Entropy-Based Proofs of Combinatorial Results on Bipartite Graphs

Igal Sason

ECE Department, Technion - Israel Institute of Technology,

Haifa, Israel

2021 IEEE International Symposium on Information Theory

(ISIT 2021)

Online Virtual Conference

July 12—20, 2021

Background

The Shannon entropy serves as a powerful tool in various combinatorial and graph-theoretic applications.

Background

The Shannon entropy serves as a powerful tool in various combinatorial and graph-theoretic applications.

These include:

- Enumeration of the independent sets in (regular) graphs;

Background

The Shannon entropy serves as a powerful tool in various combinatorial and graph-theoretic applications.

These include:

- Enumeration of the independent sets in (regular) graphs;
- Enumeration of perfect matchings in bipartite graphs;

Background

The Shannon entropy serves as a powerful tool in various combinatorial and graph-theoretic applications.

These include:

- Enumeration of the independent sets in (regular) graphs;
- Enumeration of perfect matchings in bipartite graphs;
- Moore's bound on the girth (shortest cycle) of a graph;

Background

The Shannon entropy serves as a powerful tool in various combinatorial and graph-theoretic applications.

These include:

- Enumeration of the independent sets in (regular) graphs;
- Enumeration of perfect matchings in bipartite graphs;
- Moore's bound on the girth (shortest cycle) of a graph;
- Minimal number of colors in graph coloring;

Background

The Shannon entropy serves as a powerful tool in various combinatorial and graph-theoretic applications.

These include:

- Enumeration of the independent sets in (regular) graphs;
- Enumeration of perfect matchings in bipartite graphs;
- Moore's bound on the girth (shortest cycle) of a graph;
- Minimal number of colors in graph coloring;
- Number of graph homomorphisms of a fixed graph;

Background

The Shannon entropy serves as a powerful tool in various combinatorial and graph-theoretic applications.

These include:

- Enumeration of the independent sets in (regular) graphs;
- Enumeration of perfect matchings in bipartite graphs;
- Moore's bound on the girth (shortest cycle) of a graph;
- Minimal number of colors in graph coloring;
- Number of graph homomorphisms of a fixed graph;
- Shearer's lemma on hypergraphs and applications;

Background

The Shannon entropy serves as a powerful tool in various combinatorial and graph-theoretic applications.

These include:

- Enumeration of the independent sets in (regular) graphs;
- Enumeration of perfect matchings in bipartite graphs;
- Moore's bound on the girth (shortest cycle) of a graph;
- Minimal number of colors in graph coloring;
- Number of graph homomorphisms of a fixed graph;
- Shearer's lemma on hypergraphs and applications;
- Probabilistic bounds on the number of triangles in a random graph;

Background

The Shannon entropy serves as a powerful tool in various combinatorial and graph-theoretic applications.

These include:

- Enumeration of the independent sets in (regular) graphs;
- Enumeration of perfect matchings in bipartite graphs;
- Moore's bound on the girth (shortest cycle) of a graph;
- Minimal number of colors in graph coloring;
- Number of graph homomorphisms of a fixed graph;
- Shearer's lemma on hypergraphs and applications;
- Probabilistic bounds on the number of triangles in a random graph;
- etc.

Aim of the Present Work

- Properties of bipartite graphs are of great interest in graph theory, combinatorics, modern coding theory, and information theory.
- This work considers new entropy-based proofs of some known, or otherwise refined, combinatorial bounds for bipartite graphs.

Aim of the Present Work

- Properties of bipartite graphs are of great interest in graph theory, combinatorics, modern coding theory, and information theory.
- This work considers new entropy-based proofs of some known, or otherwise refined, combinatorial bounds for bipartite graphs.

These bounds include:

1. Upper bounds on the number of the independent sets of irregular graphs;

Aim of the Present Work

- Properties of bipartite graphs are of great interest in graph theory, combinatorics, modern coding theory, and information theory.
- This work considers new entropy-based proofs of some known, or otherwise refined, combinatorial bounds for bipartite graphs.

