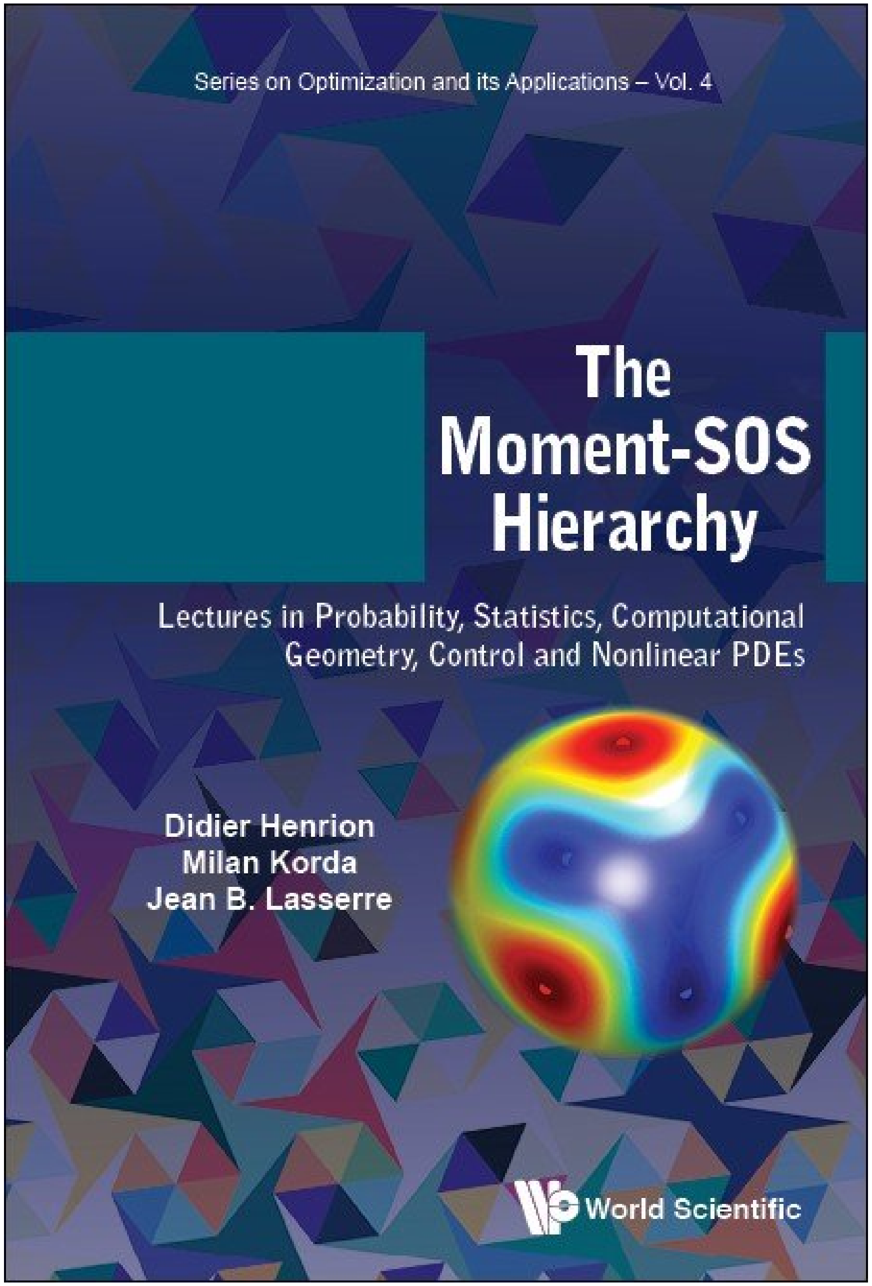
Didier Henrion

The moment-SOS hierarchy for polynomial optimization and volume approximation

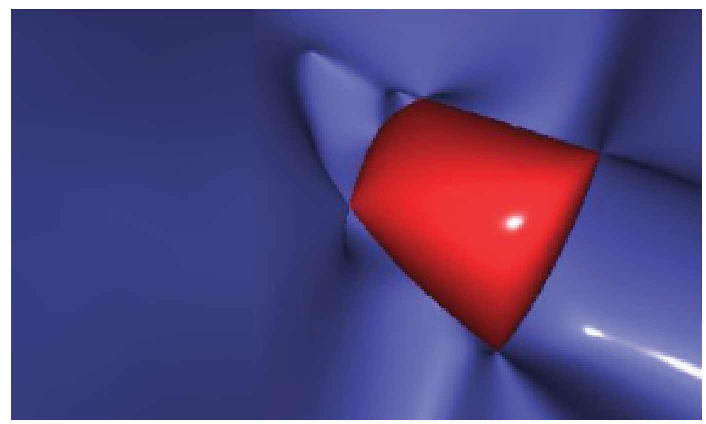
CWI Amsterdam Nov 2022



Poema

Polynomial Optimization, Efficiency through Moments and Algebra

poema-network.eu



Moment-SOS aka Lasserre hierarchy

Data science problems formulated as infinite-dimensional **linear** optimization problem on **measures**

Solved approximately with a family of **convex** (semidefinite) relaxations of increasing size indexed by relaxation order $r \in \mathbb{N}$

Based on the **duality** between the cone of positive polynomials and moments and their sum of squares (SOS) and semidefinite programming (SDP) approximations

Approximate solutions can be **extracted** from the solutions of the convex relaxations, with **convergence** guarantees

Outline

1. POP and the moment-SOS hierarchy

2. Volume approximation

1. POP and the moment-SOS hierarchy

POP = Polynomial Optimization Problem

Given multivariate real polynomials $p, p_1, \ldots, p_k$, solve **globally**

$$\begin{array}{rll} p^* := & \min_x & p(x) \\ & \text{s.t.} & x \in X := \{x \in \mathbb{R}^n : p_k(x) \geq 0,\ k = 1, \ldots, m\} \end{array}$$

where X is bounded and $p \in \mathbb{R}[x]_d$ has degree d

In general POP can be very challenging:

- p can be **nonconvex**
- X can be **nonconvex** and/or disconnected and/or discrete
