

Convergence rates of moment-sum-of-squares hierarchies for volume approximation of semialgebraic sets

Milan Korda¹, Didier Henrion^{2,3,4}

Draft of December 12, 2016

Abstract

Moment-sum-of-squares hierarchies of semidefinite programs can be used to approximate the volume of a given compact basic semialgebraic set K . The idea consists of approximating from above the indicator function of K with a sequence of polynomials of increasing degree d , so that the integrals of these polynomials generate a convergence sequence of upper bounds on the volume of K . We show that the asymptotic rate of this convergence is at least $O(1/\log \log d)$.

Keywords: moment relaxations, polynomial sums of squares, convergence rate, semidefinite programming, approximation theory.

1 Introduction

The moment-sum-of-squares hierarchy (also known as the Lasserre hierarchy) of semidefinite programs was originally introduced in the context of polynomial optimization, and later on extended in various directions, see e.g. [5] for an introductory survey and [4] for a comprehensive treatment.

In our companion note [3] we have studied the convergence rate of this hierarchy applied to optimal control. In the present note, we carry out the convergence analysis of the hierarchy tailored in [2] for approximating the volume (and all the moments of the uniform measure) of a given semi-algebraic set.

Let us first recall the main idea of [2]. Given a compact basic semi-algebraic set

$$K := \{x \in \mathbb{R}^n : g_i^K(x) \geq 0, i = 1, \dots, n_K\} \quad (1)$$

¹Department of Mechanical Engineering, University of California, Santa Barbara, CA 93106-5070. milan.korda@engineering.ucsb.edu.

²CNRS; LAAS; 7 avenue du colonel Roche, F-31400 Toulouse; France. henrion@laas.fr

³Université de Toulouse; LAAS; F-31400 Toulouse; France.

⁴Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, CZ-16626 Prague, Czech Republic.

described by polynomials $(g_i^K)_i \subset \mathbb{R}[x]$, we want to approximate numerically its volume

$$\text{vol } K := \int_K dx.$$

We assume that we are also given a compact basic semi-algebraic set

$$X := \{x \in \mathbb{R}^n : g_i^X(x) \geq 0, i = 1, \dots, n_X\} \quad (2)$$

described by polynomials $(g_i^X)_i \subset \mathbb{R}[x]$, such that K is contained in X , which is in turn contained in the unit Euclidean ball:

$$K \subset X \subset B_n := \{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}$$

and such that the moments of the uniform (a.k.a. Lebesgue) measure over X are known analytically or are easy to compute, or equivalently, such that the integral of a given polynomial over X can be evaluated easily. This is the case for instance if X is a box or a ball, e.g. $X = B_n$. The set K , on the contrary, can be a complicated (e.g., non-convex or disconnected) set for which the volume (as well as the higher Lebesgue moments) are hard to compute. In addition we invoke the following technical assumption:

Assumption 1 *The origin belongs to the interior of K , i.e., $0 \in \text{int } K$.*

This assumption can be satisfied whenever the set K has a non-empty interior by translating the set such that the origin belongs to the interior.

For notational simplicity, we let $g_0^K := g_0^X := 1$. Moreover, since K respectively X are included in the unit ball B_n , we assume without loss of generality that one of the g_i^K resp. g_i^X , $i > 0$, is equal to $1 - \sum_{k=1}^n x_k^2$. Attached to the set K is a specific convex cone called truncated quadratic module defined by

$$Q_d(K) := \left\{ \sum_{i=0}^{n_K} g_i^K s_i^K : s_i^K \in \Sigma[x], \deg(g_i^K s_i^K) \leq 2 \lfloor \frac{d}{2} \rfloor \right\}$$

where $\Sigma[x] \subset \mathbb{R}[x]$ is the semidefinite representable convex cone of polynomials that can be written as sums of squares of polynomials. Attached to the outer bounding set X , we define analogously the truncated quadratic module

$$Q_d(X) := \left\{ \sum_{i=0}^{n_X} g_i^X s_i^X : s_i^X \in \Sigma[x], \deg(g_i^X s_i^X) \leq 2 \lfloor \frac{d}{2} \rfloor \right\}.$$

