Harmonic hierarchies for polynomial optimization.

Sergio Cristancho
Mauricio Velasco
Universidad de los Andes (Colombia)

BrainPOP Seminar
November 22, 2021

Polynomial Optimization:

Denote by:

- $R := \mathbb{R}[x_1, \ldots, x_n]$ the ring of n-variate polynomials with real coefficients.
- $F \in R_{2k}$ a homogeneous polynomial of degree $2k$.
- $S \subseteq \mathbb{R}^n$ the unit sphere.

Definition.

The **polynomial optimization problem** *asks us to find*

$$\alpha^* := \min_{x \in S} F(x)$$

Polynomial Optimization:

Denote by:

- $R := \mathbb{R}[x_1, \ldots, x_n]$ the ring of n-variate polynomials with real coefficients.
- $F \in R_{2k}$ a homogeneous polynomial of degree $2k$.
- $S \subseteq \mathbb{R}^n$ the unit sphere.

Definition.

The **polynomial optimization problem** *asks us to find*

$$\alpha^* := \min_{x \in S} F(x)$$

This is a fundamental problem for several reasons: It is a model for **global**, **non-convex** optimization problems and has a **wealth of applications** initiated by work of J.B. Lasserre and P. Parrilo in the early 2000's ([20C93 according to 2020 MSC])

A solution strategy:

Assume $F(x_1, \ldots, x_n)$ is homogeneous and of even degree $2k$.

$$\alpha^* := \min_{x \in S} F(x)$$

Lemma.

The following equality holds:

$$\alpha^* := \max\left\{\lambda \in \mathbb{R} : F(x) - \lambda \|x\|_2^{2k} \geq 0 \text{ on } S\right\}$$

A solution strategy:

Assume $F(x_1, \ldots, x_n)$ is homogeneous and of even degree $2k$.

$$\alpha^* := \min_{x \in S} F(x)$$

Lemma.

The following equality holds:

$$\alpha^* := \max \left\{ \lambda \in \mathbb{R} : F(x) - \lambda \|x\|_2^{2k} \geq 0 \text{ on } S \right\}$$

Definition.

Let $P_{2k} \subseteq R_{2k}$ be the collection of homogeneous polynomials of degree $2k$ which are nonnegative on S.

A solution strategy:

Assume $F(x_1, \ldots, x_n)$ is homogeneous and of even degree $2k$.

$$\alpha^* := \min_{x \in S} F(x)$$

Lemma.

The following equality holds:

$$\alpha^* := \max \left\{ \lambda \in \mathbb{R} : F(x) - \lambda \|x\|_2^{2k} \geq 0 \text{ on } S \right\}$$

Definition.

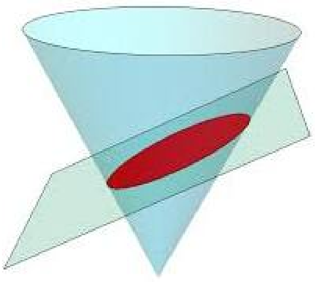
Let $P_{2k} \subseteq R_{2k}$ be the collection of homogeneous polynomials of degree $2k$ which are nonnegative on S.

$$\alpha^* := \max \left\{ \lambda \in \mathbb{R} : F(x) - \lambda \|x\|^{2k} \in P_{2k} \right\}$$

Nonnegative polynomials

Definition.

Let $P_{2k} \subseteq R_{2k}$ be the collection of homogeneous polynomials of degree $2k$ which are nonnegative on S.



The set $P_{2k} \subseteq R_{2k}$ is a closed convex cone.

Hierarchies:

A natural approach is to construct **hierarchies** approximating P_{2k},

Definition.

A hierarchy is a collection of convex cones $(A_s)_{s\in\mathbb{N}}$ satisfying the following properties:

1. *The cones A_s have fast membership algorithms.*
2. $A_s \subseteq P$
3. *The sets A_s converge to P in the sense that*

$$P^\circ \subseteq \bigcup_s A_s \subseteq P$$

Hierarchies:

A natural approach is to construct **hierarchies** approximating P_{2k},

Definition.

A hierarchy is a collection of convex cones $(A_s)_{s\in\mathbb{N}}$ satisfying the following properties:

1. *The cones A_s have fast membership algorithms.*
2. $A_s \subseteq P$
3. *The sets A_s converge to P in the sense that*

$$P^\circ \subseteq \bigcup_s A_s \subseteq P$$

Useful because if

$$\alpha_s := \sup\left\{\lambda \in \mathbb{R} : F(x) - \lambda\|x\|^{2k} \in A_s\right\}$$

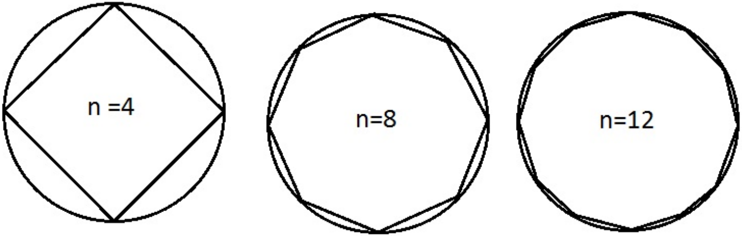
then the α_s are easily computable, $\alpha_s \leq \alpha^*$ and $\lim_{s\to\infty} \alpha_s = \alpha^*$.

Hierarchies

There exist several well-known hierarchies for P_{2k}

1. SOS hierarchies: [Parrilo],[Lasserre],[Peña, Vera, Zuluaga],[D. Henrion],...
2. SONC, SIGNOMIALS: [T. De Wolff], [V. Chandrasekaran],...
3. Polyhedral hierarchies:
   - Polya-type [Peña, Vera, Zuluaga],
   - DSOS, SDSOS: [A.A. Ahmadi, Majumdar],[A.A.Ahmadi, G. Hall],...

