Fundamentals of Convergence Theories
for Convex Relaxation Hierarchies

Levent Tunçel

Dept. of Combinatorics and Optimization,

Faculty of Mathematics,

University of Waterloo,

Canada.

`ltuncel@math.uwaterloo.ca`

September 24, 2009

1 Semidefinite Optimization

Let $\mathbb{S}^n$ denote the space of n-by-n symmetric matrices with entries in $\mathbb{R}$.

Definition 1.1 *Let* $X \in \mathbb{S}^n$*.*

X *is positive semidefinite if*

$$h^T X h \geq 0, \quad \forall h \in \mathbb{R}^n.$$

X *is positive definite if*

$$h^T X h > 0, \quad \forall h \in \mathbb{R}^n \setminus \{0\}.$$

For $A, B \in \mathbb{S}^n$, we use the trace inner-product:

$$\langle A, B \rangle := \text{Tr}(A^T B),$$

we write

$$A \succeq B$$

to mean $(A - B)$ is positive semidefinite;

$$A \succ B$$

to mean $(A - B)$ is positive definite.

Note that for $X \in \mathbb{S}^n$ all eigenvalues $\lambda_j(X)$ are real. Also,

$$X \succeq 0 \iff \lambda(X) \geq 0$$

and

$$X \succ 0 \iff \lambda(X) > 0.$$

Let $\mathcal{A} : \mathbb{S}^n \longrightarrow \mathbb{R}^m$ be a given linear transformation, $b \in \mathbb{R}^m$, $C \in \mathbb{S}^n$ are also given.

$\mathcal{A}^* : \mathbb{R}^m \longrightarrow \mathbb{S}^n$ denotes the adjoint of $\mathcal{A}$ and is defined by:

$$\langle \mathcal{A}^*(y), X \rangle = [\mathcal{A}(X)]^T y, \quad \forall X \in \mathbb{S}^n, \forall y \in \mathbb{R}^m.$$

$$\begin{array}{llrcl} (P) & \text{inf} & \langle C, X\rangle & & \\ & & \mathcal{A}(X) & = & b, \\ & & X & \succeq & 0, \end{array}$$

$$\begin{array}{llrcl} (D) & \text{sup} & b^T y & & \\ & & \mathcal{A}^*(y) & \preceq & C. \end{array}$$

Equivalently, for $A_1, A_2, \ldots, A_m \in \mathbb{S}^n$, we have

$$\begin{array}{lrcl} (P) \quad \text{inf} & \langle C, X \rangle & & \\ & \langle A_i, X \rangle & = & b_i, \quad \forall i \in \{1, 2, \ldots, m\} \\ & X & \succeq & 0, \end{array}$$

$$\begin{array}{lrcl} (D) \quad \text{sup} & b^T y & & \\ & \sum_{i=1}^{m} y_i A_i & \preceq & C. \end{array}$$

For SDP: Slater point $\bar{y} \in \mathbb{R}^m$ such that

$$\mathcal{A}^*(\bar{y}) \prec C.$$

Theorem 1.1 *(A Strong Duality Theorem) Suppose (D) has a Slater point. If the objective value of (D) is bounded from above then (P) attains its optimum value and the optimum values of (P) and (D) coincide.*

We can solve such semidefinite optimization problems

efficiently

both in terms of

- computational complexity theory and
- practical computation.

However, in terms of both computational complexity theory and practical computation the situtation in SDP is much worse than that of LP. (That is, there are very many deep, open problems.)

Why do we need extra assumptions for the duality theorem of SDP?

Consider the examples (with parameter $\gamma > 0$):

$$C := \begin{pmatrix} \gamma & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, b := \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

$$A_1 := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, A_2 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Then for every feasible solution of (D), we have $y_2 = 0$ and $y_1 \leq 0$. So, the set of feasible solutions of (D) (in the S-space) is

$$\left\{ \begin{pmatrix} \gamma & 0 & 0 \\ 0 & S_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix} : S_{22} \geq 0 \right\}.$$

Moreover, every feasible solution of (D) is optimal in (D).

For the primal, it is also true that the set of feasible solutions and the set of optimal solutions coincide, which is (in the X-space):

$$\left\{ \begin{pmatrix} 1 & 0 & X_{31} \\ 0 & 0 & 0 \\ X_{31} & 0 & X_{33} \end{pmatrix} : X_{33} \geq X_{31}^2 \right\} .$$

Even though both (P) and (D) have finite optimal objective values, both of these values are attained, there is a duality gap of γ.

$\{\alpha v : \alpha \in \mathbb{R}_+\}$ for $v \in K \setminus \{0\}$ defines a ray inside K. Such a ray $R \subseteq K$ is called an extreme ray of K if for every pair of rays $R_1, R_2 \subseteq K$, such that $R_1 + R_2 \supseteq R$ implies either $R_1 = R$ or $R_2 = R$ possibly both. The union of all extreme rays of K is denoted by Ext(K). We also use ext(K) to denote the set of normalized extreme rays of K (where each ray is represented by a single nonzero element of K). For compact, convex sets, we use the notation ext$(\cdot)$ to denote the set of extreme points of the compact, convex set.

A convex set is called pointed if it does not contain any lines.

2 Carathéodory's Theorem revisited for convex cones:

Theorem 2.1 *Let* $K \subseteq \mathbb{R}^n$ *be a pointed closed convex cone. Then for every* $\bar{x} \in K$*, there exist* $u^{(1)}, u^{(2)}, \ldots, u^{(n)} \in \text{ext}(K)$ *and* $\lambda \in \mathbb{R}^n_+$ *such that*

$$\bar{x} = \sum_{j=1}^{n} \lambda_j u^{(j)}.$$

Example: Let $K := \mathbb{S}^n_+$. Then even though $\text{dim}(K) = \frac{n(n+1)}{2}$, n extreme rays of $\mathbb{S}^n_+$ suffice (Schur/eigenvalue decomposition of $\bar{x}$).

Faces of $\mathbb{S}^n_+$:

Let $K \subseteq \mathbb{R}^d$ be a closed convex cone. A convex cone $G \subseteq K$ is a face of K if for every $u, v \in K$ such that $(u+v) \in G$, we have $u \in G$, $v \in G$.

A face G of K is exposed if there exists $a \in \mathbb{R}^d \setminus \{0\}$ such that

$$G = \{x \in K : \langle a, x\rangle = 0\} \text{ and } K \subseteq \left\{x \in \mathbb{R}^d : \langle a, x\rangle \leq 0\right\},$$

i.e., G is the intersection of K with one of its supporting hyperplanes.

