# Semidefinite hierarchies for polynomial optimization

Monique Laurent

Journées SMAI-MODE 2016, Toulouse

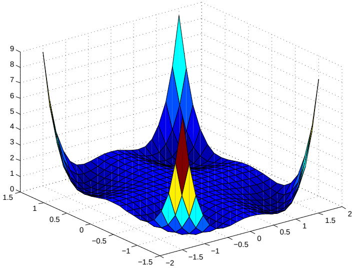
# Semidefinite hierarchies for polynomial optimization

Monique Laurent

Journées SMAI-MODE 2016, Toulouse

Based on joint works with Etienne de Klerk (Tilburg), Zhao Sun (Montreal), Jean Lasserre, Roxana Hess (Toulouse), Pablo Parrilo (MIT)

# Polynomial optimization




Minimize a polynomial function $f$ over a region

$$K = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$$

defined by polynomial inequalities

Compute: $f_{\min} = \min_{x \in K} f(x)$

This is a hard problem, even for simple sets $K$ like

- the standard simplex

$$\Delta_n = \{x \in \mathbb{R}^n : x_1, \ldots, x_n \geq 0, \ \sum_{l=1}^{n} x_i = 1\}$$

- the hypercube $Q_n = [0,1]^n$

- the unit sphere

$$S^{n-1} = \{x \in \mathbb{R}^n : \sum_{i=1}^{n} x_i^2 = 1\}$$

It captures **hard combinatorial optimization problems** like computing the stability number $\alpha(G)$ and Max-Cut.

# Polynomial optimization formulations for $\alpha(G)$

- Optimization over the simplex: [Motzkin-Straus 1965]

$$\frac{1}{\alpha(G)} = \min\ x^T(I + A_G)x \ \text{ s.t. } \sum_{v\in V} x_v = 1,\ x_v \geq 0\ (v \in V)$$

- Optimization over the hypercube: [Park-Hong 2011]

$$\alpha(G) = \max \sum_{u\in V} x_u - \sum_{uv\in E} x_u x_v \ \text{ s.t. } x \in [0,1]^n$$

- Optimization over the unit sphere: [Nesterov 2003]

$$\frac{2\sqrt{2}}{3\sqrt{3}}\sqrt{1 - \frac{1}{\alpha(G)}} = \max\ 2\sum_{ij\in\overline{G}} z_{ij} y_i y_j \ \text{ s.t. } (y,z) \in S^{n+m-1}$$

# Lower bounds for polynomial optimization

To approximate:

$$f_{\min} = \min_{x \in K} f(x), \qquad \text{where } K = \{x : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$$

use LP/SDP hierarchies

[Shor (1987), Nesterov (2000), Parrilo, Lasserre (2000–)]

- Express $f_{\min} = \sup \ \lambda \ \text{ s.t. } f(x) - \lambda \geq 0$ over $K$
- Replace nonnegativity by easier sufficient conditions

# Lower bounds for polynomial optimization

To approximate:

$$f_{\min} = \min_{x \in K} f(x), \qquad \text{where } K = \{x : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$$

use LP/SDP hierarchies

[Shor (1987), Nesterov (2000), Parrilo, Lasserre (2000–)]

- Express $f_{\min} = \sup \ \lambda$ s.t. $f(x) - \lambda \geq 0$ over $K$
- Replace nonnegativity by easier sufficient conditions

Testing whether $f$ is nonnegative is hard, but

testing whether $f$ is a sum-of-squares (SoS): $f = \sum_j g_j^2$

can be done **efficiently** using semidefinite programming (SDP)

# Lower bounds for polynomial optimization

$$f_{\min} = \min_{x \in K} f(x), \qquad \text{where } K = \{x : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$$

- Express $f_{\min} = \sup\ \lambda$ s.t. $f(x) - \lambda \geq 0$ over $K$
- Replace nonnegativity by easier sufficient conditions:

# Lower bounds for polynomial optimization

$$f_{\min} = \min_{x \in K} f(x), \qquad \text{where } K = \{x : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$$

- Express $f_{\min} = \sup \ \lambda$ s.t. $f(x) - \lambda \geq 0$ over $K$
- Replace nonnegativity by easier sufficient conditions:

(1) [LP: Handelman type] $f - \lambda = \sum_{\alpha \in \mathbb{N}^m} \lambda_\alpha \prod_{j=1}^{m} g_j^{\alpha_j}$, where $\lambda_\alpha \geq 0$

# Lower bounds for polynomial optimization

$$f_{\min} = \min_{x \in K} f(x), \qquad \text{where } K = \{x : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$$

- Express $f_{\min} = \sup\ \lambda$ s.t. $f(x) - \lambda \geq 0$ over $K$
- Replace nonnegativity by easier sufficient conditions:

(1) [LP: Handelman type] $f - \lambda = \sum_{\alpha \in \mathbb{N}^m} \lambda_\alpha \prod_{j=1}^{m} g_j^{\alpha_j}$, where $\lambda_\alpha \geq 0$

(2) [SoS: Schmüdgen type] $f - \lambda = \sum_{J \subseteq [m]} \sigma_J \prod_{j \in J} g_j$, where $\sigma_J$ SoS

# Lower bounds for polynomial optimization

$$f_{\min} = \min_{x \in K} f(x), \qquad \text{where } K = \{x : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$$

- Express $f_{\min} = \sup \ \lambda \ \text{ s.t. } f(x) - \lambda \geq 0 \text{ over } K$
- Replace nonnegativity by easier sufficient conditions:

(1) [LP: Handelman type] $\quad f - \lambda = \sum_{\alpha \in \mathbb{N}^m} \lambda_\alpha \prod_{j=1}^{m} g_j^{\alpha_j}$, where $\lambda_\alpha \geq 0$

(2) [SoS: Schmüdgen type] $\quad f - \lambda = \sum_{J \subseteq [m]} \sigma_J \prod_{j \in J} g_j$, where $\sigma_J$ SoS

(3) [SoS: Putinar type] $\quad f - \lambda = \sigma_0 + \sum_{j=1}^{m} \sigma_j g_j$, where $\sigma_j$ SoS

# Lower bounds for polynomial optimization

$$f_{\min} = \min_{x \in K} f(x), \qquad \text{where } K = \{x : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$$

- Express $f_{\min} = \sup\ \lambda$ s.t. $f(x) - \lambda \geq 0$ over $K$
- Replace nonnegativity by easier sufficient conditions:

(1) [LP: Handelman type] $f - \lambda = \sum_{\alpha \in \mathbb{N}^m} \lambda_\alpha \prod_{j=1}^{m} g_j^{\alpha_j}$, where $\lambda_\alpha \geq 0$

(2) [SoS: Schmüdgen type] $f - \lambda = \sum_{J \subseteq [m]} \sigma_J \prod_{j \in J} g_j$, where $\sigma_J$ SoS

(3) [SoS: Putinar type] $f - \lambda = \sigma_0 + \sum_{j=1}^m \sigma_j g_j$, where $\sigma_j$ SoS

- Get lower bounds $\underline{f}^{(r)}_{\mathrm{lp}}, \underline{f}^{(r)}_{\mathrm{sosS}}, \underline{f}^{(r)}_{\mathrm{sosP}}$ for $f_{\min}$ by **bounding degrees:** $\deg g^\alpha,\ \deg(\sigma_J g^J),\ \deg(\sigma_j g_j) \leq r$.

# Lower bounds for polynomial optimization

$$f_{\min} = \min_{x \in K} f(x), \qquad \text{where } K = \{x : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$$

- Express $f_{\min} = \sup \ \lambda \ \text{ s.t. } f(x) - \lambda \geq 0 \text{ over } K$
- Replace nonnegativity by easier sufficient conditions:

(1) [LP: Handelman type] $f - \lambda = \sum_{\alpha \in \mathbb{N}^m} \lambda_\alpha \prod_{j=1}^{m} g_j^{\alpha_j}$, where $\lambda_\alpha \geq 0$

(2) [SoS: Schmüdgen type] $f - \lambda = \sum_{J \subseteq [m]} \sigma_J \prod_{j \in J} g_j$, where $\sigma_J$ SoS

(3) [SoS: Putinar type] $f - \lambda = \sigma_0 + \sum_{j=1}^{m} \sigma_j g_j$, where $\sigma_j$ SoS

- Get lower bounds $\underline{f}_{\mathrm{lp}}^{(r)}, \underline{f}_{\mathrm{sosS}}^{(r)}, \underline{f}_{\mathrm{sosP}}^{(r)}$ for $f_{\min}$ by **bounding degrees:** $\deg g^\alpha, \ \deg(\sigma_J g^J), \ \deg(\sigma_j g_j) \leq r$. Clearly:

$$\{ \ \underline{f}_{\mathrm{lp}}^{(r)}, \ \underline{f}_{\mathrm{sosP}}^{(r)} \ \} \leq \underline{f}_{\mathrm{sosS}}^{(r)} \leq f_{\min}$$

# Representation results for positive polynomials

## Theorem

*Assume $K$ is compact and $f$ is* **strictly positive** *on $K$.*

# Representation results for positive polynomials

## Theorem

*Assume $K$ is compact and $f$ is* **strictly positive** *on $K$.*

(1) *[Krivine-Handelman] Assume $K$* **full-dimensional polytope** *(all $g_j$ have degree 1). Then* $f = \sum_{\alpha \in \mathbb{N}^m} \lambda_\alpha \prod_{j=1}^m g_j^{\alpha_j}$, *where* $\lambda_\alpha \geq 0$.

# Representation results for positive polynomials

## Theorem

*Assume $K$ is compact and $f$ is* **strictly positive** *on $K$.*

(1) *[Krivine-Handelman] Assume $K$* **full-dimensional polytope** *(all $g_j$ have degree 1). Then $f = \sum_{\alpha \in \mathbb{N}^m} \lambda_\alpha \prod_{j=1}^m g_j^{\alpha_j}$, where $\lambda_\alpha \geq 0$.*

(2) *[Schmüdgen 1991] Then $f = \sum_{J \subseteq [m]} \sigma_J \prod_{j \in J} g_j$, where $\sigma_J$ SoS.*

# Representation results for positive polynomials

## Theorem

*Assume $K$ is compact and $f$ is* **strictly positive** *on $K$.*

(1) *[Krivine-Handelman] Assume $K$* **full-dimensional polytope** *(all $g_j$ have degree 1). Then* $f = \sum_{\alpha \in \mathbb{N}^m} \lambda_\alpha \prod_{j=1}^m g_j^{\alpha_j}$, *where* $\lambda_\alpha \geq 0$.

