Stokes, Gibbs and volume computation of semi-algebraic sets

Matteo Tacchi[1,2], Jean Bernard Lasserre[1,3,*], Didier Henrion[1,4]

September 23, 2020

Abstract

We consider the problem of computing the Lebesgue volume of compact basic semi-algebraic sets. In full generality, it can be approximated as closely as desired by a converging hierarchy of upper bounds obtained by applying the Moment-SOS (sums of squares) methodology to a certain infinite-dimensional linear program (LP). At each step one solves a semidefinite relaxation of the LP which involves pseudo-moments up to a certain degree. Its dual computes a polynomial of same degree which approximates from above the discontinuous indicator function of the set, hence with a typical Gibbs phenomenon which results in a slow convergence of the associated numerical scheme. Drastic improvements have been observed by introducing in the initial LP additional linear moment constraints obtained from a certain application of Stokes' theorem for integration on the set. However and so far there was no rationale to explain this behavior. We provide a refined version of this extended LP formulation. When the set is the smooth super-level set of a single polynomial, we show that the dual of this refined LP has an optimal solution which is a continuous function. Therefore in this dual one now approximates a continuous function by a polynomial, hence with no Gibbs phenomenon, which explains and improves the already observed drastic acceleration of the convergence of the hierarchy. Interestingly, the technique of proof involves classical results on Poisson's partial differential equation (PDE).

Keywords: convex optimization, real algebraic geometry, multivariate integration

1 Introduction

Consider the problem of computing the Lebesgue volume of a compact basic semi-algebraic set $\mathbf{K} \subset \mathbb{R}^n$. For simplicity of exposition we will restrict to the case where $\mathbf{K}$ is the smooth super-level set $\{ \mathbf{x} : g(\mathbf{x}) \geq 0 \} \subset \mathbb{R}^n$ of a single polynomial g.

If $\mathbf{K}$ is a *convex body* then several procedures are available; see e.g. exact deterministic methods for convex polytopes [1], or non deterministic Hit-and-Run methods [17, 22] and the more recent [2, 3]. Even *approximating* $\lambda(\mathbf{K})$ by deterministic methods is still a hard problem as explained in e.g. [3] and references therein. In full generality with no specific assumption on

*The work of J.B. Lasserre was partly funded by the AI Interdisciplinary Institute ANITI through the French "Investing for the Future PIA3" program under the Grant agreement n°ANR-19-PI3A-0004.

[1]LAAS-CNRS, Université de Toulouse, France.

[2]Réseau de Transport d'Electricité (RTE), France.

[3]Institut de Mathématiques de Toulouse, Université de Toulouse, France.

[4]Faculty of Electrical Engineering, Czech Technical University in Prague, Czechia.

$\mathbf{K}$ such as convexity, the only general method available is Monte-Carlo, that is, one samples N points according to Lebesgue measure λ normalized on a simple set $\mathbf{B}$ (e.g. a box or an ellipsoid) that contains $\mathbf{K}$. If ρ_N is the proportion of points that fall into $\mathbf{K}$ then the random variable $\rho_N\,\lambda(\mathbf{B})$ provides a good estimator of $\lambda(\mathbf{K})$ with convergence guarantees as N increases. However this estimator is non deterministic and neither provides a lower bound nor an upper bound on $\lambda(\mathbf{K})$.

When $\mathbf{K}$ is a compact basic semi-algebraic set, a deterministic numerical scheme described in [9] provides a sequence $(\tau_k)_{k\in\mathbb{N}} \subset \mathbb{R}$ of upper bounds that converges to $\lambda(\mathbf{K})$ as k increases. Briefly,

$$\lambda(\mathbf{K}) \;=\; \inf_{p\in\mathbb{R}[\mathbf{x}]} \left\{ \int p\,d\lambda \;:\; p \geq \mathbb{1}_{\mathbf{K}} \text{ on } \mathbf{B} \right\} \tag{1}$$

$$\tau_k \;=\; \inf_{p\in\mathbb{R}[\mathbf{x}]_k} \left\{ \int p\,d\lambda \;:\; p \geq \mathbb{1}_{\mathbf{K}} \text{ on } \mathbf{B} \right\}, \tag{2}$$

with $\mathbf{x} \mapsto \mathbb{1}_{\mathbf{K}}(\mathbf{x}) = 1$ if $\mathbf{x} \in \mathbf{K}$ and 0 otherwise. One can notice that minimizing sequences for (1) and (2) also minimize the $L^1(\mathbf{B},\lambda)$-norm $\|p - \mathbb{1}_{\mathbf{K}}\|_1$ (with convergence to 0 in the case (1)). As the upper bound $\tau_k > \lambda(\mathbf{K})$ is obtained by restricting the search in (2) to polynomials of degree at most k, the infimum is attained and an optimal solution can be obtained by solving a semidefinite program. Of course, the size of the resulting semidefinite program increases with the degree k; for more details the interested reader is referred to [9].

Then clearly, a Gibbs phenomenon[1] takes place as one tries to approximate on $\mathbf{B}$ and from above, the discontinuous function $\mathbb{1}_{\mathbf{K}}$ by a polynomial of degree at most k. This makes the convergence of the upper bounds τ_k very slow (even for modest dimension problems). A trick was used in [9] to accelerate this convergence but at the price of loosing monotonicity of the resulting sequence.

In fact (1) is a dual of the following infinite-dimensional Linear program (LP) on measures

$$\sup_{\mu} \{ \mu(\mathbf{K}) \;:\; \mu \leq \lambda;\, \mu \in \mathcal{M}(\mathbf{K})_+ \} \tag{3}$$

(where $\mathcal{M}(\mathbf{K})_+$ is the space of finite Borel measures on $\mathbf{K}$). Its optimal value is also $\lambda(\mathbf{K})$ and is attained at the unique optimal solution $\mu^\star := \lambda_{\mathbf{K}}$ (the restriction of λ to $\mathbf{K}$).

A simple but key observation. As one knows the unique optimal solution $\mu^\star = \lambda_{\mathbf{K}}$ of (3), *any* constraint satisfied by $\mu^\star$ (in particular, linear constraints) can be included as a constraint on μ in (3) without changing the optimal value and the optimal solution. While these constraints provide additional *restrictions* in (3), they translate into additional *degrees of freedom* in the dual (hence a *relaxed* version of (1)), and therefore better approximations when passing to the finite-dimensional relaxed version of (2). A first set of such linear constraints experimented in [13] and later in [14], resulted in drastic improvements but with no clear rationale behind such improvements.

Contribution. The main message and result of this paper is that there is an appropriate set of additional linear constraints on μ in (3) such that the resulting dual (a relaxed version of (1)) has an explicit *continuous* optimal solution with value $\lambda(\mathbf{K})$. These additional linear contraints (called Stokes constraints) come from an appropriate modelling of Stokes' theorem for integration over $\mathbf{K}$, a refined version of that in [13]. Therefore the optimal continuous solution can be approximated efficiently by polynomials with no Gibbs phenomenon, by the hierarchy of semidefinite relaxations defined in [9] (adapted to these new linear constraints). Interestingly,

[1] The Gibbs phenomenon appears at a jump discontinuity when one numerically approximates a piecewise C^1 function with a polynomial function, e.g. by its Fourier series; see e.g. [20, Chapter 9].

the technique of proof and the construction of the optimal solution invoke classical results from the field of elliptic partial differential equations (PDE), namely the Lax-Milgram and Poincaré-Wirtinger inequalities as well as regularity theorems for solutions to elliptic PDEs.

Outline. In Section 2 we recall the primal-dual linear formulation of the volume problem, and we explain why the dual value is not attained, which results in a Gibbs phenomenon. In Section 3 we revisit the acceleration strategy based on Stokes' theorem, with the aim of introducing in Section 4 a more general acceleration strategy and a new primal-dual linear formulation of the volume problem. Our main result, attainment of the dual value in this new formulation, is stated as Theorem 4 at the end of Section 4. The drastic improvement in the convergence to $\lambda(\mathbf{K})$ is illustrated on a simple example of the Euclidean ball.

2 Linear reformulation of the volume problem

Consider a compact basic semi-algebraic set

$$\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) \geq 0\}$$

with $g \in \mathbb{R}[\mathbf{x}]$. We suppose that $\mathbf{K} \subset \mathbf{B}$ where $\mathbf{B}$ is a compact basic semi-algebraic set for which we know the moments $\int_{\mathbf{B}} \mathbf{x}^{\mathbf{k}}\, \mathrm{d}\mathbf{x}$ of the Lebesgue measure $\lambda_{\mathbf{B}}$, where $\mathbf{x}^{\mathbf{k}} := x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$ denotes a multivariate monomial of degree $\mathbf{k} \in \mathbb{N}^n$. We assume that

$$\mathbf{\Omega} := \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) > 0\}$$

is a nonempty open set with closure

$$\overline{\mathbf{\Omega}} = \mathbf{K},$$

and that its boundary

$$\partial\mathbf{\Omega} = \partial\mathbf{K} = \mathbf{K} \setminus \mathbf{\Omega}$$

is C^1 in the sense that it is locally the graph of a continuously differentiable function. We want to compute the Lebesgue volume of $\mathbf{K}$, i.e., the mass of the Lebesgue measure $\lambda_{\mathbf{K}}$:

$$\lambda(\mathbf{K}) := \int_{\mathbf{K}} d\mathbf{x} = \int_{\mathbb{R}^n} d\lambda_{\mathbf{K}}(\mathbf{x}).$$

If $\mathbf{X} \subset \mathbb{R}^n$ is a compact set, denote by $\mathcal{M}(\mathbf{X})$ the space of signed Borel measures on $\mathbf{X}$, which identifies with the topological dual of $C^0(\mathbf{X})$, the space of continuous functions on $\mathbf{X}$. Denote by $\mathcal{M}(\mathbf{X})_+$ the convex cone of non-negative Borel measures on $\mathbf{X}$, and by $C^0(\mathbf{X})_+$ the convex cone of non-negative continuous functions on $\mathbf{X}$.

