

- with probability 1, all pairs of types 1, 1*, 2, and 2* satisfy (10);
- with probability greater than 1/2, all pairs of type 3 satisfy (10);
- with probability greater than 1/2, all pairs of type 3* satisfy (10).

This will immediately imply that, with positive probability, (10) holds for all choices of $A \subseteq A_0$ and $B \subseteq B_0$.

Note that (10) can be equivalently written as $e_G(B_0 \setminus B, A_0 \setminus (A_0 \setminus A)) \geqslant k(|B_0 \setminus B| - |A_0 \setminus A|)$. Therefore, by symmetry, it suffices to consider types 1, 2, and 3.

Since $k \geqslant K = 10^{36}$ and $\varepsilon = 0.005$, we have $\varepsilon/4 > k^{-1/12}$. Thus, by condition (1) from Definition 4.1, every vertex in G has degree in $(1 \pm \varepsilon/4)(1+\varepsilon)k$.

Type 1 pairs: For a pair (A, B) of type 1 (with $|A| > 2|B|$), we have

$$\begin{aligned} e_G(A, B_0 \setminus B) &\geqslant \sum_{v \in A} \deg_G(v) - \sum_{v \in B} \deg_G(v) \geqslant |A| \cdot (1-\varepsilon/4)(1+\varepsilon)k - |B| \cdot (1+\varepsilon/4)(1+\varepsilon)k \\ &\geqslant k(|A| - |B|) + k((3\varepsilon/4 - \varepsilon^2/4)|A| - (5\varepsilon/4 + \varepsilon^2/4)|B|) \geqslant k(|A| - |B|). \end{aligned}$$

Type 2 pairs: For a pair (A,B) of type 2 (with $|A| \geqslant n/10$ and $|B_0 \setminus B| = n - |B| \geqslant n/2$), condition (4) from Definition 4.1 implies that

$$e_G(A, B_0 \setminus B) \geqslant (1-\varepsilon/4)\cdot(1+\varepsilon)k/n \cdot |A|(n-|B|) \geqslant k|A|\frac{n-|B|}{n} \geqslant k(|A|-|B|).$$

Type 3 pairs: Consider a pair (A,B) of type 3 (with $|B| < |A| \leqslant 2|B|$ and $|B| \leqslant n/10$). Note that if $e_G(A, B_0 \setminus B) < k(|A|-|B|)$, then

$$e_G(A,B) > \sum_{v\in A} \deg_G(v) - k(|A|-|B|) \geqslant ((1-\varepsilon/4)(1+\varepsilon)-1)k|A| + k|B| \geqslant k|B|.$$

For a set of possible edges $E \subseteq A \times B$ of size $k|B|$, let $\mathcal{E}(A,B,E)$ be the event that $E \subseteq E_G(A,B)$. Since our distribution is $((1+\varepsilon)k+1)/n$-spread, we have

$$\sum_{E\subseteq A\times B,\ |E|=k|B|} \mathbb{P}[\mathcal{E}(A,B,E)] \leqslant \binom{|A||B|}{k|B|}\cdot\left(\frac{(1+\varepsilon)k+1}{n}\right)^{k|B|} \leqslant \left(\frac{e|A|}{k}\cdot\frac{(1+\varepsilon)k+1}{n}\right)^{k|B|} \leqslant \left(\frac{6|B|}{n}\right)^{k|B|}.$$

For every $b \leqslant n/10$, we sum these probabilities over all pairs of sets (A,B) with $|B| = b$ and $b < |A| \leqslant 2b$, to obtain that

$$\sum_{\substack{|B|=b\\ b<|A|\leqslant 2b}} \sum_{\substack{E\subseteq A\times B,\\ |E|=kb}} \mathbb{P}[\mathcal{E}(A,B,E)] \leqslant \binom{n}{b}\cdot b\binom{n}{2b}\cdot\left(\frac{6b}{n}\right)^{kb} \leqslant \left(\frac{en}{b}\right)^{3b}\cdot\left(\frac{6b}{n}\right)^{kb}$$

