TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 375, Number 9, September 2022, Pages 6281–6310
https://doi.org/10.1090/tran/8643
Article electronically published on June 30, 2022

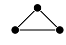
TROPICALIZATION OF GRAPH PROFILES

GRIGORIY BLEKHERMAN, ANNIE RAYMOND, MOHIT SINGH,
AND REKHA R. THOMAS

ABSTRACT. A graph profile records all possible densities of a fixed finite set of graphs. Profiles can be extremely complicated; for instance the full profile of any triple of connected graphs is not known, and little is known about hypergraph profiles. We introduce the *tropicalization* of graph and hypergraph profiles. Tropicalization is a well-studied operation in algebraic geometry, which replaces a variety (the set of real or complex solutions to a finite set of algebraic equations) with its "combinatorial shadow". We prove that the tropicalization of a graph profile is a closed convex cone, which still captures interesting combinatorial information. We explicitly compute these tropicalizations for arbitrary sets of complete and star hypergraphs. We show they are rational polyhedral cones even though the corresponding profiles are not even known to be semialgebraic in some of these cases. We then use tropicalization to prove strong restrictions on the power of the sums of squares method, equivalently Cauchy-Schwarz calculus, to test (which is weaker than certification) the validity of graph density inequalities. In particular, we show that sums of squares cannot test simple binomial graph density inequalities, or even their approximations. Small concrete examples of such inequalities are presented, and include the famous Blakley-Roy inequalities for paths of odd length. As a consequence, these simple inequalities cannot be written as a rational sum of squares of graph densities.

1. Introduction

An important tool in the study of very large graphs is to randomly sample a fixed number of small subgraphs. This methodology goes under various names such as *property testing* [12] and *subgraph sampling* [22, Section 1.3.1], and the sampling statistics are in terms of densities of the subgraphs in the given graph. There are various notions of densities. In this paper we focus on homomorphism densities, but we say a few words at the very end on consequences for other densities. A graph G has vertex set $V(G)$ and edge set $E(G)$, and is assumed to be simple, without loops or multiple edges. The *homomorphism density* of a graph H in a graph G, denoted by $t(H;G)$, is the probability that a random map from $V(H)$ to $V(G)$ is a graph homomorphism, i.e., it maps every edge of H to an edge of G.

The *graph profile* of a collection of connected graphs $\mathcal{U} = \{C_1, \ldots, C_s\}$, denoted as $\mathcal{G}_{\mathcal{U}}$, is the closure of the set of all vectors $(t(C_1;G), t(C_2;G), \ldots, t(C_s;G))$ as G varies over all graphs. For example, the graph profile of $\mathcal{U} = \Big\{$ ,  $\Big\}$ is the

Received by the editors May 21, 2020, and, in revised form, January 19, 2022, and January 19, 2022.

2020 *Mathematics Subject Classification.* Primary 05C35.

The first author was partially supported by NSF grant DMS-1352073. The second author was partially supported by NSF grant DMS-2054404. The third author was partially supported by NSF grant CCF-1717947. The fourth author was partially supported by NSF grant DMS-1719538.

well-known set in $[0,1]^2$ shown in Figure 1 (slightly distorted to better show its features) [31].

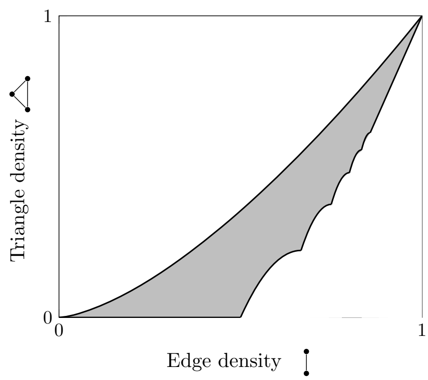


FIGURE 1. The graph profile of edge and triangle

Graph profiles are extremely complicated sets and they have been fully understood in very few cases. The study of graph profiles was initiated in [8], where it was shown that a graph profile is a closed full-dimensional subset of $[0,1]^s$ for an arbitrary s-tuple of connected graphs. However to this day, there is no triple of connected graphs for which the graph profile is fully known. For pairs of graphs, the profile $\mathcal{G}_{\mathcal{U}}$ for $\mathcal{U} = \{ \, , K_n\}$ where K_n denotes the complete graph on n vertices was determined first for $n = 3$ in [31], for $n = 4$ in [27], and for a general n in [33]. Determining the profile of $\{ \, , H\}$ where H is an arbitrary bipartite graph would involve resolving the famous *Sidorenko conjecture* which says $t(H;G) \geq t(\, ;G)^{|E(H)|}$. Despite considerable attention, this conjecture is only known for some classes of bipartite graphs [6,7,19,34]. Some two-dimensional projections of $\mathcal{G}_{\mathcal{U}}$ where $\mathcal{U} = \{ \, , \, , \, \}$ were described in [11,17]. As is evident from Figure 1, graph profiles are not necessarily convex or semialgebraic sets.

In this paper we introduce the *tropicalization* of graph profiles. Tropicalization is a very well-studied operation in real and complex algebraic geometry, which replaces a variety (the set of real or complex solutions to a finite collection of algebraic equations) with its "combinatorial shadow" [24,25]. Tropicalization of real semialgebraic sets has not been explored in as much detail [2,3,35]. As we describe below, while tropicalization loses a lot of information about a graph profile, it also keeps many of its interesting combinatorial properties.

1.1. **Tropicalization of graph profiles.** The first set of results in this paper shows that even though $\mathcal{G}_{\mathcal{U}}$ can be very complicated [16], its tropicalization denoted as $\mathrm{trop}(\mathcal{G}_{\mathcal{U}})$ is, relatively speaking, rather simple. For a set $\mathcal{S} \subseteq \mathbb{R}^s_{\geq 0}$, let $\log(\mathcal{S})$ denote the image of $\mathcal{S} \cap \mathbb{R}^s_{>0}$ under the map $\mathbf{v} \mapsto (\log_e v_1, \ldots, \log_e v_s)$. The tropicalization of $\mathcal{S}$, also known as its *logarithmic limit set*, is

$$\mathrm{trop}(\mathcal{S}) = \lim_{t \to 0} \log_{\frac{1}{t}}(\mathcal{S}).$$

It was shown in [2] that $\mathrm{trop}(\mathcal{S})$ is a closed cone, but it is not necessarily convex. We prove in Theorem 2.4 that $\mathrm{trop}(\mathcal{G}_{\mathcal{U}})$ is a closed convex cone that coincides with the closure of the conical hull of $\log_e(\mathcal{G}_\mathcal{U})$. This result extends beyond graph profiles to hypergraph profiles in Theorem 3.2, and in fact, to any set that has the *Hadamard property*, namely that if $\mathbf{u}, \mathbf{v}$ are in the set then so is their coordinate-wise Hadamard product.

Hypergraph profiles. In Section 3 we compute the tropicalization of an arbitrary k-tuple of complete graphs (and complete uniform hypergraphs) [Theorem 3.3], and an arbitrary k-tuple of star graphs (and uniform star hypergraphs) [Theorem 3.6]. Both tropicalizations are rational polyhedral cones. This is in sharp contrast to the true profile of even any triple of such graphs being quite out of reach at the moment.

Binomial graph density inequalities. The cone $\operatorname{trop}(\mathcal{G}_\mathcal{U})$ provides a perfect framework in which to study *pure binomial graph density inequalities* of the form $t(H_1;G) \geq t(H_2;G)$ where H_1 and H_2 are graphs whose connected components are contained in $\mathcal{U} = \{C_1, \dots, C_s\}$. Under the log map, the inequality $t(H_1;G) \geq t(H_2;G)$ becomes a linear inequality in the densities of its connected components. The extreme rays of the dual cone $\operatorname{trop}(\mathcal{G}_\mathcal{U})^*$ generate all of the pure binomial inequalities valid on $\mathcal{G}_\mathcal{U}$. As mentioned already, for a pair $\mathcal{U} = \{ | , H\}$ where H is bipartite, $\operatorname{trop}(\mathcal{G}_\mathcal{U})$ captures the Sidorenko conjecture for H. In this case, $\operatorname{trop}(\mathcal{G}_\mathcal{U})$ is a two-dimensional cone in $\mathbb{R}^2_{\leq 0}$, and one of its extreme rays for arbitrary H is determined in [4]. Determining the other would resolve the Sidorenko conjecture.

1.2. **Tropicalization and sums of squares.** Our next set of results uses tropicalization to show strong limitations for the *sums of squares (sos)* method, also known as *Cauchy-Schwarz calculus* [10, 14, 22, 30, 32], to prove graph density inequalities.

A finite $\mathbb{R}$-linear combination of graphs $H_1, \dots, H_s$, $a = \sum \alpha_i H_i$, is called a *graph combination*. The evaluation of a graph combination a on a graph G is $a(G) = \sum \alpha_i t(H_i;G)$, and a is said to be nonnegative, written as $a \geq 0$, if $\sum \alpha_i t(H_i;G) \geq 0$ for every G. Equivalently, $a \geq 0$ on the graph profile $\mathcal{G}_\mathcal{U}$ where $\mathcal{U} = \{C_1, \dots, C_m\}$, where $C_1, \dots, C_m$ are the connected components of $H_1, \dots, H_s$. A graph combination a is a *sum of squares (sos)* if $a = \sum [[a_j^2]]$ where a_j is a graph combination of *partially labeled graphs*.

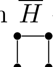
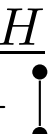
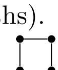
A natural certificate of nonnegativity of a graph combination is a sos expression for it, and *semidefinite programming* can be used to search for a sos expression. It was shown in [23] that every true inequality between homomorphism densities is a limit of Cauchy-Schwarz (sos) inequalities. Problem 17 in [21] asked whether every nonnegative graph combination is a sos and in particular, whether the *Blakley-Roy inequalities*, $P_k \geq |^k$, where P_k is a path of odd length k, can be certified by sos. Problem 21 in [21] asked whether every nonnegative graph combination a can be multiplied by a combination of the form $(1 + b)$ where b is sos so that the product is sos. This would certify the nonnegativity of a.

It was shown in [15] that the problem of verifying the validity of a polynomial inequality (or equivalently of a linear inequality) between homomorphism densities is undecidable. Moreover, they gave explicit examples of nonnegative graph combinations that are not sos answering the first part of Problem 17 in [21]. The class of graph combinations that become sos after multiplication with an sos were called *rational sums of squares* in [15] after Hilbert's 17th problem. The undecidability result was used to show that there exist nonnegative graph combinations that are not rational sos, thus also solving Problem 21, although no explicit example of such graph combinations was presented. In [5], we found small explicit graph density inequalities that cannot be written as a sos, and also do not become sos after multiplication by expressions of the form $1 + b$ where b is sos. A concrete instance of our results is the family of Blakley-Roy inequalities, $P_k \geq$ ^{k}, for odd k, answering Problem 21 and the second part of Problem 17 in [21].

We introduce a new notion of *sos-testable* graph combinations, which are more general than sums of squares and rational sums of squares. Roughly speaking, sos-testable graph combinations correspond to graph combinations whose nonnegativity can be recognized by sums of squares, although there is no explicit certificate of nonnegativity. We find large families of pure binomial graph density inequalities that are not sos-testable, and even their pure binomial approximations remain not sos-testable. These families include Blakley-Roy inequalities for odd paths.

Sos profiles and sos-testable functions. For a fixed positive integer d, define the *d-sos-profile*, denoted as $\mathcal{S}_d$, to be the set of all points on which all sos graph combinations $\sum[[a_j^2]]$, with all a_j having at most d edges in their constituent graphs, are nonnegative. Let the $(\mathcal{U}, d)$-sos profile $\mathcal{S}_{\mathcal{U},d}$ be the projection on $\mathcal{S}_d$ onto the graphs in $\mathcal{U}$. We prove that $\mathcal{S}_d$ is a basic, closed semialgebraic set, and that its tropicalization $\operatorname{trop}(\mathcal{S}_d)$ is a rational polyhedral cone described explicitly in Theorem 4.12.

A graph combination a is *sos-testable* if it is nonnegative on $\mathcal{S}_d$ for some d. Sos-testable functions do not have to come with an explicit certificate of nonnegativity on an sos-profile. However, in principle, since $\mathcal{S}_d$ is a semialgebraic set, nonnegativity of a graph combination on $\mathcal{S}_d$ can be verified via real quantifier elimination. We show in Theorem 4.16 that if a graph combination a becomes sos-testable after multiplication by an sos-testable graph combination b, then a was already sos-testable. The class of sos-testable functions includes sums of squares and also rational sums of squares, but is quite likely significantly larger. It is not clear at this point whether even rational sos is a bigger class than just sos.

Limitations of sos. In Section 5 we exhibit concrete families of binomial graph density inequalities that are not sos-testable, even approximately (Theorem 5.1). Namely, if $\overline{H}$ and $\underline{H}$ are two graphs with the same number of edges where $\overline{H}$ is a *trivial square* (see Section 5 for details) in which every vertex has degree p or $p + 1$, and the maximum degree in $\underline{H}$ is at most $p + 1$, then $\overline{H} - \underline{H}$ is not sos-testable. An example of such a binomial inequality would be  $-$ $^3 \geq 0$ (and in fact all Blakley-Roy inequalities for odd paths). Even more, $\overline{H}^k - \underline{H}^{k+1}$ is not sos-testable for $k \geq 2|E(\overline{H})| + 1$. In particular, ^{7} $-$ ^{24} is not sos-testable and thus cannot be written as a rational sos. The existence of non-sos-testable graph density inequalities follows from the undecidability result in [15]. However, they do not provide explicit examples, and our non-approximation results are new.

Nonnegative graph combinations admit a *Positivstellensatz*: any graph combination strictly positive on a graph profile $\mathcal{G}_{\mathcal{U}}$ is a sos [23, 26]. It follows from this that the graph profile $\mathcal{G}_{\mathcal{U}}$ is the intersection of the $(\mathcal{U}, d)$-sos profiles for all d:

$$\mathcal{G}_{\mathcal{U}} = \bigcap_d \mathcal{S}_{(\mathcal{U},d)}.$$

Perhaps surprisingly, tropicalizations of graph and sos-profiles behave rather differently in that $\mathrm{trop}(\mathcal{S}_{(\mathcal{U},d)})$ need not approach $\mathrm{trop}(\mathcal{G}_\mathcal{U})$ as $d \to \infty$. This is because tropicalizations of graph and sos-profiles only depend on an arbitrarily small neighborhood of the origin in the original set by Lemma 2.2, and sets may approach each other, while their neighborhoods of the origin do not, see for instance Example 2.3. This phenomenon plays an important role in our ability to use tropicalizations to find non-sos-testable functions. It enables us to find graph density inequalities that are not valid on any d-sos-profile while being valid on $\mathcal{G}_\mathcal{U}$.

1.3. **Open questions.** We now state some open questions raised by our results. All tropicalizations of graph profiles computed in Section 3 are rational polyhedral cones. Therefore it is natural to ask the following:

Question 1.1. Is $\mathrm{trop}(\mathcal{G}_\mathcal{U})$ a polyhedral cone for any collection $\mathcal{U}$ of connected graphs? If yes, then is it necessarily a rational polyhedral cone?

It was shown in [15] that the problem of deciding the validity of a polynomial inequality (or equivalently of a linear inequality) between homomorphism densities is undecidable. This is equivalent to saying that verifying the validity of a polynomial inequality on a graph profile is undecidable. Tropicalizations of graph profiles only carry information about pure binomial inequalities, and tropicalizations appear to be simpler than the full profile. Therefore we ask the following:

Question 1.2. Given two (not necessarily connected) graphs G_1 and G_2 is the question of whether $G_1 - G_2 \geq 0$ is a valid homomorphism density inequality decidable?

Question 1.2 is equivalent to understanding whether a given integer (or rational) point lies in the dual cone $\mathrm{trop}(G_\mathcal{U})^*$, where $\mathcal{U}$ is the set of connected components of G_1 and G_2.

1.4. **Organization of this paper.** In Section 2 we study the tropicalization $\mathrm{trop}(\mathcal{S})$ of a set $\mathcal{S} \subseteq \mathbb{R}^s_{\geq 0}$ with the Hadamard property. We prove in Lemma 2.2 that $\mathrm{trop}(\mathcal{S})$ is a closed convex cone that coincides with the closure of the conical hull of $\log_e(\mathcal{S})$ and that if the all-ones vector is present in $\mathcal{S}$, then $\mathrm{trop}(\mathcal{S})$ also coincides with the closure of the convex hull of $\log_e(\mathcal{S})$. Theorem 2.4 applies these results to graph profiles, proving that $\mathrm{trop}(\mathcal{G}_\mathcal{U})$ is a closed convex cone that coincides with both the conical and convex hull of $\log(\mathcal{G}_\mathcal{U})$.

