# Geometry of Sums of Squares Relaxations

João Gouveia[1] Tim Netzer[2]

[1] University of Washington

[2] University of Leipzig

14th July '10 / SIAM Annual Meeting

# Optimization over semialgebraic sets

## Goal

Optimize a linear objective function over

$$S = \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geq 0, \ldots, g_k(\mathbf{x}) \geq 0\}$$

where the $g_i$ are polynomials over $\mathbb{R}$.

# Optimization over semialgebraic sets

**Goal**

Optimize a linear objective function over

$$S = \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geq 0, \ldots, g_k(\mathbf{x}) \geq 0\}$$

where the $g_i$ are polynomials over $\mathbb{R}$.

- It is NP-hard

# Optimization over semialgebraic sets

**Goal**

Optimize a linear objective function over

$$S = \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geq 0, \ldots, g_k(\mathbf{x}) \geq 0\}$$

where the $g_i$ are polynomials over $\mathbb{R}$.

- It is NP-hard
- We use a hierarchy of relaxations instead

# Nonnegativity over semialgebraic sets

We start by relaxing nonnegativity over $S$.

# Nonnegativity over semialgebraic sets

We start by relaxing nonnegativity over $S$.

A classic way of certifying nonnegativity of a polynomial $p$ over $S$ is to provide a representation

$$p(\mathbf{x}) = \sigma_0(\mathbf{x}) + \sum_{i=1}^{k} \sigma_k(\mathbf{x}) g_i(\mathbf{x})$$

where the $\sigma_i$ are sums of squares of polynomials.

# Nonnegativity over semialgebraic sets

We start by relaxing nonnegativity over $S$.

A classic way of certifying nonnegativity of a polynomial $p$ over $S$ is to provide a representation

$$p(\mathbf{x}) = \sigma_0(\mathbf{x}) + \sum_{i=1}^{k} \sigma_k(\mathbf{x}) g_i(\mathbf{x})$$

where the $\sigma_i$ are sums of squares of polynomials.

We will denote by $\Sigma_d^S$ the set of all polynomials that have such a representation with $\deg(\sigma_i g_i) \leq 2d$ for all $i$.

# Convex Hulls of semialgebraic sets

We want to use this tool to approximate conv($S$).

# Convex Hulls of semialgebraic sets

We want to use this tool to approximate conv($S$). Note that

## Convex Hull

$$\overline{\text{conv}(S)} = \bigcap_{\ell \text{ linear }, \ell|_S \geq 0} \{\mathbf{x} \in \mathbb{R}^n : \ell(\mathbf{x}) \geq 0\}.$$

# Convex Hulls of semialgebraic sets

We want to use this tool to approximate conv($S$). Note that

## Convex Hull

$$\overline{\text{conv}(S)} = \bigcap_{\ell \text{ linear }, \ell|_S \geq 0} \{\mathbf{x} \in \mathbb{R}^n : \ell(\mathbf{x}) \geq 0\}.$$

We can therefore relax it by

## Lasserre Bodies

$$\mathcal{L}_d(S) = \bigcap_{\ell \text{ linear }, \ell \in \Sigma_d^S} \{\mathbf{x} \in \mathbb{R}^n : \ell(\mathbf{x}) \geq 0\}$$

which we call the $d$-th Lasserre Relaxation of $S$.

# Some remarks

- There is a (more famous) definition of the Lasserre relaxation using moments. When $S$ has nonempty interior they match.

# Some remarks

- There is a (more famous) definition of the Lasserre relaxation using moments. When $S$ has nonempty interior they match.
- In particular one can optimize over these Lasserre bodies.

# Some remarks

- There is a (more famous) definition of the Lasserre relaxation using moments. When $S$ has nonempty interior they match.
- In particular one can optimize over these Lasserre bodies.
- But it is not necessarily easy.

# Some remarks

- There is a (more famous) definition of the Lasserre relaxation using moments. When $S$ has nonempty interior they match.
- In particular one can optimize over these Lasserre bodies.
- But it is not necessarily easy.

- From the definition $\overline{\text{conv}(S)} \subseteq \mathcal{L}_d(S) \subseteq \mathcal{L}_{d-1}(S)$ for all $d$.

# Some remarks

- There is a (more famous) definition of the Lasserre relaxation using moments. When $S$ has nonempty interior they match.
- In particular one can optimize over these Lasserre bodies.
- But it is not necessarily easy.

