Counting Graph Homomorphisms by Entropy Arguments

Igal Sason
Technion – Israel Institute of Technology

EuroComb 2025
Budapest, Hungary
August 25–29, 2025

Counting Graph Homomorphisms in Bipartite Settings

Igal Sason
Technion – Israel Institute of Technology

EuroComb 2025
Budapest, Hungary
August 25–29, 2025

Setup

- F, G finite, undirected graphs without multiple edges; F - no loops.
- A *homomorphism* $\phi\colon\ \mathsf{V}(\mathsf{F}) \to \mathsf{V}(\mathsf{G})$ is adjacency-preserving (not necessarily injective).
- F and G are referred to as the source and target graphs, respectively.
- $\mathrm{Hom}(\mathsf{F}, \mathsf{G})$ = set of homomorphisms $\phi\colon\ \mathsf{V}(\mathsf{F}) \to \mathsf{V}(\mathsf{G})$.
- $\hom(\mathsf{F}, \mathsf{G}) \triangleq |\mathrm{Hom}(\mathsf{F}, \mathsf{G})|$.

The Importance of Counting Graph Homomorphisms

- Large networks & graph limits
- Computer science (CSP)
- Statistical physics (partition functions)

References

1. L. Lovász, *Large Networks and Graph Limits*, AMS, 2012.
2. C. Borgs, J. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi, "Counting graph homomorphisms," *Topics in Discrete Mathematics*, Springer, 2006.
3. Y. Zhao, *Graph Theory and Additive Combinatorics*, CUP, 2023.

Exact Counts: Core Examples

- Star $\mathrm{S}_m = \mathsf{K}_{1,m-1}$:

$$\hom(\mathrm{S}_m, \mathsf{G}) = \sum_{v \in \mathsf{V}(\mathsf{G})} d_{\mathsf{G}}(v)^{m-1}.$$

- Length-m cycle C_m:

$$\hom(\mathsf{C}_m, \mathsf{G}) = \mathrm{tr}(\mathbf{A}_{\mathsf{G}}^m) = \sum_i \lambda_i(\mathsf{G})^m,$$

where $\mathbf{A}$ is the adjacency matrix of G, and the multiset $\{\lambda_i(\mathsf{G})\}$ is its spectrum. Here, $\hom(\mathsf{C}_m, \mathsf{G})$ counts closed walks of length m in G.

- Complete bipartite source and target graphs:

$$\hom(\mathsf{K}_{p,q}, \mathsf{K}_{n_1,n_2}) = n_1^p n_2^q + n_1^q n_2^p.$$

Exact Counts: Core Examples (cont.)

- Additional closed forms highlight the tractability in special cases:

$$\hom(\mathsf{K}_{p,q}, \mathsf{C}_n) = \begin{cases} n(2^p + 2^q - 2), & n = 3 \text{ or } n \geq 5, \\ 2^{p+q+1}, & n = 4, \end{cases}$$

$$\hom(\mathsf{K}_{p,q}, \mathsf{P}_\ell) = 2 + (\ell - 2)(2^p + 2^q - 2).$$

- The tree-walk algorithm is an efficient recursive algorithm for numerically computing the exact number of homomorphisms from a tree to an arbitrary graph (Csikvári and Lin, 2014).

Stirling number of 2nd type

Let $n, k \in \mathbb{N}$ with $k \in [n]$. The Stirling number of the second kind, denoted $S(n,k)$, is defined as the number of ways to partition the set $[n]$ into k nonempty, pairwise disjoint subsets. If $k \notin [n]$, then $S(n,k) \triangleq 0$.

