Extremal no-three-in-line subsets of a modular hyperbola in the Hall–Jackson–Sudbery–Wild window

Alex Komang*

3 September 2026 — version 1.18

Abstract

Let p be an odd prime, $c \in \mathbb{F}_p^*$, and $H(c,p) = \{(x,y) : xy \equiv c \pmod p\}$ the modular hyperbola in the $2p \times 2p$ window of Hall, Jackson, Sudbery and Wild (1975), whose construction keeps $3(p-1)$ of its points with no three collinear — still the best known lower bound for the no-three-in-line problem. Whether that deletion is optimal inside its window has remained unproved, and was stated without proof by Kovács, Nagy and Szabó. We prove it, for every c, and determine the complete extremal structure: the rich lines have slope ± 1 and number $\frac{3}{2}(p-1) - s$; even the fractional LP relaxation equals $3(p-1)$; the number of maximum sets is exactly 9^s, where $s \in \{0,1,2\}$ counts the quadratic residues among $\pm c$, so the HJSW set is the unique maximum iff $p \equiv 1 \pmod 4$ and c is a non-residue; an orbit lemma gives the exact maximum $12n_2 + 10n_1 + 8n_0 + 6s$ for every position of the window; near-maximum sets are stable; and the largest no-*four*-in-line subset has exactly $\frac{7}{2}(p-1) + s$ points. Beyond one hyperbola: the union $H(1) \cup H(-1)$ satisfies $\alpha \le (115/32 + o(1))(p-1)$, refined to $(3.4482\ldots + o(1))(p-1)$ by a block decomposition along runs of consecutive squares of $\mathbb{F}_p$; and no conic, cubic graph or permutation monomial x^k reaches the hyperbola's $3(p-1)$ — closing the natural algebraic routes toward Green's Problem 72. The constants of the two final sections are fixed by explicit numerical quadrature rather than interval enclosure and are labelled accordingly.

Extended summary

Let p be an odd prime, $c \in \mathbb{F}_p^*$, and let $H(c,p) = \{(x,y) \in \mathbb{Z}^2 : xy \equiv c \pmod p\}$ be the modular hyperbola. Hall, Jackson, Sudbery and Wild (1975) observed that the $2p \times 2p$ window $G(p) = [-\frac{p-1}{2}, \frac{3p-1}{2}] \times [0, 2p-1]$ contains exactly four points of each of the $p-1$ residue classes of $H(c,p)$, at the vertices of a $p \times p$ square, and that keeping three of the four (the twelve outer blocks T_2 of the 4×4 block subdivision of $G(p)$) yields $3(p-1)$ points with no three collinear. Kovács, Nagy and Szabó (2025/26) state that this deletion is optimal, i.e. that $H(c,p) \cap T_2$ is a largest no-three-in-line subset of $H(c,p) \cap G(p)$; the value $3(p-1)$ is therefore not new. We give what is, to our knowledge, the first complete proof of this statement, for every c, and we determine the complete extremal structure: we classify all lines carrying three or more points of $H(c,p) \cap G(p)$ (they have slope ± 1, and their number is $\frac{3}{2}(p-1) - s$), we exhibit an explicit linear-programming certificate for the bound $3(p-1)$ (so even the fractional relaxation has value $3(p-1)$), and we show that the number of maximum no-three-in-line subsets is exactly 9^s, where $s \in \{0,1,2\}$ is the number of quadratic residues among c and $-c$. In particular the HJSW set is the unique maximum if and only if $p \equiv 1 \pmod 4$ and c is a quadratic non-residue. The bound does not depend on the position of the window: for *every* $2p \times 2p$ box $[x_0, x_0 + 2p) \times [y_0, y_0 + 2p)$ the modular hyperbola contains at most $3(p-1)$ points in general position, and we determine the exact maximum and the number of maximum sets for every box (an "orbit lemma": a box separates 0 or 2 of the four partner pairs of each generic orbit of the Klein four-group

*Independent researcher, Bali, Indonesia. See the AI/tool-use disclosure at the end.

generated by $(a,b) \mapsto (-b,-a)$ and $(a,b) \mapsto (b,a)$, and the maximum is $12n_2 + 10n_1 + 8n_0 + 6s$ in terms of the numbers n_2, n_1, n_0 of generic orbits with 2, 1, 0 split pairs of each kind); the HJSW box is among the boxes where $3(p-1)$ is attained. The classification is orbit-wise, which makes the problem stable: a lawful set of size $3(p-1)-t$ differs from a maximum one in at most t points (Proposition 19). The classification also gives the no-four-in-line analogue: the largest no-four-in-line subset of $H(c,p) \cap G(p)$ has exactly $\frac{7}{2}(p-1)+s$ points, so the set S_3 of Kovács–Nagy–Szabó is optimal as well. For the union of two hyperbolae $H(1) \cup H(-1)$ in the HJSW box we prove $\alpha \le 4(p-1) - 4m_8(p)$ by an explicit weighted line cover, where $m_8(p)$ (the number of eight-point diagonals) is positive for all $19 \le p \le 1500$ and equals $(1/12 + o(1))\,p$ (by Bombieri's estimate along a cubic), so that $\alpha \le (11/3 + o(1))(p-1)$; to our knowledge this is the first bound below the trivial $4(p-1)$ for a union of two hyperbolae. A second block, the residue groups with a seven-point line, is exactly bookkept (two Klein orbits meet iff their residues satisfy $e^2 \equiv d^2 \pm 4$) and improves this to $\alpha \le (115/32 + o(1))(p-1)$. In fact the whole point set decomposes into blocks indexed by the maximal runs of consecutive quadratic residues, with every row, every column and every slope-(± 1) line inside one block; the type of a run is the descent pattern of k i.i.d. uniform variables, and summing the exact optimum of each block against the descent densities gives $\alpha \le \frac{427101431}{123863040}(p-1) + o(p) = (3.4482 + o(1))(p-1)$, essentially the ceiling of what rows, columns and slope-(± 1) lines can certify (the linear relaxation with all of them is itself a direct sum over the blocks and is $\approx 3.44\,(p-1)$). For an arbitrary second hyperbola $H(k)$ a "potential" cover along the cycles of the row/column incidence graph gives an exact bound $\alpha \le 4(p-1) - (2G_8 - 2\sum_c |A_c - B_c|)/R$ in terms of the number G_8 of eight-point lines and their distribution along the cycles, and a character-sum estimate along the partner curve turns it into $\alpha \le 4(p-1) - c\sqrt{p}/\log^5 p$ for every k of multiplicative order at least $C\sqrt{p}\log^5 p$, in particular for every primitive root. Finally, no cubic graph reaches the hyperbola: for every cubic $f \in \mathbb{F}_p[x]$ and every $2p \times 2p$ box the lifted graph $Y \equiv f(X)$ has no-three-in-line subsets of size at most $(11/8 + o(1)) \cdot 2p < 3(p-1)$; for the permutation cubics ($p \equiv 2 \bmod 3$) this is an explicit cover by the lines of slope ± 1 whose cost is computed exactly on the conic of root pairs, for the other cubics a row count; the same cover handles the permutation monomials x^k, $k \in \{3,5,7\}$, with explicit rational constants ($\approx 1.40 \cdot 2p$). Adding to the ± 1-cover a matching of the "good" lines of the family $(1,3)$ (a local law made exact by a Kummer-type equidistribution lemma) improves the permutation-cubic bound to $(1.326 + o(1)) \cdot 2p$, hence every cubic graph has at most $(\frac{4}{3} + o(1)) \cdot 2p$ points; the constant is a numerical quadrature of explicit integrands, with a margin of more than ten times its error. The proofs are elementary and self-contained; independent programs check the statements about the HJSW box for all $p \le 101$ (all c for $p \le 31$), the bound $3(p-1)$ for all boxes with $p \le 61$, and the box formula and orbit lemma for all boxes with $p \le 13$ and $p \le 31$ respectively.

1 Introduction

The no-three-in-line problem asks for the largest subset of the $n \times n$ grid with no three collinear points; at most $2n$ points are possible, and the best general lower bound, $\frac{3}{2}n - o(n)$, is due to Hall, Jackson, Sudbery and Wild [1]: for $n = 2p$, p prime, they place $3(p-1)$ points of a modular hyperbola $H(c,p)$ in the window $G(p)$. Their construction rests on two facts: every line meets $H(c,p)$ in at most two residue classes, and $G(p)$ contains exactly four points of each class, at the vertices of a $p \times p$ square; keeping the three copies outside the middle block of the 4×4 block subdivision of $G(p)$ leaves no three collinear.

Kovács, Nagy and Szabó [2] recently used the same window and its refinements for the no-$(k+1)$-in-line problem. In their setting the HJSW set is the case $|W| = 1$, $k = 2$ of a general randomised construction (all four copies of every class are taken, then a minimum set R of points is deleted so that every line of slope ± 1 retains at most k points), and they state that for $|W| = 1$, $k = 2$ "Subsection 2.2 essentially details an optimal choice for the points R to be removed to gain the largest no-3-in-line set $S_2(c,p)$ inside $S(W)$" [2, App. A]. The general estimate of [2, Rem. 3.4], $|R| \geq \frac{1}{2}X$ with $X = \sum_\ell (|\ell \cap S(W)| - k)^+$, gives for $|W| = 1$, $k = 2$ the bound $|R| \geq p - 1 - s$ (see Remark 9 below), i.e. $|S| \leq 3(p-1) + s$ for every no-three-in-line $S \subseteq H(c,p) \cap G(p)$: exact when neither c nor $-c$ is a square modulo p, but one or two short of the truth otherwise. We are not aware of a proof of the exact statement in [2] or elsewhere; the purpose of this note is to give one, and to determine the whole extremal structure of the problem. To be explicit about what is new: the value $3(p-1)$ and the qualitative optimality of S_2 and S_3 are asserted in [2]; the proof, the classification of the rich lines, the LP certificate, the count 9^s with its uniqueness criterion, the exact value and count for every $2p \times 2p$ box, and the exact no-four-in-line value and count are, as far as we could determine, new.

Context. Modular hyperbolas $xy \equiv c \pmod m$ have a rich literature (see Shparlinski's survey [15]; for combinatorial-geometric questions about them, e.g. [16]); the general problem of selecting a largest subset in general position from a given point set is NP-hard and APX-hard [17], with general bounds in [18], whereas here the structure of the point set makes the problem solvable exactly. Variants of the no-three-in-line problem with exact structural answers include the extensible version [19, 20] and the torus version [21]. For $k \geq 3$ points per line the asymptotic answer kn was recently established [22], while for $k = 2$ the HJSW construction is still the best known general lower bound (see Green's Problem 72 [23]); the state of the art for exact $2n$-point configurations is documented in Flammenkamp's database [24] and in [25].

Results. Throughout, p is an odd prime, $h = (p-1)/2$, $c \in \mathbb{F}_p^*$, and

$$H(c,p) = \{(x,y) \in \mathbb{Z}^2 : \ xy \equiv c \pmod p\}, \qquad G(p) = \{-h, \dots, 3h+1\} \times \{0, \dots, 2p-1\},$$

and $\mathcal{P} := H(c,p) \cap G(p)$. A set is *lawful* (no-three-in-line, in general position) if it contains no three collinear points. A line is *rich* if it contains at least three points of $\mathcal{P}$; $\mathcal{L}$ denotes the set of rich lines. Let $s = [\, c \text{ is a QR}\,] + [\, -c \text{ is a QR}\,] \in \{0,1,2\}$ (QR = quadratic residue modulo p).

Theorem 1. *(i) Every rich line has slope $+1$ or -1; there are exactly $\frac{3}{2}(p-1) - s$ rich lines, $p-1$ of them carrying three points of $\mathcal{P}$ and $\frac{1}{2}(p-1) - s$ carrying four (Proposition 8 lists them); exactly $2s$ points of $\mathcal{P}$ lie on no rich line.*

(ii) Every lawful subset of $\mathcal{P}$ has at most $3(p-1)$ points; the bound is attained. Even the fractional relaxation (maximise $\sum_q x_q$ subject to $\sum_{q \in \ell} x_q \leq 2$ for $\ell \in \mathcal{L}$, $0 \leq x_q \leq 1$) has optimum $3(p-1)$.

(iii) The number of lawful subsets of size $3(p-1)$ is exactly 9^s. Every maximum set coincides with the HJSW set $H(c,p) \cap T_2$ outside the (at most four) classes $(a, \pm a)$ with $a^2 = \pm c$.

Corollary 2. *If $p \equiv 3 \pmod 4$ then $s = 1$ for every c: there are exactly 9 maximum sets. If $p \equiv 1 \pmod 4$ then $s = 2$ for c a QR (81 maximum sets) and $s = 0$ for c a non-residue: the HJSW set is the unique maximum. Thus $H(c,p) \cap T_2$ is the unique largest lawful subset of $H(c,p) \cap G(p)$ if and only if $p \equiv 1 \pmod 4$ and c is a quadratic non-residue.*

Theorem 3. *Let $x_0, y_0 \in \mathbb{Z}$ and $W = [x_0, x_0+2p) \times [y_0, y_0+2p)$. Every lawful subset of $H(c,p) \cap W$ has at most $3(p-1)$ points. More precisely, the maximum equals $12n_2 + 10n_1 + 8n_0 + 6s$, where n_2, n_1, n_0 are the numbers of generic orbits of the Klein four-group whose partner pairs are split by the box in the patterns $\{(0,2),(2,0)\}$, $(1,1)$, $(0,0)$ respectively ($n_2 + n_1 + n_0 = \frac{1}{4}(p-1-2s)$; no other pattern occurs), and the number of maximum sets is $144^{n_1}\, 1296^{n_0}\, 9^{s_1} 6^{s_2}$, where s_1 (resp. s_2) is the number of exceptional orbits whose pair is split (resp. shared), $s_1 + s_2 = s$. In particular $3(p-1)$ is attained iff $n_1 = n_0 = 0$; the HJSW box has $n_2 = \frac{1}{4}(p-1-2s)$, $s_1 = s$. (Maximum sets are counted as subsets of $\mathbb{Z}^2$, not up to symmetry. The pattern of an orbit and the split/shared label of an exceptional pair depend only on the orbit and the box, not on the class chosen to name the orbit.)*

The proof of Theorem 1 is a double count over the rich lines: for a lawful S, $|S| \le Z + 2|\mathcal{L}|$ where Z is the number of points on no rich line, and the identity $Z + 2|\mathcal{L}| = 3(p-1)$ follows from the classification of the rich lines. The count is exact precisely because the $2s$ points that lie on no rich line are accounted separately; this is the correction to the averaging bound of [2] mentioned above. All rich lines lie inside orbits of a Klein four-group acting on the classes, and the equality analysis reduces to two small "gadgets" (Figure 1), which gives 9^s. For a general box the rich lines still lie inside the orbits, so the maximum is a sum of orbit maxima; the orbit lemma (Section 5) says that a box splits 0 or 2 of the four partner pairs of a generic orbit, and the four possible patterns have maxima 8, 10, 12, 12 — whence the exact formula of Theorem 3. For boxes other than the HJSW box the maximum can be smaller than $3(p-1)$ (Table 1). The four-point rich lines are pairwise disjoint, which settles the no-four-in-line analogue as well (Corollary 15: the maximum is $\frac{7}{2}(p-1)+s$ in the HJSW box, attained by the set S_3 of [2]). Section 6 turns to unions of two hyperbolae: for $H(1) \cup H(-1)$ in the HJSW box, Theorem 21 gives $\alpha \le 4(p-1) - 4m_8(p)$ by an explicit line cover built from the eight-point diagonals and the two reflections of the box, and Proposition 26 evaluates $m_8(p) = (1/12 + o(1))p$; the seven-point groups add a second block (Theorem 30, Proposition 32: $\alpha \le (115/32 + o(1))(p-1)$), and Theorem 35 decomposes $\mathcal{P}_{-1}$ into blocks along the runs of consecutive quadratic residues, whose types obey a descent law (Proposition 37) and which give $\alpha \le (3.4482 + o(1))(p-1)$ (Corollary 39); Corollary 12 shows that among all conics only the shifted hyperbolae can reach $3(p-1)$ in a $2p \times 2p$ box. Section 7 treats an arbitrary second hyperbola $H(k)$: an explicit "potential" cover along the cycles of the row/column incidence graph gives $\alpha(\mathcal{P}_k) \le 4(p-1) - (2G_8 - 2\sum_c |A_c - B_c|)/R$ (Theorem 41), which is below the trivial bound whenever the eight-point lines are not too unevenly distributed along the cycles. Section 8 shows that cubic graphs $Y \equiv f(X)$ never reach the hyperbola's $3(p-1)$ in a $2p \times 2p$ box (Corollary 48: at most $(11/8 + o(1)) \cdot 2p$ points), and Section 9 extends the same ± 1-line cover to the permutation monomials x^k, $k \in \{3,5,7\}$ (Theorem 54), in the direction of Green's remark on Problem 72 of [23] that algebraic constructions modulo p should not beat $\frac{3}{2}N$; Section 10 adds a matching to the ± 1-cover and obtains the strong form $(\frac{4}{3} + o(1)) \cdot 2p$ for every cubic graph (Theorem 62, Corollary 63). Section 11 collects the computations that motivate the (unproved) expectation that unions of hyperbolae gain only $O(1)$ over $3(p-1)$.

2 Notation

Two points are *congruent* if their coordinates agree modulo p; a *class* is a congruence class, identified with an element of $\mathbb{F}_p \times \mathbb{F}_p$. Points of $H(c,p)$ have both coordinates prime to p, so the classes of $H(c,p)$ are the $p-1$ pairs $\kappa = (a,b) \in \mathbb{F}_p^* \times \mathbb{F}_p^*$ with $ab = c$. Since each coordinate range of $G(p)$ is an interval of $2p$ consecutive integers, it contains exactly two representatives of every residue class; hence every class κ of $H(c,p)$ has exactly four points in $G(p)$ and $|\mathcal{P}| = 4(p-1)$. We write (x_κ, y_κ) for the *base copy* of κ, its unique point in $\{-h,\dots,h\} \times \{1,\dots,p-1\}$ (necessarily $x_\kappa \ne 0$), and

$$\kappa(r,s) := (x_\kappa + rp,\ y_\kappa + sp), \qquad r,s \in \{0,1\},$$

for its four points in $G(p)$. Following [1, 2] we sort the classes into four *types* according to the position of the base copy:

$$A:\ x_\kappa > 0,\ y_\kappa > h; \qquad B:\ x_\kappa < 0,\ y_\kappa > h; \qquad C:\ x_\kappa > 0,\ y_\kappa \le h; \qquad D:\ x_\kappa < 0,\ y_\kappa \le h.$$

(In the notation of [2, Def. 2.3]: $\kappa(0,0) \in A_{0,0}$, $B_{-1,0}$, $C_{0,0}$, $D_{-1,0}$ respectively; the HJSW set T_2 omits, for each type, the copy $A_{0,0}$, $B_{0,0}$, $C_{0,1}$, $D_{0,1}$, i.e. $\kappa(0,0)$ for A, $\kappa(1,0)$ for B, $\kappa(0,1)$ for C, $\kappa(1,1)$ for D; the union of these omitted blocks is the middle block M of [2, §2.3]. We call this copy the *middle copy* of κ.) We write n_A, n_B, n_C, n_D for the numbers of classes of each type, so $n_A + n_B + n_C + n_D = p - 1$. Finally, for a class κ put

$$d_\kappa := x_\kappa - y_\kappa, \qquad e_\kappa := x_\kappa + y_\kappa, \qquad D_\kappa := \{x - y = d_\kappa\}, \qquad E_\kappa := \{x + y = e_\kappa + p\};$$

D_κ (the *diagonal* of κ) contains $\kappa(0,0), \kappa(1,1)$ and E_κ (the *antidiagonal*) contains $\kappa(0,1), \kappa(1,0)$. For a real number x let $x^+ = \max\{x, 0\}$.

Three involutions act on the classes of $H(c,p)$:

$$\sigma(a,b) = (-b,-a), \qquad \tau(a,b) = (b,a), \qquad \nu(a,b) = (-a,-b) = \sigma\tau(a,b) = \tau\sigma(a,b),$$

so $V = \{1, \sigma, \tau, \nu\}$ is a Klein four-group acting on the $p-1$ classes. σ has fixed classes iff $-c$ is a QR (then exactly two: $(a,-a)$ with $a^2 = -c$), τ has fixed classes iff c is a QR (two: (a,a) with $a^2 = c$), and ν has none. Put $f_\sigma := \#\{\kappa : \sigma\kappa = \kappa\} \in \{0,2\}$, $f_\tau := \#\{\kappa : \tau\kappa = \kappa\} \in \{0,2\}$; thus $f_\sigma + f_\tau = 2s$. The V-orbits are of size 4 (*generic*: four distinct classes) or of size 2 (*exceptional*: $\{\kappa, \nu\kappa\}$ with κ and $\nu\kappa$ both fixed by σ, or both fixed by τ); there are exactly s exceptional orbits.

3 Rich lines

Lemma 4. *Let ℓ be a line containing at least two points of $H(c,p)$. Then the points of $H(c,p) \cap \ell$ lie in at most two classes. More precisely, if ℓ contains a point of class κ, then the classes met by ℓ are: only κ if ℓ has slope 0 or ∞; among $\{\kappa, \sigma\kappa\}$ if ℓ has slope $+1$; among $\{\kappa, \tau\kappa\}$ if ℓ has slope -1.*

Proof. Let $(u,v) \in \mathbb{Z}^2$ be a primitive direction vector of ℓ and $(x_0, y_0) \in H(c,p) \cap \ell$. The lattice points of ℓ are $(x_0 + tu, y_0 + tv)$, $t \in \mathbb{Z}$, and such a point lies on $H(c,p)$ iff $uv\,t^2 + (x_0 v + y_0 u)\,t \equiv 0 \pmod p$ (we used $x_0 y_0 \equiv c$). The class of the point depends only on $t \bmod p$. If $uv \not\equiv 0$, the congruence has at most two roots modulo p; if exactly one of u, v is divisible by p (both cannot be, as $\gcd(u,v) = 1$), it reduces to $x_0 v\,t \equiv 0$ or $y_0 u\,t \equiv 0$ with a unit coefficient, so $t \equiv 0$: one class. For $(u,v) = (1,0), (0,1)$ this gives the claim for slopes $0, \infty$. For $(u,v) = (1,1)$ the roots are $t \equiv 0$ and $t \equiv -(x_0 + y_0)$, the latter giving the point $(-y_0, -x_0)$ modulo p, of class $\sigma\kappa$. For $(u,v) = (1,-1)$ the roots are $t \equiv 0$ and $t \equiv y_0 - x_0$, giving (y_0, x_0), of class $\tau\kappa$. $\square$

Lemma 5. *Two distinct points of one class in $G(p)$ differ by $\pm(p,0)$, $\pm(0,p)$, $\pm(p,p)$ or $\pm(p,-p)$; the line through them has slope $0, \infty, +1$ or -1 and contains no third point of that class. The point $\kappa(r,s)$ lies on the slope-$(+1)$ line $x - y = d_\kappa + (r-s)p$ and on the slope-(-1) line $x + y = e_\kappa + (r+s)p$. In particular D_κ and E_κ are the only lines of slope ± 1 carrying two points of κ.*

Proof. The four points of a class are the vertices of an axis-parallel square of side p; no three are collinear. The rest is a direct computation. $\square$

Corollary 6. *A rich line has slope $+1$ or -1. Consequently a set $S \subseteq \mathcal{P}$ is lawful iff $|S \cap \ell| \le 2$ for every $\ell \in \mathcal{L}$.*

Proof. By Lemma 4 three points of $\mathcal{P}$ on a line lie in at most two classes, so two of them are congruent, and by Lemma 5 the line has slope $0, \infty, \pm 1$. A line of slope 0 or ∞ meets a single class (Lemma 4), which has at most two points on it (Lemma 5). $\square$

Lemma 7 (position of the partner classes)**.** *For every class κ the types of $\sigma\kappa$, $\tau\kappa$ and the shifts $d_{\sigma\kappa} - d_\kappa$, $e_{\tau\kappa} - e_\kappa$ are given by the following table:*

type of κ	*type of* $\sigma\kappa$	$d_{\sigma\kappa} - d_\kappa$	*type of* $\tau\kappa$	$e_{\tau\kappa} - e_\kappa$
A	A	0	D	$-p$
B	C	$+p$	B	0
C	B	$-p$	C	0
D	D	0	A	$+p$

In particular σ preserves the types A, D and interchanges B, C; τ preserves B, C and interchanges A, D; ν interchanges $A \leftrightarrow D$ and $B \leftrightarrow C$; the fixed classes of σ lie in $A \cup D$, those of τ in $B \cup C$; and $n_A = n_D$, $n_B = n_C$.

Proof. Write $\kappa = (a, b)$ with base copy (x_κ, y_κ), so $x_\kappa \equiv a$, $y_\kappa \equiv b$. The base copy of $\sigma\kappa = (-b, -a)$ has first coordinate $-y_\kappa$ if $y_\kappa \leq h$ and $p - y_\kappa$ if $y_\kappa > h$, and second coordinate $p - x_\kappa$ if $x_\kappa > 0$ and $-x_\kappa$ if $x_\kappa < 0$ (these are the representatives of $-b$ in $[-h, h]$ and of $-a$ in $[1, p-1]$). Going through the four types: $\kappa \in A$: $(p - y_\kappa, p - x_\kappa)$, again of type A, difference $x_\kappa - y_\kappa$; $\kappa \in D$: $(-y_\kappa, -x_\kappa)$, type D, difference $x_\kappa - y_\kappa$; $\kappa \in C$: $(-y_\kappa, p - x_\kappa)$, type B, difference $x_\kappa - y_\kappa - p$; $\kappa \in B$: $(p - y_\kappa, -x_\kappa)$, type C, difference $x_\kappa - y_\kappa + p$. Similarly, the base copy of $\tau\kappa = (b, a)$ has first coordinate y_κ if $y_\kappa \leq h$ and $y_\kappa - p$ otherwise, and second coordinate x_κ if $x_\kappa > 0$ and $x_\kappa + p$ otherwise: $\kappa \in B$: $(y_\kappa - p, x_\kappa + p)$, type B, sum e_κ; $\kappa \in C$: (y_κ, x_κ), type C, sum e_κ; $\kappa \in A$: $(y_\kappa - p, x_\kappa)$, type D, sum $e_\kappa - p$; $\kappa \in D$: $(y_\kappa, x_\kappa + p)$, type A, sum $e_\kappa + p$. The remaining assertions follow. $\square$

Proposition 8. *$\mathcal{L}$ consists of the following lines, and no others; for each we list its points in $\mathcal{P}$.*

(a) For $\kappa \in A \cup D$ with $\sigma\kappa \neq \kappa$: the common diagonal $D_\kappa = D_{\sigma\kappa}$, with the four points $\kappa(0,0), \kappa(1,1), \sigma\kappa(0,0), \sigma\kappa(1,1)$.

(b) For $\kappa \in B \cup C$ with $\tau\kappa \neq \kappa$: the common antidiagonal $E_\kappa = E_{\tau\kappa}$, with the four points $\kappa(0,1), \kappa(1,0), \tau\kappa(0,1), \tau\kappa(1,0)$.

(c) For every $\kappa \in C$: D_κ with the three points $\kappa(0,0), \kappa(1,1), \sigma\kappa(1,0)$ ($\sigma\kappa \in B$); for every $\kappa \in B$: D_κ with the three points $\kappa(0,0), \kappa(1,1), \sigma\kappa(0,1)$ ($\sigma\kappa \in C$).

(d) For every $\kappa \in A$: E_κ with the three points $\kappa(0,1), \kappa(1,0), \tau\kappa(1,1)$ ($\tau\kappa \in D$); for every $\kappa \in D$: E_κ with the three points $\kappa(0,1), \kappa(1,0), \tau\kappa(0,0)$ ($\tau\kappa \in A$).

The lines listed in (a)–(d) are pairwise distinct (apart from the identifications $D_\kappa = D_{\sigma\kappa}$ in (a) and $E_\kappa = E_{\tau\kappa}$ in (b) that are built into the list), and each of them lies inside a single V-orbit of classes. Consequently

$$|\mathcal{L}| = \frac{n_A + n_D - f_\sigma}{2} + \frac{n_B + n_C - f_\tau}{2} + (n_A + n_D) + (n_B + n_C) = \frac{3}{2}(p-1) - s, \tag{1}$$

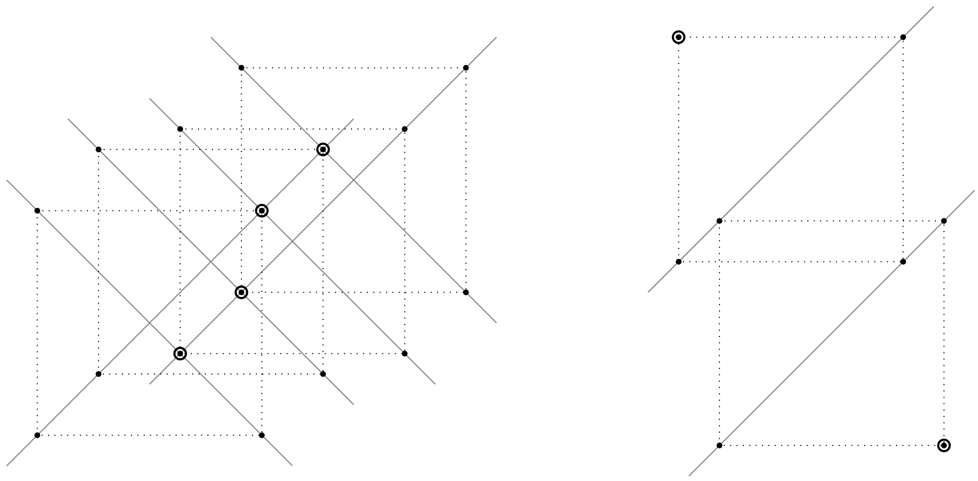
with $p - 1$ three-point lines and $\frac{1}{2}(p-1) - s$ four-point lines. Moreover, the points of $\mathcal{P}$ lying on no *rich line are exactly $\kappa(1,1)$ for σ-fixed $\kappa \in A$, $\kappa(0,0)$ for σ-fixed $\kappa \in D$, $\kappa(0,1)$ for τ-fixed $\kappa \in B$ and $\kappa(1,0)$ for τ-fixed $\kappa \in C$; their number is*

$$Z := \#\{q \in \mathcal{P} : \ q \text{ lies on no rich line}\} = f_\sigma + f_\tau = 2s. \tag{2}$$

The points lying on two *rich lines are exactly the middle copies of the classes κ whose partner ($\sigma\kappa$ for $\kappa \in A \cup D$, $\tau\kappa$ for $\kappa \in B \cup C$) is different from κ; no point lies on three rich lines.*

Proof. Let $\ell \in \mathcal{L}$. By Corollary 6 its slope is ± 1; say $+1$ (the case -1 is symmetric, with τ, e_κ, E_κ in place of σ, d_κ, D_κ). By Lemma 4 the points of $\ell \cap \mathcal{P}$ lie in two classes $\kappa, \sigma\kappa$ (a single class carries at most two points on a line), and one of the classes, say κ, has two points on ℓ; by Lemma 5 these are $\kappa(0,0), \kappa(1,1)$ and $\ell = D_\kappa$, and $\sigma\kappa \neq \kappa$. The points of $\sigma\kappa$ on D_κ are the $\sigma\kappa(r,s)$ with $d_{\sigma\kappa} + (r-s)p = d_\kappa$. By Lemma 7: for $\kappa \in A \cup D$, $r = s$, giving (a); for $\kappa \in C$, $r - s = 1$, giving $\sigma\kappa(1,0)$; for $\kappa \in B$, $r - s = -1$, giving $\sigma\kappa(0,1)$; this is (c). Conversely each listed line has at least three points.

