RECOVERING LOWER-DIMENSIONAL SEMIALGEBRAIC SUPPORT OF A MEASURE FROM ITS MOMENTS

Anonymous authors
Paper under double-blind review

ABSTRACT

Recovering probability measures from their moments has numerous applications, esp. in connection with the method of moments in statistics and optimization. In the setting where measure need not be finitely atomic, but its support is known to be compact and semialgebraic with codimension at least one, the problem is still open. We combine moment-matrix kernel information with the Christoffel–Darboux kernel to provide a discrete approximation of the support. To validate the proposed approach, we test our algorithm on analytically computed moments and pseudo-moments arising from polynomial optimization problems without unique global minimizers. This complements well-known recent work on recovery of measures with algebraic support, where the kernel of a moment matrix can reveal polynomials vanishing on the support, and on recovery of sufficiently regular full-dimensional supports, where estimators constructed by thresholding the Christoffel–Darboux kernel are known to converge asymptotically to the support.

1 INTRODUCTION

The recovery of a measure from finitely many of its moments is a classical problem (Akhiezer, 2020; Schmüdgen, 2017). It underlies the method of moments in statistics (Lindsay, 1995) and the extraction of solutions in polynomial optimization (Lasserre, 2001; Parrilo, 2003). In the second of these settings one does not, strictly speaking, observe moments at all. The moment relaxation of degree d of a polynomial optimization problem (POP) returns a *pseudo-moment* sequence, that is, the values of a linear functional which behaves on polynomials of low degree like integration against a probability measure, but which need not be the moment sequence of any measure (Lasserre et al., 2013; Laurent, 2009). The optimal values of the relaxations converge to that of the POP as the degree increases (Lasserre, 2001), but the convergence is asymptotic, and in practice only small degrees are affordable. One is therefore left with the problem of recovering the set of minimizers from a pseudo-moment sequence which is both short and inexact.

Two families of methods are available, and they rest on hypotheses of different kinds. The first is algebraic and uses the kernel of the moment matrix. Every polynomial in $\ker M_d(y)$ vanishes on $\operatorname{supp}\mu$, and so the kernel determines the Zariski closure of the support (Laurent, 2009; Wageringel, 2022; Lasserre et al., 2013). If μ is finitely atomic and the moment matrix admits a flat extension, the atoms can be computed exactly by solving an eigenvalue problem (Curto & Fialkow, 1998; Henrion & Lasserre, 2005; López Quijorna, 2021). The kernel carries no information about inequalities, however: the semicircle $\{x_1^2+x_2^2=1,\ x_2\geq 0\}$ and the full circle have the same vanishing ideal, and hence the same moment-matrix kernel in every degree. The second family is analytic and thresholds the Christoffel–Darboux (CD) kernel. The Christoffel function is relatively large on the support of the measure and decreases rapidly away from it, and the sublevel sets of the CD kernel converge in the Hausdorff metric to the support, provided that the measure is absolutely continuous and supported on a *full-dimensional* set of sufficient regularity (Lasserre & Pauwels, 2019; Pauwels et al., 2021; Vu et al., 2022; Lasserre et al., 2022). Inequalities cause no difficulty here, but the convergence results do not apply once the support has positive codimension, and this is the usual situation in polynomial optimization, where sets of minimizers generally have empty interior.

This paper is concerned with the intermediate case: supports which are compact and semialgebraic, have codimension at least one, and are not necessarily finite. We are given the moments of μ up to

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

degree $2d$, but no samples, and we wish to compute a finite set of points at small Hausdorff distance from $\operatorname{supp}\mu$. We are not aware of any existing method which does this from moments alone.

We propose a procedure, Algorithm 1, which uses the moment matrix twice, first algebraically and then analytically, and we study its behaviour numerically. Our contributions are as follows.

1. A method which combines the two techniques (Section 4). The algebraic carrier $\mathcal{Z} = V_{\mathbb{R}}(\ker M_d(y))$ is recovered first; $\mathcal{Z}$ is then sampled uniformly with respect to its intrinsic volume, by the random-slicing method of Breiding & Marigliano (2020); and only then is the *relative* CD kernel thresholded *along* $\mathcal{Z}$. The equations are treated using the kernel of the moment matrix and the inequalities using the CD kernel, each instrument being used where it is appropriate. Since κ_d is a polynomial and $\mathcal{Z}$ is algebraic, the sublevel set $\{x \in \mathcal{Z} : \kappa_d(x) \le \tau_d\}$ is already a semialgebraic description of the estimated support, of which the returned cloud of points is a sample.
2. Experiments upon exact moments and upon pseudo-moments alike (Sections 5.2 and 5.3). We work with moments computed in closed form for measures satisfying our hypotheses exactly, and with pseudo-moments produced by TSSOS (Magron & Wang, 2021a;b) for POPs whose minimizer sets are a segment and a semicircle, where the flat extension property fails and the GNS construction is of no help.
3. The ratio of the sampling errors on the ambient grid and the recovered carrier (with equal budget) grows exponentially with n, whatever function is filtered on the grid. This is shown by Theorem 1 and corroborated by experiments.

2 Related work

Vanishing ideals from moment matrices. That $\ker M_d(y)$ consists of polynomials vanishing on $\operatorname{supp}\mu$ is classical (Laurent, 2009). To pass from a numerical basis of the kernel to a set of *generators*, one must separate the new relations in each degree from polynomial multiples of relations found in lower degrees; border bases, Gröbner-type reductions and real radical computations may all be used for this purpose (Lasserre et al., 2013; Laurent & Mourrain, 2009). On truncated moment problems on varieties and semialgebraic sets see Wageringel (2022); Nie (2014).

Finitely atomic measures. When μ is a finite sum of Dirac measures, the flat extension theorem (Curto & Fialkow, 1998) gives exact recovery, as implemented in GloptiPoly (Henrion & Lasserre, 2005) and in the truncated GNS construction (López Quijorna, 2021); the extraction is stable under perturbation (Klep et al., 2018) and rests on Stickelberger's theorem (Cox, 2020), and the same ideas appear in super-resolution (Vetterli et al., 2002; de Castro & Gamboa, 2012; Duval & Peyré, 2015). None of these methods applies when μ has a non-atomic part.

Christoffel–Darboux kernels and the estimation of supports. The empirical Christoffel function of Lasserre & Pauwels (2019) estimates the support of a measure from samples; its sublevel sets converge in the Hausdorff metric under assumptions of regularity and density, and Vu et al. (2022) give rates for finite samples. Pauwels et al. (2021) develop the corresponding theory for moments, including the quotient construction by which the kernel is defined on a variety, and Lasserre et al. (2022) give a book-length account; see also Totik (2010) on asymptotics, Marx et al. (2021) on semialgebraic approximation, and Devonport et al. (2021) for an application. All the convergence results known to us assume that the support is regular closed in the ambient space, $S = \overline{\operatorname{int} S}$, and in our setting it is not. The mollified CD kernels of Bentancur et al. (2026) are designed for the lower-dimensional case; we discuss them in Section 4.3. For sampling on varieties we use Breiding & Marigliano (2020) (Section 4.2); on sampling density and topology see Niyogi et al. (2008).

3 Problem setting

3.1 Notation

We work throughout in $\mathbb{R}^n$. For $\alpha \in \mathbb{N}^n$ we write $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ and $|\alpha| = \sum_i \alpha_i$. We denote by $\mathbb{R}[x]_d$ the space of polynomials of degree at most d, by $v_d(x)$ the vector of all monomials of degree at most d, and by $\sigma(d) = \binom{n+d}{d} = \dim \mathbb{R}[x]_d$ the number of such monomials. We identify

$p \in \mathbb{R}[x]_d$ with the vector $\vec{p} \in \mathbb{R}^{\sigma(d)}$ of its coefficients in the monomial basis. Given a truncated sequence $y = (y_\alpha)_{|\alpha| \le 2d}$, the *moment matrix* $M_d(y) \in \mathbb{R}^{\sigma(d) \times \sigma(d)}$ is defined by

$$[M_d(y)]_{\alpha,\beta} = y_{\alpha+\beta}, \qquad |\alpha|, |\beta| \le d, \tag{1}$$

so that $\vec{p}^{\top} M_d(y)\, \vec{q} = \int pq \, \mathrm{d}\mu$ whenever y is the moment sequence of μ. For a set I of polynomials we write $V_{\mathbb{R}}(I) = \{x \in \mathbb{R}^n : f(x) = 0 \ \forall f \in I\}$ for its real zero set, and we write $d_{\mathrm{H}}(A, B) = \max\{\sup_{a \in A} \inf_{b \in B} \|a - b\|, \ \sup_{b \in B} \inf_{a \in A} \|a - b\|\}$ for the Hausdorff distance.

