Dense Integer Polynomials with Sparse Powers

Samuel Korsky

November 27, 2026

Abstract

For every sufficiently large d, we construct a degree-d integer polynomial with all $d+1$ coefficients nonzero whose square has at most $4\sqrt{2d} + O(d^{1/4})$ terms. For each fixed $m \geq 2$, a single dense integer polynomial of degree d can have $O_m(\sqrt{d})$ terms in every power from 2 through m. The same bound holds for $P(f)$ when P is any prescribed integer polynomial of degree m and the dense input f may depend on P.

hexagon:2611.01274v1

1 Introduction

How few terms can the square of a dense polynomial have? Write $t(g)$ for the number of nonzero coefficients of g. A degree-d polynomial is *dense* if all $d+1$ coefficients are nonzero. With positive coefficients its square has all $2d+1$ terms. With both signs, most of these terms can cancel.

Rédei asked how small $t(g^2)$ can be for a given $t(g)$. In 1947 Rényi, reporting on joint work with Kalmár and Rédei, gave a dense polynomial with 29 terms whose square has 28 [4]. A single example gives an infinite family: If g is dense of degree $k-1$ and $t(g^2) = s$, then $\prod_{i=0}^{r-1} g(x^{k^i})$ has all k^r terms, by uniqueness of base-k digits. Its square has at most s^r terms.

Erdős obtained $O(n^{1-\delta})$ terms in the square for every input size n and some $\delta > 0$, allowing rational coefficients and gaps in the exponents [3]. For dense squares, Verdenius obtained the exponents $\log 6/\log 9 \approx 0.8155$ over the integers and approximately 0.8107 over the reals [9, 1]. Unsolved Lab reports an integer exponent of approximately 0.812807 [8]. Our first construction improves the best known exponent to 0.5 and works in every sufficiently large degree.

Theorem 1.1 (Squares)**.** *For every sufficiently large d there is a dense $f \in \mathbb{Z}[x]$ of degree d with $t(f^2) \leq 4\sqrt{2d} + O(d^{1/4})$.*

FrontierMath asks for a dense integer polynomial with $t(f^2) < \sqrt{\deg f}$ [2]. Such an example would give an exponent strictly below $1/2$ along a sequence of degrees, since the base-k construction above has exponent $\log s/\log k$.

Theorem 1.2 (Several powers and prescribed compositions)**.** *For each integer $m \geq 2$ there are constants C_m and $d_0(m)$ such that the following statements hold for every $d \geq d_0(m)$.*

1. *There is one dense $F \in \mathbb{Z}[x]$ of degree d such that $t(F^j) \leq C_m\sqrt{d}$ for every $2 \leq j \leq m$.*
2. *For each $P \in \mathbb{Z}[t]$ of degree m, there is a dense $f_P \in \mathbb{Z}[x]$ of degree d such that $t(P(f_P)) \leq C_m\sqrt{d}$.*

Verdenius obtained the exponent $0.99934\ldots$ for dense real cubes [9, 1]. Coppersmith and Davenport constructed, in each degree d, a single dense real F with $t(F^j) = O_m(d^{1-\delta_m})$ for all $2 \leq j \leq m$ and some $\delta_m > 0$. They also treated compositions over $\mathbb{C}$ [1]. If $P'(a) = 0$, Taylor's formula expresses $P(a+F)$ as a constant plus a linear combination of $F^2, \ldots, F^m$. Over the integers a suitable

constant a need not exist. For example, the critical points of $P(t) = t^3 + t$ are $\pm i/\sqrt{3}$. Part 2 applies to this cubic as well.

Schinzel proved the Rényi–Erdős conjecture that $t(f^2)$ tends to infinity with $t(f)$ over $\mathbb{C}$, and proved the same for each fixed power [5]. Schinzel and Zannier give the explicit bound $t(f^m) \geq 2 + \log d / \log(4m)$ for dense f of degree $d \geq 1$ [7, Theorem 1]. Schinzel also asked about compositions, including $P(t) = t^3 - t$ [6, p. 152]. Zannier showed that if P, f are nonconstant complex polynomials and $t(P(f)) \leq L$, then $t(f) \leq B(L)$ for a bound independent of P [10, Theorem 1].