In [9] a sequence of upper bounds converging to $\lambda(\mathbf{K})$ is obtained by applying the Moment-SOS hierarchy [12] (a family of finite-dimensional convex relaxations) to approximate as closely as desired the (primal) infinite-dimensional LP on measures:

$$\begin{aligned} \max_{\mu}\ & \mu(\mathbf{K}) \\ \text{s.t. } & \mu \in \mathcal{M}(\mathbf{K})_+ \\ & \lambda_{\mathbf{B}} - \mu \in \mathcal{M}(\mathbf{B})_+ \end{aligned} \tag{4}$$

whose optimal value is $\lambda(\mathbf{K})$, attained for $\mu^\star := \lambda_{\mathbf{K}}$. The LP (4) has an infinite-dimensional LP dual on continuous functions which reads:

$$\begin{aligned} \inf_{w}\ & \int_{\mathbf{B}} w\, d\lambda \\ s.t.\ & w \in C^0(\mathbf{B})_+ \\ & w|_{\mathbf{K}} - 1 \in C^0(\mathbf{K})_+. \end{aligned} \tag{5}$$

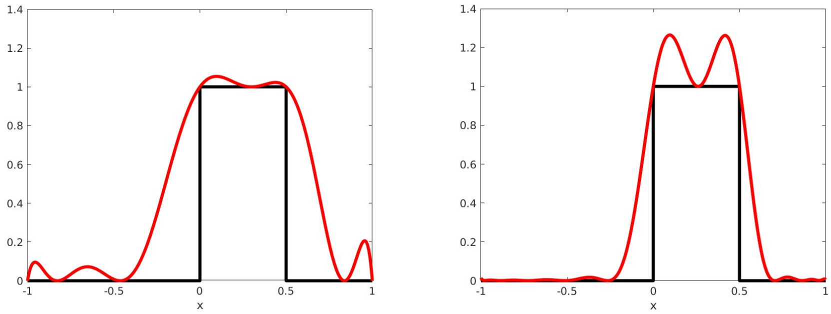

Figure 1: Gibbs effect occurring when approximating from above with a polynomial of degree 10 (left red curve) and 20 (right red curve) the indicator function of an interval (black curve).

Observe that (5) consists of approximating the discontinuous indicator function $\mathbb{1}_{\mathbf{K}}$ (equal to one on $\mathbf{K}$ and zero elsewhere) from above by continuous functions w, in minimizing the $L^1(\mathbf{B})$-norm $\|w - \mathbb{1}_{\mathbf{K}}\|_1$. Clearly the infimum $\lambda(\mathbf{K})$ is not attained.

Since $\mathbf{K}$ is generated by a polynomial g, and measures on compact sets are uniquely determined by their moments, one may apply the Moment-SOS hierarchy [12] for solving (4). The moment relaxation of (4) consists of replacing μ by finitely many of its moments $\mathbf{y}$, say up to degree $d \in \mathbb{N}$. Then the cone of moments is relaxed by a linear slice of the semidefinite cone constructed from so-called moment and localizing matrices indexed by d, as defined in e.g. [12], and which defines a semidefinite program. Therefore the dual of this semidefinite program (i.e., the dual SOS-hierarchy) is a *strenghtening* of (5) where

(i) continuous functions w are replaced with polynomials of increasing degree d, and

(ii) nonnegativity constraints are replaced with Putinar's SOS-based certificates of positivity [16] which translate to semidefinite constraints on the coefficients of polynomials; again the interest reader is referred to [9, 12] for more details.

For each fixed degree d, a valid upper bound on $\lambda(\mathbf{K})$ is computed by solving a primal-dual pair of convex semidefinite programming problems (not described here). As proved in [9] by combining Stone-Weierstrass' theorem and Putinar's Positivstellensatz [16],

(i) there is no duality gap between each primal semidefinite relaxation of the hierarchy and its dual, and

(ii) the resulting sequence of upper bounds converges to $\lambda(\mathbf{K})$ as d increases.

The main drawback of this numerical scheme is its typical slow convergence, observed already for very simple univariate examples, see e.g. [9, Figs. 4.1 and 4.5]. The best available theoretical convergence speed estimates are also pessimistic, with an asymptoptic rate of $\log \log d$ [11]. Slow convergence is mostly due to the so-called Gibbs phenomenon which is well-known in numerical analysis [20, Chapter 9]. Indeed, as already mentioned, solving (5) numerically amounts to approximating the discontinuous function $\mathbb{1}_{\mathbf{K}}$ from above with polynomials of increasing degree, which generates oscillations and overshoots and slows down the convergence, see e.g. [9, Figs. 4.2, 4.4, 4.6, 4.7, 4.10, 4.12].

Example 1. Let $\mathbf{K} := [0, 1/2] \subset \mathbf{B} := [-1, 1]$. In Figure 1 is displayed the degree-10 and degree-20 polynomials w obtained by solving the dual of SOS relaxations of problem (4). We can clearly see bumps, typical of a Gibbs phenomenon at points of discontinuity.

An idea to bypass this limitation consists of adding certain linear constraints to the finite-dimensional semidefinite relaxations, to make their optimal values larger and so closer to the optimal value $\lambda(\mathbf{K})$. Such linear constraints must be chosen appropriately:

(i) they must be *redundant* for the infinite-dimensional moment LP on measures (4), and

(ii) become *active* for its finite-dimensional relaxations.

This is the heuristic proposed in [13] to accelerate the Moment-SOS hierarchy for evaluating transcendental integrals on semi-algebraic sets. These additional linear constraints on the moments $\mathbf{y}$ of $\mu^\star$ are obtained from an application of Stokes' theorem for integration on $\mathbf{K}$, a classical result in differential geometry. It has been also observed experimentally that this heuristic accelerates significantly the convergence of the hierarchy in other applied contexts, e.g. in chance-constrained optimization problems [21].

3 Introducing Stokes constraints

In this section we explain the heuristic introduced in [13] to accelerate convergence of the Moment-SOS hierarchy by adding linear constraints on the moments of $\mu^\star$. These linear constraints are obtained from a certain application of Stokes' theorem for integration on $\mathbf{K}$.

3.1 Stokes' Theorem and its variants

Theorem 1 (Stokes' Theorem). *Let $\mathbf{\Omega} \subset \mathbb{R}^n$ be a piecewise C^1 open set. For any $(n-1)$-differential form ω on $\overline{\mathbf{\Omega}}$, it holds*

$$\int_{\partial\mathbf{\Omega}} \omega = \int_{\mathbf{\Omega}} d\omega.$$

.

Corollary 2. *In particular, for $\mathbf{u} \in C^1(\overline{\mathbf{\Omega}})^n$ and $\omega(\mathbf{x}) = \mathbf{u}(\mathbf{x}) \cdot \mathbf{n}_{\mathbf{\Omega}}(\mathbf{x})\, d\sigma(\mathbf{x})$, where the dot is the inner product, σ is the surface or Hausdorff measure on $\partial\mathbf{\Omega}$ and $\mathbf{n}_{\mathbf{\Omega}}$ is the outward pointing normal to $\partial\mathbf{\Omega}$, we obtain the Gauss formula*

$$\int_{\partial\mathbf{\Omega}} \mathbf{u}(\mathbf{x}) \cdot \mathbf{n}_{\mathbf{\Omega}}(\mathbf{x})\, d\sigma(\mathbf{x}) = \int_{\mathbf{\Omega}} \operatorname{div} \mathbf{u}(\mathbf{x})\, d\mathbf{x}. \tag{6}$$

With the choice $\mathbf{u}(\mathbf{x}) := u(\mathbf{x})\, \mathbf{e}_i$ where $u \in C^1(\overline{\mathbf{\Omega}})$ and $\mathbf{e}_i$ is the vector of $\mathbb{R}^n$ with one at entry i and zeros elsewhere, for $i = 1, \ldots, n$, we obtain the dual Gauss formula

$$\int_{\partial\mathbf{\Omega}} u(\mathbf{x})\, \mathbf{n}_{\mathbf{\Omega}}(\mathbf{x})\, d\sigma(\mathbf{x}) = \int_{\mathbf{\Omega}} \operatorname{grad} u(\mathbf{x})\, d\mathbf{x}. \tag{7}$$

