mathdb:2611.00117v1 17 Nov 2026

A ROOTED ENTROPY BOUND FOR PERFECT MATCHINGS

MYEONGJIN KANG

Abstract. Let $H_G(x)$ be the maximum entropy of a perfect-matching law with prescribed edge marginals x in a finite loopless multigraph. We prove that, for every vertex v,

$$H_G(x) \geq -\sum_e x_e \log x_e + \frac{3}{2} \sum_{e \notin \delta(v)} (1 - x_e) \log(1 - x_e).$$

On $2m$ vertices this gives a coefficient $\frac{3}{2}(1 - 1/m)$ in the unrooted complementary-entropy bound. The proof reserves the contribution of one vertex star. This makes the entropy deficit exactly additive under contraction of a tight odd cut and permits induction before a covariance second-variation argument. Boundary marginals and labelled parallel edges are included. We also obtain lower bounds for weighted partition functions and for graphs containing edge-disjoint perfect matchings.

1. Introduction

Let $G = (V, E)$ be a finite undirected loopless multigraph with $|V| = 2m \geq 2$ and at least one perfect matching. Parallel edges have distinct labels. Write $\mathcal{M}(G)$ for the matching family and $P(G) = \operatorname{conv}\{\mathbf{1}_M : M \in \mathcal{M}(G)\}$. For $x \in P(G)$, define

$$H_G(x) = \max\{\operatorname{Ent}(p) : p \text{ is a law on } \mathcal{M}(G),\ \mathbb{E}_p \mathbf{1}_M = x\},$$

$$F_G(x) = -\sum_{e \in E} x_e \log x_e, \qquad B_G(x) = -\sum_{e \in E} (1 - x_e) \log(1 - x_e).$$

All logarithms are natural, with $0 \log 0 = 0$. The maximization in H_G is essential: the conclusions do not concern every law with the given marginals. For $v \in V$, set

$$b_v(x) = -\sum_{e \in \delta(v)} (1 - x_e) \log(1 - x_e), \qquad W_{G,v}(x) = B_G(x) - b_v(x).$$

We omit the subscript G when the graph is understood.

Theorem 1.1. *For every such graph, every $x \in P(G)$ and every $v \in V$,*

$$F_G(x) - H_G(x) \leq \frac{3}{2} W_{G,v}(x). \tag{1.1}$$

Consequently,

$$H_G(x) \geq F_G(x) - \frac{3}{2}(B_G(x) - \max_v b_v(x)) \geq F_G(x) - \frac{3}{2}\left(1 - \frac{1}{m}\right) B_G(x). \tag{1.2}$$

Indeed, $\sum_v b_v = 2B$ implies $\max_v b_v \geq B/m$. There is no restriction on degrees, connectivity or edge multiplicities. When $m = 1$, both sides of (1.1) are zero.

The preprint [4] proves $H \geq F - 2(1-1/m)B$ by combining face dimension with a covariance trace. The coefficient in (1.2) is three quarters of that coefficient for every $m > 1$. The present argument uses the same finite maximum-entropy and covariance framework, but changes the quantity used in the induction. Across a tight odd cut with crossing distribution q, the unrooted deficit $F - cB - H$ acquires a positive term $c[-\sum_e (1 - q_e) \log(1 - q_e)]$. Reserving the star of a root pays this term exactly. A positive global maximum then has no nontrivial tight odd cut, so the degree equations alone give the dimension estimate needed for $c > 3/2$.

Edmonds' theorem [1] supplies the polytope description. Tight-cut contractions also underlie the dimension theory of Edmonds, Lovász and Pulleyblank [2]. In the bipartite setting, Gurvits' weighted Bethe inequality [3, Theorem 2.2], together with entropy duality, gives $H \geq F - B$. That bound has a smaller order-independent coefficient on its subclass. Our assertion concerns arbitrary graphs and retains a root and the finite order. The contraction and covariance ingredients are not new separately; their combination through the additive quantity $W_{G,v}$ proves (1.1).