(2) *[Schmüdgen 1991] Then* $f = \sum_{J \subseteq [m]} \sigma_J \prod_{j \in J} g_j$, *where $\sigma_J$ SoS.*

(3) *[Putinar 1993] Assume $\{x : g_j(x) \geq 0\}$ compact for some $j$. Then* $f = \sigma_0 + \sum_{j=1}^m \sigma_j g_j$, *where $\sigma_j$ SoS.*

# Representation results for positive polynomials

## Theorem

*Assume $K$ is compact and $f$ is* **strictly positive** *on $K$.*

(1) *[Krivine-Handelman] Assume $K$* **full-dimensional polytope** *(all $g_j$ have degree 1). Then* $f = \sum_{\alpha \in \mathbb{N}^m} \lambda_\alpha \prod_{j=1}^m g_j^{\alpha_j}$, *where* $\lambda_\alpha \geq 0$.

(2) *[Schmüdgen 1991] Then* $f = \sum_{J \subseteq [m]} \sigma_J \prod_{j \in J} g_j$, *where $\sigma_J$ SoS.*

(3) *[Putinar 1993] Assume $\{x : g_j(x) \geq 0\}$ compact for some $j$. Then* $f = \sigma_0 + \sum_{j=1}^m \sigma_j g_j$, *where $\sigma_j$ SoS.*

- **Asymptotic convergence** of the lower bounds to $f_{\min}$. [Lasserre 01]

# Representation results for positive polynomials

## Theorem

*Assume $K$ is compact and $f$ is* **strictly positive** *on $K$.*

(1) *[Krivine-Handelman] Assume $K$* **full-dimensional polytope** *(all $g_j$ have degree 1). Then $f = \sum_{\alpha \in \mathbb{N}^m} \lambda_\alpha \prod_{j=1}^m g_j^{\alpha_j}$, where $\lambda_\alpha \geq 0$.*

(2) *[Schmüdgen 1991] Then $f = \sum_{J \subseteq [m]} \sigma_J \prod_{j \in J} g_j$, where $\sigma_J$ SoS.*

(3) *[Putinar 1993] Assume $\{x : g_j(x) \geq 0\}$ compact for some $j$. Then $f = \sigma_0 + \sum_{j=1}^m \sigma_j g_j$, where $\sigma_j$ SoS.*

- **Asymptotic convergence** of the lower bounds to $f_{\min}$. [Lasserre 01]
- **Finite convergence** holds generically for SoS bounds. [Nie 14]

# Representation results for positive polynomials

**Theorem**

*Assume $K$ is compact and $f$ is* **strictly positive** *on $K$.*

(1) *[Krivine-Handelman] Assume $K$* **full-dimensional polytope** *(all $g_j$ have degree 1). Then* $f = \sum_{\alpha \in \mathbb{N}^m} \lambda_\alpha \prod_{j=1}^m g_j^{\alpha_j}$, *where $\lambda_\alpha \geq 0$.*

(2) *[Schmüdgen 1991] Then* $f = \sum_{J \subseteq [m]} \sigma_J \prod_{j \in J} g_j$, *where $\sigma_J$ SoS.*

(3) *[Putinar 1993] Assume $\{x : g_j(x) \geq 0\}$ compact for some $j$. Then* $f = \sigma_0 + \sum_{j=1}^m \sigma_j g_j$, *where $\sigma_j$ SoS.*

- **Asymptotic convergence** of the lower bounds to $f_{\min}$. [Lasserre 01]
- **Finite convergence** holds generically for SoS bounds. [Nie 14]
- What about the **rate of convergence**?

# Rate of convergence of SoS lower bounds

## Theorem

*Assume $K \subseteq (-1,1)^n$. For $f = \sum_\alpha f_\alpha x^\alpha$, set $L_f = \max_\alpha |f_\alpha| \frac{\alpha!}{|\alpha|!}$.*

- *[Schweighofer 2004]* **Analysis of Schmüdgen type bounds:**

  *There exists a constant $c > 0$ such that for any polynomial $f$ of degree $d$:*

$$f_{\min} - \underline{f}_{sosS}^{(r)} \leq c d^4 n^{2d} L_f \frac{1}{\sqrt[c]{r}} \quad \text{for } r \geq c d^c n^{cd}.$$

# Rate of convergence of SoS lower bounds

**Theorem**

*Assume $K \subseteq (-1,1)^n$. For $f = \sum_\alpha f_\alpha x^\alpha$, set $L_f = \max_\alpha |f_\alpha| \frac{\alpha!}{|\alpha|!}$.*

- *[Schweighofer 2004]* **Analysis of Schmüdgen type bounds:**
  *There exists a constant $c > 0$ such that for any polynomial $f$ of degree $d$:*
  $$f_{\min} - \underline{f}_{sosS}^{(r)} \leq cd^4 n^{2d} L_f \frac{1}{\sqrt[c]{r}} \quad \text{for } r \geq cd^c n^{cd}.$$

- *[Nie-Schweighofer 2007]* **Analysis of Putinar type bounds:**
  *There exists a constant $c' > 0$ such that for any polynomial $f$ of degree $d$:*
  $$f_{\min} - \underline{f}_{sosP}^{(r)} \leq 6d^3 n^{2d} L_f \frac{1}{\sqrt[c']{\log \frac{r}{c'}}} \quad \text{for } r \geq c' \exp((2d^2 n^d)^{c'}).$$

# Rate of convergence of SoS lower bounds

## Theorem

*Assume* $K \subseteq (-1,1)^n$. *For* $f = \sum_\alpha f_\alpha x^\alpha$, *set* $L_f = \max_\alpha |f_\alpha| \frac{\alpha!}{|\alpha|!}$.

- *[Schweighofer 2004]* **Analysis of Schmüdgen type bounds:**

  *There exists a constant $c > 0$ such that for any polynomial $f$ of degree $d$:*

$$f_{\min} - \underline{f}_{sosS}^{(r)} \leq c d^4 n^{2d} L_f \frac{1}{\sqrt[c]{r}} \quad \text{for } r \geq c d^c n^{cd}.$$

- *[Nie-Schweighofer 2007]* **Analysis of Putinar type bounds:**

  *There exists a constant $c' > 0$ such that for any polynomial $f$ of degree $d$:*

$$f_{\min} - \underline{f}_{sosP}^{(r)} \leq 6 d^3 n^{2d} L_f \frac{1}{\sqrt[c']{\log \frac{r}{c'}}} \quad \text{for } r \geq c' \exp((2d^2 n^d)^{c'}).$$

**Better results for some simple sets $K$?**

# Rate of convergence of SoS lower bounds

### Theorem

*Assume* $K \subseteq (-1,1)^n$. *For* $f = \sum_\alpha f_\alpha x^\alpha$, *set* $L_f = \max_\alpha |f_\alpha| \frac{\alpha!}{|\alpha|!}$.

- *[Schweighofer 2004]* **Analysis of Schmüdgen type bounds:**

  *There exists a constant* $c > 0$ *such that for any polynomial* $f$ *of degree* $d$:

$$f_{\min} - \underline{f}_{sosS}^{(r)} \leq c d^4 n^{2d} L_f \frac{1}{\sqrt[c]{r}} \quad \text{for } r \geq c d^c n^{cd}.$$

- *[Nie-Schweighofer 2007]* **Analysis of Putinar type bounds:**

  *There exists a constant* $c' > 0$ *such that for any polynomial* $f$ *of degree* $d$:

$$f_{\min} - \underline{f}_{sosP}^{(r)} \leq 6 d^3 n^{2d} L_f \frac{1}{\sqrt[c']{\log \frac{r}{c'}}} \quad \text{for } r \geq c' \exp((2d^2 n^d)^{c'}).$$

**Better results for some simple sets $K$?**

**Can choose $c = 1$ for Schmüdgen type bounds for simplex & cube.**

# Upper bounds for polynomial optimization

# Upper bounds for polynomial optimization

- For simple sets $K = \Delta_n$ or $Q_n$: minimize $f$ over the **rational grid points with given denominator $r$.**

# Upper bounds for polynomial optimization

- For simple sets $K = \Delta_n$ or $Q_n$: minimize $f$ over the **rational grid points with given denominator $r$.**
- For general compact sets $K$, use Lasserre idea:

$$f_{\min} = \min_{x \in K} f(x) = \min_{\mu \text{ probability measure on } K} \int_K f(x)\mu(dx)$$

# Upper bounds for polynomial optimization

- For simple sets $K = \Delta_n$ or $Q_n$: minimize $f$ over the **rational grid points with given denominator $r$.**
- For general compact sets $K$, use Lasserre idea:

$$f_{\min} = \min_{x \in K} f(x) = \min_{\mu \text{ probability measure on } K} \int_K f(x)\mu(dx)$$

and select suitable probability measures $\mu$ over $K$.

# Upper bounds for polynomial optimization

- For simple sets $K = \Delta_n$ or $Q_n$: minimize $f$ over the **rational grid points with given denominator $r$.**
- For general compact sets $K$, use Lasserre idea:

$$f_{\min} = \min_{x \in K} f(x) = \min_{\mu \text{ probability measure on } K} \int_K f(x)\mu(dx)$$

and select suitable probability measures $\mu$ over $K$.

## Theorem (Lasserre 2011)

*For $K$ compact one may use* **sum-of-squares density functions***:*

$$f_{\min} = \inf \int_K f(x)h(x)dx \quad s.t. \quad h \ SoS, \ \int_K h(x)dx = 1.$$

# Upper bounds for polynomial optimization

- For simple sets $K = \Delta_n$ or $Q_n$: minimize $f$ over the **rational grid points with given denominator $r$.**
- For general compact sets $K$, use Lasserre idea:

$$f_{\min} = \min_{x \in K} f(x) = \min_{\mu \text{ probability measure on } K} \int_K f(x)\mu(dx)$$

and select suitable probability measures $\mu$ over $K$.

## Theorem (Lasserre 2011)

*For $K$ compact one may use* **sum-of-squares density functions***:*

$$f_{\min} = \inf \int_K f(x)h(x)dx \quad s.t. \quad h \; SoS, \; \int_K h(x)dx = 1.$$

Bounding degree: $\deg(h) \leq 2r$, get upper bounds $\overline{f}^{(r)}_{\text{sos}}$ converging to $f_{\min}$.

What about the **rate of convergence**?