Proof. These are all particular cases of [10, Theorem 6.10.2]. □

3.2 Original Stokes constraints

Associated to a sequence $\mathbf{y} = (y_{\mathbf{k}})_{\mathbf{k}\in\mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$, introduce the Riesz linear functional $\mathrm{L}_{\mathbf{y}} : \mathbb{R}[\mathbf{x}] \to \mathbb{R}$ which acts on a polynomial $p := \sum_{\mathbf{k}} p_{\mathbf{k}}\, \mathbf{x}^{\mathbf{k}} \in \mathbb{R}[\mathbf{x}]$ by $\mathrm{L}_{\mathbf{y}}(p) := \sum_{\mathbf{k}} p_{\mathbf{k}}\, y_{\mathbf{k}}$. Thus, if $\mathbf{y}$ is the sequence of moments of $\lambda_{\mathbf{K}}$, i.e. $y_{\mathbf{k}} := \int_{\mathbf{K}} \mathbf{x}^{\mathbf{k}} d\mathbf{x}$ for all $\mathbf{k} \in \mathbb{N}^n$, then $\mathrm{L}_{\mathbf{y}}(p) = \int_{\mathbf{K}} p(\mathbf{x}) d\mathbf{x}$ and by (7) with $u(\mathbf{x}) := \mathbf{x}^{\mathbf{k}} g(\mathbf{x})$:

$$\begin{aligned} \mathrm{L}_{\mathbf{y}}(\operatorname{grad}(\mathbf{x}^{\mathbf{k}} g)) &= \int_{\mathbf{K}} \operatorname{grad}(\mathbf{x}^{\mathbf{k}} g(\mathbf{x}))\, d\mathbf{x} \\ &= \int_{\partial\mathbf{K}} \mathbf{x}^{\mathbf{k}} g(\mathbf{x})\, \mathbf{n}_{\mathbf{K}}(\mathbf{x})\, d\sigma(\mathbf{x}) \;=\; 0, \end{aligned}$$

since by construction g vanishes on $\partial\mathbf{K}$. Thus while in the infinite-dimensional LP (4) one may add the linear constraints

$$\int_{\mathbf{K}} \mathrm{grad}(\mathbf{x}^{\mathbf{k}} g)\, d\mu \;=\; 0 \qquad \forall \mathbf{k} \in \mathbb{N}^n,$$

without changing its optimal value $\lambda(\mathbf{K})$, on the other hand inclusion of the linear moment constraints

$$\mathrm{L}_{\mathbf{y}}(\mathrm{grad}(\mathbf{x}^{\mathbf{k}} g)) = 0\,, \quad |\mathbf{k}| \leq 2d + 1 - \deg(g) \tag{8}$$

in the moment relaxation with pseudo-moments $\mathbf{y}$ of degree at most d, will decrease the optimal value of the initial relaxation.

In practice, it was observed that adding constraints (8) dramatically speeds up the convergence of the moment-SOS hierarchy, see e.g. [13, 21]. One main goal of this paper is to provide a qualitative mathematical rationale behind this phenomenon.

3.3 Infinite-dimensional Stokes constraints

In [18], Stokes constraints were formulated in the infinite-dimensional setting, and a dual formulation was obtained in the context of the volume problem. Using (6) with $\mathbf{u} = g\mathbf{v}$ (which vanishes on $\partial\mathbf{K}$) and $\mathbf{v} \in C^1(\mathbf{K})^n$ arbitrary, yields:

$$\int_{\mathbf{K}} (\mathrm{grad}\, g(\mathbf{x}) \cdot \mathbf{v}(\mathbf{x}) + g(\mathbf{x})\, \mathrm{div}\, \mathbf{v}(\mathbf{x}))\, \mathrm{d}\mathbf{x} = \int_{\partial\mathbf{K}} g\mathbf{v}\; \mathbf{n}_{\mathbf{K}}\; d\sigma = 0\,,$$

which can be written equivalently (in the sense of distributions) as

$$(\mathrm{grad}\, g)\lambda_{\mathbf{K}} - \mathrm{grad}(g\lambda_{\mathbf{K}}) = 0\,.$$

This allows to rewrite problem (4) as

$$\begin{aligned} &\max_{\mu}\; \mu(\mathbf{K}) \\ &s.t.\; \mu \in \mathcal{M}(\mathbf{K})_+ \\ &\quad\; \lambda_{\mathbf{B}} - \mu \in \mathcal{M}(\mathbf{B})_+ \\ &\quad\; (\mathrm{grad}\, g)\mu - \mathrm{grad}(g\mu) = 0 \end{aligned} \tag{9}$$

without changing its optimal value $\lambda(\mathbf{K})$ attained at $\mu^\star = \lambda_{\mathbf{K}}$.

Using infinite-dimensional convex duality as in e.g. the proof of Theorem 2 in [6], the dual of LP (9) reads

$$\begin{aligned} &\inf_{\mathbf{v},w}\; \int_{\mathbf{B}} w\, d\lambda \\ &s.t.\; \mathbf{v} \in C^1(\mathbf{K})^n \\ &\quad\; w \in C^0(\mathbf{B})_+ \\ &\quad\; w|_{\mathbf{K}} - \mathrm{div}(g\mathbf{v}) - 1 \in C^0(\mathbf{K})_+. \end{aligned} \tag{10}$$

Crucial observation. Notice that w in (9) is *not* required to approximate $\mathbb{1}_{\mathbf{K}}$ from above anymore. Instead, it should approximate $1 + \mathrm{div}(g\mathbf{v})$ on $\mathbf{K}$ and 0 outside $\mathbf{K}$. Hence, provided that $1+\mathrm{div}(g\mathbf{v}) = 0$ on $\partial\mathbf{K}$, w might be a continuous function for some well-chosen $\mathbf{v} \in C^1(\mathbf{K})^n$, and therefore an optimal solution of (10) (i.e., the infimum is a minimum). As a result, the Gibbs phenomenon would disappear and convergence would be faster.

The issue is then to determine whether the infimum in (10) is attained or not. And if not, are there other special features of problem (10) that can be exploited to yield more efficient semidefinite relaxations ?

4 New Stokes constraints and main result

In the previous section, the Stokes constraint

$$\int_{\mathbf{K}} (\mathbf{v}(\mathbf{x}) \cdot \operatorname{grad} g(\mathbf{x}) + g(x) \operatorname{div} \mathbf{v}(\mathbf{x}))\, d\mu(\mathbf{x}) = 0$$

or equivalently (in the sense of distributions)

$$(\operatorname{grad} g)\mu - \operatorname{grad}(g\mu) = 0 \tag{11}$$

(with $\mu \in \mathcal{M}(\mathbf{K})_+$ being the Lebesgue measure on $\mathbf{K}$) was obtained as a particular case of Stokes' theorem with $\mathbf{u} = g\mathbf{v}$ in (6). Instead, we can use a more general version with $\mathbf{u}$ not in factored form, and also use the fact that $\forall \mathbf{x} \in \partial\mathbf{K}$, $0 \neq \operatorname{grad} g(\mathbf{x}) = -|\operatorname{grad} g(\mathbf{x})|\ \mathbf{n}_{\mathbf{K}}(\mathbf{x})$ (here $|\mathbf{y}| := \sqrt{\mathbf{y} \cdot \mathbf{y}}$ is the n-dimensional Euclidean norm), to obtain

$$\int_{\mathbf{K}} \operatorname{div} \mathbf{u}(\mathbf{x})\ d\mu(\mathbf{x}) = -\int_{\partial\mathbf{K}} \mathbf{u}(\mathbf{x}) \cdot \operatorname{grad} g(\mathbf{x})\ d\nu(\mathbf{x}),$$

or equivalently (in the sense of distributions)

$$\operatorname{grad} \mu = (\operatorname{grad} g)\nu, \tag{12}$$

with $\mu \in \mathcal{M}(\mathbf{K})_+$ being the Lebesgue measure on $\mathbf{K}$ and $\nu \in \mathcal{M}(\partial\mathbf{K})_+$ being the measure having density $1/|\operatorname{grad} g(\mathbf{x})|$ with respect to the $(n-1)$-dimensional Haussdorff measure σ on $\partial\mathbf{K}$. The same linear equation was used in [14] to compute moments of the Hausdorff measure. In fact, equation (12) is a generalization of equation (11) in the following sense.