Section 2 uses parity to sharpen the square bound. Section 3 uses several periods to make the powers of a single polynomial sparse simultaneously. Periods of order $\sqrt{d}$ have $O_m(\sqrt{d})$ multiples through md. Passing to an ordinary polynomial costs $O_m(\sqrt{d})$ further terms. Section 4 handles the linear term in a composition.

2 Squares

Write $S_k(x) = 1 + x + \cdots + x^{k-1}$. Modulo $z^p - 1$, each residue has p representations as an ordered sum of two residues. Hence $S_p^2 = pS_p$ and $(2S_p - p)^2 = p^2$ in this ring. For $p = 5$, the block of coefficients is $(-3, 2, 2, 2, 2)$ and its cyclic square is $(25, 0, 0, 0, 0)$. We repeat a block of this form and control the terms that appear when the repetitions end.

Proof of Theorem 1.1. Separating the two parities. Given $G = \sum g_n x^n$, form $F = \sum (-1)^{\lfloor n/2 \rfloor} g_n x^n$ by applying the repeating signs $+, +, -, -$. Writing $[x^n]$ for coefficient extraction gives

$$[x^{2k}]F^2 = (-1)^k [x^{2k}] G(x)G(-x),$$
$$[x^{2k+1}]F^2 = (-1)^k [x^{2k+1}] G(x)^2.$$

For indices summing to $2k + 1$ the two signs multiply to $(-1)^k$. For indices summing to $2k$ their product is $(-1)^k$ on even pairs and $-(-1)^k$ on odd pairs. Thus we can count even degrees in the reflected product and odd degrees in the ordinary square.

Building the polynomial. For odd k, set $A_k = (1 + x)S_k$. The geometric-sum formula gives $A_k(x)A_k(-x) = 1 - x^{2k}$. Choose odd $p > 4$, odd $q \geq 1$, and $2 \leq r \leq pq$. Put $B = A_q(x^p)$, so that

$$B(x)B(-x) = 1 - x^{2pq}, \qquad A_p B = (1 + x)(1 + x^p)S_{pq}.$$

We use $G = B(C - M)$, where $C = A_p^2 R$, $R = r - 1 + x^r$, and M is a monomial. The identity $G^2 - (BM)^2 = B^2 C(C - 2M)$ explains the choice of M. The factor $B^2 C$ contains S_{pq}^2, so a double zero of $C - 2M$ at 1 will turn S_{pq}^2 into $(1 - x^{pq})^2$. Now $R'(1) = R(1) = r$, and symmetry gives $A_p'(1)/A_p(1) = p/2$. Thus $C(1) = 4p^2 r$ and $C'(1) = (p + 1)C(1)$. Taking $M = 2p^2 r\, x^{p+1}$ makes $2M$ agree with C in both value and first derivative at 1.

To check density, write $BC = S_{pq}(1 + x)(1 + x^p)A_p R$. Before multiplication by R the positive coefficients fill $[0, pq + 2p]$. The shift by $r \leq pq$ extends this interval through $d = p(q + 2) + r$. Each coefficient of BC is at most $8pr$, the sum of the coefficients of $(1 + x)(1 + x^p)A_p R$. Every nonzero coefficient of BM is at least $2p^2 r > 8pr$. Subtraction therefore leaves every coefficient nonzero. In the interior $2p + r + 1 \leq n < pq$, BC contributes $8pr$ everywhere and BM subtracts $4p^2 r$ once per period. Up to a shift and a factor $4pr$, this is the block $2S_p - p$.

Counting terms. Since $M(-x) = M(x)$, the reflected product is

$$(1 - x^{2pq})(C(x)C(-x) - M(C(x) + C(-x)) + M^2).$$

Here $C(x)C(-x) = (1 - x^{2p})^2 R(x)R(-x)$ has $O(1)$ terms. Only the even degrees of C occur in $C(x) + C(-x)$, at most $p + \lfloor r/2 \rfloor + 1$ terms. The two shifts from $1 - x^{2pq}$ give at most $2p + r + O(1)$ even terms in F^2.

