LEVEL SETS AND NON GAUSSIAN INTEGRALS OF POSITIVELY HOMOGENEOUS FUNCTIONS

J.B. LASSERRE

ABSTRACT. We investigate various properties of the sublevel set $\{\mathbf{x} : g(\mathbf{x}) \leq 1\}$ and the integration of h on this sublevel set when g and h are positively homogeneous functions. For instance, the latter integral reduces to integrating $h \exp(-g)$ on the whole space $\mathbb{R}^n$ (a non Gaussian integral) and when g is a polynomial, then the volume of the sublevel set is a convex function of the coefficients of g. In fact, whenever h is nonnegative, the functional $\int \phi(g(\mathbf{x}))h(\mathbf{x})d\mathbf{x}$ is a convex function of g for a large class of functions $\phi : \mathbb{R}_+ \to \mathbb{R}$. We also provide a numerical approximation scheme to compute the volume or integrate h (or, equivalently to approximate the associated non Gaussian integral). We also show that finding the sublevel set $\{\mathbf{x} : g(\mathbf{x}) \leq 1\}$ of minimum volume that contains some given subset $\mathbf{K}$ is a (hard) convex optimization problem for which we also propose two convergent numerical schemes. Finally, we provide a Gaussian-like property of non Gaussian integrals for homogeneous polynomials that are sums of squares and critical points of a specific function.

1. INTRODUCTION

Positively homogeneous functions (PHF) form a wide class of functions that one encounters in many applications. As a consequence of homogeneity, they enjoy very particular properties, and among them the celebrated and very useful Euler's identity which allows to deduce additional properties of PHFs in various contexts. One goal of this paper is to show that another not well-known property of PHFs yields surprising and unexpected results, some of them already known in particular cases. Namely, we are concerned with PHFs, their sublevel sets and in particular, the integral

$$(1.1) \qquad y \mapsto I_{g,h}(y) := \int_{\{\mathbf{x} \,:\, g(\mathbf{x}) \leq y\}} h(\mathbf{x}) \, d\mathbf{x},$$

as a function $I_{g,h} : \mathbb{R}_+ \to \mathbb{R}$ when g, h are PHFs. With y fixed, we are also interested in $I_{g,h}(y)$ now as a function of g, especially when g is a nonnegative homogeneous polynomial. Nonnegative homogeneous polynomials are particularly interesting as they can be used to approximate norms; see e.g. Barvinok [4, Theorem 3.4]. Interestingly, the integral (1.1) is

---

related in a simple and remarkable manner to the non-Gaussian integral $\int_{\mathbb{R}^n} h \exp(-g)d\mathbf{x}$ and therefore, any information on either one can help understanding and evaluating the other.

Functional integrals appear frequently in quantum Physics. In the important particular case when $h = 1$, Morosov et Shakirov [17] proved that for all *forms g* of degree k,

$$(1.2) \qquad I_{g,1}(1) = \text{cte}(k) \cdot \int_{\mathbb{R}^n} \exp(-g)d\mathbf{x},$$

where the constant depends only on k. In fact, a formula of exactly the same flavor was already known for convex sets, and was the initial motivation and starting point of this paper. Namely, if $C \subset \mathbb{R}^n$ is convex, its support function σ_C is a PHF of degree 1, and the polar $C^\circ \subset \mathbb{R}^n$ of C is the convex set $\{\mathbf{x} : \sigma_C(\mathbf{x}) \leq 1\}$. Then

$$\text{vol}\,(C^\circ) = \frac{1}{n!} \int_{\mathbb{R}^n} \exp(-\sigma_C(\mathbf{x}))\,d\mathbf{x}, \qquad \forall C.$$

(See e.g. Hiriart-Urruty [3, Exercise 65, p. 243-244] and Barvinok [4, Problem 4, p. 207].) In fact, as we will see, the intriguing result (1.2) is a particular case of a more general and striking result about PHFs.

The connection between I_{g1} and $\int \exp(-g)$ is a consequence of what the authors in [17] call *action-independence* of integral discriminants. That is, for a very general class of univariate functions ϕ, and for every form g of degree d,

$$(1.3) \qquad \int \phi(g(\mathbf{x}))d\mathbf{x} = C(\phi, d) \cdot \theta(g),$$

for some function θ, where the constant $C(\phi, d)$ depends only on ϕ and d; see [17, (73)]. And so for instance, with the choices $t \mapsto \phi_1(t) = 1_{[0,1]}(t)$ and $\phi_2(t) = \exp(-t)$, and with $\mathbb{R}^n$ as integration contour,

$$\int_{\{\mathbf{x}\,:\,g(\mathbf{x})\leq 1\}} d\mathbf{x} = \frac{C(\phi_1, d)}{C(\phi_2, d)} \int_{\mathbb{R}^n} \exp(-g)\,d\mathbf{x}, \qquad \forall g,$$

whenever any of the above integrals finite. In Morosov and Shakirov [17] the motivation of the authors was not to connect I_{g1} with $\int \exp(-g)$ but rather to study a specific class of non-Gaussian integrals, the so-called *integral discriminants*. In fact, the theory of integral discriminants which lies in the field of non linear algebra, extends to integrals on more general contours of integration, and Ward identities permits to study properties of such integrals that are invariant under a change of contour of integration. Hence the goal in [17] is to provide exact formulas for integral discriminants in terms of algebraic invariants of the form g involved in the exponent; several highly non trivial examples are provided in [18, 22] and some of their results are further analyzed in Fuji [7] and Stoyanovsky [23].

Contribution. In the present paper we do not study non Gaussian integrals from the point of view of algebraic invariants as in e.g. [17, 18, 22]. We are interested in properties of the integral (1.1) (and variants as well) and its associated non-Gaussian integral $\int_{\mathbb{R}^n} h \exp(-g)d\mathbf{x}$ for PHFs that are not necessarily polynomials. Among other things, we are also interested in methods for their evaluation (or approximation). More precisely:

(a) We first extend the action-independence property (1.3) to PHFs of degree $0 \neq d \in \mathbb{R}$. Namely, given a measurable mapping $\phi : \mathbb{R}_+ \to \mathbb{R}$, and g, h PHFs of respective degree $0 \neq d, p \in \mathbb{R}$ and such that $\int |h| \exp(-g)d\mathbf{x}$ is finite,

$$(1.4) \qquad \int_{\mathbb{R}^n} \phi(g(\mathbf{x}))h(\mathbf{x}) \, d\mathbf{x} = C(\phi, d, p) \cdot \int_{\mathbb{R}^n} h \exp(-g) \, d\mathbf{x},$$

where the constant $C(\phi, d, p)$ depends only on ϕ, d, p. In particular, if the sublevel set $\{\mathbf{x} : g(\mathbf{x}) \leq 1\}$ is bounded, then for every nonnegative $y \in \mathbb{R}$,

$$(1.5) \qquad \int_{\{\mathbf{x} : g(\mathbf{x}) \leq y\}} h \, d\mathbf{x} = \frac{y^{(n+p)/d}}{\Gamma(1 + (n+p)/d)} \int_{\mathbb{R}^n} h \exp(-g) \, d\mathbf{x},$$

with Γ being the standard Gamma function[1]. And so the (Lebesgue) volume of the sublevel set $\{\mathbf{x} : g(\mathbf{x}) \leq y\}$ is given by:

$$(1.6) \qquad \mathrm{vol}\left(\{\mathbf{x} : g(\mathbf{x}) \leq y\}\right) = \frac{y^{n/d}}{\Gamma(1 + n/d)} \int_{\mathbb{R}^n} \exp(-g) \, d\mathbf{x}.$$

As already mentioned, when $h(\mathbf{x}) = 1$ for all $\mathbf{x}$, and g is a positive definite form, (1.6) was already proved in Morosov et Shakirov [17, (75) p. 39]. In fact, the same arguments as in [17] can be adapted to PHFs g and with $h \neq 1$. The same technique applies for the integral

$$I_{g_1, \ldots, g_m, h} := \int_{\{\mathbf{x} : g_k(\mathbf{x}) \leq 1, \, k=1, \ldots, m\}} h d\mathbf{x},$$

for a finite family $(g_k), h$ of PHFs with same degree $0 < d \in \mathbb{R}$.

In addition, we obtain the identity

$$\frac{\displaystyle\int_{\{\mathbf{x} : g(\mathbf{x}) \leq y\}} \exp(-g) \, h \, d\mathbf{x}}{\displaystyle\int_{\mathbb{R}^n} \exp(-g) \, h \, d\mathbf{x}} = \frac{\displaystyle\int_0^y \exp(-z) z^{(n+p)/d-1} \, dz}{\Gamma((n+p)/d)},$$

which expresses how fast $\mu(\{\mathbf{x} : g(\mathbf{x}) \leq y\})$ converges to $\mu(\mathbb{R}^n)(= \int \exp(-g)hd\mathbf{x})$ when $y \to \infty$ and μ has the non Gaussian density $h \exp(-g)$ with respect to the Lebesgue measure. It is the same speed with which the truncated (Gamma) integral $\int_0^y \exp(-z)z^{(n+p)/d-1} \, dz$ converges to $\Gamma((n+p)/d)$.

(b) We also provide **an alternative and simple proof based on Laplace transform** (in the spirit of Lasserre and Zeron [13] to provide

---

[1]The Gamma function $\Gamma : \mathbb{R} \to \mathbb{R}$ is defined by $a \mapsto \Gamma(a) := \int_0^\infty t^{a-1} \exp(-t)dt$, for every $a > -1$.

explicit expressions for certain integrals on specific domains like e.g., a simplex or an ellipsoid). This proof permits to interpret (1.2) as duality result in the usual $(\cdot, +)$-algebra which parallels duality in convex optimization, i.e., in the $(\max, +)$-algebra. This complements the list of the many remarkable properties of homogeneous functions; for instance in [12] for integration, in [9] for optimization, and in [14] for several properties of some *conjugates* of homogeneous functions.

(c) **Convexity:** (1.5) also helps analyze properties of the mapping f_h : $C_d \to \mathbb{R}$,

$$g \mapsto f_h(g) := \int_{\{\mathbf{x}\,:\,g(\mathbf{x}) \le 1\}} h(\mathbf{x})\, d\mathbf{x}, \qquad \forall\, g \in C_d,$$

where C_d is the space of nonnegative PHFs of degree $0 \ne d \in \mathbb{R}$ with bounded sublevel set $\{\mathbf{x} : g(\mathbf{x}) \le 1\}$, and h is a given PHF. Whenever h is nonnegative, f_h is strictly convex and when g is a homogeneous polynomial, we provide an explicit expression of its gradient and Hessian on the interior of its domain. This is the analogue in our context of a formula already provided in Müller et al. [19] when the integration domain is a polytope.

This convexity property is important. For instance, given $\mathbf{K} \subset \mathbb{R}^n$, it shows that the problem of computing a homogeneous polynomial g such that $\{\mathbf{x} : g(\mathbf{x}) \le 1\}$ contains $\mathbf{K}$ and has minimum volume, is a finite-dimensional convex optimization problem (but hard to solve even though it is convex)! This problem has received a lot of attention with an elegant solution in the case $d = 2$ (that is, finding the *ellipsoid* of minimum volume). Notice that here with $d > 2$ the problem is still convex even though the sublevel set $\{\mathbf{x} : g(\mathbf{x}) \le 1\}$ is *not* required to be convex. We present two numerical approximation schemes based on semidefinite programming, that provide sequences of upper and lower bounds that converge to minimum volume.

(d) **Polarity:** When g is a convex PHF of degree $d \in \mathbb{R}$ with compact sublevel set $G = \{\mathbf{x} : g(\mathbf{x}) \le 1\}$, the polar G° is a certain sublevel set of the Legendre-Fenchel conjugate g^* of g, which is itself a PHF of degree q with $1/d + 1/q = 1$ and therefore we also obtain the volume of G° in the form of non Gaussian integral, the *dual* analogue of the one obtained for G.

(e) **Numerical approximation:** Evaluating a non Gaussian integral $\int \exp(-g)$ is a very hard problem even if in some cases exact results can be obtained from some algebraic invariants of g, as in e.g. [17, 18, 22]. However, it turns out that (1.5) is indeed very useful to approximate non Gaussian integrals from polynomials. Indeed, when h is a polynomial and g is a homogeneous polynomial, we show how to approximate the non Gaussian integral $\int h \exp(-g) d\mathbf{x}$, as closely as desired, by solving a hierarchy of so-called semidefinite programs. This numerical approximation scheme complements the exact results mentioned above.

(f) **A Gaussian-like property:** Finally we also prove a (again intriguing) Gaussian-like property of homogeneous polynomials that are sums of squares (SOS). Recall that every homogeneous and SOS polynomial g of

degree $2d$ can be written as $\tilde{\mathbf{v}}_d(\mathbf{x})^T \mathbf{\Sigma} \tilde{\mathbf{v}}_d(\mathbf{x})$ for some positive semidefinite matrix $\mathbf{\Sigma}$ (not unique in general) and where $\tilde{\mathbf{v}}_d(\mathbf{x})$ is the vector of all monomials of degree d. If we now consider the function

$$\mathbf{\Sigma} \mapsto \qquad \theta(\mathbf{\Sigma}) = (\det \mathbf{\Sigma})^k \int \exp(-k\tilde{\mathbf{v}}_d(\mathbf{x})^T \mathbf{\Sigma} \tilde{\mathbf{v}}_d(\mathbf{x})),$$

where $k = n/(2d\ell(d))$ (with $\ell(d) = \binom{n+d-1}{d}$), then its critical points are the positive semidefinite matrices $\mathbf{\Sigma}$ such that $\mathbf{\Sigma}^{-1}$ is the matrix of moments of order $2d$ associated with the non Gaussian density $\exp(-k\tilde{\mathbf{v}}_d(\mathbf{x})^T \mathbf{\Sigma} \tilde{\mathbf{v}}_d(\mathbf{x}))$ (normalized to be a probability density) ... exactly like in the Gaussian case (i.e. when $d = 1$) where $\mathbf{\Sigma}^{-1}$ is a *covariance* matrix! (In the latter case, the function θ is constant, hence all points $\mathbf{\Sigma}$ are critical).

2. SUBLEVEL SETS AND NON GAUSSIAN INTEGRALS

2.1. **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 $\mathbb{N}_d^n := \{\alpha \in \mathbb{N}^n : |\alpha| \, (= \sum_i \alpha_i) = d\}$, 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 $\mathbb{R}[\mathbf{x}]_d$. Denote by $\mathcal{S}_k$ the space of $k \times k$ real symmetric matrices with scalar product $\langle \mathbf{B}, \mathbf{C} \rangle =$ trace $(\mathbf{BC})$; also, the notation $\mathbf{B} \succeq 0$ (resp. $\mathbf{B} \succ 0$) stands for $\mathbf{B}$ is positive semidefinite (resp. positive definite).