Lemma 2.2 can also be applied to hypergraph profiles (Theorem 3.2). In Section 3 we compute the tropicalizations of the profiles of an arbitrary number of hypergraphs from two families – complete hypergraphs and star hypergraphs. Both examples yield rational polyhedral cones that can be described explicitly. Their actual profiles are currently unknown.

We introduce the d-sos-profile $\mathcal{S}_d$ in Section 4 and prove that it is a basic closed semialgebraic set. Theorem 4.12 proves that $\mathrm{trop}(\mathcal{S}_d)$ is a rational polyhedral cone whose inequalities can be described explicitly by the 2×2 principal minors of a symbolic matrix. Graph profiles are contained in sos-profiles. This section also introduces the notion of sos-testable graph density inequalities. We prove in Theorem 4.16 that if b and ab are sos-testable then so is a. In particular, if a is not sos-testable it is neither a sos nor a rational sos (Corollary 4.17).

In Section 5 we find explicit families of pure binomial inequalities that are not sos-testable, even approximately.

2. Tropicalization of graph profiles

Definition 2.1. Let $\mathcal{U} = \{C_1, \ldots, C_s\}$ be a collection of connected graphs. The *graph-profile* of $\mathcal{U}$, denoted as $\mathcal{G}_{\mathcal{U}}$, is the closure of the set of vectors $(t(C_1; G), t(C_2; G), \cdots, t(C_s; G))$ as G varies over all unlabeled graphs.

For any $\mathcal{U}$, the graph profile $\mathcal{G}_{\mathcal{U}}$ is contained in $[0,1]^s$. They are highly complicated objects and very few of them are known explicitly. Note from Figure 1 that neither the graph profile, nor its convex hull, may be semialgebraic.

In this section we use tropical geometry to pass from the complicated graph profile $\mathcal{G}_{\mathcal{U}}$ to its *tropicalization*, which as we will see is a cone, and hence much easier to understand.

Let $\log : \mathbb{R}^s_{>0} \to \mathbb{R}^s$ be defined as $\log(\mathbf{v}) := (\log_e(v_1), \ldots, \log_e(v_s))$. If we need to change the base of the log from e to α then we will explicitly write $\log_\alpha$. For a set $\mathcal{S} \subseteq \mathbb{R}^s_{\geq 0}$ we define $\log(\mathcal{S}) := \log(\mathcal{S} \cap \mathbb{R}^s_{>0})$. For any set $\mathcal{S} \subseteq \mathbb{R}_{\geq 0}$ and $\tau \in (0,1)$ consider

$$\log_{\frac{1}{\tau}}(\mathcal{S}) = \frac{-1}{\log_e \tau} \log_e(\mathcal{S}).$$

The tropicalization of $\mathcal{S}$, which is also called the *logarithmic limit set* of $\mathcal{S}$, is

$$\operatorname{trop}(\mathcal{S}) := \lim_{\tau \to 0} \log_{\frac{1}{\tau}}(\mathcal{S}).$$

By [2, Proposition 2.2], $\operatorname{trop}(\mathcal{S})$ is a closed cone in $\mathbb{R}^s$. A working definition of what it means for a point $\mathbf{y}$ to lie in $\operatorname{trop}(\mathcal{S})$ is that for a sequence $\tau_k \in (0, \epsilon)$ indexed by $k \in \mathbb{N}$ converging to 0 (equivalently, any such sequence), there exists a sequence $\mathbf{y}(k) \in \mathcal{S} \cap \mathbb{R}^s_{>0}$ such that $\log_{\frac{1}{\tau_k}} \mathbf{y}(k) \to \mathbf{y}$ as $\tau_k \to 0$ [2, Proposition 2.1]. In other words, $\operatorname{trop}(\mathcal{S})$ consists of all accumulation points under the map $\log_{1/\tau_k}$ applied to $\mathcal{S} \cap \mathbb{R}^s_{>0}$ for all possible choices of sequences of bases τ_k. Note that since the log map is only defined on positive points, $\log_{\frac{1}{\tau_k}} \mathbf{y}(k)$ can exist only if $\mathbf{y}(k) \in \mathbb{R}^s_{>0}$. For sets $\mathcal{S} \subseteq \mathbb{R}^s_{\geq 0}$ such that $\mathcal{S} = \operatorname{cl}(\mathcal{S} \cap \mathbb{R}^s_{>0})$, $\operatorname{trop}(\mathcal{S})$ does not lose information carried by the points with zero coordinates in $\mathcal{S}$. The *Hadamard product* of $\mathbf{v}, \mathbf{w} \in \mathbb{R}^s$ is defined to be $\mathbf{v} \cdot \mathbf{w} = (v_1 w_1, \ldots, v_s w_s)$. We say that a set $\mathcal{S} \subseteq \mathbb{R}^s_{\geq 0}$ has the *Hadamard property* if for any two vectors $\mathbf{v}, \mathbf{w} \in \mathcal{S}$, $\mathbf{v} \cdot \mathbf{w} \in \mathcal{S}$. For a vector $\mathbf{v}$ and a positive integer k, define the kth power of $\mathbf{v}$ to be $\mathbf{v}^k := (v_1^k, \ldots, v_s^k)$.

Lemma 2.2. *Suppose $\mathcal{S} \subseteq \mathbb{R}^s_{\geq 0}$ has the Hadamard property. Then*

(1) $\operatorname{trop}(\mathcal{S})$ *is the closure of the union of all the rays from the origin through points in* $\log(\mathcal{S})$.
(2) $\operatorname{trop}(\mathcal{S})$ *is a closed convex cone and* $\operatorname{trop}(\mathcal{S}) = \operatorname{cl}(\operatorname{cone}(\log(\mathcal{S})))$, *where* $\operatorname{cone}(\cdot)$ *denotes the conical hull.*
(3) *If* $\mathbf{1} \in \mathcal{S}$, $\operatorname{trop}(\mathcal{S}) = \operatorname{cl}(\operatorname{conv}(\log(\mathcal{S})))$.
(4) *If* $\mathcal{S} \subseteq [0,1]^s$ *and* $\mathcal{S} = \operatorname{cl}(\mathcal{S} \cap [0,1)^s)$, *then for any* $\epsilon > 0$, $\operatorname{trop}(\mathcal{S})$ *is determined by the nonempty neighborhood* $\mathcal{S} \cap B(\mathbf{0}, \epsilon)$ *of* $\mathcal{S}$, *where* $B(\mathbf{0}, \epsilon)$ *is the ball with center* $\mathbf{0}$ *and radius* ϵ.

Proof.

(1) We first show that $\operatorname{trop}(\mathcal{S})$ is contained in the closure of the union of all the rays from the origin through points in $\log(\mathcal{S})$. Suppose $\mathbf{y} \in \operatorname{trop}(\mathcal{S})$. Then there exist sequences $\mathbf{y}(k) \in \mathcal{S} \cap \mathbb{R}^s_{>0}$ and $\tau_k \in (0, \epsilon)$ such that as $\tau_k \to 0$,

$$\log_{\frac{1}{\tau_k}} \mathbf{y}(k) = \frac{-1}{\log \tau_k} \log \mathbf{y}(k) \to \mathbf{y}.$$

Since $\log \mathbf{y}(k) \in \log(\mathcal{S})$ and $\frac{-1}{\log \tau_k} > 0$, we get that $\log_{\frac{1}{\tau_k}} \mathbf{y}(k)$ is in the union of all the rays from the origin through points in $\log(\mathcal{S})$ for all k, and hence, $\mathbf{y}$ is in their closure.

For the other inclusion, we want to show that, for any $\mathbf{v} \in \mathcal{S}$, the ray generated by $\log(\mathbf{v})$ is in $\operatorname{trop}(\mathcal{S})$. Since $\operatorname{trop}(\mathcal{S})$ is a cone and changing bases just rescales $\log(\mathbf{v})$, it suffices to show that $\log_e(\mathbf{v}) \in \operatorname{trop}(\mathcal{S})$. Note that $\log_e(\mathbf{v}) = \log_{e^k}(\mathbf{v}^k)$ for any $k \in \mathbb{N}$ and since $\mathcal{S}$ has the Hadamard property, $\mathbf{v}^k$ is also in $\mathcal{S}$ for all positive integers k. Thus there are sequences $\tau_k := \frac{1}{e^k}$ and $\mathbf{v}(k) := \mathbf{v}^k$ where $\tau_k \to 0$ and $\log_{\frac{1}{\tau_k}}(\mathbf{v}(k)) = \log_e(\mathbf{v})$, and so $\log_e(\mathbf{v}) \in \operatorname{trop}(\mathcal{S})$. Since $\operatorname{trop}(\mathcal{S})$ is closed, we can conclude that the closure of the union of all the rays from the origin through points in $\log(\mathcal{S})$ is contained in $\operatorname{trop}(\mathcal{S})$.

(2) We already know that $\operatorname{trop}(\mathcal{S})$ is closed. We now show that it is a convex cone. By (1), it suffices to show that for any $\log(\mathbf{v}), \log(\mathbf{w}) \in \log(\mathcal{S})$, $\alpha \log(\mathbf{v}) + \beta \log(\mathbf{w}) \in \operatorname{trop}(\mathcal{S})$ for any $\alpha, \beta \geq 0$. We can assume that $\alpha, \beta \in \mathbb{Q}$ since $\operatorname{trop}(\mathcal{S})$ is closed, and so we can further assume that α, β have the same denominator, say $\alpha = \frac{a_1}{b}$ and $\beta = \frac{a_2}{b}$ where $a_1, a_2, b \in \mathbb{N}$. From (1), we only need to show that $a_1 \log(\mathbf{v}) + a_2 \log(\mathbf{w}) \in \operatorname{trop}(\mathcal{S})$. This is equal to $\log(\mathbf{v}^{a_1} \cdot \mathbf{w}^{a_2})$ which, by the Hadamard property, is contained in $\log(\mathcal{S})$, and, from (1), thus also in the tropicalization.

From (1), we know that $\operatorname{trop}(\mathcal{S})$ is contained in the closure of the union of all the rays from the origin through points in $\log(\mathcal{S})$, which is contained in $\operatorname{cl}(\operatorname{cone}(\log(\mathcal{S})))$. To show the reverse inclusion, by (1), we know $\operatorname{trop}(\mathcal{S}) \supseteq \log(\mathcal{S})$, and since $\operatorname{trop}(\mathcal{S})$ is a closed convex cone, $\operatorname{trop}(\mathcal{S}) \supseteq \operatorname{cl}(\operatorname{cone}(\log(\mathcal{S})))$, and the result holds.

(3) If $\mathbf{1} \in \mathcal{S}$, then $\mathbf{0} = \log \mathbf{1} \in \log(\mathcal{S})$. We already saw that all positive integer multiples of a point in $\log(\mathcal{S})$ are also in $\log(\mathcal{S})$. Together these facts imply that $\operatorname{cl}(\operatorname{conv}(\log(\mathcal{S}))) = \operatorname{cl}(\operatorname{cone}(\log(\mathcal{S}))) = \operatorname{trop}(\mathcal{S})$.

(4) Since $\operatorname{trop}(\mathcal{S})$ is a closed convex cone, it is fully described by the linear inequalities that are valid on it. A linear inequality $\mathbf{a}^\top \mathbf{y} \geq \mathbf{b}^\top \mathbf{y}$ valid on $\log(\mathcal{S})$ corresponds to the binomial inequality $\mathbf{x}^{\mathbf{a}} \geq \mathbf{x}^{\mathbf{b}}$ on $\mathcal{S} \cap \mathbb{R}^s_{>0}$. Here we are setting $y_i = \log(x_i)$. Therefore, to prove the claim it suffices to argue that if a binomial inequality is valid on a small neighborhood of the origin in $\mathcal{S} \cap \mathbb{R}^s_{>0}$ then it is in fact valid on all of $\mathcal{S} \cap \mathbb{R}^s_{>0}$. Suppose there is some binomial inequality $\mathbf{x}^{\mathbf{a}} \geq \mathbf{x}^{\mathbf{b}}$ that is valid on the neighborhood and $\mathbf{v} \in \mathcal{S} \cap \mathbb{R}^s_{>0}$ violates it. Without loss of generality, we can assume that all components of $\mathbf{v}$ are less than 1. Then there is some large enough positive integer k for which $\mathbf{v}^k$ lies in the neighborhood we considered. If $\mathbf{v}^{\mathbf{a}} < \mathbf{v}^{\mathbf{b}}$ then we also have $(\mathbf{v}^k)^{\mathbf{a}} < (\mathbf{v}^k)^{\mathbf{b}}$ which is a contradiction. Thus $\operatorname{trop}(\mathcal{S})$ is determined by the behavior of $\mathcal{S}$ near the origin. □

Note that even though two sets might converge, their tropicalizations might not as seen in Example 2.3. This example also highlights the role of the neighborhood of the origin as in Lemma 2.2(4).

Example 2.3. Consider the two sets in Figure 2. On the left is a triangle and on the right a slight modification of the triangle into a quadrilateral with new vertex $(0, \varepsilon)$. The neighborhood of the origin is different for the two sets. Observe that $\mathcal{S} = \lim_{\varepsilon \to 0} \mathcal{S}_\varepsilon$ as is clear from Figure 2. On the other hand, their tropicalizations,

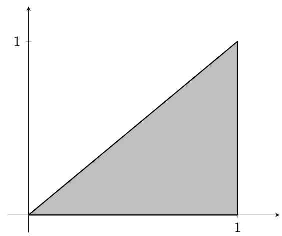


FIGURE 2. The sets $\mathcal{S}$ (left) and $\mathcal{S}_\varepsilon$ (right)

seen in Figure 3, do not, i.e., $\operatorname{trop}(\mathcal{S}) \neq \lim_{\varepsilon\to 0} \operatorname{trop}(\mathcal{S}_\varepsilon)$ since the neighborhood of the origin is different in sets $\mathcal{S}$ and $\mathcal{S}_\varepsilon$ for any $\varepsilon > 0$.

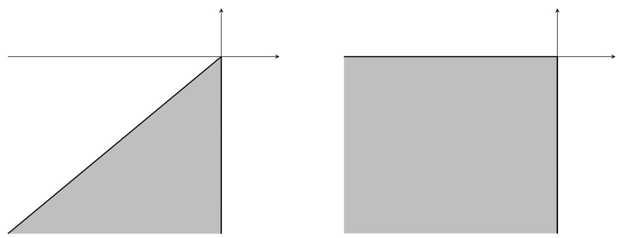


FIGURE 3. The tropicalizations of $\mathcal{S}$ (left) and $\mathcal{S}_\varepsilon$ (right)

We now apply these results to the graph profile $\mathcal{G}_\mathcal{U}$ which is known to be a connected and full-dimensional set [8]. Moreover, it is known that every $\mathbf{v} \in \mathcal{G}_\mathcal{U}$ is arbitrarily close to $(t(C_1; G), \ldots, t(C_s; G))$ for some graph G. As we will see in the proof of Theorem 4.16, one can argue that every neighborhood of $\mathbf{v}$ has a full-dimensional ball containing $\mathbf{v}$ that is contained in $\mathcal{G}_\mathcal{U}$. Therefore, there is a positive point in $\mathcal{G}_\mathcal{U}$ arbitrarily close to $\mathbf{v}$, and no information is lost by passing to $\log(\mathcal{G}_\mathcal{U})$. Also, since $\mathcal{G}_\mathcal{U}$ is contained in $[0,1]^s$, $\log(\mathcal{G}_\mathcal{U})$ and $\operatorname{trop}(\mathcal{G}_\mathcal{U})$ lie in $\mathbb{R}^s_{\leq 0}$.

Theorem 2.4. *For the graph profile $\mathcal{G}_\mathcal{U}$,*

$$\operatorname{trop}(\mathcal{G}_\mathcal{U}) = \operatorname{cl}(\operatorname{cone}(\log(\mathcal{G}_\mathcal{U}))) = \operatorname{cl}(\operatorname{conv}(\log(\mathcal{G}_\mathcal{U}))) \subseteq \mathbb{R}^s_{\leq 0}.$$

Further, for any $\epsilon > 0$, $\operatorname{trop}(\mathcal{G}_\mathcal{U})$ is determined by the nonempty neighborhood $\mathcal{G}_\mathcal{U} \cap B(\mathbf{0}, \epsilon)$ where $B(\mathbf{0}, \epsilon)$ is the ball with center $\mathbf{0}$ and radius ϵ.