The purpose of this talk is to introduce several new **polyhedral hierarchies** for approximating P and to give **quantitative bounds on the rate at which they converge**.



We call them **Harmonic Hierarchies** because the harmonic analysis on spheres (or equivalently the representation theory of $SO(n)$) plays a key role in their definition.

Outline

1. Motivation.
2. Three basic ingredients:
    1. Quadrature rules on spheres.
    2. Polynomial averaging operators.
    3. Harmonic expansion on spheres.
3. The **Construction** of Harmonic hierarchies for P_{2k}.
4. A Convergence **Theorem**.
5. Some simple polynomial optimization **examples** (computed with our Julia package for optimization via harmonic hierarchies).

1. Quadrature rules

Definition.

A **quadrature rule** *of strength* $2t$ *on the sphere* $S \subseteq \mathbb{R}^n$ *is a pair* (X, W) *where* $X \subseteq S$ *is a finite set and* $W : X \to \mathbb{R}_{>0}$ *is a function for which the following equality holds,*

$$\forall f \in R_{2t} \left(\int_S f(y) d\mu(y) = \sum_{x \in X} W(x) f(x) \right)$$

where the integral on the left-hand side is taken with respect to the $(n-1)$*-dimensional volume measure* μ *on the sphere* S*.*

1. Quadrature rules

Definition.

A **quadrature rule** *of strength* $2t$ *on the sphere* $S \subseteq \mathbb{R}^n$ *is a pair* (X, W) *where* $X \subseteq S$ *is a finite set and* $W : X \to \mathbb{R}_{>0}$ *is a function for which the following equality holds,*

$$\forall f \in R_{2t} \left(\int_S f(y) d\mu(y) = \sum_{x \in X} W(x) f(x) \right)$$

where the integral on the left-hand side is taken with respect to the $(n-1)$*-dimensional volume measure* μ *on the sphere* S*.*

Example:

$$\forall f \in R_{2t} \left(\int_{S^1} f(x, y) d\mu(y) = \sum_{j=0}^{2t+1} \frac{2\pi}{2(t+1)} f(x_j) \right)$$

where the x_j *are the vertices of a regular* $2(t+1)$*-gon.*

The main invariant of a quadrature rule (X, W) is *size* $|X|$. It is known that the minimal size ν_{2k} of a quadrature rule of weight $2k$ satisfies

$$\dim(R_k) \leq \nu_{2k} \leq \dim(R_{2k}) + 1$$

rules of minimal size are known only in few cases.

The main invariant of a quadrature rule (X, W) is *size* $|X|$. It is known that the minimal size ν_{2k} of a quadrature rule of weight $2k$ satisfies

$$\dim(R_k) \leq \nu_{2k} \leq \dim(R_{2k}) + 1$$

rules of minimal size are known only in few cases.

Example: (Gaussian Quadrature)

If X consists of the k roots of the Jacobi polynomials $P_k^{(\alpha,\beta)}(t)$ there are weights W which lead to a quadrature rule of strength $2k-1$ on $[-1,1]$,

$$\int_{-1}^{1} f(t)(1-t)^a(1+t)^b dt = \sum_{x\in X} W(x)f(x)$$

1. Quadrature rules

Theorem. (-,Cristancho)

For every integer t there exists a quadrature rule of strength $2t$ in the sphere $S \subseteq \mathbb{R}^n$ supported in $2(t+1)^{n-1}$ points.

1. Quadrature rules

Theorem. (-,Cristancho)

For every integer t there exists a quadrature rule of strength $2t$ in the sphere $S \subseteq \mathbb{R}^n$ supported in $2(t+1)^{n-1}$ points.

This quadrature rule is built inductively on dimension, starting from the very symmetric quadrature rule of the circle and using a new Gaussian quadrature at each stage.

1. Quadrature rules

Theorem. (-,Cristancho)

For every integer t there exists a quadrature rule of strength $2t$ in the sphere $S \subseteq \mathbb{R}^n$ supported in $2(t+1)^{n-1}$ points.

This quadrature rule is built inductively on dimension, starting from the very symmetric quadrature rule of the circle and using a new Gaussian quadrature at each stage. This requires computing the zeroes of Jacobi polynomials.

1. Quadrature rules

Theorem. (-,Cristancho)

For every integer t there exists a quadrature rule of strength $2t$ in the sphere $S \subseteq \mathbb{R}^n$ supported in $2(t+1)^{n-1}$ points.

This quadrature rule is built inductively on dimension, starting from the very symmetric quadrature rule of the circle and using a new Gaussian quadrature at each stage. This requires computing the zeroes of Jacobi polynomials.

Theorems (and code!) of **Alex Townsend** *makes these computations possible for high degrees.*

The blue polyhedra Q_s:

Given quadrature rules (X_d, W_d) of strength $2d$ for every d and an integer $2k$

The blue polyhedra Q_s:

Given quadrature rules (X_d, W_d) of strength $2d$ for every d and an integer $2k$

Definition.

Let Q_s be the polyhedron

$$Q_s := \{F \in R_{2k} : \forall x \in X_{2(k+s)} \, (F(x) \geq 0)\}$$

The blue polyhedra Q_s:

Given quadrature rules (X_d, W_d) of strength $2d$ for every d and an integer $2k$

Definition.

Let Q_s be the polyhedron

$$Q_s := \{F \in R_{2k} : \forall x \in X_{2(k+s)} \, (F(x) \geq 0)\}$$

This polyhedron has $\leq 2(k+s+1)^{n-1}$ facets.

2. Averaging polynomials

A multivariate polynomial can be a very complicated function, so we will pass it through a polynomial *low pass filter*...

2. Averaging polynomials

A multivariate polynomial can be a very complicated function, so we will pass it through a polynomial *low pass filter*...