A face G of K is a proper face of K if

$$\{0\} \subset G \subset K.$$

Theorem 2.2 *(a) Every proper face G of $\mathbb{S}^n_+$ is characterized by a unique subspace $L \subset \mathbb{R}^n$ such that*

$$G = \left\{X \in \mathbb{S}^n_+ : \text{Null}(X) \supseteq L\right\},$$

$$\text{relint}(G) = \left\{X \in \mathbb{S}^n_+ : \text{Null}(X) = L\right\}.$$

(b) Every proper face G of $\mathbb{S}^n_+$ is exposed.

Every proper face of $\mathbb{S}^n_+$ is isomorphic to $\mathbb{S}^k_+$ for some $k < n$.

E.g.,

$$\text{ext}(\mathbb{S}^n_+) = \left\{hh^T : h \in \mathbb{R}^n, \|h\|_2 = 1\right\}.$$

3 Helly's Theorem revisited for convex sets:

Theorem 3.1 *Let* $n \geq 2$, $r \geq n+1$ *and* $G_1, G_2, \ldots, G_r$ *be convex sets in* $\mathbb{R}^n$. *Then,* $\bigcap_{k=1}^{r} G_k \neq \emptyset$ *iff*

$$\bigcap_{k \in J} G_k \neq \emptyset, \text{ for all } J \subseteq \{1, 2, \ldots, r\} \text{ such that } |J| = n+1.$$

However, consider the following collection of closed halfspaces

$$G_k := \left\{ x \in \mathbb{R}^n : \sum_{j=1}^{n} x_j \geq k \right\}, \quad k \in \mathbb{Z}_{++}.$$

Then every finite collection of $\{G_k\}$ has a nonempty intersection; but,

$$\bigcap_{k=1}^{\infty} G_k = \emptyset.$$

However, consider the following collection of closed halfspaces

$$G_k := \left\{ x \in \mathbb{R}^n : \sum_{j=1}^{n} x_j \geq k \right\}, \quad k \in \mathbb{Z}_{++}.$$

Then every finite collection of $\{G_k\}$ has a nonempty intersection; but,

$$\bigcap_{k=1}^{\infty} G_k = \emptyset.$$

However^2,

Theorem 3.2 *Let $n \geq 2$, and $\mathcal{G}$ be a collection of compact convex sets in $\mathbb{R}^n$. Then, the intersection of the sets in $\mathcal{G}$ is nonempty iff every finite subset of $\mathcal{G}$ of size at most $(n+1)$ has nonempty intersection.*

In particular,

Theorem 3.3 *Let $n \geq 2$, and $\mathcal{G}$ be a collection of closed convex sets in $\mathbb{R}^n$. Suppose one of the sets in $\mathcal{G}$ is bounded (call it G_0). Then, the intersection of the sets in $\mathcal{G}$ is nonempty iff the intersection of every finite subset of $\mathcal{G}$ of size at most $(n+1)$ has nonempty intersection with the bounded set G_0.*

4 A Successive SDP Relaxation Method

Let $F \subset \mathbb{R}^n$ be a compact set. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function. Then the optimization problem

$$\max \{f(x) : x \in F\}$$

can be reduced to

$$\max \{x_{n+1} : f(x) \geq x_{n+1}, x \in F, \ell_{n+1} \leq x_{n+1} \leq u_{n+1}\}.$$

So, without loss of generality, we may assume that the objective function is linear. So, our problem is

$$\max \left\{ c^T x : x \in F \right\},$$

for a given vector c. Now, as in the special case of combinatorial optimization, we have

$$\max \left\{ c^T x : x \in F \right\} = \max \left\{ c^T x : x \in \text{conv}(F) \right\}.$$

We denote by D_1 the n-dimensional unit hypersphere (in the context of the next theorem this is the set of all nontrivial objective functions).

Theorem 4.1 *The set of all $c \in D_1$ for which the optimal solution of*

$$\max\left\{c^T x : x \in \text{conv}(F)\right\}$$

is not unique has zero, $(n-1)$-dimensional Hausdorff measure.

If we are able to get our hands on a convex compact set G and $G = \text{conv}(F)$, then for almost all objective function vectors c, the problem

$$\max\left\{c^T x : x \in G\right\}$$

will have a unique minimizer $\hat{x}$. Thus, $\hat{x} \in \text{ext}(G) \subseteq F$. Therefore, $\hat{x}$ will also be an optimal solution of our original (non-convex) problem

$$\max\left\{c^T x : x \in F\right\}.$$

How do we “compute” the convex hull?

First, let's choose a simple but general enough algebraic representation:

Lemma 4.1 *Every compact set $F \subset \mathbb{R}^n$ admits a representation as the feasible region of a system of quadratic inequalities.*

For this lemma, assuming F is closed is enough.

Proof: F is closed. So, $\mathbb{R}^n \setminus F$ is open in $\mathbb{R}^n$; thus, it can be written as the union of (possibly uncountably many) open balls

$$\bigcup_{\bar{x},R} \{x \in \mathbb{R}^n : \|x - \bar{x}\|_2^2 < R^2\} = \mathbb{R}^n \setminus F.$$

By taking the complement of both sides, we express F as the solution set of a system of quadratic inequalities.

$$F = \bigcap_{\bar{x},R} \{x \in \mathbb{R}^n : \|x - \bar{x}\|_2^2 \geq R^2\}.$$

■

Lemma 4.2 *For every* (Q, q, γ) *of appropriate size,*

$$\{x \in \mathbb{R}^n : x^T Q x + 2q^T x + \gamma \le 0\}$$
$$\subseteq \left\{ x : \left\langle \begin{bmatrix} \gamma & q^T \\ q & Q \end{bmatrix}, \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \right\rangle \le 0, \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \in \mathbb{S}^{n+1}_+ \right\}.$$

Equality holds above if $\begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix}$ *is forced to be rank one.*

We have the following representation theorem.

Theorem 4.2 *(Kojima, T. [2000]) Let $\mathcal{P}$ be a closed convex cone (containing $(I, 0, R)$). Then the convex sets*

$$\{x \in \mathbb{R}^n : x^T Q x + 2q^T x + \gamma \le 0, \forall (Q, q, \gamma) \in (\mathcal{P} \cap \mathcal{Q}_+)\}$$

and

$$\left\{x \in \mathbb{R}^n : \left\langle \begin{bmatrix} \gamma & q^T \\ q & Q \end{bmatrix}, \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \right\rangle \le 0, \forall (Q, q, \gamma) \in \mathcal{P}; \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \in \mathbb{S}^{n+1}_+ \right\}$$

are identical. Moreover, in the second description, we can replace $\mathcal{P}$ by its generators. Here, $\mathcal{Q}_+ := \{(Q, q, \gamma) : Q \in \mathbb{S}^n_+\}$.

Let $\mathcal{P}_F$ denote the given quadratic inequality representation of F (i.e., the input data).

