



HAL
open science

Volume of sub-level sets of polynomials

Jean B Lasserre

► **To cite this version:**

Jean B Lasserre. Volume of sub-level sets of polynomials. 18th European Control Conference (ECC 2019), Jun 2019, Naples, Italy. pp.1975-1980, 10.23919/ECC.2019.8795995 . hal-02277745

HAL Id: hal-02277745

<https://laas.hal.science/hal-02277745>

Submitted on 3 Sep 2019

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.

Volume of sub-level sets of polynomials

Jean B. Lasserre¹

Abstract—Consider the sub-level set $\mathbf{K} := \{\mathbf{x} : g(\mathbf{x}) \leq 1\}$ of a nonnegative homogeneous polynomial g . We show that its Lebesgue volume $\text{vol}(\mathbf{K})$ can be approximated as closely as desired by solving a sequence of generalized eigenvalue problems with respect to a pair of Hankel matrices of increasing size, whose entries are obtained in closed form. An extension to the non-homogeneous case is also briefly described.

I. INTRODUCTION

Let $g \in \mathbb{R}[\mathbf{x}]_t$ be a nonnegative homogeneous polynomial of degree t with associated sub-level set

$$\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) \leq 1\}. \quad (1)$$

We describe an efficient numerical scheme to approximate its Lebesgue volume $\text{vol}(\mathbf{K})$ as closely as desired. In addition of being an interesting mathematical problem on its own, computing $\text{vol}(\mathbf{K})$ has also a practical interest outside computational geometry. For instance, computing the volume of sub-level sets of polynomials can be helpful to compute the largest region of attraction in control problems as shown in Korda and Henrion [4]. In the homogeneous case it also has a direct link with computing the integral $\int \exp(-g(\mathbf{x}))d\mathbf{x}$, called an *integral discriminant* in Dolotin and Morozov [3] and Morozov and Shakirov [14]. Indeed in [14]:

$$\text{vol}(\mathbf{K}) = \frac{1}{\Gamma(1 + \frac{n+t}{2})} \int_{\mathbb{R}^n} \exp(-g(\mathbf{x})) d\mathbf{x}, \quad (2)$$

and to quote [14], *averaging with exponential weights is an important operation in statistical and quantum physics*. However, and again quoting [14], *despite simply looking, (2) remains terra incognita* (sic). However, for special cases of homogeneous polynomials, the authors in [14] have been able to obtain a closed form expression for (2) (hence equivalently for $\text{vol}(\mathbf{K})$) in terms of algebraic invariants of g . On the other hand, we prove that

$$\text{vol}(\mathbf{K}) = \lim_{d \rightarrow \infty} \lambda_{\min}(\mathbf{A}_d, \mathbf{B}_d), \quad (3)$$

where $\lambda_{\min}(\mathbf{A}_d, \mathbf{B}_d)$ is the smallest generalized eigenvalue of a pair $(\mathbf{A}_d, \mathbf{B}_d)$ of given real Hankel matrices of size $d + 1$. In addition, all entries of both Hankel matrices are easy to obtain in closed-form and the Hankel matrix \mathbf{B}_d depends only on the degree of g . Therefore, in principle the integral (2) can be approximated efficiently and as closely as desired by (linear algebra) eigenvalue routines. To the best of our knowledge this result is quite new and even if we do not provide a closed form expression of (2), this new

characterization as a limit or eigenvalue problems brings new insights.

a) Methodology: Computing (and even approximating) the Lebesgue volume of a convex body is hard (let alone non-convex bodies). Often the only possibility is to use (non deterministic) Monte Carlo type methods which provide an estimate with statistical guarantees. For a discussion on volume computation the interested reader is referred to [5] and the many references therein. However for basic semi-algebraic sets $\mathbf{K} \subset [-1, 1]^n$, Henrion et al. [5] have provided a general methodology to approximate $\text{vol}(\mathbf{K})$. It consists in solving a hierarchy $(\mathbf{Q}_d)_{d \in \mathbb{N}}$ of semidefinite programs¹ of increasing size, whose associated sequence of optimal values $(\rho_d)_{d \in \mathbb{N}}$ is monotone non increasing and converges to $\text{vol}(\mathbf{K})$. An optimal solution of \mathbf{Q}_d is a vector $\mathbf{y} \in \mathbb{R}^{s(2d)}$ (with $s(d) = \binom{n+d}{n}$) whose each coordinate y_α , $\alpha \in \mathbb{N}_{2d}^n$, approximates the α -moment of $\lambda_{\mathbf{K}}$, the restriction to \mathbf{K} of the Lebesgue measure λ on \mathbb{R}^n (and therefore y_0 approximates $\text{vol}(\mathbf{K})$ from above). An optimal solution of the dual semidefinite program \mathbf{Q}_d^* provides the coefficients $(p_\alpha)_{\alpha \in \mathbb{N}_{2d}^n}$ of a polynomial $p \in \mathbb{R}[\mathbf{x}]_{2d}$ which approximates on $[-1, 1]$ and from above, the (indicator) function $\mathbf{x} \mapsto 1_{\mathbf{K}}(\mathbf{x}) = 1$ if $\mathbf{x} \in \mathbf{K}$ and 0 otherwise. In general the convergence $\rho_d \rightarrow \text{vol}(\mathbf{K})$ is slow because of a Gibbs phenomenon² when one approximates the indicator function $1_{\mathbf{K}}$ by continuous functions. In [5] the authors have proposed a “trick” which accelerate drastically the convergence but at the price of loosing the monotone convergence $\rho_d \downarrow \text{vol}(\mathbf{K})$. Another acceleration technique was provided in [10] which still preserves monotone convergence. It uses the fact that moments of $\lambda_{\mathbf{K}}$ satisfy linear equality constraints that follows from Stokes’ theorem.