These bounds include:

1. Upper bounds on the number of the independent sets of irregular graphs;
2. Lower bounds on the minimal number of colors in a constrained edge coloring;

Aim of the Present Work

- Properties of bipartite graphs are of great interest in graph theory, combinatorics, modern coding theory, and information theory.
- This work considers new entropy-based proofs of some known, or otherwise refined, combinatorial bounds for bipartite graphs.

These bounds include:

1. Upper bounds on the number of the independent sets of irregular graphs;
2. Lower bounds on the minimal number of colors in a constrained edge coloring;
3. Lower bounds on the number of walks of a given length in bipartite graphs.

Bipartite Graph

A graph is called bipartite if it has two types of vertices, and an edge $e \in \mathsf{E}(G)$ cannot connect two vertices of the same type; we refer to the vertices $v \in \mathsf{V}(G)$ of a bipartite graph G as left and right vertices.

Bipartite Graph

A graph is called bipartite if it has two types of vertices, and an edge $e \in \mathsf{E}(G)$ cannot connect two vertices of the same type; we refer to the vertices $v \in \mathsf{V}(G)$ of a bipartite graph G as left and right vertices.

Complete Graph

- A graph G is called *complete* if every vertex $v \in \mathsf{V}(G)$ is connected to all the other vertices in $\mathsf{V}(G) \setminus \{v\}$ (and not to itself);
- A bipartite graph is called complete if every vertex is connected to all the vertices of the other type in the graph.
- A complete $(d-1)$-regular graph is denoted by K_d, having a number of vertices $\bigl|\mathsf{V}(K_d)\bigr| = d$, and a number of edges $\bigl|\mathsf{E}(K_d)\bigr| = \frac{1}{2}\, d(d-1)$.
- A complete d-regular bipartite graph is denoted by $K_{d,d}$, having a number of vertices $\bigl|\mathsf{V}(K_{d,d})\bigr| = 2d$ (i.e., d vertices of each of the two types), and a number of edges $\bigl|\mathsf{E}(K_{d,d})\bigr| = d^2$.

Independent Set

An independent set of an undirected graph G is a subset of its vertices such that none of the vertices in this subset are adjacent (i.e., none of them are joined by an edge). Let $\mathcal{I}(G)$ denote the set of all the independent sets in G, and let $\left|\mathcal{I}(G)\right|$ denote the number of independent sets in G.

Independent Set

An independent set of an undirected graph G is a subset of its vertices such that none of the vertices in this subset are adjacent (i.e., none of them are joined by an edge). Let $\mathcal{I}(G)$ denote the set of all the independent sets in G, and let $\big|\mathcal{I}(G)\big|$ denote the number of independent sets in G.

Tensor Product

The tensor product $G \times H$ of two graphs G and H is a graph such that the vertex set of $G \times H$ is the Cartesian product $\mathsf{V}(G) \times \mathsf{V}(H)$, and two vertices $(g, h), (g', h') \in \mathsf{V}(G \times H)$ are adjacent if and only if g is adjacent to g', and h is adjacent to h' (i.e., $(g, g') \in \mathsf{E}(G)$ and $(h, h') \in \mathsf{E}(H)$).

Graph K_2

The graph K_2 is specialized to two vertices that are connected by an edge. We label the two vertices in K_2 by 0 and 1.

Graph K_2

The graph K_2 is specialized to two vertices that are connected by an edge. We label the two vertices in K_2 by 0 and 1.

Bipartite Double Cover

For a graph G, the tensor product $G \times K_2$ is a bipartite graph, called the *bipartite double cover* of G. The set of vertices in $G \times K_2$ is given by

$$\mathsf{V}(G \times K_2) = \big\{(v, i) : v \in \mathsf{V}(G),\, i \in \{0, 1\}\big\},$$

and its set of edges is given by

$$\mathsf{E}(G \times K_2) = \big\{\big((u, 0), (v, 1)\big) : (u, v) \in \mathsf{E}(G)\big\}.$$

An edge $(u, v) \in \mathsf{E}(G)$ is mapped into edges $\big((u, 0), (v, 1)\big) \in \mathsf{E}(G \times K_2)$ and $\big((v, 0), (u, 1)\big) \in \mathsf{E}(G \times K_2)$ (G is undirected).

