Polytopes from Subgraph Statistics
Alexander Engström, Patrik Norén

To cite this version:

HAL Id: hal-01215115
https://hal.inria.fr/hal-01215115
Submitted on 13 Oct 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Polytopes from Subgraph Statistics

Alexander Engström† and Patrik Norén‡

1 Department of Mathematics, University of California, Berkeley, CA, USA
2 Department of Mathematics, KTH, The Royal Institute of Technology, Stockholm, Sweden

Abstract. We study polytopes that are convex hulls of vectors of subgraph densities. Many graph theoretical questions can be expressed in terms of these polytopes, and statisticians use them to understand exponential random graph models.

Relations among their Ehrhart polynomials are described, their duals are applied to certify that polynomials are non-negative, and we find some of their faces.

For the general picture we inscribe cyclic polytopes in them and calculate volumes. From the volume calculations we conjecture that a variation of the Selberg integral indexed by Schur polynomials has a combinatorial formula. We inscribe polynomially parametrized sets, called curvy zonotopes, in the polytopes and show that they approximate the polytopes arbitrarily close.

Résumé. Nous étudions les polytopes qui sont les enveloppes convexes des vecteurs des densités de sous-graphe. Beaucoup de questions théoriques de graphe peuvent être exprimées en termes de ces polytopes, et les statisticiens les utilisent pour comprendre les modèles de graphes aléatoires exponentielles.

Des relations parmi leurs polynômes d’Ehrhart sont décriptes leurs duals sont appliqués pour certifier que les polynômes sont non négatifs, et nous trouvons certaines de leurs faces.

Pour la description générale nous inscrivons les polytopes cycliques dans eux et calculons les volumes. D’après les calculs de volume, nous conjecturons qu’une variation de l’intégrale de Selberg indexées par des polynômes de Schur a une formule combinatoire. Nous inscrivons polynomialement les ensembles paramétrisés appelés “curvy zonotopes” dans les polytopes et montrons qu’ils sont arbitrairement proches de polytopes.

Keywords: polytopes, subgraph statistics, exponential random graph models, curvy zonotopes, graph limits

1 Introduction

In this paper we study polytopes from subgraph statistics. For any two graphs $F$ and $G$, the $F$-subgraph density of $G$, denoted $t(F,G)$, is the proportion of injective maps from $V(F)$ to $V(G)$ sending edges of $F$ onto edges of $G$. For any vector $\mathbf{F}$ of $d$ graphs $F_1, F_2, \ldots, F_d$; and any graph $G$, we get a point

†Alexander Engström is a Miller Research Fellow 2009-2012 at UC Berkeley, and gratefully acknowledges support from the Adolph C. and Mary Sprague Miller Institute for Basic Research in Science. alex@math.berkeley.edu
‡Patrik Norén gratefully acknowledges support from the Wallenberg foundation. pnore@math.kth.se

1365–8050 © 2011 Discrete Mathematics and Theoretical Computer Science (DMTCS), Nancy, France
\( t(F, G) = (t(F_1, G), t(F_2, G), \ldots, t(F_d, G)) \) in \( \mathbb{R}^d \). The convex hull of the collection of all such points from \( n \) vertex graphs is the \textit{polytope from subgraph statistics}

\[
P_{F;n} = \text{conv} \left\{ t(F, G) \mid G \text{ is a graph on } n \text{ vertices} \right\}.
\]

The polytope \( P_{(K_3,c_4,K_4\backslash e);6} \) is drawn in Figure 1 and in Figure 2 is a combinatorial representation of its vertices and edges. If larger examples of polytopes from subgraph statistics look anything like in Figures 1 and 2 then it would be very difficult to give an explicit facet description. And indeed many hard theorems and conjectures in extremal graph theory can be rephrased as questions about these polytopes, making a complete facet description probably impossible in general. In Figure 2 we tabulated the vertices by the actual subgraph counts and not the proportions \( t(F, G) \). This defines the lattice polytope \( P_{L,F;n} \), a rescaling of \( P_{F;n} \). It should be noted that several graphs could have the same subgraph statistics, and that

![Fig. 1: The polytope \( P_{(K_3,c_4,K_4\backslash e);6} \).](image)

even if \( t(F, G_1) \) and \( t(F, G_2) \) are different vertices on the same facet, it is not necessary that \( G_1 \) and \( G_2 \) are related in any sense, for example as subgraphs. This is illustrated in Figure 3.