Recently, Jasour et al. [6] have considered volume computation in the context of risk estimation in uncertain environments. They use an elegant “trick” which reduces computing the n -dimensional volume $\text{vol}(\mathbf{K})$ to computing $\phi([0, 1])$ for a certain pushforward measure ϕ on the real line, whose moments are known. This results in solving the hierarchy of semidefinite programs proposed in [5], but now for measures on the real line as opposed to measures on \mathbb{R}^n . Solving the corresponding hierarchy of dual semidefinite programs amounts to approximate the indicator of an interval on the real line by polynomials of increasing degree, and whose

¹A semidefinite program (SDP) is a conic convex optimization problem with a remarkable modeling power. It can be solved efficiently (in time polynomial in its input size) up to arbitrary precision fixed in advance; see e.g. Anjos and Lasserre [1]

²The Gibbs’ phenomenon appears at a jump discontinuity when one approximates a piecewise C^1 function with a continuous function, e.g. by its Fourier series.

*This work was supported by the European Research Council (ERC) under the ERC-Advanced Grant for the TAMING project

¹Jean B. Lasserre is with LAAS-CNRS and the Institute of Mathematics of Toulouse, France. lasserre@laas.fr

coefficients minimize a linear criterion. On the one hand, it yields drastic computational savings as passing from \mathbb{R}^n to \mathbb{R} is indeed a big and impressive progress. But on the other hand, the (monotone) convergence remains slow as one cannot one cannot apply the acceleration technique based on Stokes' theorem proposed e.g. in [10] because the density of ϕ is not known explicitly. In the examples provided in §III-E for comparison, we can observe this typical (very) slow convergence. However as the problem is now one-dimensional one may then solve many more steps of the resulting hierarchy of semidefinite programs provided that one works with a nice basis of polynomials (e.g. Chebyshev polynomials) to avoid numerical problems as much as possible. Interestingly, pushforward measures were also used in Magron et al. [13] to compute the Lebesgue volume of $f(\mathbf{K})$ for a polynomial mapping $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, but in this case one has to compute moments of the measure in \mathbb{R}^n whose pushforward measure is the Lebesgue measure on $f(\mathbf{K})$, and the resulting computation is still very expensive and limited to modest dimensions.

b) Contribution: We provide a simple numerical scheme to approximate $\text{vol}(\mathbf{K})$ from above with \mathbf{K} as in (1) and when g is positive and homogeneous. We are inspired by the trick of using the pushforward measure in Jasour et al. [6]. The novelty here is that by taking into account the specific nature (homogeneity) of g in (1) we can drastically simplify computations. Indeed, the hierarchy of semidefinite programs defined in [6] can be replaced (and improved significantly) with computing a sequence of scalars $(\tau_d)_{d \in \mathbb{N}}$, where each τ_d is the generalized minimum eigenvalue of two known Hankel matrices of size $d + 1$, whose entries are obtained exactly in closed-form with no numerical error. Therefore there is *no* semidefinite program to solve. Moreover, if one uses the basis of orthonormal polynomials w.r.t. the pushforward measure, then τ_d is now the minimum eigenvalue of a single real symmetric matrix of size d .

II. NOTATION AND DEFINITIONS

Let $\mathbb{R}[\mathbf{x}]$ denote the ring of polynomials in the variables $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbb{R}[\mathbf{x}]_t \subset \mathbb{R}[\mathbf{x}]$ denote the vector space of polynomials of degree at most t , hence of dimension $s(d) = \binom{n+t}{n}$. Let $\Sigma[\mathbf{x}] \subset \mathbb{R}[\mathbf{x}]$ denote the space of polynomials the are sums-of-squares (in short SOS polynomials) and let $\Sigma[\mathbf{x}]_d \subset \mathbb{R}[\mathbf{x}]_{2d}$ denote the space of SOS polynomials of degree at most $2d$. With $\alpha \in \mathbb{N}^n$ and $\mathbf{x} \in \mathbb{R}^n$, the notation \mathbf{x}^α stands for $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. Also for every $\alpha \in \mathbb{N}^n$, let $|\alpha| := \sum_i \alpha_i$ and $\mathbb{N}_d^n := \{\alpha \in \mathbb{N}^n : |\alpha| \leq d\}$.

The support of a Borel measure μ on \mathbb{R}^n is the smallest closed set Ω such that $\mu(\mathbb{R}^n \setminus \Omega) = 0$. Denote by $\mathcal{B}(\mathbf{X})$ the Borel σ -field associated with a topological space \mathbf{X} , and $\mathcal{M}(\mathbf{X})$ the space of finite Borel measures on \mathbf{X} .

Given two real symmetric matrices $\mathbf{A}, \mathbf{C} \in \mathbb{R}^{n \times n}$ denote by $\lambda_{\min}(\mathbf{A}, \mathbf{C})$ the smallest generalized eigenvalue with respect to the pair (\mathbf{A}, \mathbf{C}) , that is, the smallest scalar θ such that $\mathbf{A}\mathbf{x} = \theta \mathbf{C}\mathbf{x}$ for some non zero vector $\mathbf{x} \in \mathbb{R}^n$. When \mathbf{C} is the identity matrix then $\lambda_{\min}(\mathbf{A}, \mathbf{C})$ is just the smallest

eigenvalue of \mathbf{A} . Computing $\lambda_{\min}(\mathbf{A}, \mathbf{C})$ can be done via a pure and efficient linear algebra routine. The notation $\mathbf{A} \succeq 0$ (resp. $\mathbf{A} \succ 0$) stands for \mathbf{A} is positive semidefinite (resp. positive definite).