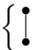
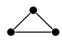
Proof. By Lemma 2.2 we just need to show that $\mathcal{G}_\mathcal{U}$ has the Hadamard property and contains $\mathbf{1}$. Equation 5.30 in [22] implies that $t(C_i; G)t(C_i; G') = t(C_i; G \times G')$ where $G \times G'$ is the categorical product of G and G'. Therefore, if $(t(C_1; G), \ldots, t(C_s; G)) \in \mathcal{G}_\mathcal{U}$ and $(t(C_1; G'), \ldots, t(C_s; G')) \in \mathcal{G}_\mathcal{U}$, then $(t(C_1; G)t(C_1; G'), \ldots, t(C_s; G)t(C_s; G'))$ $\in \mathcal{G}_\mathcal{U}$.

For the sequence of complete graphs K_n, $(t(C; K_n) : C \in \mathcal{U}) \to \mathbf{1}$ as $n \to \infty$ which implies that $\mathbf{0} = \log(\mathbf{1})$ lies in the closure of the convex hull of $\log(\mathcal{G}_\mathcal{U})$. $\square$

Theorem 2.4 implies that we obtain a great simplification in structure when we pass from the graph profile $\mathcal{G}_\mathcal{U}$ to its tropicalization, $\operatorname{trop}(\mathcal{G}_\mathcal{U})$, which is a closed convex cone. A natural next question is the following:

Question 2.5. Is $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$ a polyhedral cone for any collection $\mathcal{U}$ of connected graphs? If yes, then is it necessarily a rational polyhedral cone?

If $\mathcal{U}$ contains two graphs, then $\operatorname{trop}(\mathcal{S})$ is indeed a polyhedral cone with two extreme rays since it is a two-dimensional cone. Could the generators of the extreme rays be non-rational? No graph profile for three graphs is known. Is there a graph profile for three graphs for which $\operatorname{trop}(\mathcal{S})$ is not polyhedral?

Example 2.6. For $\mathcal{U} = \{$ ,  $\}$, we saw the graph profile $\mathcal{G}_{\mathcal{U}}$ in Figure 1.

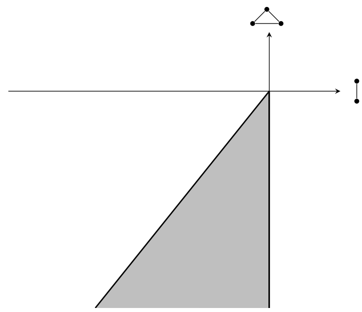
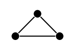
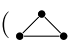
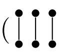
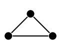
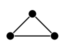
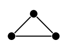


FIGURE 4. The tropicalization of the graph profile of an edge and a triangle

The curve bounding the upper part of the profile is $^3 \geq$ ^{2} which stands for the density inequality $t($$^3; G) \geq t($$^2; G)$, or equivalently, $t($$; G) \geq t($ $; G)$ for all unlabeled graphs G. Moreover, we know that  ≤ 1. Let $y_1 := \log$ and $y_2 := \log$ . Then the above binomial inequalities correspond under the log map to the linear inequalities

$$3y_1 - 2y_2 \geq 0 \text{ and } y_1 \leq 0$$

which form the cone generated by the rays $(-2, -3)$ and $(0, -1)$ in $\mathbb{R}^2_{\leq 0}$ as shown in Figure 4. In Section 3, we will see that this cone is indeed $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$.

Remark 2.7. Note that $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$ can be understood as an object recording all possible orders of growths of densities, i.e., recording whether there exists a sequence of graphs converging to that order of growth.

3. EXPLICIT TROPICALIZATION OF CLIQUES AND STARS

In this section, we explicitly compute $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$ for a finite collection $\mathcal{U}$ from two hypergraph families, cliques and stars. In both cases we will see that the tropicalizations are rational polyhedral simplicial cones.

Let $K_p^{(r)}$ be the complete r-uniform hypergraph on p vertices, i.e., the graph on p vertices where every set of r vertices forms a (hyper)edge. When $r = 2$, $K_p^{(2)}$ is simply the complete graph on p vertices. In Section 3.1, we describe $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$ when $\mathcal{U} = \{K_r^{(r)}, K_{r+1}^{(r)}, \ldots, K_l^{(r)}\}$ is the collection of complete r-uniform hypergraphs for any $r \geq 2$ and $l \geq r$.

Let $S^{(r)}(b, c)$ denote the r-uniform hypergraph with b edges all of which intersect in some set of c vertices, and nowhere else. We call such graphs, *stars with b* *branches*, and we call the intersection of all the edges the *center*. For example, $S^{(2)}(b,1) = K_{1,b}$, the complete bipartite graph with parts of size 1 and b. Note that $S^{(r)}(b,c)$ has $b(r-c)+c$ vertices. In Section 3.2, we describe $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$ when $\mathcal{U} = \{S^{(r)}(1,c), S^{(r)}(2,c), \ldots, S^{(r)}(l,c)\}$ for any $r \geq 2$, $c \leq r$ and $b \geq 1$.

We begin by showing that Theorem 2.4 also holds for r-uniform hypergraphs. We first define what we mean by the product of two hypergraphs.

Definition 3.1. Let G_1 and G_2 be two r-uniform hypergraphs. The *direct product* of G_1 and G_2 is $G_1 \times G_2$ where $V(G_1 \times G_2) = \{uv : u \in V(G_1) \text{ and } v \in V(G_2)\}$, and

$$\begin{aligned} E(G_1 \times G_2) = \{\{u_1v_1, u_2v_2, \ldots, u_rv_r\} : \{u_1, \ldots, u_r\} \in E(G_1) \\ \text{and } \{v_1, \ldots, v_r\} \in E(G_2)\}\,. \end{aligned}$$

Note that any pair of edges $\{u_1, \ldots, u_r\} \in E(G_1)$ and $\{v_1, \ldots, v_r\} \in E(G_2)$ gives rise to $r!$ different edges in $E(G_1 \times G_2)$.

Theorem 3.2. *For an r-uniform hypergraph profile $\mathcal{G}_{\mathcal{U}}$, where $|\mathcal{U}| = s$,*

$$\operatorname{trop}(\mathcal{G}_{\mathcal{U}}) = \operatorname{cl}(\operatorname{cone}(\log(\mathcal{G}_{\mathcal{U}}))) = \operatorname{cl}(\operatorname{conv}(\log(\mathcal{G}_{\mathcal{U}}))) \subseteq \mathbb{R}^s_{\leq 0}.$$

Further, for any $\epsilon > 0$, $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$ is determined by the nonempty neighborhood $\mathcal{G}_{\mathcal{U}} \cap B(\mathbf{0}, \epsilon)$ where $B(\mathbf{0}, \epsilon)$ is the ball with center $\mathbf{0}$ and radius ϵ.

Proof. We first show that $\mathcal{G}_{\mathcal{U}}$ has the Hadamard property, i.e., that

$$t(H;G)t(H;G') = t(H; G \times G')$$

for all r-uniform hypergraphs H, G, G'. Since the total number of maps from $V(H)$ to $V(G \times G')$ is $(|V(G)||V(G')|)^{|V(H)|}$, the denominators on both sides of the equation are the same. For homomorphisms $\varphi : V(H) \to V(G)$ and $\varphi' : V(H) \to V(G')$, the map $\psi : V(H) \to V(G \times G')$ such that $\psi(v) \mapsto \varphi(v)\varphi'(v)$ is also a homomorphism. Conversely, for a homomorphism $\psi : V(H) \to V(G) \times V(G')$, the projections $\varphi = \psi_G : V(H) \to V(G)$ and $\varphi' = \psi_{G'} : V(H) \to V(G')$ onto the two components $V(G)$ and $V(G')$ are homomorphisms. Thus, the numerators on both sides of the equation are also the same. Therefore, if $(t(C_1;G), \ldots, t(C_s;G)) \in \mathcal{G}_{\mathcal{U}}$ and $(t(C_1;G'), \ldots, t(C_s;G')) \in \mathcal{G}_{\mathcal{U}}$, then we have that $(t(C_1;G)t(C_1;G'), \ldots, t(C_s;G)t(C_s;G')) \in \mathcal{G}_{\mathcal{U}}$.

We now show that $\mathcal{G}_{\mathcal{U}}$ contains $\mathbf{1}$. For the sequence of complete r-uniform hypergraphs $K_n^{(r)}$, $(t(C;K_n^{(r)}) : C \in \mathcal{U}) \to \mathbf{1}$ as $n \to \infty$ which implies that $\mathbf{0} = \log(\mathbf{1})$ lies in the closure of the convex hull of $\log(\mathcal{G}_{\mathcal{U}})$.

The statement of the theorem now follows from Lemma 2.2. □

3.1. **Tropicalization of hypergraph clique profiles.** Consider the family $\mathcal{U} = \{K_r^{(r)}, K_{r+1}^{(r)}, \ldots, K_l^{(r)}\}$ of complete hypergraphs for some $r \geq 2$ and $l \geq r$. Observe that both $\mathcal{G}_{\mathcal{U}}$ and $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$ lie in $\mathbb{R}^{l-r+1}$ where the i^{th} coordinate corresponds to the graph $K_{r+i-1}^{(r)}$ for any $1 \leq i \leq l-r+1$. In Theorem 3.3, we describe the facets and extreme rays of $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$. For any $\mathbf{y} \in \mathbb{R}^{l-r+1}$, denote by $y_{K_i^{(r)}}$ the coordinate of $\mathbf{y}$ indexed by $K_i^{(r)}$ for $r \leq i \leq l$. Also, define $\mathbf{1}_{K_i^{(r)}}$ for some $r \leq i \leq l$ to be the point in $\mathbb{R}^{l-r+1}$ with 1 in the coordinate labeled by $K_i^{(r)}$ and 0 otherwise.

Theorem 3.3. *Let* $\mathcal{U} = \{K_r^{(r)}, K_{r+1}^{(r)}, \ldots, K_l^{(r)}\}$, *then*

$$\operatorname{trop}(\mathcal{G}_{\mathcal{U}}) = \left\{ \mathbf{y} \in \mathbb{R}^{l-r+1} \;:\; \begin{array}{l} y_{K_r^{(r)}} \leq 0, \\ (r+i) y_{K_{r+i-1}^{(r)}} - (r+i-1) y_{K_{r+i}^{(r)}} \geq 0 \quad \forall\, 1 \leq i \leq l-r \end{array} \right\}.$$

Moreover, the extreme rays of $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$ *are*

$$\mathbf{u}_i = -\sum_{j=i}^{l-r+1} (r+j-1) \mathbf{1}_{K_{r+j-1}^{(r)}}$$

for $1 \leq i \leq l-r+1$.

Proof. Let C be the cone on the right hand side of the equation in the theorem. First observe that $t(K_r^{(r)}, G) \leq 1$ is a valid inequality for all graphs G and thus the inequality $y_{K_r^{(r)}} \leq 0$ is valid for $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$. The Kruskal-Katona theorem [18,20] (see also [17]) in the context of graph homomorphism densities of complete graphs implies that for any integers $r \leq p < q$, $\left(t(K_p^{(r)}, G)\right)^q - \left(t(K_q^{(r)}, G)\right)^p \geq 0$ is valid for each graph G. This binomial inequality for $\mathcal{G}_{\mathcal{U}}$ implies that the inequality $q y_{K_p^{(r)}} - p y_{K_q^{(r)}} \geq 0$ is valid for $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$ for each $r \leq p < q$. Using the inequalities for each $r \leq p \leq l-1$ and $q = p+1$, we obtain that all the inequalities describing C are valid for $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$. Thus $\operatorname{trop}(\mathcal{G}_{\mathcal{U}}) \subseteq C$.

Before we prove the other containment, we show that extreme rays of C are as claimed in the theorem.

Lemma 3.4. *The extreme rays of* C *are* $\mathbf{u}_i = -\sum_{j=i}^{l-r+1} (r+j-1) \mathbf{1}_{K_{r+j-1}^{(r)}}$ *for* $i \in \{1, \ldots, l-r+1\}$.

Proof. We have $C = \{\mathbf{y} \in \mathbb{R}^{l-r+1} : M\mathbf{y} \geq 0\}$ where

$$M := \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ (r+1) & -r & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & (r+2) & -(r+1) & 0 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \\ 0 & 0 & 0 & 0 & 0 & \cdots & (l-1) & -(l-2) & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & l & -(l-1) \end{bmatrix}.$$

Since M is lower triangular square matrix of size $l-r+1$, with non-zero diagonal, it is invertible. The candidate extreme rays can be obtained by setting a subset of $l-r$ constraints at equality or equivalently, all but one of the constraints at equality. Let M_{-i} denote the matrix obtained after removing the i^{th} row of M. Then a simple check shows that solutions $M_{-i}\mathbf{y} = 0$ are exactly $\{\lambda \cdot \mathbf{u}_i : \lambda \in \mathbb{R}\}$. Since $\mathbf{u}_i \in C$, we obtain that it is an extreme ray. Since these are all the candidate extreme rays, we have the lemma. ∎

To prove $C \subseteq \operatorname{trop}(\mathcal{G}_{\mathcal{U}})$, it is enough to show that the extreme rays of C are contained in $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$ as shown in Lemma 3.5.

Lemma 3.5. *The vectors* $\mathbf{u}_i$ *are in* $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$ *for every* $i \in \{1, \ldots, l-r+1\}$.

Proof. Let $T_{m,k}^{(r)}$ be an r-uniform hypergraph on m vertices partitioned into k parts as equal in size as possible where every r vertices coming from r different parts form an edge (when $r = 2$, this graph is the Turán graph on m vertices with k parts). If

$k < r$, then $T^{(r)}_{m,k}$ is the empty graph on m vertices. Note that $T^{(r)}_{m,k}$ contains cliques of size k (and less), but no clique of size $k+1$ (or more).

Fix some $i \in \{1, \dots, l-r+1\}$. For any $\alpha \in (0,1)$ such that $\alpha n \in \mathbb{N}$, consider the hypergraph G where αn vertices form a clique and where the remaining $(1-\alpha)n$ vertices form a $T^{(r)}_{(1-\alpha)n,r+i-2}$ graph (see Figure 5). Note that $t\left(K^{(r)}_j; G\right) = \alpha^j + \frac{(r+i-2)!(1-\alpha)^j}{(r+i-2-j)!(r+i-2)^j} + O\left(\frac{1}{n}\right)$ for any $r \le j \le r+i-2$ (since cliques of those sizes can be found both in the αn-clique and in $T^{(r)}_{(1-\alpha)n,r+i-2}$) and that $t\left(K^{(r)}_j; G\right) = \alpha^j + O\left(\frac{1}{n}\right)$ for any $r+i-1 \le j \le l$ (since cliques of those size can only be found in the αn clique of G). For instance, to see the latter, note that $t\left(K^{(r)}_j; G\right) = \frac{\binom{\alpha n}{j} j!}{n^j}$ for any $r+i-1 \le j \le l$.

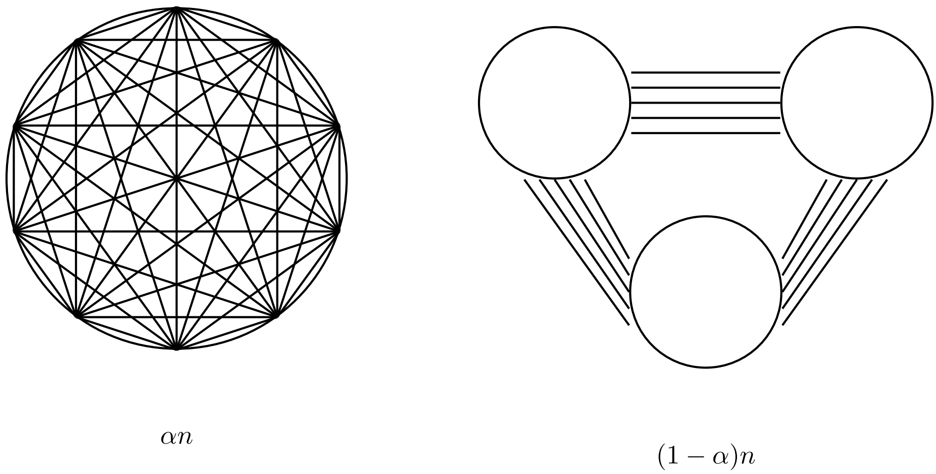


FIGURE 5. G when $i = 3$, $r = 2$

Now consider the vector $\mathbf{v} \in \mathbb{R}^{l+r-1}$ such that $v_h = \left(\frac{\log(t(K^{(r)}_{r+h-1};G))}{\log \alpha}\right)$ for each $1 \le h \le l-r+1$.As $\alpha \to 0$ and $n \to \infty$, we have $v_h \to 0$ for all $1 \le h \le i-1$ and $v_h \to (r+h-1)$ for $i \le h \le l-r+1$. Since this limit point is exactly $\mathbf{u}_i$ and in $\mathrm{trop}(\mathcal{G}_{\mathcal{U}})$, we have the lemma. □

This completes the proof of Theorem 3.3. □

Note that one can find $\mathrm{trop}(\mathcal{G}_{\mathcal{U}'})$ for $\mathcal{U}' \subset \mathcal{U}$ by projecting down $\mathrm{trop}(\mathcal{G}_{\mathcal{U}})$ on the appropriate coordinates. Moreover, a consequence of Theorem 3.3 is that any valid binomial inequality for $\mathcal{G}_{\mathcal{U}}$ is implied by the Kruskal-Katona inequalities and $t(K^{(r)}_r; G) \le 1$.

