Positive graphs

Omar Antolín Camarena, Endre Csóka, Tamás Hubai,
Gábor Lippner, László Lovász

May 2012

Abstract

We study "positive" graphs that have a nonnegative homomorphism number into every edge-weighted graph (where the edgeweights may be negative). We conjecture that all positive graphs can be obtained by taking two copies of an arbitrary simple graph and gluing them together along an independent set of nodes. We prove the conjecture for various classes of graphs including all trees. We prove a number of properties of positive graphs, including the fact that they have a homomorphic image which has at least half the original number of nodes but in which every edge has an even number of pre-images. The results, combined with a computer program, imply that the conjecture is true for all graphs up to 9 nodes.

Contents

1 Problem description

For a graph G we are going to denote the set of its vertices by $V(G)$ and the set of its edges by $E(G)$, but may simply write V and E when the it is clear from the context which graph we are talking about.

Let G and H be two simple graphs. A *homomorphism* $G \to H$ is a map $V(G) \to V(H)$ that preserves adjacency. We denote by $\hom(G,H)$ the number of homomorphisms $G \to H$. We extend this definition to graphs H whose edges are weighted by real numbers $\beta_{ij} = \beta_{ji}$ $(i,j \in V(H))$:

$$\hom(G,H) = \sum_{f\colon\ V(G)\to V(H)} \prod_{ij\in E(H)} \beta_{f(i)f(j)}.$$

(One could extend it further by allowing nodeweights, and also by allowing weights in G. Positive nodeweights in H would not give anything new; whether we get anything interesting through weighting G is not investigated in this paper.)

We call the graph G *positive* if $\hom(G,H) \geq 0$ for every edge-weighted graph H (where the edgeweights may be negative). It would be interesting to characterize these graphs; in this paper we offer a conjecture and line up supporting evidence.

We call a graph *symmetric*, if its vertices can be partitioned into three sets (S,A,B) so that S is an independent set, there is no edge between A and B, and there exists an isomorphism between the subgraphs spanned by $S\cup A$ and $S\cup B$ which fixes S.

Conjecture 1. *A graph G is positive if and only if it is symmetric.*

There is an analytic definition for graph positivity which is sometimes more convenient to work with. A *kernel* is a symmetric bounded measurable function $W : [0,1]^2 \to \mathbb{R}$. A map $p : V(G) \to [0,1]$ can be thought of as a homomorphism into W. It also naturally induces a map $p : E(G) \to [0,1]^2$. The *weight* of $p \in [0,1]^{V(G)}$ is then defined as

$$\hom(G,W,p) = \prod_{e\in E} W\big(p(e)\big) = \prod_{(a,b)\in E} W\big(p(a),p(b)\big).$$

The *homomorphism density* of a graph $G=(V,E)$ in a kernel W is defined as the expected weight of a random map:

$$t(G,W) = \int\limits_{[0,1]^V} \hom(G,W,p)\,\mathrm{d}p = \int\limits_{[0,1]^V} \prod_{e\in E} W\big(p(e)\big)\,\mathrm{d}p. \tag{1}$$

Graphs with real edge weights can be considered as kernels in a natural way: Let H be a looped-simple graph with edge weights β_{ij}; assume that $V(H) = [n] = \{1,\dots,n\}$. Split the interval $[0,1]$ into n intervals $J_1,\dots,J_n$ of equal length, and define

$$W_H(x,y) = \beta_{ij} \quad \text{for} \quad x\in J_i,\ y\in J_j.$$

Then it is easy to check that for every simple graph G and edge-weighted graph H, we have $t(G, W_H) = t(G, H)$, where $t(G, H)$ is a normalized version of homomorphism numbers between finite graphs:

$$t(G,H) = \frac{\hom(G,H)}{|V(H)|^{|V(G)|}}.$$

(For two simple graph G and H, $t(G, H)$ is the probability that a random map $V(G) \to V(H)$ is a homomorphism.)

It follows from the theory of graph limits [1, 6] that positive graphs can be equivalently be defined by the property that $t(G, W) \geq 0$ for every kernel W.

Hatami [3] studied "norming" graphs G, for which the functional $W \mapsto t(G, W)^{|E(G)|}$ is a norm on the space of kernels. Positivity is clearly a necessary condition for this (it is far from being sufficient, however). We don't know whether our Conjecture can be proved for norming graphs.

2 Results

In this section, we state our results (and prove those with simpler proofs). First, let us note that the "if" part of the conjecture is easy.

Lemma 2. *If a graph G is symmetric, then it is positive.*

Proof. For any map $p : V \to [0, 1]$ and any subset $M \subset V$ let p_M denote the restriction of p to M. Let further $G[M]$ denote the subgraph of G spanned by M.

$$t(G,W) \overset{(1)}{=} \int\limits_{[0,1]^V} \prod_{e\in E} W\big(p(e)\big)\,\mathrm{d}p = \int\limits_{[0,1]^V} \Big(\prod_{e\in G[S\cup A]} W\big(p(e)\big)\Big)\cdot\Big(\prod_{e\in G[S\cup B]} W\big(p(e)\big)\Big)\,\mathrm{d}p$$

$$= \int\limits_{[0,1]^S} \Big(\int\limits_{[0,1]^A} \prod_{e\in G[S\cup A]} W\big(p(e)\big)\,\mathrm{d}p_A\Big)\cdot\Big(\int\limits_{[0,1]^B} \prod_{e\in G[S\cup B]} W\big(p(e)\big)\,\mathrm{d}p_B\Big)\,\mathrm{d}p_S$$

$$= \int\limits_{[0,1]^S} \Big(\int\limits_{[0,1]^A} \prod_{e\in G[S\cup A]} W\big(p(e)\big)\,\mathrm{d}p_A\Big)^2\,\mathrm{d}p_S \geq= 0. \qquad \square$$

In the reverse direction, we only have partial results. We are going to prove that the conjecture is true for trees (Corollary 19) and for all graphs up to 9 nodes (see Section 5).