3.2 Problem statement

Problem 1. Let μ be a Borel measure on $\mathbb{R}^n$ supported on a compact semialgebraic set S given by

$$f_i(x) = 0, \qquad i = 1, 2, \dots, k, \tag{2}$$

$$-1 \le x_j \le 1, \qquad j = 1, 2, \dots, n, \tag{3}$$

$$g_j(x) \ge 0, \qquad j = 1, 2, \dots, l, \tag{4}$$

where (3) normalizes the ambient box. Let

$$\mathcal{Z} := \{x \in \mathbb{R}^n : f_i(x) = 0, \ i = 1, \dots, k\},$$

suppose that $\mathcal{Z}$ has dimension $s < n$, and let $\lambda := \mathcal{H}^s|_{\mathcal{Z}}$ be its intrinsic reference measure, where $\mathcal{H}^s$ denotes s-dimensional Hausdorff measure. Suppose that μ is absolutely continuous with respect to λ on S,

$$\mathrm{d}\mu(x) = w(x)\, \mathrm{d}\lambda(x), \qquad x \in S,$$

with density bounded away from zero, $w(x) \ge w_0 > 0$ on S. Suppose finally that the moments of μ up to some degree $2d$ are known,

$$y_\alpha = \int x^\alpha \, \mathrm{d}\mu(x), \qquad \alpha \in \mathbb{N}^n, \ |\alpha| \le 2d,$$

and that *no* samples from μ are available. Given $\delta > 0$, find a finite set S_δ such that

$$d_{\mathrm{H}}(S_\delta, \operatorname{supp} \mu) < \delta. \tag{5}$$

A set S_δ satisfying (5) is called a *δ-accurate approximation* of $\operatorname{supp} \mu$.

Three features of Problem 1 determine our approach: the support has empty interior in $\mathbb{R}^n$, so the full-dimensional theory of the CD kernel does not apply; S is described by equations and inequalities together; and only moments are given, so any samples must be generated by the method itself (Section 4.2).

4 Method

The procedure is set out in Algorithm 1. The moment matrix is used twice: its kernel gives the algebraic carrier $\mathcal{Z}$ (Section 4.1), and its pseudoinverse gives the relative CD kernel, which serves as a membership oracle on $\mathcal{Z}$ (Section 4.3). Between these two steps, Section 4.2 generates the points of $\mathcal{Z}$ at which the oracle is evaluated. Section 4.4 describes the output and the choice of parameters.

4.1 Step 1: the algebraic carrier from the kernel of the moment matrix

Let $f_1, \dots, f_k$ be the polynomials in (2), and let $\mathcal{Z}$ be their real zero set. The first step computes polynomials in the vanishing ideal of $\operatorname{supp} \mu$. These need not be the f_i themselves, since the ideal generated by $f_1, \dots, f_k$ need not be real radical (Lasserre et al., 2013). The basis of this step is the following classical correspondence (Laurent, 2009).

Proposition 1 (Kernel of the moment matrix and support)**.** *Let μ be a finite nonnegative Borel measure on $\mathbb{R}^n$ with moments $y_\alpha = \int x^\alpha \, \mathrm{d}\mu(x)$ up to degree $2d$, and let $M_d(y)$ be the associated moment matrix. Then for every $f \in \mathbb{R}[x]_d$,*

$$f \in \ker M_d(y) \quad \Longleftrightarrow \quad \operatorname{supp} \mu \subseteq V_{\mathbb{R}}(f) := \{x \in \mathbb{R}^n : f(x) = 0\},$$

or equivalently $\ker M_d(y) = \{f \in \mathbb{R}[x]_d : f|_{\operatorname{supp} \mu} = 0\}$.

Algorithm 1 `AlgSup`: recovery of a lower-dimensional semialgebraic support

Require: Degree d; moments $y = (y_\alpha)_{|\alpha|\le 2d}$; sampling budget N; threshold τ_d.
Ensure: Semialgebraic set S_{d,τ_d} and a finite sample $\mathcal{M}_{\mathrm{adm}} \subset S_{d,\tau_d}$.

1: **Recover the algebraic carrier $\mathcal{Z}$.**
(a) Form the moment matrix $M_d(y)$ as in (1).
(b) Compute $I_d := \ker M_d(y)$ and extract generators degree by degree.
(c) Set $\mathcal{Z} := V_{\mathbb{R}}(I_d)$.

2: **Sample the carrier.**
(a) Draw N points on $\mathcal{Z}$, uniformly with respect to $\mathcal{H}^s|_{\mathcal{Z}}$, by random slicing (Breiding & Marigliano, 2020).
(b) Denote the resulting set of points by $\mathcal{M}$.

3: **Form the relative CD kernel on $\mathcal{Z}$.**
(a) Compute the pseudoinverse $A := M_d(y)^\dagger$.
(b) Set $\kappa_d(x) := v_d(x)^\top A\, v_d(x)$ for $x \in \mathcal{Z}$.

4: **Threshold.**
(a) Evaluate κ_d on $\mathcal{M}$.
(b) Retain $\mathcal{M}_{\mathrm{adm}} := \{x \in \mathcal{M} : \kappa_d(x) \le \tau_d\}$.

5: **return** $S_{d,\tau_d} = \{x \in \mathcal{Z} : \kappa_d(x) \le \tau_d\}$ and $\mathcal{M}_{\mathrm{adm}}$.

Thus $V_{\mathbb{R}}(\ker M_d(y))$ always contains the Zariski closure of $\operatorname{supp}\mu$, and equals it once d is at least the largest degree in a generating set of the vanishing ideal; this closure lies in $\mathcal{Z}$, and may be smaller if $\operatorname{supp}\mu$ misses a component of $\mathcal{Z}$. We denote the recovered carrier again by $\mathcal{Z}$. A numerical basis of $\ker M_d(y)$ mixes generators with their polynomial multiples, so generators are extracted degree by degree (Section A); in our experiments the relations appear in the first degree in which the kernel is nontrivial, and no reduction is needed. Step 1 does not recover the inequalities (4), since $\ker M_d(y)$ depends only on the Zariski closure of $\operatorname{supp}\mu$.

4.2 Step 2: sampling the algebraic carrier

Problem 1 does not permit sampling from μ, but $\mathcal{Z}$ is given explicitly by its equations and can be sampled. We use the random-slicing method of Breiding & Marigliano (2020), implemented in `HomotopyContinuation.jl` (Breiding & Timme, 2018), which produces points distributed uniformly with respect to the intrinsic volume of $\mathcal{Z}$.

Briefly, $\mathcal{Z}$ is intersected with a random affine subspace $\mathcal{L}_{A,b} = \{x : Ax = b\}$ of complementary dimension, where $A \in \mathbb{R}^{s\times n}$ and $b \in \mathbb{R}^s$; the intersection consists of finitely many points, which are found by solving one polynomial system. If (A, b) is drawn from a Gaussian density reweighted by a weighted count of the intersection points, and one of these points is then chosen with probability proportional to the reciprocal of its weight, the resulting point is uniformly distributed on $\mathcal{Z}$ (Breiding & Marigliano, 2020). The weights and the rejection step are described in Section B.

The samples lie exactly on $\mathcal{Z}$, so no point is accepted merely because it lies near the support in the ambient space, and since they are uniform in intrinsic volume, the resolution of the sample depends only on the budget N.

4.3 Step 3: the relative Christoffel–Darboux kernel on the carrier

Let μ and $M_d(y)$ be as above. Under Problem 1, $M_d(y)$ is singular as soon as d is at least the lowest degree of a polynomial vanishing on $\operatorname{supp}\mu$. The bilinear form $\langle p, q\rangle_\mu = \int pq\,\mathrm{d}\mu$ on $\mathbb{R}[x]_d$ has nullspace

$$I_d := \ker M_d(y) = \Big\{p \in \mathbb{R}[x]_d : \int p^2\,\mathrm{d}\mu = 0\Big\},$$

which is the space already considered in Section 4.1. The quotient $\mathcal{H}_d := \mathbb{R}[x]_d/I_d$ is a finite-dimensional Hilbert space, which we identify with a space of polynomial functions on the carrier $\mathcal{Z} = V_{\mathbb{R}}(I_d)$. If $\{P_1, \dots, P_{r_d}\}$ is an orthonormal basis of $\mathcal{H}_d$, the *relative* CD kernel is $\kappa_d(x, z) = \sum_{j=1}^{r_d} P_j(x)P_j(z)$ for $x, z \in \mathcal{Z}$, and on the diagonal

$$\kappa_d(x) = \kappa_d(x, x) = v_d(x)^\top M_d(y)^\dagger v_d(x), \qquad x \in \mathcal{Z}, \tag{6}$$

where $M_d(y)^\dagger$ is the Moore–Penrose pseudoinverse; the Christoffel function is $\Lambda_d = 1/\kappa_d$. The quotient construction and the formula involving the pseudoinverse are developed in Pauwels et al. (2021); Lasserre et al. (2022). Note that (6) can be computed from the same data as Step 1, and that it is well defined, although $M_d(y)$ is singular, because the pseudoinverse is used.