Shearer's Lemma

Shearer's lemma extends the subadditivity property of Shannon entropy.

Proposition (Shearer's Lemma)

Let $X_1, \ldots, X_n$ be discrete random variables, and let $\mathcal{S}_1, \ldots, \mathcal{S}_m \subseteq [1:n]$ include every element $i \in [1:n]$ in at least $k \geq 1$ *of these subsets. Then,*

$$k \, \mathrm{H}(X^n) \leq \sum_{j=1}^{m} \mathrm{H}(X_{\mathcal{S}_j}). \tag{1}$$

Theorem (Kahn, 2001)

If G is a bipartite d-regular graph with n vertices, then

$$\bigl|\mathcal{I}(G)\bigr| \leq \bigl(2^{d+1}-1\bigr)^{\frac{n}{2d}}. \tag{2}$$

If n is an even multiple of d, then the upper bound in (2) is tight, and it is obtained by a disjoint union of $\frac{n}{2d}$ complete d-regular bipartite graphs $K_{d,d}$.

Theorem (Kahn, 2001)

If G is a bipartite d-regular graph with n vertices, then

$$\big|\mathcal{I}(G)\big| \leq \left(2^{d+1} - 1\right)^{\frac{n}{2d}}. \tag{2}$$

If n is an even multiple of d, then the upper bound in (2) is tight, and it is obtained by a disjoint union of $\frac{n}{2d}$ complete d-regular bipartite graphs $K_{d,d}$.

A different approach, not relying on IT, leads to the generalized result.

Theorem (Sah et al., 2019)

Let G be an undirected graph without isolated vertices or multiple edges connecting any pair of vertices. Let d_v be the degree of $v \in \mathsf{V}(G)$. Then,

$$\big|\mathcal{I}(G)\big| \leq \prod_{(u,v)\in \mathsf{E}(G)} \left(2^{\mathsf{d}_u} + 2^{\mathsf{d}_v} - 1\right)^{\frac{1}{\mathsf{d}_u\, \mathsf{d}_v}}, \tag{3}$$

with equality if G is a disjoint union of complete bipartite graphs.

Kahn's IT Proof (2001)

- Kahn's proof of the bipartite case of Theorem 1 made clever use of Shearer's entropy inequality.
- It remained unclear how to apply Shearer's inequality in a lossless way in the irregular case, despite some previous attempts to do so during the last decade.

Our Contributions

1. An extension of Kahn's information-theoretic proof technique to handle irregular bipartite graphs.
   - When the bipartite graph is regular on one side, but it may be irregular in the other, the extended entropy-based proof technique yields the same bound that was conjectured by Kahn and proved by Sah et al.
2. Providing a variant of the proof of Zhao's Inequality (which also involves entropy):

Theorem (Zhao 2010)

For every finite graph G:

$$\big|\mathcal{I}(G)\big|^2 \leq \big|\mathcal{I}(G \times K_2)\big|. \tag{4}$$

As an application of (4), an extension of (2) and (3) from bipartite graphs to general undirected graphs (without isolated vertices or multiple edges) was shown by Galvin & Zhao (2011).

A Recent Published Journal Paper

I. Sason, "A generalized information-theoretic approach for bounding the number of independent sets in bipartite graphs," *Entropy*, vol. 23, no. 3, paper 270, pp. 1–14, March 2021.

An entropy inequality proved by Kaced et al. (IEEE T-IT, 2018):

Theorem (Kaced et al. 2018)

Let A, X and Y be discrete random variables taking their values in the sets $\mathcal{A}, \mathcal{X}, \mathcal{Y}$, respectively, with a joint probability mass function $\mathsf{P}_{A,X,Y}$. If for every $(x, y) \in \mathcal{X} \times \mathcal{Y}$, there exists at most one element $a \in \mathcal{A}$ such that $\mathsf{P}_{A,X}(a, x)\, \mathsf{P}_{A,Y}(a, y) > 0$ then

$$\mathrm{H}(A|X) + \mathrm{H}(A|Y) \le \mathrm{H}(A). \tag{5}$$

The following is a modest generalization of the previous theorem by Kaced et al. (IEEE T-IT, 2018), which suggests an extension of their result with respect to edge coloring of bipartite graphs.

