Approximating the stability number of a graph via copositive programming

Approximating forms over the simplex

Pólya's theorem

Monique Laurent

CWI, Amsterdam

Computing the stability number $\alpha(G)$ of a graph $G = (V, E)$ via copositive programming

$$\alpha(G) \leq \vartheta(G) = \begin{array}{ll} \max & \langle J, X\rangle \\ \text{s.t.} & \langle I, X\rangle = 1 \\ & X_{ij} = 0 \ (ij \in E) \quad \text{(SDP)} \\ & X \succeq 0 \end{array}$$

Pf: S stable set $\longrightarrow X := \frac{1}{|S|}\chi^S(\chi^S)^T$ feasible for (SDP)

Hierarchy of conic conditions for matrix var. X:

- $X \succeq 0 \longrightarrow \vartheta(G)$
- $X \in \Re_+(uu^T \mid u \in \{0,1\}^n) \longrightarrow \alpha(G)$
- $X \in \Re_+(uu^T \mid u \in \Re^n_+)$ (X *completely positive*) $\longrightarrow$?

$$\alpha(G) \leq \begin{array}{ll} \max & \langle J, X\rangle \\ \text{s.t.} & \langle I, X\rangle = 1 \\ & \langle A, X\rangle = 0 \\ & X \in \mathcal{C}_n^* \end{array} \leq \begin{array}{ll} \min & \lambda \\ \text{s.t.} & \lambda I + yA - J \in \mathcal{C}_n \\ & \lambda, y \in \Re \end{array} \overset{?}{\leq} \alpha(G)$$

Motzkin-Straus' theorem [1965]:

$$\frac{1}{\alpha(G)} = \begin{array}{ll} \min & x^T(I + A)x \\ \text{s.t.} & x \in \Delta := \{x \in \Re^n_+ \mid \Sigma_{i=1}^n x_i = 1\} \end{array}$$

Equivalently: The matrix $\alpha(I + A) - J$ is copositive

Proof of Motzkin-Straus' theorem:

$$\boxed{\mu := \min_{x\in\Delta} x^T(I+A)x \stackrel{?}{=} \frac{1}{\alpha(G)}}$$

$$\boxed{\mu \stackrel{?}{\leq} \frac{1}{\alpha(G)}}$$

Let S be a stable set of size $|S| = \alpha(G)$ and $x := \frac{1}{|S|}\chi^S$. Then, $x \in \Delta$ and $x^T(I+A)x = \sum_{i\in S} x_i^2 = \dfrac{|S|}{|S|^2} = \dfrac{1}{\alpha(G)}$.

$$\boxed{\mu \stackrel{?}{\geq} \frac{1}{\alpha(G)}}$$

Let x^* be a minimizer of $x^T(I+A)x$ over Δ with support $S := \{i \mid x_i^* > 0\}$.

• If S is stable, then we are done. Indeed,

$$\mu = (x^*)^T(I+A)x^* = \sum_{i\in S}(x_i^*)^2 \geq \frac{1}{|S|}(\sum_i x_i^*)^2 \geq \frac{1}{\alpha(G)}$$

• Assume that S is not stable and let $i, j \in S$ with $ij \in E$. Set $c_0 := 1 - \sum_{k\neq i,j} x_k^*$. Fix $x_k = x_k^*$ for all $k \neq i, j$ and let x_i, x_j be variables. Then,

$$\mu = \min_{\substack{x_i,x_j\geq 0\\ x_i+x_j=1-c_0}} (x_i+x_j)^2 + c_1x_i + c_2x_j + c_3$$

has a minimizer with $x_i = 0$ or $x_j = 0$ (c_1, c_2, c_3 being constants).

Iteratively, this shows existence of a minimizer whose support is stable.

Therefore:

$$\boxed{\begin{array}{rllcll} \alpha(G) = & \max & \langle J, X\rangle & = & \min & \lambda \\ & \text{s.t.} & \langle I+A, X\rangle = 1 & & \text{s.t.} & \lambda(I+A) - J \in \mathcal{C}_n \\ & & X \in \mathcal{C}_n^* & & & \end{array}}$$

• Optimization over the cones $\mathcal{C}_n$ and $\mathcal{C}_n^*$ is hard!

• Testing whether a matrix is not copositive is an NP-complete problem [Murty Kabadi 1987]

An example of test for copositivity: [Keller]
M is copositive if and only if, for every principal submatrix A of M, if all cofactors of the last row of A are nonnegative then the determinant of A is nonnegative

• No analogue test for complete positivity ?

$\mathcal{N}_n$:= all symmetric nonnegative $n \times n$ matrices

Obviously:

$$\mathrm{PSD}_n + \mathcal{N}_n \subseteq \mathcal{C}_n, \quad \mathcal{C}_n^* \subseteq \mathrm{PSD}_n \cap \mathcal{N}_n$$

Theorem [Diananda 1962]:
Equality holds if and only if $n \leq 4$

Counterexample for $n = 5$:

$$M := \begin{pmatrix} 1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 \end{pmatrix} \in \mathcal{C}_5 \setminus (\mathrm{PSD}_5 + \mathcal{N}_5)$$

- M is copositive since

$$\begin{aligned} x^T M x &= (x_1 - x_2 + x_3 - x_4 + x_5)^2 + 4x_2x_5 + 4x_1(x_4 - x_5) \\ &= (x_1 - x_2 + x_3 + x_4 - x_5)^2 + 4x_2x_4 + 4x_3(x_5 - x_4) \end{aligned}$$

- Assume that $M = P + N$, where $P \succeq 0$ and $N \geq 0$.

