A fixed-proportion upper bound for arc colourings of square grids

Draft awaiting external review, 4 November 2026

Abstract

We present the proof manuscript of a uniform upper bound for ordinary arc colourings of integer square grids. The argument gives $\limsup_{n\to\infty} \chi_A(G_n)/n \le 494/499$, and hence $\chi_A(G_n) \le \lfloor 0.99n \rfloor$ for all sufficiently large integers n. The construction combines a geometric exchange of quadratic-curve classes, uniform estimates for a parameter-indexed residual set, and an ordered random-arc process converted into an ordinary colouring. Three appendices give the counting, probability and absorption arguments. Equation numbers are retained from the Chinese companion for direct comparison.

Contents

1 Introduction and the main theorem **2**

A The candidate set and lattice estimates **6**

B From random arcs to ordinary colourings **11**

C Geometric exchange, absorption and transfer **19**

1 Introduction and the main theorem

Let

$$G_n = \{0, 1, \ldots, n-1\}^2.$$

An *arc* is a set of points containing no three distinct collinear points. Throughout this paper, collinearity refers to lines in the real plane. An arc colouring is a partition into arcs, and $\chi_A(V)$ denotes the least number of parts in such a partition.

Theorem 1. *There is an absolute constant N_0 such that, for every integer $n \geq N_0$,*

$$\chi_A(G_n) \leq \lfloor 0.99n \rfloor. \tag{1.1}$$

More precisely,

$$\limsup_{n\to\infty} \frac{\chi_A(G_n)}{n} \leq \frac{494}{499}. \tag{1.2}$$

We prove the theorem by recolouring part of a quadratic-curve construction. Removing approximately $\delta q/2$ old colours leaves a set whose points are indexed by pairs of selected parameters. The resulting recolouring cost is quadratic in δ, whereas the saving is linear in δ. To obtain the stated numerical bound, we must keep the coefficient of the recolouring cost explicit.

The proof has three components. A geometric exchange reduces the problem to colouring a specific set P_L. A random-arc construction and a concentration argument give an ordinary colouring of P_L with at most $(56/27 + o(1))\delta^2 q$ colours. Finally, a two-prime construction transfers the saving to every sufficiently large integer side length. The geometric details are in Appendix C, the estimates for P_L in Appendix A, and the ordinary-colouring argument in Appendix B.

All logarithms are natural. Parameters described as fixed are chosen before the relevant grid size tends to infinity. Error estimates involving a root or a line are uniform in that root or line. This manuscript gives an existence proof with an unspecified threshold. The ordered random selection and potential-function method follow the approach of Pham, Le Quang-Ham and Xue [2]; their method is credited separately from the lattice estimates and its application here. This is a research manuscript prepared for further mathematical scrutiny.

1.1 The residual set

Let $q \equiv 3 \pmod 4$ be prime, and put

$$h_q = (q-1)/2, \qquad U_q = [-h_q, h_q]^2 \cap \mathbb{Z}^2.$$

Write $\mathcal{Q}$ for the nonzero squares in $\mathbb{F}_q$, and χ for the quadratic character, extended by $\chi(0) = 0$. For $a \in \mathbb{F}_q$, let $[a]$ be its unique representative in $[-h_q, h_q]$.

Given $L \subseteq \mathcal{Q}$, let P_L contain the following two points for each unordered pair of distinct parameters $b, c \in L$:

$$\left(\left[\frac{b+c}{2}\right], [\sqrt{bc}]\right), \qquad \left(\left[\frac{b+c}{2}\right], [-\sqrt{bc}]\right). \tag{3.1}$$

Both square roots are used, so the definition is independent of the choice of a square root. A point (x, y) recovers its unordered parameter pair as the roots of

$$X^2 - 2xX + y^2 = 0,$$

Thus distinct pairs give distinct points, and

$$|P_L| = |L|(|L| - 1). \tag{3.2}$$

Let P_* denote the set obtained by taking $L = \mathcal{Q}$. We shall prove in Appendix A that

$$P_* = \{(x, y) \in U_q : y \neq 0,\ \chi(2(x + y)) = \chi(2(x - y)) = 1\}. \tag{3.3}$$

The geometric construction also produces a set R_L of points to be removed from the old colours. Its role is illustrated in Figure 1. In the main grid, a centre z may lie in two forbidden triples, one on each diagonal direction. Removing the centre meets both triples. If the centre is retained, removing the two right-hand corner points meets the two triples separately. All removed points must subsequently be recoloured.

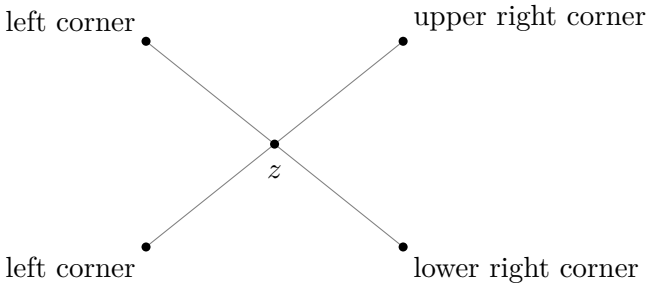


Figure 1: The two triples through a centre, drawn schematically. Their exact coordinates are given in (C.15) and (C.16).

We first implement this choice in three adjacent $q \times q$ windows. The two right-corner maps are

$$F_q^+(u, v) = (q - v, q - u), \qquad F_q^-(u, v) = (q + v, -q + u).$$

The set R_L satisfies the pointwise hitting property

$$z \in R_L \quad \text{or} \quad F_q^+(z), F_q^-(z) \in R_L. \tag{3.4}$$

The second alternative requires both images. Appendix C proves the exchange and the complete colour budget

$$\chi_A(R_L) \leq q - 1 - |L| + \chi_A(P_L) + q^{3/4+o(1)}. \tag{3.5}$$

The last term pays for the auxiliary lines and all old points displaced while absorbing them.

1.2 A suitable choice of parameters

Fix $0 < \delta \leq 1/2$, and include each parameter of $\mathcal{Q}$ in L independently with probability δ. The points of P_L share parameters, so their dependencies must be retained in the argument.

For a finite point set V, a *block* is the intersection with a real line containing at least three points of V. Distinct blocks meet in at most one point. Define the degree at a root v by

$$d_V(v) = \sum_{B \ni v} \binom{|B| - 1}{2}. \tag{4.1}$$

This counts the collinear triples containing v. Call the blocks of directions $(1, 1)$ and $(1, -1)$ *special*; all other blocks are *ordinary*. A point belongs to at most two special blocks.

Lemma 2 (Parameter selection)**.** *For fixed δ, every sufficiently large prime $q \equiv 3 \pmod 4$ admits a set $L \subseteq \mathcal{Q}$ with*

$$\begin{aligned} |L| &= (\delta/2 + o(1))q, \\ |B| &\leq (\delta^2/4 + o(1))q \qquad \textit{for ordinary blocks}, \\ |B| &\leq (\delta^2/2 + o(1))q \qquad \textit{for special blocks}, \\ \max_{v \in P_L} d_{P_L}(v) &\leq (5/64 + o(1))\delta^4 q^2 \log q. \end{aligned} \tag{4.2}$$

Appendix A proves the lemma. The line-length factor δ^2 comes from the two parameters needed for a point. Conditional on a root being present, a typical triple needs four further parameters, giving the factor δ^4. Triples with shared parameters are counted separately. The coefficient $5/64$ follows from a uniform character-sum estimate for the degree of P_*.

1.3 The ordinary-colouring lemma

Lemma 3 (Colouring at a prescribed scale)**.** *Fix $K, A > 0$. Suppose V lies in an integer square of side length $q \le Ks$, $|V| \le s^A$, and*

$$|B| \le s \quad \textit{for ordinary blocks}, \qquad |B| \le 2s \quad \textit{for special blocks}, \qquad d_V(v) \le s^2 \log s. \tag{5.1}$$

Then

$$\chi_A(V) \le (200/27 + o(1))s. \tag{5.2}$$

Appendix B constructs a random arc with the same inclusion probability

$$\rho = 27/(200s)$$

at every point, and with two-point inclusion probability at most $s^{-2503/1500}$. We repeatedly select such an arc and assign it a new colour. A martingale argument controls both the remaining block lengths and the remaining degrees. The two-point bound is needed for the latter control.

The geometric random-arc lemma requires q/s to be bounded. We therefore run it only until the scale is a fixed small fraction ξ of its initial value. An abstract colouring lemma, which has no ambient-square restriction, colours the remainder at a cost of at most $(17 + o(1))\xi s$. Taking the stage parameters and ξ sufficiently small, and then letting s tend to infinity, gives the coefficient $200/27$.

1.4 The colour budget

Apply the parameter-selection lemma and set

$$s = \frac{7}{25}\delta^2 q. \tag{6.1}$$

The line-length hypotheses hold because

$$\frac{7}{25} > \frac{1}{4}, \qquad \frac{14}{25} > \frac{1}{2}.$$

For the degree hypothesis we have the strict margin

$$\left(\frac{7}{25}\right)^2 - \frac{5}{64} = \frac{11}{40000} > 0. \tag{6.2}$$

Since δ is fixed, $\log s / \log q \to 1$, so this margin absorbs the error in (4.2). Also $q/s = 25/(7\delta^2)$ is fixed and $|P_L| \le q^2 \le s^3$ eventually. The colouring lemma therefore gives

$$\chi_A(P_L) \le \left(\frac{200}{27}\frac{7}{25} + o(1)\right)\delta^2 q = \left(\frac{56}{27} + o(1)\right)\delta^2 q. \tag{6.3}$$

Choose

$$C = \frac{499}{240}, \qquad C - \frac{56}{27} = \frac{11}{2160} > 0,$$

and fix

$$\delta = \frac{1}{4C} = \frac{60}{499}. \tag{6.4}$$

Indeed, the saving $\delta/2 - C\delta^2$ equals

$$\frac{1}{16C} - C\left(\delta - \frac{1}{4C}\right)^2$$

and is largest at this choice of δ. For all sufficiently large relevant primes, the positive margin $C - 56/27$ gives $\chi_A(P_L) \le C\delta^2 q$. Substitution into (3.5) yields

$$\begin{aligned} \chi_A(R_L) &\le \left(1 - \frac{\delta}{2} + C\delta^2 + o(1)\right) q \\ &= \left(1 - \frac{15}{499} + o(1)\right) q = \left(\frac{484}{499} + o(1)\right) q. \end{aligned} \tag{6.5}$$

The auxiliary cost $q^{3/4+o(1)}$ is included in the final $o(q)$ term.

