



HAL
open science

Computing gaussian & exponential measures of semi-algebraic sets

Jean-Bernard Lasserre

► **To cite this version:**

Jean-Bernard Lasserre. Computing gaussian & exponential measures of semi-algebraic sets. *Advances in Applied Mathematics*, 2017, 91, pp.137-163. <hal-01185641v2>

HAL Id: hal-01185641

<https://hal.science/hal-01185641v2>

Submitted on 10 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

COMPUTING GAUSSIAN & EXPONENTIAL MEASURES OF SEMI-ALGEBRAIC SETS

JEAN B. LASSERRE

ABSTRACT. We provide a numerical scheme to approximate as closely as desired the Gaussian or exponential measure $\mu(\Omega)$ of (not necessarily compact) basic semi-algebraic sets $\Omega \subset \mathbb{R}^n$. We obtain two monotone (non-increasing and non-decreasing) sequences of upper and lower bounds $(\overline{\omega}_d)$, $(\underline{\omega}_d)$, $d \in \mathbb{N}$, each converging to $\mu(\Omega)$ as $d \rightarrow \infty$. For each d , computing $\overline{\omega}_d$ or $\underline{\omega}_d$ reduces to solving a semidefinite program whose size increases with d . Some preliminary (small dimension) computational experiments are encouraging and illustrate the potential of the method. The method also works for any measure whose moments are known and which satisfies Carleman's condition.

1. INTRODUCTION

Given a basic semi-algebraic set

$$(1.1) \quad \Omega = \{ \mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geq 0, j = 1, \dots, m \},$$

for some polynomials $(g_j) \subset \mathbb{R}[\mathbf{x}]$, we want to compute (or more precisely, approximate as closely as desired) $\mu(\Omega)$ for the standard Gaussian measure on \mathbb{R}^n :

$$(1.2) \quad \mu(B) = \frac{1}{(2\pi)^{n/2}} \int_B \exp\left(-\frac{1}{2}\|\mathbf{x}\|^2\right) d\mathbf{x}, \quad \forall B \in \mathcal{B}(\mathbb{R}^n),$$

and the standard exponential measure on the positive orthant \mathbb{R}_+^n :

$$(1.3) \quad \mu(B) = \frac{1}{(2\pi)^{n/2}} \int_B \exp\left(-\sum_{i=1}^n x_i\right) d\mathbf{x}, \quad \forall B \in \mathcal{B}(\mathbb{R}_+^n).$$

This problem is ‘‘canonical’’ as for a non-centered Gaussian with density $\exp(-(\mathbf{x} - m)^T \mathbf{Q}(\mathbf{x} - m))$ (for some real symmetric positive definite matrix \mathbf{Q}) or for an exponential measure with density $\exp(-\sum_i \lambda_i x_i)$ one may always reduce the problem to the above one by an appropriate change of variable. Indeed after this change of variable the new domain is again a basic semi-algebraic set of the form (1.1).

Computing $\mu(\Omega)$ has applications in Probability & Statistics where the Gaussian measure plays a central role. In full generality with sets Ω as general as (1.1), it is a difficult and challenging problem even for rectangles Ω , e.g.:

$$(1.4) \quad \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{a_1} \int_{-\infty}^{a_2} \exp(-(x^2 + y^2 - 2\rho xy)/2(1-\rho^2)) dx dy,$$

in small dimension like $n = 2$ or $n = 3$. Indeed, citing A. Genz [10]: *bivariate and trivariate probability distributions computation are needed for many statistics*

1991 *Mathematics Subject Classification.* 28A75 49Q15 13P25 90C22 90C05 65C60 65K05.

Key words and phrases. Computational geometry; Statistics; Probability; Gaussian measure; moments; semi-algebraic sets; semidefinite programming; semidefinite relaxations.

applications ... high quality algorithms for bivariate and trivariate probability distribution computations have only more recently started to become available. For instance, Genz [10] describes techniques with high accuracy results for bivariate and trivariate “rectangles” (1.4) using sophisticated techniques to integrate Plackett’s formula. Again those efficient techniques takes are very specific as they take advantage of Plackett’s formula available for (1.4). Interestingly, a (complicated) formula in closed form is provided in Chandramouli and Ranganathan [6] via the characteristic function method.

The case of ellipsoids Ω has been investigated in the pioneering work of Ruben [20], Kotz et al. [14, 15] when studying the distribution of random variables that are quadratic forms of independent normal variables. Even in small dimension it has important applications in Astronautics where for instance $\mu(\Omega)$ can model the probability of collision between two spatial vehicles and must be computed with very good accuracy; see for instance the works by Alfano [1], Chan [5], Patera [19] and the more recent [22] which combines Laplace transform techniques with the theory of D -finite functions (see e.g. Zeilberger [21] and Salvy [23]). In doing so one obtains $\mu(\Omega)$ as a series with only nonnegative terms so that its evaluation does not involve error prone cancellations. For more details the reader is referred to [1, 5, 19, 22] and the references therein. Other applications in Astronautics and for weapon evaluation require to compute integral of bivariate Normal distributions on *convex polygons*, a case treated in Didonato et al. [7] where at an intermediate step the authors also evaluate Gaussian integrals on (unbounded) angular regions.

However the techniques developed in the above cited works and in some of the references therein, are not reproducible for more general sets Ω of the form (1.1).

It should be noted that computing exactly the Lebesgue measure of a compact convex set is a very difficult problem in general (even for a convex polytope) and in fact, even approximating the volume of a polytope within some bounds is difficult; on the other hand some *probabilistic* methods can provide good estimates. For more details on the computational complexity of volume computation the interested reader is referred to the discussion in [13] and the references therein.

Contribution. The goal of this paper is to provide a systematic numerical scheme to approximate $\mu(\Omega)$ as closely as desired at the price of solving a hierarchy of semidefinite programs¹ of increasing size. To do so we use that computing $\mu(\Omega)$ is equivalent to solving a linear program on an appropriate space of measures, an instance of the Generalized Problem of Moments. We also use that *all moments* $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, of μ can be computed efficiently. In fact, the methodology that we propose is valid for *any* measure μ on \mathbb{R}^n which satisfies Carleman’s condition and whose moments can be computed. (The gaussian and exponential measures being two notable examples of such measures as their moments are available in closed form.)

These two ingredients were already used in Henrion et al. [13] to compute the Lebesgue *volume* of a *compact* basic semi-algebraic set Ω as in (1.1), by solving a hierarchy of semidefinite programs. The resulting sequence of optimal values was shown to converge to the Lebesgue volume of Ω . A technique to improve the convergence was also provided in [13] to limit the Gibbs effect present in the initial

¹A semidefinite program is a convex conic optimization problem which can be solved at arbitraryrecision (fixed in advance) efficiently, i.e., in time polynomial in its input size; for more details the interested reader is referred to e.g. [2].

and basic version of the numerical scheme. (With this technique the convergence is indeed much faster but not monotone any more.) So the novelty in this paper with respect to [13] is threefold:

- We here consider the Gaussian and exponential measures on \mathbb{R}^n and \mathbb{R}_+^n respectively, instead of the Lebesgue measure. If the latter is appropriate in applications of computational geometry (e.g. to compute “volumes”), the former are particularly interesting for applications in Probability & Statistics (especially the Gaussian measure which plays a central role). In addition, as already mentioned the method also works for *any* measure μ on \mathbb{R}^n which satisfies Carleman’s condition and whose moments can be computed. In a few words, we show how from the sole knowledge of moments of such measures μ , one may approximate as closely as desired the measure $\mu(\Omega)$ of basic closed semi-algebraic sets $\Omega \subset \mathbb{R}^n$.

- We can handle *any* basic semi-algebraic set Ω of the form (1.1) whereas [13] is restricted to compact ones. Importantly, the method provides two monotone sequences $\bar{\omega}_d$ and $\underline{\omega}_d$, $d \in \mathbb{N}$, such that $\underline{\omega}_d \leq \mu(\Omega) \leq \bar{\omega}_d$ for all d , and each sequence converges to $\mu(\Omega)$ as $d \rightarrow \infty$. Therefore at each step d the error $\bar{\omega}_d - \mu(\Omega)$ is bounded by $\epsilon_d := \bar{\omega}_d - \underline{\omega}_d$ and $\epsilon_d \rightarrow 0$ as $d \rightarrow \infty$. In our opinion this is an important feature which to the best of our knowledge is not present in other methods of the Literature, at least at this level of generality.

It is also worth noting that the method also provides a convergent numerical scheme to approximate as closely as desired any (*a priori* fixed) number of moments of the measure μ_Ω , the restriction of μ to Ω ($\mu(\Omega)$ being only the mass of μ_Ω).

- The acceleration technique described in [13] also applies to our context of the Gaussian measure on non-compact sets Ω . But again the monotone convergence is lost. Therefore we also provide another technique of independent interest to accelerate the convergence of the numerical scheme. It uses Stokes’ Theorem for integration which permits to obtain linear relations between moments of μ_Ω .

Computational remarks. As we did in [13] for the compact case and the Lebesgue measure, the problem is formulated as an instance of the *generalized problem of moments* with polynomial data and we approximate $\mu(\Omega)$ as closely as desired by solving a hierarchy of semidefinite programs of increasing size. This procedure is implemented in the software package GloptiPoly² which for modelling convenience uses the standard basis of monomials. As the monomial basis is well-known to be a source of numerical ill-conditioning, only a limited control on the output accuracy is possible and so in this case only bounds $(\bar{\omega}_d, \underline{\omega}_d)$ with $d \leq 10$ for $n = 2, 3$ are meaningful. Therefore for simple sets like rectangles (1.4) and ellipsoids (with $n = 3$), in its present form our technique does not compete in terms of accuracy with *ad-hoc* procedures like in e.g. Genz [10] (rectangles) and the recent [22] (for ellipsoids). However, in view of the growing interest for semidefinite programming and its use in many applications, it is expected that more efficient packages will be available soon. For instance the semidefinite package SDPA [9] has now been provided with a double precision variant³. Moreover a much better accuracy could be obtained if one uses other bases for polynomials than the standard monomial

²GloptiPoly [12] is a software package for solving the Generalized Problem of Moments with polynomial data.

³The SDPA-GMP, SDPA-QD and SDPA-DD versions of the standard SDPA package; see <http://sdpa.sourceforge.net/family.html#sdpa>.

basis. For instance when μ is the Gaussian measure one should rather use the basis of Hermite polynomials, orthogonal with respect to μ .

Such issues are beyond the scope of the present paper and on the other hand, for very general sets like (1.1) in relatively small dimension, and to the best of our knowledge for the first time, our method can provide relatively good bounds $\underline{\omega}_d \leq \mu(\Omega) \leq \bar{\omega}_d$ in a reasonable amount of time, and with strong theoretical guarantees as d increases.

2. MAIN RESULT

2.1. Notation and definitions. Let $\mathbb{R}[\mathbf{x}]$ be the ring of polynomials in the variables $\mathbf{x} = (x_1, \dots, x_n)$. Denote by $\mathbb{R}[\mathbf{x}]_d \subset \mathbb{R}[\mathbf{x}]$ the vector space of polynomials of degree at most d , which forms a vector space of dimension $s(d) = \binom{n+d}{d}$, with e.g., the usual canonical basis (\mathbf{x}^α) of monomials. Also, let $\mathbb{N}_d^n := \{\alpha \in \mathbb{N}^n : \sum_i \alpha_i \leq d\}$ and denote by $\Sigma[\mathbf{x}] \subset \mathbb{R}[\mathbf{x}]$ (resp. $\Sigma[\mathbf{x}]_d \subset \mathbb{R}[\mathbf{x}]_{2d}$) the space of sums of squares (SOS) polynomials (resp. SOS polynomials of degree at most $2d$). If $f \in \mathbb{R}[\mathbf{x}]_d$, write

$$f(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}_d^n} f_\alpha \mathbf{x}^\alpha \quad \left(= \sum_{\alpha \in \mathbb{N}_d^n} f_\alpha x_1^{\alpha_1} \cdots x_n^{\alpha_n} \right),$$

in the canonical basis and denote by $\mathbf{f} = (f_\alpha) \in \mathbb{R}^{s(d)}$ its vector of coefficients. Finally, let \mathcal{S}^n denote the space of $n \times n$ real symmetric matrices, with inner product $\langle \mathbf{A}, \mathbf{B} \rangle = \text{trace } \mathbf{A}\mathbf{B}$, and where the notation $\mathbf{A} \succeq 0$ (resp. $\mathbf{A} \succ 0$) stands for \mathbf{A} is positive semidefinite.