Distinctness: if $D_\kappa = D_\lambda$ then this line of slope $+1$ meets both classes κ, λ, so $\lambda \in \{\kappa, \sigma\kappa\}$ by Lemma 4; and $D_\kappa = D_{\sigma\kappa}$ holds for $\kappa \in A \cup D$ (shift 0) but not for $\kappa \in B \cup C$ (shift $\pm p$). Likewise $E_\kappa = E_\lambda$ forces $\lambda \in \{\kappa, \tau\kappa\}$, and $E_\kappa = E_{\tau\kappa}$ holds exactly for $\kappa \in B \cup C$. A rich line meets only



generic orbit ($p = 11$, $c = 1$, classes $(2,6),(5,9) \in A$, $(6,2),(9,5) \in D$): 16 points, 6 rich lines; the four circled middle copies lie on two rich lines each

exceptional orbit (τ-fixed classes $(1,1) \in C$, $(10,10) \in B$): 8 points, 2 rich lines with 3 points each; the two circled points lie on no rich line

Figure 1: The two gadgets of the equality analysis, drawn to scale for $p = 11$, $c = 1$ (dotted squares: the four copies of a class; grey: rich lines). Generic orbit: the circled points are excluded and the other 12 are forced ($= T_2$). Exceptional orbit: the circled points are forced and each rich line keeps two of its three points, $3 \cdot 3 = 9$ choices.

the classes $\kappa, \sigma\kappa$ (or $\kappa, \tau\kappa$), which lie in one V-orbit. Counting: (a) gives $(n_A + n_D - f_\sigma)/2$ lines, (b) $(n_B + n_C - f_\tau)/2$, (c) $n_B + n_C$, (d) $n_A + n_D$, which is (1).

For the incidences of a single point, take $\kappa \in A$. Its points $\kappa(0,1), \kappa(1,0)$ lie on $E_\kappa \in \mathcal{L}$ (d), and on no line of type (a)/(c) (their slope-(+1) lines $x - y = d_\kappa \pm p$ carry no second point of κ and, by the table, no point of $\sigma\kappa$). $\kappa(1,1)$ lies on D_κ, which is rich iff $\sigma\kappa \neq \kappa$, and on $x + y = e_\kappa + 2p = e_{\tau\kappa} + 3p$, which carries no point of $\tau\kappa$: so $\kappa(1,1)$ is on one rich line if $\sigma\kappa \neq \kappa$ and on none otherwise. $\kappa(0,0)$ lies on D_κ and on $x + y = e_\kappa = e_{\tau\kappa} + p$, i.e. on $E_{\tau\kappa} \in \mathcal{L}$: two rich lines if $\sigma\kappa \neq \kappa$, one otherwise. The types D, B, C are analogous (for B: $\kappa(1,0) \in D_{\sigma\kappa}$ since $d_\kappa + p = d_{\sigma\kappa}$, while $\kappa(0,1)$ lies on $x - y = d_\kappa - p = d_{\sigma\kappa} - 2p$, which carries no point of $\sigma\kappa$; for C the roles of $(1,0), (0,1)$ are exchanged; for D: $\kappa(1,1) \in E_{\tau\kappa}$ since $e_\kappa + 2p = e_{\tau\kappa} + p$, and $\kappa(0,0)$ lies on $x + y = e_\kappa = e_{\tau\kappa} - p$). This gives (2) and the last two assertions. $\square$

4 Proof of Theorem 1 and Corollary 2

Part (i) is Corollary 6 and Proposition 8.

Upper bound. Let $S \subseteq \mathcal{P}$ be lawful. For $q \in \mathcal{P}$ let $m(q)$ be the number of rich lines through q. Every rich line carries at most two points of S (Corollary 6), hence

$$\begin{aligned} |S| &= \#\{q \in S : m(q) = 0\} + \#\{q \in S : m(q) \geq 1\} \leq Z + \sum_{q \in S} m(q) - \sum_{q \in S} (m(q) - 1)^+ \\ &= Z + \sum_{\ell \in \mathcal{L}} |S \cap \ell| - \sum_{q \in S} (m(q) - 1)^+ \leq Z + 2|\mathcal{L}|, \end{aligned} \tag{3}$$

and by (1), (2), $Z + 2|\mathcal{L}| = 2s + 3(p-1) - 2s = 3(p-1)$. The HJSW set attains the bound [1]; the equality analysis below exhibits all sets attaining it, so attainment does not depend on [1].

LP certificate. The same count is valid for fractional $x \in [0,1]^{\mathcal{P}}$ with $\sum_{q \in \ell} x_q \leq 2$ ($\ell \in \mathcal{L}$): $\sum_q x_q \leq \sum_{q \colon m(q) = 0} x_q + \sum_{\ell \in \mathcal{L}} \sum_{q \in \ell} x_q \leq Z + 2|\mathcal{L}|$. In LP-duality terms the dual solution assigns weight 1 to every rich line and weight 1 to the bound $x_q \le 1$ of each of the Z points on no rich line. This proves (ii). (The statement concerns the optimum value; we do not claim that the polytope of the relaxation is integral.)

Equality. $|S| = 3(p-1)$ holds iff all inequalities in (3) are equalities, i.e. iff (I) S contains all Z points lying on no rich line; (II) $|S \cap \ell| = 2$ for every $\ell \in \mathcal{L}$; (III) no point of S lies on two rich lines. Since every rich line lies inside one V-orbit of classes (Proposition 8), conditions (I)–(III) split into independent conditions on the orbits, and the number of maximum sets is the product of the numbers of admissible choices in the orbits. In a generic orbit $\{\kappa, \sigma\kappa, \tau\kappa, \nu\kappa\}$ all four classes have distinct partners; the six rich lines meeting the orbit (two of type (a) or (b), four of type (c) or (d)) meet no other orbit, its four middle copies have $m = 2$ and are excluded by (III), and each of the six lines then retains exactly two points, both with $m = 1$, which are forced into S by (II): on a generic orbit S coincides with $H(c,p) \cap T_2$, and there is exactly one admissible choice. An exceptional orbit $\{\kappa, \nu\kappa\}$ (κ and $\nu\kappa$ fixed by σ, or both fixed by τ) meets exactly two rich lines, each with three points of $m = 1$, and contains two points with $m = 0$: conditions (I)–(III) allow $\binom{3}{2}\binom{3}{2} = 9$ choices, and all of them are lawful by Corollary 6. There are s exceptional orbits, whence 9^s maximum sets, all agreeing with the HJSW set outside the exceptional orbits, i.e. outside the classes $(a, \pm a)$ with $a^2 = \pm c$. This proves (iii).

Corollary 2. -1 is a QR modulo p iff $p \equiv 1 \pmod 4$. If $p \equiv 3 \pmod 4$, exactly one of $c, -c$ is a QR, so $s = 1$; if $p \equiv 1 \pmod 4$, c and $-c$ have the same character, so $s \in \{0, 2\}$ according as c is a non-residue or a residue. $\square$

Remark 9. (a) In the language of [2, Constr. 3.2, Rem. 3.4] with $W = \{c\}$, $k = 2$ (so $S(W) = \mathcal{P}$ and L' is the set of lines of slope ± 1): $X_{2,p}(\{c\}) = \sum_\ell (|\ell \cap \mathcal{P}| - 2)^+ = (p-1) \cdot 1 + (\frac{1}{2}(p-1) - s) \cdot 2 = 2(p-1) - 2s$ by Theorem 1(i), so the general bound $|R| \ge \frac{1}{2}X$ gives $|R| \ge p - 1 - s$ and hence, for every lawful $S \subseteq \mathcal{P}$, $|S| \le |\mathcal{P}| - |R| \le 3(p-1) + s$ (a lawful S is a trimming, by Corollary 6). The averaging loses exactly one point per exceptional orbit, because each such orbit contains two points on no rich line and no point on two rich lines; the count (3) keeps track of these points and is exact.

(b) The theorem says that no lawful set can use the middle block M profitably: adding fourth copies to the HJSW set is impossible without removing at least as many points. Any improvement of the HJSW construction must therefore use points outside $H(c,p) \cap G(p)$ (another hyperbola, a larger window — by Theorem 3 not merely a shifted $2p \times 2p$ box — or non-hyperbola points).

(c) The proof uses only: at most two classes per line (Lemma 4), the square arrangement of the four copies, and the position of the partner classes (Lemma 7). It does not use $p \ge 7$; $p = 3$ and $p = 5$ are covered as well.

5 All $2p \times 2p$ boxes

Let $W = [x_0, x_0 + 2p) \times [y_0, y_0 + 2p)$ with $x_0, y_0 \in \mathbb{Z}$ (only x_0, y_0 modulo p matter), $\mathcal{P}_W = H(c,p) \cap W$, and let $\rho_x(u) \in [x_0, x_0 + p)$, $\rho_y(u) \in [y_0, y_0 + p)$ denote the representatives of a residue u. Every class $\kappa = (a,b)$ of $H(c,p)$ again has exactly four points $\kappa(r,s) = (\rho_x(a) + rp, \rho_y(b) + sp)$, $r, s \in \{0,1\}$, in W; put $d_\kappa = \rho_x(a) - \rho_y(b)$, $e_\kappa = \rho_x(a) + \rho_y(b)$, $D_\kappa = \{x - y = d_\kappa\}$, $E_\kappa = \{x + y = e_\kappa + p\}$. Lemmas 4, 5 and Corollary 6 hold verbatim (they do not use the position of the box), so a rich line is D_κ for a class with $\sigma\kappa \ne \kappa$ or E_κ for a class with $\tau\kappa \ne \kappa$, and its points are the two copies $\kappa(0,0), \kappa(1,1)$ (resp. $\kappa(0,1), \kappa(1,0)$) together with the copies of the partner class on the same line. Since $d_\kappa \equiv d_{\sigma\kappa} \pmod p$ and both lie in an interval of length $2p - 2$, $d_{\sigma\kappa} - d_\kappa \in \{-p, 0, p\}$; we call the σ-pair $\{\kappa, \sigma\kappa\}$ *shared* if $d_{\sigma\kappa} = d_\kappa$ (then $D_\kappa = D_{\sigma\kappa}$ carries the four points $\kappa(0,0), \kappa(1,1), \sigma\kappa(0,0), \sigma\kappa(1,1)$) and *split* otherwise (then D_κ carries $\kappa(0,0), \kappa(1,1)$ and one copy of $\sigma\kappa$, and $D_{\sigma\kappa}$ carries $\sigma\kappa(0,0), \sigma\kappa(1,1)$ and one copy of κ). Likewise a τ-pair is shared ($E_\kappa = E_{\tau\kappa}$, four points) or split (two lines with three points). In the HJSW box the σ-pairs in $A \cup D$ and the τ-pairs in $B \cup C$ are shared and the others are split (Lemma 7); in the box $[0, 2p)^2$ all pairs are shared.

Lemma 10 (orbit lemma)**.** *Let $O = \{\kappa, \sigma\kappa, \tau\kappa, \nu\kappa\}$ be a generic orbit, and let e_O (resp. f_O) be the number of split σ-pairs (resp. τ-pairs) among the two σ-pairs and two τ-pairs of O. Then $e_O + f_O \in \{0, 2\}$; in particular $e_O + f_O \le 2$.*

Proof. Write $\kappa = (a, b)$, $X_u = \rho_x(u)$, $X'_u = \rho_x(-u)$, $Y_u = \rho_y(u)$, $Y'_u = \rho_y(-u)$ for $u \in \{a, b\}$. The base copies are $P = (X_a, Y_b)$ of κ, $Q = (X'_b, Y'_a)$ of $\sigma\kappa$, $R = (X_b, Y_a)$ of $\tau\kappa$ and $T = (X'_a, Y'_b)$ of $\nu\kappa$; the σ-pairs are $\{P, Q\}$, $\{R, T\}$ and the τ-pairs $\{P, R\}$, $\{Q, T\}$. With $d(x, y) = x - y$, $e(x, y) = x + y$, the four split conditions are $E_1 : d_P \ne d_Q$, $E_2 : d_T \ne d_R$, $F_1 : e_P \ne e_R$, $F_2 : e_Q \ne e_T$; each of the four differences is $\equiv 0 \pmod p$ and lies in $[-p, p]$, hence is one of $-p, 0, p$.

Let m_x be the least multiple of p exceeding $2x_0$ and $\mu_x = m_x - 2x_0 \in [1, p]$; similarly m_y, μ_y. Since $X_u + X'_u \equiv 0 \pmod p$ and $X_u + X'_u \in [2x_0 + 1, 2x_0 + 2p - 3]$, we have $X_u + X'_u = m_x + \varepsilon_x(u)p$ with $\varepsilon_x(u) \in \{0, 1\}$, and explicitly $\varepsilon_x(u) = [X_u - x_0 > \mu_x]$; likewise $Y_u + Y'_u = m_y + \varepsilon_y(u)p$, $\varepsilon_y(u) = [Y_u - y_0 > \mu_y]$. Put $\Delta = \varepsilon_x(a) - \varepsilon_x(b)$, $\Gamma = \varepsilon_y(a) - \varepsilon_y(b)$, $A = \varepsilon_x(b) - \varepsilon_y(a)$, $B = \varepsilon_x(a) - \varepsilon_y(b)$, and define the integers θ, η by

$$(X_a + X_b) - (Y_a + Y_b) - (m_x - m_y) = \theta p, \qquad (X_b - X_a) - (Y_b - Y_a) = \eta p.$$

Substituting $X'_u = m_x + \varepsilon_x(u)p - X_u$, $Y'_u = m_y + \varepsilon_y(u)p - Y_u$ one finds

$$d_P - d_Q = (\theta - A)p, \quad d_T - d_R = (B - \theta)p, \quad e_R - e_P = \eta p, \quad e_Q - e_T = (\Gamma - \Delta - \eta)p,$$

so that $E_1 \Leftrightarrow \theta \ne A$, $E_2 \Leftrightarrow \theta \ne B$, $F_1 \Leftrightarrow \eta \ne 0$, $F_2 \Leftrightarrow \eta \ne \Gamma - \Delta$, and $\theta - A,\ \theta - B,\ \eta,\ \eta - (\Gamma - \Delta) \in \{-1, 0, 1\}$.

Parity. For $x \in \{-1, 0, 1\}$, $[x \ne 0] \equiv x \pmod 2$. Hence $E_1 + E_2 + F_1 + F_2 \equiv (\theta - A) + (\theta - B) + \eta + (\eta - \Gamma + \Delta) \equiv A + B + \Gamma - \Delta = 2\varepsilon_x(b) - 2\varepsilon_y(b) \equiv 0 \pmod 2$: the number of split pairs in O is even.

If F_1 holds, then E_1 and E_2 do not both hold. Put $s = X_a - x_0$, $t = X_b - x_0 \in [0, p-1]$ and $v = (x_0 - y_0) \bmod p$; since $Y_u \equiv X_u$, we have $Y_a - y_0 = s + v - c_a p$ and $Y_b - y_0 = t + v - c_b p$ with $c_a = [s + v \ge p]$, $c_b = [t + v \ge p]$, and therefore $\eta = c_b - c_a$. So F_1 means $c_a \ne c_b$; the substitution $a \leftrightarrow b$ exchanges $E_1 \leftrightarrow E_2$ and fixes F_1, so we may assume $c_a = 0$, $c_b = 1$, i.e. $s \le p - 1 - v < t$. Note that F_1 forces $v \ge 1$: for $v = 0$ we would have $s + v, t + v \le p - 1$ and hence $c_a = c_b = 0$. Since $\mu_x - \mu_y \equiv -2v \pmod p$ and $|\mu_x - \mu_y| \le p - 1$, we have $\mu_x - \mu_y = -2v + kp$ with $k \in \{0, 1, 2\}$, and $\theta p = (2x_0 + s + t) - (2y_0 + s + t + 2v - p) - (\mu_x + 2x_0 - \mu_y - 2y_0) = p - 2v - (\mu_x - \mu_y) = (1 - k)p$, i.e. $\theta = 1 - k$. With $\varepsilon_x(a) = [s > \mu_x]$, $\varepsilon_x(b) = [t > \mu_x]$, $\varepsilon_y(a) = [s + v > \mu_y]$, $\varepsilon_y(b) = [t + v - p > \mu_y]$ and $\mu_y = \mu_x + 2v - kp$:

- $k = 1$: $E_1 \Leftrightarrow [t > \mu_x] \ne [s > \mu_x + v - p]$ and $E_2 \Leftrightarrow [s > \mu_x] \ne [t > \mu_x + v]$. If $s > \mu_x$ then $t > \mu_x$ and $s > \mu_x + v - p$, so E_1 fails. If $s \le \mu_x$, E_2 forces $t > \mu_x + v$, so $\mu_x + v \le p - 2$ and $[t > \mu_x] = 1$, and E_1 forces $s \le \mu_x + v - p < 0$, impossible.
- $k = 0$: here $\mu_y = \mu_x + 2v \le p$, so $\varepsilon_y(b) = 0$ and $E_2 \Leftrightarrow s \le \mu_x$; then $\varepsilon_y(a) = [s > \mu_x + v] = 0$ and $E_1 \Leftrightarrow t \le \mu_x$, but $t \ge p - v > p - 2v \ge \mu_x$ (the middle inequality uses $v \ge 1$).
- $k = 2$: here $\mu_y = \mu_x + 2v - 2p \ge 1$, so $\mu_x \ge 2p + 1 - 2v$ and $\varepsilon_y(a) = 1$; $E_1 \Leftrightarrow t > \mu_x$ and $E_2 \Leftrightarrow (s > \mu_x) \vee (t \le \mu_x + v - p)$. Now $s > \mu_x$ would give $2p + 2 - 2v \le s \le p - 1 - v$, i.e. $v \ge p + 3$; so $s \le \mu_x$ and E_2 needs $t \le \mu_x + v - p < \mu_x$, contradicting E_1.

Consequently, if F_1 holds then $E_1 + E_2 + F_1 + F_2 \le 3$, and if F_1 fails then again $E_1 + E_2 + F_1 + F_2 \le 3$; by parity the sum is 0 or 2. $\square$

Lemma 11 (gadget maxima)**.** *Let O be an orbit and $\mathcal{P}_O \subseteq \mathcal{P}_W$ its 16 (or 8) points. Every rich line meeting $\mathcal{P}_O$ lies in $\mathcal{P}_O$, and the largest lawful subset of $\mathcal{P}_O$ has size and multiplicity as follows: exceptional orbit: 6 points, 9 sets if the pair is split and 6 if it is shared; generic orbit with $(e_O, f_O) = (0, 0)$: 8 points, $6^4 = 1296$ sets; $(1, 1)$: 10 points, 144 sets; $(0, 2)$ or $(2, 0)$: 12 points, one set.*

Proof. A rich line meets only a class and its σ- or τ-partner (Lemma 4), hence lies in one orbit. *Exceptional orbit* $\{\kappa, \nu\kappa\}$, say $\sigma\kappa = \kappa$: the rich lines are $E_\kappa = E_{\nu\kappa}$ (four points) if the τ-pair is shared, resp. E_κ, $E_{\nu\kappa}$ (three points each) if it is split; the remaining four (resp. two) points are free: $2 + 4 = 6$ points in $\binom{4}{2} = 6$ ways, resp. $2 + 2 + 2 = 6$ points in $3 \cdot 3 = 9$ ways. *Pattern* $(0, 0)$: the two shared diagonals and the two shared antidiagonals are four disjoint four-point lines covering $\mathcal{P}_O$: $4 \cdot 2 = 8$ points in 6^4 ways. *Pattern* $(2, 0)$ (and symmetrically $(0, 2)$): the four diagonals D_λ ($\lambda \in O$) are rich with three points each and the two antidiagonals $E_\kappa = E_{\tau\kappa}$, $E_{\sigma\kappa} = E_{\nu\kappa}$ have four points; every class λ has exactly one point of multiplicity $m = 2$, namely its off-diagonal copy lying on $D_{\sigma\lambda}$ (which is also on the antidiagonal of λ). Each D_λ carries $\lambda(0, 0), \lambda(1, 1)$ ($m = 1$) and one $m = 2$ point, each antidiagonal carries two $m = 1$ and two $m = 2$ points. By (3) restricted to O (with $Z = 0$), $|S \cap \mathcal{P}_O| \le 12$ with equality iff S avoids the four $m = 2$ points and takes both remaining points of each of the six lines — which is lawful; so the maximum is 12 and it is unique. (For the HJSW box this is the generic orbit of Section 4.) *Pattern* $(1, 1)$: the split σ-pair and the split τ-pair share exactly one class κ (any σ-edge and any τ-edge of the four-cycle $\kappa - \tau\kappa - \nu\kappa - \sigma\kappa - \kappa$ meet in one vertex). The six rich lines are

$$
\begin{array}{lllll}
A := D_\kappa, & B := E_\kappa, & A' := D_{\sigma\kappa}, & B' := E_{\tau\kappa} & \text{(three points each)}, \\
A'' := D_{\tau\kappa} = D_{\nu\kappa}, & B'' := E_{\sigma\kappa} = E_{\nu\kappa} & & & \text{(four points each)}.
\end{array}
$$

Write κ_A, κ'_A for the two copies of κ on A (these are $\kappa(0, 0), \kappa(1, 1)$) and κ_B, κ'_B for the two on B; the third point of A is a copy $\bar{\sigma}$ of $\sigma\kappa$ off its own diagonal, hence on B''; the third point of B is a copy $\bar{\tau}$ of $\tau\kappa$ off its own antidiagonal, hence on A''; the third point of A' is a copy of κ off A, i.e. one of κ_B, κ'_B — name it κ_B; the third point of B' is one of κ_A, κ'_A — name it κ_A. No other incidences occur (a rich line of slope $+1$ through a copy of κ is A or A', of slope -1 it is B or B', and $A \cap B = \emptyset$ because a common lattice point would have $x = x_\kappa + p/2$). Hence $m = 2$ exactly for the four points $\kappa_A \in A \cap B'$, $\kappa_B \in B \cap A'$, $\bar{\sigma} \in A \cap B''$, $\bar{\tau} \in B \cap A''$, and $m = 1$ for the other twelve: A and B consist of one $m = 1$ point and two $m = 2$ points, A' and B' of two $m = 1$ points and one $m = 2$ point, A'' and B'' of three $m = 1$ points and one $m = 2$ point. For lawful $S \subseteq \mathcal{P}_O$ let t be the number of $m = 2$ points in S; since every point of $\mathcal{P}_O$ lies on a rich line, $|S| = \sum_\ell |S \cap \ell| - t$ over the six lines, each term being ≤ 2. If $t = 0$ then $|S \cap A|, |S \cap B| \le 1$ and $|S| \le 1 + 1 + 2 + 2 + 2 + 2 = 10$; if $t = 1$, the $m = 2$ point lies on exactly one of A, B, the other of the two carries at most one point of S, and $|S| \le 11 - 1 = 10$; if $t \ge 2$ then $|S| \le 12 - t \le 10$. So the maximum is at most 10, and a set of size 10 makes every inequality used tight. Counting these sets by t (every set listed is lawful, since each rich line receives at most two of its points):

$t = 0$: S contains κ'_A, κ'_B, both $m = 1$ points of A' and both of B' (forced), and two of the three $m = 1$ points of A'' and of B'': $3 \cdot 3 = 9$ sets.

$t = 1$: the $m = 2$ point q is one of $\kappa_A, \bar{\sigma} \in A$ or $\kappa_B, \bar{\tau} \in B$; say $q \in A$ (the case $q \in B$ is symmetric). Then $S \cap A = \{q, \kappa'_A\}$ and $S \cap B = \{\kappa'_B\}$; the second line through q receives q and one of its $m = 1$ points, and each of the three remaining lines receives two of its $m = 1$ points. For $q = \kappa_A$: B' in 2 ways, A' forced, A'' and B'' in $3 \cdot 3$ ways — 18 sets; for $q = \bar{\sigma}$: B'' in 3 ways, A' and B' forced, A'' in 3 ways — 9 sets. With the symmetric case $q \in B$: $2 \cdot (18 + 9) = 54$ sets.

$t = 2$: the two $m = 2$ points cannot both lie on A (then $|S \cap B| \le 1$ and $|S| \le 11 - 2 = 9$), nor both on B; so one is on A and one on B, every line receives exactly two points, and $S \cap A = \{q, \kappa'_A\}$, $S \cap B = \{q', \kappa'_B\}$. For $\{q, q'\} = \{\kappa_A, \kappa_B\}$: B' and A' each receive their $m = 2$ point and one of two $m = 1$ points $(2 \cdot 2)$, A'', B'' two of three $(3 \cdot 3)$: 36 sets; for $\{\kappa_A, \bar{\tau}\}$: B' 2 ways, A'' ($\ni \bar{\tau}$) 3 ways, A' forced, B'' 3 ways: 18; for $\{\bar{\sigma}, \kappa_B\}$: symmetrically 18; for $\{\bar{\sigma}, \bar{\tau}\}$: B'' 3, A'' 3, A', B' forced: 9. Total $36 + 18 + 18 + 9 = 81$.

$t \ge 3$: $|S| \le 12 - t \le 9$: no maximum set.

Altogether $9+54+81=144$ maximum sets of size 10. (An independent brute-force enumeration of the 16-point gadgets, `slack/gadget11_check.py`, confirms the line structure and the counts $9, 54, 81$ for every $(1,1)$-orbit encountered in all boxes for $p \le 13$ and in samples of boxes for $p \le 31$.) $\square$

Proof of Theorem 3. By Lemma 11 the rich lines of $\mathcal{P}_W$ lie inside the orbits, so a set is lawful iff its intersection with every orbit is lawful, and the maximum and the number of maximum sets are the sum, resp. the product, of the orbit values. By Lemma 10 every generic orbit has pattern $(0,0)$, $(1,1)$, $(0,2)$ or $(2,0)$, and Lemma 11 gives the formula. Since $n_2+n_1+n_0 = \frac{1}{4}(p-1-2s)$, the maximum is $12n_2+10n_1+8n_0+6s \le 12(n_2+n_1+n_0)+6s = 3(p-1)$, with equality iff $n_1=n_0=0$. For the HJSW box, Lemma 7 gives $(e,f)=(0,2)$ for the A/D-orbits and $(2,0)$ for the B/C-orbits, and the exceptional pairs are split; this recovers Theorem 1(ii),(iii). $\square$

Corollary 12 (all conics)**.** *Let C be a non-degenerate conic over $\mathbb{F}_p$ and $c \ge 2$. If the projective closure of C contains both axial points at infinity* (1:0:0) *and* (0:1:0)*, then C is a shifted modular hyperbola $(x+e)(y+d)=c'$, and for $c=2$ every lawful subset of its lift into a $2p \times 2p$ box has at most $3(p-1)$ points (Theorem 3 after the shift). Otherwise every lawful subset of the lift of C into a box of side cp has at most $c(p+3)$ points; for $c=2$ and $p \ge 11$ this is less than $3(p-1)$. Hence for $p \ge 11$ no conic admits more than $3(p-1)$ lawful points in a $2p \times 2p$ box, and only the shifted modular hyperbolae can attain $3(p-1)$ (they do so exactly in the boxes described by Theorem 3, e.g. the HJSW box; a shifted hyperbola in a box where $n_1+n_0>0$ attains less, and a conic that is $\mathbb{F}_p$-affinely equivalent to a hyperbola without being one, e.g. $(x-y)y=1$, is not covered by the shift and satisfies only the bound $c(p+3)$ — for $p=7$ its lift into $G(7)$ has maximum $16<18$).*

Proof. A conic through (1:0:0) and (0:1:0) has no x^2 and no y^2 term, so its equation is $xy+dx+ey+f=0$, i.e. $(x+e)(y+d)=de-f$, a shifted hyperbola (non-degenerate iff $de \ne f$). If, say, $(1{:}0{:}0) \notin C$, then C has an x^2 term, so every horizontal line $y=b$ meets C in at most two points, and the number of residues b with $C \cap \{y=b\} \ne \emptyset$ is at most $\lceil |C(\mathbb{F}_p)|/2 \rceil + 1 \le (p+3)/2$ (only the ≤ 2 points with vertical tangent are alone on their line). In a box of side cp the lift of C therefore meets at most $c(p+3)/2$ rows, and a lawful set has at most two points per row. The symmetric argument applies if $(0{:}1{:}0) \notin C$. $\square$

(A degenerate conic is a union of at most two lines, and the lift of an $\mathbb{F}_p$-line into a box of side cp lies on $O(c\sqrt{p})$ integer lines, so it contains $O(c\sqrt{p})$ lawful points. Note that $\mathbb{F}_p$-affine equivalence is *not* a symmetry of the lifted problem: only integer translations of the box are. The observation that only the shifted hyperbolae contain both axial points at infinity was made independently by reasoning agents consulted by the author of record, who also supplied the $p=7$ example.)

Remark 13 (algorithmic consequence)*.* Selecting a largest subset in general position from an arbitrary point set is NP-hard [17]; for $H(c,p) \cap W$ Theorems 1 and 3 reduce it to bookkeeping: after listing the $p-1$ classes and their V-orbits, the maximum, the number of maximum sets and an explicit maximum set are obtained in $O(p)$ operations (the pattern of an orbit is read off from the split/shared status of its four pairs).

Remark 14*.* The double count of Section 4 gives, for every box, $|S| \le Z + 2|\mathcal{L}| = 2(p-1)+2s+2\sum_O (e_O+f_O) = 12n_2+12n_1+8n_0+6s$ (one checks $Z+2|\mathcal{L}| = 2(p-1)+2s+\#\{m=2\}$ and $\#\{m=2\} = 2\sum_O(e_O+f_O)$ exactly as in Section 4); it is off by exactly 2 on every $(1,1)$-orbit, where two rich three-point lines consist of one point of multiplicity one and two of multiplicity two.