Moment matrix: Given a real sequence $\phi = (\phi_\alpha)_{\alpha \in \mathbb{N}^n}$, let $\mathbf{M}_d(\phi)$ denote the multivariate (Hankel-type) moment matrix defined by $\mathbf{M}_d(\phi)(\alpha, \beta) = \phi_{\alpha+\beta}$ for all $\alpha, \beta \in \mathbb{N}_d^n$. For instance, in the univariate case $n = 1$, with $d = 2$, \mathbf{M}_2 is the Hankel matrix

$$\mathbf{M}_2(\phi) = \begin{bmatrix} \phi_0 & \phi_1 & \phi_2 \\ \phi_1 & \phi_2 & \phi_3 \\ \phi_2 & \phi_3 & \phi_4 \end{bmatrix}.$$

If $\phi = (\phi_j)_{j \in \mathbb{N}}$ is the moment sequence of a Borel measure ϕ on \mathbb{R} then $\mathbf{M}_d(\phi) \succeq 0$ for all $d = 0, 1, \dots$. Conversely, if $\mathbf{M}_d(\phi) \succeq 0$ for all $d \in \mathbb{N}$, then ϕ is the moment sequence of some finite Borel measure ϕ on \mathbb{R} . The converse result is not true anymore in the multivariate case.

Let ϕ, ν be two finite Borel measures on \mathbb{R} . The notation $\phi \leq \nu$ stands for $\phi(B) \leq \nu(B)$ for all $B \in \mathcal{B}(\mathbb{R})$.

Lemma 2.1: Let ϕ, ν be two finite Borel measures on \mathbb{R} with all moments $\phi = (\phi_j)_{j \in \mathbb{N}}$ and $\nu = (\nu_j)_{j \in \mathbb{N}}$ finite. Then $\phi \leq \nu$ if and only if

$$\mathbf{M}_d(\phi) \preceq \mathbf{M}_d(\nu), \quad \forall d = 0, 1, \dots$$

Proof: Only if part: $\phi \leq \nu$ implies that $\nu - \phi$ with associated sequence $\nu - \phi = (\nu_j - \phi_j)_{j \in \mathbb{N}}$ is a finite Borel measure on \mathbb{R} , and therefore:

$$\mathbf{M}_d(\nu) - \mathbf{M}_d(\phi) = \mathbf{M}_d(\nu - \phi) \succeq 0, \quad d \in \mathbb{N},$$

i.e., $\mathbf{M}_d(\nu) \succeq \mathbf{M}_d(\phi)$ for all $d \in \mathbb{N}$.

If part: If $\mathbf{M}_d(\phi) \preceq \mathbf{M}_d(\nu)$ for all $d \in \mathbb{N}$ then the sequence $\nu - \phi = (\nu_j - \phi_j)_{j \in \mathbb{N}}$ satisfies $\mathbf{M}_d(\nu - \phi) \succeq 0$ for all $d \in \mathbb{N}$. Therefore, the moment sequence $\nu - \phi = (\nu_j - \phi_j)_{j \in \mathbb{N}}$ of the possibly signed measure $\nu - \phi$ is in fact the moment sequence of a finite Borel (positive) measure on \mathbb{R} , and therefore $\nu \geq \phi$. ■

Localizing matrix: Given a real sequence $\phi = (\phi_\alpha)_{\alpha \in \mathbb{N}^n}$ and a polynomial $\mathbf{x} \mapsto p(\mathbf{x}) := \sum_\gamma p_\gamma \mathbf{x}^\gamma$, let $\mathbf{M}_d(p\phi)$ denote the real symmetric matrix defined by:

$$\mathbf{M}_d(p\phi)(\alpha, \beta) = \sum_\gamma p_\gamma \phi_{\alpha+\beta+\gamma}, \quad \alpha, \beta \in \mathbb{N}_d^n.$$

For instance, with $n = 1$, $d = 2$ and $x \mapsto p(x) = x(1 - x)$:

$$\mathbf{M}_2(p\phi) = \begin{bmatrix} \phi_1 - \phi_2 & \phi_2 - \phi_3 & \phi_3 - \phi_4 \\ \phi_2 - \phi_3 & \phi_3 - \phi_4 & \phi_4 - \phi_5 \\ \phi_3 - \phi_4 & \phi_4 - \phi_5 & \phi_5 - \phi_6 \end{bmatrix},$$

also a Hankel matrix.

Lemma 2.2: Let $x \mapsto p(x) = x(1 - x)$.

(i) If a real finite sequence $\phi = (\phi_j)_{j \leq 2d}$ satisfies $\mathbf{M}_d(\phi) \succeq 0$ and $\mathbf{M}_{d-1}(p\phi) \succeq 0$, then there is a measure μ on $[0, 1]$ whose moments $\boldsymbol{\mu} = (\mu_j)_{j \leq 2d}$ match ϕ .

(ii) If a real sequence $\phi = (\phi_j)_{j \in \mathbb{N}}$ satisfies $\mathbf{M}_d(\phi) \succeq 0$ and $\mathbf{M}_d(p\phi) \succeq 0$ for all d , then there is a measure μ on $[0, 1]$ whose moments $\boldsymbol{\mu} = (\mu_j)_{j \in \mathbb{N}}$ match ϕ .

See for instance Lasserre [11] and the many references therein.

Pushforward measure: Let $\mathbf{K} \subset \mathbb{R}^n$ be a Borel set and λ a probability measure on \mathbf{K} . Given a measurable mapping $f : \mathbf{K} \rightarrow \mathbb{R}^p$, the pushforward measure of λ on \mathbb{R}^p w.r.t. f is denoted by $\#\lambda$ and satisfies:

$$\#\lambda(B) := \lambda(f^{-1}(B)), \quad \forall B \in \mathcal{B}(\mathbb{R}^p).$$

In particular, its moments are given by:

$$\#\lambda_\alpha := \int_{\mathbb{R}^p} \mathbf{z}^\alpha \#\lambda(d\mathbf{z}) = \int_{\mathbf{K}} f(\mathbf{x})^\alpha \lambda(d\mathbf{x}), \quad (4)$$

for all $\alpha \in \mathbb{N}^p$.