Then, $P_{ij} \leq M_{ij}$ for all i, j. Therefore,

$$0 \leq P_{ii} \leq 1 \quad \forall i$$

$$P_{ij} \leq -1 \ \forall \text{ circuit edges}$$

$$|P_{ij}| \leq \sqrt{P_{ii}}\sqrt{P_{jj}} \leq 1 \quad \forall i, j$$

$$\Longrightarrow P_{ij} = -1 \text{ for all circuits edges}, \quad P_{ii} = 1 \quad \forall i$$

$\longrightarrow$ contradicting the fact that $P \succeq 0$

Parrilo's idea: Approximate the copositive cone by a hierarchy of tractable cones defined via sums of squares

Recall: M is copositive if $x^T M x \geq 0$ for all $x \geq 0$; equivalently, if

$$\boxed{p_M(x) := \sum_{i,j=1}^{n} M_{ij} x_i^2 x_j^2 \geq 0 \quad \text{for all } x \in \Re^n}$$

For an integer $r \geq 0$, define the polynomial:

$$\boxed{p_M^r(x) := \left(\sum_{i=1}^{n} x_i^2\right)^r p_M(x) = \left(\sum_{i=1}^{n} x_i^2\right)^r \left(\sum_{i,j=1}^{n} M_{ij} x_i^2 x_j^2\right)}$$

and the cones:

$$\boxed{\mathcal{L}_n^r := \{M \mid p_M^r(x) \text{ has only nonnegative coefficients}\}}$$

$$\boxed{\mathcal{K}_n^r := \{M \mid p_M^r(x) \text{ is a sum of squares of polynomials}\}}$$

Obviously:

$$\mathcal{L}_n^r \subseteq \mathcal{K}_n^r \subseteq \mathcal{C}_n$$

$$\mathcal{L}_n^0 \subseteq \ldots \subseteq \mathcal{L}_n^r \subseteq \mathcal{L}_n^{r+1} \subseteq \ldots \subseteq \mathcal{C}_n$$

$$\mathcal{K}_n^0 \subseteq \ldots \subseteq \mathcal{K}_n^r \subseteq \mathcal{K}_n^{r+1} \subseteq \ldots \subseteq \mathcal{C}_n$$

Pólya's theorem shows:

$$\boxed{\mathcal{L}_n^r \nearrow \mathcal{C}_n}$$

Pólya's theorem

Definition: A matrix M is *strictly copositive* if

$$x^T M x > 0 \ \text{ for all } x \in \Re^n_+ \setminus \{0\}$$

Equivalently, if

$$p_{\min} := \min_{x \in \Delta} x^T M x > 0.$$

Define

$$L_M := \max_{i,j=1,..,n} M_{ij}.$$

Theorem [Pólya 1928, Powers-Reznick 2000]:

$$\boxed{M \textit{ strictly copositive} \implies M \in \mathcal{L}_n^r \text{ for all } r \geq \tfrac{L_M}{p_{min}} - 2}$$

General formulation for forms of degree d

Theorem: Let $p(x) = \sum_{|\alpha|=d} p_\alpha x^\alpha$ be a form of degree d which is *positive*, i.e.,

$$p(x) > 0 \ \text{ for all } x \in \Re^n \setminus \{0\};$$

equivalently,

$$p_{\min} := \min_{x \in \Delta} p(x) > 0.$$

Set

$$L_p := \max_{\alpha} \frac{\alpha!}{d!} |p_\alpha|.$$

Then, $\left(\sum_{i=1}^n x_i\right)^r p(x)$ has only nonnegative coefficients for all

$$r \geq \binom{d}{2} \frac{L_p}{p_{\min}} - d$$

Application to the stability number (de Klerk-Pasechnik 2002)

Define the polyhedral bound:

$$\boxed{\lambda^r(G) := \min \lambda \ \text{ s.t. } \ \lambda(I + A) - J \in \mathcal{L}_n^r}$$

and the semidefinite bound:

$$\boxed{\vartheta^r(G) := \min \lambda \ \text{ s.t. } \ \lambda(I + A) - J \in \mathcal{K}_n^r}$$

$$\boxed{\alpha(G) \leq \vartheta^r(G) \leq \lambda^r(G)}$$

For any fixed r, one can compute in polynomial time the bounds $\lambda^r(G)$ and $\vartheta^r(G)$ (with arbitrary precision)

Theorem:

$$\boxed{\alpha(G) \leq \lambda^r(G) < \alpha(G) + 1 \ \text{ for } r \geq \alpha(G)^2 - 1}$$

Proof: Define $M := \alpha(1 + \epsilon)(I + A) - J$, where

$$\epsilon := \frac{1}{\alpha + 1/(\alpha - 1)} = \frac{\alpha - 1}{\alpha^2 - \alpha + 1}.$$

$$\left.\begin{array}{l} p_{\min} = \epsilon \quad \text{(by Motzkin-Straus)} \\ L_M = \alpha(1+\epsilon) - 1 = \frac{\alpha^3 - \alpha^2 + \alpha - 1}{\alpha^2 - \alpha + 1} \end{array}\right\} \Longrightarrow \frac{L_M}{p_{\min}} - 2 = \alpha^2 - 1.$$

Hence, by Powers-Reznick, for $r \geq \alpha^2 - 1$,

$$M \in \mathcal{L}_n^r \Longrightarrow \lambda^r(G) \leq \alpha(1 + \epsilon) = \alpha \frac{\alpha^2}{\alpha^2 - \alpha + 1} < \alpha + 1.$$

Do the semidefinite bounds $\vartheta^r(G)$ converge faster to $\alpha(G)$ than the linear bounds $\lambda^r(G)$?