Corollary 15 (no-four-in-line)**.** *In every box W the four-point rich lines are pairwise disjoint; hence the largest subset of $\mathcal{P}_W$ with no four collinear points has exactly $4(p-1)-N_{sh}(W)$ points, where $N_{sh}(W)$ denotes the number of shared pairs (σ-pairs and τ-pairs together; equivalently,* *the number of four-point rich lines of $\mathcal{P}_W$), and there are $4^{N_{sh}(W)}$ such subsets. By the orbit lemma $N_{sh}(W) \geq 2 \cdot \#\{$generic orbits$\} = \frac{1}{2}(p-1) - s$, so the maximum is at most $\frac{7}{2}(p-1)+s$; in the HJSW box $N_{sh} = \frac{1}{2}(p-1) - s$ and the set $S_3 = H(c,p) \cap T_3$ of [2, Constr. 2.10] (with $|S_3| = \frac{7}{2}(p-1)+s$) is a largest no-four-in-line subset of $H(c,p) \cap G(p)$, as asserted in [2, App. A]. (No line carries five points of $\mathcal{P}_W$, so the no-five-in-line maximum is $|\mathcal{P}_W| = 4(p-1)$.)*

Proof. A four-point rich line is D_κ for a shared σ-pair (points $\kappa(0,0), \kappa(1,1), \sigma\kappa(0,0), \sigma\kappa(1,1)$) or E_κ for a shared τ-pair (points with $r+s=1$). The slope-(-1) line through a point $\kappa(r,r)$ is not of the second kind: if it were E_λ then $\kappa \in \{\lambda, \tau\lambda\}$ by Lemma 4, so $E_\lambda = E_\kappa$, which contains only the copies $\kappa(0,1), \kappa(1,0)$ of κ; likewise the slope-$(+1)$ line through a point $\kappa(r, 1-r)$ is not of the first kind. Hence the four-point lines are pairwise disjoint, a no-four-in-line subset must miss at least one point of each of them, and deleting exactly one point from each is sufficient (all other lines carry at most three points). For the HJSW box, $N_{sh} = \frac{1}{2}(p-1) - s$ by Lemma 7, and $|S_3| = |S_2| + |M \cap E \cap H(c,p)| = 3(p-1) + \frac{1}{2}(p-1) + s$ because the classes fixed by σ or τ lie in E and the remaining classes split evenly between E and F [2, Obs. 2.9]. □

For the box $[0,2p)^2$ all pairs are shared ($n_2 = n_1 = 0$, $s_2 = s$): the maximum is $8n_0 + 6s = 2(p-1) + 2s$ with $1296^{n_0}6^s$ maximum sets. Table 1 shows the exact maxima for all boxes at $p = 11$, $c = 1$ (two generic orbits, $s = 1$): the values $22, 24, 26, 28, 30$ correspond to $(n_1, n_0) = (0,2), (1,1), (2,0)$ or $(0,1)$, $(1,0)$, $(0,0)$. The number of boxes attaining $3(p-1)$ depends on c (for $p = 13$: 28 of the 169 boxes for $c = 1$, 4 for $c = 2$); it always includes the HJSW box $x_0 = -h$, $y_0 \in \{0,1\}$.

Remark 16 (larger boxes). In boxes larger than $2p \times 2p$ the classes have more copies; the maximum is then at least $3(p-1)$ (equal to it for $N = 2p, 2p+1$ in the cases computed and larger from $N = 2p+2$ on), and the ratio maximum$/N$ does not exceed the HJSW value $3(p-1)/2p$ in the range computed (it is not monotone in N; for $p = 7$ the $3p \times 3p$ box ties, $27/21 = 18/14$). Exact values (MIP with zero gap, $c = 1$, box $[-h, -h+N) \times [0, N)$): for $p = 11$ and $N = 22, 23, \ldots, 33$ the maxima are 30, 30, 31, 31, 33, 36, 36, 38, 39, 40, 42, 43; for $p = 13$, $N = 26, \ldots, 39$: 36, 36, 38, 38, 38, 38, 40, 40, 43, 45, 47, 49, 51, 52; for $p = 17$, $N = 34, \ldots, 51$: 48, 48, 49, 49, 50, 51, 52, 53, 56, 56, 58, 61, 62, 63, 65, 66, 67, 69. In the $3p \times 3p$ box the maximum is $27, 43, 52, 69, 77, 93$ for $p = 7, 11, 13, 17, 19, 23$: about $4.3(p-1)$, i.e. ratio 1.29–1.35 to the side $3p$, against $3(p-1)/2p = 1.29, 1.36, 1.38, 1.41, 1.42, 1.43$ for the HJSW box at the same primes (the value depends on c: $51, 52, 54$ for $p = 13$, $c = 2, 1, 3$). In the $2p \times 4p$ box the trivial bound $4(p-1)$ (two of the four collinear copies in each column) is attained for $p = 7, 11, 13$. Data: `slack/verification/windows_2p_to_3p.txt`.

Stability: near-maximum sets are near a maximum set

The classification of Theorem 3 is in fact orbit-wise, and this makes the extremal problem for one hyperbola *stable*: a lawful set of size $3(p-1) - t$ differs from a maximum one in at most t points.

Lemma 17 (orbit decomposition)**.** *Let $O = \{\kappa, \sigma\kappa, \tau\kappa, \nu\kappa\}$ be a V-orbit of classes and $P_O \subseteq H(c,p) \cap G(p)$ the union of their lifts (16 points for a generic orbit, 8 for an exceptional one). Every line containing at least three points of $H(c,p) \cap G(p)$ has all of them in a single P_O; consequently $\alpha(H(c,p) \cap G(p)) = \sum_O \alpha(P_O)$, and a set is lawful iff its restriction to every orbit is. Moreover the incidence structure of P_O — which points lie on a common rich line — is the same for all generic orbits and all p, with rich-line profile $(4,4,3,3,3,3)$, and likewise for all exceptional orbits, with profile $(3,3)$.*

Proof. The first assertion is part of Proposition 8. For the last one, a generic orbit consists of two classes of type A and two of type D, or of two of type B and two of type C (Lemma 7: σ preserves A, D and exchanges B, C, τ the opposite), and Proposition 8 then lists its rich lines: in the first case one four-point diagonal for the two A's and one for the two D's (item (a)) and one

$x_0 + h \backslash y_0$	0	1	2	3	4	5	6	7	8	9	10
0	30	30	30	28	26	24	22	24	26	28	30
1	28	28	28	30	28	26	24	26	28	30	28
2	26	26	26	28	30	28	26	28	30	28	26
3	24	24	24	26	28	30	28	30	28	26	24
4	22	22	22	24	26	28	30	28	26	24	22
5	22	22	22	24	26	28	30	28	26	24	22
6	22	22	22	24	26	28	30	28	26	24	22
7	22	22	22	24	26	28	30	28	26	24	22
8	24	24	24	26	28	30	28	30	28	26	24
9	26	26	26	28	30	28	26	28	30	28	26
10	28	28	28	30	28	26	24	26	28	30	28

Table 1: $p = 11$, $c = 1$: exact maximum lawful subset of $H(1, 11)$ in the box $[x_0, x_0 + 22) \times [y_0, y_0 + 22)$, indexed by $x_0 + h \bmod p$ (rows; 0 is the HJSW position $x_0 = -h$) and $y_0 \bmod p$ (columns). Values (HiGHS MIP, witnesses through the law; they agree with the formula of Theorem 3) range from $2(p-1) + 2s = 22$ to $3(p-1) = 30$; the double-count bound $Z + 2|\mathcal{L}|$ equals 30 for 76 boxes but is attained for 20 (the others have a $(1, 1)$-orbit).

three-point antidiagonal for each of the four classes (item (d)); in the second case the same with diagonals and antidiagonals, and (b),(c) in place of (a),(d). In both cases the incidence structure is the stated one, and the two cases are exchanged by $(x, y) \mapsto (y, x)$. An exceptional orbit is $\{\kappa, \nu\kappa\}$ with κ fixed by σ or by τ, and Proposition 8 gives two three-point lines. $\square$

Lemma 18 (a single orbit)**.** *For a generic orbit $\alpha(P_O) = 12$ and the maximum set is* unique*; for an exceptional orbit $\alpha(P_O) = 6$ and there are exactly nine maximum sets. Moreover every lawful $S \subseteq P_O$ with $|S| = \alpha(P_O) - d$ satisfies $\min_M |S \setminus M| \leq d$, the minimum being over the maximum sets of P_O (for an exceptional orbit $\min_M |S \setminus M| = 0$: every lawful subset of size $6 - d$ is contained in a maximum set).*

Proof. By Lemma 17 there are only two incidence structures, so this is a finite verification, carried out by exhaustive enumeration of all subsets of the 16, resp. 8 points (`slack/t221/orbit_stability.py`, `slack/t221/orbit_types.py`; the values of $\alpha(P_O)$, the number of maximum sets and the bound $\min_M |S \setminus M| \leq d$ for $d \leq 3$ were confirmed for every orbit of every prime $p \leq 31$, and exactly two incidence types occur; for $d \geq 4$ the bound is trivial, since $|P_O \setminus M| = 4$, resp. 2). The value 12 and the count of maximum sets also follow from the gadget analysis of Theorem 3. $\square$

Proposition 19 (stability)**.** *Let $S \subseteq H(c, p) \cap G(p)$ be lawful with $|S| = 3(p-1) - t$. Then there is a maximum lawful set M, $|M| = 3(p-1)$, with $|S \setminus M| \leq t$; in particular $|S \cap M| \geq 3(p-1) - 2t$.*

Proof. By Lemma 17 the deficits $d_O := \alpha(P_O) - |S \cap P_O|$ are non-negative and sum to t. Choose in each orbit a maximum set M_O realising the minimum of Lemma 18; the union $M = \bigcup_O M_O$ is lawful of size $\sum_O \alpha(P_O) = 3(p-1)$ (Lemma 17) and $|S \setminus M| = \sum_O |S \cap P_O \setminus M_O| \leq \sum_O d_O = t$. $\square$

Remark 20. (a) The bound is sharp for every t in the near-maximum range: an exact computation (CP-SAT, $p = 11, 13, 17, 19, 23, 29, 31$ and $1 \leq t \leq 4$, every value proved optimal; `slack/t221/stability_h3.py`) gives

$$\max\{\min_M |S \setminus M| \ : \ S \text{ lawful}, \ |S| = 3(p-1) - t\} = \min(t, \, p - 1 - 2s),$$

the plateau value $p - 1 - 2s$ being the number of generic classes: one unit of deficit buys exactly one deviating point until every generic class has been touched. (b) The proposition is the "stability" input that a proof of $\alpha(\mathcal{P}_k) \leq 3(p-1) + O(1)$ would want: it says that the part of a lawful set on one hyperbola is, up to its own deficit, a piece of a HJSW-type configuration, so that the interaction with the second hyperbola can be analysed against a fixed structure rather than against an arbitrary set. What it does *not* give is the exchange rate: a point of the second hyperbola must be paid for by a deficit *relative to the optimum* of the first, and that is the whole conjecture. (c) The same decomposition explains the count 9^s of Theorem 3: generic orbits are rigid, exceptional ones carry nine patterns each.

6 Two hyperbolae: a first bound for $H(1) \cup H(-1)$

Theorems 1 and 3 concern one hyperbola. For a union of two hyperbolae in the same box the exact maxima are known only computationally (Section 11); the trivial bound is $4(p-1)$, because every column and every row of the box that is not $\equiv 0$ contains four candidate points (two lifts from each hyperbola) and a lawful set takes at most two of them. Here we prove, for $k = -1$, a bound below the trivial one by an explicit weighted line cover. Throughout this section

$$\mathcal{P}_{-1} := (H(1,p) \cup H(-1,p)) \cap G(p), \qquad |\mathcal{P}_{-1}| = 8(p-1),$$

and, for a slope-(± 1) integer line ℓ, $|\ell|$ denotes the number of points of $\mathcal{P}_{-1}$ on ℓ. Let $m_8(p)$ be the number of slope-(+1) lines $x - y = d'$ with $|\ell| = 8$.

Theorem 21. *For every odd prime p, every lawful subset of $\mathcal{P}_{-1}$ has at most $4(p-1) - 4\,m_8(p)$ points.*

The proof rests on the two reflections of the box. Put $R(x, y) = (p - x,\, y)$ and $R'(x, y) = (x,\, 2p-y)$. Both map $G(p) \cap \{xy \not\equiv 0\}$ onto itself (the x-range $[-h, 3h+1]$ is symmetric about $p/2$, the y-range $[1, 2p-1]$ about p; the two omitted lines $y = 0$ and $x \in \{0, p\}$ carry no candidates), both map $H(1)$ onto $H(-1)$ and vice versa (since $(p-x)y \equiv -xy$ and $x(2p-y) \equiv -xy$), and both interchange slope $+1$ and slope -1 lines: $R\{x - y = d'\} = \{x + y = p - d'\}$, $R'\{x - y = d'\} = \{x + y = 2p + d'\}$.

Lemma 22. *Every column $x \not\equiv 0$ of the box contains exactly four points of $\mathcal{P}_{-1}$: the two lifts of the $H(1)$-class $\kappa = (x, 1/x)$ and the two lifts of the $H(-1)$-class $R'(\kappa) = (x, -1/x)$. Every row $y \not\equiv 0$ contains exactly four points: the two lifts of $\kappa = (1/y, y) \in H(1)$ and the two lifts of $R(\kappa) = (-1/y, y) \in H(-1)$.*

Proof. On the column x the candidates are the $y \in [0, 2p)$ with $xy \equiv \pm 1$: two lifts of $y \equiv 1/x$ and two of $y \equiv -1/x$. Rows likewise. ☐

Lemma 23 (residue groups)**.** *Fix a residue $d \in \mathbb{F}_p$. The points of $\mathcal{P}_{-1}$ on the slope-(+1) lines $x - y \equiv d$* (mod p) *are: the four lifts of each $H(1)$-class κ with $x_\kappa - y_\kappa \equiv d$ (at most two classes, κ and $\sigma\kappa$), lying on the three lines $x - y = c_\kappa - p$, c_κ, $c_\kappa + p$ with $1, 2, 1$ points, where $c_\kappa = x_\kappa - y_\kappa$ is computed from the base copy; and the images under R of the four lifts of each $H(1)$-class λ with $x_\lambda + y_\lambda \equiv -d$ (at most two classes, λ and $\tau\lambda$), lying on the lines $x - y = c'_\lambda - p$, c'_λ, $c'_\lambda + p$ with $1, 2, 1$ points, where $c'_\lambda = -(x_\lambda + y_\lambda)$. All these centres lie in the interval $[-h-p+1,\, h-1]$ of length $2p-2$, so a residue d involves at most two distinct centres, c and $c+p$. In particular $|\ell| \le 8$ for every slope-(± 1) line, and $|\ell| = 8$ for $\ell = \{x - y = c\}$ iff four distinct classes ($\kappa, \sigma\kappa, \lambda, \tau\lambda$; none fixed by σ resp. τ) have the common centre c; then the sixteen points of these four classes lie on the three lines $x - y = c - p, c, c + p$ with $4, 8, 4$ points, and no other point of $\mathcal{P}_{-1}$ lies on these three lines. We call this a* good residue group *of slope $+1$; good residue groups of slope -1 are defined symmetrically (they are the images under R of the good groups of slope $+1$, so there are $m_8(p)$ of each slope).*

Proof. The copies of κ are $(x_\kappa + rp, y_\kappa + sp)$, on $x - y = c_\kappa + (r - s)p$: one point each for $(r, s) = (1, 0), (0, 1)$ and two for $(0, 0), (1, 1)$. A point of $H(-1)$ on $x - y = d'$ is $R(q)$ for a point q of $H(1)$ on $x + y = p - d'$; the copies of λ lie on $x + y = x_\lambda + y_\lambda + (r + s)p$, whose R-images lie on $x-y=-(x_\lambda+y_\lambda)-(r+s)p+p$, i.e. centre c'_λ with pattern $1,2,1$ (with $r+s=1$ on the centre line). Since $x_\kappa\in[-h,h]$, $y_\kappa\in[1,p-1]$, both centres lie in $[-h-p+1,h-1]$. A class fixed by σ (resp. τ) is its own partner and contributes only its own two points to the centre line, so an 8-point line needs four distinct classes with a common centre; the count of points on $c\pm p$ is then $1+1+1+1$. Only the classes with $x-y\equiv d$ (resp. $x+y\equiv -d$) have points on lines of residue d, so no other points occur. □

Lemma 24. *Let $\mathcal{G}$ be the set of all good residue groups of both slopes and let $B\subseteq\mathcal{P}_{-1}$ be the union of their points. Then (i) two distinct good groups have no point in common; (ii) $R(B)=R'(B)=B$; (iii) every row and every column of the box is either contained in B or disjoint from B.*

Proof. (i) Two good groups of the same slope have different residues, hence disjoint point sets. A good group of slope $+1$ contains, from $H(1)$, a σ-pair $\{\kappa,\sigma\kappa\}$ with $c_\kappa=c_{\sigma\kappa}$ (a *shared* pair, i.e. x_κ and $x_{\sigma\kappa}=-y_\kappa \bmod p$ have opposite signs, types A/D of Lemma 7), and from $H(-1)$ the R-images of a shared τ-pair (types B/C). A good group of slope -1 is the R-image of one of slope $+1$, so it contains from $H(1)$ a shared τ-pair and from $H(-1)$ the R-images of a shared σ-pair. Since a class cannot be both σ-shared and τ-shared (Lemma 7: the former means $\kappa\in A\cup D$, the latter $\kappa\in B\cup C$), an $H(1)$-class lies in at most one good group, and the same holds for $H(-1)$-classes by applying R. Hence the point sets are pairwise disjoint. (ii) R and R' map $\mathcal{P}_{-1}$ onto itself and slope-$(+1)$ lines onto slope-(-1) lines, so they permute the good groups. (iii) By Lemma 22 a row consists of the lifts of κ and of $R(\kappa)$; the four lifts of a class are in B or all outside B (good groups consist of whole classes), and $\kappa\subseteq B$ iff $R(\kappa)\subseteq B$ by (ii). Columns: κ and $R'(\kappa)$. □

Proof of Theorem 21. Let $S\subseteq\mathcal{P}_{-1}$ be lawful, so $|S\cap\ell|\le 2$ for every line ℓ. If a family $\mathcal{C}$ of lines covers $\mathcal{P}_{-1}$, then $|S|\le\sum_{\ell\in\mathcal{C}}|S\cap\ell|\le 2|\mathcal{C}|$. Take $\mathcal{C}=$ the three lines of every good residue group (both slopes: $3\cdot 2m_8$ lines) together with every row that is disjoint from B. A point of B lies on a line of its group (Lemma 23); a point outside B lies on a row disjoint from B (Lemma 24(iii)); so $\mathcal{C}$ covers. The rows disjoint from B contain exactly the $8(p-1)-32m_8$ points outside B, four each (Lemma 22), so there are $2(p-1)-8m_8$ of them, and $|S|\le 2\,(6m_8+2(p-1)-8m_8)=4(p-1)-4m_8$. □

(The same count with weight $\frac{1}{2}$ on the rows and on the columns disjoint from B gives the same bound as a fractional cover. Theorem 21 was found independently, on the same day, by two of the agents mentioned in the disclosure; the point-by-point verification of the cover for all $19\le p<320$ is `slack/block_cover_km1.py`, and for $p\le 101$ `slack/km1_theorem_check.py`.)

Proposition 25 (the number of eight-point lines)**.** *Let $X(u)\in[-h,h]$ denote the least absolute residue of $u\in\mathbb{F}_p^*$. Then $4\,m_8(p)$ is the number of pairs $(a,b)\in(\mathbb{F}_p^*)^2$ with*

$$ab^2+(a^2-1)b+a\equiv 0 \pmod p,\qquad a^2\not\equiv -1,\quad b^2\not\equiv 1,\qquad X(a)X(1/a)<0,\quad X(b)X(1/b)>0,$$

$$X(a)-X(1/a)+X(b)+X(1/b)\;=\;p\,([X(a)>0]-[X(b)<0]).$$

Proof. By Lemma 23 an eight-point line $x-y=c$ corresponds to a residue d with an $H(1)$-class $\kappa=(a,1/a)$, $a-1/a\equiv d$, whose σ-pair is shared and non-degenerate ($\sigma\kappa\ne\kappa$, i.e. $a^2\not\equiv -1$), and an $H(1)$-class $\lambda=(b,1/b)$, $b+1/b\equiv -d$, whose τ-pair is shared and non-degenerate ($b^2\not\equiv 1$), with $c_\kappa=c'_\lambda$. Eliminating d from $a-1/a\equiv-(b+1/b)$ gives the cubic. Writing the base copies with $Y(u)=X(u)+p[X(u)<0]$: $c_\kappa=X(a)-Y(1/a)$ and $c_{\sigma\kappa}=X(-1/a)-Y(-a)=X(a)-X(1/a)-p[X(a)>0]$, so the σ-pair is shared iff $X(a)$ and $X(1/a)$ have opposite signs, and then $c_\kappa=X(a)-X(1/a)-p[X(a)>0]$; likewise $e_\lambda=X(b)+Y(1/b)$, $e_{\tau\lambda}=X(1/b)+Y(b)$, the τ-pair is shared iff $X(b)$, $X(1/b)$ have the same sign, and then $c'_\lambda=-e_\lambda=-X(b)-X(1/b)-p[X(b)<0]$. The equality $c_\kappa=c'_\lambda$ is the displayed equation. Each line arises from exactly four pairs ($a\in\{a,-1/a\}$, $b\in\{b,1/b\}$). □

Since $a-1/a+b+1/b\equiv 0$ on the cubic, the integer $L=X(a)-X(1/a)+X(b)+X(1/b)\in[-4h,4h]$ is a multiple of p, i.e. $L\in\{-p,0,p\}$; the last condition selects one of these three values.

Proposition 26. *$m_8(p)=\frac{1}{12}p+O(\sqrt{p}\,\log^4 p)$. Consequently $\alpha(\mathcal{P}_{-1})\le\frac{11}{3}(p-1)+O(\sqrt{p}\,\log^4 p)$.*

Proof. Let $C_0=\{(a,b)\in\mathbb{F}_p^2:\ ab^2+(a^2-1)b+a=0\}$. Its only point with $a=0$ or $b=0$ is $(0,0)$.

Step 1: C_0 is absolutely irreducible and $|C_0(\mathbb{F}_p)|=p+O(\sqrt{p})$. As a polynomial in b over $K=\overline{\mathbb{F}_p}(a)$, $F=ab^2+(a^2-1)b+a$ has discriminant $\Delta=(a^2-1)^2-4a^2=a^4-6a^2+1=(a^2-3)^2-8$, whose four roots $a^2=3\pm2\sqrt{2}$ are distinct (their product is $1\ne0$ and $3\pm2\sqrt{2}$ are distinct, $p\ne2$), so Δ is not a square in K and F is irreducible over K; being primitive in b (its coefficients a, a^2-1, a have no common factor), F is irreducible in $\overline{\mathbb{F}_p}[a,b]$ (Gauss). Hence C_0 is an absolutely irreducible plane cubic and, by the Weil bound for (possibly singular) absolutely irreducible curves, $|C_0(\mathbb{F}_p)|=p+O(\sqrt{p})$ ([4, Ch. 11]; for singular curves [5]).

Step 2: the conditions of Proposition 25 are interval conditions on four linear forms. On C_0 we have $1/b=-(a-1/a+b)$, so with $x_1=X(a)$, $x_2=X(1/a)$, $x_3=X(b)$ and $T=x_1-x_2+x_3\in(-3p/2,3p/2)$ the fourth least residue is $X(1/b)=-T+p\varepsilon$ with $\varepsilon\in\{-1,0,1\}$ chosen so that $|-T+p\varepsilon|\le h$, and $L=T+X(1/b)=p\varepsilon$. The conditions $X(a)X(1/a)<0$, $X(b)X(1/b)>0$, $L=p([x_1>0]-[x_3<0])$ therefore hold iff (x_1,x_2,x_3) lies in one of the four sets

$$W_1=\{x_1>0>x_2,\ x_3>0,\ p/2<T<p\},\qquad W_2=\{x_1>0>x_2,\ x_3<0,\ 0<T<p/2\},$$
$$W_3=\{x_1<0<x_2,\ x_3>0,\ -p/2<T<0\},\quad W_4=\{x_1<0<x_2,\ x_3<0,\ -p<T<-p/2\}.$$

Indeed, for $x_1>0>x_2$, $x_3>0$ we have $T\in(0,3p/2)$; for $T<p/2$ we get $\varepsilon=0$ and $X(1/b)=-T<0$ (excluded, as $X(b)>0$), for $p/2<T<p$ we get $\varepsilon=1$, $X(1/b)=p-T>0$ and $L=p$, which is the required value $p([x_1>0]-[x_3<0])$, and for $T>p$ again $X(1/b)=p-T<0$; this gives W_1, and the other three cases are analogous. (So the condition on L is implied by the two sign conditions; it was kept in Proposition 25 because it is what the cover uses.) Now pass to the torus $\mathbb{T}^3=(\mathbb{R}/\mathbb{Z})^3$ and to the points $\mathbf{x}=(a,a^{-1},b)/p \bmod 1$, one for each $(a,b)\in C_0(\mathbb{F}_p)$ with $a\ne0$; there are $N=|C_0(\mathbb{F}_p)|-1=p+O(\sqrt{p})$ of them. With the four linear forms

$$\ell_1(\mathbf{x})=x_1,\qquad \ell_2(\mathbf{x})=x_2,\qquad \ell_3(\mathbf{x})=x_3,\qquad \ell_4(\mathbf{x})=x_1-x_2+x_3\pmod 1$$

we have $X(u)>0\iff u/p \bmod 1\in(0,\frac12)$ and $X(u)<0\iff u/p \bmod 1\in(\frac12,1)$, and inside each octant the condition on T is a condition on $T/p \bmod 1=\ell_4(\mathbf{x})$: for W_1 ($T\in(0,3p/2)$) the condition $p/2<T<p$ is $\ell_4\in(\frac12,1)$, for W_2 ($T\in(-p/2,p)$) the condition $0<T<p/2$ is $\ell_4\in(0,\frac12)$, for W_3 ($T\in(-p,p/2)$) it is $\ell_4\in(\frac12,1)$ and for W_4 ($T\in(-3p/2,0)$) it is $\ell_4\in(0,\frac12)$. Hence each W_i is $\{\mathbf{x}:\ \ell_j(\mathbf{x})\in I_{ij},\ j=1,\dots,4\}$ with $I_{ij}\in\{(0,\frac12),(\frac12,1)\}$; the endpoints are never attained ($a,b\ne0$, p odd, T an integer), so the conditions are exact. Writing $x_i=\pm\frac{p}{2}U_i$ with U_i uniform on $(0,1)$, the volume of W_i is $\frac18\Pr[1<U_1+U_2+U_3<2]=\frac18\cdot\frac23$ for W_1,W_4 and $\frac18\Pr[0<U_1+U_2-U_3<1]=\frac18\cdot\frac23$ for W_2,W_3 (the density of $U_1+U_2+U_3$ has mass $\frac16$ on $[0,1]$ and on $[2,3]$; that of $U_1+U_2-U_3$ has mass $\frac16$ on $[-1,0]$ and on $[1,2]$), in total $\sum_i\lambda(W_i)=\frac13$.

Step 3: exponential sums. For $\mathbf{h}=(h_1,h_2,h_3)\in\mathbb{Z}^3$ with $\mathbf{h}\not\equiv0\pmod p$ the rational function $f_{\mathbf{h}}=h_1a+h_2a^{-1}+h_3b$ on C_0 is non-constant: if $h_3\not\equiv0$ a constant value would give $b\in K$, contradicting the irreducibility of F over K; if $h_3\equiv0$ then $h_1a+h_2/a=c$ for infinitely many a forces $h_1\equiv h_2\equiv0$. Its degree is bounded independently of $\mathbf{h}$ and p, so it is not of the form g^p-g+c for p large, and Bombieri's estimate for exponential sums along an absolutely irreducible curve [3] (see also [4, Ch. 11]) gives

$$S(\mathbf{h}):=\sum_{\mathbf{x}}e(\mathbf{h}\cdot\mathbf{x})=\sum_{(a,b)\in C_0(\mathbb{F}_p),\,a\ne0}e\Big(\frac{h_1a+h_2a^{-1}+h_3b}{p}\Big)=O(\sqrt{p})\qquad(\mathbf{h}\not\equiv0 \bmod p)$$

with an absolute implied constant, while $S(\mathbf{0})=N$.

Step 4: counting with Selberg polynomials. Fix $H = \lfloor\sqrt{p}\rfloor$ (so $4H < p$ for $p \geq 17$; the statement is asymptotic). For an arc $I \subset \mathbb{R}/\mathbb{Z}$ there are trigonometric polynomials $S_I^- \leq \mathbf{1}_I \leq S_I^+$ of degree at most H (Selberg's polynomials, [6, Ch. 1]; [7]) with $\widehat{S_I^\pm}(0) = |I| \pm \frac{1}{H+1}$, $|\widehat{S_I^\pm}(m)| \leq \frac{1}{H+1} + \min(|I|, \frac{1}{\pi|m|})$ for $0 < |m| \leq H$, and $0 \leq S_I^+ \leq C_0$ for an absolute constant C_0. Fix i and write $\mathbf{1}_j = \mathbf{1}_{I_{ij}}$, $S_j^\pm = S_{I_{ij}}^\pm$, and $\mathbf{n}_j \in \mathbb{Z}^3$ for the coefficient vector of ℓ_j (so $\ell_j(\mathbf{x}) = \mathbf{n}_j \cdot \mathbf{x}$, $\mathbf{n}_4 = (1, -1, 1)$). The number of our points in W_i is $\sum_{\mathbf{x}} \prod_{j=1}^4 \mathbf{1}_j(\ell_j \mathbf{x})$. By the telescoping identity $\prod_j A_j - \prod_j B_j = \sum_j (\prod_{k<j} B_k)(A_j - B_j)(\prod_{k>j} A_k)$ and $0 \leq \mathbf{1}_k \leq 1$, $0 \leq S_k^+ \leq C_0$,

$$\Big|\prod_j \mathbf{1}_j(\ell_j\mathbf{x}) - \prod_j S_j^+(\ell_j\mathbf{x})\Big| \;\leq\; C_0^3 \sum_{j=1}^4 (S_j^+ - \mathbf{1}_j)(\ell_j\mathbf{x}) \;\leq\; C_0^3 \sum_{j=1}^4 (S_j^+ - S_j^-)(\ell_j\mathbf{x}),$$

and for each j, expanding $S_j^+ - S_j^-$ in its Fourier series, $\sum_{\mathbf{x}} (S_j^+ - S_j^-)(\ell_j\mathbf{x}) = \sum_{|m|\leq H} (\widehat{S_j^+}(m) - \widehat{S_j^-}(m)) S(m\mathbf{n}_j) = \frac{2N}{H+1} + O\Big(\sqrt{p} \sum_{0<|m|\leq H} \big(\frac{1}{H+1} + \frac{1}{|m|}\big)\Big) = O(\sqrt{p}\,\log p)$, because $m\mathbf{n}_j \not\equiv 0 \pmod p$ for $0 < |m| \leq H$. It remains to evaluate the smoothed count: with $\mathbf{h}(\mathbf{m}) = \sum_j m_j \mathbf{n}_j$,

$$\sum_{\mathbf{x}} \prod_{j=1}^4 S_j^+(\ell_j\mathbf{x}) = \sum_{\mathbf{m}\in\mathbb{Z}^4,\ |m_j|\leq H} \ \prod_{j=1}^4 \widehat{S_j^+}(m_j)\ S(\mathbf{h}(\mathbf{m})).$$

The terms with $\mathbf{h}(\mathbf{m}) = \mathbf{0}$ contribute $N \int_{\mathbb{T}^3} \prod_j S_j^+(\ell_j\mathbf{x})\, d\mathbf{x}$, and since each $\ell_j : \mathbb{T}^3 \to \mathbb{R}/\mathbb{Z}$ is measure preserving ($\mathbf{n}_j$ is primitive), the same telescoping gives $|\int \prod_j S_j^+(\ell_j\mathbf{x})\, d\mathbf{x} - \lambda(W_i)| \leq C_0^3 \sum_j \int_0^1 (S_j^+ - \mathbf{1}_j) = \frac{4C_0^3}{H+1}$. The terms with $\mathbf{h}(\mathbf{m}) \neq \mathbf{0}$ have $0 < \|\mathbf{h}(\mathbf{m})\|_\infty \leq 4H < p$, hence $|S(\mathbf{h}(\mathbf{m}))| = O(\sqrt{p})$, and their total is $O(\sqrt{p}\,(\sum_{|m|\leq H} |\widehat{S^+}(m)|)^4) = O(\sqrt{p}\,\log^4 p)$. Altogether the number of our points in W_i is $N\lambda(W_i) + O(N/H + \sqrt{p}\log^4 p) = N\lambda(W_i) + O(\sqrt{p}\,\log^4 p)$. Summing over i and using $N = p + O(\sqrt{p})$, $\sum_i \lambda(W_i) = \frac{1}{3}$: the number of pairs (a, b) satisfying the conditions of Proposition 25, up to the $O(1)$ excluded pairs with $a^2 \equiv -1$ or $b^2 \equiv 1$, is $\frac{1}{3}p + O(\sqrt{p}\log^4 p)$, so $m_8(p) = \frac{1}{12}p + O(\sqrt{p}\,\log^4 p)$. The last statement follows from Theorem 21. □

The data agree: $m_8(p)/p$ has mean 0.0825 over the 193 primes in $[200, 1500]$ (range 0.058–0.108; $1/12 = 0.0833$).