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 real-valued continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is homogeneous (resp. *positively* homogeneous) of degree k ($k \in \mathbb{R}$) if $f(\lambda \mathbf{x}) = \lambda^k f(\mathbf{x})$ for all $\lambda \neq 0$ (resp. for all $\lambda > 0$) and all $\mathbf{x} \in \mathbb{R}$. For instance if $\mathbf{x} \mapsto f(\mathbf{x})$ is homogeneous of degree k, then $|f|$ is positively homogeneous of degree k (but not homogeneous). Let us denote by:

- C_d the convex cone of PHFs g of degree $d \in \mathbb{R}$ with bounded sublevel set $G := \{\mathbf{x} : g(\mathbf{x}) \leq 1\}$. Notice that necessarily $d \neq 0$ and any $g \in C_d$ is nonnegative. Indeed, if $d = 0$ then $\mathbf{x} \in G$ implies $\lambda \mathbf{x} \in G$ for every $\lambda > 0$, in contradiction with boundedness of G. Similarly, assume that $g(\mathbf{x}_0) < 0$ for some some $\mathbf{x}_0$ (hence $\mathbf{x}_0 \in G$). As $g(\lambda \mathbf{x}_0) = \lambda^d g(\mathbf{x}_0) < 0$ for every $\lambda > 0$, $\lambda \mathbf{x}_0 \in G$ for every $\lambda > 0$, again in contradiction with boundedness of G.
- $\mathbf{P}[\mathbf{x}]_d \subset \mathbb{R}[\mathbf{x}]_d$, $d \in \mathbb{N}$, the convex cone of homogeneous polynomials of degree d with compact sublevel set $\{\mathbf{x} : g(\mathbf{x}) \leq 1\}$. Hence, $\mathbf{P}[\mathbf{x}]_d \subset C_d$ whenever $d \in \mathbb{N}$.

Remember that a convex function is *proper* if $f(\mathbf{x}) > -\infty$ for all $\mathbf{x}$ and $f(\mathbf{x}) < +\infty$ for some $\mathbf{x}$. A proper convex function is closed if it is lower semicontinuous; see Rockafellar [21].

In the sequel we will need the following result which extends to PHFs of degree $d \in \mathbb{R}$, one already used in [17] for forms of degree $d \in \mathbb{N}$ (and also used in Hiriart-Urruty [3] for the volume of the polar of a convex set).

Lemma 2.1. *Let $d, p \in \mathbb{R}$ with $d \neq 0$, and let $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ be a measurable function such that*

$$\int_0^\infty t^{n+p-1}\phi(t^d)\, dt < \infty,$$

and let g, h be nonnegative PHFs of degree $0 \neq d \in \mathbb{R}$ and $p \in \mathbb{R}$ respectively. The functional $\int_{\mathbb{R}^n} \phi(g(\mathbf{x}))h(\mathbf{x})\, d\mathbf{x}$ is finite if and only if $\int_{\mathbb{R}^n} h\exp(-g)d\mathbf{x} < +\infty$, in which case:

$$(2.1) \qquad \int_{\mathbb{R}^n} \phi(g(\mathbf{x}))h(\mathbf{x})\, d\mathbf{x} \;=\; \underbrace{\frac{d \displaystyle\int_0^\infty t^{n+p-1}\phi(t^d)\, dt}{\Gamma((n+p)/d)}}_{\mathrm{cte}(n,p,d,\phi)} \cdot \int_{\mathbb{R}^n} h\,\exp(-g)\, d\mathbf{x}.$$

Proof. With $\mathbf{z} = (z_1, \ldots, z_{n-1})$, do the change of variable $x_1 = t$, $x_2 = tz_1, \ldots, x_n = tz_{n-1})$ so that one may decompose $\int_{\mathbb{R}^n} \phi(g(\mathbf{x}))h(\mathbf{x})d\mathbf{x}$ into the sum

$$\int_{\mathbb{R}_+ \times \mathbb{R}^{n-1}} t^{n+p-1}\phi(t^d g(1,\mathbf{z}))h(1,\mathbf{z})\, dt\, d\mathbf{z}$$

$$+ \int_{\mathbb{R}_+ \times \mathbb{R}^{n-1}} t^{n+p-1}\phi(t^d g(-1,-\mathbf{z}))h(-1,-\mathbf{z})\, dt\, d\mathbf{z},$$

$$= \int_{\mathbb{R}^{n-1}} \left(\int_0^\infty t^{n+p-1}\phi(t^d g(1,\mathbf{z}))\, dt \right) h(1,\mathbf{z})\, d\mathbf{z}$$

$$+ \int_{\mathbb{R}^{n-1}} \left(\int_0^\infty t^{n+p-1}\phi(t^d g(-1,-\mathbf{z}))\, dt \right) h(-1,-\mathbf{z})\, d\mathbf{z},$$

where the last two integrals are obtained from the sum of the previous two by using Tonelli's Theorem (since ϕ and h are nonnegative functions) and whether or not the integrals are finite; see e.g. Dunford and Schwartz [6, Theorem 14, p. 194].

Next, with the change of variable $u = t\,g(1,\mathbf{z})^{1/d}$ and $u = t\,g(-1,-\mathbf{z})^{1/d}$

$$(2.2) \qquad \int_{\mathbb{R}^n} \phi(g(\mathbf{x}))\, h(\mathbf{x})d\mathbf{x} \;=\; \left(\int_{\mathbb{R}_+} u^{n+p-1}\phi(u^d)\, du \right) \cdot A(g,h),$$

with

$$(2.3) \qquad A(g,h) = \int_{\mathbb{R}^{n-1}} \left(\frac{h(1,\mathbf{z})}{g(1,\mathbf{z})^{(n+p)/d}} + \frac{h(-1,-\mathbf{z})}{g(-1,-\mathbf{z})^{(n+p)/d}} \right) d\mathbf{z}.$$

In particular, the choice $t \mapsto \phi(t) = \exp(-t)$ yields:

$$(2.4) \qquad A(g,h) \;=\; \frac{d}{\Gamma((n+p)/d)} \int_{\mathbb{R}^n} h \exp(-g) \, d\mathbf{x},$$

and so $A(g,h)$ is finite if and ony if $\int h \exp(-g) d\mathbf{x}$ is finite. Substituting $A(g,h)$ in (2.2) with (2.4) yields (2.1). $\qquad\qquad\qquad\qquad\qquad\square$

2.2. **Integration on sublevel sets and non Gaussian integrals.** As a consequence of Lemma 2.1 we obtain the following result.

Theorem 2.2. *Let g, h be PHFs of degree $0 \neq d \in \mathbb{R}$ and $p \in \mathbb{R}$ respectively, and such that $g \in C_d$. Then for every $y \in [0, \infty)$:*

$$(2.5) \qquad \mathrm{vol}\left(\{\mathbf{x} \,:\, g(\mathbf{x}) \leq y\}\right) \;=\; \frac{y^{n/d}}{\Gamma(1+n/d)} \int_{\mathbb{R}^n} \exp(-g) \, d\mathbf{x},$$

and

$$(2.6) \qquad \int_{\{\mathbf{x}\,:\,g(\mathbf{x}) \leq y\}} h \, d\mathbf{x} \;=\; \frac{y^{(n+p)/d}}{\Gamma(1+(n+p)/d)} \int_{\mathbb{R}^n} h \, \exp(-g) \, d\mathbf{x},$$

whenever $\int_{\mathbb{R}^n} |h| \exp(-g) d\mathbf{x}$ (or $\int_{\{\mathbf{x}\,:\,g(\mathbf{x}) \leq 1\}} |h| \, d\mathbf{x}$) is finite. And (2.6) also holds if h is replaced with $|h|$, or with $\max[0,h]$, or with $\max[0,-h]$.

Proof. Write $h = h^+ - h^-$ with $h^+(\mathbf{x}) := \max[0, h(\mathbf{x})]$, $h^-(\mathbf{x}) := \max[0, -h(\mathbf{x})]$ for all $\mathbf{x}$. Observe that $|h|$, h^+ and h^- are nonnegative and positively homogeneous of degree p, and $|h| = h^+ + h^-$, $h = h^+ - h^-$.

With A as in (2.2), using Lemma 2.1 with $\phi(t) = \exp(-t)$, yields

$$\int_{\mathbb{R}^n} |h| \, \exp(-g) \, d\mathbf{x} \;=\; \frac{\Gamma(1+(n+p)/d)}{n+p} \cdot A(g, |h|)$$

where all integrals are finite since the right-hand-side is finite. Similarly, using Lemma 2.1 with $\phi(t) = 1_{[0,1]}(t)$, yields

$$\int_{\{\mathbf{x}:g(\mathbf{x}) \leq 1\}} |h| \, d\mathbf{x} \;=\; \frac{1}{n+p} \cdot A(g, |h|),$$

and therefore,

$$\int_{\{\mathbf{x}\,:\,g(\mathbf{x}) \leq 1\}} |h| \, d\mathbf{x} \;=\; \frac{1}{\Gamma(1+(n+p)/d)} \int_{\mathbb{R}^n} |h| \, \exp(-g) \, d\mathbf{x}.$$

Next, by homogeneity,

$$\int_{\{\mathbf{x}\,:\,g(\mathbf{x}) \leq y\}} |h| \, d\mathbf{x} \;=\; y^{(n+p)/d} \int_{\{\mathbf{x}\,:\,g(\mathbf{x}) \leq 1\}} |h| \, d\mathbf{x},$$

and so

$$(2.7) \qquad \int_{\{\mathbf{x}\,:\,g(\mathbf{x}) \leq y\}} |h| \, d\mathbf{x} \;=\; \frac{y^{(n+p)/d}}{\Gamma(1+(n+p)/d)} \int_{\mathbb{R}^n} |h| \, \exp(-g) \, d\mathbf{x}.$$

Obviously, with same arguments, (2.7) also holds if we replace $|h|$ with h^+ and h^- and finiteness follows because

$$0 \le \int_{\{\mathbf{x}\,:\,g(\mathbf{x})\le y\}} h^+ \, d\mathbf{x} \le \int_{\{\mathbf{x}\,:\,g(\mathbf{x})\le y\}} |h| \, d\mathbf{x},$$

and similarly for h^-. Hence (2.7) with h in lieu of $|h|$ (i.e. (2.6)) follows by additivity since $h = h^+ - h^-$.

Finally, from what precedes, finiteness of $\int_{\{\mathbf{x}:g(\mathbf{x})\le 1\}} |h| d\mathbf{x}$ is equivalent to finiteness of $\int_{\mathbb{R}^n} |h| \exp(-g) d\mathbf{x}$. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

When $y = 1$, g is a positive definite form with $d \in \mathbb{N}$, and h is identical to the constant function 1, (2.5) is already proved in Morosov and Shakirov [17, (75) p. 39].

In particular, when g is the quadratic form $\mathbf{x} \mapsto g_2(\mathbf{x}) := \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x}$ for some real symmetric positive definite matrix $\mathbf{Q}$, one retrieves that the volume of the ellipsoid $\xi(y) := \{\mathbf{x} : g_2(\mathbf{x}) \le y\}$ is simply related to the determinant of $\mathbf{Q}$ by the formula

$$(2.8) \qquad \mathrm{vol}\,(\xi(y)) \;=\; \frac{y^{n/2}}{\Gamma(1+n/2)} \underbrace{\int_{\mathbb{R}^n} \exp(-\frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x})\,d\mathbf{x}}_{=(2\pi)^{n/2}/\sqrt{\det\mathbf{Q}}}$$

So Theorem 2.2 states that the volume of the sublevel set is simply related to the integral of $\exp(-g(\mathbf{x}))$ over the entire domain $\mathbb{R}^n$ which happens to be simply related to the determinant of $\mathbf{Q}$ when g is the quadratic form $\mathbf{x}^T\mathbf{Q}\mathbf{x}$. One goal of the theory of integral discriminants is precisely to express $\int \exp(-g)$ in terms of invariants of g when g is a form. See e.g. Morosov and Shakirov [17, 18] and Shakirov [22].

Intersection of sublevel sets. Suppose that with h being positively homogeneous of degree p, one wishes to compute the integral $\int_\Omega h(\mathbf{x})\,d\mathbf{x}$ where

$$\Omega := \{\mathbf{x} \in \mathbb{R}^n \;:\; g_k(\mathbf{x}) \le z_k, \quad k = 1, \dots, m\},$$

for some m PHFs $g_1, \dots, g_m$ of degree $0 \ne d \in \mathbb{R}$, and some (strictly) positive vector $\mathbf{z} \in \mathbb{R}^m$. Equivalently,

$$\Omega := \{\mathbf{x} \in \mathbb{R}^n \;:\; \tilde{g}_k(\mathbf{x}; \mathbf{z}) \le 1, \quad k = 1, \dots, m\},$$

for the functions $\mathbf{x} \mapsto \tilde{g}_k(\mathbf{x}; \mathbf{z}) := g_k(z_k^{-1/d}\mathbf{x})$, $k = 1, \dots, m$, which are also PHFs of degree $d \in \mathbb{R}$. Hence with no loss of generality, one may restrict to sets of the form

$$(2.9) \qquad \Omega(y) \;:=\; \{\mathbf{x} \;:\; g_k(\mathbf{x}) \le y, \quad k = 1, \dots, m\}$$

for some positive scalar $y \in \mathbb{R}$, and PHFs $g_1, \dots, g_m$ of same degree $d \in \mathbb{R}$.

Notice that $\mathbf{x} \mapsto \psi(\mathbf{x}) := \max[g_1(\mathbf{x}), \dots, g_m(\mathbf{x})]$ is a PHF of degree d.

Corollary 2.3. *Let h be a PHF of degree $p \in \mathbb{R}$, let $g_1, \ldots, g_m$ be PHFs of degree $0 \neq d \in \mathbb{R}$, and assume that the set $\{\mathbf{x} : g_k(\mathbf{x}) \leq 1,\ k = 1, \ldots, m\}$ is bounded and for every $y > 0$, let $\Omega(y)$ be as in (2.9). Then:*

$$(2.10) \qquad \operatorname{vol}(\Omega(y)) \;=\; \frac{y^{n/d}}{\Gamma(1 + n/d)} \int_{\mathbb{R}^n} \exp(-\psi)\, d\mathbf{x},$$

and

$$(2.11) \qquad \int_{\Omega(y)} h(\mathbf{x})\, d\mathbf{x} \;=\; \frac{y^{(n+p)/d}}{\Gamma(1 + (n+p)/d)} \int_{\mathbb{R}^n} h \, \exp(-\psi)\, d\mathbf{x},$$

whenever $\int_{\mathbb{R}^n} |h|\, \exp(-\psi)\, d\mathbf{x}$.

Proof. Notice that $\mathbf{x} \mapsto \psi(\mathbf{x}) := \max[g_1(\mathbf{x}), \ldots, g_m(\mathbf{x})]$ is also a PHF of degree $0 \neq d \in \mathbb{R}$, and $\Omega(y) = \{\mathbf{x} : \psi(\mathbf{x}) \leq y\}$. And so if $\Omega(y)$ is bounded then $\psi \in C_d$. Hence (2.11)-(2.10) is just (2.5)-(2.6) with h and ψ in lieu of h and g. $\qquad\square$

2.3. An alternative proof with a duality interpretation.

Next, we present an alternative proof of Theorem 2.2 that uses Laplace transform techniques and provides an interpretation of the result in an appropriate *duality* framework.