1.5 Transfer to all integers

Appendix C proves the following transfer statement. If, for every sufficiently large prime $q \equiv 3 \pmod 4$, a set R_L has the hitting property (3.4) and satisfies

$$\chi_A(R_L) \le (\gamma + o(1))q,$$

then

$$\limsup_{n\to\infty} \chi_A(G_n)/n \le (2+\gamma)/3. \tag{7.1}$$

To explain the scales, fix a small $\eta > 0$. For any sufficiently large integer n, choose a main prime

$$p = \frac{2n}{3-\eta} + o(n),$$

and a secondary prime

$$q \equiv 3 \pmod 4, \qquad q = \frac{n}{3-\eta} + o(n).$$

The prime number theorem and its fixed-modulus arithmetic-progression version suffice for these choices. After removing a translated copy of R_L, the main quadratic classes are arcs and cost approximately p colours. The removed set costs approximately γq colours, and the main zero-label points cost $o(n)$ extra colours. The fixed margin η accommodates the prime errors and the translation. Consequently,

$$\chi_A(G_n) \le \left(\frac{2+\gamma}{3-\eta} + o(1)\right) n.$$

Taking the upper limit first and then letting the fixed parameter η decrease to zero proves (7.1).

Using $\gamma = 484/499$ from (6.5), we obtain

$$\limsup_{n\to\infty} \frac{\chi_A(G_n)}{n} \le \frac{2 + 484/499}{3} = \frac{494}{499}.$$

Finally,

$$\frac{99}{100} - \frac{494}{499} = \frac{1}{49900} > 0.$$

Choose a fixed number strictly between $494/499$ and $99/100$. The definition of the upper limit bounds $\chi_A(G_n)/n$ by that number for every sufficiently large n. Since the colour count is an integer, (1.1) follows.

A The candidate set and lattice estimates

A.1 The character-sum input

We use the prime-field case of Corollary 1 of Lange and Winterhof [1]. If $f \in \mathbb{F}_q[X]$ has positive degree at most four and is not a constant times a square, then for an interval I of $H \le q$ consecutive integers,

$$\left|\sum_{t\in I} \chi(f(t))\right| \le CH^{1/2}q^{1/4}. \tag{A.1}$$

The constant is absolute under this degree bound. Translation of the interval can be absorbed into f, and making f monic multiplies the sum by a character value of absolute value one. For $H = q$, the complete Weil bound also suffices. We check the nonsquare condition at each use.

A.2 The character description of the candidate set

Write $b = r^2, c = t^2$, where b, c are distinct nonzero squares. Choose the signs so that $y = rt$. For $x = (b + c)/2$,

$$2(x + y) = (r + t)^2, \qquad 2(x - y) = (r - t)^2.$$

Both squares are nonzero because $b \ne c$, and $y \ne 0$. Thus every point of the parameter construction satisfies (3.3).

Conversely, suppose $y \ne 0$ and choose nonzero R, S with

$$R^2 = 2(x + y), \qquad S^2 = 2(x - y).$$

Set $r = (R + S)/2, t = (R - S)/2$. Then

$$rt = y \ne 0, \quad r^2 + t^2 = 2x, \quad r^2 - t^2 = RS \ne 0.$$

Hence r^2, t^2 are distinct nonzero square parameters producing the point. This proves (3.3). Put

$$F(x, y) = \frac{(1 + \chi(2(x + y)))(1 + \chi(2(x - y)))}{4}.$$

We have $0 \le \mathbf{1}_{P_*} \le F \le 1$ throughout the window. At a zero of a character argument, F need not be the indicator; only the upper bound is used.

A.3 Degrees in the full grid

Represent an unoriented primitive integer direction by (a, b) with a fixed overall sign, and call $d = \max(|a|, |b|)$ its height. There are $4\varphi(d)$ directions of height d, including $d = 1$. A line of this direction has at most $1 + \lfloor (q - 1)/d \rfloor$ points in the window. For a fixed root, the other two points can therefore be chosen in at most $q^2/(2d^2)$ ways. It follows that

$$d_{U_q}(z) \le 2q^2 \sum_{d<q} \frac{\varphi(d)}{d^2} = \left(\frac{2}{\zeta(2)} + o(1)\right) q^2 \log q. \tag{A.2}$$

For completeness, $\varphi(d)/d = \sum_{a|d} \mu(a)/a$ gives

$$\sum_{a\le x} \frac{\mu(a)}{a^2} H_{\lfloor x/a \rfloor} = \frac{\log x}{\zeta(2)} + O(1).$$

using $H_m = \log m + O(1)$ and the convergence of $\sum_a (\log a)/a^2$.

A.4 A uniform factor of one sixteenth

We prove, uniformly over roots,

$$d_{P_*}(z) \le \frac{1}{16} d_{U_q}(z) + O(q^2 \log\log q), \qquad z \in P_*. \tag{A.3}$$

Fix $z = (x_0, y_0) \in P_*$, and write $\alpha = 2(x_0 + y_0)$ and $\beta = 2(x_0 - y_0)$. Both are nonzero squares. Split the directions at

$$D_- = \frac{\sqrt{q}}{(\log q)^4}, \qquad D_+ = \sqrt{q}(\log q)^4.$$

A.4.1 Small heights

For height $d \le D_-$, write the points on the root line as $z + t(a, b)$, where t ranges over an interval of length $m \le 2q/d$. The character arguments are

$$\alpha + 2t(a + b), \qquad \beta + 2t(a - b).$$

When both slopes are nonzero and the roots are distinct, expansion of F and (A.1) give

$$m_* \le m/4 + O(m^{1/2} q^{1/4}).$$

The number of neighbour pairs on this line is consequently at most

$$\frac{1}{16}\binom{m-1}{2} + O(m^{3/2} q^{1/4} + m q^{1/2} + m). \tag{A.4}$$

The exceptional directions satisfy $a + b = 0$, $a - b = 0$, or $\alpha(a - b) = \beta(a + b) \pmod q$. Each of the first two specifies one direction. The last specifies one projective direction over the field. Two distinct primitive directions of height at most D_- cannot reduce to it, since their determinant would be a multiple of q with absolute value at most $2D_-^2 < q$. Thus there are at most three exceptional lines, contributing $O(q^2)$.

Summing the other errors using $4\varphi(d) \le 4d$ gives

$$O(q^{7/4}\sqrt{D_-} + q^{3/2} D_- + q D_- + q^2) = O(q^2).$$

A.4.2 Intermediate heights

For $D_- < d < D_+$ we use the full-grid estimate. Its entire contribution is

$$O\left(q^2 \sum_{D_- < d < D_+} \frac{1}{d}\right) = O(q^2 \log\log q).$$

which is lower order than $q^2 \log q$.

A.4.3 Large heights

For $d \ge D_+$ write the other two points as

$$z + j(a, b), \quad z + k(a, b), \qquad j, k \ne 0, \quad j \ne k.$$

Let $M = \max(|j|, |k|)$. Then $M \le J := \lfloor (q-1)/D_+ \rfloor$. Normalize the direction by $a > 0$; the vertical direction belongs to the small-height range.

Fix (j, k). The two window conditions restrict a and b separately to integer intervals, with coordinates of absolute value at most $A = (q-1)/M$. Removing the region of height less than

D_+ leaves at most three disjoint rectangles. Their side lengths are $O(A)$, and each coordinate interval contains at most q integers.

Expand the product of the two endpoint values of F. Its constant term is $1/16$; each of the other fifteen terms is the character of a nonempty subproduct of

$$\alpha + 2j(a+b), \quad \beta + 2j(a-b), \quad \alpha + 2k(a+b), \quad \beta + 2k(a-b). \tag{A.5}$$

We retain primitivity by the identity

$$\mathbf{1}_{\gcd(a,b)=1} = \sum_{m|a,\ m|b} \mu(m),$$

Set $a = mu, b = mv$. Since $a > 0$, only $1 \le m \le A$ occur. Fixing v, the four factors are nonconstant linear polynomials in u. The two plus factors have distinct roots, as do the two minus factors. Here $\alpha, \beta \neq 0$ and $j \neq k \pmod q$, the latter following from $2J < q$.

A plus root can equal a minus root for at most four residue classes of v. Each such class occurs only a bounded number of times in the interval. In every other row all four roots are distinct, so every nonempty subproduct is nonsquare. Apply (A.1) in these rows and use the trivial bound in the exceptional rows. For each rectangle and m, every nonconstant term contributes at most

$$O(q^{1/4}(A/m)^{3/2} + A/m).$$

Summing the errors over m gives

$$O(q^{1/4}A^{3/2} + A\log(2A)).$$

The constant term retains its signed Möbius sum, and therefore counts primitive directions exactly. Since the product of the two candidate indicators is bounded by the product of the corresponding F values, the candidate count is at most $1/16$ of the full-grid count plus this error.

There are at most $8M$ ordered step pairs with maximum absolute value M. Their total error is

$$O(q^{7/4}\sqrt{J} + qJ\log q) = O(q^2).$$

Divide by two to pass from ordered to unordered neighbour pairs. Combining all three ranges proves (A.3). Together with (A.2), it gives

$$\max_z d_{P_*}(z) \le \left(\frac{1}{8\zeta(2)} + o(1)\right) q^2 \log q \le (5/64 + o(1))q^2 \log q, \tag{A.6}$$

The inequality $\zeta(2) > 8/5$ already follows from the first twenty-four terms of its defining series.

A.5 The two line-length bounds

We claim that ordinary lines meet P_* in at most $q/4 + O(q^{3/4})$ points, and special lines in at most $q/2 + O(q^{3/4})$ points. Directions of height at least four already have full-grid line length at most $q/4 + 1$. It remains to examine the finitely many directions of heights one, two and three.

