hexagon:2611.00562v1

A FULL-SEQUENCE QUANTITATIVE GAP BETWEEN THE CHROMATIC AND COCHROMATIC NUMBERS OF A RANDOM GRAPH

SAMUIL PETKOV

Abstract. The cochromatic number $\zeta(G)$ of a graph G is the least number of parts in a partition of $V(G)$ into cliques and independent sets, so $\zeta(G) \le \chi(G)$. Erdős and Gimbel asked whether $\chi(G_n) - \zeta(G_n) \to \infty$ with high probability for $G_n \sim G(n, 1/2)$. Heckel answered this affirmatively for about 95% of all integers n and conjectured that the natural order of the difference along the full sequence is $n/(\ln n)^3$. We prove that, as $n \to \infty$ through all integers, $\chi(G_n) - \zeta(G_n) \ge 0.063\, n/(\ln n)^3$ with high probability. The failure probabilities are summable, so under any coupling of the graphs G_n the bound holds almost surely for all sufficiently large n.

The proof counts partitions of the vertex set into classes of five consecutive sizes, each class declared independent or complete. The two choices per class outweigh the entropy lost by restricting the class sizes, so the expected number of such partitions is large already at a number of classes that lies order $n/(\ln n)^3$ below the first-moment lower bound for the chromatic number, uniformly in the rounding phase. A second-moment argument, which controls all overlaps between two such partitions, including whole shared classes and large intersections of unequal classes, bounds the probability that such a partition exists from below. Vertex-exposure concentration, together with a coloring of the leftover vertices, then raises this to high probability at a cost of $o(n/(\ln n)^3)$ extra classes.

Introduction

A *cocoloring* of a graph G is a partition of $V(G)$ into nonempty classes, each inducing either an edgeless graph or a complete graph. The least number of classes in such a partition is the *cochromatic number* $\zeta(G)$. Since an ordinary proper coloring is a cocoloring using only independent-set classes, one always has $\zeta(G) \le \chi(G)$.

Write $[n] := \{1, \dots, n\}$, and let $G_n \sim G(n, 1/2)$ be the labeled random graph on $[n]$. All logarithms are natural unless a base is displayed. The leading asymptotics of $\chi(G_n)$ and $\zeta(G_n)$ agree, so they do not determine whether the difference between the two numbers grows. Erdős and Gimbel asked whether

$$\chi(G_n) - \zeta(G_n) \longrightarrow \infty$$

with high probability (Erdős and Gimbel, 1993, p. 263). The question was restated by Gimbel (Gimbel, 2016, Section 7.4) and is cataloged as Erdős Problem 625 (Bloom, 2026). We prove a quantitative bound along the full sequence of integers.

The expected number of independent sets of size k is

$$\mu_k = \binom{n}{k} 2^{-\binom{k}{2}}.$$

Solving $\mu_k = 1$ asymptotically gives a real-valued center $\alpha_0(n)$ (Bollobás and Erdős, 1976, p. 420). The phase δ_n is its fractional part, which records where this center lies between consecutive integer class sizes. Both quantities are defined precisely in the main theorem below.

Main theorem (Uniform lower bound and phase refinement)**.** *For every integer $n \ge 2$, let $G_n \sim G(n, 1/2)$ be the random simple graph on $[n]$, with edges present independently with probability $1/2$. All logarithms without a displayed base are natural, and all limits are as $n \to \infty$ through the integers. Set*

$$\alpha_0(n) := 2\log_2 n - 2\log_2 \log_2 n + 2\log_2(\mathrm{e}/2) + 1, \qquad \delta_n := \{\alpha_0(n)\} = \alpha_0(n) - \lfloor \alpha_0(n) \rfloor.$$

Date: 30 August 2026.

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

Key words and phrases. random graph, chromatic number, cochromatic number, second moment method, configuration model.

Thus $\delta_n \in [0,1)$ is deterministic. For $0 \le \delta \le 1$, define

$$S_+ := \{i \in \mathbb{Z} : i \ge -1\}, \qquad S_5 := \{2,3,4,5,6\}, \qquad T_\delta := 1 + \frac{2}{\log 2} - \delta.$$

Then $T_\delta \in (2,5)$. For $S \in \{S_+, S_5\}$ and $T \in (2,5)$, put

$$\mathcal{F}_S(T) := \inf_{\lambda \in \mathbb{R}} \left\{ \log \sum_{i \in S} \exp\left(\lambda i - \frac{\log 2}{2} i^2\right) - \lambda T \right\}, \qquad A_5(\delta) := \log 2 - \mathcal{F}_{S_+}(T_\delta) + \mathcal{F}_{S_5}(T_\delta).$$

The sum inside the logarithm defining $\mathcal{F}_{S_+}(T)$ converges for every $\lambda \in \mathbb{R}$; for $T \in (2,5)$, each displayed infimum is a finite real and is uniquely attained.

Then A_5 is continuous, positive and nonconstant on $[0,1]$, so

$$c_5 := \frac{(\log 2)^2}{4} \min_{0 \le \delta \le 1} A_5(\delta) > 0.$$

For every constant $c < c_5$,

$$\mathbb{P}\left(\chi(G_n) - \zeta(G_n) \ge c\, \frac{n}{(\log n)^3}\right) \longrightarrow 1.$$

Moreover, there is a deterministic sequence of nonnegative reals $(\varepsilon_n)_{n \ge 2}$ with $\varepsilon_n \to 0$ such that

$$\mathbb{P}\left(\chi(G_n) - \zeta(G_n) \ge \left[\frac{(\log 2)^2}{4} A_5(\delta_n) - \varepsilon_n\right] \frac{n}{(\log n)^3}\right) \longrightarrow 1.$$

The first bound holds for every fixed coefficient below c_5. The refinement keeps the coefficient $A_5(\delta_n)$, which depends on where the independent-set center lies between neighboring integers. Since $A_5(\delta_n)$ is at least the minimum of A_5 on $[0,1]$ and $\varepsilon_n \to 0$, the refined bound implies the uniform bound for every $c < c_5$. The constant c_5 is bounded from below in Section 11 by Lemma 5.4 (Section 5), which shows that $A_5 > \log(169/100)$ on $[0,1]$, so

$$c_5 > \frac{(\log 2)^2}{4} \log\left(\frac{169}{100}\right) = 0.0630268508\ldots.$$

Since $n/(\log n)^3 \to \infty$, both forms answer the question of Erdős and Gimbel along the full sequence of integers. Remark 11.1 shows that one may take $\varepsilon_n = O(\log\log n/\log n)$. Corollary 11.2 bounds the failure probability by a summable sequence, and Corollary 11.3 deduces the almost-sure statement for each fixed coupling.

Origin of the coefficient. Write a class size as $\alpha - i$ with $\alpha = \lfloor \alpha_0(n) \rfloor$; we call i the deficit. Up to a term affine in i, which the totals of classes and vertices fix, $-\log\big((\alpha - i)!\, 2^{\binom{\alpha - i}{2}}\big) = -\frac{\log 2}{2} i^2 + O(1/\log n)$ for fixed i (Section 3). Thus $\mathcal{F}_S(T)$ is the optimal entropy per class, less this quadratic cost, over deficit distributions on S with mean T, and T_δ is the mean deficit at the first-moment root. Declaring each class independent or complete adds $\log 2$ per class, so $A_5(\delta)$ is the net gain per class of marked five-size partitions over colorings. Dividing the total gain, about $A_5(\delta_n)\frac{\log 2}{2}\frac{n}{\log n}$, by the slope $\frac{2}{\log 2}(\log n)^2$ of the counting exponent in the number of classes gives the coefficient in the theorem (Section 5).

When $\alpha_0(n)$ crosses an integer, the size cutoff changes, and with it the optimization problem used to choose the profile and the first-moment root of that problem, the point where the optimized counting exponent vanishes. An argument that is valid for most integers can therefore miss some phases. We keep the phase explicit and make every estimate uniform up to both endpoints of the phase interval, including for integers on either side of a change of cutoff.

Relation to previous work. The first-order asymptotic for the chromatic number of a dense random graph was proved by Bollobás (Bollobás, 1988), following the early work of Grimmett and McDiarmid (Grimmett and McDiarmid, 1975); later refinements include McDiarmid (1990), Panagiotou and Steger (2009), and Heckel (2018). The cochromatic number belongs to the theory of generalized chromatic numbers associated with hereditary graph properties (Scheinerman, 1992; Bollobás and Thomason, 1995).

For the difference $\chi(G_n) - \zeta(G_n)$, Heckel and, independently, Steiner showed how a high-probability upper bound on the gap would imply concentration of the chromatic number (Heckel, 2024; Steiner, 2025). Heckel subsequently proved a near-linear lower bound for a phase-dependent set containing approximately 95% of the integers and conjectured that the natural full-sequence scale is $n/(\log n)^3$ (Heckel, 2025). We prove a lower bound of this order along the full sequence $n \to \infty$, with an explicit phase-resolved coefficient. We use two ingredients from Heckel's argument: the additional choices obtained by declaring each class independent or complete, and the strategy for raising a small success probability to high probability. Beyond these, we allow five consecutive class sizes, which makes the profile optimization uniform in the phase,

and our second-moment estimate counts compatible class declarations exactly and controls all overlaps between two partitions, including whole shared classes.

Appendix H records the companion Lean theorem and its reproducibility materials.

Proof strategy. We construct deterministic integers k_χ^- and k_{co}, and a deterministic $a_n = o(n/(\log n)^3)$, for which

(A)
$$k_\chi^- - k_{\mathrm{co}} - a_n = \left\lfloor \frac{(\log 2)^2}{4} A_5(\delta_n) + o(1) \right\rfloor \frac{n}{(\log n)^3};$$

(B)
$$\mathbb{P}\big(\chi(G_n) > k_\chi^-\big) \longrightarrow 1;$$

(C)
$$\mathbb{P}\big(\zeta(G_n) \le k_{\mathrm{co}} + a_n\big) \longrightarrow 1.$$

The intersection of the events in (B) and (C), together with (A), gives the phase-resolved assertion of the main theorem; no independence between the two events is required.

Here r_+ and r_5^{co} are first-moment roots: the numbers of classes at which the optimized logarithmic counting exponent vanishes for ordinary colorings and for marked five-size partitions, respectively (Section 3). Figure 1 shows the order of all five locations. Sections 1–3 collect standard estimates, the phase expansion and the roots, and Section 4 proves (B). Section 5 shows that $r_+ - r_5^{\mathrm{co}}$ has the order in (A) and fixes the profile and k_{co}. Sections 6–9 bound the second moment of the marked cocoloring count, and Section 10 uses that bound to construct a_n and prove (C). The final section completes (A) and combines the bounds. Long routine proofs are given in the appendices.

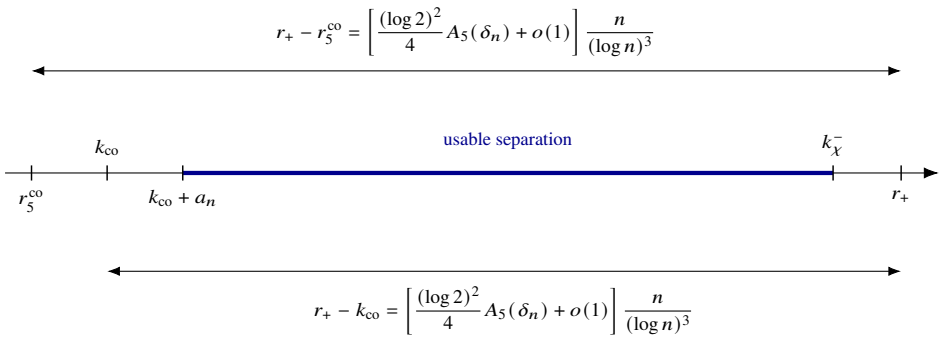


FIGURE 1. First-moment roots and the class counts used after rounding and amplification. For all sufficiently large n, $r_5^{\mathrm{co}} < k_{\mathrm{co}} < k_{\mathrm{co}} + a_n < k_\chi^- < r_+$. The thick segment is the remaining gap between the class counts. Positions are schematic and not to scale.

The two counting difficulties. Two kinds of overlap enter the second moment. The first is whole shared classes. Section 7 handles them by exact recurrences when few or almost all classes are shared, and by a negative exponential rate in the intermediate range. These estimates are uniform over the phase, including the range omitted in the earlier argument. The second is large intersections between classes that are not equal.

To handle these intersections, we form a table with one row for each class of the first partition and one column for each class of the second. Each entry is the size of the corresponding intersection. Entries larger than half the maximum class size cannot share a row or column; viewed as edges between row and column classes, they form a matching. Section 8 bounds their contribution. Section 9 treats the other entries and the compatibility of the class declarations. We retain the chosen vertex pairs throughout, so each labeled configuration is counted once. If Z counts the marked cocolorings of the chosen profile, the resulting second-moment bound is

$$\frac{\mathbb{E}Z^2}{(\mathbb{E}Z)^2} \le \exp\left\{ o\left(\frac{n}{(\log n)^4} \right) \right\}.$$

Asymptotic and counting conventions. An event holds *with high probability* if its probability tends to one as $n \to \infty$. Whenever a phase parameter is present, each occurrence of $o(1)$ denotes a deterministic error sequence that is uniform over the full phase interval, including sequences approaching either endpoint. Since δ_n is a function of n, uniformity in the phase is a property of the proofs: no constant, threshold or error sequence depends on a lower bound for δ_n or for $1 - \delta_n$. For integers $x, r \geq 0$, we use the falling factorial

$$(x)_r = x(x-1)\cdots(x-r+1), \qquad (x)_0 = 1,$$

and set $(x)_r = 0$ for $r > x$. A quotient involving falling factorials is used only after the denominator has been shown to be nonzero.

Class marks and profiles. For the counting argument, give each partition class a mark I or K, meaning that the class must be independent or complete, respectively. All classes in the selected profile have size at least two for sufficiently large n. The two requirements are therefore disjoint, and a realized cocoloring determines its marks uniquely. Forgetting the marks introduces no multiplicity.

A *profile* is a finitely supported sequence $(k_s)_{s\geq 1}$ of nonnegative integers, where k_s is the number of classes of size s (Heckel, 2025, Section 2.2). It is feasible on n vertices with k parts when

$$\sum_s k_s = k, \qquad \sum_s s k_s = n.$$

From Section 3 on, profiles are indexed by the deficit $i = \alpha - s$, where $\alpha = \lfloor \alpha_0(n) \rfloor$: there k_i is the number of classes of size $\alpha - i$, so k_2 counts classes of size $\alpha - 2$, not 2. The ordinary first moment is the expected number of profile partitions whose classes are all independent. The marked first moment is the expected number of profile partitions with one I/K declaration per class, each satisfying its independence or completeness requirement.

Overlap tables. For two ordered profile partitions, the overlap table records the sizes of all pairwise intersections of their classes, and its support graph joins two classes whose intersection has size at least two; both are defined in Section 6. Large cells and endpoint tables are introduced in Sections 8–9.

1. Independent-set counts and probability inequalities

This section records the expected independent-set counts and the standard probabilistic inequalities used throughout the paper; none of them is new.

For integers $v, s \geq 0$, with $\binom{v}{s} = 0$ when $s > v$, let

$$\mu_s(v) = \binom{v}{s} 2^{-\binom{s}{2}}, \qquad \mu_s = \mu_s(n). \tag{1.1}$$

We use the following standard inequalities, stated with the hypotheses and constants needed in the proof.

1. For integers $m \geq 1$,
$$\log(m!) = m\log m - m + \frac{1}{2}\log(2\pi m) + O(1/m). \tag{1.2}$$
Consequently, with $0 \log 0 = 0$, one has $\log(m!) = m \log m - m + O(\log(m+1))$ uniformly for $m \in \mathbb{N}_0$, with an absolute implied constant.
2. If $r \geq 1$, $t \geq 0$, and a random variable Y is a function of r independent blocks and changing one block changes Y by at most one, then
$$\mathbb{P}(|Y - \mathbb{E}[Y]| \geq t) \leq 2\exp(-2t^2/r). \tag{1.3}$$
We will use the corresponding one-sided bounds as well.
3. If $Z \geq 0$, $0 < \mathbb{E}[Z] < \infty$, and $\mathbb{E}\left[Z^2\right] < \infty$, then
$$\mathbb{P}(Z > 0) \geq \frac{\mathbb{E}[Z]^2}{\mathbb{E}\left[Z^2\right]}. \tag{1.4}$$
4. If $X \sim \mathrm{Bin}(m, 1/2)$, then
$$\mathbb{P}(X \leq m/4) \leq e^{-m/16}. \tag{1.5}$$
Indeed, exponential Markov with $t = \log 3$ bounds the probability by $[3^{1/4}(2/3)]^m \leq e^{-m/16}$.
5. For every nonnegative random variable X and $a > 0$, $\mathbb{P}(X \geq a) \leq \mathbb{E}[X]\,/a$.

The first is Stirling's estimate, the second is McDiarmid's bounded-differences inequality, and the third is the zero-threshold case of Paley–Zygmund. The fourth is the only binomial-tail estimate used below; the fifth is Markov's inequality. The bounded-differences method is due to McDiarmid (1989); the one-sided formulation used here is McDiarmid (1998, Theorem 3.1).

2. Uniform first-moment estimates across the rounding phase

This section fixes the parameters that describe the rounding phase and records, uniformly in the phase, how μ_s behaves for s near α_0. Two consequences are used later: independent sets of size $\alpha + 2$ are absent with high probability by (2.9), which caps the class sizes in Section 4; and $\mu_{\alpha-2}$ is large by (2.4), which controls the common classes in Section 7.

Put

$$q := \log 2, \qquad L := \log n, \qquad \ell := \log\log n, \qquad C := 1 + \log q - q.$$

(In Sections 6–7 and in the appendix that proves Lemma 7.1, ℓ instead counts common classes.) We shall use the elementary bounds

$$\frac{2}{3} < q < \frac{7}{10}. \tag{2.0}$$

We also use $q > 56/81$: writing $q = 2\int_0^{1/3}(1-t^2)^{-1}\,dt$ and retaining $1 + t^2$ in the integrand gives this strict bound. The location parameter α_0 defined in the introduction has the following equivalent form in natural logarithms (Heckel, 2025, Section 2.1):

$$\alpha_0 = 2\log_2 n - 2\log_2\log_2 n + 2\log_2(\mathrm{e}/2) + 1 = \frac{2(L - \ell + C)}{q} + 1. \tag{2.1}$$

Let

$$\alpha := \lfloor \alpha_0 \rfloor, \qquad \delta := \alpha_0 - \alpha, \qquad b := 1 - \delta.$$

Thus $0 \le \delta < 1$, $0 < b \le 1$, and

$$\alpha = \frac{2(L - \ell + C)}{q} + b.$$

When emphasizing dependence on n, we write $\delta_n := \delta$. We prove uniform estimates on the closed range $0 \le \delta \le 1$.

Lemma 2.1 (Uniform phase expansion and adjacent-size control)**.** *There exist a bounded continuous function $K\colon [0,1] \to \mathbb{R}$, absolute constants $C_{\mathrm{ph}}, C_2, c > 0$, and a real sequence $(E_n^{\mathrm{ph}})_{n\ge 2}$ such that, with*

$$\varepsilon_n^{\mathrm{ph}} := C_{\mathrm{ph}}\frac{1+\ell^2}{L},$$

one has $\varepsilon_n^{\mathrm{ph}} \to 0$ and, for all sufficiently large n, the following expansion holds at the induced phase $\delta_n = \alpha_0 - \lfloor \alpha_0 \rfloor$:

$$\log \mu_\alpha = \delta_n L + \left(\frac{2}{q} - \frac{1}{2} - \delta_n\right)\ell + K(\delta_n) + E_n^{\mathrm{ph}}, \qquad |E_n^{\mathrm{ph}}| \le \varepsilon_n^{\mathrm{ph}}. \tag{2.2}$$

The error bound is uniform over all induced phase values, and therefore also along sequences approaching either endpoint of $[0,1]$. Moreover, uniformly over those phase values,

$$\begin{aligned} \log \mu_{\alpha+2} &= (\delta_n - 2)L + \left(\frac{2}{q} + \frac{3}{2} - \delta_n\right)\ell + O(1), \\ \mu_{\alpha+2} &\le \exp\{-L + C_2\ell\} = o(1) \end{aligned} \tag{2.3}$$

while

$$\mu_{\alpha-2} \ge c\, n^2 (\log n)^{2/q - 5/2} \tag{2.4}$$

for all sufficiently large n.

The proof starts with Stirling's formula and bounds its remainder uniformly in the phase. Ratios between adjacent class sizes then give the upper size cap and the lower first-moment bound. Section A.1 gives both calculations.

Let $\alpha(G)$ denote the independence number of a graph G, and let $X_{\alpha+2}$ denote the number of independent sets of size $\alpha + 2$. Then $\mathbb{E}X_{\alpha+2} = \mu_{\alpha+2}$, and the event $\alpha(G_n) > \alpha + 1$ implies $X_{\alpha+2} \ge 1$. Consequently Markov's inequality and (2.3) give, uniformly along the full sequence of integers, the deterministic cap error

$$\mathbb{P}\big(\alpha(G_n) > \alpha + 1\big) \le \mu_{\alpha+2} =: \varepsilon_n^{\mathrm{cap}} \longrightarrow 0 \tag{2.9}$$

3. Real profile optimization and first-moment roots

We initially allow real class counts and maximize the logarithmic counting exponent subject to fixed totals for classes and vertices. A zero of this optimized exponent is a first-moment root. The ordinary-coloring root is used to obtain a lower bound for the chromatic number; the five-size marked-partition root determines the number of parts in our cocoloring. Section 5 compares these roots before the class counts are rounded to integers. Proposition 5.5 collects the properties of the selected integer profile. Throughout this section, n is sufficiently large that $\alpha \geq 10$. For an integer class size $u \geq 1$, put

$$d_u := 2^{\binom{u}{2}} u!.$$

Write $u = \alpha - i$. The index i records the deficit of a class size from the integer center α. We call the set of allowed deficits the support and use

$$S_+ := \{-1, 0, 1, 2, \ldots\}, \qquad S_5 := \{2, 3, 4, 5, 6\}. \tag{3.1}$$

At finite n, the unrestricted support is $S_+^{(n)} := \{-1, 0, \ldots, \alpha - 1\}$: the lower endpoint is the deterministic size cap $u \leq \alpha + 1$, and the upper endpoint comes from positivity of class sizes. For $S \in \{S_+, S_5\}$, set

$$S^{(n)} := \begin{cases} S_+^{(n)}, & S = S_+, \\ S_5, & S = S_5. \end{cases}$$

All profiles, maximizations and sums at finite n use $S^{(n)}$. Only the limiting quantities use the full support S.

We separate the class-size contribution into an affine function of the deficit and a remainder. The conservation constraints fix the affine contribution, leaving only the remainder to optimize. For $i \in S_+^{(n)}$, define that remainder by

$$h_n(i) := \begin{cases} -\dfrac{\log 2}{2} i^2 + \displaystyle\sum_{r=0}^{i-1} \log\Big(1 - \frac{r}{\alpha}\Big), & i \geq 0, \\ -\dfrac{\log 2}{2} + \log\Big(\dfrac{\alpha}{\alpha+1}\Big), & i = -1. \end{cases} \tag{3.1a}$$

The affine part is fixed by the mean constraint:

$$-\log d_{\alpha - i} = A_n + B_n i + h_n(i), \tag{3.1b}$$

where

$$A_n := -\log d_\alpha, \qquad B_n := q\alpha - \frac{q}{2} + \log \alpha, \tag{3.1c}$$

Direct subtraction gives this identity also at $i = -1$.

For $S \in \{S_+, S_5\}$ and a nonnegative real profile $\mathbf{k} = (k_i)_{i \in S^{(n)}}$, put $k := \sum_{i \in S^{(n)}} k_i$. When $k > 0$, define

$$L_{\mathbf{k}} := n \log n - n + k - \sum_{i \in S^{(n)}} k_i \log(k_i d_{\alpha - i}), \tag{3.2}$$

with the convention $0 \log 0 = 0$, and let $L_S(n, k)$ be its maximum over real $k_i \geq 0$ satisfying

$$\sum_{i \in S^{(n)}} k_i = k, \qquad \sum_{i \in S^{(n)}} (\alpha - i) k_i = n. \tag{3.3}$$

We call $k > 0$ *feasible for* S if (3.3) has a nonnegative real solution, that is, if $\min S^{(n)} \leq \alpha - n/k \leq \max S^{(n)}$. For k feasible for the support S under consideration, set

$$s := \frac{n}{k}, \qquad T := \alpha - s = \alpha - \frac{n}{k}. \tag{3.4}$$

With $p_i := k_i / k$, the constraints give $T = \sum_{i \in S^{(n)}} i p_i$; thus T is the mean deficit of the normalized profile.

For $S \in \{S_+, S_5\}$, assign positive weights to the allowed deficits. Normalizing by their sum, the partition function, gives a probability distribution. A real parameter, the tilt, changes its mean deficit; the mean map gives this mean as a function of the tilt. The finite and limiting versions are

$$Z_{n,S}(\lambda) := \sum_{i \in S^{(n)}} \exp\{\lambda i + h_n(i)\}, \qquad M_{n,S}(\lambda) := \frac{d}{d\lambda} \log Z_{n,S}(\lambda),$$

$$Z_S(\lambda) := \sum_{i \in S} \exp\Big(\lambda i - \frac{q}{2} i^2\Big), \qquad M_S(\lambda) := \frac{d}{d\lambda} \log Z_S(\lambda).$$

To obtain a prescribed mean deficit, we solve $M_{n,S}(\lambda) = T$ or $M_S(\lambda) = T$. When the solution is unique, write it as $\lambda_{n,S}(T)$ or $\lambda_S(T)$, respectively.

We need the first-moment roots only near the typical average class size. The next lemma gives a narrow interval, called the root corridor, in which the roots are unique and their slopes are controlled. It also relates the finite profile optimization to its limiting one-parameter formula.

In detail, the lemma collects every property of the optimized exponents that is used later: (i) a unique root in a narrow corridor, located by (3.6), which defines r_+ in Section 4 and r_5^{co} in Section 5; (ii) the slope (3.7), which converts differences of the exponent into differences of the number of classes in (4.3c), (5.11) and Appendix C; (iii) the exact support comparison (3.8), used in (5.9a); (iv) uniform convergence of the finite duals and tilts, which bounds ω_n^{root} in (5.9c); and (v) convergence and a positive lower bound for the five-size optimizer, used in (5.15).

Lemma 3.1 (Uniform root, slope, and finite-dual estimates)**.** *Let*

$$s_0 := \alpha_0 - 1 - \frac{2}{\log 2}, \qquad T_0 := \alpha - s_0 = 1 + \frac{2}{\log 2} - \delta. \tag{3.5}$$

There exist a fixed $A_* > 0$ and an integer $n_{\mathrm{root}} \geq 2$ such that, for every $n \geq n_{\mathrm{root}}$, every $S \in \{S_+, S_5\}$, and every $0 \leq c \leq \log 2$, the equation

$$L_S(n,k) + ck = 0$$

has a unique feasible zero $r_{S,c}$ in the corridor

$$\left| \frac{n}{r_{S,c}} - s_0 \right| \leq A_* \frac{\log\log n}{\log n}.$$

No uniqueness outside this corridor is asserted. Uniformly in S, c, and the phase,

$$\frac{n}{r_{S,c}} = s_0 + O\left(\frac{\log\log n}{\log n}\right), \qquad \alpha - \frac{n}{r_{S,c}} = T_0 + O\left(\frac{\log\log n}{\log n}\right). \tag{3.6}$$

For every fixed $A > 0$, uniformly over $S \in \{S_+, S_5\}$, $0 \leq c \leq \log 2$, and k feasible for S with $|n/k - s_0| \leq A \log\log n/\log n$,

$$\frac{\partial}{\partial k}\{L_S(n,k) + ck\} = \frac{2}{\log 2}(\log n)^2 + O_A((\log n)(\log\log n)). \tag{3.7}$$

For the normalized slope error, take the maximum over $S \in \{S_+, S_5\}$ and $0 \leq c \leq \log 2$. For each pair, take the supremum over feasible k satisfying $|n/k - s_0| \leq A \log\log n/\log n$. Thus the deterministic normalized slope error

$$\varepsilon_{n,A}^{\mathrm{slope}} := \max_{S,c} \sup_k \left| \frac{1}{(\log n)^2} \frac{\partial}{\partial k}\{L_S(n,k) + ck\} - \frac{2}{\log 2} \right|$$

satisfies $\varepsilon_{n,A}^{\mathrm{slope}} \to 0$.

For two supports $R, S \in \{S_+, S_5\}$ at a k feasible for both finite supports $R^{(n)}$ and $S^{(n)}$,

$$\frac{L_R(n,k) - L_S(n,k)}{k} = \mathcal{F}_{n,R}(T) - \mathcal{F}_{n,S}(T), \tag{3.8}$$

where the maximum below is over nonnegative weights on $S^{(n)}$ with total mass one and mean T:

$$\mathcal{F}_{n,S}(T) := \max\left[-\sum_{i \in S^{(n)}} p_i \log p_i + \sum_{i \in S^{(n)}} p_i h_n(i) \right]. \tag{3.8a}$$

Fix the compact target interval

$$K_* := \left[\frac{2}{\log 2} - \frac{1}{10}, 1 + \frac{2}{\log 2} + \frac{1}{10} \right] \subset (2,5). \tag{3.9}$$

Then, uniformly for $T \in K_*$,

$$\mathcal{F}_{n,S}(T) \longrightarrow \mathcal{F}_S(T),$$

where $\mathcal{F}_S$ is the function defined in the main theorem. Equivalently, maximize over nonnegative weights on S with total mass one, mean T, and finite second moment, $\sum_{i \in S} i^2 p_i < \infty$:

$$\mathcal{F}_S(T) = \max\left[-\sum_{i \in S} p_i \log p_i - \frac{\log 2}{2} \sum_{i \in S} i^2 p_i \right]. \tag{3.9a}$$

For every $S \in \{S_+, S_5\}$ and $T \in K_*$, the limiting mean equation $M_S(\lambda) = T$ has the unique solution $\lambda_S(T)$. There exist a fixed $\Lambda > 0$ and an integer n_{dual} such that, for every $n \geq n_{\mathrm{dual}}$, every such S, and every $T \in K_*$, the finite equation $M_{n,S}(\lambda) = T$ has the unique solution $\lambda_{n,S}(T)$ and

$$|\lambda_{n,S}(T)| + |\lambda_S(T)| \leq 2\Lambda.$$

For $n \ge n_{\mathrm{dual}}$, define the deterministic dual error

$$\varepsilon_n^{\mathrm{dual}} := \max_{S\in\{S_+,S_5\}} \sup_{T\in K_*} \big(|\mathcal{F}_{n,S}(T) - \mathcal{F}_S(T)| + |\lambda_{n,S}(T) - \lambda_S(T)|\big)$$

and set it arbitrarily, say equal to zero, for $n < n_{\mathrm{dual}}$. Then

$$\varepsilon_n^{\mathrm{dual}} = O(1/\alpha) = O(1/\log n),$$

with a phase-independent implied constant. In particular, $\varepsilon_n^{\mathrm{dual}} \to 0$.