Remark 27. (a) The bound is attained at $p = 11$: $m_8(11) = 2$ and $\alpha(\mathcal{P}_{-1}) = 32 = 4 \cdot 10 - 8$ (exact maximum, Section 11); $m_8(13) = m_8(17) = 0$ (no gain), and $m_8(p) \geq 2$ for every prime $19 \leq p \leq 1500$ (computed), so the bound is below $4(p-1)$ from $p = 19$ on: for $k = -1$ the trivial bound $4(p-1)$ is never attained once an eight-point line exists. In terms of the total number $M_8 = 2m_8$ of eight-point lines of both slopes the bound reads $4(p-1) - 2M_8$. For $p = 19, 23, 29, 37, 47$ it coincides with the best rank-1 (LP) certificate using rows, columns and slope-(±1) lines, which is $64, 80, 96, 128, 160$; for other p that LP is somewhat smaller (it also exploits 5–7-point lines with fractional weights). (b) Proposition 26 makes the bound asymptotically explicit: $\alpha(\mathcal{P}_{-1}) \leq (11/3 + o(1))(p-1)$, against the trivial $4(p-1)$ and the computed maxima $3(p-1) + O(1)$ for $p \leq 23$. The proposition quotes standard estimates (Bombieri's bound, Selberg's polynomials) with unspecified absolute constants; the theorem itself is elementary and unconditional. (Versions 1.0–1.1 of this note had the error term $O(p^{5/6}\log^3 p)$, obtained by covering the polytopes W_i with boxes; the present $O(\sqrt{p}\log^4 p)$ uses that the W_i are cut out by four linear forms, a simplification pointed out by an external referee-style report.) (c) The argument uses that R and R' are symmetries of $\mathcal{P}_{-1}$; for a general second hyperbola $H(k)$ the rows pair a class of $H(1)$ with a class of $H(k)$ that is not the image of the first under a symmetry. Section 7 replaces the block cover by a "potential" cover along the cycles of the row/column incidence graph and obtains an exact bound for every k; the case $k = -1$ is the one where all cycles have length four, and Theorem 21 is recovered.

The seven-point groups: a second block

Theorem 21 uses only the eight-point lines. The residue groups with a seven-point line give a second, smaller block, and the bookkeeping of *which* groups can be used together is exact. For a residue $d \not\equiv 0$ let G_d be the set of points of $\mathcal{P}_{-1}$ on the slope-$(+1)$ lines $x - y \equiv d \pmod p$ and G'_d the set of points on the slope-(-1) lines $x + y \equiv d$; thus $R\,G_d = G'_{-d}$, $R'G_d = G'_d$, $RR'G_d = G_{-d}$. Call d *full* if G_d contains four classes: an $H(1)$-pair $\{\kappa, \sigma\kappa\}$ with $d_\kappa \equiv d$ and the R-images of an $H(1)$-pair $\{\lambda, \tau\lambda\}$ with $e_\lambda \equiv -d$ (Lemma 23; the first pair exists iff $d^2 + 4$ is a nonzero square, the second iff $d^2 - 4$ is). Then $|G_d| = 16$, the four centres take at most the two values $c, c + p$ (Lemma 23), and the profile of line sizes of G_d is

$(4, 8, 4)$ if both pairs are shared, $\quad (3, 7, 5, 1)$ or $(1, 5, 7, 3)$ if exactly one is, $\quad (2, 6, 6, 2)$ if none is,

where the $H(1)$-pair is shared iff $\kappa \in A \cup D$ and the R-image of the τ-pair is shared iff $\lambda \in B \cup C$ (Lemma 7); the first case is the good group of Lemma 23. We call d a *seven-group* in the second case and write $m_7(p)$ for the number of seven-groups of slope $+1$ (equivalently, of seven-point lines of slope $+1$).

Lemma 28 (Klein orbits)**.** *Let $d \not\equiv 0$ and $\Omega_d := G_d \cup G_{-d} \cup G'_d \cup G'_{-d}$ $(= \Omega_{-d})$. (i) The slope-(± 1) lines carrying the points of the four groups are pairwise disjoint as subsets of $\mathcal{P}_{-1}$; hence Ω_d is their disjoint union and $|\Omega_d| = 4|G_d|$. (ii) Every row and every column of the box is contained in Ω_d or disjoint from it. (iii) If d is full and $S \subseteq \mathcal{P}_{-1}$ is lawful, then $|S \cap \Omega_d| \leq 4\sum_i \min(n_i, 2)$, where (n_i) is the profile of G_d: at most 24 for a good group and at most 28 for a seven-group.*

Proof. (i) Two lines of the same slope among the four groups are distinct (the residues d and $-d$ differ, as $d \not\equiv 0$), hence disjoint. Let $x - y = c_1$ be a line of G_d and $x + y = c_2$ a line of $G'_{-d} = R\,G_d$, so $c_2 = p - c_3$ with $c_3 \equiv d$; a common lattice point has $2x = c_1 + c_2 = p - (c_3 - c_1)$ with $c_3 - c_1 \in p\mathbb{Z}$, so either $c_1 + c_2$ is odd and there is none, or $x \equiv 0 \pmod p$ and the point is not a candidate. For a line $x + y = c_2$ of $G'_d = R'G_d$ we have $c_2 = 2p + c_3$, $c_3 \equiv d$, and a common point has $2x = c_1 + c_3 + 2p$, so $x \equiv c_1$ and $y = x - c_1 \equiv 0$: again not a candidate. The remaining pairs (lines of $G_{-d} = RR'G_d$ against those of $G'_{\pm d}$) are the images of the two cases under R or R'. (ii) The four points of a class have the same residues $x - y$ and $x + y$ modulo p, so Ω_d is a union of whole classes, and R, R' permute the four groups, so $\kappa \subseteq \Omega_d$ iff $R\kappa \subseteq \Omega_d$ iff $R'\kappa \subseteq \Omega_d$. A row is $\kappa \cup R\kappa$ and a column $\kappa \cup R'\kappa$ (Lemma 22). (iii) The four groups have the same profile (R, R' preserve line sizes), S has at most $\min(n_i, 2)$ points on a line with n_i candidates, and by (i) every point of Ω_d lies on exactly one of the sixteen lines. □

Lemma 29 (neighbours)**.** *Let $d, e \not\equiv 0$ with $e \not\equiv \pm d$. Then $\Omega_d \cap \Omega_e \neq \emptyset$ iff $e^2 \equiv d^2 \pm 4 \pmod p$. If $e^2 \equiv d^2 + 4$, the intersection consists of the $H(1)$-classes of $G_d \cup G_{-d}$ and of the $H(-1)$-classes of $G'_d \cup G'_{-d}$ (eight classes, 32 points, when d is full); if $e^2 \equiv d^2 - 4$, of the $H(-1)$-classes of $G_{\pm d}$ and the $H(1)$-classes of $G'_{\pm d}$. Two good groups d, e never have $\Omega_d \cap \Omega_e \neq \emptyset$.*

Proof. For a candidate $q = (x, y)$ put $r_+(q) = x - y$, $r_-(q) = x + y \pmod p$; then $q \in \Omega_d$ iff $\pm d \in \{r_+(q), r_-(q)\}$. For $q \in H(1)$, $r_-^2 - r_+^2 = 4xy \equiv 4$; for $q \in H(-1)$, $r_-^2 - r_+^2 \equiv -4$. If $q \in \Omega_d \cap \Omega_e$ with $e \not\equiv \pm d$, then $\{r_+(q)^2, r_-(q)^2\} = \{d^2, e^2\}$, whence $e^2 \equiv d^2 \pm 4$. Conversely, if $e^2 \equiv d^2 + 4$, an $H(1)$-point with $r_+ \equiv \pm d$ has $r_- \equiv \pm e$ automatically, and an $H(-1)$-point with $r_- \equiv \pm d$ has $r_+ \equiv \pm e$; these are exactly the points listed, and there are no others (an $H(1)$-point with $r_- \equiv \pm d$ would need $r_+^2 \equiv d^2 - 4 \equiv e^2 - 8 \not\equiv e^2$). The case $e^2 \equiv d^2 - 4$ is symmetric. The last claim is Lemma 24(i): the four groups of Ω_d are good when d is (they are images of G_d under R, R'), and distinct good groups are disjoint. □

A seven-group d is *clean* if Ω_d is disjoint from every good group and from Ω_e for every seven-group $e \not\equiv \pm d$; by Lemma 29 this concerns only the residues e with $e^2 \equiv d^2 \pm 4$. Since RR' maps seven-groups to seven-groups, they come in pairs $\pm d$ with the same orbit; let $c_7(p)$ be the number of orbits Ω_d of clean seven-groups.

Theorem 30 (seven-point blocks)**.** *For every odd prime p, every lawful subset of $\mathcal{P}_{-1}$ has at most $4(p-1) - 4\,m_8(p) - 4\,c_7(p)$ points.*

Proof. Let U be the union of the good groups B (Lemma 24) and of the orbits of the clean seven-groups; the good orbits are pairwise disjoint (Lemma 24(i)), the clean orbits are disjoint from B and from each other by definition, so $|U| = 32m_8 + 64c_7$, and U is a union of whole rows and columns (Lemma 24(iii), Lemma 28(ii)). A lawful S has at most $12m_8 + 28c_7$ points in U (24 per good orbit, Lemma 28(iii), i.e. 6 lines per pair of good groups as in the proof of Theorem 21; 28 per clean orbit), and at most two points on each of the $2(p-1) - |U|/4$ rows outside U. Hence $|S| \le 12m_8 + 28c_7 + 4(p-1) - 16m_8 - 32c_7 = 4(p-1) - 4m_8 - 4c_7$. $\square$

The number of seven-groups has the same asymptotics as the number of good groups, and a fixed proportion of them is clean.

Proposition 31. *$m_7(p) = m_8(p) + O(\sqrt{p}\log^4 p) = \frac{1}{12}p + O(\sqrt{p}\,\log^4 p)$; likewise the number of $(2,6,6,2)$-groups of slope $+1$ is $\frac{1}{12}p + O(\sqrt{p}\log^4 p)$.*

Proof. In the proof of Proposition 26 the pairs (a,b) (there $\kappa = (a, 1/a)$, λ with $b+1/b \equiv a-1/a$) are equidistributed, with error $O(\sqrt{p}\log^4 p)$, with respect to the law in which $(x_\kappa, y_\kappa, x_\lambda)$ is uniform on $(-p/2, p/2)^3$ and y_λ is determined; the event $\mathcal{A}$ “the $H(1)$-pair is shared” ($\kappa \in A \cup D$) has probability $\frac{1}{2}$ (two independent signs), $\mathcal{A} \wedge \mathcal{B}$ (both shared) has probability $\frac{1}{3}$ (the volume computed there), so $\mathcal{A} \wedge \neg\mathcal{B}$ has probability $\frac{1}{6}$. The map $(x_\kappa, y_\kappa, x_\lambda) \mapsto (x_\lambda, y_\lambda, x_\kappa)$ is measure preserving on the torus and interchanges the roles of the two pairs, so $\neg\mathcal{A} \wedge \mathcal{B}$ has probability $\frac{1}{6}$ as well and $\neg\mathcal{A} \wedge \neg\mathcal{B}$ probability $\frac{1}{3}$. Multiplying by the number $\frac{1}{4}p + O(\sqrt{p})$ of full residues gives the claims. $\square$

Proposition 32. *$c_7(p) = \frac{7}{16} \cdot \frac{m_7(p)}{2} + O(\sqrt{p}\log^6 p) = \frac{7}{384}\,p + O(\sqrt{p}\,\log^6 p)$. Consequently*

$$\alpha(\mathcal{P}_{-1}) \;\le\; \Big(\tfrac{11}{3} - \tfrac{7}{96} + o(1)\Big)(p-1) \;=\; \Big(\tfrac{115}{32} + o(1)\Big)(p-1) \;=\; (3.59375 + o(1))(p-1).$$

Proof. Let d be a full residue with pairs $\{\kappa, \sigma\kappa\}$ ($d_\kappa \equiv d$) and $R\{\lambda, \tau\lambda\}$ ($e_\lambda \equiv -d$), and write $e_+ := e_\kappa$, $e_- := d_\lambda$; then $e_+^2 \equiv d^2 + 4$, $e_-^2 \equiv d^2 - 4$ (as $x_\kappa y_\kappa \equiv x_\lambda y_\lambda \equiv 1$), and by Lemma 29 the orbits meeting Ω_d are Ω_{e_+} and Ω_{e_-}. *The neighbour e_+.* The group G_{e_+} contains the R-images of the τ-pair $\{\sigma\kappa, \nu\kappa\}$ (indeed $e_{\sigma\kappa} = e_{\nu\kappa} \equiv -e_+$), which is shared iff $\sigma\kappa \in B \cup C$ iff $\kappa \in B \cup C$ (Lemma 7: σ preserves $A \cup D$), i.e. iff the $H(1)$-pair of G_d is *not* shared; and it contains an $H(1)$-pair $\{\mu, \sigma\mu\}$ with $d_\mu \equiv e_+$ iff $d^2 + 8 = e_+^2 + 4$ is a nonzero square, shared iff $\mu \in A \cup D$. So e_+ is good or a seven-group iff $d^2 + 8$ is a nonzero square and ($\kappa \in B \cup C$ or $\mu \in A \cup D$). *The neighbour e_-.* Symmetrically, G_{e_-} contains the $H(1)$-pair $\{\lambda, \sigma\lambda\}$, shared iff $\lambda \in A \cup D$ iff the $H(-1)$-pair of G_d is not shared, and an $H(-1)$-pair $R\{\mu', \tau\mu'\}$ ($e_{\mu'} \equiv -e_-$) iff $d^2 - 8$ is a nonzero square, shared iff $\mu' \in B \cup C$; e_- is good or a seven-group iff $d^2 - 8$ is a nonzero square and ($\lambda \in A \cup D$ or $\mu' \in B \cup C$).

Now let d be a seven-group with $\kappa \in A \cup D$, $\lambda \in A \cup D$ (the case $\kappa, \lambda \in B \cup C$ is symmetric under the exchange of the two pairs, as in Proposition 31). Then Ω_{e_-} is used iff $d^2 - 8$ is a nonzero square, and Ω_{e_+} is used iff $d^2 + 8$ is a nonzero square and $\mu \in A \cup D$. We claim that, among such d for which $d^2 + 8$ is a nonzero square, the proportion with $\mu \in A \cup D$ is $\frac{1}{4} + O(p^{-1/2}\log^6 p)$, and that the two square conditions are independent of each other and of the position conditions, each holding for a proportion $\frac{1}{2} + O(p^{-1/2}\log^6 p)$. Granting this, the proportion of clean seven-groups is $(1 - \frac{1}{2}\cdot\frac{1}{4})(1 - \frac{1}{2}) = \frac{7}{16}$ up to the error, and $c_7 = \frac{7}{16} \cdot \frac{m_7}{2} + O(\sqrt{p}\log^6 p)$; the last display follows from Theorem 30 and Propositions 26, 31.

The local law. Parametrise by $a = x_\kappa \bmod p$, $l = x_\lambda \bmod p$, $t = x_\mu \bmod p$: the relations are $l + 1/l \equiv -(a - 1/a)$ and $t - 1/t \equiv a + 1/a$, i.e. (a, l, t) runs over the curve C_1 whose function field is $\mathbb{F}_p(d)(\sqrt{d^2+4}, \sqrt{d^2-4}, \sqrt{d^2+8})$, absolutely irreducible of genus 5 (three independent quadratic extensions branched at the six roots of $(d^2+4)(d^2-4)(d^2+8)$). The six coordinates

$x_\kappa, y_\kappa, x_\lambda, y_\lambda, x_\mu, y_\mu$ are the centred (resp. positive) representatives of $a, 1/a, l, 1/l, t, 1/t$; the last two of these are integer linear forms in the four free residues $a, 1/a, l, t$ (plus a multiple of p), and every nontrivial integer linear form in $a, 1/a, l, t$ is a nonconstant function on C_1 (l and t have nonzero components along the independent square roots $\sqrt{d^2-4}$, $\sqrt{d^2+8}$, and $a, 1/a, 1$ are linearly independent), so the argument of Proposition 26 (Bombieri's bound on C_1, Selberg's polynomials for the six interval conditions) shows that the number of points of $C_1(\mathbb{F}_p)$ with the six coordinates in prescribed intervals is $|C_1(\mathbb{F}_p)|\lambda + O(\sqrt{p}\log^6 p)$, λ the volume of the corresponding region for the law in which $(u_1, u_2, u_3, u_5) = (x_\kappa, y_\kappa, x_\lambda, x_\mu)/p$ is uniform on $(-\frac{1}{2}, \frac{1}{2})^4$ and $u_4 = y_\lambda/p$, $u_6 = y_\mu/p$ are the representatives of $u_1 - u_2 - u_3$ and $u_5 - u_1 - u_2$ modulo 1 (we may use centred representatives throughout, since the types depend only on the signs). Points of C_1 exist over d iff d^2+8 is a nonzero square, two for each of the pairs, so restricting to C_1 is the same as imposing that condition; the second condition (d^2-8 square) is the Kummer character $\chi(d^2-8)$, and $d^2 \pm 4, d^2 \pm 8$ are multiplicatively independent modulo squares in $\mathbb{F}_p(d)$, so the same count on the curve C_2 with function field $\mathbb{F}_p(C_1)(\sqrt{d^2-8})$ (genus 17) gives the joint frequencies; this is the independence claimed.

The volume. Condition on (u_1, u_2) with $u_1u_2 < 0$ and put $s = u_1 + u_2$, $t_0 = |u_1| + |u_2|$, $m = \min(t_0, 1-t_0)$. For u_3 uniform, $u_4 \equiv -(u_3 - (u_2 - u_1))$, so $u_3u_4 < 0$ iff u_3 and the representative of $u_3 - (u_2 - u_1)$ have the same sign, which happens with probability $1 - 2m$ ($|u_2 - u_1| = t_0$); for u_5 uniform, $u_6 \equiv u_5 - s$, and $u_5u_6 < 0$ with probability $2|s|$. Given t_0, the difference $|u_1| - |u_2| = \pm s$ is uniform on $(-m, m)$, so $\mathbb{E}[2|s| \mid t_0] = m$. The density of t_0 is $4t_0$ on $(0, \frac{1}{2})$ and $4(1-t_0)$ on $(\frac{1}{2}, 1)$; weighting by $1-2m$ gives m the density $24m(1-2m)$ on $(0, \frac{1}{2})$, whence $\mathbb{P}(u_5u_6 < 0 \mid u_1u_2 < 0,\ u_3u_4 < 0) = \mathbb{E}[m] = 24\int_0^{1/2} m^2(1-2m)\,dm = \frac{1}{4}$. □

Remark 33. (a) *Data* (`slack/t221_agents/clean_counts.py`, all 405 primes in $[100, 3000]$ and five primes up to 10^5): the proportion of clean seven-orbits is 0.4424 (pooled; $7/16 = 0.4375$), the seven-orbits meeting only good orbits, only seven-orbits, and both are 32.5%, 18.4%, 4.9% of all seven-orbits (the local law gives $21/64 = 32.8\%$, $12/64 = 18.75\%$ and $\approx 5.0\%$ by simulation, the two neighbours being positively correlated through m), and $4(p-1) - 4m_8 - 4c_7 = (3.5945)(p-1)$ pooled ($115/32 = 3.59375$). The neighbour rule of Lemma 29 and the cover of Lemma 28 were checked point by point for all primes $p \le 500$ (`slack/t221_agents/orbit_cover.py`: 837 orbits, $|S \cap \Omega| \le 28$ certified by the LP in every case, the twelve lines of sizes $7, 5, 3$ partitioning 60 of the 64 points). (b) The certificate is again a cover by lines and rows; the LP with rows, columns and all slope-(±1) lines is lower ($\approx 3.45(p-1)$ for $p \le 199$), but its saving beyond the two blocks is not localised (Section 11). A seven-orbit whose only used neighbours are seven-orbits could be kept when they are discarded (an independent set in the neighbour graph, whose components are paths since every orbit has at most two neighbours); the model puts $\approx 3/32$ more of the seven-orbits within reach, i.e. a constant near 3.58, at the price of a longer bookkeeping. Both blocks together save $\frac{1}{3} + \frac{7}{96} = \frac{39}{96}$ of p; the LP saves about $0.55p$.

The blocks: runs of consecutive quadratic residues

The two constructions above (the eight-point groups, the seven-point groups) are special cases of a decomposition of $\mathcal{P}_{-1}$ that is exact and, as it turns out, carries the whole strength of the line cover by rows, columns and slope-(±1) lines.

For a point $q = (x, y) \in \mathcal{P}_{-1}$ put $r_+(q) = x - y$ and $r_-(q) = x + y$ modulo p. Since $r_-^2 - r_+^2 = 4xy \equiv \pm 4$ according as $q \in H(1)$ or $q \in H(-1)$, the two residues of a point are linked; we index everything by

$$t(q) := \tfrac{1}{4}\, r_+(q)^2 \in \mathbb{F}_p \qquad \text{for } q \in H(1), \qquad t(q) := \tfrac{1}{4}\, r_-(q)^2 \ \text{ for } q \in H(-1),$$

so that $q \in H(1)$ has $\{\frac{1}{4}r_+^2, \frac{1}{4}r_-^2\} = \{t, t+1\}$ and $q \in H(-1)$ has $\{\frac{1}{4}r_-^2, \frac{1}{4}r_+^2\} = \{t, t+1\}$: every point determines an unordered pair of *consecutive* values $t, t+1$, both of which are squares (as $4t = r^2$). Call $\{t, t+1\}$ the *edge* of q.

Lemma 34 (edges). *Let $t \in \mathbb{F}_p$ be such that t and $t+1$ are both squares (with 0 counted as a square). The points of $\mathcal{P}_{-1}$ with edge $\{t, t+1\}$ are the four lifts of each of the $H(1)$-classes κ with $d_\kappa^2 \equiv 4t$ together with the four lifts of each of their images under R, which are $H(-1)$-classes. For $t \neq 0, -1$ there are four such $H(1)$-classes (two σ-pairs, for $d_\kappa \equiv \pm 2\sqrt{t}$), hence 32 points; for $t = 0$ there are two (the τ-fixed classes $(1,1)$, $(-1,-1)$) and for $t = -1$ two (the σ-fixed classes, present iff -1 is a square), hence 16 points. If one of $t, t+1$ is a non-square there are no such points. The reflections R and R' preserve the edge, so every row and every column of the box consists of four points with the same edge.*

Proof. An $H(1)$-class $\kappa = (a, 1/a)$ has $r_+ \equiv a - 1/a =: d_\kappa$ and $r_- \equiv a + 1/a =: e_\kappa$ with $e_\kappa^2 - d_\kappa^2 = 4$, so its edge is $\{\frac{1}{4}d_\kappa^2, \frac{1}{4}d_\kappa^2 + 1\}$; conversely $d_\kappa \equiv \pm 2\sqrt{t}$ determines a up to the two roots of $a^2 - d_\kappa a - 1$ (which exist because $d_\kappa^2 + 4 = 4(t+1)$ is a square), and those two roots form a σ-pair, so that the two signs of d_κ give four classes unless $d_\kappa \equiv -d_\kappa$, i.e. $t = 0$, when $a^2 \equiv 1$ gives the single pair $\{(1,1), (-1,-1)\}$; the case $t = -1$ is $e_\kappa \equiv 0$, i.e. $a^2 \equiv -1$. By $R(x,y) = (p - x, y)$ we have $(r_+, r_-) \mapsto (-r_-, -r_+)$ and by $R'(x,y) = (x, 2p - y)$ we have $(r_+, r_-) \mapsto (r_-, r_+)$: both preserve the pair $\{\frac{1}{4}r_+^2, \frac{1}{4}r_-^2\}$ and exchange $H(1)$ with $H(-1)$, so the classes are as stated and rows ($\kappa \cup R\kappa$) and columns ($\kappa \cup R'\kappa$) stay inside the edge (Lemma 22). ☐

Let $Q := \{t \in \mathbb{F}_p : \ t \text{ is a square}\}$ (including $t = 0$) and consider the graph on Q whose edges are the pairs $\{t, t+1\}$ of consecutive squares (Lemma 34). Its connected components are the maximal *runs* of consecutive squares $t_0, t_0 + 1, \ldots, t_0 + k - 1$ (with 0 joining the runs that end at -1 and start at 1, if -1 is a square). For a run ϱ let $B_\varrho \subseteq \mathcal{P}_{-1}$ be the union of the points of its $k - 1$ edges, so that $|B_\varrho| = 32(k-1)$ unless the run contains 0, when each of its at most two degenerate edges contributes 16 instead of 32.

Theorem 35 (block decomposition). *The sets B_ϱ, ϱ a run of consecutive squares, partition $\mathcal{P}_{-1}$; each B_ϱ is a union of rows and of columns; and every row, every column and every slope-(± 1) line of the box lies inside a single B_ϱ. Consequently*

$$\alpha(\mathcal{P}_{-1}) \ \leq \ \sum_{\varrho} \beta(B_\varrho), \tag{4}$$

where $\beta(B)$ is the largest size of an $S \subseteq B$ with at most two points on every row, column and slope-(± 1) line, and the same sum with β replaced by the value of the linear relaxation equals the value of the linear relaxation for $\mathcal{P}_{-1}$ with rows, columns and slope-(± 1) lines. (Lines of other slopes may carry three points of $\mathcal{P}_{-1}$ — two from one hyperbola and one from the other — and such a line need not stay inside a block; dropping those constraints is what makes the bound block-wise, and it only weakens it.)

Proof. Every point lies on exactly one edge (Lemma 34), and the edges of one run are pairwise disjoint sets of classes, so the B_ϱ partition $\mathcal{P}_{-1}$; rows and columns stay inside an edge, hence inside a block. A line of slope $+1$ with residue d carries the $H(1)$-classes with $d_\kappa \equiv \pm d$ — edge $\{t, t+1\}$, $t = \frac{1}{4}d^2$ — and the $H(-1)$-classes with $r_+ \equiv \pm d$, whose edge is $\{t-1, t\}$ (their $r_-^2 = d^2 - 4$); both edges contain the vertex t, so they lie in the same run and the line lies in one block. Slope -1 is symmetric under R. Since the blocks partition $\mathcal{P}_{-1}$, a lawful S satisfies $|S| = \sum_\varrho |S \cap B_\varrho|$, and each $S \cap B_\varrho$ has at most two points on every row, column and slope-(± 1) line, so $|S \cap B_\varrho| \leq \beta(B_\varrho)$; the constraint matrix of the linear relaxation over those three families is block diagonal. ☐

Lemma 36 (equidistribution along a run). *Fix $k \geq 2$ and let C_k be the affine curve $\{(u_1, \ldots, u_k) : u_{i+1}^2 = u_i^2 + 1 \ (1 \leq i < k)\}$, i.e. the multiquadratic cover $\mathbb{F}_p(t)(\sqrt{t}, \sqrt{t+1}, \ldots, \sqrt{t+k-1})$ of the t-line ($t = u_1^2$). It is absolutely irreducible of genus $1 + 2^{k-2}(k-3)$, and for every $\mathbf{h} = (h_1, \ldots, h_k) \in \mathbb{F}_p^k \setminus \{0\}$ the function $\sum_i h_i u_i$ is non-constant on C_k. Consequently, for all* *intervals $I_1, \dots, I_k \subseteq (-p/2, p/2)$ and all $\varepsilon_0, \varepsilon_k \in \{\pm 1\}$,*

$$\#\Big\{(u_i) \in C_k(\mathbb{F}_p):\ X(u_i) \in I_i\ \forall i,\ \chi(u_1^2 - 1) = \varepsilon_0,\ \chi(u_k^2 + 1) = \varepsilon_k\Big\} = \frac{p}{4}\prod_{i=1}^{k}\frac{|I_i|}{p} + O_k(\sqrt{p}\log^k p),$$

where χ is the quadratic character and X the centred representative. In particular the positions $X(u_i)/p$ are asymptotically independent and uniform on $(-\frac{1}{2}, \frac{1}{2})$, and independent of the two Legendre conditions that make the run maximal.