Start by setting the first convex relaxation for conv(F) as

$$G_0 := \left\{ x : \langle \begin{bmatrix} \gamma & q^T \\ q & Q \end{bmatrix}, \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \rangle \le 0, \forall (Q, q, \gamma) \in \mathcal{P}_F; \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \in \mathbb{S}_+^{n+1} \right\}$$

How are the new inequalities generated?

Recall that D_1 denotes the unit hypersphere in $\mathbb{R}^n$. Let

$$D_2 := \left\{ v^{(i)}, -v^{(i)} : i \in \{1, 2, \ldots, n\} \right\},$$

$\left\{ v^{(i)} : i \in \{1, 2, \ldots, n\} \right\}$ is any basis for $\mathbb{R}^n$.

Compute

$$\alpha := \min \left\{ a^T x : x \in G_k \right\}$$

and

$$\beta := \min \left\{ b^T x : x \in G_k \right\}$$

for every $a \in D_1$ and for every $b \in D_2$. Then the quadratic inequality

$$\left(a^T x - \alpha \right) \left(b^T x - \beta \right) \geq 0$$

is a valid quadratic inequality for G_k and hence a valid quadratic inequality for F.

$$\left(a^T x - \alpha\right)\left(b^T x - \beta\right) \geq 0$$

Denote by $\mathcal{P}_k$ the set of all such valid quadratic inequalities. We define G_{k+1} by the convex relaxation where $\mathcal{P}_F$ is replaced by $\mathcal{P}_F \cup \mathcal{P}_k$ (or generators of the cone $(\mathcal{P}_F \cup \mathcal{P}_k)$). This gives a very general method (SSDPR) Successive SDP Relaxation method.

$$G_{k+1} := \left\{ x : \left\langle \begin{bmatrix} \gamma & q^T \\ q & Q \end{bmatrix}, \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \right\rangle \leq 0, \forall (Q, q, \gamma) \in (\mathcal{P}_F \cup \mathcal{P}_k)\,; \right.$$

$$\left. \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \in \mathbb{S}_+^{n+1} \right\}$$

Theorem 4.3 *(Kojima, T. [2000]) With the above definitions, the sequence of convex compact sets G_k generated by SSDPR method satisfies*

(a) *for every* $k \geq 0$, $\mathrm{conv}(F) \subseteq G_{k+1} \subseteq G_k$ *(monotonicity), in fact* $G_{k+1} = G_k$ *iff* $G_k = \mathrm{conv}(F)$*;*

(b) $\displaystyle\bigcap_{k=1}^{k^*} G_k = \emptyset$ *for some finite number* k^* *if* $F = \emptyset$ *(detecting infeasibility in finitely many steps);*

(c) $\displaystyle\bigcap_{k=1}^{\infty} G_k = \mathrm{conv}(F)$ *(asymptotic convergence).*

Proof: (Proof of (c))

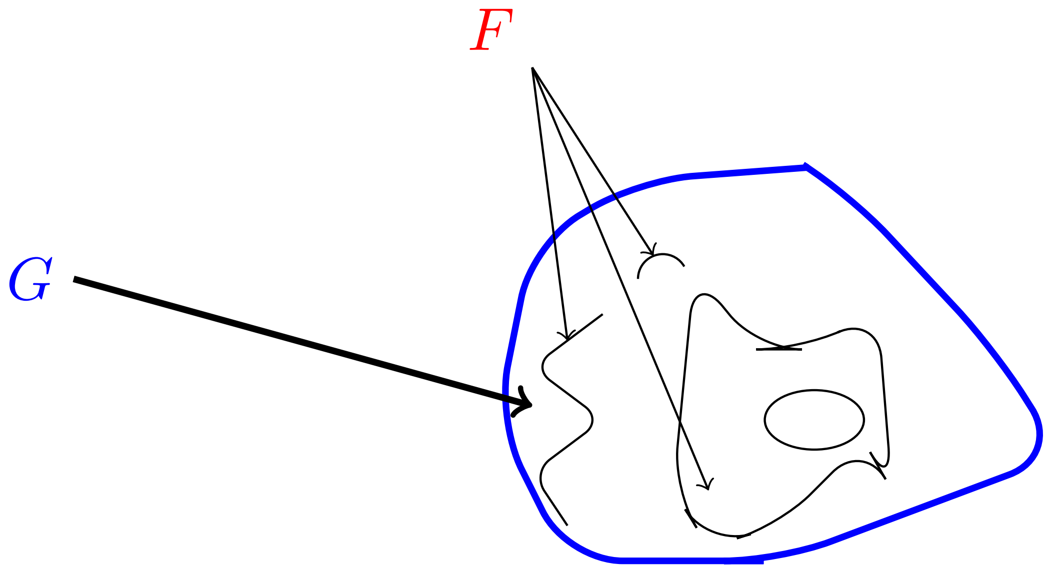

Figure 1: Suppose $G_k \to G \neq$ conv(F) (seeking a contradiction)

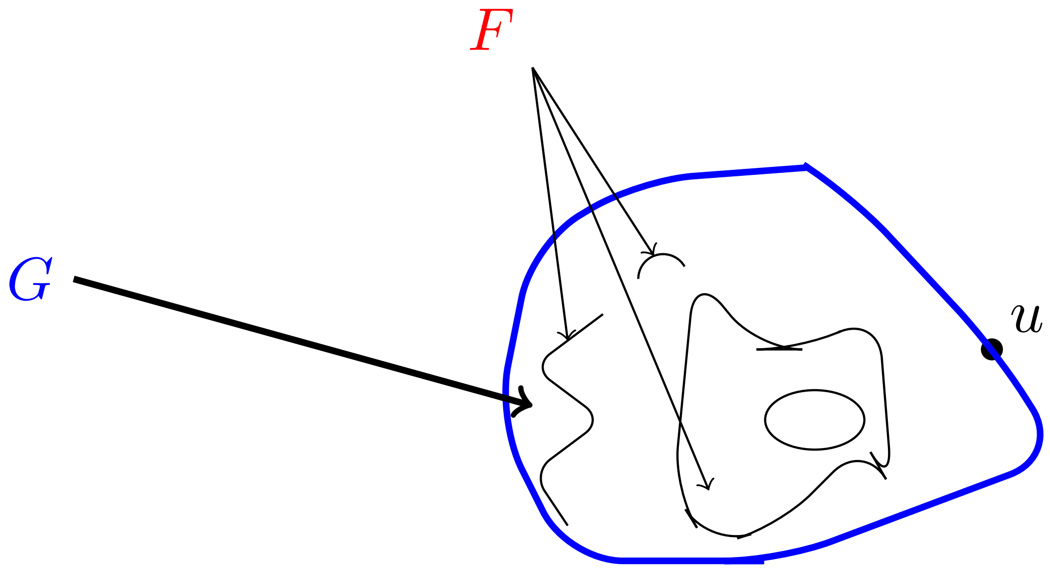

Figure 2: Then, $\exists u \in G \setminus$ conv(F)

