A Density-One Sequence with Distinct Consecutive-Block Products

Przemyslaw Chojecki

January 17, 2026

Abstract

We construct a density-1 increasing sequence $1 \le d_1 < d_2 < \cdots$ such that all consecutive-block products $\prod_{u \le i \le v} d_i$ are distinct. We implement a dyadic-stage deletion scheme in which collisions are encoded by separated-variable curve families of the form $\prod(x - a_i) = \prod(y + b_j)$. We introduce a capacity constraint enforced via the Moser–Tardos algorithmic Lovász local lemma, which yields a small deletion set at each dyadic scale hitting all "logarithmic-scale" collisions. We further show how to verify the requisite subpolynomial dependency degree in this logarithmic regime using uniform determinant-method bounds of Castryck–Cluckers–Dittmann–Nguyen and a degeneracy analysis via Bilu–Tichy and Hajdu–Tijdeman (Pell-type families).

1 Introduction

Let $d_1 < d_2 < \cdots$ be an increasing sequence of positive integers. We say that (d_i) has the *distinct consecutive-block product property* if

$$\prod_{u \le i \le v} d_i \neq \prod_{u' \le i \le v'} d_i \tag{1}$$

for all distinct pairs of intervals $[u, v] \neq [u', v']$. Equivalently, there are no nontrivial solutions to

$$d_{M-\ell+1} \cdots d_M \;=\; d_N \cdots d_{N+k} \tag{2}$$

with $M - \ell + 1 \le M < N \le N + k$ and $(\ell, k) \neq (0, 0)$.

The goal is to construct such a sequence with asymptotic density 1:

$$\lim_{n \to \infty} \frac{|\{d_i \le n\}|}{n} = 1.$$

This would settle positively an open problem by Erdős

Problem 1.1 (Problem 421, [1]). *Is there a sequence* $1 \le d_1 < d_2 < \cdots$ *with density* 1 *such that all products* $\prod_{u \le i \le v} d_i$ *are distinct?*

1.1 Main result proved in this paper

We implement a dyadic-stage deletion scheme which removes a set of integers of density 0. The arithmetic geometry input (determinant method bounds, degeneracy classification) is handled in the logarithmic regime.

Theorem 1.1 (Density-1 construction). *There exists a density-*1 *increasing sequence* (d_i) *satisfying* (1).

2 Initial sprinkling eliminates very long relations

In this section we remove, at density-0 cost, all collisions whose block lengths are *very long* relative to the maximal element. Concretely we set

$$L_{\text{long}}(n) \;:=\; \Big\lceil \exp\big((\log n)^{2/3}\big)\Big\rceil, \tag{3}$$

and we delete a sparse random set so that, in the resulting set, there are no solutions to (2) with $\min(\ell, k+1) \ge L_{\text{long}}(d_M)$ once d_M is large.

2.1 Random thinning

For $n \ge 3$ define

$$\delta(n) \;:=\; \frac{1}{(\log n)^2}.$$

Let $S \subset \mathbb{N}$ be a random set obtained by deleting each integer $n \ge 3$ independently with probability $\delta(n)$ (and always keeping $1, 2$). Set

$$A_0 := \mathbb{N} \setminus S.$$

Lemma 2.1 (Density-1 after sprinkling). *With probability* 1 *one has* $|S \cap [1, N]| = o(N)$*, hence* A_0 *has asymptotic density* 1.

Proof. Let $X_n := \mathbf{1}_{n \in S}$ for $n \ge 3$. Then $\mathbb{E}X_n = \delta(n)$ and $\mathrm{Var}(X_n) \le \delta(n)$. For dyadic $N = 2^m$,

$$\mathbb{E} \sum_{n \le N} X_n \;=\; \sum_{n \le N} \delta(n) \;\asymp\; \frac{N}{(\log N)^2}.$$

A standard second-moment/Chebyshev argument on dyadic scales (or Kolmogorov's strong law since $\sum_n \mathrm{Var}(X_n)/n^2 < \infty$) yields

$$\sum_{n \le N} X_n \;=\; \frac{N}{(\log N)^2}\,(1+o(1)) \qquad \text{a.s.}$$

In particular $|S \cap [1,N]| = o(N)$ almost surely. □

2.2 Counting candidate collisions

Fix an integer $m \ge 3$. Consider *any* increasing sequence

$$e_1 < e_2 < \cdots < e_r \le m$$

of distinct integers in $[1,m]$ (we will later take e_i to be the increasing enumeration of $A_0 \cap [1,m]$). The number of consecutive blocks $e_u \cdots e_v$ is at most $r^2 \le m^2$, so the number of ordered pairs of consecutive blocks is at most m^4.

Thus, for any fixed set A_0, the number of candidate equalities

$$\prod_{u \le i \le v} e_i \;=\; \prod_{u' \le i \le v'} e_i \tag{4}$$

among consecutive blocks inside $A_0 \cap [1,m]$ is at most m^4.