3.2. **Tropicalization of star hypergraph profiles.** We now give the tropicalization of a collection of generalized stars $\{S^{(r)}(1,c), S^{(r)}(2,c), \dots, S^{(r)}(l,c)\}$. For any $\mathbf{y} \in \mathbb{R}^l$, denote by $y_{S^{(r)}(b,c)}$ the coordinate of $\mathbf{y}$ indexed by $S^{(r)}(b,c)$ for $1 \le b \le l$. Also, define $\mathbf{1}_{S^{(r)}(b,c)}$ for some $1 \le b \le l$ to be the point in $\mathbb{R}^l$ with 1 in the coordinate labeled by $S^{(r)}(b,c)$ and 0 otherwise. Finally, let $d_{v_1,v_2,\ldots,v_k} = |\{e \in E(G) | v_1, v_2, \ldots, v_k \in e\}|$ be the *(common) degree of a set of vertices* $v_1, \ldots, v_k$.

Theorem 3.6. *Let* $\mathcal{U} = \{S^{(r)}(1,c), S^{(r)}(2,c), \ldots, S^{(r)}(l,c)\}$, *then*

$$\operatorname{trop}(\mathcal{G}_{\mathcal{U}}) = \left\{\mathbf{y} \in \mathbb{R}^l : \mathbf{a}_b^\top \mathbf{y} \geq 0, \ \forall\, 1 \leq b \leq l\right\},$$

where

$$\begin{aligned}
\mathbf{a}_1^\top \mathbf{y} &= -2y_{S^{(r)}(1,c)} + y_{S^{(r)}(2,c)}, \\
\mathbf{a}_b^\top \mathbf{y} &= y_{S^{(r)}(b-1,c)} - 2y_{S^{(r)}(b,c)} + y_{S^{(r)}(b+1,c)} \qquad \text{for } 2 \leq b \leq l-1, \text{ and} \\
\mathbf{a}_l^\top \mathbf{y} &= y_{S^{(r)}(l-1,c)} - y_{S^{(r)}(l,c)}.
\end{aligned}$$

Moreover, the extreme rays of $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$ *are* $\mathbf{u}_1, \ldots, \mathbf{u}_l$ *where*

$$\mathbf{u}_b = -\sum_{j=1}^{b} j\mathbf{1}_{S^{(r)}(j,c)} - \sum_{j=b+1}^{l} b\mathbf{1}_{S^{(r)}(j,c)}$$

for $1 \leq b \leq l$.

Proof. To calculate the homomorphism density of $S^{(r)}(b,c)$ in some graph G with n vertices, we first note that there are $n^{b(r-c)+c}$ maps from $V(S^{(r)}(b,c))$ to $V(G)$. We first decide to which c distinct vertices $v_1, v_2, \ldots, v_c \in V(G)$ to send the center and in which of $c!$ different ways to do so. Then each of the b edges of $S^{(r)}(b,c)$ can be sent to any of the $d_{v_1,v_2,\ldots,v_c}$ edges containing v_1, v_2, $\ldots$, v_c in G. For each of the b edges, there are $(r-c)!$ different orders to send the vertices in that edge that are not in the center to some chosen edge in $d_{v_1,v_2,\ldots,v_c}$. Thus, for any $b \geq 1$, $1 \leq c \leq r-1$ and $r \geq 2$, we have that

$$\begin{aligned}
t(S^{(r)}(b,c);G) &= \frac{c! \sum_{1\leq v_1<v_2<\ldots<v_c\leq n} ((r-c)! d_{v_1,v_2,\ldots,v_c})^b}{n^{b(r-c)+c}} \\
&= \frac{c!}{n^c} \sum_{1\leq v_1<v_2<\ldots<v_c\leq n} \left(\frac{(r-c)! d_{v_1,v_2,\ldots,v_c}}{n^{r-c}}\right)^b.
\end{aligned}$$

Let $\delta_{v_1,v_2,\ldots,v_c} = \frac{(r-c)! d_{v_1,v_2,\ldots,v_c}}{n^{r-c}}$. Consider the uniform measure on $\delta_{1,2,\ldots,c}$, $\ldots$, $\delta_{n-c+1,n-c+2,\ldots,n} \in [0,1]$. The bth moment of this measure is

$$\frac{\sum_{1\leq v_1<v_2<\ldots<v_c\leq n} \delta^b_{v_1,v_2,\ldots,v_c}}{\binom{n}{c}} = t(S^{(r)}(b,c);G) + O\left(\frac{1}{n}\right).$$

This connection allows us to write down binomial inequalities that are valid on $(t(S^{(r)}(b,c);G) : b = 1, \ldots, l)$. They are of the form

$$m_2 \geq m_1^2, \ \ m_{b-1}m_{b+1} \geq m_b^2, \ b = 2, \ldots, l-1, \ \ m_{l-1} \geq m_l.$$

These inequalities follow from Hölder's inequality [13, Theorem 18]. These binomial inequalities imply that the $\mathbf{a}_b^\top \mathbf{y} \geq 0$ are valid for $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$. Thus if we let

$$C = \{\mathbf{y} \in \mathbb{R}^l : \mathbf{a}_b^\top \mathbf{y} \geq 0, \ \forall\, 1 \leq b \leq l\},$$

we have $\operatorname{trop}(\mathcal{G}_{\mathcal{U}}) \subseteq C$. As in Theorem 3.3, we first characterize the extreme rays of C and then show that they are in $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$ to complete the proof.

Lemma 3.7. *The extreme rays of the cone* C *are exactly* $\mathbf{u}_b$ *for* $1 \leq b \leq l$.

Proof. Observe that $C = \{\mathbf{y} \in \mathbb{R}^l : A\mathbf{y} \geq 0\}$ where

$$A := \begin{bmatrix} -2 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & -2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 & -1 \end{bmatrix},$$

where the b^{th} row of A is $\mathbf{a}_b$. Since $A \in \mathbb{R}^{l\times l}$, the candidate extreme rays of C are obtained by setting $l-1$ of the defining constraints $\mathbf{a}_b^\top \mathbf{y} \geq 0$ to equality. Since there are exactly l constraints, it implies there are at most l extreme rays. Observe that $\mathbf{u}_i = (-1, -2, \ldots, -b, -b, \ldots, -b)$ satisfies all but the b^{th} constraint at equality. Since $\mathbf{u}_b \in C$, it is an extreme ray for each $1 \leq b \leq l$. $\square$

We will now show that $C \subseteq \operatorname{trop}(\mathcal{G}_\mathcal{U})$ by showing that $\mathbf{u}_m \in \operatorname{trop}(\mathcal{G}_\mathcal{U})$ for $1 \leq m \leq l$. This is done by exhibiting a family of graphs $\{G_n\}$ for each extreme ray $\mathbf{u}_m$ for which

$$(\log(t(S^{(r)}(b,c); G_n)) \,:\, b = 1, \ldots, l)$$

limits to this extreme ray.

Lemma 3.8. *The extreme rays of C are in $\operatorname{trop}(\mathcal{G}_\mathcal{U})$, and hence $C = \operatorname{trop}(\mathcal{G}_\mathcal{U})$.*

Proof. We say G is k-regular if $d_{v_1,v_2,\ldots,v_c} = k$ for some $k \in \mathbb{N}$ for every $v_1, v_2, \ldots, v_c \in V(G)$. Moreover, by edge density of an r-uniform hypergraph H on n vertices, we will mean the homomorphism density of an edge in H which is $\frac{r!|E(H)|}{n^r}$.

Consider an r-uniform hypergraph $G^{(r)}_{n,\rho}$ on n vertices constructed as follows: $(1-\alpha)n$ vertices form a k-regular graph with edge density ρ and the remaining αn vertices form a clique. Call the subgraph formed by the clique A, and the subgraph formed by the regular part B. Furthermore, any c vertices in A and any $r-c$ vertices in B also form an edge. The parameter α will be chosen later. Also, note that $k = \frac{((1-\alpha)n)^{r-c}\rho}{(r-c)!} + O(n^{r-c-1})$. This can be seen by adding the degrees of all sets of c vertices in B, and seeing that each edge gets counted $\binom{r}{c}$ times that way. Thus,

$$\frac{k\binom{(1-\alpha)n}{c}}{\binom{r}{c}} = |E(B)| = \frac{\rho((1-\alpha)n)^r}{r!},$$

yielding the desired relation between k and ρ.

Every set of c distinct vertices $v_1, v_2, \ldots, v_c \in A$ has degree $d_{v_1,v_2,\ldots,v_c} = \binom{n-c}{r-c}$ since that set of vertices forms an edge with any $r-c$ distinct vertices in $G^{(r)}_{n,\rho}$ different from it. Every set of c distinct vertices $v_1, v_2, \ldots, v_c \in B$ has degree $k = \frac{((1-\alpha)n)^{r-c}\rho}{(r-c)!} + O(n^{r-c-1})$ from edges fully in B and, if $r-c \geq c$, then $\binom{(1-\alpha)n-c}{r-2c}\binom{\alpha n}{c}$ gets added to that for edges going between B and A. Finally every set of c distinct vertices where i of them come from A and $c-i$ come from B where

$\max\{1, 2c-r\} \leq i \leq c-1$ has degree $\binom{\alpha n}{c-i}\binom{((1-\alpha)n}{r-2c+i}$. Thus,

$$
\begin{aligned}
&t(S^{(r)}(b,c); G^{(r)}_{n,\rho}) \\
&= \frac{c!}{n^c}\Bigg[\binom{\alpha n}{c}\left(\frac{(r-c)!\binom{n-c}{r-c}}{n^{r-c}}\right)^b \\
&\quad + \binom{(1-\alpha)n}{c}\Bigg(\frac{(r-c)!}{n^{r-c}}\cdot\Bigg(\frac{((1-\alpha)n)^{r-c}\rho}{(r-c)!} + O(n^{r-c-1}) \\
&\quad + \binom{(1-\alpha)n-c}{r-2c}\binom{\alpha n}{c}\Bigg)\Bigg)^b \\
&\quad + \sum_{i=\max\{1,2c-r\}}^{c-1}\binom{\alpha n}{i}\binom{n-\alpha n}{c-i}\left(\frac{(r-c)!\binom{\alpha n}{c-i}\binom{(1-\alpha)n}{r-2c+i}}{n^{r-c}}\right)^b\Bigg],
\end{aligned}
$$

where the second and third lines respectively come from sending the center of $S^{(r)}(b,c)$ to A and B. The fourth line comes from sending i vertices of the center of $S^{(r)}(b,c)$ to A, and the other $c-i$ vertices to the B.

As $n \to \infty$, this goes to

$$
\begin{aligned}
&\alpha^c + (1-\alpha)^c\left((1-\alpha)^{r-c}\rho + \frac{(r-c)!}{(r-2c)!}(1-\alpha)^{r-2c}\alpha^c\right)^b \\
&\quad + \sum_{i=\max\{1,2c-r\}}^{c-1}\frac{c!}{i!(c-i)!}\alpha^i(1-\alpha)^{c-i}\left(\frac{(r-c)!\alpha^{c-i}(1-\alpha)^{r-2c+i}}{(c-i)!(r-2c+i)!}\right)^b.
\end{aligned}
$$

If we choose $\alpha = \rho^{\frac{m}{c}}$, with $\rho << 1$ then the lowest degree part of $t(S^{(r)}(b,c); G^{(r)}_{n,\rho})$ is $\rho^b + \rho^m$. If $m < b$, then ρ^m dominates and if $m \geq b$, then ρ^b dominates. Thus $\lim_{n\to\infty,\rho\to 0}\frac{\log t(S^{(r)}(b,c);G^{(r)}_{n,\rho})}{\log\frac{1}{\rho}}$ equals $-b$ if $b \leq m$ and $-m$ if $b \geq m$. Thus this vector, $(-1,-2,\ldots,-m,-m\ldots,-m) = \mathbf{u}_m \in \operatorname{trop}(\mathcal{G}_{\mathcal{U}})$ as claimed. □

This completes the proof of the containment of $C \subseteq \operatorname{trop}(\mathcal{G}_{\mathcal{U}})$ and thus Theorem 3.6 holds. □

Note that one can find $\operatorname{trop}(\mathcal{G}_{\mathcal{U}'})$ for $\mathcal{U}' \subset \mathcal{U}$ by projecting down $\operatorname{trop}(\mathcal{G}_{\mathcal{U}})$ on the appropriate coordinates. Moreover, it follows that any valid binomial inequality for $\mathcal{G}_{\mathcal{U}}$ is implied by moment inequalities.

4. Sums of squares profiles and their tropicalizations

In this section we introduce *sos-profiles* which are semialgebraic sets that contain graph profiles. The main result of this section is Theorem 4.12 which shows that the tropicalization of any sos-profile is a rational polyhedral cone that can be described explicitly. We will use this description in Section 5 to show strong limitations of sums of squares in recognizing graph density inequalities.

4.1. **A general framework.** Let M be a symmetric matrix filled with monomials in the finite set of variables $x_1, \ldots, x_s$ for which no 2×2-principal minor is identically zero. For a point $\mathbf{v} \in \mathbb{R}^s_{\geq 0}$, let $M(\mathbf{v})$ denote the matrix obtained by evaluating each entry of M at $\mathbf{v}$, and consider the set

$$S_M := \{\mathbf{v} \in \mathbb{R}^s_{\geq 0} : M(\mathbf{v}) \succeq 0\}.$$

Also consider the superset of S_M

$$S_M^{2\times 2} := \{\mathbf{v} \in \mathbb{R}^s_{\geq 0} : \text{ all } 2\times 2 \text{ principal minors of } M(\mathbf{v}) \text{ are nonnegative}\}.$$

Both S_M and $S_M^{2\times 2}$ are (closed) semialgebraic sets in $\mathbb{R}^s_{\geq 0}$, and they both have the Hadamard property. Indeed, if $\mathbf{v}, \mathbf{w} \in \mathbb{R}^s_{\geq 0}$ and $M(\mathbf{v}) \succeq 0$ and $M(\mathbf{w}) \succeq 0$ then their Hadamard product which is $M(\mathbf{v}\cdot\mathbf{w})$ is also positive semidefinite. Therefore, S_M has the Hadamard property. If $\mathbf{v}, \mathbf{w} \in \mathbb{R}^s_{\geq 0} \cap S_M^{2\times 2}$, then for any 2×2 principal minor $\mathbf{x^a x^b} - \mathbf{x^{2c}}$ of M, we have that $\mathbf{v^{a+b}} \geq \mathbf{v^{2c}}$ and $\mathbf{w^{a+b}} \geq \mathbf{w^{2c}}$. Therefore, we also have $\mathbf{v^{a+b}w^{a+b}} \geq \mathbf{v^{2c}w^{2c}}$, or equivalently, $(\mathbf{v}\cdot\mathbf{w})^{\mathbf{a+b}} \geq (\mathbf{v}\cdot\mathbf{w})^{\mathbf{2c}}$. Therefore, $S_M^{2\times 2}$ has the Hadamard property.

Lemma 4.1.

(1) *The set* $\log(S_M^{2\times 2})$ *is a polyhedral cone in* $\mathbb{R}^s$.
(2) $\operatorname{trop}(S_M^{2\times 2}) = \log(S_M^{2\times 2})$.

Proof.