If the two linear character arguments have nonzero slopes and distinct roots, (A.1) gives one quarter of the full line length, with error $O(q^{3/4})$. The ordinary directions of height one are horizontal and vertical. Coincident roots then give $y = 0$ or $x = 0$ modulo q. The first is excluded by definition. The second has no candidate points because $\chi(-1) = -1$.

For ordinary directions of heights two and three, coincident roots give at most one half of the line length, with the same error. The full line has length at most $q/2 + 1$, so the desired quarter bound still holds. In a diagonal direction one character argument is constant. If it is zero or a nonsquare there are no candidate points; otherwise the remaining linear condition gives the half bound.

A.6 Concentration of sampled line lengths

Each parameter b defines a field parabola

$$\Pi_b : \quad 2bx = y^2 + b^2.$$

A candidate point belongs to the two parabolas specified by its parameter pair. For a fixed real line, regard its candidate points as edges between parameters. This is a loopless multigraph of maximum degree two, since a line meets each Π_b in at most two points. A parameter pair can give two parallel edges.

The edges split into three matchings. Indeed, every component is a path, a cycle or a double edge; alternating colours suffice except that an odd cycle may need a third. Within a matching the events use disjoint parameter pairs, so they are independent Bernoulli events of probability δ^2.

Let X be the retained line length, m the original length, and $\mu = \delta^2 m$. Hölder's inequality applied to the three matching moment-generating functions gives

$$\log \mathbb{E}e^{t(X-\mu)} \le \frac{3\mu t^2}{2(1-t)}, \quad 0 < t < 1.$$

To see the constant, write $X = \sum_{i=1}^{3} X_i$ and bound the logarithmic moment-generating function by one third of the sum for $3t(X_i - \mathbb{E}X_i)$. The Bernoulli bound and $e^{3t} - 1 - 3t \le 9t^2/[2(1-t)]$ give the displayed estimate. Exponential Markov yields

$$\Pr(X \ge \mu + u) \le \exp\left(-\frac{u^2}{6\mu + 2u}\right). \tag{A.7}$$

Take $u = \sqrt{120\mu \log q} + 40 \log q$. At most q^4 lines are determined by pairs of candidate points. A union bound therefore gives, with probability tending to one, the simultaneous estimate

$$X \le \delta^2 m + O_\delta(\sqrt{q \log q} + \log q)$$

Combining this with the deterministic line bounds proves the line-length assertions of (4.2).

A.7 Triples sharing parameters

Fix a root v and condition on its two parameters being selected. All remaining parameters stay independent. Let $N_r(v)$ count the root triples involving exactly r distinct parameters. The three-edge parameter multigraph has maximum degree two, so $r \in \{3, 4, 5, 6\}$. We have

$$N_3 \le 8, \qquad N_4 \le 16q, \qquad N_5 \le 4q^2, \qquad N_6 \le d_{P_*}(v). \tag{A.8}$$

For $r = 3$ the graph is a triangle. Put $\Phi(r, s) = ((r^2 + s^2)/2, rs)$. After choosing signs, the three points are $\Phi(r, s), \Phi(s, t)$ and either $\Phi(t, r)$ or $\Phi(t, -r)$. Doubling both coordinates, their oriented determinants are respectively

$$-2(r - s)(s - t)(t - r)(r + s + t), \qquad -2(t - r)(t + s)(r + s)(t + r - s).$$

Distinct square parameters exclude the first three factors in each product from being zero. Collinearity implies $t = -r - s$ or $t = s - r$. In terms of the three square parameters this is the symmetric relation $a^2 + b^2 + c^2 - 2ab - 2ac - 2bc = 0$. Given root parameters b, c, the third satisfies $(b + c - a)^2 = 4bc$, giving at most two choices and at most four choices for the other point signs. Hence $N_3 \le 8$.

For $r = 4$ the graph is a three-edge path or a double edge and a disjoint edge. Set $k = |\mathcal{Q}|$. If the root is the middle edge of the path, choose a new parameter adjacent to b and its point,

in at most $2k$ ways. The resulting line already meets Π_c at the root, so there is at most one further intersection, which determines the last parameter. If the root is an end edge, choosing which root parameter to extend and then the next point gives at most $4k$ choices; the new parabola again determines the last point. If the root lies in a double edge, its mate is fixed and the third point lies on a vertical line, giving at most q choices. If the root is the disjoint edge, the other two points form a sign pair on that vertical line, again at most q possibilities. Thus $N_4 \le 6k + 2q \le 16q$.

For $r = 5$ the graph is a two-edge path and a disjoint edge. If the root belongs to the path, its adjacent point lies on one of the two root parabolas, giving at most $2q$ choices; allow at most q positions for the third point on their line. If the root is the disjoint edge, choose the common parameter of the path and one of its parabola points, in at most kq ways. The line through that point and the root has at most one further intersection with this parabola. These estimates give $N_5 \le 4q^2$.

A root triple involving r parameters survives the root conditioning with probability δ^{r-2}. Consequently,

$$\mathbb{E}[d_{P_L}(v) \mid v \in P_L] \le 8\delta + 16q\delta^2 + 4q^2\delta^3 + \delta^4 d_{P_*}(v). \tag{A.9}$$

For fixed δ, the first three terms are lower order than $q^2 \log q$.

A.8 Simultaneous degree concentration

Let $c_a(v)$ bound the change in root degree caused by changing a single nonroot parameter a. Count root triples containing a second point on Π_a. In directions of height $d \le \sqrt{q}$, there are at most $4d$ directions and two parabola intersections per direction, each with at most $2q/d$ partners. The generous bound $32q^{3/2}$ covers their sum. For larger heights, the at most q points on Π_a have at most $2\sqrt{q}$ partners each. Thus we may use

$$c_a(v) \le 64q^{3/2}.$$

Each root triple involves at most four nonroot parameters, so $\sum_a c_a(v) \le 4d_{P_*}(v)$. The coarse bound $d_{P_*}(v) \le 32q^2(1 + \log q)$ gives

$$\sum_a c_a(v)^2 \le 8192q^{7/2}(1 + \log q). \tag{A.10}$$

The bounded-differences inequality gives upper-tail probability at most $\exp(-2t^2/\sum_a c_a^2)$. One proof exposes the independent parameters in order. Each Doob martingale difference has conditional range of length at most c_a, and Hoeffding's lemma bounds its exponential moment by $\exp(\lambda^2 c_a^2/8)$. Iteration and optimization give the stated inequality.

Take $t = q^{15/8}$. The failure probability is $\exp(-\Omega(q^{1/4}/\log q))$, uniformly at a root. A union bound over at most q^2 roots still tends to zero. The conditioning on the two root parameters is legitimate: roots absent from P_L require no degree control. Also $|L| = \delta(q - 1)/2 + O(q^{2/3})$ with probability tending to one by its variance bound. Combining all successful events with (A.6) and (A.9) proves the parameter-selection lemma.

A.9 A two-root lattice correlation bound

Now let V be any subset of an integer square of side length q. For distinct u, x, put

$$w(u, x) = \max\{|V \cap ux| - 2, 0\}.$$

We claim that for distinct roots u, v,

$$\sum_{x \in V \setminus uv} w(u, x)w(v, x) \le Cq^2 T(q)(1 + \log q)^2 = q^{2+o(1)}, \tag{A.11}$$

where $T(q) = \max_{1 \le m \le 4q^2} \mathrm{d}(m)$ and $\mathrm{d}(m)$ is the divisor function.

Translate u to the origin and put $h = v - u$. For $x \notin uv$, let r, s be primitive directions parallel to $x - u, x - v$. There are nonzero integers i, j with $h = ir - js$, so

$$0 \neq \det(r, s) \mid \det(r, h), \qquad 0 < |\det(r, h)| \le 4q^2.$$

Group heights dyadically as $\|r\|_\infty \in [R, 2R)$ and $\|s\|_\infty \in [S, 2S)$. Suppose first that $R \le S$. There are $O(R^2)$ choices for r and at most $2T(q)$ signed values for $\det(r, s)$. For fixed primitive r and a fixed determinant, the integer solutions are $s = s_0 + kr$, because the integer kernel is $\mathbb{Z}r$. At most $O(1 + S/R)$ lie in the height range. Thus there are $O(T(q)RS)$ direction pairs. Each nonparallel pair determines at most one intersection x. The two weights are at most q/R and q/S, so this group contributes $O(T(q)q^2)$. Interchanging roots treats $R > S$. There are $O((1 + \log q)^2)$ groups, proving (A.11).

To justify $T(q) = q^{o(1)}$, fix $\varepsilon > 0$. For sufficiently large primes ℓ, $a + 1 \le 2^a \le \ell^{\varepsilon a}$. The finitely many smaller primes contribute at most a fixed power of $1 + \log m$, eventually bounded by a constant times m^ε. Thus $\mathrm{d}(m) \ll_\varepsilon m^{2\varepsilon}$, uniformly in the required range, and ε is arbitrary.

A.10 Six-point configurations

Suppose all degrees are at most D. At a fixed root v, consider the following ordered configurations of six distinct vertices:

$$\begin{aligned} &\mathrm{I}: \quad v, u_1, u_2 \in B; \quad x, u_1, y_1 \in B_1; \quad x, u_2, y_2 \in B_2; \quad x \notin B, \\ &\mathrm{II}: \quad v, u_1, u_2 \in B; \quad x, v, y_1 \in B_1; \quad x, u_1, y_2 \in B_2; \quad x \notin B. \end{aligned}$$

There are at most $2D$ ordered choices of (u_1, u_2). For type I, apply (A.11) to (u_1, u_2) to count the completions; for type II, apply it to (v, u_1). The weights count the available witness points on the respective lines. Each type therefore has at most

$$CDq^2T(q)(1 + \log q)^2. \tag{A.12}$$

configurations. For fixed K, $q \le Ks$ and $D = s^2 \log s$, this is $s^{4+o(1)}$.

B From random arcs to ordinary colourings

This appendix proves the colouring lemma used in the main argument. We give the process and all estimates for the parameters used here. The ordered selection and potential-function framework is adapted from Pham, Le Quang-Ham and Xue [2].