We give the entropy and contraction arguments below. Apart from the classical matching-polytope theorem, no result of [4] is needed as an unproved premise of Theorem 1.1. We credit that manuscript for the proof framework and the coefficient-two comparison. Our theorem is an existence result for laws at prescribed marginals; it does not supply an efficient sampler.

2. Entropy on a matching face

Edmonds' description, in perfect-matching form, is

$$P(G) = \{x \geq 0 : x(\delta(v)) = 1 \ (v \in V), \quad x(\delta(S)) \geq 1 \ (|S| \text{ odd})\}. \tag{2.1}$$

It applies with distinct parallel-edge coordinates. It can also be deduced from the simple-graph statement by adding coordinates in each parallel class, decomposing the resulting simple-graph mean, and choosing a label conditionally on the selected class. At x, let Q_x be the face obtained by retaining all constraints tight at x, and let T_x be its direction space. The point x is in the relative interior of Q_x: small two-sided perturbations preserving the active constraints preserve every inactive inequality. Every law with mean x is supported on the matching vectors in Q_x, since each active inequality has nonnegative slack on every matching. In particular, if all s edge coordinates are positive and no odd cut with at least three vertices on each side is tight, then

$$\dim T_x \geq s - 2m. \tag{2.2}$$

Singleton cuts and their complements are already degree equations.

Lemma 2.1. *The function H is attained and continuous on $P(G)$. On the matching vectors $\mathcal{A}_x$ of Q_x, its maximizing law is positive and has the form*

$$p_a = \frac{e^{\theta \cdot a}}{\sum_{b \in \mathcal{A}_x} e^{\theta \cdot b}}. \tag{2.3}$$

Its covariance K has image T_x and diagonal $K_{ee} = x_e(1 - x_e)$.

Proof. The fiber of laws with a given mean is a nonempty compact subset of a finite simplex, so entropy attains its maximum. Mixing maximizing laws shows that H is concave. If $y_n \to x$, a convergent subsequence of maximizing laws proves $\limsup H(y_n) \leq H(x)$. For the reverse inequality, put $t = \sqrt{\|y - x\|}$ and $z = x + (y - x)/t$ for $y \neq x$ sufficiently close to x. Then $z \in P(G)$. Active inequalities remain satisfied by linearity, and all inactive inequalities remain satisfied because $z \to x$. Since $y = (1 - t)x + tz$ and $H \geq 0$, concavity gives $H(y) \geq (1 - t)H(x)$. This proves continuity.

Relative interiority supplies a law q of mean x positive on all of $\mathcal{A}_x$: mix a small amount of the uniform law on those atoms with a law whose mean lies just beyond x from their barycenter. If a maximizer p vanished on any face atom, mixing in uq would increase entropy by a positive multiple of $u \log(1/u) + O(u)$. Thus p is positive. The entropy gradient is orthogonal to the kernel of the matrix with columns $(1, a)$. It follows that $\log p_a$ is affine in a, proving (2.3).

For any parameter in (2.3), the resulting law maximizes entropy at its own mean. Indeed, nonnegativity of relative entropy gives $\mathrm{Ent}(q) \leq \log \sum_a e^{\theta \cdot a} - \theta \cdot \mathbb{E}_q a$. The mean remains in the relative interior of the same face. Finally, $w^T K w = \sum_a p_a [w \cdot (a - x)]^2$, so its kernel is $T_x^{\perp}$. The diagonal formula follows because each edge indicator is Boolean. □

mathdb:2611.00117v1 17 Nov 2026

Lemma 2.2. *Suppose $0 < x_e < 1$ for every edge. If $F - cW_v - H$ has a local maximum at x, then, with $d = \dim T_x$ and $s = |E|$,*

$$0 \geq d - s + m + c(m-1). \tag{2.4}$$

Proof. Put $J = F - cW_v$. Its Hessian A is diagonal, with

$$A_{ee} = -\frac{1}{x_e} + \frac{c\mathbf{1}_{\{e \notin \delta(v)\}}}{1 - x_e}.$$

Tilt the law in Lemma 2.1 by $\theta + ta$, and denote its mean by $y(t)$. Then $y'(0) = Ka$, $y''(0) \in T_x$, and

$$H(y(t)) = \log \sum_{b \in \mathcal{A}_x} e^{(\theta + ta)\cdot b} - (\theta + ta) \cdot y(t).$$