Lemma 3. *If $\nu \in \mathcal{M}(\partial\mathbf{K})_+$ is such that $\mu \in \mathcal{M}(\mathbf{K})_+$ satisfies* (12)*, then μ also satisfies* (11)*.*

Proof. Equation (12) means that $\int_{\mathbf{K}} \operatorname{div} \mathbf{u}(\mathbf{x}) d\mu(\mathbf{x}) + \int_{\partial\mathbf{K}} \mathbf{u}(\mathbf{x}) \operatorname{grad} g(\mathbf{x}) d\nu(\mathbf{x}) = 0$ for all $\mathbf{u} \in C^1(\mathbf{K})^n$. In particular if $\mathbf{u} = g\mathbf{v}$ for some $\mathbf{v} \in C^1(\mathbf{K})^n$ then (12) reads

$$\int_{\mathbf{K}} (\mathbf{v}(\mathbf{x}) \cdot \operatorname{grad} g(\mathbf{x}) + g(x) \operatorname{div} \mathbf{v}(\mathbf{x}))\, d\mu(\mathbf{x}) = 0,$$

which is precisely (11). □

Hence we can incorporate linear constraints (12) on μ and ν, to rewrite problem (4) as

$$\begin{aligned} \max_{\mu,\nu}\ & \mu(\mathbf{K}) \\ s.t.\ & \mu \in \mathcal{M}(\mathbf{K})_+ \\ & \nu \in \mathcal{M}(\partial\mathbf{K})_+ \\ & \lambda_{\mathbf{B}} - \mu \in \mathcal{M}(\mathbf{B})_+ \\ & (\operatorname{grad} g)\nu - \operatorname{grad} \mu = 0 \end{aligned} \tag{13}$$

without changing its optimal value $\lambda(\mathbf{K})$ attained at $\mu^\star = \lambda_{\mathbf{K}}$ and $\nu^\star = \sigma/|\operatorname{grad} g|$. Notice that LP (13) involves *two* measures μ and ν whereas LP (9) involves only one measure μ.

Next, by convex duality as in e.g. the proof of Theorem 2 in [6], the dual of (13) reads

$$\begin{aligned} \inf_{\mathbf{u},w}\ & \int_{\mathbf{B}} w\, d\lambda \\ s.t.\ & \mathbf{u} \in C^1(\mathbf{K})^n \\ & w \in C^0(\mathbf{B})_+ \\ & w|_{\mathbf{K}} - \operatorname{div} \mathbf{u} - 1 \in C^0(\mathbf{K})_+ \\ & -(\mathbf{u} \cdot \operatorname{grad} g)|_{\partial\mathbf{K}} \in C^0(\partial\mathbf{K})_+. \end{aligned} \tag{14}$$

Our main result states that the optimal value of the dual (14) is attained at some continuous function $(w, \mathbf{u}) \in C^0(\mathbf{B})_+ \times C^1(\mathbf{K})^n$. Therefore, in contrast with problem (5), there is no Gibbs phenomenon at an optimal solution of the (finite-dimensional) semidefinite strengthening associated with (14).

Let $\mathbf{\Omega}_i$, $i = 1, \ldots, N$ denote the connected components of $\mathbf{\Omega}$, and let

$$m_{\mathbf{\Omega}_i}(g) := \frac{1}{\lambda(\mathbf{\Omega}_i)} \int_{\mathbf{\Omega}_i} g \, d\lambda.$$

Theorem 4. *In dual LP* (14) *the infimum is a minimum, attained at*

$$w^\star(\mathbf{x}) := g(\mathbf{x}) \sum_{i=1}^{N} \frac{\mathbb{1}_{\mathbf{\Omega}_i}(\mathbf{x})}{m_{\mathbf{\Omega}_i}(g)}, \quad \mathbf{x} \in \mathbf{B},$$

and

$$\mathbf{u}^\star(\mathbf{x}) := \operatorname{grad} u(\mathbf{x}),$$

where u solves the Poisson PDE

$$\left\{ \begin{array}{rcll} -\Delta u(\mathbf{x}) & = & 1 - w^\star(\mathbf{x}), & \mathbf{x} \in \mathbf{\Omega} \\ \partial_{\mathbf{n}} u(\mathbf{x}) & = & 0, & \mathbf{x} \in \partial\mathbf{\Omega}. \end{array} \right.$$

Remark 1. The moment-SOS hierarchy associated to LPs (13) and (14) yields upper bounds for the volume. Theorem 4 is designed for these LPs but it has a straightforward counterpart for lower bound volume computation, obtained by replacing $\mathbf{K}$ with $\mathbf{B} \setminus \mathbf{\Omega}$ in the previous developments, i.e. computing upper bounds of $\lambda(\mathbf{B} \setminus \mathbf{\Omega})$. However, two additional technicalities should then be considered:

- This work only deals with semi-algebraic sets defined by a single polynomial; actually, it immediately generalizes to finite intersections of such semi-algebraic sets, as long as their boundaries do not intersect (i.e. here $\mathbf{K}$ should be included in the *interior* of $\mathbf{B}$): the constraints on boundaries should just be splitted between the boundaries of the intersected sets.
- This work heavily relies on the fact that the boundary of the considered set should be smooth; for this reason, computing lower bounds of the volume implies that one chooses a smooth bounding box $\mathbf{B}$ (typically a euclidean ball, ellipsoid or ℓ^p ball), which rules out simple sets like the hypercube $[-1, 1]^n$.

Upon taking into account these technicalities, Theorem 4 still holds, allowing to deterministically compute *upper and lower* bounds for the volume, with arbitrary precision. Of course in practice, one is limited by the performance of state-of-art SDP solvers.

5 Proof of main result

Theorem 4 is proved in several steps as follows:

- we show that the optimal dual solution satisfies a Poisson PDE;
- we study the Poisson PDE on a connected domain;
- we study the Poisson PDE on a union of connected domains;
- we construct an explicit optimum for problem (14).

5.1 Equivalence to a Poisson PDE

Lemma 5. *Problem* (14) *has an optimal solution iff there exist* $\mathbf{u} \in C^1(\overline{\mathbf{\Omega}})^n$ *and* $h \in C^0(\overline{\mathbf{\Omega}})_+$ *solving*

$$h = 0 \qquad \text{on } \partial\mathbf{\Omega}, \tag{15a}$$
$$-\operatorname{div}\mathbf{u} = 1 - h \qquad \text{in } \mathbf{\Omega}, \tag{15b}$$
$$\mathbf{u}\cdot\mathbf{n}_{\mathbf{\Omega}} = 0 \qquad \text{on } \partial\mathbf{\Omega}. \tag{15c}$$

Proof. Let $(\mathbf{u}, h)$ solve (15). Using (15a), one can define

$$w(\mathbf{x}) = \begin{cases} h(\mathbf{x}) & \text{if } \mathbf{x} \in \overline{\mathbf{\Omega}} \\ 0 & \text{if } \mathbf{x} \in \mathbf{B} \setminus \overline{\mathbf{\Omega}}. \end{cases}$$

Then $(\mathbf{u}, w)$ is feasible for (14) and one has

$$\begin{aligned} \int_{\mathbf{B}} w \, d\lambda &= \int_{\mathbf{\Omega}} h \, d\lambda \\ &\overset{(15b)}{=} \int_{\mathbf{\Omega}} (1 + \operatorname{div}\mathbf{u}) \, d\lambda \\ &\overset{(6)}{=} \lambda(\mathbf{\Omega}) + \int_{\partial\mathbf{\Omega}} \mathbf{u}\cdot\mathbf{n}_{\mathbf{\Omega}} \, d\sigma \\ &\overset{(15c)}{=} \lambda(\mathbf{\Omega}) \end{aligned}$$

so that $(\mathbf{u}, w)$ is optimal.

Conversely, let $(\mathbf{u}, w)$ be an optimal solution of problem (14). We know that $(\mu^\star, \nu^\star) = (\lambda_{\mathbf{\Omega}}, \sigma/|\operatorname{grad} g|)$ is optimal for problem (13). Then, duality theory ensures complementarity:

$$\int_{\mathbf{\Omega}} (w|_{\mathbf{\Omega}} - \operatorname{div}\mathbf{u} - 1) \, d\lambda = 0, \tag{16a}$$
$$\int_{\partial\mathbf{\Omega}} \mathbf{u}\cdot\frac{\operatorname{grad} g}{|\operatorname{grad} g|} \, d\sigma = 0. \tag{16b}$$

Since $w|_{\mathbf{\Omega}} - \operatorname{div}\mathbf{u} - 1$ is nonnegative, (16a) yields (15b) with $h := w|_{\mathbf{\Omega}}$. Likewise, since $-(\mathbf{u}\cdot\operatorname{grad} g)|_{\partial\mathbf{\Omega}}$ is nonnegative, (16b) yields (15c) and thus, using (6), it holds $\int_{\mathbf{\Omega}} \operatorname{div}\mathbf{u} \, d\lambda = 0$. Eventually, (16a) yields $\int_{\mathbf{\Omega}} w \, d\lambda = \lambda(\mathbf{\Omega}) = \int_{\mathbf{B}} w \, d\lambda$ by optimality of w, so that $\int_{\mathbf{B}\setminus\mathbf{\Omega}} w \, d\lambda = 0$ and, since w is nonnegative, $w|_{\mathbf{B}\setminus\mathbf{\Omega}} = 0$. Continuity of w finally allows to conclude that $w = 0$ on $\partial\mathbf{\Omega}$, which is exactly (15a). □

From Lemma 5, existence of an optimum for (14) is then equivalent to existence of a solution to (15), which we rephrase as follows, defining $f := 1 - h$ and $\mathbf{u} = \operatorname{grad} u$ with $u \in C^2(\overline{\mathbf{\Omega}})$, and where $\Delta u := \operatorname{div}\operatorname{grad} u$ is the Laplacian of u, and $\partial_{\mathbf{n}} u := \operatorname{grad} u \cdot \mathbf{n}_{\mathbf{\Omega}}$.