B.1 The random arc required

Suppose V lies in an integer square of side length $q \le Ks$. Ordinary blocks have length at most s, special blocks length at most $2s$, and a point belongs to at most two special blocks. The degree bound is

$$D = s^2 \log s.$$

We first construct a random arc J satisfying

$$\Pr(v \in J) = \rho := \frac{27}{200s}, \qquad \Pr(u, v \in J) \le \beta := s^{-2503/1500}. \tag{B.1}$$

for distinct u, v. Here K is fixed and the threshold may depend on K. The polynomial bound on the number of vertices will be used only in the colouring conversion.

B.2 Sampling and pruning

Use π for the sampling probability, to distinguish it from the main prime p. Set

$$\pi = s^{-5/6}, \quad a = \tfrac{1}{5}s^{-1/6}, \quad t_* = s^{-1/1000}, \quad \epsilon_0 = 1/1000. \tag{B.2}$$

Sample vertices independently with probability π, obtaining S. Delete sampled vertices whose sampled degree exceeds $(1+\epsilon_0)\pi^2 D$, and those lying in a sampled block longer than $(1+\epsilon_0)\pi s$ for ordinary blocks, or twice that for special blocks. Call the remaining set W.

Fix a root v and condition on $v \in S$. Distinct incident blocks are disjoint outside v, so their sampled sizes outside the root are independent. If a block has m other vertices and sampled size X there, expansion into pair indicators gives

$$\operatorname{Var}\binom{X}{2} \le \binom{m}{2}\pi^2 + 6\binom{m}{3}\pi^3.$$

Disjoint pairs have zero covariance, while pairs sharing a vertex contribute at most π^3. Since $m \le 2s$, summing over incident blocks gives

$$\operatorname{Var}(d_S(v) \mid v \in S) \le \pi^2 D + 4s\pi^3 D.$$

The conditional mean is at most $\pi^2 D$, so the probability of deletion for excessive degree is at most

$$\epsilon_0^{-2}\left(\frac{1}{\pi^2 D} + \frac{4s}{\pi D}\right) = o(1).$$

There are at most D incident blocks of size at least three. A binomial Chernoff bound and a union bound give deletion probability at most $D\exp(-c\pi s) = o(1)$ for excessive length. Here $\pi s = s^{1/6}$, and the fixed slack ϵ_0 absorbs the extra root and rounding. Hence, uniformly in v,

$$\Pr(v \in W) = \pi(1 - o(1)). \tag{B.3}$$

On W, the degree bound $D' = (1+\epsilon_0)\pi^2 D$ and the two length thresholds hold deterministically. Also any six specified distinct vertices belong to W with probability at most π^6, since $W \subseteq S$.

B.3 The ordered process

Fix an order on W and give each vertex initial weight $z_u = a$. Process a vertex x by selecting it with probability $r = z_x$. Say that a later vertex u is *affected* by x if x, u and a previously selected vertex belong to a common block.

If x is selected, set the weights of all affected later vertices to zero. If x is rejected, replace each such weight by $\min\{z_u/(1-r), t_*\}$. Leave other weights unchanged. Use fresh randomness at each step. The selected set is an arc: once a block has two selected vertices, every later vertex in that block has weight zero.

Without the cap, the update preserves each affected weight in conditional expectation, since $r \cdot 0 + (1-r)z_u/(1-r) = z_u$. The shared random decision nevertheless correlates the affected weights, and the cap can reduce their expectations. We bound the resulting loss by a potential at each root.

B.4 The potential

Set

$$\sigma = \frac{1}{1000\log s}, \qquad \Lambda = \log(t_*/a) = \frac{497}{3000}\log s + \log 5.$$

For an ordinary block take $\eta_B = 1/4, \kappa_B = 4/3$, and for a special block take $\eta_B = 1/2, \kappa_B = 2$. In both cases,

$$\kappa_B(1 - \eta_B) = 1. \tag{B.4}$$

Put $\lambda_B = \exp(-99\eta_B/(100\sigma))$.

For an unprocessed root v, let U_B be the other unprocessed vertices in an incident block, and let b_B be the number already selected in that block. A positive root weight implies $b_B \in \{0, 1\}$. Define

$$G_B = \prod_{u \in U_B} (1 + z_u/\sigma),$$

$$\Phi_{B,0} = \kappa_B \sum_{\{u,w\} \subset U_B} z_u z_w, \qquad \Phi_{B,1} = \kappa_B \sum_{u \in U_B} z_u,$$

and

$$C_v = z_v \left[1 - \sum_{B \ni v} \lambda_B G_B - \frac{\log(z_v/a) + \sum_{B \ni v} \Phi_{B,b_B}}{\Lambda} \right]_+ . \tag{B.5}$$

Here $[x]_+ = \max(x, 0)$. If the root is absent from W, has zero weight, or has previously had zero potential, its potential is kept at zero. Thus no logarithm of zero is taken. When a root with positive potential is processed, replace the potential by its selection indicator and freeze it; a root whose potential has already died retains potential zero.

The product term penalizes concentration of weight in a block. The Φ terms reserve a cost for completing a triple, and the logarithmic term controls approach to the cap. Every positive weight is at least a, so $0 \le C_v \le z_v$ before processing. Processing the root itself cannot decrease conditional expected potential, and its final potential is at most its inclusion indicator.

As $t_*/\sigma \to 0$, for large s and $0 \le z \le t_*$ we have $\log(1 + z/\sigma) \ge 0.99z/\sigma$. Thus $C_v > 0$ implies

$$z_v < t_*, \qquad M_B := \sum_{u \in U_B} z_u < \eta_B, \qquad \lambda_B G_B < 1. \tag{B.6}$$

For example, if $M_B \ge \eta_B$, then $\lambda_B G_B \ge 1$, contradicting positivity of the bracket in (B.5).

B.4.1 Initial potential

The initial total weight of a pruned ordinary block is at most $1001/5000$, and that of a special block at most $1001/2500$. The gaps below $0.99\eta_B$ are $473/10000$ and $473/5000$, respectively. Since $G_B \le \exp(M_B/\sigma)$ and there are at most D' incident blocks,

$$\sum_{B \ni v} \lambda_B G_B \le D' s^{-473/10} = o(1).$$

For the initial quadratic potential use coefficient $4/3$ for all root triples, then add the coefficient difference $2/3$ on the at most two special blocks. Their combined degree is at most $4(1 + \epsilon_0)^2\pi^2 s^2$. Multiplication by a^2 gives $O(1)$, and division by Λ makes this contribution $o(1)$. The main contribution satisfies

$$\frac{(4/3)a^2 D'}{\Lambda} \le \frac{572}{1775}.$$

Together with (B.3), this gives

$$\mathbb{E}C_v(0) \ge \left(\frac{1203}{8875} - o(1) \right) \frac{1}{s}. \tag{B.7}$$

B.5 One-step losses

Condition on W and the history immediately before x is processed. A dead root potential has no further loss, so assume $C_v > 0$. Take s large enough that $r \le t_* \le 1/4$, $\Lambda \ge 1$, $\sigma \le 1$ and $2t_*\eta_B/\sigma \le \log 2$.

Consider an incident block B not containing x, and let $A_B \subseteq U_B$ be its affected vertices. Write $e_2(z_{A_B}) = \sum_{\{u,w\}\subset A_B} z_u z_w$. For upper bounds on the updated products and Φ terms we may ignore the cap.

The two-vertex case explains the pair sum. Suppose $U_B = A_B = \{u, w\}$. Selection of x sets both weights to zero; rejection divides both by $1 - r$. Each individual expectation is unchanged, but the expected product becomes

$$r \cdot 0 + (1-r)\frac{z_u}{1-r}\frac{z_w}{1-r} = \frac{z_u z_w}{1-r}.$$

Expand $G_B = (1 + z_u/\sigma)(1 + z_w/\sigma)$. Its constant and linear terms cancel in $\mathbb{E}G'_B - G_B$, leaving

$$\mathbb{E}G'_B - G_B = \frac{1}{\sigma^2}\left(\frac{z_u z_w}{1-r} - z_u z_w\right) = \frac{r}{1-r}\frac{z_u z_w}{\sigma^2} \qquad \text{(with the cap omitted)}.$$

Restoring the cap can only decrease the updated product. If only one of the two vertices is affected, the uncapped product expectation is unchanged. This is why the general estimate counts simultaneously affected pairs.

For arbitrary A_B, product expansion gives

$$\mathbb{E}G'_B - G_B \le \frac{4rG_B}{\sigma^2} e_2(z_{A_B}), \tag{B.8}$$

$$\mathbb{E}\Phi'_{B,0} - \Phi_{B,0} \le \frac{\kappa_B r}{1-r} e_2(z_{A_B}), \qquad \mathbb{E}\Phi'_{B,1} \le \Phi_{B,1}. \tag{B.9}$$

To verify (B.8), factor the product into affected and unaffected variables. The constant term is unchanged in both branches. For degree $j \ge 1$ in the affected variables, the coefficient difference after averaging is $(1-r)^{1-j} - 1$. It is zero for $j = 1$; for $j \ge 2$ it is at most $(j-1)r(1-r)^{1-j}$. Use $j - 1 \le \binom{j}{2}$ to choose a pair first and restore the remaining terms as a product. The ratio of the inflated product to the original is at most $\exp(2rM_B/\sigma) \le 2$, and $(1-r)^{-1} \le 4/3$. The relaxed constant four follows. For (B.9), only pairs with both vertices affected can increase in expectation; a pair with one affected vertex has unchanged expectation.

B.5.1 The current vertex does not affect the root

The root weight is unchanged. If v, x have a common block, it has no selected vertex. Let $R = U_B \setminus \{x\}$. Removing x changes the expected Φ on this block to

$$\kappa_B[(1-r)e_2(z_R) + r\sum_R z_u] \le \kappa_B[e_2(z_R) + r\sum_R z_u],$$

The right-hand side is its previous quadratic potential. The product also does not increase. Apply (B.8)–(B.9) on the other blocks. Using $\lambda_B G_B \le 1$, weights at most t_* and $\kappa_B \le 2$, the conditional expected loss is at most

$$\frac{8t_*^3 r}{\sigma^2} \sum_{B\ni v,\ x\notin B} \#\{\{u_1, u_2\} \subset U_B : x \text{ affects both } u_1, u_2\}. \tag{B.10}$$

The positive-part truncation preserves this lower bound on expected potential, since $[x]_+ \ge x$. When a potential first becomes nonpositive inside the bracket, its value at that step is precisely zero, so the estimate also covers the rule that it remains dead thereafter.