Consider then the following hierarchy of convex optimization problems indexed by degree d :

$$v_d^* := \inf_{p \in \mathbb{R}[x]_d} \int_X p \quad \text{s.t.} \quad \begin{aligned} p - 1 &\in Q_d(K) \\ p &\in Q_d(X) \end{aligned} \quad (3)$$

where the decision variable belongs to $\mathbb{R}[x]_d$, the finite-dimensional vector space of polynomials of total degree less than or equal to d . In problem (3) the objective function is a linear function, the integral of p over X , for which we use the simplified notation

$$\int_X p := \int_X p(x) dx$$

here and in the remainder of the document. Since we assume that X is chosen such that the moments of the uniform measure are known, the objective function is an explicit linear function of the coefficients of p expressed in any basis of $\mathbb{R}[x]_d$. The truncated quadratic modules $Q_d(K)$ and $Q_d(X)$ are semidefinite representable, so problem (3) translates to a finite-dimensional semidefinite program. Moreover, it is shown in [2, Theorem 3.2] that the sequence $(v_d^*)_{d \in \mathbb{N}}$ converges from above to the volume of K , i.e.

$$v_d^* \geq v_{d+1}^* \geq \lim_{d \rightarrow \infty} v_d^* = \text{vol } K.$$

The goal of this note is to analyze the rate of convergence of this sequence.

2 Convergence rate

Let us first recall several notions from approximation theory. Given a bounded measurable function $f : X \rightarrow \mathbb{R}$ we define

$$\omega_f^x(t) := \sup_{\substack{y \in X \\ \text{s.t. } \|y - x\|_2 \leq t}} |f(y) - f(x)|$$

as the modulus of continuity of f at a point $x \in X$. In addition, we define

$$\bar{\omega}_f(t) := \int_X \omega_f^x(t) dx$$

as the averaged modulus of continuity of f on X . Finally we define

$$e_f(d) := \inf_{\substack{p \in \mathbb{R}[x]_d \\ \text{s.t. } p \geq f \text{ on } X}} \int_X (p - f)$$

as the error of the best approximation to f from above in the L_1 norm by polynomials of degree no more than d .

Theorem 1 ([1]) *For any bounded measurable function f and for all $d \in \mathbb{N}$ it holds*

$$e_f(d) \leq \bar{c} \bar{\omega}_f(1/d) \tag{4}$$

for some constant \bar{c} depending only on f .

Theorem 1 links the L_1 approximation error from above to the averaged modulus of continuity. Now we derive an estimate on the averaged modulus of continuity when $f = I_K$, the indicator function of K equal to 1 on K and 0 outside.

Lemma 1 For all $t \in [0, 1]$ it holds

$$\bar{\omega}_{I_K}(t) \leq c t$$

for some constant c depending only on K and X .

Proof: Let ∂K denote the boundary of K and let

$$\partial K(t) := \{x \in X : \min_{y \in \partial K} \|x - y\|_2 \leq t\}$$

denote the set of all points whose Euclidean distance to the boundary of X is no more than t . Then we have $\omega_{I_K}^x(t) = 0$ for all $x \notin \partial K(t)$. Hence

$$\bar{\omega}_{I_K}(t) = \int_X \omega_{I_K}^x(t) dx = \int_{\partial K(t)} \omega_{I_K}^x(t) dx \leq \int_{\partial K(t)} 1 dx = \text{vol} \partial K(t).$$

Therefore we need to bound the volume of $\partial K(t)$. In order to do so, we decompose

$$\partial K = \bigcup_{i=1}^{n_\partial} \partial K_i$$

where each ∂K_i is a smooth manifold of dimension no more than $n - 1$ and n_∂ is finite. Defining

$$\partial K_i(t) := \{x \in X : \min_{y \in \partial K_i} \|x - y\|_2 \leq t\},$$

Weyl's tube formula [8] yields the bound

$$\text{vol} \partial K_i(t) \leq c_i t$$

for all $t \in [0, 1]$ and for some constant c_i depending on K and X only. Hence we get

$$\text{vol} \partial K(t) \leq \sum_{i=1}^{n_\partial} \text{vol} \partial K_i(t) \leq \sum_{i=1}^{n_\partial} c_i t = c t$$

where $c := \sum_{i=1}^{n_\partial} c_i$ depends on K and X only, as desired. \square

This leads to the following crucial estimate on the rate of convergence of the best polynomial approximation from above to the indicator function of K :

Theorem 2 For all $d \in \mathbb{N}$ it holds

$$e_{I_K}(d) \leq \frac{c_1}{d} \tag{5}$$

for some constant c_1 depending only on K and X .

