52nd IEEE Conference on Decision and Control
December 10-13, 2013. Florence, Italy

A generalization of the Löwner-John's ellipsoid theorem

Jean B. Lasserre

***Abstract*— We provide the following generalization of the Löwner-John's ellipsoid theorem. Given a (non necessarily convex) compact set $\mathbf{K} \subset \mathbb{R}^n$ and an even integer $d \in \mathbb{N}$, there is a unique homogeneous polynomial g of degree d such that $\mathbf{K} \subset \mathbf{G} := \{\mathbf{x} : g(\mathbf{x}) \leq 1\}$ and $\mathbf{G}$ has minimum volume among all such sets. The symmetric case of the Lowner-John theorem is a particular case when $d = 2$, and importantly, we neither require the set $\mathbf{K}$ nor the sublevel set $\mathbf{G}$ to be convex. We also provide a numerical scheme to approximate the optimal value and the unique optimal solution as closely as desired.**

I. Introduction

Quoting Calafiore [5], *"The problem of approximating observed data with simple geometric primitives is, since the time of Gauss, of fundamental importance in many scientific endeavors"*. And for practical purposes and numerical efficiency the most commonly used are polyhedra and ellipsoids and such techniques are ubiquitous in several different area, control, statistics, computer graphics, computer vision, to mention a few. For instance:

• In *robust linear control* convex inner approximations of the stability region have been proposed in form of polytopes in Nurges [20], ellipsoids in Henrion et al. [12], and more general convex sets defined by Linear Matrix Inequalities (LMIs) in Henrion et al. [14], and Karimi et al. [15].

• In *statistics* and *pattern separation*, one is interested in the ellipsoid ξ of minimum volume covering some given k of m data points or separating two sets of data points. See e.g. [7], Rousseeuw [27], Croux et al. [7], Rousseeuw and Leroy [27]. Calafiore [5], Sun and Freund [28] and references therein.

• Other clustering techniques in *computer graphics*, *computer vision* and *pattern recognition*, use various (geometric or algebraic) distances (e.g. the equation error) and compute the best ellipsoid by minimizing an associated non linear least squares criterion. See e.g. Bookstein [4], Pratt [21], Rosin [25], Rosin and West [26], Taubin [29], Gander et al. [9], and in another context, Chernousko [6].

The ellipsoid of minimum volume Ω that contains a convex body $\mathbf{K} \subset \mathbb{R}^n$ is a classical and famous mathematical problem whose unique optimal solution (called the *Löwner-John's* ellipsoid) is characterized by contacts points on the boundary $\mathbf{K} \cap \Omega$. See Ball [2], Henk [11] and the many references therein. It reduces to a (often tractable) convex optimization problem; see e.g. Calafiore [5], Sun and Freund [28], and the many references therein.

J.B. Lasserre is with LAAS-CNRS and the Institute of Mathematics, University of Toulouse, France. lasserre@laas.fr

A more general optimal data fitting problem

In the Löwner-John problem one restricts to convex bodies $\mathbf{K}$ because for a non convex set $\mathbf{K}$ the optimal ellipsoid is also solution to the problem where $\mathbf{K}$ is replaced with its convex hull $\mathrm{co}(\mathbf{K})$. However, if one considers sets that are more general than ellipsoids, an optimal solution for $\mathbf{K}$ is not necessarily the same as for $\mathrm{co}(\mathbf{K})$, and indeed, in some applications one is interested in approximating as closely a non convex set $\mathbf{K}$. In this case a non convex approximation is sometimes highly desirable as more accurate.

For instance, in the robust control problem already alluded to above, in Henrion and Lasserre [13] we have provided an inner approximation of the stability region $\mathbf{K}$ by the sublevel set $\mathbf{G} = \{\mathbf{x} : g(\mathbf{x}) \leq 0\}$ of a non convex polynomial g. By allowing the degree of g to increase one obtains the convergence $\mathrm{vol}(\mathbf{G}) \to \mathrm{vol}(\mathbf{K})$ which is impossible with the convex polytopes, ellipsoids and LMI approximations described in [20], [12], [14], [15].

So if one considers the more general data fitting problem where $\mathbf{K}$ and/or the (outer) approximating set are allowed to be non convex, can we still infer interesting conclusions as for the Löwner-John problem? Can we also derive a practical algorithm for computing an optimal solution? The purpose of this paper is to provide results in this direction that can be seen as a non convex generalization of the Lowner-John's problem but, surprisingly, still a convex optimization problem. Namely, we address the following problem $\mathbf{P}$ which is a generalization of the Löwner-John problem:

$\mathbf{P}$*: Let* $\mathbf{K} \subset \mathbb{R}^n$ *be a compact set (not necessarily convex) and let* d *be an even integer. Find an homogeneous polynomial* g *of degree* d *such that its sublevel set* $\mathbf{G}_1 := \{\mathbf{x} : g(\mathbf{x}) \leq 1\}$ *contains* $\mathbf{K}$ *and has minimum volume among all such sub level sets with this inclusion property.*

Necessarily g is nonnegative otherwise $\mathbf{G}_1$ is not bounded. Of course, when $d = 2$ then g is convex (i.e., $\mathbf{G}_1$ is an ellipsoid) because every nonnegative quadratic form is convex, and g is an optimal solution for problem $\mathbf{P}$ with $\mathbf{K}$ or its convex hull $\mathrm{co}(\mathbf{K})$. That is, one retrieves the Löwner-John's problem. But when $d > 2$ then $\mathbf{G}_1$ is not necessarily convex. For instance, take $\mathbf{K} = \{\mathbf{x} : g(\mathbf{x}) \leq 1\}$ where g is some nonnegative homogeneous polynomial such that $\mathbf{K}$ is compact but non convex. Then g is an optimal solution for problem $\mathbf{P}$ with $\mathbf{K}$ and cannot be optimal for $\mathrm{co}(\mathbf{K})$; a two-dimensional example is $(x, y) \mapsto g(x, y) := x^2y^2 + \epsilon(x^4 + y^4)$ for $\epsilon > 0$ sufficiently small. See Figure 1 below.

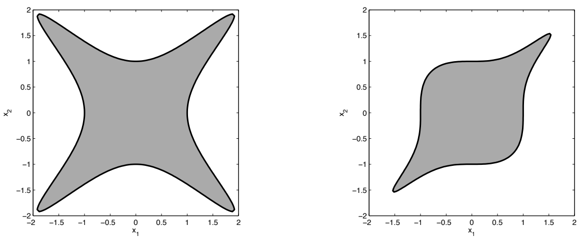

Fig. 1. $\mathbf{G}_1$ with $x^4+y^4-1.925\,x^2y^2$ and $x^6+y^6-1.925\,x^3y^3$

Contribution

Our contribution is to show that problem $\mathbf{P}$ is indeed a natural generalization of the Löwner-John problem in the sense that $\mathbf{P}$ also has a *unique* solution g^* which is also characterized by s contact points in $\mathbf{K}\cap\mathbf{G}_1$, where s is now bounded by $\binom{n+d-1}{d}$. And so when $d=2$ we retrieve the symmetric Löwner-John problem as a particular case.