Conjecture [de Klerk-Pasechnik 2002]:

$$\boxed{\vartheta^r(G) = \alpha(G) \text{ for } r \geq \alpha(G) - 1}$$

Remarks:

(i) The conjecture holds for $\alpha(G) = 2$ [dKP 2002], also for $\alpha(G) = 3$

(ii) The other SDP hierrachies (LS, Lasserre, ..) do converge in $\alpha(G)$ steps

(iii) $r = \alpha^2(G) - 1$ iterations are needed for achieving $\lfloor \lambda^r(G) \rfloor = \alpha(G)$ [Bomze - de Klerk 2002]

How does the cone $\mathcal{L}_n^r$ look like? (Bomze - de Klerk 2002):

$$M \in \mathcal{L}_n^r \overset{\text{def}}{\Longleftrightarrow} p_M^r(x) = \left(\sum_i x_i^2\right)^r \left(\sum_{i,j} M_{ij} x_i^2 x_j^2\right)$$
has nonnegative coefficients

$$\Longleftrightarrow \beta^T M \beta - \beta^T \mathrm{diag}(M) \geq 0 \text{ for all } \beta \in Z_+^n$$
with $|\beta| := \sum_{i=1}^n \beta_i = r + 2$

Use the multinomial theorem:

$$\left(\sum_{i=1}^n y_i\right)^r = \sum_{|\alpha|=r} \frac{r!}{\alpha_1! \cdots \alpha_n!} y_1^{\alpha_1} \cdots y_n^{\alpha_n!} =: \sum_{|\alpha|=r} \frac{r!}{\alpha!} y^\alpha$$

for computing the polynomial $p_M^r(x)$:

$$\boxed{p_M^r(x) = \sum_{|\beta|=r+2} \frac{r!}{\beta!} (\beta^T M \beta - \beta^T \mathrm{diag}(M))\ x^{2\beta}}$$

Proof:

$$p_M^r(x) = \left(\sum_{|\alpha|=r} \frac{r!}{\alpha!} x^{2\alpha}\right)\left(\sum_{i,j} M_{ij} x_i^2 x_j^2\right) = \sum_{\substack{|\alpha|=r \\ i,j=1,\dots,n}} \frac{r!}{\alpha!} M_{ij} x^{2\alpha+2e_i+2e_j}$$

$$= \sum_{|\beta|=r+2} \left(\sum_{i,j|\beta\geq e_i+e_j} \frac{r!}{(\beta-e_i-e_j)!} M_{ij}\right) x^{2\beta} =: \sum_{|\beta|=r+2} p_M^r(2\beta) x^{2\beta}$$

where

$$p_M^r(2\beta) = \sum_{i|\beta_i\geq 2} \frac{r!\beta_i(\beta_i-1)}{\beta!} M_{ii} + \sum_{i\neq j|\beta_i,\beta_j\geq 1} \frac{r!\beta_i\beta_j}{\beta!} M_{ij}$$

$$= \frac{r!}{\beta!}\left(\sum_i \beta_i(\beta_i-1)M_{ii} + \sum_{i\neq j} \beta_i\beta_j M_{ij}\right)$$

$$= \frac{r!}{\beta!}(\beta^T M\beta - \beta^T \mathrm{diag}(M))$$

An elementary proof of the Powers-Reznick bound for Pólya's theorem for quadratic forms

We prove:

$$\boxed{r \geq \frac{\max_i M_{ii}}{p_{\min}} - 2 \Longrightarrow \quad \begin{array}{c} \beta^T M \beta - \beta^T \mathrm{diag}(M) \geq 0 \\ \text{for all } \beta \in N^n \text{ with } \Sigma_i \beta_i = r+2. \end{array}}$$

Proof:

Let $\beta \in N^n$ with $\Sigma_i \beta_i = r+2$. Then,

$$\frac{\beta}{r+2} \in \Delta \Longrightarrow \beta^T M \beta \geq (r+2)^2 p_{\min}$$

On the other hand,

$$\begin{aligned} \beta^T \mathrm{diag}(M) = \Sigma_i \beta_i M_{ii} \ &\leq (r+2) \max_i M_{ii} \\ &= (r+2)^2 \ \tfrac{\max_i M_{ii}}{r+2} \\ &\leq (r+2)^2 \ p_{\min} \end{aligned}$$

This shows:

$$\beta^T M \beta - \beta^T \mathrm{diag}(M) \geq 0$$

How does the cone $\mathcal{K}_n^r$ look like?

$$M \in \mathcal{K}_n^r \overset{\text{def}}{\Longleftrightarrow} p_M^r(x) = \left(\sum_i x_i^2\right)^r \left(\sum_{i,j} M_{ij} x_i^2 x_j^2\right) \text{ SOS}$$

$$\Longleftrightarrow \exists\, X = (X_{\beta,\beta'})_{|\beta|,|\beta'|=r+2} \succeq 0 \text{ satisfying:}$$

$$\sum_{\beta,\beta' | \beta+\beta'=\gamma} X_{\beta,\beta'} = p_M^r(\gamma) \ (|\gamma| = 2r+4)$$

Recall: $p_M^r(x) = \sum_{|\gamma|=2r+4} p_M^r(\gamma) x^\gamma$, where

$$p_M^r(\gamma) = \begin{cases} 0 & \text{if some } \gamma_i \text{ is odd} \\ \frac{r!}{\beta!}(\beta^T M \beta - \beta^T \text{diag}(M)) & \text{if } \gamma = 2\beta \end{cases}$$

Theorem [Parrilo 2000]: $\boxed{\mathcal{K}_n^0 = \text{PSD}_n + \mathcal{N}_n}$

Proof: Let $M \in \mathcal{K}_n^0$ and $X \succeq 0$ satisfying (S).

$$X = \begin{array}{c c} & \begin{array}{cc} 2e_1 \ldots 2e_n & e_1+e_2 \ldots e_{n-1}+e_n \end{array} \\ \begin{array}{c} 2e_1 \\ \vdots \\ 2e_n \\ e_1+e_2 \\ \vdots \\ e_{n-1}+e_n \end{array} & \left(\begin{array}{cc} & \phantom{e_1+e_2 \ldots e_{n-1}+e_n} \end{array} \right) \end{array}$$

Define matrix P by $P_{ij} := X_{2e_i,2e_j}$ and $N = M - P$. Then, $P \succeq 0$ and $N \geq 0$. Indeed, $N_{ij} = M_{ij} - X_{2e_i,2e_j} \geq 0$ since, by (S),

$$X_{2e_i,2e_j} = \tfrac{1}{2}(2M_{ii}) \ \ (\gamma = 4e_i)$$
$$X_{e_i+e_j,e_i+e_j} + 2X_{2e_i,2e_j} = 2M_{ij} \ \ (\gamma = 2e_i + 2e_j)$$

Another characterization of $\mathcal{K}_n^r$ (Zualaga-Vera-Péna 2003)

Theorem: Let $q(x) = \sum_{|\alpha|=m} q_\alpha x^\alpha$ be a form of degree m. Then, $q(x_1^2, \ldots, x_n^2)$ is a sum of squares if and only if

$$q(x) = \sum_{\substack{I \subseteq \{1,\ldots,n\} \\ |I| \equiv m}} x^I q_I$$

where q_I is a form of degree $(m - |I|)/2$ which is a sum of squares, and $x^I := \prod_{i \in I} x_i$.