- there can be **several** global optimizers, maybe infinitely many

Suppose we can generate samples $x_k \in X$ for $k = 1, \ldots, N$

Consider the linear optimization problem

$$\min_{w} \sum_{k=1}^{N} p(x_k) w_k \quad \text{s.t.} \quad \sum_{k=1}^{N} w_k = 1, \; w_k \geq 0$$

so that weight w_k should be large whenever $p(x_k)$ is small

Passing to the limit

$$\min_{\mu} \int_X p(x) d\mu(x) \quad \text{s.t.} \quad \int_X d\mu(x) = 1, \; \mu \geq 0$$

the unknown μ is now a **probability measure** on X

Let $(a_i(x))_{i\in\mathbb{N}^n_d}$ denote a basis of the vector space $\mathbb{R}[x]_d$ indexed in $\mathbb{N}^n_d := \{i \in \mathbb{N}^n : \sum_{k=1}^n i_k \leq d\}$ of dimension $n_d := \binom{n+d}{n}$

The polynomial p can then be written as

$$p(x) = \sum_{i\in\mathbb{N}^n_d} p_i a_i(x)$$

and the objective function can be written as

$$\int_X p(x)d\mu(x) = \sum_{i\in\mathbb{N}^n_d} p_i y_i$$

which is a linear function of the **moments** of measure μ

$$y_i = \int_X a_i(x)d\mu(x)$$

The nonlinear POP

$$p^* = \min_x p(x) \quad \text{s.t.} \quad x \in X$$

becomes a linear problem on moments

$$p^* = \min_y \sum_i p_i y_i \quad \text{s.t.} \quad y_0 = 1,\ y \in \mathscr{M}(X)_d$$

where $\mathscr{M}(X)_d$ is the **cone of moments** on X of degree up to d

We have reformulated a nonlinear nonconvex problem
as a linear problem in a finite-dimensional convex cone