Let (\mathbf{A}_j) , $j = 0, \dots, s$, be a set of real symmetric matrices. An inequality of the form

$$(\mathbf{A}(\mathbf{x}) :=) \mathbf{A}_0 + \sum_{k=1}^s \mathbf{A}_k x_k \succeq 0, \quad \mathbf{x} \in \mathbb{R}^s,$$

is called a *Linear Matrix Inequality* (LMI) and a set of the form $\{\mathbf{x} : \mathbf{A}(\mathbf{x}) \succeq 0\}$ is the canonical form of the feasible set of semidefinite programs.

Given a real sequence $\mathbf{z} = (z_\alpha)$, $\alpha \in \mathbb{N}^n$, define the Riesz linear functional $L_{\mathbf{z}} : \mathbb{R}[\mathbf{x}] \rightarrow \mathbb{R}$ by:

$$f (= \sum_{\alpha} f_{\alpha} \mathbf{x}^{\alpha}) \mapsto L_{\mathbf{z}}(f) = \sum_{\alpha} f_{\alpha} z_{\alpha}, \quad f \in \mathbb{R}[\mathbf{x}].$$

A sequence $\mathbf{z} = (z_\alpha)$, $\alpha \in \mathbb{N}^n$, has a representing measure μ if

$$z_\alpha = \int_{\mathbb{R}^n} \mathbf{x}^\alpha d\mu, \quad \forall \alpha \in \mathbb{N}^n.$$

Moment matrix. The *moment matrix* associated with a sequence $\mathbf{z} = (z_\alpha)$, $\alpha \in \mathbb{N}^n$, is the real symmetric matrix $\mathbf{M}_d(\mathbf{z})$ with rows and columns indexed by \mathbb{N}_d^n , and whose entry (α, β) is just $z_{\alpha+\beta}$, for every $\alpha, \beta \in \mathbb{N}_d^n$. Alternatively, let $\mathbf{v}_d(\mathbf{x}) \in \mathbb{R}^{s(d)}$ be the vector (\mathbf{x}^α) , $\alpha \in \mathbb{N}_d^n$, and define the matrices $(\mathbf{B}_\alpha) \subset \mathcal{S}^{s(d)}$ by

$$(2.1) \quad \mathbf{v}_d(\mathbf{x}) \mathbf{v}_d(\mathbf{x})^T = \sum_{\alpha \in \mathbb{N}_{2d}^n} \mathbf{B}_\alpha \mathbf{x}^\alpha, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Then $\mathbf{M}_d(\mathbf{z}) = \sum_{\alpha \in \mathbb{N}_{2d}^n} z_\alpha \mathbf{B}_\alpha$. If \mathbf{z} has a representing measure μ then $\mathbf{M}_d(\mathbf{z}) \succeq 0$ because

$$\langle \mathbf{f}, \mathbf{M}_d(\mathbf{z})\mathbf{f} \rangle = \int f^2 d\mu \geq 0, \quad \forall \mathbf{f} \in \mathbb{R}^{s(d)}.$$

A measure whose all moments are finite is said to be *moment determinate* if there is no other measure with same moments. The support of a Borel measure μ on \mathbb{R}^n (denoted $\text{supp } \mu$) is the smallest closed set \mathbf{K} such that $\mu(\mathbb{R}^n \setminus \mathbf{K}) = 0$.

A sequence $\mathbf{z} = (z_\alpha)$, $\alpha \in \mathbb{N}^n$, satisfies Carleman's condition if

$$(2.2) \quad \sum_{k=1}^{\infty} L_{\mathbf{z}}(x_i^{2k})^{-1/2k} = +\infty, \quad \forall i = 1, \dots, n.$$

If a sequence $\mathbf{z} = (z_\alpha)$, $\alpha \in \mathbb{N}^n$, satisfies Carleman's condition (2.2) and $\mathbf{M}_d(\mathbf{z}) \succeq 0$ for all $d = 0, 1, \dots$, then \mathbf{z} has a representing measure on \mathbb{R}^n which is moment determinate; see e.g. [18, Proposition 3.5]. In particular a sufficient condition for a measure μ to satisfy Carleman's condition is that $\int \exp(c \sum_i |x_i|) d\mu < \infty$ for some $c > 0$.

Localizing matrix. With \mathbf{z} as above and $g \in \mathbb{R}[\mathbf{x}]$ (with $g(\mathbf{x}) = \sum_{\gamma} g_{\gamma} \mathbf{x}^{\gamma}$), the *localizing matrix* associated with \mathbf{z} and g is the real symmetric matrix $\mathbf{M}_d(g \mathbf{z})$ with rows and columns indexed by \mathbb{N}_d^n , and whose entry (α, β) is just $\sum_{\gamma} g_{\gamma} z_{\alpha+\beta+\gamma}$, for every $\alpha, \beta \in \mathbb{N}_d^n$. Alternatively, let $\mathbf{C}_\alpha \in \mathcal{S}^{s(d)}$ be defined by:

$$(2.3) \quad g(\mathbf{x}) \mathbf{v}_d(\mathbf{x}) \mathbf{v}_d(\mathbf{x})^T = \sum_{\alpha \in \mathbb{N}_{2d+\deg g}^n} \mathbf{C}_\alpha \mathbf{x}^\alpha, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Then $\mathbf{M}_d(g \mathbf{z}) = \sum_{\alpha \in \mathbb{N}_{2d+\deg g}^n} z_\alpha \mathbf{C}_\alpha$. If \mathbf{z} has a representing measure μ whose support is contained in the set $\{\mathbf{x} : g(\mathbf{x}) \geq 0\}$ then $\mathbf{M}_d(g \mathbf{z}) \succeq 0$ because

$$\langle \mathbf{f}, \mathbf{M}_d(g \mathbf{z})\mathbf{f} \rangle = \int f^2 g d\mu \geq 0, \quad \forall \mathbf{f} \in \mathbb{R}^{s(d)}.$$

Let $g_0(\mathbf{x}) = 1$ for all $\mathbf{x} \in \mathbb{R}^n$, and with a family $(g_1, \dots, g_m) \subset \mathbb{R}[\mathbf{x}]$ is associated the *quadratic module*:

$$(2.4) \quad Q(g_1, \dots, g_m) := \left\{ \sum_{j=0}^m \sigma_j g_j : \sigma_j \in \Sigma[\mathbf{x}], j = 0, \dots, m \right\},$$

and its *truncated version*

$$(2.5) \quad Q_k(g_1, \dots, g_m) := \left\{ \sum_{j=0}^m \sigma_j g_j : \sigma_j \in \Sigma[\mathbf{x}]_{k-d_j}, j = 0, \dots, m \right\},$$

where $d_j = \lceil \deg(g_j)/2 \rceil$, $j = 0, \dots, m$. Next, given a closed set $\mathcal{X} \subseteq \mathbb{R}^n$, let $C(\mathcal{X}) \subset \mathbb{R}[\mathbf{x}]$ (resp. $C_d(\mathcal{X}) \subset \mathbb{R}[\mathbf{x}]_d$) be the convex cone of polynomials (resp. polynomials of degree at most $2d$) that are nonnegative on \mathcal{X} .

For more details on the above notions of moment and localizing matrix, quadratic module, as well as their use in potential applications, the interested reader is referred to Lasserre [18].

2.2. A basic numerical scheme for upper bounds. Recall that the goal is to compute $\mu(\Omega)$, where μ is the Gaussian measure in (1.2) or the exponential measure in (1.3) and Ω the basic semi-algebraic set in (1.1). Let $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, be the moments of μ , i.e.,

$$y_\alpha = \int \mathbf{x}^\alpha d\mu, \quad \alpha \in \mathbb{N}^n,$$

which are obtained in closed form. Call μ_Ω the *restriction* of μ to Ω , that is:

$$\mu_\Omega(B) = \mu(\Omega \cap B), \quad \forall B \in \mathcal{B}(\mathbb{R}^n).$$

Let $M(\mathbb{R}^n)$ (resp. $M(\Omega)$) be the convex cone of finite (nonnegative) Borel measures on \mathbb{R}^n (resp. Ω), and consider the new abstract problem:

$$(2.6) \quad \rho = \sup_{\phi, \nu} \left\{ \int f d\phi : \phi + \nu = \mu; \quad \phi \in M(\Omega), \nu \in M(\mathbb{R}^n) \right\}.$$

Lemma 2.1. *Let $f \in \mathbb{R}[\mathbf{x}]$ be strictly positive almost everywhere on Ω . Then $(\phi^*, \nu^*) := (\mu_\Omega, \mu - \mu_\Omega)$ is the unique optimal solution of (2.6) and $\phi^*(\Omega) = \mu(\Omega)$ while $\rho = \int_\Omega f d\mu$. In particular if $f = 1$ then $\rho = \mu(\Omega)$.*

Proof. As $\phi \leq \mu$ and $f \geq 0$ on Ω , we have $\int f d\phi \leq \int_\Omega f d\mu = \int f d\mu_\Omega$, and so $\rho \leq \int f d\mu_\Omega$. On the other hand, pick $(\phi^*, \nu^*) := (\mu_\Omega, \mu - \mu_\Omega)$, to obtain $\int f d\phi^* = \int f d\mu_\Omega$, which shows that $\rho = \int f d\mu_\Omega$, and $\phi^*(\Omega) = \mu(\Omega)$.

Next, as $\phi + \nu = \mu$ every feasible solution is such that ϕ is absolutely continuous with respect to μ , and hence with respect to μ_Ω , denoted $\phi \ll \mu_\Omega$. In particular, by the Radon-Nikodym theorem [3], there exists a nonnegative measurable function g on Ω such that

$$\phi(B) = \int_B g(\mathbf{x}) d\mu_\Omega, \quad \forall B \in \mathcal{B}(\Omega),$$

and since $\phi(B) \leq \mu_\Omega(B)$ for all $B \in \mathcal{B}(\Omega)$, it follows that $g(\mathbf{x}) \leq 1$ almost everywhere on Ω . So suppose that there exists another optimal solution (ϕ, ν) with $\int f d\phi = \rho$, that is, $\int f(g-1)d\mu_\Omega = 0$. As $f > 0$ a.e. on Ω and $g \leq 1$ on Ω , this implies $\mu_\Omega(\{\mathbf{x} : g(\mathbf{x}) < 1\}) = 0$. In other words $g = 1$ a.e. on Ω which in turn implies $\phi(B) = \mu(B) = \phi^*(B)$ for all $B \in \mathcal{B}(\Omega)$. That is, $(\phi, \nu) = (\phi^*, \nu^*)$. \square

Problem (2.6) is an infinite dimensional linear program on an appropriate space of measures and its dual is the following optimization problem:

$$(2.7) \quad \rho^* = \inf_{p \in \mathbb{R}[\mathbf{x}]} \left\{ \int p d\mu : p - f \in C(\Omega); \quad p \in C(\mathbb{R}^n) \right\}$$

with $\rho^* \geq \int f d\mu_\Omega$. However neither (2.6) nor its dual (2.7) are tractable. This is why, with $d \geq d_0 := \max_j d_j$ fixed, we now consider the semidefinite program:

$$(2.8) \quad \begin{aligned} \bar{\rho}_d = \sup_{\mathbf{u}, \mathbf{v} \in \mathbb{R}[\mathbf{x}]_{2d}^*} \{ & L_{\mathbf{u}}(f) \text{ s.t. } u_\alpha + v_\alpha = y_\alpha, \quad \forall \alpha \in \mathbb{N}_{2d}^n; \\ & \mathbf{M}_d(\mathbf{u}), \mathbf{M}_d(\mathbf{v}) \succeq 0; \mathbf{M}_{d-d_j}(g_j \mathbf{u}) \succeq 0, \quad j = 1, \dots, m \}, \end{aligned}$$

which is a *relaxation* of (2.6), hence with $\bar{\rho}_d \geq \int_\Omega f d\mu$ for all d . Indeed the constraints on (\mathbf{u}, \mathbf{v}) in (2.8) are only necessary for \mathbf{u} and \mathbf{v} to be moments of some measures ϕ on Ω and ν on \mathbb{R}^n respectively, such that $\phi + \nu = \mu$.