$$\leqslant \left((6e)^3\cdot(6/10)^{k-3}\right)^b \leqslant (1/4)^b.$$

Therefore, the probability that there is a pair of type 3 violating (10) is at most

$$\sum_{b=1}^{\infty}(1/4)^b < 1/2. \qquad \square$$

Proof of Theorem 1.2. Recall that $K = 10^{36}$, $\varepsilon = 0.005$, and $K \leqslant k \leqslant 0.9n$. Let $(1+\varepsilon)k = m_{(1)} < m_{(2)} < \ldots < m_{(r)}$ be the sequence of real numbers from (6).

Then, by Lemma 4.2(b), there exists an $(m_{(r)}+1)/n$-spread distribution supported on $m_{(r)}$-nice sets. Iteratively applying Lemma 4.2(a), we obtain an $(m_{(1)}+1)/n$-spread distribution supported on $m_{(1)}$-nice sets. Next, Lemma 4.3 implies that there exists an $m_{(1)}$-nice set S containing a subset $S' \subseteq S$ with exactly k points on each horizontal and each vertical line.

We claim that S' has the desired properties. For this, we need to check that for every line $L \in \mathcal{L}$, we have $|S' \cap L| \leqslant k\cdot(w(L)+0.01)$. Let $\delta := m_{(1)}^{-1/12} \leqslant 0.001$. Since S is $m_{(1)}$-nice, condition (1) implies that for every line L with $w(L) \in (m_{(1)}^{-2/3}, 1]$,

$$|S\cap L| \leqslant (1+\delta)m_{(1)}\cdot w(L) = (1+\delta)(1+\varepsilon)k\cdot w(L) \leqslant k\cdot(w(L)+0.01).$$

On the other hand, conditions (2) and (3) imply that for every line L with $w(L) \leqslant m_{(1)}^{-2/3}$,

$$|S\cap L| \leqslant \max\left((1+\delta)m_{(1)}^{1/3}, 14\right) \leqslant 0.01\cdot k.$$

This completes the proof, because S' is a subset of S. $\square$

5. Proof of Theorem 1.1

In this section, we deduce Theorem 1.1 from Theorem 1.2 using the results of Kovács, Nagy, and Szabó [24]. The key observation behind this deduction is that the only lines with weight close to 1 are horizontal lines, vertical lines, and the lines of slope ±1 near the main diagonals of the $n\times n$ grid.

Proof of Theorem 1.1. An explicit construction [24, Proposition 3.1] takes care of the case $k \geqslant \frac{2}{3}n$. So, we may assume that $k < \frac{2}{3}n$, and hence $\frac{3}{10}k < 0.9 \cdot \frac{1}{4}n$.

First, we additionally assume that $4 \mid n$ and $10 \mid k$. As in [24, Definition 3.2], we partition the $n \times n$ grid into sixteen blocks of size $n/4 \times n/4$. Formally, for each $1 \leqslant i, j \leqslant 4$, we define

$$\mathcal{G}_{i,j} := \Big\{(x, y) \in [n]^2 : (i-1)\frac{n}{4} < x \leqslant i\frac{n}{4},\ (j-1)\frac{n}{4} < y \leqslant j\frac{n}{4}\Big\}.$$

We also set

$$k_{i,j} := \begin{cases} \frac{2}{10}k, & \text{if } i = j \text{ or } i + j = 5 \text{ (i.e., if } \mathcal{G}_{i,j} \text{ intersects one of the main diagonals);} \\ \frac{3}{10}k, & \text{otherwise,} \end{cases}$$

and let $w_{i,j}(L) := |\mathcal{G}_{i,j} \cap L|/(n/4)$ for each line $L \in \mathcal{L}$.

