Symmetric Sums of Squares

Annie Raymond
University of Washington → MSRI → University of Massachusetts

Joint with James Saunderson (Monash University),
Mohit Singh (Georgia Tech), and Rekha Thomas (UW)

November 6, 2017
Goal

Certify the nonnegativity of a symmetric polynomial over the hypercube.

**Our key result:** the runtime does not depend on the number of variables of the polynomial

1. Background
2. Our setting
3. Results
4. Flag algebras
5. Future work
Goal

Certify $p \geq 0$ over the solutions of a system of polynomial equations.
Sums of squares modulo an ideal

Goal

Certify $p \geq 0$ over the solutions of a system of polynomial equations.

Example

Show that $1 - y \geq 0$ whenever $x^2 + y^2 = 1$

$$1 - y = \left( \frac{x}{\sqrt{2}} \right)^2 + \left( \frac{y - 1}{\sqrt{2}} \right)^2 - \frac{1}{2}(x^2 + y^2 - 1)$$

$$= \frac{1}{2} \begin{pmatrix} 1 & x & y \end{pmatrix} 
\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ x \\ y \end{pmatrix} - \frac{1}{2}(x^2 + y^2 - 1)$$
Sums of squares modulo an ideal

Goal
Certify $p \geq 0$ over the solutions of a system of polynomial equations.

Example
Show that $1 - y \geq 0$ whenever $x^2 + y^2 = 1$

$$1 - y = \left(\frac{x}{\sqrt{2}}\right)^2 + \left(\frac{y - 1}{\sqrt{2}}\right)^2 - \frac{1}{2} (x^2 + y^2 - 1)$$

$$= \frac{1}{2} \begin{pmatrix} 1 & x & y \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ x \\ y \end{pmatrix} - \frac{1}{2} (x^2 + y^2 - 1)$$

- Ideal $\mathcal{I} \subseteq \mathbb{R}[x]$
- $\mathcal{V}_{\mathbb{R}}(\mathcal{I})=$its real variety
- $p$ is sos modulo $\mathcal{I}$ if $p \equiv \sum_{i=1}^{l} f_i^2 \mod \mathcal{I}$ (i.e., if $\exists h \in \mathcal{I}$ such that $p = \sum_{i=1}^{l} f_i^2 + h$)
- $p$ is $d$-sos mod $\mathcal{I}$ if $p \equiv \sum_{i=1}^{l} f_i^2 \mod \mathcal{I}$ where $\deg(f_i) \leq d \ \forall \ i$
Sums of squares modulo an ideal

Goal

Certify $p \geq 0$ over the solutions of a system of polynomial equations.

Example

Show that $1 - y \geq 0$ whenever $x^2 + y^2 = 1$

$$1 - y = \left(\frac{x}{\sqrt{2}}\right)^2 + \left(\frac{y - 1}{\sqrt{2}}\right)^2 - \frac{1}{2}(x^2 + y^2 - 1)$$

$$= \frac{1}{2} (1 \times y) \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ x \\ y \end{pmatrix} - \frac{1}{2}(x^2 + y^2 - 1)$$

- Ideal $\mathcal{I} \subseteq \mathbb{R}[x]$
- $\mathcal{V}_{\mathbb{R}}(\mathcal{I}) =$ its real variety
- $p$ is sos modulo $\mathcal{I}$ if $p \equiv \sum_{i=1}^{l} f_i^2 \mod \mathcal{I}$ (i.e., if $\exists h \in \mathcal{I}$ such that $p = \sum_{i=1}^{l} f_i^2 + h$)
- $p$ is $d$-sos mod $\mathcal{I}$ if $p \equiv \sum_{i=1}^{l} f_i^2 \mod \mathcal{I}$ where $\deg(f_i) \leq d \ \forall \ i \iff \exists Q \succeq 0$ such that $p \equiv \nu^\top Q \nu \mod \mathcal{I}$ (semidefinite programming can find $Q$ in $n^{O(d)}$-time)
Our problem

Let $\mathcal{V}_{n,k} = \{0, 1\}^{n\choose k}$ be the $k$-subset discrete hypercube coordinates indexed by $k$-element subsets of $[n]$

Goal

Minimize a symmetric* polynomial over $\mathcal{V}_{n,k}$

*symmetric = $\mathfrak{S}_n$-invariant

$s \cdot x_{i_1i_2...i_k} = x_{s(i_1)s(i_2)...s(i_k)} \forall s \in \mathfrak{S}_n$
Our problem