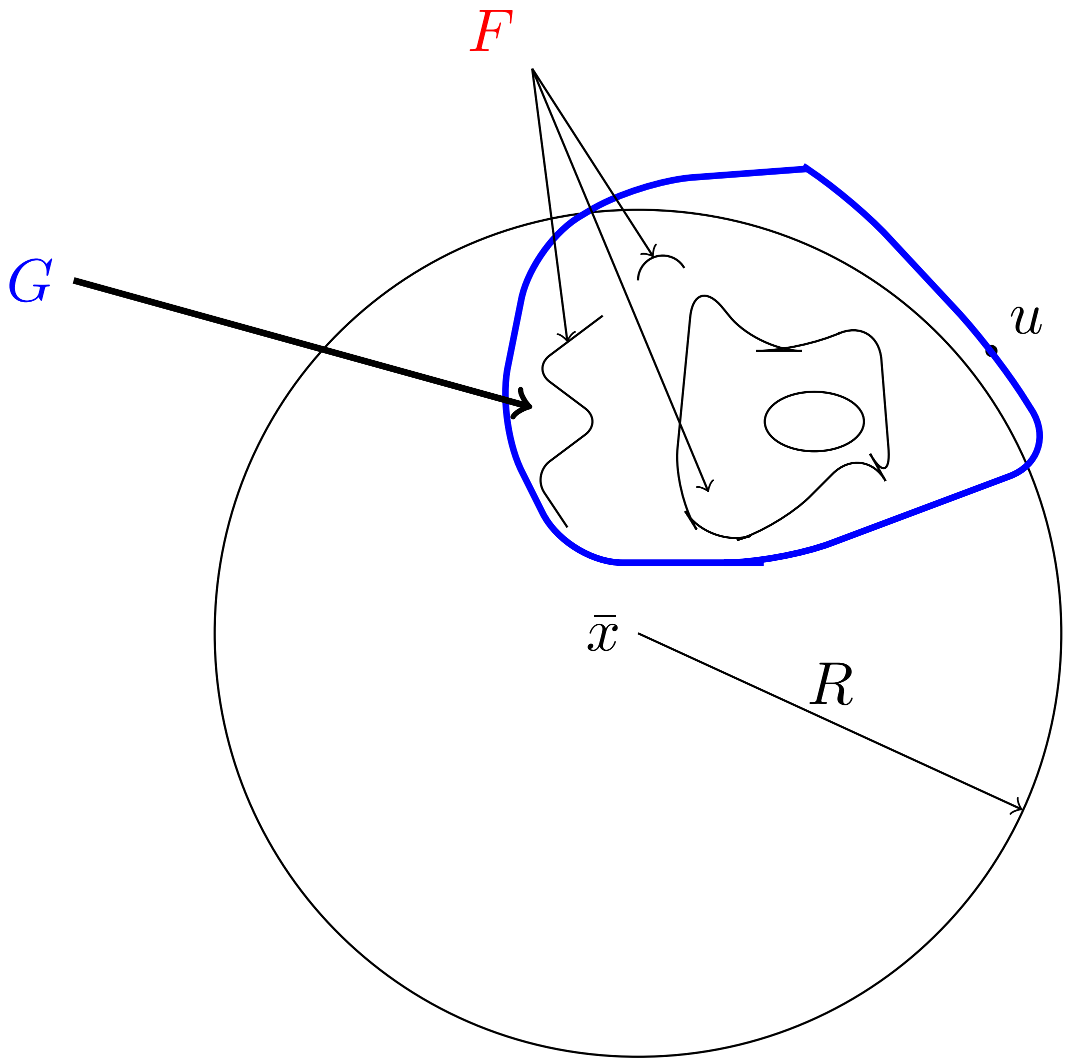

Figure 3: There exists a separating hypersphere for u, conv(F)