# Upper bounds for polynomial optimization

- For simple sets $K = \Delta_n$ or $Q_n$: minimize $f$ over the **rational grid points with given denominator $r$.**
- For general compact sets $K$, use Lasserre idea:

$$f_{\min} = \min_{x \in K} f(x) = \min_{\mu \text{ probability measure on } K} \int_K f(x)\mu(dx)$$

and select suitable probability measures $\mu$ over $K$.

## Theorem (Lasserre 2011)

*For $K$ compact one may use* **sum-of-squares density functions***:*

$$f_{\min} = \inf \int_K f(x)h(x)dx \quad s.t. \quad h \ SoS, \int_K h(x)dx = 1.$$

Bounding degree: $\deg(h) \leq 2r$, get upper bounds $\overline{f}^{(r)}_{\text{sos}}$ converging to $f_{\min}$.

What about the **rate of convergence**? **... in** $O(1/\sqrt{r})$

# This talk

**Analysis of the rate of convergence:**

- For the **simplex** $\Delta_n$ and the **cube** $Q_n$: regular grid upper bounds and the LP lower bounds.
- For $K$ **compact**: SoS-density upper bounds.
- For the **cube** $Q_n$: other upper bounds (using other density functions).

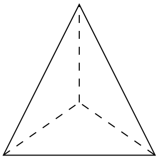
# Polynomial optimization over the simplex

**Upper bounds**: For $r \geq 1$

$$\Delta_n(r) = \{x \in \Delta_n : rx \in \mathbb{N}^n\}$$

$$f_{\min,\Delta_n(r)} = \min_{x \in \Delta_n(r)} f(x)$$

**Example:** $n = 3$:

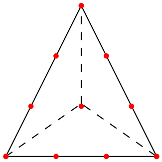


**Upper bounds**: For $r \geq 1$

$$\Delta_n(r) = \{x \in \Delta_n : rx \in \mathbb{N}^n\}$$

$$f_{\min,\Delta_n(r)} = \min_{x \in \Delta_n(r)} f(x)$$

**Example:** $n = 3$:



$r = 3$

Let $f = \sum_{|\beta|=d} f_\beta x^\beta$ homogeneous of degree $d$.

**Upper bounds**:

$$f_{\min} \leq f_{\min,\Delta_n(r)} = \min_{x \in \Delta_n(r)} f(x)$$

Let $f = \sum_{|\beta|=d} f_\beta x^\beta$ homogeneous of degree $d$.

**Upper bounds**:

$$f_{\min} \;\leq\; f_{\min,\Delta_n(r)} = \min_{x\in\Delta_n(r)} f(x) = \min_{\alpha\in\mathbb{N}^n:|\alpha|=r} f(\alpha/r)$$

Let $f = \sum_{|\beta|=d} f_\beta x^\beta$ homogeneous of degree $d$.

**Upper bounds**:

$$f_{\min} \;\leq\; f_{\min,\Delta_n(r)} = \min_{x\in\Delta_n(r)} f(x) = \min_{\alpha\in\mathbb{N}^n:|\alpha|=r} f(\alpha/r) \quad \Big(= \sum_{|\beta|=d} f_\beta \frac{\alpha^\beta}{r^d}\Big)$$

Let $f = \sum_{|\beta|=d} f_\beta x^\beta$ homogeneous of degree $d$.

**Upper bounds**:

$$f_{\min} \; \leq \; f_{\min,\Delta_n(r)} = \min_{x\in\Delta_n(r)} f(x) = \min_{\alpha\in\mathbb{N}^n:|\alpha|=r} f(\alpha/r) \quad \Big(= \sum_{|\beta|=d} f_\beta \frac{\alpha^\beta}{r^d}\Big)$$

**Lower bounds:**

$$f_{\min} \; \geq \; \underline{f}_{\rm lp}^{(r)} = \sup \; \lambda \;\; \text{s.t.} \;\; f(x) - \lambda = \underbrace{h(x)}_{\mathbb{R}_+[x]_r} + \underbrace{u(x)(1 - \sum_{i=1}^n x_i)}_{\mathbb{R}[x]_r}$$

Let $f = \sum_{|\beta|=d} f_\beta x^\beta$ homogeneous of degree $d$.

**Upper bounds**:

$$f_{\min} \le f_{\min,\Delta_n(r)} = \min_{x\in\Delta_n(r)} f(x) = \min_{\alpha\in\mathbb{N}^n:|\alpha|=r} f(\alpha/r) \quad \left(= \sum_{|\beta|=d} f_\beta \frac{\alpha^\beta}{r^d}\right)$$

**Lower bounds:**

$$f_{\min} \ge \underline{f}_{\mathrm{lp}}^{(r)} = \sup\ \lambda\ \text{ s.t. } f(x) - \lambda = \underbrace{h(x)}_{\mathbb{R}_+[x]_r} + \underbrace{u(x)(1 - \sum_{i=1}^n x_i)}_{\mathbb{R}[x]_r}$$

$$= \min_{\alpha\in\mathbb{N}^n:|\alpha|=r} \sum_{|\beta|=d} f_\beta \frac{\alpha^{\underline{\beta}}}{r^{\underline{d}}}.$$

where $r^{\underline{d}} = r(r-1)\cdots(r-d+1)$

Let $f = \sum_{|\beta|=d} f_\beta x^\beta$ homogeneous of degree $d$.

**Upper bounds**:

$$f_{\min} \;\le\; f_{\min,\Delta_n(r)} = \min_{x\in\Delta_n(r)} f(x) = \min_{\alpha\in\mathbb{N}^n:|\alpha|=r} f(\alpha/r) \quad \left(= \sum_{|\beta|=d} f_\beta \frac{\alpha^\beta}{r^d}\right)$$

**Lower bounds:**

$$f_{\min} \;\ge\; \underline{f}_{\mathrm{lp}}^{(r)} = \sup\ \lambda \ \text{ s.t. } f(x) - \lambda = \underbrace{h(x)}_{\mathbb{R}_+[x]_r} + \underbrace{u(x)(1 - \sum_{i=1}^n x_i)}_{\mathbb{R}[x]_r}$$

$$= \min_{\alpha\in\mathbb{N}^n:|\alpha|=r} \sum_{|\beta|=d} f_\beta \frac{\alpha^{\underline{\beta}}}{r^{\underline{d}}}.$$

where $r^{\underline{d}} = r(r-1)\cdots(r-d+1)$

$$\boxed{\underline{f}_{\mathrm{lp}}^{(r)} \le f_{\min} \le f_{\min,\Delta_n(r)}}$$

# Error analysis: $\underline{f}_{\rm lp}^{(r)} \le f_{\min} \le f_{\min,\Delta_n(r)}$

Theorem (De Klerk-L-Parrilo 2006)

- *For any polynomial $f$ of degree $d$:*

$$\boxed{f_{\min,\Delta_n(r)} - f_{\min},\ f_{\min} - \underline{f}_{\rm lp}^{(r)} \le \frac{C_d}{r}\,(f_{\max} - f_{\min}).}$$

# Error analysis: $\underline{f}_{\mathrm{lp}}^{(r)} \le f_{\min} \le f_{\min,\Delta_n(r)}$

## Theorem (De Klerk-L-Parrilo 2006)

- *For any polynomial $f$ of degree $d$:*

$$f_{\min,\Delta_n(r)} - f_{\min},\ f_{\min} - \underline{f}_{\mathrm{lp}}^{(r)} \le \frac{C_d}{r}\,(f_{\max} - f_{\min}).$$

- *Can compute the bounds via $|\Delta_n(r)| = O(n^r)$ function evaluations.*

# Error analysis: $\underline{f}_{\mathrm{lp}}^{(r)} \leq f_{\min} \leq f_{\min,\Delta_n(r)}$

## Theorem (De Klerk-L-Parrilo 2006)

- *For any polynomial $f$ of degree $d$:*

$$f_{\min,\Delta_n(r)} - f_{\min},\ f_{\min} - \underline{f}_{\mathrm{lp}}^{(r)} \leq \frac{C_d}{r}\,(f_{\max} - f_{\min}).$$

- *Can compute the bounds via $|\Delta_n(r)| = O(n^r)$ function evaluations.*
- *Get* **PTAS** *for minimizing a polynomial of fixed degree $d$ over $\Delta_n$.*

# Error analysis: $\underline{f}_{\mathrm{lp}}^{(r)} \leq f_{\min} \leq f_{\min,\Delta_n(r)}$

## Theorem (De Klerk-L-Parrilo 2006)

- *For any polynomial $f$ of degree $d$:*

$$f_{\min,\Delta_n(r)} - f_{\min},\ f_{\min} - \underline{f}_{\mathrm{lp}}^{(r)} \leq \frac{C_d}{r}\,(f_{\max} - f_{\min}).$$

- *Can compute the bounds via $|\Delta_n(r)| = O(n^r)$ function evaluations.*
- *Get* **PTAS** *for minimizing a polynomial of fixed degree $d$ over $\Delta_n$.*

Convergence rate in $1/r$ is **tight** for the **PTAS property:**

# Error analysis: $\underline{f}_{\rm lp}^{(r)} \le f_{\min} \le f_{\min,\Delta_n(r)}$

## Theorem (De Klerk-L-Parrilo 2006)

- *For any polynomial $f$ of degree $d$:*

$$f_{\min,\Delta_n(r)} - f_{\min},\ f_{\min} - \underline{f}_{\rm lp}^{(r)} \le \frac{C_d}{r}\ (f_{\max} - f_{\min}).$$

- *Can compute the bounds via $|\Delta_n(r)| = O(n^r)$ function evaluations.*
- *Get* **PTAS** *for minimizing a polynomial of fixed degree $d$ over $\Delta_n$.*

Convergence rate in $1/r$ is **tight** for the **PTAS property:** If $f = \sum_i x_i^2$ and $r = \frac{3n}{2}$ then: $\quad f_{\min,\Delta_n(r)} - f_{\min} = \frac{1}{6r-9}(f_{\max} - f_{\min})$

# Error analysis: $\underline{f}_{\mathrm{lp}}^{(r)} \leq f_{\min} \leq f_{\min,\Delta_n(r)}$

## Theorem (De Klerk-L-Parrilo 2006)

- *For any polynomial $f$ of degree $d$:*

$$f_{\min,\Delta_n(r)} - f_{\min}, \ f_{\min} - \underline{f}_{\mathrm{lp}}^{(r)} \leq \frac{C_d}{r} \ (f_{\max} - f_{\min}).$$

- *Can compute the bounds via $|\Delta_n(r)| = O(n^r)$ function evaluations.*
- *Get* **PTAS** *for minimizing a polynomial of fixed degree $d$ over $\Delta_n$.*

Convergence rate in $1/r$ is **tight** for the **PTAS property:** If $f = \sum_i x_i^2$ and $r = \frac{3n}{2}$ then: $f_{\min,\Delta_n(r)} - f_{\min} = \frac{1}{6r-9}(f_{\max} - f_{\min}) \ \leq \frac{n}{4r^2}$.