- From the definition $\overline{\text{conv}(S)} \subseteq \mathcal{L}_d(S) \subseteq \mathcal{L}_{d-1}(S)$ for all $d$.
- If $S$ is compact we have assymptotic convergence.

# Some remarks

- There is a (more famous) definition of the Lasserre relaxation using moments. When $S$ has nonempty interior they match.
- In particular one can optimize over these Lasserre bodies.
- But it is not necessarily easy.

- From the definition $\overline{\mathrm{conv}(S)} \subseteq \mathcal{L}_d(S) \subseteq \mathcal{L}_{d-1}(S)$ for all $d$.
- If $S$ is compact we have assymptotic convergence.
- When is the convergence (not) finite?

# Technical Lemma

To verify that we don't have finite convergence in a particular case we will use a well-known result:

**Theorem**

*If $S$ has non-empty interior, $\Sigma_d^S$ is closed for all $d$.*

# Technical Lemma

To verify that we don't have finite convergence in a particular case we will use a well-known result:

**Theorem**

*If $S$ has non-empty interior, $\Sigma_d^S$ is closed for all $d$.*

In geometrical terms this implies the following lemma:

**Lemma**

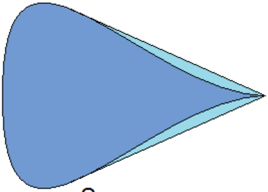
*If $S$ has non-empty interior and $\ell$ is a linear polynomial that is non-negative over $\mathcal{L}_d(S)$, then $\ell \in \Sigma_d^S$.*

which we can use to give a certificate of $\mathcal{L}_d(S) \neq \overline{\text{conv}(S)}$.

# Obstruction Lemma

## Lemma (Obstruction Lemma)

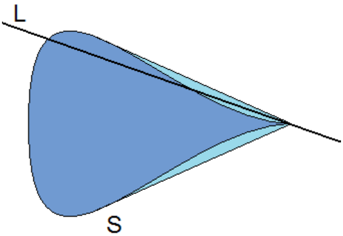
- *Let* $S = \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geq 0, \ldots, g_k(\mathbf{x}) \geq 0\}$,



# Obstruction Lemma

## Lemma (Obstruction Lemma)

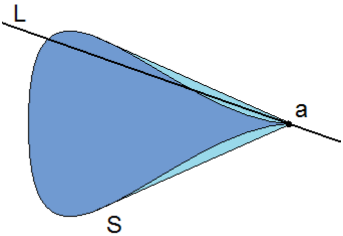
- *Let* $S = \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geq 0, \ldots, g_k(\mathbf{x}) \geq 0\}$,
- $L$ *a line in* $\mathbb{R}^n$ *s.t.* $\operatorname{int}(S \cap L) \neq \emptyset$ *relative to* $L$,



# Obstruction Lemma

## Lemma (Obstruction Lemma)

- *Let* $S = \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geq 0, \ldots, g_k(\mathbf{x}) \geq 0\}$,
- $L$ *a line in* $\mathbb{R}^n$ *s.t.* $\text{int}(S \cap L) \neq \emptyset$ *relative to* $L$,
- $a \in S$ *be in the relative boundary of* $\overline{\text{conv}(S)} \cap L$.

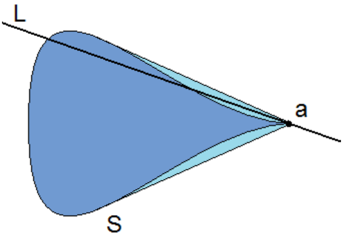


# Obstruction Lemma

## Lemma (Obstruction Lemma)

- *Let* $S = \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geq 0, \ldots, g_k(\mathbf{x}) \geq 0\}$*,*
- $L$ *a line in* $\mathbb{R}^n$ *s.t.* $\text{int}(S \cap L) \neq \emptyset$ *relative to* $L$*,*
- $a \in S$ *be in the relative boundary of* $\overline{\text{conv}(S)} \cap L$*.*

*If for all* $g_i$ *s.t.* $\underline{g_i(a) = 0}$ *we have* $\triangledown g_i(a) \perp L$ *then, for all* $d$*, we have* $\mathcal{L}_d(S) \neq \overline{\text{conv}(S)}$*.*

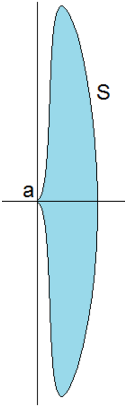


# Sketch of proof

Assume $a = \mathbf{0}$, $L = x_1$-axis, $S \cap L \subseteq \mathbb{R}_+$.