Proof. Irreducibility: the k quadratic extensions are independent because the polynomials $t, t+1, \dots, t+k-1$ are pairwise coprime and squarefree (for $p > k$), so the Galois group over $\overline{\mathbb{F}_p}(t)$ is $(\mathbb{Z}/2)^k$ and the cover is connected. Genus: the cover of degree 2^k is ramified exactly over the k finite points $t = 0, -1, \dots, -(k-1)$ and over ∞, with inertia $\mathbb{Z}/2$ at each (at ∞ the inertia is generated by the simultaneous flip of all square roots, each $\sqrt{t+i}$ having odd polar order); each such point has 2^{k-1} preimages with $e = 2$, so $2g - 2 = 2^k(-2) + (k+1)2^{k-1}$, i.e. $g = 1 + 2^{k-2}(k-3)$. Non-constancy: the group $(\mathbb{Z}/2)^k$ acts on C_k by $u_i \mapsto \pm u_i$ independently; if $\sum_i h_i u_i$ were constant, averaging over the subgroup flipping only u_j would give $h_j u_j$ constant, hence $h_j = 0$ for every j. Now expand the interval indicators in Fourier series (or Selberg polynomials) of ℓ^1-norm $O(\log p)$ each and the two Legendre conditions as $\frac{1}{2}(1 \pm \chi)$; the main term is the product of the volumes times $|C_k(\mathbb{F}_p)|/4 = p/4 + O_k(\sqrt{p})$, and every other term is a mixed character sum $\sum_{C_k(\mathbb{F}_p)} \chi\big((u_1^2 - 1)^a (u_k^2 + 1)^b\big) e\big(\sum_i h_i u_i / p\big)$ with $(a, b, \mathbf{h}) \neq 0$, which is $O_k(\sqrt{p})$ by the Bombieri–Perel'muter bound: the additive part is non-constant when $\mathbf{h} \neq 0$, and for $\mathbf{h} = 0$ the Kummer sheaf is geometrically non-trivial because $u_1^2 - 1 = t - 1$ and $u_k^2 + 1 = t + k$ have odd order at their zeros on the t-line and their zeros are unramified in C_k. □

Proposition 37 (signature densities are descent statistics)**.** *Let a component of the neighbour graph be a maximal run $t, t+1, \dots, t+k-1$ of quadratic residues, and for $1 \le i \le k-1$ let $\xi_i \in \{0, 1\}$ record whether the $H(1)$-pair carried by the edge $(t+i-1, t+i)$ shares its centre. Then, as $p \to \infty$,*

$$\Pr\big[(\xi_1, \dots, \xi_{k-1}) = \varepsilon\big] = \frac{D_k(S_\varepsilon)}{k!} + O_k(p^{-1/2}\log^k p), \qquad S_\varepsilon = \{i : \varepsilon_i = 1\},$$

where $D_k(S)$ is the number of permutations of $\{1, \dots, k\}$ with descent set exactly S. The types of the vertices are determined by ε: the vertex between the edges $i-1$ and i is a good group $(4, 8, 4)$ iff $(\xi_{i-1}, \xi_i) = (0, 1)$, a $(2, 6, 6, 2)$-group iff $(1, 0)$, and a seven-group otherwise; the end vertices are $(4, 2, 2)$ iff $\xi_1 = 1$ (left), resp. $\xi_{k-1} = 0$ (right), and $(3, 3, 1, 1)$ otherwise.

Proof. Write u_i for a square root of $t+i-1$. The $H(1)$-pair of the edge $(t+i-1, t+i)$ is $\{a, 1/a\}$ with $a = u_i + u_{i+1}$, $1/a = u_{i+1} - u_i$ (indeed $a \cdot (1/a) = u_{i+1}^2 - u_i^2 = 1$ and $a - 1/a = 2u_i$ is the residue d of the group), and the pair shares its centre iff $X(a)X(1/a) < 0$; this is independent of the choice of signs of the u_i's. Put $v_i = X(u_i)/p$. Since the sign of the centred representative of x equals the sign of $\sin 2\pi x$ and $\sin 2\pi(v+w)\,\sin 2\pi(w-v) = \frac{1}{2}(\cos 4\pi v - \cos 4\pi w)$, we get the exact identity

$$\xi_i = 1 \iff \cos 4\pi v_i < \cos 4\pi v_{i+1} \iff g(v_i) > g(v_{i+1}), \qquad g(x) := \min\big(|x|, \tfrac{1}{2} - |x|\big).$$

By Lemma 36 the v_i are asymptotically independent and uniform, hence the $g(v_i)$ are independent and uniform on $(0, \frac{1}{4})$ (the map g is four-to-one and piecewise linear), so (ξ_i) is the descent pattern of k i.i.d. uniform variables, whose law is $D_k(S)/k!$. The type dictionary is Lemma 7 together with the fact that the pair on an edge is seen as an $H(1)$-pair by its left vertex and as (the R-image of) a τ-pair by its right vertex, which exchanges $A \cup D$ with $B \cup C$. □

Corollary 38 (density of runs by signature)**.** *For every fixed $k \geq 2$ and $S \subseteq \{1, \dots, k-1\}$, the number of maximal runs $t, t+1, \dots, t+k-1$ of quadratic residues whose signature has descent set S is*

$$\frac{D_k(S)}{k!} \cdot \frac{p}{2^{k+2}} + O_k(\sqrt{p} \log^k p).$$

Consequently, if $\mathrm{sav}(\varepsilon)$ *denotes the exact saving of a component with signature ε (a bounded integer computable by a finite optimisation, Section 6), then for every K*

$$\alpha(\mathcal{P}_{-1}) \;\leq\; \Big(4 - \sum_{k=2}^{K} \frac{1}{2^{k+2}} \sum_{S \subseteq \{1,\dots,k-1\}} \frac{D_k(S)}{k!} \,\mathrm{sav}(S) + o(1)\Big)(p-1),$$

and the truncation loses at most $\sum_{k>K} 4(k-1)2^{-(k+2)}$ (the saving of a component of length k never exceeds $4(k-1)$ in the computed range). Both statements are asymptotic in p for each fixed *K: the implied constant of Lemma 36 grows like $2^k k$ against a main term $p/2^{k+2}$, so the error is negligible only once $4^k k \log^k p \ll \sqrt{p}$ — already at $k = 2$ this needs p of order 10^{12}. In particular the limit of the displayed bound as $K \to \infty$ requires K to be sent to infinity* after *p, not with it.*

Proof. The run conditions are the $k+2$ Legendre conditions $\chi(t-1) = -1$, $\chi(t) = \cdots = \chi(t+k-1) = +1$, $\chi(t+k) = -1$; by Lemma 36 (applied to the k inner ones, which define C_k, with the two outer ones as the characters $\varepsilon_0, \varepsilon_k$) each of the 2^k sign choices of the u_i contributes equally, and the signature is the descent set of $(g(v_1), \dots, g(v_k))$ by Proposition 37, whose law is uniform. The last statement is the block decomposition (Theorem 35) restricted to the components of length at most K; dropping the others only decreases the saving. □

The saving $s(\varepsilon) := |B_\varrho|/2 - \beta(B_\varrho)$ of a block depends only on its signature (equivalently, on ε): this was verified for every block occurring in the primes $p \leq 20000$. Combining Proposition 37 with the exact values of β (a finite computation: `slack/t221/fast_blocks.py` constructs the block of a run directly from the square roots $u_i = \sqrt{t+i-1}$ and `scipy` solves the integer program), and using that the number of runs of length k is $2^{-(k+2)}p + O(\sqrt{p}\log p)$ (Weil), we obtain the following.

Corollary 39. *For every $K \geq 2$,*

$$\alpha(\mathcal{P}_{-1}) \;\leq\; \Big(4 - \sum_{k=2}^{K} 2^{-(k+2)} \sum_{\varepsilon \in \{0,1\}^{k-1}} \frac{D_k(S_\varepsilon)}{k!}\, s(\varepsilon) + o(1)\Big)(p-1),$$

a finite computation for each K. The table of $s(\varepsilon)$ was computed in three independent ways: by constructing the block of every run occurring at the primes $p \leq 20000$ (`slack/t221/fast_blocks.py`), by an independent implementation through the ξ-bits at $p \leq 9000$ (`slack/sig_savings.py`), and — with no prime at all — from the standard model *of a block, in which the positions $v_i = X(u_i)/p$ are replaced by generic rationals with the prescribed descent pattern and the incidences are decided by exact arithmetic (`slack/t221/standard_model.py`); the three agree wherever they overlap (500 comparisons for $k \leq 9$, no discrepancy), and the standard model supplies every ε, including those too rare to occur at moderate primes. With $K = 9$ the sum equals* 68350729/123863040*, so*

$$\alpha(\mathcal{P}_{-1}) \;\leq\; \frac{427101431}{123863040}\,(p-1) + o(p) \;=\; (3.44817\ldots + o(1))\,(p-1).$$

Since $s(\varepsilon) \leq 4(k-1)$, the terms with $k > 9$ contribute at most $\sum_{k>9} 2^{-(k+2)}4(k-1) = 0.0196$, so the constant of this certificate lies in [3.4286, 3.4482]*; numerically it is* 3.4413…

Remark 40. (a) The blocks are small: the run through t has length k with probability $\asymp 2^{-k}$, so $\beta(B_\varrho)$ is a bounded computation, and by Weil's bound the number of runs of a given length and given *type* (the sequence of line-size profiles of the groups at the interior vertices, which are the $(4,8,4)$, $(3,7,5,1)$, $(2,6,6,2)$ groups of Section 6, the end vertices carrying two-class groups) is $c\,p + O(\sqrt{p}\log^C p)$ with an explicitly computable c — the local law of Propositions 26, 31 applied along the run. Theorem 21 is the case of a single $(4,8,4)$ interior vertex ($\beta = |B|/2 - 8$ for the two edges at it), Theorem 30 the case of a $(3,7,5,1)$ vertex ($\beta = |B|/2 - 4$). (b) Numerically (`slack/t221/`, all primes $200 \le p \le 1400$): the saving $|B_\varrho|/2 - \beta(B_\varrho)$ depends only on the type of the run (no exception among the several hundred blocks tested), and $\sum_\varrho(|B_\varrho|/2 - \beta(B_\varrho)) = 0.5587\,p$ on average over $401 \le p \le 1103$, in agreement with Corollary 39; truncating to runs of length $\le L_0$ gives 3.595 for $L_0 = 5$, 3.480 for $L_0 = 7$ and 3.452 for $L_0 = 10$. At a given prime the bound (4) is at most the value of the linear relaxation with all these lines, and strictly below it when integrality bites inside a block (3.4107 vs 3.4286 at $p = 113$, 3.3824 vs 3.3873 at $p = 137$, 3.4343 vs 3.4545 at $p = 199$, 3.4100 vs 3.4300 at $p = 401$); the *proved* constant 3.4482 of Corollary 39 is larger than the numerical value 3.4413 of the same certificate only because of the truncation at $K = 9$. *Correction (v1.15):* an earlier version of this remark quoted the pair 3.4286 vs 3.4286 at $p = 113$ and the constant 3.4616 (a $K = 8$ truncation surviving from v1.11). An independent recount — exact MILP per block, the two disputed 64-point blocks refereed by a SAT solver (satisfiable at β, unsatisfiable at $\beta + 1$), and the LP ceilings $\lfloor 26.67\rfloor = 26$, $\lfloor 50.67\rfloor = 50$ forcing the same values — gives $\sum_\varrho \beta(B_\varrho) = 382 = 3.4107\,(p-1)$ at $p = 113$: the first number of that pair was the LP value, and integrality does bite there, in three blocks of twelve. The other three pairs of this remark were confirmed exactly by the same recount, as were the four block-sum anchors and, structurally, the decomposition itself (components under rows, columns and slope-(±1) lines match the runs of consecutive squares in number, size law and total for all 57 primes $19 \le p \le 311$). (c) The method is special to $k = \pm1$. For $\mathcal{P}_k = H(1) \cup H(k)$ a point of $H(c)$ has $r_-^2 - r_+^2 = 4c$, so with $t = \frac{1}{4}r_+^2$ the points of $H(1)$ join t to $t+1$ and those of $H(k)$ join t to $t+k$: for $k = \pm1$ the two rules coincide and the neighbour graph on the squares has mean degree 1 — subcritical, with components of size $O(\log p)$ — whereas for every other k the mean degree is 2, the critical value, and the components have size about $p^{2/3}$ (at $p = 4001$: largest component 11 for $k = -1$ against 22, 32, 95, 177 for $k = 2, 3, 5, 7$; `slack/blocks_general_k.py`). So a finite table of block optima exists only for the pair $H(1) \cup H(-1)$; for a general second hyperbola the decomposition is still valid but useless, and Section 7 (a saving of order $\sqrt{p}$) is what remains. (d) There is no local competition between the two hyperbolae at all. If O is a V-orbit of $H(1)$ (Lemma 17) and $R(O)$ its mirror in $H(-1)$, then the largest lawful subset of $O \cup R(O)$ — with *all* lines, of every slope — is exactly 16 for a generic orbit (32 points) and 8 for an exceptional one (16 points), i.e. exactly the trivial bound "two per row", although each half alone admits 12, resp. 6 (`slack/orbit_union.py`, all orbits of $p \le 23$). Summing over orbits returns $4(p-1)$: the whole gain from $4(p-1)$ down to the computed $3(p-1) + O(1)$ comes from the interaction of *different* orbits through three-point lines of general slope. No certificate that is local for the orbits or for the blocks can therefore reach $3(p-1)$. (e) The decomposition explains the plateau of Section 11: no certificate built from rows, columns and slope-(±1) lines can give less than $\sum_\varrho \beta(B_\varrho)$, and this is $\approx 3.44(p-1)$, far above the computed maxima $3(p-1) + O(1)$. The remaining gap is carried by the three-point lines of the other slopes, which are invisible to the block decomposition.

7 Every second hyperbola: the potential cover

Fix $k \in \mathbb{F}_p \setminus \{0, 1\}$ and let $\mathcal{P}_k = (H(1) \cup H(k)) \cap G(p)$. Its $8(p-1)$ points fall into the classes $\kappa_a = (a, 1/a) \in H(1)$ and $\mu_b = (b, k/b) \in H(k)$, $a, b \in \mathbb{F}_p^*$, four points each (the two lifts of the x- and of the y-coordinate). Two elementary facts organise the box: every non-empty row y carries exactly the four points of $\kappa_{1/y}$ and $\mu_{k/y}$, and every non-empty column x exactly those of κ_x and μ_x. Hence, if we join two classes when they share a row ("row-edge") or a column ("column-edge"), every class has one edge of each kind and the *class graph* is a disjoint union of even cycles

$$\kappa_x - \mu_{kx} - \kappa_{kx} - \mu_{k^2x} - \cdots - \kappa_x,$$

one for each coset of $\langle k\rangle$ in $\mathbb{F}_p^*$, of length $2\,\mathrm{ord}(k)$, with the κ's at the even and the μ's at the odd positions. A row-edge stands for two rows (y and $y+p$), a column-edge for two columns; each of these lines carries two points of each of its two classes. Slope-(± 1) lines are organised in residue groups exactly as in Section 6 (the residue d groups the classes with $a - 1/a \equiv d$ and $b - k/b \equiv d$ for slope $+1$, and $a + 1/a \equiv d$, $b + k/b \equiv d$ for slope -1); a group whose three lines have $4, 8, 4$ points (an "eight-group") consists of four classes, two of each hyperbola, and each of its classes has two points on the eight-point line and one point on each four-point line. (Lemma 23 and its proof carry over with the obvious changes: the $H(k)$-classes of the slope-$(+1)$ group of residue d are $\mu_b, \mu_{-k/b}$ with $b - k/b \equiv d$, their two centres are equal or differ by exactly p, and they coincide iff $X(b)$ and $X(k/b)$ have opposite signs; when both pairs of a group share their centres the two centres coincide automatically, for every k — the sign bookkeeping of the proof of Proposition 25; verified for all $p \le 500$ and $k \in \{-1, 2, 3, 5, 7\}$, `slack/g8_general.py`.) Let $G_8 = G_8(k,p)$ be the number of eight-point lines of both slopes and $s(v) \in \{0,1,2\}$ the number of eight-groups containing the class v. On a cycle c put $A_c = \sum_{\kappa\in c} s(\kappa)$, $B_c = \sum_{\mu \in c} s(\mu)$ (so $\sum_c A_c = \sum_c B_c = 2G_8$), and let R_c be the range (maximum minus minimum) of the $L+1$ partial sums $W_{-1} = 0,\ W_i = \sum_{j\le i} \varepsilon_j s(v_j)$ $(0 \le i \le L-1)$ along the cycle $v_0, v_1, \ldots, v_{L-1}$, where $\varepsilon_j = +1$ at even and -1 at odd positions; $R = \max_c R_c$.

Theorem 41 (every second hyperbola)**.** *For every prime p and every $k \in \mathbb{F}_p \setminus \{0,1\}$ with $G_8(k,p) \ge 1$,*

$$\alpha(\mathcal{P}_k) \ \le\ 4(p-1) \ -\ \frac{2G_8 - 2\sum_c |A_c - B_c|}{R}.$$

For $k = -1$ all cycles have length four, $A_c = B_c$ and $R = 1$, and the bound is $4(p-1) - 2G_8 = 4(p-1) - 4m_8$: Theorem 21.

Proof. We exhibit a fractional cover: non-negative weights on lines such that every point receives total weight at least one; since a lawful set meets every line in at most two points, its size is at most twice the total weight. Put $t = 1/R$ and weight t on the three lines of every eight-group (cost $6tG_8$; a point of a class v receives $s(v)\,t$ from these lines). On a cycle $v_0, \ldots, v_{L-1}$ let e_i be the edge $\{v_i, v_{i+1}\}$ and give both lines of e_i the weight

$$w_i = \tfrac{1}{2} + (-1)^i \varphi_i, \qquad \varphi_i \in [-\tfrac{1}{2}, \tfrac{1}{2}].$$

Every point of v_i lies on one line of e_{i-1} and one of e_i, hence receives $w_{i-1} + w_i = 1 + (-1)^i(\varphi_i - \varphi_{i-1})$ from rows and columns; so the cover is feasible as soon as $\varphi_i - \varphi_{i-1} \ge -t\,s(v_i)$ at even i and $\varphi_i - \varphi_{i-1} \le t\,s(v_i)$ at odd i (the "potential" may drop by at most ts at a κ and rise by at most ts at a μ; moves in the other direction are free), and $w \in [0,1]$ because $|\varphi| \le \frac{1}{2}$. The rows and columns of the cycle cost $4\sum_i w_i = 2L + 4\sum_i (-1)^i \varphi_i$, and for even L the alternating sum telescopes: $\sum_i (-1)^i \varphi_i = -\sum_{i \text{ odd}} (\varphi_i - \varphi_{i-1})$ (indices mod L). Choose $\varphi_i = \varphi_0 - tW_i$ for $0 \le i \le L-1$ (full moves at every special class; the value before v_0 is $\varphi_0 - tW_{-1} = \varphi_0$); since the $L+1$ partial sums range over an interval of length $R_c \le R = 1/t$, some φ_0 keeps $|\varphi| \le \frac{1}{2}$ throughout. The traversal ends at $\varphi_{L-1} = \varphi_0 - t(A_c - B_c)$; close the cycle by one free move at v_0 (a rise at the κ v_0 if $A_c > B_c$, a drop at the μ v_{L-1} if $B_c > A_c$; both endpoints are values of the walk, so the move stays inside $[-\frac{1}{2}, \frac{1}{2}]$). The sum of the odd-position jumps is then $tB_c - t\max(0, B_c - A_c) = t\min(A_c, B_c)$, and the total cost is

$$\begin{aligned}\sum_c \bigl(2L_c - 4t\min(A_c,B_c)\bigr) + 6tG_8 &= 4(p-1) - 4t\Bigl(2G_8 - \tfrac{1}{2}\sum_c |A_c - B_c|\Bigr) + 6tG_8 \\ &= 4(p-1) - t\Bigl(2G_8 - 2\sum_c |A_c - B_c|\Bigr),\end{aligned}$$

using $\sum_c L_c = 2(p-1)$ and $\min(A,B) = \frac{1}{2}(A + B - |A - B|)$. $\square$

The bound is exact for the cover with uniform line weight $1/R$: the class-level linear program with these line weights has exactly this value for every (p,k) we computed (`slack/exact_c.py`, `slack/cyc_model.py`), e.g. $\alpha(\mathcal{P}_3) \le 3.914\,(p-1)$ at $p=199$ ($G_8=34$, $R=4$) and $\alpha(\mathcal{P}_3) \le 3.952\,(p-1)$ at $p=401$; when $\operatorname{ord}(k)$ is small the cycles are many and the imbalance term can cancel the gain ($p=401$, $k=5$: sixteen cycles, $\sum_c |A_c-B_c| = G_8$, bound $4(p-1)$).

Proposition 42. *For every $k \in \mathbb{F}_p \setminus \{0,1\}$, $G_8(k,p) = \frac{1}{6}p + O(\sqrt{p}\log^4 p)$ with an absolute implied constant.*

Proof (sketch). The proof of Proposition 26 goes through with the cubic $C_0(k):\ ab^2-(a^2-1)b-ka=0$ (from $b-k/b \equiv a-1/a$): its discriminant in b, $a^4+(4k-2)a^2+1$, has the four distinct nonzero roots $a^2 = 1-2k \pm 2\sqrt{k^2-k}$ when $k \notin \{0,1\}$, so $C_0(k)$ is absolutely irreducible; an eight-point line of slope $+1$ corresponds to four points (a,b) of $C_0(k)$ with $X(a)X(1/a)<0$ and $X(b)X(k/b)<0$, and on $C_0(k)$ the fourth residue $X(k/b)$ is determined by $X(b)-X(a)+X(1/a)$ (substitute $b \mapsto -b$ in Step 2), so the conditions are again interval conditions on the four linear forms $x_1, x_2, x_3, x_3-x_1+x_2$ of total volume $\frac{1}{3}$; slope -1 is symmetric. All constants depend only on degrees. $\square$

Lemma 43 (arc imbalance)**.** *There is an absolute constant C such that for every prime p, every $k \in \mathbb{F}_p \setminus \{0,1\}$, every coset $a_0\langle k\rangle$ and every arc $I = \{a_0k^j :\ J_1 \le j \le J_2\}$ of the corresponding cycle,*

$$\Big|\sum_{j=J_1}^{J_2} \big(s(\kappa_{a_0k^j}) - s(\mu_{a_0k^{j+1}})\big)\Big| \ \le\ C\sqrt{p}\,\log^5 p.$$

Consequently $R \le C\sqrt{p}\log^5 p$ (the range of the partial sums is the maximum of an arc sum) and $|A_c - B_c| \le C\sqrt{p}\log^5 p$ for every cycle c; if k is a primitive root, $A_c = B_c$ exactly.

Proof. (This argument is due to the first of the agents mentioned in the disclosure; its write-up with numerical checks is the file `docs/research/arc_imbalance_lemma.md` of the repository.) *Fact 0.* The own-pair conditions of the two slopes are complementary (for slope $+1$ the pair $\{a,-1/a\}$ shares its centre iff $X(a)$, $X(1/a)$ have opposite signs, for slope -1 the pair $\{a,1/a\}$ shares iff they have equal signs; likewise for μ_b with $X(b), X(k/b)$), so a class lies in an eight-group of at most one slope: $s(v) \in \{0,1\}$, $s = f_K^+ + f_K^-$ on $H(1)$-classes and $s = f_M^+ + f_M^-$ on $H(k)$-classes, where $f_K^\pm(x)$ (resp. $f_M^\pm(y)$) is the indicator that κ_x (resp. μ_y) is a class of an eight-group of slope ±1; in particular $\sum_c A_c = \sum_c B_c = 2G_8$.

Step 1 (complete character sums). For a nontrivial multiplicative character ψ of $\mathbb{F}_p^*$ we show $|\sum_x f_K^+(x)\psi(x)| \le C_1\sqrt{p}\log^4 p$ (the three other functions are identical up to signs and the roles of x, b). Now $f_K^+(x) = 1$ iff $x^2 \ne -1$, the partner quadratic $b^2-(x-1/x)b-k=0$ has two roots (its discriminant $Q(x) = (x^4+(4k-2)x^2+1)/x^2$ is a nonzero square), $X(x)X(1/x)<0$, and $X(b)X(k/b)<0$ for one (equivalently both) of the roots. On the affine curve $C_0(k):\ xb^2-(x^2-1)b-kx=0$ there are exactly two points over each such x, so $\sum_x f_K^+(x)\psi(x) = \frac{1}{2}\sum_{(x,b)\in C_0(k)(\mathbb{F}_p)} B_1(x)B_2(b)\psi(x) + O(1)$, where $B_1(x)$ is the indicator of $X(x)X(1/x)<0$ and $B_2(b)$ that of $X(b)X(k/b)<0$. Each of B_1, B_2 is a sum of two products of two interval indicators on $\mathbb{Z}/p$, and an interval indicator has a Fourier expansion $\sum_m \hat{c}(m)e(mu/p)$ with $\sum_m |\hat{c}(m)| \le \log p + 3$; hence $\sum_x f_K^+\psi = \frac{1}{2}\sum c^1_{m_1n_1}c^2_{m_2n_2}T(\psi;f) + O(1)$ with $\sum|c^1|, \sum|c^2| \le 2(\log p+3)^2$ and

$$T(\psi;f) = \sum_{(x,b)\in C_0(k)(\mathbb{F}_p)} \psi(x)\,e\big(f(x,b)/p\big), \qquad f = m_1x + n_1x^{-1} + m_2b + n_2kb^{-1}.$$

$C_0(k)$ is absolutely irreducible of genus one (its function field is $\mathbb{F}_p(x,\sqrt{Q_1(x)})$ with $Q_1 = x^4+(4k-2)x^2+1$ squarefree for $k \notin \{0,1\}$), the function x has divisor $P_0^+ + P_0^- - P_\infty^+ - P_\infty^-$ with all multiplicities ±1, so x is not a d-th power times a constant for any $d \ge 2$ and the Kummer sheaf $\mathcal{L}_\psi(x)$ is geometrically nontrivial; the Artin–Schreier sheaf $\mathcal{L}_e(f)$ has order p, coprime to the order of ψ, so $\mathcal{L}_\psi(x) \otimes \mathcal{L}_e(f)$ is nontrivial for *every* f (including $f = 0$), and f has at most four poles, all of order ≤ 1. By the Weil–Bombieri bound for mixed sums along curves ([8]; explicit form in [10, Thm. 1]: at most $(2g - 2 + \deg(f)_\infty + \#(f)_\infty + \#\,\mathrm{supp}\,\mathrm{div}(x))\sqrt{p}$ for the smooth model, here $\le 12\sqrt{p}$, plus $O(1)$ for the excluded points) $|T(\psi; f)| \le 12\sqrt{p} + 8$. Summing, $|\sum_x f_K^+ \psi| \le 2C_2\sqrt{p}(\log p + 3)^4 + O(1)$.

Step 2 (completion of the arc). Let $L = \mathrm{ord}(k)$, $r = (p-1)/L$, and $F(y) = s(\kappa_y) - s(\mu_{ky})$ (so the sum in the lemma is $\sum_{j\in I} F(a_0 k^j)$). Expanding the indicator of $I \subset \mathbb{Z}/L$ in the characters $\chi_m(k^j) = e(mj/L)$ of $\langle k\rangle$ gives $|\sum_{j\in I} F(a_0k^j) - \frac{|I|}{L}\hat{F}(0)| \le (\log L + 1)\max_{m\ne 0}|\hat{F}(m)|$ with $\hat{F}(m) = \sum_j F(a_0k^j)e(-mj/L)$; each χ_m is the restriction of exactly r characters ψ of $\mathbb{F}_p^*$, and $\hat{F}(m) = \frac{1}{r}\sum_{\psi|\langle k\rangle = \chi_{-m}} \bar{\psi}(a_0)\sum_y F(y)\psi(y)$, where for $m \ne 0$ all these ψ are nontrivial; by Step 1 (applied to $s(\kappa_y)$ and to $s(\mu_{ky})$, the latter being $\bar{\psi}(k)$ times the sum for $s(\mu_y)$) $|\hat{F}(m)| \le 4C_1\sqrt{p}\log^4 p$. Finally $\hat{F}(0) = \sum_{y\in a_0\langle k\rangle} F(y) = A_c - B_c$ is $\frac{1}{r}\sum_{\psi|\langle k\rangle=1}\bar{\psi}(a_0)\sum_y F\psi$; the term $\psi = 1$ is $(2G_8 - 2G_8)/r = 0$ by Fact 0 and the others are again $\le 4C_1\sqrt{p}\log^4 p$ (there are none if $r = 1$). Both bounds together give the lemma with $C = 8C_1 + O(1)$. □

Corollary 44. *There are absolute constants $c, C > 0$ such that for every prime p and every $k \in \mathbb{F}_p \setminus \{0, 1\}$ with $\mathrm{ord}(k) \ge C\sqrt{p}\log^5 p$,*

$$\alpha(\mathcal{P}_k) \;\le\; 4(p-1) - c\,\frac{\sqrt{p}}{\log^5 p} \qquad (p \ge p_0).$$

In particular this holds for every primitive root k, and for all but $O(p^{1/2+o(1)})$ values of k (the elements of order $\le M$ in a cyclic group of order $p-1$ number at most $\tau(p-1)\,M$).

Proof. By Proposition 42, $G_8 \ge p/7$ for $p \ge p_0$; by Lemma 43, $R \le C\sqrt{p}\log^5 p$ and $\sum_c |A_c - B_c| \le \frac{p-1}{\mathrm{ord}(k)} C\sqrt{p}\log^5 p \le G_8/2$ once $\mathrm{ord}(k) \ge 14C\sqrt{p}\log^5 p$. Theorem 41 gives $\alpha(\mathcal{P}_k) \le 4(p-1) - G_8/R \le 4(p-1) - p/(14C\sqrt{p}\log^5 p)$. □

The saving $\sqrt{p}/\log^5 p$ is what the "never blocking" potential gives from the arc bound alone; numerically the potential cover with the uniform weight $t = 1/m$, $m = 4, \ldots, 8$, saves a positive proportion of G_8 (Remark below), and Lemma 43 is numerically sharp without the logarithms ($\max_\psi |\sum f_K\psi|/\sqrt{p} \in [1.2, 2.7]$ for $p \le 2003$, arcs $\le 1.0\sqrt{p}$).

Remark 45. (a) The potential construction is exact in the following sense: for fixed line weights t_G on the eight-groups, the cheapest assignment of row and column weights is $2L - 4D$ per cycle, where D is the total deposit of a buffer of width one driven along the cycle, every κ of an eight-group withdrawing at most t_G and every μ depositing at most t_G; Theorem 41 is the case of uniform $t = 1/R$, where the buffer never blocks. (b) Numerically the uniform weight $t = 1/m$ with $m = 4, \ldots, 8$ is much better: it saves about 0.2 per eight-group on every (p, k) computed ($p \le 997$, $k = 2, 3, 5$) and also on synthetic instances with randomly placed groups, i.e. $\alpha(\mathcal{P}_k) \le (4 - c_0 + o(1))(p-1)$ with $c_0 \approx 0.03$ appears to hold for every k; a proof needs the local statistics of the eight-group classes along the cycles (a fixed-window equidistribution at the level of fibred products of $C_0(k)$), which we have not carried out. The LP with rows, columns and the eight-groups' lines saves about one per eight-group ($p \le 601$), the ceiling of certificates of this kind.

8 Cubic graphs do not reach the HJSW value

We now leave the hyperbola and ask the same question of an arbitrary cubic curve $Y \equiv f(X) \pmod{p}$. Throughout this section $f \in \mathbb{F}_p[x]$ has degree 3 and $B = [x_0, x_0 + 2p) \times [y_0, y_0 + 2p)$ is an arbitrary $2p \times 2p$ box of integers. Every residue $x \in \mathbb{F}_p$ has exactly two lifts $X \in$

$\{x_0 + r_x,\, x_0 + r_x + p\}$ in the x-window ($r_x \in [0, p)$), and $y \equiv f(x) \pmod p$ has two lifts Y in the y-window, so

$$P(f, B) \;:=\; \{(X, Y) \in B :\; Y \equiv f(X) \pmod p\}, \qquad |P(f, B)| = 4p.$$

We call $P(f, B)$ the *box lift of the cubic graph $y = f(x)$*, and write $\alpha(f, B)$ for the largest size of a subset of $P(f, B)$ with no three points on a common line. We call f a *permutation cubic* if $x \mapsto f(x)$ is a bijection of $\mathbb{F}_p$; recall that the HJSW value on a $2p \times 2p$ box has side $N = 2p$ and the HJSW/hyperbola construction attains $3(p-1) = \frac{3}{2}N - 3$ (Theorem 1).

