

A cubic variance bound for matching deficiency*

Matěj Kripner Václav Rozhoň Robert Šámal
Martin Šindelář Adrián Zámečník

Abstract

Let M be a uniformly random matching of a finite graph, let $D = \nu - \mathbb{E}|M|$ be its expected deficiency from a maximum matching, and let $V = \text{Var}(|M|)$. Kahn showed that D and V tend to infinity together, proved $D = O(V^4)$ when V is bounded below, and conjectured $D = O(V^3)$. We prove $D \leq 41V + 104V^2 + 64V^3$ for every graph, confirming the conjecture, and show that the exponent is attained by connected bipartite graphs with perfect matchings. The proof combines alternating-path identities for monomer probabilities with the sharp estimate $|V(G) - V(G - v)| \leq 1/4$.

1 Introduction

1.1 Background and the source problem

A uniformly chosen matching need not be close to a maximum matching. The quantity $D = \nu - \mathbb{E}|M|$ measures this gap, whereas $V = \text{Var}(|M|)$ measures the fluctuations around the mean. These are different features of a distribution: concentration alone does not generally force its mean to approach the largest possible value. Kahn's normal law for matchings shows that graph structure nevertheless ties them together. The question addressed here is how large the gap can be when the fluctuations are prescribed.

All graphs below are finite and simple; $\text{Vert}(H)$ denotes the vertex set of a graph H . A matching is a set of pairwise vertex-disjoint edges, and the empty matching is included. Write $Z(G)$ for the number of matchings, $\nu(G)$ for the maximum matching size, and

$$\mu(G) = \mathbb{E}|M|, \quad V(G) = \text{Var}(|M|), \quad D(G) = \nu(G) - \mu(G), \quad \sigma(G) = \sqrt{V(G)},$$

where M is uniform on all matchings of G . In the language of the monomer–dimer model, selected edges are dimers and uncovered vertices are monomers. Uniform sampling is the model at dimer activity one.

The connection with normal approximation comes from the real-rootedness theorem of Heilmann and Lieb [2]. The polynomial $Z_G(z) = \sum_M z^{|M|}$ has only negative real zeros. Factoring its normalized version into linear factors represents $|M|$ in distribution as a sum of independent Bernoulli variables. Godsil [3] used this representation to study asymptotic normality; the criterion, restated in [1, Theorem 1.1], is $V(G_n) \rightarrow \infty$. Kahn [1, Theorem 1.6] proved that this is equivalent to $D(G_n) \rightarrow \infty$. His quantitative argument passes through a weighted vertex-cover parameter and gives $D = O(V^4)$ when V is bounded below [1, Theorem 1.10 and discussion preceding (8)]. He proposed the sharper comparison below.

Conjecture 1.1 (Kahn, [1, (8)]). For every fixed $c > 0$ there is a constant C_c such that every finite simple graph with $\sigma(G) \geq c$ satisfies $D(G) \leq C_c \sigma(G)^6$.

*This work was produced within the Bolzano project (<https://bolzano.app>), using AI systems for mathematical research and manuscript preparation. The listed authors contributed to the project and the preparation of this release. The manuscript received only limited human checking.

Kahn's Example 7.2 has $D = \Theta(V^3)$, so the proposed exponent is already necessary among general graphs.

The lower bound on σ is essential. The star $K_{1,m}$ has exactly $m + 1$ matchings: the empty matching and its m single-edge matchings. Its size is Bernoulli with parameter $m/(m + 1)$, giving

$$D = \frac{1}{m + 1}, \quad V = \frac{m}{(m + 1)^2}. \quad (1)$$

Thus D and V have the same small order, while D/V^3 diverges. A uniform estimate must retain a lower-order term.

1.2 Main results

We prove Kahn's conjecture in the following uniform form.

1. Theorem 4.3 gives the uniform inequality $D \leq 41V + 104V^2 + 64V^3$, and hence Conjecture 1.1.
2. Lemma 3.1 gives $|V(G) - V(G - v)| \leq 1/4$, with the same sharp constant at every common positive dimer activity.
3. Theorem 5.4 gives $D = \Theta(V^3)$ with $V \rightarrow \infty$ for connected bipartite graphs having perfect matchings.

The last assertion locates the sharp exponent in a more restricted class than Kahn's original example. The universal deficiency inequality is proved for uniform matchings; the activity-general statement concerns the deletion lemma.

1.3 Related work

Matchings and the monomer–dimer model. Matchings are a central object of combinatorics; see Lovász and Plummer [4] for the structural theory, including the Gallai–Edmonds decomposition, and Godsil [16] for the matching polynomial. Heilmann and Lieb [2] introduced the real-rootedness of the matching polynomial in the study of monomer–dimer systems. Uniform and weighted random matchings are also central algorithmically: Jerrum and Sinclair [5] showed that the monomer–dimer Markov chain mixes rapidly, giving approximate counting and sampling of matchings at every positive activity.