Computation

The Stirling numbers of the second kind satisfy the recurrence relation

$$S(n,k) = k\,S(n-1,k) + S(n-1,k-1), \quad n,k \in \mathbb{N},\ k \leq n,$$

which yields the following closed-form expression:

$$S(n,k) = \frac{1}{k!} \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} j^n, \quad 1 \leq k \leq n.$$

Exact Expression via Stirling Numbers and Bipartite Cliques

Let G be bipartite with partite sets $\mathcal{L}, \mathcal{R}$. For $p, q \in \mathbb{N}$,

$$\hom(\mathsf{K}_{p,q}, \mathsf{G}) = \sum_{k=1}^{p} \sum_{\ell=1}^{q} k!\, \ell!\, S(p,k)\, S(q,\ell)\, \big(N_{k,\ell}(\mathsf{G}) + N_{\ell,k}(\mathsf{G})\big),$$

where

- $S(\cdot,\cdot) =$ Stirling numbers of the second kind.
- $N_{k,\ell}(\mathsf{G}) =$ number of labelled bipartite cliques with k in $\mathcal{L}$ and ℓ in $\mathcal{R}$.

Exact Expression via Stirling Numbers and Bipartite Cliques

Let G be bipartite with partite sets $\mathcal{L}, \mathcal{R}$. For $p, q \in \mathbb{N}$,

$$\hom(\mathsf{K}_{p,q}, \mathsf{G}) = \sum_{k=1}^{p} \sum_{\ell=1}^{q} k!\, \ell!\, S(p,k)\, S(q,\ell)\, \big(N_{k,\ell}(\mathsf{G}) + N_{\ell,k}(\mathsf{G})\big),$$

where

- $S(\cdot,\cdot)$ = Stirling numbers of the second kind.
- $N_{k,\ell}(\mathsf{G})$ = number of labelled bipartite cliques with k in $\mathcal{L}$ and ℓ in $\mathcal{R}$.

Idea

- The complete bipartite graph $\mathsf{K}_{p,q}$ is mapped, by a homomorphism, to a bipartite clique of G. All vertices in the same partite set of $\mathsf{K}_{p,q}$ must be mapped to non-isolated vertices in the same partite set of G.
- The factors $k!\, S(p,k)$ and $\ell!\, S(q,\ell)$ count surjective mappings from the p (resp. q) vertices of $\mathsf{K}_{p,q}$ onto k (resp. ℓ) distinct vertices in the two partite sets of G.

The main difficulty in the exact expression is the computation of the coefficients $\{N_{k,\ell}(\mathsf{G})\}$ for a large bipartite graph G as p and q grow.

The main difficulty in the exact expression is the computation of the coefficients $\{N_{k,\ell}(\mathsf{G})\}$ for a large bipartite graph G as p and q grow.

Combinatorial Lower Bound; Equality for C_4-free Targets

For bipartite G and $p, q \in \mathbb{N}$,

$$\hom(\mathsf{K}_{p,q}, \mathsf{G}) \;\geq\; \sum_{w \in \mathsf{V}(\mathsf{G})} \bigl\{ d(w)^p + d(w)^q \bigr\} \;-\; 2\,|\,\mathsf{E}(\mathsf{G})|,$$

with equality if and only if G is C_4-free.

The main difficulty in the exact expression is the computation of the coefficients $\{N_{k,\ell}(\mathsf{G})\}$ for a large bipartite graph G as p and q grow.

Combinatorial Lower Bound; Equality for C_4-free Targets

For bipartite G and $p, q \in \mathbb{N}$,

$$\hom(\mathsf{K}_{p,q}, \mathsf{G}) \;\geq\; \sum_{w \in \mathsf{V}(\mathsf{G})} \bigl\{ d(w)^p + d(w)^q \bigr\} \;-\; 2\,|\,\mathsf{E}(\mathsf{G})|,$$

with equality if and only if G is C_4-free.

Idea

- This result follows directly from the exact expression for $\hom(\mathsf{K}_{p,q}, \mathsf{G})$ together with the explicit formula for $\hom(\mathrm{S}_m, \mathsf{G})$.
- When G is C_4-free, the lower bound becomes exact.