We got interested in studying the polytopes from subgraph statistics after several questions were raised about them by Rinaldo, Fienberg and Zhou [15]. They investigated maximum likelihood estimation for exponential random graph models and realized that its behavior is closely linked to the geometry of the polytopes. For some vector \( F \) of \( d \) graphs and model parameter \( \gamma \in \mathbb{R}^d \), the probability of observing the \( n \) vertex graph \( G \) is

\[
p_G = \frac{1}{Z(\gamma)} e^{\gamma \cdot t(F, G)}
\]

where \( Z(\gamma) \) is the normalizing partition function. Given an empirical distribution of \( n \) vertex graphs, the object of a maximum likelihood estimation is to find the best parameter \( \gamma \) explaining the observations.
**Fig. 2:** A combinatorial representation of the vertices and edges of $P_{(K_5,C_4,K_4 \setminus e)}$, indexed by the actual subgraph counts.

**Fig. 3:** The graphs underlying the statistics of a piece of the polytope in Figures 1 and 2. Dotted graph edges could be included or not. Recall that the subgraphs counted are $(K_3, C_4, K_4 \setminus e)$. 
The graph vector $F$ is usually determined by the applications of the model, in the social sciences small graphs as stars and triangles are common \[16\].

Before embarking on more general results about the polytopes from subgraph statistics, we point out some easy propositions about the facets of certain polytopes. Proving them is a good exercise to get acquainted with the polytopes.

**Proposition 1.1** Let $F$ be a vector of $d$ graphs of order at most $n$, with no pairs in a subgraph relation. Then $x_i = 0$ defines a facet of $P_{F,n}$, containing $\emptyset$ for all $1 \leq i \leq d$.

**Proposition 1.2** Let $F$ be a vector of graphs including the edge $K_2$. Then the line from $1$ to $t(F, K_n \setminus e)$ is on an edge of $P_{F,n}$.

**Proposition 1.3** Let $F$ be a vector of $d$-regular graphs, and let $G_i$ be the complete graph $K_n$ with a star on $i$ edges removed from it. Then the points $t(F, G_i)$ are on a line in $P_{F,n}$. If this line is on the boundary of $P_{F,n}$, then so it is for any $n' > n$.

The first non-trivial result follows from a graph reconstruction type argument.

**Proposition 1.4** Let $F$ be a vector of graphs of order at most $n$. Then $P_{F,n''} \subseteq P_{F,n'}$ if $n'' \geq n' \geq n$.

The vertices of the lattice polytope $P_{F,n}$ are the actual subgraph counts and not the relative densities. For any lattice polytope $P$ the number of lattice points in $kP$ is the Ehrhart polynomial $E_P(k)$ (see chapter 12 of \[14\]). This is the translation of Proposition \[14\] into the lattice polytope setting.

**Proposition 1.5** Let $F$ be a vector of graphs of order $l$. Then $E_{P_{F,n''}}(\binom{n''}{l}) \leq E_{P_{F,n'}}(\binom{n'}{l})$ for all positive integers $k$, if $n'' \geq n' \geq l$.

In the proposition it is required that all graphs in $F$ are of the same order, and this can partially be generalized by adding isolated vertices to get graphs of the same order. If $F$ is the graph vector of the path on three vertices and the triangle, then

$$E_{P_{F,3}}\left(\binom{4}{3}k\right) = E_{P_{F,3}}\left(\binom{3}{3}k\right) = 8k^2 + 6k + 1$$

and

$$E_{P_{F,3}}\left(\binom{5}{3}k\right) = 50k^2 + 15k + 1 \geq 48k^2 + 13k + 1 = E_{P_{F,5}}\left(\binom{3}{3}k\right).$$

2 The spine of polytopes

Since it’s hard to understand the polytopes exactly, we now try to inscribe more accessible polytopes and varieties within them. For a vector $F$ of $m$ graphs, the spine is the generalized moment curve

$$\{(p^{e_1}, p^{e_2}, \ldots, p^{e_m}) \mid 0 \leq p \leq 1\},$$

where $e_i$ is the number of edges in $F_i$. In the Erdős–Rényi random graph model $G(n, p)$ edges are included independently with probability $p$. The expected value of $t(F, G)$ for $G \in \mathcal{G}(n, p)$ is $(p^{e_1}, p^{e_2}, \ldots, p^{e_m})$, proving the following proposition.

**Proposition 2.1** For any vector $F$ of graphs of order at most $n$ with $e$ edges, the spine $\{(p^{e_1}, p^{e_2}, \ldots, p^{e_m}) \mid 0 \leq p \leq 1\}$ is in $P_{F,n}$. 