For odd degrees, $(BM)^2$ contributes at most q terms, one for each odd i between 0 and $2q$ at exponent $2(p + 1) + ip$. The remaining contribution is

$$G^2 - (BM)^2 = (1 - x^{pq})^2 H, \qquad H = (1 + x)^2(1 + x^p)^2 R \cdot \frac{C - 2M}{(1 - x)^2}.$$

The double zero makes H a polynomial of degree $4p + 2r$. Each of its three shifts contains at most $2p + r + 1$ odd exponents. Together the two counts give $t(F^2) \le q + 8p + 4r + O(1)$.

Hitting the prescribed degree. The cost $q + 8p$ suggests $p \approx \sqrt{d/8}$ and $q \approx 8p$. Let s be the least odd integer at least $\sqrt{d/8}$. Set $p = s - 2t$, $q = 8(s + 2t) - 1$. As t increases, the cost stays at $16s - 1$, while

$$p(q + 2) = 8s^2 + s - 32t^2 - 2t.$$

Take the first integer $t \ge 0$ with $p(q+2) \le d-2$. The initial excess over $d-2$ is positive and $O(\sqrt{d})$, so $t = O(d^{1/4})$. The last step decreases the product by $64t - 30$. Thus $r = d - p(q + 2)$ is at least 2 and is $O(d^{1/4})$. For large d, both p and q are of order $\sqrt{d}$ and $r \le pq$. The resulting F has degree d and $t(F^2) \le 16s + O(d^{1/4}) = 4\sqrt{2d} + O(d^{1/4})$. □

3 Several powers at once

Start with the square and cube. Let $\nu = e^{\pi i/3}$ and take the four weights $1, i, \nu, i\nu$. Their sum is $W = (1 + i)(1 + \nu) \neq 0$. Their squares sum to zero because $i^2 = -1$, and their cubes sum to zero because $\nu^3 = -1$.

Give each weight w_i its own period p_i, with the periods pairwise coprime. Write μ_p for the set of pth roots of unity. Assign w_i to each root in $\mu_{p_i} \setminus \{1\}$. This gives four disjoint sets, and the geometric-sum formula gives

$$\sum_{\xi \in \mu_{p_i} \setminus \{1\}} \xi^n = p_i \mathbf{1}_{p_i | n} - 1. \tag{1}$$

For example, take the periods $5, 7, 11, 13$ in that order. At $n = 1$ each set contributes -1 times its weight, so the combined sums with weights w_i, w_i^2 and w_i^3 are $-W, 0, 0$. At $n = 5$ only the first set changes, and its weight is 1, so the sums are $5 - W, 5, 5$. Thus the first sum can be nonzero at every index while the other two vanish away from multiples of the periods.

For one weighted root α, the polynomial $c_\alpha S_\ell(\alpha x)$ has coefficient $c_\alpha \alpha^n$ at degree $0 \le n < \ell$. Its jth power contributes $c_\alpha^j \alpha^n [x^n] S_\ell(x)^j$ at degree n. All roots share the factor $[x^n] S_\ell^j$, so the sum of these contributions has the desired cancellations. The power of a sum also contains mixed products between distinct roots. The next lemma confines their effect to a few short intervals by changing coefficients at the two ends.

Lemma 3.1. *Let T be a set of $r \geq 1$ distinct nonzero algebraic numbers, with a nonzero algebraic weight c_α for each $\alpha \in T$. Suppose every automorphism of $\overline{\mathbb{Q}}$ over $\mathbb{Q}$ permutes the pairs (α, c_α). Write $v_j(n) = \sum_{\alpha \in T} c_\alpha^j \alpha^n$. Fix $m \geq 2$, set $H = 2(m-1)r$, and suppose $d > 2H$. With $\ell = d - H$, there is a polynomial $f \in \mathbb{Q}[x]$ of degree d satisfying $[x^n]f = v_1(n)$ for $H < n < \ell$ and having nonzero coefficients on $[0, H]$ and $[\ell, d]$. For $2 \leq j \leq m$,*

$$[x^n]f^j = v_j(n)\,[x^n]S_\ell(x)^j \qquad \text{for } n \notin I_j, \qquad I_j = \bigcup_{k=0}^{j} [k\ell, k\ell + jH]. \tag{2}$$

In particular, at most $(j+1)(jH+1)$ indices are exceptional.