Entropy-Based Lower Bound (I): Density-Only

Let G be bipartite with partite sizes n_1, n_2 and with an edge density

$$\delta \triangleq \frac{|\mathsf{E}(\mathsf{G})|}{n_1 n_2} \in [0, 1].$$

Then, for all $p, q \in \mathbb{N}$,

$$\begin{aligned} \hom(\mathsf{K}_{p,q}, \mathsf{G}) &\geq \delta^{pq} \left(n_1^p n_2^q + n_1^q n_2^p\right) \\ &= \delta^{pq} \hom(\mathsf{K}_{p,q}, \mathsf{K}_{n_1,n_2}) \end{aligned}$$

Proof Outline

- Let $\mathcal{U}$ and $\mathcal{V}$ denote the partite vertex sets of the bipartite graph G, where $|\mathcal{U}| = n_1$ and $|\mathcal{V}| = n_2$.
- Let (U, V) be a random vector taking values in $\mathcal{U} \times \mathcal{V}$, and assume that $\{U, V\}$ is distributed uniformly at random over the edge set of G.
- The joint entropy of (U, V) is given by

$$\mathrm{H}(U, V) = \log(\delta n_1 n_2).$$

Proof Outline

- Let $\mathcal{U}$ and $\mathcal{V}$ denote the partite vertex sets of the bipartite graph G, where $|\mathcal{U}| = n_1$ and $|\mathcal{V}| = n_2$.
- Let (U, V) be a random vector taking values in $\mathcal{U} \times \mathcal{V}$, and assume that $\{U, V\}$ is distributed uniformly at random over the edge set of G.
- The joint entropy of (U, V) is given by

$$\mathrm{H}(U, V) = \log(\delta n_1 n_2).$$

- Construct random vector $(\mathbf{U}^p, \mathbf{V}^q)$ as follows:
  - 1: The entries of $\mathbf{V}^q \triangleq (V_1, \ldots, V_q)$ are conditionally i.i.d. given U,
  - 2: The entries of $\mathbf{U}^p \triangleq (U_1, \ldots, U_p)$ are conditionally i.i.d. given $\mathbf{V}^q$,

  in a proper way (explicitly given in the paper) such that
  1. $U_i \sim U$ for all $i \in [p]$.
  2. $(U_i, \mathbf{V}^q) \sim (U, \mathbf{V}^q)$ and $(U_i, V_j) \sim (U, V)$ for all $i \in [p]$, $j \in [q]$.

Proof Outline

- Let $\mathcal{U}$ and $\mathcal{V}$ denote the partite vertex sets of the bipartite graph G, where $|\mathcal{U}| = n_1$ and $|\mathcal{V}| = n_2$.
- Let (U, V) be a random vector taking values in $\mathcal{U} \times \mathcal{V}$, and assume that $\{U, V\}$ is distributed uniformly at random over the edge set of G.
- The joint entropy of (U, V) is given by

$$\mathrm{H}(U, V) = \log(\delta n_1 n_2).$$

- Construct random vector $(\mathbf{U}^p, \mathbf{V}^q)$ as follows:
  - 1: The entries of $\mathbf{V}^q \triangleq (V_1, \ldots, V_q)$ are conditionally i.i.d. given U,
  - 2: The entries of $\mathbf{U}^p \triangleq (U_1, \ldots, U_p)$ are conditionally i.i.d. given $\mathbf{V}^q$,

  in a proper way (explicitly given in the paper) such that
  1. $U_i \sim U$ for all $i \in [p]$.
  2. $(U_i, \mathbf{V}^q) \sim (U, \mathbf{V}^q)$ and $(U_i, V_j) \sim (U, V)$ for all $i \in [p]$, $j \in [q]$.
- By the chain rule of the Shannon entropy, it can be verified that

$$\mathrm{H}(\mathbf{U}^p, \mathbf{V}^q) \geq \log(\delta^{pq} n_1^p n_2^q).$$

Proof Outline (cont.)

- Combine it with the uniform bound of the Shannon entropy to the natural partition of $\mathrm{Hom}(\mathsf{K}_{p,q}, \mathsf{G})$ into two disjoint subsets:
  - $\mathcal{H}_1$: homomorphisms where the p-part of $\mathsf{K}_{p,q}$ maps into the n_1-part of G and the q-part of $\mathsf{K}_{p,q}$ maps into the n_2-part of G;
  - $\mathcal{H}_2$: homomorphisms where the p-part of $\mathsf{K}_{p,q}$ maps into the n_2-part of G and the q-part of $\mathsf{K}_{p,q}$ into the n_1-part of G.