In Figure 4 is the polytope $P(\kappa_3, c_4, K_4 \backslash \{r\}; 6)$ from Figure 1 drawn with its spine. The point of Proposition 2.1 is that the spine is a generalized moment curve inside $P_{\mathbb{R}^n}$. The convex hull of a finite number of points on the spine is a cyclic polytope (this can be seen directly by using generalized van der Monde matrices instead of the ordinary one in Ziegler’s textbook derivation of the combinatorial structure of cyclic polytopes [21].) This shows that there is a cyclic polytope inscribed in $P_{\mathbb{R}^n}$. The convex hull of all of the spine is not a polytope, but its boundary can be algebraically described. In Figure 5 is the spine from Figure 4 drawn with its convex hull. Since the boundary structure of the convex hull in Figure 5 is not very clear from this angle, we include in Figure 6 the same spine with its convex hull, but from another perspective.

2.1 Volumes

Inside our polytopes we have convex hulls of generalized moment curves and their volumes bound the volumes of polytopes from subgraph statistics. For the ordinary moment curve $\{(p, p^2, \ldots, p^d) \mid 0 \leq p \leq 1\}$ the volume of its convex hull was calculated by Karlin and Shapley [12]. An interesting curiosity is that Selberg and Shapley were in Princeton at the same time, and that this volume calculation was the first application of the now famous Selberg integral [8], by then only available in Norwegian and published by Selberg in a magazine for college math teachers [18].

**Theorem 2.2** The volume of the convex hull of the $2m$–dimensional spine

$$\text{Vol}(\text{conv} \{(p^{e_1}, p^{e_2}, \ldots, p^{e_{2m}}) \mid 0 \leq p \leq 1\})$$

is

$$\frac{1}{(2m)!} \int_{[0,1]^m} S_\lambda(x_1, x_1, x_2, x_2, \ldots, x_m, x_m) \prod_{0 \leq i < j \leq m} (x_i - x_j)^4 \, dx_1 \cdots dx_m,$$
Fig. 5: The spine in Figure drawn with its convex hull.

Fig. 6: The spine with its convex hull from Figure drawn from another perspective.
where $\lambda_i = e_i - i + 1$ in the Schur polynomial $S_\lambda$, and it is assumed that $0 < e_1 < e_2 < \cdots < e_{2m}$.

For the definition and properties of the Schur polynomials we refer to Sagan [17]. For applications, it is perhaps most important that they are symmetric with non-negative coefficients, and are easy to compute. We proved Theorem 2.2 by approximating the convex hull by a cyclic polytope with $n$ vertices and show that the volume converges to the integral above. The difference in volume between the cyclic polytope on the $n$ vertices given by $p = 0, 1/(n - 1), \ldots, 1$ and the convex hull of the spine is less than $4d(e_1 + e_2 + \cdots + e_d)n^{-1}$. Using the Selberg integral formula [8, 18] we get an explicit form for consecutive $e_i$'s.

**Corollary 2.3** The volume of the convex hull of the $2m$–dimensional spine

$$\text{Vol}(\text{conv}\{(p^e, p^{e+1}, \ldots, p^{e+2m-1}) \mid 0 \leq p \leq 1\})$$

is

$$\frac{1}{(2m)!m!} \prod_{j=0}^{m-1} \frac{(2e + 2j)(2j)!}{(2e + 2(m + j) - 1)!}.$$ 

The integral in Theorem 2.2 evaluates, according to our extensive experiments, to combinatorial looking expressions like the one from Selberg’s integral formula in Corollary 2.3. Even if special cases can be treated by variations of the Selberg’s integral formula [8] there doesn’t seem to be a general formula in the literature [9].

**Conjecture 2.4** For any Schur polynomial $S_\lambda$ there is an explicit combinatorial formula for

$$\int_{[0,1]^m} S_\lambda(x_1, x_2, x_2, \ldots, x_m, x_m) \prod_{0 \leq i < j \leq m} (x_i - x_j)^4 \, dx_1 \cdots dx_m.$$ 

For the odd dimensional spines we have similar results.

### 2.2 Duality

As for polytopes there is a duality theory for convex hulls of algebraic sets [3]. The dual of the convex hull of the moment curve $\{(p, p^2, \ldots, p^n) \mid 0 \leq p \leq 1\}$ parametrizes the degree $n$ polynomials that are non-negative on the interval $[0, 1]$. The convex hulls of generalized moment curves are inside polytopes from subgraph statistics, so the polytopes can be used to certify that polynomials are non-negative.