Proof. Truncate a geometric series. For $h = 0, \ell$, put

$$A_h(x) = \sum_{\alpha \in T} \frac{c_\alpha \alpha^h}{1 - \alpha x}.$$

The coefficients of A_0 are $v_1(0), v_1(1), \ldots$. Subtracting $x^\ell A_\ell$ removes the tail starting at degree ℓ. Before taking this difference, we correct each A_h by a polynomial P_h.

Cancel the unwanted terms at each pole. Near $x = \alpha^{-1}$, use $t = 1 - \alpha x$. One summand of A_h has a pole C/t, where $C = c_\alpha \alpha^h$. The others give a Taylor series $a_0 + a_1 t + \cdots$. On squaring, a_0 produces the extra pole $2Ca_0/t$. To handle all powers through m, choose P_h to cancel $a_0, \ldots, a_{m-2}$ at every pole. Then $B_h = A_h + P_h$ satisfies

$$B_h = \frac{c_\alpha \alpha^h}{1 - \alpha x} + O((1 - \alpha x)^{m-1}) \qquad (x \longrightarrow \alpha^{-1}). \tag{3}$$

This prescribes P_h modulo $(1 - \alpha x)^{m-1}$ at each α. The moduli are pairwise coprime, so the polynomial Chinese remainder theorem gives a unique solution of degree less than $M = (m-1)r$, the degree of their product $D(x)$. The conditions are permuted by conjugation, so uniqueness gives $P_h \in \mathbb{Q}[x]$.

Fill the two ends. We may add any multiple of D to either interpolant. Choose one multiple U of degree $H = 2M$ with every coefficient nonzero. To do this, take $U(x) = D(x)S_{M+1}(tx)$. At each degree $u \leq M$ its coefficient contains the term $D(0)t^u$, and at $M < u \leq 2M$ it contains the leading coefficient of D times t^{u-M}. Each is a nonzero polynomial in t, so one integer t works for all of them. Now set

$$f = B_0 - x^\ell B_\ell = \sum_{n=0}^{\ell-1} v_1(n)x^n + P_0 - x^\ell P_\ell. \tag{4}$$

Adding AU to both P_0 and P_ℓ changes just the two disjoint intervals $[0, H]$ and $[\ell, \ell + H]$. Each coefficient there excludes at most one scalar A. Choose an integer A avoiding all these values. This also makes the coefficient at $\ell + H = d$ nonzero.

Count the exceptional indices. Expand $(B_0 - x^\ell B_\ell)^j$. At the pole α^{-1}, multiplying the pole terms in $B_0^{j-k} B_\ell^k$ gives $c_\alpha^j \alpha^{k\ell}/(1 - \alpha x)^j$. Any product using an error term from (3) has order at least $(m-1) - (j-1) = m - j \geq 0$ there. Subtract this pole term for each α. The remainder E_{jk} has no finite poles and grows at most as x^{jH} at infinity, so it is a polynomial of degree at most jH.

Summing over k now gives

$$f^j = \sum_{\alpha \in T} c_\alpha^j S_\ell(\alpha x)^j + \sum_{k=0}^{j} (-1)^k \binom{j}{k} x^{k\ell} E_{jk}(x).$$

Indeed, for each α the binomial sum of its pole terms has numerator $(1-(\alpha x)^\ell)^j$ and denominator $(1-\alpha x)^j$. The second sum is supported in I_j. At degree n the first sum contributes $[x^n]S_\ell(x)^j$ times $\sum_\alpha c_\alpha^j \alpha^n$, proving (2). □