The projection bound and non-permutation cubics

Lemma 46 (projection bound)**.** *Let $f : \mathbb{F}_p \to \mathbb{F}_p$ be any function and B any $2p \times 2p$ box. Then $\alpha(f, B) \le 4\,|f(\mathbb{F}_p)|$.*

Proof. For $c \in f(\mathbb{F}_p)$, the two horizontal lines $Y \equiv c \pmod p$ meeting B together carry $2\,|f^{-1}(c)|$ points of $P(f, B)$ (two lifts of X for each $x \in f^{-1}(c)$), and a no-three-in-line subset meets a line in at most 2 points; summing over the $2\,|f(\mathbb{F}_p)|$ occupied rows gives $\alpha(f, B) \le 2 \cdot 2\,|f(\mathbb{F}_p)| = 4\,|f(\mathbb{F}_p)|$. ☐

By Dickson's classification of the low-degree permutation polynomials over $\mathbb{F}_p$ [11, 12], the permutation cubics are exactly the maps $f(x) = a(x+\beta)^3 + \gamma$ with $a \in \mathbb{F}_p^*$, $\beta, \gamma \in \mathbb{F}_p$, and $p \equiv 2 \pmod 3$; for $p \equiv 1 \pmod 3$ no cubic is a permutation of $\mathbb{F}_p$. If f is *not* a permutation cubic, its value-set size is governed by the Galois group of $f(x) - t$ over $\mathbb{F}_p(t)$ via Chebotarev's density theorem together with the Weil bound for the corresponding curve (a standard computation; see e.g. [4, 12]): generically this group is S_3 and

$$|f(\mathbb{F}_p)| = \tfrac{2}{3}p + O(\sqrt{p}) \qquad (S_3 \text{ case}),$$

while in the degenerate cyclic case ($f(x) - t$ has Galois group A_3, which forces $p \equiv 1 \pmod 3$ and, after an affine change of variable, $f(x) = a(x+\beta)^3 + \gamma$ again, now 3-to-1) the value set is exactly

$$|f(\mathbb{F}_p)| = \tfrac{p+2}{3} \qquad (A_3 \text{ case}).$$

Lemma 46 then gives, for every non-permutation cubic and every box,

$$\alpha(f, B) \;\le\; \Big(\tfrac{8}{3} + o(1)\Big)p \quad (S_3 \text{ case}), \qquad\qquad \alpha(f, B) \;\le\; \tfrac{4}{3}\,p + O(1) \quad (A_3 \text{ case}). \tag{5}$$

Both bounds are well below $\frac{11}{4}p$, the value at which Theorem 47 below caps the permutation case; this is what makes the permutation cubics, not the generic ones, the extremal case of the problem.

Permutation cubics

By the translations $x \mapsto x + \beta$, $y \mapsto y + \gamma$ (which move the box but not α) we may assume $f(x) = ax^3$, $a \ne 0$, $p \equiv 2 \pmod 3$. Write $N_3 = \#\{c \in \mathbb{F}_p :\; at^3 - t = c \text{ has three distinct roots in } \mathbb{F}_p\}$; the analogous count N_3^- for $at^3 + t = c$ satisfies $N_3^- = N_3 + O(1)$ (both root conics below have exactly $p+1$ points), which is all we use; we write N_3 for either.

Theorem 47 (permutation cubics)**.** *Let $p \equiv 2 \pmod 3$ be prime, $a \in \mathbb{F}_p^*$, and $f(x) = ax^3$. For every $2p \times 2p$ box B,*

$$\alpha(f, B) \;\le\; 4p - \tfrac{15}{2}N_3 + O(\sqrt{p}\log^3 p) \;=\; \tfrac{11}{4}\,p + O(\sqrt{p}\log^3 p),$$

where $N_3 = (p+1)/6 + O(1)$. The implied constants are absolute, independent of a, x_0, y_0.

The proof occupies the rest of this subsection: we first classify, exactly, the lines of slope ± 1 carrying ≥ 4 points of $P(f, B)$ (rows and columns are useless, as f is a bijection); we exhibit an explicit weighted cover by these lines; and we evaluate its cost to the stated precision using the Weil bound on a genus-0 conic together with a Bombieri–Perel'muter estimate for a quadratic (Kummer) twist along the same conic.

The ± 1-lines and their classes. Fix $x \in \mathbb{F}_p$ with lifts $X \in \{r_x, r_x + p\}$, $Y \in \{s_x, s_x + p\}$ ($r_x \in [x_0, x_0 + p)$, $s_x \in [y_0, y_0 + p)$). Among the four lifted points, $Y - X \in \{s_x - r_x$ (twice), $s_x - r_x + p$, $s_x - r_x - p\}$, and $Y - X \equiv ax^3 - x =: c_+(x) \pmod p$. For $c \in \mathbb{F}_p$ let $R(c) = \{t : at^3 - t = c\}$; generically $|R(c)| \in \{0, 1, 3\}$ (at most two values of c carry a double root). The integers $s_t - r_t$, $t \in R(c)$, all lie in the length-$2p$ interval $(y_0 - x_0 - p,\ y_0 - x_0 + p)$, which contains exactly two integers $\equiv c \pmod p$ (one, for a single exceptional residue c, which we ignore), say $c' < c' + p$; call t of *class* 0 if $s_t - r_t = c'$ and of *class* 1 if $s_t - r_t = c' + p$. With n_0, n_1 roots of each class, the four lines $Y - X \in \{c' - p, c', c' + p, c' + 2p\}$ carry n_0, $2n_0 + n_1$, $n_0 + 2n_1$, n_1 points of $P(f, B)$ respectively. For $|R(c)| = 3$ this profile is $(3, 6, 3, 0)$ or $(0, 3, 6, 3)$ if the three roots share a class (call c, or its root set, **same**) and $(2, 5, 4, 1)$ or $(1, 4, 5, 2)$ if they split $2{:}1$ (**split**); for $|R(c)| \leq 1$ every such line carries ≤ 2 points. Hence the slope-$(+1)$ lines with ≥ 4 points of $P(f, B)$ are exactly: one 6-point line for every *same* 3-root residue c, and one 5-point and one 4-point line for every *split* one (plus $O(1)$ lines from the ≤ 2 double-root residues, absorbed into the error term). A direct check of which lifts of a root $t \in R(c)$ lie on which line shows: in a *same* triple the two lifts with equal shift ($X = r_t + \delta p$, $Y = s_t + \delta p$, $\delta \in \{0, 1\}$) lie on a rich line and the two mixed lifts do not; in a *split* triple all four lifts of t lie on a rich line except one, namely $(\delta, \varepsilon) = (1, 0)$ if t has class 0 and $(0, 1)$ if t has class 1.

By the symmetric computation with $X + Y \equiv ax^3 + x =: c_-(x)$, the slope-$(-1)$ lines carry the same four profiles (now the roles of the "equal" and "mixed" lifts are reversed: a *same* triple rescues its two mixed lifts, a *split* triple rescues all but the lift $(0, 0)$ or $(1, 1)$ according to the class of the exceptional root), governed by the roots of $at^3 + t = c$.

Root conics and N_3. Since $at^3 - t - c$ has no quadratic term, its roots satisfy $t_1 + t_2 + t_3 = 0$, $t_1t_2 + t_1t_3 + t_2t_3 = -1/a$; eliminating $t_3 = -t_1 - t_2$ gives the conic

$$Q_+ : \ t_1^2 + t_1t_2 + t_2^2 = 1/a \qquad \text{(likewise } Q_- : \ t_1^2 + t_1t_2 + t_2^2 = -1/a \text{ for } at^3 + t - c).$$

The quadratic form $t_1^2 + t_1t_2 + t_2^2$ is the norm form of $\mathbb{F}_{p^2}/\mathbb{F}_p$ exactly when $p \equiv 2 \pmod 3$ (its discriminant -3 is then a non-residue), so $Q_\pm$ is a smooth conic with $|Q_\pm(\mathbb{F}_p)| = p + 1$; every 3-root residue c of the corresponding cubic accounts for exactly six ordered pairs of distinct roots on $Q_\pm(\mathbb{F}_p)$, so $N_3 = (p + 1 - r)/6$ with $r \in \{0, 6\}$ the (bounded) contribution of double roots, i.e. $N_3 = (p + 1)/6 + O(1)$, and the number of residues x that occur as a root of some 3-root c (for a fixed sign) is $3N_3 = \frac{1}{2}p + O(1)$. (The ≤ 2 double-root residues contribute $O(1)$ further points and at most $O(1)$ lines with ≥ 4 points; they are absorbed in the error terms below.)

The cover and its exact cost. Put weight 1 on every ± 1-line with ≥ 4 points and weight 1 on every point of $P(f, B)$ lying on none of them; since a no-three-in-line subset S meets a weight-1 line in ≤ 2 points, $|S| \leq 2L_4 + U$ where L_4 is the number of weighted lines and U the number of singleton points. Writing A^σ, S^σ ($\sigma \in \{+, -\}$) for the numbers of *same* and *split* 3-root residues of slope σ, $A^\sigma + S^\sigma = N_3$ and

$$L_4 = \sum_\sigma (A^\sigma + 2S^\sigma).$$

For U: fix a residue x and let $m_\pm(x)$ be the indicator that x is a root of a 3-root residue of slope $\pm$ (equivalently, that the "partner" quadratic $t^2 + xt + x^2 - 1/a = 0$, resp. $t^2 + xt + x^2 + 1/a = 0$, has two distinct roots in $\mathbb{F}_p$), and $\mathrm{split}_\pm(x)$ that this triple is a split one; write $\mathrm{class}_\pm(x) \in \{0, 1\}$ for the class of x in it when $m_\pm(x)$ holds. By the lift-count above, the two equal-shift lifts $(0, 0), (1, 1)$ of x are both covered by slope-$(+1)$ lines whenever $m_+(x)$; if $m_-(x)$ and $\mathrm{split}_-(x)$, exactly one of them is covered by a slope-(-1) line (the other is the exceptional lift of x in its split triple), and if $m_-(x)$ fails or the $(-)$-triple of x is a *same* one, neither is; symmetrically for the two mixed lifts with the roles of the slopes exchanged. Summing over the two kinds of lifts of every residue x,

$$U = \sum_{x\in\mathbb{F}_p} \Big\{[\neg m_+(x)](2 - [m_- \wedge \mathrm{split}_-](x)) + [\neg m_-(x)](2 - [m_+ \wedge \mathrm{split}_+](x))\Big\}. \tag{6}$$

The proof of Theorem 47 is now the evaluation, to error $O(\sqrt{p}\log^3 p)$, of N_3, $S^{\pm}$, and the two *cross sums* $\sum_x[\neg m_+][m_- \wedge \mathrm{split}_-]$, $\sum_x[\neg m_-][m_+ \wedge \mathrm{split}_+]$ in (6).

Classes as box conditions, and equidistribution. Let $u_t = (r_t - x_0)/p \in [0,1)$ be the position of the residue t in the x-window, $v_t = (s_t - y_0)/p$ the position of $f(t)$ in the y-window, and $g = g(c) = ((c-y_0+x_0) \bmod p)/p \in [0,1)$. For a root t of $at^3-t = c$, $s_t-r_t = (y_0-x_0)+p(v_t-u_t)$ with $v_t - u_t \equiv g \pmod 1$ and $|v_t - u_t| < 1$, so $s_t - r_t = c'$ (the smaller value) iff $v_t < u_t$ iff $u_t \geq 1-g$: writing $\theta = 1-g$,

$$\mathrm{class}_+(t) = [\,u_t < \theta\,],$$

all roots of the same c compared with the same threshold $\theta \in [0,1)$ (the labelling of the two classes is immaterial to what follows); the slope-(-1) computation is identical with the threshold $\theta' = g'(c_-)$, $g'(c) = ((c - x_0 - y_0) \bmod p)/p$ (class 0 iff $u_t \leq \theta'$). Over the six-fold cover of 3-root residues by ordered root pairs $(t_1,t_2) \in Q_+(\mathbb{F}_p)$, the triple $(u_{t_1}, u_{t_2}, g) \in (\mathbb{R}/\mathbb{Z})^3$ is equidistributed with error $O(\sqrt{p})$ per nonzero Fourier coefficient: for $(h_1,h_2,h_3) \neq 0$ the function $h_1t_1 + h_2t_2 + h_3(at_1^3 - t_1)$ is non-constant on the conic Q_+ (if $h_3 \neq 0$ it has a genuine cubic term in t_1, and t_2 is algebraic of degree 2 over $\mathbb{F}_p(t_1)$ on Q_+, so it cannot cancel it; if $h_3 = 0$ a non-trivial line meets the conic in ≤ 2 points), so $\sum_{Q_+(\mathbb{F}_p)} e(\cdot/p) = O(\sqrt{p})$ by the Weil bound for low-degree rational functions on $\mathbb{P}^1$ [4]. Expanding a box indicator in a Fourier/Selberg series of ℓ^1-norm $O(\log p)$ [6, 7] turns this into an error $O(\sqrt{p}\log^3 p)$ for the count of any fixed condition on $(u_{t_1}, u_{t_2}, \theta)$ (three interval indicators); the third root satisfies $u_{t_3} \equiv -u_{t_1} - u_{t_2} - \kappa \pmod 1$ with the constant $\kappa = (3x_0 \bmod p)/p$, so any two of the three roots, together with θ, are (up to relabelling) two coordinates of a point of $Q_+(\mathbb{F}_p)$ and inherit the same equidistribution; likewise for Q_-.

The split count: only pairwise data survive. Let $A_i = [u_{t_i} < \theta]$, $i = 1,2,3$, for the three roots of a 3-root c. The indicator of *same* is $[\mathrm{all} \geq \theta] + [\mathrm{all} < \theta]$, and

$$[\mathrm{all} \geq \theta] + [\mathrm{all} < \theta] = \big(1 - \textstyle\sum_i A_i + \sum_{i<j} A_iA_j - A_1A_2A_3\big) + A_1A_2A_3 = 1 - \sum_i A_i + \sum_{i<j} A_iA_j :$$

the triple term cancels, so *same* depends only on pairwise data. Since θ is equidistributed on $[0,1)$ independently (to the stated error) of each u_{t_i} and of each pair (u_{t_i}, u_{t_j}), $E[A_i] = E[\theta] = 1/2$ and $E[A_iA_j] = E[\theta^2] = 1/3$ to error $O(1/\sqrt{p})$ per residue, summed with $O(\sqrt{p}\log^3 p)$ absolute error over all $3N_3$ roots (equivalently N_3 triples), giving

$$\#\{\text{same 3-root residues}\} = N_3\big(1 - 3\cdot\tfrac{1}{2} + 3\cdot\tfrac{1}{3}\big) + O(\sqrt{p}\log^3 p) = \tfrac{1}{2}N_3 + O(\sqrt{p}\log^3 p),$$

hence, for both slopes,

$$A^\sigma = S^\sigma = \tfrac{1}{2}N_3 + O(\sqrt{p}\log^3 p) \qquad (\sigma = +,-), \tag{7}$$

with no dependence on κ, x_0, y_0: the split proportion is exactly $1/2$ for *every* box.

The cross sums: a Kummer twist and Bombieri–Perel'muter. Write $\Delta_+(x) = 4/a - 3x^2$, the discriminant of the partner quadratic $t^2 + xt + x^2 - 1/a$, so $[\neg m_+(x)] = \frac{1}{2}(1 - \chi(\Delta_+(x)))$ for $\Delta_+(x) \neq 0$ (χ the quadratic character of $\mathbb{F}_p^*$), and by the identity above $[\mathrm{split}_-](x) = \sum_i A_i - \sum_{i<j} A_iA_j$, a function of the box-positions of the roots of $at^3 + t = c_-(x)$, i.e. of the point $(t_1,t_2) = (x, \mathrm{partner}) \in Q_-(\mathbb{F}_p)$ (each x with $m_-(x)$ gives two such points). Thus

$$\sum_x [\neg m_+(x)][m_- \wedge \mathrm{split}_-](x) = \tfrac{1}{4} \sum_{(t_1,t_2)\in Q_-(\mathbb{F}_p)} (1 - \chi(\Delta_+(t_1))) \cdot [\mathrm{split}_-](t_1,t_2),$$

where [split_] is a box condition on $(u_{t_1}, u_{t_2}, \theta')$; the term without χ equals $\frac{1}{4}\cdot 2\sum_x [m_- \wedge \text{split}_-] = \frac{1}{2}\cdot 3S^-$. The term with χ is a mixed sum, over the genus-0 curve Q_-, of the multiplicative character $\chi(\Delta_+(t_1)) = \chi(4/a - 3t_1^2)$ times additive characters of the box conditions on (t_1, t_2, c_-). The Kummer sheaf of $\chi \circ \Delta_+$ pulled back to Q_- is geometrically non-trivial: Δ_+ has simple zeros at the (at most) four points of $Q_-(\overline{\mathbb{F}_p})$ over $t_1 = \pm\sqrt{4/(3a)}$, and these are unramified for the projection $Q_- \to t_1$ because Q_- ramifies over $t_1^2 = -4/(3a) \neq 4/(3a)$ ($p \neq 2, 3$); hence $\chi \circ \Delta_+$ is not a square in the function field of Q_-, and the twisted sheaf remains geometrically non-trivial after tensoring with the Artin–Schreier sheaf of any additive character (their orders are coprime). By the Bombieri–Perel'muter bound for mixed character sums along a curve [8] (as already used for Lemma 43 above), the χ-term is $O(\sqrt{p})$ absolutely, and after the same Selberg expansion of the box conditions, $O(\sqrt{p}\log^3 p)$. Hence

$$\sum_x [\neg m_+][m_- \wedge \text{split}_-] = \tfrac{1}{2}\cdot 3S^- + O(\sqrt{p}\log^3 p),$$

and symmetrically for the other cross sum with the roles of $+, -$ exchanged.

Assembly. Since $\sum_x [\neg m_\sigma(x)] = p - 3N_3 + O(1)$ (only the $3N_3 + O(1)$ roots of 3-root residues, plus $O(1)$ from double roots, satisfy m_σ), and $S^{\bar\sigma} = \frac{1}{2}N_3 + O(\sqrt{p}\log^3 p)$ for both slopes by (7) (writing $\bar\sigma$ for the sign opposite to σ),

$$\begin{aligned} U &= \sum_\sigma \Big[2(p - 3N_3) - \tfrac{3}{2}S^{\bar\sigma}\Big] + O(\sqrt{p}\log^3 p) \\ &= 2\big[2(p - 3N_3) - \tfrac{3}{4}N_3\big] + O(\sqrt{p}\log^3 p) \;=\; 4p - \tfrac{27}{2}N_3 + O(\sqrt{p}\log^3 p), \end{aligned}$$

while by $A^\sigma + 2S^\sigma = \frac{3}{2}N_3 + O(\sqrt{p}\log^3 p)$ (again from (7)),

$$2L_4 = 2\sum_\sigma (A^\sigma + 2S^\sigma) = 2\cdot 2\cdot \tfrac{3}{2}N_3 + O(\sqrt{p}\log^3 p) = 6N_3 + O(\sqrt{p}\log^3 p).$$

Combining, and using $N_3 = (p+1)/6 + O(1)$,

$$\text{cost} = 2L_4 + U = 6N_3 + 4p - \tfrac{27}{2}N_3 + O(\sqrt{p}\log^3 p) = 4p - \tfrac{15}{2}N_3 + O(\sqrt{p}\log^3 p) = \tfrac{11}{4}p + O(\sqrt{p}\log^3 p),$$

and $\alpha(f, B) \le$ cost since S meets a weighted line in ≤ 2 points and a singleton point in ≤ 1. This proves Theorem 47. ☐

Corollary 48 (no cubic graph reaches HJSW)**.** *For every prime p, every cubic $f \in \mathbb{F}_p[x]$, and every $2p \times 2p$ box B,*

$$\alpha(f, B) \;\le\; \Big(\tfrac{11}{8} + o(1)\Big)N, \qquad N = 2p.$$

In particular $\alpha(f, B) < \frac{3}{2}N - 3$ for p large: no cubic graph in any $2p \times 2p$ box attains, or beats, the HJSW value $3(p-1)$ of the modular hyperbola.

Proof. If $p \equiv 1 \pmod 3$, or $p \equiv 2 \pmod 3$ and f is not a permutation cubic, (5) gives $\alpha(f, B) \le (\frac{8}{3} + o(1))p$; if $p \equiv 2 \pmod 3$ and f is a permutation cubic, Theorem 47 gives $\alpha(f, B) \le \frac{11}{4}p + O(\sqrt{p}\log^3 p)$. Since $\frac{8}{3} < \frac{11}{4}$, in every case $\alpha(f, B) \le (\frac{11}{4} + o(1))p = (\frac{11}{8} + o(1))N$, and $\frac{11}{8} < \frac{3}{2}$. ☐

Remark 49 (numerics and further permutation polynomials). The cover of Theorem 47 is explicit and was checked by direct computation (`slack/cube_cover4.py`): at $p = 197$ ($N_3 = 33$, prediction $4p - \frac{15}{2}N_3 = 540.5$) five boxes give costs $544, 540, 532, 539, 534$, i.e. 1.35–$1.38\,N$, with fluctuations of the expected order $O(\sqrt{p})$; comparable agreement holds at $p = 293, 401$. The cover is not the true extremal LP: solving the line-cover LP restricted to ± 1-lines already improves the ratio to about 1.32–$1.36\,N$ (fractional weights and 3-point lines help), and the LP over *all* lines with ≥ 3 points of $P(f, B)$ drops further to about $1.10\,N$, using the three-point "family" lines that Theorem 47 discards; closing this gap to a proven constant near 4/3 is open. The same method applies to any permutation polynomial f of a fixed degree d (rows/columns are again useless): the ±1-lines organize by the residues of $f(x) \mp x$, their point-profiles are governed by the class pattern of the roots of $f(t) \mp t = c$ on the corresponding (bounded-genus) curve, and the constant is an explicit function of the root-count distribution — governed by the Galois group of $f(t) \mp t - c$ over $\mathbb{F}_p(t)$, via Chebotarev, exactly as in the non-permutation case above — computed by the same kind of argument (for residues with $r \geq 4$ roots the class pattern enters beyond pairwise data). Numerically the same ±1-line cover stays below $\frac{3}{2}N$ for x^5, x^7, x^{11} and the Dickson polynomial $D_5(x,1)$ on all sampled primes and boxes (1.15–$1.48\,N$; `slack/verification/perm_poly_covers.txt`), but with a smaller margin than for x^3: it is the third root of $at^3 \mp t = c$, present for a sixth of the residues, that creates the 6- and 5-point lines and pushes the cubic constant down to $\frac{11}{8}$.

9 Permutation monomials: the same cover, the same conclusion

Section 8 used only the lines of slope ±1 and the exact geometry of their point-profiles. Nothing in that argument is special to degree three except the final bookkeeping, and here we carry it out for the permutation monomials $f(x) = x^k$, k odd, $\gcd(k, p-1) = 1$ (for even k there are none). The result is the same as for cubics: the explicit ±1-line cover has cost about $1.40\,N$ for every k, so no permutation monomial graph reaches the HJSW value. The proof isolates three "local law" statements (root counts, root positions, and the two slopes) which we prove for x^k, and a general lift-accounting lemma; the constant is an explicit rational number.

Throughout, $k \geq 3$ is odd and fixed, $p \nmid k(k-1)$, $\gcd(k, p-1) = 1$, $B = [x_0, x_0 + 2p) \times [y_0, y_0 + 2p)$, $P = P(x^k, B)$ ($4p$ points, two per row and per column), and for $\sigma \in \{+1, -1\}$ we write $g_\sigma(t) = t^k - \sigma t$, so that $Y - X \equiv g_+(x)$ and $X + Y \equiv g_-(x) \pmod p$ for the lifts of the residue x. For $c \in \mathbb{F}_p$ let $R_\sigma(c) = \{t : g_\sigma(t) = c\}$; every residue x lies in exactly one $R_\sigma(c)$ for each σ, and we write $k_\sigma(x) = |R_\sigma(g_\sigma(x))|$. Classes and thresholds are exactly as in Section 8: with $u_t = (r_t - x_0)/p$ the position of t in the x-window, $\theta_+(c) = 1 - ((c - y_0 + x_0) \bmod p)/p$ and $\theta_-(c) = ((c - x_0 - y_0) \bmod p)/p$, a root $t \in R_+(c)$ has class 0 iff $u_t \geq \theta_+(c)$ and a root $t \in R_-(c)$ has class 0 iff $u_t \leq \theta_-(c)$; for a single residue x the two thresholds are tied by the identity

$$\theta_-(g_-(x)) \;\equiv\; 1 - \theta_+(g_+(x)) + 2u_x \pmod 1, \tag{8}$$

because $g_-(x) - g_+(x) = 2x$.

The lift-accounting lemma

Lemma 50 (profiles and lifts)**.** *Let c be a residue of slope σ with $|R_\sigma(c)| = k'$ roots, n_0 of class 0 and $n_1 = k' - n_0$ of class 1. The four consecutive lines of slope σ carrying the $4k'$ lifts of these roots have $n_0,\ k' + n_0,\ k' + n_1,\ n_1$ points. For a residue x with $k_\pm = k_\pm(x)$ and $n_\pm = n_\pm(x)$ the number of roots of its residue of slope $\pm$ lying in the class of x (so $1 \leq n_\pm \leq k_\pm$), the lines through the four lifts (δ, ε) of x have the following numbers of points ($\{a, b\}$ means the two values in some order for the two lifts of the pair):*

lifts	*slope* $+1$	*slope* -1
$(0,0),(1,1)$	$k_+ + n_+,\ k_+ + n_+$	$\{n_-,\ 2k_- - n_-\}$
$(0,1),(1,0)$	$\{2k_+ - n_+,\ n_+\}$	$k_- + n_-,\ k_- + n_-$

Consequently, for the cover W_4 (weight 1 on every ±1-line with ≥ 4 points, weight 1 on every point lying on none of them),

$$\begin{aligned} U(x) &= [k_+ + n_+ \leq 3]\big([n_- \leq 3] + [2k_- - n_- \leq 3]\big) \\ &\quad + [k_- + n_- \leq 3]\big([n_+ \leq 3] + [2k_+ - n_+ \leq 3]\big), \end{aligned} \tag{9}$$

$$L = \sum_\sigma \sum_c \Big([n_0 \geq 4] + [k' + n_0 \geq 4] + [k' + n_1 \geq 4] + [n_1 \geq 4]\Big), \tag{10}$$

$$\alpha(x^k, B) \ \leq \ \mathrm{cost}(W_4) = 2L + \sum_{x \in \mathbb{F}_p} U(x). \tag{11}$$

Proof. For slope $+1$ the lifts of a class-0 root t have $Y - X = c'$ (twice, $\delta = \varepsilon$), $c' + p$ $((0,1))$ and $c' - p$ $((1,0))$, those of a class-1 root $c' + p$ (twice), $c' + 2p$ $((0,1))$ and c' $((1,0))$, exactly as in Section 8; counting gives the profile $(n_0, k' + n_0, k' + n_1, n_1)$ on the lines $c' - p, c', c' + p, c' + 2p$, and reading off the line of each lift of x gives the first column of the table (a class-0 root has its equal-shift lifts on the line with $k' + n_0 = k' + n_+$ points, its $(0,1)$-lift on the line with $k' + n_1 = 2k' - n_+$ points and its $(1,0)$-lift on the line with $n_0 = n_+$ points; for a class-1 root the same three numbers occur with the two mixed lifts exchanged). For slope -1 the roles of "equal" and "mixed" lifts are exchanged (the profile sits on $X + Y \in \{c'', c'' + p, c'' + 2p, c'' + 3p\}$ with the two mixed lifts of a root on the central lines). A lift is a singleton of W_4 iff both its lines have ≤ 3 points; summing over the two lifts of each kind gives (9), and (11) is the usual duality (a lawful set meets a line in ≤ 2 points, a point in ≤ 1). ☐

For $k' = 3$ this is the bookkeeping of Section 8 ($k_+ + n_+ \leq 3$ iff $k_+ = 1$ or $k_+ = 2$ with a split pair; the bracket $[n_- \leq 3] + [2k_- - n_- \leq 3]$ equals 2 for a same triple and 1 for a split one). Note that for $k' \geq 4$ both central lines are always rich ($k' + n_0, k' + n_1 \geq 4$).