Given a threshold τ_d, membership is decided by the sublevel set

$$S_{d,\tau_d} := \{x \in \mathcal{Z} : \kappa_d(x) \le \tau_d\}. \tag{7}$$

The optimal threshold τ_d^* is the one which minimizes $d_{\mathrm{H}}(S_{d,\tau_d}, S)$. When the moment matrix comes from a moment relaxation of a POP, an estimate $\hat{\tau}_d$ of τ_d^* is given by the largest threshold for which all the retained points $\mathcal{M}_{\mathrm{adm}}$ are feasible; we do not use this rule in the experiments. The choice is motivated by the dichotomy of the Christoffel function: κ_d remains relatively small on the support and grows quickly away from it, typically exponentially in d.

This dichotomy has been proved only under additional assumptions on S and μ. For full-dimensional supports, Lasserre & Pauwels (2019) establish Hausdorff convergence of suitable sublevel sets of the Christoffel function, and Vu et al. (2022) give rates for finite samples. These results require, in particular, that $S = \overline{\mathrm{int}(S)}$, and so they do not cover Problem 1, in which S has empty interior in $\mathbb{R}^n$. For lower-dimensional algebraic supports the relative kernel is still well defined (Pauwels et al., 2021), but in our setting no computable thresholds τ_d are known which guarantee that $d_{\mathrm{H}}(S_{d,\tau_d}, \operatorname{supp}\mu) \to 0$. We therefore use (7) as a heuristic.

Bentancur et al. (2026) introduce a mollified CD kernel and prove a dichotomy of this kind under intrinsic regularity assumptions. Its evaluation, however, requires a reference measure, mollifiers on the recovered variety and the associated integrals, none of which is supplied by the moments. We therefore use the ordinary relative CD kernel, and regard the mollified kernel as a route to guarantees in future work.

4.4 Output and choice of parameters

Algorithm 1 returns two objects. The set S_{d,τ_d} of (7) is semialgebraic by construction, since $\mathcal{Z}$ is algebraic and κ_d is a polynomial of degree $2d$. The method therefore describes the estimated support in the same terms as (2)–(4), in addition to providing a finite set of points. That set, $\mathcal{M}_{\mathrm{adm}}$, is a sample from S_{d,τ_d}, and the Hausdorff distances reported in Section 5 are computed for it.

Two parameters are not determined by the theory. There is no formula for τ_d in terms of the moments, and we choose it from the distribution of the values of κ_d on $\mathcal{M}$, which is bimodal whenever the dichotomy is pronounced (Figure 1, left). Nor can the smallest d achieving a given δ be predicted, and in any case the order of the relaxation bounds d from above. Algorithm 1 therefore does not guarantee (5) for a prescribed δ; the experiments report the value of δ attained.

5 Experiments

We test Algorithm 1 in two settings: with moments computed in closed form for measures which satisfy Problem 1 exactly (Section 5.2), and with pseudo-moment sequences obtained from POP relaxations (Section 5.3). In the second setting the hypotheses need not hold: the sequence need not be a moment sequence, and if it is one, the representing measure may be supported on any subset of the set of minimizers. We therefore regard the POP experiments as exploratory; they test whether the method extracts useful geometric information beyond the setting of exact moments.

5.1 Setup and baseline

In all experiments the support is estimated by thresholding the relative CD kernel (6); the two settings differ only in the origin of the moments. The comparison concerns the points at which the kernel

	Algorithm 1			CD, ambient grid		
Experiment	d_v	Samples	d_{H}	d_g	Samples	d_{H}
Moments in closed form (Section 5.2)						
Parabola $y = x^2,\ 0 \le x \le 1$	2	1.1×10^4	**0.014**	4	10^6	0.095
Hyperbola $xy = 1,\ 0.5 \le x \le 1$	2	4945	**0.002**	4	10^6	0.114
Hyperbola 2 $x^2 - y^2 = 1,\ 1 \le x \le 2$	2	1.4×10^4	**0.003**	4	10^6	0.096
3D cone (Section 5.2) $z^2 - x^2 - y^2 = 0,$ $z \in [-1, -0.5] \cup [0.5, 1]$	6	2×10^4	**0.090**	8	10^6	0.240
Cylinder in 3D $x^2 + y^2 = 1,\ z \in [0, 1]$	2	2.3×10^4	**0.058**	8	10^6	0.982
Pseudo-moments from POP relaxations (Section 5.3)						
Line segment (Section D) $y = x,\ -0.5 \le x \le 0.5$	1	1.1×10^4	**0.002**	2	10^6	0.003
Semicircle (Section E) $x^2 + y^2 = 1,\ y \ge 0$	2	2.2×10^4	**0.050**	4	10^6	0.270
Product of segments in 10D (Section 5.3.2)	1	5×10^4	**0.530**	2	9.7×10^6	1.32
Product measure in 4D (Section F)	2	10^4	**0.1**	4	6.2×10^6	0.246

Table 1: Support recovery by Algorithm 1 and by CD thresholding on an ambient grid. Here d_v and d_g are the kernel degrees, *Samples* is the number of points at which the kernel is evaluated, and d_{H} is the Hausdorff distance to the true support, computed against a fine discretization of $\operatorname{supp}\mu$. The grid baseline uses the higher degree and more samples in every row, and has the larger error in every row.

is evaluated. Our **carrier** estimator evaluates κ_{d_v} on $\mathcal{M} \subset \mathcal{Z}$, sampled as in Section 4.2. The **grid** estimator, which serves as the baseline, evaluates κ_{d_g} on a uniform grid in the ambient box (3); it is CD thresholding without Steps 1 and 2. Both estimators use moments of the same measure or relaxation, and in every case the grid is allowed the higher degree $d_g \ge d_v$. Off $\mathcal{Z}$ the pseudoinverse ignores the component of $v_d(x)$ in $\ker M_d(y)$; a regularized kernel built from $(M_d(y) + \varepsilon I)^{-1}$ would penalize that component, and is a natural further baseline which we have not evaluated. In each example τ_d is chosen by hand, usually as a percentile of the values of κ_d on the samples from the variety, and the reported errors depend on this choice. For each estimator we record the degree, the number of samples and the Hausdorff distance d_{H} between the accepted points and the true support, computed against a fine discretization of S. The results are collected in Table 1.

5.2 Cone moments computed in closed form

Here μ is the normalized uniform measure, with respect to intrinsic surface area, on a set which satisfies Problem 1 exactly, and its moments are computed in closed form. We describe the cone, which is the most intricate of the five examples: it has codimension one in $\mathbb{R}^3$, it has a singular point at the origin, and its support, which is the point of the experiment, has *two* connected components,

$$S = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 :\ x_3^2 - x_2^2 - x_1^2 = 0,\ x_3 \in [-1, -0.5] \cup [0.5, 1]\right\}.$$

Step 1 finds the polynomial $f_1(x) = x_3^2 - x_2^2 - x_1^2$ in $\ker M_2(y)$; this is the polynomial which defines $\mathcal{Z}$ in (2). The singular point of the cone lies outside S and is met by random slices with probability zero. The CD kernel of degree 6, which requires $M_6(y)$, is then evaluated at 2×10^4 samples on the cone and thresholded at $\tau_6 = 82.7$, and 10600 points are retained. The projection of the accepted set on the x_3-axis is $[-0.958, -0.489] \cup [0.489, 0.958]$, against the true $[-1, -0.5] \cup [0.5, 1]$. Both components are recovered, and recovered separately, although the method is not told how many there are; the error is concentrated at the outer rims, where the density is lowest (Figure 1). The Hausdorff distance $d_{\mathrm{H}}(S^{\mathrm{variety}}_{6,\tau_6}, S) = 0.09$ was computed against an ε-net of S constructed by hand, with $\varepsilon = 0.01$. Thresholding 10^6 grid points gives $d_{\mathrm{H}}(S^{\mathrm{grid}}_{8,\tau_8}, S) = 0.240$.

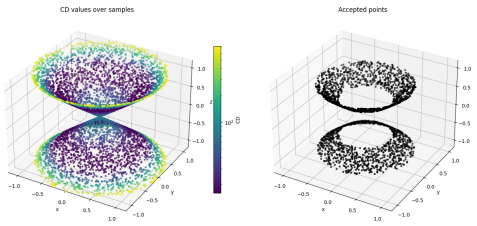

Figure 1: The cone, with moments in closed form. *Left:* values of κ_6 at 20000 points sampled uniformly on the recovered carrier $\{x_3^2 - x_2^2 - x_1^2 = 0\}$, on a logarithmic color scale; the dichotomy appears as dark inner annuli and bright rims. *Right:* the 10600 points accepted at $\tau_6 = 82.7$; both connected components are recovered.

