hexagon:2611.00582v1

ANTI-CONCENTRATION OF THE CHROMATIC NUMBER OF A DENSE RANDOM GRAPH

WILL HOWE, GRANT HUFFMAN, JENSEN KOHLMEYER, AND LIAM KRUER

Abstract. For $G_n \sim G(n, 1/2)$ and each fixed $0 < d < 1/4$, we obtain that the supremum over intervals of radius n^d of the probability that $\chi(G_n)$ lies in the interval tends to zero. This implies the failure of concentration on any fixed number of possible values and gives the eventual strict half-mass bound asked for in Erdős Problem 1156. The proof treats different supplies of large independent sets by matching sums and by one- and two-count Gaussian comparisons.

1. Statement and notation

Let G_n have vertex set $[n]$ and independent edges of probability $1/2$, and put

$$\mathcal{L}_n(r) = \sup_{t\in\mathbb{R}} \mathbb{P}\{|\chi(G_n) - t| \le r\}.$$

Theorem 1.1. *For each fixed $0 < d < 1/4$, the ordinary chromatic number of $G(n, 1/2)$ satisfies $\lim_{n\to\infty} \mathcal{L}_n(n^d) = 0$.*

This proves both assertions in the sense specified by Erdős Problem 1156 [1], as shown in the final section. The width results of Heckel [2] and Heckel–Riordan [4] give nonconcentration along subsequences. Here the required bound is for all sufficiently large orders.

Write $p = q = 1/2$, $L = \log n$, and

$$\alpha_0(n) = 2\log_2 n - 2\log_2\log_2 n + 2\log_2(e/2) + 1, \quad a = \lfloor \alpha_0(n) \rfloor, \quad \phi = \alpha_0(n) - a.$$

Let X_i count independent i-sets, with mean

$$\mu_i(x) = \binom{x}{i} q^{\binom{i}{2}}$$

at order x. In the interior let $b = a - 1$, $\mu = \mu_a(n)$, $m = \mu_b(n)$. Stirling's formula gives $\mu = n^{\phi+o(1)}$ and $m \asymp n\mu/a$, uniformly on compact phase intervals. The latter relation also follows by taking the ratio $a2^{a-1}/(n - a + 1)$. At a phase endpoint choose b by $\alpha_0(n) - b = 1 + o(1)$; then $m = n^{1+o(1)}$ and $\mu_{b-1}(n) = n^{2+o(1)}$.

1.1. Profiles and matching sums.

For a cap c, a coloring profile on x vertices has k_i classes of size $i \le c$, with $\sum k_i = j$ and $\sum i k_i = x$. Its product-graph first moment for unordered color classes is

$$\frac{x!}{\prod_i k_i! d_i^{k_i}}, \qquad d_i = i! q^{-\binom{i}{2}}. \tag{1.1}$$

Let $Z_c(y) = \sum_{i=1}^{c} e^{iy}/d_i$ and choose y by $(\log Z_c)'(y) = x/j$. The smooth log first moment is

$$L_0(x, j, c) = x\log x - x - j\log j + j + j\log Z_c(y) - xy. \tag{1.2}$$

Date: November 17, 2026.

2020 *Mathematics Subject Classification.* Primary 05C80. Secondary 05C15, 60C05.

Key words and phrases. Random graphs, chromatic number, anti-concentration, independent sets.

AI tools assisted with developing and checking arguments and with drafting this manuscript.

If $\pi_i = e^{iy}/(d_i Z_c(y))$, the log in (1.1) equals $L_0 - jD(\boldsymbol{k}/j\|\boldsymbol{\pi}) + O(a^2)$, where D is relative entropy. Stirling gives the error and the inequality from any profile. We use the rounding and profile estimates of Heckel–Panagiotou [3, Section 7] when $x \asymp n$, $c = \alpha_0(x) - O(1)$, and $x/j = c - \Theta(1)$. They give a rounded profile with log first moment $L_0 + O(a^2)$, top-class count $\Theta(n/a)$, a Gaussian tail in the gap $c - i$, and all its sizes within $O(\sqrt{L})$ of c. The number of profiles contributes at most $O(a^2)$ in the log.

For a family of large independent sets let I range over disjoint subfamilies, including the empty one. Denote by V_I the remaining vertices. If I has s size-b sets and t size-a sets, write $x = n - bs - at$, $j = K - s - t$. The log matching sum at color count K is the logarithm of

$$\sum_I \exp\{L_0(x, j, c)\}. \tag{1.3}$$

The cap is $c = b - 1$ at the endpoints and in lower phases. There we include size-b sets, and also size-a sets in the transition at $1/3$. For the fixed upper phases a version with cap b and a correction to L_0 is used in Section 4. A root is a zero of the log sum in its indicated bracket. We may assign any value on exceptional events of probability $o(1)$.

The derivatives of (1.2) are $\partial_x L_0 = \log x - y$, $\partial_j L_0 = \log Z_c(y) - \log j$. The terms involving differentiation of y cancel. In a bracket where the profile estimates apply, the derivative in K is $(1 + o(1))aL$. Inserting one size-u label into a term changes its weight, to constant factors, by

$$\frac{j}{\mu_u(x)} \frac{e^{uy}}{d_u Z_c(y)}. \tag{1.4}$$

The second factor stays between fixed positive constants for $u - c = 1, 2$, by the bounded centered tilt. For a size-b insertion, write $d_0 = bm/n$ and $q_I = (x)_b/(n)_b$, with $(x)_b = x(x-1)\cdots(x-b+1)$. Its weight multiplier is $\asymp (d_0 q_I)^{-1}$. The continuous derivative gives the discrete estimate since a step changes q_I by $1 + O(a^2/n)$. We use the cap first-moment location $x/j = \alpha_0(x) - 1 - 2/\log 2 + o(1)$ from [4, 3].

2. Planting and the lower-phase roots

Lemma 2.1. *In each of the lower phase ranges, the matching root has vanishing interval mass below the following scales.*

range	*size of a downward shift with probability* $1 - o(1)$
$\alpha_0(n) - b \to 1$	$n^{1/2 - o(1)}$
$0 < u \le \phi \le v < 1/3$	$n^{\sigma(\phi) - o(1)}$
$\phi \to 1/3$	$n^{1/3 - o(1)}$

Here $\sigma(\phi) = 0.38$ *for* $\phi \le 0.15$, *and* $\sigma(\phi) = (1 - \phi)/2$ *for* $\phi > 0.15$. *Each shift couples the original root with a root whose marginal law differs by* $o(1)$ *in total variation.*

Proof. For a fixed independent b-set, the conditional expected count of other such sets at intersection i is

$$\binom{b}{i}\binom{n-b}{b-i} q^{\binom{b}{2} - \binom{i}{2}}.$$

For $i = 2$ it is $n^{-1+\phi+o(1)}$ in the lower interior, and for $i = b - 1$ it is $n^{-1+o(1)}$. The sum of the intermediate terms is bounded using the ratio $2^i(b-i)^2/[(i+1)(n-2b+i+1)]$, which makes the interior no larger than a logarithmic factor times its endpoints. At an endpoint use $\phi = o(1)$ in these exponents. Thus $\operatorname{Var} X_b = O(m)$.

Planting one uniform size-b set means declaring its internal edges absent. The resulting graph law has density X_b/m relative to the original law, by averaging the conditional laws for individual sets.

The variation distance is $O(m^{-1/2})$. Applying the same kernels $h = \lfloor\sqrt{m}/L\rfloor$ times costs $o(1)$, by contraction of total variation. This process enlarges the label family and leaves the vertex count fixed.

We give the bounds on the matching sizes and the part of the sum used for insertion. If N_s counts matchings of size s in the size-b family, then exactly

$$\mathbb{E}N_s = \frac{m^s}{s!}\frac{(n)_{bs}}{(n)_b^s}. \tag{2.1}$$

At an endpoint put $\ell = 1 + |\log d_0| + \log b = o(b)$. At $s = \lceil 3n\ell/b^2\rceil$, the log of the falling-factorial ratio is at most $-(1 + o(1))b^2s^2/(2n)$. The other term is at most $s(\ell + O(1))$. Consequently all matchings have size at most $s_0 = \lceil 4n\ell/b^2\rceil$ with high probability, with room for one insertion. In this range $-\log q_I = O(\ell)$. The empty matching and $|Y| \le 2m$ for the label family Y bound the total log relative contribution at the empty-match root by $O(s_0\ell)$, since each insertion changes log weight by $O(\ell)$. The positive K-slope supplies a root in a bracket of width $O(s_0\ell/(aL))$, in which the class gap stays in the required compact range.

In the lower interior put $w = bs/n$, $A = 2/\log 2$, and $g(w) = w + (1 - w)\log(1 - w)$. The log of (2.1) is $n[(\phi/A)w - g(w)] + o(n)$ at a fixed positive w. Since $g(w) \ge w^2/2$, no $w \ge 0.701\phi + o(1)$ is possible with high probability. For $\phi \le 0.07$ this gives $w < 0.05$. Elsewhere in the lower interior, including approaches to $1/3$, the part of the sum with $w \ge 0.0499$ is exponentially small relative to its complement. Here are the estimates giving this claim.