Real-rooted distributions. Harper [6] introduced the method of writing a random variable whose generating polynomial has only real zeros as a sum of independent Bernoulli variables, and used it to prove asymptotic normality of Stirling numbers. Pitman [7] surveys the resulting probabilistic bounds on coefficients, means and modes of such distributions. Michelen and Sahasrabudhe show that central limit theorems persist when the zeros of the generating polynomial merely avoid a neighbourhood of 1 in a suitable sense [8]. The polynomial viewpoint supplies powerful limit theorems, but does not by itself settle the optimal comparison with deficiency. Lebowitz, Pittel, Ruelle and Speer [9, Theorem 3.1] obtain a normal-approximation error of order $V^{-1/2}$ for counting distributions whose generating polynomials have zeros in the closed left half-plane. Their framework includes matchings and more general graph-counting models. Du, Chao and Yeh [10] develop sufficient normality criteria based on coefficients, with applications to matching numbers and Laplacian coefficients. Zhang, Brunetti and Wang [11] give further criteria for graph polynomials, including the Laplacian matching polynomial and characteristic polynomials of positive semidefinite graph matrices. This line of work supplies criteria for normal approximation. The present question is the quantitative comparison between V and the deficiency $\nu - \mu$ in the uniform matching measure.

More precise estimates are available under regularity assumptions. For a d -regular graph on n vertices, Kahn and Kim [12, Theorem 1.1] show, uniformly as $d \rightarrow \infty$, that each monomer probability is asymptotic to $d^{-1/2}$ and that $V \sim n/(4\sqrt{d})$. The situation changes for linear uniform

hypergraphs: Lee [13, Theorem 1.4] disproves the corresponding pointwise prediction for every uniformity at least three. Gupta [14, Theorems 1.1–1.2] gives quantitative bounds for the average vacancy in linear hypergraphs, including an order-uniform $O(\log \log d / \log d)$ bound in the regular case. These results concern degree-dependent occupancy estimates. The bound here applies to arbitrary finite simple graphs and uses only their matching-size variance.

A separate recent development concerns another question in Kahn’s paper. Let F_M be the graph induced by the vertices uncovered by M , and set $\lambda(G) = \mathbb{E}\nu(F_M)$. Zheng et al. [15, Appendix C.7, Theorem C.7.2 and Corollary C.7.3] construct graphs with $V/\lambda \rightarrow \infty$, disproving the universal comparability asked for in Kahn’s Question 7.3. Their discussion also recalls Conjecture 1.1, which involves D instead of λ . The cubic deficiency bound and the failure of the variance–residual-matching comparison concern these distinct parameter pairs.

The proof below also uses classical tools explicitly. Godsil’s path-tree identity [16] transfers a vertex-deletion ratio of matching polynomials to a rooted tree; this identity is restated in [13, Theorem 3.1]. Here its spectral representation yields the variance-deletion estimate in Lemma 3.1. The fixed-matching estimates in Lemma 4.1 are already consequences of Kahn’s Lemma 5.1 and Corollary 4.3. Their proofs are included to keep the argument self-contained.

1.4 Why the cubic bound appears

Write $p_v = \mathbb{P}(v \text{ is uncovered})$ and $h_v = \mu(G) - \mu(G - v)$. Superimposing two independent matchings decomposes their difference into alternating paths and cycles. An odd path contributes one signed unit to the difference of their sizes. Counting these components gives the variance identity and the rooted relation $\sum_v p_v h_v = 2V$ in Lemma 2.1. Truncating an even path at a root, together with the deletion estimate, then gives

$$h_v \geq 1 - (3/2 + 2V)p_v.$$

Consequently, a vertex with sufficiently small monomer probability has $h_v \geq 1/2$: deleting it lowers the mean by a definite amount. The sum identity forces the total monomer mass of all such vertices to be at most $4V$.

For intuition, suppose $V \geq 1$ and split the vertices at a threshold of order V^{-1} . On the low-probability side L , the preceding mass bound and the path covariance identity control the expected number of internal matching edges by $O(V^2)$. On the other side H , the endpoints of every edge of a fixed matching have product probability of order at least V^{-2} . Kahn’s fixed-matching estimate bounds the sum of these products by $4V$, so $\nu(G[H]) = O(V^3)$. Counting covered vertices in L finally gives

$$D \leq \sum_{v \in L} p_v + \mathbb{E}|M \cap E(G[L])| + \nu(G[H]).$$

This directly combines terms of orders V , V^2 and V^3 . Keeping the threshold $1/(3 + 4V)$ instead of its large-variance order produces the stated polynomial for all $V \geq 0$. All monomer probabilities in this argument belong to the original graph, including when a matching is chosen inside $G[H]$.