5.3 Pseudo-moments from polynomial optimization

The set of minimizers of a POP is semialgebraic, and when it is infinite it has positive dimension; this is the setting of Problem 1. We solve the POPs by the moment relaxation (Lasserre, 2001), as implemented in TSSOS (Magron & Wang, 2021a;b). Under the usual assumptions of the moment–SOS hierarchy the relaxations converge to the global optimal value, but they do not return the minimizers: one obtains a finite pseudo-moment sequence, from which the minimizers must be located. When the objective and constraints are known, samples on $\mathcal{Z}$ could also be filtered by them directly; we do not do so, since our aim is to test what can be recovered from the pseudo-moments alone. When the flat extension property (Curto & Fialkow, 1998) holds, the GNS construction (López Quijorna, 2021) solves the extraction problem. In the examples below the flat extension property fails, and the GNS construction cannot be used.

5.3.1 The polynomial optimization setting

In each of the following examples we solve a polynomial optimization problem

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & F(x) \\ \text{s.t.} \quad & g_j(x) \geq 0, \quad j = 1, 2, \ldots, l, \end{aligned}$$

whose set of minimizers S is known in closed form. The relaxation returns a moment matrix $M_d(y)$, which we pass to Algorithm 1 as though it had been generated by a measure supported on S. With $F(x) = (x_1 - x_2)^2$ and suitable constraints g_j, the set S is a line segment (Section D); with $F(x) = (x_1^2 + x_2^2 - 1)^2$ it is a subset of the circle (Section E). In both cases S is recovered to the accuracy reported in Table 1.

5.3.2 A product of segments in $2n$ dimensions

To examine how the approach scales, we consider a family of examples in dimensions $2n \in \{2, 4, 6, 8, 10\}$. We solve, at relaxation degree 2,

$$\begin{aligned} \min_{x \in \mathbb{R}^{2n}} \quad & F(x) = \sum_{i=1}^{n} (x_{2i-1} - x_{2i})^2 \\ \text{s.t.} \quad & -0.5 \leq x_j \leq 0.5, \qquad j = 1, \ldots, 2n, \end{aligned} \tag{8}$$

whose set of minimizers is $S^n \subset \mathbb{R}^{2n}$, where $S = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 = x_2,\ -0.5 \leq x_1 \leq 0.5\}$. The objective is chosen to be quadratic, so that its evaluation requires moments only up to degree 2. A dense moment matrix $M_d(y)$ of order d has $\binom{d+2n}{d}^2$ entries, and raising the relaxation degree soon becomes prohibitively expensive. We study how the number of evaluation points needed to reach a prescribed Hausdorff accuracy grows with the dimension. The Hausdorff distances in Figure 2 are computed against a tensor-product discretization of S^n with 20 points per intrinsic coordinate.

For the stricter target $d_{\mathrm{H}} < 0.6$, the number of carrier samples required grows by about one order of magnitude for each additional intrinsic dimension over the range tested. For the grid baseline we also solve the POP at relaxation degree 3 and allow it to use κ_3. It succeeds for $2n = 2$; for

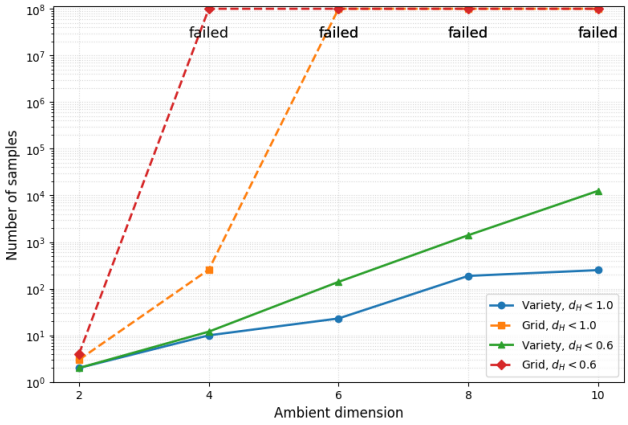

Figure 2: Number of evaluation points needed to reach Hausdorff errors $d_{\mathrm{H}} < 1.0$ and $d_{\mathrm{H}} < 0.6$. The carrier estimator uses κ_1 on $\mathcal{Z}$; the grid baseline is allowed κ_3 in $\mathbb{R}^{2n}$. Failed runs are plotted at 10^8 for display only; they indicate failure within the tested budget and are not estimates of the number of samples needed.

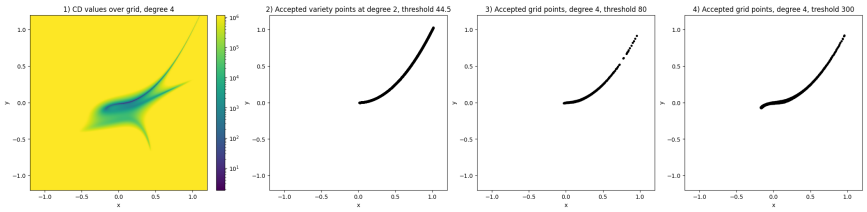

Figure 3: The parabola $y = x^2$, $0 \le x \le 1$, with moments in closed form. *1)* The CD kernel of degree 4 on the grid. *2)* Points of $\mathcal{Z}$ accepted by Algorithm 1, $d_{\mathrm{H}} = 0.014$. *3)* Grid points accepted at threshold 80; the estimate has large gaps, $d_{\mathrm{H}} = 0.095$. *4)* Grid points accepted at the higher threshold 500; the gaps are filled, but points far from the support are admitted, $d_{\mathrm{H}} = 0.17$.

$2n = 4$ it needs 256 grid points to reach $d_{\mathrm{H}} \le 1$, and with up to 10^8 grid points it does not reach $d_{\mathrm{H}} < 0.6$ for any of the thresholds tried. For $2n \ge 8$ memory is the limit: a grid with 9 points per coordinate has $9^8 \approx 4.3 \times 10^7$ points, and evaluating the degree-2 CD kernel on it needs about 115 GiB. The CD kernel of degree 1 cannot be used on the grid, because it always accepts the point $x = (-1, 1, -1, 1, \ldots)$ (see Figure 4), and so forces the Hausdorff distance to be at least $\mathrm{dist}(x, S^n) = \sqrt{2n}$. Restriction to the carrier thus reduces the dimension to be discretized from $2n$ to n.

A four-variable example with quartic terms, the product of a semicircle and a segment, is given in Section F.

5.4 The effect of restricting to the carrier

Table 1 shows that the carrier estimator has a smaller Hausdorff error than the grid in all nine experiments, and Figure 3 illustrates the reason. Samples drawn on $\mathcal{Z}$ satisfy its defining equations to machine precision, with residuals of order 10^{-15}, whereas even grid points close to $\mathcal{Z}$ have residuals of order 10^{-3} to 10^{-1}. In our experiments the region in which the CD kernel is small forms a thin neighbourhood of the support, which a grid may fail to resolve. A small threshold then leaves gaps in the estimated support; a larger threshold fills the gaps, but admits points farther from $\mathcal{Z}$ and so increases the Hausdorff error. Evaluating the kernel only on $\mathcal{Z}$ avoids this compromise.

In dimension ten (Section 5.3.2) the resolution of the grid itself becomes a limitation. Indeed, the Hausdorff error of any estimator which selects points from a grid can be bounded below, whatever function is used to make the selection. Let

$$G_m = \{-1, -1 + \tfrac{2}{m}, \ldots, 1\}^n$$

denote the uniform grid of $(m+1)^n$ points in $[-1,1]^n$.

Definition 1 (Dispersion)**.** Let $G \subset \mathbb{R}^n$ be finite and $S \subseteq \mathbb{R}^n$ compact. The *dispersion* of G relative to S is

$$\operatorname{disp}(G,S) := \sup_{y\in S} \operatorname{dist}(y,G) = \sup_{y\in S} \min_{x\in G} \|x-y\|,$$

that is, the smallest r for which G is an r-net of S (Vershynin, 2018, §4.2); it is also known as the *fill distance* (Niederreiter, 1992). It is the one-sided Hausdorff distance from S to G, and it does not depend on how points of G are subsequently selected.