Suppose that $g, h \in C_d$. Since g is nonnegative, the function $I_{g,h}$ vanishes on $(-\infty, 0]$. Its Laplace transform $\mathcal{L}_{I_{g,h}} : \mathbb{C} \to \mathbb{R}$ is the function

$$\lambda \mapsto \mathcal{L}_{I_{g,h}}(\lambda) := \int_0^\infty \exp(-\lambda y) I_{g,h}(y)\, dy,$$

and observe that

$$
\begin{aligned}
\mathcal{L}_{I_{g,h}}(\lambda) \;&=\; \int_0^\infty \exp(-\lambda y) \left(\int_{\{\mathbf{x}: g(\mathbf{x}) \leq y\}} h\, d\mathbf{x} \right) dy \\
&=\; \int_{\mathbb{R}^n} h(\mathbf{x}) \left(\int_{g(\mathbf{x})}^\infty \exp(-\lambda y)\, dy \right) d\mathbf{x} \quad \text{[by Fubini's Theorem]} \\
&=\; \frac{1}{\lambda} \int_{\mathbb{R}^n} h(\mathbf{x}) \exp(-\lambda g(\mathbf{x}))\, d\mathbf{x} \\
&=\; \frac{\lambda^{-p/d}}{\lambda} \int_{\mathbb{R}^n} h(\lambda^{1/d}\mathbf{x}) \exp(-g(\lambda^{1/d}\mathbf{x}))\, d\mathbf{x} \quad \text{[by homogeneity]} \\
&=\; \frac{1}{\lambda^{1+(n+p)/d}} \int_{\mathbb{R}^n} h(\mathbf{z}) \exp(-g(\mathbf{z}))\, d\mathbf{z} \quad \text{[by } \lambda^{-1/d}\mathbf{x} \to \mathbf{z}] \\
&=\; \frac{\displaystyle\int_{\mathbb{R}^n} h(\mathbf{z}) \exp(-g(\mathbf{z}))\, d\mathbf{z}}{\Gamma(1 + (n+p)/d)} \, \mathcal{L}_{y^{(n+p)/d}}(\lambda).
\end{aligned}
$$

And so, by uniqueness of the Laplace transform,

$$I_{g,h}(y) = \frac{y^{(n+p)/d}}{\Gamma(1 + (n+p)/d)} \int_{\mathbb{R}^n} h(\mathbf{x}) \exp(-g(\mathbf{x}))\, d\mathbf{x},$$

which is the desired result. So the above expression of $I_{g,h}(y)$ is obtained by "inverting" the Laplace transform $\mathcal{L}_{I_{g,h}}$ at the point y, which in fact, as we next see, is solving a "dual" problem.

For analogy purposes, consider, the optimization problem

$$\rho_{g,h}(y) = \sup_{\mathbf{x}} \{h(\mathbf{x}) : g(\mathbf{x}) \leq y\},$$

where y is fixed, h and $-g$ are concave and h is nonnegative. Equivalently, $\rho_{g,h}(y) = \exp(\theta_{g,h}(y))$ where $y \mapsto \theta_{g,h}(y)$ is the *optimal value function* of the optimization problem

$$\mathbf{P}_y : \quad \theta_{g,h}(y) = \sup_{\mathbf{x}} \{\ln h(\mathbf{x}) : g(\mathbf{x}) \leq y\}.$$

Associated with $\mathbf{P}_y$ is the *dual* problem $\mathbf{P}_y^* : \inf_\lambda \{\mathcal{G}(\lambda) : \lambda \geq 0\}$, where $G : \mathbb{R}_+ \to \mathbb{R}$ is the function $\lambda \mapsto G(\lambda) := \sup_{\mathbf{x}} \{\ln h(\mathbf{x}) + \lambda(y - g(\mathbf{x}))\}$. Observe that:

$$\begin{aligned}
\lambda \mapsto G(\lambda) &= \lambda y + \sup_{\mathbf{x}} \{\ln h(\mathbf{x}) - \lambda g(\mathbf{x})\} \\
&= \lambda y + \sup_{\mathbf{z}} \{\ln(\lambda^{-p/d} h(\mathbf{z})) - g(\mathbf{z})\} \quad [\text{via } \mathbf{z} = \lambda^{1/d}\mathbf{x}] \\
&= \lambda y + -\frac{p}{d} \ln \lambda + \sup_{\mathbf{z}} \{\ln h(\mathbf{z}) - g(\mathbf{z})\}.
\end{aligned}$$

And so the dual problem $\mathbf{P}_y^*$ reads

$$\begin{aligned}
\mathbf{P}_y^* : \quad \gamma &= \sup_{\mathbf{z}} \{\ln h(\mathbf{z}) - g(\mathbf{z})\} + \inf_{\lambda \geq 0} \{\lambda y - \frac{p}{d} \ln \lambda\} \\
&= \ln y^{p/d} + \ln\left(\sup_{\mathbf{z}} \{h(\mathbf{z}) \exp(-g(\mathbf{z}))\}\right) + \frac{p}{d}(1 - \ln\frac{p}{d})
\end{aligned}$$

In particular, if h is log-concave and g is convex, by a standard argument of convex optimization, $\gamma = \theta_{g,h}(y) \, (= \ln \rho_{g,h}(y))$. And therefore,

$$\rho_{g,h}(y) = \exp(\theta_{g,h}(y)) = y^{p/d} \frac{\exp(p/d)}{(p/d)^{p/d}} \sup_{\mathbf{x}} \{h(\mathbf{x}) \exp(-g(\mathbf{x}))\}$$

to compare with

$$I_{g,h}(y) = y^{(n+p)/d} \frac{1}{\Gamma(1 + (n+p)/d)} \int_{\mathbb{R}^n} h(\mathbf{x}) \exp(-g(\mathbf{x})) \, d\mathbf{x}.$$

Alternatively, the Legendre-Fenchel transform of $\theta_{g,h}(y)$ (for the concave version) is the function

$$\begin{aligned}
\theta_{g,h}^*(\lambda) &= \inf_y \{\lambda y - \theta_{g,h}(y)\} \\
&= \inf_\lambda \{\lambda y - \sup_x \{\ln h(\mathbf{x}) : g(\mathbf{x}) \leq y\}\} \\
&= \inf_{\mathbf{x}} \{\lambda g(\mathbf{x}) - \ln h(\mathbf{x})\} = -\ln\left(\sup_{\mathbf{x}} \{h(\mathbf{x}) \exp(-\lambda g(\mathbf{x}))\}\right)
\end{aligned}$$

and so when h is log-concave and g is convex, $\theta_{g,h}(y) = (\theta_{g,h}^*)^*(y)$, so that

$$
\begin{aligned}
\theta_{g,h}(y) &= \inf_\lambda \{\lambda y - \theta_{g,h}^*(\lambda)\} \\
&= \inf_\lambda \left\{\lambda y + \ln\left(\sup_\mathbf{x}\{h(\mathbf{x})\exp(-\lambda g(\mathbf{x}))\}\right)\right\} \\
&= \ln y^{p/d} + \ln\left(\sup_\mathbf{z}\{h(\mathbf{z})\exp(-g(\mathbf{z}))\}\right) + \frac{p}{d}(1 - \ln\frac{p}{d}).
\end{aligned}
$$

In summary, to the Laplace transform step

$$(2.12) \qquad \mathcal{L}_{I_{g,h}}(\lambda) = \int_{\mathbb{R}^n} \exp(-\lambda y) I_{g,h}(y)\,dy$$

in the usual $(\cdot, +)$-algebra, corresponds to the Legendre Fenchel transform

$$
\begin{aligned}
\theta_{g,h}^*(\lambda) = \inf_y\{\lambda y - \theta_{g,h}(y)\} &= -\sup_y\{-\lambda y + \theta_{g,h}(y)\} \\
&= -\ln\sup_y\{\exp(-\lambda y)\rho_{g,h}(y)\},
\end{aligned}
$$

where to emphasize the analogy, the latter term can be written

$$\ln\left(\text{``}\int\text{''}\exp(-\lambda y)\,\rho_{g,h}(y)\right),$$

i.e., the "sup" operator is integration "$\int$" in the $(\max, +)$-algebra. And with this convention

$$\exp(\theta_{g,h}^*(\lambda)) \simeq \text{``}\int\text{''}\exp(-\lambda y)\,\rho_{g,h}(y).$$

(Compare with (2.12).) Similarly, the Laplace inverse transform step

$$(2.13) \qquad I_{g,h}(y) = \int_{\omega-i\infty}^{\omega+i\infty} \exp(\lambda y)\,\mathcal{L}_{I_{g,h}}(\lambda)\,d\lambda$$

(where $\omega \in \mathbb{C}$ is on the right of all singularities of $\mathcal{L}_{I_{g,h}}$) and which yields

$$I_{g,h}(y) = \frac{y^{(n+p)/d}}{\Gamma(1 + (n+p)/d)}\int_{\mathbb{R}^n} h\exp(-g)d\mathbf{x},$$

is solving the "dual" and corresponds to the Legendre-Fenchel transform (which is involutive) applied to $\theta_{g,h}^*$

$$\exp(\theta_{g,h}(y)) = \exp(\inf_\lambda\{\lambda y - \theta_{g,h}^*(\lambda)\}) \simeq -\text{``}\int\text{''}\exp(-\lambda y)\,\exp(\theta_{g,h}^*(\lambda))\,d\lambda$$

(compare with (2.13)), and which yields

$$\exp(\theta_{g,h}(y)) = y^{p/d}\frac{\exp(p/d)}{(p/d)^{p/d}}\text{``}\int\text{''}h(\mathbf{x})\exp(-g(\mathbf{x})).$$

A rigorous analysis of the links between integration and linear optimization on a polytope has been already investigated in [10].

2.4. **Approximating non Gaussian integrals.** As we have already mentioned, in some cases the non Gaussian integral can be computed explicitly in terms of some on algebraic invariants of g; see e.g. Morosov and Shakirov [17] and Shakirov [22]. But so far there is no general formula and therefore an alternative is to seak for a numerical scheme for its evaluation, or at least, its approximation.

For this purpose, we next show that Theorem 2.2 is helpful as it provides a means to compute any moment of the measure $d\mu = \exp(-g)d\mathbf{x}$ on $\mathbb{R}^n$ by computing the same moment but now of the Lebesgue measure on the sublevel set $\{\mathbf{x} : g(\mathbf{x}) \leq 1\}$. Indeed, for every $\alpha \in \mathbb{N}^n$, letting $\mathbf{x} \mapsto h(\mathbf{x}) := \mathbf{x}^\alpha$ in Theorem 2.2, yields

$$(2.14) \qquad \int_{\mathbb{R}^n} \mathbf{x}^\alpha \exp(-g(\mathbf{x}))\, d\mathbf{x} \;=\; \Gamma(1 + (n + |\alpha|)/d) \int_{\{\mathbf{x}\,:\,g(\mathbf{x})\leq 1\}} \mathbf{x}^\alpha\, d\mathbf{x},$$

where $|\alpha| = \sum_i \alpha_i$. Have we made any progress with this equivalence?

The answer is yes. If g is a (non necessarily homogeneous) polynomial, it turns out that every moment of the Lebesgue measure on the sublevel set $\{\mathbf{x} : g(\mathbf{x}) \leq 1\}$ can be approximated as closely as desired by solving a hierarchy of semidefinite programs[2] as described in Henrion et al. [8]. In fact, for every $\alpha \in \mathbb{N}^n$ fixed, the moment

$$z_\alpha \;:=\; \int_{\{\mathbf{x}\,:\,g(\mathbf{x})\leq 1\}} \mathbf{x}^\alpha\, d\mathbf{x}$$

can be approximated to arbitrary precision $\epsilon > 0$ fixed in advance, by solving two sequences of semidefinite programs, one which provides a monotone non decreasing sequence of upper bounds u_k, $k \in \mathbb{N}$, while the other provides a monotone non increasing sequence of lower bounds ℓ_k, $k \in \mathbb{N}$. The procedure stops whenever $u_k - \ell_k < \epsilon$, in which case one may set

$$z_\alpha \approx \tilde{z}_\alpha \;:=\; (u_k + \ell_k)/2.$$

This requires to solve two sequences of semidefinite programs for each $\alpha \in \mathbb{N}^n$. In fact, if one is ready to relax the monotonicity property of the upper and lower bounds $\{u_k, \ell_k\}$, it is enough to solve a single sequence of semidefinite programs, e.g., the one defined to approximate the mass z_0. Then if $d \in \mathbb{N}$ and $\epsilon > 0$ are fixed, and k is large enough, not only $|u_k - z_0| < \epsilon$ but also from the solution of the semidefinite program at step k one obtains scalars $\tilde{z}_\alpha$ such that $|\tilde{z}_\alpha - z_\alpha| < \epsilon$, for all $\alpha \in \mathbb{N}^n$ such that $|\alpha| < d$. However, in contrast to the case of upper and lower bounds, there is no simple stopping criterion to guarantee the ϵ-approximation. For more details, the interested reader is referred to Henrion et al. [8]. And therefore, once the

---

[2]A semidefinite program is a finite-dimensional convex conic optimization problem, that up to arbitrary (fixed) precision, can be solved efficiently, i.e., in time polynomial in the input size of the problem. For more details the interested reader is referred to e.g. [24].

$\tilde{z}_\alpha$ have been computed, using Theorem 2.2, we obtain:

$$\left| \int_{\mathbb{R}^n} \mathbf{x}^\alpha \exp(-g(\mathbf{x}))\, d\mathbf{x} \, - \, \tilde{z}_\alpha \, \Gamma(1 + (n + |\alpha|)/d) \right| < \epsilon, \quad \forall \alpha \in \mathbb{N}_d^n,$$

which provides an ϵ-approximation guarantee for the non Gaussian integral $\int_{\mathbb{R}^n} \exp(-g)d\mathbf{x}$ (and more generally for the integral $\int_{\mathbb{R}^n} h \exp(-g)d\mathbf{x}$ whenever h is any polynomial).

2.5. **Sensitivity analysis and convexity.** Recall that when $d \in \mathbb{N}$, $\mathbf{P}[\mathbf{x}]_d \subset C_d$ is the convex cone of nonnegative and homogeneous polynomials of degree d, with compact sublevel set $\{\mathbf{x} : g(\mathbf{x}) \leq 1\}$. Formula (2.6) of Theorem 2.2 allows us to provide insights into the function $f : C_d \to \mathbb{R}$, defined by:

$$(2.15) \qquad g \mapsto f_h(g) := \int_{\{\mathbf{x} : g(\mathbf{x}) \leq 1\}} h(\mathbf{x})\, d\mathbf{x}, \qquad g \in C_d,$$

where h is a PHF. In particular when one wishes to see how f_h changes when some coefficient of $g \in \mathbf{P}[\mathbf{x}]_d$ varies. Notice that the restriction of f_h to $\mathbf{P}[\mathbf{x}]_d$ may be seen as a function $f_h : \mathbb{R}^{\ell(d)} \to \mathbb{R}$ of the coefficient vector of the polynomial $g \in \mathbf{P}[\mathbf{x}]_d$, where $\ell(d) = \binom{n+d-1}{d}$.

