Journal of Convex Analysis
Volume 15 (2008), No. 3, 635–654

# Computing Uniform Convex Approximations for Convex Envelopes and Convex Hulls

**R. Laraki**
*Laboratoire d'Econométrie and CNRS, Ecole Polytechnique,*
*1 rue Descartes, 75005 Paris, France*
*rida.laraki@polytechnique.edu*

**J. B. Lasserre***
*LAAS-CNRS and Institute of Mathematics, University of Toulouse,*
*7 Avenue du Colonel Roche, 31077 Toulouse, France*
*lasserre@laas.fr*

*Dedicated to the memory of Thomas Lachand-Robert.*

Received: March 7, 2006
Revised manuscript received: March 19, 2007

We provide a numerical procedure to compute uniform convex approximations $\{f_r\}$ of the convex envelope $\widehat{f}$ of a rational fraction $f$ defined on a compact basic semi-algebraic set $\mathbf{D}$. At each point $x$ of the convex hull $\mathbf{K} = \mathrm{co}(\mathbf{D})$, computing $f_r(x)$ reduces to solving a semidefinite program. We next characterize $\mathbf{K}$ in terms of the projection of a *semi-infinite* LMI, and provide outer convex approximations $\{\mathbf{K}_r\} \downarrow \mathbf{K}$. Testing whether $x \notin \mathbf{K}$ reduces to solving finitely many semidefinite programs.

## 1. Introduction

Computing the convex envelope $\widehat{f}$ of a given function $f : \mathbb{R}^n \to \mathbb{R}$ is a difficult problem. To the best of our knowledge, there is still no efficient algorithm that approximates $\widehat{f}$ by convex functions (except for the simpler univariate case). For instance, for a function $f$ on a bounded domain $\Omega$, Brighi and Chipot [4] propose triangulation methods and provide piecewise degree-1 polynomial approximations $f_h \geq \widehat{f}$, and derive estimates of $f_h - \widehat{f}$ (where $h$ measures the size of the mesh). Another possibility is to view the problem as a particular instance of the general moment problem, and use geometrical approaches as described in e.g. Anastassiou [1] or Kemperman [8]; but, as acknowledged in [1, 8], this approach is only practical for say, the univariate or bivariate cases.

Concerning convex sets, an important issue raised in Ben-Tal and Nemirovski [3], Parrilo and Sturmfels [14], is to characterize the convex sets that have a *LMI (Linear Matrix Inequalities)* or *semidefinite representation*, and called *SDr* sets in Ben-Tal and Nemirovski [3]. For instance, the epigraph of a univariate convex polynomial is a SDr set. So far, and despite some progress in particular cases (see e.g. the recent proof of the Lax conjecture by Lewis et al [13]), little is known. However, Helton and Vinnikov [6] have proved recently that *rigid convexity* is a necessary condition for a set to be SDr.

*The second author acknowledges financial support from the french ANR agency under grant NT05-3-41612.

ISSN 0944-6532 / $ 2.50 

In this paper we consider both convex envelopes and convex hulls for certain classes of functions and sets, namely *rational fractions* and compact basic *semi-algebraic* sets. In both cases, one provides relatively simple numerical uniform approximations via semidefinite programming.

**Contribution.** Our contribution is twofold: Concerning convex envelopes, we consider the class of *rational fractions* $f$ on a compact basic semi-algebraic set $\mathbf{D} \subset \mathbb{R}^n$ (and $+\infty$ outside $\mathbf{D}$). We view the problem as a particular instance of the general moment problem, and we provide an algorithm for computing *convex* and *uniform* approximations of its convex envelope $\widehat{f}$. More precisely, with $\mathbf{K} := \mathrm{co}(\mathbf{D})$ being the convex hull of $\mathbf{D}$:

(a) We provide a sequence of *convex* functions $\{f_r\}$ that converges to $\widehat{f}$ **uniformly** on any compact subset of $\mathbf{K}$ where $\widehat{f}$ is continuous, as $r$ increases[1].
(b) At each point $x \in \mathbb{R}^n$, computing $f_r(x)$ reduces to solving a semidefinite program $\mathbf{Q}_{rx}$.
(c) For every $x \in \mathrm{int}\,\mathbf{K}$, the SDP dual $\mathbf{Q}^*_{rx}$ is solvable and any optimal solution provides an element of the subgradient $\partial f_r(x)$ at the point $x \in \mathrm{int}\,\mathbf{K}$.
(d) We give a geometric condition on $\mathbf{D}$ that ensures that $\widehat{f}$ is continuous on $\mathbf{K}$. This extends the one introduced in Laraki [11] for the case where $\mathbf{D} = \mathbf{K}$.

Concerning sets, we consider the class of compact basic semi-algebraic subsets $\mathbf{D} \subset \mathbb{R}^n$, and:

(e) We characterize its *convex hull* $\mathbf{K} := \mathrm{co}(\mathbf{D})$ as the projection of a *semi-infinite* SDr set $S_\infty$, i.e., a set defined by finitely many LMIs involving matrices of infinite dimension, and countably many variables. Importantly, the LMI representation of $S_\infty$ is simple and given *directly* in terms of the data defining the original set $\mathbf{D}$.
(f) We provide outer convex approximations of $\mathbf{K}$, namely a monotone nonincreasing sequence of convex sets $\{\mathbf{K}_r\}$, with $\mathbf{K}_r \downarrow \mathbf{K}$. Each set $\mathbf{K}_r$ is the projection of a SDr set $S_r$, obtained from $S_\infty$ by "finite truncation". Then, checking whether $x \notin \mathbf{K}$ reduces to solving finitely many SDPs based on SDr sets $S_r$, until one is unfeasible, which eventually happens for some $r$. Importantly again, the LMI representation of $S_r$ is simple and given *directly* in terms of the data defining the original set $\mathbf{D}$. Other outer approximations are of course possible, like e.g. convex polytopes $\{\Omega_r\}$ containing $\mathbf{K}$, but obtaining such polytopes with $\Omega_r \downarrow \mathbf{K}$ is far from trivial.

The proof combines known technics (some of them developed by the authors). In particular, it uses Choquet's representation of probability measures on the convex compact set $\mathbf{K} := \mathrm{co}(\mathbf{D})$, together with its reformulation as an infinite-dimensional linear program $\mathbf{P}_x$, a particular instance of the generalized problem of moments. In case where $f = \frac{p}{q}$ with $p$ and $q$ two polynomials and $q$ positive on $\mathbf{D}$, one uses standard arguments to show that the dual is solvable and there is no duality gap between $\mathbf{P}_x$ and its dual $\mathbf{P}^*_x$. Assuming further that $\mathbf{D}$ is basic semi-algebraic, one follows the methodology developed in Lasserre [12] to derive semidefinite programming (SDP) relaxations of $\mathbf{P}_x$ whose optimal values form a monotone sequence converging to the optimal value of $\mathbf{P}_x$. When $\widehat{f}$ is continuous on $\mathbf{K}$, by Dini's theorem the convergence is uniform with respect to $x \in \mathbf{K}$. Conditions for continuity may be obtained using technics from Laraki [11].

[1]Determining $r$ such that $f_r$ approximates $\widehat{f}$ up to some prescribed error $\varepsilon$ is an open problem.

## 2. Notation definitions and preliminary results

In the sequel, $\mathbb{R}[y](:= \mathbb{R}[y_1, \dots, y_n])$ denotes the ring of real-valued polynomials in the variable $y = (y_1, \dots, y_n)$. Let $\overline{y_i} \in \mathbb{R}[y]$ be the natural projection on the $i$-variable that is for every $x \in \mathbb{R}^n$, $\overline{y_i}(x) = x_i$. For a real-valued symmetric matrix $M$, the notation $M \succeq 0$ stands for $M$ is positive semidefinite.

Let $\mathbf{D} \subset \mathbb{R}^n$ be compact, and denote by:

- $\mathbf{K}$, the convex hull of $\mathbf{D}$. Hence, by a theorem of Caratheodory, $\mathbf{K}$ is convex and compact; see Rockafellar [16].
- $C(\mathbf{D})$, the Banach space of real-valued continuous functions on $\mathbf{D}$, equipped with the sup-norm $\|f\| := \sup_{x\in\mathbf{D}} |f(x)|$, $f \in C(\mathbf{D})$.
- $M(\mathbf{D})\,(\simeq C(\mathbf{D})^*)$, its topological dual, i.e., the Banach space of finite *signed* Borel measures on $\mathbf{K}$, equipped with the norm of total variation.
- $M_+(\mathbf{D}) \subset M(\mathbf{D})$, the positive cone, i.e., the set of finite Borel measures on $\mathbf{D}$.
- $\Delta(\mathbf{D}) \subset M_+(\mathbf{D})$, the set of Borel probability measures on $\mathbf{D}$.
- $\widetilde{f}$, the natural extension to $\mathbb{R}^n$ of $f \in C(\mathbf{D})$, that is

$$x \mapsto \widetilde{f}(x) := \begin{cases} f(x) & \text{on } \mathbf{D} \\ +\infty & \text{on } \mathbb{R}^n \setminus \mathbf{D}. \end{cases} \tag{1}$$

  Note that $\widetilde{f}$ is lower-semicontinuous (l.s.c.), admits a minimum and its effective domain $\mathbf{D}$ is non-empty and compact (in the sequel we denote it by $\mathrm{dom}\,\widetilde{f}$).
- $\widehat{f}$ (with $f$ in $C(\mathbf{D})$), the convex envelope of $\widetilde{f}$, that is, the greatest convex function majorized by $\widetilde{f}$.

$M(\mathbf{D})$ and $C(\mathbf{D})$ form a *dual pair* of vector spaces, with duality bracket

$$\langle \sigma, f \rangle := \int_{\mathbf{D}} f\, d\sigma, \quad \sigma \in M(\mathbf{D}), f \in C(\mathbf{D}).$$

Hence, let $\tau^*$ denote the associated weak $\star$ topology; this is the coarsest topology on $M(\mathbf{D})$ for which $\sigma \to \langle \sigma, f \rangle$ is continuous for every function $f$ in $C(\mathbf{D})$.