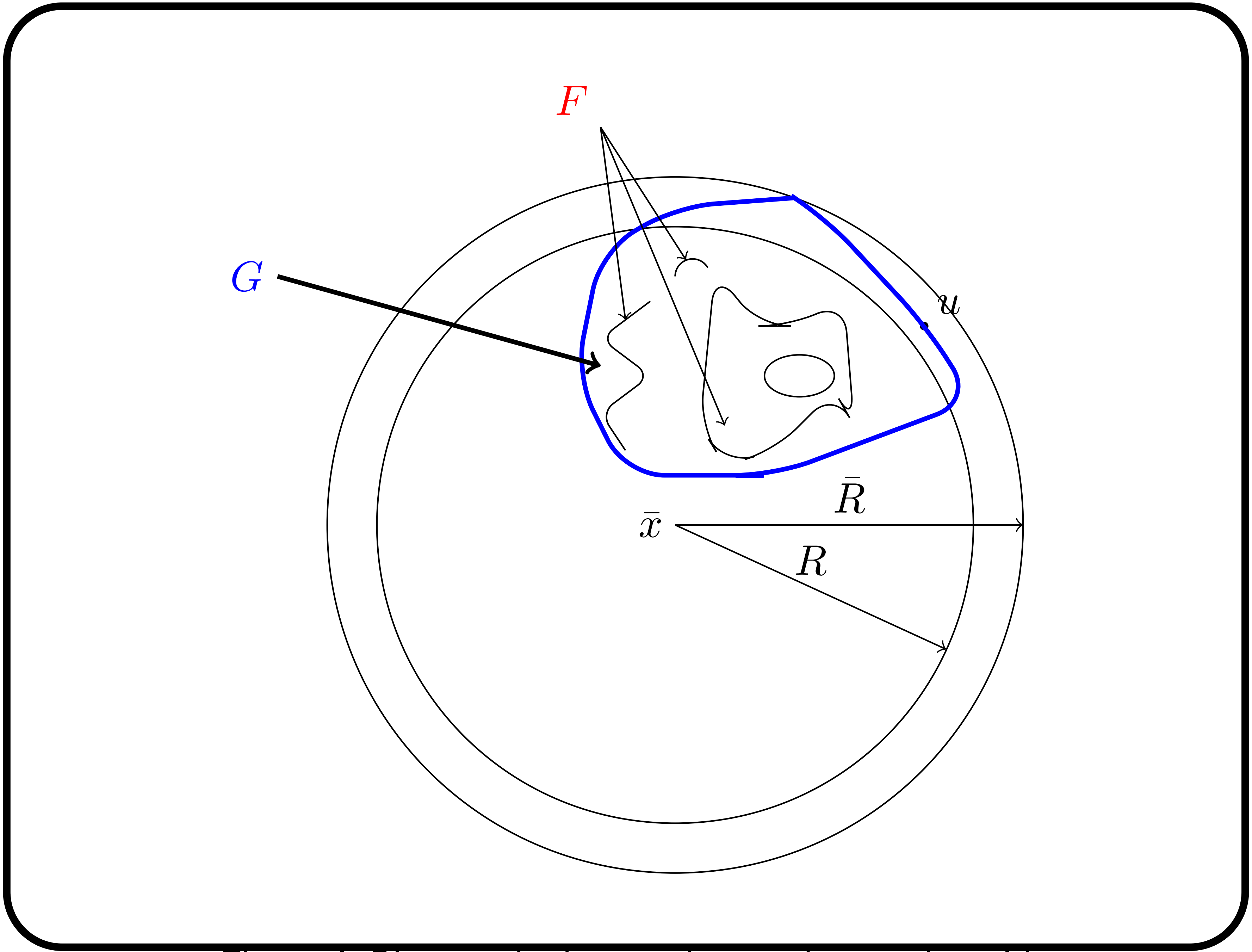

Figure 4: Blow-up the hypersphere to just enclose G

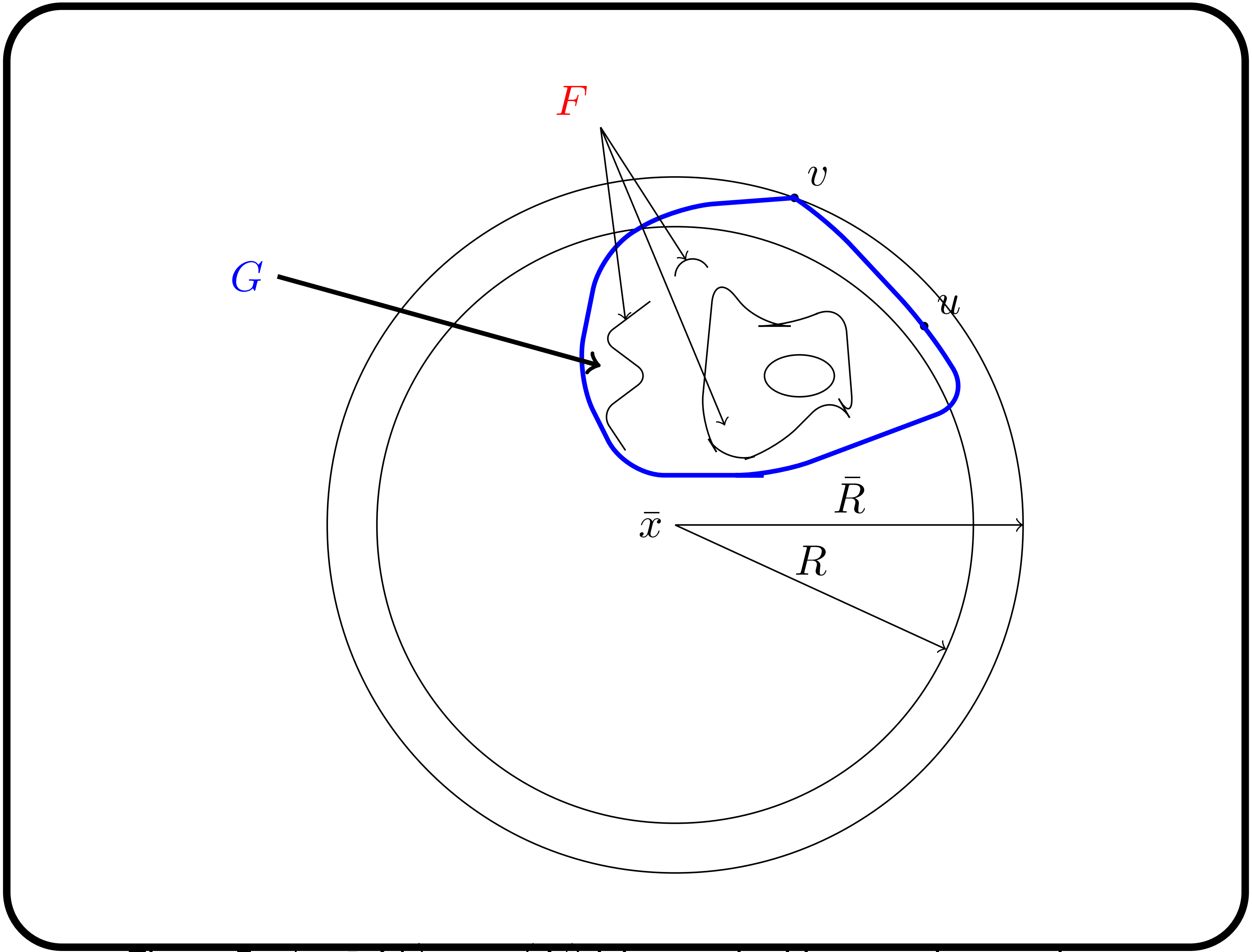

Figure 5: $\exists v \in G \setminus$ conv(F) lying on the blown-up hypersphere

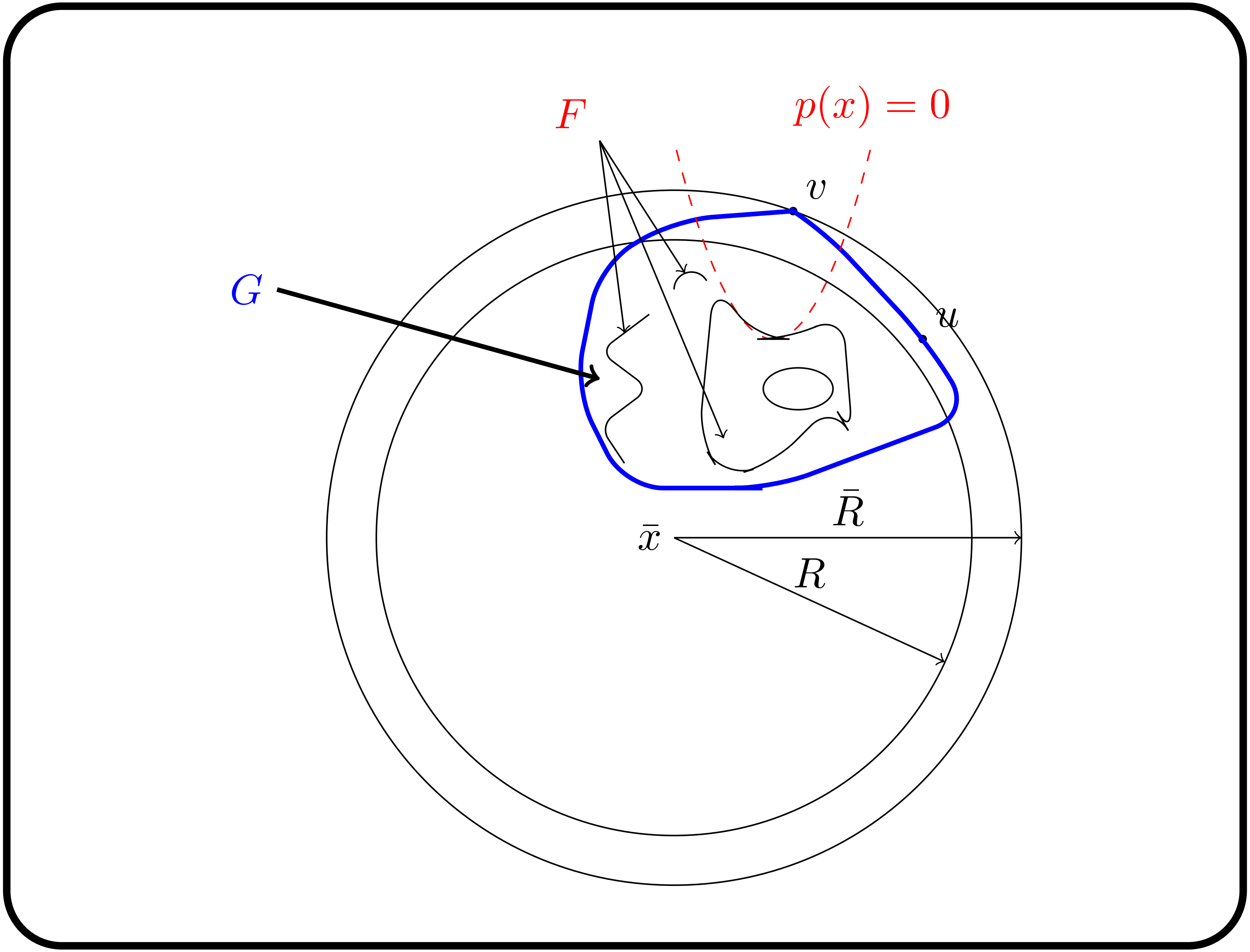

Figure 6: $\exists$ original constraint $p(x) \geq 0$ violated by v

■

5 Applications to Systems of Polynomial Inequalities

Since any system of polynomial inequalities can be reformulated as a system of quadratic inequalities, the above results can be translated to the setting of polynomial optimization problems (POP):

$$\begin{array}{rcl} \inf \quad p_0(x) & & \\ p_i(x) & \geq & 0, \quad i \in \{1, 2, \ldots, m\}, \end{array}$$

where $p_0, p_1, \ldots, p_m : \mathbb{R}^n \to \mathbb{R}$ are polynomials.

Since any system of polynomial inequalities can be reformulated as a system of quadratic inequalities, e.g., consider the system

$$\begin{aligned} x_1^4 x_2^2 + x_2^3 x_3 + x_1^5 - 1 &\geq 0, \\ 2x_1^3 - x_2^4 &\geq 0 \end{aligned}$$

$$\begin{aligned} x_1^4x_2^2 + x_2^3x_3 + x_1^5 - 1 &\geq 0, \\ 2x_1^3 - x_2^4 &\geq 0 \end{aligned}$$

which is equivalent to the quadratic system:

$$\begin{aligned} x_5x_6 + x_6x_7 + x_1x_5 - 1 &\geq 0, \\ 2x_1x_4 - x_6^2 &\geq 0, \\ x_4 &= x_1^2, \\ x_5 &= x_4^2, \\ x_6 &= x_2^2, \\ x_7 &= x_2x_3, \end{aligned}$$

...

Since any system of polynomial inequalities can be reformulated as a system of quadratic inequalities, the above results can be translated to the setting of polynomial optimization problems (POP):

$$\begin{aligned} \min \quad & p_0(x) \\ & p_i(x) \quad \geq \quad 0, \quad i \in \{1, 2, \ldots, m\}, \end{aligned}$$

where $p_0, p_1, \ldots, p_m : \mathbb{R}^n \to \mathbb{R}$ are polynomials.

There is a lot of work in the area of solving (POP) by utilizing Linear Optimization, Semidefinite Optimization and Convex Optimization techniques. See Lasserre [2001-...], Parrilo [2003-...], Laurent [2003-...], de Klerk and Pasechnik [2002-...], Peña, Vera and Zuluaga [2007], Gouveia, Parrilo, Thomas [2009], Anjos and Vera [2009].

