A Second Order Cone Representation for SONC Cones

Jie Wang (AMSS-CAS)
Joint work with Victor Magron

https://wangjie212.github.io/jiewang

August 19, SIAM AG21

Content

Certificates of nonnegativity

Problem

Given a multivariate polynomial f, decide whether f is (globally) nonnegative and certify its nonnegativity if it is.

This is a core problem in real algebraic geometry and has important applications in optimization.

Nonnegative polynomials and polynomial optimization

The unconstrained polynomial optimization problem (POP) can be formulated as follows:

$$f^* := \inf \{f(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\}.$$

Nonnegative polynomials and polynomial optimization

The unconstrained polynomial optimization problem (POP) can be formulated as follows:

$$f^* := \inf \{f(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\}.$$

It is equivalent to

$$f^* = \sup \{\lambda : f(\mathbf{x}) - \lambda \geq 0\}.$$

Sums of squares

A classical approach for certifying nonnegativity of polynomials is the use of sums of squares.

Sums of squares

Given a polynomial $f \in \mathbb{R}[\mathbf{x}] = \mathbb{R}[x_1, \ldots, x_n]$, if there exist polynomials $f_1, \ldots, f_m \in \mathbb{R}[\mathbf{x}]$ such that

$$f = \sum_{i=1}^{m} f_i^2,$$

then we say f is a sum of squares (SOS).

Remark: The computation of SOS decompositions for a given polynomial can be cast as a semidefinite program (SDP).

Not every nonnegative polynomial is an SOS

Theorem (Hilbert)

Every nonnegative polynomial is an SOS only in the univariate case, the quadratic case and the bivariate quartic case.

Not every nonnegative polynomial is an SOS

Theorem (Hilbert)

Every nonnegative polynomial is an SOS only in the univariate case, the quadratic case and the bivariate quartic case.

Except these three cases, there exist nonnegative polynomials which cannot be decomposed as an SOS.

$\triangleright$ Motzkin's polynomial: $x^4y^2 + x^2y^4 + 1 - 3x^2y^2$.

Scalability issue

Assume f has n variables, $2d$ degree, the size of SDP:

- size of PSD matrix: $\binom{n+d}{d}$
- number of equality constraints: $\binom{n+2d}{2d}$

In view of the current state of SDP solvers (e.g. Mosek), tractable polynomials are limited to $n \leq 30$ when $d = 2$ on a standard laptop.

Existing techniques to improve scalability

- Newton polytopes (Reznick, 1978)
- symmetry (Gatermann and Parrilo, 2004)
- correlative sparsity (Waki et al., 2006)
- sign-symmetry (Lofberg, 2009)
- DSOS/SDSOS (Ahmadi and Majumdar, 2018)
- term sparsity (Wang et al., 2019)

Questions

- **Question 1**: Can we have other nonnegativity certificates which exactly inherit sparsity of polynomials?

Questions

- **Question 1**: Can we have other nonnegativity certificates which exactly inherit sparsity of polynomials?
  SONC/SAGE decompositions

- **Question 2**: If the answer is yes, how can we efficiently compute such a nonnegativity certificate for a given polynomial?

Questions

- **Question 1**: Can we have other nonnegativity certificates which exactly inherit sparsity of polynomials?
  SONC/SAGE decompositions
- **Question 2**: If the answer is yes, how can we efficiently compute such a nonnegativity certificate for a given polynomial?
  Geometric/second order cone/relative entropy programming

Circuit polynomials

▷ Trellis: $\mathscr{A} \subseteq (2\mathbb{N})^n$ comprises the vertices of a simplex

Definition (Iliman and de Wolff, 2016)

Let $\mathscr{A}$ be a trellis and $f \in \mathbb{R}[\mathbf{x}]$. Then f is called a circuit polynomial if it is of the form

$$f = \sum_{\boldsymbol{\alpha} \in \mathscr{A}} c_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} - d\mathbf{x}^{\boldsymbol{\beta}},$$

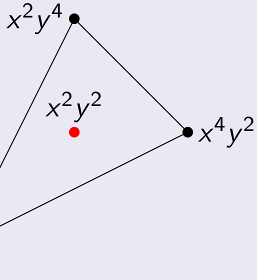
and satisfies:

1. $c_{\boldsymbol{\alpha}} > 0$ for each $\boldsymbol{\alpha} \in \mathscr{A}$;
2. $\boldsymbol{\beta} \in \mathrm{conv}(\mathscr{A})^{\circ}$.

Circuit polynomials

Example (Motzkin's polynomial)

Motzkin's polynomial $M(x,y) = x^4y^2 + x^2y^4 + 1 - 3x^2y^2$ is a nonnegative circuit polynomial.