### 2.1. Preliminaries

In this section some well known results are stated and proved (in our semi-algebraic context). With $f \in C(\mathbf{D})$, and $x \in \mathbf{K} = \mathrm{co}(\mathbf{D})$ fixed, arbitrary, consider the infinite-dimensional linear program (LP):

$$\mathrm{LP}_x : \begin{cases} & \inf_{\sigma \in M_+(\mathbf{D})} \langle \sigma, f \rangle \\ \text{s.t.} & \langle \sigma, \overline{y_i} \rangle = x_i, \quad i = 1, \dots, n \\ & \langle \sigma, 1 \rangle = 1. \end{cases} \tag{2}$$

Its optimal value is denoted by $\inf \mathrm{LP}_x$, and $\min \mathrm{LP}_x$ if the infimum is attained. Notice that $\mathrm{LP}_x$ is a particular instance of the general *moment problem*, as described in e.g. Kemperman [8, §2.6]. In particular, the set of $x \in \mathbb{R}^n$ such that $\mathrm{LP}_x$ has a feasible solution, is called the *moment space*.

**Lemma 2.1.** *Let* $\mathbf{K} = \mathrm{co}(\mathbf{D})$, $f \in C(\mathbf{D})$ *and* $\widetilde{f}$ *be as in* (1). *Then the convex envelope* $\widehat{f}$ *of* $\widetilde{f}$ *is given by:*

$$\widehat{f}(x) = \begin{cases} \min \mathrm{LP}_x, & x \in \mathbf{K}, \\ +\infty, & x \in \mathbb{R}^n \setminus \mathbf{K}, \end{cases} \tag{3}$$

*and so,* $\mathbf{K} = \mathrm{dom}\,\widehat{f}$.

**Proof.** For $x \in \mathbf{K}$, let $\Delta_x(\mathbf{D})$ be the set of probability measures $\sigma$ on $\mathbf{D}$, that are centered at $x$ (that is $\langle \sigma, \overline{y_i} \rangle = x_i$ for $i = 1, \ldots, n$). Let $\Delta_x^*(\mathbf{D}) \subset \Delta_x(\mathbf{D})$ be the subset of those probability measures that have a finite support. It is well known that

$$\widehat{f}(x) = \inf_{\sigma \in \Delta_x^*(\mathbf{D})} \langle \sigma, f \rangle, \quad \forall x \in \mathbf{K};$$

see Choquet [5] or Laraki [11]. Next, since $\Delta_x^*(\mathbf{D})$ is dense in $\Delta_x(\mathbf{D})$ with respect to the weak $\star$ topology, and $\Delta(\mathbf{D})$ is metrizable and compact with respect to the same topology (see Choquet [5]), deduce that for every $x \in \mathbf{K}$

$$\begin{aligned} \widehat{f}(x) &= \min_{\sigma \in \Delta_x(\mathbf{D})} \langle \sigma, f \rangle \\ &= \min \mathrm{LP}_x. \end{aligned}$$

If $x \notin \mathbf{K}$, there is no probability measure on $\mathbf{D}$, with finite support, and centered in $x$; therefore $\widehat{f}(x) = +\infty$. $\square$

Next, let $p, q \in \mathbb{R}[y]$, with $q > 0$ on $\mathbf{D}$, and let $f \in C(\mathbf{D})$ be defined as

$$y \mapsto f(y) = p(y)/q(y), \quad y \in \mathbf{D}. \tag{4}$$

For every $x \in \mathbf{K}$, consider the LP,

$$\mathbf{P}_x : \quad \begin{cases} & \inf_{\sigma \in M_+(\mathbf{D})} \langle \sigma, p \rangle \\ \text{s.t.} & \langle \sigma, \overline{y_i} q \rangle = x_i, \quad i = 1, \ldots, n \\ & \langle \sigma, q \rangle = 1. \end{cases} \tag{5}$$

A dual of $\mathbf{P}_x$, is the LP

$$\mathbf{P}_x^* : \quad \sup_{\gamma \in \mathbb{R}, \lambda \in \mathbb{R}^n} \{\gamma + \langle \lambda, x \rangle : \ p(y) - q(y)\langle \lambda, y \rangle \geq \gamma q(y), \quad \forall y \in \mathbf{D}\}, \tag{6}$$

where $\langle \lambda, y \rangle := \sum_{i=1}^n \lambda_i y_i$ stands for the standard inner product in $\mathbb{R}^n$. The optimal value of $\mathbf{P}_x^*$ is denoted by $\sup \mathbf{P}_x^*$ (and $\max \mathbf{P}_x^*$ if the supremum is attained). Equivalently, as $q > 0$ everywhere on $\mathbf{D}$, and $f = p/q$ on $\mathbf{D}$,

$$\mathbf{P}_x^* : \quad \sup_{\gamma \in \mathbb{R}, \lambda \in \mathbb{R}^m} \{\gamma + \langle \lambda, x \rangle : \ f(y) - \langle \lambda, y \rangle \geq \gamma, \quad \forall y \in \mathbf{D}\}. \tag{7}$$

In view of the definition (4) of $f$, notice that

$$f(y) - \langle \lambda, y \rangle \geq \gamma, \quad \forall y \in \mathbf{D} \quad \Leftrightarrow \quad \widetilde{f}(y) - \langle \lambda, y \rangle \geq \gamma, \quad \forall y \in \mathbb{R}^n.$$

Hence $\mathbf{P}_x^*$ in (7) is just the dual $\mathrm{LP}_x^*$ of $\mathrm{LP}_x$, for every $x \in \mathbf{K}$, and so $\sup \mathbf{P}_x^* = \sup \mathrm{LP}_x^*$, for every $x \in \mathbf{K}$. In fact we have the following well known result (that holds of course in a more general framework).

**Theorem 2.2.** *Let $p, q \in \mathbb{R}[x]$ with $q > 0$ on $\mathbf{D}$, and let $f$ be as in (4). Let $x \in \mathbf{K} = \mathrm{co}(\mathbf{D})$ be fixed, arbitrary, and let $\mathbf{P}_x$ and $\mathbf{P}_x^*$ be as in (5) and (7), respectively. Then $\mathbf{P}_x$ and $\mathrm{LP}_x$ are solvable and there is no duality gap, i.e.,*

$$\sup \mathbf{P}_x^* \;=\; \sup \mathrm{LP}_x^* \;=\; \min \mathrm{LP}_x \;=\; \min \mathbf{P}_x \;=\; \widehat{f}(x), \quad x \in \mathbf{K}. \tag{8}$$

**Proof.** This is a consequence of Legendre-Fenchel duality. Observe that $\mathbf{P}_x$ is equivalent to $\mathrm{LP}_x$. Indeed, with $\sigma$ an arbitrary feasible solution of $\mathbf{P}_x$, the measure $d\mu := q d\sigma$ is feasible in $\mathrm{LP}_x$, with same value. Similarly, with $\mu$ an arbitrary feasible solution of $\mathrm{LP}_x$, the measure $d\sigma := q^{-1} d\mu$, well defined on $\mathbf{D}$ because $q > 0$ on $\mathbf{D}$, is feasible in $\mathbf{P}_x$, and with same value. Finally, it is well known that $\widehat{f}$ is the Legendre-Fenchel biconjugate[2] of $\widetilde{f}$, and so $\widehat{f}(x) = \sup \mathbf{P}_x^*$, for all $x \in \mathbf{K}$. Indeed, let $f^* : \mathbb{R}^n \to \mathbb{R}$ be the Legendre-Fenchel conjugate of $\widetilde{f}$, i.e.,

$$\lambda \mapsto f^*(\lambda) := \sup_{y \in \mathbb{R}^n} : \{\langle \lambda, y \rangle - \widetilde{f}(y)\}.$$

In view of the definition of $\widetilde{f}$,

$$f^*(\lambda) \;=\; \sup_{y \in \mathbf{D}} : \{\langle \lambda, y \rangle - f(y)\},$$

and therefore,

$$\begin{aligned} \sup \mathbf{P}_x^* &= \sup_{\lambda} \{\langle \lambda, x \rangle + \inf_{y \in \mathbf{D}} \{f(y) - \langle \lambda, y \rangle\}\} \\ &= \sup_{\lambda} \{\langle \lambda, x \rangle - \sup_{y \in \mathbf{D}} \{\langle \lambda, y \rangle - f(y)\}\} \\ &= \sup_{\lambda} \{\langle \lambda, x \rangle - f^*(\lambda)\} \;=\; (f^*)^*(x) \;=\; \widehat{f}(x). \end{aligned}$$

□

We deduce the following.