For $r(t) = J(y(t)) - H(y(t))$, differentiation gives

$$\begin{aligned} r'(0) &= (\nabla J + \theta) \cdot Ka, \\ r''(0) &= (Ka)^T A(Ka) + a^T Ka + (\nabla J + \theta) \cdot y''(0). \end{aligned}$$

Stationarity for every a and $\operatorname{im} K = T_x$ cancel the last term. For each unit eigenvector of K with positive eigenvalue λ, divide $r''(0) \leq 0$ by λ and then sum. This yields $0 \geq d + \operatorname{tr}(KA)$. No commutation of the two matrices is needed. The degree equations give $\sum_e x_e = m$ and $\sum_{e \notin \delta(v)} x_e = m - 1$, whence

$$\operatorname{tr}(KA) = \sum_e [-(1 - x_e) + c\mathbf{1}_{\{e \notin \delta(v)\}} x_e] = -s + m + c(m-1).$$

□

3. The rooted contraction identity

Write $D_G = F_G - H_G$. Let S be odd with both S and its complement of size at least three, and suppose $x(\delta(S)) = 1$. Choose the side containing the root v to be S. The graph G_1 retains S and contracts its complement; G_2 retains the complement and contracts S to a vertex z. Discard edges internal to a contracted side. Keep every crossing label separately in both graphs. Let x_1, x_2 be the restricted means and put $q = x|_{\delta(S)}$.

Lemma 3.1. *For the preceding contraction,*

$$D_G(x) = D_{G_1}(x_1) + D_{G_2}(x_2), \qquad W_{G,v}(x) = W_{G_1,v}(x_1) + W_{G_2,z}(x_2). \tag{3.1}$$

Proof. A matching has a positive odd number of crossing edges. Under a law of mean x, its expected crossing count is one, so it crosses exactly once almost surely. Its two contractions are matchings sharing a crossing label L with law q. This also proves feasibility of x_1, x_2. Define

$$h(q) = -\sum_e q_e \log q_e, \qquad b(q) = -\sum_e (1 - q_e)\log(1 - q_e).$$

Conditional subadditivity gives $\operatorname{Ent}(M_1, M_2) \leq \operatorname{Ent}(M_1) + \operatorname{Ent}(M_2) - h(q)$. Conversely, take maximizing child laws, draw $L \sim q$, and draw the children independently conditional on L. Joining their common crossing edge gives a parent matching with the required mean. Zero-probability labels need no conditional law. Thus

$$\begin{aligned} H_G &= H_{G_1} + H_{G_2} - h(q), \\ F_G &= F_{G_1} + F_{G_2} - h(q), \qquad\qquad B_G = B_{G_1} + B_{G_2} - b(q). \end{aligned}$$

The first identity in (3.1) follows. The original root star is unchanged in G_1, whereas $b_z(x_2) = b(q)$. Subtracting these star contributions proves the second identity. □

mathdb:2611.00117v1 17 Nov 2026

mathdb:2611.00117v1 17 Nov 2026

4. Small graphs and the induction

Lemma 4.1. *For a probability triple p, put $h(p) = -\sum_i p_i \log p_i$ and $b(p) = -\sum_i (1-p_i)\log(1-p_i)$. Then $h(p) \le 3b(p)/2$.*

Proof. Let $f(u) = -u\log u + \frac{3}{2}(1-u)\log(1-u)$. On the boundary of the probability triangle, $\sum_i f(p_i) = -h(p)/2 \le 0$. In the interior, $f''(u) = -1/u + 3/[2(1-u)]$ changes sign at $2/5$. At a maximum the three first derivatives agree. Two coordinates above $2/5$, or one above and one equal to $2/5$, violate the second-derivative test in their difference direction. Two equal to $2/5$ violate stationarity with the third. If all are at most $2/5$, strict decrease of f' there forces the uniform triple, where the sum is $\log(8/9) < 0$.