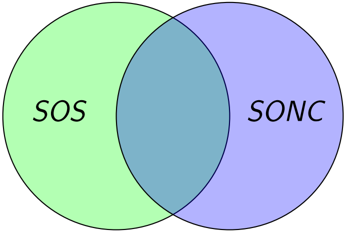


SONC decompositions

$\triangleright$ A polynomial decomposes as a sum of nonnegative circuit polynomials (SONC) $\Longrightarrow$ it is nonnegative

SONC decompositions

▷ A polynomial decomposes as a sum of nonnegative circuit polynomials (SONC) $\Longrightarrow$ it is nonnegative



SONC decompositions preserve sparsity of polynomials

For $f \in \mathbb{R}[\mathbf{x}]$, let

$$\Lambda(f) := \{\boldsymbol{\alpha} \in \mathrm{supp}(f) \mid \boldsymbol{\alpha} \in (2\mathbb{N})^n \text{ and } c_{\boldsymbol{\alpha}} > 0\}$$

and

$$\Gamma(f) := \mathrm{supp}(f) \backslash \Lambda(f)$$

such that we can write

$$f = \sum_{\boldsymbol{\alpha} \in \Lambda(f)} c_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta} \in \Gamma(f)} d_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}.$$

SONC decompositions preserve sparsity of polynomials

For $f \in \mathbb{R}[\mathbf{x}]$, let

$$\Lambda(f) := \{\boldsymbol{\alpha} \in \mathrm{supp}(f) \mid \boldsymbol{\alpha} \in (2\mathbb{N})^n \text{ and } c_{\boldsymbol{\alpha}} > 0\}$$

and

$$\Gamma(f) := \mathrm{supp}(f) \backslash \Lambda(f)$$

such that we can write

$$f = \sum_{\boldsymbol{\alpha} \in \Lambda(f)} c_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta} \in \Gamma(f)} d_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}.$$

For each $\boldsymbol{\beta} \in \Gamma(f)$, let

$$\mathscr{F}(\boldsymbol{\beta}) := \{\Delta \mid \Delta \text{ is a simplex}, \boldsymbol{\beta} \in \Delta^{\circ}, V(\Delta) \subseteq \Lambda(f)\}.$$

SONC decompositions preserve sparsity of polynomials

Theorem (Wang, 2018)

Let $f = \sum_{\alpha\in\Lambda(f)} c_{\alpha}\mathbf{x}^{\alpha} - \sum_{\beta\in\Gamma(f)} d_{\beta}\mathbf{x}^{\beta} \in \mathbb{R}[\mathbf{x}]$. If $f \in \mathrm{SONC}$, then f admits a SONC decomposition:

$$f = \sum_{\beta\in\Gamma(f)} \sum_{\Delta\in\mathscr{F}(\beta)} f_{\beta\Delta} + \sum_{\alpha\in\tilde{\mathscr{A}}} c_{\alpha}\mathbf{x}^{\alpha},$$

where $f_{\beta\Delta}$ is a nonnegative circuit polynomial supported on $V(\Delta)\cup\{\beta\}$ for each β and each Δ, and $\tilde{\mathscr{A}} = \{\alpha \in \Lambda(f) \mid \alpha \notin \bigcup_{\beta\in\Gamma(f)} \bigcup_{\Delta\in\mathscr{F}(\beta)} V(\Delta)\}$.

Remark: A similar theorem on SAGE decompositions was independently proved by Murray et al.

SONC polynomials and sums of binomial squares

Circuit polynomials and sums of binomial squares

- For a subset $M \subseteq \mathbb{N}^n$, let $\overline{A}(M) := \{\frac{1}{2}(\mathbf{u} + \boldsymbol{v}) \mid \mathbf{u} \neq \boldsymbol{v}, \mathbf{u}, \boldsymbol{v} \in M \cap (2\mathbb{N})^n\}$.
- For a trellis $\mathscr{A}$, M is an $\mathscr{A}$-mediated set if $\mathscr{A} \subseteq M \subseteq \overline{A}(M) \cup \mathscr{A}$.

Circuit polynomials and sums of binomial squares

- For a subset $M \subseteq \mathbb{N}^n$, let $\overline{A}(M) := \{\frac{1}{2}(\mathbf{u}+\boldsymbol{v}) \mid \mathbf{u} \neq \boldsymbol{v}, \mathbf{u}, \boldsymbol{v} \in M \cap (2\mathbb{N})^n\}$.
- For a trellis $\mathscr{A}$, M is an $\mathscr{A}$-mediated set if $\mathscr{A} \subseteq M \subseteq \overline{A}(M) \cup \mathscr{A}$.