Adjoining size b to the annealed profile changes the cap log expectation at the empty-match root by at most $O(n/a)$, by evaluating the Legendre minimum at the old tilt in (1.2). Thus Markov and the empty term bracket the actual root within $O(n/a^3)$. The annealed cap-b optimizer has proportion at most $1/26 + o(1)$ of its size-b classes. The limiting Gaussian on gaps $a - i \ge 1$ has mean $1 + A - \phi$. At the tilt centered at 3.5 its first six weights, in a common normalization, are $8, 32, 64, 64, 32, 8$. The omitted negative-side weights shift the mean by less than $4/208$. Since the required mean is larger, its tilt gives mass at gap 1 at most $8/208$. The entropy loss in (1.2) therefore bounds the log annealed large-matching part below the optimum by $\Omega(n/a)$, by Pinsker's inequality.

For the opposite bound on the small-matching part, take the rounded optimal top size s_1 and consider $T = m + O(\sqrt{m}L)$ independent uniform indexed labels. Its matching count has expectation differing from (2.1) by $(T)_{s_1}/m^{s_1} = \exp(o(n/a))$. Put $P_j = (n)_{bj}/(n)_b^j$. The exact relative second moment is $\mathbb{E}P_J^{-1}$, where J is the hypergeometric intersection size of two s_1-sets of indices, of mean $S = s_1^2/T$. Indeed, conditional on shared indices, fixing the shared disjoint b-sets leaves disjointness probability P_{s_1}/P_J independently on each side. For $J \le 10S$ the log cost is $O(b^2S^2/n) = n^{1-2\phi+o(1)}$. The rest is bounded by the inequality $\mathbb{P}(J \ge j) \le (eS/j)^j$ and $-\log P_j \le 1.05b^2j^2/(2n)$. The function $\log(j/S) - 1.05b^2j/(2n)$ is concave in this interval. At $10S$ its value exceeds 2 for large n, and at s_1 the first term is $(\phi + o(1))L$ and the subtracted term is at most $0.061L$. These estimates apply when $\phi \ge 0.07$. Thus Paley–Zygmund gives probability $\exp(-n^{1-2\phi+o(1)})$ of attaining the needed small-matching mass within $o(n/a)$ in the log. Changing one label costs $O(L)$ in the log by (1.4). Bounded differences with variance proxy $O(TL^2)$ place the mean within another $n^{1-\phi/2+o(1)} = o(n/a)$, and make a fixed positive loss of order n/a exponentially unlikely.

The transfer of this bound to the actual family costs $n^{2\phi+o(1)}$ label edits. To see the scale, couple the indicators to independent indicators one at a time. The previous indicators on disjoint pair supports are jointly independent of the current indicator. The discrepancy is bounded by twice the sum over prior neighbors of their joint success probability and the product of the individual probabilities. Summing the intersection bounds gives $O(m^2a^4/n^2 + ma^{O(1)}/n)$. Conditional on the independent count, the labels are uniform; repetition in indexed sampling has negligible probability. The log edit cost $n^{2\phi+o(1)}O(L)$ is $o(n/a)$ in the stated range. Markov for the large-matching part and the small-matching lower bound hold simultaneously over integer K in the bracket; the slope extends the result to real K. Including

size-a labels at the transition introduces a factor at most $\exp(O(X_a L)) = \exp(n^{1/3+o(1)})$, which leaves the $\Omega(n/a)$ margin.

Now take K at the initial root. At each stage of the planting, a uniform proposed label is disjoint from a given old matching with probability q_I. In (1.4) this cancels q_I^{-1}. Therefore the relative increase U from extending old matchings, retaining the small-matching portion when appropriate, has conditional mean $\gg d_0^{-1}$. Duplicates change this by a negligible amount because their probability is at most $2m/\binom{n}{b}$. The increase is bounded by $O((d_0 q_{\min})^{-1})$, using the smallest disjointness probability among the matchings used. The chord bound for $\log(1+U)$ gives conditional mean $\gg \min(d_0^{-1}, q_{\min})$. At the endpoints this is $n^{-o(1)}$. When $\phi \le 0.07$, $-\log q_{\min} \le (2.2\phi + o(1))L$ by $w \le 0.701\phi + o(1)$. For the remaining small-matching range $w < 0.05$, $-\log q_{\min} \le (0.15 + o(1))L$, since $-A\log(0.95) < 0.15$.

The good bounds hold throughout the monotone trajectory on the initial and terminal good events, by bounding small mass from below by its initial value and large mass from above by its terminal value. Stop the lower-increment estimates at the first failure and thereafter substitute their deterministic drift. For increments divided by a deterministic $O(L)$ upper bound, $e^{-D} \le 1 - (1-e^{-1})D$ on $0 \le D \le 1$. Iterating the conditional estimate makes a total gain smaller than a fixed fraction of the sum of the drifts unlikely. At the endpoints the gain in the log is $n^{1/2-o(1)}$; in the lower interior and transition the exponents are those in the statement, after dividing by the uniform slope $O(aL)$. In particular when $\phi > 0.15$ the exponent is $(1+\phi)/2 - \phi = (1-\phi)/2$.

If a variable Z' has law within $o(1)$ of Z in total variation and $Z' \le Z - h_n$ with probability $1 - o(1)$, comparison of their distribution functions bounds the mass of every interval of length less than h_n by $o(1)$, using left limits for the lower endpoint. Apply this with the two coupled roots. □

3. From the lower-phase roots to ordinary coloring

We use the following accuracy estimates for the roots of (1.3). Write K_{mat} for the single-size root and K_{mix} for the two-size root. At a phase endpoint,

$$|\chi - K_{\mathrm{mat}}| \le n^{\epsilon}$$

with probability $1-o(1)$ for each $\epsilon > 0$. In a compact lower-interior phase interval, the same comparison holds with error $n^{\gamma(\phi)}$, where $\gamma(\phi) = (\max(\phi, 0.15) + \sigma(\phi))/2$. At the transition $\phi \to 1/3$, the comparison is $|\chi - K_{\mathrm{mix}}| \le n^{1/4}$ with probability $1 - o(1)$. We describe the estimates for proper profiles that enter these comparisons.

3.1. **Conditional first moments.** Condition on the actual entire size-b label family. Let H be the graph of pairs forced absent by these labels. The conditional law is the product law on the remaining pairs, conditioned on the increasing event that every other b-set fails to be independent. For a proposed partition π, with f pairs required absent inside its classes and h_π of those pairs in H, Harris gives the bound q^{f-h_π} on its conditional probability. Summing over partitions uses the slot average of 2^{h_π}. One expands

$$2^{h_\pi} = \prod_{e \in H} \bigl(1 + \mathbf{1}_{\{e \text{ is internal to } \pi\}}\bigr).$$

If a connected marked component has j vertices, it must lie in a single class. The slot factor for disjoint components of sizes $j_1, \dots, j_t$ is bounded by $\prod_\nu k(c)_{j_\nu}/(x)_{\sum j_\nu}$ for k classes. Retaining the denominator matters for large components.

At the endpoints one may remove the vertices of labels sharing a pair and of triples with three separate intersection vertices. Their number is $n^{o(1)}$, by the pair count and the triple bound $O(m^3 b^6/n^3)$ for pair-disjoint supports. Elsewhere the intersections $3, \dots, b-2$ are absent with high probability by the intersection ratio in Lemma 2.1. Pairs meeting in $b-1$ vertices form isolated hulls of size $b+1$, and components joined by a shared pair have size at most two for $\phi < 1/3$ with fixed margin, or at

most three for $\phi \to 1/3$. Indeed an ordinary intersection-two pair has total expected count $n^{2\phi+o(1)}$; each additional distinct ordinary label meeting one of a bounded number of those labels in a pair has conditional contribution $n^{-1+\phi+o(1)}$. Hence triples are excluded if $3\phi < 1$ with fixed margin, and quadruples if $4\phi < 2$ with fixed margin. The same count for a close hull begins with expected number $n^{\phi+o(1)}$. An outside label meeting the hull beyond one vertex either has bounded intersection with it or extends its $(b-1)$-core. These contributions are $O(mb^4/n^2 + nq^{b-1})$ per hull. For constraints involving distinct components, one first restricts to disjoint required pair sets before using their product probabilities.

The following counts control the connected-component expansion. The maximum degree is $n^{\phi+o(1)}$ in the lower interior and transition, by the factorial-moment estimate $(O(mb/n))^r/r!$ for r distinct components at a vertex. For a base label or close hull, outside adjacency into it and common neighbors of an edge away from its supporting bases have bounded size with high probability. To bound a large family of such witnesses, either fix one supporting base and prescribe two vertices of a further base, costing $n^{-1+\phi+o(1)}$, or prescribe two endpoint vertices and count two joined partner bases, costing $n^{-1+2\phi+o(1)}$. Ten witnesses with distinct components suffice for a union bound in either case. Ambiguous and missing hull pairs have bounded degree by the same factorial count for nontrivial components through a fixed point. If M is the largest base intersection of a j-set, summing codegrees gives

$$2e_H^2/j - e_H \le Me_H + O(j^2), \qquad e_H \le Mj/2 + O(j^{3/2}).$$

Away from a large base intersection these estimates, together with $n(e^2\Delta_H)^{j-1}$ connected supports and the factor x^{-j}, give geometric decay for growing j. For bounded j the activity is at most $n^{2-j+(j-1)\phi+o(1)}$. Near a single base, adding $r = j - M$ outside vertices costs at most $(O(bj\Delta_H/x)2^{r/2+O(1)})^r$; the inside activity is bounded by

$$k\,|Y|\binom{b}{M}(c)_M x^{-M} 2^{\binom{M}{2}}.$$

For $M = b - h$, $h \ge 1$ fixed, this is $n^{1-h+d(w)+o(1)}$, where $d(w) = -A\log(1-w)$; for bounded M it is $n^{2+\phi-M+o(1)}$. The ratios $2^M(b-M)(c-M)/[(M+1)x]$ bound the intervening sizes by their endpoint contributions.