The sharpness construction uses a different imbalance. Two copies of K_{t^4, t^4+t^3} are joined by t^3 disjoint edges between their larger parts. A perfect matching uses every joining edge, but a uniform matching uses only $O(t)$ of them on average. This leaves deficiency of order t^3 , while the exact conditional distribution gives variance of order t . Theorem 5.4 makes both scales precise.

Finally, the next proposition explains why a linear small-variance term is intrinsic to a uniform reformulation of the conjecture.

Proposition 1.2. *An absolute bound $D(G) \leq CV(G)^3$ for graphs with $V(G) \geq 1$ is equivalent, up to its constant, to an absolute bound $D(G) \leq C'(V(G) + V(G)^3)$ for all graphs.*

Proof. The second bound gives the first with constant $2C'$. Conversely, deficiency and variance are additive under disjoint unions: the matching distribution of a disjoint union is the product of its

component distributions. If $0 < V(G) < 1$, take $k = \lceil 1/V(G) \rceil$ disjoint copies of G . Their total variance lies in $[1, 2)$, and the assumed bound gives

$$kD(G) \leq C(kV(G))^3, \quad D(G) \leq Ck^2V(G)^3 \leq 4CV(G).$$

For $V(G) \geq 1$ the assumed bound already suffices. If $V(G) = 0$, the graph is edgeless: any edge would give both the empty matching and a single-edge matching positive probability. Thus $D(G) = 0$ in the remaining case. Combining the three cases gives the global estimate with $C' = 4C$. $\square$

2 Alternating paths and monomer probabilities

For a path P in G , define

$$q_G(P) = \frac{Z(G - \text{Vert}(P))}{Z(G)}.$$

Paths are simple and identified with their reversals. Odd paths have positive odd length; even paths in sums below have positive even length. For distinct vertices u, v , put

$$p_v = \frac{Z(G - v)}{Z(G)}, \quad w_{uv} = \frac{Z(G - u - v)}{Z(G)}, \quad h_v = \mu(G) - \mu(G - v).$$

If I_v is the indicator that v is unmatched, then $p_v = \mathbb{E}I_v$ and $w_{uv} = \mathbb{E}(I_u I_v)$. Every p_v is positive. If uv is an edge, the operation of adjoining uv is a bijection from matchings of $G - u - v$ to matchings containing uv , so also

$$w_{uv} = \mathbb{P}(uv \in M). \quad (2)$$

Write c_{uv}^+ and c_{uv}^- for the sums of $q_G(P)^2$ over, respectively, odd and positive even paths with endpoints u, v .

Lemma 2.1. *The following identities hold:*

$$V(G) = \sum_{P \text{ odd}} q_G(P)^2, \quad (3)$$

$$w_{uv} - p_u p_v = c_{uv}^+ - c_{uv}^-, \quad (4)$$

$$p_v h_v = \sum_{\substack{P \text{ odd} \\ v \text{ endpoint}}} q_G(P)^2, \quad (5)$$

$$p_v(1 - p_v - h_v) = \sum_{\substack{P \text{ positive even} \\ v \text{ endpoint}}} q_G(P)^2. \quad (6)$$

In particular, $0 \leq h_v \leq 1 - p_v$ and $\sum_v p_v h_v = 2V(G)$.

Proof. Let M_1, M_2 be independent uniform matchings. Their multiset union consists of alternating paths, alternating even cycles, doubled edges and isolated vertices. Conditional on this union, each nontrivial path and each cycle has two equiprobable color assignments, independently of the others. Doubled edges have fixed assignments. For a specified positive-length path P , either assignment fixes the two matchings on its vertices, and outside those vertices the matchings are arbitrary. Therefore the probability that P is a component of the union is $2q_G(P)^2$.