We state and prove a number of properties of positive graphs. Each of these is of course satisfied by symmetric graphs.

Lemma 3. *If G is positive, then G has an even number of edges.*

Proof. Otherwise, choosing W to be the constant -1 kernel we get $t(G, W) = (-1)^{|E(G)|} = -1$. $\square$

We call a homomorphism *even* if the preimage of each edge is has even cardinality.

Lemma 4. *If G is positive, then there exists an even homomorphism of G into itself.*

Proof. Let H be obtained from G by assigning random ± 1 weights to its edges, and let f be a random map $V(G) \to V(H)$. Then $\mathsf{E}_f(\hom(G, H, f)) = t(G, H) \geq 0$, and $t(G, H) > 0$ if all weights are 1, so $\mathsf{E}_H \mathsf{E}_f(\hom(G, H, f)) > 0$. Hence there is a f for which $\mathsf{E}_H(\hom(G, H, f)) > 0$. But clearly $\mathsf{E}_H(\hom(G, H, f)) = 0$ unless f is an even homomorphism of G into itself. $\square$

Let K_n denote the complete graph on the vertex set $[n]$, where $n \geq |V(G)|$.

Theorem 5. *If a graph G is positive, then there exists an even homomorphism $f:\ G \to K_n$ so that $\big|f(V(G))\big| \geq \frac{1}{2}\big|V(G)\big|$.*

We will prove this theorem in Section 4.

There are certain operations on graphs that preserve symmetry. Every such operation should also preserve positiveness. We are going to prove three results of this kind; such results are also useful in proving the conjecture for small graphs.

We need some basic properties of the homomorphism density function: Let G_1 and G_2 be two simple graphs, and let G_1G_2 denote their disjoint union. Then for every kernel W

$$t(G_1G_2, W) = t(G_1, W)t(G_2, W). \tag{2}$$

For two looped-simple graphs G_1 and G_2, we denote by $G_1 \times G_2$ their *categorical product*, defined by

$$\begin{aligned} &V(G_1 \times G_2) = V(G_1) \times V(G_2), \\ &E(G_1 \times G_2) = \big\{\big((i_1, i_2), (j_1, j_2)\big) :\ (i_1, j_1) \in E(G_1), (i_2, j_2) \in E(G_2)\big\}. \end{aligned}$$

We note that if at least one of G_1 and G_2 is simple (has no loops) then so is the product. The quantity $t(G_1 \times G_2, W)$ cannot be expressed as simply as (2), but the following formula will be good enough for us. For a kernel W

and looped-simple graph H, let us define the function $W^H : ([0,1]^V)^2 \to \mathcal{R}$ by

$$W^H\big((x_1,\dots,x_k),(y_1,\dots,y_k)\big) = \prod_{(i,j)\in E(H)} W(x_i,y_j) \tag{3}$$

(every non-loop edge of H contributes two factors in this product). Then we have

$$t(G\times H, W) = t(G, W^H). \tag{4}$$

The following lemma implies that it is enough to prove the conjecture for connected graphs.

Lemma 6. *A graph G is positive if and only if every connected graph that occurs among the connected components of G an odd number of times is positive.*

Proof. The "if" part is obvious by (2). To prove the converse, let $G_1,\dots,G_m$ be the connected components of a positive graph G. We may assume that these connected components are different and non-positive, since omitting a positive component or two isomorphic components does not change the positivity of G. We want to show that $m = 0$. Suppose that $m \ge 1$.

Claim 1. *We can choose kernels $W_1,\dots,W_m$ so that $t(G_i,W_i) < 0$ and $t(G_i,W_j) \ne t(G_j,W_j)$ for $i \ne j$.*

For every i there is a kernel W_i such that $t(G_i,W_i) < 0$, since G_i is not positive. Next we show that for every $i \ne j$ there is a kernel W_{ij} such that $t(G_i,W_{ij}) \ne t(G_j,W_{ij})$. If $|V(G_i)| \ne |V(G_j)|$ then the kernel $W_{ij} = \mathbb{1}(x,y \le 1/2)$ does the job, as in this case, due to the connectivity of the graphs, $t(G_i,W_{ij}) = (1/2)^{|V(G_i)|}$. So we may suppose that $|V(G_i)| = |V(G_j)|$. Then by [4, Theorem 5.29] there is a simple graph H such that $\hom(G_i,H) \ne \hom(G_j,H)$, and hence we can choose $W_{ij} = W_H$.

Let us denote $\underline{x} = (x_1,\dots,x_m)$ and define $W'_j(\underline{x}) = W_j + \sum_{i\ne j} x_i W_{ij}$. Expanding the product in the definition of $t(-,-)$ one easily sees that $Q_j(\underline{x}) = t(G_i, W'_j(\underline{x}))$ $(i = 1,\dots,m)$ are all different polynomials in the variables $\underline{x}$, and hence their values are all different for a generic choice of $\underline{x}$. If $\underline{x}$ is chosen close to $\underline{0}$, then $t(G_j, W'_j(\underline{x})) < 0$, and hence we can replace W_j by $W'_j(\underline{x})$. This proves the Claim.