However, just testing whether $y \in \mathscr{M}(X)_d$ is difficult

Not much is known about the geometry of this cone

No efficient barrier function is known

... so we will content ourselves with **approximations**

The cone of moments $\mathscr{M}(X)_d$ is dual to the cone $\mathscr{P}(X)_d$ of polynomials of degree up to d which are positive on X, we denote this by $\mathscr{M}(X)_d = \mathscr{P}(X)'_d$

Let us construct a family of **inner approximations** to $\mathscr{P}(X)_d$

Since $X := \{x \in \mathbb{R}^n : p_k(x) \geq 0,\ k = 1, \ldots, m\}$ is bounded, we can assume that $p_1(x) = R^2 - \sum_{i=1}^n x_i^2$ for R large enough

Let $p_0(x) := 1$ and for $r \geq d$ define the convex cone

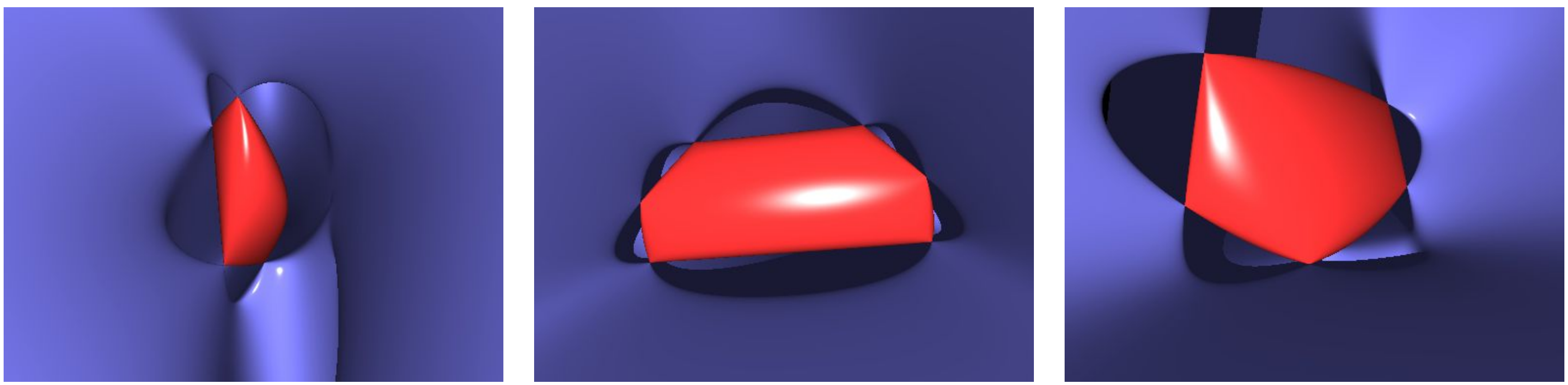
$$\mathscr{Q}(X)_r := \{q \in \mathbb{R}[x]_d : q = \sum_{k=0}^{m} \underbrace{s_k p_k}_{\in \mathbb{R}[x]_r},\ s_k\ \text{SOS}\}$$

where SOS stands for **sum of squares**

Lemma: By construction $\mathscr{Q}(X)_r \subset \mathscr{Q}(X)_{r+1} \subset \mathscr{P}(X)_d$

Polynomial SOS

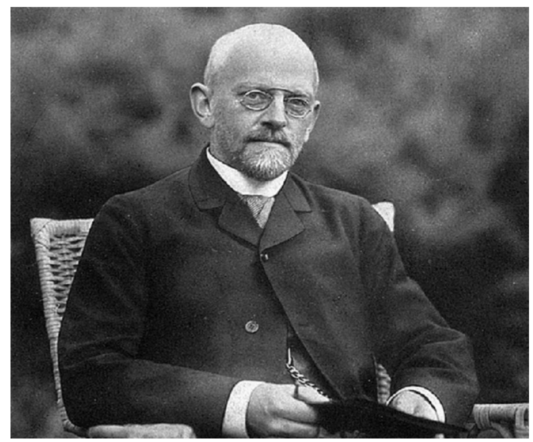
Lemma: Deciding whether a polynomial is SOS reduces to semidefinite programming (SDP) i.e. optimization over linear matrix inequalities (LMI)