B.5.2 The current vertex affects the root

The common block B_* now contains one selected vertex. Selection of x kills the root weight. On rejection we first show that the updated root weight is exactly $z_v/(1-r)$, without capping. Indeed,

$$\sum \Phi \geq \kappa_{B_*} r \geq \frac{4}{3} r \geq -\log(1-r),$$

and positivity of (B.5) implies $\log(z_v/a) - \log(1-r) < \Lambda$, as required.

Put $M_R = M_{B_*} - r$. On the rejection branch, the common-block product satisfies

$$\log(G'_{B_*}/G_{B_*}) \leq -\frac{0.99r}{\sigma} + \frac{rM_R}{\sigma(1-r)} \leq 0,$$

because $M_R/(1-r) \leq (1/2)/(3/4) = 2/3 < 0.99$. Its linear potential decreases by at least

$$\kappa_{B_*}\left(M_{B_*} - \frac{M_{B_*} - r}{1-r}\right) = \frac{\kappa_{B_*} r(1 - M_{B_*})}{1-r} \geq \frac{r}{1-r} \geq -\log(1-r).$$

using (B.4). This pays for the increase in the root's logarithmic term.

For any other incident block, rejection increases the product by at most $4rG_B \sum_{A_B} z_u/\sigma$. The corresponding increase in the Φ term is at most $10r \sum_{A_B} z_u$. The first uses the product-ratio bound two. For the second, the linear term is divided by $1-r$. Bound the quadratic term using $(1-r)^{-2} - 1 \leq 4r$ together with $M_B \leq 1/2$. Since $(1-r)z'_v = z_v$, the conditional expected loss is at most

$$\frac{16t_*^2 r}{\sigma} \sum_{B \ni v,\ B \neq B_*} \#\{u \in U_B : x \text{ affects } u\}. \tag{B.11}$$

B.6 Charging the losses to configurations

For each term in (B.10), choose a previously selected witness for each of the two affected vertices. This gives a type I configuration of Appendix A. For (B.11), call the affected vertex u_1, choose an auxiliary vertex u_2 in its root block, and choose selected witnesses for v and u_1. This gives type II. The block has at least three vertices in W, so the auxiliary exists; it need not be unprocessed.

The six vertices are distinct. The current vertex x is outside the root block. If the two witnesses were equal, the two witness lines through x would coincide and force x onto the root block. A witness equal to another specified root-block vertex gives the same contradiction. The remaining coincidences are excluded by the processing order or by definition. Choose the first available witnesses and auxiliary in a fixed order. Each ordered configuration is charged at most once in a run, since it specifies x.

Given W containing the six vertices, the witnesses must be selected before x. The tower property gives

$$\mathbb{E}[z_x \mathbf{1}_{\{y_1,y_2\} \subset J} \mid W] = \Pr(x, y_1, y_2 \in J \mid W) \leq t_*^3.$$

The witness event is measurable before processing x. The bound uses the conditional selection cap at three successive processing times, without requiring independence.

The six vertices belong to W with probability at most π^6, and Appendix A bounds the number of configurations by $s^{4+o(1)}$. Summing all expected losses therefore gives

$$O\left(s^{4+o(1)}\pi^6\left(\frac{t_*^6}{\sigma^2} + \frac{t_*^5}{\sigma}\right)\right) = o(1/s). \tag{B.12}$$

Indeed, $\pi^6 = s^{-5}$, and after division by $1/s$ the two terms are $s^{-6/1000+o(1)}$ and $s^{-5/1000+o(1)}$. Their negative powers absorb the divisor and logarithmic factors for fixed K.

Initial expected potential minus all losses is bounded above by the final inclusion probability. Thus

$$\Pr(v \in J_0) \geq (1203/8875 - o(1))/s.$$

Since $1203/8875 - 27/200 = 39/71000 > 0$, we can make all marginals exactly ρ by independently retaining a selected vertex v with probability ρ/ρ_v, where ρ_v is its actual inclusion probability. This preserves the arc property and decreases pair probabilities. Both vertices must lie in the initial sample, of probability at most π^2; conditional on W, their joint selection probability is at most t_*^2. Hence

$$\Pr(u, v \in J) \leq \pi^2 t_*^2 = s^{-2503/1500}.$$

This proves (B.1). The thinning establishes existence and does not require an efficient procedure for computing the ρ_v.

B.7 An abstract lemma for the final remainder

We also need a lemma for an arbitrary finite linear block system, with no ambient-square assumption. If the maximum block size is at most $2s$ and all degrees are at most $D = s^2 \log s$, there is a random arc with marginal $1/(17s)$ and pair probability at most $s^{-19/10}$.

B.7.1 Abstract configuration counts

Let $w(u, x)$ be the common-block size minus two, or zero if there is no block of size at least three containing both. Then $\sum_x w(u, x) = 2d(u) \leq 2D$. If $v \notin B$, the blocks joining v to distinct $x \in B$ are distinct, by linearity. Therefore

$$\sum_{x \in B} w(v, x)^2 \leq 2D, \quad \sum_{x \in B} w(v, x) \leq \sqrt{2|B|D}.$$

For roots u, v, group points outside their common block by blocks through u, and put $\alpha_B = |B| - 2$. Blocks with $\alpha_B < D^{1/3}$ contribute at most $2D^{4/3}$. For the others, use $|B| \leq 3\alpha_B$ to bound the contribution by

$$\sqrt{6D} \sum_{\alpha_B \geq D^{1/3}} \alpha_B^{3/2} \leq \sqrt{6D}\, D^{-1/6} \sum_B \alpha_B^2 \leq 2\sqrt{6} D^{4/3}.$$

Thus the two-root correlation sum is less than $7D^{4/3}$. Choosing the root's ordered neighbour pair in at most $2D$ ways shows that each type of six-point configuration has at most $14D^{7/3}$ possibilities.

B.7.2 Parameters and potential estimates

Use the same process with

$$\pi = s^{-14/15}, \quad a = s^{-1/15}/12, \quad t_* = s^{-1/60}, \quad \epsilon_0 = 1/100, \quad \sigma = (256 \log s)^{-1}.$$

For every block take $\eta = 3/8$, $\kappa = 2$, $\lambda = \exp(-\eta/(2\sigma)) = s^{-48}$ and $\Lambda = (\log s)/20 + \log 12$. Prune blocks above $2(101/100)\pi s$. The variance argument is unchanged, and $\pi s = s^{1/15} \to \infty$, so the root-retention probability is still $\pi(1 - o(1))$.

Use (B.5), now with these constants. The inequality $\log(1 + z/\sigma) \geq z/(2\sigma)$ gives $M_B < 3/8$ when the potential is positive. The common-block product estimate needs $M_R/(1 - r) \leq 1/2$, which follows from $(3/8)/(3/4) = 1/2$. The linear-potential compensation needs $\kappa(1 - \eta) \geq 1$, also valid. All other estimates above use the same inequalities and give (B.10)–(B.11) with unchanged upper constants.

The initial block weight is at most 101/600, so its product charge is bounded by

$$D's^{-48+256(101/600)} = O(s^{-358/75}\log s) = o(1).$$

The quadratic charge is at most $2a^2D'/\Lambda \le 101/360$. Initial expected potential is therefore at least $(259/4320 - o(1))/s$. The abstract configuration count bounds the total loss by

$$256\pi^6 D^{7/3}(t_*^6/\sigma^2 + t_*^5/\sigma) = o(1/s);$$

after division by $1/s$, the two terms are $O(s^{-1/30}\log^{13/3} s)$ and $O(s^{-1/60}\log^{10/3} s)$. Since $259/4320 - 1/17 = 83/73440 > 0$, the same thinning gives the exact marginal. The pair bound is $\pi^2 t_*^2 = s^{-19/10}$, proving the abstract random-arc lemma.

B.8 Converting random arcs into colour classes

Consider a general conversion with maximum block length $2s$ and maximum degree $D = s^2\log s$. Suppose each induced system at this scale admits a random arc with exact marginal $\rho = c/s$ and pair bound $\beta = s^{-1-\varepsilon}$, where $c > 0$ and $0 < \varepsilon < 1$ are fixed. For the geometric lemma track ordinary lengths s and special lengths $2s$ separately.

Fix $\theta, h > 0$, and run

$$T = \lceil(\theta + h)/\rho\rceil$$

rounds. At each round use fresh randomness on the remaining vertices and assign the selected arc a fresh colour. Within a stage the scale stays fixed; deletion only improves its hypotheses. All probability bounds below are conditional on the preceding history.

B.8.1 A one-sided martingale inequality

Suppose the increments X_t of an adapted process have nonpositive conditional means, satisfy $X_t \le b$, and have conditional second moments whose sum is at most a deterministic v_*. Then

$$\Pr(M_T - M_0 \ge u) \le \exp\left(-\frac{u^2}{2(v_* + bu/3)}\right). \tag{B.13}$$

For $x \le b$ and $0 < \lambda < 3/b$,

$$e^{\lambda x} \le 1 + \lambda x + \frac{\lambda^2 x^2}{2(1 - \lambda b/3)}.$$

For negative x, use $e^t \le 1 + t + t^2/2$; for positive x, expand the series and use $j! \ge 2\cdot 3^{j-2}$. Write $V_t = \mathbb{E}(X_t^2 \mid \mathcal{F}_{t-1})$ and $g = \lambda^2/[2(1 - \lambda b/3)]$. Conditional expectation gives $\mathbb{E}(e^{\lambda X_t} \mid \mathcal{F}_{t-1}) \le e^{gV_t}$. Hence $\exp(\lambda(M_t - M_0) - g\sum_{i\le t} V_i)$ is a nonnegative supermartingale starting at one. Since $\sum V_i \le v_*$, the moment-generating function is at most e^{gv_*}. Markov's inequality with $\lambda = u/(v_* + bu/3)$ proves (B.13). The degenerate case $v_* = 0$ follows directly or by a limit.