Computational complexity. The semidefinite program (2.8) has $2 \binom{n+2d}{n}$ variables and the matrices $\mathbf{M}_d(\mathbf{u}), \mathbf{M}_d(\mathbf{v})$ are $\binom{n+d}{n} \times \binom{n+d}{n}$. In principle most SDP solvers use interior point methods, considered to be the most efficient. So in view of the present status of state-of-the-art semidefinite solvers, this methodology is still limited to problems in small dimension n (say $n \leq 3, 4$). Indeed even though semidefinite programs (SDPs) are convex optimization problems, the computational complexity of interior point methods is still too high to solve large size SDPs on today's computers. For more details on interior point methods, their analysis and refinements, see e.g. [2] and [27].

Remark 2.2. In case where μ is the exponential measure (1.3) then the support of the measure ϕ in (2.8) is in fact $\Omega \cap \text{supp } \mu = \Omega \cap \mathbb{R}_+^n$, which is the same as replacing Ω with $\Omega \cap \mathbb{R}_+^n$, i.e.,

$$(2.9) \quad \Omega := \Omega \cap \mathbb{R}_+^n = \{\mathbf{x} : g_j(\mathbf{x}) \geq 0, j = 1, m+n\},$$

by introducing the additional polynomials $\mathbf{x} \mapsto g_{m+i}(\mathbf{x}) = x_i, i = 1, \dots, n$.

Then in (2.8) the “support” constraints $\mathbf{M}_{d-d_j}(g_j \mathbf{u}) \succeq 0, j = 1, \dots, m$, now become $\mathbf{M}_{d-d_j}(g_j \mathbf{u}) \succeq 0, j = 1, \dots, m+n$.

The dual of (2.8) is also a semidefinite program, which in compact form reads:

$$(2.10) \quad \bar{\sigma}_d^* = \inf_{p \in \mathbb{R}[\mathbf{x}]_{2d}} \left\{ \int p d\mu : p - f \in Q_{2d}(\Omega); \quad p \in \Sigma[\mathbf{x}]_d \right\},$$

and as expected, is a *strengthening* of (2.7). Indeed one has replaced the nonnegativity constraint $p - f \geq 0$ on Ω (resp. $p \geq 0$ on \mathbb{R}^n) with the stronger membership $(p - f) \in Q_{2d}(\Omega)$ (resp. the stronger membership $p \in \Sigma[\mathbf{x}]_d$).

One main result of this paper is the following:

Theorem 2.3. Let μ be as in (1.2) or (1.3). With $\Omega \subset \mathbb{R}^n$ as in (1.1) (or as in (2.9) when μ is the exponential measure), assume that both Ω and $\text{supp } \mu \setminus \Omega$ have nonempty interior and let $f \in \mathbb{R}[\mathbf{x}]$ be strictly positive a.e. on Ω . Then:

(a) Both problems (2.8) and (2.10) have an optimal solution and $\bar{\sigma}_d = \bar{\sigma}_d^*$ for every $d \geq d_0$.

(b) The sequence $(\bar{\sigma}_d), d \in \mathbb{N}$, is monotone non-increasing and $\bar{\sigma}_d \rightarrow \int_{\Omega} f d\mu$ as $d \rightarrow \infty$. In addition if $(\mathbf{u}^d, \mathbf{v}^d)$ is an optimal solution of (2.8) then $u_0^d \rightarrow \mu(\Omega)$ as $d \rightarrow \infty$.

(c) If $f = 1$ then the sequence $(\bar{\sigma}_d), d \in \mathbb{N}$, is monotone non-increasing and $\bar{\sigma}_d \rightarrow \mu(\Omega)$ as $d \rightarrow \infty$.

Proof. (a) Recall that μ_{Ω} is the restriction of μ to Ω , and let $\mathbf{u} = (y_{\alpha}^{\Omega}), \alpha \in \mathbb{N}_{2d}^n$, be the moments of μ_{Ω} , up to order $2d$. Similarly let $\mathbf{v} = (y_{\alpha} - u_{\alpha}), \alpha \in \mathbb{N}_{2d}^n$, be the moments of $\mu - \mu_{\Omega}$. Hence (\mathbf{u}, \mathbf{v}) is a feasible for (2.8). In addition, as both Ω and $\text{supp } \mu \setminus \Omega$ have nonempty interior, $\mathbf{M}_d(\mathbf{u}) \succ 0$ and $\mathbf{M}_d(\mathbf{v}) \succ 0$. Indeed otherwise there would exist $0 \neq p \in \mathbb{R}[\mathbf{x}]_d$ such that

$$0 = \langle \mathbf{p}, \mathbf{M}_d(\mathbf{u}) \mathbf{p} \rangle = \int p^2 d\mu \Rightarrow p = 0 \text{ a.e. on } \Omega.$$

But then $p = 0$ since it has to vanish on some open set of Ω . The same argument also works for \mathbf{v} . Hence Slater's condition holds for (2.8) since (\mathbf{u}, \mathbf{v}) is strictly feasible solution. By a standard result in convex optimization, this in turn implies

$\bar{\sigma}_d = \bar{\sigma}_d^*$ and moreover (2.10) has an optimal solution if $\bar{\sigma}_d^*$ is finite. We next prove that (2.8) has an optimal solution and therefore so does (2.10). Observe that from the constraint $u_\alpha + v_\alpha = y_\alpha$ we deduce that $u_0 \leq y_0$, $v_0 \leq y_0$. Moreover all (nonnegative) diagonal elements of $\mathbf{M}_d(\mathbf{u})$ and $\mathbf{M}_d(\mathbf{v})$ are dominated by those of $\mathbf{M}_d(\mathbf{y})$. In particular:

$$\max_i [L_{\mathbf{u}}(x_i^{2d})] \leq \max_i [L_{\mathbf{y}}(x_i^{2d})]; \quad \max_i [L_{\mathbf{v}}(x_i^{2d})] \leq \max_i [L_{\mathbf{y}}(x_i^{2d})].$$

From Lasserre [16, Proposition 2.38], $|u_\alpha| \leq \tau_d$ and $|v_\alpha| \leq \tau_d$, for all $\alpha \in \mathbb{N}_{2d}^n$ (where $\tau_d := \max[y_0, \max_i [L_{\mathbf{y}}(x_i^{2d})]]$). Therefore the (closed) feasible of (2.8) is bounded, hence compact, which in turn implies that (2.8) has an optimal solution.

(b) That the sequence $(\bar{\sigma}_d)$, $d \in \mathbb{N}$, is monotone non-increasing is straightforward. Next, let $(\mathbf{u}^d, \mathbf{v}^d)$ be an optimal solution of (2.8). Notice that from the proof of (a) we have seen that $|u_\alpha^d| \leq \tau_d$ and $|v_\alpha^d| \leq \tau_d$ for all $\alpha \in \mathbb{N}_{2d}^n$. Completing with zeros we now consider \mathbf{u}^d and \mathbf{v}^d as infinite vectors indexed by \mathbb{N}^n . Therefore, for each infinite sequence $\mathbf{u}^d = (u_\alpha^d)$, $\alpha \in \mathbb{N}^n$, it holds:

$$u_0^d \leq y_0 \quad \text{and} \quad 2j - 1 \leq |\alpha| \leq 2j \Rightarrow |u_\alpha^d| \leq \tau_j, \quad \forall j = 1, \dots,$$

Similarly

$$v_0^d \leq y_0 \quad \text{and} \quad 2j - 1 \leq |\alpha| \leq 2j \Rightarrow |v_\alpha^d| \leq \tau_j, \quad \forall j = 1, \dots,$$

Hence by a standard argument there exists a subsequence d_k and two infinite sequences $\mathbf{u}^* = (u_\alpha^*)$ and $\mathbf{v}^* = (v_\alpha^*)$, $\alpha \in \mathbb{N}^n$, such that

$$(2.11) \quad \text{for every } \alpha \in \mathbb{N}^n, \quad u_\alpha^{d_k} \rightarrow u_\alpha^* \quad \text{and} \quad v_\alpha^{d_k} \rightarrow v_\alpha^*, \quad \text{as } k \rightarrow \infty.$$

In particular $\lim_{k \rightarrow \infty} \bar{\sigma}_{d_k} = \lim_{k \rightarrow \infty} L_{\mathbf{u}^{d_k}}(f) = L_{\mathbf{u}^*}(f)$, and as the sequence $(\bar{\sigma}_d)$, $d \in \mathbb{N}$, is monotone:

$$(2.12) \quad \int_{\Omega} f d\mu \leq \lim_{d \rightarrow \infty} \bar{\sigma}_d = \lim_{k \rightarrow \infty} \bar{\sigma}_{d_k} = L_{\mathbf{u}^*}(f) \quad \text{and} \quad u_\alpha^* + v_\alpha^* = y_\alpha, \quad \forall \alpha \in \mathbb{N}^n.$$

Then the convergence (2.11) implies that for each fixed k , $\mathbf{M}_k(\mathbf{u}^*) \succeq 0$, $\mathbf{M}_k(\mathbf{v}^*) \succeq 0$, and $\mathbf{M}_k(g_j \mathbf{u}^*) \succeq 0$ for all $j = 1, \dots, m$. Next observe that the moment sequence \mathbf{y} of the Gaussian or exponential measure μ satisfies Carleman's condition (2.2) and so μ is moment determinate. But from $L_{\mathbf{u}^*}(x_i^{2d}) \leq L_{\mathbf{y}}(x_i^{2d})$ and $L_{\mathbf{v}^*}(x_i^{2d}) \leq L_{\mathbf{y}}(x_i^{2d})$, we deduce that both \mathbf{u}^* and \mathbf{v}^* also satisfy Carleman's condition. This latter fact combined with $\mathbf{M}_k(\mathbf{u}^*) \succeq 0$ and $\mathbf{M}_k(\mathbf{v}^*) \succeq 0$ for all k , yields that \mathbf{u}^* and \mathbf{v}^* are the moment sequences of some finite Borel measures ϕ^* and ν^* on \mathbb{R}^n ; see for instance Lasserre [18, Proposition 3.5]

Next, as \mathbf{u}^* satisfies $L_{\mathbf{u}^*}(x_i^{2d}) \leq L_{\mathbf{y}}(x_i^{2d})$ and μ is the Gaussian or exponential measure, there is some $M > 0$ such that $L_{\mathbf{u}^*}(x_i^{2k}) \leq M(2k)!$ for all k . As in addition $\mathbf{M}_k(g_j \mathbf{u}^*) \succeq 0$, for all k , and all $j = 1, \dots, m$, by Lasserre [17, Theorem 2.2, p. 2494], the support of ϕ^* is contained in Ω . Hence $\phi^* \in M(\Omega)$ and $\nu^* \in M(\mathbb{R}^n)$. Moreover from (2.12),

$$\int \mathbf{x}^\alpha d(\phi^* + \nu^*) = \int \mathbf{x}^\alpha d\mu, \quad \forall \alpha \in \mathbb{N}^n,$$

and as μ is moment determinate it follows that $\phi^* + \nu^* = \mu$. Hence the pair (ϕ^*, ν^*) is feasible for problem (2.6) with value $\int f d\phi^* = L_{\mathbf{u}^*}(f) \geq \int_{\Omega} f d\mu = \rho$, which proves that (ϕ^*, ν^*) is an optimal solution of problem (2.6), and so $\int f d\phi^* = \int_{\Omega} f d\mu$.

(c) When $f = 1$ then $\bar{\sigma}_d = u_0^d$ and by (2.11) $u_0^{d_k} \rightarrow \mu(\Omega)$ as $k \rightarrow \infty$. Therefore by monotonicity of the sequence $(\bar{\sigma}_d)$, the result follows. \square

Remark 2.4. (i) The fact that Ω is a basic semi-algebraic set of the form $\{\mathbf{x} : g_j(\mathbf{x}) \geq 0, j = 1, \dots, m\}$ is exploited in the construction of the semidefinite relaxation (2.8) through the psd constraints $\mathbf{M}_{d-d_j}(g_j \mathbf{u}) \succeq 0, j = 1, \dots, m$, on the localizing matrices associated with the (g_j) that define Ω .