Corollary: $M \in \mathcal{K}_n^r$, i.e., $p_M^r(x)$ is a sum of squares if and only if

$$(\sum_i x_i)^r (\sum_{i,j} M_{ij} x_i x_j) = \sum_{I \subseteq \{1,\ldots,n\} \mid |I| \equiv r+2} x^I q_I$$

where q_I is a form of degree $(r + 2 - |I|)/2$ which is a sum of squares

Proof: Assume that $q(x_1^2, \ldots, x_n^2)$ is SOS. Then,

$$q(x_1^2, \ldots, x_n^2) = \sum_\ell \phi_\ell(x)^2 = \sum_\ell \left(\sum_{|\alpha|=m} \phi_{\ell,\alpha} x^\alpha \right)^2$$
$$= \sum_\ell \sum_{|\alpha|, |\alpha'|=m} \phi_{\ell,\alpha} \phi_{\ell,\alpha'} x^\alpha x^{\alpha'}$$

Define the *odd support* of $\alpha \in Z_+^n$ as $O(\alpha) := \{i \mid \alpha_i \text{ is odd}\}$. We can restrict the summation to pairs (α, α') having the same odd support. Therefore,

$$
\begin{aligned}
q(x_1^2,\ldots,x_n^2) &= \sum_\ell \sum_{I\subseteq\{1,\ldots,n\}} \sum_{\alpha,\alpha'|O(\alpha)=O(\alpha')=I} \phi_{\ell,\alpha}\phi_{\ell,\alpha'}x^\alpha x^{\alpha'} \\
&= \sum_\ell \sum_I \left(\sum_{\alpha|O(\alpha)=I} \phi_{\ell,\alpha}x^\alpha\right)^2 \\
&= \sum_\ell \sum_{I||I|\equiv m} \left(\sum_{\alpha|O(\alpha)=I} \phi_{\ell,\alpha}x^{\alpha-\chi^I}\right)^2 (x^I)^2 \\
&= \sum_\ell \sum_{I||I|\equiv m} \left(\sum_{\alpha|O(\alpha)=I} \phi_{\ell,\alpha}(x^{(\alpha-\chi^I)/2})^2\right)^2 (x^I)^2.
\end{aligned}
$$

Therefore,

$$
q(x) = \sum_\ell \sum_{I||I|\equiv m} x^I \left(\sum_{\alpha|O(\alpha)=I} \phi_{\ell,\alpha}x^{(\alpha-\chi^I)/2}\right).
$$

$$\boxed{m=2}$$

$$
q(x_1^2,\ldots,x_n^2) \text{ SOS } \iff q(x) = q_0 + \sum_{1\le i<j\le n} x_ix_jq_{ij}
$$

where q_0 is a form of degree 2 which is SOS, $q_{ij}\in\Re_+$

$$
\iff q(x) = x^T(P+N)x \quad \text{where } P\succeq 0,\ N\ge 0
$$

$$\boxed{m=3}$$

$$
q(x_1^2,\ldots,x_n^2) \text{ SOS } \iff q(x) = \sum_{i=1}^n x_iq_i + \sum_{1\le i<j<k\le n} x_ix_jx_kq_{ijk}
$$

where q_i is a form of degree 2 which is SOS and $q_{ijk}\in\Re_+$

$$
\iff q(x) = \sum_{i=1}^n x_ix^T(P^i+N^i)x \quad \text{where } P^i\succeq 0,\ N^i\ge 0
$$

Theorem [Parrilo, de Klerk and Bomze]:

$$M \in \mathcal{K}_n^1 \Longleftrightarrow \qquad \exists \Lambda_1, \ldots, \Lambda_n \text{ s.t.}$$

$$\begin{cases} M - \Lambda_i \succeq 0 & (\forall i) \\ \Lambda_{ii}^i = 0 & (\forall i) \\ \Lambda_{jj}^i + \Lambda_{ij}^i + \Lambda_{ij}^j = 0 & (\forall i \neq j) \\ \Lambda_{jk}^i + \Lambda_{ik}^j + \Lambda_{ij}^k \geq 0 & (\forall i \neq j \neq k) \end{cases}$$

Dual formulation for the bound $\vartheta^r(G)$

$$\begin{aligned}
\vartheta^r(G) &= \min \lambda \text{ s.t. } \lambda(I+A) - J \in \mathcal{K}_n^r \\
&\overset{(1)}{=} \min \lambda \text{ s.t. } \langle \lambda(I+A) - J, C\rangle \geq 0 \ \ \forall C \in (\mathcal{K}_n^r)^* \\
&= \min \lambda \text{ s.t. } \lambda\langle I+A, C\rangle \geq \langle J, C\rangle \ \ \forall C \in (\mathcal{K}_n^r)^* \\
&\overset{(2)}{=} \max \langle J, C\rangle \text{ s.t. } \langle I+A, C\rangle = 1,\ C \in (\mathcal{K}_n^r)^*
\end{aligned}$$

Equality (1) holds since $\mathcal{K}_n^{(r)} = ((\mathcal{K}_n^{(r)})^*)^*$, as $\mathcal{K}_n^r$ is closed.