# Error analysis: $\underline{f}_{\rm lp}^{(r)} \le f_{\min} \le f_{\min,\Delta_n(r)}$

## Theorem (De Klerk-L-Parrilo 2006)

- *For any polynomial $f$ of degree $d$:*

$$f_{\min,\Delta_n(r)} - f_{\min},\ f_{\min} - \underline{f}_{\rm lp}^{(r)} \le \frac{C_d}{r}\,(f_{\max} - f_{\min}).$$

- *Can compute the bounds via $|\Delta_n(r)| = O(n^r)$ function evaluations.*
- *Get* **PTAS** *for minimizing a polynomial of fixed degree $d$ over $\Delta_n$.*

Convergence rate in $1/r$ is **tight** for the **PTAS property:** If $f = \sum_i x_i^2$ and $r = \frac{3n}{2}$ then: $\quad f_{\min,\Delta_n(r)} - f_{\min} = \frac{1}{6r-9}(f_{\max} - f_{\min}) \le \frac{n}{4r^2}$.

## Theorem

- *[De Klerk-L-Sun-Vera 2015]* $f_{\min,\Delta_n(r)} - f_{\min} \le \frac{C_f}{r^2}$.

# Error analysis: $\underline{f}_{\rm lp}^{(r)} \le f_{\min} \le f_{\min,\Delta_n(r)}$

## Theorem (De Klerk-L-Parrilo 2006)

- *For any polynomial $f$ of degree $d$:*

$$f_{\min,\Delta_n(r)} - f_{\min}, \; f_{\min} - \underline{f}_{\rm lp}^{(r)} \le \frac{C_d}{r}\,(f_{\max} - f_{\min}).$$

- *Can compute the bounds via $|\Delta_n(r)| = O(n^r)$ function evaluations.*
- *Get* **PTAS** *for minimizing a polynomial of fixed degree $d$ over $\Delta_n$.*

Convergence rate in $1/r$ is **tight** for the **PTAS property:** If $f = \sum_i x_i^2$ and $r = \frac{3n}{2}$ then: $\quad f_{\min,\Delta_n(r)} - f_{\min} = \frac{1}{6r-9}(f_{\max} - f_{\min}) \le \frac{n}{4r^2}$.

## Theorem

- *[De Klerk-L-Sun-Vera 2015]* $f_{\min,\Delta_n(r)} - f_{\min} \le \frac{C_f}{r^2}$.
- *[De Klerk-L-Sun 2015] May choose $C_f = mC_d(f_{\max} - f_{\min})$, if $f$ has a* **rational minimizer** *with denominator $m$.*

# Key idea for the $1/r$ convergence rate of $f_{\min,\Delta_n(r)}$

- Use the **Bernstein approximation** of $f$ of order $r$:

$$B_r(f)(x) = \sum_{\alpha \in \mathbb{N}^n : |\alpha| = r} f\left(\frac{\alpha}{r}\right) \frac{r!}{\alpha!} x^{\alpha}.$$

- So $B_r(f)(x)$ is the **average value** of $f$ over the grid points in $\Delta_n(r)$.

-

$$f_{\min,\Delta_n(r)} - f_{\min} \leq \min_{x \in \Delta_n} B_r(f)(x) - f_{\min}$$

# Key idea for the $1/r$ convergence rate of $f_{\min,\Delta_n(r)}$

- Use the **Bernstein approximation** of $f$ of order $r$:

$$B_r(f)(x) = \sum_{\alpha \in \mathbb{N}^n : |\alpha| = r} f\Big(\frac{\alpha}{r}\Big) \frac{r!}{\alpha!} x^{\alpha}.$$

- So $B_r(f)(x)$ is the **average value** of $f$ over the grid points in $\Delta_n(r)$.

- $$\boxed{f_{\min,\Delta_n(r)} - f_{\min} \leq \min_{x \in \Delta_n} B_r(f)(x) - f_{\min} \leq \max_{x \in \Delta_n} B_r(f)(x) - f(x)}$$

- Using properties of Bernstein approximations, one can show:

$$\max_{x \in \Delta_n} B_r(f)(x) - f(x) \leq \frac{C_d}{r}.$$

[De Klerk-L-Sun 2014]

# Polynomial optimization over the hypercube

**Upper bounds**:

$$f_{\min} \;\leq\; f_{\min,Q_n(r)} = \min_{x\in Q_n(r)} f(x) = \min_{\alpha\in\mathbb{N}^n:\alpha_i\leq r} f(\alpha/r)$$

**Upper bounds**:

$$f_{\min} \ \leq \ f_{\min, Q_n(r)} = \min_{x \in Q_n(r)} f(x) = \min_{\alpha \in \mathbb{N}^n : \alpha_i \leq r} f(\alpha/r)$$

**Lower bounds:**

$$f_{\min} \ \geq \ \underline{f}_{\mathrm{lp}}^{(r)} = \sup \ \lambda \ \text{ s.t. } \ f(x) - \lambda = \sum_{\alpha, \gamma \in \mathbb{N}^n} \underbrace{\underbrace{\lambda_{\alpha,\gamma}}_{\geq 0} x^{\alpha} (1-x)^{\gamma}}_{\text{degree } \leq r}$$

**Upper bounds**:

$$f_{\min} \;\leq\; f_{\min,Q_n(r)} = \min_{x\in Q_n(r)} f(x) = \min_{\alpha\in\mathbb{N}^n:\alpha_i\leq r} f(\alpha/r)$$

**Lower bounds:**

$$f_{\min} \;\geq\; \underline{f}_{\mathrm{lp}}^{(r)} = \sup\ \lambda\ \text{ s.t. } f(x)-\lambda = \sum_{\alpha,\gamma\in\mathbb{N}^n} \underbrace{\underbrace{\lambda_{\alpha,\gamma}}_{\geq 0} x^{\alpha}(1-x)^{\gamma}}_{\text{degree } \leq r}$$

$$\boxed{\underline{f}_{\mathrm{lp}}^{(r)} \;\leq\; f_{\min} \;\leq\; f_{\min,Q_r(d)}}$$

**Upper bounds**:

$$f_{\min} \;\leq\; f_{\min,Q_n(r)} = \min_{x\in Q_n(r)} f(x) = \min_{\alpha\in\mathbb{N}^n:\alpha_i\leq r} f(\alpha/r)$$

**Lower bounds:**

$$f_{\min} \;\geq\; \underline{f}_{\rm lp}^{(r)} = \sup\ \lambda \ \text{ s.t. } f(x)-\lambda = \sum_{\alpha,\gamma\in\mathbb{N}^n} \underbrace{\underbrace{\lambda_{\alpha,\gamma}}_{\geq 0} x^\alpha (1-x)^\gamma}_{\text{degree } \leq r}$$

$$\boxed{\underline{f}_{\rm lp}^{(r)} \;\leq\; f_{\min} \;\leq\; f_{\min,Q_r(d)}}$$

## Theorem

- *[De Klerk-L 2010]* *For $r \geq dn$* $\quad f_{\min} - \underline{f}_{\rm lp}^{(r)} \;\leq\; \binom{d+1}{3}\frac{n^{d+1}L_f}{r}.$

**Upper bounds**:

$$f_{\min} \;\le\; f_{\min,Q_n(r)} = \min_{x\in Q_n(r)} f(x) = \min_{\alpha\in\mathbb{N}^n:\alpha_i\le r} f(\alpha/r)$$

**Lower bounds:**

$$f_{\min} \;\ge\; \underline{f}_{\rm lp}^{(r)} = \sup\ \lambda\ \text{ s.t. } f(x)-\lambda = \sum_{\alpha,\gamma\in\mathbb{N}^n} \underbrace{\underbrace{\lambda_{\alpha,\gamma}}_{\ge 0} x^\alpha(1-x)^\gamma}_{\text{degree } \le r}$$

$$\boxed{\underline{f}_{\rm lp}^{(r)} \;\le\; f_{\min} \;\le\; f_{\min,Q_r(d)}}$$

## Theorem

- *[De Klerk-L 2010]* *For* $r\ge dn$ $\quad f_{\min}-\underline{f}_{\rm lp}^{(r)} \;\le\; \binom{d+1}{3}\frac{n^{d+1}L_f}{r}.$
- *[De Klerk-Lasserre-L-Sun 2015]* *For* $r\ge 1$ $\quad f_{\min,Q_n(r)}-f_{\min} \;\le\; \frac{C_f}{r^2}.$

**Upper bounds**:

$$f_{\min} \;\leq\; f_{\min,Q_n(r)} = \min_{x\in Q_n(r)} f(x) = \min_{\alpha\in\mathbb{N}^n:\alpha_i\leq r} f(\alpha/r)$$

**Lower bounds:**

$$f_{\min} \;\geq\; \underline{f}_{\rm lp}^{(r)} = \sup\ \lambda\ \text{ s.t. } f(x)-\lambda = \sum_{\alpha,\gamma\in\mathbb{N}^n} \underbrace{\underbrace{\lambda_{\alpha,\gamma}}_{\geq 0} x^\alpha(1-x)^\gamma}_{\text{degree } \leq r}$$

$$\boxed{\underline{f}_{\rm lp}^{(r)} \;\leq\; f_{\min} \;\leq\; f_{\min,Q_r(d)}}$$

## Theorem

- *[De Klerk-L 2010]* *For* $r\geq dn$ $\quad f_{\min}-\underline{f}_{\rm lp}^{(r)} \;\leq\; \binom{d+1}{3}\frac{n^{d+1}L_f}{r}$.
- *[De Klerk-Lasserre-L-Sun 2015]* *For* $r\geq 1$ $\quad f_{\min,Q_n(r)}-f_{\min} \;\leq\; \frac{C_f}{r^2}$.