Proposition

Let A, X, Y be discrete random variables taking values in sets $\mathcal{A}, \mathcal{X}, \mathcal{Y}$, respectively. Then,

$$\mathrm{H}(A|X) + \mathrm{H}(A|Y) \leq \mathrm{H}(A) + \log m, \tag{6}$$

where

$$m \triangleq \sup_{(x,y) \in \mathcal{X} \times \mathcal{Y}} \Bigl| \bigl\{ a \in \mathcal{A} : \, \mathsf{P}_{A,X}(a,x) \, \mathsf{P}_{A,Y}(a,y) > 0 \bigr\} \Bigr|. \tag{7}$$

Proposition 13 is useful if $\log m < \mathrm{H}(A)$; otherwise, (6) is trivial.

Corollary: Rich Edge Coloring of Bipartite Graphs

Consider

- a bipartite graph G with minimal left and right degrees that are equal to d_{L} and d_{R}, respectively.
- an edge coloring of G where
  1. every two edges sharing a node have different colors;
  2. for all pairs of left node v_{L} and right node v_{R} in $\mathsf{V}(G)$, there are at most m colors touching both v_{L} and v_{R}.

Then, the number of colors in such an edge coloring of G is at least $\frac{d_{\mathsf{L}}\, d_{\mathsf{R}}}{m}$.

Number of Walks of a Given Length in Bipartite Graphs

- Lower bounds on the number of walks of a given length in bipartite graphs rely on the work by Alon, Hoory and Linial on the Moore bound and its extension (2002).
- Its later information-theoretic formulation is due to Babu and Radhakrishnan (2014).

Number of Walks of a Given Length in Bipartite Graphs

- Lower bounds on the number of walks of a given length in bipartite graphs rely on the work by Alon, Hoory and Linial on the Moore bound and its extension (2002).
- Its later information-theoretic formulation is due to Babu and Radhakrishnan (2014).

Contribution

We introduce in this work refined bounds which are expressed in terms of Shannon entropies of probability mass functions that are induced by the degree distributions of the bipartite graph.

Lower Bounds on the Number of Walks of a Given Length

Proposition

Let G be a bipartite graph with a disjoint partition of its vertex set $\mathsf{V}(G)$ to sets of left and right vertices $\mathcal{U}$ and $\mathcal{V}$, respectively, with $|\mathcal{U}| = m$ and $|\mathcal{V}| = n$.

Lower Bounds on the Number of Walks of a Given Length

Proposition

Let G be a bipartite graph with a disjoint partition of its vertex set $\mathsf{V}(G)$ to sets of left and right vertices $\mathcal{U}$ and $\mathcal{V}$, respectively, with $|\mathcal{U}| = m$ and $|\mathcal{V}| = n$. Let $\mathcal{P}_k$ be the set of all walks of a given length $k \in \mathbb{N}$, where edges may be repeated.

Lower Bounds on the Number of Walks of a Given Length

Proposition

Let G be a bipartite graph with a disjoint partition of its vertex set $\mathsf{V}(G)$ to sets of left and right vertices $\mathcal{U}$ and $\mathcal{V}$, respectively, with $|\mathcal{U}| = m$ and $|\mathcal{V}| = n$. Let $\mathcal{P}_k$ be the set of all walks of a given length $k \in \mathbb{N}$, where edges may be repeated. Let d_r denote the degree of a vertex $r \in \mathsf{V}(G)$, and let P and Q be PMFs defined, respectively, on $\mathcal{U}$ and $\mathcal{V}$ as follows:

$$\mathsf{P}(u) \triangleq \frac{\mathsf{d}_u}{|\,\mathsf{E}(G)|}, \quad u \in \mathcal{U}, \tag{8}$$

$$\mathsf{Q}(v) \triangleq \frac{\mathsf{d}_v}{|\,\mathsf{E}(G)|}, \quad v \in \mathcal{V}. \tag{9}$$

Lower Bounds on the Number of Walks of a Given Length (cont.)