Let $W_0 = 1$ denote the identically 1 kernel. For nonnegative integers $k_0,\dots,k_m$, construct a kernel $W_{k_0,\dots,k_m}$ by arranging k_i rescaled copies of W_i for each i on the "diagonal". Then

$$t(G_1\dots G_m, W_{k_0,\dots,k_m}) \overset{(2)}{=} \Big(\sum k_i\Big)^{-\sum|V(G_j)|} \prod_{j=1}^{m}\Big(\sum_{i=0}^{m} k_i t(G_j,W_i)\Big).$$

We know that this expression is nonnegative for every choice of the k_i. Since the right hand side is homogeneous in $k_0, \dots, k_m$, it follows that

$$\prod_{j=1}^{m}\Big(1+\sum_{i=1}^{m} x_i t(G_j, W_i)\Big) \geq 0 \tag{5}$$

for every $x_1, \dots, x_m \geq 0$. But the m linear forms $\ell_j(x) = 1+\sum_{i=1}^m x_i t(G_j, W_i)$ are different by the choice of the W_i, and each of them vanishes on some point of the positive orthant since $t(G_j, W_j) < 0$. Hence there is a point $x \in \mathbb{R}_+^m$ where the first linear form vanishes but the other forms do not. In a small neighborhood of this point the product (5) changes sign, which is a contradiction. □

Proposition 7. *If G is a positive simple graph and H is any looped-simple graph, then $G \times H$ is positive.*

Proof. Immediate from (4). □

Let $G(r)$ be the graph obtained from G by replacing each node with r twins of it. Then $G(r) \cong G \times K_r^\circ$, where K_r° is the complete r-graph with a loop added at every node. Hence we get:

Corollary 8. *If G is a positive simple graph, then so is $G(r)$ for every positive integer r.*

As a third result of this kind, we will show that the subgraph of a positive graph spanned by nodes with a given degree is also positive (Corollary 17). This proof, however, is more technical and is given in the next section. Unfortunately, these tools do not help us much for regular graphs G.

3 Subgraphs of positive graphs

In this section we develop a technique to show that one can partition the vertices of a positive graph in a certain way so that subgraphs spanned by each part are also positive. The main idea is to limit, over what maps $p : V \to [0,1]$ one has to average to check positivity. Using this idea recursively we can finally reduce to maps that take each partition to disjoint subsets of $[0,1]$. This in turn allows us to conclude positivity of the spanned subgraphs.

To this end, first we have to introduce the notion of $\mathcal{F}$-positivity. Let $G = (V, E)$ be a simple graph. For a measurable subset $\mathcal{F} \subseteq [0,1]^V$ and a bounded measurable weight function $\omega : [0,1] \to (0, \infty)$, we define

$$t(G, W, \omega, \mathcal{F}) = \int_{p\in\mathcal{F}} \hom(G, W, \omega, p)\, \mathrm{d}p, \tag{6}$$

where the *weight* of a $p : V \to [0,1]$ is

$$\hom(G, W, \omega, p) = \prod_{v \in V} \omega\big(p(v)\big) \prod_{e \in E} W\big(p(e)\big) \tag{7}$$

With the measure μ with density function ω (i.e., $\mu(X) = \int_X \omega$), we can write this as

$$t(G, W, \omega, \mathcal{F}) = \int\limits_{\mathcal{F}} \prod_{e \in E} W\big(p(e)\big)\, \mathrm{d}\mu^V(p). \tag{8}$$

We say that G is *$\mathcal{F}$-positive* if for every kernel W and function ω as above, we have $t(G, W, \omega, \mathcal{F}) \geq 0$. It is easy to see that G is $[0,1]^V$-positive if and only if it is positive.

We say that $\mathcal{F}_1, \mathcal{F}_2 \subseteq [0,1]^V$ are *equivalent* if there exists a bijection $\varphi : [0,1] \to [0,1]$ such that both φ and φ^{-1} are measurable, and $p \in \mathcal{F}_1 \Leftrightarrow \varphi(p) \in \mathcal{F}_2$, where $\varphi(p)(v) = \varphi\big(p(v)\big)$.

Lemma 9. *If $\mathcal{F}_1$ and $\mathcal{F}_2$ are equivalent, then G is $\mathcal{F}_1$-positive if and only if it is $\mathcal{F}_2$-positive.*

Proof. Let φ denote the bijection in the definition of the equivalence. For a kernel W and weight function ω, define the kernel $W^\varphi(x,y) = W\big(\varphi(x), \varphi(y)\big)$, and weight function $\omega^\varphi(x) = \omega\big(\varphi(x)\big)$, and let μ and μ_φ denote the measures defined by ω and ω^φ, respectively. With this notation,

$$\begin{aligned} t(G, W^\varphi, \omega^\varphi, \mathcal{F}_2) &= \int\limits_{\mathcal{F}_2} \prod_{e \in E} W^\varphi\big(p(e)\big)\, \mathrm{d}\mu_\varphi^V(p) \\ &= \int\limits_{\mathcal{F}_1} \prod_{e \in E} W\big(p(e)\big)\, \mathrm{d}\mu^V(p) = t(G, W, \omega, \mathcal{F}_1). \end{aligned}$$

This shows that if G is $\mathcal{F}_2$-positive if and only if it is $\mathcal{F}_1$-positive. $\square$

For a nonnegative kernel $W : [0,1]^2 \to [0,1]$ (these are also called *graphons*), function $\omega : [0,1] \to [0,\infty)$, and $\mathcal{F} \subseteq [0,1]^V$, define

$$s = s(G, W, \omega, \mathcal{F}) = \sup_{p \in \mathcal{F}} \Big(\prod_{v \in V} \omega\big(p(v)\big) \cdot \prod_{e \in E} W\big(p(e)\big) \Big), \tag{9}$$

and

$$\mathcal{F}_{max} = \Big\{ p \in \mathcal{F} : \prod_{v \in V} \omega\big(p(v)\big) \cdot \prod_{e \in E} W\big(p(e)\big) = s \Big\}.$$

If the Lebesgue measure $\lambda(\mathcal{F}_{max}) > 0$, then we say that $\mathcal{F}_{max}$ is *emphasizable* from $\mathcal{F}$, and (W, ω) *emphasizes* it.