**but** $|Q_n(r)|=(r+1)^n$ **...**

# Measure-based upper bounds

# Upper bounds using SoS densities

Lasserre (2011) proved: For $K$ compact

$$f_{\min} = \min_{x \in K} f(x) = \min_{\mu \text{ probability measure on } K} \int_K f(x) d\mu$$

$$= \inf \int_K f(x) h(x) dx \quad \text{s.t.} \quad h \text{ SoS}, \ \int_K h(x) dx = 1.$$

# Upper bounds using SoS densities

Lasserre (2011) proved: For $K$ compact

$$f_{\min} = \min_{x \in K} f(x) = \min_{\mu \text{ probability measure on } K} \int_K f(x) d\mu$$

$$= \inf \int_K f(x) h(x) dx \quad \text{s.t.} \quad h \text{ SoS}, \ \int_K h(x) dx = 1.$$

$$\overline{f}^{(r)}_{\text{sos}} = \min \int_K f(x) h(x) dx \quad \text{s.t.} \quad h \text{ SoS}, \ \deg(h) \leq r, \ \int_K h(x) dx = 1.$$

# Upper bounds using SoS densities

Lasserre (2011) proved: For $K$ compact

$$f_{\min} = \min_{x \in K} f(x) = \min_{\mu \text{ probability measure on } K} \int_K f(x) d\mu$$

$$= \inf \int_K f(x) h(x) dx \quad \text{s.t.} \quad h \text{ SoS}, \ \int_K h(x) dx = 1.$$

$$\overline{f}^{(r)}_{\text{sos}} = \min \int_K f(x) h(x) dx \quad \text{s.t.} \quad h \text{ SoS}, \ \deg(h) \leq r, \ \int_K h(x) dx = 1.$$

- Compute $\overline{f}^{(r)}_{\text{sos}}$ as **generalized eigenvalue problem**
- Need **moments** $\int_K x^\alpha dx$:

# Upper bounds using SoS densities

Lasserre (2011) proved: For $K$ compact

$$f_{\min} = \min_{x \in K} f(x) = \min_{\mu \text{ probability measure on } K} \int_K f(x) d\mu$$

$$= \inf \int_K f(x)h(x)dx \quad \text{s.t.} \quad h \text{ SoS}, \int_K h(x)dx = 1.$$

$$\overline{f}^{(r)}_{\text{sos}} = \min \int_K f(x)h(x)dx \quad \text{s.t.} \quad h \text{ SoS}, \ \deg(h) \leq r, \ \int_K h(x)dx = 1.$$

- Compute $\overline{f}^{(r)}_{\text{sos}}$ as **generalized eigenvalue problem**
- Need **moments** $\int_K x^\alpha dx$: known for simplex, hypercube, sphere,...

# Upper bounds using SoS densities

Lasserre (2011) proved: For $K$ compact

$$f_{\min} = \min_{x \in K} f(x) = \min_{\mu \text{ probability measure on } K} \int_K f(x) d\mu$$

$$= \inf \int_K f(x) h(x) dx \quad \text{s.t.} \quad h \text{ SoS}, \ \int_K h(x) dx = 1.$$

$$\overline{f}^{(r)}_{\text{sos}} = \min \int_K f(x) h(x) dx \quad \text{s.t.} \quad h \text{ SoS}, \ \deg(h) \leq r, \ \int_K h(x) dx = 1.$$

- Compute $\overline{f}^{(r)}_{\text{sos}}$ as **generalized eigenvalue problem**
- Need **moments** $\int_K x^\alpha dx$: known for simplex, hypercube, sphere,...

## Theorem (De Klerk-L-Sun 2015)

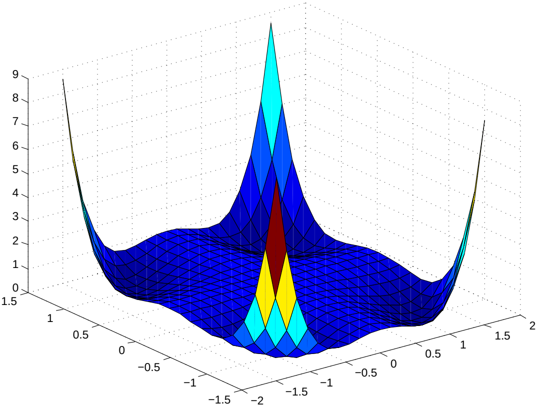
*Assume $K$ is compact and 'nice' (e.g. convex body) and $f$ has Lipschitz constant $M_f$. There exist constants $C_K > 0$ and $r_K \geq 1$ such that*

$$\overline{f}^{(r)}_{sos} - f_{\min} \leq \frac{C_K M_f}{\sqrt{r}} \qquad \forall r \geq r_K.$$

# Example: Motzkin polynomial on $K = [-2, 2]^2$

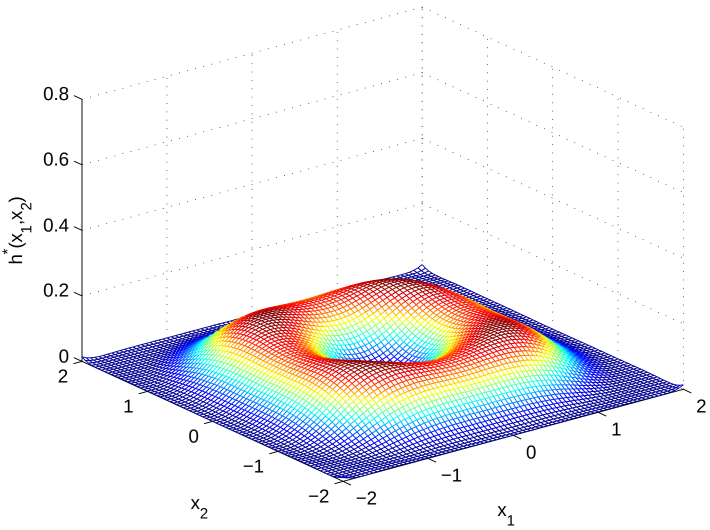
$$f(x_1, x_2) = x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2 + 1$$

Global minimizers: $(-1, -1)$, $(-1, 1)$, $(1, -1)$, $(1, 1)$.



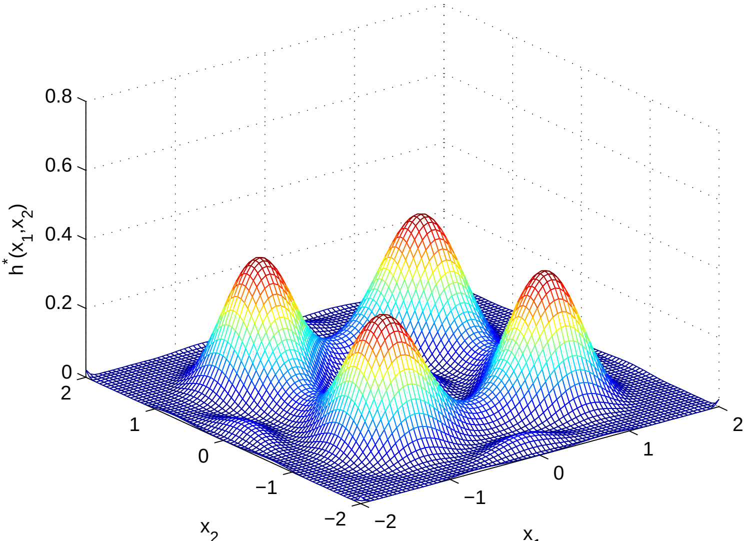
# Example: Motzkin polynomial on $[-2,2]^2$ (ctd.)

Optimal SoS density $h$ of degree 12



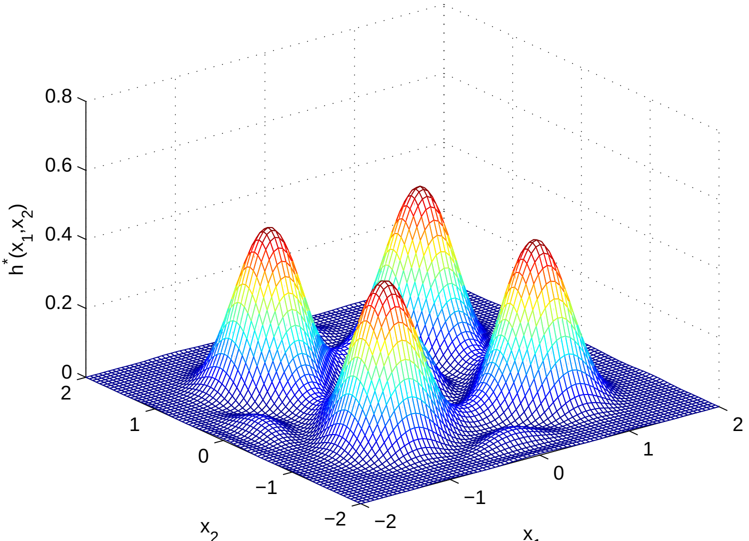
# Example: Motzkin polynomial on $[-2, 2]^2$ (ctd.)

Optimal SoS density $h$ of degree 16



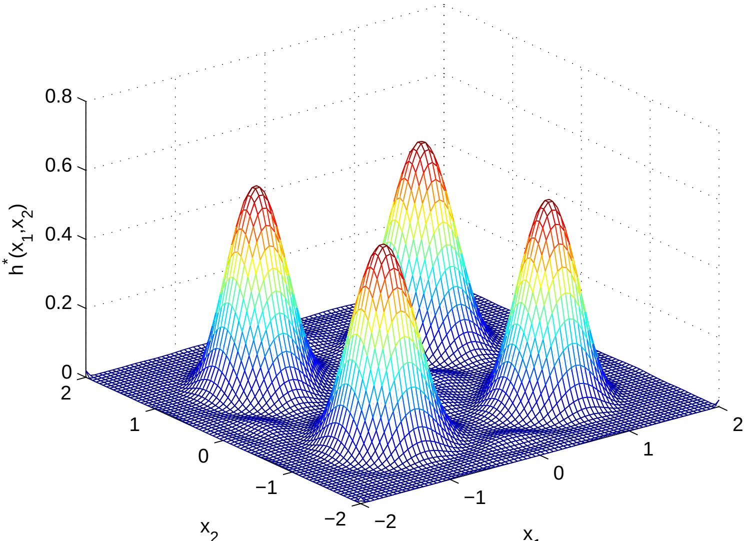
# Example: Motzkin polynomial on $[-2, 2]^2$ (ctd.)

Optimal SoS density *h* of degree 20



# Example: Motzkin polynomial on $[-2, 2]^2$ (ctd.)

Optimal SoS density $h$ of degree 24



# Convergence analysis: sketch of proof

- Let **a** be a global minimizer of $f$ in $K$.
- **Want:** A SoS polynomial $h$ of degree $2r$ which "looks like" the delta function at **a**.