Theorem 1. *Let $S \subseteq \mathcal{Z} \cap [-1,1]^n$ be compact.*

(i) *For every m and every $\hat{S} \subseteq G_m$, $d_{\mathrm{H}}(\hat{S},S) \geq \operatorname{disp}(G_m,S)$.*

(ii) *Let $\mathcal{M} \subset \mathcal{Z}$ be finite and $\hat{S} = \{x \in \mathcal{M} : \operatorname{dist}(x,S) \leq \rho\}$. If $\rho \geq \operatorname{disp}(\mathcal{M},S)$, then $d_{\mathrm{H}}(\hat{S},S) \leq \rho$.*

(iii) *Let $S_n = \{t\mathbf{1} : |t| \leq \frac{1}{2}\} \subset \mathbb{R}^n$, the set of minimizers of $\sum_{i<j}(x_i-x_j)^2$ over $[-\frac{1}{2},\frac{1}{2}]^n$, with carrier $\mathcal{Z} = \mathbb{R}\mathbf{1}$. Then $\operatorname{disp}(G_m,S_n) \geq \sqrt{n}/m$ for every $m \geq 1$. On the other hand, if $\mathcal{M}$ consists of $N \geq 4\ln(N/\epsilon)$ points drawn uniformly from $\mathcal{Z} \cap [-1,1]^n$ and $\hat{S} = \mathcal{M} \cap S_n$, then with probability at least $1-\epsilon$*

$$d_{\mathrm{H}}(\hat{S},S_n) \;\leq\; 4\sqrt{n}\,\frac{\ln(N/\epsilon)}{N}.$$

Theorem 1 concerns only the choice of evaluation points. Part (i) holds for every selection rule, and in particular for CD thresholding; parts (ii) and (iii) use the ideal rule $\operatorname{dist}(x,S) \leq \rho$, which is not available in practice, so they bound what the carrier can achieve, not what Algorithm 1 achieves. Whether CD thresholding comes close to the ideal rule is the question tested by the experiments; S_n is a simplified model of Section 5.3.2. Thus, for the same budget of $N = (m+1)^n$ evaluations, the two errors on S_n satisfy $d_{\mathrm{H}}^{\mathrm{grid}}/d_{\mathrm{H}}^{\mathrm{carrier}} \geq (m+1)^n \big/ \big(4m\ln((m+1)^n/\epsilon)\big)$, a ratio which grows exponentially in n for every fixed m. To bring the error below a prescribed $\delta < \sqrt{n} = \operatorname{diam} S_n$, the grid needs at least $(1+\sqrt{n}/\delta)^n$ evaluations, and the carrier needs $O\big(\sqrt{n}\,\delta^{-1}\ln(\sqrt{n}/\delta\epsilon)\big)$. The count is of evaluations, not of computing time: each sample on $\mathcal{Z}$ costs one polynomial-system solve. The proof is in Section C; part (iii) does not depend on the grid being cubic (Remark 3).

6 Conclusion and limitations

We have given a method for estimating, from its moments alone, a support which is compact, semialgebraic and of positive codimension. The kernel of the moment matrix recovers the equations, and the relative Christoffel–Darboux kernel, evaluated on a uniform sample of the recovered carrier, recovers the inequalities. In our experiments CD thresholding on an ambient grid has the larger Hausdorff error in every case, although it is run at a higher degree; the margin is negligible for the segment and large elsewhere. For a model problem, and with an ideal selection rule on the carrier, Theorem 1 shows that the gap in attainable error grows exponentially with the dimension.

Several questions remain open. Step 1 is exact for true moments, but for the pseudo-moments of Section 5.3 the kernel of $M_d(y)$ is computed numerically and its singular-value cut-off is set by inspection; a perturbation theory along the lines of Klep et al. (2018) is needed. Step 3 relies on a dichotomy proved for full-dimensional supports (Lasserre & Pauwels, 2019; Vu et al., 2022) and, for the mollified kernel, on varieties (Bentancur et al., 2026); for the ordinary relative kernel in codimension at least one we know of no computable thresholds for which $d_{\mathrm{H}}(S_{d,\tau_d}, \operatorname{supp}\mu) \to 0$, hence no rate and no principled choice of degree or threshold, and we have not studied the sensitivity of the results to τ_d. The sampler assumes a smooth carrier, and the pseudoinverse of order $\sigma(d) = \binom{n+d}{d}$ dominates the cost.

AI use statement

Large language models were used only as an aid to writing and formatting, that is, for copy-editing and for LaTeX. The mathematical statements, the algorithm, the implementation and the experimental results are the authors' own, and the authors take full responsibility for the content of this paper.

Reproducibility statement

Algorithm 1 is stated in full, and the parameters needed to reproduce Table 1 (the kernel degrees, the thresholds, the sampling budgets and the defining equations of each example) are given in Section 5 and Section G. The two external components, the random-slicing sampler of Breiding & Marigliano (2020) in `HomotopyContinuation.jl` (Breiding & Timme, 2018) and the moment relaxations of TSSOS (Magron & Wang, 2021a;b), are both publicly available.

References

N.I. Akhiezer. *The Classical Moment Problem and Some Related Questions in Analysis*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 2020. ISBN 9781611976397. URL https://books.google.cz/books?id=6c0MEAAAQBAJ.

Leandro Bentancur, Didier Henrion, and Mauricio Velasco. Mollified christoffel-darboux kernels and density recovery on varieties, 2026.

Paul Breiding and Orlando Marigliano. Random points on an algebraic manifold. *SIAM Journal on Mathematics of Data Science*, 2(3):683–704, 2020. doi: 10.1137/19M1271178. URL https://doi.org/10.1137/19M1271178.

Paul Breiding and Sascha Timme. Homotopycontinuation.jl: A package for homotopy continuation in julia. In *Mathematical Software – ICMS 2018*, volume 10931 of *Lecture Notes in Computer Science*, pp. 458–465. Springer, 2018. doi: 10.1007/978-3-319-96418-8_54.

David A. Cox. Stickelberger and the Eigenvalue Theorem, 2020.

Raúl E. Curto and Lawrence A. Fialkow. *Flat Extensions of Positive Moment Matrices: Recursively Generated Relations*, volume 648 of *Memoirs of the American Mathematical Society*. American Mathematical Society, 1998. ISBN 978-1-4704-0237-2.

Yohann de Castro and Fabrice Gamboa. Exact reconstruction using beurling minimal extrapolation. *Journal of Mathematical Analysis and Applications*, 395(1):336–354, 2012. ISSN 0022-247X. doi: https://doi.org/10.1016/j.jmaa.2012.05.011.

Alex Devonport, Forest Yang, Laurent El Ghaoui, and Murat Arcak. Data-driven reachability analysis and support set estimation with Christoffel functions, 2021.

Vincent Duval and Gabriel Peyré. Exact support recovery for sparse spikes deconvolution. *Foundations of Computational Mathematics*, 15(5):1315–1355, 2015. doi: 10.1007/s10208-014-9228-6.

Didier Henrion and Jean-Bernard Lasserre. Detecting global optimality and extracting solutions in gloptipoly. In *Positive Polynomials in Control*, volume 312 of *Lecture Notes in Control and Information Science*, pp. 293–310. Springer, Berlin, Heidelberg, 2005. doi: 10.1007/10997703_15.

Igor Klep, Janez Povh, and Jurij Volčič. Minimizer extraction in polynomial optimization is robust. *SIAM Journal on Optimization*, 28(4):3177–3207, 2018. doi: 10.1137/17M1152061.

Jean B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001. doi: 10.1137/S1052623400366802.

Jean B Lasserre and Edouard Pauwels. The empirical Christoffel function with applications in data analysis. *Advances in Computational Mathematics*, 45(3):1439–1468, 2019.

Jean B. Lasserre, Monique Laurent, Bernard Mourrain, Philipp Rostalski, and Philippe Trébuchet. Moment matrices, border bases and real radical computation. *Journal of Symbolic Computation*, 51:63–85, 2013. doi: 10.1016/j.jsc.2012.06.015.

Jean Bernard Lasserre, Edouard Pauwels, and Mihai Putinar. *The Christoffel–Darboux Kernel for Data Analysis*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 2022.

Monique Laurent. Sums of Squares, Moment Matrices and Optimization Over Polynomials. In Mihai Putinar and Seth Sullivant (eds.), *Emerging Applications of Algebraic Geometry*, volume 149 of *The IMA Volumes in Mathematics and its Applications*, pp. 157–270. Springer, New York, NY, 2009. ISBN 978-0-387-09686-5. doi: 10.1007/978-0-387-09686-5_7. URL https://doi.org/10.1007/978-0-387-09686-5_7.

Monique Laurent and Bernard Mourrain. A generalized flat extension theorem for moment matrices. *Archiv der Mathematik*, 93(1):87–98, 2009. doi: 10.1007/s00013-009-0017-4.

Bruce G. Lindsay. *Mixture Models: Theory, Geometry and Applications*. Lecture Notes. Institute of Mathematical Statistics, 1995.

María López Quijorna. Detecting optimality and extracting solutions in polynomial optimization with the truncated gns construction. *Journal of Global Optimization*, 81(3):559–598, 2021. doi: 10.1007/s10898-020-00987-9.

Victor Magron and Jie Wang. Tssos: a julia library to exploit sparsity for large-scale polynomial optimization. *arXiv preprint arXiv:2103.00915*, 2021a. doi: 10.48550/arXiv.2103.00915.

Victor Magron and Jie Wang. TSSOS: a Julia library to exploit sparsity for large-scale polynomial optimization. In *Effective Methods in Algebraic Geometry*, Tromso, Norway, June 2021b. doi: 10.48550/arXiv.2103.00915.

Swann Marx, Edouard Pauwels, Tillmann Weisser, Didier Henrion, and Jean Bernard Lasserre. Semi-algebraic approximation using christoffel–darboux kernel. *Constructive Approximation*, 54: 391–429, 2021. doi: 10.1007/s00365-021-09535-4. URL https://doi.org/10.1007/s00365-021-09535-4.

Jiawang Nie. The $\mathcal{A}$-truncated K-moment problem. *Foundations of Computational Mathematics*, 14(6):1243–1276, 2014. doi: 10.1007/s10208-014-9225-9. URL https://doi.org/10.1007/s10208-014-9225-9.