A version of Stokes' theorem

Let $\Omega \subset \mathbb{R}^n$ be an open subset with boundary $\partial\Omega$ and let $\mathbf{x} \mapsto \mathbf{X}(\mathbf{x})$ be a given vector field. Then under suitable smoothness assumptions,

$$\begin{aligned} \int_{\Omega} \text{Div}(\mathbf{X})f(\mathbf{x}) d\mathbf{x} + \int_{\Omega} \langle \mathbf{X}, \nabla f(\mathbf{x}) \rangle d\mathbf{x} \\ = \int_{\partial\Omega} \langle \vec{n}_{\mathbf{x}}, \mathbf{X} \rangle f(\mathbf{x}) d\sigma, \end{aligned} \quad (5)$$

where $\vec{n}_{\mathbf{x}}$ is the outward pointing normal to Ω at $\mathbf{x} \in \partial\Omega$, and σ is the $(n-1)$ -dimensional Hausdorff measure on the boundary $\partial\Omega$; see e.g. Taylor [15, Proposition 3.2, p. 128].

III. MAIN RESULT

Let $\mathbf{B} := [-1, 1]^n$ and $\mathbf{K} \subset \mathbf{B}$ be as in (1) with $\partial\mathbf{K} \subset \{\mathbf{x} : g(\mathbf{x}) = 1\}$. Let λ be the Lebesgue measure on \mathbf{B} normalized to a probability measure so that $\text{vol}(\mathbf{K}) = 2^n \lambda(\mathbf{K})$. Let g in (1) be a nonnegative and homogeneous polynomial of degree t . That is, $g(\lambda\mathbf{x}) = \lambda^t g(\mathbf{x})$ for all $\lambda \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^n$. Let $\bar{b}_g := \max\{g(\mathbf{x}) : \mathbf{x} \in \mathbf{B}\}$ and notice that as g is nonnegative with $g(0) = 0$, $g(\mathbf{B}) = [0, \bar{b}_g]$.

A. Reduction to the real line

We next follow an elegant idea of Jasour et al. [6], adapted to the present context. It reduces the computation of $\text{vol}(\mathbf{K})$ (in \mathbb{R}^n) to a certain volume computation in \mathbb{R} , by using a particular pushforward measure of the Lebesgue measure λ on \mathbf{B} .

Let $\#\lambda$ be the pushforward on the positive half line of λ , by the polynomial mapping $g : \mathbf{B} \rightarrow [0, \bar{b}_g]$. The support of $\#\lambda$ is the interval $I := [0, \bar{b}_g] \subset \mathbb{R}$. Then in view of (4):

$$\#\lambda_k := \int_I z^k \#\lambda(dz) = \int_{\mathbf{B}} g(\mathbf{x})^k \lambda(d\mathbf{x}), \quad (6)$$

for all $k = 0, 1, \dots$. All scalars $(\#\lambda_k)_{k \in \mathbb{N}}$ can be obtained in closed form as g is a polynomial and λ is the (normalized) Lebesgue measure on \mathbf{B} . Namely, writing the expansion

$$\mathbf{x} \mapsto g(\mathbf{x})^k = \sum_{\alpha \in \mathbb{N}_{kd}^n} g_{k\alpha} \mathbf{x}^\alpha,$$

for some coefficients $(g_{k\alpha})$, one obtains:

$$\#\lambda_k = 2^{-n} \sum_{\alpha \in \mathbb{N}_{kd}^n} g_{k\alpha} \left(\prod_{i=1}^n \frac{(1 - (-1)^{\alpha_i + 1})}{\alpha_i + 1} \right), \quad (7)$$

for $k = 0, 1, \dots$. Next observe that $2^{-n} \text{vol}(\mathbf{K}) = \#\lambda(g(\mathbf{K}))$ and note that $g(\mathbf{K}) = [0, 1]$. Therefore following the recipe introduced in Henrion et al. [5], and with $S := [0, 1] \subset [0, \bar{b}_g]$:

$$\#\lambda(S) = \max_{\phi \in \mathcal{M}(S)} \{\phi(S) : \phi \leq \#\lambda\}. \quad (8)$$

Let ϕ^* be the measure on the real line which is the restriction to S of the pushforward measure $\#\lambda$, i.e.,

$$\phi^*(B) := \#\lambda(B \cap S), \quad \forall B \in \mathcal{B}(\mathbb{R}). \quad (9)$$

Then ϕ^* is the unique optimal solution of (8) and therefore $\phi^*(S) = \#\lambda(S)$; see e.g. Henrion et al. [5]. Then to approximate $\phi^*(S)$ from above, one may solve the hierarchy of semidefinite relaxations:

$$\begin{aligned} \rho_d = \max_{\phi} \{ \phi_0 : 0 \preceq \mathbf{M}_d(\phi) \preceq \mathbf{M}_d(\#\lambda) \\ \mathbf{M}_{d-1}(x(1-x)\phi) \succeq 0 \}, \end{aligned} \quad (10)$$

where $\phi = (\phi_j)_{j \leq 2d}$, $\mathbf{M}_d(\#\lambda)$ is the (Hankel) moment matrix (with moments up to order $2d$) associated with the pushforward measure $\#\lambda$, and $\mathbf{M}_d(\phi)$ (resp. $\mathbf{M}_{d-1}(z(1-z)\phi)$) is the Hankel moment (resp. localizing) matrix (with moments up to order $2d$) associated with the sequence ϕ and the polynomial $x \mapsto p(x) = x(1-x)$; see §II. Indeed (10) is a relaxation of (8) and the sequence $(\rho_d)_{d \in \mathbb{N}}$ is monotone non increasing and converges to $\phi^*(S) = \#\lambda(S)$ from above; see e.g. [5]. The dual of (10) is the semidefinite program

$$\begin{aligned} \rho_d^* = \max_{p \in \mathbb{R}[x]_{2d}} \{ \int p d\#\lambda : p - 1 = \sigma + \psi x(1-x) \\ p, \sigma \in \Sigma[x]_d; \psi \in \Sigma[x]_{d-1} \}, \end{aligned}$$

and if \mathbf{K} has nonempty interior then $\rho_d^* = \rho_d$.