# Convergence analysis: sketch of proof

- Let **a** be a global minimizer of $f$ in $K$.
- **Want:** A SoS polynomial $h$ of degree $2r$ which "looks like" the delta function at **a**.
- **Idea:** Approximate the **Gaussian distribution** centered at **a** with (suitable) variance $\sigma$:

$$G_a(x) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\|x - \mathbf{a}\|^2}{2\sigma^2}\right)$$

# Convergence analysis: sketch of proof

- Let **a** be a global minimizer of $f$ in $K$.
- **Want:** A SoS polynomial $h$ of degree $2r$ which "looks like" the delta function at **a**.
- **Idea:** Approximate the **Gaussian distribution** centered at **a** with (suitable) variance $\sigma$:

$$G_a(x) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\|x-\mathbf{a}\|^2}{2\sigma^2}\right)$$

- by its **truncation of its Taylor expansion at degree** $2r$:

$$H_{r,\mathbf{a}}(x) = \frac{1}{(2\pi\sigma^2)^{n/2}} \sum_{k=0}^{2r} \frac{1}{k!}\left(-\frac{\|x-\mathbf{a}\|^2}{2\sigma^2}\right)^k .$$

# Convergence analysis: sketch of proof

- Let $\mathbf{a}$ be a global minimizer of $f$ in $K$.
- **Want:** A SoS polynomial $h$ of degree $2r$ which "looks like" the delta function at $\mathbf{a}$.
- **Idea:** Approximate the **Gaussian distribution** centered at $\mathbf{a}$ with (suitable) variance $\sigma$:

$$G_a(x) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\|x-\mathbf{a}\|^2}{2\sigma^2}\right)$$

- by its **truncation of its Taylor expansion at degree** $2r$:

$$H_{r,\mathbf{a}}(x) = \frac{1}{(2\pi\sigma^2)^{n/2}} \sum_{k=0}^{2r} \frac{1}{k!} \left(-\frac{\|x-\mathbf{a}\|^2}{2\sigma^2}\right)^k .$$

- **Fact:** $H_{r,\mathbf{a}}(x)$ **is a SoS polynomial of degree** $4r$,

# Convergence analysis: sketch of proof

- Let $\mathbf{a}$ be a global minimizer of $f$ in $K$.
- **Want:** A SoS polynomial $h$ of degree $2r$ which "looks like" the delta function at $\mathbf{a}$.
- **Idea:** Approximate the **Gaussian distribution** centered at $\mathbf{a}$ with (suitable) variance $\sigma$:

$$G_a(x) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\|x-\mathbf{a}\|^2}{2\sigma^2}\right)$$

- by its **truncation of its Taylor expansion at degree** $2r$:

$$H_{r,\mathbf{a}}(x) = \frac{1}{(2\pi\sigma^2)^{n/2}} \sum_{k=0}^{2r} \frac{1}{k!} \left(-\frac{\|x-\mathbf{a}\|^2}{2\sigma^2}\right)^k .$$

- **Fact:** $H_{r,\mathbf{a}}(x)$ **is a SoS polynomial of degree** $4r$, because the univariate polynomial $p_r(t) = \sum_{k=0}^{2r} \frac{(-t)^k}{k!}$ is nonnegative and thus SoS:

$$0 \leq p_r(t) - \exp(-t)$$

# Convergence analysis: sketch of proof

- Let $\mathbf{a}$ be a global minimizer of $f$ in $K$.
- **Want:** A SoS polynomial $h$ of degree $2r$ which "looks like" the delta function at $\mathbf{a}$.
- **Idea:** Approximate the **Gaussian distribution** centered at $\mathbf{a}$ with (suitable) variance $\sigma$:

$$G_a(x) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\|x-\mathbf{a}\|^2}{2\sigma^2}\right)$$

- by its **truncation of its Taylor expansion at degree** $2r$:

$$H_{r,\mathbf{a}}(x) = \frac{1}{(2\pi\sigma^2)^{n/2}} \sum_{k=0}^{2r} \frac{1}{k!}\left(-\frac{\|x-\mathbf{a}\|^2}{2\sigma^2}\right)^k.$$

- **Fact:** $H_{r,\mathbf{a}}(x)$ **is a SoS polynomial of degree** $4r$, because the univariate polynomial $p_r(t) = \sum_{k=0}^{2r} \frac{(-t)^k}{k!}$ is nonnegative and thus SoS:

$$0 \leq p_r(t) - \exp(-t) \;\leq \frac{t^{2r+1}}{(2r+1)!}.$$

# Convergence analysis

- Define the **normalizing constant:** $c^r_{\mathbf{a}} \int_K H_{r,\mathbf{a}}(x) dx = 1.$

# Convergence analysis

- Define the **normalizing constant:** $c_{\mathbf{a}}^r \int_K H_{r,\mathbf{a}}(x)dx = 1$.
- $K$ is **nice at a** if there exist constants $\eta_K$ and $\epsilon_K$ such that

$$\text{Vol}(B_\epsilon(\mathbf{a}) \cap K) \geq \eta_K \text{Vol}(B_\epsilon(\mathbf{a})) \quad \forall 0 < \epsilon \leq \epsilon_K.$$

Will be used to control the constant $c_{\mathbf{a}}^r$.

# Convergence analysis

- Define the **normalizing constant:** $c_{\mathbf{a}}^r \int_K H_{r,\mathbf{a}}(x)dx = 1$.
- $K$ is **nice at a** if there exist constants $\eta_K$ and $\epsilon_K$ such that

$$\text{Vol}(B_\epsilon(\mathbf{a}) \cap K) \geq \eta_K \text{Vol}(B_\epsilon(\mathbf{a})) \quad \forall 0 < \epsilon \leq \epsilon_K.$$

Will be used to control the constant $c_{\mathbf{a}}^r$.

- The analysis will work when selecting: $\boxed{\sigma \sim \frac{1}{\sqrt{2r+1}}}$

# Convergence analysis

- Define the **normalizing constant:** $c_{\mathbf{a}}^r \int_K H_{r,\mathbf{a}}(x)dx = 1$.
- $K$ is **nice at a** if there exist constants $\eta_K$ and $\epsilon_K$ such that

$$\mathrm{Vol}(B_\epsilon(\mathbf{a}) \cap K) \geq \eta_K \mathrm{Vol}(B_\epsilon(\mathbf{a})) \quad \forall 0 < \epsilon \leq \epsilon_K.$$

Will be used to control the constant $c_{\mathbf{a}}^r$.

- The analysis will work when selecting: $\sigma \sim \frac{1}{\sqrt{2r+1}}$
- **Main result:** If $f$ has Lipschitz constant $M_f$ and $K$ is **nice at a** then

$$\int_K f(x) c_{\mathbf{a}}^r H_{r,\mathbf{a}}(x)dx - f_{\min} \leq \frac{C_K M_f}{\sqrt{r}}.$$

## Other measure-based upper bounds for the hypercube:

- Handelman type densities

- Schmüdgen type densities

# Using Handelman type densities for $K = [0,1]^n$

For $K = [0,1]^n$, consider the upper bound:

$$\overline{f}^{(r)}_{\mathrm{lp}} = \min \int_K f(x)h(x)dx \quad \text{s.t. } h(x) = \sum_{\alpha,\beta \in \mathbb{N}^n} \underbrace{\underbrace{\lambda_{\alpha,\beta}}_{\geq 0} x^\alpha (1-x)^\beta}_{\text{degree } = r}, \int_K h(x)dx = 1.$$

# Using Handelman type densities for $K = [0,1]^n$

For $K = [0,1]^n$, consider the upper bound:

$$\overline{f}_{\mathrm{lp}}^{(r)} = \min \int_K f(x)h(x)dx \ \text{ s.t. } h(x) = \underbrace{\sum_{\alpha,\beta\in\mathbb{N}^n} \underbrace{\lambda_{\alpha,\beta}}_{\geq 0} x^\alpha (1-x)^\beta}_{\text{degree } = r}, \int_K h(x)dx = 1.$$

## Theorem (De Klerk-Lasserre-L-Sun 2015)

- $\overline{f}_{\mathrm{lp}}^{(r)}$ *needs* $O(n^r)$ **elementary computations:**

$$\overline{f}_{\mathrm{lp}}^{(r)} = \min_{|\alpha+\beta|=r} \frac{\int_K f(x)x^\alpha(1-x)^\beta dx}{\int_K x^\alpha (1-x)^\beta dx}$$

# Using Handelman type densities for $K = [0,1]^n$

For $K = [0,1]^n$, consider the upper bound:

$$\overline{f}^{(r)}_{\rm lp} = \min \int_K f(x)h(x)dx \;\text{ s.t. } h(x) = \sum_{\alpha,\beta\in\mathbb{N}^n} \underbrace{\underbrace{\lambda_{\alpha,\beta}}_{\geq 0} x^\alpha(1-x)^\beta}_{\text{degree } = r}, \int_K h(x)dx = 1.$$

## Theorem (De Klerk-Lasserre-L-Sun 2015)

- $\overline{f}^{(r)}_{\rm lp}$ *needs* $O(n^r)$ **elementary computations:**

$$\overline{f}^{(r)}_{\rm lp} = \min_{|\alpha+\beta|=r} \frac{\int_K f(x)x^\alpha(1-x)^\beta dx}{\int_K x^\alpha(1-x)^\beta dx}$$

- **Convergence rate:** *For* $r \geq r_0$

$$\overline{f}^{(r)}_{\rm lp} - f_{\min} \leq \frac{C_f}{\sqrt{r}},$$

# Using Handelman type densities for $K = [0,1]^n$

For $K = [0,1]^n$, consider the upper bound:

$$\overline{f}_{\mathrm{lp}}^{(r)} = \min \int_K f(x)h(x)dx \quad \text{s.t.} \quad h(x) = \sum_{\alpha,\beta\in\mathbb{N}^n} \underbrace{\underbrace{\lambda_{\alpha,\beta}}_{\geq 0} x^\alpha (1-x)^\beta}_{\text{degree } = r}, \int_K h(x)dx = 1.$$

## Theorem (De Klerk-Lasserre-L-Sun 2015)

- $\overline{f}_{\mathrm{lp}}^{(r)}$ *needs* $O(n^r)$ **elementary computations:**

$$\overline{f}_{\mathrm{lp}}^{(r)} = \min_{|\alpha+\beta|=r} \frac{\int_K f(x)x^\alpha(1-x)^\beta dx}{\int_K x^\alpha (1-x)^\beta dx}$$