2.3 Borel–Cantelli: no very long collisions survive

Theorem 2.2 (Very long relations are eliminated)**.** *With probability* 1*, for all sufficiently large integers m, there is* no *solution to* (2) *inside the increasing enumeration* (d_i) *of* A_0 *with maximal term* $d_M = m$ *and*

$$\min(\ell, k+1) \;\ge\; L_{\mathrm{long}}(m).$$

Equivalently: in the set A_0 *produced by sprinkling, all consecutive-block product collisions have* $\min(\ell, k+1) < L_{\mathrm{long}}(d_M)$ *once* d_M *is large.*

Proof. Fix $m \ge 3$. Consider any *specific* candidate equality (4) in $[1,m]$ in which both blocks have length at least $L_{\mathrm{long}}(m)$ (so the equality involves at least $2L_{\mathrm{long}}(m)$ distinct integers, all $\le m$). The probability that *all* those integers survive the sprinkling is at most

$$\prod_{n \in \text{involved integers}} (1-\delta(n)) \;\le\; (1-\delta(m))^{2L_{\mathrm{long}}(m)} \;\le\; \exp\Big(-\frac{2L_{\mathrm{long}}(m)}{(\log m)^2}\Big).$$

As discussed above, the number of candidate equalities (4) among consecutive blocks in $A_0 \cap [1, m]$ is at most m^4 (uniformly for all realizations of A_0). Hence, by a union bound, the probability that there exists *any* surviving collision with maximal term m and both blocks of length $\geq L_{\text{long}}(m)$ is at most

$$m^4 \exp\Big(- \frac{2L_{\text{long}}(m)}{(\log m)^2} \Big) \;\leq\; m^4 \exp\Big(- \exp\big((\log m)^{2/3}/2 \big) \Big),$$

for all large m. The series $\sum_{m \geq 3}$ of these upper bounds converges. Therefore, by the Borel–Cantelli lemma, with probability 1 only finitely many m admit such a surviving collision. This proves the theorem. □

Remark 2.3 (How this plugs into the main construction). From now on, we *fix* one realization of A_0 for which Theorem 2.2 and Lemma 2.1 hold.

We then run the dyadic Moser–Tardos construction (Sections 6.5 and 6.6 together with the short-edge cleanup) inside A_0 to eliminate the remaining collisions with $\min(\ell, k+1) < L_{\text{long}}(d_M)$.

Since A_0 already has density 1 and the additional deletions have density 0, the final set retains density 1 and has *no* collisions at all.

Corollary 2.4 (Sprinkling kills polylog-long collisions)**.** *With probability* 1*, there exists* m_0 *such that for all* $m \geq m_0$ *there is no collision* (2) *inside the increasing enumeration of* A_0 *whose maximal element equals* m *and whose lengths satisfy*

$$\max(\ell, k+1) \;\geq\; 10(\log m)^3.$$

Proof. Fix m.

Any equality of consecutive-block products may be reduced (by cancelling the overlap of the two index-intervals) to an equality of *disjoint* consecutive blocks

$$d_{M-\ell+1} \cdots d_M \;=\; d_N \cdots d_{N+k}, \qquad M < N,$$

so that the two blocks use disjoint sets of terms. In particular, such a collision involves at least $\ell + (k+1) \geq \max(\ell, k+1)$ *distinct* integers $\leq m$. Therefore, under the hypothesis $\max(\ell, k+1) \geq 10(\log m)^3$, the collision involves at least $10(\log m)^3$ distinct integers that must all survive the sprinkling.

Since $\delta(n) = 1/(\log n)^2$ and $\delta(n) \geq \delta(m)$ for $n \leq m$, this survival probability is at most

$$(1 - \delta(m))^{10(\log m)^3} \leq \exp\Big(- \frac{10(\log m)^3}{(\log m)^2} \Big) = m^{-10}.$$

There are at most m^4 candidate ordered pairs of consecutive blocks among integers $\le m$, so by union bound the probability that any such collision survives is at most $m^4 \cdot m^{-10} = m^{-6}$, which is summable. Apply Borel–Cantelli. $\square$

Remark 2.5. After fixing a realization of A_0 for which Corollary 2.4 holds, it suffices at dyadic scale X to eliminate collisions with $\max(\ell, k+1) \le L(X) := 10(\log X)^3$; there is no "intermediate" regime.

3 Dyadic stage architecture

We build a set $A \subset \mathbb{N}$ by deleting, at each dyadic scale $X = 2^m$, a set $D_X \subset [X, 2X)$ and setting

$$A := \mathbb{N} \setminus \bigcup_{m \ge 0} D_{2^m}.$$

Let (d_i) be the increasing enumeration of A.

We work inside the sprinkled density-1 set A_0 from Section 2. At every stage we only consider integers in $A_0 \cap [X, 2X)$ as eligible (all integers in $\mathbb{N} \setminus A_0$ are pre-deleted). Thus all subsequent deletion sets are subsets of A_0.

A key stability property is that if we delete at the stage of the maximal element of a potential collision, then that collision cannot reappear later (since all future elements exceed the maximal element). Thus we target collisions according to their maximal value.

3.1 A logarithmic cut-off

Fix X dyadic and define

$$L = L(X) := \lceil C_0(\log X)^3 \rceil, \qquad T = T(X) := \lceil (\log X)^{2/3} \rceil, \qquad p := \frac{T}{10L}.$$

The choice $L = (\log X)^3$ is large enough to cover all "logarithmic scale" collisions treated by the arithmetic geometry input below; the exponent 3 is not optimized.

3.2 Scale goal after sprinkling (what remains to kill at scale X)

Fix dyadic X. We work inside the sprinkled density-1 set A_0 from Section 2 and write (d_i) for the increasing enumeration of the current set.

Recall that the sprinkling theorem (Theorem 2.2) guarantees that, for all sufficiently large maximal terms d_M, there are *no* collisions (2) in which *both* blocks have length at least

$$L_{\mathrm{long}}(d_M) = \Big\lceil \exp\big((\log d_M)^{2/3}\big)\Big\rceil.$$

Therefore, at scale X it suffices to eliminate collisions with maximal term $d_M \in [X, 2X)$ in which at least one side is *not* very long, i.e.

$$\text{there is no solution to (2) with } d_M \in [X, 2X) \text{ and } \min(\ell, k+1) < L_{\mathrm{long}}(X). \tag{5}$$

We achieve (5) by splitting into two subregimes:

1. **Geometric regime:** $\max(\ell, k+1) \le L(X)$, where

$$L(X) := \lceil C_0 (\log X)^3 \rceil \ll L_{\mathrm{long}}(X),$$

so collisions are encoded by split-product curves $F_{\mathbf{a},\mathbf{b}}(x, y) = 0$ with controlled shifts, and are handled by CCDN/Bombieri–Pila plus the degenerate (Pell-type) analysis.

