499 Final Report WN24

Zizhao Ding, Bohang Yu, Jason Zheng, Zhengjia Zhuo

April 23, 2024

Chapter 1

Flag Algebra

Flag algebra is a powerful tool used to study homomorphism density in graphs and many related topics. We provide an introduction to the theory of flag algebra, and some common applications.

1.1 Basic Notations

REF: Casey Qi's paper on flag algebra,

Definition 1.1.1 (order of G:|G|). The order of a graph G is the number of vertices in G. In this report, we write $|G| = |V(G)|$ and shall also refer to it as the size of G.

Definition 1.1.2 (subgraph induced by graph G, $G|_U$). Let $G = (V, E)$ be a graph and $U \subseteq V$ be a subset of the vertices in G. *The subgraph of G induced by U*, which we will denote as $G|_U$, is a graph whose vertex set is U and for all $u_1, u_2 \in U, u_1u_2$ is an edge in $G|_U$ if and only if u_1u_2 is an edge in G.

Note. This definition of induced subgraph is simply taking out the corresponding subpart of G.

Notation (families of (unlabeled) isomorphic copies of F in G,$C_F(G)$). $C_F(G)$ is the set of m-element subsets of $V(G)$ that induce a subgraph isomorphic to F.

1.2 Graph Density Definition Summary

Definition 1.2.1 (induced density of graph F in G, p(F,G)). Let F and G be two graphs with $m = |V(F)| \leq |V(G)| = n$, if $C_F(G)$ is the set of m-element subsets of $V(G)$ that induce a subgraph isomorphic to F. Then,*The induced density of F in $G, p(F, G)$*, is

$$p(F, G) = \frac{|C_F(G)|}{\binom{|G|}{|F|}}$$

Note. Note that here, as long as there exist an automorphism from F to a induced subgraph H of G, then it will have one and only one contribution to $C_F G$. Therefore, it actually counts the "unlabelled copies of F in G". We will introduce the $t_{ind}(F, G)$ in the following that each counts the "labelled" copies.

Definition 1.2.2 (induced homomorphism density). Let $\varphi : V(F) \to V(G)$ be a randomly chosen injective map. There are a total of $\binom{|G|}{|F|}|F|\,!$ such maps. Define $t_{ind}(F,G)$ to be the probability that φ is an injective homomorphism that also preserves non-adjacency, which we call an induced homomorphism density. Then we have

$$t_{ind}(F,G)=\frac{p(F;G)}{\left(S_{|F|}:\operatorname{Aut}(F)\right)},$$

where $\left(S_{|F|}:\operatorname{Aut}(F)\right)$ is the index of the automorphism group of F in the symmetric group $S_{|F|}$ and $C_F(G)$ is the set of m-element subsets of $V(G)$ that induce a subgraph isomorphic to F.

Definition 1.2.3 (edge density). The edge density of a graph G is defined as

$$\rho(G)=p\left(K_2,G\right)=\frac{|E(G)|}{\binom{|G|}{2}}$$

This is the total number of edges in G divided by the number of all possible edges in G.

Definition 1.2.4 ($\mathcal{H}-free$). Let $\mathcal{H}$ be a collection of graphs. A graph G is $\mathcal{H}$-free if $p(F,G)=0$ for all graphs F in $\mathcal{H}$. Equivalently, G is $\mathcal{H}$-free if G doesn't contain any F in $\mathcal{H}$ as an induced subgraph.

1.3 Types and Flags

Definition 1.3.1 (Type σ (is just a H-free graph)). Fix a collection of forbidden graphs $\mathcal{H}$ and a non-negative integer k. A *type* σ *of size* k is a $\mathcal{H}$-free graph with $V(\sigma)=[k]$. The empty type is denoted as $\emptyset$.

Remark. In other words, a type σ is a graph with a bijection from its vertices to the labels in $[k]$. We can think of it as a fully labelled graph.

Definition 1.3.2 (σ-flag (F,θ), is a H free graph + embedding of type σ to F). A σ-flag is a pair (F,θ), where F is a $\mathcal{H}$-free graph and θ is an embedding of σ into F. We will drop the θ when not emphasizing the embedding and refer to it as the σ-flag F.

Note. Here $\theta: V(\sigma)\to V(F)$ a injective graph homomorphism.

A σ-flag is essentially a partially labeled graph that is $\mathcal{H}$-free and whose labelled part is a copy of σ.

Remark. Observe that a $\emptyset$-flag is just a $\mathcal{H}$-free graph with unlabelled vertices and that a type σ of size k can be seen as a σ-flag with θ being the identity on $[k]$.

Notation (subflag). If $F=(G,\theta)$ is a σ-flag and $\operatorname{im}(\theta)\subseteq U\subseteq V(G)$), we will use $\left.F\right|_U$ to denote the subflag $(\left.G\right|_U,\theta)$.

Definition 1.3.3 (isomorphic σ-flag). σ-flags $F=(G,\theta)$ and $F'=(G',\theta')$ are *isomorphic* if there is a graph isomorphism $f:V(G)\to V\left(G'\right)$ between G and G' such that $f(\theta(i))=\theta'(i)$ for all $i\in[|\sigma|]$. In other words, F and F' are isomorphic if $G\cong G'$ and the type σ is preserved in this isomorphism. We write $(F,\theta)\cong(F',\theta')$ or simply $F\cong F'$.

Note. A simple way to think about whether two sigma-flags are isomorphic, is to first label the images of theta in G and G', then see if there exist isomoprhisms between G and G' successfully projects the correct images of the labeled vertices.

Definition 1.3.4 ($\mathcal{F}^\sigma$,$\mathcal{F}^\sigma_l$)**.** Define $\mathcal{F}^\sigma$ to be the set of all σ-flags up to isomorphism and $\mathcal{F}^\sigma_l$ the set of all σ-flags of size $l \geq |\sigma|$, up to isomorphism. Note that $\mathcal{F}^\sigma_{|\sigma|}$ is the set of the single element $(\sigma, \mathrm{Id}_\sigma)$.

Notation. We will write $1_\sigma = (\sigma, \mathrm{Id}_\sigma)$. Let $\mathcal{G}$ be the set of all $\mathcal{H}$-free graphs up to isomorphism. Observe that $\mathcal{F}^\emptyset = \mathcal{G}$.

Definition 1.3.5 (degenerate)**.** Type σ is degenerate if $\mathcal{F}^\sigma$ is finite. If σ is non-degenerate, then $|\mathcal{F}^\sigma_l| \geq 1$ for all $l \geq |\sigma|$.

1.4 Flag Densities

We will keep with the tradition of using the bold font to denote randomly chosen objects.

Definition 1.4.1 (flag density)**.** Fix a type σ of size k and let $l, l_1, \ldots, l_t \geq k$ be integers that satisfy

$$(l_1 - k) + (l_2 - k) + \ldots + (l_t - k) \leq l - k \tag{1.1}$$

For σ-flags $F = (G, \theta), F_1, F_2, \ldots, F_t$ of size $l, l_1, l_2, \ldots, l_t$, respectively, we define the quantity $p(F_1, F_2, \ldots, F_t; F)$ in $[0, 1]$ as follows. Choose sets $\mathbf{V_1}, \mathbf{V_2}, \ldots, \mathbf{V_t}$ from $V(G)$ uniformly at random such that $V_i \cap V_j = \mathrm{im}(\theta)$ for all $i \neq j \in [t]$ and that $|V_i| = l_i$ for all $i \in [t]$. This is possible because of Equation 1.1 and it is equivalent to choosing disjoint sets of size $l_1 - k, l_2 - k, \ldots, l_t - k$ randomly from $V(G) \setminus \mathrm{im}(\theta)$ and taking the union of each of these sets and $\mathrm{im}(\theta)$ to form $V_1, V_2, \ldots, V_t$. We let $p(F_1, F_2, \ldots, F_t; F)$ denote the probability that for all $i \in [t]$, $F|_{V_i} \cong F_i$. Note that when $t = 1, p(F_1; F)$ represents the density of F_1 in F. We will write $p(F_1, F)$ instead of $p(F_1; F)$.