Set $\Delta X = |M_1| - |M_2|$ and $\Delta I_v = I_v(M_1) - I_v(M_2)$. Only odd path components contribute to ΔX . Each contributes an independent symmetric sign, giving $2V(G) = \mathbb{E}(\Delta X)^2 = \mathbb{E}(\text{number of odd path components})$. This proves (3). The difference ΔI_v is nonzero precisely at endpoints of path components. At the two endpoints of an odd path these differences have the same sign, and at those of an even path they have opposite signs. Different components have zero expected product. Thus

$$\text{Cov}(I_u, I_v) = \frac{1}{2} \mathbb{E}(\Delta I_u \Delta I_v) = c_{uv}^+ - c_{uv}^-,$$

which is (4). An odd component with endpoint v contributes -1 to $\Delta X \Delta I_v$; all other contributions have mean zero. Consequently

$$\text{Cov}(|M|, I_v) = - \sum_{P \text{ odd, } v \text{ endpoint}} q_G(P)^2.$$

Conditional on $I_v = 1$, the matching is uniform in $G - v$, so the same covariance equals $p_v(\mu(G - v) - \mu(G)) = -p_v h_v$. Finally, $\text{Var}(I_v) = \frac{1}{2} \mathbb{E}(\Delta I_v)^2$ is the sum over all positive paths with endpoint v . Subtraction proves (6). Nonnegativity gives the bounds on h_v , and summing (5) counts each odd path twice. $\square$

3 Deleting one vertex

The next bound will permit the use of (3) in a smaller graph without losing control of the ambient variance.

Lemma 3.1. *For every graph G and vertex v ,*

$$|V(G) - V(G - v)| \leq \frac{1}{4}.$$

The same inequality holds when matchings in both graphs are sampled with weight $z^{|M|}$, for any common activity $z > 0$.

Proof. Let $Z_H(z) = \sum_M z^{|M|}$ and use $\mu_H(z), V_H(z)$ for the corresponding mean and variance. With $\partial = z d/dz$,

$$\partial \log Z_H(z) = \mu_H(z), \quad \partial \mu_H(z) = V_H(z).$$

Put $p(z) = Z_{G-v}(z)/Z_G(z)$ and $h(z) = \mu_G(z) - \mu_{G-v}(z)$. Then $h = -\partial \log p$ and $\partial h = V_G(z) - V_{G-v}(z)$.

We obtain a useful positive representation of p directly. The matching recurrence at a vertex x of H gives

$$\frac{Z_{H-x}(z)}{Z_H(z)} = \left(1 + z \sum_{y \in N_H(x)} \frac{Z_{H-x-y}(z)}{Z_{H-x}(z)} \right)^{-1}. \quad (7)$$

Construct the rooted tree $T = T(G, v)$ whose vertices are simple paths starting at v , with adjacency defined by extending a path by one vertex. Its root r is the length-zero path. The branch starting through u is $T(G - v, u)$. Induction using (7) and factorization over tree branches proves $p(z) = Z_{T-r}(z)/Z_T(z)$: both root ratios obey the same recurrence, and the single-vertex ratio is 1.

Let A be the adjacency matrix of T . A Schur complement at the root shows that the root entry of $(I + i\sqrt{z}A)^{-1}$ obeys the same recurrence. Indeed, for branch matrices A_j and branch roots r_j , it equals

$$\left(1 + z \sum_j [(I + i\sqrt{z}A_j)^{-1}]_{r_j r_j} \right)^{-1}.$$

These inverses exist because A and the A_j are real symmetric. Induction therefore identifies this entry with $p(z)$. Choose a real orthonormal eigenbasis of A , write its eigenvalues as θ_j , and let a_j be the squared root coordinates of the eigenvectors. Thus $a_j \geq 0$ and $\sum_j a_j = 1$. Taking real parts of the spectral decomposition gives, with $\lambda_j = \theta_j^2$,

$$p(z) = \sum_j \frac{a_j}{1 + z\lambda_j}. \quad (8)$$

Define

$$\alpha_j = \frac{a_j/(1 + z\lambda_j)}{p(z)}, \quad r_j = \frac{z\lambda_j}{1 + z\lambda_j}.$$

The α_j are probabilities and $0 \leq r_j \leq 1$. Differentiating (8) yields

$$h = \sum_j \alpha_j r_j, \quad \partial \alpha_j = \alpha_j(h - r_j), \quad \partial r_j = r_j(1 - r_j).$$

Consequently,

$$\partial h = h + h^2 - 2 \sum_j \alpha_j r_j^2 = h(1 - h) - 2 \text{Var}_\alpha(r).$$

Since $r_j^2 \leq r_j$, its variance lies between 0 and $h(1 - h)$. It follows that $|\partial h| \leq h(1 - h) \leq 1/4$, as required. $\square$

Remark 3.2. The constant $1/4$ in Lemma 3.1 is attained when $G = K_2$: the matching-size variance is $1/4$, and deleting a vertex leaves variance zero. Deletion need not decrease variance. For the two-edge star, $V(K_{1,2}) = 2/9$, whereas deleting a leaf leaves K_2 with variance $1/4$. Thus the sign of the difference cannot be fixed, even for trees.