Equality (2) holds since $\langle I+A, C\rangle \geq 0$, as any $C \in (\mathcal{K}_n^r)^*$ is nonnegative and psd.

How does the dual cone of $\mathcal{K}_n^r$ look like?

The dual cone of $\mathcal{K}_n^r$

Definition: Given $y = (y_\gamma)_{|\gamma|=2r+4}$, define its *moment matrix*

$$M_{r+2}(y) := \sum_\gamma y_\gamma A_\gamma = (y_{\beta+\beta'})_{|\beta|,|\beta'|=r+2}$$

and its $n \times n$ *'compression' matrix*

$$C(y) := \left(\sum_{|\alpha|=r} \frac{r!}{\alpha!} y_{2\alpha+2e_i+2e_j} \right)_{i,j=1}^n$$

Define the cones:

$$\boxed{\mathcal{M}_{r+2} := \{y = (y_\gamma)_{|\gamma|=2r+4} \mid M_{r+2}(y) \succeq 0\}}$$

$$\boxed{\mathcal{C}_n^r := \{C(y) \mid y \in \mathcal{M}_{r+2}\}}$$

Recall:

$$\boxed{(\mathcal{M}_{r+2})^* = \text{cone of forms of degree } 2r+4 \text{ that are SOS}}$$

Theorem: $\boxed{\mathcal{C}_n^r = (\mathcal{K}_n^r)^*}$

To see it, recall: For a symmetric $n \times n$ matrix M,

$$p_M^r(x) = \left(\sum_i x_i^2\right)^r \left(\sum_{i,j} M_{ij} x_i^2 x_j^2\right) =: \sum_{|\gamma|=2r+4} p_M^r(\gamma) x^\gamma$$

$$\sum_{i,j} M_{ij} \sum_{|\alpha|=r} \frac{r!}{\alpha!} x^{2\alpha+2e_i+2e_j}$$

$$\boxed{\mathcal{X}_M := \{X \mid \langle X, A_\gamma \rangle = p_M^r(\gamma) \ (|\gamma| = 2r+4)\}}$$

$$\boxed{M \in \mathcal{K}_n^r \text{ (i.e., } p_M^r(x) \text{ SOS)} \iff \mathcal{X}_M \cap \mathrm{PSD}_n \neq \emptyset}$$

Lemma: Given $y = (y_\gamma)_{|\gamma|=2r+4}$ and $X \in \mathcal{X}_M$,

$$\boxed{\langle M, C(y) \rangle = \langle X, M_{r+2}(y) \rangle = y^T p_M^r}$$

Proof:

$$\begin{aligned}\langle X, M_{r+2}(y) \rangle &= \langle X, \sum_\gamma y_\gamma A_\gamma \rangle = \sum_\gamma y_\gamma \langle X, A_\gamma \rangle \\ &= \sum_\gamma y_\gamma p_M^r(\gamma) = y^T p_M^r \\ &= \sum_{i,j} M_{ij} \sum_{|\alpha|=r} \frac{r!}{\alpha!} y_{2\alpha+2e_i+2e_j} = \langle M, C(y) \rangle\end{aligned}$$

Corollary: $(\mathcal{C}_n^r)^* = \mathcal{K}_n^r$, $(\mathcal{C}_n^r)^* = \mathcal{K}_n^r$

Proof: Enough to show: $(\mathcal{C}_n^r)^* = \mathcal{K}_n^r$, as $\mathcal{C}_n^r$ is closed

$$\begin{aligned}M \in (\mathcal{C}_n^r)^* &\iff \langle M, C(y) \rangle \geq 0 \ \forall C \in \mathcal{C}_n^r \\ &\iff y^T p_M^r \geq 0 \ \forall y \in \mathcal{M}_{r+2} \\ &\iff p_M^r \text{ SOS, i.e, } M \in \mathcal{K}_n^r\end{aligned}$$

How does the cone $\mathcal{C}_n^r$ look like?

$$y = (y_\gamma)_{|\gamma|=2r+4} \rightsquigarrow M_{r+2}(y)$$

For each $|\alpha| = r$, define $M^\alpha(y)$:= the $n \times n$ principal submatrix of $M_{r+2}(y)$ indexed by $\alpha + 2e_1, \ldots, \alpha + 2e_n$

$$\boxed{C(y) = \sum_{|\alpha|=r} \frac{r!}{\alpha!} M^\alpha(y)}$$

Case: $r = 0$

$$\begin{array}{c} \\ 2e_1 \\ \vdots \\ 2e_n \\ e_1+e_2 \\ \vdots \\ e_{n-1}+e_n \end{array} \begin{array}{c} 2e_1 \quad \ldots 2e_n \quad e_1+e_2 \ldots e_{n-1}+e_n \\ \left(\begin{array}{ccc} & & \\ & & \\ & & \\ & & \\ & & \\ & & \end{array} \right) \end{array} = M_{r+2}(y)$$

$$\rightsquigarrow C(y) \succeq 0, \;\; C(y) \geq 0$$

New semidefinite relaxations for the stable set polytope

$$\begin{array}{llcllcll}
\max & \langle J, X\rangle & \overset{(2)}{\leq} & \max & \langle J, X\rangle & \overset{(1)}{=} & \min & \lambda \\
\text{s.t.} & \langle I, X\rangle = 1 & & \text{s.t.} & \langle I+A, X\rangle = 1 & & \text{s.t.} & \lambda(I+A) - J \in \mathcal{K}_n^r \\
& \langle A, X\rangle = 0 & & & & & & \| \\
& X \in \mathcal{C}_n^r & & & X \in (\mathcal{K}_n^r)^* & & & \vartheta^r(G)
\end{array}$$

$$\begin{array}{lcllcll}
\alpha(G) & \overset{(4)}{\leq} & \max & \langle I, Z\rangle & \overset{(3)}{\leq} & \max & \langle J, X\rangle \\
& & \text{s.t.} & Z - \operatorname{diag}(Z)\operatorname{diag}(Z)^T \succeq 0 & & \text{s.t.} & \langle I, X\rangle = 1 \\
& & & \langle A, Z\rangle = 0 & & & \langle A, X\rangle = 0 \\
& & & Z \in \mathcal{C}_n^r & & & X \in \mathcal{C}_n^r
\end{array}$$