Note. The definition follows the intuition in the sense that if we fix the graph σ and randomly take |F| - k vertices(do not include any vertice of σ), how possible could we get (almost disjoint) copies of F1, F2,....Fk. (Here almost disjoint means that intersection between any two of F1,F2,....Fk could be and only be σ)

Note. The goal is finding out the flag density of a sequence of (fixed) smaller σ flags F1,F2, F3,...Ft, with respect to a bigger σ flag F $= (G, \theta)$. The fixed variables are k, l1, l2,...lt and a type σ of size k.

1. We first get V1, V2,....Vt by choosing disjoint sets of size $l_1 - k, l_2 - k, \ldots, l_t - k$ randomly from $V(G) \setminus \mathrm{im}(\theta)$ and taking the union of each of these sets and $\mathrm{im}(\theta)$ to form $V_1, V_2, \ldots, V_t$.
2. We then check if each subgraph of F induced by V_i is isomorphic σ-flag to Fi.

Lemma 1.4.1 (Chain Rule)**.** Let σ be a type of size k, $F_i \in \mathcal{F}^\sigma_{l_i}$ for all $1 \leq i \leq t$, and $F \in \mathcal{F}^\sigma_l$. Then for all $1 \leq s \leq t$ and $\tilde{l} \leq l$ such that $(l_1 - k) + \ldots + (l_s - k) \leq \tilde{l} - k$ and $(\tilde{l} - k) + (l_{s+1} - k) + \ldots + (l_t - k) \leq l - k$, we have

$$p(F_1, \ldots, F_t; F) = \sum_{\tilde{F} \in \mathcal{F}^\sigma_{\tilde{l}}} p\left(F_1, \ldots, F_s; \tilde{F}\right) p\left(\tilde{F}, F_{s+1}, \ldots, F_t; F\right).$$

Proof. (sketch)

The tuple $(\mathbf{V_1}, \mathbf{V_2}, \ldots, \mathbf{V_t})$ of the sets in Definition 1.4.1 can be generated in multiple ways. One way, as shown in the definition, is to choose sets $\mathbf{V_1}, \mathbf{V_2}, \ldots, \mathbf{V_t}$ uniformly at random from $V(G)$ such that they all share $\mathrm{im}(\theta)$ but are otherwise disjoint with size l_i respectively. An alternative way is to choose sets $\tilde{\mathbf{V}}, \mathbf{V_{s+1}}, \ldots, \mathbf{V_t}$ uniformly at random from $V(G)$ in the same way as above first. Then choose sets $\mathbf{V_1}, \mathbf{V_2}, \ldots, \mathbf{V_s}$ uniformly at random within $\tilde{\mathbf{V}}$. This also gives a uniform distribution of $(\mathbf{V_1}, \mathbf{V_2}, \ldots, \mathbf{V_t})$ in $V(G)$. Let S denote the event that $F_{V_i} \cong F_i$ for all $i \in [t]$ and $S_{\tilde{F}}$ denote the event that $\tilde{F}_{V_i} \cong F_i$ for all $i \in [s]$. Then the law of total probability gives

$$P(S) = \sum_{\tilde{F} \in \mathcal{F}^\sigma_{\tilde{l}}} P\left(S \mid S_{\tilde{F}}\right) P\left(S_{\tilde{F}}\right) \tag{1.2}$$

Note that $P(S) = \rho\left(F_1, \ldots, F_t; F\right)$ and $P\left(S_{\tilde{F}}\right) = \rho\left(F_1, \ldots, F_s; \tilde{F}\right)$ follow from our set up. Also observe that $\rho\left(F_1, \ldots, F_s; \tilde{F}\right) \rho\left(\tilde{F}, F_{s+1}, \ldots, F_t; F\right) = P\left(S \cap S_{\tilde{F}}\right)$, which means that $P\left(S \mid S_{\tilde{F}}\right) = \rho\left(\tilde{F}, F_{s+1}, \ldots, F_t; F\right)$. Plugging these into Equation 1.2 gets the desired result. ⊛

Corollary 1.4.1. For all $\tilde{l} \leq l$, we have

$$p\left(F_1, \ldots, F_t; F\right) = \sum_{\tilde{F} \in \mathcal{F}^\sigma_{\tilde{l}}} p\left(F_1, \ldots, F_t; \tilde{F}\right) p(\tilde{F}; F).$$

Proof. Follows directly from chain rule when we take $s = t$. ⊛

1.5 Limit Densities

Definition 1.5.1 (convergence of σ-flag sequence)**.** Fix a type σ, a sequence of σ-flags $\{F_n\}_{n\in\mathbb{N}}$, where $F_n \in \mathcal{F}^\sigma_{l_n}$, is convergent if it is increasing (i.e. the sequence $\{l_n\}_{n\in\mathbb{N}}$ is strictly increasing), and $\lim_{n\to\infty} p\left(F, F_n\right)$ exists for every $F \in \mathcal{F}^\sigma$.

Notation. For $F_1 \in \mathcal{F}^\sigma_{l_1}$ and $F_2 \in \mathcal{F}^\sigma_{l_2}$ with $l_1 > l_2$, we let $p\left(F_1, F_2\right) = 0$.

As to what such a limit actually is, flag algebra gives a minimalist answer.

Definition 1.5.2. Let $\phi : \mathcal{F}^\sigma \to [0, 1]$ be a function. If $\lim_{n\to\infty} p\left(F, F_n\right) = \phi(F)$ for every $F \in \mathcal{F}^\sigma$, then we say the sequence of σ-flags $\{F_n\}_{n\in\mathbb{N}}$ converges to ϕ and ϕ is called a limit functional.

Remark. Since $|\mathcal{F}^\sigma_n|$ is finite for every $n \in \mathbb{N}, \mathcal{F}^\sigma = \bigcup_{n\in\mathbb{N}} \mathcal{F}^\sigma_n$ is countable. We can think of every σ-flag F_n as a point p^{F_n} in the infinite-dimensional space $[0, 1]^{\mathcal{F}^\sigma}$ defined by $p^{F_n}(F) = p\left(F, F_n\right)$ for every $F \in \mathcal{F}^\sigma$. Therefore, an increasing sequence of σ-flags $\{F_n\}_{n\in\mathbb{N}}$ is convergent if and only if the sequence $\left\{p^{F_n}\right\}_{n\in\mathbb{N}}$ is convergent in $[0, 1]^{\mathcal{F}^\sigma}$ in the product topology.

Theorem 1.5.1. Every increasing sequence of σ-flags contains a convergent subsequence.