Large components use at least $b - 20$ vertices of a base. By (2.1) applied to shortened labels, a disjoint collection of them covers at most $(0.24 + o(1))n$ vertices, and at the endpoints it covers $o(n)$. Marking 21 deleted vertices per base costs $\exp(O((n/b)\log b)) = \exp(o(n))$, which does not change the large-deletion first-moment bound. The falling-factorial correction for z_1 large-component and z_2 other vertices is at most $\exp(xg(z_1/x) + 2z_2)$. For $x \ge 0.95n$ the first exponent is at most $0.15z_1$. Charging e per large component bounds the contribution with $z_1 > n^{4/5}$ by $\exp(-n^{4/5-o(1)})$ times the exponential series of activities. For smaller z_1 the correction coefficient is $O(n^{-1/5})$, and then marking all components restricts to total occupancy at most $n^{2/5}$ with negligible error. In this range the correction is $1 + o(1)$. These bounds give log slot moment $n^{o(1)}$ at the endpoints and $n^{\max(\phi,d(w))+o(1)}$ in the fixed lower interior at $w < 0.05$. At larger w, a union bound over all residues using the independent-label count and the label edit bound restricts the nearly contained pool to $n^{\max(1,1+\phi-d(w))+o(1)}$. With $z_1/x < 0.314+o(1)$ the denominator correction is at most $e^{0.183z_1+2z_2}$, giving log error at most $n^{0.99+o(1)}$, absorbed by the $\Omega(n/a)$ loss in the matching sum.

At the transition the leading selected-edge activity has to cancel. Write $N_x = \binom{x}{2}$ and $\kappa = m\binom{b}{2}/\binom{n}{2}$. On the exact label event, forcing the cap pairs in π gives the upper probability

$$q^{f-h_\pi}\exp[-\kappa(f-h_\pi) + o(1)]. \tag{3.1}$$

The extra term is the first-order change of the avoidance probability of other size-b labels when $f - h_\pi$ new pairs are absent. For a candidate meeting one newly forced pair it has multiplier q^{-1}. The

multiple-pair terms are bounded by the connected expansion just described, using the bounded off-base codegrees for the three-vertex supports at marked pairs. A cap class containing a missing close-hull pair has conditional probability zero. For the remaining terms the changed-candidate overlap error gains $n^{-1+o(1)}$ for the prescribed pair with an exterior vertex. These estimates give the absolute $o(1)$ error in (3.1).

The exact size-two activity in the slot sum is $\lambda_V = f|H[V]|/N_x$ up to $o(1)$. Uniformly over residues $\big||H[V]| - \kappa N_x\big| \le n^{1+\phi/2+o(1)}$. This follows by Hoeffding for the independent indexed family, with each summand bounded by $\binom{b}{2}$, followed by the union bound over 2^n residues and the label edit and duplicate-pair bounds. Thus the leading term cancels $-\kappa f$ within $n^{1/6+o(1)}$ in the log. The higher connected activities at $w < 0.05$ contribute $n^{d(w)+o(1)}$, where $d(w) < 0.15$. Both errors are smaller than the log deficit at $K_{\mathrm{mix}} - n^{1/5}/2$.

At the negative thresholds in all three ranges the slope supplies a log deficit $\Omega(aL\mathcal{E})$, where $\mathcal{E} = n^{\epsilon}, n^{\gamma(\phi)}, n^{1/5}$, respectively. The error estimates above are smaller than this deficit. Markov excludes the tested color count. Smaller colorings can be refined, and the allowance for splitting exceptional larger classes is $n^{o(1)}$ at an endpoint, $n^{\phi+o(1)}$ in the fixed lower interior, and zero at the transition, where size-a classes are included and size $a+1$ is absent with high probability. This gives the lower comparisons to ordinary coloring.

3.2. **Weighted second moments on the residues.** At a positive threshold choose one pair of matching sizes (s,t) accounting for the largest term in (1.3). There are polynomially many possibilities. If $N_{s,t}$ counts its matchings and E_p is the product expectation for the rounded cap profile on each residue, then $\log(N_{s,t}E_p) \ge c_1 aL\mathcal{E}$ for some $c_1 > 0$. Here $\mathcal{E} = n^{\epsilon}, n^{\gamma(\phi)}, n^{1/4}$ in the three ranges. At the endpoints, trimming labels that share a pair costs $n^{o(1)}$ in the log. At the transition the close and higher-component trim costs $n^{0.15+o(1)}$, from the per-label insertion bound $O(d_0^{-1}n^{d(0.05)+o(1)})$ and $n^{1/3+o(1)}$ trimmed labels. Thus the margin survives the trims. The conditioning on the actual labels retains the entire original family.

We record the avoidance calculation used to compare weighted moments. In a product Bernoulli law let an event A_i require all coordinates of a specified support to be one, and declare two events adjacent when their supports intersect. Let $p_i = \mathbb{P}(A_i)$ and Δ_h be the sum of joint probabilities of connected h-tuples without repetitions, counted without ordering. If $\max_i \sum_{j\sim i} p_j \le 1/2$ and $\max_i p_i \le 1/2$, then

$$\log \mathbb{P}(\text{no } A_i) = -\sum_i p_i + \sum_{j<i,\ j\sim i} (\mathbb{P}(A_iA_j) - p_ip_j) + O\left(\sum_i p_i^2 + (\max_i p_i)\Delta_2 + \Delta_3\right). \tag{3.2}$$

To derive it, divide the predecessors of i into T, its neighbors, and U, its nonneighbors. With D_J the event avoiding J, its hazard is $p_i\mathbb{P}(D_T \mid A_i, D_U)/\mathbb{P}(D_T \mid D_U)$. Bonferroni expands the numerator as $1 - \sum_{j\in T}\mathbb{P}(A_j \mid A_i)$, and the denominator as $1 - \sum_{j\in T} p_j$, to the required accuracy. The error in each single-neighbor term comes from the members of U adjacent to that neighbor: condition first on all other members of U, then use the union bound and Harris for the remaining ones. After multiplying by p_i these errors and the Bonferroni pair errors are bounded by connected triple probabilities. Expanding the quotient adds only connected triple terms and the diagonal terms $O((\max p_i)\Delta_2)$. The hazard is at most p_i by Harris, so summing the log approximation costs $O(\sum p_i^2)$. This proves (3.2).

For the endpoint and fixed lower-interior product comparison, let $c = b-1$, k_c be the number of top classes, $Q_c = \binom{c}{2}$, $\beta = k_c/\mu_c(x)$, $S = k_c^2/\mu_c(x)$, and $N_x = \binom{x}{2}$. If M_V is the number of edges on the residue V, use the weight

$$W_V = \exp\big(\theta(M_V - pN_x) - \beta(X_c(V) - \mu_c(x))\big), \qquad \theta = (f - k_cQ_c)/(qN_x).$$

It is increasing in present edges. Attach to each c-set a private tag of probability $z_A = 1 - e^{-\gamma_A}$, where γ_A is the sum of the exponents of all weights containing that set. The penalty on its independence indicator

is precisely the probability of avoiding its active event. Put $L_z = \sum_A z_A q^{Q_c}$, $V_e = \sum_{A \supset e} z_A q^{Q_c}$. When a pair set R is forced absent and the unforced odds receive a tilt θ_e, the log of the tag expectation is bounded above by

$$-L_z + p \sum_e \theta_e V_e - \frac{p}{q} \sum_{e \in R} V_e + \frac{p}{2q} \sum_e V_e^2 - \sum_{A \text{ fully forced}} \gamma_A + o(1). \tag{3.3}$$