Lasserre uses the connections to Putinar's Theorem [1993]:

Theorem 5.1 *Suppose* $F := \{x \in \mathbb{R}^n : p_i(x) \geq 0, i \in \{1, 2, \ldots, m\}\}$ *is compact, the polynomials* p_i *have even degree, and their highest degree homogeneous parts do not have common zeroes in* $\mathbb{R}^n$ *except 0. Then every polynomial that is positive on* F *can be written as a nonnegative combination of polynomials of the form*

$$[h_0(x)]^2 + \sum_{i=1}^{m} [h_i(x)]^2 \, p_i(x).$$

Perhaps one of the most fundamental problems here is the *K-moment problem* which is, given $K \subset \mathbb{R}^n$ to decide when a real valued function f of set of monomials in n variables is a moment function $\int_K x^m d\mu$ for some nonnegative Borel measure μ on K. Schmüdgen [1991] characterized the solutions to the K-moment problem (called *K-moment sequences*) for all compact semi-algebraic sets K in terms of the positive definiteness of matrices arising from the moment functions. Schmüdgen's proof utilizes Positivstellensatz in proving the above-mentioned algebraic fact. The result also generalizes many other preexisting beautiful results such as Handelman's Theorem [1988]; some of these connections are old and they generalize results some of which go all the way back to Minkowski in the late 1800's.

Positivstellensatz (Stengle [1974]):

$F := \{x \in \mathbb{R}^n : p_i(x) \geq 0, i \in \{1, 2, \ldots, m\}\} = \emptyset$ iff there exist $s_0, s_J, \ldots \in$ SoS$(n, *)$ such that

$$\sum_{J \subseteq \{1,2,\ldots,m\}} s_J \prod_{i \in J} p_i = -1.$$

Parrilo uses the Positivstellensatz and gets SoS certificates!

Given $x \in \mathbb{R}^n$ and polynomial $f : \mathbb{R}^n \to \mathbb{R}$ of degree $2d$, let

$$h(x) := [1, x_1, x_2, \ldots, x_n, x_1^2, x_1x_2, x_1x_3, \ldots, x_2^2, \ldots, x_n^d]^T \in \mathbb{R}^N,$$

where $N := \binom{n+d}{d}$. We are interested in

$$\mathcal{F}(f) := \left\{ X \in \mathbb{S}^N : [h(x)]^T X h(x) = f(x) \right\}.$$

The following well-known fact connects SoS and semidefinite optimization.