For S_5, both optimizers are unique; denote them by $p_{n,i}(T)$ and $p_i(T)$. Then

$$\sup_{T\in K_*} \max_{2\le i\le 6} |p_{n,i}(T) - p_i(T)| = O(1/\alpha). \tag{3.9b}$$

Moreover, there are $\eta_* > 0$ and $n_{\mathrm{int}} \in \mathbb{N}$ such that

$$p_{n,i}(T) \ge \eta_* \qquad (n \ge n_{\mathrm{int}},\ T \in K_*,\ 2 \le i \le 6).$$

All displayed bounds and thresholds are uniform over the induced phases δ_n, including sequences approaching either endpoint.

Identity (3.8) is exact for a simple reason. By (3.1b) and (3.3), $\sum_i k_i(A_n + B_n i) = kA_n + B_n(\alpha k - n)$ for every admissible profile, and writing $k_i = kp_i$ in (3.2) gives, for k feasible for S,

$$L_S(n,k) = n\log n - n + k - k\log k + kA_n + B_n(\alpha k - n) + k\,\mathcal{F}_{n,S}\Big(\alpha - \frac{n}{k}\Big). \tag{3.9c}$$

Only the last term depends on the support; (A.8) is (3.9c) divided by k.

Section A.2 proves the lemma. It first bounds the finite tilts uniformly on K_*. Stirling's formula then locates the roots and controls their slopes.

4. A uniform lower bound for the chromatic number χ

This section proves estimate (B) of the introduction. A first-moment bound excludes colorings with $k_\chi^- + 1$ classes of size at most $\alpha + 1$: the integer k_χ^- lies about $\log n$ below the root r_+, and the slope (3.7) makes the counting exponent fall by order $(\log n)^3$ over that distance, which outweighs the $\exp\{O((\log n)^2)\}$ profiles. The size cap is then removed at the cost of one class, because independent sets of size $\alpha + 2$ are rare by (2.9).

Let $r_+(n) = r_{S_+,0}$ be the unrestricted-support corridor zero from Lemma 3.1, write $L_+(n,k) := L_{S_+}(n,k)$, and put

$$k_\chi^- := \lfloor r_+(n)\rfloor - \lceil \log n\rceil. \tag{4.1}$$

We choose this deterministic integer a logarithmic distance below the ordinary first-moment root. Over that distance, the slope estimate makes the counting exponent decrease by more than the error incurred when we sum over profiles. Since $r_+(n) = \Theta(n/\log n)$ uniformly in the phase, one has $1 \le k_\chi^- < n$ for all sufficiently large n.

For an integer profile $\mathbf{k} = (k_i)_{-1\le i\le \alpha-1}$ feasible on n vertices, let $X_{\mathbf{k}}$ be the number of unordered proper colorings having exactly k_i classes of size $\alpha - i$. Direct enumeration, as in Heckel (2025, Section 2.3), gives

$$\mathbb{E}[X_{\mathbf{k}}] = \frac{n!}{\prod_i ((\alpha-i)!)^{k_i} k_i!} 2^{-\sum_i k_i \binom{\alpha-i}{2}}. \tag{4.2}$$

For an integer k, let $\mathcal{K}_{n,k,\alpha+1}$ be the set of these profiles satisfying

$$\sum_i k_i = k, \qquad \sum_i (\alpha - i)k_i = n,$$

and define

$$E_{n,k,\alpha+1} := \sum_{\mathbf{k}\in\mathcal{K}_{n,k,\alpha+1}} \mathbb{E}[X_{\mathbf{k}}]\,.$$

Thus $E_{n,k,\alpha+1}$ is the expected number of unordered proper colorings with exactly k nonempty parts, each of size at most $\alpha + 1$. When the feasible profile set is empty, we use the extended-real convention $\log E_{n,k,\alpha+1} = -\infty$.

Summing the capped coloring profiles. There are $\alpha + 1 = O(\log n)$ profile coordinates, and each coordinate lies between zero and n. Hence

$$|\mathcal{K}_{n,k,\alpha+1}| \le (n+1)^{\alpha+1} = \exp\{O((\log n)^2)\} \tag{4.3}$$

uniformly in k and in the phase. Keep the class-size factorials inside $d_{\alpha-i}$ exact, and put

$$R(m) := \log(m!) - (m \log m - m), \qquad R(0) := 0.$$

For every feasible profile, exact cancellation in (4.2) gives

$$\log \mathbb{E}[X_{\mathbf{k}}] - L_{\mathbf{k}} = R(n) - \sum_i R(k_i).$$

The uniform estimate (1.2) therefore gives

$$\log \mathbb{E}[X_{\mathbf{k}}] \le L_{\mathbf{k}} + O((\log n)^2).$$

The error is uniform even when some $k_i = 0$: there are $O(\log n)$ profile coordinates and $|R(k_i)| = O(\log(n+1))$. For every k feasible for $S_+^{(n)}$, maximizing over the feasible profile and then summing (4.3) yields

$$\log E_{n,k,\alpha+1} \le L_+(n,k) + O((\log n)^2). \tag{4.3a}$$

Displacement below the root. Set

$$\Delta_n := r_+(n) - k_\chi^-.$$

The floor and ceiling in (4.1) give the deterministic bounds

$$\log n \le \Delta_n < \log n + 2. \tag{4.3b}$$

We next verify that the entire interval $[k_\chi^-, r_+(n)]$ lies in the derivative corridor of Lemma 3.1. Write $s(k) = n/k$. Uniformly on this interval, $k = \Theta(n/\log n)$ and $s(k) = \Theta(\log n)$. Therefore

$$|s(k_\chi^-) - s(r_+)| \le \sup_{k \in [k_\chi^-, r_+]} \frac{n}{k^2}\,\Delta_n = O\left(\frac{(\log n)^3}{n}\right) = o\left(\frac{\log\log n}{\log n}\right).$$

Thus, after one phase-independent threshold, every point of the interval satisfies

$$\left|\frac{n}{k} - s_0\right| \le (A_* + 1)\frac{\log\log n}{\log n}.$$

The same estimates place $T(k) = \alpha - n/k$ in $K_* \subset (-1, \alpha - 1)$ throughout this interval, so every point is feasible for S_+. Equation (3.7), with $A = A_* + 1$, $S = S_+$, and $c = 0$, consequently gives an absolute $c_{\mathrm{sl}} > 0$ such that

$$\frac{d}{dk} L_+(n,k) \ge c_{\mathrm{sl}} (\log n)^2 \qquad (k_\chi^- \le k \le r_+) \tag{4.3c}$$

for all sufficiently large n, uniformly in the phase. The size-cap argument below uses this one class above k_χ^-. Put $k' := k_\chi^- + 1$. By (4.3b), $k' \in [k_\chi^-, r_+]$ and $r_+ - k' \ge \log n - 1$, so $L_+(n, r_+) = 0$, the mean-value theorem and (4.3c) give $L_+(n,k') \le -c_{\mathrm{sl}}(\log n)^2(\log n - 1)$. With (4.3a), there is an absolute $c_1 > 0$ such that

$$E_{n,k',\alpha+1} \le e^{-c_1 (\log n)^3} \tag{4.4}$$

(the left side is zero if its feasible profile set is empty).

Removing the size cap with one additional class. For an integer s, let X_s count independent s-subsets of G_n, and put $\mu_s = \mathbb{E}X_s = \binom{n}{s} 2^{-\binom{s}{2}}$. For the size-cap argument, it is enough to rule out the presence of two distinct independent sets of the exceptional size. The next bound therefore estimates their pair count, allowing a single exceptional set.

Lemma 4.1 (An independent-set pair bound)**.** *For integers $1 \le s < n$, put*

$$a = \frac{s^2\, 2^{s/4}}{n - s}, \qquad b = \sqrt{2}\, sn\, 2^{-3s/4}.$$

Then

$$\mathbb{E}[X_s(X_s - 1)] \le \mu_s^2 e^a + \mu_s(e^b - 1).$$

Section A.3 proves the lemma. It splits the ordered pairs by the size of their intersection: a union bound handles intersections of size at most $s/2$, and a conditional first moment handles larger ones.

For any graph on n vertices, if $2 \le s < n$ and $X_s \le 1$, no independent set has more than s vertices. Otherwise an independent $(s+1)$-set would contain at least two independent s-sets. Moreover, a proper coloring has at most one class of size s. If such a class is present, split off one vertex as a singleton. Thus, whenever $1 \le k < n$, a proper coloring with at most k classes yields one with at most $k+1$ classes, all of size at most $s-1$. Since the target class count does not exceed the number of vertices, further splitting gives exactly $k+1$ nonempty classes.

Apply this observation with $s = \alpha + 2$ and $k = k_\chi^-$, using the same n and cutoff throughout. If $X_s \le 1$ and $\chi(G_n) \le k_\chi^-$, the observation gives a proper coloring with exactly $k_\chi^- + 1$ nonempty classes, all of size at most $\alpha + 1$. Its profile lies in $\mathcal{K}_{n,k_\chi^-+1,\alpha+1}$, so $\sum_{\mathbf{k}\in\mathcal{K}_{n,k_\chi^-+1,\alpha+1}} X_{\mathbf{k}} \ge 1$, and Markov's inequality bounds the probability of this event by $E_{n,k_\chi^-+1,\alpha+1}$. Since $X_s(X_s-1) \ge 2$ on $\{X_s \ge 2\}$, Markov's inequality also gives $\mathbb{P}(X_s \ge 2) \le \frac{1}{2}\mathbb{E}[X_s(X_s-1)]$, which Lemma 4.1 bounds. Hence

$$\mathbb{P}(\chi(G_n) \le k_\chi^-) \le \frac{\mu_s^2 e^a + \mu_s(e^b-1)}{2} + E_{n,k_\chi^-+1,\alpha+1}.$$

By (4.4), we have proved

$$\mathbb{P}(\chi(G_n) \le k_\chi^-) \le \frac{\mu_s^2 e^a + \mu_s(e^b-1)}{2} + e^{-c_1(\log n)^3}, \qquad s = \alpha + 2. \tag{4.5}$$

The phase formula gives $2^s = \Theta(n^2/(\log n)^2)$ uniformly, hence $a = O((\log n)^{3/2}/\sqrt{n})$ and $b = O((\log n)^{5/2}/\sqrt{n})$. Together with $\mu_{\alpha+2} \to 0$ from (2.9), this proves

$$\mathbb{P}(\chi(G_n) > k_\chi^-) \longrightarrow 1. \tag{4.6}$$

The sharper form of (4.5) will also yield a summable failure bound below. Finally, (4.3b) records the unchanged location precision:

$$|k_\chi^- - r_+(n)| < \log n + 2 = o\left(\frac{n}{(\log n)^3}\right). \tag{4.7}$$

5. Five-size entropy gain and the selected integer profile

Restricting the allowed deficits to $S_5 = \{2,3,4,5,6\}$ reduces the number of available profiles, but makes the later overlap estimates uniform. We compare the resulting entropy loss per part with $\log 2$, the gain from the 2^k class declarations. The loss is smaller, and Lemma 5.3 gives a uniform margin for the resulting positive net gain. The slope estimate in Lemma 3.1 then converts this gain in the counting exponent into a separation between the roots.

5.1. **The phase coefficient and entropy bound.** For $S \in \{S_+, S_5\}$ write $w_i(\lambda) := \exp\bigl(\lambda i - \frac{\log 2}{2} i^2\bigr)$, so that $Z_S(\lambda) = \sum_{i\in S} w_i(\lambda)$ and $M_S(\lambda) = \sum_{i\in S} i\, w_i(\lambda)/Z_S(\lambda) = (\log Z_S)'(\lambda)$ as in Section 3.

The Gaussian factor makes both sums converge. Moreover, $M_S'(\lambda) = \mathrm{Var}_{S,\lambda}(i) > 0$, and the endpoint limits of M_S are the endpoints of the convex hull of S. Thus, for every $T \in (2,5)$, there is a unique tilt $\lambda_S(T)$ with $M_S(\lambda_S(T)) = T$. The Gibbs identity (A.7) proves that the optimizer in (3.9a) is

$$p_i = \frac{w_i(\lambda_S(T))}{Z_S(\lambda_S(T))} = \frac{e^{\lambda_S(T) i - \log 2 i^2/2}}{\sum_{j\in S} e^{\lambda_S(T) j - \log 2 j^2/2}}. \tag{5.1}$$

This also gives the dual representation

$$\mathcal{F}_S(T) = \inf_{\lambda\in\mathbb{R}} \{\log Z_S(\lambda) - \lambda T\} = \log Z_S(\lambda_S(T)) - \lambda_S(T) T.$$

These facts give the side assertions of the main theorem. By (2.0), $T_\delta = 1 + 2/q - \delta \in [2/q, 1+2/q] \subset (14/5, 4) \subset (2,5)$ for $0 \le \delta \le 1$. For $T \in (2,5)$, the map $\lambda \mapsto \log Z_S(\lambda) - \lambda T$ is strictly convex, with derivative $M_S(\lambda) - T$, so its infimum $\mathcal{F}_S(T)$ is finite and attained only at $\lambda_S(T)$. Continuity and nonconstancy of A_5 are proved in Lemma 5.1, and positivity in Lemma 5.3.

Since $T_0 = 1 + 2/\log 2 - \delta = T_\delta$, introduce the entropy loss D_5 and the uniform bound γ_5 by

$$D_5(\delta) := \mathcal{F}_{S_+}(T_\delta) - \mathcal{F}_{S_5}(T_\delta), \qquad A_5(\delta) = \log 2 - D_5(\delta), \qquad \gamma_5 := \log\frac{8750}{6681}. \tag{5.2}$$

To find the smallest phase gain, compare the two optimizing distributions at the same mean deficit. The five-size distribution has smaller variance than the unrestricted one. The following lemma translates this comparison into strict concavity of the phase gain.

Lemma 5.1 (Curvature and minimum of the phase gain)**.** *For $S \in \{S_+, S_5\}$ and $T \in (2,5)$, let*

$$V_S(T) := \sum_{i \in S} (i-T)^2 \frac{w_i(\lambda_S(T))}{Z_S(\lambda_S(T))}.$$

Then $0 < V_{S_5}(T) < V_{S_+}(T)$. The function A_5 extends to a twice continuously differentiable function on an open interval containing $[0,1]$, and

$$A_5'(\delta) = \lambda_{S_5}(T_\delta) - \lambda_{S_+}(T_\delta), \qquad A_5''(\delta) = \frac{1}{V_{S_+}(T_\delta)} - \frac{1}{V_{S_5}(T_\delta)} < 0.$$

In particular, A_5 is strictly concave and nonconstant on $[0,1]$, and

$$\min_{0 \le \delta \le 1} A_5(\delta) = \min\{A_5(0), A_5(1)\}.$$

Section B.1 proves the lemma by differentiating the dual values twice in T and comparing the variances of the two optimizing distributions.

For comparison only, extend the same tilted definitions to $S_4 := \{2,3,4,5\}$, and put

$$A_4(\delta) := q - \mathcal{F}_{S_+}(T_\delta) + \mathcal{F}_{S_4}(T_\delta), \qquad D_4(\delta) := q - A_4(\delta).$$

This four-size calculation gives an explicit reference bound for the entropy loss. We use it to obtain the stronger five-size bound.

Lemma 5.2 (Four-size comparison certificate)**.** *For $0 \le \delta \le 1$,*

$$0 \le D_4(\delta) < \log(153/100), \qquad A_4(\delta) > \log(200/153).$$

Its proof in Section B.2 uses only the four-size optimizer; none of the profiles selected below uses that support.

Lemma 5.3 (Uniform entropy bound and strict five-size gain)**.** *For every $0 \le \delta \le 1$,*

$$A_5(\delta) - A_4(\delta) > \log(525/524), \qquad 0 \le D_5(\delta) < \log(6681/4375), \qquad A_5(\delta) > \gamma_5. \tag{5.3}$$

Section B.3 proves the lemma. It brackets the five-size tilt at T_δ and bounds from below the optimal mass at deficit six, which is what S_5 gains over S_4.

Lemma 5.3 suffices for every positivity argument below. The lower bound for the uniform constant in the main theorem uses the following sharper bound. By Lemma 5.1, the minimum of A_5 on $[0,1]$ is attained at an endpoint. Section B.6 records numerical values of $A_5(0)$ and $A_5(1)$; they are not used in any proof.

Lemma 5.4 also implies $A_5 > \gamma_5$. We keep Lemmas 5.2 and 5.3 because their proofs do not use the numerical endpoint table of Section B.5, and because the bound $A_4 > \log(200/153)$ of Lemma 5.2 gives the coefficient $((\log 2)^2/4)\log(200/153)$ of the released Lean theorem (Appendix H).

Lemma 5.4 (A sharp uniform bound for the phase gain)**.** *For every $0 \le \delta \le 1$,*

$$A_5(\delta) > \log \frac{169}{100}.$$

Section B.4 proves the lemma. It bounds the five-size dual from below by convexity and the unrestricted dual from above at an explicit tilt, on two cells that cover the phase interval; the resulting inequalities reduce to rational evaluations at the cell endpoints (Table 1).

5.2. **The gain from class marks and root separation.** The factor 2^k comes from the one-partition identity in Heckel (2025, Proposition 6). We recall its short proof before using it to compare the roots. For a fixed partition into k classes, assigning each class the declaration "independent" or "complete" multiplies its first moment by 2^k: every declaration prescribes all internal edge bits, and all 2^k declarations have probability $2^{-\sum \binom{u}{2}}$. The declarations are disjoint because every allowed class size is at least two for all sufficiently large n. Put

$$\Phi_n(k) := L_{S_5}(n,k) + (\log 2)k, \tag{5.9}$$

and let $r_5^{\mathrm{co}} := r_{S_5,\log 2}$ be its unique corridor zero from Lemma 3.1. To compare the two roots, evaluate Φ_n at the unrestricted root. Let

$$T_+(n) := \alpha - \frac{n}{r_+(n)}$$

and introduce the finite-n support loss

$$D_{5,n}(T) := \mathcal{F}_{n,S_+}(T) - \mathcal{F}_{n,S_5}(T).$$

By (3.6) and the fixed margin in (3.9), one has $T_+(n) \in K_* \subset (2,5)$ for all sufficiently large n, uniformly in the phase. Thus r_+ is feasible for both supports before (3.8) is applied. Since $L_{S_+}(n, r_+) = 0$, the exact support-comparison identity (3.8) gives

$$\begin{aligned}\Phi_n(r_+) &= L_{S_5}(n, r_+) - L_{S_+}(n, r_+) + (\log 2) r_+ \\ &= r_+\{\log 2 - D_{5,n}(T_+(n))\}.\end{aligned} \tag{5.9a}$$

For finite n, this identity expresses the marked counting exponent at the unrestricted root in terms of the loss from restricting the support. We next replace that finite loss by its phase-dependent limit.

There are two errors in this replacement: the finite dual values differ from their limits, and the root target $T_+(n)$ differs from the phase-center target T_0. For the latter, define

$$\varepsilon_n^{\text{target}} := |T_+(n) - T_0|.$$

By (3.6), uniformly in the phase,

$$\varepsilon_n^{\text{target}} = O\left(\frac{\log\log n}{\log n}\right). \tag{5.9b}$$

The limiting dual value is differentiable, with

$$\mathcal{F}_S'(T) = -\lambda_S(T),$$

because the envelope theorem cancels the derivative of the optimizing tilt. Section 3 bounds the limiting tilts by one constant on K_*, so the two limiting dual values have a common Lipschitz constant C_{Lip}. The deterministic finite-dual error from Section 3 satisfies

$$\sup_{T \in K_*} |\mathcal{F}_{n,S}(T) - \mathcal{F}_S(T)| \le \varepsilon_n^{\text{dual}} \qquad (S = S_+, S_5).$$

Consequently the single root error used in the final theorem,

$$\omega_n^{\text{root}} := 2\varepsilon_n^{\text{dual}} + 2C_{\text{Lip}}\varepsilon_n^{\text{target}} \longrightarrow 0 \tag{5.9c}$$

is deterministic and phase-uniform. Indeed,

$$\begin{aligned}|D_{5,n}(T_+(n)) - D_5(\delta)| \le\ & |\mathcal{F}_{n,S_+}(T_+) - \mathcal{F}_{S_+}(T_+)| \\ & + |\mathcal{F}_{n,S_5}(T_+) - \mathcal{F}_{S_5}(T_+)| \\ & + |\mathcal{F}_{S_+}(T_+) - \mathcal{F}_{S_+}(T_0)| \\ & + |\mathcal{F}_{S_5}(T_+) - \mathcal{F}_{S_5}(T_0)| \\ \le\ & \omega_n^{\text{root}}.\end{aligned}$$

The quantitative finite-dual estimate in Section 3 and (5.9b) give

$$\omega_n^{\text{root}} = O\left(\frac{\log\log n}{\log n}\right),$$

with a phase-independent implied constant.

Substitution into (5.9a) yields the estimate used in the final theorem,

$$L_{S_5}(n, r_+) + (\log 2) r_+ = r_+\{\log 2 - D_5(\delta) + O(\omega_n^{\text{root}})\} = r_+\{\log 2 - D_5(\delta) + o(1)\}. \tag{5.10}$$

Finally, Lemma 5.3 gives $\log 2 - D_5(\delta) \ge \gamma_5 > 0$. Hence (5.10) implies $\Phi_n(r_+) > 0$ for all sufficiently large n, uniformly in the phase. The normalized slope error in (3.7) tends to zero, so the derivative of Φ_n is positive throughout the common root corridor after one phase-independent threshold. Since $\Phi_n(r_5^{\text{co}}) = 0$, this proves

$$r_5^{\text{co}} < r_+.$$

The exponent gain becomes a root separation by division by the slope. Both roots lie in (3.6), so the slope estimate holds throughout the interval between them. The mean-value theorem gives a point $\xi_n \in (r_5^{\text{co}}, r_+)$ such that

$$\Phi_n(r_+) - \Phi_n(r_5^{\text{co}}) = \Phi_n'(\xi_n)(r_+ - r_5^{\text{co}}).$$

The corridor estimate (3.6) gives

$$r_+ = \frac{\log 2}{2}\frac{n}{\log n}\left\{1 + O\left(\frac{\log\log n}{\log n}\right)\right\},$$

Indeed, (2.1) and (3.5) give $s_0 = \alpha_0 - 1 - 2/q = (2/q)(L - \ell + C - 1) = (2/q)L\{1 + O(\ell/L)\}$, where $L = \log n$ and $\ell = \log\log n$, and (3.6) gives $n/r_+ = s_0 + O(\ell/L)$; dividing n by n/r_+ proves the display. Equation (3.7), uniformly at ξ_n, gives

$$\Phi_n'(\xi_n) = \frac{2}{\log 2}(\log n)^2 + O((\log n)(\log\log n)).$$

Since $\Phi_n(r_5^{\mathrm{co}}) = 0$, (5.10) and the two displays above give

$$r_+ - r_5^{\mathrm{co}} = \frac{\Phi_n(r_+)}{\Phi_n'(\xi_n)} = \frac{\frac{\log 2}{2}\frac{n}{\log n}\{A_5(\delta_n) + O(\omega_n^{\mathrm{root}})\}}{\frac{2}{\log 2}(\log n)^2}\Big\{1 + O\Big(\frac{\log\log n}{\log n}\Big)\Big\}.$$

As A_5 is bounded and $\omega_n^{\mathrm{root}} = O(\log\log n/\log n)$, this is

$$r_+ - r_5^{\mathrm{co}} = \frac{(\log 2)^2}{4} A_5(\delta_n)\frac{n}{(\log n)^3} + O\Big(\frac{n\log\log n}{(\log n)^4}\Big). \tag{5.11}$$

The implied constant is independent of the phase: A_5 is bounded on $[0,1]$, and all three relative errors above are uniform. This remainder concerns the analytic roots. Rounding and amplification introduce further errors in the probabilistic bound.

For the fixed-offset selection below, we use the following coarse consequence of (5.11): for all sufficiently large n,

$$r_+ - r_5^{\mathrm{co}} \geq \frac{(\log 2)^2\gamma_5}{8}\frac{n}{(\log n)^3}. \tag{5.12}$$

5.3. **An exact integer profile above the first-moment root for marked partitions.** Choose an integer a fixed distance above the first-moment root for marked partitions:

$$k_{\mathrm{co}} = \lceil r_5^{\mathrm{co}} \rceil + 16. \tag{5.13}$$

Thus

$$16 \leq k_{\mathrm{co}} - r_5^{\mathrm{co}} < 17.$$

By (5.12), $k_{\mathrm{co}} < r_+$ for all sufficiently large n. Consequently the common corridor in Lemma 3.1 contains the entire interval from r_5^{co} to k_{co}. Every real point of this interval therefore has target deficit mean in the fixed compact interval $K_* \subset (2,5)$ and is feasible for the real S_5 optimization. In particular this holds at the selected integer before the type counts are rounded. The convenient choice 16 leaves a fixed explicit margin in (5.19).

We now round the optimizing real class counts to obtain the *selected integer profile*. The rounding must preserve both finite-n conservation constraints: the total number of classes and the total number of vertices.

At $k = k_{\mathrm{co}}$, let $p_i^{(n)}$ be the normalized proportions of the finite-n maximizing profile for L_{S_5}. Let $\lambda_{n,\mathrm{def}}$ denote the Lagrange multiplier for the deficit-mean constraint, with the sign convention that its contribution is $e^{\lambda_{n,\mathrm{def}} i}$. The Lagrange equations and the decomposition (3.1b) then give

$$p_i^{(n)} \;\propto\; d_{\alpha-i}^{-1} e^{\lambda_{n,\mathrm{def}} i} \;\propto\; e^{\widehat{\lambda}_n i + h_n(i)}, \qquad \widehat{\lambda}_n := B_n + \lambda_{n,\mathrm{def}}. \tag{5.14}$$

The finite-support tilt $\widehat{\lambda}_n$ includes the affine contribution $B_n i$. The target deficit mean is $\alpha - n/k_{\mathrm{co}}$. By uniqueness in Lemma 3.1, $\widehat{\lambda}_n$ is the finite-support tilt $\lambda_{n,S_5}(\alpha - n/k_{\mathrm{co}})$ introduced there. The corridor estimate and $k_{\mathrm{co}} - r_5^{\mathrm{co}} = O(1)$ give, uniformly in the phase,

$$\Big|\alpha - \frac{n}{k_{\mathrm{co}}} - T_0\Big| = O\Big(\frac{\log\log n}{\log n}\Big) + O\Big(\frac{(\log n)^2}{n}\Big) \longrightarrow 0.$$

Lemma 3.1 therefore makes these proportions converge uniformly to (5.1), with $\widehat{\lambda}_n - \lambda_{S_5}(T_0) \to 0$ uniformly in the phase. Equivalently, the computation leading to (3.9c) shows that, for a real profile with k_{co} parts and n vertices, $L_{\mathbf{k}}$ equals a function of k_{co} and n alone plus k_{co} times the bracket in (3.8a) at $p_i = k_i/k_{\mathrm{co}}$, with $S = S_5$ and mean $T = \alpha - n/k_{\mathrm{co}}$. Hence $p_i^{(n)} = p_{n,i}(T)$, the finite optimizer of Lemma 3.1, and (5.14) is its Gibbs form. Since $\alpha - n/k_{\mathrm{co}} \in K_*$, the last assertion of Lemma 3.1 gives (5.15) with $c_p := \eta_*$:

$$\min_{2\leq i\leq 6} p_i^{(n)} \geq c_p > 0 \tag{5.15}$$

uniformly in the phase.

For each sufficiently large n, fix a deterministic choice of preliminary integers $\widetilde{k}_i$ with $|\widetilde{k}_i - k_{\mathrm{co}} p_i^{(n)}| \leq 1$. The estimates below hold for every such choice. Define the signed constraint errors

$$e_0 = \sum_i \widetilde{k}_i - k_{\mathrm{co}}, \qquad e_1 = \sum_i i\widetilde{k}_i - (\alpha k_{\mathrm{co}} - n),$$

so $e_0, e_1 = O(1)$, and add

$$\Delta k_2 = e_1 - 3e_0, \qquad \Delta k_3 = 2e_0 - e_1. \tag{5.16}$$

Set $\Delta k_4 = \Delta k_5 = \Delta k_6 = 0$, $k_i := \widetilde{k}_i + \Delta k_i$ for $2 \le i \le 6$, and $\mathbf{k} := (k_2, k_3, k_4, k_5, k_6)$.

The correction removes both constraint errors by changing only $O(1)$ counts. Appendix C checks positivity and shows that the loss in the counting exponent is $O(1/k_{\mathrm{co}})$.

The resulting exact integer vector satisfies

$$u_i = \alpha - i, \qquad k_i = \Theta(n/\log n) \quad (2 \le i \le 6). \tag{5.17}$$

Let $Z_{\mathbf{k}}^{\mathrm{sgn}}$ be the number of pairs consisting of a partition with exactly k_i classes of size u_i for $2 \le i \le 6$ and a declaration of each class as independent or complete, such that every declaration is realized by G_n. Direct counting gives the exact identity

$$\mathbb{E}\big[Z_{\mathbf{k}}^{\mathrm{sgn}}\big] = 2^{k_{\mathrm{co}}} \frac{n!}{\prod_i (u_i!)^{k_i} k_i!} 2^{-\sum_i k_i \binom{u_i}{2}}, \tag{5.18}$$

Keeping every $u_i!$ exact and applying (1.2) only to $n!$ and the five profile-multiplicity factorials $k_i!$ then gives

$$\log \mathbb{E}\big[Z_{\mathbf{k}}^{\mathrm{sgn}}\big] = L_{S_5}(n, k_{\mathrm{co}}) + \log 2k_{\mathrm{co}} + O(\log n) \ge c_Z (\log n)^2 \tag{5.19}$$

for a phase-independent $c_Z > 0$. Appendix C verifies that the fixed displacement above the root exceeds the rounding and Stirling losses. Every u_i tends to infinity, so the independent and complete declarations of one part are disjoint. Thus $Z_{\mathbf{k}}^{\mathrm{sgn}} > 0$ implies an actual cocoloring with k_{co} parts.