- **Convergence rate:** *For* $r \geq r_0$

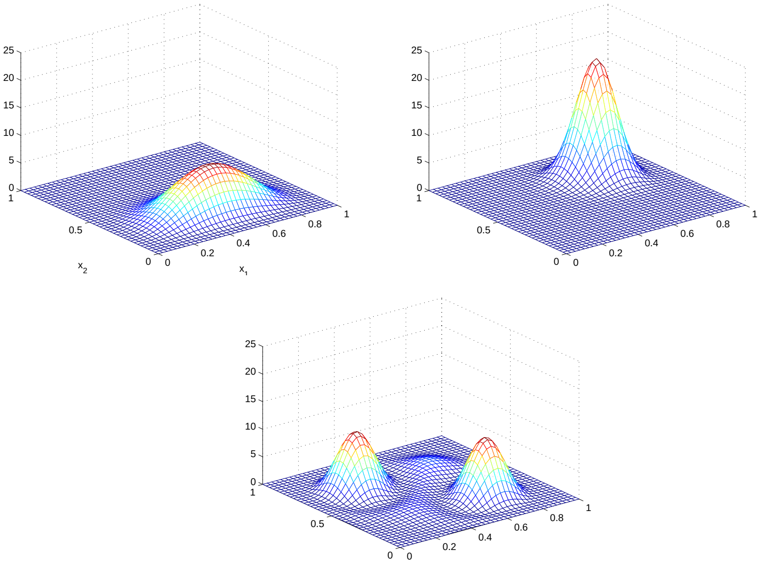
$$\overline{f}_{\mathrm{lp}}^{(r)} - f_{\min} \leq \frac{C_f}{\sqrt{r}}, \qquad \leq \frac{C'_f}{r} \quad \textit{if } f \textit{ has a } \textbf{rational minimizer}.$$

$\rightsquigarrow$ Link to the **beta distribution**

Motzkin polynomial:

$$f(x,y) = (4x-2)^4(4y-2)^2 + (4x-2)^2(4y-2)^4 - 3(4x-2)^2(4y-2)^2 + 1$$

over $K = [0,1]^2$: Handelman-type densities (deg 24, 50) & SOS (deg 24)



# Using Schmüdgen type densities for $K = [-1,1]^n$

For $K = [-1,1]^n$ consider the upper bound:

$$\overline{f}^{(r)}_{\text{sosS}} = \min \int_K f(x)h(x)d\mu_n \quad \text{s.t. } h(x) = \underbrace{\sum_{I\subseteq[n]} \underbrace{\sigma_I}_{\text{SoS}}(1-x^2)^I}_{\text{degree } \leq r}, \quad \int_K h(x)d\mu_n = 1$$

# Using Schmüdgen type densities for $K = [-1,1]^n$

For $K = [-1,1]^n$ consider the upper bound:

$$\overline{f}^{(r)}_{\text{sosS}} = \min \int_K f(x)h(x)d\mu_n \ \text{ s.t. } h(x) = \underbrace{\sum_{I \subseteq [n]} \underbrace{\sigma_I}_{\text{SoS}} (1-x^2)^I}_{\text{degree } \leq r}, \quad \int_K h(x)d\mu_n = 1$$

where $d\mu_n = \left(\prod_{i=1}^n \pi\sqrt{1-x_i}\right)^{-1} dx.$

# Using Schmüdgen type densities for $K = [-1,1]^n$

For $K = [-1,1]^n$ consider the upper bound:

$$\overline{f}^{(r)}_{\text{sosS}} = \min \int_K f(x)h(x)d\mu_n \ \text{ s.t. } h(x) = \underbrace{\sum_{I\subseteq[n]} \underbrace{\sigma_I}_{\text{SoS}} (1-x^2)^I}_{\text{degree } \leq r}, \quad \int_K h(x)d\mu_n = 1$$

where $d\mu_n = \left(\prod_{i=1}^n \pi\sqrt{1-x_i}\right)^{-1} dx.$

## Theorem (De Klerk-Hess-L 2016)

- $\overline{f}^{(r)}_{sosS}$ *can be computed as* **generalized eigenvalue** *problem:*

$$\overline{f}^{(r)}_{sosS} = \min_{I\subseteq[n]} \left\{ \max_{\lambda} \lambda \ \ s.t. \ A^{(I)} - \lambda B^{(I)} \succeq 0 \right\}.$$

# Using Schmüdgen type densities for $K = [-1,1]^n$

For $K = [-1,1]^n$ consider the upper bound:

$$\overline{f}^{(r)}_{\text{sosS}} = \min \int_K f(x)h(x)d\mu_n \text{ s.t. } h(x) = \underbrace{\sum_{I \subseteq [n]} \underbrace{\sigma_I}_{\text{SoS}} (1-x^2)^I}_{\text{degree } \leq r}, \quad \int_K h(x)d\mu_n = 1$$

where $d\mu_n = \left(\prod_{i=1}^n \pi\sqrt{1-x_i}\right)^{-1} dx.$

## Theorem (De Klerk-Hess-L 2016)

- $\overline{f}^{(r)}_{sosS}$ *can be computed as* **generalized eigenvalue** *problem:*

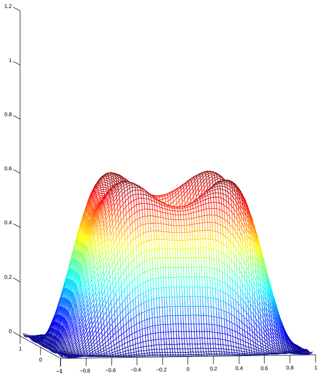
$$\overline{f}^{(r)}_{sosS} = \min_{I \subseteq [n]} \left\{ \max_{\lambda} \lambda \ \ s.t. \ A^{(I)} - \lambda B^{(I)} \succeq 0 \right\}.$$

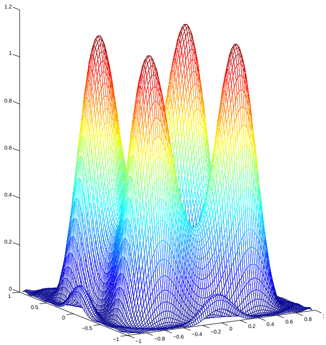
- **Convergence rate**: *For* $r \geq r_0$

$$\overline{f}^{(r)}_{sosS} - f_{\min} \leq \frac{C_f}{r^2}.$$

Motzkin polynomial: $f(x,y) = 64(x^4y^2 + x^2y^4) - 48x^2y^2 + 1$
over $K = [-1,1]^2$

Optimal Schmüdgen type densities of degree 12, 16:





# Key idea: use the Polynomial Kernel Method

Following [Weisse-Alvermann-Fehske 2006] use **approximations** of the delta function at **a** by taking its convolution with the **Jackson kernel:**

# Key idea: use the Polynomial Kernel Method

Following [Weisse-Alvermann-Fehske 2006] use **approximations** of the delta function at **a** by taking its convolution with the **Jackson kernel:**

$$\delta_{\mathbf{a}}^{(r)}(x) = 1 + 2\sum_{k=1}^{r} g_k^{(r)} T_k(\mathbf{a}) T_k(x)$$

# Key idea: use the Polynomial Kernel Method

Following [Weisse-Alvermann-Fehske 2006] use **approximations** of the delta function at **a** by taking its convolution with the **Jackson kernel:**

$$\delta_{\mathbf{a}}^{(r)}(x) = 1 + 2\sum_{k=1}^{r} g_k^{(r)}\, T_k(\mathbf{a})\, T_k(x)$$

- $T_k(x) = \cos(k \arccos x)$: **Tchebyshev polynomials**

# Key idea: use the Polynomial Kernel Method

Following [Weisse-Alvermann-Fehske 2006] use **approximations** of the delta function at **a** by taking its convolution with the **Jackson kernel:**

$$\delta_{\mathbf{a}}^{(r)}(x) = 1 + 2\sum_{k=1}^{r} g_k^{(r)}\, T_k(\mathbf{a})\, T_k(x)$$

- $T_k(x) = \cos(k \arccos x)$: **Tchebyshev polynomials** $\rightsquigarrow$ orthogonal basis of $\mathbb{R}[x]$ for inner product: $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)d\mu_1$.

# Key idea: use the Polynomial Kernel Method

Following [Weisse-Alvermann-Fehske 2006] use **approximations** of the delta function at **a** by taking its convolution with the **Jackson kernel:**

$$\delta_{\mathbf{a}}^{(r)}(x) = 1 + 2\sum_{k=1}^{r} g_k^{(r)} T_k(\mathbf{a}) T_k(x)$$

- $T_k(x) = \cos(k \arccos x)$: **Tchebyshev polynomials** $\rightsquigarrow$ orthogonal basis of $\mathbb{R}[x]$ for inner product: $\langle f, g \rangle = \int_{-1}^{1} f(x) g(x) d\mu_1$.
- $\vartheta_r = \frac{\pi}{r+2}$, $g_k^{(r)} = \frac{1}{r+2}((r+2-k)\cos(k\vartheta_r) + \frac{\sin(k\vartheta_r)}{\sin\vartheta_r}\cos\vartheta_r)$,

# Key idea: use the Polynomial Kernel Method

Following [Weisse-Alvermann-Fehske 2006] use **approximations** of the delta function at **a** by taking its convolution with the **Jackson kernel:**

$$\delta_{\mathbf{a}}^{(r)}(x) = 1 + 2\sum_{k=1}^{r} g_k^{(r)}\, T_k(\mathbf{a})\, T_k(x)$$

- $T_k(x) = \cos(k \arccos x)$: **Tchebyshev polynomials** $\rightsquigarrow$ orthogonal basis of $\mathbb{R}[x]$ for inner product: $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)d\mu_1$.
- $\vartheta_r = \frac{\pi}{r+2}$, $g_k^{(r)} = \frac{1}{r+2}((r+2-k)\cos(k\vartheta_r) + \frac{\sin(k\vartheta_r)}{\sin\vartheta_r}\cos\vartheta_r)$, so that:
- $\delta_{\mathbf{a}}^{(r)}(x)$ is a polynomial (of degree $r$) **density** for $\mu_1$ on $[-1, 1]$.