(ii) We emphasize that the set Ω is *not* assumed to be compact. So Theorem 2.3 extends the methodology proposed in [13] in two directions: Firstly, we now consider the Gaussian or exponential measure instead of the Lebesgue measure and secondly, we now allow arbitrary (possibly non-compact) basic semi-algebraic sets Ω (whereas in [13] Ω has to be compact) .

(iii) Solving (2.10) has a simple interpretation: Namely when $f = 1$ one tries to approximate the indicator function $\mathbf{x} \mapsto 1_{\Omega}(\mathbf{x})$ ($= 1$ if $\mathbf{x} \in \Omega$ and 0 otherwise) by polynomials of increasing degree. It is well-known that a Gibbs effect occurs at points of the boundary of Ω . So if we choose $\mathbf{x} \mapsto f(\mathbf{x}) := \prod_{j=1}^m g_j(\mathbf{x})$ (so that f is continuous, nonnegative on Ω , and vanishes on the boundary of Ω) then the function $\mathbf{x} \mapsto f(\mathbf{x})1_{\Omega}(\mathbf{x})$ is continuous, hence much easier to approximate by polynomials. For compact sets Ω and the Lebesgue measure μ , it has been observed on [13] that this strategy strongly limits the Gibbs effect and yields drastic improvements on the convergence to $\mu(\Omega)$ (but this convergence is not monotone any more).

However, when $f = 1$ the monotone convergence $\bar{\sigma}_d \rightarrow \mu(\Omega)$ as $d \rightarrow \infty$ is an important attractive feature of the method because when convergence has not taken place yet, one still has the useful information that $\mu(\Omega) \leq \bar{\sigma}_d$, which is important in some applications.

We next show how even with $f = 1$ one may still improve the convergence $\bar{\sigma}_d \rightarrow \mu(\Omega)$ significantly (and thus still keep a monotone sequence of upper bounds on $\mu(\Omega)$). But before we show how to get a converging sequence of lower bounds on $\mu(\Omega)$ with same techniques.

2.3. Lower bounds on $\mu(\Omega)$. Recall that if μ is the exponential measure then for practical computation in (2.8) one replaces Ω with $\Omega := \Omega \cap \text{supp } \mu = \Omega \cap \mathbb{R}_+^n$, that is, Ω is defined in (2.9); see Remark 2.2.

Let $\Omega^c := \text{supp } \mu \setminus \Omega$ be the complement of Ω in the support of μ . As μ is absolutely continuous with respect to the Lebesgue measure, the set Ω^c can be written as

$$(2.13) \quad \Omega^c = \cup_{\ell=1}^s \Omega_{\ell}^c \quad \text{with} \quad \mu(\Omega^c) = \sum_{\ell=1}^s \mu(\Omega_{\ell}^c),$$

where each Ω_{ℓ}^c is a closed basic semi-algebraic set, $\ell = 1, \dots, s$, and the overlaps between the sets Ω_{ℓ}^c have μ -measure zero. Of course the decomposition (2.13) is not unique. For instance if μ is the Gaussian measure and $\Omega = \{\mathbf{x} : g_j(\mathbf{x}) \geq 0, j = 1, 2\}$ then $\Omega^c = \Omega_1^c \cup \Omega_2^c$ with:

$$\Omega_1^c = \{\mathbf{x} : g_1(\mathbf{x}) \leq 0\}; \quad \Omega_2^c := \{\mathbf{x} : g_1(\mathbf{x}) \geq 0; g_2(\mathbf{x}) \leq 0\}.$$

On the other hand if μ is the exponential measure and $\Omega = \{\mathbf{x} \geq 0 : g_j(\mathbf{x}) \geq 0, j = 1, 2\}$ then $\Omega^c = \Omega_1^c \cup \Omega_2^c$ with:

$$\Omega_1^c = \{\mathbf{x} \geq 0 : g_1(\mathbf{x}) \leq 0\}; \quad \Omega_2^c := \{\mathbf{x} \geq 0 : g_1(\mathbf{x}) \geq 0; g_2(\mathbf{x}) \leq 0\}.$$

Then write

$$(2.14) \quad \Omega_\ell^c = \{\mathbf{x} : g_{\ell j}(\mathbf{x}) \geq 0, j = 1, \dots, m_\ell\}, \quad \ell = 1, \dots, s.$$

for some integer m_ℓ and some polynomials $(g_{\ell j}) \subset \mathbb{R}[\mathbf{x}]$, $j = 1, \dots, m_\ell$. Again let $d_j := \lceil (\deg g_{\ell j})/2 \rceil$ and $d_0 = \max_j d_j$.

Corollary 2.5. *Let $\Omega \subset \mathbb{R}^n$ be as in (1.1) (or as in (2.9) when μ is the exponential measure). Assume that both Ω and Ω^c have nonempty interior and let Ω^c be as in (2.13).*

(a) *If $f = 1$ then for each $\ell = 1, \dots, s$, let $\bar{\sigma}_{\ell d}^c$ be the optimal value of the semidefinite program (2.8) with $g_{\ell j}$ in lieu of g_j (and m_ℓ in lieu of m), and for every $d \geq d_0$, let :*

$$(2.15) \quad \underline{\sigma}_d := \mu(\mathbb{R}^n) - \left(\sum_{\ell=1}^s \bar{\sigma}_{\ell d}^c \right).$$

Then the sequence $\underline{\sigma}_d$, $d \in \mathbb{N}$, is monotone non-decreasing with $\mu(\Omega) \geq \underline{\sigma}_d$ for all $d \geq d_0$ and $\underline{\sigma}_d \rightarrow \mu(\Omega)$ as $d \rightarrow \infty$.

(b) *If $f \neq 1$ and f vanishes on $\partial\Omega$, then for each $\ell = 1, \dots, s$, let $\bar{\sigma}_{\ell d}^c$ be the optimal value of the semidefinite program (2.8) with f^2 and $g_{\ell j}$ in lieu of f and g_j (and m_ℓ in lieu of m), and for every $d \geq d_0$, let :*

$$(2.16) \quad \underline{\sigma}_d := \int f^2 d\mu - \left(\sum_{\ell=1}^s \bar{\sigma}_{\ell d}^c \right).$$

Then the sequence $\underline{\sigma}_d$, $d \in \mathbb{N}$, is monotone non-decreasing with $\int_\Omega f^2 d\mu \geq \underline{\sigma}_d$ for all $d \geq d_0$ and $\underline{\sigma}_d \rightarrow \int_\Omega f^2 d\mu$ as $d \rightarrow \infty$.

Proof. (a) By construction of the semidefinite program (2.8), $\bar{\sigma}_{\ell d}^c \geq \mu(\Omega_\ell^c)$ for every $d \geq d_0$ and every $\ell = 1, \dots, s$ and so

$$\sum_{\ell=1}^s \bar{\sigma}_{\ell d}^c \geq \sum_{\ell=1}^s \mu(\Omega_\ell^c) = \mu(\cup_{\ell} \Omega_\ell^c) = \mu(\Omega^c),$$

which in turn implies

$$(2.17) \quad \underline{\sigma}_d = \mu(\mathbb{R}^n) - \sum_{\ell=1}^s \bar{\sigma}_{\ell d}^c \leq \mu(\mathbb{R}^n) - \mu(\Omega^c) = \mu(\Omega), \quad \forall d.$$

Next from Theorem 2.3(b) one has $\bar{\sigma}_{\ell d}^c \rightarrow \mu(\Omega_\ell^c)$ as $d \rightarrow \infty$, for every $\ell = 1, \dots, s$. Therefore using (2.13) and taking limit in (2.17) as $d \rightarrow \infty$, yields the desired result.

(b) Similarly, by construction of the semidefinite program (2.8), $\bar{\sigma}_{\ell d}^c \geq \int_{\Omega_\ell^c} f^2 d\mu$ for every $d \geq d_0$ and every $\ell = 1, \dots, s$ and so

$$\sum_{\ell=1}^s \bar{\sigma}_{\ell d}^c \geq \sum_{\ell=1}^s \int_{\Omega_\ell^c} f^2 d\mu \geq \int_{\cup_{\ell} \Omega_\ell^c} f^2 d\mu = \int_{\Omega^c} f^2 d\mu,$$

which in turn implies

$$(2.18) \quad \underline{\sigma}_d = \int f^2 d\mu - \sum_{\ell=1}^s \bar{\sigma}_{\ell d}^c \leq \int f^2 d\mu - \int_{\Omega^c} f^2 d\mu = \int_\Omega f^2 d\mu, \quad \forall d.$$

Next from Theorem 2.3(b) one has $\bar{\sigma}_{\ell d}^c \rightarrow \mu(\Omega_\ell^c)$ as $d \rightarrow \infty$, for every $\ell = 1, \dots, s$. Therefore using (2.13) and taking limit in (2.17) as $d \rightarrow \infty$, yields the desired result. \square

As a consequence of Theorem 2.3 and Corollary 2.5, we finally obtain:

Corollary 2.6. *Let Ω be as in (1.1) (or as in (2.9) when μ is the exponential measure) and $\bar{\sigma}_d$ be as in (2.8). If $f = 1$ then*

$$(2.19) \quad \underline{\sigma}_d \leq \mu(\Omega) \leq \bar{\sigma}_d, \quad \forall d \geq d_0 \quad \text{and} \quad \lim_{d \rightarrow \infty} \underline{\sigma}_d = \mu(\Omega) = \lim_{d \rightarrow \infty} \bar{\sigma}_d.$$

where $\underline{\sigma}_d$ is defined in (2.15).

Similarly let $f \in \mathbb{R}[\mathbf{x}]$ vanish on $\partial\Omega$. If $\bar{\sigma}_d$ in (2.8) is computed with f^2 (in lieu of f) then

$$(2.20) \quad \underline{\sigma}_d \leq \int_{\Omega} f^2 d\mu \leq \bar{\sigma}_d, \quad \forall d \geq d_0 \quad \text{and} \quad \lim_{d \rightarrow \infty} \underline{\sigma}_d = \int_{\Omega} f^2 d\mu = \lim_{d \rightarrow \infty} \bar{\sigma}_d.$$

where $\underline{\sigma}_d$ is defined in (2.16).

Why can (2.20) be also potentially interesting? Recall that any optimal solution \mathbf{u}^d of (2.8) is such that $u_0^d \rightarrow \mu(\Omega)$ as $d \rightarrow \infty$. So if in (2.20) $\bar{\sigma}_d - \underline{\sigma}_d$ is small then it is a good indication that $u_0^d \approx \mu(\Omega)$.

3. IMPROVING CONVERGENCE

Corollary 2.6 states nothing on the rate of convergence for the sequences $(\bar{\sigma}_d)$ and $(\underline{\sigma}_d)$, $d \in \mathbb{N}$, of upper and lower bounds on $\mu(\Omega)$. As already observed in [13] for compact sets Ω and the Lebesgue measure, when $f = 1$ the convergence appears to be rather slow. On the other hand, if f is positive on Ω and vanishes on $\partial\Omega$ the convergence is significantly faster but one loses the non-increasing (resp. non-decreasing) monotonicity of the sequence $\bar{\sigma}_d$ (resp. $\underline{\sigma}_d$). Keeping the monotonicity is important because no matter how close to $\mu(\Omega)$ is $\bar{\sigma}_d$ or $\underline{\sigma}_d$, one still has the useful information that $\underline{\sigma}_d \leq \mu(\Omega) \leq \bar{\sigma}_d$ for all d .

In this section we show how to improve significantly this convergence while keeping the monotonicity of the sequences of upper and lower bounds. To do this we use Stokes' Theorem for integration and in the sequel, to avoid technicalities, we assume that $\Omega \subset \mathbb{R}^n$ is the closure of its interior, i.e., $\Omega = \overline{\text{int}(\Omega)}$.