Theorem (Reznick, 1989; Iliman and de Wolff, 2016)

Let $f = \sum_{\boldsymbol{\alpha}\in\mathscr{A}} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - d\mathbf{x}^{\boldsymbol{\beta}} \in \mathbb{R}[\mathbf{x}]$ be a nonnegative circuit polynomial with $\mathscr{A}$ a trellis. Then f is a sum of binomial squares if and only if there exists an $\mathscr{A}$-mediated set containing $\boldsymbol{\beta}$. More specifically, suppose that $\boldsymbol{\beta}$ belongs to an $\mathscr{A}$-mediated set $M = \{\mathbf{u}_i\}_{i=1}^s$. For each $\mathbf{u}_i \in M\backslash\mathscr{A}$, let $\mathbf{u}_i = \frac{1}{2}(\mathbf{u}_{p(i)} + \mathbf{u}_{q(i)})$. Then f is a sum of binomial squares and $f = \sum_{\mathbf{u}_i\in M\backslash\mathscr{A}}(a_i\mathbf{x}^{\frac{1}{2}\mathbf{u}_{p(i)}} - b_i\mathbf{x}^{\frac{1}{2}\mathbf{u}_{q(i)}})^2$, $a_i, b_i \in \mathbb{R}$.

Theorem (Reznick, 1989; Iliman and de Wolff, 2016)

Theorem (Reznick, 1989; Iliman and de Wolff, 2016) inspires us to leverage sums of binomial squares to compute SONC decompositions. However, there are two obstacles regarding this:

- There may not exist such an $\mathscr{A}$-mediated set containing a given lattice point;
- Even if such a set exists, there is no existing efficient algorithm to compute it.

$\mathscr{A}$-rational mediated sets

- For $M \subseteq \mathbb{Q}^n$, let $\widetilde{A}(M) := \{\frac{1}{2}(\mathbf{u}+\boldsymbol{v}) \mid \mathbf{u} \neq \boldsymbol{v}, \mathbf{u}, \boldsymbol{v} \in M\}$.
- Let $\mathscr{A}$ be a trellis. We say that M is an $\mathscr{A}$-rational mediated set if $\mathscr{A} \subseteq M \subseteq \widetilde{A}(M) \cup \mathscr{A}$.

$\mathscr{A}$-rational mediated sets

- For $M \subseteq \mathbb{Q}^n$, let $\widetilde{A}(M) := \{\frac{1}{2}(\mathbf{u} + \boldsymbol{v}) \mid \mathbf{u} \neq \boldsymbol{v}, \mathbf{u}, \boldsymbol{v} \in M\}$.
- Let $\mathscr{A}$ be a trellis. We say that M is an $\mathscr{A}$-rational mediated set if $\mathscr{A} \subseteq M \subseteq \widetilde{A}(M) \cup \mathscr{A}$.

Theorem (Wang and Magron, 2020)

Given a trellis $\mathscr{A}$ and a lattice point $\boldsymbol{\beta} \in \mathrm{conv}(\mathscr{A})^{\circ}$, there is an algorithm to compute an $\mathscr{A}$-rational mediated set $M_{\mathscr{A}\boldsymbol{\beta}}$ containing $\boldsymbol{\beta}$ such that the denominators of coordinates of points in $M_{\mathscr{A}\boldsymbol{\beta}}$ are odd numbers and the numerators of coordinates of points in $M_{\mathscr{A}\boldsymbol{\beta}} \backslash \{\boldsymbol{\beta}\}$ are even numbers.

The one dimensional case

- For a sequence of natural numbers $A = \{0, q_1, \ldots, q_m, p\}$, if every q_i is an average of two distinct numbers in A, then we say A is a mediated sequence.
- $A = \{0, 2, 4, 5, 8, 11\}$ is a mediated sequence.
- $N(\frac{q}{p})$: the minimum length of mediated sequences containing $0 < q < p$
- $N(\frac{q}{p}) < \frac{1}{2}(\log_2(p) + \frac{3}{2})^2$
- **Conjecture**: $N(\frac{q}{p}) = \lceil \log_2(p) \rceil + 2$ for any $0 < q < p, (p, q) = 1$

Circuit polynomials and sums of binomial squares

Theorem (Wang and Magron, 2020)