Theorem 5.2 *Let $\bar{z} \in \mathbb{R}$. Then $[f(x) - \bar{z}]$ is SoS iff*

$$\{X \in \mathcal{F}(f) : X \succeq \bar{z} e_1 e_1^T\} \neq \emptyset.$$

Let Pos(n, d) denote the set of nonnegative polynomials of degree d in n variables.

Let SoS(n, d) denote the set of polynomials (of n variables) of degree d that are sums of squares.

Proposition 5.1 (Hilbert, 1888) *For every $n \geq 1$,*

$$\text{Pos}(n, 2) = \text{SoS}(n, 2).$$

Suppose $F = \emptyset$. Then, (by Kojima-T. convergence theorem) for some finite k,

$$G_k = \left\{ x : \langle \begin{bmatrix} \gamma & q^T \\ q & Q \end{bmatrix}, \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \rangle \geq 0, \forall (Q, q, \gamma) \in \mathcal{P}; \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \in \mathbb{S}^{n+1}_{+} \right\}$$

$$= \left\{ x \in \mathbb{R}^n : x^T Q x + 2q^T x + \gamma \geq 0, \forall (Q, q, \gamma) \in (\mathcal{P} \cap \mathcal{Q}_-) \right\} = \emptyset.$$

Helly's Theorem implies, there exist $(n+1)$ inequalities, say given by, $P_1, P_2, \ldots, P_{n+1} \in (\mathcal{P} \cap \mathcal{Q}_-)$ such that

$$\left\{ x \in \mathbb{R}^n : \langle P_i, \left[\begin{array}{cc} 1 & x^T \\ x & xx^T \end{array} \right] \rangle \geq 0, \forall i;\ \ R - x^T x \geq 0 \right\} = \emptyset.$$

Now, let’s use SDP duality theory on the above. The underlying SDP is infeasible, its dual is

$$\begin{array}{rrcl} \text{sup} & \eta & & \\ & \begin{bmatrix} \eta & 0 \\ 0 & 0 \end{bmatrix} + \sum_{i=1}^{n+2} y_i P_i & \preceq & 0, \\ & y & \geq & 0. \end{array}$$

Since $P_{n+2} = \begin{bmatrix} R & 0 \\ 0 & -I \end{bmatrix}$, the dual has a Slater point.

Whence (by the Strong Duality Thm.), there exists $\bar{y} \geq 0$ such that

$$\sum_{i=1}^{n+2} \bar{y}_i(-P_i) \succeq \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Therefore (by Hilbert's Proposition),

$$\left[\sum_{i=1}^{n+2} \bar{y}_i(-p_i) - 1\right] \in \mathsf{SoS}(n, 2).$$

We have the certificate of infeasibility:

$$-1 = \sum_{i=1}^{n+2} \bar{y}_i p_i + \mathsf{SoS}(n, 2).$$

Therefore,

$$\left[\sum_{i=1}^{n+2} \bar{y}_i(-p_i) - 1\right] \in \text{SoS}(n, 2).$$

We have the certificate of infeasibility:

$$-1 = \sum_{i=1}^{n+2} \bar{y}_i p_i + \text{SoS}(n, 2).$$

But, some p_i may be derived recursively (may not be directly from the initial formulation).

Recall that D_1 denotes the unit hypersphere in $\mathbb{R}^n$ (we may take a finite mesh here). Let

$$D_2 := \left\{ v^{(i)}, -v^{(i)} : i \in \{1, 2, \ldots, n\} \right\},$$

$\left\{ v^{(i)} : i \in \{1, 2, \ldots, n\} \right\}$ is any basis for $\mathbb{R}^n$.

Compute

$$\alpha := \min \left\{ a^T x : x \in G_k \right\}$$

and

$$\beta := \min \left\{ b^T x : x \in G_k \right\}$$

for every $a \in D_1$ and for every $b \in D_2$. Then the quadratic inequality

$$\left(a^T x - \alpha \right) \left(b^T x - \beta \right) \geq 0$$

is a valid quadratic inequality for G_k and hence a valid quadratic inequality for F.

Note that the valid inequality $\left(a^T x - \alpha\right)\left(b^T x - \beta\right) \geq 0$ has the SoS certificate

$$\begin{aligned} &\left(a^T x - \alpha\right)\left(b^T x - \beta\right) \\ &= \left(\sum_i y_i^{(\alpha)} p_i + \mathsf{SoS}_\alpha(n,2)\right)\left(\sum_i y_i^{(\beta)} p_i + \mathsf{SoS}_\beta(n,2)\right), \end{aligned}$$

where $y^{(\alpha)}, y^{(\beta)} \geq 0$.

This leads to an algorithmic/constructive/optimization proof of Positivstellensatz type result with control over the degrees of the certificates and their *signatures*.

Recall,

Positivstellensatz (Stengle [1974]):

$F := \{x \in \mathbb{R}^n : p_i(x) \geq 0, i \in \{1, 2, \ldots, m\}\} = \emptyset$ iff there exist $s_0, s_J, \ldots \in$ SoS$(n, *)$ such that

$$g = \sum_{J \subseteq \{1,2,\ldots,m\}} s_J \prod_{i \in J} p_i = -1.$$

So far, we proved:

Theorem 5.3 *(Slumdog Positivstellensatz) Assume that in the initial system every polynomial is a quadratic. Then, $F = \emptyset$ iff there exist $s_0, s_J, \ldots \in$* SoS$(n, *)$ *and a finite k such that*

$$g = \sum_{J \subseteq \{1,2,\ldots,r(n,k)\}} s_J \prod_{i \in J} p_i^{z(i,J)} = -1,$$

where $r(n,k) \leq (n+1)n^{2k}$ and $z(i,J) \leq 2^{2^k}$.

But, the even powers of p_i can be absorbed by the s_J. So, the Slumdog SoS certificate is no worse than the original certificate. (That is, wlog, $z(i, J) \leq 1, \forall i, \forall J$.)

Proof: (A new, optimization theory proof of Stengle's Positivstellensatz)

1. Write the POP system as a system of quadratic inequalities (QIS), using *reduction equations and auxiliary variables*.
2. Apply the SCRM and the finite convergence Theorem of Kojima-T. [2000]. After k iterations of the SCRM, we have $C_k = \emptyset$.
3. Using Helly's Theorem, extract a subsystem of size $(n + 2)$.
4. Apply the SDP Duality Theorem, obtain an SoS certificate for $C_k = \emptyset$.
5. Using the recursive definition of SCRM$^{(*)}$ and Carathéodory's Theorem, convert the SoS certificate to an SoS certificate for (QIS).
6. Using the reduction equations convert the SoS certificate for QIS to an SoS certificate for the original POP.

■

(*) To obtain the SoS certificate in intermediate steps, we use the characterization of the faces of the cone of symmetric positive semidefinite matrices.

6 Final Remarks for this Part

Laurent [2003] showed that for combinatorial optimization problems, N_+ operator of Lovász and Schrijver is no stronger than the corresponding operators in Lasserre's (and hence Parrilo's) construction.