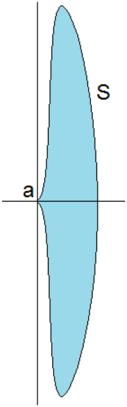
# Sketch of proof

Assume $a = \mathbf{0}$, $L = x_1$-axis, $S \cap L \subseteq \mathbb{R}_+$.



# Sketch of proof

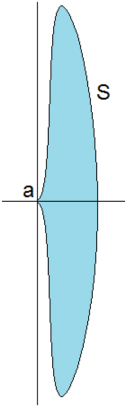
Assume $a = \mathbf{0}$, $L = x_1$-axis, $S \cap L \subseteq \mathbb{R}_+$.




- Let $g_i^*(x) = g_i(x, 0, \ldots, 0)$, and $S^* = \{x \in \mathbb{R} : g_i^*(x) \geq 0, i = 1, ..., k\}$.

# Sketch of proof

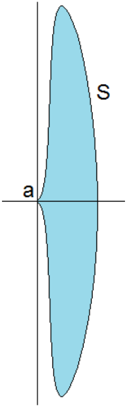
Assume $a = \mathbf{0}$, $L = x_1$-axis, $S \cap L \subseteq \mathbb{R}_+$.




- Let $g_i^*(x) = g_i(x, 0, \ldots, 0)$, and $S^* = \{x \in \mathbb{R} : g_i^*(x) \geq 0, i = 1, ..., k\}$.
- Note that $\mathcal{L}_d(S^*) \subseteq \mathcal{L}_d(S) \cap x_1$-axis.

# Sketch of proof

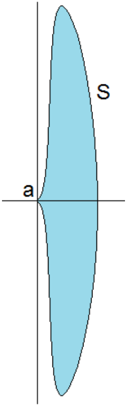
Assume $a = \mathbf{0}$, $L = x_1$-axis, $S \cap L \subseteq \mathbb{R}_+$.




- Let $g_i^*(x) = g_i(x, 0, \ldots, 0)$, and $S^* = \{x \in \mathbb{R} : g_i^*(x) \geq 0, i = 1, ..., k\}$.
- Note that $\mathcal{L}_d(S^*) \subseteq \mathcal{L}_d(S) \cap x_1$-axis.
- Let $i \in I$ if $g_i(0) = 0$ and $i \in J$ otherwise. By hypothesis $(g_i^*)'(0) = g_i^*(0) = 0$ for $i \in I$.

# Sketch of proof

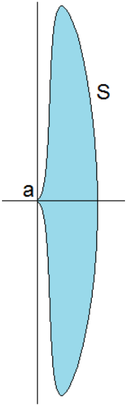
Assume $a = \mathbf{0}$, $L = x_1$-axis, $S \cap L \subseteq \mathbb{R}_+$.




- Let $g_i^*(x) = g_i(x, 0, \ldots, 0)$, and $S^* = \{x \in \mathbb{R} : g_i^*(x) \geq 0, i = 1, ..., k\}$.
- Note that $\mathcal{L}_d(S^*) \subseteq \mathcal{L}_d(S) \cap x_1$-axis.
- $(g_i^*)'(0) = g_i^*(0) = 0$ for $i \in I$.

# Sketch of proof

Assume $a = \mathbf{0}$, $L = x_1$-axis, $S \cap L \subseteq \mathbb{R}_+$.



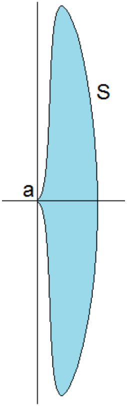

- Let $g_i^*(x) = g_i(x, 0, \ldots, 0)$, and $S^* = \{x \in \mathbb{R} : g_i^*(x) \geq 0, i = 1, ..., k\}$.
- Note that $\mathcal{L}_d(S^*) \subseteq \mathcal{L}_d(S) \cap x_1$-axis.
- $(g_i^*)'(0) = g_i^*(0) = 0$ for $i \in I$.
- Suppose $\ell(x) = x$ was valid in $\mathcal{L}_d(S^*)$:

$$\ell(x) = \sigma_0 + \sum_{i \in I} \sigma_i g_i^* + \sum_{j \in J} \sigma_j g_j^*,$$

by evaluating at 0, $\sigma_0(0) = \sigma_j(0) = 0$.