1) If k is odd, then

$$\begin{aligned} \left|\mathcal{P}_k\right| &\geq \left|\,\mathsf{E}(G)\right|^k \exp\left(-\tfrac{1}{2}(k-1)[\mathrm{H}(P)+\mathrm{H}(Q)]\right) && (10) \\ &\geq \frac{\left|\,\mathsf{E}(G)\right|^k}{(mn)^{\frac{k-1}{2}}}. && (11) \end{aligned}$$

Lower Bounds on the Number of Walks of a Given Length (cont.)

1. If k is odd, then

$$
\begin{aligned}
\left|\mathcal{P}_k\right| &\geq |\,\mathsf{E}(G)|^k \exp\!\left(-\tfrac{1}{2}(k-1)[\mathrm{H}(P)+\mathrm{H}(Q)]\right) && (10) \\
&\geq \frac{|\,\mathsf{E}(G)|^k}{(mn)^{\frac{k-1}{2}}}. && (11)
\end{aligned}
$$

2. If k is even, then

$$
\begin{aligned}
\left|\mathcal{P}_k\right| &\geq |\,\mathsf{E}(G)|^k \exp\!\left(-(\tfrac{1}{2}k-1)[\mathrm{H}(P)+\mathrm{H}(Q)]\right) \\
&\quad \cdot \exp\!\left(-\min\{\mathrm{H}(P),\mathrm{H}(Q)\}\right) && (12) \\
&\geq \frac{|\,\mathsf{E}(G)|^k}{(mn)^{\frac{k}{2}-1}\,\min\{m,n\}}. && (13)
\end{aligned}
$$

Lower Bounds on the Number of Walks of a Given Length (cont.)

1) If k is odd, then

$$\begin{aligned} \big|\mathcal{P}_k\big| &\geq \; |\,\mathsf{E}(G)|^k \exp\big(-\tfrac{1}{2}(k-1)[\mathrm{H}(P)+\mathrm{H}(Q)]\big) && (10) \\ &\geq \; \frac{|\,\mathsf{E}(G)|^k}{(mn)^{\frac{k-1}{2}}}. && (11) \end{aligned}$$

2) If k is even, then

$$\begin{aligned} \big|\mathcal{P}_k\big| &\geq \; |\,\mathsf{E}(G)|^k \exp\big(-(\tfrac{1}{2}k-1)[\mathrm{H}(P)+\mathrm{H}(Q)]\big) \\ &\quad \cdot \exp\big(-\min\{\mathrm{H}(P),\mathrm{H}(Q)\}\big) && (12) \\ &\geq \; \frac{|\,\mathsf{E}(G)|^k}{(mn)^{\frac{k}{2}-1}\,\min\{m,n\}}. && (13) \end{aligned}$$

Equalities in (10) and (12) hold if the bipartite graph G is regular.

Independent Sets

It is left for future work to study if our analysis (I.S., Entropy, March '21) can be applied to yield bounds on the size of a random independent set.

Independent Sets

It is left for future work to study if our analysis (I.S., Entropy, March '21) can be applied to yield bounds on the size of a random independent set.

Number of Trails and Paths of a Given Length (cont.)

- In a paper by Alon, Hoory and Linial (2002), a certain non-returning walk was considered for graphs of minimum degree at least 2.
- It is left for a future study to examine an adaptation of our analysis to yield similar bounds on the number of
  - k-length trails (i.e., walks with no repeated edges);
  - k-length paths (i.e., walks with no repeated edges & vertices).

Independent Sets

It is left for future work to study if our analysis (I.S., Entropy, March '21) can be applied to yield bounds on the size of a random independent set.

Number of Trails and Paths of a Given Length (cont.)

- In a paper by Alon, Hoory and Linial (2002), a certain non-returning walk was considered for graphs of minimum degree at least 2.
- It is left for a future study to examine an adaptation of our analysis to yield similar bounds on the number of
  - k-length trails (i.e., walks with no repeated edges);
  - k-length paths (i.e., walks with no repeated edges & vertices).

Thanks for your attention.