2 Flag algebras

Let denote by $\mathcal{G}$ the set of all graphs (in general it can be any "reasonable" class of discrete structures) and by $\mathcal{G}_n$ the set of graphs on n vertices. Consider the set $\mathbb{R}\mathcal{G}$ of finite linear combinations of graphs with real coefficients with the natural operations of addition and multiplication by a scalar. For a graphon W define a homomorphism f_W from $\mathbb{R}\mathcal{G}$ to $\mathbb{R}$ by setting

$$f_W\left(\sum_{i\in I}\alpha_i H_i\right) = \sum_{i\in I}\alpha_i d(H_i, W)$$

for every element $\sum_{i\in I}\alpha_i H_i$ of $\mathbb{R}\mathcal{G}$. It is straightforward to verify that f_W is indeed a homomorphism from $\mathbb{R}\mathcal{G}$ to $\mathbb{R}$.

Observe that for every graph H and every integer $n \geq |H|$, it holds that

$$d(H,W) = \sum_{G\in\mathcal{G}_n} d(H,G)d(G,W). \tag{1}$$

This identity just says that the probability that the W-random graph of order $|H|$ is isomorphic to H is equal to the probability that the graph obtained from the W-random graph of order n by removing random $n-|H|$ vertices is isomorphic to H, which follows from the definition of a W-random graph. This identity implies that

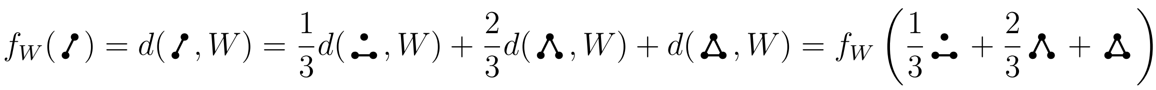
$$f_W\left(H - \sum_{G\in\mathcal{G}_n} d(H,G)G\right) = 0$$

for every graph H and every integer $n \geq |H|$. Let $\mathcal{A}_0$ be the the linear subspace of $\mathbb{R}\mathcal{G}$ generated by elements

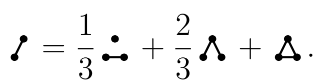
$$H - \sum_{G\in\mathcal{G}_n} d(H,G)G$$

for all graphs H and all $n \geq |H|$, and let $\mathcal{A}$ be the algebra $\mathbb{R}\mathcal{G}$ factored by $\mathcal{A}_0$, i.e., the algebra $\mathcal{A}$ is formed by the equivalence classes of the form $a + \mathcal{A}_0$ for $a \in \mathbb{R}\mathcal{G}$. Since $f_W(a) = 0$ for all $a \in \mathcal{A}_0$, the homomorphism f_W yields a homomorphism from $\mathcal{A}$ to $\mathbb{R}$.

Slightly abusing the notation, we will think of the elements of $\mathcal{A}$ as formal linear combinations of graphs with real coefficients and consider those belonging to the same equivalence class as equal. For brevity, we will also identify $a \in \mathcal{A}$ with $f_W(a)$, where W is any graphon. For example, since



holds for every graphon W, we briefly write



We want to define an operation of multiplication on the elements of $\mathcal{A}$ in a way that $f_W(a \cdot b) = f_W(a)f_W(b)$ for any $a, b \in \mathcal{A}$, i.e., the mapping f_W stays a homomorphism.

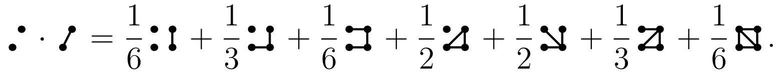
Consider two graphs H_1 and H_2. Observe that if we take the W-random graph of order $|H_1|+|H_2|$, consider its induced subgraphs arising from a random choice of disjoint subsets of $|H_1|$ and $|H_2|$ vertices, then we get the W-random graphs of orders $|H_1|$ and $|H_2|$. It follows that

$$d(H_1, W)d(H_2, W) = \sum_{G \in \mathcal{G}_{|H_1|+|H_2|}} d(H_1, H_2, G)d(G, W),$$