This fixes a deterministic profile for the rest of the proof. The next proposition collects the exact constraints, the first-moment bound and the root separation that the second-moment argument will use.

Proposition 5.5 (The selected five-size profile)**.** *For each sufficiently large n, choose k_{co} by (5.13) and $\mathbf{k} = (k_2, k_3, k_4, k_5, k_6)$ by the deterministic rounding and correction above. There are phase-independent constants $c_* > 0$, $c_Z > 0$ and an integer n_0 such that, for every integer $n \ge n_0$, the following hold.*

(i) *With $u_i = \alpha - i$ for $2 \le i \le 6$, the counts k_i are positive integers and satisfy exactly*

$$\sum_{i=2}^{6} k_i = k_{\mathrm{co}}, \qquad \sum_{i=2}^{6} u_i k_i = n.$$

Moreover, $k_i/k_{\mathrm{co}} \ge c_*$ for every i, and $k_{\mathrm{co}} \asymp n/\log n$. The mean deficit satisfies

$$T := \alpha - \frac{n}{k_{\mathrm{co}}} = T_0 + O\left(\frac{\log\log n}{\log n}\right) \le 39/10.$$

(ii) *For $G_n \sim G(n, 1/2)$, the marked cocoloring count defined in (5.18) obeys*

$$\log \mathbb{E}\big[Z_{\mathbf{k}}^{\mathrm{sgn}}\big] \ge c_Z (\log n)^2.$$

A positive count gives a cocoloring with exactly k_{co} parts.

(iii) *The ordinary and signed roots satisfy*

$$r_+(n) - r_5^{\mathrm{co}}(n) = \frac{(\log 2)^2}{4} A_5(\delta_n) \frac{n}{(\log n)^3} + O\left(\frac{n \log\log n}{(\log n)^4}\right).$$

For the integer chromatic lower location k_χ^- in (4.1),

$$\begin{aligned} k_\chi^- - k_{\mathrm{co}} &= \left[\frac{(\log 2)^2}{4} A_5(\delta) + o(1)\right] \frac{n}{(\log n)^3} \\ &\ge \left[\frac{(\log 2)^2 \gamma_5}{4} - o(1)\right] \frac{n}{(\log n)^3}. \end{aligned} \tag{5.20}$$

Here $\delta = \delta_n$. The choices, constants and error sequences are deterministic. The implied constants and vanishing errors are uniform over the induced phases, including integer sequences approaching either endpoint.

Proof. Equation (5.16) removes both rounding errors exactly. Appendix C gives $k_i \ge (c_p/2) k_{\mathrm{co}}$ after one phase-independent threshold; take $c_* = c_p/2$. The root corridor and (5.13) give the order of k_{co} and the displayed estimate for T. Since $q > 56/81$, one has $T_0 \le 1 + 2/q < 109/28 = 39/10 - 1/140$. The displayed error therefore gives $T \le 39/10$ after one phase-independent threshold. This window concerns the selected target, not every point of K_*. Equation (5.19) proves the first-moment bound, and the paragraph following it identifies a realized marked cocoloring with a cocoloring. The root separation is (5.11). Finally, (4.1) and (5.13) change its two roots by $O(\log n)$ and $O(1)$ respectively, giving (5.20). □

6. Second moment of the marked cocoloring count

A large first moment alone does not give a useful probability of a cocoloring. The second-moment estimates in Sections 6–9 supply that probability through a deterministic exponent Λ_n satisfying

$$\frac{\mathbb{E}[(Z_{\mathbf{k}}^{\mathrm{sgn}})^2]}{(\mathbb{E}Z_{\mathbf{k}}^{\mathrm{sgn}})^2} \le e^{\Lambda_n}$$

and give the scale hierarchy

$$k_{\mathrm{co}} \asymp \frac{n}{\log n}, \qquad r_+ - r_5^{\mathrm{co}} \asymp \frac{n}{(\log n)^3}, \qquad \Lambda_n = o\left(\frac{n}{(\log n)^4}\right).$$

Once the last estimate is available, its amplification cost (by vertex exposure the uncovered set has expected size at most $\sqrt{(n-1)\Lambda_n/2}$, (10.5), and (10.3) colors any vertex set S with $C_0|S|/\log n + n^{1/3}$ colors) satisfies

$$\frac{\sqrt{n\Lambda_n}}{\log n} = o\left(\frac{n}{(\log n)^3}\right),$$

which is smaller than the root separation.

Fix the deterministic five-size profile from Proposition 5.5. The marked cocoloring count has a large first moment. To bound from below the probability that it is positive, we must determine which edge constraints two partitions share and whether their class marks, which we also call signs, prescribe consistent constraints. We begin by counting the compatible pairs of markings. For partitions sharing exactly ℓ whole classes, Heckel (2025, Proposition 6) used the bound $2^{2k-\ell}$ to transfer selected ordinary-coloring estimates to cocolorings. Here we retain the full overlap matrix. Lemma 6.1 counts the compatible markings using the connected components formed by cells of size at least two, including cells that are whole classes in both partitions.

Overlap law and compatible signs. Temporarily distinguish the class slots within each of the five types. This multiplies the marked partition count by the deterministic factor $\prod_i k_i!$, and therefore leaves its normalized second moment unchanged. Let $\mathcal{I}_{\mathrm{row}}$ and $\mathcal{I}_{\mathrm{col}}$ be the row- and column-slot index sets. Choose two independent uniform ordered partitions of the profile, write them as $(V_a)_{a\in\mathcal{I}_{\mathrm{row}}}$ and $(W_b)_{b\in\mathcal{I}_{\mathrm{col}}}$, and set $s_a := |V_a|$ and $t_b := |W_b|$. Define

$$\mathbf{s} := (s_a)_{a\in\mathcal{I}_{\mathrm{row}}}, \qquad \mathbf{t} := (t_b)_{b\in\mathcal{I}_{\mathrm{col}}},$$

and let $\mathcal{R}(\mathbf{s},\mathbf{t})$ be the finite set of nonnegative integer matrices with row margins $\mathbf{s}$ and column margins $\mathbf{t}$. Put

$$r_{ab} = |V_a \cap W_b|. \tag{6.1}$$

The overlap matrix has the exact law

$$p(r) = \frac{\prod_a s_a! \prod_b t_b!}{n! \prod_{a,b} r_{ab}!}, \tag{6.2}$$

Both degree lists are the multiset with k_i copies of u_i, $2 \le i \le 6$. We can obtain the same law by placing s_a stubs at each row slot and t_b stubs at each column slot, then choosing a uniform perfect matching of the two n-stub sets. Each fixed ordered column partition arises from the same number, $\prod_b t_b!$, of column-stub orderings. Thus this configuration model induces exactly the law (6.2).

Let $H(r)$ be the simple bipartite graph whose edges are cells with $r_{ab} \ge 2$, omitting isolated slot vertices. Write $H := H(r)$ and

$$W = \sum_{a,b} \binom{r_{ab}}{2}, \qquad \beta(H) = |E(H)| - |V(H)| + c(H). \tag{6.3}$$

We use the standard convention $c(\varnothing) = 0$, so that $\beta(\varnothing) = 0$.

Figure 2 illustrates how cells of size at least two determine the support graph, and how its cycle-space factor combines with the local cell factors.

For a fixed pair of partitions, divide the expected number of compatible pairs of valid markings by the square of the one-partition expectation. The next lemma expresses this normalized factor in terms of the overlap table.

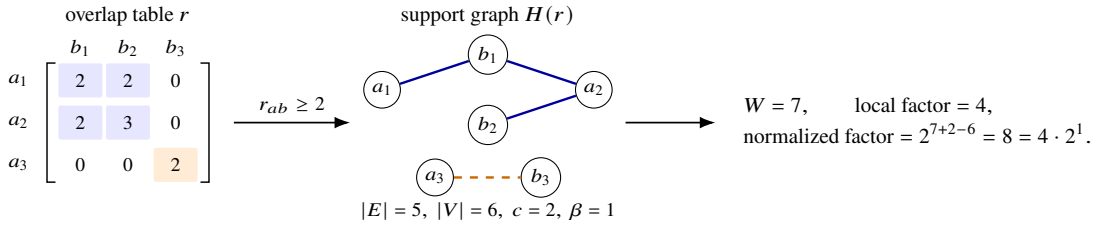


FIGURE 2. From an overlap table to the signed support factor. The shaded entries are precisely those with multiplicity at least two. Their support consists of one four-cycle (solid blue) and one isolated edge (dashed orange), so its binary cycle space has dimension one. Compatible signs are constant on each connected component; the colors and line styles mark the two components and carry no sign information. The displayed arithmetic verifies (6.4) for this example.

Lemma 6.1 (Exact sign sum)**.** *For every overlap matrix $r \in \mathcal{R}(\mathbf{s}, \mathbf{t})$, the normalized local factor of the corresponding pair of partitions is*

$$A_\zeta(r) = 2^{W+c(H)-|V(H)|} = \left(\prod_{a,b} g(r_{ab})\right) 2^{\beta(H)}, \tag{6.4}$$

where

$$g(0) = g(1) = g(2) = 1, \qquad g(x) = 2^{\binom{x}{2}-1} \quad (x \geq 3). \tag{6.5}$$

Consequently

$$\frac{\mathbb{E}\left[(Z_{\mathbf{k}}^{\mathrm{sgn}})^2\right]}{\mathbb{E}\left[Z_{\mathbf{k}}^{\mathrm{sgn}}\right]^2} = \sum_{r \in \mathcal{R}(\mathbf{s},\mathbf{t})} p(r) A_\zeta(r). \tag{6.6}$$

Here the sum ranges over the matrix set $\mathcal{R}(\mathbf{s}, \mathbf{t})$ defined above.

Proof. Fix the two partitions. Let $B = \sum_a \binom{s_a}{2}$ be the number of edge indicators prescribed by each partition. A cell of size at least two contains an edge whose presence is prescribed by both its row and its column. Their signs must therefore agree. Repeating this condition along the support edges forces a constant sign on each component of H. Conversely, every edge bit prescribed by both partitions lies in a cell of size at least two, so signs that are constant on the components of H prescribe it consistently.

The number of compatible pairs of sign assignments, including the free signs on slots isolated from H, is

$$2^{2k_{\mathrm{co}}-|V(H)|+c(H)}.$$

For each compatible pair, W edge bits are prescribed twice, so the probability of all constraints is $2^{-(2B-W)}$. Division by the square of the expected number of valid markings of one fixed partition $(2^{k_{\mathrm{co}}}2^{-B})^2$ gives

$$\frac{2^{2k_{\mathrm{co}}-|V(H)|+c(H)}2^{-(2B-W)}}{(2^{k_{\mathrm{co}}}2^{-B})^2} = 2^{W+c(H)-|V(H)|},$$

which is the first expression in (6.4).

For every support edge, split $\binom{r}{2}$ as $\left(\binom{r}{2} - 1\right) + 1$. Since $|E| - |V| + c = \beta$,

$$\prod_{a,b} g(r_{ab}) = 2^{W-|E(H)|},$$

and multiplying by $2^{\beta(H)}$ gives the second expression.

To average, let $\mathcal{P}$ be the finite set of ordered partitions with the selected slot sizes, and for $P \in \mathcal{P}$ let Z_P be the number of valid markings of P, so that $\mathbb{E}Z_P = 2^{k_{\mathrm{co}}}2^{-B}$ for every P. By the computation above, $\mathbb{E}[Z_P Z_Q] = (2^{k_{\mathrm{co}}}2^{-B})^2 A_\zeta(r(P,Q))$, where $r(P,Q)$ is the overlap matrix of P and Q. Expanding the square of $\sum_P Z_P$ therefore gives

$$\frac{\mathbb{E}[(\sum_P Z_P)^2]}{(\mathbb{E}\sum_P Z_P)^2} = \frac{1}{|\mathcal{P}|^2} \sum_{P,Q \in \mathcal{P}} A_\zeta(r(P,Q)),$$

the average of $A_\zeta(r(P,Q))$ over independent uniform P and Q. Averaging over (6.2) proves (6.6). □

Call an edge set even if every vertex has even degree in that edge set. The identity

$$2^{\beta(H)} = \#\{F \subseteq E(H) : \deg_F(v) \text{ is even for every } v\} \tag{6.7}$$

counts these sets: they form the binary cycle space, whose dimension is $\beta(H)$. In Section 9 we will restrict an even set to the edges outside the exposed large-cell matching. That restriction is injective, allowing us to bound the cycle-space and local-cell contributions together.

Joint multiplicity bounds for overlap cells. The overlap cells are dependent, so separate one-cell probability bounds are not enough. We instead count stub pairs that meet all the prescribed cell demands at once. The next estimate bounds that joint event; it is used only in Section 9.

Lemma 6.2 (Joint prescribed-cell bound)**.** *Let $\mathcal{I}$ and $\mathcal{J}$ be finite row- and column-index sets, and let $(d_a)_{a\in\mathcal{I}}$ and $(d'_b)_{b\in\mathcal{J}}$ be nonnegative integer degree lists satisfying*

$$\sum_{a\in\mathcal{I}} d_a = \sum_{b\in\mathcal{J}} d'_b = m_0.$$

Choose a uniform perfect matching of the two m_0-stub sets, write $\mathbb{P}_{\text{match}}$ for its law, and let r_{ab} be the number of matched pairs in cell (a,b). For a finite set $D \subseteq \mathcal{I}\times\mathcal{J}$, prescribe integer demands $x_{ab} \geq 1$ on D. Extend $x_{ab} = 0$ off D, and put

$$x = \sum_D x_{ab}, \quad D_a = \sum_b x_{ab}, \quad D'_b = \sum_a x_{ab}.$$

In the feasible case

$$x \leq m_0, \qquad D_a \leq d_a \ \ (\forall a), \qquad D'_b \leq d'_b \ \ (\forall b),$$

we have

$$\mathbb{P}_{\text{match}}(r_{ab} \geq x_{ab} \text{ for every } (a,b)\in D) \leq \frac{\prod_a (d_a)_{D_a} \prod_b (d'_b)_{D'_b}}{(m_0)_x \prod_{(a,b)\in D} x_{ab}!}. \tag{6.8}$$

If any feasibility condition fails, the prescribed event is empty and is handled separately. In particular, in all cases with $m_0 > 0$, with $\mathrm{e} := \exp(1)$ and $\theta_{ab} = \mathrm{e}\, d_a d'_b/m_0$, its probability is at most

$$\prod_{(a,b)\in D} \frac{\theta_{ab}^{x_{ab}}}{x_{ab}!}. \tag{6.9}$$

Proof. If a feasibility condition fails, no matching realizes all demands. In the feasible case, assigning disjoint row stubs to the demanded cells gives

$$\prod_a \frac{(d_a)_{D_a}}{\prod_b x_{ab}!}$$

choices. Independently selecting the column stubs gives

$$\prod_b \frac{(d'_b)_{D'_b}}{\prod_a x_{ab}!}$$

choices. For each cell ab, the selected row and column stubs can be paired in $x_{ab}!$ ways. Hence the product over all cells cancels one of the two copies of $\prod_{(a,b)\in D} x_{ab}!$, and the total number of choices of prescribed stub pairs is

$$\frac{\prod_a (d_a)_{D_a} \prod_b (d'_b)_{D'_b}}{\prod_{(a,b)\in D} x_{ab}!}.$$

A fixed set of stub pairs occurs with probability $1/(m_0)_x$: after its x pairs are exposed, the remaining matching is uniform on the unused stubs. Summing over the possible prescribed pair sets gives (6.8) by a union bound.

It remains to simplify the global denominator. For $1 \leq x \leq m_0$,

$$\binom{m_0}{x} = \prod_{j=0}^{x-1} \frac{m_0 - j}{x - j} \geq \left(\frac{m_0}{x}\right)^x, \qquad x! \geq \left(\frac{x}{\mathrm{e}}\right)^x.$$

The factorial inequality follows, for example, from $\sum_{j=1}^{x} \log j \geq \int_1^x \log t\, dt = x \log x - x + 1$. Consequently

$$(m_0)_x = x! \binom{m_0}{x} \geq (m_0/\mathrm{e})^x, \tag{6.10}$$

while $(m_0)_0 = 1 = (m_0/\mathrm{e})^0$ covers $x = 0$. Combining (6.10) with $(d)_r \leq d^r$ proves (6.9). Equation (6.8) retains the single global falling factorial before the coarser bound (6.9) is taken. □

7. Weights of common whole classes

Fix the profile of Proposition 5.5. A specified collection of whole classes common to both partitions is a *partial diagonal*. In each selected pair, the row and column classes have the same vertex set. We call such a choice a *selection*.

For each vector of selected class counts, let $D(\ell)$ be the sum of the weights over its selections. Section 8 uses these common-class weights as the reference for large overlap cells. We prove that their sum over all subprofiles is $1 + o(1)$ uniformly in the phase.

For a selected common subprofile $\ell = (\ell_i)$, with $i \in \{2, 3, 4, 5, 6\}$, put

$$\ell_\bullet = \sum_i \ell_i, \qquad m = \sum_i u_i \ell_i, \qquad b_\ell = \sum_i \binom{u_i}{2} \ell_i. \tag{7.1}$$

Here $0 \leq \ell_i \leq k_i$ coordinatewise. Since the full profile has $\sum_i u_i k_i = n$, this automatically gives $m \leq n$; thus every factorial in the counting formula below has its ordinary combinatorial meaning.

Marked weights. Summed over all choices of an m-vertex support in $[n]$, the first moment of the marked count of these partial partitions is

$$Y_\ell^{\mathrm{sgn}} = 2^{\ell_\bullet} \frac{n!}{(n-m)! \prod_i (u_i!)^{\ell_i} \ell_i!} 2^{-b_\ell}. \tag{7.2}$$

For each type i, choose ℓ_i row slots, ℓ_i column slots and a bijection between them, in $\binom{k_i}{\ell_i}^2 \ell_i! = (k_i)_{\ell_i}^2 / \ell_i!$ ways. For two independent uniform ordered partitions, the matched slots have identical vertex sets with probability $\prod_i (u_i!)^{\ell_i} / (n)_m$. Each such common class is an isolated edge of $H(r)$, so by (6.4) it carries the factor $g(u_i) = 2^{\binom{u_i}{2} - 1}$. The product, summed over the slot choices, is the exact *marked* common-subprofile weight

$$D(\ell) = \prod_i \frac{(k_i)_{\ell_i}^2 \, (u_i!\, g(u_i))^{\ell_i}}{\ell_i!} \cdot \frac{1}{(n)_m} = \frac{\prod_i \binom{k_i}{\ell_i}^2}{Y_\ell^{\mathrm{sgn}}}. \tag{7.3}$$

Equivalently, $D(\ell)$ is the expected total weight, for two independent uniform ordered partitions, of the selections of ℓ_i row classes and ℓ_i column classes of each type that are paired with identical vertex sets, where a selection has weight $\prod_i g(u_i)^{\ell_i}$, the product of the local factors of its common classes.

The weight $D(\ell)$ accounts only for the selection of whole common classes and their local signed factors. The other cell factors and the cycle-space factor remain to be estimated in Sections 8–9.

Different selections need not be disjoint: an overlap containing several common blocks contributes to each corresponding selection. The sums below are sums of these nonnegative weights, so no disjointness is used. In particular, $D(0) = 1$ and $D(\mathbf{k}) = 1/\mathbb{E}\big[Z_{\mathbf{k}}^{\mathrm{sgn}}\big]$. Section D.1 derives two exact recurrences for these weights: a forward one, which adds one common class to a selection, and a reverse one, which starts from the complete profile and removes classes.

In Heckel's admissible-phase argument, the corresponding control of partial profiles enters through the tame-coloring estimates of Heckel and Panagiotou (2023, Lemmas 5.1–5.3, 6.3–6.5, and 7.20), as used in Heckel (2025, Proposition 5); their hypotheses are the source of the lower μ_α-window in that theorem. The next lemma supplies the uniform five-size estimate needed here, including both corners and every intermediate mass.

Lemma 7.1 (All common subprofiles)**.** *For the selected integer profile, uniformly over the full phase interval,*

$$\sum_{0 \leq \ell_i \leq k_i} D(\ell) = 1 + o(1), \tag{7.7}$$

where the error is bounded by one deterministic sequence tending to zero.

Proof outline. The empty selection contributes one. We bound all other terms in three ranges: the forward recurrence (D.1) handles small selected mass, a negative leading rate handles intermediate mass, and the reverse recurrence (D.3) handles a nearly full selected profile. In the last range, the large first moment of the complete profile supplies the decay.

The main point in the intermediate range is that the leading exponential rate is negative. Put $K = k_{\mathrm{co}}$, $p_i = k_i/K$, $y_i = \ell_i/K$, and $z_i = p_i - y_i$. Define

$$Y = \sum_i y_i, \qquad R = \sum_i z_i = 1 - Y, \qquad T = \sum_i i p_i, \qquad I_r = \sum_i i z_i.$$

Thus Y is the selected class fraction and R is the residual class fraction. The exact weight formula and the factorial estimates in Section D.3 give, when the residual vertex fraction $\rho = (n-m)/n$ exceeds $1/32$,

$$\log D(\ell) \le K\alpha\Phi_T + C_{\mathrm{diag}} KY \log(\mathrm{e}/Y) + C_{\mathrm{diag}} \log n, \qquad \Phi_T = R\log R + \frac{\log 2}{2}(I_r - TR),$$

with a phase-independent constant C_{diag}. All deficits lie in $\{2,3,4,5,6\}$, so

$$I_r - TR \le (6-T)R, \qquad I_r - TR \le (T-2)(1-R).$$

Multiplying by $1-R$ and R, respectively, and adding gives $I_r - TR \le 4R(1-R)$. Also $\rho > 1/32$ implies $R \ge 1/64$ eventually, and Proposition 5.5 gives $T \le 39/10$. Lemma D.1 therefore yields $\Phi_T \le -Y/10000$ throughout this range.

Put $\eta := \log\log n/(32\log n)$. In the intermediate range $m > \eta n$, $n - m > n/32$, one has $Y \ge \eta/2$ eventually, so the two error terms are $o(K\alpha Y)$ uniformly in the phase (see (D.20)). Hence $\log D(\ell) \le -K\alpha\eta/40000$ eventually, and since there are at most $(K+1)^5$ subprofiles and $K\alpha\eta/\log n \to \infty$, their total contribution tends to zero. The two recurrences cover the other ranges.

Section D.3 gives the three ranges deterministic, phase-independent error sequences $\varepsilon_n^{\mathrm{empty}}$, $\varepsilon_n^{\mathrm{central}}$, $\varepsilon_n^{\mathrm{full}} \to 0$ and proves the quantitative form of (7.7) used in Section 9:

$$1 \le \sum_{0 \le \ell_i \le k_i} D(\ell) \le e^{\varepsilon_n^{\mathrm{empty}}} + \varepsilon_n^{\mathrm{central}} + \varepsilon_n^{\mathrm{full}}. \tag{7.7a}$$

8. Large overlap cells and endpoint comparisons

A large overlap cell need not contain a whole class. Our reference is full containment, where every stub of the smaller endpoint class is matched inside the cell. For a fixed set of large cells, we first sum over their labeled stub realizations. The remaining costs come from missing stubs in these cells and unequal endpoint class sizes. We sum over the missing-stub deficits, then group the full-containment weights by endpoint type to control the unequal sizes. The outcome is Proposition 8.6: summed over all large-cell configurations, these weights are at most $\exp\{O(n^{3/4}\log n)\}$, well inside the budget $o(n/(\log n)^4)$ for Λ_n from Section 5.

The common-class weights from Section 7 have total mass $1 + o(1)$. Here we sum the weights of cells larger than half the class-size cap; Section 9 bounds the remaining local factors and the cycle-space factor.

8.1. Large cells and their labeled realizations.

Let $U := \alpha - 2$ be the largest class size in the fixed five-size profile and put

$$R_0 := \left\lfloor \frac{U}{2} \right\rfloor.$$

A cell is *large* if its multiplicity exceeds R_0. Two large cells cannot share a row or a column: otherwise that row or column would contain more than U stubs, because $2(R_0+1) > U$. Hence all large cells form a canonical bipartite matching, denoted by M.

The matching M records which pairs of class slots form large cells. A labeled stub realization specifies the individual stub pairs within those cells. The endpoints of M are disjoint, so choosing pairs in one cell of M uses none of the stubs available to another. The realization counts therefore multiply.

For $e \in M$, let s_e and t_e be its endpoint block sizes and set

$$m_e := \min\{s_e, t_e\}, \qquad d_e := |s_e - t_e|, \qquad h_e := m_e - j_e,$$

where j_e is the realized multiplicity. Thus m_e is the full-containment multiplicity, h_e is the deficit, and $j_e = m_e - h_e$. Since $j_e > U/2$ and $m_e \le U$,

$$2h_e < m_e. \tag{8.1}$$

Put $J := \sum_{e\in M} j_e$.

For fixed (M, j), let $\mathfrak{F}(M, j)$ be the finite set of labeled partial stub matchings containing exactly J pairs, all in cells of M, with exactly j_e pairs in each cell $e \in M$. Thus $\mathfrak{F}(M, j)$ records the labeled pairs,

not just their cell counts. Let Π be the uniform perfect matching of the two n-stub sets from Section 6; a fixed set of J stub pairs lies in Π with probability $1/(n)_J$. For $0 \le j_e \le \min\{s_e, t_e\}$ put

$$w(M, j) := \sum_{S \in \mathfrak{F}(M,j)} \mathbb{P}(S \subseteq \Pi) \prod_{e \in M} g(j_e).$$

A size-j partial matching between labeled blocks of sizes s and t can be chosen in

$$\binom{s}{j}\binom{t}{j} j! = \frac{(s)_j (t)_j}{j!}$$

ways: choose its two endpoint sets and then the bijection between them.

Proposition 8.1 (Aggregate weight of a fixed cell matching)**.** *Let M be a fixed matching support arising from two copies of the fixed feasible n-vertex profile. For a multiplicity vector $j = (j_e)_{e \in M}$ satisfying $0 \le j_e \le \min\{s_e, t_e\}$ for every $e \in M$, one has $J \le n$, and the total configuration probability multiplied by the exposed local factor, summed over $\mathfrak{F}(M, j)$, is*

$$w(M, j) = \frac{\prod_{e \in M} (s_e)_{j_e} (t_e)_{j_e}}{(n)_J \prod_{e \in M} j_e!} \prod_{e \in M} g(j_e). \tag{8.2}$$

Proof. Because M is a matching, distinct selected cells use disjoint row and column stub sets. Therefore

$$|\mathfrak{F}(M, j)| = \prod_{e \in M} \frac{(s_e)_{j_e} (t_e)_{j_e}}{j_e!}.$$

A prescribed set of J stub pairs occurs with probability $(n)_J^{-1}$ in the uniform bipartite configuration matching. Multiplying this probability by the displayed cardinality and by the exposed local factor $\prod_{e \in M} g(j_e)$ gives (8.2). The sum is over prescribed partial matchings and can count one full matching more than once. Section 9 enforces exact cell multiplicities through the residual-completion condition. □

A pair (M, j) is *feasible* when M is a matching between classes and

$$R_0 < j_e \le m_e \qquad (e \in M).$$

For a feasible pair (M, j), write

$$w_{\mathrm{hi}}(M, j) := w(M, j).$$

The same exposed weight appears in the exact conditional decomposition of Section 9.

Remark 8.2. The matching hypothesis is essential. If two selected cells shared a row or a column, their endpoint choices would compete for the same stubs and the one-cell realization counts would not multiply.

Figure 3 shows how we split the overlap table. To recover the full labeled matching, we also retain the exposed stub pairs, as in Proposition 9.1. The last line separates the exposed weight w_{hi}, defined above, from the residual factor $\mathcal{A}$, which accounts for the unexposed matching and is defined in Section 9. We bound these two factors separately.

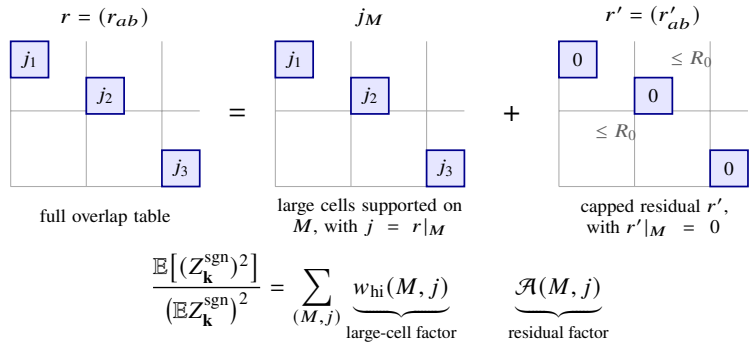


FIGURE 3. Canonical large-cell extraction. The highlighted cells $M = \{(a, b) : r_{ab} > R_0\}$ form a matching, with multiplicities $j = r|_M$. Extending j by zero gives $r = j_M + r'$, where r' vanishes on M and is capped by R_0 elsewhere. The displayed identity is Proposition 9.1.

Proposition 9.1 lifts this table decomposition to labeled configuration matchings by retaining the selected-pair set $S \in \mathfrak{F}(M, j)$ and the residual matching Π_{res}. This gives the normalized factorization shown in the figure. Blank cells have no displayed multiplicity. Outside M, the original cell equals its residual cell in r', is bounded by R_0, and has zero entry in j_M.