Proof Outline (cont.)

- Combine it with the uniform bound of the Shannon entropy to the natural partition of $\mathrm{Hom}(\mathsf{K}_{p,q}, \mathsf{G})$ into two disjoint subsets:
  - $\mathcal{H}_1$: homomorphisms where the p-part of $\mathsf{K}_{p,q}$ maps into the n_1-part of G and the q-part of $\mathsf{K}_{p,q}$ maps into the n_2-part of G;
  - $\mathcal{H}_2$: homomorphisms where the p-part of $\mathsf{K}_{p,q}$ maps into the n_2-part of G and the q-part of $\mathsf{K}_{p,q}$ into the n_1-part of G.
- Overall, this gives

$$\log |\mathcal{H}_1| \geq \mathrm{H}(\mathbf{U}^p, \mathbf{V}^q) \geq \log(\delta^{pq} n_1^p n_2^q).$$

and by symmetry, interchanging p and q (or n_1 and n_2) gives

$$\log |\mathcal{H}_2| \geq \log(\delta^{pq} n_1^q n_2^p).$$

Proof Outline (cont.)

- Combine it with the uniform bound of the Shannon entropy to the natural partition of $\mathrm{Hom}(\mathsf{K}_{p,q}, \mathsf{G})$ into two disjoint subsets:
  - $\mathcal{H}_1$: homomorphisms where the p-part of $\mathsf{K}_{p,q}$ maps into the n_1-part of G and the q-part of $\mathsf{K}_{p,q}$ maps into the n_2-part of G;
  - $\mathcal{H}_2$: homomorphisms where the p-part of $\mathsf{K}_{p,q}$ maps into the n_2-part of G and the q-part of $\mathsf{K}_{p,q}$ into the n_1-part of G.
- Overall, this gives

$$\log |\mathcal{H}_1| \geq \mathrm{H}(\mathbf{U}^p, \mathbf{V}^q) \geq \log(\delta^{pq} n_1^p n_2^q).$$

and by symmetry, interchanging p and q (or n_1 and n_2) gives

$$\log |\mathcal{H}_2| \geq \log(\delta^{pq} n_1^q n_2^p).$$

- Finally,

$$\begin{aligned} \mathrm{hom}(\mathsf{K}_{p,q}, \mathsf{G}) &= |\mathcal{H}_1| + |\mathcal{H}_2| \\ &\geq \delta^{pq}(n_1^p n_2^q + n_1^q n_2^p) = \delta^{pq}\, \mathrm{hom}(\mathsf{K}_{p,q}, \mathsf{K}_{n_1,n_2}). \end{aligned}$$

Comparison to Sidorenko Inequality (for complete bipartite sources)

If G is a bipartite graph with edge density δ that has no isolated vertices, then

$$\hom(\mathsf{K}_{p,q}, \mathsf{G}) \ \geq \ (2\delta)^{pq} \, (n_1 + n_2)^{p+q-2pq} \, (n_1 n_2)^{pq}.$$

Our density-only entropy bound improves this for every bipartite graph G *(proof by invoking Jensen's inequality).*

Refined Entropy-Based Lower Bound (with Degree Profiles)

Let $\mathcal{U}, \mathcal{V}$ be the partite sets ($|\mathcal{U}| = n_1, |\mathcal{V}| = n_2$), and let