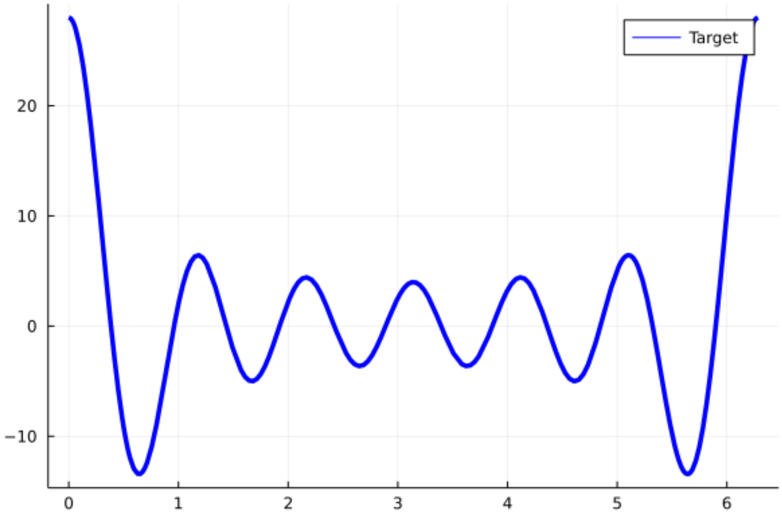
Definition.

Let $g(t)$ be a univariate polynomial which is nonnegative on $[-1,1]$. Define $\Gamma_g : R \to R$ via $\Gamma_g(f(x)) = h(x)$ where

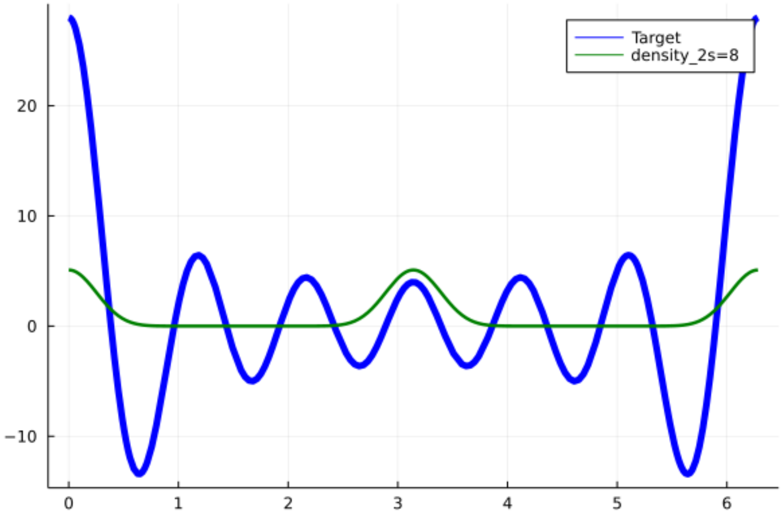
$$h(x) = \int_S g(\langle x, y \rangle) f(y) d\mu(y)$$

where μ is the $(n-1)$-dimensional volume measure.