Proof of Theorem 1.2, part 1. Choose weights and periods. For each $k = 2, \dots, m$, let $\nu_k = e^{\pi i/k}$. Take as weights $w_1, \dots, w_s$ all $s = 2^{m-1}$ products $\prod_{k \in S} \nu_k$ over subsets $S \subseteq \{2, \dots, m\}$, including the empty product and retaining repetitions. Expanding the products below gives

$$W = \sum_i w_i = \prod_{k=2}^{m} (1+\nu_k) \neq 0, \qquad \sum_i w_i^j = \prod_{k=2}^{m} (1+\nu_k^j) = 0 \quad (2 \le j \le m). \tag{5}$$

The vanishing factor is the one with $k = j$. For the periods, fix the first s primes q_i and take p_i to be the least power of q_i exceeding $\sqrt{d}$. Then $\sqrt{d} < p_i \le q_i\sqrt{d}$, and the p_i are pairwise coprime.

Make the coefficients rational. Let K be any number field containing $\nu_2, \dots, \nu_m$, and let $e = [K : \mathbb{Q}]$. Take one copy of the roots for each embedding $\sigma : K \hookrightarrow \mathbb{C}$. To separate these copies, choose a primitive element θ of K over $\mathbb{Q}$ and set $\lambda = 1 + \eta\theta$ for a small positive rational η. Use the points $\sigma(\lambda)\xi$ in copy σ. Two different roots ξ give different limits as $\eta \to 0$. For the same root and distinct embeddings, the linear terms in η differ because the conjugates of θ are distinct. There are finitely many pairs, so a single η makes all points distinct and $\lambda \neq 0$.

Give each point $\sigma(\lambda)\xi$ with $\xi \in \mu_{p_i} \setminus \{1\}$ weight $\sigma(\gamma w_i)$ for a common $\gamma \in K$. Conjugation permutes the pairs. By (1) their weighted sum is

$$v_j(n) = \mathrm{Tr}_{K/\mathbb{Q}}\left(\gamma^j \lambda^n \left(\sum_{i : p_i \mid n} p_i w_i^j - \sum_i w_i^j\right)\right), \tag{6}$$

where the trace sums over the e embeddings. For $1 \le n \le d$ at most one p_i divides n, since any two have product greater than d. At $j = 1$ the expression in parentheses is either $-W$ or $p_i w_i - W$. Both are nonzero for large d, since $|w_i| = 1$ and $p_i > s \ge |W|$. Thus $v_1(n) = \mathrm{Tr}(\gamma z_n)$ with $z_n \neq 0$. This linear form is nonzero, as its value at z_n^{-1} is e. Taking $\gamma = 1 + t\theta + \cdots + t^{e-1}\theta^{e-1}$ turns each form into a nonzero polynomial in t because $1, \theta, \dots, \theta^{e-1}$ span K. One integer t avoids their finitely many roots.

Count the surviving terms. For $2 \le j \le m$, the last sum in (6) vanishes by (5). Thus $v_j(n) = 0$ unless some p_i divides n. There are $r = e\sum_i (p_i - 1) = O_{m,e}(\sqrt{d})$ roots, so $H = 2(m-1)r = O_{m,e}(\sqrt{d})$ and $d > 2H$ for large d. Lemma 3.1 gives a dense f of degree d and

$$t(f^j) \le (j+1)(jH+1) + \sum_{i=1}^{s} (1 + \lfloor jd/p_i \rfloor) = O_{m,e}(\sqrt{d}). \tag{7}$$

The first term counts the exceptional intervals and the sum counts multiples of the periods. Taking $K = \mathbb{Q}(\nu_2, \dots, \nu_m)$ makes e depend only on m. Multiplying f by a common denominator gives the desired $F \in \mathbb{Z}[x]$. □

4 Prescribed compositions

Even when $F^2, \dots, F^m$ are sparse, the linear term of $P(F)$ can contribute $d+1$ terms. For $P(t) = t^3 + t$, shifting by c changes that term to $(3c^2+1)F$. We seek a short J for which JF is sparse, then choose c so that $J \mid 3c^2+1$. For general P we need $J \mid P'(c)$.