**Corollary 2.3.** *Let $x \in \mathbf{K}$ be fixed, arbitrary, and let $\mathbf{P}_x^*$ be as in (7).*

(*a*) $\mathbf{P}_x^*$ *is solvable if and only if $\partial \widehat{f}(x) \neq \emptyset$*[3]*, in which case any optimal solution $(\lambda^*, \gamma^*)$ satisfies:*

$$\lambda^* \in \partial \widehat{f}(x), \qquad \textit{and} \qquad \gamma^* \;=\; -f^*(\lambda^*). \tag{9}$$

(*b*) *If $f$ is a rational fraction on $\mathbf{D}$ as in (4), then $\partial \widehat{f}(x) \neq \emptyset$ for every $x$ in $\mathbf{K}$ so that, in this case $\mathbf{P}_x^*$ is solvable and (a) holds for every $x$ in $\mathbf{K}$.*

**Proof.** The first part is standard. Suppose that for some $x \in \mathbf{K}$, $\mathbf{P}_x^*$ is solvable (that is, the supremum is achieved, say at $\lambda^*(x)$ and $\gamma^*(x)$). Then, for every $y \in \mathbf{K}$,

$$\begin{aligned} \widehat{f}(x) &= \langle \lambda^*(x), x \rangle + \gamma^*(x) \\ \widehat{f}(y) &= \sup_{\lambda,\gamma} \{\langle \lambda, y \rangle + \gamma : f(z) - \langle \lambda, z \rangle \geq \gamma, \;\; \forall z \in \mathbf{D}\}, \quad y \in \mathbf{K} \\ &\geq \langle \lambda^*(x), y \rangle + \gamma^*(x), \end{aligned}$$

[2]See Section 2 in Benoist and Hiriart-Urruty [2]

[3]$\partial \widehat{f}(x) \neq \emptyset$ at least for every $x$ in the relative interior of $\mathbf{K}$ (see Rockafellar [16], Theorem 23.4)

and in view of (3), the latter inequality also holds for every $y$ in $\mathbb{R}^n$; therefore,

$$\widehat{f}(y) - \widehat{f}(x) \geq \langle \lambda^*(x), y - x \rangle, \quad \forall y \in \mathbb{R}^n.$$

Hence, $\lambda^*(x) \in \partial \widehat{f}(x)$. Finally, from the standard Legendre-Fenchel equality, we deduce that $\gamma^*(x) = -f^*(\lambda^*(x))$ where $f^*$ is the Legendre-Fenchel conjugate of $\widetilde{f}$. Conversely, if $\lambda^*(x) \in \partial \widehat{f}(x)$ then, by the Legendre-Fenchel equality we have,

$$\widehat{f}(x) = \langle \lambda^*(x), x \rangle - f^*(\lambda^*(x)).$$

Therefore, we have:

$$\sup \mathbf{P}_x^* \; = \; \widehat{f}(x) \; = \; \langle \lambda^*(x), x \rangle - f^*(\lambda^*(x)). \tag{10}$$

Next, from

$$f^*(\lambda^*(x)) \; = \; \sup_{y \in \mathbf{D}} \langle \lambda^*(x), y \rangle - f(y),$$

we have

$$-f^*(\lambda^*(x)) \; \leq \; f(y) - \langle \lambda^*(x), y \rangle, \quad \forall y \in \mathbf{D},$$

which shows that the pair $(\lambda^*(x), -f^*(\lambda^*(x)))$ is a feasible solution of $\mathbf{P}_x^*$, and in view of (10), an optimal solution.

Now, if $f$ is a rational fraction on $\mathbf{D}$, then it is differentiable and Lipschitz on $\mathbf{D}$ so that, from Theorem 3.6 in Benoist and Hiriart-Urruty [2], $\partial \widehat{f}(x)$ is uniformly bounded as $x$ varies on the relative interior of $\mathbf{K}$. Since $\widehat{f}$ is l.s.c. (see below), we deduce that $\partial \widehat{f}(x) \neq \emptyset$ for every $x$ in $\mathbf{K}$. Actually, let $x \in \mathbf{K}$ and let $x_n$ be a sequence in the relative interior of $\mathbf{K}$ that converges to $x$ and let $\lambda_n \in \partial \widehat{f}(x_n)$ such that $\lambda_n \to \lambda$ (which is possible (passing to a subsequence if needed) since $\partial \widehat{f}(x_n)$ is uniformly bounded). Hence, for every $y$ in $\mathbf{R}^n$,

$$\widehat{f}(y) \geq \widehat{f}(x_n) + \langle \lambda_n, y - x_n \rangle \,,$$

so that,

$$\begin{aligned} \widehat{f}(y) \;&\geq\; \liminf_{n \to \infty} \widehat{f}(x_n) + \langle \lambda, y - x \rangle \\ &\geq\; \widehat{f}(x) + \langle \lambda, y - x \rangle \end{aligned}$$

consequently, $\lambda \in \partial \widehat{f}(x)$, the desired result. □

In other words, any optimal solution of $\mathbf{P}_x^*$ provides an element of the subgradient of $\widehat{f}$ at the point $x$. Corollary 2.3 should be viewed as a refinement for convex envelopes of rational fractions, of Theorem 2.20 in Kemperman [8, p. 28] for the general moment problem, where strong duality results are obtained for the interior of the moment space (here $\operatorname{int} \mathbf{K}$) only.

### 2.2. On the preservation of continuity

As we will construct a sequence $\{f_r\}$ that approximates $\widehat{f}$ uniformly on compact sets where $\widehat{f}$ is continuous, it is natural to investigate conditions on the data which ensure that $\widehat{f}$ is continuous everywhere on its domain.

In 1969, Kruskal [9] provided an example showing that this is not always true. Kruskal constructs a compact, convex basic semi-algebraic set $\mathbf{K}$ for which the set of extreme points is *not* closed. Then he exhibits a polynomial $f$ (of degree 2) on $\mathbf{K}$ whose convex envelope $\widehat{f}$ is discontinuous on $\mathbf{K}$. Therefore, that $\mathbf{K}$ is a basic semi-algebraic set does *not* guarantee that $\widehat{f}$ is continuous on $\mathbf{K}$ whenever $f$ is. This example also shows that restricting attention to polynomials does not help in getting continuity of $\widehat{f}$. This issue was recently addressed in Laraki [11] with a complete answer to the Kruskal's main observation in the case where $\mathbf{D}$ is convex. It is shown for example that when $\mathbf{D}$ is a polytope or is an euclidean ball (two basic semi-algebraic sets) then continuity is preserved. By adapting a condition in Laraki [11] we can obtain a necessary and sufficient condition for the preservation of continuity when $\mathbf{D}$ is not convex.

**Definition 2.4.** The compact set $\mathbf{D}$ of $\mathbb{R}^n$ is Splitting-Continuous if and only if $x \mapsto \Delta_x(\mathbf{D})$ is continuous when $\Delta(\mathbf{D})$ is equipped with the weak $\star$ topology.

**Lemma 2.5.** *Let $f$ be a continuous function on a compact $\mathbf{D}$ of $\mathbb{R}^n$. Then, $\widehat{f}$ is l.s.c. on $\mathbf{K}$ and is continuous on any compact $\overline{\mathbf{K}}$ that is strictly included in the relative interior of $\mathbf{K}$. Moreover, $\mathbf{D}$ is Splitting-Continuous if and only if $\widehat{f}$ is continuous on $\mathbf{K}$, for every $f$ which is the restriction on $\mathbf{D}$ of some continuous function on $\mathbf{K}$.*

**Proof.** As $\widehat{f}$ is convex, then by Theorem 10.1 in Rockafellar [16], it is continuous on the relative interior of $\mathbf{K}$. In addition, $\widehat{f}$ is l.s.c. on $\mathbf{K}$ because $f$ is continuous and the correspondence $x \mapsto \Delta_x(\mathbf{D})$ is upper-semicontinuous (u.s.c.); see e.g. Laraki and Sudderth [10, Theor. 6].

Again, from [10, Theor. 6], $x \mapsto \Delta_x(\mathbf{D})$ is continuous if and only if $\widehat{f}$ is continuous on $\mathbf{K}$, whenever $f$ is the restriction on $\mathbf{D}$ of some continuous function on $\mathbf{K}$. □

## 3. Uniform convex approximations of $\widehat{f}$ by SDP-relaxations

In this section, we assume that $f$ is defined as in (4) for some polynomials $p, q \in \mathbb{R}[x]$, with $q > 0$ on $\mathbf{D}$, where $\mathbf{D} \subset \mathbb{R}^n$ is a compact basic semi-algebraic set defined by

$$\mathbf{D} := \{x \in \mathbb{R}^n : \; g_j(x) \geq 0, \quad j = 1, \ldots, m\}, \tag{11}$$

for some polynomials $\{g_j\} \subset \mathbb{R}[x]$. Depending on its parity, let $2r_j - 1$ or $2r_j$ be the total degree of $g_j$, for all $j = 1, \ldots, m$. Similarly, let $2r_p, 2r_q$ or $2r_p - 1, 2r_q - 1$ be the total degree of $p$ and $q$ respectively.

We next provide a sequence $\{f_r\}_r$ of functions such that for every $r$:

- $f_r$ is convex with $\operatorname{dom} f_r = \mathbf{K}_r \supset \mathbf{K}$;
- $f_r \leq \widehat{f}$ and for every $x \in \mathbf{K}$, $f_r(x) \uparrow \widehat{f}(x)$ as $r \to \infty$. In fact, we even have

$$\lim_{r \to \infty} \|\widehat{f} - f_r\|_{\overline{\mathbf{K}}} \to 0,$$

that is, $f_r$ converges to $\widehat{f}$ *uniformly* on any compact $\overline{\mathbf{K}} \subset \mathbf{K}$ where $\widehat{f}$ is continuous. Consequently, if $\mathbf{D}$ is Splitting-Continuous and if $q > 0$ on $\mathbf{K}$, then one obtains uniform convergence on $\mathbf{K}$. Also, if $\overline{\mathbf{K}}$ is strictly included in the relative interior of $\mathbf{K}$ then we also obtain uniform convergence on $\overline{\mathbf{K}}$.

To do this we first introduce some additional notation.

### 3.1. Notation and definitions

Let $y = \{y_\alpha\}_{\alpha\in\mathbb{N}^n}$ be a sequence indexed in the canonical basis $\{z^\alpha\}$ of $\mathbb{R}[z]$, and let $L_y : \mathbb{R}[z] \to \mathbb{R}$ be the linear functional defined by

$$h\,(:= \sum_{\alpha\in\mathbb{N}^n} h_\alpha z^\alpha) \;\mapsto L_y(h) \;:=\; \sum_{\alpha\in\mathbb{N}^n} h_\alpha y_\alpha.$$

Let $\mathcal{P}_k \subset \mathbb{R}[z]$ be the space of polynomials of total degree less than $k$, and let $r_0 := \max[r_p,\, r_q+1,\, r_1,\ldots,r_m]$. Then for $r \geq r_0$, consider the optimization problem:

$$\mathbf{Q}_{rx}: \quad \begin{cases} & \inf_y L_y(p) \\ \text{s.t.} & L_y(z_i q) \;=\; x_i, \quad i=1,\ldots,n, \\ & L_y(h^2) \geq 0, \quad \forall\, h \in \mathcal{P}_r, \\ & L_y(h^2 g_j) \geq 0, \quad \forall\, h \in \mathcal{P}_{r-r_j},:\; j=1,\ldots,m, \\ & L_y(q) = 1. \end{cases} \tag{12}$$

Problem $\mathbf{Q}_{rx}$ is a convex optimization problem, in fact, a so-called *semidefinite programming* problem, called a SDP-relaxation of $\mathbf{P}_x$. For more details on semidefinite programming and its applications, the reader is referred to Vandenberghe and Boyd [18].