8.2. **Summing the deficits.** We first fix the matching M and sum over the multiplicities of its cells, comparing each large cell with full containment. Write

$$\mathbf{m} := (m_e)_{e\in M}, \qquad w_{\mathrm{full}}(M) := w(M, (m_e)_{e\in M}), \qquad J_* := \sum_{e\in M} m_e, \qquad h_\bullet := J_* - J = \sum_{e\in M} h_e.$$

We label the five selected class sizes by *types* $i \in \{0, 1, 2, 3, 4\}$: type i has size $u_i := \alpha - 2 - i$ and $\kappa_i := k_{i+2}$ classes. Through Section 9, u_i is this shifted size. Assume $\alpha > 8$; since $\alpha > \alpha_0 - 1$, this holds beyond a phase-independent threshold. For $i, j \in \{0, 1, 2, 3, 4\}$ let $m_{ij} := \min\{u_i, u_j\}$ and put

$$\rho_{ij} := \frac{nm_{ij}}{2^{\lfloor (3m_{ij}-1)/4\rfloor}}, \qquad \rho_\Sigma := \sum_{i,j=0}^{4} \rho_{ij}.$$

Proposition 8.3 (Fixed-support all-deficit bound)**.** *Write* $\mathbf{m} = (m_e)_{e\in M}$*. The sum below runs over vectors* $h = (h_e)_{e\in M}$ *of nonnegative integers satisfying* $m_e - h_e > R_0$ *for every* $e \in M$*. If* $\rho_\Sigma \le 1$*, then*

$$\sum_h w(M, \mathbf{m} - h) \le w_{\mathrm{full}}(M)\,(1 + (\alpha + 1)\rho_\Sigma)^{|M|}. \tag{8.3}$$

Section E.2 proves the proposition. Lowering the multiplicity of a large cell with endpoint sizes $m \le m + d$ from m to $m - h$ multiplies its realization count and local factor by at most $m^h 2^{-hm+h(h+1)/2}$, while the single global falling factorial changes by at most a factor n^h. Since $2h < m$, one has $hm - h(h + 1)/2 \ge h\lfloor(3m - 1)/4\rfloor$, so each missing stub costs at most a factor $nm2^{-\lfloor(3m-1)/4\rfloor} \le \rho_\Sigma$, and $\rho_\Sigma = O(n^{-1/4}\log n)$ by (8.14). Summing the at most $\alpha + 1$ deficits of each cell independently gives the proposition.

Define $K := \sum_{i=0}^{4} \kappa_i = k_{\mathrm{co}}$, the number of row-class slots. Projection onto the row endpoint is injective on a block matching, and therefore

$$|M| \le K. \tag{8.4}$$

8.3. **Regrouping by endpoint table.** After summing the deficits, we must sum the full-containment reference weights. These depend on how many matched class pairs have each pair of endpoint types. Record those counts in an endpoint table and group the matching supports by that table. We can then compare its row and column margins with the common-class weights from Section 7.

For a block matching M, let

$$\ell_{ij}(M) := \#\{e \in M : e \text{ joins row type } i \text{ to column type } j\}, \qquad L(M) := (\ell_{ij}(M))_{0\le i,j\le 4}.$$

With this type indexing, for $L = (\ell_{ij}) \in \mathbb{N}_0^{5\times 5}$ write

$$r = (r_i)_{i=0}^4, \qquad r_i = \sum_j \ell_{ij}, \qquad c = (c_j)_{j=0}^4, \qquad c_j = \sum_i \ell_{ij}.$$

Call L *feasible* if $r_i \le \kappa_i$ and $c_j \le \kappa_j$ for every i, j. Such a table is realized by choosing the indicated row and column block slots and pairing them cell by cell; conversely, every block matching has a feasible table. For a feasible endpoint table L, define

$$W(L) := \sum_{M:\, L(M)=L} w_{\mathrm{full}}(M).$$

Partitioning the finite set of block matchings by $L(M)$ gives

$$\sum_M w_{\mathrm{full}}(M) = \sum_{L \text{ feasible}} W(L). \tag{8.5}$$

Whenever $\rho_\Sigma \le 1$, combining Proposition 8.3, (8.4), and (8.5) yields, after enlarging the nonnegative sum from canonical large-cell matchings to all block matchings,

$$\sum_{(M,j) \text{ feasible}} w_{\mathrm{hi}} \le \left(\sum_{L \text{ feasible}} W(L)\right)(1 + (\alpha + 1)\rho_\Sigma)^K. \tag{8.6}$$

The finite argument through (8.6) applies to any fixed set of endpoint types for which the selected cells form a matching. The five consecutive sizes enter through the numerical deficit bound and the endpoint-table comparison below.

8.4. **Comparison with the common-class weights.** When the endpoint sizes agree, full containment means that a whole class is common to both partitions, the case estimated in Section 7. For a general endpoint table, we compare its weight with the common-class weights of the row and column margins, including the corresponding multinomial factors. The reference value is their geometric mean; unequal endpoint sizes introduce a small correction.

For a full-containment table $L = (\ell_{ij})$, set

$$r = (r_i)_{i=0}^4, \qquad r_i = \sum_j \ell_{ij}, \quad c = (c_j)_{j=0}^4, \quad c_j = \sum_i \ell_{ij},$$

$$J(L) = \sum_{i,j} m_{ij}\ell_{ij}.$$

The number of block matchings with table L is

$$\frac{\prod_i (\kappa_i)_{r_i} \prod_j (\kappa_j)_{c_j}}{\prod_{i,j} \ell_{ij}!}.$$

Indeed, row and column slot allocations initially produce two copies of the cell-factorial denominator, while the bijections between the selected slots inside each type cell contribute one copy back. Thus the exact reference weight is

$$W(L) = \frac{\prod_i (\kappa_i)_{r_i} \prod_j (\kappa_j)_{c_j}}{\prod_{i,j} \ell_{ij}!} \frac{1}{(n)_{J(L)}} \prod_{i,j} \left[\frac{(u_i)_{m_{ij}} (u_j)_{m_{ij}}}{m_{ij}!} g(m_{ij}) \right]^{\ell_{ij}}. \tag{8.7}$$

Let

$$\mathcal{V} := \{v = (v_i)_{i=0}^4 \in \mathbb{N}_0^5 : 0 \le v_i \le \kappa_i \text{ for every } i\}.$$

For $v \in \mathcal{V}$, define

$$D(v) = \frac{\prod_i (\kappa_i)_{v_i}^2}{\prod_i v_i!} \frac{\prod_i [u_i! g(u_i)]^{v_i}}{(n)_{m(v)}}, \qquad m(v) = \sum_i u_i v_i. \tag{8.8}$$

This is the common-subprofile weight (7.3) with $\ell_{i+2} = v_i$: since $\kappa_i = k_{i+2}$ and $u_i = \alpha - (i+2)$, (8.8) is the first expression in (7.3). Below, $D(r)$ and $D(c)$ evaluate (8.8) on the margins in $0 \le v_i \le \kappa_i$.

For the endpoint pair (i, j) put $d_{ij} := |u_i - u_j|$, so its endpoint sizes are m_{ij} and $m_{ij} + d_{ij}$ as in Lemma E.1, and define

$$Q_{ij} := (n+1)^{d_{ij}/2} \frac{\sqrt{(m_{ij} + d_{ij})_{d_{ij}}}}{d_{ij}!} 2^{-\{d_{ij} m_{ij} + \binom{d_{ij}}{2}\}/2}.$$

Thus $Q_{ii} = 1$. Since $m_{ij} + d_{ij} \le U$, $m_{ij} \ge U - 4$ and $d_{ij} \le 4$,

$$Q_{ij} \le \frac{(\tau_n^{\text{end}})^{d_{ij}}}{d_{ij}!}, \qquad \tau_n^{\text{end}} = 4\sqrt{\frac{(n+1)U}{2^U}} = O\left(\frac{(\log n)^{3/2}}{\sqrt{n}}\right). \tag{8.9}$$

Let

$$A_L = \frac{\prod_i r_i!}{\prod_{i,j} \ell_{ij}!}, \qquad C_L = \frac{\prod_j c_j!}{\prod_{i,j} \ell_{ij}!}, \qquad Q^L = \prod_{i,j} Q_{ij}^{\ell_{ij}}.$$

Lemma 8.4 (Endpoint-table product comparison)**.** *Assume $\alpha > 8$ and that $(\kappa_i, u_i)_{i=0}^4$ is the fixed feasible profile, so $\sum_{i=0}^4 \kappa_i u_i = n$. Then every feasible endpoint table L satisfies*

$$W(L)^2 \le \big(D(r) A_L Q^L\big)\big(D(c) C_L Q^L\big). \tag{8.10}$$

Section E.3 proves the lemma. After division by $(D(r)A_L)(D(c)C_L)$, the profile multiplicities and table factorials cancel, each cell of endpoint type (i, j) leaves the factor $Q_{ij}^2/(n+1)^{d_{ij}}$, and the remaining ratio of falling factorials is at most $(n+1)^{\sum_{i,j} d_{ij}\ell_{ij}}$.

Proposition 8.5 (Endpoint-table sum)**.** *There is an absolute constant $C_{\text{end}} > 0$ such that, for all sufficiently large n, uniformly in the phase,*

$$\sum_{L \text{ feasible}} W(L) \le (1 + 4\tau_n^{\text{end}})^K \sum_{v \in \mathcal{V}} D(v) \le \exp\left\{C_{\text{end}} \sqrt{n \log n}\right\}. \tag{8.11}$$

All sums in this proposition are finite.

Proof. For all sufficiently large n, one has $\tau_n^{\mathrm{end}} \le 1$ by (8.9); hence $Q_{ij} \le \tau_n^{\mathrm{end}}$ whenever $i \ne j$. By the arithmetic–geometric mean inequality and (8.10),

$$2W(L) \le D(r)A_L Q^L + D(c)C_L Q^L.$$

Sum the row term first, keeping its margin r fixed. We can drop the column constraints and sum over all nonnegative integer tables with row margin r, since $A_L Q^L \ge 0$. The multinomial theorem then gives

$$\sum_{\substack{L \text{ feasible} \\ \mathrm{row}(L)=r}} A_L Q^L \le \sum_{\substack{L \in \mathbb{N}_0^{5\times 5} \\ \mathrm{row}(L)=r}} A_L Q^L = \prod_i \left(\sum_j Q_{ij} \right)^{r_i} \le (1 + 4\tau_n^{\mathrm{end}})^K.$$

Multiply by $D(r)$ and sum over the row margin r. The result is at most $(1 + 4\tau_n^{\mathrm{end}})^K \sum_{v \in \mathcal{V}} D(v)$, with v a dummy subprofile variable. Exchanging rows and columns gives the same bound for the second term. Thus the endpoint-table choices cost only the displayed correction factor. The remaining sum over margins is the common-subprofile sum of Section 7: $\sum_{v \in \mathcal{V}} D(v) = 1 + o(1)$ uniformly in the phase, with all common-class overlaps included. Finally, $K = \Theta(n/\log n)$ and (8.9) imply $\tau_n^{\mathrm{end}} K = O(\sqrt{n \log n})$. □

8.5. **The large-cell bound across the phase.** By (2.1) and $\alpha_0 - 1 < \alpha \le \alpha_0$, the integer center for the independent-set size satisfies, uniformly in the phase,

$$\alpha = \left(\frac{2}{\log 2} + o(1) \right) \log n.$$

In particular, for all sufficiently large n,

$$\frac{5}{2} \log n \le \alpha. \tag{8.12}$$

Every endpoint minimum satisfies $m_{ij} \ge \alpha - 6$, and, since $4\lfloor x/4 \rfloor \ge x - 3$ for integers x and $3m_{ij} - 1 \ge 3\alpha - 19$,

$$3\alpha - 22 \le 4 \left\lfloor \frac{3m_{ij} - 1}{4} \right\rfloor.$$

Using $\log 2 > 2/3$, $3\alpha - 22 > 0$ and (8.12), we obtain

$$\frac{5}{4} \log n - \frac{11}{3} \le (\log 2) \left\lfloor \frac{3m_{ij} - 1}{4} \right\rfloor. \tag{8.13}$$

Hence

$$2^{\lfloor (3m_{ij}-1)/4 \rfloor} \ge e^{-11/3} n^{5/4}.$$

Since every $m_{ij} = O(\log n)$,

$$\rho_\Sigma = O\left(\frac{\log n}{n^{1/4}} \right) \longrightarrow 0. \tag{8.14}$$

Thus $\rho_\Sigma \le 1$ for all sufficiently large n.

Let Σ_n^{hi} denote the sum of exposed weights:

$$\Sigma_n^{\mathrm{hi}} := \sum_{(M,j) \text{ feasible}} w_{\mathrm{hi}}(M, j),$$

where the sum is over the feasible pairs (M, j) defined above.

Proposition 8.6 (Bound for the exposed weights)**.** *There is a phase-independent deterministic sequence $\varepsilon_n^{\mathrm{hi}} \ge 0$ with $\varepsilon_n^{\mathrm{hi}} \to 0$ such that, for all sufficiently large n,*

$$\Sigma_n^{\mathrm{hi}} \le \exp\left\{ \varepsilon_n^{\mathrm{hi}} \frac{n}{(\log n)^4} \right\}. \tag{8.15}$$

Proof. The deficit factor in (8.6) satisfies

$$\log\left(1 + (\alpha + 1)\rho_\Sigma\right)^K \le K(\alpha + 1)\rho_\Sigma = O(n^{3/4} \log n).$$

Proposition 8.5 contributes $O(\sqrt{n \log n})$ to the logarithm. Both quantities are $o(n/(\log n)^4)$. Combining these estimates with (8.6) proves the proposition. □

9. Residual factors and the second-moment bound

We keep the selected profile of Proposition 5.5 and the endpoint indexing from Section 8. The exposed matching M has pairs of class slots as its edges. The labeled stub pairs chosen inside those cells are denoted by S.

Fix a feasible pair (M, j) and condition on a labeled realization of its exposed stub pairs being present. The unused row and column stubs are then matched uniformly. An allowed completion adds no pair in an exposed cell and respects the residual cap elsewhere. We estimate the local and cycle-space factors jointly, since they depend on this same residual matching, and include the indicator that the completion is allowed. A bound uniform over all feasible pairs (M, j) can then be taken outside the exposed-weight sum from Section 8.

Throughout this section, write $Z := Z_{\mathbf{k}}^{\mathrm{sgn}}$ for the selected marked cocoloring count.

9.1. **Exact decomposition after exposing the large cells.** Fix a feasible pair (M, j) and recall $J = \sum_{e\in M} j_e$. Put

$$m_0 := n - J.$$

Let $(d_a)_a$ and $(d'_b)_b$ be the residual row and column degrees. Their two sums equal m_0, and every degree is at most the phase cap U. Explicitly,

$$d_a = s_a - \sum_{b:(a,b)\in M} j_{ab}, \qquad d'_b = t_b - \sum_{a:(a,b)\in M} j_{ab}.$$

The unexposed pairs form a uniform bipartite configuration matching with these residual degrees; when $m_0 = 0$, this means the unique empty matching. Write $\mathbb{P}_{\mathrm{res}}$ and $\mathbb{E}_{\mathrm{res}}$ for its law and expectation. Let r'_{ab} be its cell counts and let H_{res} be the simple support graph of cells with $r'_{ab} \geq 2$. The event $\mathcal{E}(M, j)$ specifies which residual matchings complete the chosen large cells. Outside the exposed matching it imposes the cap $r'_{ab} \leq \lfloor U/2 \rfloor$; in a cell of M, no further pair is allowed. The exposed data already include every pair in each large cell, so an additional pair would change its multiplicity. Thus

$$\mathcal{E}(M, j) = \{r'_{ab} = 0 \ ((a, b) \in M), \quad r'_{ab} \leq R_0 \ ((a, b) \notin M)\}.$$

More explicitly, with Π as in Section 8.1, for every $S \in \mathfrak{F}(M, j)$ and every event $\mathcal{B}$ of the residual matching,

$$\mathbb{P}\{S \subseteq \Pi, \ \Pi \setminus S \in \mathcal{B}\} = \frac{1}{(n)_J}\, \mathbb{P}_{\mathrm{res}}(\mathcal{B}). \tag{9.1}$$

Different labeled realizations leave the same residual degree lists. A blockwise bijection between the unused row stubs, together with one between the unused column stubs, identifies their residual matching spaces and preserves every cell count. The residual expectation therefore depends on M and the degree lists alone. Define

$$\mathcal{A}(M, j) := \mathbb{E}_{\mathrm{res}}\left[\left(\prod_{a,b} g(r'_{ab})\right) 2^{\beta(M\cup H_{\mathrm{res}})} \mathbf{1}_{\mathcal{E}(M,j)}\right]. \tag{9.2}$$

Only the residual configuration matching is random in this expectation.

Proposition 9.1 (Canonical exact overlap decomposition)**.** *Let n be finite and let the fixed selected profile be feasible, with phase cap $U \geq 2$. Then*

$$\frac{\mathbb{E}Z^2}{(\mathbb{E}Z)^2} = \sum_{(M,j)} w_{\mathrm{hi}}(M, j)\, \mathcal{A}(M, j), \tag{9.3}$$

where the sum ranges over the finite family of feasible pairs (M, j).

Section F.1 proves the proposition by a bijection. A full matching Π determines its large cells $M(\Pi) = \{(a, b) : r_{ab}(\Pi) > R_0\}$, which form a matching because $2(R_0 + 1) > U$; their multiplicities $j(\Pi)$; the set $S(\Pi) \in \mathfrak{F}(M, j)$ of its stub pairs in those cells; and the residual matching $\Pi \setminus S(\Pi)$, which satisfies $\mathcal{E}(M, j)$. Conversely, every feasible (M, j), every $S \in \mathfrak{F}(M, j)$ and every residual matching in $\mathcal{E}(M, j)$ glue to exactly one full matching. For fixed (M, j, S), (9.1) makes the total contribution $(n)_J^{-1} \prod_{e\in M} g(j_e)\, \mathcal{A}(M, j)$, and summing over S gives $w_{\mathrm{hi}}(M, j)\mathcal{A}(M, j)$.

The case $m_0 = 0$ is complete without introducing cell activities. The residual matching is empty, every residual local factor is one, and $H_{\mathrm{res}} = \varnothing$. Since M is a matching, $\beta(M) = 0$, so

$$\mathcal{A}(M, j) = 1 \qquad (m_0 = 0). \tag{9.4}$$

In all subsequent subsections we assume $m_0 > 0$. In view of Proposition 8.6, it remains to bound $\mathcal{A}(M, j)$ uniformly in the positive-residual case.

9.2. **Threshold expansion and cell activities.** We expand the integrand of (9.2) into products of threshold indicators, bound each product by the joint estimate of Lemma 6.2, and collect the per-cell contributions into two activities $\lambda_{ab} \le q_{ab}$.

For every row–column pair (a, b), put

$$\theta_{ab} := \frac{\mathrm{e}\, d_a d'_b}{m_0}. \tag{9.5}$$

This intensity is defined for cells of M as well as for cells outside M. Both activities are set to zero on M. Let

$$\Delta_x := g(x) - g(x-1),$$

so $\Delta_x \ge 0$: indeed $\Delta_3 = 3$, while $g(x)/g(x-1) = 2^{x-1} > 1$ for $x \ge 4$. For $(a, b) \notin M$, define

$$\lambda_{ab} := \sum_{x=3}^{R_0} \Delta_x \frac{\theta_{ab}^x}{x!}, \qquad q_{ab} := \frac{\theta_{ab}^2}{2} + \lambda_{ab}. \tag{9.6}$$

On M, set $\lambda_{ab} = q_{ab} = 0$. The first activity accounts for increments of the local factor above one. The second also includes the threshold-two requirement for an edge of the residual support.

For a finite array of nonnegative integer demands $x = (x_{ab})$ supported outside M, apply Lemma 6.2 to the residual degree lists and $\mathrm{supp}(x) := \{(a, b) : x_{ab} > 0\}$. It gives, in all cases,

$$\mathbb{P}_{\mathrm{res}}(r'_{ab} \ge x_{ab} \text{ for all } a, b) \le \prod_{(a,b)\in \mathrm{supp}(x)} \frac{\theta_{ab}^{x_{ab}}}{x_{ab}!}. \tag{9.7}$$

Because this estimate is joint, it applies simultaneously to cells sharing a row or a column; no independence between cells is asserted.

For $0 \le r \le R_0$,

$$g(r) = 1 + \sum_{x=3}^{R_0} \Delta_x \mathbf{1}_{\{r\ge x\}}.$$

Using the row- and column-slot sets defined in Section 6, put $E_0 := (\mathcal{I}_{\mathrm{row}} \times \mathcal{I}_{\mathrm{col}}) \setminus M$, and let $\mathfrak{E}(M)$ be the family of even subsets of $M \cup E_0$. For an edge $e = (a, b)$, write $r'_e := r'_{ab}$, $\theta_e := \theta_{ab}$, $\lambda_e := \lambda_{ab}$, and $q_e := q_{ab}$. For $F \in \mathfrak{E}(M)$, define

$$\Phi_F(r') := \prod_{e\in F\setminus M} \left(\mathbf{1}_{\{r'_e\ge 2\}} + \sum_{x=3}^{R_0} \Delta_x \mathbf{1}_{\{r'_e\ge x\}} \right) \times \prod_{e\in E_0\setminus F} \left(1 + \sum_{x=3}^{R_0} \Delta_x \mathbf{1}_{\{r'_e\ge x\}} \right).$$

On the event $\mathcal{E}(M, j)$, expansion of the cycle-space cardinality and of all local factors gives the exact identity

$$\left(\prod_{a,b} g(r'_{ab}) \right) 2^{\beta(M\cup H_{\mathrm{res}})} = \sum_{F\in\mathfrak{E}(M)} \Phi_F(r').$$

To verify it, note that evenness of an edge set does not depend on the ambient graph and that $E(H_{\mathrm{res}}) \subseteq E_0$ on $\mathcal{E}(M, j)$. By (6.7), $2^{\beta(M\cup H_{\mathrm{res}})}$ counts the $F \in \mathfrak{E}(M)$ with $r'_e \ge 2$ for all $e \in F \setminus M$:

$$2^{\beta(M\cup H_{\mathrm{res}})} = \sum_{F\in\mathfrak{E}(M)} \prod_{e\in F\setminus M} \mathbf{1}_{\{r'_e\ge 2\}}.$$

Multiply by $\prod_{a,b} g(r'_{ab})$, whose factors on M equal $g(0) = 1$; expand each factor on E_0 by the formula for $g(r)$ (valid as $r'_e \le R_0$), and use $\mathbf{1}_{\{r\ge 2\}} g(r) = \mathbf{1}_{\{r\ge 2\}} + \sum_{x=3}^{R_0} \Delta_x \mathbf{1}_{\{r\ge x\}}$.

This identity uses both parts of $\mathcal{E}(M, j)$: off-M cells are capped at R_0, and r' vanishes on M. We next multiply by $\mathbf{1}_{\mathcal{E}(M,j)}$ and bound the expectation term by term. On $\mathcal{E}(M, j)^c$ we use only this expansion, without identifying it with the uncapped local factor.

Lemma 9.2 (Fixed even-set expansion)**.** *For every $F \in \mathfrak{E}(M)$,*

$$\mathbb{E}_{\mathrm{res}}\left[\Phi_F(r')\mathbf{1}_{\mathcal{E}(M,j)}\right] \le \prod_{e\in F\setminus M} q_e \prod_{e\in E_0\setminus F} (1 + \lambda_e).$$

Proof. Expand Φ_F by choosing one summand from each factor. For a cell in $F \setminus M$, choose either the threshold-two term or one higher increment, never both. For a cell outside $F \cup M$, choose either the constant term or one higher increment. Apply (9.7) once to all the demands in each resulting product. The threshold-two term contributes $\theta_e^2/2$; the higher increments sum to λ_e; and the constant choice contributes one. Every product is nonnegative and $\mathbf{1}_{\mathcal{E}(M,j)} \le 1$, so dropping the indicator of $\mathcal{E}(M, j)$ can only increase its expectation. □

Summing over F and then inserting the missing factors $1 + \lambda_e \ge 1$ gives

$$\mathcal{A}(M, j) \le \left(\prod_{e \in E_0} (1 + \lambda_e)\right) \sum_{F \in \mathfrak{E}(M)} \prod_{e \in F \setminus M} q_e. \tag{9.8}$$

9.3. **Restricting even edge sets outside the exposed matching.** To bound the sum over even edge sets, we first discard their edges in the exposed matching, which have unit weight. The retained edges determine the entire even set. We can therefore bound the sum by allowing every subset of the edges outside the matching and expanding the resulting product.

Lemma 9.3 (Even-set restriction bound)**.** *Let H be a finite graph, let $M \subseteq E(H)$ be a matching, and let $q_e \ge 0$ for $e \in E(H) \setminus M$. Then*

$$\sum_{\substack{F \subseteq E(H) \\ F \text{ even}}} \prod_{e \in F \setminus M} q_e \le \prod_{e \in E(H) \setminus M} (1 + q_e).$$

Proof. Two even edge sets with the same restriction outside M have an even symmetric difference contained in M. A nonempty subset of a matching has a vertex of degree one, so the difference must be empty. Restriction outside M is therefore injective. Enlarge the resulting sum to all subsets of $E(H) \setminus M$ and expand the product. □

Apply the lemma to the complete bipartite graph on the row and column slots, with edge set $M \cup E_0$. It gives

$$\sum_{F \in \mathfrak{E}(M)} \prod_{e \in F \setminus M} q_e \le \prod_{e \in E_0} (1 + q_e). \tag{9.9}$$

Since $0 \le \lambda_e \le q_e$, equations (9.8) and (9.9), together with $1 + x \le e^x$, imply

$$\mathcal{A}(M, j) \le \exp\left(2 \sum_{e \in E_0} q_e\right). \tag{9.10}$$

Thus both the local factors and the cycle-space cardinality are controlled by one total residual activity.

9.4. **Large residual mass: a quadratic activity bound.** We next bound the sum of the residual activities. When many stubs remain unmatched, each cell intensity is small relative to the class-size cap. In this range, the contribution from higher multiplicities is bounded by a constant times the square of the intensity.

Assume

$$2^U \le m_0^3. \tag{9.11}$$

Then $m_0 \ge 2^{U/3}$ and, by the degree cap,

$$\theta_{ab} \le \mathrm{e} U^2 2^{-U/3}.$$

Lemma 9.4 (Quadratic activity bound)**.** *There is an absolute constant C_0 such that, under (9.11),*

$$q_{ab} \le C_0 \theta_{ab}^2 \tag{9.12}$$

for every residual cell outside M.

The quadratic activity estimate is proved in Section F.2; summing it over the residual degree lists gives the bound below.

Because θ_{ab} was defined for every pair, the cell intensities satisfy the exact identity

$$\sum_{a,b} \theta_{ab}^2 = \frac{\mathrm{e}^2}{m_0^2} \left(\sum_a d_a^2\right) \left(\sum_b (d'_b)^2\right).$$

Since every degree is at most U and both degree sums equal m_0,

$$\sum_a d_a^2 \le U m_0, \qquad \sum_b (d'_b)^2 \le U m_0.$$

On M the activities are zero, while outside M the preceding lemma applies. Therefore

$$\sum_{e\in E_0} q_e \le C_{\Sigma q} U^2 \qquad \text{with } C_{\Sigma q} := \mathrm{e}^2 C_0. \tag{9.13}$$

Together with (9.10), this proves

$$\mathcal{A}(M, j) \le \exp(2C_{\Sigma q} U^2) \qquad \text{if } 2^U \le m_0^3. \tag{9.14}$$

9.5. **Small residual mass: a pointwise bound.** When few stubs remain unmatched, the quadratic activity estimate may not apply. We instead bound the contribution of each residual matching directly. Suppose $2^U > m_0^3$. Then

$$m_0 < 2^{U/3} \le 2^{\lceil U/3\rceil}. \tag{9.15}$$

The pointwise argument uses only the canonical cap and the matching property, so it also applies outside this regime. Retaining the cap gives a sharper estimate. On $\mathcal{E}(M, j)$, write

$$H := H_{\mathrm{res}}, \qquad W' := \sum_{a,b} \binom{r'_{ab}}{2}.$$

Every support edge is a cell of multiplicity at least two, so the local factor, including $g(2) = 1$, is exactly

$$\prod_{a,b} g(r'_{ab}) = 2^{W' - |E(H)|}.$$

Restriction of even edge sets outside the matching M is injective by the argument following Lemma 9.3. No residual pair lies in an exposed cell, so $E(H) \cap M = \varnothing$; hence

$$2^{\beta(M\cup H)} \le 2^{|E(H)|}.$$

The edge-cardinality factors cancel. Since $U \ge 2$, the cap $R_0 = \lfloor U/2\rfloor$ is at least one, and for $0 \le r \le R_0$,

$$\binom{r}{2} \le \frac{R_0 - 1}{2} r.$$

Summing over cells, using that their residual multiplicities sum to m_0, gives

$$\left(\prod_{a,b} g(r'_{ab})\right) 2^{\beta(M\cup H)} \le 2^{W'} \le 2^{(R_0-1)m_0/2}.$$

The residual factor integrand vanishes outside $\mathcal{E}(M, j)$. Taking expectations proves

$$\mathcal{A}(M, j) \le 2^{(R_0-1)m_0/2} \qquad \text{for every } m_0 \ge 0. \tag{9.16}$$

We use this bound when the positive residual mass is small. It also covers zero residual mass, empty residual support, and an empty canonical event.

The following proposition collects the bounds for zero, large and small residual mass. The threshold depends on the class-size cap U, rather than on the number of class-size types.

Proposition 9.5 (Uniform residual factor)**.** *For the fixed feasible selected profile with $U \ge 2$, there is an absolute constant $C_{\mathrm{att}} > 0$ such that every feasible pair (M, j) satisfies*

$$\mathcal{A}(M, j) \le \begin{cases} 1, & m_0 = 0, \\ \exp(C_{\mathrm{att}} U^2), & m_0 > 0 \text{ and } 2^U \le m_0^3, \\ 2^{(R_0-1)m_0/2}, & m_0 > 0 \text{ and } 2^U > m_0^3. \end{cases}$$

Proof. The three cases are exhaustive. Their bounds are (9.4), (9.14), and (9.16), with $C_{\mathrm{att}} := 2C_{\Sigma q}$. □

9.6. **The global second-moment exponent.** Insert the bounds for exposed weights and residual factors into the exact second-moment decomposition. Their logarithms add. We keep track of the resulting exponent to verify that it is small enough for the amplification step.

Combine the three deterministic partial-diagonal errors introduced in Section 7 and estimated in Section D.3 by setting

$$\varepsilon_n^{\mathrm{pd}} := e^{\varepsilon_n^{\mathrm{empty}}} - 1 + \varepsilon_n^{\mathrm{central}} + \varepsilon_n^{\mathrm{full}}. \tag{9.17}$$

Then $\varepsilon_n^{\mathrm{pd}} \to 0$ uniformly in the phase, and the disjoint three-range assembly gives

$$\sum_{v\in\mathcal{V}} D(v) \le 1 + \varepsilon_n^{\mathrm{pd}}. \tag{9.18}$$

All sums in this subsection are finite and range over the domains defined in Sections 7–9.