Before proceeding further we need the following result. Recall that the support $\operatorname{supp}\mu$ of a Borel measure on $\mathbb{R}^n$ is the smallest closed set A such that $\mu(\mathbb{R}^n \setminus A) = 0$. Let $C_d^0 := \{g \in C_d : g \text{ is continuous}\} \subset C_d$.

Lemma 2.4. *Let μ be a non trivial σ-finite Borel measure on $\mathbb{R}^n$ and let $\Theta_\mu \subset \mathbb{R}[\mathbf{x}]_d$ be the convex cone of polynomials g of degree at most d such that $\int \exp(-g)d\mu < \infty$. Then:*

(a) With $d \in \mathbb{R}$, the function $f : C_d \to \mathbb{R}$, with $g \mapsto f(g) := \int \exp(-g)d\mu$, is convex (and strictly convex on C_d^0 if $\operatorname{supp}\mu = \mathbb{R}^n$).

(b) If $\mu(O) > 0$ for some open set $O \subset \mathbb{R}^n$ and if $d \in \mathbb{N}$, then f is strictly convex and twice differentiable on $\operatorname{int}(\Theta_\mu)$, with:

$$(2.16) \qquad \frac{\partial f(g)}{\partial g_\alpha} = \int \mathbf{x}^\alpha \exp(-g)\, d\mu, \qquad \forall \alpha, \ |\alpha| \leq d.$$

$$(2.17) \qquad \frac{\partial^2 f(g)}{\partial g_\alpha \partial g_\beta} = \int \mathbf{x}^{\alpha+\beta} \exp(-g)\, d\mu, \qquad \forall \alpha, \beta, \ |\alpha|, |\beta| \leq d.$$

Proof. (a) Observe that f is nonnegative. Let $\alpha \in [0, 1]$ and let $g, q \in C_d$. To prove $f(\alpha g + (1-\alpha)q) \leq \alpha f(g) + (1-\alpha)f(q)$, we only need consider the case where $f(g), f(q) < +\infty$, for which we have

$$f(\alpha g + (1-\alpha)q) = \int \exp(-\alpha g - (1-\alpha)q)\, d\mu.$$

By convexity of $u \mapsto \exp(-u)$,

$$\begin{aligned} f(\alpha g + (1-\alpha)q) &\leq \int \left[\alpha \exp(-g) + (1-\alpha)\exp(-q) \right] d\mu \\ &= \alpha f(g) + (1-\alpha)f(q), \end{aligned}$$

and so f is convex. Now, in view of the strict convexity of $u \mapsto \exp(-u)$, equality may occur only if $g(\mathbf{x}) = q(\mathbf{x})$, μ-almost everywhere. If $g, q \in C_d^0$, the set $\Delta := \{\mathbf{x} : g(\mathbf{x}) - q(\mathbf{x}) = 0\}$ is closed and so if $\Delta \neq \mathbb{R}^n$ then $\mu(\Delta) < \mu(\mathbb{R}^n)$ because $\mathrm{supp}\,\mu = \mathbb{R}^n$. Therefore, equality occurs only if $g = q$ so that f is strictly convex on C_d^0.

(b) Next, if $d \in \mathbb{N}$ and $g \in \mathrm{int}(\Theta_\mu)$, write g in the canonical basis as $g(\mathbf{x}) = \sum_\alpha g_\alpha \mathbf{x}^\alpha$. For every $\alpha \in \mathbb{N}_d^n$, 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 $t \geq 0$,

$$\frac{f(g + te_\alpha) - f(g)}{t} = \int \exp(-g) \left(\underbrace{\frac{\exp(-t\mathbf{x}^\alpha) - 1}{t}}_{\psi(t,\mathbf{x})} \right) d\mu(\mathbf{x})$$

Notice that for every $\mathbf{x}$, by convexity of the function $t \mapsto \exp(-t\mathbf{x}^\alpha)$,

$$\lim_{t \downarrow 0} \psi(t, \mathbf{x}) = \inf_{t \geq 0} \psi(t, \mathbf{x}) = \exp(-t\mathbf{x}^\alpha)'_{|t=0} = -\mathbf{x}^\alpha,$$

because for every $\mathbf{x}$, the function $t \mapsto \psi(t, \mathbf{x})$ is nondecreasing; see e.g. Rockafellar [21, Theorem 23.1]. Hence, the one-sided directional derivative $f'(g; e_\alpha)$ in the direction e_α satisfies

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

where the third equality follows from the Extended Monotone Convergence Theorem [2, 1.6.7]. Indeed for all $t < t_0$ with t_0 sufficiently small, the function $\psi(t, \cdot)$ is bounded above by $\psi(t_0, \cdot)$ and $\int \exp(-g)\psi(t_0, \mathbf{x})d\mu < \infty$. Similarly, for every $t > 0$

$$\frac{f(g - te_\alpha) - f(g)}{t} = \int \exp(-g) \underbrace{\frac{\exp(t\mathbf{x}^\alpha) - 1}{t}}_{\xi(t,\mathbf{x})} d\mu(\mathbf{x}),$$

and by convexity of the function $t \mapsto \exp(t\mathbf{x}^\alpha)$

$$\lim_{t \downarrow 0} \xi(t, \mathbf{x}) = \inf_{t \geq 0} \xi(t, \mathbf{x}) = \exp(t\mathbf{x}^\alpha)'_{|t=0} = \mathbf{x}^\alpha.$$

Therefore, with exactly same arguments as before,

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

and so

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

for every α with $|\alpha| \leq d$, which yields (2.16). Similar arguments can used for the Hessian $\nabla^2 f(g)$ which yields (2.17).

So the Hessian $\nabla^2 f(g)$ is the matrix $\mathbf{M}_d \in \mathcal{S}_{\ell(d)}$ whose rows and columns are indexed in the set $\Gamma_d := \{\alpha \in \mathbb{N}^n : |\alpha| = d\}$ and with entries

$$\mathbf{M}_d(\alpha, \beta) = \int_{\mathbb{R}^n} \mathbf{x}^{\alpha+\beta} \underbrace{\exp(-g) \, d\mu}_{d\nu}, \qquad \alpha, \beta \in \Gamma_d,$$

i.e., $\mathbf{M}_d$ is the matrix of $2d$-moments of the finite Borel measure ν. Let $\mathbf{h} \in \mathbb{R}^{\ell(d)}$ be the coefficient vector of a non trivial and arbitrary homogeneous polynomial $h \in \mathbb{R}[\mathbf{x}]_d$. Then

$$\langle \mathbf{h}, \mathbf{M}_d \mathbf{h} \rangle \; \left(= \int_{\mathbb{R}^n} h(\mathbf{x})^2 \, d\nu(\mathbf{x}) \right) > 0$$

because $\mu(O) > 0$ (hence $\nu(O) > 0$) on some open set $O \subset \mathbb{R}^n$. Therefore $\nabla^2 f(g) \succ 0$, which in turn implies that f is strictly convex on $\mathrm{int}(\Theta_\mu)$. $\quad\square$

Corollary 2.5. *Let h be a PHF of degree $p \in \mathbb{R}$ and with $0 \neq d \in \mathbb{R}$, consider the function $f_h : C_d \to \mathbb{R}$ defined by:*

$$(2.18) \qquad g \mapsto f_h(g) := \int_{\{\mathbf{x}:g(\mathbf{x})\leq 1\}} h(\mathbf{x}) \, d\mathbf{x}, \qquad \forall g \in C_d.$$

The function f_h is a PHF of degree $-(n+p)/d$ and convex whenever h is nonnegative (and strictly convex if $h > 0$ on $\mathbb{R}^n \setminus \{0\}$). In addition, if $d \in \mathbb{N}$ and h is continuous with $\int |h| \exp(-g) d\mathbf{x} < \infty$, then:

   (a) *f_h is twice differentiable on $\mathrm{int}(\mathbf{P}[\mathbf{x}]_d)$, and for every $\alpha, \beta \in \mathbb{N}_d^n$:*

$$(2.19) \quad \frac{\partial f_h(g)}{\partial g_\alpha} = \frac{-1}{\Gamma(1+(n+p)/d)} \int_{\mathbb{R}^n} \mathbf{x}^\alpha \, h(\mathbf{x}) \exp(-g(\mathbf{x})) \, d\mathbf{x}$$

$$(2.20) \qquad\qquad = \frac{-\Gamma(2+(n+p)/d)}{\Gamma(1+(n+p)/d)} \int_{\{\mathbf{x}\,:\,g(\mathbf{x})\leq 1\}} \mathbf{x}^\alpha \, h(\mathbf{x}) \, d\mathbf{x}.$$

$$(2.21) \quad \frac{\partial^2 f_h(g)}{\partial g_\alpha \partial g_\beta} = \frac{1}{\Gamma(1+(n+p)/d)} \int_{\mathbb{R}^n} \mathbf{x}^{\alpha+\beta} \, h(\mathbf{x}) \exp(-g(\mathbf{x})) \, d\mathbf{x}.$$

(b) *If h is non trivial and nonnegative, then f_h is strictly convex on $\mathrm{int}(\mathbf{P}[\mathbf{x}]_d)$, and its Hessian $\nabla^2 f_h(g)$ is the matrix of $2d$-moments of the measure*

$$\nu(B) = \frac{1}{\Gamma(1+(n+p)/d)} \int_B h \, \exp(-g) \, d\mathbf{x}, \qquad B \in \mathcal{B}.$$

Proof. With $\lambda > 0$ and $g \in C_d$,

$$f_h(\lambda g) = \int_{\{\mathbf{x}:\lambda g(\mathbf{x})\leq 1\}} h \, d\mathbf{x} = \int_{\{\mathbf{x}:g(\lambda^{1/d}x)\leq 1\}} h \, d\mathbf{x},$$

and so, doing the change of variable $\mathbf{z} = \lambda^{1/d}\mathbf{x}$, one obtains $f_h(\lambda g) = \lambda^{-(n+p)/d} f_h(g)$, i.e., f_h is a PHF of degree $-(n+p)/d$.

Next, first consider the case where h is nonnegative, and let μ be the σ-finite measure defined by $\mu(B) := \int_B h(\mathbf{x})d\mathbf{x}$ for every Borel set B of $\mathbb{R}^n$. By Theorem 2.2, whenever $g \in C_d$ and $f_h(g)$ is finite,

$$f_h(g) = \frac{1}{\Gamma(1 + (n+p)/d)} \int_{\mathbb{R}^n} \exp(-g)\, d\mu.$$

Hence by Lemma 2.4, f_h is convex (and strictly convex if $h > 0$ on $\mathbb{R}^n \setminus \{0\}$).

In addition, if $d \in \mathbb{N}$ and h is continuous, $f_h(g) < \infty$ if $g \in \text{int}(\mathbf{P}[\mathbf{x}]_d)$ (and so $g \in \text{int}(\Theta_\mu)$). Moreover, h being non trivial, nonnegative and continuous, $h > 0$ on some open set O and so $\mu(O) > 0$. Therefore, by Lemma 2.4(b), f_h is twice differentiable and strictly convex on $\text{int}(\mathbf{P}[\mathbf{x}]_d)$ and (2.16)-(2.17) yield (2.19) and (2.21), while (2.20) follows from (2.19) and Theorem 2.2. That $\nabla^2 f_h(g)$ is the matrix of $2d$-moments of the measure $d\nu = h \exp(-g)d\mathbf{x}$ follows from (2.21). This proves (b).

To prove (a) when h is not nonnegative, write $h = h^+ - h^-$ with $h^+ := \max[0, h]$ and $h^- := \max[0, -h]$. Both h^+ and h^- are continuous PHFs of degree p, and nonnegative. Moreover, $f_h = f_{h^+} - f_{h^-}$ and so applying (b) to f_{h^+} and f_{h^-}, yields (a) by additivity.

$\square$

Remark 2.6. *(a) Notice that proving convexity of f_h directly from its definition (2.18) is not obvious at all whereas it becomes much easier when using Theorem 2.2.*

(b) In Lemma 2.4, differentiability of f on the convex cone C_d should be now in the sense of Gâteaux-differentiablity, not explored here.

We end up with the following relatively surprising results which even though are again particular cases of Lemma 2.1, deserve special mention.

Lemma 2.7. *Let $y \geq 0$ be fixed, let h be a PHF of degree $p \in \mathbb{R}$ and let $\xi, \psi : \mathbb{R}_+ \to \mathbb{R}$ be measurable functions such that*

$$\int_{\{t : \psi(t^d) \leq y\}} t^{n+p-1} \xi(t^d)\, dt \ < \ +\infty.$$

Let $f_h : C_d \to \mathbb{R}$, $0 \neq d \in \mathbb{R}$, be the function:

$$g \mapsto f_h(g) := \int_{\{\mathbf{x} : \psi(g(\mathbf{x})) \leq y\}} \xi(g(\mathbf{x}))\, h(\mathbf{x})\, d\mathbf{x}, \qquad g \in C_d.$$

Then whenever $\int |h| \exp(-g)d\mathbf{x} < +\infty$, $f_h(g)$ is finite, and

$$(2.22) \qquad f_h(g) = \underbrace{\frac{d \displaystyle\int_{\{t : \psi(t^d) \leq y\}} t^{n+p-1} \xi(t^d)\, dt}{\Gamma((n+p)/d)}}_{\text{cte}(y, p, d, \xi, \psi)} \cdot \int_{\mathbb{R}^n} h \exp(-g)d\mathbf{x},$$

where the constant depends only on ξ, ψ, d, p, y and neither on g nor h. Therefore, f_h is convex whenever h is nonnegative (and strictly convex on C_d^0 whenever $h > 0$ on $\mathbb{R}^n \setminus \{0\}$). In particular,

$$(2.23) \qquad \int_{\mathbb{R}^n} gh \, \exp(-g) \, d\mathbf{x} \;=\; \frac{n+p}{d} \int_{\mathbb{R}^n} h \, \exp(-g) \, d\mathbf{x}$$

$$(2.24) \qquad \int_{\{\mathbf{x}\,:\,g(\mathbf{x})\leq 1\}} gh \, d\mathbf{x} \;=\; \frac{1}{\Gamma((n+p)/d)} \int_{\mathbb{R}^n} h \, \exp(-g) \, d\mathbf{x}$$

$$(2.25) \qquad \frac{\displaystyle\int_{\{\mathbf{x}\,:\,g(\mathbf{x})\leq y\}} \exp(-g) \, d\mathbf{x}}{\displaystyle\int_{\mathbb{R}^n} \exp(-g) \, d\mathbf{x}} \;=\; \frac{\displaystyle\int_0^y \exp(-z) z^{n/d-1} \, dz}{\Gamma(n/d)}$$

$$(2.26) \qquad \int_{\{\mathbf{x}\,:\,g(\mathbf{x})\leq y\}} \exp(g) \, d\mathbf{x} \;=\; \frac{\displaystyle\int_{\mathbb{R}^n} \exp(-g) \, d\mathbf{x}}{\Gamma(n/d)} \int_0^y \exp(z) z^{n/d-1} \, dz.$$

Proof. We first assume that h is nonnegative and $\int h \exp(-g) < +\infty$, so that (2.22) follows from Lemma 2.1 with $t \mapsto \phi(t) := \xi(t) I_{[0,y]}(\psi(t))$, which yields $f_h(g) = \mathrm{cte}(y, p, d, \xi, \psi) \cdot A(g, h)$, with $A(\cdot, \cdot)$ as in (2.3) and

$$\mathrm{cte}(y, p, d, \xi, \psi) \;=\; \int_{\{t:\psi(t^d)\leq y\}} t^{n+p-1} \xi(t^d) \, dt,$$

and the result follows by recalling that with $\phi(t) = \exp(-t)$ one had already obtained in (2.4)

$$A(g, h) \;=\; \frac{d}{\Gamma((n+p)/d)} \int_{\mathbb{R}^n} h \, \exp(-g) \, d\mathbf{x} \; (< \, +\infty).$$

When h is not nonnegative, writing $h = h^+ - h^-$ where both h^+ and h^- are also PHFs of degree p, the result follows by additivity since $\int |h| \exp(-g) d\mathbf{x} < +\infty$ only if both $\int h^+ \exp(-g) d\mathbf{x}$ and $\int h^- \exp(-g) d\mathbf{x}$ are finite.