Indeed, given $y = \{y_\alpha\}$, let $M_r(y)$ be the *moment* matrix associated with $y$, that is, the rows and columns of $M_r(y)$ are indexed in the canonical basis of $\mathcal{P}_r$, and the entry $(\alpha,\beta)$ is defined by $M_r(y)(\alpha,\beta) = L_y(z^{\alpha+\beta}) = y_{\alpha+\beta}$, for all $\alpha,\beta \in \mathbb{N}^n$, with $|\alpha|,|\beta| \leq r$. Then

$$L_y(h^2) \geq 0, \quad \forall h \in \mathcal{P}_r \quad \Leftrightarrow \quad M_r(y) \succeq 0.$$

Similarly, writing

$$z \;\mapsto\; g_j(z) \;:=\; \sum_{\gamma\in\mathbb{N}^n} (g_j)_\gamma z^\gamma, \quad j=1,\ldots,m,$$

the *localizing* matrix $M_r(g_j y)$ associated with $y$ and $g_j \in \mathbb{R}[z]$, is the matrix also indexed in the canonical basis of $\mathcal{P}_r$, and whose entry $(\alpha,\beta)$ is defined by

$$M_r(g_j y)(\alpha,\beta) \;=\; L_y(g_j(z) z^{\alpha+\beta}) \;=\; \sum_{\gamma\in\mathbb{N}^n} y_{\alpha+\beta+\gamma}(g_j)_\gamma,$$

for all $\alpha,\beta \in \mathbb{N}$, with $|\alpha|,|\beta| \leq r$. Then, for every $j = 1,\ldots,m$,

$$L_y(g_j h^2) \geq 0, \quad \forall h \in \mathcal{P}_r \quad \Leftrightarrow \quad M_r(g_j y) \succeq 0.$$

Observe that if $y$ has a *representing measure* $\mu_y$, i.e. if

$$y_\alpha \;=\; \int x^\alpha \, d\mu_y, \quad \forall \alpha \in \mathbb{N}^n,$$

then, with $h \in \mathcal{P}_r$, and denoting by $\mathbf{h} = \{h_\alpha\} \in \mathbb{R}^{s(r)}$ its vector of coefficients in the canonical basis,

$$\langle \mathbf{h}, M_r(g_j y)\mathbf{h}\rangle \;=\; \int h^2 g_j \, d\mu_y.$$

Therefore, if $\mu_y$ has its support contained in the level set $\{x \in \mathbb{R}^n : \; g_j(x) \geq 0\}$, one has $M_r(g_j y) \succeq 0$.

One also denotes by $M_\infty(y)$ and $M_\infty(g_j y)$ the (obvious) respective "infinite" versions of $M_r(y)$ and $M_r(g_j y)$, i.e., moment and localizing matrices with countably many rows and columns indexed in the canonical basis $\{z_\alpha\}$, and involving *all* the moment variables $y$ (as opposed to finitely many in $M_r(y)$ and $M_r(g_j y)$).

For more details on moment and localizing matrices, the reader is referred to e.g. Lasserre [12].

### 3.2. SDP-relaxations

Hence, using the above notation, the optimization problem $\mathbf{Q}_{rx}$ defined in (12) is just the SDP

$$\mathbf{Q}_{rx} : \quad \begin{cases} & \inf_y L_y(p) \\ \text{s.t.} & L_y(z_i q) = \; x_i, \quad i = 1, \ldots, n \\ & M_r(y) \succeq 0, \\ & M_{r-r_j}(g_j y) \succeq 0, \quad j = 1, \ldots, m, \\ & L_y(q) = 1, \end{cases} \tag{13}$$

with optimal value denoted $\inf \mathbf{Q}_{rx}$, and $\min \mathbf{Q}_{rx}$ if the infimum is attained.

Writing $M_r(y) = \sum_\alpha B_\alpha y_\alpha$, and $M_{r-r_j}(g_j y) = \sum_\alpha C^j_\alpha y_\alpha$, for appropriate symmetric matrices $\{B_\alpha, C^j_\alpha\}$, the dual of $\mathbf{Q}_{rx}$ is the SDP

$$\mathbf{Q}^*_{rx} : \quad \begin{cases} & \sup\limits_{\lambda,\gamma,X,Z_j} \; \gamma + \langle \lambda, x\rangle \\ \text{s.t.} & \langle B_\alpha, X\rangle + \sum\limits_{j=1}^m \langle C^j_\alpha, Z_j\rangle + \gamma q_\alpha + \sum\limits_{i=1}^n \lambda_i (z_i q)_\alpha = p_\alpha, \quad |\alpha| \leq 2r \\ & X, Z_j \succeq 0. \end{cases} \tag{14}$$

In fact, letting $g_0 \equiv 1$, and using the spectral decompositions $X = \sum_l u_{0l} u_{0l}^t$ and $Z_j = \sum_l u_{jl} u_{jl}^t$ for some vectors $\{u_{jl}\}$, $j = 0, \ldots, m$, one may write $\mathbf{Q}^*_{rx}$ as:

$$\mathbf{Q}^*_{rx} : \quad \begin{cases} & \sup\limits_{\gamma,\lambda,\{u_j\}} \; \gamma + \langle \lambda, x\rangle \\ \text{s.t.} & p - \gamma\, q - \langle \lambda, z\rangle\, q = \sum\limits_{j=0}^m u_j \, g_j \\ & u_j \text{ s.o.s.}, \deg u_j \, g_j \leq 2r, \quad j = 0, \ldots, m \end{cases} \tag{15}$$

(where s.o.s. stands for *sum of squares*); see for instance the derivation in Lasserre [12]. We next make the following assumption on the polynomials $\{g_j\} \subset \mathbb{R}[z]$ that define the set $\mathbf{D}$ in (11).

**Assumption 3.1.** There is a polynomial $u \in \mathbb{R}[z]$ which can written

$$u \;=\; u_0 + \sum_{j=1}^{m} u_j\, g_j, \tag{16}$$

for some family of s.o.s. polynomials $\{u_j\}_{j=0}^m \subset \mathbb{R}[z]$, and whose level set $\{z \in \mathbb{R}^n : u(z) \geq 0\}$ is compact.

Assumption 3.1 is not very restrictive. For instance, it is satisfied if:

- all the $g_j$'s are linear (and so, $\mathbf{D}$ is a convex polytope; see Putinar [15]), or if
- the level set $\{z \in \mathbb{R}^n \,:\, g_j(z) \geq 0\}$ is compact, for some $j \in \{1, \ldots, m\}$.

Moreover, if one knows some $M \in \mathbb{R}$ such that the compact set $\mathbf{D}$ is contained in the ball $\{z \in \mathbb{R}^n \,:\, \|z\| \leq M\}$, then it suffices to add the redundant quadratic constraint $M^2 - \|z\|^2 \geq 0$ in the definition (11) of $\mathbf{D}$, and Assumption 3.1 holds true. In some problems, computing such a constant may be costly.

Under Assumption 3.1, every polynomial $v \in \mathbb{R}[x]$, strictly positive on $\mathbf{D}$, can be written as

$$v \;=\; v_0 + \sum_{j=1}^{m} v_j\, g_j,$$

for some family of s.o.s. polynomials $\{v_j\}_{j=0}^m \subset \mathbb{R}[x]$. This is Putinar's Positivstellensatz, a refinement of Schmüdgen's Positivstellensatz (see Putinar [15] and Jacobi and Prestel [7]). Then we have the following result:

**Theorem 3.2.** *Let $\mathbf{D}$ be as in* (11)*, and let Assumption* 3.1 *hold. Let $f$ be as in* (4) *with $p, q \in \mathbb{R}[z]$, and with $q > 0$ on $\mathbf{D}$. Let $\widehat{f}$ be as in* (3)*, and with $x \in \mathbf{K} = \mathrm{co}(\mathbf{D})$ fixed, consider the SDP-relaxations $\{\mathbf{Q}_{rx}\}$ defined in* (12)

($a$) *The function $f_r : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ defined by*

$$x \;\mapsto\; f_r(x) \;:=\; \inf \mathbf{Q}_{rx}, \quad x \in \mathbb{R}^n, \tag{17}$$

*is convex, and as $r \to \infty$, $f_r(x) \uparrow \widehat{f}(x)$ pointwise, for all $x \in \mathbb{R}^n$.*

($b$) *If $\mathbf{K}$ has a nonempty interior $\mathrm{int}\,\mathbf{K}$, then*

$$\sup \mathbf{Q}^*_{rx} \;=\; \max \mathbf{Q}^*_{rx} = \inf \mathbf{Q}_{rx} \;=\; f_r(x), \quad x \in \mathrm{int}\,\mathbf{K}, \tag{18}$$

*and for every optimal solution $(\lambda^*_r, \gamma^*_r)$ of $\mathbf{Q}^*_{rx}$,*

$$f_r(y) - f_r(x) \;\geq\; \langle \lambda^*_r, y - x\rangle, \quad \forall\, y \in \mathbb{R}^n,$$

*that is, $\lambda^*_r \in \partial f_r(x)$.*

**Proof.** ($a$) By standard weak duality, $\sup \mathbf{Q}^*_{rx} \leq \inf \mathbf{Q}_{rx} \leq \widehat{f}(x)$ for all $r \in \mathbb{N}$, and all $x \in \mathbb{R}^n$. Next, let $x \in \mathbf{K}$ be fixed, arbitrary. From Theorem 2.2 and Corollary 2.3, $\mathbf{P}^*_x$ is solvable, and $\sup \mathbf{P}^*_x = \max \mathbf{P}^*_x = \widehat{f}(x)$ for all $x \in \mathbf{K}$. Therefore, from the definition of $\mathbf{P}^*_x$ in (6), there is some $(\gamma^*, \lambda^*) \in \mathbb{R} \times \mathbb{R}^n$ such that

$$p(y) - q(y)\langle \lambda^*, y\rangle - \gamma^* q(y) \geq 0, \quad y \in \mathbf{D},$$

and $\gamma^* - \langle \lambda^*, x\rangle = \widehat{f}(x)$.