2. **Intermediate regime:**(empty by Corollary 2.4) $L(X) < \max(\ell, k + 1) < L_{\mathrm{long}}(X)$, where we treat collisions by the same capacity-constrained hitting method, but with window length $L_{\mathrm{long}}(X)$ and capacity $U = (\log X)^2$, or by iterating the geometric regime on a sparse sequence of auxiliary cutoffs tending to $L_{\mathrm{long}}(X)$.

In particular, once the geometric and intermediate regimes are eliminated at each dyadic scale, Theorem 2.2 handles the complementary very-long case and the final sequence has no collisions at all.

4 Encoding collisions by separated-variable curves

Let $x = d_M$ and $y = d_{N+k}$ be the endpoints of a collision (2). Define the *shift vectors* by

$$a_i := x - d_{M-i} \quad (1 \le i \le \ell - 1), \qquad b_j := d_{N+j} - d_N \quad (1 \le j \le k).$$

Then (2) can be written in the *split product form*

$$x(x - a_1) \cdots (x - a_{\ell-1}) = y(y + b_1) \cdots (y + b_k). \tag{6}$$

For fixed (ℓ, k) and shifts $(\mathbf{a}, \mathbf{b})$, this is the integer point condition on the algebraic curve

$$F_{\mathbf{a},\mathbf{b}}(x,y) = 0, \qquad F_{\mathbf{a},\mathbf{b}}(x,y) := \prod_{i=0}^{\ell-1}(x - a_i) - \prod_{j=0}^{k}(y + b_j),$$

with $a_0 := 0$ and $b_0 := 0$.

4.1 Capacity implies bounded local shifts

To make this useful, we enforce a *capacity constraint* on the deletion set D_X:

$$\forall I \subset [X, 2X) \text{ an interval of integers with } |I| = L, \qquad |D_X \cap I| \le T. \quad (7)$$

Lemma 4.1 (Capacity $\Rightarrow$ bounded local gaps and bounded shifts)**.** *Assume* (7) *at scale* X *and let* $A_X := [X, 2X) \setminus D_X$*. Then* A_X *has no gap of length* $\ge L$ *in the integer line, and for any collision* (2) *with* $d_M \in [X, 2X)$ *and* $\max(\ell, k+1) \le L$ *one has*

$$\max_i a_i \ll L + T, \qquad \max_j b_j \ll L + T,$$

with implied constant absolute.

Proof. If there were a gap of length $\ge L$ inside $[X, 2X)$, some length-L interval would be entirely deleted, contradicting (7) since then $|D_X \cap I| = L > T$.

For the shifts, note that by (7), every integer interval of length L contains at least $L - T$ surviving integers. Hence for any integer interval J of length $L + 2T$, by covering J with two overlapping length-L intervals, we see J contains at least $(L + 2T) - 2T = L$ surviving integers. Therefore, moving left from x in the integer line by $L + 2T$ we encounter at least L elements of A, which implies the previous $\ell \le L$ elements in the sequence lie within distance $\ll L + T$ in value. The same argument applies forward on the y-side. □

Thus, in the logarithmic regime, collisions correspond to integer solutions of $F_{\mathbf{a},\mathbf{b}}(x, y) = 0$ with bounded shifts $\|\mathbf{a}\|_\infty, \|\mathbf{b}\|_\infty \ll L$.

5 Diophantine input: counting integer points on the collision curves

5.1 Nondegenerate curves: determinant method bounds

We use bounds of Bombieri–Pila for fixed degree, and the strengthened polynomial-in-degree bounds of Castryck–Cluckers–Dittmann–Nguyen (CCDN) in the growing-degree range.

Theorem 5.1 (Bombieri–Pila)**.** *Let $C \subset \mathbb{A}^2$ be an irreducible plane curve over $\mathbb{Q}$ of degree $d \geq 2$. Then for any $\varepsilon > 0$,*

$$\#\big(C(\mathbb{Z}) \cap [-B, B]^2\big) \ \ll_{d,\varepsilon} \ B^{1/d+\varepsilon}.$$

Theorem 5.2 (CCDN, polynomial dependence on d)**.** *Let $C \subset \mathbb{A}^2$ be an irreducible plane curve over $\mathbb{Q}$ of degree $d \geq 2$. Then*

$$\#\big(C(\mathbb{Z}) \cap [-B, B]^2\big) \ \ll \ d^{O(1)}\, B^{1/d}\, (\log B + d),$$

with an absolute implied constant and polynomial dependence on d.

Remark 5.3. We use Theorem 5.2 only in the range $d \geq d_0(X)$ so that $B^{1/d} = X^{o(1)}$. For small fixed d we rely on Theorem 5.1.

5.2 Degeneracy classification and Pell-type families

The collision polynomial $F_{\mathbf{a},\mathbf{b}}$ has split linear factors in each variable. Degenerate families (where the curve contains a low-genus component with infinitely many points) are governed by the Bilu–Tichy classification of $f(x) = g(y)$ and its refinements for simple rational roots. In the important $m = 2$ branch, Pell-type parametrizations yield at most $O(\log X)$ points in a dyadic box.

Theorem 5.4 (Bilu–Tichy)**.** *Let $f, g \in \mathbb{Q}[x]$ be nonconstant polynomials. If $f(x) = g(y)$ has infinitely many integer solutions, then (f, g) factors through a* standard pair *in the Bilu–Tichy list.*

Theorem 5.5 (Hajdu–Tijdeman: simple rational roots)**.** *Let $f, g \in \mathbb{Q}[x]$ have only simple rational roots. If $f(x) = g(y)$ has infinitely many solutions with bounded denominator, then (f, g) lies in a strongly restricted set of families related to Prouhet–Tarry–Escott tuples; in particular, the non-line cases in the $m = 2$ branch are Pell-type.*

We summarize the consequence we need.