For one partition this is an equality to the stated error. For two partitions the upper estimate retains generic candidates having at most one forced pair, and the fully forced top classes. Each fully forced candidate is contained in a row or column of the overlap array, so its tag is counted exactly. The mean of a generic candidate changes linearly from forcing a pair or perturbing its probability. The error bound for the aggregate second-order terms is $O((\mu_c(n) \max z_A) a^6/n^2)$. Among pairs sharing two vertices the contribution is the covariance term in (3.3); the forced-pair edits change it by a relative $O(a^3/n)$. For other intersections, the bound $(a^2/(n-a))^i q^{-\binom{i}{2}}$ handles moderate i, and for overlap $c - r > 3c/4$ the conditional extension bound is $\max z_A \binom{c}{r} n^r q^{r(2c-r-1)/2}$. In the triple bound, a third candidate meeting the union of two small-overlap seeds either has r outside vertices or has points on both exclusive sides. Its unforced pairs give factors $q^{r(2c-r-1)/2}$ and q^{ij} if the exclusive counts are i, j. With both sides containing at least $c/8$, the latter gives exponential loss in c^2. Otherwise orient toward the larger side and count choices by $c^{r+2j+O(1)} n^r$, with the larger exclusive side of size at least $c/4 - O(1)$ when the outside and smaller counts are each below $c/8$. This gives a power saving for $r + j \geq 1$. A close starting pair with union within eleven vertices of a top set has at most a polylogarithmic extension bound, and the tail of more distant close pairs beats $n^{4+o(1)}$ for an unrestricted union extension. At $w < 0.05$ in the fixed lower range, the ambient bounds are $\mu_c(n) \max z_A \leq n^{1.15+o(1)}$ and $\mu_c(n)(\max z_A)^2 \leq n^{0.24+o(1)}$. They leave a strict power saving in these estimates. For one partition the omitted non-forced candidates with more than one forced pair have active mean $o(1)$. For near-full intersection with one class, the choices of $r \geq 1$ outside vertices must pay $q^{r(c-r)}$ and a tag. For moderate intersections use the connected expansion of internal class occupancies, beginning with two forced pairs rather than one. Harris then restores them with relative error $1 + o(1)$. This gives (3.3) by (3.2).

The edge tilt and the last quadratic edge term in (3.3) combine into

$$-p \sum_{e \in R} (\theta_e + V_e/q) + \tfrac{pq}{2} \sum_e (\theta_e + V_e/q)^2.$$

On a single residue $V_e = \beta \mu_c(x) Q_c / N_x$ up to relative $O(\beta)$, with the same formula proportional to the power of the weight. The effective tilt is $\Theta = f/(qN_x)$. A single proper partition weighted by W_V^ℓ, $1/2 \leq \ell \leq 2$, therefore has expectation

$$q^f \exp\left[\left(\frac{pf^2}{qN_x} + S\right)(\ell^2/2 - \ell) + o(1)\right].$$

For two residues V, V' write $C = V \cap V'$, $f_{i,C}$ for each partition's internal pairs in C, $k_{i,C}$ for its top classes contained in C, j_c for their identical top classes and g for their common internal pairs. Relative to the square of the single mean, the two-term upper bound is q^{-g} times the exponential of

$$-p\Theta(f_{1,C} + f_{2,C} - 2g) + pq\Theta^2 \binom{|C|}{2} + \beta^2 \mu_c(|C|) - \beta(k_{1,C} + k_{2,C} - 2j_c) + o(1). \tag{3.4}$$

Indeed the joint tag mean has a cross term $-\beta^2 \mu_c(|C|)$ in L_z. The forced-class correction removes the duplicated classes once, and includes both tag weights on each fully forced top class lying in C. The cubic error $\mu_c(x)\beta^3 = o(1)$ is uniform. Truncating at $\log W_V \leq \sqrt{S + a^2} L$ loses $o(1)$ of the weighted single mean by the single-power formula at $\ell = 1 + (S + a^2)^{-1/2}$.

For clarity, the slot count from two partitions is as follows. Restrict to the mutually relevant pairs of Heckel–Panagiotou [3, Sections 4–6], after singleton completion off each residue. Their conditions hold

with weighted probability $1 - o(1)$: the witness bounds use the Gaussian profile tail and $O(a^2)$ pairs per witness, so an $O(a/n)$ change in edge probabilities preserves them. The conditions are increasing in the free present edges, and Harris preserves their probability under the additional tags. Apart from $O(L^3)$ exceptional rows, a nonidentical row has only $O(\sqrt{L})$ vertices in cells of size at least two, by the composition bound. For an indicated collection of such cells of sizes i_ν, its slot factor is bounded by

$$\frac{\prod_\nu \binom{d_{\mathrm{row},\nu}}{i_\nu}(d_{\mathrm{column},\nu})_{i_\nu}}{(x)_{\sum_\nu i_\nu}}.$$

If selected cells share a column, $(d)_{\sum i_\nu} \le \prod (d)_{i_\nu}$ justifies the bound. On sparse pairs with identical fraction below L^{-3}, the total number of vertices in nontrivial cells is $o(x)$. Charging e per exact cell bounds the range above $n^{1/10}$ by $\exp(-n^{1/10-o(1)})$. This beats $S \le n^{0.09+o(1)}$ in the fixed lower interior, and $S = n^{o(1)}$ at the endpoints. Below the cutoff the whole falling denominator is $x^{\sum i_\nu}(1 + o(1))$. Expand the weight q^{-g} in nonnegative cell increments. The size-two coefficient is p/q, and an identical c-cell includes the extra factor $e^{2\beta}$. The activities sum to $(p/q)f_{1,C}f/N_x$ for size two and $k_{1,C}k_c/\mu_c(x) + o(1)$ at size c. Intermediate sizes contribute $o(1)$, by the ratio $2^i(c-i)^2/[(i+1)x]$ and the bounds at sizes 3 and $c-1$. A permutation transposition changes the local pair counts by $O(a)$ and the local top-class counts by at most two. Deviations beyond $n^{3/5}$ have probability $\exp(-n^{1/5-o(1)})$, and Hölder with moment $1 + a^{-2}$ absorbs the exceptional contribution. Replacing the local counts in (3.4) by these typical values makes the two cell activities cancel the negative edge and tag contributions. The relative sparse moment is at most $1 + o(1)$ uniformly on the two residues.

For intermediate overlaps, if ℓ_i identical classes of size i have total size h, extracting those slots costs at most

$$\frac{1}{(x)_h}\prod_i \frac{(k_i)_{\ell_i}^2}{\ell_i!}(i!q^{-\binom{i}{2}})^{\ell_i}. \tag{3.5}$$

Every residual nonexceptional nontrivial cell has size $O(\sqrt{L})$. The same slot bound pays $\exp(O(L^5))$ for the total residual cell weight including the exceptional rows. When h/x is small but at least L^{-3}, [3, Lemmata 3.5–3.6] bound (3.5) by $\exp(-\Omega(h))$ after class selection. At the endpoints the compact middle overlaps follow from [3, (7.33)]. For the fixed lower range, the compact middle uses their Lemma 3.3 and the actual residue mean. Write $\tau = 1 - h/x$, and measure gaps from cap c. The actual mean and reference mean lie in $[1.55, 2.04]$; the tilted gap law has mass at zero at most $0.27 + o(1)$ and excess means above $5, 4, 3$ at most $0.0016, 0.019, 0.135$, up to $o(1)$. These follow by bounding the discrete Gaussian tilt by centers 1 and 2; at center 2 the successive terms for those excess tails are bounded by geometric ratios $0.09, 0.177, 0.354$ from first terms $2^{-8}, 2^{-9/2}, 1/4$, divided by a normalizer above 2.9. The full exponent per vertex is at least $-0.005 + o(1)$ from $[g(w) - (\phi/A)w]/(1-w)$. With omitted mean gap $\bar{\imath}$ and reference gap D^*, the added partial-profile exponent is $\tau[-\log\tau - (\bar{\imath} - D^*)/A]$. Apply $\bar{\imath} \le r + C_r/\tau + o(1)$ on $[0.01, 0.1]$, $[0.1, 0.2]$, $[0.2, 0.4]$ for $r = 5, 4, 3$ respectively. The right-end lower bounds on the bracket are $1, 0.68, 0.17$, each sufficient after multiplication by the left endpoint. For $h/x \le 0.6$ the lower bound

$$-\tau\log\tau - \frac{2.04}{A}(1-\tau) + \frac{(1-\tau-0.27)_+}{A}$$

is positive with a linear margin at $1 - \tau = 0$, and concavity applies between $1 - \tau = 0.27$ and 0.6. This gives the compact exponential loss. At τ sufficiently small, the remaining class mean exceeds its new reference by order $\log(1/\tau)$, since the Gaussian gap moment controls the leftover gap by $\log_2(1/\tau) + O(1)$, while the reference decreases by $2\log_2(1/\tau) + O(1)$. Factoring (3.5) through the residual complete profile and applying [3, Lemma 3.4] then gives the loss $\exp(-\Omega(\tau x))$ after class selection.

For near-identity pairs one uses this extraction together with the placement of missing vertices. If $D = |V \setminus V'|$, every such vertex must be in the leftover slots of the first partition. The factor is at most $(\tau x)_D/(x)_D \le \tau^D$. When $\tau < n^{-1/100}$, a direct row bound gives $\exp(-\Omega(DL))$. Each missing vertex outside a scrambled row is a required exceptional choice, costing $n^{-1+o(1)}$. Regrouping s' scrambled rows costs at most $(2^c (s')^c k)^{s'}/s'!$ before lost-pair weights. The loss of pairs supplies exponent at most $-(\log 2)b^2 s'/2 + O(s'b\sqrt{1+\log(k/s')})$ by the Gaussian row-gap bound. This pays $\Omega(s'a\log(k/s'))$ in the net exponent, including the vertices missing in these rows. At an endpoint use the same count for $\tau < \sqrt{\ell/b}$; then the loss $\Omega(D\log(1/\tau))$ suffices. Keeping one truncated weight pays at most $\exp(n^{o(1)})/E_p$ at the endpoints, or $\exp(n^{0.09+o(1)})/E_p$ in the fixed lower regime, before the missing-vertex penalty.

