Notice that in general the quantities $[\![a \cdot b]\!]$ and $[\![a]\!] \cdot [\![b]\!]$ for $a, b \in \mathcal{A}^R$ can differ significantly, even if $a = b$. For example, for the previously considered graphon $U =$  we have

$$[\![e_1 \cdot e_1]\!] = [\![G^{1}_{12,13} + G^{1}_{12,13,23}]\!] = \frac{1}{3} P_3 + K_3 = \frac{1}{3},$$

while

$$[\![e_1]\!]^2 = K_2^{\,2} = \frac{25}{81}.$$

A general and useful bound is the following Cauchy-Schwarz inequality for flags.

Theorem 5. *For any labeled graph R and any flags a, $b \in \mathcal{A}^R$*

$$[\![a \cdot b]\!]^2 \leq [\![a^2]\!] \cdot [\![b^2]\!] .$$

In particular

$$[\![a]\!]^2 \leq [\![a^2]\!] \cdot [\![R]\!] .$$

To illustrate the just introduced concepts, we prove Mantel's theorem, which for graphons says that $K_3 = 0$ implies $K_2 \leq \frac{1}{2}$. Starting with an inequality $[\![(\bar e_1 - e_1)^2]\!] \geq 0$, valid because a square is always non-negative, we obtain

$$\begin{aligned} 0 \leq [\![(\bar e_1 - e_1)^2]\!] &= [\![\bar e_1 \cdot \bar e_1 - 2\,\bar e_1 \cdot e_1 + e_1 \cdot e_1]\!] \\ &= [\![G^{1}_{\emptyset} + G^{1}_{23} - G^{1}_{12} - G^{1}_{12,23} + G^{1}_{12,13} + G^{1}_{12,13,23}]\!] \\ &= \bar K_3 + \frac{1}{3} K_2\cup K_1 - \frac{2}{3} K_2\cup K_1 - \frac{2}{3} P_3 + \frac{1}{3} P_3 + K_3 \\ &= \bar K_3 - \frac{1}{3} K_2\cup K_1 - \frac{1}{3} P_3 + K_3 . \end{aligned}$$

Dividing this inequality by 2 and adding to the equality $K_2 = \frac{1}{3} K_2\cup K_1 + \frac{2}{3} P_3 + K_3$ we get

$$K_2 \leq \frac{1}{2} \bar K_3 + \frac{1}{6} K_2\cup K_1 + \frac{1}{2} P_3 + \frac{3}{2} K_3 \leq \frac{1}{2} \left(\bar K_3 + K_2\cup K_1 + P_3 + K_3 \right) + K_3 ,$$

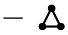
which for $K_3 = 0$ implies that $K_2 \leq \frac{1}{2}$ as needed.

The above prove shows Mantel's theorem only for graphons, but one can obtain from it a proof of Mantel's theorem for all graphs. Assume that there exists a counterexample to Mantel's theorem, i.e., a k-vertex graph G that does not contain triangles and has at least $\frac{k^2+1}{4}$ edges. Consider a sequence $G[n]$ of blow-ups of G. Each graph $G[n]$ does not contain triangles, has kn vertices and at least $\frac{k^2+1}{4} n^2$ edges, so its edge density is at least $\frac{k^2+1}{2k^2}$. From compactness, this sequence of graphs contains a convergent subsequence (one can even show that it is convergent by itself) whose limiting graphon has triangle density equal to 0 and edge density strictly larger than $\frac{1}{2}$, which contradicts the proven Mantel's theorem for graphons.

Additionally, observe that it follows from the proof that the equality in Mantel's theorem implies $K_2\cup K_1 = 0$ and $\bar e_1 = e_1$, which means that the extremal graphon corresponds to sequence of complete bipartite graphs which are regular (up to lower error terms).

The crucial ingredient used in the proof of Mantel's theorem was the inequality $\frac{1}{2} [\![(\bar e_1 - e_1)^2]\!] \geq 0$ expressed as inequality on flags on 3 vertices. One may wonder why exactly this inequality was used. The important advantage of applying flag algebra is the fact that one can use computer optimization in order to find inequalities giving the best estimate. We will briefly explain the idea behind this approach.

2.1 Automatic approach

We focus on maximizing some linear expression of flags g in some graph class $\mathcal{G}$. For example, in Mantel's theorem $g =$  and $\mathcal{G}$ is a class of triangle-free graphs, while in Goodman's theorem $g = 1 -$  $-$ and $\mathcal{G}$ is a class of all graphs (we have the minus sign here, because we want to find the minimum, not the maximum of  $+$ ).