B.8.2 Block-length contraction

Fix an initial block. Let Z_t be its remaining size and Y_t the number removed in round t. Since the selected set is an arc, $0 \le Y_t \le 2$, and, conditionally on $\mathcal{F}_{t-1}$,

$$\mathbb{E}(Y_t \mid \mathcal{F}_{t-1}) = \rho Z_{t-1}, \qquad \mathbb{E}Y_t^2 \le 2\rho Z_{t-1} \le 4\rho s.$$

Thus $(1-\rho)^{-t}Z_t$ is a martingale. Its normalization is bounded during the stage by a constant depending only on θ, h. Its positive increments and conditional second moments are $O(1)$, and the summed second moments are $O(s)$.

Put $r_0 = e^{-\theta}$. If an ordinary block ends longer than $r_0 s$, its normalized value exceeds the initial upper bound s by at least $(e^h - 1 + o(1))s$. For a special block the target is $2r_0 s$ and the same fixed relative gap applies. By (B.13), the failure probability for each block is $\exp(-\Omega_{\theta,h}(s))$.

B.8.3 Tracking neighbour pairs after the root is coloured

For every root present at the start of the stage define

$$Z_t(v) = \sum_{B \ni v} \binom{|(B \setminus \{v\}) \cap V_t|}{2},$$

where V_t is the uncoloured set. Continue to use this quantity even if v has been coloured. It equals the actual degree whenever v survives, and avoids conditioning on survival.

Each remaining neighbour pair is destroyed with probability between $2\rho - \beta$ and 2ρ. Thus the one-round loss Y satisfies, conditionally on the history,

$$(2\rho - \beta) Z_{t-1} \le \mathbb{E}Y \le 2\rho Z_{t-1}.$$

Give a neighbour u weight w_u equal to the number of these pairs containing it. Linearity gives $w_u \le 2s$ and $\sum_u w_u = 2Z_{t-1}$. Since $Y \le \sum_u w_u \mathbf{1}_{u \in J}$,

$$\mathbb{E}Y^2 \le \rho \sum_u w_u^2 + \beta (\sum_u w_u)^2 \le 4\rho s D + 4\beta D^2. \tag{B.14}$$

Let $\alpha = 1 - (2\rho - \beta)$. Then $\alpha^{-t} Z_t$ is a supermartingale with positive increments at most $O(\rho D)$. Its stage-total conditional second moment is

$$O(sD + (\beta/\rho) D^2).$$

Here $\rho^2 \le \beta$ eventually because $\varepsilon < 1$. If the final degree exceeds $(r_0 s)^2 \log(r_0 s)$, its normalized value exceeds D by at least $(e^{2h} - 1 + o(1))D$. Inequality (B.13) bounds the failure probability by

$$\exp\left(-\Omega_{\theta,h}\left(\frac{1}{s/D + \beta/\rho + \rho}\right)\right) \le \exp(-\Omega_{\theta,h}(s^{\varepsilon})).$$

There are at most $\binom{|V|}{2}/3$ blocks of size at least three and at most $|V|$ roots. A union bound gives simultaneous contraction failure probability at most

$$O(|V|^2) \exp(-c_{\theta,h} s^{\varepsilon}). \tag{B.15}$$

B.9 Stage costs and the limiting coefficient

First apply the abstract random-arc lemma with $c = 1/17$ and $\varepsilon = 9/10$. Suppose the initial number of vertices is $N \le s^A$. Reduce the scale successively by the factor r_0 until it first drops below $R = C_{\theta,h}(1 + \log^3(N+2))$. Above R, choose a successful stage; (B.15) makes its failure probability less than one when the constant is sufficiently large. This is a finite existence induction on stages.

At termination the remaining degree is at most $R^2 \log R$. Greedy colouring uses at most $\lfloor R^2 \log R \rfloor + 1$ further colours, since each incident triple forbids at most one colour when its final vertex is processed. This costs $O(\log^6 N \log\log(N+3)) = o(s)$. Summing the stage costs gives

$$\chi \le \left(\frac{\theta + h}{c(1 - e^{-\theta})} + o(1)\right) s.$$

For any prescribed accuracy, choose a sufficiently small fixed θ, set $h = \theta^2$, and then let s grow. The coefficient tends to $1/c$. The abstract colouring cost is therefore $(17 + o(1))s$.

Now use the geometric random-arc lemma with $c = 27/200$ and $\varepsilon = 1003/1500$. Fix $\xi > 0$ and stop when the scale first falls below ξs. There are only a fixed number of stages, and throughout them $q/s_i \le Ke^{\theta}/\xi$. Thus the geometric lemma remains applicable with a fixed ambient constant. Colour the tail by the abstract bound, at cost at most $(17 + o(1))\xi s$. A polynomial bound on the initial vertex count remains polynomial in the tail scale, since ξ is fixed. Total cost is at most

$$\left(\frac{\theta + h}{(27/200)(1 - e^{-\theta})} + 17\xi + o(1)\right) s.$$

Choose fixed small θ, $h = \theta^2$ and ξ in that order of approximation, and then take s sufficiently large. We obtain

$$\chi_A(V) \le (200/27 + o(1))s.$$

Every selected arc has received one actual colour, and the terminal set has been fully coloured. This proves the ordinary-colouring lemma.

C Geometric exchange, absorption and transfer

This appendix proves (3.5) and (7.1). For the main zero-label points, the argument below gives a repair cost $q^{1/2+o(1)}$. This is sufficient for the fixed-proportion theorem and makes a stronger historical repair estimate unnecessary here.

C.1 Four-point conflicts in adjacent windows

Let $q \equiv 3 \pmod 4$, put $h_q = (q-1)/2$, and let $\mathcal{D}$ be the nonzero nonsquares. Define

$$C_d = \{(x, y) \in U_q : x^2 - y^2 = d \pmod q\}, \quad V_d = C_d \cup ((q, q) + C_d) \cup ((q, -q) + C_d).$$

Reduction modulo q is injective on each window. A real grid line, written with primitive integer coefficients, still reduces to a field line. Since the nondegenerate curve $x^2 - y^2 = d$ has at most two intersections with a field line, any three collinear points in V_d must contain a repeated modular image.

Repeated images in the central and upper windows differ by (q, q); those in the central and lower windows differ by $(q, -q)$. Along either direction, restriction of $x^2 - y^2$ is a nonconstant linear polynomial. If its linear term vanishes, its constant value is zero, whereas $d \neq 0$. Thus each such field line contains at most one curve point. The representative in the third window is not on the repeated pair's line, so these pairs give no triple.

The remaining repeated pairs lie in the upper and lower windows and differ by $(0, 2q)$. Their vertical line misses the central window, since $q - h_q > h_q$. Two nonzero roots of $y^2 = x^2 - d$ give exactly four points,

$$(q + x, q + y), (q + x, q - y), (q + x, -q + y), (q + x, -q - y), \quad 1 \le y \le h_q. \tag{C.1}$$

A zero root yields only two points. Thus all forbidden triples are contained in these disjoint four-point branches, and deleting any two points from each branch makes every old class an arc.

C.2 Parabola colours and the pointwise exchange

For $b \in \mathcal{Q}$ consider

$$\Gamma_b : \ y = (x^2 + b^2)/(2b),$$

and its representative graph S_b in U_q. Its copies in the two corner windows form a set $\mathcal{A}_b$. This is an arc. Three distinct modular images would contradict the parabola's two-intersection property. A repeated image is a vertical pair of copies, and that vertical line contains no third graph point because the graph has one value at each horizontal coordinate.

Let $l = |L|$ and retain parameters $\mathcal{Q} \setminus L$. Whenever a branch is covered by a retained parabola at (x, y) or $(x, -y)$, remove the corresponding two corner copies from the old colour and give them the parabola colour. Each parameter colour is a subset of $\mathcal{A}_b$.

For a branch with $x \neq 0$, membership in Γ_b is equivalent to

$$b^2 - 2yb + x^2 = 0.$$

Because $x^2 - y^2$ is a nonsquare and -1 is a nonsquare, $y^2 - x^2$ is a nonzero square. The equation has distinct nonzero roots b, c with

$$b + c = 2y, \quad bc = x^2, \quad x^2 - y^2 = -(b - c)^2/4. \tag{C.2}$$

Their quadratic characters agree. Replacing y by $-y$ negates both roots, so exactly one pair consists of squares. The branch is uncovered precisely when both those square parameters lie in L. Each unordered pair gives the two roots of $x^2 = bc$, hence two branches. There are exactly $2\binom{l}{2} = l(l-1)$ uncovered nonaxis branches. For $x = 0$ the unique nonzero parameter is $2y$; exactly one of it and its negative is square. This gives exactly l additional uncovered branches.

Define

$$P_L^0 = \{([b/2], 0) : b \in L\}, \qquad P = P_L \cup P_L^0.$$

and let $Z = \{(x, y) \in U_q : x^2 - y^2 = 0 \pmod q\}$. Set

$$R_0 = Z \cup \bigcup_{d \in \mathcal{D}} V_d, \quad E = F_q^+(P) \cup F_q^-(P), \quad R_L = (R_0 \setminus E) \cup P. \tag{C.3}$$

To verify the exchange, fix b, c and put $u = [(b + c)/2]$, $v = [\sqrt{bc}]$. The four central points $(\pm u, \pm v)$ have the nonzero square label $(b - c)^2/4$. The two corner images of each point contribute one point to each of the two uncovered branches with horizontal coordinates $q + v, q - v$. The four central points account for all four positions in each branch. Our choice $(u, v), (u, -v)$ therefore removes exactly two points per branch. An axis point $([b/2], 0)$ also removes exactly two points from its corresponding branch. The parameter recovery in (C.2) shows that distinct parameter objects do not overlap in branches.