Harald Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*, volume 63 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, Philadelphia, 1992. doi: 10.1137/1.9781611970081.

Partha Niyogi, Stephen Smale, and Shmuel Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete and Computational Geometry*, 39(1–3):419–441, 2008. doi: 10.1007/s00454-008-9053-2.

Pablo A Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical programming*, 96(2):293–320, 2003. doi: 10.1007/s10107-003-0387-5.

Edouard Pauwels, Mihai Putinar, and Jean-Bernard Lasserre. Data analysis from empirical moments and the christoffel function. *Foundations of Computational Mathematics*, 21(1):243–273, 2021. doi: 10.1007/s10208-020-09451-2.

Konrad Schmüdgen. *The Moment Problem*, volume 277 of *Graduate Texts in Mathematics*. Springer International Publishing AG, Cham, Switzerland, 2017. ISBN 978-3-319-64545-2. doi: 10.1007/978-3-319-64546-9.

Vilmos Totik. Christoffel functions on curves and domains. *Transactions of the American Mathematical Society*, 362(4):2053–2087, 2010. doi: 10.1090/s0002-9947-09-05059-4.

Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*, volume 47 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, 2018. doi: 10.1017/9781108231596.

Martin Vetterli, Paola Marziliano, and Thierry Blu. Sampling signals with finite rate of innovation. *IEEE Transactions on Signal Processing*, 50(6):1417–1428, 2002.

Mai Trang Vu, François Bachoc, and Edouard Pauwels. Rate of convergence for geometric inference based on the empirical Christoffel function. *ESAIM: Probability and Statistics*, 26:171–207, 2022. doi: 10.1051/ps/2022003.

Markus Wageringel. Truncated moment problems on positive-dimensional algebraic varieties. *arXiv preprint arXiv:2203.01269*, 2022. doi: 10.48550/arXiv.2203.01269.

A From kernel elements to generators

Remark 1. Even when the vanishing ideal $I(\operatorname{supp}\mu)$ is generated by a single polynomial f of degree d, the polynomial f can be read off directly from the kernel of the moment matrix only in the first degree in which a nontrivial relation appears. Indeed, if $I(\operatorname{supp}\mu)\cap\mathbb{R}[x]_{d-1}=\{0\}$ and $I(\operatorname{supp}\mu)\cap\mathbb{R}[x]_d=\operatorname{span}\{f\}$, then $\dim\ker M_d(y)=1$, and f may be recovered, up to a scalar factor, from a basis vector of $\ker M_d(y)$. In higher degrees the kernel also contains polynomial multiples of relations already found: $f, x_1f,\dots,x_nf\in\ker M_{d+1}(y)$, so that $\dim\ker M_{d+1}(y)>1$ although the vanishing ideal may still have a single generator. More generally, a numerical basis of $\ker M_d(y)$ consists of arbitrary linear combinations of generators of the vanishing ideal and of their polynomial multiples.

To extract generators one proceeds degree by degree, separating the new relations in each degree from those generated in lower degrees. Equivalently, in degree d one removes from $\ker M_d(y)$ the subspace generated by polynomial multiples of kernel elements of lower degree, and the remaining independent relations give new generators. In computational algebra this is done by row reduction with respect to a monomial ordering, by Gröbner-basis techniques, or by border-basis methods; see, for instance, Laurent & Mourrain (2009); Lasserre et al. (2013).

Remark 2 (Several generators). The same reasoning applies, under suitable assumptions, when $\mathcal{Z}$ is defined by several polynomials. Suppose that the first nontrivial relations in $I(\operatorname{supp}\mu)$ occur in degree d, and that

$$I(\operatorname{supp}\mu)\cap\mathbb{R}[x]_d=\operatorname{span}\{f_1,\dots,f_k\}.$$

Then $\dim\ker M_d(y)=k$. A numerical nullspace computation may return a different basis $v_j=\sum_{i=1}^k C_{ji}f_i$ with C invertible, but then $\{v_1=\cdots=v_k=0\}=\{f_1=\cdots=f_k=0\}$, and the same algebraic set is obtained. In polynomial optimization the f_i often involve disjoint cliques of variables, and each independent regular equation then typically lowers the local dimension by one. Since samples are drawn only on the recovered carrier, each such equation removes, before CD thresholding, one ambient dimension in which spurious samples could otherwise lie.

B The random-slicing sampler

We recall the construction of Breiding & Marigliano (2020) used in Section 4.2. Suppose that $\mathcal{Z}$ is a smooth algebraic manifold of dimension s, defined by $F=(f_1,\dots,f_k)$ as in (2), and of positive, finite volume with respect to the induced volume form. For $A\in\mathbb{R}^{s\times n}$ and $b\in\mathbb{R}^s$ put $\mathcal{L}_{A,b}=\{x\in\mathbb{R}^n : Ax=b\}$, and define

$$\bar{1}(A,b)=\sum_{x\in\mathcal{Z}\cap\mathcal{L}_{A,b}}\frac{1}{\alpha(x)},\qquad \bar{1}:\mathbb{R}^{s\times n}\times\mathbb{R}^s\to\mathbb{R}, \tag{9}$$

where $\alpha(x)>0$ is a normalizing factor computed from the Jacobian of F at x. If φ denotes the standard Gaussian density on $\mathbb{R}^{s\times n}\times\mathbb{R}^s$, the integral-geometric identity $\mathbb{E}_\varphi[\bar{1}]=\operatorname{vol}(\mathcal{Z})$ holds. The sampler draws (A,b) from the density

$$\psi(A,b)=\frac{\varphi(A,b)}{\operatorname{vol}(\mathcal{Z})}\sum_{x\in\mathcal{Z}\cap\mathcal{L}_{A,b}}\alpha(x)^{-1}$$

by rejection: one proposes $(A,b)\sim\varphi$ and accepts it with probability $\eta\,\bar{1}(A,b)$, where the constant η depends only on $\deg\mathcal{Z}$ and on a bound for $\|x\|$ over $\mathcal{Z}$ (Breiding & Marigliano, 2020). Each evaluation of $\bar{1}$ requires the solution of one polynomial system. Given an accepted (A,b), one returns $x\in\mathcal{Z}\cap\mathcal{L}_{A,b}$ with probability $\alpha(x)^{-1}/\bar{1}(A,b)$, and the resulting point is uniformly distributed on $\mathcal{Z}$.

C PROOF OF THEOREM 1

Throughout, $\operatorname{dist}(y, G) = \min_{x \in G} \|x - y\|$, and $d_{\mathrm{H}}(\emptyset, S) = +\infty$.

(i). For $\hat{S} \subseteq G_m$ and $y \in S$ we have $\operatorname{dist}(y, \hat{S}) \geq \operatorname{dist}(y, G_m)$. Hence $d_{\mathrm{H}}(\hat{S}, S) \geq \sup_{y \in S} \operatorname{dist}(y, \hat{S}) \geq \operatorname{disp}(G_m, S)$. In particular, this applies to $\hat{S} = G_m \cap \{\kappa_d \leq \tau\}$ for every d and τ.

(ii). Let $y \in S$. By hypothesis there is $x \in \mathcal{M}$ with $\|x - y\| \leq \operatorname{disp}(\mathcal{M}, S) \leq \rho$; then $\operatorname{dist}(x, S) \leq \rho$, so that $x \in \hat{S}$ and $\operatorname{dist}(y, \hat{S}) \leq \rho$. Conversely, every $x \in \hat{S}$ satisfies $\operatorname{dist}(x, S) \leq \rho$ by construction. Both one-sided terms of d_{H} are therefore at most ρ.

(iii): recovery of the carrier. An affine form $c_0 + \sum_i c_i x_i$ vanishes on S_n if and only if $c_0 = 0$ and $\sum_i c_i = 0$, so the affine part of $\ker M_1(y)$ is spanned by $x_i - x_1$, $i = 2, \ldots, n$, whose common zero set is the line $\mathbb{R}\mathbf{1}$. Step 1 therefore returns $\mathcal{Z}$ exactly in degree one, as in the case $n = 2$ of Section D.

(iii): the grid. The coordinates of points of G_m are $-1 + 2j/m$, $j = 0, \ldots, m$, and the centres of its cells have coordinates $-1 + (2j + 1)/m$, $j = 0, \ldots, m - 1$. Put $\gamma = 0$ if m is odd and $\gamma = 1/m$ if m is even. In either case γ is a coordinate of a cell centre and $|\gamma| \leq \frac{1}{2}$, so that $c = \gamma\mathbf{1} \in S_n$. Each coordinate of c is at distance exactly $1/m$ from the nearest grid coordinate, whence $\operatorname{dist}(c, G_m) = \sqrt{n}/m$ and $\operatorname{disp}(G_m, S_n) \geq \sqrt{n}/m$. Relative to $\operatorname{diam} S_n = \sqrt{n}$, the error of the grid is thus at least $1/m$, that is, $1/(N^{1/n} - 1)$ with $N = (m + 1)^n$ evaluations; it tends to zero only if the budget grows faster than every exponential in n.