Hence, with $\epsilon > 0$ fixed, arbitrary, $\gamma^* - \epsilon - \langle \lambda^*, x\rangle = \widehat{f}(x) - \epsilon$, and

$$p(y) - q(y)\langle \lambda^*, y\rangle - (\gamma^* - \epsilon)q(y) \geq \epsilon q(y) > 0, \quad y \in \mathbf{D}.$$

Therefore, under Assumption 3.1, the polynomial $p - q\langle \lambda^*, y\rangle - (\gamma^* - \epsilon)q$, which is strictly positive on $\mathbf{D}$, can be written

$$p(y) - q(y)\langle \lambda^*, y\rangle - (\gamma^* - \epsilon)q(y) \;=\; \sum_{j=0}^{m} u_j g_j,$$

for some s.o.s. polynomials $\{u_j\}_{j=0}^m \subset \mathbb{R}[x]$. But then, $(\gamma^* - \epsilon, \lambda^*, \{u_j\})$ is a feasible solution of $\mathbf{Q}^*_{rx}$ as soon as $r \geq r_\epsilon := \max_{j=0,1,\ldots,m} \deg(u_j g_j)$, and with value $\gamma^* - \epsilon - \langle \lambda^*, x\rangle = \widehat{f}(x) - \epsilon$. Hence, for every $\epsilon > 0$,

$$\widehat{f}(x) - \epsilon \;\leq\; \sup \mathbf{Q}^*_{rx} \;\leq \inf \mathbf{Q}_{rx} \;\leq\; \widehat{f}(x), \quad r \geq r_\epsilon.$$

Hence we obtain the convergence $f_r(x) \uparrow \widehat{f}(x)$ for all $x \in \mathbf{K}$.

Next, let $x \notin \mathbf{K}$ so that $\widehat{f}(x) = +\infty$. From the proof of Theorem 2.2, we have seen that $\sup \mathbf{P}^*_x = \widehat{f}(x)$ for *all* $x \in \mathbb{R}^n$. Therefore, with $M > 0$ fixed, arbitrarily large, one may find $\lambda \in \mathbb{R}^n, \gamma \in \mathbb{R}$ such that

$$M \leq \langle \lambda, x\rangle + \gamma \quad \text{and} \quad f(y) + \langle \lambda, y\rangle \geq \gamma, \quad \forall y \in \mathbf{D}.$$

Hence

$$f(y) + \langle \lambda, y\rangle - \gamma + \epsilon > 0, \quad \forall y \in \mathbf{D}.$$

Therefore, as $q > 0$ on $\mathbf{D}$, the polynomial $g := p + \langle \lambda, y\rangle q - (\gamma - \epsilon)q$ is positive on $\mathbf{D}$. By Putinar Positivstellensatz [15], it may be written

$$g \;=\; u_0 + \sum_{j=1}^{m} u_j\, g_j$$

for some family of s.o.s. polynomials $\{u_j\}_{j=0}^m \subset \mathbb{R}[x]$. But then, with $2r_M \geq \max[\deg u_0, \deg u_j g_j]$, the 3-uplet $(\lambda, \gamma - \epsilon, \{u_j\})$ is a feasible solution of $\mathbf{Q}^*_{rx}$, whenever $r \geq r_M$, and with value $M - \epsilon$. And so, as $M$ was arbitrarily large, $\sup \mathbf{P}^*_{rx} \to +\infty = \widehat{f}(x)$, as $r \to \infty$. Hence, we also obtain $f_r(x) \uparrow \widehat{f}(x)$ for $x \notin \mathbf{K}$.

Let us prove that $f_r$ is convex. It follows from its definition $f_r(x) = \inf \mathbf{Q}_{rx}$ for all $x \in \mathbb{R}^n$, and the definition (13) of the SDP $\mathbf{Q}_{rx}$. Observe that for all $r$ sufficiently large, say $r \geq r'_0$, $\inf \mathbf{Q}_{rx} > -\infty$ for all $x \in \mathbb{R}^n$, because $\sup \mathbf{Q}^*_{rx} \geq -1$, for all $x \in \mathbb{R}^n$. Indeed, with $\gamma = -1, \lambda = 0$, the polynomial $p + q = p - q\gamma - q\langle \lambda, y\rangle$ is positive on $\mathbf{D}$, and therefore, by Putinar Positivstellensatz [15], $p + q = u_0 + \sum_j u_j g_j$, for some family of s.o.s. polynomials $\{u_j\}_0^m$. Therefore, $(-1, 0, \{u_j\})$ is feasible for $\mathbf{Q}^*_{rx}$ with value $-1$, whenever $2r'_0 \geq \max[\deg u_0, \deg u_j g_j]$. Next, let $x := \alpha u + (1-\alpha)v$, with $u, v \in \mathbb{R}^n$, and $0 \leq \alpha \leq 1$. As we want to prove

$$\inf \mathbf{Q}_{r(\alpha u + (1-\alpha)v)} \;\leq\; \alpha \inf \mathbf{Q}_{ru} + (1-\alpha) \inf \mathbf{Q}_{rv}, \quad 0 \leq \alpha \leq 1,$$

we may restrict to $u, v \in \mathbb{R}^n$ such that $\inf \mathbf{Q}_{ru}, \inf \mathbf{Q}_{rv} < +\infty$. So, let $y_u$ (resp. $y_v$) be feasible for $\mathbf{Q}_{ru}$ (resp. $\mathbf{Q}_{rv}$), and with respective values $\inf \mathbf{Q}_{ru} + \epsilon$, $\inf \mathbf{Q}_{rv} + \epsilon$. As the matrices $M_r(y), M_{r-r_j}(g_j y)$ are all linear in $y$, and $y \mapsto L_y(\bullet)$ is linear in $y$ as well, $y := \alpha y_u + (1-\alpha) y_v$ is feasible for $\mathbf{Q}_{rx}$, with value $\alpha \inf \mathbf{Q}_{ru} + (1-\alpha) \inf \mathbf{Q}_{rv} + \epsilon$. Therefore,

$$\inf \mathbf{Q}_{rx} = \inf \mathbf{Q}_{r(\alpha u + (1-\alpha) v)} \leq \alpha \inf \mathbf{Q}_{ru} + (1-\alpha) \inf \mathbf{Q}_{rv} + \epsilon, \quad \forall \epsilon > 0,$$

and letting $\epsilon \to 0$ yields the result.

($b$) Let $\mathbf{K}$ be with a nonempty interior $\operatorname{int} \mathbf{K}$, and let $x \in \operatorname{int} \mathbf{K}$. Let $\mu$ be the probability measure uniformly distributed on the ball $B_x := \{y \in \mathbf{K} : \|y - x\| \leq \delta\} \subset \mathbf{K}$.

Hence, $\int z_i d\mu = x_i$ for all $i = 1, \ldots, n$. Next, define the measure $\nu$ to be $d\nu = q^{-1} d\mu$ so that $\int q d\nu = 1$, and $\int z_i d\mu = \int z_i q d\nu = x_i$ for all $i = 1, \ldots, n$, Take for $y = \{y_\alpha\}$, the vector of moments of the measure $\nu$. As $\nu$ has a density, and is supported on $\mathbf{K}$, it follows that $M_r(y) \succ 0$ and $M_r(q_j y) \succ 0$, $j = 1, \ldots, m$, for all $r$. Therefore, $y$ is a *strictly* feasible solution of $\mathbf{Q}_{rx}$, i.e., Slater's condition holds, which in turn implies the absence of a duality gap between $\mathbf{Q}_{rx}$ and its dual $\mathbf{Q}^*_{rx}$ ($\sup \mathbf{Q}^*_{rx} = \inf \mathbf{Q}_{rx}$). In addition, as $\inf \mathbf{Q}_{rx} > -\infty$, we get $\sup \mathbf{Q}^*_{rx} = \max \mathbf{Q}^*_{rx} = \inf \mathbf{Q}_{rx}$, which is (18).

So, as $\mathbf{Q}^*_{rx}$ is solvable, let $(\gamma^*_r, \lambda^*_r, \{u^*_j\})$ be an optimal solution, that is, $f_r(x) = \gamma^*_r - \langle \lambda^*_r, x \rangle$ and

$$p(z) - \gamma^*_r q(z) - q(z) \langle \lambda^*_r, y \rangle = \sum_{j=0}^{m} u^*_j(z) g_j(z), \quad \forall\, z \in \mathbb{R}^n.$$

Therefore, one has

$$\begin{aligned} f_r(x) &= \gamma^*_r + \langle \lambda^*_r, x \rangle \\ f_r(y) &= \sup_{\gamma, \lambda, u} \{\gamma + \langle \lambda, y \rangle \,:\, p(z) - \gamma q(z) - q(z) \langle \lambda, z \rangle = \sum_{j=0}^{m} u_j(z) g_j(z), : \forall z \in \mathbb{R}^n\} \\ &\geq \gamma^*_r + \langle \lambda^*_r, y \rangle, \quad \forall y \in \mathbb{R}^n, \end{aligned}$$

from which we get

$$f_r(y) - f_r(x) \geq \langle \lambda^*_r, y - x \rangle, \quad \forall y \in \mathbb{R}^n,$$

that is, $\lambda^*_r \in \partial f_r(x)$, the desired result. □

As a consequence, we also get:

**Corollary 3.3.** *Let* $\mathbf{D}$ *be as in* (11), *and let Assumption* 3.1 *hold. Let* $f$ *and* $\widehat{f}$ *be as in* (4) *and* (3) *respectively, and let* $f_r : \mathbf{K} \to \mathbb{R}$, *be as in Theorem* 3.2.