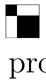
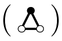
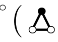
where $d(H_1, H_2, G)$ denotes the probability that a random partition of the vertex set of G into sets of $|H_1|$ and $|H_2|$ vertices is such that the subgraph of G induced by the first set is H_1 and the subgraph induced by the second set is H_2. In other words, $d(H_1, H_2, G)$ is the number of such partitions divided by the number of all partitions. Hence, we define

$$H_1 \cdot H_2 = \sum_{G \in \mathcal{G}_{|H_1|+|H_2|}} d(H_1, H_2, G)G$$

for any two graphs H_1 and H_2 and linearly extend it to $\mathbb{R}\mathcal{G}$ and $\mathcal{A}$. For example, it holds that



We now want to extend this concept to partially labeled graphs, so called *rooted graphs*, where some of the vertices are distinguishable. Consider a labeled graph R, called *root*, and let $\mathcal{G}^R$ be the set of all graphs with $|R|$ labeled vertices such that the labeled vertices induce the graph R, and let $\mathcal{G}^R_n$ for $n \geq |R|$ be the set of all graphs in $\mathcal{G}^R$ of order n. The elements of $\mathcal{G}^R$ will be depicted in a way that labeled vertices are drawn with empty circles and other vertices with full circles, moreover, vertices with the same label will be always drawn at the same position. For example, the graphs  and  contained in $\mathcal{G}_3^{\circ\,\circ}$ are different graphs.

Consider a labeled graph R with vertices $r_1, \ldots, r_k$, a graphon W having positive density of R with disregarded labels, and $x_1, \ldots, x_k \in [0, 1]$. If the probability that the W-random graph of order k with $r_1, \ldots, r_k$ chosen as $x_1, \ldots, x_k \in [0, 1]$ is R is positive, then we set $f_W^{x_1,\ldots,x_k}(H)$ for $H \in \mathcal{G}^R$ to be the probability that the W-random graph of order $|H|$ is H conditioned on the vertices $r_1, \ldots, r_k$ being chosen to be $x_1, \ldots, x_k$ and the graph sampled on these vertices being R. We linearly extend $f_W^{x_1,\ldots,x_k}$ from $\mathcal{G}^R$ to $\mathbb{R}\mathcal{G}^R$. Finally, we define f_W^R as the probability distribution of functions $f_W^{x_1,\ldots,x_k}$ under condition that $x_1, \ldots, x_k$ forms R. As an example consider the graphon $U =$  and observe that $f_U^{1/6,1/6}($$) = 1/3$, $f_U^{2/3,2/3}($$) = 2/3$, while $f_U^{\circ\!-\!\circ}($$) = 1/3$ with probability $1/5$ and $f_U^{\circ\!-\!\circ}($$) = 2/3$ with probability $4/5$. In any expressions involving f_W^R we keep the same roots, for example we write $f_U^{\circ}($$) + f_U^{\circ}($$) = 1$.

Extending the definition of density of graphs, for two rooted graphs G and H in $\mathcal{G}^R$ we define $d(H, G)$ as the probability that random $|H| - |R|$ unlabeled vertices of G together with R form the graph H. Observe that for every graphon W, labeled graph R, $H \in \mathcal{G}^R$ and $n \geq |H|$ it holds

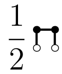
$$f_W^R(H) = \sum_{G \in \mathcal{G}_n^R} d(H, G) f_W^R(G).$$

Hence, similarly as for unrooted graphs, we consider the subspace $\mathcal{A}_0^R$ of $\mathbb{R}\mathcal{G}^R$ generated by $H - \sum_{G \in \mathcal{G}_n^R} d(H, G)G$ for all $H \in \mathcal{G}^R$ and $n \geq |H|$, and define $\mathcal{A}^R$ to be $\mathbb{R}\mathcal{G}^R$ factored by $\mathcal{A}_0^R$. Note that $\mathcal{A}^R$ may be viewed as $\mathcal{A}$ if R is a graph without vertices.