(1) A point $\mathbf{y} \in \log(S_M^{2\times 2})$ if and only if $\mathbf{y} = \log(\mathbf{v})$ for some $\mathbf{v} \in S_M^{2\times 2} \cap \mathbb{R}^s_{>0}$. A 2×2 principal minor of M evaluated at $\mathbf{v} \in \mathbb{R}^s_{>0}$ is of the form $\mathbf{v^a v^b} - \mathbf{v^{2c}}$. For a $\mathbf{v} \in \mathbb{R}^s_{>0}$, $\mathbf{v^{a+b}} \geq \mathbf{v^{2c}}$ if and only if $\sum_{i=1}^s (a_i + b_i - 2c_i)\log(v_i) \geq 0$. Thus $\log(S_M^{2\times 2})$ is the polyhedral cone in $\mathbb{R}^s$ defined by the linear inequalities $\sum_{i=1}^s (a_i + b_i - 2c_i) y_i \geq 0$ obtained from the 2×2 principal minors of M.
(2) We already showed that $S_M^{2\times 2}$ has the Hadamard property. Since $\log(S_M^{2\times 2})$ is a polyhedral cone, it coincides with the closure of both its cone hull and convex hull, and so by Lemma 2.2, $\operatorname{trop}(S_M^{2\times 2}) = \log(S_M^{2\times 2})$. $\square$

Corollary 4.2. *The dual cone,* $\operatorname{trop}(S_M^{2\times 2})^* \subset \mathbb{R}^s$, *is the rational polyhedral cone generated by the vectors*

$$\left\{ \mathbf{a} + \mathbf{b} - 2\mathbf{c} \, : \, \begin{pmatrix} \mathbf{x^a} & \mathbf{x^c} \\ \mathbf{x^c} & \mathbf{x^b} \end{pmatrix} \textit{ is a } 2\times 2 \textit{ principal submatrix of } M \right\}.$$

Lemma 4.3. $\operatorname{trop}(S_M) = \operatorname{cl}(\operatorname{conv}(\log(S_M)))$.

Proof. The all-ones matrix $M(\mathbf{1})$ is positive semidefinite and hence $\mathbf{1} \in S_M$. The result now follows from S_M having the Hadamard property and Lemma 2.2. $\square$

Theorem 4.4. *Let M be a symmetric matrix filled with monomials in $x_1, \ldots, x_s$ such that no 2×2 minor of M is identically 0. Suppose also that $S_M^{2\times 2}$ has interior in $\mathbb{R}^s_{>0}$. Then* $\operatorname{trop}(S_M) = \operatorname{trop}(S_M^{2\times 2})$.

Proof. Since $S_M \subseteq S_M^{2\times 2}$ we have that

$$\operatorname{trop}(S_M) = \operatorname{cl}(\operatorname{conv}(\log(S_M)) \subseteq \operatorname{cl}(\operatorname{conv}(\log(S_M^{2\times 2})) = \operatorname{trop}(S_M^{2\times 2}).$$

To show the reverse containment, we use the following strategy. Pick a $\mathbf{v} \in \operatorname{int}(S_M^{2\times 2})$, or equivalently, $\log(\mathbf{v}) \in \operatorname{int}(\log(S_M^{2\times 2})) = \operatorname{int}(\operatorname{trop}(S_M^{2\times 2}))$. Then show that for a large enough positive integer k, $k\log(\mathbf{v})$ lies in $\log(S_M) \subseteq \operatorname{trop}(S_M)$. Since $\operatorname{trop}(S_M)$ is a cone, it must follow that $\log(\mathbf{v}) \in \operatorname{trop}(S_M)$. Thus we have that $\operatorname{int}(\operatorname{trop}(S_M^{2\times 2})) \subseteq \operatorname{trop}(S_M)$ which means that $\operatorname{trop}(S_M^{2\times 2}) \subseteq \operatorname{trop}(S_M)$ since the tropicalizations are closed sets.

Consider $\log(\mathbf{v})$ in the interior of $\log(S_M^{2\times 2})$. Then $\mathbf{v}$ lies in the interior of $S_M^{2\times 2}$ and all 2×2 principal minors of $M(\mathbf{v})$ are strictly positive. Since $\log(S_M^{2\times 2})$ is a cone, for any $k > 0$ and integer, $k \log(\mathbf{v}) = \log(\mathbf{v}^k)$ lies in $\log(S_M^{2\times 2})$. We will now argue that if k is large enough then $\mathbf{v}^k$ is also in S_M or equivalently, that all principal minors of $M(\mathbf{v}^k)$ are positive. Recall that no 2×2 principal minor of M is identically zero. A 2×2 principal minor of $M(\mathbf{v}^k)$, namely $\mathbf{v}^{k(\mathbf{a}+\mathbf{b})} - \mathbf{v}^{k(2\mathbf{c})} = (\mathbf{v}^{\mathbf{a}+\mathbf{b}})^k - (\mathbf{v}^{2\mathbf{c}})^k$, is positive since $\mathbf{v}^{\mathbf{a}+\mathbf{b}} > -\mathbf{v}^{2\mathbf{c}}$. Now consider an $l \times l$ principal minor of $M(\mathbf{v}^k)$, and a term T in the Laplace expansion of its determinant indexed by a nonidentity permutation. Replace every non-diagonal entry in T from position $(i,j), i \neq j$ by the product of the square roots of the diagonal entries in positions (i,i) and (j,j) in this $l\times l$ principal minor. Let T' denote the modification of T obtained by replacing all non-diagonal terms in T as above. Since all 2×2 minors of $M(\mathbf{v}^k)$ are positive, we get that $T' > T$. Note that T' is the product of diagonal entries in the $l \times l$ principal minor we are considering. Since there are only finitely many terms in the Laplace expansion of the determinant of the $l \times l$ principal minor, we can choose k large enough to ensure that T' is so much bigger than the other terms making the entire determinant positive. $\square$

Example 4.5. We now give an example to illustrate the necessity of the condition that no 2×2 principal minor of M should be identically zero. Consider the matrix

$$\begin{pmatrix} x & x^2 & x^2 \\ x^2 & x^3 & x^2 \\ x^2 & x^2 & 1 \end{pmatrix}$$

in which the upper left 2×2 minor is identically 0. The values of x for which all 2×2 principal minors are nonnegative is precisely $0 \leq x \leq 1$. Therefore, $S_M^{2\times 2} = [0,1]$ and $\operatorname{trop}(S_M^{2\times 2}) = \mathbb{R}_{\leq 0}$. The determinant of this matrix is $-(x-1)^2 x^5$, so the only values that make the matrix positive semidefinite are $x = 0$ and $x = 1$. Therefore, $\operatorname{trop}(S_M)$ is the origin which is strictly contained in $\operatorname{trop}(S_M^{2\times 2})$.

The main take away from what we have so far is that even though S_M might be strictly contained in $S_M^{2\times 2}$, their tropicalizations agree and form a polyhedral cone with an explicit inequality description given by the 2×2 principal minors of M.

4.2. **Specialization to graphs.** We now specialize the above results to the case of graphs. For this we begin with a few definitions about the *gluing algebra* of graphs, the reader is referred to Lovász [22] for a broader exposition. A graph is *partially labeled* if a subset of its vertices are labeled with elements of $\mathbb{N} := \{1,2,3,\ldots\}$ such that no vertex receives more than one label. If no vertices of H are labeled then H is *unlabeled*. Let $\mathcal{A}$ denote the vector space of all formal finite $\mathbb{R}$-linear combinations of partially labeled graphs without isolated vertices, including the empty graph with no vertices which we denote as 1. We call an element $a = \sum \alpha_i H_i$ of $\mathcal{A}$ a *graph combination*, each $\alpha_i H_i$ a *term* of a, and each H_i a *constituent graph* of a. Let $\mathcal{A}_\emptyset$ denote the subspace of $\mathcal{A}$ spanned by unlabeled graphs. We view elements $a \in \mathcal{A}_\emptyset$ as functions that can be evaluated on unlabeled graphs G via homomorphism densities, namely $t(a;G) = \sum \alpha_i t(H_i;G)$. An element $a = \sum \alpha_i H_i$ of $\mathcal{A}_\emptyset$ is called *nonnegative* if $\sum \alpha_i t(H_i;G) \geq 0$ for all graphs G.

The vector space $\mathcal{A}$ has a product defined as follows. For two labeled graphs H_1 and H_2, form the new labeled graph $H_1 H_2$ by gluing together the vertices in the two graphs with the same label, and keeping only one copy of any edge that may have doubled in the process. Equipped with this product, $\mathcal{A}$ becomes an $\mathbb{R}$-algebra with the empty graph 1 as its multiplicative identity. The algebra $\mathcal{A}$ admits a simple linear map into $\mathcal{A}_\emptyset$ that removes the labels in a graph combination to create a graph combination of unlabeled graphs. We call this map *unlabeling* and denote it by $[[\cdot]]$. A *sum of squares (sos)* in $\mathcal{A}_\emptyset$ is a finite sum of unlabeled squares of graph combinations $a_i \in \mathcal{A}$, namely, $\sum[[a_i^2]]$. A sum of squares is a nonnegative graph combination.

A *d-sos graph combination* is an sos $a = \sum[[a_j^2]]$ where each constituent graph in a_j is partially labeled and has at most d edges. This means that every constituent graph of a has at most $2d$ edges. For a fixed $d \in \mathbb{N}$, it follows from results of [29] (see also [28]) that any d-sos graph combination can be written using only finitely many, say $\ell(d)$, labels. Let $\mathcal{B}_d$ denote the set containing the empty graph 1 with no vertices, and all partially labeled graphs with labels $1, \ldots, \ell(d)$, and at most d edges and no isolated vertices. Define

$$\mathcal{V}_d = \{[[ab]] \text{ connected } : a, b \in \mathcal{B}_d\} \backslash \{1\}.$$

A d-sos graph combination is a sum of squares of graph combinations in the span of $\mathcal{B}_d$. i.e., if $a = \sum[[a_j^2]]$ is d-sos then $a_j \in \operatorname{span}(\mathcal{B}_d)$. Also, any term in a d-sos graph combination a is a monomial in the elements of $\mathcal{V}_d$ (including constant terms) and each constituent graph in a has at most $2d$ edges.

Definition 4.6.

(1) The *d-sos-profile*, denoted $\mathcal{S}_d$, is the set of all $\mathbf{v} \in \mathbb{R}^s_{\geq 0}$ such that $a(\mathbf{v}) \geq 0$ for all d-sos graph combinations a and where $s = |\mathcal{V}_d|$.

(2) The *$(\mathcal{U}, d)$-sos-profile* denoted as $\mathcal{S}_{\mathcal{U},d}$ is the projection of $\mathcal{S}_d$ on coordinates corresponding to graphs in $\mathcal{U}$.

Let $\mathcal{M}_d$ be the *moment matrix* of size $|\mathcal{B}_d| \times |\mathcal{B}_d|$ which is defined as the matrix with rows and columns indexed by the graphs in $\mathcal{B}_d$ and whose (A, B)-entry is $[[AB]]$. Such a matrix is called a *connection matrix* in [22]. Every entry in $\mathcal{M}_d$ is a monomial in the elements of $\mathcal{V}_d$ (including 1) and the corresponding graph has at most $2d$ edges. Further, the entries of $\mathcal{M}_d$ and the monomials that appear in d-sos graph combinations are the same.

Lemma 4.7. *The d-sos-profile $\mathcal{S}_d = \{\mathbf{v} \in \mathbb{R}^s_{\geq 0} : \mathcal{M}_d(\mathbf{v}) \succeq 0\} = S_{\mathcal{M}_d}$.*

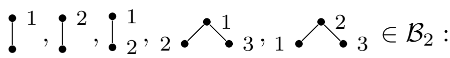
Proof. For any $Q \succeq 0$, $\langle \mathcal{M}_d, Q\rangle$ is a d-sos graph combination $\sum_j [[a_j^2]]$ since $a_j \in \operatorname{span}(\mathcal{B}_d)$. Conversely, any d-sos graph combination can be written as $\langle \mathcal{M}_d, Q\rangle$ for some $Q \succeq 0$. Therefore, $\mathbf{v} \in \mathbb{R}^s_{\geq 0}$ lies in $\mathcal{S}_d$ if and only if $\langle \mathcal{M}_d(\mathbf{v}), Q\rangle \geq 0$ for all $Q \succeq 0$ which happens if and only if $\mathcal{M}_d(\mathbf{v}) \succeq 0$. $\square$

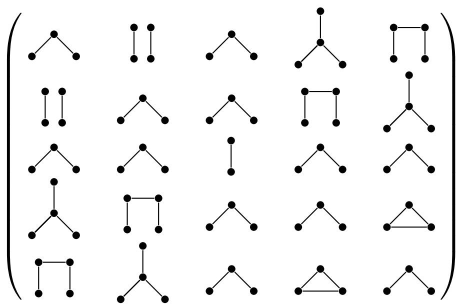
Lemma 4.7 shows that the sos-profile $\mathcal{S}_d$ is a semialgebraic set in $\mathbb{R}^s_{\geq 0}$ since the condition $\mathcal{M}_d(\mathbf{v}) \succeq 0$ is equivalent to the (finitely many) principal minors of $\mathcal{M}_d(\mathbf{v})$ being nonnegative, and each such minor is a polynomial in the elements of $\mathcal{V}_d$.

We now note the connection between the graph profile $\mathcal{G}_{\mathcal{V}_d}$ and the d-sos-profile $\mathcal{S}_d$. Recall that $\mathcal{G}_{\mathcal{V}_d}$ is the closure of all points $(t(C; G) : C \in \mathcal{V}_d)$ as G varies over all unlabeled graphs. Tautologically, $\mathcal{G}_{\mathcal{V}_d}$ is also the set of all points in $\mathbb{R}^s_{\geq 0}$ on which all nonnegative graph combinations in the polynomial ring $\mathbb{R}[\mathcal{V}_d]$ are nonnegative. Since there is no bound to the number of edges in the constituent graphs of such nonnegative graph combinations, the profile $\mathcal{G}_{\mathcal{V}_d}$ may not be semialgebraic, and as mentioned in the introduction, the profile of edge and triangle is not semialgebraic. A d-sos graph combination is also a nonnegative graph combination in $\mathbb{R}[\mathcal{V}_d]$, but one in which the constituent graphs cannot have more than $2d$ edges. Therefore, we immediately get that the graph profile $\mathcal{G}_{\mathcal{V}_d}$ is contained in $\mathcal{S}_d$. Indeed, if $\mathbf{v} \in \mathbb{R}^s_{\geq 0} \cap \mathcal{G}_{\mathcal{V}_d}$, then $\mathbf{v}\mathbf{v}^\top = \mathcal{M}_d(\mathbf{v}) \succeq 0$. If we are given a specific set of finite connected graphs $\mathcal{U}$, then $\mathcal{U} \subseteq \mathcal{V}_d$ for d large enough. Certainly, any finite connected graph $H \in \mathcal{U}$ can be written as the symmetrized product of two partially labeled finite graphs, namely the empty graph and an unlabeled copy of H itself, which is in $\mathcal{V}_d$ for any $d \geq |E(H)|$. We then have that $\mathcal{G}_{\mathcal{U}}$ is a projection of the graph profile $\mathcal{G}_{\mathcal{V}_d}$, and contained in the $(\mathcal{U}, d)$-sos-profile $\mathcal{S}_{\mathcal{U},d}$.

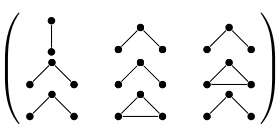
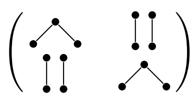
Example 4.8. Consider the graph profile $\mathcal{G}_{\mathcal{U}}$ for $\mathcal{U} = \{ , \}$ from Figure 1. We will show that $(0.7, 0.12)$ is not in $\mathcal{G}_{\mathcal{U}}$. We know that $\mathcal{G}_{\mathcal{U}}$ is contained in $\mathcal{S}_{\mathcal{U},d}$ for all $d \geq 2$.