*Then* $f_r$ *is l.s.c. Moreover, for every compact* $\overline{\mathbf{K}} \subset \mathbf{K}$ *on which* $\widehat{f}$ *is continuous,*

$$\lim_{r \to \infty} \sup_{x \in \overline{\mathbf{K}}} |\widehat{f}(x) - f_r(x)| = 0, \tag{19}$$

*that is, the monotone nondecreasing sequence* $\{f_r\}$ *converges to* $\widehat{f}$, *uniformly on every compact on which* $\widehat{f}$ *is continuous. If in addition,* $\mathbf{D}$ *is Splitting-continuous and if* $q > 0$ *on* $\mathbf{K}$, *then the convergence is uniform on* $\mathbf{K}$.

**Proof.** Lower semicontinuity of $f_r$ may be obtained using Laraki and Sudderth [10], and is due to the facts that:

- the objective function of $\mathbf{Q}_{rx}$ does not depend on $x$, and
- the feasible set of $\mathbf{Q}_{rx}$ as a multifunction of $x$, is u.s.c. in the sense of Kuratowski.

By Theorem 3.2, we already have that $f_r \uparrow \widehat{f}$ on $\overline{\mathbf{K}}$.

(1) The convergence $f_r \uparrow \widehat{f}$ on $\overline{\mathbf{K}}$, (2) that $f_r$ is l.s.c., (3) that the limit $\widehat{f}$ is continuous, and finally (4) that $\overline{\mathbf{K}}$ is compact, imply by Dini's theorem that the convergence is *uniform* on $\overline{\mathbf{K}}$. □

**Remark 3.4.** If Assumption 3.1 does not hold, then in the SDP-relaxation $\mathbf{Q}_{rx}$ in (12), one replaces the $m$ LMI constraints $L_y(g_j h^2) \geq 0$ for all $h \in \mathcal{P}_{r-r_j}$, with the $2^m$ LMI constraints

$$L_y(g_J h^2) \geq 0, \quad \forall h \in \mathcal{P}_{r-r_J}, \quad \forall J \subseteq \{1, \dots, m\},$$

where $g_J := \prod_{j \in J} g_j$, and $r_J = \deg g_J$, for all $J \subseteq \{1, \dots, m\}$ (and $g_\emptyset \equiv 1$). Indeed, Theorem 3.2 and Corollary 3.3 remain valid with $f_r(x) := \inf \mathbf{Q}_{rx}$, for all $x \in \mathbb{R}^n$ (with the newly defined $\mathbf{Q}_{rx}$). In the proof, one now invokes Schmüdgen's Positivstellensatz [17] (instead of Putinar's Positivstellensatz [15]) which states that every polynomial $v$, strictly positive on $\mathbf{D}$, can be written as

$$v = \sum_{J \subset \{1,\dots,m\}} v_J \, g_J, \quad [(\text{compare with } (16))]$$

for some family of s.o.s. polynomials $\{v_J\} \subset \mathbb{R}[x]$; see Schmüdgen [17].

**Example 3.5.** Consider the bivariate rational function $f : [-1,1]^2 \to \mathbb{R}$:

$$(x,y) \mapsto f(x,y) := \frac{(x^2 - 1/4)(y^2 - 1/4)}{1 + x^2 + y^2}, \quad (x,y) \in [-1,1]^2,$$

displayed in Figure 3.5. The approximation $f_r(x)$ was computed using the software GloptiPoly3[4] at every point $x$ of a $30 \times 30$ grid of $[-1,1]^2$. In Figure 3.5 we have displayed $(f_3 - f_2)$ whose maximum value is 0.0023 and which is zero at most points, except in a small region around the four corners of the $[-1,1]^2$ box. The difference $f_4 - f_3$ is also displayed in one such region and one may see a significant improvement, as the maximum value is now about $10^{-5}$ and many more zeros in that region. Finally, the maximum value of $f_5 - f_4$ on the corner $[-0.8, -1] \times [0.8, 1]$ is about $3.10^{-8}$. Therefore, $f_4$ is already very close to the convex envelope $\widehat{f}$. In other words, a very good approximation is already obtained at the third relaxation $\mathbf{Q}_4$ (and even $\mathbf{Q}_3$ or $\mathbf{Q}_2$ for most points).

**Example 3.6.** Next consider the bivariate rational function $f : [-1,1]^2 \to \mathbb{R}$:

$$(x,y) \mapsto f(x,y) := \frac{xy}{1 + x^2 + y^2}, \quad (x,y) \in [-1,1]^2,$$

on $[-1,1]^2$ displayed in Figure 3.6, with $f_3$ as well. As for Example 3.5, in Figure 3.6

[4]For solving the SDP-relaxation $\mathbf{Q}_{rx}$ in (13), we have used the new version of the Software GloptiPoly3, dedicated to solving the generalized problem of moments; see www.laas.fr/∼henrion/software/gloptipoly. The authors wishes to thank D. Henrion for helpful discussions.

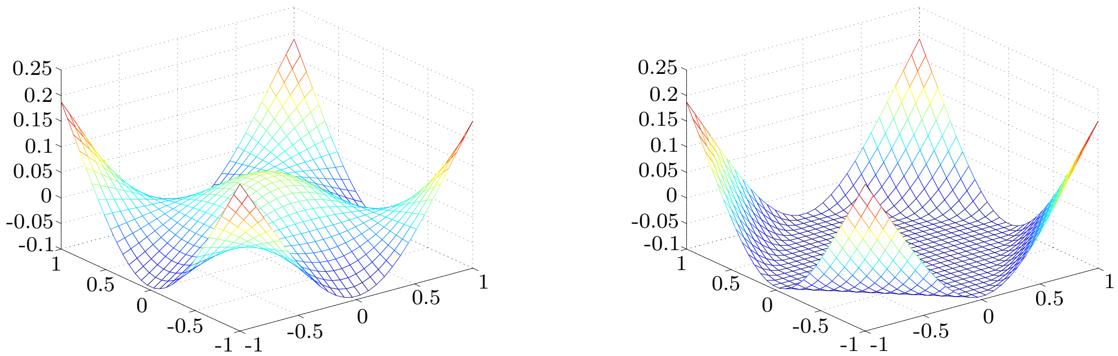


Figure 3.1: Example 3.5, $f$ and $f_3$ on $[-1,1]^2$

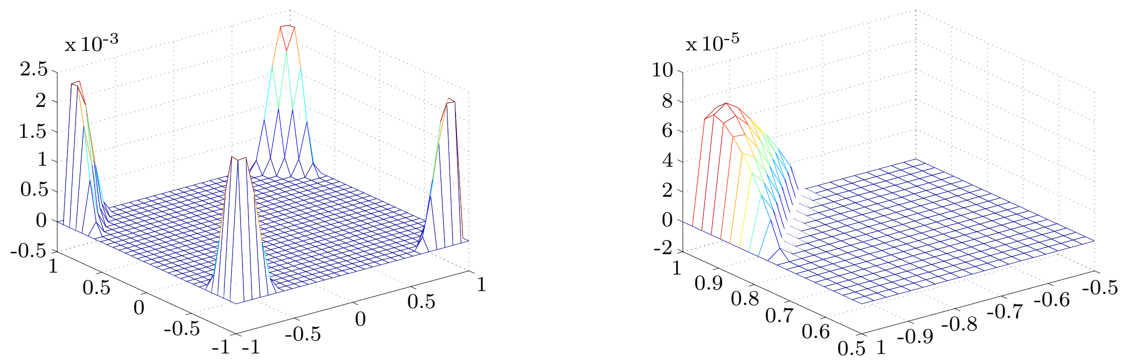


Figure 3.2: Example 3.5, $f_3 - f_2$ on $[-1,1]^2$ and $f_4 - f_3$ on one corner

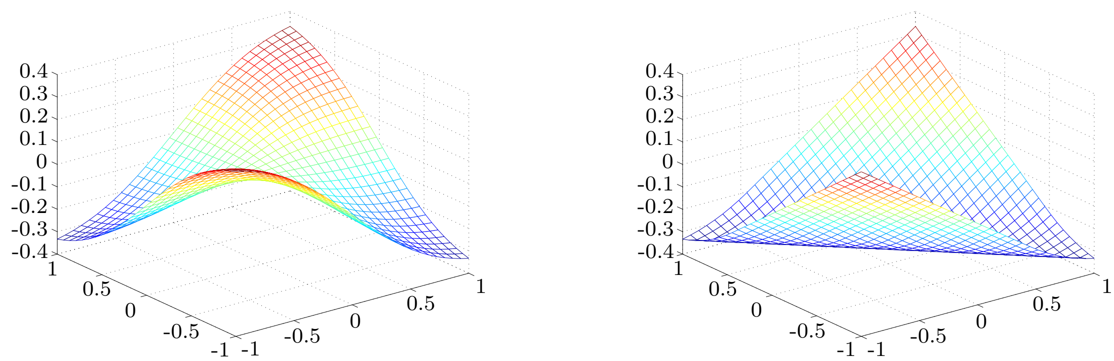


Figure 3.3: Example 3.6, $f$ and $f_3$ on $[-1,1]^2$

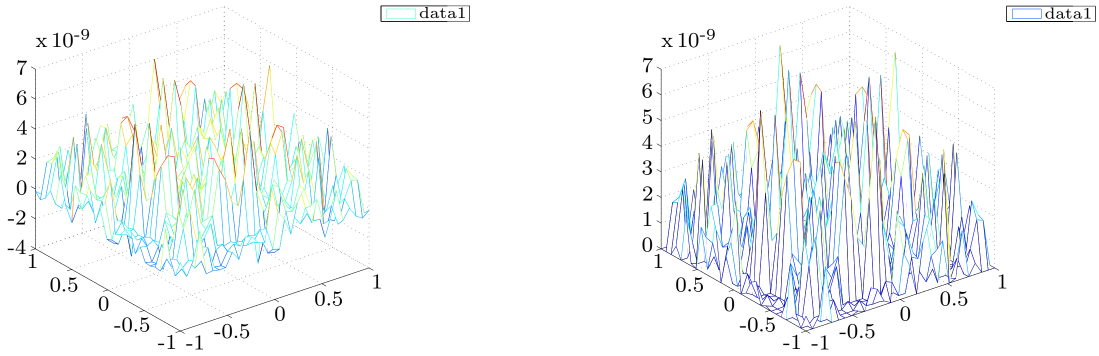


Figure 3.4: Example 3.5, $f_3 - f_2$ and $(f - 3 - f_2)^+$ on $[-1,1]^2$

we have displayed $(f_3 - f_2)$ which is of the order $10^{-9}$. That explains why again for a few values of $x \in [-1,1]^2$ one may have $f_3(x) \leq f_2(x)$ as we are at the limit of machine precision. It also means that again $f_2$ provides a very good approximation of the convex envelope $\hat{f}$, that is, a very good approximation is already obtained at the first relaxation (here $\mathbf{Q}_2$)!