(1) Shown before

(2) $(\mathcal{K}_n^r)^* = \mathcal{C}_n^r$ and disaggregate the constraint $\langle I + A, X\rangle = 1$

(3) Z feasible for the left program $\Longrightarrow X := \frac{1}{\langle I,Z\rangle} Z$ feasible for the right program with value $\langle J, X\rangle \geq \langle I, Z\rangle$, since

$$Z - \operatorname{diag}(Z)\operatorname{diag}(Z)^T \succeq 0 \Longrightarrow \langle J, Z\rangle \geq \langle J, \operatorname{diag}(Z)\operatorname{diag}(Z)^T\rangle = \langle I, Z\rangle^2$$

(4) follows from the following:

Proposition: The set

$$\begin{aligned} \mathrm{TH}^r(G) := \{x \in \Re^n \mid\ & x = \operatorname{diag}(Z) \text{ for some } Z \in \mathcal{C}_n^r \text{ s.t.} \\ & Z - \operatorname{diag}(Z)\operatorname{diag}(Z)^T \succeq 0,\ \langle Z, A\rangle = 0\} \end{aligned}$$

is a semidefinite relaxation of the stable set polytope.

Proof:

Remarks - Open questions about $\vartheta^r(G)$, $\mathrm{TH}^r(G)$

1) For $r = 0$

$$\mathcal{K}_n^0 = \mathrm{PSD}_n + \mathcal{N}_n, \quad \mathcal{C}_n^0 = (\mathcal{K}_n^0)^* = \mathrm{PSD}_n \cap \mathcal{N}_n$$

$$\begin{array}{rll} \vartheta^0(G) = & \max & \langle J, X\rangle \qquad =: \vartheta'(G) \\ & \text{s.t.} & \langle A, X\rangle = 0 \\ & & \langle I, X\rangle = 1 \\ & & X \in \mathcal{C}_n^0 = \mathrm{PSD}_n \cap \mathcal{N}_n \end{array}$$

$\vartheta'(G)$ is the strengthening of $\vartheta(G)$ from [Schrijver 79]

2) What is the smallest r for which equality: $\mathrm{TH}^r(G) = \mathrm{STAB}(G)$ hold? Does equality hold for $r = \alpha(G) - 1$?

3) How does the semidefinite relaxation $\mathrm{TH}^r(G)$ compare with the Lasserre-based semidefinite relaxations ?

4) Weak points of the cones $\mathcal{K}_n^r$ and $\mathcal{C}_n^r$: not closed under adding zero-row, under positive diagonal scaling, ...

$$M := \begin{pmatrix} 1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 \end{pmatrix} \in \mathcal{K}_5^1, \ \begin{pmatrix} 0 & 0 \\ 0 & M \end{pmatrix} \notin \mathcal{K}_6^1$$

$$M \in \mathcal{K}_n^1 \not\Longrightarrow \mathrm{diag}(d) M \mathrm{diag}(d) \in \mathcal{K}_n^r$$

$$(\sum_i x_i^2)^r (\sum_{i,j} M_{ij} x_i^2 x_j^2) \text{ SOS } \not\Longrightarrow (\sum_i x_i^2)^r (\sum_{i,j} M_{ij} d_i d_j x_i^2 x_j^2) \text{ SOS}$$

Minimization of forms on the simplex Δ and Pólya's theorem

For a form $p(x)$ of degree d, consider the problem:

$$\boxed{p_{\min} := \min_{x\in\Delta} p(x)}$$

and, for $r \geq 0$, the lower bound:

$$\boxed{\begin{array}{l} p^r := \max\ \lambda\ \text{ s.t. } \ (\sum_{i=1}^n x_i)^r \left(p(x) - \lambda \left(\sum_{i=1}^n x_i\right)^d\right) \\ \qquad\qquad\qquad\qquad\qquad \text{has nonnegative coefficients} \end{array}}$$

$$\boxed{p^r \leq p_{\min}}$$

Pólya's theorem: Assume p is *strictly positive*:

$$p(x) > 0 \ \text{ for all } x \in \Re^n \setminus \{0\};$$

equivalently, $p_{\min} > 0$. Then,

$$p^r \geq 0 \ \text{ for some } r \geq 0;$$

equivalently, $(\sum_{i=1}^n x_i)^r p(x)$ has nonnegative coefficients for some $r \geq 0$.

Corollary: $p^r \nearrow p_{\min}$ as $r \to \infty$

Our objective:

- prove Pólya's theorem and Powers-Reznick bound on r
- give estimates on $p^r - p_{\min}$ by Faybusovich [2003]
- PTAS for the minimization on the simplex of forms of degree 2 [Bomze-de Klerk 2002] and degree 3 [de Klerk-L-Parrilo]

A PTAS for the minimization of degree 2 forms on the simplex (Bomze - de Klerk 2002)

$$p(x) := x^T M x$$

Definition: For $r \geq 0$, define

$$\Delta_{r+2} := \{x \in \Delta \mid (r+2)x \text{ integral}\},$$

the set of grid points in Δ with denominator $r+2$. Set

$$p_{\min,r+2} := \min_{x \in \Delta_{r+2}} p(x), \quad p_{\max} := \max_{x \in \Delta} p(x)$$

$$\boxed{p^r \leq p_{\min} \leq p_{\min,r+2} \leq p_{\max}}$$

Theorem: $\boxed{p_{\min,r+2} - p_{\min} \leq \frac{1}{r+2}(p_{\max} - p_{\min})}$

Claim 1: $\boxed{p^r = \min_{x \in \Delta_{r+2}} \frac{r+2}{r+1}\left(x^T M x - \frac{1}{r+2} x^T \mathrm{diag}(M)\right).}$