3.1. Stokes helps. As we already know in advance that $(\mu_\Omega, \mu - \mu_\Omega)$ is an optimal solution of the infinite-dimensional linear program (2.6), any additional information on the moments of μ_Ω will be helpful if when included as additional constraints in the relaxation (2.8), one still has a semidefinite program to solve. Fortunately for basic semi-algebraic sets Ω this is possible. Indeed suppose for the moment that Ω is compact with smooth boundary $\partial\Omega$ and let X be some given vector field. Let θ_μ be the density of the Gaussian or exponential measure μ and $f \in \mathbb{R}[\mathbf{x}]$ a given polynomial. (In case where μ is the exponential measure then one replaces Ω with $\Omega \cap \mathbb{R}_+^n$.) By Stokes' Theorem, for every $\alpha \in \mathbb{N}^n$:

$$\begin{aligned} & \int_{\Omega} \text{Div}(X) \mathbf{x}^\alpha f(\mathbf{x}) \theta_\mu(\mathbf{x}) d\mathbf{x} + \int_{\Omega} \langle X, \nabla(\mathbf{x}^\alpha f(\mathbf{x}) \theta_\mu(\mathbf{x})) \rangle d\mathbf{x} \\ &= \int_{\partial\Omega} \langle X, \vec{n}_{\mathbf{x}} \rangle \mathbf{x}^\alpha f(\mathbf{x}) \theta_\mu(\mathbf{x}) d\sigma \end{aligned}$$

(where $\vec{n}_{\mathbf{x}}$ is the outward pointing normal at $\mathbf{x} \in \partial\Omega$ and σ is the $(n-1)$ -Hausdorff measure on $\partial\Omega$). In fact the above identity holds even if the boundary is algebraic and not smooth everywhere⁴. Therefore if $f \in \mathbb{R}[\mathbf{x}]$ vanishes on the boundary $\partial\Omega$ then

$$(3.1) \quad \int_{\Omega} \operatorname{Div}(X) \mathbf{x}^{\alpha} f(\mathbf{x}) \theta_{\mu}(\mathbf{x}) d\mathbf{x} + \int_{\Omega} \langle X, \nabla(\mathbf{x}^{\alpha} f(\mathbf{x}) \theta_{\mu}(\mathbf{x})) \rangle d\mathbf{x} = 0.$$

So for instance with $X := e_i$ (with $e_i = (\delta_{j=i}) \in \mathbb{R}^n$) (and since $\operatorname{Div}(X) = 0$):

$$(3.2) \quad \int_{\Omega} \frac{\partial(\mathbf{x}^{\alpha} f(\mathbf{x}) \theta_{\mu}(\mathbf{x}))}{\partial x_i} d\mathbf{x} = 0, \quad \forall \alpha \in \mathbb{N}^n,$$

Equivalently, introduce the polynomials $p_{i,\alpha} \in \mathbb{R}[\mathbf{x}]$, $i = 1, \dots, n$ (of degree $d_{\alpha} = \deg f + |\alpha| + 1$ if μ is the Gaussian measure and $d_{\alpha} = \deg f + |\alpha|$ if μ is the exponential measure), defined by:

$$(3.3) \quad \mathbf{x} \mapsto p_{i,\alpha}(\mathbf{x}) := \frac{\partial(\mathbf{x}^{\alpha} f)}{\partial x_i} - \mathbf{x}^{\alpha} f(\mathbf{x}) x_i, \quad \forall \alpha \in \mathbb{N}^n,$$

if μ is the Gaussian measure (1.2) and by

$$(3.4) \quad \mathbf{x} \mapsto p_{i,\alpha}(\mathbf{x}) := \frac{\partial(\mathbf{x}^{\alpha} f)}{\partial x_i} - \mathbf{x}^{\alpha} f(\mathbf{x}) \quad \forall \alpha \in \mathbb{N}^n,$$

if μ is the exponential measure (1.3). Then we have the equality constraints

$$(3.5) \quad \int p_{i,\alpha} d\mu_{\Omega} = 0, \quad \forall \alpha \in \mathbb{N}^n, i = 1, \dots, n,$$

which defines a *linear* constraint on the moments of μ_{Ω} .

3.2. Extension to non-compact sets Ω . We next show that (3.2) (or, equivalently (3.5)) extends to non-compact semi-algebraic sets Ω .

Lemma 3.1. *Let $\Omega \subset \mathbb{R}^n$ be as in (1.1) (not necessarily compact) or as in (2.9) if μ is the exponential measure (1.3). Let $f \in \mathbb{R}[\mathbf{x}]$ vanish on the boundary $\partial\Omega$. Then with $p_{i,\alpha} \in \mathbb{R}[\mathbf{x}]$ as in (3.3),*

$$(3.6) \quad \int_{\Omega} \frac{\partial(\mathbf{x}^{\alpha} f \exp(-\|\mathbf{x}\|^2/2))}{\partial x_i} d\mathbf{x} = \int_{\Omega} p_{i,\alpha} d\mu = 0,$$

for every $\alpha \in \mathbb{N}^n$ and $i = 1, \dots, n$. Similarly, with $p_{i,\alpha} \in \mathbb{R}[\mathbf{x}]$ as in (3.4),

$$(3.7) \quad \int_{\Omega} \frac{\partial(\mathbf{x}^{\alpha} f \exp(-\sum_{i=1}^n x_i))}{\partial x_i} d\mathbf{x} = \int_{\Omega} p_{i,\alpha} d\mu = 0,$$

for every $\alpha \in \mathbb{N}^n$ and $i = 1, \dots, n$.

Proof. We only consider the case of the Gaussian measure as the proof for the exponential measure uses similar arguments. For $M > 0$ arbitrary, fixed, let $\mathbf{B}_M := \{\mathbf{x} : \|\mathbf{x}\|^2 \leq M^2\}$ and consider the compact set $\Omega_M := \Omega \cap \mathbf{B}_M$. Write $\partial\Omega_M := \partial\Omega_M^1 \cup \partial\Omega_M^2$, where

$$\partial\Omega_M^1 := \partial\Omega \cap \{\mathbf{x} : \|\mathbf{x}\|^2 \leq M^2\}; \quad \partial\Omega_M^2 := \Omega \cap \{\mathbf{x} : \|\mathbf{x}\|^2 = M^2\}.$$

⁴A generalization of Stokes' Theorem is valid for smooth differential forms (see Whitney [26, Theorem 14A], or Federer [8]). Such generalization of Stokes' Theorem has already been used in other contexts like e.g. in Barvinok [4] to provide formulae for exponential integrals on polyhedra (hence with "corners" on the boundary).

Even though the boundary $\partial\Omega_M$ is not smooth everywhere, Stokes' Theorem still applies (see previous footnote) and so if $X = e_i$ and if f vanishes on $\partial\Omega$, (3.2) reads

$$\begin{aligned} \int_{\Omega_M} \frac{\partial(\mathbf{x}^\alpha f \exp(-\|\mathbf{x}\|^2/2))}{\partial x_i} d\mathbf{x} &= \int_{\Omega_M} p_{i,\alpha} d\mu \\ &= \underbrace{\int_{\partial\Omega_M^1} \langle e_i, \vec{n}_{\mathbf{x}} \rangle \mathbf{x}^\alpha f \exp(-\|\mathbf{x}\|^2/2) d\sigma}_{=0} \\ &\quad + \int_{\partial\Omega_M^2} \langle e_i, \vec{n}_{\mathbf{x}} \rangle \mathbf{x}^\alpha f \exp(-\|\mathbf{x}\|^2/2) d\sigma \\ &= \frac{\exp(-M^2/2)}{M} \int_{\partial\Omega_M^2} x_i \mathbf{x}^\alpha f d\sigma \end{aligned}$$

and so

$$\left| \int_{\Omega_M} p_{i,\alpha} d\mu \right| \leq \frac{2\pi^{\frac{n+1}{2}} \|f\|_1 M^{|\alpha|+n+\deg f}}{\Gamma((n+1)/2)} \exp(-M^2/2),$$

because $\sup\{|f(\mathbf{x})| : \|\mathbf{x}\|^2 = M^2\} \leq \|f\|_1 \sup\{|\mathbf{x}^\alpha| : \mathbf{x} \in \mathbf{B}_M\} \leq \|f\|_1 M^{\deg f}$, with $\|f\|_1 = \sum_\alpha |f_\alpha|$. This in turn implies

$$(3.8) \quad \lim_{M \rightarrow \infty} \int_{\Omega_M} p_{i,\alpha} d\mu = 0, \quad \forall \alpha \in \mathbb{N}^n; \quad i = 1, \dots, n.$$

It remains to show that

$$\int_{\Omega_M} p_{i,\alpha} d\mu \rightarrow \int_{\Omega} p_{i,\alpha} d\mu \quad \text{as } M \rightarrow \infty.$$

Observe that

$$|p_{i,\alpha}(\mathbf{x})| I_{\Omega_M}(\mathbf{x}) \exp(-\|\mathbf{x}\|^2/2) \leq |p_{i,\alpha}(\mathbf{x})| \exp(-\|\mathbf{x}\|^2/2), \quad \forall \mathbf{x} \in \mathbb{R}^n,$$

$$\text{and } \int |p_{i,\alpha}(\mathbf{x})| \exp(-\|\mathbf{x}\|^2/2) d\mathbf{x} < \infty.$$

In addition, for every $\mathbf{x} \in \mathbb{R}^n$:

$$p_{i,\alpha}(\mathbf{x}) I_{\Omega_M}(\mathbf{x}) \exp(-\|\mathbf{x}\|^2/2) \rightarrow p_{i,\alpha}(\mathbf{x}) I_{\Omega}(\mathbf{x}) \exp(-\|\mathbf{x}\|^2/2) \quad \text{as } M \rightarrow \infty.$$

Therefore by (3.8) and invoking the Dominated Convergence Theorem [3] one obtains the desired result

$$\begin{aligned} 0 &= \lim_{M \rightarrow \infty} \int_{\Omega_M} p_{i,\alpha} d\mu = \lim_{M \rightarrow \infty} \int p_{i,\alpha}(\mathbf{x}) I_{\Omega_M}(\mathbf{x}) \exp(-\|\mathbf{x}\|^2/2) d\mathbf{x} \\ &= \int p_{i,\alpha}(\mathbf{x}) I_{\Omega}(\mathbf{x}) \exp(-\|\mathbf{x}\|^2/2) d\mathbf{x} = \int_{\Omega} p_{i,\alpha} d\mu. \end{aligned}$$

□

So with $p_{i,\alpha} \in \mathbb{R}[\mathbf{x}]$ as in (3.3) (or (3.4)) and $r(d) := 2d - 1$ (or $r(d) := 2d$), we can now consider the new semidefinite program:

$$(3.9) \quad \begin{aligned} \bar{\omega}_d &= \sup_{\mathbf{u}, \mathbf{v} \in \mathbb{R}[\mathbf{x}]_{2d}^*} \{ u_0 : \text{s.t. } u_\alpha + v_\alpha = y_\alpha, \quad \forall \alpha \in \mathbb{N}_{2d}^n; \\ &\quad \mathbf{M}_d(\mathbf{u}), \mathbf{M}_d(\mathbf{v}) \succeq 0; \mathbf{M}_{d-d_j}(g_j \mathbf{u}) \succeq 0, \quad j = 1, \dots, m, \\ &\quad L_{\mathbf{u}}(p_{i,\alpha}) = 0, \quad \forall \alpha \in \mathbb{N}_{r(d)}^n; i = 1, \dots, n \}, \end{aligned}$$

which is a relaxation of the infinite-dimensional linear program (2.6) (with $f = 1$). The dual of (3.9) is the semidefinite program:

$$(3.10) \quad \begin{aligned} \bar{\omega}_d^* = \inf_{p \in \mathbb{R}[\mathbf{x}]_{2d}, \boldsymbol{\theta}} \quad & \left\{ \int p d\mu : \text{s.t. } \boldsymbol{\theta} \in \mathbb{R}^{nm(d)}, p \in \Sigma[\mathbf{x}]_d; \right. \\ & \left. p + \sum_{i=1}^n \sum_{\alpha \in \mathbb{N}_{m(d)}^n} \theta_{i\alpha} p_{i,\alpha} - 1 \in Q_{2d}(\boldsymbol{\Omega}), \right. \end{aligned}$$

where $m(d) = \binom{n+r(d)}{n}$ and $\boldsymbol{\theta} = (\theta_{i,\alpha}) \in \mathbb{R}^{nm(d)}$ is the vector of dual variables associated with the equality constraints $L_{\mathbf{u}}(p_{i,\alpha}) = 0$, $\alpha \in \mathbb{N}_{r(d)}^n$, $i = 1, \dots, n$.