Definition 1.5.3. Let $\mathbb{R}\mathcal{F}^\sigma$ be the real vector space of all formal, finite linear combinations of σ-flags with real coefficients and let ϕ be a limit functional. We can extend ϕ linearly to $\phi' : \mathbb{R}\mathcal{F}^\sigma \to \mathbb{R}$ by defining

$$\phi'\left(\sum_{i=1}^{t} c_i F_i\right) = \sum_{i=1}^{t} c_i \phi\left(F_i\right)$$

where $c_i \in \mathbb{R}$ and $F_i \in \mathcal{F}^\sigma$ for all $i \in [t]$. We will write $\phi' = \phi$. Then (3.7) gives

$$\phi(F) = \sum_{\tilde{F} \in \mathcal{F}^\sigma_{\tilde{l}}} p(F, \tilde{F})\phi(\tilde{F}) = \phi\left(\sum_{\tilde{F} \in \mathcal{F}^\sigma_{\tilde{l}}} p(F, \tilde{F})\tilde{F}\right),$$

which means that

$$F - \sum_{\tilde{F} \in \mathcal{F}^\sigma_{\tilde{l}}} p(F, \tilde{F})\tilde{F}$$

is in the kernel of ϕ. Let $\mathcal{K}^\sigma$ be the linear subspace generated by all elements of the form (3.13), where $F \in \mathcal{F}^\sigma_l$, and $|\sigma| \le l \le \tilde{l}$. Let $\mathcal{A}^\sigma$ be the quotient space $\mathcal{A}^\sigma = \mathbb{R}\mathcal{F}^\sigma / \mathcal{K}^\sigma$. Define a bilinear map $(\mathbb{R}\mathcal{F}^\sigma) \otimes (\mathbb{R}\mathcal{F}^\sigma) \to \mathcal{A}^\sigma$, where for all $f_1, f_2 \in \mathbb{R}\mathcal{F}^\sigma, f_1 \otimes f_2 \mapsto f_1 \cdot f_2$, as follows. Take any $F_1, F_2 \in \mathcal{F}^\sigma$ and choose $|l| \ge |l_1| + |l_2| - |\sigma|$. Define

$$F_1 \cdot F_2 = \sum_{F \in \mathcal{F}^\sigma_l} p\left(F_1, F_2, F\right) F$$

and extend this product onto $(\mathbb{R}\mathcal{F}^\sigma) \otimes (\mathbb{R}\mathcal{F}^\sigma) \to \mathcal{A}^\sigma$ by linearity.

1.6 The Flag Algebra

This section may be unintuitive at first, but is crucial to understanding why flag algebra is useful.

1.6.1 Basic Definitions

Definition 1.6.1 (convergent sequence of δ-flags). Fix a type σ, a sequence of σ-flags $\{F_n\}_{n\in\mathbb{N}}$, where $F_n \in \mathcal{F}^\sigma_{l_n}$, is convergent if the size diverges (i.e. the sequence $\{l_n\}_{n\in\mathbb{N}}$ is strictly increasing), and $\lim_{n\to\infty} p\left(F, F_n\right)$ exists for every $F \in \mathcal{F}^\sigma$. Convention: For $F_1 \in \mathcal{F}^\sigma_{l_1}$ and $F_2 \in \mathcal{F}^\sigma_{l_2}$ with $l_1 > l_2$, we let $p\left(F_1, F_2\right) = 0$.

As to what such a limit actually is, flag algebra gives a minimalist answer.

Definition 1.6.2. Let $\phi : \mathcal{F}^\sigma \to [0, 1]$ be a function. If $\lim_{n\to\infty} p\left(F, F_n\right) = \phi(F)$ for every $F \in \mathcal{F}^\sigma$, then we say the sequence of σ-flags $\{F_n\}_{n\in\mathbb{N}}$ converges to ϕ and ϕ is called a limit functional.

For the rest of this section, fix ϕ a limit functional on a flag σ.

1.6.2 The Algebra A^σ

Now, we study the behavior of ϕ, as an attempt to classify all such limit functionals. We want to be able to perform linear algebra on graphs, flags, and flag densities, so we allow ϕ to act on real linear combinations of flags, not just individual flags.

$$\phi : \mathbb{R}F^\sigma \to \mathbb{R}, \phi(\sum c_i F_i) = \sum c_i \phi(F_i) \tag{1.3}$$

Also, we now define a multiplication of flags. This definition may seem unintuitive at first. But, the reason we define multiplication this way is because we want to preserve multiplicity of ϕ. Explicitly, we want that

$$\phi(F \cdot G) = \phi(F) \cdot \phi(G) \tag{1.4}$$

You will notice that the definition of multiplication looks suspiciously similar to the chain rule. This is not a coincidence.

Definition 1.6.3. Let F_1, F_2 σ-flags, not necessarily the same size. $F_1 \cdot F_2 = \sum_{F \in F^\sigma_l} p(F_1, F_2, F)F$ for some $l \in \mathbb{N}$ large enough.

Is this well defined? Intuitively, because the chain rule "controls for sizes" of flags in F^σ_l, we would want this to be well defined with respect to l. Let's check.

Let $n < n'$.

$$\sum_{F \in F_n^\sigma} p(F_1, F_2; F)F - \sum_{F' \in F_{n'}^\sigma} p(F_1, F_2; F')F' \tag{1.5}$$

$$= \sum_{F \in F_n^\sigma} p(F_1, F_2; F)F - \sum_{F' \in F_{n'}^\sigma} (\sum_{F \in F_n^\sigma} p(F_1, F_2; F)p(F, F'))F' \tag{1.6}$$

$$= \sum_{F \in F_n^\sigma} p(F_1, F_2; F)F - \sum_{F \in F_n^\sigma} p(F_1, F_2, F) \sum_{F' \in F_{n'}^\sigma} p(F_1, F_2; F)p(F, F')F' \tag{1.7}$$

$$= \sum_{F \in F_n^\sigma} p(F_1, F_2; F)(F - \sum_{F' \in F_{n'}^\sigma} p(F, F')F') \tag{1.8}$$

Unfortunately, it is not well defined, but we differ by a "flag of limit density 0". We would like to modify our algebraic structure to allow for this difference. Note the term in the bottommost equation: We have that

$$\phi(F - \sum_{F' \in F_{n'}^\sigma} p(F, F')F') = 0 \tag{1.9}$$

by the chain rule. Exercise: How should we modify the domain/codomain of this formulation to make multiplication well-defined?

Answer: Define $K^\sigma =$ the vector space spanned by all elements of the form $F - \sum_{F' \in F_{n'}^\sigma} p(F, F')F'$, for all l, and all F. Now ,we define $A^\sigma = \mathbb{R}F^\sigma / K^\sigma$, the modular algebra. Now, we have a map $\cdot : A^\sigma \otimes A^\sigma \to A^\sigma$ that is well defined and preserves some very nice algebraic properties.

Theorem 1.6.1. ϕ is multiplicative. $\phi(f \cdot g) = \phi(f) \cdot \phi(g)$ for $f, g \in A^\sigma$.

Proof. By linearity, it suffices to prove this for flags F, G.

$$\phi(F \cdot G) = \phi(\sum_{H \in F_l^\sigma} p(F, G, H)H) \tag{1.10}$$

$$= \sum_{H \in F_l^\sigma} p(F, G, H)\phi(H) \tag{1.11}$$

$$= \sum_{H \in F_l^\sigma} p(F, G, H) lim_{n \to \infty} p(H, A_n) \text{ for some convergent } (A_n) \tag{1.12}$$

$$= lim_{n \to \infty} p(F, G, A_n) \tag{1.13}$$

$$= lim_{n \to \infty} p(F, A_n) lim_{n \to \infty} p(G, A_n) \text{ by multiplicity} \tag{1.14}$$

$$= \phi(F) \cdot \phi(G) \tag{1.15}$$

■

Theorem 1.6.2. Define $Hom(A^\sigma, \mathbb{R})$ =functions satisfying $\phi(af + bg) = a \cdot \phi(f) + b \cdot \phi(g), \phi(f \cdot g) = \phi(f) \cdot \phi(g), \phi(1_\sigma) = 1$. $Hom^+(A^\sigma, \mathbb{R})$, the homomorphisms mapping each flag to a nonnegative value, is exactly the set of limit functionals ϕ.