To show the above result we prove another crucial result, namely that the Lebesgue-volume function $g\mapsto f(g):=\mathrm{vol}(\mathbf{G}_1)$ is a strictly convex function of the coefficients of g, which is far from being obvious from its definition. And importantly, neither $\mathbf{K}$ nor $\mathbf{G}_1$ are required to be convex!

Finally, we also provide a numerical scheme to approximate the optimal value as closely as desired. It consists of solving a hierarchy of convex optimization problems with a strictly convex objective function and a feasible set defined by Linear Matrix Inequalities (LMIs).

II. Notation, definitions and preliminary results

A. Notation and definitions

Let $\mathbb{R}[\mathbf{x}]$ be the ring of polynomials in the variables $\mathbf{x}=(x_1,\ldots,x_n)$ and let $\mathbb{R}[\mathbf{x}]_d$ be the vector space of polynomials of degree at most d (whose dimension is $s(d):=\binom{n+d}{n}$). For every $d\in\mathbb{N}$, let $\mathbb{N}^n_d:=\{\alpha\in\mathbb{N}^n:|\alpha|\,(=\sum_i\alpha_i)\leq d\}$, and let $\mathbf{v}_d(\mathbf{x})=(\mathbf{x}^\alpha)$, $\alpha\in\mathbb{N}^n$, be the vector of monomials of the canonical basis $(\mathbf{x}^\alpha)$ of $\mathbb{R}[\mathbf{x}]_d$.

A polynomial $f\in\mathbb{R}[\mathbf{x}]_d$ is written

$$\mathbf{x}\mapsto f(\mathbf{x})\;=\;\sum_{\alpha\in\mathbb{N}^n} f_\alpha\,\mathbf{x}^\alpha,$$

for some vector of coefficients $\mathbf{f}=(f_\alpha)\in\mathbb{R}^{s(d)}$. A polynomial $f\in\mathbb{R}[\mathbf{x}]_d$ is homogeneous of degree d if $f(\lambda\mathbf{x})=\lambda^d f(\mathbf{x})$ for all $\mathbf{x}\in\mathbb{R}^n$ and all $\lambda\in\mathbb{R}$.

Let us denote by $\mathbf{P}[\mathbf{x}]_d\subset\mathbb{R}[\mathbf{x}]_d$, $d\in\mathbb{N}$, the set of nonnegative and homogeneous polynomials of degree d such that their sublevel set $\mathbf{G}_1:=\{\mathbf{x}:g(\mathbf{x})\leq 1\}$ has finite Lebesgue volume (denoted $\mathrm{vol}(\mathbf{G}_1)$). Notice that necessarily d is even and $0\notin\mathbf{P}[\mathbf{x}]_d$.

For $d\in\mathbb{N}$ and a closed set $\mathbf{K}\subset\mathbb{R}^n$, denote by $C_d(\mathbf{K})$ the convex cone of all polynomials of degree d that are nonnegative on $\mathbf{K}$, and denote by $\mathcal{M}(\mathbf{K})$ the Banach space of signed Borel measures with support contained in $\mathbf{K}$ (equipped with the total variation norm). Let $M(\mathbf{K})\subset\mathcal{M}(\mathbf{K})$ be the convex cone of finite Borel measures on $\mathbf{K}$.

In the Euclidean space $\mathbb{R}^n$ we denote by $\langle\cdot,\cdot\rangle$ the usual duality bracket.

B. Some preliminary results

We first have the following result:

Proposition 1: The set $\mathbf{P}[\mathbf{x}]_d$ is a convex cone.

Proof: Let $g,h\in\mathbf{P}[\mathbf{x}]_d$ with associated sublevel sets $\mathbf{G}_1=\{\mathbf{x}:g(\mathbf{x})\leq 1\}$ and $\mathbf{H}_1=\{\mathbf{x}:h(\mathbf{x})\leq 1\}$. For $\lambda\in(0,1)$, consider the nonnegative homogeneous polynomial $\theta:=\lambda g+(1-\lambda)h\in\mathbb{R}[\mathbf{x}]_d$, with associated sublevel set

$$\Theta_1\,:=\,\{\mathbf{x}:\theta(\mathbf{x})\leq 1\}\,=\,\{\mathbf{x}:\lambda g(\mathbf{x})+(1-\lambda)h(\mathbf{x})\leq 1\}.$$

Write $\Theta_1=\Theta_1^1\cup\Theta_1^2$ where $\Theta_1^1=\Theta_1\cap\{\mathbf{x}:g(\mathbf{x})\geq h(\mathbf{x})\}$ and $\Theta_1^2=\Theta_1\cap\{\mathbf{x}:g(\mathbf{x})<h(\mathbf{x})\}$. Observe that $\mathbf{x}\in\Theta_1^1$ implies $h(\mathbf{x})\leq 1$ and so $\Theta_1^1\subset\mathbf{H}_1$. Similarly $\mathbf{x}\in\Theta_1^2$ implies $g(\mathbf{x})\leq 1$ and so $\Theta_1^2\subset\mathbf{G}_1$. Therefore $\mathrm{vol}(\Theta_1)\leq\mathrm{vol}(\mathbf{G}_1)+\mathrm{vol}(\mathbf{H}_1)<\infty$. And so $\theta\in\mathbf{P}[\mathbf{x}]_d$. ■

Next, with $y\in\mathbb{R}$ and $g\in\mathbb{R}[\mathbf{x}]$ let $\mathbf{G}_y:=\{\mathbf{x}:g(\mathbf{x})\leq y\}$.