From Lemma 4.7, $\mathcal{S}_d$ contains all the points $\mathbf{v}$ that make $\mathcal{M}_d(\mathbf{v}) \succeq 0$. To show that $(0.7, 0.12) \notin \mathcal{G}_{\mathcal{U}}$, we can instead prove $(0.7, 0.12) \notin \mathcal{S}_{\mathcal{U},2}$ by observing that any point $\mathbf{v} \in \mathbb{R}^s$ where the component is 0.7 and the component is 0.12 is such that $\mathcal{M}_2(\mathbf{v}) \not\succeq 0$. Indeed, consider the principal submatrix of $\mathcal{M}_2$ corresponding to the graphs



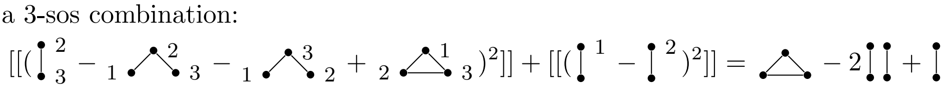


The bottom 3×3 principal minor and the top 2×2 principal minor,

det  and det  ,

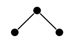
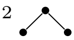
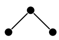
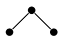
yield that $-2\,{}^3 + 0.94\,{}^2 - 0.01008 \geq 0$ and ${}^2 - 0.7^4 \geq 0$ when $= 0.7$ and $= 0.12$. However, there is no value of in $[0, 1]$ that satisfies both inequalities. Thus the point $(0.7, 0.12) \notin \mathcal{S}_{\mathcal{U},2}$.

Alternatively, to show that $(0.7, 0.12) \notin \mathcal{G}_{\mathcal{U}}$, we could have shown that $(0.7, 0.12) \notin \mathcal{S}_{\mathcal{U},3}$ by thinking about $\mathcal{S}_3$ as the set of points that evaluate nonnegatively on all 3-sos graph combinations. As seen in [22], the Goodman bound can be written as a 3-sos combination:



so $-2 + \geq 0$ is a valid inequality for $\mathcal{S}_3$ and $\mathcal{S}_{\mathcal{U},3}$. Since $0.12 - 2 \cdot 0.7^2 + 0.7 < 0$, $(0.7, 0.12) \notin \mathcal{S}_{\mathcal{U},3}$.

Finally, we note that $\mathcal{S}_{\mathcal{U},3}$ strictly contains $\mathcal{G}_{\mathcal{U}}$ in this case, that is, there are points $\mathbf{v}$ such that $\mathcal{M}_3(\mathbf{v}) \succeq 0$, but such that the projection of $\mathbf{v}$ on the coordinates corresponding to graphs in $\mathcal{G}_{\mathcal{U}}$ is not in $\mathcal{G}_{\mathcal{U}}$. It was shown in [28] that one needs sums of squares of arbitrarily high degree to carve out $\mathcal{G}_{\mathcal{U}}$ when $\mathcal{U} = \{ , \}$.

Lemma 4.7 suggests that we may be able to apply Theorem 4.4 to $\mathcal{S}_d$. The first hurdle is that some principal minors of $\mathcal{M}_d$ might be identically zero. For example, the principal minor with rows (columns) indexed by and  is ^{2} ^{2} $-$   $= 0$.

Say that two partially labeled graphs are isomorphic if they are isomorphic as labeled graphs. In particular, two isomorphic labeled graphs have the same labels.

Lemma 4.9. *A 2×2-principal minor in $\mathcal{M}_d$ is identically zero if and only if the corresponding rows (columns) correspond to two graphs with isomorphic labeled components.*

Proof. Consider a 2×2-principal minor in $\mathcal{M}_d$ corresponding to the rows and columns indexed by H_1 and H_2 where the labels of H_1 and H_2 (if any) are contained in some finite set L. The minor is thus equal to $[[H_1^2]][[H_2^2]] - [[H_1H_2]][[H_1H_2]]$. We now show that this expression is a sum of squares. Let $\tilde{H}_i$ be the same graph as H_i for $i \in \{1, 2\}$ but where any label l becomes $l + |L|$. Note that the label sets of H_i and $\tilde{H}_j$ do not intersect for $i, j \in \{1, 2\}$, that $[[H_i^2]] = [[\tilde{H}_i^2]]$ for $i \in \{1, 2\}$ and $[[H_1H_2]] = [[\tilde{H}_1\tilde{H}_2]]$. Thus the minor is equal to

$$\begin{aligned}\frac{1}{2}[[(H_1\tilde{H}_2 - \tilde{H}_1H_2)^2]] &= \frac{1}{2}[[H_1^2\tilde{H}_2^2 - 2H_1H_2\tilde{H}_1\tilde{H}_2 + \tilde{H}_1^2H_2^2]] \\ &= \frac{1}{2}\left([[H_1^2]][[\tilde{H}_2^2]] - 2[[H_1H_2]][[\tilde{H}_1\tilde{H}_2]] + [[\tilde{H}_1^2]][[H_2^2]]\right) \\ &= \frac{1}{2}\left([[H_1^2]][[H_2^2]] - 2[[H_1H_2]][[H_1H_2]] + [[H_1^2]][[H_2^2]]\right)\end{aligned}$$

as desired. Indeed, to go from the first to the second line, one simply needs to note that the symmetrization of products of graphs that have no labels in common is equal to the product of the symmetrization of those graphs.

Thus, if the minor is identically zero, $[[(H_1\tilde{H}_2 - \tilde{H}_1H_2)^2]] = 0$. From Lemma 2.3 of [5], the only way this can be so is if $H_1\tilde{H}_2 = \tilde{H}_1H_2$. Let $H_1 = H_1^lH_1^u$ and $H_2 = H_2^lH_2^u$ where H_i^l is the graph H_i restricted to components that contain at least one label, and H_i^u is the graph H_i restricted to components that are unlabeled. Then $\tilde{H}_i = \tilde{H}_i^lH_i^u$. Thus, $H_1\tilde{H}_2 = \tilde{H}_1H_2$ is equivalent to $H_1^lH_1^u\tilde{H}_2^lH_2^u = \tilde{H}_1^lH_1^uH_2^lH_2^u$ which is equivalent to $H_1^l\tilde{H}_2^l = \tilde{H}_1^lH_2^l$.

Suppose H_1 and H_2 are two graphs where the labeled components are isomorphic, i.e., $H_1^l = H_2^l$. Note that this includes the case when both H_1 and H_2 are fully unlabeled. Observe that $\tilde{H}_1^l = \tilde{H}_2^l$. Thus, $H_1^l\tilde{H}_2^l = \tilde{H}_1^lH_2^l$ and the corresponding 2×2-minor is identically zero.

Suppose now that the labeled components of H_1 and H_2 are not isomorphic, i.e., $H_1^l \neq H_2^l$ and $\tilde{H}_1^l \neq \tilde{H}_2^l$. (Note that this implies that at least one of those two graphs contain a label, otherwise, their labeled parts would be the same.) Moreover, note that $\tilde{H}_1^l \neq H_2^l$ and $H_1^l \neq \tilde{H}_2^l$ as the labels come from non-intersecting label sets. Thus it is impossible that $H_1^l\tilde{H}_2^l = \tilde{H}_1^lH_2^l$, and the corresponding 2×2-minor cannot be identically zero. $\square$

To avoid 2×2 principal minors that are identically 0 in $\mathcal{M}_d$, we restrict $\mathcal{B}_d$ to $\tilde{\mathcal{B}}_d$, the subset containing the empty graph 1, and all partially labeled graphs without unlabeled connected components and isolated vertices. Every graph in $\tilde{\mathcal{B}}_d$ still has at most d edges. Furthermore, let $\tilde{\mathcal{M}}_d$ be the moment matrix for $\tilde{\mathcal{B}}_d$. Then Corollary 4.10 follows from Lemma 4.9.

Corollary 4.10. *No 2×2-principal minor in the symmetric matrix $\tilde{\mathcal{M}}_d$ is identically zero.*

We now replace $\mathcal{M}_d$ with $\tilde{\mathcal{M}}_d$ in Lemma 4.7.

Lemma 4.11. *The d-sos-profile $\mathcal{S}_d$ coincides with $S_{\tilde{\mathcal{M}}_d} = \{\mathbf{v} \in \mathbb{R}^s_{\geq 0} : \tilde{\mathcal{M}}_d(\mathbf{v}) \succeq 0\}$.*

Proof. By Lemma 4.7 it suffices to argue that for $\mathbf{v} \geq 0$, $\tilde{\mathcal{M}}_d(\mathbf{v}) \succeq 0$ if and only if $\mathcal{M}_d(\mathbf{v}) \succeq 0$. Since $\tilde{\mathcal{M}}_d$ is a principal submatrix of $\mathcal{M}_d$, $\mathcal{M}_d(\mathbf{v}) \succeq 0$ implies $\tilde{\mathcal{M}}_d(\mathbf{v}) \succeq 0$.

Consider a graph $F = F^u F^l$ where F^u is unlabeled and every component in F^l has at least one label. Then the row of $\mathcal{M}_d$ indexed by F is $(F^u[[F^l H]] : H \in \mathcal{B}_d)$. Therefore the corresponding row of $\mathcal{M}_d(\mathbf{v})$ is $(F^u(\mathbf{v})[[F^l H]](\mathbf{v}) : H \in \mathcal{B}_d)$. This means that every term in a principal minor of $\mathcal{M}_d(\mathbf{v})$ (expanded in terms of permutations) involving the row indexed by F contains the common factor $F^u(\mathbf{v})$ which is nonnegative. Factoring this out, we obtain a principal minor of $\tilde{\mathcal{M}}_d(\mathbf{v})$. Thus if $\tilde{\mathcal{M}}_d(\mathbf{v}) \succeq 0$ then $\mathcal{M}_d(\mathbf{v}) \succeq 0$. $\square$

The second requirement in Theorem 4.4 is that $S_{\tilde{\mathcal{M}}_d^{2\times 2}}$ has an interior. By Lemma 4.7, $\mathcal{G}_{\mathcal{V}_d} \subseteq \mathcal{S}_d \subseteq S_{\tilde{\mathcal{M}}_d^{2\times 2}}$ and since $\mathcal{G}_{\mathcal{V}_d}$ has an interior [8] so do $\mathcal{S}_d$ and $S_{\tilde{\mathcal{M}}_d^{2\times 2}}$. We can now apply Theorem 4.4.

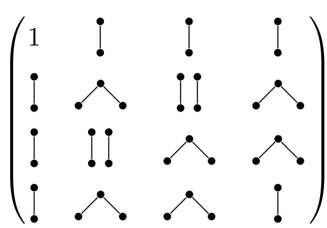
Theorem 4.12. *The d-sos-profile $\mathcal{S}_d$ is a basic semialgebraic set in $\mathbb{R}^s_{\geq 0}$ containing the graph profile $\mathcal{G}_{\mathcal{V}_d}$, and its tropicalization, $\operatorname{trop}(\mathcal{S}_d)$, is the rational polyhedral cone $\log(S^{2\times 2}_{\tilde{\mathcal{M}}_d})$.*

Corollary 4.13. *The $(\mathcal{U}, d)$-sos-profile $\mathcal{S}_{\mathcal{U},d}$ is a semialgebraic set containing the graph profile $\mathcal{G}_{\mathcal{U}}$ for all d. Its tropicalization, $\operatorname{trop}(\mathcal{S}_{\mathcal{U},d})$, is the projection of the rational polyhedral cone $\log(S^{2\times 2}_{\tilde{\mathcal{M}}_d})$ onto the coordinates indexed by $\mathcal{U}$. In particular, $\operatorname{trop}(\mathcal{S}_{\mathcal{U},d})$ is also a rational polyhedral cone.*

Example 4.14. Let $d = 1$. Then $\mathcal{V}_1 = \{ , \}$ is the set of all unlabeled connected graphs that can be obtained as a product of two partially labeled graphs with at most one edge. The 1-sos-profile $\mathcal{S}_1$ is the set of all $\mathbf{v} \in \mathbb{R}^2_{\geq 0}$ that evaluate nonnegatively on all 1-sos polynomials in $\mathbb{R}[,]$. The graph profile $\mathcal{G}_{\mathcal{V}_1}$ is shown in Figure 6. The lower bound consists of dn-regular graphs with $d \in [0, 1]$. The upper bound on the right consists of a clique on αn vertices for some $\alpha \in [\frac{1}{2}, 1]$, and the upper bound on the left consists of the complement of such graphs [1].

We will show that $\operatorname{trop}(\mathcal{G}_{\mathcal{V}_1}) = \operatorname{trop}(\mathcal{S}_1)$.

Consider $\tilde{\mathcal{B}}_1 = \{1, {}^{1}, {}^{2}, {}^{1}_{2}\}$. Then

$\tilde{\mathcal{M}}_1 =$  .

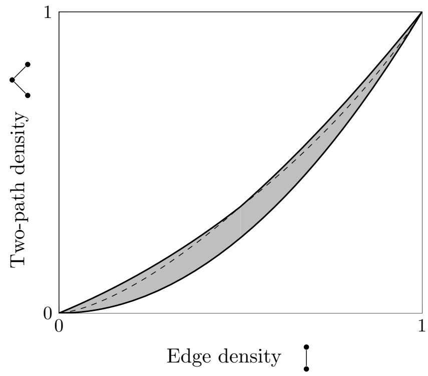


FIGURE 6. The graph profile of $\mathcal{V}_1$

After removing redundancies, the six 2×2 principal minors of $\tilde{\mathcal{M}}_1$ yield the following two inequalities:

$$ - {}^2 \geq 0 \text{ and } - \geq 0.$$

Therefore, $\log(S_{\tilde{\mathcal{M}}_1^{2\times 2}}) = \operatorname{trop}(\mathcal{S}_1)$ is the cone generated by the rays $(-1,-1)$ and $(-1,-2)$. As we saw in the star example of Section 3.2, this cone coincides with $\operatorname{trop}(\mathcal{G}_{\mathcal{V}_1})$ when $\mathcal{V}_1 = \{ , \}$.

4.3. **Sos-testable graph combinations.** We now introduce the notion of an sos-testable graph combination which plays an important role in Section 5.

Definition 4.15. A graph combination a is *sos-testable* if $a \geq 0$ on a d-sos-profile $\mathcal{S}_d$ for some d.

Theorem 4.16. *Let a, b, c be graph combinations such that $b \neq 0$ and c are sos-testable. If $ab = c$ then a is sos-testable.*

Proof. Let $\mathcal{U} = \{F_1, \ldots, F_k\}$ be the set of connected components of graphs in a, b and c. There exist $d_1, d_2 \in \mathbb{N}$ such that b is nonnegative on $\mathcal{S}_{\mathcal{U},d_1}$ and c is nonnegative on $\mathcal{S}_{\mathcal{U},d_2}$. Let $d = \max\{d_1, d_2\}$. We will prove that $a \geq 0$ on $\mathcal{S}_{\mathcal{U},d}$ making it sos-testable.

We first argue that every neighborhood of $\mathbf{1}$ has a ball contained in $\mathcal{G}_{\mathcal{U}}$, i.e., for every $r > 0$ there exists an $\tilde{r} > 0$, and $\mathbf{w} \in \mathbb{R}^{|\mathcal{U}|}$ such that $B(\mathbf{w}, \tilde{r}) \subseteq \mathcal{G}_{\mathcal{U}} \cap B(\mathbf{1}, r)$. Here $B(\mathbf{w}, \tilde{r})$ denotes the closed ball of radius $\tilde{r}$ around $\mathbf{w}$. From [9, Theorem 1], we have that there exists $\mathbf{z} \in \mathbb{R}^{|\mathcal{U}|}$ and $\epsilon > 0$ such that $B(\mathbf{z}, \epsilon) \subseteq \mathcal{G}_{\mathcal{U}}$. Thus, for every $\mathbf{y} \in B(\mathbf{z}, \epsilon)$ there exists a sequence of graphs $G_1, \ldots, G_n, \ldots$, where $|V(G_n)| = n$, such that $\lim_{n\to\infty}(t(F_1; G_n), \ldots, t(F_k; G_n)) = \mathbf{y}$. Now fix $r > 0$. Then consider the graph sequence H_n which consists of $G_{\delta n}$ along with a disjoint copy of $K_{(1-\delta)n}$ for each n, where $\delta > 0$ will be fixed later and we ignore integrality issues with δn. For an F_i, we have $t(F_i, H_n) = (1-\delta)^{|V(F_i)|} + \delta^{|V(F_i)|} t(F_i, G_{\delta n})$. Thus $\lim_{n\to\infty} t(F_i, H_n) = (1-\delta)^{|V(F_i)|} + \delta^{|V(F_i)|} y_i$. Let $\phi : \mathbb{R}^{|\mathcal{U}|} \to \mathbb{R}^{|\mathcal{U}|}$ denote the map where

$$\phi(\mathbf{y}) = \left((1-\delta)^{|V(F_1)|} + \delta^{|V(F_1)|} y_1, \ldots, (1-\delta)^{|V(F_k)|} + \delta^{|V(F_k)|} y_k \right).$$

By construction, $\phi(\mathbf{y}) \in \mathcal{G}_{\mathcal{U}}$. If we set $\delta = \frac{r}{\max_i |V(F_i)|}$ then $(1-\delta)^{|V(F_i)|} \geq 1 - \delta|V(F_i)| \geq 1 - r$. Thus $\phi(\mathbf{y}) \geq 1 - r$ and $\phi(\mathbf{y}) \in B(\mathbf{1}, r)$. Therefore, we have $\phi(B(\mathbf{z},\epsilon)) \subseteq \mathcal{G}_\mathcal{U} \cap B(\mathbf{1}, r)$. Moreover, the Jacobian of ϕ is just the diagonal matrix with i^{th} diagonal entry $\delta^{|V(F_i)|}$ and thus has a non-zero determinant. Thus $\phi(B(\mathbf{z},\epsilon))$ is full-dimensional and contains a ball $B(\mathbf{w}, \tilde{r})$ for some $\mathbf{w} \in \mathbb{R}^\mathcal{U}, \tilde{r} > 0$. Therefore, we have $B(\mathbf{w}, \tilde{r}) \subseteq \mathcal{G}_\mathcal{U} \cap B(\mathbf{1}, r)$ as claimed.