This is the approach advocated by Jasour et al. [6] and indeed this reduction of the initial (Lebesgue) volume computation in \mathbb{R}^n

$$\text{vol}(\mathbf{K}) = \max_{\phi \in \mathcal{M}(\mathbf{K})} \{\phi(\mathbf{K}) : \phi \leq \lambda\} \quad (11)$$

to instead compute $\#\lambda([0, 1])$ (in \mathbb{R}) by solving (8) is quite interesting as it yields drastic computational savings; in fact, solving the multivariate analogues for (11) of the univariate semidefinite relaxations (10) for (8), becomes rapidly impossible even for moderate d , except for problems of modest dimension (say e.g. $n \leq 4$).

However, even in the one-dimensional case the convergence $\rho_d \downarrow \#\lambda(S)$ as $d \rightarrow \infty$, is very slow in general and numerical problems are expected for large values of d . To partially remedy this problem the authors of [6] suggest to express moment and localizing matrices in (10) in the Chebyshev basis (as opposed to the standard monomial basis). This allows to solve a larger number of relaxations but it does not change the typical slow convergence. The trick based on Stokes' theorem used in [10] cannot be used here because the dominating (or reference) measure $\#\lambda$ in (8) is *not* the Lebesgue measure λ anymore (as in (11)). On the other hand the trick to accelerate convergence used in [5] can still be used, that is, in (10) one now maximizes

$L_\phi(x(1-x)) = \phi_1 - \phi_2$ instead of ϕ_0 . If $\phi^d = (\phi_j^d)_{j \leq 2d}$ is an optimal solution of (10) then $\phi_0^d \rightarrow \#\lambda(S)$ as d increases but one loses the monotone convergence from above.

B. Stokes helps

We show that in the particular case where g is positive and homogeneous then one can avoid solving the hierarchy (10) and instead solve a hierarchy of simple generalized eigenvalue problems with *no* semidefinite program to solve and with a much faster convergence.

Let ϕ^* be the Borel measure on \mathbb{R} defined in (9) with moments $\phi^* = (\phi_j^*)_{j \in \mathbb{N}}$. In particular

$$\phi_0^* = \phi^*(S) = \#\lambda(S) = \text{vol}(\mathbf{K}).$$

Define \mathbf{M}_d^* to be the Hankel (moment) matrix with entries:

$$\mathbf{M}_d^*(k, \ell) := \frac{n}{n + (k + \ell - 2)t}, \quad k, \ell = 1, 2, \dots, d + 1 \quad (12)$$

Similarly, define $\mathbf{M}_{d,x(1-x)}^*$ to be the Hankel matrix with entries:

$$\mathbf{M}_{d,x(1-x)}^*(k, \ell) := \frac{n}{n + (k + \ell - 1)t} - \frac{n}{n + (k + \ell)t}, \quad (13)$$

for all $k, \ell = 1, 2, \dots, d + 1$.

Theorem 3.1: For each $d \in \mathbb{N}$, let \mathbf{M}_d^* be as in (12) and let $\mathbf{M}_d(\#\lambda)$ be the Hankel moment matrix associated with $\#\lambda$ (hence with sequence of moments as in (7)). Then :

$$\text{vol}(\mathbf{K}) = \phi_0^* = \lim_{d \rightarrow \infty} \lambda_{\min}(\mathbf{M}_d(\#\lambda), \mathbf{M}_d^*), \quad (14)$$

i.e., ϕ_0^* is the limit of a sequence of minimum generalized eigenvalues for the pairs $(\mathbf{M}_d(\#\lambda), \mathbf{M}_d^*)$, $d \in \mathbb{N}$.

For a proof the interested reader is referred to Lasserre [7].

C. A convergent hierarchy of upper bounds

Therefore to approximate $\text{vol}(\mathbf{K})$ from above, one proceeds as follows. Start with $d = 1$ and then

- Compute all moments of $\#\lambda$ up to order $2d$ by (7).
- Compute $\tau_d := \lambda_{\min}(\mathbf{M}_d(\#\lambda), \mathbf{M}_d^*)$
- set $d = d + 1$ and repeat.

This produces the required monotone sequence of upper bounds $(\tau_d)_{d \in \mathbb{N}}$ on ϕ_0^* , which converges to $\phi_0^* = 2^{-n} \text{vol}(\mathbf{K})$ as d increases. Next, from the proof in [7] it turns out that $\phi_0^* \mathbf{M}_d^*$ (resp. $\phi_0^* \mathbf{M}_{d,x(1-x)}^*$) is the moment (resp. localizing) matrix associated with the measure ϕ^* on $[0, 1]$. Recall that:

$$\begin{aligned} \tau_d &:= \lambda_{\min}(\mathbf{M}_d(\#\lambda), \mathbf{M}_d^*) \\ &= \max \{ \tau : \tau \mathbf{M}_d^* \preceq \mathbf{M}_d(\#\lambda) \}, \end{aligned} \quad (15)$$

The following result shows that $\tau_d \leq \rho_d$.