### 3.3. The univariate case

In the univariate case, simplifications occur. Let $\mathbf{D} \subset \mathbb{R}$ be the interval $[a,b]$, that is, $\mathbf{D}$ has the representation

$$\mathbf{D} := \{x \in \mathbf{R} : g(x) \geq 0\}, \text{ with } x \mapsto g(x) := (b-x)(x-a), \quad x \in \mathbb{R}. \tag{20}$$

**Theorem 3.7.** *Let* $\mathbf{D}$ *be as in* (20)*,* $p, q \in \mathbf{R}[x]$*, with* $q > 0$ *on* $\mathbf{D}$*, and let* $f, \widehat{f}$ *be as in* (4) *and* (3) *respectively. Then, with* $2r \geq \max[\deg p, 1 + \deg q]$*, let* $\mathbf{Q}_{rx}$ *be the SDP-relaxation defined in* (12)*. Then:*

$$\widehat{f}(x) = \inf \mathbf{Q}_{rx}, \quad x \in \mathbf{K}. \tag{21}$$

**Proof.** Recall that when $\mathbf{D}$ is convex and compact, then for every $x \in \mathbf{D}$, $\widehat{f}(x) = \sup \mathbf{P}_x^*$, with $\mathbf{P}_x^*$ as defined in (6). Next, in the univariate case, a polynomial $h \in \mathbb{R}[x]$ of degree $2r$ or $2r-1$, is nonnegative on $\mathbb{K}$ if and only if $h = h_0 + h_1 g$, for some s.o.s. polynomials $h_0, h_1 \in \mathbb{R}[x]$, *and* with $\deg h_0, h_1 g \leq 2r$. This is in contrast with the multivariate case, where the degree in Putinar's representation (16) is not known in advance. Therefore, let $2r \geq \max[\deg p, 1 + \deg q]$. The polynomial $p - \gamma q - \langle \lambda, y \rangle q$ (of degree $\leq 2r$) is nonnegative on $\mathbf{D}$ if and only if

$$p - \gamma q - \langle \lambda, y \rangle q = u_0 + u_1 g,$$

for some s.o.s. polynomials $u_0, u_1 \in \mathbf{R}[x]$, with $\deg u_0, u_1 g \leq 2r$. Therefore, as $q > 0$ on $\mathbf{K}$, and recalling the definition of $\mathbf{P}_x^*$ in (6),

$$\begin{aligned}
f(y) - \langle \lambda, y \rangle \geq \gamma, \quad \forall y \in \mathbf{K} &\Leftrightarrow \\
p(y) - q(y)\,\langle \lambda, y \rangle \geq \gamma\, q(y), \quad \forall y \in \mathbf{K} &\Leftrightarrow \\
p(y) - q(y)\,\langle \lambda, y \rangle - \gamma\, q(y) \geq 0, \quad \forall y \in \mathbf{K} &\Leftrightarrow \\
p - q\,\langle \lambda, y \rangle - \gamma\, q = u_0 + u_1\, g,&
\end{aligned}$$

for some s.o.s. polynomials $u_0, u_1 \in \mathbb{R}[x]$, with $\deg u_0, u_1 g \leq 2r$. Therefore, $\mathbf{Q}_{rx}^*$ is identical to $\mathbf{P}_x^*$, from which the result follows. □

So, in the univariate case, the SDP-relaxation $\mathbf{Q}_{rx}$ is *exact*, that is, the value at $x \in \mathbf{K}$ of the convex envelope $\widehat{f}$, is easily obtained by solving a *single* SDP.

## 4. The convex hull of a compact basic semi-algebraic set

An important question stated in Ben-Tal and Nemirovski [3, §4.2 and §4.10.2], Parrilo and Sturmfels [14], and not settled yet, is to characterize the convex subsets of $\mathbb{R}^n$ that are *semidefinite representable* (written SDr), or equivalently, have an *LMI representation*; that is, subsets $\Omega \subset \mathbb{R}^n$ of the form

$$\Omega = \{x \in \mathbb{R}^n : \ M_0 + \sum_{i=1}^{n} M_i x_i \succeq 0\},$$

for some family $\{M_i\}_{i=0}^n$ of real symmetric matrices. In other words, a SDr set is the *feasible set* of a system of LMI's (Linear Matrix Inequalities), and powerful techniques are now available to solve SDPs. For instance, the epigraph of a univariate convex polynomial is SDr; see [3, p. 292]. Recently, Helton and Vinnikov [6] have proved that *rigid convexity* (as defined in [6]) is a necessary condition for a convex set to be SDr.

In this section, we are concerned with a (large) class of convex sets, namely the convex hull of an arbitrary compact basic semi-algebraic set, i.e., the convex hull $\mathbf{K} = \mathrm{co}(\mathbf{D})$ of a compact set $\mathbf{D}$ defined by finitely many polynomial inequalities, as in (11). We will show that:

- $\mathbf{K}$ is the *projection* of a *semi-infinite* SDr set $S_\infty$, that is, $S_\infty$ is defined by finitely many LMIs involving matrices with countably many rows and columns, and involving countably many variables. Importantly, the LMI representation of the set $S_\infty$ is given directly in terms of the data, i.e., in terms of the polynomials $g_j$'s that define the set $\mathbf{D}$.
- $\mathbf{K}$ can be approximated by a monotone nonincreasing sequence of convex sets $\{\mathbf{K}_r\}$ (with $\mathbf{K}_r \supset \mathbf{K}$ for all $r$), that are *projections* of SDr sets $S_r$. Each SDr set $S_r$ is a "finite truncation" of $S_\infty$, and therefore, also has a specific LMI representation, directly in terms of the data defining the set $\mathbf{D}$. In other words, $\{\mathbf{K}_r\}$ is a converging sequence of *outer convex approximations* of $\mathbf{K}$, i.e. $\mathbf{K}_r \downarrow \mathbf{K}$ as $r \to \infty$. Detecting whether a point $x \in \mathbf{R}^n$ belongs to $\mathbf{K}_r$ reduces to solving a single SDP that involves the SDr set $S_r$.

With $\mathbf{D} \subset \mathbb{R}^n$ as in (11), define the $2^m$ polynomials

$$x \mapsto g_J(x) := \prod_{j \in J} g_j, \quad \forall\, J \subseteq \{1, \ldots, m\}, \tag{22}$$

of total degree $2r_J$ or $2r_J - 1$, and with the convention that $g_\emptyset \equiv 1$.

Let $M_r(g_J y) \in \mathbb{R}^{s(r) \times s(r)}$ be the localizing matrix associated with the polynomial $g_J$, and a sequence $y$, for all $J \subseteq \{1, \ldots, m\}$, and all $r = 0, 1, \ldots$; see also §3.1 for the definition of the infinite matrix $M_\infty(g_J y)$.

Define $S_\infty \subset \mathbb{R}^\infty$ by:

$$S_\infty := \{ y \in \mathbb{R}^\infty : y_0 = 1;\ M_\infty(g_J y) \succeq 0, \quad \forall J \subseteq \{1,\ldots,m\}\}, \tag{23}$$

The set $S_\infty$ is a *semi-infinite* SDr set as it is defined by $2^m$ LMIs whose matrices have countably many rows and columns, and with countably many variables.

If Assumption 3.1 holds, one may instead use the simpler semi-infinite SDr set

$$S'_\infty := \{ y \in \mathbb{R}^\infty : y_0 = 1;\ M_\infty(y) \succeq 0,\ M_\infty(g_j y) \succeq 0, \quad \forall j = 1,\ldots,m\}.$$

Similarly, define $\mathbf{K}_\infty \subset \mathbb{R}^n$ by:

$$\mathbf{K}_\infty := \{x \in \mathbb{R}^n : \exists\, y \in S_\infty : \text{ s.t. } L_y(z_i) = x_i, i = 1,\ldots,n\}. \tag{24}$$

**Lemma 4.1.** *Let $\mathbf{D} \subset \mathbb{R}^n$ be as in* (11) *and compact, and let $\mathbf{K}_\infty$ be as in* (24)*. Then $\mathbf{K}_\infty = \mathbf{K} = \mathrm{co}(\mathbf{D})$.*

**Proof.** If $x \in \mathrm{co}(\mathbf{D}) = \mathbf{K}$, then $x_i = \int z_i\, d\mu$, $\forall i = 1,\ldots,n$, for some probability measure $\mu$ with support contained in $\mathbf{D}$. Let $y$ be the vector of moments of $\mu$, well defined because $\mu$ has compact support. Then we necessarily have $y_0 = 1$, and $M_r(g_J y) \succeq 0$ for all $r$ and all $J \subseteq \{1,\ldots,m\}$; see §3.1. Equivalently, $M_\infty(g_J y) \succeq 0$, for all $J \subseteq \{1,\ldots,m\}$, and so $y \in S_\infty$. From, $\int z_i\, d\mu, = L_y(z_i), : \forall i = 1,\ldots,n$, we conclude that $x \in \mathbf{K}_\infty$, and so $\mathbf{K} \subseteq \mathbf{K}_\infty$.