To sum the penalty over different matchings, first trim the close hulls and at the endpoints trim all pair-overlapping labels. At an endpoint two matchings differing in r labels per side have at least $br/2$ vertices of either residue missing from the other. The expected number of failures is bounded at each r by

$$(em/r)^{2r} 2^{br} \left(\frac{br/n}{1-br/n}\right)^{br/2}.$$

The two varying unions in the bare count are independent uniform br-sets. Only edge-disjoint deterministic labels are used when taking product probabilities, and then this restriction is dropped in the numerical count. The sum tends to zero by $br/n = o(1)$ and $\log d_0 = o(b)$. In the fixed lower case the same argument with $br/n \le 0.05$ pays at most 2^r for possible intersection-two partnerships. The joint-overlap exponent per covered vertex is below -0.77 at 0.05 before the label entropy, which contributes at most $(\log 2)\phi + (2/b)\log(n/(br)) + o(1)$. The coefficient of $\log(n/(br))$ remains negative. The loss $br/4$ suffices after allowing the $n^{\phi+o(1)}$ close labels when the distance exceeds a fixed multiple of that number. Smaller distances cost $\exp(n^{\phi+o(1)})$. In particular $0.01^{br/4}$ pays more than $3.3rL$, which beats the $2r\log|Y|$ choice bound. The near-identity total divided by $N_{s,t}E_p$ is $o(1)$ at the positive threshold, by the log margin and $\gamma(\phi) > \max(\phi, 0.15)$.

At the transition the moment comparison takes place directly in the graph conditioned on the actual labels. Use the edge-count weight with tilt $f/(qN_x)$ and restrict to profiles whose hole edges are separated in their classes, with at most one internal hole edge per class or base and at most $n^{2/5}$ in total. A single weighted mean is

$$(1+o(1))q^f e^{-pf^2/(2qN_x)} 2^{h_\pi} e^{-\lambda_V}.$$

In the slot expansion of 2^{h_π}, the retained edges are disjoint and use distinct classes. Up to $n^{2/5}$ edges the slot denominator is its power times $1+o(1)$. The exponential series of these edges has mean λ_V, and in its independent sampling the probabilities of two choices sharing a class, an endpoint, or a base are bounded by $O(n^{4/5}a/n)$, $O(n^{4/5}\Delta_H/|H[V]|)$, and $O(n^{4/5}mb^4/|H[V]|^2)$. All vanish. Finishing the uniform placement costs $o(1)$ in relative mass: one-endpoint and occupied-class additional holes contribute $O(n^{2/5}\Delta_H a/n + |H|n^{2/5}a^2/n^2)$, and two additional holes in one class contribute $O(n\Delta_H^2 a^2/n^2 + |H|^2 a^3/n^3)$. The removed missing pairs have probability $O(a/n)$ each of occupying a common class. These estimates give the single partition normalization $e^{\lambda_V + o(1)}$ in the slot average.

The sparse pair bound uses $j \ge 3$ cells of total activity $o(1)$ since $\mu_c(x) \ge n^{2.18}$. In the marked-hole expansion the size-two terms are the ordinary overlap activity $(p/q)f_{1,C}f/N_x$ and $\lambda_{V'}$. A mixed marked component containing a hole and a nontrivial within-row cell of size at least three has vanishing activity, by the same slot ratios and the cap of one marked pair per row. The exponent $-pf^2\binom{|C|}{2}/(qN_x^2)$ from the pair edge tilt cancels the ordinary overlap activity after typical-count replacement. The one-partition normalization cancels the remaining λ-terms. Charging e per cell cuts occupancy above $n^{9/20}$ at cost $\exp(-n^{9/20-o(1)})$, which absorbs the crude extra $\exp(O(n^{2/5}))$ hole factor. The intermediate calculation uses (3.5) with the same Gaussian estimates, here with upper reference gap 1.701. It gives

$\exp(-\Omega(n/L^3))$ through identical fraction 0.99. At τ between $n^{-.01}$ and .01, conditioning that the D missing vertices lie in leftover slots changes the normalized hole factor by at most $\exp(Dn^{-.65+o(1)} + o(1))$. To see this, edges with neither endpoint missing change total activity by $O(D\lambda_V/x)$; those with one or two missing endpoints cost $O(D\Delta_H c/(x - D) + D\Delta_H c/(\tau x))$. All denominators may be replaced by powers for the retained $O(n^{2/5})$ marked edges. Thus the factor τ^D remains sufficient. Below $\tau = n^{-.01}$ the same row regrouping pays $\exp(-\Omega(DL))$ including the additional hole marks in nonidentical rows. Mixed matching separation gives $D \geq br/2$ for distance r per side after trimming: the product count of differing unions uses the overlap bound above, and the size-a unions have $o(n)$ total vertices and only boundedly many intersections per seed apart from the trimmed components. The resulting factor $.01^{br/2}$ already beats the $\exp((8/3 + o(1))rL)$ choices. Truncation keeps the other weight at cost $\exp(n^{1/5} + O(L^5))/E_p$, which is absorbed by $\log(N_{s,t}E_p) \gg aLn^{1/4}$. Thus the relative second moment is $1 + o(1)$ in the conditional graph, and appending the known larger classes gives the upper transition comparison.

3.3. **Repair in the product comparison.** At the endpoints and the fixed lower interior, the weighted count has relative second moment $1 + o(1)$ in the independent product graph. Biasing the graph by this count changes its law by $o(1)$ in total variation, by Cauchy–Schwarz. On this biased graph sample a term proportionally to its weight. Its marginal is uniform over matching choices and profile slots because the product term normalizers agree by permutation symmetry. Place the known label pairs absent, and couple the remaining product graph to the exact label-conditioned graph. The increasing conditioning exposes each next present edge with probability $p/(1 - qt)$, where t is the conditional chance that replacing an absent edge by a present edge is required to avoid a further label. The same uniform random variable for both edge exposures gives a monotone coupling. The probability of an added edge is at most qt. An added edge on a chosen color class requires some other candidate label to have all its other pairs absent. Harris bounds this probability, also under the increasing single-term weight on the free coordinates before truncation. The truncation loses relative $o(1)$.

For a candidate S and a partition π, let $r_\pi(S)$ be its number of internal class pairs. The union bound for the repair mean is

$$\sum_{|S|=b} 2^{-\binom{b}{2}+e_H(S)}\mathbb{E}_\pi[r_\pi(S)2^{r_\pi(S)}].$$

The first sum of probabilities is $O(m)$ by the same connected expansion with slot factor $(b)_j/n^j$; whole-base supports now contribute a bounded activity. In the slot average for fixed S, restrict first to occupancies at most $3b/4$. The marked-pair activity is $O(b^3/x)$, higher activities are summable, and $r2^r \leq 2b^2(2^r - 1)$. A class containing $c - h > 3b/4$ is unique. The $h = 0$ contribution is at most $O(kb/\mu_c(x))$, and each $h > 0$ gains the factor $(bx\,2^{-(c-h-1)})^h$. Thus the slot average is at most $n^{-1+o(1)} + n^{1+o(1)}/\mu_c(x)$. The repair mean is $n^{o(1)}$ at an endpoint and $n^{\max(\phi,d(w))+o(1)}$ in the lower interior. Markov shows that there are fewer than half the allowed additional colors in damaged pairs with high probability. One fresh color per pair suffices. Appending the selected independent classes now gives the claimed upper comparisons.

By Lemma 2.1, the endpoint error can be taken smaller than any fixed window exponent below $1/2$. In the lower interior $\sigma(\phi)$ exceeds both $\gamma(\phi)$ and $1/4$ by a fixed margin on any compact interval. At the transition the shift exponent $1/3$ exceeds $1/4$. Enlarging a chromatic interval by the corresponding comparison error gives the endpoint bound for every $d < 1/2$ and the lower-interior and transition bounds at $d = 1/4$.

4. The fixed upper phases

Fix $1/2 < u \le \phi \le v < 1$. In this range

$$\chi(G_n) = K_n - \frac{1-\phi}{a}(X_a - \mu) + o_{\mathbb{P}}\big(\sqrt{\mu}/a\big) \tag{4.1}$$

for a deterministic K_n. We give the construction of the center and the comparison for the actual graph.

4.1. **Largest-set planting.** For a specified a-set conditional on its independence, the expected number of other independent a-sets meeting it in at least two vertices is $O(\mu a^4/n^2)+n^{-1+o(1)}$, by the intersection estimate in Lemma 2.1. Multiplying by μ gives $o(1)$. The sequential indicator comparison therefore couples the entire size-a label family with independent indicators with probability $1-o(1)$. Their count is within $o(1)$ total variation of a Poisson variable of mean μ.