Proposition 3.2: For each $d \in \mathbb{N}$, let ρ_d (resp. τ_d) be as in (10) (resp. (15)). Then $\rho_d \geq \tau_d$.

Proof: Consider the sequence $\mu = (\mu_j)_{j \leq 2d}$ defined by:

$$\mu_j = \tau_d \frac{n}{n + jt}, \quad j \leq 2d,$$

so that $\mu_0 = \tau_d$ and $\tau_d \mathbf{M}_d^* = \mathbf{M}_d(\mu)$. Then from (15), $0 \preceq \mathbf{M}_d(\mu) \preceq \mathbf{M}_d(\#\lambda)$. Similarly $\tau_d \mathbf{M}_{d-1,x(1-x)} =$

$\mathbf{M}_{d-1}(x(1-x)\mu) \succeq 0$. In other words, the sequence μ is a feasible solution of (10), which implies $\mu_0 (= \tau_d) \leq \rho_d$. ■ Hence the above eigenvalue procedure (which involves no semidefinite program) provides a monotone sequence of upper bounds on $\phi_0^* = \text{vol}(\mathbf{K})$ that is always better than the corresponding sequence of upper bounds $(\rho_d)_{d \in \mathbb{N}}$ obtained by solving the hierarchy of semidefinite relaxations (10) proposed in [6]. Notice also that the matrix \mathbf{M}_d^* depends only on the degree of g and not on g itself.

Remark 3.3: (i) Alternatively one may also use an orthonormal basis associated with the sequence of moments $(n/(n + jt))_{j \in \mathbb{N}}$. In this case the new moment matrix $\hat{\mathbf{M}}_d^*$ expressed in this basis is the identity matrix and τ_d is now the minimum eigenvalue of the new moment $\hat{\mathbf{M}}_d(\#\lambda)$ expressed in this basis. Notice that the orthonormal basis does not depend on g (only on its degree).

(ii) Another possibility is to simply use the basis of Chebyshev polynomials. Then computing τ_d (still a generalized eigenvalue problem) involves matrices with much better numerical conditioning.

D. Some comments

a) The moments of $\#\lambda$: Theoretically the (off line) computation of moments of the pushforward measure $\#\lambda$ is not difficult since they can be obtained in closed form via (7). However, in practice such computation can be demanding if g has relatively high degree and/or n is large. Indeed, g^k is a homogeneous polynomial of degree tk in n variables, with possibly many non zero coefficients out of potentially $\binom{n-1+tk}{tk}$ in the monomial basis. Therefore for large values of d , some care is needed to obtain $\mathbf{M}_d(\#\lambda)$ efficiently and accurately; see Remark 3.3.

b) The box \mathbf{B} : As for the method in [5], the smaller is the box \mathbf{B} that contains \mathbf{K} , the better is the efficiency of the method. The smallest box is $\mathbf{B} := \prod_{i=1}^n [\underline{w}_i, \overline{w}_i]$, with:

$$\overline{w}_i := \max \{x_i : \mathbf{x} \in \mathbf{K}\}; \quad \underline{w}_i := \min \{x_i : \mathbf{x} \in \mathbf{K}\} \quad (16)$$

for all $i = 1, \dots, n$. But of course upper bounds on \overline{w}_i (resp. lower bounds on \underline{w}_i) are also fine, and in particular those obtained by solving the first semidefinite relaxation of the Moment-SOS hierarchy described in [12], applied to the above polynomial optimization problems.

c) A variant: One may also consider a variant which consists of using $\tilde{g} := g^{1/p}$ with p integer in lieu of g . Indeed $\mathbf{K} = \{\mathbf{x} : g \leq 1\} = \{\mathbf{x} : \tilde{g} \leq 1\}$. However the function \tilde{g} which is homogeneous of degree t/p is not a polynomial any more and therefore the moments $(\#\lambda_k)_{k \in \mathbb{N}}$ in (6) are not available in closed form. On the other hand as the integral (6) deals with Lebesgue measure on a box, it can be approximated by using Monte-Carlo methods or by Gauss cubatures. Surprisingly, it turns out that this variant can be quite efficient (especially with $p = kt$, $k \in \mathbb{N}$, which makes \tilde{g} homogeneous of degree $1/k$).

E. Some numerical examples

a) The Euclidean ball: To compare τ_d computed in (15) with ρ_d computed in solving the semidefinite program (10)

proposed in [6], we have considered a favorable case for (10). We chose \mathbf{K} to be the Euclidean unit ball $\{\mathbf{x} : \|\mathbf{x}\| \leq 1\}$ with Lebesgue volume $\pi^{n/2}/\Gamma(1+n/2)$ and the ball \mathbf{B} that contains \mathbf{K} is the smallest one, i.e., $\mathbf{B} = [-1, 1]^n$. Indeed, the smaller is the ball \mathbf{B} , the better are the upper bounds ρ_d in (10).

Start with $n = 2$, $g = \|\mathbf{x}\|^2 = x_1^2 + x_2^2$, and $\mathbf{B} = [-1, 1]^2$, so that $\text{vol}(\mathbf{K}) = \pi$. Then:

$$\mathbf{M}_1^* = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix}; \quad \mathbf{M}_1(\#\lambda) = \begin{bmatrix} 1 & 2/3 \\ 2/3 & 28/45 \end{bmatrix}$$

This yields $4 \cdot \tau_1 \approx 3.20$ which is already a good upper bound on π whereas $4 \cdot \rho_1 = 4$. Next with $d = 2$ we obtain $4 \cdot \tau_2 \approx 3.1440$ while $4 \cdot \rho_2 = 3.8928$. Hence $4\tau_2$ already provides a very good upper bound on π with only moments of order 4. To appreciate the difference in speed of convergence between ρ_d and τ_d , Table I displays both values τ_d and ρ_d in the case of $n = 4$ variables and $d = 1, \dots, 5$. While the convergence $\tau_d \rightarrow 4.9348$ is quite fast with a relative error of 0.03% at step $d = 5$, the convergence $\rho_d \rightarrow 4.9348$ is extremely slow as $\rho_5 \approx 8.499$ only; see Figure 1.