1.7 Downward Operator

So far, we have always been dealing with flags of a fixed type σ. But, in practice we often want to convert between flags of different types. So, we define the downward operator.

Definition 1.7.1. Fix a type σ with $|\sigma| = k$, and an injection $\eta : [k'] \to [k]$, and let $F = (G, \theta)$ be a σ-flag. Then, let $q_{\sigma,\eta}(F)$ be the probability that a random injection $\alpha : [k] \to V_G$ satisfies $(G, \theta) \cong F$, given that $(G, \theta\eta) \cong F_{|\eta}$.

Definition 1.7.2. Downward operator. Define the downward operator $[[\cdot]]_{\sigma,\eta} : F^\sigma \to \mathbb{R}F^{\sigma|\eta}$ to be

$$[[F]]_{\sigma,\eta} = q_{\sigma,\eta}(F) \cdot F|_\eta. \tag{1.16}$$

Extend this map linearly to $\mathbb{R}F^\sigma \to \mathbb{R}F^{\sigma|\eta}$.

Theorem 1.7.1. $[[\cdot]]_{\sigma,\eta}$ takes K^σ to $K^{\sigma|_\eta}$ and is thus linear from A^σ to $A^{\sigma|_\eta}$.

Note, the coefficient $q_{\sigma,\eta}$ is necessary to preserve linearity between the flag algebras of A^σ and $A^{\sigma|\eta}$. Intuitively, it "controls for" the respective densities of the flags σ, $\sigma|\eta$.

Proposition 1.7.1. $[[\cdot]]_{\sigma,\eta} : F^\sigma \to \mathbb{R}F^{\sigma|\eta}$ is not surjective on $F^{\sigma|\eta}$ up to $\mathbb{R}$-spans.

Proof. Let $\sigma = K_5$, $\sigma|\eta = K_4$. Consider the $\sigma|\eta$-flag F with 10 vertices and minimal edges. F is not in the image of $[[\cdot]]$ because $[[\cdot]]$ must preserve the edges of some K_5. ■

Example. See Casey Qi's example on page 14, or the presentation from 4/3/24.

1.8 Flag Algebra in Sidorenko's Conjecture

In the following sections, all graphs are assumed to be finite and simple.

Notation ($\mathcal{M}_n$ set of all simple n vertice graphs up to an isomorphism). For every natural number n,let $\mathcal{M}_n$ denote the set of all simple graphs on n vertices up to an isomorphism.

Notation (G and complement G^*). For a graph G, let $V(G)$ and $E(G)$, respectively denote the set of the vertices and the edges of G. The complement of G is denoted by G^*.

Theorem 1.8.1 (chain rule in sidorenko conjecture). The homomorphism density of a graph H in a graph G, denoted by $t(H;G)$, is the probability that a random map from the vertices of H to the vertices of G is a graph homomorphism, that is it maps every edge of H to an edge of G. If $H \in \mathcal{M}_\ell, G \in \mathcal{M}_n$, and $\ell \leq n$, then $t_0(H;G)$ denotes the probability that a random injective map from $V(H)$ to $V(G)$ is a graph homomorphism, and $p(H,G)$ denotes the probability that a random set of ℓ vertices of G induces a graph isomorphic to H. We have the following chain rule (cf. [Raz07, Lemma 2.2]):

$$t_0(H;G) = \sum_{F \in \mathcal{M}_\ell} t_0(H;F)p(F,G),$$

where $|V(H)| \leq \ell \leq |V(G)|$.

Recall the definition of common graph.

Definition 1.8.1. A graph H is called common if

$$\liminf_{n\to\infty} \min_{G\in\mathcal{M}_n} \left(t(H;G) + t\left(H;G^*\right)\right) \geq 2^{1-|E(H)|}.$$

Remark. It is easy to see that as $n \to \infty$, for a random graph G on n vertices, we have, with high probability, $t(H;G) + t\left(H;G^*\right) = 2^{1-|E(H)|} \pm o(1)$. Thus, H is common if the total number of copies of H in every graph and its complement asymptotically minimizes for random graphs. Note also that since $t(H;G)$ and $t_0(H;G)$ are asymptotically equal (again, as $n \to \infty$), one could use $t_0(H;G)$ in place of $t(H;G)$ in (2.2), and this is what Hatami will do in his proof.

Ref: NON-THREE-COLORABLE COMMON GRAPHS EXIST, HAMED HATAMI

In fact, Hatami solved one special case, the 5-wheel W5, of sidorenko conjecture in his paper, using flag-algebra. The goal of this section is roughly present what Hatami have done with the tool of flag algebra.

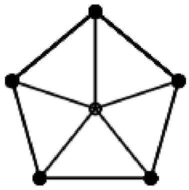

Figure 1.1: the five wheels W5

Notation. In the following sections, all graphs are assumed to be finite and simple

Notation ($\mathcal{M}_n$ set of all simple n vertice graphs up to an isomorphism)**.** For every natural number n,let $\mathcal{M}_n$ denote the set of all simple graphs on n vertices up to an isomorphism.

Notation (G and complement G^*)**.** For a graph G, let $V(G)$ and $E(G)$, respectively denote the set of the vertices and the edges of G. The complement of G is denoted by G^*.

Theorem 1.8.2 (chain rule in sidorenko conjecture)**.** The homomorphism density of a graph H in a graph G, denoted by $t(H;G)$, is the probability that a random map from the vertices of H to the vertices of G is a graph homomorphism, that is it maps every edge of H to an edge of G. If $H \in \mathcal{M}_\ell, G \in \mathcal{M}_n$, and $\ell \leq n$, then $t_0(H;G)$ denotes the probability that a random injective map from $V(H)$ to $V(G)$ is a graph homomorphism, and $p(H,G)$ denotes the probability that a random set of ℓ vertices of G induces a graph isomorphic to H. We have the following chain rule (cf. [Raz07, Lemma 2.2]):

$$t_0(H;G) = \sum_{F \in \mathcal{M}_\ell} t_0(H;F)p(F,G),$$

where $|V(H)| \leq \ell \leq |V(G)|$.

Recall the definition of common graph.

Definition 1.8.2. A graph H is called common if

$$\liminf_{n\to\infty} \min_{G \in \mathcal{M}_n} \left(t(H;G) + t\left(H;G^*\right)\right) \geq 2^{1-|E(H)|}.$$

Remark. It is easy to see that as $n \to \infty$, for a random graph G on n vertices, we have, with high probability, $t(H;G) + t\left(H;G^*\right) = 2^{1-|E(H)|} \pm o(1)$. Thus, H is common if the total number of copies of H in every graph and its complement asymptotically minimizes for random graphs. Note also that since $t(H;G)$ and $t_0(H;G)$ are asymptotically equal (again, as $n \to \infty$), one could use $t_0(H;G)$ in place of $t(H;G)$ in (2.2), and this is what Hatami will do in his proof.

Here is a rough view of Hatami's paper.