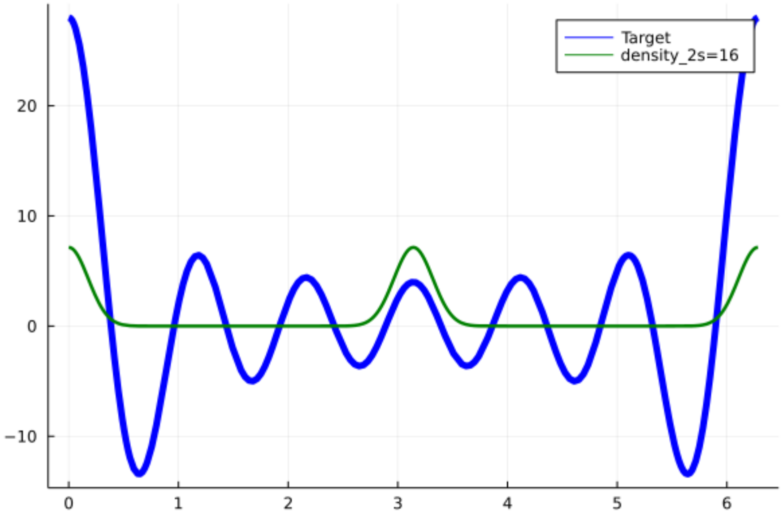
2. Averaging polynomials



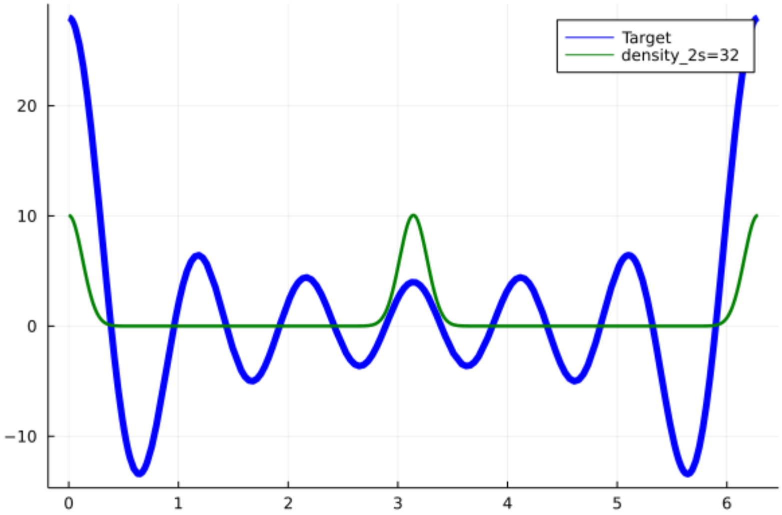
2. Averaging polynomials



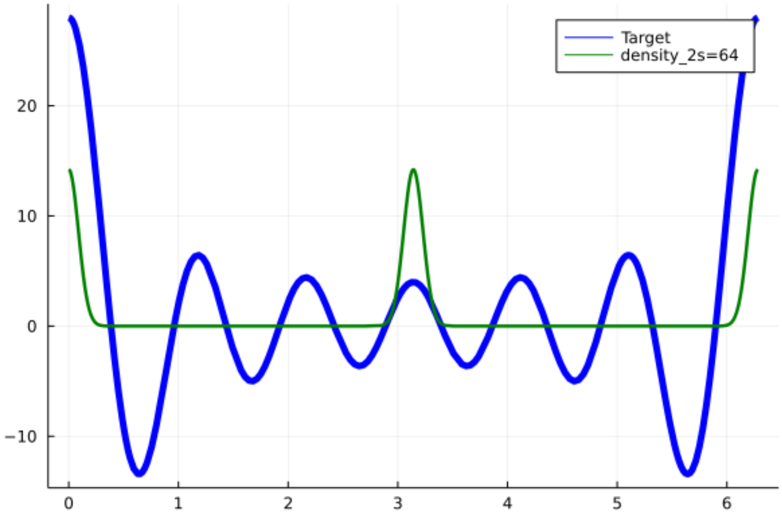
2. Averaging polynomials



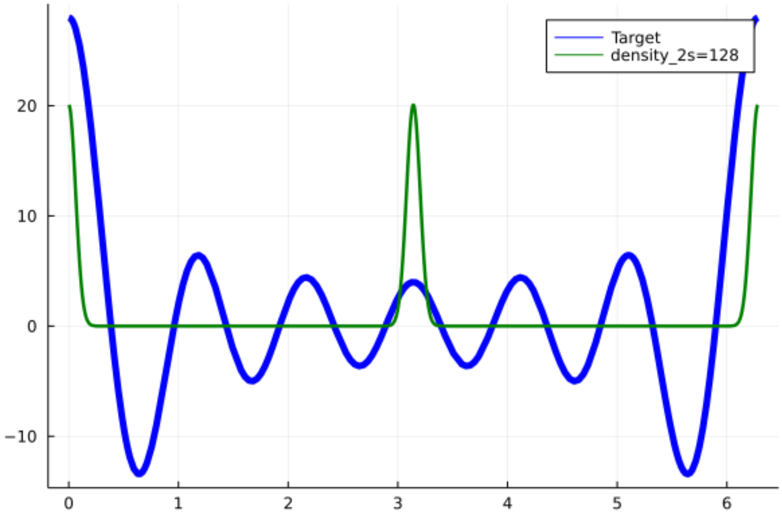
2. Averaging polynomials



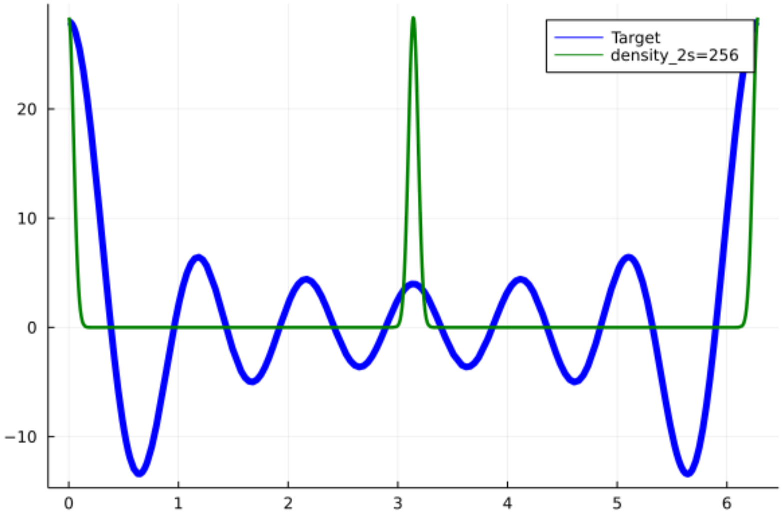
2. Averaging polynomials



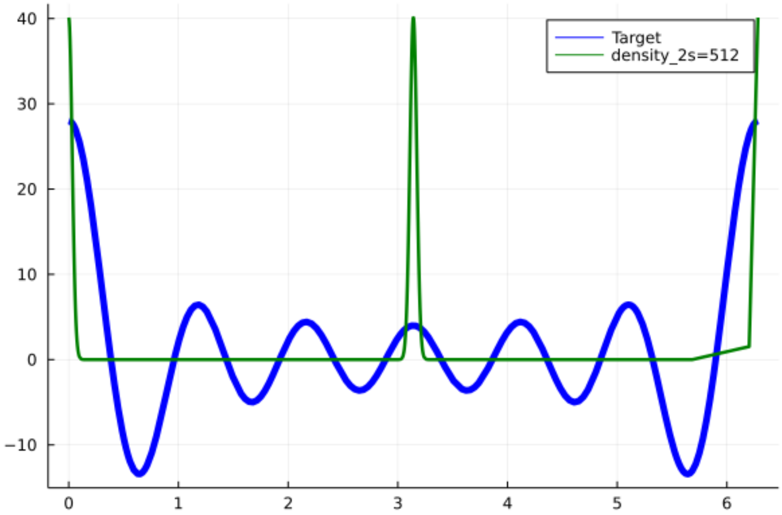
2. Averaging polynomials



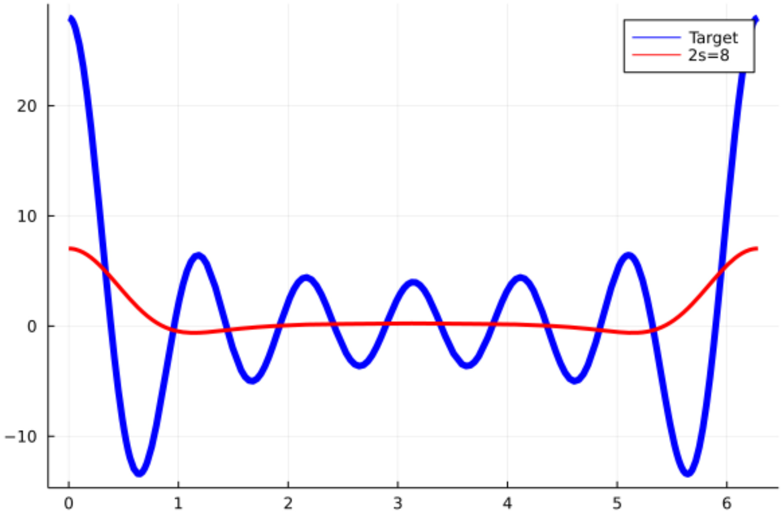
2. Averaging polynomials



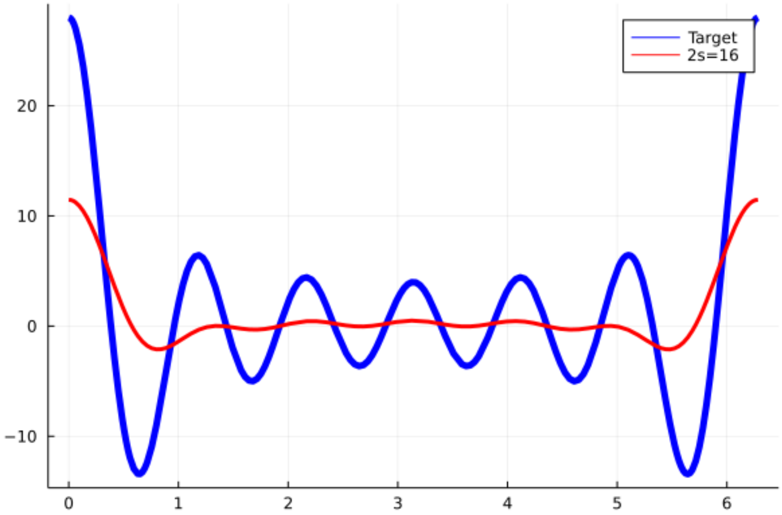
2. Averaging polynomials



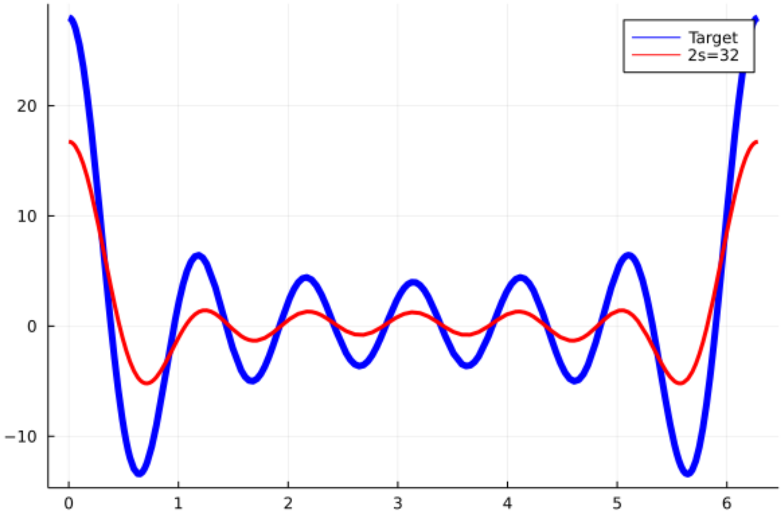
2. Averaging polynomials



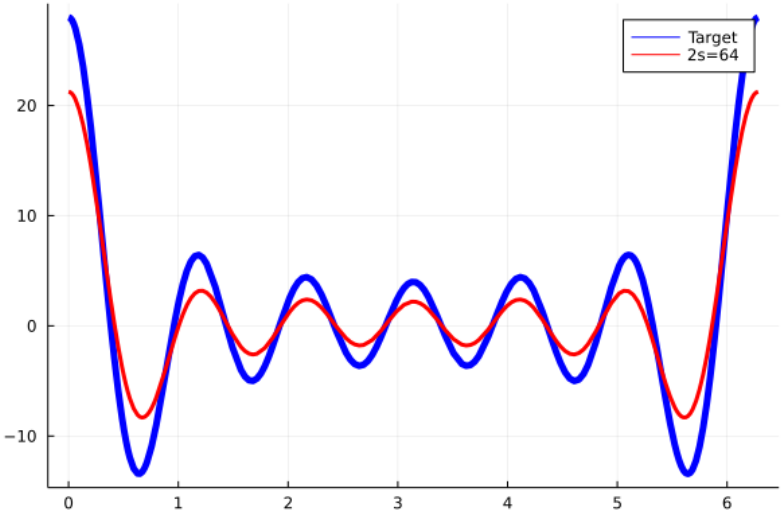