Lemma 6. *If there exist* $u \in C^2(\overline{\mathbf{\Omega}})^n$ *and* $f \in C^0(\overline{\mathbf{\Omega}})$ *solving*

$$-\Delta u = f \qquad \text{in } \mathbf{\Omega}, \tag{17a}$$
$$\partial_{\mathbf{n}} u = 0 \qquad \text{on } \partial\mathbf{\Omega}, \tag{17b}$$
$$f \leq 1 \qquad \text{in } \mathbf{\Omega}, \tag{17c}$$
$$f = 1 \qquad \text{on } \partial\mathbf{\Omega}, \tag{17d}$$

then problem (14) *has an optimal solution.*

This rephrazing is a Poisson PDE (17a) with Neumann boundary condition (17b), whose source term f is a parameter subject to constraints (17c) and (17d).

Remark 2 (Loss of generality). Looking for $\mathbf{u}$ under the form $\mathbf{u} = \operatorname{grad} u$ makes us loose the equivalence. Indeed, while (14) and (15) are equivalent, existence of a solution to (17) is only a sufficient condition for existence of an optimum for (14), since (15) might have only solutions $\mathbf{u}$ that are not gradients.

Remark 3 (Invariant set for gradient flow). From a dynamical systems point of view, the constraint in (14) which states that the inner product of $\mathbf{u} = \operatorname{grad} u$ with $\operatorname{grad} g$ is non-positive on $\partial\boldsymbol{\Omega}$, means that we are looking for a velocity field or control $\mathbf{u}$ in the form of the gradient of a potential u such that $\overline{\boldsymbol{\Omega}}$ is an invariant set for the solutions $t \in \mathbb{R} \mapsto \mathbf{x}(t) \in \mathbb{R}^n$ of the Cauchy problem

$$\dot{\mathbf{x}}(t) = -\operatorname{grad} u(\mathbf{x}(t)), \quad \mathbf{x}(0) \in \mathbf{B}$$

after what we just have to define $h := 1 + \Delta u$ on $\boldsymbol{\Omega}$.

5.2 Poisson PDE on a connected domain

It remains to prove existence of solutions to problem (17). First, notice that PDE (17a) together with its boundary condition (17b) enforces an important constraint on the source term f, namely its mean must vanish:

$$\int_{\boldsymbol{\Omega}} f \, d\lambda = 0. \tag{18}$$

Indeed, if (f, u) solves (17), then

$$\begin{aligned} \int_{\boldsymbol{\Omega}} f \, d\lambda &\overset{(17a)}{=} - \int_{\boldsymbol{\Omega}} \Delta u \, d\lambda \\ &\overset{(6)}{=} - \int_{\partial\boldsymbol{\Omega}} \operatorname{grad} u \cdot \mathbf{n}_{\boldsymbol{\Omega}} \, d\sigma \\ &= - \int_{\partial\boldsymbol{\Omega}} \partial_{\mathbf{n}} u \, d\sigma \overset{(17b)}{=} 0\,. \end{aligned}$$

Moreover, the following holds.

Lemma 7 (Existence on a connected domain)**.** *Suppose that $\boldsymbol{\Omega}$ is connected. Let the source term $f \in L^2(\boldsymbol{\Omega}) \cap C^\infty(\boldsymbol{\Omega})$ have zero mean on $\boldsymbol{\Omega}$. Then there exists $u \in C^\infty(\boldsymbol{\Omega})$ satisfying* (17a) *and* (17b).

Proof. First let us rephrase the Poisson PDE with Neumann boundary condition under a variational form. The problem reduces to finding $u \in H^1(\boldsymbol{\Omega})$ such that for any $v \in H^1(\boldsymbol{\Omega})$ one has

$$\int_{\boldsymbol{\Omega}} \operatorname{grad} u \cdot \operatorname{grad} v \, d\lambda = \int_{\boldsymbol{\Omega}} f \; v \, d\lambda. \tag{19}$$

Then, since $f \in L^2(\boldsymbol{\Omega})$, the interior H^2-regularity theorem (see [5, Theorem 1 in Section 6.3.1]) ensures that $u \in H^2_{loc}(\boldsymbol{\Omega})$, and Green's theorem writes, for all $v \in H^1(\boldsymbol{\Omega})$:

$$\int_{\partial\boldsymbol{\Omega}} v \; \partial_{\mathbf{n}} u \, d\sigma = \int_{\boldsymbol{\Omega}} v \Delta u \, d\lambda + \int_{\boldsymbol{\Omega}} \operatorname{grad} u \cdot \operatorname{grad} v \, d\lambda = \int_{\boldsymbol{\Omega}} (\Delta u + f) \; v \, d\lambda.$$

Especially, for $v \in C_c^\infty(\boldsymbol{\Omega}) \subset H^1(\boldsymbol{\Omega})$ the left hand side is zero, and by density of $C_c^\infty(\boldsymbol{\Omega})$ in $L^2(\boldsymbol{\Omega})$, we deduce that $-\Delta u = f$ for the $L^2(\boldsymbol{\Omega})$ Hilbert topology and then almost everywhere in $\boldsymbol{\Omega}$. The left hand side is then zero for any $v \in H^1(\mathbf{\Omega})$ and especially for any $v \in C^\infty(\overline{\mathbf{\Omega}}) \subset H^1(\mathbf{\Omega})$, so that, again by density of $C^\infty(\partial\mathbf{\Omega})$ in $L^2(\mathbf{\Omega})$, it holds $\partial_{\mathbf{n}} u = 0$ in $L^2(\partial\mathbf{\Omega})$ and then almost everywhere on $\partial\mathbf{\Omega}$.

Eventually, the interior C^∞-regularity theorem (see [5, Theorem 3 in section 6.3.1]) ensures that since $f \in C^\infty(\mathbf{\Omega})$, $u \in C^\infty(\mathbf{\Omega})$ and we obtain the announced result: u is a smooth strong solution of the Poisson PDE.

Next we invoke Lax-Milgram's theorem which provides existence and uniqueness of a solution to a given PDE (see e.g. [5, Section 6.2.1]). In our context the goal is to solve (19) for which it is clear that if u is a solution then any $\hat{u} := u + C$, $C \in \mathbb{R}$ is also solution, which makes it impossible to obtain uniqueness of the solution in $H^1(\mathbf{\Omega})$. We thus restrict ourselves to the hyperplane of zero-mean functions

$$\mathcal{H} := \left\{ u \in H^1(\mathbf{\Omega}) : \int_{\mathbf{\Omega}} u \, d\lambda = 0 \right\}$$

which is closed by continuity of the Lebesgue integral, so that $\mathcal{H}$ is a Hilbert space for the scalar product

$$\langle u|v\rangle_{\mathcal{H}} := \langle u|v\rangle_{H^1(\mathbf{\Omega})} = \int_{\mathbf{\Omega}} (uv + \operatorname{grad} u \cdot \operatorname{grad} v) \, d\lambda.$$

We then define the applications

$$\mathscr{B} : \begin{cases} \mathcal{H} \times \mathcal{H} & \longrightarrow \quad \mathbb{R} \\ (u, v) & \longmapsto \quad \displaystyle\int_{\mathbf{\Omega}} \operatorname{grad} u \cdot \operatorname{grad} v \, d\lambda \end{cases}$$

and

$$\mathscr{L} : \begin{cases} \mathcal{H} & \longrightarrow \quad \mathbb{R} \\ v & \longmapsto \quad \displaystyle\int_{\mathbf{\Omega}} f \; v \, d\lambda. \end{cases}$$

The Lax-Milgram theorem then states that if $\mathscr{L}$ and $\mathscr{B}$ are continuous and if $\mathscr{B}$ is moreover coercive, then there is a unique $u \in \mathcal{H}$ so that $\mathscr{L} = \mathscr{B}(u, \cdot)$, which is the announced equality. Let us show that these hypotheses are met.

- Continuity of $\mathscr{L}$. Since $\mathscr{L}$ is a linear operator, it is sufficient to show that it is bounded. Let $v \in \mathcal{H}$. Then, Hölder's inequality yields

$$\begin{aligned} |\mathscr{L}(v)| &= \left| \int_{\mathbf{\Omega}} f \; v \, d\lambda \right| \\ &\leq \|f\|_{L^2(\mathbf{\Omega})} \|v\|_{L^2(\mathbf{\Omega})} \\ &\leq \|f\|_{L^2(\mathbf{\Omega})} \|v\|_{\mathcal{H}} \end{aligned}$$

  because $\|v\|_{\mathcal{H}} = \sqrt{\|v\|^2_{L^2(\mathbf{\Omega})} + \|\operatorname{grad} v\|^2_{L^2(\mathbf{\Omega})}} \geq \|v\|_{L^2(\mathbf{\Omega})}$. Thus, $\mathscr{L}$ is a bounded operator and $|||\mathscr{L}||| = \|f\|_{L^2(\mathbf{\Omega})}$ (equality is obtained by taking $v = f \in \mathcal{H}$, made possible by (18)).