Again, if μ is the exponential measure (1.3) then $\boldsymbol{\Omega}$ is as in (2.9) and so in (3.9) the constraints $\mathbf{M}_{d-d_j}(g_j \mathbf{u}) \succeq 0$, $j = 1, \dots, m$, become $\mathbf{M}_{d-d_j}(g_j \mathbf{u}) \succeq 0$, $j = 1, \dots, m+n$; see Remark 2.2.

Theorem 3.2. *Let $\boldsymbol{\Omega} \subset \mathbb{R}^n$ be as in (1.1) (or as in (2.9) if μ is the exponential measure) and assume that both $\boldsymbol{\Omega}$ and $\text{supp } \mu \setminus \boldsymbol{\Omega}$ have nonempty interior. Then:*

(a) *Both problems (3.9) and its dual have an optimal solution and $\bar{\omega}_d = \bar{\omega}_d^*$ for every $d \geq \max_j d_j$.*

(b) *The sequence $(\bar{\omega}_d)$, $d \in \mathbb{N}$, is monotone non-increasing and $\bar{\omega}_d \rightarrow \mu(\boldsymbol{\Omega})$ as $d \rightarrow \infty$.*

Proof. (a) follows from an almost verbatim copy of the proof of Theorem 2.3. Concerning (b) it suffices to observe that $\bar{\omega}_d \geq \bar{\omega}_d \geq \mu(\boldsymbol{\Omega})$ for every $d \geq d_0$. Therefore the convergence $\bar{\omega}_d \rightarrow \mu(\boldsymbol{\Omega})$ is a consequence of the convergence $\bar{\omega}_d \rightarrow \mu(\boldsymbol{\Omega})$. Finally the convergence is monotone since by construction the sequence $(\bar{\omega}_d)$, $d \in \mathbb{N}$, is monotone non-increasing. \square

Next, exactly as we did in §2.3 we can provide an associated sequence of lower bounds $(\underline{\omega}_d)$, $d \in \mathbb{N}$, by considering the complement $\boldsymbol{\Omega}^c := \text{supp } \mu \setminus \boldsymbol{\Omega}$ in $\text{supp } \mu$ (with $\boldsymbol{\Omega} := \boldsymbol{\Omega} \cap \text{supp } \mu$) and its decomposition (2.13). Then the analogue of Corollary 2.5 reads:

Corollary 3.3. *Let $\boldsymbol{\Omega} \subset \mathbb{R}^n$ be as in (1.1). Assume that both $\boldsymbol{\Omega}$ and $\boldsymbol{\Omega}^c$ have nonempty interior and let $\boldsymbol{\Omega}^c$ be as in (2.13). For each $\ell = 1, \dots, s$, let $\bar{\omega}_{\ell d}^c$ be the optimal value of the semidefinite program (3.9) with $g_{\ell j}$ in lieu of g_j and m_{ℓ} in lieu of m . For every $d \geq d_0$ let :*

$$(3.11) \quad \underline{\omega}_d := \mu(\mathbb{R}^n) - \left(\sum_{\ell=1}^s \bar{\omega}_{\ell d}^c \right).$$

Then the sequence $(\underline{\omega}_d)$, $d \in \mathbb{N}$, is monotone non-decreasing with $\mu(\boldsymbol{\Omega}) \geq \underline{\omega}_d$ for all $d \geq d_0$ and $\underline{\omega}_d \rightarrow \mu(\boldsymbol{\Omega})$ as $d \rightarrow \infty$.

The proof being similar to that of Corollary 2.5 is omitted. So again but now by Theorem 3.2 and Corollary 3.3 one has :

$$(3.12) \quad \underline{\omega}_d \leq \mu(\boldsymbol{\Omega}) \leq \bar{\omega}_d, \quad \forall d \geq d_0 \quad \text{and} \quad \lim_{d \rightarrow \infty} \underline{\omega}_d = \mu(\boldsymbol{\Omega}) = \lim_{d \rightarrow \infty} \bar{\omega}_d.$$

4. SOME NUMERICAL EXPERIMENTS

In this section we provide some examples to illustrate the methodology developed in §3. The examples are all 2D-examples as 3D-examples would have required some help for implementation in GloptiPoly. Also and importantly, for simplicity of

implementation of the semidefinite programs (3.9) we have chosen to express the moment and localizing matrices $\mathbf{M}_d(\mathbf{z})$ and $\mathbf{M}_{d-d_j}(g_j \mathbf{z})$ in the standard monomial basis (\mathbf{x}^α) , $\alpha \in \mathbb{N}^n$, which is the worst choice from a numerical point of view. Indeed, it is well known that the Hankel-type moment and localizing matrices matrix become rapidly ill-conditioned as their size increases. A much better choice would be the basis of Hermite polynomials that are orthogonal with respect to the Gaussian measure on \mathbb{R}^n (and similarly for the exponential measure). However this would require a non trivial implementation work beyond the scope of the present paper. Nonetheless the 2D-examples provided below already give some indications on the potential of the method since in many cases (but not all) good approximations $\underline{\omega}_d \leq \mu(\Omega) \leq \overline{\omega}_d$ are obtained with relatively small d (say $d = 8, 9$ or 10).

4.1. Numerical experiments for the Gaussian measure. We have considered the Gaussian measures $d\mu = \exp(-\|\mathbf{x}\|^2/\sigma^2)$ for the three values $\sigma^2 = 0.5, 0.8, 1$. In a first set of experiments we have tested the methodology on some examples of compact sets $\Omega \subset \mathbb{R}^2$ where we could compute a very good approximate value of $\mu(\Omega)$ by other methods. For each example we have compared the hierarchy of semidefinite programs (2.8) with the hierarchy (3.9). Then we have tested the two hierarchies on some simple example of non-compact sets $\Omega \subset \mathbb{R}^2$ for which the value $\mu(\Omega)$ can be obtained exactly.

On compact sets $\Omega \subset \mathbb{R}^2$.

Example 1. Consider non-centered Euclidean balls $\Omega := \{\mathbf{x} : \|\mathbf{x} - \mathbf{u}\|^2 \leq 1\}$ with

$$\mathbf{u} = (0, 0), \quad (0.1, 0.1), \quad (0.5, 0.5), \quad (0.1, 0.5), \quad (1, 0),$$

that is,

$$\Omega = \{\mathbf{x} : g(\mathbf{x}) \leq 1\}, \quad \text{with} \quad \mathbf{x} \mapsto g(\mathbf{x}) = (x_1 - u_1)^2 + (x_2 - u_2)^2.$$

We first show how the formulation (3.9) with $f = 1$ drastically improves convergence when comparing with the initial formulation (2.8) with $f = 1$ and $f = 1 - g$ (which vanishes on $\partial\Omega$) for the example where $\sigma = 0.5$ and $\mathbf{u} = (0.5, 0.5)$. Indeed in Figure 1 one may see that the convergence is very slow for the hierarchy (2.8) with $f = 1$ (left) whereas it is very fast for the hierarchy (3.9) (right). In Figure 2 are plotted the values $\overline{\omega}_d$ and the values u_0^d obtained in (2.8) with $f = 1 - g$. One can see that the convergence $u_0^d \rightarrow \mu(\Omega)$ of the hierarchy (2.8) with $f = 1 - g$ is not monotone and not as good as (3.9). (The yellow horizontal line is the value of $\mu(\Omega)$, exact up to 6 digits.) Iteration 1 was run with $d = 3$ (i.e. with moments up to order 6) and $d + 1, \dots, 9$ (i.e., with moments up of order 20). With $d = 7$ (i.e. with moments up to order 14) the optimal value $s_7 = 0.573324$ obtained in (3.9) is already very close to the value 0.573132 (exact up to 6 digits).

For some other value of the parameters σ and \mathbf{u} , Figure 3 displays the difference $\overline{\omega}_d - \mu(\Omega)$ for the hierarchy (2.8) with $f = 1$ and $\overline{\omega}_d - \mu(\Omega)$ for the hierarchy (3.9), whereas Figure 4 displays the respective relative errors.

Results obtained for several combinations of the parameters σ and \mathbf{u} are displayed in Tables 1, 2 and 3.

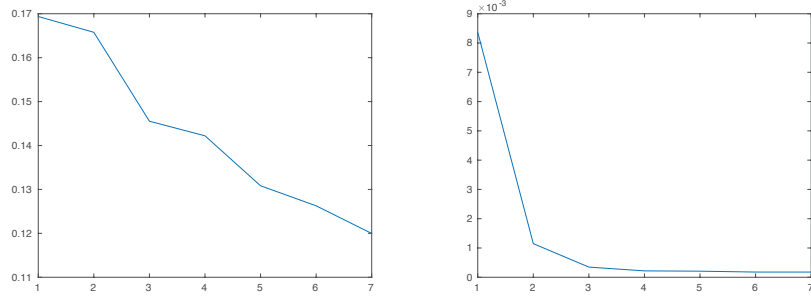


FIGURE 1. Example 1: $\sigma = 0.5$, $\mathbf{u} = (0.5, 0.5)$; Comparing $\bar{\sigma}_d - \mu(\Omega)$ with $f = 1$ (left) and $\bar{\omega}_d - \mu(\Omega)$ (right).

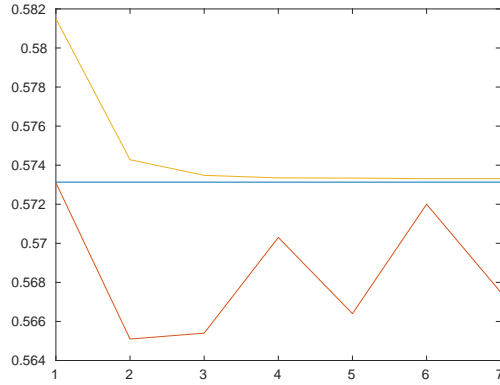


FIGURE 2. Example 1: $\sigma = 0.5$, $\mathbf{u} = (0.5, 0.5)$; Comparing (3.9) and (2.8) with $f = 1 - g$; the horizontal line is the exact value.

Example 2. Consider a set $\Omega := \{\mathbf{x} \in \mathbb{R}^2 : g(\mathbf{x}) \leq 1\}$ with $\mathbf{x} \mapsto g(\mathbf{x}) := (\mathbf{x} - \mathbf{u})^T \mathbf{A} (\mathbf{x} - \mathbf{u})$ with:

$$\mathbf{A} := \begin{bmatrix} 0.4 & 0.1 \\ 0.1 & -0.4 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} 0.4 & 0.1 \\ 0.1 & -0.4 \end{bmatrix}$$

and $\mathbf{u} = (0.1, 0.5)$, $\mathbf{u} = (0.5, 0.1)$. The results displayed in Table 4 show that good approximations can be obtained with relatively small $d = 9$ when $\sigma \leq 1$.