Lemma 1: Let $g\in\mathbf{P}[\mathbf{x}]_d$. Then for every $y\geq 0$:

$$\mathrm{vol}(\mathbf{G}_y)\;=\;\frac{y^{n/d}}{\Gamma(1+n/d)}\int_{\mathbb{R}^n}\exp(-g(\mathbf{x}))\,d\mathbf{x}. \qquad (1)$$

Proof: As $g\in\mathbf{P}[\mathbf{x}]_d$, and using homogeneity, $\mathrm{vol}(\mathbf{G}_1)<\infty$ implies $\mathrm{vol}(\mathbf{G}_y)<\infty$ for every $y\geq 0$. Let $f:\mathbb{R}\to\mathbb{R}$ be the function $y\mapsto f(y):=\mathrm{vol}(\mathbf{G}_y)$. Since g is nonnegative, the function f vanishes on $(-\infty,0]$. Its Laplace transform $\mathcal{L}_f:\mathbb{C}\to\mathbb{C}$ is the function

$$\lambda\mapsto\mathcal{L}_f(\lambda):=\int_0^\infty\exp(-\lambda y)f(y)\,dy,\qquad \Re\,\lambda\,>\,0.$$

Observe that whenever $\lambda\in\mathbb{R}$ with $\lambda>0$, using a Fubini interchange and homogeneity,

$$\begin{aligned}\mathcal{L}_f(\lambda)\;&=\;\frac{1}{\lambda^{1+n/d}}\int_{\mathbb{R}^n}\exp(-g(\mathbf{z}))\,d\mathbf{z}\\ &=\;\frac{\int_{\mathbb{R}^n}\exp(-g(\mathbf{z}))\,d\mathbf{z}}{\Gamma(1+n/d)}\,\mathcal{L}_{y^{n/d}}(\lambda).\end{aligned}$$

And so, by analyticity of the Laplace transform $\mathcal{L}_f$,

$$f(y)\;=\;\frac{y^{n/d}}{\Gamma(1+n/d)}\int_{\mathbb{R}^n}\exp(-g(\mathbf{x}))\,d\mathbf{x},\qquad y\geq 0,$$

which is the desired result. ■

And we also conclude:

Corollary 1: $g\in\mathbf{P}[\mathbf{x}]_d\Longleftrightarrow\displaystyle\int_{\mathbb{R}^n}\exp(-g(\mathbf{x}))d\mathbf{x}<\infty.$

Formula (1) relating the Lebesgue volume $\mathbf{G}_1$ with $\int\exp(-g)$ is already proved (with a different argument) in Morozov and Shakirov [18], [19] where the goal of the authors is to try to express the non Gaussian integral $\int\exp(-g)$ in terms of algebraic invariants of g.

Sensitivity analysis and convexity

We now investigate some properties of the function:

$$g\mapsto f(g)\;:=\;\mathrm{vol}(\mathbf{G}_1)\;=\;\int_{\mathbf{G}_1}d\mathbf{x},\quad g\in\mathbf{P}(\mathbf{x}]_d, \qquad (2)$$

i.e., a function of the vector $\mathbf{g}=(g_\alpha)\in\mathbb{R}^{s(d)}$ of coefficients of g in the canonical basis of $\mathbb{R}[\mathbf{x}]_d$.

Theorem 1: The *Lebesgue-volume* function $f : \mathbf{P}(\mathbf{x})_d \to \mathbb{R}$ defined in (2) is strictly convex with gradient ∇f and Hessian $\nabla^2 f$ given by:

$$\frac{\partial f(g)}{\partial g_\alpha} = \frac{-1}{\Gamma(1+n/d)} \int_{\mathbb{R}^n} \mathbf{x}^\alpha \exp(-g(\mathbf{x}))\, d\mathbf{x}, \quad (3)$$

for all $\alpha \in \mathbb{N}^n_d$, $|\alpha| = d$.

$$\frac{\partial^2 f(g)}{\partial g_\alpha \partial g_\beta} = \frac{1}{\Gamma(1+n/d)} \int_{\mathbb{R}^n} \mathbf{x}^{\alpha+\beta} \exp(-g(\mathbf{x}))\, d\mathbf{x}, \quad (4)$$

for all $\alpha, \beta \in \mathbb{N}^n_d$, $|\alpha| = |\beta| = d$.

Moreover, we also have

$$\int_{\mathbb{R}^n} g(\mathbf{x}) \exp(-g(\mathbf{x}))\, d\mathbf{x} = \frac{n}{d} \int_{\mathbb{R}^n} \exp(-g(\mathbf{x}))\, d\mathbf{x}. \quad (5)$$

Proof: Recall that by Lemma 1 $f(g) = \Gamma(1 + n(d))^{-1} \int_{\mathbb{R}^n} \exp(-g(\mathbf{x}))d\mathbf{x}$. Let $p, q \in \mathbf{P}[\mathbf{x}]_d$ and $\alpha \in [0, 1]$. By convexity of $u \mapsto \exp(-u)$,

$$\begin{aligned} f(\alpha g + (1-\alpha)q) &\leq \int_{\mathbb{R}^n} [\, \alpha \exp(-g(\mathbf{x})) \\ &\quad +(1-\alpha)\exp(-q(\mathbf{x}))\,]\, d\mathbf{x} \\ &= \alpha f(g) + (1-\alpha) f(q), \end{aligned}$$

and so f is convex. Next, in view of the strict convexity of $u \mapsto \exp(-u)$, equality may occur only if $g(\mathbf{x}) = q(\mathbf{x})$ almost everywhere, which implies $g = q$ and which in turn implies strict convexity of f.

To obtain (3)-(4) one takes partial derivatives under the integral sign, which in this context is allowed. Indeed, write g in the canonical basis as $g(\mathbf{x}) = \sum_\alpha g_\alpha \mathbf{x}^\alpha$. For every $\alpha \in \mathbb{N}^n_d$, let $e_\alpha = (e_\alpha(\beta)) \in \mathbb{R}^{s(d)}$ be such that $e_\alpha(\beta) = \delta_{\beta=\alpha}$ (with δ being the Kronecker symbol). Then for every $t \geq 0$,

$$\frac{f(g+te_\alpha) - f(g)}{t} = \int_{\mathbb{R}^n} \exp(-g) (\underbrace{\frac{\exp(-t\mathbf{x}^\alpha) - 1}{t}}_{\psi(t,\mathbf{x})}) d\mathbf{x}.$$

Notice that for every $\mathbf{x}$, by convexity of the function $t \mapsto \exp(-t\mathbf{x}^\alpha)$, $\lim_{t\downarrow 0} \psi(t, \mathbf{x}) = -\mathbf{x}^\alpha$. Hence, the one-sided directional derivative $f'(g; \mathrm{e}_\alpha)$ in the direction e_α satisfies

$$\begin{aligned} f'(g; e_\alpha) &= \lim_{t\downarrow 0} \frac{f(g+te_\alpha) - f(g)}{t} \\ &= \int_{\mathbb{R}^n} -\mathbf{x}^\alpha \exp(-g)\, d\mathbf{x}, \end{aligned}$$

where we have used the Extended Monotone Convergence Theorem. With exactly same arguments as before,

$$\begin{aligned} f'(g; -e_\alpha) &= \lim_{t\downarrow 0} \frac{f(g-te_\alpha) - f(g)}{t} \\ &= \int_{\mathbb{R}^n} \mathbf{x}^\alpha \exp(-g)\, d\mathbf{x} = -f'(g; e_\alpha), \end{aligned}$$

and so
$$\frac{\partial f(g)}{\partial g_\alpha} = -\int_{\mathbb{R}^n} \mathbf{x}^\alpha \exp(-g)\, d\mathbf{x},$$

for every α with $|\alpha| \leq d$, which yields (3). Similar arguments can used for the Hessian $\nabla^2 f(g)$ which yields (4).