- Continuity of $\mathscr{B}$. Since $\mathscr{B}$ is a bilinear operator, it is sufficient to show that it is bounded. Let $u, v \in \mathcal{H}$. Again, Hölder's inequality yields

$$\begin{aligned} |\mathscr{B}(u, v)| &= \left| \int_{\mathbf{\Omega}} \operatorname{grad} u \cdot \operatorname{grad} v \, d\lambda \right| \\ &\leq \|\operatorname{grad} u\|_{L^2(\mathbf{\Omega})} \|\operatorname{grad} v\|_{L^2(\mathbf{\Omega})} \\ &\leq \|\operatorname{grad} u\|_{\mathcal{H}} \|\operatorname{grad} v\|_{\mathcal{H}} \end{aligned}$$

because $\|v\|_{\mathcal{H}} = \sqrt{\|v\|^2_{L^2(\mathbf{\Omega})} + \|\operatorname{grad} v\|^2_{L^2(\mathbf{\Omega})}} \geq \|\operatorname{grad} v\|_{L^2(\mathbf{\Omega})}$. Then, $\mathscr{B}$ is a bounded bilinear operator.

- Coercivity of $\mathscr{B}$. First, let us recall the following classical result, proved e.g. in [5, Theorem 1 in Section 5.8.1].

  Lemma 8 (Poincaré-Wirtinger inequality). *Let $\mathbf{\Omega} \subset \mathbb{R}^n$ be a bounded, connected, C^1 open set. There is a constant $C_{\mathbf{\Omega}} \geq 0$ such that for any $u \in H^1(\mathbf{\Omega})$:*

  $$\|u - m_{\mathbf{\Omega}}(u)\|_{L^2(\mathbf{\Omega})} \leq C_{\mathbf{\Omega}} \|\operatorname{grad} u\|_{L^2(\mathbf{\Omega})^n}$$

  where $m_{\mathbf{\Omega}}(u) := \frac{1}{\lambda(\mathbf{\Omega})} \int_{\mathbf{\Omega}} u \, d\lambda$.

  Now let us look for a constant $C \in \mathbb{R}$ such that $\|u\|_{\mathcal{H}} \leq C \;\; \mathscr{B}(u, u)$. Let $u \in \mathcal{H}$. Then, since $\partial\mathbf{\Omega}$ is locally Lipschitz, we can use Lemma 8:

  $$\begin{aligned}
  \mathscr{B}(u, u) &= \int_{\mathbf{\Omega}} |\operatorname{grad} u|_2^2 \, d\lambda \\
  &= \|\operatorname{grad} u\|^2_{L^2(\mathbf{\Omega})} \\
  &\geq \frac{\|\operatorname{grad} u\|^2_{L^2(\mathbf{\Omega})} + \|u - m_{\mathbf{\Omega}}(u)\|^2_{L^2(\mathbf{\Omega})}}{1 + C_{\mathbf{\Omega}}^2} \\
  &= \frac{\|u\|^2_{\mathcal{H}}}{1 + C_{\mathbf{\Omega}}^2}
  \end{aligned}$$

  since $u \in \mathcal{H}$ implies that $m_{\mathbf{\Omega}}(u) := \frac{1}{\lambda(\mathbf{\Omega})} \int_{\mathbf{\Omega}} u \, d\lambda = 0$.

Thus, the conditions of the Lax-Milgram theorem are satisfied, which gives us a unique $u \in \mathcal{H}$ such that for all $v \in \mathcal{H}$, equation (19) holds. To conclude, we still need to extend this property to functions v that have nonzero mean. Let $v \in H^1(\mathbf{\Omega})$, not necessary in $\mathcal{H}$. We define $\hat{v} := v - m_{\mathbf{\Omega}}(v)$, so that $\hat{v} \in \mathcal{H}$ and $\operatorname{grad} v = \operatorname{grad} \hat{v}$. Then,

$$\begin{aligned}
\int_{\mathbf{\Omega}} f \; v \, d\lambda &= \int_{\mathbf{\Omega}} f \; \hat{v} \, d\lambda + m_{\mathbf{K}}(v) \int_{\mathbf{\Omega}} f \, d\lambda \\
&\overset{(18)}{=} \int_{\mathbf{\Omega}} f \; \hat{v} \, d\lambda \\
&\overset{(19)}{=} \int_{\mathbf{\Omega}} \operatorname{grad} u \cdot \operatorname{grad} \hat{v} \, d\lambda \\
&= \int_{\mathbf{\Omega}} \operatorname{grad} u \cdot \operatorname{grad} v \, d\lambda,
\end{aligned}$$

which concludes the solution of the variational formulation and the proof of Lemma 7. □

5.3 Poisson PDE with boundary regularity on a union of connected domains

In Lemma 7, we assumed that $\mathbf{\Omega}$ is connected, so that we could apply the Poincaré-Wirtinger inequality to use the Lax-Milgram theorem, obtaining both existence and uniqueness of a solution in a well-chosen space. However, we are not interested in the uniqueness property, and we would like to tackle non-connected sets. Since $\mathbf{\Omega}$ is a semi-algebraic set, it has a finite number of connected components $\mathbf{\Omega}_1, \ldots, \mathbf{\Omega}_N$.

Lemma 9. *Let the source term $f \in L^2(\mathbf{\Omega}) \cap C^\infty(\mathbf{\Omega})$ have zero mean on each connected component of $\mathbf{\Omega}$. There there exists $u \in C^\infty(\mathbf{\Omega})$ solving* (17a) *and*(17b)*.*

Proof. Let $i = 1, \ldots, N$. Since $\int_{\mathbf{\Omega}_i} f \, d\lambda = 0$, we can apply the result of Theorem 7 replacing $\mathbf{\Omega}$ with $\mathbf{\Omega}_i$ to obtain $u_i \in C^\infty(\mathbf{\Omega}_i)$ such that $-\Delta u_i = f$ in $\mathbf{\Omega}_i$ and $\partial_{\mathbf{n}} u_i = 0$ on $\partial \mathbf{\Omega}_i$.

Then, we notice that since $\partial \mathbf{\Omega}$ is C^1, the $\mathbf{\Omega}_i$ cannot be mutually tangent, so that $\partial \mathbf{\Omega} = \bigsqcup_{i=1}^N \partial \mathbf{\Omega}_i$. Thus, for any $\mathbf{x} \in \mathbf{\Omega}$, the following sum has exactly one non-zero term:

$$u := \sum_{i=1}^N \mathbb{1}_{\overline{\mathbf{\Omega}}_i} u_i.$$

By definition of the $\mathbf{\Omega}_i$ as the connected components of $\mathbf{\Omega}$, $u \in C^\infty(\mathbf{\Omega})$.

Let $\mathbf{x} \in \mathbf{\Omega}$. There is an i such that $\mathbf{x} \in \mathbf{\Omega}_i$, so that $u = u_i$ on a neighbourhood of $\mathbf{x}$. Thus, $-\Delta u(\mathbf{x}) = -\Delta u_i(\mathbf{x}) = f(\mathbf{x})$.

Let $\mathbf{x} \in \partial \mathbf{\Omega}$. There is an i such that $\mathbf{x} \in \partial \mathbf{\Omega}_i$, so that $u = u_i$ on a neighbourhood of $\mathbf{x}$ in $\mathbf{\Omega}$. Thus, $\partial_{\mathbf{n}} u(\mathbf{x}) = \partial_{\mathbf{n}} u_i(\mathbf{x}) = 0$. □

In Section 5.2 we have proved that under suitable conditions on the source term f, equations (17a) and (17b) have a solution $u \in C^\infty(\mathbf{\Omega})$. However, the existence of an optimum for problem (14) requires u to be in $C^1(\mathbf{K})$: we need to establish regularity at the boundary. For this, an additional assumption on $\mathbf{\Omega}$ is needed to state the following corollary to Lemma 7.

Lemma 10. *Let the source term $f \in C^\infty(\overline{\mathbf{\Omega}})$ have zero mean on each connected component of $\mathbf{\Omega}$. Suppose that $\partial \mathbf{\Omega}$ is C^∞. Then, there exists $u \in C^\infty(\overline{\mathbf{\Omega}})$ solving* (17a) *and*(17b)*.*

Proof. First, since $L^2(\mathbf{\Omega}) \cap C^\infty(\mathbf{\Omega}) \subset C^\infty(\overline{\mathbf{\Omega}})$, we can use Lemma 9 to get a suitable $u \in C^\infty(\mathbf{\Omega})$. The only thing that remains to be proved is the regularity of u on $\partial \mathbf{\Omega}$. For this, we use the boundary C^∞-regularity theorem [5, Theorem 6 in Section 6.3.2]: since $f \in C^\infty(\overline{\mathbf{\Omega}})$ and $\partial \mathbf{\Omega}$ is C^∞, we conclude that $u \in C^\infty(\overline{\mathbf{\Omega}})$. □

Remark 4. Assuming that $\partial \mathbf{\Omega}$ is C^∞ instead of C^1 is actually without loss of generality since $\mathbf{\Omega}$ is a semi-algebraic set: as soon as $\partial \mathbf{\Omega}$ is locally the graph of a C^1 function, it is smooth.

Remark 5. Lemma 10 automatically enforces $-\Delta u = 1$ on $\partial \mathbf{\Omega}$, which is crucial for the continuity of the optimization variable w.