Let $\mathcal{V}_{n,k} = \{0, 1\}^{n\choose k}$ be the $k$-subset discrete hypercube $\rightarrow$ coordinates indexed by $k$-element subsets of $[n]$

Goal

Minimize a symmetric* polynomial over $\mathcal{V}_{n,k}$

*symmetric $\Rightarrow \mathfrak{S}_n$-invariant

$\mathfrak{s} \cdot x_{i_1i_2...i_k} = x_{s(i_1)s(i_2)...s(i_k)} \forall \mathfrak{s} \in \mathfrak{S}_n$

How?
By finding sos certificates over $\mathcal{V}_{n,k}$ that exploit symmetry, i.e., that we can find in a runtime independent of $n$.

$k = 1$: see Blekherman, Gouveia, Pfeiffer (2014)

$k \geq 2$: ?
Examples of such problems

- **Turán-type problem**
  Given a fixed graph $H$, determine the limiting edge density of a $H$-free graph on $n$ vertices as $n \to \infty$
Examples of such problems

- **Turán-type problem**
  Given a fixed graph $H$, determine the limiting edge density of a $H$-free graph on $n$ vertices as $n \to \infty$

- **Ramsey-type problem**
  Color the edges of $K_n$ ruby or sapphire. Find the smallest $n$ for which you are guaranteed a ruby clique of size $r$ or a sapphire clique of size $s$
Examples of such problems

- **Turán-type problem**
  Given a fixed graph $H$, determine the limiting edge density of a $H$-free graph on $n$ vertices as $n \to \infty$

- **Ramsey-type problem**
  Color the edges of $K_n$ ruby or sapphire. Find the smallest $n$ for which you are guaranteed a ruby clique of size $r$ or a sapphire clique of size $s$

Focus on $\mathcal{V}_n := \mathcal{V}_{n,2} = \{0, 1\}^{\binom{n}{2}}$

$\to$ coordinates are indexed by pairs $ij$, $1 \leq i < j \leq n$
Passing to optimization - Turán-type problem

Example

Forbidding triangles in a graph on \(n\) vertices, find

\[
\max \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} x_{ij}
\]

s.t. \(x_{ij}^2 = x_{ij}\) \(\forall 1 \leq i < j \leq n\)

\(x_{ij}x_{jk}x_{ik} = 0\) \(\forall 1 \leq i < j < k \leq n\)

In particular, show that this is at most \(\frac{1}{2} + O\left(\frac{1}{n}\right)\)

→ show that \(\frac{1}{2} + O\left(\frac{1}{n}\right) - \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} x_{ij} \geq 0\)
Example (continued)

Find $Q \succeq 0$ and $d \in \mathbb{Z}^+$ such that

$$
\frac{1}{2} + O \left(\frac{1}{n}\right) - \frac{1}{n^2} \sum_{1 \leq i < j \leq n} x_{ij} \equiv v^\top Q v \mod \mathcal{I}
$$

where $v =$ vector of basis elements of $(\mathbb{R}[x]/\mathcal{I})_d$ and

$$
\mathcal{I} = \langle x_{ij}^2 - x_{ij} \forall 1 \leq i < j \leq n, \\
x_{ij}x_{jk}x_{ik} \forall 1 \leq i < j < k \leq n \rangle
$$
Issue with passing to optimization - Turán-type problem

Example (continued)

Find $Q \succeq 0$ and $d \in \mathbb{Z}^+$ such that

$$
\frac{1}{2} + O\left(\frac{1}{n}\right) - \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} x_{ij} \equiv v^\top Qv \mod \mathcal{I}
$$

where $v =$vector of basis elements of $(\mathbb{R}[x]/\mathcal{I})_d$ and

$$
\mathcal{I} = \langle x_{ij}^2 - x_{ij} \ \forall 1 \leq i < j \leq n, \quad x_{ij}x_{jk}x_{ik} \ \forall 1 \leq i < j < k \leq n \rangle
$$

Can we do this with semidefinite programming?
The runtime would be $\binom{n}{2}^{O(d)}$
Issue with passing to optimization - Turán-type problem

Example (continued)

Find $Q \succeq 0$ and $d \in \mathbb{Z}^+$ such that

$$
\frac{1}{2} + O \left( \frac{1}{n} \right) - \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} x_{ij} \equiv v^\top Q v \quad \text{mod } \mathcal{I}
$$

where $v =$ vector of basis elements of $(\mathbb{R}[x]/\mathcal{I})_d$ and

$$
\mathcal{I} = \langle x_{ij}^2 - x_{ij} \forall 1 \leq i < j \leq n, \quad x_{ij} x_{jk} x_{ik} \forall 1 \leq i < j < k \leq n \rangle
$$