**Proposition 2.5** Let $P$ be a polytope containing the generalized moment curve $\{(p^{e_1}, p^{e_2}, \ldots, p^{e_d}) \mid 0 \leq p \leq 1\}$. If $(c_1, c_2, \ldots, c_d) \cdot v \geq -1$ for all vertices $v$ of $P$ then the polynomial $1 + c_1x^{e_1} + c_2x^{e_2} + \cdots + c_dx^{e_d}$ is non-negative on the interval $[0, 1]$. 

Our running example $P(K_5, c_1, K_4 \setminus e)$ in Figure 1 is perhaps not the most interesting polytope to certify non-negativity with, but we will use it in an example anyways. The polynomial $p(x) = 1 - \frac{16}{7}x^3 + \frac{11}{2}x^4 - \frac{1}{7}x^5$ is non-negative on $[0, 1]$ since $(-\frac{16}{7}, \frac{11}{2}, -\frac{1}{7}) \cdot v \geq -1$ for all vertices $v$ of $P(K_5, c_1, K_4 \setminus e)$. Note that the point $(-\frac{16}{7}, \frac{11}{2}, -\frac{1}{7})$ is dual to the facet with vertices $(8/20, 10/45, 16/90), (10/20, 15/45, 30/90), (5/20, 3/45, 6/90)$, which one can find using Figure 2.
3 Curvy zonotopes

In the previous section we used the spine to construct a cyclic polytope inscribed in the polytopes under investigation. We didn’t prove any theorem about the ratio between the volume of the polytope and its inscribed cyclic polytope, and perhaps it could be arbitrarily bad for some large graph vectors $\textbf{F}$. We got the spine as the expected value of subgraph densities for different $p$ in the Erdős-Rényi random graph model $G(n, p)$. Lovász and Szegedy [13] introduced an extremely general random graph model $G(n, W)$ where $W$ is any symmetric measurable function from $[0, 1]^2$ to $[0, 1]$, and hiding in the background of this section, and essential for all our proofs, is the theory of graph limits. But we state our results without this machinery to make them more accessible.

We remark that the route to proving these results is not to use the graph limit results of Lovász and Szegedy [13] right off. It was realized by Diaconis and Janson [6] and Aldous [2] that the theory of graph limits is a reinterpretation of instances of old very abstract results on exchangeable sequences by Aldous [1] and Hoover [10]. Fortunately there is a new textbook by Kallenberg [11] that covers the relevant probability theory in chapter 7.

Before introducing the curvy zonotopes that generalizes spines, we define the limit object

$$P_{\textbf{F};\infty} = \bigcap_{n' \geq n} P_{\textbf{F};n'},$$

where $n$ is some integer not smaller than the order of any graph in $\textbf{F}$. Note that $P_{\textbf{F};\infty}$ is closed and convex, and should be viewed as the limit of the infinite sequence

$$P_{\textbf{F};n'} \supseteq P_{\textbf{F};n'+1} \supseteq P_{\textbf{F};n'+2} \supseteq \cdots.$$  

The first non-trivial result on $P_{\textbf{F};\infty}$ was proved by Bollobás [4, 5].

**Proposition 3.1** The limit object $P_{(K_2, K_m); \infty}$ is the convex hull of $(1, 1)$ and

$$\left\{ \left( 1 - \frac{1}{k}, \frac{m!}{k^m} \binom{k}{m} \right) \mid k = 1, 2, 3, \ldots \right\}.$$  

Considering Turán’s theorem it isn’t strange that complete $k$-partite graphs are around: the number $\frac{m!}{k^m} \binom{k}{m}$ is the limit of $t((K_2, K_m); K_t,t,\ldots,t)$ as the $k$-partite graph $K_t,t,\ldots,t$ grows larger as $t \to \infty$. The limit object $P_{\textbf{F};\infty}$ is contained in $P_{\textbf{F};n}$ in a fairly strong sense.

**Theorem 3.2** Let $x \neq 0, 1$ be a point in $P_{\textbf{F};\infty}$ for some vector of graphs $\textbf{F}$. If $n$ is not smaller than the order of some graph in $\textbf{F}$, then there is a polytope $P_x \subset P_{\textbf{F};n}$ with $x$ in its interior.

Note that the polytope $P_x$ could be of lower dimension than $P_{\textbf{F};n}$.