Proposition 5.6 (Pell-type degeneracy is dyadically sparse)**.** *Let*

$$F(x,y) = f(x) - g(y), \qquad f(x) = \prod_{i=0}^{\ell-1}(x - a_i), \quad g(y) = \prod_{j=0}^{k}(y + b_j),$$

with $a_i, b_j \in \mathbb{Z}$ distinct and $a_0 = b_0 = 0$, $a_i, b_j \geq 0$. Assume that $F(x,y) = 0$ has a degenerate (genus 0 or 1) irreducible component with infinitely many integer points, and consider only solutions with $x, y > 0$ and $x < y$. Then for every dyadic X,

$$\#\{(x,y) \in [X, 2X]^2 \cap \mathbb{Z}^2 : \; F(x,y) = 0, \; x < y\} \; \ll \; \log X.$$

Proof. By Theorem 5.4, if $F(x,y) = 0$ has infinitely many integer points, then f and g factor through a standard pair after composing with a polynomial φ.

Because f and g are split with simple rational roots, Theorem 5.5 applies and forces the base degree $m \in \{1, 2\}$. If $m = 1$, then the standard-pair mechanism is affine (translate/reflection) and produces an affine line component. In the collision setting we only count points with $x < y$ and $x, y > 0$, and Lemma 6.4 shows such line components contribute no admissible points.

If $m = 2$, then Hajdu–Tijdeman show (see [6, §6 and §9], in particular Theorem 6.1 and Theorem 9.1 together with the Pell-driven examples) that the infinite families arise from Pell-type parametrizations: there exist integers $D > 0$ and $E \neq 0$ and rational functions (indeed polynomials after clearing denominators) Φ, Ψ of bounded degree such that all sufficiently large integer solutions on the degenerate component are obtained from solutions $(u, v) \in \mathbb{Z}^2$ to

$$u^2 - Dv^2 = E$$

via $(x, y) = (\Phi(u,v), \Psi(u,v))$ (up to finitely many exceptional points).

The solutions of a Pell equation grow exponentially in the index (e.g. via the fundamental unit), hence the number of solutions with $|u|, |v| \leq X^{O(1)}$ is $O(\log X)$. Since Φ, Ψ have fixed bounded degree, the constraint $(x, y) \in [X, 2X]^2$ forces $|u|, |v| \leq X^{O(1)}$, so there are at most $O(\log X)$ admissible pairs (x, y) in the dyadic box. This proves the proposition. □

Remark 5.7. In the nondegenerate case we will apply Theorem 5.2 to every irreducible component of degree ≥ 2; affine line components can be ignored for collision counting in the region $x < y$ by Lemma 6.4.

6 Capacity-constrained hitting via Moser–Tardos

Fix a dyadic scale X and write

$$V_X := [X, 2X) \cap A_0$$

for the eligible integers at this scale. We expose independent Bernoulli variables

$$(\xi_n)_{n \in V_X}, \qquad \xi_n = 1 \text{ meaning "delete } n \text{ at scale } X\text{"},$$

and we set

$$B_X := \{n \in V_X : \ \xi_n = 1\}, \qquad D_X := B_X.$$

The goal at scale X is to choose an outcome of (ξ_n) for which (i) no collision with $\max(\ell, k+1) \leq L(X)$ survives with maximal endpoint in $[X, 2X)$, and (ii) the deletions satisfy a local capacity constraint on every length-$L(X)$ interval.

6.1 Freezing "consecutive" structure via local patterns

Fix a dyadic scale X and write $V_X := [X, 2X) \cap A_0$ for the eligible integers at this scale. We expose independent Bernoulli variables $(\xi_n)_{n \in V_X}$, where $\xi_n = 1$ means "delete n" (at scale X) and $\xi_n = 0$ means "keep n".

The main issue is that the notion of "consecutive in the surviving sequence" depends on the realization of (ξ_n). Under the capacity constraint (7), however, the t-th predecessor of a point in the surviving set is determined by the pattern of (ξ_n) inside a bounded window.

For an integer $y \in [X/2, 2X]$ define the local window

$$W_y := [y - (L + 2T),\, y) \cap V_X,$$

and write $\sigma_y \in \{0,1\}^{W_y}$ for the restriction $(\xi_n)_{n \in W_y}$. Under (7), every such window contains at least L kept points, so for each $1 \leq t \leq L$ the t-th predecessor of y in the surviving set inside V_X is a *deterministic function* of σ_y.

Lemma 6.1 (Local predecessor functional)**.** *Assume the capacity constraint* (7) *at scale X. Then there exist deterministic maps*

$$\mathrm{Pred}_t : \{0,1\}^{W_y} \to \mathbb{Z} \qquad (1 \leq t \leq L)$$

such that for every realization of (ξ_n) *satisfying* (7) *and every* $y \in [X, 2X)$*, the t-th predecessor of y among the kept points in V_X equals* $\mathrm{Pred}_t(\sigma_y)$*. Consequently the full shift vector* $(b_1, \dots, b_k)$ *for any right block of length* $k + 1 \leq L$ *ending at y is determined by* σ_y*.*

Proof. Under (7), every interval of length $L+2T$ contains at least L kept points (by covering it with $O(1)$ intervals of length L and using $|D_X \cap I| \le T$). Thus the t-th predecessor of y lies inside W_y, and is obtained by scanning leftward through W_y and selecting the t-th position with $\xi_n = 0$. This depends only on σ_y. □