Can we do this with semidefinite programming?
The runtime would be $\binom{n}{2} O(d) \to \infty$ as $n \to \infty$. 
Foreshadowing

Example

The following is a sos proof of Mantel’s theorem

\[
(1 \ q_1) \begin{pmatrix}
\frac{(n-1)^2}{2} & -\frac{2(n-1)}{n} \\
-\frac{2(n-1)}{n} & \frac{8}{n^2}
\end{pmatrix} \begin{pmatrix}
1 \\
q_1
\end{pmatrix} + \text{sym} \left( (q_2) \left( \frac{8}{n^2} \right) (q_2) \right)
\]

where \( q_1 = \sum_{i<j} x_{ij} \) and \( q_2 = \sum_{i<j} x_{ij} - \frac{n-2}{2} \sum_{i=1}^{n-1} x_{in} \)

Key features of desired sos certificates:

- exploits symmetry
- constant size
- entries are functions of \( n \)
Representation theory needed for exploiting symmetry

\[ (\mathbb{R}[x]/\mathcal{I})_d =: V = \bigoplus_{\lambda \vdash n} V_{\lambda} \] isotypic decomposition

- partition \( \lambda = (5, 3, 3, 1) \) for \( n = 12 \)
Representation theory needed for exploiting symmetry

- \((\mathbb{R}[x]/\mathcal{I})_d =: V = \bigoplus_{\lambda \vdash n} V_\lambda\) isotropic decomposition
  - partition \(\lambda = (5, 3, 3, 1)\) for \(n = 12\)
- \(V_\lambda = \bigoplus_{\tau_\lambda} W_{\tau_\lambda}\)
  - shape of \(\lambda\): standard tableau \(\tau_\lambda\):
    - \(\mathcal{R}_{\tau_\lambda}\) := row group of \(\tau_\lambda\) (fixes the rows of \(\tau_\lambda\))
    - \(W_{\tau_\lambda} := (V_\lambda)^{\mathcal{R}_{\tau_\lambda}}\) = subspace of \(V_\lambda\) fixed by \(\mathcal{R}_{\tau_\lambda}\)
    - \(n_\lambda :=\) number of standard tableaux of shape \(\lambda\)
    - \(m_\lambda :=\) dimension of \(W_{\tau_\lambda}\)
Representation theory needed for exploiting symmetry

- $(\mathbb{R}[x]/\mathcal{I})_d =: V = \bigoplus_{\lambda \vdash n} V_{\lambda}$ isotypic decomposition
  - partition $\lambda = (5, 3, 3, 1)$ for $n = 12$

- $V_{\lambda} = \bigoplus_{\tau_{\lambda}} W_{\tau_{\lambda}}$
  - shape of $\lambda$: standard tableau $\tau_{\lambda}$:
    $\begin{array}{cccc}
    1 & 4 & 5 & 6 \\
    2 & 7 & 10 \\
    3 & 8 & 12 \\
    11
    \end{array}$
  - $\mathcal{R}_{\tau_{\lambda}} :=$ row group of $\tau_{\lambda}$ (fixes the rows of $\tau_{\lambda}$)
  - $W_{\tau_{\lambda}} := (V_{\lambda})^{\mathcal{R}_{\tau_{\lambda}}} =$ subspace of $V_{\lambda}$ fixed by $\mathcal{R}_{\tau_{\lambda}}$
  - $n_{\lambda} :=$ number of standard tableaux of shape $\lambda$
  - $m_{\lambda} :=$ dimension of $W_{\tau_{\lambda}}$

$$V = \bigoplus_{\lambda \vdash n} \bigoplus_{\tau_{\lambda}} W_{\tau_{\lambda}}$$

Note: $\dim(V) = \sum_{\lambda \vdash n} m_{\lambda} n_{\lambda}$
Gatermann-Parrilo symmetry-reduction technique

Recall: $p$ d-sos mod $\mathcal{I} \iff \exists \ Q \succeq 0 \ \text{s.t.} \ p \equiv v^\top Qv \ \text{mod} \ \mathcal{I}$

where $v =$vector of basis elements of $(\mathbb{R}[x]/\mathcal{I})_d$
Gatermann-Parrilo symmetry-reduction technique

Recall: $p$ d-sos mod $\mathcal{I} \iff \exists \ Q \succeq 0$ s.t. $p \equiv \nu^\top Q \nu$ mod $\mathcal{I}$

where $\nu =$vector of basis elements of $(\mathbb{R}[x]/\mathcal{I})_d$

**Theorem (Gatermann-Parrilo, 2004)**

For each $\lambda$, fix $\tau_\lambda$ and find a symmetry-adapted basis $\{b_{1\lambda}^\top, \ldots, b_{m_\lambda}^\top\}$ for $W_{\tau_\lambda}$.