Lemma 10. *If G is $\mathcal{F}_1$-positive and $\mathcal{F}_2$ is emphasizable from $\mathcal{F}_1$, then G is $\mathcal{F}_2$-positive.*

Proof. Suppose that (U, τ) emphasizes $\mathcal{F}_2$ from $\mathcal{F}_1$, and let $s = s(G, U, \tau, \mathcal{F}_1)$. Assume that G is not $\mathcal{F}_2$-positive, then there exists a kernel W and a weight function ω with $t(G, W, \omega, \mathcal{F}_2) < 0$. Consider the kernel $W_n = U^n W$ and weight function $\omega_n = s^{-n/|V|}\tau^n\omega$. Then

$$\prod_{v\in V}\omega_n\big(p(v)\big)\cdot\prod_{e\in E}W_n\big(p(e)\big) = \Big(\prod_{v\in V}\omega\big(p(v)\big)\cdot\prod_{e\in E}W\big(p(e)\big)\Big)\cdot a(p)^n,$$

where

$$a(p) = \frac{1}{s}\prod_{v\in V}\tau\big(p(v)\big)\cdot\prod_{e\in E}U\big(p(e)\big)\begin{cases} = 1 & \text{if } p\in\mathcal{F}_2,\\ < 1 & \text{otherwise.}\end{cases}$$

Thus (by the dominated convergence theorem)

$$\begin{aligned} t(G, W_n, \omega_n, \mathcal{F}_1) &= \int_{\mathcal{F}_1}\prod_{v\in V}\omega_n\big(p(v)\big)\cdot\prod_{e\in E}W_n\big(p(e)\big)\,\mathrm{d}p \\ &\to \int_{\mathcal{F}_2}\prod_{v\in V}\omega\big(p(v)\big)\cdot\prod_{e\in E}W\big(p(e)\big)\,\mathrm{d}p = t(G, W, \omega, \mathcal{F}_2) < 0,\end{aligned}$$

which implies that G is not $\mathcal{F}_1$-positive. □

For a partition $\mathcal{P}$ of $[0,1]$ into a finite number of sets with positive measure and a function $\pi: V\to\mathcal{P}$, we call the box $\mathcal{F}(\pi) = \{p\in[0,1]^V : p(v)\in\pi(v)\ \forall v\in V\}$ a *partition-box.* Equivalently, a partition-box is a product set $\prod_{v\in V}S_v$, where the sets $S_v\subseteq[0,1]$ are measurable, and either $S_u\cap S_v=\emptyset$ or $S_u = S_v$ for all $u, v\in V$.

A partition $\mathcal{N}$ of V is *positive* if for any partition $\mathcal{P}$ as above, and any $\pi: V\to\mathcal{P}$ such that $\pi^{-1}(\mathcal{P}) = \mathcal{N}$, G is $\mathcal{F}(\pi)$-positive.

Lemma 11. *If $\mathcal{F}_1\supseteq\mathcal{F}_2$ are partition-boxes, and G is $\mathcal{F}_2$-positive, then it is $\mathcal{F}_1$-positive.*

Proof. Let $\mathcal{F}_i$ be a product of classes of partition $\mathcal{P}_i$; we may assume that $\mathcal{P}_2$ refines $\mathcal{P}_1$. For $P\in\mathcal{P}_2$, let $\overline{P}$ denote the class of $\mathcal{P}_1$ containing P. Since every definition is invariant under measure preserving automorphisms of $[0,1]$, we may assume that every partition class of $\mathcal{P}_1$ and $\mathcal{P}_2$ is an interval.

Consider any kernel W and any weight function ω. Let $\varphi: [0,1]\to[0,1]$ be the function that maps each $P\in\mathcal{P}_2$ onto $\overline{P}$ in a monotone increasing and affine way. The map φ is measure-preserving, because for each $A\subseteq Q\in\mathcal{P}_1$,

$$\lambda\big(\varphi^{-1}(A)\big) = \sum_{\substack{P\in\mathcal{P}_2\\ P\subseteq Q}}\lambda\big(\varphi^{-1}(A)\cap P\big) = \sum_{\substack{P\in\mathcal{P}_2\\ P\subseteq Q}}\lambda(A)\frac{\lambda(P)}{\lambda(Q)} = \lambda(A). \tag{10}$$

Applying φ coordinate-by-coordinate we get a measure preserving map $\psi:\ [0,1]^V \to [0,1]^V$. Then $\psi' = \psi|_{\mathcal{F}_2}$ is an affine bijection from $\mathcal{F}_2$ onto $\mathcal{F}_1$, and clearly $\det(\psi') > 0$. Hence

$$\begin{aligned}
t(G, W^\varphi, \omega^\varphi, \mathcal{F}_2) &\overset{(6)}{=} \int_{\mathcal{F}_2} \prod_{v\in V} \omega^\varphi\big(p(v)\big) \cdot \prod_{e\in E} W^\varphi\big(p(e)\big)\,\mathrm{d}p \\
&= \int_{\mathcal{F}_2} \prod_{v\in V} \omega\big((\psi'(p))(v)\big) \cdot \prod_{e\in E} W\big((\psi'(p))(e)\big)\,\mathrm{d}p \\
&= \det(\psi')^{-1} \cdot \int_{\mathcal{F}_1} \prod_{v\in V} \omega\big(p(v)\big) \cdot \prod_{e\in E} W\big(p(e)\big)\,\mathrm{d}p \\
&\overset{(6)}{=} \det(\psi')^{-1} \cdot t(G, W, \omega, \mathcal{F}_1).
\end{aligned}$$

Since G is $\mathcal{F}_2$-positive, the left hand side is positive, and hence $t(G, W, \omega, \mathcal{F}_1) \geq 0$, proving that G is $\mathcal{F}_1$-positive. □