(iii): the carrier. We have $\mathcal{Z} \cap [-1, 1]^n = \{t\mathbf{1} : |t| \leq 1\}$, and a uniform point of this set is $t\mathbf{1}$ with t uniform on $[-1, 1]$. Partition $[-\frac{1}{2}, \frac{1}{2}]$ into K intervals of length $1/K$. A sample falls in a given interval with probability $1/(2K)$, so the probability that some interval receives none of the N samples is at most $K(1 - \frac{1}{2K})^N \leq Ke^{-N/(2K)}$. On the complementary event every $t \in [-\frac{1}{2}, \frac{1}{2}]$ lies within $1/K$ of some sample $t_i \in [-\frac{1}{2}, \frac{1}{2}]$; the point $t_i\mathbf{1}$ belongs to S_n, and hence to $\hat{S}$, and $\operatorname{dist}(t\mathbf{1}, \hat{S}) \leq \sqrt{n}/K$. Since $\hat{S} \subseteq S_n$, the other one-sided term vanishes. Hence $d_{\mathrm{H}}(\hat{S}, S_n) \leq \sqrt{n}/K$ on that event. Take $K = \lfloor N/(2\ln(N/\epsilon)) \rfloor$. Then $e^{-N/(2K)} \leq \epsilon/N$, so the exceptional probability is at most $\epsilon K/N \leq \epsilon$; and $N \geq 4\ln(N/\epsilon)$ gives $K \geq N/(4\ln(N/\epsilon))$, whence $\sqrt{n}/K \leq 4\sqrt{n}\ln(N/\epsilon)/N$.

Consequences. At $N = (m + 1)^n$ the quotient of the two bounds is $(\sqrt{n}/m)\big/\big(4\sqrt{n}\ln(N/\epsilon)/N\big) = N/(4m\ln(N/\epsilon))$, which is the ratio stated after Theorem 1. For a target $\delta < \sqrt{n}$, (i) and (iii) require $\sqrt{n}/m \leq \delta$, that is, $m \geq \sqrt{n}/\delta > 1$ and $N = (m + 1)^n \geq (1 + \sqrt{n}/\delta)^n$. For the carrier, put $A = 4\sqrt{n}/\delta$ and $N = \lceil 2A\ln(2A/\epsilon) \rceil$; since $\ln\ln(2A/\epsilon) \leq \ln(2A/\epsilon)$, we have $A\ln(N/\epsilon) \leq N$, and so the bound in (iii) is at most δ.

Remark 3 (Arbitrary ambient designs). The lower bound does not depend on the grid being cubic. If $G \subset [-1, 1]^n$ is any set of N points and ω_n is the volume of the unit ball, then the N balls of radius r centred at the points of G cannot cover the box unless $N\omega_n r^n \geq 2^n$. Hence some $c \in [-1, 1]^n$ satisfies $\operatorname{dist}(c, G) \geq 2(N\omega_n)^{-1/n}$, and $\omega_n^{-1/n} \sim \sqrt{n/(2\pi e)}$. Any segment S through c, with $\mathcal{Z}$ the line containing it, then satisfies $\operatorname{disp}(G, S) \geq 2(N\omega_n)^{-1/n}$, and (i) applies. Thus any fixed set of N evaluation points in the ambient space has dispersion of order $N^{-1/n}$, whereas N uniform points on $\mathcal{Z}$ have dispersion of order $N^{-1/s}$, up to a logarithmic factor (Remark 4).

Remark 4 (General carriers). The bound for the carrier in (iii) is the one-dimensional case of a standard covering argument. Suppose that $\mathcal{H}^s(\mathcal{Z} \cap B(y, r)) \geq c\, r^s$ for all $y \in S$ and $r \leq r_0$, and let S be covered by $\mathcal{N}$ balls of radius $\delta/2 \leq r_0$ with centres in S. Then N uniform samples from $\mathcal{Z} \cap [-1, 1]^n$, a set of $\mathcal{H}^s$-measure V, satisfy $\operatorname{disp}(\mathcal{M}, S) \leq \delta$ with probability at least $1 - \epsilon$ as soon as $N \geq \frac{V}{c(\delta/2)^s} \ln \frac{\mathcal{N}}{\epsilon}$ (cf. Niyogi et al., 2008), and (ii) then bounds d_{H} by $\max(\delta, \rho)$. The budget grows like δ^{-s}, compared with δ^{-n} for the grid.

D THE LINE SEGMENT

The polynomial optimization problem for the line segment of Section 5.3.1, which we solve at relaxation order 2, is

$$\begin{aligned} \min_{x_1,x_2\in\mathbb{R}} \quad & F(x_1,x_2) = (x_1-x_2)^2 \\ \text{s.t.} \quad & x_2+0.5\geq 0, \quad 0.5-x_2\geq 0, \quad x_1+0.5\geq 0, \quad 0.5-x_1\geq 0, \end{aligned} \tag{10}$$

and its set of minimizers is $S = \{(x_1,x_2)\in\mathbb{R}^2 : x_1 = x_2,\ -0.5\leq x_1\leq 0.5\}$. The kernel of $M_1(y)$ is spanned by the vector $(0,\ 0.70710678,\ -0.70710678)$, which gives, up to a scalar factor, the polynomial $f_1(x) = x_2 - x_1$ of (2) defining the carrier $\mathcal{Z}$. The line is sampled over the box $[-1.2, 1.2]^2$ with `HomotopyContinuation.jl` (Breiding & Timme, 2018), and the grid consists of 1000×1000 points in $[-1,1]^2$. We now examine the behaviour of the CD kernel as the degree d increases.

Degree one. The CD kernel κ_1 of (6) is evaluated at 1.1×10^4 samples on $\mathcal{Z}$ and at 10^6 grid points. With the threshold $\tau_1 = 3.36$, 4873 points $S^{\text{variety}}_{1,\tau_1}$ of $\mathcal{Z}$ and 659528 grid points $S^{\text{grid}}_{1,\tau_1}$ are accepted, and $d_{\text{H}}(S^{\text{variety}}_{1,\tau_1}, S) = 0.002$, whereas $d_{\text{H}}(S^{\text{grid}}_{1,\tau_1}, S) = \sqrt{2}$. The kernel κ_1 and the accepted points are shown in Figure 4.

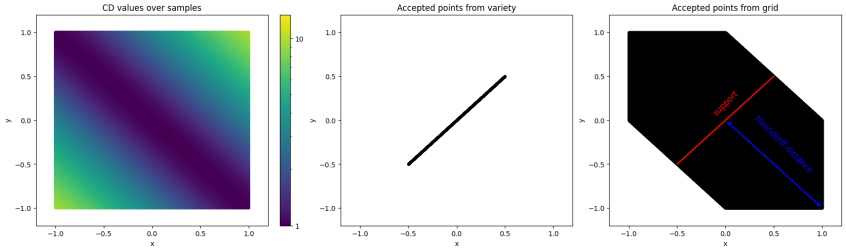

Figure 4: Thresholding the CD kernel κ_1 on the variety and on the grid. *Left:* values of κ_1 on the 1000×1000 grid. *Middle:* points of the variety accepted at $\tau_1 = 3.36$. *Right:* grid points accepted at $\tau_1 = 3.36$.

Degrees two and three. Since $S^{\text{variety}}_{1,\tau_1}$ already approximates S sufficiently well, we evaluate the kernels κ_2 and κ_3 on the grid only. The computation of κ_3 requires $M_3(y)$, and so (10) must be solved again at relaxation order 3. The Hausdorff distances are $d_{\text{H}}(S^{\text{grid}}_{2,\tau_2}, S) = 1.2$ and $d_{\text{H}}(S^{\text{grid}}_{3,\tau_3}, S) = 0.006$; the estimates are shown in Figure 5.

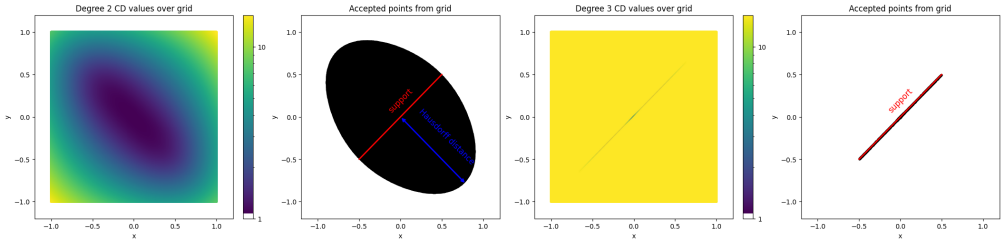

Figure 5: The CD kernels κ_2 and κ_3, evaluated and thresholded on the 1000×1000 grid. *Left:* values of κ_2. *Middle left:* grid points accepted at $\tau_2 = 3.36$. *Middle right:* values of κ_3. *Right:* grid points accepted at $\tau_3 = 5.0$.