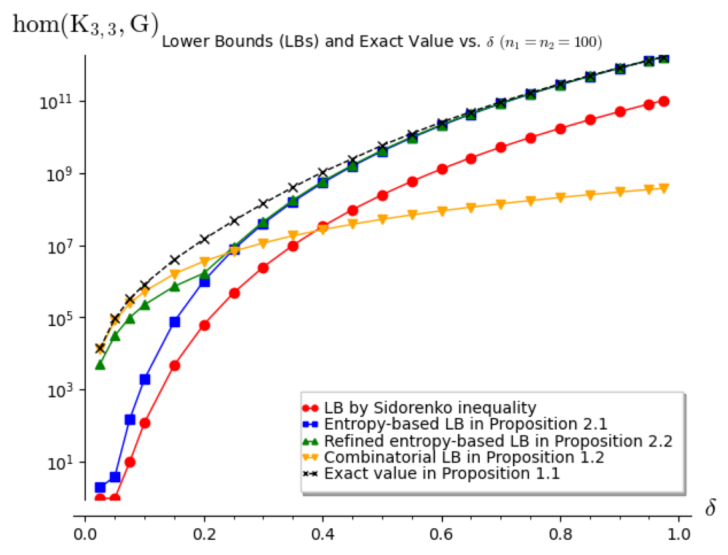
$$x \triangleq \sum_{k=1}^{n_1} \frac{d_k^{(\mathcal{U})}}{|\,\mathsf{E}(\mathsf{G})|} \log\left(\frac{|\,\mathsf{E}(\mathsf{G})|}{d_k^{(\mathcal{U})}}\right) = \mathrm{H}(U), \quad y \triangleq \sum_{k=1}^{n_2} \frac{d_k^{(\mathcal{V})}}{|\,\mathsf{E}(\mathsf{G})|} \log\left(\frac{|\,\mathsf{E}(\mathsf{G})|}{d_k^{(\mathcal{V})}}\right) = \mathrm{H}(V).$$

Then for all $p, q \in \mathbb{N}$,

$$\begin{aligned}
&\hom(\mathsf{K}_{p,q}, \mathsf{G}) \\
&\geq \max\Big\{(\delta n_1 n_2)^{pq} e^{-p(q-1)x - q(p-1)y},\ (\delta n_1 n_2)^{q} e^{-(q-1)x},\ (\delta n_1 n_2)^{p} e^{-(p-1)y}\Big\} \\
&\quad + \max\Big\{(\delta n_1 n_2)^{pq} e^{-q(p-1)x - p(q-1)y},\ (\delta n_1 n_2)^{p} e^{-(p-1)x},\ (\delta n_1 n_2)^{q} e^{-(q-1)y}\Big\}.
\end{aligned}$$

- Captures irregularity via entropy of the degree distributions of the vertices in each partite set; refines and strengthens the density-only lower bound.

Numerical Evidence



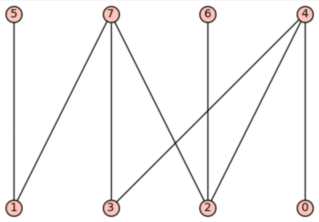
Exact values of $\hom(\mathsf{K}_{3,3}, \mathsf{G})$ vs. lower bounds for random bipartite G.

Beyond Complete Bipartite Sources

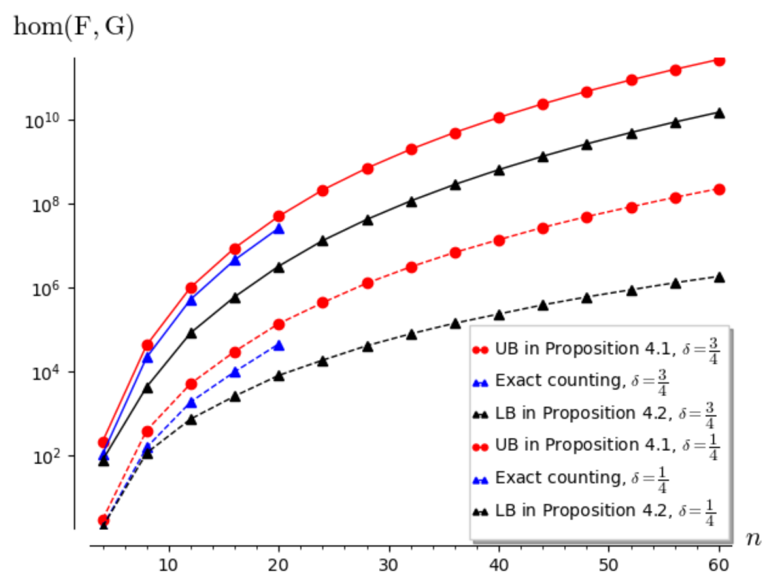
- Extend to general bipartite F $\to$ G: lower bounds derived from the $K_{p,q}$ case + new auxiliary lemmas.
- Upper bounds via reverse Sidorenko (Sah–Sawhney–Stoner–Zhao, 2020) + exact-count expression for complete bipartite sources, including its simplification for C_4-free targets.