The same comparison can be made jointly with the graph. Fix a proposed family Y of at most 2μ distinct labels. Call an a-set neutral if it uses none of the pairs forced by Y. Avoidance of neutral sets is an increasing event independent of the pairs to be forced. By Harris the mean of unplanted nonneutral sets under this conditioning is at most $O(\mu|Y|a^4/n^2)$. After planting, the mean of further nonneutral sets, averaged over the uniform proposed family of a fixed size, is $o(1)$. To bound it, multiply the individual seed multipliers for each candidate and expand in nonnegative increments. For a fixed selection of j distinct labels the sampling factor $(|Y|)_j/(\binom{n}{a})_j$ is at most $(|Y|/\binom{n}{a})^j$. Thus the total extension bound is $2\mu(e^{2\lambda}-1) = o(1)$, where $\lambda = O(\mu a^4/n^2) + n^{-1+o(1)}$ is the overlap sum above. Harris preserves this upper bound under neutral avoidance. Since duplicates and a count above 2μ have probability $o(1)$, we may therefore work with independent Poisson seeds planted in an independent no-a-set background, with joint variation error $o(1)$. There are no larger classes with high probability, since $\mu_{a+1}(n) = n^{-1+\phi+o(1)}$.

For s chosen seeds set $x = n - as$, $j = K - s$. Let $J_b(x)$ be the smooth number of top classes at the cap-b first-moment root and $S_*(x) = J_b(x)^2/\mu_b(x)$. Use the log matching sum with log weights

$$H_K(s) = L_0(x, j, b) - S_*(x)/2.$$

The K-slope is $(1 + o(1))aL$, and $\exp(-\partial_s H_K(s)) = \Theta(a\mu/n)$ uniformly on a common bracket about the single-size root at $s = \lfloor\mu\rfloor$. Indeed (1.4) applies with the size-a weight continued one step above the cap, and differentiating the cap root gives $J_b'(x) = O(1/a)$ and $S_*'(x) = O(aS_*/n)$; the latter contributes $o(1)$ per insertion. Also $x = n - O(a\mu)$, so $\mu_a(x)/\mu = 1 + o(1)$. These bounds give the single-size root slope $-(1-\phi+o(1))/a$. For $|Y| = \lfloor\mu\rfloor$ independent indexed seeds, the expectation of the log matching sum has a unique zero K_n in the bracket. Indeed the empty term gives the lower edge of the log-sum bracket, and the upper bound uses the largest single-size term times $2^{|Y|}$.

Two independent seeds meet with probability $\pi \le a^2/n$. Specifying a tree of h seed indices has edge-event probability π^{h-1}, by removing tree leaves. The component of a fixed seed has size at least h with probability at most $(e^2|Y|\pi)^{h-1}$. Replacing one seed leaves the log sum unchanged if both versions of the seed are isolated. Otherwise its change is $O((h + h')L)$ for the sizes of the two affected components, by holding other components fixed and using the weight difference per insertion. Efron–Stein bounds the variance by $O(|Y|^2\pi L^2) = o(\mu L^2)$. An isolated insertion increases the log sum by $(1 - \phi + o(1))L$, uniformly by the derivative estimate. A nonisolated insertion has expected contribution $O(|Y|\pi L) = o(L)$. Telescoping this drift over the Poisson fluctuation of the seed count and dividing by the K-slope prove that the random root differs from the right side of (4.1) by $o_{\mathbb{P}}(\sqrt{\mu}/a)$.

4.2. **Comparison with the root.** All components have bounded size with high probability by the tree estimate and $v < 1$. The seed–vertex incidence graph has no cycles: an incidence cycle of length $2h$ requires two distinct vertices in each of h seeds and has expected count bounded by $(2\mu a^2/n)^h$. Sum over h up to the component bound. This permits exposing seed sets one at a time, each meeting at

most one already exposed vertex. The number of intersecting pairs of seeds is $O_{\mathbb{P}}(\mu^2 a^2/n)$, by the pair expectation (pair indicators sharing one index are also uncorrelated by symmetry).

For the lower comparison use the background event $X_b \le m + \sqrt{m}L$, which has probability tending to one by the overlap variance bound and Harris under no-a avoidance. In a cap profile let k be its number of b-classes and $\beta = k/m$, and let R' be the required absent pairs still unforced by unselected seeds. If J top classes use a supplied pair, at least $k - J$ top classes are fully forced when R' is forced. Use the inequality obtained by multiplying by $e^{-\beta X_b}$ on the count event, and retain neutral a-sets and neutral tagged b-sets. Their omitted mean and pair correction are $O(a^2)$; the fully forced classes give a factor $e^{-\beta(k-J)}$. Dividing by the background probability $e^{-\mu+o(1)}$ and expanding $1 - e^{-\beta}$ gives the upper bound on the properness mean

$$q^{|R'|} \exp\Big(-k^2/(2m) + \beta J + O(a^2 + \beta\sqrt{m}L + 1)\Big).$$

The error is $o(\sqrt{\mu}L)$, since $\mu > n^{1/2+\eta}$ for some $\eta > 0$.

Under uniform slot placement, expose the unselected seeds in the forest order. At moderate class occupancies the excess multiplier for $2^{|R\setminus R'|}e^{\beta J}$ is $O(a^3/n)$ per seed, by the slot count for two or more vertices of a seed in a class. A near-full class gives $O(a/\mu)$ per seed, allowing the one known slot: losing r vertices from occupancy b gains $2^{-r(b-r-1)}$, which pays the additional $(an)^r$ choices for $b - r > 3a/4$. Thus the iterated cost is $\exp(O(a))$. Changing m to $\mu_b(x)$ costs $O(a)$ as well. The entropy loss from (1.2) absorbs the difference between k and the smooth top count at cost $O(j^3/\mu_b(x)^2)$, by completing the square in $|k - j\pi_b|$ using Pinsker. This yields $e^{H_K(s)+o(\sqrt{\mu}L)}$ as a uniform upper bound for the coloring mean after summing profiles. Summing over every seed matching and using the negative log-margin below the root proves the lower bound on ordinary χ.

For the upper comparison, the deficit of a matching from the largest option in a fixed component costs a factor $O(a\mu/n)$ per missing seed. Since the component sizes are bounded, at least half the positive-threshold weight lies at deficit $O(\mu^2 a/n)$ from the maximum matching size. Choose a size there with N_s matchings and E the fixed-edge expectation for the rounded cap profile on the residue of order x. Then

$$\log(N_s E) - S/2 \gg \sqrt{\mu}L, \qquad S = k^2/\mu_b(|C_0|), \tag{4.2}$$

where C_0 is the common core outside all seeds. The margin constant depends on the fixed multiple of $\sqrt{\mu}/a$ by which we exceed the root. Indeed $x - |C_0| = O_{\mathbb{P}}(\mu^2 a^3/n)$, so replacement of $\mu_b(x)$ by $\mu_b(|C_0|)$ costs only $o(1)$ in S.

Use one weight for all residues, defined on the no-a background. With $m_C = \mu_b(|C_0|)$, $\beta = k/m_C$, $Q_b = \binom{b}{2}$, $N = \binom{n}{2}$, and M the background edge count, take

$$W = \exp\left(\frac{f - \beta m_C Q_b}{qN}(M - pN) - \beta(X_b(C_0) - m_C) - \beta m_C \frac{a\mu}{m}\right).$$

The last term compensates for containments of tagged b-sets in a-sets. Indeed (3.2) gives covariance $a\mu L_z/m$ for these containments. Each named top class also gives a special possible a-extension to an outside vertex, with summed mean $\nu = k(n - b)q^b$. Their supports are stars on the new vertex. The contribution of pairs of such special events is $o(1)$, by $O(n^2 q^{3b/2})$ for small intersections and $O(nL^3 q^{b+1})$ for the $O(L^3)$ exceptional nearly-identical top pairs allowed by mutual relevance. Generic–special pairs gain the probability of meeting a specified star pair. The remaining means and covariances are as in (3.3), with these two a-set contributions. Here $m_C\beta^3 = o(1)$, the difference in the edge sums between C_0 and the whole domain costs $O(a^2(x - |C_0|)/n + a^3\mu/n) = o(1)$, and the omitted fully forced tags cost $O(\beta(x - |C_0|)) = o(1)$. The single expectation is $Ee^{-S/2-\nu+o(1)}$; the single-power estimate gives a truncation at $\log W \le H = \sqrt{S + a^2}L$ with relative loss $o(1)$.

For a pair with g shared internal pairs and j_b identical top classes, the relative bound beyond q^{-g} is

$$\exp\left(-S + (2\beta + (n-b)q^b)j_b - \frac{pf^2}{qN} + \frac{2pfg}{qN} + o(1)\right). \tag{4.3}$$

In fact the fixed-edge exponent for $2f - g$ forbidden pairs relative to two single terms is $-p[(2f - g)^2 - 2f^2]/(2qN)$. The product tilt contributes $pg^2/(2qN)$ and cancels the quadratic term in g. The remaining tag and containment terms give the first two terms of (4.3).