Proof: $(\sum_i x_i)^r (p(x) - \lambda(\sum_i x_i)^2) = (\sum_i x_i)^r \; x^T(M - \lambda J)x$

$$= \sum_{|\beta|=r+2} \frac{r!}{\beta!} x^\beta (\beta^T (M - \lambda J)\beta - \beta^T \mathrm{diag}(M - \lambda J))$$

$$= \sum_{|\beta|=r+2} \frac{r!}{\beta!} x^\beta \left(\beta^T M \beta - \beta^T \mathrm{diag}(M) - \lambda(r+1)(r+2)\right)$$

$$\begin{aligned} \Longrightarrow \; p^r &= \min_{|\beta|=r+2} \frac{1}{(r+1)(r+2)} (\beta^T M \beta - \beta^T \mathrm{diag}(M)) \\ &= \min_{x \in \Delta_{r+2}} \frac{r+2}{r+1}\left(x^T M x - \frac{1}{r+2} x^T \mathrm{diag}(M)\right) \end{aligned}$$

as $x := \frac{1}{r+2}\beta \in \Delta_{r+2}$.

Corollary 2: $\boxed{p^r = \frac{r+2}{r+1} \min_{x \in \Delta_{r+2}} (p(x) - \phi(x))}$

after setting:

$$\boxed{\phi(x) := \frac{1}{r+2}\, x^T \mathrm{diag}(M), \quad \phi_{\max} := \max_{x \in \Delta} \phi(x)}$$

Corollary 3: $\boxed{p_{\min} \geq \frac{r+2}{r+1}(p_{\min,r+2} - \phi_{\max})}$

Claim 2. $\phi_{\max} \leq \frac{1}{r+2} p_{\max}$

Proof: For $x \in \Delta$,

$$\phi(x) = \frac{1}{r+2} \sum_i M_{ii} x_i \leq \frac{1}{r+2} p_{\max}$$

since, for all i,

$$p(e_i) = e_i^T M e_i = M_{ii} \leq p_{\max}.$$

Back to the general case: $p(x) = \sum_{|\beta|=d} p_\beta x^\beta$ is a degree d form

Some notation:

$$\Delta_{r+d} := \{x \in \Delta \mid (r+d)x \text{ integral}\}$$

$$p_{\min,k} := \min_{x \in \Delta_k} p(x)$$

$$L_p := \max_{|\beta|=d} \frac{\beta!}{d!} |p_\beta|$$

$$w_r(d) := \prod_{i=1}^{d-1} \left(1 - \frac{i}{r+d}\right) = \frac{(r+d)!}{r!(r+d)^d}$$

Note:

$$1 - \binom{d}{2}\frac{1}{r+d} \leq w_r(d) \leq 1, \text{ implying } \lim_{r\to\infty} w_r(d) = 1$$

Obviously:

$$p^r \leq p_{\min} \leq p_{\min,r+d} \leq p_{\max}$$

We prove:

(1) $p^r \geq \frac{1}{w_r(d)}\left(p_{\min} - L_p\binom{d}{2}\frac{1}{r+d}\right)$

(2) Hence, $p^r \geq 0$ for $r > \binom{d}{2}\frac{L_p}{p_{\min}} - d$

(3) $p_{\min} - p^r \leq \left(\frac{1}{w_r(d)} - 1\right)(L_p - p_{\min})$

(4) $p_{\min,r+d} - p_{\min} \leq (1 - w_r(d))(L_p - p_{\min})$

(1), (2) by [Powers-Reznick 2000], (2) by [Faybusovich 2003]

Claim 1: $\boxed{p^r = \min_{x\in\Delta_{r+2}} \frac{1}{w_r(d)} \sum_{|\beta|=d} p_\beta \prod_{i=1}^{n} (x_i)^{\beta_i}_{\frac{1}{r+d}}}$

setting, for y, t reals, m integer:

$$\boxed{(y)_t^m := y(y-t)\cdots(y-(m-1)t) = \prod_{i=0}^{m-1}(y-it)}$$

Thus, $(y)_t^m = y^m$ for $t = 0$

Corollary: If $p_{\min} > 0$, then $p^r \geq 0$ for r large enough.

Proof: Compute the coefficients of the polynomial:

$$(\sum_{i=1}^{n} x_i)^r \left(p(x) - \left(\sum_{i=1}^{n} x_i\right)^d\right) =: \sum_{|\gamma|=r+d} A_\gamma x^\gamma$$

First:

$$(\sum_{i=1}^{n} x_i)^r p(x) = \left(\sum_{|\alpha|=r} \frac{r!}{\alpha!} x^\alpha\right) (\sum_{|\beta|=d} p_\beta x^\beta) = \sum_{|\gamma|=r+d} \left(\sum_{\substack{|\alpha|=r, |\beta|=d \\ \alpha+\beta=\gamma}} \frac{r!}{\alpha!} p_\beta\right) x^\gamma$$

Hence:

$$A_\gamma = \sum_{|\beta|=d, \beta\leq\gamma} \frac{r!}{(\gamma-\beta)!}\, p_\beta = \frac{r!(r+d)^d}{\gamma!} \sum_{|\beta|=d,\beta\leq\gamma} p_\beta\, \frac{\gamma!}{(\gamma-\beta)!(r+d)^d}$$

$$= \frac{r!(r+d)^d}{\gamma!} \sum_{|\beta|=d,\beta\leq\gamma} p_\beta \prod_{i=1}^{n} \frac{\gamma_i!}{(\gamma_i-\beta_i)!(r+d)^{\beta_i}} = \frac{r!(r+d)^d}{\gamma!} \sum_{|\beta|=d} p_\beta \prod_{i=1}^{n} \frac{\gamma_i(\gamma_i-1)\cdots(\gamma_i-\beta_i+1)}{(r+d)^{\beta_i}}$$

$$= \frac{r!(r+d)^d}{\gamma!} \sum_{|\beta|=d} p_\beta \prod_{i=1}^{n} \left(\frac{\gamma_i}{r+d}\right)^{\beta_i}_{1/(r+d)}$$