5.4 Explicit optimum for volume computation with Stokes constraints

Our optimization problem does not feature only the Poisson PDE with Neumann condition: it also includes constraints (17c) and (17d) on the source term. Consequently, a function $f \in C^\infty(\overline{\mathbf{\Omega}})$ with zero integral over any connected component of $\mathbf{\Omega}$ and satisfying (17c) and (17d) remains to be constructed. We keep the notations of Lemma 9 and suggest as candidate

$$\mathbf{x} \mapsto f(\mathbf{x}) := 1 - g(\mathbf{x}) \sum_{i=1}^N \frac{\mathbb{1}_{\mathbf{\Omega}_i}(\mathbf{x})}{m_{\mathbf{\Omega}_i}(g)}. \tag{20}$$

By definition, $g = 0$ on $\partial \mathbf{\Omega}$, so that (17d) automatically holds. Moreover, both g and $\mathbb{1}_{\mathbf{\Omega}_i}$ are nonnegative on $\mathbf{K}$, so that (17c) also holds.

In terms of regularity, f is polynomial on each connected component of $\mathbf{\Omega}$ and since g smoothly vanishes on $\partial \mathbf{\Omega}$, $f \in C^\infty(\mathbf{K})$.

Eventually, let $i \in 1, \ldots, N$ so that $\mathbf{\Omega}_i$ is a connected component of $\mathbf{\Omega}$. Then, by definition, $\partial\mathbf{\Omega}_i \subset \partial\mathbf{\Omega}$, and one has

$$\begin{aligned}\int_{\mathbf{\Omega}_i} f \, d\lambda \; &= \int_{\mathbf{\Omega}_i} \left(1 - g(\mathbf{x}) \sum_{i=1}^{N} \frac{\mathbb{1}_{\mathbf{\Omega}_i}(\mathbf{x})}{m_{\mathbf{\Omega}_i}(g)}\right) \mathrm{d}\mathbf{x} \\ &= \lambda(\mathbf{\Omega}_i) - \frac{1}{m_{\mathbf{\Omega}_i}(g)} \int_{\mathbf{\Omega}_i} g(\mathbf{x}) \, \mathrm{d}\mathbf{x} \; = \; 0,\end{aligned}$$

since by definition $m_{\mathbf{\Omega}_i}(g) = \frac{1}{\lambda(\mathbf{\Omega}_i)} \int_{\mathbf{\Omega}_i} g(\mathbf{x}) \, \mathrm{d}\mathbf{x}$.

We finally obtain our couple (u, f) solution to problem (17) with f defined in (20) and u given by Lemma 10. Then we retrieve the couple $(\mathbf{u}, h)$ solution to problem (15) by defining $\mathbf{u} := \operatorname{grad} u$ and for all $\mathbf{x} \in \mathbf{\Omega}$:

$$h(\mathbf{x}) := 1 - f(\mathbf{x}) = g(\mathbf{x}) \sum_{i=1}^{N} \frac{\mathbb{1}_{\mathbf{\Omega}_i}(\mathbf{x})}{m_{\mathbf{\Omega}_i}(g)}.$$

Eventually, the optimization problem (14) has a (global) minimizer $(\mathbf{u}, w)$ with, for all $\mathbf{x} \in \mathbf{B}$,

$$w(\mathbf{x}) = g(\mathbf{x}) \sum_{i=1}^{N} \frac{\mathbb{1}_{\mathbf{\Omega}_i}(\mathbf{x})}{m_{\mathbf{\Omega}_i}(g)}.$$

Indeed, one can check that

$$\begin{aligned}\int_{\mathbf{B}} w \, d\lambda \; &= \sum_{i=1}^{N} \frac{1}{m_{\mathbf{\Omega}_i}(g)} \int_{\mathbf{\Omega}_i} g \, d\lambda \\ &= \sum_{i=1}^{N} \lambda(\mathbf{\Omega}_i) \; = \; \lambda(\mathbf{\Omega}) \; = \; \lambda(\mathbf{K}),\end{aligned}$$

which concludes the proof of Theorem 4.

6 Examples

To illustrate how efficient can be the introduction of Stokes constraints for volume computation, we consider the simple setting where $\mathbf{K}$ is a Euclidean ball included in $\mathbf{B}$ the unit Euclidean ball. Indeed drastic improvements on the convergence are observed. All numerical examples were processed on a standard laptop computer under the Matlab environment with the SOS parser of YALMIP [15], the moment parser GloptiPoly [8] and the semidefinite programming solver of MOSEK [4].

6.1 Practical implementation

Following the Moment-SOS hierarchy methodology for volume computation as described in [9], in the (finite-dimensional) degree d semidefinite strengthening of dual problem (14):

- $w \in \mathbb{R}[\mathbf{x}]_d$ and $\mathbf{u} \in \mathbb{R}[\mathbf{x}]_d^n$ are polynomials of degree at most d;
- the positivity constraint $w \in C^0(\mathbf{B})_+$ is replaced with a Putinar certificate of positivity on $\mathbf{B}$, that is:

  $$w(\mathbf{x}) \; = \; \sigma_0(\mathbf{x}) + \sigma_1(\mathbf{x})(1 - |x|^2) \,, \quad \forall \mathbf{x} \in \mathbb{R}^n \,,$$

  where σ_0 (resp. σ_1) is an SOS polynomial of degree at most $2d$ (resp. $2d - 2$);

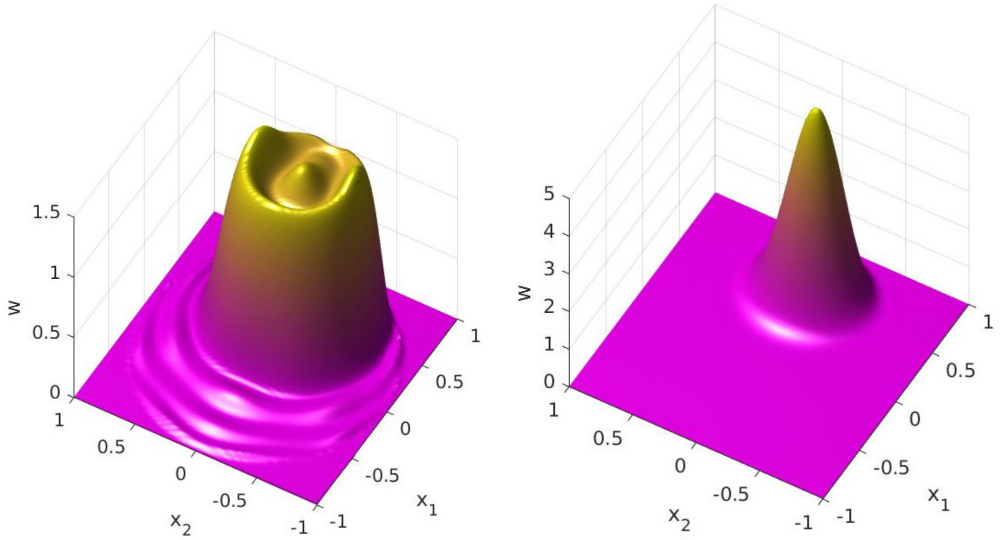

Figure 2: Degree 16 polynomial approximations obtained without Stokes constraints (left) and with Stokes constraints (right).

- the positivity constraint $w|_{\mathbf{K}} - \operatorname{div} \mathbf{u} - 1 \in C^0(\mathbf{K})_+$ is replaced with a Putinar certificate of positivity on $\mathbf{K}$, that is:

$$w(\mathbf{x}) - \operatorname{div} \mathbf{u}(\mathbf{x}) - 1 \;=\; \psi_0(\mathbf{x}) + \psi_1(\mathbf{x})\, g(\mathbf{x})\,, \quad \forall \mathbf{x} \in \mathbb{R}^n\,,$$

  where ψ_0 (resp. ψ_1) is an SOS polynomial of degree at most $2d$ (resp. $2d - \deg(g)$);

- the positivity constraint $(\mathbf{u} \cdot \operatorname{grad} g)|_{\partial\mathbf{K}} \in C^0(\partial\mathbf{K})_+$ is replaced with a Putinar certificate of positivity on $\partial\mathbf{K}$, that is:

$$-\mathbf{u}(\mathbf{x}) \cdot \operatorname{grad} g(\mathbf{x}) \;=\; \eta_0(\mathbf{x}) + \eta_1(\mathbf{x})\, g(\mathbf{x})\,, \quad \forall \mathbf{x} \in \mathbb{R}^n\,,$$

  where η_0 is an SOS polynomial of degree at most $2d$ and η_1 is a polynomial of degree at most $2d - \deg(g)$;

- the linear criterion $\int_{\mathbf{B}} w d\lambda$ translates into linear criterion on the vector of coefficients of w, as $\int_{\mathbf{B}} \mathbf{x}^\alpha \, d\lambda$ is available in closed-form.

The above identities define linear constraints on the coefficients of all the uknown polynomials. Next, stating that some of these polynomials must be SOS translate into semidefinite constraints on their respective unknown Gram matrices. The resulting optimization problem is a semidefinite program; for more details the interested reader is referred to e.g. [9].