Proof: Follows immediately from Theorem 1 and Lemma 1 with $c_1 := \bar{c} c$. \square

We need the following assumption, which we conjecture is without loss of generality:

Assumption 2 (Finite Gibbs phenomenon) For $d \in \mathbb{N}$ there is a sequence

$$p_d \in \underset{p \in \mathbb{R}[x]_d}{\operatorname{argmin}} \int_X (p - I_K) \\ \text{s.t. } p \geq I_K \text{ on } X$$

and a finite constant c_G such that $\max_{x \in X} p_d(x) \leq c_G$.

In addition to the results from approximation theory, we need the following result which is a consequence of the fundamental result of [6] and of [3, Corollary 1]:

Theorem 3 Let $p \in \mathbb{R}[x]$ be strictly positive on a set $S \in \{K, X\}$ and let Assumption 1 hold. Then $p \in Q_d(S)$ whenever

$$d \geq c_2^S \exp \left[\left(k(\deg p)(\deg p)^2 n^{\deg p} \frac{\max_{x \in S} p(x)}{\min_{x \in S} p(x)} \right)^{c_2^S} \right]$$

for some constant c_2^S depending only the algebraic description of S and with $k(d) = 3^{d+1} r^d$, where

$$r := \frac{1}{\sup\{s > 0 \mid [-s, s]^n \subset K\}}, \quad (6)$$

which is the reciprocal value of the length of the side of the largest box centered around the origin included in K^1 .

Now we are in position to prove our main result:

Theorem 4 For all $\varepsilon > 0$ it holds $v_d^* - \operatorname{vol} K < \varepsilon$ whenever

$$d \geq c_2 \exp \left[\left(\frac{3c_3(\varepsilon)^2 (3rn)^{c_3(\varepsilon)} (2c_G \operatorname{vol} B_n + \varepsilon)}{\varepsilon} \right)^{c_2} \right] = O \left(\exp \left[\frac{1}{\varepsilon^{3c_2}} (3rn)^{\frac{2c_1}{\varepsilon}} \right] \right) \quad (7)$$

where $c_3(\varepsilon) := \lceil 2c_1/\varepsilon \rceil$, r is defined in (6) and the constants c_1 and c_2 depend only on X and K .

Proof: Let $\varepsilon > 0$ be fixed and decompose $\varepsilon = \varepsilon_1 + \varepsilon_2$ with $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$. Let $p \in \mathbb{R}[x]_d$ be any polynomial feasible in (3) for some $d \geq 1$. Then we have

$$v_d^* - \operatorname{vol} K = v_d^* - \int_X I_K \leq \int_X (p - I_K) \leq \int_X (\tilde{p}_{\tilde{d}} - I_K) + \int_X |p - \tilde{p}_{\tilde{d}}|$$

where $\tilde{p}_{\tilde{d}}$ is an optimal polynomial approximation to I_K from above of degree \tilde{d} satisfying Assumption 2, i.e., $\int_X (\tilde{p}_{\tilde{d}} - I_K) \leq c_1/\tilde{d}$ and $\max_{x \in X} \|\tilde{p}_{\tilde{d}}(x)\| \leq c_G$. Selecting $\tilde{d} = \lceil c_1/\varepsilon_1 \rceil$ we obtain

$$v_d^* - \operatorname{vol} K \leq \varepsilon_1 + \int_X |f - \tilde{p}_{\tilde{d}}|$$

¹By assumption $K \subset X$ and hence it is sufficient to take K in the definition of r in (6) rather than S .