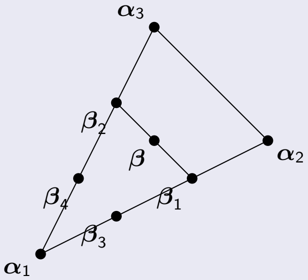
Let $f=\sum_{\boldsymbol{\alpha}\in\mathscr{A}} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}}-d\mathbf{x}^{\boldsymbol{\beta}}\in\mathbb{R}[\mathbf{x}]$ be a circuit polynomial and assume that $M_{\mathscr{A}\boldsymbol{\beta}}=\{\mathbf{u}_i\}_{i=1}^s$ is an $\mathscr{A}$-rational mediated set containing $\boldsymbol{\beta}$ such that the denominators of coordinates of points in $M_{\mathscr{A}\boldsymbol{\beta}}$ are odd numbers and the numerators of coordinates of points in $M_{\mathscr{A}\boldsymbol{\beta}}\backslash\{\boldsymbol{\beta}\}$ are even numbers. For each $\mathbf{u}_i\in M_{\mathscr{A}\boldsymbol{\beta}}\backslash\mathscr{A}$, let $\mathbf{u}_i=\frac{1}{2}(\mathbf{u}_{p(i)}+\mathbf{u}_{q(i)})$. Then f is nonnegative if and only if f can be written as

$$f=\sum_{\mathbf{u}_i\in M_{\mathscr{A}\boldsymbol{\beta}}\backslash\mathscr{A}}(a_i\mathbf{x}^{\frac{1}{2}\mathbf{u}_{p(i)}}-b_i\mathbf{x}^{\frac{1}{2}\mathbf{u}_{q(i)}})^2, a_i, b_i\in\mathbb{R}.$$

An example

Example

Let $f = x^4y^2 + x^2y^4 + 1 - 3x^2y^2$ be Motzkin's polynomial and $\mathscr{A} = \{\boldsymbol{\alpha}_1 = (0,0), \boldsymbol{\alpha}_2 = (4,2), \boldsymbol{\alpha}_3 = (2,4)\}$, $\boldsymbol{\beta} = (2,2)$. Then $M = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\beta}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{\beta}_4\}$ is an $\mathscr{A}$-rational mediated set containing $\boldsymbol{\beta}$.



By a simple computation, we obtain

$f = \frac{3}{2}(x^{\frac{2}{3}}y^{\frac{4}{3}} - x^{\frac{4}{3}}y^{\frac{2}{3}})^2 + (xy^2 - x^{\frac{1}{3}}y^{\frac{2}{3}})^2 + \frac{1}{2}(x^{\frac{2}{3}}y^{\frac{4}{3}} - 1)^2 + (x^2y - x^{\frac{2}{3}}y^{\frac{1}{3}})^2 + \frac{1}{2}(x^{\frac{4}{3}}y^{\frac{2}{3}} - 1)^2.$

SONC polynomials and sums of binomial squares

Theorem (Wang and Magron, 2020)

Let $f=\sum_{\boldsymbol{\alpha}\in\Lambda(f)}c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}}-\sum_{\boldsymbol{\beta}\in\Gamma(f)}d_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}}\in\mathbb{R}[\mathbf{x}]$. For every $\boldsymbol{\beta}\in\Gamma(f)$ and every $\Delta\in\mathscr{F}(\boldsymbol{\beta})$, let $M_{\boldsymbol{\beta}\Delta}$ be a $V(\Delta)$-rational mediated set containing $\boldsymbol{\beta}$ such that the denominators of coordinates of points in $M_{\boldsymbol{\beta}\Delta}$ are odd numbers and the numerators of coordinates of points in $M_{\boldsymbol{\beta}\Delta}\backslash\{\boldsymbol{\beta}\}$ are even numbers. Let $M=\bigcup_{\boldsymbol{\beta}\in\Gamma(f)}\bigcup_{\Delta\in\mathscr{F}(\boldsymbol{\beta})}M_{\boldsymbol{\beta}\Delta}$. For each $\mathbf{u}\in M\backslash\Lambda(f)$, let $\mathbf{u}=\frac{1}{2}(\boldsymbol{v}_{\mathbf{u}}+\mathbf{w}_{\mathbf{u}}),\boldsymbol{v}_{\mathbf{u}}\neq\mathbf{w}_{\mathbf{u}}\in M$. Let $\tilde{\mathscr{A}}=\{\boldsymbol{\alpha}\in\Lambda(f)\mid\boldsymbol{\alpha}\notin\bigcup_{\boldsymbol{\beta}\in\Gamma(f)}\bigcup_{\Delta\in\mathscr{F}(\boldsymbol{\beta})}V(\Delta)\}$. Then $f\in\mathrm{SONC}$ if and only if f can be written as

$$f=\sum_{\mathbf{u}\in M\backslash\Lambda(f)}(a_{\mathbf{u}}\mathbf{x}^{\frac{1}{2}\boldsymbol{v}_{\mathbf{u}}}-b_{\mathbf{u}}\mathbf{x}^{\frac{1}{2}\mathbf{w}_{\mathbf{u}}})^2+\sum_{\boldsymbol{\alpha}\in\tilde{\mathscr{A}}}c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}},a_{\mathbf{u}},b_{\mathbf{u}}\in\mathbb{R}.$$