Second:

$$(\sum_i x_i)^{r+d} = \sum_{|\beta|=r+d} \frac{(r+d)!}{\beta!} x^\beta$$

Hence:

$$p^r = \max \lambda \text{ s.t. } A_\gamma \geq \lambda \tfrac{(r+d)!}{\gamma!} \text{ for all } |\gamma| = r+d$$

$$= \min_{|\gamma|=r+d} \tfrac{\gamma!}{(r+d)!} A_\gamma = \min_{|\gamma|=r+d} \tfrac{r!(r+d)^d}{(r+d)!} \sum_{|\beta|=d} p_\beta \prod_{i=1}^n \left(\tfrac{\gamma_i}{r+d}\right)^{\beta_i}_{\frac{1}{r+d}}$$

$$= \min_{x \in \Delta_{r+d}} \tfrac{1}{w_r(d)} \sum_{|\beta|=d} p_\beta \prod_{i=1}^n (x_i)^{\beta_i}_{\frac{1}{r+d}}$$

$$\boxed{\phi(x) := p(x) - \sum_{|\beta|=d} p_\beta \prod_{i=1}^{n} (x)^{\beta_i}_{\frac{1}{r+d}} = \sum_{|\beta|=d} p_\beta \left(x^\beta - \prod_{i=1}^{n} (x)^{\beta_i}_{\frac{1}{r+d}} \right)}$$

$$\boxed{p^r = \min_{x \in \Delta_{r+d}} \frac{1}{w_r(d)} (p(x) - \phi(x))}$$

Hence:

$$p^r \geq \frac{1}{w_r(d)} (p_{\min} - \phi_{\max})$$

$$p_{\min} \geq \frac{1}{w_r(d)} (p_{\min,r+d} - \phi_{\max})$$

implying:

$$p_{\min} - p^r \leq (1 - \frac{1}{w_r(d)}) p_{\min} + \frac{1}{w_r(d)} \phi_{\max}$$

$$p_{\min,r+d} - p_{\min} \leq w_r(d) \left((1 - \frac{1}{w_r(d)}) p_{\min} + \frac{1}{w_r(d)} \phi_{\max} \right)$$

Claim 2: $\boxed{\phi_{\max} \leq L_p (1 - w_r(d))}$

Corollary:

$$p_{\min} - p^r \leq \left(\frac{1}{w_r(d)} - 1 \right) (L_p - p_{\min})$$

$$p_{\min,r+d} - p_{\min} \leq (1 - w_r(d))(L_p - p_{\min})$$

Proof of Claim 2: From the definition of L_p,

$$\phi(x) \leq L_p \sum_{|\beta|=d} \frac{d!}{\beta!}(x^\beta - (x)^\beta_{\frac{1}{r+d}})$$

$$= L_p\left(\left(\sum_i x_i\right)^d - \left(\sum_i x_i\right)^d_{\frac{1}{r+d}}\right) = L_p\left((1)^d - (1)^d_{\frac{1}{r+d}}\right)$$

$$= L_p(1 - w_r(d))$$

using the **multinomial identity:**

$$\left(\sum_{i=1}^n x_i\right)^d = \sum_{|\beta|=d} \frac{d!}{\beta!} x_1^\beta \cdots x_n^\beta$$

and the **Vandermonde-Chu identity:**

$$\left(\sum_{i=1}^n x_i\right)^d_t = \sum_{|\beta|=r} \frac{d!}{\beta!}(x_1)^\beta_t \cdots (x_n)^\beta_t$$

A PTAS for the minimization of degree 3 forms on the simplex

Idea: Bound $\phi_{\max}$ in terms of $p_{\min}$ and $p_{\max}$ only

Claim 3: For a degree 3 form,

$$\boxed{\phi_{\max} \leq \tfrac{4}{r+3}\, p_{\max} - \tfrac{r+5}{(r+3)^2}\, p_{\min}}$$

Corollary: $\boxed{p_{\min,r+3} - p_{\min} \leq \tfrac{4}{r+3}(p_{\max} - p_{\min})}$

Proof of Claim 3: Set $t := \frac{1}{r+3}$.
One can verify that

$$\phi(x) = \sum_{i=1}^{n} p_{3e_i}(3tx_i^2 - 2t^2 x_i) + \sum_{1\leq i<j\leq n} (p_{2e_i+e_j} + p_{e_i+2e_j}) t x_i x_j.$$

Evaluating $p(x)$ at the simplex points e_i and $\frac{1}{2}(e_i + e_j)$ yields:

$$p_{\min} \leq p_{3e_i} \leq p_{\max}$$

$$p_{3e_i} + p_{3e_j} + p_{2e_i+e_j} + p_{e_i+2e_j} \leq 8p_{\max}$$

This implies:

$$\begin{aligned} \sum_{1\leq i<j\leq n} (p_{2e_i+e_j} + p_{e_i+2e_j}) x_i x_j \; &\leq \sum_{i<j} (8p_{\max} - p_{3e_i} - p_{3e_j}) x_i x_j \\ &= 8p_{\max} \sum_{i<j} x_i x_j - \sum_i p_{3e_i} x_i (1 - x_i) \end{aligned}$$

Therefore,

$$\phi(x) \leq \underbrace{8tp_{\max} \sum_{i<j} x_i x_j + 4t \sum_i p_{3e_i} x_i^2}_{\phi_1} - \underbrace{t(2t+1) \sum_i p_{3e_i} x_i}_{\phi_2}$$

As $p_{3e_i} \leq p_{\max}$ and $\sum_i x_i = 1$,

$$\phi_1 \leq 4tp_{\max} = \frac{4}{r+3}p_{\max}$$

As $p_{3e_i} \geq p_{\min}$,

$$\phi_2 \geq t(2t+1)p_{\min} = -\frac{r+5}{(r+3)^2}p_{\min}$$