Theorem 3.1. *The 5-wheel W_5 is common.*

Proof. Let $\widehat{W}_5 \in \mathcal{A}^0$ be the element that counts the injective homomorphism density of the 5-wheel, that is

$$\widehat{W}_5 \stackrel{\text{def}}{=} \sum_{F \in \mathcal{M}_6} t_0(W_5, F)F.$$

We shall prove that

$$\widehat{W}_5 + \widehat{W}_5^* \geq 2^{-9}\,, \tag{3.1}$$

where the inequality $\leq$ in the algebra $\mathcal{A}^0$ is defined in [Raz07, Definition 6]. An alternate interpretation of this inequality [Raz07, Corollary 3.4] is that

$$\liminf_{n\to\infty} \min_{G\in\mathcal{M}_n} (p(\widehat{W}_5, G) + p(\widehat{W}_5^*, G)) \geq 2^{-9}.$$

Since $p(\widehat{W}_5, G) = \sum_{F\in\mathcal{M}_6} t_0(\widehat{W}_5; F)p(F; G) = t_0(\widehat{W}_5; G)$ by (2.1), and, likewise, $p(\widehat{W}_5^*, G) = p(\widehat{W}_5, G^*) = t_0(\widehat{W}_5; G^*)$, (3.1) implies Theorem 3.1.

We now give a proof of (3.1). To this end we work with suitable quadratic forms $Q_{\sigma_i}^{+/-}$ defined by symmetric matrices $M_{\sigma_i}^{+/-}$ and vectors $\mathbf{g}_i^{+/-}$ in the algebras $\mathcal{A}^{\sigma_i}$. The numerical values of the matrices $M_{\sigma_i}^{+/-}$ and vectors $\mathbf{g}_i^{+/-}$ are given in the appendix. It is essential that all the matrices $M_{\sigma_i}^{+/-}$ are positive definite which can be verified using any general mathematical software. Next we define

$$R := \left(\sum_{i=0}^{4} [\![Q_{\sigma_i}^{+}(\mathbf{g}_i^{+})]\!]_{\sigma_i}\right) + [\![Q_{\sigma_1}^{-}(\mathbf{g}_1^{-})]\!]_{\sigma_1} + [\![Q_{\sigma_4}^{-}(\mathbf{g}_4^{-})]\!]_{\sigma_4}\,.$$

We claim that

$$\widehat{W}_5 + \widehat{W}_5^* = 2^{-9} + R + R^*. \tag{3.2}$$

All the terms in (3.2) can be expressed as linear combinations of graphs from $\mathcal{M}_6$ and thus checking (3.2) amounts to checking the coefficients of the 156 flags from $\mathcal{M}_6$. We offer a C-code available at `http://kam.mff.cuni.cz/~kral/wheel` that verifies the equality (3.2).

By [Raz07, Theorem 3.14], we have

$$\left(\sum_{i=0}^{4} [\![Q_{\sigma_i}^{+}(\mathbf{g}_i^{+})]\!]_{\sigma_i}\right) + [\![Q_{\sigma_1}^{-}(\mathbf{g}_1^{-})]\!]_{\sigma_1} + [\![Q_{\sigma_4}^{-}(\mathbf{g}_4^{-})]\!]_{\sigma_4} \geq 0\,.$$

Therefore, (3.2) implies (3.1). □

Theorem 3.1 shows that a typical random graph $G = G_{n,\frac{1}{2}}$ asymptotically minimizes the quantity $t(W_5; G) + t(W_5; G^*)$. Extending our method, we convinced ourselves that $G_{n,\frac{1}{2}}$ is essentially the only minimizer of $t(W_5; G) + t(W_5; G^*)$. In terms of flag algebras

Figure 1.2: the five wheels W5

Chapter 2

A Brief Summary of Sidorenko's Paper

The goal of this section is give a brief view of the results in Sidorenko's original 1993 paper, in which he proved several results leading to the famous conjecture.

2.1 Basic Notations

Definition 2.1.1 (Bipartite Graph G; Reflection G*). For a bipartite graph $G = (V_1, V_2, E)$, denote $v_i(G) = |V_i|\,(i = 1, 2), e(G) = |E|$. If $E = V_1 \times V_2$, the graph G is called complete bipartite and is denoted by $K_{m,n}$ where $m = v_1(G), n = v_2(G)$. Denote by G^* the bipartite graph (V_2, V_1, E^*), where $(w, v) \in E^*$ if and only if $(u, w) \in E$.

Definition 2.1.2 (Independent Graph). Graphs are called *independent* if they have no common vertices.

Definition 2.1.3 (Graph Sum and Product (of bipartite graphs)). Let $G' = (V_1', V_2', E')$ and $G'' = (V_1'', V_2'', E'')$ be independent, then graphs $G' + G''$ and $G' \times G''$ are defined as follows:

$$G' + G'' = (V_1' \cup V_1'', V_2' \cup V_2'', E' \cup E''),$$
$$G' \times G'' = (V_1' \cup V_1'', V_2' \cup V_2'', E' \cup E'' \cup (V_1' \times V_2'') \cup (V_1'' \times V_2')).$$

Note. Graph sum is simple joining together the edge set and the vertex set. Graph product is graph sum with additional edges, such that the corresponding pairs of vertex sets become complete bipartite graph.

Definition 2.1.4 (number of connected components, k(G)). Let $k = k(G)$ be the maximal integer such that the bipartite graph G can be represented as $G = G_1 + G_2 + \ldots + G_k$. It is easy to see that this maximal representation is unique. The independent graphs $G_1, G_2, \ldots, G_k$ in the representation are called connected components of G. If $k(G) = 1$, the graph G is called connected.

Note. k(G) simply represents the number of (maximal) connected components.

It is conjectured that a bipartite graph belongs to the class $\mathcal{F}$ if it satisfies Conditions A, B

2.2 Conjecture and Related Inequalities

Denote by $\mathcal{F}$ the class of bipartite graphs $G = (\{u_1, \ldots, u_m\}, \{w_1, \ldots, w_n\}, E)$ which satisfy the following conditions.

Definition 2.2.1 (Condition A). $|E| \geq m, |E| \geq n$. Inequalities for functionals generated by bipartite graphs

Definition 2.2.2 (Condition B). For any spaces Ω, Λ and any functions $h \in K(\Omega \otimes \Lambda)$, $f, f_1, \ldots, f_m \in K(\Omega), g, g_1, \ldots, g_n \in K(\Lambda)$

$$\int \prod_{(u_i,w_j)\in E} h(x_i, y_j) \prod_{i=1}^{m} f_i(x_i) \prod_{j=1}^{n} g_j(y_j)\, \mathrm{d}\mu^m\, \mathrm{d}\nu^n \left(\int f(x)\mathrm{d}\mu\right)^{|E|-m} \left(\int g(y)\mathrm{d}\nu\right)^{|E|-n}$$

$$\geq \left(\int h(x,y) \left(f(x)^{|E|-m} g(y)^{|E|-n} \prod_{i=1}^{m} f_i(x) \prod_{j=1}^{n} g_j(y)\right)^{1/|E|} \mathrm{d}\mu \mathrm{d}\nu\right)^{|E|}$$

Note. Here condition A simple states that we want the edges number to be geq to size of both vertex classes of the bipartite graph.

Definition 2.2.3 (Sidorenko Conjecture).

$$\int \prod_{(u_i,w_j)\in E} h(x_i, y_j)\, \mathrm{d}\mu^m\, \mathrm{d}\nu^n \geq \left(\int h(x,y)\mathrm{d}\mu\mathrm{d}\nu\right)^{|E|} \mathrm{d}\mu(X)^{m-|E|}\mathrm{d}\nu(Y)^{n-|E|}.$$

Note. This is just the same definition as we seen in the Lovasz Book.