Second order cones

An n-dimensional (rotated) second order cone (SOC) is defined by

$$\mathcal{Q} := \{\mathbf{x} \in \mathbb{R}^m : ||A\mathbf{x} + \mathbf{b}||_2 \leq \mathbf{c}^T\mathbf{x} + d\},$$

where $A \in \mathbb{R}^{(n-1)\times m}, \mathbf{b} \in \mathbb{R}^{n-1}, \mathbf{c} \in \mathbb{R}^m, d \in \mathbb{R}$.

Second order cones

An n-dimensional (rotated) second order cone (SOC) is defined by

$$\mathcal{Q} := \{\mathbf{x} \in \mathbb{R}^m : ||A\mathbf{x} + \mathbf{b}||_2 \le \mathbf{c}^T\mathbf{x} + d\},$$

where $A \in \mathbb{R}^{(n-1)\times m}, \mathbf{b} \in \mathbb{R}^{n-1}, \mathbf{c} \in \mathbb{R}^m, d \in \mathbb{R}$.

Example

$$\mathbb{S}_+^2 := \{\begin{bmatrix} a & b \\ b & c \end{bmatrix} \in \mathbb{R}^{2\times 2} \mid \begin{bmatrix} a & b \\ b & c \end{bmatrix} \text{ is positive semidefinite}\}$$

is a 3-dimensional second order cone.

Remark: The optimization problem over second order cones can be solved more efficiently than semidefinite programming.

Second order cone lifts of convex cones

$\mathcal{Q}^k = \mathcal{Q} \times \cdots \times \mathcal{Q}$: the Cartesian product of k copies of a second order cone $\mathcal{Q}$

Definition

A convex cone $C \subseteq \mathbb{R}^m$ has a second order cone lift of size k (or simply a $\mathcal{Q}^k$-lift) if it can be written as the projection of a slice of $\mathcal{Q}^k$, that is, there is a subspace L of $\mathcal{Q}^k$ and a linear map $\pi\colon \mathcal{Q}^k \to \mathbb{R}^m$ such that $C = \pi(\mathcal{Q}^k \cap L)$.

Second order cone lifts of convex cones

$\mathcal{Q}^k = \mathcal{Q} \times \cdots \times \mathcal{Q}$: the Cartesian product of k copies of a second order cone $\mathcal{Q}$

Definition

A convex cone $C \subseteq \mathbb{R}^m$ has a second order cone lift of size k (or simply a $\mathcal{Q}^k$-lift) if it can be written as the projection of a slice of $\mathcal{Q}^k$, that is, there is a subspace L of $\mathcal{Q}^k$ and a linear map $\pi\colon \mathcal{Q}^k \to \mathbb{R}^m$ such that $C = \pi(\mathcal{Q}^k \cap L)$.

Theorem (Fawzi, 2018)

The cone $\mathrm{SOS}_{n,2d}$ does not admit any second order cone lift except in the case $(n, 2d) = (1, 2)$.

$(\mathbb{S}_+^2)^k$-lifts of SONC cones

Given $\mathscr{A} \subseteq (2\mathbb{N})^n$, $\mathscr{B}_1 \subseteq \mathrm{conv}(\mathscr{A}) \cap (2\mathbb{N})^n$ and $\mathscr{B}_2 \subseteq \mathrm{conv}(\mathscr{A}) \cap (\mathbb{N}^n \backslash (2\mathbb{N})^n)$ with $\mathscr{A} \cap \mathscr{B}_1 = \varnothing$, define the SONC cone supported on $\mathscr{A}, \mathscr{B}_1, \mathscr{B}_2$ as

$$\begin{aligned}\mathrm{SONC}_{\mathscr{A},\mathscr{B}_1,\mathscr{B}_2} :=& \{(\mathbf{c}_{\mathscr{A}}, \mathbf{d}_{\mathscr{B}_1}, \mathbf{d}_{\mathscr{B}_2}) \in \mathbb{R}_+^{|\mathscr{A}|} \times \mathbb{R}_+^{|\mathscr{B}_1|} \times \mathbb{R}^{|\mathscr{B}_2|} \\ & \mid \sum_{\boldsymbol{\alpha} \in \mathscr{A}} c_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta} \in \mathscr{B}_1 \cup \mathscr{B}_2} d_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}} \in \mathrm{SONC}\}.\end{aligned}$$