# Sketch of proof

Assume $a = \mathbf{0}$, $L = x_1$-axis, $S \cap L \subseteq \mathbb{R}_+$.



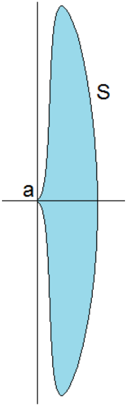

- Let $g_i^*(x) = g_i(x, 0, \ldots, 0)$, and $S^* = \{x \in \mathbb{R} : g_i^*(x) \geq 0, i = 1, ..., k\}$.
- Note that $\mathcal{L}_d(S^*) \subseteq \mathcal{L}_d(S) \cap x_1$-axis.
- $(g_i^*)'(0) = g_i^*(0) = 0$ for $i \in I$.
- $\sigma_0(0) = \sigma_j(0) = 0$ for all $j \in J$.
- Suppose $\ell(x) = x$ was valid in $\mathcal{L}_d(S^*)$:

$$\ell(x) = \sigma_0 + \sum_{i \in I} \sigma_i g_i^* + \sum_{j \in J} \sigma_j g_j^*,$$

by evaluating at 0, $\sigma_0(0) = \sigma_j(0) = 0$.

# Sketch of proof

Assume $a = \mathbf{0}$, $L = x_1$-axis, $S \cap L \subseteq \mathbb{R}_+$.

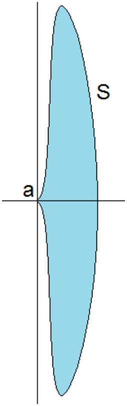



- Let $g_i^*(x) = g_i(x, 0, \ldots, 0)$, and $S^* = \{x \in \mathbb{R} : g_i^*(x) \geq 0, i = 1, ..., k\}$.
- Note that $\mathcal{L}_d(S^*) \subseteq \mathcal{L}_d(S) \cap x_1$-axis.
- $(g_i^*)'(0) = g_i^*(0) = 0$ for $i \in I$.
- $\sigma_0(0) = \sigma_j(0) = 0$ for all $j \in J$.
- Suppose $\ell(x) = x$ was valid in $\mathcal{L}_d(S^*)$:

$$1 = (\sigma_0)' + \sum_{l=1,\ldots,k} ((\sigma_l)' g_l^* + \sigma_l (g_l^*)'),$$

# Sketch of proof

Assume $a = \mathbf{0}$, $L = x_1$-axis, $S \cap L \subseteq \mathbb{R}_+$.




- Let $g_i^*(x) = g_i(x, 0, \ldots, 0)$, and $S^* = \{x \in \mathbb{R} : g_i^*(x) \geq 0, i = 1, ..., k\}$.
- Note that $\mathcal{L}_d(S^*) \subseteq \mathcal{L}_d(S) \cap x_1$-axis.
- $(g_i^*)'(0) = g_i^*(0) = 0$ for $i \in I$.
- $\sigma_0(0) = \sigma_j(0) = 0$ for all $j \in J$.
- Suppose $\ell(x) = x$ was valid in $\mathcal{L}_d(S^*)$:

$$1 = (\sigma_0)' + \sum_{l=1,\ldots,k} ((\sigma_l)' g_l^* + \sigma_l (g_l^*)'),$$

by evaluating at 0, $1 = 0$. $\square$

# Singularities

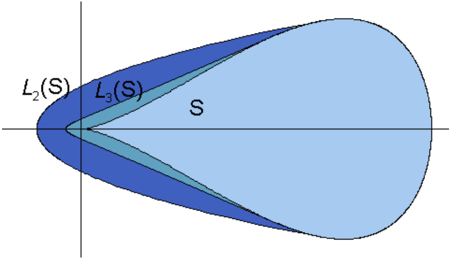
**Corollary**

*If $S$ has non-empty interior and there exists a point $a \in S$ that is on the boundary of $\text{conv}(S)$ s.t. all $g_i$ verifying $g_i(a) = 0$ are singular at $a$, then we have $\mathcal{L}_d(S) \neq \overline{\text{conv}(S)}$ for all $d$.*