The sparse cell count now uses a cutoff n^ζ with $1 - u < \zeta < 1/2$. As before size two contributes $pf^2/(qN) + o(1)$. The top activity is $S + o(1)$, from its coefficient $e^{2\beta+(n-b)q^b} - q^{b-1}$. The intervening sizes sum to $o(1)$. Thus the sparse pair sum is at most $1 + o(1)$ after (4.3); the large occupancy tail beats its crude $\exp(S + O(a^2))$ factor. Intermediate identical masses use (3.5) and [3, Section 7], with $\mu_b(x) \geq x^{1.05}$. Up to a remaining size $x_t \geq 2n^{1-c_0}$, $0 < c_0 < (1 - v)/3$, the leftover profile bound supplies $\exp(-\Omega(x_t))$ near unit identical mass and beats the $\exp(O(\mu L) + S)$ errors. This follows by the same comparison of the leftover Gaussian gap and the decrease in the reference mean as above.

On the very similar pairs, require at most D_0 seed-union vertices in any class, where $(D_0+1)(1-v) > 1$, losing a fraction $o(1)$ by the slot union bound. A missing vertex outside a scrambled row costs $n^{-1+o(1)}$. Scrambled rows cost $\exp(-\Omega(L^2))$ each by the same regrouping count and have at most D_0 missing vertices. This gives a penalty $n^{-D/2}$ for D missing vertices. If two equal-size matchings differ in h components, at least $h/2$ differing seeds on each side contribute at least $a - O(1)$ vertices each outside the other matching, by the forest and the bounded component size. Summing over at most $\binom{|Y|}{h}2^{O(h)}$ competitors keeps total penalty $1 + o(1)$. With one weight truncated, the similar contribution is bounded by $n^{O(1)}e^{S/2+H+O(a^2)}/(N_s E) = o(1)$ by (4.2). The relative second moment is $1 + o(1)$ and gives a cap coloring on one residue, to which we append its matching. This proves the upper comparison. Equation (4.1) now follows since the root comparison holds for every fixed positive multiple of $\sqrt{\mu}/a$ and X_a has the Poisson limit. Its Gaussian fluctuation scale is $n^{\phi/2+o(1)}$, giving $\mathcal{L}_n(n^{1/4}) = o(1)$ in the fixed upper range.

5. Two independent-set counts

Fix $1/3 < u \leq \phi \leq v < 2/3$. In this range there are two contributing Gaussian fluctuations. To describe them let $K_*(x)$ be the cap-b root of $L_0(x, K_*(x), b) = S_*(x)/2$, with S_* defined as above. Put $A_n = \partial_j L_0(n, K_*(n), b)$, $w_b = \sqrt{S_*(n)}/A_n$, and $w_a = (1 - \phi)\sqrt{\mu}/a$. With $F_n(t) = t + K_*(n - at)$, the desired comparison is

$$\chi(G_n) = F_n(\mu) - w_a\frac{X_a - \mu}{\sqrt{\mu}} - w_b\frac{X_b - m}{\sqrt{m}} + o_{\mathbb{P}}\big((w_a^2 + w_b^2)^{1/2}\big). \tag{5.1}$$

5.1. **A vertex-dependent weight.** We first work in the no-a background. For a rounded cap-b profile put $k = k_b$, $\beta = k/m$, $S = k^2/m$, $R_3 = m\beta^3$, and $B = b^2S^2/n$. In this range $na^C/\mu^3 = o(1)$ for every fixed C, $S = o(\sqrt{m})$, and $B + R_3 + n\beta^2a^C = o(\sqrt{S})$. A constant tag exponent would leave a nonnegligible term of order R_3. To cancel it, let U_v be independent nonnegative two-point variables with mean $\beta/(4b)$ and variance β^2. Their nonzero value is $(4b + 1/(4b))\beta$. Set $\Gamma_A = 3\beta/4 + \sum_{v\in A} U_v$ for size-b labels. With $z_A = 1 - e^{-\Gamma_A}$, $L_0' = q^{Q_b}\sum_A z_A$, $V_e = q^{Q_b}\sum_{A\supset e} z_A$, $D_e = V_e - Q_bL_0'/N$, use the graph weight

$$W = \mathbb{E}_U \exp\left[\begin{array}{c}\dfrac{f - Q_bL_0'}{qN}(M - pN) - \displaystyle\sum_A \Gamma_A \mathbf{1}_{\{A \text{ independent}\}} \\ + L_0' - \dfrac{a\mu}{m}L_0' - \dfrac{p}{2q}\displaystyle\sum_e D_e^2\end{array}\right].$$

The last two terms compensate the containment and the centered edge covariance. Since $z_A \le \Gamma_A \le (4b^2+1)\beta$ and $\beta = O(1/\mu)$, every tag is $O(b^2/\mu)$. The connected triple and moderate-overlap estimates for (3.3) allow such fixed powers of b and retain strict polynomial margins. The same special a-extension term as in the fixed upper computation remains. A single proper-profile mean is $(1+o(1))Em_1e^{-\nu}$, with $m_1 = e^{-3\beta k/4}(\mathbb{E}e^{-U_v})^{bk}$ and $\log m_1 = -S + O(n\beta^2)$.

For two copies of the field under a proper pair, let j be the number of identical top classes and D the size of the intersection of the two top vertex unions. The field tilt at a vertex is by $e^{-c_vU_v}$ with c_v its number of distinct top classes in the two partitions. Thus $\sum c_v = b(2k-j)$ and $\sum c_v^2 = b(2k-j)+2(D-bj)$. Expanding the forced factors and dividing by the square of the single factor gives log term

$$-2S + 2\beta j + \beta^2[b(k-j) + 2(D-bj)] + o(1).$$

For $j \le 10S$, the cross-tag mean $\sum_A q^{Q_b} z_A(U) z_A(U')$ in this field law is $S - (b+1)R_3 - 4B + o(1)$. In this calculation $\mathbb{E}\Gamma_A = \beta - \beta^2 \sum_{v\in A} c_v + O(a^{O(1)}\beta^3)$, and $\mathbb{E}\Gamma_A^2 = (b+1)\beta^2 + O(a^{O(1)}\beta^3)$. For the exponential mean, its two centered linear sums and its centered diagonal bilinear sum each contribute variance $B+o(1)$. Their cross covariances vanish before the local exponential tilt. The local summands have absolute bound $O(a^2/\mu)$, giving log correction $3B/2 + o(1)$. The off-diagonal term factors into a product of the two centered field sums with coefficient $O(\mu a/n)$. Their squared range parameters are $O(na^2/\mu^2)$, giving exponential moment $1 + o(1)$ even with mean displacements $O(na^C/\mu^2)$. The centered cubic remainder has sum of squared differences $O(na^C/\mu^4)$ and mean displacement $O(na^C/\mu^3)$, and the centered spatial correction has squared differences $O(a^8/n)$ in total and mean displacement $O(a^C/\mu)$. All contribute $o(1)$ to the log. Combining these expansions gives the relative field factor

$$\exp\Big(-S + 2\beta j - R_3 + 2\beta^2 D - \tfrac{5}{2}B + o(1)\Big). \tag{5.2}$$

The fixed-edge compensation adds the same $-pf^2/(qN) + 2pfg/(qN)$ as in (4.3).

Extract identical classes by (3.5). On the sparse region $h \le n/L^3$, all remaining cells have size at most $b-1$. Their total activity is $O(a^2)$ in size two and $o(1)$ at larger sizes. Indeed the upper endpoint uses $\mu_{b-1}(n) = n^{2+\phi+o(1)}$. The cutoff n^ζ, $1-2u < \zeta < 1/2$, has tail $\exp(-n^{\zeta-o(1)})$ and absorbs the residual $\exp(O(B+a^2))$ terms. In the small region use nonnegative witnesses from the graph expansion of $(1+p/q)^{\binom{i}{2}}$ in a cell of size i. After assigning the witness vertices, the free permutation is uniform. The remaining top–top size is hypergeometric. Its mean together with already assigned vertices is $(bk)^2/n + O(h+z_w)$ for total witness size z_w. At $2\beta^2$ its centered exponential moment is $1+o(1)$ by $n\beta^4 = o(1)$. The $O(\beta^2 z_w)$ charge changes cell activities by $o(1)$. Thus integrating the D term adds $2B + O(\beta^2 h) + o(1)$ in the log. The size-two activity cancels $-pf^2/(qN)$ and $fg/N = o(1)$ in this region.

For $j \le 10S$ the top part of (3.5) is

$$\frac{S^j}{j!}\exp\left[\left(\frac{b^2}{2n} - \frac{1}{k}\right)j^2 + o(1)\right].$$

For lower sizes the sum of activities is $o(1)$, also after charging $O(\beta^2 h)$. Under a Poisson variable of mean S, replacing $2\beta j + (b^2/(2n) - 1/k)j^2$ by its value at S costs relative $o(1)$ in the exponential average. The Lipschitz constant on $j \le 10S$ is $O(a/\mu)$, and $S(a/\mu)^2 = o(1)$. The value is $R_3 + B/2$, exactly canceling the remaining terms in (5.2). The Poisson tail $j > 10S$ absorbs the crude $\exp(S + o(S) + O(a^2))$ factor. Intermediate identical masses again use (3.5), the profile inequalities of [3, Sections 3, 6–7], and the leftover gap bound. The loss $\exp(-\Omega(n-h))$ near unit identical mass beats $\exp(S+o(S))$ as long as $n - h \ge n^{1-c_0}$ with $c_0 < 1/3$.