Let us fix some integer n (at least of the size of the largest graph in g) and denote by $\mathcal{G}_n$ the family of all graphs in $\mathcal{G}$ on n vertices up to isomorphism. If n is not too large, then we can generate all flags in $\mathcal{G}_n$ and express g using identity (1) as $g = \sum_{H \in \mathcal{G}_n} g_H H$. In order to find a useful inequality, as in the discussed example, consider a labeled graph R, a vector v composed of all possible flags $F_i \in \mathcal{G}_m^R$ of order $m \leq (n + |R|)/2$, and assign some real coefficients q_{ij} to each pair of these flags. If the matrix $Q = (q_{ij})_{F_i, F_j \in \mathcal{G}_m^R}$ is positive semidefinite, then the inequality

$$vQv^T = \sum_{F_i, F_j \in \mathcal{G}_m^R} q_{ij} F_i \cdot F_j \geq 0$$

holds for each choice of $|R|$ labeled vertices. Using the definition of multiplication (2) we obtain a linear inequality on rooted flags on $2m - |R| \leq n$ vertices. Now, by averaging over a uniformly random choice of the root R we deduce an inequality of the form

$$[\![vQv^T]\!] = \sum_{H \in \mathcal{G}_n} c_H H \geq 0 \tag{3}$$

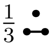
for some coefficients c_H depending on the root R, the integer m and the used positive semidefinite matrix Q.

Using it we obtain

$$g \leq \sum_{H \in \mathcal{G}_n} (g_H + c_H) H \leq \max_H (g_H + c_H) \sum_{H \in \mathcal{G}_n} = \max_H (g_H + c_H. \tag{4}$$

Since some of the coefficients c_H may be negative, for an appropriate choice of the root R, the order m and the positive semidefinite matrix Q, we can obtain a much better bound than the trivial one $g \leq \max_H g_H$. Furthermore, we can consider t choices of (R_i, m_i, Q_i), where each R_i is a labeled graph, each $m_i \leq (n + |R_i|)/2$ is an integer, and each Q_i is a positive semidefinite matrix of dimension $|\mathcal{G}_{m_i}^{R_i}| \times |\mathcal{G}_{m_i}^{R_i}|$. Defining the coefficients c_H as the sum of the coefficients obtained for each such triple, we infer a similar bound.

Therefore, we are left with the problem to minimize the bound $\max_H(g_H + c_H)$ from (4) over all positive semidefinite matrices Q_i used in the determination of the coefficients c_H. An upshot of the method is that this can be stated as a semidefinite programming problem (SDP): given any particular choice of the roots R_i and orders m_i, find positive semidefinite matrices Q_i that minimize $\max_H(g_H + c_H)$ in (4). Later we explain how one can set up an SDP solver.

Before we state this as a semidefinite problem, let us see how it works on a specific example. We will once again prove Mantel's theorem, i.e., find the best upper bound for the density of $g =$  in a class of triangle-free graphs. We use $n = 3$, express  as $\frac{1}{3}$  $+ \frac{2}{3}$ , consider the root R consisting of a single labeled vertex and take the vector

v of all flags on $m = 2$ vertices, i.e., $v = (F_0, F_1)$. Our task now is to find the smallest possible upper bound for

$$\frac{1}{3} K_2\cup K_1 + \frac{2}{3} P_3 + [\![(F_0, F_1) Q (F_0, F_1)^T]\!] \,,$$

where Q is some positive semidefinite matrix of size 2×2. Denoting the entries of Q by q_{ij} we obtain

$$\begin{aligned} [\![(F_0, F_1) Q (F_0, F_1)^T]\!] &= q_{00} [\![F_0 \cdot F_0]\!] + 2q_{01} [\![F_0 \cdot F_1]\!] + q_{11} [\![F_1 \cdot F_1]\!] \\ &= q_{00} [\![R_1 + R_2]\!] + q_{01} [\![R_3 + R_4]\!] + q_{11} [\![R_5]\!] \\ &= q_{00} \left(\overline{K_3} + \frac{1}{3} K_2\cup K_1 \right) + q_{01} \left(\frac{2}{3} K_2\cup K_1 + \frac{2}{3} P_3 \right) + q_{11} \left(\frac{1}{3} P_3 \right) \\ &= q_{00}\, \overline{K_3} + \left(\frac{1}{3} q_{00} + \frac{2}{3} q_{01} \right) K_2\cup K_1 + \left(\frac{2}{3} q_{01} + \frac{1}{3} q_{11} \right) P_3 . \end{aligned}$$

Altogether, we deduce that $\frac{1}{3} K_2\cup K_1 + \frac{2}{3} P_3 + [\![(F_0, F_1) Q (F_0, F_1)^T]\!]$ is equal to

$$q_{00}\, \overline{K_3} + \left(\frac{1}{3} + \frac{1}{3} q_{00} + \frac{2}{3} q_{01} \right) K_2\cup K_1 + \left(\frac{2}{3} + \frac{2}{3} q_{01} + \frac{1}{3} q_{11} \right) P_3 ,$$

which is at most

$$\max \left\{ q_{00}, \frac{1}{3} + \frac{1}{3} q_{00} + \frac{2}{3} q_{01}, \frac{2}{3} + \frac{2}{3} q_{01} + \frac{1}{3} q_{11} \right\} (\overline{K_3} + K_2\cup K_1 + P_3) .$$

Therefore, our final task is to minimize

$$\max \left\{ q_{00}, \frac{1}{3} + \frac{1}{3} q_{00} + \frac{2}{3} q_{01}, \frac{2}{3} + \frac{2}{3} q_{01} + \frac{1}{3} q_{11} \right\} \tag{5}$$

over all positive semidefinite matrices of size 2×2, which can be done by semidefinite programming.