**Definition 3.3** Let $\textbf{F}$ be a vector of $d$ graphs and $n$ a positive integer. The curvy zonotope is

$$Z_{\textbf{F};n} = \left\{ (p_{F_1,n}(x), p_{F_2,n}(x), \ldots, p_{F_d,n}(x) \mid x \in [0, 1]^{n^2} \right\}$$

where

$$p_{F,n}(x_1, x_2, \ldots, x_{nn}) = \frac{1}{n|\mathcal{F}|} \sum_{\phi: V(\mathcal{F}) \to [n]} \prod_{ij \in E(\mathcal{F})} x_{\phi(i)\phi(j)}.$$
When all the polynomials $p_{F;n}(x)$ are linear $Z_{F;n}$ is a zonotope. But usually they are polynomials and $Z_{F;n}$ is a curvy zonotope. The curvy zonotope $Z(K_3,C_4,K_4\setminus e);2$, which looks like a melted toblerone, is drawn in Figure 7. As drawn in the Figure 7 the curvy zonotope is inscribed in the polytope $P(K_3,C_4,K_4\setminus e);6$. But it can actually be inscribed in the limit object $P(K_3,C_4,K_4\setminus e);\infty$ according to this proposition.

**Proposition 3.4** For any vector $F$ of graphs and positive integer $n$, $Z_{F;n} \subseteq P_{F;\infty}$.

Together with the following proposition, this shows that the polytopes from subgraph statistics and their limits are full dimensional.

**Proposition 3.5** For any vector $F$ of $d$ graphs and positive integer $n$, $Z_{F;n}$ is homeomorphic to a $d$-dimensional ball.

The spine is inside the curvy zonotope, just set $x_{11} = \cdots = x_{nn}$ in all the polynomials to recover it.

The convex hull of the spine is possible to describe very explicit by taking the limit of cyclic polytopes and then Gale’s evenness condition define the facets. For curvy zonotopes the convex hull is not as easily described, but the situation is fairly good. From an algebraic geometry perspective, calculating the algebraic boundary of the convex hull of a set parametrized by polynomials, is a nice situation [19, 20].

Since the curvy zonotopes are in the polytopes from subgraph statistics, so are their convex hulls. In Figure 8 is the convex hull of the curvy zonotope in Figure 7. Theorem 3.6 is our main result on curvy zonotopes, it states that their convex hulls converge towards the limit object $P_{F;\infty}$. In the core of the proof is a method for constructing convergent graph sequences from [13] by the weak Szemeredi regularity lemma as in [7].

**Theorem 3.6** Let $F$ be a vector of $d$ graphs on at most $e$ edges. If $\varepsilon > 0$ and $n > \left[2^{1600d^3}e^2\varepsilon^{-2}\right]$ then

$$0 \leq \text{Vol}(P_{F;\infty}) - \text{Vol(} \text{conv } Z_{F;n}) < \varepsilon.$$
We end this section by stating a result for the readers who knows graph limits as presented in [13]: The convex hull of $t(F, W)$ for all symmetric measurable function $W : [0, 1]^2 \rightarrow [0, 1]$ is $P_{F; \infty}$.

4 Conjectures about the limit object $P_{F; \infty}$.

We have previously inscribed cyclic polytopes in the limit objects. We will now define another cyclic polytope and conjecture that a particular class of limit objects actually are cyclic polytopes.

As described in Proposition 3.1 and the following discussion, the vertices of $P(K_2, K_n; \infty)$ are given by the limits of complete $k$-equipartite graphs. It is not hard to see that $P(K_2, K_n; \infty)$ is a cyclic polytope, and we believe that this is true in a more general setting.

For positive integers $e_1 < e_2 < \ldots < e_m$, define the tail $s^e : [0, 1] \rightarrow [0, 1]^m$ by

$$s^e_i(x) = \prod_{j=1}^{e_i} (1 - jx).$$

**Proposition 4.1** The convex hull of any finite set of points on a tail is a cyclic polytope.

**Conjecture 4.2** Let $e_1 < e_2 < \ldots < e_m$ be positive integers and $s^e$ their tail. The convex hull of 1 and \{s^e(1/k) | k = 1, 2, 3, \ldots\} is $P(K_{e_1}, K_{e_2}, \ldots, K_{e_m}; \infty)$.

The conjectured vertex description of $P(K_{e_1}, K_{e_2}, \ldots, K_{e_m}; \infty)$ also gives a facet description since it’s essentially a cyclic polytope. If we would chop off the vertex 1 from the convex hull described in Conjecture 4.2 with an hyperplane, then the remaining convex set would be an ordinary polytope. We believe this is true in the following general form.
**Conjecture 4.3** For any vector $F$ of graphs there is a positive integer $m$, such that for any $\varepsilon > 0$, the limit object $P_{F;\infty}$ can be chopped down to a polytope with a finite number of vertices, by using $m$ hyperplans to remove at most a volume $\varepsilon$.

**References**