Lemma 6.2 (Global capacity implies enough predecessors)**.** *On the event $\bigcap_I \neg \mathsf{B}_I$, every interval $J \subset [X/2, 2X)$ of length L contains at least $L-T$ integers of $([X/2,2X) \cap A_0) \setminus (D_{<X} \cup B_X)$. In particular, for every $u \in [X, 2X)$ the window W_u contains at least L kept points, so the maps $\mathrm{Pred}_t(\sigma_u)$ in Lemma 6.1 are well-defined.*

Proof. If $|J| = L$ then $\neg \mathsf{B}_J$ gives $|(D_{<X} \cup B_X) \cap J| \le T$; subtract from L. For W_u (length $L+2T$), cover it by $O(1)$ length-L subintervals and use the previous bound. □

We now define collision events by quantifying over *patterns* rather than over the a posteriori enumeration (d_i).

For each $y \in [X, 2X)$ and each $k \le L-1$, each admissible local pattern σ_y determines a right polynomial

$$g_{(\sigma_y,k)}(Y) := \prod_{j=0}^{k} \big(Y + b_j(\sigma_y)\big),$$

and similarly a local pattern σ_x at x determines a left polynomial

$$f_{(\sigma_x,\ell)}(X) := \prod_{i=0}^{\ell-1} \big(X - a_i(\sigma_x)\big).$$

Therefore the separated-variable curve

$$F_{(\sigma_x,\ell),(\sigma_y,k)}(x,y) := f_{(\sigma_x,\ell)}(x) - g_{(\sigma_y,k)}(y)$$

is *fixed once the patterns are fixed.* This allows us to pre-enumerate all potential collision constraints as MT/LLL bad events.

6.2 Short-degree collisions: sparse and capacity-respectable

Set the degree threshold

$$d_0 = d_0(X) := \left\lceil (\log X)^{1/3} \log\log X \right\rceil .$$

Let $\mathcal{E}_X^{<d_0}$ denote the family of collision edges E arising from solutions of (6) with $\max(\ell, k+1) < d_0$ and $d_M \in [X, 2X)$.

Lemma 6.3 (Short-degree witness sparsity)**.** *Assume* (7) *at scale* X*. Then the number of solutions* $(x, y) \in [X, 2X]^2$ *to collision equations* (6) *with* $\max(\ell, k+1) < d_0$ *is* $X^{1/2+o(1)}$*. In particular, the number of distinct endpoints* $x \in [X, 2X)$ *appearing in such solutions is* $o(X)$*.*

Proof. By Lemma 4.1, all relevant shifts are $\ll L$. For each fixed (ℓ, k) with $\max(\ell, k+1) < d_0$ and each fixed shift-pattern $(\mathbf{a}, \mathbf{b})$, the curve $F_{\mathbf{a},\mathbf{b}}(x, y) = 0$ has degree $d = \max(\ell, k+1) < d_0$. If it has a line component, we treat it by Lemma 6.4. Otherwise, Bombieri–Pila (Theorem 5.1) gives

$$\#\{(x, y) \in [X, 2X]^2 : \ F_{\mathbf{a},\mathbf{b}}(x, y) = 0\} \ \ll \ X^{1/d+o(1)} \ \leq \ X^{1/2+o(1)}.$$

Now we sum over patterns: the capacity constraint implies there are at most $X^{o(1)}$ possible local shift-patterns at each endpoint, hence at most $X^{o(1)}$ relevant $(\mathbf{a}, \mathbf{b})$ overall in the dyadic box. Summing yields $X^{1/2+o(1)}$ total solutions. The endpoint bound follows. □

6.3 Line components are irrelevant for genuine collisions

In the collision setting we always have $M < N$ after cancelling any common factors, hence

$$x = d_M < d_N \leq y = d_{N+k},$$

so any genuine collision point satisfies $x < y$.

Lemma 6.4 (Affine line components contribute no collision points)**.** *Let*

$$f(x) = \prod_{i=0}^{\ell-1}(x - a_i), \qquad g(y) = \prod_{j=0}^{k}(y + b_j),$$

with $a_i \geq 0$*,* $b_j \geq 0$*, and* $a_0 = b_0 = 0$*, and set* $F(x, y) = f(x) - g(y)$*. If* F *has an affine line component over* $\mathbb{Q}$*, then every point* $(x, y) \in \mathbb{Z}^2$ *on that line satisfies either* $y \leq x$ *or* $y \leq 0$*. In particular, there are no integer points on any line component with* $x, y > 0$ *and* $x < y$*.*

Proof. If F has a $\mathbb{Q}$-line component, then (since f, g are monic with simple roots) it forces a polynomial identity

$$f(x) \equiv g(x + t) \quad \text{or} \quad f(x) \equiv g(-x + t)$$

for some integer t, corresponding to the lines $y = x + t$ and $y = -x + t$ respectively.

(i) If $f(x) \equiv g(x+t)$, then the roots of f are $\{a_i\}$ and the roots of $g(x+t)$ are $\{-t-b_j\}$. Since $b_0 = 0$, we have $-t \in \{a_i\} \subset [0,\infty)$, hence $t \le 0$ and so $y = x + t \le x$. Thus this line contains no points with $x < y$.

(ii) If $f(x) \equiv g(-x+t)$, then the roots of $g(-x+t)$ are $\{t+b_j\}$. Since 0 is a root of f (as $a_0 = 0$), there exists j with $t + b_j = 0$, so $t \le 0$. On the other hand $b_0 = 0$ implies $t + b_0 = t$ is a root of $g(-x+t)$, hence also a root of f, so $t \in \{a_i\} \subset [0,\infty)$ and $t \ge 0$. Therefore $t = 0$ and the line is $y = -x$, which has no points with $x, y > 0$. □

Remark 6.5. Consequently, when counting collision-producing integer points with $x < y$ in dyadic boxes, one may ignore line components entirely: they contribute zero admissible points.