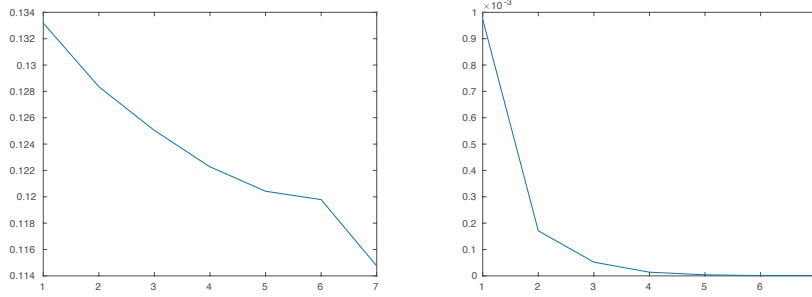


FIGURE 3. Example 1: $\bar{\omega}_d - \mu(\Omega)$ (scheme (2.8) with $f = 1$, left) and $\underline{\omega}_d - \mu(\Omega)$ (scheme (3.9) right)

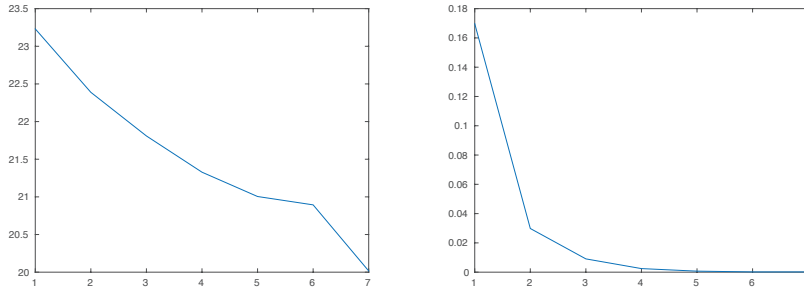


FIGURE 4. Comparing the relative errors $100(\bar{\omega}_d - \mu(\Omega))/\mu(\Omega)$ (scheme (2.8) with $f = 1$, left) with $100(\underline{\omega}_d - \mu(\Omega))/\mu(\Omega)$ (scheme (3.9), right)

	$\mathbf{u} = (0, 0)$	$\mathbf{u} = (0.1, 0.1)$	$\mathbf{u} = (0.1, 0.5)$	$\mathbf{u} = (0.5, 0.5)$	$\mathbf{u} = (1, 0)$
$\bar{\omega}_7$	0.782888	0.781978	0.783345	0.753521	0.441636
$\underline{\omega}_7$	0.782842	0.781932	0.783238	0.753492	0.441631
ϵ_7	0.005%	0.006%	0.01%	0.003%	0.001%

TABLE 1. Values of $\bar{\omega}_7$, $\underline{\omega}_7$ and $\epsilon_7 := 100(\bar{\omega}_7 - \underline{\omega}_7)/\underline{\omega}_7$ with $\sigma = .5$

	$\mathbf{u} = (0, 0)$	$\mathbf{u} = (0.1, 0.1)$	$\mathbf{u} = (0.1, 0.5)$	$\mathbf{u} = (0.5, 0.5)$	$\mathbf{u} = (1, 0)$
$\bar{\omega}_7$	1.906281	1.900152	1.920172	1.770550	1.183305
$\underline{\omega}_7$	1.905856	1.899708	1.919838	1.770276	1.182967
ϵ_7	0.02%	0.02%	0.017%	0.015%	0.028%

TABLE 2. Values of $\bar{\omega}_7$, $\underline{\omega}_7$ and $\epsilon_7 := 100(\bar{\omega}_7 - \underline{\omega}_7)/\underline{\omega}_7$ with $\sigma = .8$

Non-compact sets $\Omega \subset \mathbb{R}^2$. We next provide some examples of non-compact sets $\Omega \subset \mathbb{R}^2$ and show that the hierarchy (3.9) can still provide good results.

	$\mathbf{u} = (0, 0)$	$\mathbf{u} = (0.1, 0.1)$	$\mathbf{u} = (0.1, 0.5)$	$\mathbf{u} = (0.5, 0.5)$	$\mathbf{u} = (1, 0)$
$\bar{\omega}_7$	2.811312	2.802918	2.843411	2.624901	2.947715
$\underline{\omega}_7$	2.806004	2.797471	2.839425	2.621340	2.946099
ϵ_7	0.19%	0.19%	0.14%	0.13%	0.05%

TABLE 3. Values of $\bar{\omega}_7$, $\underline{\omega}_7$ and $\epsilon_7 := 100(\bar{\omega}_7 - \underline{\omega}_7)/\underline{\omega}_7$ with $\sigma = .999$

$\mathbf{u} = (0.1, 0.5)$			
σ	$\bar{\omega}_9$	$\underline{\omega}_9$	$(\bar{\omega}_9 - \underline{\omega}_9)/\underline{\omega}_9$
1	1.728640	1.657139	4.3%
0.8	1.305168	1.298370	0.5%
0.5	0.648429	0.648415	0.002%
$\mathbf{u} = (0.5, 0.1)$			
σ	$\bar{\omega}_9$	$\underline{\omega}_9$	$(\bar{\omega}_9 - \underline{\omega}_9)/\underline{\omega}_9$
1	1.798363	1.730477	3.9%
0.8	1.401565	1.394053	0.53%
0.5	0.724022	0.724005	0.002%

TABLE 4. Example 2: $\Omega = \{\mathbf{x} : (\mathbf{x} - \mathbf{u})^T \mathbf{A}(\mathbf{x} - \mathbf{u}) \leq 1\}$; compact case.

Example 3. Complement of Euclidean balls: We have first considered evaluating $\mu(\Omega^c)$ for the complement $\Omega^c := \{\mathbf{x} : \|\mathbf{x} - \mathbf{u}\|^2 \geq 1\}$ of the sets Ω considered in Example 1. As the lower bound $\underline{\omega}_d$ on $\mu(\Omega)$ in Example 1 was computed via an upper bound on $\mu(\Omega^c)$, the results in Table 1, Table 2 and Table 3 clearly indicate that good upper and lower bounds are also obtained for Ω^c .

Example 4. Non-convex quadratics: Let $\mathbf{u} = (0.1, 0.5)$, $\mathbf{u} = (0.5, 1)$,

$$\mathbf{A} := \begin{bmatrix} 0.4 & 0.1 \\ 0.1 & -0.4 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & -8 \end{bmatrix} \begin{bmatrix} 0.4 & 0.1 \\ 0.1 & -0.4 \end{bmatrix},$$

and $\mathbf{x} \mapsto g(\mathbf{x}) = (\mathbf{x} - \mathbf{u})^T \mathbf{A}(\mathbf{x} - \mathbf{u})$ so that the set $\Omega := \{\mathbf{x} : g(\mathbf{x}) \leq 1\}$ is non-compact. The results displayed in Table 5 show that good approximations can be obtained with relatively small $d = 9$ when $\sigma \leq 1$.

$\mathbf{u} = (0.5, 0.1)$			
σ	$\bar{\omega}_9$	$\underline{\omega}_9$	$(\bar{\omega}_9 - \underline{\omega}_9)/\underline{\omega}_9$
1	2.829605	2.824718	0.17%
0.8	1.876731	1.876609	0.006%
$\mathbf{u} = (0.1, 0.5)$			
σ	$\bar{\omega}_9$	$\underline{\omega}_9$	$(\bar{\omega}_9 - \underline{\omega}_9)/\underline{\omega}_9$
1	2.989832	2.986599	0.10%
0.8	1.969188	1.969103	0.004%

TABLE 5. $\Omega = \{\mathbf{x} : (\mathbf{x} - \mathbf{u})^T \mathbf{A}(\mathbf{x} - \mathbf{u}) \leq 1\}$; non-compact case.

Example 5. Half-spaces: In this example we consider the half-space $\Omega := \{\mathbf{x} \in \mathbb{R}^2 : x_1 + 2x_2 \geq 1\}$, and compute the upper and lower bounds $\bar{\omega}_d$ and $\underline{\omega}_d$ for $d = 8$.

the results displayed in Table 6 show that good approximations can be obtained with relatively small $d = 8$ when $\sigma \leq 1$.

σ	$\bar{\omega}_8$	$\underline{\omega}_8$	$(\bar{\omega}_8 - \underline{\omega}_8)/\underline{\omega}_8$
1.0	0.828105	0.827800	0.03%
0.8	0.4314786	0.431473	0.001%
0.5	0.080858	0.0808578	0.0003%

TABLE 6. $\Omega := \{\mathbf{x} \in \mathbb{R}^2 : x_1 + 2x_2 \geq 1\}$

Example 6. In this example we consider the cone $\Omega := \{\mathbf{x} \in \mathbb{R}^2 : x_1 + 2x_2 \leq -0.5; x_1 \geq -0.8\}$, with apex $\mathbf{x} = (-0.8, 0.15)$, and compute the upper and lower bounds $\bar{\omega}_d$ and $\underline{\omega}_d$ for $d = 10$.

σ	$\bar{\omega}_{10}$	$\underline{\omega}_{10}$	$(\bar{\omega}_{10} - \underline{\omega}_{10})/\underline{\omega}_{10}$
0.5	0.208606	0.19410	7%
0.4	0.108321	0.10768	0.6%
0.3	0.041300	0.0411906	0.26%

TABLE 7. $\Omega := \{\mathbf{x} \in \mathbb{R}^2 : x_1 + 2x_2 \leq -0.5; x_1 \geq -0.8\}$

Table 7 shows that when σ is relatively small then a good approximation can be obtained by using no more than $2d = 20$ moments. On the other hand for the cone $\Omega := \{\mathbf{x} : \mathbf{x} \geq 0\}$, even with $\sigma = 0.3$ the approximation with $d = 10$ is rather rough as we have $\bar{\omega}_{10} = 0.11172 \geq \mu(\Omega) \approx 0.070685$ and lower bound being negative is not informative. It seems that this is because the origin (the apex of the cone Ω) is in a region where the density $\exp(-h)$ is maximum. In this case higher order moments are needed for a better approximation. This in turn requires some care in solving the associated semidefinite programs (3.9). Indeed, when expressed in the standard monomial basis (\mathbf{x}^α) , $\alpha \in \mathbb{N}$, the moment and localizing matrices in (3.9) become ill-conditioned. An interesting issue of further investigation and beyond the scope of this paper is to consider alternative bases, e.g. the basis of Hermite polynomials which are orthogonal with respect to the Gaussian measure.

4.2. Numerical experiments for the exponential measure. In this second set of experiments we have chosen the exponential measure with density $\mathbf{x} \mapsto \exp(-\lambda \sum_i x_i)$ on the positive orthant $\mathbb{R}_+^n = \{\mathbf{x} : \mathbf{x} \geq 0\}$, for some real scalar $\lambda > 0$.

Compact sets.

Example 7. Ellipsoids: As in Example 2, consider the set $\Omega := \{\mathbf{x} \in \mathbb{R}^2 : g(\mathbf{x}) \leq 1\}$ with $\mathbf{x} \mapsto g(\mathbf{x}) := (\mathbf{x} - \mathbf{u})^T \mathbf{A} (\mathbf{x} - \mathbf{u})$ with:

$$\mathbf{A} := \begin{bmatrix} 0.4 & 0.1 \\ 0.1 & -0.4 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} 0.4 & 0.1 \\ 0.1 & -0.4 \end{bmatrix}$$

and $\mathbf{u} = (0.1, 0.5)$, $\mathbf{u} = (0.5, 0.1)$.

$\mathbf{u} = (0.1, 0.5)$			
λ	$\bar{\omega}_9$	$\underline{\omega}_9$	$(\bar{\omega}_9 - \underline{\omega}_9)/\underline{\omega}_9$
4	0.061906	0.060311	2.64%
4.5	0.049080	0.048498	1.2%
5	0.039844	0.039619	0.56%
$\mathbf{u} = (0.5, 0.1)$			
σ	$\bar{\omega}_9$	$\underline{\omega}_9$	$(\bar{\omega}_9 - \underline{\omega}_9)/\underline{\omega}_9$
4	0.061394	0.059065	3.94%
4.5	0.048757	0.047922	1.74%
5	0.039622	0.039283	0.86%

TABLE 8. $\Omega = \{\mathbf{x} : (\mathbf{x} - \mathbf{u})^T \mathbf{A}(\mathbf{x} - \mathbf{u}) \leq 1\}$; compact case.

Example 8. Simplex: Let $\Omega := \{\mathbf{x} \in \mathbb{R}_+^2 : 3x_1 + x_2 \leq 1\}$. In this case it is easy to compute the exact value $\mu(\Omega)$ by Laplace transform techniques, which gives

$$(4.1) \quad \mu(\Omega) = \frac{1}{\lambda^2} \left(1 + \frac{\exp(-p)}{2} - \frac{3 \exp(-p/3)}{2} \right).$$

Table 9 displays results obtained with $d = 8$.