Corollary 12. *If $\mathcal{N}_2$ refines $\mathcal{N}_1$ and $\mathcal{N}_2$ is positive, then $\mathcal{N}_1$ is positive as well.*

Lemma 13. *Suppose that $\mathcal{F}_1$ is a partition-box defined by a partition $\mathcal{P}$ and function π_1. Let $Q \in \mathcal{P}$ and let U be the union of an arbitrary set of classes of $\mathcal{P}$. Let θ be a positive number but not an integer. Split Q into two parts with positive measure, Q^+ and Q^-. Let $\deg(v, U)$ denote the number of neighbors u of v with $\pi_1(u) \subseteq U$. Define*

$$\pi_2(v) = \begin{cases} \pi_1(v) & \text{if } \pi_1(v) \neq Q, \\ Q^+ & \text{if } \pi_1(v) = Q \text{ and } \deg(v, U) > \theta, \\ Q^- & \text{if } \pi_1(v) = Q \text{ and } \deg(v, U) < \theta, \end{cases}$$

and let $\mathcal{F}_2$ be the corresponding partition-box. Then there exists a pair (W, ω) emphasizing $\mathcal{F}_2$ from $\mathcal{F}_1$.

Proof. Clearly, $\lambda(\mathcal{F}_2) > 0$. First, suppose that $Q \not\subseteq U$. Let W be 2 in $Q^+ \times U$ and in $U \times Q^+$, and 1 everywhere else. Let $\omega(x)$ be $2^{-\theta}$ if $x \in Q^+$ and 1 otherwise. It is easy to see that the weight of a $p \in \mathcal{F}_1$ is 2^a, where $a = \sum_{v\in p^{-1}(Q^+)} \big(\deg(v, U) - \theta\big)$. This expression is maximal if and only if $p \in \mathcal{F}_2$.

In the case when $Q \subset U$ the only difference is that one has to let $W = 4$ in the intersection $Q^+ \times U \cap U \times Q^+$. □

Corollary 14. *If $\mathcal{N}_1$ is a positive partition of the vertex set, U is an arbitrary union of classes, Q is a single class, $\theta > 0$ is not an integer, and $\mathcal{N}_2$ is obtained from $\mathcal{N}_1$ by splitting Q according to whether (by abuse of notation)* $\deg(v, U) > \theta$ *or not for each vertex $v \in Q$, then $\mathcal{N}_2$ is also positive.*

We can use Corollary 14 iteratively: we start with the trivial partition, and refine it so that it remains positive. This is essentially the 1-dimensional Weisfeiler–Lehman algorithm, which classifies vertices recursively, see e. g. [2] It starts splitting vertices into classes according to their degree. Then in each step it refines the existing classes according to the number of neighbors in each of the current classes. The analogy will be clear from the proofs below. There is a non-iterative description of the resulting partition, and this is what we are going to describe next.

The *walk-tree* of a rooted graph (G, v) is the following infinite rooted tree $R(G, v)$: its nodes are all finite walks starting from v, its root is the 0-length walk, and the parent of any other walk is obtained by deleting its last node. The *walk-tree partition* $\mathcal{R}$ is the partition of V in which two nodes $u, v \in V$ belong to the same class if and only if $R(G, u) \cong R(G, v)$.

Proposition 15. *If a graph G is positive, then its walk-tree partition is also positive.*

Proof. Let the k-neighborhood of r in $R(G, r)$ be denoted by $R_k(G, r)$. The *k-walk-tree partition* $\mathcal{R}_k$ is the partition of V in which two nodes $u, v \in V$ belong to the same class if and only if $R_k(G, u) \cong R_k(G, v)$. Clearly, if for two vertices $R(G, u) \neq R(G, v)$ then there is a $k = k(u, v)$ such that $R_k(G, u) \neq R_k(G, v)$. Since V is finite, choosing $k_0 = \max_{u,v\in V} k(u, v)$ we see that $\mathcal{R}_{k_0} = \mathcal{R}$. Thus we are done if we show that $\mathcal{R}_k$ is positive for every k.

We prove this by induction. If $k = 0$ then $\mathcal{R}_0$ is the trivial partition, hence the assertion follows from the positivity of G. Now let us assume that the statement is true for k. Clearly, $R_{k+1}(G, v)$ is determined by the *neighborhood profile*, the multi-set $\{R_k(G, u) : u \sim v\}$. Using Corollary 14, we separate each class Q into subclasses so that $u, v \in Q$ end up in the same class if and only if their neighborhood profiles are the same. The new partition is exactly $\mathcal{R}_{k+1}$. □

Corollary 16. *Let $G(V, E)$ be a positive graph, and let $S \subset V$ be the union of some classes of the walk-tree partition. Then $G[S]$ is also positive.*

Proof. By Corollary 12 the partition $\mathcal{N} = \{S, V \setminus S\}$ is positive. Let $\mathcal{P} = \{[0, 1/2], (1/2, 1]\}$ and define $\pi : V \to \mathcal{P}$ by $\pi(v) = [0, 1/2]$ if and only if

$v \in S$. Suppose that $G[S]$ is negative as demonstrated by some W. Let us define

$$W'(x,y) = \begin{cases} W(2x,2y) & : x,y \in [0,1/2] \\ 1 & : \text{otherwise} \end{cases}$$

Then $t(G, W', 1, \mathcal{F}(\pi)) < 0$ contradicting the positivity of the partition $\mathcal{N}$. □

Corollary 17. *If G is positive, then for each k the subgraph of G spanned by all nodes with degree k is positive as well.* □

Corollary 18. *For each odd k the number of nodes of G with degree k must be even.*