Lemma 3.3. *For every vertex v ,*

$$h_v \geq 1 - p_v(1 + 2V(G - v)) \geq 1 - (\frac{3}{2} + 2V(G))p_v.$$

Proof. Deleting endpoint v from a positive even path P produces an odd path Q in $G - v$, and $q_G(P) = p_v q_{G-v}(Q)$. Each unoriented Q has at most two preimages, obtained by adjoining v at one of its endpoints. Equations (3) and (6) give

$$p_v(1 - p_v - h_v) \leq 2p_v^2 V(G - v).$$

Divide by $p_v > 0$ and apply Lemma 3.1. $\square$

4 A monomer partition and the deficiency bound

The first inequality in the next lemma is Kahn's fixed-matching bound [1, Lemma 5.1]. Its short conditioning proof is included. The second inequality also follows from his Corollary 4.3; the path proof below keeps the argument within the identities already established here.

Lemma 4.1. *For every fixed matching J in G ,*

$$\sum_{uv \in J} w_{uv} \leq 2V(G), \quad \sum_{uv \in J} p_u p_v \leq 4V(G).$$

Proof. Condition on the exact outside matching $R = M \setminus J$. An edge of J is available if neither endpoint meets R . Every subset of available edges is an equally likely completion, because J is a matching. If $A(R)$ is their number, then $|M| \mid R \sim |R| + \text{Bin}(A(R), 1/2)$, whence $V(G) \geq \mathbb{E}A(R)/4$. An edge $uv \in J$ is available exactly when it is selected or both endpoints are unmatched. Both disjoint events have probability w_{uv} , so $\mathbb{E}A(R) = 2 \sum_{uv \in J} w_{uv}$.

By (4), $p_u p_v \leq w_{uv} + c_{uv}^-$. To bound the even-path terms, orient each edge of J once and for all. For an even path with endpoints forming an oriented edge $u \rightarrow v$ in J , delete v . The resulting path Q is odd and $q_G(P) \leq q_G(Q)$, since every matching avoiding $\text{Vert}(P)$ also avoids $\text{Vert}(Q)$. For a given unoriented Q , choose which endpoint is u . Its partner v is forced by J , and the only possible preimage appends v to the other endpoint of Q . There are at most two such choices. Hence $\sum_{uv \in J} c_{uv}^- \leq 2 \sum_{Q \text{ odd}} q_G(Q)^2 = 2V(G)$. Combining the two estimates proves the second assertion. $\square$

Lemma 4.2. *For any vertex set S , let $a(S) = \sum_{v \in S} p_v$. Then*

$$\mathbb{E}|M \cap E(G[S])| \leq \frac{1}{2}a(S)^2 + V(G).$$

Proof. Equations (2) and (4) give $w_{uv} \leq p_u p_v + c_{uv}^+$ for every edge. The sum of $p_u p_v$ over edges in S is at most $a(S)^2/2$. In the remaining sum each odd path is counted at most once, since it has a unique pair of endpoints. Its total is at most $V(G)$ by (3). $\square$

The set of high-probability vertices can be arbitrarily large: isolated vertices all have monomer probability one. The appropriate quantity to control is its matching number, as supplied by Lemma 4.1.

Theorem 4.3. *Every finite simple graph satisfies*

$$D(G) \leq 41V(G) + 104V(G)^2 + 64V(G)^3 \leq 116(V(G) + V(G)^3). \quad (9)$$

In particular, $D(G) \leq 209\sigma(G)^6$ if $V(G) \geq 1$. More generally, for $\sigma(G) \geq c > 0$,

$$D(G) \leq (41c^{-4} + 104c^{-2} + 64)\sigma(G)^6.$$

Proof. Write $V = V(G)$ and set $B = 3/2 + 2V$, $\tau = 1/(2B)$. Partition the vertices as

$$L = \{v : p_v \leq \tau\}, \quad H = \text{Vert}(G) \setminus L, \quad A = \sum_{v \in L} p_v.$$

By Lemma 3.3, $h_v \geq 1/2$ on L . Lemma 2.1 therefore gives

$$A \leq 2 \sum_{v \in L} p_v h_v \leq 4V. \quad (10)$$

Take a maximum matching J in $G[H]$. Its endpoints have ambient monomer probability exceeding τ , so Lemma 4.1 implies

$$\tau^2 \nu(G[H]) \leq \sum_{uv \in J} p_u p_v \leq 4V, \quad \nu(G[H]) \leq 16B^2 V. \quad (11)$$