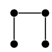
Now suppose there exists $\mathbf{x} \in \mathcal{S}_{\mathcal{U},d}$ such that $a(\mathbf{x}) < 0$. Since $\mathcal{S}_{\mathcal{U},d}$ has the Hadamard property and $\mathcal{G}_\mathcal{U} \subset \mathcal{S}_{\mathcal{U},d}$, we see that any neighborhood of $\mathbf{x}$ in $\mathcal{S}_{\mathcal{U},d}$ also contains a closed ball by applying the Hadamard property to $\mathbf{x}$ and the closed ball in the neighborhood of $\mathbf{1}$. Since $a(\mathbf{x}) < 0$ and a is polynomial function in the coordinates indexed by $\mathcal{U}$, and therefore continuous, it follows that there exists $\tilde{\mathbf{x}} \in \mathcal{S}_{\mathcal{U},d}$ such that $a(\tilde{\mathbf{x}}) < 0$ and a closed ball B around $\tilde{\mathbf{x}}$ is contained in $\mathcal{S}_{\mathcal{U},d}$. Since b and c are sos-testable, we have $b(\mathbf{u}), c(\mathbf{u}) \geq 0$ for all $\mathbf{u} \in B$. Moreover, since $b \neq 0$, there exists $\hat{\mathbf{x}} \in B$ such that $a(\hat{\mathbf{x}}) < 0$ and $b(\hat{\mathbf{x}}) > 0$. Then we have $ab(\hat{\mathbf{x}}) < 0$ while $c(\hat{\mathbf{x}}) \geq 0$ which is a contradiction. □

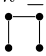
Corollary 4.17. *If a graph combination a is not sos-testable, it is not a rational sos.*

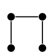
Proof. For the sake of contradiction, suppose a is a rational sos, i.e., $a = \frac{c}{b}$ where $b \neq 0$ and b and c are sos. Then we have $ab = c$. Moreover, b and c are sos-testable since they are sos. Thus a must also be sos-testable from Theorem 4.16 which is a contradiction. □

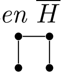
5. Limitations of sums of squares

In this section we will use the d-sos-profile $\mathcal{S}_d$ and the $(\mathcal{U}, d)$-sos-profile $\mathcal{S}_{\mathcal{U},d}$ defined in Section 4 to show that there are simple binomial graph density inequalities that are not sos-testable. This means that sums of squares do not recognize these inequalities.

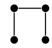
Following [5], we call an unlabeled graph H, a *trivial square*, if whenever $H = [[F^2]]$ then F must be a fully labeled copy of H. For example,  is a trivial square.

Theorem 5.1. *Let $\overline{H}$ and $\underline{H}$ be two graphs with the same number of edges where the former is a trivial square in which every vertex has degree p or $p+1$ for some integer $p \geq 1$, and the degree of any vertex in $\underline{H}$ is at most $p+1$. Then for any $k \geq 2|E(\underline{H})| + 1$, the inequality $\overline{H}^k - \underline{H}^{k+1} \geq 0$ is not sos-testable. In particular,*  $^{k} -$  $^{3(k+1)}$ *is not sos-testable for $k \geq 7$.*

One can also show that  $^{k} -$ $^{3k+1}$ is not sos-testable, for k big enough, by using a similar strategy as the one described below. Theorem 5.1 has Corollary 5.2.

Corollary 5.2. *Let $\overline{H}$ and $\underline{H}$ be two graphs with the same number of edges where the former is a trivial square in which every vertex has degree p or $p+1$ for some integer $p \geq 1$, and the degree of any vertex in $\underline{H}$ is at most $p+1$. Then $\overline{H} - \underline{H}$ is not sos-testable and cannot be written as a rational sos. In particular,*  $-$  ^{3} *is not sos-testable and cannot be written as a rational sos.*

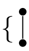
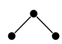
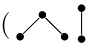
Proof. Observe that if $\overline{H} - \underline{H}$ is sos-testable then $\overline{H}^k - \underline{H}^k$ is sos-testable for every integer $k \geq 1$. Furthermore, $\overline{H}^k - \underline{H}^{k+1}$ is also sos-testable for every integer $k \geq 1$ since graph densities lie in $[0, 1]$. Therefore, if $\overline{H}$ and $\underline{H}$ satisfy the conditions of Theorem 5.1, $\overline{H} - \underline{H}$ is not sos-testable. By Corollary 4.17 we then have that $\overline{H} - \underline{H}$

is not a rational sos. For the last claim, observe that $\overline{H} =$  and $\underline{H} =$ ^{3} satisfies the conditions of Theorem 5.1. $\square$

Before we prove Theorem 5.1 we make a few definitions. Recall that every term in a d-sos graph combination is a constant times a monomial in the elements of $\mathcal{V}_d$. Therefore, the coordinates of vectors in both $\mathcal{S}_d$ and $\operatorname{trop}(\mathcal{S}_d)$ are indexed by graphs in $\mathcal{V}_d$, and $|\mathcal{V}_d| = s$. In what follows we will assume that we have fixed an ordering of the elements of $\mathcal{V}_d$. For any $\mathbf{x} \in \mathbb{R}^s$ and $H \in \mathcal{V}_d$, denote by x_H the coordinate of $\mathbf{x}$ indexed by H. Also, define $\mathbf{1}_H$ to be the point in tropical space with 1 in the coordinate labeled by H and 0 otherwise, the *indicator vector* of H.

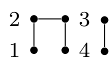
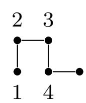
Definition 5.3. Let G be an unlabeled graph with factorization $G = C_1^{\alpha_1} \cdots C_s^{\alpha_s}$ into connected unlabeled graphs $C_i \in \mathcal{V}_d$. Define $\boldsymbol{\alpha}(G)$ to be the point in tropical space recording the exponents in the factorization of G:

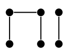
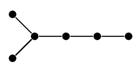
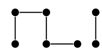
$$\boldsymbol{\alpha}(G) = \sum_{i=1}^{s} \alpha_i \mathbf{1}_{G_i}.$$

Example 5.4. Consider $\mathcal{V}_1 = \{$ ,  $\}$. Then $\boldsymbol{\alpha}($  $^3) = (3, 0)$, $\boldsymbol{\alpha}($  $^2) = (2, 1)$.

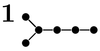
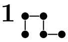
Definition 5.5. For a pair of partially labeled graphs A, B, define

$$m(A, B) = \boldsymbol{\alpha}([[A^2]][[B^2]]) - \boldsymbol{\alpha}([[AB]]^2).$$

Example 5.6. Consider $A =$  and $B =$ . Then

$[[A^2]] =$  ^{2}, $[[B^2]] =$  and $[[AB]] =$ .

So

$m(A, B) = \mathbf{1}$  $+ \mathbf{1}$  $- 2 \cdot \mathbf{1}$ .

Proof strategy for Theorem 5.1. We need to show that for every $d \geq 1$, the inequality

$$(\overline{H}^k - \underline{H}^{k+1})(\mathbf{x}) \geq 0 \tag{5.1}$$

is not valid for $\mathcal{S}_d$. Suppose (5.1) is not valid, then for each fixed $d \geq 1$, we have a point $\mathbf{z} \in \operatorname{trop}(\mathcal{S}_d)$ such that $\langle \mathbf{z}, \boldsymbol{\alpha}(\overline{H}^k) - \boldsymbol{\alpha}(\underline{H}^{k+1}) \rangle < 0$. Since $\operatorname{trop}(\mathcal{S}_d) = \operatorname{cl}(\operatorname{conv}(\log(\mathcal{S}_d)))$ and the inequality is strict, we can assume that $\mathbf{z} \in \log(\mathcal{S}_d)$. This means there exists $\mathbf{x} \in \mathcal{S}_d$ such that $z_C = \log(x_C)$ for all $C \in \mathcal{V}_d$. Exponentiating, we get that $(\overline{H}^k - \underline{H}^{k+1})(\mathbf{x}) < 0$. Since for each $d \geq 1$ there is such a $\mathbf{x}$ and $\mathbf{z}$, it follows that $\overline{H}^k - \underline{H}^{k+1}$ is not sos-testable and we are done.

Thus our task is to show that (5.1) is not a valid constraint for $\operatorname{trop}(\mathcal{S}_d)$. For the sake of contradiction, assume it is valid for some d, or equivalently that $\boldsymbol{\alpha}(\overline{H}^k) - \boldsymbol{\alpha}(\underline{H}^{k+1})$ is in the dual cone to $\operatorname{trop}(\mathcal{S}_d) = \log(\mathcal{S}^{2\times 2}_{\tilde{\mathcal{M}}_d})$. By Corollary 4.2, the dual cone is generated by the vectors $\{m(A, B) : A, B \in \tilde{\mathcal{B}}_d\}$. Thus

$$\boldsymbol{\alpha}(\overline{H}^k) - \boldsymbol{\alpha}(\underline{H}^{k+1}) = \sum_{A,B \in \tilde{\mathcal{B}}_d} \lambda_{A,B}\, m(A, B), \tag{5.2}$$

where $\lambda_{A,B} \geq 0$. We will now proceed in steps to derive a contradiction.

We first exhibit a point $\mathbf{y} \in \operatorname{trop}(\mathcal{S}_d)$ such that $\langle \mathbf{y}, \boldsymbol{\alpha}(\overline{H}^k) - \boldsymbol{\alpha}(\underline{H}^{k+1})\rangle$ is small. We prove this via Lemma 5.8 and Lemma 5.9.

Definition 5.7. For an unlabeled graph F, let $\delta_i(F)$ be the degree of vertex i. Define $L(F) := \sum g(\delta_i(F))$ where $g : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\leq 0}$ such that $g(0) = 0$, g is non-increasing and g is convex.

Lemma 5.8. *The point* $\mathbf{y} = (L(C))_{C \in \mathcal{V}_d}$ *lies in* $\operatorname{trop}(\mathcal{S}_d)$.

Proof. Since the dual cone to $\operatorname{trop}(\mathcal{S}_d)$ is spanned by $\{m(A,B) : A, B \in \tilde{\mathcal{B}}_d\}$, it is enough to check that $\langle \mathbf{y}, m(A,B)\rangle \geq 0$ for all $A, B \in \tilde{\mathcal{B}}_d$. We look carefully at $\langle \mathbf{y}, m(A,B)\rangle$.

Since $y_C = L(C) = \sum g(\delta_i(C))$, it suffices to understand the contribution of different types of vertices to $\langle \mathbf{y}, m(A,B)\rangle$. An unlabeled vertex of A (resp. B) of degree d leads to two vertices in A^2 (resp. B^2) both of degree d, and one vertex of degree d in AB. Such a vertex contributes $g(d) + g(d) - 2g(d) = 0$ to $\langle \mathbf{y}, m(A,B)\rangle$.

If a label is used in only one graph, say in A, and the vertex with this label has s fully labeled edges and t partially labeled edges, then in AB we get a vertex of degree $s+t$, while in A^2 we get a vertex of degree $s+2t$ and this vertex has no impact on B^2. The total contribution of such a vertex is thus $g(s+2t) - 2g(s+t)$. Note that $s + 2t \leq 2(s+t)$ so $g(s+2t) \geq g(2(s+t)) \geq 2g(s+t)$ where the first inequality follows since g is non-increasing and the second follows from the convexity of g and $g(0) = 0$.

The last case is if a label is used in both graphs. Suppose that in A it is adjacent to t_1 partially labeled edges, s_1 fully labeled edges that are also in B and u_1 fully labeled edges that are not in B. Similarly, suppose that this vertex in B is adjacent to t_2 partially labeled edges, s_2 fully labeled edges that are also in A and u_2 fully labeled edges that are not in A. Note that, by definition, $s_1 = s_2$. Then in A^2 we get a vertex with degree $2t_1 + s_1 + u_1$, and in B^2, a vertex with degree $2t_2 + s_2 + u_2$. In AB, we get a vertex of degree $t_1 + t_2 + s_1 + u_1 + u_2$. The total contribution of such a vertex to $\langle \mathbf{y}, m(A,B)\rangle$ is thus

$$\begin{aligned} &g(2t_1 + s_1 + u_1) + g(2t_2 + s_1 + u_2) - 2g(t_1 + t_2 + s_1 + u_1 + u_2) \\ &\qquad \geq 2g(t_1 + t_2 + s_1 + u_1/2 + u_2/2) - 2g(t_1 + t_2 + s_1 + u_1 + u_2) \geq 0. \end{aligned}$$

□

From now on, we let g be given by $g(m) = -m$ for $0 \leq m \leq p + \frac{3}{2}$ and $g(m) = -(p + \frac{3}{2})$ for $m > p + \frac{3}{2}$. Note that $g(0) = 0$, and g is convex and non-increasing. This g has the following effect on $L(F)$ for an unlabeled graph F: if the maximum degree of a vertex in F is $p+1$, then $g(\delta_i(F)) = -\delta_i(F)$ and hence $L(F) = \sum g(\delta_i(F)) = -2|E(F)|$.

Lemma 5.9. *We have* $\langle \mathbf{y}, \boldsymbol{\alpha}(\overline{H}^k) - \boldsymbol{\alpha}(\underline{H}^{k+1})\rangle = |E(\underline{H})|$.

Proof. By definition of $\overline{H}$ and $\underline{H}$, both have the same number of edges, and each vertex has degree at most $p+1$. Therefore, $\langle \mathbf{y}, \boldsymbol{\alpha}(\overline{H}^k)\rangle = -2k|E(\overline{H})|$. Similarly, $\langle \mathbf{y}, \boldsymbol{\alpha}(\underline{H}^{k+1})\rangle = -2(k+1)|E(\underline{H})|$, and the result holds. □

We know that any connected component of $\overline{H}$ is a trivial square and that each appears only once in $\overline{H}$ since $\overline{H}$ is a trivial square. Since $\overline{H}$ and $\underline{H}$ are distinct, there must exist a connected component C in $\overline{H}$ not in $\underline{H}$.

Lemma 5.10. *For every $A, B \in \tilde{\mathcal{B}}_d$ such that the Cth coordinate of $m(A,B)$ is positive, $\langle \mathbf{y}, m(A,B)\rangle \geq \frac{1}{2}(z_A + z_B) > 0$ where z_A is the number of fully labeled copies of C that appear in A but not in AB, and z_B is the number of fully labeled copies of C that appear in B but not in AB. Moreover, we also have $\langle \mathbf{1}_C, m(A,B)\rangle \leq z_A + z_B$.*

Proof. If the Cth component of $m(A,B)$ is positive, then this means that $[[A^2]][[B^2]]$ must contain at least one copy of C. Since C is a trivial square, at least A or B must contain either an unlabeled copy of C or a fully labeled copy of C.

Suppose A contains l_A fully labeled copies of C that appear in AB but not in B, and similarly, suppose B contains l_B fully labeled copies of C that appear in AB but not in A. Suppose there are l_{AB} fully labeled copies of C that appear in both A and B (and thus also in AB). Suppose there are z_A fully labeled copies of C that appear in A that do not appear in AB, and similarly, z_B fully labeled copies of C that appear in B but not in AB. Finally, suppose A and B respectively contain u_A and u_B unlabeled copies of C. Then the Cth component of $\boldsymbol{\alpha}([[A^2]][[B^2]])$ is $l_A + l_{AB} + z_A + 2u_A + l_B + l_{AB} + z_B + 2u_B$ and the Cth component of $\boldsymbol{\alpha}([[AB]]^2)$ is $2(l_A + l_{AB} + l_B + u_A + u_B)$. Thus the Cth component of $m(A,B)$ is $z_A + z_B - l_A - l_B$, and we have $\langle \mathbf{1}_C, m(A,B)\rangle \leq z_A + z_B$ proving the second claim in the lemma.