Finally, (2.23)-(2.26) are special cases of (2.22) with respective choices $t \mapsto \psi(t) := 0, \xi(t) := \exp(-t)$, then $t \mapsto \psi(t) := t, \xi(t) = t$ and finally, $t \mapsto \psi(t) = t, \xi(t) := \exp(-t)$ and $t \mapsto \xi(t) := \exp(t)$.

At last, when h is nonnegative, convexity and strict convexity follow from Lemma 2.4(a) with $d\mu = h d\mathbf{x}$ (and so $\mathrm{supp}\,\mu = \mathbb{R}^n$ if $h > 0$). $\qquad \square$

So Lemma 2.7 shows that f is convex provided that h is nonnegative and no matter how the functions ξ and ψ behave!

Next, if μ is the non Gaussian measure $d\mu = \exp(-g) d\mathbf{x}$ on $\mathbb{R}^n$, then (2.25) shows how fast $\mu(\{\mathbf{x} : g(\mathbf{x}) \leq y\})$ converges to the non Gaussian integral $\int_{\mathbb{R}^n} \exp(-g) d\mathbf{x}$ as $y \to \infty$. It converges as fast as the one-dimensional integral $\int_0^y t^{n/d-1} \exp(-t) dt$ converges to the Gamma function $\Gamma(n/d)$.

2.6. **Polarity.** We here investigate the *polar* G° set of the sublevel set $G :=$ $\{\mathbf{x} : g(\mathbf{x}) \leq 1\}$ assumed to be compact and when g is a proper closed convex PHF. In this case G is a convex body, and in fact g is a *gauge*.

The polar C° of a set $C \subset \mathbb{R}^n$ is the convex set defined by

$$C^\circ = \{\mathbf{x} \in \mathbb{R}^n : \sigma_C(\mathbf{x}) \leq 1\} \quad \text{with } \sigma_C(\mathbf{x}) := \sup_{\mathbf{y}} \{\langle \mathbf{x}, \mathbf{y} \rangle : \mathbf{y} \in C\}.$$

and $(C^\circ)^\circ$ is the smallest convex balanced set that contains C.

Recall that the Legendre-Fenchel *conjugate* $f^* : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty, -\infty\}$ of $f : \mathbb{R}^n \to \mathbb{R}$ is defined by

$$f^*(\mathbf{u}) := \sup_{\mathbf{x}} \{\langle \mathbf{u}, \mathbf{x} \rangle - f(\mathbf{x})\} \qquad \mathbf{u} \in \mathbb{R}^n.$$

The conjugate g^* of a PHF g of degree $d \in \mathbb{R}$ is tself a PHF of degree q with $\frac{1}{d} + \frac{1}{q} = 1$ (where d does not need to be positive); see e.g. Lasserre [14].

Proposition 2.8. *Let g be closed proper convex PHF of degree $1 < d \in \mathbb{R}$, and let $G := \{\mathbf{x} : g(\mathbf{x}) \leq 1/d\}$. Then with $\frac{1}{d} + \frac{1}{q} = 1$,*

$$(2.27) \qquad\qquad G^\circ = \{\mathbf{x} \in \mathbb{R}^n : g^*(\mathbf{x}) \leq 1/q\},$$

where g^* is the Legendre-Fenchel conjugate of g. In other words, G° is a sublevel set the PHF g^* of degree q. Moreover, if G is bounded then

$$(2.28) \qquad \text{vol}(G^\circ) = \frac{1}{q^{n/q}\Gamma(1 + n/q)} \int_{\mathbb{R}^n} \exp(-g^*)\, d\mathbf{x}.$$

Proof. (2.27) is from Rockafellar [21, Corollary 15.3.2] and (2.28) follows from Theorem 2.2 applied to the PHF g^* and with $y = 1/q$. $\square$

Example 1. *Let $x \mapsto g(x) := |x|^3$ if $x > 0$ and $+\infty$ otherwise. Hence, $d = 3$, $q = 3/2$, $G = [-1, 1]$, $\sigma_G(x) = |x|$, and $g^*(x) = \frac{2}{3\sqrt{3}}|x|^{3/2}$. One retrieves that $G^\circ = G = [-1, 1] = \{x : g^*(x) \leq (d-1)/d^q\}$.*

Example 2. *Let $x \mapsto g(x) := |x|$ so that $G = [-1, 1]$. As $g^*(x) = 0$ if $x \in [-1, 1]$ and $+\infty$ otherwise (a PHF of degree 0) one may check that indeed $G^\circ = [-1, 1] = G$ and (2.27) holds although $d = 1$ and g is not strictly convex.*

Example 3. *Let $\mathbf{x} \mapsto g(\mathbf{x}) := x_1^4 + x_2^4$ and $G = \{\mathbf{x} : x_1^4 + x_2^4 \leq 1/4\}$. Then $g^*(\mathbf{x}) = 3(x_1^{4/3} + x_2^{4/3})/4^{4/3}$ (a PHF of degree 4/3) and $G^\circ = \{\mathbf{x} : x_1^{4/3} + x_2^{4/3} \leq 1/4^{1/3}\}$.*

2.7. **A variational property of homogeneous polynomials.** We end up this section with an intriguing variational property of homogeneous polynomials that are sums of squares.

Let $\tilde{\mathbf{v}}_d(\mathbf{x})$ be the vector of all monomials $(\mathbf{x}^\alpha)$ of degree d and let $g \in \mathbb{R}[\mathbf{x}]_{2d}$ be homogeneous and a sum of squares, that is,

$$(2.29) \qquad g(\mathbf{x}) = -\frac{1}{2}\tilde{\mathbf{v}}_d(\mathbf{x})^T \boldsymbol{\Sigma} \tilde{\mathbf{v}}_d(\mathbf{x}), \qquad \mathbf{x} \in \mathbb{R}^n,$$

for some real symmetric $\ell(d) \times \ell(d)$ matrix $\boldsymbol{\Sigma}$ which is positive definite (denoted $\boldsymbol{\Sigma} \succ 0$). If $d = 1$ it is well-known that

$$\int_{\mathbb{R}^n} \exp(-g) d\mathbf{x} = \frac{(2\pi)^{n/2}}{\sqrt{\det \boldsymbol{\Sigma}}},$$

and

$$\int_{\mathbb{R}^n} \tilde{\mathbf{v}}_d(\mathbf{x}) \tilde{\mathbf{v}}_d^T(\mathbf{x}) \exp(-g) d\mathbf{x} = \frac{(2\pi)^{n/2}}{\sqrt{\det \boldsymbol{\Sigma}}} \boldsymbol{\Sigma}^{-1},$$

that is, $\boldsymbol{\Sigma}^{-1}$ is the covariance matrix associated with the Gaussian probability density $(2\pi)^{-n/2} \sqrt{\det \boldsymbol{\Sigma}} \exp(-g)$.

When $d > 1$ the non Gaussian integral can be still expressed as a (possibly complicated) combination of several algebraic invariants of g, but in general not in terms of the single algebraic invariant $\det \boldsymbol{\Sigma}$.

It turns out that $\det \boldsymbol{\Sigma}$ and $\int \exp(-\frac{1}{2} \mathbf{v}_d(\mathbf{x})^T \boldsymbol{\Sigma} \mathbf{v}_d(\mathbf{x})) d\mathbf{x}$ are still related in the following Gaussian-like manner. Let $\mathcal{S}_{\ell(d)}^{++} \subset \mathcal{S}_{\ell(d)}$ be the convex cone of $\ell(d) \times \ell(d)$ positive definite matrices, and let $\theta_d : \mathcal{S}_{\ell(d)}^{++} \to \mathbb{R}$ be defined by:

$$(2.30) \quad \theta_d(\boldsymbol{\Sigma}) := (\det \boldsymbol{\Sigma})^k \int_{\mathbb{R}^n} \exp(-k \tilde{\mathbf{v}}_d(\mathbf{x})^T \boldsymbol{\Sigma} \, \tilde{\mathbf{v}}_d(\mathbf{x})) \, d\mathbf{x}, \qquad \boldsymbol{\Sigma} \in \mathcal{S}_{\ell(d)}^{++},$$

where $k = n/(2d\ell(d))$ and let

$$\mathbf{M}_d(\boldsymbol{\Sigma}) := \frac{\displaystyle\int_{\mathbb{R}^n} \tilde{\mathbf{v}}_d(\mathbf{x}) \tilde{\mathbf{v}}_d(\mathbf{x})^T \exp(-k \tilde{\mathbf{v}}_d(\mathbf{x})^T \boldsymbol{\Sigma} \, \tilde{\mathbf{v}}_d(\mathbf{x})) \, d\mathbf{x}}{\displaystyle\int_{\mathbb{R}^n} \exp(-k \tilde{\mathbf{v}}_d(\mathbf{x})^T \boldsymbol{\Sigma} \, \tilde{\mathbf{v}}_d(\mathbf{x})) \, d\mathbf{x}},$$

be the matrix of moments of order d, associated with the non Gaussian probability measure

$$(2.31) \qquad \mu(B) := \frac{\displaystyle\int_B \exp(-k \tilde{\mathbf{v}}_d(\mathbf{x})^T \boldsymbol{\Sigma} \tilde{\mathbf{v}}_d(\mathbf{x})) \, d\mathbf{x}}{\displaystyle\int_{\mathbb{R}^n} \exp(-k \tilde{\mathbf{v}}_d(\mathbf{x})^T \boldsymbol{\Sigma} \tilde{\mathbf{v}}_d(\mathbf{x})) \, d\mathbf{x}}, \qquad B \in \mathcal{B}.$$

Observe that θ_d is nonnegative and positively homogeneous of degree 0; therefore θ_d is constant in any fixed direction $\boldsymbol{\Sigma}$. In particular, if $d = 1$ then $k = 1/2$, μ is a Gaussian probability measure, $\mathbf{M}_1(\boldsymbol{\Sigma})$ is the associated covariance matrix $\boldsymbol{\Sigma}^{-1}$ and $\theta_d(\boldsymbol{\Sigma})$ is constant. In fact,

Lemma 2.9. *Let θ_d be the function defined in (2.30). Then $\langle \mathbf{M}_d(\boldsymbol{\Sigma}), \boldsymbol{\Sigma} \rangle = \ell(d)$ for all $\boldsymbol{\Sigma}$ in the domain of θ_d and $\nabla \theta_d(\boldsymbol{\Sigma}) = 0$ if*

$$(2.32) \qquad\qquad \mathbf{M}_d(\boldsymbol{\Sigma}) = \boldsymbol{\Sigma}^{-1}.$$

Proof. Let $g(\mathbf{x}) = k\tilde{\mathbf{v}}_d(\mathbf{x})\boldsymbol{\Sigma}\tilde{\mathbf{v}}_d(\mathbf{x})$ so that

$$
\begin{aligned}
\langle \mathbf{M}_d(\boldsymbol{\Sigma}), \boldsymbol{\Sigma} \rangle \int_{\mathbb{R}^n} \exp(-g)\, d\mathbf{x} &= \left\langle \iint_{\mathbb{R}^n} \tilde{\mathbf{v}}_d(\mathbf{x})\tilde{\mathbf{v}}_d(\mathbf{x})^T \exp(-g)\, d\mathbf{x}, \boldsymbol{\Sigma} \right\rangle \\
&= k^{-1} \int_{\mathbb{R}^n} k\tilde{\mathbf{v}}_d(\mathbf{x})^T \boldsymbol{\Sigma}\tilde{\mathbf{v}}_d(\mathbf{x}) \exp(-g)\, d\mathbf{x} \\
&= k^{-1} \int_{\mathbb{R}^n} g\, \exp(-g)\, d\mathbf{x} \\
&= \frac{k^{-1}n}{2d} \int_{\mathbb{R}^n} \exp(-g)\, d\mathbf{x} \quad \text{[by (2.23)]},
\end{aligned}
$$

which yields the desired result $\langle \mathbf{M}_d(\boldsymbol{\Sigma}), \boldsymbol{\Sigma} \rangle = \ell(d)$. Next, write the gradient $\nabla\theta(\boldsymbol{\Sigma})$ in the form $A_1 + A_2$ with

$$
\begin{aligned}
A_1 &= \nabla((\det\boldsymbol{\Sigma})^k) \int_{\mathbb{R}^n} \exp(-k\tilde{\mathbf{v}}_d(\mathbf{x})^T \boldsymbol{\Sigma}\tilde{\mathbf{v}}_d(\mathbf{x}))\, d\mathbf{x} \\
&= k(\det\boldsymbol{\Sigma})^{k-1}\boldsymbol{\Sigma}^{\mathbb{A}} \int_{\mathbb{R}^n} \exp(-k\tilde{\mathbf{v}}_d(\mathbf{x})^T \boldsymbol{\Sigma}\tilde{\mathbf{v}}_d(\mathbf{x}))\, d\mathbf{x} \\
&= k\frac{\boldsymbol{\Sigma}^{\mathbb{A}}}{\det\boldsymbol{\Sigma}}(\det\boldsymbol{\Sigma})^k \int_{\mathbb{R}^n} \exp(-k\tilde{\mathbf{v}}_d(\mathbf{x})^T \boldsymbol{\Sigma}\tilde{\mathbf{v}}_d(\mathbf{x}))\, d\mathbf{x} \\
&= k\boldsymbol{\Sigma}^{-1} \theta_d(\boldsymbol{\Sigma}),
\end{aligned}
$$

where $\boldsymbol{\Sigma}^{\mathbb{A}}$ is the *adjugate* of $\boldsymbol{\Sigma}$ (see e.g. [5, p. 411]), and

$$
\begin{aligned}
A_2 &= (\det\boldsymbol{\Sigma})^k \nabla \left(\int_{\mathbb{R}^n} \exp(-k\tilde{\mathbf{v}}_d(\mathbf{x})^T \boldsymbol{\Sigma}\tilde{\mathbf{v}}_d(\mathbf{x}))\, d\mathbf{x} \right) \\
&= -k(\det\boldsymbol{\Sigma})^k \int_{\mathbb{R}^n} \tilde{\mathbf{v}}_d(\mathbf{x})\tilde{\mathbf{v}}_d(\mathbf{x})^T \exp(-k\tilde{\mathbf{v}}_d(\mathbf{x})^T \boldsymbol{\Sigma}\tilde{\mathbf{v}}_d(\mathbf{x}))\, d\mathbf{x} \\
&= -k\mathbf{M}_d(\boldsymbol{\Sigma})\, \theta_d(\boldsymbol{\Sigma}).
\end{aligned}
$$

This yields $A_1 + A_2 = k\theta_d(\boldsymbol{\Sigma})(\boldsymbol{\Sigma}^{-1} - \mathbf{M}_d(\boldsymbol{\Sigma}))$ and so $\nabla\theta_d(\boldsymbol{\Sigma}) = 0$ if $\mathbf{M}_d(\boldsymbol{\Sigma}) = \boldsymbol{\Sigma}^{-1}$. $\square$

Lemma 2.9 states that for all critical points $\boldsymbol{\Sigma}$, or equivalently for all critical SOS homogeneous polynomials g of the function θ_d (assuming that at least one such critical point exists), their associated non Gaussian measure $d\mu = \exp(-g)d\mu$ (rescaled to a probability measure) has the Gaussian-like property that $\boldsymbol{\Sigma}^{-1}$ is the "d-covariance" matrix of μ!