All probabilities here are those of G , not of $G[H]$. Write M_{LL}, M_{LH}, M_{HH} for matching edges in the two parts and across them. Lemma 4.2 and (10) give

$$\mathbb{E}|M_{LL}| \leq 8V^2 + V. \quad (12)$$

Every matching has at most $|L|$ edges not entirely in H , hence $\nu(G) \leq |L| + \nu(G[H])$. Counting covered vertices in L also gives

$$|L| - A = 2\mathbb{E}|M_{LL}| + \mathbb{E}|M_{LH}|,$$

and therefore

$$\mu(G) = |L| - A - \mathbb{E}|M_{LL}| + \mathbb{E}|M_{HH}|.$$

Subtracting and using (10)–(12),

$$\begin{aligned} D(G) &\leq A + \mathbb{E}|M_{LL}| + \nu(G[H]) \\ &\leq 4V + (8V^2 + V) + 16(3/2 + 2V)^2 V \\ &= 41V + 104V^2 + 64V^3. \end{aligned}$$

The inequality $2V^2 \leq V + V^3$ gives the second bound in (9). If $V \geq c^2$, then $V \leq c^{-4}V^3$ and $V^2 \leq c^{-2}V^3$, giving the last assertion. No division by V occurred, so edgeless graphs are included. $\square$

5 A connected bipartite sharpness family

We give a quantitative construction to locate the optimal exponent within a more restricted class than required in Conjecture 1.1.

5.1 A one-dimensional variance estimate

Lemma 5.1. *Suppose X is supported on $\{0, \dots, N\}$ with positive probabilities $p_0, \dots, p_N$, and set $B(r) = (r+1)p_{r+1}/p_r$ for $r < N$, $B(N) = 0$. If B is nonincreasing, then $\text{Var } X \leq \mathbb{E}X$. If also $|B(r) - B(s)| \leq L|r - s|$, then $\text{Var } X \geq \mathbb{E}X/(1+L)$.*

Proof. Changing index in finite sums yields $\mathbb{E}B(X) = \mathbb{E}X$ and $\mathbb{E}[XB(X)] = \mathbb{E}[X(X-1)]$, so $\text{Var } X = \mathbb{E}X + \text{Cov}(X, B(X))$. For an independent copy X' ,

$$\text{Cov}(X, B(X)) = \frac{1}{2}\mathbb{E}[(X - X')(B(X) - B(X'))].$$

Monotonicity makes this nonpositive, and the Lipschitz condition bounds it below by $-L \text{Var } X$. Rearrangement proves both assertions. $\square$

Fix $b \geq 1$. In a uniform matching of $K_{b,b+d}$, let R_d count unmatched vertices in the part of size b , and let $r_d = \mathbb{E}R_d$, $v_d = \text{Var } R_d$. Denote the matching count by $Z_b(d)$, with $Z_0(d) = 1$.

Lemma 5.2. *For $d \geq 0$, put $L_d = (b+d+1)/(d+1)^2$. Then*

$$\frac{r_d}{1+L_d} \leq v_d \leq r_d, \quad (13)$$

$$\frac{b}{d+2+b/(d+1)} \leq r_d \leq \frac{b}{d+1}. \quad (14)$$

For $d \geq 1$,

$$r_{d-1} - r_d = \frac{v_d}{d+r_d} \in [0, 1]. \quad (15)$$

Moreover, $h_d = Z_b(d-1)/Z_b(d)$ is nondecreasing in $d \geq 1$ and

$$h_d = \frac{d+r_d}{b+d} = \frac{1}{1+r_{d-1}}. \quad (16)$$

Proof. The number of matchings with $R_d = r$ is

$$a_r(d) = \frac{b!(b+d)!}{r!(b-r)!(d+r)!}, \quad 0 \leq r \leq b.$$

Its birth function is $B_d(r) = (b-r)/(d+r+1)$. This is decreasing, and for $r < b$ its consecutive decrement is $(b+d+1)/((d+r+1)(d+r+2)) \leq L_d$. The same formula at $r = b-1$ includes the endpoint $B_d(b) = 0$. Lemma 5.1 proves (13), and $r_d = \mathbb{E}B_d(R_d) \leq b/(d+1)$. Alternatively changing index gives

$$\mathbb{E}[R_d(d+R_d)] = b - r_d, \quad b = (d+1)r_d + r_d^2 + v_d.$$

Use $v_d \leq r_d$ and $r_d \leq b/(d+1)$ on the right to obtain the lower bound in (14). For $d \geq 1$, $a_r(d-1)/a_r(d) = (d+r)/(b+d)$. Normalizing shows that the law at $d-1$ is the law at d tilted by $(d+R_d)/(d+r_d)$. Its mean differs by $v_d/(d+r_d)$, proving (15). The same coefficient identity proves the first expression in (16). For the second, delete a specified vertex in the larger part of $K_{b,b+d}$. The matching recurrence gives

$$Z_b(d) = Z_b(d-1) + bZ_{b-1}(d).$$

In $K_{b,b+d-1}$ a specified small-part vertex is unmatched with probability r_{d-1}/b , giving $Z_{b-1}(d)/Z_b(d-1) = r_{d-1}/b$. Thus $Z_b(d)/Z_b(d-1) = 1 + r_{d-1}$. Finally (15) says r_d is nonincreasing, so h_d is nondecreasing. $\square$