To obtain (5) observe that $g \mapsto h(g) := \int \exp(-g) d\mathbf{x}$, $g \in \mathbf{P}[\mathbf{x}]_d$, is a positively homogeneous function of degree $-n/d$, continuously differentiable. Hence (5) follows from (3) combined with Euler's identity, ∎

Notice that convexity of f is not obvious at all from its definition (2) whereas it becomes transparent when using formula (1).

C. The dual cone of $C_d(\mathbf{K})$

For a convex cone $C \subset \mathbb{R}^n$, the convex cone

$$C^* := \{\mathbf{y} : \langle \mathbf{y}, \mathbf{x} \rangle \geq 0 \quad \forall \mathbf{x} \in C\},$$

is the *dual* cone of C, and if C is closed then $(C^*)^* = C$.

Recall that for a set $\mathbf{K} \subset \mathbb{R}^n$, $C_d(\mathbf{K})$ denotes the convex cone of polynomials of degree at most d which are nonnegative on $\mathbf{K}$. We say that a vector $\mathbf{y} \in \mathbb{R}^{s(d)}$ has a representing measure (or is a d-truncated moment sequence) if there exists a finite Borel measure ϕ such that

$$y_\alpha = \int_{\mathbb{R}^n} \mathbf{x}^\alpha\, d\phi, \qquad \forall \alpha \in \mathbb{N}^n_d.$$

Lemma 2: Let $\mathbf{K} \subset \mathbb{R}^n$ be compact. For every $d \in \mathbb{N}$, the dual cone $C_d(\mathbf{K})^*$ is the convex cone

$$\Delta_d := \left\{ \left(\int_{\mathbf{K}} \mathbf{x}^\alpha\, d\phi \right),\ \alpha \in \mathbb{N}^n_d \ :\ \phi \in M(\mathbf{K}) \right\}, \quad (6)$$

i.e., the convex cone of vectors of $\mathbb{R}^{s(d)}$ which have a representing measure with support contained in $\mathbf{K}$.

III. Main result

Solving problem $\mathbf{P}$ described in the introduction, is equivalent to solving the convex optimization problem $\mathcal{P}$:

$$\rho = \inf_{g \in \mathbf{P}[\mathbf{x}]_d} \left\{ \int_{\mathbb{R}^n} \exp(-g)\, d\mathbf{x} \ :\ 1 - g \in C_d(\mathbf{K}) \right\}. \quad (7)$$

Proposition 2: Problem $\mathbf{P}$ has an optimal solution if and only if problem $\mathcal{P}$ in (7) has an optimal solution. Moreover, $\mathcal{P}$ is a finite-dimensional convex optimization problem.

Proof: By Lemma 1, whenever $\mathbf{G}_1$ has finite Lebesgue volume,

$$\mathrm{vol}\,(\mathbf{G}_1) = \frac{1}{\Gamma(1+n/d)} \int_{\mathbb{R}^n} \exp(-g)\, d\mathbf{x}.$$

Moreover, $\mathbf{G}_1$ contains $\mathbf{K}$ if and only if $1 - g \in C_d(\mathbf{K})$, and so $\mathbf{P}$ has an optimal solution $g^* \in \mathbf{P}[\mathbf{x}]_d$ if and only if g^* is an optimal solution of $\mathcal{P}$ (with value $\mathrm{vol}(\mathbf{G}_1)\Gamma(1 + n/2(d))$). Now since $g \mapsto \int_{\mathbb{R}^n} \exp(-g)d\mathbf{x}$ is strictly convex (by Lemma 1) and both $C_d(\mathbf{K})$ and $\mathbf{P}[\mathbf{x}]_d$ are convex cones, problem $\mathcal{P}$ is a finite-dimensional convex optimization problem. ∎

We now can state the first result of this paper: Recall that $M(\mathbf{K})$ is the convex cone of finite Borel measures on $\mathbf{K}$ and let $\mathbb{R}^{++} = \{\lambda \in \mathbb{R} : \lambda > 0\}$.

Theorem 2: Let $\mathbf{K} \subset \mathbb{R}^n$ be compact with nonempty interior and consider the optimization problem $\mathcal{P}$ in (7).

(a) $\mathcal{P}$ has a unique optimal solution $g^* \in \mathbf{P}[\mathbf{x}]_d$.

(b) Let $g^* \in \mathbf{P}[\mathbf{x}]_d$ be the unique optimal solution of $\mathcal{P}$. There exists a finite Borel measure $\mu^* \in M(\mathbf{K})$ such that

$$\int_{\mathbb{R}^n} \mathbf{x}^\alpha \exp(-g^*) d\mathbf{x} = \int_{\mathbf{K}} \mathbf{x}^\alpha \, d\mu^*, \quad \forall |\alpha| = d \quad (8)$$

$$\int_{\mathbf{K}} (1 - g^*) \, d\mu^* = 0 \quad (9)$$

$$\mu^*(\mathbf{K}) = \frac{n}{d} \int_{\mathbb{R}^n} \exp(-g^*) \, d\mathbf{x}. \quad (10)$$

In particular, μ^* is supported on the set $V := \{\mathbf{x} \in \mathbf{K} : g^*(\mathbf{x}) = 1\}$ $(= \mathbf{K} \cap \mathbf{G}_1^*)$ and in fact, μ^* can be substituted with another measure $\nu^* \in M(\mathbf{K})$ supported on at most $\binom{n+d-1}{d}$ contact points of V.

(c) Conversely, if $g^* \in \mathbb{R}[\mathbf{x}]_d$ is homogeneous with $1 - g^* \in C_d(\mathbf{K})$, and there exist points $(\mathbf{x}(i), \lambda_i) \in \mathbf{K} \times \mathbb{R}^{++}$, $i = 1, \ldots, s$, such that $g^*(\mathbf{x}(i)) = 1$ for all $i = 1, \ldots, s$, and

$$\int_{\mathbb{R}^n} \mathbf{x}^\alpha \exp(-g^*) \, d\mathbf{x} = \sum_{i=1}^{s} \lambda_i \, \mathbf{x}(i)^\alpha, \qquad |\alpha| = d,$$

then g^* is the unique optimal solution of problem $\mathcal{P}$.

Proof: (a) As $\mathcal{P}$ is a minimization problem, its feasible set $\{g \in \mathbf{P}[\mathbf{x}]_d : 1 - g \in C_d(\mathbf{K})\}$ can be replaced by the smaller set

$$F = \begin{array}{ll} \{g \in \mathbf{P}[\mathbf{x}]_d & : \displaystyle\int_{\mathbb{R}^n} \exp(-g) d\mathbf{x} \leq \int_{\mathbb{R}^n} \exp(-g_0) d\mathbf{x}, \\ & \quad 1 - g \in C_d(\mathbf{K})\} \end{array}$$

for some $g_0 \in \mathbf{P}[\mathbf{x}]_d$. The set F is a closed convex set since the convex function $g \mapsto \int_{\mathbb{R}^n} \exp(-g) d\mathbf{x}$ is continuous on the interior of its domain.