λ	$\mu(\Omega)$	$\bar{\omega}_8$	$\underline{\omega}_8$	$(\bar{\omega}_8 - \underline{\omega}_8)/\underline{\omega}_8$	$(\bar{\omega}_8 - \mu(\Omega))/\mu(\Omega)$
5	0.028802	0.029771	0.028086	6%	3.36%
6	0.022173	0.022605	0.021790	3.74%	1.95%

TABLE 9. Ex. 8: $\Omega = \{\mathbf{x} \geq 0 : 3x_1 + x_2 \leq 1\}$; compact case.

Non-compact sets.

Example 9. Half-space: Let Ω be the non-compact set $\{\mathbf{x} \in \mathbb{R}_+^2 : 3x_1 + x_2 \geq 1\}$. From (4.1) we have

$$\mu(\Omega) = \mu(\mathbb{R}^2) - \mu(\Omega^c) = \frac{1}{\lambda^2} \left(\frac{3 \exp(-p/3)}{2} - \frac{\exp(-p)}{2} \right).$$

Results for $d = 8$ are displayed in Table 10. For $d = 9$ some numerical instability occurred.

λ	$\mu(\Omega)$	$\bar{\omega}_8$	$\underline{\omega}_8$	$(\bar{\omega}_8 - \underline{\omega}_8)/\underline{\omega}_8$	$(\bar{\omega}_8 - \mu(\Omega))/\mu(\Omega)$
5	0.011197	0.011913	0.010228	16.4%	6.4%
6	0.005604	0.005986	0.005171	15.7%	6.8%

TABLE 10. Ex 9: $\Omega = \{\mathbf{x} \geq 0; 3x_1 + x_2 \geq 1\}$

Example 10. Hyperbola: Let $\Omega := \{\mathbf{x} \in \mathbb{R}_+^2 : x_1 x_2 \leq 0.1\}$. Results for $d = 8$ are displayed in Table 10. As in Example 9, some numerical instability occurred for $d = 9$.

$\Omega = \{\mathbf{x} \geq 0; x_1 x_2 \leq 0.1\}$			
λ	$\bar{\omega}_8$	$\underline{\omega}_8$	$(\bar{\omega}_8 - \underline{\omega}_8)/\underline{\omega}_8$
5	0.004328	0.003970	9%
6	0.001701	0.001616	5.2
$\Omega = \{\mathbf{x} \geq 0; x_1 x_2 \geq 0.1\}$			
λ	$\bar{\omega}_8$	$\underline{\omega}_8$	$(\bar{\omega}_8 - \underline{\omega}_8)/\underline{\omega}_8$
5	0.036015	0.035635	1%
6	0.026166	0.026068	0.37%

TABLE 11. $\Omega = \{\mathbf{x} \geq 0 : x_1 x_2 \leq 0.1\}$ and Ω^c

Example 11. Non-convex quadratics: Let $\Omega := \{\mathbf{x} \geq 0 : g(\mathbf{x}) \leq 0.05\}$ with $\mathbf{x} \mapsto g(\mathbf{x}) := (\mathbf{x} - \mathbf{u})^T \mathbf{A} (\mathbf{x} - \mathbf{u})$ with:

$$\mathbf{A} := \begin{bmatrix} 0.4 & 0.1 \\ 0.1 & -0.4 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 10 \end{bmatrix} \begin{bmatrix} 0.4 & 0.1 \\ 0.1 & -0.4 \end{bmatrix}$$

and $\mathbf{u} = (0, 0)$, $\mathbf{u} = (-0.1, 0.1)$ and $\mathbf{u} = (0.1, -0.1)$. The results displayed in Table 12 show that for the same value of $d = 9$ the quality of the approximation can be sensitive to data. For instance it is much worse with $\mathbf{u} = (-0.1, 0.1)$ than with $\mathbf{u} = (0, 0)$ or $\mathbf{u} = (-0.1, 0.1)$.

$\mathbf{u} = (0, 0)$			
λ	$\bar{\omega}_9$	$\underline{\omega}_9$	$(\bar{\omega}_9 - \underline{\omega}_9)/\underline{\omega}_9$
8	0.013444	0.012065	11%
9	0.010961	0.009890	10%
$\mathbf{u} = (-0.1, 0.1)$			
λ	$\bar{\omega}_9$	$\underline{\omega}_9$	$(\bar{\omega}_9 - \underline{\omega}_9)/\underline{\omega}_9$
8	0.014877	0.014288	4.1%
9	0.011925	0.011512	3.6%
$\mathbf{u} = (0.1, -0.1)$			
λ	$\bar{\omega}_9$	$\underline{\omega}_9$	$(\bar{\omega}_9 - \underline{\omega}_9)/\underline{\omega}_9$
8	0.009464	0.006009	57%
8	0.007933	0.004929	60%

TABLE 12. $\Omega = \{\mathbf{x} \geq 0 : (\mathbf{x} - \mathbf{u})^T \mathbf{A} (\mathbf{x} - \mathbf{u}) \leq 0.05\}$;

5. CONCLUDING REMARKS

We have provided a systematic numerical scheme to approximate the Gaussian or exponential measure $\mu(\Omega)$ of a class of semi-algebraic sets $\Omega \subset \mathbb{R}^2$. In principle $\mu(\Omega)$ can be approximated as closely as desired by solving a hierarchy of semidefinite programs of increasing size d . Of course the size of the resulting semidefinite programs increases with d . Namely the semidefinite program (3.9) has $2\binom{n+2d}{n}$ variables and both (real symmetric) moment matrices $\mathbf{M}_d(\mathbf{u})$ and $\mathbf{M}_d(\mathbf{v})$ are of size $\binom{n+d}{n} \times \binom{n+d}{n}$.

So as long as one is interested in $2D$ or $3D$ problems the size is not the most serious problem. Indeed when $\mu(\Omega)$ is not very small compared to the mass of μ

(and $\sigma < 1$ for the gaussian measure and $\lambda > 5$ for the exponential measure), good results can be expected for reasonable values of d (as shown in some illustrative 2D-examples). On the other hand if one needs high values of d (e.g. when $\sigma \gg 1$ or $\lambda > 1$) then the precision of the SDP solvers can become a serious issue as the semidefinite programs (3.9) become ill-conditioned.

In particular in such conditions the choice of the monomial basis (\mathbf{x}^α) , $\alpha \in \mathbb{N}^n$, in which to express the moment and localizing matrices $\mathbf{M}_d(\mathbf{u})$, $\mathbf{M}_d(g_j \mathbf{u})$ is not appropriate. At this preliminary validation stage of the methodology it was done for simplicity and easyness of implementation (in order to use the software package GloptiPoly [12]). To avoid ill-conditioning a more appropriate strategy is to use the basis of *Hermite polynomials*, orthogonal with respect to the Gaussian measure μ . With such a choice one may expect to be able to solve (3.9) for higher values of d and so obtain better upper and lower bounds.

In any case it is also worth mentioning that even if one is forced to stop the hierarchy (3.9) at a relatively small value of d , one has still obtained a non trivial finite sequence of upper and lower bounds on $\mu(\Omega)$ for non-trivial (and possibly non-compact) sets Ω .

Acknowledgements. Research funded by by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement ERC-ADG 666981 TAMING)”

REFERENCES

- [1] S. Alfano. A numerical implementation of spherical objet collision probability, *J. Astro. Sci.* **53**, pp., 2005.
- [2] M. Anjos, J.B. Lasserre. *Handbook of Semidefinite, Conic and Polynomial Optimization*, M. Anjos and J.B. Lasserre Eds., Springer-Verlag, New York, 2012.
- [3] R. Ash. *Real Analysis and Probability*, Academic Press Inc., San Diego, 1972.
- [4] A.I. Barvinok. Exponential integrals and sums over convex polyhedra, *Funct. Anal. Appl.* **26**, pp; 127–129, 1992.
- [5] F. K. Chan. *Spacecraft Collision Probability*, IAA. The Aerospace Press, 2008.
- [6] R. Chandramouli, N. Ranganathan. Computing the Bivariate Gaussian Probability Integral, *IEEE Sign. Proc. Letters* **6**, pp. 129–130, 1999.
- [7] A.R Didonato, M.P. Jarnagin, R.K. Hageman. Computation of the integral of the bivariate normal distribution over convex polygons, *SIAM J. Sci. Stat. Comput.* **1**, pp. 179–186.
- [8] H. Federer. The Gauss-Green theorem, *Trans. Amer. Math. Soc.* **58**, pp. 44–76, 1945.
- [9] K. Fujisawa, M. Fukuda, M. Kojima, K. Nakata. Numerical Evaluation of SDPA (Semidefinite Programming Algorithm), In: *High Performance Optimization*, Applied Optimization Volume **33**, Springer, New York, 2000, pp. 267–301.
- [10] A. Genz. Numerical computation of Rectangular bivariate and trivariate Normal and t probabilities, *Statistics and Computing* **14**, pp. 151–160, 2004.
- [11] A. Genz, J. Monahan. A Stochastic Algorithm for High Dimensional Integrals over Unbounded Regions with Gaussian Weight, *J. Comp. Appl. Math.* **112**, pp. 71–81, 1999.
- [12] D. Henrion D., J.B. Lasserre, J. Lofberg. Gloptipoly 3: moments, optimization and semidefinite programming, *Optim. Methods & Softwares* **24**, pp. 761–779, 2009
- [13] Henrion D., Lasserre J.B., Savorgnan C. Approximate volume and integration for basic semi-algebraic sets, *SIAM Review* **51**, pp. 722–743, 2009.
- [14] S. Kotz, N.L. Johnson, D.W. Boyd. Series representation of distribution of quadratic forms in normal variables. I. Central case, *Annals Math. Stat.* **38**, pp. 823–837, 1967.
- [15] S. Kotz, N.L. Johnson, D.W. Boyd. Series Representations of Distributions of Quadratic Forms in Normal Variables II. Non-Central Case, *Annals Math. Stat.* **38**, pp. 838–848, 1967.
- [16] J.B. Lasserre. *An Introduction to Polynomial and Semi-Agebraic Optimization*, Cambridge University Press, Cambridge, 2015.

- [17] J.B. Lasserre. The K-moment problem for continuous functionals, *Trans. Amer. Math. Soc.* **365**, pp. 2489–2504, 2013.
- [18] J.B. Lasserre. *Moments, Positive Polynomials and Their Applications*, Imperial College Press, London, 2009.
- [19] R. P. Patera. General method for calculating satellite conjunction probability, *J. Guidance Control & Dynamics* **24**, pp. 2001.
- [20] H. Ruben. Probability Content of Regions Under Spherical Normal Distributions, IV: The Distribution of Homogeneous and Non-Homogeneous Quadratic Functions of Normal Variables, *Annals Math. Stat.* **33**, pp. 542–570, 1962.
- [21] B. Salvy. D-finiteness: Algorithms and applications. In *Proceedings of the 18th International Symposium on Symbolic and Algebraic Computation*, Beijing, China, July 24–27, 2005, Manuel Kauers Editor, ACM Press, 2005, pp. 2–3.
- [22] R. Serra, D. Arzelier, M. Joldes, J.B. Lasserre, A. Rondepierre, B. Salvy. Fast and accurate computation of orbital collision probability for short-term encounters, *J. Guidance Control & Dynamics* **39**, pp. 1–13, 2016.
- [23] D. Zeilberger. A holonomic systems approach to special functions identities. *J. Comput. Appl. Math.* **32**, pp. 321–368, 1990.
- [24] A Fortran 90 Program for Evaluation of Multivariate Normal and Multivariate t Integrals Over Convex Regions, *J. Stat. Software* **3**, No 4, 1999.
- [25] Numerical Computation of Multivariate Normal and Multivariate t Probabilities over Ellipsoidal Regions, *J. Stat. Software* **6**, No 8, 2001.
- [26] H. Whitney. *Geometric Integration Theory*, Princeton University Press, Princeton, 1957.
- [27] S. Wright. *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia, PA, 1997.

LAAS-CNRS AND INSTITUTE OF MATHEMATICS, UNIVERSITY OF TOULOUSE, LAAS, 7 AVENUE
DU COLONEL ROCHE, 31077 TOULOUSE CÉDEX 4, FRANCE
E-mail address: `lasserre@laas.fr`