For each i, j, we apply Theorem 1.2 (with parameters $n/4$ and $k_{i,j}$) to obtain a subset $S_{i,j} \subseteq \mathcal{G}_{i,j}$ with exactly $k_{i,j}$ points on each horizontal and each vertical line, and at most $k_{i,j} \cdot (w_{i,j}(L) + 0.01)$ points on every other line L. Let $S := \bigsqcup_{1 \leqslant i,j \leqslant 4} S_{i,j}$ be the union of all these sets.

Since each $S_{i,j}$ contains exactly $k_{i,j}$ points on each horizontal line and each vertical line, S contains exactly k points on each horizontal and each vertical line. Crucially, by [24, Lemma 3.5], for every line L which is not horizontal or vertical, we have

$$\sum_{1 \leqslant i,j \leqslant 4} k_{i,j} w_{i,j}(L) \leqslant \frac{4}{5}k,$$

so

$$|S \cap L| = \sum_{1 \leqslant i,j \leqslant 4} |S_{i,j} \cap L| \leqslant \sum_{1 \leqslant i,j \leqslant 4} k_{i,j}(w_{i,j}(L) + 0.01) \leqslant \frac{4}{5}k + 0.01 \sum_{1 \leqslant i,j \leqslant 4} k_{i,j} = 0.84 \cdot k.$$

Since $k \geqslant 10^{37}$, this implies, in particular, that $|S \cap L| \leqslant k - 15$.

This completes the proof of Theorem 1.1 in the case $4 \mid n$ and $10 \mid k$. The general case now follows from [24, Lemma 3.12 and Lemma 3.13], making use of the "reserve" of 15 points per line. $\square$

6. Concluding remarks

Smaller values of k? We did not try to optimise the constant $K_0 = 10^{37}$ in Theorem 1.1, and our argument provides a lot of room for optimisation. However, making any progress on the classical no-three-in-line problem (i.e., the case $k = 2$) would likely require new ideas. Note that, even for "light" lines, our argument can only guarantee that they contain at most $C = 14$ points of our set (see condition (3) in Definition 4.1).

Partitioning the grid into no-$(k+1)$-in-line sets. In [18], Jain and Pham constructed an optimally spread probability distribution over decompositions of the edge-set of $K_{n,n}$ into perfect matchings (i.e., over Latin squares of order n). It is plausible that, modifying our approach in a suitable way, one also might be able to prove that for a sufficiently large k and every n divisible by k, the grid $[n]^2$ can be partitioned into n/k sets of size kn, each containing no $k + 1$ points on a line.

Exact bound in higher dimensions? A natural further direction is to obtain an exact bound in Theorem 1.3, analogous to Theorem 1.1. In order to achieve this using our approach, one would need to come up with a suitable version of the regularisation step (Lemma 4.3). Note that a set $S' \subseteq [n]^d$ that contains exactly k points in each axis-aligned affine subspace of dimension t is a (multipartite) $(n, d, d-t, k)$-design, and that one would need to find such set S' inside a "random-like" configuration $S \subseteq [n]^d$ that contains $(1 + o(1))k$ points in each axis-aligned affine subspace of dimension t. This appears to be related to the challenging problem of finding designs inside Erdős–Rényi random hypergraphs of appropriate density (see [6, 18, 19, 20, 27] for partial results in this direction).

References

[1] Noga Alon and Joel H. Spencer. *The Probabilistic Method*. Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, 4th edition, 2016.

[2] David Brent Anderson. Update on the no-three-in-line problem. *Journal of Combinatorial Theory, Series A*, 27(3):365–366, 1979.

[3] Natalie Behague, Daniel Il'kovič, and Richard Montgomery. A proof of the Kim-Vu sandwich conjecture. Private communication.

[4] Peter Brass and Christian Knauer. On counting point-hyperplane incidences. *Computational Geometry*, 25(1):13–20, 2003. European Workshop on Computational Geometry - CG01.

[5] Peter Brass, William O. J. Moser, and János Pach. Lattice point problems. In *Research Problems in Discrete Geometry*, pages 417–433. Springer, New York, 2005.

[6] Michelle Delcourt, Tom Kelly, and Luke Postle. Thresholds for $(n, q, 2)$-steiner systems via refined absorption, 2024. Preprint, arXiv:2402.17858.