$(\mathbb{S}_+^2)^k$-lifts of SONC cones

Given $\mathscr{A} \subseteq (2\mathbb{N})^n$, $\mathscr{B}_1 \subseteq \mathrm{conv}(\mathscr{A}) \cap (2\mathbb{N})^n$ and $\mathscr{B}_2 \subseteq \mathrm{conv}(\mathscr{A}) \cap (\mathbb{N}^n \backslash (2\mathbb{N})^n)$ with $\mathscr{A} \cap \mathscr{B}_1 = \varnothing$, define the SONC cone supported on $\mathscr{A}, \mathscr{B}_1, \mathscr{B}_2$ as

$$\mathrm{SONC}_{\mathscr{A},\mathscr{B}_1,\mathscr{B}_2} := \{(\mathbf{c}_{\mathscr{A}}, \mathbf{d}_{\mathscr{B}_1}, \mathbf{d}_{\mathscr{B}_2}) \in \mathbb{R}_+^{|\mathscr{A}|} \times \mathbb{R}_+^{|\mathscr{B}_1|} \times \mathbb{R}^{|\mathscr{B}_2|} \mid \sum_{\boldsymbol{\alpha} \in \mathscr{A}} c_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta} \in \mathscr{B}_1 \cup \mathscr{B}_2} d_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}} \in \mathrm{SONC}\}.$$

Theorem (Wang and Magron, 2020)

For $\mathscr{A} \subseteq (2\mathbb{N})^n$, $\mathscr{B}_1 \subseteq \mathrm{conv}(\mathscr{A}) \cap (2\mathbb{N})^n$ and $\mathscr{B}_2 \subseteq \mathrm{conv}(\mathscr{A}) \cap (\mathbb{N}^n \backslash (2\mathbb{N})^n)$ with $\mathscr{A} \cap \mathscr{B}_1 = \varnothing$, the SONC cone $\mathrm{SONC}_{\mathscr{A},\mathscr{B}_1,\mathscr{B}_2}$ admits an $(\mathbb{S}_+^2)^k$-lift for some $k \in \mathbb{N}$.

SONC optimization via second order cone programming

SONC optimization

Consider the unconstrained polynomial optimization problem:

$$f^* := \begin{cases} \text{sup} & \lambda \\ \mathrm{s.t.} & f(\mathbf{x}) - \lambda \geq 0. \end{cases}$$

SONC optimization

Consider the unconstrained polynomial optimization problem:

$$f^* := \begin{cases} \text{sup} & \lambda \\ \mathrm{s.t.} & f(\mathbf{x}) - \lambda \geq 0. \end{cases}$$

Replacing the nonnegativity condition by SONC to obtain:

$$f_{sonc} := \begin{cases} \text{sup} & \lambda \\ \mathrm{s.t.} & f(\mathbf{x}) - \lambda \in \mathrm{SONC}. \end{cases}$$

PN-polynomials

Suppose $f = \sum_{\boldsymbol{\alpha}\in\Lambda(f)} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta}\in\Gamma(f)} d_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}} \in \mathbb{R}[\mathbf{x}]$. If $d_{\boldsymbol{\beta}} > 0$ for all $\boldsymbol{\beta} \in \Gamma(f)$, then we call f a PN-polynomial.

PN-polynomials

Suppose $f = \sum_{\boldsymbol{\alpha}\in\Lambda(f)} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta}\in\Gamma(f)} d_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}} \in \mathbb{R}[\mathbf{x}]$. If $d_{\boldsymbol{\beta}} > 0$ for all $\boldsymbol{\beta} \in \Gamma(f)$, then we call f a PN-polynomial.

For a PN-polynomial f, we have

$$f(\mathbf{x}) \geq 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n \iff f(\mathbf{x}) \geq 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n_+$$

PN-polynomials

Suppose $f = \sum_{\boldsymbol{\alpha} \in \Lambda(f)} c_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta} \in \Gamma(f)} d_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}} \in \mathbb{R}[\mathbf{x}]$. If $d_{\boldsymbol{\beta}} > 0$ for all $\boldsymbol{\beta} \in \Gamma(f)$, then we call f a PN-polynomial.

For a PN-polynomial f, we have

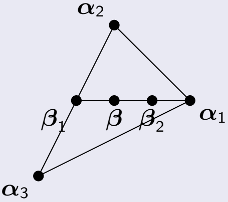
$$f(\mathbf{x}) \geq 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n \iff f(\mathbf{x}) \geq 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n_+$$

$\triangleright$ To represent a SONC PN-polynomial as a sum of binomial squares, we do not require the denominators of coordinates of points in rational mediated sets to be odd. This allows us to decrease the number of binomial squares.