The remaining possibility has two equal smaller coordinates and one larger coordinate. Write it as $(t, t, 1-2t)$, $0 < t < 1/3$, and set $g(t) = 2f(t) + f(1-2t)$. With $u = (1-2t)/t \ge 1$,

$$\frac{1}{2}g'(t) = k(u) = \log u - \frac{3}{2}\log\frac{u+1}{2}, \qquad k'(u) = \frac{1-u/2}{u(u+1)}.$$

The function k increases from zero, then decreases to minus infinity, with exactly one further zero. As t increases, g therefore first decreases and then increases. Its maximum on $[0, 1/3]$ is an endpoint value, either 0 or $\log(8/9)$. This proves the lemma. □

Lemma 4.2. *Collapse parallel classes to a simple graph, and let y_a be the sum of x_e in class a. Then $D_G(x) = D_{\mathrm{simple}}(y)$ and $W_{G,v}(x) \ge W_{\mathrm{simple},v}(y)$.*

Proof. When $y_a > 0$, let $r^a_e = x_e/y_a$ be its conditional label law. The chain rule and conditional subadditivity give

$$H_G(x) \le H_{\mathrm{simple}}(y) + \sum_a y_a h(r^a).$$

For clarity, the entropy of the labels conditional on the collapsed matching is at most the sum of their conditional entropies. Concavity then bounds the contribution of class a by $y_a h(r^a)$. Choose a maximizing collapsed law and independently assign labels with these conditional probabilities to attain equality. The same identity holds for F, so the deficits agree. For $\beta(u) = -(1-u)\log(1-u)$, concavity and $\beta(0) = 0$ give $\sum_{e\in a}\beta(x_e) \ge \beta(y_a)$. Each class is entirely inside or outside the root star, proving the assertion for W. Classes of zero mass contribute zero throughout. □

On two vertices $H = F$ and $W_v = 0$. On four vertices use Lemma 4.2. A simple graph has at most the three perfect-matching types of K_4. Their probabilities p_1, p_2, p_3 are fixed by the mean, and each type contains two edges. Thus $H = h(p)$, $F = 2h(p)$, and $W_v = b(p)$ at every root. Missing types have probability zero. Lemma 4.1 proves Theorem 1.1 at these orders.

Proof of Theorem 1.1. Fix $c = 3/2 + \varepsilon$ with $\varepsilon > 0$. We prove $D_G \le cW_{G,v}$ simultaneously for all graphs, feasible means and roots by induction on m. The two small orders were just proved. Suppose $m \ge 3$ and a counterexample exists. By continuity and compactness, $D_G - cW_{G,v}$ has a positive global maximum, attained at some x.

Delete zero-coordinate edges. The point remains a global maximum on the resulting graph, since every mean there extends by zeros to the original graph with the same objective. If $x_e = 1$, that edge belongs to every matching in the law and exhausts both endpoint degrees. Delete it and its endpoints. This preserves D and B. If the root survives, it also preserves W_v, and induction contradicts positivity. If the root is deleted, its star contribution was zero, so $W_v = B$. The rooted bound at any vertex of the remaining graph gives $D \le c(B - b_w) \le cB$, again a contradiction. If nothing remains, $D = B = 0$. Hence every retained coordinate is strictly between zero and one.

If a nontrivial odd cut is tight, both contracted graphs have smaller even order. Induction and Lemma 3.1 give $D_G - cW_{G,v} \le 0$, a contradiction. Thus no such cut is tight at the positive

maximum. Equations (2.2) and (2.4) now imply

$$0 \geq d - s + m + c(m-1) \geq (c-1)m - c = \frac{m-3}{2} + \varepsilon(m-1) > 0.$$

This contradiction completes the induction. For fixed G, x, v, let $\varepsilon \downarrow 0$ to obtain (1.1). The limiting step is needed at $m = 3$, where the dimension estimate alone gives a non-strict inequality at $c = 3/2$. The consequence (1.2) was established in the introduction. □

5. Weighted and counting consequences

For finite real edge weights θ_e, put

$$Z_\theta = \sum_{M \in \mathcal{M}(G)} e^{\theta \cdot \mathbf{1}_M}, \qquad Q_\theta = \max_{x \in P(G)} \{F(x) + \theta \cdot x\}.$$