Next, $\mathrm{int}(C_d(\mathbf{K})^*) \neq \emptyset$ because $\mathbf{K}$ has nonempty interior. So let $\mathbf{z} = (z_\alpha)$, $\alpha \in \mathbb{N}_d^n$, be a (fixed) element of $\mathrm{int}(C_d(\mathbf{K})^*)$ (hence $z_0 > 0$). Then the constraint $1 - g \in C_d(\mathbf{K})$ implies $\langle \mathbf{z}, 1 - g \rangle \geq 0$, i.e., $\langle \mathbf{z}, g \rangle \leq z_0$. On the other hand, being an element of $\mathbf{P}[\mathbf{x}]_d$, g is nonnegative and in particular $g \in C_d(\mathbf{K})$. But then by Corollary I.1.6 in Faraut et Korányi [8, p. 4], the set $\{g \in C_d(\mathbf{K}) : \langle \mathbf{z}, g \rangle \leq z_0\}$ is compact. Therefore, the set F is a compact convex set. Finally, since $g \mapsto \int_{\mathbb{R}^n} \exp(-g(\mathbf{x})) d\mathbf{x}$ is strictly convex, it is continuous on the interior of its domain and so it is continuous on F. Hence problem $\mathcal{P}$ has a unique optimal solution $g^* \in \mathbf{P}[\mathbf{x}]_d$.

(b) We may and will consider any homogeneous polynomial g as an element of $\mathbb{R}[\mathbf{x}]_d$ with $g_\alpha = 0$ whenever $|\alpha| < d$. And so Problem $\mathcal{P}$ is equivalent to the problem

$$\mathcal{P}' : \quad \left\{ \begin{array}{lll} \rho = \displaystyle\inf_{g \in \mathbb{R}[\mathbf{x}]_d} & \displaystyle\int_{\mathbb{R}^n} \exp(-g(\mathbf{x})) \, d\mathbf{x} \\ \text{s.t.} & g_\alpha = 0, \quad \forall \, \alpha \in \mathbb{N}_d^n; \; |\alpha| < d \\ & 1 - g \in C_d(\mathbf{K}), \end{array} \right. \quad (11)$$

where we replaced $g \in \mathbf{P}[\mathbf{x}]_d$ with the equivalent constraints $g \in \mathbb{R}[\mathbf{x}]_d$ and $g_\alpha := 0$ for all $\alpha \in \mathbb{N}_d^n$ with $|\alpha| < d$. Next, doing the change of variable $h = 1 - g$, $\mathcal{P}$' reads:

$$\mathcal{P}' : \quad \left\{ \begin{array}{lll} \rho = \displaystyle\inf_{h \in \mathbb{R}[\mathbf{x}]_d} & \displaystyle\int_{\mathbb{R}^n} \exp(h(\mathbf{x}) - 1) \, d\mathbf{x} \\ \text{s.t.} & h_\alpha = 0, \quad \forall \, \alpha \in \mathbb{N}_d^n; \; 0 < |\alpha| < d \\ & h_0 = 1 \\ & h \in C_d(\mathbf{K}), \end{array} \right. \quad (12)$$

As $\mathbf{K}$ is compact, there exists $\theta \in \mathbf{P}[\mathbf{x}]_d$ such that $1 - \theta \in \mathrm{int}(C_d(\mathbf{K}))$, i.e., Slater's condition holds for the convex optimization problem $\mathcal{P}'$. Indeed, choose $\theta := M^{-1} \|\mathbf{x}\|^d$ for $M > 0$ sufficiently large so that $1 - \theta > 0$ on $\mathbf{K}$. Hence with $\|g\|_1$ denoting the ℓ_1-norm of the coefficient vector of g (in $\mathbb{R}[\mathbf{x}]_d$), there exists $\epsilon > 0$ such that for every $h \in B(\theta, \epsilon) (:= \{h \in \mathbb{R}[\mathbf{x}]_d : \|\theta - h\|_1 < \epsilon\})$, the polynomial $1 - h$ is (strictly) positive on $\mathbf{K}$.

Therefore, the unique optimal solution $(1 - g^*) =: h^* \in \mathbb{R}[\mathbf{x}]_d$ of $\mathcal{P}$' in (12) satisfies the Karush-Kuhn-Tucker (KKT) -optimality conditions, which for (12) read:

$$\int_{\mathbb{R}^n} \mathbf{x}^\alpha \, \exp(h^*(\mathbf{x}) - 1) \, d\mathbf{x} = y_\alpha^*, \quad \forall |\alpha| = d \quad (13)$$

$$\int_{\mathbb{R}^n} \mathbf{x}^\alpha \, \exp(h^*(\mathbf{x}) - 1) \, d\mathbf{x} + \gamma_\alpha = y_\alpha^*, \quad \forall \, |\alpha| < d \quad (14)$$

$$\langle h^*, \mathbf{y}^* \rangle = 0; \quad h_0^* = 1; \; h_\alpha^* = 0, \quad \forall \, 0 < |\alpha| < d \quad (15)$$

for some $\mathbf{y}^* = (y_\alpha^*)$, $\alpha \in \mathbb{N}_d^n$, in the dual cone $C_d(\mathbf{K})^* \subset \mathbb{R}^{s(d)}$ of $C_d(\mathbf{K})$, and some vector $\gamma = (\gamma_\alpha)$, $0 < |\alpha| < d$. By Lemma 2, there exists $\mu^* \in M(\mathbf{K})$ such that

$$y_\alpha^* = \int_{\mathbf{K}} \mathbf{x}^\alpha \, d\mu^*, \qquad \forall \alpha \in \mathbb{N}_d^n,$$

and so (8) is just (13) restated in terms of μ^*.

Next, the condition $\langle h^*, \mathbf{y}^* \rangle = 0$ (or equivalently, $\langle 1 - g^*, \mathbf{y}^* \rangle = 0$), reads: $\int_{\mathbb{R}^n} (1 - g^*) \, d\mu^* = 0$, which combined with $1 - g^* \in C_{2k}(\mathbf{K})$ and $\mu^* \in M(\mathbf{K})$, implies that μ^* is supported on $\mathbf{K} \cap \{\mathbf{x} : g^*(\mathbf{x}) = 1\} = \mathbf{K} \cap \mathbf{G}_1^*$.