Solved approximately efficiently with primal-dual interior-point

SOS and positivity

Theorem (Hilbert 1888): $\mathscr{Q}(\mathbb{R}^n)_{2d} = \mathscr{P}(\mathbb{R}^n)_{2d}$ if and only if $n = 1$ or $d = 1$ or $n = d = 2$



Hilbert's 17th problem at ICM Paris 1900

Motzkin's 1965 example

$$1 - 3x_1^2x_2^2 + x_1^4x_2^2 + x_1^2x_2^4 \in \mathscr{P}(\mathbb{R}^2)_6 \backslash \mathscr{Q}(\mathbb{R}^2)_6$$

Theorem (Putinar 1993): $\overline{\mathscr{Q}(X)_\infty} = \mathscr{P}(X)_d$

In words, every positive polynomial on a compact semialgebraic set can be approximated arbitrary well by SOS polynomials

For the moment cone we have the dual statement

Theorem : $\mathscr{Q}(X)'_r \supset \mathscr{Q}(X)'_{r+1} \supset \overline{\mathscr{Q}(X)'_\infty} = \mathscr{M}(X)_d$

and hence we can approximate elements of the moment cone as closely as desired with **SDP outer approximations**

Moment hierarchy

For our POP

$$p^* = \min_x p(x) \quad \text{s.t.} \quad x \in X$$

formulated as a moment problem

$$p^* = \min_y \sum_i p_i y_i \quad \text{s.t.} \quad y_0 = 1,\ y \in \mathscr{M}(X)_d$$

we now have a hierarchy of SDP problems of increasing size

$$p^*_r = \min_y \sum_i p_i y_i \quad \text{s.t.} \quad y_0 = 1,\ y \in \mathscr{Q}(X)'_r$$

Since $\mathscr{Q}(X)'_r \supset \mathscr{M}(X)_d$ the value p^*_r is a lower bound on p^*, we say that the SDP problem is a **relaxation** of the moment problem, and it yields a vector of **pseudo-moments**, i.e. not necessarily coming from a measure

SOS hierarchy

And hence for the dual POP

$$d^* = \max_{p_{\text{low}}} \quad p_{\text{low}} \quad \text{s.t.} \quad p(x) \geq p_{\text{low}} \quad \forall x \in X$$

formulated as a polynomial positivity problem

$$d^* = \max_{p_{\text{low}}} \quad p_{\text{low}} \quad \text{s.t.} \quad p(x) - p_{\text{low}} \in \mathscr{P}(X)_d$$

we have a hierarchy of SDP problems of increasing size

$$d_r^* = \max_{p_{\text{low}}} \quad p_{\text{low}} \quad \text{s.t.} \quad p(x) - p_{\text{low}} \in \mathscr{Q}(X)_r$$

The **SOS** constraint implies positivity on X

Theorem (Lasserre 2001): $p_r^* = d_r^* \leq p_{r+1}^* \leq \cdots \leq p_\infty^* = p^*$

Finite convergence

Theorem (Nie 2014): Generically $\exists r < \infty$ such that $p_r^* = p^*$

In other words, a vanishing small random perturbation of the input data of a given POP ensures **finite convergence** of the Lasserre hierarchy

We have linear algebra conditions on the moments to ensure finite convergence and **extract** global minimizers

We can also use the Christoffel-Darboux kernel to approximate the **variety** of global minimizers from the moments

Define the **moment matrix** of degree $2d$

$$M_d(y) := \int_X a(x)a(x)^\top d\mu(x)$$

as a symmetric matrix linear function of the moments y of μ

> **Theorem (Flat Extension, Curto & Fialkow 1991):** If the rank of $M_r(y)$ does not increase when r increases, then the moment relaxation is exact