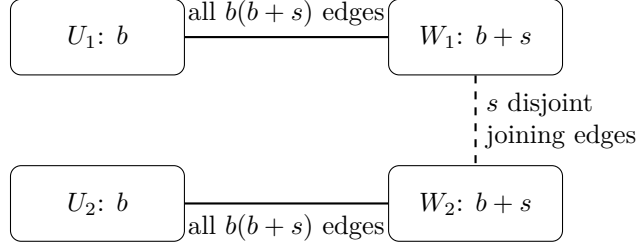


Figure 1: The graph $G_{b,s}$. The boxes represent vertex classes, not single vertices. Its bipartition is $(U_1 \cup W_2, U_2 \cup W_1)$.

5.2 Joining two imbalanced complete bipartite graphs

Take two disjoint copies with parts (U_1, W_1) and (U_2, W_2) of $K_{b,b+s}$, and add s disjoint edges between W_1 and W_2 . Call the resulting graph $G_{b,s}$. For $b, s \geq 1$ it is connected and bipartite, with bipartition $(U_1 \cup W_2, U_2 \cup W_1)$. Figure 1 shows the four vertex classes. Selecting all joining edges leaves a $K_{b,b}$ in each copy, giving a perfect matching and $\nu(G_{b,s}) = 2b + s$.

Proposition 5.3. *Put $A = sh_s^2$, with h_s from (16). Then*

$$s - A \leq D(G_{b,s}) \leq s + \frac{2b}{s+1} + A,$$

$$V(G_{b,s}) \leq \frac{2b}{s+1} + 3A.$$

If K counts selected joining edges and $f(k) = k - 2r_{s-k}$, then

$$V(G_{b,s}) = 2\mathbb{E}v_{s-K} + \text{Var}(f(K)), \quad \text{Var } K \leq \mathbb{E}K \leq A. \quad (17)$$

Proof. Conditional on an exact set of k joining edges, the residual matching is a pair of independent uniform matchings of $K_{b,b+s-k}$. This conditional law depends only on k , so it remains a product law conditional on $K = k$. Moreover,

$$\mathbb{P}(K = k) \propto \binom{s}{k} Z_b(s-k)^2.$$

Its birth function is $(s-k)h_{s-k}^2$ for $k < s$, and zero at $k = s$. It is nonincreasing by Lemma 5.2; Lemma 5.1 gives $\text{Var } K \leq \mathbb{E}K \leq sh_s^2 = A$. Let R_1, R_2 count unmatched vertices in U_1, U_2 . Since $|M| = 2b + K - R_1 - R_2$,

$$D = s - \mathbb{E}K + 2\mathbb{E}r_{s-K}.$$

By (15), $r_s \leq r_{s-k} \leq r_s + k$, giving the asserted deficiency bounds using (14). Conditional variance gives the equality in (17). Also (15) implies $f(k+1) - f(k) = 1 - 2(r_{s-k-1} - r_{s-k}) \in [-1, 1]$. Thus f is 1-Lipschitz and, using an independent copy of K , $\text{Var}(f(K)) \leq \text{Var } K$. By (13), the right side of (17) is at most $2\mathbb{E}r_{s-K} + \text{Var } K \leq 2r_s + 3\mathbb{E}K$, proving the variance bound. $\square$

Theorem 5.4. *For integers $t \geq 2$, let $G_t = G_{t^4, t^3}$. These connected simple bipartite graphs have perfect matchings and satisfy*

$$t^3 - t \leq D(G_t) \leq t^3 + 3t, \quad \frac{t}{6} \leq V(G_t) \leq 5t.$$

Consequently $D(G_t) = \Theta(V(G_t)^3) = \Theta(\sigma(G_t)^6)$. No exponent smaller than six suffices even within this graph class.