Proof. Otherwise, consider the partition-box $\mathcal{F}$ that separates the vertices of G with degree d to class $A = [0, 1/2]$ and the other vertices to $\bar{A} = (1/2, 1]$. Consider the kernel W which is -1 between A and $\bar{A}$ and 1 in the other two cells. Then for each map $p \in [0,1]^V$, the total degree of the nodes mapped into class A is odd, so there is an odd number of edges between A and $\bar{A}$. So the weight of p is -1, therefore $t(G, W, 1, \mathcal{F}) = -\lambda(\mathcal{F}) < 0$. □

Corollary 19. *Conjecture 1 is true for trees.*

Proof. From the walk-tree of a vertex v of the tree G, we can easily decode the rooted tree (G, v). We call a vertex *central* if it cuts G into components with at most $|V|/2$ nodes. There can be either one central node or two neighboring central nodes of G. If there are two of them, then their walk-trees are different from the walk-trees of every other node. But these two points span a graph with a single edge, which is not positive, therefore Corollary 16 implies that neither is G. If there is only one central node, then consider the walk-trees of its neighbors. If there is an even number of each kind, then G is symmetric (and is thus positive by Lemma 2). Otherwise we can find two classes (one consist of the central node, the other consists of an odd number of its neighbors) whose union spans a graph with an odd number of edges, hence it is negative by Lemma 3. □

4 Homomorphic images of positive graphs

The main goal of this section is to prove Theorem 5. In what follows, let n be an integer. For a homomorphism $f : G \to K_n$, we call an edge $e \in E(K_n)$ *f-odd* if $\left|f^{-1}(e)\right|$ is odd. We call a vertex $v \in V(K_n)$ *f-odd* if there exists an

f-odd edge incident with v. Let $E_{\mathrm{odd}}(f)$ and $V_{\mathrm{odd}}(f)$ denote the set of f-odd edges and nodes of K_n, respectively, and define

$$r(f) = \big|V(G)\big| - \big|f\big(V(G)\big)\big| + \frac{1}{2}|V_{\mathrm{odd}}(f)|. \tag{11}$$

Lemma 20. *Let $G_i = (V_i, E_i)$ $(i = 1, 2)$ be two graphs, let $f:\ G_1G_2 \to K_n$, and let $f_i:\ G_i \to K_n$ denote the restriction of f to V_i. Then $r(f) \geq r(f_1) + r(f_2)$.*

Proof. Clearly $|V(G)| = |V_1| + |V_2|$ and $|V(f(G))| = |f(V_1)| + |f(V_2)| - |f(V_1) \cap f(V_2)|$. Furthermore, $E_{\mathrm{odd}}(f) = E_{\mathrm{odd}}(f_1) \triangle E_{\mathrm{odd}}(f_2)$, which implies that $V_{\mathrm{odd}}(f) \supseteq V_{\mathrm{odd}}(f_1) \triangle V_{\mathrm{odd}}(f_2)$. Hence

$$\begin{aligned} |V_{\mathrm{odd}}(f)| &\geq |V_{\mathrm{odd}}(f_1)| + |V_{\mathrm{odd}}(f_2)| - 2|V_{\mathrm{odd}}(f_1) \cap V_{\mathrm{odd}}(f_2)| \\ &\geq |V_{\mathrm{odd}}(f_1)| + |V_{\mathrm{odd}}(f_2)| - 2|f(V_1) \cap f(V_2)|. \end{aligned}$$

Substituting these expressions in (11), the lemma follows. □

Let G^k denote the disjoint union of k copies of a graph G. This lemma implies that if $f:\ G^k \to K_n$ is any homomorphism and $f_i:\ G \to K_n$ denotes the restriction of f to the i-th copy of G, then

$$r(f) \geq \sum_{i=1}^{k} r(f_i). \tag{12}$$

We define two parameters of a graph G:

$$\bar{r}(G) = \min\big\{r(f)\big|n \in \mathbb{N}, f:\ G \to K_n\big\} \tag{13}$$

and

$$q(G) = \min\Big\{|V(G)| - |f(V(G))|\ \Big|\ n \in \mathbb{N}, f:\ G \to K_n \text{ is even}\Big\}. \tag{14}$$

If there is no even homomorphism from G to K_n for any n then we define $q(G) = \infty$. Since $q(G) = \min\big\{r(f)\big|n \in \mathbb{N}, f:\ G \to K_n \text{ is even}\big\}$, it follows that

$$q(G) \geq \bar{r}(G). \tag{15}$$

Furthermore, considering any injective $f:\ G \to K_n$, we see that

$$\bar{r}(G) \leq r(f) = \big|V(G)\big| - \big|f\big(V(G)\big)\big| + \frac{1}{2}\big|f\big(V(G)\big)\big| = \frac{1}{2}\big|V(G)\big|. \tag{16}$$

Lemma 21.

$$\bar{r}(G^k) = k\bar{r}(G). \tag{17}$$

Proof. For one direction, take an $f: G^k \to K_n$ with $r(f) = \bar{r}(G^k)$. Then

$$\bar{r}(G^k) = r(f) \overset{(12)}{\geq} \sum_{i=1}^{k} r(f_i) \overset{(13)}{\geq} \sum_{i=1}^{k} \bar{r}(G) = k \cdot \bar{r}(G).$$

For the other direction, let us choose each f_i so that $r(f_i) = \bar{r}(G)$ and the images $f_i(G)$ are pairwise disjoint. Then

$$\bar{r}(G^k) \overset{(13)}{\leq} r(f) = \sum_{i=1}^{k} r(f_i) = \sum_{i=1}^{k} \bar{r}(G) = k \cdot \bar{r}(G). \qquad \square$$

Lemma 22.

$$q(G^2) = \bar{r}(G^2). \tag{18}$$

Proof. We already know by (15) that $q(G^2) \geq \bar{r}(G^2)$. For the other direction, we define $f: G^2 \to K_n$ as follows. We choose f_1 so that $r(f_1) = \bar{r}(G)$. Consider all points $v_1, v_2, ..., v_l$ in $f_1\big(V(G)\big)$ which are not f_1-odd. Let us choose pairwise different nodes $v_1', v_2', ..., v_l'$ disjointly from $f_1\big(V(G)\big)$. Now we choose f_2 so that for each $x \in V(G)$, if $f_1(x)$ is an f_1-odd point, then $f_2(x) = f_1(x)$, and if $f_1(x) = v_i$, then $f_2(i) = v_i'$.