Therefore, for combinatorial optimization problems, the n-step convergence proof just follows from the d-step convergence proof of $N_{(j)}$ operators of Balas and Balas-Ceria-Cornuéjols (BCC).

Our techniques lead to convergence proofs for the general setting that do not rely on Putinar's Theorem or the Positivstellensatz. Hence the SSDPR serves a role analogous to the BCC operator.

E.g., the convergence of Parrilo's hierarchy also follows from the convergence theorem of Kojima, T. [2000].

7 Homogeneous Cones, Hyperbolic Cones, Facial Structure, Some Research Directions

A homogeneous polynomial $p : \mathbb{R}^d \mapsto \mathbb{R}$ is *hyperbolic in the direction* $h \in \mathbb{R}^d$, if $p(h) > 0$ and the univariate polynomial (in $t \in \mathbb{R}$)

$$p(x + th)$$

has only real roots for every $x \in \mathbb{R}^d$.

A convex cone K is a *hyperbolic cone* if it is

$$\left\{x \in \mathbb{R}^d : p(x + th) \neq 0, \ \forall t \in \mathbb{R}_+\right\}$$

for a polynomial p which is hyperbolic in the direction $h \in \mathbb{R}^d$.

A pointed closed convex cone K with nonempty interior in $\mathbb{R}^n$ is *homogeneous* if the group Aut(K) of nonsingular, linear maps on $\mathbb{R}^n$ keeping K invariant acts transitively on int(K). Homogeneous cones make up a proper subset of hyperbolic cones. (See Gindikin [1992] Tube Domains and the Cauchy Problem...). In fact, all homogeneous cones admit LMI representations. (See Chua [2004] who uses Vinberg's T-algebra characterization [1960s] of homogeneous cones.)

Definition 7.1 *$G \subset \mathbb{R}^d$ is said to admit a lifted-LMI representation if there exists $\mathcal{L} : \mathbb{R}^d \bigoplus \mathbb{R}^m \rightarrow \mathbb{R}^n$ a linear map such that*

$$x \in \text{int}(G) \iff \mathcal{L}(x, u) \succ 0 \text{ for some } u \in \mathbb{R}^m.$$

Note that a convex cone G admits a lifted-LMI representation iff its dual cone G^* does. In addition to this symmetry property, lifted-LMI representations cover a larger class of convex sets than the LMI representations and in terms of optimization algorithms which are used to solve the underlying optimization problems, lifted-LMI representations provide no additional difficulty (at least from a theoretical viewpoint).

G admits a *poly-time, lifted-LMI representation* if

$$\max\{m, n\} = O(\text{poly}(d)),$$

where poly$(\cdot)$ is a polynomial.

Open Problem: Characterize all d-dimensional, pointed, closed, convex cones in $\mathbb{R}^d$ which admit a lifted-LMI representation. I believe that this set of cones strictly contain the hyperbolic cones.

Open Problem: Characterize all d-dimensional, pointed, closed, convex cones in $\mathbb{R}^d$ which admit a poly-time, lifted-LMI representation.

Most specifically:

Open Problem: Are all Hyperbolic Feasibility Problems polynomial-time equivalent to LMI problems?

This last question needs some definitions and clarifications.

Definition 7.2 *Let* $p_1, p_2, \dots, p_m : \mathbb{R}^d \rightarrow \mathbb{R}$ *be given polynomials. Then the problem "does there exist* $x \in \mathbb{R}^d$ *such that* $p_i(x) \geq 0, \forall i \in \{1, 2, \dots, m\}$ *is a* Hyperbolic Feasibility Problem (HFP) *if every* p_i *is a hyperbolic polynomial.*

Next, we define the size(HFP). The "size" should involve the basic complexity measures needed to bound the amount of computational effort required (in the Blum-Shub-Smale real computation model [**?**]) to "solve" HFP to $\epsilon \in (0,1)$ accuracy using some general class of well-established algorithms. For instance, let η denote the maximum degree of polynomials $p_1, p_2, \ldots, p_m$, we can define

$$\text{size}(HFP) := \max\{m, d, \eta, \ln(1/\epsilon), \ln(R)\},$$

where $R > 1$ denotes the volume of a given ellipsoid E_0 which determines the region in which we will decide the solvability of HFP. I.e., our problem is to find $\bar{x} \in E_0$ satisfying all the inequalities. We require that after

$$\text{poly}(\text{size}(HLP)) \text{ operations}$$

the algorithm either outputs $\bar{x} \in \mathbb{R}^d$ such that

$p_i(\bar{x}) \geq 0, \forall i \in \{1, 2, \ldots, m\}$ or it outputs

“there does not exist a ball of volume at least ϵ which is contained in

$$E_0 \cap \left\{ x \in \mathbb{R}^d : p_i(x) \geq 0, \forall i \in \{1, 2, \ldots, m\} \right\} .''$$

In this context, when we say *HFP is polynomial-time equivalent to LMI* we mean that for every HFP (with m, d, η, R and a given $\epsilon \in (0, 1)$), we can explicitly describe an LMI such that

- the formulated LMI can be solved to ϵ accuracy in time poly $\left(\text{size}(HFP)\right)$,
- solving the LMI within accuracy ϵ, solves the original HFP.

Open Problem: In general, develop the theory of spectrahedral combinatorics/analysis (somewhat parallel to polyhedral combinatorics but not completely analogous).

A related, more specific task: Provide an elegant and useful characterization of the extreme rays of hyperbolic cones in terms of the defining polynomial.

Some preliminary findings by Myklebust [2008].

Homogeneous cones have been also classified via a recursive construction using bilinear forms $B(\cdot,\cdot)$ on simpler homogeneous cones. (These are sometimes called Siegel Domains or Siegel Cones.)

$$\begin{aligned} SC(K,B) := \text{cl}\Big\{ \quad & (x,u,t) \in \mathbb{R}^n \oplus \mathbb{R}^p \oplus \mathbb{R} : t > 0, \\ & \left[x - \tfrac{B(u,u)}{t}\right] \in K \Big\}. \end{aligned}$$

Construction starts from $K := \mathbb{R}_+$.

Extreme rays of homogeneous cones have been characterized:

Theorem 7.1 (Truong and T. 2004) *Let K be a homogeneous cone and let B be a homogeneous K-bilinear symmetric form. Then*

$$\begin{aligned}\text{ext}(SC(K,B)) &= \{(x,0,0) \in \mathbb{R}^n \oplus \mathbb{R}^p \oplus \mathbb{R} \;:\; x \in \text{ext}(K)\} \\ &\cup \left\{ \frac{(B(u,u),u,1)}{\|(B(u,u),u,1)\|} \;:\; u \in \mathbb{R}^p \right\}.\end{aligned}$$