Next, let $s := \sum_{|\alpha| = d} g_\alpha^* y_\alpha^* (= y_0^*)$. From $\langle 1 - g^*, \mu^* \rangle = 0$, the measure $s^{-1} \mu^* =: \psi$ is a probability measure supported on $\mathbf{K} \cap \mathbf{G}_1^*$, and satisfies $\int \mathbf{x}^\alpha d\psi = s^{-1} y_\alpha^*$ for all $|\alpha| = d$.

Hence from a result of Mulholland and Rogers [23], there exists an atomic measure $\nu^* \in M(\mathbf{K})$ supported on $\mathbf{K} \cap \mathbf{G}_1^*$ such that

$$\int_{\mathbf{K} \cap \mathbf{G}_1^*} \mathbf{x}^\alpha d\nu^*(\mathbf{x}) = \int_{\mathbf{K} \cap \mathbf{G}_1^*} \mathbf{x}^\alpha d\psi(\mathbf{x}) = s^{-1} \, y_\alpha^*$$

for all α such that $|\alpha| = d$; see e.g. Anastassiou [1, Theorem 2.1.1, p. 39]. As $g^*(\mathbf{x}) = 1$ for all $\mathbf{x}$ in the support of ν^*, we conclude that

$$1 = s^{-1} \sum_{|\alpha| = d} g_\alpha^* \, y_\alpha^* = \int_{\mathbf{K} \cap \mathbf{G}_1^*} g^* d\nu^* = \int_{\mathbf{K} \cap \mathbf{G}_1^*} d\nu^*,$$

i.e., ν^* is a probability measure and therefore may be chosen to be supported on at most $\binom{n+d-1}{d}$ points in $\mathbf{K} \cap \mathbf{G}_1^*$; see [1, Theorem 2.1.1]. Hence in (8) the measure μ^* can be substituted with the atomic measure $s \, \nu^*$ supported on at most $\binom{n+d-1}{d}$ contact points in $\mathbf{K} \cap \mathbf{G}_1^*$.

To obtain $\mu^*(\mathbf{K}) = \frac{n}{d}\int_{\mathbb{R}^n}\exp(-g^*)$, multiply both sides of (13)-(14) by h^*_α for every $\alpha \neq 0$, sum up and use $\langle h^*, \mathbf{y}^*\rangle = 0$ to obtain

$$\begin{aligned} -y^*_0 &= \sum_{\alpha\neq 0} h^*_\alpha\, y^*_\alpha \,=\, \int_{\mathbb{R}^n}(h^*-1)\exp(h^*-1)\,d\mathbf{x} \\ &= -\int_{\mathbb{R}^n} g^*\exp(-g^*)\,d\mathbf{x} \,=\, -\frac{n}{d}\int \exp(-g^*)\,d\mathbf{x}, \end{aligned}$$

where we have also used (5).

(c) Let $\mu^* := \sum_{i=1}^s \lambda_i \delta_{\mathbf{x}(i)}$ where $\delta_{\mathbf{x}(i)}$ is the Dirac measure at the point $\mathbf{x}(i) \in \mathbf{K}$, $i = 1,\ldots,s$. Next, let $y^*_\alpha := \int \mathbf{x}^\alpha d\mu^*$ for all $\alpha \in \mathbb{N}^n_d$, so that $\mathbf{y}^* \in C_d(\mathbf{K})^*$. In particular $\mathbf{y}^*$ and g^* satisfy

$$\langle 1-g^*, \mathbf{y}^*\rangle \,=\, \int_{\mathbf{K}} (1-g^*)d\mu^* \,=\, 0,$$

because $g^*(\mathbf{x}(i)) = 1$ for all $i = 1,\ldots,s$. In other words, the pair $(g^*, \mathbf{y}^*)$ satisfies the Karush-Kuhn-Tucker (KKT) optimality conditions associated with the convex problem $\mathcal{P}$. But since Slater's condition holds for $\mathcal{P}$, those conditions are also sufficient for g^* to be an optimal solution of $\mathcal{P}$, the desired result. ■

Notice that neither $\mathbf{K}$ nor $\mathbf{G}^*_1$ are required to be convex.

IV. A NUMERICAL SCHEME

The finite-dimensional convex optimization problem $\mathcal{P}$ in (7) is hard to solve for mainly two reasons:

- From Theorem 1, the gradient and Hessian of the (strictly) convex objective function $g \mapsto \int \exp(-g)$ requires evaluating integrals of the form

$$\int_{\mathbb{R}^n} \mathbf{x}^\alpha \exp(-g(\mathbf{x}))\,d\mathbf{x},$$

  for every $\alpha \in \mathbb{N}^n_d$, a difficult and challenging problem.
- The convex cone $C_d(\mathbf{K})$ has no *exact* and *tractable* representation to handle the constraint $1-g \in C_d(\mathbf{K})$ in problem (7).

However, below we outline a numerical scheme to approximate to any desired ϵ-accuracy (with $\epsilon > 0$):

- the optimal value ρ of (7),

- the optimal solution $g^* \in \mathbf{P}[\mathbf{x}]_d$ of $\mathcal{P}$ obtained in Theorem 2.

A. Concerning gradient and Hessian evaluation

To approximate the gradient and Hessian of the objective function we will use the following result:

Lemma 3: Let $g \in \mathbf{P}[\mathbf{x}]_d$ and let $\mathbf{G}_1 = \{\mathbf{x} : g(\mathbf{x}) \leq 1\}$. Then for every $\alpha \in \mathbb{N}^n$

$$\int_{\mathbb{R}^n} \mathbf{x}^\alpha \exp(-g)\,d\mathbf{x} \,=\, \Gamma(1+\frac{n+|\alpha|}{d})\int_{\mathbf{G}_1} \mathbf{x}^\alpha\,d\mathbf{x}. \quad (16)$$

The proof being identical to that of Lemma 1 is omitted. In Henrion et al. [16] we have provided a hierarchy of semidefinite programs to approximate as closely as desired any finite moment sequence (z_α) defined by:

$$z_\alpha \,=\, \int_\Omega \mathbf{x}^\alpha\,d\mathbf{x}, \quad \alpha \in \mathbb{N}^n_\ell.$$

where ℓ is fixed, arbitrary, and Ω is a compact basic semi-algebraic set of the form $\{\mathbf{x} : \theta_j(\mathbf{x}) \geq 0,\ j = 1,\ldots,m\}$ for some polynomials $(\theta_j) \subset \mathbb{R}[\mathbf{x}]$.

Hence given a current iterate $g \in \mathbf{P}[\mathbf{x}]_d$, by using the methodology described in [16], [17] one may approximate as closely as desired value at g of the objective function of problem (7), as well as its gradient and Hessian.