Hence it suffices to find a $d \geq 0$ and a polynomial p feasible in (3) for this d such that $\int_X |p - \tilde{p}_d| \leq \varepsilon_2$. Let $p := \tilde{p}_d + \varepsilon_2/\text{vol } B_n$, then

$$\int_X |p - \tilde{p}_d| = \int_X \frac{\varepsilon_2}{\text{vol } B_n} \leq \frac{\varepsilon_2}{\text{vol } B_n} \int_{B_n} = \varepsilon_2.$$

In addition, since $\tilde{p}_d \geq I_K$ on X , we have $p \geq \varepsilon_2/\text{vol } B_n > 0$ on X and $p - 1 \geq \varepsilon_2/\text{vol } B_n > 0$ on K and hence p is feasible in (3) for some $d \geq 0$ [7]. To bound the degree d we invoke Theorem 3 on the two constraints of problem (3). Set $c_2 := \max(c_2^X, c_2^K)$, where the constants c_2^X and c_2^K are the constants from Theorem 3 (note that these constants depend only on the algebraic description of the sets, not on the polynomial p). Since $\max_{x \in K} |p(x) - 1| \leq \max_{x \in X} |p(x)| \leq c_G + \varepsilon_2/\text{vol } B_n$ and $d_p(\varepsilon_1) := \deg p = \tilde{d} = \lceil c_1/\varepsilon_1 \rceil$, Theorem 3 applied to the two constraints of the problem (3) yields

$$d \geq c_2 \exp \left[\left(\frac{k(d_p(\varepsilon_1)) d_p(\varepsilon_1)^2 n^{d_p(\varepsilon_1)} (c_G + \varepsilon_2/\text{vol } B_n)}{\varepsilon_2/\text{vol } B_n} \right)^{c_2} \right]$$

with $k(d) = 3^{d+1} r^d$ and r is defined in (6). Setting $\varepsilon_1 = \varepsilon_2 = \varepsilon/2$ and $c_3(\varepsilon) := \lceil 2c_1/\varepsilon \rceil$, we obtain (7). □

Corollary 1 *It holds $v_d^* - \text{vol } K = O(1/\log \log d)$.*

Proof: Follows by inverting the asymptotic expression $O(\exp[\frac{1}{\varepsilon^{3c_2}} (3rn)^{\frac{2c_1}{\varepsilon}}])$ using the fact that $(3rn)^{\frac{4c_1}{\varepsilon}} \geq \frac{1}{\varepsilon^{3c_2}} (3rn)^{\frac{2c_1}{\varepsilon}}$ for small ε . □

3 Concluding remarks

Our doubly logarithmic asymptotic convergence rate is likely to be pessimistic, even though we suspect the actual convergence rate to be sublinear in d , see for example [2, Fig. 4.5] which reports on the values of the moment-sums-of-squares hierarchy for an elementary univariate interval length estimation problem. In this case we could compute numerically the values for degrees up to 100 thanks to the use of Chebyshev polynomials. Beyond univariate polynomials, it is however difficult to study experimentally the convergence rate since the number of variables in the semidefinite program grows polynomially in the degree d , but with an exponent equal to the dimension of the set K .

References

- [1] E. S. Bhatta. Interpolating operators for multiapproximation. *Journal of Mathematics and Statistics*, 6:240-245, 2010.
- [2] D. Henrion, J. B. Lasserre, C. Savorgnan. Approximate volume and integration for basic semialgebraic sets. *SIAM Review* 51(4):722-743, 2009.

- [3] M. Korda, D. Henrion, C. N. Jones. Convergence rates of moment-sum-of-squares hierarchies for optimal control problems. [arXiv:1609.02762](https://arxiv.org/abs/1609.02762), 9 Sept. 2016.
- [4] J. B. Lasserre. Moments, positive polynomials and their applications. Imperial College Press, London, UK, 2010.
- [5] M. Laurent. Sums of squares, moment matrices and polynomial optimization. In M. Putinar, S. Sullivan (eds.). Emerging applications of algebraic geometry, Vol. 149 of IMA Volumes in Mathematics and its Applications, Springer, Berlin, 2009.
- [6] J. Nie, M. Schweighofer. On the complexity of Putinar's Positivstellensatz. *Journal of Complexity* 23:135-150, 2007.
- [7] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, 42:969-984, 1993.
- [8] H. Weyl. On the volume of tubes. *American Journal of Mathematics*, 61:461-472, 1939.