[7] H. E. Dudeney. *Amusements in Mathematics*. Nelson, London, 1917.

[8] David Eppstein. *Forbidden Configurations in Discrete Geometry*. Cambridge University Press, 2018.

[9] David Eppstein. Random no-three-in-line sets, 2018. Blog post.

[10] Achim Flammenkamp. Progress in the no-three-in-line problem. *J. Combin. Theory Ser. A*, 60(2):305–311, 1992.

[11] Achim Flammenkamp. Progress in the no-three-in-line problem. II. *J. Combin. Theory Ser. A*, 81(1):108–113, 1998.

[12] Pu Gao, Mikhail Isaev, and Brendan McKay. Kim–Vu's sandwich conjecture is true for $d \gg \log^4 n$, 2020. Preprint, arXiv:2011.09449.

[13] Anubhab Ghosal, Ritesh Goenka, and Peter Keevash. On subsets of lattice cubes avoiding affine and spherical degeneracies. 2025. Preprint, arXiv:2509.06935.

[14] Ben Green. 100 open problems. Manuscript.

[15] Richard K. Guy and Patrick A. Kelly. The no-three-in-line problem. *Canadian Mathematical Bulletin*, 11(4):527–531, 1968.

[16] Bernhard Haeupler, Barna Saha, and Aravind Srinivasan. New constructive aspects of the Lovász local lemma. *J. ACM*, 58(6):Art. 28, 28, 2011.

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

[18] Vishesh Jain and Huy Tuan Pham. Optimal thresholds for Latin squares, Steiner triple systems, and edge colorings. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1425–1436. SIAM, Philadelphia, PA, 2024.

[19] Dong Yeap Kang, Tom Kelly, Daniela Kühn, Abhishek Methuku, and Deryk Osthus. Thresholds for Latin squares and Steiner triple systems: bounds within a logarithmic factor. *Trans. Amer. Math. Soc.*, 376(9):6623–6662, 2023.

[20] Peter Keevash. The optimal edge-colouring threshold, 2022. Preprint, arXiv:2212.04397.

[21] J. H. Kim and V. H. Vu. Sandwiching random graphs: universality between random graph models. *Adv. Math.*, 188(2):444–469, 2004.

[22] Tereza Klimošová, Christian Reiher, Andrzej Ruciński, and Matas Šileikis. Sandwiching biregular random graphs. *Combin. Probab. Comput.*, 32(1):1–44, 2023.

[23] Benedek Kovács, Zoltán Lóránt Nagy, and Dávid R. Szabó. Randomised algebraic constructions for the no-$(k + 1)$-in-line problem. 2025. Preprint, arXiv:2508.07632.

[24] Benedek Kovács, Zoltán Lóránt Nagy, and Dávid R. Szabó. Settling the no-$(k+1)$-in-line problem when k is not small. 2025. Preprint, arXiv:2502.00176.

[25] Hanno Lefmann. Extensions of the no-three-in-line problem. Preprint, 2012.

[26] L. Lovász. *Combinatorial Problems and Exercises*. AMS Chelsea Publishing, Providence, RI, 2nd edition, 2007.

[27] Ashwin Sah, Mehtaab Sawhney, and Michael Simkin. Threshold for Steiner triple systems. *Geom. Funct. Anal.*, 33(4):1141–1172, 2023.

[28] Wolfgang M. Schmidt. Asymptotic formulae for point lattices of bounded determinant and subspaces of bounded height. *Duke Math. J.*, 35:327–339, 1968.

[29] Benny Sudakov and István Tomon. Evasive sets, covering by subspaces, and point-hyperplane incidences. *Discrete Comput. Geom.*, 72(3):1333–1347, 2024.

Institute of Science and Technology Austria (ISTA)

Email address: aleksandr.grebennikov@ist.ac.at

Institute of Science and Technology Austria (ISTA)

Email address: matthew.kwan@ist.ac.at