With τ_n^{end} from (8.9), put

$$\begin{aligned}\Gamma_n^{\mathrm{skel}} &:= K\log(1+(\alpha+1)\rho_\Sigma)\\ &\quad + K\log\Big(1+4\tau_n^{\mathrm{end}}\Big)+\log\Big(1+\varepsilon_n^{\mathrm{pd}}\Big).\end{aligned} \tag{9.19}$$

Inserting (9.18) into the first inequality of (8.11), and the result into (8.6), gives, for all sufficiently large n,

$$\Sigma_n^{\mathrm{hi}} \le e^{\Gamma_n^{\mathrm{skel}}}. \tag{9.20}$$

As in the proofs of Propositions 8.5 and 8.6, the three terms of Γ_n^{skel} are $O(n^{3/4}\log n)$, $O(\sqrt{n\log n})$ and $o(1)$; so $\varepsilon_n^{\mathrm{skel}} := (\log n)^4\Gamma_n^{\mathrm{skel}}/n \to 0$.

Let C_{att} be the constant from Proposition 9.5. Define

$$\Gamma_n^{\mathrm{att}} := \max\left\{C_{\mathrm{att}}U^2,\ \frac{\log 2}{2}(R_0-1)2^{U/3}\right\}. \tag{9.21}$$

Proposition 9.5 implies the uniform bound

$$\mathcal{A}(M,j) \le e^{\Gamma_n^{\mathrm{att}}} \tag{9.22}$$

for every feasible pair (M,j). For large residual mass we use (9.14). For small positive residual mass, $m_0 < 2^{U/3}$ and

$$2^{(R_0-1)m_0/2} = \exp\left(\frac{\log 2}{2}(R_0-1)m_0\right) \le \exp\left(\frac{\log 2}{2}(R_0-1)2^{U/3}\right).$$

The zero-residual case contributes one.

By (2.1) and $\alpha_0 - 1 < \alpha \le \alpha_0$, uniformly in the phase,

$$U = O(\log n), \qquad 2^U = \Theta\left(\frac{n^2}{(\log n)^2}\right),$$

so

$$\Gamma_n^{\mathrm{att}} = O\Big((\log n)^2 + n^{2/3}(\log n)^{1/3}\Big) = o\left(\frac{n}{(\log n)^4}\right).$$

Thus the deterministic normalized residual factor error

$$\varepsilon_n^{\mathrm{att}} := \frac{(\log n)^4}{n}\Gamma_n^{\mathrm{att}} \longrightarrow 0. \tag{9.23}$$

Proposition 9.6 (Explicit normalized second-moment bound)**.** *For the marked cocoloring count $Z = Z_{\mathbf{k}}^{\mathrm{sgn}}$ from Proposition 5.5, define*

$$\Lambda_n := \Gamma_n^{\mathrm{skel}} + \Gamma_n^{\mathrm{att}} = \Big(\varepsilon_n^{\mathrm{skel}} + \varepsilon_n^{\mathrm{att}}\Big)\frac{n}{(\log n)^4}. \tag{9.24}$$

Then

$$\Lambda_n = O(n^{3/4}\log n) = o\left(\frac{n}{(\log n)^4}\right)$$

and for all sufficiently large n,

$$1 \le \frac{\mathbb{E}Z^2}{(\mathbb{E}Z)^2} \le e^{\Lambda_n}. \tag{9.25}$$

Proof. Insert (9.22) into the exact decomposition (9.3). Because all summands are nonnegative, the uniform residual factor may be taken outside the finite sum over (M,j). Equation (9.20) then gives

$$\frac{\mathbb{E}Z^2}{(\mathbb{E}Z)^2} \le e^{\Gamma_n^{\mathrm{att}}}\Sigma_n^{\mathrm{hi}} \le e^{\Gamma_n^{\mathrm{att}}+\Gamma_n^{\mathrm{skel}}} = e^{\Lambda_n}.$$

The lower bound follows from $\mathrm{Var}(Z) \ge 0$. The component estimates above give the stated $O(n^{3/4}\log n)$ bound for Λ_n, while the bound on $\varepsilon_n^{\mathrm{skel}}$ above and (9.23) give the small-o bound. □

The contribution Γ_n^{skel} comes from common classes, deficits and endpoint tables; Γ_n^{att} comes from residual local factors and the cycle-space factor. Applying Paley–Zygmund with their sum gives the probability bound used in the next section.

This section proves estimate (C) of the introduction: with high probability, $\zeta(G_n) \le k_{\mathrm{co}} + a_n$ for a deterministic sequence $a_n = o(n/(\log n)^3)$.

Proposition 9.6 and (1.4) give the probability bound

$$\mathbb{P}\big(Z_{\mathbf{k}}^{\mathrm{sgn}} > 0\big) \ge e^{-\Lambda_n}. \tag{10.1}$$

On the event $Z_{\mathbf{k}}^{\mathrm{sgn}} > 0$, a realized marked partition gives a cocoloring with k_{co} classes. Hence

$$\mathbb{P}(\zeta(G_n) \le k_{\mathrm{co}}) \ge e^{-\Lambda_n}. \tag{10.2}$$

The probability in (10.2) may be small. Following Heckel (2025, Theorem 1), we use vertex-exposure concentration, as in Scott (2017, Theorem 1), to find a large induced subgraph with the desired cocoloring. We then color the vertices outside it. Because this leftover set depends on the graph, we first prove a coloring bound that holds simultaneously for every vertex set. The second lemma uses that bound to control the mean number of additional classes. Markov's inequality then gives an error of $o(n/(\log n)^3)$, smaller than the root separation. Section G.2 gives a stronger explicit tail bound.

Coloring every leftover vertex set.

Lemma 10.1 (Simultaneous leftover coloring)**.** *There are absolute constants $C_0, c > 0$ such that, for all sufficiently large n, with probability at least $1 - e^{-c\sqrt{n}}$, every $S \subseteq [n]$ satisfies*

$$\chi(G_n[S]) \le C_0 \frac{|S|}{\log n} + n^{1/3}. \tag{10.3}$$

Section G.1 proves the lemma by the classical greedy argument. By (1.5) and a union bound, outside an event of probability at most $e^{-c\sqrt{n}}$ every set of $\lceil n^{1/4} \rceil$ vertices spans at least a quarter of its pairs in the complement of G_n, and averaging extends this to every larger set. In a set of at least $n^{1/3}$ vertices, repeatedly choosing a vertex of maximum complement degree and keeping its complement neighbors yields an independent set of G_n whose size is at least a constant multiple of $\log n$. Removing such sets one at a time, and coloring the fewer than $n^{1/3}$ remaining vertices singly, gives (10.3).

Mean excess above the class-count bound. We bound the mean number of vertices left outside a cocolorable set, then color those vertices on the simultaneous density event.

Lemma 10.2 (Mean bound for the cochromatic number)**.** *For all sufficiently large n, every deterministic integer $k \ge 0$ and every deterministic finite real $\Lambda \ge 0$ satisfying*

$$\mathbb{P}(\zeta(G_n) \le k) \ge e^{-\Lambda}$$

obey the following bound. Put $u_0 := \lceil n^{1/4} \rceil$ and

$$\varepsilon_n^{\mathrm{dens}} := \min\left\{1, \binom{n}{u_0} \exp\left(-\frac{\binom{u_0}{2}}{16}\right)\right\}.$$

With $x_+ := \max\{x, 0\}$ and the same absolute C_0 as in Lemma 10.1,

$$\mathbb{E}[(\zeta(G_n) - k)_+] \le \frac{C_0}{\log n}\sqrt{\frac{(n-1)\Lambda}{2}} + n^{1/3} + n\varepsilon_n^{\mathrm{dens}}.$$

The eventual threshold is independent of k, Λ, and the phase.

Proof. Let $\mathcal{D}_n$ be the density event used in the proof of Lemma 10.1: every u_0-set has edge density at least $1/4$ in the complement of G_n. By (1.5) and a union bound,

$$\mathbb{P}\big(\mathcal{D}_n^c\big) \le \varepsilon_n^{\mathrm{dens}}.$$

On $\mathcal{D}_n$, (10.3) holds for every vertex set.

To apply concentration, we measure the largest vertex set that can be cocolored with at most the given number of classes. For $W \subseteq [n]$, let $\zeta(W) = \zeta(G_n[W])$, with $\zeta(\varnothing) = 0$, and define

$$S_k = \max\{|W| : W \subseteq [n],\ \zeta(W) \le k\}. \tag{10.4}$$

Expose the graph in $n - 1$ independent vertex blocks, placing in block v the edges from its vertex to earlier vertices. Two outcomes that differ only in block v agree on all edges away from that vertex. Take a maximizing set in the first outcome. If it contains the changed vertex, delete it and restrict the cocoloring to the remaining set; otherwise keep the set. Either way, it is feasible in the second outcome and has lost at

most one vertex. Interchanging the outcomes gives the reverse comparison. Thus S_k changes by at most one, and (1.3) applies (McDiarmid, 1998, Theorem 3.1).

Since $S_k = n$ exactly when $\zeta(G_n) \le k$, the assumed probability bound and the one-sided upper-tail form of (1.3) give

$$e^{-\Lambda} \le \mathbb{P}(S_k - \mathbb{E}[S_k] \ge n - \mathbb{E}[S_k]) \le \exp\left\{-\frac{2(n - \mathbb{E}[S_k])^2}{n-1}\right\}.$$

Taking logarithms bounds the mean number of uncovered vertices:

$$n - \mathbb{E}[S_k] \le \sqrt{(n-1)\Lambda/2}. \tag{10.5}$$

Choose a maximizing set W and put $V_{\text{left}} = [n] \setminus W$. Combining a cocoloring of $G_n[W]$ using at most k parts with an ordinary coloring of $G_n[V_{\text{left}}]$ gives

$$\zeta(G_n) \le k + \chi(G_n[V_{\text{left}}]). \tag{10.6}$$

On $\mathcal{D}_n$, (10.3) applies to the graph-dependent set V_{left}, so by (10.6) $(\zeta(G_n) - k)_+ \le C_0(n - S_k)/\log n + n^{1/3}$; off $\mathcal{D}_n$, $(\zeta(G_n) - k)_+ \le n$. Since $n - S_k \ge 0$,

$$\mathbb{E}[(\zeta(G_n) - k)_+] \le \frac{C_0}{\log n}\big(n - \mathbb{E}[S_k]\big) + n^{1/3} + n\,\mathbb{P}\big(\mathcal{D}_n^c\big),$$

and (10.5) with $\mathbb{P}\big(\mathcal{D}_n^c\big) \le \varepsilon_n^{\text{dens}}$ proves the assertion. Finally,

$$\log\binom{n}{u_0} = O(n^{1/4}\log n), \qquad \binom{u_0}{2} = (1 + o(1))\sqrt{n}/2,$$

so $\varepsilon_n^{\text{dens}} \le e^{-c\sqrt{n}}$ eventually for one absolute $c > 0$. □

For the probability bound (10.2), the explicit second-moment bound satisfies $\Lambda_n = o(n/(\log n)^4)$. Consequently,

$$\mathbb{E}[(\zeta(G_n) - k_{\text{co}})_+] = o\left(\frac{n}{(\log n)^3}\right).$$

This also gives $\mathbb{E}[\zeta(G_n)] \le k_{\text{co}} + o(n/(\log n)^3)$, uniformly along the full integer sequence.

Applying the probability bound. The mean bound gives the high-probability conclusion by Markov's inequality. Put $s_n = n/(\log n)^3$ and let

$$b_n := \frac{C_0}{\log n}\sqrt{\frac{(n-1)\Lambda_n}{2}} + n^{1/3} + n\varepsilon_n^{\text{dens}}, \qquad a_n := \sqrt{b_n s_n}. \tag{10.7}$$

Here $b_n > 0$ is a deterministic upper bound for $\mathbb{E}(\zeta(G_n) - k_{\text{co}})_+$. Since $\Lambda_n = o(n/(\log n)^4)$ and $\varepsilon_n^{\text{dens}} \le e^{-c\sqrt{n}}$,

$$b_n = o(s_n), \qquad a_n = o(s_n). \tag{10.8}$$

Markov's inequality now gives

$$\mathbb{P}\big(\zeta(G_n) > k_{\text{co}} + a_n\big) \le \frac{b_n}{a_n} = \sqrt{\frac{b_n}{s_n}} \longrightarrow 0. \tag{10.9}$$

All these sequences are deterministic, and the bounds hold uniformly in the phase. The stronger tail estimate is Lemma G.1 in Section G.2.

Uniformity across the phase. The root estimates are uniform on K_*, and the errors above have deterministic bounds independent of the phase. The residual estimates cover all residual masses, and the leftover-coloring event holds for every vertex set simultaneously. All eventual conditions, including $\alpha > 8$, $U \ge 2$, $\rho_\Sigma \le 1$ and $\tau_n^{\text{end}} \le 1$, therefore hold after one deterministic threshold. Section G.2 records the corresponding uniformity for the explicit tail bound.

11. The phase-resolved gap and a uniform constant

We now combine the three estimates (A)–(C) from the introduction. The first-moment roots give a deterministic separation. With high probability, the chromatic number lies above its lower location and the cochromatic number lies below the amplified upper location. On the intersection of these events, their separation gives the phase-resolved gap. Continuity on the closed phase interval then gives a uniform constant.

Proof of the main theorem. Take n beyond the common deterministic threshold at the end of Section 10. Proposition 5.5(iii) gives, uniformly in the phase,

$$k_\chi^- - k_{\mathrm{co}} = \left[\frac{(\log 2)^2}{4} A_5(\delta_n) + o(1)\right] \frac{n}{(\log n)^3}. \tag{11.1}$$

It combines (5.11) with $|k_\chi^- - r_+(n)| < \log n + 2$ from (4.7) and $16 \le k_{\mathrm{co}} - r_5^{\mathrm{co}} < 17$; the rounding (5.16) changes only the class-size counts, not k_{co}. By (4.6), $\mathbb{P}(\chi(G_n) > k_\chi^-) \to 1$.

The amplification result (10.9) gives a deterministic sequence $a_n = o(n/(\log n)^3)$ such that

$$\mathbb{P}\big(\zeta(G_n) \le k_{\mathrm{co}} + a_n\big) \longrightarrow 1.$$

Figure 1 in the introduction shows the deterministic locations used in this comparison.

To make the final error explicit, put

$$s_n := \frac{n}{(\log n)^3}, \qquad \theta_n := \frac{k_\chi^- - k_{\mathrm{co}}}{s_n} - \frac{(\log 2)^2}{4} A_5(\delta_n), \qquad \eta_n := \frac{a_n}{s_n}.$$

By (11.1) and (10.9), $\theta_n \to 0$ and $\eta_n \to 0$. Hence the nonnegative deterministic error

$$\varepsilon_n^{\mathrm{gap}} := |\theta_n| + \eta_n$$

tends to zero. On the event

$$\{\chi(G_n) > k_\chi^-\} \cap \{\zeta(G_n) \le k_{\mathrm{co}} + a_n\},$$

$$\chi(G_n) - \zeta(G_n) > k_\chi^- - k_{\mathrm{co}} - a_n \ge \left[\frac{(\log 2)^2}{4} A_5(\delta_n) - \varepsilon_n^{\mathrm{gap}}\right] s_n.$$

By (4.6) and (10.9), a union bound shows that this intersection has probability $1 - o(1)$. Consequently,

$$\mathbb{P}\left(\chi(G_n) - \zeta(G_n) \ge \left[\frac{(\log 2)^2}{4} A_5(\delta_n) - \varepsilon_n^{\mathrm{gap}}\right] \frac{n}{(\log n)^3}\right) \longrightarrow 1. \tag{11.2}$$

On this tail, set $\varepsilon_n := \varepsilon_n^{\mathrm{gap}}$ and define the finitely many preceding terms arbitrarily as nonnegative reals. Thus $(\varepsilon_n)_{n\ge 2}$ is deterministic, nonnegative, and tends to zero.

The uniform coefficient.

Put $q := \log 2$. Lemma 5.1 shows that A_5 is continuous and nonconstant on $[0,1]$ and locates its smallest value at an endpoint, while Lemma 5.3 shows that $A_5 > \gamma_5 > 0$. Hence

$$m_5 := \min\{A_5(0), A_5(1)\} = \min_{0\le\delta\le 1} A_5(\delta) > 0, \qquad c_5 = \frac{q^2}{4} m_5 > 0.$$

Fix $c < c_5$. Since $\varepsilon_n^{\mathrm{gap}} \to 0$, eventually $\varepsilon_n^{\mathrm{gap}} < c_5 - c$. Therefore the coefficient in (11.2) is at least $q^2 A_5(\delta_n)/4 - \varepsilon_n^{\mathrm{gap}} \ge c_5 - \varepsilon_n^{\mathrm{gap}} > c$ for all sufficiently large n. It follows that

$$\mathbb{P}\left(\chi(G_n) - \zeta(G_n) \ge c\,\frac{n}{(\log n)^3}\right) \longrightarrow 1.$$

This proves the main theorem.

To bound the constant from below, Lemma 5.4 gives $m_5 > \log(169/100)$. Consequently

$$c_5 > \frac{(\log 2)^2}{4} \log\left(\frac{169}{100}\right) = 0.0630268508\ldots,$$

and the uniform bound holds in particular with $c = \frac{1}{4}(\log 2)^2 \log(169/100)$. Since $n/(\log n)^3 \to \infty$, the chromatic–cochromatic difference tends to infinity with high probability along the full sequence of integers. We have proved the lower-bound direction of the conjectured $n/(\log n)^3$ scale; a matching upper bound and the optimal constant remain open. □

The remainder rate and summable failure probabilities.

Remark 11.1 (Rate of the phase-resolved remainder). The deterministic error in the main theorem may be chosen so that

$$\varepsilon_n = O\left(\frac{\log\log n}{\log n}\right),$$

with a phase-independent implied constant. Indeed, put $L = \log n$ and $\ell = \log\log n$. The explicit bounds in Section 9.6 give $\Lambda_n = O(n^{3/4}L)$. Hence the quantities in (10.7) satisfy

$$b_n = O(n^{7/8}L^{-1/2}), \qquad a_n = O(n^{15/16}L^{-7/4}) = o(n\ell/L^4).$$

Combining (5.11), (4.1), and (5.13) therefore gives

$$\theta_n = O(\ell/L), \qquad \eta_n = o(\ell/L).$$

The error $\varepsilon_n^{\mathrm{gap}} = |\theta_n| + \eta_n$ used above has the asserted rate. Equivalently, the phase-resolved lower bound has deterministic remainder $O(n \log\log n/(\log n)^4)$. All estimates are uniform over the phase.

The pair bound in Section 4 sharpens the probability estimate enough to sum the failures over all n.

Corollary 11.2 (A summable bound on the failure probability)**.** *Put $q = \log 2$, $\beta = 2/q + 1/2$, $L = \log n$, and*

$$w_n := \frac{L^\beta}{n}\left(\frac{L}{n}\right)^{1-\delta_n} = n^{\delta_n - 2} L^{\beta+1-\delta_n}.$$

There are constants $C_1, C_2 > 0$ and an integer N, all independent of the phase, such that every integer $n \geq N$ satisfies

$$\mathbb{P}\left(\chi(G_n) - \zeta(G_n) < \left\lceil \frac{q^2}{4} A_5(\delta_n) - C_1 \frac{\log\log n}{\log n} \right\rceil \frac{n}{(\log n)^3}\right) \leq C_2 \left(w_n^2 + \frac{w_n L^{5/2}}{\sqrt{n}} + n^{-2}\right).$$

The right-hand side is $O(n^{-3/2}(\log n)^{\beta+5/2})$, uniformly in the phase, and is summable over n. For each fixed $c < c_5$, the uniform lower bound in the main theorem with coefficient c has the same failure bound, after increasing N.

Section G.3 proves the corollary. It applies the explicit tail bound of Lemma G.1 with tail parameter $2\log n$, together with the sharper chromatic failure bound (4.5) and the expansion (2.3).

Corollary 11.3 (Almost-sure eventual gap)**.** *Fix any joint law of $(G_n)_{n\geq 1}$ whose n-th marginal is $G(n, 1/2)$. With probability one, for every sufficiently large n,*

$$\chi(G_n) - \zeta(G_n) \geq \left\lceil \frac{(\log 2)^2}{4} A_5(\delta_n) - C_1 \frac{\log\log n}{\log n} \right\rceil \frac{n}{(\log n)^3}.$$

For each fixed $c < c_5$, the uniform lower bound in the main theorem with coefficient c also holds almost surely for every sufficiently large n.

Proof. Under the fixed joint law, let F_n be failure of the phase-resolved bound. The preceding corollary bounds each marginal failure probability and gives $\sum_n \mathbb{P}(F_n) < \infty$. The first Borel–Cantelli lemma therefore implies that only finitely many F_n occur. Independence across n is not required. The eventual threshold may depend on the realization; the assertion is for each fixed joint law. □

AI assistance disclosure. OpenAI GPT-5.6 Sol, accessed through ChatGPT and Codex, was used to probe the written proof for counterexamples, check symbolic consistency, and generate candidate Lean code. Codex also assisted with supplementary proof arguments for the finite-dual approximation and residual overlap bounds. Aristotle, Harmonic's automated theorem-proving system (Achim et al., 2025), performed automated Lean proof search and a clean-environment replay. Every formal output incorporated into the released development was reviewed and rebuilt under the pinned Lean toolchain. Machine-checked claims rest on the Lean kernel, and manuscript-only claims on the written proof. Samuil Petkov directed the work, selected the claims and arguments, reviewed the manuscript and formalization, and assumes responsibility for the accuracy, integrity, citations, and final submission.

References

Tudor Achim, Alex Best, Alberto Bietti, Kevin Der, Mathïs Fédérico, Sergei Gukov, Daniel Halpern-Leistner, Kirsten Henningsgard, Yury Kudryashov, Alexander Meiburg, Martin Michelsen, Riley Patterson, Eric Rodriguez, Laura Scharff, Vikram Shanker, Vladmir Sicca, Hari Sowrirajan, Aidan Swope, Matyas Tamas, Vlad Tenev, Jonathan Thomm, Harold Williams, and Lawrence Wu. Aristotle: IMO-level automated theorem proving, 2025. URL https://doi.org/10.48550/arXiv.2510.01346.

Thomas F. Bloom. Erdős problem #625. https://www.erdosproblems.com/625, 2026. Accessed 16 July 2026.

Béla Bollobás. The chromatic number of random graphs. *Combinatorica*, 8(1):49–55, 1988. doi: 10.1007/BF02122551. URL https://doi.org/10.1007/BF02122551.

Béla Bollobás and Paul Erdős. Cliques in random graphs. *Mathematical Proceedings of the Cambridge Philosophical Society*, 80(3): 419–427, 1976. doi: 10.1017/S0305004100053056. URL https://doi.org/10.1017/S0305004100053056.

Béla Bollobás and Andrew Thomason. Generalized chromatic numbers of random graphs. *Random Structures & Algorithms*, 6(2–3): 353–356, 1995. doi: 10.1002/rsa.3240060222. URL https://doi.org/10.1002/rsa.3240060222.

Paul Erdős and John Gimbel. Some problems and results in cochromatic theory. In John Gimbel, John W. Kennedy, and Louis V. Quintas, editors, *Quo Vadis, Graph Theory?*, volume 55 of *Annals of Discrete Mathematics*, pages 261–264. North-Holland/Elsevier, 1993. doi: 10.1016/S0167-5060(08)70393-5. URL https://doi.org/10.1016/S0167-5060(08)70393-5.

John Gimbel. Some of my favorite coloring problems for graphs and digraphs. In Ralucca Gera, Stephen T. Hedetniemi, and Craig Larson, editors, *Graph Theory: Favorite Conjectures and Open Problems – 1*, Problem Books in Mathematics, pages 95–108. Springer, Cham, 2016. ISBN 978-3-319-31940-7. doi: 10.1007/978-3-319-31940-7_7. URL https://doi.org/10.1007/978-3-319-31940-7_7.

Geoffrey R. Grimmett and Colin J. H. McDiarmid. On colouring random graphs. *Mathematical Proceedings of the Cambridge Philosophical Society*, 77(2):313–324, 1975. doi: 10.1017/S0305004100051124. URL https://doi.org/10.1017/S0305004100051124.

Annika Heckel. The chromatic number of dense random graphs. *Random Structures & Algorithms*, 53(1):140–182, 2018. doi: 10.1002/rsa.20757. URL https://doi.org/10.1002/rsa.20757.

Annika Heckel. On a question of Erdős and Gimbel on the cochromatic number. *The Electronic Journal of Combinatorics*, 31(4): P4.72, 2024. doi: 10.37236/13346. URL https://doi.org/10.37236/13346.

Annika Heckel. The difference between the chromatic and the cochromatic number of a random graph, 2025. URL https://arxiv.org/abs/2409.17614. Version 2, revised 19 February 2025; first submitted in 2024.

Annika Heckel and Konstantinos Panagiotou. Colouring random graphs: Tame colourings, 2023. URL https://arxiv.org/abs/2306.07253.

Colin McDiarmid. On the method of bounded differences. In Johannes Siemons, editor, *Surveys in Combinatorics, 1989*, volume 141 of *London Mathematical Society Lecture Note Series*, pages 148–188. Cambridge University Press, 1989. doi: 10.1017/CBO9781107359949.008. URL https://doi.org/10.1017/CBO9781107359949.008.

Colin McDiarmid. On the chromatic number of random graphs. *Random Structures & Algorithms*, 1(4):435–442, 1990. doi: 10.1002/rsa.3240010404. URL https://doi.org/10.1002/rsa.3240010404.

Colin McDiarmid. Concentration. In Michel Habib, Colin McDiarmid, Jorge Ramirez-Alfonsin, and Bruce Reed, editors, *Probabilistic Methods for Algorithmic Discrete Mathematics*, volume 16 of *Algorithms and Combinatorics*, pages 195–248. Springer, Berlin, Heidelberg, 1998. doi: 10.1007/978-3-662-12788-9_6. URL https://cgm.cs.mcgill.ca/~breed/conc/colin.pdf. Theorem 3.1; author preprint, pp. 10–11.

Konstantinos Panagiotou and Angelika Steger. A note on the chromatic number of a dense random graph. *Discrete Mathematics*, 309 (10):3420–3423, 2009. doi: 10.1016/j.disc.2008.09.019. URL https://doi.org/10.1016/j.disc.2008.09.019.

Samuil Petkov. Erdős Problem 625: Lean 4 formalization and reproducibility record. Software companion at exact Git revisions, 2026. URL https://github.com/SamPetkov/Erdos/commit/31bbe00c529a996bdb61b880120d71240172d18f. Formal source revision 824e4b609466d2e26b216a76ecf103184dac2663; replay and archive revision 31bbe00c529a996bdb61b880120d71240172d18f; Lean and Mathlib v4.31.0.

Edward R. Scheinerman. Generalized chromatic numbers of random graphs. *SIAM Journal on Discrete Mathematics*, 5(1):74–80, 1992. doi: 10.1137/0405006. URL https://doi.org/10.1137/0405006.

Alex Scott. On the concentration of the chromatic number of random graphs, 2017. URL https://arxiv.org/abs/0806.0178. Version 2; first submitted in 2008.

Raphael Steiner. On the difference between the chromatic and cochromatic number. *SIAM Journal on Discrete Mathematics*, 39(4): 2268–2274, 2025. doi: 10.1137/24M1715180. URL https://doi.org/10.1137/24M1715180. Published online 17 November 2025; arXiv:2408.02400.

Yaming Yu. Relative log-concavity and a pair of triangle inequalities. *Bernoulli*, 16(2):459–470, 2010. doi: 10.3150/09-BEJ216. URL https://doi.org/10.3150/09-BEJ216.

Appendix A. Uniform first-moment estimates

This appendix proves the three first-moment lemmas whose proofs were deferred: the phase expansion of Lemma 2.1, the root, slope and dual estimates of Lemma 3.1, and the independent-set pair bound of Lemma 4.1. The first two are used throughout the paper, and the third removes the size cap in Section 4; in its proof, as there, X_s counts the independent s-subsets of G_n and $\mu_s = \mathbb{E}X_s$. We use the notation of Section 2: $q = \log 2$, $L = \log n$, $\ell = \log\log n$ (not a number of common classes), $C = 1 + \log q - q$, $\alpha = \lfloor \alpha_0 \rfloor$, $\delta = \alpha_0 - \alpha$ and $b = 1 - \delta$. In the proof of Lemma 3.1, n is large enough that $\alpha \geq 10$, as in Section 3. Every threshold and implied constant below is independent of the phase, including along sequences approaching either endpoint of $[0, 1]$.

A.1. Uniform Stirling expansion and adjacent-size ratios.

Proof of Lemma 2.1. We first control the remainder in the logarithmic count at the integer center. Exact adjacent-size ratios will then give the upper and lower size estimates. Because $\alpha = O(L)$ uniformly in the phase, one has $\alpha < n/2$ for all sufficiently large n. For $0 \leq x \leq 1/2$, $-2x \leq \log(1-x) \leq -x$. Hence

$$\log(n)_\alpha = \alpha L + \sum_{j=0}^{\alpha-1} \log\left(1 - \frac{j}{n}\right) = \alpha L + R_{n,\alpha}, \qquad |R_{n,\alpha}| \leq \frac{\alpha^2}{n}.$$

Using the uniform Stirling remainder

$$\log(\alpha!) = \alpha \log \alpha - \alpha + \frac{1}{2}\log(2\pi\alpha) + O(1/\alpha)$$

in

$$\log \mu_\alpha = \log(n)_\alpha - \log(\alpha!) - q\binom{\alpha}{2}$$

gives

$$\log \mu_\alpha = \alpha\left(L - \log\alpha + 1 - \frac{q}{2}(\alpha - 1)\right) - \frac{1}{2}\log(2\pi\alpha) + O(1/L). \tag{A.1}$$

All implied constants and remainder bounds here are uniform in δ.