3. SUBLEVEL SET OF MINIMUM VOLUME CONTAINING A COMPACT SET

If $\mathbf{K} \subset \mathbb{R}^n$ is a convex body, computing the ellipsoid of minimum volume that contains $\mathbf{K}$ is a classical problem which has an optimal solution called the *Löwner-John* ellipsoid; see e.g. Barvinok [4, p. 209]. In this section we consider the following generalization:

P: *Find a homogeneous polynomial g of degree $2d$ such that its sublevel set $G := \{\mathbf{x} : g(\mathbf{x}) \leq 1\}$ contains $\mathbf{K}$ and has minimum volume among all such sublevel sets with this inclusion property.*

With $\mathbf{K} \subset \mathbb{R}^n$, let $C_{2d}(\mathbf{K}) \subset \mathbb{R}[\mathbf{x}]_{2d}$ be the convex cone of polynomials of degree at most $2d$ that are nonnegative on $\mathbf{K}$. Recall that $\mathbf{P}[\mathbf{x}]_{2d} \subset \mathbb{R}[\mathbf{x}]_{2d}$ is the convex cone of homogeneous polynomials of degree $2d$ with compact sublevel set $\{\mathbf{x} : g(\mathbf{x}) \leq 1\}$. We next show that problem $\mathbf{P}$ is a convex optimization problem:

Proposition 3.1. *The minimum volume of a sublevel set* $\{\mathbf{x} : g(\mathbf{x}) \leq 1\}$, $g \in \mathbf{P}[\mathbf{x}]_{2d}$*, that contains* $\mathbf{K} \subset \mathbb{R}^n$ *is* $\rho/\Gamma(1 + n/2d)$ *where* ρ *is the optimal value of the finite-dimensional convex optimization problem:*

$$(3.1) \qquad \mathcal{P} : \qquad \rho = \inf_{g \in \mathbf{P}[\mathbf{x}]_{2d}} \left\{ \int_{\mathbb{R}^n} \exp(-g) \, d\mathbf{x} \; : \; 1 - g \in C_{2d}(\mathbf{K}) \right\}.$$

Proof. From Theorem 2.2

$$\mathrm{vol}\left(\{\mathbf{x} : g(\mathbf{x}) \leq 1\}\right) = \frac{1}{\Gamma(1 + n/2d)} \int_{\mathbb{R}^n} \exp(-g) \, d\mathbf{x}.$$

Moreover, the sublevel set $\{\mathbf{x} : g(\mathbf{x}) \leq 1\}$ contains $\mathbf{K}$ if and only if $1 - g \in C_{2d}(\mathbf{K})$, and so $\rho/\Gamma(1 + n/2d)$ in (3.1) is the minimum value of all volumes of sublevels sets $\{\mathbf{x} : g(\mathbf{x}) \leq 1\}$, $g \in \mathbf{P}[\mathbf{x}]_{2d}$, that contain $\mathbf{K}$. Now since $g \mapsto \int_{\mathbb{R}^n} \exp(-g) d\mathbf{x}$ is strictly convex (see Corollary 2.5(b)) and $C_{2d}(\mathbf{K})$ is a convex cone, problem $\mathcal{P}$ is a finite-dimensional convex optimization problem. $\qquad \square$

We also have the following characterization of an optimal solution of $\mathcal{P}$ when it exists. Let $M(\mathbf{K})$ be the convex cone of finite Borel measures on $\mathbf{K}$.

Theorem 3.2. *Let* $\mathbf{K} \subset \mathbb{R}^n$ *be compact and consider the convex optimization problem* $\mathcal{P}$ *in (3.1).*
   (a) *Suppose that* $g^* \in \mathbf{P}[\mathbf{x}]_{2d}$ *is an optimal solution of* $\mathcal{P}$*. Then there exists* $\mu^* \in M(\mathbf{K})$ *such that*

$$(3.2) \quad \int_{\mathbb{R}^n} \mathbf{x}^\alpha \exp(-g^*) d\mathbf{x} = \int_{\mathbf{K}} \mathbf{x}^\alpha \, d\mu^*, \quad \forall |\alpha| = 2d; \quad \int_{\mathbf{K}} (1 - g^*) \, d\mu^* = 0.$$

In particular, μ^* *is supported on the real variety* $V := \{\mathbf{x} \in \mathbf{K} : g(\mathbf{x}) = 1\}$ *and in fact,* μ^* *can be substituted with another measure* $\nu^* \in M(\mathbf{K})$ *supported on at most* $\binom{n+2d-1}{2d} + 1$ *points of* V*.*
   (b) *Conversely, if* $g^* \in \mathbf{P}[\mathbf{x}]_{2d}$ *and* $\mu^* \in M(\mathbf{K})$ *satisfy (3.2) then* g^* *is an optimal solution of* $\mathcal{P}$*.*

Proof. (a) We may and will consider g^* as an element of $\mathbb{R}[\mathbf{x}]_{2d}$ with $g_\beta^* = 0$ whenever $|\beta| < 2d$. As $\mathbf{K}$ is compact, there exists $\theta \in \mathbf{P}[\mathbf{x}]_{2d}$ such that $1 - \theta \in \mathrm{int}\, C_{2d}(\mathbf{K})$, i.e., Slater's condition holds for $\mathcal{P}$. Indeed, choose $\theta := M^{-1}\|\mathbf{x}\|^{2d}$ 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 $\mathbb{R}[\mathbf{x}]_{2d}$), there exists $\epsilon > 0$ such that for every $h \in B(\theta, \epsilon)(:= \{h \in \mathbb{R}[\mathbf{x}]_{2d} : \|\theta - h\|_1 < \epsilon\})$, $1 - h > 0$ on $\mathbf{K}$.

Therefore, the optimal solution g^* satisfies the KKT-optimality conditions, which read:

$$(3.3) \qquad \int_{\mathbb{R}^n} \mathbf{x}^\alpha \exp(-g^*)\, d\mathbf{x} = y^*_\alpha, \quad \forall |\alpha| = 2d; \quad \langle 1 - g^*, \mathbf{y}^* \rangle = 0,$$

for some $\mathbf{y}^* = (y^*_\alpha)$, $\alpha \in \mathbb{N}^n_{2d}$, an element of the dual cone $C_{2d}(\mathbf{K})^* \subset \mathbb{R}^{s(2d)}$ of $C_{2d}(\mathbf{K})$. By Lemma 4.1 in §4.3,

$$C_{2d}(\mathbf{K})^* = \{ \mathbf{y} \in \mathbb{R}^{s(2d)} \ : \ \exists \mu^* \in M(\mathbf{K}) \text{ s.t. } y_\alpha = \int_{\mathbf{K}} \mathbf{x}^\alpha\, d\mu^*, \ \alpha \in \mathbb{N}^n_{2d} \},$$

and so (3.2) is just (3.3) restated in terms of μ^*. Finally, the last statement follows from Anastassiou [1, Theorem 2.1.1, p. 39] applied to the $\binom{n+2d-1}{2d}$ equality constraints of (3.2).

(b) As Slater's condition holds for $\mathcal{P}$, the KKT-optimality conditions (3.2) are sufficient to ensure that g^* is an optimal solution on $\mathcal{P}$. $\qquad \square$

Theorem 3.2 states that when (3.2) holds, there is an optimal solution g^* to $\mathcal{P}$ such that $g(\mathbf{x}) = 1$ on (at most) $\binom{n+2d-1}{2d} + 1$ points of $\mathbf{K}$, the analogue for $d > 1$ of the well-known property of the Löwner-John ellipsoid in the case $d = 1$.

Even though being convex and finite-dimensional, $\mathcal{P}$ is by no means easy to solve because there is no simple and computationally tractable way of describing the convex cone $C_{2d}(\mathbf{K})$. However, there are cases where one may provide a sequence of inner or outer approximations that both converge to $C_{2d}(\mathbf{K})$. One such case is when one knows all moments of a finite Borel measure whose support is exactly $\mathbf{K}$, and another case is when $\mathbf{K} = \{ \mathbf{x} : g_k(\mathbf{x}) \geq 0, \ k = 1, \dots, m \}$ for some polynomials $(g_k) \subset \mathbb{R}[\mathbf{x}]$, i.e., $\mathbf{K}$ is a compact basic semi-algebraic set.

3.1. **Lower bounds via inner approximations.** Suppose that one knows all moments $\mathbf{z} = (z_\alpha)$, $\alpha \in \mathbb{N}^n$, of a finite Borel measure μ on $\mathbf{K}$, i.e.,

$$z_\alpha = \int_{\mathbf{K}} \mathbf{x}^\alpha\, d\mu(\mathbf{x}), \qquad \forall \alpha \in \mathbb{N}^n,$$

whose support is exactly $\mathbf{K}$. For every $k \in \mathbb{N}$ and $p \in \mathbb{R}[\mathbf{x}]$, let $\mathbf{M}_k(p, \mathbf{z})$ be the localizing matrix with respect to the polynomial p and the moment sequence $\mathbf{z}$, that is, $\mathbf{M}_k(p, \mathbf{z})$ is the $s(k) \times s(k)$ real symmetric matrix with rows and columns indexed in the canonical basis $(\mathbf{x}^\alpha)$, $\alpha \in \mathbb{N}^n_k$, of $\mathbb{R}[\mathbf{x}]_k$, and with entries

$$\begin{aligned}
\mathbf{M}_k(p, \mathbf{z})(\alpha, \beta) &= \int_{\mathbf{K}} p(\mathbf{x})\, \mathbf{x}^{\alpha+\beta}\, d\mu(\mathbf{x}), \qquad \forall \alpha, \beta \in \mathbb{N}^n_k \\
&= \sum_\gamma p_\gamma\, z_{\alpha+\beta+\gamma}
\end{aligned}$$

(when $p(\mathbf{x}) = \sum_\gamma p_\gamma \mathbf{x}^\gamma$). We recall the following result.

Lemma 3.3. *([15, Theorem 3.2]) Let $\mathbf{K} \subset \mathbb{R}^n$ be compact and let μ be a finite Borel measure with support $\mathbf{K}$ and with moments $\mathbf{z} = (z_\alpha)$, $\alpha \in \mathbb{N}^n$. Then $p \in \mathbb{R}[\mathbf{x}]$ is nonnegative on $\mathbf{K}$ if and only if $\mathbf{M}_k(p, \mathbf{z}) \succeq 0$ for every $k = 0, 1, \ldots$.*

In view of Lemma 3.3, a natural idea is to relax the "difficult" constraint $1 - g \in C_{2d}(\mathbf{K})$ in (3.1) to $\mathbf{M}_k(1 - g, \mathbf{z}) \succeq 0$ for fixed k, and then let $k \to \infty$. Indeed, for every fixed k, the latter is much easier to handle as it defines a spectrahedron[3] $\Delta_k \subset \mathbb{R}^{\ell(2d)}$ on the coefficients of the homogeneous polynomial g. And so one obtains a hierarchy of convex relaxations of $\mathcal{P}$ by minimizing the (strictly) convex function $g \mapsto \int \exp(-g)$ on the spectrahedra Δ_k, $k \in \mathbb{N}$, which yields a monotone nondecreasing sequence of *lower bounds* $\rho_k \leq \rho$, $k \in \mathbb{N}$, on ρ. Of course the larger k (i.e., the larger the size of the localizing matrix $\mathbf{M}_k(1 - g, \mathbf{z})$) the better the lower bound ρ_k (but also the harder the problem).

Theorem 3.4. *Let $\mathbf{K} \subset \mathbb{R}^n$ be compact with nonempty interior and consider the finite-dimensional convex optimization problem:*

$$(3.4) \qquad \mathcal{P}_k: \qquad \rho_k = \inf_{g \in \mathbf{P}[\mathbf{x}]_{2d}} \left\{ \int_{\mathbb{R}^n} \exp(-g) \, d\mathbf{x} \; : \; \mathbf{M}_k(1 - g, \mathbf{z}) \succeq 0 \right\}.$$

The sequence (ρ_k), $k \in \mathbb{N}$, is monotone nondecreasing with $\rho_k \leq \rho$ and:

(a) *If $\mathcal{P}_k$ has an optimal solution $g_k^* \in \mathbf{P}[\mathbf{x}]_{2d}$ then there exists a SOS polynomial $\sigma_k^* \in \Sigma[\mathbf{x}]_k$ such that*

$$(3.5) \qquad \int_{\mathbb{R}^n} \mathbf{x}^\alpha \exp(-g_k^*) \, d\mathbf{x} \;\; = \;\; \int_{\mathbf{K}} \mathbf{x}^\alpha \, \sigma_k^* \, d\mu, \qquad \forall |\alpha| = 2d.$$

$$(3.6) \qquad \int_{\mathbf{K}} (1 - g_k^*) \, \sigma_k^* \, d\mu \;\; = \;\; 0,$$

and $\rho_k = \frac{2d}{n} \int_{\mathbf{K}} \sigma_k^* d\mu$.

(b) *Conversely if (3.5)-(3.6) holds for some $g_k^* \in \mathbf{P}[\mathbf{x}]_{2d}$ and some $\sigma_k^* \in \Sigma[\mathbf{x}]_k$ then g_k^* is an optimal solution of $\mathcal{P}_k$.*

(c) *If in addition,*

$$\sup_k \sup_{|\alpha|=2d} |g_{k\alpha}^*| \; < \; M,$$

for some $M > 0$, then $\rho_k \to \rho$ and $\mathcal{P}$ has an optimal solution g^*.

For a proof see §4.1. So problem $\mathcal{P}_k$ amounts to minimize a strictly convex function on a spectrahedron of $\mathbb{R}^{\ell(2d)}$. For instance, one may use *interior point* methods and minimize the standard log-barrier function

$$g \mapsto \phi_\nu(g) := \int_{\mathbb{R}^n} \exp(-g) d\mathbf{x} - \frac{1}{\nu} \log \det(\mathbf{M}_k(1 - g, \mathbf{z})),$$

---

[3]A spectrahedron is a convex set that can be formed by intersecting the cone of positive semidefinite matrices $\mathcal{S}_n$ with a linear affine subspace.

with parameter ν, and let $\nu \to \infty$. For more details on log-barrier methods, the reader is referred to e.g. Wright [25].