B. Concerning the convex cone $C_d(\mathbf{K})$

We here assume that the compact set $\mathbf{K} \subset \mathbb{R}^n$ is the basic semi-algebraic set

$$\mathbf{K} \,=\, \{\mathbf{x} \in \mathbb{R}^n_+ \,:\, w_j(\mathbf{x}) \geq 0,\ j = 1,\ldots,s\}, \qquad (17)$$

for some given polynomials $(w_j) \subset \mathbb{R}[\mathbf{x}]$. Denote by $\Sigma_k \subset \mathbb{R}[\mathbf{x}]_{2k}$ the convex cone of SOS (sum of squares) polynomials of degree at most $2k$, and let w_0 be the constant polynomial equal to 1, and $v_j := \lceil \deg(w_j)/2\rceil$, $j = 0,\ldots,s$.

With k fixed, arbitrary, we now replace the condition $1-g \in C_d(\mathbf{K})$ with the stronger condition $1-g \in \mathcal{C}_k (\subset C_d(\mathbf{K}))$ where

$$\mathcal{C}_k \,=\, \left\{\sum_{j=0}^s \sigma_j\,w_j \,:\, \sigma_j \in \Sigma_{k-v_j},\ j = 0,1,\ldots,s\right\}. \quad (18)$$

It turns out that membership in $\mathcal{C}_k$ translates into LMIs on the coefficients of the polynomials g and σ_j's; see e.g. [17]. Note that if $\mathbf{K}$ has nonempty interior then $\mathcal{C}_k$ is closed.

Assumption 1 (Archimedean assumption): There exist $M > 0$ and $k \in \mathbb{N}$ such that the quadratic polynomial $\mathbf{x} \mapsto M - \|\mathbf{x}\|^2$ belongs to $\mathcal{C}_k$.

Under Assumption 1 $C_d(\mathbf{K}) = \overline{\bigcup_{k\geq 0} \mathcal{C}_k}$, that is, the family of convex cones $(\mathcal{C}_k)$, $k \in \mathbb{N}$, provide a sequence of (nested) *inner approximations* of $C_d(\mathbf{K})$.

C. A numerical scheme

In view of the above it is natural to consider the following hierarchy of convex optimization problems $(\mathcal{P}_k)$, $k \in \mathbb{N}$:

$$\begin{array}{rl} \rho_k = \displaystyle\min_{g,\sigma_j} & \displaystyle\int_{\mathbb{R}^n} \exp(-g)\,d\mathbf{x} \\ \text{s.t.} & 1-g = \displaystyle\sum_{j=0}^s \sigma_j w_j \\ & g_\alpha = 0, \quad \forall |\alpha| < d \\ & \sigma_j \in \Sigma_{k-v_j},\ j = 0,\ldots,s. \end{array} \qquad (19)$$

Of course the sequence (ρ_k), $k \in \mathbb{N}$, is monotone non increasing and $\rho_k \geq \rho$ for all k. For each fixed $k \in \mathbb{N}$, $\mathcal{P}_k$ is a convex optimization problem which consists of minimizing a strictly convex function under LMI constraints.

From Corollary 1, $\int_{\mathbb{R}^n}\exp(-g)d\mathbf{x} < \infty$ if and only if $g \in \mathbf{P}[\mathbf{x}]_d$ and so the objective function also acts as a barrier for the convex cone $\mathbf{P}[\mathbf{x}]_d$. Therefore, to solve $\mathcal{P}_k$ one may use first-order or second-order (local minimization) algorithms, starting from an initial $g_0 \in \mathbf{P}[\mathbf{x}]_d$. At each step of such an algorithm one may use the methodology described in [16] to approximate the gradient and Hessian of the objective function.

Theorem 3: Let $\mathbf{K}$ in (17) be compact with nonempty interior and let Assumption 1 hold. Then there exists k_0 such that for every $k \geq k_0$, problem $\mathcal{P}_k$ in (19) has a unique optimal solution $g_k^* \in \mathbf{P}[\mathbf{x}]_d$.

Proof: Firstly $\mathcal{P}_k$ has a feasible solution for sufficiently large k. Let us consider the polynomial $\mathbf{x} \mapsto g_0(\mathbf{x}) = \sum_{i=1}^n x_i^d$ which belongs to $\mathbf{P}[\mathbf{x}]_d$. Then as $\mathbf{K}$ is compact, $M - g_0 > 0$ on $\mathbf{K}$ for some M and so $1 - g_0/M \in \mathcal{C}_k$ for some k_0 (and hence for all $k \geq k_0$). Hence g_0/M is a feasible solution for $\mathcal{P}_k$ for all $k \geq k_0$. Of course as $\mathcal{C}_k \subset C_d(\mathbf{K})$ every feasible solution g satisfies $0 \leq g \leq 1$ on $\mathbf{K}$. So proceeding as in the proof of Theorem 2 and using the fact that $\mathcal{C}_k$ is closed, the set

$$\{g \in \mathbf{P}[\mathbf{x}]_d \cap \mathcal{C}_k : \int_{\mathbb{R}^n} \exp(-g) d\mathbf{x} \leq \int_{\mathbb{R}^n} \exp(-\frac{g_0}{M}) d\mathbf{x}\},$$

is compact. And as the objective function is strictly convex, the optimal solution $g_k^* \in \mathbf{P}[\mathbf{x}]_d \cap \mathcal{C}_k$ is unique (but the representation of $1 - g_k^*$ in (19) is not unique in general). ■

We now consider the asymptotic behavior of the solution of (19) as $k \to \infty$.

Theorem 4: Let $\mathbf{K}$ in (17) be compact with nonempty interior and let Assumption 1 hold. If ρ (resp. ρ_k) is the optimal value of $\mathcal{P}$ (resp. $\mathcal{P}_k$) then $\rho = \lim_{k\to\infty} \rho_k$. Moreover, for every $k \geq k_0$, let $g_k^* \in \mathbf{P}[\mathbf{x}]_d$ be the unique optimal solution of $\mathcal{P}_k$. Then as $k \to \infty$, $g_k^* \to g^*$ where g^* is the unique optimal solution of $\mathcal{P}$.

Proof: By Theorem 2, $\mathcal{P}$ has a unique optimal solution $g^* \in \mathbf{P}[\mathbf{x}]_d$. Let $\epsilon > 0$ be fixed, arbitrary. As $1 - g^* \in C_d(\mathbf{K})$, the polynomial $1 - g/(1+\epsilon)$ is strictly positive on $\mathbf{K}$, and so by Putinar's theorem [22], $1 - g^*/(1+\epsilon)$ belongs to $\mathcal{C}_k$ for all $k \geq k_\epsilon$ for some integer k_ϵ. Hence the polynomial $g^*/(1+\epsilon) \in \mathbf{P}[\mathbf{x}]_d$ is a feasible solution of $\mathcal{P}_k$ for all $k \geq k_\epsilon$. Moreover, by homogeneity,

$$\begin{aligned}\int_{\mathbb{R}^n} \exp(-\frac{g^*}{1+\epsilon})\, d\mathbf{x} &= (1+\epsilon)^{n/d} \int_{\mathbb{R}^n} \exp(-g^*)\, d\mathbf{x} \\ &= (1+\epsilon)^{n/d} \rho.\end{aligned}$$

This shows that $\rho_k \leq (1+\epsilon)^{n/d}\rho$ for all $k \geq k_\epsilon$. Combining this with $\rho_k \geq \rho$ and the fact that $\epsilon > 0$ was arbitrary, yields the desired result.