Full version: arXiv:2508.06977.

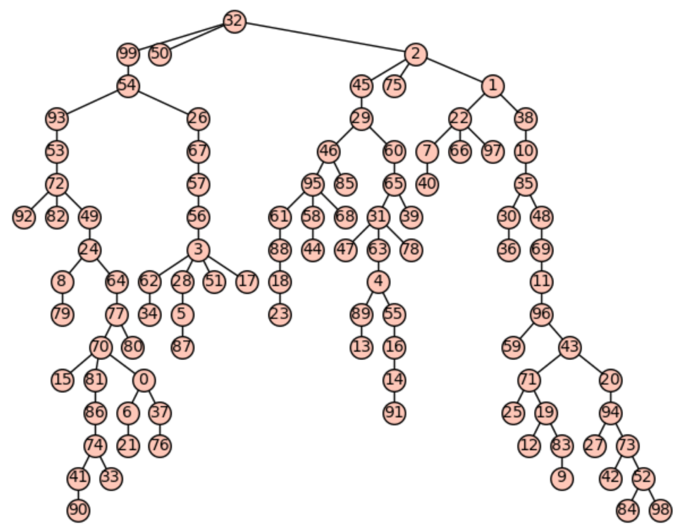
Example 1: Fixed Bipartite Source Graph F



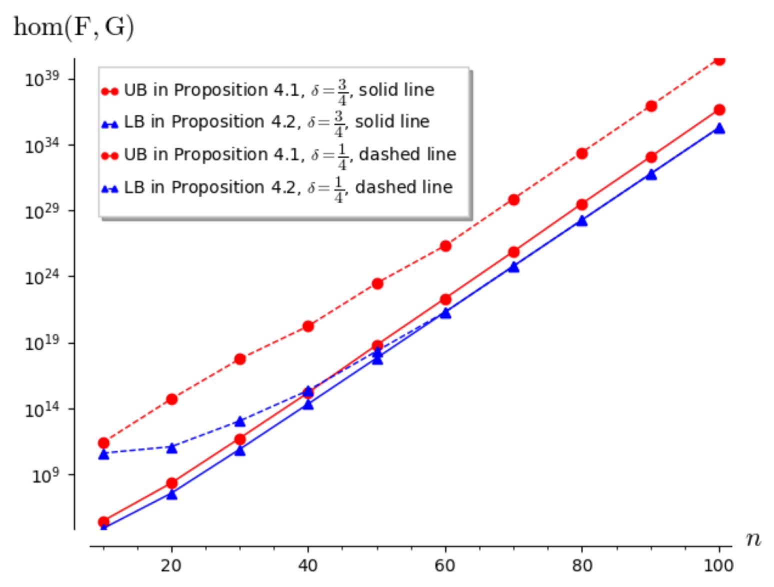
Example 1: Fixed Bipartite Source Graph F



Example 2: Fixed Target Graph G (Tree on 100 vertices)



Example 2: Fixed Target Graph G (Tree on 100 vertices)



Summary

- We study homomorphism counts from bipartite source graphs to bipartite target graphs.
- Combinatorial and two entropy-based lower bounds are derived for complete bipartite source graphs.
- Both entropy-based bounds improve upon the inequality implied by Sidorenko's conjecture for complete bipartite graphs.
- These lower bounds, combined with new auxiliary results, yield general bounds on homomorphism counts between arbitrary bipartite graphs.
- A known reverse Sidorenko inequality (by Sah, Sawhney, Stoner, and Zhao, 2020) is used to derive corresponding upper bounds.
- Numerical comparisons with exact counts in tractable cases support the effectiveness of the proposed computable bounds.
- Full paper version: I.S., "Counting graph homomorphisms in bipartite settings," submitted, August 2025. https://arxiv.org/abs/2508.06977