E THE SEMICIRCLE

The polynomial optimization problem for the semicircle of Section 5.3.1, which we solve at relaxation order 2, is

$$\begin{aligned} \min_{x_1,x_2\in\mathbb{R}} \quad & F(x_1,x_2) = (x_1^2+x_2^2-1)^2 \\ \text{s.t.} \quad & x_2 \geq 0, \quad 1-x_2 \geq 0, \end{aligned}$$

and its set of minimizers is $S = \{(x_1,x_2)\in\mathbb{R}^2 : x_1^2+x_2^2 = 1,\ x_2 \geq 0\}$. In degree 2 the kernel of the moment matrix contains the vector $(-0.57734851,\ 0,\ 0,\ 0.57734864,\ 0,\ 0.57735366)$, that is, the polynomial $x_1^2+x_2^2-1$. The kernel can give no more than this, since the vanishing ideal of the semicircle equals that of the *whole* circle, and no moment-matrix kernel in any degree will produce the constraint $x_2 \geq 0$. We evaluate κ_2 at 2.2×10^4 samples on the circle and threshold at $\tau_2^{\text{variety}} = 9.77$. The extreme points of the 1.1×10^4 accepted samples are $(-1.00000363,\ -1.7\times10^{-4})$ and $(1.00000364,\ 1.2\times10^{-4})$, and $d_{\text{H}}(S_{2,\tau_2}^{\text{variety}}, S) = 0.05$. The lower half of the circle is rejected by the CD kernel alone, κ_2 increasing by more than an order of magnitude across $x_2 = 0$ (Figure 6, left). Thresholding the 10^6 grid points in degree 2 with $\tau_2^{\text{grid}} = 4.64$ gives $d_{\text{H}}(S_{2,\tau_2}^{\text{grid}}, S) = 0.69$, and thresholding in degree 4 with $\tau_4^{\text{grid}} = 30$ gives $d_{\text{H}}(S_{4,\tau_4}^{\text{grid}}, S) = 0.19$ (Figure 7).

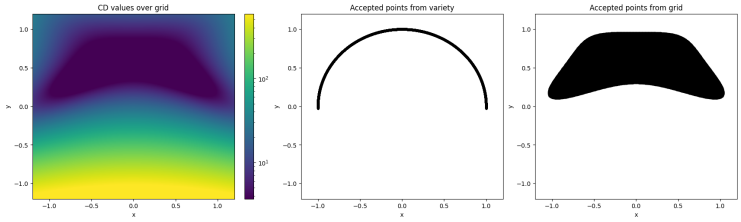

Figure 6: The CD kernel κ_2 for the semicircle. *Left:* values of κ_2 on the 1000×1000 grid, on a logarithmic color scale; the kernel separates the two halves of the circle, although both have the same vanishing ideal. *Middle:* the 1.1×10^4 accepted points of the variety, $d_{\text{H}}(S_{2,\tau_2}^{\text{variety}}, S) = 0.05$. *Right:* the 2×10^5 accepted grid points, $d_{\text{H}}(S_{2,\tau_2}^{\text{grid}}, S) = 0.69$.

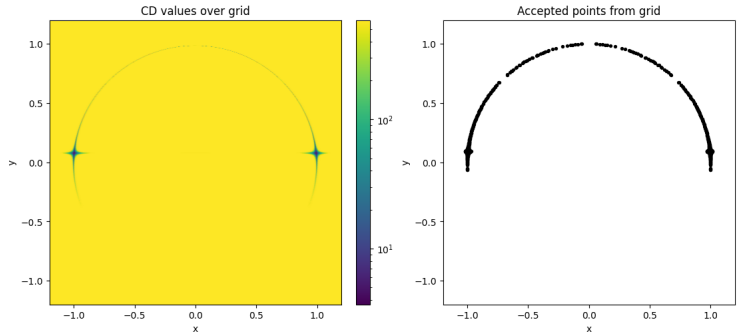

Figure 7: The CD kernel κ_4 for the semicircle. *Left:* values of κ_4 on the 1000×1000 grid, on a logarithmic color scale. *Right:* the accepted grid points, $d_{\text{H}}(S_{4,\tau_4}^{\text{grid}}, S) = 0.19$.

F THE PRODUCT OF A SEMICIRCLE AND A SEGMENT

This example is set up in $\mathbb{R}^4$. To include terms of higher degree in the objective while keeping the moment matrix small, we take four variables and solve, at relaxation degree 4,

$$\begin{aligned} \min_{x\in\mathbb{R}^4} \quad & F(x) = (1-x_1^2-x_2^2)^2 + (x_3-x_4)^2 \\ \text{s.t.} \quad & x_2 \geq 0, \quad x_3 \geq -0.5, \quad x_4 \leq 0.5, \end{aligned}$$

whose set of minimizers is

$$S = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1,\ x_2 \geq 0\} \times \{(x_3, x_4) \in \mathbb{R}^2 : x_3 = x_4,\ -0.5 \leq x_3 \leq 0.5\} \subset \mathbb{R}^4.$$

In degree 2 the kernel of $M_2(y)$ is two-dimensional, spanned by the coefficient vectors of $x_1^2 + x_2^2 - 1$ and $x_3 - x_4$. The homotopy-continuation sampler, applied to these two polynomials, gives 10^4 samples, at which we evaluate the CD kernel κ_2 of degree 2. We also evaluate the CD kernel κ_4 of degree 4 at 6×10^6 grid points. Thresholding at $\tau_2^{\text{variety}} = 14$ and $\tau_4^{\text{grid}} = 20$ results into $S_{2,\tau_2}^{\text{variety}}$ consisting of 2676 points and $S_{4,\tau_4}^{\text{grid}}$ which has 324 points. We report the Hausdorff distances computed against a 0.01-dense discretization of S:

$$d_{\mathrm{H}}\big(S_{2,\tau_2}^{\text{variety}}, S\big) = 0.1, \quad d_{\mathrm{H}}\big(S_{4,\tau_4}^{\text{grid}}, S\big) = 0.246.$$

The projection of samples accepted on the variety is depicted in Figure 8.

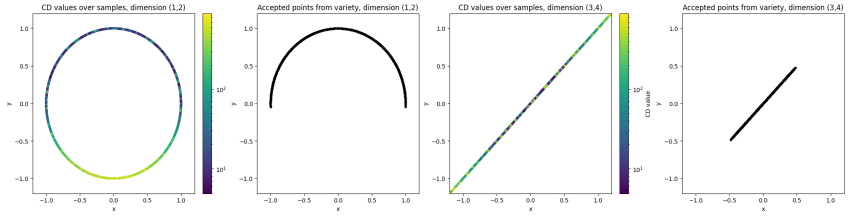

Figure 8: CD kernel κ_2 and the accepted variety points. *Left:* κ_2 evaluated on the variety, projected onto dimensions 1 and 2. *Middle left:* The accepted points projected onto dimensions 1 and 2. *Middle right:* κ_2 evaluated on the variety, projected onto dimensions 3 and 4. *Right:* The accepted points projected onto dimensions 3 and 4.

G Experimental details

Examples with moments in closed form. For each of the five experiments of this kind in Table 1, μ is the uniform measure with respect to $\mathcal{H}^s|_Z$ on the indicated subset of the indicated variety, except for the second hyperbola, where μ is the image of the uniform measure on $t \in [-T, T]$, $T = \operatorname{arcosh} 2$, under $t \mapsto (\cosh t, \sinh t)$; its density with respect to arc length is bounded above and below, as Problem 1 requires. The moments up to degree $2d_v$ are computed in closed form. The carrier is recovered from $\ker M_d(y)$ at the smallest d for which the kernel is nontrivial, which is $d = 2$ for the conics, the cone and the cylinder.

Examples from polynomial optimization. The POPs for the segment and the semicircle in Section 5.3 are solved with TSSOS (Magron & Wang, 2021a;b) at relaxation degree 2, and the resulting pseudo-moment sequence is used as y in Algorithm 1 without further processing. Neither pseudo-moment matrix has the flat extension property, so that neither the extraction procedure of Curto & Fialkow (1998) nor the GNS construction of López Quijorna (2021) is available.

Reproducibility. All figures were generated using Jupyter notebooks included with the reproducibility code. We list below the notebook used for each figure. Figure 1 was generated by `analytic_cone_extraction.ipynb`. Figure 2 was generated using `2nd_line_extraction_TSSOS.ipynb`, run separately for each dimension; the resulting data used in the figure are stored in the same notebook. Figure 3 was generated by `analytic_parabola.ipynb`. Figures 4 and 5 were generated by `segment_extraction_TSSOS_degree_1.ipynb` and `segment_extraction_TSSOS_degree_2.ipynb`. Figures 6 and 7 were generated by `semicircle_extraction_TSSOS_1.ipynb` and `semicircle_extraction_TSSOS_2.ipynb`. Figure 8 was produced by `4d_circle_extraction_TSSOS_1.ipynb`