6.4 Long-degree collisions: Moser–Tardos

Let $\mathcal{E}_X^{\ge d_0}$ be the remaining collision edge family with $\max(\ell, k+1) \in [d_0, L]$ and $d_M \in [X, 2X)$. We now choose a second deletion set B_X (randomly) to hit every edge in $\mathcal{E}_X^{\ge d_0}$ while preserving capacity.

Lemma 6.6 (Uniform $X^{o(1)}$ point bound for $d \ge d_0$)**.** *Let $F_{\mathbf{a},\mathbf{b}}$ be as above with degree $d = \max(\ell, k+1) \ge d_0$. Then*

$$\#\{(x,y) \in [X,2X]^2 \cap \mathbb{Z}^2 : \ F_{\mathbf{a},\mathbf{b}}(x,y) = 0\} \ \le \ X^{o(1)}.$$

We count only integer points with $x < y$; any affine line component contributes none by Lemma 6.4.

If $F_{\mathbf{a},\mathbf{b}}$ is degenerate but has no line component, then the same set has size $O(\log X)$.

Proof. If $F_{\mathbf{a},\mathbf{b}}$ is nondegenerate, apply CCDN (Theorem 5.2) to each irreducible component of degree ≥ 2:

$$\# \ll d^{O(1)} X^{1/d} (\log X + d).$$

Since $d \ge d_0 = (\log X)^{1/3} \log\log X$,

$$X^{1/d} = \exp\Big(\frac{\log X}{d}\Big) \le \exp\Big(\frac{(\log X)^{2/3}}{\log\log X}\Big) = X^{o(1)},$$

and $d^{O(1)}(\log X + d) = \exp(O(\log\log X))$ is absorbed into $X^{o(1)}$. If degenerate without a line component, use Proposition 5.6. □

Remark 6.7 (Uniformity for pattern-parametrized collision curves). In the Moser–Tardos formulation, each admissible pattern pair (σ_x, σ_y) and length pair (ℓ, k) determines a *fixed* split-product polynomial

$$F_{(\sigma_x,\ell),(\sigma_y,k)}(X,Y) = \prod_{i=0}^{\ell-1}(X - a_i) - \prod_{j=0}^{k}(Y + b_j)$$

with integer shifts a_i, b_j bounded by $O(L + T)$ (by the capacity constraint and Lemma 4.1). In particular, its degree satisfies

$$d = \max(\ell, k+1) \ \leq \ L(X) = 10(\log X)^3.$$

Therefore the dyadic integer-point bounds invoked in Lemma 6.6 apply *uniformly* to all such polynomials: CCDN yields a bound of the form $d^{O(1)}X^{1/d}(\log X + d)^{O(1)}$, which is $X^{o(1)}$ for $d \geq d_0(X)$, and the degenerate non-line families are dyadically sparse by Proposition 5.6.

Lemma 6.8 (Uniform dyadic bound for all admissible pattern-curves)**.** *Fix dyadic X and $d_0(X) \leq d \leq L(X)$. Then for every admissible pattern-curve $F_{(\sigma_x,\ell),(\sigma_y,k)}$ of degree d,*

$$\#\{(x,y) \in [X,2X]^2 \cap \mathbb{Z}^2 : \ F_{(\sigma_x,\ell),(\sigma_y,k)}(x,y) = 0\} \ \leq \ X^{o(1)},$$

and if the curve is degenerate but non-linear, the bound improves to $O(\log X)$. We count only integer points with $x < y$; any affine line component contributes none by Lemma 6.4.

Proof. This is exactly Lemma 6.6, with the observation (from Remark 6.7) that every admissible pattern-curve is a split-product polynomial of degree $\leq L(X)$ with bounded shifts, so the CCDN/Pell inputs apply uniformly. □

6.5 Collision cylinder events and dependency degree

Fix dyadic X, and write $V_X := [X, 2X) \cap A_0$ for the eligible variables at this scale. Let $L = L(X)$ and $T = T(X)$, and let $p := T/(10L)$.

Local windows. Fix dyadic X and set

$$L := L(X) = 10(\log X)^3, \qquad T := T(X) = \lceil (\log X)^{2/3} \rceil, \qquad p := \frac{T}{10L}.$$

For $u \in [X/2, 2X)$ define the (left) window

$$W_u := [u - (L + 2T),\ u) \cap \big([X/2, 2X) \cap A_0\big).$$

Only the variables ξ_n with $n \in W_u \cap V_X$ are random at scale X; all ξ_n with $n \in W_u \cap [X/2, X)$ were fixed at earlier scales and are treated as constants.

Two-sided capacity events. Let $D_{<X} := \bigcup_{Y<X} D_Y$ be the deletions fixed at smaller dyadic scales, and recall $D_X = B_X$ at scale X. For each integer interval $I \subset [X/2, 2X)$ of length L define the *overflow* and *underflow* events

$$\mathsf{B}_I^+ := \{|(D_{<X} \cup B_X) \cap I| > T\}, \qquad \mathsf{B}_I^- := \{|(D_{<X} \cup B_X) \cap I| < T/100\}.$$

Let B_I denote the union $\mathsf{B}_I^+ \cup \mathsf{B}_I^-$.

Admissible local patterns. For $u \in [X/2, 2X)$ write σ_u for a pattern on $W_u \cap V_X$, i.e.

$$\sigma_u \in \{0,1\}^{W_u \cap V_X}.$$

Call σ_u (L,T)-*admissible* if for every integer interval $J \subset [u - (L + 2T), u)$ of length L one has

$$\frac{T}{200} \le \sum_{n \in J \cap V_X} \sigma_u(n) \le \frac{T}{2}.$$

Let Σ_u denote the set of (L,T)-admissible patterns at u.