We next define multiplication on elements of $\mathcal{A}^R$ in a way analogous to those for unrooted graphs. If H_1, H_2 and G from $\mathcal{G}^R$ are such graphs that $|G| = |H_1| + |H_2| - |R|$, we define $d(H_1, H_2, G)$ to be the probability that a random partition of the unlabeled vertices of G to sets of sizes $|H_1| - |R|$ and $|H_2| - |R|$ is such that the first set together with R induces H_1 and the second set together with R induces H_2. We set

$$H_1 \cdot H_2 = \sum_{G \in \mathcal{G}^R_{|H_1|+|H_2|-|R|}} d(H_1, H_2, G) G \tag{2}$$

for any two graphs H_1 and H_2 from $\mathcal{G}^R$ and linearly extend it to $\mathbb{R}\mathcal{G}^R$ and $\mathcal{A}$. For example, it holds that

 $\cdot$  $= \frac{1}{2}$  $+ \frac{1}{2}$ .

The defined algebra on $\mathcal{A}^R$ is called *flag algebra* and its elements are called *flags*. Note that it coincides with the definition of $\mathcal{A}$ if R is a graph without vertices.

Our goal now is to define a linear mapping $[\![\cdot]\!]$ from $\mathcal{A}^R$ to $\mathcal{A}$ that corresponds to averaging over all possible placements of the roots that forms R, i.e., for every $a \in \mathcal{A}^R$ and graphon W with $d(R_0, W) > 0$ satisfying

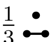
$$f_W([\![a]\!]) = f_W([\![R]\!]) \, \mathbb{E} f^R_W(a).$$

For a graph $H \in \mathcal{G}^R$, let $H_0 \in \mathcal{G}$ be the graph obtained from H by unlabeling the vertices of R. We define $q(H)$ to be the probability that choosing the vertices of R randomly among the vertices of H_0 yields H. Formally speaking,

$$q(H) = \frac{(|H| - |R|)!}{|H|!} \cdot \frac{|\operatorname{Aut}(H_0)|}{|\operatorname{Aut}(H)|}.$$

We set

$$[\![H]\!] = q(H) H_0$$

and linearly extend it to $\mathbb{R}\mathcal{G}^R$. For example, $[\![$  $]\!] = \frac{2}{3}$  and $[\![$  $]\!] = \frac{1}{3}$ . It is easy to see that $[\![\cdot]\!]$ is well defined as a mapping from $\mathcal{A}^R$ to $\mathcal{A}$. Note that for every $H \in \mathcal{G}^R$ and graphon W with $d(R_0, W) > 0$ it holds

$$\begin{aligned} f_W([\![R]\!]) \mathbb{E} f^R_W(H) &= \frac{(|H| - |R|)!}{|\operatorname{Aut}(H)|} \int_{[0,1]^{|H|}} \prod_{v_i v_j \in E(H)} W(x_i, x_j) \prod_{v_i v_j \notin E(H)} (1 - W(x_i, x_j)) \, \mathrm{d}x_1 \cdots \mathrm{d}x_{|H|} \\ &= \frac{(|H| - |R|)!}{|H|!} \cdot \frac{|\operatorname{Aut}(H_0)|}{|\operatorname{Aut}(H)|} \cdot d(H_0, W) \\ &= q(H) d(H_0, W) = f_W([\![H]\!]) \end{aligned}$$

as wanted.

Similarly as for unrooted graphs, we identify $H \in \mathcal{A}^R$ with $f^R_W(H)$. In particular, for $a \in \mathcal{A}^R$ and $x \in \mathbb{R}$ we write $a \geq x$ if $f^R_W(a) \geq x$ holds with probability 1 for every graphon W such that f^R_W is defined ($d(R_0, W) > 0$). From linearity of expectation it holds that if $a \geq 0$, then $[\![a]\!] \geq 0$. Similarly, if $a = 0$, then $[\![a]\!] = 0$, and more generally, if $a = x$ or $a \geq x$, then $[\![a]\!] = x \, [\![R]\!]$ or $[\![a]\!] \geq x \, [\![R]\!]$, respectively.