Write

$$\alpha = \frac{2L}{q}(1 + x_n), \qquad x_n := \frac{-\ell + C + qb/2}{L}.$$

Uniformly for the actual range $0 < b \le 1$ (and on its closure), one has $|x_n| \le 1/2$ for all sufficiently large n. The Taylor formula $\log(1 + x) = x + O(x^2)$ therefore gives

$$\log\alpha = \ell + q - \log q + \frac{-\ell + C + qb/2}{L} + O\left(\frac{\ell^2}{L^2}\right).$$

Since

$$\frac{q}{2}(\alpha - 1) = L - \ell + C - \frac{q\delta}{2},$$

we obtain the uniform expansion

$$L - \log\alpha + 1 - \frac{q}{2}(\alpha - 1) = \frac{q\delta}{2} + \frac{\ell - C - qb/2}{L} + O\left(\frac{\ell^2}{L^2}\right). \tag{A.2}$$

Multiplying by $\alpha = 2(L - \ell + C)/q + b$ and collecting terms yields

$$\begin{aligned}\alpha\left(L - \log\alpha + 1 - \frac{q}{2}(\alpha - 1)\right) = \delta L + \left(\frac{2}{q} - \delta\right)\ell \\ + \delta C + \frac{q}{2}\delta(1 - \delta) - \frac{2C}{q} - (1 - \delta) + O\left(\frac{1 + \ell^2}{L}\right).\end{aligned}$$

Also

$$-\frac{1}{2}\log(2\pi\alpha) = -\frac{1}{2}\ell - \frac{1}{2}\log(2\pi) - \frac{q}{2} + \frac{1}{2}\log q + O(\ell/L).$$

Thus (2.2) holds with

$$\begin{aligned}K(\delta) := \delta C + \frac{q}{2}\delta(1 - \delta) - \frac{2C}{q} - (1 - \delta) \\ - \frac{1}{2}\log(2\pi) - \frac{q}{2} + \frac{1}{2}\log q.\end{aligned} \tag{A.3}$$

This function is continuous and therefore bounded on $[0, 1]$. Combining the remainders in (A.1), (A.2) and the expansion of $\log(2\pi\alpha)$ bounds their absolute value by $C'(1 + \ell^2)/L$ for an absolute C'. Choose $C_{\mathrm{ph}} \ge C'$; then $|E_n^{\mathrm{ph}}| \le \varepsilon_n^{\mathrm{ph}}$ as required in (2.2), with E_n^{ph} denoting that combined remainder.

For integers $1 \le s < n$, the adjacent-size ratios are exact:

$$\frac{\mu_{s+1}}{\mu_s} = \frac{n - s}{s + 1}2^{-s}, \qquad \frac{\mu_{s-1}}{\mu_s} = \frac{s}{n - s + 1}2^{s-1}. \tag{A.4}$$

The displayed formula for α gives the exact phase representation

$$2^\alpha = \exp(q\alpha) = \exp(2C + qb)\frac{n^2}{L^2}.$$

Since $0 < b \le 1$, the multiplicative factor $\exp(2C + qb)$ is bounded above and below by positive absolute constants. Therefore, for every fixed M, uniformly over integers s with $|s - \alpha| \le M$,

$$\frac{\mu_{s+1}}{\mu_s} = \Theta(L/n), \qquad \frac{\mu_{s-1}}{\mu_s} = \Theta(n/L).$$

In particular,

$$\mu_{\alpha+2} = \mu_\alpha\,\Theta(L^2/n^2), \qquad \mu_{\alpha-2} = \mu_\alpha\,\Theta(n^2/L^2). \tag{A.5}$$

Taking logarithms and using (2.2) proves the first line of (2.3). Since $\delta \le 1$, its leading term is at most $-L$, and the coefficient of ℓ is uniformly bounded; this proves the second line.

For the lower adjacent size, (2.2) and (A.5) give

$$\log\mu_{\alpha-2} = 2L + \left(\frac{2}{q} - \frac{5}{2}\right)\ell + \delta(L - \ell) + K(\delta) + O(1).$$

For large n, $L - \ell > 0$, while K is bounded below. Exponentiating gives (2.4), with one absolute constant c and one threshold valid for the full phase interval. □

A.2. **Proof of the uniform root estimates.**

Proof of Lemma 3.1. We first compare the finite and limiting duals, which proves (3.9)–(3.9b); then we prove the identity (3.8); finally we locate the roots and bound the slopes, which proves (3.6)–(3.7). The quantities h_n, $S^{(n)}$, $Z_{n,S}$, $M_{n,S}$, $\lambda_{n,S}$ and $\mathcal{F}_{n,S}$ depend on n only through the integer α, and T ranges over the fixed interval K_*; so the first two steps are uniform in the phase automatically. In the last step the phase enters only through $b = 1 - \delta \in [0, 1]$ in (2.1), that is, through s_0 and $T_0 \in [2/q, 1 + 2/q]$.

Dividing (3.2) by $k = n/s$ and writing $p_i = k_i/k$ gives the formula below. All sums in this formula are over $S^{(n)}$. The maximum is over nonnegative weights with total mass one and mean deficit $\alpha - s$.

$$\frac{L_S(n, n/s)}{n/s} = (s-1)L - s + 1 + \log s + \max\left[-\sum_i p_i \log p_i - \sum_i p_i \log d_{\alpha-i}\right].$$

The definition of the remainder in (3.1a) gives

$$h_n(i) = -\frac{q}{2}i^2 + O(i^2/\alpha) \quad \text{for every fixed } i, \qquad h_n(i) \le -\frac{q}{2}i^2 \quad (i \ge -1). \tag{A.6}$$

The first estimate controls each fixed finite set of deficits uniformly. The second applies to the whole finite support.

Finite and limiting duals and optimizing tilts.

We first choose the exponential tilt so that the weighted distribution has the required mean. The target center satisfies $2/q \le T_0 \le 1 + 2/q$, so it lies in the interior of K_*. Every target in K_* is also strictly inside the feasible mean range of both supports.

Choose $\Lambda > 0$ so that

$$M_S(-\Lambda) < \min K_* < \max K_* < M_S(\Lambda) \qquad (S = S_+, S_5).$$

The limiting mean maps are strictly increasing, and their endpoint limits bracket K_*, so this choice is possible. Before inverting the mean maps, we compare the finite weights with their limiting Gaussian weights. Write

$$w_i(\lambda) := \exp\Big(\lambda i - \frac{q}{2}i^2\Big), \qquad \rho_n(i) := -h_n(i) - \frac{q}{2}i^2$$

for indices in the finite support. Formula (3.1a) shows that this loss in the exponent is nonnegative, and

$$0 \le \rho_n(-1) = \log(1 + 1/\alpha) \le 1/\alpha.$$

For $0 \le i \le \alpha/2$, use $-\log(1-x) \le 2x$ on $0 \le x \le 1/2$ to get

$$0 \le \rho_n(i) = -\sum_{r=0}^{i-1} \log\Big(1 - \frac{r}{\alpha}\Big) \le \frac{i(i-1)}{\alpha} \le \frac{i^2}{\alpha}.$$

Since $0 \le 1 - e^{-\rho_n(i)} \le \rho_n(i)$, for $m = 0, 1, 2$ the retained low indices satisfy

$$\sum_{\substack{i \in S^{(n)} \\ i \le \alpha/2}} |i|^m \left| e^{\lambda i + h_n(i)} - w_i(\lambda) \right| \le \frac{C_m}{\alpha} \qquad (|\lambda| \le \Lambda).$$

Here $|i|^0$ means one. The constants are finite because every polynomial times $\exp(\Lambda|i| - qi^2/2)$ is summable.

For the remaining indices, (A.6) bounds the difference by the limiting weight itself. The same tail also contains every index omitted by the finite cutoff. For sufficiently large α,

$$\sum_{i > \alpha/2} i^m e^{\Lambda i - qi^2/2} \le \sum_{i > \alpha/2} i^m e^{-qi^2/4} \le C_m e^{-c\alpha^2}.$$

Comparison with a Gaussian integral proves the last bound: the fixed polynomial can be absorbed by weakening the Gaussian exponent. The tail for S_5 is eventually empty. Combining the low-index and tail bounds, the zeroth, first and second unnormalized tilted moments differ by $O(1/\alpha)$ uniformly on $[-\Lambda, \Lambda]$, for both supports. The zeroth moments also have a common positive lower bound. Indeed, retain the $i = 2$ term, which occurs in both finite supports; the tilt interval and $h_n(2)$ are bounded. We may therefore take logarithms and ratios to obtain, uniformly on this interval,

$$\begin{aligned} \log Z_{n,S} - \log Z_S &= O(1/\alpha), \\ M_{n,S} - M_S &= O(1/\alpha), \\ M'_{n,S} - M'_S &= O(1/\alpha). \end{aligned}$$

For all sufficiently large n, the finite support $S^{(n)}$ contains at least two indices, and hence $M'_{n,S}(\lambda) = \mathrm{Var}_{n,S,\lambda}(i) > 0$ for every $\lambda \in \mathbb{R}$. Likewise, $M'_S(\lambda) = \mathrm{Var}_{S,\lambda}(i) > 0$ for every λ. Thus

both mean maps are globally strictly increasing. Take the minimum below over $S \in \{S_+, S_5\}$ and $|\lambda| \le \Lambda$. Continuity and compactness give

$$v_* := \frac{1}{2} \min_{S,\lambda} M_S'(\lambda) > 0.$$

For all sufficiently large n, one has $M_{n,S}'(\lambda) \ge v_*$ on the same interval, and the finite mean maps also bracket K_*. Hence, for every $T \in K_*$, the equations

$$M_{n,S}(\lambda_{n,S}(T)) = T, \qquad M_S(\lambda_S(T)) = T$$

have unique solutions in $[-\Lambda, \Lambda]$. The derivative lower bound makes the inverse mean maps uniformly v_*^{-1}-Lipschitz. Consequently

$$\sup_{T \in K_*} |\lambda_{n,S}(T) - \lambda_S(T)| \le v_*^{-1} \sup_{|\lambda| \le \Lambda} |M_{n,S}(\lambda) - M_S(\lambda)| = O(1/\alpha).$$

For the finite supports, the entropy functional in (3.8a) is strictly concave on its feasible simplex. Its unique optimizer is

$$p_{n,S,T}(i) = \frac{\exp\{\lambda_{n,S}(T) i + h_n(i)\}}{Z_{n,S}(\lambda_{n,S}(T))},$$

and its dual value is exactly

$$\mathcal{F}_{n,S}(T) = \log Z_{n,S}(\lambda_{n,S}(T)) - \lambda_{n,S}(T) T.$$

For the countable support S_+, strict concavity alone does not establish attainment. We verify it using the tilted distribution. Fix $T \in K_*$, write $\lambda = \lambda_{S_+}(T)$, and set

$$\pi_\lambda(i) := \frac{\exp\{\lambda i - q i^2/2\}}{Z_{S_+}(\lambda)}.$$

Write

$$D_{\mathrm{KL}}(p \| \pi_\lambda) := \sum_i p_i \log \frac{p_i}{\pi_\lambda(i)},$$

with the convention that a summand with $p_i = 0$ is zero. The tilted distribution defined above has mean T. For every feasible probability vector p with finite quadratic moment,

$$-\sum_i p_i \log p_i - \frac{q}{2} \sum_i i^2 p_i = \log Z_{S_+}(\lambda) - \lambda T - D_{\mathrm{KL}}(p \| \pi_\lambda). \tag{A.7}$$

The identity follows first for finitely supported p and then by truncation; finite quadratic moment also makes the entropy finite by comparison with a Gaussian law. We interpret the objective as $-\infty$ when the quadratic moment diverges. Since relative entropy is nonnegative and vanishes only at $p = \pi_\lambda$, (A.7) proves finiteness, attainment, uniqueness, and the dual formula for S_+. The same identity also recovers the finite-support case. The dual objectives for the finite and limiting supports attain their minima in the same tilt interval. Their difference at a fixed tilt is $\log Z_{n,S} - \log Z_S$, so the difference of their minimum values is at most the uniform norm of this quantity. Together with the tilt estimate, this proves the uniform convergence $\mathcal{F}_{n,S} \to \mathcal{F}_S$ on K_* and

$$\varepsilon_n^{\mathrm{dual}} = O(1/\alpha).$$

On the finite support S_5, substitute the exponent and tilt error bounds into the optimizer formula to obtain the coordinatewise rate (3.9b). The limiting coordinates are positive and continuous in T on the compact set K_*. They therefore have a positive minimum, and the coordinatewise rate transfers this lower bound to the finite optimizers.

Exact cancellation between supports.

We next return to the counting exponent. Substitution of (3.1b) gives

$$\begin{aligned} \Psi_{n,S}(s) &:= \frac{L_S(n, n/s)}{n/s} \\ &= (s-1)L - s + 1 + \log s + A_n + B_n(\alpha - s) + \mathcal{F}_{n,S}(\alpha - s). \end{aligned} \tag{A.8}$$

Fix a class count k feasible for both supports. They then have the same mean class size s and mean deficit $T = \alpha - s$. The affine terms in (A.8) cancel, leaving exactly the difference of the two finite dual values in (3.8). This comparison uses the same numbers of parts and vertices on both sides, before taking either limiting dual.

The exponent at the phase center and its nearby slope.

To locate the zero, we estimate the exponent at the phase center and its derivative nearby. At $s = s_0$, let $T_0 = \alpha - s_0$. Since $A_n = -q\binom{\alpha}{2} - \log(\alpha!)$ and $\alpha = s_0 + T_0$, the terms of (A.8) involving A_n and B_n satisfy

$$-q\binom{\alpha}{2} + \left(q\alpha - \frac{q}{2}\right)T_0 = -\frac{q}{2}s_0^2 + \frac{q}{2}s_0 + \frac{q}{2}T_0^2, \qquad (T_0 - \alpha)\log\alpha + \alpha = s_0(1 - \log\alpha) + T_0.$$

Applying (1.2) to $\log(\alpha!)$ and adding $(s_0 - 1)L - s_0 + 1 + \log s_0$ gives

$$\begin{aligned}\Psi_{n,S}(s_0) - \mathcal{F}_{n,S}(T_0) = s_0\left(L - \log\alpha - \frac{q}{2}s_0 + \frac{q}{2}\right) - L + T_0 + 1 + \log s_0 + \frac{q}{2}T_0^2 \\ - \frac{1}{2}\log(2\pi\alpha) + O(1/L).\end{aligned} \tag{A.9}$$

Equation (2.1) gives the exact identity

$$\frac{qs_0}{2} = L - \ell + \log q - q.$$

Since $\alpha = s_0 + T_0$ and T_0 is uniformly bounded,

$$\log\alpha = \log s_0 + \frac{T_0}{s_0} + O(L^{-2}), \qquad \log s_0 = \ell + q - \log q + O(\ell/L).$$

Consequently

$$\begin{aligned}s_0\left(L - \log\alpha - \frac{q}{2}s_0 + \frac{q}{2}\right) &= -T_0 + \frac{q}{2}s_0 + O(\ell) \\ &= -T_0 + L - \ell + \log q - q + O(\ell).\end{aligned}$$

Substitution in (A.9) cancels the terms L and T_0 and leaves $O(\ell)$ uniformly in the phase. The dual term in (A.8) is $O(1)$ uniformly on K_*, by the Gaussian majorant above. Therefore

$$\Psi_{n,S}(s_0) = O(\ell) \tag{A.10}$$

uniformly for both supports.

The phase formula for α and the expansion of $\log\alpha$ also give

$$B_n = 2L - \ell + O(1) \tag{A.11}$$

uniformly in δ. Since $M'_{n,S} \geq v_* > 0$ on $[-\Lambda, \Lambda]$, the inverse function theorem makes $\lambda_{n,S}$ differentiable on K_*. Differentiating $\mathcal{F}_{n,S}(T) = \log Z_{n,S}(\lambda_{n,S}(T)) - \lambda_{n,S}(T)T$ and using $M_{n,S}(\lambda_{n,S}(T)) = T$ gives

$$\frac{d}{dT}\mathcal{F}_{n,S}(T) = -\lambda_{n,S}(T).$$

Differentiating (A.8), with $T = \alpha - s$, yields

$$\Psi'_{n,S}(s) = L - 1 + s^{-1} - B_n + \lambda_{n,S}(T) = -L + O_A(\ell) \tag{A.12}$$

uniformly whenever $|s - s_0| \leq A\ell/L$. Indeed, the image of this corridor under $T = \alpha - s$ lies inside K_* for all sufficiently large n, because K_* has a fixed margin around the complete range of T_0.

Roots and slopes in the phase corridor.

Set

$$\Psi_{n,S,c}(s) := \Psi_{n,S}(s) + c.$$

By (A.10), there is an absolute C_0 such that $|\Psi_{n,S,c}(s_0)| \leq C_0\ell$ for both supports and every $0 \leq c \leq q$. By (A.12), after one phase-independent threshold,

$$\Psi'_{n,S,c}(s) \leq -\frac{1}{2}L$$

throughout each fixed corridor. Choose $A_* > 2C_0$. Integration gives

$$\begin{aligned}\Psi_{n,S,c}\left(s_0 - A_*\frac{\ell}{L}\right) &\geq \left(\frac{A_*}{2} - C_0\right)\ell > 0, \\ \Psi_{n,S,c}\left(s_0 + A_*\frac{\ell}{L}\right) &\leq \left(C_0 - \frac{A_*}{2}\right)\ell < 0.\end{aligned}$$

The opposite endpoint signs give a zero, and strict decrease in s makes it unique on this interval. Since $k = n/s > 0$, this gives the *corridor* root $r_{S,c}$ of $L_S(n,k) + ck = 0$ in (3.6). The uniqueness claim is restricted to this corridor.

Finally,

$$L_S(n,k) + ck = k\Psi_{n,S,c}(s), \qquad \frac{ds}{dk} = -\frac{s}{k}.$$

Hence

$$\begin{aligned}\frac{\partial}{\partial k}\{L_S(n,k)+ck\} &= \Psi_{n,S,c}(s) - s\Psi'_{n,S}(s)\\ &= \frac{2}{q}L^2 + O_A(L\ell),\end{aligned} \tag{A.13}$$

because $s = (2/q)L + O_A(\ell)$, (A.12) holds, and $\Psi_{n,S,c}(s) = O_A(\ell)$ throughout the corridor. This proves (3.7) and $\varepsilon^{\mathrm{slope}}_{n,A} \to 0$ with one threshold for the full phase interval. □

A.3. **Proof of the independent-set pair bound.**

Proof of Lemma 4.1. Write $(x)_j = x(x-1)\cdots(x-j+1)$, with $(x)_0 = 1$. For independently and uniformly sampled s-subsets S, T of $[n]$, let $p_j = \mathbb{P}(|S \cap T| = j)$. Counting ordered distinct pairs gives

$$\mathbb{E}[X_s(X_s-1)] = \mu_s^2 \sum_{j=0}^{s-1} p_j\, 2^{\binom{j}{2}}.$$

For $0 \le j \le s/2$, a union bound over j-subsets of S gives

$$p_j \le \binom{s}{j}\frac{(s)_j}{(n)_j} \le \frac{1}{j!}\left(\frac{s^2}{n-s}\right)^j.$$

Since $2^{(j-1)/2} \le 2^{s/4}$, this part of the sum is at most e^a.

For $j > s/2$, write $r = s - j$, so $1 \le r < s/2$. Conditional on a fixed S being independent in G_n, the expected number of independent partners with this overlap is

$$\binom{s}{r}\binom{n-s}{r}2^{-r(2s-r-1)/2} \le \frac{b^r}{(r!)^2}.$$

Indeed, the partner requires $r(2s-r-1)/2$ additional absent edges, and $s - (r+1)/2 \ge 3s/4 - 1/2$. Summing over the first set contributes μ_s times this conditional expectation. The large-overlap contribution is therefore at most $\mu_s \sum_{r\ge1} b^r/(r!)^2 \le \mu_s(e^b - 1)$. The two ranges exhaust the distinct pairs; the diagonal $j = s$ is excluded. □

Appendix B. Bounds for the five-size phase gain

This appendix proves the curvature of the phase gain (Lemma 5.1), the four-size comparison (Lemma 5.2) and the two lower bounds for A_5 (Lemmas 5.3 and 5.4). It also records numerical values of A_5 at the endpoints of the phase interval, which are not used in any proof. We use the notation of Section 5: $q = \log 2$, $w_i(\lambda) = \exp(\lambda i - qi^2/2)$ and, for $S \in \{S_+, S_4, S_5\}$, the partition function Z_S, the mean map M_S, the tilt $\lambda_S(T)$ and the dual value $\mathcal{F}_S(T)$. By (2.0), $T_\delta = 1 + 2/q - \delta \in [2/q, 1+2/q] \subset (14/5, 4)$ for $0 \le \delta \le 1$. All arguments are deterministic and involve no limit in n.

B.1. **Proof of the curvature of the phase gain.**

Proof of Lemma 5.1. The Gaussian factor permits termwise differentiation locally in λ. Since $M'_S(\lambda_S(T)) = V_S(T) > 0$, differentiating the inverse mean equation and the dual value gives

$$\lambda'_S(T) = \frac{1}{V_S(T)}, \qquad \mathcal{F}'_S(T) = -\lambda_S(T), \qquad \mathcal{F}''_S(T) = -\frac{1}{V_S(T)}.$$

For the variance comparison, fix T and write p_+ and p_5 for the two probability vectors in (5.1), extending p_5 by zero off S_5. On S_5,

$$\frac{p_5(i)}{p_+(i)} = \frac{Z_{S_+}(\lambda_{S_+}(T))}{Z_{S_5}(\lambda_{S_5}(T))}e^{(\lambda_{S_5}(T)-\lambda_{S_+}(T))i}.$$

Within the five-size support, this ratio is monotone in the index. Thus the indices where $d_i := p_5(i) - p_+(i) > 0$ form a consecutive block $\{j, \dots, k\} \subseteq S_5$. This block is nonempty because p_+ has positive mass outside S_5 and both vectors have total mass one. Put $a = j - 1/2$ and $b = k + 1/2$. Then $(i-a)(i-b)d_i \le 0$ for every $i \in S_+$: the quadratic factor is negative on the block and positive at all other integer indices. The inequality is strict at $i = j$. Both vectors have mass one and mean T, so the constant and linear terms of the quadratic cancel in the sum:

$$V_{S_5}(T) - V_{S_+}(T) = \sum_{i\in S_+} i^2 d_i = \sum_{i\in S_+}(i-a)(i-b)d_i < 0.$$

All sums converge absolutely. The nonstrict inequality is also a standard consequence of the equal-mean relative log-concavity comparison (Yu, 2010, Theorem 12); the quadratic argument above verifies strictness here.

Use $T'_\delta = -1$ in the derivative formulas above. Since $T_\delta \in (2,5)$ on a neighborhood of $[0,1]$, the resulting formulas remain valid at both endpoints. The minimum is attained at an endpoint because the phase gain is strictly concave. □

B.2. **Tilt brackets and tails for the entropy bound.**

Proof of Lemma 5.2. Fix $\delta \in [0,1]$, put $T_0 = 1 + 2/\log 2 - \delta$, and abbreviate

$$\lambda_4 := \lambda_{S_4}(T_0).$$

We compare the two dual values by evaluating the unrestricted dual at the four-size optimizer. First we bound the tilt; these bounds will control the weights omitted from the four-size support. Put $\vartheta = 2^{-1/2}$. At tilt $2\log 2$, the weights w_i, $i = 2,3,4,5$, divided by the first, are $1, \vartheta, 1/4, \vartheta/16$. Hence

$$M_{S_4}(2\log 2) = 2 + \frac{19\vartheta/16 + 1/2}{5/4 + 17\vartheta/16} < \frac{14}{5} < \frac{2}{\log 2} \le T_0, \tag{B.1}$$

because four-fifths of the denominator minus the numerator is $(40 - 27\vartheta)/80 > 0$, and $\log 2 < 7/10$ by (2.0).

At $\lambda = 9\log 2/2$, put $t = 2^{-1/8}$. The identity

$$\frac{9\log 2}{2} i - \frac{\log 2}{2} i^2 = \frac{\log 2}{8}\{81 - (2i-9)^2\}$$

gives

$$w_i(9\log 2/2) = 2^{81/8} t^{(2i-9)^2}.$$

Thus the four weights, in the order $i = 2,3,4,5$, are proportional to

$$t^{25}, \quad t^9, \quad t, \quad t.$$

The numerator of $M_{S_4}(9\log 2/2) - 4$ is therefore

$$t - t^9 - 2t^{25} = t(1 - t^8 - 2t^{24}) = t/4 > 0, \tag{B.2}$$

and the denominator is positive. Hence this mean is greater than four, whereas $T_0 \le 1 + 2/\log 2 < 4$. Since M_{S_4} is strictly increasing, the two mean comparisons give

$$2\log 2 < \lambda_4 < 9\log 2/2. \tag{B.3}$$

For a tilt λ, define the omitted low- and high-deficit ratios

$$R_{\mathrm{low}}(\lambda) := \frac{\sum_{i=-1}^{1} w_i(\lambda)}{Z_{S_4}(\lambda)}, \qquad R_{\mathrm{high}}(\lambda) := \frac{\sum_{i\ge 6} w_i(\lambda)}{Z_{S_4}(\lambda)}.$$

Let $M_A(\lambda)$ denote the mean of i under the weights $w_i(\lambda)$ restricted to an index set A. Differentiation gives

$$\frac{d}{d\lambda}\log R_{\mathrm{low}}(\lambda) = M_{\{-1,0,1\}}(\lambda) - M_{S_4}(\lambda) \le 1 - 2 = -1,$$
$$\frac{d}{d\lambda}\log R_{\mathrm{high}}(\lambda) = M_{\{6,7,\dots\}}(\lambda) - M_{S_4}(\lambda) \ge 6 - 5 = 1.$$

Thus R_{low} decreases and R_{high} increases on $\mathbb{R}$. To bound their sum throughout the tilt interval, we split the interval at the intermediate reference tilt and use the appropriate endpoint bound for each ratio. The calculations below give:

$$R_{\mathrm{low}}(2\log 2) < \frac{51}{100}, \quad R_{\mathrm{high}}(3\log 2) < \frac{1}{50}, \quad R_{\mathrm{low}}(3\log 2) < \frac{3}{25}, \quad R_{\mathrm{high}}(9\log 2/2) < \frac{1}{4}. \tag{B.4}$$

At $2\log 2$,

$$R_{\mathrm{low}}(2\log 2) = \frac{\vartheta + 1/4 + \vartheta/16}{1 + \vartheta + 1/4 + \vartheta/16} < 51/100.$$

At $\lambda = 3\log 2$,

$$w_i(3\log 2) = \exp\left((\log 2)\left(3i - \frac{i^2}{2}\right)\right) = 2^{9/2} 2^{-(i-3)^2/2}.$$

After canceling the common factor $2^{9/2}$, the weight at index i is $\vartheta^{(i-3)^2}$. Thus the four S_4-weights, for $i = 2, 3, 4, 5$, are

$$\vartheta, \quad 1, \quad \vartheta, \quad \vartheta^4,$$

and

$$2^{-9/2} Z_{S_4}(3\log 2) = 1 + 2\vartheta + \vartheta^4 = \frac{5}{4} + 2\vartheta.$$

For the omitted low indices $i = -1, 0, 1$, the weights are

$$(\vartheta^{16}, \vartheta^9, \vartheta^4) = \left(\frac{1}{256}, \frac{\vartheta}{16}, \frac{1}{4}\right).$$

For the omitted high indices $i \geq 6$, they are

$$\vartheta^9 + \vartheta^{16} + \sum_{r \geq 5} \vartheta^{r^2}.$$

The first term in the remaining sum is $\vartheta^{25} = \vartheta/4096$, and the ratio between successive terms is at most $\vartheta^{11} < 1/32$. Therefore

$$\sum_{r \geq 5} \vartheta^{r^2} \leq \frac{\vartheta^{25}}{1 - 1/32} < \frac{1}{3968}.$$

It follows that

$$R_{\text{low}}(3\log 2) = \frac{1/256 + \vartheta/16 + 1/4}{5/4 + 2\vartheta} < \frac{3}{25},$$

and

$$R_{\text{high}}(3\log 2) \leq \frac{\vartheta/16 + 1/256 + 1/3968}{5/4 + 2\vartheta} < \frac{1}{50},$$

where the final inequalities follow from $7/10 < \vartheta < 71/100$.

Finally, at $\lambda = 9\log 2/2$,

$$\frac{w_i(9\log 2/2)}{2^{81/8}} = t^{(2i-9)^2}.$$

After dividing by the common factor t of the $i = 4, 5$ weights, the high indices $i = 6, 7, 8$ contribute t^8, t^{24}, t^{48}. From $i = 8$ onward the ratio between successive terms is at most $1/16$, so the tail beginning with t^{48} is less than $1/60$. The normalized S_4 denominator is $2 + t^8 + t^{24}$. Hence $R_{\text{high}}(9\log 2/2) < 1/4$ follows from

$$3(t^8 + t^{24}) + 4/60 < 2, \tag{B.5}$$

because $t^8 = 1/2$ and $t^{24} = 1/8$.

If $\lambda_4 \leq 3\log 2$, monotonicity and (B.4) give $R_{\text{low}}(\lambda_4) + R_{\text{high}}(\lambda_4) < 53/100$; if $\lambda_4 \geq 3\log 2$, they give $R_{\text{low}}(\lambda_4) + R_{\text{high}}(\lambda_4) < 37/100$. By the dual representation in Section 5,

$$\mathcal{F}_{S_4}(T_0) = \log Z_{S_4}(\lambda_4) - \lambda_4 T_0,$$

whereas evaluating the S_+ dual function at the same parameter gives

$$\mathcal{F}_{S_+}(T_0) \leq \log Z_{S_+}(\lambda_4) - \lambda_4 T_0.$$

Subtracting these two relations yields

$$\begin{aligned} D_4(\delta) &\leq \log \frac{Z_{S_+}(\lambda_4)}{Z_{S_4}(\lambda_4)} \\ &= \log\bigl(1 + R_{\text{low}}(\lambda_4) + R_{\text{high}}(\lambda_4)\bigr) < \log(153/100). \end{aligned}$$

Since S_4 is a subset of S_+, the loss is nonnegative. Subtracting from $\log 2$ completes the proof. □

B.3. **Proof of the uniform entropy bound.**

Proof of Lemma 5.3. Write $\vartheta = 2^{-1/2} < 3/4$. At tilt $2q$, the five weights, after division by their first weight, are

$$1, \quad \vartheta, \quad 1/4, \quad \vartheta/16, \quad 1/256.$$

Consequently

$$M_{S_5}(2q) = 2 + \frac{19\vartheta/16 + 33/64}{321/256 + 17\vartheta/16} < 14/5.$$

Indeed, four-fifths of the denominator minus the numerator is $(39 - 27\vartheta)/80 > 0$. At tilt $4q$, symmetry about deficit four gives $M_{S_5}(4q) = 4$. Since $T_\delta \in [2/q, 1 + 2/q] \subset (14/5, 4)$, strict monotonicity of the mean gives $2q < \lambda_{S_5}(T_\delta) < 4q$.