6.2 Bivariate disk

Let us first illustrate Theorem 4 for computing the area of the disk $\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^2 : g(\mathbf{x}) = 1/4 - (x_1 - 1/2)^2 - x_2^2 \geq 0\}$ included in the unit disk $\mathbf{B} := \{\mathbf{x} \in \mathbb{R}^2 : 1 - x_1^2 - x_2^2 \geq 0\}$.

The degree $d = 16$ polynomial approximation w obtained by solving the SOS relaxation of linear problem (5) is represented at the left of Figure 2. We can see bumps and ripples typical

n	d	without Stokes	with Stokes
3	4	88% (0.03s)	18% (0.04s)
3	8	57% (0.16s)	1.0% (0.44s)
3	12	47% (1.97s)	0.0% (4.63s)
3	16	43% (23.9s)	0.0% (30.1s)
3	20	41% (142s)	0.0% (206s)

Table 1: Relative errors (%) and computational times (in brackets in seconds) for solving moment relaxations of increasing degrees d approximating the volume of ball of dimension $n = 3$.

n	d	without Stokes	with Stokes
1	10	17% (0.05s)	0.0% (0.03s)
2	10	35% (0.09s)	0.2% (0.25s)
3	10	56% (0.52s)	0.3% (1.19s)
4	10	72% (9.74s)	0.4% (22.8s)
5	10	79% (150s)	0.6% (669s)

n	d	without Stokes	with Stokes
6	4	190% (0.25s)	45.1% (1.03s)
7	4	203% (0.32s)	60.0% (4.88s)
8	4	221% (0.42s)	78.6% (8.45s)
9	4	245% (1.15s)	102% (45.1s)
10	4	278% (3.10s)	131% (176s)

Table 2: Relative errors (%) and computational times (in brackets in seconds) for solving the degree $d = 10$ (left) and $d = 4$ (right) moment relaxation approximating the volume of a ball of increasing dimensions n.

of a Gibbs phenomenon, since the polynomial should approximate from above the discontinuous indicator function $\mathbb{1}_{\mathbf{K}}$ as closely as possible. A rather loose upper bound of 1.1626 is obtained on the volume $\lambda(\mathbf{K}) = \frac{\pi}{4} \approx 0.7854$.

In comparison, the degree $d = 16$ polynomial approximation w obtained by solving the SOS relaxation of linear problem (14) is represented at the right of Figure 2. As expected from the proof of Theorem 4, the poynomial should approximate from above the continuous function $g\mathbb{1}_{\mathbf{K}}\, \lambda(\mathbf{K})/(\int g\lambda_{\mathbf{K}})$. The resulting polynomial approximation is smoother and yields a much improved upper bound of 0.7870.

6.3 Higher dimensions

In Table 1 we report on the dramatic acceleration brought by Stokes constraints in the case of the Euclidean ball $\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^3 : g(\mathbf{x}) = (3/4)^2 - |\mathbf{x}|^2 \geq 0\}$ of dimension $n = 3$ included in the unit ball $\mathbf{B}$. We specify the relative errors on the bounds obtained by solving moment relaxations with and without Stokes constraints, together with the computational times (in seconds), for a relaxation degree d ranging from 4 to 20. We observe that tight bounds are obtained already at low degrees with Stokes constraints, sharply contrasting with the loose bounds obtained without Stokes constraints. However, we see also that the inclusion of Stokes constraints has a computational price.

In Table 2 we report the relative errors on the bounds obtained with and without Stokes constraints, together with the computational times (in seconds), for a relaxation degree equal to $d = 10$ (left) resp. $d = 4$ (right) and for dimension n ranging from 1 to 5 (left) resp. from 6 to 10 (right). When $d = 10$ and $n = 5$ the semidefinite relaxation features 6006 pseudo-moments without Stokes constraints, and 12194 pseudo-moments with Stokes constraints. We see that introducing Stokes constraints incurs a computational cost, to be compromised with the expected quality of the bounds.

Higher dimensional problems can be addressed only if the problem description has some sparsity structure, as explained in [18]. Also, depending on the geometry of the problem, and for larger values of the relaxation degree, alternative polynomial bases may be preferable numerically than the monomial basis which is used by default in Moment and SOS parsers.

7 Conclusion

In this paper we proposed a new primal-dual infinite-dimensional linear formulation of the problem of computing the volume of a smooth semi-algebraic set generated by a single polynomial, generalizing the approach of [9] while still allowing the application of the moment-SOS hierarchy. The new dual formulation contains redundant linear constraints arising from Stokes's Theorem, generalizing the heuristic of [13]. A striking property of this new formulation is that the dual value is attained, contrary to the original formulation. As a consequence, the corresponding dual SOS hierarchy does not suffer from the Gibbs phenomenon, thereby accelerating the convergence.

Numerical experiments (not reported here) reveal that the values obtained with the new Stokes constraints (with a general vector field) are closely matching the values obtained with the original Stokes constraints of [13] (with the generating polynomial factoring the vector field). It may be then expected that the original and new Stokes constraints are equivalent. However at this stage we have not been able to prove equivalence.

The proof of dual attainment builds upon classical tools from linear PDE analysis, thereby building up a new bridge between infinite-dimensional convex duality and PDE theory, in the context of the moment-SOS hierarchy. We expect that these ideas can be exploited to prove regularity properties of linear reformulations of other problems in data science, beyond volume approximation. For example, it would be desirable to design Stokes constraints tailored to the infinite-dimensional linear reformulation of the region of attraction problem [6] or its sparse version [19].

Acknowledgment

This work benefited from discussions with Swann Marx.

References

[1] B. Büeler, A. Enge, K. Fukuda. Exact Volume Computation for Polytopes: A Practical Study. In Kalai G. Ziegler G.M. (Eds) Polytopes - Combinatorics and Computation, DMV Seminar, Birhäuser, 29:131-154, 2000.

[2] B. Cousins, S. Vempala. A practical volume algorithm. Math. Program. Comput. 8:133-160, 2016.

[3] D. Dadush, S. Vempala. Near-optimal deterministic algorithms for volume computation via M-ellipsoids, PNAS 110(48):19237-19245, 2013.

[4] J. Dahl. Semidefinite optimization using MOSEK. ISMP, Berlin, 2012.

[5] L. C. Evans. Partial Differential Equations. 2nd edition, AMS, 2010.

[6] D. Henrion, M. Korda. Convex computation of the region of attraction of polynomial control systems. IEEE Trans. Automatic Control, 59(2):297-312, 2014.

[7] D. Henrion, M. Korda, J. B. Lasserre. The Moment-SOS Hierarchy - Lectures in Probability, Statistics, Computational Geometry, Control and Nonlinear PDEs. Series on Optimization and Its Applications, World Scientific, 2020.

[8] D. Henrion, J. B. Lasserre, J. Löfberg. GloptiPoly 3: moments, optimization and semidefinite programming Optimization Methods and Software 24(4-5):761-779, 2009.

[9] D. Henrion, J. B. Lasserre, C. Savorgnan. Approximate volume and integration for basic semialgebraic sets. SIAM Review 51(4):722–743, 2009.

[10] J. H. Hubbard, B. Burke Hubbard. Vector Calculus, Linear Algebra and Differential Forms (A Unified Approach), 3rd edition, Pearson, 2007.

[11] M. Korda, D. Henrion. Convergence rates of moment-sum-of-squares hierarchies for volume approximation of semialgebraic sets. Optimization Letters, 12(3):453-442, 2018.

[12] J. B. Lasserre. Moments, Positive Polynomials and Their Applications. Imperial College Press, 2010.

[13] J. B. Lasserre. Computing Gaussian and exponential measures of semi-algebraic sets. Adv. Appl. Math., 91:137-163, 2017.

[14] J. B. Lasserre, V. Magron. Computing the Hausdorff boundary measure of semi-algebraic sets. `arXiv:2001.07598`, to appear in SIAM J. Appl. Alg. Geometry, 2020.

[15] J. Löfberg. YALMIP : A Toolbox for Modeling and Optimization in MATLAB. IEEE CACSD, Taiwan, 2004.

[16] M. Putinar. Positive polynomials on compact semi-algebraic sets. Indiana Univ. Math. J. 42, 1993.

[17] R. Smith. The hit-and-run sampler: A globally reaching Markov chain sampler for generating arbitrary multivariate distribution. In J. M. Charnes, D. J. Morrice, D. T. Brunner, J. J Swain (Eds). Proceedings of the 1996 Winter Simulation Conference, 260-264, 1996.

[18] M. Tacchi, T. Weisser, J. B. Lasserre, D. Henrion. Exploiting sparsity in semi-algebraic set volume computation. `arXiv:1902.02976`, submitted for publication, 2019.

[19] M. Tacchi, C. Cardozo, D. Henrion, J. B. Lasserre. Approximating regions of attraction of a sparse polynomial differential system. IFAC World Congress, Berlin, 2020.

[20] L. N. Trefethen. Approximation Theory and Approximation Practice. SIAM, 2013.

[21] T. Weisser. Computing Approximations and Generalized Solutions using Moments and Positive Polynomials. PhD thesis, University of Toulouse, 2018.

[22] Z. B. Zabinsky, R.L. Smith. Hit-and-Run methods. In S. Gass, M. C. Fu (Eds.). Encyclopedia of Operations Research and Management Science, 2013.