PN-polynomials

An example

Let $f = x^4y^2 + x^2y^4 + 1 - 3x^2y^2$ be Motzkin's polynomial and $\mathscr{A} = \{\boldsymbol{\alpha}_1 = (4,2), \boldsymbol{\alpha}_2 = (2,4), \boldsymbol{\alpha}_3 = (0,0)\}$, $\boldsymbol{\beta} = (2,2)$. Then $\boldsymbol{\beta} = \frac{1}{3}\boldsymbol{\alpha}_1 + \frac{1}{3}\boldsymbol{\alpha}_2 + \frac{1}{3}\boldsymbol{\alpha}_3 = \frac{1}{3}\boldsymbol{\alpha}_1 + \frac{2}{3}(\frac{1}{2}\boldsymbol{\alpha}_2 + \frac{1}{2}\boldsymbol{\alpha}_3)$. Let $\boldsymbol{\beta}_1 = \frac{1}{2}\boldsymbol{\alpha}_2 + \frac{1}{2}\boldsymbol{\alpha}_3$ such that $\boldsymbol{\beta} = \frac{1}{3}\boldsymbol{\alpha}_1 + \frac{2}{3}\boldsymbol{\beta}_1$. Let $\boldsymbol{\beta}_2 = \frac{2}{3}\boldsymbol{\alpha}_1 + \frac{1}{3}\boldsymbol{\beta}_1$. It is easy to check that $M = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\beta}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2\}$ is an $\mathscr{A}$-rational mediated set containing $\boldsymbol{\beta}$.



By a simple computation, we obtain $f = (1 - xy^2)^2 + 2(x^{\frac{1}{2}}y - x^{\frac{3}{2}}y)^2 + (xy - x^2y)^2$. Here we represent f as a sum of three binomial squares with rational exponents.

Converting to PN-polynomials

Let $f = \sum_{\boldsymbol{\alpha} \in \Lambda(f)} c_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta} \in \Gamma(f)} d_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}$ and let
$\tilde{f} = \sum_{\boldsymbol{\alpha} \in \Lambda(f)} c_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta} \in \Gamma(f)} |d_{\boldsymbol{\beta}}| \mathbf{x}^{\boldsymbol{\beta}}$ be its associated PN-polynomial.

Converting to PN-polynomials

Let $f = \sum_{\boldsymbol{\alpha}\in\Lambda(f)} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta}\in\Gamma(f)} d_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}}$ and let
$\tilde{f} = \sum_{\boldsymbol{\alpha}\in\Lambda(f)} c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\beta}\in\Gamma(f)} |d_{\boldsymbol{\beta}}|\mathbf{x}^{\boldsymbol{\beta}}$ be its associated PN-polynomial.

Fact: $f \in \mathrm{SONC} \iff \tilde{f} \in \mathrm{SONC}$.

We can replace f by $\tilde{f}$ without changing the optimum:

$$f_{sonc} = \begin{cases} \sup & \lambda \\ \mathrm{s.t.} & \tilde{f}(\mathbf{x}) - \lambda \in \mathrm{SONC}. \end{cases}$$

SONC optimization via second order cone programming

Suppose $f=\sum_{\boldsymbol{\alpha}\in\Lambda(f)}c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}}-\sum_{\boldsymbol{\beta}\in\Gamma(f)}d_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}}\in\mathbb{R}[\mathbf{x}]$. Let $\{(\mathscr{A}_k,\boldsymbol{\beta}_k)\}_{k=1}^l$ be a circuit cover with $\mathscr{A}_k\subseteq\Lambda(f),\forall k$ and $\Gamma(f)\subseteq\{\boldsymbol{\beta}_k\}_{k=1}^l$.

For each k, let M_k be an $\mathscr{A}_k$-rational mediated set containing $\boldsymbol{\beta}_k$ and $s_k=\#M_k\backslash\mathscr{A}_k$. For each $\mathbf{u}_i^k\in M_k\backslash\mathscr{A}_k$, let us write $\mathbf{u}_i^k=\frac{1}{2}(\boldsymbol{v}_i^k+\mathbf{w}_i^k)$. Then we can relax the SONC optimization problem to a second order cone program (SOCP):

$$f_{socp}:=\begin{cases}\sup & \lambda\\ \text{s.t.} & \tilde{f}(\mathbf{x})-\lambda=\sum_{k=1}^l\sum_{i=1}^{s_k}(2a_i^k\mathbf{x}^{\boldsymbol{v}_i^k}+b_i^k\mathbf{x}^{\mathbf{w}_i^k}-2c_i^k\mathbf{x}^{\mathbf{u}_i^k})+\sum_{\boldsymbol{\alpha}\in\tilde{\mathscr{A}}}c_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}},\\ & (a_i^k,b_i^k,c_i^k)\in\mathcal{Q},\quad\forall i,k,\end{cases}$$

where $\mathcal{Q}$ is a 3-dimensional second order cone.