Let p_4, p_5 be the optimizers at this same mean, extending p_4 by zero at deficit six. For $\lambda_5 := \lambda_{S_5}(T_\delta)$ one has $\log p_5(i) = \lambda_5 i - \frac{q}{2} i^2 - \log Z_{S_5}(\lambda_5)$; since p_4 has mean T_δ and attains $\mathcal{F}_{S_4}(T_\delta)$, this gives

$$\mathcal{F}_{S_5}(T_\delta) - \mathcal{F}_{S_4}(T_\delta) = \sum_{i=2}^{5} p_4(i) \log \frac{p_4(i)}{p_5(i)}.$$

Put $r_i = p_5(i)/(1 - p_5(6))$ for $i = 2, 3, 4, 5$. These numbers sum to one. The inequality $\log x \le x - 1$ gives $\sum_i p_4(i) \log(p_4(i)/r_i) \ge 0$, so the displayed expression is at least $-\log(1 - p_5(6))$. The mass at six increases strictly with the tilt, since its derivative is $p_5(6)(6 - M_{S_5}) > 0$. Hence

$$p_5(6) > \frac{1}{321 + 272\vartheta} > \frac{1}{525}.$$

The unrestricted dual cancels in $A_5 - A_4$, proving the first inequality. Combining it with Lemma 5.2 gives $A_5 > \log(200/153) + \log(525/524) = \log(8750/6681) = \gamma_5$. Finally $S_5 \subset S_+$ gives $D_5 \ge 0$, and $D_5 = q - A_5$ gives its upper bound. □

B.4. **Proof of the sharp bound for the phase gain.**

Proof of Lemma 5.4. For $y > 0$ put $\lambda(y) := q/2 + \log y$. Then $w_i(\lambda(y)) = y^i 2^{-i(i-1)/2}$ for every integer i. Write

$$P(y) := Z_{S_5}(\lambda(y)) = \frac{y^2}{2} + \frac{y^3}{8} + \frac{y^4}{64} + \frac{y^5}{1024} + \frac{y^6}{32768}, \qquad M(y) := M_{S_5}(\lambda(y)).$$

For rational y, both $P(y)$ and $M(y)$ are rational. Moreover, M is increasing, because M_{S_5} and λ are. For $y \le 513/50$ one has $M(y) < 4$ (see Table 1), so $\lambda(y)$ solves the mean equation at $M(y)$ and $\mathcal{F}_{S_5}(M(y)) = \log P(y) - \lambda(y) M(y)$.

The five-size dual. The function $\lambda \mapsto \log Z_{S_5}(\lambda)$ is convex with derivative $M_{S_5}(\lambda)$. Its tangent line at $\lambda(y)$ gives $\log Z_{S_5}(\lambda) - \lambda M(y) \ge \mathcal{F}_{S_5}(M(y))$ for every real λ. Let $y_a < y_b \le 513/50$ and $T = \theta M(y_a) + (1 - \theta) M(y_b)$ with $\theta \in [0, 1]$. Writing $\log Z_{S_5}(\lambda) - \lambda T$ as the same convex combination of the two tangent expressions and taking the infimum over λ gives

$$\mathcal{F}_{S_5}(T) \ge \theta \mathcal{F}_{S_5}(M(y_a)) + (1 - \theta) \mathcal{F}_{S_5}(M(y_b)).$$

(This is the concavity of $\mathcal{F}_{S_5}$ from the proof of Lemma 5.1.)

The unrestricted dual. Let $0 < z \le 16$. For $i \ge 7$, $w_{i+1}(\lambda(z))/w_i(\lambda(z)) = z 2^{-i} \le 1/8$, so the terms with $i \ge 7$ sum to at most $\frac{8}{7} w_7(\lambda(z)) = \frac{8}{7} z^7 2^{-21}$. Hence

$$Z_{S_+}(\lambda(z)) \le U(z) := \frac{1}{2z} + 1 + z + P(z) + \frac{8z^7}{7 \cdot 2^{21}},$$

and evaluating the infimum at $\lambda(z)$ gives $\mathcal{F}_{S_+}(T) \le \log U(z) - \lambda(z) T$.

Reduction to the endpoints. If $T_\delta = \theta M(y_a) + (1 - \theta) M(y_b)$, the two bounds give

$$A_5(\delta) \ge q - \log U(z) + \lambda(z) T_\delta + \theta \mathcal{F}_{S_5}(M(y_a)) + (1 - \theta) \mathcal{F}_{S_5}(M(y_b)).$$

The right-hand side is affine in θ. At $\theta = 1$ it equals $\log\big(2P(y_a)/U(z)\big) + M(y_a) \log(z/y_a)$, and at $\theta = 0$ it is the same expression with y_b in place of y_a. Thus

$$A_5(\delta) - \log \frac{169}{100} \ge \min\{E(y_a, z), E(y_b, z)\}, \qquad E(y, z) := \log \frac{200 P(y)}{169\, U(z)} + M(y) \log \frac{z}{y},$$

for every δ with $M(y_a) \le T_\delta \le M(y_b)$.

Two cells. Use the cells

$$(y_a, y_b, z) = \Big(\frac{1907}{500}, \frac{703}{100}, \frac{527}{100}\Big), \qquad (y_a, y_b, z) = \Big(\frac{703}{100}, \frac{513}{50}, \frac{213}{25}\Big).$$

The endpoint values after this proof show that $M(1907/500) < 2/q$ and $M(513/50) > 1 + 2/q$, and that E is positive at the four endpoint pairs. Since $T_\delta \in [2/q, 1 + 2/q]$ for $\delta \in [0, 1]$ and M is increasing, every T_δ lies in one of the two cells. □

B.5. **Endpoint values for the sharp phase-gain bound.** The values below complete the proof of Lemma 5.4 and use its notation. For rational y and z, the numbers $P(y)$, $M(y)$ and $U(z)$ are rational, so only the logarithms require estimates. Every logarithm below is bounded from below by

$$\log m \geq \log c + 2\sum_{j=0}^{5} \frac{u^{2j+1}}{2j+1} \qquad \text{whenever } m \geq c\,\frac{1+u}{1-u},\ 0 \leq u < 1,$$

where c is the largest element of $\{1/4, 1/2, 1, 2\}$ not exceeding m and $u := (m - c)/(m + c)$. For $c \neq 1$ we use $0.6931471803 < \log 2 < 0.6931471808$; these follow from $\log 2 = 2\sum_{j\geq 0} 3^{-2j-1}/(2j + 1)$ by keeping nine terms and bounding the remaining terms by a geometric series of ratio $1/9$.

Breakpoints. The bounds for $\log 2$ give $2/\log 2 > 2.8853900$ and $1 + 2/\log 2 < 3.8853901$. Direct evaluation gives $M(1907/500) < 2.8853656$ and $M(513/50) > 3.8856023$. This proves $M(1907/500) < 2/q$ and $M(513/50) > 1 + 2/q$.

Endpoint values. Table 1 lists, for each endpoint pair, a lower bound for each logarithm in $E(y, z)$ and a bound for $M(y)$. The bound for $M(y)$ is a lower bound when $\log(z/y) > 0$ and an upper bound when $\log(z/y) < 0$, so the last column, computed from the displayed entries, is a lower bound for $E(y, z)$.

Table 1. Endpoint values for Lemma 5.4.

y	z	bound for $M(y)$	$\log \frac{200P(y)}{169\,U(z)} \geq$	$\log \frac{z}{y} \geq$	$E(y,z) \geq$
1907/500	527/100	≥ 2.8853655	−0.9318854	0.3233518	0.00110
703/100	527/100	≤ 3.4578900	0.9981835	−0.2881564	0.00177
703/100	213/25	≥ 3.4578899	−0.5607627	0.1922296	0.10394
513/50	213/25	≤ 3.8856024	0.8263169	−0.1858365	0.10423

All four lower bounds for E are positive, which completes the proof of Lemma 5.4. The same cells, endpoint inequalities and logarithm bounds are checked by the Lean kernel; see Appendix H.

B.6. **Numerical values at the endpoints of the phase interval.** By Lemma 5.1, the minimum of A_5 on $[0, 1]$ is attained at an endpoint. As a numerical observation that is not used in any proof: solving the saddle-point equations for $\mathcal{F}_{S_+}$ and $\mathcal{F}_{S_5}$ in 40-digit floating-point arithmetic, with the sum over S_+ truncated at $i \leq 399$, gives $A_5(0) \approx 0.659497$ and $A_5(1) \approx 0.525995$. The minimum therefore appears to be $A_5(1)$, and the bound of Lemma 5.4 appears to be within 0.0013 of it.

Appendix C. Rounding the selected profile and its first moment

This appendix completes the proof of Proposition 5.5. It checks that the correction (5.16) keeps every count positive and costs $O(1/k_{\mathrm{co}})$ in the counting exponent, and that the fixed offset in (5.13) leaves the margin needed for (5.19). We use the notation of Section 5, including $p_i^{(n)}$, $\widetilde{k}_i$, e_0, e_1, Φ_n and the constant c_p of (5.15); every threshold is independent of the phase.

Rounding and first-moment estimates in Proposition 5.5. The correction satisfies $\Delta k_2 + \Delta k_3 = -e_0$ and $2\Delta k_2 + 3\Delta k_3 = -e_1$, so it removes both constraint errors by changing $O(1)$ counts. To estimate the resulting loss in the counting exponent, put $\mathbf{k}^* = (k_{\mathrm{co}} p_i^{(n)})_{i=2}^{6}$ and let $\Delta^{\mathrm{tot}} = \mathbf{k} - \mathbf{k}^*$ be the total rounding and correction displacement. If $f_i := \widetilde{k}_i - k_i^*$ for $i = 4, 5, 6$, so that $|f_i| \leq 1$, the two conservation equations give exactly

$$\Delta^{\mathrm{tot}} = (f_4 + 2f_5 + 3f_6,\ -2f_4 - 3f_5 - 4f_6,\ f_4, f_5, f_6).$$

Thus $\|\Delta^{\mathrm{tot}}\|_\infty \leq 9$ and $\|\Delta^{\mathrm{tot}}\|_2^2 \leq 120$. Set $C_\Delta = 9$. The displacement preserves both conservation equations, so it lies in the tangent space of the feasible affine plane.

To apply Taylor's formula, we must keep the segment from the optimizer to the corrected profile inside the positive orthant. By (5.15), $k_i^* \geq c_p k_{\mathrm{co}}$; hence, uniformly in $t \in [0, 1]$ and in the phase,

$$k_i^* + t\Delta_i^{\mathrm{tot}} \geq c_p k_{\mathrm{co}} - C_\Delta \geq \frac{c_p}{2} k_{\mathrm{co}}$$

for all sufficiently large n. In particular the correction preserves positivity. At the exact constrained optimizer the linear term in every tangent direction vanishes, while at a positive vector x the Hessian of the entropy term is $-\operatorname{diag}(1/x_i)$. Its operator norm along the rounding segment is therefore at most $2/(c_p k_{\mathrm{co}})$. Taylor's formula with integral remainder gives

$$|L_{\mathbf{k}^*+\Delta^{\mathrm{tot}}} - L_{\mathbf{k}^*}| \le \frac{1}{2} \sup_{0\le t\le 1} \left\|\nabla^2 L_{\mathbf{k}^*+t\Delta^{\mathrm{tot}}}\right\|_{\mathrm{op}} \|\Delta^{\mathrm{tot}}\|_2^2 = O(1/k_{\mathrm{co}}).$$

The fixed displacement above the root provides the margin: $\Phi_n(r_5^{\mathrm{co}}) = 0$ and $16 \le k_{\mathrm{co}} - r_5^{\mathrm{co}} < 17$. By Equation (3.7), with $S = S_5$ and $c = \log 2$, there is a phase-independent $c_1 > 0$ such that $\Phi_n'(k) \ge c_1(\log n)^2$ throughout this interval for all sufficiently large n. Integration therefore gives

$$L_{S_5}(n, k_{\mathrm{co}}) + (\log 2)k_{\mathrm{co}} = \Phi_n(k_{\mathrm{co}}) \ge 16c_1(\log n)^2.$$

Rounding within the constraint plane loses $O(k_{\mathrm{co}}^{-1})$ in the exponent, and the six Stirling approximations contribute $O(\log n)$ in total. Both errors are smaller than the margin just obtained, so (5.19) holds with $c_Z := 8c_1$ for all sufficiently large n. □

Appendix D. Common whole classes

This appendix proves Lemma 7.1: the common-class weights of Section 7 sum to $1 + o(1)$. It derives exact one-class recurrences for these weights, proves an elementary rate inequality for five consecutive deficits, and then bounds three ranges of selected mass. We use the deficit indexing of Section 7: types are $i \in \{2, \dots, 6\}$ with class sizes $u_i = \alpha - i$ (not the shifted sizes of Sections 8–9), and $\ell = (\ell_i)$ is a selected common subprofile, not $\log\log n$, which we write out. The profile $\mathbf{k}$ is that of Proposition 5.5, so $k_i/k_{\mathrm{co}} \ge c_*$ and $\alpha - n/k_{\mathrm{co}} \le 39/10$. All thresholds and error sequences are deterministic and independent of the phase.

D.1. **Recurrences for the common-class weights.** The weights $D(\ell)$ of (7.3) satisfy two exact one-class recurrences. Whenever $\ell_i < k_i$, one may add a common block of type i. Let e_i denote the ith coordinate vector. Since $n - m = \sum_j u_j(k_j - \ell_j) \ge u_i$, the denominator below is nonzero.

Moreover,

$$\frac{\binom{k_i}{\ell_i+1}^2}{\binom{k_i}{\ell_i}^2} = \frac{(k_i - \ell_i)^2}{(\ell_i + 1)^2},$$

while (7.2) gives

$$\frac{Y^{\mathrm{sgn}}_{\ell+e_i}}{Y^{\mathrm{sgn}}_{\ell}} = \frac{2(n-m)_{u_i}}{u_i!(\ell_i+1)2^{\binom{u_i}{2}}} = \frac{2\mu_{u_i}(n-m)}{\ell_i + 1}.$$

Dividing these two ratios gives

$$\frac{D(\ell + e_i)}{D(\ell)} = \frac{(k_i - \ell_i)^2}{2(\ell_i + 1)\mu_{u_i}(n-m)}. \tag{D.1}$$

At the opposite endpoint, put $h = \mathbf{k} - \ell$, $h_\bullet = \sum_i h_i$, $v = \sum_i u_i h_i$, and $b_h = \sum_i \binom{u_i}{2} h_i$. Define the finite first-moment factor

$$M_h(v) := 2^{h_\bullet} \frac{v!}{\prod_i (u_i!)^{h_i} h_i!} 2^{-b_h}.$$

Exact factorial cancellation gives

$$D(\mathbf{k} - h) = \frac{B(h)}{\mathbb{E}\left[Z^{\mathrm{sgn}}_{\mathbf{k}}\right]}, \qquad B(h) = \left(\prod_i \binom{k_i}{h_i}\right) M_h(v), \tag{D.2}$$

and, whenever $h_i < k_i$,

$$\frac{\binom{k_i}{h_i+1}}{\binom{k_i}{h_i}} = \frac{k_i - h_i}{h_i + 1}, \qquad \frac{M_{h+e_i}(v + u_i)}{M_h(v)} = \frac{2\mu_{u_i}(v + u_i)}{h_i + 1}.$$

Therefore

$$\frac{B(h + e_i)}{B(h)} = \frac{2(k_i - h_i)\mu_{u_i}(v + u_i)}{(h_i + 1)^2}. \tag{D.3}$$

D.2. **A uniform rate for five deficit sizes.** The intermediate range rests on the following elementary inequality; in its application below, R is the residual class fraction, T the mean deficit and $D = I_r - TR$.

Lemma D.1 (A uniform rate for five deficit sizes)**.** *Let* $1/64 \le R \le 1$, $T \le 39/10$, *and* $D \le 4R(1-R)$, $D \le (T-2)(1-R)$. *Then*

$$R\log R + \frac{q}{2}D \le -\frac{1-R}{10000}.$$

Proof. Put $a = 12/25$ and $L_0 = 111176/151959$. With $t = 13/37$,

$$-\log a = 2\int_0^t \frac{du}{1-u^2} > 2(t + t^3/3) = L_0.$$

For $R \le a$, concavity of log on the chord from R to 1 gives $\log R \le (1-R)\log a/(1-a)$. Using $q < 7/10$ and $R \ge 1/64$,

$$R\log R + (q/2)D \le -R(1-R)\left(\frac{L_0}{1-a} - \frac{7}{5}\right) \le -\frac{5287}{48626880}(1-R).$$

For $R \ge a$, convexity of $x\log x$ on $[a,1]$ gives $R\log R \le (1-R)a\log a/(1-a)$. Therefore

$$R\log R + (q/2)D \le -(1-R)\left(\frac{aL_0}{1-a} - \frac{133}{200}\right) = -\frac{104751}{10130600}(1-R).$$

Both rational constants exceed $1/10000$. These arguments include $R = 1$, where the desired bound is zero. □

D.3. **Three-range estimates for common subprofiles.**

Proof of Lemma 7.1. We use the original deficit indexing of Section 7 throughout this proof. Set

$$\eta := \frac{\log\log n}{32\log n}.$$

For all sufficiently large n we have $\eta < 31/32$. Split the selected mass into three disjoint ranges: $m \le \eta n$; $m > \eta n$ with $n - m > n/32$; and $n - m \le n/32$. The empty subprofile in the first range contributes one. We show that every other contribution vanishes in total, using recurrence (D.1) for the first range, a rate estimate for the second, and reverse recurrence (D.3) for the third.

Small selected mass.

Let

$$\xi_i = \frac{k_i^2}{2\mu_{u_i}(n)}, \qquad \Xi_{\text{empty}} = \sum_i \xi_i. \tag{D.4}$$

For $2 \le i \le 6$ one has $u_i \le \alpha - 2$, and (A.4) with the estimate after it gives $\mu_{s-1} > \mu_s$ for $|s-\alpha| \le 6$ and large n; hence $\mu_{u_i}(n) \ge \mu_{\alpha-2}$. With (2.4) and $k_i = \Theta(n/\log n)$ there is a phase-independent C_ξ such that

$$\Xi_{\text{empty}} \le C_\xi (\log n)^{-(2/\log 2 - 1/2)}. \tag{D.5}$$

For $m \le \eta n$, we compare every denominator in the recurrence with the corresponding first moment on the full ambient vertex set.

Fix a final subprofile ℓ with selected mass $m \le \eta n$, and expose its blocks in any order. At an intermediate selected mass $m' \le m$, the exact ratio of first moments is

$$\frac{\mu_{u_i}(n)}{\mu_{u_i}(n-m')} = \frac{(n)_{u_i}}{(n-m')_{u_i}} = \prod_{r=0}^{u_i-1} \frac{n-r}{n-m'-r}.$$

Uniformly in the phase, $u_i = O(\log n)$ whereas $\eta n \gg \log n$. Hence, for all sufficiently large n, $m' + r \le 2\eta n$ throughout this product, and therefore

$$\frac{\mu_{u_i}(n)}{\mu_{u_i}(n-m')} \le (1-2\eta)^{-u_i}.$$

The phase expansion and $\log 2 > 2/3$ give $u_i \le 3\log n$ eventually. Since $\eta \le 1/4$ and $-\log(1-2\eta) \le 4\eta$,

$$(1-2\eta)^{-u_i} \le \exp(4\eta u_i) \le \exp\left(\frac{3}{8}\log\log n\right) = (\log n)^{3/8}. \tag{D.6}$$

Every threshold in this estimate is independent of the phase.

Suppose the current partial profile is a and the next selected block has type i. The exact recurrence (D.1), the inequality $(k_i - a_i)^2 \le k_i^2$, and (D.6) give

$$\frac{D(a+e_i)}{D(a)} \le \frac{(\log n)^{3/8}\xi_i}{a_i + 1}.$$

Multiplying these one-block bounds gives the same estimate for every exposure order:

$$D(\ell) \le \prod_i \frac{\left((\log n)^{3/8}\xi_i\right)^{\ell_i}}{\ell_i!}.$$

All summands are nonnegative. We may therefore sum over every nonnegative integer choice of the five type counts, rather than only the selected-mass range. Expanding the resulting exponential series yields

$$1 \le \sum_{m \le \eta n} D(\ell) \le \exp\Big((\log n)^{3/8}\Xi_{\text{empty}}\Big). \tag{D.7}$$

The lower bound is the empty marked subprofile.

The exponent tends to zero with a fixed power margin. By (D.5),

$$(\log n)^{3/8}\Xi_{\text{empty}} = O\Big((\log n)^{-(2/q-7/8)}\Big),$$

and $q < 7/10$ gives

$$\frac{2}{q} - \frac{7}{8} > \frac{20}{7} - \frac{7}{8} = \frac{111}{56} > 0.$$

Thus, with the phase-independent $\varepsilon_n^{\text{empty}} := C_\xi (\log n)^{-(2/q-7/8)} \to 0$,

$$1 \le \sum_{m \le \eta n} D(\ell) \le e^{\varepsilon_n^{\text{empty}}} = 1 + o(1).$$

Intermediate selected mass.

Use the normalized variables $p_i, y_i, z_i, Y, R, T, I_r$ of Section 7 and put $I := \sum_i i y_i$, so that $I_r = T - I$ and $T = \alpha - n/k_{\text{co}}$. The residual vertex fraction is

$$\rho = \frac{n-m}{n} = R + \frac{I - TY}{\alpha - T}. \tag{D.8}$$

A uniform estimate for the marked weights.

Before approximating any factorial, equations (7.2) and (7.3) give the exact identity

$$\begin{aligned}\log D(\ell) = 2\sum_i \{\log(k_i!) - \log(\ell_i!) - \log((k_i-\ell_i)!)\} \\ - \ell_\bullet \log 2 - \log(n!) + \log((n-m)!) + \sum_i \ell_i \log(u_i!) + \sum_i \log(\ell_i!) + (\log 2) b_\ell.\end{aligned}$$

Some selected or remaining type counts may vanish. We therefore use Stirling's estimate in a form valid also at zero: for an absolute constant $C_S > 0$,

$$\log(t!) = t\log t - t + \sigma(t), \qquad |\sigma(t)| \le C_S \log(t+2) \qquad (t \in \mathbb{N}_0),$$

with the convention $0 \log 0 = 0$. Apply this estimate to the factorials that count vertices and classes: $n!$, $(n-m)!$, $k_i!$, $\ell_i!$, and $(k_i - \ell_i)!$. Keep the class-size factorials $u_i!$ exact, because their coefficients depend on the subprofile. We approximate only a fixed number of factorials, each with argument at most n. Their total remainder is therefore $O(\log n)$ uniformly, even when some ℓ_i or $k_i - \ell_i$ is zero.

Set $K = k_{\text{co}}$, so $k_i = Kp_i$, $\ell_i = Ky_i$, and $k_i - \ell_i = Kz_i$. The profile-multiplicity factorials then contribute

$$\begin{aligned}2\sum_i \log(k_i!) - \sum_i \log(\ell_i!) - 2\sum_i \log((k_i-\ell_i)!) \\ = K\left[2\sum_i p_i \log p_i - \sum_i y_i \log y_i - 2\sum_i z_i \log z_i - Y + Y\log K\right] + O(\log n).\end{aligned}$$

Since $m = K(\alpha Y - I) = K\sum_i u_i y_i$, $n = K(\alpha - T)$, and $n - m = n\rho$, the two ambient factorials contribute

$$-\log(n!) + \log((n-m)!) = n\rho\log\rho + K\sum_i y_i\big(u_i - u_i \log n\big) + O(\log n).$$

Finally, $\ell_\bullet = K\sum_i y_i$ and $b_\ell = K\sum_i y_i \binom{u_i}{2}$, so the class-size and sign terms are exactly

$$K\sum_i y_i \left[\log(u_i!) + (\log 2)\binom{u_i}{2} - \log 2\right].$$

Adding these three displayed contributions and distributing $Y \log K = \sum_i y_i \log K$ gives

$$\begin{aligned} \log D(\ell) &= n\rho \log \rho \\ &\quad + k_{\mathrm{co}} \sum_i \{2p_i \log p_i - 2(p_i - y_i)\log(p_i - y_i) - y_i \log y_i - y_i + y_i E_i\} \\ &\quad + O(\log n), \end{aligned} \tag{D.9}$$

where

$$E_i = \log k_{\mathrm{co}} + \log(u_i!) + u_i - u_i \log n + (\log 2)\binom{u_i}{2} - \log 2. \tag{D.10}$$

The entropy bound must remain valid when a type is entirely selected or entirely unselected. By Proposition 5.5(i), $p_i \geq c_* > 0$. There is therefore an absolute C_{H} such that, for $c_* \leq p \leq 1$ and $0 \leq y \leq p$,

$$2p \log p - 2(p-y)\log(p-y) - y \log y - y \leq C_{\mathrm{H}} y \log\left(\frac{\mathrm{e}}{y}\right). \tag{D.11}$$

Here and below, $y \log(\mathrm{e}/y)$ takes its continuous value 0 at $y = 0$; the analogous convention applies to $Y \log(5\mathrm{e}/Y)$. Indeed, when $y \leq p/2$, the mean-value theorem bounds $p \log p - (p-y)\log(p-y)$ by $C_{\mathrm{H}} y$, while $-y \log y$ supplies the only singular term. When $y \geq p/2$, one has $y \geq c_*/2$, and the assertion follows by compactness. Finally,

$$\sum_i y_i \log\left(\frac{\mathrm{e}}{y_i}\right) \leq Y \log\left(\frac{5\mathrm{e}}{Y}\right),$$

by the entropy bound on five coordinates. The resulting sum is $O(Y \log(\mathrm{e}/Y))$ uniformly.

Since $u_{i+1} = u_i - 1$, exact subtraction in (D.10) gives

$$\begin{aligned} E_{i+1} - E_i &= \log n - \log(\alpha - i) - (\log 2)(\alpha - i) + \log 2 - 1 \\ &= -\frac{\log 2}{2}\alpha + O(1), \end{aligned} \tag{D.12}$$

uniformly for $2 \leq i \leq 5$. For the second line, use $(\log 2)\alpha/2 = \log n - \log\log n + O(1)$ and $\log(\alpha - i) = \log\log n + O(1)$.

The type coefficients vary strongly, but we only need an upper bound for their average over the full profile. The complete first moment gives this bound. Let $\bar{E} = \sum_i p_i E_i$. Taking $\ell = \mathbf{k}$ in (D.9), so that $\rho = 0$, $z = 0$ and $D(\mathbf{k}) = 1/\mathbb{E}\big[Z_{\mathbf{k}}^{\mathrm{sgn}}\big]$, and dividing by k_{co} (the error $O(\log n)/k_{\mathrm{co}}$ is $o(1)$) gives

$$-\frac{1}{k_{\mathrm{co}}} \log \mathbb{E}\big[Z_{\mathbf{k}}^{\mathrm{sgn}}\big] = \sum_i p_i \log p_i - 1 + \bar{E} + o(1). \tag{D.13}$$

The phase-uniform growth estimate (5.19) is eventually positive. Solving (D.13) for $\bar{E}$ and using $-\sum_i p_i \log p_i \leq \log 5$ therefore gives

$$\bar{E} = 1 - \sum_i p_i \log p_i - \frac{1}{k_{\mathrm{co}}} \log \mathbb{E} Z_{\mathbf{k}}^{\mathrm{sgn}} + o(1) \leq 1 + \log 5 + o(1).$$

In particular, $\bar{E} \leq C_E$ for one phase-independent constant. To expose the affine structure in (D.12), set

$$F_i := E_i + \frac{(\log 2)\alpha}{2} i.$$

Then $F_{i+1} - F_i = O(1)$. Because the support has only five consecutive indices,

$$F_i = \sum_j p_j F_j + O(1) = \bar{E} + \frac{(\log 2)\alpha}{2} T + O(1).$$

Consequently

$$E_i = \bar{E} - \frac{(\log 2)\alpha}{2}(i - T) + O(1),$$

and therefore

$$\sum_i y_i E_i \leq \frac{(\log 2)\alpha}{2}(TY - I) + O(Y). \tag{D.14}$$

Only the upper bound on $\bar{E}$ is used here.

The two bounds are expressed in different variables: the ambient term uses the remaining vertex fraction ρ, while the coefficient bound uses the remaining class fraction R. We now compare these fractions. Since $2 \leq i \leq 6$ and $2 \leq T \leq 6$,

$$|I - TY| = \left|\sum_i (i - T) y_i\right| \leq 4Y,$$

so (D.8) gives

$$|\rho - R| \le \frac{4Y}{\alpha - T} = O(Y/\alpha).$$

If $\rho \ge 1/32$, then $Y \le 1$ and the preceding bound implies $R \ge 1/64$ for all sufficiently large n. Hence $x \mapsto x \log x$ is uniformly Lipschitz on the interval between R and ρ. Moreover, $-R \log R \le 1 - R = Y$. Choose a fixed Lipschitz constant $C_{\mathrm{lip}} > 0$ for $x \mapsto x \log x$ on $[1/64, 1]$. Then

$$\begin{aligned} &|n\rho \log \rho - k_{\mathrm{co}} \alpha R \log R| \\ &\quad \le k_{\mathrm{co}}(\alpha - T) C_{\mathrm{lip}} |\rho - R| + k_{\mathrm{co}} T |R \log R| \le (4C_{\mathrm{lip}} + 6) k_{\mathrm{co}} Y, \end{aligned}$$

which proves

$$n\rho \log \rho = k_{\mathrm{co}} \alpha R \log R + O(k_{\mathrm{co}} Y). \tag{D.15}$$

Under the standing condition $\rho \ge 1/32$, combining (D.9), (D.11), (D.14), and (D.15) yields one absolute constant C_{diag} such that

$$\log D(\ell) \le k_{\mathrm{co}} \alpha \Phi_T(z) + C_{\mathrm{diag}} k_{\mathrm{co}} Y \log\left(\frac{\mathrm{e}}{Y}\right) + C_{\mathrm{diag}} \log n, \tag{D.16}$$

where

$$\Phi_T(z) = R \log R + \frac{\log 2}{2}(I_r - TR). \tag{D.17}$$

All constants and thresholds in this bound are independent of the phase.

A negative bound for the leading rate.

For $\rho \ge 1/32$ we have $R \ge 1/64$ eventually, and Proposition 5.5(i) gives $T \le 39/10$. The two inequalities for $I_r - TR$ displayed after Lemma 7.1 and Lemma D.1 therefore give

$$\Phi_T(z) \le -\frac{Y}{10000}. \tag{D.18}$$

Summing the intermediate profiles.