Originally R_0 has the hitting property: a centre of zero or nonsquare label is already in R_0, whereas the two images of a square-labelled centre have nonsquare labels. If the exchange deletes an image, its unique central preimage is in P and is added to R_L. Hence

$$\forall z \in U_q : \quad z \in R_L \text{ or } F_q^+(z), F_q^-(z) \in R_L. \tag{C.4}$$

Covered branches are repaired using retained parabola colours; uncovered branches have two points removed by the exchange. The remaining old classes are arcs. We have used $(q-1)/2$ old colours and $(q-1)/2 - l$ parameter colours. The uncoloured part is $Z \cup P_L^0 \cup P_L$. The set P_L is handled in the main proof. The other points lie on $y = x, y = -x, y = 0$, and the simultaneous absorption argument below costs $q^{3/4+o(1)}$. It will establish

$$\chi_A(R_L) \leq q - 1 - l + \chi_A(P_L) + q^{3/4+o(1)}. \tag{C.5}$$

C.3 Restricted matchings

Consider a bipartite graph with M vertices on each side, missing at most $d < M/2$ edges at each vertex. It has a perfect matching. Indeed, a nonempty left set of size at most $M - d$ has at least $M - d$ neighbours, while a larger set has size greater than d and therefore meets the neighbourhood of every right vertex. Hall's condition follows; the empty set is immediate.

Choose uniformly among allowed perfect matchings. For an allowed edge e, a matching containing e has at least $M - 1 - 2d$ other edges permitting a two-edge switch that removes e. A matching without e has at most one inverse switch. Double counting gives

$$\Pr(e \text{ is used}) \le \frac{1}{M - 2d}.$$

Conditional on compatible specified edges, the remaining matching is uniform on the remaining graph. If the original size is m, the missing degree is at most $m/8$, and at most $m/4$ edges are specified, every denominator is at least $m/2$. Thus any $k \le m/4$ compatible specified edges satisfy

$$\Pr(e_1, \dots, e_k \text{ are all used}) \le (2/m)^k. \tag{C.6}$$

Incompatible assignments have probability zero. We choose different groups independently, but use this joint bound within each group.

C.4 Witness bounds for axes and an exterior line

All coordinates in this subsection lie in $[-Cq, Cq]^2$, with fixed C, and there are $m \ge \eta_0 q$ old colours, with fixed $\eta_0 > 0$. Each old class is already an arc on a specified nonzero modular conic. A *witness* is a new point collinear with two same-colour old points. Let $w(A, d)$ count the union of old endpoints of colour d witnessed by A.

C.4.1 Zero axes of a product conic

Suppose the old equation is $xy = d \neq 0$ and the new line is $y = 0$. Fix an old point (x, y). Partners of the same modular image number $O_C(1)$, and their lines each meet the new line at most once. For distinct modular images, the modular intercept relation gives $a = x + x' - kq$, with $O_C(1)$ possible integers k. Integer collinearity gives the exact identity

$$(x' - kq)y' = (x - kq)y \neq 0. \tag{C.7}$$

Its nonzero right-hand side has size $O_C(q^2)$. The divisor bound therefore gives $q^{o(1)}$ partners. Thus each old point has at most $D_0 = q^{o(1)}$ single-new-point witnesses. Each old class has $O_C(q)$ points, so every column sum of the witness matrix is $q^{1+o(1)}$.

Fix $A = (a, 0)$, first separating the $O_C(1)$ points with $a \equiv 0 \pmod q$. For a distinct-image chord write $P = A + i(u, v)$ and $P' = A + j(u, v)$, with primitive direction, $u > 0$ and $v \neq 0$. Equal labels imply

$$(i + j)u = kq - a \neq 0.$$

There are $O_C(1)$ values of k and $q^{o(1)}$ divisor choices for u. Once u is fixed, so is $i + j$, and $|v| \le O_C(q)/\max(|i|, |j|)$. Harmonic summation over i gives $q^{1+o(1)}$ endpoints. Repeated-image chords have only $O_C(1)$ primitive directions, each with $O_C(q)$ endpoints. This also bounds the row sum by $q^{1+o(1)}$. Interchanging coordinates treats the vertical axis.

The same estimates hold for a fixed number of lines $x = kq$ or $y = kq$: translate the supporting line to the corresponding axis. Modular labels are unchanged and the coordinate constant only increases by a fixed amount. Modular-origin exceptions on these lines still number $O_C(1)$.

C.4.2 An exterior line

Now suppose old labels satisfy $x^2 - y^2 = d$ and $\chi(-d) = -1$, and the new line is $x = 0$. In particular every old point has nonzero horizontal coordinate modulo q. We prove a single-old-point witness bound $q^{1/2+o(1)}$ and row and column bounds $q^{3/2+o(1)}$.

We use the standard additive Weil bound for bounded-degree rational functions with a nonzero simple pole, and the square-root bound for Kloosterman sums; suitable statements appear in [3, 4]. A simple pole excludes the exceptional form $g^q - g +$ constant. Complete nontrivial sums are $O(\sqrt{q})$. A nonzero linear phase on a split hyperbola has the same bound.

The sum of absolute values of the unnormalized discrete Fourier coefficients of an integer interval of length at most q is $O(q \log q)$. This follows by summing the geometric-series bound $\min\{|I|, O(q/|t|)\}$. Fourier completion in two coordinates therefore gives

$$\#\{(j, v) \in I_1 \times I_2 : j = R(v) \pmod q\} \ll_C |I_1||I_2|/q + \sqrt{q} \log^2 q, \tag{C.8}$$

where poles are omitted and

$$R(v) = \frac{2xu - 2yv}{v^2 - u^2} = \frac{x - y}{v - u} - \frac{x + y}{v + u}, \quad u \neq 0, \quad x^2 - y^2 \neq 0 \pmod q.$$

The poles are noncancelling. The zero frequency contributes the area term; nonzero frequencies give the error. An interval of length $O_C(q)$ is split into a bounded number of intervals of length at most q.

A second consequence, for $k, a \neq 0 \pmod q$, is

$$\#\{(u, v) \in [-O_C(R), O_C(R)]^2 : k(u^2 - v^2) - 2av = 0 \pmod q\} \ll_C R^2/q + \sqrt{q} \log^2 q. \tag{C.9}$$

Indeed, the conic is equivalent to $(kv + a - ku)(kv + a + ku) = a^2 \neq 0$. This invertible affine change of variables sends every nonzero linear phase to a nonzero linear phase on the hyperbola. Fourier completion is performed on interval indicators in the original coordinates.

Fix an old point $P = (x, y)$. Write a distinct-image partner as $P + j(u, v)$ with primitive direction and $u > 0$. The integral intercept $y - xv/u$ implies $u \mid x$, giving $q^{o(1)}$ possibilities. The common-label relation becomes $j = R(v) \pmod q$. If $v = \pm u \pmod q$, the original equation would require $2u(x \mp y) = 0$, impossible here, so the poles give no missing witnesses.

The coordinate bounds give $|j| \leq O_C(q)/u$ and $|jv| \leq O_C(q)$. Group $|j|$ dyadically at size J; each group lies in a bounded number of rectangles of side lengths $O(J), O_C(q/J)$. By (C.8), each group contributes $O_C(\sqrt{q} \log^2 q)$ pairs. Multiply by the divisor choices and the $O(\log q)$ groups. The single-point witness bound is $q^{1/2+o(1)}$. Repeated-image partners add only a bounded contribution. Since an old class has $O_C(q)$ points, its column sum is at most $q^{3/2+o(1)}$.

Next fix $A = (0, a)$ after removing the bounded number with $a = 0 \pmod q$. For a distinct-image chord write its endpoints as $A + i(u, v), A + j(u, v)$. With direction height r, $|i|, |j| \leq O_C(q/r)$ and

$$(i + j)(u^2 - v^2) - 2av = 0 \pmod q. \tag{C.10}$$

For $r \leq \sqrt{q}$, use $O(r)$ directions of height r and $O_C(q/r)$ endpoints per line. The total is $O_C(q^{3/2})$.

In a dyadic range $R \leq r < 2R$ with $R > \sqrt{q}$, set $k = i + j$. There are $O_C(q/R)$ possibilities and $|k| < q$ for large q. If $k = 0$, then $v = 0 \pmod q$. The primitive direction with $v = 0$ was already counted at small height; nonzero multiples of q give only $O_C(q)$ endpoints in total. For $k \neq 0$, apply (C.9) to count directions and multiply by $O_C(q/R)$ endpoints per direction. The contribution of this range is

$$O_C\left(q + \frac{q^{5/2} \log^2 q}{R^2}\right).$$

Summing over dyadic ranges gives $q^{3/2+o(1)}$. This counts endpoint unions, so no further partner factor is needed. Repeated-image chords contribute $O_C(q)$. The row bound follows.

C.5 A general planar recolouring estimate

For a set of M points with at most ℓ collinear,

$$\chi_A \le C\{\sqrt{M\log(M+2)} + \ell\log(M+2) + 1\}. \tag{C.11}$$

We derive the estimate from the Szemerédi–Trotter incidence theorem [5]: the number of lines containing at least $j \ge 2$ points of an N-point set is $O(N^2/j^3 + N/j)$. Grouping line lengths dyadically and multiplying by the triples per line gives

$$T \le C_1[N^2\log(\ell+2) + N\ell^2].$$

Choose $A = C_2(N\log(\ell+2)+\ell^2+1)$ so that $T \le NA$. Sample independently with probability $\nu = \min\{1/2, 1/\sqrt{4A}\}$, then delete one vertex of each surviving triple. The resulting arc has expected size at least $N\nu - T\nu^3 \ge 3N\nu/4$. Thus some arc has size at least

$$c\frac{N}{\sqrt{N\log(\ell+2)} + \ell + 1}$$

Repeated removal of such arcs reduces a remainder of size N to size at most $N/2$ in $O(\sqrt{N\log(\ell+2)} + \ell + 1)$ colours. Summing over successive dyadic sizes, the square-root terms form a geometric series, while the ℓ term occurs $O(\log(M+2))$ times. This proves (C.11).

C.6 Simultaneous absorption

We use a single argument for both witness bounds. Suppose there are a fixed number of supporting lines, no old point lies on any of them, and each supports at most $2m$ new points. Assume the single-old-point witness bound is $D_0 = q^{\vartheta+o(1)}$ and each row and column sum is at most qF, where $F = q^{\vartheta+o(1)}$. We need only $\vartheta = 0$ and $1/2$. First separate line intersections and the modular-origin exceptions, whose number is bounded.

Split each line into two groups of size at most m, and pad each group with dummy vertices to size m. There are a fixed number g of groups. Allow only assignments with $w(A,d) \le J$, where J is a sufficiently large fixed multiple of F. The row and column bounds ensure missing degree at most $m/8$. Independently in each group choose a uniform allowed perfect matching. Each old colour receives at most two new points on each line and at most g new points overall.