Proof of Theorem 1.2, part 2. Find a short multiplier. Adjoin one root a of P' to the field K in Section 3 before constructing F. This increases $e = [K : \mathbb{Q}]$ by a factor of at most $m-1$, so (7) still gives $O_m(\sqrt{d})$ terms in every power from 2 to m. Put

$$J(x) = \prod_{\sigma : K \hookrightarrow \mathbb{C}} (1 - \sigma(\lambda)x) \in \mathbb{Q}[x].$$

To bound $t(JF)$, use (4) before clearing denominators. The endpoint corrections have $O_m(\sqrt{d})$ terms. In (6), the sum over i with $p_i \mid n$ likewise contributes only $O_m(\sqrt{d})$ positions to the truncation. Multiplication by J increases these counts by at most a factor of $e+1$. The remaining part is

$$-\sum_{\sigma} \sigma(\gamma W) S_\ell(\sigma(\lambda)x).$$

In each geometric sum J cancels the denominator, leaving $1 - (\sigma(\lambda)x)^\ell$ times a polynomial of degree $e-1$. This gives at most $2e$ terms per summand, so $t(JF) = O_m(\sqrt{d})$.

Choose the shift. The e conjugates $\sigma(\lambda)$ are distinct and nonzero. Interpolate $c(x)$ of degree at most e with $c(0) = 0$ and $c(1/\sigma(\lambda)) = \sigma(a)$ for each σ. The data are permuted by conjugation, so $c \in \mathbb{Q}[x]$. At each root of J, c takes a root of P'. Since the roots of J are simple, $J \mid P'(c)$.

Choose an integer A for which $g = c + AF$ is dense of degree d. Each coefficient excludes at most one A, since F is dense and $\deg c \le e < d$ for large d. Taylor expansion gives

$$P(g) = P(c) + A \cdot \frac{P'(c)}{J} \cdot (JF) + \sum_{j=2}^{m} \frac{A^j}{j!} \cdot P^{(j)}(c)\, F^j.$$

Each multiplier and $P(c)$ has degree at most me. Thus each product uses only $O_m(1)$ shifts of a polynomial with $O_m(\sqrt{d})$ terms. Finally $g(0) = AF(0)$ is an integer. If N is a positive common denominator of the coefficients of g, then $f_P(x) = g(Nx)$ is a dense integer polynomial. The identity $P(f_P(x)) = (P \circ g)(Nx)$ preserves every nonzero exponent of the composition. Both the constant and the degree threshold depend only on m. □

Open questions.

1. Can a dense integer polynomial f satisfy $t(f^2) < \sqrt{\deg f}$?
2. Do there exist an integer $m \ge 3$ and a constant $c > 0$ such that every dense integer polynomial f of sufficiently large degree d satisfies $\max_{2 \le j \le m} t(f^j) \ge d^c$?

References

[1] D. Coppersmith and J. H. Davenport, Polynomials whose powers are sparse, *Acta Arith.* **58** (1991), 79–87.

[2] Epoch AI, Dense polynomials whose squares are sparse, *FrontierMath Open Problems*, 2026.

[3] P. Erdős, On the number of terms of the square of a polynomial, *Nieuw Arch. Wisk.* (2) **23** (1949), 63–65.

[4] A. Rényi, On the minimal number of terms of the square of a polynomial, *Hungarica Acta Math.* **1** (1947), no. 2, 30–34.

[5] A. Schinzel, On the number of terms of a power of a polynomial, *Acta Arith.* **49** (1987), 55–70.

[6] A. Schinzel, Solved and unsolved problems on polynomials, in *Panoramas of Mathematics*, Banach Center Publ. **34** (1995), 149–159.

[7] A. Schinzel and U. Zannier, On the number of terms of a power of a polynomial, *Rend. Lincei Mat. Appl.* **20** (2009), 95–98.

[8] Unsolved Lab, Dense polynomials, sparse squares, research report, September 2026.

[9] W. Verdenius, On the number of terms of the square and the cube of polynomials, *Indag. Math.* **11** (1949), 459–465.

[10] U. Zannier, On composite lacunary polynomials and the proof of a conjecture of Schinzel, *Invent. Math.* **174** (2008), 127–138.