Sidorenko conjectured that this holds for any bipartite graph G, any spaces Ω, Λ, and any function $h \in K(\Omega \otimes \Lambda)$. This conjecture is formulated with respect to graphons, but we also consider the corresponding statement for graphs.

Remark. Obviously, Sidorenko Conjecture is a particular case of Condition B when functions $f, f_1, \ldots, f_m, g, g_1, \ldots g_n$ are constants equal to the unit. However, the domain of definition is wider for inequality Sidorenko Conjecture, since Condition A is not required (see Remark 1 below). Let us mention another cases of inequality Condition B. If we set

$$f(x) = \left(\int h(x,y) \left(g(y)^{|E|-n} \prod_{j=1}^{n} g_j(y)\right)^{1/|E|} \mathrm{d}\nu\right)^{|E|/m} \left(\prod_{i=1}^{m} f_i(x)\right)^{1/m}$$

the integral in the second factor of the left-hand side of Condition B coincides with the integral in the right-hand side. It gives the inequality

$$\int \prod_{(u_i,w_j)\in E} h(x_i, y_j) \prod_{i=1}^{m} f_i(x_i) \prod_{j=1}^{n} g_j(y_j)\, \mathrm{d}\mu^m\, \mathrm{d}\nu^n \left(\int g(y)\mathrm{d}\nu\right)^{|E|-n}$$

2.3 Survey of Results

Theorem 2.3.1. Let a graph G satisfy Condition A. If $v_1(G) \leq 3$ or $v_2(G) \leq 3$ then $G \in \mathcal{F}$.

Actually, we may prove the same assertion with $v_1(G) \leq 4$ or $v_2(G) \leq 4$, but we omit this because of the complexity of expressions in the proof.

Theorem 2.3.2. Let a bipartite graph G'' be obtained from a graph G by adding a new vertex and

a new edge which joins this vertex to some vertex a of the graph G. If G belongs to $\mathcal{F}$ then G'' also belongs to $\mathcal{F}$.

Corollary 2.3.1. Any tree with more than one vertex belongs to the class $\mathcal{F}$.

The definition of the class $\mathcal{F}$ is symmetric with respect to the colours of vertices. This implies the following assertions.

Proposition 2.3.1. If $G \in \mathcal{F}$ then $G^* \in \mathcal{F}$.

Proposition 2.3.2. If independent bipartite graphs G' and G'' belong to $\mathcal{F}$, then $G' + G''$ also belongs to $\mathcal{F}$.

Taking into account Remark 1 from Section 2, we get the following corollary.

Corollary 2.3.2. If a graph satisfies Condition A and all of its connected components belong to $\mathcal{F}$, then the graph belongs to $\mathcal{F}$ as well.

It follows from Theorems 2 and 4 that all trees and forests (that satisfy Condition A) belong to the class $\mathcal{F}$. For this case, inequality (1) was already proved in [13] with an additional assumption that $\Omega = \Lambda$ and the function h is symmetric.

Proposition 2.3.3. If $H \in \mathcal{F}$ then $H \times K_{p,q} \in \mathcal{F}$ with any $p, q \in \{0, 1, 2, \ldots\}$.

Theorem 5 implies that all complete bipartite graphs belong to the class $\mathcal{F}$. We would like to point out that the proof of inequality (1) for these graphs is implicitly contained in [6].

Theorem 2.3.3. Let a graph G belong to $\mathcal{F}$. Let us mark some of its vertices (their colours are not important) such that each edge has at most one marked end. Now we take k independent copies of G and, for each marked vertex, we identify (glue) all of its k images. Then the resulting graph G' belongs to the class $\mathcal{F}$.

Note. This is essentially considering independent sets of vertices, and allowing a graph to be "folded" with respect to its independent copies. See a later section for a more detailed explanation of "folding".

Theorem 2.3.4. Let a graph G belong to $\mathcal{F}$. Let us mark some of its vertices (their colours are not important) such that each edge has at most one marked end. Now we take k independent copies of G and, for each marked vertex, we identify (glue) all of its k images. Then the resulting graph G' belongs to the class $\mathcal{F}$.

Theorem 2.3.5. Let us consider the graphs G and G' from the formulation of Theorem 6. Let us choose an unmarked vertex a of colour i in the graph G. Let us consider a graph G_1 which belongs to $\mathcal{F}$ and satisfies the inequality $v_i\left(G_1\right) \leq k$. We identify (glue) each of its vertices of colour i with one of the k images of the vertex a in the graph G'. Then the resulting graph G'' belongs to the class $\mathcal{F}$.

Theorem 2.3.6. Let a bipartite graph G be a forest. The equality in (1) is attained if and only if the two following conditions hold simultaneously: (a) if $v_2(G) > 1$ then the function $\varphi(x) = \int h(x, y) \mathrm{d}\nu(y)$ is equal to a constant for almost all x with respect to the measure μ; (b) if $v_1(G) > 1$ then the function $\psi(y) = \int h(x, y) \mathrm{d}\mu(x)$ is equal to a constant for almost all y with respect to the measure ν.

Chapter 3

Entropy

We now move away from flag algebra to consider a more analytic technique to study graph densities.

3.1 Definitions and Basics

Definition : Given a discrete random variable X taking values in S, with $p_s := P(X = s)$, its entropy is defined to be

$$H(X) := \sum_{s \in S} -p_s \log_2 p_s$$

(3.1)

(by convention if $p_s = 0$ then the corresponding summand is set to zero).

Lemma 3.1.1. Uniform bound.

$$H(X) \leq \log_2 |\operatorname{support}(X)|,$$

with equality if and only if X is uniformly distributed.

Proof. Let function $f(x) = -x \log_2 x$ is concave for $x \in [0, 1]$. Let $S = \operatorname{support}(X)$. Then

$$H(X) = \sum_{s \in S} f(p_s) \leq |S| f\left(\frac{1}{|S|} \sum_{s \in S} p_s\right) = |S| f\left(\frac{1}{|S|}\right) = \log_2 |S|.$$

We write $H(X, Y)$ for the entropy of the joint random variables (X, Y), i.e., letting $Z = (X, Y)$,

$$H(X, Y) := H(Z) = \sum_{(x,y)} -P(X = x, Y = y) \log_2 P(X = x, Y = y).$$

Note that $H(X, Y) = H(X) + H(Y) \quad$ if X and Y are independent.

■

Definition 3.1.1. (Conditional entropy). Given jointly distributed random variables X and Y, define

$$\begin{aligned} H(X \mid Y) &:= E_y[H(X \mid Y = y)] \\ &= \sum_y P(Y = y) H(X \mid Y = y) \\ &= \sum_y P(Y = y) \sum_x -P(X = x \mid Y = y) \log_2 P(X = x \mid Y = y) \end{aligned}$$

Each line unpacks the previous line. In the summations, x and y range over the supports of X and Y respectively.