Lemma 6.9 (Pattern count under two-sided capacity)**.** *For every $u \in [X/2, 2X)$ one has*

$$|\Sigma_u| \le \sum_{j \le CT} \binom{|W_u|}{j} = \exp\big(O(T \log(L/T))\big) = X^{o(1)}.$$

Proof. Since $|W_u| = O(L)$, the upper bound $\sum_{n \in J \cap V_X} \sigma_u(n) \le T/2$ on every length-L subwindow implies that σ_u has Hamming weight $\le CT$ on $W_u \cap V_X$ for an absolute C (cover W_u by $O(1)$ length-L intervals). Count such patterns and use $\binom{M}{j} \le (eM/j)^j$ with $M = O(L)$. □

Collision cylinder events. Fix endpoints $(x, y) \in [X/2, 2X]^2 \cap \mathbb{Z}^2$ with $x < y$, and fix lengths $1 \le \ell \le L$, $1 \le k+1 \le L$. For patterns $\sigma_x \in \Sigma_x$ and $\sigma_y \in \Sigma_y$, Lemma 6.1 determines the predecessor lists and hence the shift vectors (a_i) and (b_j), and thus determines a fixed split-product polynomial

$$F_{(\sigma_x,\ell),(\sigma_y,k)}(X, Y) \in \mathbb{Z}[X, Y]$$

as in (6).

Define the collision index set $\mathcal{T}_X$ to consist of all tuples

$$\tau = (x, y, \ell, k, \sigma_x, \sigma_y)$$

with $x < y$, ℓ, k as above, $\sigma_x \in \Sigma_x$, $\sigma_y \in \Sigma_y$, and such that

$$F_{(\sigma_x,\ell),(\sigma_y,k)}(x,y) = 0. \tag{8}$$

For each $\tau \in \mathcal{T}_X$, define the *collision cylinder event* A_τ as the event that the restriction of (ξ_n) to $W_x \cup W_y$ coincides with the specified patterns:

$$\mathsf{A}_\tau \ := \ \{\ (\xi_n)_{n\in W_x} = \sigma_x \ \text{ and } \ (\xi_n)_{n\in W_y} = \sigma_y\ \}.$$

(Under A_τ the predecessor lists at x and y are exactly those encoded by the patterns, so (8) means that the corresponding consecutive blocks collide.)

Dependency graph. Two bad events are adjacent if they depend on a common variable ξ_n.

Proposition 6.10 (Subpolynomial dependency degree)**.** *For all sufficiently large dyadic X, the maximum dependency degree among the events*

$$\{\mathsf{A}_\tau : \tau \in \mathcal{T}_X\} \ \cup \ \{\mathsf{B}_I : |I| = L\}$$

is $X^{o(1)}$.

Proof. Fix a variable ξ_n with $n \in V_X$.

(i) Capacity events. The variable ξ_n appears in B_I only when $n \in I$. There are $O(L)$ length-L intervals $I \subset [X/2, 2X)$ containing n.

(ii) Cylinder events. If ξ_n appears in A_τ, then $n \in W_x \cup W_y$. For fixed n, there are only $O(L)$ choices of an endpoint $u \in [X/2, 2X)$ with $n \in W_u$ (because W_u has length $O(L)$). Hence there are $O(L^2)$ possible endpoint pairs (x,y) with $n \in W_x \cup W_y$. For each such (x,y) there are at most L^2 choices of (ℓ,k), and at most $|\Sigma_x||\Sigma_y| = X^{o(1)}$ choices of (σ_x,σ_y) by Lemma 6.9. Therefore the number of tuples τ with n appearing in A_τ is at most

$$O(L^2)\cdot L^2 \cdot X^{o(1)} \ = \ X^{o(1)}.$$

Thus the number of bad events adjacent to ξ_n is $X^{o(1)}$, and the same holds for the maximum dependency degree. □

6.6 Moser–Tardos (pattern-event formulation)

Lemma 6.11 (Coverage: collisions force a cylinder event)**.** *Fix dyadic X and assume $\bigcap_I \neg\mathsf{B}_I$ (global capacity). Let A be the resulting kept set after deleting $D_{<X} \cup B_X$ from A_0. If there exists a collision* (2) *in the increasing*

enumeration of A whose maximal element lies in $[X, 2X)$ and whose lengths satisfy $\max(\ell, k+1) \le L = L(X)$*, then there exists a tuple*

$$\tau = (x, y, \ell, k, \sigma_x, \sigma_y) \in \mathcal{T}_X$$

such that the corresponding cylinder event A_τ *occurs.*

Proof. Let y be the maximal element of the collision, so $y \in [X, 2X)$, and let x be the endpoint of the other block (so $(x, y) \in [X/2, 2X]^2$ after shrinking X by a factor 2 if needed). Because $\max(\ell, k+1) \le L$ and global capacity holds, Lemma 6.2 implies that the predecessor lists of length ℓ below x and length $k+1$ below y are determined by the restrictions $\sigma_x := (\xi_n)_{n \in W_x}$ and $\sigma_y := (\xi_n)_{n \in W_y}$.