# Key idea: use the Polynomial Kernel Method

Following [Weisse-Alvermann-Fehske 2006] use **approximations** of the delta function at **a** by taking its convolution with the **Jackson kernel:**

$$\delta_{\mathbf{a}}^{(r)}(x) = 1 + 2\sum_{k=1}^{r} g_k^{(r)}\, T_k(\mathbf{a})\, T_k(x)$$

- $T_k(x) = \cos(k \arccos x)$: **Tchebyshev polynomials** $\rightsquigarrow$ orthogonal basis of $\mathbb{R}[x]$ for inner product: $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)d\mu_1$.
- $\vartheta_r = \frac{\pi}{r+2}$, $g_k^{(r)} = \frac{1}{r+2}((r+2-k)\cos(k\vartheta_r) + \frac{\sin(k\vartheta_r)}{\sin\vartheta_r}\cos\vartheta_r)$, so that:
- $\delta_{\mathbf{a}}^{(r)}(x)$ is a polynomial (of degree $r$) **density** for $\mu_1$ on $[-1,1]$.

  **Hence:** $\delta_{\mathbf{a}}^{(r)}(x) = \sigma_0 + \sigma_1(1-x^2)$, with $\sigma_0, \sigma_1$ SoS of degree $\leq r$.

# Key idea: use the Polynomial Kernel Method

Following [Weisse-Alvermann-Fehske 2006] use **approximations** of the delta function at **a** by taking its convolution with the **Jackson kernel:**

$$\delta_{\mathbf{a}}^{(r)}(x) = 1 + 2\sum_{k=1}^{r} g_k^{(r)}\, T_k(\mathbf{a})\, T_k(x)$$

- $T_k(x) = \cos(k \arccos x)$: **Tchebyshev polynomials** $\rightsquigarrow$ orthogonal basis of $\mathbb{R}[x]$ for inner product: $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)d\mu_1$.
- $\vartheta_r = \frac{\pi}{r+2}$, $g_k^{(r)} = \frac{1}{r+2}((r+2-k)\cos(k\vartheta_r) + \frac{\sin(k\vartheta_r)}{\sin\vartheta_r}\cos\vartheta_r)$, so that:
- $\delta_{\mathbf{a}}^{(r)}(x)$ is a polynomial (of degree $r$) **density** for $\mu_1$ on $[-1, 1]$.

  **Hence:** $\delta_{\mathbf{a}}^{(r)}(x) = \sigma_0 + \sigma_1(1 - x^2)$, with $\sigma_0, \sigma_1$ SoS of degree $\leq r$.
- For $k = 1$ $\quad |g_1^{(r)} - 1| = O(1/r^2)$.

# Key idea: use the Polynomial Kernel Method

Following [Weisse-Alvermann-Fehske 2006] use **approximations** of the delta function at **a** by taking its convolution with the **Jackson kernel:**

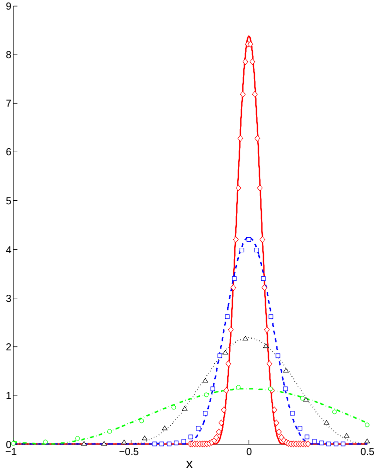
$$\delta_{\mathbf{a}}^{(r)}(x) = 1 + 2\sum_{k=1}^{r} g_k^{(r)}\, T_k(\mathbf{a})\, T_k(x)$$

- $T_k(x) = \cos(k \arccos x)$: **Tchebyshev polynomials** $\rightsquigarrow$ orthogonal basis of $\mathbb{R}[x]$ for inner product: $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)d\mu_1$.
- $\vartheta_r = \frac{\pi}{r+2}$, $g_k^{(r)} = \frac{1}{r+2}((r+2-k)\cos(k\vartheta_r) + \frac{\sin(k\vartheta_r)}{\sin\vartheta_r}\cos\vartheta_r)$, so that:
- $\delta_{\mathbf{a}}^{(r)}(x)$ is a polynomial (of degree $r$) **density** for $\mu_1$ on $[-1,1]$.

  **Hence:** $\delta_{\mathbf{a}}^{(r)}(x) = \sigma_0 + \sigma_1(1-x^2)$, with $\sigma_0, \sigma_1$ SoS of degree $\leq r$.
- For $k = 1$ $\quad |g_1^{(r)} - 1| = O(1/r^2)$.

$\rightsquigarrow$ rate of convergence in $1/r^2$ for $\overline{f}_{\text{sosS}}^{(r)}$.

Kernel approximations $\delta_0^{(r)}(x)$ of the Dirac at degree $r = 8, 16, 32, 64$:




$$\frac{\delta_0^{(r)}(x)}{\pi\sqrt{1-x^2}} \sim \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad \text{with} \quad \sigma \sim \frac{\pi}{r+2}.$$

# Concluding remarks

- Discrepancy between **theory and practice** for the upper bounds:

# Concluding remarks

- Discrepancy between **theory and practice** for the upper bounds:

| r | Matyas | | Three-Hump | Camel | Motzkin | |
|---|---|---|---|---|---|---|
| | Value | Time (sec.) | Value | Time (sec.) | Value | Time (sec.) |
| 2 | 8.26667 | 0.000739 | 265.774 | 0.000742 | 4.2 | 0.000719 |
| 6 | 4.28172 | 0.000072 | 29.0005 | 0.000066 | 1.06147 | 0.000080 |
| 12 | 2.99563 | 0.000263 | 4.43983 | 0.000263 | 0.801069 | 0.000208 |
| 18 | 1.83356 | 0.000655 | 2.55032 | 0.000586 | 0.565553 | 0.000766 |
| 24 | 1.11785 | 0.001753 | 1.2775 | 0.001693 | 0.406076 | 0.001712 |
| 30 | 0.8524 | 0.002270 | 1.0185 | 0.002936 | 0.3004 | 0.002351 |
| 36 | 0.5760 | 0.005510 | 0.7113 | 0.004882 | 0.2300 | 0.006060 |
| 40 | 0.4815 | 0.006975 | 0.6064 | 0.007031 | 0.1817 | 0.007686 |

Matyas: $f = (x_1 + 2x_2 - 7)^2 + (2x_1 + x_2 - 5)^2, \quad K = [-10, 10]^2$.

Three-Hump Camel: $f = 2x_1^2 - 1.05x_1^4 + \frac{1}{6}x_1^6 + x_1x_2 + x_2^2, \; K = [-5, 5]^2$.

Motzkin: $f = x_1^4x_2^2 + x_1^2x_2^4 - 3x_1^2x_2^2 + 1, \quad K = [-2, 2]^2$.

$f_{\min} = 0$, bounds $\overline{f}^{(r)}_{\text{sos}}$ with SoS density.

# Concluding remarks

- Discrepancy between **theory and practice** for the upper bounds,

  and also for the **lower vs. upper bounds:**

  Better theoretical convergence results for the upper bounds, but slower in practice...

# Concluding remarks

- Discrepancy between **theory and practice** for the upper bounds,

  and also for the **lower vs. upper bounds:**

  Better theoretical convergence results for the upper bounds, but slower in practice...

- **Problem:** Show **better convergence rates** for the lower bounds.

  New techniques needed ...

# Concluding remarks

- One can use the upper bounds to generate good feasible solutions using sampling.

| r | $\overline{f}^{(r)}_{\text{sos}}$ | Mean | Variance | Minimum | Sample Size |
|---|---|---|---|---|---|
| 2 | 265.774 | 216.773 | 177142.0 | 0.106854 | 20 |
| | | 261.23 | 193466.0 | 0.11705 | 1000 |
| 4 | 29.0005 | 28.0344 | 2964.85 | 1.1718 | 20 |
| | | 27.712 | 6712.8 | 0.014255 | 1000 |
| 14 | 4.43983 | 3.96711 | 20.3193 | 0.260331 | 20 |
| | | 3.7911 | 57.847 | 0.0076111 | 1000 |
| 22 | 1.71275 | 1.30757 | 1.90985 | 0.0320489 | 20 |
| | | 1.6379 | 7.2518 | 0.0021144 | 1000 |
| 24 | 1.27749 | 0.841194 | 0.914514 | 0.0369565 | 20 |
| | | 1.2105 | 2.3 | 0.0005154 | 1000 |
| Uniform Sample | | 304.032 | 163021.0 | 1.65885 | 20 |
| | | 243.216 | 183724.0 | 0.00975034 | 1000 |

SoS upper bounds for the Three-Hamp Camel function:
$f = 2x_1^2 - 1.05x_1^4 + \frac{1}{6}x_1^6 + x_1x_2 + x_2^2$ over $K = [-5, 5]^2$.

Thank you

# Based on the papers

- *Improved convergence rates for Lasserre-type hierarchies of upper bounds for box-constrained polynomial optimization.* With E. de Klerk, R. Hess, arXiv:1603.03329.
- *Bound-constrained polynomial optimization using only elementary calculations*. With E. de Klerk, J. Lasserre and Z. Sun, arXiv:1507.04404.
- *Convergence analysis for Lasserre's measure-based hierarchy of upper bounds for polynomial optimization.* With E. de Klerk and Z. Sun, arXiv:1411.6867
- *On the convergence rate of grid search for polynomial optimization over the simplex.* With E. de Klerk, Z. Sun, J. Vera, Opt. Letters, 2016.
- *An alternative proof of a PTAS for fixed-degree polynomial optimization over the simplex.* With E. de Klerk and Z. Sun. Math.Prog. 2014.
- *A PTAS for the minimization of polynomials of fixed degree over the simplex.* With E. De Klerk and P. Parrilo. TCS, 2006.

# Optimization over the unit sphere

SOS lower bounds:

$$\underline{f}_{\mathrm{sosS}}^{(r)} = \sup \ \lambda \ \text{ s.t. } \ f(x) - \lambda = \underbrace{\sigma_0}_{\text{SOS of degree } 2r} + u(1 - \sum_{i=1}^{n} x_i^2).$$

## Theorem

*Let $f$ be a homogeneous polynomial of even degree.*

- *[Faybusovich 2003] $f_{\min} - \underline{f}_{sosS}^{(r)} = O(\frac{1}{r})$ for $r = O(n)$.*
- *In addition, Doherty-Wehner (2013) construct measure-based upper bounds with the same performance guarantee.*
- *Parrilo-Wehner announced convergence in $O(\frac{1}{r^2})$ for both upper and lower bounds.*