2. Averaging polynomials



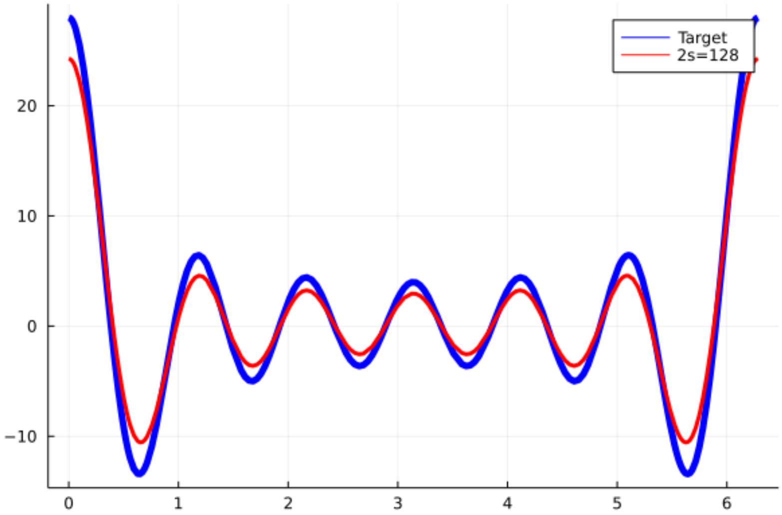
2. Averaging polynomials



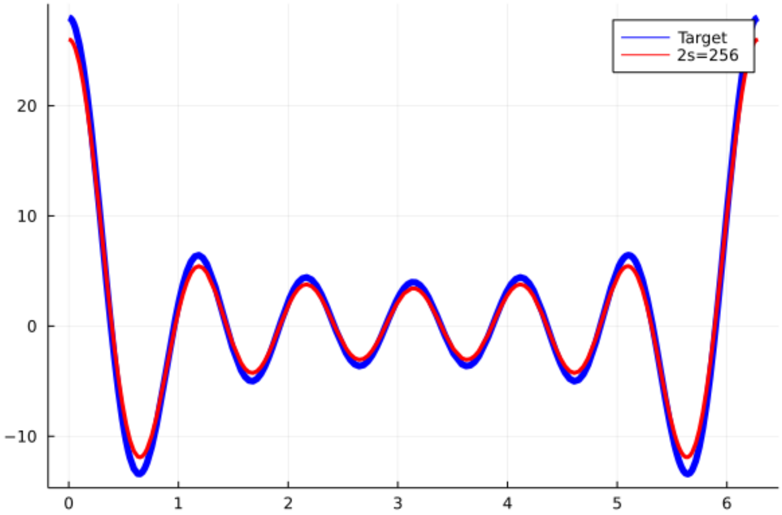
2. Averaging polynomials



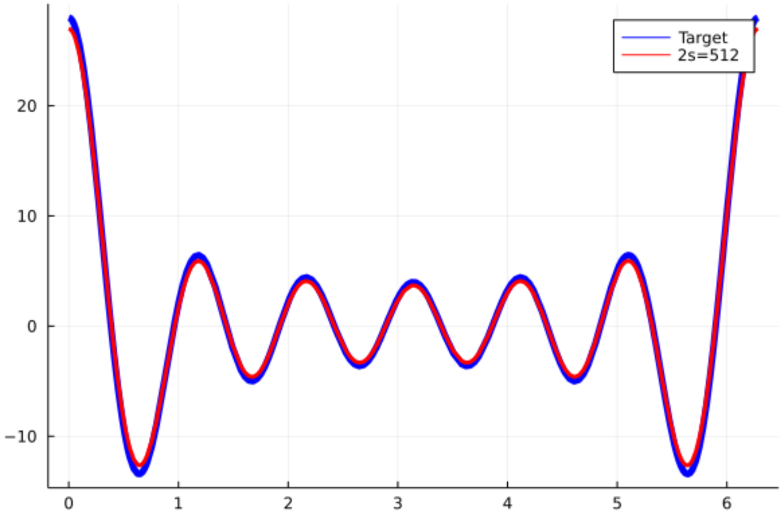
2. Averaging polynomials



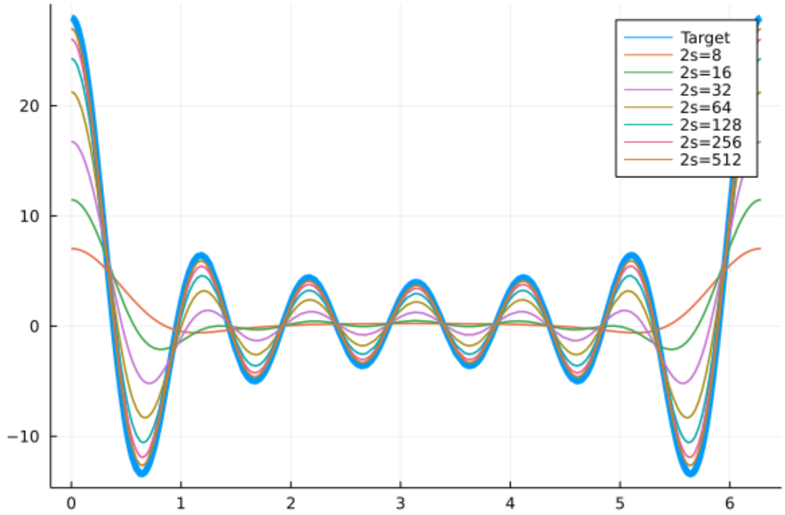
2. Averaging polynomials



2. Averaging polynomials



2. Averaging polynomials



Quadrature and Averaging

We are given Quadrature rules (X_d, W_d) of strength $2d$ and an integer $2k$. Recall

Quadrature and Averaging

We are given Quadrature rules (X_d, W_d) of strength $2d$ and an integer $2k$. Recall

Definition.

Let Q_s be the polyhedron

$$Q_s := \{F \in R_{2k} : \forall x \in X_{2(k+s)} \, (F(x) \geq 0)\}$$

Quadrature and Averaging

We are given Quadrature rules (X_d, W_d) of strength $2d$ and an integer $2k$. Recall

Definition.

Let Q_s *be the polyhedron*

$$Q_s := \{F \in R_{2k} : \forall x \in X_{2(k+s)} (F(x) \geq 0)\}$$

We will map the polyhedron Q_s into P with our averaging map Γ_g

Lemma. (-,Cristancho)

If $f \in Q_s$ *then* $\Gamma_s(f)$ *is a nonnegative polynomial.*

Lemma. (-,Cristancho)

If $f \in Q_s$ then $\Gamma_s(f)$ is a nonnegative polynomial.

Proof.

Recall $\Gamma_s(f(x)) = h(x)$.

$$h(x) = \int_S \frac{\langle x, y\rangle^{2s}}{N} f(y) d\mu(y)$$

Lemma. (-,Cristancho)

If $f \in Q_s$ then $\Gamma_s(f)$ is a nonnegative polynomial.