Fact: $f_{socp}\le f_{sonc}\le f^*$

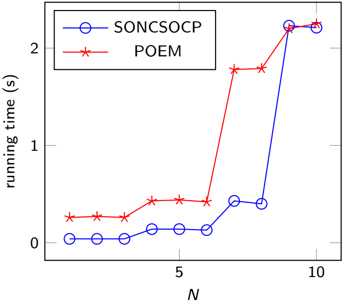
Experimental settings

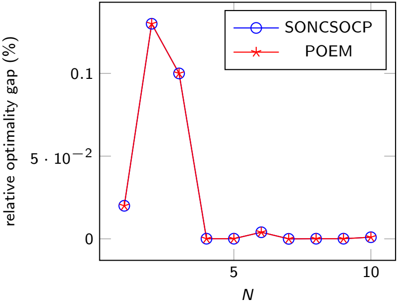
- SONCSOCP: our tool for SONC optimization via SOCP with Mosek as an SOCP solver
- POEM: Seidler and de Wolff's tool for SONC optimization with ECOS as a geometric programming solver
- Benchmarks: Random polynomials generated by Seidler and de Wolff
- Relative optimality gap: $\frac{|f_{up}-f_{lb}|}{|f_{up}|}$, where f_{up} is a local minimum provided by a local solver and f_{lb} is the lower bound given by SONCSOCP or POEM

Results for random polynomials with standard simplex Newton polytopes

Take $N = 10$ polynomials

Number of variables: $10 \sim 40$, degree: $40 \sim 60$, number of terms: $20 \sim 100$

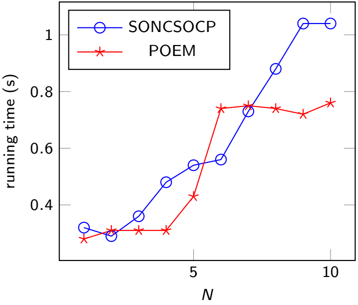


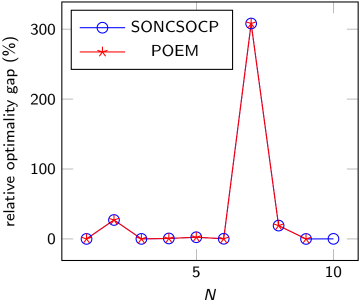


Results for random polynomials with general simplex Newton polytopes

Take $N = 10$ polynomials

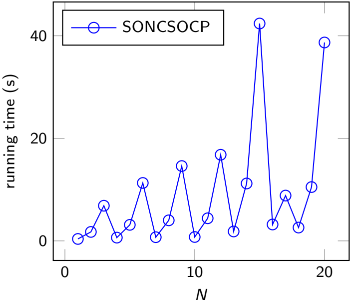
Number of variables: 10, degree: $20 \sim 60$, number of terms: $20 \sim 30$

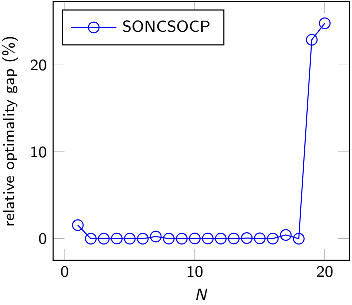




Results for random polynomials with arbitrary Newton polytopes

Take $N = 20$ polynomials
Number of variables: 10, degree: $20 \sim 50$, number of terms: $30 \sim 300$





Summary

- SONC decompositions provide a new way for certifying nonnegativity of sparse polynomials and for (unconstrained) sparse polynomial optimization.
- Each SONC cone admits a second order cone representation.
- We are able to solve SONC optimization via SOCP.

Main References

S. Iliman, T. de Wolff, *Amoebas, nonnegative polynomials and sums of squares supported on circuits*, Res. Math. Sci., 2016.

J. Wang, *Nonnegative polynomials and circuit polynomials*, arXiv:1804.09455, 2018.

J. Wang and V. Magron, *A Second Order Cone Characterization for Sums of Nonnegative Circuits*, in Proceedings of the 45th International Symposium on Symbolic and Algebraic Computation, 2020.

V. Magron and J. Wang, *SONC Optimization and Exact Nonnegativity Certificates via Second-Order Cone Programming*, Journal of Symbolic Computation, 2021

Thank you for your attention!