To complete the proof of the first claim, we can assume that $z_A + z_B \geq 1$ since the Cth component of $m(A,B)$ is assumed to be strictly positive. Without loss of generality, assume $z_A \geq 1$, i.e., there is a fully labeled copy of C in A, say C_l with labels $1, 2, \ldots, r$, that does not appear in AB. For this to be the case, B must contain at least some labeled vertex $b \in \{1, 2, \ldots, r\}$ that is adjacent to some partially or fully labeled edge not in C_l.

Recall that from Lemma 5.8, $\langle \mathbf{y}, m(A,B)\rangle \geq 0$ and it can be obtained as a sum of contributions of different vertices separately, each of which is nonnegative. We show that b will contribute at least $\frac{1}{2}$ to $\langle \mathbf{y}, m(A,B)\rangle$.

Indeed, in A, we know that b is adjacent to $t_1 = 0$ partially labeled edges and suppose it is adjacent to s_1 fully labeled edges that are also in B and u_1 fully labeled edges that are not in B, where $s_1 + t_1 \in \{p, p+1\}$ by definition of $\overline{H}$. Further, in B, suppose b is adjacent to t_2 partially labeled edges, $s_2 = s_1$ fully labeled edges that are also in A and u_2 fully labeled edges that are not in A. We know that $t_2 + u_2 \geq 1$. Note that this implies that b in AB will always contribute at least one more edge than in A. So the contribution of b to $\langle \mathbf{y}, m(A,B)\rangle$ is at least $g(2t_1 + s_1 + u_1) + g(2t_2 + s_1 + u_2) - 2g(t_1 + t_2 + s_1 + u_1 + u_2) = g(s_1 + u_1) + g(2t_2 + s_1 + u_2) - 2g(t_2 + s_1 + u_1 + u_2)$. Let's consider a few cases.

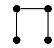
Either b in B is adjacent to at least $p+2$ edges, i.e., $2t_2 + s_1 + u_2 \geq p+2$ and $g(2t_2 + s_1 + u_2) = -(p + \frac{3}{2})$, but b in AB is adjacent to at most $p+1$ edges. In that case, b in A can only be adjacent to p edges, i.e., $s_1 + u_1 = p$ since $u_2 + t_2 \geq 1$. So b in AB is adjacent to exactly $p+1$ edges and the contribution of b to $\langle \mathbf{y}, m(A,B)\rangle$ is $-p - (p + \frac{3}{2}) - 2(-(p+1)) = \frac{1}{2}$.

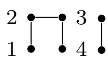
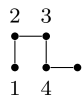
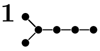
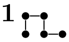
If b in B is adjacent to at least $p+2$ edges and b in AB is also adjacent to at least $p+2$ edges, then the contribution is at least $-(p+1) - (p + \frac{3}{2}) - 2(-(p + \frac{3}{2})) = \frac{1}{2}$.

Otherwise, if b in B is adjacent to at most $p+1$ edges and b in A is adjacent to p edge, then b in AB is adjacent to at least $p+1$ edges, so the contribution is at least $-p - (p+1) - 2(-(p+1)) = 1$. On the other hand, if b in B is adjacent to at most $p+1$ edges and b in A is adjacent to $p+1$ edges, then b in AB is adjacent to at least $p+2$ edges, so the contribution is at least $-(p+1) - (p+1) - 2(-(p + \frac{3}{2})) = 1$.

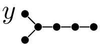
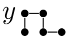
Finally, we know that every other vertex of C_l in A contributes at least zero. Thus the contribution of C_l to $\langle \mathbf{y}, m(A,B)\rangle$ is at least $\frac{1}{2}$.

Note that the same argument holds for every one of the $z_A + z_B$ fully labeled copies of C in A or B that do not appear in AB. Thus, we get a total contribution of at least $\frac{1}{2}(z_A + z_B)$ to $\langle \mathbf{y}, m(A,B)\rangle$. $\square$

Example 5.11. Consider $C =$  which is a connected component of $\overline{H} =$  but not $\underline{H} =$ ^{3}. Here $p+1 = 2$, and so $g(1) = -1, g(2) = -2$ and $g(3) = -2.5$. We saw that

if $A =$  and $B =$ , then $m(A,B) = \mathbf{1}$ $+ \mathbf{1}$ $- 2 \cdot \mathbf{1}$.

The component indexed by C in $m(A,B)$ is positive, and

y $= -6$, y $= -9.5$ and y $= -8$.

Thus, as proved in Lemma 5.10,

$$\langle \mathbf{y}, m(A,B)\rangle = -6 - 9.5 - 2(-8) = 0.5 > 0.$$

Proof of Theorem 5.1. We will contradict Equation (5.2) by a simple counting argument.

Let C be a connected component of $\overline{H}$ that does not appear in $\underline{H}$. Equating the coordinate indexed by C on both sides of Equation (5.2), we get

$$k \leq \sum_{(A,B)\in\mathcal{I}^+} \lambda_{AB}\langle m(A,B), \mathbf{1}_C\rangle,$$

where $\mathcal{I}^+$ indexes the pairs (A,B) such that $A, B \in \tilde{\mathcal{B}}_d$ and $\langle \mathbf{1}_C, m(A,B)\rangle > 0$. From Lemma 5.10, we know that $\langle \mathbf{1}_C, m(A,B)\rangle \leq z_A + z_B$, so

$$k \leq \sum_{(A,B)\in\mathcal{I}^+} \lambda_{AB}(z_A + z_B).$$

Recall that $\langle \mathbf{y}, \boldsymbol{\alpha}(\overline{H}^k) - \boldsymbol{\alpha}(\underline{H}^{k+1})\rangle = |E(\underline{H})|$ from Lemma 5.9, so

$$\begin{aligned}
|E(\underline{H})| &= \left\langle \mathbf{y}, \sum_{A,B\in\tilde{\mathcal{B}}_d} \lambda_{AB} m(A,B) \right\rangle \\
&\geq \sum_{(A,B)\in\mathcal{I}^+} \lambda_{AB}\langle \mathbf{y}, m(A,B)\rangle \\
&\geq \sum_{(A,B)\in\mathcal{I}^+} \lambda_{AB} \cdot \frac{1}{2}(z_A + z_B) \\
&\geq \frac{1}{2}k,
\end{aligned}$$

where the second line follows from the first because $\langle \mathbf{y}, m(A,B)\rangle \geq 0$ and $\lambda_{AB} \geq 0$ for all $A, B \in \tilde{\mathcal{B}}_d$; the third follows from the second by Lemma 5.10; and the last implication follows from $\sum_{(A,B)\in\mathcal{I}^+} \lambda_{AB}(z_A + z_B) \geq k$. Therefore, if Equation (5.2) holds, then $k \leq 2|E(\underline{H})|$. This implies that if $k \geq 2|E(\underline{H})| + 1$ as assumed in Theorem 5.1, then Equation (5.2) cannot hold, completing the proof of the theorem. $\square$

From the third section of [5], we know that the existence of a sos certificate in the gluing algebra we presented here is equivalent to the existence of a sos certificate in Lovász-Szegedy's gluing algebra [23], Hatami-Norine's gluing algebra [15] and Razborov's flag algebra [30]. Since the existence of a rational sos certificate for some graph combination a relies on the existence of two sos b and c such that $ab = c$, we get Corollary 5.12.

Corollary 5.12. *The binomial graph combinations* $\overline{H}^k - \underline{H}^{k+1}$ *for* $k \geq 2|E(\overline{H})|+1$, *and* $\overline{H} - \underline{H}$, *satisfying the conditions in Theorem 5.1 are not rational sos in Lovász-Szegedy's gluing algebra, Hatami-Norine's gluing algebra or Razborov's flag algebra.*

Note that $\overline{H} - \underline{H}$ needs to be translated to induced densities first in the case of Razborov's flag algebra.

References

[1] R. Ahlswede and G. O. H. Katona, *Graphs with maximal number of adjacent pairs of edges*, Acta Math. Acad. Sci. Hungar. **32** (1978), no. 1-2, 97–120, DOI 10.1007/BF01902206. MR505076

[2] Daniele Alessandrini, *Logarithmic limit sets of real semi-algebraic sets*, Adv. Geom. **13** (2013), no. 1, 155–190, DOI 10.1515/advgeom-2012-0020. MR3011539

[3] Xavier Allamigeon, Stéphane Gaubert, and Mateusz Skomra, *Tropical spectrahedra*, Discrete Comput. Geom. **63** (2020), no. 3, 507–548, DOI 10.1007/s00454-020-00176-1. MR4074332

[4] Grigoriy Blekherman and Shyamal Patel, *Threshold graphs maximize homomorphism densities*, Preprint, arXiv:2002.12117, 2020.

[5] Grigoriy Blekherman, Annie Raymond, Mohit Singh, and Rekha R. Thomas, *Simple graph density inequalities with no sum of squares proofs*, Combinatorica **40** (2020), no. 4, 455–471, DOI 10.1007/s00493-019-4124-y. MR4150878

[6] David Conlon, Jacob Fox, and Benny Sudakov, *An approximate version of Sidorenko's conjecture*, Geom. Funct. Anal. **20** (2010), no. 6, 1354–1366, DOI 10.1007/s00039-010-0097-0. MR2738996

[7] David Conlon, Jeong Han Kim, Choongbum Lee, and Joonkyung Lee, *Some advances on Sidorenko's conjecture*, J. Lond. Math. Soc. (2) **98** (2018), no. 3, 593–608, DOI 10.1112/jlms.12142. MR3893193

[8] Paul Erdős, László Lovász, and Joel Spencer, *Strong independence of graphcopy functions*, Graph theory and related topics (Proc. Conf., Univ. Waterloo, Waterloo, Ont., 1977), Academic Press, New York-London, 1979, pp. 165–172. MR538044

[9] Paul Erdős, László Lovász, and Joel Spencer, *Strong independence of graphcopy functions*, Graph Theory and Related Topics (1979), pp. 165–172.

[10] Victor Falgas-Ravry and Emil R. Vaughan, *Applications of the semi-definite method to the Turán density problem for 3-graphs*, Combin. Probab. Comput. **22** (2013), no. 1, 21–54, DOI 10.1017/S0963548312000508. MR3002572

[11] Roman Glebov, Andrzej Grzesik, Ping Hu, Tamas Hubai, Daniel Kral, and Jan Volec, *Densities of 3-vertex graphs*, Preprint, arXiv:1610.02446, 2016.

[12] Oded Goldreich, Shafi Goldwasser, and Dana Ron, *Property testing and its connection to learning and approximation*, J. ACM **45** (1998), no. 4, 653–750, DOI 10.1145/285055.285060. MR1675099

[13] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1988. Reprint of the 1952 edition. MR944909

[14] Hamed Hatami, Jan Hladký, Daniel Kráľ, Serguei Norine, and Alexander Razborov, *On the number of pentagons in triangle-free graphs*, J. Combin. Theory Ser. A **120** (2013), no. 3, 722–732, DOI 10.1016/j.jcta.2012.12.008. MR3007147

[15] Hamed Hatami and Serguei Norine, *Undecidability of linear inequalities in graph homomorphism densities*, J. Amer. Math. Soc. **24** (2011), no. 2, 547–565, DOI 10.1090/S0894-0347-2010-00687-X. MR2748400

[16] Hamed Hatami and Sergey Norin, *On the boundary of the region defined by homomorphism densities*, J. Comb. **10** (2019), no. 2, 203–219. MR3912211

[17] Hao Huang, Nati Linial, Humberto Naves, Yuval Peled, and Benny Sudakov, *On the 3-local profiles of graphs*, J. Graph Theory **76** (2014), no. 3, 236–248, DOI 10.1002/jgt.21762. MR3200287

[18] G. Katona, *A theorem of finite sets*, Theory of graphs (Proc. Colloq., Tihany, 1966), Academic Press, New York, 1968, pp. 187–207. MR0290982

[19] Jeong Han Kim, Choongbum Lee, and Joonkyung Lee, *Two approaches to Sidorenko's conjecture*, Trans. Amer. Math. Soc. **368** (2016), no. 7, 5057–5074, DOI 10.1090/tran/6487. MR3456171

[20] Joseph B. Kruskal, *The number of simplices in a complex*, Mathematical optimization techniques, Univ. California Press, Berkeley, Calif., 1963, pp. 251–278. MR0154827

[21] László Lovász, *Graph homomorphisms: open problems*, 2008. Available at `http://web.cs.elte.hu/~lovasz/problems.pdf`.

[22] László Lovász, *Large networks and graph limits*, American Mathematical Society Colloquium Publications, vol. 60, American Mathematical Society, Providence, RI, 2012, DOI 10.1090/coll/060. MR3012035

[23] László Lovász and Balázs Szegedy, *Random graphons and a weak Positivstellensatz for graphs*, J. Graph Theory **70** (2012), no. 2, 214–225, DOI 10.1002/jgt.20611. MR2921000

[24] Diane Maclagan and Bernd Sturmfels, *Introduction to tropical geometry*, Graduate Studies in Mathematics, vol. 161, American Mathematical Society, Providence, RI, 2015, DOI 10.1090/gsm/161. MR3287221

[25] Grigory Mikhalkin, *Enumerative tropical algebraic geometry in* $\mathbb{R}^2$, J. Amer. Math. Soc. **18** (2005), no. 2, 313–377, DOI 10.1090/S0894-0347-05-00477-7. MR2137980

[26] Tim Netzer and Andreas Thom, *Positivstellensätze for quantum multigraphs*, J. Algebra **422** (2015), 504–519, DOI 10.1016/j.jalgebra.2014.08.053. MR3272089

[27] V. Nikiforov, *The number of cliques in graphs of given order and size*, Trans. Amer. Math. Soc. **363** (2011), no. 3, 1599–1618, DOI 10.1090/S0002-9947-2010-05189-X. MR2737279

[28] Aaron Potechin, *Sum of squares lower bounds from symmetry and a good story*, 10th Innovations in Theoretical Computer Science, LIPIcs. Leibniz Int. Proc. Inform., vol. 124, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2019, pp. Art. No. 61, 20. MR3899855

[29] Annie Raymond, Mohit Singh, and Rekha R. Thomas, *Symmetry in Turán sums of squares polynomials from flag algebras*, Algebr. Comb. **1** (2018), no. 2, 249–274, DOI 10.5802/alco. MR3856524

[30] Alexander A. Razborov, *Flag algebras*, J. Symbolic Logic **72** (2007), no. 4, 1239–1282, DOI 10.2178/jsl/1203350785. MR2371204

[31] Alexander A. Razborov, *On the minimal density of triangles in graphs*, Combin. Probab. Comput. **17** (2008), no. 4, 603–618, DOI 10.1017/S0963548308009085. MR2433944

[32] Alexander A. Razborov, *On 3-hypergraphs with forbidden 4-vertex configurations*, SIAM J. Discrete Math. **24** (2010), no. 3, 946–963, DOI 10.1137/090747476. MR2680226

[33] Christian Reiher, *The clique density theorem*, Ann. of Math. (2) **184** (2016), no. 3, 683–707, DOI 10.4007/annals.2016.184.3.1. MR3549620

[34] Balazs Szegedy, *An information theoretic approach to Sidorenko's conjecture*, Preprint, `arXiv:1406.6738`, 2014.

[35] Josephine Yu, *Tropicalizing the positive semidefinite cone*, Proc. Amer. Math. Soc. **143** (2015), no. 5, 1891–1895, DOI 10.1090/S0002-9939-2014-12428-2. MR3314099

School of Mathematics, Georgia Institute of Technology, 686 Cherry Street, Atlanta, Georgia 30332
Email address: greg@math.gatech.edu

Department of Mathematics and Statistics, Lederle Graduate Research Tower, 1623D, University of Massachusetts Amherst, 710 N. Pleasant Street, Amherst, Massachusetts 01003
Email address: raymond@math.umass.edu

H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, 755 Ferst Drive, NW, Atlanta, Georgia 30332
Email address: mohit.singh@isye.gatech.edu

Department of Mathematics, University of Washington, Box 354350, Seattle, Washington 98195
Email address: rrthomas@uw.edu