Corollary 5.1. *For every graph in Theorem 1.1,*

$$\log Z_\theta \geq \max_{x \in P(G)} \{F(x) + \theta \cdot x - \tfrac{3}{2}(B(x) - \max_v b_v(x))\}, \tag{5.1}$$

$$Q_\theta - \tfrac{3}{2}(m-1) \leq \log Z_\theta \leq Q_\theta. \tag{5.2}$$

Proof. The Gibbs variational principle, obtained by comparing any matching law with the law proportional to $e^{\theta \cdot \mathbf{1}_M}$, gives $\log Z_\theta = \max_x (H(x) + \theta \cdot x)$. For any matching law, order the edges and apply the entropy chain rule to each vertex star. Conditioning on all earlier edges can only decrease each term, and every edge lies in two stars. Thus $2\,\mathrm{Ent}(M) \leq \sum_v \mathrm{Ent}(M \cap \delta(v)) = 2F(x)$, proving $H \leq F$. This gives the upper bound; Theorem 1.1 gives (5.1). Finally, $-(1-u)\log(1-u) \leq u$ for $0 \leq u \leq 1$, so $W_v \leq \sum_{e \notin \delta(v)} x_e = m - 1$. Evaluate the rooted lower bound at a maximizer for Q_θ. □

Corollary 5.2. *If G has $k \geq 2$ edge-disjoint perfect matchings, then*

$$|\mathcal{M}(G)| \geq k^m (1 - 1/k)^{\frac{3}{2}(m-1)(k-1)} \geq k^m e^{-\frac{3}{2}(m-1)}. \tag{5.3}$$

The same bounds hold if G is k-regular and every odd cut has at least k edges.

Proof. The uniform mixture of the k disjoint matchings has marginal $1/k$ on their union, with km edges and k edges incident to every vertex. It therefore has $F = m \log k$ and $W_v = (m-1)(k-1)\log(k/(k-1))$. Use $\log|\mathcal{M}(G)| \geq H(x)$ and the rooted bound. The second inequality follows from $\log(1 + 1/(k-1)) \leq 1/(k-1)$. In the regular case the constant vector $x_e = 1/k$ satisfies (2.1), so the same computation applies. □

After removing zero edges and forced matching edges, one may instead apply the theorem to each remaining connected component. Entropy at the prescribed mean is additive over components: subadditivity gives one direction and the product of maximizing laws gives the other. This yields the corresponding sum of rooted bounds and can improve the estimate from the original vertex count.

We do not claim that $3/2$ is optimal. Already on K_4, the uniform mixture of its three matchings has $D = \log 3$ and $W_v = 2\log(3/2)$. Hence any universal coefficient in (1.1) is at least $\log 3/[2\log(3/2)] \approx 1.35476$. The counting corollary is a consequence of the prescribed-mean theorem; it need not improve bounds specialized to bipartite graphs or other restricted families.

Acknowledgements

The author used OpenAI language models for proof development, literature searches, computational checks, and manuscript preparation.

mathdb:2611.00117v1 17 Nov 2026

References

[1] J. Edmonds, Maximum matching and a polyhedron with 0,1-vertices, *J. Res. Nat. Bur. Standards Sect. B* **69B** (1965), 125–130. doi:10.6028/jres.069B.013.

[2] J. Edmonds, W. R. Pulleyblank and L. Lovász, Brick decompositions and the matching rank of graphs, *Combinatorica* **2** (1982), 247–274. doi:10.1007/BF02579233.

[3] L. Gurvits, Unleashing the power of Schrijver's permanental inequality with the help of the Bethe approximation, arXiv:1106.2844v11, 2012.

[4] OpenAI, *Entropy and face dimension of the perfect-matching polytope*, September 23, 2026, preprint in the `openai/math` repository, revision `fd4aeeb2ee4fc729c18d98444fed42fd0529eeeb`, pinned manuscript, Theorem 1.1 and Sections 2–3.

Independent researcher

mathdb:2611.00117v1 17 Nov 2026