Lemma 3.1.2. (Chain rule). $H(X,Y) = H(X) + H(Y \mid X)$

Proof. Proof. Writing $p(x,y) = P(X = x, Y = y)$, etc., we have by Bayes's law,

$$p(x \mid y)p(y) = p(x,y),$$

(3.2)

and so

$$\begin{aligned} H(X \mid Y) := \mathbf{E}_y[H(X \mid Y = y)] &= \sum_y -p(y) \sum_x p(x \mid y) \log_2 p(x \mid y) \\ &= \sum_{x,y} -p(x,y) \log_2 \frac{p(x,y)}{p(y)} \\ &= \sum_{x,y} -p(x,y) \log_2 p(x,y) + \sum_y p(y) \log_2 p(y) \\ &= H(X,Y) - H(Y) \end{aligned}$$

■

Intuitively, the conditional entropy $H(X \mid Y)$ measures the amount of additional information in X not contained in Y.

Some important special cases:

- if $X = Y$, or $X = f(Y)$, then $H(X \mid Y) = 0$.
- If X and Y are independent, then $H(X \mid Y) = H(X)$
- If X and Y are conditionally independent on Z, then $H(X \mid Y, Z) = H(X \mid Z)$.

3.2 Conjectures and Important Results

We now survey several relevant results.

We will construct a probability distribution μ on $\mathrm{Hom}(F,G)$, the set of all graph homomorphisms $F \to G$. We are trying to prove a lower bound on $\mathrm{hom}(F,G)$ instead of an upper bound. we take μ to be carefully constructed distribution, and apply the upper bound.

$$H(\mu) \le \log_2 | \text{ support } \mu | = \log_2 \mathrm{hom}(F,G).$$

We are trying to show that

$$\frac{\mathrm{hom}(F,G)}{v(G)^{v(F)}} \ge \left(\frac{2e(G)}{v(G)^2}\right)^{e(F)}.$$

So we would like to find a probability distribution μ on $\mathrm{Hom}(F,G)$ satisfying

$$H(\mu) \ge e(F) \log_2(2e(G)) - (2e(F) - v(F)) \log_2 v(G).$$

Theorem 3.2.1. (Blakey and Roy 1965). Sidorenko's conjecture holds if F is a tree

Proof. Proof: Let us explain the proof when F is a path on 4 vertices. The same proof extends to all trees F.
We choose randomly a walk $XYZW$ in G as follows: - XY is a uniform random edge of G (by this we mean first choosing an edge of G uniformly at random, and then let X be a uniformly chosen

endpoint of this edge, and then Y the other endpoint); - Z is a uniform random neighbor of Y; - W is a uniform random neighbor of Z.
Key observation: YZ is distributed as a uniform random edge of G, and likewise with ZW
Indeed, conditioned on the choice of Y, the vertices X and Z are both independent and uniform neighbors of Y, so XY and YZ are uniformly distributed.
Also, the conditional independence observation implies that

$$H(Z \mid X, Y) = H(Z \mid Y) \quad \text{and} \quad H(W \mid X, Y, Z) = H(W \mid Z)$$

and futhermore both quantities are equal to $H(Y \mid X)$ since XY, YZ, ZW are each distributed as a uniform random edge.
Thus

$$\begin{aligned} H(X, Y, Z, W) &= H(X) + H(Y \mid X) + H(Z \mid X, Y) + H(W \mid X, Y, Z) && \text{[chain rule]} \\ &= H(X) + H(Y \mid X) + H(Z \mid Y) + H(W \mid Z) && \text{[conditional independence]} \\ &= H(X) + 3H(Y \mid X) && \text{[chain rule]} \\ &= 3H(X, Y) - 2H(X) \\ &\geq 3\log_2(2e(G)) - 2\log_2 v(G) \end{aligned}$$

In the final step we used $H(X, Y) = \log_2(2e(G))$ since XY is uniformly distributed among edges, and $H(X) \leq \log_2 |\text{support}(X)| = \log_2 v(G)$. This proves conjecture and hence the theorem for a path of 4 vertices. (As long as the final expression has the "right form" and none of the steps are lossy, the proof should work out.)
This proof easily generalizes to all trees.

■

Theorem 3.2.2. Sidorenko's conjecture holds for all complete bipartite graphs.

Proof. Following the same framework as earlier, let us demonstrate the result for $F = K_{2,2}$. The same proof extends to all $K_{s,t}$. We will pick a random tuple $(X_1, X_2, Y_1, Y_2) \in V(G)^4$ with $X_iY_j \in E(G)$ for all i, j as follows. - X_1Y_1 is a uniform random edge; - Y_2 is a uniform random neighbor of X_1; - X_2 is a conditionally independent copy of X_1 given (Y_1, Y_2).

The last point deserves more attention. Note that we are not simply uniformly randomly choosing a common neighbor of Y_1 and Y_2 as one might naively attempt. Instead, one can think of the first two steps as generating a distribution for (X_1, Y_1, Y_2)-according to this distribution, we first generate (Y_1, Y_2) according to its marginal, and then produce two conditionally independent copies of X_1.

From the previous proof (applied to a 2-edge path), we see that $H(X_1, Y_1, Y_2) \geq 2H(X_1, Y_1) - H(X_1) \geq 2\log_2(2e(G)) - \log_2 v(G)$.

So we have
$$\begin{aligned} &H(X_1, X_2, Y_1, Y_2) \\ &= H(Y_1, Y_2) + H(X_1, X_2 \mid Y_1, Y_2) && \text{[chain rule]} \\ &= H(Y_1, Y_2) + 2H(X_1 \mid Y_1, Y_2) && \text{[conditional independence]} \\ &= 2H(X_1, Y_1, Y_2) - H(Y_1, Y_2) && \text{[chain rule]} \\ &\geq 2(2\log_2(2e(G)) - \log_2 v(G)) - 2\log_2 v(G). && \text{[prev. ineq. and uniform bound]} \\ &= 4\log(2e(G)) - 4\log_2 v(G). \end{aligned}$$

■

Theorem 3.2.3. (Conlon, Fox, Sudakov 2010). Sidorenko's conjecture holds for a bipartite graph that has a vertex adjacent to all vertices in the other part.

Proof. Let us illustrate the proof for the following graph. The proof extends to the general case. Let us choose a random tuple $(X_0, X_1, X_2, Y_1, Y_2, Y_3) \in V(G)^6$ as follows: - X_0Y_1 is a uniform random edge; - Y_2 and Y_3 are independent uniform random neighbors of X_0; - X_1 is a conditionally independent copy of X_0 given (Y_1, Y_2); - X_2 is a conditionally independent copy of X_0 given (Y_2, Y_3). (as well as other symmetric versions.) Some important properties of this distribution: - X_0, X_1, X_2 are conditionally independent given (Y_1, Y_2, Y_3); - X_1 and (X_0, Y_3, X_2) are conditionally independent given (Y_1, Y_2); - The distribution of (X_0, Y_1, Y_2) is identical to the distribution of (X_1, Y_1, Y_2).