# Singularities

## Corollary

*If $S$ has non-empty interior and there exists a point $a \in S$ that is on the boundary of* $\mathrm{conv}(S)$ *s.t. all $g_i$ verifying $g_i(a) = 0$ are singular at $a$, then we have $\mathcal{L}_d(S) \neq \overline{\mathrm{conv}(S)}$ for all $d$.*




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

# Faces of a convex set

## Definition

- A face of a convex set $S$ is a convex subset $F \subseteq S$ s.t. if $a, b \in S$ and $]a, b[ \cap F \neq \emptyset$ then $a, b \in F$.

# Faces of a convex set

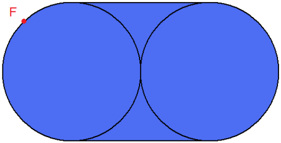
## Definition

- A face of a convex set $S$ is a convex subset $F \subseteq S$ s.t. if $a, b \in S$ and $]a, b[ \cap F \neq \emptyset$ then $a, b \in F$.
- A face $F$ is said to be exposed if there exists an hyperplane $H$ such that $F = H \cap S$.

# Faces of a convex set

## Definition

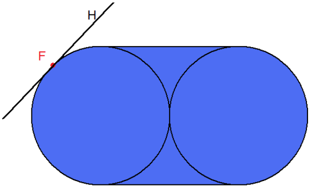
- A face of a convex set $S$ is a convex subset $F \subseteq S$ s.t. if $a, b \in S$ and $]a, b[ \cap F \neq \emptyset$ then $a, b \in F$.
- A face $F$ is said to be exposed if there exists an hyperplane $H$ such that $F = H \cap S$.



# Faces of a convex set

## Definition

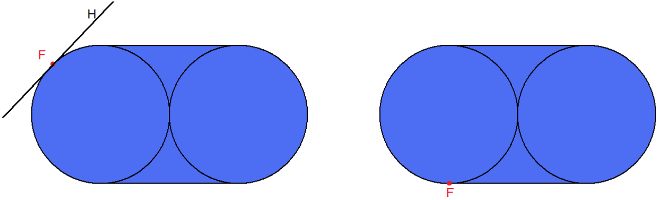
- A face of a convex set $S$ is a convex subset $F \subseteq S$ s.t. if $a, b \in S$ and $]a, b[ \cap F \neq \emptyset$ then $a, b \in F$.
- A face $F$ is said to be exposed if there exists an hyperplane $H$ such that $F = H \cap S$.



# Faces of a convex set

## Definition

- A face of a convex set $S$ is a convex subset $F \subseteq S$ s.t. if $a, b \in S$ and $]a, b[ \cap F \neq \emptyset$ then $a, b \in F$.
- A face $F$ is said to be exposed if there exists an hyperplane $H$ such that $F = H \cap S$.



# Non-exposed faces

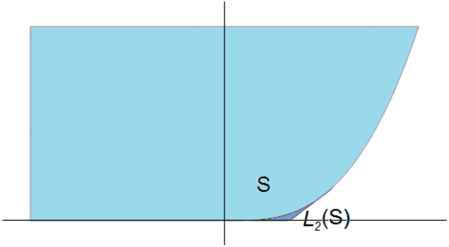
**Corollary (Netzer-Plaumann-Schweighofer)**

*Suppose S is convex and has non-empty interior. If S has a non-exposed face then* $\mathcal{L}_d(S) \neq \overline{\mathrm{conv}(S)}$ *for all* $d$.

# Non-exposed faces

**Corollary (Netzer-Plaumann-Schweighofer)**

*Suppose S is convex and has non-empty interior. If S has a non-exposed face then* $\mathcal{L}_d(S) \neq \overline{\text{conv}(S)}$ *for all d.*




$$g_1(x,y) = y - x^3,\ g_2(x,y) = y,\ g_3(x,y) = x+1,\ g_4(x,y) = 1-y$$

# Questions from NPS

**Question 1**

Can $\mathcal{L}_d(S)$ have non-exposed faces?

# Questions from NPS

**Question 1**

Can $\mathcal{L}_d(S)$ have non-exposed faces?