3.2. Upper bounds via outer approximations.

Let $\mathbf{K} \subset \mathbb{R}^n$ be the compact basic semi-algebraic set defined by

$$\mathbf{K} = \{\mathbf{x} \: : \: u_j(\mathbf{x}) \geq 0, \quad j = 1, \ldots, m\},$$

for some polynomials $(u_j) \subset \mathbb{R}[\mathbf{x}]$. As $\mathbf{K}$ is compact, with no loss of generality we may and will suppose that $u_1(\mathbf{x}) = M - \|\mathbf{x}\|^2$ for some M large enough. Let

$$Q^* = \{\sum_{j=0}^m \sigma_j u_j \: : \: u_j \in \Sigma[\mathbf{x}], \: j = 0, \ldots, m\},$$

be the quadratic module of $\mathbb{R}[\mathbf{x}]$ generated by the u_j's (with $u_0 = 1$). Q^* is Archimedean because the quadratic polynomial $\mathbf{x} \mapsto M - \|\mathbf{x}\|^2$ belongs to Q^*. If one defines

$$Q_k^* := \{\sum_{j=0}^m \sigma_j u_j \: : \: u_j \in \Sigma[\mathbf{x}]; \quad \deg \sigma_j u_j \leq 2k, \: j = 0, \ldots, m\}, \quad k \in \mathbb{N},$$

(a subset of Q^* with a degree bound on the SOS weights σ_j) then this time one may replace the "difficult" constraint $1 - g \in C_{2d}(\mathbf{K})$ by the stronger constraint $1 - g \in Q_k^*$ for fixed $k \in \mathbb{N}$.

The convex cone Q_k^* is the dual cone of the closed pointed convex cone

$$(3.7) \quad Q_k := \{\mathbf{y} \in \mathbb{R}^{s(2k)} \: : \: \mathbf{M}_k(\mathbf{y}) \succeq 0, \quad \mathbf{M}_{k-v_j}(u_j, \mathbf{y}) \succeq 0, \: j = 1, \ldots, m\},$$

where $\mathbf{M}_k(\mathbf{y})$ (resp. $\mathbf{M}_{k-v_j}(u_j, \mathbf{y})$) is the moment matrix associated with the sequence $\mathbf{y}$ (resp. the localizing matrix associated with $\mathbf{y}$ and the polynomial u_j). Hence, Q_k^* has a nonempty interior; see e.g. Rockafellar [21].

Then solving the problem

$$(3.8) \qquad \mathcal{P}_k' : \quad \inf_{g \in \mathbf{P}[\mathbf{x}]_{2d}} \left\{ \int_{\mathbb{R}^n} \exp(-g) d\mathbf{x} \: : \: 1 - g \in Q_k^* \right\}, \quad k \in \mathbb{N},$$

$k \geq d$, now provides a monotone sequence of *upper bounds* $\rho_k' \geq \rho$, $k \in \mathbb{N}$.

Theorem 3.5. *Let $\mathbf{K} \subset \mathbb{R}^n$ be compact with nonempty interior and consider the finite-dimensional convex optimization problem (3.8), $k \geq d$.*

(a) The sequence (ρ_k'), $d \leq k \in \mathbb{N}$, is monotone nonincreasing with $\rho_k' \to \rho$ as $k \to \infty$. Moreover there exists an optimal solution $g_k^* \in \mathbf{P}[\mathbf{x}]_{2d}$.

In addition, assume that there exists $g_0 \in \mathbf{P}[\mathbf{x}]_{2d}$ such that $1 - g_0 \in \mathrm{int}\, Q_{k_0}^*$ for some $k_0 \geq d$. Then:

(b) If $k \geq k_0$ and $g_k^* \in \mathbf{P}[\mathbf{x}]_{2d}$ is an optimal solution of $\mathcal{P}_k'$, there exists a vector $\mathbf{y}^k \in Q_k$ such that:

$$(3.9) \qquad 0 = \langle 1 - g_k^*, \mathbf{y}^k \rangle; \quad y_\alpha^k = \int_{\mathbb{R}^n} \mathbf{x}^\alpha \exp(-g_k^*) \, d\mathbf{x}, \quad \forall |\alpha| = 2d.$$

(c) *Conversely, if $k \geq k_0$ and $(\mathbf{y}^k, 1 - g_k^*) \in Q_k \times Q_k^*$ satisfy (3.9), then g_k^* is an optimal solution of $\mathcal{P}_k'$.*

(d) *Let $(g_k^*) \subset \mathbf{P}[\mathbf{x}]_{2d}$, $k \geq k_0$, be a sequence of optimal solutions. If*

$$\sup_k \ \sup_{|\alpha|=2d} |g_{k\alpha}^*| \ < \ N,$$

for some $N > 0$, then every accumulation point g^* of the sequence (g_k^*), $k \in \mathbb{N}$, is an optimal solution of $\mathcal{P}$.

For proof see §4.2.

4. PROOFS

4.1. **Proof of Theorem 3.4.**

Proof. (a) That the sequence (ρ_k) is monotone non decreasing is straightforward since the constraints of $\mathcal{P}_k$ are more and more restrictive as k increases. Next, as $\mathbf{K}$ is compact there exists $M, \delta > 0$ such that $M - \|\mathbf{x}\|^{2d} > \delta$ for all $\mathbf{x} \in \mathbf{K}$. With $g_0 \in \Sigma[\mathbf{x}]_d$ being the polynomial $\mathbf{x} \mapsto M^{-1}\|\mathbf{x}\|^{2d}$, one has $1 - g_0(\mathbf{x}) > \delta/M$ for all $\mathbf{x} \in \mathbf{K}$ and so $\mathbf{M}_k(1 - g_0, \mathbf{z}) \succ 0$. Indeed, for every $0 \neq f \in \mathbb{R}[\mathbf{x}]_k$,

$$\langle \mathbf{f}, \mathbf{M}_k(1 - g_0, \mathbf{z})\, \mathbf{f} \rangle \ = \ \int_{\mathbf{K}} f^2(1 - g_0)\, d\mu \ \geq \ \frac{\delta}{M} \int_{\mathbf{K}} f^2 d\mu \ > \ 0.$$

where the last inequality is because $\mathbf{K}$ has nonempty interior and $\operatorname{supp} \mu = \mathbf{K}$. Observe also that $g_0 \in \mathbf{P}[\mathbf{x}]_{2d}$ and so g_0 is a strictly feasible solution of $\mathcal{P}_k$, that is, Slater's condition[4] holds for $\mathcal{P}_k$. So the Karush-Kuhn-Tucker (KKT) optimality conditions at a point g_k^* are necessary and sufficient for g_k^* to be a (global) minimizer of $\mathcal{P}_k$. Therefore, there exists $0 \preceq \Delta \in \mathcal{S}_{s(k)}$ such that:

$$(4.1) \qquad \int_{\mathbb{R}^n} -\mathbf{x}^\alpha \exp(-g_k^*)\, d\mathbf{x} + \langle \Delta, \mathbf{M}_k(\mathbf{x}^\alpha, \mathbf{z}) \rangle \ = \ 0, \qquad \forall |\alpha| = 2d,$$

and

$$(4.2) \qquad\qquad \langle \mathbf{M}_d(1 - g_k^*, \mathbf{z}), \Delta \rangle \ = \ 0,$$

where one has written $\mathbf{M}_k(g_k^* \mathbf{z}) = \sum_\alpha g_{k\alpha}^* \mathbf{M}_k(\mathbf{x}^\alpha, \mathbf{z})$ (with $\mathbf{M}_k(\mathbf{x}^\alpha, \mathbf{z})$ being the localizing matrix with respect to the polynomial $\mathbf{x}^\alpha$ and the moment sequence $\mathbf{z}$). From its spectral decomposition, $\Delta = \sum_j q_j q_j^T$ for some vectors (q_j), which yields

$$\langle \Delta, \mathbf{M}_k(\mathbf{x}^\alpha, \mathbf{z}) \rangle \ = \ \int_{\mathbf{K}} \sigma_k^*(\mathbf{x})\, \mathbf{x}^\alpha\, d\mu, \qquad \forall |\alpha| = 2d,$$

---

[4]For a convex optimization $\inf_{\mathbf{x}}\{f(\mathbf{x}) \ : \ h_k(\mathbf{x}) \geq 0, \ k = 1, \ldots, m\}$ Slater's condition holds if there exists $\mathbf{x}_0$ such that $h_k(\mathbf{x}_0) > 0$ for all k. In this case the Karush-Kuhn-Tucker (KKT) optimality conditions are necessary and sufficient.

where $\sigma_k^* = \sum_j (q_j^T \mathbf{v}_k(\mathbf{x}))^2 \in \Sigma_k$ is a SOS polynomial. And so (4.1) yields (3.5), and the complementary condition (4.2) yields (3.6). Next, multiplying both sides of (4.1) with $g_{k\alpha}^*$ and summing up yields:

$$\int_{\mathbb{R}^n} g_k^* \exp(-g_k^*) \, d\mathbf{x} \left(= \frac{n}{2d} \int_{\mathbb{R}^n} \exp(-g_k^*) \, d\mathbf{x} \right) = \langle \Delta, \mathbf{M}_k(g_k^*, \mathbf{z}) \rangle$$
$$= \int_{\mathbf{K}} \sigma_k^* g_k^* \, d\mu,$$
$$= \int_{\mathbf{K}} \sigma_k^* \, d\mu \quad [\text{by (4.2)}]$$

(b) The converse is because under Slater's condition, the KKT optimality conditions are also sufficient for g_k^* to be a minimizer.

(c) Write $g_k^*(\mathbf{x}) = \sum_{|\alpha|=2d} g_{k\alpha}^* \mathbf{x}^\alpha$. As $\sup_k \sup_{|\alpha|=2d} |g_{k\alpha}^*| < M$, there exists a subsequence (k_i), $i \in \mathbb{N}$, and a vector $g^* \in \mathbb{R}^{\ell(2d)}$ such that for every $\alpha \in \mathbb{N}^n$ with $|\alpha| = 2d$, $g_{k_i\alpha}^* \to g_\alpha^*$ as $i \to \infty$.

Notice that $\mathbf{x} \mapsto g^*(\mathbf{x}) = \limsup_{i\to\infty} g_{k_i}^*(\mathbf{x}) = \liminf_{i\to\infty} g_{k_i}^*(\mathbf{x})$. Next, as $\rho \geq \rho_k$ for every k, and $\exp(-g_k^*) \geq 0$, by Fatou's Lemma (see e.g. Ash [2])

$$\rho \geq \liminf_{i\to\infty} \rho_{k_i} = \liminf_{i\to\infty} \int_{\mathbb{R}^n} \exp(-g_{k_i}^*) \, d\mathbf{x}$$
$$\geq \int_{\mathbb{R}^n} \liminf_{i\to\infty} \exp(-g_{k_i}^*) \, d\mathbf{x}$$
$$= \int_{\mathbb{R}^n} \exp(-\limsup_i g_{k_i}^*) \, d\mathbf{x} = \int_{\mathbb{R}^n} \exp(-g^*) \, d\mathbf{x}.$$

On the other hand, observe that for every $k \in \mathbb{N}$, $\mathbf{M}_k(1 - g_k^*, \mathbf{z}) \succeq 0$ implies $\mathbf{M}_j(1 - g_k^*, \mathbf{z}) \succeq 0$ for all $j \leq k$. Hence, let j be fixed arbitrary so that $\mathbf{M}_j(1 - g_k^*, \mathbf{z}) \succeq 0$, for all sufficiently large k. For every $f \in \mathbb{R}[\mathbf{x}]_j$, $f^2 g_k$ is uniformly bounded on $\mathbf{K}$ and so by Fatou's Lemma

$$0 \leq \limsup_{i\to\infty} \int_{\mathbf{K}} f^2 (1 - g_{k_i}^*) \, d\mu = -\liminf_{i\to\infty} \int_{\mathbf{K}} f^2 (g_{k_i}^* - 1) \, d\mu$$
$$\leq \int_{\mathbf{K}} -\liminf_{i\to\infty} f^2 (g_{k_i}^* - 1) \, d\mu$$
$$= \int_{\mathbf{K}} f^2 (1 - g^*) \, d\mu.$$

As $f \in \mathbb{R}[\mathbf{x}]_j$ was arbitrary, $\mathbf{M}_j(1 - g^*, \mathbf{z}) \succeq 0$; and as j was arbitrary, $\mathbf{M}_j(1 - g^*, \mathbf{z}) \succeq 0$, for all $j = 0, 1, \dots$ But by Lemma 3.3, this implies $1 - g^* \geq 0$ on $\mathbf{K}$, and so g^* is a feasible solution for $\mathcal{P}$. Combining with $\rho \geq \int_{\mathbb{R}^n} \exp(-g^*) d\mathbf{x}$ yields that g^* is an optimal solution of $\mathcal{P}$. $\square$

4.2. Proof of Theorem 3.5.

Proof. (a) The sequence (ρ_k') is monotone non increasing because $Q_k^* \subset Q_{k+1}^*$ for every $k \in \mathbb{N}$. Next, let $\epsilon > 0$ be fixed, and let $g \in \mathbf{P}[\mathbf{x}]_{2d}$ be such that

$\rho \leq \int_{\mathbb{R}^n} \exp(-g) d\mathbf{x} \leq \rho + \epsilon$. Observe that $\tilde{g} := (1 + \epsilon)^{-1} g \in \mathbf{P}[\mathbf{x}]_{2d}$ and

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

and $1 - \tilde{g} = \frac{1 + \epsilon - g}{1 + \epsilon} > 0$ on $\mathbf{K}$. Therefore, as Q^* is Archimedean, by Putinar's Positivstellensatz [20], $1 - \tilde{g} \in Q^*$, that is, $1 - \tilde{g} \in Q_{k_1}^*$ for some k_1. Hence $\tilde{g}$ is a feasible solution of $\mathcal{P}_{k_1}$ which implies $\rho \leq \rho'_{k_1} \leq (1 + \epsilon)^{n/d} (\rho + \epsilon)$. As the sequence is monotone and $\epsilon > 0$ was arbitrary, we obtain the desired convergence $\rho'_k \to \rho$.