Assume

$$m > \eta n, \qquad n - m > n/32. \tag{D.19}$$

Then $\rho > 1/32$. Moreover,

$$\frac{m}{n} = \frac{\alpha Y - I}{\alpha - T} \le \frac{\alpha Y}{\alpha - T}$$

and $m/n > \eta$ imply that, for all sufficiently large n,

$$Y > \eta \frac{\alpha - T}{\alpha} \ge \frac{\eta}{2}.$$

Define the deterministic error ratio

$$\varepsilon_n^{\mathrm{diag}} := \frac{C_{\mathrm{diag}} \log(2\mathrm{e}/\eta)}{\alpha} + \frac{2C_{\mathrm{diag}} \log n}{k_{\mathrm{co}} \alpha \eta}. \tag{D.20}$$

Since

$$\alpha = \Theta(\log n), \qquad k_{\mathrm{co}} = \Theta(n/\log n), \qquad \eta = \Theta(\log\log n/\log n),$$

one has $\varepsilon_n^{\mathrm{diag}} \to 0$, uniformly in the phase. Because $Y \ge \eta/2$, equations (D.16) and (D.18) give

$$\log D(\ell) \le -k_{\mathrm{co}} \alpha Y \left(\frac{1}{10000} - \varepsilon_n^{\mathrm{diag}}\right).$$

For all sufficiently large n, the parenthesis is at least $1/20000$. Moreover, $\alpha\eta = \Theta(\log\log n)$. Hence there is a phase-independent $c > 0$ such that

$$D(\ell) \le \exp\{-c k_{\mathrm{co}} \log\log n\}. \tag{D.21}$$

Fix phase-independent constants $0 < c_K < C_K$ such that, eventually,

$$c_K \frac{n}{\log n} \le k_{\mathrm{co}} \le C_K \frac{n}{\log n}.$$

Define the phase-independent envelope

$$\varepsilon_n^{\mathrm{central}} := \left(C_K \frac{n}{\log n} + 1\right)^5 \exp\left\{-c c_K \frac{n}{\log n} \log\log n\right\}.$$

There are at most $(k_{\mathrm{co}} + 1)^5$ subprofiles, so the total central-range contribution is at most $\varepsilon_n^{\mathrm{central}}$. Since $n \log\log n/\log n \gg \log n$, this deterministic sequence tends to zero.

Nearly full selected mass.

Put $h = \mathbf{k} - \ell$ and write

$$v(h) := \sum_i u_i h_i = n - m.$$

The nearly full range, also called the full corner, is

$$v(h) \le \frac{n}{32}. \tag{D.22}$$

In this range we use the exact reverse recurrence (D.3), keeping the ambient falling factorials unchanged.

The remaining vertex set has small first moments for each of the five block sizes. For every integer $0 \le w \le n$, the ratio below is zero when $w < u_i$; otherwise

$$\frac{\mu_{u_i}(w)}{\mu_{u_i}(n)} = \frac{(w)_{u_i}}{(n)_{u_i}} = \prod_{r=0}^{u_i-1} \frac{w-r}{n-r} \le \left(\frac{w}{n}\right)^{u_i}.$$

Equations (2.2) and (A.4) give $\mu_{u_i}(n) \le n^{7+o(1)}$ uniformly in the phase. Moreover, $u_i = (2/\log 2 + o(1)) \log n$, and $\log 32 = 5 \log 2$. Consequently, uniformly for $w \le n/32$,

$$\mu_{u_i}(w) \le n^{7+o(1)} 32^{-u_i} = n^{-3+o(1)}. \tag{D.23}$$

In particular, there is one phase-independent threshold after which

$$\mu_{u_i}(w) \le n^{-5/2} \qquad (w \le n/32,\ 2 \le i \le 6). \tag{D.24}$$

We use the weaker exponent $-5/2$ so that the next recurrence bound has a single threshold valid for every phase.

Fix a final residual profile h satisfying (D.22), and expose its blocks in any order. Let a be the current residual profile and $v(a) = \sum_i u_i a_i$. If the next block has type i, then $a_i < h_i$ and therefore

$$v(a) + u_i \le v(h) \le n/32.$$

The exact recurrence (D.3), $k_i - a_i \le k_i \le n$, and (D.24) now give

$$\frac{B(a+e_i)}{B(a)} = \frac{2(k_i - a_i)\mu_{u_i}(v(a)+u_i)}{(a_i+1)^2} \le 2n^{-3/2} < 1$$

for all sufficiently large n. By convention $M_{\mathbf{0}}(0) = 1$, and hence $B(0) = 1$. Thus B never increases along the path from 0 to h, and

$$B(h) \le B(0) = 1. \tag{D.25}$$

By the exact complementary identity (D.2) and the phase-uniform first-moment growth estimate (5.19),

$$D(\mathbf{k} - h) = \frac{B(h)}{\mathbb{E}\left[Z_{\mathbf{k}}^{\mathrm{sgn}}\right]} \le \exp\{-c_Z (\log n)^2\}. \tag{D.26}$$

There are at most $(k_{\mathrm{co}} + 1)^5$ residual profiles, so

$$\sum_{v(h) \le n/32} D(\mathbf{k} - h) \le \varepsilon_n^{\mathrm{full}} := \left(C_K \frac{n}{\log n} + 1\right)^5 e^{-c_Z (\log n)^2} \longrightarrow 0.$$

Summing the three ranges.

Since $\eta < 31/32$, every ℓ with $n - m \le n/32$ has $m > \eta n$; hence the three ranges at the start of the proof partition the box $0 \le \ell_i \le k_i$, and the three bounds give $1 \le \sum_\ell D(\ell) \le e^{\varepsilon_n^{\mathrm{empty}}} + \varepsilon_n^{\mathrm{central}} + \varepsilon_n^{\mathrm{full}} = 1 + o(1)$ with deterministic phase-independent errors. This proves (7.7a), and hence (7.7). □

Appendix E. Large overlap cells

This appendix proves Proposition 8.3 and Lemma 8.4, the two estimates of Section 8 whose proofs are routine calculations. We keep the notation of that section: $U = \alpha - 2$, $R_0 = \lfloor U/2 \rfloor$, the canonical matching M of large cells, the weights $w(M, j)$ and $w_{\mathrm{full}}(M)$, and, for $e \in M$, the full-containment multiplicity m_e, the endpoint-size difference d_e and the deficit h_e, which satisfy $2h_e < m_e$ by (8.1). Types are indexed by $i \in \{0, \dots, 4\}$, with shifted sizes $u_i = \alpha - 2 - i$ and $\kappa_i = k_{i+2}$ classes, and $\alpha > 8$, which holds beyond a phase-independent threshold.

E.1. **Deficit ratios and the ambient denominator.** For endpoint sizes m and $m + d$, define

$$R_{m,d}(h) := \frac{\binom{m}{h}}{(d+1)(d+2)\cdots(d+h)} 2^{-hm + h(h+1)/2}, \qquad R_{m,d}(0) := 1. \tag{E.1}$$

The product in the denominator is empty when $h = 0$.

Lemma E.1 (Exact one-cell deficit ratio)**.** *For endpoint sizes m and $m+d$, define the local numerator*

$$f_{m,d}(j) := \frac{(m)_j (m+d)_j}{j!}\, g(j).$$

For integers $m \geq 1$, $d \geq 0$, and $h \geq 0$ with $2h < m$,

$$\frac{f_{m,d}(m-h)}{f_{m,d}(m)} = R_{m,d}(h).$$

Proof. The ratio of the partial matching counts is

$$\frac{(m)_{m-h}(m+d)_{m-h}}{(m-h)!}\,\frac{m!}{(m)_m (m+d)_m} = \frac{\binom{m}{h}}{(d+1)(d+2)\cdots(d+h)}.$$

The local factors satisfy

$$\frac{g(m-h)}{g(m)} = 2^{-hm+h(h+1)/2}.$$

Multiplying the two identities proves the claim. □

Recall $\mathbf{m}$, $w_{\mathrm{full}}(M)$, J_* and $h_\bullet$ from Section 8.2. After the one-cell ratios have been extracted, the only factor that still couples different cells is the ambient denominator:

$$\frac{(n)_{J_*}}{(n)_J} = (n-J)_{h_\bullet} \leq n^{h_\bullet}. \tag{E.2}$$

Lemma E.2 (Aggregate deficit comparison)**.** *For every $h = (h_e)_{e\in M} \in \mathbb{N}_0^M$ satisfying $2h_e < m_e$ for each $e \in M$,*

$$w(M, \mathbf{m}-h) \leq w_{\mathrm{full}}(M) \prod_{e\in M} n^{h_e} R_{m_e,d_e}(h_e). \tag{E.3}$$

Proof. Apply Lemma E.1 in each selected cell, apply (E.2) once, and use $n^{h_\bullet} = \prod_{e\in M} n^{h_e}$. □

E.2. **Summing all positive deficits.** For a selected cell e, enlarge the exact positive-deficit range to

$$A_e := \{h \in \mathbb{N} : h \geq 1,\ 2h < m_e\}.$$

This range may include deficits that no longer give a large cell. Adding their nonnegative weights only increases the upper bound.

For any nonnegative $\varphi_e(h)$, set $\varphi_e(0) := 1$. Expanding the finite product gives the optional-choice identity

$$\sum_{\omega \in \prod_{e\in M}(\{0\}\cup A_e)} \prod_{e\in M} \varphi_e(\omega_e) = \prod_{e\in M}\left(1 + \sum_{h\in A_e} \varphi_e(h)\right),$$

where each ω_e is either the zero-deficit choice, of weight one, or one positive deficit $h \in A_e$.

For $2h < m$, integrality gives $2h+1 \leq m$. Hence

$$\frac{h+1}{2} \leq \frac{m+1}{4} \leq m - \left\lfloor \frac{3m-1}{4} \right\rfloor.$$

Multiplying by h yields

$$h\left\lfloor \frac{3m-1}{4} \right\rfloor \leq hm - \frac{h(h+1)}{2}. \tag{E.4}$$

Together with $\binom{m}{h} \leq m^h$ and (E.1), this gives

$$n^h R_{m,d}(h) \leq \left(\frac{nm}{2^{\lfloor (3m-1)/4 \rfloor}}\right)^h. \tag{E.5}$$

With the types u_i, κ_i and the quantities ρ_{ij}, ρ_Σ of Section 8.2, every local base in (E.5) is at most ρ_Σ.

Lemma E.3 (Coarse local deficit sum)**.** *Suppose $\rho_\Sigma \leq 1$. Then, for every selected cell e,*

$$\sum_{h\in A_e} n^h R_{m_e,d_e}(h) \leq (\alpha+1)\rho_\Sigma.$$

Proof. If e has endpoint type (i,j), then (E.5) bounds its term of deficit h by ρ_{ij}^h. Since $h \geq 1$ and $\rho_{ij} \leq \rho_\Sigma \leq 1$, this is at most ρ_Σ. Moreover, $|A_e| \leq m_e \leq \alpha + 1$. Summation proves the bound. □

Proof of Proposition 8.3. By (8.1), every admissible h_e lies in $\{0\} \cup A_e$. Use Lemma E.2, enlarge the deficit range to A_e, apply the optional-choice identity above, and then use Lemma E.3. □

E.3. **Proof of the endpoint-table product comparison.**

Proof of Lemma 8.4. Fix the table and divide $W(L)^2$ by $(D(r)A_L)(D(c)C_L)$. The profile multiplicities and table factorials cancel, leaving one ratio for each endpoint pair and one ratio of falling factorials from the matching probabilities. We compute the cell ratios first, then bound the remaining matching ratio.

Fix one endpoint pair and abbreviate $m = m_{ij}$ and $d = d_{ij}$, so its endpoint sizes are m and $m + d$. Its local atom in (8.7) is

$$\frac{(m)_m(m+d)_m}{m!}g(m) = \frac{(m+d)!}{d!}g(m).$$

Consequently

$$\frac{\left[\frac{(m)_m(m+d)_m}{m!}g(m)\right]^2}{m!g(m)\,(m+d)!g(m+d)} = \frac{(m+d)_d}{(d!)^2}2^{-dm-\binom{d}{2}} = \frac{Q_{ij}^2}{(n+1)^d}.$$

Under the hypothesis, $m \geq \alpha - 6 \geq 3$, so in particular $m \geq 2$. Hence we may use $g(m+d)/g(m) = 2^{dm+\binom{d}{2}}$; the identity is also valid for $d = 0$.

Moreover, $m(r), m(c) \geq J(L)$ and

$$m(r) + m(c) - 2J(L) = \sum_{i,j} d_{ij}\ell_{ij}.$$

Therefore, since $(n)_x/(n)_y = (n-y)_{x-y} \leq n^{x-y}$ for integers $0 \leq y \leq x \leq n$,

$$\frac{(n)_{m(r)}(n)_{m(c)}}{(n)_{J(L)}^2} \leq (n+1)^{m(r)+m(c)-2J(L)}.$$

Combining this inequality with the local ratio cell by cell gives

$$\begin{aligned}\frac{W(L)^2}{(D(r)A_L)(D(c)C_L)} &\leq (n+1)^{m(r)+m(c)-2J(L)}\prod_{i,j}\left(\frac{Q_{ij}^2}{(n+1)^{d_{ij}}}\right)^{\ell_{ij}} \\ &= \prod_{i,j} Q_{ij}^{2\ell_{ij}} = (Q^L)^2,\end{aligned}$$

where the equality uses $m(r) + m(c) - 2J(L) = \sum_{i,j} d_{ij}\ell_{ij}$. This proves (8.10). □

Appendix F. Residual factors

This appendix proves the exact decomposition of Proposition 9.1 and the quadratic activity bound of Lemma 9.4. We keep the notation of Sections 6, 8 and 9: the shifted type sizes $u_i = \alpha - 2 - i$, the cap $U = \alpha - 2 \geq 2$, $R_0 = \lfloor U/2 \rfloor$, the uniform configuration matching Π, the feasible pairs (M, j) with $J = \sum_{e\in M} j_e$, the residual mass $m_0 = n - J$ and the event $\mathcal{E}(M, j)$. Lemma 9.4 concerns residual cells outside M when $m_0 > 0$ and $2^U \leq m_0^3$, as in (9.11).

F.1. **Proof of the exact overlap decomposition.**

Proof of Proposition 9.1. Let Π be a full bipartite configuration matching and let $r_{ab}(\Pi)$ be its cell counts. Define

$$M(\Pi) := \{(a,b) : r_{ab}(\Pi) > R_0\}, \qquad j_{ab}(\Pi) := r_{ab}(\Pi) \quad ((a,b) \in M(\Pi)).$$

Since $2(R_0 + 1) > U$, the degree cap makes $M(\Pi)$ a matching. Moreover, $j_{ab} > \lfloor U/2 \rfloor$ and $U \geq 2$ imply $j_{ab} \geq 2$, so every cell of $M(\Pi)$ is an edge of the support graph from Section 6. Let $S(\Pi)$ be the set of all stub pairs of Π lying in cells of $M(\Pi)$, and put $\Pi_{\mathrm{res}} := \Pi \setminus S(\Pi)$. Then

$$\Pi \longmapsto \big(M(\Pi), j(\Pi), S(\Pi), \Pi_{\mathrm{res}}\big)$$

takes values in the feasible pairs (M, j), their labeled realizations, and residual matchings satisfying $\mathcal{E}(M, j)$. It is injective because $\Pi = S(\Pi) \cup \Pi_{\mathrm{res}}$.

Conversely, take a feasible (M, j), an $S \in \mathfrak{F}(M, j)$, and a perfect matching of the unused stubs satisfying $\mathcal{E}(M, j)$. Their disjoint union is a full matching. The condition $r'_{ab} = 0$ on M leaves multiplicity $j_e > R_0$ in each $e \in M$, and the residual cap leaves every cell outside M with multiplicity at most R_0. Thus its canonical large-cell data are exactly (M, j), and this construction is the inverse of the displayed map.

By (6.6) and the configuration-model form of (6.2), $\mathbb{E}Z^2/(\mathbb{E}Z)^2 = \mathbb{E}[A_\zeta(r(\Pi))]$. Under the bijection, $r(\Pi) = j_M + r'$, where r' counts the pairs of Π_{res} and vanishes on M. Hence $E(H(r(\Pi))) = M \cup E(H_{\mathrm{res}})$, and (6.4) factors as

$$A_\zeta(r(\Pi)) = \prod_{e\in M} g(j_e)\cdot\Big(\prod_{a,b} g(r'_{ab})\Big)2^{\beta(M\cup H_{\mathrm{res}})}.$$

For fixed (M, j, S), (9.1) turns the sum over $\Pi_{\mathrm{res}} \in \mathcal{E}(M, j)$ into $(n)_j^{-1} \prod_{e \in M} g(j_e)\, \mathcal{A}(M, j)$; summing over $S \in \mathfrak{F}(M, j)$ gives $w_{\mathrm{hi}}(M, j)\mathcal{A}(M, j)$. The bijection therefore yields (9.3) and counts every full matching once. □

F.2. **Quadratic bounds for residual cell activities.**

Proof of Lemma 9.4. Fix a cell and write $\theta = \theta_{ab}$. If $\theta = 0$, then $q_{ab} = 0$, so the claim holds. Assume $\theta > 0$. If $R_0 < 3$, the higher-reward sum is empty, so $q_{ab} = \theta^2/2$ and the claim is immediate. We may therefore assume $R_0 \geq 3$. To control the contribution from higher multiplicities, use $0 \leq \Delta_x \leq g(x)$ and bound

$$a_x := g(x)\frac{\theta^x}{x!}, \qquad x \geq 3.$$

For $x \geq 3$,

$$\frac{a_{x+1}}{a_x} = \frac{2^x\theta}{x+1}, \qquad \frac{a_{x+2}/a_{x+1}}{a_{x+1}/a_x} = 2\frac{x+1}{x+2} > 1.$$

The successive ratios increase, so (a_x) is log-convex. Its maximum on $[3, R_0]$ is therefore attained at an endpoint, giving

$$\lambda_{ab} \leq R_0(a_3 + a_{R_0}).$$

Now $a_3 = (2/3)\theta^3$. Since $R_0\theta = O(U^3 2^{-U/3})$, one has $R_0 a_3 \leq \theta^2$ for all sufficiently large U.

For the other endpoint, using $R_0 = \lfloor U/2 \rfloor$ and $\theta \leq \mathrm{e}U^2 2^{-U/3}$ gives

$$\log_2 \theta \leq \log_2 \mathrm{e} + 2\log_2 U - \frac{U}{3}.$$

For sufficiently large U, one has $R_0 \geq 3$ and $g(R_0) = 2^{\binom{R_0}{2}-1}$. Hence

$$\begin{aligned}
\log_2\left(\frac{a_{R_0}}{\theta^2}\right) &= \binom{R_0}{2} - 1 + (R_0 - 2)\log_2\theta - \log_2(R_0!) \\
&\leq \left(\frac{U^2}{8} + O(U)\right) - \left(\frac{U^2}{6} + O(U)\right) + O(U\log U) \\
&= -\frac{U^2}{24} + O(U\log U).
\end{aligned}$$

The positive quadratic term comes from the growth of $g(R_0)$; the larger negative term comes from the small intensity in θ^{R_0-2}. The factorial only improves the bound. This estimate is uniform for $0 < \theta \leq \mathrm{e}U^2 2^{-U/3}$. Adding $\log_2 R_0 = O(\log U)$ does not change the leading negative term, so $R_0 a_{R_0} \leq \theta^2$ for all sufficiently large U. Combining the two endpoint bounds gives $\lambda_{ab} \leq 2\theta^2$, and hence $q_{ab} \leq (5/2)\theta^2$.

Choose an integer U_0 beyond which the preceding two endpoint bounds hold. For $1 \leq U < U_0$, define

$$q_U(\theta) := \frac{\theta^2}{2} + \sum_{x=3}^{\lfloor U/2 \rfloor} \Delta_x \frac{\theta^x}{x!}, \qquad \Theta_U := \mathrm{e}U^2 2^{-U/3},$$

where an empty sum is zero. The quotient $q_U(\theta)/\theta^2$ extends continuously to $\theta = 0$ with value $1/2$. Therefore

$$C_{\mathrm{small}} := \max_{1 \leq U < U_0} \max_{0 \leq \theta \leq \Theta_U} \frac{q_U(\theta)}{\theta^2} < \infty.$$

Taking $C_0 = \max\{5/2, C_{\mathrm{small}}\}$ proves the lemma for every U. □

Appendix G. Amplification and failure probabilities

This appendix proves the simultaneous leftover coloring of Lemma 10.1, states and proves an explicit tail form of the amplification step, and uses it to prove Corollary 11.2. The tail form is not needed for the main theorem: it gives the summable failure bound, and the released Lean proof uses it (Appendix H). All thresholds and constants are deterministic and independent of the phase.

G.1. **Proof of the simultaneous leftover coloring.**

Proof of Lemma 10.1. We work in the complement H of G_n, where a clique is an independent set in the original graph. We will find such a clique in every sufficiently large vertex set. Put $u_0 = \lceil n^{1/4} \rceil$. For a fixed u_0-set, (1.5) bounds the probability that its H-edge density is below $1/4$ by $\exp\{-\binom{u_0}{2}/16\}$. There are $\binom{n}{u_0}$ such sets, and for all sufficiently large n,

$$\binom{n}{u_0}\exp\left(-\frac{\binom{u_0}{2}}{16}\right) \leq \exp\left\{u_0\log(\mathrm{e}n/u_0) - \frac{u_0(u_0-1)}{32}\right\} \leq e^{-c\sqrt{n}}$$

for an absolute $c > 0$: the positive term is $O(n^{1/4} \log n)$, whereas $u_0(u_0 - 1) = (1 + o(1))\sqrt{n}$. Thus, outside an event of probability at most $e^{-c\sqrt{n}}$, every u_0-set has density at least $1/4$. For edge counts, the subscript specifies the graph and the argument specifies the vertex set. If S has size $s \geq u_0$, double-counting each edge of $H[S]$ over the u_0-subsets containing it gives

$$\frac{1}{\binom{s}{u_0}} \sum_{\substack{T \subseteq S \\ |T| = u_0}} \frac{e_H(T)}{\binom{u_0}{2}} = \frac{e_H(S)}{\binom{s}{2}}.$$

Therefore every larger set also has density at least $1/4$.

Let S_0 be any vertex set of size at least $n^{1/3}$. While $|S_t| \geq u_0$, choose $v_t \in S_t$ of maximum degree in $H[S_t]$ and retain its neighbors in the current set. Define $S_{t+1} := N_H(v_t) \cap S_t$, and put $s_t := |S_t|$. Density at least $1/4$ implies $|N_H(v_t) \cap S_t| \geq (s_t - 1)/4$. Hence

$$s_{t+1} \geq (s_t - 1)/4, \qquad s_t \geq 4^{-t} s_0 - 1/3. \tag{G.1}$$

The sets are nested and $S_{t+1} \subseteq N_H(v_t)$. Thus, whenever $j > t$,

$$v_j \in S_j \subseteq S_{t+1} \subseteq N_H(v_t).$$

Every later choice is therefore adjacent in H to every earlier one, so the chosen vertices form a clique in H. Since $s_0 \geq \lceil n^{1/3} \rceil$, for every $0 \leq t \leq \lfloor \log n/(13 \log 4) \rfloor$, (G.1) gives

$$s_t \geq n^{1/3 - 1/13} - \frac{1}{3} = n^{10/39} - \frac{1}{3} \geq n^{1/4}$$

for all sufficiently large n. Since s_t is an integer, this gives $s_t \geq u_0$, so all these iterations are permitted. Thus the procedure constructs at least $c_{\mathrm{ind}} \log n$ vertices for an absolute $c_{\mathrm{ind}} > 0$, and these vertices form an independent set in G_n.

On the same density event, take any S and repeatedly remove one of these independent sets. Assign it a new color, and stop when fewer than $n^{1/3}$ vertices remain. Color those vertices singly. Each removed set has at least $c_{\mathrm{ind}} \log n$ vertices, so this uses at most

$$\frac{|S|}{c_{\mathrm{ind}} \log n} + n^{1/3}.$$

After enlarging the absolute constant C_0, this is (10.3), simultaneously for every S. □

G.2. **An explicit tail bound for amplification.** The mean bound in Section 10 gives the high-probability conclusion. Here we retain a free tail parameter in the concentration bound, as needed for the summable failure estimate.

Lemma G.1 (Amplification by vertex exposure)**.** *There are absolute constants $C, c > 0$ and a deterministic sequence $\varepsilon_n^{\mathrm{left}} \geq 0$ such that $\varepsilon_n^{\mathrm{left}} \leq e^{-c\sqrt{n}}$ for all sufficiently large n, and the following holds. Suppose deterministic integers $k_n \geq 0$ and deterministic reals $\Lambda_n \geq 0$ satisfy, for all sufficiently large n,*

$$\mathbb{P}(\zeta(G_n) \leq k_n) \geq e^{-\Lambda_n}. \tag{G.2}$$

For every deterministic choice $r = r(n) > 0$, with the same C and $\varepsilon_n^{\mathrm{left}}$ independent of k_n, Λ_n, r, the bound

$$\mathbb{P}\left(\zeta(G_n) > k_n + C\left(\frac{\sqrt{n\Lambda_n} + \sqrt{nr}}{\log n} + n^{1/3} + 1\right)\right) \leq e^{-r} + \varepsilon_n^{\mathrm{left}}. \tag{G.3}$$

holds at every sufficiently large n for which (G.2) holds.

Proof. Let $\mathcal{D}_n$ be the density event in the proof of Lemma 10.2; on it, (10.3) holds for every vertex set. Set $\varepsilon_n^{\mathrm{left}} := \mathbb{P}(\mathcal{D}_n^c) \leq \varepsilon_n^{\mathrm{dens}} \leq e^{-c\sqrt{n}}$ for all sufficiently large n.

Let S_{k_n} be the maximum cocolorable-set size defined in (10.4). The vertex-exposure argument in the proof of Lemma 10.2 gives (10.5), with $k = k_n$ and $\Lambda = \Lambda_n$.

The lower tail with radius $\sqrt{(n-1)r/2}$ gives, outside an event of probability at most e^{-r},

$$n - S_{k_n} \leq \sqrt{(n-1)\Lambda_n/2} + \sqrt{(n-1)r/2}. \tag{G.4}$$

Choose a maximizing set and its complement V_{left} as in (10.6).

The leftover set V_{left} depends on the graph. The event $\mathcal{D}_n$ nevertheless controls its coloring, because it controls every vertex set simultaneously. Intersect this event with the event in (G.4), and apply (10.3) with $|V_{\text{left}}| = n - S_{k_n}$ in (10.6). After enlarging the absolute constant C, the additional number of parts has the form in (G.3). A union bound gives exceptional probability at most $e^{-r} + \varepsilon_n^{\text{left}}$. □

G.3. **Proof of the summable failure bound.**

Proof of Corollary 11.2. Write $\ell = \log\log n$. Apply Lemma G.1 to (10.2) with $r = 2L$. Since $\Lambda_n = O(n^{3/4}L)$, its additional class allowance

$$t_n := C\left(\frac{\sqrt{n\Lambda_n} + \sqrt{2nL}}{L} + n^{1/3} + 1\right) = O(n^{7/8}L^{-1/2}) = o(n\ell/L^4)$$

satisfies

$$\mathbb{P}(\zeta(G_n) > k_{\text{co}} + t_n) \le n^{-2} + e^{-c\sqrt{n}}.$$

Here the second term is the simultaneous leftover-coloring failure probability in Lemma 10.1.

For the chromatic lower event, the uniform expansion (2.3) gives

$$\log \mu_{\alpha+2} = -L + \beta\ell + (\delta_n - 1)(L - \ell) + O(1), \qquad \mu_{\alpha+2} \le C_{\text{cap}} w_n$$

for a phase-independent constant C_{cap}. In (4.5), the estimates for a, b imply $e^a = O(1)$ and $e^b - 1 = O(L^{5/2}/\sqrt{n})$. Thus

$$\mathbb{P}(\chi(G_n) \le k_\chi^-) \le C\left(w_n^2 + \frac{w_n L^{5/2}}{\sqrt{n}}\right) + e^{-c_1 L^3}.$$

Equations (5.11), (4.1), and (5.13), together with the bound on t_n, give, after one phase-independent threshold,

$$k_\chi^- - k_{\text{co}} - t_n \ge \frac{q^2}{4} A_5(\delta_n)\frac{n}{L^3} - C_1\frac{n\ell}{L^4}.$$

On the chromatic lower event and the cocoloring upper event, the gap is larger than the left-hand side. A union bound, with both exponential terms absorbed into n^{-2}, proves the displayed failure bound. Since $0 \le \delta_n < 1$ and $L/n < 1$,

$$w_n \le \frac{L^\beta}{n}, \qquad w_n^2 + \frac{w_n L^{5/2}}{\sqrt{n}} + n^{-2} = O\left(n^{-3/2}L^{\beta+5/2}\right).$$

This envelope is summable. Finally, for $c < c_5$, the positive margin $c_5 - c \le \frac{q^2}{4}A_5(\delta_n) - c$ absorbs $C_1\ell/L$, proving the assertion for the uniform bound. □

Appendix H. The Lean theorem and reproducibility records

A companion Lean 4 development (Petkov, 2026) kernel-checks a weaker full-sequence consequence of the uniform bound in the main theorem, with formal coefficient $((\log 2)^2/4)\log(200/153)$. The top-level declaration is `Erdos625.erdos625`; its type is `Erdos625.Erdos625Statement`. The following results are proved in this manuscript but lie outside the scope of the released Lean theorem: the stronger five-size coefficient, the phase-resolved refinement involving $A_5(\delta_n)$, the quantified remainder in Remark 11.1, and the failure-probability estimate in Corollary 11.2 together with its almost-sure consequence in Corollary 11.3. The Lean proof uses the tail bound of Lemma G.1. The mean-positive-excess bound in Lemma 10.2 is proved here but is not asserted as a separately released Lean theorem. The formal source revision `824e4b6` recorded in Petkov (2026) fixes Lean and Mathlib at version 4.31.0 and contains the modular sources together with their generated single-file closure. A separate replay revision records the trust-boundary check and a replay in a clean environment; the commands, checksums and transcripts needed to reproduce that replay are recorded in the replay record for the released theorem. A later revision of the same repository, `b654d1c`, also kernel-checks the entropy bounds of Lemmas 5.2, 5.3 and 5.4, including $A_5(\delta) > \gamma_5$ and $A_5(\delta) > \log(169/100)$ for every $\delta \in [0,1]$; the last bound is `Erdos625.FiveSizePhaseGainCertificate.log_169_div_100_lt_phaseGain5`. The released theorem `Erdos625.erdos625` does not use these lemmas, and its coefficient is unchanged.

École normale supérieure, Université PSL, Paris, France; samuil.petkov@ens.psl.eu