d	$d=1$	$d=2$	$d=3$	$d=4$	$d=5$
ρ_d	12.19	11.075	9.163	8.878	8.499
τ_d	6.839	5.309	5.001	4.945	4.936

TABLE I
 $n = 4$, $\rho^* = 4.9348$; ρ_d VERSUS τ_d

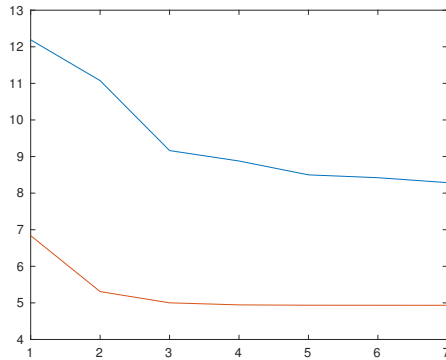


Fig. 1. $n = 4$; Comparing τ_d (red below) with ρ_d (blue above)

We next provide results for the same problem but now in larger dimensions $n = 8, 9$ in Table III and Table IV respectively. From inspection we can observe a fast and regular decrease in the value $2^n \tau_d$ as d increases, and similarly for the relative error.

d	$d=2$	$d=3$	$d=4$	$d=5$	$d=6$
τ_d	5.309	5.001	4.945	4.936	4.935
error	7.58%	1.35%	0.22%	0.03%	0.004%

TABLE II
 $n = 4$, $\rho^* = 4.9348$; τ_d AND RELATIVE ERROR

d	$d=4$	$d=5$	$d=6$	$d=7$	$d=8$
$2^n \tau_d$	5.569	4.639	4.272	4.133	4.083
error	37%	14%	5.26%	1.83%	0.60%

TABLE III
 $n = 8$, $\rho^* = 4.0587$; τ_d AND RELATIVE ERROR

d	$d=4$	$d=5$	$d=6$	$d=7$	$d=8$
$2^n \tau_d$	5.935	4.413	3.764	3.485	3.369
error	79%	33.8%	14.1%	5.6%	2.15%

TABLE IV
 $n = 9$, $\rho^* = 3.298$; τ_d AND RELATIVE ERROR

For $n = 10$ and $d = 8$, we have encountered numerical problems because the Hankel matrix $\mathbf{M}_8(\#\lambda)$ is ill-conditioned and then one should use another basis of polynomials in which to express the matrices \mathbf{M}_8^* and $\mathbf{M}_8(\#\lambda)$; see Remark 3.3.

b) *More examples:* : With $n = 2$ consider the set:

$$\mathbf{K}_c := \{\mathbf{x} \in \mathbb{R}^2 : x_1^4 + x_2^4 - c x_1^2 x_2^2\}, \quad c \in [0, 2),$$

where c is fixed, and we consider the values $c = 0, 1, 1.5$. In Table V we display results for τ_d computed in (15), and for the optimal value ρ_d of the semidefinite relaxations described in [10] for solving (11) (which also include Stokes constraints to accelerate convergence). The box \mathbf{B} was chosen according to (16) and the bounds $\underline{w}_i = -\overline{w}_i$ are tight.

As one can see in Table V, with $d = 8$, ρ_d and τ_d are very close to each other (and close to $\text{vol}(\mathbf{K})$ which is not known exactly). Also, remarkably, with $d = 1$ the bound τ_d is already quite good since the relative error between τ_1 and τ_8 is about 0.36%, 0.07%, and 1.58% respectively.

Again it is important to emphasize that for the same index d , in (15) we solve a generalized eigenvalue problem with two Hankel matrices of size $d + 1$ (independently of the number of variables n), whereas to obtain ρ_d in [10] we need solve a semidefinite program involving a moment matrix of size $\binom{n+d}{n}$. (For relatively large d , the latter method in [10] is possible only for small dimension n .)

$c=0.0$; $\text{vol}(\mathbf{K}) \approx 3.708$					
d	$d=1$	$d=2$	$d=4$	$d=6$	$d=8$
τ_d	3.7221	3.7217	3.7151	3.7093	3.7087
ρ_d		5.4553	3.9402	3.7256	3.7092
$c=1$; $\text{vol}(\mathbf{K}) \approx 4.313$					
d	$d=1$	$d=2$	$d=4$	$d=6$	$d=8$
τ_d	4.3164	4.3161	4.3158	4.3145	4.3134
ρ_d		4.6656	4.3571	4.3157	4.3131
$c=1.5$; $\text{vol}(\mathbf{K}) \approx 4.947$					
d	$d=1$	$d=2$	$d=4$	$d=6$	$d=8$
τ_d	5.0315	5.0294	4.9838	4.9597	4.9519
ρ_d		6.0516	5.1260	4.9773	4.9494

TABLE V
 $n = 2$, τ_d VERSUS ρ_d FOR $c = 0, 1, 1.5$

Finally, we have also considered the variant in §III-D(c) which consists of using the homogeneous function $\tilde{g} = g^{1/kt}$ of degree $1/k$ in lieu of g . In Table VI one can see that with

only $d = 1$ and $d = 2$ one obtains very good approximations for $k = 1, \dots, 4$, significantly better than with the original polynomial g .