2.2 Setting up the SDP

For fixed real symmetric matrices $C, A_1, A_2, \ldots, A_k$ of size $N \times N$ and a vector $b = (b_1, b_2, \ldots, b_k)$, the *semidefinite programming problem* (SDP) is a problem of the form

$$\begin{aligned} \max \quad & \mathrm{Tr}(CX) \\ \text{subject to} \quad & \mathrm{Tr}(A_i X) = b_i, \ \text{for } i \in \{1, 2, \ldots, k\}, \text{ and} \\ & X \text{ is positive semidefinite.} \end{aligned}$$

Such form is used by SDP solvers available online, like CSDP.

In order to express our minimization problem in (4) as an instance of SDP we take N equal to $1 + |\mathcal{G}_n| + |\mathcal{G}_m^R|$ and consider $k := |\mathcal{G}_n|$ symmetric matrices $A_1, A_2, \ldots, A_k$ and a matrix C, each consisting of 3 blocks: a 1×1 block, a diagonal block of size $|\mathcal{G}_n|$ and a block of size $|\mathcal{G}_m^R| \times |\mathcal{G}_m^R|$. In the matrix C we set the only entry of the first block to -1 and everything else to 0. For each graph H_i in $\mathcal{G}_n$ we create one matrix A_i by setting the entry in the first block to -1, the diagonal entry on position i in the second block to 1, and the values coming from $\llbracket vQv^T \rrbracket$ in the third block. As we assume that Q is positive semidefinite, in particular symmetric, we can put the corresponding values of $\llbracket vQv^T \rrbracket$ also symmetrically. Finally, for $i \in [|\mathcal{G}_n|]$ let the value of b_i be equal to $-g_H$, where H is the i-th graph in $\mathcal{G}_n$.

Let us denote the entries of the matrix X by x_0 in the first block, by x_i for $i \in [k]$ in the second diagonal block, and by q_{ij} for $i, j \in [|\mathcal{G}^R_m|]\}$ in the last block forming the matrix Q. The condition that X is positive semidefinite is equivalent to saying that the matrix Q is positive semidefinite and $x_i \geq 0$ for $i \in \{0, 1, 2, \ldots, k\}$.

By this assignment, the SDP objective function $\max \operatorname{Tr}(CX)$ is just maximizing $-x_0$, so minimizing x_0. For the i-th graph $H \in \mathcal{G}_n$, the assumption $\operatorname{Tr}(A_i X) = b_i$ turns into the equality $-x_0 + x_i + c_H = -g_H$, which means that $x_0 \geq g_H + c_H$ as $x_i \geq 0$. This is exactly saying that x_0 is the upper bound for the value of $\max_{H \in \mathcal{G}_n}(g_H + c_H)$ in (4). Thus, the SDP is indeed minimizing the wanted upper bound.

In our example, we have $N = 1 + 3 + 2 = 6$, $k = 3$, the matrices

$$C = \begin{pmatrix} -1 & & & & & \\ & 0 & & & & \\ & & 0 & & & \\ & & & 0 & & \\ & & & & 0 & 0 \\ & & & & 0 & 0 \end{pmatrix}, \qquad A_1 = \begin{pmatrix} -1 & & & & & \\ & 1 & & & & \\ & & 0 & & & \\ & & & 0 & & \\ & & & & 1 & 0 \\ & & & & 0 & 0 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} -1 & & & & & \\ & 0 & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \frac{1}{3} & \frac{1}{3} \\ & & & & \frac{1}{3} & 0 \end{pmatrix}, \qquad A_3 = \begin{pmatrix} -1 & & & & & \\ & 0 & & & & \\ & & 0 & & & \\ & & & 1 & & \\ & & & & 0 & \frac{1}{3} \\ & & & & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

and the vector $b = (0, -\frac{1}{3}, -\frac{2}{3})$.

With this assignment, the objective function $\max \operatorname{Tr}(CX)$ turns into minimization of x_0, while the assumptions $\operatorname{Tr}(A_i X) = b_i$ translate into

$$\begin{cases} -x_0 + x_1 + q_{00} = 0 \\ -x_0 + x_2 + \dfrac{1}{3} q_{00} + \dfrac{2}{3} q_{01} = -\dfrac{1}{3} \\ -x_0 + x_3 + \dfrac{2}{3} q_{01} + \dfrac{1}{3} q_{11} = -\dfrac{2}{3}. \end{cases} \tag{6}$$

Using $x_i \geq 0$, it gives

$$\max\left\{q_{00}, \frac{1}{3} + \frac{1}{3} q_{00} + \frac{2}{3} q_{01}, \frac{2}{3} + \frac{2}{3} q_{01} + \frac{1}{3} q_{11}\right\} \leq x_0.$$

Thus, our SDP is indeed minimizing the expression (5), as wanted.