Next, let $\mathbf{y} \in \mathrm{int}(C_d(\mathbf{K})^*)$ be as in the proof of Theorem 2. From $1 - g_k^* \in \mathcal{C}_k$ we also obtain $\langle \mathbf{y}, 1 - g_k^* \rangle \geq 0$, i.e.,

$$y_0 \geq \langle \mathbf{y}, g_k^* \rangle, \quad \forall k \geq k_0,$$

As the set $\{g \in C_d(\mathbf{K}) : \langle \mathbf{y}, g \rangle \leq y_0\}$ is compact, there exists a subsequence (k_ℓ), $\ell \in \mathbb{N}$, and $\tilde{g} \in C_d(\mathbf{K})$ such that $g_{k_\ell}^* \to \tilde{g}$ as $\ell \to \infty$. Moreover, again by Fatou's lemma,

$$\rho \,=\, \lim_{\ell\to\infty} \rho_{k_\ell} \,=\, \lim_{\ell\to\infty} \int_{\mathbb{R}^n} \exp(-g_{k_\ell}^*(\mathbf{x})) d\mathbf{x} \,\geq$$

$$\int_{\mathbb{R}^n} \liminf_{\ell\to\infty} \exp(-g_{k_\ell}^*(\mathbf{x})) d\mathbf{x} \,=\, \int_{\mathbb{R}^n} \exp(-\tilde{g}(\mathbf{x})) d\mathbf{x} \,=\, \rho,$$

which proves that $\tilde{g}$ is an optimal solution of $\mathcal{P}$, and by uniqueness $\tilde{g} = g^*$. As the converging subsequence (g_{k_ℓ}) was arbitrary, the whole sequence (g_k^*) converges to g^*. ■

References

[1] G.A. Anastassiou. *Moments in Probability and Approximation Theory*, Longman Scientific & Technical, UK, 1993.

[2] K. Ball. Ellipsoids of maximal volume in convex bodies, Geom. Dedicata **41** (1992), pp. 241–250.

[3] J. Bastero and M. Romance. John's decomposition of the identity in the non-convex case, Positivity **6** (2002), pp. 1–16.

[4] F.L. Bookstein. Fitting conic sections to scattered data, *Comp. Graph. Image. Process.* **9** (1979), pp. 56–71.

[5] G. Calafiore. Approximation of n-dimnsional data using spherical and ellipsoidal primitives, *IEEE Trans. Syst. Man. Cyb.* **32** (2002), pp. 269–2768.

[6] F.L. Chernousko. Guaranteed estimates of undetermined quantities by means of ellipsoids, *Sov. Math. Dodkl.* **21** (1980), pp. 396–399.

[7] C. Croux, G. Haesbroeck and P.J. Rousseeuw. Location adjustment for the minimum volume ellipsoid estimator, *Stat. Comput.* **12** (2002), pp. 191–200.

[8] J. Faraut and A. Korányi. *Analysis on Symmetric Cones*, Clarendon Press, Oxford, 1994.

[9] W. Gander, G.H. Golub and R. Strebel. Least-squares fitting of circles and ellipses, *BIT* **34** (1994), pp. 558–578.

[10] A. Giannopoulos, I. Perissinaki, A. Tsolomitis. A. John's theorem for an arbitrary pair of convex bodies, Geom. Dedicata **84** (2001), pp. 6379.

[11] M. Henk. Löwner-John ellipsoids, Documenta Math., to appear.

[12] D. Henrion, D. Peaucelle, D. Arzelier and M.Sebek. Ellipsoidal approximation of the stability domain of a polynomial, *IEEE Trans. Aut. Control* **48** (2003), pp. 2255–2259.

[13] D. Henrion and J.B. Lasserre. Inner approximations for polynomial matrix inequalities and robust stability regions, *IEEE Trans. Aut. Control* **57** (2012), pp. 1456–1467.

[14] D. Henrion, M. Sebek and V. Kucera. Positive polynomials and and robust stabilization with fixed-order controllers, *IEEE Trans. Aut. Control* **48** (2003), pp. 1178–1186.

[15] A. Karimi, H. Khatibi and R. Longchamp. Robust control of polytopic systems by convex optimization, *Automatica* **43** (2007), pp. 1395–1402.

[16] D. Henrion, J.B. Lasserre, C. Savorgnan. Approximate volume and integration of basic semi-algebraic sets, SIAM Review **51** (2009), pp. 722–743.

[17] J.B. Lasserre. *Moments, Positive Polynomials and Their Applications*, Imperial College, London, 2009.

[18] A. Morosov and S. Shakirov. New and old results in resultant theory, `arXiv:0911.5278v1`, 2009.

[19] A. Morosov and S. Shakirov. Introduction to integral discriminants, J. High Energy Phys. **12** (2009), `arXiv:0911.5278v1`, 2009.

[20] U. Nurges. Robust pole assignment via reflection coefficientsof polynomials, *Automatica* **42** (2006), pp. 1223–1230.

[21] V. Pratt. Direct least squares fittingof algebraic surfaces, *ACM J. Comp. Graph.* **21** (1987).

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

[23] H.P. Mulholland and C.A. Rogers, Representation theorems for distribution functions, Proc. London Math. Soc. **8** (1958), pp. 177–223.

[24] J.B. Rosen. Pattern separation by convex programming techniques, *J. Math. Anal. Appl.* **10** (1965), pp. 123–1324.

[25] P.L. Rosin. A note on the least squares fitting of ellipses, *Pattern Recog. Letters* **14** (1993), pp. 799–808.

[26] P.L. Rosin and G.A. West. Nonparametric segmentation of curves into various representations, *IEEE Trans. Pattern Anal. Machine Intell.* **17** (1995), pp. 1140–1153.

[27] P.J. Rousseeuw and A.M. Leroy. *Robust Regression and Outlier Detection*, John Wiley, New York (1987).

[28] P. Sun and R. Freund. Computation of minimum-volume covering ellipsoids, *Oper. Res.* **52** (2004), pp. 690–706.

[29] G. Taubin. Estimation of planar curves, surfaces and nonplanar space curves defined by implicit equations, with applications to to edge and range image segmentation, *IEEE Trans. Pattern Anal. Machine Intell.* **13** (1991), pp. 1115–1138.