Next, let $\mathbf{w} = (w_\alpha)$, $\alpha \in \mathbb{N}^n$, be the sequence of moment of the Lebesgue measure on $\mathbf{K}$, and let $g \in \mathbb{R}[\mathbf{x}]_{2d}$ be an arbitrary feasible solution. Let $\mathbf{M}_k(\mathbf{w})$ (resp. $\mathbf{M}_{k-v_j}(u_j \, \mathbf{w})$) be the moment (resp. localizing) matrix associated with $\mathbf{w}$ and u_j. Hence, $\mathbf{M}_d(\mathbf{w}) \succ 0$ and $\mathbf{M}_{d-v_j}(u_j \, \mathbf{w}) \succ 0$, $j = 1, \ldots, m$, because $\mathbf{K}$ has nonempty interior. The constraint $1 - g \in Q_k(u)$ which reads $(1 - g) = \sigma_0 + \sum_{j=1}^m \sigma_j u_j$ for some SOS polynomials (σ_j) has the equivalent form

$$
(4.3) \qquad (1 - g)_\alpha = \langle \mathbf{X}_0, \mathbf{B}_\alpha \rangle + \sum_{j=1}^m \langle \mathbf{X}_j, \mathbf{C}_{j\alpha} \rangle, \quad \forall \alpha \in \mathbb{N}_{2k}^n,
$$

for some appropriate real symmetric $s(2(k - v_j)) \times s(2(k - v_j))$ matrices $\mathbf{X}_j \succeq 0$, (and where $\mathbf{B}_\alpha, \mathbf{C}_{j\alpha}$ are given real symmetric matrices). The vector of coefficients of the SOS polynomial $\sigma_j \in \Sigma[\mathbf{x}]$ is obtained from the eigenvectors of $\mathbf{X}_j$, $j = 0, \ldots, m$. For more details see e.g. [11]. Multiplying (4.3) by w_α and summing up yield:

$$
\begin{aligned}
\int_{\mathbf{K}} (1 - g) \, d\mathbf{x} &= \langle \mathbf{X}_0, \sum_\alpha \mathbf{B}_\alpha w_\alpha \rangle + \sum_{j=1}^m \langle \mathbf{X}_j, \sum_\alpha \mathbf{C}_{j\alpha} w_\alpha \rangle \\
&= \langle \mathbf{X}_0, \mathbf{M}_k(\mathbf{w}) \rangle + \sum_{j=1}^m \langle \mathbf{X}_j, \mathbf{M}_{k-v_j}(u_j \, \mathbf{w}) \rangle
\end{aligned}
$$

Observe that every feasible solution $g \in \mathbb{R}[\mathbf{x}]_{2d}$ is nonnegative otherwise $\int \exp(-g) dx$ is not bounded. And so $\exp(-g) \geq 1 - g$ on $\mathbf{K}$ because $g \geq 0$. Therefore, for any minimizing sequence $(g_n, \mathbf{X}_j^n)$ of (3.8),

$$
\int_{\mathbf{K}} (1 - g_n) \, d\mathbf{x} \leq \int_{\mathbf{K}} \exp(-g_n) \, d\mathbf{x} \leq \int_{\mathbb{R}^n} \exp(-g_n) \, d\mathbf{x} \leq \int_{\mathbb{R}^n} \exp(-g_0) \, d\mathbf{x}.
$$

And so

$$
\sup_n \left\{ \langle \mathbf{X}_0^n, \mathbf{M}_k(\mathbf{w}) \rangle + \sum_{j=1}^m \langle \mathbf{X}_j^n, \mathbf{M}_{k-v_j}(u_j \, \mathbf{w}) \rangle \right\} \leq \int_{\mathbb{R}^n} \exp(-g_0) \, d\mathbf{x}.
$$

As $\mathbf{M}_k(\mathbf{w}), \mathbf{M}_{k-v_j}(u_j \, \mathbf{w}) \succ 0$ and $\mathbf{X}_j^n \succeq 0$, $j = 1, \ldots, m$, all matrices $\mathbf{X}_j^n$ are uniformly bounded. Hence one may extract a subsequence (n_ℓ) such that

$\mathbf{X}_j^{n_\ell} \to \mathbf{X}_j^* \succeq 0$, $j = 0, \dots, n$, as $\ell \to \infty$. And so for every $\alpha \in \mathbb{N}_{2k}^n$,

$$\langle \mathbf{X}_0^*, \mathbf{B}_\alpha \rangle + \sum_{j=1}^m \langle \mathbf{X}_j^*, \mathbf{C}_{j\alpha} \rangle = \lim_{\ell \to \infty} (1 - g_{n_\ell})_\alpha =: (1 - g^*)_\alpha,$$

for some homogeneous polynomial $g^* \in \mathbb{R}[\mathbf{x}]_{2d}$ (as all coefficients g_α^* with $|\alpha| \neq 2d$ vanish). In other words $(g_{n_\ell})_\alpha \to g_\alpha^*$ as $\ell \to \infty$, for all $\alpha \in \mathbb{N}_{2k}^n$. As $(1 - g_{n_\ell}) \geq 0$ on $\mathbf{K}$ one also has $(1 - g^*) \geq 0$ on $\mathbf{K}$. Finally, since we also have the pointwise convergence $g_{n_\ell}(\mathbf{x}) \to g^*(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$, invoking Fatou's lemma yields

$$\rho_k' = \lim_{\ell \to \infty} \int_{\mathbb{R}^n} \exp(-g_{n_\ell})\, d\mathbf{x} \geq \int_{\mathbb{R}^n} \liminf_{\ell \to \infty} \exp(-g_{n_\ell})\, d\mathbf{x}$$

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

which proves that g^* is an optimal solution of (3.8).

(b) $1 - g_0 \in \text{int}\, Q_k^*$ for all $k \geq k_0$ because $1 - g_0 \in \text{int}\, Q_{k_0}^*$ and $Q_k^* \supset Q_{k_0}^*$. Hence Slater's condition holds for $\mathcal{P}_k'$ whenever $k \geq k_0$. Therefore, if g_k^* is an optimal solution of $\mathcal{P}_k$, there exists $\mathbf{y}^k \in Q_k$ such that the KKT-optimality condition hold, which yields (3.9).

(c) Follows from the fact that under Slater's condition, the KKT-optimality conditions are sufficient for g_k^* to be an optimal solution of $\mathcal{P}_k'$.

(d) If $\sup_k \sup_{|\alpha|=2d} |g_{k\alpha}^*| < N$, let g^* be an accumulation point, i.e., a limit point of some subsequence $(g_{k_i}^*)$, $i \in \mathbb{N}$, i.e., such that for every α with $|\alpha| = 2d$, $g_{k_i\alpha}^* \to g_\alpha^*$ as $i \to \infty$. Let $g^* \in \mathbb{R}[\mathbf{x}]_{2d}$ be the homogeneous polynomial with coefficients g_α^*, $|\alpha| = 2d$. For every $\mathbf{x} \in \mathbf{K}$, $(1 - g_{k_i}(\mathbf{x})) \geq 0$ because $1 - g_{k_i}^* \in Q_{k_i}^*$ for every i. Therefore, with $\mathbf{x} \in \mathbf{K}$ fixed,

$$0 \leq 1 - \lim_{i \to \infty} g_{k_i}^*(\mathbf{x}) = 1 - g^*(\mathbf{x}),$$

and so, as $\mathbf{x} \in \mathbf{K}$ is arbitrary, $1 - g^* \in C_{2d}(\mathbf{K})$, which implies that $1 - g^*$ is a feasible solution of $\mathcal{P}$. Next, from (a) and by Fatou's Lemma,

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

$$\geq \int_{\mathbb{R}^n} \liminf_{i \to \infty} \exp(-g_{k_i}^*)\, d\mathbf{x}$$

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

and so as g^* is feasible for $\mathcal{P}$, it is an optimal solution of $\mathcal{P}$. $\qquad\square$

4.3. **The dual cone of $C_d(\mathbf{K})$.** Recall that $M(\mathbf{K})$ is the space of finite Borel measures on $\mathbf{K}$. Let $C_\infty(\mathbf{K}) \subset \mathbb{R}[\mathbf{x}]$ be the space of polynomials nonnegative on $\mathbf{K} \subseteq \mathbb{R}^n$ and let $\Delta_\infty \subset \mathbb{R}[\mathbf{x}]^*$ be the set:

$$(4.4) \qquad \Delta_\infty := \left\{ \left(\int_\mathbf{K} \mathbf{x}^\alpha\, d\phi \right), \alpha \in \mathbb{N}^n : \phi \in M(\mathbf{K}) \right\}.$$

It is known that when $\mathbf{K} \subseteq \mathbb{R}^n$ is closed then $\Delta_\infty^* = C_\infty(\mathbf{K})$ and $\Delta_\infty = C_\infty(\mathbf{K})^*$. See e.g. the proof in Laurent [16, Prop. 4.5] for $\mathbf{K} = \mathbb{R}^n$, which also works for any closed set $\mathbf{K} \subset \mathbb{R}^n$. The following result is a truncated version when $\mathbf{K}$ is compact.

Lemma 4.1. *Let* $\mathbf{K} \subset \mathbb{R}^n$ *be compact. For every* $d \in \mathbb{N}$*, the dual* $C_d(\mathbf{K})^*$ *of* $C_d(\mathbf{K})$ *is the set:*

$$(4.5) \qquad \Delta_d := \left\{ \left(\int_{\mathbf{K}} \mathbf{x}^\alpha \, d\phi \right), \, \alpha \in \mathbb{N}_d^n \, : \, \phi \in M(\mathbf{K}) \right\}.$$

Proof. For every $\mathbf{y} = (y_\alpha) \in \Delta_d$ and $f \in C_d(\mathbf{K})$ with coefficient vector $\mathbf{f} \in \mathbb{R}^{s(d)}$:

$$(4.6) \qquad \langle \mathbf{y}, \mathbf{f} \rangle = \sum_{\alpha \in \mathbb{N}_d^n} f_\alpha y_\alpha = \sum_{\alpha \in \mathbb{N}_d^n} \int_{\mathbf{K}} f_\alpha \mathbf{x}^\alpha \, d\phi = \int_{\mathbf{K}} f \, d\phi \geq 0.$$

Since (4.6) holds for all $f \in C_d(\mathbf{K})$ and $\mathbf{y} \in \Delta_d$, then necessarily $\Delta_d \subseteq C_d(\mathbf{K})^*$ and similarly, $C_d(\mathbf{K}) \subseteq \Delta_d^*$. Next,

$$\begin{aligned} \Delta_d^* &= \left\{ \mathbf{f} \in \mathbb{R}^{s(d)} \, : \, \langle \mathbf{f}, \mathbf{y} \rangle \geq 0 \quad \forall \mathbf{y} \in \Delta_d \right\} \\ &= \left\{ f \in \mathbb{R}[\mathbf{x}]_d \, : \, \int_{\mathbf{K}} f \, d\phi \geq 0 \quad \forall \phi \in M(\mathbf{K}) \right\} \\ &\Rightarrow \Delta_d^* \subseteq C_d(\mathbf{K}), \end{aligned}$$

and so $\Delta_d^* = C_d(\mathbf{K})$. Hence the result follows if one proves that Δ_d is closed, because then $C_d(\mathbf{K})^* = (\Delta_d^*)^* = \Delta_d$, the desired result. So let $(\mathbf{y}^k) \subset \Delta_d$, $k \in \mathbb{N}$, with $\mathbf{y}^k \to \mathbf{y}$ as $k \to \infty$. Equivalently, $\int_{\mathbf{K}} \mathbf{x}^\alpha d\phi_k \to y_\alpha$ for all $\alpha \in \mathbb{N}_d^n$. In particular, the convergence $y_0^k \to y_0$ implies that the sequence of measures (ϕ_k), $k \in \mathbb{N}$, is bounded, that is, $\sup_k \phi_k(\mathbf{K}) < M$ for some $M > 0$. As $\mathbf{K}$ is compact, the unit ball of $M(\mathbf{K})$ is sequentially compact in the weak $\star$ topology $\sigma(M(\mathbf{K}), C(\mathbf{K}))$ where $C(\mathbf{K})$ is the space of continuous functions on $\mathbf{K}$. Hence there is a finite Borel measure $\phi \in M(\mathbf{K})$ and a subsequence (k_i) such that $\int_{\mathbf{K}} g \, d\phi_{k_i} \to \int_{\mathbf{K}} g \, d\phi$ as $i \to \infty$, for all $g \in C(\mathbf{K})$. In particular, for every $\alpha \in \mathbb{N}_d^n$,

$$y_\alpha = \lim_{k \to \infty} y_\alpha^k = \lim_{i \to \infty} y_\alpha^{k_i} = \lim_{i \to \infty} \int_{\mathbf{K}} \mathbf{x}^\alpha d\phi_{k_i} = \int_{\mathbf{K}} \mathbf{x}^\alpha d\phi,$$

which shows that $\mathbf{y} \in \Delta_d$, and so Δ_d is closed. $\qquad\square$

REFERENCES

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

[2] R.B. Ash. *Real Analysis and Probability*, Academic Press Inc., Boston, 1972.

[3] D. Azé, J.-B. Hiriart-Urruty. *Analyse Variationnelle et Optimisation*, CEPADUES editions, 2010.

[4] A. Barvinok. *A Course in Convexity*, American Mathematical Society, Providence, RI, 2002.

[5] D.S. Bernstein. *Matrix Mathematics*, Princeton University Press, Princeton, 2005.

[6] N. Dunford and J.T. Schwartz. *Linear Operators: Part I: General Theory*, Interscience Publishers, Inc., New York, 1958.

[7] K. Fujii. Beyond the Gaussian, Sigma **7** (2011).

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

[9] J.B. Lasserre, J-B. Hiriart-Urruty. (2002). Mathematical properties of optimization problems defined by positively homogeneous functions, J. Optim. Theor. Appl. **112** (2002), pp. 31—52.

[10] J.B. Lasserre. *Linear and Integer Programming vs Linear Integration and Counting*, Springer, New York, 2009.

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

[12] J.B. Lasserre. Integration and homogeneous functions Proc. Amer. Math. Soc. **127** (1999), pp. 813–818.

[13] J.B. Lasserre and E.S. Zeron. Solving a class of multivariable integration problems via Laplace techniques, Appl. Math. (Warsaw) **28**, pp. 391–405.

[14] J.B. Lasserre. Homogeneous functions and conjugacy, J. Convex Anal. **5** (1998), pp. 397–403.

[15] J.B. Lasserre. A new look at nonnegativity on closed sets and polynomial opimization, SIAM J. Optim. **21** (2011), pp. 864–885.

[16] M. Laurent. Sums of squares, moment matrices and optimization over polynomials, in *Emerging Applications of Algebraic Geometry*, M. Putinar and S. Sullivant (Eds.), Springer Verlag, New York, 2009

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

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

[19] C. Müller, A. Feuer, G. Goodwin. Derivative of an integral over a convex polytope, Appl. Math. Letters **24** (2011), 1120–1123

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

[21] R.T. Rockafellar. *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.

[22] S.R. Shakirov. Nonperturbative approach to finite-dimensional non-Gaussian integrals, Theor. Math. Physics **163** (2010), pp.804–812

[23] A.V. Stoyanovsky. On integral of exponent of a homogeneous polynomial. `arXiv:11030514v4`, 2011

[24] H. Wolkowicz, R. Saigal and L. Vandenberghe. *Handbook of Semidefinite Programming*, H. Wolkowicz, R. Saigal and L. Vandenberghe (Eds.), Kluwer Academic Publishers, Norwell, 2000.

[25] S. Wright. On the convergence of the Newton/log-barrier method, Math. Program. Ser. A **90** (2001), Ser. A, pp. 71–100.

LAAS-CNRS and Institute of Mathematics, University of Toulouse, 7 Avenue du Colonel Roche, 310777 Toulouse Cedex 4, France.
E-mail address: `lasserre@laas.fr`