Simultaneously delete both old endpoints of every monochromatic triple with two old and one new point. Also delete the old point of every monochromatic triple with one old and two new points. Call the old deletion set B. The first type deletes at most gJ points per colour. Each new pair determines a line meeting an old arc in at most two points, so the second deletes at most $2\binom{g}{2}$. Hence

$$|B| \le m(gJ + 2\binom{g}{2}) = q^{1+\vartheta+o(1)}. \tag{C.12}$$

To control line lengths in B, fix an old point. Its first-type deletion has at most D_0 single-new-point witnesses. Its second-type witnesses use two different supporting lines, because the old point is on neither line. Choosing one new point determines the other as an intersection, so there are $O_C(q)$ such pairs.

For $k \le m/4$ specified points of different old colours, expand their deletion witnesses. If one new point is assigned two different colours, the term has probability zero. Otherwise

each group has at most k specified edges; the two edges of a pair witness belong to different supporting lines. Applying (C.6) and summing over witnesses gives

$$\Pr(P_1,\dots,P_k\in B)\le\left(\frac{2D_0}{m}+\frac{O_C(q)}{m^2}\right)^k\le(W/q)^k,\quad W=q^{\vartheta+o(1)}. \tag{C.13}$$

Each old colour has at most two points on any line. A line containing $2L$ deleted points therefore supplies L points of different old colours. A line in the bounded box has $O_C(q)$ lattice points and there are $O_C(q^4)$ candidate lines. Taking $L=\lceil K(W+\log q)\rceil$ with sufficiently large K yields

$$\Pr(\exists\ell:|B\cap\ell|\ge 2L)\le O_C(q^4)(O_C(e)W/L)^L=o(1).$$

Since $\vartheta\le 1/2$, eventually $L\le m/4$, as required by the matching bound.

It remains to handle triples consisting only of new points. Each supporting line has at most two new points of a colour. If two points of a triple lie on the same supporting line, the third must be a previously separated intersection. All remaining candidates use three different supporting lines. The first two points determine the third, so there are $O_C(q^2)$ candidates. Each candidate is monochromatic with probability at most $m(2/m)^3=8/m^2$. Their expected number is $O_C(1)$; with probability tending to one there are at most $q^{1/10}$ such triples. Remove all their endpoints and colour them individually at cost $O(q^{1/10})$.

Choose a matching realization satisfying both successful events. By (C.11)–(C.12), recolouring the old deletion set costs at most

$$q^{(1+\vartheta)/2+o(1)}+q^{\vartheta+o(1)}=q^{(1+\vartheta)/2+o(1)}. \tag{C.14}$$

The surviving old colours are arcs: old triples were absent to begin with, mixed triples were removed through their old endpoints, and the all-new triples have just been removed. All repair palettes are disjoint from the old colours.

Apply this argument to the three lines of the exchange. A nonsquare old label d excludes old points on $y=\pm x$ and $y=0$. The real-invertible integer transformation $(u,v)=(x-y,x+y)$ makes the first two lines the zero axes of $uv=d$. Its image is a sublattice, for which full-lattice upper bounds remain valid. For the third line interchange x,y: the label becomes $-d$ and the exterior-line condition is $\chi(-(-d))=\chi(d)=-1$. The combined witness estimates therefore allow $\vartheta=1/2$.

Here $m=(q-1)/2$. Each diagonal has $q-1=2m$ nonorigin points, and P_L^0 has at most m points. The capacity conditions hold, all coordinates are $O(q)$, and the added cost is $q^{3/4+o(1)}$. This completes (C.5). For the main-grid product classes on a fixed number of zero-axis lines, the same argument with $\vartheta=0$ gives extra cost $q^{1/2+o(1)}=o(q)$, subject to the same per-line capacity.

C.7 The main-grid conflicts

Let $p=2h+1$ be an odd prime, $a\ge 1$, $2a\le h$, and $e\in\{0,1\}$. Put

$$t=2a-e,\quad n=p+t,\quad S=[-h-a,h+a-e]^2\cap\mathbb{Z}^2,\quad f(x,y)=(x+h+e)y\pmod p.$$

Then S is a translate of G_n. First consider nonzero old labels. In coordinates $(X,Y)=(x+h+e,y)$, they lie on $XY=c\neq 0$. A field line meets this curve at most twice. Each coordinate interval has fewer than $2p$ integers, so representatives of a single modular point form at most the four corners of a rectangle and cannot include three collinear points. Thus a same-colour real triple contains a repeated-image pair and a third point of a different image. The repeated pair has horizontal, vertical or diagonal direction. Horizontal and vertical directions are excluded by the one-intersection property of a nonzero product conic.

A repeated pair of slope $+1$ has the form

$$(X-h-e, Y-h-e), \quad (X+h+1-e, Y+h+1-e), \quad X, Y \in [e-a, a-1].$$

The other modular intersection has possible representatives $(-Y+jp, -X+jp)$. Since $p-a > h+a$, only $j=0$ lies in S. Writing this point as $z=(u,v)$ gives the triple

$$\{z, (-v-h-e, -u-h-e), (h+1-e-v, h+1-e-u)\}. \tag{C.15}$$

For slope -1, the repeated pair is $(X-h-e, Y+h+1-e)$, $(X+h+1-e, Y-h-e)$. The other representatives are $(Y+1-2e+jp, X-jp)$, again with only $j=0$ allowed. The triple is

$$\{z, (v-h-e, u+h+e), (h+1-e+v, -h-1+e+u)\}. \tag{C.16}$$

All centres lie in $U=[-a, a-e]^2$. With $H=h+1-e$, the right corner points are $F_H^+(z)$ and $F_H^-(z)$. The repeated-direction argument exhausts all nonzero same-colour triples. Hence removal of a set that contains each centre or both its specified right images makes every nonzero old colour an arc.

C.8 Choosing two primes for every integer side length

Fix $\eta \in (0,1)$. Given a sufficiently large integer n, take the least prime $p \geq 2n/(3-\eta)$, choose e for which $a=(n-p+e)/2$ is an integer, and set $H=(p-1)/2+1-e$. Take the least prime $q \equiv 3 \pmod 4$ with $q \geq H$, and put $b=H-q$.

The prime number theorem and its fixed-modulus arithmetic-progression version give

$$p=\frac{2n}{3-\eta}+o(n), \quad q=\frac{n}{3-\eta}+o(n), \quad b=o(q), \quad t=n-p=(1-\eta)q+o(q). \tag{C.17}$$

The needed form of the latter is stated, for example, in [6]. For any fixed $\epsilon > 0$ the interval $[x, (1+\epsilon)x]$ has asymptotically $\epsilon x/(2\log x)$ primes congruent to three modulo four. Allowing any fixed ϵ proves that the least subsequent such prime is $x+o(x)$.

Since $t=(1-\eta)q+o(q)$ and $h=q+o(q)$, eventually $2a \leq h$. The central square U has side length $t+1$, whereas U_q has side length q. Their difference is $\eta q+o(q)$, larger than the offset $|b|=o(q)$. More explicitly, eventually

$$q-(t+1) \geq 2|b|+2,$$

which ensures $U \subseteq U_q+(b,0)$ coordinate by coordinate, including the half-unit centre shift due to parity.

Let $T_b(x,y)=(x+b,y)$. Direct substitution gives the exact integer identities

$$F_H^+ T_b = T_b F_q^+, \qquad F_H^- T_b = T_b F_q^-.$$

Thus $T_b R_L$ has the required hitting property on the actual centres. Restricting it to S does not increase its chromatic number and still meets every actual triple, whose vertices all lie in S.

C.9 Paying for the zero label and completing the transfer

After this removal, the $p-1$ nonzero old classes are arcs. In product coordinates they lie in a fixed multiple of a p-square. The zero-labelled points lie on lines $x=kp$ or $y=kp$. Each coordinate interval has length less than $2p$, so there are at most four supporting lines. Each contains at most n points, and $n/p \to (3-\eta)/2 < 3/2$, so eventually $n \leq 2(p-1)$.

Apply the zero-axis absorption result with $m = p - 1$; it costs $p^{1/2+o(1)} = o(n)$ extra colours. Zero-labelled points already removed and recoloured need no further payment.

If $\chi_A(R_L) \le (\gamma + o(1))q$, the full budget is

$$\chi_A(G_n) \le p - 1 + o(n) + \chi_A(R_L) \le \left(\frac{2+\gamma}{3-\eta} + o(1)\right) n.$$

Take the upper limit for each fixed $\eta > 0$, then let $\eta \downarrow 0$. This proves the transfer coefficient $(2+\gamma)/3$ used in the main theorem.

References

[1] T. Lange and A. Winterhof, *Incomplete character sums over finite fields and their application to the interpolation of the discrete logarithm by Boolean functions*, Acta Arithmetica **101** (2002), 223–229. https://www.impan.pl/shop/en/publication/transaction/download/product/82707.

[2] T. Pham, Le Quang-Ham and B. Xue, *General-position sets and linear hypergraphs*, arXiv:2610.34254v1 (2026), preprint. https://arxiv.org/abs/2610.34254v1.

[3] T. Bayer, T. Mészáros, L. Rónyai and T. Szabó, *Subgraphs of the projective norm graph*, author manuscript (2021), Appendix, Theorem 6. https://page.mi.fu-berlin.de/szabo/PDF/SubgraphsPNG.pdf.

[4] V. Blomer, É. Fouvry, E. Kowalski, P. Michel and D. Milićević, *Some applications of smooth bilinear forms with Kloosterman sums*, author manuscript, p. 1. https://people.math.ethz.ch/~kowalski/complement.pdf.

[5] E. Szemerédi and W. T. Trotter, Jr., *Extremal problems in discrete geometry*, Combinatorica **3** (1983), 381–392. https://trotter.math.gatech.edu/papers/38.pdf.

[6] A. V. Sutherland, *Sieve theory and small gaps between primes: Introduction*, Oberwolfach lecture notes, 6 July 2015, p. 3. https://math.mit.edu/~drew/PrimeGapsOberwolfach1.pdf.