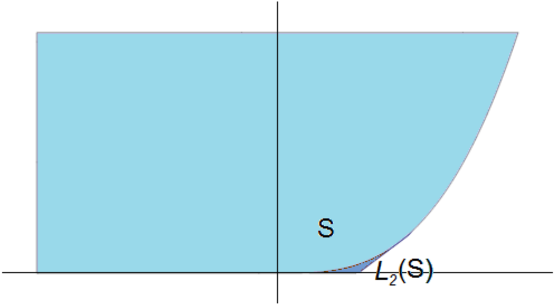
Yes.

# Questions from NPS

**Question 1**

Can $\mathcal{L}_d(S)$ have non-exposed faces?

Yes.

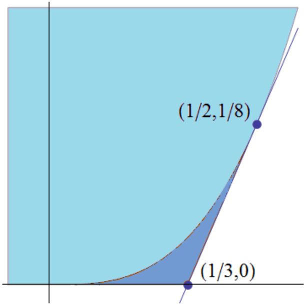


# Questions from NPS

## Question 1

Can $\mathcal{L}_d(S)$ have non-exposed faces?

Yes.



# Questions from NPS

## Question 2

If we drop convexity of $S$ and just assume $\text{conv}(S)$ has non-exposed faces can we reach the same conclusion?

# Questions from NPS

**Question 2**

If we drop convexity of $S$ and just assume $\mathrm{conv}(S)$ has non-exposed faces can we reach the same conclusion?

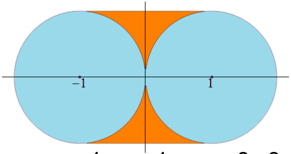
No.

# Questions from NPS

## Question 2

If we drop convexity of $S$ and just assume $\text{conv}(S)$ has non-exposed faces can we reach the same conclusion?

No.



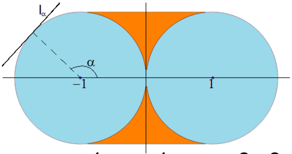

$$g(x,y) = -x^4 - y^4 - 2x^2y^2 + 4X^2$$

# Questions from NPS

## Question 2

If we drop convexity of $S$ and just assume conv$(S)$ has non-exposed faces can we reach the same conclusion?

No.




$$g(x, y) = -x^4 - y^4 - 2x^2y^2 + 4X^2$$

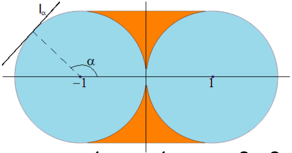
$$l_\alpha = 1 - cos(\alpha) - cos(\alpha)x - sin(\alpha)Y$$

# Questions from NPS

## Question 2

If we drop convexity of $S$ and just assume conv($S$) has non-exposed faces can we reach the same conclusion?

No.




$$g(x,y) = -x^4 - y^4 - 2x^2y^2 + 4X^2$$

$$l_\alpha = 1 - cos(\alpha) - cos(\alpha)x - sin(\alpha)Y$$

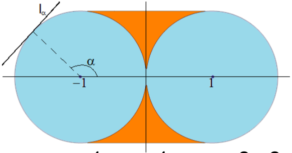
$(8-8\cos(\alpha))l_\alpha =$

# Questions from NPS

## Question 2

If we drop convexity of $S$ and just assume conv($S$) has non-exposed faces can we reach the same conclusion?

No.




$$g(x,y) = -x^4 - y^4 - 2x^2y^2 + 4X^2$$

$$l_\alpha = 1 - cos(\alpha) - cos(\alpha)x - sin(\alpha)Y$$

$$(8-8\cos(\alpha))l_\alpha = g(x,y) + (x^2+y^2-2+2\cos(\alpha))^2 +$$
$$+\left(2\sqrt{1-\cos(\alpha)}(y-\sin(\alpha))\right)^2 + \left(2\sqrt{-\cos(\alpha)}(x-\cos(\alpha)+1)\right)^2$$

# Open Questions

**Question 1**

What other obstructions exist to the finite convergence of the Lasserre hierarchy?

# Open Questions

### Question 1

What other obstructions exist to the finite convergence of the Lasserre hierarchy?

### Question 2

Give an example of a smooth, convex, basic semialgebraic set with no non-exposed faces such that the Lasserre hierarchy does not converge finitely. Is there such an example?

# Reference

**Positive Polynomials and Projections of Spectrahedra**,
João Gouveia and Tim Netzer, *arXiv:0911.2750*

# The End

Thank You