Global solutions extracted by linear algebra, as implemented in our Matlab interface GloptiPoly (2002)

homepages.laas.fr/henrion

SUMS OF SQUARES, MOMENT MATRICES AND OPTIMIZATION OVER POLYNOMIALS

MONIQUE LAURENT*

Updated version: February 6, 2010

Abstract. We consider the problem of minimizing a polynomial over a semialgebraic set defined by polynomial equations and inequalities, which is NP-hard in general. Hierarchies of semidefinite relaxations have been proposed in the literature, involving positive semidefinite moment matrices and the dual theory of sums of squares of polynomials. We present these hierarchies of approximations and their main properties: asymptotic/finite convergence, optimality certificate, and extraction of global optimum solutions. We review the mathematical tools underlying these properties, in particular, some sums of squares representation results for positive polynomials, some results about moment matrices (in particular, of Curto and Fialkow), and the algebraic eigenvalue method for solving zero-dimensional systems of polynomial equations. We try whenever possible to provide detailed proofs and background.

Key words. positive polynomial, sum of squares of polynomials, moment problem, polynomial optimization, semidefinite programming

homepages.cwi.nl/∼monique

2. Volume Approximation

Given a polynomial p, we want to compute the volume or Lebesgue measure of the compact basic semialgebraic set

$$X := \{x \in \mathbb{R}^n : p(x) \geq 0\}$$

included in the unit Euclidean ball B.

Existing algorithmic approaches include:
- sampling (for convex sets),
- computer algebra (real algebraic geometry, symbolic integration, numerical analytic continuation),
- moment-SOS hierarchy (convex optimization)

Linear optimization problems in duality:

$$\begin{array}{ll} \text{sup}_\mu & \displaystyle\int_X \mu = \mu(X) \\ \text{s.t.} & 1 - \mu \geq 0 \text{ on } B \\ & \mu \geq 0 \text{ on } X \end{array} \qquad \begin{array}{ll} \text{inf}_v & \displaystyle\int_B v = \|v\|_{\mathscr{L}^1(B)} \\ \text{s.t.} & v \geq 0 \text{ on } B \\ & v - 1 \geq 0 \text{ on } X \end{array}$$

Theorem: Primal and dual values are vol X

[D. Henrion, J. B. Lasserre, C. Savorgnan. Approximate volume and integration for basic semialgebraic sets. SIAM Review 51(4), 2009]

Primal

$$\begin{array}{ll} \text{sup}_\mu & \int_X \mu \\ \text{s.t.} & 1 - \mu \geq 0 \text{ on } B \\ & \mu \geq 0 \text{ on } X \end{array}$$

$\mu =$ **measure** on X

Optimal solution $\mu^* = I_X$

Dual

$$\begin{array}{ll} \inf_v & \int_B v \\ \text{s.t.} & v \geq 0 \text{ on } B \\ & v - 1 \geq 0 \text{ on } X \end{array}$$

$v =$ **function** on B

Optimal solution $v^* = I_X$
(however discontinuous)

The primal problem on measures

$$\begin{array}{ll} \text{sup}_\mu & \int_X \mu \\ \text{s.t.} & 1-\mu \geq 0 \text{ on } B \\ & \mu \geq 0 \text{ on } X \end{array}$$

can be written as

$$\begin{array}{ll} \text{sup}_{\mu,\nu} & \int_X \mu \\ \text{s.t.} & \mu + \nu = I_B \\ & \mu \geq 0 \text{ on } X \\ & \nu \geq 0 \text{ on } B \end{array}$$

or equivalently as a primal problem on **moments**

$$\begin{array}{ll} \text{sup}_{y,z} & y_0 \\ \text{s.t.} & y_i + z_i = \displaystyle\int_B a_i(x)dx, \ i \in \mathbb{N}^n \\ & y \in \mathscr{M}(X), \ z \in \mathscr{M}(B) \end{array}$$

upon denoting $y_i := \displaystyle\int_X a_i(x)d\mu(x)$ and $z_i := \displaystyle\int_X a_i(x)d\nu(x)$

The dual problem on functions

$$\begin{array}{ll} \inf_v & \int_B v \\ \text{s.t.} & v \geq 0 \text{ on } B \\ & v - 1 \geq 0 \text{ on } X \end{array}$$

can be written as a dual problem on **positive polynomials**

$$\begin{array}{ll} \inf_v & \sum_i v_i \int_B a_i(x)dx \\ \text{s.t.} & v := \sum_i v_i a_i(x) \in \mathscr{P}(B) \\ & v - 1 \in \mathscr{P}(X) \end{array}$$