If $p$ is symmetric and d-sos mod $\mathcal{I}$, then

$$p \equiv \sum_{\lambda \vdash n} \text{sym}(b^\top Q_\lambda b),$$

where $b = (b_{1\lambda}^\top, \ldots, b_{m_\lambda}^\top)^\top$ and $Q_\lambda \succeq 0$ has size $m_\lambda \times m_\lambda$.

**Gain:** size of SDP is $\sum_{\lambda \vdash n} m_\lambda$ instead of $\sum_{\lambda \vdash n} m_\lambda n_\lambda$
Recall: \( p \ d\text{-sos mod } I \iff \exists Q \succeq 0 \text{ s.t. } p \equiv \nu^\top Q\nu \mod I \)

where \( \nu \) = vector of basis elements of \((\mathbb{R}[x]/I)_d\)

**Theorem (Gatermann-Parrilo, 2004)**

For each \( \lambda \), fix \( \tau_\lambda \) and find a symmetry-adapted basis \( \{b_{1\lambda}^{\tau}, \ldots, b_{m_\lambda\lambda}^{\tau}\} \) for \( W_{\tau_\lambda} \).

If \( p \) is symmetric and \( d\text{-sos mod } I \), then

\[
p = \sum_{\lambda \vdash n} \text{sym}(b^\top Q_\lambda b),
\]

where \( b = (b_{1\lambda}^{\tau}, \ldots, b_{m_\lambda\lambda}^{\tau})^\top \) and \( Q_\lambda \succeq 0 \) has size \( m_\lambda \times m_\lambda \).

Gain: size of SDP is \( \sum_{\lambda \vdash n} m_\lambda \) instead of \( \sum_{\lambda \vdash n} m_\lambda n_\lambda \)

→ how much smaller is the size of this SDP?
Gatermann-Parrilo symmetry-reduction technique

Recall: \( p \ d\text{-sos mod } \mathcal{I} \leftrightarrow \exists \ Q \succeq 0 \text{ s.t. } p \equiv v^\top Q v \mod \mathcal{I} \)
where \( v = \text{vector of basis elements of } (\mathbb{R}[x]/\mathcal{I})_d \)

**Theorem (Gatermann-Parrilo, 2004)**

For each \( \lambda \), fix \( \tau_\lambda \) and find a **symmetry-adapted basis** \( \{ b_1^{\tau_\lambda}, \ldots, b_{m_\lambda}^{\tau_\lambda} \} \) for \( W_{\tau_\lambda} \).

\( \rightarrow \) **complexity of the algorithm depends on** \( n \)

If \( p \) is symmetric and \( d\text{-sos mod } \mathcal{I} \), then

\[
p = \sum_{\lambda \vdash n} \text{sym}(b^\top Q_\lambda b),
\]

where \( b = (b_1^{\tau_\lambda}, \ldots, b_{m_\lambda}^{\tau_\lambda})^\top \) and \( Q_\lambda \succeq 0 \) has size \( m_\lambda \times m_\lambda \).

Gain: size of SDP is \( \sum_{\lambda \vdash n} m_\lambda \) instead of \( \sum_{\lambda \vdash n} m_\lambda n_\lambda \)

\( \rightarrow \) **how much smaller is the size of this SDP?**
**Theorem (RSST, 2016)**

If $p$ is symmetric and $d$-sos, then it has a symmetry-reduced sos certificate that can be obtained by solving a SDP of size independent of $n$ by keeping only a few partitions in Gatermann-Parrilo.
Theorem (RSST, 2016)

If \( p \) is symmetric and \( d\)-sos, then it has a symmetry-reduced sos certificate that can be obtained by solving a SDP of size independent of \( n \) by keeping only a few partitions in Gatermann-Parrilo.