5.2. **The count cutoff.** Let $Y_b = (X_b - m)/\sqrt{m}$ in the no-a background. Weighting one term or the sparse pair estimate by e^{-tY_b}, $t \ge 0$, replaces each tag exponent by that exponent plus $t/\sqrt{m}$. Omitted

b-candidates have mean $O(ma^3/n + a^C) = o(\sqrt{m})$. The low occupancies follow by the two-forced-pair slot count. At intermediate occupancy, prescribe the bounded number of classes or exceptional class unions containing a fixed fraction of a candidate. If they cover $s > b/3$ of its vertices with each intersection at most $3b/4 + 1$, there remain at least $b(b - s)/2 + \Omega(b^2)$ free pairs. The quadratic loss pays the selection cost. Near a full class the outside vertices instead pay their cross pairs. In exceptional paired-class unions an outside vertex has at most two forced neighbors in that union. Wholly internal candidates have only polylogarithmic stock. Harris bounds the omitted mean in the weighted law by the same order. The normalized log Laplace transform changes by

$$t^2/2 + t(L_0' - k_R)/\sqrt{m} + o(1).$$

The second term has mean and centered log moment $o(1)$, by the field ranges, the forced-class count k_R, and $j \le 10S = o(\sqrt{m})$. The transforms identify the standard normal by first tilting at a fixed negative Laplace argument and then undoing the tilt on compact intervals. In particular, retaining $Y_b \ge z$ for fixed z retains the factor $1 - \Phi(z) + o(1)$ in the normalized single and sparse pair estimates.

The similar pairs need a count-dependent truncation. In the one-partition expectation of $W^2 e^{\beta(X_b - m)}$, the actual tag exponent for two independent copies is $\Gamma_U + \Gamma_{U'} - \beta \ge \beta/2$. For $\delta_U = \Gamma_U - \beta$, use the identity

$$(1 - e^{-\Gamma_U}) + (1 - e^{-\Gamma_{U'}}) - (1 - e^{-(\Gamma_U + \Gamma_{U'} - \beta)}) = 1 - e^{-\beta} + e^{-\beta}(1 - e^{-\delta_U})(1 - e^{-\delta_{U'}}).$$

After subtracting βm, the scalar term is $-S/2 + O(R_3)$, and all field terms cost $O(B + n\beta^2 + a^C) = o(\sqrt{S})$. Thus Markov under the weighted single law permits the cutoff

$$W \le \exp(-S/2 - z\sqrt{S} + \kappa\sqrt{S}) \quad \text{on } Y_b \ge z$$

for any fixed $\kappa > 0$, with relative loss $o(1)$. On similar pairs use this bound on one of the two weights and the one-copy properness estimate on the other. Its factor relative to squared normalization is at most $\exp(S + o(\sqrt{S}))$. The similar-pair count of [3, Lemma 6.5] thus gives a negligible contribution when $\log E \ge S/2 - z\sqrt{S} + 2\kappa\sqrt{S}$. The weighted second moment, with one common count event, certifies a coloring on all but $o(1)$ of the event $Y_b \ge z$. Both the first moment fraction and the normalized pair upper bound are $1 - \Phi(z) + o(1)$. A finite grid for z, tightness, and shrinking grid width give the upper comparison at $K_* - w_b Y_b + o_{\mathbb{P}}(w_b)$.

For the lower comparison on $Y_b \le z$, tilt a proposed profile with q_b top classes by $\gamma = q_b/m$, retaining neutral tags and the fully forced classes as in the fixed upper comparison. The log penalty is $-q_b^2/(2m) + zq_b/\sqrt{m}$, with error $O(a^2 + m\gamma^3)$ and the contribution of classes touched by omitted seeds. Averaging over them costs $\exp(O(a))$ by the nonnegative subset expansion under sampling without replacement. Entropy absorbs deviations of q_b from its smooth optimum at cost $O(j_{\mathrm{col}}^3/m^2 + j_{\mathrm{col}}/m) = o(\sqrt{S})$, with j_{col} the tested number of residue colors. The first moment excludes cap profiles below the corresponding count-dependent threshold.

5.3. **Transfer to the ordinary chromatic number.** Use the joint largest-set planting comparison above. A maximum seed matching loses $O_{\mathbb{P}}(\mu^2 a^2/n) = o_{\mathbb{P}}(\sqrt{\mu})$ labels when $v < 2/3$. It can be selected independently of the background. The residue at order $x = n - as$ has its own centering $\mu_b(x)$. The removed size-b count has variance $O(m - \mu_b(x)) = o(m)$ by the overlapping-pair bound. Hence the residue and full centered counts differ by $o_{\mathbb{P}}(\sqrt{m})$. The background residue law is within $o(1)$ of the no-a conditioning at order x. Indeed the extra likelihood for avoiding a-sets crossing the deleted vertices is $\exp(-(\mu_a(n) - \mu_a(x)) + o_{\mathbb{P}}(1))$ under the residue law, from Harris and the pair bound in (3.2). The normalizer has that same value to relative $1 + o(1)$, so the mean-one likelihood converges in L^1. The cap upper bound then applies to this chosen residue, and its seed classes are appended.

For the lower bound include every seed matching. Omitting one seed costs at least $\log(n/(Ca\mu))$ in log first moment by (1.4) and the corrected cap location. Summing all omission choices therefore

costs at most $(1 + Ca\mu/n)^T$ for T seeds. Its logarithm is $o(A_n(w_a^2 + w_b^2)^{1/2})$, since $v < 2/3$. The first moment for all these choices gives the ordinary lower bound with the stated error. No larger classes occur with high probability.

For X_b the scalar Poisson approximation in the product law follows from the dependency graph of shared vertex pairs. The joint-probability sum divided by m is $O(ma^4/n^2) + n^{-1+o(1)} = o(1)$. The no-$a$ background has the same centered Gaussian limit, from the tag transform or the joint planting comparison. The seed count is independent of the background. Adding the seeds changes X_b by $o_{\mathbb{P}}(\sqrt{m})$: the inside-seed contributions are $O(a\mu)$, and the noncontainment multiplier has expected cost $O(m\mu a^{O(1)}/n^2) + n^{o(1)}$, by the intersection sum. Thus $(X_a - \mu)/\sqrt{\mu}$ and $(X_b - m)/\sqrt{m}$ have independent standard normal limits. Differentiating the cap equation gives $F_n'(t) = -(1 - \phi + o(1))/a$ for t within $O(\sqrt{\mu}L)$ of μ, by the derivatives following (1.2). Combining the two comparisons yields (5.1).

Let $W_n = (w_a^2 + w_b^2)^{1/2}$. The weighted standardized sum converges to a standard normal, since along any subsequence the two normalized deterministic coefficients have a convergent further subsequence with squared sum one. As $k \asymp n/a$, $m \asymp n\mu/a$, and $A_n \asymp aL$, one has $w_a w_b \asymp \sqrt{n}/(a^{5/2}L)$. Consequently $W_n \geq n^{1/4-o(1)}$, giving $\mathcal{L}_n(n^d) = o(1)$ on the fixed two-count range whenever $d < 1/4$.

6. The all-orders limit

For a normalized variable converging to a continuous distribution, uniform convergence of its distribution functions implies vanishing maximum mass on intervals of length tending to zero. Apply this observation to the two Gaussian comparisons. Fix $0 < d < 1/4$. If $\mathcal{L}_n(n^d)$ stayed bounded below on a subsequence, pass to one with $\phi \to \phi_* \in [0, 1]$. At 0 or 1, choose $b = a - 1$ or $b = a$ respectively, and use the endpoint comparison. For $0 < \phi_* < 1/3$ use a compact lower interval, and for $\phi_* = 1/3$ use the mixed comparison. If $1/3 < \phi_* \leq 1/2$, use a compact interval with lower endpoint above $1/3$ and upper endpoint below $2/3$. If $1/2 < \phi_* < 1$, use the fixed upper comparison. Every case contradicts the lower bound on the subsequence. Hence $\mathcal{L}_n(n^d) \to 0$ along all orders.

For any sequence of deterministic sets of at most C integers, the mass of each is bounded by $C\mathcal{L}_n(0) \to 0$. This answers the first question of Erdős Problem 1156 negatively, including for separated values. For the second, choose $\omega(n) = n^d$ for any fixed $0 < d < 1/4$. The half-mass bound then holds strictly for every center at every sufficiently large order. Any divergent positive ω eventually at most this choice also works.

References

[1] T. F. Bloom, *Erdős Problem #1156*, https://www.erdosproblems.com/1156.
[2] A. Heckel, *Non-concentration of the chromatic number of a random graph*, arXiv:1906.11808.
[3] A. Heckel and K. Panagiotou, *Colouring random graphs: Tame colourings*, arXiv:2306.07253v3, 2024.
[4] A. Heckel and O. Riordan, *How does the chromatic number of a random graph vary?*, J. Lond. Math. Soc. (2) (2023). doi:10.1112/jlms.12794.

Email address: howew@purdue.edu

Email address: huffma97@purdue.edu

Email address: jenwin@purdue.edu

Email address: kruerl@purdue.edu