Linear optimization problems in duality:

$$\begin{aligned} p^* = \ & \sup_{y,z} y_0 \\ \text{s.t.} \ & y_i + z_i = \int_B a_i(x)dx, \ i \in \mathbb{N}^n \\ & y \in \mathscr{M}(X), \ z \in \mathscr{M}(B) \end{aligned}$$

$$\begin{aligned} d^* = \ & \inf_v \sum_{i \in \mathbb{N}^n} v_i \int_B a_i(x)dx \\ \text{s.t.} \ & v \in \mathscr{P}(B), \ v - 1 \in \mathscr{P}(X) \end{aligned}$$

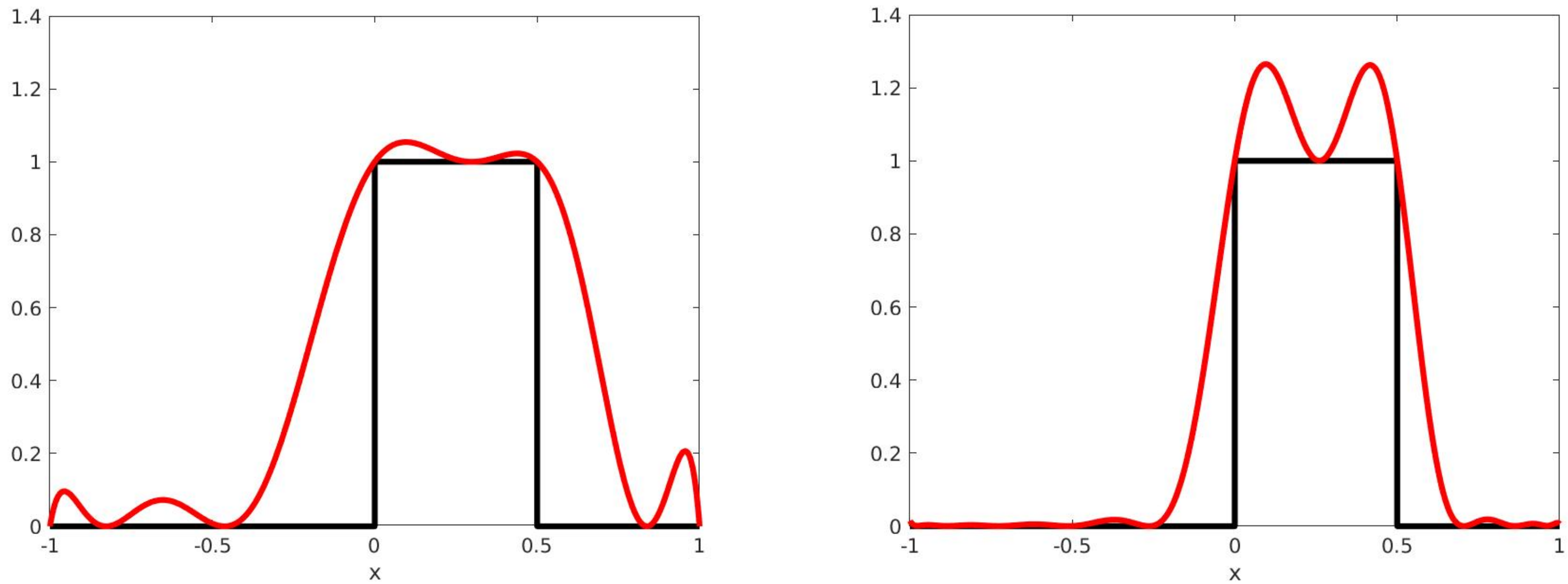
can be solved with the moment-SOS hierarchy:

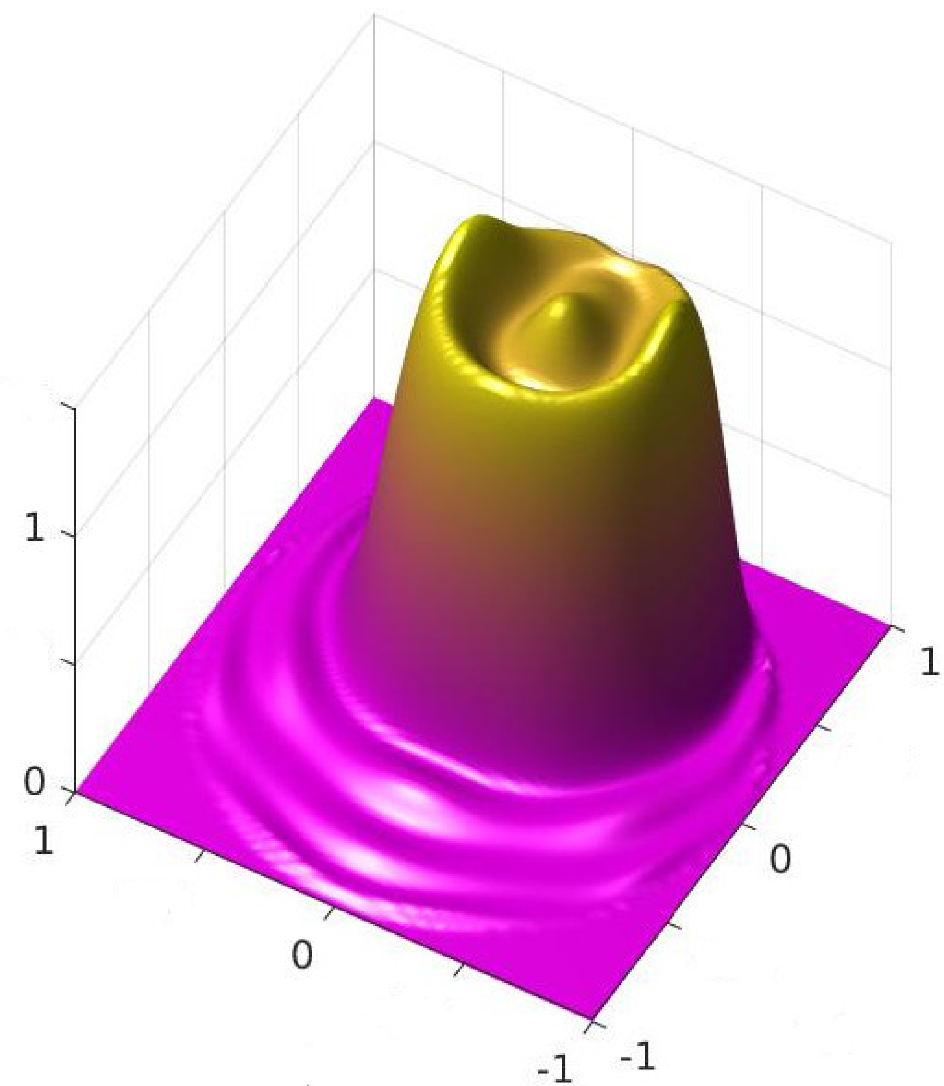
$$\begin{aligned} p_r^* = \ & \sup_{y,z} y_0 \\ \text{s.t.} \ & y_i + z_i = \int_B a_i(x)dx, \ i \in \mathbb{N}_r^n \\ & y \in \mathscr{Q}(X)'_r, \ z \in \mathscr{Q}(B)'_r \end{aligned}$$

$$\begin{aligned} d_r^* = \ & \inf_v \sum_{i \in \mathbb{N}_r^n} v_i \int_B a_i(x)dx \\ \text{s.t.} \ & v \in \mathscr{Q}(B)_r, \ v - 1 \in \mathscr{Q}(X)_r \end{aligned}$$

Thm: $p_r^* = d_r^* \geq p_{r+1}^* = d_{r+1}^* \geq p_\infty^* = d_\infty^* = p^* = d^* = \text{vol}\, X$

Dual polynomial SOS approximation suffers from the Gibbs effect





1
0
1
0
-1
1
0
-1