Define τ by these endpoints, lengths, and patterns. By construction, $F_{(\sigma_x, \ell),(\sigma_y, k)}(x, y) = 0$ (it is exactly the collision equality written in split-product form using the predecessor-determined shifts), so $\tau \in \mathcal{T}_X$. Finally, since σ_x, σ_y are the actual restrictions of (ξ_n) on W_x, W_y, the cylinder event A_τ occurs. □

Theorem 6.12 (Moser–Tardos at scale X)**.** *For all sufficiently large dyadic X there exists $B_X \subset V_X = [X, 2X) \cap A_0$ such that:*

1. *No collision cylinder event occurs:* A_τ *fails for every $\tau \in \mathcal{T}_X$. In particular, there is no collision* (6) *with* $\max(\ell, k+1) \le L(X)$ *and endpoints in $[X/2, 2X]$ consistent with the capacity constraint.*

2. *No overflow occurs: for every integer interval $I \subset [X, 2X)$ of length L one has $|B_X \cap I| \le T/2$.*

3. $|B_X| = o(X)$*.*

Remark 6.13. If $|B_X \cap I| \le T$ for every length-L interval $I \subset [X, 2X)$, then

$$|B_X| \;\le\; \Big(\lceil X/L \rceil + 1\Big)T \;=\; o(X),$$

so the global density loss is automatically 0.

Proof. Choose independent Bernoulli variables $(\xi_n)_{n \in V_X}$ with $\mathbb{P}(\xi_n = 1) = p := T/(10L)$, and set $B_X = \{n \in V_X : \xi_n = 1\}$.

(A) Probabilities of bad events. For any length-L interval I, we have

$$\mathbb{E}|B_X \cap I| \;=\; p|I \cap V_X| \;\asymp\; pL \;=\; T/10.$$

Hence Chernoff gives

$$\mathbb{P}(\mathsf{B}_I^+) \le \exp(-cT), \qquad \mathbb{P}(\mathsf{B}_I^-) \le \exp(-cT)$$

for an absolute constant $c > 0$, and therefore $\mathbb{P}(\mathsf{B}_I) \le 2\exp(-cT)$.

For a collision cylinder event A_τ, the pattern constraints force

$$\sum_{n \in J \cap V_X} \xi_n \ \ge\ T/200$$

for each length-L subinterval J of $[x-(L+2T), x)$ and similarly on the y-side. In particular, the total number of ones fixed by A_τ on $(W_x \cup W_y) \cap V_X$ is $\Omega(T)$. Writing $m := |(W_x \cup W_y) \cap V_X| = O(L)$ and $s := \#\{n : \xi_n = 1\}$ fixed by A_τ, we get

$$\mathbb{P}(\mathsf{A}_\tau) \ =\ p^s(1-p)^{m-s} \ \le\ p^s \ \le\ \left(\frac{T}{10L}\right)^{c_0 T} \ =\ \exp\bigl(-\Omega(T\log(L/T))\bigr)$$

for some absolute $c_0 > 0$.

(B) Apply the asymmetric Lovász local lemma. Set LLL-weights

$$x(\mathsf{A}_\tau) := \exp\bigl(-\Omega(T\log(L/T))\bigr), \qquad x(\mathsf{B}_I) := \exp(-\Omega(T)).$$

By Proposition 6.10, each bad event has $X^{o(1)}$ neighbors, while $x(\mathsf{A}_\tau) = \exp(-\Omega(T\log(L/T)))$ and $x(\mathsf{B}_I) = \exp(-\Omega(T))$. Since $T\log(L/T) \to \infty$, we have

$$\sum_{\mathsf{F} \sim \mathsf{E}} x(\mathsf{F}) \ =\ o(1),$$

and the asymmetric LLL inequalities hold for large X. The Moser–Tardos theorem then yields an assignment with no bad events. This gives (1) and (2).

(C) Size bound. Since $\mathbb{E}|B_X| = p|V_X| \ll pX = X \cdot T/(10L) = o(X)$, we may fix an output assignment satisfying (1) and (2) and also $|B_X| = o(X)$ (e.g. by Markov on the output distribution). □

7 Completion of the proof of Theorem 1.1

Fix dyadic X. Choose B_X by Theorem 6.12 and set $D_X := B_X$. By construction, all collisions with $\max(\ell, k+1) \le L(X)$ and maximal element in $[X, 2X)$ are removed, and the two-sided capacity constraint holds on every length-$L(X)$ interval.

Summing dyadically, since $|D_X| = o(X)$ one has

$$\Big| \bigcup_{2^m \le N} D_{2^m} \Big| = o(N),$$

so the resulting set A has density 1.

Finally, Theorem 2.2 (from the initial sprinkling step) eliminates all collisions in which both blocks have length at least $L_{\text{long}}(d_M)$ once d_M is large. The dyadic Moser–Tardos construction eliminates all remaining collisions (with at least one block shorter than $L_{\text{long}}(d_M)$) at density-0 cost. Therefore the final increasing enumeration (d_i) satisfies (1). This proves Theorem 1.1. □

References

[1] T. Bloom, Erdős Problem #421 `https://www.erdosproblems.com/421`

[2] E. Bombieri and J. Pila, The number of integral points on arcs and ovals, *Duke Math. J.* **59** (1989), no. 2, 337–357. (link)

[3] W. Castryck, R. Cluckers, P. Dittmann, and K. H. Nguyen, The dimension growth conjecture, polynomial in the degree and without logarithmic factors, *Algebra Number Theory* **14** (2020), no. 8, 2261–2294. (arXiv:1904.13109)

[4] R. A. Moser and G. Tardos, A constructive proof of the general Lovász local lemma, *J. ACM* **57** (2010), no. 2, Art. 11. (arXiv:0903.0544)

[5] Y. F. Bilu and R. F. Tichy, The Diophantine equation $f(x) = g(y)$, *Acta Arith.* **95** (2000), no. 3, 261–288. (pdf)

[6] L. Hajdu and R. Tijdeman, The Diophantine equation $f(x) = g(y)$ for polynomials with simple rational roots, preprint (arXiv:2204.12345). (arXiv:2204.12345)

[7] A. Hildebrand, Integers without large prime factors, *J. Théor. Nombres Bordeaux* **5** (1993), 411–484. (eudml)