We have

$$\begin{aligned} & H\left(X_0, X_1, X_2, Y_1, Y_2, Y_3\right) \\ & = H\left(X_0, X_1, X_2 \mid Y_1, Y_2, Y_3\right) + H\left(Y_1, Y_2, Y_3\right) && \text{[chain rule]} \\ & = H\left(X_0 \mid Y_1, Y_2, Y_3\right) + H\left(X_1 \mid Y_1, Y_2, Y_3\right) + H\left(X_2 \mid Y_1, Y_2, Y_3\right) + H\left(Y_1, Y_2, Y_3\right) && \text{[conditional independence]} \\ & = H\left(X_0 \mid Y_1, Y_2, Y_3\right) + H\left(X_1 \mid Y_1, Y_2\right) + H\left(X_2 \mid Y_2, Y_3\right) + H\left(Y_1, Y_2, Y_3\right) && \text{[conditional independence]} \\ & = H\left(X_0, Y_1, Y_2, Y_3\right) + H\left(X_1, Y_1, Y_2\right) + H\left(X_2, Y_2, Y_3\right) - H\left(Y_1, Y_2\right) - H\left(Y_2, Y_3\right) && \text{[chain rule]} \end{aligned}$$

■

The proof of Theorem 10.3.3 actually lower bounds the first three terms:

$$\begin{aligned} H\left(X_0, Y_1, Y_2, Y_3\right) & \geq 3 \log_2(2e(G)) - 2\log_2 v(G) \\ H\left(X_1, Y_1, Y_2\right) & \geq 2 \log_2(2e(G)) - \log_2 v(G) \\ H\left(X_2, Y_2, Y_3\right) & \geq 2 \log_2(2e(G)) - \log_2 v(G) \end{aligned}$$

We can apply the uniform support bound on the remaining terms.

$$H\left(Y_1, Y_2\right) = H\left(Y_2, Y_3\right) \leq 2\log_2 v(G).$$

(3.3)

Putting everything together, we have

$$H\left(X_0, X_1, X_2, Y_1, Y_2, Y_3\right) \geq 7\log_2(2e(G)) - 8\log_2 v(G),$$

(3.4)

thereby verifying the conjecture.

3.3 Folding

We now present an application of entropy. Let A be a Sidorenko graph, and let W be a set of independent vertex in a, and let U be a vertex set and $U = V(A) \setminus W$, and we can label each vertex in W, U, $W = \{w_1, w_2 ... w_n\}, U = \{u_1, u_2 ... u_k\}$.
Finally, we set $U_i = \{u_{i1}, u_{i2} ... u_{ik}\}$.
Then we define a new graph G such that $V(G) = W \cup \bigcup_{i=1}^{k} U_i$ and $E(G)$ defined as follows:
Fix $v, u \in E(G)$, there are 4 possiblities:

1. $v, u \in W$, then they are not connected.
2. $v \in W, u \in U_i$ then let $v = w_x, u = u_{iy}$, if $(w_x, u_y) \in E(A)$, we say that v, u are connected.
3. $v \in U_i, u \in W$, the rules are the same with order of v, u reversed.
4. $v \in U_i, u \in U_j$, let $v = U_{ix}, u = U_{jy}$, we say that $(u, v) \in E(G)$, if $(u_x, u_y) \in E(A)$ and $i = j$.

We say that this graph G is Sidorenko. Note that we can see that G is also bipartite.

We will prove this result by using the graphon formulation of Sidorenko conjecture. We can see:

$$\begin{aligned}
\int_{[0,1]^{|V(G)|}} \prod_{i,j\in E(G)} H(x_i,x_j) &= \int_{[0,1]^{|W|}} \int_{[0,1]^{|V(G)-W|}} \prod_{i,j\in E(G)} H(x_i,x_j) \\
&= \int_{[0,1]^{|W|}} \int_{[0,1]^{|V(G)-W|}} \prod_{i,j\in E(G)} H(x_i,x_j) \\
&= \int_{[0,1]^{|W|}} \int_{[0,1]^{|V(G)-W|}} \prod_{n=1}^{k} \prod_{i,j\in E(G),x_j\in U_n} H(x_i,x_j) \\
&= \int_{[0,1]^{|W|}} \prod_{n=1}^{k} \int_{[0,1]^{|V(G)-W|}} \prod_{i,j\in E(G),x_j\in U_n} H(x_i,x_j) \\
&= \int_{[0,1]^{|W|}} (\int_{[0,1]^{|V(G)-W|}} \prod_{i,j\in E(G),x_j\in U_1} H(x_i,x_j))^k \ (\dagger)
\end{aligned}$$

Let's define

$$f = \int_{[0,1]^{|V(G)-W|}} \prod_{i,j\in E(G),x_j\in U_1} H(x_i,x_j)$$

By Jenson's inequality, we can see that $\int_{[0,1]^{|W|}} f^k \geq (\int_{[0,1]^{|W|}} f)^k$. Therefore,

$$\begin{aligned}
(\dagger) &\geq (\int_{[0,1]^{|W|}} f)^k \\
&= (\int_{[0,1]^{|W|}} f)^k \\
&= (\int_{[0,1]^{|W|}} \int_{[0,1]^{|V(G)-W|}} \prod_{i,j\in E(G),x_j\in U_1} H(x_i,x_j))^k \\
&= (\int_{[0,1]^{|V(G)|}} \prod_{i,j\in E(G),x_j\in U_1} H(x_i,x_j))^k \\
&= (\int_{[0,1]^{|V(G)-W-U_1|}} \int_{[0,1]^{|W+U_1|}} \prod_{i,j\in E(G),x_j\in U_1} H(x_i,x_j))^k \\
&= (\int_{[0,1]^{|V(G)-W-U_1|}} \int_{[0,1]^{|W+U_1|}} \prod_{i,j\in E(A),x_j\in U} H(x_i,x_j))^k \\
&= (\int_{[0,1]^{|V(G)-W-U_1|}} \int_{[0,1]^{|W+U_1|}} \prod_{i,j\in E(A)} H(x_i,x_j))^k \\
&= (\int_{[0,1]^{|V(G)-W-U_1|}} (\int_{[0,1]} H)^{E(A)})^k \\
&= ((\int_{[0,1]} H)^{E(A)})^k \\
&= (\int_{[0,1]} H)^{E(A)\cdot k} \\
&= (\int_{[0,1]} H)^{E(G)}
\end{aligned}$$

Chapter 4

Appendix: Turan Density

Definition 4.0.1 (Turan Density)**.** Let $\mathcal{H}$ be a collection of graphs and n a positive integer. The Turán number of $\mathcal{H}$ is the maximum number of edges a $\mathcal{H}$-free graph on n vertices can have and is expressed as

$$\mathrm{ex}(n, \mathcal{H}) = \max\{|E(G)| : G \text{ is } \mathcal{H}\text{-free}, |G| = n\}.$$

Note. If we normalize the Turán number of $\mathcal{H}$ by the number of all possible edges in a graph of size n, we would get the maximal edge density of a $\mathcal{H}$-free graph of size n. By taking the limit as n goes to infinity, we would get the induced Turán density of $\mathcal{H}$:

$$\pi(\mathcal{H}) = \lim_{n \to \infty} \frac{\mathrm{ex}(n, \mathcal{H})}{\begin{pmatrix} n \\ 2 \end{pmatrix}}$$

Note. Here just consider $\pi(\mathcal{H})$ as the asymptotic (maximal) edge density of graphs such that it is H-free (as number of vertices goes to inf).

We can extend this asymptotic maximal edge density to that of an arbitrary subgraph F.

Definition 4.0.2. Let $\mathcal{H}$ be a collection of graphs and n a positive integer. Let $C_F(G)$ be the set of m-element subsets of $V(G)$ that induce a subgraph isomorphic to F. Then the Turán F-number of $\mathcal{H}$ is

$$\mathrm{ex}_F(n, \mathcal{H}) = \max\left\{|C_F(G)| : G \text{ is } \mathcal{H}\text{-free}, |G| = n\right\}$$

Accordingly, the induced Turán F-density of $\mathcal{H}$ is defined as

$$\pi_F(\mathcal{H}) = \lim_{n \to \infty} \frac{\mathrm{ex}_F(n, \mathcal{H})}{\begin{pmatrix} n \\ |F| \end{pmatrix}}$$