Proof.

Recall $\Gamma_s(f(x)) = h(x)$.

$$h(x) = \int_S \frac{\langle x, y\rangle^{2s}}{N} f(y)\,d\mu(y)$$

$$\int_S \frac{\langle x, y\rangle^{2s}}{N} f(y)\,d\mu(y) = \sum_{j=1}^{|X_{s+k}|} \frac{\langle x, y_j\rangle^{2s}}{N} f(y_j) w_j$$

Lemma. (-,Cristancho)

If $f \in Q_s$ then $\Gamma_s(f)$ is a nonnegative polynomial.

Proof.

Recall $\Gamma_s(f(x)) = h(x)$.

$$h(x) = \int_S \frac{\langle x, y\rangle^{2s}}{N} f(y) d\mu(y)$$

$$\int_S \frac{\langle x, y\rangle^{2s}}{N} f(y) d\mu(y) = \sum_{j=1}^{|X_{s+k}|} \frac{\langle x, y_j\rangle^{2s}}{N} f(y_j) w_j$$

Since $f \in Q_s$ the expression is a sum of even powers of linear forms with nonnegative coefficients. □

The maps Γ_g have the following remarkable symmetry property

Lemma.

For every $T \in SO(n)$ we have

$$\Gamma_g\left(f(Tx)\right) = h(Tx)$$

in other words Γ_g is a morphism of $SO(n)$ representations.

3.Harmonic expansions

The maps Γ_g have the following remarkable symmetry property

Lemma.

For every $T \in SO(n)$ we have

$$\Gamma_g\left(f(Tx)\right) = h(Tx)$$

in other words Γ_g is a morphism of $SO(n)$ representations.

As a result, due to the machinery of representation theory, all the maps Γ_g become simultaneously diagonal in a natural basis.

The maps Γ_g have the following remarkable symmetry property

Lemma.

For every $T \in SO(n)$ we have

$$\Gamma_g(f(Tx)) = h(Tx)$$

in other words Γ_g is a morphism of $SO(n)$ representations.

As a result, due to the machinery of representation theory, all the maps Γ_g become simultaneously diagonal in a natural basis. This basis allows us to understand the operators Γ_g and allows us to compute them for very high-degrees in practice.

3.Harmonic expansions

Every homogeneous polynomial $f \in R_{2k}$ *can be written uniquely in its* **harmonic expansion** *as*

$$f = \|x\|^{2k} f_0 + \|x\|^{2(k-1)} f_2 + \|x\|^{2(k-2)} f_4 + \cdots + f_{2k}$$

where the f_{2j} *are homogeneous harmonic polynomials (i.e.* $\Delta f_{2j} \equiv 0$*) of degree* $2j$*.*

3.Harmonic expansions

Every homogeneous polynomial $f \in R_{2k}$ can be written uniquely in its **harmonic expansion** *as*

$$f = \|x\|^{2k} f_0 + \|x\|^{2(k-1)} f_2 + \|x\|^{2(k-2)} f_4 + \cdots + f_{2k}$$

where the f_{2j} are homogeneous harmonic polynomials (i.e. $\Delta f_{2j} \equiv 0$) of degree $2j$.

Equivalently, the homogeneous polynomials decompose, as $SO(n)$-representations as:

$$R_{2k} = H_{2k} \oplus \|x\|^2 H_{2(k-1)} \oplus \|x\|^4 H_{2(k-2)} \oplus \cdots \oplus \|x\|^{2k} H_0$$

where the H_j are the vector spaces of homogeneous harmonic polynomials of degree j.

3.Harmonic expansions

Every homogeneous polynomial $f \in R_{2k}$ can be written uniquely in its **harmonic expansion** *as*

$$f = \|x\|^{2k} f_0 + \|x\|^{2(k-1)} f_2 + \|x\|^{2(k-2)} f_4 + \cdots + f_{2k}$$

where the f_{2j} are homogeneous harmonic polynomials (i.e. $\Delta f_{2j} \equiv 0$) of degree $2j$.

Equivalently, the homogeneous polynomials decompose, as $SO(n)$-representations as:

$$R_{2k} = H_{2k} \oplus \|x\|^2 H_{2(k-1)} \oplus \|x\|^4 H_{2(k-2)} \oplus \cdots \oplus \|x\|^{2k} H_0$$

where the H_j are the vector spaces of homogeneous harmonic polynomials of degree j.

This is the decomposition of R_{2k} into pairwise non-isomorphic irreducible representations of $SO(n)$.

3.Harmonic expansions

In this decomposition the operator Γ_g becomes diagonal...

Lemma. (Funk-Hecke)

Assume $g(t) = \sum_{j=0}^{n} \lambda_{2j}^{g} \phi_{2j}(t)$ *is the unique expression of* $g(t)$ *as linear combination of (suitably normalized) Gegenbauer polynomials. If*

$$f = \|x\|^{2k} f_0 + \|x\|^{2(k-1)} f_2 + \|x\|^{2(k-2)} f_4 + \cdots + f_{2k}$$

is the unique harmonic expansion for $f \in R_{2k}$ *then we have*

$$\hat{\Gamma}_g(f) = \lambda_0^g \|x\|^{2k} f_0 + \lambda_2^g \|x\|^{2(k-1)} f_2 + \lambda_4^g \|x\|^{2(k-2)} f_4 + \cdots + \lambda_{2k}^g f_{2k}.$$

3.Harmonic expansions

$$\hat{\Gamma}_g(f) = \lambda_0^g \|x\|^{2k} f_0 + \lambda_2^g \|x\|^{2(k-1)} f_2 + \lambda_4^g \|x\|^{2(k-2)} f_4 + \cdots + \lambda_{2k}^g f_{2k}.$$

Definition.