If an edge $e \in E(K_n)$ is incident to a v_i, then $\left|f_1^{-1}(e)\right|$ is even and $f_2^{-1}(e) = \emptyset$. If e is incident to a v_i', then $\left|f_2^{-1}(e)\right|$ is even and $f_1^{-1}(e) = \emptyset$. If e is not incident to any v_i or v_i', then $\left|f_1^{-1}(e)\right| = \left|f_2^{-1}(e)\right|$. Therefore f is even. Thus,

$$q(G^2) \overset{(14)}{\leq} r(f) \overset{(11)}{=} \left|V(G^2)\right| - \left|f(V(G^2))\right|$$

$$= 2\left|V(G)\right| - \Big|f_1\big(V(G)\big)\Big| - \Big|f_2\big(V(G)\big)\Big| + \Big|f_1\big(V(G)\big) \cap f_2\big(V(G)\big)\Big|$$

$$= 2\left|V(G)\right| - 2\Big|f_1\big(V(G)\big)\Big| + |V_{\mathrm{odd}}(f_1)| \overset{(11)}{=} 2r(f_1) = 2\bar{r}(G) \overset{(17)}{=} \bar{r}(G^2). \qquad \square$$

Let K_n^w denote K_n equipped with an edge-weighting $w: E(K_n) \to \{-1, 1\}$. Let the stochastic variable $K_n^{\pm 1}$ denote K_n^w with a uniform random w.

Lemma 23. *For a fixed graph G, and $n \to \infty$,*

$$\mathsf{E}\big(t(G, K_n^{\pm 1})\big) = \begin{cases} \Theta\big(n^{-q(G)}\big) & \textit{if } q(G) < \infty \\ 0 & \textit{otherwise.} \end{cases}$$

Proof. If an edge e is f-odd, then changing the weight on e changes the sign of the homomorphism, therefore $\mathsf{E}_w\big(\hom(G, K_n^w, f)\big) = 0$. On the other hand, if f is even, then for all w, $\hom(G, K_n^w, f) = 1$. Therefore, taking a uniformly random homomorphism $f:\ G \to K_n$,

$$\begin{aligned}\mathsf{E}\big(t(G, K_n^{\pm 1})\big) &= \mathsf{E}_w\Big(\mathsf{E}_f\big(\hom(G, K_n^w, f)\big)\Big) = \mathsf{E}_f\Big(\mathsf{E}_w\big(\hom(G, K_n^w, f)\big)\Big)\\ &= \mathsf{P}(f \text{ is even}).\end{aligned}$$

If $q(G) = \infty$ we are done. Otherwise we have

$$\mathsf{P}(f \text{ is even}) \leq \mathsf{P}\Big(\big|V(G)\big| - \big|f(V(G))\big| \geq q(G)\Big) = O(n^{-q(G)}).$$

On the other hand, consider an even homomorphism $g:\ G \to K_n$ with $r(g) = q(G)$. For each subset $H \subset V(K_n)$ of size $|H| = |g(V(G))|$ there is a permutation σ_H on $V(K_n)$ that maps $g(V(G))$ bijectively to H. Then $f_H = \sigma(g(x))$ is also an even homomorphism, and clearly $f_{H_1} \neq f_{H_2}$ unless $H_1 = H_2$. Thus there are at least $\binom{n}{|g(V(G)|}$ different even homomorphisms $f:\ G \to K_n$. Therefore

$$\begin{aligned}\mathsf{P}(f \text{ is even}) &\geq \mathsf{P}(f = f_H \text{ for some } H) = \binom{n}{|g(V(G))|}\Big/ n^{|V(G)|}\\ &= \Omega(n^{-q(G)}).\end{aligned}$$

$\square$

Now let us turn to the proof of Theorem 5. Assume that G is positive, then the random variable $X = t(G, K_n^{\pm 1})$ is nonnegative. Applying the Cauchy-Schwartz inequality to $X^{1/2}$ and $X^{3/2}$ we get that

$$\mathsf{E}(X) \cdot \mathsf{E}(X^3) \geq \mathsf{E}(X^2)^2. \tag{19}$$

Here

$$\mathsf{E}(X^k) = \mathsf{E}\big(t(G, K_n^{\pm 1})^k\big) \overset{(2)}{=} \mathsf{E}\big(t(G^k, K_n^{\pm 1})\big) = \Theta\big(n^{-q(G^k)}\big),$$

so (19) shows that $n^{-q(G)} \cdot n^{-q(G^3)} = \Omega\big(n^{-2q(G^2)}\big)$, thus $q(G) + q(G^3) \leq 2q(G^2)$. Hence

$$4\bar{r}(G) \overset{(17)}{=} \bar{r}(G) + \bar{r}(G^3) \overset{(15)}{\leq} q(G) + q(G^3) \leq 2q(G^2) \overset{(18)}{=} 2\bar{r}(G^2) \overset{(17)}{=} 4\bar{r}(G). \tag{20}$$

All expressions in (20) must be equal, therefore $\bar{r}(G) = q(G)$.