Proof. Set $b = t^4$, $s = t^3$. We have $r_s \leq t$, and

$$h_s = \frac{s + r_s}{b + s} \leq \frac{t^3 + t}{t^4 + t^3} \leq \frac{1}{t}.$$

Hence $A = sh_s^2 \leq t$, and Proposition 5.3 yields the deficiency bounds and $V(G_t) \leq 5t$. For the lower bound, $\mathbb{E}K \leq t$ gives $\mathbb{P}(K \leq s/2) \geq 1 - 2/t^2 \geq 1/2$. On this event $d = s - K$ lies in $[s/2, s]$. Lemma 5.2 gives

$$r_d \geq \frac{t^4}{t^3 + 2 + 2t} \geq \frac{t}{2}, \quad L_d \leq \frac{4(t^4 + t^3 + 1)}{t^6} \leq \frac{25}{16} < 2.$$

Here $t^3 \geq 2 + 2t$ for $t \geq 2$, and the upper bound on L_d decreases with t . Thus $v_d \geq r_d/(1 + L_d) \geq t/6$. Equation (17) implies $V(G_t) \geq 2\mathbb{P}(K \leq s/2)t/6 \geq t/6$. For fixed $\alpha < 6$ these estimates give $D(G_t)/\sigma(G_t)^\alpha = \Theta(t^{3-\alpha/2}) \rightarrow \infty$. $\square$

6 Concluding remarks

The lower-order term in (9) is necessary by the star calculation (1). The constants have not been optimized: the argument uses coarse path multiplicities and a fixed probability threshold. The theorem establishes the exponent in Kahn's conjecture and the bipartite construction establishes its sharpness under connectivity and perfect-matching restrictions. It does not assert an analogous cubic bound for arbitrary matching activities; only the vertex-deletion lemma was proved in that generality.

References

- [1] J. Kahn, A normal law for matchings, *Combinatorica* **20** (2000), 339–391. doi:10.1007/PL00009835.
- [2] O. J. Heilmann and E. H. Lieb, Theory of monomer-dimer systems, *Communications in Mathematical Physics* **25** (1972), 190–232. doi:10.1007/BF01877590.
- [3] C. D. Godsil, Matching behaviour is asymptotically normal, *Combinatorica* **1** (1981), 369–376. doi:10.1007/BF02579458.
- [4] L. Lovász and M. D. Plummer, *Matching Theory*, North-Holland Mathematics Studies **121**, North-Holland, Amsterdam, 1986; reprinted by AMS Chelsea Publishing, 2009.
- [5] M. Jerrum and A. Sinclair, Approximating the permanent, *SIAM J. Comput.* **18** (1989), 1149–1178. <https://doi.org/10.1137/0218077>.
- [6] L. H. Harper, Stirling behavior is asymptotically normal, *Ann. Math. Statist.* **38** (1967), 410–414. <https://doi.org/10.1214/aoms/1177698956>.
- [7] J. Pitman, Probabilistic bounds on the coefficients of polynomials with only real zeros, *J. Combin. Theory Ser. A* **77** (1997), 279–303. <https://doi.org/10.1006/jcta.1997.2747>.
- [8] M. Michelen and J. Sahasrabudhe, Central limit theorems and the geometry of polynomials, *J. Eur. Math. Soc.* **28** (2026), no. 5, 2261–2305. <https://doi.org/10.4171/JEMS/1530>.
- [9] J. L. Lebowitz, B. Pittel, D. Ruelle and E. R. Speer, Central limit theorems, Lee–Yang zeros, and graph-counting polynomials, *Journal of Combinatorial Theory, Series A* **141** (2016), 147–183. doi:10.1016/j.jcta.2016.02.009.

- [10] Z. Du, T.-W. Chao and Y.-N. Yeh, Several improved asymptotic normality criteria and their applications to graph polynomials, *Discrete Mathematics* **346** (2023), no. 5, 113303. doi:10.1016/j.disc.2022.113303.
- [11] H. Zhang, M. Brunetti and J. Wang, Asymptotic normality for the coefficients of graph polynomials, *Mediterranean Journal of Mathematics* **23** (2026), 188. doi:10.1007/s00009-026-03186-0.
- [12] J. Kahn and J. H. Kim, Random matchings in regular graphs, *Combinatorica* **18** (1998), 201–226. <https://sites.math.rutgers.edu/~jkahn/RM.pdf>.
- [13] H. Lee, Random matchings in linear hypergraphs, *Forum of Mathematics, Sigma* **14** (2026), e117. doi:10.1017/fms.2026.10253.
- [14] A. Gupta, A rate for the average vacancy of uniformly random matchings in linear hypergraphs, arXiv:2609.26133v1 (2026). <https://arxiv.org/abs/2609.26133>.
- [15] Z. Zheng, S. Zhang, J. Avigad, P. Tetali and S. Welleck, The Problem Is the Problem: Towards Scalable Mathematical Discovery, arXiv:2608.16977v1 (2026), Appendix C.7. <https://arxiv.org/abs/2608.16977>.
- [16] C. D. Godsil, Matchings and walks in graphs, *Journal of Graph Theory* **5** (1981), 285–297. doi:10.1002/jgt.3190050310.