The three local laws for x^k

Lemma 51 (root counts)**.** *Let $p \nmid k(k-1)$. For each σ and $0 \leq j \leq k$, $\#\{c \in \mathbb{F}_p : |R_\sigma(c)| = j\} = P_j\, p + O_k(\sqrt{p})$ with $P_j = \frac{1}{j!}\sum_{i=0}^{k-j} \frac{(-1)^i}{i!}$ (the proportion of permutations of k letters with j fixed points; $P_{k-1} = 0$). In particular the number of residues x with $k_\sigma(x) = j$ is $jP_j\, p + O_k(\sqrt{p})$.*

Proof. The map $t \mapsto g_\sigma(t)$ has the $k-1$ simple critical points $t^{k-1} = \sigma/k$ with the pairwise distinct critical values $t\,(\sigma/k - \sigma)$, so its geometric monodromy group is generated by transpositions and is transitive ($g_\sigma(t) - c$ is irreducible over $\overline{\mathbb{F}_p}(c)$, being linear in c): it is S_k, hence so is the arithmetic monodromy group. Chebotarev's density theorem for the function field $\mathbb{F}_p(c)$ (with the Weil error term, the covering curve having genus bounded in terms of k) gives the count of c whose Frobenius has j fixed points, i.e. j rational roots. ☐

Lemma 52 (root positions)**.** *Let $1 \leq j \leq k$ and $\sigma \in \{\pm1\}$. Over the ordered j-tuples $(t_1, \ldots, t_j)$ of distinct roots of a common residue c of slope σ (one for each c and each ordered choice of j distinct elements of $R_\sigma(c)$), the vector $(u_{t_1}, \ldots, u_{t_j}, \theta_\sigma(c)) \in (\mathbb{R}/\mathbb{Z})^{j+1}$ is equidistributed with error $O_k(\sqrt{p}\log^{j+1} p)$ for box conditions, except that for $j = k$ the positions satisfy the one linear relation $\sum_i u_{t_i} \equiv -kx_0/p$ (and are equidistributed on that hyperplane). Consequently, for a residue with $j < k$ roots the number n_0 of class-0 roots is $\mathrm{Bin}(j, \theta)$ with θ uniform, i.e. uniform on $\{0, \ldots, j\}$, and for a random member x of such a residue $n_\sigma(x) = 1 + \mathrm{Bin}(j-1, q_\sigma)$, where q_σ is the probability that a partner has the class of x: $q_+ = 1 - \theta_+$ if $u_x \geq \theta_+$ and $q_+ = \theta_+$ otherwise; $q_- = \theta_-$ if $u_x \leq \theta_-$ and $q_- = 1 - \theta_-$ otherwise.*

Proof. The variety V_j of ordered j-tuples of distinct roots of $g_\sigma(t) = c$ is a curve whose components correspond to the orbits of the geometric monodromy group S_k on such tuples; S_k is j-transitive, so V_j is absolutely irreducible. For $(h_1, \ldots, h_j, h_0) \neq 0$ the function $\sum_i h_i t_i + h_0 g_\sigma(t_1)$ is non-constant on V_j when $j < k$: the fibre of $V_j \to V_{j-1}$ (forget t_j) has $k - j + 1 \geq 2$ points, so t_j is not a rational function of $(t_1, \dots, t_{j-1})$; inductively all $h_i = 0$ for $i \geq 2$, and then $h_1 t_1 + h_0 g_\sigma(t_1)$ is non-constant unless $h_1 = h_0 = 0$. For $j = k$ the same argument leaves exactly the relation $\sum_i t_i = 0$ (the coefficient of t^{k-1} in g_σ). Bombieri's bound on the curve V_j [4] gives $O_k(\sqrt{p})$ for each such exponential sum, and the Fourier expansion of the $j+1$ interval indicators (or Selberg's polynomials [6, 7]) gives the box-count with error $O_k(\sqrt{p}\log^{j+1} p)$. The consequences follow by integrating in θ: $\int_0^1 \binom{j}{m}\theta^m(1-\theta)^{j-m}\,d\theta = 1/(j+1)$. $\square$

Lemma 53 (the two slopes). *Let $D_\pm(x) = k^k(x^k \mp x)^{k-1} \mp (k-1)^{k-1}$ (up to a constant, the discriminant of $g_\pm(t) - g_\pm(x)$ as a polynomial in t), and let p not divide $k(k-1)\cdot \mathrm{Res}(D_+, D_-)$. Then, over $x \in \mathbb{F}_p$, the pairs $(k_+(x), k_-(x))$ are asymptotically independent with the laws of Lemma 51 (size-biased), and, given $(u_x, \theta_+(g_+(x)))$, the positions of the partners of x on the two slopes are independent and uniform, with the tie* (8) *for the second threshold; all with the error terms of Lemma 52.*

Proof. The Galois closures $L_\pm$ over $\mathbb{F}_p(x)$ of $g_\pm(t) - g_\pm(x)$ are ramified (over the affine x-line) exactly at the zeros of $D_\pm$; since $\mathrm{Res}(D_+, D_-) \not\equiv 0$, these finite branch loci are disjoint, so $L_+ \cap L_-$ is unramified over the affine line and tame at ∞ (its degree divides $k! < p$), hence trivial: L_+ and L_- are linearly disjoint and the fibre product W of the two root varieties over the x-line is absolutely irreducible. Chebotarev on the compositum gives the joint law of the two Frobenius classes (independence of the counts), and Bombieri on W gives the joint equidistribution of x, $g_+(x)$ and the partner positions of both slopes (no linear relation between them can hold, as in Lemma 52, by linear disjointness); the tie (8) is an identity, not a hypothesis. Finally $\mathrm{Res}(D_+, D_-) \neq 0$ for *every* odd k: a common root x of D_+ and D_- (over $\overline{\mathbb{Q}}$; the resultant is an integer) would satisfy $(x^k+x)^{k-1} = -(x^k-x)^{k-1}$, so $x^k + x = \zeta(x^k - x)$ with $\zeta^{k-1} = -1$, whence $x \neq 0$ and $x^{k-1} = -(1+\zeta)/(1-\zeta)$; then $(x^k-x)^{k-1} = x^{k-1}(x^{k-1}-1)^{k-1} = \pm 2^{k-1}(1+\zeta)/(1-\zeta)^k$ is a unit at every prime $\ell \mid k$ (the roots of unity $\zeta, -\zeta \neq 1$ have order dividing $2(k-1)$, coprime to ℓ, so $1 \mp \zeta$ are ℓ-adic units), whereas $D_+(x) = 0$ demands $(x^k - x)^{k-1} = (k-1)^{k-1}/k^k$, of negative ℓ-adic valuation — a contradiction. (Symbolically, for $k = 5$ the prime factors of the resultant below 5000 are $2, 5, 7, 97, 127$; for $k = 7$: $2, 3, 5, 7, 193, 1327$.) $\square$

The theorem

For the rencontres numbers P_j of Lemma 51 define

$$\begin{aligned} C_k \;:=\; & 2\sum_\sigma \sum_j P_j\, \mathbb{E}_{n_0\sim \mathrm{U}\{0..j\}}\big[[n_0 \geq 4] + [j + n_0 \geq 4] + [2j - n_0 \geq 4] + [j - n_0 \geq 4]\big] \\ & + \int_0^1\!\!\int_0^1 \mathbb{E}\big[U(x) \mid u_x = u,\ \theta_+ = 1 - g\big]\,du\,dg, \end{aligned} \tag{12}$$

where inside the expectation $k_\pm$ have the size-biased laws jP_j, $n_\pm = 1 + \mathrm{Bin}(k_\pm - 1, q_\pm)$ with $q_\pm$ as in Lemma 52 and $\theta_- = 1 - \theta_+ + 2u \bmod 1$, and $U(x)$ is (9). The integrand is a piecewise polynomial in (u, g) (four regions cut out by $g + 2u = 1$, $g + 2u = 2$, $u + g = 1$), so C_k is an explicit rational number:

$$C_3 = \tfrac{11}{4} = 2.75, \qquad C_5 = \tfrac{28183}{10080} = 2.79593\ldots, \qquad C_7 = \tfrac{15265237}{5443200} = 2.80446\ldots,$$

$$C_9 = \tfrac{94048104091}{33530112000} = 2.80488\ldots, \qquad C_{11} = 2.80489\ldots$$

(`slack/pp_constant_exact.py`; for $k = 3$ the two summands are $\frac{1}{2}\cdot 2$ and $\frac{7}{4}$, as in Section 8). As $k \to \infty$ the rencontres numbers converge to the Poisson law $e^{-1}/j!$ ($|P_j - e^{-1}/j!| \leq 1/(j!\,(k-j+1)!)$), and $C_k \to C_\infty = 2.80488\ldots$ (the value of (12) for the Poisson law); since the first summand of (12) is 16-Lipschitz in $\ell^1(P)$ and the second is 8-Lipschitz in ℓ^1 of the size-biased law $(jP_j)_j$, one gets $|C_k - C_\infty| \leq 2^{k+3}(k+5)/(k+1)!$, which is < 0.001 for $k \geq 11$; together with the exact values for $k = 7, 9$ this gives $C_k < 2.81$ for every odd $k \geq 7$.

Theorem 54 (permutation monomials). *Let $k \ge 3$ be odd, and let p be a prime with* $\gcd(k, p-1) = 1$ *and $p \nmid k(k-1)\mathrm{Res}(D_+, D_-)$ (a non-zero integer, Lemma 53). Then for every $2p \times 2p$ box B,*

$$\alpha(x^k, B) \;\le\; (C_k + \varepsilon_k + o(1))\, p, \qquad 0 \le \varepsilon_k \le \frac{2}{(k-1)!} + \frac{8}{k!}, \quad \varepsilon_3 = 0,$$

as $p \to \infty$ (k fixed); numerically $C_5 + \varepsilon_5 \le 2.946$, $C_7 + \varepsilon_7 \le 2.809$ and $C_k + \varepsilon_k < 2.81$ for all odd $k \ge 9$. In particular $\alpha(x^k, B) \le (1.474 + o(1)) \cdot 2p < \frac{3}{2} \cdot 2p$ for every odd k: no permutation monomial graph reaches the HJSW value.

Proof. By Lemma 50, $\alpha \le 2L + \sum_x U(x)$. The count L is a sum over residues of a function of (j, n_0); by Lemmas 51 and 52 the residues with $j < k$ roots contribute the first term of (12) times p, up to $O_k(\sqrt{p}\log^{k+1} p)$; the residues with $j = k$ roots have density $P_k = 1/k!$ per slope, and for them only the two outer indicators $[n_0 \ge 4] + [n_1 \ge 4]$ can deviate from the model (the central lines are rich for $k \ge 4$), so L deviates by at most $2 \cdot 2P_k p = 4p/k!$ and $2L$ by at most $8p/k!$ — this is the second part of ε_k. For $\sum_x U(x)$: by Lemma 53 the joint law of (k_+, n_+, k_-, n_-) for a random residue x is, up to the error terms, the one described after (12), except for the residues x lying in a k-root residue of some slope (density $kP_k = 1/(k-1)!$ per slope), where the partner positions obey the linear relation of Lemma 52; for such x the only factor of (9) that can be affected is a single $[n_\sigma \le 3]$ (the other factor $[k_\sigma + n_\sigma \le 3]$ vanishes for $k_\sigma = k \ge 4$), so replacing it by 1 changes $\sum_x U(x)$ by at most $2/(k-1)! \cdot p$ — the first part of ε_k. All remaining terms are the integral in (12) times p, with error $O_k(\sqrt{p}\log^{k+1} p)$. For $k = 3$, k-root residues are exactly the three-root residues, and Section 8 shows that only pairwise data enter, whence $\varepsilon_3 = 0$ and $C_3 = 11/4$. The numerical values of C_k follow from the exact integration of the piecewise polynomial integrand. □

Corollary 55 (generic permutation polynomials). *Let $k \ge 3$ and let $f \in \mathbb{F}_p[x]$ be a permutation polynomial of degree k such that for both $g_\pm(t) = f(t) \mp t$: (i) the map $t \mapsto g_\pm(t)$ has monodromy group S_k (for instance, $g'_\pm$ has $k-1$ distinct simple zeros with pairwise distinct critical values, and $p \nmid k$), and (ii) the discriminants* $\mathrm{disc}_t(g_+(t) - g_+(x))$ *and* $\mathrm{disc}_t(g_-(t) - g_-(x))$ *are coprime polynomials in x. Then for every $2p \times 2p$ box, $\alpha(f, B) \le (C_k + \varepsilon_k + o(1))\, p$ with the same C_k and ε_k as in Theorem 54 (the implied constants depending on k and on the family, uniformly in the box).*

Proof. Nothing in Lemmas 50–53 or in the proof of Theorem 54 used the shape of f beyond (i) and (ii): the classes, thresholds and the tie (8) are defined for any f (as $g_-(x) - g_+(x) = 2x$), the root-count laws are the rencontres numbers as soon as the monodromy is S_k, the j-tuple curves are irreducible by j-transitivity, the only linear relation is $\sum_i t_i = \text{const}$ for $j = k$ (covered by ε_k), and (ii) gives the linear disjointness of the two slopes. Hence the constant depends on k only. □

For an arbitrary permutation polynomial (no assumption on the monodromy) the same proof gives $\alpha(f, B) \le (C(P^+, P^-) + \varepsilon + o(1))p$ with the root-count laws $P^\pm$ of $f \mp x$ in place of the rencontres numbers, provided the corresponding root varieties are irreducible and the two slopes are linearly disjoint; whether every permutation polynomial of degree ≥ 3 satisfies $\alpha \le (\frac{3}{2} - c_k)N$ is open (see (c) below).

Remark 56 (numerics; the true constant; other bijections). (a) The bookkeeping of Lemma 50 evaluates $\mathrm{cost}(W_4)$ in time $O(p)$ without enumerating lines (`slack/pp_exact_UL_np.py`). At $p \approx 10^7$ (three primes, box $(0,0)$) it gives $\sum_x U(x)/p = 1.7500 \pm 0.0002$ for $k = 3$ (the exact $7/4$) and 1.7631 ± 0.0002 for $k = 5$, against the value $17767/10080 = 1.76260$ of the tied-threshold law in (12) — the small excess is the k-root correction ε_5, so the true asymptotic constant for x^5 is $2.797p \approx 1.3985\,N$; a model with independent slopes would give $1519/864 = 1.7581$, which the data exclude. At $p \le 887$ the cost of W_4 fluctuates by a few percent (for x^5: mean $1.388\,N$, range 1.36–1.41

over seven primes; for x^7: mean $1.415\,N$; `docs/research/g_agents/x5_x7_numerics.md`); the LP over all ± 1-lines is ≈ 1.32–$1.34\,N$ and the LP over all lines with ≥ 3 points ≈ 1.07–$1.10\,N$, as for cubics. (b) Why the hyperbola $x \mapsto 1/x$ sits exactly at $3/2$ in this framework: every residue has exactly two roots on each slope (no lines with 5 or 6 points), and since $c = t_1 + t_2$ is a *linear* function of the root pair, the same-class probability of a pair is $1/2$ rather than the $2/3$ of Lemma 52; the cover W_4 then costs $2 \cdot \frac{p}{2} + 2p = 3p$ whatever the correlation between the two slopes (a residue contributes two covered lifts for each slope on which its pair is same), in accordance with Theorem 1. (Its two slopes are in fact maximally dependent — the partners of x are $\mp 1/x$ and the pair of slope $+1$ shares its centre iff that of slope -1 does not — so Lemma 53 would fail for it; this affects only the double-coverage bookkeeping, not the value $3p$.) For x^k the residues with three or more roots and the same-probability $2/3$ of the pairs are what push the constant below $3/2$. (c) For a general permutation polynomial f of degree k the same proof gives $\alpha \leq (C(P^+, P^-) + \varepsilon + o(1))p$ with the root-count laws $P^\pm$ of $f \mp x$ in place of the rencontres numbers, as soon as the two covers $t \mapsto f(t) \mp t$ have monodromy S_k and disjoint finite branch loci; and if $f - x$ is not a permutation polynomial, at least a positive proportion of its residues carry ≥ 2 roots (Wan's bound $|V(f-x)| \leq p - (p-1)/k$ [13]), while for p large in terms of k no polynomial of degree ≥ 2 is a complete mapping [14] — so a linear saving from ± 1-lines is available for every permutation polynomial of bounded degree, though with a constant that depends on the family.

10 The strong form for cubic graphs: W_4 plus a matching

Theorem 47 used only the lines of slopes ± 1 and gave $\frac{11}{8}N$ for the permutation cubics; the fractional cover LP over all lines gives about $1.10\,N$ (Section 8, remark), so the three-point lines of the other slopes carry a lot. Here we show how to use them rigorously and cheaply: after the ± 1-cover we take a *matching* of three-point lines all of whose points are still uncovered. Only the number of such lines and the number of intersecting pairs enter, and both are counted by the same local-law machinery (Lemma 59). The outcome is $\alpha \leq (1.33 - \varepsilon)N$ for the permutation cubics and hence the bound $(\frac{4}{3} + o(1))N$ for *every* cubic graph.

Three-point lines and the certificate

We keep the notation of Section 8: $f(x) = ax^3$, $p \equiv 2 \pmod 3$, box $B = [x_0, x_0{+}2p) \times [y_0, y_0{+}2p)$, $P = P(f, B)$, lifts $(r_x + \delta p,\ s_x + \varepsilon p)$; W_4 is the cover "weight 1 on every ± 1-line with ≥ 4 points", L_4 their number, $S_0 \subset P$ the set of points on none of them ($2L_4 + |S_0| = \frac{11}{4}p + O(\sqrt{p}\log^3 p)$, Theorem 47). Three residues x_1, x_2, x_3 are collinear on the residue curve iff $x_1 + x_2 + x_3 = 0$ (they are then the roots of $ax^3 - mx - c$). A three-point line of P with a direction of slope $\notin \{0, \infty, \pm 1\}$ has three distinct residues; we call it a line of *family* $(1,3)$ if its points are $q,\ q+d,\ q+3d$ for an integer vector $d = (u^*, v^*)$: writing $z = 3^{-1}u^* \bmod p$, its residues are $-4z,\ -z,\ 5z$ (they sum to 0) and $v^* \equiv f(-z) - f(-4z) = 63az^3 \pmod p$; conversely every $u \in \mathbb{F}_p^*$, every lift of the base residue $-4z$ and every pair of representatives (u^*, v^*) with $q + d, q + 3d \in B$ gives such a line, and distinct data give distinct lines. (In the general description of the three-point lines by steps (i, j) — points $q, q + id, q + jd$, base residue $-(i + j)u/3$ — this is the family $(1, 3)$; the family $(1, 2)$, the lines through the four lifts of the centre $(0, f(0))$, is useless for what follows since all its lines meet in four points, and it is excluded throughout.) A three-point line is *good* if its three points lie in S_0.

Lemma 57 (the certificate)**.** *If M is a set of pairwise disjoint good three-point lines, then $\alpha(f, B) \leq 2L_4 + |S_0| - |M| = (\frac{11}{4} + o(1))p - |M|$.*

Proof. Weight 1 on the ± 1-lines of W_4, weight 1 on the lines of M, weight 1 on the points of S_0 not on a line of M: every point of P receives weight ≥ 1, the lines of M are disjoint and inside S_0, and a lawful set meets a line in ≤ 2 points; the cost is $2L_4 + 2|M| + (|S_0| - 3|M|)$. ☐

Lemma 58 (matching)**.** *Let G be the set of good lines of family $(1,3)$ and Π the set of unordered pairs of lines of G with a common point. Then G contains a set of $|G|-|\Pi|$ pairwise disjoint lines.*

Proof. Delete one line from each pair of Π. $\square$

The local law and the two constants

Fix a residue x with position $\pi = ((x-x_0) \bmod p)/p$ in the x-window and thresholds $\theta_+(x) = 1-((f(x)-x-y_0+x_0) \bmod p)/p$, $\theta_-(x) = ((f(x)+x-x_0-y_0) \bmod p)/p$ (Section 8). Whether a lift (δ,ε) of x is in S_0 depends, besides $(\pi,\theta_+,\theta_-,\delta,\varepsilon)$, only on the following data of x: whether the partner conics $t^2+xt+x^2 \mp 1/a = 0$ have roots (the Legendre symbols of $\Delta_\pm = \pm 4/a - 3x^2$), and the positions of the roots (t_2 and $t_3 = -x-t_2$ on each slope), through the classes and the line sizes of Lemma 50. Averaging over these data with their limiting law — fair independent coins for the two symbols, and a uniform independent position of t_2 on each slope — defines the *uncovered probability*

$$U(\pi,\theta_+,\theta_-,\delta,\varepsilon;\ x_0/p) \in [0,1],$$

an explicit piecewise polynomial (it factorises as $P_+ \cdot P_-$, the two slopes being independent; for $\delta=\varepsilon$ the factor P_+ equals $\frac{1}{2}$ exactly, since the slope-$(+1)$ line through such a lift has $k_+ + n_+ \geq 4$ points iff $k_+ = 3$). A line of family $(1,3)$ is described by $s = u^*/p \in (-\frac{2}{3},\frac{2}{3})$, the residue class of u^* modulo 3 (which fixes $z/p = (s+r)/3$ as a real number), $w' = (az^3 \bmod p)/p \in [0,1)$, the lift (δ_1,ε_1) of its base point and the choice of representative v^*; the positions of its residues are $\{a_m(s+r)/3 - x_0/p\}$ with $(a_m) = (-4,-1,5)$, the positions of $f(x_m)$ are $\{a_m^3 w'\}$, and $v/p = \{63w'\}$; the box constraints for $q+d, q+3d$ then determine the admissible (u^*,v^*) and the lifts (δ_m,ε_m) of the other two points. We define

$$\gamma_1(x_0/p) := \sum_{\delta_1,\varepsilon_1} \frac{1}{3}\sum_{r=0}^{2} \int ds \int_0^1 dw' \sum_{v^*\ \text{admissible}} \prod_{m=1}^{3} U(\pi_m,\theta_{+,m},\theta_{-,m},\delta_m,\varepsilon_m;x_0/p), \tag{13}$$

$$\gamma_2(x_0/p,y_0/p) := \text{the analogous integral over pairs of family-}(1,3)\text{ lines with a common point,} \tag{14}$$

where in (14) the second line is determined by the first and the roles (m,m') of the common point ($z' = (a_m/a_{m'})z$; for the ratios $\frac{1}{4},\frac{1}{5},\frac{4}{5},\frac{5}{4}$ the digits of z modulo 4 and 5 and of az^3 modulo 64 and 125 enter as further uniform independent variables), and the integrand is the product of the five residues' uncovered probabilities. All variables are explicit; the integrands are piecewise polynomial in (s,w').

The counting input is the equidistribution lemma for a family of three-point lines (second solver):

Lemma 59 (equidistribution along a family of three-point lines)**.** *Let $p \equiv 2 \pmod 3$, $f(x) = ax^3$ with $a \in \mathbb{F}_p^*$, and let (i,j) be a family of three-point lines with span $J \geq 3$ (so that the three residues of a line are $\alpha_m u$, $m = 1,2,3$, with $\alpha_m \in \{-(i+j)/3, (2i-j)/3, (2j-i)/3\}$ and pairwise distinct squares α_m^2), with direction (u^*,v^*), $v \equiv c_1u^3 + au$, $c_1 = (i^2-ij+j^2)/3$. For $\sigma \in \{\pm 1\}$ and $m = 1,2,3$ let $\Delta_{\sigma,m}(u) = 4\sigma/a - 3\alpha_m^2u^2$ be the discriminant of the partner conic $t^2+\alpha_m ut+\alpha_m^2u^2-\sigma/a = 0$ of the residue $\alpha_m u$, and let $t_{\sigma,m}$ denote one of its roots (the other is $-\alpha_m u - t_{\sigma,m}$). Let F be a fixed finite Boolean combination of the following conditions: interval conditions on the least residues of u (equivalently on the representative u^*), of $v(u)$, of the linear forms $\alpha_m u$, and of the roots $t_{\sigma,m}$ and $-\alpha_m u - t_{\sigma,m}$; and the six membership conditions "$\Delta_{\sigma,m}(u)$ is a nonzero square". Then*

$$\#\{u \in \mathbb{F}_p^*:\ F\} \ =\ p\,\mu(F) + O_{i,j}(\sqrt{p}\,\log^{16} p),$$

where $\mu(F)$ is the probability of F under the following "local law": u^*/p is uniform on its interval, the residues $\alpha_m u$ are the deterministic multiples $\alpha_m u^* \bmod p$, $v(u)/p$ is an independent uniform variable, the six membership bits are independent fair coins, and, conditionally on membership, each root $t_{\sigma,m}/p$ is an independent uniform variable (the second root being $-\alpha_m u - t_{\sigma,m}$). The implied constant depends only on (i,j); the statement is uniform in a and in the box, and holds for p larger than a constant depending on (i,j).

Proof. (1) The fibre products. The six discriminants are quadratics in u; the roots of $\Delta_{\sigma,m}$ are $\pm(2/\alpha_m)\sqrt{\sigma/(3a)}$. For fixed σ the three root sets are pairwise disjoint because the α_m^2 are pairwise distinct; the root sets of $\Delta_{+,m}$ and $\Delta_{-,m'}$ coincide only if $\alpha_m^2 \equiv -\alpha_{m'}^2 \pmod p$, a congruence between fixed rationals which we exclude by taking p large. Hence the six $\Delta_{\sigma,m}$ are pairwise coprime squarefree polynomials, their classes in $\overline{\mathbb{F}_p}(u)^*/(\overline{\mathbb{F}_p}(u)^*)^2$ are multiplicatively independent, and for every set S of pairs (σ,m) the fibre product X_S of the conics indexed by S over the u-line (function field $\overline{\mathbb{F}_p}(u,\sqrt{\Delta_s} : s\in S)$) is an absolutely irreducible curve of degree $2^{|S|}$ over the u-line, of genus bounded in terms of $|S|$ (at most twelve branch points), whose $\mathbb{F}_p$-points over a given u number $2^{|S|}$ when all $\Delta_s(u)$, $s\in S$, are nonzero squares and 0 when one of them is a non-square (the $O(1)$ values of u with some $\Delta_s(u)=0$ are discarded). Moreover the square roots $\sqrt{\Delta_s}$, $s\in S$, are linearly independent over $\overline{\mathbb{F}_p}(u)$, and for S' disjoint from S the function $\prod_{s\in S'}\Delta_s$ is not a square in $\overline{\mathbb{F}_p}(X_S)$.

(2) Reduction to character sums. Fix a membership pattern, with member set S and non-member set $\bar S$. Writing $[\Delta_s(u) \text{ is a nonzero square}] = \frac12(1+\chi(\Delta_s(u)))$ and its negation as $\frac12(1-\chi(\Delta_s(u)))$ (χ the quadratic character), and every interval condition on a least residue as a Fourier series on $\mathbb{Z}/p$ with ℓ^1-norm $\le \log p + 3$ (or as Selberg polynomials of degree $H=\lfloor\sqrt p\rfloor$, [6, 7]), the count of u satisfying F with this membership pattern equals $2^{-|S|}\sum_{P\in X_S(\mathbb{F}_p)}\prod_{s\in\bar S}\frac12(1-\chi(\Delta_s(u(P))))\cdot\prod(\text{interval indicators})$, where the roots t_s, $s\in S$, are the coordinates of P. Expanding, the count is a combination, with total coefficient mass $O(\log^{16} p)$, of the sums

$$T(\mathbf{h},S') = \sum_{P\in X_S(\mathbb{F}_p)} \chi\Big(\prod_{s\in S'}\Delta_s(u)\Big)\, e\Big(\tfrac1p(h_u u + h_v v(u) + \textstyle\sum_m h_m\alpha_m u + \displaystyle\sum_{s\in S}(h_s t_s + h'_s(-\alpha_{m(s)}u - t_s)))\Big), \qquad S'\subset\bar S,\ \|\mathbf{h}\|_\infty \le H.$$

(3) Main terms. If $S'=\emptyset$ and the additive form is identically zero on X_S, then $T=|X_S(\mathbb{F}_p)| = 2^{|S|}p + O(\sqrt p)$. The additive form equals $h_v c_1 u^3 + (h_u + h_v a + \sum_m h_m\alpha_m - \sum_s h'_s\alpha_{m(s)})u + \sum_s (h_s - h'_s)t_s$; it vanishes identically iff $h_v\equiv 0$ (as $c_1\not\equiv 0$), $h_s = h'_s$ for all $s\in S$ (by the linear independence of the $\sqrt{\Delta_s}$: $\sum_s(h_s-h'_s)t_s = \frac12\sum_s(h_s-h'_s)\sqrt{\Delta_s} +$ a polynomial in u), and the coefficient of u is $\equiv 0 \pmod p$. These frequency vectors are exactly the annihilators of the linear relations of the local law μ (the residues $\alpha_m u$ are fixed multiples of u, and the second root is $-\alpha_m u - t_s$); summing their Fourier coefficients gives $p\,\mu(F)$ up to the truncation error $O(p/H\cdot\log^{16}p) = O(\sqrt p\log^{16}p)$, and the independence of the memberships and of the roots in μ is the statement that no other frequency vector produces a main term.

(4) Error terms. If $S'\neq\emptyset$, the Kummer sheaf of $\chi(\prod_{S'}\Delta_s)$ pulled back to X_S is geometrically non-trivial by (1), and stays so after tensoring with any Artin–Schreier sheaf (coprime orders); by the Bombieri–Perel'muter bound [8, 10] $|T(\mathbf{h},S')|\le C(i,j)\sqrt p$ uniformly in $\mathbf{h}$. If $S'=\emptyset$ and the additive form is not identically zero on X_S, it is a non-constant function of bounded degree on the absolutely irreducible curve X_S (not of the form $g^p - g + c$ for p large), and Bombieri's bound [3] gives $|T|\le C(i,j)\sqrt p$. Multiplying by the coefficient masses gives the error $O_{i,j}(\sqrt p\log^{16}p)$. $\square$

Remark 60. The same proof applies to a pair of lines of two families sharing a point (the configurations of the second moment): the second line's parameter is a fixed rational multiple of u, its residues are again fixed multiples of u, and the (up to) ten partner conics of the five residues give a fibre product of the same kind; residues that coincide up to sign between the two lines only reduce the number of independent discriminants. The exceptional primes (dividing a numerator of $\alpha_m^2 \pm \alpha_{m'}^2$, or c_1, or a resultant of two discriminants) are finitely many for each configuration.

Proposition 61 (the counts)**.** *As $p \to \infty$, uniformly in the box: $\#\{$good lines of family $(1,3)\} = \gamma_1(x_0/p)\,p + O(\sqrt{p}\log^{16} p)$ and $\#\{$pairs of good lines of family $(1,3)$ with a common point$\} = \gamma_2(x_0/p, y_0/p)\,p + O(\sqrt{p}\log^{16} p)$.*

Proof. Both counts are sums, over the residues u, of Boolean combinations of interval conditions on u^*, $v(u)$, the linear forms $a_m z$, the roots of the partner conics of the three (resp. five) residues and their $-x-t$ companions, and of the six (resp. ten) membership conditions "$\Delta_{\sigma,m}(u)$ is a non-zero square". Lemma 59 (with the remark following it for pairs) says that each such count is $p \cdot \mu(F) + O(\sqrt{p}\log^{16} p)$, where μ is the local law: u^* uniform, the linear forms deterministic multiples, v^* (i.e. w') uniform and independent, the memberships fair independent coins, the roots independent uniform. Integrating the resulting product of indicators over the partner data gives exactly the U-products in (13)–(14). $\square$

Numerical values. The functions U, γ_1, γ_2 were implemented (`slack/g39_agents/`) and evaluated by quadrature on grids of 600×600 points in (s, w'); the line count itself (all U replaced by 1) reproduces the exact number of lines of family $(1,3)$ at $p = 12809$ to 0.1% ($1.733p$ vs $1.731p$ for the box $(0,0)$; $1.800p$ vs $1.801p$ for $x_0 = -(p-1)/2$). One finds that γ_1 depends on x_0/p only,

$$\gamma_1(x_0/p) \in [0.1287,\ 0.1362] \qquad (\text{minimum at } x_0/p = 0,\ \text{maximum at } x_0/p = \tfrac{1}{2}),$$

(evaluated on the grid $x_0/p,\ y_0/p \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$) and $\gamma_2 \approx 0.028$–0.031 at the three box positions of Table 2; the exact counts at $p = 3203$ and $p = 12809$ agree with these values within the expected $O(p^{-1/2})$ fluctuations. Hence, at every sampled box position,

$$\gamma_1 - \gamma_2 \ \geq\ 0.098 \ >\ \tfrac{1}{12} = 0.0833\ldots \tag{15}$$

(the constants of the table are the outcome of a numerical quadrature of explicit piecewise-polynomial integrands, not of a rigorous enclosure; the margin in (15) is more than ten times the quadrature error and the finite-p discrepancies; γ_1, γ_2 are continuous in the box position, their sampled variations are 0.0075 and 0.003, and even a safety margin of 0.005 on γ_2 for unsampled positions leaves $\gamma_1 - \gamma_2 \geq 0.093 > \frac{1}{12}$, i.e. the constant 2.657 in place of 2.652 in Theorem 62).

p, box	lines/p	good/p	γ_1	pairs/p	γ_2	(good$-$pairs)/p	threshold
3203, $(0,0)$	1.730	0.1318	0.1287	0.0212	0.028	0.1105	0.0895
3203, $(700,-2100)$	1.802	0.1339	0.1363	0.0253	0.030	0.1086	0.0720
12809, $(0,0)$	1.731	0.1313	0.1287	0.0275	0.028	0.1038	0.0847
12809, $(-6404,0)$	1.801	0.1365	0.1361	0.0315	0.030	0.1050	0.0830

Table 2: Family $(1,3)$: exact counts (`slack/cube_matching_fast.py`) versus the local-law constants; "threshold" $= (2L_4 + |S_0| - \frac{8}{3}p)/p$ is the matching size the finite-p certificate needs to reach $\frac{4}{3}N$ (asymptotically $\frac{1}{12}$).