c=0				
k/c	k = 1	k = 2	k = 3	k = 4
d = 1	3.7089	3.7098	3.7101	3.7102
d = 2	3.7077	3.7080	3.7084	3.7087
c=1				
	k = 1	k = 2	k = 3	k = 4
d = 1	4.3136	4.3138	4.3136	4.3136
d = 2	4.3123	4.3128	d = 4.3131	4.3133
c=1.5				
	k = 1	k = 2	k = 3	k = 4
d = 1	4.9477	4.9483	4.9489	4.9493
d = 2	4.9476	4.9473	4.9473	4.9475

TABLE VI
 $n = 2$, VARIANT $\tilde{g} = g^{1/kt}$ FOR $c = 0, 1, 1.5$

IV. EXTENSION TO THE NON-HOMOGENEOUS CASE

‘ Let $\mathbf{K} \subset \mathbb{R}^n$ (with $\mathbf{K} \subset (-1, 1)^n$ possibly after re-scaling) be defined by:

$$\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n : 0 \leq g(\mathbf{x}) \leq 1\}, \quad (17)$$

where $g \in \mathbb{R}[\mathbf{x}]_t$ is not necessarily homogeneous. Write

$$\mathbf{x} \mapsto g(\mathbf{x}) = \sum_{k=1}^t g_k(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^n,$$

where each $g_k \in \mathbb{R}[\mathbf{x}]_k$ is homogeneous of degree k , and consider the mapping $G : \mathbf{B} \rightarrow \mathbb{R}^t$,

$$\mathbf{x} \mapsto G(\mathbf{x}) = (g_k(\mathbf{x}))_{k=1}^t \in \mathbb{R}^t, \quad \mathbf{x} \in \mathbb{R}^n.$$

Let $\#\lambda \in \mathcal{M}(\mathbb{R}^t)$ be the pushforward of the Lebesgue probability measure λ on \mathbf{B} . Its support is $G(\mathbf{B})$ and as before all moments of $\#\lambda$ can be obtained easily in closed-form. In addition $\mathbf{G}(\mathbf{K}) = \{\mathbf{z} \in G(\mathbf{B}) : 0 \leq \sum_{k=1}^t z_t \leq 1\} =: S$. Then:

$$2^n \text{vol}(\mathbf{K}) = \#\lambda(S) = \sup_{\phi \in \mathcal{M}(S)} \{\phi(S) : \phi \leq \#\lambda\}. \quad (18)$$

So one has replaced the n -dimensional Lebesgue volume computation $\lambda(\mathbf{K})$ with the t -dimensional volume computation $\#\lambda(S)$, which is potentially interesting whenever $t \ll n$. Then one may use homogeneity of the g_k 's to include additional linear moment equality constraints on ϕ (again from an application of Stokes' theorem). Those additional constraints are crucial to accelerate the otherwise slow convergence of the SDP-relaxations associated with solving (18). However, this time we cannot reduce the initial problem to an eigenvalue problem as in the homogeneous case. For more details the interested reader is referred to [7].

V. CONCLUSION

We have presented a new methodology to approximate as closely as desired the Lebesgue volume of the sub-level set $\{\mathbf{x} : g(\mathbf{x}) \leq 1\}$ of a positive homogeneous n -variate polynomial g . The novelty with respect to [6] is that by using Stokes' theorem and exploiting the homogeneity of g , we are able to reduce the problem to solving a hierarchy of generalized eigenvalue problems for Hankel matrices of increasing size, with *no* optimization involved. The efficiency of the proposed extension to the non-homogeneous case remains to be validated, a topic of further investigation.

REFERENCES

- [1] Anjos M., Lasserre J.B. (Eds.), Handbook of Semidefinite, Conic and Polynomial Optimization. New York: Springer, 2012.
- [2] C.F. Denkl, Y. Xu, Orthogonal Polynomial in Several Variables. Cambridge: Cambridge University Press, UK, 2014.
- [3] V. Dolotin and A. Morozov, Introduction to Non-Linear Algebra. Singapore: World Scientific, 2007.
- [4] D. Henrion, M. Korda, Convex computation of the region of attraction of polynomial control systems, IEEE Trans. Auto. Control, vol 59, No (2), pp. 297–312, 2014.
- [5] D. Henrion, J.B. Lasserre, and C. Savorgnan, Approximate volume and integration for basic semialgebraic sets, SIAM Review 51(2009), pp. 722–743.
- [6] A. Jasour, A. Hofmann, and B.C. Williams, Moment-Sum-Of-Squares Approach For Fast Risk Estimation In Uncertain Environments, arXiv:1810.01577, 2018.
- [7] J.B. Lasserre, Volume of sub-level sets of homogeneous polynomials, arXiv:1810.10224.
- [8] J.B. Lasserre, Level sets and non Gaussian integrals of positively homogeneous functions, Int. Game Theory Review, vol 17, No 1, 2015.
- [9] J.B. Lasserre, A generalization of Löwner-John's ellipsoid theorem, Math. Program., vol 152, pp. 559–591, 2015.
- [10] J.B. Lasserre, Computing Gaussian & exponential measures of semi-algebraic sets, Adv. Appl. Math., vol 91, pp. 137–163, 2017.
- [11] J.B. Lasserre, Moments, Positive Polynomials and Their Applications. London: Imperial College Press, 2010.
- [12] J.B. Lasserre. The Moment-SOS hierarchy, Proceedings of the ICM 2018, Rio de Janeiro, 2018.
- [13] V. Magron, D. Henrion, and J.B. Lasserre, Semidefinite approximations of projections and polynomial images of semi-algebraic sets, SIAM J. Optim., vol 25, pp. 2143–2164, 2015
- [14] A. Morozov and S. Shakirov, Introduction to integral discriminants, J. High Energy Phys., vol 12, No. 12, 2009
- [15] M.E. Taylor, Partial Differential Equations: Theory. New York: Springer-Verlag, 1996.