The **Frobenius threshold** *of the map* $\hat{\Gamma}_g$ *in degree* $2k$ *is the number*

$$\tau_{2k}(g) := \sqrt{\sum_{j=0}^{k} \dim(H_{2j}) \left(\frac{1}{\lambda_{2j}^g} - 1\right)^2}$$

Construction. Harmonic Hierarchies for P_{2k}

Given:

1. Quadrature rules (X_t, W_t) of strength t on S for every even integer t.
2. A sequence of even univariate polynomials $(g_s(t))_{s\in\mathbb{N}}$ which are nonnegative on the interval $[-1, 1]$.

Definition.

The **Linear Harmonic Hierarchy** $(A_s)_{s\in\mathbb{N}}$ *in degree* $2k$ *determined by* (1) *and* (2) *is given by*

$$A_s := \hat{\Gamma}_{g_s}(Q_s) \subseteq P_{2k}$$

A quantitative convergence Theorem

Theorem. (-,Cristancho)

The following statements hold:

1. *If* $f \in R_{2k}$ *satisfies the inequality*

$$\min_{x \in X_{2(k+d_s)}} f(x) > \frac{\|f\|_2}{\sqrt{\mu(S)}} \tau_{2k}(g_s),$$

   then $f \in A_s$*.*

2. *If* $\lim_{s\to\infty} \tau_{2k}(g_s) = 0$*, then every strictly positive polynomial is contained in some* A_s *and in particular the hierachy is convergent in the sense that the following inclusions hold*

$$P_{2k}^{\circ} \subseteq \bigcup_{s=0}^{\infty} A_s \subseteq P_{2k}$$

An example: pure powers HH

Example:

Let $g_s(t) = \frac{t^{2s}}{\int_S y_1^{2s} dy_1}$.

Corollary.

The inequality

$$\frac{1+\frac{n}{2}}{s} + O\left(\frac{1}{s^2}\right) \leq \tau_{2k}(g_s) \leq \frac{k^2+\frac{kn}{2}}{s} + O\left(\frac{1}{s^2}\right).$$

holds. In particular, the resulting linear harmonic hierarchy converges.

An example: pure powers HH

Remark.

The previous Corollary implies Polya's Theorem (Reznick's proof explained by Blekherman).

An example: pure powers HH

Remark.

The previous Corollary implies Polya's Theorem (Reznick's proof explained by Blekherman).

Remark.

There exist better g_s which can ensure convergence rates $\simeq 1/s^2$ [Kang-Fawzi].

An example: pure powers HH

Remark.

The previous Corollary implies Polya's Theorem (Reznick's proof explained by Blekherman).

Remark.

*There exist better g_s which can ensure convergence rates $\simeq 1/s^2$ [Kang-Fawzi]. Bounds for harmonic hierarchies (**linear**) are as good as the best bound for SOS-hierachies (**SDP**).*

An example: pure powers HH

Remark.

The previous Corollary implies Polya's Theorem (Reznick's proof explained by Blekherman).

Remark.

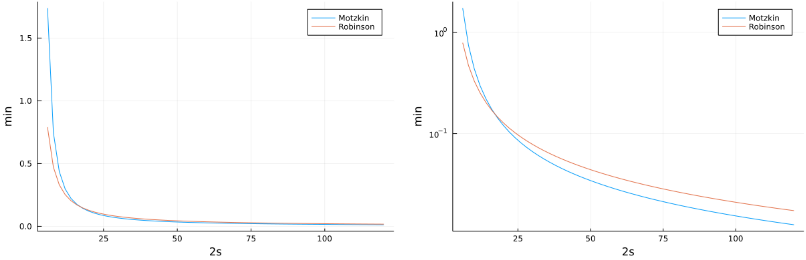
*There exist better g_s which can ensure convergence rates $\simeq 1/s^2$ [Kang-Fawzi]. Bounds for harmonic hierarchies (**linear**) are as good as the best bound for SOS-hierachies (**SDP**).*

Theorem. (-, Cristancho)

The problem of minimizing $\tau_{2k}(g)$ among all valid $g(t)$ of degree $\leq 2s$ is a convex optimization problem.

Polynomial optimization examples

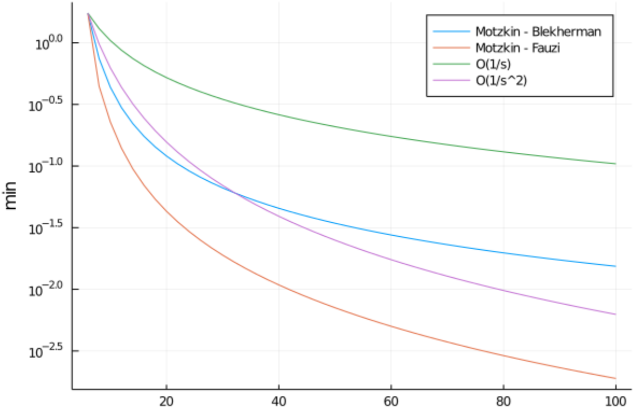
The Motzkin polynomial and the Robinson form are nonnegative ternary sextics with zeroes. We minimize them with harmonic hierarchies...



$$m(x, y, z) = x^4y^2 + x^2y^4 + z^6 - 3x^2y^2z^2$$

$$r(x, y, z) = x^6 + y^6 + z^6 - x^4y^2 - x^4z^2 - y^4z^2 - x^2z^4 - y^2z^4 + 3x^2y^2z^2$$

Pure powers vs Kang-Fawzi averaging...



$$m(x, y, z) = x^4y^2 + x^2y^4 + z^6 - 3x^2y^2z^2$$

What didn't fit...

1. Since $A_s \subseteq P_{2k}$ the duals $A_s^* \supseteq P_{2k}^*$ define converging **harmonic hierarchies for moments**.
2. Establishing the practical performance of harmonic hierarchies in areas of interest (and comparison with other hierarchies) is the subject of ongoing work...