Finally, for an even $f:\ G \to K_n$ with $\big|V(G)\big| - \big|f(V(G))\big| = q(G)$, we have

$$\frac{1}{2}\big|V(G)\big| \overset{(16)}{\geq} \bar{r}(G) = q(G) = \big|V(G)\big| - \big|f(V(G))\big|,$$

therefore $\big|f\big(V(G)\big)\big| \geq \frac{1}{2}\big|V(G)\big|$. $\square$

5 Computational results

We checked Conjecture 1 for all graphs on at most 9 vertices using the previous results and a computer program. Starting from the list of nonisomorphic graphs, we filtered out those who violated one of our conditions for being a minimal counterexample. In particular we performed the following tests:

1. Check whether the graph is symmetric, by exhaustive search enumerating all possible involutions of the vertices. If the graph is symmetric, it is not a counterexample.

2. Calculate the number of homomorphisms into graphs represented by 1×1, 2×2 or 3×3 matrices of small integers. (Checking 1×1 matrices is just the same as checking whether or not the number of edges is even.) If we get a negative homomorphism count, the graph is negative and therefore it is not a counterexample.

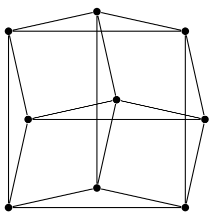
3. Calculate the number of homomorphisms into graphs represented by symbolic 3×3 and 4×4 matrices and perform local minimization on the resulting polynomial from randomly chosen points. Once we reach a negative value, we can conclude that the graph is negative.

4. Partition the vertices of the graph in such a way that two vertices belong to the same class if and only if they produce the same walk-tree (1-dimensional Weisfeiler–Lehman algorithm). Check for all proper subsets of the set of classes whether their union spans an asymmetric subgraph. If we find such a subgraph, the graph is not a minimal counterexample: either the subgraph is not positive and by Corollary 16 the original graph is not positive either, or the subgraph is positive, and therefore we have a smaller counterexample.

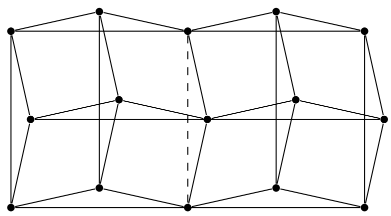
5. Consider only those homomorphisms which map all vertices in the ith class of the partition into vertices $3i+1$, $3i+2$ and $3i+3$ of the target graph represented by a symbolic matrix. If we get a negative homomorphism count, the graph is negative by Proposition 15. (In this case we work with a $3k \times 3k$ matrix where k denotes the number of classes of the walk-tree partition, but the resulting polynomial still has a manageable size because we only count a small subset of homomorphisms. Note that if one of the classes consists of a single vertex, we only need one corresponding vertex in the target graph.)

The tests were performed in such an order that the faster and more efficient ones were run first, restricting the later ones to the set of remaining graphs. For example, in step 4, we start with checking whether any of the classes spans an odd number of edges, or whether the number of edges between any two classes is odd. We used the `SAGE` computer-algebra system for our calculations and rewritten the speed-critical parts in `C` using `nauty` for isomorphism checking, `mpfi` for interval arithmetics and Jean-Sébastien Roy's `tnc` package for nonlinear optimization.

Our automated tests left only one graph on 9 vertices as a possible minimal counterexample, the graph on left:

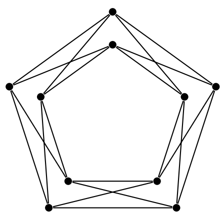


G_1

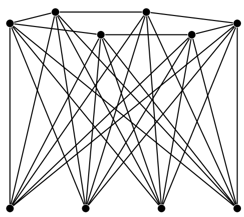


H

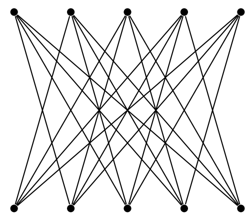
The non-positivity of this graph was checked manually by counting the number of homomorphisms into the graph on the right (where the dashed edge has weight -1 and all other edges have weight 1). This leaves only the following three of the 12 293 435 graphs on at most 10 vertices as candidates for a minimal counterexample:



G_2



G_3



G_4

Note that all three graphs are regular, as is the case for all remaining graphs on 11 vertices. We have found step 5 of the algorithm quite effective at excluding graphs with nontrivial walk-tree partitions.

Acknowledgement. The conjecture in this paper was the subject of a research group at the American Institute of Mathematics workshop "Graph and hypergraph limits", Palo Alto, CA, August 15–19, 2011. We are grateful for the inspiration from all those who took part in the discussions of this research group, in particular to Sergey Norin and Oleg Pikhurko.

Further research on the topic of this paper was supported by ERC Grant No. 227701 and NSF under agreement No. DMS-0835373. Any opinions and conclusions expressed in this material are those of the authors and do not necessarily reflect the views of the NSF or of the ERC.

References

[1] C. Borgs, J.T. Chayes, L. Lovász, V.T. Sós, and K. Vesztergombi: Convergent Graph Sequences I: Subgraph frequencies, metric properties, and testing, *Advances in Math.* **219** (2008), 1801–1851.

[2] B. L. Douglas: The Weisfeiler-Lehman Method and Graph Isomorphism Testing, `http://arxiv.org/abs/1101.5211`

[3] H. Hatami: Graph norms and Sidorenko's conjecture, *Israel J. Math.* **175** (2010), 125–150.

[4] L. Lovász: Large Networks and Graph Limits, *Colloquium Publications* 2012, 475pp.

[5] L. Lovász: Subgraph densities in signed graphons and the local Sidorenko conjecture, `http://arxiv.org/abs/1004.3026`

[6] L. Lovász, B. Szegedy: Limits of dense graph sequences, *J. Comb. Theory B* **96** (2006), 933–957.