Theorem 62 (permutation cubics, strong form; numerically assisted)**.** *Let $p \equiv 2 \pmod 3$, $a \in \mathbb{F}_p^*$, $f(x) = ax^3$ and B any $2p \times 2p$ box. Then*

$$\alpha(f,B) \ \leq\ \Big(\tfrac{11}{4} - \gamma_1(x_0/p) + \gamma_2(x_0/p, y_0/p) + o(1)\Big)p \ \leq\ (2.652 + o(1))\,p = (1.326 + o(1))\,N,$$

in particular $\alpha(f,B) < (\frac{4}{3} - 0.007)N$ for p large.

Proof. Lemma 58 and Proposition 61 give a set M of $(\gamma_1 - \gamma_2 - o(1))p$ pairwise disjoint good lines of family $(1,3)$; Lemma 57 and (15) give the bound. $\square$

Corollary 63 (Conjecture G3)**.** *For every prime p, every cubic $f \in \mathbb{F}_p[x]$ and every $2p \times 2p$ box B, $\alpha(f, B) \le (\frac{4}{3} + o(1)) \cdot 2p$; for the permutation cubics the constant is $\le 1.326 < \frac{4}{3}$.*

Proof. Non-permutation cubics: Lemma 46 ($|f(\mathbb{F}_p)| \le (\frac{2}{3} + o(1))p$, equality for the S_3 cubics); permutation cubics: Theorem 62. □

Remark 64. (a) The greedy matching over the good lines of *all* families gives $\approx 1.18\,N$ at $p \le 797$ and ≈ 1.28–$1.30\,N$ with the families of span ≤ 6–8 (log `cube_matching_fast.txt`); the LP over all lines is $\approx 1.10\,N$. The single family $(1, 3)$ with the crude bound $|G| - |\Pi|$ is what one can currently prove with a finite computation; the constant 2.652 is not optimised. (b) The same certificate applies to the permutation polynomials of Section 9 (three-point lines exist for every degree ≥ 3); the family structure of Lemma 59 would have to be redone for each degree. (c) A rigorous enclosure of $\gamma_1 - \gamma_2$ (interval arithmetic on the piecewise-polynomial integrands) would make Theorem 62 independent of floating-point quadrature; the margin 0.015 in (15) leaves ample room.

11 Computational observations beyond one hyperbola

Nothing in this section is a theorem: it records finite computations (all witnesses verified by an independent checker of all triples; "exact" means that a solver proved optimality — SAT: unsatisfiability at the next size, MIP/CP-SAT: zero gap — and "solver-proved" upper bounds carry no independent certificate) and the picture they suggest.

Two hyperbolae in the HJSW box. For $\mathcal{P}_k = (H(1, p) \cup H(k, p)) \cap G(p)$ ($8(p-1)$ candidates; trivial bound $4(p - 1)$, Section 6) the maximum lawful subsets are:

p	$3(p-1)$	k	max	gain	LP relax.	method / all k
11	30	3	35	+5	36.00	MIP, exact; over all k: 32–35 ($k = -1$: 32)
13	36	2	41	+5	47.54	SAT (two encodings), MIP, exact; all k: 38–41 ($k = -1$: 40)
17	48	−1	54	+6	60.15	MIP, exact; all k: 49–54
19	54	−1	59	+5	63.64	MIP, exact; $k = 2, 3$: 57
23	66	−1	70–74	$\ge +4$	—	CP-SAT 12 h: 70 (bound 78); MIP 8 h: 70 (bound 74); $k = 2$: 68–80, $k = 3$: 69–78
29	84	−1	≥ 84	≥ 0	—	CP-SAT 8 h: 84 (bound 92); Theorem 21: ≤ 96

Observations. (i) The gain over $3(p - 1)$ is between +1 and +6 for *every* k at $p = 11, 13, 17$ and shows no growth up to $p = 19$; the exact value at $p = 23$ is not settled. As an explicit falsifiable conjecture: $\alpha(\mathcal{P}_k) \le 3(p-1)+6$ for all $p \ge 11$ and all $k \notin \{0, 1\}$ (consistent with all exact values; at $p = 23$ the range 70–74 still allows +8; a certified witness with gain ≥ 7 would refute it). What the data exclude quantitatively: fitting gain $= a + c\,(p - 1)$ to the exact values gives $c \approx 0.06$ for the pair and $c \approx 0.05$ for a union of four hyperbolae, whereas a union reaching $3(p - 1)$'s trivial counterpart — or beating the HJSW construction — would need $c \approx 0.44$; so the measurements exclude the effect size in question by almost an order of magnitude, but they do *not* separate a bounded gain from a linear one of slope ≤ 0.1, and every prime for which the exact value is known is small. A cautionary companion inside our own data: the deficit between the block bound of Section 6 and the truth is $0, 8, 5, 5$ at $p = 11, 13, 17, 19$ — visually constant — although it provably grows like $0.44\,(p - 1)$, since the two ends are $3.4413\,(p - 1)$ and $3(p - 1) + O(1)$. (ii) The LP relaxation with all rich lines is far from the truth (column "LP"; $\approx 4(p-1) - \varepsilon$), because for two hyperbolae the fractional 2-factor $x \equiv \frac{1}{2}$ is almost feasible; certificates need lines with at least five candidates (Section 6). In terms of the saving below the trivial $4(p - 1)$, the natural certificates line up as follows (all for $k = -1$): a certificate local to one V-orbit saves nothing at all (Remark 40(d)); one local to a *pair* of orbits saves at most $0.2\,(p-1)$; the block decomposition saves $0.55\,(p - 1)$ (limit 0.559); the linear relaxation with all rich lines saves 0.6–$0.65\,(p - 1)$;

adding the level-2 moment constraints gains a further 0.03–0.05 $(p-1)$; and the truth is a saving of 1.0 $(p-1)$. The gap between the local methods and the truth is stable and closes under none of the natural strengthenings. The whole class of covering certificates is bounded away from 3: the LP with *all* rich lines gives 3.348 $(p-1)$ at $p=71$ and 3.406 $(p-1)$ at $p=101$ (against the exact values $\approx 3.3\,(p-1)$ at $p \leq 19$), and the level-2 moment relaxation (a positive semidefinite matrix indexed by the candidates, with the three-point lines lifted) closes only 8–17% of the gap between that LP and the truth at $p=13, 17, 19$, a fraction that decreases with p (`docs/research/certificate_hierarchy.md`). Rank-1 certificates using only rows, columns and slope-(± 1) lines give $\approx 3.45\,(p-1)$ for $k=-1$ $(19 \leq p \leq 199)$ and 3.6–3.9 $(p-1)$ for $k=2,3$ $(p \leq 199)$, and the integer optimum of that restricted system equals its LP value up to 2 ($k=-1$, $p \leq 59$; `slack/verification/ip1_km1.log`): so no local inequality of any rank on those lines can certify less; the remaining loss down to $3(p-1)+O(1)$ is carried by lines of unboundedly many slopes. The anatomy of the rank-1 optimum for $k=-1$ is explicit: the slope-(± 1) lines with at least three points fall into residue groups (all lines $x \mp y \equiv c \pmod p$) of four classes with line sizes $(4,8,4)$ [both the σ-pair of κ and the τ-pair of λ share their centres], $(3,7,5,1)$ [exactly one shares] or $(2,6,6,2)$ [neither], with densities $\to \frac{1}{12}, \frac{1}{12}, \frac{1}{6}$ per slope (each type is a polytope condition on C_0 as in Proposition 26; `slack/km1_lines_5678.py`). The optimum takes the three lines of every $(4,8,4)$ group (this is Theorem 21: saving 2 per group), the lines $(7,5,3)$ of the $(3,7,5,1)$ groups (locally certifiable saving 1 per group and slope), and fractional weights on the six-point lines, whose saving has no local certificate: forcing the 8- and 7-groups and optimising the rest reproduces the LP value for $p \leq 113$ (`slack/block_cover2.py`, `slack/verification/lp1_types_p113_137_199.txt`). A block lemma for the $(3,7,5,1)$ groups would therefore improve Theorem 21 to about 3.6 $(p-1)$; the LP constant ≈ 3.45 is not reachable by local arguments, and neither constant is the truth. (iii) Structure of the exact maxima: they are balanced between the two hyperbolae ($23+18$ at $p=13$, $30+29$ at $p=19$), use one or two copies of almost every class, and the two copies of a class are usually in one column or one row (at $p=19$ the optimum consists of 23 such vertical pairs and 13 single copies; at $p=17$ some classes use three copies). For $k=-1$ the points fall into $(p-1)/2$ "orbit blocks" (the 16 points of $\{\kappa, \nu\kappa, R\kappa, R'\kappa\}$ form a full 4×4 grid on its rows and columns), the trivial bound is 8 per block, and the optima have ≈ 6 per block on average with the deficits spread over most blocks; pairs and triples of blocks essentially never conflict, so the loss is a global effect. Restricted models (only vertical/horizontal copy pairs; or y-periodic sets) reach $3(p-1)+O(1)$ but not the optimum.

The vertical-pair model. For $k=-1$ consider the sets $S(r)$, $r \in \{0,1\}^{p-1}$, in which the class $\kappa_a = (a, 1/a)$ contributes both lifts in the column $X(a)+r_a p$ and the class $(a, -1/a)$ both lifts in the other column $X(a)+(1-r_a)p$; every $S(r)$ has $4(p-1)$ points, two in every row and column. Let $T(r)$ be the number of collinear triples of $S(r)$. Three facts are proved: (i) with $\varepsilon_a = (-1)^{r_a}$, T is a quadratic form $T = E + \varepsilon^{\mathsf{T}} C \varepsilon$ ($S(1-r) = R'S(r)$ and R' is a symmetry of $\mathcal{P}_{-1}$, so the odd-degree Walsh coefficients vanish); here $8E$ is the number of collinear triples of $\mathcal{P}_{-1}$ with three distinct classes, and C is an explicit symmetric matrix with entries in $\frac{1}{8}\mathbb{Z}$ and zero diagonal. (ii) Every such collinear triple is a sub-triple of a symmetric quadruple q, $q+s_1D$, $q+s_2D$, $q+(s_1+s_2)D$ (D an integer direction, $uv \equiv 2e/(s_1 s_2)$, $e = \pm 1$) whose points lie alternately on the two hyperbolae, and the family (s_1, s_2) occurs iff $s_1^2 + s_2^2$ is a square modulo p (the product xy along a line is a quadratic in the parameter, symmetric about the midpoint of two equal-sign points). (iii) Consequently $E \geq c\,(p-1)\log\log p$ for all large p: for coprime (s_1, s_2) with $s_1 + s_2 \leq p^{1/4-\epsilon}$ and $\chi(s_1^2+s_2^2) = 1$ the Kloosterman count of directions $|u|, |v| \leq p/(s_1+s_2)$ with $uv \equiv 2e/(s_1 s_2)$ gives $\gg p/(s_1+s_2)^2$ box-lifted triples, and squares modulo p are never absent from a dyadic scale of Gaussian norms (if all norms in $[M, 2M)$ were non-residues, their pairwise products, of norm in $[M^2, 4M^2)$, would be residues, and the multiplicative energy of Gaussian integers, $\ll M^2 \log M$, makes them $\gg M^2/\log M$ in number). Hence $\min_r T(r) \geq E + (p-1)\lambda_{\min}(C)$, and numerically this spectral bound is $\geq 1.9\,(p-1)$ for $11 \leq p \leq 19$ (true minima $24, 28, 42, 48$; 80 at $p = 23$ and 136 at $p = 29$ against spectral bounds 64 and 129), $\geq 3.6\,(p-1)$ for $41 \leq p \leq 71$, and 8.3, 8.5, 10.0, 14.5, 15.3 $(p-1)$ at $p = 199, 499, 997, 1999, 4001$ $(E/(p-1) = 17.2, \ldots, 28.1$; `slack/vp_lines.c`, `vp_circulant.py`): every orientation carries many collinear triples, with a certificate that is not a union bound. What is missing for a theorem is a Wigner-type edge estimate $\lambda_{\min}(C) \geq -K\sqrt{\operatorname{tr} C^2/(p-1)}$ (numerically $\lambda_{\min} = -(1.9 \ldots 2.1)\sqrt{\operatorname{tr} C^2/(p-1)}$ for all $p \leq 4001$ computed, $\sim -\sqrt{\log p}$, semicircle-like normalised moments; $\operatorname{tr} C^2/(p-1) = O(E/(p-1))$ is proved), which would give $\min_r T \geq (1-o(1))E$; the trace method reduces it to a power-saving cancellation of signed sums over closed chains of such quadruples — an arithmetic statement we could not settle (`docs/research/spectral_programme.md`, `model_theorem_conditional.md`, `slack/vp_*.py`).

Other unions. A parabola $y = x^2 + k$ together with $H(1)$ in the same box: exact maxima $34, 39, 51, 58$ for $p = 11, 13, 17, 19$ (+4, +3, +3, +4; MIP). Three hyperbolae $H(1) \cup H(2) \cup H(3)$ ($12(p-1)$ candidates): 37 at $p = 11$ (exact, MIP; +7), between 43 and 48 at $p = 13$ and between 56 and 63 at $p = 17$ (CP-SAT 4 h, not closed). Two hyperbolae with different moduli, $H(1,p) \cup H(1,q)$ in the $2p \times 2p$ box (each with its own HJSW shift): the exact maxima (MIP) exceed $3(p-1)$ by 1, 2 or 3 in all eleven pairs computed, $7 \leq p < q \leq 29$ (e.g. 32 vs 30 for $(11, 13)$, 68 vs 66 for $(23, 29)$), and the number of collinear triples with points on both hyperbolae grows like $N \log N$ in the range computed ($N \leq 586$; ratio to $N \ln N$ near 11), as for random sets of the same density. Larger boxes for one hyperbola: Remark 16.

A heuristic. Fix the HJSW set S_2 of $H(1,p)$ and call a point q of $H(k,p) \cap G(p)$ blocked if some line through q carries two points of S_2. The average number of blocking pairs per point grows like $\log p$ (about 3.1 at $p = 11$ and 7 at $p = 101$ for $k \in \{2, 3, -1\}$), and the number of unblocked points is 0–2 for every $p \leq 101$ tested (for $k = 2$ at $p = 5$ and a few other (p,k) it is 4); this depends on the chosen maximum set of $H(1)$ (over the 9^s maximum sets the number of unblocked points varies between 0 and 2 at $p = 13$, $k = 2$) and does not lower-bound the deletion cost, but it explains why the second hyperbola contributes little: freeing a point of $H(k)$ costs deletions in $H(1)$. Whether the gain is $O(1)$, $O(\log p)$ or $O(\sqrt{p})$ cannot be decided from the data (+5, +5, +6, +5); Theorem 21 shows that for $k = -1$ the gain over $3(p-1)$ is at most $(\frac{2}{3} + o(1))(p-1)$, and no more is proved. (Definitions and data: `slack/`, `docs/research/slack_experiments.md`, `docs/research/pair_bound_notes.md`.)

12 Independent computer verification

The theorem has a short self-contained proof; the following check is an independent sanity test of the bookkeeping (the classification of the rich lines, the exceptional classes and the count of the maximum sets), performed by a program that trusts no part of the proof. The script `slack/hjsw_window_check.py` of the repository recomputes, from the point set alone: all lines with at least three points of $\mathcal{P}$ (found by brute force over all pairs of points; all have slope ± 1, three or four points, and lie inside one V-orbit); the identity $2|\mathcal{L}| + Z = 3(p-1)$; the exact maximum by exhaustive search inside each orbit (at most 2^{16} subsets), assembling a maximum witness which is passed through the independent verifier `saturation.the_law` in grid coordinates; and the number of maximum sets, compared with 9^s. All checks pass for every prime $3 \leq p \leq 101$ (all $c \in \mathbb{F}_p^*$ for $p \leq 31$; $c \in \{1, 2, 3, -1, -2, (p+1)/2\}$ for larger p); log: `slack/verification/hjsw_window_check_all.log`. For Theorem 3 the script `slack/hjsw_boxes_check.py` (log `slack/verification/hjsw_boxes_check.log`) checks, for every box $(x_0, y_0 \bmod p)$ and every c: the formula $Z + 2|\mathcal{L}| = 2(p-1) + 2s + 2\sum_O (e_O + f_O)$ against brute-force line enumeration (random boxes, $p \leq 19$); $Z + 2|\mathcal{L}| \leq 3(p-1)$ for all primes $p \leq 61$ (with equality for some boxes); the orbit lemma orbit by orbit ($(e_O, f_O) \in \{(0,0), (0,2), (1,1), (2,0)\}$

only) for all $p \leq 31$; every identity used in the proof of Lemma 10 ($\theta = c_a + c_b - k$, $\eta = c_b - c_a$, the four split conditions, parity, and "$F_1 \Rightarrow \neg(E_1 \wedge E_2)$") on random instances up to $p = 199$; the maxima and multiplicities of Lemma 11 orbit by orbit (exhaustive search over the 2^{16} subsets) for all $p \leq 13$, all c and all boxes; and the formula $12n_2 + 10n_1 + 8n_0 + 6s$ against MIP optima on random boxes with $p \leq 31$. The exact maxima of Table 1 are MIP optima (HiGHS) with witnesses through the law. The values agree with the exact maxima obtained earlier in the project by SAT solving (kissat/CaDiCaL, two encodings, witnesses through the law): $18, 30, 36, 48, 54, 66$ for $p = 7, 11, 13, 17, 19, 23$. Two external LLM-assisted audits (ChatGPT deep-research reports commissioned by the author of record on v0.1 and on v0.5 of this note; kept in `docs/reviews/`) re-derived the point sets and the rich lines with clean-room implementations and confirmed $3(p-1)$, 9^s, the LP value and the no-four values for all c with $p \leq 31$ (samples for larger p), all box patterns and gadget counts through $p = 13$, and the shifted-box table for $p = 11$; the second audit also identified a gap in the earlier presentation of Lemma 11 (the $(1,1)$ count), which is now proved above.

For Sections 6–10: the block cover of Theorem 21 is checked point by point by `slack/km1_theorem_check.py` and `slack/block_cover_km1.py` ($p < 320$); the seven-point blocks (Lemma 28, Lemma 29) by `slack/t221_agents/orbit_cover.py` ($p \leq 500$) and `slack/t221/neighbours_check.py`, the orbit decomposition and the finite orbit verification behind Proposition 19 (`slack/t221/orbit_decomp_one.py`, `slack/t221/orbit_stability.py`, all orbits of all $p \leq 31$); the block decomposition (Theorem 35), the independence of the saving from anything but the signature and the table of block optima by `slack/t221/fast_blocks.py` and `slack/sig_savings.py` (two independent implementations, 28 common signatures, no discrepancy) and the convolution with the descent densities by `slack/constant_km1.py` and `slack/t221/constant.py`; the clean counts and the local law of Proposition 32 by `slack/t221_agents/clean_counts.py` ($p \leq 3000$ and to 10^5) and `slack/t221/model_mc.py`; the potential cover of Theorem 41 and its value $4(p-1) - (2G_8 - 2\sum_c |A_c - B_c|)/R$ are recomputed and compared with the class-level linear program by `slack/exact_c.py` and `slack/cyc_model.py` (the LP with the uniform line weight $1/R$ has exactly this value in every case tried, and the class-level LP agrees with the point-level LP); the arc-imbalance sums are tabulated by `slack/arc_imbalance.py` and the mixed sums along $C_0(k)$ by `slack/mixed_sums_curve.py` ($p \leq 4001$). For Section 8: the line profiles and lift rules were verified by brute force for $p \in \{101, 131, 197, 293\}$, $a \in \{1, 2\}$ and three boxes (`slack/verification/g35_verify.txt`), the cost of the explicit cover by `slack/cube_cover4.py` (Remark), and the local law (same/split $= 1/2$, independence of the slopes) by `slack/cube_local_law.py` and by exact counts up to $p = 100019$ (`slack/g_agents/group_types_output.txt`). For Section 9: the lift table of Lemma 50 (`slack/g_agents/profile_table_output.txt`), the root-count laws and the cross-slope independence (`slack/pp_joint_counts.py`), the exact rational constants C_k (`slack/pp_constant_exact.py`) and the $O(p)$ evaluation of the cover's cost at $p \approx 10^7$ (`slack/pp_exact_UL_np.py`); the discriminant resultants for $k = 3, 5, 7$ symbolically. For Section 10: the local law U and the constants γ_1, γ_2 (`slack/g39_agents/`, outputs `slack/verification/g39_*.txt`) against exact counts at $p = 3203$ and $p = 12809$, and an independent point-level check of the certificate W_4 plus matching (`slack/cube_w4_matching_check.py`). A checklist of what to verify and with which script is `docs/AUDIT_CHECKLIST_v1.7.md`.

Repository (public): `https://github.com/iwasborninbali/saturation`.
This version of the note is the immutable tag `hjsw-note-v1.17` (v1.0: the theorems as first verified; v1.1: Section 11 on the vertical-pair model; v1.2: the error term of Proposition 26; v1.4: Section 7, every second hyperbola; v1.5: Section 8, cubic graphs; v1.6: Section 9, permutation monomials; v1.7: Corollary 55; v1.8: verification paragraph, remark corrections; v1.9: Section 10, the strong form for cubic graphs; v1.10: the seven-point blocks, Theorem 30; v1.11: the block decomposition along runs of quadratic residues, Theorem 35; v1.11.1: correction of one comparison in the abstract; v1.11.2: the exact constant 17865173/5160960 with the rare signatures

tabulated; v1.12: stability of one hyperbola, Proposition 19; v1.13: the standard model of a block and the constant 427101431/123863040; v1.14: calibration of the empirical readings after the external research report; v1.15: erratum in remark (b) of the blocks section — at $p = 113$ the quoted pair was the LP value, the exact block sum is 382, and the stale $K = 8$ constant 3.4616 is replaced by the $K = 9$ value 3.4482; v1.16: a ten-line abstract in place of the page-long one, whose body survives as the extended summary, the strong-form theorem labelled numerically assisted, and this version list put in order; v1.17: the two edits that v1.16's changelog claimed — the numerically-assisted label and Green's URL — were in fact absent from the text, the label edit having been applied to the wrong file, and are applied here; the claim-without-edit is recorded, not erased): file `paper/hjsw_window.tex`; scripts in `slack/`: `hjsw_window_check.py`, `hjsw_boxes_check.py`, `gadget11_check.py`, `km1_theorem_check.py`, `block_cover_km1.py`, `km1_lines.py`, `cyc_model.py`, `exact_c.py`, `arc_imbalance.py`, `t221/`, `t221_agents/`; logs in `slack/verification/`; checksums and the list of tags in `docs/RELEASES.md`.

p	c	$\lvert\mathcal{P}\rvert$	$\lvert\mathcal{L}\rvert$	Z	max	#max sets
7	1	24	8	2	18	9
11	1	40	14	2	30	9
13	1	48	16	4	36	81
13	2	48	18	0	36	1
17	1	64	22	4	48	81
19	1	72	26	2	54	9
23	1	88	32	2	66	9
101	1	400	148	4	300	81

Table 3: Sample values from the verification log ($\lvert\mathcal{L}\rvert = \frac{3}{2}(p-1) - s$, $Z = 2s$).

AI/tool-use disclosure and authorship statement. The theorems, their proofs, the verification programs and the text of this note were produced by autonomous large-language-model agents (Anthropic Claude, model *Claude Fable 5*, August 2026) working in a shared repository: one agent found and proved Theorems 1 and 3 and wrote the first version of the note; Theorem 21 was found independently, on the same day, by two of the agents; Theorem 41 (the potential cover) is due to the second agent and Lemma 43 to the first; the seven-point blocks (Theorem 30, Proposition 32) are due to the second agent, Proposition 31 to the first; the block decomposition (Theorem 35) is due to the second agent and the descent law (Lemma 36, Proposition 37) to the first; Sections 8, 9 and 10 were found and proved by the first agent and checked by the second (Lemma 59 is due to the second); the computational setup and the observations of Section 11 come from all three; the note was revised by the agents after two external LLM-assisted audits and one literature report commissioned by the author of record (the report suggested the simplification of Step 4 in the proof of Proposition 26). The author of record posed the question, directed and organised the work, commissioned the external audits, decided what to claim and publish, and takes responsibility for the content; the extent of his independent verification of the proofs is recorded in `docs/HUMAN_AUDIT.md` of the repository and will be stated precisely in any submitted version. The AI systems are not authors. All code, logs and the correspondence of the agents are public in the repository.

Appendix: what the hyperbola route cannot give

Write $n = 2p$ for the side of the grid, so that the trivial upper bound is $2n = 4p$ and the HJSW set has $3(p-1) = \frac{3}{2}n - 3$ points. The results of this note, read as statements about constructions, say the following.

(a) *One hyperbola, any $2p \times 2p$ window.* Every no-three-in-line subset of $H(c,p) \cap W$ has at most $3(p-1) = \frac{3}{2}n - 3$ points (Theorem 3), with equality exactly for the windows with $n_1 = n_0 = 0$, the HJSW window among them (Theorem 1); the maximum sets are classified and their number is 9^s in the HJSW window and $144^{n_1}1296^{n_0}9^{s_1}6^{s_2}$ in general. The barrier $\frac{3}{2}$ is therefore exact for this route, not approximate.

(b) *Two hyperbolae.* For $H(1) \cup H(-1)$ in the HJSW window, $\alpha \leq 4(p-1) - 4m_8(p) \leq (\frac{11}{3} + o(1))(p-1) < \frac{11}{6}\,n$ (Theorem 21, Proposition 26; $m_8(p) > 0$ for $19 \leq p \leq 1500$), and the block decomposition improves this to $(3.4482 + o(1))(p-1) \approx 1.724\,n$ (Corollary 39). For an arbitrary second hyperbola $H(k)$ with k of multiplicative order at least $C\sqrt{p}\log^5 p$, $\alpha \leq 4(p-1) - c\sqrt{p}/\log^5 p$ (Theorem 41) — below the trivial bound, but only barely.

(c) *Other conics and cubics.* Among conics only the shifted hyperbolae reach $3(p-1)$ in a $2p \times 2p$ window (Corollary 12); every cubic graph $Y \equiv f(X)$ has at most $(\frac{4}{3} + o(1))\,n$ points in general position (Theorem 62), and the permutation monomials x^k, $k \in \{3,5,7\}$, at most $\approx 1.40\,n$ (Theorem 54).

What this does not cover. (1) Unions of three or more hyperbolae, and pairs $H(k), H(k')$ with k of small multiplicative order in an arbitrary window: only the computations of Section 11, which suggest a gain of $O(1)$ over $3(p-1)$ but prove nothing. (2) Windows other than $2p \times 2p$: already a $(2p+1) \times 2p$ window at $p = 5$ admits $13 > 12$ points of one hyperbola, and for larger windows only MIP values are known (Table 1). (3) Composite moduli: Lemma 4 fails (for $m = 9$ the rich lines have slopes $-1, \frac{1}{2}, 1, 2$), so everything above is for prime p. (4) Points outside the algebraic set: by Remark 9(b) any improvement of the HJSW construction must use another hyperbola, a larger window or non-hyperbola points, and about the last two the theorems are silent. (5) Curves of degree at least four. The claims of (a)–(c) were re-derived by a second, independently written program for all c with $p \leq 13$ and, for Theorem 3, for all windows and all c with $p \leq 7$ (repository `lemma-atelier`, lemma 006).

References

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

[2] B. Kovács, Z. L. Nagy, D. R. Szabó, Randomised algebraic constructions for the no-$(k+1)$-in-line problem, arXiv:2508.07632 (v1 11 Aug 2025, v2 21 Jul 2026); Advances in Combinatorics 2026:7, doi:10.19086/aic.2026.7 (to appear).

[3] E. Bombieri, On exponential sums in finite fields, Amer. J. Math. 88 (1966) 71–105.

[4] H. Iwaniec and E. Kowalski, Analytic Number Theory, Amer. Math. Soc. Colloq. Publ. 53, 2004 (Chapter 11).

[5] Y. Aubry and M. Perret, A Weil theorem for singular curves, in: Arithmetic, Geometry and Coding Theory (Luminy, 1993), de Gruyter, 1996, 1–7.

[6] H. L. Montgomery, Ten Lectures on the Interface between Analytic Number Theory and Harmonic Analysis, CBMS Regional Conf. Ser. in Math. 84, Amer. Math. Soc., 1994 (Chapter 1).

[7] J. D. Vaaler, Some extremal functions in Fourier analysis, Bull. Amer. Math. Soc. 12 (1985) 183–216.

[8] G. I. Perel'muter, Estimate of a sum along an algebraic curve, Mat. Zametki 5 (1969) 373–380.

[9] M. Perret, Multiplicative character sums and Kummer coverings, Acta Arith. 59 (1991) 279–290.

[10] F. N. Castro and C. J. Moreno, Mixed exponential sums over finite fields, Proc. Amer. Math. Soc. 128 (2000) 2529–2537.

[11] L. E. Dickson, The analytic representation of substitutions on a power of a prime number of letters with a discussion of the linear group, Ann. of Math. 11 (1897) 65–120.

[12] R. Lidl and H. Niederreiter, Finite Fields, 2nd ed., Encyclopedia of Mathematics and Its Applications 20, Cambridge Univ. Press, 1997 (Chapter 7).

[13] D. Wan, A p-adic lifting lemma and its applications to permutation polynomials, in: Finite Fields, Coding Theory, and Advances in Communications and Computing (Las Vegas 1991), Lecture Notes in Pure and Appl. Math. 141, Dekker, 1993, 209–216.

[14] H. Niederreiter and K. H. Robinson, Complete mappings of finite fields, J. Austral. Math. Soc. Ser. A 33 (1982) 197–212.

[15] I. E. Shparlinski, Modular hyperbolas, Jpn. J. Math. 7 (2012) 235–294; arXiv:1103.2879.

[16] M. R. Khan, R. Magner, S. Senger, A. Winterhof, Two combinatorial geometric problems involving modular hyperbolas, arXiv:1304.6943.

[17] V. Froese, I. Kanj, A. Nichterlein, R. Niedermeier, Finding points in general position, Internat. J. Comput. Geom. Appl. 27 (2017) 277–296; arXiv:1508.01097.

[18] M. S. Payne, D. R. Wood, On the general position subset selection problem, SIAM J. Discrete Math. 27 (2013) 1727–1733.

[19] D. T. Nagy, Z. L. Nagy, R. Woodroofe, The extensible no-three-in-line problem, European J. Combin. 114 (2023) 103796.

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

[21] C. Y. Ku, K. B. Wong, On no-three-in-line problem on m-dimensional torus, Graphs Combin. 34 (2018) 355–364.

[22] A. Ghosal, R. Goenka, A. Grebennikov, P. Keevash, M. Kwan, H. T. Pham, No-$(k+1)$-in-line problem for $k \geq 3$, arXiv:2607.05255 (2026).

[23] B. Green, 100 open problems, manuscript (version of December 2025), Problem 72, `https://people.maths.ox.ac.uk/greenbj/papers/open-problems.pdf`.

[24] A. Flammenkamp, The no-three-in-line problem, database and tables, `https://wwwhomes.uni-bielefeld.de/achim/no3in/readme.html`.

[25] T. Prellberg, Constraint satisfaction programming for the no-three-in-line problem, arXiv:2602.07751 (2026).