Example

In the sos proof of Mantel’s theorem

\[
\begin{pmatrix}
1 & q_1
\end{pmatrix}
\begin{pmatrix}
\frac{(n-1)^2}{2} & -\frac{2(n-1)}{n}
\end{pmatrix}
\begin{pmatrix}
1 \\
q_1
\end{pmatrix}
+ \text{sym}\left( (q_2) \left( \frac{8}{n^2} \right) (q_2) \right)
\]

\[\rightarrow \text{kept partitions } (n) = \underbrace{\text{\(n\) partitions}}_{\text{\(n\) partitions}} \text{ and } (n - 1, 1) = \underbrace{\text{\(n-1\) partitions}}_{\text{\(n-1\) partitions}}\]
Bypassing symmetry-adapted basis

**Theorem (RSST, 2016)**

In Gatermann-Parrilo, instead of a symmetry-adapted basis, one can use

- a spanning set for $W_{\tau \lambda}$ for $\lambda \geq_{\text{lex}} n - 2d$.
- of size independent of $n$
- that is easy to generate
Bypassing symmetry-adapted basis

Theorem (RSST, 2016)

In Gatermann-Parrilo, instead of a symmetry-adapted basis, one can use

- a spanning set for $W_{\tau,\lambda}$ for $\lambda \geq \text{lex}^{n-2d}$

- of size independent of $n$

- that is easy to generate

Examples of spanning sets containing $W_{\tau,\lambda}$

- $\text{sym}_{\tau,\lambda}(x^m) := \frac{1}{|R_{\tau,\lambda}|} \sum_{\mathbf{s} \in R_{\tau,\lambda}} \mathbf{s} \cdot x^m$

- an appropriate Möbius transformation
Razborov’s flag algebras for Turán-type problems

Use flags (=partially labelled graphs) to certify a symmetric inequality that gives a good upper bound for Turán-type problems

Key features:
- sums of squares of graph densities
- $n$ disappears
- asymptotic results for dense graphs

Theorem (Razborov, 2010)
If $\mathcal{A} = \{K_4^3\}$, then $\max_{G:|V(G)| \to \infty} d(G) \leq 0.561666$.
If $\mathcal{A} = \{K_4^3, H_1\}$, then $\max_{G:|V(G)| \to \infty} d(G) = 5/9$. 
Connection of spanning sets to flag algebras

**Theorem (RSST, 2016)**

Flags provide spanning sets for $W_{\tau,\lambda}$ of size independent of $n$.

*If $p$ is symmetric and d-sos, then its nonnegativity can be established through flags on $kd$ vertices (even in restricted cases).*

**Theorem (R., Singh, Thomas, 2015)**

*Every flag sos polynomial of degree $kd$ can be written as a succinct d-sos.*

**Theorem (RSST, 2016)**

*Flag methods are equivalent to standard symmetry-reduction methods for finding sos certificates over discrete hypercubes.*
Consequences of this connection

Corollary (RSST, 2016)

It is possible to use flags for a fixed $n$, not just asymptotic situations.
Consequences of this connection

**Corollary (RSST, 2016)**

*It is possible to use flags for a fixed n, not just asymptotic situations*

**Corollary (RSST, 2016)**

*It is possible to use flags for extremal graph theoretic problems in the sparse setting.*
Consequences of this connection

**Corollary (RSST, 2016)**

*It is possible to use flags for a fixed $n$, not just asymptotic situations.*

**Corollary (RSST, 2016)**

*It is possible to use flags for extremal graph theoretic problems in the sparse setting.*

**Corollary (RSST, 2016)**

*There exists a family of symmetric nonnegative polynomials of fixed degree that cannot be certified exactly with any fixed set of flags, namely*

\[
\frac{1}{(\frac{n}{2})^2} \left( \sum_{e \in E(K_n)} x_e - \left\lfloor \frac{n}{2} \right\rfloor \right) \left( \sum_{e \in E(K_n)} x_e - \left\lfloor \frac{n}{2} \right\rfloor - 1 \right) + O\left( \frac{1}{n^2} \right)
\]
Open problems

- Find a concrete family of nonnegative polynomials on $\binom{n}{k}$ variables that one cannot approximate up to an error of order $O\left(\frac{1}{n}\right)$ with finitely many flags or with sums of squares of fixed degree.
- Provide certificates for open problems over $\mathcal{V}_{n,k}$ using symmetric sums of squares.
Open problems

- Find a concrete family of nonnegative polynomials on $\binom{n}{k}$ variables that one cannot approximate up to an error of order $O\left(\frac{1}{n}\right)$ with finitely many flags or with sums of squares of fixed degree.
- Provide certificates for open problems over $V_{n,k}$ using symmetric sums of squares.

Thank you!