Conversely, let $x \in \mathbf{K}_\infty$. Then, there exists $y \in S_\infty$ such that $y_0 = 1$ and $L_y(z_i) = x_i$, for all $i = 1,\ldots,n$. As $M_\infty(g_J y) \succeq 0$, for all $J \subseteq \{1,\ldots,m\}$, then by Schmüdgen Positivstellensatz [17], $y$ is the vector of moments of some probability measure $\mu_y$, with support contained in $\mathbf{D}$. Next,

$$L_y(z_i) = x_i,\ \forall i = 1,\ldots,n \quad \Leftrightarrow \quad x_i = \int z_i\, d\mu_y,\ \forall i = 1,\ldots,n,$$

which proves that $x \in \mathrm{co}(\mathbf{D}) = \mathbf{K}$. Therefore, $\mathbf{K}_\infty \subseteq \mathbf{K}$, and the result follows. □

So, Lemma 4.1 states that the convex hull $\mathbf{K}$ of any compact basic semi-algebraic set $\mathbf{D}$, is the *projection* on the variables $y_\alpha$ with $|\alpha| = 1$, of the semi-infinite SDr set $S_\infty$ (recall that for every $\alpha \in \mathbb{N}^n$, $|\alpha| = \sum_{i=1}^n \alpha_i$). However, the set $S_\infty$ is *not* described by finite-dimensional LMIs, because we have countably many variables $y_\alpha$, and matrices with infinitely many rows and columns.

We next provide outer approximations $\{\mathbf{K}_r\}$ of $\mathbf{K}$, which are projections of SDr sets $\{S_r\}$, with $S_r \supset S_\infty$, for all $r$, and $\mathbf{K}_r \downarrow \mathbf{K}$, as $r \to \infty$.

With $r \geq r_0$, let $S_r \subset \mathbb{R}^{s(2r)}$ be defined as:

$$S_r := \{y \in \mathbb{R}^{s(2r)} : y_0 = 1;\ M_{r-r_J}(g_J y) \succeq 0, \ \forall J \subseteq \{1,\ldots,m\}\}. \tag{25}$$

Notice that $S_r$ is a SDr set obtained from $S_\infty$ by "finite" truncation. Indeed, $S_r$ contains finitely many variables $y_\alpha$, namely those with $|\alpha| \leq 2r$. And $M_r(g_J y)$ is a finite truncation of the infinite matrix $M_\infty(g_j y)$; see §3.1.

As for $S_\infty$, under Assumption 3.1, $S_r$ in (25) may be replaced with the (simpler) SDr set

$$S'_r := \{y \in \mathbb{R}^{s(2r)} : y_0 = 1;\; M_r(y) \succeq 0, M_{r-r_j}(g_j y) \succeq 0, \quad j = 1, \ldots, m\}.$$

Next, let

$$\mathbf{K}_r := \{x \in \mathbb{R}^n : \exists\, y \in S_r : \text{ s.t. } L_y(z_i) = x_i, i = 1, \ldots, n\}. \tag{26}$$

Equivalently,

$$\mathbf{K}_r := \left\{ x \in \mathbb{R}^n : \exists\, y \in \mathbb{R}^{s(2r)} : \begin{cases} L_y(z_i) = x_i, & i = 1, \ldots, n \\ M_{r-r_J}(g_J y) \succeq 0, & J \subseteq \{1, \ldots, m\} \\ y_0 = 1. \end{cases} \right\}. \tag{27}$$

In view of the meaning of $L_y(z_i)$, $\mathbf{K}_r$ is the projection on $\mathbb{R}^n$ of the SDr set $S_r \subset \mathbb{R}^{s(2r)}$ defined in (25) (on the $n$ variables $y_\alpha$, with $|\alpha| = 1$); Obviously, $\{\mathbf{K}_r\}$ forms a *nested* sequence of sets, and we have

$$\mathbf{K}_{r_0} \supset \mathbf{K}_{r_0+1} \ldots \supset \mathbf{K}_r \ldots \supset \mathbf{K}. \tag{28}$$

Let $f$ be the identity on $\mathbf{D}$ ($f = 1$ on $\mathbf{D}$). Then its convex envelope $\widehat{f}$ is given by

$$\widehat{f}(x) = \begin{cases} 1, & x \in \mathbf{K}, \\ +\infty & x \in \mathbf{R}^n \setminus \mathbf{K}. \end{cases}$$

Note that $\widehat{f}$ is clearly a continuous function on $\mathbf{K}$.

On $\mathbf{D}$, write $f = 1 = p/q$ with $p = q \equiv 1$, so that with $r \geq r_0 := \max_J r_J$, the SDP-relaxation $\mathbf{Q}_{rx}$ defined in (12) and in Remark 3.4, now reads

$$\mathbf{Q}_{rx} : \inf_y \{\, y_0 : \, y \in S_r;\; L_y(z_i) = x_i, \, i = 1, \ldots, n\}, \quad x \in \mathbb{R}^n, \tag{29}$$

and so, for all $r \geq r_0$,

$$\inf \mathbf{Q}_{rx} = \begin{cases} 1, & \text{if } x \in \mathbf{K}_r \\ +\infty, & \text{otherwise.} \end{cases} \tag{30}$$

Next, let $f_r : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be the function $x \mapsto f_r(x) := \inf \mathbf{Q}_{rx}$, with obvious domain $\mathbf{K}_r$.

**Corollary 4.2.** *Let $\mathbf{D} \subset \mathbb{R}^n$ be compact and defined as in* (11), *and let* $\mathbf{K} := \mathrm{co}(\mathbf{D})$.

($a$) *If $x \notin \mathbf{K}$, then $f_r(x) = +\infty$ whenever $r \geq r_x$, for some integer $r_x$.*

($b$) *With $\mathbf{K}_r$ being as in* (27), *$\mathbf{K}_r \downarrow \mathbf{K}$ as $r \to \infty$.*

**Proof.** ($a$) By Theorem 3.2($b$) and Remark 3.4, $f_r$ is convex and $f_r(x) \uparrow \widehat{f}(x)$, for all $x \in \mathbb{R}^n$. If $x \in \mathbf{K}$ then $f_r(x) = 1$ for all $r$. If $x \notin \mathbf{K}$ then $f_r(x) = 1$ if $x \in \mathbf{K}_r$, and $+\infty$ outside $\mathbf{K}_r$. But as $f_r(x) \uparrow \widehat{f}(x) = +\infty$, there is some $r_x$ such that $f_r(x) = +\infty$, for all $r \geq r_x$, the desired result.

$(b)$ As $\{\mathbf{K}_r\}$ is a nonincreasing nested sequence and $\mathbf{K} \subset \mathbf{K}_r$ for all $r$, one has

$$\mathbf{K}_r \downarrow \mathbf{K}^* := \bigcap_{r=0}^{\infty} \mathbf{K}_r \supset \mathbf{K}.$$

It suffices to show that $\mathbf{K}^* \subseteq \mathbf{K}$, which we prove by contradiction. Let $x \in \mathbf{K}^*$, and suppose that $x \notin \mathbf{K}$. By $(a)$, we must have $f_r(x) = +\infty$ whenever $r \geq r_x$, for some integer $r_x$. In other words, $x \notin \mathbf{K}_r$ whenever $r \geq r_x$. But then, $x \notin \mathbf{K}^*$, in contradiction with our hypothesis. $\square$

Corollary 4.2 provides us with a means to test whether $x \notin \mathbf{K}$. Indeed, it suffices to solve the SDP-relaxation $\mathbf{Q}_{rx}$ defined in (29), until $\inf \mathbf{Q}_{rx} = +\infty$ for some $r$ (which means that $x \notin \mathbf{K}_r$ for all $r \geq r_x$), which eventually happens if $x \notin \mathbf{K}$.

## References


[1] G. A. Anastassiou: Moments in Probability and Approximation Theory, Longman, Harlow (1993).

[2] J. Benoist, J.-B. Hiriart-Urruty: What is the subdifferential of the closed convex hull of a function?, SIAM J. Math. Anal. 27 (1996) 1661–1679.

[3] A. Ben-Tal, A. Nemirovski: Lectures on Modern Convex Optimization, SIAM, Philadelphia (2001).

[4] B. Brighi, M. Chipot: Approximated convex envelope of a function, SIAM J. Numer. Anal. 31 (1994) 128–148.

[5] G. Choquet: Lectures on Analysis. I: Integration and Topological Vector Spaces, W. A. Benjamin, New York (1969).

[6] J. W. Helton, V. Vinnikov: Linear matrix inequality representation of sets, Commun. Pure Appl. Math. 60 (2007) 654–674.

[7] T. Jacobi, A. Prestel: Distinguished representations of strictly positive polynomials, J. Reine. Angew. Math. 532 (2001) 223–235.

[8] J. H. B. Kemperman: Geometry of the moment problem, in: Moments in Mathematics, H. J. Landau (ed.), Proc. Symp. Appl. Math. 37, American Mathematical Society, Providence (1987) 16–53.

[9] J. B. Kruskal: Two convex counterexamples: a discontinuous envelope function and a non differentiable nearest-point mapping, Proc. Amer. Math. Soc. 23 (1969) 697–703.

[10] R. Laraki, W. D. Sudderth: The preservation of continuity and Lipschitz continuity by optimal reward operators, Math. Oper. Res. 29 (2004) 672–685.

[11] R. Laraki: On the regularity of the convexification operator on a compact set, J. Convex Analysis 11 (2004) 209–234.

[12] J. B. Lasserre: Global optimization with polynomials and the problem of moments, SIAM J. Optim. 11 (2001) 796–817.

[13] A. S. Lewis, P. Parrilo, M. V. Ramana: The Lax conjecture is true, Proc. Amer. Math. Soc. 133 (2005) 2495–2499.

[14] P. Parrilo, B. Sturmfels: Minimizing polynomial functions, in: Algorithmic and Quantitative Real Algebraic Geometry, S. Basu, L. Gonzalez-Vega (eds.), American Mathematical Society, Providence (2003) 83–100.

[15] M. Putinar: Positive polynomials on compact semi-algebraic sets, Indiana Univ. Math. J. 42 (1993) 969–984.

[16] R. T. Rockafellar: Convex Analysis, Princeton University Press, Princeton (1970).

[17] K. Schmüdgen: The $\mathbb{K}$-moment problem for compact semi-algebraic sets, Math. Ann. 289 (1991) 203–206.

[18] L. Vandenberghe, S. Boyd: Semidefinite programming, SIAM Rev. 38 (1996) 49–95.